



CONSEJO DE BIOQUÍMICOS  
DE LA PROVINCIA DE JUJUY

# Diplomatura en Control de Calidad para Laboratorio de Análisis Clínicos

## Herramientas de estadística

Disertante:

**Dra Valeria Pfaffen**

*Facultad de Ciencias Químicas, Universidad  
Nacional de Córdoba (FCQ-UNC)*

Director:

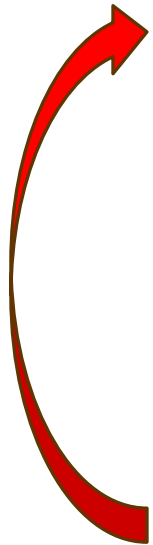
Bioq. Esp. César Collino

# ESTADISTICA

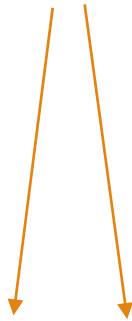
- ✓ Recoger
- ✓ Sintetizar
- ✓ Analizar

Datos cuya característica fundamental es la **VARIABILIDAD** para extraer conclusiones

Inferencia



POBLACION



¿Cuántos?  
¿Cuáles?

MUESTRA

ESTADISTICA  
DESCRIPTIVA

- Tablas
- Representaciones
- Síntesis de los datos

# **Análisis estadístico de los datos.**

## **Variables aleatorias**

Ejemplo: Valor del Hematocrito en sangre.

- ✓ Determinaciones repetitivas de una misma muestra
- ✓ Determinación en dos personas normales.
- ✓ Determinación en la misma persona en dos momentos diferentes.

## **¿Porqué son aleatorias?**

Variaciones:

\*Instrumento    \*Analista    \*Muestra    \*Ruido

# Variables (según su naturaleza)

- ✓ **Cuantitativas:** Continuas: medición en escala continua, se puede medir cualquier valor dentro de un intervalo (pH, concentraciones, edad, hematocrito)  
Discretas: cantidades que se pueden contar (cantidad de colonias)
- ✓ **Categóricas:** Ordinales: grados (severo, moderado, leve)  
Nominales: no ordenados (sexo, presencia o no de enfermedad, grupo sanguíneo, colores)

El tipo de variable determina el método de medida y el test estadístico a emplear.

# Variables (según el objetivo de estudio)

## ✓ Variables de exposición:

Permiten medir los factores estudiados:

Variable dependiente

Variable independiente

Ejemplo: recuento de glóbulos rojos ( variable ..... ) en pacientes con distintos tipos de anemia ( variable ..... ) Nominales ?

Numéricas?

*Puede haber más de una.  
Determina el tipo de test a utilizar.*

# Concepto de población

**Conjunto de elementos acotados en el tiempo y en un espacio determinado, con alguna característica común medible**

✓ Incluye todos los datos

✓ Puede ser: *Finita*: el número de elementos es N  
*Infinita*

✓ Parámetro: propiedad o característica de interés. No varía o varía muy poco en el tiempo!!!. ( $\mu$ = media poblacional) . No se puede medir.

# Concepto de muestra

**Parte de la población bajo estudio.**

**Tamaño muestral =  $n$**

**$\bar{x}$  = media muestral**

**Estimador de la media poblacional**



# **Análisis estadístico:**

1. Estadística descriptiva
2. Inferencia estadística

# ESTADISTICA DESCRIPTIVA

1. Presentación de los datos. Tablas.  
Distribución de frecuencias
2. Medidas de posición y dispersión
3. Representaciones Gráficas.

# Presentación de los datos

## Tablas. Distribuciones de frecuencias.

**Grupo Sanguíneo, número de hermanos y edad de 500 pacientes de un hospital**

Alumno Nro.	Grupo sanguíneo	Número de hermanos	Edades
1	A	0	70
2	B	3	67
-	-	-	-
-	-	-	-
-	-	-	-
-	-	-	-
500	AB	2	71

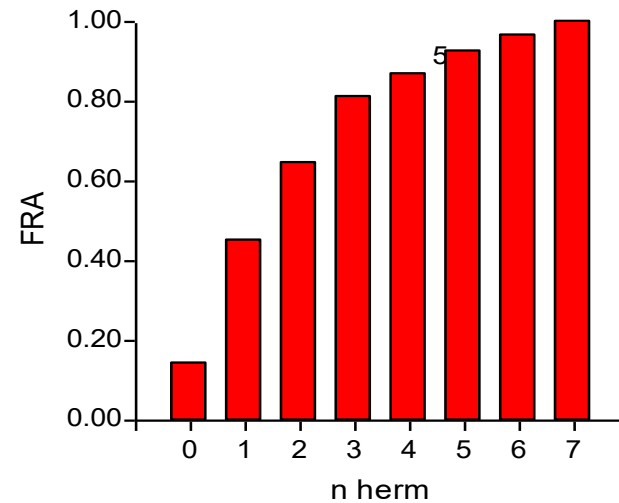
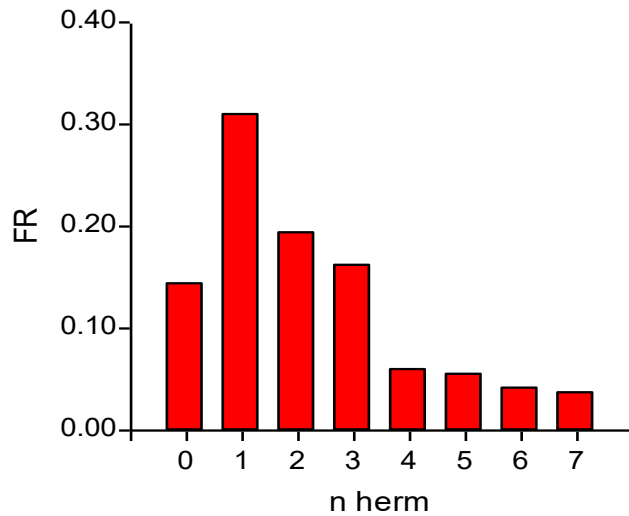
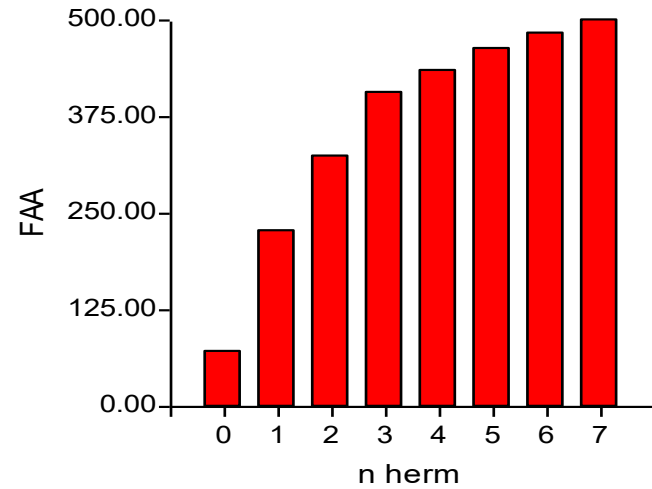
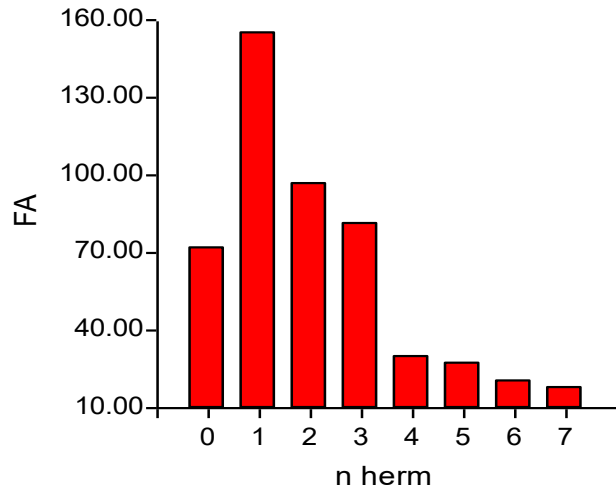
## Distribución del grupo sanguíneo en 500 pacientes

Grupo sanguíneo	Frecuencia absoluta	Frecuencia relativa	Porcentaje
A	150	0.30	30%
B	75	0.15	15%
AB	25	0.05	5%
O	250	0.50	50%
Total	500	1.00	100%

## Distribución del número de hermanos (excluido él mismo) de una muestra de 500 pacientes

Número de hermanos	FA	FAA	FR	FRA	%
0	72	72	0.144	0.144	14.4
1	155	227	0.310	0.454	31.0
2	97	324	0.194	0.648	19.4
3	81	405	0.162	0.810	16.2
4	30	435	0.060	0.870	6.0
5	27	462	0.054	0.924	5.4
6	20	482	0.040	0.964	4.0
Más de 6	18	500	0.036	1.000	3.6
total	500	500	1.000	1.000	100

# Diagramas de Frecuencias



# Datos cuantitativos continuos. Hacer intervalos.

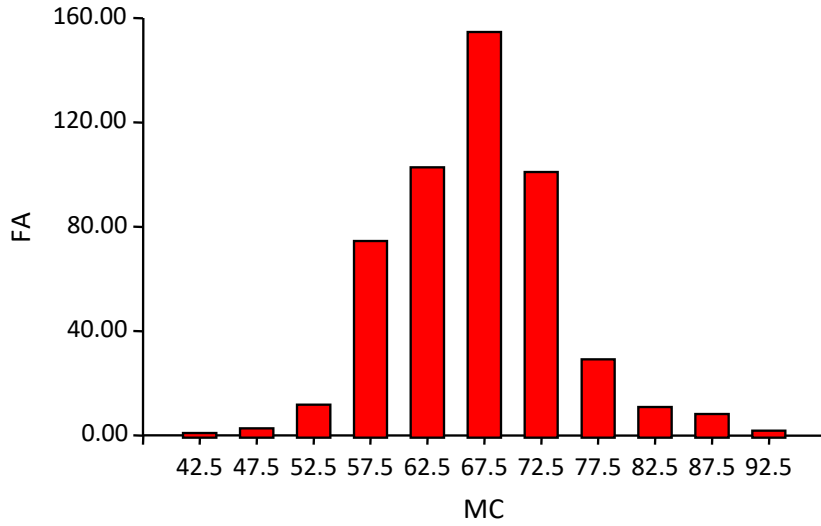
## Distribución de edades de una muestra de 500 pacientes

