

ESTADÍSTICA II	SEPTIEMBRE 2008. RESERVA
Código de la Carrera 65	Código de la Asignatura 201

PRIMERA PARTE: CUESTIONES TEÓRICO-CONCEPTUALES

1. En el cálculo de probabilidades de una variable discreta y continua ¿existen diferencias? Razone la respuesta.

Respuesta.-

Para una variable discreta $X = \{x_i, i=1, 2, 3, \dots\}$, existe una función de probabilidad $P[X = x_i]$, cumpliéndose que $\sum_{\forall i} P[X = x_i] = 1$. La función de distribución $F(x) = P[X \leq x] = \sum_{x_i \leq x} P[X = x_i]$.

Para una variable continua X existe una función de densidad f tal que la función de distribución $F(x) = P[X \leq x] = \int_{-\infty}^x f(x) dx$

2. Concepto de población, muestra y estadístico muestral. Ponga un ejemplo.

Respuesta.-

Población es el conjunto de individuos que son objeto del estudio estadístico, de forma que existe una variable aleatoria X asignando un número real a cada uno de los individuos de la población.

Una muestra, de tamaño n , es un conjunto de n observaciones de la variable aleatoria. $\{X_1, X_2, X_3, \dots, X_n\}$.

Un estadístico muestral es una función de las variables $X_1, X_2, X_3, \dots, X_n$ que componen la muestra.

Por ejemplo, una población puede ser el conjunto de clientes de unos grandes almacenes durante el año 2008, siendo X el gasto efectuado por cada cliente. Elegimos n clientes y anotamos sus respectivos gastos $\{X_1, X_2, X_3, \dots, X_n\}$. Un estadístico muestral sería por

ejemplo, la media muestral $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$

3. El Teorema de Chebychev, explique su significado.

Respuesta.-

Si X es una variable aleatoria, tal que $E(X) = \mu$ y $\text{Var}(X) = \sigma^2$, entonces, $\forall k > 0$ se cumple que $P[|X - \mu| \geq k] \leq \frac{\sigma^2}{k^2}$. Significa que la probabilidad de que un valor de la variable

se aleje de la media de la población una distancia superior a k , está acotada por el cociente $\frac{\sigma^2}{k^2}$, es decir, es menor cuanto menor sea la varianza o cuanto mayor sea el cuadrado de k .

4. ¿Cuál es el objetivo de un contraste de hipótesis? Razone la respuesta.

Respuesta.-

Es confirmar o rechazar una determinada hipótesis efectuada sobre el valor de un parámetro poblacional o sobre la distribución de una variable aleatoria. Para cierto estadístico, determinaremos el conjunto de valores que admitiremos para aceptar la hipótesis (región de aceptación) y el conjunto complementario (región de rechazo). Para efectuar el contraste, se elige una muestra y se calcula el valor que proporciona para el estadístico en cuestión, lo que nos permite aceptar la hipótesis o bien rechazarla y, en su caso, aceptar una hipótesis alternativa.

PROBLEMAS

1.- Un estudio realizado establece que el número de hijos que tienen las parejas de 25-35 años viene dado por:

$$P(X = x, Y = y) = K(x + y) \quad x = 1, 2, 3 \quad e \quad y = 1, 2$$

- a) Calcular el valor de la constante K .
- b) Poner las probabilidades en forma de tabla de doble entrada y hallar las correspondientes distribuciones de probabilidad marginales.
- c) Calcular las distribuciones de probabilidad condicionadas para ambas variables.

Solución.-

a) Debe cumplirse que $1 = \sum_{\forall x, y} K(x + y) = K[(1+1)+(1+2)+(2+1)+(2+2)+(3+1)+(3+2)] =$

$= 21K$, luego $K = \frac{1}{21}$.

b)

	Y	1	2	P_i = P[X=x]
X				
1		$\frac{2}{21}$	$\frac{3}{21}$	$\frac{5}{21}$
2		$\frac{3}{21}$	$\frac{4}{21}$	$\frac{7}{21}$
3		$\frac{4}{21}$	$\frac{5}{21}$	$\frac{9}{21}$
P_j = P[Y=y]		$\frac{9}{21}$	$\frac{12}{21}$	

Las funciones de probabilidad marginales serían:

$$P_i = P[X=x] = \sum_{\forall y} P[X = x, Y = y] = \sum_{\forall y} \frac{1}{21}(x + y) = \frac{1}{21}[(x + 1) + (x + 2)] = \frac{2x + 3}{21}$$

$$P_j = P[Y=y] = \sum_{\forall x} P[X = x, Y = y] = \sum_{\forall x} \frac{1}{21}(x + y) = \frac{1}{21}[(1 + y) + (2 + y) + (3 + y)] = \frac{3y + 6}{21}$$

c) Las funciones de probabilidad condicionadas:

$$P[X=x/Y=y] = \frac{P[X = x, Y = y]}{P[Y = y]} = \frac{\frac{1}{21}(x + y)}{\frac{1}{21}(3y + 6)} = \frac{x + y}{3y + 6} \rightarrow$$

X	P[X=x/Y=1]	P[X=x/Y=2]
1	$\frac{2}{9}$	$\frac{3}{12}$
2	$\frac{3}{9}$	$\frac{4}{12}$
3	$\frac{4}{9}$	$\frac{5}{12}$

$$P[Y=y/X=x] = \frac{P[X=x, Y=y]}{P[X=x]} = \frac{\frac{1}{21}(x+y)}{\frac{1}{21}(2x+3)} = \frac{x+y}{2x+3}$$

Y	1	2
P[X=1/Y=y]	$\frac{2}{5}$	$\frac{3}{5}$
P[X=2/Y=y]	$\frac{3}{7}$	$\frac{4}{7}$
P[X=x/Y=y]	$\frac{4}{9}$	$\frac{5}{9}$

2.- Sea una variable aleatoria que se distribuye según $N(\mu, 3)$, donde la media es desconocida, y se quiere decidir entre la hipótesis $H_0: \mu = 110$ frente a la alternativa $H_1: \mu = 130$. Para ello se toma una muestra aleatoria simple de 81 observaciones y se establece la siguiente regla de decisión: “si $\bar{X} \geq 114$ rechazaremos H_0 ”. Se pide:

- Obtener los errores de tipo I y II.
- Obtener la potencia del contraste.

Solución.-

a) La probabilidad α de error de tipo I es la de rechazar H_0 , siendo cierta, en cuyo caso \bar{X} se distribuye normal $N\left(110, \frac{3}{\sqrt{81}}\right)$. Tendremos pues:

$$\alpha = P[\bar{X} \geq 114 / \mu = 110] = (\text{tipificando}) = P[Z \geq 12] = 0$$

La probabilidad β de error de tipo II es la de aceptar H_0 siendo falsa, es decir siendo cierta H_1 , en cuyo caso \bar{X} se distribuye normal $N\left(130, \frac{3}{\sqrt{81}}\right)$. Tendremos pues:

$$\beta = P[\bar{X} < 114 / \mu = 130] = (\text{tipificando}) = P[Z < -48] = 0$$

b) La potencia del contraste $P_C(\mu)$ es la probabilidad de rechazar H_0 , según los distintos valores de μ . Se tendrá:

$$P_C(\mu) = \begin{cases} P[\bar{X} \geq 114] = \alpha = 0, & \text{si } \mu = 110 \\ P[\bar{X} \geq 114] = 1 - P[\bar{X} < 114] = 1 - \beta = 1, & \text{si } \mu = 130 \end{cases}$$