HOMEWORK 1

Use the weights of freshmen males in September to construct a frequency distribution. Begin with a lower class limit of 50 kg and use a class width of 10 kg.

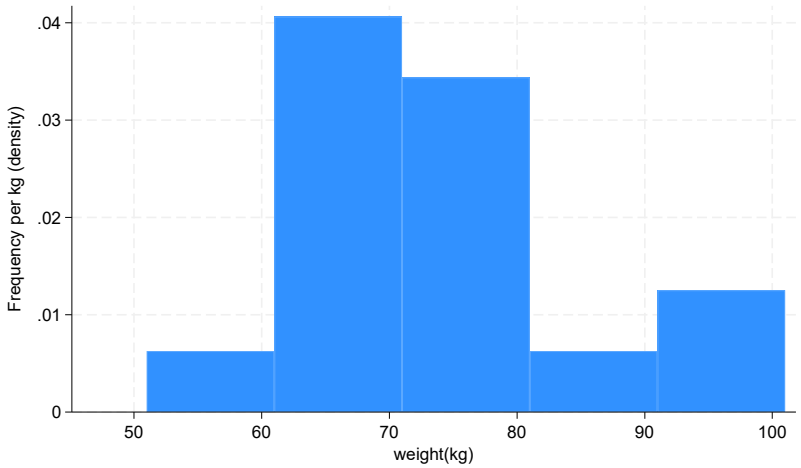
52	69	72	78
56	69	72	78
62	69	73	81
63	69	75	88
64	69	75	92
66	70	75	94
66	70	76	96
67	72	78	97

1. Build the frequency table with relative and cumulative frequencies

Weight class	**Frequency**	**Relative frequency**	**Cumulative frequency**	Density Frequency per kg
50-|60	2	0.063	2	0.063/10=0.0063
60-|70	13	0.406	15	0.0406
70-|80	11	0.344	26	0.0344
80-|90	2	0.063	28	0.0063
90-|100	4	0.125	32	0.0125
Total	32	1	32	

(quiz 1, 2,3)

2. Draw a histogram to show the distribution of weights



quiz 4: 3 histograms are correct, on the Y axis you can have: frequency OR relative frequency OR density (frequency per kg), the last one is better as it is the only one that accounts for intervals with different length!

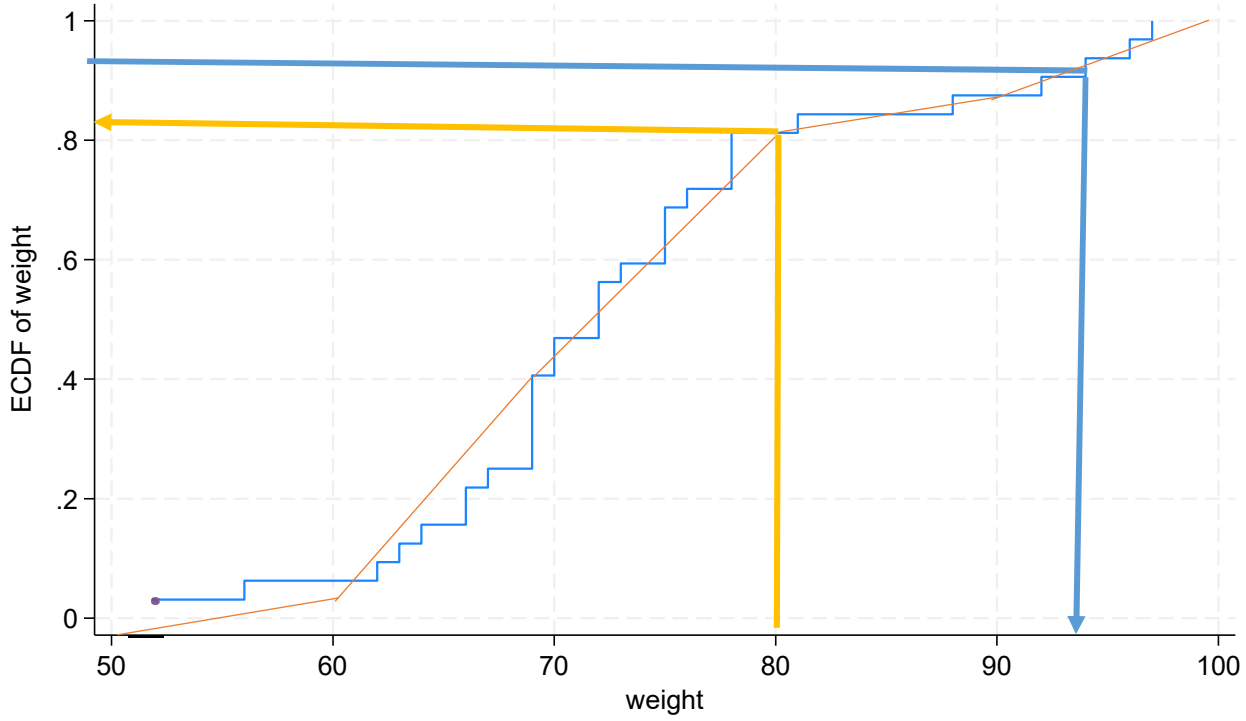
3. Compute cumulative relative frequencies

