

Laura C. Polola, Liliana S. Pagano, Gabriela I. Beordi,
Mónica A. Bortolotto, Sandra E. Rodríguez y Pablo F. López

ESTADÍSTICA PRACTICA

Usos y aplicaciones en las Ciencias Económicas



DEPARTAMENTO DE
CIENCIAS ECONÓMICAS

**Laura C. Polola, Liliana S. Pagano, Gabriela I. Beordi,
Mónica A. Bortolotto, Sandra E. Rodríguez y Pablo F. López**

ESTADÍSTICA PRÁCTICA

Usos y aplicaciones en las Ciencias Económicas

Departamento de Ciencias Económicas

Cátedra de Estadística (2407)

2022



Universidad Nacional de La Matanza

Estadística práctica : usos y aplicaciones en las Ciencias Económicas / Laura C.
Polola ... [et al.]. - 1a ed. - San Justo : Universidad Nacional de La Matanza, 2022.
Libro digital, PDF

Archivo Digital: descarga y online
ISBN 978-987-8931-41-8

1. Economía. 2. Estadísticas. I. Polola, Laura C.
CDD 330.071

© Universidad Nacional de La Matanza, 2022
Florencio Varela 1903 (B1754JEC)
San Justo / Buenos Aires / Argentina
Telefax: (54-11) 4480-8900
editorial@unlam.edu.ar
www.unlam.edu.ar

ISBN: 978-987-8931-41-8

Diseño: Editorial UNLaM

Hecho el depósito que marca la ley 11.723
Prohibida su reproducción total o parcial
Derechos reservados

ÍNDICE

Introducción	7
Temas de Estadística	9
Red conceptual de la asignatura	13
Práctica 1: introducción a la estadística - procesamiento de los datos	
GPS 1: orientaciones importantes para no perderse	15
Glosario 1	16
Problemas y situaciones prácticas 1	16
Respuestas de las actividades de la práctica 1.....	19
Práctica 2: tipos de medidas	
GPS 2: orientaciones importantes para no perderse	35
Glosario 2	36
Problemas y situaciones prácticas 2	37
Respuestas de las actividades de la práctica 2.....	50
Práctica 3: números índices	
GPS 3: orientaciones importantes para no perderse	59
Glosario 3	59
Problemas y situaciones prácticas 3	60
Respuestas de las actividades de la práctica 3.....	63
Práctica 4: series de tiempo	
GPS 4: orientaciones importantes para no perderse	69
Glosario 4	69
Problemas y situaciones prácticas 4	70
Respuestas de las actividades de la práctica 4.....	75
Práctica 5: cálculo de probabilidades	
GPS 5: orientaciones importantes para no perderse	83
Glosario 5	84
Problemas y situaciones prácticas 5	84
Respuestas de las actividades de la práctica 5.....	91
Práctica 6: variable aleatoria	
GPS 6: orientaciones importantes para no perderse	97
Glosario 6	97
Problemas y situaciones prácticas 6	98
Respuestas de las actividades de la práctica 6.....	104
Práctica 7: distribuciones especiales de probabilidad. Variables discretas	
GPS 7: orientaciones importantes para no perderse	109
Glosario 7	110
Problemas y situaciones prácticas	110
Respuestas de las actividades de la práctica 7.....	116

Práctica 8: distribuciones especiales de probabilidad. Variables continuas	
GPS 8: orientaciones importantes para no perderse	119
Glosario 8	120
Problemas y situaciones prácticas 8	120
Respuestas de las actividades de la práctica 8.....	127
Práctica 9: distribuciones de estadísticos muestrales	
GPS 9: orientaciones importantes para no perderse	129
Glosario 9	129
Problemas y situaciones prácticas 9	130
Respuestas de las actividades de la práctica 9.....	132
Práctica 10: estimación de parámetros	
GPS 10: orientaciones importantes para no perderse	135
Glosario 10	135
Problemas y situaciones prácticas 10	136
Respuestas de las actividades de la práctica 10.....	141
Práctica 11: prueba de hipótesis	
GPS 11: orientaciones importantes para no perderse.....	145
Glosario 11	145
Problemas y situaciones prácticas 11	146
Respuestas de las actividades de la práctica 11	150
Práctica 12: análisis de regresión y de correlación	
GPS 12: orientaciones importantes para no perderse	155
Glosario 12	155
Problemas y situaciones prácticas 12	156
Respuestas de las actividades de la práctica 12.....	160
Bibliografía.....	165
Anexo	169
Apéndices	171

Introducción

¿Qué es la estadística?

La estadística se estructuró, como disciplina científica, en el siglo XIX, pero ya se conocía y se aplicaba en forma rudimentaria desde la Antigüedad, donde se realizaban relevamientos de población y de recursos. Se puede afirmar que el quehacer estadístico es tan viejo como la historia registrada.

La estadística puede definirse de diversas maneras y actualmente podría caracterizarse como la ciencia de la recolección y análisis de datos para la toma de decisiones, que transforme datos en información. En su método comienza presentando técnicas de diseño y recolección de datos respecto a un fenómeno y luego, mediante la estadística descriptiva, se resume lo medular de la información. La inferencia estadística extiende las conclusiones obtenidas de la muestra a la población de la que ella es parte, además de postular modelos que mejor se ajusten a los datos.

La cátedra de Estadística, que reúne a los autores de este material, viene trabajando desde hace tiempo sobre estrategias de enseñanza de la disciplina desde una visión fundamentalmente aplicada y dinámica, a fin de robustecer la capacidad de transferir los aprendizajes que logren obtenerse mediante el uso en situaciones concretas de técnicas estadísticas, probabilísticas y de inferencia, a la actividad profesional y académica, que en el futuro no lejano se prevé que abordarán sus estudiantes.

A raíz de estas experiencias, se destacan las principales características de esta obra, que son fundamentalmente guiar al alumno en un recorrido que replique diferentes momentos de un trabajo de investigación donde se aplican cálculos y procesamientos estadísticos variados.

Durante el presente curso introduciremos técnicas estadísticas básicas como el diseño y aplicación de modelos para los distintos tipos de variables, métodos de estimación univariado, contraste de hipótesis, análisis de asociación lineal entre variables y pronóstico.

Todos estos contenidos se articulan en función de cómo se desarrollan a lo largo de un proceso de investigación, siguiendo su orden natural de uso y aplicación cuando se realizan trabajos de campo que aportan evidencias del estado de situación de un determinado hecho de interés, y con ellas se generan conclusiones que pueden requerir diferentes grados de tratamiento de la información recopilada, según sean los objetivos del estudio.

Este material contempla, como contexto general donde se enmarcan las actividades de aplicación estadística, diversas temáticas referidas a cuestiones económicas, comerciales y administrativas acordes a las carreras donde se incluye la materia, que son las de Contador Público, Administración, Comercio Internacional y Economía.

Por último, y por ser de gran valor para nosotros, es preciso decir que todo el material didáctico recopilado en este volumen tuvo su origen e inspiración en el trabajo llevado adelante desde la fundación de la Universidad Nacional de La Matanza hasta el presente, en el dictado de esta materia para las primeras carreras de departamento y luego las que se fueron incorporando, por parte de docentes que crearon un estilo y dejaron su impronta que caracteriza a esta cátedra, y esta es fundamentalmente abordar la tarea docente con todo lo que eso implica: una gran responsabilidad en lo que a la enseñanza se refiere, en la motivación por investigar incansablemente acerca de metodologías y estrategias de enseñanza, en la elaboración de materiales novedosos para poder apoyar a los alumnos tanto en la disponibilidad de contenidos teóricos como prácticos, en formato impreso, virtual y digital desde mucho antes que la pandemia lo hiciera obligatorio y universal.

Por eso y por mucho más, vaya un especial agradecimiento a María Eugenia Ángel y Graciela Fernández, reconocidas como docentes fundadoras de la Universidad, como así también a Silvia Brunetti, Mónica González Camus y Enrique Borgna, porque a través de su obra y ejemplo a lo largo de los años, nos siguen acompañando.

Temas de Estadística

Extracto de contenidos presentes en el programa oficial de la materia

Introducción a la estadística

Concepto de estadística. Campos del conocimiento donde se originó y se aplica. Utilización de la investigación estadística en la economía. Encuestas. Etapas del análisis estadístico. Población y muestra. La unidad de observación. Variables cualitativas y cuantitativas. Estadística descriptiva e inferencial.

Procesamiento de los datos

Recopilación y codificación de los datos. Tabulación de la información: matriz de datos. Distribuciones de frecuencias para los distintos tipos de variables: frecuencias absolutas, relativas, porcentuales, acumuladas. Representación gráfica: tortas, barras, bastones, escalones. Arreglo de datos ordenados. Datos agrupados en intervalos de clase; gráficos: el histograma de Pearson y la ojiva de Galton; su contexto histórico y disciplinar. Casos en los que se necesita agrupar en intervalos. Polígono de frecuencias. Tablas de distribución de frecuencias conjuntas; gráficos.

Tipos de medidas

Medidas de posición: media (postulado de la media aritmética de Gauss), mediana, modo, cuartiles, deciles, percentiles. Medidas de dispersión o de variabilidad: rango, varianza de Gauss, desvío típico o estándar, desvío medio de Laplace, desvío intercuartil. Coeficiente de variación de Pearson. Para cada una de las medidas: definición, propiedades, interpretación, relaciones, determinación analítica y posibilidad de determinación gráfica en algunos casos. Medidas de intensidad: cálculo e interpretación de razones y proporciones. Nociones sobre asimetría y curtosis. Gráfico de caja y bigotes (*box-plot*) de Tukey: construcción, análisis e interpretación y comparación de distribuciones de datos.

Números índices

Definiciones, aplicaciones y propiedades. Clasificación. Problemas implicados en el cálculo de números índices. Enlaces y cadenas relativas. Cambio del período base en los números índices. Índices simples y compuestos. Índices de precio, cantidad y valor. Números índice de Laspeyres, Paasche y Fisher. Contexto histórico y disciplinar. Índices generales de precios del INDEC: IPC, IPIM e ICC. Índices especiales. Índice deflactor. Ingreso real. Poder adquisitivo del dinero.

Series de tiempo

Nociones básicas. Componentes en el dominio de tiempo: secular, cíclica, estacional e irregular. Gráficos. Análisis de las componentes: obtención de la recta de tendencia secular (método de los mínimos cuadrados), cálculo de los índices estacionales (método del promedio móvil) y de la fluctuación cíclica (método del porcentaje de la tendencia y de los residuales cíclicos relativos). Pronosticación.

Probabilidad

Experimentos determinísticos y aleatorios. Espacios muestrales. Sucesos aleatorios. Sucesos elementales, suceso cierto y sucesos imposibles. Operaciones entre sucesos. Relaciones de exclusión y de independencia estadística entre sucesos aleatorios. Probabilidad: definición clásica de Laplace y frecuencial de Von Mises. Álgebra de las probabilidades: axiomas y teoremas. Contexto histórico y disciplinar. Tipos de probabilidad: total, compuesta y condicional. Asignación de probabilidad a espacios muestrales finitos y equiprobables. Tablas de contingencia: probabilidades compuestas o conjuntas y probabilidades marginales o totales. Diagrama de árbol: probabilidades normales totales y condicionales. Teorema de la probabilidad total. Teorema de Bayes o de la probabilidad condicional de las causas.

Variable aleatoria

Concepto. Variables aleatorias discretas y continuas. Funciones de probabilidad y de distribución de una variable aleatoria discreta, propiedades. Funciones de densidad de probabilidad y de distribución de una variable aleatoria continua, propiedades. Esperanza matemática: definición y propiedades. Varianza y desvío estándar: definición y propiedades. Cálculo de esperanza y varianza de variables aleatorias discretas y continuas. Gráficos para funciones de probabilidad y de distribución.

Modelos especiales de probabilidad

Para variables aleatorias discretas: binomial, hipergeométrico y de Poisson. Aproximación de la distribución Binomial por Poisson. Para variables aleatorias continuas: normal, normal estándar, uniforme y exponencial. Para cada caso: características del modelo, fenómenos al que responde, posibilidad de aplicación, funciones de probabilidad y de distribución, cálculo e interpretación de la esperanza y varianza. Uso de tablas.

Información muestral

Representatividad. Diferencia entre censos y muestras. Técnicas de selección de muestras. Distintos tipos de muestreo. Estadísticos muestrales, concepto y propiedades.

Distribución de estadísticos muestrales

Distribuciones: la media, la proporción y la varianza muestrales. Propiedades. Error estándar. Propiedades de un buen estimador. Muestreo en poblaciones normales: nociones sobre las distribuciones “t” de *Student* y Ji-Cuadrado de Helmer. La distribución normal como límite de la distribución de los estadísticos muestrales. Teorema central del límite.

Estimación de parámetros por intervalos de confianza

Concepto. Fundamentos. Nivel de confianza. Precisión de la estimación. Estimación de la media, la proporción, la varianza y el desvío estándar de una población. Determinación del tamaño de la muestra. Aplicaciones del teorema central del límite.

Ensayo de hipótesis

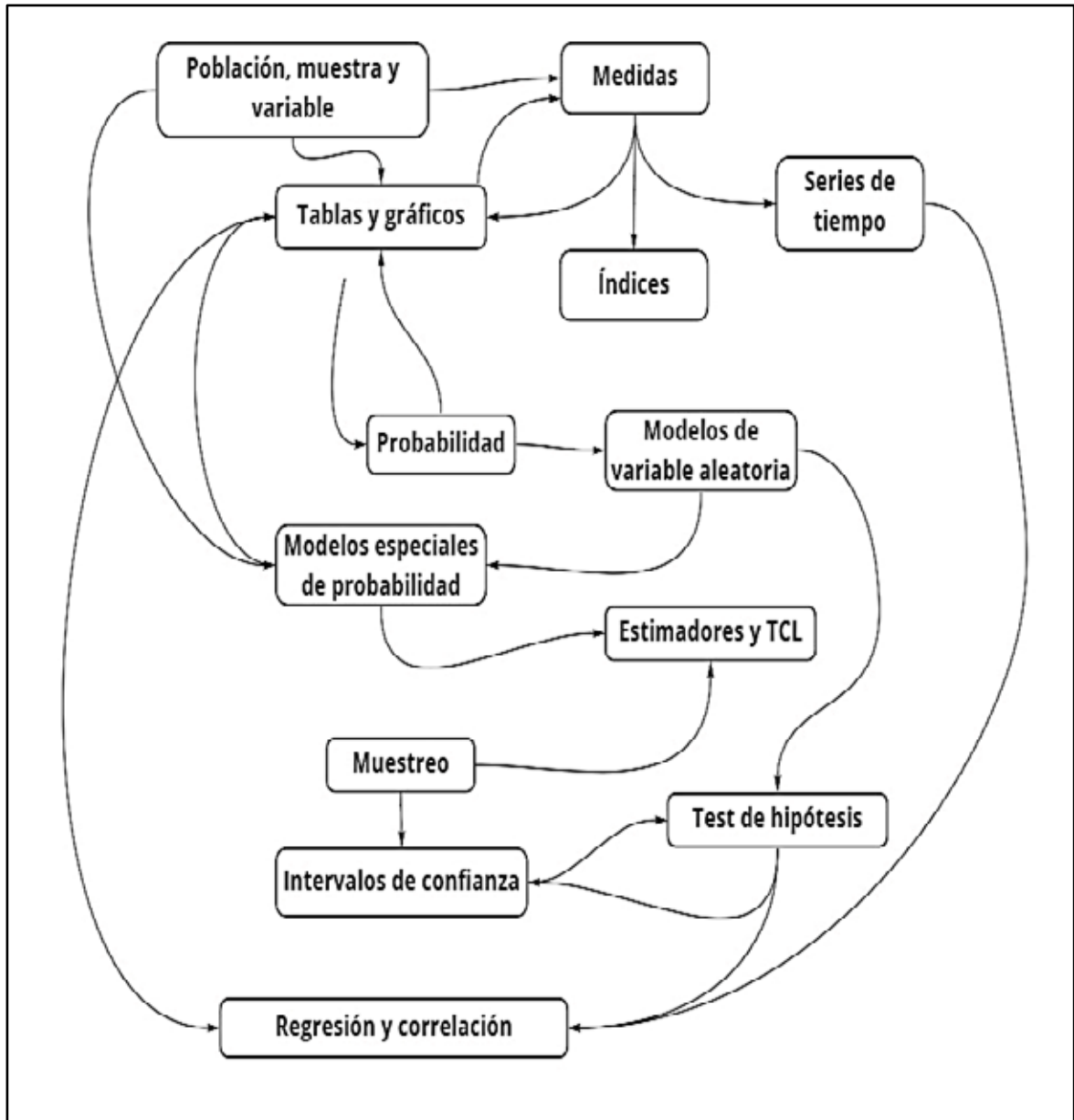
Concepto general. Fundamentos. Hipótesis estadísticas simples y compuestas. Hipótesis nula y alternativa. Punto crítico. Nivel de significación. Tipos de errores. Región crítica o de rechazo de la hipótesis nula y región de aceptación o de no rechazo. Estadísticos de prueba y

reglas de decisión. Procedimientos. Pruebas para la media, la varianza y las proporciones poblacionales. Aplicaciones del teorema central del límite. Pruebas de una o dos colas.

Análisis de regresión y de correlación

Conceptos. Regresión lineal simple. Diagrama de dispersión. Análisis de correlación: correlación lineal, cálculo de los coeficientes de determinación y de correlación; su significado. Análisis de regresión simple: recta de regresión por el método de cuadrados mínimos, límites de aplicabilidad del modelo. Regresión no lineal. Nociones de regresión múltiple. Prueba de hipótesis sobre el coeficiente de correlación y los parámetros de la regresión.

Red conceptual de la asignatura



PRÁCTICA 1

Introducción a la estadística - procesamiento de los datos

En esta práctica se tratarán diversas maneras de organizar los datos y preparar el material de base para que, en adelante, sea posible obtener medidas o indicadores estadísticos. Cada forma de organización implica la caracterización de distintos espacios donde a posteriori se obtendrán las medidas estadísticas.

GPS ESTADÍSTICO 1

Orientaciones importantes para no perderse:

- *La cantidad de información que pueda extraerse de una variable tiene relación directa con la cantidad de espacios abstractos de análisis que ella permita.*
- *Comparativamente, las variables cuantitativas pueden dar mayor cantidad de información que las cualitativas ordenables y estas más que las no ordenables.*
- *Las cualitativas no ordenables poseen solo dos espacios: el de sus valores y sus frecuencias que solo permiten determinar su presencia o peso dentro del conjunto de datos; para las ordenables se agrega su ordenamiento en formato de arreglo ordenado y aparece la noción de acumulación y, por último, en las cuantitativas se incorpora el espacio de números reales donde será posible medir sus distancias relativas como descriptores de su dispersión/concentración.*
- *Los gráficos elaborados permiten tener una idea de conjunto complementando la lectura de la información que se extrae del grupo de datos relevados. Su objetivo principal es maximizar la extracción de la información contenida en los datos empíricos.*
- *Cuando se habla de población no necesariamente esta es un conjunto de personas. Se podrían estudiar, por ejemplo, todas las unidades económicas de la Argentina (como en el caso de los censos nacionales económicos). En este caso la población comprendería desde kioscos, almacenes, supermercados, talleres automotrices, talleres y fábricas textiles hasta grandes industrias siderúrgicas y petroleras de la Argentina donde cada una de esas unidades económicas será una unidad de observación y/o de análisis.*
- *Si el relevamiento se realizara a solo una parte de esa población se trataría de un muestreo y no de un censo.*

Glosario 1

Población objetivo: es la colección de todos los miembros concebibles del grupo que deseamos estudiar.

Unidad de observación estadística: es el objeto o individuo sobre el cual se realiza la observación, se relaciona íntimamente con el grado de detalle con que se planifica estudiar la población.

Muestra representativa: es un subconjunto de la población que debe reflejar sus características.

Variable estadística: cualquier aspecto o característica que varía de un individuo a otro y que se desee estudiar en la población. Se desprende de algún/os atributo/s relevado/s.

Matriz de datos: arreglo que incluye conjuntamente los atributos relevados y las unidades observadas.

Datos crudos o brutos: datos colectados en un relevamiento y no organizados de forma alguna.

Arreglo de datos ordenados: son los datos colectados ordenados en forma ascendente o descendente según los valores numéricos o categorías ordenables de una variable.

Frecuencia simple: cantidad de veces que aparece un valor de la variable, o su porcentaje, o su cantidad relativa en un conjunto de datos.

Frecuencia acumulada: cantidad o porcentaje de datos ordenados que se acumulan hasta un cierto valor de la variable inclusive.

Distribución de frecuencias: tabla donde se listan todos los valores de la variable acompañados de sus correspondientes frecuencias.

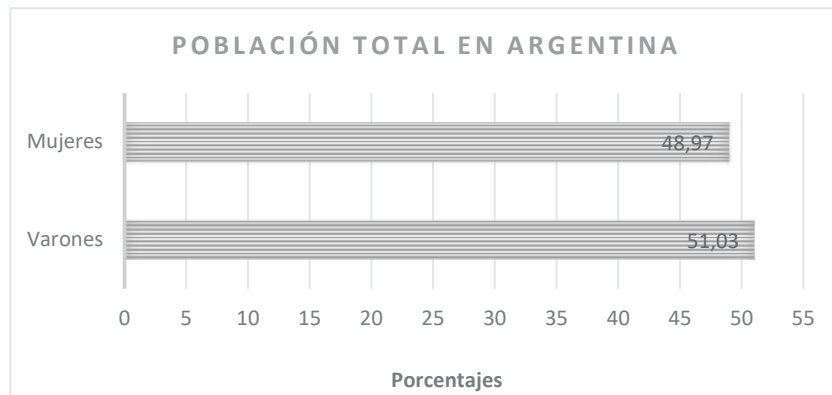
Problemas y situaciones prácticas 1

1. Según publica www.economis.com.ar el 8/8/2019, producto de la mala cosecha de 2018, la Argentina fue desplazada del *ranking* de los 10 países más exportadores de productos agrícolas.

Ranking 2018 (en miles de millones de dólares)					
Puesto	País o comunidad	Volumen	Puesto	País o comunidad	Volumen
1. ^{er}	Unión Europea	681	6. ^o	Indonesia	46
2. ^o	Estados Unidos	172	7. ^o	Tailandia	44
3. ^{er}	Brasil	93	8. ^o	India	39
4. ^o	China	83	9. ^o	Australia	38
5. ^o	Canadá	69	10. ^o	México	35

- a) Identificar y clasificar la variable analizada.
- b) Construir una tabla de distribución de frecuencias.
- c) Graficar.

2. El gráfico siguiente presenta estadísticas demográficas de la Argentina en 2014 donde la población total era de 42.669.500 habitantes (Fuentes: INDEC y Ministerio de Salud de la República Argentina).



- Identificar y clasificar la variable analizada y realizar otro tipo de gráfico con la información disponible.
- Calcular las frecuencias absolutas simples y armar una tabla de distribución de frecuencias.

3- La información que sigue fue extraída de la sección Infográficos Estadísticos publicados en la página web de la Comisión Económica para América Latina (CEPAL). Fuente: https://estadisticas.cepal.org/cepalstat/WEB_CEPALSTAT/Library/CEPALSTAT/Infograficos_pdf/Inf_10_es.pdf



- Identificar las variables intervinientes y clasificarlas.
- Compilar toda la información en una tabla de distribución conjunta.

c) Realizar otro tipo de gráfico con la información disponible.

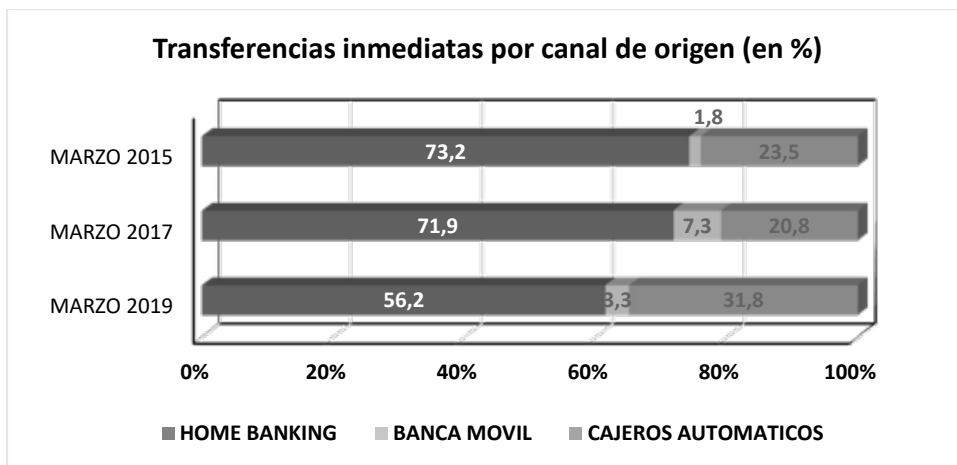
4- La tabla muestra la cantidad de usuarios de servicios bancarios en Argentina a través de distintos dispositivos en un determinado período, en millones de usuarios (Fuente: Clarín, consultado el 26/4/2019).

	Home banking	Banca móvil
2015	5,2	1,5
2016	6,0	3,1
2017	6,6	5,1
2018	7,2	7,1
2019	7,9	10,0

- Identificar y clasificar las variables intervinientes.
- Realizar los gráficos que correspondan con la información disponible.
- Extraer al menos tres conclusiones.

5. En el gráfico siguiente se presentan las transferencias inmediatas realizadas según el canal de origen, en porcentajes. Luego de su análisis:

- Identificar y clasificar las variables intervinientes.
- Compilar la información en una tabla de distribución conjunta.
- Realizar otro tipo de gráfico.
- Sacar conclusiones.



6- Los datos siguientes corresponden a una muestra de los empleados de una empresa tomados al azar, según la planta donde trabajan y el nivel de conformidad con sus condiciones laborales.

Planta	BA	R	C	BA	R	BA	C	BA	C	R
Nivel de conformidad	PC	C	PC	C	C	PC	C	C	MC	PC

Planta	C	C	R	BA	BA	BA	C	R	BA	R
Nivel de conformidad	C	MC	C	PC	MC	C	C	PC	PC	MC

Referencias sobre Plantas

BA: Buenos Aires; R: Rosario; C: Córdoba

Referencias sobre conformidad

C: Conforme; PC: Poco Conforme; MC: Muy Conforme

- Identificar y clasificar las variables en estudio.
- Realizar una tabla de distribución conjunta.
- Graficar.

7- En un centro comercial barrial se realiza un estudio sobre un conjunto de negocios minoristas acerca de la cantidad de empleados que posee cada uno, registrándose la siguiente información.

Empleados	0	1	2	3	4	5
Cantidad de establecimientos	2	4	5	6	4	2

- Identificar la variable en estudio y clasificar. ¿Cuál es el tamaño de la muestra?
- Obtener las frecuencias relativas simples y acumuladas.
- Realizar un gráfico para cada una de las frecuencias obtenidas.
- Extraer conclusiones de la observación de los gráficos.

8- El departamento de personal de una empresa registró la cantidad de días de ausentismo debido a causas personales de sus 115 empleados, en el período determinado, obteniéndose:

N.º de días	0	1	2	3	4	5	6	7	8	10	Total
N.º de empleados	5	19	27	35	10	8	6	3	1	1	115

- Representar gráficamente las frecuencias porcentuales simples y las acumuladas.
- De la observación de las frecuencias acumuladas extraer conclusiones.

9- Los datos que se presentan a continuación corresponden a la cantidad de préstamos otorgados, en una semana determinada, por 56 entidades financieras, cuya finalidad fue la adquisición de viviendas de uso particular.

2	9	14	19	4	7	14	8	13	3	1	19	5	5
15	6	1	8	5	6	10	11	20	7	2	8	15	10
6	11	5	15	1	9	8	8	21	17	17	6	5	1
2	1	4	22	20	10	6	11	3	2	3	10	10	4

- Listar los datos de menor a mayor.
- Confeccionar la tabla de distribución de frecuencias.
- Graficar las frecuencias porcentuales simples y acumuladas.
- Determinar *gráficamente* el porcentaje de entidades que otorgaron:
 - menos de 12 préstamos para la adquisición de viviendas.
 - al menos 8 préstamos para la adquisición de viviendas.
 - más de 10 y hasta 20 préstamos.
 - a lo sumo 15 préstamos.

- e) Agrupar los datos en siete intervalos de igual amplitud y elaborar la tabla correspondiente.
- f) Graficar las frecuencias porcentuales simples y acumuladas.
- g) Determinar gráficamente los mismos porcentajes solicitados en d).
- h) Comparar los gráficos realizados en c) y en d).
- i) Comparar los resultados obtenidos en d) con los obtenidos en g).

10- Los montos otorgados en 50 préstamos personales (en miles de pesos), por una determinada entidad bancaria, son los siguientes:

90,32	50,15	40,52	190,00	120,00	120,78	250,40	50,85	160,50	120,19
100,00	50,54	90,73	60,60	70,20	130,88	80,51	30,29	140,23	70,27
30,56	110,90	30,00	160,10	150,25	100,00	180,90	90,35	50,92	60,55
220,27	90,54	210,12	40,45	70,84	80,70	60,30	300,00	30,34	150,90
200,50	60,85	170,40	90,29	100,09	60,45	80,65	16,60	80,71	130,29

- a) Listar los datos de menor a mayor.
 - b) Agrupar los datos en 5 intervalos y en 10 intervalos, y hacer coincidir en los dos agrupamientos el límite inferior del primer intervalo y el límite superior del último.
 - c) Para cada caso construir la distribución de frecuencias correspondiente.
 - d) Realizar los histogramas y los polígonos de frecuencias porcentuales y compararlos.
- 11- El tiempo de espera (en minutos) en un restaurante de *fast food* en 20 servicios se resume en la siguiente tabla:

Tiempo	0-5	5-10	10-15	15-20	20-25	25-30
Cantidad	2	5	6	4	2	1

- a) Completar la tabla de distribución de frecuencias.
 - b) Representar las frecuencias simples y acumuladas.
 - c) Determinar
 - i) el tiempo de espera más frecuente;
 - ii) el tiempo máximo de espera de la mitad de los pedidos.
- 12- La siguiente información corresponde a las horas semanales trabajadas por el total de obreros de cierta industria.

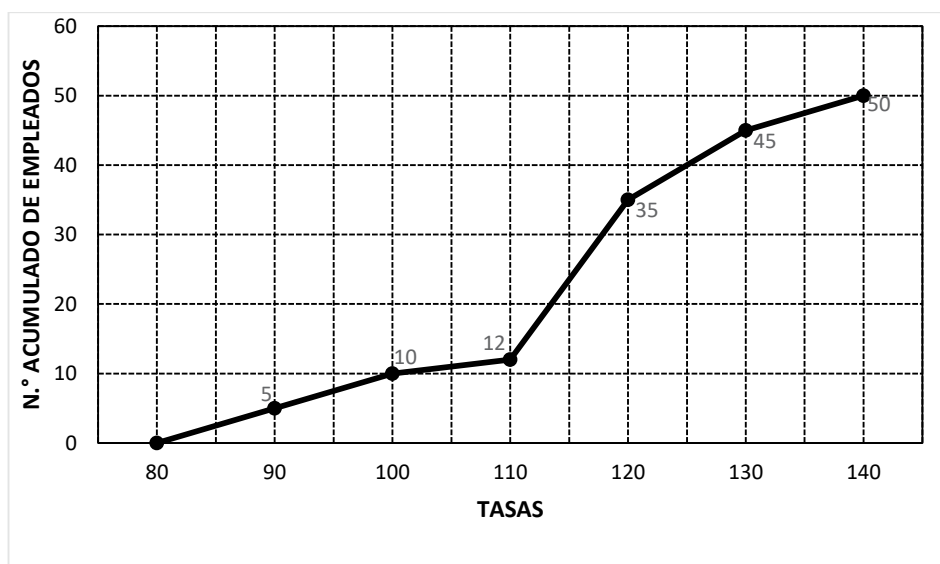
H SEMANALES	30-32	32-34	34-36	36-38	38-40	40-42
N.º DE OBREROS	5	22	30	28	10	2

- a) Completar la tabla de distribución de frecuencias.
- b) Representar gráficamente los distintos tipos de frecuencias simples.
- c) Discutir las propiedades que tiene cada uno de los gráficos (sugerencia: comparar áreas).
- d) Representar las frecuencias porcentuales acumuladas.
- e) Determinar gráficamente el número máximo de horas trabajadas por
 - i) la mitad de los empleados,

- ii) la cuarta parte de los obreros y
- iii) las tres cuartas partes de los obreros.

13- En una empresa multinacional a cada trabajador se le otorga una tasa de producción que representa su eficiencia en el trabajo. Las tasas se organizaron en una distribución de frecuencias y luego se representaron en forma de un polígono de frecuencias acumuladas.

- a) Marcar en el gráfico la tasa máxima de producción que posee el 50 % de los empleados.
- b) Reconstruir la tabla de distribución de frecuencias.
- c) Confeccionar el histograma correspondiente.
- d) Marcar en el gráfico la tasa que predomina.
- e) Encontrar el porcentaje de empleados que poseen una tasa de producción inferior 110.
- f) Hallar el porcentaje de empleados que poseen una tasa de producción superior a 130.
- g) ¿Cuál es aproximadamente la tasa de producción otorgada como máximo al 75 % de los empleados?
- h) Elaborar un breve informe que compile e interprete los resultados obtenidos a modo de conclusión general.



Como se evidencia en esta situación, el gráfico de la ojiva de Galton -que conlleva el concepto de acumulación de frecuencias- permite extraer información del conjunto de datos desde otro ángulo.

A pesar de que el concepto de acumulación parece surgir naturalmente a posteriori del de frecuencia simple, en la realidad histórica se formalizó antes: el gráfico de la Ojiva de Galton precedió al histograma de Pearson en más de una década.

En el procesamiento de la información se utilizan técnicas básicas para llevar a cabo un estudio descriptivo, se puede hablar de dos tipos diferenciados de técnicas, unas utilizadas para variables cualitativas y otras para variables cuantitativas.

14- Analizar si las siguientes afirmaciones son verdaderas *Siempre - A veces - Nunca*. Justificar y dar algún ejemplo donde no se cumplan aquellas que no son siempre ciertas.

- Si una variable adopta valores numéricos es cuantitativa.
- Si una variable es cualitativa no es ordenable.
- Una variable continua no puede trabajarse como discreta (discretización).
- Una variable discreta puede trabajarse como continua.
- De una tabla de frecuencias de una variable discreta (o discretizada) pueden recuperarse los datos originales.
- De una tabla de frecuencias de una variable agrupada en intervalos de clase pueden recuperarse los datos originales aproximados.
- Desde cualquier gráfico es posible reconstruir la tabla de frecuencias que le dio origen.

15- La información consignada en la matriz de datos que aparece a continuación, corresponden a una muestra aleatoria de 40 empresas pymes del conurbano bonaerense, extraída del Censo Nacional Económico 2004/2005.

Orden	Tipo	Rubro	Antigüedad	Endeudamiento	Cantidad de personal	Puestos de trabajo 2007-2009
1	P	S	1	M	128	M
2	P	I	6	M	150	D
3	R	A	8	N	19	A
4	R	C	4	M	170	M
5	P	I	8	MA	112	M
6	P	A	3	A	140	D
7	R	I	10	B	20	A
8	P	S	8	MA	23	D
9	P	C	4	M	114	M
10	P	A	0	MA	148	A
11	R	A	2	M	154	A
12	P	I	4	B	100	D
13	P	C	5	MA	270	D
14	R	I	6	A	142	A
15	P	A	7	MA	105	M
16	R	I	2	A	160	M
17	P	C	9	M	22	M
18	R	I	17	B	5	M
19	P	S	11	M	120	D
20	P	I	6	N	137	A
21	P	I	7	M	80	D
22	R	A	4	A	109	M
23	R	C	25	MA	110	A
24	P	S	11	B	144	A
25	P	C	4	M	120	A

26	R	I	7	M	110	M
27	P	S	9	A	98	M
28	P	S	6	B	67	D
29	P	C	10	A	115	D
30	P	C	1	N	21	M
31	R	S	5	N	109	M
32	R	I	2	M	125	M
33	R	A	8	MA	34	A
34	P	A	0	B	99	A
35	R	I	15	M	132	D
36	P	S	8	N	165	A
37	P	S	6	M	38	M
38	R	I	3	A	25	M
39	R	A	5	M	9	M
40	P	S	2	N	115	A

Para las diferentes características observadas se indican las referencias correspondientes:

Tipo de pyme: R: recuperada; P: Privada

Rubro: A: Agrícola; C: Comercial; I: Industrial; S: Servicios.

Antigüedad: en años al 2005.

Nivel de endeudamiento: N: Ninguno; B: Bajo; M: Medio; A: Alto; MA: Muy Alto.

Puestos de trabajo durante 2003-2005: D: Disminuyó; M: Mantuvo; A: Aumentó.

- Identificar y clasificar todas las variables cuyos valores están volcados en la matriz.
- Para cada una de las variables confeccionar la tabla de distribución de frecuencias adecuada y graficar convenientemente.
- Elegir dos variables y realizar una tabla de distribución de frecuencias conjuntas. Graficar de dos formas diferentes y extraer conclusiones de los gráficos realizados.
- Elaborar un breve informe donde se desarrollen e interpreten los resultados observados en el análisis realizado.

RESPUESTAS DE LAS ACTIVIDADES DE LA PRÁCTICA 1

Introducción a la estadística - procesamiento de datos

1- a) Variable:

X: los 10 países más exportadores de productos agrícolas en 2018 (variable cualitativa no ordenable).

b) Tabla de distribución de frecuencias:

X	fi	fr	f%
Unión Europea	681	0.5239	52.39
Estados Unidos	172	0.1323	13.23
Brasil	93	0.0715	7.15
China	83	0.0639	6.39
Canadá	69	0.0531	5.31
Indonesia	46	0.0354	3.54
Tailandia	44	0.0338	3.38
India	39	0.0300	3.00
Australia	38	0.0292	2.92
México	35	0.0268	2.69
	1300	1.0000	100.00

c) Gráfico de barras.

2. a) Variable:

X: sexo de las personas en Argentina en 2014 (variable cualitativa no ordenable)

Otro gráfico: torta.

b) Tabla de distribución de frecuencias:

X	fi	Fr	f%
Varones	20895254	0.4897	48.97
Mujeres	21774246	0.5103	51.03
	42669500	1	100

3- a) Mediante gráficos de torta segmentados se muestra la información conjunta de las siguientes variables:

Variable 1: “*Categoría ocupacional de la población ocupada urbana de 25 a 59 años de América Latina*”.

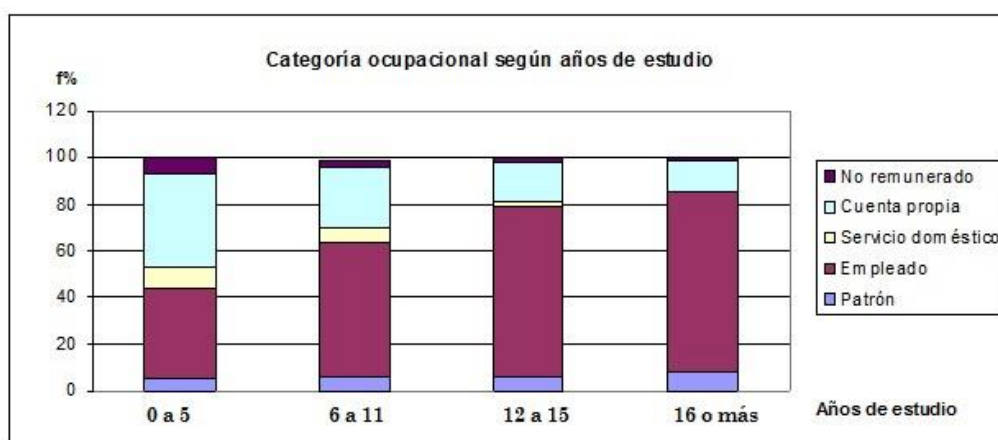
Variable 2: “*Cantidad de años de estudio de la población ocupada urbana de 25 a 59 años de América Latina*”.

b) Tabla de distribución conjunta.

Categoría ocupacional	Años de estudio			
	0 a 5	6 a 11	12 a 15	16 o más
Patrón	5	6	6	8
Empleado	39	58	73	78
Servicio doméstico	9	6	2	0
Cuenta propia	40	26	17	13
No remunerado	7	3	2	1
Subtotal	100	99	100	100

Importante: todas las columnas deberían sumar 100, probablemente falló el ajuste por redondeo de los porcentajes en el grupo que tienen entre 6 y 11 años de estudio. Esto no puede corregirse si no se cuenta con los datos primarios, o sea con la información relevada originalmente.

- c) Pueden utilizarse para mostrar la información disponible los gráficos de barras segmentadas (o apiladas) o el de barras adyacentes.



4. a) Variables

X_1 : años (variable cualitativa ordenable)

X_2 : dispositivos de acceso a servicios bancarios (variable cualitativa no ordenable)

b) Gráfico de barras segmentadas o adyacentes.

5. a) X_1 : períodos de tiempo (variable cualitativa ordenable)

X_2 : canal de origen de las transferencias inmediatas.
(variable cualitativa no ordenable)

b) Tabla de distribución de frecuencias.

	<i>Home banking</i>	Banca móvil	Cajeros automáticos	
Marzo 2015	73.2	3.3	23.5	100
Marzo 2017	71.9	7.3	20.8	100
Marzo 2019	56.2	12	31.8	100

c) Gráfico de barras adyacentes.

6- a) Variables

X_1 : planta (variable cualitativa no ordenable)

X_2 : nivel de conformidad (variable cualitativa ordenable)

b) Ver enunciado del ejercicio 2) de la práctica 2

c) Gráfico de barras segmentadas o adyacentes.

7- a) Variable: “Cantidad de empleados que posee un negocio minorista en el centro comercial barrial”. Variable cuantitativa discreta

Tamaño de la muestra: $n = 23$

b) Frecuencias relativas simples y acumuladas tabuladas.

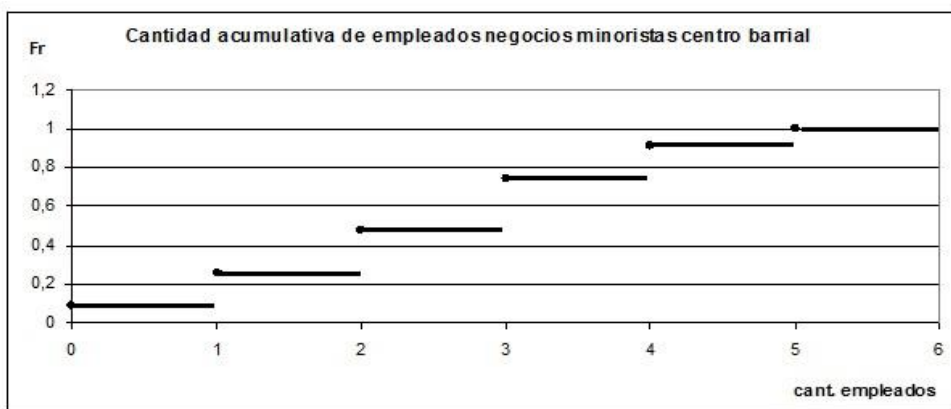
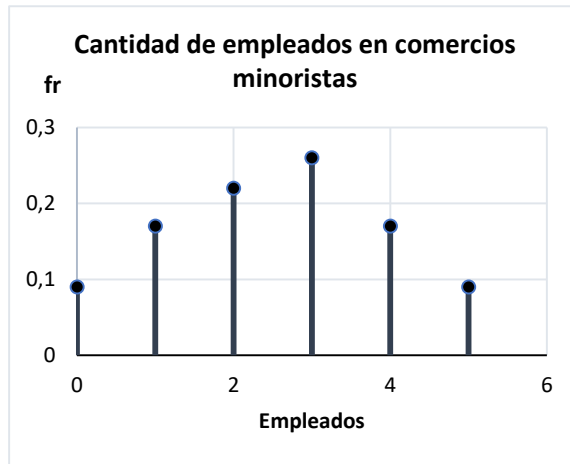
Cant. empleados	0	1	2	3	4	5	Totales
fi	2	4	5	6	4	2	23
fr	0,09	0,17	0,22	0,26	0,17	0,09	1,00
Fr	0,09	0,26	0,48	0,74	0,91	1,00	

c) Para las frecuencias relativas simples:

Gráfico de bastones

Para las frecuencias relativas acumuladas:

Gráfico escalonado



d) Conclusiones de la observación de los gráficos.

- ✓ En el centro comercial barrial predominan los negocios que poseen 3 empleados.
- ✓ En el centro comercial barrial hay la misma proporción de negocios que no poseen empleados que la de los que tienen 5 empleados.
- ✓ La distribución de la cantidad de empleados acumula más rápidamente de 0 a 3 empleados por negocio minorista y después más lentamente (los escalones crecen rápidamente y luego más lentamente pues al final están menos distanciados unos de otros).

8- a) Variable: “Cantidad de días de ausentismo por causas personales por empleado”. Variable cuantitativa discreta. Gráfico de bastones o de puntos (frecuencias porcentuales simples) y gráfico de escalones (frecuencias acumuladas).

b) Posibles conclusiones desde las frecuencias acumuladas:

- ✓ En el departamento de personal de la empresa del ejemplo hay un predominio de empleados que cuentan en un año con 3 días de ausentismo por causas personales.

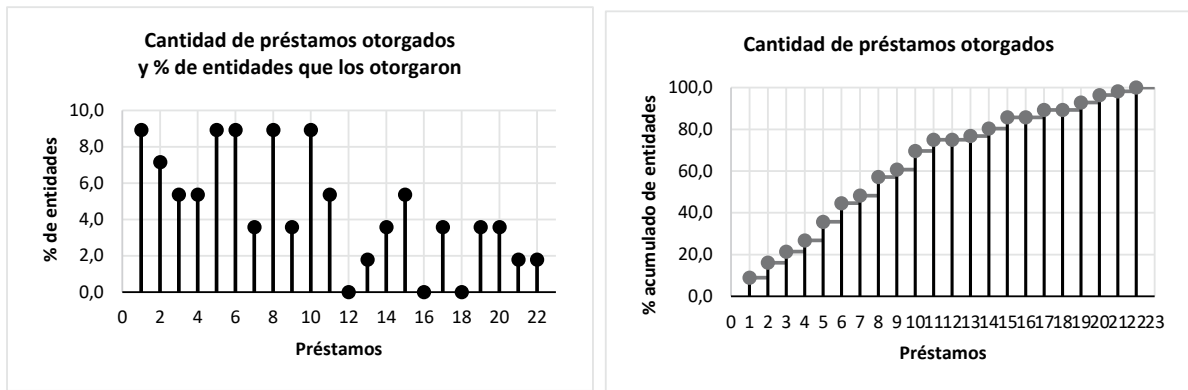
- ✓ El 44,35 % de los empleados faltó como máximo dos días por causas personales.
- ✓ El 4,35 % de los empleados no faltó nunca por causas personales.
- ✓ El 1,74 % de los empleados faltó como mínimo ocho días por causas personales (100-98,26).

9- Variable: “Cantidad de préstamos para la adquisición de viviendas de uso particular otorgados por semana”.

a) Lista de datos ordenados de menor a mayor.

1	1	1	1	1	2	2	2	2	3	3	3	4	4	4
5	5	5	5	5	6	6	6	6	6	7	7	8	8	8
8	8	9	9	10	10	10	10	10	11	11	11	13	14	14
15	15	15	17	17	19	19	20	20	21	22				

c) Gráficos de las frecuencias porcentuales simples y acumuladas.



d) Los porcentajes solicitados provienen de las frecuencias acumuladas porcentuales, por lo cual hay que ir al gráfico de escalones (o a la ojiva en el agrupamiento en intervalos) para determinarlas.

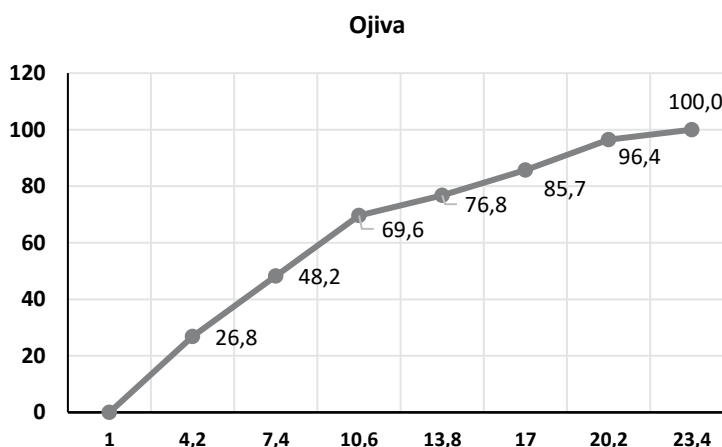
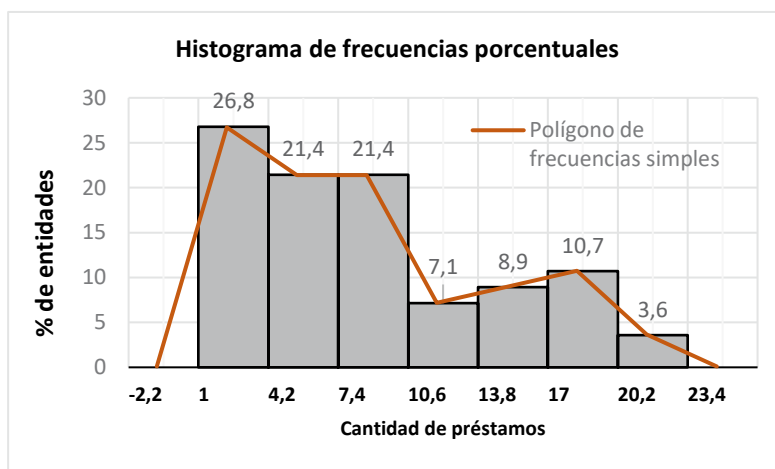
- i) El 76 % de las entidades financieras aproximadamente otorgaron menos de 12 préstamos para la adquisición de viviendas. Justificación: menos de 12 préstamos significa que 12 no es de interés por lo que se desplaza verticalmente desde 11 hasta tocar el gráfico y allí se mueve horizontalmente hasta encontrar el eje de las F%, y donde la horizontal corta ese eje se ubica el porcentaje buscado que es aproximadamente 76 %.
- ii) El 42 % de las entidades financieras aproximadamente otorgaron al menos (como mínimo) 8 préstamos para la adquisición de viviendas (100-58).
- iii) El 27 % de las entidades financieras aproximadamente otorgaron más de 10 y hasta 20 préstamos para la adquisición de viviendas. Justificación: más de 10 y hasta 20 préstamos queda en la zona comprendida por dos verticales que levantemos en 10 y en 20, y el porcentaje será la diferencia del correspondiente a 20 menos el de 10.

iv) El 88 % de las entidades financieras aproximadamente otorgaron a lo sumo 15 préstamos para la adquisición de viviendas. Justificación: a lo sumo es menor o igual a 15, significa que el dato 15 pertenece al escalón que va de 15 hasta 17 pues estos son cerrados en el límite inferior de cada intervalo, por eso se sube con la flecha hasta el escalón siguiente al que termina en 15.

