Probabilidad y Estadística
Segundo del grado en Ingeniería de Telecomunicaciones, UAM, 2018-2019

Examen de la convocatoria extraordinaria, 17-6-2019

1. Una prueba diagnóstica para un cierto tipo de cáncer tiene probabilidad del 98 % de resultar positiva si el paciente tiene cáncer. El 97 % de los individuos sin cáncer dan prueba negativa. Se elige un individuo al azar en una población de personas, de las cuales el 1 % tienen dicho tipo de cáncer. Calcula:

(a) La probabilidad de que ese individuo dé positivo en la prueba y tenga cáncer.

(b) Si sabemos que el individuo ha dado resultado positivo, ¿cuál es la probabilidad de que tenga realmente cáncer?

2. a) Se dispone de una serie de datos $(x_1, \dots, x_{100})$, cuya media muestral es $\overline{x}$. Se va a añadir un dato más a la muestra. ¿Qué condición habrá de cumplir ese nuevo dato para que la nueva media muestral (la de la muestra con 101 datos) sea, al menos, una unidad mayor que la original?

b) Disponemos de una serie (doble) de datos $((x_1, y_1), \dots, (x_{50}, y_{50}))$. La ecuación de la recta de regresión de esta serie es $y = x + 1$. A la serie se le añade el dato (3,4). ¿Cuál será la ecuación de la recta de regresión de la serie con los 51 datos?

3. La variable X sigue una distribución $\mathcal{N}(\mu, \sigma)$. Se sabe que $\mathbf{P}(X > 1) = 50\,\%$, y que $\mathbf{E}(X^2) = 5$. ¿Cuánto vale $\mathbf{P}(X > 2)$?

4. Se dispone de un dado piramidal, cuyas cuatro caras van rotuladas con los números del 1 al 4. Se sabe que la cara con el 1 sale con probabilidad 10 %, la del 2 con probabilidad 30 % y la del 3 con probabilidad 20 %.

a) Se lanza 10 veces seguidas el dado. En media, ¿qué puntuación total se obtendrá?

b) Se lanza 200 veces seguidas el dado. Calcula la probabilidad de que la puntuación total que se obtenga sea ≥ 600.

5. La variable X sigue una geométrica de parámetro $p = 1/4$. La variable Y sigue una binomial $\text{BIN}(10, 1/3)$. Las variables X e Y son independientes.

a) Considera la variable aleatoria $Z = X - 3Y$. Calcula $\mathbf{V}(Z)$.

b) Calcula $\mathbf{P}(X + Y = 3)$.

6. a) Un experimento aleatorio consiste en sortear un punto dentro del círculo de radio 1 (es decir, los puntos (x, y) del plano tales que $x^2 + y^2 \leq 1$) con probabilidad *uniforme*. Sea (X, Y) el par de variables aleatorias que representa el resultado del experimento. Escribe la fórmula de la función de densidad conjunta $f_{(X,Y)}(x, y)$.

b) El par de variables aleatorias (X, Y) tiene función de densidad conjunta

$$f_{(X,Y)}(x,y) = \begin{cases} 15xy^2 & \text{si } 0 \leq y \leq x \leq 1, \\ 0 & \text{en cualquier otro caso.} \end{cases}$$

Calcula la densidad marginal $f_Y(y)$ de la variable Y.

7. La variable X sigue una distribución geométrica de parámetro $p \in (0,1)$.
Se dispone de una muestra aleatoria de X de tamaño 40, cuya media muestral es 4 y cuya cuasidesviación típica muestral es 3.25.

Halla la estimación de p por máxima verosimilitud a que da lugar esa muestra.

8. Para un grupo de 41 alumnos de un instituto, se observaron las calificaciones obtenidas en las EvAU, y se obtuvo una media muestral de 7.9 y una cuasidesviación típica muestral de 1.7. Para otro grupo de 25 estudiantes procedentes de otro instituto, se constató que la media muestral era de 7.1 y la cuasidesviación típica muestral era 1.3.
Contesta a las siguientes preguntas, detallando las hipótesis utilizadas.
a) Halla el intervalo de confianza para el cociente de las varianzas al nivel de confianza del 90 %.

b) ¿Aporta la muestra evidencia estadística suficiente como para concluir, con nivel de significación del 5 %, que la nota media en el primer instituto es mayor que en el segundo?

