ESTADISTICA GENERAL

- PRINCIPALES DISTRIBUCIONES CONTINUAS
- Profesor: Celso Gonzales

OBJETIVOS

- **Describir las características de las distribuciones de probabilidad : Normal, Ji-cuadrado, t de student y F de snedecor.**
- **Calcular las probabilidades de las distribuciones de probabilidad: Normal, Ji-cuadrado, t de student y F de snedecor, utilizando tablas de estas distribuciones.**
- **Definir y elaborar una distribución de muestreo de medias muestrales.**
- **Explicar el teorema del limite central**

INTRODUCCIÓN

DISTRIBUCIONES CONTINUAS IMPORTANTES

DISTRIBUCIÓN NORMAL O DE GAUSS

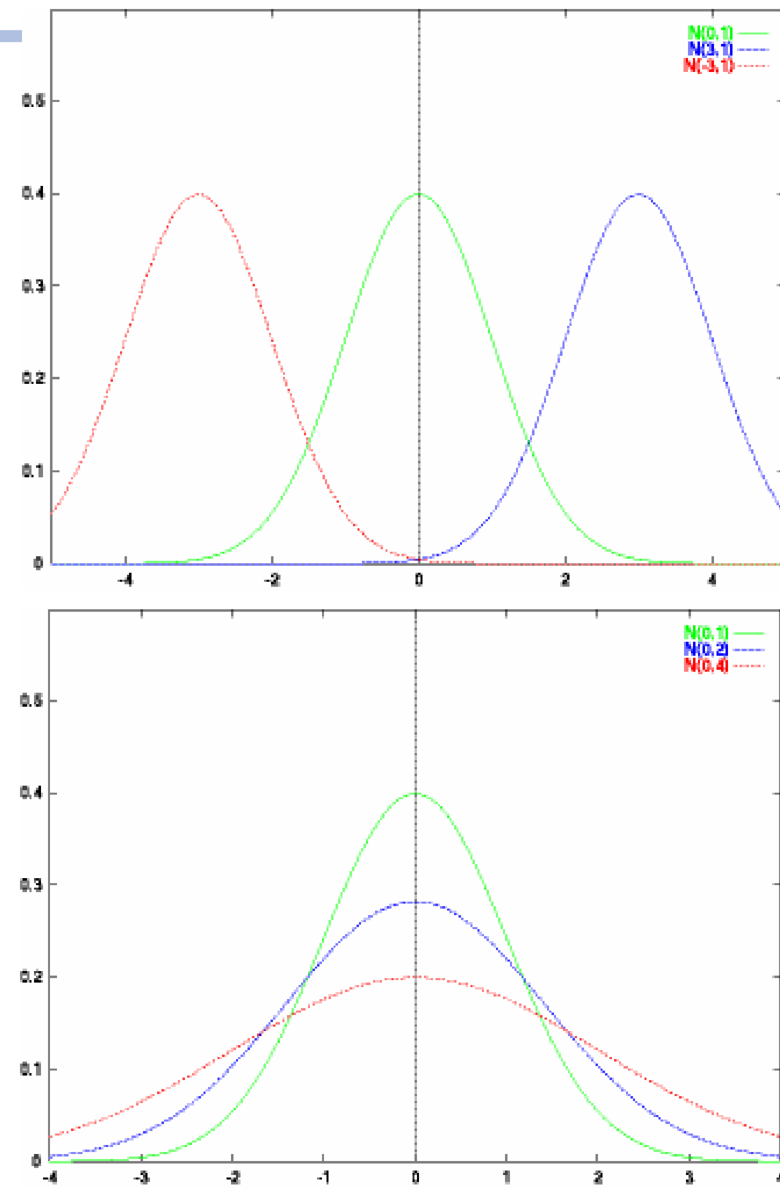
- Aparece de manera natural:
  - Errores de medida,Contenido neto, Diámetros, longitud,..
  - Distribuciones binomiales con n grande (n>30) y ‘p ni pequeño’ (np>5) ‘ni grande’ (nq>5).