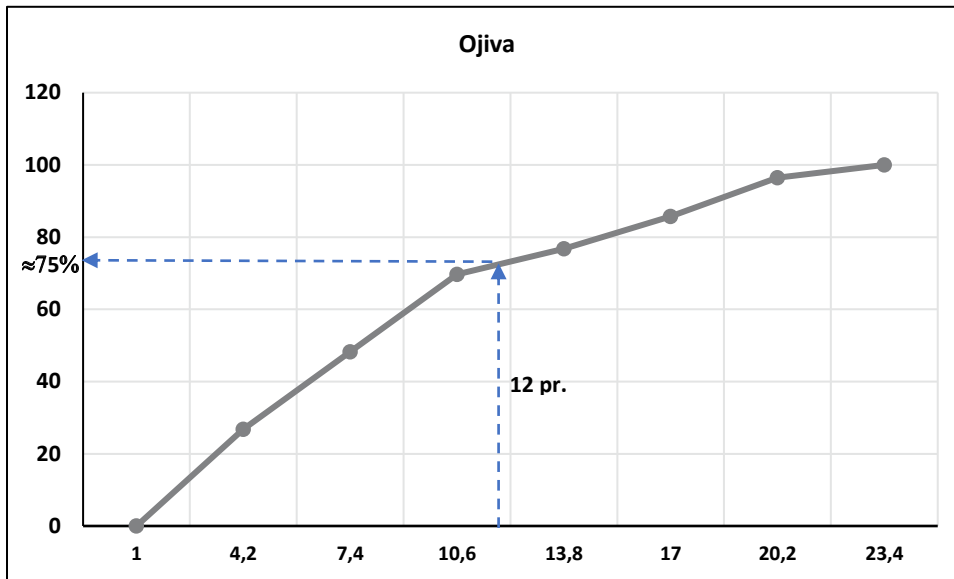
e) Los intervalos que hemos considerado son los siguientes y la tabla ya se vio como se arma:

[1; 4,2) , [4,2; 7,4) ; [7,4; 10,6) , [10,6; 13,8) , [13,8; 17) , [17; 20,2) , [20,2; 23,4)

f) Gráficos: histograma, polígono de frecuencias (frecuencias porcentuales simples) y ojiva (frecuencias porcentuales acumuladas):



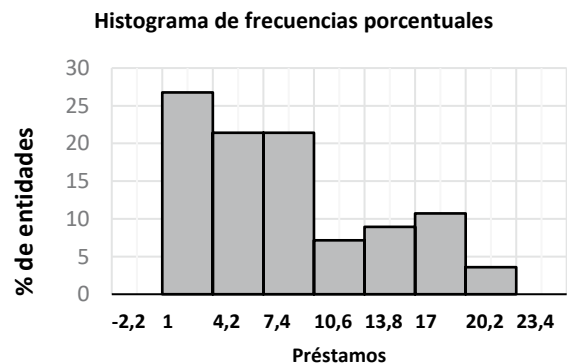
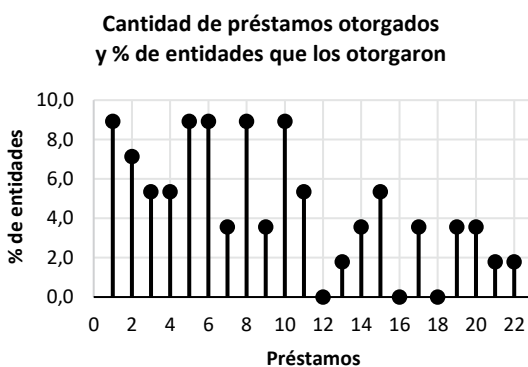
g) Determinar gráficamente los mismos porcentajes solicitados en d).



En la ojiva fue más difícil posicionar exactamente el 12 y la lectura en el eje vertical no es la misma, aproximamos con 75 %, según la escala que usen les va a dar un valor cercano, no necesariamente el mismo.

Respuesta: el 75 % de las entidades financieras aproximadamente otorgaron menos de 12 préstamos para la adquisición de viviendas.

- h) En ambos ítems c) y d) los gráficos muestran la misma “forma general” de la distribución, pero en el agrupamiento en intervalos se pierde detalle de la información, por ejemplo, de que ninguna entidad financiera otorgó 12, 16 o 18 préstamos, como se observa en sendos gráficos.



Lo mismo ocurre con las frecuencias acumulativas, en los gráficos bajo el tratamiento continuo se ve la forma, pero son menos precisos al determinar medidas *por eso solo se los utiliza para visualizar la distribución general, pero no para calcular medidas.*

Únicamente cuando no se tienen los datos originales y solo se dispone de su organización de intervalos, se aplicarán las fórmulas para interpolar valores con el objetivo de obtener aproximaciones de las medidas.

- i) Los porcentajes obtenidos a partir de la ojiva **no siempre** son iguales a los determinados en el gráfico de escalones correspondiente a los valores listados, que son los verdaderos.

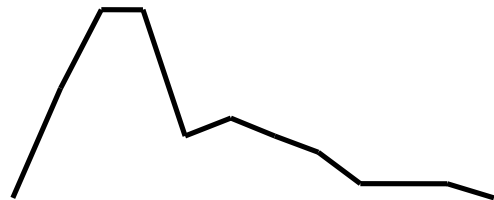
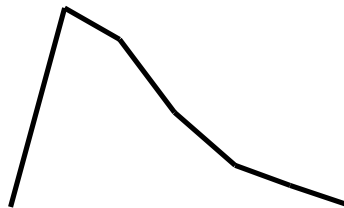
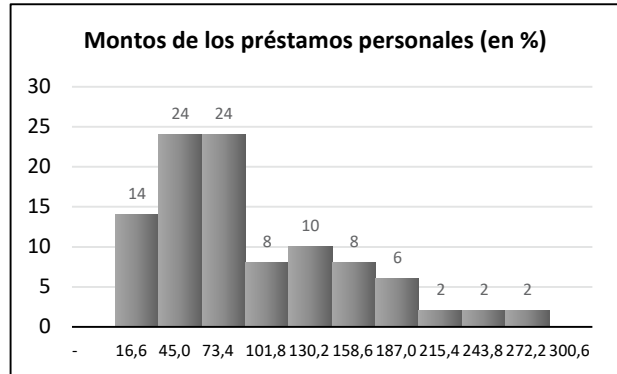
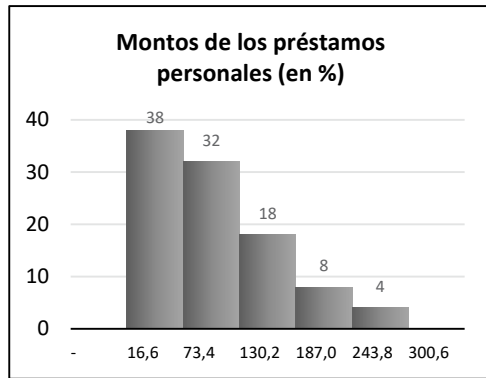
10- a) Lista de datos ordenados y tablas de frecuencias agrupados en 5 y 10 intervalos.

16,6	30	30,29	30,34	30,56	40,45	40,52	50,15	50,54	50,85
50,92	60,3	60,45	60,55	60,6	60,85	70,2	70,27	70,84	80,51
80,65	80,7	80,71	90,29	90,32	90,35	90,54	90,73	100	100
100,09	110,9	120	120,19	120,78	130,29	130,88	140,23	150,25	150,9
160,1	160,5	170,4	180,9	190	200,5	210,12	220,27	250,4	300

Intervalos	x_i	f_i	$f\%$	$F\%$
16,6-73,4	45,0	19	38	38
73,4-130,2	101,8	16	32	70
130,2-187	158,6	9	18	88
187-243,8	215,4	4	8	96
243,8-300,6	272,2	2	4	100
n=		50	100	

Intervalos	x_i	f_i	$f\%$	$F\%$
16,6-45	30,8	7	14	14
45-73,4	59,2	12	24	38
73,4-101,8	87,6	12	24	62
101,8-130,2	116,0	4	8	70
130,2-158,6	144,4	5	10	80
158,6-187	172,8	4	8	88
187-215,4	201,2	3	6	94
215,4-243,8	229,6	1	2	96
243,8-272,2	258,0	1	2	98
272,2-300,6	286,4	1	2	100
n=		50	100	

d) Gráficos de los histogramas y los polígonos de frecuencias porcentuales.



Se puede observar en los gráficos que, a mayor cantidad de intervalos, más suave será la curva que determine el polígono de frecuencias, en el de menor cantidad de intervalos los saltos son más abruptos tanto en el histograma como en el polígono de frecuencias.

11- a) Tabla de distribución de frecuencias:

Tiempo	fi	fr	f%	Fi	Fr	F%
0-5	2	0.1	10	2	0.1	10
5-10	5	0.25	25	7	0.35	35
10-15	6	0.3	30	13	0.65	65
15-20	4	0.2	20	17	0.85	85
20-25	2	0.1	10	19	0.95	95
25-30	1	0.05	5	20	1	100
Total:	20	1	100			

b) Gráficos:

Con las frecuencias simples: histograma, polígono de frecuencias.

Con las frecuencias acumuladas: ojiva

c) i) $I. Mo = [10,15)$

Los tiempos de espera más frecuentes se encuentran entre 10 y 15 minutos o, podríamos decir, alrededor de 12.5 minutos.

ii) En la ojiva de las F% se traza una horizontal que pasa por 50 hasta cortar al gráfico y se proyecta en el eje X.

12- Similar al ejercicio 11)

c) El histograma y el polígono de frecuencias encierran la misma área. Considerando como unidad de medida el ancho del intervalo, para los distintos tipos de frecuencia se verifica que:

$$A_{fi} = n$$

$$A_{fr} = 1 \text{ (propiedad importante que retomaremos en la práctica 6)}$$

$$A_{f\%} = 100$$

13- a) 50 % de los empleados = 25 empleados.

Se traza una horizontal por 25 hasta cortar al gráfico.

b) Tabla de distribución de frecuencias.

Tasa	fi	f%	Fi	F%
80-90	5	10	5	10
90-100	5	10	10	20
100-110	2	4	12	24
110-120	23	46	35	70
120-130	10	20	45	90
130-140	5	10	50	100
Total:	50	100		

d) I. Mo = [110,120)

Las tasas de producción más frecuentes se encuentran entre 110 y 120 o alrededor de 115.

e) 24 %

f) 10 %

g) Similar a a)

14-

a) A veces

d) A veces

g) Siempre

b) A veces

e) Siempre

c) A veces

f) Nunca

PRÁCTICA 2

Medidas estadísticas

“Toda ciencia es medición, toda medición es estadística”.
Hermann L.F. de Helmholtz (físico y fisiólogo alemán-S. XIX)

En esta práctica se comenzará a medir sobre el conjunto de datos con la finalidad de determinar sus características principales.

Las medidas que se determinarán son los denominados indicadores estadísticos del grupo de datos. La interpretación que se realice de cada uno de estos indicadores debe referirse estrictamente al grupo de datos tratados.

GPS ESTADÍSTICO 2

Orientaciones importantes para no perderse

- *Cualquier resumen estadístico, como las medidas descriptivas o indicadores, se refiere a todo el conjunto de datos y no a alguno de los individuos en particular.*
- *Si el conjunto de datos constituye una población entonces los indicadores serán poblacionales y se los denomina parámetros. Si es una muestra entonces serán indicadores muestrales y se denominan estadísticos.*
- *Con respecto a las medidas se puede decir que:*
 - ✓ *La media o promedio solo es aplicable a las variables cuantitativas porque se determina, calculándola, en el espacio de los números reales.*
 - ✓ *Todo grupo de datos ordenables tiene mediana, dado que ella se determina en el espacio de los datos ordenados.*
 - ✓ *La moda puede determinarse para cualquier tipo de variable, cualitativa o cuantitativa, dado que ella se define en el espacio de los valores o categorías.*
 - ✓ *Las medidas de intensidad son aplicables a cualquier tipo de variable, cuantitativa o cualitativas dado que su espacio de definición es el espacio de las frecuencias. Pueden expresarse indistintamente como coeficientes o como porcentajes.*
- *El agrupamiento de los datos en forma de intervalos de clase permite construir el histograma y tener una noción de conjunto del grupo de datos y en algunos casos resolver la detección de la moda.*
- *No es conveniente su uso para el cálculo del resto de las medidas: media, mediana, percentiles, varianza, etc., porque virtualmente los datos se “matan” al ubicarlos en los intervalos y eso en estadística debe considerarse como deterioro de la información.*

De ser necesario calcularlas utilizando el agrupamiento en intervalos hay que tener presente que esas medidas serán de “menor calidad” que las que se podrían calcular si se contara con los datos originales.

Glosario 2

Moda o modo: es el valor de la variable que tiene mayor frecuencia simple.

Mediana: es el lugar geométrico del espacio de los datos ordenados que lo divide en dos partes iguales, representando cada una de ellas hasta el 50 % de los datos.

Percentil k: es el lugar geométrico del espacio de los datos ordenados, que deja por debajo el k% de los datos.

Datos centrales: son el 50 % del total y ocupan la zona central de una distribución (la caja del *box-plot*).

Datos adyacentes: son los que se encuentran en las inmediaciones de la caja, por fuera de ella y a ambos lados hasta una distancia de 1,5. $(Q_3 - Q_1)$.

Datos externos (*outliers moderados*): son los que se localizan a una distancia de entre $1,5 \cdot (Q_3 - Q_1)$ y $3 \cdot (Q_3 - Q_1)$ a ambos lados de la caja.

Datos lejanos (*outliers severos*): son los que se encuentran más allá de $3 \cdot (Q_3 - Q_1)$ de la caja.

Marca de clase: punto medio del intervalo de clase.

Límite inferior: es el menor valor numérico de un intervalo de clase.

Límite superior: es el mayor valor numérico que supera a todos los valores de un intervalo de clase (no es el mayor valor numérico del intervalo ya que este es abierto en su extremo superior).

Proporción: mide cuánto incide o “pesa” un valor o grupo de valores de la variable en el total.

Razón: medida comparativa del peso o presencia de un valor o grupo de valores, contra otro valor o grupo de valores de una variable.

Rango o recorrido de una distribución: diferencia entre el valor máximo y el valor mínimo de una variable cuantitativa. Mide la dispersión absoluta de la distribución sin tomar en cuenta los valores intermedios.

Rango intercuartílico: diferencia entre el tercer y el primer cuartil. Mide la dispersión absoluta del 50 % central de la distribución sin tomar en cuenta los valores de los datos que caen dentro de la caja.

Desvío de un dato: diferencia entre el valor del dato y la media de toda la distribución.

Varianza: promedio del cuadrado de los desvíos. Mide la dispersión de una distribución teniendo en cuenta todos los datos. Su magnitud es el cuadrado de la magnitud de la variable.

Desvío estándar: es la raíz cuadrada de la varianza. Su magnitud es igual a la de la variable. En distribuciones simétricas describe un entorno alrededor de la media que encierra aproximadamente las 2/3 partes del total de valores de la variable.

Coefficiente de variación: es la relación entre el desvío estándar y la media. Mide dispersión relativa. Es un coeficiente adimensional y sirve sobre todo para comparar, dentro de un mismo grupo de unidades de observación, la dispersión de variables que tienen distinta magnitud.

Problemas y situaciones prácticas 2

- 1- En un informe del INDEC se muestran datos relevados de 400 grandes empresas nacionales y extranjeras que brindaron información relevante respecto a la situación de la competitividad y las perspectivas y obstáculos a la inversión en distintos sectores. El relevamiento se realizó a través de la Encuesta Nacional sobre Competitividad e Inversión (ENCI, 31/01/2014). Accesible en: https://www.indec.gob.ar/ftp/indecinforma/nuevaWEB/cuadros/16/enci_31_01_14.pdf. A continuación, se presenta un resumen de parte de la información brindada en formato de gráfico o tabla.

Razones que motivan la nueva inversión	%	Factores relevantes frente a competidores del exterior (producto principal)	%
Reponer maquinaria y equipo por desgaste	15,9	Precio	36,1
Mejorar la calidad de los equipos / servicios que produce	12,7	Problemas de escala	5,2
Incrementar la productividad	17,9	Calidad	0,8
Desarrollar nuevos productos / servicios	9,6	Problemas de abastecimiento	12,9
Responder a incrementos en la demanda interna	9,4	Plazos de entrega	4,0
Eficiencia energética	7,6	Costos de entrega	11,2
Incrementar la cuota de mercado / conquistar nuevos mercados locales	6,3	Financiamiento	7,6
Alcanzar nuevos mercados de exportación	5,9	Condiciones y forma de pago	10,0
Sustitución de insumos importados	5,7	Otros	12,0
Responder a incrementos en la demanda externa	4,6		
Expandirse hacia otras actividades productivas	2,7		
Otras	1,7		
Fuente: Encuesta Nacional de Competitividad e Inversión		Fuente: Encuesta Nacional de Competitividad e Inversión	

Tabla 1

Tabla 2

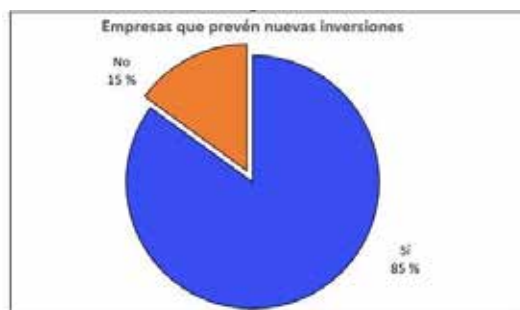


Gráfico 1

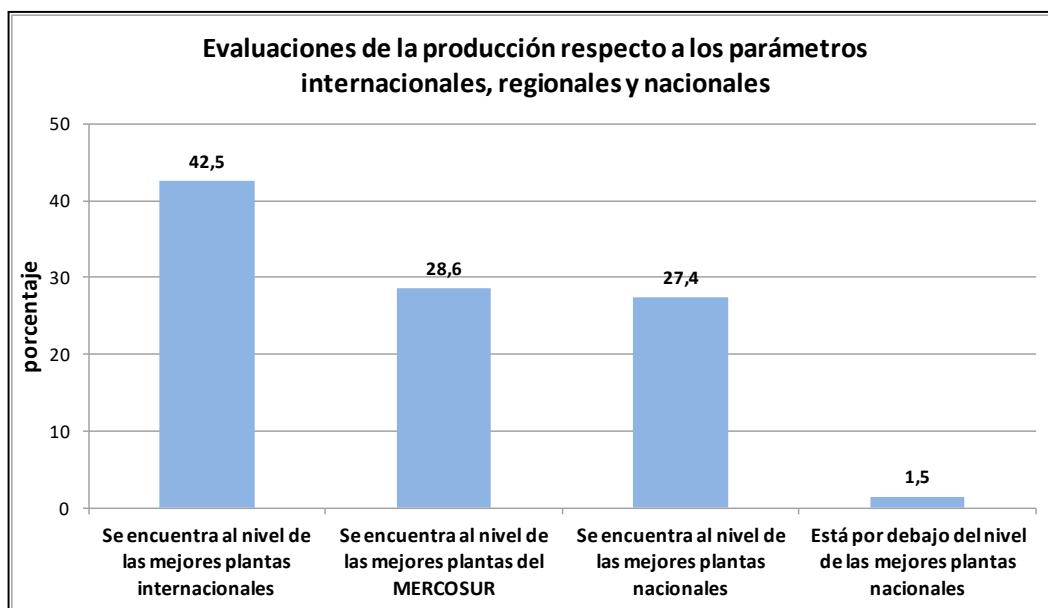


Gráfico 2

Factores relevantes frente a competidores del exterior (producto principal), según grupo de actividad (en %)									
Grupo de actividad	Precio	Problemas de escala / falta de capacidad instalada	Calidad	Problemas de abastecimiento de materias primas e insumos	Plazos de entrega	Costos de entrega (empaquete, flete, seguros, etc.)	Financiamiento	Condiciones y formas de pago	Otros
Alimentos, bebida y tabaco	9,7	1,6	0,0	4,1	0,8	6,8	2,4	2,9	4,9
Textiles, productos de vestir y calzados	2,0	0,4	0,0	0,8	0,4	0,0	0,0	0,0	0,5
Papel y productos de madera	2,0	0,4	0,0	0,4	0,0	0,8	0,4	0,5	0,0
Químicos, farmacéuticos y plástico	8,9	1,6	0,4	4,8	2,0	2,0	1,6	2,9	3,7
Productos metalúrgicos y minerales	3,7	0,0	0,4	0,0	0,0	0,0	0,8	1,3	0,5
Maquinarias y equipos	3,3	0,4	0,0	1,2	0,3	0,0	1,6	1,2	0,5
Vehículos y máquinas de transporte	4,5	0,4	0,0	0,8	0,4	1,6	0,8	0,9	0,4
Energía y comercio	0,4	0,0	0,0	0,8	0,0	0,0	0,0	0,0	0,8
Transporte y logística	1,6	0,4	0,0	0,0	0,0	0,0	0,0	0,5	0,9
Total	36,1	5,2	0,8	12,9	4,0	11,2	7,6	10,0	12,0

Fuente: Encuesta Nacional de Competitividad e Inversión

Tabla 3

Responder indicando la medida y el tipo de medida utilizada.

a) ¿Cuál es la proporción de empresas que opinan que el factor predominante frente a los competidores del exterior es el precio?

- b) ¿Qué porcentaje de empresas prevén nuevas inversiones respecto a las que no las prevén?
- c) ¿Cuál es la razón que motiva la nueva inversión que predomina?
- d) ¿Cuál fue el grupo de actividad y factor relevante frente a competidores extranjeros que poseen mayor frecuencia?
- e) ¿Cuál es la evaluación de la producción respecto a parámetros internacionales, regionales y nacionales que predominan?
- f) Por cada empresa que considera que el factor relevante frente a competidores extranjeros es el financiamiento, ¿cuántas opinan que es el precio?

2- Con los datos de la muestra del ejercicio 6 de P1 se ha confeccionado la siguiente tabla de distribución de frecuencias conjunta:

	PC	C	MC	Total
Bs. As.	4	3	1	8
Rosario	2	3	1	6
Córdoba	1	3	2	6
Total	7	9	4	20

- a) En la planta de Buenos Aires, ¿cuál es la proporción de empleados que están por lo menos conformes con la empresa?
- b) En la planta de Rosario, ¿cuál es la proporción de empleados que están a lo sumo conformes con la empresa?
- c) Hallar la proporción de empleados que trabajan en Rosario y están poco conformes con la empresa.
- d) Hallar la razón entre los empleados que están poco conformes con la empresa en las plantas de Buenos Aires y Rosario.
- e) En la planta de Córdoba, por cada empleado conforme, ¿cuántos están muy conformes?
- f) Hallar la moda bivariada e interpretar.

3- Para la variable “cantidad de empleados” del ejercicio 7 de la P1 se obtuvo la siguiente tabla de distribución de frecuencias:

x	fi	f%	F%
0	2	9	9
1	4	17	26
2	5	22	48
3	6	26	74
4	4	17	91
5	2	9	100

Determinar, identificando en cada caso la medida utilizada:

- a) La proporción de negocios que poseen al menos tres empleados.
- b) Por cada establecimiento que posee como mucho un empleado, ¿cuántos poseen al menos dos?
- c) El número medio de empleados por establecimiento.

- d) La cantidad de empleados que poseen como mínimo el 50 % de los negocios.
- e) La cantidad de empleados predominante para los establecimientos de la muestra.
- f) Cuántos empleados posee el 75 % de los establecimientos como mínimo.

4- Una compañía tiene un crédito abierto con un banco. El préstamo mostró los siguientes saldos mensuales finales, en miles de dólares, durante el año:

33,040 25,050 19,570 16,035 31,930 21,020 19,850 16,035 28,570 21,020 18,500 15,210.

Se sabe que la compañía será elegida para recibir una tasa reducida de interés si su saldo mensual promedio es mayor de USD 21,500; ¿cumple la compañía con tal requisito?

5- Un vendedor viajante hizo cinco viajes durante los meses de junio y julio de 2006. El número de días y el valor de las ventas de cada viaje son los siguientes:

Viaje	N.º de días	Valor de las ventas en \$	Ventas por día en \$
1	7	175.000	25.000
2	8	320.000	40.000
3	10	300.000	30.000
4	3	84.000	28.000
5	9	387.000	43.000

El gerente de ventas criticó la actuación del vendedor porque sus ventas medias por día solo ascendieron a \$33.200. Pero el vendedor dijo que el gerente de ventas estaba equivocado porque sus ventas medias por día fueron superiores a \$34.000. ¿Quién está en lo correcto?

6- En el ejercicio 2 de la P1 se indicaron las siguientes estadísticas demográficas de la Argentina en 2014:

Población total: 42.669.500 habitantes

Varones: 48,97 %

Mujeres: 51,03 %



Si la edad media de vida es 74,46 y 81,09 años para varones y mujeres respectivamente, hallar la edad media de vida en Argentina.

7- Discutir y analizar el grado de veracidad de los siguientes enunciados.

- a) El lugar geométrico de la media aritmética en el eje real es el centro de la distribución de datos.
- b) Todos los conjuntos de datos tienen promedio.
- c) En el cálculo de la media aritmética influyen los valores extremos.
- d) El valor del promedio supera a los valores que toman la mitad de los datos ordenados.

- 8- Colocar verdadero (V) o falso (F) según corresponda. En caso de ser falso explicar por qué.
- Para determinar la mediana se tiene en cuenta el valor de cada uno de los datos.
 - El valor que más se repite en un conjunto de datos se denomina media aritmética.
 - Para un conjunto ordenado de 50 observaciones, la posición de la mediana será el valor de la vigésimo quinta observación en el arreglo.
 - Los valores extremos de una distribución no influyen en la determinación de la mediana.
- 9- Una fábrica de electrodomésticos, para atender a la demanda de un mes determinado requirió de diferentes clases de mano de obra en las siguientes cantidades:

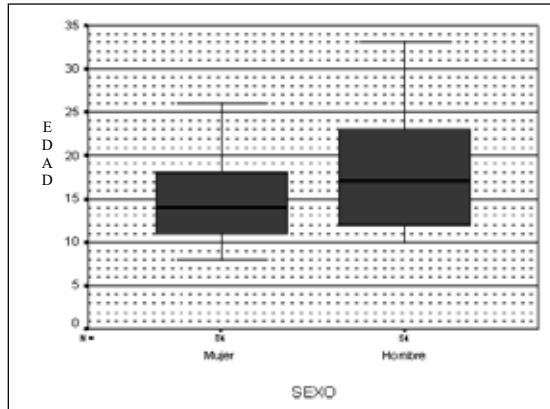
Mano de obra	Cantidad de horas	Tasa salarial (\$/hora)
Calificada	70	250
Semicalificada	130	120
No calificada	400	50

- ¿Cuál fue el costo medio de la mano de obra por hora? y ¿con qué varianza?
 - ¿Por qué es necesario elevar al cuadrado las diferencias entre la media y los valores de los que se obtuvo el promedio, cuando se calcula la varianza de la población?
 - Calcular la proporción de mano de obra semicalificada.
 - ¿Cuál es la razón entre la mano de obra calificada y la mano de obra no calificada?
 - ¿Cuál es la razón entre la mano de obra semicalificada y la calificada?
- 10- En un centro médico se observa especialmente la calidad de atención que reciben los pacientes. En un estudio sobre el tiempo de espera para ser atendidos se ha registrado la siguiente información:

Tiempo (min.)	Pacientes
0-5	12
5-10	19
10-15	28
15-20	36
20-25	25
25-30	17
30-35	14
35-40	9

- ¿Cuál es el tiempo de espera más habitual?
- ¿Hay diferencia entre el tiempo promedio de espera y el tiempo “central” del conjunto de datos?
- Calcular las medidas de dispersión.
- ¿El 65 % de los pacientes cuánto esperan como mínimo para ser atendidos?
- ¿Cuál es la espera máxima del 80 % de los pacientes?
- Calcular la amplitud intercuartil e interpretar.
- ¿Es simétrica la distribución?

- 11- Para realizar un estudio de mercado sobre determinado producto se consultó a una muestra de 200 personas de distintas edades, los siguientes *box-plots* se construyeron teniendo en cuenta la edad y el sexo de las personas que resultaron satisfechas con el producto. Analizarlos y extraer conclusiones.



12- Analizar en cada uno de los siguientes *box-plots*:

- la simetría de la distribución que representan,
- la distribución de los porcentajes de datos,
- la existencia de valores raros,
- el rango de la distribución y la ubicación de las vallas.

Gráfico 1

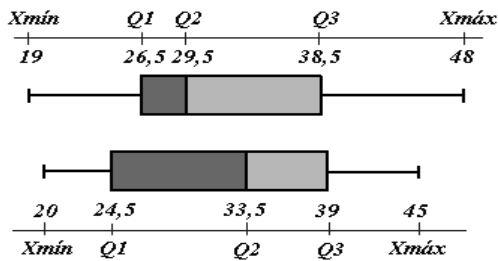


Gráfico 2

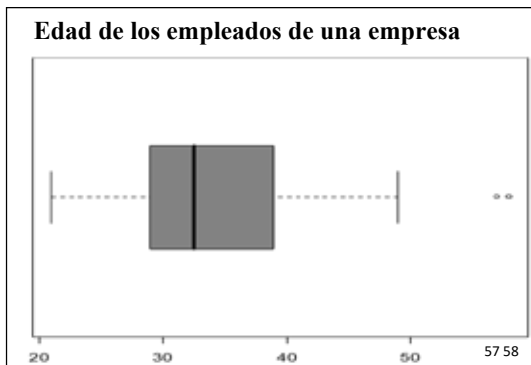
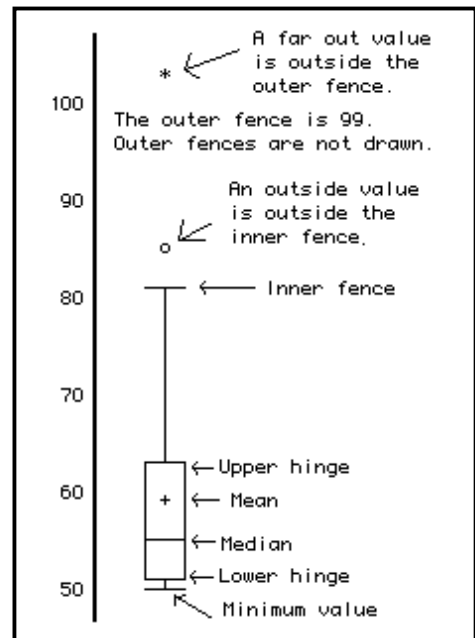
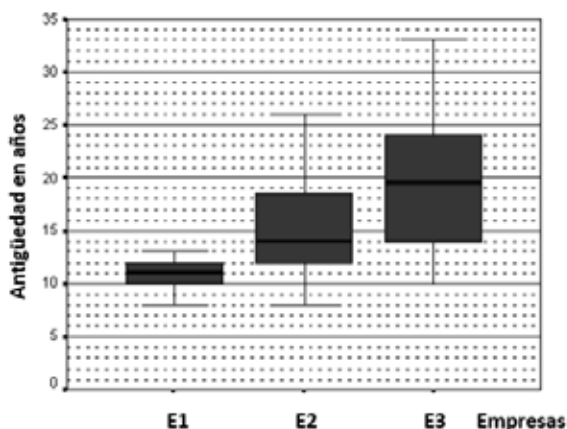


Gráfico 3

13- Con la finalidad de analizar la permanencia de empleados en los puestos de trabajo, se estudiaron tres empresas similares por la importancia y cantidad de años de actividad en el

mercado. El estudio consistió en consultar a todos los empleados de dichas empresas sobre la antigüedad que poseían en estas, los resultados se volcaron en los siguientes *box-plots*. Analizar los *box-plots* y extraer conclusiones.



Paréntesis histórico

A comienzos del siglo XIX el astrónomo Adolfo Quetelet –considerado el fundador de la estadística moderna– aplicó a las ciencias sociales los métodos estadísticos hasta entonces utilizados en las ciencias naturales, y contribuyó a la ampliación del campo de la estadística.

A partir de mediados de 1960, con el análisis exploratorio de datos (EDA), desarrollado por J. Tukey y otros, surgió un nuevo enfoque en el tratamiento de datos cimentado en el uso de la informática donde el soporte tecnológico permitió sostener una gran masa de datos y procesarlos en tiempo real, y contribuyó así al mejoramiento de la calidad de la información resultante.

El EDA condujo además a la aparición de productos gráficos tales como el diagrama de caja y bigotes, entre otros.

14- Los datos ordenados para la variable “cantidad de personal” del ejercicio 15 de la P1 son:

5	9	19	20	21	22	23	25	34	38
67	80	98	99	100	105	109	109	110	110
112	114	115	115	120	120	125	128	132	137
140	142	144	148	150	154	160	165	170	270

- Determinar la cantidad de personal predominante en las pymes.
- Calcular la cantidad media de empleados y el desvío.
- Determinar la cantidad mediana de empleados e interpretarla.
- Analizar el tipo de asimetría de la distribución de la cantidad de personal.
- Graficar un *box-plot* e identificar si en alguna pyme la cantidad de empleados es atípica.
- ¿Entre qué valores se encuentra el 50 % central de la cantidad de empleados?

15- Para la variable “cantidad de préstamos otorgados” del ejercicio 9 de la P1, los datos ordenados son los siguientes:

1	1	1	1	1	2	2
2	2	3	3	3	4	4
4	5	5	5	5	5	6
6	6	6	6	7	7	8
8	8	8	8	9	9	10
10	10	10	10	11	11	11
13	14	14	15	15	15	17
17	19	19	20	20	21	22

- a) ¿Cuántos préstamos otorgaron como máximo el 65 % de las entidades financieras?
- b) ¿Cuántos préstamos otorgaron como mínimo el 75 % de las entidades financieras?
- c) ¿Cuántas entidades otorgaron como mínimo 12 préstamos?
- d) ¿Cuántas entidades otorgaron menos de 10 préstamos?
- e) Graficar e interpretar el *box-plot*.
- f) Calcular las medidas necesarias y analizar la simetría de la distribución

16- Para la variable “horas semanales” del ejercicio 12 de la P1 se tiene la siguiente tabla de distribución de frecuencias:

H semanales	fi	f%	F%
30-32	5	5.15	5.15
32-34	22	22.78	27.83
34-36	30	30.93	58.76
36-38	28	28.87	87.63
38-40	10	10.31	97.94
40-42	2	2.06	100.00

- a) Hallar el número medio de horas trabajadas y el desvío.
- b) Hallar la razón entre los empleados que trabajan como mínimo 32 h y los que trabajan menos de 32 h.
- c) ¿Cuál es el máximo de horas trabajadas por el 40 % de los empleados?
- d) ¿Superior a cuánto es el número de horas trabajadas por el 80 % de los empleados?
- e) ¿Inferior a cuánto es el número de horas trabajadas por el 60 % de los empleados?
- f) ¿Cuál es el mínimo de horas trabajadas por el 25 % de los empleados?
- g) ¿Qué porcentaje de empleados trabajan
 - i) menos de 35,5 horas?
 - ii) más de 37,25 horas?
 - iii) entre 35,5 y 37,25 horas?

17- Una compañía está considerando la conveniencia de implantar dos programas de capacitación para sus empleados. Se seleccionaron dos grupos y a ambos se les impartió capacitación para realizar la misma tarea. En el grupo de trabajadores en que se implementó el programa A se requirieron en promedio 32,11 horas para capacitar a cada empleado, con una varianza

de 68,09 h². En el grupo en donde se llevó a cabo el programa B se necesitó un promedio de 19,75 horas para capacitar a cada empleado, con una varianza de 71,14 h². ¿Cuál programa resultó menos heterogéneo en sus resultados?

18- En una empresa se estudiaron los ingresos anuales de los ejecutivos y el de los empleados no calificados. Los resultados indicaron que el ingreso anual medio de los ejecutivos es de \$86.000 con un desvío de \$5.000 y el ingreso anual medio de los empleados no calificados es de \$7200 con un desvío de \$200. ¿Existe mayor consistencia en los sueldos de los trabajadores no calificados? Justificar la respuesta.

19-Para las variables “antigüedad” y “cantidad de personal” de la matriz del ejercicio 15 de la P1 se tiene:

	Promedio	Desvío
Antigüedad	6,475	4,814
Cantidad de personal	101,600	56,420

Determinar bajo qué variable la muestra es más heterogénea y justificar.

20- Se analizan dos comisiones de estudiantes que rindieron el mismo examen de Historia donde aparecían 12 preguntas y para aprobar se requerían al menos 6 correctas. En cada caso la cantidad de respuestas correctas observadas fueron:

Curso A	4	7	5	6	8	3	11	7	6	2	6	7	10	6	5	3	6	1	8
Curso B	5	8	9	5	4	9	12	4	4	3	6	5	8	5	7	6	7	3	9

- Para cada curso obtener el arreglo ordenado de datos.
- Calcular para cada curso las tres medidas de posición, comparar e interpretar.
- Calcular la proporción de aprobados en cada curso y comparar relativamente la cantidad de aprobados de ambos cursos.
- Comparar la homogeneidad de ambos conjuntos de datos.
- Obtener el *box-plot* en cada caso y extraer conclusiones.
- ¿Qué porcentaje de alumnos aprobados contestaron más de 7 preguntas correctamente?

21- Analizar si las siguientes afirmaciones son verdaderas *Siempre - A veces - Nunca*. Justificar y para las que no son verdaderas siempre pensar ejemplos donde no se cumplan.

- Un conjunto de datos muy asimétrico queda bien representado por una medida como la media aritmética.
- La mediana es la medida más adecuada de posición de un conjunto de datos.
- La varianza tiene como principal aplicación cuantificar la dispersión absoluta de los datos respecto a la media.
- El desvío puede interpretarse más adecuadamente en el contexto de un problema o situación por trabajar en las mismas unidades que la variable.
- El coeficiente de variación permite decidir si una distribución es o no homogénea.
- Si en un *box-plot* la mediana está ubicada en el centro de la caja eso es equivalente a que la distribución es simétrica.

Paréntesis reflexivo

En esta era llamada “la era de la información” la estadística se viene tornando en una herramienta cada vez más popular. La comunicación -pública y privada- se ha embebido de datos y gráficos estadísticos para “demostrar” aquello que se está informando y es por esto mismo que se hace imprescindible aprender a “leer y comprender” información estadística.

Desde el discurso muchas veces hay cifras ponderadas como muy grandes -o muy pequeñas- y estas apreciaciones eventualmente pueden no ser ingenuas ni inocentes, por eso también es preciso tomar los recaudos adecuados como, por ejemplo, revisar si se indican valores de referencia cuando de comparaciones se trata, o como en el caso de las proporciones, ver el marco o universo que representa el total de los casos considerados.

Es por esto que en estadística todos los resultados deben interpretarse, ya que cada uno de ellos informará de diferentes formas el estado de una situación de acuerdo con el contexto y significado atribuido a ese resultado.

Mix de problemas de procesamiento de datos y medidas estadísticas (P1 y P2)

22- Con los datos de la matriz del ejercicio 15 de la P1 se confeccionaron las siguientes tablas de distribución de frecuencias.

Nivel de endeudamiento	Cantidad de empresas
N	6
B	6
M	14
A	7
MA	7

Rubro	Cantidad de empresas
A	8
C	9
I	13
S	10

- Completar las tablas y graficar.
- ¿Cuántas pymes de las encuestadas tienen un endeudamiento a lo sumo bajo?
- ¿Qué nivel de endeudamiento tienen como máximo la mitad de las pymes?
- ¿Qué porcentaje de esas empresas tienen un endeudamiento superior a medio?
- ¿Cuál es el rubro que predomina en la muestra?
- Por cada pyme agrícola, ¿cuántas pymes industriales hay en la muestra?

23- En un molino arrocero, el arroz se comercializa en bolsas de distinto peso. En la tabla se presentan los contenidos (en kilos) de las bolsas de un embarque

Contenido	60	70	80	90	100
N.º de envases	146	105	84	75	50

- Identificar la variable en estudio y obtener el tamaño de la muestra.
- Completar la tabla de distribución de frecuencias y graficar.

- c) Calcular la proporción de envases con a lo sumo 80 kg.
- d) Por cada envase con más de 80 kg, calcular el número de envases que pesan a lo sumo 80 kg.
- e) ¿Cuál es el peso que predomina?
- f) ¿Cuál es el peso máximo del 50 % de las bolsas de ese embarque?
- g) Calcular el peso medio de las bolsas y el desvío estándar.
- h) Analizar simetría o asimetría de la distribución.
- i) ¿Entre qué valores se encuentra el 50 % central de los pesos?
- j) Si para otra muestra (muestra B) el peso medio es 65 kg con un desvío de 12,5 kg, ¿para qué muestra la variable “peso” es más homogénea?
- k) Si el precio medio de las bolsas es 91 USD con un desvío de 13 USD, ¿bajo qué variable (entre peso y precio) la muestra es más homogénea?

24- En la tabla se presenta la cantidad de libros técnicos vendidos por día en una librería para una muestra de 50 días (muestra A).

Libros	100-105	105-110	110-115	115-120	120-125	125-130
Días	2	4	6	8	18	12

- a) Completar la tabla de distribución de frecuencias y graficar.
- b) Calcular la proporción de días con por lo menos 110 ventas.
- c) Calcular la razón del número de días con por lo menos 110 ventas respecto del número de días con menos de 110 ventas. Interpretar.
- d) Calcular las medidas de tendencia central y analizar la simetría o asimetría de la distribución.
- e) Si para una muestra B la cantidad media de libros vendidos es 203 con una varianza de 104 libros², ¿cuál es la muestra menos homogénea?
- f) ¿Cuál es el número máximo de ventas en el 60 % de los días?
- g) ¿Cuántos libros se vendieron por lo menos en el 35 % de los días?
- h) ¿Cuál es la proporción de días con al menos 107 ventas?
- i) ¿Cuál es la proporción de días con entre 107 y 118 libros vendidos?

25- Los datos tabulados corresponden a terrenos fiscales de CABA que pasaron a manos privadas en 2017 y 2018 (Fuente: La Nación, 2/05/19).

- a) Identificar y clasificar las variables analizadas.
- b) Para cada variable listar los datos de menor a mayor.
- c) Para cada variable completar las tablas de distribución de frecuencias y graficar las frecuencias porcentuales, simples y acumuladas.
- d) Obtener conclusiones.

Superficie en m ²	Valor de la venta en millones de USD
16765,62	151,500
3200,00	50,100

3200,00	50,000
3200,00	40,500
7828,00	40,400
3200,00	40,200
5927,00	33,000
5887,00	26,000
3188,00	25,050
3954,00	23,025
3920,00	20,000
3100,00	15,000
5589,00	13,600
1647,00	8,150
1487,00	7,250
897,00	5,600
1183,00	5,600
897,00	5,450
705,00	4,550
727,00	4,500
751,00	4,050
370,00	2,200
619,00	0,360
742,00	0,300

- e) Para cada variable decidir si hay valores atípicos.
 f) Calcular la proporción de ventas de por lo menos 15 millones de USD.
 g) Por cada venta menor a 15 millones de USD, ¿cuántas hay de al menos 15 millones de USD?
 h) Para la variable superficie calcular los cuartiles e interpretar.
 i) ¿Bajo qué variable la muestra es más homogénea?

Valor (en millones de USD)	N.º de terrenos
0 – 20	
20 – 40	
40 – 60	
60 – 80	
80 – 100	
100 – 120	
120 – 140	

Superficie (en miles de m ²)	N.º de terrenos
0 – 2.5	
2.5 – 5	
5 – 7.5	
7.5 – 10	
10 – 12.5	
12.5 – 15	
15 – 17.5	

26- Se seleccionó una muestra aleatoria de los empleados de una empresa y sus edades clasificadas según el género son:

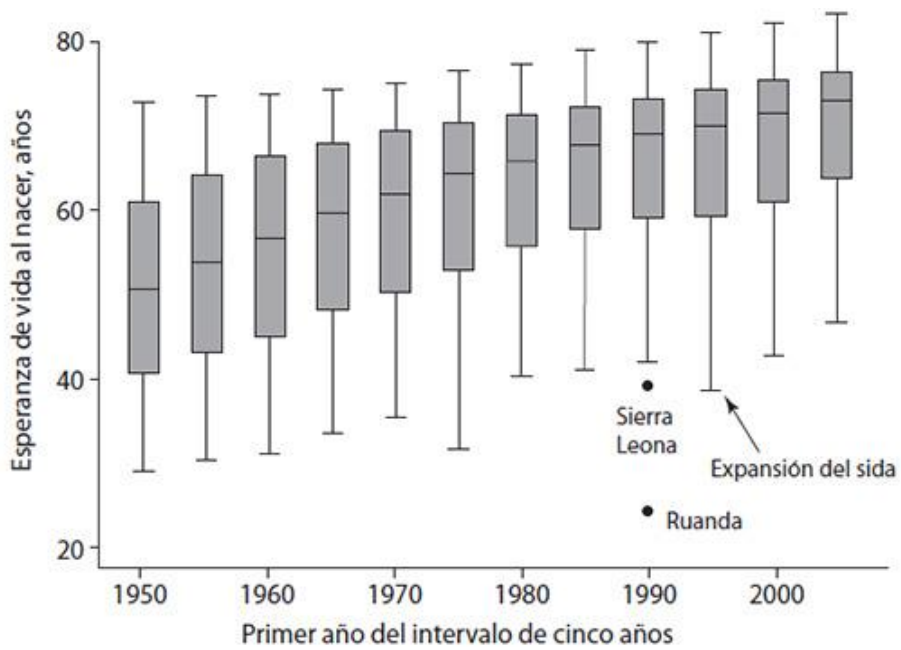
Hombres:	Mujeres:
18, 18, 19, 20, 20, 20, 21, 21, 22, 27, 34, 35, 36, 38, 40, 42, 43, 45, 49, 51, 51, 53, 56, 57, 59, 59	18, 18, 18, 18, 20, 21, 23, 24, 24, 25, 25, 26, 26, 27, 28, 28, 29, 29, 30, 30, 35, 40, 45, 50, 55

- a) Calcular la edad media de las mujeres, la edad media de los hombres, la edad media del total de los empleados y comparar.
 b) Calcular las medidas necesarias y analizar la simetría de las edades de las mujeres.
 c) Calcular la proporción de mujeres que tiene al menos 23 años.
 d) Por cada mujer con al menos 23 años, ¿Cuántas tienen menos de 23?
 e) Calcular la edad máxima del 80 % de las mujeres.
 f) Las edades de las mujeres, ¿presentan valores raros?

- g) Construir los *box-plots* correspondientes a las edades de hombres y mujeres y comparar.
- h) Decidir cuál es la muestra más homogénea.
- i) Agrupar los datos de hombres y mujeres sin distinción de género en 6 intervalos y completar la siguiente tabla, graficar convenientemente las frecuencias simples y acumuladas y extraer conclusiones.

Edades	fi	fr	f%	F%

27- Extraer conclusiones de la siguiente serie cronológica de gráficos de caja correspondientes a períodos de 5 años donde se relevó para diferentes países la esperanza de vida al nacer.



Esperanza de vida y su dispersión en el mundo

RESPUESTAS DE LAS ACTIVIDADES DE LA PRÁCTICA 2

Medidas estadísticas

- 1- a) $P_{\text{factor precio}} = 0,361$
El 36,1 % de las empresas opinan que el factor predominante frente a los competidores del exterior es el precio.
- b) $R_{\text{nuevas inv. Respecto no las prevén}} = 340/60 * 100 = 566,67$
Las empresas que prevén nuevas inversiones es el 566,67 % de las empresas que no las prevén.
- c) $M_o = \text{incrementar la productividad.}$
La razón predominante que motiva la nueva inversión de las empresas es incrementar la productividad.
- d) $M_o = \text{Alimentos, bebidas y tabaco y el "factor precio"}$.
El grupo de actividad "alimentos, bebidas y tabaco" y el "factor precio" frente a competidores extranjeros posee mayor frecuencia.
Interpretación: predominan las empresas dedicadas al grupo alimentos, bebidas y tabaco que tienen como factor relevante al precio frente a competidores extranjeros.
- e) $M_o = \text{la empresa se encuentra al nivel de las mejores plantas internacionales.}$
Predomina la evaluación de la producción que señala que la empresa se encuentra al nivel de las mejores plantas internacionales.
- f) $R_{\text{factor rel. precio respecto factor rel. financiamiento}} = 36,1/7,6 = 4,75$
Por cada empresa que considera que el factor relevante frente a competidores extranjeros es el financiamiento 4,75 empresas opinan que es el precio dicho factor.

Medidas y tipo de medida utilizadas en los ítems anteriores:

- a) Proporción. Medida de intensidad. b) Razón. Medida de intensidad. c) Modo. Medida de posición (tendencia central). d) Modo conjunto. Medida de posición. e) Modo. Medida de posición. f) Razón. Medida de intensidad.

- 2- a) $P = \frac{4}{8} \times 100 = 50 \%$
b) $P = \frac{5}{6} \times 100 = 83,33 \%$
c) $P = \frac{2}{20} \times 100 = 10 \%$
d) $R = \frac{4}{2} = 2$
e) $R = \frac{2}{3} = 0,67$

En la planta de Córdoba, por cada empleado conforme hay 0.67 empleados muy conformes.

- f) $M_o: \text{BA y PC}$

En la muestra predominan los empleados que trabajan en la planta de Buenos Aires y están poco conformes con la empresa.

- 3- a) $P(\text{negocios que poseen al menos tres empleados}) = P_{3 \text{ o más empleados}} = 12/23 = 0,52$

- En el centro comercial barrial la proporción de negocios que poseen al menos tres empleados es 0,52. Medida: Proporción.
- b) Por cada establecimiento que posee como mucho un empleado hay 2,83 establecimientos poseen al menos dos. Medida: razón ($R = 17/6 = 2,83$).
 - c) El número medio de empleados por establecimiento es 2,52 personas. Medida: media o promedio.
 - d) El 50 % de los negocios posee como mínimo 3 empleados. Medida: Mediana o Q2
 - e) La cantidad de empleados que predominan en la muestra es de 3 por establecimiento. Medida: modo.
 - f) El 75 % de los establecimientos, en el centro comercial, posee 1 o más empleados. Medida: Q_1 o P_{25}
- 4- Sí cumple con el requisito pues el promedio es 22.153 USD.
 - 5- El vendedor está en lo correcto pues en promedio vendió \$34.216,21
 - 6- $\bar{X} = 74.46 \times 0.4897 + 81.09 \times 0.5103 = 77.84$
La edad media de vida de los argentinos es 77.84 años.
 - 7- a) Falso. Es verdadera solamente cuando la distribución de los datos es simétrica porque en ese caso $\bar{X} = Me$
b) Falso, si la variable es cualitativa no puede determinarse el promedio.
c) Verdadero.
d) Falso. A veces es verdadera cuando la distribución es asimétrica a derecha.
 - 8- a) Falso, para determinar la Me sólo se tiene en cuenta el valor central (n impar) o los dos valores centrales (n par).
b) Falso, se denomina modo o moda.
c) Falso, la Me será el promedio de la vigésimo quinta y de la vigésimo sexta observación en el arreglo ordenado.
d) Verdadero.
 - 9- a) El costo medio de la mano de obra por hora para la fabricación del electrodoméstico es de \$88,50 con una varianza de 4.253,17
b) Es necesario elevar al cuadrado las diferencias entre la media y los valores de los que se obtuvo el promedio, cuando se calcula la varianza para evitar la suma cero de las diferencias entre cada dato y la media, $X_i - \bar{X}$.
c) La proporción de mano de obra semicalificada es del 21,67 %
d) La razón de la mano de obra no calificada respecto de la mano de obra calificada es 5,71
Interpretación: en la fábrica de electrodomésticos, durante ese mes, se utiliza por cada hora de mano de obra calificada 5,1 horas de mano de obra no calificada.
e) La razón de la mano de obra semicalificada respecto de la calificada es 1,86

Interpretación: en la fábrica de electrodomésticos, durante ese mes, se utiliza por cada hora de mano de obra calificada 1,86 horas de mano de obra semicalificada.

