

**Registro de conceptos estadísticos como recurso
para una alfabetización estadística**

55B/142

Directora:

María Eugenia Ángel

Investigadores:

Graciela Fernández

Laura Polola

Mario Enrique Borgna

Liliana Sandra Pagano

Silvia Brunetti

Miriam Ecalte

Informe Final

Diciembre 2010

Registrado en Dirección Nacional de Derecho de Autor
Expediente N° 963.575 el 28/9/2011
Todos los Derechos Reservados SECyT-UNLaM

1- Informe técnico - académico

1.1. Resumen

Con el objetivo de elaborar un registro de términos y conceptos estadísticos, se desarrolla el presente trabajo que consiste en identificar y clasificar conceptos, estableciendo similitudes y diferencias de interpretaciones de acuerdo al contexto de aplicación académico-profesional.

Se trata de elaborar un lenguaje estadístico a través del cuál se puedan resignificar algunos conceptos básicos como así también analizar y describir herramientas estadísticas para su mejor aprovechamiento.

Se ha realizado una revisión y análisis bibliográfico de material estadístico a fin de localizar algunos usos y aplicaciones de diversos términos, así como ubicar posibles interpretaciones erróneas o ambiguas como situaciones donde se podía detectar cierta manipulación de la información.

El registro con los términos seleccionados es el producto final de la presente investigación, elaborado a partir de la confluencia del material procesado con los estándares teóricos vigentes.

Palabras claves: *registro estadístico – herramientas estadísticas – técnicas estadísticas – conceptos estadísticos – significados estadísticos.*

1.2. Introducción

Llegará un día en el que el razonamiento estadístico será tan necesario para el ciudadano como ahora lo es la habilidad de leer y escribir.

H. G. Wells (1866-1946)

En el complicado mundo de hoy pocas cuestiones están claras y no sujetas a controversia. Para entender y poder emitir una opinión sobre un determinado tema se debe recoger información, es decir, contar con datos. Y, para interpretarlos es conveniente poseer conocimientos básicos de Estadística, que es la ciencia que brinda herramientas adecuadas para extraer conclusiones válidas de conjuntos de datos.

La influencia de la estadística como disciplina transversal en las demás ciencias y su utilización creciente más allá del ámbito científico -facilitada y potenciada por las nuevas tecnologías en el tratamiento estadístico de la información- requiere establecer un referente conceptual que permita su buen uso y aplicación; de ahí la necesidad de construir un puente entre la teoría formal y la práctica responsable. Es decir, en la nueva situación sociocultural del mundo actual, se instala la necesidad de una alfabetización estadística orientada a determinar qué conocimientos, valores y habilidades necesita desarrollar el ciudadano de hoy para desempeñarse socialmente con pericia.

Con el presente trabajo se ha iniciado una reestructuración, una renovación de los términos, conceptos y herramientas estadísticas desde lo formal pero sin dejar de lado la función comunicativa del lenguaje que permita remediar en parte su desconocimiento o usos erróneos. El objetivo general de la investigación fue precisamente la elaboración de un registro de términos y conceptos estadísticos donde se unificaran y jerarquizaran las definiciones a partir de las similitudes y diferencias de interpretaciones de acuerdo a su contexto y aplicación. Para lograrlo se identificaron, compilaron, compendiaron, puntualizaron, analizaron, describieron y sintetizaron los términos, conceptos y herramientas estadísticas seleccionadas.

Los estadísticos han invadido todas las ramas de la ciencia con una rapidez de conquista que sólo tiene como rivales a Atila, a Mahoma y al escarabajo de Colorado.

Maurice Kendall (estadístico británico)

Se ha hecho mención de la presencia de la información como un contexto donde el hombre se encuentra inmerso y se hace imperiosa su interpretación, manejo, análisis y tratamiento para sacar conclusiones, pero ¿qué se entiende por información?

Se puede decir que información es todo mensaje oral o escrito posible de impactar sobre la percepción de un sujeto y su estado cognoscitivo. Cuando la persona logra interpretar correctamente un mensaje se dice que se ha comunicado satisfactoriamente, que se ha logrado interpretar correctamente la información, y, su importancia radica en que se reduce la incertidumbre que rodea a la situación que se ha planteado permitiendo una mejor toma de decisiones frente a la misma.

Caben algunos interrogantes, ¿por qué se resalta la situación actual de las sociedades modernas?, ¿esto no ocurría antes?, ¿qué es lo que ha cambiado?

A nuestras sociedades se las denomina Sociedades de la Información porque desde las últimas décadas del siglo XX, la disponibilidad de las nuevas tecnologías de la información y de la

comunicación ha permitido un cambio a nivel global de las economías de los países gracias precisamente a concebirla como un recurso de producción. La información se ha transformado en un medio estratégico, tanto para las empresas como para los gobiernos y los ciudadanos, por eso la importancia de que el trabajador, el estudiante y los profesionales en general se encuentren capacitados para la interpretación y tratamiento de la misma.

*“La información permite construir conocimiento útil para la acción, entendido como la interacción permanente de experiencias, valores, información y saber hacer, que genera, en forma cíclica, nuevas informaciones, experiencias y conocimientos”.*¹

Se ha establecido un nuevo paradigma, que ha generado una mayor demanda de los sistemas de información estadística. Estos sistemas están compuestos por datos provenientes de las más diversas fuentes, integrándolas y relacionándolas para cubrir la diversificación de la demanda.

Por ello se puede decir que todo trabajador y estudiante se ha convertido en un *usuario* de información estadística. La diversidad que presentan estos usuarios obedece a su capacidad de abordar los fenómenos económicos y sociales desde diferentes características, como ser: la magnitud o dimensión de los fenómenos, estructura y forma en que se desagregan sus componentes, distribución espacio- temporal y las relaciones entre fenómenos.

La información estadística si bien se enmarca en un entorno de aproximaciones matemáticas, es la que hace posible el acceso al conocimiento de la realidad, interpretación de la forma en que se comportan los fenómenos, la predicción de su conducta futura siempre y cuando cuente, necesariamente, con los atributos de calidad, veracidad y precisión.

El extraer información de los datos de campo, poder representarla tabular y/o gráficamente, permite darle sentido a esos datos. Lo mismo ocurre al contemplar la realidad desde la perspectiva de su modelización surgiendo la comprensión de las cualidades o características del mundo real.

Debido al caudal de aplicaciones y a la complejidad que puede involucrar el trabajo con la información estadística, se puede prever un incremento de la demanda de herramientas de fácil comprensión y manejo como punto de partida. Dentro de este contexto surgió la necesidad de elaborar un registro de conceptos y procedimientos de diversa complejidad atendiendo a los requerimientos de la sociedad de hoy.

¹ Cuadernos del SIEL Sistema de Información Estadística Local, Serie 1. Materiales institucionales, conceptuales y metodológicos Cultura Estadística y Gestión Local Instituto Nacional de Estadística y Censos Dirección Nacional de Planificación y Coordinación Estadística Julio 2007.

1.3. Desarrollo de la investigación

A partir de la decisión de elaborar el registro de conceptos estadísticos en función de los requerimientos, se realizaron diversos relevamientos basados en la observación y búsqueda de carencias y necesidades de los usuarios en relación al conocimiento estadístico.

Relevamiento de información secundaria y sus resultados

La primera tarea correspondió al tratamiento de información secundaria mediante la recopilación de terminología estadística utilizada en medios de difusión, revistas de divulgación general, páginas de Internet y textos de Estadística y de otras ramas del conocimiento incluyendo textos escolares.

El trabajo se centró en seleccionar, registrar y analizar información referida a la utilización de términos y conceptos estadísticos.

- ***Revisión de libros de texto escolares***

Una de las primeras acciones fue trabajar con los textos escolares de los últimos años del nivel primario y primeros del secundario debido a que es en esos niveles de escolaridad donde se realiza el primer acercamiento a la estadística.

A continuación se mencionan las observaciones realizadas a diferentes textos donde aparecen algunos conceptos de manera incompleta o parcial o conceptualizaciones gemelas que, por un exceso de simplificación, se identifican como iguales cuando no lo son. Las mismas fueron organizadas en bloques teniendo en cuenta la fuente de donde fueron extraídas la cual se cita al final de cada uno de ellos.

- ✓ El concepto de *suceso* aparece relacionado con cada resultado posible de un experimento aleatorio siendo muy restrictiva la definición, ya que no se menciona en ningún momento que un suceso puede tener más de un elemento. Se identifica el concepto de suceso en términos generales con el de suceso elemental.
- ✓ Se define *probabilidad* (siguiendo la definición clásica de Laplace) como el cociente entre el número de casos favorables y número de casos posibles omitiendo que los espacios muestrales deben ser finitos y equiprobables.²
- ✓ Se trabaja con el concepto de *variable* sin definirla, no se la reconoce como característica que se desea estudiar en la población.³

² Ed. Estrada Manual de 6º grado. 2006

³ Matemática de 7º grado (1º E.S.B.). Ed. Puerto de Palos. 2005.

- ✓ Las *medidas* no son definidas como concepto, sólo se trabajan desde su cálculo numérico y no desde su interpretación enmarcada en el contexto de un problema, no se justifica la necesidad de su cálculo y de su aplicación.
- ✓ La noción de *mediana* pierde su característica esencial al sólo definirse como valor central sin otra consideración que la complete y justifique.
- ✓ La única diferenciación entre *mediana* y *media aritmética* que se realiza es a través de su forma de cálculo.
- ✓ Las *medidas de dispersión* sólo se calculan sin explicar qué indican y sin vincularlas con las medidas de posición.⁴

- ✓ Los conceptos de *muestra* y *población* no se distinguen claramente.
- ✓ A todas las medidas se las denomina *parámetros estadísticos* perdiéndose la distinción entre los conceptos de parámetro poblacional y estadístico muestral, y cómo están relacionados.
- ✓ Se presentan las *medidas de tendencia central* para resumir la información contenida en los datos y no se dan sus respectivas interpretaciones.
- ✓ Las *variables cualitativas* no se mencionan aún cuando prevalecen en la información de uso corriente, por ejemplo en los medios de comunicación.
- ✓ Se presentan las *medidas de dispersión*, aclarando solamente que permiten completar la información que proporcionan los datos, pero no se trabaja sobre su construcción y concepto.⁵

- ✓ Se presentan los *parámetros estadísticos* de centralización y de dispersión sin dar sus interpretaciones.
- ✓ Se consideran los conceptos de *muestra* y de *población* en forma indistinta.
- ✓ Se plantean problemas que involucran *propiedades y relaciones* no vistas entre las medidas.⁶
- ✓ Se denomina *parámetros estadísticos* a los valores que representan a los datos de una *muestra* o de una *población* en forma indistinta.⁷

- ✓ Se denomina *suceso* a cada uno de los resultados del experimento, dando una definición restrictiva e incompleta del concepto.
- ✓ Se define *muestra* y *población*, pero a todas las medidas se las denomina *parámetros estadísticos*.⁸

- ✓ A las *variables cualitativas* se las identifica como caracteres pero no variables.⁹

- ✓ Se define *estadística descriptiva* como la parte de la estadística que trata de analizar un grupo dado sin incluir la distinción de los conceptos de *población* y *muestra*.
- ✓ A las *variables cualitativas* sólo se las trata en su cuantificación numérica.
- ✓ A las *medidas* se las denomina estadísticos de centralización o posición, se presenta sólo el cálculo y no se aclara que se está trabajando con muestras. Ni se mencionan sus propiedades, interpretaciones y relaciones.

⁴ Textos 8° y 9° ((2° y 3° E.S.B.): Matemática I, Ed. Santillana, 1995. Matemática 8, Ed. Nuevas Propuestas. Matemática 8, Ed. Tinta Fresca, 2005. Matemática 9, Ed. Puerto de Palos, 2001.

⁵ Ed. Kapelusz de 9° año.

⁶ Ed. Santillana de 9° año.

⁷ Ed A-Z de 9° año .

⁸ Ed. Kapelusz de 1° año.

⁹ Ed. Kapelusz de 1° año.

- ✓ Para calcular *probabilidades* se dan dos métodos: recurriendo al experimento, usando la frecuencia relativa y utilizando la fórmula de Laplace, sin especificar las condiciones de aplicación de cada una de las definiciones.
- ✓ Se presenta la *probabilidad total, condicional* y la *regla de la multiplicación* desde una definición incompleta de probabilidad.¹⁰

En resumen, en algunos de los textos revisados aparecen definiciones próximas a las que se manejan en los libros de Estadística y que pueden considerarse universalmente aceptadas. Esto sucede con la definición de suceso o la definición clásica de probabilidad, que se presentan *sólo en una de sus formas* o *dejando de lado el contexto de trabajo*, ya que no se tienen en cuenta características del fenómeno o experimento como, por ejemplo, que hay sucesos con más de un elemento o que todos los resultados del experimento sean igualmente probables de ocurrir. Para el caso de la noción empírica de probabilidad, para espacios muestrales no equiprobables, se enuncia una aproximación al concepto como “...*la frecuencia relativa de un suceso al realizar el experimento un número muy grande de veces...*” justificando esta práctica con la ley de los grandes números, “*a medida que aumenta el número de repeticiones del experimento, el valor de la frecuencia relativa del evento se aproxima más a la probabilidad del mismo*”. El valor de esta propiedad no se agota en lo que afirma sino que además introduce la noción de límite de manera concreta.

En los temas trabajados en estadística descriptiva es reiterativa una tendencia muy marcada al cálculo matemático de las medidas en detrimento de la conceptualización e interpretación.

Por otro lado, es para destacar que en el material examinado, aparecen referencias históricas sobre las primeras experiencias de recolección de datos, siendo ésta una buena herramienta didáctica complementaria y un elemento para la motivación de la curiosidad de valor *per se*.

Muchas observaciones se repitieron en diferentes textos, de esta manera la frecuencia con que aparecen también resulta interesante a la hora de la detección de las carencias o inconvenientes en la presentación de los conceptos estadísticos. De esta forma los conceptos pilares donde se han detectado las mayores dificultades son:

- Población y muestra.
- Individuo.
- Variables cualitativas y cuantitativas discretas y continuas.
- Tabla de frecuencias, frecuencias absolutas y relativas.
- Media, Moda y Mediana.
- Varianza y desvío absoluto, medio y típico.
- Azar y aleatoriedad.

¹⁰ Ed. A-Z, Funciones y estadística De Simone, Turner.

- Espacio muestral y sucesos. Sucesos especiales.
- Probabilidad.

En general, puede decirse que los conceptos expresados aparecen de manera incompleta o confusa sin llegar a su formalización a través de una visión más integral de los mismos. La trascendencia de estos conceptos es fundamental como base para la construcción del conocimiento estadístico, y por esta razón constituyen el punto de partida en la elaboración del registro de términos estadísticos.

Estos conceptos–nodos permiten establecer redes conceptuales que involucran un buen número de nociones, técnicas y procedimientos que podrán consultarse en el registro.

- ***Revisión de libros universitarios básicos***

En los libros universitarios revisados, los de uso más frecuente, no se han detectado nociones confusas o contradictorias, sin embargo estos textos pueden agruparse en dos grandes grupos no excluyentes: los de estadística aplicada y los enfocados más a desarrollos teóricos. En relación a los primeros es común apreciar una tendencia calculista sobre la interpretativa o analítica; mientras que en los más teóricos se nota poca presencia de situaciones concretas donde se puedan aplicar los contenidos.

En ambos casos se pierde el objetivo de la transferencia, aplicación y comprensión acabada de los conceptos tratados.

En estos libros suelen aparecer distintos nombres para un mismo concepto como ser: *Docimasia de Hipótesis* –se utiliza en el libro de Probabilidad y aplicaciones estadísticas de P. Meyer– el nombre que puede darse a la metodología utilizada para resolver casos de verificación de hipótesis, también denominada prueba o ensayo de hipótesis.

Al teorema de *probabilidad total* se lo denomina también regla de eliminación.

Definir los conceptos estadísticos identificándolos con los nombres que reciben sus referentes en la teoría de conjuntos es otro ejemplo, tal como se utiliza en los módulos de estadística y probabilidades finitas de la Maestría en Enseñanza de la Matemática de la UNCU.

Al mismo tiempo el término *promedio*, que se refiere indistintamente a media aritmética, mediana, centro de amplitud y moda, aunque cada uno presenta un panorama diferente sobre la representación de los datos, usualmente en los textos se refiere a la media aritmética.