- Está caracterizada por dos parámetros: La media, μ, y la desviación estandar, σ.

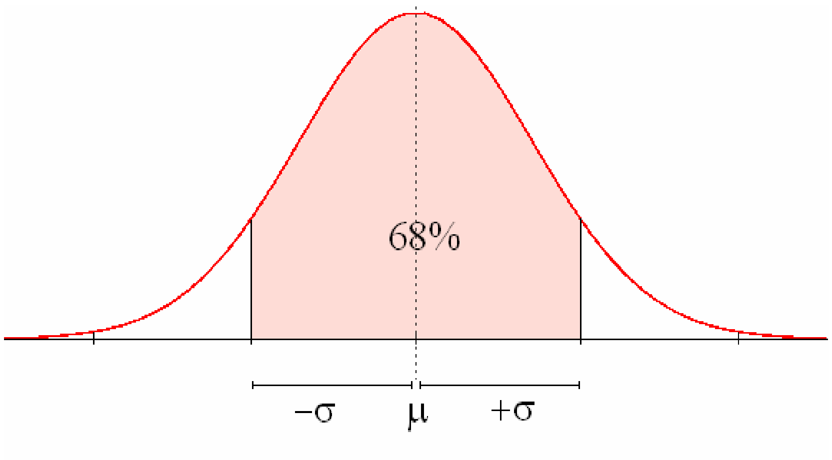
- Su función de densidad es:

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

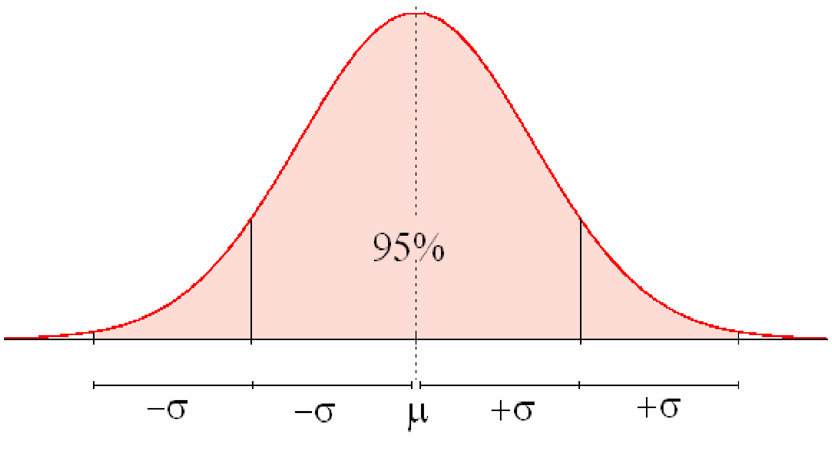
N(μ, σ): Interpretación geométrica



- Entre la media y una desviación estándar a tenemos siempre la misma probabilidad: aprox. 68%



- Entre la media y dos desviaciones estándar aprox. 95%



CARACTERÍSTICAS

- Es de forma acampanada.
- Es simétrica, mesocúrtica y unimodal.
- Es asintótica
- Los puntos de inflexión de la función de densidad están a distancia **σ** de μ.
- Si tomamos intervalos centrados en μ, y cuyos extremos están…
  - a distancia **σ**, → tenemos probabilidad **68%**
  - a distancia **2 σ**, → tenemos probabilidad **95%**
  - a distancia **2'5 σ** → tenemos probabilidad **99%**

DISTRIBUCIÓN NORMAL ESTANDAR

Todas las distribuciones normales N(μ, σ), pueden ponerse mediante una traslación μ, y un cambio de escala σ, como **N(0,1)**. Esta distribución especial se llama **normal estándar**.

