

## Revisión Teórica

### Seleccionar la/s respuesta/s correcta/s en cada caso

(a) Una hipótesis estadística es:

- ☐ Cualquier afirmación acerca de la distribución de una variable en una población.
- ☐ Una región del espacio muestral que lleva asociada una alternativa.
- ☐ Una regla de decisión que nos indica cuando debe aceptarse o rechazarse el valor de un parámetro.
- ☐ Un valor que se asigna provisionalmente a un estadístico en tanto no se demuestre cuál es su verdadero valor.
- ☐ ninguna de las anteriores es correcta.

(b) Un test de hipótesis estadístico es:

- ☐ Una región del espacio paramétrico.
- ☐ Una afirmación acerca de la población.
- ☐ Una regla de decisión.
- ☐ Un estadístico, función de la muestra.
- ☐ ninguna de las anteriores es correcta.

(c) La región crítica de un test es:

- ☐ Un subconjunto del espacio muestral tal que si la muestra pertenece a él se rechaza la hipótesis nula.
- ☐ Un subconjunto del espacio muestral tal que si la muestra pertenece a él se rechaza la hipótesis falsa.
- ☐ Una región del espacio paramétrico tal que si el parámetro pertenece a ella se rechaza la hipótesis nula.
- ☐ Una región del espacio paramétrico tal que si el parámetro pertenece a ella se rechaza la hipótesis falsa.
- ☐ ninguna de las anteriores es correcta.

(d) La hipótesis nula del Análisis de la Varianza afirma:

- ☐ Existen diferencias entre las medias muestrales.
- ☐ No existen diferencias entre las medias de las subpoblaciones.

- ☐ Existen diferencias entre las medias de las subpoblaciones.
- ☐ No existen diferencias entre las medias muestrales
- ☐ Ninguna de las anteriores es correcta.

(e) La hipótesis nula del test de Levene establece:

- ☐ Lo mismo que la hipótesis nula del test de Bartlett.
- ☐ Que las medias poblacionales de los grupos son iguales.
- ☐ Que las varianzas poblacionales de los grupos son iguales.
- ☐ Son correctas la primera y la tercer.
- ☐ Son correctas la primera y la segunda.

(f) El rechazo de la hipótesis nula del test de Kruskal Wallis indica:

- ☐ Que las medias de los grupos son distintas si las dispersiones son similares.
- ☐ Que las medianas de los grupos son distintas si las distribuciones son similares.
- ☐ Que las dispersiones son distintas aunque las medias pueden ser iguales.
- ☐ Que las distribuciones son distintas si las medianas son iguales.
- ☐ Ninguna de las anteriores es correcta.

(g) El análisis de clusters tiene por objetivo seleccionar grupos que:

- ☐ Maximicen la varianza dentro del grupo y minimicen la varianza entre grupos.
- ☐ Minimicen la varianza dentro y entre los grupos.
- ☐ Maximicen la varianza dentro y entre los grupos.
- ☐ Minimicen la varianza dentro del grupo y maximicen la varianza entre grupos.
- ☐ Ninguna de las anteriores es correcta.

(h) Los conglomerados elegidos por el método dependen de:

- ☐ Las variables elegidas para el análisis.
- ☐ La distancia establecida entre las observaciones.
- ☐ La metodología elegida de medir distancia entre clusters.
- ☐ Todas las anteriores son correctas.
- ☐ Ninguna de las anteriores es correcta.

(i) Para medir o cuantificar la capacidad discriminante de un método es conveniente:

- ☐ Construir una tabla de confusión con las observaciones utilizadas para entrenar el método.
  - ☐ Construir una tabla de confusión con observaciones no utilizadas para entrenar el método.
  - ☐ Extraer las observaciones de una y validar el método construido sin ella para clasificarla a esa.
  - ☐ Se puede elegir la segunda o la tercera según el tamaño de la base disponible.
  - ☐ Todas las anteriores son correctas.
- (j) Los supuestos del análisis discriminante cuadrático de Fisher son:
- ☐ Normalidad de la distribución de las observaciones.
  - ☐ Homocedasticidad de la distribución de las observaciones.
  - ☐ Homocedasticidad de la distribución e independencia de las observaciones.
  - ☐ Ninguna de las anteriores es correcta.
  - ☐ La primera y la segunda son correctas.

### Indicar el valor de verdad de las siguientes premisas

- (a) ☐ V ☐ F El análisis de cluster se puede aplicar a variables y también a observaciones.
- (b) ☐ V ☐ F La hipótesis nula del test M de Box establece que las matrices de covarianza de los grupos son iguales.
- (c) ☐ V ☐ F El test de análisis de la varianza se basa en el estadístico F que es el cociente entre dos variables Chi cuadrado.
- (d) ☐ V ☐ F Se denomina centroide de un grupo al valor de la función discriminante evaluado el vector de medias del grupo.
- (e) ☐ V ☐ F El análisis discriminante es adecuado cuando no se ha rechazado la hipótesis de igualdad de vectores medios de las poblaciones en estudio.
- (f) ☐ V ☐ F Una función de similaridad satisface que es positiva, simétrica y cumple la propiedad triangular.
- (g) ☐ V ☐ F La distribución T de Student es la versión univariada de la distribución de Wishart.
- (h) ☐ V ☐ F Si el coeficiente de una variable es alto en la función discriminante implica que no es una variable que permita discriminar entre los grupos.
- (i) ☐ V ☐ F Si un vector  $X=(x_1, x_2, \dots, x_q, x_{q+1}, \dots, x_p)$  tiene distribución normal multivariada, cualquier combinación lineal de sus componentes tiene distribución normal.

- (j) ☐ V ☐ F El coeficiente de correlación cofenética establece el nivel de asociación entre las distancias originales y las representadas por el árbol binario.

## Responder brevemente a las siguientes preguntas

- (a) ¿Cuándo es adecuado utilizar el análisis de perfiles?
- (b) ¿Cómo se clasifica una nueva observación cuando en el análisis discriminante hay más de dos grupos?
- (c) ¿Qué ventaja tiene el método de “probabilidades a posteriori” para el análisis discriminante?
- (d) Establezca al menos dos diferencias importantes entre el análisis discriminante y el análisis de cluster.
- (e) Explique la diferencia entre segmentación jerárquica y no jerárquica. Señale las características de los métodos divisivos y aglomerativos.