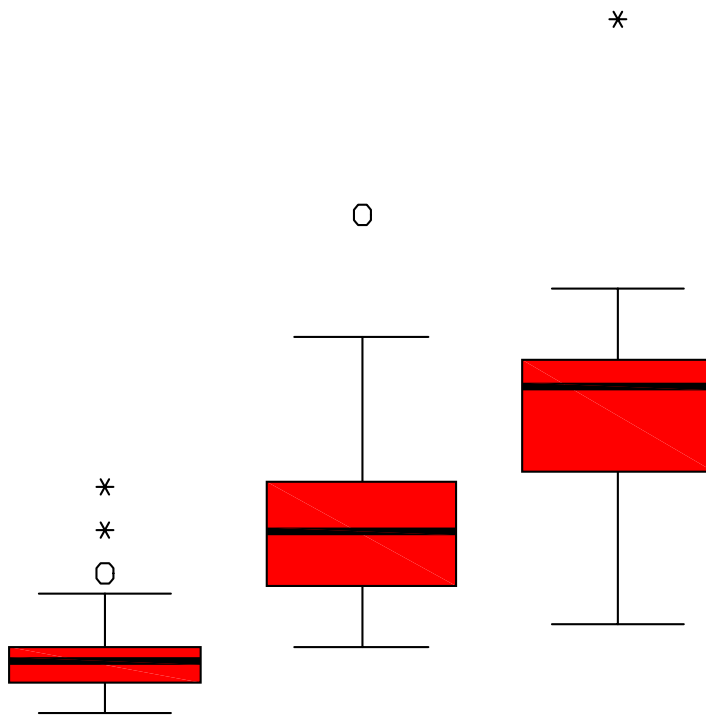


# ESTADÍSTICA APLICADA EN LAS CIENCIAS SOCIALES Y HUMANAS

## ESTADÍSTICA I



César N. AGUIRRE M. Fernanda NIÑO Eduardo F. SIMONETTI



EDITORIAL UNIVERSITARIA DE MISIONES

**San Luis 1870**

Posadas - Misiones – Tel-Fax: (03752) 428601

Correos electrónicos:

edunam-admini@arnet.com.ar

edunam-direccion@arnet.com.ar

edunam-produccion@arnet.com.ar

edunam-ventas@arnet.com.ar

**Colección:** Cuadernos de Cátedra

**Coordinación de la edición:** Nicolás Capaccio

**Tapa:** Francisco Sánchez

**Compaginación y armado de interiores:** Amelia E. Morgenstern

**Corrección:** Hedda Giraudó - Amelia E. Morgenstern

ISBN 987-9121-98-8

1ª reimpresión

Impreso en Argentina

©Editorial Universitaria

Universidad Nacional de Misiones

Posadas, 2005

Aguirre, César

Estadística aplicada en las ciencias sociales y humanas / César Aguirre; M. Fernanda Niño y Eduardo F. Simonetti; coordinado por Rodolfo Nicolás Capaccio - 1a ed. 1a reimp. - Posadas: Univ. Nacional de Misiones-Editorial Universitaria, 2005.

240 p.; 30x21 cm. (Cuadernos de Cátedra)

ISBN 987-9121-98-8

1. Sociología I. Niño, M. Fernanda, II. Simonetti, Eduardo F., III. Rodolfo Nicolás Capaccio, coord. IV. Título

CDD 301.

## PRESENTACIÓN

### ¿POR QUÉ SABER DE ESTADÍSTICA?

La Estadística ha ganado reconocimiento como disciplina importante en la formación profesional universitaria de diferentes campos del conocimiento. Sus aportes a esta formación ocurren en dos niveles: el *primero* como disciplina contributiva a la preparación científica de los estudiantes, proporcionando los conocimientos indispensables e irremplazables en la producción, tratamiento y análisis de datos cuantitativos. El *segundo* nivel se manifiesta en el plano de la formación general de los individuos que deben desenvolverse en el mundo actual, intensamente conectado por las modernas telecomunicaciones, y cada vez más interdependiente en lo económico y social. "En este nuevo mundo es importante ser capaz de orientarse en la red de información disponible, mucha de la cual es cuantitativa. El ciudadano debe moverse entre masas de datos cuantitativos que muchas veces son contradictorios y requieren de un mínimo de conciencia sobre la forma en que tales datos se recogen, organizan, analizan e interpretan. Como resultado se precisan nuevas habilidades<sup>1</sup>.

