TITULACIÓN

LICENCIATURA EN A.D.E.

TÉCNICAS ESTADÍSTICAS DE CONTROL DE CALIDAD (12249)

Mª Isabel López Rodríguez
Dpto. Economía Aplicada

CURSO ACADÉMICO 2013/2014

TEMA 2: CONCEPTOS ESTADÍSTICOS BÁSICOS

2.1. INTRODUCCIÓN

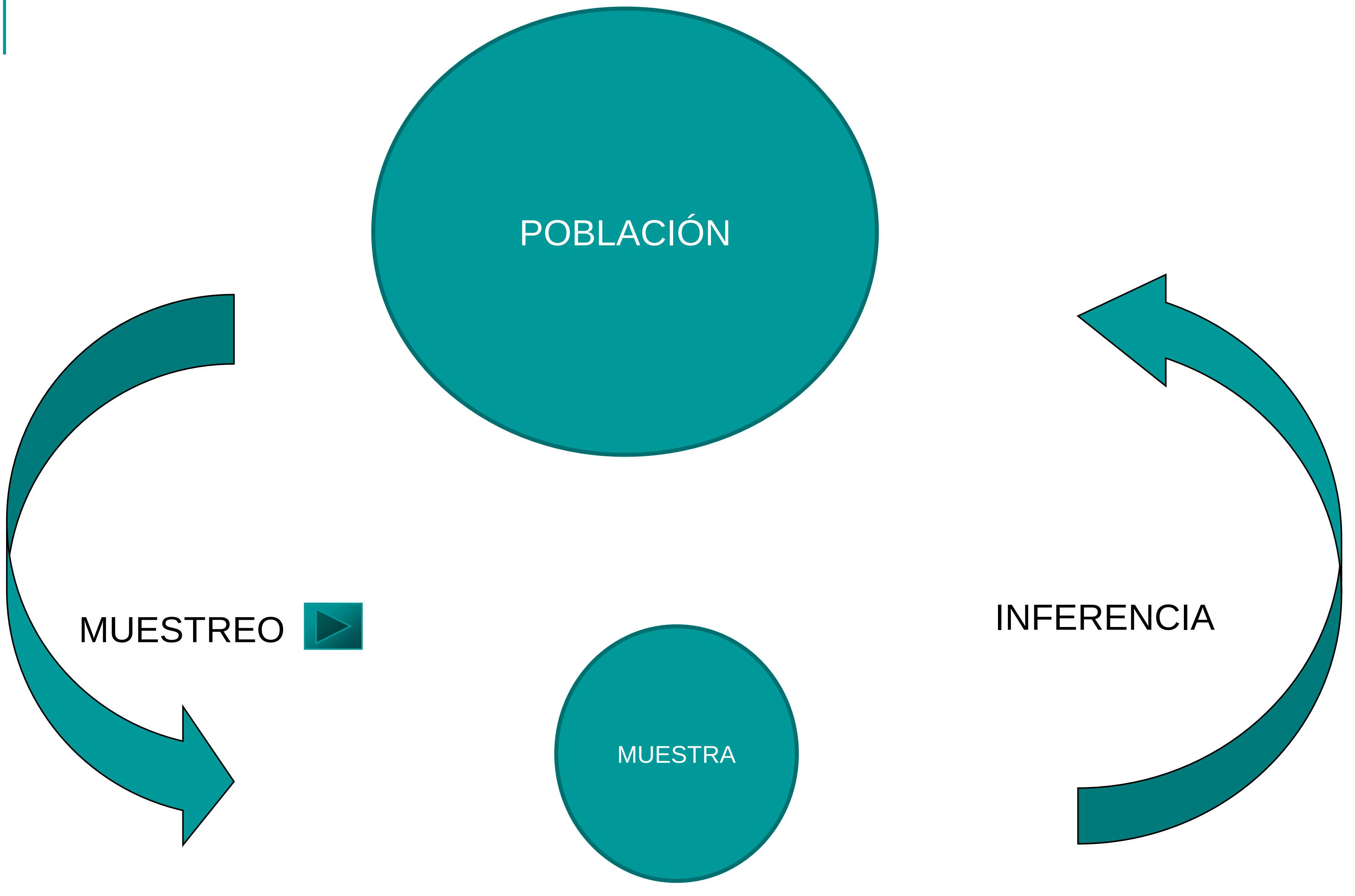
2.2. ESTADÍSTICA DESCRIPTIVA

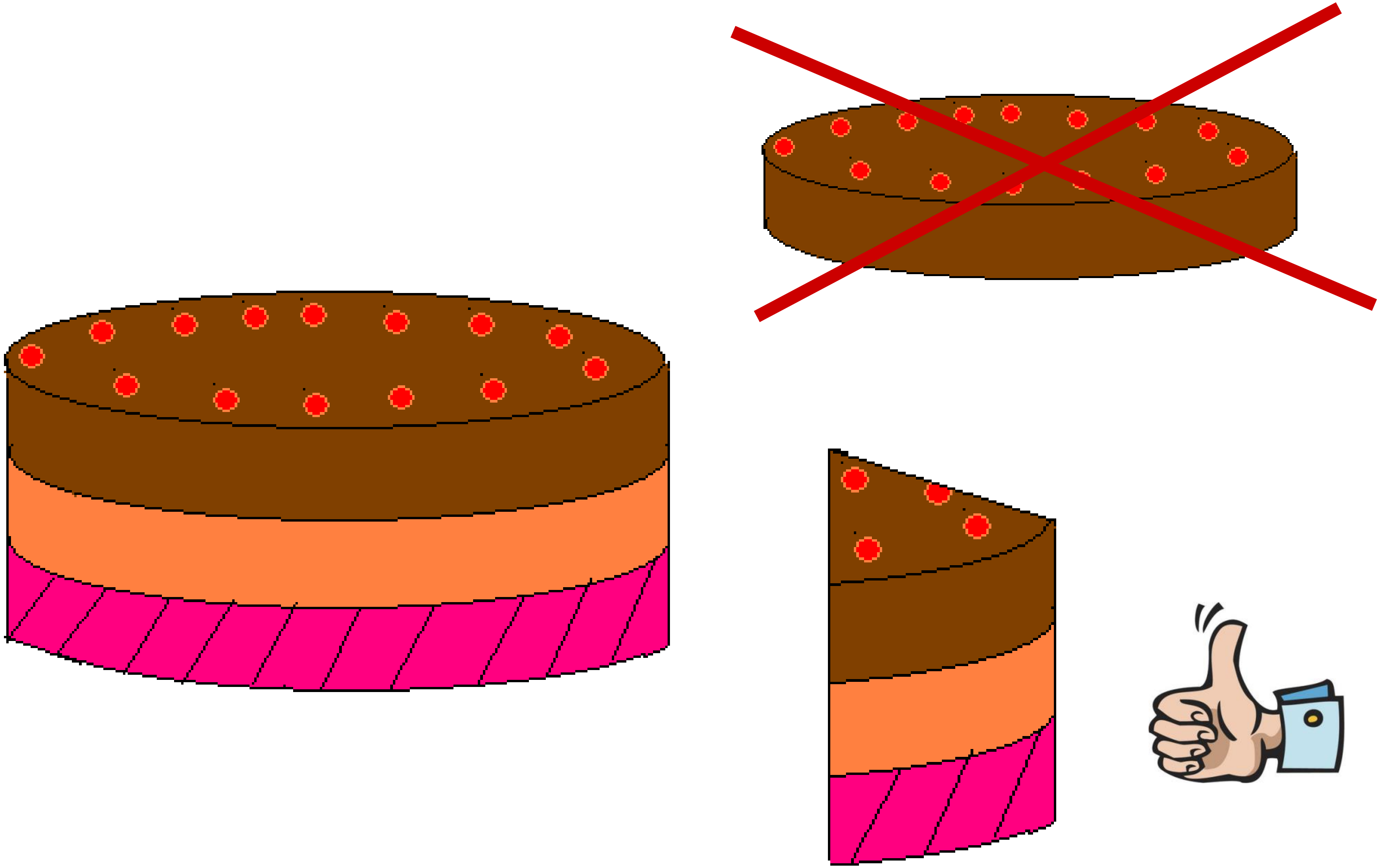
2.3. CÁLCULO DE PROBABILIDADES

2.4. INFERENCIA ESTADÍSTICA

2.1. INTRODUCCIÓN

ESTADÍSTICA DESCRIPTIVA → TEORÍA O CÁLCULO DE PROBABILIDADES → INFERENCIA ESTADÍSTICA


POBLACIÓN
MUESTREO
INFERENCIA
MUESTRA



2. 2. ESTADÍSTICA DESCRIPTIVA

OBJETIVO: analizar los datos, de modo que las conclusiones a las que se llegue no sobrepasen los límites de éstos. No se aplican a la población.

- Características que pueden ser observadas en un conjunto de elementos
  - Características medibles: Variables
    - Variables discretas: número de artículos vendidos en un día, nº. piezas defectuosas fabricadas en un mes, etc.
    - Variables continuas: longitud de unos tornillos, espesor planchas de acero, etc.
  - Características no medibles
    - Atributos: tipos de defectos en la fabricación de un artículo, modelos producidos de un automóvil, etc.
    - Ordinales: satisfacción de los clientes en el servicio postventa, clasificación de los proveedores de un producto, etc.

ESTADÍSTICA DESCRIPTIVA

Variables estadísticas

-DISTRIBUCIÓN DE FRECUENCIAS

-MEDIDAS DE REDUCCIÓN DE DATOS

- MEDIDAS DE POSICIÓN: media aritmética $(\overline{x})$
- MEDIDAS DE DISPERSIÓN: varianza (**S^2**), desviación típica (**S**), recorrido

ESTADÍSTICA DESCRIPTIVA

Ejemplo 1

Se ha preguntado a un grupo de 200 habitantes de la ciudad de Valencia el número de veces que cogen el transporte publico en un día laborable. Las respuestas se recogen en la siguiente tabla

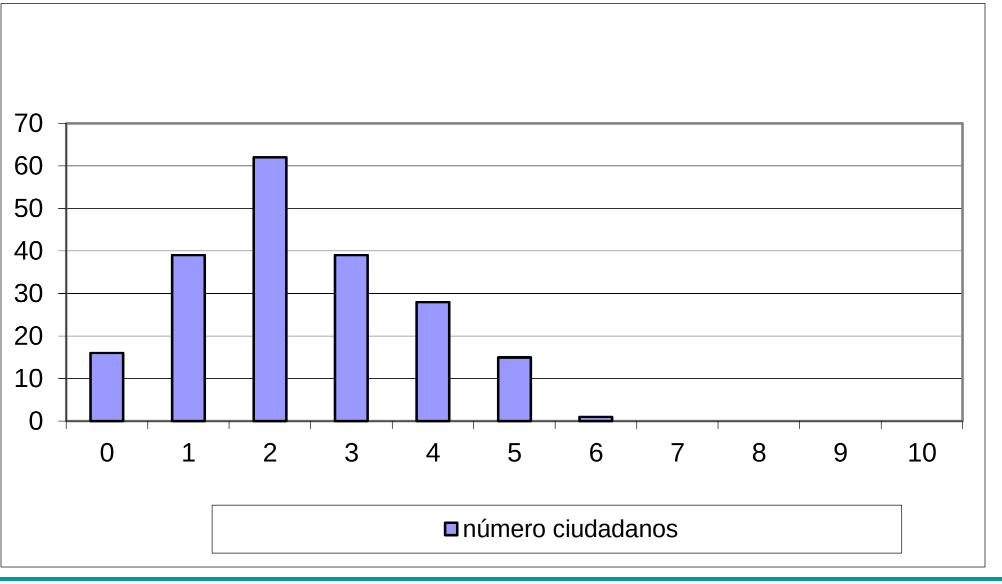
0	2	3	1	4	2	0	2	0	1
4	2	1	3	2	0	0	2	5	1
3	2	4	3	2	1	1	1	5	1
0	0	1	1	2	2	0	1	2	3
4	0	2	0	1	3	2	2	1	1
1	0	2	3	4	5	3	3	2	1
5	2	3	3	1	1	0	0	2	1
4	2	5	4	3	2	2	1	0	1
6	2	3	4	5	4	2	1	2	3
5	2	1	3	3	4	4	1	2	0
3	2	1	5	4	2	1	3	1	2
5	4	3	2	1	3	2	1	3	2
2	2	4	3	3	1	2	3	2	1
2	2	4	2	2	5	3	2	2	4
3	2	3	2	2	4	1	1	2	1
2	5	2	3	2	3	4	4	2	3
4	4	3	3	2	2	1	2	2	3
3	2	4	5	3	5	3	4	4	1
2	4	2	5	2	3	1	2	4	1
5	3	2	0	1	2	3	4	2	4

Ejemplo 2

Tabla de frecuencias absolutas

nº veces al día que cogen autobús Xi	nº Ciudadanos ni frecuencia absoluta
0	16
1	39
2	62
3	39
4	28
5	15
6	1
	200

Diagrama de barras



ESTADÍSTICA DESCRIPTIVA

Ejemplo 3

Un segundo estudio se realiza en la ciudad de Benidorm. Las respuestas de las 200 personas entrevistadas a la pregunta de cuantas veces cogen el autobús se recoge en la siguiente tabla.

