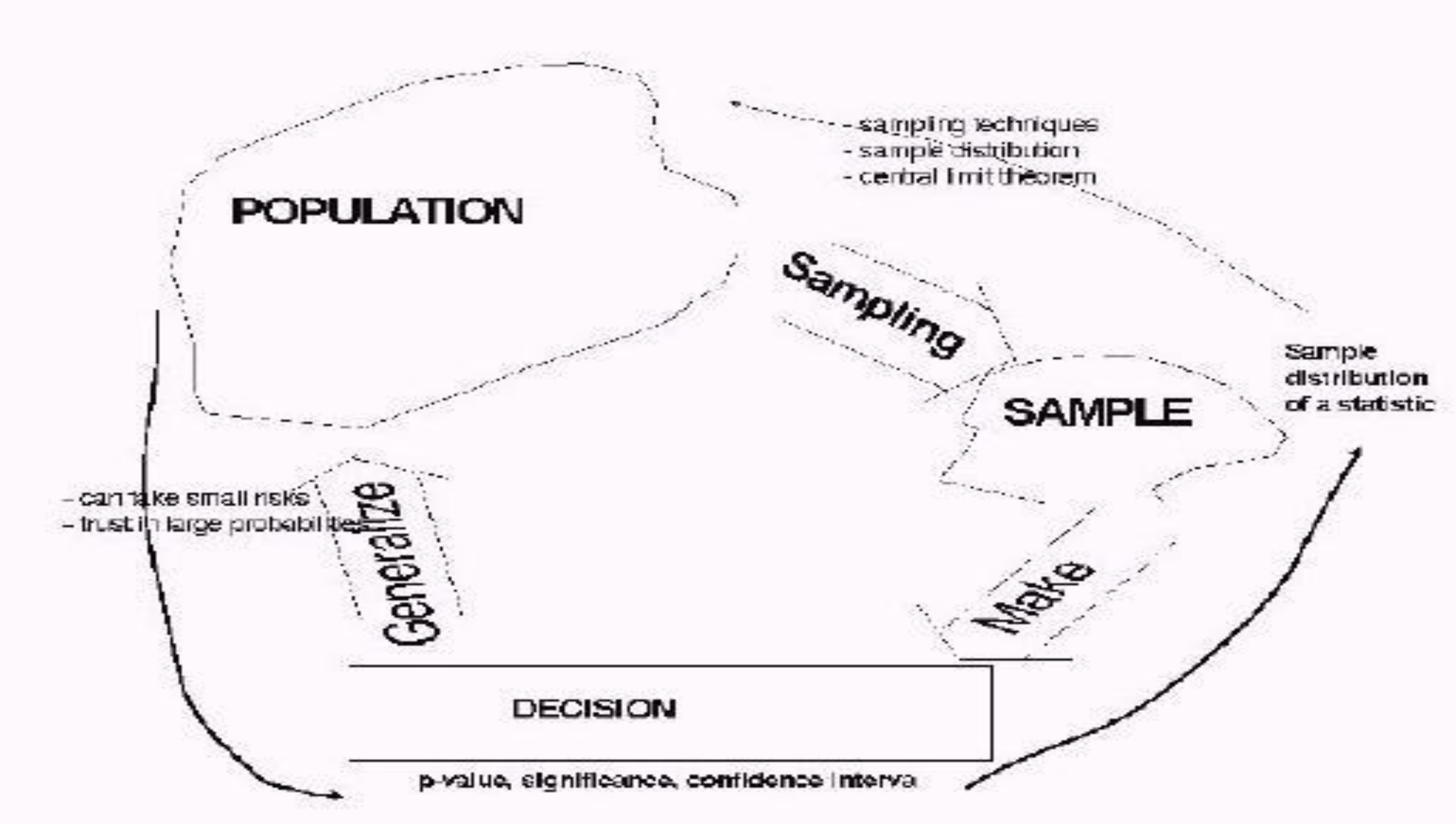
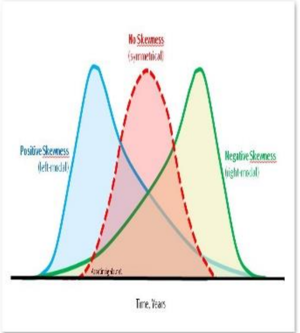
Estimación



Osca Centeno Mora

Preámbulo

- A partir del presente tema iniciamos el componente de la Inferencia Estadística.
- La Inferencia Estadística es de suma importancia en el estudio del análisis de la información: siempre vamos a querer inferir cierto tema de una población "X", o hacer cierta prueba para cierta población.
- Desde un punto de vista más práctico, la Inferencia Estadística nos permitirá responder ciertas incógnitas que poseemos, sin tener que pasar por el estudio de toda una población.



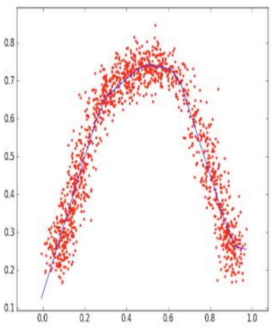
- Veremos que todas las técnicas que utilizaremos requieren del uso de las probabilidades, en especial de la suposición de que los datos se distribuyen según una determinada función de probabilidad.

Preámbulo

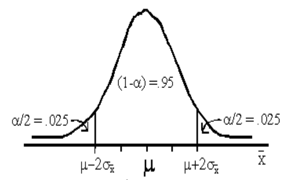
Un caso de la aplicación de la estimación en el campo de salud es la prueba del hemograma completo.

Preámbulo

- Aunque el presente curso cubre únicamente los temas de estimación y prueba de hipótesis (estos temas son la base): el análisis de la asociación, comparaciones de medias, tablas de contingencia y otros, se apoyan en estos elementos.



- La estimación y la prueba de hipótesis, son dos formas o pensamientos distintos de hacer Inferencia Estadístico: el primero se fundamento en un post pensamiento, mientras que el segundo en un pre pensamiento.



- Veamos primero la estimación dentro de la Inferencia Estadística, enfocándonos en el promedio ($\bar{x}$) como medida para datos cuantitativos, y en la proporción (p) como medida para datos cualitativos.

Índice

Índice

Introducción

- En todo estudio o investigación existe un campo de referencia, universo o **población** al cual se desea generalizar los resultados o conclusiones que se obtengan.

- Por **población** o universo se entiende el total de personas, objetos o mediciones que tienen una característica común.

- "El total de todas las observaciones correspondientes a una cierta característica". Gómez. M.

Introducción

- Las investigaciones y experimentos se llevan a cabo con el propósito de llegar a conclusiones o leyes de carácter general, es decir, buscando resultados que se puedan generalizar a todos los elementos de la población.

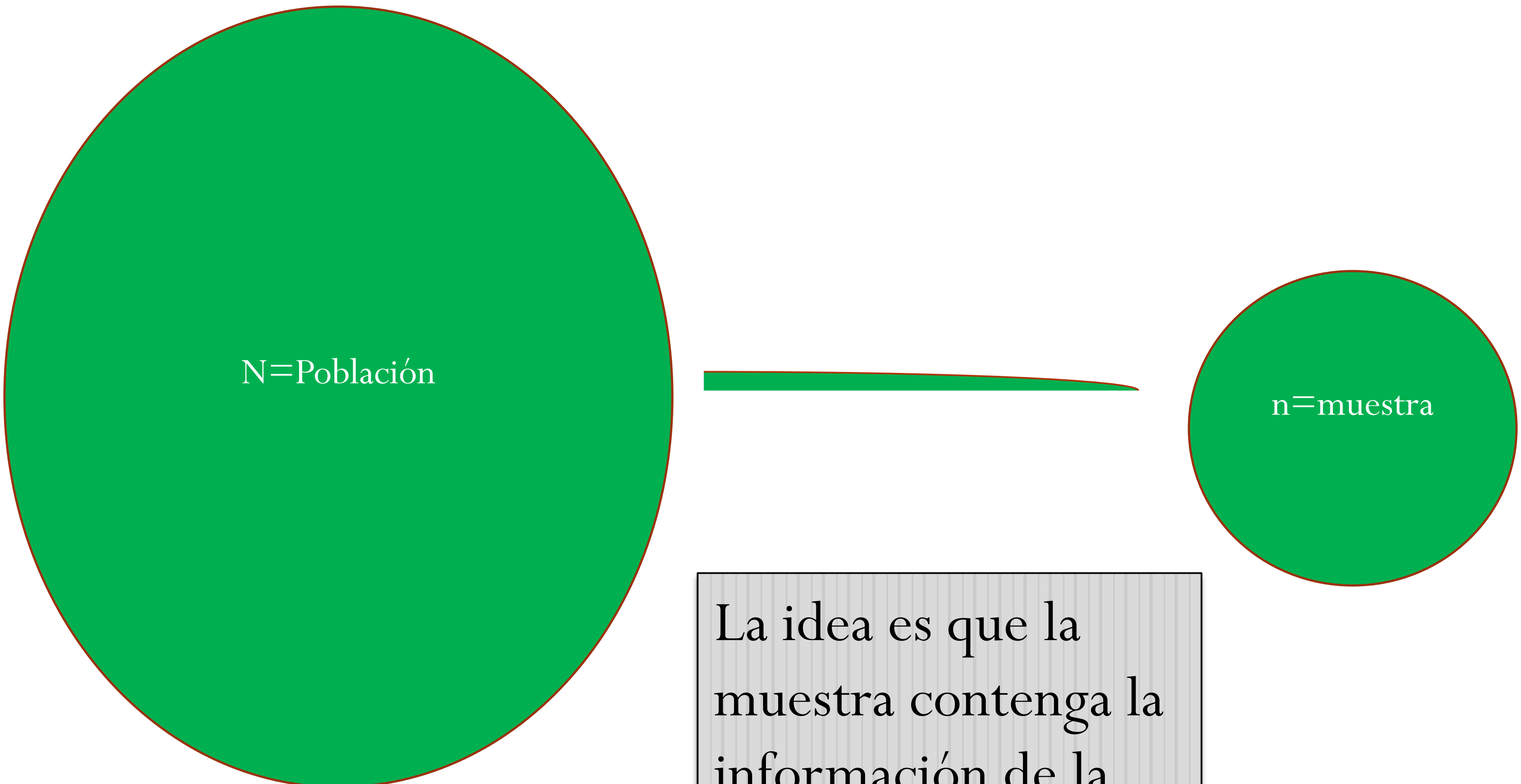
- El sentido común indica que lo mejor sería "censar" a todos los elementos que constituyen la población.

- Sin embargo, esto resulta muchas veces imposible por cuestiones de administración, tiempo, grandeza y dinero (recursos).

- Como no se puede "censar" toda la población, se toma una muestra y los resultados de ella se generalizan a toda la población.

Introducción

- Generalización de la información:



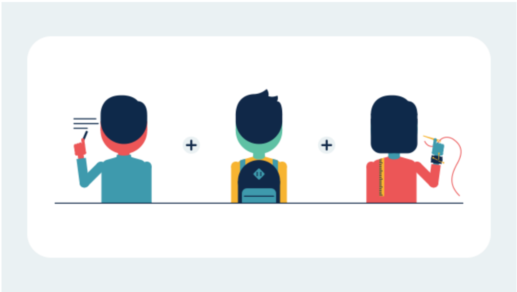
La idea es que la muestra contenga la información de la población