Intervalo de clase	FA	FAA	FR	FRA	%	Marca de clase
$x \leq 45$	1	1	0.002	0.002	0.2	42.5
$45 \leq x < 50$	3	4	0.006	0.008	0.6	47.5
$50 \leq x < 55$	12	16	0.024	0.032	2.4	52.5
$55 \leq x < 60$	75	91	0.150	0.182	15.0	57.5
$60 \leq x < 65$	103	194	0.206	0.388	20.6	62.5
$65 \leq x < 70$	155	349	0.310	0.698	31.0	67.5
$70 \leq x < 75$	101	450	0.202	0.900	20.2	72.5
$75 \leq x < 80$	29	479	0.058	0.958	5.8	77.5
$80 \leq x < 85$	11	490	0.022	0.980	2.2	82.5
$85 \leq x < 90$	8	498	0.016	0.996	1.6	87.5
$x \geq 90$	2	500	0.004	1.000	0.4	92.5
total	500	500	1.000	1.000	100.0	-

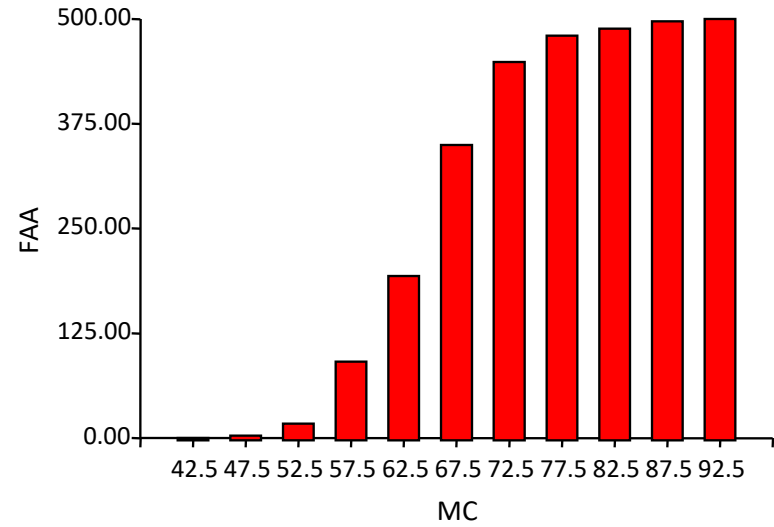
**FAA= Frecuencias absolutas acumuladas . Suma de las frecuencias absolutas.**  
**FRA= Frecuencias relativas acumuladas . Suma de las frecuencias relativas.**

# Diagramas de Frecuencias

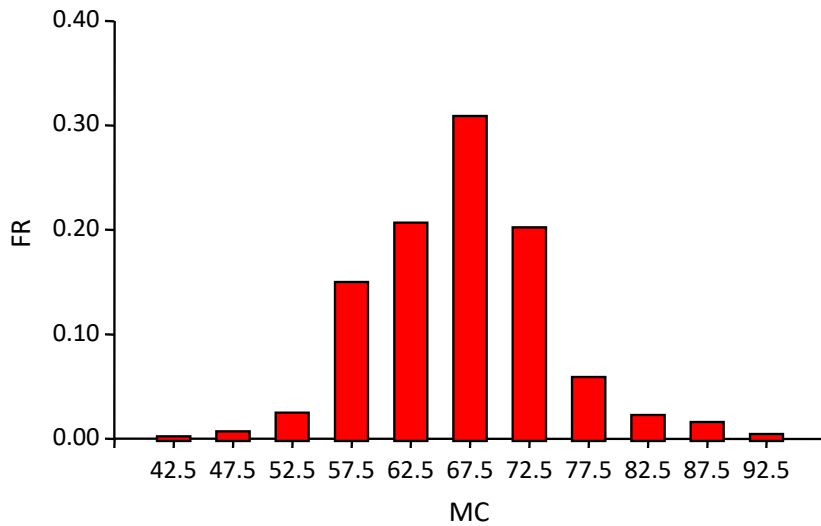
*Edades*



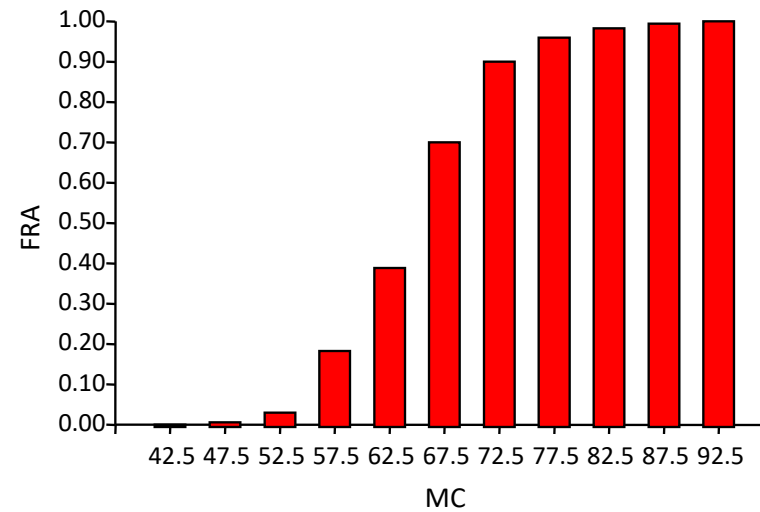
*Edades*



*Edades*



*Edades*





# Medidas de posición:

- ✓ **Moda:** Valor más frecuente. Puede haber más de una  
Tiene sentido definirla cuando hay muchos datos y están agrupados en intervalos de clase.
- ✓ **Mediana:** Deja tantas observaciones por encima como por debajo de ella.  
Es el valor central o el promedio de los dos valores centrales
- ✓ **Media:** Promedio de los valores

## ✓ *Percentiles o cuantiles*

$p_1, p_2 \dots p_i \dots p_{99}$

valores que dejan a su izquierda el  $i\%$  de los datos.

## ✓ *Cuartiles:*

$p_{25}, p_{50}, p_{75}$  ; dividen a la muestra en cuatro partes iguales

# Medidas de dispersión

✓ **Rango:** Es la diferencia entre el valor más alto y el más bajo de la variable

## ***Ventajas***

- Fácil de calcular
- Iguales unidades que los datos de origen

## ***Desventajas***

- Considera solo dos valores de la muestra

Muestra 1: 1,1,1,1,1,5

Muestra 2: 1,2,3,3,4,5

- En general aumenta con el tamaño de la muestra

✓ **Desviación Estandar:** Desvío de cada dato respecto de una media, se elevan al cuadrado se suman, se divide por  $n$  o por  $n-1$  y se saca la raíz para que tengan la misma unidad que los datos.

**Ventajas** —————> Tiene las mismas unidades que los datos.

✓ **Varianza:** Es el cuadrado de la desviación estándar

**Ventajas**

Es mas fácil de ser tratada matemáticamente

Ambas utilizan el valor de la media como centro para calcular la dispersión

✓ **Coeficiente de Variación:**  $CV = (s/\bar{x}) * 100$

**Ventajas** —————> Permite comparar desviaciones estandar

# Cómo se simbolizan las medidas de posición, dispersión segun sean poblacionales o muestrales

	Parámetro Estimador	
	<i>Población</i>	<i>Muestra</i>
<i>Media</i>	$\mu$	$\bar{x}$
<i>Desviación estándar</i>	$\sigma$	$s$
<i>Varianza</i>	$\sigma^2$	$s^2$

# Gráficos de los datos

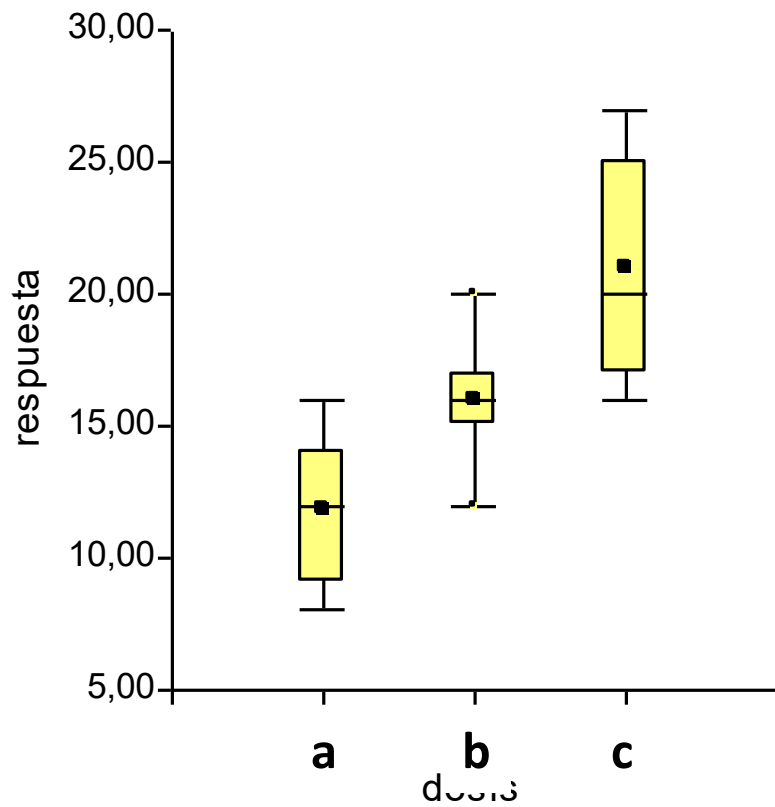
- 1) Gráfico de barras
- 2) Gráfico de puntos
- 3) Gráfico de cajas
- 4) Gráfico de densidad de puntos
- 5) Diagrama de dispersión