- 10- a) El tiempo de espera más habitual es 17,5 minutos.
b) Sí, hay diferencia pues el tiempo promedio de espera es $\bar{X} = 18,59$ minutos y el tiempo central es $Me = 17,92$ minutos.
c) Medidas de dispersión: $S = 9,44$ minutos; $S^2 = 89,20$ minutos² y $CV(x) = 50,78$ %
d) $P_{35} = 14,46$ minutos. El 65 % de los pacientes esperan como mínimo 14,46 minutos para ser atendidos.
e) $P_{80} = 27,35$ minutos. El 80 % de los pacientes espera como máximo 27,35 minutos.
f) Amplitud intercuartil $d = Q_3 - Q_1 = 25 - 11,61 = 13,39$ minutos.
Interpretación: el 50 % central del tiempo de espera de los pacientes para ser atendidos presenta una diferencia de 13,39 minutos. El rango del tiempo de espera del 50 % central de los pacientes es de 13,39 minutos.
g) No es simétrica la distribución, presenta una asimetría a derecha.

11- Análisis de los *box-plots* y algunas conclusiones.

Distribución de la edad de las mujeres satisfechas con el producto:

$$X_{\min} = 8 \quad X_{\max} = 26 \quad Q_1 = 11 \quad Q_2 = 14 \quad Q_3 = 18$$

se determinan d y las vallas y se encuentra que

- *No existen edades raras* en el grupo de mujeres encuestadas.

Distribución de la edad de los hombres satisfechos con el producto:

$$X_{\min} = 10 \quad X_{\max} = 33 \quad Q_1 = 12 \quad Q_2 = 17 \quad Q_3 = 23$$

se determinan d y las vallas y se encuentra que

- *No existen valores raros u outliers* entre los varones encuestados.
- Ambas distribuciones son asimétricas hacia la derecha.
- El grupo de mujeres que resultó satisfecho con el producto en general es más joven que el de los hombres que dijeron estar satisfechos con este.
- El 50 % central de las mujeres tienen edades comprendidas entre 11 y 18 años mientras que la mitad central de los hombres varían entre los 12 y 23 años.
- El cuarto de mayor edad de los hombres está más disperso que el cuarto de mayor edad de las mujeres.

12- Análisis de *box-plots*

Gráfico N.º 1:

El *box-plot* superior indica una asimetría hacia la derecha, en cambio, el inferior una asimetría hacia la izquierda.

El 50 % central de la distribución superior es menos disperso que el de la inferior. Allí mismo, el primer cuarto de la distribución superior está menos concentrado que el segundo, en cambio, el *box-plot* inferior evidencia lo contrario.

Con respecto a los bigotes, el 25 % inicial y también el final de la distribución que representa el *box-plot* superior se encuentra más disperso que para el inferior.

Ninguna de las distribuciones presenta valores raros.

Los datos representados en el primer gráfico tienen un rango o amplitud total mayor al del segundo marcando diferencias en su dispersión absoluta. (29 y 25 respectivamente).

Gráfico N.º 2:

Este *box-plot* muestra una asimetría hacia la derecha, la $Me < \bar{X}$ pues $55 < 60$.

El 50% central de los datos se encuentra entre 51 y 63 aproximadamente. Dentro de la mitad central de los datos el primer cuarto es menos disperso que el segundo.

El 25 % inferior de la distribución es menos disperso que el 25 % superior, el cual está comprendido entre 63 y 102 aproximadamente. El 25 % de los datos que presentan valores más bajos se encuentra mucho más concentrados o sea con menor dispersión absoluta que el 25 % más alto.

La distribución representada posee dos valores raros que son, aproximadamente en 85 y 102 (102 es extremo lejano).

El rango de la distribución es aproximadamente 52 (entre 50 y 102).

Las vallas superiores se encuentran, la interna en 81 y la externa en 99 aproximadamente. Se pueden calcular con la distancia intercuartil.

Gráfico N.º 3:

La distribución de las edades de los empleados de una empresa presenta una asimetría hacia la derecha.

El 50 % central de la distribución de las edades está comprendido entre 28 y 38 años aproximadamente.

Las edades 57 y 58 son raras.

El rango de edades de los empleados de la empresa es aproximadamente 36 años.

13- El *box-plot* de la empresa E1 muestra una asimetría a la izquierda mientras que los otros dos la asimetría es a la derecha.

En la empresa E1, si bien los empleados tienen menos antigüedad que en las otras dos, es la que presenta menor dispersión (los valores están más concentrados, comprendidos entre 8 y 13 años). El menor rango de valores implica que hay una gran rotación de empleados.

En la empresa E2 los empleados tienen antigüedades que van de los 8 años a 26 años, hay una menor rotación de los trabajadores que en la E1.

La antigüedad de los empleados de la empresa E3 está comprendida entre 10 y 33 años.

La mitad de los trabajadores de E1 como máximo tiene 11 años en la empresa mientras que en la E2 la mitad como máximo tienen 14 años de antigüedad y en la E3 el 50 % de los trabajadores tienen hasta 19 años de antigüedad.

El 100 % de la antigüedad de los trabajadores de E1 es menor que la mitad de la de los de la E2 y que un cuarto de la E3.

El 50 % central de las antigüedades de los empleados de E3 es más disperso que el 50 % correspondiente de las otras dos empresas. Dentro del cuarto de trabajadores con mayor cantidad de años en la E3 hay mayor dispersión que el correspondiente a la E2 y este a su vez mayor que el de la E1.

14- a) Corresponde agrupar datos en intervalos, por ejemplo:

Cant. emp.	5-50	50-95	95-140	140-185	185-230	230-275
N.º pymes	10	2	18	9	0	1

$$I.Mo = [95, 140)$$

$$Mo = 117,5$$

Interpretación: basándonos en las observaciones se concluye que predominan las pymes que tienen entre 98 y 137 empleados.

$$b) \bar{X} = 101,6$$

$$s = 56,42$$

Estos resultados se obtuvieron los datos sin agrupar, con datos agrupados se obtienen medidas de menor calidad.

$$c) Me = \frac{110+112}{2} = 111$$

La mitad de las empresas tienen como máximo 110 empleados, la otra mitad como mínimo 112.

d) Siendo $\bar{X} < Me < Mo$, la distribución es asimétrica a la izquierda.

$$e) Q_1 = 52,5 \quad Q_3 = 138,5 \quad RI = 86 \quad d = 129$$

$$VII = 52,5 - 129 = -76,5$$

$$VIS = 138,5 + 129 = 267,5$$

Por lo tanto, 270 es una cantidad de empleados atípica.

f) El 50 % central de los datos (los 20 datos centrales) se encuentran entre 67 y 137

15- Con respecto al ejercicio 9 de la P1.

a) El 65 % de las entidades financieras otorgaron como máximo 10 préstamos

b) El 75 % de las entidades financieras otorgaron como mínimo 4 préstamos.

c) Catorce entidades otorgaron como mínimo 12 préstamos.

d) Treinta y cuatro entidades otorgaron menos de 10 préstamos.

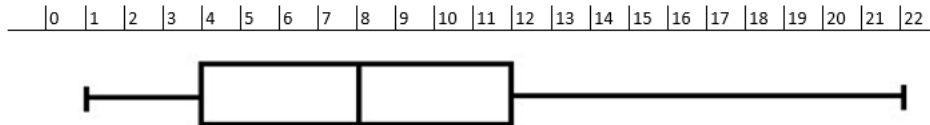
e) A partir del arreglo de datos ordenados del ítem a) se obtiene el *box-plot* identificando:

$$X_{\min.} = 1 \quad X_{\max.} = 22 \quad Q_1 = 4 \quad Me = 8 \quad Q_3 = 12 \quad d = 8$$

Vallas: $V_{ii} = 4 - 1,5 \cdot 8 = \text{negativo}$ por lo que no existen valores raros en la parte inferior de la distribución.

$V_{si} = 12 + 12 = 24$ supera al $X_{m\acute{a}x.}$ por lo que no existen valores raros en la parte superior de la distribución.

En el *box-plot* de la distribución del número de préstamos otorgados por las entidades financieras se observa una asimetría a la derecha.



El 50 % central de las entidades financieras otorgaron entre 4 y 11 préstamos.

El 25 % de las entidades financieras que menos cantidad de préstamos concedieron como mucho otorgaron 4 préstamos, este cuarto de la distribución tiene menos dispersión absoluta que el cuarto superior.

El 25 % de las entidades financieras que más préstamos otorgó como mínimo autorizaron 13 préstamos.

f) $Me = 8$ préstamos

$\bar{X} = 8,66$ préstamos

$Mo = 2,6$ (considerando $I.Mo = [1; 4,2)$)

Siendo $Mo < Me < \bar{X}$, la distribución es asimétrica a la derecha.

16- Utilizando los datos y los gráficos confeccionados en el ejercicio 12 de la P1.

a) $\bar{X} = 35,45\$$ $s = 2,264\$$

b) Por cada obrero que trabaja menos de 32 h hay 18,4 obreros que trabajan 32 h o más.

c) $P_{40} = 34 + 2 \frac{40 - 27,83}{30,93} = 34,79$

El 40 % de los empleados trabaja como máximo 34,79 horas.

d) $P_{20} = 33,3$

El 80 % de los empleados trabaja más de 33,3 horas.

e) $P_{60} = 36,09$

El 60 % de los empleados trabaja menos de 36,09 horas.

f) $P_{75} = 37,12$

el 25 % de los empleados trabaja como mínimo 37,12 horas.

g) Porcentaje de obreros que trabajan:

i) menos de 35,5 h.

$$35,5 = PK = 34 + 2x(K\% - 27,83\%) / 30,93 \%$$

$$\text{despejando } K\%: K\% = (35,5 - 34) \times 30,93 / 2 + 27,83 = 51,03 \%$$

El 51,03 % de los obreros trabajan menos de 35,5 horas.

ii) más de 37,25 horas.

$$37,25 = PK = 36 + 2x(K\% - 58,76\%) / 28,87 \%$$

$$\text{despejando } K\%: K\% = 76,8 \%$$

$$\text{Respuesta: } 100 - K\% = 23,2 \%$$

El 23,2 % de los obreros trabajan más de 37,25 horas.

iii) trabajan entre 35,5 y 37,25 horas.

$$K_{ii} - K_i = 76,8 \% - 51,03 \% = 25,77 \%$$

El 25,77 % de los obreros trabajan entre 35,5 y 37,25 horas.

17- El programa del grupo A resultó menos heterogéneo en sus resultados.

(CVGrupo A= 25,69 % y CVGrupo B= 42,68 %).

18- Sí, existe mayor consistencia en los sueldos de los trabajadores no calificados pues su CV es menor. (CVEjecutivos=5,81 %; CVNo calificados= 2,78 %).

19- $CV_{ant} = \frac{4.814}{6.475} = 0,743$

$$CV_{ant} = \frac{56.42}{101.6} = 0,553$$

La muestra es más heterogénea bajo la variable “antigüedad”.

20- a) Arreglo ordenado de datos:

Curso A	1	2	3	3	4	5	5	6	6	6	6	6	7	7	7	8	8	10	11
Curso B	3	3	4	4	4	5	5	5	5	6	6	7	7	8	8	9	9	9	12

b) Calcular para cada curso las tres medidas de posición, comparar e interpretar.

Curso A: $M_o = 6$ $M_e = 6$ y $\bar{X} = 5,84$ Curso B: $M_o = 5$ $M_e = 6$ y $\bar{X} = 6,26$

Si bien en ambos grupos la mitad de los estudiantes como máximo respondieron en forma correcta 6 preguntas, en el curso A predominaron los que respondieron bien 6 preguntas y en el B los que lo hicieron en 5 de las 12 preguntas. La distribución de la cantidad de respuestas correctas de la comisión A es asimétrica a izquierda y la de la B hacia la derecha pues los promedios de la 1.ª resultaron menores que las otras dos medidas y en el 2.º superó a ambas medidas.

c) Calcular la proporción de aprobados en cada curso y comparar relativamente la cantidad de aprobados de ambos cursos.

Proporción de aprobados: A) $P = 0,63$ B) $P = 0,53$

Razón de alumnos aprobados de curso A respecto a los aprobados del B: $R = 1,2$

Por cada alumno del curso B que aprobó el examen de historia en el A aprobaron 1,2 alumnos. Es decir que hubo un 20 % más de aprobados en el curso A que en el B.

d) Comparar la homogeneidad de ambos conjuntos de datos.

Curso A: $CV(x) = 43,57 \%$ Curso B: $CV(x) = 38,71 \%$

Ambas distribuciones no son homogéneas. La distribución de la cantidad de respuestas correctas de la comisión A es más heterogénea que la de la comisión B.

e) Obtener el *box-plot* en cada caso y extraer conclusiones.

- f) ¿Qué porcentaje de alumnos aprobados contestaron más de 7 preguntas correctamente?
El 45,45 % de los estudiantes aprobados contestaron más de 7 preguntas correctamente (aprobados de ambas comisiones) $P = 10/22 = 0,4545$

21- Analizar si las siguientes afirmaciones son verdaderas siempre-a veces-nunca. Justificar.

- Un conjunto de datos muy asimétrico queda bien representado por una medida como la media aritmética. Nunca, en ese caso la mediana es más representativa que el promedio.
- La mediana es la medida más adecuada de posición de un conjunto de datos. A veces, cuando los datos son muy asimétricos.
- La varianza tiene como principal aplicación cuantificar la dispersión absoluta de los datos respecto a la media. Siempre.
- El desvío puede interpretarse más adecuadamente en el contexto de un problema o situación por trabajar en las mismas unidades que la variable. Siempre.
- El coeficiente de variación permite decidir si una distribución es o no homogénea. A veces, algunos autores lo utilizan así, pero en la cátedra lo utilizamos para comparar la homogeneidad de dos o más distribuciones: a menor CV mayor homogeneidad.
- Si en un *box-plot* la mediana está ubicada en el centro de la caja eso es equivalente a que la distribución es simétrica. A veces, si el *box-plot* es asimétrico entonces la distribución también lo es, pero *box-plot* simétrico no implica que la distribución también lo sea.

Mix de ejercicios de prácticas 1 y 2 (algunas respuestas)

22- b) 12

- $Me = M$
- $P = 14/40 \cdot 100 = 35 \%$
- $Mo = I$
- $R_{I/A} = 13/8 = 1,625$

23- a) Variable: peso de las bolsas

- $n = 460$
- 72,8 %
- $R = 335/125 = 2,68$
- $Mo = 60$
- $Me = 70$
- $\bar{x} = 75,17$
 $s = 13,66$
- Siendo $Mo < Me < \bar{x}$, la distribución es asimétrica a derecha
- $Q_1 = 60$
 $Q_3 = 90$
El 50 % central de las bolsas pesa entre 60 y 90 kg.
- $CV_A = \frac{13,66}{75,17} = 0,18$ $CV_B = \frac{12,5}{65} = 0,19$
La muestra más homogénea es la A (la que tiene menor CV)

$$k) CV_{precio} = \frac{13}{91} = 0,14$$

La muestra es más homogénea bajo la variable “precio”.

24- b) $P = 44/50 \cdot 100 = 88 \%$

c) $R = 44/6 = 7,33$

Por cada día con menos de 110 ventas hubo 7,33 con 110 o más.

d) $Mo = 122,5$

$$Me = 120 + 5 \cdot (50-40)/36 = 121,4$$

$$\bar{x} = 119,7$$

Siendo $\bar{x} < Me < Mo$, la distribución es asimétrica a izquierda

e) $CV_A = \frac{7.01}{199.7} = 0,059$

$$CV_B = \frac{10.2}{203} = 0,050$$

La muestra A es menos homogénea.

f) $P_{60} = 120 + 5 \cdot (60-40)/36 = 122,78$

En el 60 % de los días se vendieron como máximo 123 libros.

g) $P_{65} = 120 + 5 \cdot (65-40)/36 = 123,5$

En el 35 % de los días se vendieron como mínimo 124 libros.

h) $107 = 105 + 5 \cdot (K-4)/8$. Despejando $K = 7,2 \%$

Respuesta: $100-7,2 = 92,8 \%$

i) $118 = 115 + 5 \cdot (K - 24)/16$. Despejando, $K = 33,6 \%$

Respuesta: $33,6 - 7,2 = 26,4 \%$

25- a) X_1 : superficie X_2 : ventas

e) Para la variable “ventas”

$$Q_1 = 5 \quad Q_3 = 36,6 \quad RI = 31,6 \quad VIS = 83,85$$

Por lo tanto, 151,5 millones de dólares es un valor atípico.

f) $P = 50 \%$

g) $R = 1$

h) $Q_1 = 746,5$ (es el promedio entre el cuarto y quinto valor de los datos ordenados).

La superficie máxima del 25 % de los terrenos vendidos es 742. Y la mínima del 75 % restante es 751

26- $\bar{x}_h = 36,69$

$$\bar{x}_m = 28,48$$

$$\bar{x} = 32,67 \quad \text{Se verifica } \bar{x} = \frac{36.69 \cdot 26 + 28.48 \cdot 25}{51} = 32,6$$

PRÁCTICA 3

Números índices

El número índice es una medida de intensidad y el objetivo de su cálculo es el de lograr indicadores que pongan de manifiesto, en forma cuantitativa y abreviada, las variaciones relativas de los valores que adopta un fenómeno en dos o más circunstancias distintas en el tiempo o en el espacio (en su acepción más amplia).

GPS ESTADÍSTICO 3

Orientaciones importantes para no perderse:

- *Si bien cada medida de posición es única para un mismo conjunto de datos, es el fenómeno en estudio el que debe incidir en la decisión de elegir cuál o cuáles son las más indicadas u óptimas para representarlo.*
- *Para las medidas de intensidad, en particular para los números índice, no solo incide el fenómeno, sino que además la propiedad de unicidad no se cumple.*
- *Por lo tanto, la elección de una medida en el marco de un problema de estudio está influenciada por subjetividades, hecho que conlleva el riesgo de manipulación estadística y/o la mala información.*

Glosario 3

Índice simple: compara un valor consigo mismo, en dos momentos distintos en el tiempo o en dos situaciones distintas en el espacio (en un sentido amplio).

Índice compuesto: compara un grupo de valores con respecto a sí mismo, en dos momentos distintos en el tiempo o en dos situaciones distintas en el espacio.

Media de relativos simples (de precios, cantidades o valores): es un índice compuesto y resulta de promediar los relativos simples de cada valor.

Índice de agregados no ponderados: es un índice compuesto que resulta del cociente entre las sumas o agregaciones (de precios o cantidades) correspondientes a las dos situaciones distintas.

Índice de agregados ponderados: es un índice compuesto de precios que resulta del cociente entre las sumas o agregaciones de precios -ponderados con cantidades- correspondientes a las

dos situaciones distintas. Según el criterio de ponderación se califican en índice de Laspeyres, de Paasche, de Fisher, etc.

Índice deflactor: es cualquier índice general de precios (IPC, IPIM, ICC) usado para corregir el valor monetario de una magnitud económica, lo que permite efectuar comparaciones a lo largo del tiempo (por ejemplo, el salario).

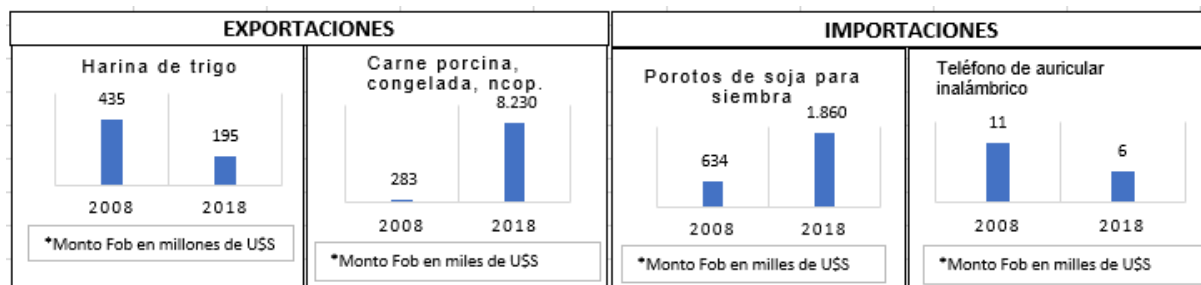
IPC: índice de precios al consumidor. Se usa para analizar la evolución del costo de vida.

IPIM: índice de precios industriales mayoristas. En realidad, es un sistema de varios índices (SIPIM).

ICC: índice de costos de la construcción. En rigor no es un índice de precios sino de costos.

Problemas y situaciones prácticas 3

1- La siguiente información referida al comercio exterior de Argentina se encuentra a disposición en la página del INDEC: <https://comex.indec.gob.ar>, en la sección de consulta de comercio exterior de bienes.



Analizar la información y extraer conclusiones.

2- La siguiente tabla muestra los volúmenes anuales exportados de productos lácteos de la Argentina (en toneladas) publicado por www.senasa.gob.ar en la sección: informes y estadísticas.

	2013	2014	2016	2017	2018
Leche polvo entera	148.159	132.723	91.371	58.535	69.081
Leche polvo otras	30.987	32.201	20.180	16.735	9.780
Leche polvo descrema	25.748	16.531	11.943	13.523	7.980
Leche fluida	402	746	340	304	375
Total	205.296	182.201	123.834	89.097	87.216

Nota: los datos del año 2018 corresponden a una anualización, la información del año 2015 no se encontraba disponible.

Se pide:

a) Calcular para cada tipo de leche exportada los relativos simples de cada año respecto al anterior e interpretar los resultados de las variaciones porcentuales obtenidas.

- b) Calcular el índice compuesto de agregados no ponderados: i) del año 2017 respecto a 2013 y ii) del año 2018 respecto a 2016
- c) Calcular el índice compuesto a partir del promedio de relativos simples: i) del año 2017 respecto a 2013 y ii) del año 2018 respecto a 2016
- d) Compara los resultados obtenidos en b) y c).

3- A continuación, se muestra la cantidad de transferencias de automóviles correspondientes a 5 provincias en los períodos indicados.

	1.º sem. 2017	2.º sem. 2017	1.º sem. 2018	2.º sem. 2018	1.º sem. 2019
Bs. As.	258.495	351.648	323.869	314.365	305.383
CABA	65.500	91.699	82.812	81.071	76.747
Córdoba	73.563	104.050	97.499	97.439	94.431
Entre Ríos	27.158	36.471	35.120	33.617	33.463
Mendoza	33.982	48.274	46.630	44.436	43.598
Santa Fe	61.965	88.232	81.153	78.559	74.557

Fuente: Cámara del Comercio Automotor, Estadísticas, www.cca.org.ar

Se pide:

- a) Calcular para cada provincia los relativos simples de cada semestre respecto al anterior e interpretar los resultados de las variaciones porcentuales obtenidas.
- b) Calcular el índice compuesto de agregados no ponderados: i) del 1.º semestre 2019 respecto al 1.º semestre 2017 y ii) del 2.º semestre 2018 respecto al 2.º semestre 2017.
- c) Calcular el índice compuesto a partir del promedio de relativos simples: i) del 1.º semestre 2019 respecto al 1.º semestre 2017 y ii) del 2.º semestre 2018 respecto al 2.º semestre 2017

4- Un fabricante asiático de artículos de cotillón recopiló la siguiente información sobre las ventas realizadas en nuestro país.

Tipos/año	Precios anuales promedios (miles)			Unidades vendidas (x 1.000)		
	2016	2017	2018	2016	2017	2018
Vasos	51	62	65	50	60	65
Copas	45	53	56	80	85	88
Platos	53	65	67	75	82	85
Cubiertos	48	54	58	45	47	52

Se pide:

- a) Calcular los índices de Paasche, Laspeyres y Fisher para 2017 usando como período base 2016. Compararlos.
- b) Calcular los índices de Paasche, Laspeyres y Fisher para 2018 usando como período base 2016. Compararlos.

- 5- Se seleccionaron tres familias bonaerenses y se les consultó sobre el total de kWh anuales consumidos y el precio del kWh promedio anual que se encuentra detallado en la factura del servicio eléctrico en los últimos tres años.

La información brindada se compiló en el siguiente cuadro:

	Total anual de kWh consumido (en miles)			Precio anual promedio x kWh		
	2016	2017	2018	2016	2017	2018
Familia A	6,68	6,75	6,92	0,043	0,76	1,35
Familia B	7,25	6,92	7,12	0,046	0,77	1,49
Familia C	7,05	7,25	7,7	0,047	0,75	1,52

Calcular los índices de Paasche, Laspeyres y Fisher para 2018 usando como período base 2016. Compararlos.

- 6- El índice de precios internos al por mayor (IPIM) y el salario de los empleados madereros división aglomerado para los siguientes semestres fueron:

Calcular las variaciones porcentuales de los salarios reales e interpretar los valores correspondientes a:

- Diciembre 2018 respecto a diciembre 2016
- Junio 2019 respecto a junio 2017
- Diciembre 2017 respecto de junio 2016

	IPIM* (Base dic. 2016 = 100)	1.º oficial - aglomerado (salario por hora)
Jun-16	94,22	55,91
Dic-16	100	64,15
Jun-17	111,85	77,17
Dic-17	124,79	85,59
Jun-18	144,81	73,29
Dic-18	184,26	104,84
Jun-19	225,53	124,74

Datos de la Unión de Sindicatos de la Industria Maderera de la República Argentina y de la Federación Argentina de Consejos profesionales de Ciencias Económicas

- 7- El índice de precios al consumidor y el promedio de salarios por hora en ciertas industrias seleccionadas son:

Año	IPIM	Salarios por hora (en \$)		
		Servicios	Comercio al menudeo	Manufactura
2016	100,00	55,09	42,26	65,15
2017	124,79	76,64	55,75	87,79
2018	184,26	98,57	67,80	104,84
2019	225,53	129,98	90,40	137,10

- ¿Qué ocurrió a los salarios reales de un empleado representativo en cada una de las tres industrias seleccionadas, entre el período 2016 y los períodos 2017 y 2018?
- ¿Qué sucedió al poder adquisitivo del dinero entre el año 2017 y el año 2019?
- El ingreso mensual de José Gómez en 2017 fue de \$20.040 y de \$31.900 en 2019. Explicar qué le pasó a su ingreso real en dinero.

RESPUESTAS DE LAS ACTIVIDADES DE LA PRÁCTICA 3

Números índices

1- $I = 195/435 \cdot 100 = 44,83 \%$

Diferencia: $44,83 \% - 100 \% = -55,17 \%$

Interpretación: se observa un decrecimiento en las exportaciones de la harina del 55,17 % comparando 2018 con base (con respecto a) 2008

$I = 8230/283 \cdot 100 = 2908,13 \%$ $2908,13 \% - 100 \% = 2808,13 \%$

Interpretación: se observa un crecimiento en las exportaciones de la carne del 2808,13 % comparando 2018 con base (con respecto a) 2008

$I = 1860/634 \cdot 100 = 293,38 \%$ $293,38 \% - 100 \% = 193,38 \%$

Interpretación: se observa un crecimiento en las importaciones de la soja del 193,38 % comparando 2018 con base (con respecto a) 2008

$I = 6/11 \cdot 100 = 54,56 \%$ $54,56 \% - 100 \% = -45,44 \%$

Interpretación: se observa un decrecimiento en las importaciones de los teléfonos del 45,44 % comparando 2018 con base (con respecto a) 2008

2- a)

	2013/13	2014/13	2016/14	2017/16	2018/17
Leche polvo entera	100	89,58	68,84	64,06	118,07
Leche polvo otras	100	103,92	62,67	82,93	58,44
Leche polvo descrema	100	64,20	72,25	113,23	59,01
Leche fluida	100	185,57	45,58	89,41	123,36

	2013/13	2014/13	2016/14	2017/16	2018/17
Leche polvo entera	0	-10,42 %	-31,16 %	-35,94 %	18,03 %
Leche polvo otras	0	3,92 %	-37,33 %	-17,07 %	-41,56 %
Leche polvo descrema	0	-35,80 %	-27,75 %	13,23 %	-40,99 %
Leche fluida	0	85,57 %	-54,42 %	-10,59 %	23,36 %

Conclusiones:

-Leche en polvo entera:

2014-2013: en el año 2014 las exportaciones de leche en polvo entera disminuyeron un 10,42 % con respecto al 2013

2016-2014: en el año 2016 las exportaciones de leche en polvo entera disminuyeron un 31,16 % con respecto al 2014

2017-2016: en el año 2017 las exportaciones de leche en polvo entera disminuyeron un 35,94 % con respecto al 2016

2018-2017: en el año 2018 las exportaciones de leche en polvo entera aumentaron un 18,03 % con respecto al 2017

-Leche en polvo otras:

2014-2013: en el año 2014 las exportaciones de leche en polvo otras aumentaron un 3,92 % con respecto al 2013

2016-2014: en el año 2016 las exportaciones de leche en polvo otras disminuyeron un 37,33 % con respecto al 2014

2017-2016: en el año 2017 las exportaciones de leche en polvo otras disminuyeron un 17,07 % con respecto al 2016

2018-2017: en el año 2018 las exportaciones de leche en polvo otras disminuyeron un 41,56 % con respecto al 2017

-Leche en polvo descremada:

2014-2013: en el año 2014 las exportaciones de leche en polvo descremada disminuyeron un 35.80 % con respecto al 2013

2016-2014: en el año 2016 las exportaciones de leche en polvo descremada disminuyeron un 27.75 % con respecto al 2014

2017-2016: en el año 2017 las exportaciones de leche en polvo descremada aumentaron un 13.23 % con respecto al 2016

2018-2017: en el año 2018 las exportaciones de leche en polvo descremada disminuyeron un 40.99 % con respecto al 2017

-Leche fluida:

2014-2013: en el año 2014 las exportaciones de leche fluida aumentaron un 85.57 % con respecto al 2013

2016-2014: en el año 2016 las exportaciones de leche fluida disminuyeron un 54.42 % con respecto al 2014

2017-2016: en el año 2017 las exportaciones de leche fluida disminuyeron un 10.59 % con respecto al 2016

2018-2017: en el año 2018 las exportaciones de leche fluida aumentaron un 23.36 % con respecto al 2017

b) i) $I = 89097 / 205296 \cdot 100 = 43,4 \%$

En el año 2017 el total de las exportaciones de los 4 tipos de leches disminuyeron un 56,60 % con respecto al 2013

ii) $I = 70,43 \%$

En el año 2018 el total de las exportaciones de los 4 tipos de leches disminuyeron un 29,57 % con respecto al 2016

c) i) $I = 58535 / 148159 \cdot 100 = 39,51 \%$

$I = 16.735 / 30.987 \cdot 100 = 54,01 \%$

$I = 13.523 / 25.748 \cdot 100 = 52,52 \%$

$I = 304 / 402 \cdot 100 = 75,62 \%$

$I = (39,50 + 54 + 52,52 + 75,52) / 4 = 55,41$

$55,41 - 100 = -44,59 \%$

En el año 2017 el total de las exportaciones de los 4 tipos de leches disminuyeron un 44,59 % con respecto al 2013

ii) $I = 75,30 \%$

En el año 2018 el total de las exportaciones de los 4 tipos de leches disminuyeron un 24,70 % con respecto al 2016

d) El total de volúmenes anuales exportados de los 4 tipos de leches, entre 2017, cayó un 56,60 % mientras que en promedio los volúmenes anuales exportados de los 4 tipos de leches, disminuyó el 44,59 %. El total de volúmenes anuales exportados de los 4 tipos de leches, entre 2018 y 2016, cayó un 29,57 % mientras que en promedio los volúmenes anuales exportados de los 4 tipos de leches disminuyeron el 24,70 %

3- a)

	1.er/1.er 17	2.º/1.er 17	1.er 18/2.º 17	2.º/1.º 18	1.er 19/2.º 18
BA	100	136,04	92,10	97,07	97,14
CABA	100	140,00	90,31	97,90	94,67
ER	100	141,44	93,70	99,94	96,91
M	100	134,29	96,30	95,72	99,54
SF	100	142,06	96,59	95,29	98,11

	1.er/1.er 17	2.º/1.er 17	1.er 18/2.º 17	2.º/1.º 18	1.er 19/2.º 18
BA	0	36,04	-7,90	-2,93	-2,86
CABA	0	40,00	-9,69	-2,10	-5,33
ER	0	41,44	-6,30	-0,06	-3,09
M	0	34,29	-3,70	-4,28	-0,46
SF	0	42,06	-3,41	-4,71	-1,89

b) $I = 120,65 \%$

c) $I = 90,16 \%$

4- a)
$$I_P = \frac{(62.60)+(53.85)+(65.82)+(54.47)}{(51.60)+(45.85)+(53.82)+(48.47)} \cdot 100 = 119,32 \%$$

$$I_L = \frac{(62.50)+(53.80)+(65.75)+(54.45)}{(51.50)+(45.80)+(53.75)+(48.45)} \cdot 100 = 119,21 \%$$

$$I_F = \sqrt{119,21 \cdot 119,32} = 119,26 \%$$

b) $I_P = 125,13 \%$

$I_L = 125,07 \%$

$I_F = 125,10 \%$

5- $I_P = 3207,24 \%$

$$I_L = 3207,31 \%$$

$$I_F = 3207,275 \%$$

6- a) $SR_{Dic18/Dic16} = \frac{104,84}{64,15 \cdot 1,8426} \cdot 100 = 88,70$ $II_{Dic18/Dic16} = 184,26/100 =$ índice deflactor

El salario real de un 1.º oficial aglomerado por hora disminuyó un (88,7-100=-11,3 %) 11,3 % de diciembre 2018 con respecto a diciembre 2016

b) $SR_{Jun19/Jun17} = \frac{124,74}{77,17 \cdot 2,016} = 80,18 \%$ $II_{Jun19/Jun17} = 225,53/111,85 = 2,016$

El salario real de un 1.º oficial aglomerado por hora disminuyó un 19,82 % (80,18-100= -19,82) de junio 2019 con respecto a junio 2017

c) $SR_{Dic17/Jun16} = \frac{85,59}{55,91 \cdot 1,3245} \cdot 100 = 115,58 \%$

$$II_{Dic17/Jun16} = 124,79/94,22 = 1,3245$$

El salario real de un 1.º oficial aglomerado por hora **augmentó** un 15,58 % de diciembre 2017 con respecto a junio 2016

7- a) $SR_{17/16} = \frac{76,64}{55,09 \cdot 1,2479} \cdot 100\% = 111,48 \%$ $II_{17/16} = 124,79/100 = 1,2479$ í. deflactor

El salario real de un empleado representativo del sector servicios aumentó un 11,48 % en el 2017 con respecto al 2016

$$SR_{18/16} = \frac{98,57}{55,09 \cdot 1,8426} \cdot 100\% = 97,10 \%$$
 $II_{18/16} = 184,26/100 = 1,8426$

El salario real de un empleado representativo del sector Servicios disminuyó un 2,90 % en el 2018 con respecto al 2016

$$SR_{17/16} = \frac{55,75}{42,26 \cdot 1,2479} \cdot 100 \% = 105,71 \%$$
 $II = 184,26/100 = 1,8426$

$$SR_{17/16} = \frac{67,80}{42,26 \cdot 1,8426} \cdot 100 \% = 87,08 \%$$
 $II = 87,08/100 = 0,8708$

El salario real de un empleado representativo del sector comercio al menudeo aumentó un 5,71 % en el 2017 con respecto al 2016, mientras que disminuyó un 12,92 % entre el período 2018 y 2016

Manufactura:

$$SR_{17/16} = \frac{87,79}{65,15 \cdot 1,2479} \cdot 100 \% = 107,98 \%$$
 $II = 124,79/100 = 1,2479$

$$SR_{18/16} = \frac{104,84}{65,15 \cdot 1,8426} \cdot 100 \% = 87,33 \%$$
 $II = 184,26/100 = 1,8426$

El salario real de un empleado representativo del sector manufactura aumentó un 7,98 % en el 2017 con respecto al 2016, mientras que disminuyó un 12,67 % entre el período 2018 y 2016

b) $II_{19/17} = 225,53 / 124,79 \cdot 100 = 180,73 \% \quad 180,73 - 100 = 80,73 \%$
El poder adquisitivo del 2019 con respecto al 2017 aumentó un 80,73 %

c) $SR_{19/17} = \frac{31900}{20040 \times 1,8073} \cdot 100 \% = 88,08 \% \quad II = 225,23 / 124,79 = 1,8073 \quad 88,08 - 100 = -11,92$
El ingreso real del señor José Gómez disminuyó un 11,92 % en el año 2019 con respecto al 2017

PRÁCTICA 4

Series de tiempo

Pronosticar es un desafío que se aplica en una multiplicidad de escenarios, donde a partir del estudio del comportamiento de las variables de interés a lo largo del tiempo, se trata de establecer patrones que permitan proyectarlas a futuro con menor grado de incertidumbre y éste es el principal objetivo del estudio de series temporales o cronológicas.

Áreas que involucran conocimientos de planeación y control de operaciones, mercadotecnia, economía, especulación y riesgos financieros, presupuestos empresariales y gubernamentales, demografía y estrategias para estados de crisis son los principales demandantes de estas técnicas, de ahí su trascendencia [...]. (Diebold, 1999, cap. 1, pp. 1-2)

GPS ESTADÍSTICO 4

Orientaciones importantes para no perderse

- *La función de las series de tiempo es describir, en forma sintética, el comportamiento de una variable a lo largo del tiempo y generar predicciones.*
- *La importancia de determinar sus componentes radica en poder controlar sus principales variaciones para lograr pronósticos de mejor calidad.*
- *Existen muchos métodos de análisis de series temporales y depende del tipo de variable en estudio y del contexto de aplicación para definir cuál de ellos es más adecuado.*
- *La decisión de qué modelo de análisis es más conveniente aplicar requiere de un conocimiento profundo de las características de cada uno de ellos, de ahí que ese trabajo dependa de especialistas.*

Glosario 4

Componentes de una serie de tiempo: partes en las que se desglosa una serie de tiempo para su análisis.

Tendencia: componente que modela el comportamiento de la serie a largo plazo.

Componentes cíclicas: describen –si los hubiere- el comportamiento de los ciclos periódicos de distinta longitud que aparecen en la serie. Es decir que modelan los fenómenos de naturaleza periódica que inciden en la serie.

Componente estacional: nombre de la componente cíclica cuya longitud es de un año.

Componente irregular: contiene los ciclos y variaciones no periódicos y el ruido estadístico.

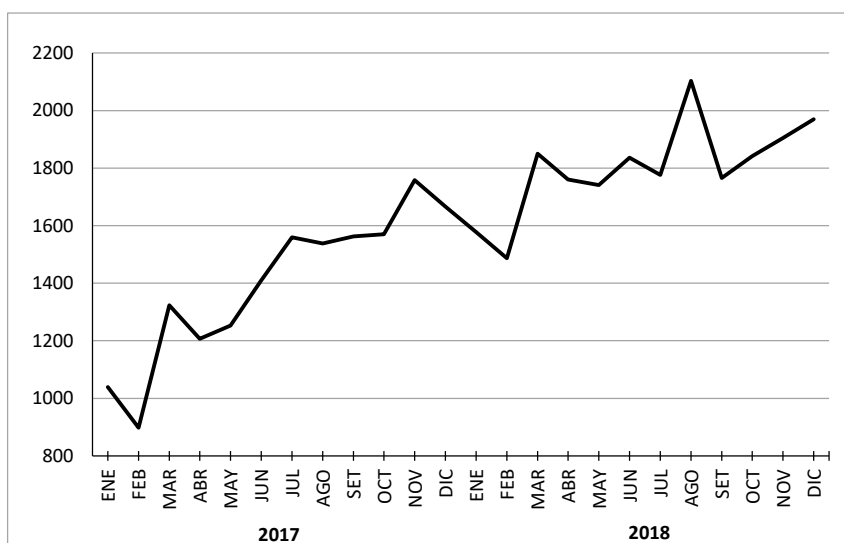
Recta de tendencia: modelo matemático que describe la tendencia de la serie.

Índices o coeficientes cíclicos: conjunto de valores puntuales que describen el ciclo modelado. Si la serie tuviera más de un ciclo periódico, para cada uno de ellos existirá un conjunto de índices descriptores. Cuando los períodos son intra anuales se les llama también *índices estacionales*.

Problemas y situaciones prácticas 4

- 1- Un analista necesita establecer un modelo que le permita estimar los montos correspondientes a importaciones mensuales de insumos para determinado rubro industrial, a partir de la información registrada en la tabla y el gráfico correspondiente, en millones de dólares.

Mes	2017	2018
Enero	1039	1578
Febrero	898	1487
Marzo	1323	1850
Abril	1207	1760
Mayo	1253	1741
Junio	1410	1836
Julio	1559	1776
Agosto	1538	2103
Septiembre	1563	1766
Octubre	1570	1841
Noviembre	1758	1904
Diciembre	1666	1970



- Determinar la recta de tendencia y trazarla en el gráfico.
- Filtrar la serie por tendencia, graficar los valores obtenidos y observar si presenta componente periódica. Justificar.
- Analizar, en ambos casos, los residuos relativos y determinar en qué año se produjo la máxima fluctuación.
- Estimar los montos que habrían sido requeridos para importaciones de los meses de marzo, julio y diciembre de 2019

Paréntesis técnico

En las **series de tiempo** la **variable** estadística en estudio está en función del tiempo, por lo que se la representa en el **eje de ordenadas**.

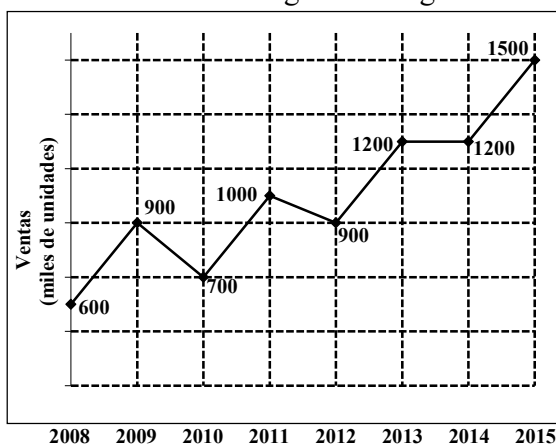
- 2- Las ventas de productos para la construcción de Construmar SA desde 2005 (en millones de dólares) fueron:

Año	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Ventas	1246	1325	1480	1350	1150	1021	1310	1375	990	1020	1115

- Graficar los datos.
- Determinar la ecuación de la tendencia por cuadrados mínimos y graficarla.
- Calcular los residuos relativos entre los valores estimados \hat{y} y los observados.
- Graficar y verificar si existe alguna componente periódica. Modelarla.
- Interpretar los ciclos modelados.
- Analizar los residuos y ubicar en qué año se produjo la máxima fluctuación.
- Pronosticar las ventas para los años 2018 y 2020

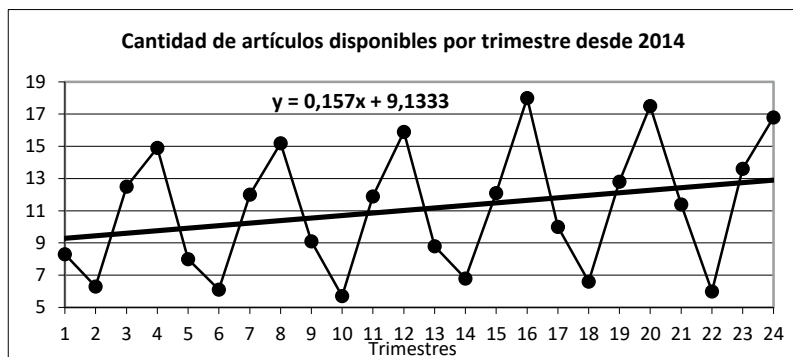
- 3- Las ventas de una empresa, que se inician en 2008 se muestran en el siguiente diagrama:

- Reconstruir la tabla de datos a partir del gráfico.
- Estimar la ecuación de la tendencia para la serie de ventas.
- Calcular y graficar los residuos relativos entre los valores estimados y los observados.
- ¿Dónde se presenta la mayor variación?
- Verificar si existe alguna componente cíclica periódica y establecer su longitud.
- Calcular los correspondientes índices cíclicos para cada fracción del ciclo.
- Pronosticar las ventas para 2017 y 2019



- 4- Una fábrica de autopartes hace inventario de sus productos disponibles cada trimestre, en miles de unidades. Los datos de cada trimestre a partir de 2014 se indican en la tabla y gráfico siguientes:

Año	I	II	III	IV
2014	8,3	6,3	12,5	14,9
2015	8,0	6,1	12,0	15,2
2016	9,1	5,7	11,7	15,7
2017	8,8	6,8	12,1	18,0
2018	10,0	6,6	12,8	17,5
2019	11,4	6,0	13,6	16,8



A partir de la recta de tendencia que figura en el diagrama:

- Interpretar la pendiente y la ordenada al origen.
- Calcular los porcentajes de la tendencia.
- Filtrar la serie por tendencia y analizar los residuos.
- Verificar la presencia de ciclos periódicos anuales. Modelar e interpretar los ciclos.
- Calcular los correspondientes índices cíclicos para cada fracción del ciclo.
- Observar si se presenta alguna componente de mayor longitud. Comentar.
- Con los índices calculados proyectar los inventarios para el primer trimestre y el cuarto trimestre del 2020

5- Para analizar la evolución de las erogaciones de un departamento de una empresa, se tomaron los siguientes datos que expresan, en millones de pesos, los gastos trimestrales de los cuatro años que figuran en la tabla:

- Hallar la recta de tendencia y componente cíclica si corresponde.
- Estimar los gastos para el tercer cuatrimestre de 2019

	2015	2016	2017	2018
1.º trimestre	30	40	48	55
2.º trimestre	28	37	44	48
3.º trimestre	27	35	42	46
4.º trimestre	32	41	50	56

6- Las ventas de una empresa en millones de pesos entre los años 2011 y 2017 se estiman por tendencia mediante $\hat{y} = 12,57 + 1,46x$. La serie presenta una variación cíclica bianual que se modela con los siguientes índices: I (años pares) = 90 e I (años impares) = 110

- Calcular el aumento anual promedio de ventas.
- Estimar la variación porcentual de 2019 respecto de 2018 en ventas.

7- La siguiente serie se refiere al PBI per cápita de Argentina para el período 2011-2018. Esta variable surge del cociente entre el PBI y la población total del país para un año dado.

Año	2011	2012	2013	2014	2015	2016	2017	2018
PBI per cápita	10883	10674,9	10784	10398	10568	10239	10404	10043

Fuente: base

de datos Banco Mundial,

<https://datos.bancomundial.org/indicador/NY.GDP.PCAP.KD?end=2018&locations=AR&start=2008>

- Obtener la tendencia lineal para la serie e interpretar la pendiente.
- Graficar la serie y la tendencia obtenida en un mismo diagrama cartesiano.
- Obtener el porcentaje o filtro por tendencia y graficar convenientemente.
- Estudiar si la serie presenta variación cíclica.
- Calcular la variación del PBI per cápita entre 2018 y 2011
- Calcular cual sería la variación entre 2020 y 2019 del PBI *per cápita*.
- Identificar las diferencias en la información utilizada de los puntos c y d.

8- La siguiente tabla muestra la variación porcentual de cada año del PBI argentino (variación PBI Argentina USD constantes de 2010).

Año	2012	2013	2014	2015	2016	2017	2018
Variación %	-1,03	2,41	-2,51	2,73	-2,08	2,67	-2,48

Fuente: base de datos Banco Mundial,

<https://datos.bancomundial.org/indicador/NY.GDP.MKTP.KD.ZG?locations=AR>

- Determinar la recta de tendencia y la variación cíclica, si la hay, en la serie.
- Pronosticar la variación porcentual del PBI para el año 2020
- Pronosticar la variación porcentual del PBI para el año 2032
- Identificar las diferencias entre la recta de tendencia de este ejercicio y el anterior. Tratándose de la variable PBI e) ¿Por qué la pendiente es positiva en una serie y negativa en la otra, si están definidas para el mismo período?

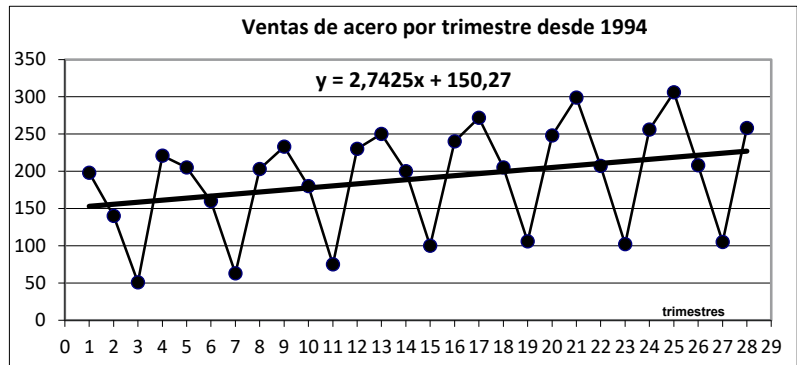
9- Los promedios anuales de las cuentas bancarias en millones de pesos de un banco extranjero están dados en la siguiente tabla:

Tipo de cuentas/años	1992	1994	1996	1998	2000	2002	2004
Cuentas de ahorro	19500	25000	27500	31900	33400	19500	26700
Cuentas de inversiones	5500	6200	7400	7900	8100	6000	7500

- Hallar la tendencia a largo plazo de las cuentas de ahorro de este banco.
- Calcular los porcentajes de la tendencia y determinar en qué período se encuentra la mayor fluctuación.
- Analizar si hay alguna componente cíclica periódica, en caso afirmativo decir de qué longitud y modelarla.

