

Estadística descriptiva

18 de junio de 2020

Se compararon tres dietas respecto al control de azúcar en la sangre en pacientes diabéticos. En el archivo dietas.txt se encuentran los valores de glucosa para las tres dietas consideradas (A, B, C), que contienen las lecturas de glucosa en la sangre de los pacientes. Es deseable que el paciente tenga valores entre $80 - 110 \text{ mg/dl}$.

a) Para cada una de las tres dietas calcule medidas de centralidad: la media, la mediana, la media α -podada para $\alpha = 0.1, 0.2$. Para cada dieta compare los valores obtenidos de las cuatro medidas de posición. Si observa alguna diferencia, ¿a qué podría deberse? Mirando las medidas resumen ¿qué dieta no parece ser muy saludable?

b) Calcule medidas de dispersión: el desvío estándar, la distancia intercuartiles y la madecada una de las dietas. Compare los valores de dispersión obtenidos. ¿Cuál de las dietas parece ser la más estable?

c) Obtenga los percentiles 10, 25, 50, 75 y 90. Compare los valores de los percentiles obtenidos entre las distintas dietas. Calcule la distancia intercuartil y el rango de las distancia dietas.

d) Construya diagramas de tallo-hoja e histogramas que permitan visualizar los valores de glucosa para cada dieta. Compare la distribución de glucosa. ¿Alguna de ellas parece bimodal? ¿En alguna de ellas parece haber valores alejados? ¿Las dietas mantienen a los pacientes en los valores deseados? ¿La distribución de glucosa es asimétrica en alguno de los grupos? ¿En algún caso el ajuste normal parece razonable?

e) Grafique los box-plots correspondientes. ¿Cómo se compara la información que dan estos gráficos con la obtenida con los histogramas? En base a los gráficos obtenidos, discuta simetría, presencia de outliers y compare dispersiones nuevamente.

f) Grafique los qqplots correspondientes. ¿En algún caso el ajuste normal parece razonable?

g) En base a todo el análisis anterior, ¿cuál le parece la dieta más aconsejable?