El ciudadano medio de hoy necesita reunir conocimientos que le sirvan para decodificar la cuantiosa información que recibe por diferentes medios, que le permitan juzgar la calidad de esa información, que le sean útiles para comprender ideas expresadas por otros y formar las propias, que le sirvan para construir y comunicar su propia información.

La Estadística es la disciplina que aporta los recursos culturales y prácticos que el ciudadano necesita para desenvolverse en la nueva sociedad de la información, y la enseñanza de la disciplina tiene el fin de generar y difundir una "cultura estadística" dirigida a dotar al ciudadano común de tales recursos y habilidades.

### OBJETIVOS GENERALES

El curso de estadística aplicada tiene el propósito de *promover la formación de usuarios competentes* de esta disciplina y sus herramientas. Usuarios con habilidades que le permitan *vincular los conocimientos estadísticos con la solución de problemas* de su campo disciplinar o profesional específico.

Es decir, se trata de desarrollar en los alumnos la capacidad de **abordar y resolver, desde la Estadística, problemas de producción de información** con fines científicos o de tomar decisiones.

Los **objetivos generales** derivados de este propósito de formación, son:

1. Promover en los alumnos el desarrollo de habilidades intelectuales del *pensamiento-razonamiento estadístico*. Ello supone abordar la solución de un problema de trabajo desde los siguientes elementos básicos:
  - la *necesidad de datos* para alcanzar una correcta comprensión del fenómeno o problema en estudio (*impulso estadístico*);
  - la idea de la "*transnumeración*", entendida como la habilidad de construir y producir *datos como representaciones de aspectos de un sistema real* para lograr una *mejor comprensión* de dicho sistema (datos que capturan información significativa sobre elementos del sistema);
  - la necesidad de la *observación masiva* (cantidad numerosa de datos: *datos estadísticos*) como fundamento del *análisis estadístico* de los fenómenos;
  - la presencia de la *variación* en los datos (y en los fenómenos que ellos representan) y el *error e incertidumbre* como consecuencia de ella;

---

<sup>1</sup> OTTAVIANI, M. G. (1999): *Notas sobre los Desarrollos y Perspectivas en Educación Estadística*. International Association for Statistical Education -IASE-.

- la necesidad del *resumen, descripción y modelización* de la variación.
- 2. Capacidad para *formalizar un problema de trabajo* o investigación en términos (preguntas) estadísticos.
- 3. Capacitar para *la construcción y obtención* de los datos que requiere la solución de un problema.
- 4. *Conocer herramientas* estadísticas para el tratamiento y análisis de datos y *comprender sus fundamentos* lógicos, limitaciones, propiedades, etc.
- 5. Manejar la *tecnología auxiliar disponible* para la aplicación de las *herramientas estadísticas*.
- 6. *Integrar* apropiadamente diferentes *herramientas estadísticas* en el análisis de un problema (estrategia de abordaje estadístico del problema).
- 7. *Interpretar los resultados* estadísticos en el contexto del problema de estudio y *comunicar los hallazgos* o respuestas a las preguntas iniciales (información-comunicación).

## **ENFOQUE DEL CURSO**

Tratándose de un Curso de Estadística para “*no estadísticos*” y considerando los objetivos generales citados anteriormente, la estrategia pedagógica a seguir **enfatizará la comprensión conceptual de los contenidos** a desarrollar. Es decir, se privilegiará la *conceptualización de la Estadística*, por sobre la demostración y el tratamiento matemático de sus conocimientos.