Vale aclarar que hasta el término *estadística* presenta distintas acepciones, puede significar la ciencia propiamente dicha como así también el resultado del trabajo realizado por un estadístico. Y la expresión *estadístico*, por su parte, puede identificarse con la persona especializada en estadística como así también con medidas obtenidas a partir de muestras.

En algunas áreas, como ser Metodología de la Investigación Social se pudo observar el uso indistinto de los conceptos de unidad de observación y unidad de análisis siendo éstos diferentes para ciertos autores, como R. Sierra Bravo¹¹, quien sostiene que la unidad de análisis corresponde a colecciones de individuos –o conglomerados- y las unidades de observación en general pueden ser individuos particulares.

En los casos expuestos, se observan distintos términos que se utilizan para indicar significados iguales - *significados idénticos*- y expresiones idénticas para designar distintas nociones muchas veces sin las debidas aclaraciones para el lector novato.

Otra observación encontrada en estos libros es que en general se emiten conclusiones taxativas en los procedimientos inferenciales –ambientes no determinísticos– donde el uso de los verbos en *modo potencial* podría ayudar a conservar la noción de incertidumbre que deben acompañar a estas conclusiones.

- ***Revisión de páginas de Internet de uso público***

A continuación se sintetizan algunas observaciones realizadas en las páginas de Internet analizadas en las que se encontraron errores en ciertos conceptos.

Cada bloque corresponde a una fuente que se cita al final de cada uno de ellos.

- ✓ Se presenta un ejemplo de *muestra representativa* que contradice a la definición del concepto suministrado. Se toma como muestra de la población de todos los estudiantes de la Universidad de Málaga, al conjunto de los alumnos de la Facultad de Medicina, que no constituye una muestra representativa de dicha población.
- ✓ Se afirma erróneamente que la *mediana* de una variable discreta es siempre un valor de la variable estudiada.¹²
- ✓ Se considera a la *estimación* como la operación por la que se asigna a un [parámetro](#) de la [población](#) el mismo valor que a un [estadístico](#) calculado a partir de una [muestra](#), y a la estimación por

¹¹ Sierra Bravo, R. Técnicas de Investigación Social. Teoría y Ejercicios. Decimosegunda Edición. 1998. Editorial Paraninfo.

¹² *Bioestadística: Métodos y Aplicaciones*. [Bioestadística](#). Facultad de Medicina. [Universidad de Málaga](#). ISBN: 847496-653-1. <http://www.bioestadistica.uma.es/libro/node2.htm>

intervalo como una operación arriesgada. De esta forma sólo se tiene en cuenta como válida la estimación puntual de parámetros.

- ✓ Se acota el campo de la *Estadística Inferencial* sólo a las operaciones de estimación sin incluir las pruebas de hipótesis.
- ✓ Se hace referencia a la *distribución de Poisson* sólo como el caso límite de la Binomial.
- ✓ Se sostiene que la *proporción* es sinónimo de [frecuencia relativa](#), es decir que no la trata como una medida de intensidad o tasa y que el sesgo es sinónimo de asimetría.
- ✓ Se afirma que la distribución t de Student es una distribución que sigue la estimación de la desviación típica.¹³

- ✓ Se refiere a la *dicotomía* como el proceso de categorización de una variable en sus dos modalidades posibles. Cuando existen variables factibles de dicotomizar a pesar de presentar más de dos categorías.
- ✓ Hay imprecisión al puntualizar a la *mediana* como el valor central que divide a la población en dos subpoblaciones iguales. Cuando en realidad la igualdad sólo debe referirse a la cantidad de elementos y no a atributos esenciales.
- ✓ En el procedimiento para trazar un *polígono de frecuencias* se omiten los tramos inicial y final de dicha representación gráfica.
- ✓ Se toma al *rango* como la situación de un dato respecto a una distribución.
- ✓ Se circunscribe la variabilidad a la medida rango o recorrido de una variable en estudio.¹⁴

- ✓ Se identifica al concepto *atributo* sólo con variables cualitativas.
- ✓ Al especificar lo que representa un *espacio muestral* no se aclara que se refiere a los resultados de un experimento aleatorio.
- ✓ Se afirma que la *estimación* es el valor específico observado de un estimador, dejando de lado la estimación por intervalos de confianza.
- ✓ Se considera que la *Inferencia estadística* es el proceso de generalizar los resultados de la muestra a los resultados de la población, cuando la inferencia no es sólo una generalización.¹⁵

- ✓ Se representa al *polígono de frecuencias* en forma errónea sin tener en cuenta la marca de clase de los intervalos previo al primero y posterior al último con frecuencia cero, comenzando en el límite inferior del primer intervalo y terminando en el límite superior del último intervalo de clase.
- ✓ Se traza mal la *ojiva*, uniendo las marcas de clase y no partiendo de cero en el límite inferior del primer intervalo.
- ✓ Al precisar el concepto de *parámetro* estadístico no lo diferencia del estadístico muestral.
- ✓ No se hace una descripción adecuada del *histograma*.¹⁶

- ✓ Se define de dos maneras distintas a los *percentiles*, una de ellas como cada uno de los 99 segmentos que se toman al dividir una muestra o un conjunto de elementos ordenados por cien

¹³ [Diccionario de Estadística Práctica](http://www.hojamat.es/estadistica/diccio/diccestad.htm) <http://www.hojamat.es/estadistica/diccio/diccestad.htm>.

¹⁴ Fernando Valdes - Universidad Romulo Gallegos ([Robert Ponterio](#) - Project Director and HTML). [SUNY Cortland Dept. of International Communications and Culture](#). GLOSARIO

<http://maristascoruna.wikispaces.com/file/view/Apuntes+Estadistica+P4.pdf>

¹⁵ <http://www.so.ucr.ac.cr/Enlaces/Estadistica/problemas/Diccionario.html>

¹⁶ http://www.vitutor.com/estadistica/descriptiva/a_2.html

partes de igual frecuencia. Y otra como puntos característicos de la distribución que dividen a la distribución en cien partes.¹⁷

- ✓ Se define el concepto de *Aleatorio* afirmando que no sigue un patrón particular que se pueda describir directamente por ecuaciones. Se basa más en la probabilidad. El término aleatorio representa una idea que debe ser expresada en términos del concepto de probabilidad. Se trata de una definición circular porque a partir de la probabilidad se intenta medir la aleatoriedad de un suceso o fenómeno.
- ✓ Se define *Atributos* como característica cualitativa de un objeto o individuo tal como sexo, país de origen, estado marital. Es una definición acotada pues no contempla las características cuantitativas como por ejemplo la edad es un atributo de la unidad de análisis y no es una característica cualitativa.
- ✓ Se considera la *Variable cuantitativa discreta* como aquella a la que se les puede asociar un número entero, es decir, aquella que por su naturaleza no admiten un fraccionamiento de la unidad, por ejemplo número de hermanos, páginas de un libro, etc. Variable que tiene un número limitado de valores. El sexo, por ejemplo, sería una variable discreta pero cuidado, no lo es.
- ✓ Se define a la *Variable dependiente* como aquella que teóricamente puede tomar cualquier valor en una escala de medidas, ya sea entero o fraccionario cuando esa definición correspondería a una variable cuantitativa y que se trata de una Variable aleatoria o estocástica cuando no es posible hacer medidas numéricas, serían aquellas que son susceptibles de clasificación.¹⁸ Concepto erróneo pues confunde las variables aleatorias con las variables cualitativas o categóricas.

En síntesis, en las páginas consultadas se han detectado imprecisiones en definiciones de conceptos y en representaciones gráficas. Se encontró además que no se define el concepto de media aritmética y sólo se indica el procedimiento algebraico para su determinación, quedando de lado el significado, su interpretación y sus limitaciones.

Muchas de las imprecisiones detectadas coinciden con las encontradas en los libros de texto de la escuela.

• ***Revisión de información periodística***

Debido al gran volumen de información periodística que incluye datos estadísticos que son interpretados libremente se han seleccionado algunos casos interesantes, a modo de ejemplo, como los siguientes:

- ✓ En el artículo del diario Clarín del 26 de julio de 2009, bajo el título “*Inmovilidad social*” dice: “*Casi 4 de cada 10 chicos que nacen en un lugar ubicado en el quintil más bajo de la pirámide socioeconómica (el 20% más pobre) permanece en esa condición el resto de sus vidas. En contraposición, sólo un 22% de los hijos de padres que están entre el 20% más rico de la Argentina sigue estando en ese quintil en su vida adulta*”. Aquí se considera *quintil* como el

¹⁷ http://es.wikipedia.org/wiki/Medidas_de_posici%C3%B3n_no_central

¹⁸ Diccionario de Estadística Práctica. Universidad Interamericana de Panama. Lic. Admin de Recursos Humanos. 2 de febrero de 2007.

porcentaje de datos que deja por debajo ese valor (una franja) y no como un punto en el espacio de los datos ordenados, sin especificar su valor. Además no se aclara qué variable se analizó para determinar las regiones mencionadas de la pirámide. Por otro lado, la inmovilidad social que se estudia, aparece cuantificada en función de los datos analizados, pero de ninguna manera pueden asimilarse como números que establecen parámetros, salvo que se apliquen técnicas de inferencia apropiadas. Esto no se aclara y puede llevar a conclusiones equivocadas.

- ✓ Juan Ignacio Piovani en el XXV Congreso ALAS, Porto Alegre, Brasil en agosto de 2005, en su ponencia expresa que ha tratado de mostrar *“que las ideas estereotípicas con respecto a la estadística y sus aplicaciones en las ciencias sociales tienen evidentemente algún fundamento; pero como sucede con los estereotipos, son visiones simplificadas y estandarizadas de una trama mucho más compleja y que detrás de la denuncia legítima de los límites que dichos instrumentos tienen se halla la ignorancia y el desinterés por aprender algo que demanda esfuerzo, y que dentro de ciertos límites —y concebido como instrumental a los fines cognoscitivos de la ciencia— podría resultar de mucha utilidad (Marradi, 1997). “Yule se esforzó por denunciar los efectos nefastos de la absolutización de la medición en la investigación social y económica. Para él muchos fenómenos no eran medibles y debían por lo tanto ser analizados, aún en el marco de un abordaje estadístico, cualitativamente. Fue justamente él quien inició el desarrollo de técnicas estadísticas adecuadas para el tratamiento de variables cualitativas”*. Así, las herramientas estadísticas suelen quedar “tironeadas” entre un uso dogmático, ritualista y acrítico, *Ya en la década de 1930, el destacado sociólogo de la Escuela de Chicago Znaniecki advertía sobre los riesgos de “sustituir los métodos intelectuales por técnicas tabulares, y de esta forma eliminar el pensamiento teórico del proceso de investigación científica”*. También cita a Pearson quien en 1933 *se lamentaba por el carácter “purista” y alejado de los problemas de la práctica científica que la estadística —y sus aplicaciones— estaba adquiriendo. Para él ella debía ser eminentemente “aplicada”, y tal definición daba prioridad a la consideración de los problemas de investigación, y no a los instrumentos. La estadística era justamente instrumental, un medio al que recurrir en la labor científica, y no un fin en sí mismo. Pearson temía, como Znaniecki, que la investigación —apelando a la estadística como instrumento que todo lo puede— quedara en manos de los que llamaba “técnicos” de la ciencia, sujetos carentes de aquella “imaginación disciplinada” que caracteriza al verdadero científico*”. Esto implica, obviamente, problematizar los instrumentos que se utiliza en la investigación. Dice Piovani que cuando están dadas las condiciones para su uso, no hay ninguna razón para no usarlas; pero cuando no están dadas tales condiciones, no hay ninguna razón para usarlas.¹⁹
- ✓ En *Los retos de la cultura estadística* de Carmen Batanero, Universidad de Granada, España, 2002, la autora se refiere a la reflexión epistemológica sobre el significado de los conceptos, procedimientos (en general objetos) particulares. La especialista afirma que *“los organismos oficiales de estadística se han concienciado de la necesidad de hacer llegar los estudios que realizan en forma comprensible a todos los ciudadanos y al mismo tiempo mejorar la imagen pública de la estadística. Además de poner el énfasis en la provisión de información y consejo para el gobierno y uso profesional y en investigación, las organizaciones estadísticas oficiales se interesan en proporcionar información a la sociedad, como un todo”*. En vistas de la difusión y

¹⁹ Juan Ignacio Piovani. XXV Congreso ALAS. Universidade Federal do Rio Grande do Sul. Porto Alegre, 22-26 de agosto de 2005. “La estadística y el sentido común de los sociólogos”. *Università degli studi di Bologna, Representación en Argentina. Universidad Nacional de La Plata*.

necesidad de trabajar sobre este tema, menciona una página con la iniciativa de una revista para la educación estadística.²⁰

- ✓ En el artículo titulado *Peligrosas estadísticas y dudosas encuestas* de Rosa Montero, escritora española, publicado por el diario Los Andes, 17 de octubre de 2004, Mendoza, Argentina, la autora presenta situaciones que conducen a resultados absurdos o engañosos alertando acerca de los efectos del mal uso de la Estadística o de interpretación errónea de Información estadística. Muestra *abusos* en el uso de la Estadística y afirma que las encuestas son poco confiables pues dependen tanto de la formulación de las preguntas como de la selección de los encuestados²¹.
- ✓ En *La media no es el mensaje*, artículo de Stephen Jay Gould, biólogo y prolífico divulgador científico, él relata su propia experiencia dado que enfermó gravemente de un cáncer muy raro y agresivo en 1982. Buscó información estadística sobre la enfermedad y se enteró –con espanto– que, una vez detectado, la media de supervivencia era de 8 meses. En base a su incansable espíritu curioso pudo profundizar sobre cómo fueron obtenidas las estadísticas halladas y comprendió que su vida podría no terminar en el tiempo medio. El ya sabía que una *media* puede extraerse, por ejemplo, de una muestra de pacientes muy homogénea, en la que todos mueren aproximadamente al cabo de 8 meses, pero también puede tratarse de una curva muy pronunciada, en la que algunos enfermos mueran enseguida y otros vivan mucho tiempo. Todo indicaba que dado su calidad de vida Gould pertenecía al grupo de mayor sobrevida, de hecho vivió 20 años más y murió de otra enfermedad. De esta manera, conocer las propiedades de la media aritmética permiten lograr una correcta interpretación, ya que en ciertas distribuciones no es representativa del conjunto de datos y, sin saberlo, se puede llegar a conclusiones fatalistas.

Las publicaciones analizadas corroboran que existe un denominador común que hace que las estadísticas no siempre sean bien vistas en algunos ámbitos y circunstancias. No quedan dudas que la imagen de la estadística se encuentra deteriorada y de acuerdo a estas experiencias puede deberse al mal empleo o, en algunos casos inclusive, al uso malintencionado que se hace de sus resultados para mostrar realidades que no son.

El mal uso de la estadística, en general puede resumirse como la falta de conocimiento del alcance de la información estadística, el manejo indiscriminado de datos cuantitativos, que si no se expresan responsablemente pueden resultar con derivaciones alarmantes –o hasta devastadoras- si se habla, por ejemplo, de enfermedades, crisis económicas, catástrofes climáticas y tendencias de todo tipo.

Por las dificultades y errores encontrados es recomendable la divulgación masiva de los conocimientos básicos estadísticos a los alumnos, profesionales y trabajadores en general porque las consecuencias de la ignorancia estadística queda patente en las publicaciones revisadas. Por otro lado, cabe agregar que la liviandad con que pueden extraerse predicciones

²⁰ <http://www.ugr.es/local/batanero/sergroup.htm>.

²¹ *Aún las encuestas más exactas y fiables pueden conducir a graves errores según cómo se lea... a veces incluso pueden matar*”, afirma ejemplificándolo con el ejemplo de un enfermo terminal de cáncer.

a partir de hechos ocurridos denota una hipersimplificación en la interpretación de las causas que lo producen, sin tener presente que existen factores que no pueden manejarse pero si tal vez controlarse, así como múltiples factores que podrían corregirse para evitar consecuencias no esperadas como tragedias de cualquier índole.

- ***Antecedentes en investigación sobre la formación de conceptos estadísticos***

I. Evaluación del Conocimiento Estadístico en la Formación Inicial del Profesorado realizado en la Universidad de Lleida (España). Informe elaborado por Assumpta Estrada²²

Mediante estudios de campo realizados sobre la educación estadística de profesores en formación se obtuvieron resultados que permiten ver cómo hay situaciones recurrentes más allá del origen y el contexto de las observaciones realizadas.

