Bioestadística y uso de software científico

TEMA 2

PROBABILIDAD

DISTRIBUCIONES DE PROBABILIDAD

Índice

Conceptos generales de probabilidad

- La probabilidad de que algo ocurra es la proporción de veces que ocurriría si repitiéramos el experimento un número grande de veces.

$$P(A) = \lim_{n \to \infty} \frac{\textit{número de veces que se produce el resultado } A}{n}$$

Conceptos generales de probabilidad

- La probabilidad de que algo ocurra es la proporción de veces que ocurriría si repitiéramos el experimento un número grande de veces.
  - Definición frecuentista.
  - No dice nada sobre lo que ocurriría si el experimento se repite pocas veces.
  - No dice nada sobre lo que ocurrirá la próxima vez

Conceptos generales de probabilidad

- En una muestra:

- La probabilidad de que algo ocurra es la proporción de veces que ocurre en la muestra.

$$P(A) = \frac{\text{número de veces que se produce el resultado } A}{n}$$

Axiomas de la probabilidad

1. La probabilidad de cualquier suceso toma valores entre 0 y 1:

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

Axiomas de la probabilidad

2. La probabilidad del suceso seguro es igual a 1:

$$P(A) + P(no\ A) = P(E) = 1$$

Axiomas de la probabilidad

Dos sucesos son incompatibles si no pueden ocurrir simultáneamente:

$$P(A \cap B) = P(A \ y \ B) = 0$$

3. La probabilidad de que ocurra uno de los dos sucesos incompatibles es la suma de la probabilidad de cada uno de ellos:

$$P(A \cup B) = P(A \ ó \ B) = P(A) + P(B)$$

Una propiedad importante

Si dos sucesos son compatibles:

$$P(A \cap B) > 0$$

Entonces, la probabilidad de que ocurra uno de los dos sucesos es:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Una propiedad importante

	Fuman	No fuman	Total
Mujeres	20	80	100
Varones	40	160	200
Total	60	240	300

$$P(mujer) = 100/300 = 1/3$$

$$P(\text{varón}) = 200/300 = 2/3$$

$$P(fumador) = 60/300 = 0,2$$

$$P(no\ \ fumador) = 240/300 = 0,8$$

Una propiedad importante

	Fuman	No fuman	Total
Mujeres	20	80	100
Varones	40	160	200
Total	60	240	300

$$P(\text{mujer y varón}) = 0 \rightarrow \textit{sucesos incompatibles}$$

$$P(\text{mujer o varón}) = P(mujer) + P(\text{varón}) = 1/3 + 2/3 = 1$$

Una propiedad importante

	Fuman	No fuman	Total
Mujeres	20	80	100
Varones	40	160	200
Total	60	240	300

$$P(\text{mujer y fumador}) = 20/300 \rightarrow \textit{sucesos compatibles}$$

$$P(\text{mujer o fumador}) = P(mujer) + P(fumador) - P(mujer\ y\ fumador) =$$

$$= 1/3 + 0,2 - 20/300 = 140/300 = 0,4\overset{\frown}{6}$$

Índice

- Probabilidad
  - Definición de probabilidad
  - Probabilidad condicionada
  - Sucesos independientes
  - Teorema de Bayes
- Distribuciones de probabilidad
  - Distribución normal
  - Distribución binomial
  - Distribución de Poisson

Probabilidad condicionada

- Probabilidad condicionada P(A|B):
  - Probabilidad de que ocurra A habiendo ocurrido B
    - A=cara en la moneda 1; B=cara en la moneda 2
    - A=infarto; B=elevación del ST
    - A=infarto en el décimo paciente que llega a urgencias; B=infarto en el primer paciente que llega a urgencias.

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

Probabilidad condicionada

	Diabetes	No diabetes	Total
Glucosuria	60	8	68
No glucosuria	140	792	932
Total	200	800	1000

$$P(\text{glucosuria} \mid diabetes) = \frac{P(\text{glucosuria} \cap \text{diabetes})}{P(diabetes)}$$

$$= \frac{60/1000}{200/1000} = 0,3$$

Probabilidad condicionada

- ¿Por qué es importante la probabilidad condicionada?
  - Comparar dos tratamientos:
    - A=curación; B1=tratamiento con penicilina; B2=tratamiento con eritromicina

$$¿P(A \mid B1) = P(A \mid B2)?$$

Probabilidad condicionada

- ¿Por qué es importante la probabilidad condicionada?
  - Comparar dos pruebas diagnósticas:
    - A1=dar + en una radiografía; A2=dar + en un TAC; B=tener cáncer de pulmón

$$¿P(A1 \mid B) = P(A2 \mid B)?$$

Probabilidad condicionada

- ¿Por qué es importante la probabilidad condicionada?
  - Identificar factores de riesgo:
    - A=tener un infarto; B1=tener colesterol alto; B2=no tener colesterol alto

$$¿P(A \mid B1) = P(A \mid B2)?$$

Índice

- Probabilidad
  - Definición de probabilidad
  - Probabilidad condicionada
  - Sucesos independientes
  - Teorema de Bayes
- Distribuiciones de probabilidad
  - Distribución normal
  - Distribución binomial
  - Distribución de Poisson

Concepto de independencia

- Independencia:
  - Conocer si B ha ocurrido NO aporta ninguna información sobre si ocurre A

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

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

Concepto de independencia

	Diabetes	No diabetes	Total
Glucosuria	60	8	68
No glucosuria	140	792	932
Total	200	800	1000

$$P(\text{glucosuria} \mid \textit{diabetes}) = 0,3$$

$$P(\text{glucosuria}) = \frac{68}{1000} = 0,068$$

$$P(\text{glucosuria} \mid \textit{diabetes}) \neq P(\text{glucosuria}) \rightarrow$$

$$\rightarrow \text{glucosuria y diabetes NO son independientes}$$

Concepto de independencia

	Fuman	No fuman	Total
Mujeres	20	80	100
Varones	40	160	200
Total	60	240	300

$$P(\text{mujer} \mid fumador) = 20/60 = 1/3$$

$$P(\text{mujer}) = \frac{100}{300} = \frac{1}{3}$$

$$P(\text{mujer} \mid fumador) = P(mujer) \rightarrow$$

$$\rightarrow mujer\ \ y\ \ fumador\ \text{SÍ}\ son\ independientes$$

Un error muy frecuente

- Creer que una racha de sucesos independientes determina el siguiente resultado:
  - Me ha salido mal 5 veces seguidas, a la siguiente tengo más posibilidades de que me salga bien.
  - Ha caído tres veces en rojo, a la siguiente tiene más probabilidades de caer en negro.
  - He tenido tres niñas, ahora me toca un niño.
  - Se me han infectado los tres últimos pacientes, el siguiente tiene menos posibilidades de infectarse

Índice

- Probabilidad
  - Definición de probabilidad
  - Probabilidad condicionada
  - Sucesos independientes
  - Teorema de Bayes
- Distribuiciones de probabilidad
  - Distribución normal
  - Distribución binomial
  - Distribución de Poisson

Teorema de Bayes

	Diabetes	No diabetes	Total
Glucosuria	60	8	68
No glucosuria	140	792	932
Total	200	800	1000

$$P(D \mid G) = \frac{P(D)P(G \mid D)}{P(D)P(G \mid D) + P(\bar{D})P(G \mid \bar{D})}$$

$$P(D \mid G) = \frac{\frac{200}{1000} \times \frac{60}{200}}{\frac{200}{1000} \times \frac{60}{200} + \frac{800}{1000} \times \frac{8}{800}} = \frac{60}{68} = 0{,}882$$

Teorema de Bayes

	Diabetes	Intolerancia a la glucosa	No diabetes
Glucosuria			
No glucosuria			

$$P(D \mid G) = \frac{P(D)P(G \mid D)}{P(D)P(G \mid D) + P(I)P(G \mid I) + P(ND)P(G \mid ND)}$$

Teorema de Bayes

B_1, B_2,..., B_n sucesos mutuamente excluyentes, de los cuales uno debe ocurrir.

