

# Introducción a la Estadística y al Análisis de Datos



UCR – ECCI

CI-0115 Probabilidad y Estadística

Prof. Kryscia Daviana Ramírez Benavides



# Fundamentos de la Estadística

- La **estadística** recoge, ordena y analiza datos para estudiar las características o el comportamiento de un colectivo.
  - Es el conjunto de métodos científicos ligados a la toma, organización, recopilación, presentación y análisis de datos, tanto para la deducción de conclusiones como para tomar decisiones razonables de acuerdo con tales análisis.
- Ejemplos de estudios estadísticos:
  - Opinión de los costarricenses sobre el TLC.
  - El color favorito de los alumnos de una clase.
  - Número de goles marcados por los equipos de fútbol de Primera División.
  - Estatura de los alumnos de la clase de BD de la ECCI.

## Fundamentos de la Estadística (cont.)

- Terminología estadística:
  - **Población:** Conjunto de todos los elementos que cumplen una determinada característica. Conjunto de todos los valores de una variable aleatoria. Los elementos de la población se llaman individuos o unidades estadísticas.
  - **Muestra:** Cualquier subconjunto de la población.
  - **Carácter estadístico:** Es una propiedad que permite clasificar a los individuos de una población. Se distinguen dos tipos:
    - **Cuantitativo:** Son aquellos que se pueden medir.
    - **Cualitativos:** Son aquellos que no se pueden medir.
  - **Variable estadística:** Es el aspecto que se va a estudiar. Si se puede medir se llama variable cuantitativa si no se pueden medir se llama variable cualitativa.



## Fundamentos de la Estadística (cont.)

- La **variable estadística** es una propiedad característica de la población que estamos interesados en estudiar.
- Tipos de variables estadísticas:
  - **Cualitativa:** No se expresa mediante un número. Por ejemplo, la opinión de los costarricenses sobre el TLC.
  - **Cuantitativa:** Se expresa mediante un número, hay dos tipos:
    - **Cuantitativa Discreta:** Sólo admite valores aislados, toma un número determinado de valores. Por ejemplo, el número de goles marcados por los equipos de fútbol de Primera División.
    - **Cuantitativa Continua:** Puede admitir cualquier valor dentro de un intervalo, puede tomar cualquier valor entre los valores dados. Por ejemplo, la estatura de los alumnos de la clase de BD de la ECCI.



## Fundamentos de la Estadística (cont.)

- Ejemplos: Variable Estadística Cuantitativa Discreta

Calificaciones	Número de Alumnos
1	0
2	1
3	0
4	4
5	5
6	3
7	3
8	2
9	1
10	1

## Fundamentos de la Estadística (cont.)

- Ejemplos: Variable Estadística Cuantitativa Continua

Talla	Número de Alumnos
[155,160[	5
[160,165[	2
[165,170[	3
[170,175[	4
[175,180[	3
[180,185[	3



## Fundamentos de la Estadística (cont.)

- Una **variable no aleatoria** (asociada al resultado de una experiencia que sí produce el mismo resultado) está caracterizada por un valor para cada condición.
- Una **variable aleatoria** está caracterizada por la llamada función densidad de probabilidad, a partir de la cual se obtienen las probabilidades para sus posibles valores para cada condición.



## Fundamentos de la Estadística (cont.)

- Los objetivos se pueden entender en términos de encontrar y describir las variables de interés y las relaciones entre ellas, para el problema en estudio.
- La estadística es la ciencia que estudia los métodos que permiten realizar este proceso para variables aleatorias, y se divide en dos áreas:
  - **Estadística descriptiva:** Trata de describir las variables aleatorias en las muestras.
  - **Estadística inductiva o inferencial:** Trata de la generalización hacia las poblaciones de los resultados obtenidos en las muestras y de las condiciones bajo las cuales estas conclusiones son válidas.





