**Excel Project #3: Interval Estimation**

**Due Date:**

|  |  |
| --- | --- |
| **Confidence Intervals**  **with known**  *CONFIDENCE + SIGNIFICANCE = 100%*  Information Needed:  ~ sample mean  ~ number of samples  ~ significance level  Calculate:  ~ Standard Error  ~ Z-score | **Confidence Intervals**  **with unknown**  *CONFIDENCE + SIGNIFICANCE = 100%*  Information Needed:  ~ sample mean  s ~ sample standard deviation  ~ number of samples  ~ significance level  Calculate:  ~ Standard Error  **~ t-score** |

Introduction: In this Excel Project, we will explore the idea of interval estimation. With interval estimation, we can infer the boundaries of an interval likely to contain the value of the true mean of an entire population based on the limited information of a sample of data. The relevant formulae and summaries are provided in the table above.

Instructions: Navigate to *D2L > Course Content > Course Materials > Excel Project > Excel Project #3.* Open the Excel spreadsheet entitled '*EP3\_yourname.xls'.* Save the file to your computer and replace your name in the title where appropriate. In this spreadsheet, you will find two columns of data. The first column represents the age of actresses when they won the Academy Award for Best Actress. The second column represents the age of actors when they won the Academy Aware for Best Actor. You will be estimating the true mean of each dataset, under various assumptions.

When you are finished with the problems, save the spreadsheet and upload it to the *Excel Project #3* drop box on D2L.

**Part 1: Standard Deviation Known**

Instructions: For the first part, assume the standard deviation of each dataset is known prior to the observation being taken. For the actresses, assume and for the actors, assume

1. Using the column that represents the age of actors when they won an Oscar, calculate the following summary statistics and save them in the appropriate cell:

a. The sample mean

b. The sample variance

c. The sample standard deviation

d. The number of observations.

2. Calculate the *margin of error* by following the steps in this problem. Choose a significance of 10% and save it in the appropriate cell.

a. Use this significance to calculate the critical Z-score with the formula,



*Note:* this formula assumes the significance is stored in cell *K7*. Make sure to reference the significance; do not input the significance of 10% directly into the norm.s.inv() function. We will be modifying the significance to see how it affects the confidence interval in the next problem. It will be much easier to see make modifications if we only have to change one cell and then let the cell references take care of everything else!

b. Calculate the standard error with the formula,

c. Use the results of *part a* and *part b* to calculate the *margin of error*

d. Add and subtract the MOE to the sample mean to find the upper and lower bounds of the 90% confidence interval for the true mean. Save the results in separate cells. Label them appropriately.

i. Adjust the significance to 5%. Save the new results in separate cells. Label them appropriately.

ii. Adjust the significance to 1%. Save the new results in separate cells. Label them appropriately.

Iii. Using the results of this problem, how does changing the significance/confidence level affect the width of the confidence interval? Save your answer in a cell and label it.

3. Repeat #1 and #2 for the column representing the age of actresses when they won an Oscar.

**Part 1: Standard Deviation Unknown**

Instructions: For the first part, assume the standard deviation of each dataset is known prior to the observation being taken. For the actresses, assume and for the actors, assume

1. Using the column that represents the age of actors when they won an Oscar, calculate the following summary statistics and save them in the appropriate cell:

a. The sample mean

b. The sample variance

c. The sample standard deviation

d. The number of observations.

2. Calculate the *margin of error* by following the steps in this problem. Choose a significance of 10% and save it in the appropriate cell.

a. Use this significance to calculate the critical t-score with the formula,



*Note:* This formula assumes the significance is stored in cell *G7* and the number of observations in stored in cell *E13.* Recall, the t-distribution requires the additional parameters, *the degrees of freedom,* which for our purposes will always be one less than the number of observations,  *n – 1*.

i. In a cell, explain why we need the *t-distribution* for this calculation instead of the *normal distribution*.

ii. Compare the critical t-score to the critical z-score calculated in the previous part. Explain the difference. Save your answer in a cell.

b. Calculate the standard error with the formula,

c. Use the results of *part a* and *part b* to calculate the *margin of error*

d. Add and subtract the MOE to the sample mean to find the upper and lower bounds of the 90% confidence interval for the true mean. Save the results in separate cells. Label them appropriately.

3. Repeat #1 and #2 for the column representing the age of actresses when they won an Oscar.