- Conociendo la probabilidad de cada B_i y la probabilidad de otro suceso A condicionada a cada B_i

- Calcular la probabilidad de B_j condicionada a A

$$P(B_j \mid A) = \frac{P(B_j)P(A \mid B_j)}{\sum_{i=1}^{n} P(B_i)P(A \mid B_i)}$$

Teorema de Bayes

¿Por qué es importante el teorema de Bayes?

- El 75% de los pacientes con apendicitis tienen fiebre, dolor en fosa ilíaca derecha y vómitos

- Llega un paciente con fiebre, dolor en fosa ilíaca derecha y vómitos, ¿cuál es la probabilidad de que tenga una apendicitis?

Teorema de Bayes

¿Por qué es importante el teorema de Bayes?

- El 75% de los pacientes con apendicitis tienen fiebre, dolor en fosa ilíaca derecha y vómitos

$$P(F+D+V \mid Apendicitis)$$

- Llega un paciente con fiebre, dolor en fosa ilíaca derecha y vómitos, ¿cuál es la probabilidad de que tenga una apendicitis?

$$P(Apendicitis \mid F+D+V)$$

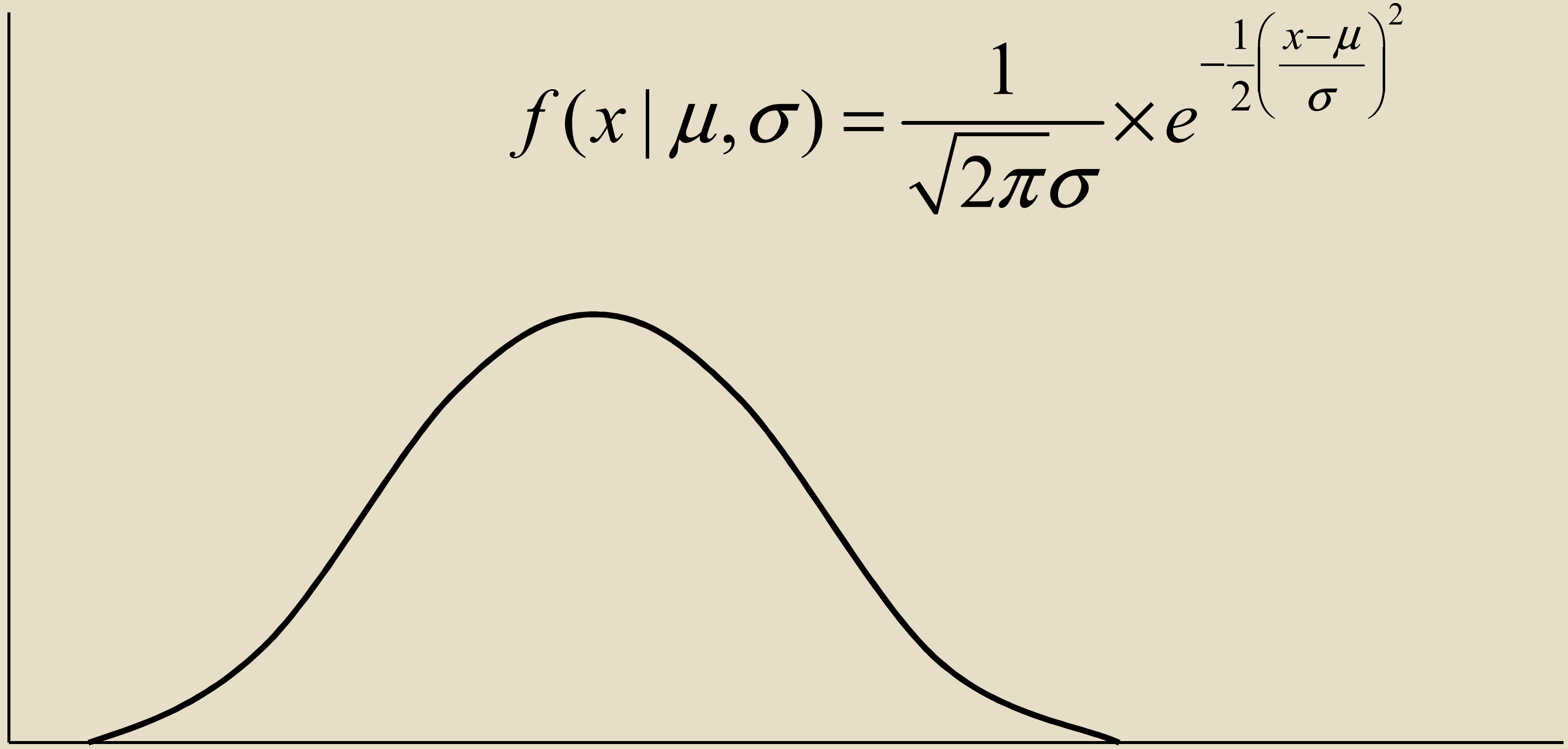
Índice

- Probabilidad
  - Definición de probabilidad
  - Probabilidad condicionada
  - Sucesos independientes
  - Teorema de Bayes
- Distribuiciones de probabilidad
  - Distribución normal
  - Distribución binomial
  - Distribución de Poisson

Distribuciones de probabilidad

- "Curva de Gauss"

