

Análisis Cuantitativo I

Introducción y Conceptos Básicos

Carlos Cardona

Universidad del Rosario

27 de enero de 2017

Introducción

- ¿Sobre qué trata este curso?
 - Investigación Social.
 - Estadística.
 - Su intersección e interacción.

¿Por qué este curso?

- 1 Para aquéllos no involucrados en la investigación:
 - Todos los días estamos expuestos a información que requiere un conocimiento básico de estadística. (Publicidad, noticias, campañas políticas, encuestas de opinión.)
- 2 Para aquéllos involucrados en la investigación:
 - La estadística fortalece el proceso investigativo. En muchos escenarios provee validez a las conclusiones encontradas.
 - Permite encontrar evidencia empírica a teorías formuladas con base en información cualitativa.

Algunas preguntas de investigación

- Entre algunas de las preguntas que la estadística puede ayudar a responder se encuentran:
 - ¿Qué efecto tiene la radio sobre la participación electoral de los ciudadanos?
 - ¿La forma de agricultura tradicional practicada en la era pre-industrial tiene relación con los roles de género en la actualidad?
 - ¿Existe discriminación racial en el mercado laboral?

¿Qué es la Estadística?

- La investigación en muchos campos requiere la recolección de información con el propósito de responder, por ejemplo, si la radio tiene algún efecto sobre el comportamiento político.
- Al final de este proceso, el investigador usualmente termina con una gran cantidad de medidas tales como la tasa participación política, la tasa de abstención, el número de estaciones de radio, etc.
- *La estadística se refiere al conjunto de procedimientos que permiten organizar, resumir, interpretar y analizar estas características.*

Población (1)

- La investigación empieza con una pregunta específica a un grupo de *individuos*.
- Por ejemplo, el gobierno de un país está interesado en evaluar si la introducción de la radio en sus ciudades tiene algún efecto sobre el porcentaje de abstención.
- En este caso, el investigador estaría interesado en todas las ciudades del país.
- Otros ejemplos serian...

Población (2)

Definición

Una población es el conjunto completo de *sujetos* de interés para una pregunta de investigación en particular. La letra N define el tamaño de la población.

Muestra (1)

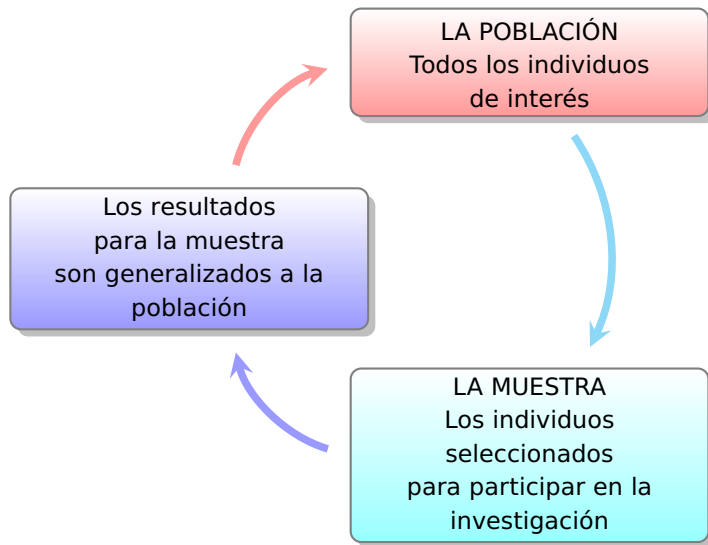
- Una población puede ser extremadamente grande (e.g., la población de todo el mundo.). La investigación podría ser más específica y estar interesada en la población de este salón.
- Claramente, el tamaño de la población puede variar según como el investigador la defina.
- Usualmente, las poblaciones tienden a ser grandes por lo que el investigador no puede examinar cada individuo. Por tanto, los investigadores seleccionan un subconjunto de la población, más pequeño y manejable.

Muestra (2)

Definición

Una muestra es un conjunto de individuos seleccionados de la población, con el fin de representar la población en una investigación. Usualmente se utiliza la letra n para definir el tamaño de la muestra.

La relación entre Población y Muestra



Variables

- Típicamente, el investigador está interesado en características específicas de los individuos de la población o en factores externos que los influyen.
- Por ejemplo, podría estar interesado en la influencia del conflicto sobre la participación política.
- Al cambiar los niveles de conflicto, ¿cambia también la participación política?
- La estadística provee una manera de lidiar con la *variabilidad*, ya que permite explicar la variación de una característica a partir de la variación de otras.