También se pondrá mucho énfasis en **vincular los conceptos y herramientas con la solución de problemas reales de investigación o de decisiones**. De ahí que **el cálculo estadístico no constituye una actividad central** del curso y se realizará –únicamente- con propósitos pedagógicos.

En este primer nivel del Curso de Estadística se presentarán con la mayor profundización posible, temas relativos a un análisis descriptivo de los datos (Estadística Descriptiva). En esta pretensión, resulta básico e insoslayable el abordaje de los siguientes tópicos:

- ✓ La Investigación Estadística (Unidad 1)
- ✓ Organización y Descripción Inicial de los Datos (Unidad 2)
- ✓ Los valores que Caracterizan al Conjunto de Datos (Unidad 3)
- ✓ Análisis de la Variación y Asimetría (Unidad 4)
- ✓ Estudio de la Relación entre Variables (Unidad 5)
- ✓ Los Números Índices (Unidad 6)

Este Curso de Estadística ha sido pensado como una **propuesta no presencial** de formación. Por ello, los contenidos y las actividades han sido organizados y producidos de tal manera que el trabajo pueda auto-administrarse sin grandes dificultades.

El presente material incluye, para cada unidad, lo que se dio en llamar “**Notas de Cátedra**”, en las que se **desarrollan los conceptos teóricos centrales**, orientando el aprendizaje<sup>2</sup>. Además, las “Notas de Cátedra” remiten periódicamente a las “**Guías de actividades**”, en las que se proponen consignas de trabajo (teóricas y/o prácticas) que facilitan la comprensión de los conceptos tratados y le plantean situaciones concretas de análisis de datos que favorecen el desarrollo de las habilidades necesarias para el tratamiento de datos y la producción de información.

Al final de las “**Guías de Actividades**” se presenta un propuesta de trabajo denominada “**Evaluación Parcial de la Unidad....**”. Se trata de una actividad de síntesis de los conocimientos teóricos y prácticos desarrollados en la unidad, mediante la cual se podrán evaluar los avances en sus conocimientos.

---

<sup>2</sup> En cada unidad, se recomienda bibliografía complementaria.

## **ORIENTACIONES PARA EL USO DEL MATERIAL**

El material incluye señales que intentan orientar la lectura y subrayar aquellas cuestiones en las que se debe poner especial atención. También ofrece esquemas y gráficos que sintetizan un conjunto de conceptos y las relaciones que se pueden establecer entre ellos.

Además de los recursos utilizados tradicionalmente para destacar algún aspecto parcial de la presentación (uso de negritas y/o cursivas), en las Notas de Cátedra se han utilizado un conjunto de íconos que señalan partes diferentes -en cuanto a su naturaleza- del desarrollo teórico. Así, aparecerán:



En general, los temas se presentan a partir de situaciones que ponen en evidencia la necesidad de nuevas herramientas de análisis. Estas situaciones se traducen en preguntas de investigación y estadísticas, las que requieren el uso de herramientas específicas para encontrar una respuesta. El icono señala el carácter de **planteo general** del texto.



Destaca en el texto los **conceptos y definiciones**.



Indica el desarrollo de un **ejemplo** donde se utilizan los conceptos presentados.



Enfatiza lo **importante** de algunas cuestiones consideradas en el desarrollo del tema.



Señala la parte del texto donde se realiza la **interpretación** de los resultados estadísticos obtenidos.



Advierte sobre la necesidad de **hacer un alto en la lectura** y realizar la actividad que se indica.



Señala las **consignas de trabajo** a realizar, en las Guías de Actividades de cada unidad.

En todas las unidades se han incluido **esquemas** de los contenidos tratados, una síntesis de "**lo visto**", una enumeración de los **conceptos centrales** y las **habilidades** que se pretendieron transmitir. El propósito de incluir estos recursos es **brindarle una mirada global** sobre la unidad y simultáneamente **destacar las ideas centrales que la estructuran**, así como las habilidades que se asocian a esas ideas en la práctica. Consecuentemente, le recomendamos especial atención a estas diferentes formas de síntesis de cada unidad, ya que constituyen otra forma de aproximación a los conceptos desarrollados y le permitirán reorientar una segunda lectura del material.