10- Los volúmenes de las ventas de acero de una fábrica (en millones de dólares), por trimestre desde 2013, se muestran a continuación (tabla y gráfico):

Año	I	II	III	IV
2013	198	140	51	221
2014	205	160	63	203
2015	233	180	75	230
2016	250	200	100	240
2017	272	205	106	248
2018	299	207	102	256
2019	306	208	105	258



- A partir de la ecuación de mínimos cuadrados de los volúmenes de ventas de acero de la fábrica (adjunta en el gráfico) filtrar la serie por tendencia y analizar los residuos en la tabla modelo que se adjunta al final.
- Verificar la presencia de ciclos periódicos anuales.
- Modelar e interpretar los ciclos, si los hay.
- Calcular los promedios de un año de longitud.
- Calcular los índices cíclicos para cada fracción de año en la que se encuentran los datos.
- Proyectar las ventas para el último trimestre del 2020 y el primer trimestre del 2021

Tabla modelo de procesamiento de series de tiempo

Período	x	Serie	Tendencia	Filtro	Residuo	Variación	Período	ÍC
		Y	\hat{Y}	$Y/\hat{Y}(x100\%)$	Filtro - 100 %	Relativa al período anterior	I-II-...	IC (per.)
2013-I	1	198						
II	2	140						
III	3	51						
IV	4	221						
2014-I	5	205						
II	6							
III	7							
IV	8							
2015-I	9							
II	10							
III	11							
IV	12							
2016-I	13							
II	14							
III	15							
IV	16							
2017-I	17							
II	18							
III	19							
IV	20							

2018-I	21							
II	22							
III	23							
IV	24							
2019-I	25							
II	26							
III	27							
IV	28							
Pronósticos		--						
		--						

RESPUESTAS DE LAS ACTIVIDADES DE LA PRÁCTICA 4

Series de tiempo

1- a) $\hat{y} = 1131,6594 + 37,4539x$ $\frac{y}{\hat{y}} \cdot 100\%$

T	X	Y	\hat{Y}	$\frac{Y}{\hat{Y}} \cdot 100\%$	$\frac{Y}{\hat{Y}} \cdot 100\% - 100\%$
2017 -Enero	1	1039	1169,11	88,87	-11,13
Febrero	2	898	1206,57	74,43	-25,57
Marzo	3	1323	1244,02	106,35	6,35
Abril	4	1207	1281,47	94,19	-5,81
Mayo	5	1253	1318,93	95,00	-5,00
Junio	6	1410	1356,38	103,95	3,95
Julio	7	1559	1393,84	111,85	11,85
Agosto	8	1538	1431,29	107,46	7,46
Septiembre	9	1563	1468,74	106,42	6,42
Octubre	10	1570	1506,20	104,24	4,24
Noviembre	11	1758	1543,65	113,89	13,89
Diciembre	12	1666	1581,11	105,37	5,37
2018 -Enero	13	1578	1618,56	97,49	-2,50
Febrero	14	1487	1656,01	89,79	-10,26
Marzo	15	1850	1693,47	109,24	9,25
Abril	16	1760	1730,92	101,68	1,68
Mayo	17	1741	1768,37	98,45	-1,54
Junio	18	1836	1805,83	101,67	1,67
Julio	19	1776	1843,28	96,35	-3,65
Agosto	20	2103	1880,74	111,82	15,82
Septiembre	21	1766	1918,19	92,07	-7,96
Octubre	22	1841	19556,65	94,14	-5,85

Noviembre	23	1904	1993,10	95,53	-4,47
Diciembre	24	1970	2030,55	97,02	-2,98

b) No presenta componente cíclica periódica.

c) Máxima fluctuación: febrero del año 2017

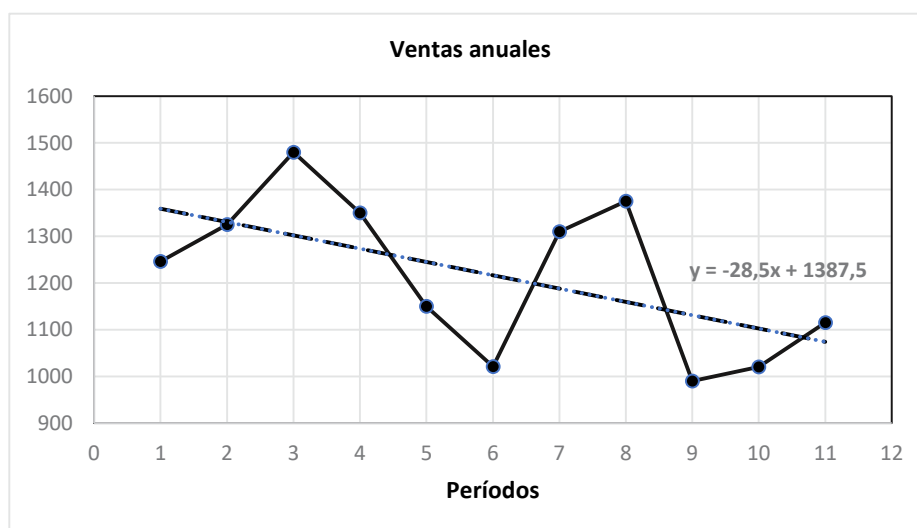
En enero, febrero, abril y mayo de 2017 y en enero febrero, mayo, julio y de septiembre a diciembre de 2018, el valor real estuvo por debajo del valor estimado por tendencia, y ocurrió lo contrario en los restantes meses.

d) Marzo 2019 $x=27$ $\hat{y} = 2142915$

Julio 2019 $x=31$ $\hat{y} = 2292,7303$

Diciembre 2019 $x=36$ $\hat{y} = 2480$

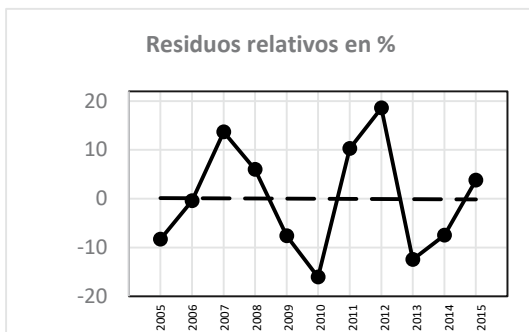
2- a)



b) $a=1387,54$ $b=-28,5$ $\hat{y}=1387,54 - 28,5 \cdot x$

T	X	Y	\hat{Y}	$\frac{Y}{\hat{Y}} \cdot 100\%$	$\frac{Y}{\hat{Y}} \cdot 100\% - 100\%$
2005	1	1246	1359,05	91,68	-8,32
2006	2	1325	1330,55	99,58	-0,42
2007	3	1480	1302,05	113,67	13,67
2008	4	1350	1273,55	106,00	6,00
2009	5	1150	1245,05	92,37	-7,63
2010	6	1021	1216,55	83,93	-16,07
2011	7	1310	1188,05	110,26	10,26
2012	8	1375	1159,55	118,58	18,58
2013	9	990	1131,05	87,53	-12,47
2014	10	1020	1102,55	92,51	-7,49
2015	11	1115	1074,05	103,81	3,81

c)



e) No presenta ciclos.

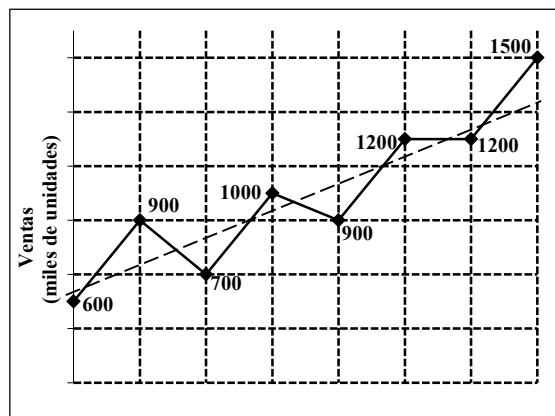
f) La máxima fluctuación se produjo en el año 2012. El valor real estuvo por debajo del valor estimado en los años: 2005, 2006, 2009, 2010, 2013, 2014. En los restantes años ocurrió lo contrario.

g) Para 2018, $x=14$ $\hat{y}=1387,54 - 28,5 \cdot 14 = 988,54$ millones de dólares

Para 2020, $x=16$ $\hat{y}=1387,54 - 28,5 \cdot 16 = 931,54$ millones de dólares

3- a)

T	X	Y
2008	1	600
2009	2	900
2010	3	700
2011	4	1000
2012	5	900
2013	6	1200
2014	7	1200
2015	8	1500



b) $\hat{y} = 507,1429 + 109,5238 \cdot x$

c)

T	X	Y	\hat{Y}	$\frac{Y}{\hat{Y}} \cdot 100 \%$	$\frac{Y}{\hat{Y}} \cdot 100 \% - 100 \%$	Per	IC
2008	1	600	616,67	97,30	-2,70	I	90
2009	2	900	726,19	123,93	23,93	II	110
2010	3	700	835,71	83,76	-16,24	I	90
2011	4	1000	945,24	105,79	5,79	II	110
2012	5	900	1054,76	85,33	-14,67	I	90
2013	6	1200	1164,29	103,07	3,07	II	110
2014	7	1200	1273,81	94,21	-5,79	I	90
2015	8	1500	1383,33	108,43	8,43	II	110

- d) La mayor variación se registró en el año 2009
 e) Como el movimiento se repite cada dos años, la longitud del ciclo es 2 años.
 f) $I=90$ $I=110$
 g) Para el año 2017, $x=10$, $\hat{y}(10)=1602,38$

$$P = \frac{1602,38 \cdot 110}{100} = 1762,618 \text{ mil unidades vendidas.}$$

Para el año 2019, $x=12$, e $\hat{y}(12)=1821,43$

$$P = \frac{1821,43 \cdot 110}{100} = 2003,57 \text{ mil unidades vendidas.}$$

4- a) $\hat{y} = 9,1123 + 0,1573x$

$b=0,1573$, por cada trimestre la cantidad disponible de juguetes aumenta en promedio 0,1573 mil juguetes.

$a=9,1123$ la cantidad de juguetes disponible en el IV trimestre de 2013 según este modelo sería de 9,1123 mil juguetes.

b)

T	X	Y	\hat{Y}	$\frac{Y}{\hat{Y}} \cdot 100 \%$	$\frac{Y}{\hat{Y}} \cdot 100 \% - 100 \%$	Per	IC
2014 I	1	8,3	9,27	89,54	-10,46	I	85
II	2	6,3	9,43	66,81	-3,19	II	57
III	3	12,5	9,58	130,48	30,48	III	112
IV	4	14,9	9,74	152,98	52,98	IV	146
2015 I	5	8	9,9	80,81	-19,19	I	85
II	6	6,1	10,06	60,64	-39,36	II	57
III	7	12	10,21	117,53	17,53	III	112
IV	8	15,2	10,37	146,58	46,58	IV	146
2016 I	9	9,1	10,53	86,42	-13,58	I	85
II	10	5,7	10,69	53,32	-46,68	II	57
III	11	11,7	10,84	107,93	7,93	III	112
IV	12	15,7	11	142,73	42,73	IV	146
2017 I	13	8,8	11,58	75,99	-24,01	I	85
II	14	6,8	11,32	60,07	-39,93	II	57
III	15	12,1	11,47	105,49	5,49	III	112
IV	16	18	11,63	154,77	54,77	IV	146
2018 I	17	10	11,79	84,82	-15,18	I	85
II	18	6,6	11,94	55,28	-44,72	II	57
III	19	12,8	12,1	105,79	5,79	III	112
IV	20	17,5	12,26	142,74	42,74	IV	146
2019 I	21	11,4	12,42	91,79	-8,21	I	85
II	22	6	12,57	47,73	-52,27	II	57
III	23	13,6	12,73	106,83	6,83	III	112
IV	24	16,8	12,89	130,33	30,33	IV	146

- c) En los dos primeros trimestres de cada año el valor observado estuvo por debajo de la tendencia, y ocurrió lo contrario en los dos segundos trimestres.
d) Existen ciclos de longitud en 4 trimestres.
e) 85, 57, 112, 146
f) Si, en el IV trimestre de 2017
g) En promedio se dispondría de
 $x=25 \quad P=13,05.85/100=11,0925$ miles de juguetes según esta tendencia.
 $x=28 \quad P=13,52.146/100=19,7392$ miles de juguetes según esta tendencia.

5- a) $\hat{y} = 36,95 + 0,4985.x$

T	X	Y	\hat{Y}	$\frac{Y}{\hat{Y}} \cdot 100\%$	$\frac{Y}{\hat{Y}} \cdot 100\% - 100\%$	Per	IC
I trim. 2015	1	30	37,45	80,11	-19,89	I	72
II	2	40	37,95	105,4	5,4	II	94
III	3	48	38,45	124,84	24,84	III	111
IV	4	55	38,94	141,24	41,24	IV	123
I trim. 2016	5	28	39,44	70,99	29,01	I	72
II	6	37	39,94	92,64	-7,36	II	94
III	7	44	40,44	108,8	8,8	III	111
IV	8	48	40,94	117,24	17,24	IV	123
I trim. 2017	9	27	41,44	65,15	-34,85	I	72
II	10	35	41,94	83,45	-16,55	II	94
III	11	42	42,43	98,99	-1,01	III	111
IV	12	46	42,93	107,15	7,15	IV	123
I trim. 2018	13	32	43,43	73,68	-26,32	I	72
II	14	41	43,93	93,33	-6,67	II	94
III	15	50	44,43	112,54	12,54	III	111
IV	16	56	44,93	124,64	24,64	IV	123

Se presentan ciclos de 4 trimestres de longitud.

b) Para el I trimestre 2019, $x=17$, $P=32,706$

6- $\hat{y} = 12,57 + 1,46.x$

a) 1,46

b) Para 2018 $x=8$, entonces el pronóstico es:

$$P = (12,57 + 1,46.8).90/100 = 21,825$$

$$\text{Para } 2019 \quad x=9, \quad P = (12,57 + 1,46.9).110/100 = 28,281$$

$$I = 28,281/21,825.100 = 129,58\% \quad \text{Variación} = 29,58\%$$

7- a) $\hat{y} = 10965,32 - 103,57.x$

b = - 103,57 El PBI *per cápita* cae en promedio 103,57 dólares por año.

c)

T	x	Y	\hat{Y}	$Y/\hat{Y} \cdot 100$	L	lc
2011	1	10.883,00	10.861,75	100,20	I	101
2012	2	10.675,00	10.758,18	99,23	II	99
2013	3	10.784,00	10.654,61	101,21	I	101
2014	4	10.398,00	10.551,04	98,55	II	99
2015	5	10.568,00	10.447,46	101,15	I	101
2016	6	10.239,00	10.343,89	98,99	II	99
2017	7	10.404,00	10.240,32	101,60	I	101
2018	8	10.043,00	10.136,75	99,08	II	99

d) La longitud del ciclo es de 2 años.

e) $I=10043/10404 \cdot 100=96,53\%$ Variación: -3,47 %

$I=9830,311071/10133,51036 \cdot 100=97,01\%$ Variación -2,99 %

8- a) $\hat{y} = 0,44 - 0,12 \cdot x$

T	X	Y	\hat{Y}	$Y/\hat{Y} \cdot 100$
2012	1	-1,03	0,32	-321,875
2013	2	2,41	0,20	1205
2014	3	-2,51	0,08	3137,5
2015	4	2,73	-0,04	-6825
2016	5	-2,08	-0,16	1300
2017	6	2,67	-0,28	953,57
2018	7	-2,48	-0,41	604,88

No presenta ciclos

b) $\hat{y}(9) = 0,44 - 0,12 \cdot 9 = -0,64$

c) 2032 está lejos de recorrido de la variable.

9-a) $\hat{y} = 23857,14 + 589,29 \cdot x$

T	X	Y	\hat{Y}	$\frac{Y}{\hat{Y}} \cdot 100\%$	$\frac{Y}{\hat{Y}} \cdot 100\% - 100\%$
1992	1	19500	2446,43	79,77	-20,23
1994	2	25000	25035,71	99,86	-0,14
1996	3	27500	25625	107,32	7,32
1998	4	31900	26214,29	121,69	21,69
2000	5	33400	26803,57	124,6	24,6
2002	6	19500	27392,86	71,19	-28,81
2004	7	26700	27982,14	95,42	-4,58

b) Mayor fluctuación en 2002

c) No.

10-a) $\hat{y} = 150,27 + 2,74.x$

T	X	Y	\hat{Y}	$Y/\hat{Y} \cdot 100$
I trim. 2013	1	198	153,01	129,4
II	2	140	155,75	89,89
III	3	51	158,50	111,23
IV	4	221	161,24	137,06
I trim. 2014	5	205	163,98	125,02
II	6	160	166,72	95,97
III	7	63	169,47	37,17
IV	8	203	172,21	117,88
I trim. 2015	9	233	174,95	133,18
II	10	180	177,69	101,3
III	11	75	180,44	41,57
IV	12	230	183,18	125,56
I trim. 2016	13	250	185,92	134,47
II	14	200	188,66	106,01
III	15	100	191,41	52,24
IV	16	240	194,15	123,62
I trim. 2017	17	272	196,89	138,15
II	18	205	199,63	102,69
III	19	106	202,38	52,38
IV	20	248	205,12	120,9
I trim. 2018	21	299	207,86	143,85
II	22	207	210,60	98,29
III	23	102	213,35	47,81
IV	24	256	216,09	118,47
I trim. 2019	25	306	218,83	139,83
II	26	208	221,57	93,87
III	27	105	224,32	46,81
IV	28	258	227,06	113,63

b) No presenta ciclos.

f) Para $x=32$, $P=238,03$

Para $x=33$, $P=24,77$

PRÁCTICA 5

Cálculo de probabilidades

La teoría de la probabilidad es la herramienta conceptual necesaria para abordar con fundamento los problemas de la estadística inferencial.

En la inferencia estadística, las muestras representativas de una población son el material básico de trabajo y para poder extraerlas adecuadamente se utilizan los experimentos aleatorios compuestos por repetición de uno simple. Por otro lado, el estudio de la asociación y/o relación causa efecto entre variables se apoya en los experimentos compuestos bivariados.

Fue en el siglo XVII donde se desarrolló el cálculo de probabilidad como disciplina científica, paralelamente al desarrollo de la estadística, pero en forma independiente. Se podría decir que a fines del siglo XIX se comenzó a utilizar el concepto de azar, indeterminismo o aleatoriedad.

GPS ESTADÍSTICO 5

Orientaciones para no perderse

- *La probabilidad cuantifica la posibilidad de ocurrencia de un hecho. Poder reconocer que la incertidumbre en algunos casos es insalvable habla de la capacidad para poder comprender que no todo es medible o previsible. Hay limitaciones que la ciencia debió aceptar en lo que a comprensión y “dominio” se refiere.*
- *Cuando estudiamos la probabilidad de que ocurra “esto o aquello” no necesariamente se está pidiendo que no pasen ambas cosas juntas, ya que basta que se dé una de ellas para que ese evento suceda. Por el contrario, si queremos ver cuan probable es que “esto y aquello” ocurra, es imprescindible que ambas cosas sucedan a la vez. Si esto es imposible, esos sucesos resultan mutuamente excluyentes.*
- *Es importante diferenciar los conceptos de exclusión y de independencia; estos son dos conceptos opuestos, dos puntos de vista diferentes para estudiar la relación entre sucesos aleatorios.*
- *La exclusión entre sucesos indica que ellos no ocurren simultáneamente, es decir que su intersección es vacía entonces la probabilidad de la intersección entre los sucesos es cero.*
- *La independencia entre sucesos indica que la ocurrencia de uno de ellos no incide en la ocurrencia del otro; en este caso la probabilidad de la intersección entre los sucesos es el producto de la probabilidad de cada uno de ellos.*

- *Cuando dos sucesos son mutuamente excluyentes son fuertemente dependientes.*
- *Cuando dos sucesos son dependientes no necesariamente serán mutuamente excluyentes.*

Glosario 5

Experimento aleatorio, probabilístico o estocástico: es un experimento o fenómeno cuyo resultado no puede predecirse.

Espacio muestral: conjunto de todos los resultados posibles de un experimento aleatorio.

Suceso aleatorio: cualquier subconjunto de un espacio muestral.

Probabilidad: es la medida de la ocurrencia de un suceso aleatorio.

Equiprobabilidad: es la igualdad de oportunidad que tiene cualquier elemento de una población, respecto de los otros elementos, de ser seleccionado para integrar una muestra.

Sucesos elementales: son los subconjuntos unitarios seleccionados de un espacio muestral.

Suceso cierto: es el subconjunto que coincide con el espacio muestral.

Suceso imposible: es el suceso que no tiene ningún elemento.

Problemas y situaciones prácticas 5

1- Se tira un dado una vez y se lee el número de la cara superior.

- ¿Cuál es el espacio de los sucesos?
- ¿Los sucesos elementales son equiprobables?
- ¿Por qué los resultados son mutuamente excluyentes?
- ¿Cuál es la probabilidad de que sea un dos? ¿Cuál es el supuesto para poder calcularla?

Paréntesis histórico

Los primeros estudios sobre probabilidad estaban intrínsecamente unidos con los juegos de azar, y buscaban la máxima rentabilidad monetaria o la mayor utilidad. Este objetivo estuvo vigente hasta que apareció la obra de Daniel Bernoulli, hacia 1730.

La condición de equiprobabilidad es introducida por Laplace como un recurso para compensar el déficit en la formulación clásica de la probabilidad.

Las obras de Laplace y de otros matemáticos como Poisson, Gauss o Galton, proveyeron al cálculo de probabilidades de recursos matemáticos que lo llevaron a un grado de perfeccionamiento que lo ha hecho apto para las aplicaciones a diversos campos de la ciencia y muy especialmente a la estadística.

2- De los 18637 inscriptos en el curso de ingreso de la UNLaM se sabe que agrupados por área se inscribieron 6234 para Ciencias de la Salud, 3538 para Ciencias Económicas, 3529 para Humanidades y Ciencias Sociales, 3094 para Ingeniería e Investigaciones tecnológicas y

2242 para Derecho y Cs. Políticas. Si se selecciona un aspirante al azar, calcular la probabilidad de que se haya inscripto en alguna carrera del área de:

- a) Salud
- b) Económicas
- c) Derecho o Humanidades
- d) Alguna que no sea de Ingeniería
- e) Comparar los resultados obtenidos en los ítems anteriores y explicar las relaciones que puedan establecerse entre ellos.

3- Una caja de lápices perdidos contiene 10 lápices, 4 de ellos son rojos, 2 verdes, 1 amarillo y 3 azules. Si se extrae un lápiz al azar, calcular la probabilidad de sacar:

- a) un lápiz rojo
- b) un lápiz azul
- c) un lápiz que no sea verde
- d) uno rojo o uno verde

4- De 300 estudiantes que cursan primer año de la Lic. en Administración, 200 están inscriptos en Contabilidad y 100 en Matemática. Estas cifras de inscripción incluyen 30 estudiantes que están inscriptos en ambas materias. ¿Cuál es la probabilidad de que un alumno elegido al azar:

- a) no esté inscripto en ninguna de las dos materias?
- b) esté inscripto únicamente en Contabilidad?
- c) esté inscripto únicamente en Matemática?
- d) esté inscripto en una sola de las dos materias?
- e) esté inscripto en alguna de las dos materias?

5- El propietario de una tienda de música sabe que el 30 % de sus clientes pide ayuda a los dependientes y que el 20 % hace una compra antes de abandonar el local. Además, sabe que el 15 % de los clientes pide ayuda y hace una compra.

- a) ¿Cuál es la probabilidad de que un cliente haga al menos una de estas dos cosas?
- b) Si se consideran los sucesos “el cliente pide ayuda” y “el cliente hace una compra”, responder a las siguientes preguntas y justificar las respuestas en términos de probabilidades de los sucesos relevantes.
 - i) ¿Son los dos sucesos mutuamente excluyentes?
 - ii) ¿Los dos sucesos cubren exhaustivamente el espacio muestral?

6- De un grupo de 9 personas, dos son inglesas, tres francesas y cuatro italianas. Se eligen 2 personas aleatoriamente, ¿cuál es la probabilidad de que las dos sean italianas?

7- Si de la caja del ejercicio 3 se extraen 2 lápices, calcular la probabilidad de sacar:

- a) dos lápices verdes
- b) uno verde y uno amarillo
- c) al menos uno rojo
- d) como máximo uno azul
- e) ninguno rojo

- 8- En una empresa trabajan 200 personas de las cuales 120 son hombres, según el departamento de personal 150 de los empleados residen en la CABA y el resto en el Gran Bs. As. y que 58 personas del total son mujeres que viven en la CABA.
Completar la siguiente tabla conjunta que resume la información.

	Mujer	Hombre	
CABA	58		150
Bs. As.			
		120	200

- a) Si se selecciona aleatoriamente un empleado cuál es la probabilidad de que
- i) sea mujer,
 - ii) resida en la CABA,
 - iii) sea mujer o resida en el Gran Bs. As.,
 - iv) no resida en el Gran Bs. As.
 - v) sea mujer y resida en CABA,
 - vi) sea mujer si reside en CABA.
 - vii) que resida en Gran Bs. As. sabiendo que es mujer.
- b) Indicar que tipo de probabilidad se calculó en cada caso.
- c) Si se seleccionan *dos* empleados al azar, calcular la probabilidad de que:
- i) ambos sean mujeres,
 - ii) ambos residan en CABA,
 - iii) a lo sumo uno resida en el Gran Bs. As.
 - iv) al menos uno sea hombre,
 - v) ninguno sea hombre.
- 9- Una industria tiene dos plantas cuyos empleados se distribuyen como se indica en la siguiente tabla:

ÁREA	Planta 1	Planta 2
Administrativa	12	7
Operativa	200	250
Logística	50	32

- a) Si se elige un empleado al azar calcular la probabilidad de que:
- i) pertenezca al área administrativa,
 - ii) no sea de logística y pertenezca a la planta 1,
 - iii) sea del área operativa o de la planta 2,
 - iv) sea del área administrativa si pertenece a la planta 1,
 - v) no sea de la planta 2,
 - vi) sea de la planta 1 si pertenece al área administrativa.
- b) Si se seleccionan dos empleados al azar calcular la probabilidad de que:
- i) ambos sean del área operativa,
 - ii) al menos uno pertenezca a la planta 2

10- Una empresa de prospección petrolera encuentra petróleo o gas en el 10 % de sus perforaciones. Si la empresa estuviera programando perforar dos pozos en cuencas geológicas totalmente independientes,

- describir el espacio muestral para ese programa de trabajo.
- Completar la siguiente tabla (determinar las probabilidades faltantes).

Resultado de la 1. ^{ra} perforación	Resultado de la 2. ^a perforación	Probabilidad
Se encontró (petróleo o gas)	Se encontró (petróleo o gas)	0,01
Se encontró	No se encontró	
No se encontró	Se encontró	
No se encontró	No se encontró	

- Calcular la probabilidad de que la compañía encuentre gas o petróleo en por lo menos una de las dos perforaciones.

11- Un tirador acierta el 93 % de sus tiros, si este realiza tres tiros:

- Describir el espacio muestral del experimento.
- ¿Cuál es la probabilidad de que acierte las tres veces?
- ¿Cuál es la probabilidad de que no acierte ninguna vez?
- ¿Cuál es la probabilidad de que acierte por lo menos una vez?

12- En los controles realizados sobre los asientos contables efectuados por un empleado de una importante firma, con la finalidad de evaluarlo debido a las quejas recibidas, se detectaron errores de tipeo y de cálculo de los datos en la proporción del 8 % y 5 % respectivamente. Si se supone que los errores mencionados son independientes entre sí, determinar la probabilidad de que al seleccionar al azar un asiento cualquiera del empleado, en el mismo:

- no se encuentre ningún error,
- se encuentren los dos tipos de error,
- se encuentre solo una clase de error.

13- Repetir el cálculo de las probabilidades calculadas en los ejercicios 6, 7, 8c) y 9b) suponiendo que la segunda selección se realiza *con reposición*.

14- En el experimento de arrojar un dado del ejercicio 1, considerar los sucesos: A: sale un dos y B: sale un número par.

- ¿Tienen A y B la misma probabilidad?
- ¿Son A y B mutuamente excluyentes?
- ¿Son A y B independientes?
- Justificar las respuestas.
- Calcular $P(A/B)$ y $P(B/A)$ y explicar para estos sucesos en particular, el significado de las probabilidades calculadas.

- 15- El 80 % de los alumnos de primer año aprueban Contabilidad y el 70 % aprueba Matemática, estas cifras incluyen un 60 % que aprueban ambas materias. ¿Cuál es la probabilidad de que un alumno apruebe Matemática sabiendo que no aprobó Contabilidad?
- 16- El 30 % de los empleados de una empresa son solteros, de estos hay un 60 % que vive a menos de 2 km de su trabajo. El 20 % de los que no son solteros habitan a más de 2 km de su trabajo. Si se selecciona un empleado al azar:
- ¿Cuál es la probabilidad de que viva a menos de 2 km?
 - ¿Cuál es la probabilidad de que no sea soltero si vive a menos de 2 km?
 - ¿Cuál es la probabilidad de que viva a más de 2 km o sea soltero?
- 17- En un supermercado el 70 % de las ventas de yerba mate son de la marca AMANECER y el resto son de otras marcas. El 45 % de los paquetes de la marca AMANECER son de $\frac{1}{2}$ kg y de las otras marcas, el 30 % son de este tamaño. Si se selecciona un paquete al azar calcular la probabilidad de que:
- no sea de la marca AMANECER y no pese $\frac{1}{2}$ kg
 - sea de $\frac{1}{2}$ kg
 - no pese $\frac{1}{2}$ kg o no sea de la marca AMANECER
 - sea de la marca AMANECER si pesa $\frac{1}{2}$ kg
- 18- El 30 % de los créditos de un banco son para vivienda, el 45 % para industria y el 25 % para consumo diverso. Resultan morosos el 24 % de los créditos para vivienda, el 18 % de los créditos para industria y el 60 % de los créditos para consumo.
- Calcular la probabilidad de que se pague a término un crédito elegido al azar.
 - Si el crédito seleccionado fuese pagado a término, ¿cuál es la probabilidad que haya sido otorgado para la adquisición de una vivienda?
- 19- El 58 % de los alumnos de un secundario faltan al menos 2 veces al mes, de estos el 40 % tiene al menos una materia baja y el 10 % del resto también.
- ¿Qué porcentaje de los alumnos tiene alguna materia baja?
 - Si se elige un alumno que tiene alguna materia baja, ¿cuál es la probabilidad de que haya faltado menos de 2 veces?
- 20- En un programa de entrenamiento a nivel gerencial, 80 % de los asistentes son mujeres, el 90 % de las mujeres son egresadas universitarias y el 85 % de los hombres también. Si se selecciona al azar una de las personas en entrenamiento. Calcular la probabilidad de que:
- Se trate de una mujer que no asistió a la universidad.
 - Se trate de una mujer si no asistió a la universidad.
 - No haya asistido a la universidad si es mujer.
 - Realice un diagrama de árbol que muestre todas las probabilidades normales, condicionales y conjuntas.
 - ¿Cuánto suman todas las probabilidades conjuntas? Justificar el resultado.

- 21-** Tres refinerías producen alcohol siendo su producción diaria de 100.000, 200.000 y 1.300.000 litros respectivamente. El porcentaje inutilizable para la venta es del 3 %, 5 % y 10 % respectivamente. Si se toma una muestra al azar de la producción de alcohol, calcular la probabilidad de que:
- no sea apta para la venta,
 - sea de la primera refinería sabiendo que es apta para la venta.
- 22-** Una fábrica posee dos máquinas que producen el mismo artículo en iguales proporciones. El 1 % de los artículos que produce la primera máquina presenta algún defecto y de la otra, el 2 %
- ¿Cuál es la probabilidad de seleccionar un artículo defectuoso?
 - Si se seleccionó un artículo bueno, ¿cuál es la probabilidad de que provenga de la primera máquina? ¿Y de la segunda?
- 23-** Una compañía de seguros de automóviles ha asegurado 35.000 conductores de clase (muy buenos conductores); 50.000 conductores de la clase B (medianamente buenos conductores) y a 15.000 conductores de la clase C (malos conductores). La probabilidad de que un conductor de la clase A, B o C tenga uno o más accidentes, durante un año es de 0,01; 0,04 y 0,15 respectivamente. La compañía vende a la señora García una póliza de seguro y en un año tiene un accidente. ¿Cuál es la probabilidad de que la señora sea:
- un conductor de la clase A?
 - un conductor de clase B?
- 24-** En una empresa el 60 % de los empleados son obreros, el 20 % son administrativos y el resto de maestranza. El 10 % de los obreros y el 5 % de los de maestranza piden alguna licencia al año, además se sabe que la proporción total de licencias que piden los empleados en el año es del 12 %
- Para organizar la información realizar un diagrama de árbol, una tabla de contingencia y realizar un diagrama de Venn que refleje la situación indicada en el problema.
 - Si se selecciona un empleado que pidió alguna licencia calcular la probabilidad de que este sea un empleado administrativo.
- 25-** Se sabe que en una fábrica el 68 % de los empleados son hombres. El 70 % de los hombres percibe asignaciones familiares en su salario. Se sabe además que el 58 % de los empleados recibe las asignaciones en su salario. Si se selecciona un empleado al azar
- ¿Cuál es la probabilidad de que reciba salario familiar si es mujer?
 - ¿Cuál es la probabilidad de reciba asignación familiar o sea mujer?
- 26-** Sean A y B dos sucesos tales que $P(A) = 0,37$; $P(B) = 0,908$ y $P(A \cup B) = 0,98$. Hallar $P(A/B)$ y $P(B/A)$.
- 27-** Una empresa turística promociona tres destinos: Mendoza, Cataratas del Iguazú y Calafate con una capacidad de 40, 55 y 50 pasajes para cada destino respectivamente, el 80 % de los turistas que viajan a Mendoza solicita una base simple, como el 65 % de los de Cataratas y

90 % de los de Calafate también. Si se selecciona un turista al azar, calcular la probabilidad de que:

- a) Viaje a Cataratas.
- b) Haya solicitado una base simple.
- c) Sabiendo que pidió base doble, viaje a Mendoza.

28- Si dos sucesos M y N cumplen con: $P(M/N) = 0,50$; $P(N/M) = 0,65$ y $P(M \cap N) = 0,30$. Calcular $P(M \cup N)$.

29- El 62 % de los ingresantes de la carrera de Contador Público son varones, se sabe que de ellos el 48 % cursó el colegio secundario con orientación contable. Del *total* de los alumnos que ingresan a esta carrera el 60 % cursó el secundario con esta orientación. Si se selecciona un alumno aleatoriamente calcular la probabilidad de que sea:

- a) Egresado de un colegio secundario con orientación contable si es mujer.
- b) Mujer si cursó el secundario con otra orientación.
- c) Sea varón o haya cursado el secundario con una orientación distinta a la contable.

30- El 40 % de los usuarios de telefonía móvil corresponden a la empresa MOVILE y el resto a otras compañías, de los usuarios de MOVILE se ha registrado que el 50 % de ellos realiza algún reclamo en el mes y que la proporción de los que realizan algún reclamo o pertenecen a la compañía MOVILE representan un 45 %

- a) Si un usuario realizó algún reclamo en el mes, calcular la probabilidad de que no pertenezca a la compañía MOVILE.
- b) Si el usuario seleccionado no es de la compañía MOVILE, ¿cuál es la probabilidad de que no realice ningún reclamo?

31- a) Completar el siguiente cuadro teniendo en cuenta la información disponible para cada caso:

Caso	P(A)	P(B)	P(A ∩ B)	P(A ∪ B)	P(A/B)	P(B/A)	A y B son sucesos		
							equiprobables	independientes	mutuamente excluyentes
i	0,35	0,25	0,11						
ii	0,56	0,62					NO	SI	NO
iii	0,05						SI	NO	SI
iv		0,18	0,054				NO	SI	NO
v			0,28	0,84			SI	NO	NO
vi			0,12		0,66		NO	SI	NO
vii		0,23		0,65			NO	NO	SI
viii			0,27		0,52	0,72			

b) Analizar:

- i. ¿Pueden dos sucesos ser equiprobables y mutuamente excluyentes?
- ii. ¿Pueden dos sucesos ser independientes y mutuamente excluyentes?
- iii. ¿Pueden dos sucesos ser independientes y equiprobables?

iv. ¿Pueden dos sucesos ser independientes, equiprobables y mutuamente excluyentes?

32- Si $P(A) = 1/4$, $P(B/A) = 1/2$ y $P(A/B) = 1/4$, decidir si es verdadero o falso y justificar que:

$$P(A/B) + P(\bar{A}/\bar{B}) = 1$$

33- De las ventas de un comercio de electrodomésticos se sabe que el 40 % son en efectivo, el 25 % paga con tarjeta de crédito y el 35 % restante pide un crédito a sola firma. El 70 % de los que pagan con tarjeta y el 90 % de los que pide un crédito a sola firma realizan la compra en 2 o más cuotas. Si se selecciona una venta calcular la probabilidad de que:

- Pague en 2 o más cuotas.
- Sabiendo que realizó un único pago, este sea en efectivo.

34- Si dos sucesos A y B son independientes, verificar que se cumple la siguiente igualdad:

$$\frac{P(A/B) \cdot P(B/A)}{P(A \cap B)} = 1$$

RESPUESTAS DE LAS ACTIVIDADES DE LA PRÁCTICA 5

Cálculo de probabilidades

- $E = \{1, 2, 3, 4, 5, 6\}$
 - Si.
 - Porque la intersección entre pares de sucesos elementales es vacía.
 - $B = \{2\}$ $P(B) = 1/6 = 0,1667$, es necesario que el espacio muestral sea finito y equiprobable.
- $P(S) = 6234/18637 = 0,334$
 - $P(E) = 3538/18637 = 0,19$
 - $P(D \cup H) = P(D) + P(H) = 2242/18637 + 3529/18637 = 0,31$
 - $P(\bar{I}) = 1 - P(I) = 1 - 3094/18637 = 0,834$
 - $P(\bar{I}) = P(S) + P(E) + P(DUH) = 0,834$
- $P(R) = \frac{4}{10} = 0,4$
 - $P(A) = \frac{3}{10} = 0,3$
 - $P(V^c) = 0,8$
 - $P(RUV) = 0,6$
- $P((MUC)^c) = 0,1$
 - $P(C-M) = 0,5667$
 - $P(M-C) = 0,2333$

d) $P((C-M) \cup (M-C)) = 0,8$

e) $P(MUC) = 1 - 0,1 = 0,9$

5- a) $P(\text{al menos } 1) = P(AUC) = P(A) + P(C) + P(A \cap C) = 0,3 + 0,2 - 0,15 = 0,35$

b) i) No, porque $P(A \cap C) \neq \emptyset$

ii) No, porque $P(AUC) = 0,35 \neq 1$

6- $P(I_t \cap I_t) = P(I_t / I_t)$. $P(I_t) = 4/9 \cdot 3/8 = 0,1666$

7- a) $P(V \cap V) = P(V) \cdot P(V/V) = 0,0222$

b) $P(V \cap A) + P(A \cap V) = 0,044$

c) $P(FR \cap R) + P(R \cap V) + P(R \cap A) + P(R \cap Az) + P(V \cap R) + P(A \cap R) + P(A \cap R) + P(Az \cap R) = 0,666$

d) $1 - P(Az \cap Az) = 0,933333$

e) $P(V \cap V) + P(V \cap A) + P(V \cap Az) + P(A \cap V) + P(A \cap Az) + P(Az \cap V) + P(Az \cap A) + P(Az \cap Az) = 0,3333$

8- a) i) $P(M) = 80/200 = 0,40$

ii) $P(C) = 150/200 = 0,75$

iii) $P(M \cup B) = 80/200 + 50/200 - 22/200 = 0,4 + 0,25 - 0,11 = 0,54$

iv) $P(B^c) = 1 - P(B) = 1 - 50/200 = 1 - 0,25 = 0,75$ $P(C) = 150/200 = 0,75$

v) $P(M \cap C) = 58/200 = 0,29$

vi) $P(M/C) = 58/150 = 0,29/0,75 = 0,3867$

vii) $P(B/M) = 22/80 = 0,11/0,4 = 0,275$

b) i) $P(M) = 0,40$

ii) $P(C) = 0,75$

iii) $P(M \cup B) = 0,4 + 0,25 - 0,11 = 0,54$

iv) $P(B^c) = 1 - 0,25 = 0,75$

v) $P(M \cap C) = 0,29$

vi) $P(M/C) = 0,29/0,75 = 0,3867$

vii) $P(B/M) = 0,11/0,4 = 0,275$

simple total
simple total
unión total
del complemento
conjunta
condicional
condicional

c) i) $P(M \cap M) = 80/200 \cdot 79/199 = 0,1588$

ii) $P(C \cap C) = 150/200 \cdot 149/199 = 0,2977$

iii) $1 - P(B \cap B) = 1 - \frac{50}{200} \cdot \frac{49}{199} = 0,9384$

iv) $1 - P(M \cap M) = 1 - 80/200 \cdot 79/199 = 0,8412$

v) $P(M \cap M) = 80/200 \cdot 79/199 = 0,1588$

9- a) i) $P(A) = 19/551 = 0,0344$

ii) $P(L^c \cap P1) = 212/551 = 0,384$

iii) $P(O \cup P2) = P(O) + P(P2) - P(O \cap P2) = 450/551 + 289/551 - 250/551 = 0,8875$

iv) $P(A/P1) = 12/262 = 0,045$

v) $P(P2^c) = 1 - P(P2) = P(P1) = 262/551 = 0,4754$

vi) $P(P1/A) = 12/19 = 0,63$

- b) i) $P(O \cap O) = 450/551.449/550 = 0,66$
 ii) $1 - P(\overline{P2} \cap \overline{P2}) = 1 - 262/551.261/550 = 0,78$

10- a) $E = \{(E, E), (E, \overline{E}), (\overline{E}, E), (\overline{E}, \overline{E})\}$

b)

Resultado de la 1.ª perforación	Resultado de la 2.ª perforación	Probabilidad
Se encontró (petróleo o gas)	Se encontró (petr. o gas)	0,01
Se encontró	No se encontró	0,09
No se encontró	Se encontró	0,09
No se encontró	No se encontró	0,81

c) $P((\overline{E} \cap \overline{E})^c) = 1 - 0,81 = 0,19$

11- a) $E = \{(A, A, A), (A, A, NA), (A, NA, A), (A, NA, NA), (NA, AA), (NA, A, NA), (NA, NA, A), (NA, NA, NA)\}$

b) $P(A \cap A \cap A) = 0,804$

c) $P(NA \cap NA \cap NA) = 0,0003$

d) $1 - P(NA \cap NA \cap NA) = 0,9997$

12- a) $P(\text{ningún error}) = 0,874$

b) $P(2 \text{ errores}) = 0,004$

c) $P(\text{solo 1 error}) = 0,122$

13- Para el 6- 0,198

Para el 7- a) 0,04 b) 0,04 c) 0,64 d) 0,91 e) 0,36

Para el 8- c) i. 0,16 ii. 0,5625 iii. 0,9375 iv. 0,84 v. 0,16

Para el 9- b) i. 0,667 ii. 0,774

14- a) no

b) no

c) no

e) $P(A/B) = (1/6) / (3/6) = 1/3$ $P(B/A) = (1/6) / (1/6) = 1$

15- $P(M/NC) = 0,1/0,2 = 0,5$

16- a) $P(V) = 0,74$

b) $P(\overline{S}/V) = 0,757$

c) $P(\overline{V} \cup S) = 0,44$

17- a) $P(A \cap O) = 0,21$

b) $P(M) = 0,405$

c) $P(M \cup A) = 0,685$

- d) $P(A/M) = 0,777$
- 18- a) $P(NM) = 0,697$
 b) $P(V/NM) = 0,228/0,697 = 0,327$
- 19- a) $P(B) = 0,274$, el 27,4 %
 b) $P(M/B) = 0,232/0,274 = 0,8467$
- 20- a) $P(M \cap \bar{U}) = 0,08$
 b) $P(M/\bar{U}) = 0,08/0,11 = 0,7273$
 c) $P(\bar{U}/M) = 0,08/0,80 = 0,1$
- 21- a) $P(I) = 0,089375$
 b) $P(R1/U) = 0,060625/0,910625 = 0,06575$
 $P(R2/U) = 0,11875/0,910625 = 0,1304049$
- 22- a) $P(D) = 0,015$
 b) $P(M1/\bar{D}) = 0,005/0,985 = 0,0051$
 c) $P(M2/\bar{D}) = 0,01/0,985 = 0,0102$
- 23- a) $P(A/a) = 0,076$
 b) $P(B/a) = 0,435$
- 24- a) Diagramas
 b) $P(A/L) = 0,05/0,12 = 0,416$
- 25- a) $P(F/M) = 0,104/0,32 = 0,325$
 b) $P(A.F \cup M) = 0,58 + 0,32 - 0,104 = 0,796$
- 26- $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0,98$ despejando $P(A \cap B) = 0,298$
 $P(A/B) = P(A \cap B) / P(B) = 0,298/0,908 = 0,328$
 $P(B/A) = 0,298/0,37 = 0,805$
- 27- a) $P(CI) = 0,379$
 b) $P(S) = 0,777$
 c) $P(M/NO S) = 0,0552/0,2223 = 0,248$
- 28- $P(M/N) = P(M \cap N) / P(N) = 0,3$ despejamos
 $P(N) = 0,3/0,5 = 0,6$
 $P(N/M) = P(M \cap N) / P(M) = 0,65$ despejamos
 $P(M) = 0,3/0,650,462$
- 29- a) $P(O.C/M) = 0,3024/0,38 = 0,796$
 b) $P(M/N \cup O.C) = 0,0776/0,4 = 0,194$

c) $P(V \cup \text{no O.C}) = 0,62 + 0,4 - 0,3224 = 0,6976$

30- $P(R \cup M) = 0,45 = P(R) + P(M) - P(R \cap M)$ despejamos

$P(R) = 0,45 - 0,4 + 0,2 = 0,25$ Por diferencia

$P(O \cap R) = 0,05$ y $P(O \cap NR) = 0,55$

a) $P(O/R) = 0,05/0,25 = 0,2$

b) $P(NR/O) = 0,55/0,6 = 0,917$

31- a)

Caso	P(A)	P(B)	P(A∩B)	P(A∪B)	P(A/B)	P(B/A)	A y B son sucesos			
							equiprobables	independientes	mutuamente excluyentes	excluyentes
i	0,35	0,25	0,11	0,49	0,44	0,3143	NO	NO	NO	
ii	0,56	0,62	0,3472	0,8328	0,56	0,62	NO	SI	NO	
iii	0,05	0,05	0	0,1	0	0	SI	NO	SI	
iv	0,3	0,18	0,054	0,426	0,3	0,18	NO	SI	NO	
v	0,56	0,56	0,28	0,84	0,5	0,5	SI	NO	NO	
vi	0,66	0,1818	0,12	0,7218	0,66	0,1818	NO	SI	NO	
vii	0,42	0,23	0	0,65	0	0	NO	NO	SI	
viii	0,375	0,5192	0,27	0,6242	0,52	0,72	NO	NO	NO	

b) i. Si

ii. No

iii. Si

iv. No

32- $P\left(\frac{\bar{A}}{\bar{B}}\right) = \frac{P(\bar{A} \cap \bar{B})}{P(\bar{B})} = \frac{P(\overline{A \cup B})}{1 - P(B)} = \frac{1 - P(A \cup B)}{1 - P(B)}$

$P(B/A) = P(A \cap B) / P(A) = 0,5$ despejamos $P(A \cap B) = 1/8$

$P(A/B) = P(A \cap B) / P(B) = 1/4$ despejamos $P(B) = P(A \cap B) / P(A/B) = 1/2$

$P(A \cup B) = 0,625$

$P\left(\frac{\bar{A}}{\bar{B}}\right) = \frac{P(\bar{A} \cap \bar{B})}{P(\bar{B})} = \frac{P(\overline{A \cup B})}{1 - P(B)} = \frac{1 - 0,625}{1 - \frac{1}{2}} = 0,75$

$P(A/B) + P(\bar{A}/\bar{B}) = 1/4 + 0,75 = 1$ verdadero

33- a) $P(D) = 0,49$

b) $P(E/ND) = 0,4/0,51 = 0,784$

34- Si A y B son independientes $\frac{P(A/B) \cdot P(B/A)}{P(A \cap B)} = \frac{P(A) \cdot P(B)}{P(A) \cdot P(B)} = 1$

PRÁCTICA 6

Variable aleatoria

Una variable aleatoria es una asignación unívoca entre los sucesos resultantes de un experimento aleatorio y un conjunto numérico que se define con el objetivo de cuantificarlos. Esta síntesis numérica es la que permite, junto con el modelo probabilístico, determinar algunas características del comportamiento del fenómeno en estudio.

GPS ESTADÍSTICO 6

Orientaciones importantes para no perderse

- *Los valores de una función de probabilidad de una variable aleatoria discreta son probabilidades, pero, los de una función de densidad de probabilidad de una variable aleatoria continua no son probabilidades.*
- *La función de densidad permite calcular la probabilidad de cualquier intervalo del recorrido de la variable aleatoria continua a través del área encerrada por la función y los extremos de dicho intervalo, de ahí que cualquier probabilidad puntual para este tipo de variables es cero.*
- *La función de distribución mide la probabilidad acumulada a izquierda de un valor determinado de la variable, por lo que sus resultados siempre son probabilidades:
 $F(t)=P(X\leq t)$*

Glosario 6

Variable aleatoria o estocástica: es una función que asigna valores reales a cada suceso de un experimento aleatorio.

Variable aleatoria discreta: su imagen o recorrido es un subconjunto de los números enteros.

Variable aleatoria continua: su imagen o recorrido es un subconjunto de los números reales.

Función de probabilidad: asigna una probabilidad a cada valor de una variable aleatoria discreta.

Función de densidad de probabilidad: permite calcular las probabilidades para cualquier intervalo del recorrido de una variable aleatoria continua.

Función de distribución: asigna la probabilidad acumulada hasta él a cada valor de la variable aleatoria.

Problemas y situaciones prácticas 6

1- Los posibles valores que puede asumir una variable aleatoria discreta son:

$$R(x): \{0, 1, 2, 3, 4\}$$

¿Cuáles de las siguientes funciones es la función de probabilidad que le corresponde?

Justifique su respuesta, explicando también porqué rechaza a las otras dos.

$h_1(x):$	0,25	0,20	0,15	0,10	0,05
$h_2(x):$	0,30	0,10	0,15	0,25	0,30
$h_3(x):$	0,40	0,20	0,20	0,15	0,05

2- El número de personas que esperan ser atendidas en una farmacia es una variable aleatoria que tiene la siguiente función de probabilidad:

x	1	2	3	4	5	6
h(x)	0,05	0,2		0,35		0,1

- Completar la tabla sabiendo que la probabilidad de que haya 3 y 5 clientes esperando son iguales.
 - Calcular la probabilidad de tener al menos 4 clientes esperando.
 - Determinar la esperanza de la variable e interpretar en el contexto del problema.
- 3- Una pastelería ofrece tortas con decoración especial para cumpleaños, bodas y otras ocasiones; y también tiene pasteles comunes para la venta. En la tabla siguiente se proporciona el número total de tortas vendidas por día y los porcentajes correspondientes.