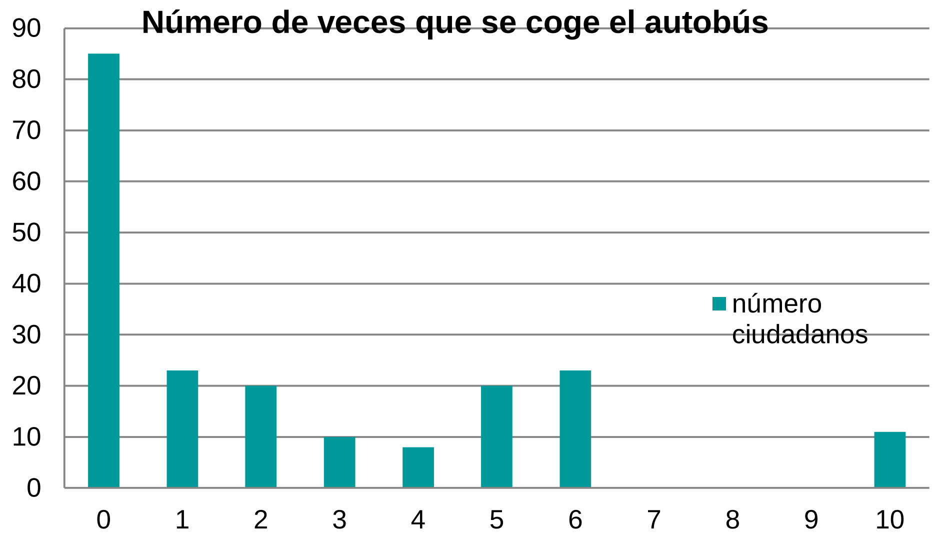
2	0	0	0	5	5	5	0	0	0
4	0	1	10	10	0	0	0	5	0
3	0	6	5	0	0	0	0	6	1
0	0	1	1	6	10	0	1	1	0
4	0	0	0	1	0	2	0	1	1
1	0	2	3	4	5	0	10	0	1
6	2	5	3	1	1	0	0	0	0
6	2	5	6	5	2	2	1	5	0
10	2	3	6	2	6	2	1	4	5
5	0	0	5	3	6	6	1	0	0
3	2	0	5	4	0	0	0	0	0
0	6	10	2	1	0	2	0	0	0
2	2	4	5	5	1	0	0	0	0
2	2	4	0	1	0	0	0	0	6
3	2	0	0	0	6	0	1	0	1
2	0	0	0	0	0	5	5	0	10
4	6	3	0	0	0	1	0	0	5
3	0	6	6	10	6	0	6	6	0
0	6	0	0	0	2	1	10	6	0
5	3	0	0	0	0	10	6	10	6

Tabla de frecuencias absolutas

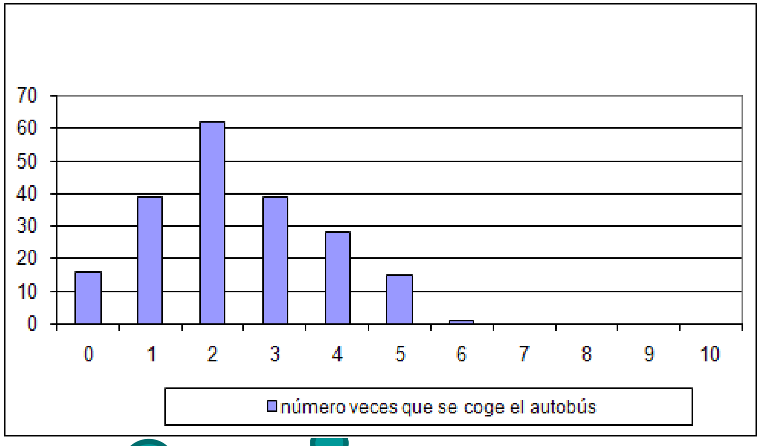
Ejemplo 3

nº veces al día que cogen autobús Yi	nº Ciudadanos ni frecuencia absoluta
0	85
1	23
2	20
3	10
4	8
5	20
6	23
10	11
	200

Diagrama de barras



Ejemplo 2

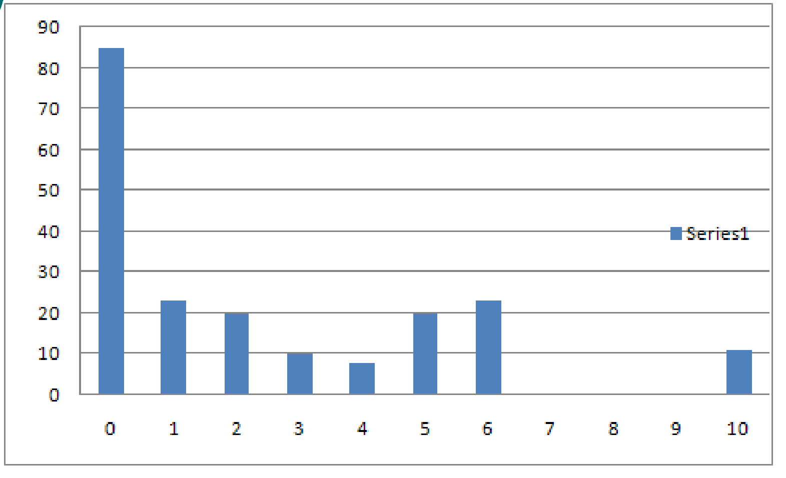


$$\overline{X} = \frac{\sum_{i=1}^{200} X_i}{200} = 2{,}365 = \text{PROMEDIO}(\quad)$$

Distinta distribución

Misma media

Ejemplo 3



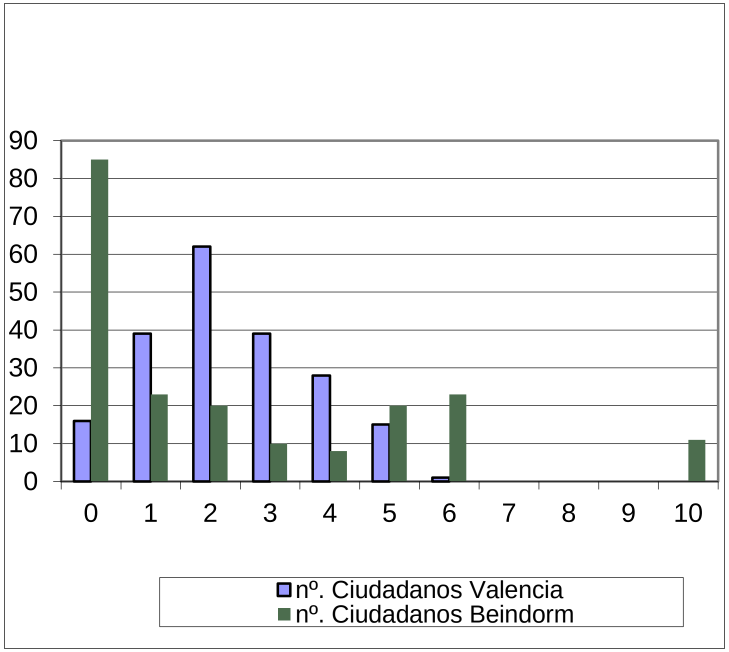
$$\overline{Y} = \frac{\sum_{i=1}^{200} Y_i}{200} = 2{,}365 = \text{PROMEDIO}(\quad)$$

Ejemplo 2

$$S_X^2 = \frac{\sum_{i=1}^{200}\left(X_i - \overline{X}\right)^2}{200} = VAR.P() = 1{,}891\,\text{días}^2$$

$$s_X = \sqrt{\frac{\sum_{i=1}^{200}\left(X_i - \overline{X}\right)^2}{200}} = DESVEST.P() = 1{,}3754\,\text{días}$$

$$\mathrm{Re}\,corrido = MAX() - MIN() = 6$$



Ejemplo 3

$$S_Y^2 = \frac{\sum_{i=1}^{200}\left(Y_i - \overline{Y}\right)^2}{200} = VAR.P() = 8{,}1518\,\text{días}^2$$

$$s_Y = \sqrt{\frac{\sum_{i=1}^{200}\left(Y_i - \overline{Y}\right)^2}{200}} = DESVEST.P() = 2{,}855\,\text{días}$$

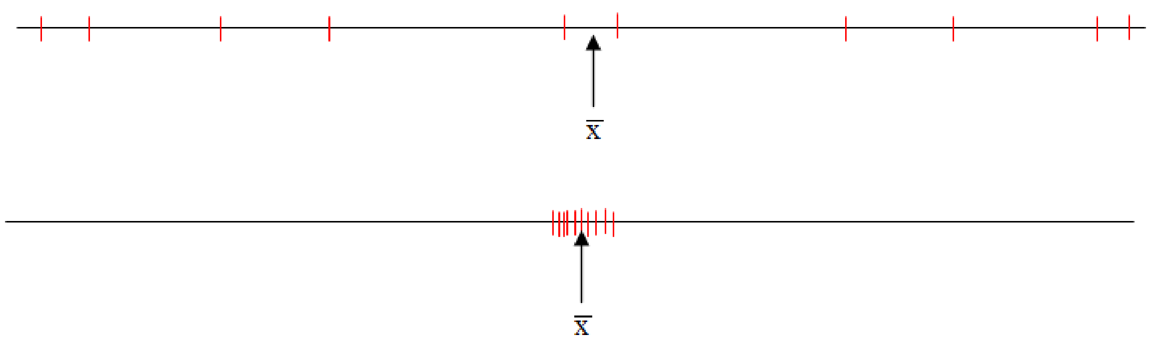
$$\mathrm{Re}\,corrido = MAX() - MIN() = 10$$

MEDIDAS DE POSICIÓN: MEDIA ARITMÉTICA

Si x_1, x_2, ……..x_N son los N valores de la variable estadística:

$$\overline{x} = \frac{x_1 + x_2 + \ldots\ldots + x_N}{N} = \frac{\sum_{i=1}^{N} x_i}{N}$$

¿Representa bien la media a dichos valores?



ESTADÍSTICA DESCRIPTIVA

MEDIDAS DE DISPERSIÓN: VARIANZA, DESVIACIÓN TÍPICA, RECORRIDO

Si x_1, x_2, ……..x_N son los N valores de la variable estadística:

- *VARIANZA:*

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

=VAR.P (   )

- *DESVIACIÓN TÍPICA:*

$$S_x = +\sqrt{S_x^2} \geq 0$$

- *RECORRIDO:*

Recorrido: $R_e = X_{max} - X_{min}$

ESTADÍSTICA DESCRIPTIVA

Ejemplo 4

Se analizan los pedidos efectuados por la empresa LUNASPUMA a una sección comercial de la misma empresa a lo largo de 360 días. El importe de los pedidos en euros se recoge en la siguiente tabla:

700	1000	700	1200	1200	800	700	1000	450	1200
500	800	500	450	450	700	800	800	500	900
700	1200	900	700	700	600	450	1200	900	700
600	600	500	800	800	700	900	600	1000	450
900	450	450	300	300	800	300	450	800	500
1200	300	600	600	600	1000	700	300	1200	1000
700	300	900	1200	700	600	800	300	600	1000
700	500	1200	800	300	450	1000	500	450	800
800	900	450	1000	450	500	300	900	300	700
1200	450	700	600	1000	300	1000	450	300	600
1000	450	800	700	800	900	800	450	500	450
600	300	1200	500	1200	1000	1200	300	900	300
900	1000	1000	300	600	600	600	1000	450	300
1200	300	600	900	450	900	450	300	450	500
450	500	900	1000	300	1200	300	500	300	900
700	450	700	800	300	450	300	450	1000	450
800	600	450	1200	500	700	500	600	300	450
700	900	700	700	900	800	900	900	500	300
600	1200	800	600	450	1200	450	1200	700	1000
700	450	1200	900	450	1000	450	450	600	300
800	700	450	1200	300	600	300	700	900	500
700	800	700	450	1000	900	1000	800	1200	450
500	1200	800	700	300	1200	300	300	450	600
450	1000	300	800	500	450	500	600	700	900
600	600	600	1000	450	700	450	1200	800	700
900	900	1200	600	600	800	600	800	300	450
1200	700	800	450	900	450	900	1000	600	700
450	450	1000	500	1200	700	1200	900	1200	800
700	700	600	300	450	800	450	450	800	300
800	800	450	900	700	1200	700	450	700	600
450	700	500	1000	800	1000	800	300	600	1200
900	500	300	800	300	600	300	1000	450	800
300	700	900	1000	600	900	600	300	500	1000
700	800	1000	700	700	1200	700	1000	300	600
800	1000	800	450	800	900	800	1200	900	900
700	700	1000	300	1000	300	1000	900	1000	700

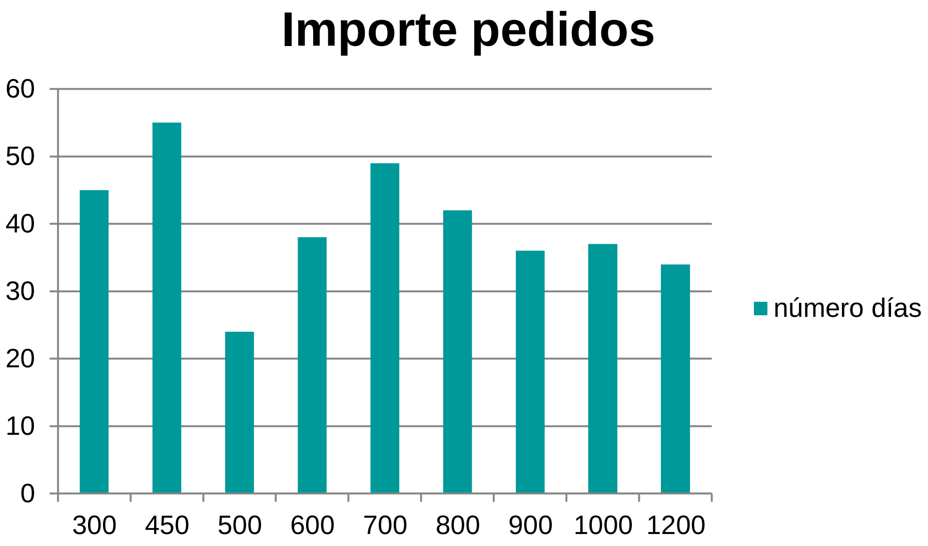
a) Definir y clasificar la variable objeto de estudio.
b) Tabla distribución de frecuencia absoluta y gráfico de barras.
c) Calcular el recorrido del importe de los pedidos.
d) Obtener la cuantía media del importe de los pedidos.
e) Deducir la desviación típica del importe de los pedidos e interpretar el resultado obtenido.

