Bioestadística

REPASO PRIMER PARCIAL

José Aurelio Pina Romero
Ja.pina@ua.es
Bioestadística – Grado Enfermería
UA- Departamento de Enfermería

Universitat d'Alacant
Universidad de Alicante

CUESTIONES – V/F

1) El nivel de glucosa en sangre (mg/dl) de un grupo de pacientes mayores de 65 años es una variable cualitativa.
2) El nivel de glucosa en sangre (mg/dl) de un grupo de pacientes mayores de 65 años es una variable cuantitativa continua.
3) Los diagramas de barras se utiliza para representar datos de variables cualitativas o discretas.
4) Un histograma se utiliza para representar datos de variables cuantitativas continuas.
5) Un diagrama de sectores se utiliza para representar datos de variables cuantitativas discretas.

Clasificación OMS	Menopausia	
	No	si
Normal	189	280
Osteopenia	108	359
Osteoporosis	6	58

Se han elegido a 1000 mujeres de una población. Y tenemos los siguientes resultados.

1) La probabilidad de tener osteoporosis y ser menopaúsica es 58/1000
2) La probabilidad de tener osteoporosis si es menopaúsica es de 58/1000
3) La probabilidad de ser menopaúsica o tener osteoporosis es 761/100
4) La probabilidad de ser menopaúsica o tener osteoporosis es 703/100

Pruebas diagnósticas

	Resultados de la enfermedad	
Resultados de la prueba	E	$\bar{E}$
+	a	b
-	c	d

Sensibilidad = P(+ | Enfermo)= a/(a+c)
Falsos negativos = P(- | E) = c / (a+c)

Especificidad = P(- | Sano) = d / (b+d)
Falsos positivos = P(+ I Sano) = b / (b+d)

Valor predictivo positivo(VPP)=P(Enfermo|+) = a/(a+b)
Valor predictivo negativo(VPN) = P(Sano | -) = d /(c+d)

CUESTIONES – V/F

1) El falso positivo de una prueba diagnóstica es la probabilidad de que la prueba proporciones un resultado positivo sobre un individuo sano.
2) Falso positivo = 1- Especificidad
3) Falso negativo = 1- Sensibilidad
4) Falso positivo = 1 – Sensibilidad
5) Falso negativo = 1 – Especificidad
6) A mayor prevalencia de una patología en una determinada población, mayor valor predictivo positivo.
7) A menor prevalencia de una patología en una determinada población, mayor valor predictivo negativo.

$$p(E/+) = \frac{p(+/E)\,.\,p(E)}{p(+)} = \frac{p(+/E)\,.\,p(E)}{p(+/E)\,.\,p(E) + p(+/\bar{E})\,.\,p(\bar{E})}$$

o, equivalentemente:

$$\text{Valor predictivo positivo} = \frac{\text{Sensibilidad} \times \text{Prevalencia}}{\text{Sensibilidad} \times \text{Prevalencia} + (1 - \text{Especificidad}) \times (1 - \text{Prevalencia})}$$

y

$$p(\bar{E}/-) = \frac{p(-/\bar{E})\,.\,p(\bar{E})}{p(-)} = \frac{p(-/\bar{E})\,.\,p(\bar{E})}{p(-/\bar{E})\,.\,p(\bar{E}) + p(-/E)\,.\,p(E)}$$

o, equivalentemente:

$$\text{Valor predictivo negativo} = \frac{\text{Especificidad} \times (1 - \text{Prevalencia})}{\text{Especificidad} \times (1 - \text{Prevalencia}) + (1 - \text{Sensibilidad}) \times \text{Prevalencia}}$$

Recordemos que los datos disponibles eran:

ESPACIO MUESTRAL ENFERMEDAD-PRUEBA		Resultados en enfermedad		Total
		E	_	
Resultados	+	190	120	310
de la prueba	-	10	680	690
diagnóstica	Total	200	800	1000

por lo que tendremos:

$$\text{Sensibilidad} = p(+ / E) = \frac{190}{200} = 0{,}95$$

$$\text{Especificidad} = p(- / \bar{E}) = \frac{680}{800} = 0{,}85$$

$$\text{Falso positivo} = p(+ / \bar{E}) = \frac{120}{800} = 0{,}15$$

$$\text{Falso negativo} = p(- / E) = \frac{10}{200} = 0{,}05$$

Los valores predictivos dependerán de la prevalencia, que en este caso se estima a partir del dato frecuencial 0,5%, como p(E) = 0,005. En ese caso tendremos:

$$\text{Valor predictivo} + = p(E / +) = \frac{0{,}95 \,.\, 0{,}005}{0{,}95 \,.\, 0{,}005 + 0{,}15 \,.\, 0{,}995} = 0{,}0308$$

$$\text{Valor predictivo} - = p(\bar{E} / -) = \frac{0{,}85 \,.\, 0{,}995}{0{,}85 \,.\, 0{,}995 + 0{,}05 \,.\, 0{,}005} = 0{,}9997$$

$$\text{Valor predictivo} + = p(E / +) = \frac{0{,}95\ .\ 0{,}30}{0{,}95\ .\ 0{,}30 + 0{,}15\ .\ 0{,}70} = 0{,}7308$$

$$\text{Valor predictivo} - = p(\bar{E} / -) = \frac{0{,}85\ .\ 0{,}70}{0{,}85\ .\ 0{,}70 + 0{,}05\ .\ 0{,}30} = 0{,}9754$$

CUESTIONES – V/F

1) En el modelo de probabilidad Normal la media no coincide con la mediana, ni con la moda.
2) En el modelo de probabilidad Normal la media, la mediana y la moda coinciden.
3) La función de probabilidad de una variable aleatoria discreta no puede tomar valores negativos.
4) La función de densidad de una variable aleatoria continua nos permite conocer la probabilidad de la variable tome valores entre cualesquiera dos valores.
5) Si $A \cap B \neq 0 \rightarrow P(A \cap B) = 0$
6) En un modelo de probabilidad Normal (170, 10), la probabilidad de que un sujeto mida exactamente 2 metros ($P(X=200)$) es cero.

➢ Una variable aleatoria continua, X, sigue una distribución normal de parámetros λ y σ

$$\boxed{X \to N(\mu\,;\sigma)}$$

si su función de densidad es:

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

$$\text{donde}\quad \begin{cases} -\infty < x < \infty \\ -\infty < \mu < \infty \\ \sigma > 0 \end{cases}$$

▪ Puede comprobarse que se verifica:

$$\int_{-\infty}^{+\infty} f(x)\,dx = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx = 1$$