## Fundamentos de la Estadística (cont.)

- Existen dos formas de hacer inferencia estadística:
  - **La estimación de parámetros:** A partir de una muestra se proponen dos objetivos:
    - Obtener valores aproximados de parámetros poblacionales con la **estimación puntual** (medidas de ciertas muestras).
    - La **estimación por intervalos de confianza** tiene por objeto proporcionar, a partir de la información recogida en la muestra, un **intervalo que contenga con alto nivel de confianza (probabilidad), al parámetro** objeto de nuestro interés.
      - A partir de dicho intervalo obtendremos una medida del error máximo cometido al aproximar puntualmente el parámetro.



## Fundamentos de la Estadística (cont.)

- Existen dos formas de hacer inferencia estadística:
  - **Las pruebas de hipótesis:** Si se quiere decidir entre dos hipótesis que afectan a un cierto parámetro de la población, a partir de la información de la muestra se usa el **contraste de hipótesis**, cuando se obtiene por una de estas dos hipótesis, se habrá de conocer una medida del error cometido, es decir, cuantas veces de cada cien se cometió un error.
    - Es importante saber cómo se escribirían las hipótesis que se quieren contrastar:
      - **H<sub>0</sub>** se llama **hipótesis nula** y es lo contrario de lo que se sospecha que va a ocurrir.
      - **H<sub>1</sub>** se llama **hipótesis alternativa** y es lo que se sospecha que va a ser cierto.



## Fundamentos de la Estadística (cont.)

- En la inferencia estadística hay varios métodos, pero en cualquier caso es necesario utilizar una muestra que represente a la población, esto se consigue con las **técnicas de muestreo**.
- Para estudiar una población se puede hacer de dos formas:
  - Un **estudio exhaustivo o censo** es cuando la investigación se ha hecho sobre todos y cada uno de los elementos que constituyen la población.
  - Un **estudio por muestreo** cuando la investigación se ha realizado única y exclusivamente sobre una muestra (un subconjunto limitado y convenientemente seleccionado de la población).



## Fundamentos de la Estadística (cont.)

- Distinguimos dos tipos fundamentales de muestreo:
  - **Muestreo no probabilístico:** En este tipo de muestreo, puede haber clara influencia de la persona o personas que seleccionan la muestra o simplemente se realiza atendiendo a razones de comodidad.
    - Salvo en situaciones muy concretas en la que los errores cometidos no son grandes, debido a la homogeneidad de la población, en general no es un tipo de muestreo riguroso y científico, dado que no todos los elementos de la población pueden formar parte de la muestra.
    - Por ejemplo, si hacemos una encuesta telefónica por la mañana, las personas que no tienen teléfono o que están trabajando, no podrán formar parte de la muestra.





## Fundamentos de la Estadística (cont.)

- Distinguimos dos tipos fundamentales de muestreo:
  - **Muestreo probabilístico:** En este tipo de muestreo, todos los individuos de la población pueden formar parte de la muestra, tienen probabilidad positiva de formar parte de la muestra.
    - Por lo tanto es el tipo de muestreo que deberemos utilizar en nuestras investigaciones, por ser el riguroso y científico.



## Fundamentos de la Estadística (cont.)

- En el contexto de muestreo probabilístico, existen varias posibilidades de obtención de una muestra:
  - **Muestreo aleatorio simple:** Todos los individuos tienen la misma probabilidad de ser seleccionados. La selección de la muestra puede realizarse a través de cualquier mecanismo probabilístico en el que todos los elementos tengan las mismas opciones de salir.
    - Por ejemplo, uno de estos mecanismos es utilizar una tabla de números aleatorios, o un generador de números aleatorios comprendidos entre cero y uno, y multiplicarlos por el tamaño de la población.



## Fundamentos de la Estadística (cont.)

- En el contexto de muestreo probabilístico, existen varias posibilidades de obtención de una muestra:
  - **Muestreo aleatorio estratificado:** Es frecuente que cuando se realiza un estudio interese estudiar una serie de subpoblaciones (estratos) en la población, siendo importante que en la muestra haya representación de todos y cada uno de los estratos considerados. El muestreo aleatorio simple no nos garantiza que tal cosa ocurra. Para evitar esto, se saca una muestra de cada uno de los estratos.