Ejemplo 4

Tabla frecuencias

Importe pedidos en euros	nº de dias
300	45
450	55
500	24
600	38
700	49
800	42
900	36
1000	37
1200	34
	360

Gráfico barras



$$\overline{X} = 697,6€$$

$$S_X^2 = 71293€^2$$

$$S_X = 267€$$

$$Re = 900€$$

2.3. CÁLCULO DE PROBABILIDADES

- Experimentos:
  - Deterministas o Causales
  - Aleatorios
    - Variables aleatorias: discretas y continuas

-DISTRIBUCIÓN DE PROBABILIDAD

-MEDIDAS DE REDUCCIÓN

- MEDIDAS DE POSICIÓN: esperanza matemática (μ)
- MEDIDAS DE DISPERSIÓN: varianza (σ^2), desviación típica (σ), recorrido

CÁLCULO DE PROBABILIDADES

- Modelos de distribuciones de probabilidad: idealizaciones de la realidad que tipifican en familias las distribuciones de probabilidad

  - Discretos: Binomial y Poisson

  - Continuos: Normal o de Gauss

Los modelos van a actuar de puente entre
lo observado (muestra) y lo desconocido (población)

CÁLCULO DE PROBABILIDADES

MODELO BINOMIAL: CARACTERÍSTICAS

- **Repetición del experimento: n veces**
- **Resultados dicotómicos (éxito/fracaso) con p=probabilidad de éxito y q=probabilidad de fracaso**
- **Independencia de resultados en las n pruebas**
- **Variable aleatoria: X= nº de éxitos obtenidos en las** n **pruebas**

$$X \in \{0, 1, 2, \ldots\ldots\ldots n\}$$

NOTACIÓN: X~B(n, p)

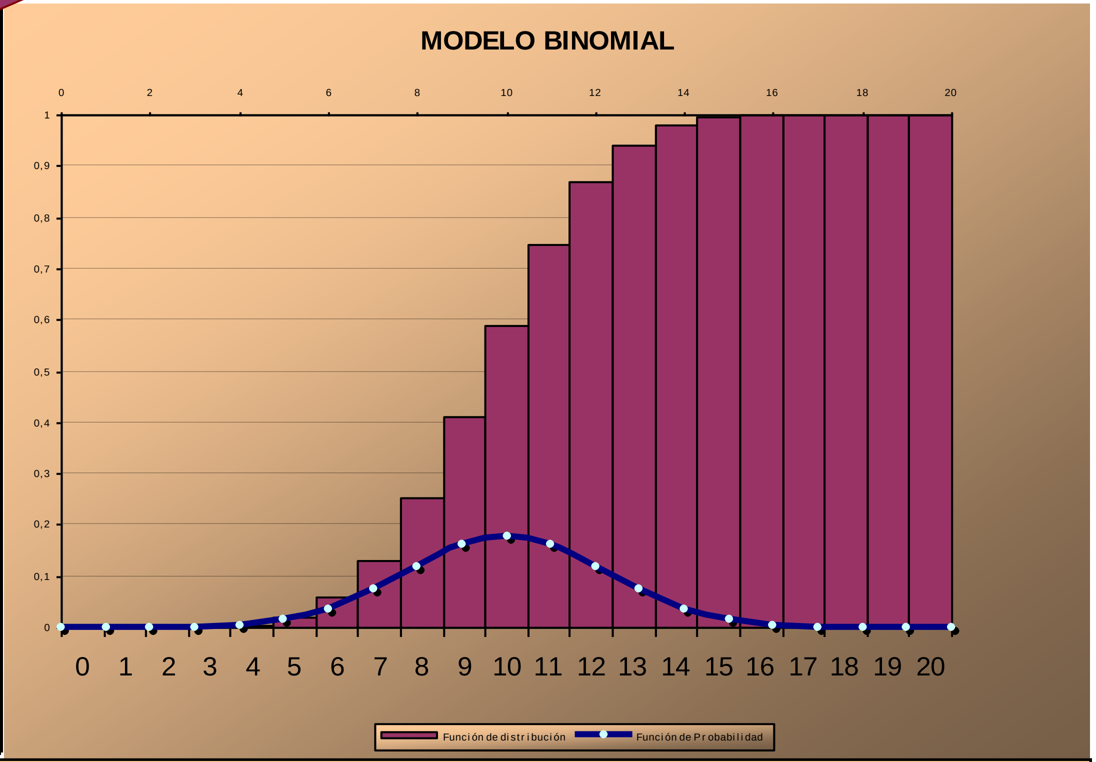
CÁLCULO DE PROBABILIDADES

=DISTR.BINOM.N(X;n;p;FALSO)

=DISTR.BINOM.N(X;n;p;VERDADERO)

MODELO BINOMIAL

Número de pruebas	X=Número éxitos	Función de Probabilidad	Función de Distribución
20	0	0,0000	9,53674E-07
	1	0,0000	2,00272E-05
Probabilidad éxito p	2	0,0002	0,000201225
0,5	3	0,0011	0,001288414
	4	0,0046	0,005908966
	5	0,0148	0,020694733
	6	0,0370	0,057659149
	7	0,0739	0,131587982
	8	0,1201	0,251722336
	9	0,1602	0,411901474
	10	0,1762	0,588098526
	11	0,1602	0,748277664
	12	0,1201	0,868412018
	13	0,0739	0,942340851
	14	0,0370	0,979305267
	15	0,0148	0,994091034
	16	0,0046	0,998711586
	17	0,0011	0,999798775
	18	0,0002	0,999979973
	19	0,0000	0,999999046
	20	0,0000	1



RESULTADOS DE INTERÉS

Valor medio de B(n,p) μ **= n.p= 20.0,5 =10**

Varianza de B(n; p): $\sigma^2 = n\,p\,q = 20 \cdot 0{,}5 \cdot 0{,}5 = 5$

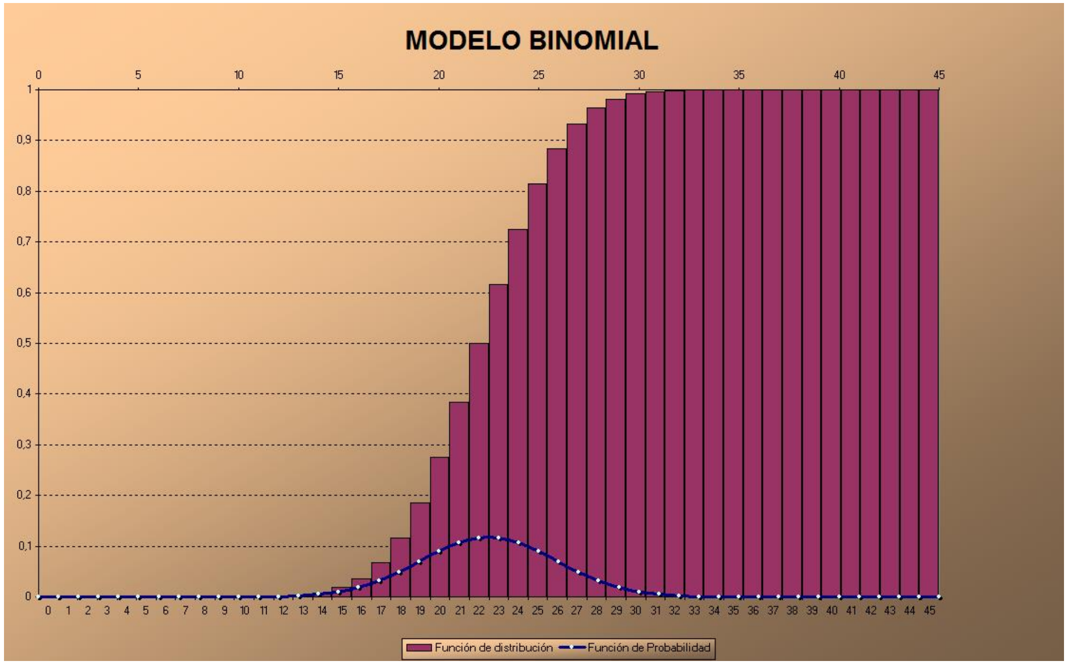
Ejemplo 5

Se sabe que el 50% de los habitantes de Valencia capital esperan habitualmente el autobús más de 15 minutos. Si se seleccionan al azar 45 habitantes de dicha ciudad y se les pregunta acerca del tiempo de espera habitual en la parada del autobús:

a) ¿Cuántos de ellos se espera que contesten que están en dicha situación (tiempo de espera habitual superior a los 15 minutos)?

b) ¿Cuál es la probabilidad de que 30 de ellos contesten que esperan habitualmente al autobús más de 15 minutos?

c) ¿Cuál es la probabilidad de que más de 30 de los encuestados contesten que esperan habitualmente al autobús más de 15 minutos?



a) 22,5

b) 0’0098

c) 0’008

MODELO POISSON: CARACTERÍSTICAS

- **Un suceso puede ocurrir, o no, una o más veces en un intervalo de tiempo o espacio**
- **Probabilidad de ocurrencia de un suceso en un intervalo de tiempo o espacio es independiente del intervalo considerado (intervalos de igual dimensión)**
- **Si el intervalo es extremadamente pequeño, la probabilidad de que no ocurra ningún suceso es próxima a 1, la de que ocurra un suceso es muy pequeña y la de que ocurran 2 es prácticamente 0**
- **Variable aleatoria: X= nº de sucesos que ocurren en una unidad de tiempo o espacio**

$$X \in \{0, 1, 2, ...\}$$

λ= nº medio de sucesos que ocurren en una unidad de tiempo o espacio

NOTACIÓN: $X \sim \mathcal{P}(\lambda)$

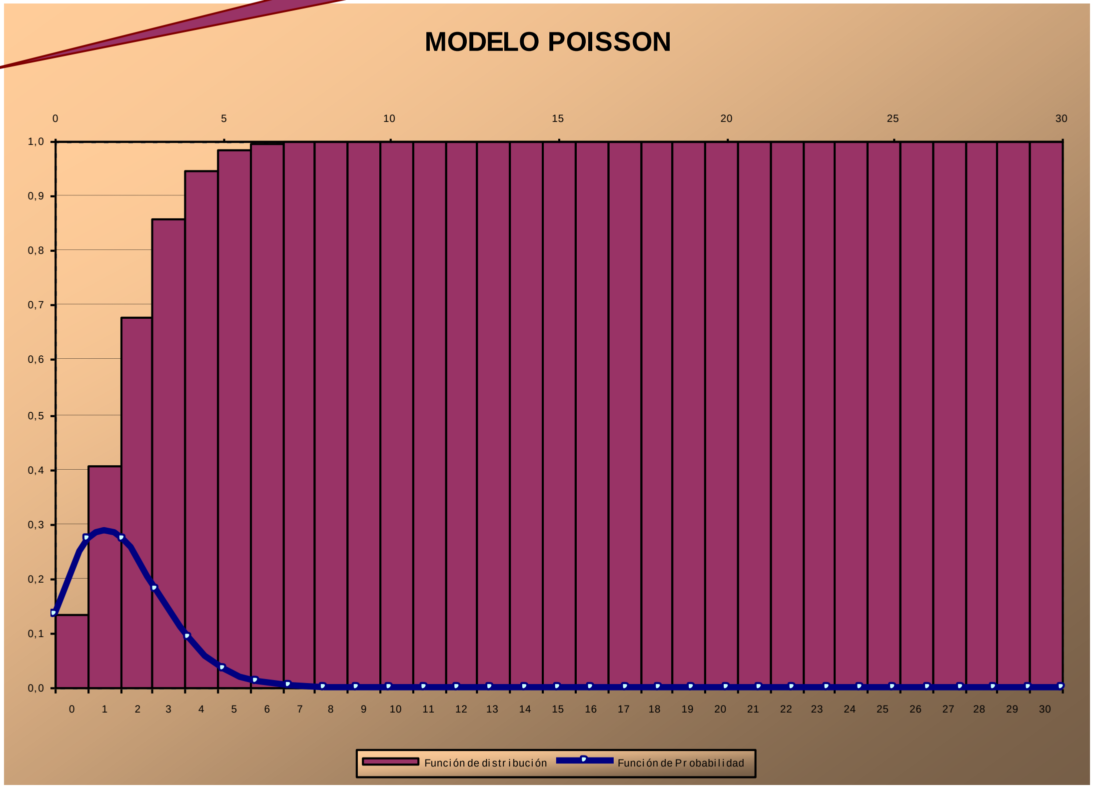
CÁLCULO DE PROBABILIDADES

=POISSON.DIST(X;λ;FALSO)

=POISSON.DIST(X;λ;VERDADERO)

