Problem Set 5
Solutions

Economics 310- Professor Menzie D. Chinn
Fall 2003
University of Wisconsin-Madison

8.70 a. The power of a test increases when:

1. The distance between the null and alternative values of μ increases.
2. The value of α increases.
3. The sample size increases.

b. The power of a test is equal to $1 - \beta$. As β increases, the power decreases.

8.82 a. It would be necessary to assume that the population has a normal distribution.

b. H_0: $\sigma^2 = 1$
H_a: $\sigma^2 > 1$

The test statistic is $\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{6(4.84)}{1} = 29.04$

The rejection region requires $\alpha = .05$ in the upper tail of the χ^2 distribution with df $= n - 1 = 7 - 1 = 6$. From Table VII, Appendix B, $\chi^2_{.05} = 12.5916$. The rejection region is $\chi^2 > 12.5916$.

Since the observed value of the test statistic falls in the rejection region ($29.04 > 12.5916$), H_0 is rejected. There is sufficient evidence to indicate that the variance is greater than 1 at $\alpha = .05$.

c. H_0: $\sigma^2 = 1$
H_a: $\sigma^2 \neq 1$

The test statistic is $\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{6(4.84)}{1} = 29.04$

The rejection region requires $\alpha/2 = .025$ in the upper tail of the χ^2 distribution with df $= n - 1 = 7 - 1 = 6$. From Table VII, Appendix B, $\chi^2_{.975} = 1.237347$ and $\chi^2_{.025} = 14.4494$. The rejection region is $\chi^2 < 1.237347$ or $\chi^2 > 14.4494$.

Since the observed value of the test statistic falls in the rejection region ($29.04 > 14.4494$), H_0 is rejected. There is sufficient evidence to indicate that the variance is not equal to 1 at $\alpha = .05$.

9.2 a. $\mu_{\bar{x}_1} = \mu_1 = 12$ $\sigma_{\bar{x}_1} = \frac{\sigma_1}{\sqrt{n_1}} = \frac{4}{\sqrt{64}} = .5$

b. $\mu_{\bar{x}_2} = \mu_2 = 10$ $\sigma_{\bar{x}_2} = \frac{\sigma_2}{\sqrt{n_2}} = \frac{3}{\sqrt{64}} = .375$

c. $\mu_{\bar{x}_1-\bar{x}_2} = \mu_1 - \mu_2 = 12 - 10 = 2$

$$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{4^2}{64} + \frac{3^2}{64}} = \sqrt{\frac{25}{64}} = \frac{5}{8} = .625$$

d. Since $n_1 \geq 30$ and $n_2 \geq 30$, the sampling distribution of $\bar{x}_1 - \bar{x}_2$ is approximately normal by the Central Limit Theorem.

9.4 Assumptions about the two populations:

1. Both sampled populations have relative frequency distributions that are approximately normal.
2. The population variances are equal.

Assumptions about the two samples:

The samples are randomly and independently selected from the population.

9.6 a. $s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(25 - 1)120 + (25 - 1)100}{25 + 25 - 2} = \frac{5280}{48} = 110$

b. $s_p^2 = \frac{(20 - 1)12 + (10 - 1)20}{20 + 10 - 2} = \frac{408}{28} = 14.5714$

c. $s_p^2 = \frac{(6 - 1).15 + (10 - 1).2}{6 + 10 - 2} = \frac{2.55}{14} = .1821$

d. $s_p^2 = \frac{(16 - 1)3000 + (17 - 1)2500}{16 + 17 - 2} = \frac{85,000}{31} = 2741.9355$

e. s_p^2 falls near the variance with the larger sample size.

9.20 a. To determine if the mean annual percentage turnover for U.S. plants exceeds that for Japanese plants, we test:

$$H_0\text{: } \mu_1 - \mu_2 = 0$$
$$H_a\text{: } \mu_1 - \mu_2 > 0$$

The test statistic is $t = 4.46$ (from printout).

The rejection region requires $\alpha = .05$ in the upper tail of the t-distribution with df = $n_1 + n_2 - 2 = 5 + 5 - 2 = 8$. From Table VI, Appendix B, $t_{.05} = 1.860$. The rejection region is $t > 1.860$.