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

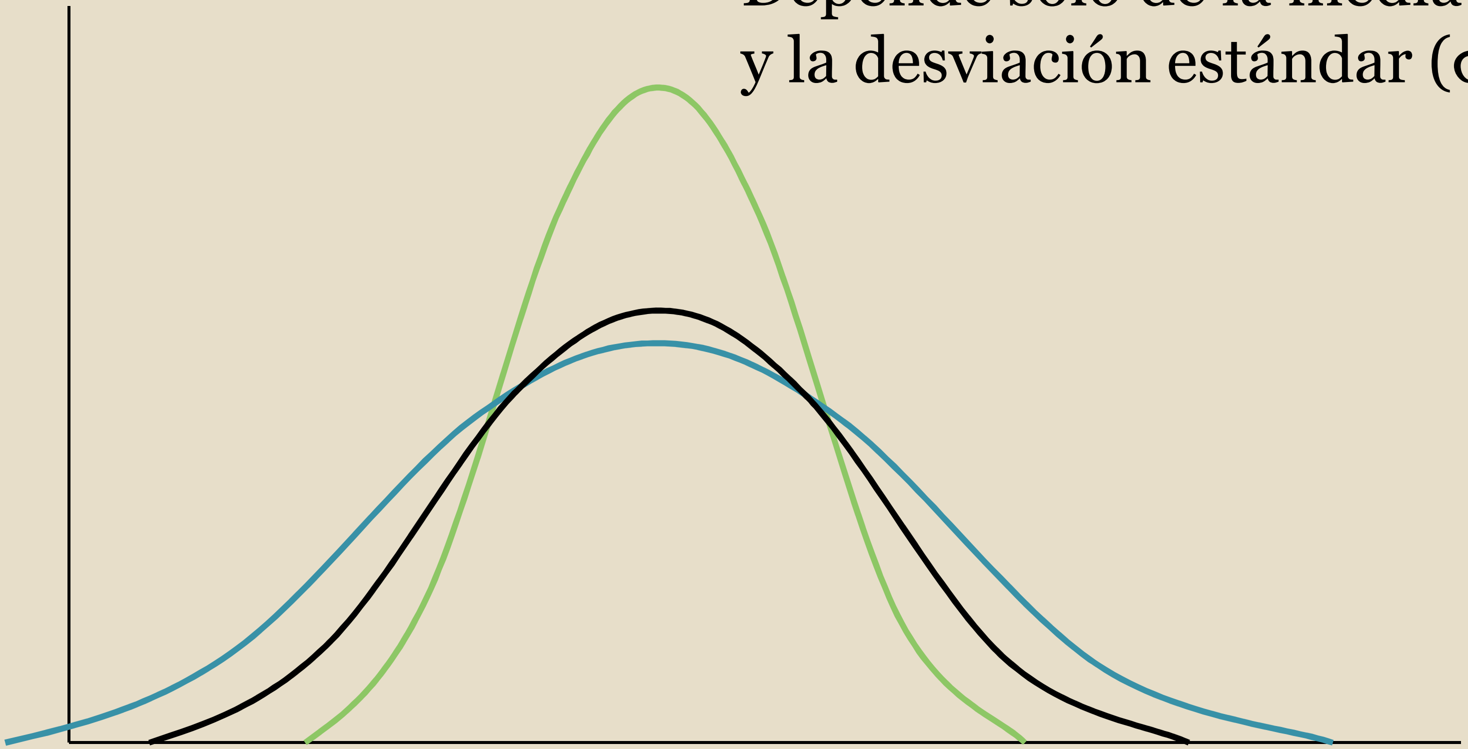


Distribución normal

Depende sólo de la media (μ)
y la desviación estándar (σ)

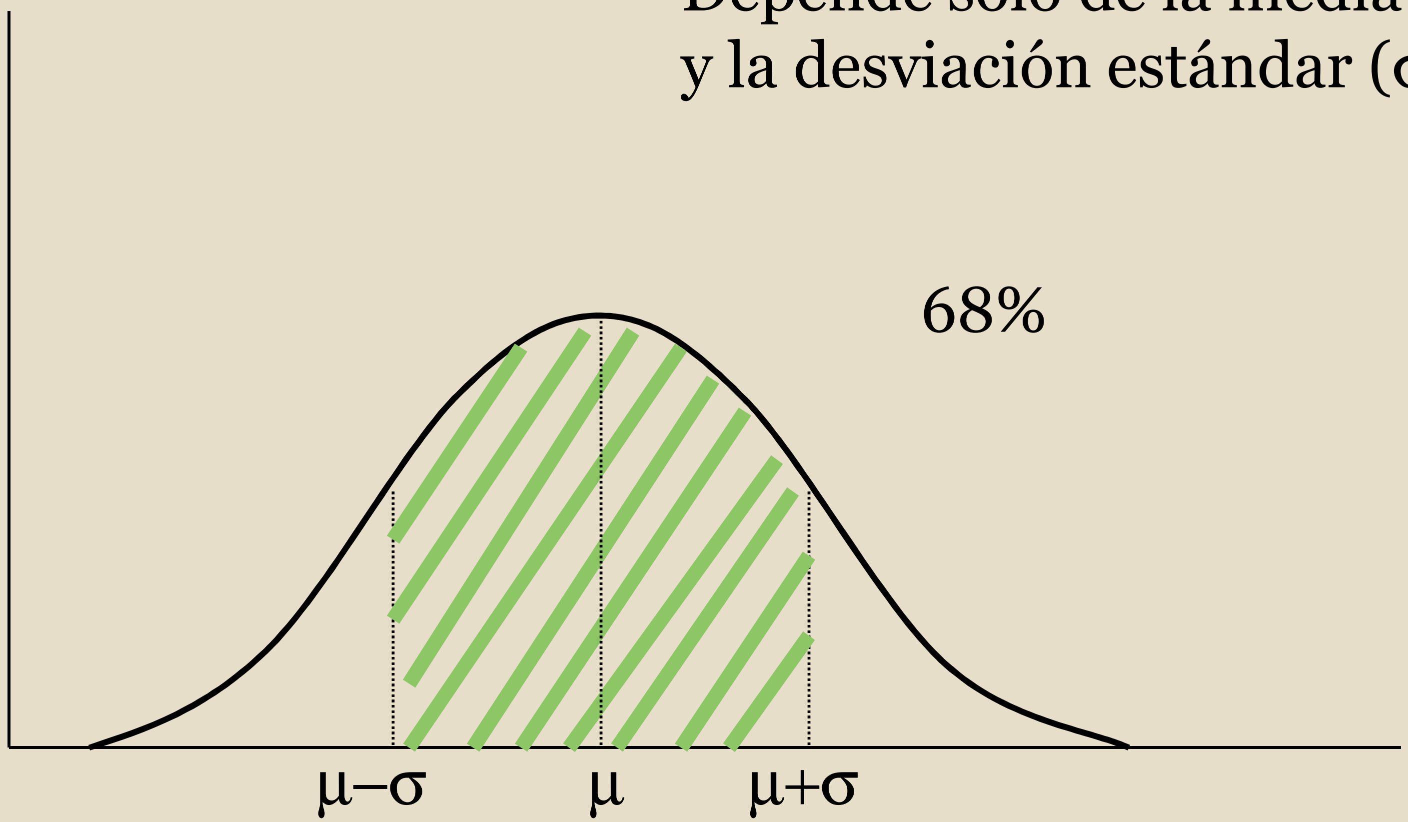
Distribución normal

Depende sólo de la media (μ)
y la desviación estándar (σ)



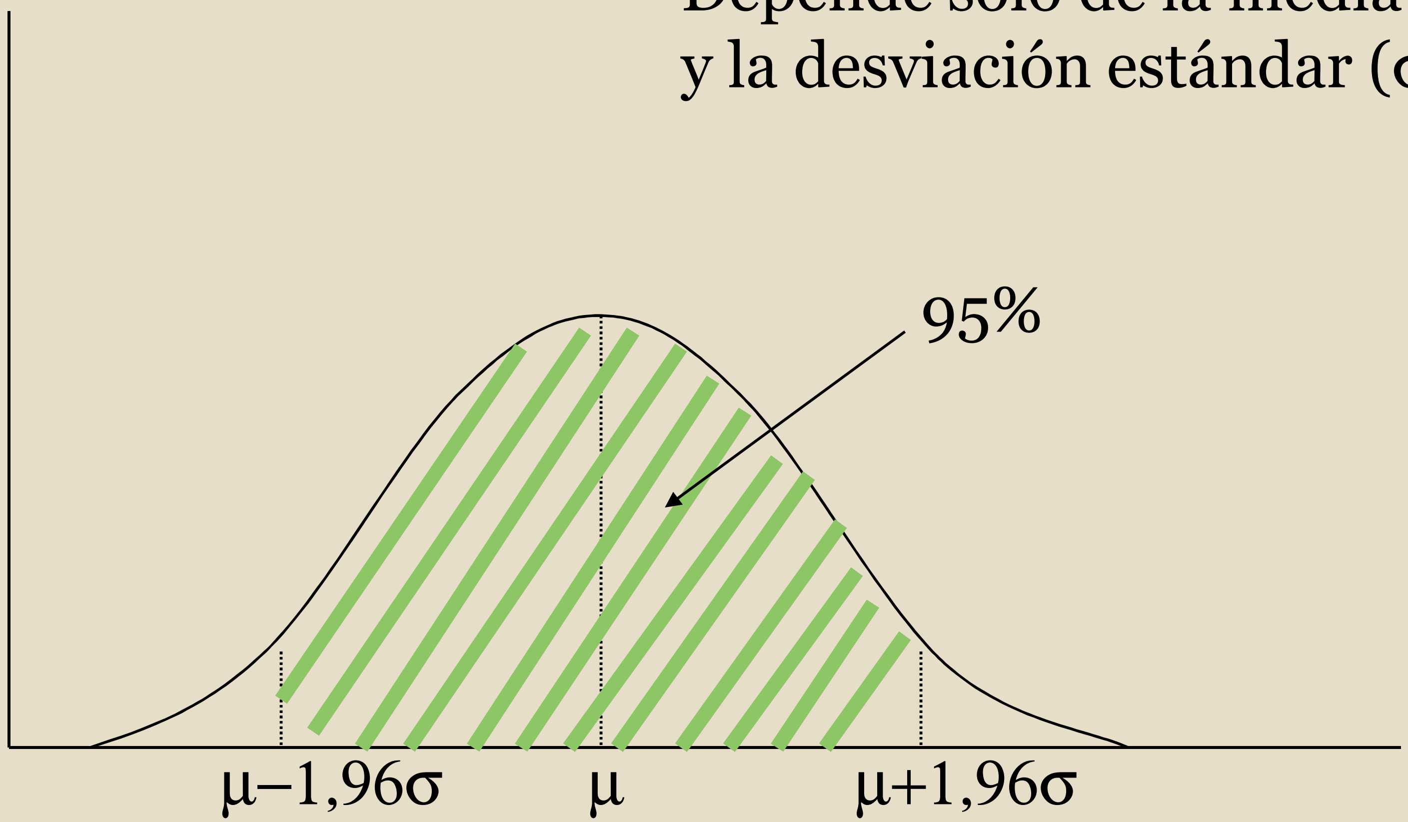
Distribución normal

Depende sólo de la media (μ)
y la desviación estándar (σ)



