Statistical Inference Theory
Lesson 40
Small Sampling Theory

40.1 - What is the Student t Distribution for $\overline{X}$?

40.1 - Problem 1:
Step 1: Go to the top of the t distribution table D.

Step 2: Go across until you find $t_{.49}$.

Step 3: Move down until you reach the line df = 29 - 1 =28.

Step 4: From table D $t_{.49} = 2.467$.

40.1 - Problem 2:
We use the formula $\overline{X} = \mu \pm t\sigma_{\overline{X}}$.

Step 1: Since, $\sigma_{\overline{x}} = \frac{1}{\sqrt{9}} \approx 0.33$.

Step 2: For 80%, we look up $\frac{0.8}{2} = 0.40$ and df = 10 - 1 = 9 in the t - distribution table D.

Therefore, $t_{.40} = 1.383$.

Step 3: $\overline{X} = \mu \pm t_{.45}\sigma_{\overline{X}} = 1 \pm 1.383(0.33) \approx 1 \pm 0.46$

Step 4: All sample means lie in the interval $0.54 \le \overline{X} \le 1.46$ with 80% confidence.

40.2 - Estimating μ.

40.2 - Problem 1:

We use the formula $\overline{X} - ts_{\overline{X}} \le \mu \le \overline{X} + ts_{\overline{X}}$.

Step 1: Since we want a 95% confidence, we use $\frac{0.95}{2} = 0.475$ Therefore, $t = t_{.475.}$

Step 2: For N = 12, df = 12 - 1 = 11

Step 3: $s_{\overline{X}} = \dfrac{0.15}{\sqrt{12 - 1}} \approx 0.045$

Step 4: For df = 11 and $t_{.475}$ the t distribution table D gives t = 2.201.

Step 5: Since $\overline{X}$ =11.3, the above formula gives $11.3 - (2.2)(0.045) \leq \mu \leq 11.3 + (2.2)(0.045)$.

Step 6: From step 5 we have approximately $11.2 \leq \mu \leq 11.4$.

40.3- Statistical Decision Theory

40.3 - Problem 1:

➤(a).

Since we are interested in determining if the new feed will significantly increase production of eggs,

$H_o : \mu = 38$, (there is no increase in egg production).

$H_a : \mu > 38$ (there is a significant increase in egg production).

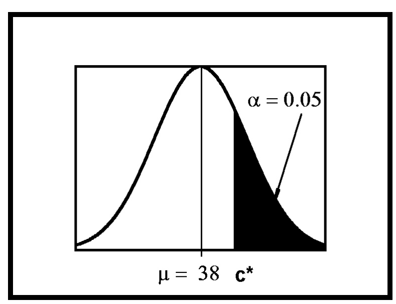
➤(b).

Since we are testing $\mu > 38$, we first state the decision rule as:

If $\overline{X} \geq c^* > 38$ then accept that the new feed is effect in increasing production. If $\overline{X} < c^*$ then assume the new feed does not increase production.

Step 1: We use the formula $c^* = 38 + t\sigma_{\overline{X}}$.

b.



Step 2: $\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{N - 1}}$

Step 3: Since N = 10 and σ = 2.5 .

then $\sigma_{\overline{X}} = \dfrac{2.5}{\sqrt{10 - 1}} \approx 0.83.$

Step 4: $c^* = 38 + t\sigma_{\overline{X}} = 38 + t(0.83)$

Step 5: To find t we need df = 10 - 1 = 9 and since α = 0.05, we find $t_{.45}$ in table D: $t_{.45} = 1.833$.

Step 6: $c^* = 38 + t\sigma_{\overline{X}} = 38 + t(0.83) = 38 + 1.833(0.83) \approx 39.52$

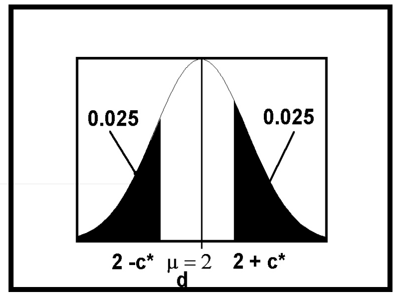
Step 7:The decision rule reads now:

If $\overline{X} \geq 39.52$ then accept that the new feed is effect in increasing production. If $\overline{X} < 39.52$ then assume the new feed does not increase production.

➤(c).
Since our sample resulted in an average $\overline{X} = 40 > 39.52$, we reject H_0, and accept the new feed increases production of eggs.

40.4 - The T distribution of the difference of means for small samples.

b



Problem 4.1:
➤(a).
Since we wish to check to seek if there is a significant difference in average number of eggs laid, we state the following:

H_o: $\mu_d = 2$, (there is a 2 egg difference in the average number of eggs laid).

H_a: $\mu_d \neq 2$, (there is not a 2 egg difference in the average number of eggs laid.)

➤(b).
This is a two sided test. We first write the decision rule as follows

Assume $\overline{X}_d = \overline{X}_2 - \overline{X}_1$. *If* $2 - c^* \leq \overline{X}_d \leq 2 + c^*$ *then reject* H_a*; otherwise reject* H_0 *and accept* H_a.

Here we assume H_o: $\mu_d = 2$.