Introducción

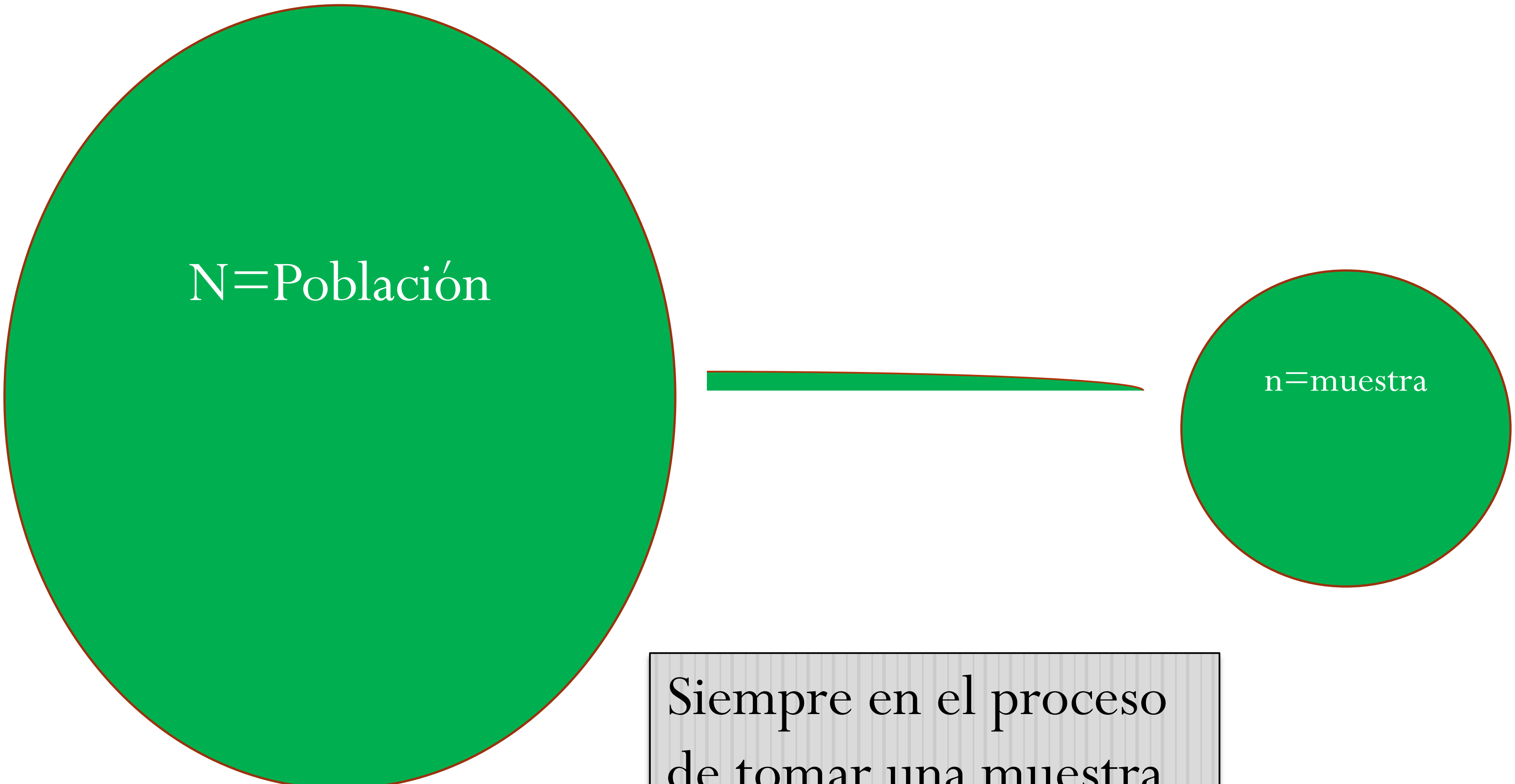
- La razón más importante de la muestra es la no necesidad de estudiar toda la población.

- Los resultados que arrojaría una muestra bien seleccionada de tamaño razonable, serán suficientemente precisos para alcanzar los fines prácticos: generalizar a la población.

- ***Inferencia estadística:*** "es el proceso mediante el cual se generalizan los resultados observados en una muestra aleatoria, a la población de la cual se extrajo (población de interés)." Gomez, M.



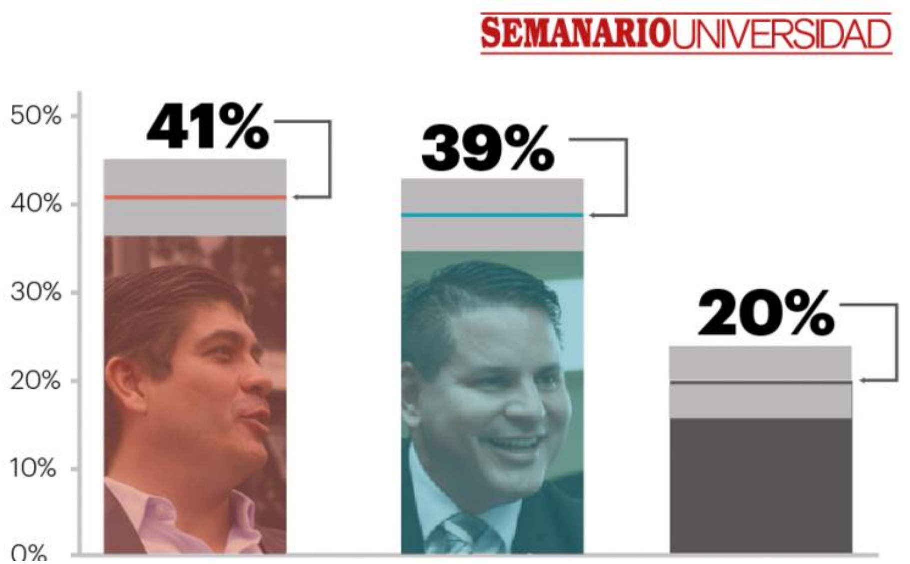
Introducción



Siempre en el proceso de tomar una muestra se hace:

<u>Inferencia Estadística</u>

Introducción



Índice

1 Introducción

2 Características de la muestra

Características de la muestra

- Para tener resultados válidos de generalización, se dice que la muestra debe ser aleatoria o probabilística.

- Esto es, los elementos que constituyen la muestran fueron seleccionados utilizando mecanismos que da a cada elemento de la población una probabilidad conocida de ser incluida en la muestra, y estos fueron seleccionadas al azar.

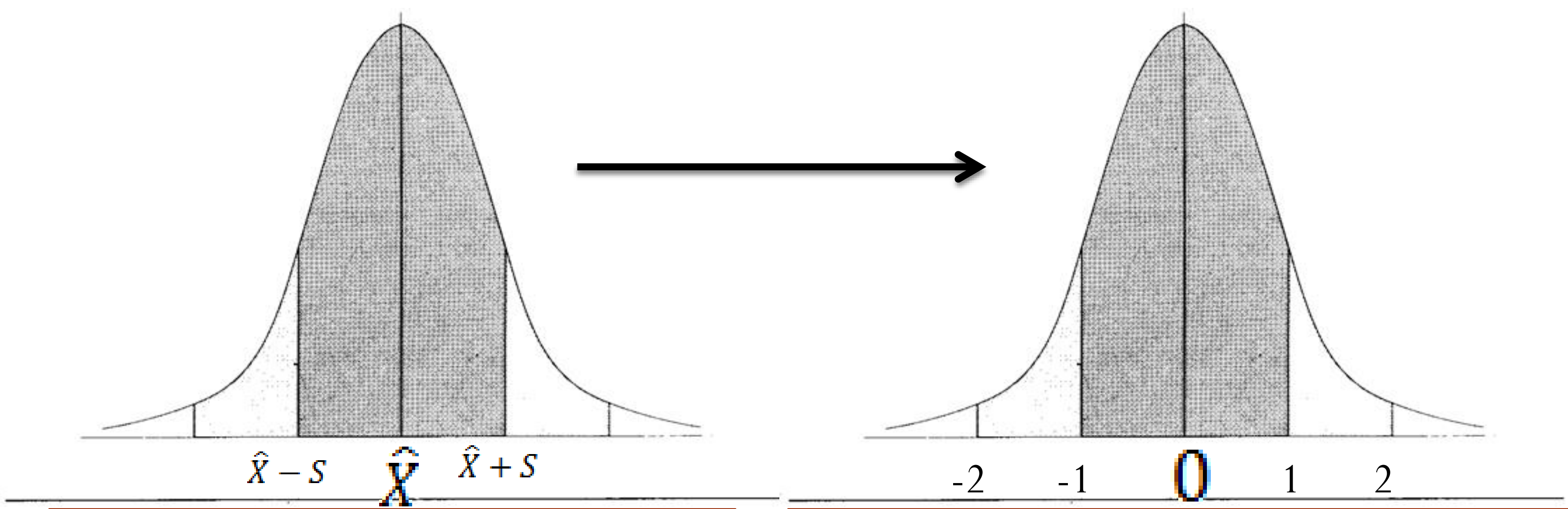
- Al ser la muestra aleatoria, esto permite que esta posea una distribución probabilística y es posible aplicar la teoría de las probabilidades para evaluar la confiabilidad de las inferencias.

Aplicación de la curva normal estándar
N(0,1)

Características de la muestra

n=muestra

Al ser una muestra aleatoria, se puede aplicar la teoría de probabilidad para obtener inferencias confiables



Información de la muestra

Aplicación de las probabilidades

El error asociado a la muestra

- Cada muestra posee un error asociado a la generalización (predicho probabilísticamente).

- Esto es, como no se toma en cuenta a todos los individuos de la población, entonces hay una cierta diferencia entre el valor verdadero que si quiere (el de la población), y el que se obtuvo con la muestra.

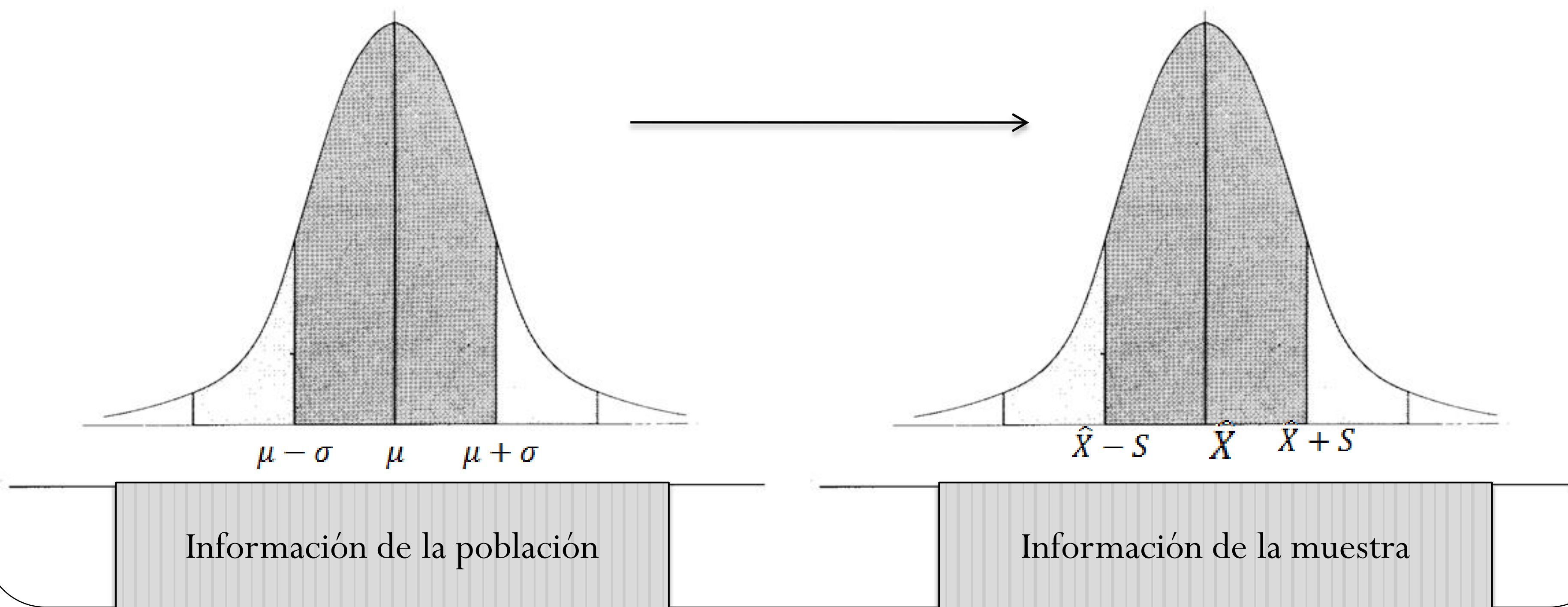
- Mediante la teoría de probabilidades, si y sólo si se tomó una muestra aleatoria, es posible, es posible medir ese error entre el valor real de la población y el de la muestra.

- A este error se le conoce como **error de muestreo**.

El error asociado a la muestra

La diferencia entre el valor real de la población población y el de la muestra se denominada error de muestreo.

Para el caso del promedio: $\mu - \bar{x} = \text{error de muestreo}$



Muestra no aleatorio o al azar

- Si la muestra es seleccionada en forma intencional o de juicio, los errores no tendrán un comportamiento aleatorio, sino que dependerán de los prejuicios y de tendencias conscientes o inconscientes del que hizo la selección.
- En estas circunstancias no es posible aplicar la teoría de las probabilidades para evaluar las inferencias o generalizaciones.
- Al no tener una forma objetiva de evaluar los errores de muestreo, la inferencia estadística no puede ser aplicada a las muestras no aleatorias.
- Por lo tanto, por esta y otras razones, se debe aplicar una muestra aleatoria.

Muestra no aleatorio o al azar

- La preferencia por las muestras aleatorias para la realización de inferencias obedece a tres razones básicas:

1. Eliminan los sesgos o errores de selección.
2. Producen errores aleatorios de muestreo que son medibles utilizando modelos probabilísticos.
3. El error de muestreo puede hacerse tan pequeño como se quiera, aumentando el tamaño de la muestra.