***En qué caso se usa cada uno?***

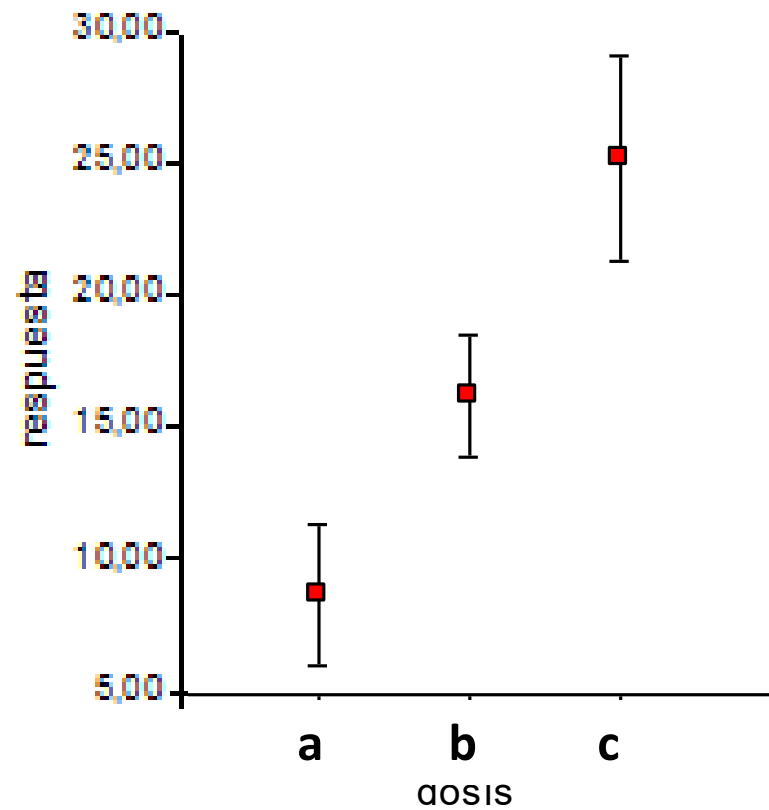
**La siguiente tabla muestra los resultados de un experimento de respuesta a una dosis, realizado con 5 animales a los que se le aplicaron 3 dosis de antibiótico diferentes.**

<b>dosis</b>	<b>Respuesta</b>
a	8, 12, 9, 14, 6
b	16, 20, 12, 15, 17
c	20, 17, 25, 27, 16

***Que gráfico haría??***



***Grafico de cajas***



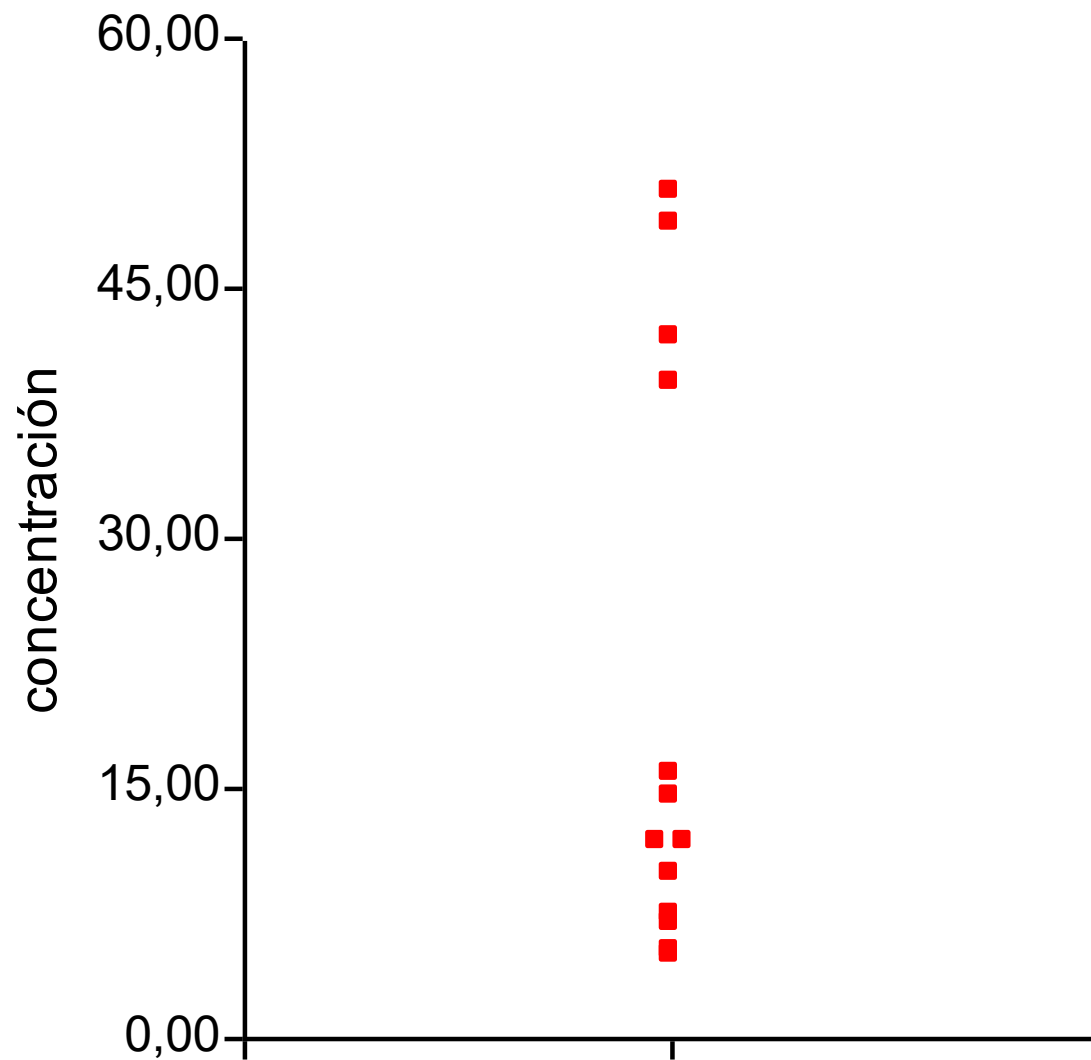
***Grafico de puntos.***



**Los siguientes valores de contenido de un metabolito en la Sangre de un paciente en 13 extracciones diferentes durante el día:**

11,6	39,2	4,9	7,3	50,6	9,8	11,6	6,7	42,1	14,4	5,1	48,8	15,9
------	------	-----	-----	------	-----	------	-----	------	------	-----	------	------

*Los datos están informados en mg/L.  
Haga un gráfico de densidad de puntos y analice los resultados.*

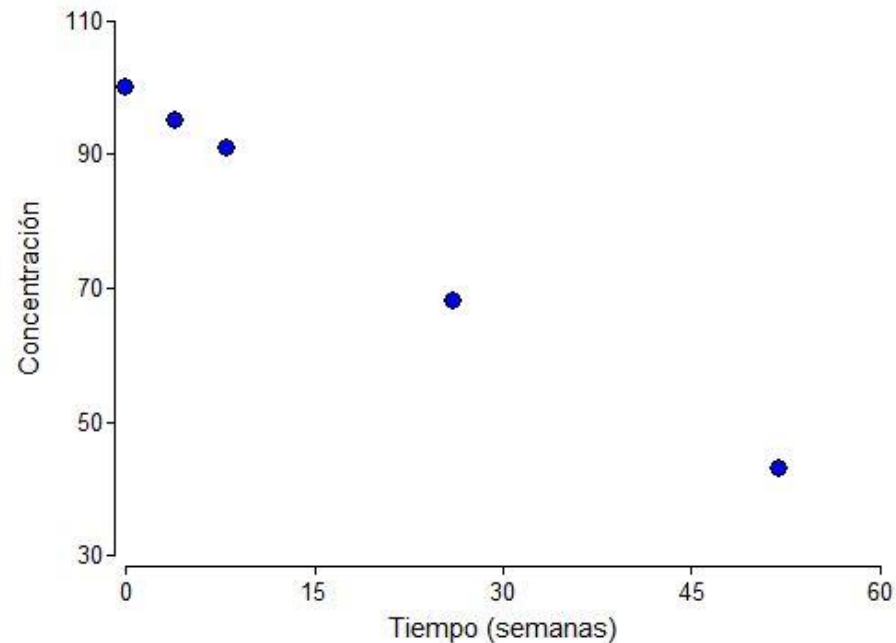


Se midió la concentración de una droga en solución en función del tiempo

*Representar la concentración en función del tiempo.*

Tiempo (semanas)	Concentración
0	100
4	95
8	91
26	68
52	43

*Diagrama de dispersión*



**La siguiente tabla muestra las frecuencias de niveles de colesterol en sangre de una población de 1067 pacientes**

Clase	LI	LS	MC	FA	FR	FAA	FRA
1	100	135	117.5	13	0.01	13	0.01
2	135	170	152.5	150	0.14	163	0.15
3	170	205	187.5	442	0.41	605	0.57
4	205	240	222.5	299	0.28	904	0.85
5	240	275	257.5	115	0.11	1019	0.96
6	275	310	292.5	34	0.03	1053	0.99
7	310	345	327.5	9	0.01	1062	1
8	345	380	362.5	5	0	1067	1

*Cual es percentil 85 y su significado.*

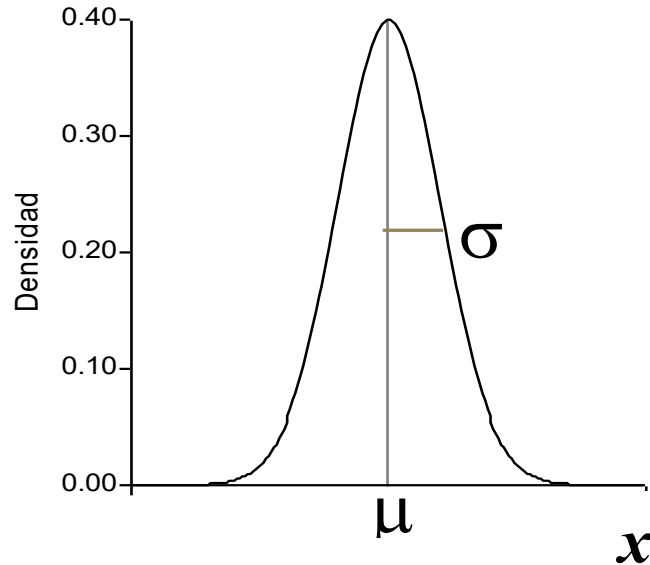
*Cuál es el valor aproximado de la mediana*