N.º de tortas vendidas por día	10	11	12	13	14
Porcentajes	15	35	25	15	10

Para un día cualquiera tomado al azar:

- Armar la tabla de distribución de probabilidad incluyendo la función de distribución acumulativa y graficar.
 - ¿Cuál es la probabilidad de que se vendan al menos 12 tortas?
 - Calcular el número de tortas que se espera vender y el desvío estándar.
- 4- La cantidad de pedidos que realiza una pizzería hasta las 18 h es una variable aleatoria con función de probabilidad dada por:

x	0	1	2	3	4
----------	----------	----------	----------	----------	----------

$h(x)$	0,3	0,2			0,05
--------	-----	-----	--	--	------

- Completar la tabla si se sabe que la probabilidad de recibir al menos 3 pedidos es 0,15.
- Si un día se recibieron no más de 2 pedidos en ese horario, calcular la probabilidad de que se hayan hecho por lo menos 1.
- Determinar el número esperado de pedidos.

- 5- La cantidad de artículos defectuosos que produce una máquina en un día es una variable aleatoria cuya función de probabilidad está representada en la siguiente tabla:

X	0	1	2	3	4
$h(x)$		0,25	0,12	0,13	

- Completar la tabla sabiendo que la probabilidad de no obtener artículos defectuosos es el cuádruple de tener 4 defectuosos.
- Calcular la probabilidad de que se produzcan más de 3 artículos defectuosos.
- Calcular $E(X)$ y $V(X)$.

- 6- Se ha determinado que el número de camiones que llegan cada hora a un depósito tiene la siguiente distribución de probabilidad:

N.º de camiones	0	1	2	3	4	5	6
Probabilidad	0,05	0,10	0,15			0,10	0,01

Para una hora cualquiera tomada al azar:

- Calcular $P(3)$ y $P(4)$ sabiendo que la probabilidad de que lleguen a lo sumo tres camiones es de 0,55
- Calcular el número esperado de llegadas de camiones y la varianza.

- 7- La tabla muestra la demanda diaria, en miles de unidades, de cierto artículo en un determinado período de tiempo.

Unidades demandadas	3	4	7	10	12	14
Probabilidad	0,06	0,08		0,20	0,24	

Para un día cualquiera tomado al azar:

- Completar la tabla sabiendo que la demanda de a lo sumo 7 mil unidades es igual a 0,20
- Calcular la demanda esperada del artículo e interpretar el resultado.

- 8- Una industria desea instalar una nueva planta y está decidiendo entre dos lugares posibles. Se sabe por investigaciones previas que en la primera localidad se puede lograr una ganancia de USD 2.000.000 si tienen éxito y una pérdida de USD 200.000 luego de un año. En la otra localidad la ganancia anual podría ser una suma de USD 2.500.000 y la pérdida de USD 500.000

Si la probabilidad de tener éxito en la industria es de 0,5 para la primera localidad y de 0,6 para la segunda, se desea saber dónde conviene más instalar la sucursal.

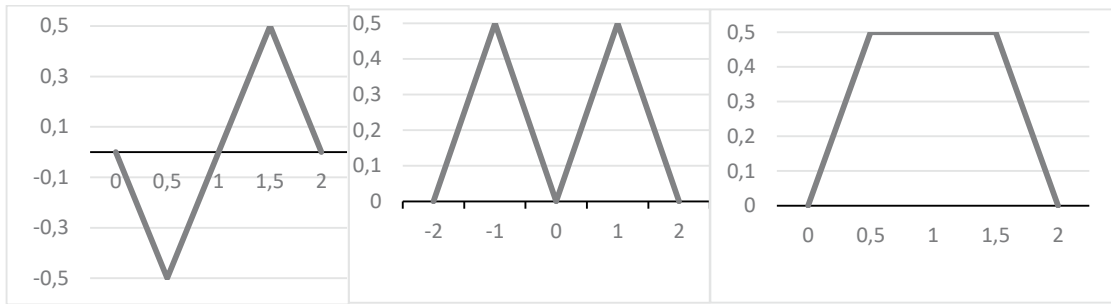
- 9- La probabilidad de que una casa de cierto tipo quede destruida por un incendio en cualquier período de 12 meses es 0,005. Una compañía de seguros ofrece vender al propietario de tal casa una póliza de seguro contra incendio por 100.000 dólares y a un año por una prima de 3.000 dólares. ¿Cuál es la ganancia esperada de la compañía para una casa elegida al azar entre las aseguradas?
- 10- Una fábrica que produce lámparas led verifica que en el control de calidad el 0.2 % de estas presenta alguna falla. Al producir un artículo defectuoso el fabricante pierde \$1, mientras que el artículo en perfecto estado le brinda una utilidad de \$5. ¿Cuál es la ganancia esperada por artículo?
- 11- La probabilidad de que un hombre de 50 años, en un año, no tenga un accidente grave de trabajo que lo inhabilite de por vida, es de 0,995. ¿Qué prima debería cobrarle la compañía de seguros a cualquier hombre de 50 años para una póliza de seguros por accidente de USD 100.000 por un año, si la prima debe incluir un beneficio?
- 12- Una inversión especulativa de USD 1.500, puede valer USD 1.000, USD 2.000 o USD 5.000 a fin de año. Las probabilidades de estos valores son 0,25; 0,60 y 0,15 respectivamente.
- ¿Cuál es la ganancia esperada en esta inversión?
 - ¿Cuál es el valor esperado de la inversión a fin de año?
- 13- Utilizando la información del ejercicio 9,
- calcular la ganancia esperada por la compañía suponiendo que asegura 135 casas del mismo tipo.
 - ¿Qué propiedad de la esperanza aplicó en el ítem anterior? Justificar.
 - Explicar la diferencia entre el experimento de este ejercicio comparado con el experimento del ejercicio 9
- 14- Utilizando la información del ejercicio 10 de la practica 5, donde la probabilidad de encontrar gas o petróleo en una perforación era de 0,10 si la empresa perforara 3 pozos,
- hallar la distribución de probabilidad de la cantidad de pozos donde se encontró gas o petróleo.
 - Calcular la esperanza e interpretar.
- 15- El departamento de alumnos de una universidad informó que la cantidad de materias en las que se anota un alumno regular por cuatrimestre para una carrera determinada puede ser 1, 2 o 3 con la siguiente distribución de probabilidad:

Cantidad de materias	1	2	3
$h(x)$	0,3	0,5	0,2

- a) Determinar la distribución de probabilidad de la cantidad de materias en que se anota un alumno en 2 cuatrimestres suponiendo que dicha cantidad es independiente de un cuatrimestre a otro.
- b) ¿Cuál es la probabilidad de que un alumno curse como máximo 4 materias al año?
- c) ¿Cuántas materias se espera que curse el alumno en el año?
- 16-** Utilizando la información del ejercicio 7 de la práctica 5 (ente 10 lápices, hay 4 rojos, 2 verdes, 1 amarillo y 3 azules).
- a) Armar la tabla de distribución de probabilidades de la cantidad de lápices rojos que podrían extraerse al seleccionar dos lápices.
- b) ¿Cuál es la probabilidad de hallar al menos un lápiz rojo?
- c) ¿Cuántos lápices rojos esperan extraerse?
- 17-** Una imprenta observa que el 90 % de las páginas están impresas correctamente, de aquellas páginas que no están bien impresas, el 20 % presenta solo fallas de alineación, el 40 % presenta solo colores tenues y el resto ambos defectos.
- a) Encontrar la distribución de probabilidad que describa la cantidad de defectos por hoja.
- b) Hallar la esperanza e interpretar.
- 18-** En un centro médico se observa especialmente la calidad de atención que reciben los pacientes. En un estudio sobre el tiempo de espera para ser atendidos se ha registrado la siguiente información:

Tiempo	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40
Pacientes	12	19	28	36	25	17	14	9

- a) A partir de dicha información construir la función de densidad de la v. a. tiempo de espera.
- b) Si se selecciona un paciente al azar hallar la probabilidad de que espere:
- menos de 15 minutos,
 - al menos 25 minutos,
 - entre 10 y 37 minutos,
 - más de 5 minutos si esperó menos de 20
- c) Calcular la esperanza e interpretar.
- 19-** Los siguientes gráficos corresponden a diferentes funciones reales. Decidir cuáles de ellos podrían ser una función de densidad de probabilidad asociada a una variable aleatoria continua y explicar por qué.



20- La siguiente función describe el consumo bimestral de energía eléctrica en miles kW de una empresa:

$$f(x) = \begin{cases} \frac{1}{2} - \frac{1}{8}x & \text{para } 0 < x < 4 \\ 0 & \text{para otro caso} \end{cases}$$

- Verificar si es una función de densidad.
- Calcular la probabilidad de que un bimestre la empresa consuma:
 - más de 3000 kW, ii) menos de 300 kW, iii) entre 1000 y 2500 kW, iv) menos de 2000 kW si sabiendo que ya gastó más de 1500 kW
- Calcular el consumo esperado y el consumo mediano e interpretar.
- Hallar la varianza.

21- Sea x una variable aleatoria cuya función de densidad es:

$$f(x) = \begin{cases} k \cdot x & \text{si } 0 < x < 5 \\ 0 & \text{para otros valores} \end{cases}$$

- Hallar el valor de k para que $f(x)$ sea función de densidad.
- Calcular i) $P(x < 1)$, ii) $P(2 < x < 4)$, iii) $P(x > 2)$
- Hallar la función de distribución acumulativa.
- Calcular las probabilidades de b) usando $F(x)$
- Calcular $E(x)$, Me y $V(x)$

22- El tiempo de impresión de una página medido en segundos es una variable aleatoria cuya función de densidad es:

$$f(x) = \begin{cases} 2x + 1 & \text{si } 0,5 < x < 1,5 \\ 0 & \text{para otros valores} \end{cases}$$

- Verificar que es una función de densidad.
- Calcular la función de distribución acumulativa.
- Calcular la probabilidad de que una hoja se imprima en: i) más de 1 segundo, ii) menos de 0,7 segundos.

d) Hallar $E(X)$ y Me e interpretar.

23- El tiempo de vida, en días de uso continuo, de una batería es una variable aleatoria con función de distribución de probabilidad:

$$F_x(t) = \begin{cases} 0 & \text{si } t < 1 \\ -\frac{t^2}{4} + \frac{3}{2}t - \frac{5}{4} & \text{si } 1 \leq t < 3 \\ 1 & \text{si } t \geq 3 \end{cases}$$

- Calcular la probabilidad de que una batería seleccionada al azar funcione por lo menos durante dos días.
- Calcular la probabilidad de que una batería seleccionada al azar funcione entre dos y tres días sabiendo que lleva funcionando a lo sumo dos días y medio.

24- El tiempo en minutos, requerido para la fabricación de un vaso de vidrio es una variable aleatoria cuya función de densidad es:

$$f(x) = \begin{cases} a \cdot (x^2 - 3 \cdot x) & \text{si } 0 \leq x < 3 \\ 0 & \text{para otro } x \end{cases}$$

- Calcular el valor de la constante a para que $f(x)$ sea una función de densidad.
- Calcular la esperanza e interpretar el valor obtenido.
- ¿Qué porcentaje de los vasos se producen en menos de 2 minutos?

25- El diámetro, en milímetros de ciertos ejes es una variable aleatoria cuya función de densidad es:

$$f(x) = \begin{cases} a(x - 3)^2 & \text{para } 1\text{mm} < x < 5\text{mm} \\ 0 & \text{sino} \end{cases}$$

Calcular:

- El valor que debe tomar la constante “ a ” para que f sea una función de densidad.
- La probabilidad de que el diámetro supere los 4mm si se sabe que se encuentra entre 2 y 5 mm.

26- Sea x una variable aleatoria cuya función de densidad es:

$$f(x) = \begin{cases} cx^3 & 0 < x < 2 \\ 0 & \text{en otro caso} \end{cases}$$

Hallar:

- El valor de la constante “ c ” para que f sea una función de densidad.
- La función de distribución $F(t)$.
- La esperanza, la mediana y la varianza.

d) La probabilidad de que la variable esté a, como mínimo, 0,5 unidades de distancia de 0,9

RESPUESTAS DE LAS ACTIVIDADES DE LA PRÁCTICA 6

Variable aleatoria

1- $h_3(x)$ es una función de probabilidad puntual.

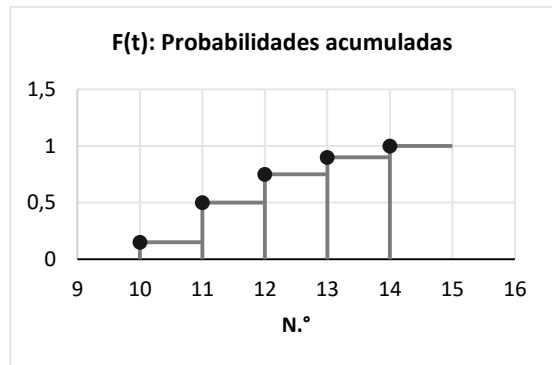
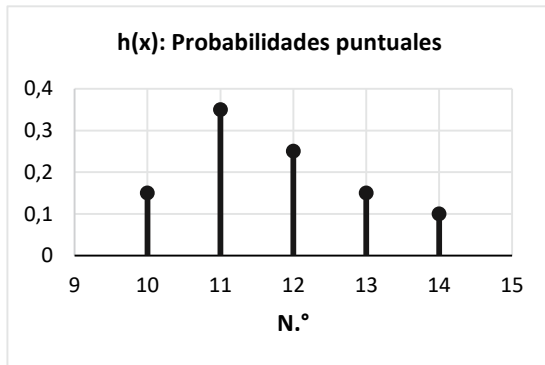
2- a) $P(X=3) = P(X=5) = 0,15$

b) $P(X \geq 4) = 0,6$

c) $E(X) = 3,65$

3- a)

X	10	11	12	13	14
$h(x)$	0,15	0,35	0,25	0,15	0,10
$F(t)$	0,15	0,50	0,75	0,90	1



b) $P(X \geq 12) = 0,50$

c) $E(X) = 11,7$ $DS(X) = 1,187$

4- a) $P(X=3) = 0,10$ $P(X=2) = 0,35$

b) $P(X \geq 1/X \leq 2) = 0,76471$

c) $E(X) = 1,7$

5- a) $P(X=4) = 0,1$ $P(X=0) = 0,4$

b) $P(X > 3) = 0,1$

c) $E(X) = 1,28$ $V(X) = 1,86$

6- a) $P(X=3) = 0,25$

$P(X=4) = 0,34$

b) $E(X)=3,07$ $V(X)=1,8251$

7- a) $P(X=7) = 0,06$ $P(X=14) = 0,36$
 c) $E(X)=10,84$

8- En la segunda, dado que: $E(X_1)= 900000$ y $E(X_2)=1300000$

9- $E(X) = 2500$

10- $E(X)=4,998$

11- Para que $E(X) > 0$ (o sea que haya alguna ganancia) debe ser la Prima > 500 USD

12- $E(X_1)=700$ USD $E(X_2)=2.200$ USD

13- a) $E(135.X)= 135.2500$
 b) $E(a.X)=a.E(X)$

14- a)

X	h(x)	F(t)
0	0,729	0,729
1	0,243	0,972
2	0,027	0,999
3	0,001	1
	1	

b) $E(X)= 0,3$

15- a)

X	h(x)	F(t)
2	0,09	0,09
3	0,30	0,39
4	0,37	0,76
5	0,20	0,96
6	0,04	1

b) $P(X \leq 4) = 0,76$
 c) $E(X)=3,8$

16- a)

X	h(x)	F(t)
0	0,3333	0,3333

1	0,5334	0,8667
2	0,1333	1
	1	

b) $P(X \geq 1) = 0,6667$

c) $E(X) = 0,8$

17- a)

x	h(x)	F(t)
0	0,90	0,90
1	0,06	0,96
2	0,04	1

b) $E(X) = 0,14$

18- b) i) $P(X < 15) = 0,36875$

ii) $P(X \geq 25) = 0,25$

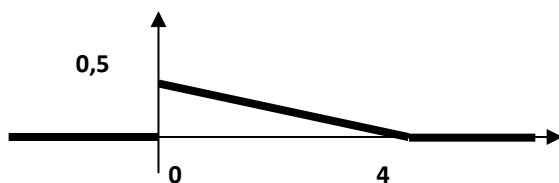
iii) $P(10 < X < 37) = 0,7725$

iv) $P(X > 5 / X < 20) = 0,8723$

c) $E(X) = 18,7$

19- El segundo gráfico.

20- a) $f(x) = 0,5 - 0,125 \cdot x$ para $0 < x < 4$



se observa en el gráfico que $f(x) \geq 0$

y el área=1 calculada con integrales o haciendo $b \cdot h / 2$

b) i) $P(x > 3) = 0,0625$

ii) $P(X < 0,3) = 0,14437$

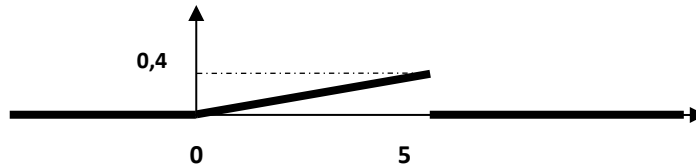
iii) $P(1 < X < 2,5) = 0,421875$

iv) $P(X < 2 / X > 1,5) = 0,36$

c) $E(X) = 2,67$ y $Me = 1,17$ ($Me = 6,83$ no verifica porque se va del dominio).

d) $V(X) = 0,89$

21- a) $k = 0,08$

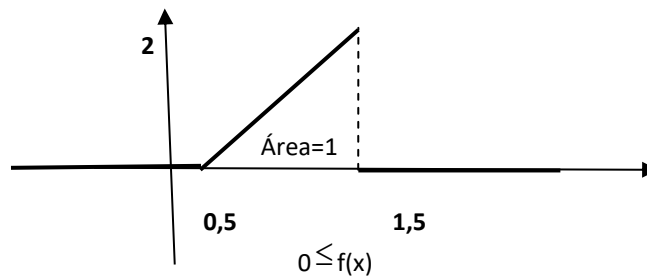


- b) i) $P(x < 1) = 0,04$
 ii) $P(2 < x < 4) = 0,48$
 iii) $P(x > 2) = 0,84$

$$c) F_x(t) = \begin{cases} 0 & \text{si } t < 0 \\ 0,04t^2 & \text{si } 0 \leq t < 5 \\ 1 & \text{si } t \geq 5 \end{cases}$$

- d) i) $P(x < 1) = P(x \leq 1) = F(1) = 0,04$
 ii) $P(2 < x < 4) = F(4) - F(2) = 0,48$
 iii) $P(x > 2) = 1 - F(2) = 0,84$
 e) $E(X) = 3,33$ $V(X) = 1,41$ $Me = 3,535$

22- a) $f(x) = 2x - 1$



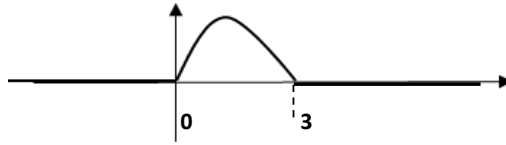
$$b) F_x(t) = \begin{cases} 0 & \text{si } t < 0 \\ t^2 - t + 0,25 & \text{si } 0,5 \leq t < 1,5 \\ 1 & \text{si } t \geq 1,5 \end{cases}$$

- c) i) $P(X > 1) = 1 - F(1) = 0,75$
 ii) $P(X < 0,7) = F(0,7) = 0,04$
 d) $E(X) = 1,17$ y $Me = 1,2$ porque $-0,2$ es un valor donde $f(x)$ es nula (no ocurre).

23- a) $P(X \geq 2) = 1 - F(2) = 0,25$

b) $P(2 < X < 3 / X \leq 2,5) = 0,2$

24- a) $a = -2/9$

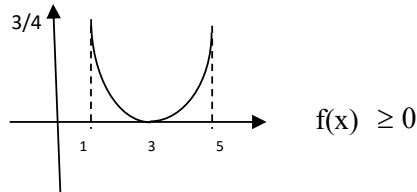


b) $E(X)=3/2=1,5$

c) $P(X<2)=20/27=0,74$ El 74 %

25- a) $a=3/16$

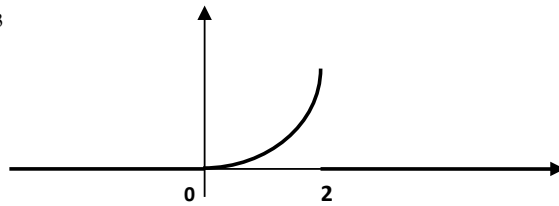
$f(x)=3/16.(x-3)^2$



b) $P(X > 4 / 2 < X \leq 5)=0,4375$

26- a) $c=1/4$

$f(x)=1/4.x^3$



$$b) F_x(t) = \begin{cases} 0 & \text{si } t < 0 \\ \frac{1}{8}t^4 & \text{si } 0 \leq t < 2 \\ 1 & \text{si } t \geq 2 \end{cases}$$

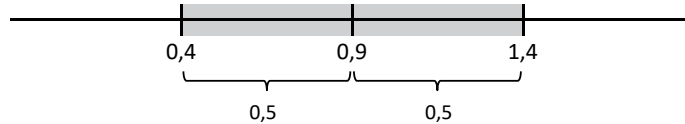
c) $E(X)= 1,6$

$Me=1,68$

(El valor -1,68 se descarta)

$V(X)=0,11$

d)



$P(0,4<X<1,4)= F(1,4)-F(0,4) = 0,523$

PRÁCTICA 7

Distribuciones especiales de probabilidad

Un ejemplo notable de modelización estadística a partir de un problema práctico son las distribuciones de probabilidad, que permiten describir en forma sintética y formal el comportamiento de las distribuciones empíricas de datos estadísticos y hacer predicciones sobre este.

Variables discretas

Modelos especiales de probabilidad para variables aleatorias discretas: *distribución binomial, hipergeométrica y de Poisson. Aproximación de la distribución binomial de Poisson.*

GPS ESTADÍSTICO 7

Orientaciones importantes para no perderse

- *La distribución binomial formalmente se aplica a observaciones sucesivas donde sus resultados son independientes entre sí. Si se realizan selecciones con reposición esto queda garantizado.*
- *Igualmente se puede aplicar cuando una muestra proviene de una población de gran tamaño, aunque la extracción se realice sin reposición, hecho que en otras circunstancias implicaría la obtención de resultados dependientes. La independencia en caso de trabajar sobre lotes muy extensos se verifica de manera práctica o “empírica” pero no teórica.*
- *Por lo anterior se concluye que la distribución hipergeométrica tiende a convertirse en (o sea que converge a) una binomial al considerar poblaciones de gran tamaño y al asumir el valor de $p=R/N$ que se mantiene casi constante en las sucesivas observaciones (el ser casi constante permite usar la aproximación).*
- *La distribución de Poisson tiene la particularidad de producir resultados independientes si los sucesos se observan en intervalos continuos no superpuestos.*

Glosario 7

Experimento de Bernoulli: experimento aleatorio *simple* que tiene solo dos resultados posibles, donde se identifica uno como *éxito* y como *fracaso* al otro.

Experimento binomial: es un experimento compuesto por repeticiones de un experimento simple de Bernoulli y donde las n repeticiones son *independientes*.

Variable aleatoria binomial: cuenta la cantidad de éxitos en n repeticiones independientes de un experimento de Bernoulli.

Experimento hipergeométrico: es un experimento compuesto por repetición de un experimento simple de Bernoulli y donde las n repeticiones no son *independientes*.

Variable aleatoria hipergeométrica: cuenta la cantidad de éxitos en n repeticiones dependientes de un experimento de Bernoulli.

Proceso aleatorio de Poisson: es un proceso de naturaleza binomial donde los éxitos ocurren a lo largo de (o sea asociados a) un intervalo continuo y el fracaso es la no ocurrencia de éxito.

Variable aleatoria de Poisson: cuenta la cantidad de éxitos ocurridos en intervalos continuos.

Problemas y situaciones prácticas 7

- **Distribución binomial**

1- Si se tira una moneda 6 veces,

- a) ¿Cuál es la probabilidad de obtener una cara?
- b) ¿Cuál es la probabilidad de obtener al menos una cara?
- c) Calcular el número esperado de caras.

2- En la sección de computación de un hipermercado, la probabilidad de que un cliente que pregunta acerca de los precios de varios artículos compre alguno es de 0,3. Si se detecta que 18 clientes consultan por los precios en la sección de accesorios para PC y celulares por hora,

- a) ¿Cuál es la probabilidad de que al menos 4 clientes compren algo?
- b) ¿Cuántas personas se espera que compren algo?

3- Una empresa de prospección petrolera encuentra gas o petróleo en el 10 % de sus perforaciones. Si la empresa perfora tres pozos en cuencas geológicas totalmente independientes,

- a) Encontrar la distribución de probabilidad de que la compañía encuentre gas o petróleo.
- b) Graficar.
- c) Calcular la probabilidad de que la compañía encuentre gas o petróleo en por lo menos una de las tres perforaciones.

- 4- Una empresa que arma computadoras vende en el mismo día 12 equipos. La experiencia previa indica que la probabilidad de que este tipo de equipos esté funcionando correctamente 1 año después de la venta, es decir, antes de vencerse su garantía, es de 0,95. Calcular la probabilidad de que al vencerse la garantía:
- 2 equipos estén fuera de servicio.
 - A lo sumo esté un solo equipo fuera de servicio.
 - Al menos 5 equipos funcionen correctamente.
 - Calcular el valor esperado y el desvío. Interpretar.
- 5- En cierta empresa se ha encontrado que el 25 % de las operaciones comerciales que se realizan no producen el mínimo de ganancia deseado, determinar la probabilidad de que, si se seleccionan al azar 5 de dichas operaciones comerciales, siendo estas estadísticamente independientes, resulte que:
- 2 estén por debajo del nivel de ganancia deseado.
 - Todas produzcan un nivel de ganancia satisfactorio.
 - Al menos 1 no produzca el nivel de ganancia deseado.
- 6- Un vendedor de seguros sabe que la oportunidad de vender una póliza es mayor mientras más contactos realice con clientes potenciales. Si la probabilidad de que una persona compre una póliza de seguro después de la visita es constante e igual a 0,25 y si el conjunto de visita constituye un conjunto independiente de ensayos, ¿cuántos compradores potenciales debe visitar el vendedor para que la probabilidad de vender por lo menos una póliza sea de 0,80 como mínimo?
- 7- Una línea aérea ha observado que 5 % de los pasajeros que tienen reservaciones para cierto vuelo no se presentan a tomar el avión. Si la aerolínea vende 160 boletos para un vuelo con solamente 155 asientos, ¿cuál es la probabilidad de que todos los pasajeros que se presentan con una reservación y deseen viajar obtengan asiento?

• **Distribución hipergeométrica**

- 8- Con el propósito de verificar la exactitud de sus estados financieros, las compañías tienen auditores permanentes para verificar los asientos contables. Suponiendo que hubiera 6 asientos erróneos en un grupo de 20 asientos y un auditor verifica 4 asientos elegidos al azar de esos 20:
- Encontrar la distribución de probabilidad del número de errores detectados.
 - Graficar.
 - Calcular la probabilidad de que el auditor encuentre menos de un error.
- 9- Un supermercado tiene 10 impresoras de las cuales 3 han presentado algún tipo de falla. La compañía de reparaciones selecciona al azar 5 de ellas. ¿Cuál es la probabilidad de que ninguna de las 5 máquinas presente falla?

- 10- En el relevamiento previo a un censo de viviendas se han detectado 15 casas y 18 departamentos en una manzana en particular. ¿Cuál es la probabilidad de que en una muestra de 6 viviendas de dicha manzana:
- a) aparezcan 4 casas?
 - b) aparezca igual número de casas que de departamentos?
- 11- En un depósito hay 55 paquetes de hilo, todos los paquetes son del mismo tamaño, están herméticamente cerrados y ninguno tiene identificación. Veinte de los paquetes contienen hilo de uso industrial y el resto hilo de uso comercial. Si se seleccionan al azar y juntos 8 paquetes de hilo calcular la probabilidad de que entre ellos halla:
- a) Un paquete de hilo comercial.
 - b) La misma cantidad de paquetes de hilo industrial que de hilo comercial.
- 12- Un gerente selecciona al azar 3 personas del conjunto de 10 empleados de su departamento, para asignarlos a un estudio de clasificación de sueldos. Suponiendo que se asignó a 4 de los empleados a un proyecto similar, determinar la probabilidad de que exactamente 2 de los 3 empleados hayan tenido experiencia previa.
- 13- En una convención nacional se reunieron 40 representantes de las Cámaras de Comercio de las distintas provincias argentinas, para elegir la autoridad máxima del organismo entre dos candidatos. Si entre los representantes 25 apoyan al candidato A y 15 al B y seleccionan aleatoriamente 5 de ellos, ¿cuál es la probabilidad de que, entre estos cinco, por lo menos 2 apoyen al candidato A?
- 14- Una empresa tiene 15 vehículos para reparto en forma de *delivery*. Si en un control se detecta que 6 de los vehículos están en falta con la verificación y se seleccionan 4 vehículos al azar, ¿cuál es la probabilidad que dos de los vehículos tengan frenos defectuosos?

- **Distribución de Poisson**

- 15- Se estima que, entre las 18 h y las 21 h, arriban a cierto puesto de peaje 120 automóviles por hora. Calcular la probabilidad de que:
- a) Llegue un auto en un período de 2 minutos.
 - b) Lleguen menos de 2 autos en un período de 3 minutos.
- 16- La llegada de barcos pesqueros a puerto Madryn, en determinada época del año, corresponde a un proceso de Poisson caracterizado por una intensidad media de dos arribos por hora. Calcular la probabilidad de que:
- a) Transcurran 4 horas sin que se produzca la llegada de un barco pesquero.
 - b) Que como máximo lleguen 3 barcos pesqueros durante un intervalo de 3,5 horas.
- 17- En una estación de servicio se sabe que en promedio se reciben 3 billetes falsos por mes.
- a) ¿Cuál es la probabilidad de que en 15 días se reciban 3 billetes falsos?

- b) Si hoy alguien pagó con un billete falso, ¿cuál es la probabilidad de que pasen 10 días sin que vuelva a ocurrir esto?
- c) ¿Cuál es la probabilidad de que transcurra una semana sin recibir ningún billete falso?

18- En promedio, cada rollo de 500 metros de acero laminado tiene 2 defectos. Un defecto es una raspadura o alguna otra irregularidad que afectaría el uso de ese segmento de la hoja de acero en el producto terminado. ¿Cuál es la probabilidad de que un segmento específico de 100 metros tenga al menos 2 defectos?

19- La deficiencia en el número de glóbulos rojos en sangre puede determinarse mediante el examen de una muestra de sangre. Se considera que para personas normales en un mililitro de sangre se encuentran en promedio 5 millones de glóbulos rojos, ¿cuál es la probabilidad de que una muestra de 3 mililitros de una persona normal contenga 14 millones de glóbulos rojos?

20- Se sabe que un líquido contiene bacterias a razón de 2 bacterias por cm^3 . Calcular la probabilidad de que en una muestra de 3 cm^3 no haya ninguna bacteria.

- **Mix de distribuciones de variables discretas**

21-En una empresa 80 de los 150 empleados tienen más de 5 años de antigüedad en su puesto de trabajo. Si se eligen aleatoriamente 4 empleados de este grupo, ¿cuál es la probabilidad de que al menos 2 de ellos tengan más de 5 años de antigüedad en su puesto de trabajo?

22- Una compañía de seguros está considerando la adición de cobertura para una enfermedad relativamente rara en el campo de los seguros médicos mayores. La probabilidad de que una persona elegida al azar tenga esa enfermedad es de 0,0015 y en el grupo asegurado existen 2.750 personas.

- a) ¿Cuál es el número esperado de personas que tienen la enfermedad?
- b) ¿Cuál es la probabilidad de que ninguna persona de las 2.750 tenga la enfermedad?

23- Se sabe que la probabilidad de ganar en un juego de cartas *on-line* es de 0,0025. Si en una exhibición participan simultáneamente 12.000 jugadores de todo el mundo,

- a) ¿cuál es la probabilidad de que haya algún ganador?
- b) ¿y de que ganen por lo menos 5 personas?
- c) ¿el número esperado de ganadores?

24- Una editorial realiza una campaña telefónica con el propósito de aumentar la cantidad de suscriptos para una nueva revista que lanzó al mercado. La experiencia previa indica que 2 de cada 50 personas que reciben la llamada se suscribe a la revista. Si en un día dado, 20 personas reciben la llamada telefónica,

- a) ¿cuál es la probabilidad de que por lo menos 3 de ellas se suscriban?
- b) ¿cuál es el número esperado de personas que se suscriben a la revista?

- 25- Un club de automovilistas comienza una campaña telefónica para aumentar el número de miembros. En base a experiencias previas, se sabe que 1 de cada 20 personas que reciben la llamada se une al club. Si en un día 25 personas reciben la llamada telefónica:
- Calcular la probabilidad de que por lo menos 2 de ellas se inscriban al club.
 - Calcular la esperanza matemática e interpretar.
- 26- En un grupo de 160 postulantes para un trabajo, 100 no están oficialmente matriculados. Al elegir 30 de ellos al azar para responder un *test*, ¿cuál es la probabilidad de que a lo sumo 5 de ellos no estén oficialmente matriculados? Indicar el modelo utilizado. Justificar.
- 27- La probabilidad de que un empleado de un banco cometa un error en el ingreso de su clave es 0,001. Si se seleccionan 120 empleados al azar de diferentes entidades bancarias, ¿cuál es la probabilidad de que al menos 4 empleados comentan un error al ingresar su clave?
- 28- Una inmobiliaria tiene datos acerca de 40 inversores que se atrasaron en el pago de las cuotas de un grupo de 100 inversionistas. Si seleccionan al azar 30 personas de dicho grupo, cuál es la probabilidad de que a lo sumo 4 personas se atrasen en el pago de las cuotas.
- 29- Los cuatro focos de un automóvil fallan, cada uno con probabilidad 0,05, en forma independiente, durante un recorrido de 10.000 kilómetros. El automóvil puede funcionar sin riesgos en la oscuridad de una ruta mientras funcionen sin fallar por lo menos dos focos:
- ¿Cuál es la probabilidad de que el automóvil pueda avanzar sin riesgos en la oscuridad?
 - Calcular el valor esperado, el desvío e interpretar.
- 30- Una fábrica de artefactos electrónicos entrega su producto en lotes de gran tamaño. El comprador habitual desea rechazar los lotes que contienen una proporción alta de artefactos defectuosos. A tal fin selecciona al azar y controla 15 de ellos de cada lote. Si encuentra 3 o más defectuosos rechaza el lote entero.
- ¿Cuál es la probabilidad de que el lote sea rechazado si contiene el 5 % de artefactos defectuosos?
 - ¿Cuál es el número esperado de artefactos defectuosos en dicha muestra?
- 31- Indicar si los siguientes enunciados son V o F. Justificar.
- En un experimento binomial los sucesivos resultados son independientes entre sí por lo que la probabilidad de éxito se mantiene constante.
 - Una variable con distribución de Poisson permite medir el tiempo de espera entre dos hechos sucesivos.
 - En un experimento hipergeométrico los sucesivos resultados dependen entre sí por lo que la probabilidad de éxito cambia de ensayo a ensayo.

32- Los empleados de una compañía se encuentran separados en dos categorías: monotributistas y autónomos. La siguiente tabla conjunta indica la proporción de empleados en cada división clasificados por su relación con su empleador:

	Relación de dependencia	Independiente
Monotributista	0,08	0,12
Autónomo	0,24	0,56

- Si se seleccionan 5 empleados al azar, calcular la probabilidad de que al menos uno sea monotributista si trabaja en relación de dependencia.
- Si se seleccionan 10 empleados monotributistas al azar, calcular la probabilidad de que a lo sumo uno trabaje en forma independiente.

33- ¿Cuál distribución de probabilidad (binomial, hipergeométrica o de Poisson) es la que tiene mayores posibilidades de utilizarse con las siguientes variables?

- El número de autos que pasan por el semáforo ubicado en el cruce de las dos avenidas que limitan el predio de la universidad entre las 18:40 horas y las 19:15 horas.
- El número de tubos fluorescentes defectuosos que habría entre los tomados, de una partida de 100, para reemplazar a los quemados de las distintas aulas.
- La cantidad de alumnos de la UNLaM ausentes durante una semana cualquiera del primer cuatrimestre.
- De una muestra de 10 estudiantes de la UNLaM, se cuenta el número de alumnos que tienen trabajo.

34- En base a una encuesta efectuada entre los empleados de una empresa textil se obtuvieron las siguientes frecuencias relativas, donde las variables que se cruzaron son el género y la condición de ser o no sostén principal de su hogar.

	Masculino	Femenino
Es principal sostén de su hogar	0,34	0,23
No es principal sostén de su hogar	0,14	0,29

Si se eligen 6 empleados al azar,

- Calcular la probabilidad de que uno o más sean el principal sostén de su hogar si corresponden al género masculino.
- Calcular el número esperado de hombres que son el principal sostén de su hogar.

35- La fotocopiadora del negocio “XT” requiere en promedio una reparación cada 700 horas de trabajo y se prevé que en el próximo cuatrimestre se trabajarán cuatro meses, de lunes a sábado, en jornadas de 12 horas.

- ¿Cuál es la probabilidad de que la fotocopiadora no requiera reparaciones en todo el cuatrimestre?
- ¿Cuál es la probabilidad de requerir, a lo sumo, tres reparaciones en el cuatrimestre?
- ¿Cuántas reparaciones se espera que necesite tener la máquina en un cuatrimestre?

RESPUESTAS DE LAS ACTIVIDADES DE LA PRÁCTICA 7

Distribuciones especiales discretas de probabilidad

Distribución binomial

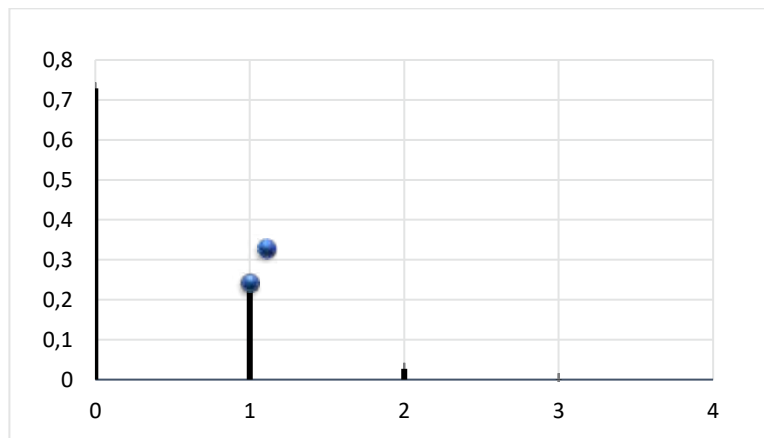
- 1- a) 0,09375
b) 0,984375
c) 3

- 2- a) 0,83545
b) 5,4

- 3- a)

x	0	1	2	3
h(x)	0,729	0,243	0,027	0,001
F(x)	0,729	0,972	0,999	1

- b)



Se usa el gráfico de escalones para representar $F(X)$.

- c) 0,271

- 4- a) 0,09879
b) 0,88164
c) 0,99999
d) 11,4 / 0,75

- b) 0,2373
c) 0,76269

- 6- 5,59

- 5- a) 0,26367

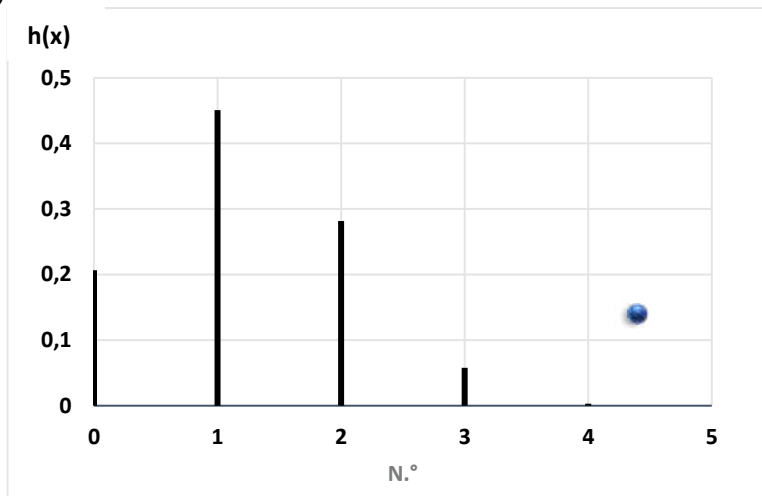
- 7- 0

Distribución hipergeométrica

8- a)

X	0	1	2	3	4
h(x)	0,2066	0,4508	0,2817	0,0578	0,0031
F(x)	0,2066	0,6574	0,9391	0,9969	1

b)



Se puede realizar un gráfico de escalones para representar $F(X)$.

c) 0,2066

9- 0,08333

12- 0,3

10- a) 0,18856

b) 0,33522

13- 0,94357

11- a) 0,00223

b) 0,20835

14- 0,3956

Distribución de Poisson

15- a) 0,07326

b) 0,01735

18- 0,06155

16- a) 0,00033

b) 0,06657

19- 0,10244

20- 0,00248

17- a) 0,12551

b) 0,36788

c) 0,49658

Mix de distribuciones de variables discretas

21- 0,73

22- 0,0161

23- a) 1
b) 0,99999
c) 30

24- a) 0,04386
b) 0,8

25- a) 0,35762
b) 1,25

26- 0,000000343

27- 0,00000785

28- 0,00151

29- a) 0,99952
b) 3,8 / 0,44

30- a) 0,0362
b) 0,75

31- a) V b) F c) V

32- a) 0,76269
b) 0,00168

33- a) Poisson
b) Binomial
c) Poisson
d) Hipergeométrica

34- a) 0,99940
b) 4,26

35- a) 0,1466
b) 0,87126
c) 1,92

PRÁCTICA 8

Distribuciones especiales de probabilidad Variables continuas

Uno de los problemas de gran interés entre algunos matemáticos de los siglos XVIII y XIX (como Lagrange y Gauss) fue el del ajuste de curvas a los datos, dando surgimiento a los modelos de las distribuciones continuas de probabilidad.

Las distribuciones de probabilidad como modelo son la base de la inferencia estadística y es en las distribuciones binomial y normal en las que se apoya la inferencia clásica.

Otros posibles modelos de distribución son la distribución uniforme -que provee la base teórica para estudiar la propagación de los errores de redondeo- y la distribución exponencial –íntimamente vinculada a un proceso de Poisson– que se utiliza en modelos de Costos y en investigación operativa.

Modelos especiales de probabilidad para variables aleatorias continuas: *distribuciones uniforme, exponencial, normal y normal estándar.*

GPS ESTADÍSTICO 8

Orientaciones importantes para no perderse

- *Los modelos probabilísticos son una representación formal de un fenómeno real, por eso la experiencia en general se describe aproximadamente mediante dichos modelos teóricos.*
- *Para determinar qué modelo continuo describe el comportamiento de una variable que surge de un fenómeno en estudio, es importante reconocer las características que cada modelo expresa. De ahí que muchas veces en la práctica se asume que, para calcular probabilidades, una variable se “ajusta” a un modelo específico en función de su naturaleza u origen.*
- *Los parámetros de las distribuciones teóricas pueden determinarse de manera aproximada a través de las técnicas de inferencia que se trabajarán más adelante.*
- *Existen muchos más modelos continuos que los que se trabajarán en esta práctica y los que se presentan se consideran como representantes de “familias” de distribuciones.*

Glosario 8

Modelo uniforme: expresa, en el contexto de las variables continuas, el concepto de equiprobabilidad ya que, para intervalos de igual longitud sin importar su ubicación en el rango de definición de la variable, la probabilidad es la misma.

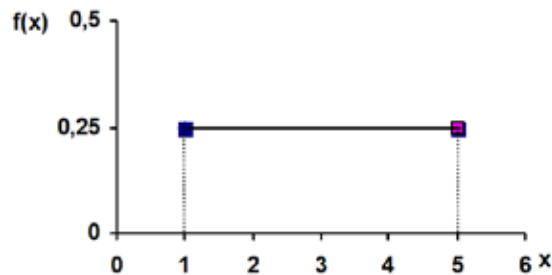
Modelo exponencial: representa una familia muy extensa de distribuciones que poseen un carácter asimétrico y se aplican especialmente para medir tiempos de vida útil de artículos eléctricos o electrónicos. Es un modelo que representa formalmente los tiempos de espera entre dos hechos consecutivos enmarcados en un proceso de Poisson.

Modelo normal: se aplica para distribuciones simétricas y dado su origen, se usa especialmente para modelar variables del tipo “pesos y medidas” sobre todo en procesos de crecimiento y desarrollo de seres vivos en el ámbito de la agricultura, ganadería, entre otros.

Problemas y situaciones prácticas

- **Distribución uniforme**

1- Sabiendo que el siguiente gráfico corresponde a una función de densidad de probabilidad de una variable aleatoria continua, donde x indica la concentración de un contaminante en una sustancia:



- Encontrar la expresión de la función de densidad de probabilidad $f(x)$.
- Calcular: $P(1,5 < X < 3)$ y $P(X > 2 / 1,5 < X < 3)$ e interpretar.
- Calcular $E(x)$ y Me e interpretar los resultados.

2- Si X es una variable aleatoria continua y la función que mejor describe su comportamiento es de la forma:

$$f(x) = \begin{cases} k & \text{para } 0 < x < 2 \\ 0 & \text{para otros valores de } x \end{cases}$$

- Determinar el valor de k para que $f(x)$ sea una función de densidad de probabilidad.
- Calcular $P(X < 1,5)$.
- Calcular la esperanza, desvío y mediana.

- 3- Los litros de café vendidos diariamente en un *fast food* tienen distribución uniforme variando entre 15 y 25 litros.
- Calcular la probabilidad de que en cierto día se vendan a lo sumo 22 litros.
 - Calcular la probabilidad de que se vendan por lo menos 17,5 litros en un día cualquiera.
 - ¿Cuántos litros se vendieron como máximo en el 45 % de los días?
- 4- El tiempo de un viaje (ida y vuelta) de los camiones que transportan arena hacia una obra en construcción en una carretera está distribuido uniformemente en un intervalo de 45 a 60 minutos. ¿Cuál es la probabilidad de que la duración de un viaje cualquiera sea mayor a 55 minutos si el destino del viaje es un lugar de la carretera que se encuentra a más de 50 minutos?
- 5- La cantidad de alimento balanceado consumida para alimentar animales en una granja se distribuye uniformemente entre 60 y 70 kilogramos semanales. ¿Cuál es la probabilidad de que en una semana elegida al azar se consuman entre 65 y 68 kilogramos?
- 6- El número 3,2 se pudo obtener por el redondeo de algún número comprendido entre 3,15 y 3,25. Cualquier número real del intervalo tiene la misma probabilidad que cualquier otro de ser el generador del número 3,2 por redondeo.
- Construir la función de densidad de la variable aleatoria continua con un recorrido de valores reales con dos decimales que por redondeo generen el número 3,2.
 - Calcular el valor esperado de la variable.
- 7- El contenido de sodio en la elaboración de un producto nuevo lanzado al mercado se distribuye uniformemente entre 0,2 % y 0,5 % de sodio en su composición.
- Calcular la probabilidad de que un producto elegido al azar contenga al menos 0,3 % de sodio.
 - Calcular las medidas posibles e interpretarlas.

- **Distribución exponencial**

- 8- La vida útil de un cierto componente de una computadora tiene una media de 1.000 horas.
- Calcular la probabilidad de que un componente de ese tipo dure menos de 1.000 horas, más de 1.200 horas y entre 800 y 1.100 horas.
 - Una computadora utiliza 3 componentes de ese tipo que funcionan independientemente. La máquina trabaja mientras funcionen los 3. ¿Cuál es la probabilidad de que la computadora opere por lo menos 600 h? y ¿1.200 h?
- 9- Un cierto tubo electrónico tiene una vida útil media de 500 h. ¿Cuántas horas de funcionamiento deben especificarse para que con 0,90 de probabilidad funcione por lo menos durante ese lapso?

- 10- El tiempo (en horas) que tarda un gerente en entrevistar a un aspirante para un trabajo tiene una distribución exponencial con parámetro 3. Los aspirantes están programados en intervalos de 1/4 de hora, empezando a las 8:00 a. m., y llegan exactamente a tiempo. Cuando el aspirante con una cita a las 8:15 a. m. llega a la oficina del gerente, ¿cuál es la probabilidad de que tenga que esperar para poder ver al gerente?
- 11- El tiempo de reposición de los cuatro focos de un automóvil sigue una distribución exponencial con media de 10 minutos.
- ¿Cuál es la probabilidad de que se tarde a lo sumo 5 minutos en cambiar cada foco?
 - ¿Cuál es el tiempo mínimo de reposición de los focos del 10 % de los automóviles?
- 12- Los tiempos obtención de frutos maduros se distribuyen exponencialmente desde la estación del año en que se produce su mayor producción, con un promedio de 10 unidades por día.
- Si se elige una planta al azar, ¿cuál es la probabilidad de que tarde a lo sumo 2 días hasta el próximo fruto maduro?
 - Calcular las medidas posibles e interpretarlas.

- **Distribución normal**

- 13- Se ha encontrado que el tiempo de servicio que se requiere por persona en una caja bancaria tiene distribución aproximadamente normal con media $\mu = 130$ segundos y desvío $\sigma = 43$ segundos.
- ¿Cuál es la probabilidad de que una persona elegida al azar requiera menos de 100 segundos para terminar sus transacciones?
 - ¿Cuál es la probabilidad de que una persona elegida al azar pase entre 2 y 3 minutos en la caja bancaria?
 - ¿Cuánto tiempo tardaron como máximo el 90 % de las personas?
 - ¿Qué medida está implicada en la pregunta anterior?
 - ¿Entre las personas que tardaron menos de 130 segundos, qué porcentaje tardó más de 100 segundos?
- 14- El gobierno está interesado en analizar la situación de los trabajadores del sector de comercio a fin de distribuir un plan de ayuda social. Sabe que el salario de dichos trabajadores se distribuye normalmente, con una media de 21 mil pesos y un desvío de 10 mil pesos.
- Calcular el porcentaje de trabajadores que tienen salarios comprendidos entre \$15 mil y \$25 mil.
 - Calcular el salario máximo del 75 % de los trabajadores.
- 15- La experiencia de una empresa al aplicar una prueba a universitarios recién egresados reveló que la puntuación media de la prueba fue de 500 con un desvío de 50. La distribución de los datos es aproximadamente normal.
- Basándose en el desempeño de los universitarios en la prueba, el director de personal plantea no tener en cuenta a las personas que tengan puntuaciones de 400 o menos en la prueba. ¿Qué porcentaje de postulantes no se tendrá en cuenta?

