STAT 4164
Stat Methods III

Homework 1 Solutions

1. (**Extra Problem**) By the definition of $F(n_1, n_2)$ distribution, we have

$$F = \frac{\frac{(n_1-1)S_X^2}{\sigma_1^2}/(n_1-1)}{\frac{(n_2-1)S_Y^2}{\sigma_2^2}/(n_2-1)} \sim F(n_1, n_2).$$

Note that F can be rewritten as

$$F = \frac{\sigma_2^2 S_X^2}{\sigma_1^2 S_Y^2} \sim F(n_1, n_2).$$

It follows that

$$\mathrm{P}\left\{F^{\alpha/2}(n_1, n_2) \le F \le F^{1-\alpha/2}(n_1, n_2)\right\} = 1 - \alpha,$$

where $F^{\alpha/2}(n_1, n_2)$ and $F^{1-\alpha/2}(n_1, n_2)$ are the $\alpha/2$ and $1-\alpha/2$ quantile of the $F(n_1, n_2)$ distribution. After bringing the expression of F, it becomes

$$\mathrm{P}\left\{F^{\alpha/2}(n_1, n_2) \le \frac{\sigma_2^2 S_X^2}{\sigma_1^2 S_Y^2} \le F^{1-\alpha/2}(n_1, n_2)\right\} = 1 - \alpha.$$

Or equivalently,

$$\mathrm{P}\left\{\frac{S_Y^2}{S_X^2} F^{\alpha/2}(n_1, n_2) \le \frac{\sigma_2^2}{\sigma_1^2} \le \frac{S_Y^2}{S_X^2} F^{1-\alpha/2}(n_1, n_2)\right\} = 1 - \alpha.$$

Namely, a $100 \times (1-\alpha)\%$ confidence interval for $\frac{\sigma_2^2}{\sigma_1^2}$ is

$$\left(\frac{S_Y^2}{S_X^2} F^{\alpha/2}(n_1, n_2), \quad \frac{S_Y^2}{S_X^2} F^{1-\alpha/2}(n_1, n_2)\right).$$

2. (**Extra Problem**) Based on the given sample data, it can be found that $n_1 = 6$, $\bar{x} = 12.29$, $s_x = 2.126$, and $n_2 = 10$, $\bar{y} = 26.92$, $s_y = 3.038$.

   (a) To test H_0: $\sigma_1^2 = \sigma_2^2$ vs. H_a at the significance level $\alpha = .05$: $\sigma_1^2 \ne \sigma_2^2$, use the test statistic

$$F = \frac{s_y^2}{s_x^2} = \frac{3.038^2}{2.126^2} = 2.04.$$

   Here, for the convenience of using the F table given in the appendix, we always place the larger variance on the numerator and the smaller one on the denominator.

Decision Rule: reject H_0 if $F > F^{1-\alpha}(n_1 - 1, n_2 - 1) = F^{.95}(5, 9) = 3.48$.

Conclusion: since the observed F is 2.04, which does not fall into the rejection region, we can NOT reject the null hypothesis H_0 at $\alpha = .05$. We may conclude that the data does not provide sufficient evidence to support that the two normal populations have different variance.

(b) To test H_0: $\mu_1 = \mu_2$ vs. H_a: $\mu_1 \neq \mu_2$ at $\alpha = .05$, the pooled t test is used. We need the following three statistical assumptions:

i. Two independent random samples are available.
ii. Both populations are normal.
iii. Two populations must have the same common variance.

We first compute the pooled estimate for the common variance:

$$s_p^2 = \frac{(n_1 - 1)s_x^2 + (n_2 - 1)s_y^2}{n_1 + n_2 - 2} = \frac{(6-1) \times 2.126^2 + (10-1) \times 3.038^2}{6 + 10 - 2} = 7.547.$$

Then the pooled t test statistic is

$$t = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\sqrt{s_p^2 \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(12.29 - 26.92) - 0}{\sqrt{7.54753 \times \left(\frac{1}{6} + \frac{1}{10}\right)}} = -10.31.$$

Decision Rule: reject H_0 if $|t| > t^{1-\alpha/2}(n_1 + n_2 - 2) = t^{.975}(14) = 2.14$.

Conclusion: the null hypothesis should be rejected at $\alpha = .05$ and we conclude that significant difference does exist between two population means.