**Los Autores:**

**AGUIRRE, César Norberto**

Estadístico (Universidad Nacional de Rosario -Argentina-), postgrado en Estadística y Cuentas Nacionales (Instituto de Estudios Sociales de La Haya -Holanda-), Especialización en Administración Estratégica Universitaria (Universidad de Quebec -Canadá-), Especialización en Administración Estratégica de Negocios (Universidad Nacional de Misiones).

Profesor Regular Titular de Estadística (Facultad de Humanidades y Ciencias Sociales - UNaM-), Profesor Titular de Estadística (Maestría de Gestión Pública -UNaM-), ex Docente de Cursos de Postgrado en Análisis Exploratorio de Datos (Programa PRESTA, Universidad Libre de Bruselas-Unión Europea).

**NIÑO, María Fernanda**

Profesora de Matemática, Física y Cosmografía, Inst. Sup. del Profesorado "Pbro. Dr. Antonio Saenz". Maestría en Docencia Universitaria, Fac. de Ingeniería-Univ. Nac. Misiones (etapa elaboración de tesis).

Ayudante de Primera (Regular) de Estadística (Fac. de Humanidades y Ciencias Sociales de la UNaM). Ex docente de Cursos de Posgrado de Métodos Estadísticos Multivariados Aplicados a las Ciencias Humanas Y Sociales, del Programa PRESTA (Univ. Libre de Bruselas), auspiciado por la Unión Europea. Docente en el Curso de Estadística Aplicada, (Maestría en Gestión Pública de la Facultad de Ciencias Económicas-UNaM). Ex docente tutor del Seminario "Metodología y Técnicas de la Investigación Social" (Maestría en Gerencia y Administración de Proyectos Sociales -UNaM-)

**SIMONETTI, Eduardo Francisco**

Estadístico (Universidad Nacional de Rosario -Argentina-), Master en Desarrollo Económico para América Latina, Universidad Internacional de Andalucía – Sede Iberoamericana de La Rábida (España).

Profesor Titular Regular de "Indicadores Socioeconómicos", (Facultad de Humanidades y Ciencias Sociales de la UNaM). Docente en el Curso de Estadística Aplicada, (Maestría en Gestión Pública de la Facultad de Ciencias Económicas-UNaM). Docente del Seminario "Sistemas de Información y Herramientas Informáticas para la Gestión de Programas Sociales" (Maestría en Gerencia y Administración de Proyectos Sociales -UNaM-).

---

 ÍNDICE
 

---

**Unidad 1: La Investigación Estadística**

	<i>Página</i>
<b>1. Introducción</b> .....	11
<b>2. Intentando Definir la Estadística</b> .....	11
<b>3. Problema de Trabajo e Investigación Estadística</b> .....	12
3.1. Las Preguntas de Investigación .....	13
3.2. Las Preguntas estadísticas .....	13
<b>4. Los Datos</b> .....	13
<b>5. Las Variables</b> .....	15
<b>6. Conjunto de Datos: Datos Estadísticos</b> .....	17
<b>7. Fuentes de Datos</b> .....	18
<b>8. Investigación por Censo y por Muestra</b> .....	19
<b>9. ¿Qué Hemos Visto?</b> .....	20
<b>Esquema-La Estadística en el Proceso de Investigación</b> .....	21
<b>Esquema – Estructura del Curso – Estadística Descriptiva</b> .....	22
<b>Bibliografía</b> .....	23