Índice

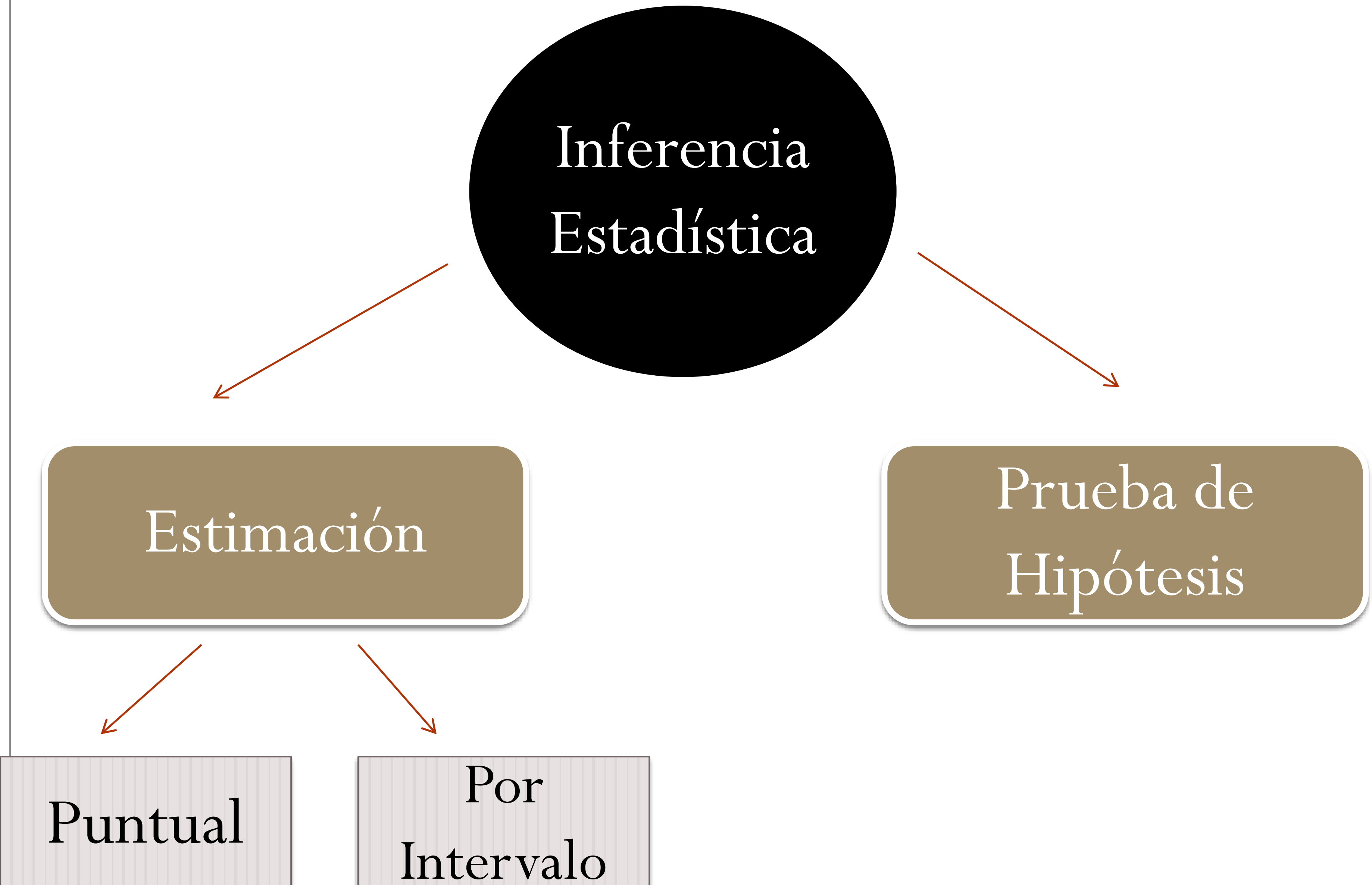
Inferencia estadística

- Dentro de la inferencia estadística se distinguen dos tipos de problemas:

1. Uno en que interesa ***estimar (estimación)*** características o propiedades de la población –promedio, variancia- de los elementos de ciertas características.
2. El otro es cuando interesa someter a ***prueba de hipótesis*** cierta característica de la población.

- Mientras que en la estimación interesa saber el rango de valores que podría tener cierto parámetro, en la prueba de hipótesis interesa si un determinado parámetro podría tener un cierto valor.

Inferencia estadística



Parámetros y Estimadores

- Las medidas que representan características o propiedades de la población, y que fueron obtenidas para toda la población, se denominan “valores poblacionales” o “Parámetros”. Algunos a citar son:

Parámetro	Definición
μ	Promedio de la población
σ^2	Variancia de la población
σ	Desviación estándar de la población
Me	Mediana de la población
P	Proporción de la población
ρ	Coeficiente de población

Parámetros y Estimadores

- Generalmente, los valores poblacionales no son conocidos y una de las tareas básicas de la inferencia estadística es su ***estimación***.

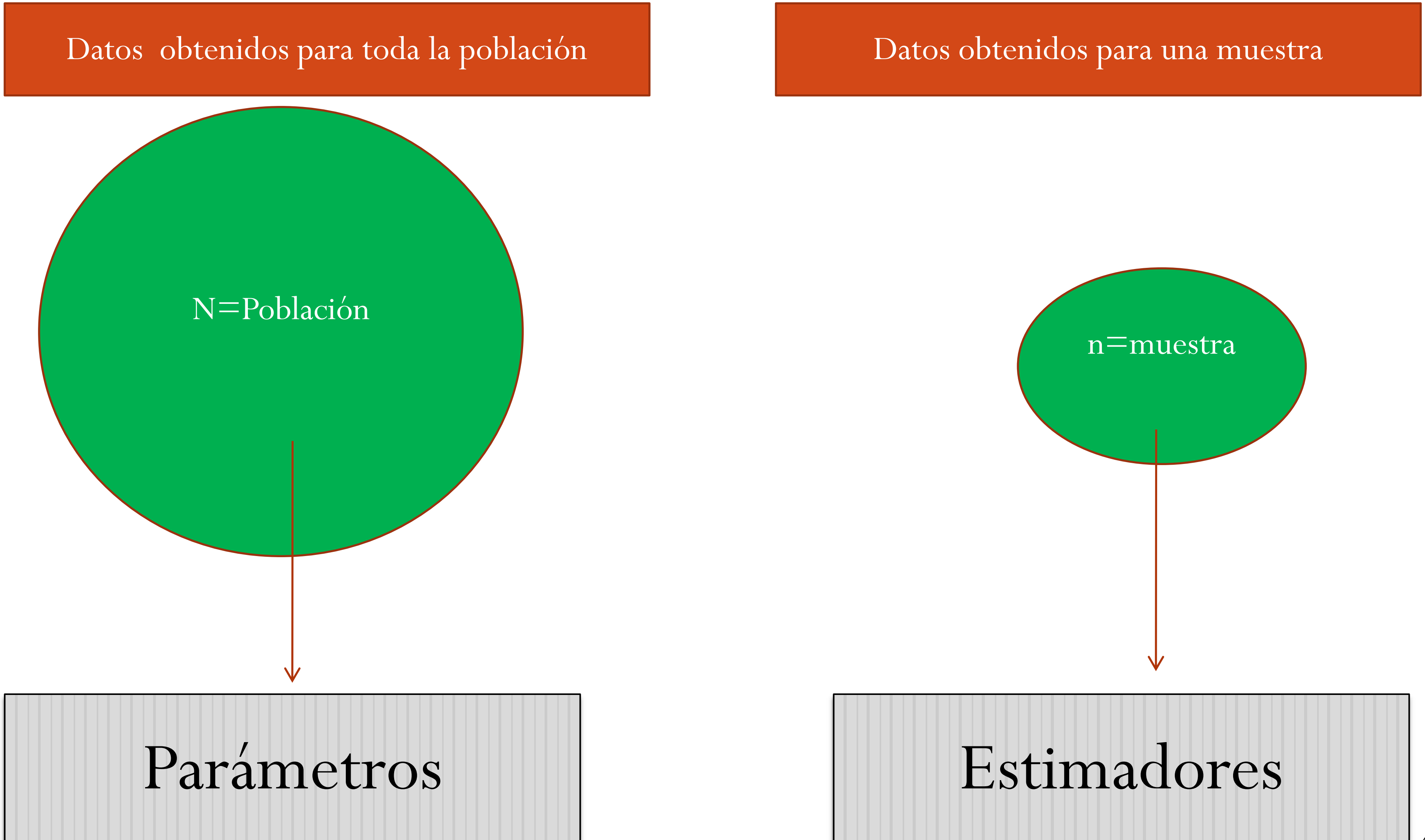
- La estimación es el cálculo a partir de los datos de la muestra, de un valor numérico que será considerado, para efectos prácticos, como el valor de la población.

- Entonces, ***un estimador*** es una función de los valores de la muestra que se utiliza para aproximar o estimar el valor poblacional.

- La ***estimación*** es el valor numérico producido por una muestra específica (es el proceso de cálculo con un respectivo valor).

Parámetros y Estimadores

- Resumiendo:



Parámetros y Estimadores

- Los estimadores más famosos son los del promedio (x barra), la variancia (S^2) y el de la desviación estándar (S). A continuación se presentan algunos parámetros y estimadores

ESTADÍSTICO	PARÁMETRO	ESTIMADOR
Media aritmética	μ	$\bar{x}$
Desviación estándar	σ_x	s
Variancia	σ^2	s^2
Proporción	P	$\hat{p}$
Coeficiente de correlación	ρ	r

Índice

Estimación puntual del parámetro μ

- Para estimar el valor del promedio de cierta población (μ) se podría utilizar el estimador $\bar{x}$. Esto sería una estimación puntual de μ.

- Ejemplo: se quiere estimar el promedio de edad de todos las personas de la Universidad Z (μ). Se selecciona una única muestra para estimar el promedio ($\bar{x}$), y el resultado fue de 25.7 años.

$$\begin{tikzcd}
\mu \arrow[r] \arrow[d] & \boxed{?????} \arrow[d] \\
\bar{x} \arrow[r] & \boxed{25.7}
\end{tikzcd}$$

Estimación puntual del parámetro μ

- Sin embargo, antes se había discutido que, al no tomar toda la información de la población, se podría producir una diferencia importante entre el valor real (μ), y el del estimador ($\bar{x}$).

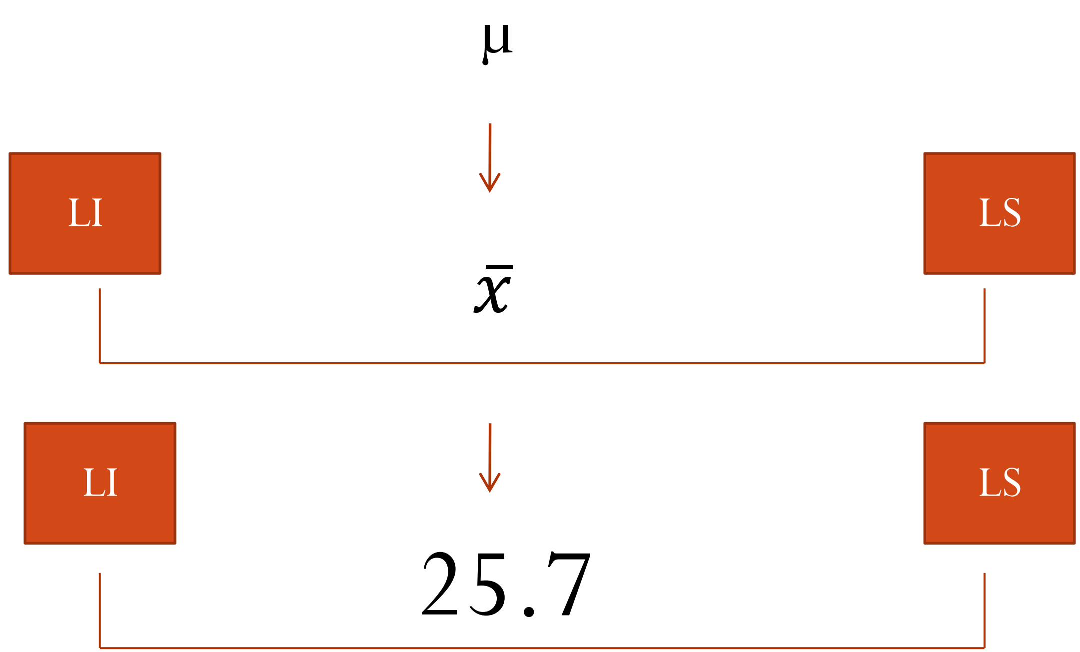
Error de muestreo: $\mu - X = E.M.$

- De ahí la importancia de contar con una idea acerca del error asociado con la estimación que se ha realizado.

- La estimación puntual es insuficiente si no se posee alguna medida de confianza.

Estimación puntual del parámetro μ

- El problema se resuelve utilizando la ***estimación por intervalos***, o sea construyendo, a partir de la estimación puntual, un intervalo dentro de cuyos límites, se espera, con una determinada grado de confianza, que esté contenido el verdadero valor poblacional.



Estimación puntual del parámetro μ

- Para nuestro presente caso, la estimación del promedio (25.7 años), podría estar muy alejada del valor real de la población de interés (μ), o poco alejada, no se sabe.

