STAT 571, Solution for Assignment #9

November 12, 2003

1. (a.) inexpensive instrument: $\bar{x}_1 = 5.887$, $s_1^2 = 2.0164$, $n_1 = 8$,
   expensive instrument: $\bar{x}_2 = 6.8231$, $s_2^2 = 0.0569$, $n_2 = 13$,
   Test $H_0 : \mu_1 - \mu_2 = 0$ versus $H_A : \mu_1 - \mu_2 \neq 0$.
   (i) Assuming equal variances.
   $s_p^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2} = \frac{(8-1)2.0164+(13-1)0.0569}{8+13-2} = \frac{14.7976}{19} = 0.7788$,
   $s_p = \sqrt{.7788} = .882$,
   $s_{(\bar{X}_1-\bar{X}_2)} = \sqrt{s_p^2(\frac{1}{n_1} + \frac{1}{n_2})} = 0.397$.
   Test statistic $t = \frac{\bar{X}_1-\bar{X}_2}{s_{(\bar{X}_1-\bar{X}_2)}} = \frac{5.887-6.8231}{0.397} = -2.358$.
   $.02 <$ p-value $= 2P(T > 2.358) < .05$ (degrees of freedom of the t-distribution is 19).
   (ii) Assuming unequal variances.
   $s_{(\bar{X}_1-\bar{X}_2)} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0.506$,
   Test statistic $t' = \frac{\bar{X}_1-\bar{X}_2}{s_{(\bar{X}_1-\bar{X}_2)}} = \frac{5.887-6.8231}{0.506} = -1.85$
   adjusted degrees of freedom $adf = \frac{(vr_1+vr_2)^2}{(\frac{vr_1^2}{n_1-1})+(\frac{vr_2^2}{n_2-1})} = 7.24 = 7$,
   where $vr_1 = \frac{s_1^2}{n_1} = .252$, $vr_2 = \frac{s_2^2}{n_2} = .0044$
   $.10 <$ p-value $= 2P(T > 1.85) < .20$, using T-distribution with degrees of fredom 7.

   (b.) R output:

   ```
   Neither of the QQnorm plots display serious deviation from the
   normality assumption.
   ```

i.

```
 > t.test(inexpensive,expensive,var.equal=TRUE)

        Two Sample t-test

data:  inexpensive and expensive
t = -2.3597, df = 19, p-value = 0.02914
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.7654302 -0.1057237
sample estimates:
mean of x mean of y
 5.887500  6.823077
```

ii.

```
> t.test(inexpensive,expensive)

        Welch Two Sample t-test

data:  inexpensive and expensive
t = -1.8479, df = 7.244, p-value = 0.1057
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.1246107  0.2534568
sample estimates:
mean of x mean of y
 5.887500  6.823077
```

(c.) Use Levene's method to test the assumption of equal variances($\sigma_1^2 = \sigma_2^2$):

1) The median for inexpensive instrument is 5.8, and the median for expensive instrument is 6.8.

2) Calculate the absolute value of all deviations from the median. We get

inexpensive instrument: 1.3 2.3 0.1 2.1 1.4 0.4 0.1 0.8

expensive instrument: 0.1 0.2 0.4 0.5 0 0.1 0.2 0.2 0.3 0 0.1 0 0.2

3) Delete one value 0 in expensive instrument.

inexpensive instrument: 1.3 2.3 0.1 2.1 1.4 0.4 0.1 0.8

expensive instrument: 0.1 0.2 0.4 0.5 0.1 0.2 0.2 0.3 0 0.1 0 0.2

4) Perform a T-test for comparing the means of the two lists of numbers(with variances assumed equal). We obtain a p-value=.0026. Thus there's strong evidence that the variances σ_1^2 and σ_2^2 are not equal.

(d.) According to part(c), the t'-test assuming unequal variance is more appropriate. Since the expensive sample has a larger n, the pooled variance in the first test will be 'pulled' more in that direction.

2. (a). p_1 is the proportion of abnormal seeds early in the season,
p_2 is the proportion of abnormal seeds early in the season.
$H_0 : p_1 = p_2, \quad H_A : p_1 \neq p_2$.
Use Z statistic

$$Z = \frac{\hat{p_1} - \hat{p_2}}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}}.$$

From the data, we calculate

$$\hat{p_1} = \frac{25}{66} = 0.3788, \hat{p_2} = \frac{32}{115} = 0.2783, \hat{p} = \frac{25+32}{66+115} = 0.3149.$$

The one we've observed is: $z = 1.40$, then,
p-value $= 2\ P(Z \geq 1.40) = 2(0.0808) = 0.1616$.