Es decir:

$$z = \frac{x - \mu}{\sigma}$$

USO DE LA TABLA

Ejemplos

1. Una lámpara tiene intensidad luminosa según una distribución normal de media 3000 candelas y desviación estándar 50 candelas. Encuentre el limite inferior de especificación para que solo el 0,5 % de las lámparas estén por debajo de dicho limite.

2. Una característica de calidad de un proceso de manufactura que se mide está distribuida normalmente con media 40 y desviación estándar de 13 y los límites de especificación son 41 ± 5. Suponiendo que se puede retrabajar un articulo que excede el límite superior de especificación y que se debe rechazar tal artículo si está por debajo del límite inferior de especificación, ¿que porcentaje de artículos para retrabajo y rechazo origina el proceso?

DISTRIBUCION DE LA MEDIA MUESTRAL

Objetivos:

Al finalizar esta unidad el estudiante será capaz:

- **Definir y elaborar una distribución de muestreo para la media de la muestra.**
- **Explicar el teorema del limite central.**
- **Definir y elaborar una distribución de muestreo para la diferencia medias muestrales.**
- **Definir y elaborar una distribución de muestreo para la proporción muestral.**

MUESTREO

Proceso que nos permite la extracción de una muestra a partir de una población.

Hay dos tipos básicos de muestreo:

Muestreo Probabilistico

Muestreo No probabilistico.

VALOR ESTADISTICO:

Es cualquier cantidad cuyo valor se puede calcular a partir de datos muestrales.

MUESTRA ALEATORIA

Las variables aleatorias X_1, …..X_n forman una muestra aleatoria si:

1. independientes.
2. X_i tiene la misma distribución de probabilidad

DISTRIBUCION DE LA MEDIA MUESTRAL

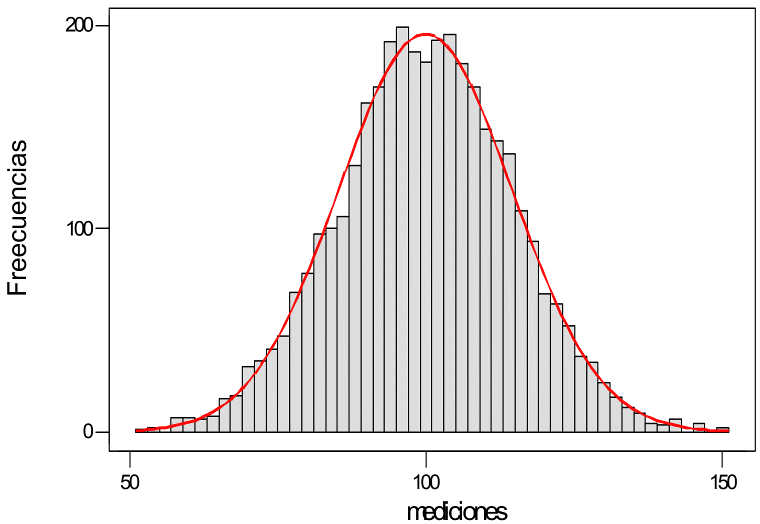
Distribución de probabilidad de todas las medias posibles de un tamaño muestra dado.

Muestreo	Media	Varianza	Error estándar
Con reemplazo	$\mu_{\bar{x}} = \mu$	$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$	$\sigma_{\bar{x}} = \sqrt{\frac{\sigma^2}{n}}$
Sin reemplazo	$\mu_{\bar{x}} = \mu$	$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$	$\sigma_{\bar{x}} = \sqrt{\frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)}$

$$Z = \frac{\bar{x} - E\left(\bar{x}\right)}{\sqrt{v\left(\bar{x}\right)}}$$

EJEMPLO

Distribución muestral de la media.



