Question 1

1.

Let the number of classes be:

$$K = 15$$

Then we have the number of individuals:

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

Answer

- $n = 9465$

2.

Let the individual be:

$$x_i, \;\; i \in \{1, 2, ..., n\}$$

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:

$$y_i = 28.349\, x_i, \;\; i \in \{1, 2, ..., n\}$$

Then the class becomes:

$$Z_{yk} = 28.349\, Z_k, \;\; k \in \{1, 2, ..., K\}$$

Then the mean becomes:

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

Then the variance becomes:

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

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 proportions of the distances of the median to the classes surrounding:

$$p_{[6,i_{c_m}]} = \frac{i_m - \sum_{k=1}^{6} N_k}{N_7} = \frac{1\,684}{2\,240} \approx 0.752$$
$$p_{[i_{c_m},7]} = \frac{\sum_{k=1}^{7} N_k - i_m}{N_7} = \frac{556}{2\,240} \approx 0.248$$

Then we have the median:

$$m = p_{[6,i_{c_m}]} Z_7 + p_{[i_{c_m},7]} Z_6 = \frac{235\,232}{2\,240} \approx 105.014$$

As well as the second quartile:

$$Q_2 = m = \frac{235\,232}{2\,240} \approx 105.014$$

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

$$i_{Q_1} = \frac{i_m + 1}{2} = 2\,367$$
$$i_{Q_3} = \frac{i_m + n}{2} = 7\,099$$

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 proportions of the distances of the quartiles to the classes surrounding:

$$p_{[5,i_{c_{Q_1}}]} = \frac{i_{Q_1} - \sum_{k=1}^{5} N_k}{N_6} = \frac{1\,047}{1\,729} \approx 0.606$$
$$p_{[i_{c_{Q_1}},6]} = \frac{\sum_{k=1}^{6} N_k - i_{Q_1}}{N_6} = \frac{682}{1\,729} \approx 0.394$$

$$p_{[7,i_{c_{Q_3}}]} = \frac{i_{Q_3} - \sum_{k=1}^{7} N_k}{N_8} = \frac{1\,810}{2\,007} \approx 0.902$$
$$p_{[i_{c_{Q_3}},8]} = \frac{\sum_{k=1}^{8} N_k - i_{Q_3}}{N_8} = \frac{197}{2\,007} \approx 0.098$$

Then we have the first quartile and the third quartile:

$$Q_1 = p_{[5,i_{c_{Q_1}}]} Z_6 + p_{[i_{c_{Q_1}},6]} Z_5 = \frac{165\,715}{1\,729} \approx 95.844$$

$$Q_3 = p_{[7,i_{c_{Q_3}}]} Z_8 + p_{[i_{c_{Q_3}},8]} Z_7 = \frac{229\,229}{2\,007} \approx 114.215$$

Answer

- $m_p = \frac{235\,232}{2\,240} \approx 105.014$
- $Q_1 = \frac{165\,715}{1\,729} \approx 95.844$
- $Q_2 = \frac{235\,232}{2\,240} \approx 105.014$
- $Q_3 = \frac{229\,229}{2\,007} \approx 114.215$