En particular se trata de un trabajo de evaluación del conocimiento estadístico en la formación inicial del profesorado. De ahí surge como primera y más importante recomendación la necesidad de potenciar la formación estadística de los futuros profesores.

El objetivo del trabajo fue el de obtener información sobre cuáles son los principales errores y dificultades de los profesores en formación en conceptos estadísticos elementales (incluidos en los contenidos del nivel primario) como ser *promedio, probabilidad, frecuencia, dispersión, asociación, muestreo y simetría, interpretación de gráficos, efectos de valores atípicos*, entre otros.

Un dato a considerar es que se diferenció a los participantes por especialidad en la formación docente y por los años de estudio.

Las conclusiones obtenidas del trabajo indicaron que si bien la especialidad influyó sobre los resultados, tanto como para concluir que los de áreas como Lenguas o Educación Inicial fueron los de menor porcentaje de aciertos y los de Educación Especial o Educación Física los de mayores porcentajes de aciertos, los años de estudio no cambiaron significativamente los resultados.

La explicación que da la investigadora a estos resultados es que posiblemente esto se deba a que la enseñanza que reciben se basa principalmente en el cálculo, la aplicación rutinaria de fórmulas y la representación gráfica, siendo raras las actividades interpretativas similares a las propuestas en el cuestionario.

²² <http://web.udl.es/usuaris/z4084849/docs/uno45.pdf>

Los resultados generales obtenidos sobre todos los docentes indagados se encuentran resumidos en la siguiente tabla donde aparecen los conceptos evaluados y el porcentaje de respuestas correctas.

Tabla 1. Contenido evaluado y resultado en los ítems

Ítem	Contenido evaluado	Porcentaje de aciertos	Principales errores
1	Media, valor atípico.	47	Confundir media y moda. No detectar valores atípicos.
2	Probabilidad.	76	Interpretación cualitativa.
3	Probabilidad frecuencial.	60	Enfoque en el resultado.
4	Media, mediana, moda, valor atípico.	72	No considerar el contexto.
5	Comparación de muestras. Interpretación de gráficos.	73	Estimación incorrecta de promedios. Basarse sólo en el rango. Basarse sólo en datos aislados.
6	Muestreo.	43	No apreciar el tamaño muestreo. Confusión correlación/ causalidad.
7	Promedios, inversión algoritmo media.	33	No invertir algoritmo. Posición de promedios en distribuciones no simétricas.
8	Tamaño de muestra. Variabilidad.	73	Equiprobabilidad. No apreciar el tamaño de la muestra.
9	Estimación y métodos de muestreo.	60	No percibir el sesgo en muestreo. No apreciar el tamaño de muestra. No distinguir métodos de muestreo.

II. Inferencia Estadística y Lenguaje: Un Estudio En Bachillerato realizado en la Universidad de La Laguna (España) por Israel García Alonso y Juan Antonio García Cruz.²³

En este estudio se trabajó sobre varios términos relacionados con la Inferencia Estadística que aparecen en los libros de texto de secundaria. La tarea se centró en comparar el significado de dichos términos con el que se les asigna en los libros de texto.

Los investigadores se fundamentan en que en la construcción del conocimiento matemático un vehículo importante es el lenguaje, y cuando se requiere un alto nivel de abstracción acompañado de una considerable tecnificación, se hace necesario el conocimiento y buen uso de términos específicos. Es este el caso de la Inferencia Estadística.

Sostienen también que en el aula de matemáticas se distinguen dos tipos de contexto: contexto cotidiano y contexto matemático. Se entiende por contexto cotidiano a aquel en el que se lleva a cabo la comunicación habitual, mientras que el contexto matemático es propio de la

²³ <http://dialnet.unirioja.es/servlet/articulo?codigo=2697036>

matemática como ciencia formal. Se pudo comprobar cómo algunos textos dan definiciones de los términos más relacionadas a su noción informal o cotidiana que a la formal ó matemática.

Como criterio de ordenamiento de los conceptos se realizó una clasificación según su significado para poder realizar la comparación mencionada.

Según esta premisa, los autores clasifican los términos de la siguiente forma: 1) términos con el mismo significado en ambos contextos; 2) términos con distinto significado en ambos contextos, y 3) términos propios del contexto matemático. Los estudiantes no deberían tener dificultad con la categoría 1, mientras que la categoría 3 necesariamente debe ser definida pues son términos que los estudiantes no conocen ni tienen en su vocabulario. La categoría 2 puede dar lugar a más dificultades para los estudiantes.

Los resultados de la clasificación de los términos fueron los siguientes:

1. *Mismo significado en ambos contextos*: Estadística, Población, Individuo, Tamaño de la muestra.
2. *Distinto significado en ambos contextos*: Media, Muestra, Estimación, Inferir, Distribución, Probabilidad, Representativo, Riesgo Significativo.
3. *Significado propio en el contexto matemático*: Parámetro, Estadístico, Muestreo aleatorio, Media muestral, Media poblacional, Nivel de confianza, Desviación típica, Estadística inductiva, Estadística hipotético deductiva, Error máximo admisible, Normal, Sesgado, Eficiente, Proporción muestral, Contraste de Hipótesis, Hipótesis Nula, Hipótesis Alternativa, Nivel de Significación, Error de Tipo I, Error de Tipo II

La primera categoría no deberían presentar dificultad a los estudiantes, por aparecer de manera habitual en el lenguaje cotidiano. En algunas ocasiones se añaden términos en la definición modificando sensiblemente su significado. La segunda categoría son los términos con distinto significado en ambos contextos y a veces pasa desapercibido para estudiantes y profesores. Este hecho puede provocar que el estudiante aprenda este concepto matemático con errores. Como entre los textos analizados predominan los que utilizan las definiciones dadas en el contexto cotidiano, en vez de usar la definición del contexto matemático, se plantea una situación delicada ya que el libro tendría que presentar a los estudiantes los conceptos matemáticos definidos formalmente. La tercera categoría son términos que en general el estudiante desconoce. Se observó que muchos de los textos no los definen, porque suponen que los estudiantes los conocen o por considerar que de esta forma se puede simplificar el proceso de enseñanza y aprendizaje. En otros casos, hemos observado que

algunos textos introducen términos propios de la Inferencia Estadística que luego no vuelven a utilizar, o que utilizan el concepto con otra expresión. Situación que entendemos podría provocar confusión identificando al texto como inconsistente, por no continuar utilizando aquellos términos que define. Puede decirse, entonces, que algunos libros de texto no cumplen con uno de los objetivos que deben cumplir, esto es incluir las definiciones de los términos utilizados en el contexto matemático. En otras ocasiones, el libro de texto modifica la definición, o simplemente enuncia una definición errónea lo cual es contraproducente para los estudiantes porque éstos no poseen los conocimientos necesarios para identificar el error.

En estadística inferencial, son muchos los conceptos nuevos que el estudiante debe comprender correctamente para así entender el razonamiento y el procedimiento necesario para resolver correctamente las técnicas de inferencia.

En varias investigaciones realizadas en universidades españolas analizadas se reiteran las dificultades detectadas en las otras publicaciones. Como notables pueden mencionarse los errores de interpretación y aplicación de medidas estadísticas, y como fuera explicado en cada uno de los relevamientos comentados, generalmente se encuentran asociados a la falta de conocimiento y entrenamiento en su utilización. El mero cálculo de éstas conduce a la identificación de una medida con su fórmula más que con su significado, desaprovechando la información que brinda del conjunto de datos, como por ejemplo si la media aritmética es adecuada o no para representarlo.

Otro punto recurrente es la falta de consideración del contexto de trabajo lo cual no favorece a la comprensión de los conceptos tratados, incurriéndose en lo que se podría denominar un error didáctico. Cabe aclarar que en la enseñanza de la estadística es propicio incluir ejemplos y actividades utilizando situaciones concretas asociadas a temas interesantes para el alumnado, pudiéndose desarrollar a partir de ellas conceptos y razonamientos estadísticos y hasta probabilísticos para, en una segunda etapa, formalizarlos con toda la rigurosidad del contexto matemático, logrando así que lo formal adquiriera sentido en lo concreto y a su vez, que lo concreto se revalide desde lo formal. Esta doble tarea permitiría trabajar sobre la cuestionada imagen de la estadística y ahondar en su red conceptual para dar una correcta utilización de los recursos que brinda.

Relevamiento de información primaria y sus resultados

Con la finalidad de complementar el relevamiento y análisis efectuado, se recurrió a fuentes primarias de información, realizando entrevistas y encuestas.

• *Resultados de entrevistas realizadas a usuarios de la estadística*

Se realizaron entrevistas a 40 profesionales de diversas áreas distribuidos de la siguiente manera: 8 abogados (20%), 12 del área contable (30%), 10 físicos o ingenieros (25%) y 10 de las ciencias sociales (25%).

Las entrevistas fueron realizadas por los investigadores del equipo, el 45% de los entrevistados se desempeñan como docentes universitarios (2 abogados, 3 contadores, 3 lic. en administración, 4 físicos, 1 ingeniero, 3 sociólogos y 2 trabajadores sociales) y el 55% restante realiza el ejercicio profesional en distintos ámbitos.

Los resultados más destacados de dichas entrevistas fueron:

- 7 de los 8 de los abogados consultan tablas comparativas de legislación, gráficos porcentuales tanto de barras como circulares relacionados con análisis cualitativos y registros de datos históricos. Muy pocos manejan alguna medida estadística y sólo utilizan tasas.
- Todos los contadores y administrativos utilizan tablas y gráficos porcentuales e interpretan medidas previamente calculadas.
Menos de la mitad de ellos manifestó realizar análisis de números índices y series de tiempo y ninguno usa herramientas relacionadas con la inferencia estadística.
- Todos los físicos e ingenieros han utilizado o utilizan los distintos tipos de gráficos y medidas estadísticas.
8 de los 10 realizaron o realizan comparaciones, algún tipo de prueba de hipótesis y estimaciones.
- Con respecto a las Ciencias Sociales, todos los profesionales del área trabajan con tablas y gráficos a partir de datos tanto primarios como secundarios, utilizan indicadores y relacionan variables fundamentalmente cualitativas.

Entre los entrevistados sólo investigan los físicos (3 de los 4) y el 75% de los profesionales de las ciencias sociales, precisamente los que afirmaron utilizar herramientas relacionadas con la inferencia estadística.

Ninguno de los entrevistados ha realizado algún curso de capacitación en estadística y salvo los abogados todos los demás profesionales la han estudiado como materia en la carrera de grado.

El 90% de los entrevistados expusieron la necesidad de ampliar sus conocimientos en esta área.

- ***Resultados de las Encuestas a profesionales de la salud***

Complementando las entrevistas realizadas y con el objetivo de indagar en los posibles usos de las distintas herramientas estadísticas, se encuestaron 55 profesionales del área de la salud distribuidos de la siguiente manera: 18 médicos (33%), 12 terapeutas ocupacionales (22%), 9 psicólogos (16%), 6 enfermeros (12%) y 10 de otras profesiones afines (como ser sicopedagogos, kinesiólogos, etc.) (18%).

El relevamiento de las encuestas fue realizado por alumnos de la carrera de Licenciatura en Terapia Ocupacional de la Universidad Nacional de Quilmes, e implementadas en los diversos ámbitos donde se desempeñan laboralmente: hospitales, organismos gubernamentales y no gubernamentales, hogares, etc.

Del análisis de las encuestas se pueden destacar los siguientes aspectos:

- El 67% de los profesionales encuestados se desempeña en instituciones no hospitalarias.
- El 30% ha realizado algún curso relacionado con la estadística pero todos ellos han tenido en la carrera alguna materia perteneciente a esta área (como ser bioestadística o estadística aplicada).
- El 43% realiza algún tipo de investigación o de trabajo de campo relacionado con la investigación.
- Sólo el 26% accedió al estudio de herramientas estadísticas en algún curso de capacitación realizado aunque el 55% las utiliza en el ejercicio de su profesión.

Las herramientas estadísticas comúnmente utilizadas son los gráficos, las tablas, las medidas de posición moda, media, mediana y percentiles, la desviación estándar, las

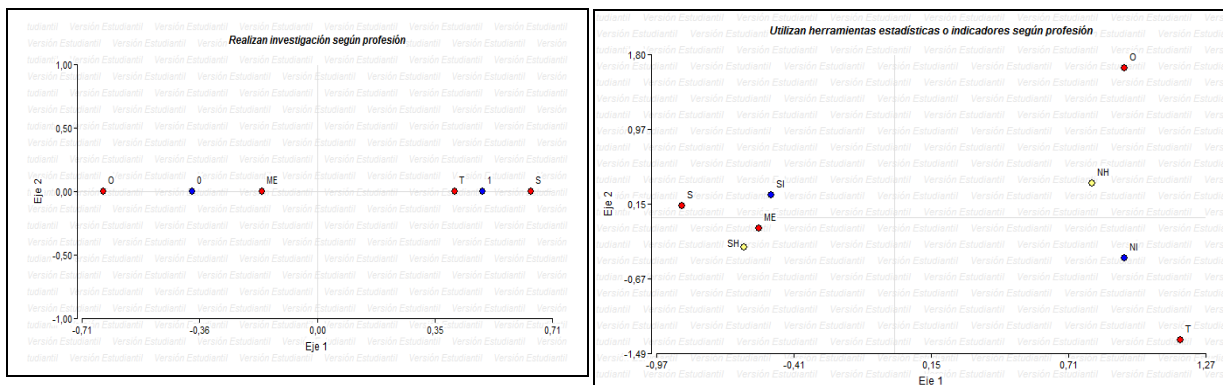
proporciones y razones y en menor grado (sólo dos de ellos declararon usarlas) las pruebas chi cuadrado y el análisis de correlación y de regresión.

- El 68% utiliza indicadores o escalas de medición prediseñadas.

Con respecto a los indicadores o escalas los profesionales de la salud señalaron a las tablas de percentiles, la escala Glasgow, las tasas y a distintos tipos de test como de uso más frecuente.

Al analizar la encuesta en su conjunto, se puede observar que existe evidencia estadística (con un nivel de significación superior al 0,022%) de la relación entre la profesión y el trabajo de investigación entre los encuestados, investigan el 67% de los terapistas ocupacionales, el 75% de los psicólogos, el 35% de los médicos o enfermeros y sólo el 11% de los de otras profesiones. Sin embargo son los médicos y los psicólogos quienes se caracterizan por la utilización de las herramientas estadísticas e indicadores.

En los siguientes gráficos (biplots) obtenidos del análisis de correspondencias pueden visualizarse las observaciones anteriores.



Donde: 0 es no investigan, 1 es investigan, S psicólogos, T terapistas, ME médicos o enfermeros y O otros, SH utiliza herramientas estadísticas, NH no utiliza herramientas estadísticas, SI utiliza indicadores, NI no utiliza indicadores.

Finalmente y con respecto a las observaciones volcadas por los encuestados, aproximadamente el 50% de ellos opina que necesita ayuda externa en la interpretación de determinadas tablas y gráficos y de los resultados obtenidos por algunos indicadores.

• **Consulta a alumnos**

Se indagó sobre los conceptos y técnicas estadísticas que utilizan los alumnos en sus ámbitos laborales. Esta búsqueda se llevó a cabo con aquellos alumnos del Departamento de Ciencias Económicas de la UNLaM que cursaron la materia Estadística, en el primer cuatrimestre de 2010, entre las respuestas más destacadas figuran:

- Uso de intervalos de confianza como medio para el control de calidad en la empresa para comparar los resultados con los registros ya establecidos.
Cabe aclarar que en la práctica cotidiana, tratando el tema concerniente a la estimación de parámetros estadísticos mediante intervalos de confianza, se reiteran errores frecuentes en la interpretación, como la confusión entre la estimación puntual de parámetros y la estimación mediante intervalos de confianza; errores en la interpretación de los mismos y trabas en la conceptualización del tema.
- Definición y clasificación de variables adecuadas y su tabulación para el estudio estadístico acerca de la caída de reservaciones realizadas para abordar un avión luego de haber sido reservado el pasaje.
- Confección de encuestas con la intención de obtener información en el área de educación estudiando opciones alternativas a preguntas abiertas para disminuir la dificultad de recolectar, clasificar o utilizar los datos para obtener conclusiones.
- Resumen y presentación de información recopilada mediante cuestionarios sobre la utilidad de un producto a lanzar en el mercado mediante un muestreo aleatorio.
- Utilización de números índices, estudio del poder adquisitivo mediante el índice de precios al consumidor.
- Confección de tablas y gráficos adecuados para mejorar la manera de presentar información ya procesada.
- Estudio de series temporales y detección de patrones de comportamiento de variables de diversa índole, incluidos los juegos de azar de realización periódica o semanal.