Distribución de la población (dist. Normal): N =3600

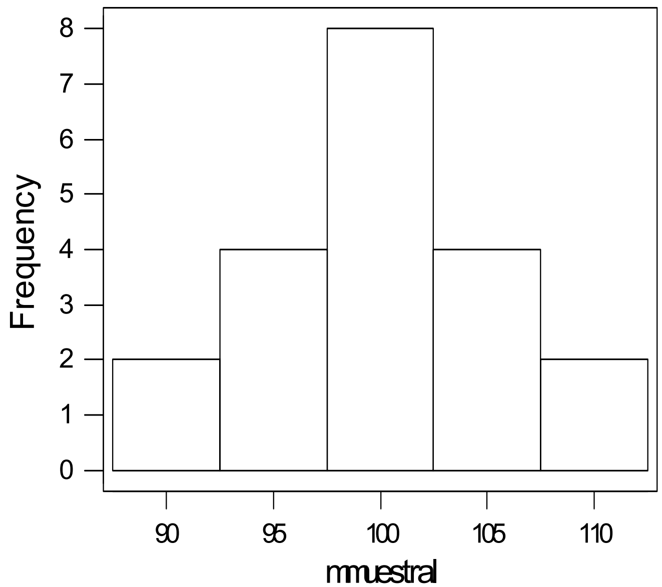
Media=100

Varianza=225

Desv.estándar=15

EJEMPLO

Distribución muestral de la media.



Muestras	Promedio
1	109.15
2	94.27
3	101.59
4	99.86
5	103.86
6	92.34
7	108.54
8	95.16
9	104.07
10	103.60
11	92.41
12	97.18
13	98.40
14	101.36
15	98.63
16	98.24
17	102.89
18	97.90
19	100.33
20	97.08

Distribución muestral de la media:

Tamaño de muestra=10

Media=100

(Varianza=22,4375)

Error estandar = $\sqrt{22.4375} = 4.74$

EJEMPLO

Un equipo de empacado de un proceso de fabricación rellena cajas de cereal de 368 g. de tal forma que la cantidad de cereal por caja tiene una distribución normal con una media de 368 g. y una desviación estándar de 15 g. Si se selecciona una muestra de 16 cajas de las miles que se rellenan cada día y se calcula el peso promedio, ¿cuál es la probabilidad que esté entre 365 g. y 368 g.?

EJEMPLO

La duración en horas de un electrodoméstico comprado a la empresa A, es una variable aleatoria $N\left(1200, \sigma_A\right)$. El 95% de los electrodomésticos duran entre 1180 y 1220 horas. Si se extraen 200 m.a.s de 7 electrodomésticos cada una:

a. ¿Calcular la desviación estándar de la duración de electrodomésticos?

b. ¿cuál es la media, error estándar de la distribución de la media muestral?

c. ¿Qué probabilidad existe de que la media muestral supere las 1200 horas?

d. ¿Cuántas muestras superan las 1200 horas?

TEOREMA DEL LÍMITE CENTRAL

Dada una v.a. **cualquiera**, si extraemos muestras de tamaño n, y calculamos los promedios muestrales, entonces:

- dichos promedios tienen distribución aproximadamente normal;
- Las aproximaciones anteriores se hacen exactas cuando n tiende a infinito.
- Este teorema justifica la importancia de la distribución normal.

EJEMPLO

Sea X una variable aleatoria que representa el contenido real en onzas de una lata de café de 1 lb. La distribución de probabilidad de x es:

$$
\begin{aligned}
f(x) &= \frac{1}{1,5} \qquad 15,5 \le x \le 17 \\
&= 0 \qquad \textit{para otros valores}
\end{aligned}
$$

A. Si se toman muestras aleatorias de 36 latas de café y se mide el contenido real promedio.

B. Hallar la media y error estándar de la distribución de la media muestral.

C. Si se toma una muestra aleatoria de 36 latas, ¿ calcular la probabilidad de que el contenido promedio de café en la muestra contenga menos de 16 onzas?