Weight class	Relative frequency	Relative cumulative frequency
50-	0.063	0.063
60-	0.406	0.469
70-	0.344	0.813
80-	0.063	0.876
90-	0.125	1
Total	1	1

(quiz 5)

Quiz 5: Please note that the relative cumulative frequency at 90 kg is 0.876 (or 87.6%) as this is the frequency cumulated up to the end of the interval (80 to 90 included). This means that 87.6% of the sample weights less than 90 kg. You might also find it by individual data (28 out of 32 freshmen males weights less than 90 kg).

4. Draw the cumulative distribution function



Blu line: cumulative relative frequency using individual data

quiz 6: yellow arrow-> 80 kg corresponds to a cumulative relative frequency of 81.3% (the exact value I got from individual data 100*26/32). This can also be identified using aggregated data.

quiz 7: blu arrow-> the 90th percentile has, by definition, a cumulative relative frequency of 0.9. In the example it corresponds to 92 kg approximately (a visual approx. is ok).

5. Calculate the sample mean

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{52 + 56 + \cdots + 96 + 97}{32} = \frac{2353}{32} = 73.5$$

(quiz 8)

6. Calculate the standard deviation

$n = 32$

$\sum x_i = 2353$

$\sum x_i^2 = 52^2 + 56^2 + \cdots + 96^2 + 97^2 = 2704 + 3136 + \cdots + 9216 + 9409 = 176569$

$$s = \sqrt{\frac{1}{n-1}\sum(x_i - \bar{x})^2} = \sqrt{\frac{1}{n-1}\left(\sum x_i^2 - \frac{(\sum x_i)^2}{n}\right)} = \sqrt{\frac{1}{32-}\left(176569 - \frac{2353^2}{32}\right)} = \sqrt{\frac{3550}{31}} = 10.7$$

(quiz 9 e 10)

7. Calculate the mode

$mode = 69$

(quiz 11)

8. Calculate quartiles and indicate median

$first\ quartile: q1 = 67$

$median: \frac{32}{2} = 16 \rightarrow 16th = 72; \frac{32}{2} + 1 = 17 \rightarrow 17th = 72 \rightarrow q2 = 72$

$third\ quartile: q3 = 78$

(quiz 12,13,14)

9. Identify minimum and maximum value and calculate range

$min = 52$

$max = 97$

$range = 97 - 52 = 45$

(quiz 15)

10. Draw a box and whisker plot (also called box-plot)

5 ingredients: MIN=52, BOX from 67 to 78, MAX= 97 , median=72

Outliers below 51 and above 94 kg :

First quartile=67

Third quartile=78

Interquartile range = 78-67=11

*67-11*1.5=50.5*

*78+11*1.5=94.5*

11. Answer to question in the course web page