9. Para estimar la proporción de personas que opinan que la temporada final de Juego de Tronos es una castaña, se va a realizar ua encuesta.

a) Se desea estimar el valor de esa proporción con un error máximo del 0.1 %, con confianza del 95 %. ¿Cuál será el número mínimo de personas a las que habrá que encuestar?

b) Se ha decidido, finalmente, preguntar 1450 personas, de las que un 54 % opinaban que, efectivamente, la temporada final era mala. A la vista de estos datos, ¿hay evidencia estadística suficiente como para concluir que más de la mitad de la población considera mala esa última temporada? (argumenta con p-valores).

10. En una serie de 100 observaciones de una cierta magnitud, se han obtenido los siguientes resultados:

rangos	entre 0 y 1	entre 1 y 2	entre 2 y 3	entre 3 y 4
Número de resultados	15	18	22	45

Se pide contratar la hipótesis de que esos datos son muestras aleatorias de una variable X con función de densidad $f(x) = x/8$ si $0 \leq x \leq 4$, y $f(x) = 0$ en otro caso.
(Argumenta con p-valores).

- **Algunos valores de percentiles de la normal estándar:**

α	5 %	4.5 %	4.0 %	3.5 %	3.0 %	2.5 %	2.0 %	1.5 %	1.0 %	0.5 %
z_α	1.645	1.695	1.751	1.812	1.881	1.960	2.054	2.170	2.326	2.576

• Algunos valores de la función Φ de distribución de la normal estándar:

x	-3	-2	-1	0	1	2	3
$\Phi(x)$	0.135 %	2.275 %	15.866 %	50.000 %	84.134 %	97.725 %	99.865 %

• Algunos valores de percentiles de la χ^2:

α	5 %	4.5 %	4.0 %	3.5 %	3.0 %	2.5 %	2.0 %	1.5 %	1.0 %	0.5 %
$\chi^2_{\{2;\alpha\}}$	5.991	6.202	6.438	6.705	7.013	7.378	7.824	8.399	9.210	10.597
$\chi^2_{\{3;\alpha\}}$	7.815	8.049	8.311	8.607	8.947	9.348	9.837	10.465	11.345	12.838
$\chi^2_{\{4;\alpha\}}$	9.488	9.742	10.026	10.345	10.712	11.143	11.668	12.339	13.277	14.860

• Algunos percentiles de la t de Student con n grados de libertad:

n	58	59	60	61	62	63	64	65
$t_{\{n;5\,\%\}}$	1.6716	1.6711	1.6706	1.6702	1.6698	1.6694	1.6690	1.6686

• Algunos percentiles de la F de Fisher (n_1 y n_2 grados de libertad):

$$F_{\{24,40;5\,\%\}} = 1.793, \quad F_{\{24,40;95\,\%\}} = 0.529, \quad F_{\{40,24;5\,\%\}} = 1.892, \quad F_{\{40,24;95\,\%\}} = 0.558.$$

• Una variable X sigue una **geométrica** de parámetro p si toma los valores $\{1, 2, 3, \dots\}$ con probabilidades

$$\mathbf{P}(X = j) = p(1-p)^{j-1} \qquad \text{para } j = 1, 2, 3, \dots$$

Se tiene que $\mathbf{E}(X) = 1/p$ y $\mathbf{V}(X) = (1-p)/p^2$.

• Una variable X sigue una **binomial** de parámetros n y p si toma los valores $\{0, 1, \dots, n\}$ con probabilidades

$$\mathbf{P}(X = j) = \binom{n}{j} p^j \, (1-p)^{n-j} \qquad \text{para } j = 0, 1, \dots, n.$$

Se tiene que $\mathbf{E}(X) = np$ y $\mathbf{V}(X) = np(1-p)$.