

PEC 3

UOC

Introducción

La empresa encuestadora aZigna-2 ha entrevistado a una muestra de 129880 pasajeros que prefirieron coger un avión, al nuevo tren de alta velocidad, para llegar a Badajoz.

Los datos de la empresa aZigna-2 proceden del siguiente enlace de [Kaggle](#). El fichero para realizar la PEC 3 es “*data_pac3*” y lo encontraréis en formato csv.

Las variables que se encuentran en el dataset son las siguientes:

- *ID* : Identificador único de pasajeros.
- *Type.of.Travel* : Finalidad del vuelo (negocios/personal).
- *Class* : Tipo de vuelo (Economy Plus, Economy, Business).
- *Flight.Distance* : Distancia de vuelo en millas.

Os puede ser útil consultar el siguiente material:

- Manuales de Intervalo de Confianza y del Teorema del Límite Central.
- Actividades Resueltas del Reto 3 (Intervalo de Confianza y del Teorema del Límite Central).

Hay que entregar la práctica en fichero pdf o html (exportando el resultado final a pdf o html por ejemplo). Se recomienda generar el informe con Rmarkdown que genera automáticamente el pdf/html a entregar.

Pregunta 1 (resolver con R). (3.5 puntos).

Debido a la crisis energética y sanitaria, la aerolínea que estamos analizando decidió renovar los filtros de aire de los aviones, comprando unos nuevos en Amiexpress. Se ha notificado que estos filtros fallan a veces, pero por suerte solo ocurre en el 40% de las ocasiones. Si la aerolínea solo ha renovado 5 filtros:

- a) ¿Qué distribución normal se aproxima a esta distribución binomial? (1.25 puntos).

#Espacio para solución

- b) Calcule $P(X < 1)$ mediante la aproximación normal de la binomial. (1.25 puntos).

#Espacio para solución

- c) En este caso, ¿la aproximación normal es adecuada? Explique las razones. (1 punto).

Espacio para solución

Pregunta 2 (resolver con R). (3.5 puntos).

La empresa aZigna-2 publica todos los años información relevante para mejorar los servicios de las aerolíneas. Este año ha publicado que los pasajeros de aviones de tipología Business recorren una distancia media vuelo, distribuida de manera normal, con una $\mu = 1670$ km y una $\sigma = 150$ km. Nuestra aerolínea de estudio trabaja con este tipo de modelo de aviones y opina que la media es mayor de 1670 km.

- a) Se desea realizar un estudio para estimar (con un nivel de confianza del 90%) la distancia media por vuelo de tipología Business con un margen de error de 10 km. Suponiendo que la desviación estándar de los vuelos de tipología Business es $\sigma = 150$ km, ¿qué tamaño de muestra debería utilizar para este estudio? (1.25 puntos).

#Espacio para solución

- b) Finalmente se llevó a cabo el estudio usando los datos proporcionados en el dataset *data_pac3*, ignorando la recomendación sobre el tamaño de la muestra. Construya un intervalo de confianza del 95% para la distancia media poblacional de los vuelos de tipología Business (**Suponga normalidad y que σ es desconocida**). (1.25 puntos).

Nota: Los km vienen en la variable “Flight.Distance” y la tipología de vuelos Business corresponde con la variable “Class” y el nivel Business.

#Espacio para solución

- c) ¿Se podría rechazar la afirmación de la empresa aZigna-2 de que la distancia media de los pasajeros en vuelos con aviones de tipología Business es de 1670 km basándose en el intervalo de confianza que ha construido anteriormente? (1 punto).

Espacio para solución

Pregunta 3 (resolver con R). (3 puntos).

Se están analizando los vuelos con una distancia recorrida estrictamente mayor de 1000 km (variable “Flight.Distance”).

- a) Encuentre un intervalo de confianza al 99% para la proporción de vuelos con una distancia recorrida estrictamente mayor de 1000 km. Haga este apartado siguiendo las fórmulas de las notas de estudio, y usando R para hacer los diferentes cálculos. (1.5 puntos).

#Espacio para solución

- b) Si queremos calcular el intervalo de confianza del apartado anterior con un nivel de confianza del 99% y con una precisión de 0.001, ¿qué tamaño de muestra necesitamos? (1.5 puntos).

#Espacio para solución