- b) Debido al número limitado de vacantes este año, los solicitantes con puntuaciones entre 400 y 485 se pondrán “en espera”. Si solicitaron trabajo en la empresa un total de 1.000 personas, ¿cuántos quedarán en la clasificación de “en espera”?
- c) La dirección está considerando aceptar a las personas cuya puntuación esté en el 6 % superior de la distribución y contratarlas directamente para un puesto de responsabilidad. ¿Cuál es la puntuación más baja que debe tener un egresado para calificar en un puesto de responsabilidad?
- d) Los que calificaron para un puesto de responsabilidad, ¿qué porcentaje representan del grupo de los que entraron a trabajar en la empresa?
- 16-** Se observó durante un largo período que la cantidad semanal gastada en el mantenimiento y en las reparaciones en cierta fábrica (en miles de pesos) tiene aproximadamente una distribución normal con una media de \$40 y una desviación estándar de \$2
- a) Si el presupuesto para la próxima semana es de \$45.000, ¿cuál es la probabilidad de que los costos reales sean mayores que la cantidad presupuestaria?
- b) ¿De cuánto tendría que ser el presupuesto en reparaciones semanales y mantenimiento, para que la cantidad presupuestada solamente se rebase con una probabilidad de 0,1?
- 17-** En una planta industrial, el consumo mensual de combustible es una variable aleatoria normal con media de 20.000 litros y desviación estándar de 2.500 litros. ¿Qué capacidad debe tener un tanque para satisfacer el consumo de un mes cualquiera con una probabilidad de 0,95?
- 18-** Un vendedor de autos usados encontró que el tiempo transcurrido, hasta que sea necesaria una gran reparación en los autos que vende, tiene una distribución normal con una media de 10 meses y un desvío de 3 meses. Para las próximas ventas, ¿de cuánto tendrá que ser el tiempo de garantía de los autos si para fijarla el vendedor tiene en cuenta el 5 % de los que más rápido requirieron una gran reparación?
- 19-** En una empresa los ingresos mensuales del personal (en miles de pesos) tienen una distribución normal. En promedio los trabajadores cobran \$55 mil con una desviación de \$15 mil.
- a) ¿Qué proporción de empleados cobra más de \$65 mil?
- b) En el departamento de compras el salario mínimo es de \$30 mil. ¿Cuál es la probabilidad de que un empleado tomado al azar entre los que pertenecen a dicho departamento cobre menos de \$50 mil?
- c) ¿Cuál es el salario superado por el 30 % de los empleados de la empresa? Identificar la medida utilizada.
- 20-** El consumo mensual de detergente en un restaurante tiene distribución normal con $\sigma = 6,5$ litros y el 70 % de los meses se consumen a lo sumo 48,6 litros.
- a) Identificando el tipo de información poblacional con que se cuenta, calcular el consumo medio mensual de detergente en el restaurante.
- b) Si para un mes cualquiera se compraran 50 litros de detergente, calcular la probabilidad de que en ese mes el consumo supere los 35 litros.

21- El consumo semanal de alimentos, en un comedor estudiantil, es una variable aleatoria que tiene distribución normal, con desvío igual a 25 kg. En el 80 % de las semanas el consumo fue superior a 150 kg.

- a) ¿Cuál fue el consumo medio semanal?
- b) ¿Cuánto se deberá tener en depósito en una semana cualquiera si se pretende que en ella no falte alimento, con una probabilidad de 0,90?
- c) Si se cree que la próxima semana podría ser una de las que el consumo es superior a la media, ¿cuál es la probabilidad de que se consuma menos de 250 kg?

22- El volumen diario de producción de cierto artículo tiene distribución normal. Se sabe que el 80 % de los días se producen menos de 160 toneladas de dicho artículo y que el 15 % de los días se producen entre 160 y 180 toneladas.

- a) Identificar la información poblacional con que se cuenta.
- b) Calcular el promedio diario de producción y el desvío estándar.
- c) Para un día cualquiera, ¿cuál es la probabilidad de que la producción supere a la mediana?
- d) ¿Entre qué valores estuvo la producción en el 50 % central de los días?

• Relación entre distribuciones de Poisson y exponencial

23- La llegada de barcos pesqueros a puerto Madryn, en determinada época del año, corresponde a un proceso de Poisson caracterizado por una intensidad media de dos arribos por hora. Luego del arribo de un barco al puerto:

- a) ¿Cuánto tiempo se espera que transcurra hasta la llegada del próximo barco?
- b) ¿Cuál es la probabilidad de que el siguiente ingreso se produzca antes de 1 hora?

24- En promedio, cada rollo de 500 metros de acero laminado tiene 2 defectos. Luego de la ocurrencia de un defecto:

- a) ¿Cuántos metros sin defectos se espera fabricar?
- b) ¿Cuál es la probabilidad de que se puedan fabricar 100 metros sin defecto?
- c) Con una probabilidad de 0,95, ¿cuántos metros podrían fabricarse sin defecto?

25- Una distribuidora mayorista comprobó que cada 5 días hábiles recibe en promedio 3 pedidos de embarque de cierto artículo.

- a) Teniendo en cuenta que el tiempo para reponer un embarque en depósito es de 1 día, despachado un pedido ¿con qué probabilidad el siguiente llegará después de ese lapso?
- b) ¿Con qué probabilidad el siguiente pedido será antes de lo esperado?
- c) Con una probabilidad de 0,90, ¿de cuánto tiempo se dispone entre dos pedidos?

• Mix de distribuciones 1

¹ Los problemas 26 a 31 corresponden a convergencia en distribución y suma de variables que, si bien pueden resolverse con las aplicaciones digitales sin inconvenientes, se retomarán y profundizarán en Estadística Económica Especializada para quienes sigan la carrera de Economía.

- 26-** Sea X una variable obtenida de un experimento binomial con:
- Si $n = 30$ y $p = 0,49$
 - Usando la distribución binomial, calcular: $P(12 < x < 15)$; $P(12 < x \leq 15)$
 - Obtener los valores de $\mu = E(x)$ y $\sigma^2 = V(x)$ y utilizando la distribución normal y calcular las siguientes probabilidades: $P(12 < x < 15)$; $P(12 < x \leq 15)$
 - Comparar los resultados anteriores.
 - Si $n = 300$ y $p = 0,49$, hallar $\mu = E(x)$ y $\sigma^2 = V(x)$ y calcular $P(137 < x < 150)$
- 27-** Una compañía está pensando adoptar un sistema más flexible de horario para sus empleados. La idea es que cada empleado adopte el horario que más le convenga, siempre y cuando trabaje 40 horas semanales. Los resultados de un estudio piloto mostraron que el 47 % de los empleados no están a favor del nuevo sistema. Para verificar dicho resultado se toma una muestra de 200 empleados. Usando la aproximación normal de la binomial calcular la probabilidad de que:
- No más de 80 de los 200 empleados se opongan.
 - Al menos 70 empleados se opongan al nuevo sistema de horarios sabiendo que se opusieron menos de 100
- 28-** Suponga que 40.000 de los 105.000 miembros de un sindicato están a favor de la huelga, si 100 miembros distintos son seleccionados al azar, calcular la probabilidad de que al menos 60 miembros de la muestra no estén a favor de la huelga.
- 29-** La experiencia indica que solo el 42 % de los empleados de una empresa recibe bonificación por presentismo. Suponiendo que se seleccionen 450 empleados verificar si reciben la bonificación por presentismo, encontrar la probabilidad de que:
- 160 o más la reciban.
 - Menos de 280 no reciban dicha bonificación.
- 30-** Una empresa editorial debe separar el 34 % de los libros que imprime para vender como segunda selección. Se le ha encargado a dicha editorial un tiraje de 1.450 ejemplares:
- ¿Cuántos ejemplares se espera que sean separados?
 - ¿Cuál es la probabilidad de que se separen menos de 300 ejemplares?
- 31-** Según lo informado por una empresa consultora de estadística la participación en el mercado de las heladeras FRIOX es del 55 %. Se estima que en el próximo trimestre la demanda global de heladeras puede ser de 1.850
- ¿Cuál es la probabilidad de que se vendan a lo sumo 1.000 heladeras FRIOX?
 - ¿Cuál es la cantidad de ventas de heladeras FRIOX que será superada con probabilidad de 0,97?
- 32-** El tiempo que dura una batería en funcionamiento tiene distribución exponencial de parámetro 0,005
- Calcular el tiempo medio de funcionamiento de la batería.

- b) Calcular la probabilidad de que la batería dure más de 160 h.
- c) Calcular la probabilidad de que el tiempo de vida de la batería sea de a lo sumo 220 h.
- d) ¿Cuántas horas podrá funcionar la batería por lo menos con probabilidad 0,3?

- 33- En una aerolínea, el tiempo para atender a los pasajeros sin billete en el mostrador de aeropuerto sigue una distribución exponencial con una media de 5 minutos.
- a) ¿Cuál es la probabilidad de un tiempo de atención menor a 2,5 minutos?
 - b) ¿Cuál es el tiempo máximo que deben esperar los pasajeros, con una probabilidad de 0,90?

- 34- Los siguientes datos correspondientes a 16 jóvenes preseleccionados para una competencia nacional de triatlón correspondientes al distrito escolar de La Matanza.

Altura (en cm)	Género		Totales
	F	M	
150-153	1	0	1
153-156	0	1	1
156-159	4	3	7
159-162	1	5	6
162-165	1	0	1
Totales	7	9	16

- a) Si se seleccionan 5 varones al azar, ¿cuál es la probabilidad que las alturas de todos ellos sean inferiores a 156 cm?
- b) Si se seleccionan aleatoriamente en toda la región metropolitana 100 jóvenes de por lo menos 159 cm de altura, ¿cuál es la probabilidad de que al menos 10 de ellos sean mujeres? Suponer que las proporciones por categoría se mantienen para poder aplicarlas en el conjunto de mayor tamaño.

- 35- Una encuesta aplicada a ejecutivos se enfocó sobre la permanencia en la empresa donde trabajan. Una de las preguntas planteadas fue: “Si otra empresa le hiciera una oferta igual o mejor a la de su puesto actual, ¿aceptaría la propuesta?”. Las respuestas de los 200 ejecutivos a los que se les tomó la encuesta se clasificaron en forma cruzada con su tiempo de servicio en la empresa:

	Tiempo de servicio			
	Menos de 1 año	De 1 a 5 años	De 5 a 10 años	Más de 10 años
Se quedaría	8	30	10	72
No se quedaría	20	20	10	30

- a) Si se seleccionan 20 personas de la empresa al azar ¿cuál es la probabilidad que todas se queden en la empresa al menos 5 años?
- b) Si la seleccionan aleatoriamente 120 personas de la empresa que llevan más de 1 año de servicio, ¿cuál es la probabilidad de que al menos 60 se queden en la empresa?

- 36- Sea X una variable aleatoria con distribución uniforme sobre el intervalo $[a ; b]$. Si $E(X) = 10$ y $V(X) = 12$, encontrar los valores de a y de b .

- 37- Sea X una variable aleatoria con distribución uniforme sobre el intervalo $[3; 5]$.
- a) Calcular la probabilidad de que X tome un valor tal que difiera de la media en, a lo sumo, un desvío estándar.

b) ¿Puede tomar X un valor que se encuentre a 2 desvíos estándar de la media?

38- En una tela las fallas se distribuyen según un proceso de Poisson, a razón de 1 falla cada 15 metros. ¿Cuál es la probabilidad de que la distancia entre la 4.^a falla y la 5.^a falla sea mayor a un metro?

39- A una central de teléfonos llegan 12 llamadas por minuto, siguiendo una distribución de Poisson. ¿Cuál es la probabilidad de que en menos de 1 minuto se reciba la siguiente llamada?

40- El número de clientes, en promedio, que llegan por minuto a solicitar servicio a un banco es de 5 siguiendo un proceso de Poisson. ¿Cuál es la probabilidad de que dos clientes tarden de 30 a 45 segundos en llegar el próximo cliente al banco?

RESPUESTAS DE LAS ACTIVIDADES DE LA PRÁCTICA 8

Distribuciones especiales continuas de probabilidades

Distribución uniforme

1- a) $f(x) = \begin{cases} 0,25 & x \in [1; 5] \\ 0 & x \notin [1; 5] \end{cases}$

b) 0,375 / 0,66667

c) 3 / 3

2- a) 0,5 b) 0,75

c) 1 / 1 / 0,577

3- a) 0,7

b) 0,75

c) 19,5

4- 0,5

5- 0,3

6- a) $f(x) = \begin{cases} 10 \dots x \in [3,15; 3,25] \\ 0 \dots x \notin [3,15; 3,25] \end{cases}$

b) 3,2

7- a) 0,66667

b) 0,35 / 0,35 / 0,087 / 0,0075

Distribución exponencial

8- a) 0,63212 / 0,30119 / 0,11646

b) 0,1653 / 0,02732

9- 52,68

10- 0,47237

11- a) 0,39347

b) 23,02

12- a) 0,9999

b) 0,1 / 0,1 / 0,01

Distribución normal

13- a) 0,24196

b) 0,46793

c) 185,04

d) Percentil 90

PRÁCTICA 9

Distribución de estadísticos muestrales

Del estudio del comportamiento de los estadísticos muestrales surgen propiedades que los relacionan intrínsecamente con los respectivos parámetros.

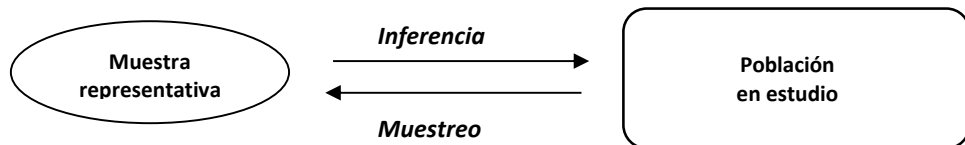
Las curvas de distribución resultantes de ese estudio son claves para fundamentar las soluciones a los problemas que plantea la estadística inferencial.

Información muestral como variables aleatorias. Distribución de estadísticos muestrales.

GPS ESTADÍSTICO 9

Orientaciones importantes para no perderse

- *La inferencia estadística es el procedimiento por el cual se extrapolan o extienden (no generalizan) los resultados de una muestra representativa a la población en estudio.*



- *Para hacer inferencia es preciso conocer los estadísticos muestrales y sus distribuciones.*
- *Bajo ciertas condiciones \bar{X} y p , media y proporción muestrales respectivamente, siguen una distribución aproximadamente normal para muestras suficientemente grandes (corolario del teorema central del límite).*
- *Los parámetros de las distribuciones de probabilidad de \bar{X} son $E(\bar{X}) = \mu$ y $V(\bar{X}) = \frac{\sigma^2}{n}$ y para p son $E(p) = P$ y $V(p) = \frac{P(1-P)}{n}$ donde μ y σ^2 son la esperanza y varianza poblacional de la variable en estudio y P es la proporción poblacional de éxitos estudiada.*

Glosario 9

Estadístico: medida estadística obtenida de datos muestrales.

Parámetro: constante obtenida de datos poblacionales.

Error estándar: desvío estándar de la distribución de un estadístico alrededor del parámetro correspondiente.

Problemas y situaciones prácticas 9

- **Distribución de la media muestral**

- 1- Una empresa lleva a cabo una prueba para seleccionar nuevos empleados. Por la experiencia de pruebas anteriores, se sabe que las puntuaciones siguen una distribución normal de media 80 y desviación típica 25
 - a) ¿Qué porcentaje de candidatos obtendrá entre 75 y 100 puntos?
 - b) ¿Cuál será el puntaje máximo del 75 % de los aspirantes?
 - c) En una muestra aleatoria de 25 empleados, ¿cuál es la probabilidad de que el puntaje medio sea al menos de 75 puntos?
- 2- El valor promedio depositado por cuenta en una entidad bancaria es de \$30.000 con un desvío de \$4.000. Si se planifica realizar una muestra aleatoria de 36 cuentas, ¿cuál es la probabilidad de que la media de esa muestra:
 - a) sea inferior a \$29.500?,
 - b) se encuentre entre \$29.000 y \$30.200 y
 - c) ¿cuántas cuentas se deberán seleccionar para que el promedio de los depósitos de ellas resulte superior a \$29.500 con una probabilidad de 0,95?
- 3- El gobierno está interesado en analizar la situación de los trabajadores del sector comercio a fin de distribuir un plan de ayuda social. Se sabe que el salario de dichos trabajadores se distribuye normalmente, con una media de \$36.000 y un desvío de \$3.000
 - a) Calcular el porcentaje de trabajadores que tienen salarios de entre \$35.500 y \$36.450
 - b) Si se selecciona una muestra de 120 trabajadores, ¿cuál es la probabilidad de que el salario medio de la muestra esté comprendido entre \$35.500 y \$36.450?
- 4- Se sabe que las ventas diarias realizadas por una empresa están distribuidas normalmente con una media de \$22.500 y un desvío de \$3.069,41. De las 1.815 ventas realizadas en el mes, se extrae una muestra de 16 facturas.
 - a) ¿Cuál es la probabilidad de que la media de la muestra sea superior a \$24.000?
 - b) ¿Cuál es la probabilidad de que la media de la muestra difiera de la media poblacional en más de \$1.300?
- 5- El contenido de las latas de arvejas de una determinada marca tiene distribución normal con peso medio de 201 g y desvío de 8 g. Un cliente desea comprar una partida de cajas, pero la decisión la tomará en base a una muestra de 36 latas. Si el peso medio de la muestra fuese superior a 200 g comprará la partida.
 - a) ¿Cuál es la probabilidad de que ello ocurra?
 - b) ¿Cuál debería ser el tamaño mínimo de la muestra para que el cliente compre la partida con una probabilidad de 0,98?

- 6- Un hipermercado de la zona de San Justo ha implementado un sistema de ventas por medio de una tarjeta de compras. Los usuarios no siempre abonan el resumen mensual en fecha. El gerente de ventas ha estudiado la situación encontrando que el promedio de los días de atraso en el pago mensual de la tarjeta de compra es de 60 días con un desvío de 14 días; selecciona al azar 70 fichas para averiguar la probabilidad de que la media de los días de atraso no supere los 55 días. ¿Cuál es dicha probabilidad?
- 7- Un kiosco de revistas de un barrio de la Ciudad de Bs. As. tiene 120 clientes, cada uno en promedio compra mensualmente por valor de \$650 con un desvío de \$56,70. ¿Cuál es la probabilidad de que una muestra de 40 clientes arroje un promedio de compra mensual comprendido entre \$630 y \$670?

- **Distribución de la proporción muestral**

- 8- Una inmobiliaria de Morón ha cambiado su política de alquileres introduciendo una opción de alquiler a sola firma, sin garante. Actualmente el 55 % de los departamentos que ofrece son con esta modalidad. Del archivo de alquileres de la inmobiliaria se tomó una muestra aleatoria de 180 fichas. ¿Cuál es la probabilidad de que la proporción de inquilinos sin garante esté comprendida entre el 48 % y el 60 %?
- 9- En la población de los aspirantes a ingresar en la carrera de Educación Física de la UNLaM, formada por 745 alumnos, hay 327 que han practicado un deporte en forma regular. Se saca una muestra de 90 alumnos. ¿Cuál es la probabilidad de que la proporción de alumnos que han practicado un deporte en la muestra sea inferior a 0,35?
- 10- En una entidad financiera el 48 % de los cheques presentados al cobro se abonan de inmediato y sin verificación de saldo según sea su monto. Si se toman aleatoriamente 250 cheques, ¿cuál es la probabilidad de que como máximo el 40 % de los cheques de la muestra no necesiten verificación de saldo?
- 11- Responder los siguientes interrogantes:
- a) ¿Cuál es la diferencia entre censo y muestreo?, ¿y entre población y muestra?
 - b) ¿Qué significa que una muestra sea representativa?
 - c) Los estudiantes de la carrera Lic. en Administración, ¿pueden constituir una muestra representativa de la población de la UNLaM? ¿Por qué?
 - d) ¿Para seleccionar una muestra representativa de los habitantes de una determinada ciudad, se puede utilizar la guía telefónica? ¿Por qué?
 - e) ¿Qué significa que una muestra sea aleatoria?

- **Para resolver**

- 12- Un análisis de la duración de las llamadas telefónicas locales hechas desde la oficina de una empresa muestra que el tiempo de duración de las llamadas es una variable aleatoria que

tiene distribución aproximadamente normal con una media de 1,7 minutos y un desvío estándar de 0,2 minutos.

- a) ¿Qué porcentaje de estas llamadas dura menos de 2 minutos?
- b) Si se registran, se miden las llamadas, y se toma una muestra aleatoria de 49 de ellas, ¿cuál es la probabilidad de que el promedio de estas dure más de 2 minutos?

13- El 34 % del material que recibe una empresa de reciclaje es papel. Si se toma una muestra de 1.500 kg de material a reciclar:

- a) ¿Cuántos kg se espera que sean de papel?
- b) ¿Cuál es la probabilidad de que menos de 300 kg de la muestra sea papel? Sugerencia, ¿se tiene la proporción muestral p ?

RESPUESTAS DE LAS ACTIVIDADES DE LA PRÁCTICA 9

Distribución de estimadores

Distribución de la media muestral

- 1-
 - a) El 36,74 % de los candidatos obtendrá entre 75 y 100 puntos.
 - b) El puntaje máximo del 75 % de los aspirantes es 97 puntos.
 - c) 0,84134
- 2-
 - a) 0,22663
 - b) 0,5511
 - c) Se deberán seleccionar aproximadamente 175 cuentas para que el promedio de los depósitos de ellas resulte superior a \$29.500 con una probabilidad de 0,95
- 3-
 - a) El 12,711 % de los trabajadores tienen salarios entre \$35.500 y \$36.450
 - b) 0,91588
- 4-
 - a) 0,02559
 - b) 0,09102
- 5-
 - a) 0,77337
 - b) El tamaño mínimo de la muestra para que el cliente compre la partida con una probabilidad de 0,98 es 268.96 (aproximadamente 269).
- 6- 0,00139
- 7- 0,99988

Distribución de la proporción muestral

8- 0,8805

9- Con $p=327/745=0,44$, la probabilidad de que la proporción de alumnos que han practicado un deporte en la muestra sea inferior a 0,35 es 0,04272 (4,272 %).

10- La probabilidad de que como máximo el 40 % de los cheques de la muestra no necesiten verificación de saldo es 0,00570 (0,57 %).

11- a) Se realiza un censo cuando se examinan todos los elementos de la población que se investiga mientras que en un muestreo se seleccionan unos cuantos elementos (una muestra) de un grupo de datos (una población) para ser examinados.

b) Una muestra es representativa cuando siendo un subconjunto de la población refleja las características de esta (población).

c) No, porque no refleja todas las características de los alumnos de la UNLaM.

d) No, porque no contendría las características de los habitantes que no poseen teléfono.

e) Que una muestra sea aleatoria significa que ha sido obtenida por medio de un muestreo aleatorio simple, es decir, que hace uso de la simulación de un sorteo mediante números aleatorios, como el sorteo de la lotería nacional.

Para resolver

12- a) El 99,319 % de las llamadas dura menos de 2 minutos.

b) La probabilidad de que el promedio de duración de las llamadas sea de más de 2 minutos es aprox. 0,00002

13- a) Se espera que 510 kg sean de papel.

b) Aproximadamente 0

PRÁCTICA 10

Estimación de parámetros - intervalos de confianza

Los resultados o conclusiones que se obtienen al hacer inferencias nunca son ciertos, siempre están provistos de algún grado de confianza o de algún margen de error que se obtiene a través de la probabilidad y de la distribución de probabilidad de los estimadores. Por esto la teoría de la probabilidad es la que da fundamento y sustento a la inferencia estadística.

La inferencia estadística aborda básicamente dos tipos de problemas: la estimación de parámetros y la prueba de hipótesis.

GPS ESTADÍSTICO 10

Orientaciones importantes para no perderse

- *La estimación por intervalos de confianza permite, a través de los resultados de una muestra, obtener y con una cierta probabilidad asociada, establecer entre qué valores podría encontrarse un parámetro poblacional.*
- *En general, puede verificarse que en los procesos de fabricación en los que intervienen variables de medición como ser: peso, longitud, capacidad, etc., la distribución de estas variables es aproximadamente normal.*
- *La distribución Chi-cuadrado surgió a partir de los problemas de muestreo en poblaciones normales. La distribución T fue diseñada para problemas de muestreo donde la varianza muestral siguiera una distribución Chi-cuadrado. Por tal motivo, la distribución T tanto como la Chi solo pueden aplicarse al muestreo en poblaciones normales.*

Glosario 10

Nivel de confianza, $1 - \alpha$: es la probabilidad que tiene un intervalo de confianza, construido a partir de una muestra, de contener al parámetro poblacional estimado.

Error máximo: distancia máxima entre el estadístico muestral y el parámetro poblacional a estimar, con una probabilidad asociada.

Problemas y situaciones prácticas 10

- 1- A fin de controlar el proceso de llenado de un producto alimenticio, se controló durante 16 días al azar el desperdicio en gramos resultante del proceso, obteniéndose los siguientes resultados:

512	510	495	508	504	505	496	512
514	493	496	506	502	509	497	505

Desde la experiencia en esta labor se ha establecido el desvío de la variable estudiada en 5 g.

- ¿Podría asumirse que el desperdicio en gramos de todo el proceso se distribuye normalmente? Discutirlo.
 - Obtener un intervalo de confianza del 95 % para la media general del desperdicio diario.
 - Obtener un intervalo de confianza del 99 % para el desperdicio medio diario del proceso.
 - Decidir cuál de las dos estimaciones anteriores es más precisa y justificar.
 - ¿De qué tamaño debería ser una nueva muestra para que el error máximo asociado a una nueva estimación sea dos tercios del obtenido en b) utilizando el mismo nivel de confianza?
 - Si no se conociera la dispersión estándar, calcular el intervalo de confianza del 95 % para la media del desperdicio diario.
 - Comparar f) y b) y analizar cuál de las dos estimaciones sería más ventajosa.
- 2- La tabla muestra el peso para los contenidos de diez unidades de miel de abeja envasada, tomadas aleatoriamente de una partida grande. Si el proceso de envasado de miel se distribuye con un desvío de 4,30 gramos según lo especifica el manual de la máquina:

Envase	1	2	3	4	5	6	7	8	9	10
Peso (g)	493	501	495	496	492	505	494	496	503	502

- Estimar el peso medio de toda la partida de envases de miel con una confianza de 0,95
 - Indicar qué información poblacional se tuvo en cuenta para realizar la estimación de a).
 - Para mejorar la estimación, ¿qué es lo más conveniente?
 - ¿Cómo procedería para resolver c)? Realizar el procedimiento decidido con un nuevo valor posible.
- 3- Si se tienen las siguientes expresiones de los límites de un intervalo de confianza, obtener el nivel de confianza asociado al intervalo (recordar que el desvío de \bar{X} es $\sigma_{\bar{X}} = \sigma/\sqrt{n}$).
- a) $\bar{X} \pm 1,6 \sigma_{\bar{X}}$ b) $\bar{X} \pm 1,8 \sigma_{\bar{X}}$ c) $\bar{X} \pm 2,4 \sigma_{\bar{X}}$
- 4- a) ¿Con qué nivel de confianza se obtuvo el intervalo [95,90; 104,10] para la media μ si se sabe que el desvío poblacional es 10 y el tamaño de la muestra 16?
- b) ¿Qué suposición sobre la variable estudiada debió tenerse en cuenta para la obtención del intervalo de a)?

- 5- a) ¿Con qué confianza el intervalo $[268,3 ; 294,3]$ contiene al parámetro media poblacional si fue calculado de una muestra de 16 elementos donde la varianza de la muestra s^2 es de 300?
 b) ¿Qué suposición sobre la variable estudiada se tuvo en cuenta para la obtención del intervalo de a)?

- 6- En la tabla se indica el número de clientes que entraron, en un día determinado, a 5 farmacias elegidas aleatoriamente en una ciudad.

Farmacia	1	2	3	4	5
N.º de clientes	580	430	612	498	591

Si la cantidad diaria de clientes que concurren a las farmacias tiene distribución normal con un desvío histórico de 120,3 clientes:

- a) Calcular el error máximo de la estimación para un intervalo del promedio diario de clientes que concurren a todas las farmacias de la ciudad, con una confianza de 0,99
 b) Calcular el error máximo de la estimación, pero desconociendo el desvío de la población con una confianza del 99 %
 c) Comparar a) y b) y explicar las diferencias.
- 7- Indicar, para cada uno de los siguientes enunciados, si es verdadero o falso. Si es falso, dar un contraejemplo e indicar la respuesta correcta:
 a) Un parámetro es una medida que se calcula teniendo en cuenta todos los elementos que componen una muestra seleccionada aleatoriamente.
 b) La variable \bar{X} tiene siempre distribución normal.
 c) Para reducir la amplitud de un intervalo de confianza, hacerlo más preciso, es conveniente reducir el nivel de confianza.
 d) La manera más efectiva de mejorar una estimación por intervalos de confianza de, por ejemplo, la media poblacional es aumentar el tamaño de la muestra.
- 8- Una gran tienda desea estimar con un nivel de confianza del 90 % y un error máximo de \$250 la verdadera media del valor de las compras mensuales de sus clientes de cuentas corrientes. ¿Cuál debe ser el tamaño de la muestra que el negocio tome de sus registros para satisfacer las especificaciones, si se sabe que la desviación estándar es de \$1.036?
- 9- Se planifica un estudio para determinar el tiempo promedio que los niños de nivel preescolar ven televisión. Un estudio piloto previo indicó que el tiempo medio por semana fue de 20 h, con un desvío de 4 h. Se desea estimar el tiempo medio con un error máximo de un cuarto de hora, y con un nivel de confianza del 98 %. ¿Cuántos niños de nivel preescolar deben incluirse en la investigación?
- 10- Se desea realizar un muestreo para determinar el ingreso familiar medio anual en una determinada región. Con la finalidad de tener alguna información sobre la población se tomó de un estudio previo la desviación estándar del ingreso familiar anual que resultó de \$6.035.

¿Cuántas familias deberán ser entrevistadas como mínimo para que, con base en la información del estudio previo y con una confianza del 95 %, pueda resultar que la media muestral no difiera de la media poblacional en más de \$1.200?

- 11- Una muestra de 60 cajas de papel oficina seleccionadas aleatoriamente de un gran envío de cajas de dicho papel tiene un peso medio de 700 kg y una desviación típica de 12 kg.
- Si se usa la media muestral para estimar la media de todo el envío, ¿qué se puede decir con respecto al posible error $\bar{X} - \mu$, con una probabilidad de 0,95? Tener en cuenta que $|\bar{X} - \mu| < a$
 - Si el error se redujera a la mitad sin modificar el nivel de confianza; ¿cuál sería un tamaño de muestra adecuado?
- 12- A un conjunto de 100 administradores de publicaciones en redes sociales se les preguntó su opinión sobre la influencia de la publicidad que aparece allí sobre la decisión de los usuarios de adquirir los productos promocionados. En total 54 de las personas consultadas afirmaron en base a su experiencia, que el público en general resulta inducido a comprar el producto luego de ver su publicidad. Construir e interpretar un intervalo del 98 % para la proporción de usuarios que serían inducidos a comprar productos promocionados en las redes sociales.
- 13- En una encuesta de opinión un candidato obtiene 216 votos sobre 400 encuestados.
- Hallar un intervalo de confianza del 95 % para la verdadera proporción de votantes a favor del candidato.
 - Según el resultado obtenido en a), ¿podría decirse que hay un 95 % de confianza de que el candidato podría ganar por mayoría?
 - ¿De qué tamaño debería ser como mínimo la muestra para que el intervalo resultante garantice con probabilidad del 95 % que el candidato sea elegido, suponiendo mismo valor de la p muestral inicial? Para que el candidato sea elegido por mayoría debe superar el 50 % de los votos.
- 14- De un lote de 200 lámparas led se extrajo una muestra de 36 unidades para estimar la vida útil media en horas de uso continuo. De la muestra se obtuvo una duración media de 850 horas y un desvío de 25 horas.
- Estimar la duración media de todas las lámparas con el 95 % de confianza.
 - Con el 90 % de confianza estimar la desviación estándar del lote.
- 15- Los siguientes datos corresponden a una muestra de 410 facturas de ventas seleccionadas al azar en un mes en un negocio de repuestos para máquinas agrícolas.

Importe (miles de \$)	8-8,5	8,5-9	9-9,5	9,5-10	10-10,5	10,5-11	11-11,5	11,5-12
N.º de ventas	25	43	81	97	78	42	27	17

- Estimar el monto medio de todas las ventas del mes con una confianza de 0,95
- Estimar la proporción de todas las ventas del mes con un monto inferior a \$10.500 con una confianza del 90 %

- c) ¿Cómo se debe resolver el punto anterior si fuera para montos inferiores a \$11.200? ¿Y superiores a \$9.800?
- 16-** La sección de pruebas de una fábrica de autos decide estudiar el consumo de nafta de un nuevo modelo. Para ello realiza la experiencia 60 veces para un mismo recorrido, y obtiene una media de 10,5 litros con un desvío de 1,6 litros.
- Estimar, para ese modelo de autos, el consumo medio de nafta y el desvío con una confianza del 90 %
 - Si la estimación de la media se realiza con un nivel de confianza mayor al 90 %, ¿se pueden obtener mejores resultados?
 - Si para estimar la media se requiere reducir el error máximo a 0,10; con un mismo nivel de confianza, ¿cuál sería el tamaño de la muestra necesario?
- 17-** Para mejorar la estimación por intervalos de confianza de un parámetro poblacional hay que (decidir en cada opción si es verdadera o falsa la afirmación y justificar):
- aumentar el nivel de confianza
 - disminuir los grados de libertad
 - agrandar el tamaño de la muestra
 - disminuir el nivel de confianza
 - incorporar información adicional sobre la población
- 18-** La permanencia de un grupo de rock en el primer lugar de los “Diez más votados”, de una FM de la Ciudad de Buenos Aires, depende de la cantidad de llamadas telefónicas realizadas por los fans a dicho programa. La FM en cuestión solicitó a una empresa dedicada a la medición de *rating* que realice un estudio sobre la cantidad de votos que capta Miranda; la respuesta de la empresa fue que al menos el 80 % de los radioescuchas del mencionado programa, que suelen realizar votaciones telefónicas, le otorgan su voto al conjunto. Para verificar dicha afirmación se tomó una muestra de 550 votantes de los cuales 425 votaron por Miranda. Estimar por intervalo, con un nivel de confianza de 0,98, la verdadera proporción de radioescuchas que votaron por Miranda.
- 19-** Los siguientes datos corresponden a los pagos en concepto de impuestos a las ganancias, los cuales se distribuyen normalmente, que fueron realizados por 6 ejecutivos en el último período fiscal: 120.000 – 125.000 – 122.000 – 115.000 – 118.000 – 116.000
- Estime con una confianza del 95 % el monto medio de los pagos en concepto de impuesto a las ganancias por este tipo de contribuyentes.
 - Si la estimación de la media se realiza con un nivel de confianza mayor al 95 %, ¿qué conclusión puede obtener? Justifique su respuesta.
 - Estimar el desvío estándar con una confianza del 99 %
- 20-** Las producciones diarias de un producto químico para una semana cualquiera fueron 950, 1.010, 1.025, 936 y 1.009 toneladas.
- Estimar, para todos los días de producción, las toneladas promedio diaria del producto químico, con un nivel de confianza del 98 %
 - Estimar la varianza de la producción del producto químico por día, con un nivel de confianza del 90 %

c) ¿Qué supuestos se requieren para que sea válido el procedimiento aplicado para realizar las estimaciones a) y b)?

21- Se seleccionó una muestra aleatoria de 7 departamentos de un edificio consultando a sus dueños sobre el monto total bimestral gastado en impuestos, y se obtuvo la siguiente información (en pesos): 4.000 – 3.500 – 3.300 – 4.200 – 3.250 – 2.500 – 2.600

Si se sabe que la distribución de la cantidad gastada en impuestos en todos los departamentos del edificio es normal,

a) estimar la cantidad promedio y el desvío estándar de dinero que se gasta en todos los departamentos de ese edificio, con un nivel de confianza de 98 %

b) Si la estimación de la media se realiza con un nivel de confianza menor al 98 %, ¿qué conclusión puede obtener? Justifique su respuesta.

22- El matrimonio Mascetti posee naranjales en la zona de San Pedro. Al señor Mascetti le preocupa que en los últimos cinco años las heladas hayan dañado los 2.500 naranjos de su propiedad. A fin de averiguar el daño causado a los árboles, ha muestreado el número de naranjas producidas en 51 de ellos y ha observado que la producción promedio fue de 580 naranjas por árbol, con un desvío estándar de 40 naranjas por árbol.

a) Estimar por intervalo de confianza del 95 % el número medio de naranjas por árbol en todo el naranjal.

b) Estimar el desvío estándar poblacional con un nivel de confianza del 85 %

23- Los siguientes datos corresponden a una muestra de 410 facturas de ventas seleccionadas al azar en un mes en un negocio de repuestos del automotor.

Importe en \$	500-550	550-600	600-650	650-700	700-750	750-800	800-850	850-900
Cant. de ventas	25	43	81	97	78	42	27	17

a) Estimar el monto medio de todas las ventas del mes con una confianza de 0,95

b) Estimar la proporción de todas las ventas del mes con un monto inferior \$690 con una confianza del 90 %

Mix de estimación por intervalos

24- En base a la muestra de alumnos pasantes del Departamento de Ciencias Económicas resumidos en la Matriz 1 del Anexo.

a) Estimar por intervalo de confianza del 95 % la media del promedio de notas de las materias aprobadas de todos los pasantes del Departamento.

b) Estimar por intervalo de confianza del 90 % el desvío del promedio de notas de las materias aprobadas de todos los pasantes del Departamento.

c) Hallar los intervalos de confianza del 95 % y 99 % para la proporción de todos los alumnos pasantes con promedio superior a 8 puntos.

d) Comparar y discutir los resultados obtenidos en ambas estimaciones de c).

e) Estimar con una confianza del 85 % la proporción de alumnos pasantes que tienen un nivel de supervisión de a lo sumo moderado.

- 25- En la Matriz 2 se presentan los montos de los salarios de 60 empleados de una empresa extranjera de turismo, seleccionados al azar.
- Estimar mediante un intervalo del 90 % de confianza, el salario medio de los empleados de la empresa.
 - Utilizando un nivel de confianza del 85 %, estimar la proporción de empleados con salario superior a USD 2.500
 - Estimar mediante un intervalo del 80 % de confianza, la varianza del monto de los salarios de los empleados de la empresa.
 - Explicar por qué es necesario que los montos de los salarios de los empleados se distribuyan normalmente para a) y c).
 - Discutir y fundamentar si en todos los intervalos calculados anteriormente puede aplicarse el concepto de error máximo.
- 26- Mediante un intervalo de confianza del 95 % se estimar la varianza de X= peso escurrido de arvejas en latas para exportación. Primero, se toma una muestra aleatoria de tamaño $n_1 = 31$ de la cual resulta una varianza muestral de 9 g^2 . Con estos datos se obtiene el intervalo buscado (A, B). En una segunda instancia, y con otro nivel de confianza, se vuelve a tomar una muestra aleatoria con $n_2 = 28$, de la cual resulta una varianza muestral de 10 g^2 y el intervalo de confianza obtenido es (A, C), o sea, que el extremo inferior A del primer intervalo es igual al del segundo. ¿Cuál es el nivel de confianza del segundo intervalo?

RESPUESTAS DE LAS ACTIVIDADES DE LA PRÁCTICA 10

Estimación de parámetros-intervalos de confianza

- 1- a) No necesariamente.
- b) $I_{\mu} = (504 - 1,96 \cdot \frac{5}{\sqrt{16}}; 504 + 1,96 \cdot \frac{5}{\sqrt{16}}) = (504 - 2,45; 504 + 2,45) = (501,55; 506,45)$
- c) $I_{\mu} = (504 - 2,58 \cdot \frac{5}{\sqrt{16}}; 504 + 2,58 \cdot \frac{5}{\sqrt{16}}) = (504 - 3,225; 504 + 3,225) = (500,775; 507,225)$
- d) La estimación de la parte b es más precisa que la de la parte c.
- e) $n=36$
- f) $I_{\mu} = (504 - 2,13 \cdot \frac{6,8}{\sqrt{16}}; 504 + 2,13 \cdot \frac{6,8}{\sqrt{16}}) = (504 - 3,62; 504 + 3,62) = (500,38; 507,62)$
- g) En base a una muestra, con un valor de desvío poblacional, y con un nivel de confianza del 95 % la estimación hallada en b) es más precisa que la f), aunque es necesario contar con el desvío poblacional, que no siempre es dado.
- 2- a) $I_{\mu} = (497,7 - 1,96 \cdot \frac{4,30}{\sqrt{10}}; 497,7 + 1,96 \cdot \frac{4,30}{\sqrt{10}}) = (497,7 - 2,66; 497,7 + 2,66) = (495,04; 500,36)$
- b) Desvío poblacional.
- c) Aumentar el tamaño de la muestra.
- d) con $1-\alpha=0,99$ $I_{\mu} = (497,7 - 2,58 \cdot \frac{4,30}{\sqrt{10}}; 497,7 + 2,58 \cdot \frac{4,30}{\sqrt{10}}) = (497,7 - 3,51; 497,7 + 3,51) =$

$$= (494,19;501,21)$$

- 3- a) $1-\alpha=0,8904$
 b) $1-\alpha=0,92814$
 c) $1-\alpha=0,9836$
- 4- a) $1-\alpha=0,899$
 b) Distribución aproximadamente normal.
- 5- a) $1-\alpha=0,99$
 b) Distribución aproximadamente normal.
- 6- a) $\varepsilon=2,58 \cdot \frac{120,3}{\sqrt{5}}=138,8$
 b) $\varepsilon=4,6 \cdot \frac{76,2}{\sqrt{5}}=156,75$
 c) En a) se conoce el desvío poblacional, se utiliza la tabla normal; b) se supone desconocido el desvío, el tamaño de la muestra es chica, se utiliza la tabla t.
- 7- a) F, es un estadístico.
 b) F, si el tamaño de la muestra es grande
 c) F, agrandar el tamaño de la muestra.
 d) V
- 8- $n=(1,65 \cdot 1036/250)^2=47$
- 9- $n=(2,33 \cdot 4/0,25)^2=1390$
- 10- $n=(1,96 \cdot 6035/1200)^2=97,16$ aprox. 98
- 11- a) $|\bar{X} - \mu| < \varepsilon = 1,96 \cdot \frac{12}{\sqrt{60}} = 3,036$
 b) $n=(1,96 \cdot 12/1,518)^2=240$
- 12- $I_P = (0,54 - 2,33 \cdot \sqrt{\frac{0,54 \cdot 0,46}{100}}; 0,54 + 2,33 \cdot \sqrt{\frac{0,54 \cdot 0,46}{100}}) = (0,54 - 0,12; 0,54 + 0,12) = (0,42; 0,66)$
- 13- a) $I_P = (0,54 - 1,96 \cdot \sqrt{\frac{0,54 \cdot 0,46}{400}}; 0,54 + 1,96 \cdot \sqrt{\frac{0,54 \cdot 0,46}{400}}) = (0,54 - 0,05; 0,54 + 0,05) = (0,49; 0,59)$
 b) No.
 c) Con $\varepsilon=0,04$ $n=(1,96 \cdot 0,5/0,04)^2=600,25$ aprox. 601
- 14- a) $I_\mu = (850 - 1,96 \cdot \frac{25}{\sqrt{36}}; 850 + 1,96 \cdot \frac{25}{\sqrt{36}}) = (850 - 8,17; 850 + 8,17) = (841,83; 858,17)$
 b) $I_\sigma = (\sqrt{\frac{35.625}{49,8}}; \sqrt{\frac{35.625}{22,47}}) = (20,96; 94,11)$

- 15- a) $I_{\mu} = (9,83 - 1,96 \cdot \frac{0,87}{\sqrt{410}} ; 9,83 + 1,96 \cdot \frac{0,87}{\sqrt{410}}) = (9,75; 9,91)$
- b) Con $p = 324/410 = 0,79$ $I_p = (0,79 - 1,65 \cdot \sqrt{\frac{0,79 \cdot 0,21}{410}} ; 0,79 + 1,65 \cdot \sqrt{\frac{0,79 \cdot 0,21}{410}}) = (0,76; 0,83)$
- c) i) $11,2 = 11 + 0,5 \cdot \frac{k - 89,27}{6,58}$ $k = 92\%$ $p = 0,92$
- $I_p = (0,92 - 1,65 \cdot \sqrt{\frac{0,92 \cdot 0,08}{410}} ; 0,92 + 1,65 \cdot \sqrt{\frac{0,92 \cdot 0,08}{410}}) = (0,90 ; 0,94)$
- ii) $9,8 = 9,5 + 0,5 \cdot \frac{k - 36,34}{23,66}$ $k = 50,54\%$ $\text{aprox. } p = 0,49$
- $I_p = (0,49 - 1,65 \cdot \sqrt{\frac{0,49 \cdot 0,51}{410}} ; 0,49 + 1,65 \cdot \sqrt{\frac{0,49 \cdot 0,51}{410}}) = (0,45; 0,53)$
- 16- a) $I_{\mu} = (10,5 - 1,65 \cdot \frac{1,6}{\sqrt{60}} ; 10,5 + 1,65 \cdot \frac{1,6}{\sqrt{60}}) = (10,16; 10,84)$
- $I_{\sigma} = (\sqrt{\frac{59,2,56}{41,49}} ; \sqrt{\frac{59,2,56}{76,78}}) = (1,42; 1,90)$
- b) A mayor nivel de confianza, menor precisión.
- c) $n = (1,65 \cdot 1,6/0,10)^2 = 697$
- 17- F-F-V-V
- 18- Con $P = 425/550 = 0,77$ $I_p = (0,77 - 2,33 \cdot \sqrt{\frac{0,77 \cdot 0,23}{550}} ; 0,77 + 2,33 \cdot \sqrt{\frac{0,77 \cdot 0,23}{550}}) = (0,73; 0,81)$
- 19- a) $I_{\mu} = (119333,33 - 2,57 \cdot \frac{3777,12}{\sqrt{6}} ; 119333,33 + 2,57 \cdot \frac{3777,12}{\sqrt{6}}) = (115370,38; 122371,0755)$
- b) A mayor nivel de confianza, menor precisión.
- c) $I_{\sigma} = (\sqrt{\frac{5,1426635,49}{16,75}} ; \sqrt{\frac{5,1426356,49}{0,41}}) = (2063,66; 13190)$
- 20- a) $I_{\mu} = (986 - 3,75 \cdot \frac{40,07}{\sqrt{5}} ; 986 + 3,75 \cdot \frac{40,07}{\sqrt{5}}) = (918,8; 1053,2)$
- b) $I_{\sigma}^2 = (\frac{4,1605,6}{9,49} ; \frac{4,1605,6}{0,71}) = (676,75; 9045,63)$
- c) Población con distribución aproximadamente normal.
- 21- a) $I_{\mu} = (3335,71 - 3,14 \cdot \frac{640,78}{\sqrt{7}} ; 3335,71 + 3,14 \cdot \frac{640,78}{\sqrt{7}}) = (2575,23; 4096,19)$
- $I_{\sigma} = (\sqrt{\frac{6,410599,0084}{16,81}} ; \sqrt{\frac{6,410599,0084}{0,872}}) = (382,83; 1680,34)$
- b) A menor nivel de confianza, mayor precisión
- 22- a) $I_{\mu} = (580 - 1,96 \cdot \frac{40}{\sqrt{51}} ; 580 + 1,96 \cdot \frac{40}{\sqrt{51}}) = (569,02; 590,98)$
- b) $I_{\sigma} = (\sqrt{\frac{50,1600}{67,5}} ; \sqrt{\frac{50,1600}{34,76}}) = (34,43; 47,97)$

23- a) $I_{\mu} = (683,41 - 1,96 \cdot \frac{86,62}{\sqrt{410}} ; 683,41 + 1,96 \cdot \frac{86,62}{\sqrt{410}}) = (675,02; 691,79)$
 b) $690 = 650 + 50 \cdot \frac{k - 36,34}{23,66}$ $k = 55,27\%$ $p = 0,55$

$$I_p = (0,55 - 1,65 \cdot \sqrt{\frac{0,55 \cdot 0,45}{410}} ; 0,55 + 1,65 \cdot \sqrt{\frac{0,55 \cdot 0,45}{410}}) = (0,51; 0,59)$$

Mix de estimación por intervalos

24- a) $I_{\mu} = (8,0492 - 1,96 \cdot \frac{0,591}{\sqrt{50}} ; 8,0492 + 1,96 \cdot \frac{0,591}{\sqrt{50}}) = (7,885; 8,213)$

b) $I_{\sigma} = (\sqrt{\frac{49,0,3493}{66,34}} ; \sqrt{\frac{49,0,3493}{33,93}}) = (0,5079; 0,7102)$

c) $I_p = (0,5 - 1,64 \cdot \sqrt{\frac{0,5 \cdot 0,5}{50}} ; 0,5 + 1,64 \cdot \sqrt{\frac{0,5 \cdot 0,5}{50}}) = (0,3614 ; 0,6386)$

$I_p = (0,5 - 2,58 \cdot \sqrt{\frac{0,5 \cdot 0,5}{50}} ; 0,5 + 2,58 \cdot \sqrt{\frac{0,5 \cdot 0,5}{50}}) = (0,3176 ; 0,6824)$

d) Es más preciso el intervalo del 95% de confianza

e) $I_p = (0,88 - 1,44 \cdot \sqrt{\frac{0,88 \cdot 0,12}{50}} ; 0,88 + 1,44 \cdot \sqrt{\frac{0,88 \cdot 0,12}{50}}) = (0,8138; 0,9462)$

25- a) $I_{\mu} = (2356,96 - 1,64 \cdot \frac{900,19}{\sqrt{60}} ; 2356,96 + 1,64 \cdot \frac{900,19}{\sqrt{60}}) = (2166,37; 2547,55)$

b) $I_p = (0,45 - 1,44 \cdot \sqrt{\frac{0,45 \cdot 0,55}{60}} ; 0,45 + 1,44 \cdot \sqrt{\frac{0,45 \cdot 0,55}{60}}) = (0,3575 ; 0,5425)$

c) $I_{\sigma}^2 = (\frac{59,810344,71}{73,28} ; \frac{59,810344,71}{45,58}) = (652427,75; 1048924,19)$

26- $n_1 = 31$ $\sigma^2 = 9$ (A, B)
 $n_2 = 28$ $\sigma^2 = 10$ (A, C)

$$A = \frac{30,9}{\chi_{30;0,975}^2} = \frac{27,10}{\chi_{27;1-\alpha/2}^2} \text{ (Estadístico de prueba)}$$

$46,98 = \chi_{27;1-\alpha/2}^2$ $1-\alpha/2 = 0,99$ $\alpha = 0,02$ $1-\alpha = 0,98$

Práctica 11

Prueba de hipótesis

La prueba de hipótesis es otra técnica de inferencia estadística y en este tipo de situaciones de trabajo, si bien el parámetro es desconocido se posee algún valor referencial previo sobre él. El valor de referencia o valor hipotético puede ser alguna información basada en el comportamiento histórico, algún supuesto, alguna condición que debería cumplirse (por ejemplo, en el control de calidad) o algún valor establecido arbitrariamente.