MODELO POISSON $\mathcal{P}(\lambda)$

Valor de la media λ	X=Valor variable	Función de Probabilidad	Función de Distribución
2	0	0,1353	0,1353
	1	0,2707	0,4060
	2	0,2707	0,6767
	3	0,1804	0,8571
	4	0,0902	0,9473
	5	0,0361	0,9834
	6	0,0120	0,9955
	7	0,0034	0,9989
	8	0,0009	0,9998
	9	0,0002	1,0000
	10	0,0000	1,0000
	11	0,0000	1,0000
	12	0,0000	1,0000
	13	0,0000	1,0000
	14	0,0000	1,0000
	15	0,0000	1,0000
	16	0,0000	1,0000
	17	0,0000	1,0000
	18	0,0000	1,0000
	19	0,0000	1,0000
	20	0,0000	1,0000
	21	0,0000	1,0000
	22	0,0000	1,0000
	23	0,0000	1,0000
	24	0,0000	1,0000
	25	0,0000	1,0000
	26	0,0000	1,0000
	27	0,0000	1,0000
	28	0,0000	1,0000
	29	0,0000	1,0000
	30	0,0000	1,0000



RESULTADOS DE INTERES

Valor medio de $\mathcal{P}(\lambda)$: $\mu = \lambda = 2$

Varianza de $\mathcal{P}(\lambda)$: $\sigma^2 = \lambda = 2$

SUMA DE K POISSON INDEPENDIENTES

$\mathcal{P}(\lambda_1) + \mathcal{P}(\lambda_2) + \mathcal{P}(\lambda_3) + \ldots\ldots\ldots\ldots + \mathcal{P}(\lambda_k) = \mathcal{P}(\lambda_1+\lambda_2+\lambda_3+\ldots+\lambda_k)$

Ejemplo 6

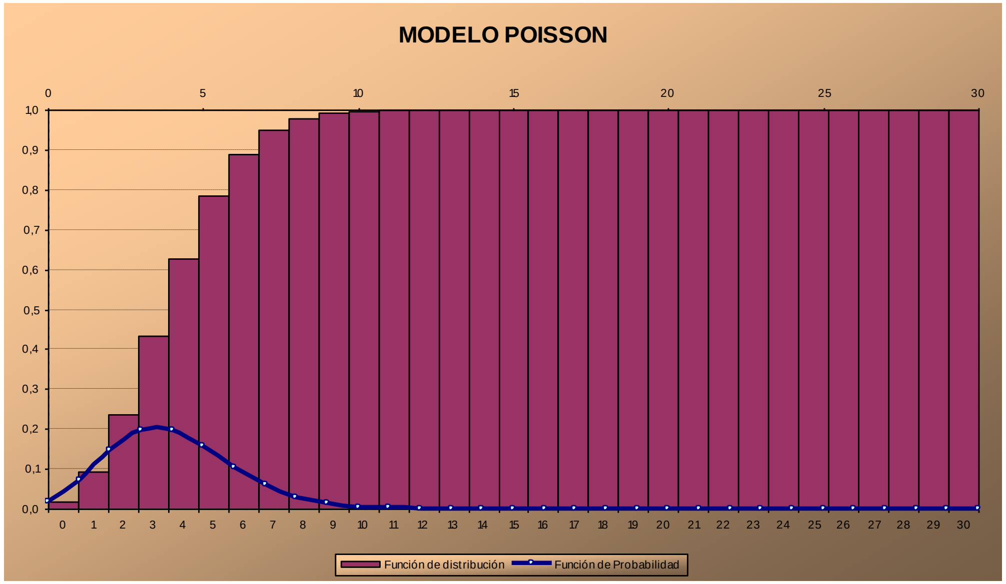
Si el número de averías de los autobuses de la EMT con más de 5 años de servicio sigue un modelo Poisson con media 4 averías anuales

a) ¿Cuál es el porcentaje de autobuses con dicha antigüedad que tienen 2 averías en un año?

b) ¿Cuál es el porcentaje de autobuses con dicha antigüedad que tienen menos de 2 averías en un año?

c) Suponiendo independiente el número de averías año a año, obtener la probabilidad de que un autobús con más de 5 años de servicio tenga un total de 10 averías en los próximos 3 años.

CÁLCULO DE PROBABILIDADES



a) 14’65%

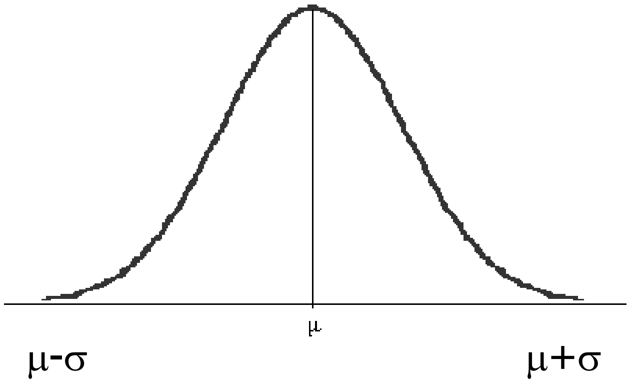
b) 9’16%

c) 0’1048

MODELO (DISTRIBUCIÓN) NORMAL O DE GAUSS

- **X= variable aleatoria de *tipo continuo* sigue una distribución Normal de parámetros μ y σ (con σ>0), si su función de densidad es:**

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



NOTACIÓN: X~N(μ, σ)

CÁLCULO DE PROBABILIDADES

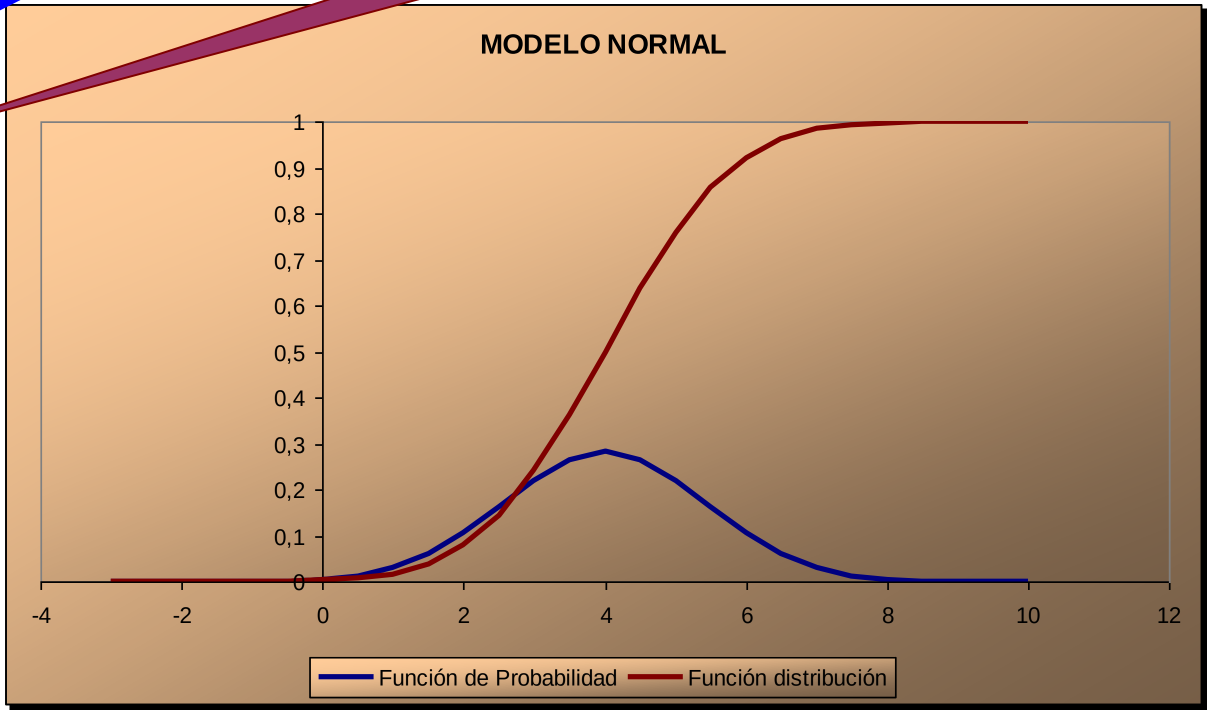
=DISTR.NORM.N(X;μ;σ;FALSO)

=DISTR.NORM.N(X;μ;σ;VERDADERO)

MODELO NORMAL N(μ

Media (μ)	Varianza (σ^2)
4	2

X=Valor variable	Función de Probabilidad	Función de Distribución
-3	1,34986E-06	3,71549E-07
-2,5	7,29726E-06	2,15139E-06
-2	3,48133E-05	1,10452E-05
-1,5	0,000146569	5,0311E-05
-1	0,000544571	0,000203476
-0,5	0,00178558	0,000731358
0	0,005166746	0,002338867
0,5	0,013193749	0,006664164
1	0,029732572	0,016947427
1,5	0,059130281	0,038549936
2	0,103776874	0,078649604
2,5	0,160732767	0,144422183
3	0,219695645	0,239750061
3,5	0,265003532	0,361836805
4	0,282094792	0,5
4,5	0,265003532	0,638163195
5	0,219695645	0,760249939
5,5	0,160732767	0,855577817
6	0,103776874	0,921350396
6,5	0,059130281	0,961450064
7	0,029732572	0,983052573
7,5	0,013193749	0,993335836
8	0,005166746	0,997661133
8,5	0,00178558	0,999268642
9	0,000544571	0,999796524
9,5	0,000146569	0,999949689
10	3,48133E-05	0,999988955



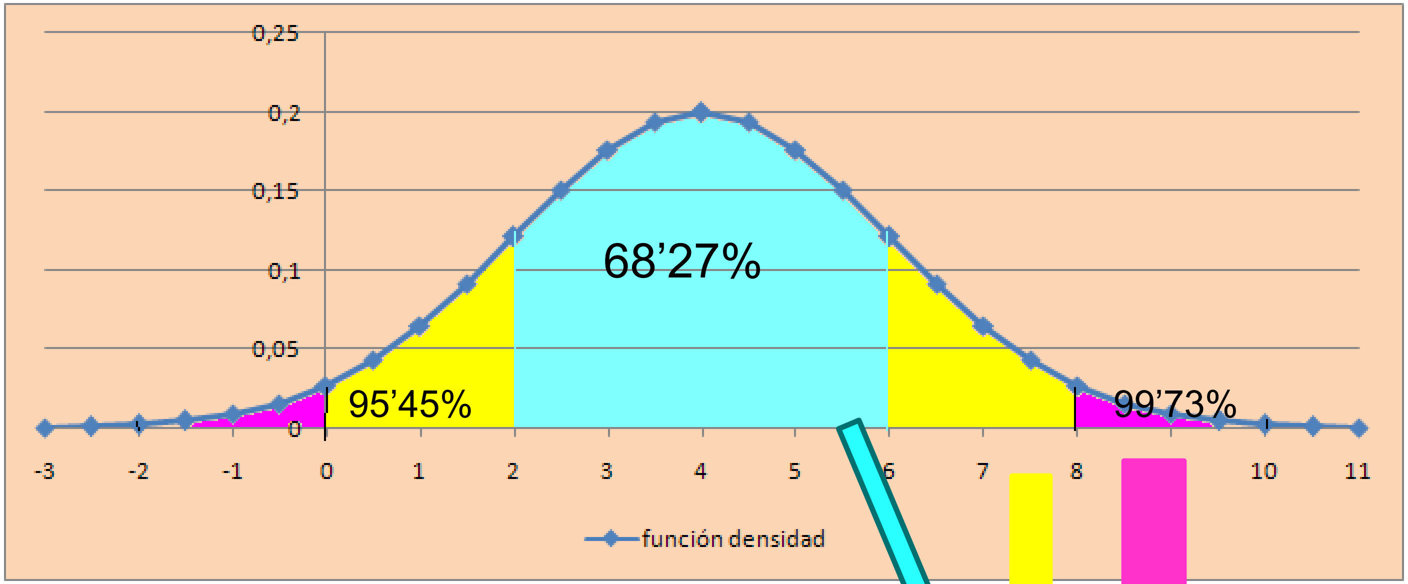
RESULTADOS DE INTERÉS

p(a <X < b) = F(b)- F (a)

		F(b)		F(a)	p(a <X < b)
k	μ+ k σ	p(x < μ + k σ)	μ- k σ	p(x < μ - k σ)	p(μ-k σ <x < μ+ k σ)
0	4	0,5	4	0,5	0,0
0,5	4,7071	0,6915	3,2929	0,3085	0,3829
1	5,4142	0,8413	2,5858	0,1587	0,6827
1,5	6,1213	0,9332	1,8787	0,0668	0,8664
2	6,8284	0,9772	1,1716	0,0228	0,9545
2,5	7,5355	0,9938	0,4645	0,0062	0,9876
3	8,2426	0,9987	-0,2426	0,0013	0,9973

CÁLCULO DE PROBABILIDADES

$N(\mu=4,\ \sigma = 2)$



$$P(\mu-\sigma<X<\mu+\sigma)=0{,}6827$$

$$P(\mu-2\sigma<X<\mu+2\sigma)=0{,}9545$$

$$P(\mu-3\sigma<X<\mu+3\sigma)=0{,}9973$$

Ejemplo 7

El tiempo que tarda en realizar su recorrido la línea 1 del metro de Valencia sigue una distribución Normal con media 30 minutos y desviación típica 5 minutos.