## Fundamentos de la Estadística (cont.)

- En el contexto de muestreo probabilístico, existen varias posibilidades de obtención de una muestra:
  - **Muestreo aleatorio estratificado:** Hay dos conceptos básicos:
    - **Estratificación:** El criterio a seguir en la formación de los estratos será formarlos de tal manera que haya la máxima homogeneidad en relación a la variable a estudio dentro de cada estrato y la máxima heterogeneidad entre los estratos.
    - **Afijación:** Reparto del tamaño de la muestra en los diferentes estratos o subpoblaciones. Existen varios criterios entre los que se destaca:
      - **Afijación igual:** Todos los estratos tienen el mismo número de elementos en la muestra.
      - **Afijación proporcional:** Cada estrato tiene un número de elementos en la muestra proporcional a su tamaño.





## Papel de la Probabilidad

- La probabilidad permite cuantificar la fuerza o “confianza” en las conclusiones.
  - Las conclusiones se pueden resumir en una declaración de probabilidad o valor  $P$ .
- El puente entre los datos y la conclusión está basado en los fundamentos de la inferencia estadística, la teoría de la distribución y las distribuciones de muestreos.
- Es importante prestar atención a las medidas de posición y variabilidad de la muestra.
  - Tanto el análisis exploratorio o intuitivo de datos como la inferencia estadística formal dependen de estas medidas.



## Medidas de Posición

- Las medidas de posición en un conjunto de datos están diseñadas para proporcionar al analista alguna medida cuantitativa de dónde el centro de los datos en una muestra.
- Una medida usada es la **moda de la muestra**; esta medida se asocia al valor más común, más típico o que ocurre más frecuentemente; se puede definir formalmente como el valor al cual corresponde la mayor frecuencia.
  - La principal limitación está en el hecho de que requiere un número suficiente de observaciones para que se manifieste o se defina más claramente.
  - En algunos casos puede no existir, no estar definida; e incluso si existe puede no ser única.
  - Esta medida se aplica a variables cualitativas y cuantitativas.

## Medidas de Posición (cont.)

- Una medida obvia y muy útil es la **media de la muestra** (media aritmética). La media es simplemente un promedio numérico.
- Suponga que las observaciones en una muestra son  $x_1, x_2, \dots, x_n$ . Entonces, la media simple de la muestra es

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

## Medidas de Posición (cont.)

- Suponga que las observaciones en una muestra se agrupan en  $x_1, x_2, \dots, x_k$  y c/u tiene una frecuencia  $f_1, f_2, \dots, f_k$ . Entonces, la **media ponderada** de la muestra es

$$\bar{x} = \sum_{i=1}^k \frac{x_i f_i}{n} = \frac{x_1 f_1 + x_2 f_2 + \dots + x_k f_k}{n}$$

$$n = f_1 + f_2 + \dots + f_k = \sum_{i=1}^k f_i$$



## Medidas de Posición (cont.)

- Otra medida importante es la **mediana de la muestra**, el propósito de esta medida es reflejar la tendencia central de la muestra de manera que no esté influida por los valores extremos.
  - Se puede definir como el valor central de una serie de variables estadísticas ordenadas.
- Dado que las observaciones de una muestra son  $x_1, x_2, \dots, x_n$ , acomodados en orden creciente de magnitud, la mediana de la muestra es

$$\tilde{x} = \begin{cases} x_{(n+1)/2} & \text{si } n \text{ es impar} \\ x_{n/2}, x_{(n/2)+1} & \text{si } n \text{ es par} \end{cases}$$

## Medidas de Posición (cont.)

- Ejemplo: Supongamos que el conjunto de datos es 1.7, 2.2, 3.9, 3.11 y 14.7.