*Cual es la moda*

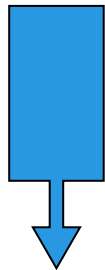
*Que porcentaje de individuos tiene valores superiores a 275 mg %*

# Distribución normal y distribución normal estándar de los datos

$\mu$  y  $\sigma$  verdaderos parámetros poblacionales, NO SE CONOCEN EXACTAMENTE

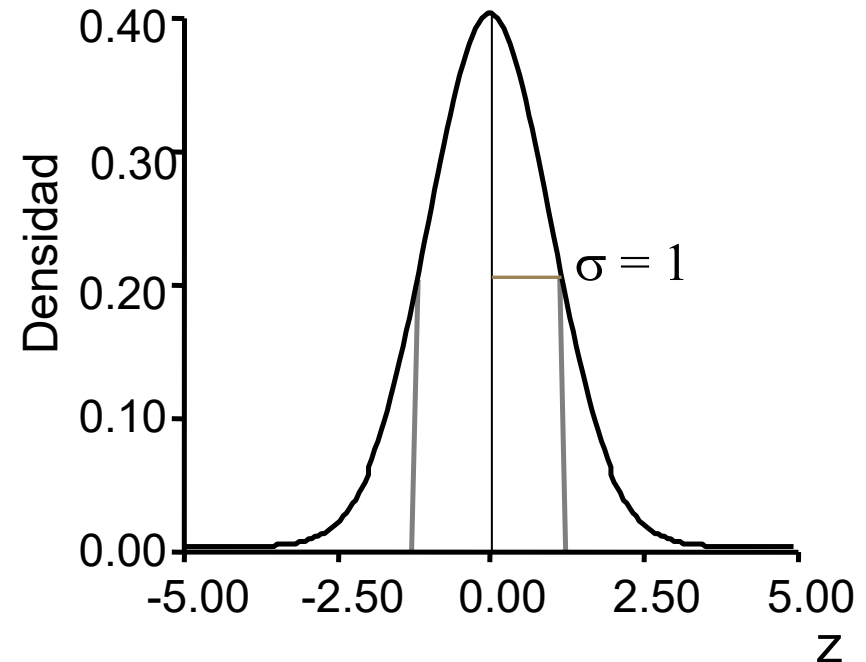


$$X \sim N(\mu, \sigma)$$



$$Z \sim N(0, 1)$$

$$Z = \frac{X - \mu}{\sigma}$$



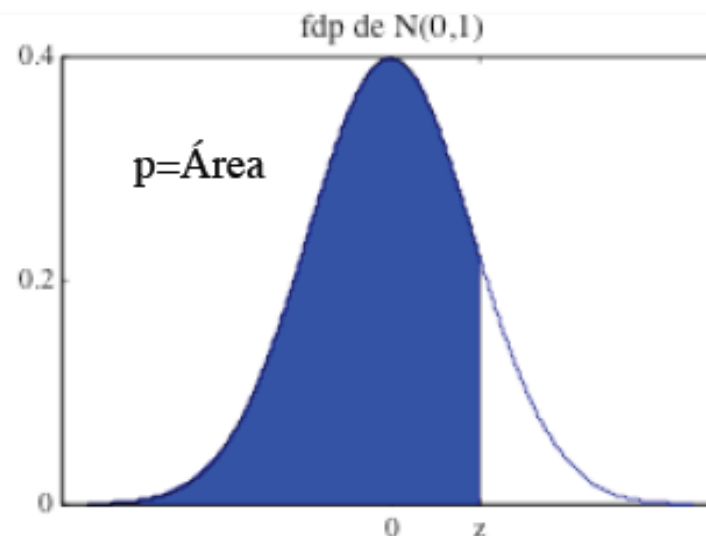
## Distribución normal estándar

$Z \sim N(0,1)$

Tabla de la función de distribución:

$$P(Z \leq z) = p$$

En la tabla figuran los valores de probabilidad acumulada  $p$  en función de  $z$ .



<b>z</b>	<b>0.00</b>	<b>0.01</b>	<b>0.02</b>	<b>0.03</b>	<b>0.04</b>	<b>0.05</b>	<b>0.06</b>	<b>0.07</b>	<b>0.08</b>	<b>0.09</b>
<b>0.0</b>	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
<b>0.1</b>	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
<b>0.2</b>	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
<b>0.3</b>	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
<b>0.4</b>	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
<b>0.5</b>	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
<b>0.6</b>	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
<b>0.7</b>	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
<b>0.8</b>	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
<b>0.9</b>	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
<b>1.0</b>	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
<b>1.1</b>	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
<b>1.2</b>	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
<b>1.3</b>	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
<b>1.4</b>	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
<b>1.5</b>	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441

<b>1.6</b>	<b>0.9452</b>	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
<b>1.7</b>	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
<b>1.8</b>	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
<b>1.9</b>	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
<b>2.0</b>	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
<b>2.1</b>	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
<b>2.2</b>	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
<b>2.3</b>	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
<b>2.4</b>	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
<b>2.5</b>	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
<b>2.6</b>	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
<b>2.7</b>	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
<b>2.8</b>	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
<b>2.9</b>	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
<b>3.0</b>	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
<b>3.1</b>	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
<b>3.2</b>	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
<b>3.3</b>	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
<b>3.4</b>	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
<b>3.5</b>	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
<b>3.6</b>	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	<b>0.9999</b>

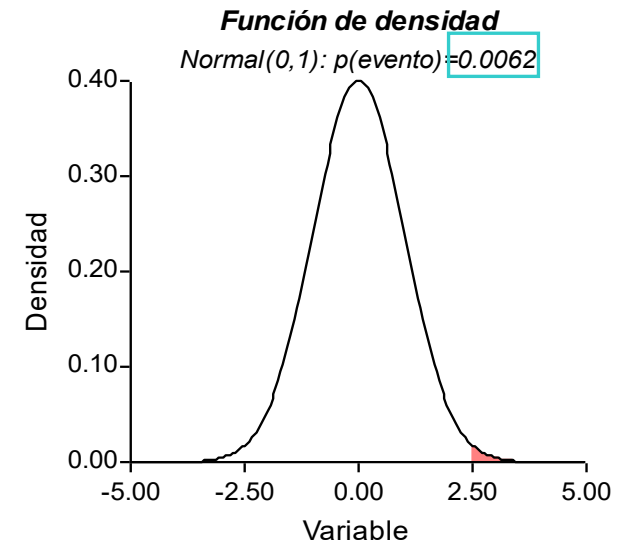
**Observaciones:** Si  $X \sim N(\mu, \sigma)$ , entonces:  $Z = (X - \mu) / \sigma$  sigue una distribución  $N(0,1)$

# Ejemplo

Se determinó la hemoglobina con un método colorimétrico de un grupo de 200 individuos que padecen anemia, obteniéndose un promedio de 10 g/dL, con un desvío estándar de 0,1 g/dL. Se pide calcular la probabilidad de encontrar un paciente elegido al azar cuyo valor sea mayor de 10,25 g/dL

$$Z = \frac{x - \mu}{\sigma} = (10.25 - 10)/0.1 = 2.5$$

1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999



**Observaciones:** Si  $X \sim N(\mu, \sigma)$ , entonces:  $Z = (X - \mu)/\sigma$  sigue una distribución  $N(0,1)$



# Muestra aleatoria

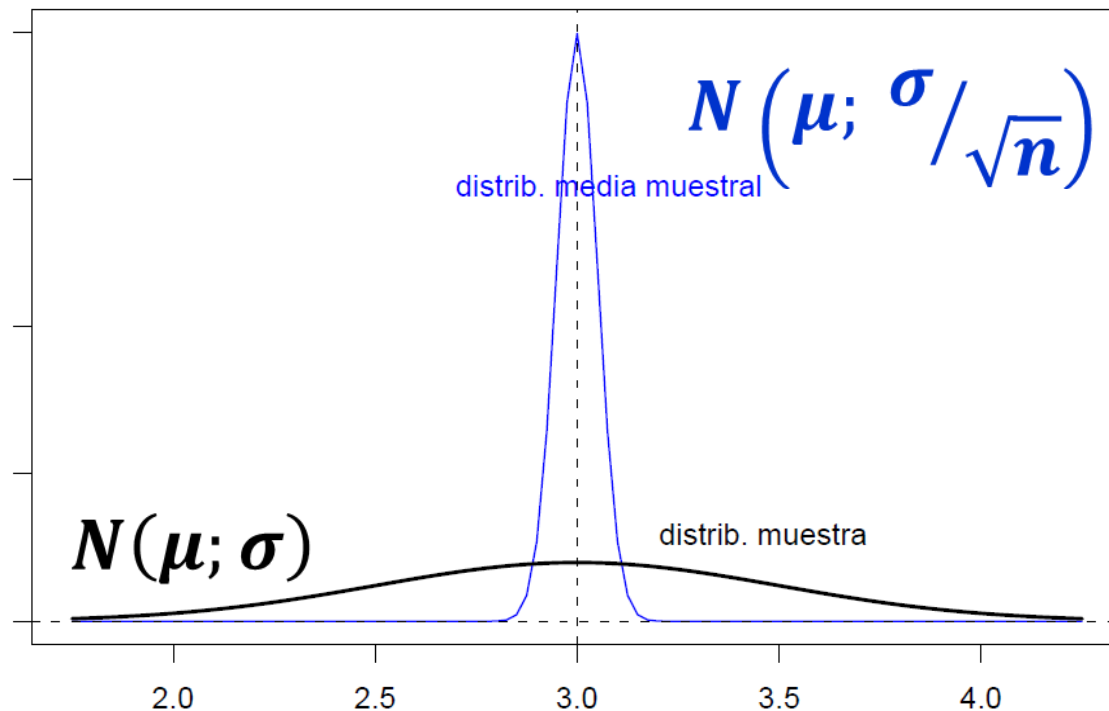
Dada una variable  $x$ , una muestra aleatoria es un conjunto  $n$  de unidades experimentales aleatorias, independientes e idénticamente distribuidas a las cuales se les medirá esa variable  $x$

