Quantitative Methods II
Assignment 4

September 18, 2011

Solutions

Problem: The t-test: Is 30 enough? ⟦10⟧

Many statistics textbooks suggest that when the sample size, n, is greater than 30, one can safely use the z-test. Previously, we showed that this was not the case if we did not know the population variance. Here, we test to determine if it is the case if the measurements are not Normally distributed.

To determine this, we produce histograms of a Monte Carlo experiment with the z-test, where the measurement is Exponentially distributed. If the test is appropriate, then we would expect the histograms to have uniform height. Departures from uniformity indicate inappropriateness for the test.

Sample size = 30. Examining the histogram of the Monte Carlo experiment (Figure 1), we see that the first bar is lower than its expected height. As such, this test will reject less often than it should. Performing a Kolmogorov-Smirnov test supports these findings that the distribution of the p-values is not Uniform ($D = 0.0044, p \ll 0.0001$). Thus, we should not use the z-test under these circumstances.

Sample size = 50. A sample of size 50 gives a the same answer. Examining the histogram of this Monte Carlo experiment (Figure 2), we see that the first bar is lower than its expected height. As such, this test appears to be inappropriate under these circumstances. Performing a Kolmogorov-Smirnov test supports these findings that the distribution of the p-values is not sufficiently Uniform ($D = 0.0031, p \ll 0.0001$). Thus, it would be inacceptable to use the z-test under these circumstances.

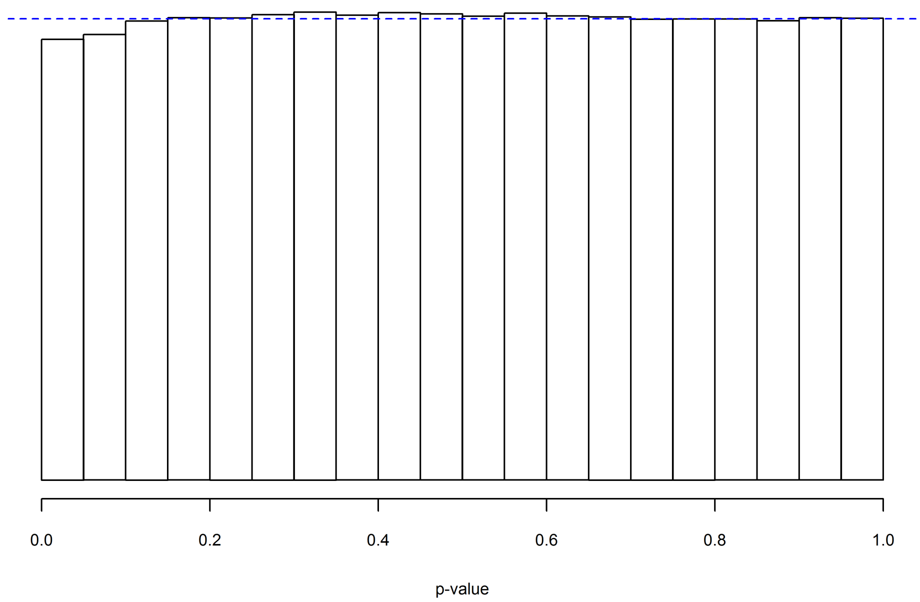

Figure 1. *Results of a Monte Carlo experiment with one million draws from an Exponential distribution (rate parameter 1), with a sample size of 30.*

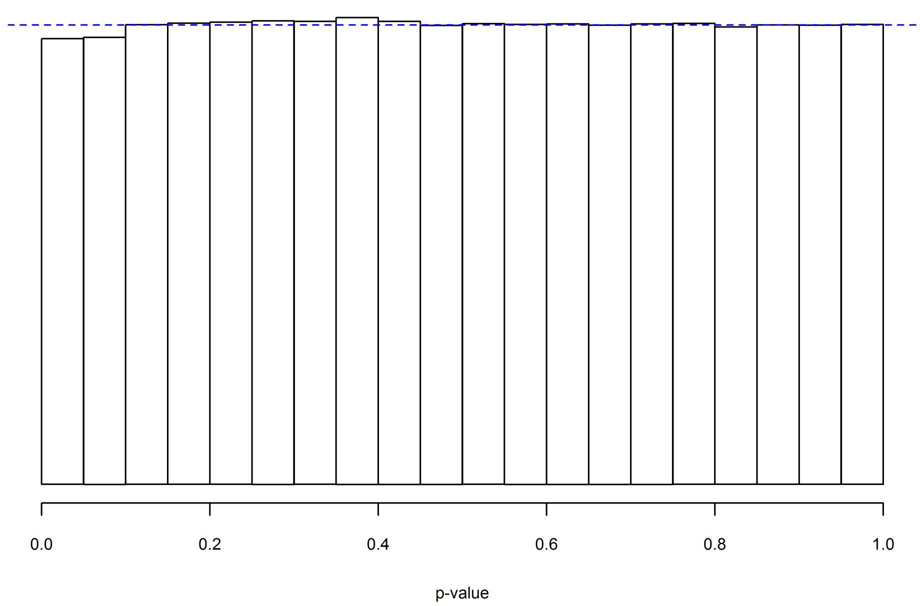

Figure 2. *Results of a Monte Carlo experiment with one million draws from an Exponential distribution (rate parameter 1), with a sample size of 50.*