❖ **Media**

$$\boxed{E[X] = \mu}$$

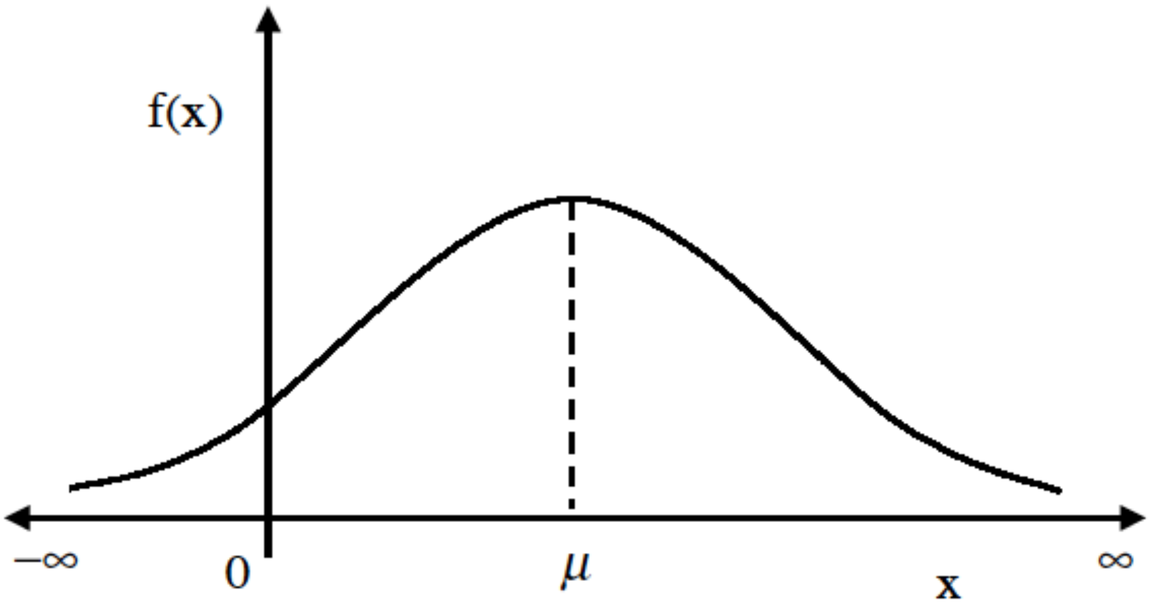
❖ **Varianza**

$$\boxed{Var[X] = \sigma^2}$$

❖ **Desviación típica**

$$\sigma = \sqrt{Var[X]}$$

◈ Campana de Gauss



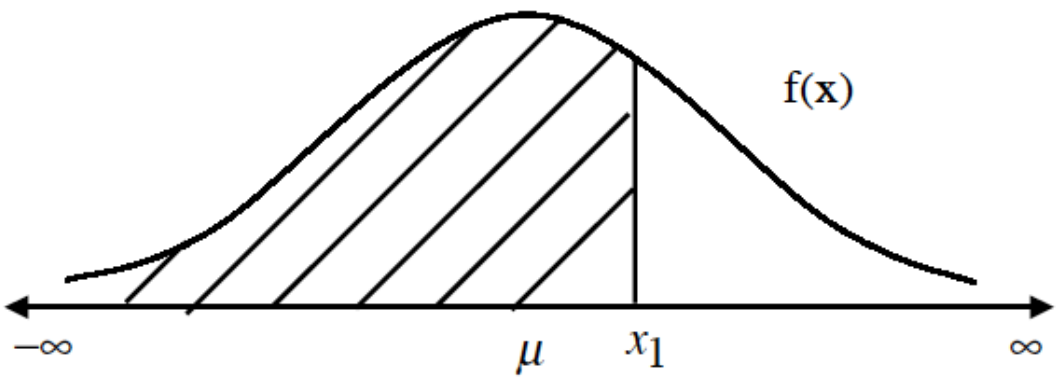
$$\int_{-\infty}^{+\infty} f(x)\,dx = 1$$

- **Se verifica:**
  - ◈ **La curva es simétrica respecto a μ**
  - ◈ **La media, la moda y la mediana coinciden**

❖ Función de distribución

$$F(x_1) = P(X \le x_1) = \int_{-\infty}^{x_1} f(x)\,dx =$$

$$= \int_{-\infty}^{x_1} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx$$



Área a la izquierda del punto *x_1*

❖ 4.1 Distribución binomial

◈ 4.1.1 Definición. Ejemplos

- Sea un experimento aleatorio en el que sólo puedan darse dos posibilidades: que ocurra un determinado suceso A, que llamaremos éxito, o que no ocurra dicho suceso, o sea que ocurra su complementario, que llamaremos fracaso, $\bar{\mathbf{A}}$.

