



Universidad de Buenos Aires
Facultad de Ciencias Sociales
Carrera de Sociología

Documento de Cátedra Nº 91

EJERCICIOS DE ESTADÍSTICA PARA METODOLOGÍA II.

Ma. Alejandra Otamendi
(2014)

Cátedra de Metodología y Técnicas de la Investigación Social
Profesora Titular: Ruth Sautu

Este Documento de Cátedra forma parte de una serie que tiene como propósito contribuir a la formación de alumnos de la Carrera de Sociología de la Universidad de Buenos Aires en cuestiones vinculadas al diseño y realización de investigaciones científicas en el campo de las ciencias sociales. Su contenido complementa los textos de metodología de lectura obligatoria y optativa incluidos en cada uno de los tres cursos que se dictan en la Carrera.

Estos documentos son material de uso interno y no pueden ser incorporados a ediciones impresas ni reproducidos comercialmente.

La Cátedra solicita a los usuarios de estos Documentos que citen a sus autores indicando las referencias completas, es decir: autores, fecha, título, número y tipo de documento (traducción, resumen, elaboración propia, etc.). En los casos en que el usuario utilice sólo parte del documento, haciendo referencia a algunos de los autores/obras originales allí incluidos, consignar que fue tomado de nuestro Documento de Cátedra. Por ejemplo:

E. O. Wrigth (1985), *Classes*, London: Verso, citado en Documento de Cátedra II.1., Plotno, G., Lederman, F. & Krause, M. (2007) "Escalas Ocupacionales".

EJERCICIOS DE ESTADISTICA PARA METODOLOGIA II.

EJERCICIO 1¹. Se midió la altura (en m) y el peso (en kg) de una muestra de 420 niños en un país pobre del sur de Asia. Se encontró que el promedio de peso / altura es de 38,1, con una desviación estándar de 8,2.

Un médico le informa que una relación peso / altura o por debajo de 27 podría ser un indicador de desnutrición severa. Suponiendo que la distribución de la muestra de relaciones de peso / talla es normal, ¿qué porcentaje de los niños es probable que sea severamente desnutridos en la muestra? Explicar el razonamiento que subyace a la respuesta.

- Construya un intervalo de confianza del 99% (IC) para la media de la población y explique su procedimiento.
- ¿Cuál es el propósito de formar un IC para una media poblacional? ¿Qué se aprende mediante un IC que no se puede aprender mediante la realización de una prueba de hipótesis?
- Explique por qué los IC son más amplios cuando usamos (a) mayor nivel de confianza, como en este caso, (b) un tamaño pequeño de muestra.
- Discuta la teoría estadística para las inferencias realizadas en el punto a) para la media poblacional. ¿Se asume normalidad de la distribución de la población? ¿Y de la distribución de la muestra? ¿Por qué sí o por qué no?

EJERCICIO 2². Por ley, una planta industrial puede descargar no más de 500 litros de aguas residuales por hora, en promedio, en un lago vecino. Sobre la base de otras infracciones que han notado, un grupo de acción ambiental considera que este límite se está excediendo. El seguimiento de la planta es caro, y sólo una pequeña muestra es posible. Una muestra aleatoria de cuatro horas se selecciona durante un período de una semana.

Una impresión de computadora muestra los resultados:

Variable	N	Media	DE	EE de la Media
Residuos	4	1000	400	200

- Si no conoce la forma de la distribución de la variable “residuos” en la población, ¿cómo debe proceder y en base a qué supuesto teórico?
- Para estimar la media de la población, ¿con qué distribución muestral debe trabajar y por qué? Descríbala.
- Calcule la estimación de la media poblacional con un intervalo con un nivel de confianza del 95%.
- Calcule otro IC con un nivel de significancia de 0,001. Compare con los resultados del punto a). ¿Qué gana y qué pierde en cada caso?

¹ Adaptado de Agresti, A. y Finlay, B. (2009). *Statistical Methods for the Social Sciences*, 4th ed. Mexico: Pearson Prentice Hall.

² Ídem.

EJERCICIO 3. Si quisiera construir una muestra al azar simple de estudiantes de Sociología de la UBA que fuera representativa por sexo con las siguientes características, ¿qué tamaño tendría?

