Stat 4473 – Data Analysis
HOW P-VALUES ARE CALCULATED
TEST STATISTICS

A hypothesis test uses the sample data in the form of a test statistic.

- The test is based on a statistic that compares the value of the parameter stated by the null hypothesis with an estimate of the parameter from the sample data.

- Values of the test statistic that are large in magnitude indicate that the estimate is far from the parameter value specified by H_o. These values give evidence against H_o. (The alternative hypothesis determines which directions count against H_o.)

- Example 1: Fire Insurance

  H_o: $\mu = 4.7$
  H_a: $\mu < 4.7$

  μ = average distance from homes to the nearest fire department

  test statistic is $t = \dfrac{\overline{X} - 4.7}{s/\sqrt{n}} \sim t(n-1)$

  Values of t that are large in the negative direction give good evidence that μ is not 4.7, but rather less than 4.7. (We are talking about values of t that are in the far left tail of the t-distribution.)

  p-value = probability of obtaining sample results as extreme or even more extreme than those obtained

  = $P(t \leq t_{obs})$

  A low p-value suggests that sample results like those obtained (or even more extreme than those obtained) rarely occur when H_o is true, and so we conclude that H_o is not true but rather H_a is true. That is, reject H_o in favor of H_a.

- Example 2: Study Habits for Older Students (SSHA test)

  H_o: $\mu = 115$
  H_a: $\mu > 115$

  μ = average SSHA score for older students

  test statistic is $t = \dfrac{\overline{X} - 115}{s/\sqrt{n}} \sim t(n-1)$

  Values of t that are large in the positive direction give good evidence that μ is not 115, but rather greater than 115. (We are talking about values of t that are in the far right tail of the t-distribution.)

  p-value = probability of obtaining sample results as extreme or even more extreme than those obtained

  = $P(t \geq t_{obs})$

  A low p-value suggests that sample results like those obtained (or even more extreme than those obtained) rarely occur when H_o is true, and so we conclude that H_o is not true but rather H_a is true. That is, reject H_o in favor of H_a.

- Example 3: Fill Amount at the Dairy

H_o: $\mu = 32$
H_a: $\mu \neq 32$

μ = average fill amount in 32 ounce cartons

test statistic is $t = \dfrac{\overline{X} - 32}{s/\sqrt{n}} \sim t(n-1)$

Values of *t* that are large in the negative direction OR large in the positive direction give good evidence that μ is not 32, but rather has deviated from 32. (We are talking about values of *t* that are in the far left tail of the *t*-distribution OR in the far right tail of the *t*-distribution.)

p-value = probability of obtaining sample results as extreme or even more extreme than those obtained

= $2\,P(\, t \geq |\, t_{obs} |\,)$

A low p-value suggests that sample results like those obtained (or even more extreme than those obtained) rarely occur when H_o is true, and so we conclude that H_o is not true but rather H_a is true. That is, reject H_o in favor of H_a.

H_o: population parameter = hypothesized value p-value = 2 P($t \geq |t_{obs}|$)
H_a: population parameter ≠ hypothesized value

H_o: population parameter = hypothesized value p-value = P($t \geq t_{obs}$)
H_a: population parameter > hypothesized value

H_o: population parameter = hypothesized value p-value = P($t \leq t_{obs}$)
H_a: population parameter < hypothesized value

for hypothesis tests involving μ

$$t = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} \sim t(n - 1)$$

for hypothesis tests involving μ_d

$$t = \frac{\bar{X}_d - 0}{\frac{s_d}{\sqrt{n}}} \sim t(n - 1)$$

for hypothesis tests involving $\mu_1 - \mu_2$

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t(df)$$

$$\textit{where} \quad df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2}$$