- La moda, la media y la mediana de la muestra son:

$$n = 5 \quad \text{es impar}$$

$$Mo = \emptyset$$

$$\bar{x} = \frac{1.7 + 2.2 + 3.9 + 3.11 + 14.7}{5} = 5.122$$

$$\tilde{x} = x_{(5+1)/2} = x_3 = 3.9$$

## Medidas de Posición (cont.)

- La media está influida de manera considerable por la presencia de la observación extrema, mientras que la mediana hace énfasis en el verdadero “centro” del conjunto de datos.
- Las propiedades de la media son:
  - Si se multiplica la media por el número de observaciones se obtiene la suma de las observaciones.
  - Si a cada una de las observaciones se le resta la media, y luego se suman sus desviaciones (o diferencias) la suma resulta igual a 0.
  - Si se suma (o se resta) una constante  $b$  a cada una de las observaciones, el promedio queda aumentado (o disminuido) en esa constante  $b$ .
  - Si se multiplica (o se divide) cada una de las observaciones por una constante  $b$ , el promedio queda multiplicado (o dividido) por esa constante  $b$ .



## Medidas de Variabilidad

- La variabilidad de una muestra juega un papel muy importante en el análisis de datos.
  - La variabilidad de un proceso y de un producto es un hecho real en los sistemas científicos y de ingeniería.
  - La variabilidad en valores de población y datos de una muestra es un hecho real.
- Una variabilidad grande entre las observaciones de una muestra a menudo puede “limpiar” cualquier efecto que el ingeniero puede tratar de detectar.
- Las medidas de posición en una muestra no proporcionan un resumen apropiado de la naturaleza de un conjunto de datos.



## Medidas de Variabilidad (cont.)

- Al igual que hay muchas medidas de tendencia central o de posición, hay muchas medidas de dispersión o variabilidad.
- Una de las más simples es el **rango de la muestra** (recorrido o amplitud), el cual puede ser muy útil sobre control estadístico de calidad.

$$x_{\max} - x_{\min}$$

- Valores mínimo y máximo de una variable estadística.
- Aprecia la variabilidad al considerar los valores extremos de la muestra.
- Otra forma de dar el rango de la muestra es decir: recorrido  $x_{\min}$  a  $x_{\max}$ .

## Medidas de Variabilidad (cont.)

- La más utilizada es la **desviación estándar de la muestra**. Supongamos que los valores de una muestra son  $x_1, x_2, \dots, x_n$ , la desviación estándar es:

$$s = \sqrt{\sum_{i=1}^n \frac{(x_i - \bar{x})^2}{(n-1)}}$$

- Medida estadística que muestra que tan dispersos están los valores en una colección de datos.
- Término estadístico que mide la desviación de los datos de una muestra o de una población respecto a su valor medio.
- Indica cuánto se alejan, en promedio, las observaciones de la media del conjunto.

## Medidas de Variabilidad (cont.)

- La **varianza** de la muestra se calcula como:

$$s^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{(n-1)}$$

- Medida estadística que muestra la variabilidad de un valor; a mayor varianza, mayores variaciones con respecto al promedio y en consecuencia, mayor volatilidad.
  - Es la media de las diferencias al cuadrado de  $n$  puntuaciones respecto a su media aritmética.
- Una variabilidad grande en un conjunto de datos produce valores relativamente grandes de  $(x - \bar{x})^2$  y por ello una varianza de la muestra grande.



## Medidas de Variabilidad (cont.)

- La cantidad  $n - 1$  a menudo se denomina **grados de libertad asociados con la varianza** estimada.
- Los grados de libertad representan el número de piezas de información independientes disponibles para calcular la variabilidad.