GPS ESTADÍSTICO 11

Orientaciones importantes para no perderse

- *En las pruebas paramétricas un valor hipotético del parámetro será sometido a prueba a raíz de alguna sospecha o indicio de cambio o variación.*
- *Dado que no hay certezas en las conclusiones de las pruebas de hipótesis estadísticas, es conveniente el uso del tiempo potencial (podría, tendría, etc.) tanto en la formulación de las preguntas como al arribar a las conclusiones.*
- *En las pruebas de hipótesis para la varianza, solo las conclusiones pueden extenderse al desvío. Es decir que no existe por sí misma una prueba para el desvío.*
- *Si en una prueba de hipótesis, cualquiera sea el estadístico de prueba, cae en un entorno próximo al punto crítico se sugiere: no decidir y aumentar el tamaño de la muestra para realizar una nueva prueba.*

Glosario 11

Hipótesis nula: es la que se pone a prueba.

Hipótesis alternativa: expresa lo que la experiencia u observación indican (indicio o evidencia muestral).

Error tipo I: es rechazar una hipótesis nula si fuera verdadera.

Error tipo II: es no rechazar (o aceptar) una hipótesis nula si fuera falsa.

Nivel de significación, α : es la probabilidad de cometer un error de tipo I.

Problemas y situaciones prácticas 11

- 1- Una inmobiliaria se especializa en el alquiler de predios para siembra en la provincia de Entre Ríos. Sus registros indican que el tiempo medio de espera para lograr un alquiler es de 90 días, sin embargo, debido a la alta competencia, se cree que ahora el tiempo medio de espera ha variado significativamente. Por tal motivo, se analizan 100 campos que la agencia ha alquilado recientemente, elegidos al azar, resultando un tiempo de espera hasta la firma del contrato de 94 días con un desvío de 22 días.
- Plantear las hipótesis estadísticas del problema y justificar su elección.
 - Acerca de la distribución del estimador a utilizar, ¿se cuenta con la información poblacional necesaria?
 - ¿Bajo qué suposición sobre la variable en estudio se puede resolver la prueba?
 - Fijando un nivel de significación de 0,10, enunciar la regla de decisión.
 - Expresar el significado de α en el contexto del problema.
 - Decidir si la sospecha de la agencia es válida. ¿En qué se basa la decisión?
- 2- En relación con la situación planteada en el ejercicio 1 de la P10, realizar una prueba de hipótesis para la media del desperdicio diario de producto, con un nivel de significación de 0,02, para analizar la posibilidad de que este sea efectivamente de 500 gramos.
- 3- Una empresa de estufas ha puesto en promoción un calefactor para que se modifique su precio. Antes de la promoción, el precio promedio del producto era de \$15.104, con una desviación estándar de \$1.805. La empresa muestrea a 36 de sus minoristas después de iniciada la promoción y encuentra que el precio medio ahora es de \$14.679. A un nivel de significación de 2 %, ¿tiene razones la empresa para creer que el precio promedio ha disminuido?
- 4- Para medir el nivel de contenido de nicotina de una marca de cigarrillos se seleccionó una muestra de 36 unidades y se obtuvieron los siguientes resultados:

Cant. de nicotina	16	17	18	19	20	21	22	23
Cant. de unidades	2	3	7	9	7	4	2	2

Si el contenido medio de nicotina para esa marca de cigarrillos tuviera que ser menor que 20, ¿qué opinaría usted en función de la información muestral?

- 5- Un fabricante ofrece baterías (pilas) para transistor que tienen una duración promedio de 4.000 horas. Un comprador cualquiera adquiere 36 pilas y encuentra que la duración promedio es de 3.600 horas, resultado que lo obliga a pensar que fue engañado ya que la duración promedio de las 36 pilas fue inferior a la ofrecida por el fabricante. El desvío de la duración de estas pilas adquiridas es de 985 horas. Al nivel de 5 %, se quiere saber si el fabricante está ofreciendo un producto de menor calidad.

- 6- Si la Cámara Farmacéutica tiene como promedio histórico un número de 615 clientes diarios que asisten a las farmacias de esa ciudad,
- ¿Estaría la muestra del ejercicio 6 de la P10 indicando un cambio? Usar un nivel de significación de 0,05 y el valor de σ .
 - ¿Representa algún inconveniente que la muestra sea de tamaño $n=5$? Justificar la respuesta.
- 7- Una compañía aseguradora empezará una extensa campaña publicitaria para vender seguros de vida en una determinada ciudad si considera que el ingreso medio mensual de todas las familias es inferior a \$60.000. Una muestra aleatoria de 41 familias de esa ciudad dio por resultado un ingreso promedio de \$54.600 y una desviación estándar de \$12.000
- Tomando como base la evidencia de la muestra, ¿se dan las condiciones para empezar la campaña al nivel del 1 %?
 - ¿Puede la conclusión a la que llegó ser errónea utilizando la evidencia de la muestra? ¿De qué tipo de error se trataría? Justifique.
 - ¿Es verdadera la siguiente afirmación? En una prueba de hipótesis, la amplitud de la zona de no rechazo de la H_0 está determinada por el nivel de significación de la prueba. Justifique la respuesta.
- 8- El fabricante de máquinas para envasar insumos industriales secos, afirma que la máquina envasa en promedio 250 gramos de producto por paquete con un desvío de 8 gramos. ¿Está en lo cierto el fabricante en su afirmación si una muestra de 45 bolsas arrojó una media de 245 gramos con un desvío de 10 gramos? Sugerencia: testear primero la dispersión y en función de la decisión, usar la información que resulte más confiable.
- 9- Un proceso químico ha producido históricamente un promedio de 1.000 toneladas diarias de un determinado producto. A partir de la muestra de cinco días que figura en el ejercicio 20 de la P10, ¿se podría concluir que la producción promedio diaria sería significativamente menor que 1.000 toneladas y que por lo tanto algo anda mal en el proceso?
- Efectuar la prueba asignando 0,05 a la probabilidad de rechazar que el promedio es de 1.000 toneladas suponiendo que lo fuere.
 - ¿Qué supuestos se requieren para que sea válido el procedimiento aplicado para realizar la prueba de hipótesis?
- 10- A partir del muestreo realizado por el señor Mascetti, datos del ejercicio 22 de la P10,
- Si el rendimiento promedio histórico de naranjas por árbol era de 600, ¿qué puede decir el señor Mascetti sobre la posible existencia de un daño con un nivel de significación del 10 %?
 - ¿Puede la conclusión a la que llegó (en base a la evidencia de esta muestra) ser errónea? ¿De qué tipo de error se trataría? Justifique.
- 11- Con un mayor nivel de significación en una prueba de hipótesis, la probabilidad de rechazar una hipótesis nula (H_0) que en realidad fuera verdadera es (marcar la/s opción/es correcta/s):
- disminuye
 - permanece inalterada
 - aumenta

- 12- Un vendedor de coches nuevos calcula que su compañía tiene que promediar más de 4,8 % de ganancias en las ventas de los autos nuevos asignados. Una muestra aleatoria de 70 coches vendidos proporcionó una media y una desviación estándar del porcentaje de ganancias por vehículo de 4,87 % y 3,9 % respectivamente. ¿Aportan los datos evidencia suficiente que indique que la política del gerente de ventas al aprobar los precios generaría una ganancia media superior a 4,8 % por coche?
- 13- Si un método de lectura veloz es efectivo en el 72 % de los casos y un nuevo método ensayado ha mejorado la lectura de 154 de los primeros 205 estudiantes que lo tomaron. ¿Se puede concluir que el último método daría en general mejores resultados? Usar $\alpha=0,02$
- 14- Un análisis de control de calidad debe mantener la proporción de artículos defectuosos en una línea de ensamble por debajo del 5 %. De una muestra de 175 artículos, 8 están defectuosos, ¿debe pararse el proceso o se deja continuar?
- 15- A partir de la información muestral suministrada en el ejercicio 12 de la P10, ¿es correcto concluir que existe evidencia suficiente de que la mayoría de las personas serían inducidas a comprar un producto promocionado? Si $\alpha = 0,05$
- 16- Una muestra aleatoria de 150 compradores tomada durante una liquidación muestra que el 95 % son mujeres. ¿Justificaría esta información la conclusión de que la proporción de compradoras podría ser significativamente mayor al 90 %? Utilizar una probabilidad de error de tipo 1 de 0,10
- 17- ¿Constituyen los datos de la muestra de oyentes del programa “Los diez más votados”, del ejercicio 18 de la P10, evidencia suficiente para contradecir que al menos el 80 % de ellos votan por el conjunto Miranda como sostiene la empresa especializada?
- 18- La empresa distribuidora de las heladeras FRIOX afirma que como mínimo el 40 % de las heladeras vendidas son blancas. Decidir si es aceptable esta afirmación sabiendo que en una muestra de 300 heladeras vendidas 140 fueron blancas. Utilizar un nivel de significación del 5 %
- 19- Un fabricante de salsa de tomate está en proceso de decidir si produce o no una versión picante. El departamento de investigación de mercados de la compañía usó una encuesta nacional a 6.000 hogares y encontró que 335 de ellos comprarían la salsa. Un estudio anterior mucho más extenso mostró que el 5 % de las casas comprarían esa marca. Con un nivel de significación del 2 %, ¿debe la compañía concluir que hay un incremento en el interés por el sabor picante?
- 20- En un proceso de llenado, la tolerancia para el peso de los recipientes es de 8 gramos. Para cumplir este requisito, el desvío estándar en el peso debe ser de 2 gramos. Los pesos de 25 recipientes seleccionados al azar dieron como resultado una desviación estándar de 2,8 gramos. Si los pesos se encuentran normalmente distribuidos, determinar si la varianza de estos

sería significativamente diferente del valor necesario. Utilice 0,02 como probabilidad de cometer un error del tipo 1

- 21- Utilizando los datos del ejercicio 13 de la P10, ¿estaría en lo correcto el candidato al afirmar que ganaría por mayoría, si obtiene 216 apoyos en una muestra de 400 personas? $\alpha = 0,05$.
- 22- El desvío estándar de las piezas fabricadas por una máquina es de 4 mm. Existe la posibilidad de comprar una máquina similar, más moderna y con un desvío notablemente menor (según el importador). Sin embargo, antes de emprender la compra, se desea solicitar permiso al importador para fabricar como prueba 35 piezas, las cuales brindaron una varianza de 11,85 mm². Utilizando una probabilidad de incurrir en un error de tipo I del 10 %, ¿sería verdad la afirmación del importador?
- 23- Un fabricante de almohadas requiere que la desviación estándar del peso del relleno de la producción no exceda significativamente los 5 g. Con una probabilidad de error tipo I de 0,01 y con base en una muestra aleatoria de tamaño 40 y desvío estándar de 6,8 g probar si puede decirse que se está cumpliendo el requisito del fabricante.

Mix de pruebas de hipótesis

- 24- En base a la muestra de alumnos pasantes del Departamento de Ciencias Económicas resumidos en la Matriz 1 del Anexo.
- a) Con un nivel de significación de 0,01, ¿es posible afirmar que el promedio de notas de las materias aprobadas de los alumnos pasantes del Departamento de Ciencias Económicas podría ser superior a 8 puntos?
 - b) Con un nivel de significación del 5 %, ¿puede considerarse que el desvío de los promedios de las notas superaría significativamente los 0,50 puntos?
 - c) Existe evidencia suficiente para concluir que la mayoría de los alumnos se moviliza en transporte público (o sea colectivo o tren)? Utilizar un nivel $\alpha=0,06$
- 25- En base a la muestra de empleados de la empresa de turismo internacional resumidos en la Matriz 2 del Anexo:
- a) Con un nivel de significación de 5 %, ¿es posible afirmar que el sueldo de los empleados no supera los USD 2.500?
 - b) ¿Existe evidencia suficiente para concluir que su desvío es superior a los USD 820? Recordar que no existe prueba de hipótesis para el desvío σ .
 - c) ¿Puede decirse que en esta empresa predominan los empleados hombres? Probarlo con un 10 % de significación.

RESPUESTAS DE LAS ACTIVIDADES DE LA PRÁCTICA 11

Prueba de hipótesis

1- a) $H_0) \mu = 90$ días $H_a) \mu > 90$ días

b) Con $\alpha = 0,10$, $t = \frac{94 - 90}{\frac{22}{\sqrt{100}}} = 1,82$ y $t_{n-1; 1-\alpha} = t_{99; 0,90} = 1,29$, se rechaza H_0 a favor de H_a ,

entonces $\mu > 90$ días, el tiempo medio de espera ha variado con $\alpha = 0,10$

c) Como no es dato la varianza poblacional, debemos suponer que la población de los días de venta de las granjas se distribuye en forma normal o aproximadamente normal.

d) Hay una probabilidad de 0,10 de rechazar que el tiempo medio de espera para lograr un alquiler es de 90 días siendo verdadero.

e) De acuerdo con la muestra de 100 granjas, el tiempo medio de venta ha aumentado, con un nivel de significación del 10 %. Se podría decir que la sospecha de la agencia es válida.

2- $H_0) \mu = 500$ g $H_a) \mu > 500$ g

Con $\alpha = 0,02$, $z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{504 - 500}{\frac{5}{\sqrt{16}}} = 3,2$ y $Z_{0,98} = 2,05$, se rechaza H_0 a favor de H_a , en-

tonces el peso medio de todos los paquetes envasados podría ser significativamente superior a 500 g.

3- $H_0) \mu = 15104$ $H_a) \mu < 15104$

Con $\alpha = 0,02$, $z = \frac{14679 - 15104}{\frac{1805}{\sqrt{36}}} = -1,41$ y $z_{0,02} = -2,05$, no se rechazó H_0 a favor de H_a ,

entonces la empresa no tendría razones para creer que el precio promedio habría disminuido al 2 % de significación.

4- $H_0) \mu = 20$ $H_a) \mu < 20$

Con $\alpha = 0,05$, $t = \frac{19,28 - 20}{\frac{1,75}{\sqrt{36}}} = -2,47$ y $t_{35; 0,05} = -1,69$, se rechaza H_0 a favor de H_a ,

entonces, el nivel medio de nicotina de todos los cigarrillos podría ser significativamente inferior a 20 al 5 % de significación.

5- $H_0) \mu = 4000$ $H_a) \mu < 4000$

Con $\alpha = 0,05$, $Z = \frac{4000 - 3600}{\frac{985}{\sqrt{36}}} = -2,44$ y $z_{0,05} = -1,64$, se rechaza H_0 a favor de H_a , entonces

existe evidencia suficiente para suponer que el fabricante está ofreciendo un producto de menor calidad al 5 % de significación.

6-a) $H_0) \mu=615$

$H_a) \mu < 615$

Con $\alpha=0,05$, $Z = \frac{542,2 - 615}{\frac{120,3}{\sqrt{5}}} = -1,35$ y $z_{0,05} = -1,64$, no se rechaza H_0 a favor de H_a , en-

tonces no existe evidencia suficiente para suponer que la muestra estaría ofreciendo un cambio al 5 % de significación.

b) No, la distribución de la población es normal.

7- a) $H_0) \mu = 60000$

$H_a) \mu < 60000$

Con $\alpha = 0,05$ $T = \frac{60000 - 54600}{\frac{12000}{\sqrt{41}}} = -2,88$ y $t = -1,68$, se rechaza que el ingreso medio

mensual de todas las familias es igual a \$60.000 en favor de que sería significativamente inferior a los \$60.000. Sí, la prueba podría ser errónea.

b) Se trataría del error de tipo I, es decir que nos podría haber tocado en la selección una muestra de probabilidad baja, es decir un grupo de las familias de menores ingresos.

c) Es falso, la amplitud de la zona de rechazo está dada por el nivel de significación de la prueba.

8- $H_0) \sigma^2 = 64$

$H_a) \sigma^2 > 64$

Con $\alpha = 0,05$ $X^2 = 44.100/64 = 68,75$ y $\chi^2_{44; 0,95} = 60,48$ se rechaza H_0 a favor de H_a ,

entonces no estaría en lo cierto el fabricante al afirmar que la máquina envasa con desvío de 8 gramos al 5 % de significación.

9- a) $H_0) \mu = 1.000$ $H_a) \mu < 1.000$

Con $\alpha=0,05$, $t = \frac{986 - 1000}{\frac{40,07}{\sqrt{5}}} = -0,78$ y $t_{4;0,05} = -2,132$, la producción diaria sería significa-

tivamente menor a 1000 toneladas con un nivel de significación del 5 %

c) Población aproximadamente normal.

10- a) $H_0) \mu = 600$ $H_a) \mu > 600$

Con $\alpha=0,10$, $t = \frac{580 - 600}{\frac{40}{\sqrt{51}}} = -3,57$ y $t_{50;0,10} = -1,30$, el señor Mascetti podrá decir que

existe un daño significativo causado a los árboles en base a la muestra de 51 árboles y con un nivel significación del 10 %

b) Sí, la prueba podría ser errónea, se trataría de un error de tipo I, es decir que podría habernos tocado en la selección una muestra de probabilidad baja, es decir que podría haber un menor nivel de daños en el árbol.

11- c) Aumenta.

12- $H_0) \mu = 4,8 \%$

$H_a) \mu > 4,8 \%$

Con $\alpha = 0,05$ $t = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}} = \frac{4,87 - 4,8}{\frac{3,9}{\sqrt{70}}} = 0,15$ y $t_{69; 0,95} = 1,67$ o $Z_{0,95} = 1,64$, los datos no

aportan evidencia suficiente que indique que la política del gerente de ventas generaría una ganancia media superior a 4,8 % con un nivel de significación del 5 %

13- $H_0) P = 0,72$ $H_a) P > 0,72$

Con $\alpha = 0,05$

$$z = \frac{p - P_0}{\sqrt{\frac{P_0 \cdot (1 - P_0)}{n}}} = \frac{0,75 - 0,72}{\sqrt{\frac{0,72 \cdot 0,28}{205}}} = \frac{0,03}{0,03136} = 0,96$$

y $z_{0,95} = 1,64$, en base a la muestra de 205 estudiantes y con un nivel de significación del 5 %, podría pensarse que con el nuevo método la proporción de estudiantes no habría mejorado. Es decir, el último método no daría mejores resultados.

14- $H_0) P = 0,05$ $H_a) P < 0,05$

Con $\alpha = 0,05$

$$z = \frac{p - P_0}{\sqrt{\frac{P_0 \cdot (1 - P_0)}{n}}} = \frac{0,0457 - 0,05}{\sqrt{\frac{0,05 \cdot 0,95}{175}}} = \frac{-0,0043}{0,016475} = -0,26$$

y $z_{0,05} = -1,64$, de acuerdo con la muestra seleccionada para el análisis de control de calidad se puede decir que la proporción de artículos defectuosos no está por debajo del 5 %. Con un nivel de significación del 5 % se podría decidir parar el proceso.

15- $H_0) P = 0,50$ $H_a) P > 0,50$

$$\text{Con } \alpha = 0,05, z = \frac{p - P_0}{\sqrt{\frac{P_0 \cdot (1 - P_0)}{n}}} = \frac{0,54 - 0,50}{\sqrt{\frac{0,50 \cdot 0,50}{100}}} = 0,80$$

y $z_{0,95} = 1,64$, a partir de la información muestral no es correcta la afirmación de los patrocinadores del programa. Con un nivel de significación del 5 %, se podría decir que no hay evidencia de que la mayoría de las personas estarían inducidas a comprar el producto.

16- $H_0) P = 0,90$

$H_a) P > 0,90$

Con $\alpha = 0,10$

$$z = \frac{p - P_0}{\sqrt{\frac{P_0 \cdot (1 - P_0)}{n}}} = \frac{0,95 - 0,90}{\sqrt{\frac{0,90 \cdot 0,10}{150}}} = \frac{0,05}{0,02449} = 2,04$$

y $z_{0,90} = 1,28$, de acuerdo con la muestra de 150 compradores y con una probabilidad de error de tipo I de 0,10 la proporción de compradoras sí sería significativamente mayor al 90 %

17- $H_0) P = 0,80$ $H_a) P < 0,80$

Con $\alpha = 0,20$, $Z = \frac{0,77 - 0,80}{\sqrt{\frac{0,80 \cdot 0,20}{550}}} = -1,76$ y $Z_{0,20} = -1,29$, los datos de la muestra de radioes-

cuchas del programa “Los diez más votados” sí constituyen evidencia suficiente para contradecir que al menos el 80 % de los radioescuchas votan por el conjunto Miranda como sostiene la empresa especializada.

18- $H_0) P = 0,40$ $H_a) P > 0,40$

Con $\alpha = 0,05$, $Z = \frac{0,47 - 0,40}{\sqrt{\frac{0,40 \cdot 0,60}{300}}} = 2,48$ y $Z_{0,95} = 1,64$, sí sería factible que más del 40 % de

las heladeras vendidas son blancas, a partir de una muestra de 300 heladeras con el 5 % de significación.

19- $H_0) P = 0,05$ $H_a) P > 0,05$

Con $\alpha = 0,02$, $z = 2,07$ y $z_{0,98} = 2,06$, como el estadístico de prueba cae tan cerca del valor crítico se recomienda no decidir y, si fuera posible, agrandar el tamaño de la muestra.

20- $H_0) \sigma^2 = 4$ $H_a) \sigma^2 > 4$

Con $\alpha = 0,02$

$$\chi^2 = \frac{(n - 1) \cdot s^2}{\sigma^2} = \frac{24,7,84}{4} = 47,04$$

Como $\chi_{n-1; \alpha}^2 = \chi_{24; 0,98}^2 = 36,415$ de acuerdo con la muestra de 25 recipientes y con un nivel de significación del 2 %, la varianza sería significativamente superior al valor requerido en el proceso de llenado, si es diferente del valor necesario.

21- $H_0) P = 0,50$ $H_a) P > 0,50$

Con $\alpha = 0,05$, $z = \frac{0,54 - 0,5}{\sqrt{\frac{0,5 \cdot 0,5}{400}}} = 1,6$ y $z_{0,95} = 1,64$, no existe evidencia suficiente para

afirmar que el candidato ganaría por mayoría al 5 % de significación.

22- $H_0) \sigma^2 = 16$ $H_a) \sigma^2 < 16$

Con $\alpha = 0,10$

$$X^2 = \frac{(n-1) \cdot s^2}{\sigma^2} = \frac{34.11,85}{16} = 25,18$$

Como $\chi_{n-1;\alpha}^2 = \chi_{34;0,10}^2 = 23,95$, no sería verdadera la afirmación del importador al 10 % de significación.

23- Ho) $\sigma^2 = 25$ Ha) $\sigma^2 > 25$

Con $\alpha = 0,01$ $X^2 = \frac{(n-1) \cdot s^2}{\sigma^2} = \frac{39.46,24}{25} = 72,13$ rechazo H_0

Como $\chi_{n-1;\alpha}^2 = \chi_{39;0,99}^2 = 62,428$ de acuerdo con la muestra de tamaño 40 y al 1 % de significación, la varianza sería significativamente superior al valor requerido por lo cual no se cumpliría el requisito de que la desviación estándar del peso del relleno de la producción no exceda significativamente los 5 g.

24- a) Ho) $\mu = 8$ Ha) $\mu > 8$

Con $\alpha = 0,01$, $T = 0,59$ y $t_{49;0,99} = 2,4$ no hay evidencia de que el promedio sea mayor que 8.

b) Ho) $\sigma^2 = 0,5^2$ Ha) $\sigma^2 > 0,5^2$

Con $\alpha = 0,05$, $X^2 = 68,23$ y $\chi_{49;0,95}^2 = 66,34$, el desvío sería superior a 0.5

c) Ho) $P = 0,5$ Ha) $P > 0,5$

Con $\alpha = 0,06$ $Z = 1,98$ y $z_{0,94} = 1,56$, la mayoría viaja en transporte público

25- Población $\mu_0 = 2500$ $\sigma_0 = 820$ $P_0 = 0,5$

Muestra $n = 60$ $\bar{X} = 2356,96$ $s = 900,19$ $p = \frac{33}{60} = 0,55$

a) Ho) $\mu = 2500$ Ha) $\mu < 2500$

Con $\alpha = 0,05$, $T = -1,23$ y $t_{59;0,05} = -1,67$, el sueldo medio sería de \$2.500

b) Ho) $\sigma^2 = 820^2$ Ha) $\sigma^2 > 820^2$

Con $\alpha = 0,05$, $X^2 = 71,1$ y $\chi_{59;0,95}^2 = 77,93$, el desvío no sería superior a 820.

c) Ho) $P = 0,5$ Ha) $P > 0,5$

Con $\alpha = 0,10$, $Z = 0,77$ y $z_{0,99} = 1,28$, no hay evidencia de que predominen los hombres.

PRÁCTICA 12

Análisis de regresión y de correlación

El análisis de regresión entre variables económicas o de otra índole consiste en elaborar un modelo de relación entre ellas.

El análisis de correlación entre variables consiste en determinar el grado de asociación lineal que poseen.

GPS ESTADÍSTICO 12

Orientaciones importantes para no perderse

- *En un estudio de regresión y correlación es preciso ser cuidadoso acerca de los valores elegidos de la variable independiente o regresora X a los que se aplicará el método a fin de pronosticar la variable regresiva Y . Para esto es preciso tener presente la extensión y ubicación del conjunto de valores muestrales de X y no ir mucho más allá de las cercanías de este.*
- *Los valores muestrales de los coeficientes dan una idea del grado de asociación entre las variables X e Y , pero no resultan información suficiente para concluir de manera taxativa sobre la calidad de su relación. De ahí que sea necesario el planteo de una prueba de hipótesis para poder llegar a conclusiones con otro sustento, basadas en procesos de inferencia estadística.*

Glosario 12

Coefficiente de regresión: es la pendiente de la recta de regresión, indica en cuánto cambia la variable dependiente por cada unidad de cambio en la variable independiente.

Coefficiente de determinación: indica en qué porcentaje la variable dependiente se encuentra relacionada con la variable independiente. Tiene un recorrido de 0 a 1

Coefficiente de correlación: mide la fuerza y el sentido de la relación lineal entre las variables. Tiene un recorrido de -1 a 1 . El sentido (o signo) es una característica que comparte con el coeficiente de regresión.

Problemas y situaciones prácticas 12

1- En un estudio de distintos fondos comunes de inversión se desarrolló un procedimiento consistente en construir la recta característica para cada fondo, esta recta explica la rentabilidad de cada fondo en función de la rentabilidad promedio del mercado. Si la pendiente de dicha recta es significativamente distinta de cero, se dice que el fondo es muy sensible a los cambios y es, por lo tanto, una inversión riesgosa. Si, por lo contrario, la pendiente de la recta es muy cercana a cero, se considera una inversión estable.

Se analizó para el período 2012-2016 la rentabilidad del Fondo común de inversión Redifon (variable y) y el promedio del mercado, obteniéndose los siguientes resultados:

$$\begin{aligned} \sum X &= 30 \\ \sum Y &= 17,60 \end{aligned}$$

$$\begin{aligned} \sum X_i^2 &= 220,00 \\ \sum Y_i^2 &= 66,88 \end{aligned}$$

$$\sum X_i Y_i = 119,6$$

- Encontrar la recta que explique la renta del Fondo Redifon en función de la renta media.
- Decidir si existe correlación lineal entre la rentabilidad promedio del mercado y la del Fondo Redifon con un nivel de significación del 10 %
- Estimar la rentabilidad del Fondo Redifon para una rentabilidad media del mercado de 8,0, si fuera posible.

2- Se tomó una muestra aleatoria de 12 casas construidas, y la información recopilada se presenta en el siguiente cuadro:

	Tamaño del lote (en decenas de m ²)	Costo de construcción (en miles de USD)
1	5	31,6
2	7	32,4
3	10	41,7
4	10	50,2
5	12	46,2
6	20	58,5
7	22	59,3
8	15	48,4
9	30	63,7
10	40	85,3
11	12	53,4
12	15	54,5

- ¿Cuál es la variable explicativa y cuál es la variable explicada? Ajustar los datos a un modelo lineal e interpretar los coeficientes resultantes.
- A un nivel de significación del 5 % realizar una prueba adecuada para decidir si un aumento en el costo de la construcción se corresponde con un aumento en el tamaño del lote.
- ¿Cuál es el costo de construcción esperado para un lote de 150 m²?

- 3- En una fábrica de zapatos se quiere estudiar si el costo total incurrido para la cantidad de unidades (pares de zapatos) y las unidades producidas en ese día se encuentran relacionadas linealmente, para ello se tomó una muestra de 15 días y se recolectó la siguiente información:

	Unidades producidas	Costo total
1	27	5.200
2	20	4.400
3	21	4.480
4	28	5.220
5	29	5.840
6	24	4.680
7	22	4.620
8	24	4.880
9	28	5.780
10	26	4.900
11	21	4.640
12	29	5.920
13	27	5.880
14	24	4.860
15	29	5.500

- Ajustar los datos a un modelo lineal e interpretar los coeficientes resultantes.
 - Realizar una prueba de hipótesis para la correlación con $\alpha = 0,05$ e interpretar.
 - Si se fabricaran 25 unidades de zapatos, estimar el costo total.
- 4- (*) En 1962 el economista norteamericano Arthur [Okun](#) planteó un modelo macroeconómico que se conoce hoy en día como la “ley de Okun” que establece que existe una relación lineal entre el cambio en la tasa de desempleo y la tasa de crecimiento del Producto Bruto Interno (PBI) real.
El siguiente cuadro muestra datos sobre desempleo y crecimiento económico en los Estados Unidos durante el período 1966-95
- Con estos datos estimar el modelo lineal y explicar el significado de los coeficientes obtenidos.
 - Determinar el punto donde la recta de regresión corta el eje x para este caso, y explicar su significado en términos del modelo.
 - Realizar una prueba de hipótesis con un nivel de significación del 5 %

	Tasa de desempleo (%)	Crecimiento PBI real (%)		Tasa de desempleo (%)	Crecimiento PBI real (%)
1966	3,6	6,0	1981	7,5	1,8
1967	3,7	2,6	1982	9,5	-2,2
1968	3,4	4,1	1983	9,5	3,9
1969	3,4	2,7	1984	7,4	6,2
1970	4,8	0,0	1985	7,1	3,2
1971	5,8	3,1	1986	6,9	2,9
1972	5,5	4,8	1987	6,1	3,1
1973	4,8	5,2	1988	5,4	3,9
1974	5,5	-0,6	1989	5,2	2,5
1975	8,3	-0,8	1990	5,6	0,8
1976	7,6	4,9	1991	6,8	-1,2
1977	6,9	4,5	1992	7,5	3,3
1978	6,0	4,8	1993	6,9	3,1
1979	5,8	2,5	1994	6,0	4,1
1980	7,0	-0,5	1995	5,5	2,0

- 5- Los datos de la siguiente tabla corresponden a un estudio sobre la relación ente el grado de entrenamiento y el desempeño posterior en una carrera de 10 km. Se evaluaron 14 mujeres entrenadas. El grado de entrenamiento se mide mediante la “potencia aeróbica máxima” (ml/kg min.) alcanzada y el desempeño posterior mediante el tiempo empleado en completar 10 km durante una competencia.

Or- den:	1	2	3	4	5	6	7	8	9	10	11	12	13	14
X	47,17	47,41	47,88	51,05	51,32	52,18	52,37	52,83	53,31	53,93	55,29	57,91	57,94	61,32
Y	47,83	46,03	45,6	48,55	42,37	43,93	44,9	40,03	42,03	45,12	39,8	44,9	41,32	39,37

Variables: potencia aeróbica máxima X y tiempo en minutos Y

- Ajustar los datos a un modelo lineal y explicar el significado de los coeficientes obtenidos.
 - Realizar una prueba de hipótesis para la correlación con un nivel de significación del 4 %
 - ¿Cuál es el tiempo esperado en minutos para una potencia aeróbica máxima de 65 y de 75?
- 6- En una empresa se quiere estudiar si el monto de los viáticos mensuales pagados a sus corredores (en miles de pesos) y el volumen de ventas mensuales (en millones de pesos) generales están relacionados linealmente. Una muestra de 7 meses arrojó los siguientes datos:

$$\hat{y} = 11,20 + 0,78 x \quad y \quad r^2 = 0,63$$

- Interpretar los coeficientes de regresión, correlación y determinación en el contexto del problema.
- A un nivel de significación del 5 % realizar una prueba adecuada para decidir si el aumento en los viáticos mensuales pagados a sus corredores se corresponde en general con el aumento en el volumen de ventas mensuales.
- Estimar el monto de los viáticos mensuales pagados a sus corredores para un volumen de ventas mensuales de 3 millones de pesos.

- 7- El gerente de ventas de una compañía se está preparando para una reunión de ventas, y le gustaría mostrar al grupo de vendedores la forma como se relaciona el número de visitas mensuales a clientes con el monto mensual de los pedidos que reciben. De sus registros recolectó los siguientes datos correspondientes a 10 meses elegidos al azar.

N.º de visitas	5	4	6	7	8	1	3	4	1	3
Pedidos (miles de \$)	4,8	6,1	12,3	13,7	15,7	3,2	5,3	7,8	1,3	7,6

- a) Calcular los coeficientes de correlación y de determinación. Interpretar.
 b) A partir de los datos, ¿puede llegarse a la conclusión de que, en general, conforme se incrementa el número de visitas, también se aumentaría el monto de pedidos? Realizar una prueba estadística adecuada con un nivel de significación del 2 %
 c) Estimar el monto de los pedidos si en un mes cualquiera elegido al azar se programaran 2 visitas. Ídem si se programaran 7 visitas.
- 8- Una financiera otorga préstamos para la compra de autos, casas, departamentos, etc.; las cuotas son fijas y dependen del monto del préstamo. La siguiente tabla consigna una muestra de cuatro préstamos con sus respectivas cuotas expresados ambos en miles de pesos:

Monto:	7	15	25	60
Cuota:	0,4	2,1	2,6	2,9

- a) ¿Qué porcentaje de la variación de las cuotas está expresado por el modelo de regresión lineal?
 b) Calcule la recta de regresión que crea conveniente y grafíquela.
 c) Probar si hay correlación lineal significativa entre las variables, utilizando $\alpha = 0,01$
 d) ¿Cuál es el valor esperado de cuota para un monto de \$45.000?
 e) A la luz de la conclusión de la prueba c): ¿es pertinente hacer la estimación d)? Justificar.
- 9- Se relevó a 20 niños para registrar su edad y la cantidad de horas diarias que dedican a jugar con una consola la de juegos y se obtuvo la siguiente ecuación:

$$\hat{y} = -0.9972 + 0,4185 x \quad y \quad r^2 = 0,5348$$

- a) ¿Qué porcentaje de la variación en la cantidad de horas diarias que dedican al juego queda explicado por la variación en la edad de los niños?
 b) Con un nivel de significación del 5 % realizar una prueba adecuada para decidir si un aumento en la edad de los niños se corresponde con un aumento en el valor de las horas diarias dedicadas al juego.
 c) Estimar la cantidad de horas que permanece un niño jugando a la consola de juegos si tiene 8 años.
- 10- Explicar por qué son falsas las siguientes afirmaciones:
 a) El coeficiente de determinación indica si la pendiente de la recta de regresión es positiva o negativa.
 b) El coeficiente de regresión mide la fuerza de asociación entre las variables.

- c) El coeficiente de correlación puede tener un valor absoluto mayor que uno.
- d) El coeficiente de correlación mide el porcentaje de la variación de “y” en función de “x” explicada por el modelo de regresión lineal.

11- La siguiente tabla muestra la variación del PBI y de las importaciones de Argentina entre los años 2008 y 2015. La estructura económica argentina se caracteriza por mostrar un estrangulamiento del comercio exterior llamado “restricción externa” que consiste en un aumento más que proporcional de las importaciones cuando aumenta el nivel de actividad (PBI). Históricamente, cuando el PBI argentino creció en un promedio del 1 %, las importaciones lo hicieron entre un 3 % y 4,5 %

Se pide:

- a) Hallar la ecuación de la recta de regresión.
- b) Calcular en qué proporción crecen las importaciones por cada punto porcentual en que lo hace el PBI en el período dado.
- c) A un nivel de significación del 5 % realizar una prueba para decidir si un aumento del PBI se corresponde con un aumento de las importaciones.

RESPUESTAS DE LAS ACTIVIDADES DE LA PRÁCTICA 12

Análisis de regresión y de correlación

Año	PBI	Importaciones
2008	4,1	13,6
2009	-5,9	-18,4
2010	10	35,2
2011	6	22
2012	-1,03	-4,7
2013	2,41	3,9
2014	-2,51	-11,5
2015	2,73	4,7

1- a) $\hat{y}=1,42+0,35.x$

b) $r = \frac{14}{\sqrt{50.4,928}} = 0,9972$ y $r^2=0,9944$

$H_0: \rho = 0$ $H_A: \rho > 0$

Con $\alpha=0,10$, $t = r \cdot \frac{\sqrt{n-2}}{\sqrt{1-r^2}} = 0,9972 \cdot \frac{\sqrt{5-2}}{\sqrt{1-0,9944}} =$

23,097 y $t_{3;0,90}=1,64$, Rechazo H_0 .

Conclusión: existe evidencia suficiente para afirmar que existiría relación lineal significativa entre la rentabilidad promedio del mercado y la del Fondo Redifon al 10 % de significación.

c) $\hat{y}=1,42+0,35.8=4,22$

Se estima una rentabilidad del 4,22 del Fondo Redifon para una rentabilidad media del mercado de 8,0

- 2- a) Tamaño y costo, variables explicativa y explicada respectivamente.

$$a=29,8228 \quad b=1,35013 \quad r=0,9449 \quad r^2=0,89259$$

Coefficiente de regresión: $b= 1,35$. En la muestra de 12 casa, por cada decena de m^2 de lote, el costo de construcción aumenta en 1,35 miles de dólares.

Coefficiente de correlación: $r = 0,9449$. Como el valor es cercano a 1 podemos decir que habría una muy alta correlación lineal directa entre las variables.

Coefficiente de determinación: $r^2= 0,8926$. El 89,26 % de las variaciones en el costo de la construcción quedaron explicadas por las variaciones en el tamaño del lote. El 10,74 % restante se debe a otras causas.

b) $H_0) \rho = 0 \quad H_a) \rho > 0$

$$\text{Con } \alpha=0,10, T=0,944769. \frac{\sqrt{12-2}}{\sqrt{1-0,892589}} = 9,116 \text{ y } t_{10;0,95}=1,81, \text{ Rechazo } H_0$$

Existe evidencia suficiente para afirmar que existiría relación lineal significativa entre el tamaño y el costo al 10 % de significación.

c) $\hat{Y}_{(15)} = 29,82 + 1,35 \cdot 15 = 50,11$

Se espera un costo de construcción de 50,11 miles de dólares con un lote de $150 m^2$

3- a) $a=1255,76 \quad b=152,94 \quad r=0,9021 \quad r^2=0,8138$.

Coefficiente de regresión: $b= 152,94$. En la muestra de 15 días, por cada unidad producida, el costo total aumenta en \$ 152,94

Coefficiente de correlación: $r = 0,9021$. Como el valor es cercano a 1 podemos decir que habría una alta correlación lineal directa entre las variables.

Coefficiente de determinación: $r^2= 0,8138$. El 81,38 % de las variaciones producidas en el costo total quedaron explicadas por una variación en las unidades producidas. El 18,62 % se debe a otras causas.

b) $H_0) \rho = 0 \quad H_a) \rho > 0$

$$\text{Con } \alpha=0,05, t=0,902128. \frac{\sqrt{15-2}}{\sqrt{1-0,8138499}} = 7,5386 \text{ y } t_{13;0,95}=1,77, \text{ Rechazo } H_0$$

Existe evidencia suficiente para afirmar que existiría relación lineal significativa entre las unidades producidas y el costo al 5 % de significación.

c) $\hat{y}(25)=5079,22$

4- a) $a=4,9402 \quad b=-0,3649 \quad r=-0,2638 \quad r^2=0,0696$

Coefficiente de regresión: $b= -0,36$ En la muestra de 30 años, por cada 1 % que aumenta la tasa de desempleo, la tasa de crecimiento del PBI disminuye en un 0,36 %

Coefficiente de correlación: $r = -0,26$ habría una baja correlación lineal inversa entre las variables.

Coefficiente de determinación: $r^2 = 0,07$ El 7 % de las variaciones producidas en la tasa de crecimiento del PBI quedaron explicadas por una variación en la tasa de desempleo. El 93 % restante se debe a otras causas.

b) $x=13,54$ Cuando la tasa de desempleo es 13,72 % la tasa de crecimiento del PBI es 0 %

c) $H_0) \rho = 0$ $H_a) \rho < 0$

$$\text{Con } \alpha=0,05, t = -0,2638. \frac{\sqrt{30-2}}{\sqrt{1-0,06959}} = -1,4472 \text{ y } t_{28;0,05} = -1,7, \text{ No rechazo } H_0$$

No existe evidencia suficiente para afirmar que existiría relación lineal significativa entre la tasa de desempleo y el crecimiento del PBI al 5 % de significación.

5- a) $a=68,49$ $b = -0,47$ $r = -0,6598$ $r^2 = 0,4353$

Coefficiente de regresión: $b = -0,4679$ En la muestra de 14 mujeres entrenadas, por cada 1 ml/kg Min. que aumenta la potencia aeróbica máxima, el tiempo en completar 10 km disminuye en 0,4679 minutos.

Coefficiente de correlación: $r = -0,6598$ habría una alta correlación lineal inversa entre las variables.

Coefficiente de determinación: $r^2 = 0,4353$ El 43,53 % de las variaciones producidas en la en el tiempo en completar 10 km queda explicada por una variación en la potencia aeróbica. El 56,47 % restante se debe a otras causas.

b) $H_0) \rho = 0$ $H_a) \rho < 0$

$$\text{Con } \alpha=0,04, t = -0,65971175. \frac{\sqrt{14-2}}{\sqrt{1-0,435271612}} = -3,04 \text{ y } t_{12;0,04} = -1,78, \text{ Rechazo } H_0$$

Existe evidencia suficiente para afirmar que existiría relación lineal significativa entre la potencia aeróbica máxima y el tiempo en minutos al 4 % de significación.

c) $\hat{y}(65) = 38,08$

No conviene usar este modelo para 75 por está lejos del recorrido de la variable.

6- a) $a=11,2$ $b=0,78$ $r=0,7937$ $r^2=0,63$

Coefficiente de regresión: $b = 0,78$. En los siete meses analizados por cada millón de pesos en que aumenta el volumen de las ventas, los viáticos se incrementan en \$780

Coefficiente de correlación: $r = 0,7937$ habría una alta correlación lineal directa entre las variables.

Coefficiente de determinación: $r^2 = 0,63$ El 63 % de las variaciones en los viáticos pagados a los corredores son explicados por las variaciones del volumen de las ventas realizadas para la empresa y el 37 % restante se debe a otras causas.

b) $H_0) \rho = 0$ $H_a) \rho > 0$

Con $\alpha=0,05$, $t=0,793725393 \cdot \frac{\sqrt{7-2}}{\sqrt{1-0,63}} = 2,918$ y $t_{5;0,95}=2,02$, Rechazo H_0

Existe evidencia suficiente para afirmar que existiría relación lineal directa entre el volumen de las ventas y el monto en los viáticos al 5 % de significación.

c) $\hat{y}(3)=13,54$ Para un volumen de ventas mensuales de 3 millones de pesos se estima un monto mensual de viáticos de 13.540 pesos.

- 7- a) $r=0,919$: la asociación lineal es muy buena (de alto grado) y directa.
 $r^2=0,845$: el 84,5 % de la variación de y con respecto a la variación de x queda bien explicada por este modelo.

b) $H_0: \rho = 0$ $H_a: \rho > 0$

Con $\alpha=0,02$, $t=0,919 \cdot \frac{\sqrt{10-2}}{\sqrt{1-0,845}} = 6,6$ y $t_{8;0,98}= 2,31$, Rechazo H_0

Existe evidencia suficiente para afirmar que conforme se incrementa el número de visitas, también se aumentaría el monto de pedidos con un 2 % de significación.

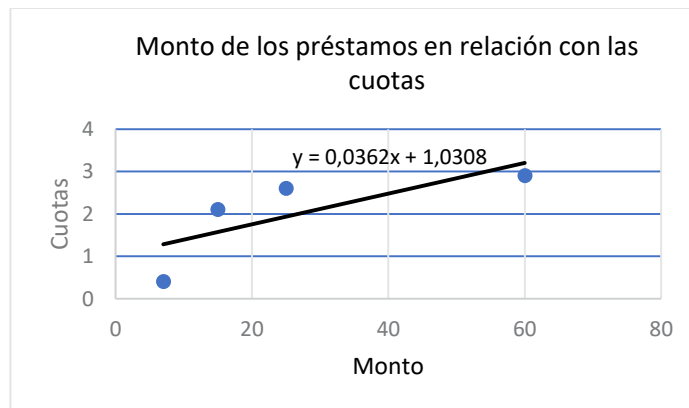
c) $a= 0,0455$ $b= 1,842$

$\hat{y}(2)=3,73$ Si en un mes cualquiera se programan dos visitas, se esperaría que los pedidos asciendan a \$ 3,73 miles de pesos.

$\hat{y}(7)=12,94$ Si en un mes cualquiera se programan siete visitas, se espera que los pedidos sean de \$ 12,94 miles de pesos.

- 8- a) El 57,44 %

b) $\hat{y}=1,03+0,036 x$



c) $H_0) \rho = 0$ $H_a) \rho > 0$

Con $\alpha=0,01$, $T=0,7579 \cdot \frac{\sqrt{4-2}}{\sqrt{1-0,5744}} = 1,6429$ y $t_{2;0,99}=6,96$, No rechazo H_0

No existe evidencia suficiente para afirmar que habría relación lineal significativa entre los montos y las cuotas al 1 % de significación.

- d) No conviene estimarlo mediante este modelo.
- e) No.

9- a) El 53,48 %

b) $H_0) \rho = 0$ $H_a) \rho > 0$

$$\text{Con } \alpha=0,05, T=0,7313 \frac{\sqrt{20-2}}{\sqrt{1-0,5348}} = 4,55 \text{ y } 18_{2;0,95}=1,73, \text{ Rechazo } H_0$$

Existe evidencia suficiente para afirmar que existiría relación lineal directa entre la edad y las horas dedicadas al juego al 5 % de significación.

c) $\hat{y}(8) = 2,35$ Un niño de 8 años pasaría 2,35 horas jugando con la consola de juegos.

10- a) El coeficiente de **determinación** indica si la pendiente de la recta de regresión es positiva o negativa. >>>>Es el de **correlación**

b) El coeficiente de **regresión** mide la fuerza de asociación entre las variables.

>>>>Es el de **correlación**

c) El coeficiente de **correlación** puede tener un valor absoluto **mayor** que uno.

>>>>Debe ser **menor o igual a 1**

d) El coeficiente de **correlación** mide el porcentaje de la variación de “y” en función de “x” explicada por el modelo de regresión lineal.

>>>>Es el de **determinación**

11- a) $\hat{y} = -1,27 + 3,48 x$

b) 3,48 % Por cada un 1 % que aumenta el PBI, las importaciones aumentan en un 3,48 %

c) $r^2 = 0,9792$ $r = 0,9895$

$H_0) \rho = 0$ $H_a) \rho > 0$

$$\text{Con } \alpha=0,05, T=0,9895 \frac{\sqrt{8-2}}{\sqrt{1-0,9792}} = 16,8 \text{ y } t_{6;0,95} = 1,94, \text{ Rechazo } H_0$$

Existe evidencia suficiente para afirmar que un aumento del PBI se corresponde con un aumento de las importaciones al 5 % de significación.

Bibliografía

- Ángel, M., Borgna, E., Brunetti, S., Elizondo, S., Fernández, G., González Camus, M. y Polola, L. (1996). *Estadística. Ciencias Económicas. Ejercicios Resueltos*. C&C.
- Berenson, M. y Levin, D. (2001). *Estadística para administración* (2a ed.). Pearson Educación.
- Canavos, G. (1998). *Probabilidad y Estadística. Aplicaciones y métodos*. McGraw-Hill.
- Cochran, W. (1998). *Técnicas de muestreo* (2a ed.). Trillas.
- Diebold, F. (1999). *Elementos de pronósticos*. Ediciones Paraninfo.
- Freund, J., Williams, F. y Perles, B. (1990). *Estadística para Administración*. Prentice Hall Hispanoamericana.
- Hildebrand, D. y Lyman, O. (1998). *Estadística Aplicada a la administración y a la economía*. Addison Wesley Iberoamericana.
- Kazmier L. y Díaz Mata, A. (1998). *Estadística aplicada a Administración y Economía*. McGraw-Hill.
- Levin, R. y Rubin, D. (2004). *Estadística para Administración y Economía*. Pearson Educación.
- Mason R. Lind, D. y Marchal, W. (2004). *Estadística para Administración y Economía*. Alfaomega.
- Mendenhall W. y Reinmuth, J. (1998). *Estadística para Administración y Economía*. Iberoamérica.
- Stevenson W. (1981). *Estadística para Administración y Economía*. Harla.