Step 1: $\overline{X}_d = \overline{X}_2 - \overline{X}_1 = 58.2 - 55.6 = 2.6$

Step 2: $\sigma_d = \sqrt{\dfrac{N_1 s_1^2 + N_2 s_2^2)(N_1 + N_2)}{(N_1 + N_2 - 2)N_1 N_2}} = \sqrt{\dfrac{[(4)2.3^2 + (6)1.5^2](4 + 6)}{(4 + 6 - 2)(4)(6)}} \approx 1.35$

Step 3: We use the formula $c^* = t\sigma_d = t(1.35)$.

Step 4: Since $\alpha = 0.05$, and we have a two-sided test, we look up the t value from the t distribution table D for $4 + 6 - 2 = 8$ degrees of freedom: $t_{0.475} = 2.306$.

Step 5: $c^* = t\sigma_d = 2.306(1.35) \approx 3.11$

Step 6: We restate the decision rule:

If $2 - 3.11 = -1.11 \leq \overline{X}_d \leq 2 + 3.11 = 5.11$ then reject H_a; otherwise reject H_0 and accept H_a.

➤(c).
Since $\overline{X}_d = 2.6$ we reject H_a.

Supplementary Problems

1.
➤a.
Either the average age of all English majors is 21.5 or it is not 21.5 years old.

$H_o : \mu = 21.5$
$H_a : \mu \neq 21.5$

➤b.
For a type I error, we assume $\mu = 21.5$. Therefore, the decision rule will read:

If $21.5 - c^* \leq \overline{X} \leq 21.5 + c^*$ reject H_a; otherwise reject H_0 and accept H_a.

$$c* = t\sigma_{\overline{X}}, \text{ where } \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{N-1}}.$$

To compute σ, we first need to compute

$$\overline{X} = \frac{19 + 17 + 21 + 20 + 19 + 21 + 18 + 22 + 17 + 17}{10} = 19.1 .$$

$$\sigma^2 = [(19 - 19.1)^2 + (17 - 19.1)^2 + (21 - 19.1)^2 + (20 - 19.1)^2 + (19 - 19.1)^2 + (21 - 19.1)^2 +$$

$$(18 - 19.1)^2 + (22 - 19.1)^2 + (17 - 19.1)^2 + (17 - 19.1)^2]/ 10 = 3.09$$

$\sigma = 1.76$

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{N-1}} = \frac{1.76}{\sqrt{9}} \approx 0.59$$

$c^* = 0.59t$

For $\alpha/2 = 0.10/2 = 0.05$, from Table D, $t = 1.833$, for 9 degrees of freedom.

$c^* = 0.59(1.833) \approx 1.08$

Therefore, the decision rule reads:

If $21.5 - 1.08 = 20.42 \leq \overline{X} \leq 21.5 + 1.08 = 22.58$ reject H_a; otherwise reject H_0 and accept H_a.

➤c.
Since $\overline{X} = 19.1 < 20.42$, falls out of the interval, we reject H_0 and accept H_a. Our conclusion is from this sample the average age of English majors is not 21.5 years old.

2.
➤(a).

Let $\overline{X}_d = 1,894 - 1,768 = 126$ years

$$\sigma_d = \sqrt{\frac{N_1 s_1^2 + N_2 s_2^2)(N_1 + N_2)}{(N_1 + N_2 - 2)N_1 N_2}} = \sqrt{\frac{[(5)153^2 + (4)210^2](5 + 4)}{(5 + 4 - 2)(5)(4)}} \approx 137.35 .$$

Assume μ_1 is the average age determined by lab 1 and μ_2 is the average determined by lab 2. If the tests are consistent then $\mu_d = \mu_1 - \mu_2 = 0$.

Therefore, the decision rule will be If $-c^* \leq \overline{X}_d \leq c^*$, then conclude the 2 lab tests are consistent with each other.

$c^* = t\sigma_d = 137.35t$

For $\alpha/2 = 0.025$, then form table D, $t.475 = 2.365$.

$c^* = 137.35(2.365) = 324.83$

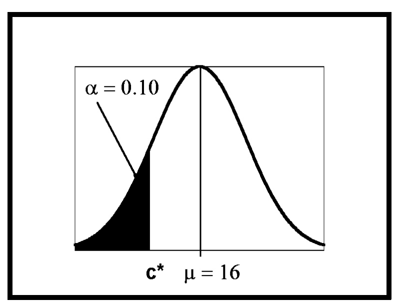
The decision rule reads: If $-324.83 \leq \overline{X}_d \leq 324.83$ then conclude the 2 lab tests are consistent with each other.

➤(b).
Since $-324.83 \leq \overline{X}_d = 126 \leq 324.83$, we conclude the 2 lab tests are consistent with each other.

3.
➤a.
Since the consumer is suspects that the machine is under filing

$H_o : \mu = 16$
$H_a : \mu < 16$.

c



➤b.
Since the degrees of freedom is 8 - 1 -= 7 and the type I error is 0.10, from Table D, we have $t_{0.40} = -1.415$.

➤c.
$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{N}} = \frac{0.31}{\sqrt{8 - 1}} \approx 0.12$$

$$c* = \mu + t\sigma_{\overline{X}} = 16 - 1.415(0.12) \approx 15.83$$

Since $\overline{X} = 15.6 < 15.83$, we do not reject her claim but accept it.