Note. *The amount of discrepancy between the expected and the experienced heights to that first bar are not substantively significant. Thus, while the test is technically inappropriate, the adjustment you would want to make to the calculated p-value is not large in either case.*

Problem: Firefighting deaths are a roll of the die ⟦15⟧

Station 1 uses professionals to fight fires; Station 2 uses volunteers. In all other aspects, these two stations are equivalent. This offers a perfect time to examine the effect of professionalism on deaths during firefighting.

If we designate X_1 as the number of years Station 1 suffered a loss of life, and X_2 as the number of years Station 2 suffered a loss of life, with the associated probabilities being π_1 and π_2, then our null hypothesis that there is no difference between the two stations is

$$\pi_1 = \pi_2$$

As the number of years with a loss of life is distributed as a Binomial random variable, the number of data points is small ($n = 20$), and the probabilities are not close to 0.500, we cannot use the usual t-test. We could use a non-parametric test, but there is a loss of power associated with such tests. There is also a loss of efficiency, since non-parametric tests will discard information we already know.

To create a good test under these circumstances, we will use Monte Carlo methods (and ten million trials) to determine the distribution of our test statistic, $TS = X_1 - X_2$.

Figure 3 is a histogram of the empirical distribution of the test statistic. Note that it is symmetric about zero. Also note that the usual two-sided 95% confidence interval for this distribution is $(-6, 6)$. Thus, since our observed test statistic ($TS = 8 - 4 = 4$) is within the 95% confidence interval, we fail to reject the null hypothesis and conclude that there is no evidnce of a significant difference between the professional and the volunteer firefighters in terms of deaths.

If we wanted to use a p-value, we could write: The data supports the null hypothesis. Thus, we conclude that there is no significant difference (in terms of deaths) between professional and volunteer fire stations ($TS = 4, p = 0.3874$).

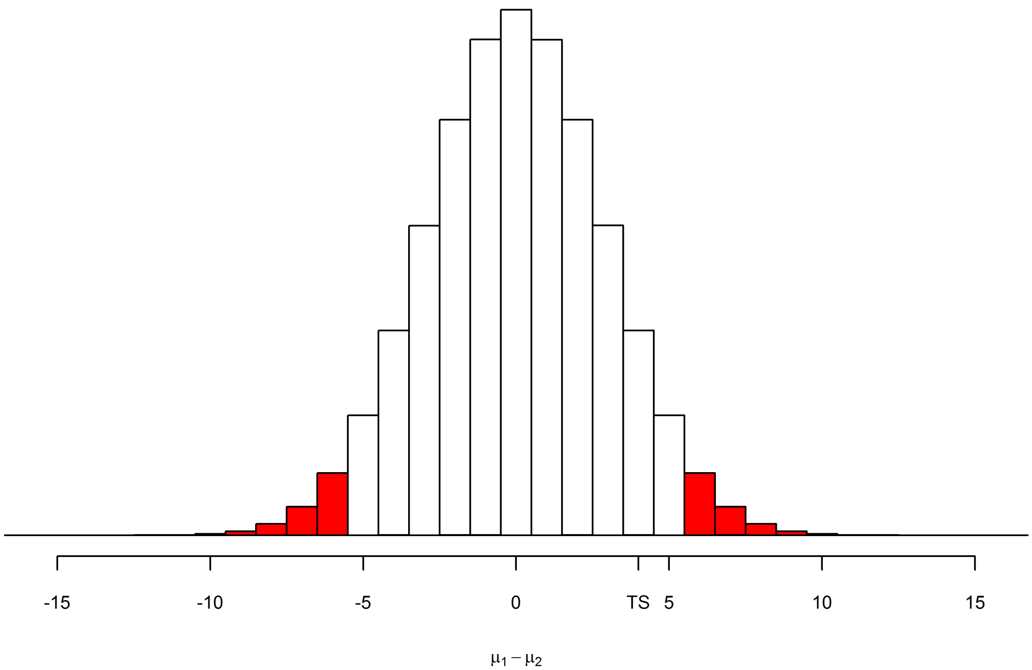

Figure 3. *A histogram of the test statistic resulting from ten million trials in a Monte Carlo experiment.*