*Si una variable aleatoria sigue una distribución normal con parámetros  $(\mu, \sigma)$  entonces la media muestral sigue una distribución normal con parámetros  $(\mu, \sigma / \sqrt{n})$*

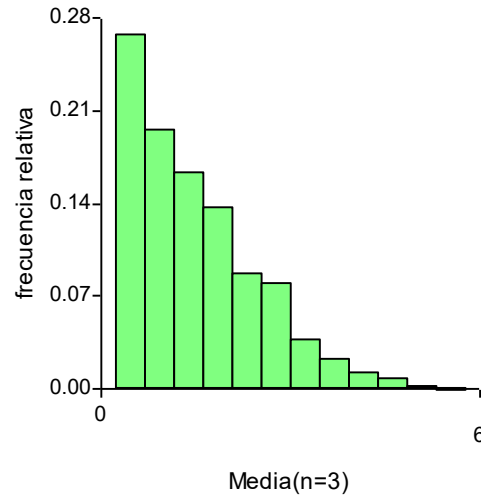
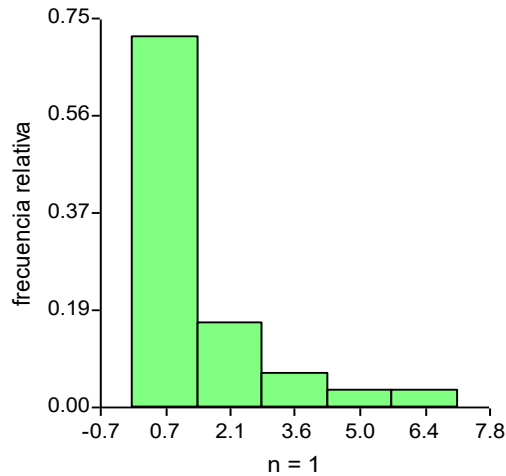
# Distribuciones de la población y distribución muestral de la media

$\sigma$  ( $DE_X$ ): mide variabilidad entre sujetos

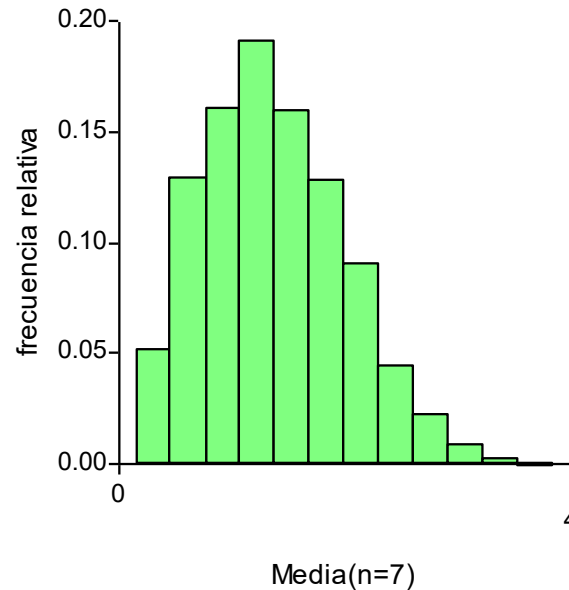
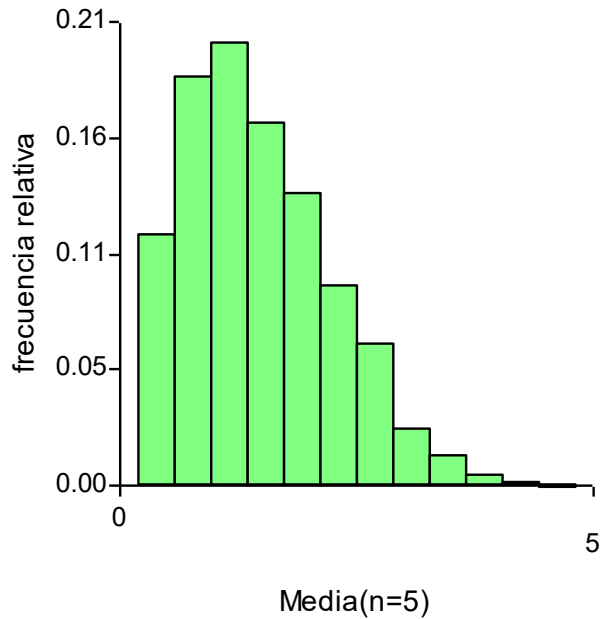
$\sigma/\sqrt{n}$  ( $EE_{\bar{X}}$ ): mide variabilidad entre muestras,  
cuánta variabilidad cabe esperar en muestras futuras

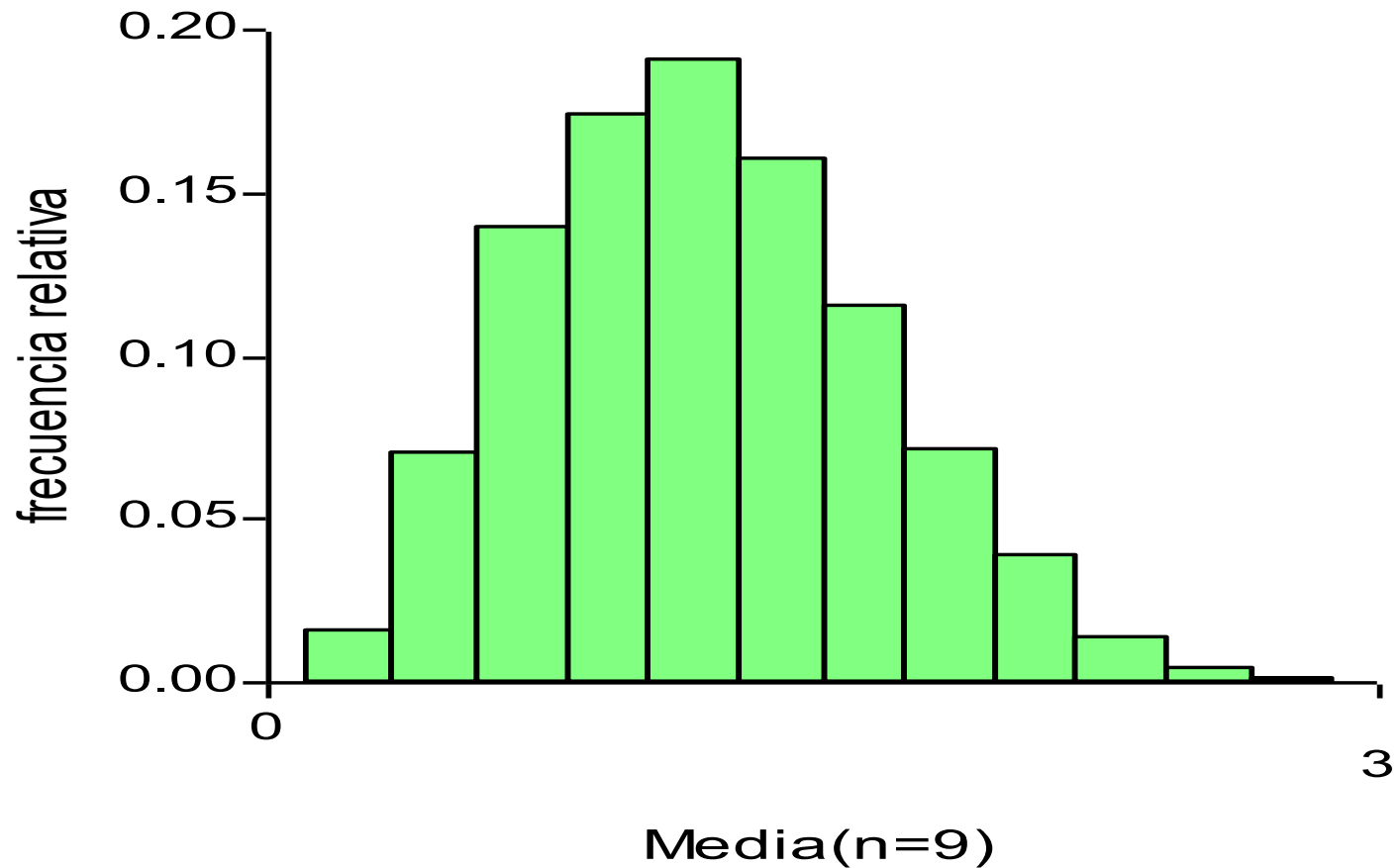


# Distribución de la media muestral



*¿Qué se observa al aumentar  $n$ ?*





***Teorema del límite central:*** la distribución de medias es aproximadamente normal, independientemente de la forma de la distribución que tengan las muestras seleccionadas, para tamaños de muestra suficientemente grandes ( $n > 30$ )

