Hypothesis Test and Comparison
STAT-UB.0001 – Statistics for Business Control

Test on a Population Mean

1. (Adapted from Stine and Foster, 4M 16.2). Does stock in IBM return a different amount on average than T-Bills? We will attempt to answer this question by using a dataset of the 264 monthly returns from IBM between 1990 and 2011. Over this period, the mean of the monthly IBM returns was 1.26% and the standard deviation was 8.27%. We will take as given that the expected monthly returns from investing in T-Bills is 0.3%.
   (a) What is the sample? What are the sample mean and standard deviation?

   (b) What is the relevant population? What are the interpretations of population mean and standard deviation?

   (c) What are the null and alternative hypotheses for testing whether or not IBM gives a different expected return from T-Bills (0.3%)?

(d) Use an appropriate test statistic to summarize the evidence against the null hypothesis.

(e) If the null hypothesis were true (there were no difference in expected monthly returns between IBM and T-Bills) what would be the chance of observing data at least as extreme as observed?

(f) Is there compelling evidence (at significance level 5%) of a difference in expected monthly returns between IBM and T-Bills?

(g) What assumptions do you need for the test to be valid? Are these assumptions plausible?

Test Statistic and Observed Significance Level (p-value)

2. In each of the following examples, for the hypothesis test with

$$H_0 : \mu = \mu_0$$
$$H_a : \mu \neq \mu_0$$

find the test statistic (t) and the p-value.

(a) $\mu_0 = 5$; $\bar{x} = 7$; $s = 10$; $n = 36$.

(b) $\mu_0 = 90$; $\bar{x} = 50$; $s = 200$; $n = 64$.

(c) $\mu_0 = 50$; $\bar{x} = 49.4$; $s = 2$; $n = 100$.

3. For each example from problem 2:

(a) Indicate whether a level 5% test would reject H_0.

(b) Indicate whether a level 1% test would reject H_0.

4. Before Facebook's recent redesign, the mean number of ad clicks per day was 100K. In the 49 days after the redesign, the mean number of ad clicks per day was 105K and the standard deviation was 35K. Is there significant evidence that the redesign affected the expected number of ad clicks? Perform a test at the 5% level.

   (a) What is the sample? What is the population?

   (b) What are the null and alternative hypotheses?

   (c) What is the test statistic?

   (d) Approximately what is the p-value?

   (e) What assumptions are you making?

   (f) What is α? What is the result of the test?

Types of Errors

5. In a hypothesis test, our decision will either be "reject H_0" or "do not reject H_0". Under what situations will each of these decisions be in error?

6. We reject H_0 when the p-value is below α.
   (a) If H_0 is true, what is the probability of making a Type I error?

   (b) If H_0 is false, what is the probability of *not* making a Type II error?

More p-values

7. Suppose we perform a hypothesis test and we observe a p-value of $p = .02$. True or false: There is a 2% chance that the null hypothesis is true.

8. Suppose we perform a hypothesis test and we observe a p-value of $p = .02$. True or false: If we reject the null hypothesis, then there is a 2% chance of making a type I error.

9. Suppose we perform a hypothesis test and we observe a T test statistic $t = -2.02$, corresponding to a p-value of $p = .02$. True or false: If we were to repeat the experiment and the null hypothesis were actually true, then there would be a 2% chance of observing a test statistic at least as extreme as $t = -2.02$.

Confidence Intervals for Comparing Means

10. Recall the class survey. Seventeen female and thirty male students filled out the survey, reporting (among other variables) their GMAT scores and interest levels in the course. We will use this data to compare females and males.

    (a) What are the relevant populations?

    (b) For the 14 female respondents who reported their GMAT scores, the mean was 721 and the standard deviation was 27. For the 28 male respondents, the mean was 720 and the standard deviation was 39. Find a 95% confidence interval for the difference in population means.

    (c) For the 17 female respondents who reported their interest levels in the course (1–10), the mean was 5.8 and the standard deviation was 1.8. For the 30 male respondents, the mean was 6.3 and the standard deviation was 2.1. Find a 95% confidence interval for the difference in population means.

    (d) For the confidence intervals you constructed in parts (b) and (c) to be valid, what assumptions need to be satisfied? How could you check these assumptions?

Hypothesis Tests for Comparing Means

11. Consider again the class survey data. We will use the data to evaluate whether or not there is a significant difference between the female and the male population means.

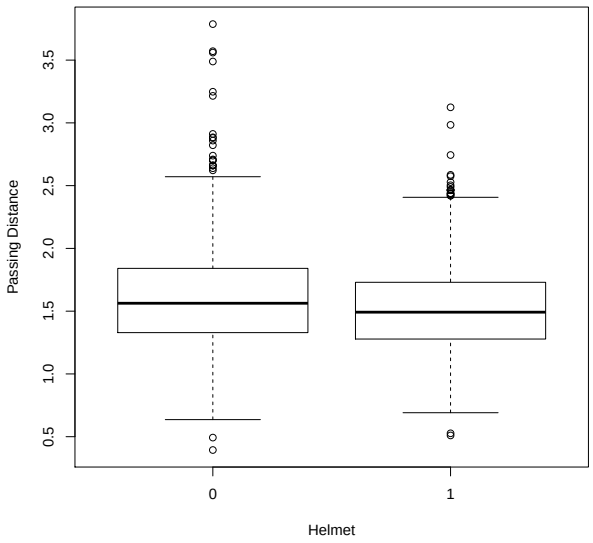
    (a) For the 14 female respondents who reported their GMAT scores, the mean was 721 and the standard deviation was 27. For the 28 male respondents, the mean was 720 and the standard deviation was 39. If the population means were equal what would be the chance of seeing a difference in sample means as large as observed?

    (b) For the 27 female respondents who reported their interest levels in the course (1–10), the mean was 5.8 and the standard deviation was 1.8. For the 30 male respondents, the mean was 6.3 and the standard deviation was 2.1. If the population means were equal what would be the chance of seeing a difference in sample means as large as observed?

    (c) What is the relationship between the confidence intervals in Question 10 and your answers to parts (a) and (b)?

Case Study: Bicycle Passing Distance

12. Here are boxplots of the passing distances (in meters) for a bike rider with and without a helmet. Is there evidence that the passing distance differs when the rider has a helmet?



Here are the sample statistics for the passing distance without a helmet: $n_1 = 1206$, $\bar{x}_1 = 1.61$, $s_1 = 0.405$. Here are the sample Here are the sample statistics for the passing distance with a helmet: $n_2 = 1149$, $\bar{x}_2 = 1.52$, $s_2 = 0.354$.

Formulate the problem as a hypothesis test, using significance level 5%.

(a) What are the populations?

(b) What are the null and alternative hypotheses?

(c) What are the samples?

(d) What is the test statistic?

(e) Approximately what is the p-value and the result of the test?

(f) Find a 95% confidence interval for the difference in passing difference with and without a helmet.