Variables

Definición

Una variable es una característica o condición que cambia o tiene diferentes valores para *individuos* en una muestra o población.

- Para demostrar cambios en las variables, es necesario realizar mediciones de las variables que se están examinando.
- El conjunto completo de valores se denomina base de datos.

¿Cómo luce una base de datos?

Individuo	Sexo	Estado Marital	Edad	Ingreso Anual
1	M	1	23	5150000
2	M	0	37	2500000
3	H	1	47	685000
4	M	1	61	455000
5	H	1	30	2350000
6	H	0	21	0
7	H	1	55	3600000
8	M	0	27	5250000

Parámetros y Estadísticos

- Antes, definimos población y muestra en términos de los *individuos*. Sin embargo, también nos podemos referir a población o muestra de los valores de las variables.
- Cada muestra (o población) de individuos produce una muestra (o población) de valores!

Definición

Un *parámetro* es un valor numérico que describe una población. Un *estadístico* es un valor numérico que describe una muestra.

Tipos de Métodos Estadísticos

Estadística Descriptiva

Es el conjunto de métodos usados para resumir, organizar y simplificar datos.

Estadística Inferencial

Consiste en técnicas que permiten realizar predicciones acerca de características de una población, basadas en información en una muestra de esa población.



Para el 2015, el 75% de los jóvenes empleados eran trabajadores asalariados y el 22,1% eran trabajadores independientes.

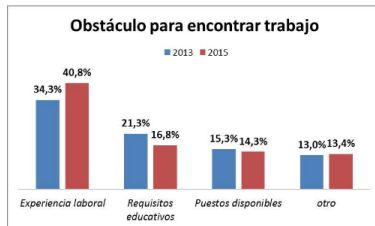
Independencia laboral incrementó levemente en los jóvenes, sin embargo, la falta de experiencia sigue siendo un obstáculo para conseguir empleo, según el Dane.

ECONOMÍA | 4/10/2016 12:15:00 AM

El 40,8% de los jóvenes colombianos dice no conseguir trabajo por falta de experiencia

Para el año 2015, la expectativa por los cargos directivos disminuyó y el interés por la

El 40,8% de los jóvenes desempleados afirmó que el principal obstáculo para encontrar trabajo fue no tener suficiente experiencia laboral, seguido del 16,8% que consideró que los requisitos del trabajo son mayores que la educación o formación recibida y el 14,3% reportó que no hay suficientes puestos de trabajo disponibles.



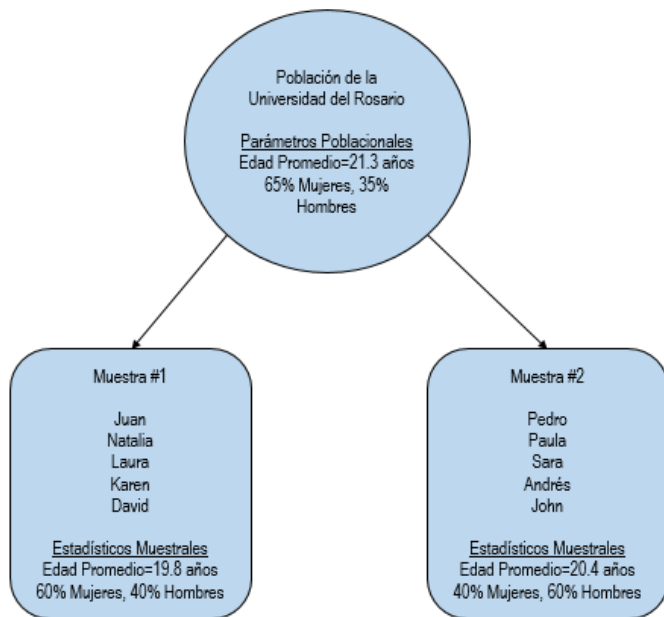
Error de Muestreo (1)

- Un problema de usar muestras, es que una muestra provee sólo información limitada de la población.
- A pesar de que las muestras generalmente son representativas de sus poblaciones, no se espera que una muestra de una perspectiva perfectamente precisa de la población entera.
- Por tanto, existe cierta discrepancia entre el estadístico muestral y el parámetro poblacional.

Error de Muestreo (2)

Definición

El error de muestreo es la discrepancia, o error, que existe entre un estadístico muestral y su correspondiente parámetro poblacional.