## Medidas de Variabilidad (cont.)

- Suponga que las observaciones en una muestra se agrupan en  $x_1, x_2, \dots, x_n$  y c/u tiene una frecuencia  $f_1, f_2, \dots, f_n$ . Entonces, la desviación estándar y la varianza de la muestra son:

$$s = \sqrt{\frac{\sum_{i=1}^k (x_i - \bar{x})^2 f_i}{\left(\sum_{i=1}^k f_i\right) - 1}}$$
$$s^2 = \frac{\sum_{i=1}^k (x_i - \bar{x})^2 f_i}{\left(\sum_{i=1}^k f_i\right) - 1}$$

## Medidas de Variabilidad (cont.)

- La **covarianza** de la muestra se calcula como:

$$s_{xy} = \sum_{i=1}^n \sum_{j=1}^k \frac{(x_i - \bar{x})(y_j - \bar{y})(n_{ij})}{n}$$

- Medida estadística que muestra la relación entre dos variables; una covarianza 0 indica que no hay relación alguna, mientras que valores altos implican que el valor de una está muy ligado al valor de la otra variable.

## Medidas de Variabilidad (cont.)

- Ejemplo: Supongamos que se tiene el conjunto de datos es 5, 17, 6 y 4.

- El promedio de la muestra es:  $\bar{x} = \frac{5+17+6+4}{4} = \frac{32}{4} = 8$

- La varianza de la muestra y la desviación estándar de la muestra son:

$$s^2 = \frac{(5-8)^2 + (17-8)^2 + (6-8)^2 + (4-8)^2}{3}$$

$$s^2 = \frac{(-3)^2 + (9)^2 + (-2)^2 + (-4)^2}{3} = \frac{110}{3} = 36.667$$

$$s = \sqrt{\frac{(5-8)^2 + (17-8)^2 + (6-8)^2 + (4-8)^2}{3}}$$

$$s = \sqrt{\frac{(-3)^2 + (9)^2 + (-2)^2 + (-4)^2}{3}} = \sqrt{\frac{110}{3}} = \sqrt{36.667} = 6.055$$

## Medidas de Variabilidad (cont.)

- En el ejemplo anterior, las cantidades dentro de los paréntesis suman cero:

$$\sum_{i=1}^n (x_i - \bar{x}) = (5 - 8) + (17 - 8) + (6 - 8) + (4 - 8) = -3 + 9 - 2 - 4 = 0$$

- Entonces, el cálculo de una varianza de la muestra no implica  $n$  desviaciones cuadradas independientes de la media  $\bar{x}$ .
- De hecho, como el valor de  $x - \bar{x}$  está determinado por los primeros  $n - 1$  valores, se dice que estas son  $n - 1$  “piezas de información” que producen  $s^2$ .
- Por ello hay  $n - 1$  grados de libertad en lugar de  $n$  grados de libertad para calcular la varianza de la muestra.





## Datos Discretos y Continuos

- Los datos reunidos en las observaciones o experimentos pueden ser **discretos** o **continuos**, dependiendo del área de aplicación.
  - Los datos discretos son un número finito o infinito numerable de elementos.
    - Se puede tener un experimento de combinación de fármacos, donde se puede encontrar datos binarios por naturaleza (el paciente responde o no).
  - Los datos continuos son un número infinito no numerable de elementos.
    - Se puede tener un experimento de querer maximizar determinada producción, la cual puede ser en porcentaje, o gramos/libras, medida en un continuo.



## Datos Discretos y Continuos (cont.)

- Los datos están asociados a una variable aleatoria, la cual es una función que asocia un número real, perfectamente definido, a cada punto muestral.
  - Las v.a. definidas sobre espacios muestrales discretos se llaman v.a. discretas y las definidas sobre espacios muestrales continuos se llaman continuas.
  - Una v.a. puede ser continua, aunque sólo se pueda acceder a un subconjunto finito de valores; por ejemplo, la presión arterial es una v.a. continua, pero sólo se puede acceder a un conjunto finito de valores por la limitación de los aparatos de medida.
  - En general, las medidas dan lugar a v.a. continuas y los conteos a v.a. discretas.