Material complementario de consulta: investigaciones realizadas por la cátedra

- Ángel, M., Fernández, G., Polola, L., Borgna, M., Pagano, L., Brunetti, S. y Ecalle, M. (2006). *Génesis y evolución histórica de los conceptos de probabilidad y estadística como herramienta metodológica* (Proyecto Proince Cód. 55B107). Secretaría de Investigaciones del Departamento de Ciencias Económicas, UNLaM. https://economicas.unlam.edu.ar/descargas/5_b107.pdf
- Ángel, M., Fernández, G., Polola, L., Borgna, M., Pagano, L., Brunetti, S. y Ecalle, M. (2008). *La historia como eje metodológico para el tratamiento de conceptos estadísticos en el*

aula (Proyecto Proince Cód. 64B129). Secretaría de Investigaciones del Departamento de Ciencias Económicas, UNLaM. https://economicas.unlam.edu.ar/descargas/645_b129.pdf

Ángel, M., Fernández, G., Polola, L., Borgna, M., Pagano, L., Brunetti, S. y Ecalle, M. (2010). *Registro de conceptos estadísticos como recurso para una alfabetización estadística* (Proyecto Proince. Cód. 55B142). Secretaría de Investigaciones del Departamento de Ciencias Económicas, UNLaM. https://economicas.unlam.edu.ar/descargas/4_b142.pdf

Ángel, M., Polola, L. Pagano, L., Beordi, G., Bortolotto, M., Rodríguez, S. y López, P. (2019). *Relevamiento de tendencias actuales en la enseñanza de la Estadística aplicada a las Ciencias Económicas. Registro y análisis del uso de herramientas tecnológicas en la práctica formativa como factor de innovación* (Proyecto Proince. Cód. 55B210). Secretaría de Investigaciones del Departamento de Ciencias Económicas, UNLaM. https://economicas.unlam.edu.ar/descargas/4_B210.pdf

Rumbo, J., Polola, L. y Ecalle, M. (1998). *El papel del razonamiento lógico en la educación matemática universitaria. Aportes metodológicos para el diseño de programas* (Proyecto Proince Cód. 55B033). Secretaría de Investigaciones del Departamento de Ciencias Económicas, UNLaM. https://economicas.unlam.edu.ar/descargas/5_B033.pdf

Webgrafía

Banco Mundial. (2022). Crecimiento del PBI (% anual) - Argentina. <https://datos.bancomundial.org/indicador/NY.GDP.MKTP.KD.ZG?locations=AR>

Banco Mundial. (2022). *PBI per cápita (USD a precios constantes de 2010) - Argentina*. <https://datos.bancomundial.org/indicador/NY.GDP.PCAP.KD?end=2018&locations=AR&start=2008>

Cámara del Comercio Automotor: CAA. (2022). *Home* [página web]. www.cca.org.ar

Clarín. (26 de abril de 2019). *Home* [página web]. <https://www.clarin.com/>

Comisión Económica para América Latina. (2010). *Infográficos estadísticos de América Latina y el Caribe*. CEPALSTAT. https://estadisticas.cepal.org/cepalstat/WEB_CEPALSTAT/Library/CEPALSTAT/Infograficos_pdf/Inf_10_es.pdf

El economista. (8 de agosto de 2019). *Home* [página web]. www.economis.com.ar

Federación Argentina de Consejos profesionales de Ciencias Económicas: FACPCE. (2022). *Home* [página web]. <https://www.facpce.org.ar/>

Instituto Nacional de Estadística y Censos. (2005). *Censo Nacional Económico 2004/2005: síntesis metodológica*. https://sitioanterior.indec.gob.ar/economico2005/CNE04_metodologia_040810.pdf

Instituto Nacional de Estadística y Censos. (2014). Mesa de Competitividad Sistémica Encuesta Nacional sobre Competitividad e Inversión. https://www.indec.gob.ar/ftp/indecinforma/nuevaWEB/cuadros/16/enci_31_01_14.pdf

Instituto Nacional de Estadística y Censos: INDEC. (2014). *Home* [página web]. <https://www.indec.gob.ar/>

La Nación. (2019). *Home* [página web]. <https://www.lanacion.com.ar/>

Servicio Nacional de Sanidad y Calidad Agroalimentaria: SENASA. (2022). *Home* [página web]. <https://www.argentina.gob.ar/senasa>

Unión de Sindicatos de la Industria Maderera de la República Argentina: USIMRA. (2022). *Home* [página web]. <http://www.usimra.com.ar/>

ANEXO

Matriz 1: Características de 50 alumnos pasantes del Departamento de Ciencias Económicas, seleccionados aleatoriamente

Orden	Género	Edad	Est. civil	Promedio	Transporte	Rapidez	Supervisión	Materias	Orden	Género	Edad	Est. civil	Promedio	Transporte	Rapidez	Supervisión	Materias
1	F	19	S	7,50	C	M	Mod.	3	26	F	20	S	8,45	C	L	Mod.	2
2	F	28	C	7,75	C	R	Mín.	2	27	F	30	D	7,75	C	M	Mín.	1
3	F	21	S	7,25	M	M	Mod.	3	28	M	24	C	8,15	C	M	Mod.	3
4	M	23	S	8,00	C	L	Máx.	1	29	M	23	S	9,00	A	R	Mín.	2
5	F	25	S	7,60	B	R	Mín.	3	30	F	22	S	8,45	A	R	Mín.	4
6	M	22	C	7,85	C	L	Máx.	3	31	F	29	C	8,05	C	M	Mod.	3
7	F	20	S	7,33	C	M	Máx.	3	32	M	23	S	8,00	C	R	Mín.	2
8	F	34	C	8,10	C	M	Mod.	2	33	F	33	C	7,66	A	R	Mín.	3
9	F	36	S	7,75	A	R	Mín.	3	34	F	37	D	8,25	A	L	Mod.	2
10	F	29	S	8,45	C	L	Máx.	3	35	M	28	S	7,55	M	L	Mod.	2
11	M	21	S	9,15	C	M	Mod.	2	36	F	23	C	7,45	C	L	Mod.	3
12	M	22	S	7,05	C	R	Mín.	3	37	F	25	C	7,66	A	L	Mod.	2
13	M	22	S	8,15	C	M	Mod.	4	38	F	31	S	7,15	A	M	Mín.	1
14	M	23	C	8,50	A	M	Mín.	1	39	M	39	C	8,10	C	R	Mín.	3
15	M	19	S	8,65	C	M	Mín.	3	40	M	24	S	8,66	C	L	Máx.	2
16	M	20	S	8,80	T	R	Mín.	2	41	F	20	S	7,50	C	M	Mod.	2
17	M	19	S	9,05	C	R	Mín.	4	42	M	25	S	8,95	C	R	Mín.	3
18	M	21	S	8,66	AP	L	Mod.	3	43	M	22	S	7,28	A	M	Mod.	3
19	M	26	C	7,50	C	R	Mín.	4	44	F	27	C	7,33	A	M	Mín.	1
20	F	21	S	8,55	T	R	Mín.	3	45	F	24	S	7,95	B	M	Mín.	3
21	M	20	S	8,36	A	L	Mod.	3	46	F	23	S	7,66	C	R	Mín.	2
22	M	19	S	8,65	C	M	Mod.	3	47	M	21	S	9,25	C	L	Mod.	2
23	M	21	C	8,05	C	R	Mín.	3	48	F	29	S	7,90	A	M	Mod.	1
24	M	25	C	9,05	C	L	Mod.	0	49	M	27	S	7,85	C	L	Máx.	2
25	M	22	S	7,25	T	M	Mín.	2	50	M	32	C	7,46	A	R	Mín.	1

Principales características para estudiar:

- **Género.** F: femenino ; M: masculino
- **Estado civil.** C: casado; S: soltero; D: divorciado.
- **Transporte** para llegar a la universidad. C: colectivo; A: auto; M: moto; B: bicicleta; AP: a pie; T: tren.
- **Rapidez** en el cumplimiento de las tareas como pasante. L: lento ; M: moderado ; R: rápido
- Nivel de **supervisión** requerida para realizar su labor. Mín.: mínima ; Mod.: moderada ; Máx.: máxima.
- Cantidad media de **materias** que cursa por cuatrimestre.

Matriz 2: Los datos corresponden a una muestra aleatoria de 60 empleados de una empresa de turismo internacional.

Orden	Sexo	Edad	Salario en USD	N.º hijos	Estado civil	Orden	Sexo	Edad	Salario en USD	N.º hijos	Estado civil
1	F	22	1556,21	2	Casado	31	M	43	3029,36	3	Casado
2	F	24	973,70	1	Casado	32	M	40	4400,34	4	Casado
3	F	23	2593,24	0	Soltero	33	M	20	1436,90	0	Soltero
4	M	21	1645,23	1	Soltero	34	F	22	1647,38	0	Soltero
5	M	48	2493,99	5	Casado	35	M	23	2479,21	2	Casado
6	M	26	2103,66	2	Casado	36	M	29	2941,67	2	Casado
7	F	22	3347,67	1	Soltero	37	F	20	1900,63	1	Soltero
8	F	19	957,61	0	Casado	38	F	30	2613,62	1	Divorciado
9	M	30	2679,05	3	Soltero	39	F	28	3462,85	3	Casado
10	M	25	2939,51	2	Soltero	40	M	22	2149,33	0	Casado
11	F	37	3262,70	5	Casado	41	F	22	1965,36	0	Soltero
12	M	22	460,60	0	Soltero	42	M	21	1360,36	0	Soltero
13	M	29	3047,83	4	Casado	43	F	22	1395,84	0	Soltero
14	F	31	2808,90	1	Casado	44	M	23	2321,64	0	Soltero
15	F	31	2472,32	4	Casado	45	M	22	1553,05	1	Casado
16	M	22	935,19	0	Soltero	46	F	23	1471,09	1	Casado
17	F	39	3005,47	7	Casado	47	F	21	1953,56	1	Soltero
18	M	51	3236,89	0	Soltero	48	M	32	2803,12	2	Soltero
19	M	25	3346,70	1	Casado	49	F	33	2243,07	1	Casado
20	M	20	1115,87	0	Soltero	50	M	27	1623,00	1	Soltero
21	M	25	2835,38	2	Divorciado	51	M	34	2559,53	4	Casado
22	F	24	1425,67	1	Soltero	52	M	40	3424,21	5	Casado
23	M	24	2246,43	0	Casado	53	F	36	2738,20	4	Divorciado
24	M	32	2586,76	3	Divorciado	54	F	24	1840,89	0	Soltero
25	F	30	3171,68	5	Casado	55	F	53	4793,92	5	Casado
26	F	24	1546,50	1	Soltero	56	M	35	2938,80	3	Casado
27	M	34	3663,44	4	Casado	57	F	24	1372,10	2	Casado
28	M	22	1937,40	0	Soltero	58	F	20	1664,92	0	Soltero
29	M	31	4010,86	3	Casado	59	F	27	1573,35	3	Casado
30	M	33	3329,85	1	Divorciado	60	M	36	2024,09	4	Casado

APÉNDICE A

Tabla A. Tabla normal estándar $F(z) = P(Z \leq z)$

z	F(z)	z	F(z)	z	F(z)	z	F(z)	z	F(z)	z	F(z)
-∞	0,00000	-3,49	0,00024	-2,97	0,00149	-2,45	0,00714	-1,93	0,02680	-1,41	0,07927
...	0,00000	-3,48	0,00025	-2,96	0,00154	-2,44	0,00734	-1,92	0,02743	-1,4	0,08076
-3,99	0,00003	-3,47	0,00026	-2,95	0,00159	-2,43	0,00755	-1,91	0,02807	-1,39	0,08226
-3,98	0,00003	-3,46	0,00027	-2,94	0,00164	-2,42	0,00776	-1,9	0,02872	-1,38	0,08379
-3,97	0,00004	-3,45	0,00028	-2,93	0,00169	-2,41	0,00798	-1,89	0,02938	-1,37	0,08534
-3,96	0,00004	-3,44	0,00029	-2,92	0,00175	-2,4	0,00820	-1,88	0,03005	-1,36	0,08691
-3,95	0,00004	-3,43	0,00030	-2,91	0,00181	-2,39	0,00842	-1,87	0,03074	-1,35	0,08851
-3,94	0,00004	-3,42	0,00031	-2,9	0,00187	-2,38	0,00866	-1,86	0,03144	-1,34	0,09012
-3,93	0,00004	-3,41	0,00032	-2,89	0,00193	-2,37	0,00889	-1,85	0,03216	-1,33	0,09176
-3,92	0,00004	-3,4	0,00034	-2,88	0,00199	-2,36	0,00914	-1,84	0,03288	-1,32	0,09342
-3,91	0,00005	-3,39	0,00035	-2,87	0,00205	-2,35	0,00939	-1,83	0,03362	-1,31	0,09510
-3,9	0,00005	-3,38	0,00036	-2,86	0,00212	-2,34	0,00964	-1,82	0,03438	-1,3	0,09680
-3,89	0,00005	-3,37	0,00038	-2,85	0,00219	-2,33	0,00990	-1,81	0,03515	-1,29	0,09853
-3,88	0,00005	-3,36	0,00039	-2,84	0,00226	-2,32	0,01017	-1,8	0,03593	-1,28	0,10027
-3,87	0,00005	-3,35	0,00040	-2,83	0,00233	-2,31	0,01044	-1,79	0,03673	-1,27	0,10204
-3,86	0,00006	-3,34	0,00042	-2,82	0,00240	-2,3	0,01072	-1,78	0,03754	-1,26	0,10383
-3,85	0,00006	-3,33	0,00043	-2,81	0,00248	-2,29	0,01101	-1,77	0,03836	-1,25	0,10565
-3,84	0,00006	-3,32	0,00045	-2,8	0,00256	-2,28	0,01130	-1,76	0,03920	-1,24	0,10749
-3,83	0,00006	-3,31	0,00047	-2,79	0,00264	-2,27	0,01160	-1,75	0,04006	-1,23	0,10935
-3,82	0,00007	-3,3	0,00048	-2,78	0,00272	-2,26	0,01191	-1,74	0,04093	-1,22	0,11123
-3,81	0,00007	-3,29	0,00050	-2,77	0,00280	-2,25	0,01222	-1,73	0,04182	-1,21	0,11314
-3,8	0,00007	-3,28	0,00052	-2,76	0,00289	-2,24	0,01255	-1,72	0,04272	-1,2	0,11507
-3,79	0,00008	-3,27	0,00054	-2,75	0,00298	-2,23	0,01287	-1,71	0,04363	-1,19	0,11702
-3,78	0,00008	-3,26	0,00056	-2,74	0,00307	-2,22	0,01321	-1,7	0,04457	-1,18	0,11900
-3,77	0,00008	-3,25	0,00058	-2,73	0,00317	-2,21	0,01355	-1,69	0,04551	-1,17	0,12100
-3,76	0,00008	-3,24	0,00060	-2,72	0,00326	-2,2	0,01390	-1,68	0,04648	-1,16	0,12302
-3,75	0,00009	-3,23	0,00062	-2,71	0,00336	-2,19	0,01426	-1,67	0,04746	-1,15	0,12507
-3,74	0,00009	-3,22	0,00064	-2,7	0,00347	-2,18	0,01463	-1,66	0,04846	-1,14	0,12714
-3,73	0,00010	-3,21	0,00066	-2,69	0,00357	-2,17	0,01500	-1,65	0,04947	-1,13	0,12924
-3,72	0,00010	-3,2	0,00069	-2,68	0,00368	-2,16	0,01539	-1,64	0,05050	-1,12	0,13136
-3,71	0,00010	-3,19	0,00071	-2,67	0,00379	-2,15	0,01578	-1,63	0,05155	-1,11	0,13350
-3,7	0,00011	-3,18	0,00074	-2,66	0,00391	-2,14	0,01618	-1,62	0,05262	-1,1	0,13567
-3,69	0,00011	-3,17	0,00076	-2,65	0,00402	-2,13	0,01659	-1,61	0,05370	-1,09	0,13786
-3,68	0,00012	-3,16	0,00079	-2,64	0,00415	-2,12	0,01700	-1,6	0,05480	-1,08	0,14007
-3,67	0,00012	-3,15	0,00082	-2,63	0,00427	-2,11	0,01743	-1,59	0,05592	-1,07	0,14231
-3,66	0,00013	-3,14	0,00084	-2,62	0,00440	-2,1	0,01786	-1,58	0,05705	-1,06	0,14457
-3,65	0,00013	-3,13	0,00087	-2,61	0,00453	-2,09	0,01831	-1,57	0,05821	-1,05	0,14686
-3,64	0,00014	-3,12	0,00090	-2,6	0,00466	-2,08	0,01876	-1,56	0,05938	-1,04	0,14917
-3,63	0,00014	-3,11	0,00094	-2,59	0,00480	-2,07	0,01923	-1,55	0,06057	-1,03	0,15151
-3,62	0,00015	-3,1	0,00097	-2,58	0,00494	-2,06	0,01970	-1,54	0,06178	-1,02	0,15386
-3,61	0,00015	-3,09	0,00100	-2,57	0,00508	-2,05	0,02018	-1,53	0,06301	-1,01	0,15625
-3,6	0,00016	-3,08	0,00104	-2,56	0,00523	-2,04	0,02068	-1,52	0,06426	-1	0,15866
-3,59	0,00017	-3,07	0,00107	-2,55	0,00539	-2,03	0,02118	-1,51	0,06552	-0,99	0,16109
-3,58	0,00017	-3,06	0,00111	-2,54	0,00554	-2,02	0,02169	-1,5	0,06681	-0,98	0,16354
-3,57	0,00018	-3,05	0,00114	-2,53	0,00570	-2,01	0,02222	-1,49	0,06811	-0,97	0,16602
-3,56	0,00019	-3,04	0,00118	-2,52	0,00587	-2	0,02275	-1,48	0,06944	-0,96	0,16853
-3,55	0,00019	-3,03	0,00122	-2,51	0,00604	-1,99	0,02330	-1,47	0,07078	-0,95	0,17106
-3,54	0,00020	-3,02	0,00126	-2,5	0,00621	-1,98	0,02385	-1,46	0,07215	-0,94	0,17361
-3,53	0,00021	-3,01	0,00131	-2,49	0,00639	-1,97	0,02442	-1,45	0,07353	-0,93	0,17619
-3,52	0,00022	-3	0,00135	-2,48	0,00657	-1,96	0,02500	-1,44	0,07493	-0,92	0,17879
-3,51	0,00022	-2,99	0,00139	-2,47	0,00676	-1,95	0,02559	-1,43	0,07636	-0,91	0,18141
-3,5	0,00023	-2,98	0,00144	-2,46	0,00695	-1,94	0,02619	-1,42	0,07780	-0,9	0,18406

-0,89	0,18673	-0,34	0,36693	0,21	0,58317	0,76	0,77637	1,31	0,90490	1,86	0,96856
-0,88	0,18943	-0,33	0,37070	0,22	0,58706	0,77	0,77935	1,32	0,90658	1,87	0,96926
-0,87	0,19215	-0,32	0,37448	0,23	0,59095	0,78	0,78230	1,33	0,90824	1,88	0,96995
-0,86	0,19489	-0,31	0,37828	0,24	0,59483	0,79	0,78524	1,34	0,90988	1,89	0,97062
-0,85	0,19766	-0,3	0,38209	0,25	0,59871	0,8	0,78814	1,35	0,91149	1,9	0,97128
-0,84	0,20045	-0,29	0,38591	0,26	0,60257	0,81	0,79103	1,36	0,91309	1,91	0,97193
-0,83	0,20327	-0,28	0,38974	0,27	0,60642	0,82	0,79389	1,37	0,91466	1,92	0,97257
-0,82	0,20611	-0,27	0,39358	0,28	0,61026	0,83	0,79673	1,38	0,91621	1,93	0,9732
-0,81	0,20897	-0,26	0,39743	0,29	0,61409	0,84	0,79955	1,39	0,91774	1,94	0,97381
-0,8	0,21186	-0,25	0,40129	0,3	0,61791	0,85	0,80234	1,4	0,91924	1,95	0,97441
-0,79	0,21476	-0,24	0,40517	0,31	0,62172	0,86	0,80511	1,41	0,92073	1,96	0,97500
-0,78	0,21770	-0,23	0,40905	0,32	0,62552	0,87	0,80785	1,42	0,92220	1,97	0,97558
-0,77	0,22065	-0,22	0,41294	0,33	0,62930	0,88	0,81057	1,43	0,92364	1,98	0,97615
-0,76	0,22363	-0,21	0,41683	0,34	0,63307	0,89	0,81327	1,44	0,92507	1,99	0,97670
-0,75	0,22663	-0,2	0,42074	0,35	0,63683	0,9	0,81594	1,45	0,92647	2	0,97725
-0,74	0,22965	-0,19	0,42465	0,36	0,64058	0,91	0,81859	1,46	0,92785	2,01	0,97778
-0,73	0,23270	-0,18	0,42858	0,37	0,64431	0,92	0,82121	1,47	0,92922	2,02	0,97831
-0,72	0,23576	-0,17	0,43251	0,38	0,64803	0,93	0,82381	1,48	0,93056	2,03	0,97882
-0,71	0,23885	-0,16	0,43644	0,39	0,65173	0,94	0,82639	1,49	0,93189	2,04	0,97932
-0,7	0,24196	-0,15	0,44038	0,4	0,65542	0,95	0,82894	1,5	0,93319	2,05	0,97982
-0,69	0,24510	-0,14	0,44433	0,41	0,65910	0,96	0,83147	1,51	0,93448	2,06	0,98030
-0,68	0,24825	-0,13	0,44828	0,42	0,66276	0,97	0,83398	1,52	0,93574	2,07	0,98077
-0,67	0,25143	-0,12	0,45224	0,43	0,66640	0,98	0,83646	1,53	0,93699	2,08	0,98124
-0,66	0,25463	-0,11	0,45620	0,44	0,67003	0,99	0,83891	1,54	0,93822	2,09	0,98169
-0,65	0,25785	-0,1	0,46017	0,45	0,67364	1	0,84134	1,55	0,93943	2,1	0,98214
-0,64	0,26109	-0,09	0,46414	0,46	0,67724	1,01	0,84375	1,56	0,94062	2,11	0,98257
-0,63	0,26435	-0,08	0,46812	0,47	0,68082	1,02	0,84614	1,57	0,94179	2,12	0,98300
-0,62	0,26763	-0,07	0,47210	0,48	0,68439	1,03	0,84849	1,58	0,94295	2,13	0,98341
-0,61	0,27093	-0,06	0,47608	0,49	0,68793	1,04	0,85083	1,59	0,94408	2,14	0,98382
-0,6	0,27425	-0,05	0,48006	0,5	0,69146	1,05	0,85314	1,6	0,94520	2,15	0,98422
-0,59	0,27760	-0,04	0,48405	0,51	0,69497	1,06	0,85543	1,61	0,94630	2,16	0,98461
-0,58	0,28096	-0,03	0,48803	0,52	0,69847	1,07	0,85769	1,62	0,94738	2,17	0,98500
-0,57	0,28434	-0,02	0,49202	0,53	0,70194	1,08	0,85993	1,63	0,94845	2,18	0,98537
-0,56	0,28774	-0,01	0,49601	0,54	0,70540	1,09	0,86214	1,64	0,94950	2,19	0,98574
-0,55	0,29116	0	0,50000	0,55	0,70884	1,1	0,86433	1,65	0,95053	2,2	0,98610
-0,54	0,29460	0,01	0,50399	0,56	0,71226	1,11	0,86650	1,66	0,95154	2,21	0,98645
-0,53	0,29806	0,02	0,50798	0,57	0,71566	1,12	0,86864	1,67	0,95254	2,22	0,98679
-0,52	0,30153	0,03	0,51197	0,58	0,71904	1,13	0,87076	1,68	0,95352	2,23	0,98713
-0,51	0,30503	0,04	0,51595	0,59	0,72240	1,14	0,87286	1,69	0,95449	2,24	0,98745
-0,5	0,30854	0,05	0,51994	0,6	0,72575	1,15	0,87493	1,7	0,95543	2,25	0,98778
-0,49	0,31207	0,06	0,52392	0,61	0,72907	1,16	0,87698	1,71	0,95637	2,26	0,98809
-0,48	0,31561	0,07	0,5279	0,62	0,73237	1,17	0,87900	1,72	0,95728	2,27	0,98840
-0,47	0,31918	0,08	0,53188	0,63	0,73565	1,18	0,88100	1,73	0,95818	2,28	0,98870
-0,46	0,32276	0,09	0,53586	0,64	0,73891	1,19	0,88298	1,74	0,95907	2,29	0,98899
-0,45	0,32636	0,1	0,53983	0,65	0,74215	1,2	0,88493	1,75	0,95994	2,3	0,98928
-0,44	0,32997	0,11	0,54380	0,66	0,74537	1,21	0,88686	1,76	0,96080	2,31	0,98956
-0,43	0,33360	0,12	0,54776	0,67	0,74857	1,22	0,88877	1,77	0,96164	2,32	0,98983
-0,42	0,33724	0,13	0,55172	0,68	0,75175	1,23	0,89065	1,78	0,96246	2,33	0,99010
-0,41	0,34090	0,14	0,55567	0,69	0,75490	1,24	0,89251	1,79	0,96327	2,34	0,99036
-0,4	0,34458	0,15	0,55962	0,7	0,75804	1,25	0,89435	1,8	0,96407	2,35	0,99061
-0,39	0,34827	0,16	0,56356	0,71	0,76115	1,26	0,89617	1,81	0,96485	2,36	0,99086
-0,38	0,35197	0,17	0,56749	0,72	0,76424	1,27	0,89796	1,82	0,96562	2,37	0,99111
-0,37	0,35569	0,18	0,57142	0,73	0,76730	1,28	0,89973	1,83	0,96638	2,38	0,99134
-0,36	0,35942	0,19	0,57535	0,74	0,77035	1,29	0,90147	1,84	0,96712	2,39	0,99158
-0,35	0,36317	0,2	0,57926	0,75	0,77337	1,3	0,90320	1,85	0,96784	2,4	0,99180

2,41	0,99202	2,68	0,99632	2,95	0,99841	3,21	0,99934	3,48	0,99975	3,75	0,99991
2,42	0,99224	2,69	0,99643	2,96	0,99846	3,22	0,99936	3,49	0,99976	3,76	0,99992

2,43	0,99245	2,7	0,99653	2,97	0,99851	3,23	0,99938	3,5	0,99977	3,77	0,99992
2,44	0,99266	2,71	0,99664	2,98	0,99856	3,24	0,99940	3,51	0,99978	3,78	0,99992
2,45	0,99286	2,72	0,99674	2,99	0,99861	3,25	0,99942	3,52	0,99978	3,79	0,99992
2,46	0,99305	2,73	0,99683	3	0,99865	3,26	0,99944	3,53	0,99979	3,8	0,99993
2,47	0,99324	2,74	0,99693	3,01	0,99869	3,27	0,99946	3,54	0,99980	3,81	0,99993
2,48	0,99343	2,75	0,99702	3,02	0,99874	3,28	0,99948	3,55	0,99981	3,82	0,99993
2,49	0,99361	2,76	0,99711	3,03	0,99878	3,29	0,99950	3,56	0,99981	3,83	0,99994
2,5	0,99379	2,77	0,99720	2,94	0,99836	3,3	0,99952	3,57	0,99982	3,84	0,99994
2,51	0,99396	2,78	0,99728	3,04	0,99882	3,31	0,99953	3,58	0,99983	3,85	0,99994
2,52	0,99413	2,79	0,99736	3,05	0,99886	3,32	0,99955	3,59	0,99983	3,86	0,99994
2,53	0,99430	2,8	0,99744	3,06	0,99889	3,33	0,99957	3,6	0,99984	3,87	0,99995
2,54	0,99446	2,81	0,99752	3,07	0,99893	3,34	0,99958	3,61	0,99985	3,88	0,99995
2,55	0,99461	2,82	0,99760	3,08	0,99896	3,35	0,99960	3,62	0,99985	3,89	0,99995
2,56	0,99477	2,83	0,99767	3,09	0,99900	3,36	0,99961	3,63	0,99986	3,9	0,99995
2,57	0,99492	2,84	0,99774	3,1	0,99903	3,37	0,99962	3,64	0,99986	3,91	0,99995
2,58	0,99506	2,85	0,99781	3,11	0,99906	3,38	0,99964	3,65	0,99987	3,92	0,99996
2,59	0,99520	2,86	0,99788	3,12	0,99910	3,39	0,99965	3,66	0,99987	3,93	0,99996
2,6	0,99534	2,87	0,99795	3,13	0,99913	3,4	0,99966	3,67	0,99988	3,94	0,99996
2,61	0,99547	2,88	0,99801	3,14	0,99916	3,41	0,99968	3,68	0,99988	3,95	0,99996
2,62	0,99560	2,89	0,99807	3,15	0,99918	3,42	0,99969	3,69	0,99989	3,96	0,99996
2,63	0,99573	2,9	0,99813	3,16	0,99921	3,43	0,99970	3,7	0,99989	3,97	0,99996
2,64	0,99585	2,91	0,99819	3,17	0,99924	3,44	0,99971	3,71	0,99990	3,98	0,99997
2,65	0,99598	2,92	0,99825	3,18	0,99926	3,45	0,99972	3,72	0,99990	3,99	0,99997
2,66	0,99609	2,93	0,99831	3,19	0,99929	3,46	0,99973	3,73	0,99990	...	0,99999
2,67	0,99621	2,94	0,99836	3,2	0,99931	3,47	0,99974	3,74	0,99991	+∞	1,00000

APÉNDICE B

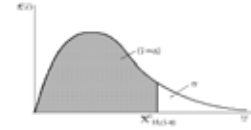


Tabla B. Tabla Chi cuadrado para $F(x)=1-\alpha$

g.l.	0,005	0,01	0,025	0,05	0,1	0,2	0,3	0,5	0,7	0,8	0,9	0,95	0,975	0,99	0,995
1	0	0	0	0	0,02	0,06	0,15	0,45	1,07	1,64	2,71	3,84	5,02	6,63	7,88
2	0,01	0,02	0,05	0,1	0,21	0,45	0,71	1,39	2,41	3,22	4,61	5,99	7,38	9,21	10,6
3	0,07	0,11	0,22	0,35	0,58	1,01	1,42	2,37	3,66	4,64	6,25	7,81	9,35	11,34	12,84
4	0,21	0,3	0,48	0,71	1,06	1,65	2,19	3,36	4,88	5,99	7,78	9,49	11,14	13,28	14,86
5	0,41	0,55	0,83	1,15	1,61	2,34	3	4,35	6,06	7,29	9,24	11,07	12,83	15,09	16,75
6	0,68	0,87	1,24	1,64	2,2	3,07	3,83	5,35	7,23	8,56	10,64	12,59	14,45	16,81	18,55
7	0,99	1,24	1,69	2,17	2,83	3,82	4,67	6,35	8,38	9,8	12,02	14,07	16,01	18,48	20,28
8	1,34	1,65	2,18	2,73	3,49	4,59	5,53	7,34	9,52	11,03	13,36	15,51	17,53	20,09	21,95
9	1,73	2,09	2,7	3,33	4,17	5,38	6,39	8,34	10,66	12,24	14,68	16,92	19,02	21,67	23,59
10	2,16	2,56	3,25	3,94	4,87	6,18	7,27	9,34	11,78	13,44	15,99	18,31	20,48	23,21	25,19
11	2,6	3,05	3,82	4,57	5,58	6,99	8,15	10,34	12,9	14,63	17,28	19,68	21,92	24,72	26,76
12	3,07	3,57	4,4	5,23	6,3	7,81	9,03	11,34	14,01	15,81	18,55	21,03	23,34	26,22	28,3
13	3,57	4,11	5,01	5,89	7,04	8,63	9,93	12,34	15,12	16,98	19,81	22,36	24,74	27,69	29,82
14	4,07	4,66	5,63	6,57	7,79	9,47	10,82	13,34	16,22	18,15	21,06	23,68	26,12	29,14	31,32
15	4,6	5,23	6,26	7,26	8,55	10,31	11,72	14,34	17,32	19,31	22,31	25	27,49	30,58	32,8
16	5,14	5,81	6,91	7,96	9,31	11,15	12,62	15,34	18,42	20,47	23,54	26,3	28,85	32	34,27
17	5,7	6,41	7,56	8,67	10,09	12	13,53	16,34	19,51	21,61	24,77	27,59	30,19	33,41	35,72
18	6,26	7,01	8,23	9,39	10,86	12,86	14,44	17,34	20,6	22,76	25,99	28,87	31,53	34,81	37,16
19	6,84	7,63	8,91	10,12	11,65	13,72	15,35	18,34	21,69	23,9	27,2	30,14	32,85	36,19	38,58
20	7,43	8,26	9,59	10,85	12,44	14,58	16,27	19,34	22,77	25,04	28,41	31,41	34,17	37,57	40
21	8,03	8,9	10,28	11,59	13,24	15,44	17,18	20,34	23,86	26,17	29,62	32,67	35,48	38,93	41,4
22	8,64	9,54	10,98	12,34	14,04	16,31	18,1	21,34	24,94	27,3	30,81	33,92	36,78	40,29	42,8
23	9,26	10,2	11,69	13,09	14,85	17,19	19,02	22,34	26,02	28,43	32,01	35,17	38,08	41,64	44,18
24	9,89	10,86	12,4	13,85	15,66	18,06	19,94	23,34	27,1	29,55	33,2	36,42	39,36	42,98	45,56
25	10,52	11,52	13,12	14,61	16,47	18,94	20,87	24,34	28,17	30,68	34,38	37,65	40,65	44,31	46,93
26	11,16	12,2	13,84	15,38	17,29	19,82	21,79	25,34	29,25	31,79	35,56	38,89	41,92	45,64	48,29
27	11,81	12,88	14,57	16,15	18,11	20,7	22,72	26,34	30,32	32,91	36,74	40,11	43,19	46,96	49,64
28	12,46	13,56	15,31	16,93	18,94	21,59	23,65	27,34	31,39	34,03	37,92	41,34	44,46	48,28	50,99
29	13,12	14,26	16,05	17,71	19,77	22,48	24,58	28,34	32,46	35,14	39,09	42,56	45,72	49,59	52,34
30	13,79	14,95	16,79	18,49	20,6	23,36	25,51	29,34	33,53	36,25	40,26	43,77	46,98	50,89	53,67
31	14,46	15,66	17,54	19,28	21,43	24,26	26,44	30,34	34,6	37,36	41,42	44,99	48,23	52,19	55
32	15,13	16,36	18,29	20,07	22,27	25,15	27,37	31,34	35,66	38,47	42,58	46,19	49,48	53,49	56,33
33	15,82	17,07	19,05	20,87	23,11	26,04	28,31	32,34	36,73	39,57	43,75	47,4	50,73	54,78	57,65
34	16,5	17,79	19,81	21,66	23,95	26,94	29,24	33,34	37,8	40,68	44,9	48,6	51,97	56,06	58,96
35	17,19	18,51	20,57	22,47	24,8	27,84	30,18	34,34	38,86	41,78	46,06	49,8	53,2	57,34	60,27
36	17,89	19,23	21,34	23,27	25,64	28,73	31,12	35,34	39,92	42,88	47,21	51	54,44	58,62	61,58
37	18,59	19,96	22,11	24,07	26,49	29,64	32,05	36,34	40,98	43,98	48,36	52,19	55,67	59,89	62,88
38	19,29	20,69	22,88	24,88	27,34	30,54	32,99	37,34	42,05	45,08	49,51	53,38	56,9	61,16	64,18
39	20	21,43	23,65	25,7	28,2	31,44	33,93	38,34	43,11	46,17	50,66	54,57	58,12	62,43	65,48
40	20,71	22,16	24,43	26,51	29,05	32,34	34,87	39,34	44,16	47,27	51,81	55,76	59,34	63,69	66,77

41	21,42	22,91	25,21	27,33	29,91	33,25	35,81	40,34	45,22	48,36	52,95	56,94	60,56	64,95	68,05
42	22,14	23,65	26	28,14	30,77	34,16	36,75	41,34	46,28	49,46	54,09	58,12	61,78	66,21	69,34
43	22,86	24,4	26,79	28,96	31,63	35,07	37,7	42,34	47,34	50,55	55,23	59,3	62,99	67,46	70,62
44	23,58	25,15	27,57	29,79	32,49	35,97	38,64	43,34	48,4	51,64	56,37	60,48	64,2	68,71	71,89
45	24,31	25,9	28,37	30,61	33,35	36,88	39,58	44,34	49,45	52,73	57,51	61,66	65,41	69,96	73,17
46	25,04	26,66	29,16	31,44	34,22	37,8	40,53	45,34	50,51	53,82	58,64	62,83	66,62	71,2	74,44
47	25,77	27,42	29,96	32,27	35,08	38,71	41,47	46,34	51,56	54,91	59,77	64	67,82	72,44	75,7
48	26,51	28,18	30,75	33,1	35,95	39,62	42,42	47,34	52,62	55,99	60,91	65,17	69,02	73,68	76,97
49	27,25	28,94	31,55	33,93	36,82	40,53	43,37	48,33	53,67	57,08	62,04	66,34	70,22	74,92	78,23
50	27,99	29,71	32,36	34,76	37,69	41,45	44,31	49,33	54,72	58,16	63,17	67,5	71,42	76,15	79,49
51	28,73	30,48	33,16	35,6	38,56	42,36	45,26	50,33	55,78	59,25	64,3	68,67	72,62	77,39	80,75
52	29,48	31,25	33,97	36,44	39,43	43,28	46,21	51,33	56,83	60,33	65,42	69,83	73,81	78,62	82
53	30,23	32,02	34,78	37,28	40,31	44,2	47,16	52,33	57,88	61,41	66,55	70,99	75	79,84	83,25
54	30,98	32,79	35,59	38,12	41,18	45,12	48,11	53,33	58,93	62,5	67,67	72,15	76,19	81,07	84,5
55	31,73	33,57	36,4	38,96	42,06	46,04	49,06	54,33	59,98	63,58	68,8	73,31	77,38	82,29	85,75
56	32,49	34,35	37,21	39,8	42,94	46,96	50,01	55,33	61,03	64,66	69,92	74,47	78,57	83,51	86,99
57	33,25	35,13	38,03	40,65	43,82	47,88	50,96	56,33	62,08	65,74	71,04	75,62	79,75	84,73	88,24
58	34,01	35,91	38,84	41,49	44,7	48,8	51,91	57,33	63,13	66,82	72,16	76,78	80,94	85,95	89,48
59	34,77	36,7	39,66	42,34	45,58	49,72	52,86	58,33	64,18	67,89	73,28	77,93	82,12	87,17	90,72
60	35,53	37,48	40,48	43,19	46,46	50,64	53,81	59,33	65,23	68,97	74,4	79,08	83,3	88,38	91,95
61	36,3	38,27	41,3	44,04	47,34	51,56	54,76	60,33	66,27	70,05	75,51	80,23	84,48	89,59	93,19
62	37,07	39,06	42,13	44,89	48,23	52,49	55,71	61,33	67,32	71,13	76,63	81,38	85,65	90,8	94,42
63	37,84	39,86	42,95	45,74	49,11	53,41	56,67	62,33	68,37	72,2	77,75	82,53	86,83	92,01	95,65
64	38,61	40,65	43,78	46,59	50	54,34	57,62	63,33	69,42	73,28	78,86	83,68	88	93,22	96,88
65	39,38	41,44	44,6	47,45	50,88	55,26	58,57	64,33	70,46	74,35	79,97	84,82	89,18	94,42	98,11
66	40,16	42,24	45,43	48,31	51,77	56,19	59,53	65,33	71,51	75,42	81,09	85,96	90,35	95,63	99,33
67	40,94	43,04	46,26	49,16	52,66	57,11	60,48	66,33	72,55	76,5	82,2	87,11	91,52	96,83	100,55
68	41,71	43,84	47,09	50,02	53,55	58,04	61,44	67,33	73,6	77,57	83,31	88,25	92,69	98,03	101,78
69	42,49	44,64	47,92	50,88	54,44	58,97	62,39	68,33	74,64	78,64	84,42	89,39	93,86	99,23	103
70	43,28	45,44	48,76	51,74	55,33	59,9	63,35	69,33	75,69	79,71	85,53	90,53	95,02	100,43	104,21
71	44,06	46,25	49,59	52,6	56,22	60,83	64,3	70,33	76,73	80,79	86,64	91,67	96,19	101,62	105,43
72	44,84	47,05	50,43	53,46	57,11	61,76	65,26	71,33	77,78	81,86	87,74	92,81	97,35	102,82	106,65
73	45,63	47,86	51,26	54,33	58,01	62,69	66,21	72,33	78,82	82,93	88,85	93,95	98,52	104,01	107,86
74	46,42	48,67	52,1	55,19	58,9	63,62	67,17	73,33	79,86	84	89,96	95,08	99,68	105,2	109,07
75	47,21	49,48	52,94	56,05	59,79	64,55	68,13	74,33	80,91	85,07	91,06	96,22	100,84	106,39	110,29
76	48	50,29	53,78	56,92	60,69	65,48	69,08	75,33	81,95	86,13	92,17	97,35	102	107,58	111,5
77	48,79	51,1	54,62	57,79	61,59	66,41	70,04	76,33	82,99	87,2	93,27	98,48	103,16	108,77	112,7
78	49,58	51,91	55,47	58,65	62,48	67,34	71	77,33	84,04	88,27	94,37	99,62	104,32	109,96	113,91
79	50,38	52,72	56,31	59,52	63,38	68,27	71,96	78,33	85,08	89,34	95,48	100,75	105,47	111,14	115,12
80	51,17	53,54	57,15	60,39	64,28	69,21	72,92	79,33	86,12	90,41	96,58	101,88	106,63	112,33	116,32
81	51,97	54,36	58	61,26	65,18	70,14	73,87	80,33	87,16	91,47	97,68	103,01	107,78	113,51	117,52
82	52,77	55,17	58,84	62,13	66,08	71,07	74,83	81,33	88,2	92,54	98,78	104,14	108,94	114,69	118,73
83	53,57	55,99	59,69	63	66,98	72,01	75,79	82,33	89,24	93,6	99,88	105,27	110,09	115,88	119,93
84	54,37	56,81	60,54	63,88	67,88	72,94	76,75	83,33	90,28	94,67	100,98	106,39	111,24	117,06	121,13
85	55,17	57,63	61,39	64,75	68,78	73,88	77,71	84,33	91,32	95,73	102,08	107,52	112,39	118,24	122,32
86	55,97	58,46	62,24	65,62	69,68	74,81	78,67	85,33	92,37	96,8	103,18	108,65	113,54	119,41	123,52

APÉNDICE C

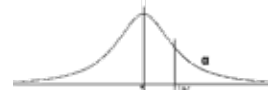


Tabla C. Tabla T-Student para $F(t) = 1 - \alpha$

g.l.	0,001	0,005	0,01	0,025	0,05	0,1	0,2	0,8	0,9	0,95	0,975	0,99	0,995	0,999	g.l.
1	-318,31	-63,66	-31,82	-12,71	-6,31	-3,08	-1,38	1,38	3,08	6,31	12,71	31,82	63,66	318,31	1
2	-22,33	-9,92	-6,96	-4,30	-2,92	-1,89	-1,06	1,06	1,89	2,92	4,30	6,96	9,92	22,33	2
3	-10,21	-5,84	-4,54	-3,18	-2,35	-1,64	-0,98	0,98	1,64	2,35	3,18	4,54	5,84	10,21	3
4	-7,17	-4,60	-3,75	-2,78	-2,13	-1,53	-0,94	0,94	1,53	2,13	2,78	3,75	4,60	7,17	4
5	-5,89	-4,03	-3,36	-2,57	-2,02	-1,48	-0,92	0,92	1,48	2,02	2,57	3,36	4,03	5,89	5
6	-5,21	-3,71	-3,14	-2,45	-1,94	-1,44	-0,91	0,91	1,44	1,94	2,45	3,14	3,71	5,21	6
7	-4,79	-3,50	-3,00	-2,36	-1,89	-1,41	-0,90	0,90	1,41	1,89	2,36	3,00	3,50	4,79	7
8	-4,50	-3,36	-2,90	-2,31	-1,86	-1,40	-0,89	0,89	1,40	1,86	2,31	2,90	3,36	4,50	8
9	-4,30	-3,25	-2,82	-2,26	-1,83	-1,38	-0,88	0,88	1,38	1,83	2,26	2,82	3,25	4,30	9
10	-4,14	-3,17	-2,76	-2,23	-1,81	-1,37	-0,88	0,88	1,37	1,81	2,23	2,76	3,17	4,14	10
11	-4,02	-3,11	-2,72	-2,20	-1,80	-1,36	-0,88	0,88	1,36	1,80	2,20	2,72	3,11	4,02	11
12	-3,93	-3,05	-2,68	-2,18	-1,78	-1,36	-0,87	0,87	1,36	1,78	2,18	2,68	3,05	3,93	12
13	-3,85	-3,01	-2,65	-2,16	-1,77	-1,35	-0,87	0,87	1,35	1,77	2,16	2,65	3,01	3,85	13
14	-3,79	-2,98	-2,62	-2,14	-1,76	-1,35	-0,87	0,87	1,35	1,76	2,14	2,62	2,98	3,79	14
15	-3,73	-2,95	-2,60	-2,13	-1,75	-1,34	-0,87	0,87	1,34	1,75	2,13	2,60	2,95	3,73	15
16	-3,69	-2,92	-2,58	-2,12	-1,75	-1,34	-0,86	0,86	1,34	1,75	2,12	2,58	2,92	3,69	16
17	-3,65	-2,90	-2,57	-2,11	-1,74	-1,33	-0,86	0,86	1,33	1,74	2,11	2,57	2,90	3,65	17
18	-3,61	-2,88	-2,55	-2,10	-1,73	-1,33	-0,86	0,86	1,33	1,73	2,10	2,55	2,88	3,61	18
19	-3,58	-2,86	-2,54	-2,09	-1,73	-1,33	-0,86	0,86	1,33	1,73	2,09	2,54	2,86	3,58	19
20	-3,55	-2,85	-2,53	-2,09	-1,72	-1,33	-0,86	0,86	1,33	1,72	2,09	2,53	2,85	3,55	20
21	-3,53	-2,83	-2,52	-2,08	-1,72	-1,32	-0,86	0,86	1,32	1,72	2,08	2,52	2,83	3,53	21
22	-3,50	-2,82	-2,51	-2,07	-1,72	-1,32	-0,86	0,86	1,32	1,72	2,07	2,51	2,82	3,50	22
23	-3,48	-2,81	-2,50	-2,07	-1,71	-1,32	-0,86	0,86	1,32	1,71	2,07	2,50	2,81	3,48	23
24	-3,47	-2,80	-2,49	-2,06	-1,71	-1,32	-0,86	0,86	1,32	1,71	2,06	2,49	2,80	3,47	24
25	-3,45	-2,79	-2,49	-2,06	-1,71	-1,32	-0,86	0,86	1,32	1,71	2,06	2,49	2,79	3,45	25
26	-3,43	-2,78	-2,48	-2,06	-1,71	-1,31	-0,86	0,86	1,31	1,71	2,06	2,48	2,78	3,43	26
27	-3,42	-2,77	-2,47	-2,05	-1,70	-1,31	-0,86	0,86	1,31	1,70	2,05	2,47	2,77	3,42	27
28	-3,41	-2,76	-2,47	-2,05	-1,70	-1,31	-0,85	0,85	1,31	1,70	2,05	2,47	2,76	3,41	28
29	-3,40	-2,76	-2,46	-2,05	-1,70	-1,31	-0,85	0,85	1,31	1,70	2,05	2,46	2,76	3,40	29
30	-3,39	-2,75	-2,46	-2,04	-1,70	-1,31	-0,85	0,85	1,31	1,70	2,04	2,46	2,75	3,39	30
31	-3,37	-2,74	-2,45	-2,04	-1,70	-1,31	-0,85	0,85	1,31	1,70	2,04	2,45	2,74	3,37	31
32	-3,37	-2,74	-2,45	-2,04	-1,69	-1,31	-0,85	0,85	1,31	1,69	2,04	2,45	2,74	3,37	32
33	-3,36	-2,73	-2,44	-2,03	-1,69	-1,31	-0,85	0,85	1,31	1,69	2,03	2,44	2,73	3,36	33
34	-3,35	-2,73	-2,44	-2,03	-1,69	-1,31	-0,85	0,85	1,31	1,69	2,03	2,44	2,73	3,35	34
35	-3,34	-2,72	-2,44	-2,03	-1,69	-1,31	-0,85	0,85	1,31	1,69	2,03	2,44	2,72	3,34	35
36	-3,33	-2,72	-2,43	-2,03	-1,69	-1,31	-0,85	0,85	1,31	1,69	2,03	2,43	2,72	3,33	36
37	-3,33	-2,72	-2,43	-2,03	-1,69	-1,30	-0,85	0,85	1,30	1,69	2,03	2,43	2,72	3,33	37
38	-3,32	-2,71	-2,43	-2,02	-1,69	-1,30	-0,85	0,85	1,30	1,69	2,02	2,43	2,71	3,32	38
39	-3,31	-2,71	-2,43	-2,02	-1,68	-1,30	-0,85	0,85	1,30	1,68	2,02	2,43	2,71	3,31	39
40	-3,31	-2,70	-2,42	-2,02	-1,68	-1,30	-0,85	0,85	1,30	1,68	2,02	2,42	2,70	3,31	40
41	-3,30	-2,70	-2,42	-2,02	-1,68	-1,30	-0,85	0,85	1,30	1,68	2,02	2,42	2,70	3,30	41
42	-3,30	-2,70	-2,42	-2,02	-1,68	-1,30	-0,85	0,85	1,30	1,68	2,02	2,42	2,70	3,30	42
43	-3,29	-2,70	-2,42	-2,02	-1,68	-1,30	-0,85	0,85	1,30	1,68	2,02	2,42	2,70	3,29	43
44	-3,29	-2,69	-2,41	-2,02	-1,68	-1,30	-0,85	0,85	1,30	1,68	2,02	2,41	2,69	3,29	44
45	-3,28	-2,69	-2,41	-2,01	-1,68	-1,30	-0,85	0,85	1,30	1,68	2,01	2,41	2,69	3,28	45
46	-3,28	-2,69	-2,41	-2,01	-1,68	-1,30	-0,85	0,85	1,30	1,68	2,01	2,41	2,69	3,28	46
47	-3,27	-2,68	-2,41	-2,01	-1,68	-1,30	-0,85	0,85	1,30	1,68	2,01	2,41	2,68	3,27	47
48	-3,27	-2,68	-2,41	-2,01	-1,68	-1,30	-0,85	0,85	1,30	1,68	2,01	2,41	2,68	3,27	48
49	-3,27	-2,68	-2,40	-2,01	-1,68	-1,30	-0,85	0,85	1,30	1,68	2,01	2,40	2,68	3,27	49
50	-3,26	-2,68	-2,40	-2,01	-1,68	-1,30	-0,85	0,85	1,30	1,68	2,01	2,40	2,68	3,26	50
51	-3,26	-2,68	-2,40	-2,01	-1,68	-1,30	-0,85	0,85	1,30	1,68	2,01	2,40	2,68	3,26	51
52	-3,25	-2,67	-2,40	-2,01	-1,67	-1,30	-0,85	0,85	1,30	1,67	2,01	2,40	2,67	3,25	52

g.l.	0,001	0,005	0,01	0,025	0,05	0,1	0,2	0,8	0,9	0,95	0,975	0,99	0,995	0,999	g.l.
53	-3,25	-2,67	-2,40	-2,01	-1,67	-1,30	-0,85	0,85	1,30	1,67	2,01	2,40	2,67	3,25	53
54	-3,25	-2,67	-2,40	-2,00	-1,67	-1,30	-0,85	0,85	1,30	1,67	2,00	2,40	2,67	3,25	54
55	-3,25	-2,67	-2,40	-2,00	-1,67	-1,30	-0,85	0,85	1,30	1,67	2,00	2,40	2,67	3,25	55
56	-3,24	-2,67	-2,39	-2,00	-1,67	-1,30	-0,85	0,85	1,30	1,67	2,00	2,39	2,67	3,24	56
59	-3,23	-2,66	-2,39	-2,00	-1,67	-1,30	-0,85	0,85	1,30	1,67	2,00	2,39	2,66	3,23	59
63	-3,22	-2,66	-2,39	-2,00	-1,67	-1,30	-0,85	0,85	1,30	1,67	2,00	2,39	2,66	3,22	63
68	-3,21	-2,65	-2,38	-2,00	-1,67	-1,29	-0,85	0,85	1,29	1,67	2,00	2,38	2,65	3,21	68
74	-3,20	-2,64	-2,38	-1,99	-1,67	-1,29	-0,85	0,85	1,29	1,67	1,99	2,38	2,64	3,20	74
98	-3,18	-2,63	-2,37	-1,98	-1,66	-1,29	-0,85	0,85	1,29	1,66	1,98	2,37	2,63	3,18	98
99	-3,17	-2,63	-2,36	-1,98	-1,66	-1,29	-0,85	0,85	1,29	1,66	1,98	2,36	2,63	3,17	99