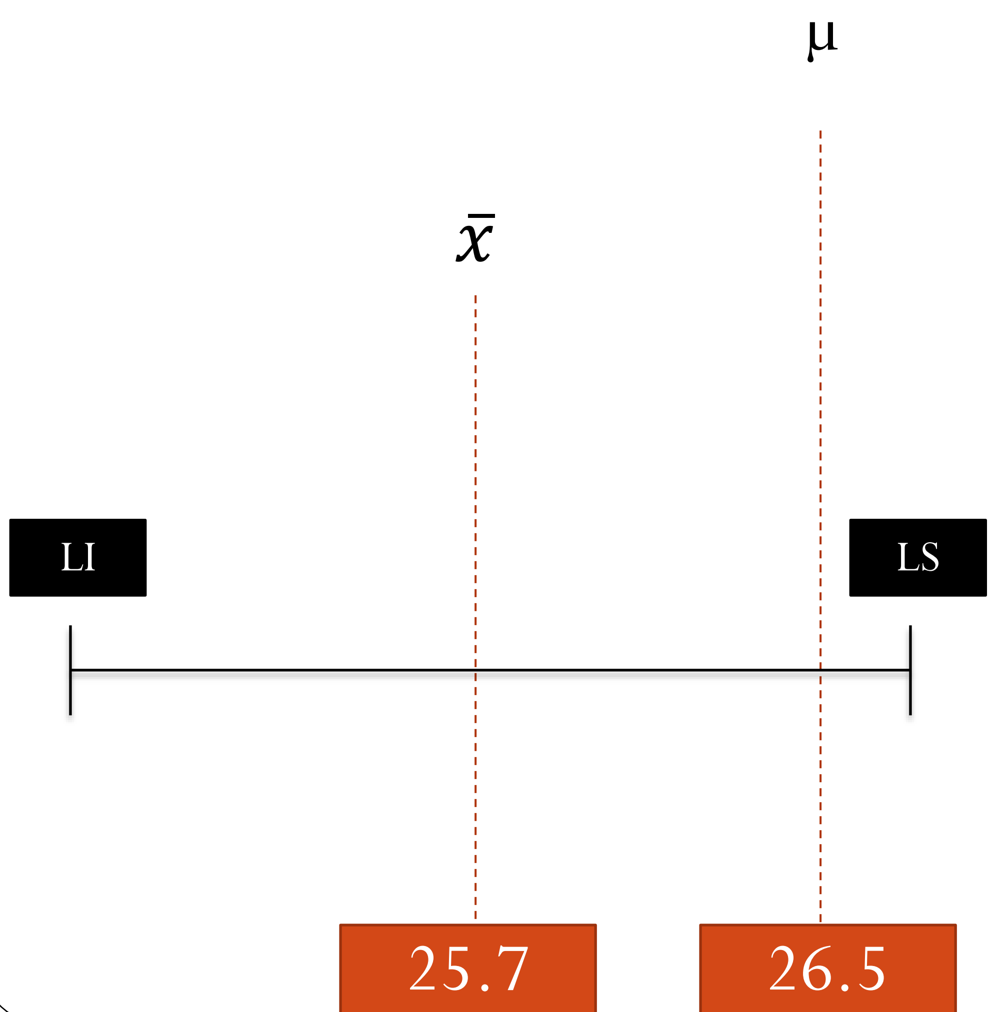
- Para superar esta dificultad, esta incertidumbre, puede usarse la estimación por intervalos.

- Al decir que el valor de 25.7 años puede estar entre dos valor (llamados límites inferior y superior), se tendría una mayor confianza de reportar el resultado.

- De ahí la gran importancia de la estimación por intervalos.

Estimación puntual del parámetro μ

- La utilidad de los intervalos de confianza:



Aunque el estimador no es igual que el parámetro, con el cálculo de los intervalos de confianza, tenemos que el intervalo está conteniendo al valor poblacional μ

Índice

Índice

1. Introducción
2. Características de la muestra
3. Inferencia estadística
4. Estimación puntual y por intervalos
5. Límites de confianza
   - Variable cuantitativa
   - Variable cuantitativa

IMPORTANTE 1

IMPORTANTE 1

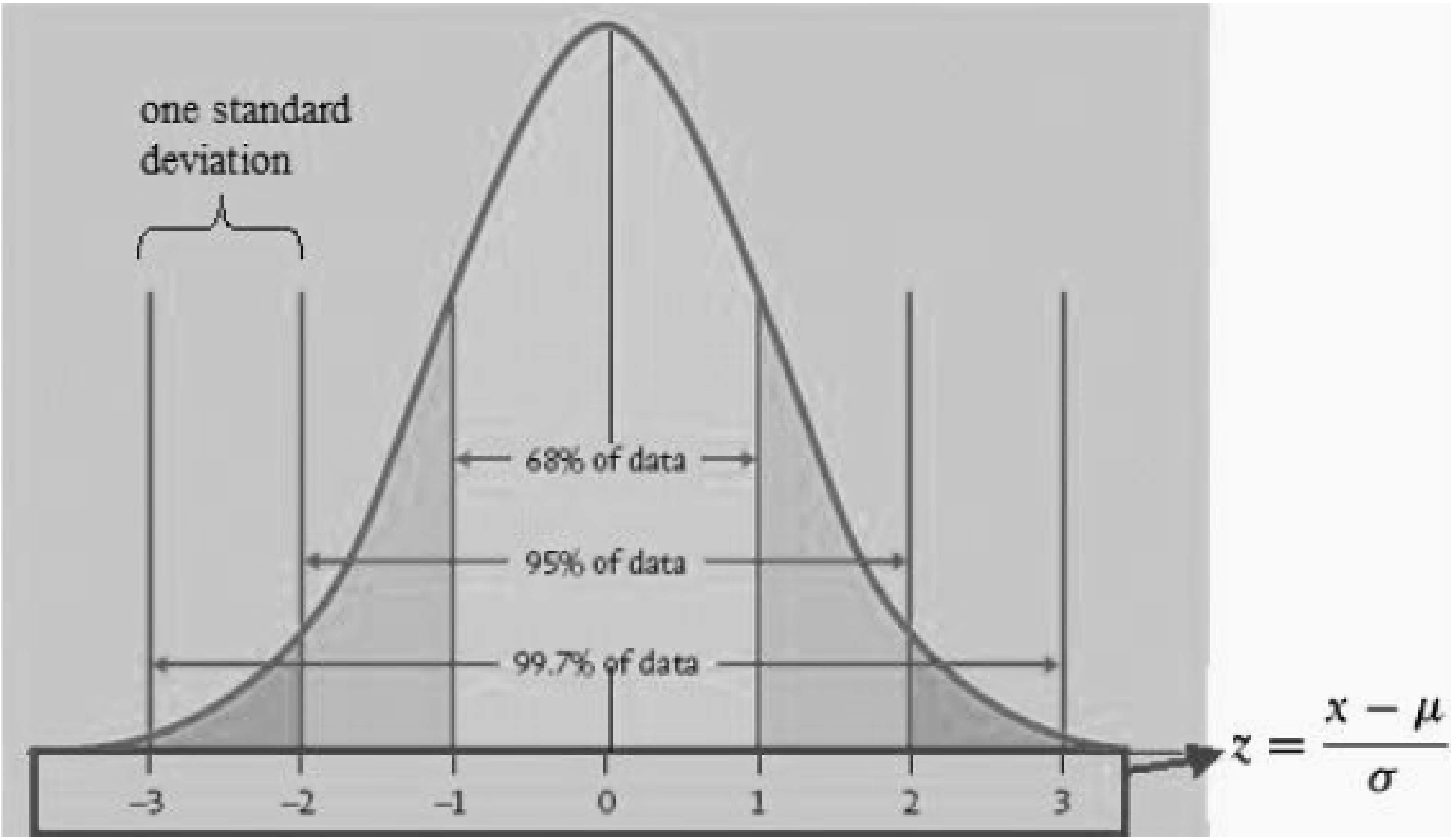
- En el análisis de la Inferencia Estadística, debemos suponer que nuestras población, bajo la característica que estamos analizando, se comportan o se le asocia un tipo de función de distribución.

- Por ejemplo, podríamos suponer que las consultas poseen una distribución normal, o que los salarios promedio de los doctores poseen también una distribución normal.

- Este punto es muy importante, dado que los valores del nivel de confianza corresponde a la función de distribución que estamos suponiendo para la característica de análisis.

- Por lo tanto, lo primero es saber la distribución de los datos, aunque para el presente curso siempre se distribuirán de forma normal.

IMPORTANTE 1



IMPORTANTE

- En la obtención de los límites de confianza, estos se obtienen a partir de cálculo de los intervalos de confianza. Un Intervalo de Confianza (IC) siempre tendrá la siguiente forma:

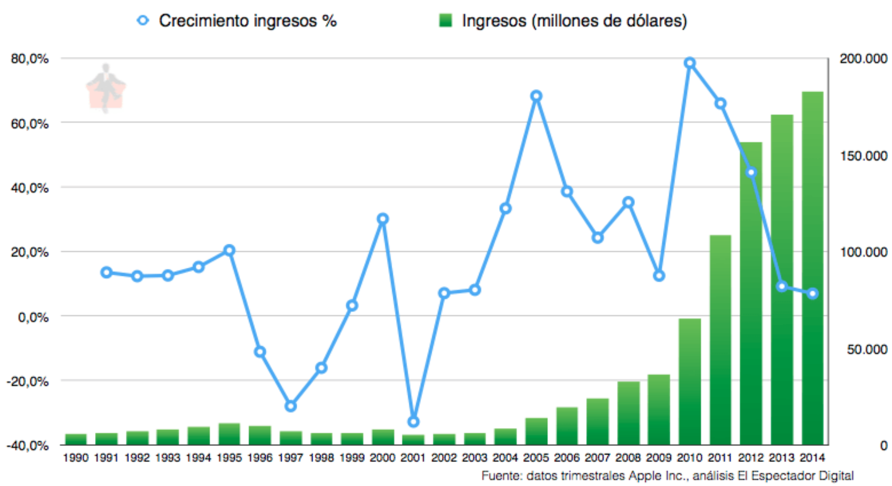
$$IC = estimador \pm Nivel\ de\ confianza * error\ de\ muestreo$$

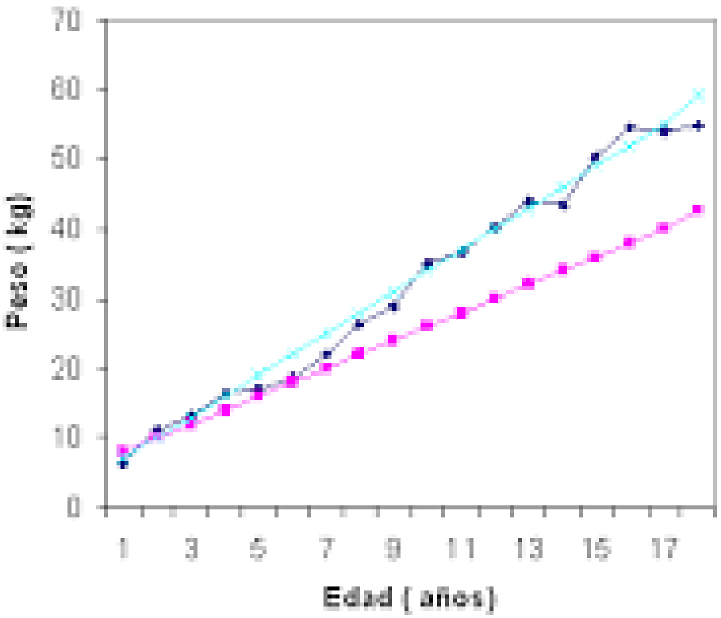
$IC =$ [Límite Interior ; Límite Superior]

- Nótese que el intervalo de confianza se constituye de 3 partes, y que estas requieren de las medidas de posición y variabilidad (Estadística Descriptiva),y de las funciones de probabilidad (Probabilidades).

- Aunque veremos la modalidad para variables cuantitativas refiriéndonos al promedio ($\bar{x}$), también estudiaremos el caso de las variables cualitativas para la proporción (p) .

Límites para variables cuantitativas





¿Qué tipo de análisis realizamos con variables cuantitativas?

Límites de confianza para μ

- Supóngase que la distribución de la población es una curva normal con media μ y variancia σ^2 (desviación estándar σ), y de esta se tomó una muestra de tamaño n.

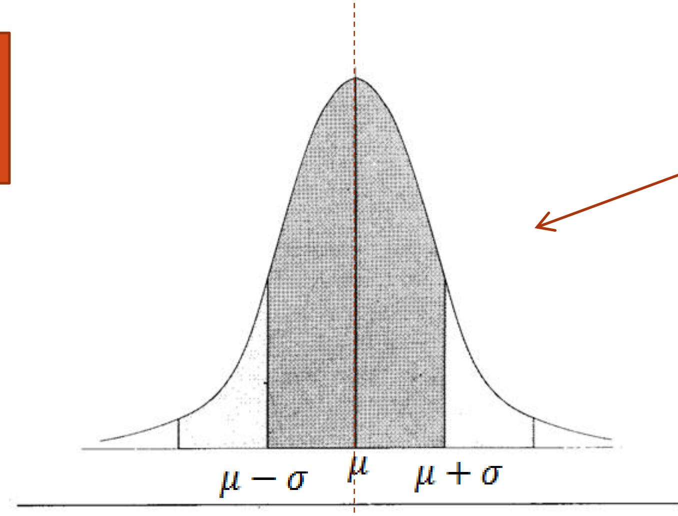
- Ahora, se desea saber la posible distribución alrededor de la media muestral ($\bar{x}$) que se obtuvo.

