Curso Práctico de Bioestadística Con Herramientas De Excel

Fabrizio Marcillo Morla MBA

barcillo@gmail.com
(593-9) 4194239

Fabrizio Marcillo Morla

- Guayaquil, 1966.
- BSc. Acuicultura. (ESPOL 1991).
- Magister en Administración de Empresas. (ESPOL, 1996).
- Profesor ESPOL desde el 2001.
- 20 años experiencia profesional:
  - Producción.
  - Administración.
  - Finanzas.
  - Investigación.
  - Consultorías.

Otras Publicaciones del mismo autor en Repositorio ESPOL

Capitulo 1

Conceptos Generales

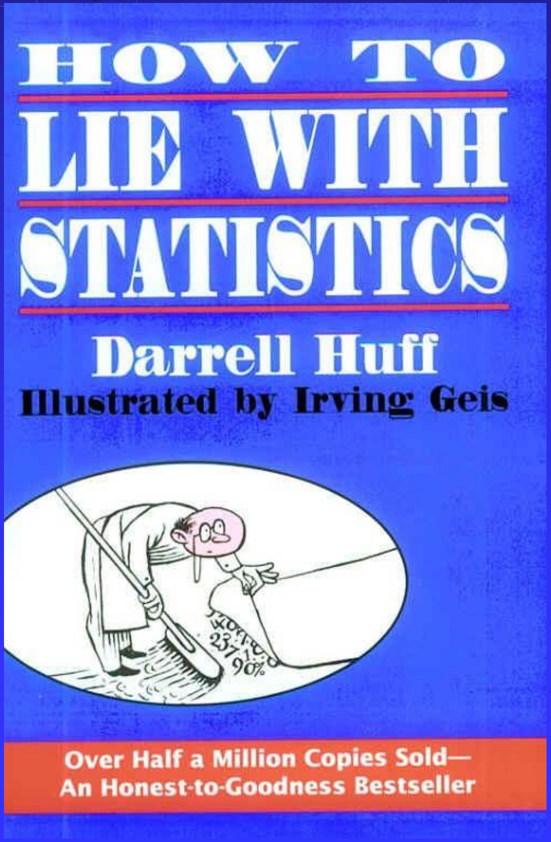
Que es Estadistica?

La ciencia pura y aplicada (no exacta), que crea, desarrolla y aplica técnicas de modo que pueda evaluarse la incertidumbre.

- Ciencia: "un conjunto de conocimientos comprobados y sistematizados".
- Pura: Por que estudia ciertos procesos teóricos.
- Aplicada: En cuanto se encarga de resolver problemas específicos.
- No exacta: No podemos obtener un resultado único, si no probabilidades de resultados esperados.

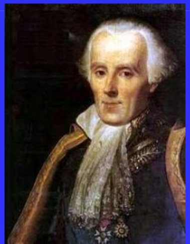
Otros Conceptos

- Estadísticas: *statísticum collegium*
  - Consejo de estado:
- Estadístico: Persona que al tener la cabeza en un horno y los pies en hielo dice “en promedio me siento muy bien”?
- Bestias Salvajes?
- No se puede generalizar?



Historia

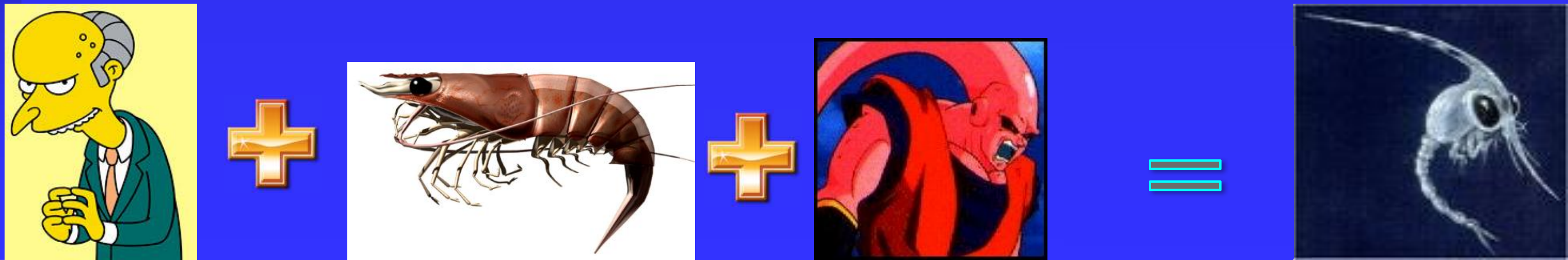
- Origenes en juegos de azar y censos.
- Siglo XVIII desarrolló teoría de Probabilidades (Gauss, Laplace, Bayes, etc). Discusión filosófica sigue hasta ahora.
- Muchas teorías, principalmente de carácter biológico como las de Mendel o Darwin tuvieron bases estadísticas
- Mayoría de métodos modernos se desarrollaron desde mediados del siglo XIX y principios del XX (Pearson, Student, Fisher…), principalmente para uso en biología, agricultura y genética



Porque BIOestadistica?

- “Muchos” biologistas desconfian de las matematicas: “los seres vivos son impredecibles”; “Camarones no se comportan como deberían”.
  - FIMA – QUIBIO????
- Métodos estadísticos desarrollados justo para ser usados en ciencias biológicas: Toman en cuenta variabilidad propia de poblaciones naturales en sus cálculos y tablas.



Tipos de Estadistica

- Descriptiva: Trata resumir e interpretar datos para poder describir una población.
  - Enumeración, organización y representación gráfica de los datos.
- Inferencial: Usa la teoría de la probabilidad para extraer conclusiones acerca de una población, a partir de la información incompleta de los datos obtenidos en una muestra:
  - Estimación
  - Comparativa
  - Predictiva

Aplicaciones de la Estadística

- Obtener una muestra.
- Resumir datos.
- Haciendo inferencias de una población, basado en los resultados de la muestra.
- Obteniendo un modelo más simple para un grupo de datos.

Variables

Una propiedad con respecto a la cual los individuos de una muestra o una población se diferencian en algo verificable.

- Ciertas "características" que presentan variación.