---

**Unidad 2: Organización y Descripción Inicial de los Datos**

<b>1. Los Datos y la Información</b> .....	25
<b>2. La Primera Organización de los Datos: la Matriz de Datos</b> .....	25
<b>3. El Análisis de la Matriz de Datos</b> .....	28
<b>4. Las Distribuciones de Frecuencias en el Análisis Univariado</b> .....	29
4.1. Variables categóricas .....	30
- el recurso numérico .....	30
- el recurso gráfico .....	31
4.2. Variables numéricas.....	32
4.2.1. Variables numéricas con pocos valores diferentes.....	32
- el recurso numérico .....	32
- el recurso gráfico .....	33
4.2.2. Variables numéricas con muchos valores diferentes.....	34
- el recurso numérico .....	34
- el recurso gráfico .....	39
4.3. Transformaciones de las frecuencias absolutas.....	42
4.3.1. Las frecuencias relativas .....	42
4.3.2. Las frecuencias acumuladas .....	43
4.3.3. La curva de Lorenz y el índice de Gini .....	45
4.4. Otras consideraciones sobre los recursos gráficos .....	52
4.5. Esquema – Tipos de gráficos univariados.....	55
<b>5. ¿Qué Hemos Visto?</b> .....	56
<b>Esquema – El Análisis de Datos: Distribuciones de Frecuencias</b> .....	57
<b>Bibliografía</b> .....	58

---

**Unidad 3: Los Valores que Caracterizan al Conjunto de Datos**

	<i>Página</i>
<b>1. ¿Por qué son Necesarios?</b> .....	59
<b>2. ¿Cuáles Son?</b> .....	60
<b>3. Media Aritmética</b> .....	60
3.1. Principales Propiedades de $\bar{x}$ .....	61
3.2. Cálculo de la Media .....	64
3.2.1. Datos sin resumir.....	64
3.2.2. Datos agrupados en arreglo de frecuencias.....	64
3.2.3. Datos agrupados en una distribución con intervalos.....	65
<b>4. La Mediana</b> .....	66
4.1. Principales propiedades de Ma.....	67
4.2. Determinación de la Ma .....	68
4.2.1. Datos numéricos sin resumir.....	68
4.2.2. Datos numéricos en arreglo de frecuencias .....	69
4.2.3. Datos numéricos en una distribución con intervalos .....	70
4.2.4. Datos categóricos ordinales .....	71
<b>5. El Modo</b> .....	72
5.1. Principales Propiedades del Mo.....	73
5.2. Determinación del Mo.....	73
5.2.1. Para arreglos de frecuencias y datos categóricos .....	73
5.2.2. Para una distribución con intervalos.....	74
<b>6. Cuartiles, Deciles, Centiles</b> .....	75
6.1. Los Cuartiles .....	76
6.2. Los Deciles .....	77
6.3. Los Centiles .....	78
6.4. Curva de Lorenz asociada a las medidas de posición.....	78
<b>7. ¿Cómo Integrar estas Medidas de Resumen?</b> .....	80
7.1. El resumen de los cinco números.....	80
7.2. El diagrama de Caja ( <i>Box-plot</i> ).....	81
<b>8. ¿Qué Hemos Visto?</b> .....	82
<b>Esquema – Valores que Caracterizan un Conjunto de Datos</b> .....	83
<b>Bibliografía</b> .....	84

**Unidad 4: Análisis de la Variación y Asimetría**

<b>1. ¿Por qué Evaluar la Variabilidad y la Asimetría?</b> .....	85
<b>2. ¿Cómo Medir la Variabilidad?</b> .....	86
2.1. Para variables numéricas .....	86
2.1.1. Las medidas absolutas .....	87
A) El Rango, Amplitud o Recorrido .....	87
B) El Rango Intercuartil.....	88
C) Desviación Media.....	88
D) Desviación Mediana.....	90
E) Variancia y Desviación estándar .....	91
2.1.2. Las medidas relativas .....	92
F) Coeficiente de variación .....	93
G) Coeficiente de Desviación Media .....	94