DISTRIBUCION DE UNA DIFERENCIA DE MEDIAS MUESTRALES

$$z = \frac{\left(\bar{x}_1 - \bar{x}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

EJEMPLO

Un proceso envasa conservas con dos maquinas (A y B). Los pesos de las conservas envasados tanto por la maquina A y B siguen una distribución normal. La Maquina A envasa conservas con peso medio de 970 g y una varianza de 144 g^2, mientras que la maquina B los envasa con un peso medio de 980 g y una varianza de 256 g^2

a. Si se eligen aleoriamente 36 conservas de cada maquina, ¿Cuál es la probabilidad de que exactamente uno de ellos tenga un peso promedio muestral mayor de 965 g.?

b. Si se eligen 32 artículos por la máquina A y 48 artículos producidos por la máquina B. Hallar la probabilidad de que el peso promedio miuestral de B sea mayor que el promedio muestral de la maquina A en por lo menos 8 g?

DISTRIBUCIÓN DE LA PROPORCIÓN MUESTRAL.

Distribución de probabilidad de todas las proporciones posibles de un tamaño muestra dado.

Muestreo	Media	Varianza	Error estándar
Con reemplazo	$\mu_p = \pi$	$\sigma_p^2 = \frac{\pi(1-\pi)}{n}$	$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}}$
Sin reemplazo	$\mu_p = \pi$	$\sigma_p^2 = \frac{\pi(1-\pi)}{n}\left(\frac{N-n}{N-1}\right)$	$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}\left(\frac{N-n}{N-1}\right)}$

A: Muestreo con reemplazo

- P: proporción de éxitos en la muestra

$$f(p) = \binom{n}{np} (\pi)^{np} (1-\pi)^{n-np}$$

B: Muestreo sin reemplazo

- P: proporción de éxitos en la muestra

$$g(p) = \frac{\binom{A}{np}\binom{B}{n-np}}{\binom{N}{n}}$$

Aproximación a la Distribución normal

$$Z = \frac{p - \pi}{\sqrt{\dfrac{\pi(1-\pi)}{n}}}$$

EJEMPLO

Cada media hora se saca una muestra aleatoria de 100 unidades de un proceso de producción. La proporción de productos no conforme fabricado es 0,03.

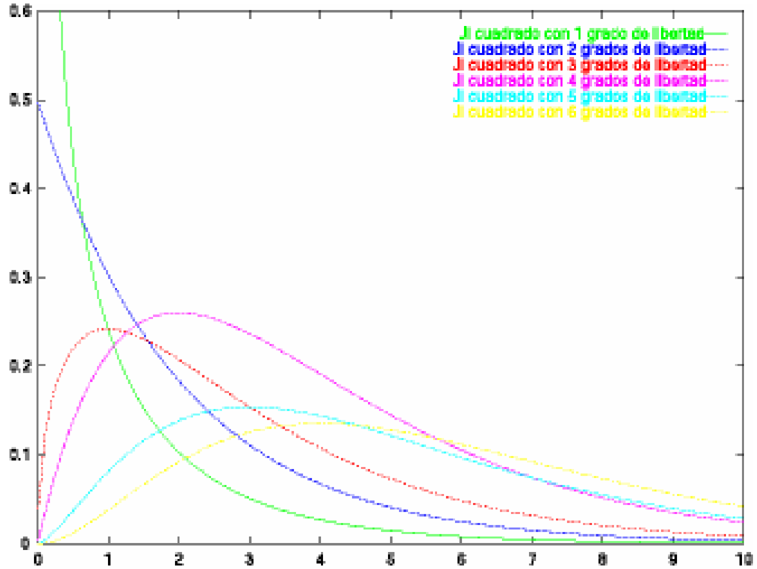
- ¿Cuál es la probabilidad de que p(proporción muestral de productos no conformes) es a lo mas 4 % ?
- Si la muestra que se saca cada media hora es de tamaño 10 ¿ cuál sería la distribución exacta de la proporción muestral de productos no conformes?. Así mismo: a media, error estándar y coeficiente de variación de la proporción muestral de productos no conformes.