Tipos de Variables (1)

Clasificación por su escala de medición

- Cualitativas
  - <u>Dicotómicas-binarias</u>
    - Sexo: masculino o femenino.
    - Status de empleo: empleado o desempleado.
  - <u>Ordinales</u>
    - Nivel socioeconómico: alto, medio o bajo.
    - Índice de lípidos: 1, 2 ,3 ,4 , 5, 6
  - <u>Nominales</u>
    - Sitio de residencia: centro, sur, norte, este, oeste
    - Estado civil: soltero, casado, viudo, divorciado, unión libre, T.L.A.
  - <u>Variables de Intervalo</u>:
    - Fiebre: Si (>37º C), No (<37º C)
- Cuantitativas
  - <u>Discretas</u>
    - Número de hijos: 1,2,3,4.
  - <u>Continuas</u>
    - Nivel de glucosa en sangre: 110 mg/dl, 145 mg/dl.
    - Peso: 10 g, 11 g, 10.5 g, 10.1 g, 10.05 g, 10.001 g…

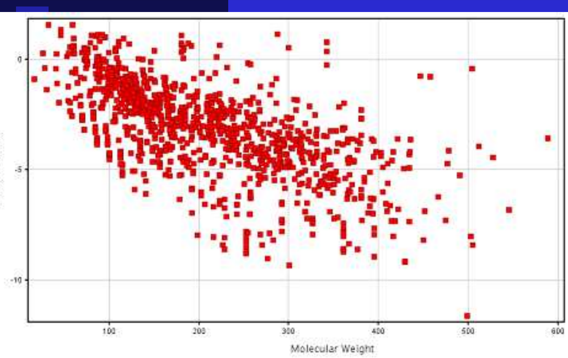
Tipos de Variables Cuantitativas

- Discretas: su conjunto de posibles valores son fijos, y no pueden tomar valores intermedios:
  - Número de peces en un acuario.
- Continuas: su conjunto de posibles valores puede alcanzar un número infinito entre dos valores cuales quiera:
  - Longitud, Peso

Tipos de Variables (2)

Clasificación por su relación

- Independientes: Aquellas cuyo valor no depende de otra variable.
- Dependientes: Aquellas cuyo valor va a depender de otra variable.



Dependiente de que?

$$X = \begin{cases} 1, & \text{if a 1 is rolled,} \\ 2, & \text{if a 2 is rolled,} \\ 3, & \text{if a 3 is rolled,} \\ 4, & \text{if a 4 is rolled,} \\ 5, & \text{if a 5 is rolled,} \\ 6, & \text{if a 6 is rolled.} \end{cases}$$

V. Intermedias y V. confusoras?

Ejemplos de Variables

De un ejemplo de cada una en su campo

- Cualitativas
  - Dicotómicas-binarias
  - Ordinales
  - Nominales
  - Variables de Intervalo:
- Cuantitativas
  - Discretas
  - Continuas

- Independientes y Dependientes

Valores, Datos, etc.

Llamamos Valores, Datos u Pbservaciones a cualquier valor numérico o cualitativo que mida una variable.

Los valores experimentales que va a tomar una variable determinada.

Variable: Peso

Valores: 125 lbs, 145 lbs, 180 lbs

Población

- El grupo de individuos bajo estudio sobre el que deseamos hacer alguna inferencia.
  - Conjunto de objetos, mediciones u observaciones del que tomamos muestra.
- Puede ser finita o infinita, dependiendo de su tamaño.
- Tamaño de población (número total de los individuos que la conforman) se lo denota con la letra N.

Ejemplos de poblaciones

- 20 camarones en una pecera.
- Todos los camarones de una piscina.
- Todos los camarones posibles a ser cultivados bajo cierto tratamiento.
- Todos los camarones del mundo.

Límites de población dependen de como la definamos nosotros acorde con nuestras necesidades.

Antes de empezar cualquier proceso estadístico es necesario Definir claramente la población o poblaciones bajo estudio.

Población y Muestra

- Conociendo la distribución de frecuencias de alguna característica (variable) de la población, es posible describirla por medio de una función de densidad, la cual a su vez vendrá caracterizada por ciertos parámetros.
- Problema es que al ser población muy grande, resulta más conveniente estudiar un subconjunto de ella (muestra) y decir que representa mas o menos fielmente (representativo) a la población total.

Muestra

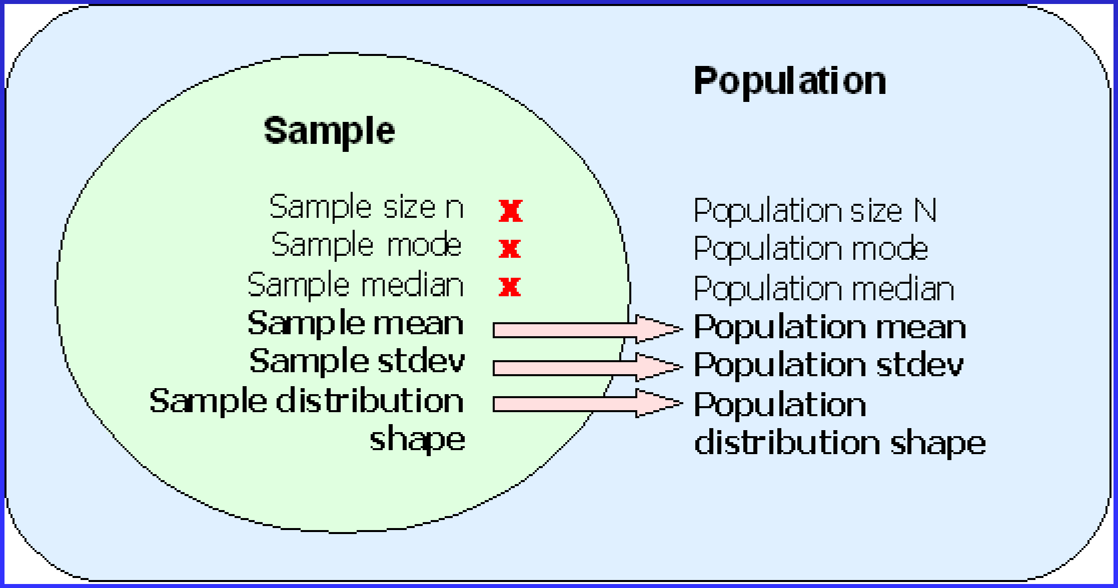
- Sea una variable aleatoria dada (X); los valores de esta variable aleatoria (X_1, X_2,...X_n) forman una muestra de la variable X, si ellas son independientes entre sí y siguen la misma distribución de X:
  - Representa fielmente a X.
- Partimos suposición: muestra es porción de población que la representa fielmente.
- No tomamos en cuenta muestreos mal realizados.
- Tamaño de la muestra se lo denota como n.

Muestra

- Igual que población, debe definirse correctamente antes de empezar estudio.
  - 20 camarones en una pecera.
  - Todos los camarones de una piscina.
  - Todos los camarones posibles a ser cultivados bajo cierto tratamiento.
  - Todos los camarones del mundo.
- Pueden representar una muestra de una población mayor.

Objetivos de Muestreo

- Obtener información sobre distribuciones de frecuencia de la población (distribución de probabilidad) o más preciso: de los parámetros poblacionales que describen dicha distribución de probabilidad.