# Modelado Estadístico, Inspección Científica y Diagnósticos Gráficos

- Muy a menudo el resultado final de un análisis estadístico es la estimación de los parámetros de un **modelo postulado**.
- Un modelo estadístico no es determinista, sino debe involucrar algunos aspectos probabilistas.
  - Una forma de modelo es la fundamentación de las **suposiciones** que hace el analista.
  - El formalismo del análisis implica nociones de prueba de hipótesis.
- El tipo de modelo que se utiliza para describir los datos a menudo depende del objetivo del experimento, y la estructura del modelo debe aprovecharse del insumo científico no estadístico.



## Modelado Estadístico, Inspección Científica y Diagnósticos Gráficos (cont.)

- La selección del modelo representa una suposición fundamental sobre la que se basa la inferencia estadística resultante.
- A veces, los gráficos o el análisis exploratorio de los datos puede enseñar al analista algo no recuperado del análisis formal.
  - Casi cualquier análisis formal requiere suposiciones que se desarrollan a partir del modelo de datos.
  - Los gráficos pueden resaltar bien la violación de suposiciones que de otra forma no se notarían.





## Modelado Estadístico, Inspección Científica y Diagnósticos Gráficos (cont.)

- **Proceso de un estudio estadístico:** Se debe tener claro que la información estadística se resume en tablas y gráficos, pero antes, hay que seguir una serie de pasos:
  - **Decidir qué se quiere estudiar:** Establecer claramente la variable estadística, acotando, claramente, la población de estudio.
  - **Elaborar una encuesta de recolección de datos:** Crear una serie de preguntas breves y claras con respuesta cerrada.
  - **Recolección de datos:** Realizar las encuestas, procurando preguntar a todo tipo de personas.
  - **Organizar los datos:** Ordenar, se pasan a papel o computador y se confeccionan primero las tablas y luego los gráficos.

## Tablas de Referencia

- La tabulación de los datos consiste en confeccionar una tabla en la que aparecen bien organizados los valores de la variables que se están estudiando, junto con otros datos:
  - **Frecuencia absoluta ( $n_j$ ):** Número de individuos que toma cada valor.
  - **Frecuencia relativa ( $f_j$ ):** Resultado de dividir la frecuencia absoluta entre el total de la población  $\rightarrow f_j = n_j/N$ .
  - **Frecuencia absoluta acumulada ( $N_j$ ):** Suma de las frecuencias absolutas de los valores menores o iguales que él (sólo tiene sentido para variables estadísticas cuantitativas).
  - **Frecuencia relativa acumulada ( $F_j$ ):** Suma de las frecuencias relativas de los valores menores o iguales que él.

## Tablas de Referencia (cont.)

- Ejemplo: Número de días a la semana que practican deporte 20 alumnos de la clase de IO de la ECCI: 1, 4, 4, 1, 1, 3, 4, 3, 1, 1, 3, 2, 1, 4, 1, 1, 2, 3, 3, 3.

Valores	$n_i$	$f_i$	$N_i$	$F_i$
1	8	$8/20 = 0.4$	8	0.4
2	2	$2/20 = 0.1$	$8+2 = 10$	$0.4+0.1 = 0.5$
3	6	$6/20 = 0.3$	$10+6 = 16$	$0.5+0.3 = 0.8$
4	4	$4/20 = 0.2$	$16+4 = 20$	$0.8+0.2 = 1$
<b>Total</b>	<b>20</b>	<b><math>20/20 = 1</math></b>	---	---



## Tablas de Referencia (cont.)

- Con variables cuantitativas continuas, o cuando el número de valores que toma la variable es grande, se agrupan los valores en intervalos.



## Tablas de Referencia (cont.)

- Ejemplo: Estatura en cm de 20 alumnos de la clase de IO de la ECCI: 154, 158, 162, 148, 163, 153, 159, 180, 165, 168, 156, 148, 162, 157, 153, 158, 147, 165, 166, 175. El valor mínimo es 147, el valor máximo es 180. Podemos tomar como amplitud del intervalo: 10cm (se podría tomar cualquier otro valor).

