

## Stat Support Activity: Randomized Simulation For Tail and P-value Calculations

### Notes on Randomized Simulation (or Randomization):

- **Randomized Simulation (or Randomization):** Having the computer create thousands of simulated random samples based on the null hypothesis of the same size as the original random sample. Calculate the statistic for each of the thousands of randomization samples and put the thousands of statistics on the same randomization distribution.
- **Tail and P-value:** Randomization Distributions can be created for a sample statistic like the sample mean or sample proportion. So, you can judge tail significance and calculate P-values without Test Statistics or Theoretical Distributions.
- **Conditions:** Since a randomized simulation is often not using the theoretical distribution, you can use randomization when theoretical distribution conditions are not met. Like bootstrapping, randomization requires only an original unbiased random sample from a large population.
- **Creating Tail and judging Significance without a Critical Value or Test Statistic:** Once the randomization distribution for a statistic is created, you simply put in the significance level as the proportion in the tail or tails. Instead of seeing if your test statistic falls in the tail, you can see if your original real random sample statistic falls in the tail. For example: Instead of creating a T-test statistic and seeing if the T-test statistic falls in the tail, we can just see if the original sample mean falls in the tail. Instead of creating a Z-test statistic and seeing if the Z-test statistic falls in the tail, we can just see if the original sample proportion falls in the tail. Either way, being in the tail indicates that the original sample statistic significantly disagrees with the null hypothesis.
- **Calculating the P-value without a Test Statistic:** Once the randomization distribution for a statistic is created, you can use it to calculate the P-value. Remember, the P-value is the probability of getting the original sample statistic or more extreme because of sampling distribution if the null hypothesis is true. Notice the definition does not say anything about test statistics. The randomization distribution is a graph of the expected sampling variability if the null hypothesis is true. So, all we have to do is find the tail probability corresponding to the original sample statistic like the original sample mean or the original sample proportion.
- **Sampling Variability:** Like Bootstrapping, there will be slight variations in the randomized simulations. Your P-value, tail, center, and standard error may be slightly different than other people doing the same distribution. Random samples are always different.
- **Advanced Data Notes:** Hypothesis tests with complicated data like ANOVA, Goodness of Fit, or Categorical Association Tests do not really use one sample statistic that represents the sample. So, they calculate the test statistic ( $F$  or  $\chi^2$ ) for the original sample. Then they make a distribution of random samples based on the null hypothesis and calculate thousands of test statistics to create the simulated randomization distribution. They can approximate critical value or values using the tail. They can also approximate the P-value using the original sample test statistic ( $F$  or  $\chi^2$ ).

## Stat Support Activity: Randomized Simulation For Tail and P-value Calculations

### Problems

1.

Go to [www.lock5stat.com](http://www.lock5stat.com) and open StatKey. Under the “Randomization Hypothesis Test” menu on the top left, open “Test for Single Mean”. Click the built-in data set that says “Body Temperature”. You will see the original random sample of 50 adults with a sample mean body temperature of 98.26 degrees Fahrenheit on the top right. We want to use this data to test the claim that mean body temperature for all humans is less than 98.6 degrees Fahrenheit. Notice this is a left tailed test. We will use a 5% significance level. Instead of calculating the test statistic and the student T-distribution, we will use randomized simulation and the actual sample mean 98.26 degrees Fahrenheit to judge significance in the tail and find the P-value.

$H_0: \mu = 98.6$  degrees Fahrenheit

$H_A: \mu < 98.6$  degrees Fahrenheit (Claim)