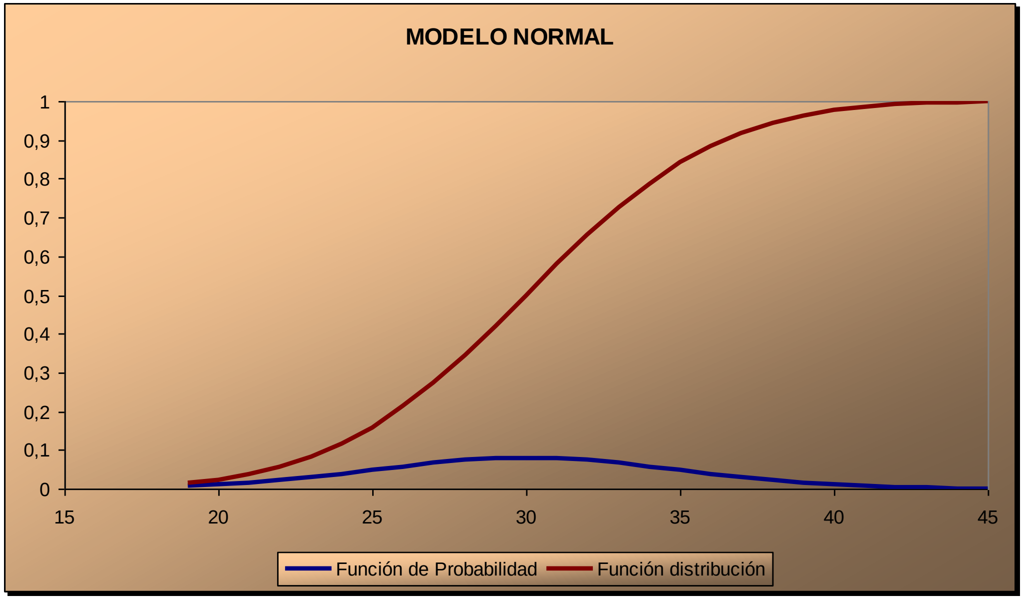
a) Obtener la probabilidad de que un metro de dicha línea:

a.1) Invierta más de 30 minutos en realizar el recorrido

a.2) Emplee entre 25 y 39 minutos en realizar el recorrido

b) ¿Cuál es el intervalo de tiempo que el 99’73% de los metros de dicha línea invierten en realizar el recorrido?

CÁLCULO DE PROBABILIDADES



a.1) 0’5

a.2) 0’8054

b) [15min., 45min]

RELACIÓN ENTRE LOS MODELOS: TEOREMAS DE CONVERGENCIA

BINOMIAL-POISSON

Si X~B(n, p) con *n suficientemente grande y p suficientemente pequeño* ⇒ **X~ $\mathscr{P}(\lambda=np)$**

$n \geq 50$ y $np \leq 5$

BINOMIAL-NORMAL

Si X~B(n, p) con n suficientemente grande ⇒ **X~N($\mu=np$, $\sigma^2=npq$)**

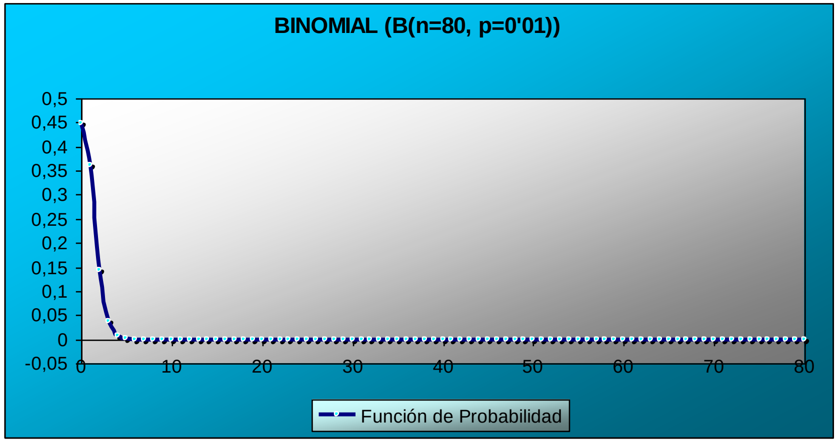
$n \geq 50$ y $np > 5$

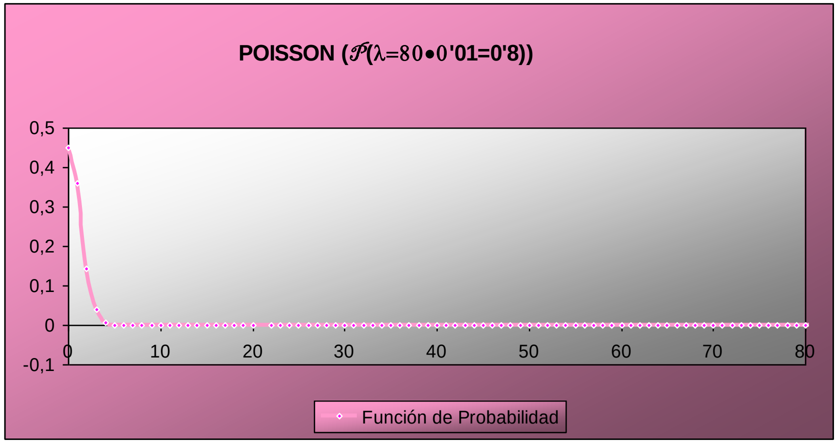
POISSON-NORMAL

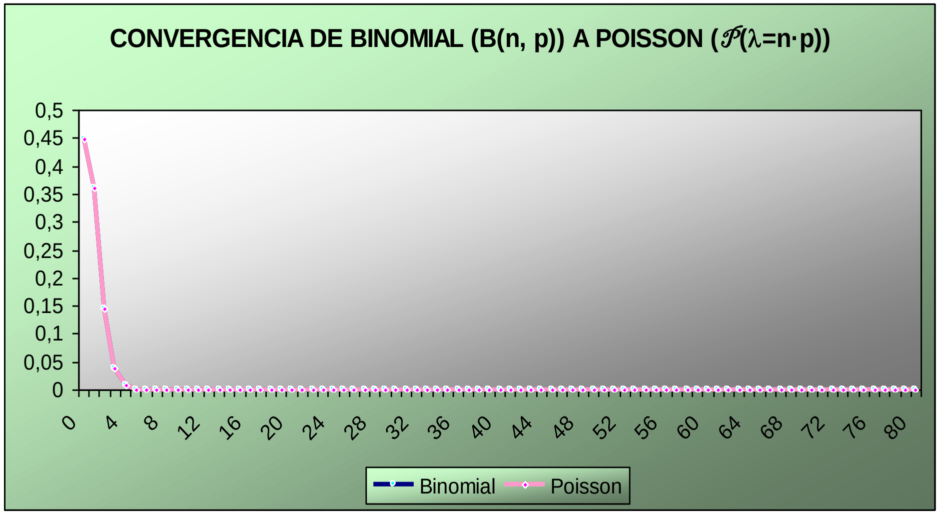
Si X~$\mathscr{P}(\lambda)$ y λ *suficientemente grande* ⇒ **X~N($\mu=\lambda$, $\sigma^2= \lambda$)**