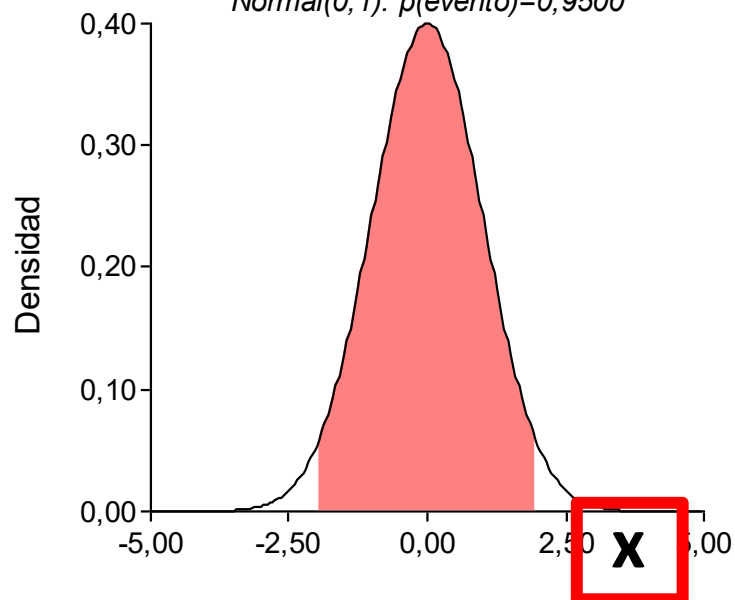
# Proceso de estandarización

$$X \sim N(\mu, \sigma)$$

$$Z = \frac{X - \mu}{\sigma}$$

*Función de densidad*

Normal(0, 1):  $p(\text{evento})=0,9500$



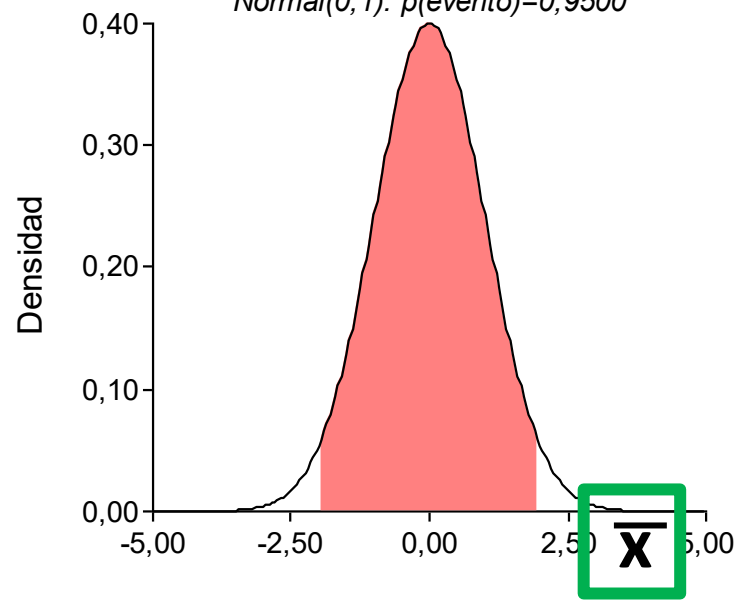
$$Z \sim N(0, 1)$$

$$\bar{X} \sim N\left(\mu; \sigma/\sqrt{n}\right)$$

$$z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

*Función de densidad*

Normal(0, 1):  $p(\text{evento})=0,9500$



# Inferencia estadística:

***Existen dos clases de experimentos:***

- ✓ *Experimentos diseñados para estimar algún parámetro o propiedad poblacional*

## **ESTIMACIÓN. INTERVALOS DE CONFIANZA**

- ✓ *Experimentos comparativos, donde dos o más tratamientos o condiciones experimentales se desean comparar. Permiten corroborar las hipótesis planteadas sobre parámetros poblacionales.*

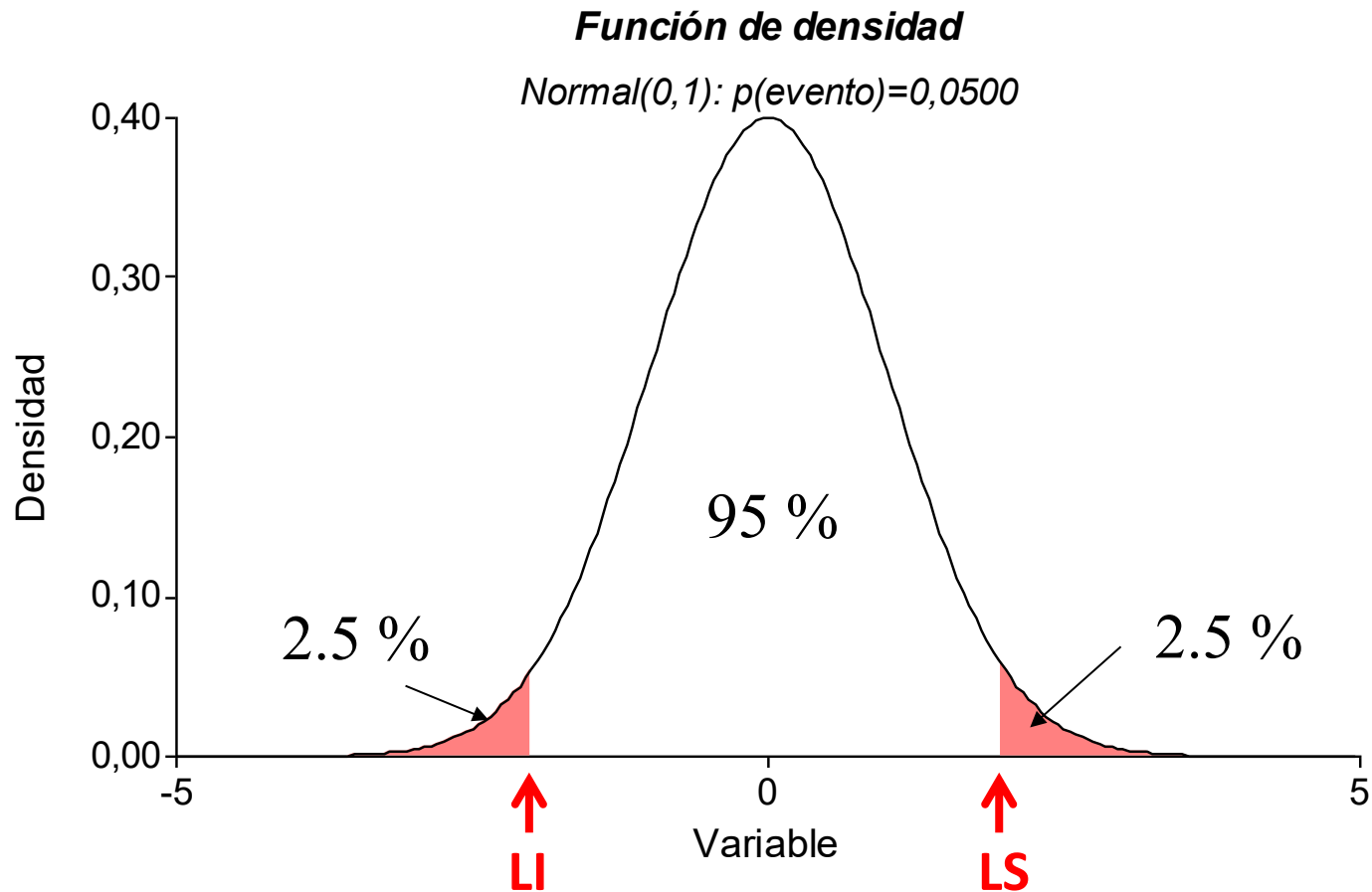
## **TEST DE HIPOTESIS**

# Estimación por Intervalos de confianza

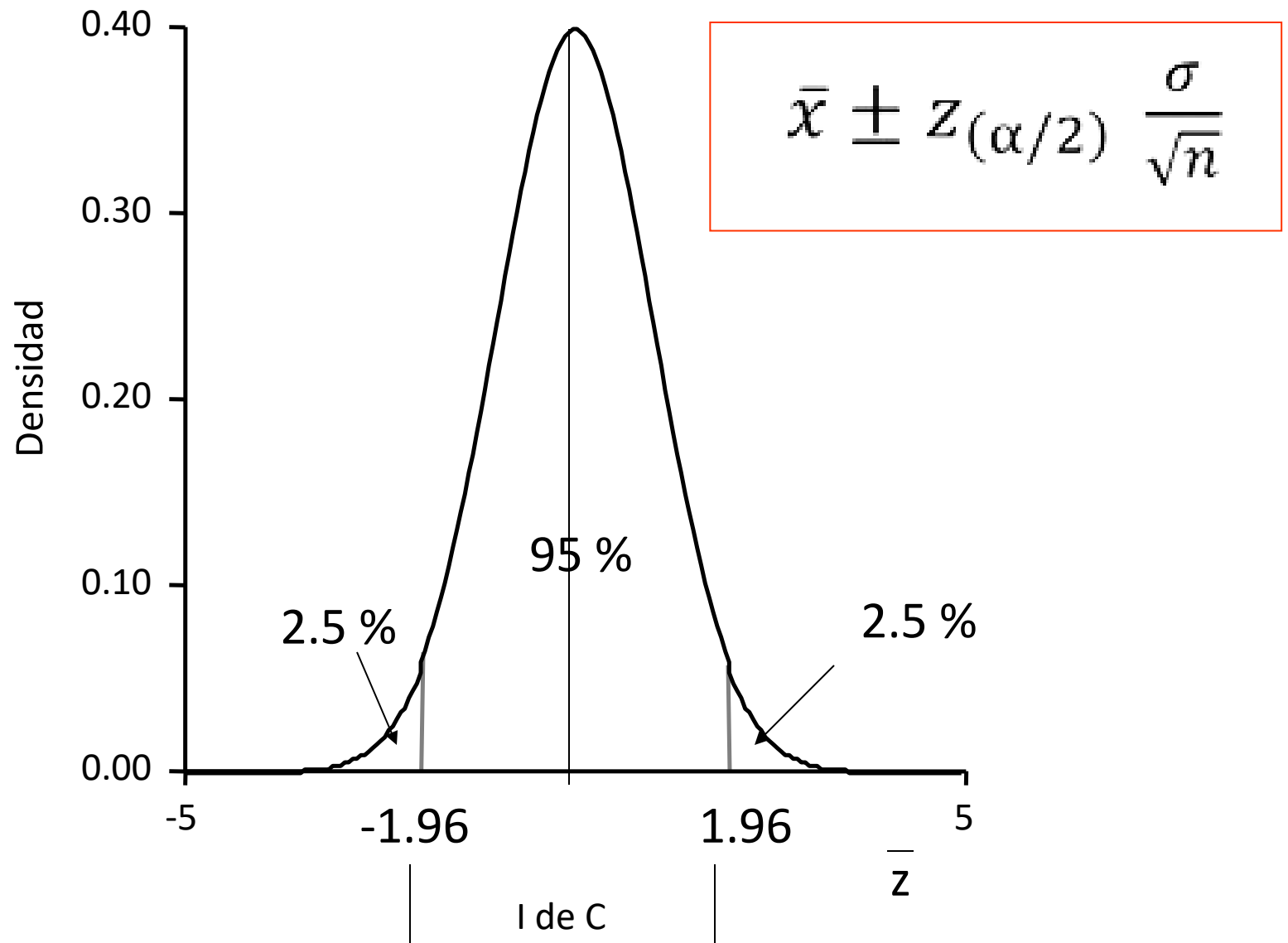
El IC esta definido por un limite inferior y un limite superior con una dada probabilidad