De la indagación realizada se obtuvo que en el ámbito laboral se analiza información, se trabaja con *intervalos de confianza*, *números índices*, *tablas de frecuencias*, *series de tiempo* y *tablas conjuntas*; en general los alumnos manifestaron realizar estas actividades sólo desde la operación y en muchos casos sin comprender el porqué. El relacionar, en el aula, los conceptos y las herramientas estadísticas utilizadas en el trabajo con los tópicos de la materia permitió a los alumnos, apreciar su utilidad y la aplicación de los contenidos estudiados en una situación concreta facilitando de esta forma su comprensión capitalizando así los aportes realizados por éstos.

1.4. Conclusión

Puede decirse que durante el siglo XX la estadística ha experimentado una nueva expansión, un crecimiento, debido a la incorporación del uso de ideas estadísticas en diferentes disciplinas, como puede verse en revistas científicas y en la creciente inclusión de especialistas en la materia en los equipos de investigación y de trabajo interdisciplinarios convirtiéndose en una de las ciencias metodológicas fundamentales y base del método científico experimental permitiéndole a su vez un desarrollo notable, desarrollo potenciado por los adelantos de la informática que le ha permitido trabajar con grandes masas de datos en tiempo real. Sin embargo, es notoria la falta de una cultura estadística en los ciudadanos inclusive a niveles académicos como lo puntualizara C. Batanero (2001)²⁴, *la estadística se usa incorrectamente, no se comprenden conceptos aparentemente elementales y no hay una valoración suficiente del trabajo del estadístico*, dentro de los equipos de investigación. Señalando así la existencia de una problemática educativa que tiene su raíz en que la inclusión de la estadística en la escuela, no es todavía un hecho. Los alumnos llegan a la universidad sin los conocimientos estadísticos elementales y es preciso comenzar el programa repitiendo los contenidos de estadística que debieran haberse asimilado en la escuela. Así, los profesores universitarios han de acelerar los tiempos de las explicaciones y actividades prácticas sin llegar muchas veces a la formalización acabada de los temas dejando de lado algunas demostraciones o razonamientos que podrían llevar al alumno a comprender mejor la metodología de la estadística. En esas condiciones, el alumno será incapaz de aplicar en su futura vida profesional las herramientas que le brinda esta ciencia.

Por otro lado, el avance de la informática y la masificación de internet ponen a los profesores frente a un nuevo escenario, los alumnos no sólo aprenden estadística en las horas lectivas ni en los centros de enseñanza, sino que se ponen en contacto con información estadística en la prensa, medios de comunicación, internet o en los textos de otras asignaturas. Se puede afirmar que internet está empezando a modificar las relaciones docentes–alumnos, con o sin participación voluntaria de los profesores. Es evidente que los profesores –en los diversos niveles educativos– han de aceptar la rapidez del cambio tecnológico y que deberían implicarse en él, si se quiere apuntar, de algún modo, a la educación estadística, y además, crear una verdadera cultura estadística en la sociedad.

²⁴ Batanero, Carmen, 2001, conferencia organizada por IASE sobre la formación de investigadores.

Dada la nueva situación socio-cultural, la *educación estadística* debería centrarse en los datos y en el razonamiento estadístico, revalorizando la comprensión e interpretación de los resultados estadísticos y difundiendo los conocimientos básicos (habilidad para manejar cifras y problemas cuantitativos; pensar en términos de probabilidades; producción, procesamiento y presentación de datos; presencia de la variabilidad en datos cuantitativos, su cuantificación y explicación).

Es importante también destacar otras componentes de la cultura estadística que son necesarias para un buen desenvolvimiento de los ciudadanos, a saber: comprensión del sentido de la información, reconocimiento de la naturaleza y limitaciones de la información estadística, distinción entre cifras adecuadas e inadecuadas, conocimiento de las consecuencias de la exactitud de los resultados, utilización de diversos procesos de muestreo, entre otros.

La justificación de estas necesidades se la encuentra en la sociedad actual, que se ha vuelto cada vez más compleja y que se halla en continuo cambio, en la misma se utiliza cotidianamente *información estadística* en ámbitos tan diferentes como el político, económico, social, científico y educativo. Poder interpretarla correctamente, así como emplear el razonamiento y los métodos estadísticos para analizar los fenómenos y poder tomar decisiones, es una necesidad para todos los miembros de la sociedad.

Desde un punto de vista global, es explícita la lucha contra el analfabetismo en el mundo, donde se destaca especialmente el rol de la UNESCO, quien ha realizado estudios, a partir de los cuales se concluye que existen carencias en lectoescritura y escasas competencias cuantitativas, así como de conocimientos numéricos básicos. Se afirma además, que los *conocimientos estadísticos esenciales* no han tenido un protagonismo para destacar en el plan de lucha para frenar el analfabetismo, a pesar de su evidente importancia central para la vida cotidiana de la gente y para su inserción en el mundo laboral.

Es claro que para lograr la *alfabetización estadística*, la enseñanza es fundamental y en su planificación hay que tener en cuenta que ciertos conceptos estadísticos son altamente sofisticados y que conllevan cierta dificultad en su aprendizaje; que habría que aprovechar las nuevas tecnología para la enseñanza como recurso para agilizar el proceso, sin descuidar el tener en cuenta los libros de texto y principalmente la formación de profesores.

En función de estas consideraciones, se puso especial énfasis en la tarea realizada en esta investigación en la detección de las falencias que deben atenderse y superarse como punto de partida para una mejor conceptualización de los términos y conceptos estadísticos.

Del análisis de la información relevada, a continuación se resumen las consideraciones más notorias:

- En los libros de texto escolares se observaron carencias o inconvenientes en la presentación de algunos conceptos estadísticos porque aparecían de manera incompleta o confusa sin su formalización correspondiente.
- En los libros universitarios básicos se utilizan distintos términos para indicar significados iguales y expresiones idénticas para designar distintas nociones.
- En las páginas de internet se han detectado imprecisiones tanto en las definiciones de algunos conceptos como en algunas representaciones gráficas.
- En la información periodística seleccionada se observaron algunos ejemplos de información errónea o de dudosa interpretación.
- Los conceptos o términos básicos más conflictivos detectados en cuanto a su presentación y la precisión de su significado son: variables, medidas (modo, media aritmética, mediana, percentiles, quintiles, entre otras), aleatoriedad, muestreo, inferencia en general y pruebas de hipótesis en particular.

Cabe aclarar que en principio se consideraron categorías que agrupaban a los conceptos en base tanto a la experiencia en el aula como a la cotidiana de los integrantes del equipo de investigación. Una de ellas se identificó como *usos y costumbres errados*, instalados fuertemente en el lenguaje y expresiones cotidianas como así también la mención frecuente de términos y técnicas, incluyendo referencias estadísticas y aplicaciones formuladas de manera incorrecta o incompletas. Otra de las categorías iniciales fue la de términos *diferentes* pero con *significados idénticos* a fin de detectar y superar ambigüedades o definiciones confusas. Y por último se había pensado agrupar conceptos y términos pasibles de ser utilizados sin conocimiento cabal de su significado o interpretación, cayendo en falacias de distinto tenor o bien, con fines dudosos siguiendo intereses particulares y no como herramienta metodológica formal dentro de la categoría *manipulación*.

Con el avance de las tareas de relevamiento la tipificación preliminar fue modelándose, revisada y ajustada dejando de ser una división estricta para pasar a ser una serie de caracteres que fueron presentándose sucesivamente, confirmando la existencia de estos usos de los conceptos y técnicas estadísticas, siendo que se repitieron frecuentemente. De esta manera se justifica la previsión de tales categorías pero ha quedado superada la clasificación de los conceptos como pertenecientes a una u otra, echando luz sobre los temas y significados que

debían incluirse, clarificarse y especificar su significado e interpretación en el registro que se iba confeccionando.

A partir de la elaboración de definiciones formales, aceptadas universalmente se han compendiado una importante serie de conceptos, términos, calificativos y técnicas estadísticas de uso más frecuente dando lugar a un registro alfabético de los mismos, que incluye términos que denotan conceptos como también métodos operativos con la finalidad de que puedan convertirse en insumos, tanto para la actividad educativa como laboral de cualquier usuario de estadística.

El registro resultante está organizado alfabéticamente y en forma dinámica materializado mediante links ad-hoc entre los conceptos vinculados por identidad de significado o por pertenecer a una misma red conceptual.

1.5. Registro de términos y conceptos estadísticos

Aleatorio: cualidad de estar sujeto al azar. Nombre equivalente: estocástico.

Amplitud intercuartílica: [Rango intercuartílico](#).

Análisis bivariado: análisis estadístico que involucra el estudio de dos [variables](#) en forma conjunta.

Análisis de correlación: consiste en determinar el grado o fuerza de asociación que existe entre [variables cuantitativas](#).

Análisis de regresión: consiste en elaborar un modelo matemático para formular la relación entre dos o más [variables cuantitativas](#). El modelo puede ser lineal, cuadrático, exponencial, logarítmico, entre otros.

Análisis de regresión lineal bivariado: es el [análisis de regresión](#) entre dos variables donde el modelo matemático que describe la relación es lineal. Se pueden visualizar los patrones que indican cómo están relacionadas las variables en el [diagrama de dispersión](#).

Análisis Exploratorio de Datos (EDA): es un enfoque en el abordaje descriptivo de los datos cuya filosofía hace hincapié en la preservación del dato y de la información subyacente en él. Desarrollado por varios autores desde mediados de la década de 1960, sus principales lineamientos quedaron plasmados en *Exploratory Data Analysis* de J. Tukey (1977). Su aplicabilidad está altamente relacionada con las capacidades logradas por informática que permitió sostener una gran masa de datos y procesarlos en tiempo real, contribuyendo así al mejoramiento de la calidad de la [información estadística](#) resultante.

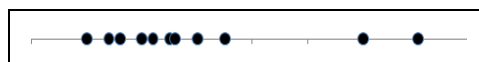
Arreglo de datos ordenados: es una matriz fila o columna que contiene a los datos ordenados según los valores numéricos o categorías ordenables de una [variable](#), en forma ascendente o descendente.

Asimetría: la distribución de los valores de una [variable cuantitativa](#) es asimetría cuando presenta una menor concentración de dichos valores en sólo uno de sus extremos o colas. Una distribución puede presentar asimetría a izquierda o asimetría a derecha.

Asimetría a izquierda: una distribución presenta [asimetría](#) a izquierda cuando los valores de la [variable](#) más alejados de la zona central se ubican a su izquierda.

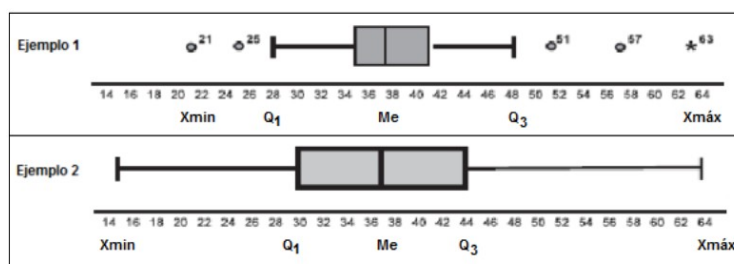


Asimetría a derecha: una distribución presenta asimetría a derecha cuando los valores de la *variable* más alejados de la zona central se ubican a su derecha



Atributo: cualidad o propiedad de la unidad de observación.

Box-Plot: es una representación gráfica de una variable cuantitativa que se construye en base a cinco elementos: el valor mínimo, el valor máximo, la mediana, el primer cuartil y el tercer cuartil de los datos. En él quedan simbolizados los datos centrales, datos adyacentes y datos raros (si hubiera). Nombre completo original: **Box-and-whisker plots**.



Censo: es el relevamiento de todas las unidades de observación de la población objetivo.

Coefficiente cíclico: es un índice cíclico expresado en forma de coeficiente.

Coefficiente de correlación (de Pearson): en el análisis de correlación, mide la fuerza y el sentido de la relación lineal entre dos variables. Su valor es un número real comprendido entre -1 y 1 . El valor 0 indica que no hay correlación lineal entre las variables, el valor -1 que existe una correlación lineal inversa perfecta y $+1$ una correlación lineal directa perfecta. El sentido (signo) es información que comparte con el coeficiente de regresión. Símbolos: r (estadístico muestral) y ρ (parámetro poblacional).

Coefficiente de desigualdad de Gini: es un indicador de la desigualdad en el reparto de un recurso, basándose en la distribución uniforme como modelo teórico en el reparto equitativo del recurso. Su valor es un número real comprendido entre 0 y 1 . El valor 0 indica una igualdad absoluta y el 1 una total desigualdad en la distribución del recurso. Nombre equivalente: coeficiente de concentración. Símbolo: G .

Coefficiente de determinación: en el análisis de regresión, indica en qué porcentaje las variaciones de una de las variables cuantitativas es explicada por las variaciones de la otra. Su valor es un número real comprendido entre 0 a 1 . Símbolos: r^2 o R^2 .

Coefficiente de regresión: en el análisis de regresión lineal bivariado, es la pendiente de la recta de regresión, indica en cuánto cambia en promedio la variable explicada por cada

unidad de cambio en la variable explicativa. Símbolos: b (estadístico muestral) y β (parámetro poblacional).

Coefficiente de variación (de Pearson): en una variable cuantitativa, es un coeficiente adimensional que se obtiene mediante el cociente entre el desvío estándar y la media aritmética. En un mismo grupo de unidades de observación, permite comparar la dispersión – relativa– entre las distribuciones de las correspondientes variables.

Coefficientes de asimetría: miden, con distintos criterios, el grado de asimetría de una distribución con respecto a una medida de tendencia central. Un coeficiente cero indica que la distribución es simétrica, un valor positivo indica que la distribución es asimétrica a derecha y uno negativo que la distribución es asimétrica a izquierda.

Coefficientes de curtosis: miden, con distintos criterios, el grado de curtosis de una distribución frecuencias o de una distribución de probabilidades con respecto a la distribución normal.

Componente cíclica: en las series de tiempo, describe el comportamiento de un ciclo periódico de una determinada longitud, si lo hubiere. En una misma serie de tiempo podrían superponerse ciclos periódicos de distintas longitudes.

Componente determinística: de un modelo estadístico, explica matemáticamente la ley que subyace detrás del comportamiento de las variables que participan en el modelo. Se la suele denominar modelo determinístico.

Componentes de una serie de tiempo: partes en las que se desglosa una serie de tiempo para su análisis. En una serie tratada en el dominio del tiempo las componentes son: tendencia, ciclos periódicos, ciclos no periódicos y ruido estadístico.

Componente estacional: de una serie de tiempo, es la componente cíclica cuya longitud temporal es de un año.

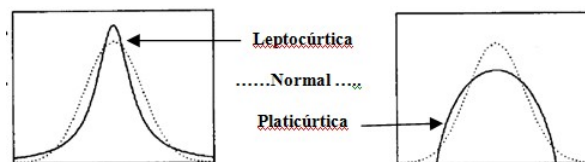
Componente estocástica: describe la parte aleatoria del modelo estadístico y se la identifica con el modelo aleatorio. Nombre equivalente: modelo estocástico.

Componente irregular: de una serie de tiempo, comprende los ciclos no periódicos y el ruido estadístico.

Covarianza: es una medida de dispersión que mide la variación conjunta entre dos variables. Su valor puede ser cero, positivo o negativo y está relacionado con el grado de dependencia lineal que puede existir entre las dos variables. Símbolos: S_{xy} (estadístico muestral) y σ_{xy} (parámetro poblacional).

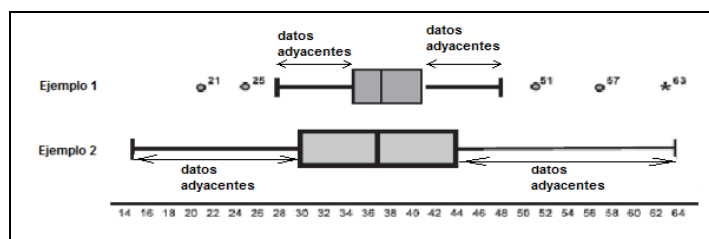
Cuartil: cada una de las tres marcas geométricas que dividen al arreglo de datos ordenado en cuatro partes iguales. El primer cuartil deja por debajo de sí la cuarta parte y por encima las tres cuartas partes del conjunto de datos. El segundo cuartil coincide con la mediana y el tercer cuartil deja por debajo de sí las tres cuartas partes de los datos y por encima la cuarta parte restante. Símbolos: Q_1 , Q_2 y Q_3 .