Distribucion de Frecuencias (introduccion)

- Operación en que dividimos un conjunto de datos, en varios grupos, mostrando el número de elementos en cada grupo.
- Más tarde veremos los detalles.
- Ahora importante entender el concepto y su relación con la distribución de probabilidad.
- Archivo: Ejercicio01 - Distribucion de Frecuencias.xlsx

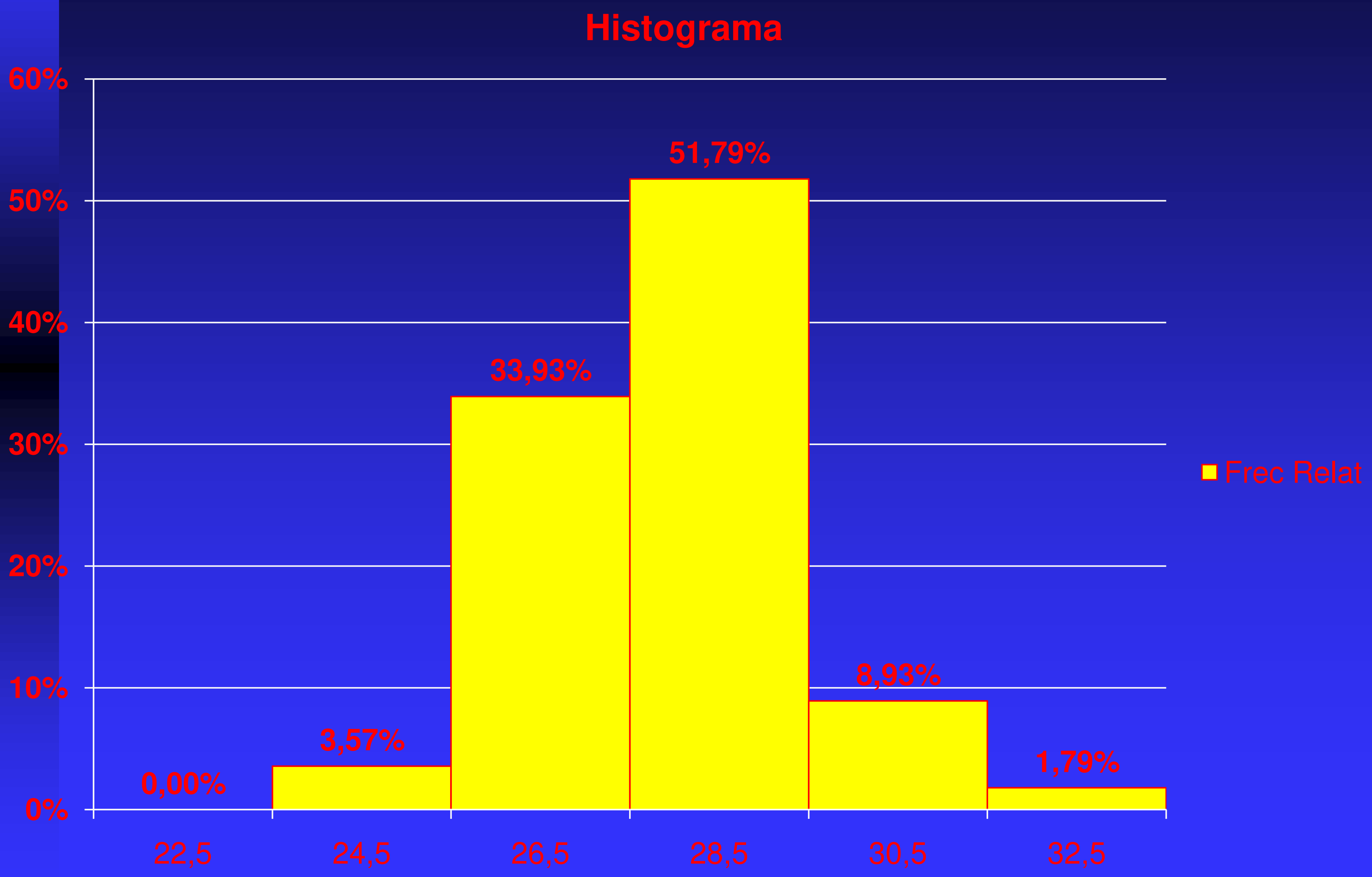
Datos longitud cefálica *O. niloticus*

25.3	26.3	27.0	27.7	28.2	29.0	29.5
25.5	26.4	27.0	28.0	28.2	29.3	29.5
26.0	26.4	27.0	28.0	28.4	29.3	29.7
26.0	26.6	27.0	28.0	28.5	29.3	29.8
26.0	26.8	27.3	28.0	28.9	29.4	30.2
26.0	27.0	27.6	28.0	28.9	29.4	31.0
26.0	27.0	27.6	28.0	29.0	29.5	31.0
26.1	27.0	27.6	28.1	29.0	29.5	33.4