Distribución normal

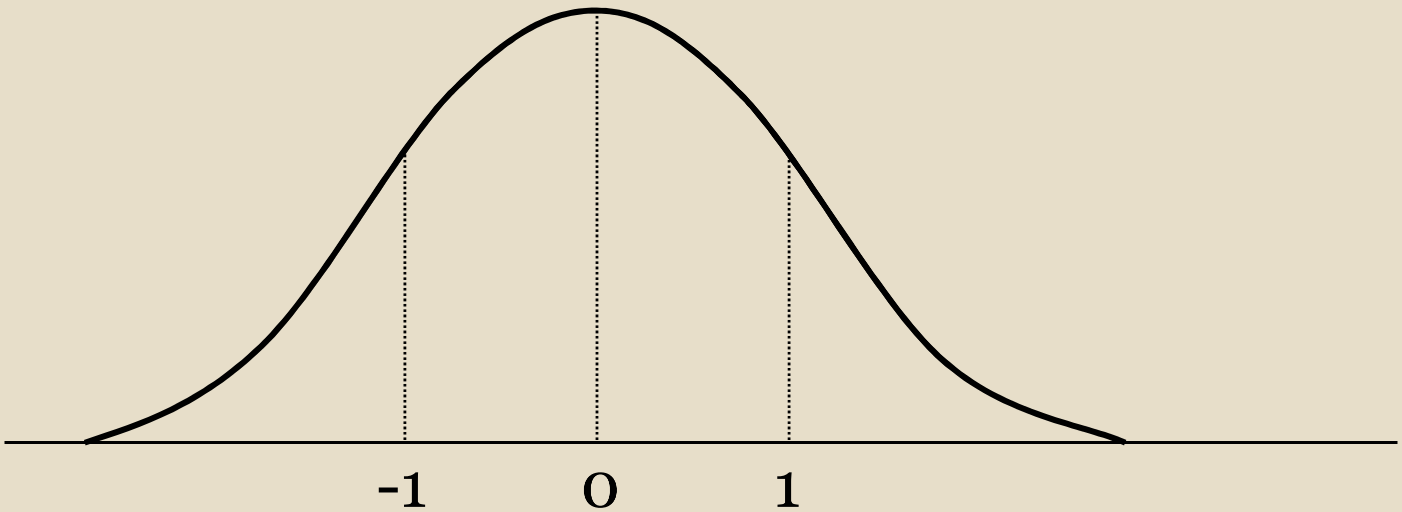
Depende sólo de la media (μ) y la desviación estándar (σ)



Distribución normal estándar (tipificada)

Ver Tabla

$\mu = 0$
$\sigma = 1$



Distribución normal: importancia

- Distribución de errores.
- Distribución límite de la binomial y la de Poisson cuando el tamaño muestral es grande.
- Teorema central del límite

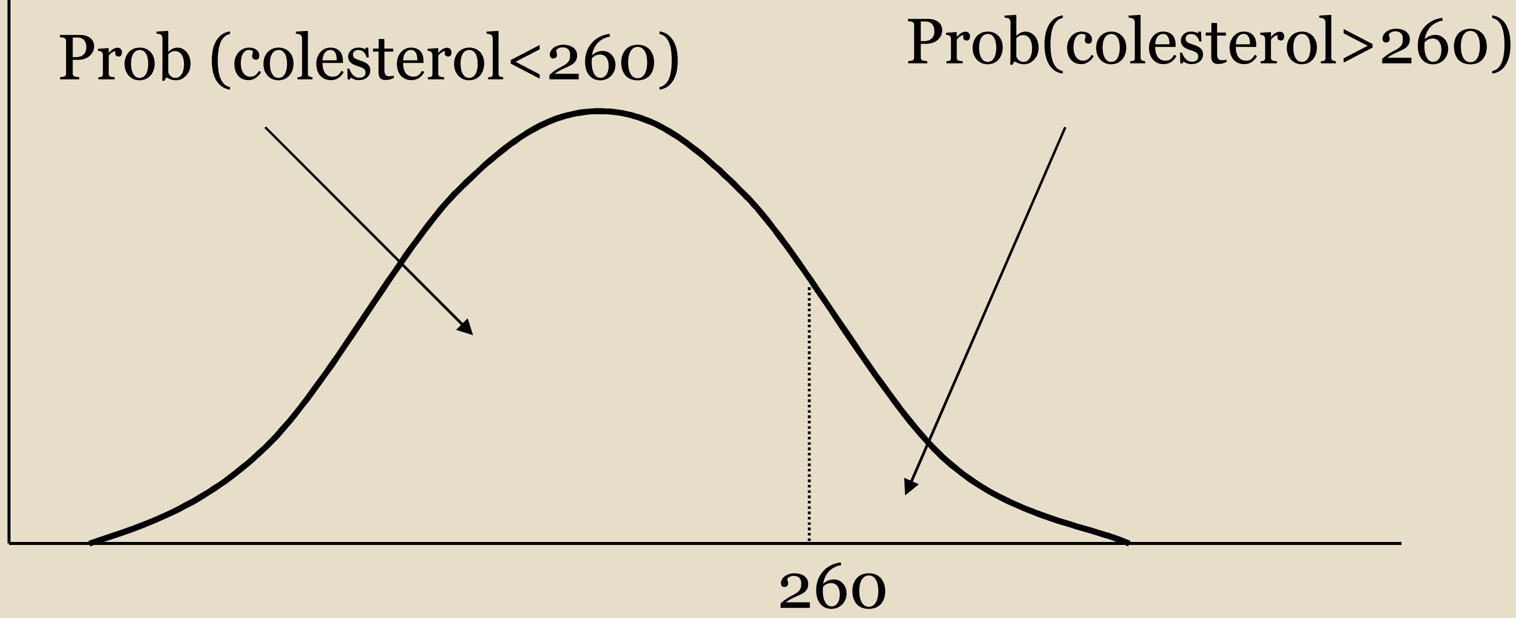
Distribución normal

- Teorema central del límite:

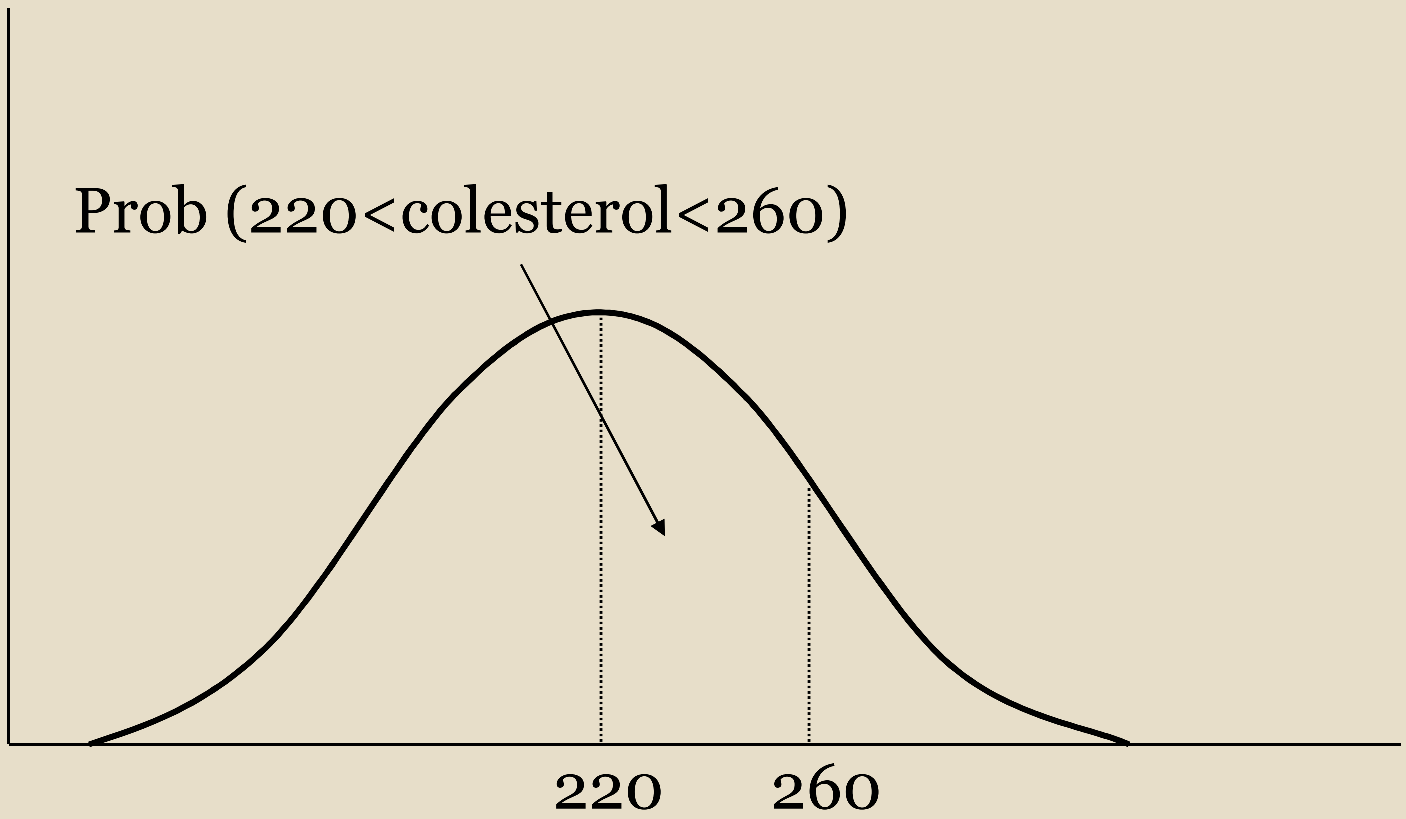
  Sea una variable NO necesariamente normal, de media μ y desviación estándar σ. La media en el muestreo es normal de media μ y desviación estándar (error estándar de la media) $\sigma / n^{1/2}$

Distribución normal

Distribución del colesterol sérico en una población adulta

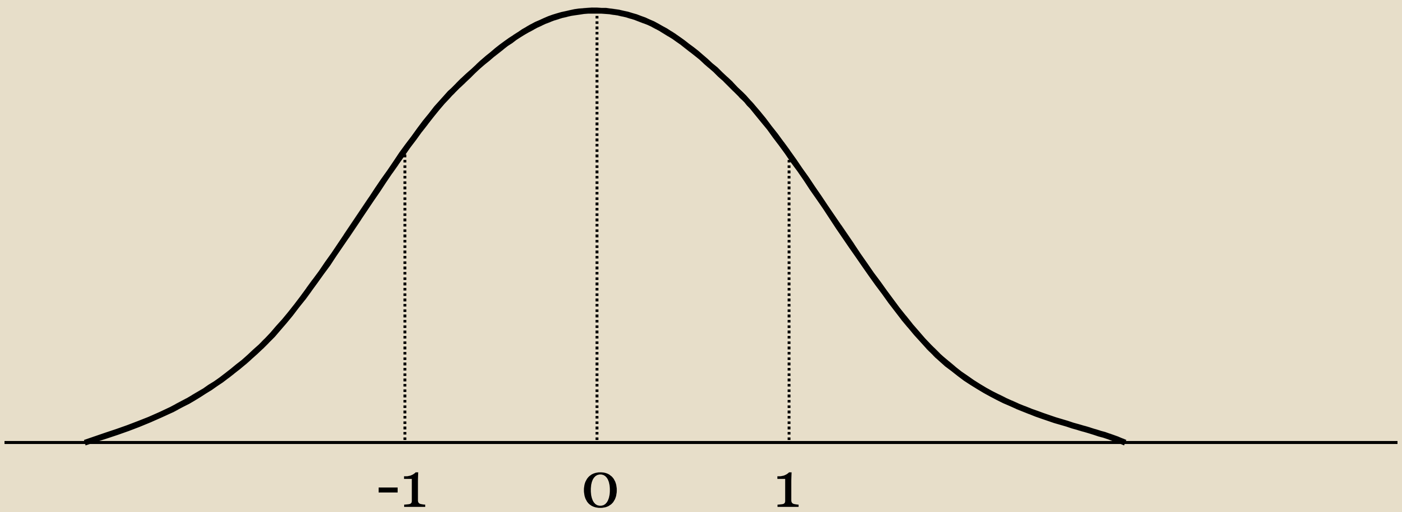


Distribución normal



Distribución normal estándar (tipificada)

$\mu = 0$
$\sigma = 1$



Uso de la tabla de la distribución normal

- ¿Cómo se responde a la siguiente pregunta?

  El colesterol se mide en una población con media 196 mg/dL y desviación estándar 46 mg/dL. ¿Cuál es la probabilidad de que un individuo tenga <166 mg/dL?

Uso de la tabla de la distribución normal

- ¿Cómo se responde a la siguiente pregunta?

  El colesterol se mide en una población con media 196 mg/dL y desviación estándar 46 mg/dL. ¿Cuál es la probabilidad de que un individuo tenga <166 mg/dL?

1. Calcular el valor z
2. Consultar en la tabla la probabilidad correspondiente a z:

   http://es.wikibooks.org/wiki/Tablas_estad%C3%ADsticas/Distribuci%C3%B3n_normal

¿Cómo se transforma una distribución normal en la distribución normal estándar?

- Cálculo de valores estándar (z):
  - Sustraer la media del valor original.
  - Dividir entre la desviación estándar.
  - El resultado (z) se consulta en la tabla de la distribución normal estándar para conocer la probabilidad.

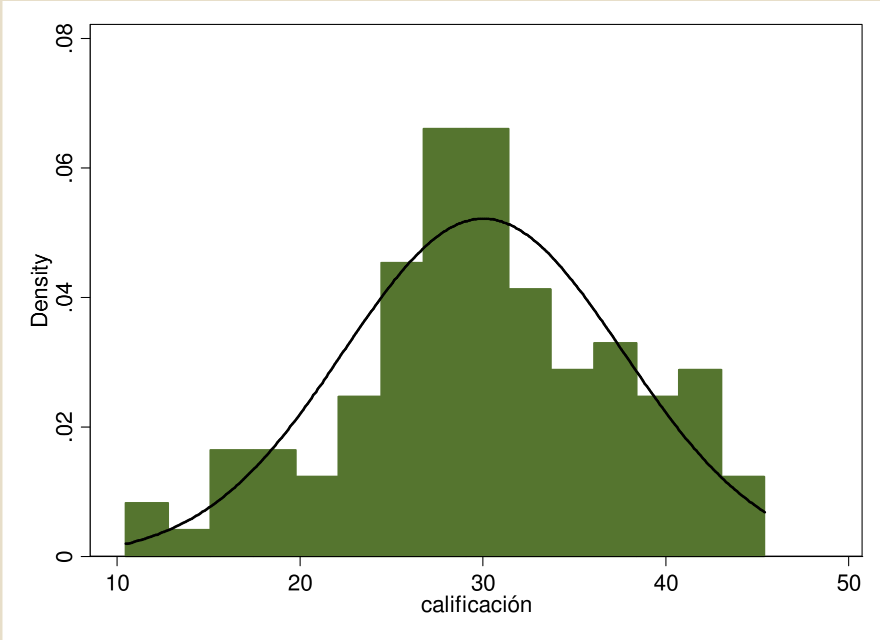
¿Cómo se transforma una distribución normal en la distribución normal estándar?