CÁLCULO DE PROBABILIDADES

APROXIMACIÓN DE BINOMIAL A POISSON

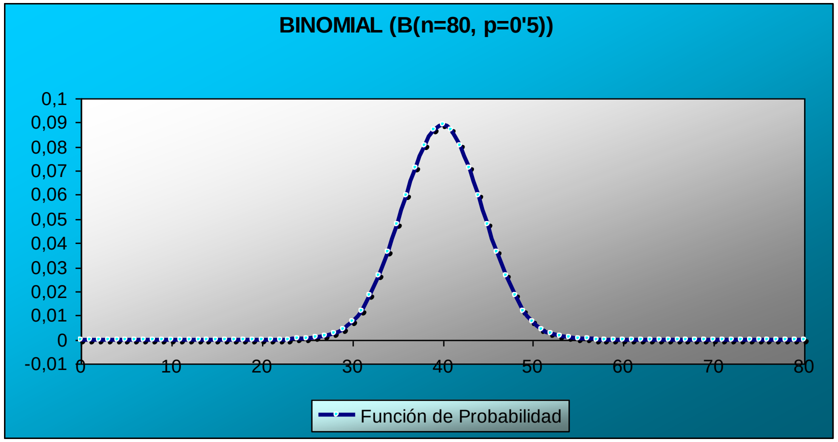


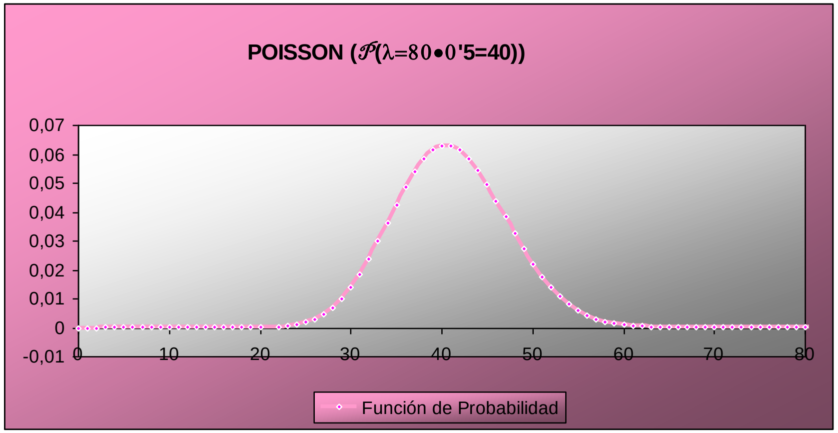


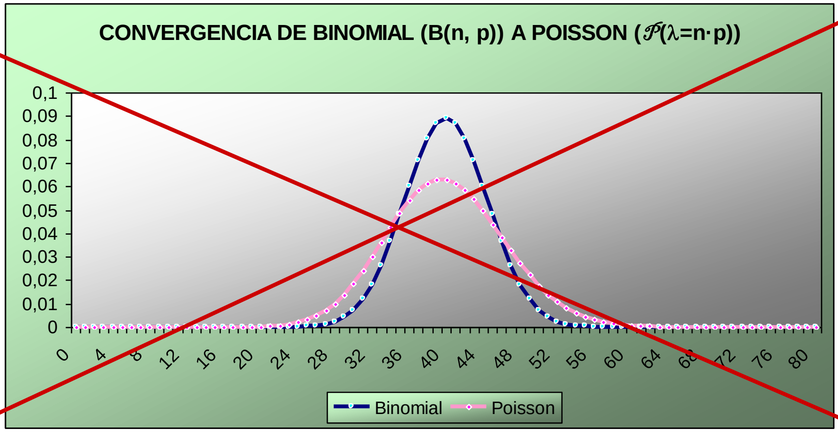


CÁLCULO DE PROBABILIDADES

APROXIMACIÓN DE BINOMIAL A POISSON

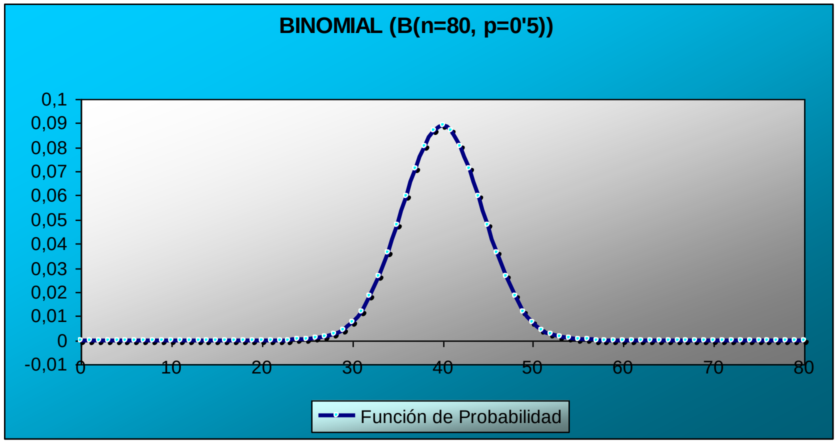


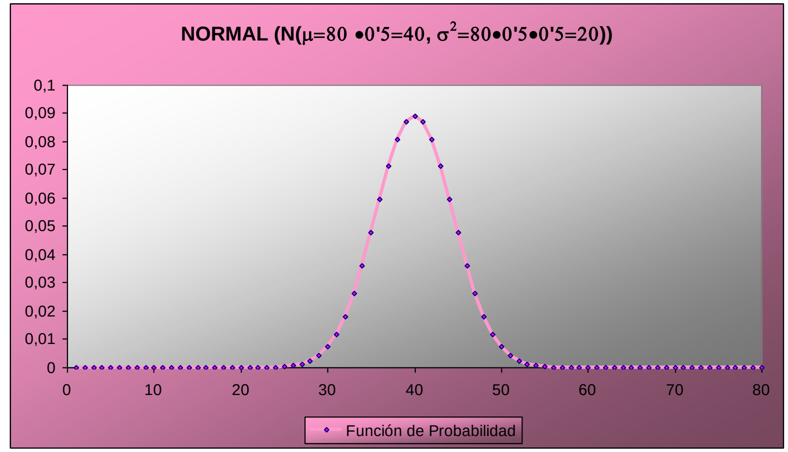


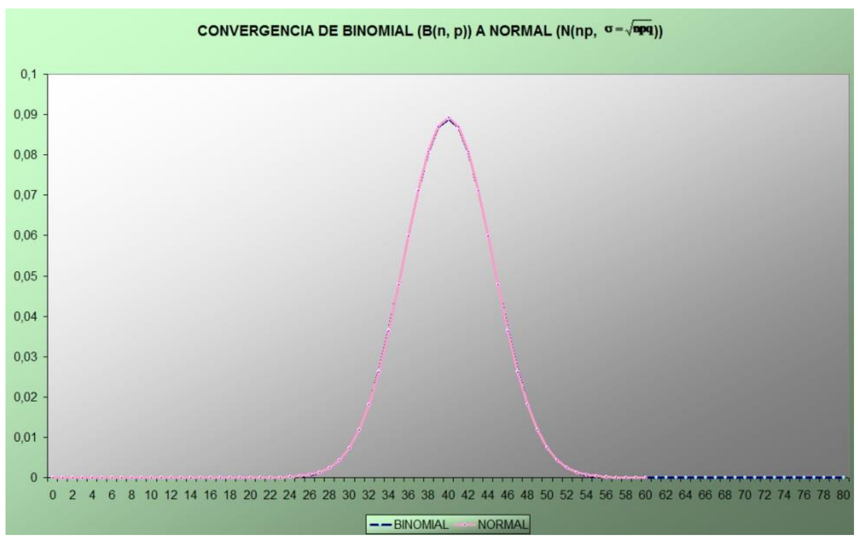


CÁLCULO DE PROBABILIDADES

APROXIMACIÓN DE BINOMIAL A NORMAL

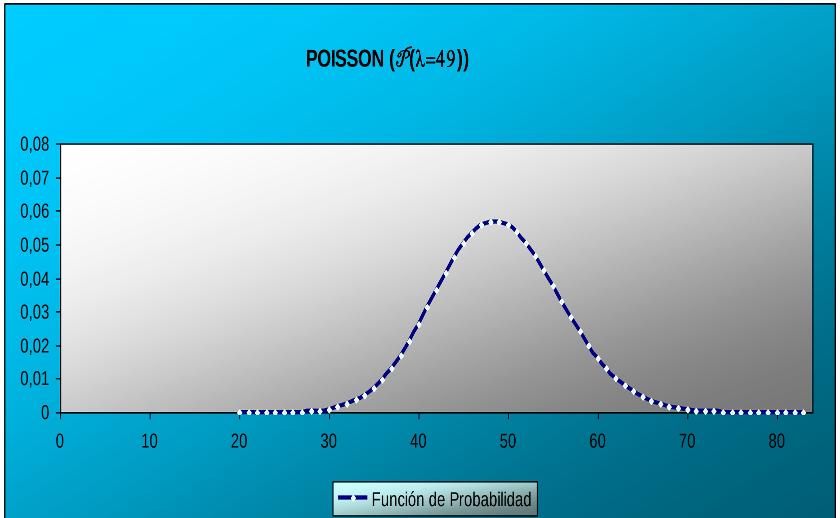


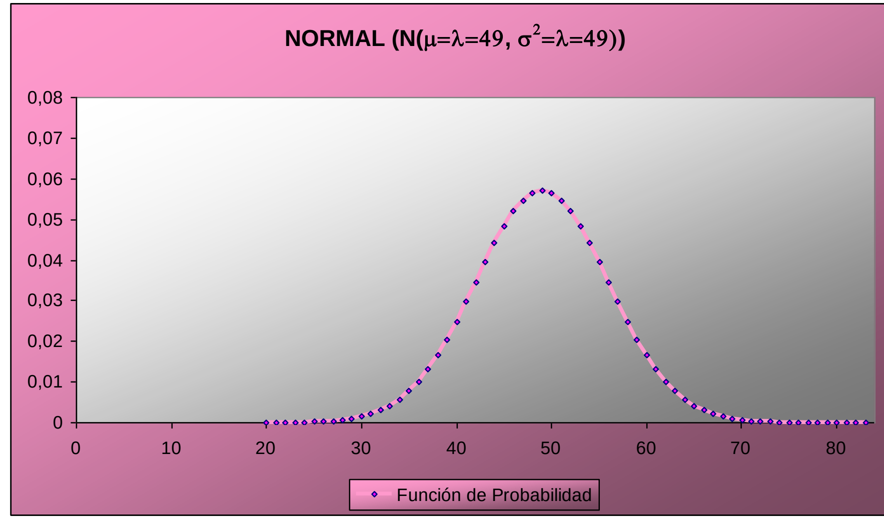


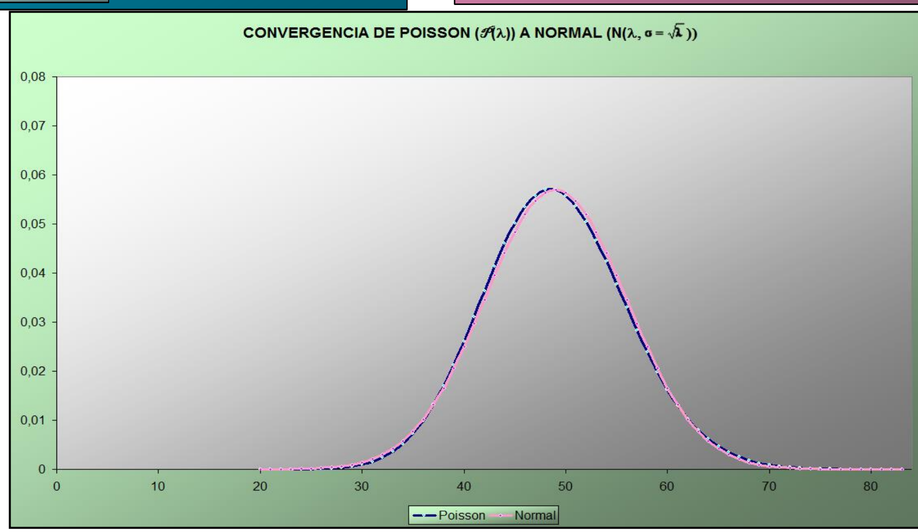


CÁLCULO DE PROBABILIDADES

APROXIMACIÓN DE POISSON A NORMAL

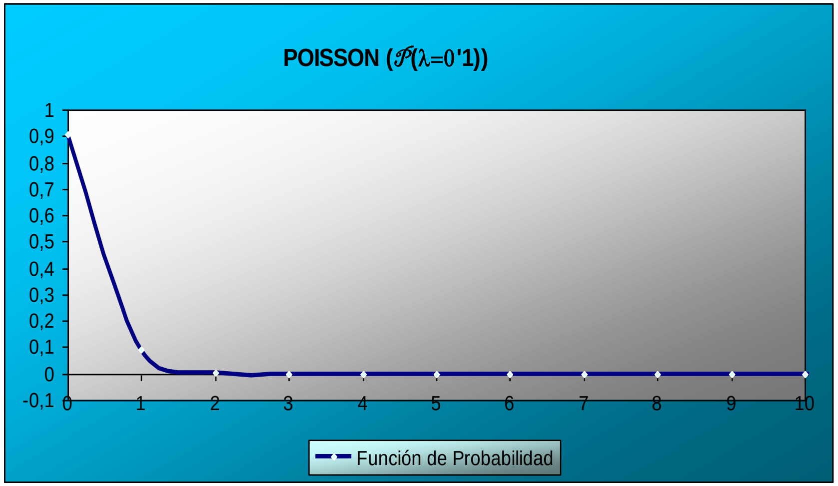


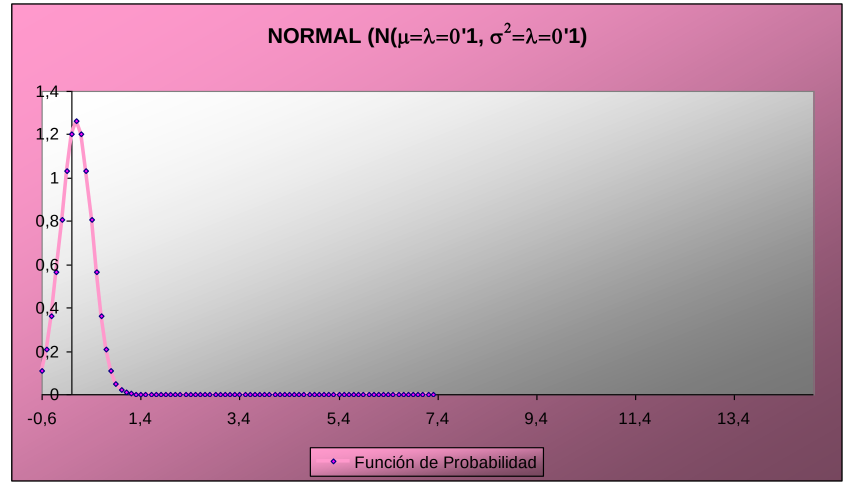


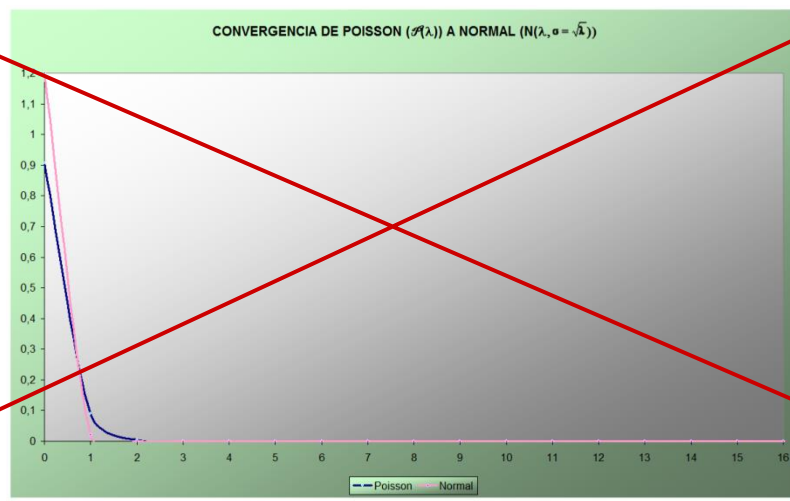


CÁLCULO DE PROBABILIDADES

APROXIMACIÓN DE POISSON A NORMAL







Ejemplo 8

En un proceso de fabricación de piezas, necesarias para la elaboración de lámparas de diseño, se sabe que el porcentaje de piezas defectuosas fabricadas es el 1%. Obtener:

a) Probabilidad de que si se seleccionan 6000 piezas más de 60 resulten defectuosas.

b) Probabilidad de que si se consideran, al azar, 70 piezas fabricadas más de 1 sea defectuosa.

Ejemplo 9

A partir de los datos del ejemplo anterior, si se seleccionasen al azar lotes de 6000 piezas, y se contasen el número de piezas defectuosas en cada lote, obtener cual sería, en el 99’73% de los lotes, el número mínimo y máximo de piezas defectuosas contabilizadas (intervalo centrado).

Ejemplo 8

a) X=v. a. nº piezas defectuosas de las 6000 seleccionadas

$X \sim B(n=6000;\ p=0'01)$

np=60>5 $\longrightarrow$ se puede aproximar a $N(\mu=np=60,\ \sigma^2=npq=59'4)$

$\boxed{p(X>60)=0'5}$

b) X=v. a. nº piezas defectuosas de las 70 seleccionadas

$X \sim B(n=70;\ p=0'01)$

np=0'7<5 $\longrightarrow$ se puede aproximar a $\mathscr{P}(\lambda=0'7)$

$\boxed{p(X>1)=0'1558}$

Ejemplo 9

X=v. a. nº piezas defectuosas de las 6000 seleccionadas

$X \sim B(n=6000;\ p=0'01)$

np=60>5 → se puede aproximar a N(μ=np=60, σ^2=npq=59'4)

el 99'73% de los datos se encuentra en el intervalo [μ-3σ; μ-3σ], en este caso

$$\mu - 3 \cdot \sigma = 60 - 3 \cdot \sqrt{59'4} = 36'88 \cong \boxed{37}$$

$$\mu + 3 \cdot \sigma = 60 + 3 \cdot \sqrt{59'4} = 83'12 \cong \boxed{83}$$

Ejemplo 10

El número de llamadas recibidas en una hora en el servicio de urgencias de un hospital sigue una distribución de Poisson con $\sigma^2=64$. Obtener la probabilidad de que en una hora seleccionada al azar se reciban, en dicho servicio, menos de 88 llamadas.

Ejemplo 11

A partir de los datos del ejemplo anterior, si se seleccionan al azar franjas horarias de una hora y se contabilizan el número de llamadas recibidas en cada una de ellas ¿Cuál sería el porcentaje de dichas franjas en las que el número de llamadas recibidas oscilara entre las 40 y las 88?.

Ejemplo 10

En Poissson $\mu=\lambda=\sigma^2$

X=v. a. nº llamadas/hora → $X \sim \mathcal{P}(\lambda=64)$

$\lambda=64$ es suficientemente grande → se puede aproximar a $N(\mu=\lambda=64,\ \sigma^2=\lambda=64)$

$p(X<88)=0'9987$

Ejemplo 11

X=v. a. nº llamadas/hora → $X \sim \mathscr{P}(\lambda=64)$ → λ suficientemente grande

↓

se puede aproximar a $N(\mu=64,\ \sigma^2=64)$

↓

el 99'73% de los datos se encuentran en el intervalo $[\mu-3\sigma;\ \mu-3\sigma]$. En este caso

$$\left.\begin{aligned}\mu-3\cdot\sigma&=64-3\cdot\sqrt{64}=\boxed{40}\\ \mu+3\cdot\sigma&=64+3\cdot\sqrt{64}=\boxed{88}\end{aligned}\right\}$$

→ En el 99'73% de las franjas de una hora el número de llamadas oscila entre las 40 y las 88

TEOREMA CENTRAL DEL LÍMITE (T.C.L.)