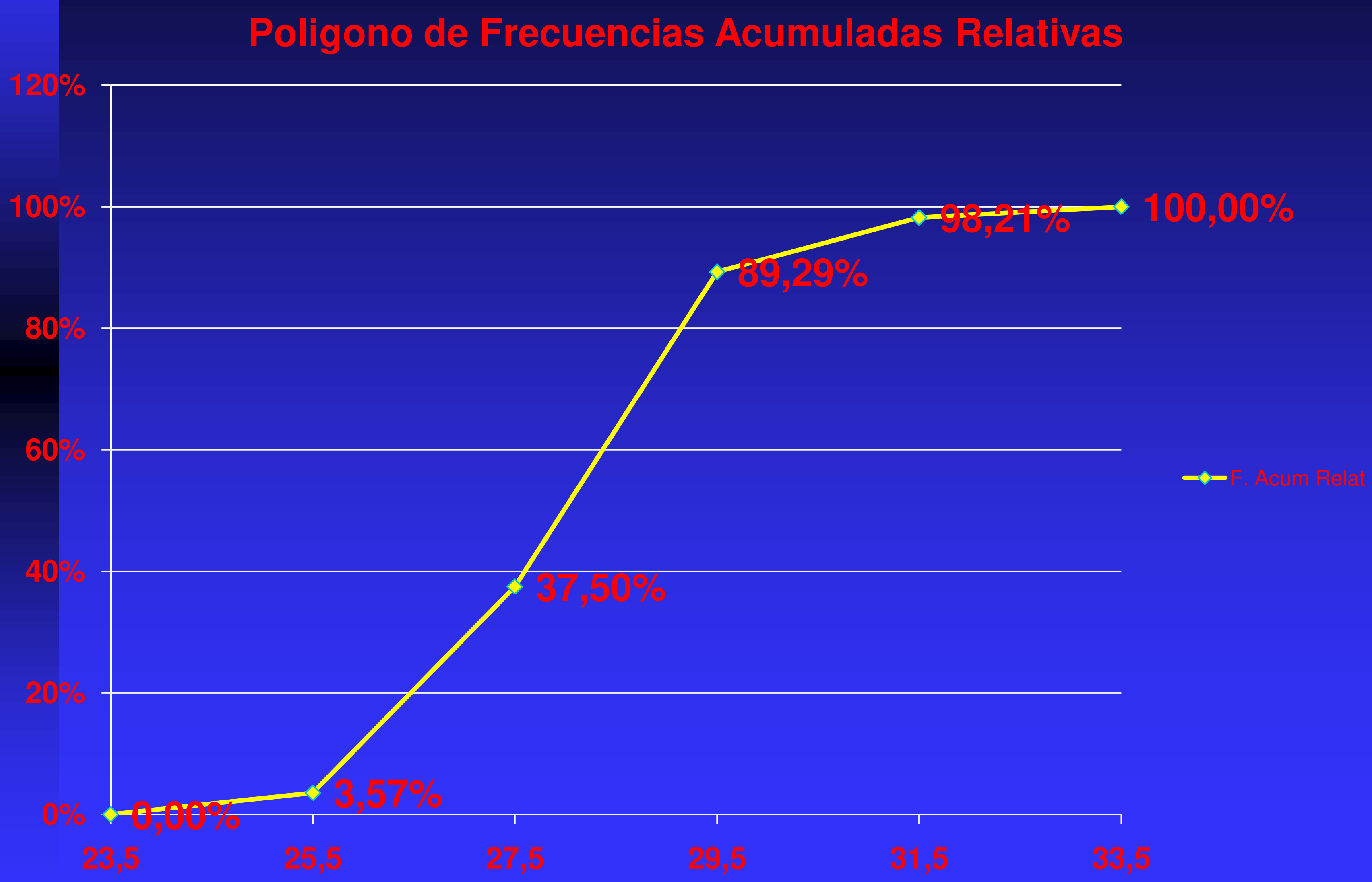
Tabla de Frecuencias

Int Repres		Int Real lim inf	Int Real lim sup	Frec	Frec Relat	Frec Acum	F. Acum Relat	Marca Clae
22	23	21.5	23.5	0	0.00%	0	0.00%	22.5
24	25	23.5	25.5	2	3.57%	2	3.57%	24.5
26	27	25.5	27.5	19	33.93%	21	37.50%	26.5
28	29	27.5	29.5	29	51.79%	50	89.29%	28.5
30	31	29.5	31.5	5	8.93%	55	98.21%	30.5
32	33	31.5	33.5	1	1.79%	56	100.00%	32.5
				56	100.00%			

Histograma


Histograma
60%
50%
40%
30%
20%
10%
0%
0,00%
3,57%
33,93%
51,79%
8,93%
1,79%
22,5
24,5
26,5
28,5
30,5
32,5
Frec Relat

Poligono de frecuencias



Ejercicio en Grupo

- 5 Grupos de 3 persona
- 2 Dados por grupo
- 1 hoja de Excel con 3 columnas :
  - 1 / cada dado
  - 1 suma de dados
- 2 personas lanzas al mismo tiempo pero por separado los dados. 60 veces
- Para cada dado y suma hacer:
  - Tabla de frecuencia
  - Histograma de frecuencia relativa
  - Poligono de frecuencia acumulada
  - Analizar

Teoria de Probabilidades

- Originó en juegos de azar
- O talvez antes?
- Todos jugamos al riesgo dia a dia.
- Varios enfoques filosóficos a probabilidad:
  - Teoria Clasica
  - Frecuentismo
  - Bayesiana
  - etc

Probabilidad

- Eventos que son comunes o improbables son aquellos cuya probabilidad de ocurrencia son grandes o pequeñas, respectivamente.
- Dia a dia calculamos "al ojo" la probabilidad de todas los sucesos que nos rodean
- Determinamos que tan "común" o "raras" son.
  - En Esmeraldas no es "común" encontrar un nativo rubio y ojos azules, en Suecia si.
  - Basado en "muestras" de Suecos y Esmeraldeños, sin necesidad de ver todos los esmeraldeños y suecos.
- Problema de este método al "ojímetro“: carecemos de un término preciso para describir la probabilidad.

Probabilidad

- Estadísticos reemplazan como "con dificultad", "pudo" o "casi con seguridad" por número de 0 a 1, que indica de forma precisa que tan probable o improbable es el evento.
- Haciendo inferencias sobre una población a partir de muestras no podemos esperar llegar siempre a resultados correctos.
- Estadística ofrece procedimientos para saber cuántas veces acertamos "en promedio". (enunciados probabilísticos).

Espacio Muestreal

- El conjunto universal de una población
- Todos los valores posibles que nuestra variable aleatoria puede tomar
  - Todas las formas en que podemos sacar 4 bolas de una funda que contenga 8 bolas rojas y 2 blancas
  - De cuantas formas puede caer un dado
  - Todas las posibles supervivencias que podamos obtener en un cultivo
  - Todos los posibles climas que puedan haber en un día determinado

Probabilidad Clasica

- Si un evento puede ocurrir de N maneras mutuamente exclusivas e igualmente posibles, y si n de ellas tienen una característica E, entonces, la posibilidad de ocurrencia de E es la fracción n/N y se indica por:

$$\mathrm{P}(E) = \frac{n}{N}$$

- Funciona bien con espacio muestreal pequeño y conocido, y en donde todas las N maneras sean igualmente posibles.

Probabilidad Frecuentista

- Probabilidad de un evento es su frecuencia relativa a lo largo del tiempo.