Curtosis o kurtosis: carácter vinculado a la forma que adopta la distribución de datos comparándola con la distribución normal de igual varianza. La distribución normal de referencia se denomina mesocúrtica; una distribución es platicúrtica si tiene un menor porcentaje de datos que la normal en un entorno de un desvío estándar alrededor de la media (menor concentración por lo tanto un menor apuntamiento en la zona central); y es leptocúrtica si presenta mayor porcentaje de datos en un entorno de un desvío estándar alrededor de la media (mayor concentración, es decir, mayor apuntamiento en la zona central).



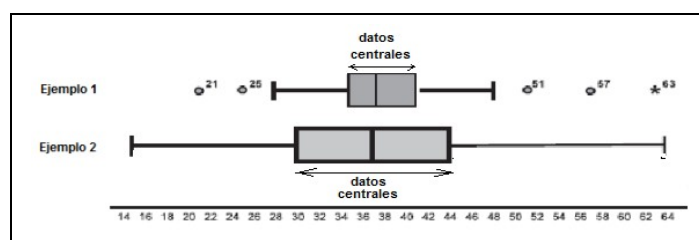
Dato: es el resultado obtenido al observar un atributo en una unidad de observación.

Dato adyacente: todo dato que se encuentra en las inmediaciones de la caja del box-plot, por fuera de ella y a ambos lados hasta una distancia de una vez y media el rango intercuartílico.

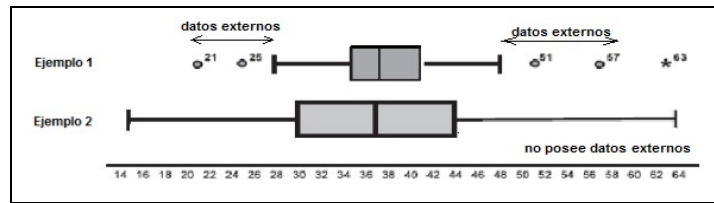


Dato bivariado: es el resultado obtenido al observar simultáneamente dos atributos en una unidad de observación.

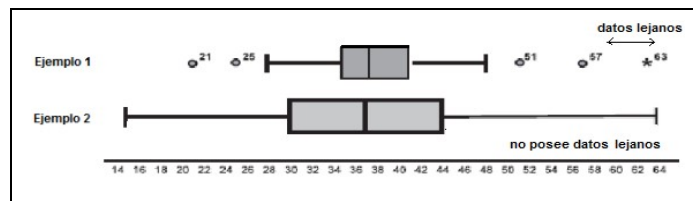
Dato central: todo dato que se ubica entre el primer y tercer cuartil. El conjunto de los datos centrales de una distribución constituye el 50% central del total y en el box-plot se representa gráficamente mediante la caja.



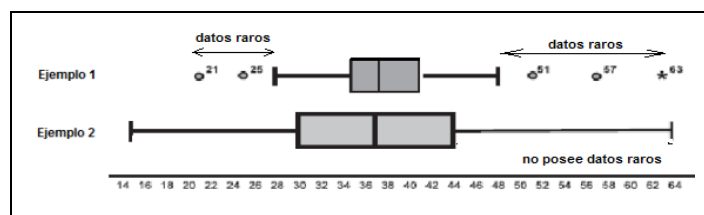
Dato externo: todo *dato raro* que se localiza a una distancia cuyo valor está comprendido entre una vez y media y tres veces el *rango intercuartílico* a ambos lados de la caja del *box-plot*.



Dato lejano: todo *dato raro* que se encuentra más allá de tres veces el *rango intercuartílico* de la caja del *box-plot*.



Dato raro: todo *dato* que se localiza a una distancia superior a una vez y media el *rango intercuartílico*, a ambos lados de la caja del *box-plot*. Un dato raro puede ser *externo* o *lejano*. Término equivalente: dato atípico. En la voz inglesa: outlier.



Datos crudos: *dato* colectados en un relevamiento que no han sido organizados o procesados. Nombre equivalente: datos en bruto.

Decil: cada una de las nueve marcas geométricas que dividen al *arreglo de datos ordenado* en diez partes iguales. Símbolos: D_i con i de 1 a 9.

Definición clásica de probabilidad (de Laplace): en un *experimento aleatorio*, la *probabilidad* de un *suceso* es el cociente entre el número de resultados favorables al suceso y el número de todos los resultados posibles del experimento. La condición de aplicabilidad de esta probabilidad es el cumplimiento del principio de la *equiprobabilidad*.

Definición frecuencial de probabilidad: la *frecuencia relativa* de un suceso dado observada bajo determinadas circunstancias puede considerarse, si permanecen dichas circunstancias, como la *probabilidad* de dicho suceso en un *experimento aleatorio*. Si un experimento aleatorio se realiza n veces bajo las mismas condiciones y un mismo suceso ocurrió $r \leq n$

veces entonces, una estimación de la probabilidad del suceso es r/n ; la verdadera probabilidad del suceso es el valor límite de r/n cuando n se aproxima al infinito.

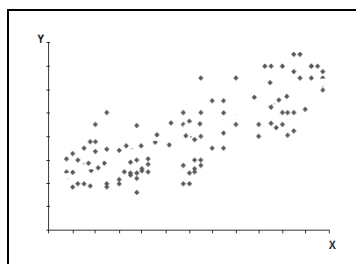
Desvío: diferencia entre el valor de un dato y una medida central de referencia de la distribución de datos.

Desvío estándar: mide la dispersión de los datos respecto de la media aritmética. Su magnitud es igual a la de la variable. Se obtiene como la raíz cuadrada de la varianza. En las distribuciones mesocúrticas describe un entorno alrededor de la media que encierra aproximadamente las dos terceras partes de todos los datos. Términos equivalentes: desviación estándar, desviación típica. Símbolos: s (estadístico muestral) y σ (parámetro poblacional).

Desvío medio (de Laplace): mide la dispersión, es el promedio de los valores absolutos de los desvíos de cada valor de la variable respecto a la media aritmética. Nombre equivalente: desviación media.

Diagrama de caja y bigotes: Box-Plot

Diagrama de dispersión: gráfico que permite visualizar la disposición conjunta de datos cuantitativos en el análisis bivariado.



Diseño de muestra: lineamientos para realizar un muestreo. Contempla las particularidades del fenómeno en estudio, la población objetivo asociada, el tamaño de la muestra y el método de muestreo adecuado.

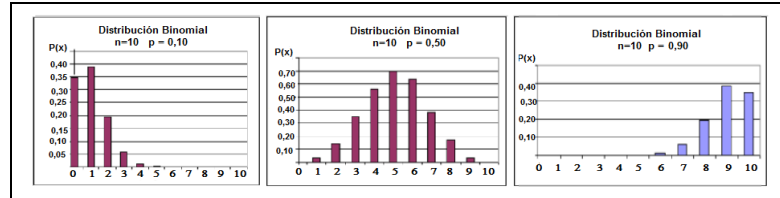
Dispersión: en las variables cuantitativas, expresa la disposición de los datos alrededor de una medida central de referencia (media aritmética, mediana o moda, según el criterio) o entre los dos valores extremos. La noción de dispersión se equipara con las de variedad, diversidad y heterogeneidad de un conjunto de datos.

Distancia intercuartílica: Rango intercuartílico

Distribución de Bernoulli: es la distribución de probabilidades de una variable aleatoria de Bernoulli con parámetro P . Donde P es la probabilidad de éxito y $(1-P)$ la probabilidad de fracaso. Su esperanza es $E(x) = P$ y su varianza es $V(x) = P \cdot (1-P)$.

Distribución binomial: es la distribución de probabilidades de una variable aleatoria Binomial con parámetros n y P . Su esperanza es $E(x) = n \cdot P$ y su varianza $V(x) = n \cdot P \cdot (1 - P)$.

Su función de probabilidad es $h(r) = P(X=r) = nCr \cdot P^r \cdot (1-P)^{n-r}$ para r perteneciente al recorrido de la variable. Su representación gráfica es:



Distribución de frecuencias: tabla donde se listan todos los valores o categorías de una variable junto con sus correspondientes frecuencias.

Distribución de frecuencias conjuntas: tabla de doble entrada donde se listan los valores de dos variables junto con sus frecuencias conjuntas absolutas o frecuencias conjuntas relativas o frecuencias conjuntas porcentuales. Nombre equivalente: tabla de distribución conjunta.

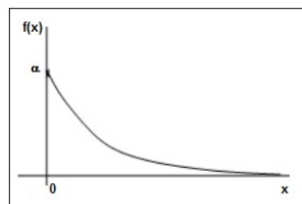
Distribución de Poisson: es la distribución de probabilidad de una variable aleatoria de Poisson. Su esperanza $E(x)$ es el parámetro. $\lambda = E(x) = \alpha \cdot t$ y la varianza $V(x) = \lambda$.

Su función de probabilidad $h(r) = P(X=r) = \frac{\lambda^r \cdot e^{-\lambda}}{r!}$ donde r es la cantidad de éxitos en un intervalo de medida t .

Distribución de probabilidad de una variable aleatoria discreta: tabla donde se listan todos los valores del recorrido de la variable aleatoria discreta junto con su función de probabilidad y su función de distribución acumulativa.

Distribución exponencial: es la distribución de probabilidad de una variable aleatoria exponencial. Su parámetro es α . Su esperanza es $E(x) = \frac{1}{\alpha}$ y su varianza $V(x) = \frac{1}{\alpha^2}$.

Su función de densidad de probabilidad es $f(x) = \begin{cases} \alpha e^{-\alpha x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$ y su representación gráfica

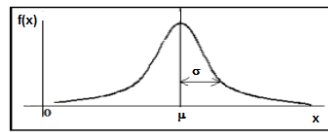


Distribución Hipergeométrica: es la distribución de probabilidad de una variable aleatoria hipergeométrica. Sus parámetros son la cantidad de éxitos en la población (R), el tamaño de la muestra (n) y el de la población (N). Su esperanza es $E(x) = n \cdot \frac{R}{N}$ y su varianza $V(x) = n \cdot \frac{R}{N} \cdot \frac{N-R}{N}$.

Su función de probabilidad es $P(x = r) = \frac{\binom{R}{r} \binom{N-R}{n-r}}{\binom{N}{n}}$ para todo r en el rango de la variable.

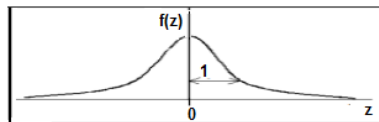
Distribución Normal: es la distribución de probabilidad de una variable aleatoria normal. Sus parámetros son μ y σ^2 . Su esperanza es $E(x) = \mu$ y su varianza $V(x) = \sigma^2$.

La función de densidad de probabilidad es $f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2}$ y su representación gráfica



Distribución Normal estándar: es la distribución de probabilidad de una variable aleatoria normal de media cero $\mu = 0$ y varianza uno $\sigma^2 = 1$. La función de densidad de probabilidad de

la distribución normal estándar es $f(z) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{1}{2} \cdot z^2}$ y su representación gráfica

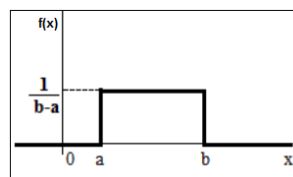


Distribución uniforme: es la distribución de probabilidad de una variable aleatoria uniforme continua que asigna probabilidades iguales a intervalos de igual amplitud. Su parámetro es el

rango de la distribución. Su esperanza es: $E(x) = \frac{a+b}{2}$ y su varianza $V(x) = \frac{(b-a)^2}{12}$.

Su función de densidad de probabilidad es $f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a; b] \\ 0 & \text{si } x \notin [a; b] \end{cases}$ y su representación

gráfica



Equiprobabilidad: es un principio de la probabilidad que enuncia la igualdad de oportunidad en la selección que debe tener cualquier elemento de una población respecto de los otros elementos.

Encuesta: instrumento diseñado para recolectar información primaria mediante preguntas abiertas o cerradas.

Error estadístico: es la diferencia entre un estadístico muestral y el parámetro poblacional.

Error de estimación: es el error estadístico cuando la referencia es el estadístico conocido.

Error estándar: término con que se nombra al desvío estándar de la distribución de un estadístico alrededor del parámetro correspondiente.

Error máximo: es el máximo error estadístico que puede ocurrir al asociársele una cierta probabilidad.

Error máximo del muestreo: respecto del parámetro conocido, es el error máximo al determinar un estadístico.

Error máximo de la estimación: respecto del estadístico determinado, es el error máximo al estimar un parámetro.

Error tipo I: es el error estadístico que se comete en una prueba de hipótesis al rechazar una hipótesis nula si fuera verdadera.

Error tipo II: es el error estadístico que se comete en una prueba de hipótesis al no rechazar una hipótesis nula si fuera falsa.

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

Esperanza: es el valor esperado de una variable aleatoria X , la media teórica, es un parámetro. Símbolo: $E(X)$. Término equivalente: esperanza matemática.

Esperanza de una variable aleatoria discreta X : es $E(X) = \sum_{x \in \mathfrak{R}(x)} x \cdot h(x)$ donde $h(x)$ es la función de probabilidad de X .

Esperanza de una variable aleatoria continua X : es $E(X) = \int_{-\infty}^{\infty} x \cdot f(x)$ donde $f(x)$ es la función de densidad de probabilidad de X .

Estadística: La función principal de la estadística es elaborar principios y métodos que nos ayuden a tomar decisiones frente a la incertidumbre. (Chou Ya Lun, 1991).

Es el arte y la ciencia de extraer información útil de datos empíricos (Hampel F.R., 1986).

Ciencia encargada de recolectar, analizar, presentar e interpretar datos. (Enciclopedia Británica, 2010).

Estadística descriptiva: área de la [estadística](#) que elabora las técnicas y métodos para la descripción de un conjunto de datos a partir de la determinación de una serie de indicadores que permiten caracterizarlo.

Estadística inferencial: área de la [estadística](#) que, basada en la teoría de la probabilidad, se encarga de proyectar a la población objetivo las características de una muestra representativa de la misma.

Estadísticas: hechos reunidos y cuantificados sistemáticamente en cualquier campo del conocimiento.

Estadístico: cualquier [medida](#) o [indicador](#) muestral. Algunos símbolos particulares son: \bar{X} ([media aritmética](#) muestral), s ([desvío estándar](#) muestral) y s^2 ([varianza](#) muestral).

Estadístico de prueba: en una [prueba de hipótesis](#) es el [estadístico](#) muestral [estandarizado](#) según la distribución utilizada para la prueba.

Estandarización: procedimiento mediante el cual la distribución de una [variable aleatoria](#) se ajusta a una distribución típica o estándar.

Estandarización normal: si X es una variable aleatoria distribuida normalmente su [estandarización](#) resulta de: $Z = \frac{X - \mu_x}{\sigma_x}$, donde Z es una [variable aleatoria](#) con [distribución normal estándar](#).

Estimación de parámetros: es un procedimiento de la [estadística inferencial](#) por medio del cual, partiendo de [estadísticos](#) muestrales, se obtienen conclusiones acerca de los [parámetros](#) poblacionales.

Estimación por intervalos de confianza: es la [estimación de un parámetro](#) mediante un [intervalo de confianza](#).

Estimación puntual: es adoptar como [parámetro](#) poblacional el valor del [estimador](#) muestral.

Estimador: [estadístico](#) utilizado en la [inferencia estadística](#). Un estimador óptimo es [insesgado](#), eficiente, consistente y suficiente. Símbolo: $\hat{\theta}$.

Estimador consistente: es el estimador que tiende al valor del parámetro al aumentar la muestra.

Estimador eficiente: es el estimador de menor varianza dentro del conjunto de todos los estimadores posibles.

Estimador insesgado: es el estimador cuyo valor esperado coincide con el parámetro.
Símbolo: $E(\hat{\theta}) = \theta$.

Estimador suficiente: es el estimador que transmite la máxima información posible de una muestra utilizando todos sus elementos.

Estocástico: aleatorio. =

Evento aleatorio: suceso aleatorio.

Experimento aleatorio: experimento donde no se puede determinar a priori cuál va a ser su resultado. Términos equivalentes: experimento probabilístico o estocástico, fenómeno aleatorio.

Experimento binomial: es un experimento aleatorio compuesto por repetición de un experimento simple de *Bernoulli*, donde las n repeticiones son *independientes*.