- Cálculo de valores estándar (z). Ejemplo:
- El colesterol se mide en una población con media 196 mg/dL y desviación estándar 46 mg/dL. ¿Cuál es la probabilidad de que un individuo tenga <166 mg/dL?
  - 166 – 196 = -30
  - -30 / 46 = -0,65
  - Prob(x>-0,65) = Prob(x<0,65) = 0,742
  - Prob(x<-0,65) = 1 – Prob(x>-0,65) = 1-0,742 = 0,258

Pruebas de normalidad

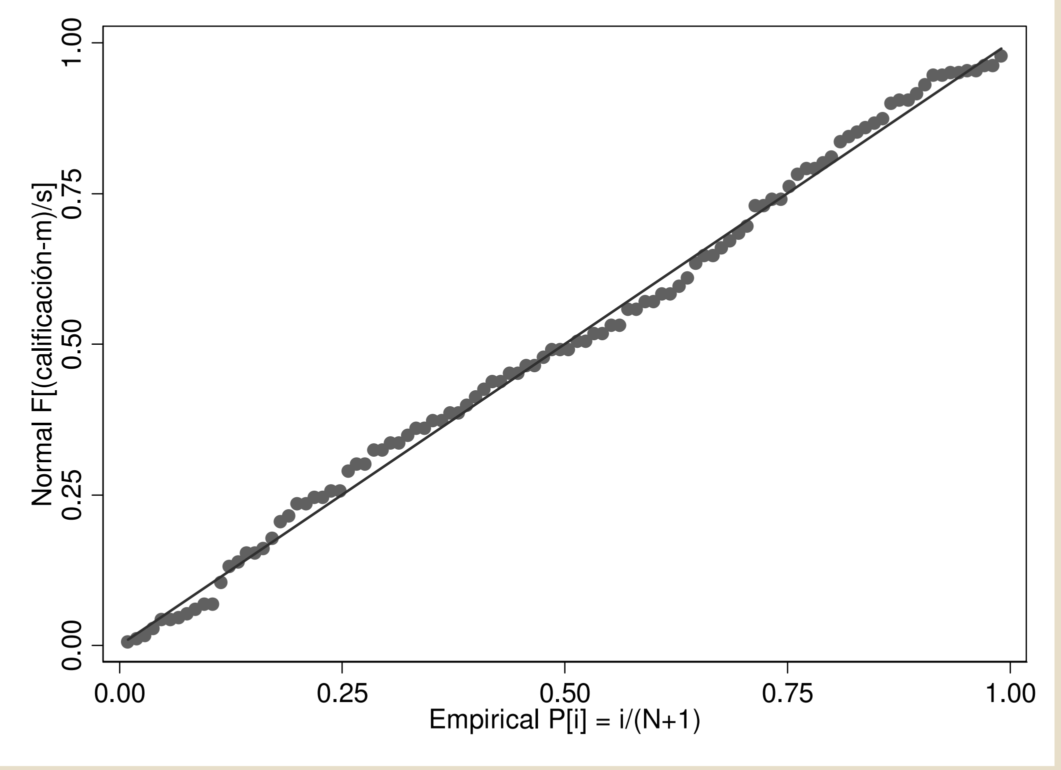
- Test de Shapiro-Wilk
- Test de Shapiro-Francia
- Test de Kolmogorov-Smirnov

Pruebas de normalidad



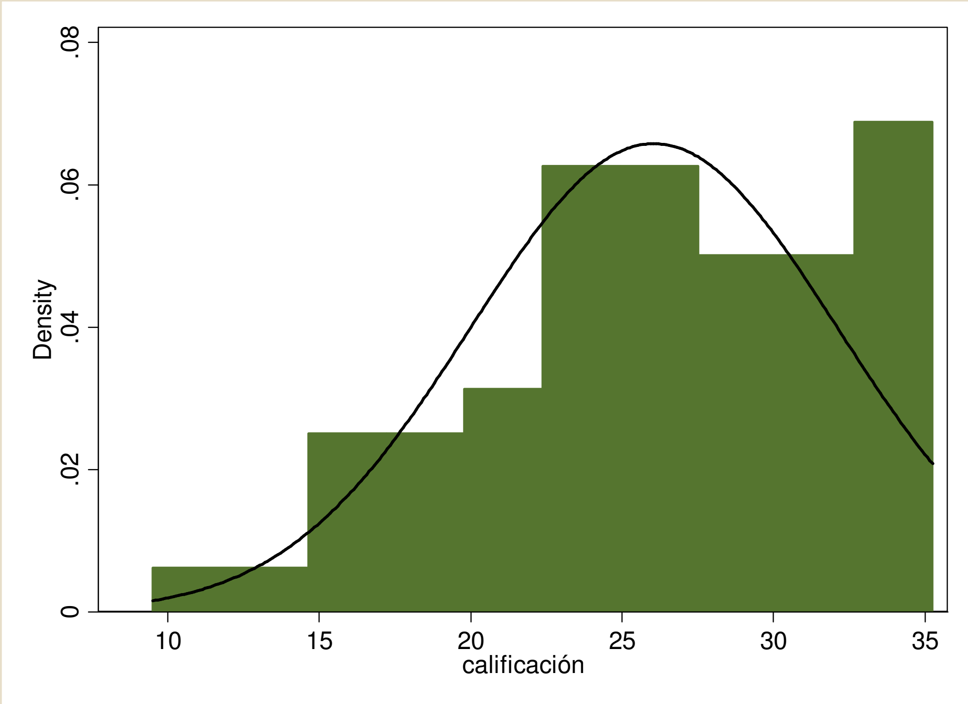
Resultados en el test de Epidemiología, junio 2009/10