	<i>Página</i>
H) Coeficiente de Desviación Mediana .....	94
2.2. Dispersión para variables categóricas .....	94
<b>3. ¿Cómo Medir la Asimetría?</b> .....	<b>97</b>
3.1. Coeficiente de Asimetría de Pearson .....	98
3.2. Coeficiente intercuartílico de Bowley .....	99
<b>4. ¿Qué Hemos Visto?</b> .....	<b>101</b>
<b>Esquema – Medidas de Dispersión y Asimetría</b> .....	<b>102</b>
<b>Bibliografía</b> .....	<b>103</b>

## Unidad 5: Estudio de la Relación entre Variables

<b>1. ¿Por qué Estudiar la Relación entre Variables?</b> .....	<b>105</b>
<b>2. La Relación entre Variables Categóricas</b> .....	<b>108</b>
2.1. El recurso numérico .....	108
2.2. El recurso gráfico .....	114
2.2.1. Gráficos compuestos.....	114
2.2.2. Gráficos de partes componentes.....	115
<b>3. La Relación entre Variables Categóricas y Numéricas</b> .....	<b>116</b>
3.1. El recurso numérico .....	116
3.2. El recurso gráfico.....	119
<b>4. La Relación entre Variables Numéricas</b> .....	<b>120</b>
4.1. El recurso gráfico.....	120
4.2. El recurso numérico .....	124
4.2.1. El análisis de regresión lineal simple.....	124
4.2.2. El coeficiente de correlación lineal de Pearson.....	127
<b>5. ¿Qué Hemos Visto?</b> .....	<b>129</b>
<b>Esquema – Estudio de la Relación entre Variables</b> .....	<b>130</b>
<b>Bibliografía</b> .....	<b>131</b>

## Unidad 6: Los Números Índices

<b>1. ¿Qué son y cuál es su utilidad?</b> .....	<b>133</b>
<b>2. Los Números Índices Simples</b> .....	<b>134</b>
2.1. El Relativo Simple de Base Fija.....	134
2.2. El Relativo Simples de Eslabón.....	135
2.3. El Relativo Simple en Cadena.....	136
<b>3. Los Números Índices Compuestos</b> .....	<b>137</b>
3.1. El Índice de Agregados no Ponderados .....	138
3.2. El Índice de Promedio de Relativos no Ponderados .....	139
3.3. Los Índices de Agregados Ponderados .....	141
3.3.1. El índice de Laspeyres.....	141
3.3.2. El índice de Paasche .....	143
3.4. Los Índices de Promedios Ponderados de Relativos.....	144
3.4.1. El índice promedio ponderado de relativos de Laspeyres.....	144
3.4.2. El índice promedio ponderado de relativos de Paasche .....	145
<b>4. Algunas Consideraciones Especiales Temas Especiales</b> .....	<b>146</b>
4.1. El Índice de Valor .....	146

	<i>Página</i>
4.2. El Cambio de Base de un Número Índice .....	147
4.3. El Empalme de Dos Números Índices Solapados.....	148
4.4. Procedimiento de Números Índices en Cadena.....	149
4.5. La Deflación de una Serie.....	149
<b>5. Problemas en la Construcción de los Números Índices .....</b>	<b>150</b>
5.1. La Selección de la Muestra .....	150
5.2. La Elección del Período Base .....	151
5.3. La Ponderación Adecuada .....	151
5.4. La Selección del Promedio.....	151
5.5. Los Cambios de Producto.....	151
<b>6. ¿Qué Hemos Visto? .....</b>	<b>152</b>
<b>Bibliografía .....</b>	<b>153</b>

---

**Anexo: GUÍA DE ACTIVIDADES**

Unidad 1: La Investigación Estadística.....	157
Unidad 2: Organización y Descripción Inicial de los Datos .....	167
Unidad 3: Los Valores que Caracterizan al Conjunto de Datos.....	177
Unidad 4: Análisis de la Variación y Asimetría.....	185
Unidad 5: El Estudio de la Relación entre Variables.....	189
Unidad 6: Los Números Índices .....	199

---

Bibliografía General.....	157
---------------------------	-----

---