Para cada parámetro estadístico se calcula de una forma determinada

# Determinar el intervalo de confianza al 95% para $\mu$ de una muestra con distribución normal







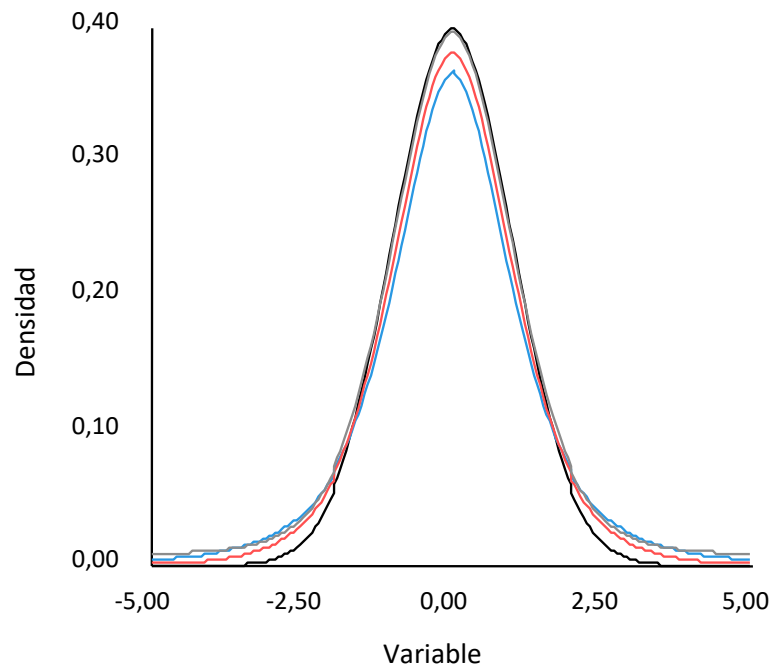
***Transformar  $\bar{z}$  en  $\bar{x}$  y establecer el intervalo***

Que ocurre cuando  $n < 30$  y  $\sigma$  es desconocido

**Distribucion Normal: Estadístico  $z$**  
$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

**Distribucion  $t$  : Estadístico  $t$**

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$



Normal (0,1)

T student (3gl)

T student (6gl)

T student (10gl)

$$IC_{95\%} = media \pm t_{crit(n-1 gl; \alpha=0,05)} EE$$

$$\text{Estimador} \pm \text{Estadístico} \times \text{EE}$$

Para  $\sigma$  conocido

$$\bar{x} \pm Z(\alpha/2) \frac{\sigma}{\sqrt{n}}$$

Para  $\sigma$  desconocido  
y muestras grandes

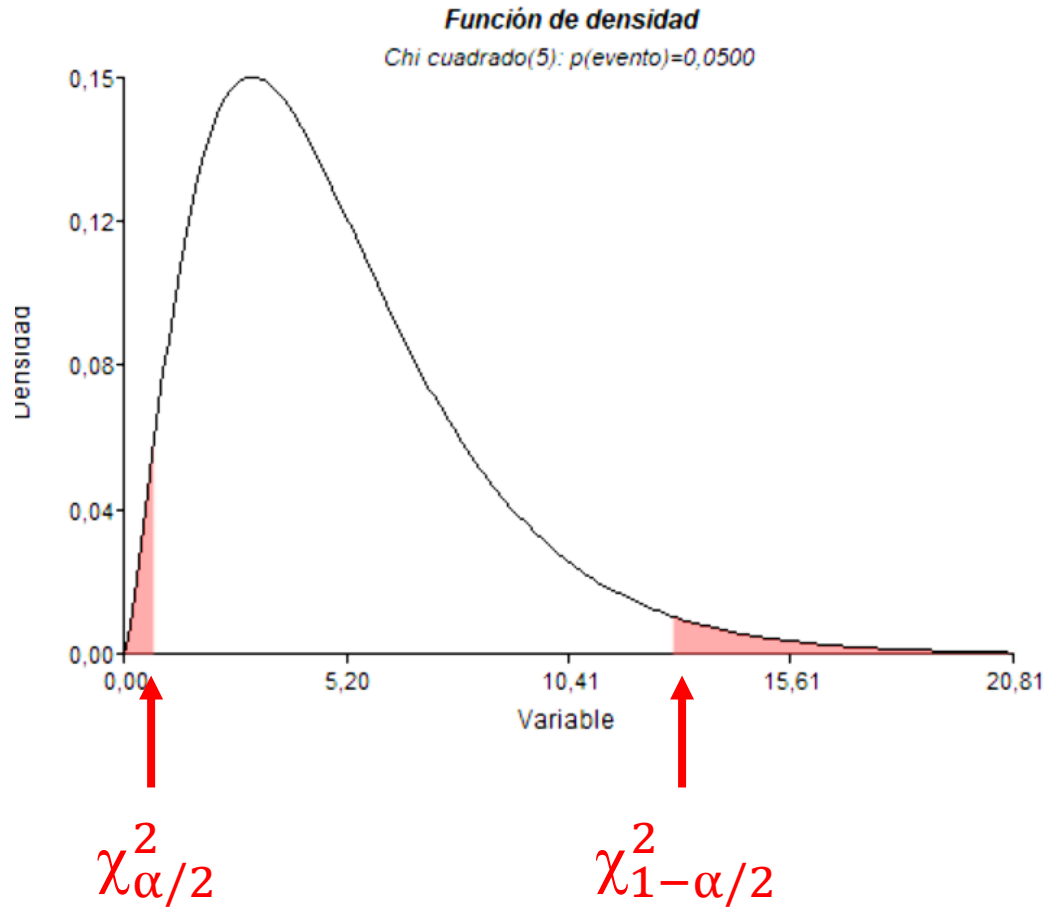
$$\bar{x} \pm Z(\alpha/2) \frac{s}{\sqrt{n}}$$

Para  $\sigma$  desconocido y  $n < 30$

$$\bar{x} \pm t_{(n-1), (\alpha/2)} \frac{s}{\sqrt{n}}$$

con  $n-1$  grados de libertad

# Estimación por Intervalos de confianza para la $\sigma^2$



$$LI = s^2 \frac{(n-1)}{\chi^2_{\frac{\alpha}{2}, (n-1)}}$$

$$LS = s^2 \frac{(n-1)}{\chi^2_{1-\frac{\alpha}{2}, (n-1)}}$$

***Se realizó un estudio caso-control para estudiar la incidencia de candidiasis en pacientes diagnosticados como positivos el virus de inmunodeficiencia humana (diagnosticada por el método de ELISA y confirmada a través de la prueba de Western Blot). Para este estudio se evaluó como variable el conteo de linfocitos y de la subpoblación CD4 (mm<sup>3</sup>). En la tabla además se registró la edad de los pacientes:***

Pacientes	HIV	Candidiasis	Leucocitos	CD4	Edad
1	pos	si	3500	400	48
2	pos	si	2500	100	36
3	pos	si	1500	120	29
4	pos	no	3000	300	50
5	pos	no	5000	250	81
6	pos	si	3400	99	23
7	pos	si	1250	130	37
8	neg	si	5500	280	43
9	neg	no	5800	560	67
10	neg	si	7900	490	78
11	neg	no	7500	390	49
12	neg	no	6500	550	63
13	neg	no	3780	520	82
14	neg	no	6300	430	87
15	neg	no	5400	480	66

*Calcular el IC para el conteo de linfocitos al 95% para pacientes HIV positivos y negativos*

## Intervalos de confianza

*Bilateral*

*Estimación paramétrica*

HIV	Variable	Parámetro	Estimación	E.E.	n	LI(95%)	LS(95%)
neg	Leucocitos	Media	6085,00	457,49	8	5003,21	7166,79
pos	Leucocitos	Media	2878,57	484,75	7	1692,43	4064,71

## Intervalos de confianza

*Bilateral*

*Estimación paramétrica*

HIV	Variable	Parámetro	Estimación	E.E.	n	LI(95%))	LS(95%))
neg	Leucocitos	Varianza	1674371,43	894989,17	8	731953,60	6935804,44
pos	Leucocitos	Varianza	1644880,95	949672,46	7	683025,11	7976184,92

# Test de hipótesis

**Procedimientos para aceptar o rechazar una hipótesis referida a un parámetro o característica poblacional**

*Se hacen proposiciones (hipotesis) sobre los parámetros poblacionales y se trabaja con muestras para concluir*

***La hipótesis que se somete a prueba se  
llama  $H_0$  o Hipótesis nula***

***La Hipótesis que se acepta si  $H_0$  es  
rechazada se denomina  $H_1$***



# Método o test paramétrico

## Variables continuas

- 1) *Plantear  $H_0$  y  $H_1$*
- 2) *Plantear el experimento a realizar*
- 3) *Definir el estadístico de contraste*
- 4) *Establecer el nivel de significación de la prueba*
- 5) *Establecer los criterios para rechazar  $H_0$  (rango estadístico de contraste)*
- 6) *Calcular el estadístico de contraste y compararlo con el valor crítico*

Se desea saber si el valor medio de colesterol de una cierta localidad del país **es el mismo** que la media en pacientes sanos, cuyo valor es 200 mg/dL ( $\sigma = 16$  mg/dL).

