

Revisión Teórica

Indicar el valor de verdad de las siguientes premisas

- (a) Los outliers sólo pueden ser detectados a través de métodos univariados o bivariados.
- (b) En el ACP siempre la suma de los autovalores de la matriz de varianzas y covarianzas es igual a la suma de las varianzas
- (c) La prueba de independencia testea la existencia de asociación entre variables categóricas. Rechaza la hipótesis de independencia cuando el estadístico toma un valor “grande” (p-valor menor que 0.05)
- (d) El ACP se basa en el análisis de la matriz de varianzas y covarianzas.
- (e) Uno de los criterios para determinar con cuántas componentes quedarse en ACP consiste en tomar un número proporcional a la cantidad de variables en el problema
- (f) Hacer ACP con la matriz de correlaciones es lo mismo que hacerlo con la matriz de covarianzas utilizando variables estandarizadas
- (g) El vector de medias de una matriz de datos de $n \times p$ tiene n componentes
- (h) El AFC requiere que las variables sean independientes para que tenga sentido aplicarse
- (i) En AFC, la suma de los autovalores es igual al valor de Chi cuadrado
- (j) El análisis de correspondencias se aplica solamente cuando los perfiles fila y columna son idénticos o paralelos.

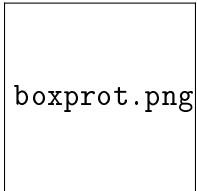
1 Responder las siguientes preguntas

- (a) Cuándo se utiliza la distancia Chi cuadrado? Qué propiedad particular la hace útil en ese contexto?
- (b) Establezca dos similitudes y dos diferencias entre los test de homogeneidad y los de Independencia?
- (c) Con qué objetivo se calculan los residuos de las celdas de un test de independencia?
- (d) Qué cuantifica la inercia en un análisis de correspondencias?

- (e) **Cómo deben tratarse los registros de una base que se consideren outliers? Si se desea detectarlos en forma multivariada cómo conviene hacerlo?**
- (f) **Para que se realizan transformaciones por fila? Y por columna?**
- (g) **Qué características tienen las matrices de covarianzas? Y las de correlación?**
- (h) **Como deben interpretarse los biplot de ACP?**
- (i) **Como deben interpretarse los biplot simétricos de AC?**

Revisión Práctica

1. La matriz de datos contiene el consumo per cápita de diferentes tipos de proteínas(vacuna, cerdo, frutas y vegetales, pescado, huevo, leche, cereal, embutido) en un conjunto de países. Albania, Austria, Bélgica, Bulgaria, Checoslovaquia, Dinamarca, Alemania E, Finlandia, Francia, Grecia, Hungría, Irlanda, Italia, P.Bajos, Noruega, Polonia, Portugal, Rumania, España, Suecia, Suiza, Inglaterra, Rusia, AlemaniaO.



boxprot.png