Ass –ump-TIONS

- I see a linear model
- I see assumptions in epsilon

X2

- Independent Identically Distributed
- Normal Zero sigma squared.

$$\varepsilon_i \overset{iid}{\sim} N(0,\sigma^2)$$

Ass 4 - helps

1. MS 7.118 - pg 364
2. MS 7.120 - pg 365
3. MS 7.128 - pg 367
4. MS 8.24 - pg 390
5. MS 8.28 - pg 392
6. MS 8.44 - pg 401
7. MS 8.84 - pg 425

Qu 1: 7.118

The point estimated for the true proportion of flightless birds for the extinct species is
$\hat{p} = \frac{Y}{n} = \frac{21}{38} = 0.5526$.

The point estimated for the true proportion of flightless birds for the nonextinct species is
$\hat{p} = \frac{Y}{n} = \frac{7}{78} = 0.0897$.

For Normal approximation to be good enough (CLT)

- $np \geq 4, nq \geq 4$

Q2: 7.120

$$\left(\overline{y}_1-\overline{y}_2\right)\pm z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\Rightarrow\left(1{,}312-1{,}352\right)\pm1.645\sqrt{\frac{422^2}{100}+\frac{271^2}{47}}$$

$$\Rightarrow -40\pm95.118\Rightarrow\left(-135.118,\ 55.118\right)$$

Q2 (cont)

For confidence coefficient 0.90, $\alpha = 0.10$ and $\alpha/2 = 0.10/2 = 0.05$. Using a computer package with $\nu_1 = n_1 - 1 = 100 - 1 = 99$ and $\nu_2 = n_2 - 1 = 47 - 1 = 46$ degrees of freedom, $F_{0.05,(99,46)} = 1.54818$ and $F_{0.05,(46,99)} = 1.49194$. The 90% confidence interval is:

$$\frac{422^2}{271^2} \cdot \frac{1}{1.54818} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{422^2}{271^2}(1.49194) \Rightarrow 1.566 \le \frac{\sigma_1^2}{\sigma_2^2} \le 3.618$$

Q3: 7.128

a. Y has a normal distribution with $\mu = 0$ and σ.

$$Z = \frac{Y - \mu}{\sigma} = \frac{Y - 0}{\sigma} = \frac{Y}{\sigma}$$ has a standard normal distribution

Q3 (cont)

$Z^2 = \frac{Y^2}{\sigma^2}$ will have a χ^2 distribution with 1 degree of freedom

Q3: Cont

$$P\left(\chi^2_{1-\alpha/2} \le \frac{Y^2}{\sigma^2} \le \chi^2_{\alpha/2}\right) = 1-\alpha$$

$$= P\left(\frac{1}{\chi^2_{1-\alpha/2}} \ge \frac{\sigma^2}{Y^2} \ge \frac{1}{\chi^2_{\alpha/2}}\right) \quad = P\left(\frac{1}{\chi^2_{\alpha/2}} \le \frac{\sigma^2}{Y^2} \le \frac{1}{\chi^2_{1-\alpha/2}}\right)$$

$$= P\left(\frac{Y^2}{\chi^2_{\alpha/2}} \le \sigma^2 \le \frac{Y^2}{\chi^2_{1-\alpha/2}}\right)$$

Q4: 8.24 (one sample t-test)

a. Let μ = mean surface roughness of coated interior pipe used in oil fields. To determine if this mean differs from 2 micrometers, we test:

$$H_0 : \mu = 2$$
$$H_a : \mu \neq 2$$

Q5: 8:28

To determine if the mean DOC value differs from 15, we test:

$$H_0 : \mu = 15$$
$$H_a : \mu \neq 15$$

The test statistic is $t = \dfrac{\overline{y} - \mu_0}{s/\sqrt{n}} = \dfrac{14.52 - 15}{12.96/\sqrt{25}} = -0.185$.

Q5 (cont)

b. We must find the rejection region in part a in terms of $\overline{y}$. We know

$$t = \frac{\overline{y} - \mu_0}{s/\sqrt{n}} \Rightarrow t\left(\frac{s}{\sqrt{n}}\right) = \overline{y} - \mu_0 \Rightarrow \overline{y} = \mu_0 + t\left(\frac{s}{\sqrt{n}}\right)$$

For $t = -1.711$, $\overline{y} = \mu_0 + t\left(\frac{s}{\sqrt{n}}\right) = 15 - 1.711\left(\frac{12.96}{\sqrt{25}}\right) = 10.565$

For $t = 1.711$, $\overline{y} = \mu_0 + t\left(\frac{s}{\sqrt{n}}\right) = 15 + 1.711\left(\frac{12.96}{\sqrt{25}}\right) = 19.435$

Q5 (cont)

Thus, we would reject H_0 if $\overline{y} < 10.565$ or $\overline{y} > 19.435$.

