



**-ESTADÍSTICA DESCRIPTIVA E INTRODUCCIÓN A LA PROBABILIDAD-
Prueba Temas 1-2**

*Doble Grado en Ingeniería Informática y Matemáticas
22 de abril de 2021*

1. [3 puntos] Indicar si son verdaderas o falsas las siguientes afirmaciones, justificando la respuesta:

- 1.1. De la distribución de la variable X ='calificación en un examen' observada sobre 50 alumnos se sabe lo siguiente: sus observaciones se agruparon en los siguientes intervalos de clase: $[0,5]$, $(5,7]$, $(7,9]$ y $(9,10]$; y los percentiles 30, 70 y 90 fueron 5, 7 y 9 puntos, respectivamente.
- a) El número de alumnos observados cuyas calificaciones oscilan entre $(5,7]$ duplica al de alumnos observados cuyas calificaciones oscilan entre $(7,9]$.
- b) La desviación típica de X vale 5.7 puntos.
- 1.2. Si dos variables estadísticas X e Y son independientes, $m_{02}=5$ y $m_{30}=4$, entonces $m_{32}=18$.
- 1.3. Si a una nube de puntos se le ajusta por mínimos cuadrados una recta de Y/X que pase por el origen, entonces la pendiente de esa recta vale m_{11}/m_{20} .
- 1.4. El coeficiente de correlación lineal entre dos variables X e $Y'=-3Y+2$ es el mismo que entre X e Y .
- 1.5. En una distribución simétrica, siempre coinciden la media aritmética, la mediana y la moda.

2. [2 puntos] Sea (X,Y) una variable estadística bidimensional con distribución de frecuencias dada por $\{(x_i, y_j); n_{ij}\}_{i=1, \dots, k; j=1, \dots, p}$. Obtener la media y varianza de la distribución marginal de X en función de las medias y varianzas de las distribuciones condicionadas de X a cada valor de Y , justificando todos los pasos de la demostración.

3. [5 puntos] Se realiza un estudio para observar el tiempo que tardan en resolver un problema unos escolares que han seguido un curso de formación por módulos. Se observa el número de módulos que han superado (X) junto con el tiempo en minutos que tardan en resolver el problema (Y).

| $X \setminus Y$ | [1-9] | (9-21] | (21-39] |
|-----------------|-------|--------|---------|
| 2 | 0 | 1 | 5 |
| 4 | 0 | 5 | 5 |
| 5 | 5 | 3 | 0 |
| 8 | 15 | 1 | 0 |

- a) Sabiendo que $\sigma_Y^2 = 104.6875$, ¿qué valor medio es más representativo, el de X o el de Y ?
- b) Para los estudiantes que superan menos de 6 módulos, ¿qué porcentaje tarda menos de 18 minutos en resolver el problema? ¿Cuál es el tiempo de respuesta más frecuente?
- c) Calcular el número mínimo de módulos del 40% de los estudiantes que más módulos superan.
- d) Sabiendo que $m_{11}=57.5$, estimar el valor de Y cuando $X=4$ mediante una recta de regresión mínimo cuadrática y dar una medida de la bondad de la predicción.
- e) Ajustar a los datos un modelo de regresión hiperbólico para predecir el tiempo de respuesta conociendo el número de módulos superados y predecir el tiempo de respuesta de un estudiante que superó 4 módulos.
- f) Para predecir el tiempo de respuesta, ¿qué modelo de regresión es más adecuado, el lineal o el hiperbólico?