- Por teoría sabemos entonces que la distribución muestral tendrá una media de $\bar{x}$ y una variancia de σ^2/n (desviación estándar $\sigma/\sqrt{n}$).

- Si la curva poblacional tiene una distribución $N(\mu,\sigma)$, la muestral tiene una curva de $N(\bar{x},\sigma/\sqrt{n})$

Límites de confianza para μ

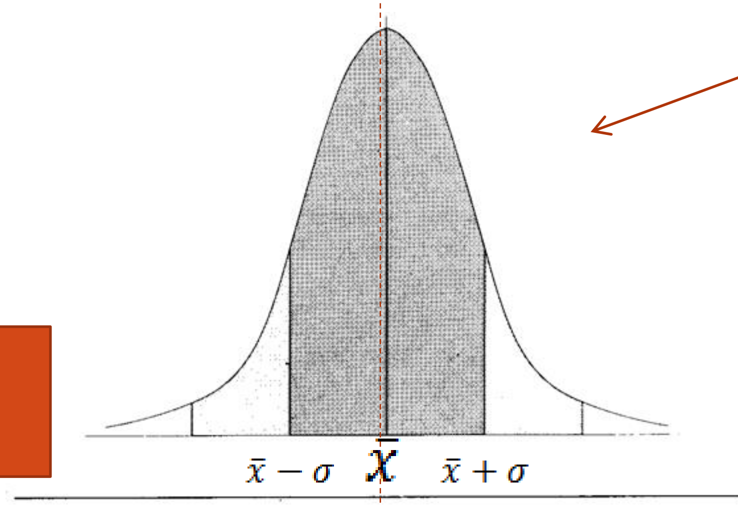
Distribución de la población



$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Se extrae muestra de tamaña “n”

$$\frac{1}{\frac{\sigma}{\sqrt{n}}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\bar{x}}{\sigma/\sqrt{n}}\right)^2}$$



Distribución de la muestra

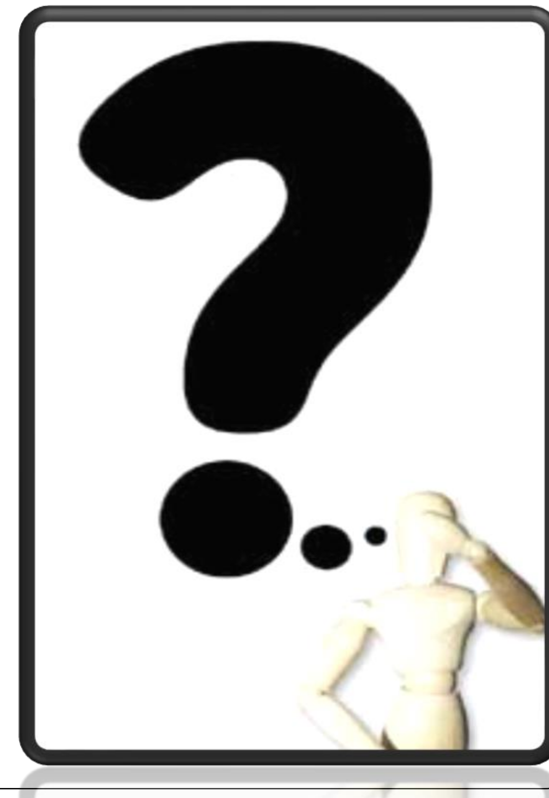
Límites de confianza para μ

Estadísticos	Población	Muestra
Posición central (promedio)	μ	$\bar{x}$
Variabilidad (desviación estándar)	σ	$\frac{\sigma}{\sqrt{n}}$

Límites de confianza para μ

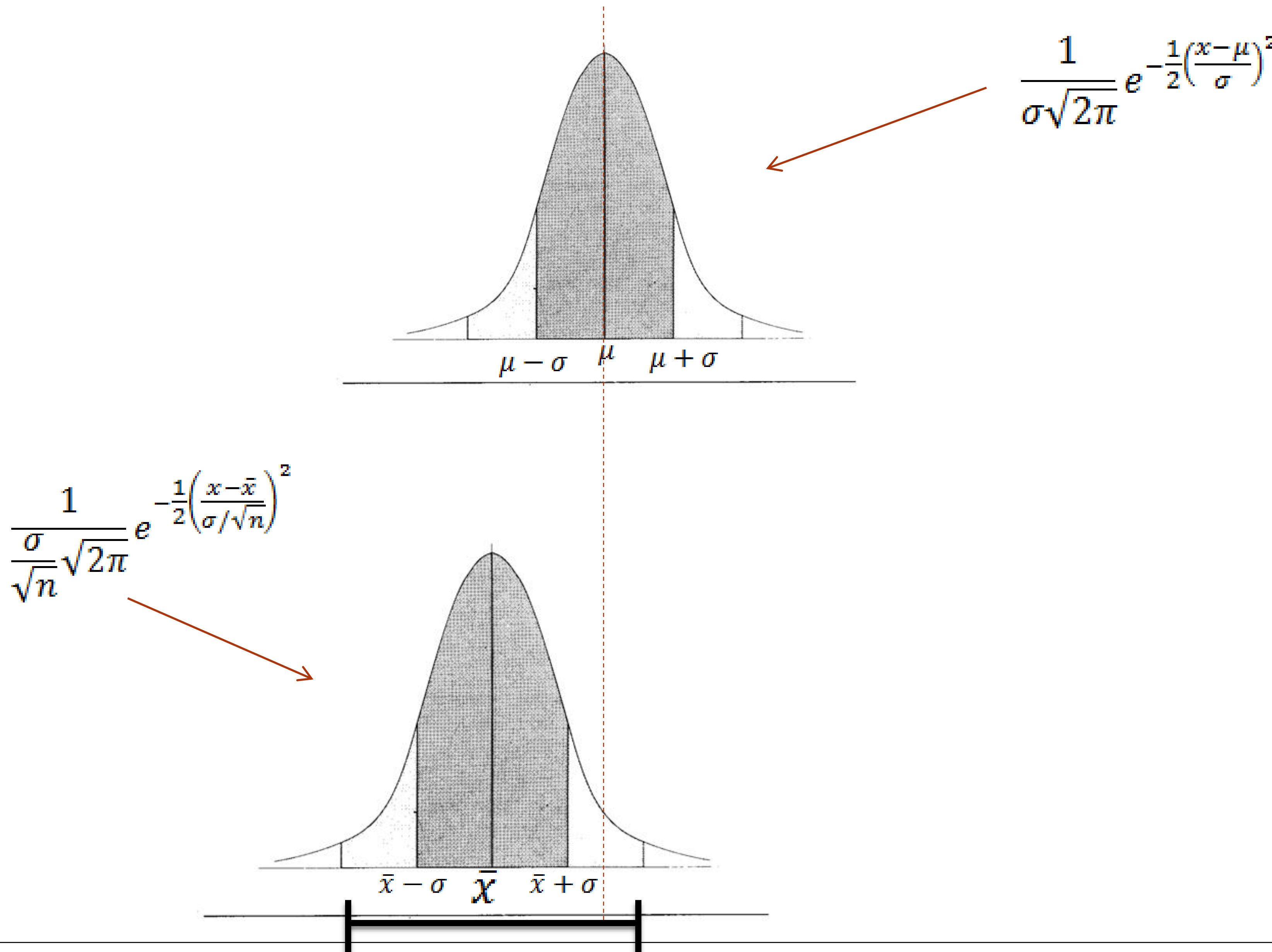
- Lo deseable es que el valor del promedio muestral ($\bar{x}$), fuera exactamente igual como el del poblacional (μ), como se dio en la filmina anterior.

- Como casi nunca se sabe el valor verdadero de μ, entonces lo ideal sería brindar posibles límites de confianza en donde se podría esperar que estos límites contengan al parámetro μ, con un determinado error probabilístico asociado.



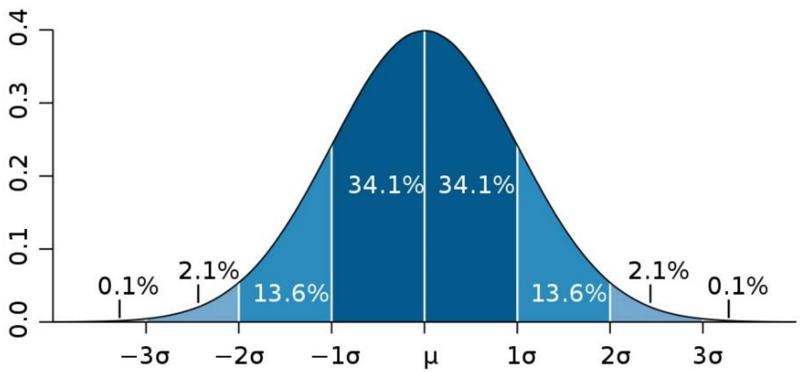
- Para este error probabilístico utilizaremos la curva normal estándar y sus valores de Z.

Límites de confianza para μ

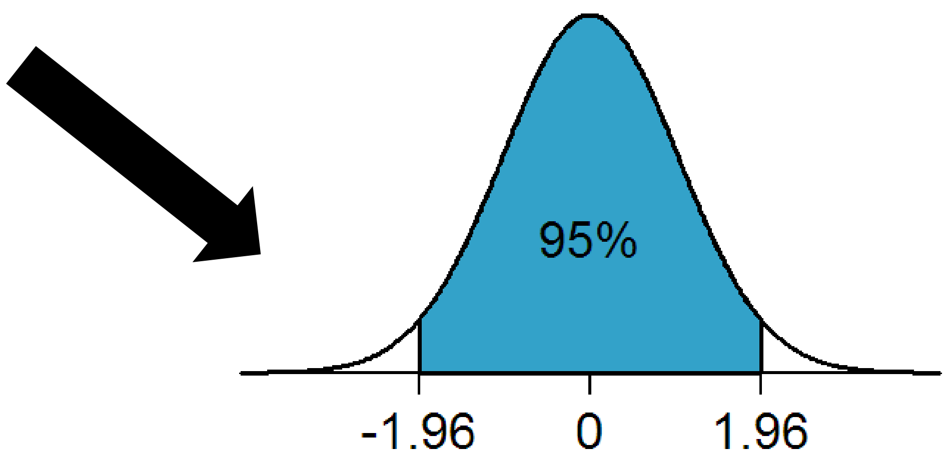


Límites de confianza para μ

- Ahora nos preguntamos: ¿dentro de qué valores se encontrará el 95% de los valores de $\bar{x}$? Recordemos que por probabilidades sabíamos que $P(-1.96 \leq z \leq 1.96) = 0{,}95$

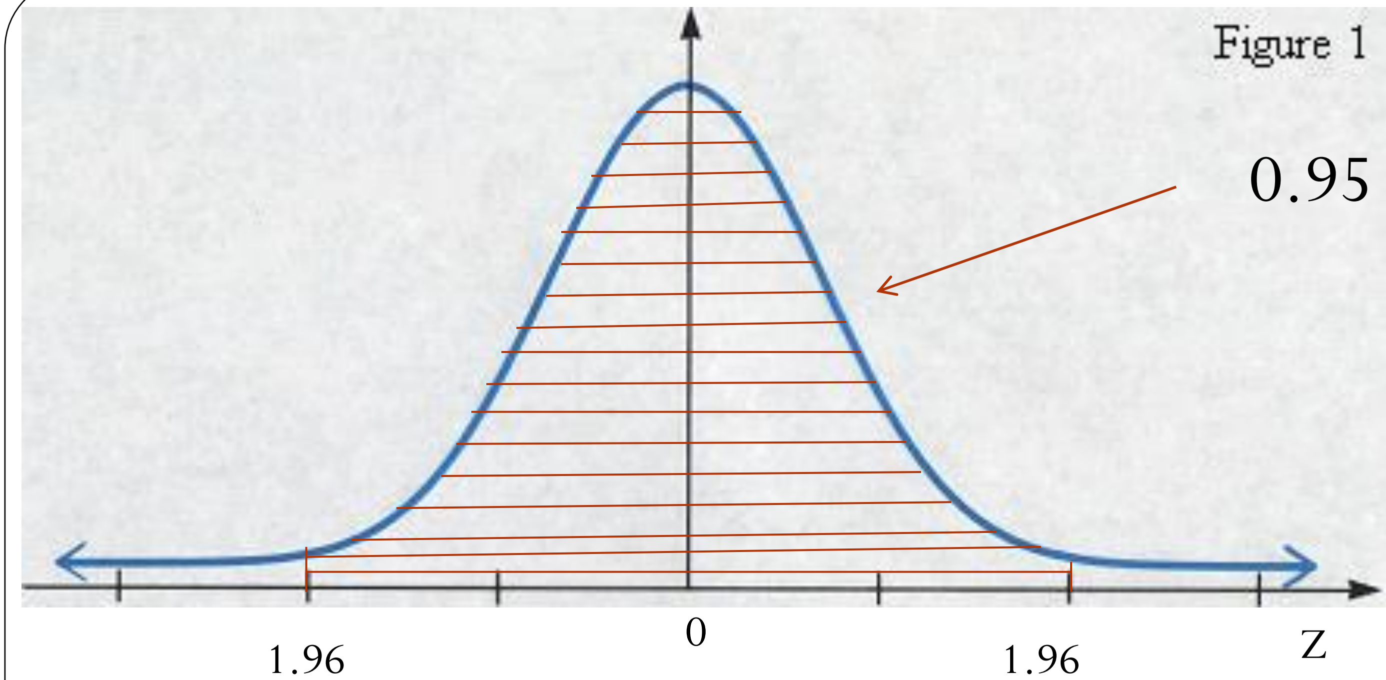


¿Cuánto acumulamos para $\pm\, 2\sigma$?