We want to find

$$P(\overline{y} < 10.565 \mid \mu_a = 14) + P(\overline{y} > 19.435 \mid \mu_a = 14)$$

$$= P\left(t < \frac{10.565 - 14}{\frac{12.96}{\sqrt{25}}} \right) + P\left(t > \frac{19.435 - 14}{\frac{12.96}{\sqrt{25}}} \right) = P(t < -1.33) + P(t > 2.10)$$

$$= 0.0980 + 0.0232 = 0.1212$$

Q6: 8.44 (2 sample t-test)

8.44 Let μ_1 = mean oxon/thion ratio for foggy days and μ_2 = mean oxon/thion ratio for cloudy/clear days.

Q6 (cont)

Some preliminary calculations are:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(8-1)0.1186^2 + (4-1)0.1865^2}{8+4-2} = 0.02028$$

To determine if the mean oxon/thion ratios differ for foggy and cloudy/clear days, we test:

$$H_0 : \mu_1 - \mu_2 = 0$$
$$H_a : \mu_1 - \mu_2 \neq 0$$

The test statistic is $t = \dfrac{(\bar{y}_1 - \bar{y}_1) - D_0}{\sqrt{s_p^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} = \dfrac{(0.2738 - 0.4521) - 0}{\sqrt{0.02028\left(\dfrac{1}{8} + \dfrac{1}{4}\right)}} = -2.045$.

Q7: 8.84 (Just use var.test())

a. Let σ_1^2 = the equality of heat rate variance for traditional gas turbines and σ_2^2 = the equality of heat rate variance for aeroderivative augmented gas turbines. To determine if there is a difference in the variation of the two gas turbine types, we test:

$$H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1$$

$$H_a : \frac{\sigma_1^2}{\sigma_2^2} \neq 1$$

To get the rejection acceptance/reject region

```
qf(c(alpha/2,1-alpha/2), df1,df2)
```

```
> qf(c(0.05/2,1-0.05/2),6,38)
[1] 0.1991693 2.7633350
```

The test statistic is $F = \dfrac{\text{Larger sample variance}}{\text{Smaller sample variance}} = \dfrac{2{,}652^2}{1{,}279^2} = 4.299$.

The rejection region requires $\alpha / 2 = 0.05 / 2 = 0.025$ in the upper tail of the F distribution. Using a computer, with $\nu_1 = n_2 - 1 = 7 - 1 = 6$ and $\nu_2 = n_1 - 1 = 39 - 1 = 38$ degrees of freedom, $F_{0.025} = 2.763$. The rejection region is $F > 2.763$.

Same as previous method (a)

b. Let σ_1^2 = the equality of heat rate variance for advanced gas turbines and σ_2^2 = the equality of heat rate variance for aeroderivative augmented gas turbines. To determine if there is a difference in the variation of the two gas turbine types, we test:

```
> qf(c(0.05/2,1-0.05/2),6,20)
[1] 0.1934834 3.1283400
```

The rejection region requires $\alpha/2 = 0.05/2 = 0.025$ in the upper tail of the F distribution. From Table 11, Appendix B, with $\nu_1 = n_2 - 1 = 7 - 1 = 6$ and $\nu_2 = n_1 - 1 = 21 - 1 = 20$ degrees of freedom, $F_{0.025} = 3.13$. The rejection region is $F > 3.13$.

Q8: 8.99 (var.test() – this is a repeat of Q7)

a. Let σ_1^2 = variance of the number of ant species in the Dry Steppe and σ_2^2 = variance of the number of ant species in the Gobi Dessert. To determine if there is a difference in the variation at the two locations, we test:

$$H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1$$

$$H_a: \frac{\sigma_1^2}{\sigma_2^2} \neq 1$$

Q9: 8.104 (paired samples)

To determine if a difference exists between the mean throughput rates of human and automated methods, we test:

$$H_0 : \mu_1 - \mu_2 = 0$$
$$H_a : \mu_1 - \mu_2 \neq 0$$

The test statistic is $t = \dfrac{\bar{d} - D_0}{\frac{s_d}{\sqrt{n}}} = \dfrac{-32.6 - 0}{\frac{35.0}{\sqrt{8}}} = -2.63$.

The rejection region requires $\alpha / 2 = 0.05 / 2 = 0.025$ in each tail of the t distribution. From Table 7, Appendix B, with $df = n - 1 = 8 - 1 = 7$, $t_{0.025} = 2.365$. The rejection region is $t < -2.365$ or $t > 2.365$.