Experimento de Bernoulli: experimento aleatorio simple que tiene sólo dos resultados posibles (o sus valores se han dicotomizado), donde se identifica como *éxito* a uno y *fracaso* al otro.

Experimento hipergeométrico: es un experimento aleatorio compuesto por repetición de un experimento simple de Bernoulli, donde las n repeticiones son dependientes.

Experimento normal: es un experimento aleatorio donde la mayor parte de sus resultados numéricos reales se concentra alrededor de un valor esperado y el resto, al alejarse de la zona central, van disminuyendo su concentración paulatinamente hacia ambos extremos, adoptando además la distribución de todos los resultados una forma acampanada.

Experimento probabilístico: experimento aleatorio

Experimento uniforme: es un experimento aleatorio donde todos los resultados numéricos reales tienen una concentración pareja dentro de un intervalo cerrado.

Fenómeno aleatorio: experimento aleatorio.

Fractil: lugar geométrico en el arreglo de datos ordenados que lo divide en dos partes y permite determinar el máximo de la fracción inferior o el mínimo de la fracción restante.

Frecuencia absoluta acumulada: cantidad de datos ordenados que se acumulan hasta un cierto valor de la variable, incluyendo a todos los que le anteceden.

Frecuencia absoluta simple: cantidad de veces que aparece un valor de la variable en el grupo de datos estudiado. Símbolo: f

Frecuencia conjunta absoluta: cantidad de veces con que, en un grupo de datos bivariados, aparecen simultáneamente los valores correspondientes a cada una de las dos variable intervinientes.

Frecuencia conjunta porcentual: peso porcentual con que, en un grupo de datos bivariados, aparecen simultáneamente los valores correspondientes a cada una de las dos variable intervinientes.

Frecuencia conjunta relativa: peso relativo con que, en un grupo de datos bivariados, aparecen simultáneamente los valores correspondientes a cada una de las dos variables intervinientes.

Frecuencia porcentual acumulada: cantidad porcentual de datos ordenados que se acumulan hasta un cierto valor de la variable, incluyendo a todos los que le anteceden.

Frecuencia porcentual simple: peso porcentual de un valor de la variable respecto al total. Símbolo: $f_{\%} = f_r \cdot 100$

Frecuencia relativa acumulada: cantidad relativa de datos ordenados que se acumulan hasta un cierto valor de la variable, incluyendo a todos los que le anteceden.

Frecuencia relativa simple: peso relativo de un valor de la variable respecto a la cantidad total de los datos. Símbolo: $f_r = \frac{f}{n}$

Función de densidad de probabilidad: es una función real que modela la distribución de probabilidad de una variable aleatoria continua. La función de densidad de probabilidad siempre toma valores no negativos y el área comprendida bajo su gráfico sobre todos los valores de x siempre es igual a 1. Esta función no asigna probabilidades pero sí permite calcularlas mediante su función de distribución acumulativa para cualquier intervalo del recorrido de la variable. Símbolo: $f(x)$

Función de distribución acumulativa: para un valor real x de una variable aleatoria X , calcula la probabilidad acumulada hasta él. Símbolo: $F(x)$, donde: $F(x) = P(X \leq x)$.

Función de probabilidad: es una función que asigna una probabilidad a cada valor de una variable aleatoria discreta. Símbolo: $h(x)$

Gráfico de barras agrupadas: es un gráfico bivariado adecuado para representar la información de una tabla conjunta, utilizando cualquiera de las frecuencias simples.

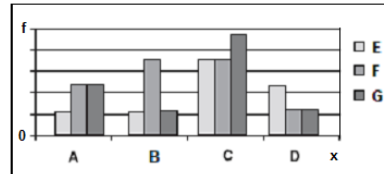


Gráfico de barras segmentadas: es un gráfico bivariado adecuado para representar la información de una tabla conjunta, que permite visualizar para cada valor de una de las variables las particiones según los valores de la otra, utilizando cualquiera de sus frecuencias simples.

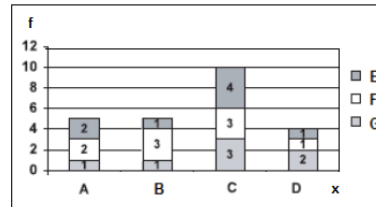


Gráfico de barras separadas: gráfico aplicable a una variable cualitativa para representar cualquiera de sus frecuencias simples.

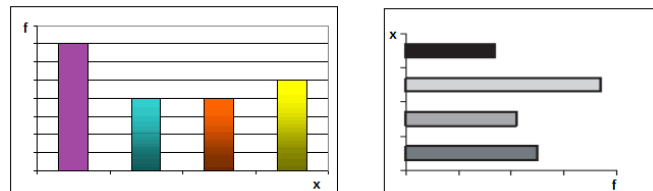


Gráfico de bastones: corresponde a cualquiera de las frecuencias simples de una variable cuantitativa bajo tratamiento discreto.

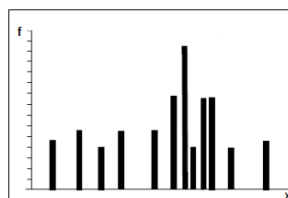


Gráfico de Caja y Bigotes: Box-Plot

Gráfico de torta: está dividido en sectores donde cada porción representa la proporción con que cada valor de la variable incide en el total. Puede aplicarse a cualquier tipo de variable. Términos equivalentes: gráfico circular o de pastel.

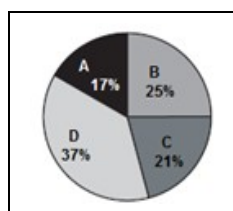
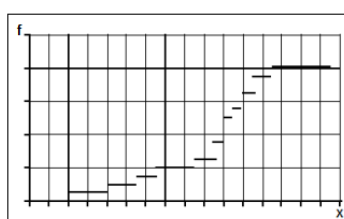


Gráfico escalonado (o de escalones): corresponde a cualquiera de las frecuencias acumuladas de una variable cuantitativa tratada en forma discreta.

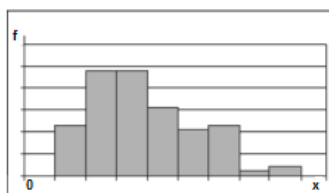


Hipótesis alternativa: es la hipótesis estadística que se opone a la hipótesis nula. Puede existir más de una hipótesis alternativa que se oponga a la misma hipótesis nula. Símbolo: H_A ó $H_1, H_2 ..H_i$.

Hipótesis estadística: supuesto, proposición o enunciado respecto a alguna característica de la población objetivo basado en la evidencia muestral.

Hipótesis nula: es la hipótesis estadística que se somete a prueba cuando se realiza una prueba de hipótesis. Símbolo: H_0 . Término equivalente: hipótesis principal

Histograma de Pearson: gráfico de barras adyacentes que puede construirse con cualquiera de las frecuencias simples y que se aplica a variables cuantitativas organizadas en intervalos de clase.



Indicador estadístico: es una medida estadística interpretada en un contexto particular concreto de índole económica, social, etc.

Índice: es la medida de intensidad que permite medir la variación relativa de la ocurrencia de los valores de una variable entre una situación dada y otra de referencia o base. Las

situaciones pueden ser temporales, espaciales o de cualquier otra índole. Para variables económicas el término ocurrencia se refiere tanto a la frecuencia, como a la cantidad o al precio de un producto o un grupo de productos. Término equivalente: número índice.

Índice compuesto: índice referido a un grupo de valores de la variable.

Índice de agregados de precios ponderados: índice compuesto que resulta del cociente entre las sumas o agregaciones de precios ponderados por cantidades.

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

Índice de costos de la construcción: es un índice compuesto general que mide la evolución de los costos de los materiales de la construcción. Acrónimo: ICC.

Índice de Fisher: índice de agregados de precios ponderados que se obtiene como la media geométrica entre el Índice de Laspeyres y el de Paasche. En símbolos: I_F . Se calcula mediante la fórmula: $I_F = \sqrt{I_L \cdot I_P}$

Índice deflactor: 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).

Índice de Laspeyres: índice de agregados de precios ponderados donde como criterio de ponderación se utilizan las cantidades de la situación base. En símbolos: I_L

$$I_L = \frac{\sum p_{ji} \cdot q_{j0}}{\sum p_{j0} \cdot q_{j0}} \cdot 100 \quad \text{Donde: } p_{j0}: \text{ precio del ítem "j" en el periodo base "o". } p_{ji}: \text{ precio del ítem "j" en el periodo "i"}$$

Índice de Paasche: : índice de agregados de precios ponderados donde como criterio de ponderación se utilizan las cantidades de la situación dada. En símbolos: I_P

$$I_P = \frac{\sum p_{ji} \cdot q_{ji}}{\sum p_{j0} \cdot q_{ji}} \cdot 100 \quad \text{Donde: } p_{j0}: \text{ precio del ítem "j" en el periodo base "o". } p_{ji}: \text{ precio del ítem "j" en el periodo "i"}$$

Índice de precios al consumidor: índice compuesto que se usa para analizar la evolución del costo de vida a partir de los precios de una canasta de productos y servicios. Acrónimo: IPC.

Índice de precios industriales mayoristas: índice compuesto que mide la evolución de los precios de los productos industriales de origen nacional e importado excepto los de la construcción. En realidad es un sistema de varios índices. Acrónimos: IPIM, SIPIM

Índices cíclicos: conjunto de índices con los que se describe una determinada componente cíclica de una serie de tiempo. Si la serie tuviera más de una componente cíclica, para cada una de ellas habría un conjunto de índice cíclicos.

Índice simple: índices referido a un solo valor de la variable.

Inferencia estadística: procedimiento por medio del cual se proyectan hacia la población objetivo los resultados de una muestra representativa.

Información estadística: toda información basada en datos empíricos recolectados y/o tratados con metodología estadística. Comprende tanto la información muestral como la información población.

Instituto Nacional de Estadística y Censos: organismo coordinador del Sistema Estadístico Nacional. Acrónimo: INDEC.

Instrumentos de medición: herramientas para la recolección de los datos como ser: encuestas, archivos, fichas, entrevistas, información brindada por organismos oficiales, aparatos de medición, etc. Se utilizan tanto en censos como en trabajo sobre muestras.

Intervalo de confianza: intervalo real en el recorrido de la variable que contiene el valor del parámetro con un cierto nivel de confianza.

Intervalos de clase: secciones en que puede dividirse el recorrido de una variable cuantitativa al agrupar los datos.

Ley de los grandes números: es un conjunto de leyes estadísticas que expresan la convergencia de estadísticos muestrales o sus distribuciones de probabilidad cuando crece el número de observaciones ($n \rightarrow \infty$).

Marca de clase: es un punto, usualmente el punto medio, elegido como representante del intervalo de clase.

Matriz de datos: es un arreglo bidimensional que incluye conjuntamente los atributos relevados y las unidades observadas, y donde cada fila representa una unidad de observación y cada columna un atributo variable, o a la inversa.

Media aritmética: en variables cuantitativas es la medida de posición que representa el punto de equilibrio o centro de masas del conjunto de datos. Se la determina mediante cálculo en el dominio de los números reales dividiendo la suma de todos los valores de los datos observados, por la cantidad total de datos. Sinónimos: promedio, media (a secas). Símbolos: \bar{x} (media aritmética muestral) y μ (media aritmética poblacional).

Media de índices simples (de precios o cantidades): es un índice compuesto y resulta de promediar los índices simples correspondientes a cada valor de la variable. Nombres equivalentes: promedio de relativos simples o media de relativos simples.

Media de relativos simples: [media de índices simples](#)

Media geométrica: es la n -ésima raíz del producto de los valores correspondientes a n datos numéricos.

Mediana: como [medida de posición](#) es el lugar geométrico del [arreglo de datos ordenados](#), que lo divide en dos partes iguales. Símbolo: Me

Medida estadística: toda medida que permite caracterizar un conjunto de datos en estudio.

Medidas de dispersión: [medidas estadísticas](#) que miden, bajo distintos criterios, el grado de [dispersión](#). Algunas medidas de dispersión son: [rango](#), [rango intercuartílico](#), [desvío estándar](#), [desvío medio](#), [varianza](#) y [covarianza](#). Término equivalente: medidas de variabilidad.

Medidas de forma: son medidas que permiten caracterizar, con distintos criterios, la forma asumida por una distribución de datos, como el [Coeficiente de asimetría](#) y el [Coeficiente de curtosis](#).

Medidas de intensidad: [medidas estadísticas](#) que miden de varias maneras la relación entre los distintos valores de una variable o de un mismo valor de la misma pero en diferentes situaciones (espaciales o temporales). Son de este tipo de medidas los [números índices](#), las [razones](#) y las [proporciones](#). Nombre equivalente: tasas._

Medidas de posición: [medidas estadísticas](#) que determinan, con distintos criterios, ubicaciones unívocas en el conjunto de datos. Son ejemplos de ellas: [media aritmética](#), [moda](#), [mediana](#), [fractiles](#), [percentiles](#), [deciles](#) y [cuartiles](#).

Medidas de tendencia central: permiten identificar la zona de centralidad en una distribución de datos. Ellas son la [media](#), la [moda](#) y la [mediana](#). En el caso de una variable cuantitativa su distribución se considera simétrica cuando esas tres medidas coinciden.

Medidas de variabilidad: [Medidas de dispersión](#).

Metodología estadística: conjunto de principios y reglas de procedimientos propios de la [estadística](#), que atañen a la recolección de datos, su organización, la extracción de información útil y su empleo en la inferencia.

Métodos de muestreo: los métodos pueden ser probabilístico o no probabilísticos. En los métodos de [muestreo probabilístico](#) se hace intervenir el azar en la selección de cada una de las unidades de observación integrantes de la muestra cosa que no ocurre en los no probabilísticos.

Moda: es el valor o categoría de la variable que ocurre con mayor frecuencia. En un conjunto de datos puede no haber moda o puede haber una o más. Si hay moda, es el valor o valores de la variable que predomina/n o prevalece/n respecto de los demás. Sinónimo: modo. Símbolo: Mo.

Modelo aleatorio: [componente estocástica](#).

Modelo determinístico: [componente determinística](#).

Modelo estadístico: en un fenómeno susceptible de ser estudiado estadísticamente, es un modelo que contiene una [componente determinística](#) y una [componente estocástica](#).

Modelos matemáticos de una serie cronológica: analizada en el dominio del tiempo, describen cómo se vinculan las componentes de la serie. Dependiendo de las características del fenómeno en estudio, el modelo matemático puede ser multiplicativo o aditivo según si interactúan o no entre sí las componentes.

Muestra: es un subconjunto de la [población objetivo](#).

Muestra representativa: es una [muestra](#) recolectada aplicando un método de [muestreo probabilístico](#) adecuado lo que garantizaría que la muestra pueda reflejar las características de la población objetivo.

Muestreo: es el proceso conducente a la obtención de una [muestra](#).

Muestreo al azar simple: [muestreo probabilístico](#) mediante el cual se seleccionan las [unidades de observación](#) directamente desde la [población](#).

Muestreo estratificado: [muestreo probabilístico](#) que se aplica a poblaciones que dada la información que se tiene de ellas se pueden dividir en grupos o clases denominadas estratos, en cada estrato se selecciona una submuestra y la agregación de todas ellas conforman la muestra definitiva. La selección de los elementos de cada submuestra puede ser proporcional o no al tamaño del estrato.

Muestreo por conglomerados: [muestreo probabilístico](#) que se aplica en poblaciones que se subdividen en partes denominadas conglomerados donde cada uno refleja a la población. La muestra se conforma por la elección al azar de unidades de observación de cada conglomerado que fue seleccionado aleatoriamente.

Muestreo probabilístico: [muestreo](#) basado en el principio de la [equiprobabilidad](#) al momento de la elección de cada una de las [unidades de observación](#) que integrarán la [muestra representativa](#) de la población. Algunos de los métodos de muestreo probabilístico son:

muestreo al azar simple, muestreo sistemático o secuencial, muestreo estratificado y muestreo por conglomerados.

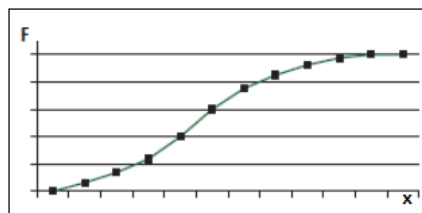
Muestreo sistemático o secuencial: muestreo probabilístico que se aplica a poblaciones resultantes de fenómenos que se desarrollan con un cierto orden temporal o espacial, sobre el que se fija una secuencia para la selección de las unidades de observación. Tanto la longitud de la secuencia como el inicio de la misma se determinan aleatoriamente.