$$\lim_{n\to\infty} \frac{n_a}{n} = p$$

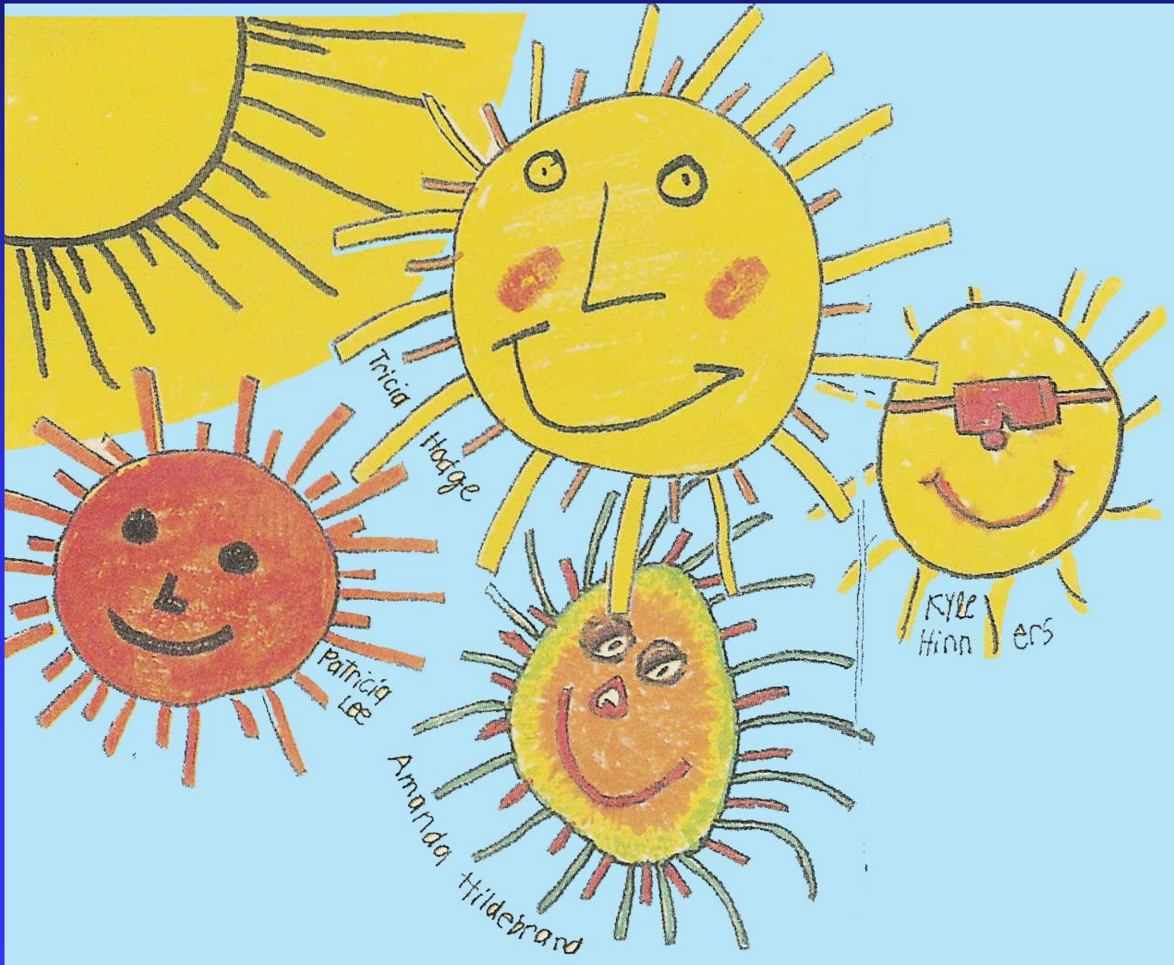
- Probabilidad de obtener “cara” al lanzar una moneda es 0.5: No porque se la calcula matemáticamente, sino porque esto ocurre al lanzarla muchas veces.
- No se puede repetir experimento infinitas veces.
- Al repetirlo pocas veces da distinta probabilidad.
- Error de probabilidad es una probabilidad… bis…

Probabilidad

- La probabilidad que un carro sea robado en Guayaquil puede ser calculada en función al número de carros robados en y al número de carros en Guayaquil.
  - Aseguradoras usan esto, para calcular el valor esperado a pagar. +costos +utilidad = prima.
- Probabilidad que en cierta camaronera una corrida a 130.000 Pl/Ha alcance 15 gr. en 120 días puede ser calculada con base en veces que se ha logrado en condiciones similares

Ejercicio Individual

- Calcular la posibilidad de que el sol salga mañana.



Teoremas Basicos (1)

- La probabilidad de un evento cualquiera va a estar en el rango de cero a uno. Esto quiere decir que no existen probabilidades negativas ni mayores de 100%

$$0 \leq P(E) \leq 1$$

Teoremas Basicos (2)

- La suma de la probabilidad de ocurrencia de un evento mas la probabilidad de no ocurrencia del mismo es igual a uno.

$$P(E) + P(\neg E) = 1$$

- Probabilidad de que salga 1 en lanzamiento de dados es 1/6
- Ocurrencia de que no salga 1 es:

$$P(\neg 1) = 1 - 1/6 = 5/6$$

Teoremas Basicos (3)

- La probabilidad de ocurrencia de dos eventos independientes es igual al producto de la ocurrencia de cada uno.

$$P(A\ B) = P(A) \times P(B)$$

- Probabilidad de que al lanzar dos dados salga 1 y 2:
- $P(1) = 1/6$; $P(2) = 1/6$

$$P(1 \text{ y } 2) = 1/6 \times 1/6 = 1/36$$

Teoremas Basicos (4)

- Para dos eventos cualesquiera A y B, la probabilidad de que ocurra A o B viene dado, por la probabilidad de que ocurra A, mas la probabilidad de que ocurra B, menos la probabilidad de que ocurran ambos.

  $$P(A \text{ o } B) = P(A) + P(B) - P(AB)$$

- Probabilidad que al lanzar dos dados obtenga solo un 1 o un 2:
- $P(1) = 1/6$; $P(2) = 1/6$

  $$P(1 \text{ o } 2) = 1/6 + 1/6 - (1/6 \times 1/6) = 11/36$$

Teoremas Basicos

- Si dos eventos son mutuamente excluyentes, P(AB) será 0 y la probabilidad de ocurrencia de ambos será :

$$P(A \text{ o } B) = P(A) + P(B)$$

- Probabilidad de que al lanzar un dado obtenga 1 o 2:
- $P(1) = 1/6$; $P(2) = 1/6$

$$P(1 \text{ o } 2) = 1/6 + 1/6 = 2/6$$

Valor Esperado

- Llamamos valor esperado al valor probable que podemos obtener al repetir cierto evento.
- Va a estar asociado a la probabilidad de ocurrencia de cada opción del mismo, y al valor que tomará la variable cada caso.
- Ejemplo:
  - Probabilidad de que ganemos al apostar a un número en la ruleta es 1/37 = 0.27.
  - Premio obtenido es 35 veces la apuesta
  - Calcule la esperanza de ganar en la ruleta apostando US$1,000.

Valor Esperado

- P(Ganar) = 1/37
- P(Perder)= 1- 1/37 = 36/37
- Valor a Ganar = $35,000
- Valor a Perder = $1,000

- Esperanza de Ganancia:

E(G) = P(ganar)xValor Ganar + P(perder)*xValor Perder

E(G) = 1/37 x $35,000 + 36/37 x - $1,000

E(ganancia) = $946 - $973 = - $27

- Si jugamos a la ruleta, apostando toda la noche a un número $1,000; la esperanza que tenemos es de perder “en promedio” $27 cada vez.

Ejercicio

- Usted se Encuentra en el programa “Haga negocios conmigo”.
- Polito le presenta 3 puertas:
- Detrás de una hay un flamante ferrari rojo descapotable ultimo modelo.
- Detrás de las otras dos un pectol

Haga Negocio Conmigo

- Usted debe de escoger una puerta.
- Luego de que la ha escogido, El Eterno Perdedor abrirá de las otras dos, la que contenga un pectol.
- En este momento usted podrá escoger: mantenerse con la misma puerta inicial, o cambiar por la otra puerta.

Que escogería y porque?

Parámetros

- Mayoría de investigaciones estadísticas quieren hacer inferencias a partir de la información contenida en muestras aleatorias sobre la población de donde fueron obtenidas.
- Gralmente inferencias sobre los parámetros poblacionales (ej: media μ y varianza σ^2). Que describen a la población.
  - Se usa letras griegas.(μ,ν, λ,σ, ρ, etc.).
- Definimos parámetros como ciertas medidas que describen a la población.
- A los parámetros en general los podemos definir como θ.

