

Question 1: Statistical Analysis of a Sample Population

Problem Statement

A sample consists of 120 males and 80 females. The table below shows the distribution of individuals by gender and age.

Part 1: Total Number of Individuals Younger Than 20

Let the number of males (n_{male}) and females (n_{female}) in the sample be:

$$\begin{aligned}n_{\text{male}} &= 120, \\n_{\text{female}} &= 80.\end{aligned}$$

Thus, the total sample size is:

$$n = n_{\text{male}} + n_{\text{female}} = 200.$$

Let the number of age groups be:

$$K = 4.$$

Define the relative frequencies of males (p_{male_k}) and females (p_{female_k}) in each group:

$$\begin{aligned}p_{\text{male}_1} &= 10\%, & p_{\text{male}_2} &= 10\%, & \dots, & p_{\text{male}_K} &= 50\%, \\p_{\text{female}_1} &= 20\%, & p_{\text{female}_2} &= 20\%, & \dots, & p_{\text{female}_K} &= 30\%.\end{aligned}$$

The absolute frequency of males and females in each group is given by:

$$\begin{aligned}N_{\text{male}_k} &= n_{\text{male}}p_{\text{male}_k}, & k &\in \{1, 2, \dots, K\}, \\N_{\text{female}_k} &= n_{\text{female}}p_{\text{female}_k}, & k &\in \{1, 2, \dots, K\}.\end{aligned}$$

The total absolute frequency is:

$$N_k = N_{\text{male}_k} + N_{\text{female}_k}, \quad k \in \{1, 2, \dots, K\}.$$

The number of people younger than 20 is:

$$N_{\#\{y/o \in \{0,1,\dots,19\}\}} = N_1 = 28.$$

Answer for Part 1:

$$\boxed{N_{\#\{y/o \in \{0,1,\dots,19\}\}} = 28}.$$

Part 2: Percentage of Individuals Aged 50 or Older

The relative frequency for the total population is:

$$p_k = \frac{N_k}{n}, \quad k \in \{1, 2, \dots, K\}.$$

The percentage of individuals aged 50 or older is:

$$p_{\frac{\#\{y/o \in \{50,51,\dots,89\}\}}{\#\{y/o \in \{0,1,\dots,89\}\}}} = p_4 = 42\%.$$

Answer for Part 2:

$$\boxed{p_{\frac{\#\{y/o \in \{50,51,\dots,89\}\}}{\#\{y/o \in \{0,1,\dots,89\}\}}} = 42\%}.$$

Part 3: Number of Males Aged 30 or Older

The number of males aged 30 years or older is:

$$N_{\text{male}_{\#\{y/o \in \{30,31,\dots,89\}\}}} = \sum_{k=3}^K N_{\text{male}_k} = 96.$$

Answer for Part 3:

$$\boxed{N_{\text{male}_{\#\{y/o \in \{30,31,\dots,89\}\}}} = 96}.$$

Part 4: Median Age Calculation

Define the age class intervals:

$$\begin{aligned} z_1 &= y/o \in \{0, 1, \dots, 19\}, \\ z_2 &= y/o \in \{20, 21, \dots, 29\}, \\ &\dots \\ z_K &= y/o \in \{50, 51, \dots, 89\}, \quad k \in \{1, 2, \dots, K\}. \end{aligned}$$

The median position is at:

$$p_m = 50\%.$$

Since:

$$\sum_{k=1}^2 p_k = 28\% < p_m < 58\% = \sum_{k=1}^3 p_k,$$

the median falls within z_3 . The lower boundary of z_3 is:

$$L_{z_3} = \frac{29 + 30}{2} = 29.5.$$

The class width is:

$$C_{z_3} = \frac{(49 + 50) - (29 + 30)}{2} = 20.$$

Assuming a uniform distribution within z_3 , the median is calculated as:

$$m = L_{z_3} + \left(\frac{p_m - \sum_{k=1}^2 p_k}{p_3} \right) C_{z_3}.$$

Substituting values:

$$m = 29.5 + \left(\frac{50\% - 28\%}{30\%} \right) \times 20 = \frac{265}{6} \approx 44.167.$$

Answer for Part 4:

$$m = \frac{265}{6} \approx 44.167.$$