Intervalos	$n_i$	$f_i$	$N_i$	$F_i$
[145,155[	6	$6/20 = 0.3$	6	0.3
[155,165[	9	$9/20 = 0.45$	$6+9 = 15$	$0.3+0.45 = 0.75$
[165,175[	3	$3/20 = 0.15$	$15+3 = 18$	$0.75+0.15 = 0.9$
[175,185[	2	$2/20 = 0.1$	$18+2 = 20$	$0.9+0.1 = 1$
<b>Total</b>	<b>20</b>	<b><math>20/20 = 1</math></b>	---	---

# Gráficos Estadísticos

- Existen varios tipos de gráficos estadísticos, siendo cada uno de ellos adecuados a un tipo de variable, según el siguiente esquema:
  - Variables cualitativas → Diagrama de sectores.
  - Variables cuantitativas discretas → Diagramas de barras.
  - Variables cuantitativas continuas → Histograma de frecuencias.
- Herramienta de gráficos y diagramas:
  - [Catálogo de Visualización de Datos](#)
  - [How To Think Visually Using Visual Analogies](#)
  - [Data Viz Project](#)



## Gráficos Estadísticos (cont.)

- **Diagrama de sectores:** Consiste en dibujar en un círculo tantos sectores como valores toma la variable estadística, de forma que el ángulo de cada sector es proporcional a la frecuencia (absoluta o relativa).
- **Diagrama de barras:** Consiste en unas barras estrechas que se sitúan sobre los valores puntuales de la variable, que están situadas a lo largo de una recta numérica.



## Gráficos Estadísticos (cont.)

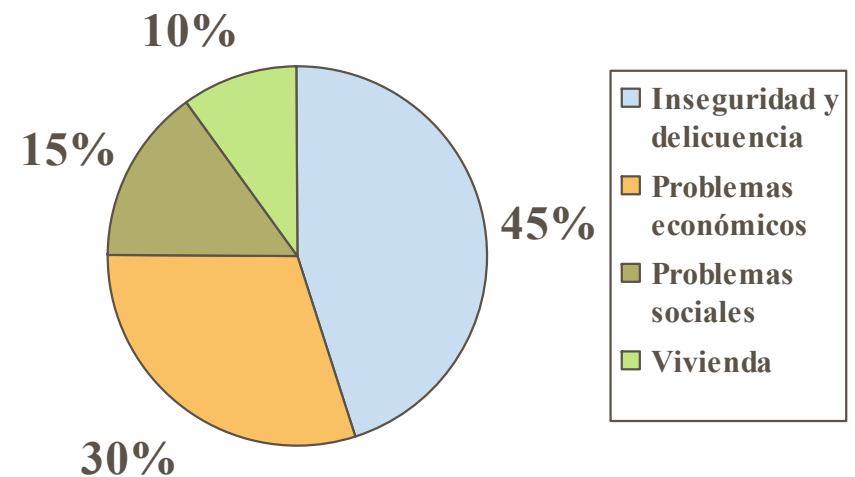
- **Histograma de frecuencias:** Es similar a un diagrama de barras, sólo que en este caso, las barras ocupan todo el ancho del intervalo al que van asociadas, pudiendo estar pegadas unas con otras (algo que nunca podía pasar en un diagrama de barras).
  - Realmente, a diferencia de lo que pasaba en un diagrama de barras, en los histogramas de frecuencias, el área de cada rectángulo debe ser proporcional a la frecuencia relativa, lo que pasa es que si tomamos la precaución de tomar todos los intervalos con la misma amplitud, entonces no tendremos que preocuparnos, por ser la razón de proporcionalidad siempre la misma: la amplitud del intervalo.



