Chapter 18 Part Two: Hypothesis Testing

The idea behind hypothesis testing for means is similar to the hypothesis test for proportions. So our process is also similar.

1. Write null and alternative hypotheses
2. Check assumptions
3. Find sampling distribution
4. calculate a test statistic
5. Find a p-value
6. List your decision
7. state your conclusion

Step One: Write the Null and Alternative Hypotheses

Null Hypothesis: Remember, this represents the status quo or the claim we would like to test. So for means, the null hypothesis is typically of this form:

$$H_0: \mu = \mu_0$$

Alternative Hypothesis: There are three possibilities for alternative hypotheses. We pick each one based on the type of claim we are wanting to test:

- $H_A: \mu < \mu_0$
- $H_A: \mu > \mu_0$
- $H_A: \mu \neq \mu_0$

Step Two: Check Assumptions

Similar to our discussion of confidence interval estimation, we need to check our conditions before we can make concrete conclusions about the sampling distribution of our sampe mean.

- Random sample
- 10% condition
- Nearly normal

Step Three: Determine the Sampling Distribution

Recall from Chapter 15 Part Two, that if all three of the conditions listed above are true, then the sampling distribution can $\bar{Y} \sim N(\mu_0, \frac{\sigma}{\sqrt{n}})$

Step Four: Calculate Test Statistic

$$t = \frac{\bar{y} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Assuming the null hypothesis is true, this test statistic will follow a t-distribution with n-1 df.

Step Five: Find a p-value

Recall the definition of a p-value from our previous discussion of hypothesis testing of a proportion. A p-value is the probability of getting a value of the test statistic that is as or more extreme than the observed value calculated form the sample data ssuming the null hypothesis is true. There are three cases we must consider:

Find using JMP

- $H_A: \mu < \mu_0$
  p-value = the area less than the test statistic in a t-distribution with n-1 df
- $H_A: \mu > \mu_0$
  p-value = the area greater than the test statistic in a t-distribution with n-1 df.
- $H_A: \mu \neq \mu_0$
  p-value = the area in the two tails outside of $-|t|$ and $|t|$ in a t-distribution with n-1 degrees of freedom.

Step Six: List Your Decision

Small p-value

- Means the value of the sample mean was unlikely to occur given the null hypothesis is true.
- strong evidence against the null hypothesis
- If we have a small p-value we can reject the null hypothesis (more in a bit)

Large p-value

- Means the value of the sample mean was unlikely to occur given the null hypothesis is true
- Weak to no evidence against the null hypothesis
- Does not mean there is evidence for the null hypothesis
- If we have a large p-value we fail to reject the null hypothesis. (More in a bits

P-value	Evidence (against Ho)
Greater than .10	Little to no evidence
Between .05 and .10	Weak evidence
Between .01 and .05	Moderate Evidence
Less than .01	Strong evidence

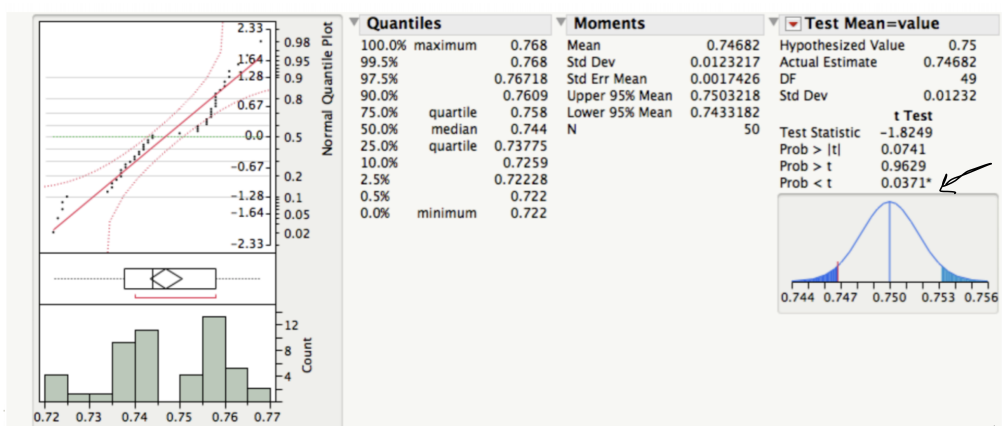
Step Seven: Conclusion/(E)xplain

After making a decision based on our p-value and test statistic, we need to make a conclusion in a meaningful way.

Your conclusions should include:

- the context of the problem
- the population mean (described in words)
- whether or not there is evidence for the alternative hypothesis.

Example *A simple random sample of 50 stainless steel metal screws is obtained from Crown Bolt, Inc., and the length of each screw is measured. The packaging indicates that the length of the screws is 0.75 in. There is concern that the mean length of the screws is less than the package suggests. We want to test to see if there is evidence to support this concern (claim). Conduct an appropriate hypothesis test.*



Quantiles

100.0%	maximum	0.768
99.5%		0.768
97.5%		0.76718
90.0%		0.7609
75.0%	quartile	0.758
50.0%	median	0.744
25.0%	quartile	0.73775
10.0%		0.7259
2.5%		0.72228
0.5%		0.722
0.0%	minimum	0.722

Moments

Mean	0.74682
Std Dev	0.0123217
Std Err Mean	0.0017426
Upper 95% Mean	0.7503218
Lower 95% Mean	0.7433182
N	50

Test Mean=value

Hypothesized Value	0.75
Actual Estimate	0.74682
DF	49
Std Dev	0.01232

t Test

Test Statistic	-1.8249
Prob > |t|	0.0741
Prob > t	0.9629
Prob < t	0.0371*

1) $H_0: \mu = 0.75in$

$H_A: \mu < 0.75in$

2) R-S: random sample taken

10%: probably more than 50 * 10 so met

N-N: Normal Q-Q plot looks O.K.

3) $\bar{x} \sim N(.75, \frac{.012}{\sqrt{50}})$

4) $t = \frac{.7468 - .75}{\frac{.012}{\sqrt{50}}} =$

5) p-value = 0.0371

6) moderate evidence against the null.

7) we have moderate evidence to suggest that the average screw length is less than 0.75in.