Estadísticos

- Para hacer tales inferencias utilizaremos los estadísticos muestreales o estimadores de los parámetros (ej: promedio o media aritmetica $\bar{x}$ y varianza muestreal s^2)
  - Valores calculadas con base en observaciones de la muestra.
- Definimos estadístico como una medida que describe a la muestra, y que sirve para estimar los parámetros.
- A los estadísticos en general los podemos definir como θ_n.

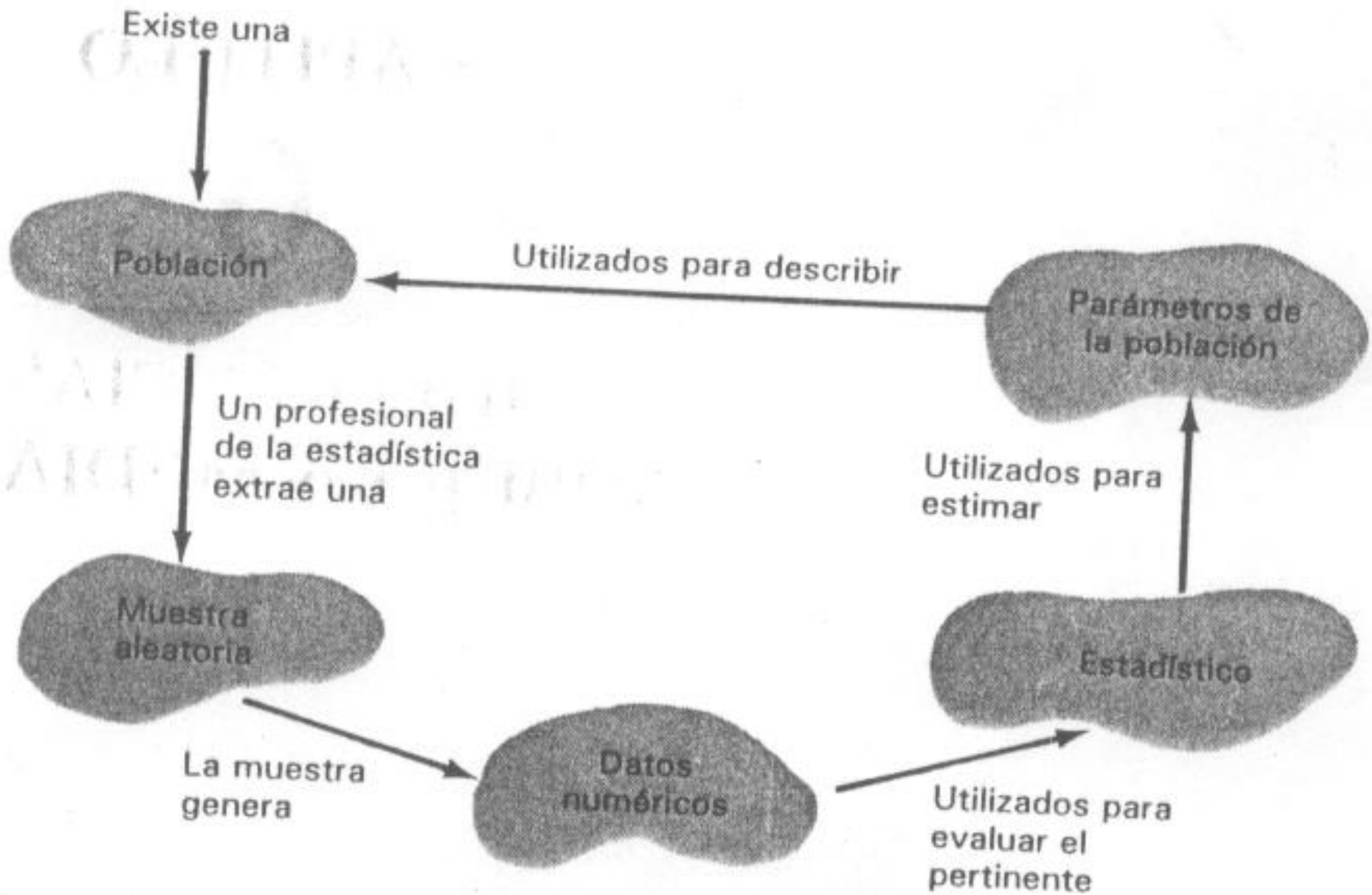


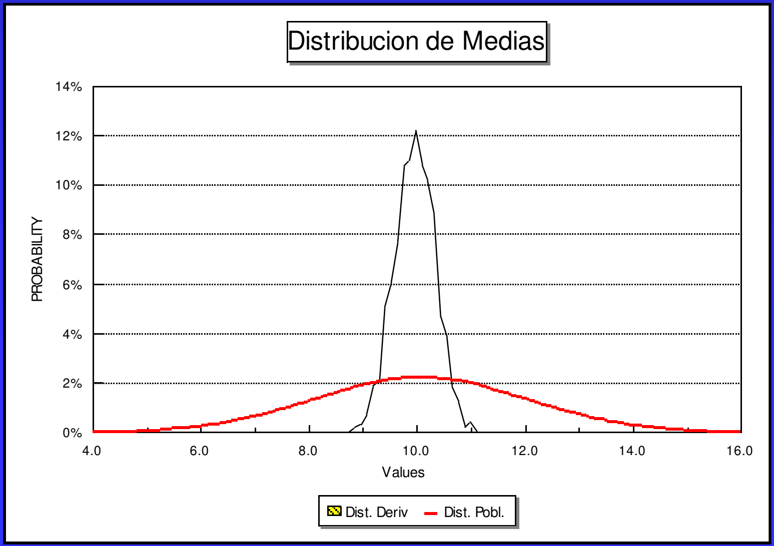
Figura 6.1
Esquema de cómo se lleva a cabo un estudio estadístico.

Estadísticos vs. Parámetros

- Importante diferencia entre estadístico y parámetro: una las bases de estadística.
- A pesar que estadísticos se usan para representar o estimar parámetros, probabilidad de que sean exactamente iguales es 0.

Ej: Promedio $\bar{x}$

- Variable aleatoria. Distribución de probabilidad (muestreo) depende mecanismo muestreo.
- Algunos valores $\bar{x}$ estarán cerca de μ, y otros alejados (para arriba o abajo).
- Al tomar varias muestras, queremos tener los $\bar{x}$ concentrados cerca a μ, y que el promedio de $\bar{x}$ esté muy cercano a μ.



Estimadores Insesgados Eficientes

- Queremos seleccionar un estimador y un plan de muestreo que:

1. Nos asegure que la esperanza de el estimador sea el parámetro ($E(\theta_0) = \theta$) Insesgado

2. La varianza del estimador tenga la menor varianza posible ($\sigma^2(\theta_0) \rightarrow$ sea baja) Eficiente

- De dos estadísticos θ_1 y θ_2, el que tenga menor varianza será el mas eficiente.