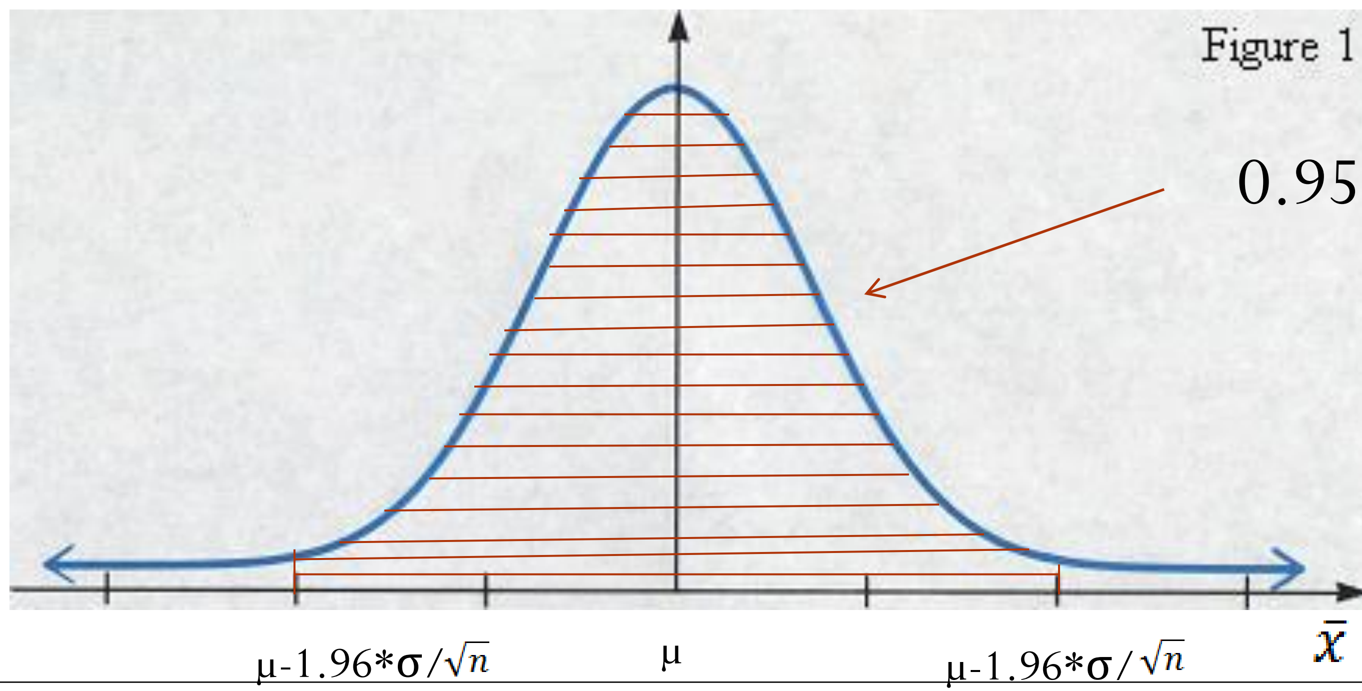
Límites de confianza para μ

- La gran pregunta acá es, sé de antemano que $P(-1.96 \leq z \leq 1.96) = 0{,}95$ para una curva $N(0{,}1)$, pero, ¿cómo entonces podría de forma general para saberlo para una $N(\mu, \sigma)$?



Normal Estándar

Normal con μ y σ



Límites de confianza para μ

- La probabilidad de que $\bar{x}$ asuma un valor en el intervalo

$\mu \pm 1.96*\sigma/\sqrt{n}$ es de 0.95.

- Esto puede expresarse simbólicamente en la siguiente forma:

$$P\left[\mu - 1.96 * \frac{\sigma}{\sqrt{n}} \leq \bar{x} \leq \mu + 1.96 * \frac{\sigma}{\sqrt{n}}\right] = 0.95$$

- De la expresión anterior se deriva la fórmula para el cálculo de un intervalo de confianza del 95% para μ. Mediante a una sencilla sencilla manipulación algebraica obtendríamos los intervalos del confianza con un 95%

$$P\left[\bar{x} - 1.96 * \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + 1.96 * \frac{\sigma}{\sqrt{n}}\right] = 0.95$$

Límites de confianza para μ

- Los límites de confianza se generan de la presente fórmula:

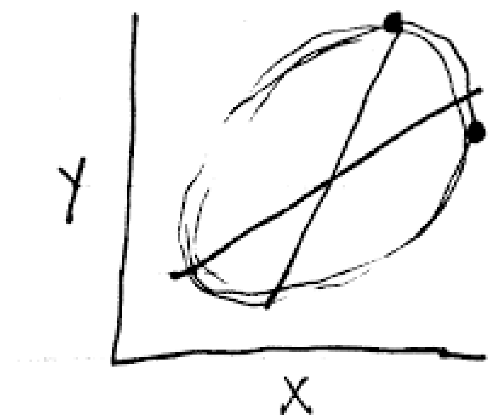
$$P\left[\underbrace{\bar{x} - 1.96 * \frac{\sigma}{\sqrt{n}}}_{\text{LI: Límite Inferior}} \leq \mu \leq \underbrace{\bar{x} + 1.96 * \frac{\sigma}{\sqrt{n}}}_{\text{LS: Límite Superior}}\right] = 0.95$$

$$P(LI \ \leq \mu \leq LS) = 0{,}95$$

Esquema de los límites de confianza para μ

De forma general, el intervalo de confianza se expresa como:

$$IC = [LI; LS] = \hat{x} \pm Z_{1-\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}}$$



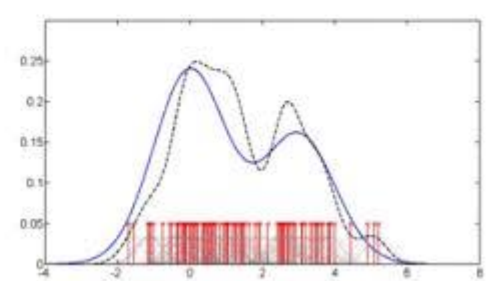
Para el caso de trabajar con nivel de confianza del $95\% \ (\alpha = 0{,}05)$

LI = límite inferior $= \hat{x} - Z_{1-\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}}$

LS = límite superior $= \hat{x} + Z_{1-\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}}$

El objetivo de la estimación por intervalos es calcular los límites de confianza (el LI y LS), para decir que, con una probabilidad del 95%, el intervalo va a contener al verdadero promedio poblacional μ.

En el desarrollo del IC solemos primero poner lo valores, la respectiva formula, calcular los límites, y finalmente interpretar el resultado.



Ejemplo de estimación por intervalos

- Se quiere estimar el promedio de edad (μ) de todos las personas de la Universidad Z.

- Se selecciona una muestra (n) de tamaño 50 para estimar el promedio de edad (μ), y el resultado fue de 25.7 años ($\bar{x}$), con una desviación estándar de 12 (σ).

- Dado que no se seleccionó a todas las personas de la Universidad Z, para tener más confianza en las estimaciones, se quiere utilizar los intervalos de confianza con un 95% de confiabilidad, esto es, con 5% de aceptar un error.

μ= ?????	σ = 12	Confianza=95%
$\bar{x}$ = 25.7	n= 50	Error=5%

Ejemplo de estimación por intervalos

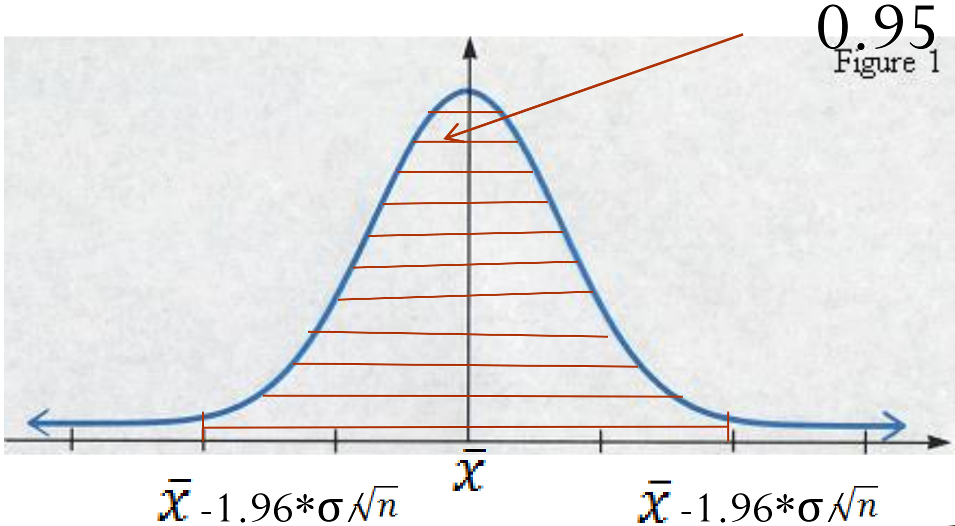
- Entonces necesitamos calcular los intervalos de confianza.

$$Límite\ inferior: LI = \bar{x} - 1.96 * \frac{\sigma}{\sqrt{n}}$$

$$Límite\ superior: LS = \bar{x} - 1.96 * \frac{\sigma}{\sqrt{n}}$$

- Y sabemos que, este límite tiene un 95% de confianza, por la fórmula antes vista:

$$P\left[\bar{x} - 1.96 * \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + 1.96 * \frac{\sigma}{\sqrt{n}}\right] =$$



Ejemplo de estimación por intervalos

- Solución:

$$Límite\ inferior: LI = 25.7 - 1.96 * \frac{12}{\sqrt{50}} = 22.37$$

$$Límite\ superior: LI = 25.7 + 1.96 * \frac{12}{\sqrt{50}} = 29.03$$

Corroborar:

$$\frac{22.37 + 29.03}{2} = 25.7$$

$$Límite\ inferior = 22.37$$

$$Límite\ superior = 29.03$$

Intervalo de confianza:

[22.37 ; 29.03]

- Interpretación: Se tiene una confianza del 95% de que intervalo calculado contenga al verdadero valor de la edad promedio en la Universidad Z.

Ejercicios… (error de $\alpha = 0,05$)

1. Se pretende estimar el numero promedio de latidos del corazón por minuto para cierta población. Se encontró que el numero promedio de latidos por minuto para 49 personas era de 90. Considere que esos 49 pacientes constituyen una muestra aleatoria y que la población sigue una distribución normal, con una desviación estándar de 10.

2. Se pretende estimar la concentración media de bilirrubina indirecta en el suero en niños de cuatro días de nacidos. La media para una muestra de 16 niños es de 5.98 mg/100 cc. Considérese que la concentración de bilirrubina en los niños de cuatro días de nacidos sigue una distribución aproximadamente normal con una desviación estándar de 3.5 mg/IOO cc.

Ejercicios… resolución de 1.