Pruebas de normalidad



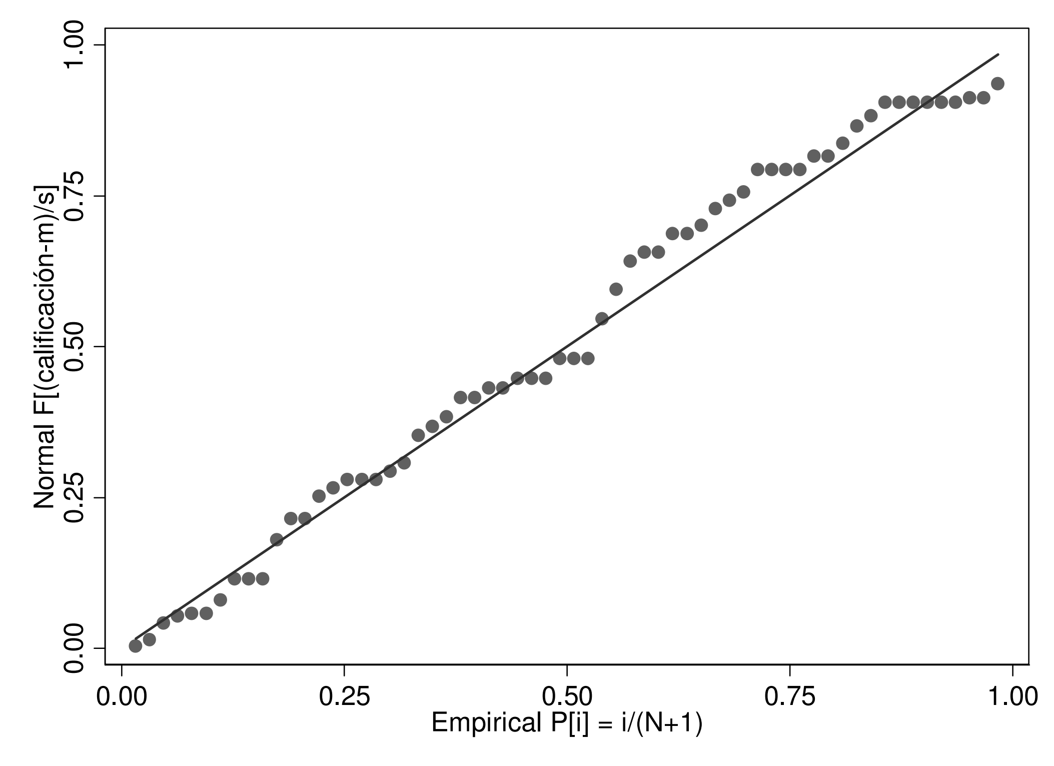
Resultados en el test de Epidemiología, junio 2009/10

Pruebas de normalidad



Resultados en el test de M. Preventiva, junio 2009/10

Pruebas de normalidad



Resultados en el test de M. Preventiva, junio 2009/10

Índice

- Probabilidad
  - Definición de probabilidad
  - Probabilidad condicionada
  - Sucesos independientes
  - Teorema de Bayes
- Distribuiciones de probabilidad
  - Distribución normal
  - Distribución binomial
  - Distribución de Poisson

Distribución binomial

- Variable "Recién nacido": Varón = 1, Mujer = 0.
- Probabilidad de que sea varón: $p = 0{,}52$
- Probabilidad de que sea mujer: $1 – p = 0{,}48$

- ¿Probabilidad de que de dos r.n. los dos sean varones?
- VV VM MV MM
- $P^2 = 0{,}52^2 = 0{,}2704$

Distribución binomial

- Probabilidad de que de tres r.n. dos sean varones y uno mujer:

VVV VVM VMV MVV
VMM MVM MMV MMM

$$3 \cdot P^2 (1 - P) = 3 \cdot 0{,}52^2 \cdot 0{,}48 = 0{,}39$$

Distribución binomial

- Probabilidad de que de tres r.n. dos sean varones y uno mujer:

VVV	**P^3**	**0,14**
VVM	**$P^2 (1 – P)$**	**0,13**
VMV	**$P^2 (1 – P)$**	**0,13**
MVV	**$P^2 (1 – P)$**	**0,13**
VMM	**$P (1 – P)^2$**	**0,12**
MVM	**$P (1 – P)^2$**	**0,12**
MMV	**$P (1 – P)^2$**	**0,12**
MMM	**$(1 – P)^3$**	**0,11**

Distribución binomial

- Probabilidad de que de tres r.n. dos sean varones y uno mujer:

VVV	P^3	0,14
VVM	$P^2 (1 - P)$	0,13
VMV	$P^2 (1 - P)$	0,13
MVV	$P^2 (1 - P)$	0,13
VMM	$P (1 - P)^2$	0,12
MVM	$P (1 - P)^2$	0,12
MMV	$P (1 - P)^2$	0,12
MMM	$(1 - P)^3$	0,11

0,39

Distribución binomial

- Una enfermedad tiene prevalencia *p*. Si seleccionamos al azar 100 personas, ¿cuál es la probabilidad de que haya 10 con la enfermedad?

$$\binom{100}{10} p^{10} \times (1-p)^{90}$$

Distribución binomial

Distribución binomial de parámetros *n* y *p*:

- En una variable dicotómica la probabilidad de obtener el resultado A es p.
- Se repite el experimento n veces.
- ¿Cuál es la probabilidad de que ocurran x resultados A?

$$\binom{n}{x} p^x \times (1-p)^{n-x}$$

Distribución binomial

Distribución binomial de parámetros *n* y *p*:

- En una variable dicotómica la probabilidad de obtener el resultado A es *p*.
- Se repite el experimento *n* veces.
- ¿Cuál es la probabilidad de que ocurran *x o menos* resultados A?

$$\sum_{i=0}^{x} \binom{n}{i} p^{i} \times (1-p)^{n-i}$$

Distribución binomial

Distribución binomial de parámetros *n* y *p*:

- En una variable dicotómica la probabilidad de obtener el resultado A es p.
- Se repite el experimento n veces.
- ¿Cuál es la probabilidad de que ocurran <u>*más de x*</u> resultados A?

$$1-\sum_{i=0}^{x}\binom{n}{i}p^{i}\times(1-p)^{n-i}$$

Distribución binomial

Distribución binomial de parámetros *n* y *p*:

- En una variable dicotómica la probabilidad de obtener el resultado A es p.
- Se repite el experimento n veces.
- ¿Cuál es el número medio de resultados A?

$$\mu = np$$

Distribución binomial

<u>Distribución binomial de parámetros *n* y *p*:</u>

- En una variable dicotómica la probabilidad de obtener el resultado A es *p*.
- Se repite el experimento *n* veces.
- ¿Cuál es la varianza de resultados A?

$$\sigma^2 = np(1 - p)$$

Distribución binomial

Distribución binomial de parámetros *n* y *p*:

- En una variable dicotómica la probabilidad de obtener el resultado A es *p*.
- Se repite el experimento *n* veces.
- ¿Cuál es la desviación típica de resultados A?

$$\sigma = \sqrt{np(1-p)}$$

Distribución binomial

Distribución binomial de parámetros *n* y *p*:

- En las familias de 3 hijos, ¿cuál es el número medio de varones? (nota: p(varón)=0,52)
- Es una binomial con n=3, p = 0,52

$$\mu = np = 3 \times 0,52 = 1,56$$

$$\sigma^2 = np(1-p) = 3 \times 0,52 \times 0,48 = 0,7488$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{3 \times 0,52 \times 0,48} = 0,865$$

Distribución binomial

- Se realiza un examen de test con 50 preguntas, con 5 respuestas alternativas de las que sólo 1 es correcta.
- ¿Cuál es la probabilidad de que un alumno que no tiene ni idea acierte *exactamente 25 preguntas*?

Distribución binomial

- Se realiza un examen de test con 50 preguntas, con 5 respuestas alternativas de las que sólo 1 es correcta.
- ¿Cuál es la probabilidad de que un alumno que no tiene ni idea acierte <u>*exactamente 25 preguntas*</u>?
- Se trata de una binomial con n=50 y p=1/5=0,2

$$\binom{50}{25} \times 0,2^{25} \times (1-0,2)^{50-25} = 0.000002$$

Distribución binomial

- Se realiza un examen de test con 50 preguntas, con 5 respuestas alternativas de las que sólo 1 es correcta.
- ¿Cuál es la probabilidad de que un alumno que no tiene ni idea acierte *25 preguntas o menos*?