Error de Estimación

- Conociendo el estadístico θ_0 usado, y su distribución de probabilidad, podemos evaluar su error de estimación.
- "El valor absoluto de la diferencia entre el estadístico y el parámetro" (E= $|\theta_0 - \theta|$).
- No sabemos exactamente cuanto es (desconocemos parámetro θ),
- Podemos encontrar límites entre los cuales existe una probabilidad de que se encuentre el parámetro θ:
  - P($|\theta_0 - \theta|$) $\leq 1-\alpha$.

Estadísticos de Centralización

Ejercicio02a - Estadisticos.xlsx

- Media poblacional μ:
  - La media aritmética de datos de toda la población
  - Representa esperanza matemática de variable aleatoria:

$$\mu = \frac{1}{N}\sum_{1}^{N} x_i$$

- Este parámetro no lo conocemos, y no lo conoceremos nunca a no ser que muestreáramos la población completa.
- Para estimarlo usamos el estadístico promedio o media muestreal $\bar{x}$.

Estadísticos de Centralización

- Promedio o media poblacional $\bar{x}$:
  - La media aritmética de los datos de la muestra

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

- Al ser μ la esperanza matemática de los $\bar{x}$, esta puede calcularse también de la siguiente forma:

$$\mu = \frac{n_j}{N}\sum_{1}^{k} \bar{x}_j$$

  - j es j-esimo grupo de un total de k grupos
  - n_j es el número de individuos en el j-esimo grupo
  - $\bar{x}_j$ es la media del j-esimo grupo

Estadísticos de Centralización

- Promedio ponderado $\hat{x}$

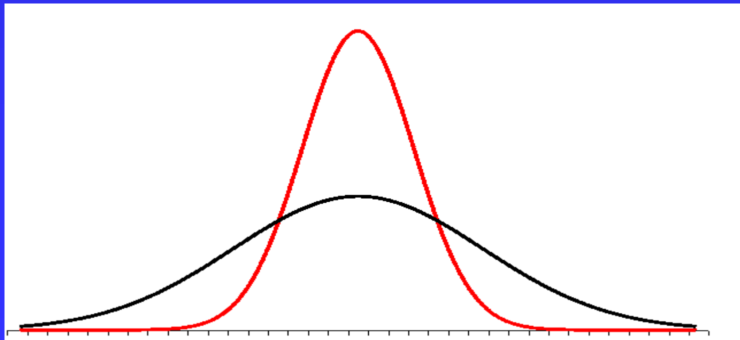
$$\hat{x} = \frac{n_j}{n} \sum_{1}^{k} \overline{x}_j$$

Estadísticos de Centralización

- Moda: Marca de clase del intervalo con mayor frecuencia
  - Aproximadamente: Valor que mas encontramos en nuestro muestreo.
- Mediana: valor más cercano a la mitad si los ordenamos, o valor con igual número de datos mayores que menores a él.
  - Valor del dato número (n+1)/2 cuando n es impar
  - Media del dato # (n/2) y el dato # (n/2 +1) cuando n es par.

Estadísticos de Dispersión

- Medidas de centralización dan una idea de hacia dónde están distribuidos nuestros datos, pero no de cómo están distribuidos.
- Probabilidad de dato igual a la media tiende a 0
- Media de posibles valores un dado 3.5
- Cruce de Rio; Pies en horno, cabeza refrigerador
- Dos poblaciones con igual media pero dispersión de datos distinta: Poblaciones distintas

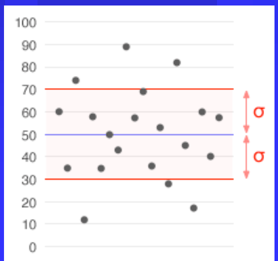


Estadísticos de Dispersión

- Parámetro varianza poblacional σ^2 :
  - Promedio de cuadrados de las desviaciones de los valores de una variable en población con respecto a media poblacional

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$$

  - x_i-μ es distancia de cada punto a la media
  - Se eleva al cuadrado porque si no distancias positivas y negativas se anularían dando 0



Varianza

- Fisher (1918) "The Correlation Between Relatives on the Supposition of Mendelian Inheritance"
  - El gran cuerpo de las estadísticas disponibles nos muestran que las desviaciones de una medida humana de su media siguen muy de cerca la ley normal de los errores, y, por tanto, que la variabilidad puede ser medida de manera uniforme por la desviación estándar correspondiente a la raíz cuadrada de la media del cuadrado del error.
  - Cuando hay dos causas de variabilidad independientes, capaces de producir en una distribución poblacional de otra manera uniforme, con desviaciones estándar θ1 y θ2, se encuentra que la distribución, cuando ambas causas actúan juntas, tiene una desviación estándar $\sqrt{\theta_1^2 + \theta_2^2}$
  - Por tanto, es conveniente en el análisis de las causas de la variabilidad, trabajar con el cuadrado de la desviación estándar como la medida de la variabilidad. Vamos a llamar esta cantidad, la varianza

Propiedades de la Varianza (1)

1. Es positiva (^{2})
2. Es en distintas unidades que la variable (^{2})
3. No varía por localización. Sumar constante a todos los datos: misma varianza. Var(x + a) = Var(x)
4. Si se multiplica todos los datos por una constante, varianza se multiplica por constante^2 Var(ax) = a^{2}Var(x)
5. La varianza de la suma de variables aleatorias es igual a la suma de sus varianzas + 2 veces su covarianza.

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X,Y),$$

6. Generalizando para N Variables

$$Y = \sum_{i=1}^{N} X_i$$

$$\mathrm{Var}(Y) = \sum_{i=1}^{N} \mathrm{Var}(X_i) + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j).$$

Propiedades de la Varianza (2)

7. Varianza = Promedio de cuadrados – el cuadrado del promedio

$$Var(X) = 1/N\ \Sigma x_i^2 - \bar{x}^2$$

8. La varianza de la suma de variables aleatorias independientes es igual a la suma de sus varianzas:

$$Var(X + Y) = Var(X) + Var(Y)$$

generalizando:

$$\operatorname{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \operatorname{Var}(X_i).$$

9. Si las variables independientes tienen la misma varianza, la varianza de su promedio puede transformarse multiplicando por $(1/n)^2$ (4).

$$\operatorname{Var}(\overline{X}) = \operatorname{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \operatorname{Var}(X_i) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}.$$