- con un nivel de confianza del 95% y un término de error de $\pm 4\%$. Además, suponga que desconoce la variabilidad de sexo de la población en estudio.
- lo mismo que en a), pero con un nivel de confianza del 99%. Compare los resultados con a) e interprete.
- con un nivel de confianza del 95% y un término de error de $\pm 3\%$, con variabilidad desconocida en la población. Compare los resultados con a) e interprete.
- con un nivel de confianza del 95% y un término de error de $\pm 4\%$, pero sabe a partir de un Censo estudiantil anterior que la proporción de mujeres es de 0,6. Compare los resultados con a) e interprete.

Utilice la Tabla 3.1. de Baranger (1992:78) y recuerde la fórmula para calcular el tamaño de muestra para una variable categórica:

$$n = \frac{(pq)(t_{\alpha})^2}{\text{ter.error}^2}$$

Donde:

n = tamaño de la muestra requerido

t_{α} = puntuación t que corresponde al nivel de confianza y significancia establecidos

p = p de la categoría de éxito en la muestra

q = p de la categoría de fracaso en la muestra

Término del error = precisión deseada en los resultados que se van a reportar

EJERCICIO 4. Se realizó una encuesta a una muestra representativa de estudiantes de Trabajo Social de UBA en el primer cuatrimestre del 2012. A partir de la información obtenida, se construyó el siguiente cuadro:

Condición de los estudiantes según sexo de los alumnos de Trabajo Social de la UBA, 2012 (en absolutos)

Condición de estudiante	Sexo		Total
	Masculino	Femenino	
Sólo estudian	65	255	320
Estudian y trabajan part-time	35	65	100
Estudian y trabajan tiempo completo	50	30	80
Total	150	350	500

Fuente: datos ficticios.

- Postule una hipótesis entre las variables y realice la lectura del cuadro. Saque conclusiones.
- Usted quiere saber si la relación que encontró en la muestra, existe en la población de estudio con cierto nivel de confianza y margen de error. ¿Qué procedimiento debe realizar? ¿Por qué?
- Realice la prueba correspondiente para un nivel de confianza de 95%. Interprete los resultados.

- d) Realice la misma prueba con un nivel de significancia de 0,01. Interprete los resultados y compárelos con el punto anterior.
- e) Si quisiera conocer la fuerza de la relación, ¿qué coeficiente de asociación utilizaría y por qué?
- f) ¿Cuáles son los criterios que debe tener en cuenta para seleccionar un coeficiente de asociación?
-

Fórmulas

$$\bar{z} = \frac{Y - \bar{Y}}{s}$$

$$\mu = \bar{Y} \pm z_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

$$\mu = \bar{Y} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n-1}}$$

$$S.E._{\bar{Y}} = \frac{s_Y}{\sqrt{n}}$$

$$n = \frac{(pq)(t_\alpha)^2}{ter.error^2}$$

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

$$f_e = \frac{\text{Total fila} * \text{total columna}}{\text{Total (n)}}$$

Tabla con medidas de asociación

Medida de asociación	Tabla	Nivel de medición	Observaciones
Phi	2 x 2	Nominales	Son medidas basadas en el estadístico chi cuadrado. Toman valores entre 0 y 1 que indican mínimo y máximo grado de asociación.
C coeficiente de contingencia	r x c	Nominales	No alcanza la unidad, aunque hubiera asociación perfecta. Depende del máximo del tamaño de la tabla.
V de Cramer	r x c	Nominales	Tiende a subestimar la asociación. Son medidas simétricas.
Lambda	r x c	Nominales	Medida asimétrica. Puede tomar El mínimo valor en tablas con asociación.
Gamma	r x c	Ordinales	Toma valores entre -1 y 1 que indican máximo grado de asociación negativa y positiva respectivamente. Puede alcanzar valores extremos en tablas en las que la asociación no es total. Medida simétrica. En un cuadro de 2 x 2 es el mismo resultado que se obtendría si se hubiese calculado Q.
Tau b de Kendall	r x c	Ordinales	Medida simétrica. Únicamente alcanza valores extremos en tablas con asociación total. Pero si r es distinto de c no puede alcanzar 1.
Tau c de Kendall	r x c	Ordinales	Medida simétrica. Puede alcanzar valores extremos en caso de que r sea distinto de c, pero tiende a subestimar la asociación.
D de Somers	r x c	Ordinales	Medida asimétrica. Puede alcanzar valores extremos en tablas donde la asociación no es total.