

1. Dadas dos variables aleatorias discretas,  $X$  e  $Y$ , y dada su distribución de probabilidad conjunta que aparece en la tabla, se pide:
  - a. ¿Cumple la distribución conjunta las propiedades de una distribución de probabilidades?
  - b. ¿Cuál es la probabilidad de  $P(X = x_1)$ ?
  - c. ¿Cuáles son las distribuciones marginales de cada  $P(X = x)$  y  $P(Y = y)$ ?
    - i. Distribución marginal: distribución de probabilidad sobre un subconjunto de las variables aleatorias del espacio probabilístico
  - d. ¿Verifican las distribuciones marginales las propiedades de una distribución de probabilidades?

	$x_1$	$x_2$	$x_3$	$x_4$	$P(Y)$
$y_1$	2/16	1/16	1/16	1/16	
$y_2$	1/16	2/16	2/16	1/16	
$y_3$	1/16	1/16	1/16	0	
$y_4$	0	2/16	0	0	
$P(X)$					

2. Utilizando el conjunto de datos `weather.nominal.practica` que se proporciona, determinar la clasificación Naive Bayes de las siguientes instancias, utilizando la estimación de máxima verosimilitud (frecuencial) y sin utilizar ninguna herramienta de minería de datos:
  - a.  $X_1 < \text{sunny, cool, normal, false} >$
  - b.  $X_2 < \text{overcast, hot, high, true} >$
3. Utilizando Weka y el clasificador Naive Bayes determinar la clasificación de los ejemplos anteriores:
  - a. ¿Coincide con la clasificación calculada en el ejercicio anterior?
4. Entrenar, con Weka, un clasificador Naive Bayes para el conjunto de datos `weather.nominal.practica`
  - a. Estimar la tasa de error cometida por el clasificador utilizando validación cruzada de 10 particiones.
  - b. Examinar la salida proporcionada por el *Explorer* y determinar cómo está estimando esta implementación de Naive Bayes los parámetros del clasificador.
5. El conjunto de datos `weather.nominal.X6` se ha generado repitiendo cada instancia del conjunto `weather.nominal.practica` seis veces. Entrenar con Weka un clasificador Naive Bayes para este conjunto de datos:
  - a. Estimar la tasa de error cometida por el clasificador utilizando validación cruzada de 10 particiones.

- b. Compare esta tasa de error con la estimada en el ejercicio anterior y discuta los resultados.