

Inferencia Estadística: Tarea 1

Análisis Descriptivo

Fecha de entrega: 21 de febrero

1. Considera la base de datos *AlumnosInferencia.xlsx* y realiza lo siguiente:
 - (a) Para cada una de las columnas especifica el tipo de variable y la escala de medición.
 - (b) Realiza un análisis descriptivo de la variable *Promedio*. Para ello deberás presentar y analizar las medidas numéricas y gráficas adecuadas.
 - (c) ¿Existe alguna diferencia de promedios entre los dos turnos? Justifica adecuadamente tu respuesta.
 - (d) ¿Existe alguna diferencia de promedios entre géneros? Justifica adecuadamente tu respuesta.
 - (e) ¿Existe alguna diferencia de promedios entre los diferentes tipos de carreras? Justifica adecuadamente tu respuesta.
2. En el archivo *cucos.txt* se presentan 50 mediciones (en mm) de huevos de cucos. El objetivo es realizar un análisis descriptivo de los datos. Para esto se requiere:
 - (a) Obtener la media, la desviación estándar y cualquier otro tipo de medida numérica que consideres oportuna (e.g. curtosis, coeficiente de asimetría, cuantiles, rango, etc.)
 - (b) Obtén el histograma y el diagrama de caja y bigotes.
 - (c) ¿Qué puedes decir de estos datos?
3. Considera la Tabla 1, en donde se presenta la tabla de frecuencias para los pesos registrados de 985 varones de la Ciudad de México.
 - (a) A partir de las frecuencias registradas, dibuja un histograma de los datos.
 - (b) ¿Consideras que los pesos se puedan asumir que vengan de una distribución normal? Justifica tu respuesta.

Tabla 1: Tabla de frecuencias.

Peso	Frecuencia	Peso	Frecuencia
40 -	4	90 -	17
45 -	19	95 -	12
50 -	49	100 -	12
55 -	98	105 -	9
60 -	200	110 -	7
65 -	195	115 -	7
70 -	145	120 -	6
75 -	100	125 -	5
80 -	62	130 -	0
85 -	37	135 - 140	1

4. Los datos en la Tabla 2 corresponden al peso máximo (en toneladas) soportado por un conjunto de cables de acero producidos en una fábrica.
- Represente los datos en un histograma. Utilice intervalos de clase de longitud 1.
 - Obtenga el diagrama de caja y brazos de los datos.
 - ¿Cuál es el peso aproximado que soporta al menos el 75% de los cables?
 - Indique en el diagrama de caja el valor del inciso anterior.
 - Identifique si hay presencia de valores atípicos.

Tabla 2: Pesos máximo soportado por un conjunto de cables de acero.

10.1	12.2	9.3	12.4	13.7	11.1	13.3
10.8	11.6	10.1	11.2	11.4	11.8	7.1
12.2	12.6	9.2	14.2	10.5	12.3	11.2

5. En un ensayo clínico se seleccionaron 1000 voluntarios mexicanos y se registró su grupo sanguíneo. Los datos están contenidos en el archivo *tipoSangre.txt*, donde los valores de la variable *tipo* corresponden a los siguientes grupos sanguíneos: $1 = O^+$, $2 = A^+$, $3 = B^+$, $4 = AB^+$, $5 = O^-$, $6 = A^-$, $7 = B^-$ y $8 = AB^-$. Considerando la variable *tipo* como una variable categórica realice lo siguiente:
- Calcule la tabla de frecuencias absolutas y relativas.
 - Construya la gráfica de barras.

- (c) Obtenga la gráfica de pie/pastel/dona.
 - (d) ¿Qué se puede decir sobre la distribución de los tipos de sangre? ¿Puede dar una aproximación para la proporción de la población mexicana que tiene grupo sanguíneo B^+ ?
6. Para reducir los costos del proceso productivo en una empresa de manufactura de gimnasios se midieron los tiempos de operación de 120 empleados. La información se encuentra en el archivo *tiemposOperacion.txt*. Realiza un análisis descriptivos completo sobre estos datos, considerando medidas numéricas así como gráficas de interés. Haz especial énfasis en la forma de la distribución y escribe tus conclusiones y realiza posibles hipótesis de los datos.