Nivel de confianza: es la probabilidad asociada al error máximo de la estimación. Símbolo: $1 - \alpha$.

Nivel de significación: de una prueba de hipótesis, es la probabilidad asociada a la región crítica y en consecuencia la probabilidad de cometer un error de tipo I. Símbolo: α .

Número índice: índice.

Ojiva de Galton: es la representación gráfica acumulada de una variable cuantitativa, con tratamiento continuo de las frecuencias acumuladas absolutas o relativas o porcentuales. Término equivalente: Polígono de frecuencias acumulativas.

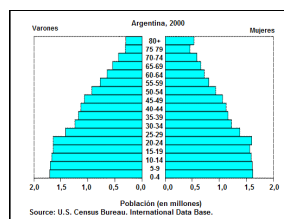


Outlier: dato raro.

Parámetro: indicador poblacional. Símbolo genérico: θ . Algunos símbolos particulares son: μ (media aritmética poblacional), σ (desvío estándar poblacional) y σ^2 (varianza poblacional).

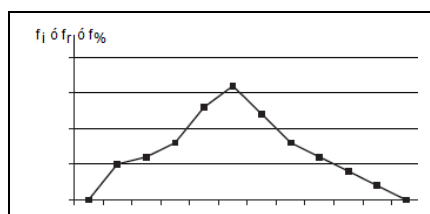
Percentil k: es la marca geométrica en el arreglo de datos ordenado, que deja por debajo de él, el k% de los datos. El percentil es el fractil expresado en porcentaje. Símbolo: P_k .

Pirámide poblacional: es un gráfico bivariado que involucra a las variables sexo y edad de una población.



Población objetivo: totalidad de individuos o elementos a estudiar; o todas las manifestaciones producto de un fenómeno susceptible de ser estudiado estadísticamente. Término equivalente: universo.

Polígono de frecuencias: es el gráfico determinado por los segmentos que unen las marcas de clase de los intervalos correspondientes a la organización de la distribución en intervalos de clase, considerado agregándole las marcas de clase del intervalo previo al primero y del siguiente al último con frecuencia cero, por no pertenecer al rango de valores de la variable. El área que encierra el polígono de frecuencias es igual al área del histograma del mismo agrupamiento.



Predicción estadística: es el resultado de aplicar métodos de interpolación/extrapolación a las componentes de un modelo estadístico.

Probabilidad: es la medida de la ocurrencia de un suceso aleatorio. Es un número real comprendido entre 0 y 1 (incluidos ambos). Se expresa como: $P(A)$ e indica la probabilidad del suceso A

Probabilidad compuesta: probabilidad conjunta.

Probabilidad condicional: es la probabilidad de ocurrencia de un suceso aleatorio condicionada por la ocurrencia de otro u otros sucesos. Se expresa: $P(A/B)$ e indica la probabilidad del suceso A si el suceso B ocurriera.

Probabilidad conjunta: es la probabilidad de que dos o más suceso aleatorio ocurran simultáneamente. En símbolos: $P(A \text{ y } B)$ o también $P(A \cap B)$. Nombre equivalente: probabilidad compuesta.

Probabilidades marginales: son las probabilidades totales en una tabla de contingencias.

Probabilidad total: es la probabilidad del suceso resultante de la unión de dos o más suceso aleatorio.

Proceso de Poisson: es un experimento aleatorio de naturaleza binomial, donde los éxitos ocurren a lo largo del continuo y el fracaso es la no ocurrencia de éxito. Está caracterizado por el número α , que es la intensidad media de ocurrencias de éxito por unidad de continuo. Nombre equivalente: proceso poissoniano.

Promedio: este concepto es equivalente al de [media aritmética](#), aunque en algunos casos su uso involucra a todas las [medidas de tendencia central](#) en forma indistinta.

Promedio de relativos simples: [media de índices simples](#).

Pronóstico estadístico: [predicción](#) realizada en base a una [serie de tiempo](#).

Proporción: [medida de intensidad](#) que representa el peso que tiene en el conjunto de datos, un valor o un grupo de valores de una o más variables. Es decir que indica cuánto incide ese valor o grupo de valores en el total. Se expresa como un coeficiente y se suele interpretar como un porcentaje.

Prueba de hipótesis: procedimiento de la [inferencia estadística](#) mediante el cual se pone a prueba una hipótesis de trabajo denominada [hipótesis nula](#) frente a una [hipótesis alternativa](#). La prueba de hipótesis se basa en la evidencia muestral y se fundamenta en la teoría de la probabilidad. Una prueba de hipótesis concluye con la decisión de rechazar o no rechazar la hipótesis nula con determinado [nivel de significación](#). Nombres equivalentes: test de hipótesis o docimasia estadística.

Prueba de hipótesis paramétrica: [prueba de hipótesis](#) que somete a prueba un valor supuesto del parámetro poblacional cuyo verdadero valor es en realidad desconocido utilizando un [estimador](#) muestral de distribución conocida.

p-valor: es el área de probabilidad determinada por el [estadístico de prueba](#) y ubicada hacia su derecha o hacia su izquierda dependiendo de la ubicación de la [región crítica](#) de la [prueba de hipótesis](#). En la voz inglesa: p-value.

p-value: [p-valor](#)

Quintil: cada una de las cuatro marcas geométricas que dividen al [arreglo de datos ordenado](#) en cinco partes iguales. Término equivalente: quintilo.

Rango: es una [medida de dispersión](#) que mide la amplitud o extensión total del recorrido de una [variable cuantitativa](#). Al tener en cuenta sólo los valores máximo y mínimo de la variable no es un estimador suficiente. Nombre equivalente: amplitud de la distribución. En símbolo: $R_g = X_{\max} - X_{\min}$

Rango intercuartílico: es una [medida de dispersión](#) que mide la amplitud de la fracción del recorrido de una [variable cuantitativa](#) que se extiende entre el primero y tercer [cuartil](#). En símbolos: $d = Q_3 - Q_1$. Nombres equivalentes: amplitud intercuartílica o distancia intercuartílica.

Razón: medida de intensidad que representa el peso que tiene sobre un determinado valor o grupo de valores otro valor o grupo de valores de una o más variables. Es decir que indica cuánto incide un valor o grupo de valores de una variable sobre otro.

Razonamiento estadístico: es el razonamiento que permite al individuo reconocer la necesidad de los datos, lograr comprender un fenómeno real a través de su modelización mediante tablas, gráficos, indicadores, etc. y poder explicar a otros los significados extraídos del modelo; diferenciando los datos propiamente dichos del modelo de los mismos. Además le permite percibir y comprender la variación que hay y se transmite en los datos, así como la incertidumbre originada por la variación no explicada. La estadística permite hacer predicciones, buscar explicaciones y causas de la variación y aprender del contexto.

Recorrido de una variable: conjunto de valores que asume una variable cuantitativa.

Recta de tendencia: es un modelo matemático que permite describir la tendencia de una serie de tiempo.

Relativo simple: índice simple.

Región crítica: en una prueba de hipótesis es la región del recorrido del estadístico muestral asociada con la hipótesis alternativa. Se la suele denominar también zona de rechazo de la hipótesis nula.

Serie cronológica: serie de tiempo

Serie de tiempo: es el conjunto de los datos bivariados de una variable estadística registrados secuencialmente en el tiempo. Nombre equivalente: serie cronológica.

Sistema Estadístico Nacional Argentino: organismo interprovincial coordinado por el INDEC e integrado por todos los organismos oficiales productores de estadísticas del país. Su funcionamiento está regulado por ley (N° 17.622/68) y se estructura sobre la base de la descentralización. Acrónimo: SEN.

Suceso aleatorio: cualquier subconjunto de un espacio muestral. Término equivalente: evento aleatorio.

Suceso cierto: suceso aleatorio del que se tiene certeza de que va a ocurrir. Coincide con el espacio muestral.

Suceso elemental: suceso aleatorio unitario de un espacio muestral.

Suceso imposible: suceso aleatorio del que se tiene certeza de que no va a ocurrir.

Sucesos excluyentes: sucesos aleatorios que no pueden ocurrir simultáneamente. Nombre equivalente: sucesos mutuamente excluyentes.

Sucesos independientes: dos sucesos aleatorios son independientes cuando ocurrencia de uno de ellos no incide en la ocurrencia del otro.

Sucesos mutuamente excluyentes: sucesos excluyentes.

Tabla de contingencias: es una distribución de frecuencias conjuntas relativas que en virtud de la definición frecuencial de probabilidad puede aplicarse a un determinado experimento aleatorio.

Tabla de distribución conjunta: distribución de frecuencias conjuntas.

Tasa: medida de intensidad

Técnica de muestreo: método de muestreo.

Tendencia: es la componente que modela el comportamiento de la serie de tiempo a largo plazo. Nombre equivalente: tendencia secular.

Teorema Central del Límite (TCL): formaliza el comportamiento asintóticamente normal, bajo determinadas condiciones, de la distribución de la suma de un gran número de variables aleatorias, en particular el de las variables aleatorias media muestral y proporción muestral.

Test de hipótesis: prueba de hipótesis.

Unidad de análisis estadística: conjunto de unidades de observación agrupadas bajo un determinado criterio. En muchos casos se la utiliza como nombre equivalente al de unidad de observación.

Unidad de observación: cada individuo o elemento de la población objetivo.

Valor atípico: dato raro

Valor esperado de una variable aleatoria: esperanza

Valor p: p-valor

Variable: variable estadística

Variable aleatoria: asigna valores numéricos reales a los sucesos definidos en el espacio muestral asociado a un experimento aleatorio. Es decir que es una función del espacio muestral sobre el espacio de los números reales. Nombre equivalente: variable estocástica.

Variable aleatoria continua: variable aleatoria cuyo recorrido es un intervalo de números reales.

Variable aleatoria discreta: variable aleatoria cuyo recorrido es un subconjunto de los números enteros.

Variable Bernoulli: en el experimento de Bernoulli, es la variable aleatoria discreta que asigna 0 al fracaso y 1 al éxito. Su recorrido $\mathfrak{R}(x) = \{0; 1\}$.

Variable Binomial: es una variable aleatoria discreta que cuenta la cantidad de éxitos en n repeticiones independientes de un experimento de Bernoulli. Su recorrido es el intervalo entero $[0, n]$.

Variable cualitativa o categórica: variable estadística que adopta valores no numéricos que usualmente se denominan categorías.

Variable cualitativa no ordenable: variable cualitativa cuyos valores no son intrínsecamente pasibles de ordenarse.

Variable cualitativa ordenable: variable cualitativa cuyos valores son intrínsecamente pasibles de ser ordenados

Variable cuantitativa: variable estadística que adopta valores numéricos.

Variable cuantitativa continua: variable cuantitativa cuyos valores son números reales que provienen de algún mecanismo de medición.

Variable cuantitativa discreta: variable cuantitativa cuyos valores provienen de un mecanismo de conteo; en consecuencia asume un número finito o infinito numerable de valores.

Variable dependiente: variable cuyos valores dependen de los valores de otra variable.

Variable de Poisson: en un proceso de Poisson, es la variable aleatoria discreta que cuenta la cantidad de éxitos en un intervalo continuo de medida t . Su recorrido es el conjunto de los números naturales y el cero.

Variable dicotómica: variable estadística que por su propia naturaleza adopta sólo dos valores o categorías.

Variable estadística: se construye a partir de uno o varios atributos variables observados en las unidades de observación.

Variable exponencial: es la variable aleatoria continua, que a partir del último éxito ocurrido dentro de un proceso de Poisson, mide el intervalo continuo que transcurriría hasta el siguiente éxito. Su recorrido es el conjunto de los números reales mayores o iguales a cero.

Variable hipergeométrica: es una variable aleatoria discreta que cuenta el número de éxitos obtenidos en n repeticiones dependientes de un experimento de Bernoulli. En particular es adecuada para el muestreo sin reposición en poblaciones pequeñas. Su recorrido es el intervalo entero $[0; n]$.

Variable independiente: variable cuyos valores no dependen de los valores de otra variable.

Variable normal: es la variable aleatoria continua que surge de un experimento normal y cuyo recorrido es el conjunto de los números reales.

Variable uniforme: es la variable aleatoria continua asociada a un experimento uniforme y cuyo recorrido es un intervalo cerrado en el conjunto de los números reales.

Varianza (de Gauss): es una medida de dispersión que se determina como la media aritmética de los cuadrados de los desvíos. Su magnitud es el cuadrado de la magnitud de la variable. Nombres equivalentes: desviación cuadrática media o variancia. Símbolos: s^2 (estadístico muestral) y σ^2 (parámetro poblacional).

Bibliografía

- Ardila A., El lenguaje Matemático y el Usual, como Mediador de la Comunicación. Acta Latinoamericana de Matemática Educativa, Vol 15, Universidad de Panamá. (1991), 1161-1171.
- Argibay, Juan Carlos. “Subjetividad y procesos cognitivos”. Subj. procesos cogn. Vol.13 no.1 Ciudad Autónoma de Buenos Aires. ARTÍCULOS ORIGINALES. Muestra en investigación cuantitativa. *The sample in quantitative research. Versión* ISSN 1852-7310. (ene./jun. 2009)
- Batanero, Carmen. “Los retos de la cultura estadística”. Congreso de Estadística. Caseros, Argentina, (2003).
- Batanero, Carmen. Didáctica de la Estadística. GEEUG (Grupo de Educación Estadística Universitaria de Granada). Publica: Grupo de Investigación Estadística, Departamento de Didáctica de la Matemática, Universidad de Granada. Granada. (2001).
- Batanero, Carmen. Conferencia organizada por IASE sobre la formación de investigadores. (2001)
- Bressan, Ana y Bressan, Oscar. “Probabilidad y estadística: como trabajar con niños y jóvenes”. Ed. Novedades Educativas. Argentina. (2008).
- Chou. Ya-Lun. “Análisis Estadístico”. Ed. McGraw-Hill. Segunda Edición. (1992)
- Diccionario de Estadística Práctica. Universidad Interamericana de Panama. Lic. Admin de Recursos Humanos. (2007)
- Ferreiro, Osvaldo y Fernández de la Reguera, Pedro. “La estadística, una ciencia de la controversia”. Artículo publicado en la Revista Universitaria N° 25 del Instituto de Matemáticas y Física .Universidad de Talca. Chile. (1988)
- Guzmán, Miguel de. “Enfoque heurístico de la enseñanza de la matemática, Aspectos didácticos de matemáticas”. Publicaciones del Instituto de Ciencias de la Educación de la Univ. de Zaragoza. (1985)
- Guzmán, Miguel de. “Para pensar mejor”. Ed. Labor, Barcelona. (1991)
- Kreyszig, Erwin. “Introducción a la Estadística Matemática. Principios y métodos”. Ed. LIMUSINA, S.A. México, Quinta reimpression, 1981.
- Meyer, Paul. “Probabilidad y aplicaciones estadísticas”. Ed. Addison-Wesley-Iberoamericana. Edición revisada USA. (1992)
- MODULO 1. ESTADISTICA DESCRIPTIVA. Glosario. SEQC. (2009)
- Moroney. M. S. Hechos y estadística. Ed. EUDEBA. (1979).
- Juan Ignacio Piovani, Juan Ignacio. XXV Congreso ALAS. Universidade Federal do Rio Grande do Sul. Porto Alegre. “La estadística y el sentido común de los sociólogos”. *Università degli studi di Bologna, Representación en Argentina. Universidad Nacional de La Plata.* (22-26 de agosto de 2005).
- Pozo, José I. “Comprender y transformar la enseñanza”. Ed. Morata. Madrid. (1993)
- Santaló Luis A.; Gálvez, Grecia; Charnay, Roland; Brousseau, Guy; Lerner, Delia y Sadosky, Patricia. Parra, Cecilia y Saiz, Irma (compiladoras). “Didáctica de matemáticas”. Aportes y reflexiones. Ed. Paidós Educador. Argentina. 5ta reimpression, (1997).
- Santaló, L.A., “La educación matemática, hoy”. Ed. Teide, Barcelona. (1975)

- Sierra Bravo, R. “Técnicas de Investigación Social. Teoría y Ejercicios”. Decimosegunda Edición. Ed. Paraninfo. (1998)
- SIEL (Sistema de Información Estadística Local), cuadernos. Serie 1. Materiales institucionales, conceptuales y metodológicos Cultura Estadística y Gestión Local Instituto Nacional de Estadística y Censos. Dirección Nacional de Planificación y Coordinación Estadística. (Julio 2007).
- Toranzos, Fausto I. “Teoría Estadística y Aplicaciones”. Ediciones Macchi. Reimpresión. (1997)
- Tukey, J. “Exploratory data analysis”. Ed. Addison Wesley. (1977)
- Valdes, Fernando. “Comprensión y Uso de la Estadística”. Universidad Rómulo Gallegos (Robert Ponterio – Project Direct and HTML). SUNY Cortland Dept. Glosario. (1998).