Distribución binomial

- Se realiza un examen de test con 50 preguntas, con 5 respuestas alternativas de las que sólo 1 es correcta.
- ¿Cuál es la probabilidad de que un alumno que no tiene ni idea acierte *25 preguntas o menos*?
- Se trata de una binomial con n=50 y p=1/5=0,2

$$\sum_{i=0}^{25} \binom{50}{i} \times 0,2^{i} \times (1-0,2)^{50-i} = 0,999998$$

Distribución binomial

- Se realiza un examen de test con 50 preguntas, con 5 respuestas alternativas de las que sólo 1 es correcta.
- ¿Cuál es la probabilidad de que un alumno que no tiene ni idea acierte <u>*más de 25 preguntas*</u>?
- Se trata de una binomial con n=50 y p=1/5=0,2

Distribución binomial

- Se realiza un examen de test con 50 preguntas, con 5 respuestas alternativas de las que sólo 1 es correcta.
- ¿Cuál es la probabilidad de que un alumno que no tiene ni idea acierte *más de 25 preguntas*?
- Se trata de una binomial con n=50 y p=1/5=0,2

$$1-\sum_{i=0}^{25}\binom{50}{i}\times 0,2^{i}\times(1-0,2)^{50-i}=0,000002$$

Distribución binomial

- En el mismo examen, un alumno ha respondido (gracias a sus extensos conocimientos) a 40 preguntas. Decide contestar al azar las otras 10.
- ¿Cuál es la probabilidad de que acierte *2 preguntas o menos*?
- Se trata de una binomial con $n=10$ y $p=1/5=0,2$

Distribución binomial

- En el mismo examen, un alumno ha respondido (gracias a sus extensos conocimientos) a 40 preguntas. Decide contestar al azar las otras 10.
- ¿Cuál es la probabilidad de que acierte *2 preguntas o menos*?
- Se trata de una binomial con n=10 y p=1/5=0,2

$$\sum_{i=0}^{2} \binom{10}{i} \times 0,2^{i} \times (1-0,2)^{10-i} = 0,678$$

Distribución binomial

- Forma de la distribución binomial

Índice

- Probabilidad
  - Definición de probabilidad
  - Probabilidad condicionada
  - Sucesos independientes
  - Teorema de Bayes
- Distribuiciones de probabilidad
  - Distribución normal
  - Distribución binomial
  - Distribución de Poisson

Distribución de Poisson

- Combinación de sucesos:
  - Binomiales
  - Independientes
  - Con probabilidad constante en:
    - El tiempo
    - El espacio

Distribución de Poisson

- En una masa radiactiva se producen 20 desintegraciones de media por minuto.
- ¿Cuál es la probabilidad de que en un minuto determinado se produzcan exactamente 15 desintegraciones?

$$P(x = 15 \mid \lambda = 20) = \frac{e^{-20} \times 20^{15}}{15!} = 0,0516$$

Distribución de Poisson

- En una población, el número medio de casos de leucemia es 5 al año.
- En 2011, se producen 8 leucemias.
- ¿Se trata de un resultado anormal?

$$P(x=8 \mid \lambda=5)=\frac{e^{-5}\times 5^{8}}{8!}=0,0653$$

$$P(x\geq 8 \mid \lambda=5)=1-\sum_{i=0}^{7}\frac{e^{-5}\times 5^{i}}{i!}=0,1334$$

Distribución de Poisson

- Distribución de Poisson de parámetro λ:

$$P(x \mid \lambda) = \frac{e^{-\lambda} \times \lambda^{x}}{x!}$$

$$P(x < X \mid \lambda) = \sum_{i=0}^{X-1} \frac{e^{-\lambda} \times \lambda^{i}}{i!}$$

$$P(x \geq X \mid \lambda) = 1 - \sum_{i=0}^{X-1} \frac{e^{-\lambda} \times \lambda^{i}}{i!}$$

Distribución de Poisson

- Distribución de Poisson de parámetro λ

$$\mu = \lambda$$

$$\sigma^2 = \lambda$$

$$\sigma = \sqrt{\lambda}$$

Distribución de Poisson

- En el cromosoma 5 se producen 4 mutaciones al año. ¿Cuál será la media y dispersión de las mutaciones en la población?

$$\mu = \lambda = 4$$

$$\sigma^2 = \lambda = 4$$

$$\sigma = \sqrt{\lambda} = 2$$

Distribución de Poisson

- Cada año se producen en Cantabria 4 casos de infección neumocócica invasiva en niños.
- En 2011 se pone en marcha un programa de vacunación.
- Al año siguiente, se producen sólo 2 casos.
- ¿Puede hablarse de éxito de la vacuna?

Distribución de Poisson

- Habitual: 4 casos/año
- Un año después de la vacuna: 2 casos
- Antes de poner la vacuna, ¿cuál era la probabilidad de que un año hubiera 2 casos o ninguno?

$$\left.\begin{aligned} P(x=2\,|\,\lambda=4) &= \frac{e^{-4}\times 4^{2}}{2!} = 0,1465 \\ P(x=1\,|\,\lambda=4) &= \frac{e^{-4}\times 4^{1}}{1!} = 0,0733 \\ P(x=0\,|\,\lambda=4) &= \frac{e^{-4}\times 4^{0}}{0!} = 0,0183 \end{aligned}\right\} P(x\leq 2\,|\,\lambda=4) = 0,2381$$

Distribución de Poisson

- ¿Cuántos leucocitos hay en un campo?
- ¿Cuántos hematíes hay en un ml de sangre?
- ¿Cuántos enfermos se atienden en urgencias cada día?
- ¿Cuántos infartos se producen en un año?
- ¿Cuántos reemplazos valvulares sobreviven cada año?
- ¿Cuántos trasplantes hepáticos sobreviven al cabo de cinco años?
- ¿Cuántas mutaciones se producen en un cromosoma en 50 años?

Aproximación a la binomial usando Poisson

- En una distribución binomial de parámetros n y p, n es muy grande y p es pequeño, entonces las probabilidades de la binomial pueden aproximarse con las probabilidades de una distribución de Poisson de parámetro $\lambda = np$

$$P(x \mid n, p) = \binom{n}{x} p^x \times (1-p)^{n-x} \approx \frac{e^{-np}(np)^x}{x!}$$

Aproximación a la binomial usando Poisson

- La probabilidad de que una persona porte el gen de la fibrosis quística es 0,001.
- En una población de 10000 habitantes, ¿cuál es la probabilidad de encontrar 5 portadores de este gen?

$$P(5 \mid 10000; 0,001) = \binom{10000}{5} \times 0,001^5 \times 0,999^{10000-5} \approx$$

$$\frac{e^{-10000 \times 0,001} \times (10000 \times 0,001)^5}{5!} = \frac{e^{-10} \times 10^5}{5!} = 0,0378$$

Aproximación a la binomial usando Poisson

- ¿Qué tal se aproxima la binomial con Poisson?

Aproximación a binomial usando normal

- Una variable binomial de parámetros *n* y *p*, si n es grande, tiende a una distribución normal con $\mu=np$ y $\sigma^2=np(1-p)$

$$\lim_{n\to\infty} B(n;p) = N(np, np(1-p))$$

Aproximación a Poisson usando normal

- ¿Qué tal se aproxima una distribución de Poisson usando una normal?

$$\lim_{n \to \infty} P(\lambda) = N(\lambda, \lambda)$$

Distribución binomial

- Se realiza un examen de test con 50 preguntas, con 5 respuestas alternativas de las que sólo 1 es correcta.
- ¿Cuál es la probabilidad de que un alumno que no tiene ni idea acierte <u>*25 preguntas o más*</u>?
- Se trata de una binomial con n=50 y p=1/5=0,2

$$1-\sum_{i=0}^{25}\binom{50}{i}\times 0,2^{i}\times(1-0,2)^{50-i}=0,000002$$

Distribución binomial

- Se trata de una binomial con n=50 y p=1/5=0,2
- Tiende a una normal:

$$N(np, np(1-p)) = N(10, 8)$$

- Calcular el valor z:

$$z = \frac{25-10}{\sqrt{8}} = 5,3033$$

- Mirar en la tabla normal: Prob(x>z)

$$\mathrm{P}(x > 5,3033) = 0,00000006$$

Enlaces de interés

- Wikipedia:
  - Probabilidad condicionada
  - Distribución normal
  - Tabla de la distribución normal
  - Distribución binomial
  - Distribución de Poisson
  - Teorema de Bayes