Valores

Nivel de confianza $(1-\alpha)$: 95% -- **Tamaño de muestra**: $n = 49$ --
Promedio muestral: $\hat{x} = 90$ -- **Desviación estándar muestral**: $\sigma = 10$ --
Error estándar muestral = $\frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{49}}$

Fórmula

Inervalo de confianza (CI) = $IC = \hat{x} \pm Z_{1-\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}}$ → $IC = \hat{x} \pm 1.96 * \frac{\sigma}{\sqrt{n}}$

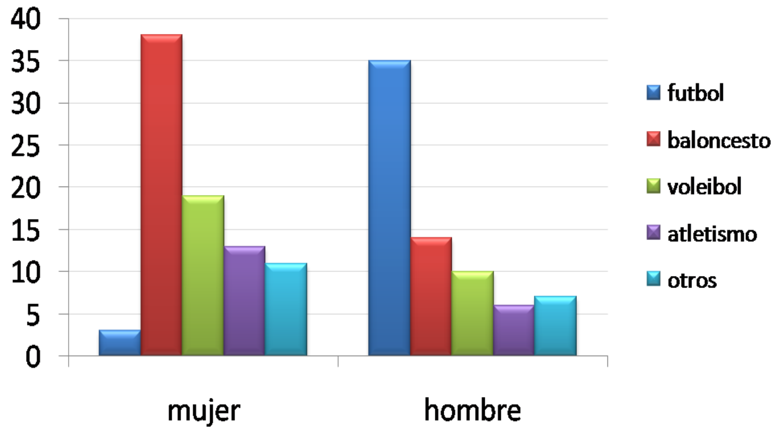
Resolución

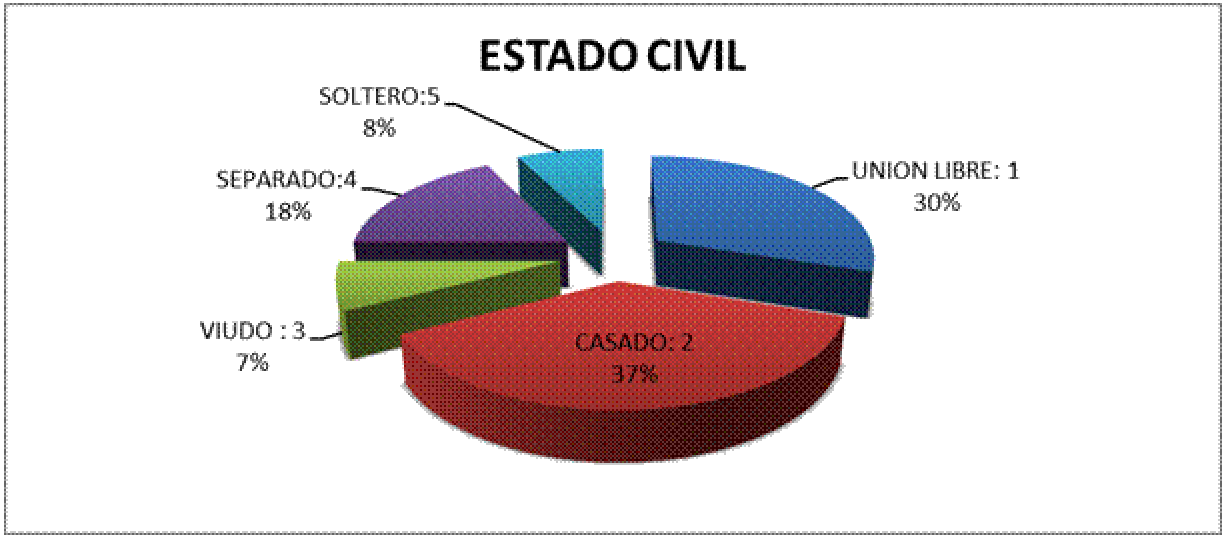
$$\boldsymbol{Límite\ inferior} = 90 - 1.96 * \frac{10}{\sqrt{49}} = 87.2$$
$$\boldsymbol{Límite\ superior} = 90 + 1.96 * \frac{10}{\sqrt{49}} = 92.8 \rightarrow IC = [87.2\ ;\ 92.8\ ]$$

Interpretación

Se tiene una confianza del 95% de que el intervalo de latidas por minutos calculado $[87.2\ ;\ 92.8\ ]$, esté conteniendo al valor poblacional de los latidos por minuto en la población de estudio.

Límites para variables cualitativas





¿Qué tipo de análisis realizamos con variables cualitativas?

Límites de confianza para P

- Supóngase ahora que la distribución de la población es una curva normal con media "P" y variancia "P(1-P)", (donde (1-P) también es conocida en ciertas ocasiones como "Q")y de esta se tomó de una muestra de tamaño n.

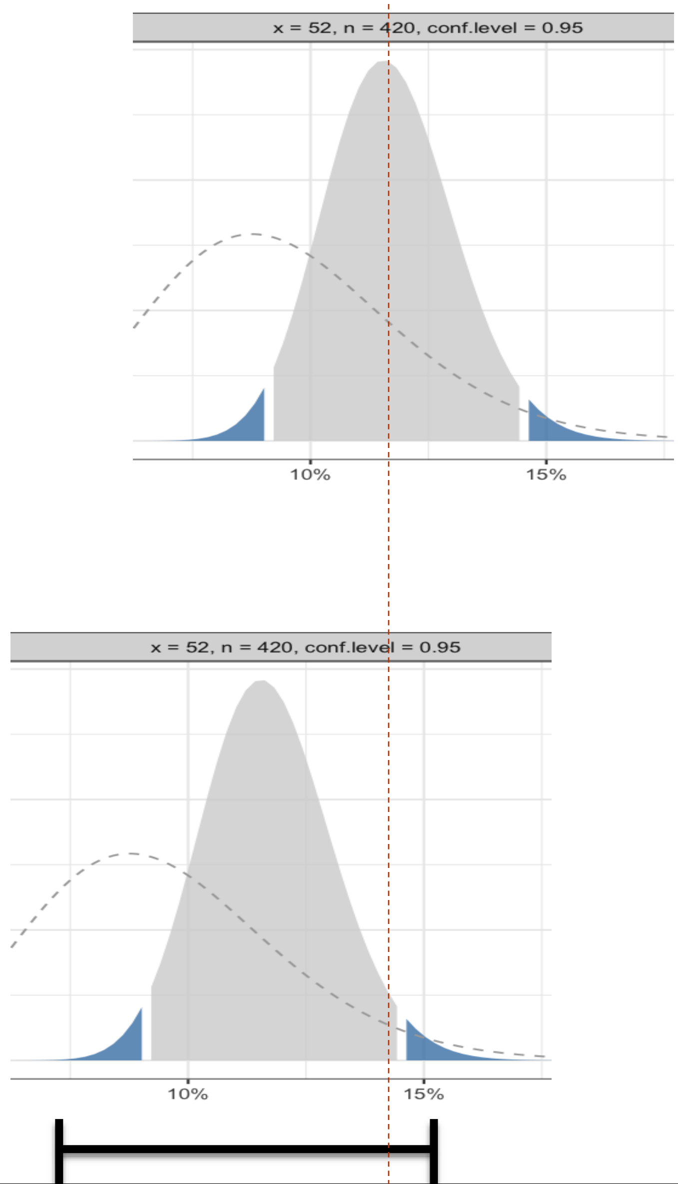
- Ahora, se desea saber si la posible distribución alrededor de la media muestral "***p***" que se obtuvo.

- Por teoría sabemos entonces que la distribución muestral tendrá una media de "***p***"y una variancia de $\frac{p(1-p)}{n}$ (la desviación

estándar estaría presentar por la fórmula $\sqrt{\frac{p(1-p)}{n}}$).

- Si la curva poblacional tiene una distribución $N\ (P, \sqrt{P(P-1)}$, la muestral tiene una curva de $N\ (p, \sqrt{\frac{p(1-p)}{n}})$

Límites de confianza para P



$$\frac{1}{\sqrt{P(1-P)2\pi}}e^{-\frac{1}{2}\left(\frac{p-P}{\sqrt{P(1-P)}}\right)^2}$$

$$\frac{1}{\sqrt{\frac{P(1-P)}{n}2\pi}}e^{-\frac{1}{2}\left(\frac{p-\bar{p}}{\sqrt{P(1-P)/n}}\right)^2}$$

Límites de confianza para p

Estadísticos	**Población**	**Muestra**
Posición central (promedio)	P	p
Variabilidad (desviación estándar)	$\sqrt{P(1-P)}$	$\sqrt{\dfrac{p(1-p)}{n}}$

$p: proporción\ establecida\ (Ej: p = 0.7)$
$1 - p = q: el\ compelemente\ (Ej: 1 - p = q = 0.3)$

Límites de confianza para P

- Muchas preguntas de interés para el analista en salud tiene relación con las proporciones de población:

1. ¿Qué proporción de pacientes que reciben un tipo especial de tratamiento se recuperan?
2. ¿Qué proporción de alguna población tiene cierta enfermedad?
3. ¿Qué proporción de una población es inmune a cierta enfermedad?

- Todas las anteriores son variables cualitativas que se suelen analizar mediante proporciones.

Límites de confianza para P

- Para estimar la proporción de una población se procede en la misma forma que cuando se estima el promedio de una población.

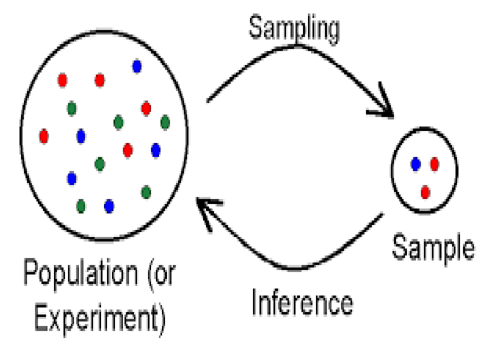
- Se extrae una muestra de la población, de interés y se calcula su proporción "***p***". Esta se utiliza como el estimador puntual para la proporción de la población.

- Como antes, el intervalo de confianza se obtiene mediante de la misma forma:

$$IC = estimador \ \pm (coeficiente\ de\ confiabilidad) * (error\ estándar)$$

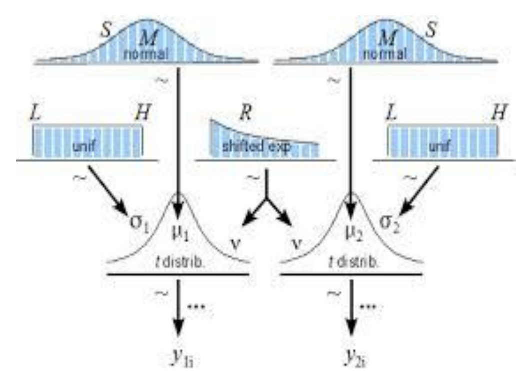
Límites de confianza para P: estimadores

- Para el caso de la estimación puntual para el promedio, este se suele representar como "*p*".



- Para obtener los intervalos de confianza, se termina el coeficiente de confiabilidad como $z_{1-\frac{\alpha}{2}}$
- El error estándar se determina como $\sqrt{\frac{p(1-p)}{n}}$
- El intervalo de confianza de la proporción se determina mediante la siguiente fórmula:

$$p \pm z_{1-\frac{\alpha}{2}} * \sqrt{\frac{p(1-p)}{n}}$$



Límites de confianza para P

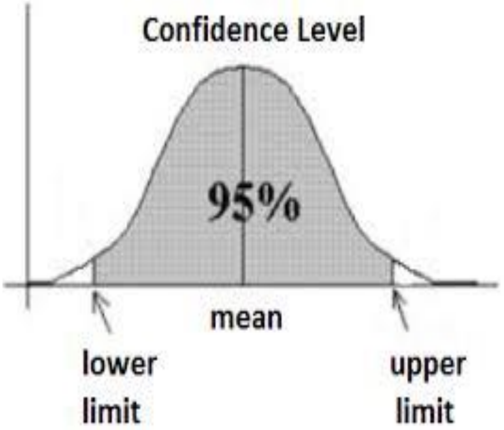
- Para cierto centro de salud, una muestra de 591 paciente internados en la sección de psiquiátrica, 204 admitieron que consumieron por lo menos un tipo de droga al menos una vez durante su vida. Se presente construir un intervalo de confianza del 95% para la proporción de individuos que consumieron droga durante su vida en la población muestrada para los internos del centro de salud.

- ¿Cuál es el supuesto que se utiliza?
- ¿Intervalo para qué tipo de variable?
- ¿Cuál sería el estimador puntual?
- ¿Cuál sería el valor del coeficiente de confiabilidad?
- ¿Cuál sería el intervalo de confianza?

Límites de confianza para P

- La mejor estimación puntual para la proporción sería:

$$p = \frac{204}{591} = 0.3452$$



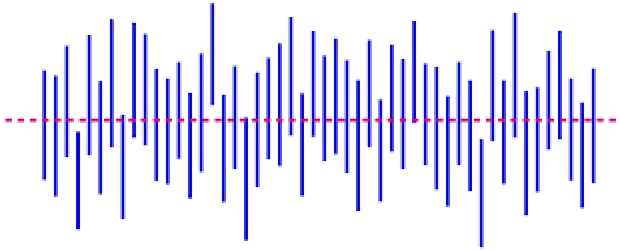
- El coeficiente de confiabilidad para un nivel de confianza del 95%, es de 1.96.
- La estimación del error estándar sería $\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.3452*0.6548}{591}} = 0.01956$

- El intervalo de confianza estaría dado cómo sigue:

$$IC\text{: } 0.3452 \pm 1.96 * (0.01956)$$

$$IC\text{: } 0.3452 \pm 0.0383$$

$$IC\text{: } [0.3069; 0.3835]$$



Límites de confianza para P

- ¿Cómo interpretamos el anterior intervalo de confianza (IC: $[0.3069; 0.3835]$)?

Índice

Interpretación de los intervalos de confianza

Entre las siguientes dos interpretaciones, cuál prefieren… :

1. Se tiene una confianza del 95% de que el salario promedio población esté entre los valores de ₡ 800 000 y el ₡ 1 200 000.

2. Se tiene una confianza del 95% de que el intervalo calculado del salario de ₡ 800 000 a ₡ 1 200 000, esté conteniendo el salario promedio poblacional.

Interpretación de los intervalos de confianza

¿Cómo se interpreta un intervalo de confianza?