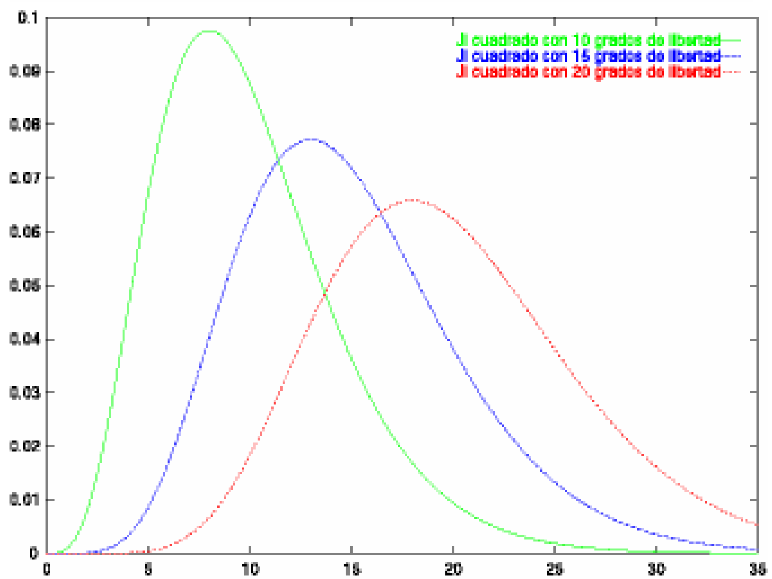
CHI CUADRADO

- Tiene un sólo parámetro denominado grados de libertad.

- La función de densidad es asimétrica positiva.
- El rango de la variable considera sólo los valores positivos.

- La función de densidad se hace más simétrica incluso casi gausiana cuando aumenta el número de grados de libertad.
- Normalmente consideraremos anómalos aquellos valores de la variable de la "cola de la derecha".
- Si $X \sim \chi^2_{(m)}$, entonces: E(X) =m V(X)= 2m





USO DE LA TABLA

GL	0,010	0,020	0,025		0,980	0,990
1	0,000	0,001	0,001		5,412	6,635
2	0,020	0,040	0,103		7,824	9,210
3	0,115	0,185	0,216		9,837	11,345
10	2,558	3,059	3,247		21,161	23,209
11	3,053	3,609	3,816		22,618	24,725

P(Y<3,247)=0.025

P(Y<9,837)=0.980

DISTRIBUCION DE LA VARIABLE ALEATORIA S^2

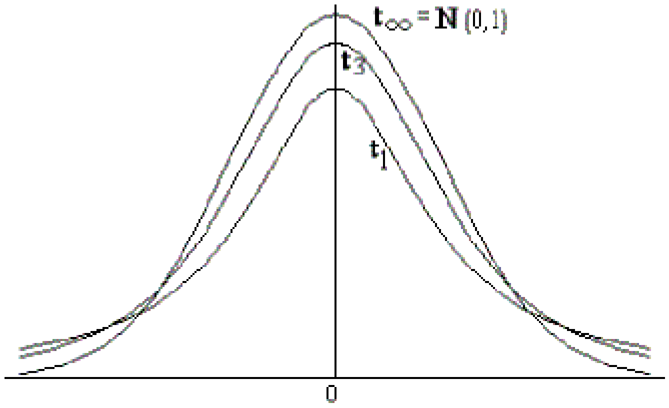
Sea $X_1,\ldots,X_n$ una muestra aleatoria de una distribución normal con media μ y desviación estandar σ, entonces la v.a:

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

T DE STUDENT

- Tiene un parámetro denominado grados de libertad.
- Cuando aumentan los grados de libertad, más se acerca a N(0,1).
- Es simétrica con respecto al cero.
- Se consideran valores anómalos los que se alejan de cero (positivos o negativos).
- Si X~t(m), entonces: E(X)=0 y

$$V(X)=\frac{m}{m-2}$$



TEOREMA

Si las v.as Z~N(0,1) y V~ $\chi^2_{(m)}$ son independientes, entonces la v.a X

$$X = \frac{Z}{\sqrt{\dfrac{V}{m}}} \sim t_{(m)}$$

USO DE LA TABLA

GL	0.005	0.01	0.025		0.8	0.9
1	-63.657	-31.821	-12.706		1.376	3.078
2	-9.925	-6.965	-4.303		1.061	1.886
3	-5.841	-4.541	-3.182		0.978	1.638
10	-3.169	-2.764	-2.228		0.879	1.372
11	-3.106	-2.718	-2.201		0.876	1.363

$$P(t_{10}<-2.228)=0.025$$

$$P(t_3<0.978)=0.8$$

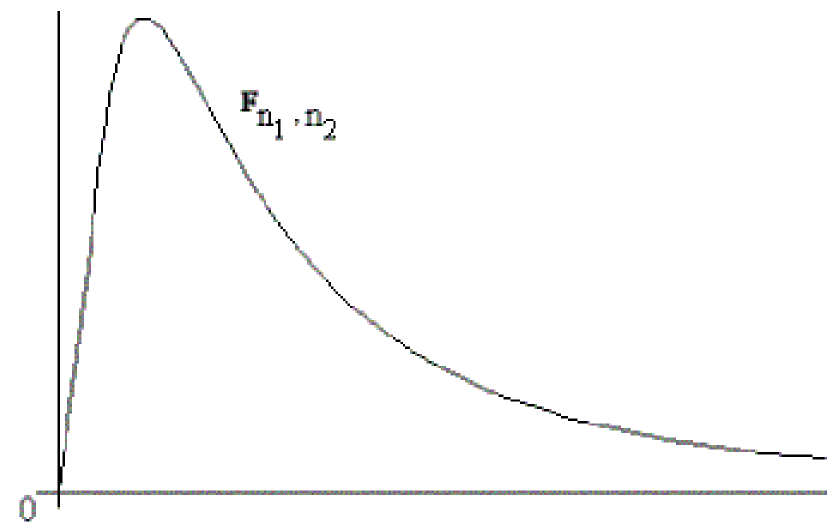
DISTRIBUCION DE LA MEDIA MUESTRAL $\overline{X}$

Sea $X_1,\ldots,X_n$ una muestra aleatoria de una distribución normal con media μ y desviación estandar σ, entonces la v.a:

$$Y = \frac{\left(\overline{X} - \mu\right)}{\frac{S}{\sqrt{n}}} \sim t_{(n-1)}$$

F DE SNEDECOR

- Tiene dos parámetros denominados grados de libertad.
- Sólo toma valores positivos. Es asimétrica.
- Es asintótica respecto al eje horizontal en su parte positiva.
- Normalmente se consideran valores anómalos los de la cola de la derecha.
- Las distribuciones F(n,m) tienden a ser simétricas cuando n y m son suficientes grandes.



USO DE LA TABLA

G.L del Numerador

GLD	..	6	7	..	19	20
1		234,0	236,8		247,7	248,0
		5858,9	5928,3		6200	6208,7
..						
7		3,87	3,79		3,45	3,45
		7,19	6,99		6,18	6,16
..						
11		3,09	3,01		2,56	2,54
		5,07	4,89		3,88	3,86

P(F<6,99)=0.99

P(F<3,45)=0.95

DISTRIBUCION DE LA RAZON DE VARIANZAS MUESTRALES $\frac{S_1^2}{S_2^2}$

Sea $X_1,\ldots,X_n$ una m.a de una distribución normal con varianza σ^2_1,y $Y_1,\ldots,Y_n$ otra m.a(independiente de las X_i) de una distribución normal con varianza σ^2_2 entonces la v.a:

$$\frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \sim F_{(n-1,m-1)}$$