2017 Statistics Review
John Glenn College of Public Affairs

1.Governments can be divided into three different types – unitary, federal, and confederations – depending upon where the concentration of power is located. This would be considered which level of measurement (nominal, ordinal, interval, or ratio)?

Nominal (or possibly ordinal if the variable of interest related to centralization of power)

2. Which measure of central tendency cuts a distribution in half when the scores area arranged from high to low?

Median

3. A distribution of income is highly skewed. Which measure of central tendency is likely best for purposes of characterizing income?

Median

4. A local city is considering building a new light rail system to alleviate downtown traffic. To find out how many people would make use of the system, a government worker interviewed a random sample of commuters. The following are the number of miles driven each week by 10 people who said they would use the train instead of driving to work:

15 35 30 42 18 75 60 45 20 90

Calculate the a) range b) mean c) median d) variance and e) standard deviation

Range = 90-15 = 75
Mean = (15+35+30+42+18+75+60+45+20+90)/10 = 43
Median: Arrange values sequentially and find midpoint
15 18 20 30 35 42 45 60 75 90
Midpoint is between 35 and 42
Calculate average between (35 + 42)/2 = 38.5
Variance:

$$\frac{\sum(x_i - \bar{x})^2}{(n-1)} = \frac{\sum(x_i - 43)^2}{(10-1)} = \frac{(15-43)^2 + (35-43)^2 + \cdots + (90-43)^2}{(10-1)}$$
$$= 633.11$$

$$\sqrt{s^2} = \sqrt{633.11} = 25.16$$

5. An analyst has collected the following data on the number of complaints about garbage collection made over the past 10 years. Calculate the mean, median, and mode. Which measure of central tendency are you most likely to use, and why?

Year	Number of complaints
2000	34
2001	42
2002	24
2003	43
2004	20
2005	45
2006	28
2007	42
2008	38
2009	40

Mean = 35.6; Median = 39; Mode = 42

6. You wanted to investigate the productivity of two departments: Econ and Public Affairs (the numbers are fake). You recorded the numbers of publications per year for each department, as the following table shows.

Table 1. Number of Publications in Two Departments

	Econ	Public Affairs
Max	20	24
Min	12	10
Mean	16	12
Standard Error	5	2
Variance	20	40
Sample Size (N)	9	14

a. Set up a two-tailed hypothesis test for this analysis.

H_0 (Null hypothesis): the number of publications in the Econ department is the same as the number of publications in the PA department.
H_1 (Alternative hypothesis): the number of publications in the Econ department is different from number of publications in the PA department.

b. Calculate the 95% confidence interval for each department in terms of the average number of publications per year.

Econ:

Confidence Interval= Mean+/- (SE*Z-Score) = [16-5*1.96, 16+5*1.96] = [6.2, 25.8]

Public Affairs:

Confidence Interval= Mean+/- (SE*Z-Score) = [12-2*1.96, 12+2*1.96] = [8.08, 15.92]

c. How will you make the conclusion about the hypothesis testing based on the 95% confidence interval you calculated in part b?

We fail to reject the null hypothesis. The number of publications in the Econ department is not different from the number of publications in the PA department.

d. Based on the 95% confidence interval you calculated in part b, is it possible that p-value is higher than 0.05?

The p-value must be higher than 0.05.