## Gráficos Estadísticos (cont.)

- En un estudio de Febrero de 2006, la juventud costarricense afirma que los principales problemas que existen en la sociedad son los siguientes:

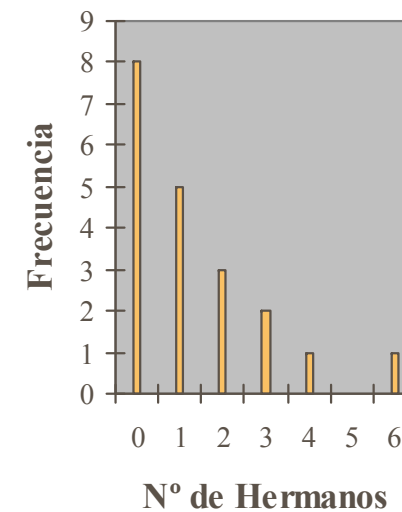
Problema	Porcentaje
Inseguridad y delincuencia	45%
Problemas económicos	30%
Problemas sociales	15%
Vivienda	10%



## Gráficos Estadísticos (cont.)

- El número de hermanos de los alumnos del curso se distribuye de la siguiente manera:

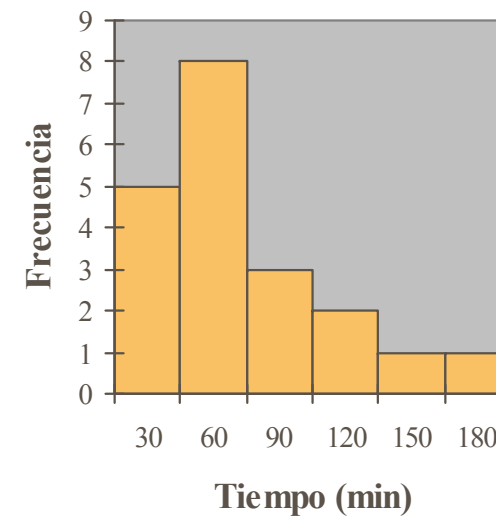
Nº de Hermanos	Frecuencia
0	8
1	5
2	3
3	2
4	1
5	0
6	1



## Gráficos Estadísticos (cont.)

- Tiempo que emplean los alumnos del curso en ir desde casa a la universidad se distribuye de la siguiente manera:

Tiempo (min)	Frecuencia
[0,30[	5
[30,60[	8
[60,90[	3
[90,120[	2
[120,150[	1
[150,180[	1



## Referencias Bibliográficas

- Walpole, R.E.; Myers, R.H.; Myers, S.L. & Ye, K. “Probabilidad y estadística para ingeniería y ciencias”. Octava Edición. Pearson Prentice-Hall. México, 2007.
- Material docente de la Unidad de Bioestadística Clínica. URL: [http://www.hrc.es/bioest/M\\_docente.html](http://www.hrc.es/bioest/M_docente.html).
- Iniciación a la Estadística. URL: [http://descartes.cnice.mecd.es/Estadistica/iniciacion\\_estadistica\\_fjgarcia/FGG990\\_UD.htm](http://descartes.cnice.mecd.es/Estadistica/iniciacion_estadistica_fjgarcia/FGG990_UD.htm).
- Estadística Descriptiva y Probabilidad. URL: <http://thales.cica.es/rd/Recursos/rd99/ed99-0278-01/indice.html>.



## Referencias Bibliográficas (cont.)

- Estadística Descriptiva. URL:  
[http://descartes.cnice.mecd.es/Estadistica/Estadistica\\_descriptiva\\_1/estadistica\\_indice.htm](http://descartes.cnice.mecd.es/Estadistica/Estadistica_descriptiva_1/estadistica_indice.htm).
- Estadística Inferencial. URL:  
[http://descartes.cnice.mecd.es/Estadistica/Muestreo\\_Inferencia\\_Estadistica/inferencia\\_estadistica.html](http://descartes.cnice.mecd.es/Estadistica/Muestreo_Inferencia_Estadistica/inferencia_estadistica.html).