(b). 99 % C. I. for $p_1 - p_2$ is given by

$$(\hat{p_1} - \hat{p_2}) \pm Z_{0.005}\sqrt{\frac{\hat{p_1}(1-\hat{p_1})}{n_1} + \frac{\hat{p_2}(1-\hat{p_2})}{n_2}}$$

$$-0.0878 \leq p_1 - p_2 \leq 0.2888.$$

(c). Assumptions:

1. Both samples constitute independent binomial experiements;
2. We assume independence of the trials from one sample to the other;
3. The sample size must be large enough so that we can use the normal approximations.

3.

(a). From the stem-leaf plot we know that both are not from normal distribution.

(b). 1) $n_1 = 12$, $n_2 = 15$.

2)

$$\begin{aligned} T^* &= \text{ sum of ranks in the small group} \\ &= 130.5 \end{aligned}$$

3) $T^{**} = n_1(n_1 + n_2 + 1) - T^* = 205.5$

4) $T = min(T^*, T^{**}) = 130.5$

5) From the table, at $\alpha = 5\%$, reject Ho if $T \leq 127$,
$0.05 < p - value$. We have no evidence to reject Ho.

4. a. $S_p^2 = \frac{S_1^2+S_2^2+S_3^2+S_4^2}{4} = 5.95$.
$V^2 = \frac{(N-4)S_p^2}{\sigma^2} \sim \chi^2_{N-4}, N = 4(8) = 32.$
95 % CI for σ^2 is

$$\frac{(32-4)(5.95)}{\chi^2_{28,0.025}} \leq \sigma^2 \leq \frac{(32-4)(5.95)}{\chi^2_{28,0.975}}$$

$$3.747 \leq \sigma^2 \leq 10.882.$$

b.

$$s_p^2 = 3.25, 2.047 \leq \sigma^2 \leq 5.944,$$

$$s_p^2 = 9.41, 5.928 \leq \sigma^2 \leq 17.211,$$

$$s_p^2 = 6.46, 4.070 \leq \sigma^2 \leq 11.8115.$$

Since s_4^2 is the smallest one, s_p^2 is the smallest when we assign 26 observations to the fourth group; While s_2^2 is the largest one, so s_p^2 is the largest when we assign 22 observations to the second group; and s_p^2 for the third assignment is the middle one.

5. a. ANOVA table

Source	df	SS	MS
Treatment	3	5521.67	1840.56
Error	8	2956	369.5
Total	11	8477.67	

b. 1. Assumption

(a) Independence assumption: Observations within a treatment and across all treatments are independent

(b) Normal assumption: $X_{ij} \sim N(\mu_i, \sigma_i^2) \quad i = 1, \ldots, k, \quad j = 1, \ldots, n_i$

(c) Equal variance assumption: $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2$

2. Notations

$$\begin{aligned} \text{SSTot} &= \sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}..)^2 = \sum_{\text{all obs}}(x_{ij})^2 - \frac{(x..)^2}{N} \\ \text{SSTrt} &= \sum_{i=1}^{k} n_i(\bar{x}. - \bar{x}..)^2 = \sum_{i=1}^{k}\frac{1}{n_i}(x_{i}.)^2 - \frac{(x..)^2}{N} \\ \text{SSErr} &= \sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_{i}.)^2 = \sum_{i=1}^{k}(n_i - 1)s_i^2 \end{aligned}$$

3. Model : $X_{ij} = \mu_i + e_{ij}$ or $X_{ij} = \mu + \alpha_i + e_{ij}$, where $e_{ij} \sim N(0, \sigma^2)$

c. Hypothesis $H_0 : \mu_1 = \mu_2 = ... = \mu_k$ $\quad H_A$:not all μ_i are equal.
Test statistics : $F = \frac{1840.56}{369.5} = 4.98$
p-value $= P(F_{3,8} \geq 4.98)$
From the table, $F_{3,8,0.05} = 4.07, F_{3,8,0.01} = 7.59$.
So $0.01 < p - value < 0.05$.
We have moderate evidence to reject Ho.

6.

$$\begin{aligned} & \text{power}(\mu_2 = 20, \mu_1 = 28) \\ = \; & P(\text{reject Ho}|\mu_2 = 20, \mu_1 = 28) \\ = \; & P(\bar{X}_1 - \bar{X}_2 > 6|\mu_2 = 20, \mu_1 = 28) \\ = \; & P(Z \geq -.96) \\ = \; & .8315 \end{aligned}$$