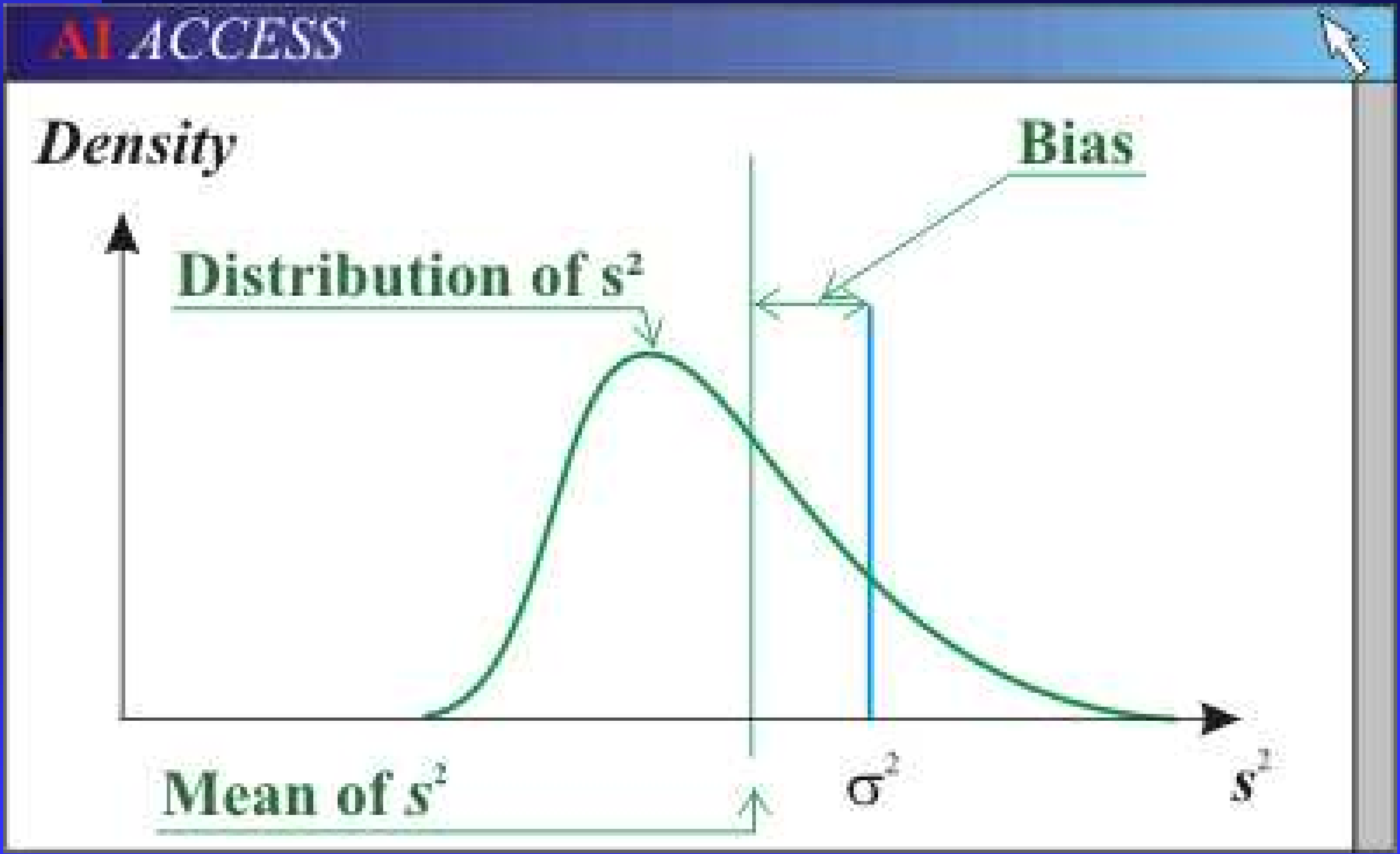
*Recordar este σ^2/n para teorema central del limite

Estadísticos de Dispersión

- Varianza empírica s^2 es el estadístico mediante el cual hacemos estimaciones de nuestro parámetro varianza poblacional.
- Ya que s^2 sería estimador sesgado de σ^2 si la dividimos para n, se la divide por n-1:

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

- A medida que tamaño de la muestra (n) aumenta, sesgo entre σ^2 y s^2 disminuye


AI ACCESS
Density
Bias
Distribution of s²
Mean of s²
σ²
s²

Estadísticos de Dispersión

- La desviación típica o desviación estándar (σ o s), es la raíz cuadrada positiva de la varianza. s es estimador sesgado de σ.
- El rango es la diferencia entre el valor del mayor dato y el valor del menor dato.
- Desviación media: promedio de las desviaciones absolutas respecto al promedio: $DM = \Sigma|xi - \bar{x}|/n$
- Error típico de la media: estima s para la distribución de $\bar{x}$: $S_{\bar{x}} = s / \sqrt{n}$
- Coeficiente de variación: expresión porcentual de variación (sin unidades): $CV = s \times 100 / \bar{x}$

Est. Disp. Usan 1 decimal más que la muestra

Introduccion al Excel Como Herramienta Estadistica

- Versatilidad:
  - Hoja de calculo
  - Base de datos
  - Diagramador
  - Lenguaje de programación
  - Análisis de datos
- Modelo de objeto:
  - Aplication
  - Workbook
  - Worksheet
  - Range
  - Otros objetos:
- Otros Objetos:
  - Row
  - Column
  - Cell
  - Area
  - Rangos con Nombre

Interfaz de Usuario?

- Ventanas
- Menu de Excel 2003 y anteriores
- Cinta de opciones Excel 2007
- Barra(s) de Herramientas
- Barra de formulas
- Cuadro de nombres
- Barra de estado
- Macros
- Complementos
- Archivos personales
- Entrada de datos e interfaz con el usuario

Tipos de Datos/Objetos?

- Texto
- Números
- Formatos
- Fórmulas
  - Referencias absolutas y relativas
- Funciones
- Matrices
- Referencias Remotas
- Comentarios
- Gráficos
- Tablas y gráficos dinámicos
- Otros objetos

Operaciones Básicas?

- Desplazamiento
- Teclas de acceso rápido
- Funciones Mouse
- Selección
- Direcciones relativas y absolutas
- Nombres de rango
- Copiar, Cortar. Pegar , Pegado Especial.
- Asistentes
- Personalización

Funciones Estadisticas

- Muchas
- Muy Utiles
- Algunas no se para que son o no las he usado
- Aplasten F1 y lean de que se tratan
- Pruebenlas y comparenlas con calculos “manuales”.
- Revisaremos las mas frecuentes.
- Existen rutas alternas en Excel

Herramientas de Analisis de Datos

- Complemento de Excel
- Existen otros complementos estadisticos de terceros
- Tienen sus ventajas y sus limitaciones
- Existen rutas alternas en Excel

Herramientas No Estadisticas Utiles Para la Estadistica

- Modelo de Hoja de Calculo
- Formulas
- Funciones no Estadisticas
- Ordenar, filtrar.
- Graficos
- Tablas Dinamicas y Graficos
- ODBC, conecciones y otros datos externos
- Macros

Ejercicio Practico

- Calcular en Ejercicio02b - Estadisticos.xlsx:
  - Suma
  - n
  - $\bar{x}$
  - Moda
  - Mediana
  - s^2 y s
  - Maximo, mínimo y rango
  - Error típico, coeficiente de variación
  - Analisis de Datos / Estadistica Descriptiva

Usar Formula y Función