Handout for Lecture 2

Empirical Distribution and Measures of Central Tendency

ECON 340: Economic Research Methods | Instructor: Div Bhagia

1. You rolled a six-sided die 100 times and noted down how many times each of the six outcomes was realized. Fill in the rest of the table below:

Outcome	Count (n_k)	Relative frequency (f_k)	Cumulative frequency (F_k)
1	18	0.18	0.18
2	18	0.18	0.36
3	12	0.12	0.48
4	16	0.16	0.64
5	21	0.21	0.85
6	15	0.15	1
Total	**100**	**1**	

Note that

$$f_k = \frac{n_k}{n} = \frac{\text{observations in category } k}{\text{total observations}}$$

(a) How many times did you get a die face with a value of at most 3? 48

(b) Are the proportions close to what you would have predicted?

Yes, we would have predicted each outcome's frequency to be close to 1/6 (or 0.16).

2. Find the mean and median for: $3, 4, 1, 6, 8$

$$mean = \frac{3 + 4 + 1 + 6 + 9}{5} = \frac{22}{5} = 4.4$$

Arrange in ascending order, then the middle number is the median: $1, 3, \mathbf{4}, 6, 8$.

3. Amongst the mean and the median, which one is more affected by outliers? Explain.

   The mean is more sensitive to outliers because all values are used in its calculation, unlike the median, which is the middle value and less affected by extreme values.

4. We asked a sample of 10 individuals whether they like icecream or not. We then created a variable X that takes value 1 if the individual likes icecream, and 0 otherwise. Here is the data we collected:

$$1, 1, 0, 0, 0, 1, 0, 1, 1, 1$$

   (a) How many individuals like icecream in our sample? 6

   (b) What proportion of individuals like icecream in our sample? 6/10 = 0.6

   (c) Use the frequency distribution table and the following formula to calculate the mean of X.

$$\bar{X} = \frac{\sum_{k=1}^{K} n_k X_k}{n} = \sum_{k=1}^{K} f_k X_k = 0.6$$

X_k	n_k	f_k	$X_k f_k$
1	6	0.6	0.6
0	4	0.4	0
			0.6

5. We have the following data on shoe sizes (X_i) for four individuals.

$$X = \{8, 6, 6, 8\}$$

(a) Calculate the mean:

$$\mu = \frac{\sum_{i=1}^{N} X_i}{N} = \frac{8+6+6+8}{4} = \frac{28}{4} = 7$$

(b) Calculate the weighted mean with weights $w = \{1, 1, 1, 1\}$.

$$\mu_{Weighted} = \frac{\sum_{i=1}^{N} w_i X_i}{\sum_{i=1}^{N} w_i} = \frac{1\times 8 + 1\times 6 + 1\times 6 + 1\times 8}{1+1+1+1} = 7$$

(c) Calculate the weighted mean with weights $w = \{1, 2, 2, 1\}$.

$$\mu_{Weighted} = \frac{\sum_{i=1}^{N} w_i X_i}{\sum_{i=1}^{N} w_i} = \frac{1\times 8 + 2\times 6 + 2\times 6 + 1\times 8}{1+2+2+1} = 6.66$$

(d) Calculate the weighted mean with weights $w = \{0.5, 0, 0, 0.5\}$.

$$\mu_{Weighted} = \frac{\sum_{i=1}^{N} w_i X_i}{\sum_{i=1}^{N} w_i} = \frac{0.5\times 8 + 0\times 6 + 0\times 6 + 0.5\times 8}{0.5+0+0+0.5} = 8$$