1) *Plantear  $H_0$  y  $H_1$*

$$H_0: \mu = 200 \text{ mg/dL}$$

$$H_1: \mu \neq 200 \text{ mg/dL}$$

$$H_1: \mu < 200 \text{ mg/dL}$$

$$H_1: \mu > 200 \text{ mg/dL}$$

## *2) Plantear el experimento a realizar*

*Se plantea tomar una muestra de 25 individuos de la localidad a evaluar para hacer la determinación de colesterol*

## *3) Definir el estadístico de contraste*

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

## *4) Establecer el nivel de significación de la prueba*

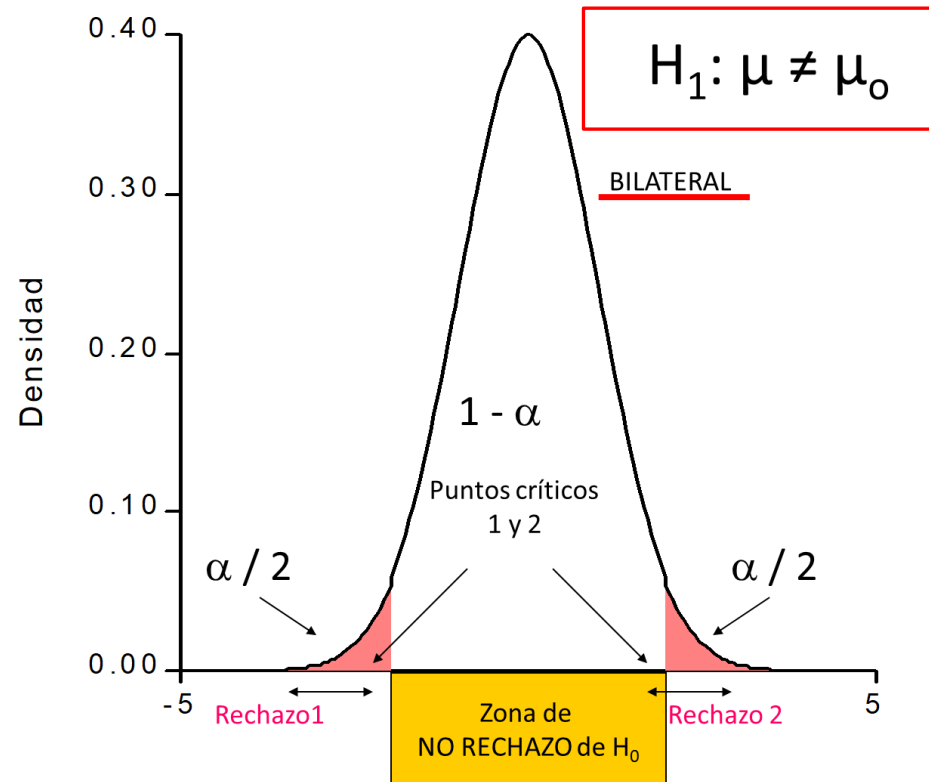
$$\alpha=0,05$$

## 5) Establecer los criterios para rechazar $H_0$ (rango estadístico de contraste)

Valores críticos  $\begin{cases} z_{\alpha/2} \\ -z_{\alpha/2} \end{cases}$

Si  $\alpha=0.05$

$-1,96 > z$   
o  
 $1,96 < z$   
Zona de rechazo



## 6) Calcular el estadístico de contraste y compararlo con el valor crítico

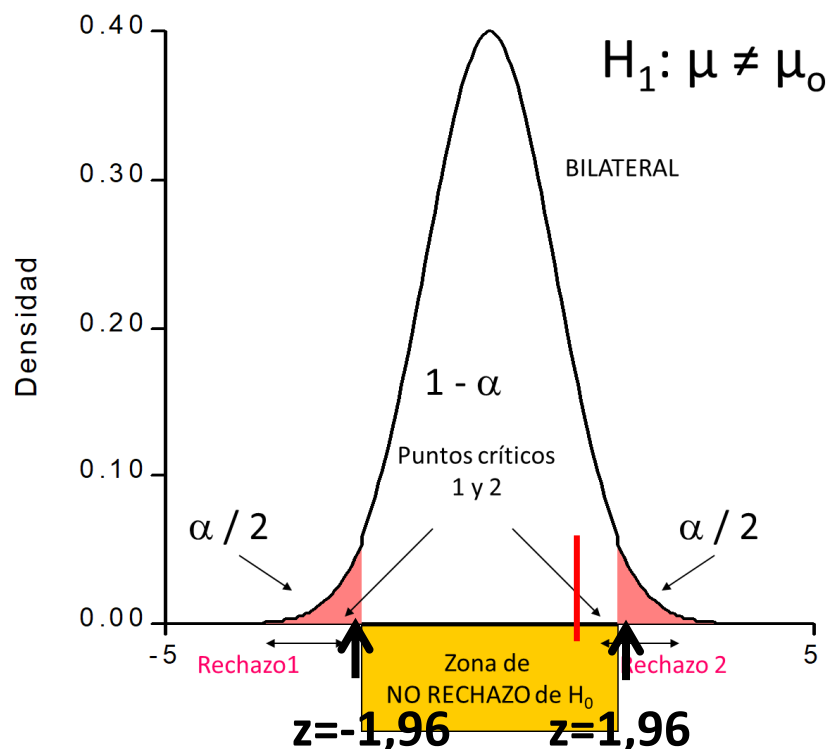
Se obtuvo un valor medio de colesterol de la muestra de 25 individuos de la localidad a evaluar de 204 mg/dL.

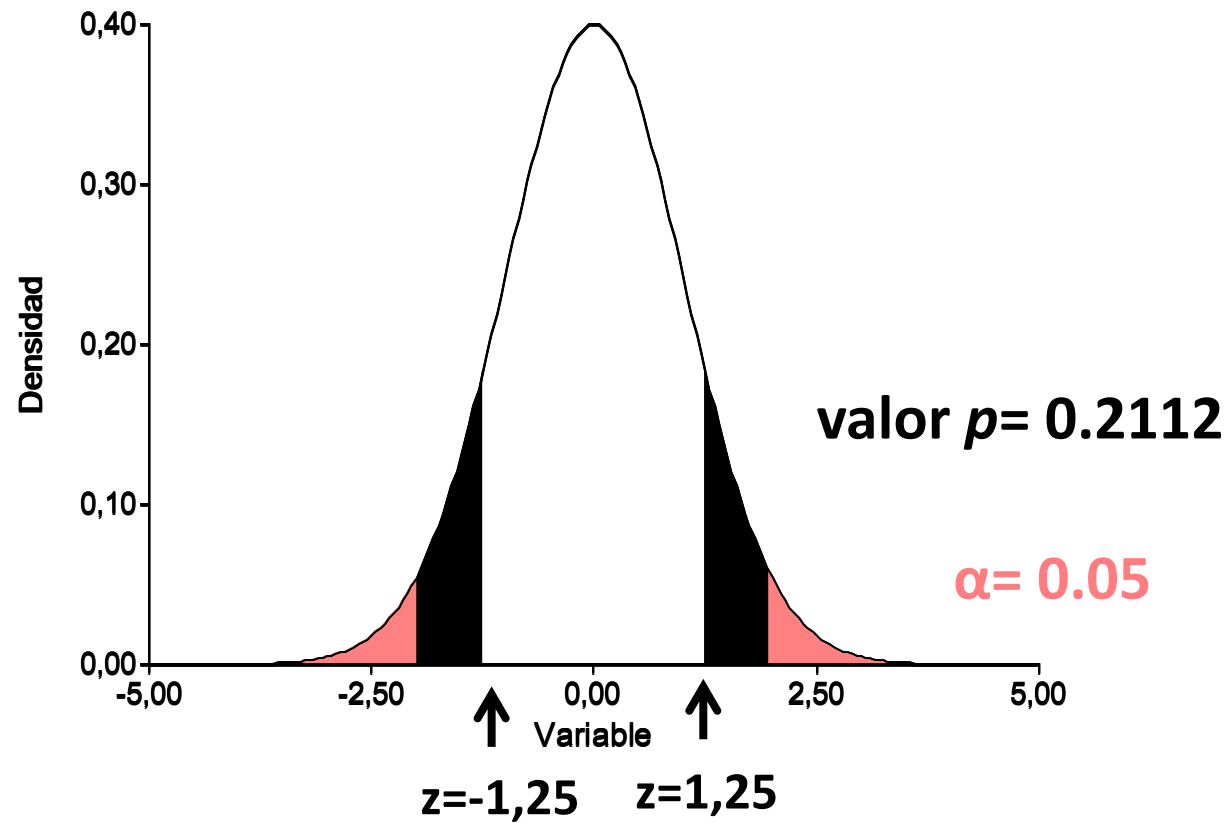
El valor con que se contrasta es 200 mg/dL de una población con  $\sigma = 16$  mg/dL.

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

$$Z = \frac{204 - 200}{16 / \sqrt{25}} = 1.25$$

**Criterio de rechazo:**  
 $-1,96 > z$  ó  $1,96 < z$





**Bilateral**

Si el valor observado del estadístico pertenece a  
la zona de rechazo: se rechaza  $H_0$

Si no pertenece a la zona de rechazo se dice:  
no tengo evidencia suficiente para rechazar  $H_0$

*Todos los tests estadísticos se basan en calcular un estadístico y compararlo con un valor teórico (valor crítico) que separa la zona de aceptación de  $H_0$  de la zona de rechazo.*

*Si el estadístico es mayor que el valor crítico se rechazará  $H_0$ . Esto es equivalente a decir que el  $p$  es menor que el nivel de significancia de la prueba ( normalmente 0,05)*