Libros de texto escolares

- De Simone, Turner., Funciones y estadística Ed. A-Z
- Estrada de 6º grado. Ed. Estrada. 2006.
- Latorre, L.; Spivak, L. ; Kaczor, P. y de Elizondo, M.C., Manual de 9º año. Ed. Santillana.
- Manual de 1º año. Ed. Kapelusz. 2006
- Manual de 7º grado. Ed. Puerto de Palos 2005.
- Manuales de 8º y 9º año. Ed. Santillana, Ed. Nuevas Propuestas. Ed. Tinta Fresca. Ed. Puerto de Palos.
- Matemática de 7º grado (1º E.S.B.). Ed. Puerto de Palos. 2005
- Textos 8º y 9º ((2º y 3º E.S.B.): Matemática I, Ed. Santillana, 1995. Matemática 8, Ed. Nuevas Propuestas. Matemática 8, Ed. Tinta Fresca, 2005. Matemática 9, Ed. Puerto de Palos, 2001
- Semino, S. ; Englebert, S. Y Pedemonti, S., Manual de 9º año. Ed A-Z.
- Seveso de Larotonda, J. ; Wykowski, Ana R. y Ferrarini, G., Manual de 9º año. Ed. Kapelusz.

Material digitalizado

<http://www.bioestadistica.uma.es/libro/node2.htm>

<http://www.hojamat.es/estadistica/diccio/diccestad.htm>.

<http://www.ucm.es/info/nomadas/>.

<http://maristascoruna.wikispaces.com/file/view/Apuntes+Estadistica+P4.pdf>

<http://web.cortland.edu/flteach/stats/>

<http://maristascoruna.wikispaces.com/file/view/Apuntes+Estadistica+P4.pdf>

<http://www.so.ucr.ac.cr/Enlaces/Estadistica/problemas/Diccionario.html>

http://www.vitutor.com/estadistica/descriptiva/a_2.html

http://es.wikipedia.org/wiki/Medidas_de_posici%C3%B3n_no_central

<http://www.ugr.es/local/batanero/sergroup.htm>

<http://web.udl.es/usuarios/z4084849/docs/uno45.pdf>

<http://dialnet.unirioja.es/servlet/articulo?codigo=2697036>

http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S1852-73102009000100001&lng=es&nrm=iso

Anexo

Instrumentos para recabar información de potenciales usuarios de herramientas estadísticas

Guía para la entrevista de alumnos de distintas carreras de la UNLaM

Edad:

Sexo:

Carrera que cursa:

Trabaja actualmente o trabajó? Si No

Indicar lugar/es

Actividad que desempeña o desempeñó:

Si en su trabajo, ha utilizado o utiliza herramientas estadísticas, completar el siguiente cuadro.

Herramientas	Indicar de qué tipo	Detallar resultados obtenidos
Cálculos		
Gráficos		
Tablas		
Otras (detallar)		

En caso de haber tenido dificultades en la utilización de las herramientas anteriores, indicar cuáles y de qué tipo.

Formulario guía para entrevista de usuarios comunes de herramientas estadísticas

Datos personales

Edad: Sexo: Nivel máximo de instrucción alcanzado:

En las siguientes opciones marcar con una X según corresponda:

Capacitación que posee:

Está estudiando alguna carrera de grado o de posgrado? Si No

En caso afirmativo la carrera que sigue es:

Indique el nivel de la misma: secundario terciario universitario

Está realizando en la actualidad algún curso de perfeccionamiento? Si No

En caso afirmativo indicar cuál o cuáles:

Realizó algún curso de estadística como estudiante o como especialización laboral?

No Sí ¿En qué condición?

Indicar cuál:.

Situación laboral:

Rubro de la fábrica, empresa o entidad donde desempeña su labor:

Antigüedad en el trabajo:

Puesto que ocupa o tareas que desempeña

Herramientas o cálculos estadísticos que necesita o maneja:

Emplea algún concepto o herramienta estadística para realizar sus tareas?

Estadística descriptiva:

No Sí ¿Señale cuál/es de los siguientes conceptos maneja?

Tablas ¿cómo?

Gráficos ¿cuáles?

Medidas ¿cuáles?

¿en qué contexto las emplea?

¿cómo las utiliza?

Inferencia Estadística:

Estimación de parámetros, ¿de qué parámetros?

Pruebas de Hipótesis, ¿de qué tipo?

¿cuáles.

Análisis de regresión – correlación, ¿en qué situación/es?

2- Producción científico – tecnológica

Trabajo sobre enseñanza de la estadística en la universidad.

Presentado y aprobado por el Congreso Internacional de Pedagogía Universitaria (CIPU).

Organizado por la UBA.

Fecha de presentación de la ponencia: 8 de setiembre de 2009.

Publicado en el Libro de Resúmenes. Primer Congreso Internacional de Pedagogía Universitaria. Editorial Eudeba. Septiembre 2009. Buenos Aires, Argentina. ISBN 978-950-23-1682-6.

Autores: Graciela Fernández, Laura Polola y María Eugenia Ángel.

3- Proyecto de Investigación

CODIGO...55B/142.....

1. PROYECTO DE INVESTIGACION

Título del proyecto: REGISTRO DE CONCEPTOS ESTADÍSTICOS COMO RECURSO PARA UNA ALFABETIZACIÓN ESTADÍSTICA

Unidad Ejecutoria: Univ. Nac. de La Matanza Departamento:Ciencias Económicas Carrera, Cátedra, etc.:Estadística Grupo de investigación: Dirección: Florencio Varela 1903 – San Justo. Cod. Postal: 1754. Tel.: 4480-8954

Investigadores Miembros del Equipo : Nombre y apellido : Graciela Fernández Nombre y apellido : Laura Polola Nombre y apellido : Enrique Borgna Nombre y apellido : Liliana Pagano Nombre y apellido : Silvia Brunetti Nombre y apellido : Miriam Ecalte

Director y Subdirector : Nombre y apellido : María Eugenia Ángel Titulo: Profesora de Matemática Legajo UNLM : 511 Categoría Docente: Profesor Titular Dedicación : Exclusiva Dirección Particular : Av. Rivadavia. 13230. Ramos. Mejía Tel. : 4658.- 2286.....
--

RESUMEN : Mediante un trabajo de buceo exploratorio se recopilará información estadística, procesada o no, de medios de difusión y comunicación –informes periodísticos, libros, publicaciones en general, Internet, su empleo a través de herramientas informáticas- para detectar usos y aplicaciones de términos y conceptos estadísticos según el contexto, como también se relevarán diferencias y similitudes de significado para mismas expresiones del lenguaje empleado a fin de identificar necesidades y potenciales destinatarios de los resultados de la presente investigación. A partir de éstos se elaborará un estudio intensivo de comparación y contrastación para reordenarlos, resignificarlos, clarificarlos y establecer relaciones entre ellos y con su raíz teórica formal tomando como referente textos y desarrollos universitarios institucionalizados ya que el lenguaje aplicado a trabajos y desarrollos formales o científicos requiere formas estandarizadas tanto a nivel académico como laboral. Desde esta tarea de compilación de datos y su análisis minucioso se realizará un primer registro de términos y conceptos universalmente utilizados en Estadística para su mejor y más acabado uso y comprensión y para la mejor interpretación de la información estadística tanto desde lo comunicacional como desde lo instrumental para su empleo en diversos ámbitos de estudio y también de trabajo, no necesariamente académico o educacional.

PLAN DE INVESTIGACION

Reflexiones que motivaron la investigación

A partir de la investigación y de la práctica educativa, desde siempre y naturalmente ha existido el interés por conocer cómo aprendemos los seres humanos para poder enriquecer la experiencia de enseñar y aprender. Dada la creciente influencia de las nuevas tecnologías en el tratamiento estadístico de la información es preciso contar con un referente conceptual que permita su buen uso y aplicación y que establezca un puente entre la teoría formal y la práctica responsable. En estos términos aparece la noción de alfabetización estadística asociada a determinar qué conocimientos, valores y habilidades necesita desarrollar el ciudadano de hoy para desempeñarse con solvencia en las sociedades actuales donde los medios de comunicación y las nuevas tecnologías actúan como herramientas en la construcción social de la cultura.

El uso de la tecnología expandió aún más la utilización de la estadística como herramienta cotidiana de trabajo. La divulgación de sistemas y programas informáticos en diversos ámbitos como los educacionales y laborales y la escasa o nula formación estadística de muchos usuarios plantean la necesidad de que el hombre común adquiriera suficiente formación estadística. Concretamente, de esta manera se estaría apuntando a una reestructuración, una renovación de los términos y conceptos estadísticos que tomara en cuenta a los interesados sin dejar de lado la función comunicativa del lenguaje.

Como cada rutina utilizada por los sujetos está llena de ecos y reflejos de otros usos y costumbres, la propuesta de este trabajo es la de remediar, en parte, el desconocimiento que se tiene de algunos términos y conceptos estadísticos de uso más frecuente. Se tratará de establecer cuáles son los más utilizados y cuáles son los vicios más comunes de encontrar en pos de enfatizar su manera más adecuada y llamar la atención sobre su mal empleo, si lo hubiere. Para tal fin se prevé confeccionar un “manual estadístico para el usuario” donde se asienten los términos y conceptos analizados con sus significados científicos -particularmente el académico- y sus aplicaciones, señalando las formas incorrectas más comunes. Esta sistematización y transmisión sería un elemento clave para organizar y modular la actividad formativa que se persigue del individuo.

Objetivo General

El objetivo es la elaboración de un registro de términos y conceptos estadísticos donde se unifiquen y jerarquicen las definiciones a partir de las similitudes y diferencias de interpretaciones de acuerdo a su contexto y aplicación.

Objetivos operacionales

- Identificar y clasificar conceptos a analizar.
- Compilar usos y aplicaciones del lenguaje estadístico.
- Compendiar posibles identificaciones entre conceptos que pueden parecer diferentes pero con idéntica raíz teórico-conceptual.
- Puntualizar diferencias entre conceptos que pueden confundirse o que se identifican sin ser equivalentes.
- Analizar y describir conceptualmente las herramientas estadísticas para un mejor aprovechamiento de sus potenciales resultados.
- Elaborar un registro de los resultados para su utilización en el estudio de la Estadística tanto para la formación básica, profesional, docente como así también para la formación del usuario de recursos estadísticos de diversa índole con fines laborales

Etapas del proceso de investigación

Esta investigación se desarrollará en varias etapas, que se encuentran relacionadas con el cumplimiento de los objetivos planteados.

- 1- Recopilación de terminología estadística utilizada en medios gráficos o audiovisuales de difusión, en paquetes informáticos de uso popular y textos de otras ramas del conocimiento.
- 2- Selección y clasificación de los términos y conceptos a analizar.
- 3- Reconocimiento de las distintas aplicaciones de los términos y conceptos seleccionados (su utilización)
- 4- Elección del material bibliográfico académico de referencia.
- 5 - Selección de definiciones formales de los términos y conceptos recopilados.
- 6- Comparación entre el uso común y la referencia formal institucionalizada (cruce de información)
- 7- Unificación y/o jerarquización del uso y significado de los términos y conceptos.
- 8- Construcción de un registro.

La metodología de trabajo es cualitativa basada en datos secundarios extraídos de diversas fuentes especializadas de las distintas áreas del conocimiento.

Cronograma de actividades:

El siguiente es un cronograma del tiempo estimado para las distintas actividades o etapas a realizar:

Actividad	Tiempo estimado (en meses)
1- Recopilación de terminología estadística utilizada en medios gráficos o audiovisuales de difusión, en paquetes informáticos de uso popular y textos de otras ramas del conocimiento.	8
2- Selección y clasificación de los términos y conceptos a analizar.	8
3- Reconocimiento de las distintas aplicaciones de los términos y conceptos seleccionados (su utilización)	8
4- Elección del material bibliográfico académico de referencia.	4
5 - Selección de definiciones formales de los términos y conceptos recopilados.	5
6- Comparación entre el uso común y la referencia formal institucionalizada (cruce de información)	6
7- Unificación y/o jerarquización del uso y significado de los términos y conceptos.	7
8- Construcción de un registro.	9

DESARROLLO DE LA INVESTIGACIÓN A TRAVÉS DEL TIEMPO (EN MESES)

ACTIVIDAD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
1	■	■	■	■	■	■	■	■																	
2		■	■	■	■	■	■	■	■																
3			■	■	■	■	■	■	■	■															
4									■	■	■														
5										■	■	■	■	■											
6												■	■	■	■	■	■								
7															■	■	■	■	■	■	■	■	■	■	■
8																	■	■	■	■	■	■	■	■	■

BIBLIOGRAFÍA BÁSICA

- A. Ardila, *El lenguaje Matemático y el Usual, como Mediador de la Comunicación*. Acta Latinoamericana de Matemática Educativa, Vol 15, Universidad de Panamá. (1991), 1161-1171.
- Batanero, Carmen. "Los retos de la cultura estadística". Congreso de Estadística. Caseros, Argentina, (2003).
- Batanero, Carmen. *Didáctica de la Estadística*. GEEUG (Grupo de Educación Estadística Universitaria de Granada). Publica: Grupo de Investigación Estadística, Departamento de Didáctica de la Matemática, Universidad de Granada. Granada. (2001).
- Bressan, Ana y Bressan, Oscar. *Probabilidad y estadística: como trabajar con niños y jóvenes*. Ed. Novedades Educativas. Argentina. (2008).
- Ferreiro, Osvaldo y Fernández de la Reguera, Pedro, *La estadística, una ciencia de la controversia*, Artículo publicado en la Revista Universitaria N° 25 del Instituto de Matemáticas y Física .Universidad de Talca. Chile. (1988)
- Guzmán, Miguel de, *Enfoque heurístico de la enseñanza de la matemática*, Aspectos didácticos de matemáticas 1 (1985), Publicaciones del Instituto de Ciencias de la Educación de la Univ. de Zaragoza.
- Guzmán, Miguel de, *Para pensar mejor* (Labor, Barcelona, 1991)
- Kreyszig, Erwin. *Introducción a la Estadística Matemática. Principios y métodos*. Ed. LIMUSINA, S.A. México, Quinta reimpresión, 1981.
- Moroney. M. S. *Hechos y estadística*. Ed. EUDEBA. (1979).
- Pozo, José I. (1993): *Comprender y transformar la enseñanza*. Ed. Morata. Madrid.
- Santaló Luis A.; Gálvez, Grecia; Charnay, Roland; Brousseau, Guy; Lerner, Delia y Sadosky, Patricia. Parra, Cecilia y Saiz, Irma (compiladoras). *Didáctica de matemáticas*. Aportes y reflexiones. Editorial Paidós Educador. Argentina. 5ta reimpresión, 1997.
- Santaló, L.A., *La educación matemática, hoy* (Teide, Barcelona, 1975)
- Toranzos, Fausto I. *Teoría Estadística y Aplicaciones*. Ediciones Macchi. 1997
- Tukey, J. *Exploratory data analysis*. Ed. Addison Wesley. (1977)

ESTADO ACTUAL DEL TEMA PROPUESTO

Actualmente existen varios diccionarios estadísticos, algunos de ellos disponibles en la web. Este material factible de utilizar presenta disparidades y contradicciones y en ninguno de los casos se intenta realizar una unificación y/o jerarquización del uso y significado de los términos y conceptos estadísticos como se pretende realizar en el presente trabajo.

TRANSFERENCIA

La transferencia del trabajo se realizará con la utilización del registro resultante de la investigación. Este registro podrá ser utilizado por estudiantes, docentes, usuarios de paquetes estadísticos y responsables de la producción de información basada en la estadística (investigadores, periodistas o comunicadores sociales)

VINCULACION DE LOS RESULTADOS OBTENIDOS CON LAS ACTIVIDADES DOCENTES

Como aplicación directa se utilizará el registro en la cátedra de estadística de la UNLaM como soporte terminológico. Extensivamente, el registro quedará a disposición para ser utilizado como complemento en cualquier otra materia que requiera de los conceptos o términos tratados.