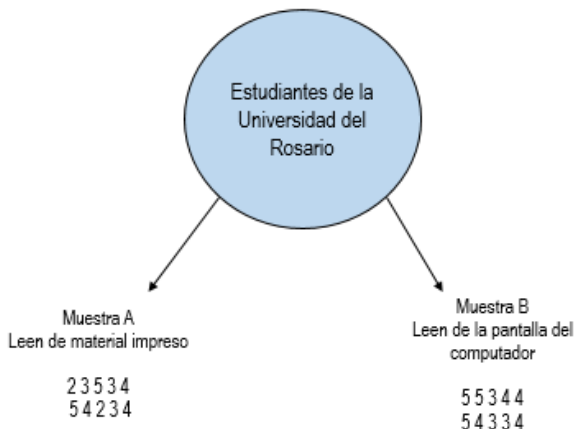


El rol de la estadística en la investigación

- La Universidad del Rosario quiere evaluar dos métodos de estudio diferentes, para así descubrir cuál es más eficiente.
- El propósito es demostrar si los estudiantes aprenden mejor leyendo de material impreso o de la pantalla de un computador.
- Dos muestras son seleccionadas de la población de estudiantes de la universidad.

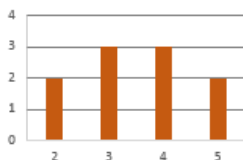
El rol de la estadística en la investigación

Paso 1: Experimento y Recolección de Datos

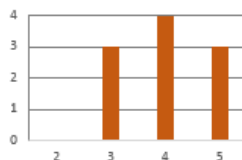


El rol de la estadística en la investigación

Paso 2: Estadística Descriptiva



Nota Promedio A=3.5



Nota Promedio B=4

El rol de la estadística en la investigación

Paso 3: Estadística Inferencial

- Los datos de la muestra presentan una diferencia de 5 décimas entre los dos métodos de estudio. Sin embargo, hay dos maneras de interpretar los resultados:
 - 1 Realmente no existe diferencia entre los dos métodos, y la diferencia muestral es debido al error de muestreo.
 - 2 Existe una diferencia entre los dos métodos, y la muestra refleja la diferencia de manera precisa.
- El objetivo de la estadística inferencial es ayudar a los investigadores a decidir entre ambas interpretaciones.

Constructo

- Algunas variables como altura, peso y edad son entidades concretas y bien definidas que pueden ser observadas y medidas de manera directa.
- Por otro lado, muchas otras variables son características internas que no pueden ser cuantificadas ni observadas directamente (e.g., la inteligencia, la ansiedad, la democracia, etc.).

Definición

Un constructo es un atributo interno o característica que no puede ser directamente observada, pero es útil para describir y explicar a los *individuos*.

Definición Operacional

- Aunque los constructos son características internas no observables, es posible medir y observar atributos que son análogos de estos constructos.
- Estos atributos externos pueden ser usados para crear una representación del constructo. Por ejemplo, el IQ para la inteligencia.

Definición

Una definición operacional identifica un instrumento con el cual describir un constructo interno.

Tipos de Variables

Según su valor

- 1 Discreta: Consiste en categorías separadas e indivisibles. No existen valores entre dos categorías vecinas.
- 2 Continua: Consiste de un número infinito de posibles valores. Además, es divisible en un número infinito de partes.

Tabla 1. Algunos ejemplos

Discreta	Continua
Número de Hijos	Altura
Sexo	Nota de Clase
Carrera	Ingreso

Tipos de Variables

Según su escala

- Escala Nominal: Consiste en un conjunto de categorías que tienen diferentes nombres.
 - ① Estado Civil.
 - ② Carrera.
- Escala Ordinal: Consiste en un conjunto de categorías que están organizadas.
 - ① Satisfacción con el gobierno.
 - ② Estrato Socioeconómico.
- Escala Proporcional o de Razón: Consiste en valores que tienen las propiedades aritméticas.
 - ① Ingreso laboral.
 - ② Número de hijos.

Distribuciones de Frecuencias (1)

- Luego de examinar el tipo de variables que se tiene, el siguiente paso consiste en analizar cada variable por separado.
- ¿Cuántos sujetos caen en cada categoría o en cada valor?

Definición

Una distribución de frecuencia es una tabulación organizada del número de sujetos localizados en cada categoría en la escala de medida.

Distribuciones de Frecuencias (2)

- Una distribución de frecuencias puede ser estructurada como una tabla o como una gráfica.
- Aún así, la distribución presenta los dos mismos elementos:
 - 1 El conjunto de categorías.
 - 2 Un registro de la frecuencia, o del número de sujetos en cada categoría.
- Una distribución de frecuencias presenta el panorama de cómo los valores de los sujetos se distribuyen sobre la escala de medida.