- Make sure the null hypothesis box says “ $\mu = 98.6$ ” (This is the basis of the randomized simulation.)
- Click on “Generate 1000 Samples” at least three times. This creates the simulated distribution based on the null hypothesis. Each of the 3000+ randomization samples have the same sample size (50) as the original sample of 50 people, but are based on sample means we would expect to get if the null hypothesis was true and  $\mu = 98.6$ .
- Is the mean (center) of the randomization distribution close to 98.6? (It should be.)
- What is the standard error for the distribution? (Answers will vary slightly due to sampling variability.)
- Since this is a left tailed test, click the box that says “Left Tail”. The proportion in the tail corresponds to the significance level 0.05 so put in 0.05 in the top left box of the distribution. What is the temperature in degrees Fahrenheit in the bottom left box of the distribution corresponding to the 5% significance level? This is where the tail begins to determine significance.
- Draw a picture of the randomization distribution labeling the left tail, the significance level and where the tail starts.
- Does the original sample mean of 98.26 degrees Fahrenheit fall in the left tail or not?
- Is the original sample mean of 98.26 degrees Fahrenheit significantly less than the population mean 98.6 degrees Fahrenheit in the null hypothesis? How do you know?  
Now let’s calculate the approximate P-value. The P-value will be the probability in the tail corresponding to the original sample mean 98.26 degrees Fahrenheit. So, type in 98.26 in the lower left box. The upper left box is now the P-value. What is the P-value? (Answers will vary slightly due to sampling variability.)
- Convert your P-value into a percent.
- Is your P-value lower than the 5% significance level?
- Could the original sample mean 98.26 have occurred because of sampling variability or is that very unlikely?
- Should we “Reject  $H_0$ ” or “Fail to reject  $H_0$ ”? (The original random sample data did pass all conditions.)
- Write the formal conclusion statement for the hypothesis test.
- Use the standard error in (d) to calculate the T-test statistic using the formula

$$T = \frac{(\text{Sample Mean} - \text{Population Mean})}{\text{Standard Error}}. \text{ (Answers will vary slightly due to sampling variability.)}$$

## Stat Support Activity: Randomized Simulation For Tail and P-value Calculations

2.

Go to [www.lock5stat.com](http://www.lock5stat.com) and open StatKey. Under the “Randomization Hypothesis Test” menu on the top left, open “Test for Single Proportion”. Click the built-in data set that says “Election Poll Support Obama”. You will see the original random sample of 1057 adults with a sample proportion of 0.552 on the top right. We want to use this data to test the claim that more than 50% of voters will vote for Obama. Notice this is a right tailed test. We will use a 5% significance level. Instead of calculating the test statistic and the standard normal Z distribution, we will use randomized simulation and the actual sample proportion of 0.552 to judge significance in the tail and find the P-value.

$$H_0: p = 0.5$$

$$H_A: p > 0.5 \text{ (Claim)}$$

- Make sure the null hypothesis box says “ $p = 0.5$ ” (This is the basis of the randomized simulation.)
- Click on “Generate 1000 Samples” at least three times. This creates the simulated distribution based on the null hypothesis. Each of the 3000+ randomization samples have the same sample size (1057) as the original sample of 1057 people, but are based on sample proportions we would expect to get if the null hypothesis was true and  $p = 0.5$  (50%).
- Is the mean (center) of the randomization distribution close to 0.5? (It should be.)
- What is the standard error for the distribution? (*Answers will vary slightly due to sampling variability.*)
- Since this is a right tailed test, click the box that says “Right Tail”. The proportion in the tail corresponds to the significance level 0.01 so put in 0.01 in the top right box of the distribution. What is the proportion in the bottom right box of the distribution corresponding to 1% significance level? This is where the tail begins to determine significance.
- Draw a picture of the randomization distribution labeling the right tail, the significance level and the proportion where tail starts.
- Does the original sample proportion 0.552 fall in the right tail or not?
- Is the original sample proportion 0.552 significantly higher than the population proportion 0.5 (50%) in the null hypothesis? How do you know?

Now let’s calculate the approximate P-value. The P-value will be the probability in the tail corresponding to the original sample proportion 0.552. So, type in 0.552 in the lower right box. The upper right box is now the P-value. What is the P-value? (*Answers will vary slightly due to sampling variability.*)

- Convert your P-value into a percent.
- Is your P-value lower than the 1% significance level?
- Could the original sample proportion 0.552 have occurred because of sampling variability or is that very unlikely?
- Should we “Reject  $H_0$ ” or “Fail to reject  $H_0$ ”? (The original random sample data did pass all conditions.)
- Write the formal conclusion statement for the hypothesis test.
- Use the standard error in (d) to calculate the Z-test statistic using the formula

$$Z = \frac{(\text{Sample Proportion} - \text{Population Proportion})}{\text{Standard Error}}. \text{ (*Answers will vary slightly due to sampling variability.*)}$$