- Se conoce la probabilidad de ocurrencia del suceso A, y por lo tanto la de su complementario:

$$P(A)=p\,;\quad P\left(\overline{A}\right)=1-p=q$$

➢ Se repite el experimento n veces en las mismas condiciones (independencia). Se define la variable aleatoria Binomial :

➢ X: “nº de veces que ocurre el suceso A (nº éxitos) en n realizaciones independientes del experimento”

✓ Por lo tanto, X: 0, 1, 2, 3, ……n

$$\boxed{X \to B(n; p)}$$

Función de probabilidad

$$P[X = k] = \binom{n}{k} p^k q^{n-k}, \; 0 \leq k \leq n$$

Media: μ = E(x) =n p

Varianza: σ^2 = V(x)= n p q

- Nº de aprobados si se presentan 80 alumnos a un examen
- Nº de familias con un solo hijo en una población de 120 familias
- Nº de reacciones negativas ante un fármaco administrado a 40 pacientes
- Nº de accidentes de tráfico si han circulado 1200 automóviles
- Nº de semillas que germinan de las 20 semillas que se han plantado en suelos de idéntica composición

❖ 4.2. Distribución de Poisson

◈ 4.2.1 Definición. Ejemplos

- Se define la variable aleatoria X como el número de sucesos que ocurren en un intervalo continuo de tiempo, longitud o espacio, de un tamaño determinado.
- Sea λ el número medio de sucesos que ocurren en estos intervalos.
- La variable aleatoria así definida sigue una distribución de Poisson de parámetro λ

Función de probabilidad

$$P[X=k]=e^{-\lambda}\frac{\lambda^k}{k!},\ k=0,1,2,\ldots$$

media =E(x)=λ

varianza= V(x)= λ

♦ Ejemplos

- Nº de leucocitos en una gota de sangre

- Nº de veces que una planta de energía nuclear emite gases radiactivos en un periodo de tres meses

PROBLEMA 1

El nivel de colesterol en una población tiene distribución normal, con media 200 y desviación típica 10. Calcule

1) La probabilidad de que una persona de la población tenga un nivel de colesterol inferior a 210.
2) La probabilidad de que una persona de la población tenga un nivel de colesterol entre 180 y 220.
3) ¿Qué valor del nivel de colesterol es superado únicamente por el 20% de los sujetos de la población?
4) Si la población es de 250 000 habitantes. ¿Cuántas personas tienen un nivel de colesterol inferior a 210?

PROBLEMA 2

Tenemos los valores de IMC (Índice de masa corporal), medido en kg/m2, de 15 pacientes:

18, 19, 18, 23, 26, 24, 21, 20, 23, 31, 29, 17, 27, 25, 16

1) Calcula la media y la mediana (P_{50})
2) Calcule la moda y el rango
3) Calcule el valor que es superado únicamente por el 30% de los sujetos.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$r_{Md} = \frac{n+1}{2}$$

$$CV = \frac{s}{\bar{x}}(x100)$$

$$p_q = (1-f)x_i + fx_{i+1}$$

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 = \frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2$$

$$S = \sqrt{S^2}$$

$$r_q = \frac{q}{100}(n+1)$$

f parte fraccionaria de r_q

PROBLEMA 3

Se sabe que el 35% de los diabéticos tipo II son tratados con insulina, si en una consulta de enfermería se dispone de 6 dosis de insulina y entran 15 sujetos.

1) Número esperado de insulinodependientes, y su desviación típica.
2) ¿Cuál es la probabilidad exacta de utilizar las seis dosis?
3) ¿Cuál es la probabilidad de utilizara entre 2 y 4 dosis.
4) ¿Cuál es la probabilidad de que se utilicen más de 6 dosis?

PROBLEMA 4

Para detectar el parásito del paludismo existe un test de respuesta inmediata que produce un 2 % de falsos positivos y un 4 % de falsos negativos. En una determinada región de África se sabe que hay un 32 % de personas con paludismo. Se pide:

1. La sensibilidad y especificidad de la prueba.
2. Los valores predictivos de la prueba.

Sensibilidad = P(+ | Enfermo)= a/(a+c)
Falsos negativos = P(- | E) = c / (a+c)

Especificidad = P(- | Sano) = d / (b+d)
Falsos positivos = P(+ I Sano) = b / (b+d)

Valor predictivo positivo(VPP)=P(Enfermo|+) = a/(a+b)
Valor predictivo negativo(VPN) = P(Sano | -) = d /(c+d)