

Question 1

A sample of newborn male infants had their birth weight measured in ounces. The data was then grouped into classes as follows, where x_k represents the class midpoint and N_k represents the frequency of the k -th class

1.

Let the number of classes be:

$$K = 15$$

Then we have the number of individuals:

$$n = \sum_{k=1}^K N_k = 2 + 6 + \dots + 1 = 9\,465$$

Answer

- $n = 9\,465$

2.

Let the individual be:

$$x_i, \quad i \in \{1, 2, \dots, n\}$$

Let the relative frequency be:

$$p_k = \frac{N_k}{n}$$

Then we have the mean:

$$\bar{x} = \sum_{k=1}^K p_k Z_k = \frac{1}{n} \sum_{k=1}^K N_k Z_k = \frac{1\,035\,467}{9\,465} \approx 109.400$$

Then we have the variance:

$$\sigma^2 = \sum_{k=1}^K p_k (Z_k - \bar{x})^2 = \frac{1}{n} \sum_{k=1}^K N_k (Z_k - \bar{x})^2 = \frac{1\,748\,957}{9\,465} \approx 184.782$$

Answer

- $\bar{x} = \frac{1\,035\,467}{9\,465} \approx 109.400$
 - $\sigma^2 = \frac{1\,748\,957}{9\,465} \approx 184.782$
-

3.

The individual becomes:

$$y_i = 28.349 x_i, \quad i \in \{1, 2, \dots, n\}$$

Then the class becomes:

$$Z_{yk} = 28.349 Z_k, \quad k \in \{1, 2, \dots, K\}$$

Then the mean becomes:

$$\begin{aligned} \bar{y} &= \frac{1}{n} \sum_{k=1}^K N_k Z_{yk} = 28.349 \left(\frac{1}{n} \sum_{k=1}^K N_k Z_k \right) \\ &= 28.349 \bar{x} = \frac{29\,354\,453\,983}{9\,465\,000} \approx 3\,101.369 \end{aligned}$$

Then the variance becomes:

$$\begin{aligned} \sigma_y^2 &= \frac{1}{n} \sum_{k=1}^K N_k (Z_{yk} - \bar{y})^2 = 28.349^2 \left(\frac{1}{n} \sum_{k=1}^K N_k (Z_k - \bar{x})^2 \right) \\ &= 28.349^2 \sigma^2 = \frac{140\,557\,692\,831}{946\,500} \approx 148\,502.581 \end{aligned}$$

Answer

- $\bar{y} = 28.349 \bar{x} = \frac{29\,354\,453\,983}{9\,465\,000} \approx 3\,101.369$
 - $\sigma_y^2 = 28.349^2 \sigma^2 = \frac{140\,557\,692\,831}{946\,500} \approx 148\,502.581$
-

4.

The index of the median:

$$i_m = \frac{n+1}{2} = 4\,733$$

Then we know the classes surrounding the median:

$$\sum_{k=1}^6 N_k = 3\,049 < i_m < 5\,289 = \sum_{k=1}^7 N_k$$

Then we have the median:

$$m = Z_6 + (Z_7 - Z_6) \left(\frac{i_m - \sum_{k=1}^6 N_k}{N_7} \right) = \frac{7\,351}{70} \approx 105.014$$

As well as the second quartile:

$$Q_2 = m = \frac{7\,351}{70} \approx 105.014$$

The index of the first quartile (Q_1) and the index of the third quartile (Q_3):

$$i_{Q_1} = \frac{n}{4} + \frac{1}{2} = 2\,366.75$$

$$i_{Q_3} = \frac{3}{4}n + \frac{1}{2} = 7\,099.25$$

Then we know the classes surrounding the quartiles:

$$\sum_{k=1}^5 N_k = 1\,320 < i_{Q_1} < 3\,049 = \sum_{k=1}^6 N_k$$

$$\sum_{k=1}^7 N_k = 5\,289 < i_{Q_3} < 7\,296 = \sum_{k=1}^8 N_k$$

Then we have the first quartile and the third quartile:

$$Q_1 = Z_5 + (Z_6 - Z_5) \left(\frac{i_{Q_1} - \sum_{k=1}^5 N_k}{N_6} \right) \approx 96.449$$

$$Q_3 = Z_7 + (Z_8 - Z_7) \left(\frac{i_{Q_3} - \sum_{k=1}^7 N_k}{N_8} \right) \approx 114.216$$

Answer

- $m_p = \frac{235\,232}{2\,240} \approx 105.014$
 - $Q_1 \approx 96.449$
 - $Q_2 = \frac{7\,351}{70} \approx 105.014$
 - $Q_3 \approx 114.216$
-