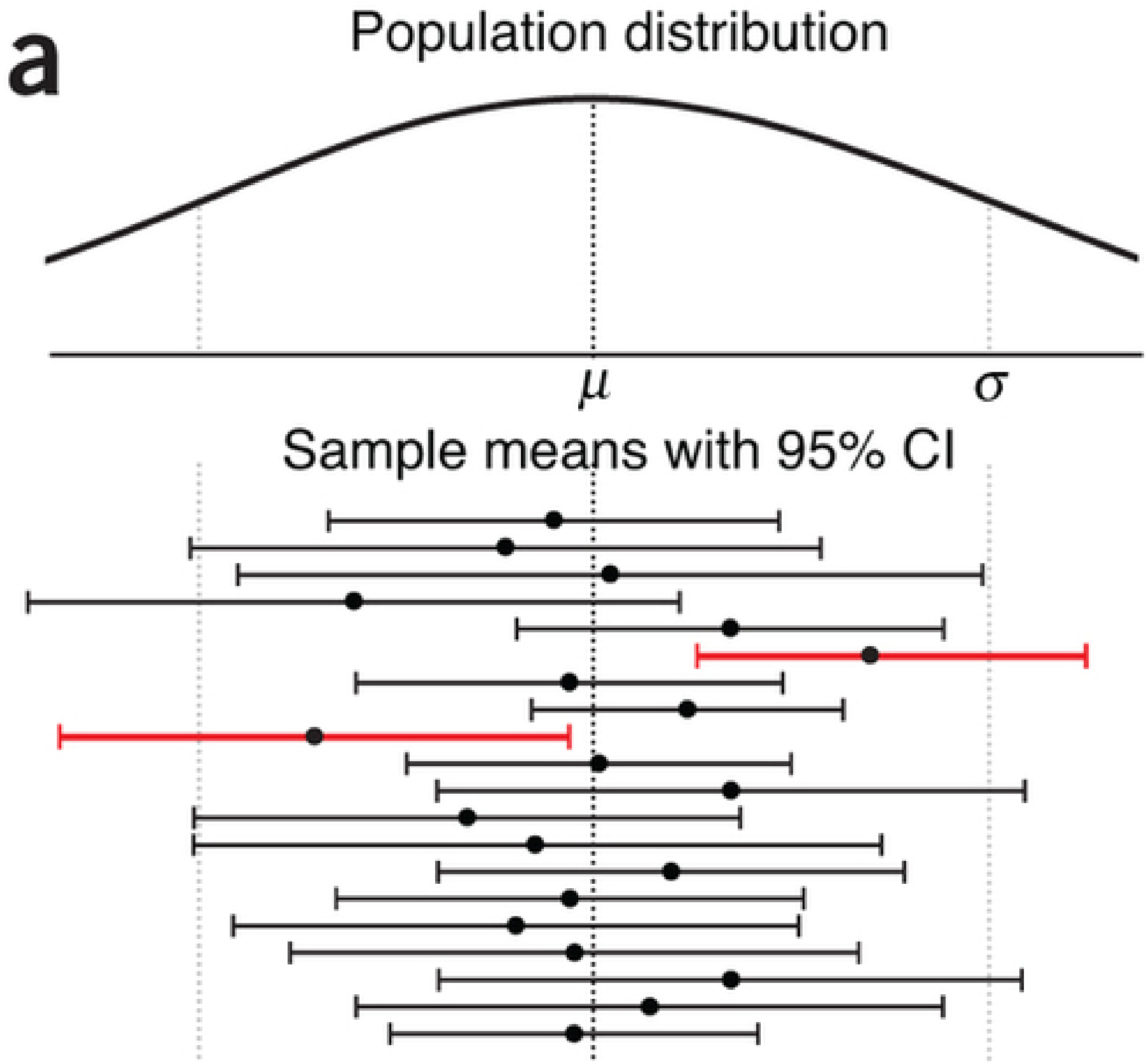
¿Se entiende bien la definición anterior?

¿ Se entiende que se pone de manifiesto que el término aleatorio es el intervalo y no el parámetro?

¿Se entiende bien el término del nivel de confianza, por ejemplo el 95% de confianza?

El intervalo de confianza suele ser la interpretación más equivocada…

Interpretación de los intervalos de confianza



Interpretación de los intervalos de confianza

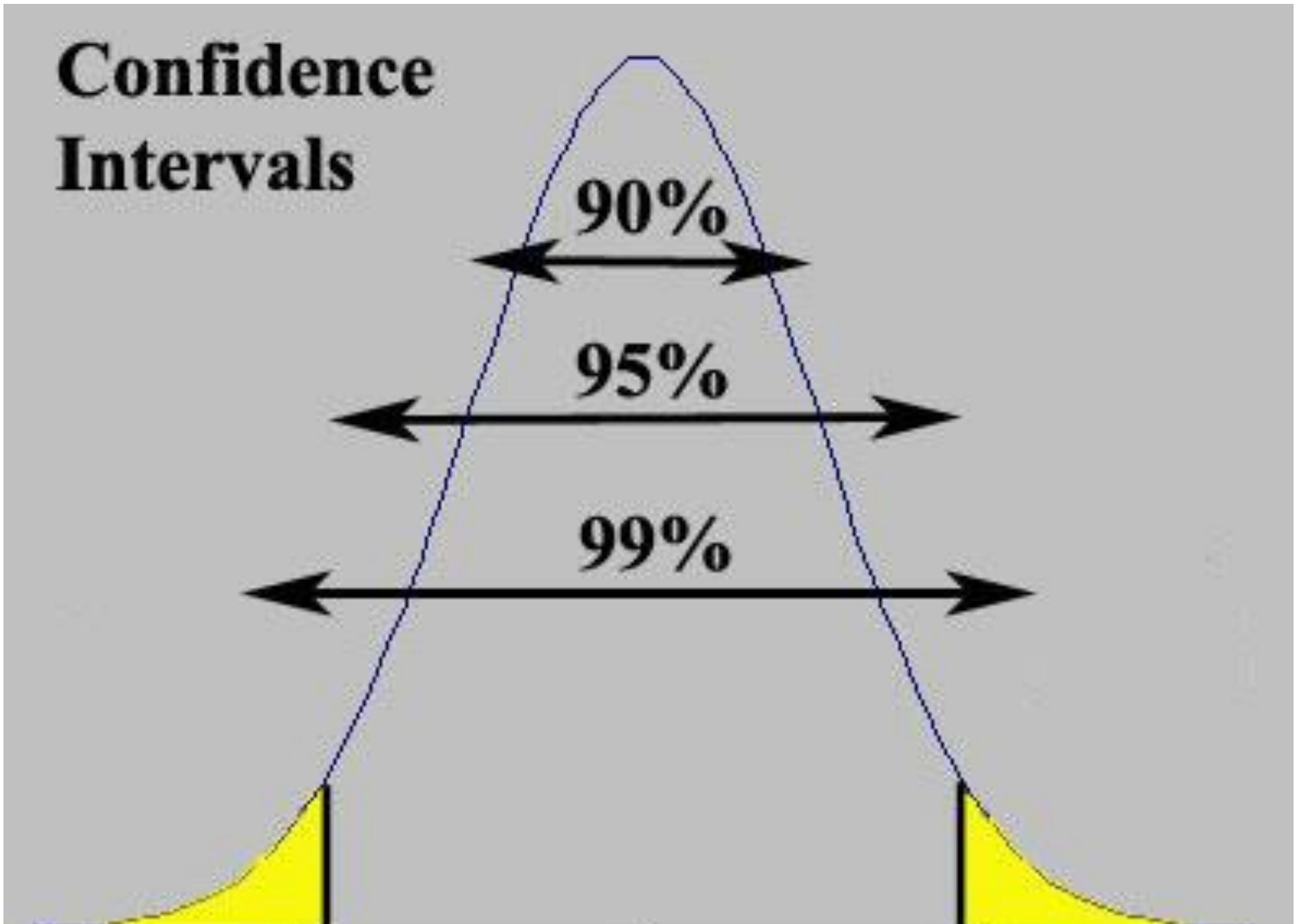
Interpretación probabilística: en la muestreo repetido, de una población con distribución normal y desviación estándar conocida 100(1-α) por ciento de todos los intervalos de la forma $\bar{x} \pm z_{1-\frac{\alpha}{2}}\,\sigma_{\bar{x}}$ incluyen a la larga la media de la población μ.

Interpretación práctica: cuando se hace un muestreo a partir de poblaciones que siguen una distribución normal y con desviación estándar conocida, existe un 100(1-α) por ciento de confianza de que el intervalo calculado $\bar{x} \pm z_{1-\frac{\alpha}{2}}\,\sigma_{\bar{x}}$, contiene la media de la población μ.

Nivel de confianza y tamaño del intervalo

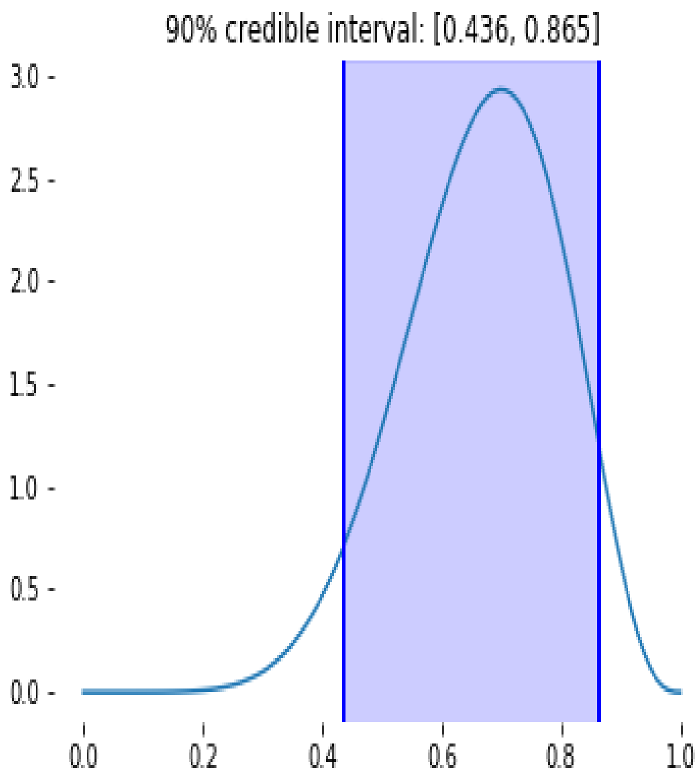
¿Cuándo se cambia el nivel de confianza, el intervalo es más grande o más pequeño?

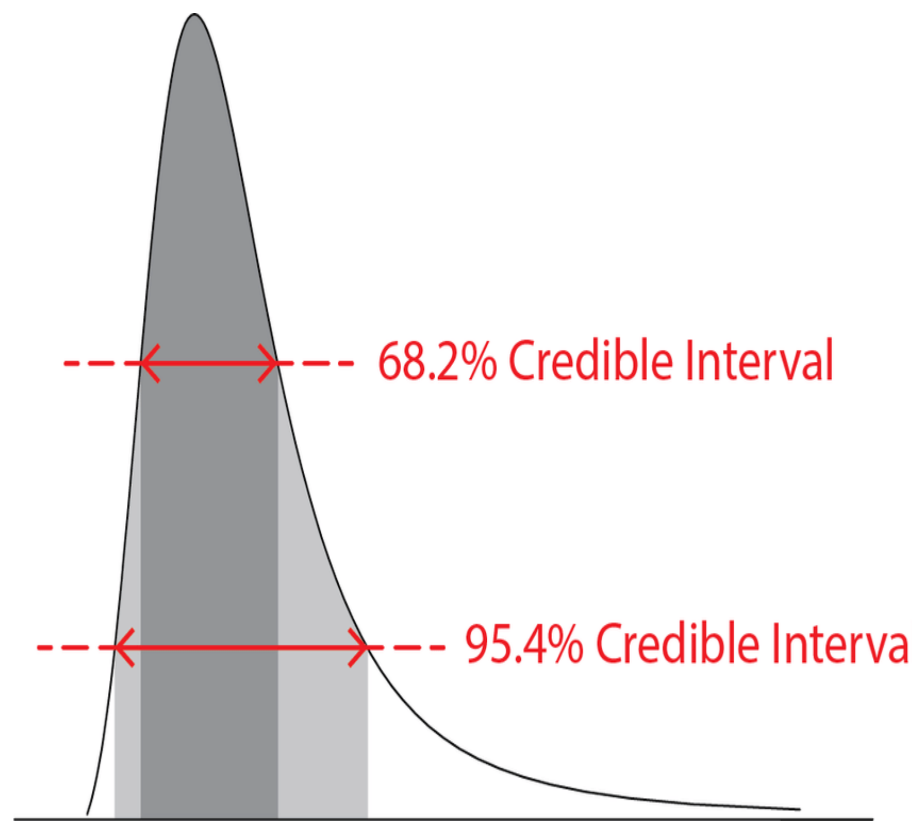
Nivel de confianza y tamaño del intervalo



Intervalo de credibilidad

Es posible que en la vida se topen no con un intervalo de confianza, sino con un intervalo de credibilidad. ¿Qué es? Es similar, pero con un método de estimación Bayesiano que utiliza información previa. Su interpretación es similar a la vista en este curso.





Reseñas finales

- La utilidad de la estimación por intervalos se utiliza en todos los campos en donde se aplica la estadística.

- En general, está es la técnica más utilizada.

- El conocer un rango de valores que podría tener cierto parámetro es de gran utilidad, ya que no siempre se puede der del todo preciso.

- Gran parte de la teoría de probabilidades, la curva normal y la estandarización residen en la estimación por intervalos.


arte
FIN