Sean $X_1, X_2, \ldots\ldots X_n$ ***v. a independientes igualmente distribuidas***

$$X_i \sim D(\mu, \boxed{\sigma}) \quad \forall i \in \{1, 2, \ldots\ldots n\} \text{ con } \sigma \neq 0$$

1ª Consecuencia

Si se define $S_n = \sum_{i=1}^{n} X_i$ → **Para n suficientemente grande**

$$S_n \sim N\left(n\mu, \sigma\sqrt{n}\right)$$

2ª Consecuencia

Si se define $M_n = \sum_{i=1}^{n} \frac{X_i}{n} = \overline{X} = \text{media muestral}$ → **Para n suficientemente grande**

$$\overline{X} \sim N\left(\mu, \boxed{\frac{\sigma}{\sqrt{n}}}\right)$$

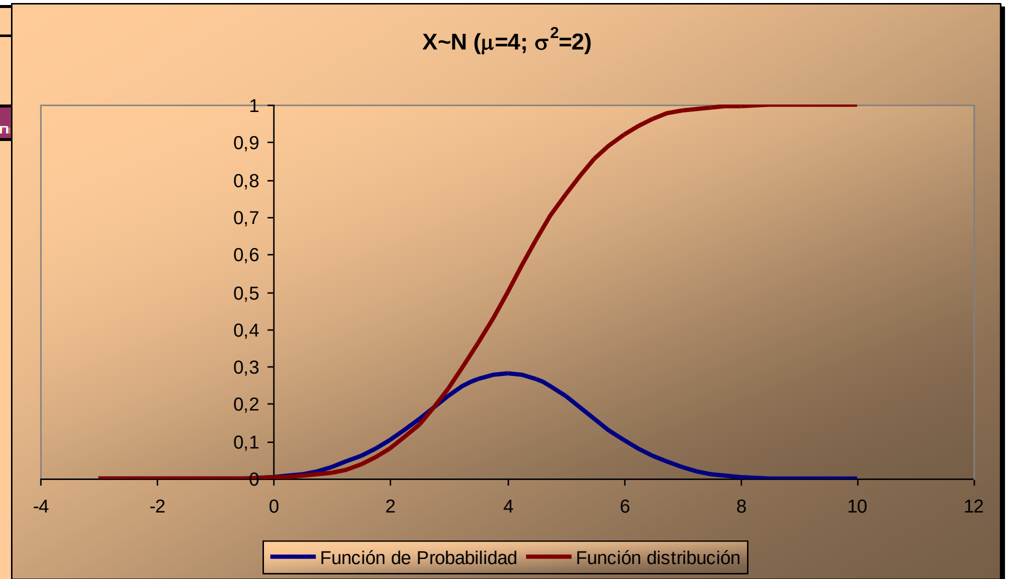
CÁLCULO DE PROBABILIDADES

Ejemplo 12

MODELO NORMAL $N(\mu, \sigma)$

Media (μ)	Varianza (σ^2)
4	2

X=Valor variable	Función de Probabilidad	Función de Distribución
-3	1,34986E-06	3,71549E-07
-2,5	7,29726E-06	2,15139E-06
-2	3,48133E-05	1,10452E-05
-1,5	0,000146569	5,0311E-05
-1	0,000544571	0,000203476
-0,5	0,00178558	0,000731358
0	0,005166746	0,002338867
0,5	0,013193749	0,006664164
1	0,029732572	0,016947427
1,5	0,059130281	0,038549936
2	0,103776874	0,078649604
2,5	0,160732767	0,144422183
3	0,219695645	0,239750061
3,5	0,265003532	0,361836805
4	0,282094792	0,5
4,5	0,265003532	0,638163195
5	0,219695645	0,760249939
5,5	0,160732767	0,855577817
6	0,103776874	0,921350396
6,5	0,059130281	0,961450064
7	0,029732572	0,983052573
7,5	0,013193749	0,993335836
8	0,005166746	0,997661133
8,5	0,00178558	0,999268642
9	0,000544571	0,999796524
9,5	0,000146569	0,999949689
10	3,48133E-05	0,999988955

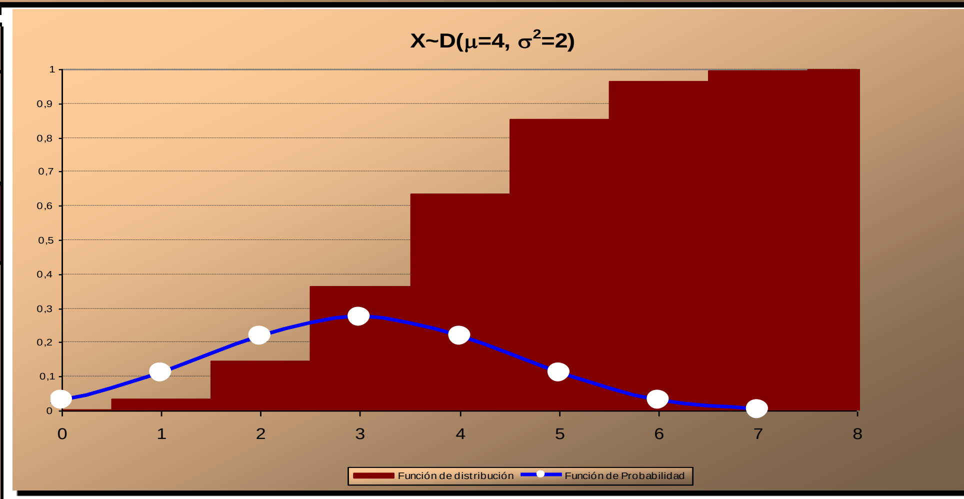


Ejemplo 13

MODELO DISCRETO $D(\mu, \sigma)$

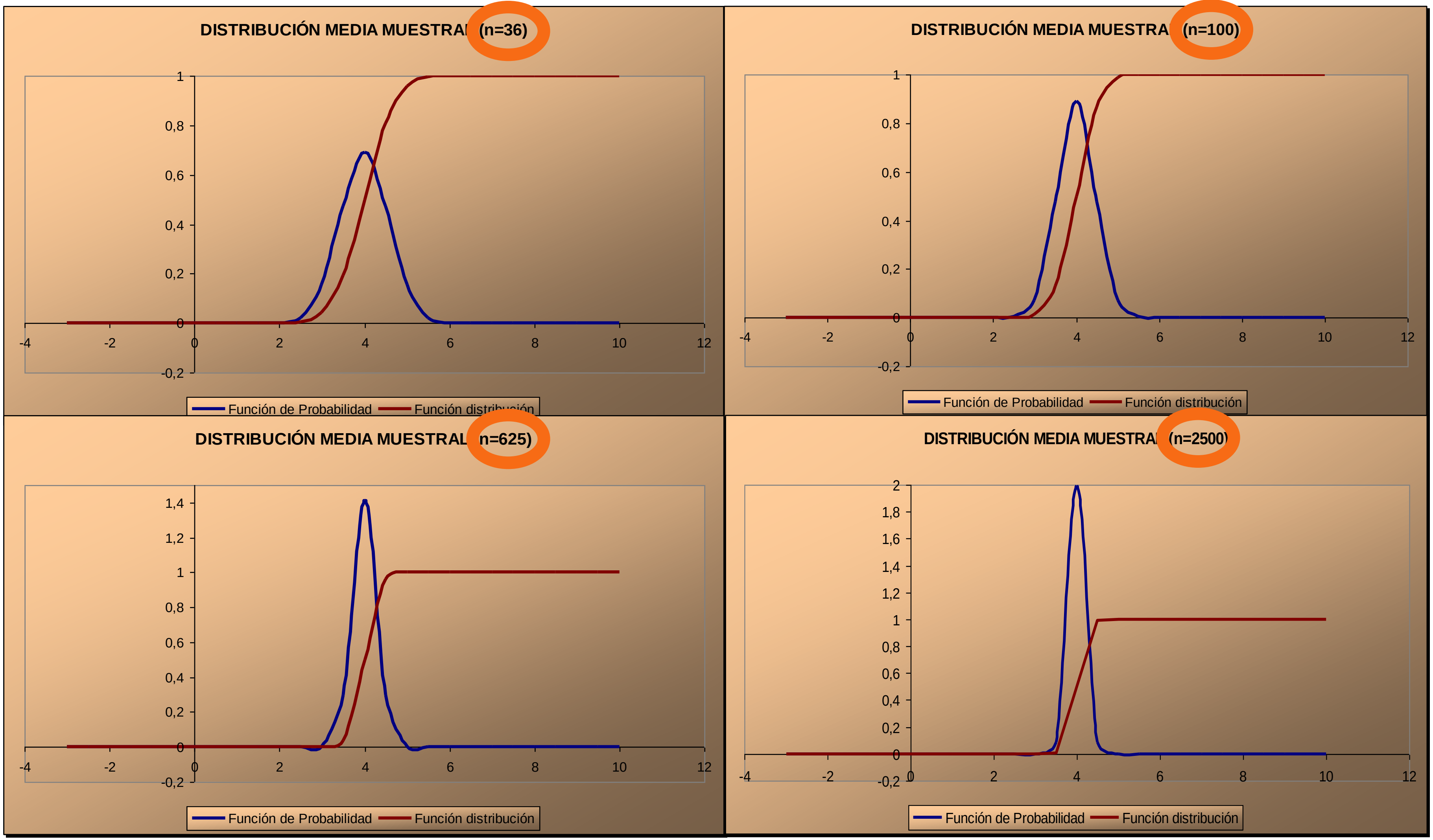
Media (μ)	Varianza (σ^2)
4	2

X=Valor variable	Función de Probabilidad	Función de Distribución
0	0,00390625	0,00390625
1	0,03125	0,03515625
2	0,109375	0,14453125
3	0,21875	0,36328125
4	0,2734375	0,63671875
5	0,21875	0,85546875
6	0,109375	0,96484375
7	0,03125	0,99609375
8	0,00390625	1



CÁLCULO DE PROBABILIDADES

Ejemplo 14



Ejemplo 15

Se sabe que el beneficio mensual de cierta empresa sigue una distribución con μ=500 (miles de €) y σ= 100 (miles de €). Suponiendo independientes los beneficios mes a mes, obtener:

a) Probabilidad de que en los próximos 3 años (36 meses hábiles) el beneficio total de dicha empresa supere los 19.200.000 €?

b) Probabilidad de que en los próximos 3 años (36 meses hábiles) el beneficio medio mensual de dicha empresa sea superior a 550.000 €?

CÁLCULO DE PROBABILIDADES

Ejemplo 15

a) Xi=v. a. beneficio/mes i-ésimo → Xi ~ D(μ=500; σ=100) e independientes

Y= v.a. beneficio total/36 meses hábiles → $Y = X_1 + X_2 + \ldots\ldots + X_{36} = \sum_{i=1}^{36} X_i$

1ª Consecuencia del TCL

$$Y \sim N\left(n\mu, \sigma\sqrt{n}\right) = N\left(36 \cdot 500;\ 100\sqrt{36}\right) = N(18.000; 600)$$

p(Y>19.200)=0’0227

b) Z=v. a. beneficio medio/en 36 meses hábiles → $Z = \frac{X_1 + X_2 + \ldots\ldots + X_{36}}{36} = \sum_{i=1}^{36} \frac{X_i}{36}$

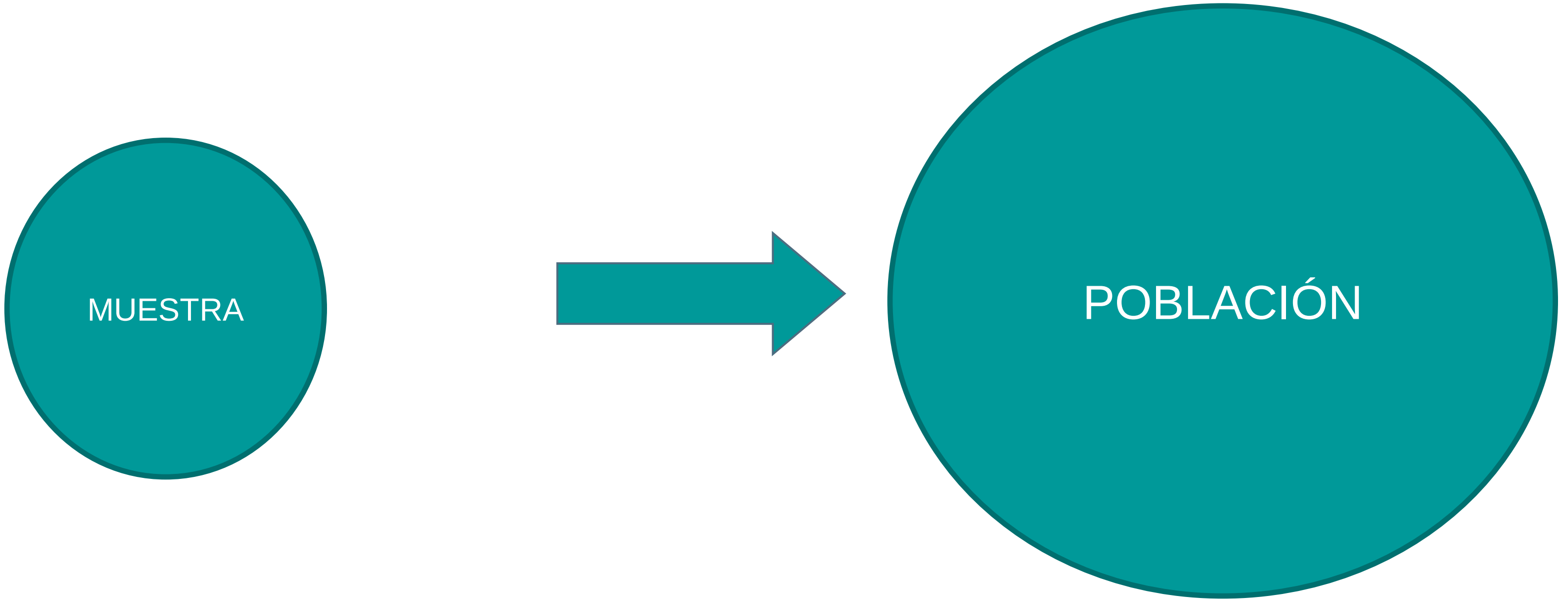
2ª Consecuencia del TCL

$$Z \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) = N\left(500;\ \frac{100}{\sqrt{36}}\right) = N(500; 16'\widehat{6})$$

p(Z>550)=0’0013

2.4. INFERENCIA ESTADÍSTICA

OBJETIVO: Obtener conclusiones de la población (todo el colectivo) a partir de los valores de una muestra (subconjunto de la población).