Tabla de Distribución de Frecuencias

- Supongamos que tenemos la variable X , que se define como la variable *nota de un quiz de esta clase*.
- $X = \{3 \ 3 \ 4 \ 5 \ 4 \ 4 \ 1 \ 4 \ 5 \ 3\}$

Tabla 2. Distribución de Frecuencia

X	f
5	2
4	4
3	3
1	1

- Cabe resaltar que $\sum f = N$.

Proporción y Porcentaje

- Adicional a las dos columnas básicas de la diapositiva anterior, hay otras medidas que describen la distribución de una variable.
- Las dos más comunes son la proporción y el porcentaje.
- La proporción mide la fracción del total del grupo que está asociado a ese valor.

$$p = \frac{f}{N}$$

- El porcentaje es la proporción multiplicada por 100. Su interpretación es la más sencilla, por tanto, es la más presentada en las tablas de distribución.

$$\text{porcentaje} = p * 100 = \frac{f}{N} * 100$$

- Al añadirle estas dos últimas medidas, la tabla 2 queda de la siguiente manera:

Tabla 3

X	f	p	%
5	2	0.2	20 %
4	4	0.4	40 %
3	3	0.3	30 %
1	1	0.1	10 %

- Con respecto a variables nominales u ordinales, la diferencia es que en vez de ser valores puntuales se tienen categorías.

Agrupación de Datos de Razón

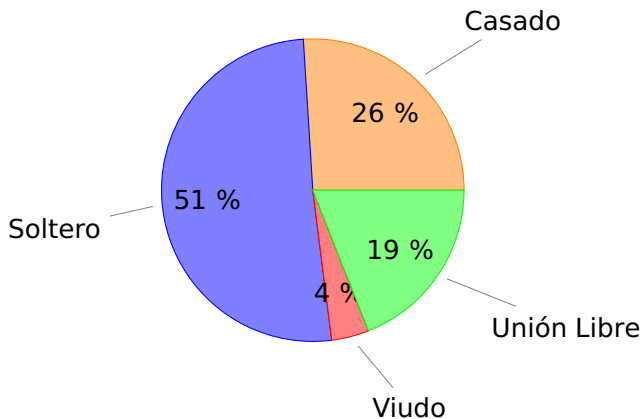
- En muchas ocasiones, el mínimo y el máximo de los valores de una variable distan mucho entre sí. Por ejemplo, la edades en un censo.
- Los valores varían entre 0 y 100 años, lo cual da 100 puntuaciones.
- Por este motivo, se combinan las edades en categorías de, por ejemplo, 10 años.

Tabla 4

X=Grupo de Edad	f	%
0-9	5	10.4 %
10-19	6	11.5 %
20-29	15	31.2 %
.	.	.
.	.	.
90 y mayor	8	13.8 %

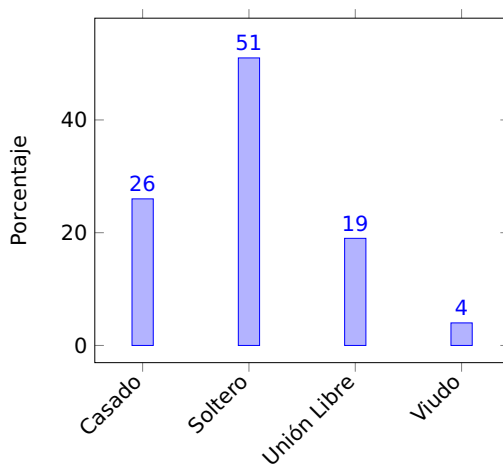
Graficar Variables Nominales u Ordinales

El gráfico de pastel



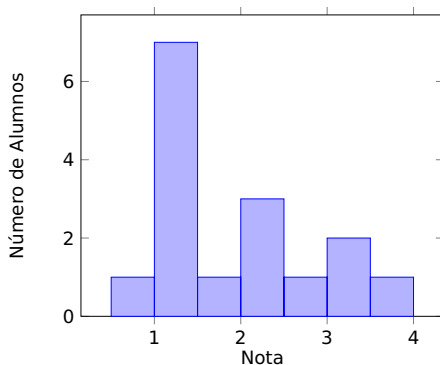
Graficar Variables Nominales u Ordinales

El gráfico de barras



Graficar Variables de Razón

El histograma



Graficar Variables de Razón

Tabla 5

Notas de Clase	f
0.5	1
1	7
1.5	1
2	3
2.5	1
3	2
3.5	1