INFERENCIA ESTADÍSTICA

Inferencia: → Resolución de dos bloques de problemas

- La estimación
  - Por punto (puntual)
  - Por intervalo
- La contrastación
  - Paramétrica
  - No Paramétrica

INFERENCIA ESTADÍSTICA

ESTIMACIÓN PUNTUAL

$$\mu \longrightarrow \overline{X}$$

$$\sigma^2 \longrightarrow S^2$$

c$_2$ y d$_2$: valores tabulados dependientes del tamaño muestral

$$\sigma \longrightarrow \frac{S}{c_2} \quad \text{o bien} \quad \frac{R}{d_2}$$

$$p \longrightarrow \hat{p}$$

Ejemplo 16

Se desea estimar el número medio de piezas defectuosas fabricadas diariamente por una máquina en cierta cadena de producción. Para ello contabilizan las piezas defectuosas fabricadas por dicha máquina a lo largo de 10 días laborables, obteniéndose los siguientes resultados:

8, 7, 10, 8, 8, 15, 15, 16, 19, 10.

Obtener, a partir de los mismos, la estimación puntual buscada.

Ejemplo 17

Con los datos del ejemplo anterior, y considerando el estimador dependiente de la desviación típica muestral, obtener la estimación puntual de la desviación típica del número de piezas fabricadas diariamente.

Ejemplo 18

Con los datos del ejemplo 16, y considerando el estimador dependiente del recorrido de la muestra, obtener la estimación puntual de la desviación típica del número de piezas fabricadas diariamente.

INFERENCIA ESTADÍSTICA

Ejemplo 16

Estimador de μ $\longrightarrow$ $\overline{X}$

En este caso $\dfrac{8 + \ldots\ldots\ldots\ldots\ldots + 10}{10} = 11'6$

Ejemplo 17

Estimador de σ dependiente de S $\longrightarrow$ $\dfrac{S}{c_2}$

En este caso $\dfrac{4'02988}{0'9727} = 4'1429$

Ejemplo 18

Estimador de σ dependiente de R $\longrightarrow$ $\dfrac{R}{d_2}$

En este caso $\dfrac{12}{3'078} = 3'8986$

INFERENCIA ESTADÍSTICA

ESTIMACIÓN PUNTUAL

EN CONTROL ESTADÍSTICO DE LA CALIDAD

$$\mu \longrightarrow \bar{\bar{X}}$$

$$\sigma \longrightarrow \left\{ \begin{array}{l} \dfrac{\bar{S}}{c_2} \\[2ex] \dfrac{\bar{R}}{d_2} \end{array} \right.$$

CONTRASTACIÓN DE HIPÓTESIS

- Hipótesis Estadística
- Contraste de Hipótesis Estadística
- Tipos de Contrastes de Hipótesis:
  - Paramétricos
  - No Paramétricos
- Tipos de Hipótesis Estadísticas:
  - Hipótesis Simple
  - Hipótesis Compuesta

EJEMPLO 19

Una PYME dedicada a la elaboración de tortas imperiales tiene en su cadena de producción una única máquina dosificadora de la cantidad de almendras que será utilizada en la confección de cada torta imperial (en la composición de éstas consta que la cantidad de almendra utilizada en cada una ha sido de 100 gramos). Se quiere constatar si la máquina está bien regulada, esto es si por TÉRMINO MEDIO el peso de cada dosis de almendras es de 100 gramos.

Paramétrico si X=v. a. peso/bolsa pertenece a una familia paramétrica conocida

No paramétrico si X=v. a. peso/bolsa no pertenece a una familia paramétrica conocida

H_o: $\mu=100$

H_1: $\mu\neq100$

- Error de Tipo I y Error de Tipo II

	Rechazar H_o	No rechazar H_o
H_o cierta	Fallo (Error Tipo I)	Acierto
H_o falsa	Acierto	Fallo (Error Tipo II)

Error de Tipo I (o de 1ª especie)=Rechazar H_o/ H_o cierta
Error de Tipo II (o de 2ª especie)=No Rechazar H_o/ H_o falsa

- Nivel de Significación=α= p (Error de Tipo I)=
  =p(Rechazar H_o/ H_o cierta)

- Potencia del Test =p(Rechazar H_o/ H_o falsa)=1-β

  (β=p(Error de Tipo II)=p(No Rechazar H_o/ H_o falsa))

EJEMPLO 20

Error de tipo I: se rechaza que la máquina está bien regulada cuando sí lo está.

Consecuencias: **Parada de producción**, **Costes** derivados del ajuste de una máquina que no necesitaba ser ajustada, **Pedidos** no entregados a tiempo, **Reclamaciones** clientes…………..

Error de tipo II: no se rechaza que la máquina está bien regulada cuando no lo está.

Consecuencias: **Sobrecoste**, **Reclamaciones** clientes…………………………

TÉCNICA GENERAL

- **P.1.** Considerar una variable aleatoria (v.a.) T, que dependa de la muestra y del parámetro acerca del que se ha efectuado la hipótesis, de forma que la distribución de T sea conocida supuesta cierta la hipótesis nula.
- **P.2.** Particionar el espacio de posibles valores de T en dos regiones: R_o (región de aceptación) y R_1 (región crítica o de rechazo)
- **P.3.** Obtener el valor de T para la muestra obtenida: T_o
- **P.4.** Aplicar la Regla de Decisión:
  - Si $T_o \in R_o \rightarrow$ No Rechazar H_o
  - Si $T_o \in R_1 \rightarrow$ Rechazar H_o

EJEMPLO 21

Suponiendo que X se distribuye N(μ, σ=5) y selecciona una muestra de 9 bolsas, cuyos pesos son:

99, 110, 106, 98, 106, 98, 99, 108, 109

¿Se rechaza o no que la máquina está regulada para α=1%?

EJEMPLO 21

Se obtiene que

$$T_o = \frac{103'\widehat{6} - 100}{5/\sqrt{9}} = 2'2$$

y

Ro=[-2’58, 2’58]

Así, dado que $T_o \in R_o$, NO SE RECHAZA que la máquina está regulada para α=1% ($\cong$ no se rechaza H_0:μ=100)

ENFOQUE DEL P-VALOR (P-VALUE)

Proporcionado por paquetes estadísticos y algunas hojas de cálculo
(aunque se puede obtener sin necesidad de utilizarlos)

P-VALOR (P-VALUE): mínimo valor de significación que nos lleva al rechazo de $H_{o,}$, es decir, si:

☹ Si $\alpha > p \Rightarrow$ se rechaza H_o

☺ Si $\alpha < p \Rightarrow$ no se rechaza H_o

VENTAJA FRENTE A TÉCNICA GENERAL: se puede tomar decisión para cualquier nivel de significación α sin necesidad de obtener explícitamente la región de aceptación (R_O)

EJEMPLO 22

Introducidos los datos anteriores en un paquete estadístico se obtiene que:

p-value=0'0278

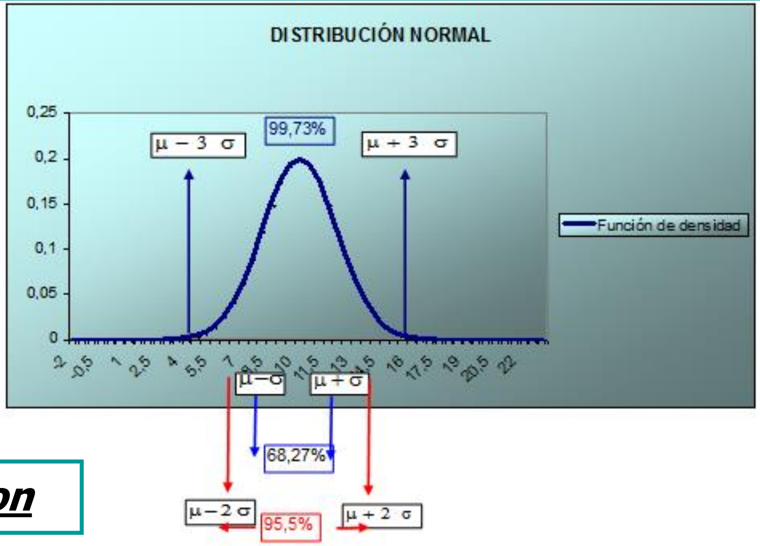
a) ¿Qué decisión se tomara para un nivel de significación α=1%?

b) ¿Y si α=5%?

SOLUCIÓN: a) NO RECHAZAR que la máquina está bien regulada ($\cong$ no se rechaza $H_0:\mu=100$).

b) RECHAZAR que la máquina está bien regulada ($\cong$ se rechaza $H_0:\mu=100$).

DISTRIBUCIONES DERIVADAS DE LA NORMAL



Distribución χ^2 de Pearson

$$X = Z_1^2 + Z_2^2 + \dots\dots + Z_n^2 \sim \chi_n^2$$

(siendo Z i ~ N(0, 1) e independientes)

Distribución t de Student

$$t = \frac{X}{\sqrt{\frac{Y}{n}}} \sim t_n$$

(siendo X~ N(0, 1); Y ~ χ_n^2 ; X e Y independientes)

Distribución F de Snedecor

$$F = \frac{X/m}{Y/n} = \frac{n}{m} \cdot \frac{X}{Y} \sim F_{m,n}$$

(siendo X~ χ_m^2 ; Y ~ χ_n^2; X e Y independientes)

GRADOS DE LIBERTAD

INFERENCIA ESTADÍSTICA

Analizar previamente la adecuación del modelo Normal

- Contrastes no paramétricos (χ^2 de Pearson, Kolmogorov)
- Métodos gráficos (histograma, papel probabilístico Normal)

VARIABLES Y DISTRIBUCIONES	*APLICACIONES*
$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0, 1)$	Contraste de hipótesis sobre la media de una población Normal con varianza conocida
$\frac{\overline{X}-\mu}{S} \cdot \sqrt{n-1} \sim t_{n-1}$	Contraste de hipótesis sobre la media de una población Normal con varianza desconocida
$\frac{\overline{X}-\overline{Y}-(\mu_X-\mu_Y)}{\sqrt{\frac{\sigma_X^2}{n_X}+\frac{\sigma_Y^2}{n_Y}}} \sim N(0; 1)$	Contraste de hipótesis sobre la diferencia de medias de dos poblaciones Normales con varianzas conocidas (muestras independientes)
$\frac{n_X}{n_Y} \cdot \frac{n_Y-1}{n_X-1} \cdot \frac{\sigma_y^2}{\sigma_x^2} \cdot \frac{S_x^2}{S_y^2} \sim F_{n_X-1;\, n_Y-1}$	Contraste de hipótesis sobre la razón de varianzas de dos poblaciones Normales (muestras independientes)

COMPARAR MEDIAS

COMPARAR VARIANZAS

TABLA

n	c_2	d_2
2	0'7979	1'128
3	0'8862	1'693
4	0'9213	2'059
5	0'9400	2'326
6	0'9515	2'534
7	0'9594	2'704
8	0'9650	2'847
9	0'9690	2'970
10	0'9727	3'078
11	0'9754	3'173
12	0'9776	3'258
13	0'9794	3'336
14	0'9810	3'407
15	0'9823	3'472
16	0'9835	3'532
17	0'9845	3'588
18	0'9854	3'640
19	0'9862	3'689
20	0'9869	3'735
21	0'9876	3'778
22	0'9882	3'819
23	0'9887	3'858
24	0'9892	3'895
25	0'9896	3'931

***GLOSARIO DE FUNCIONES EN EXCEL*:**

- ♦Media aritmética: =PROMEDIO( )
- ♦Varianza de la muestra: =VAR.P( )
- ♦Desviación típica de la muestra: =DESVEST.P( )
- ♦Recorrido: =MAX( )-MIN( )
- ♦Frecuencias: =FRECUENCIA(datos, grupos),
- ♦Función de Probabilidad (modelo Binomial): =DISTR.BINOM.N(X;n;p;FALSO)
- ♦Función de Distribución (modelo Binomial): =DISTR.BINOM.N(X;n;p;VERDADERO)
- ♦Función de Probabilidad (modelo Poisson): =POISSON.DIST(X;λ;FALSO)
- ♦Función de Distribución (modelo Poisson): =POISSON.DIST(X;λ;VERDADERO)
- ♦Función de Probabilidad (modelo Normal): =DISTR.NORM.N(X;μ;σ;FALSO)
- ♦Función de Distribución (modelo Normal): =DISTR.NORM.N(X;μ;σ;VERDADERO)