Since the observed value of the test statistic falls in the rejection region ($t = 4.46 > 1.860$), H_0 is rejected. There is sufficient evidence to indicate the mean annual percentage turnover for U.S. plants exceeds that for Japanese plants at $\alpha = .05$.

b. The observed significance is $.0031/2 = .00155$.

Since the p-value is so small, there is evidence to reject H_0 for $\alpha > .005$.

c. The necessary assumptions are:

1. Both sampled populations are approximately normal.
2. The population variances are equal.
3. The samples are randomly and independently sampled.

There is no indication that the populations are not normal. Both sample variances are similar, so there is no evidence the population variances are unequal. There is no indication the assumptions are not valid.

9.28 a. $H_0\text{: } \mu_1 - \mu_2 = 0$
$H_a\text{: } \mu_1 - \mu_2 < 0$

The rejection region requires $\alpha = .10$ in the lower tail of the t-distribution with df = $n_D - 1 = 18 - 1 = 17$. From Table VI, Appendix B, $t_{.10} = 1.333$. The rejection region is $t < -1.333$.

b. $H_0\text{: } \mu_1 - \mu_2 = 0$
$H_a\text{: } \mu_1 - \mu_2 < 0$

The test statistic is $t = \dfrac{\bar{x}_D - 0}{\dfrac{s_D}{\sqrt{n_D}}} = \dfrac{-3.5 - 0}{\dfrac{\sqrt{21}}{\sqrt{18}}} = -3.24$

The rejection region is $t < -1.333$. (Refer to part a).

Since the observed value of the test statistic falls in the rejection region ($t = -3.24 < -1.333$), H_0 is rejected. There is sufficient evidence to indicate $\mu_1 - \mu_2 < 0$ at $\alpha = .10$.

c. The necessary assumptions are:

1. The population of differences is normal.
2. The differences are randomly selected.

d. For confidence coefficient .90, $\alpha = 1 - .90 = .10$ and $\alpha/2 = .10/2 = .05$. From Table VI, Appendix B, with df = 17, $t_{.05} = 1.740$. The confidence interval is:

$$\bar{x}_D \pm t_{.05}\frac{s_D}{\sqrt{n_D}} \Rightarrow -3.5 \pm 1.740\frac{\sqrt{21}}{\sqrt{18}} \Rightarrow -3.5 \pm 1.88 \Rightarrow (-5.38, -1.62)$$

e. The confidence interval provides more information since it gives an interval of possible values for the difference between the population means.

9.34 Some preliminary calculations are:

Working Days	Difference (Design 1 − Design 2)
8/16	−53
8/17	−271
8/18	−206
8/19	−266
8/20	−213
8/23	−183
8/24	−118
8/25	−87

$$\bar{x}_D = \frac{\sum x_D}{n_D} = \frac{-1{,}397}{8} = -174.625$$

$$s_D^2 = \frac{\sum x_D^2 - \frac{\left(\sum x_D\right)^2}{n_D}}{n_D - 1} = \frac{289{,}793 - \frac{(-1{,}397)^2}{8}}{8 - 1} = 6{,}548.839$$

$$s_D = \sqrt{s_D^2} = \sqrt{6{,}548.839} = 80.925$$

a. For confidence coefficient .95, $\alpha = .05$ and $\alpha/2 = .05/2 = .025$. From Table VI, Appendix B, with df $= n_D - 1 = 8 - 1 = 7$, $t_{.025} = 2.365$. The 95% confidence interval is:

$$\bar{x}_D \pm t_{.025}\frac{s_D}{\sqrt{n_D}} \Rightarrow -174.625 \pm 2.365\frac{80.925}{\sqrt{8}} \Rightarrow -174.625 \pm 67.666$$

$$\Rightarrow (-242.291, -106.959)$$

We are 95% confident that the difference in mean daily output of the two designs is between −242.291 and −106.959.

b. We must assume that the population of differences is normal and that the sample of differences is randomly selected.

c. Yes. Since 0 is not contained in the confidence interval and the endpoints are both negative, there is evidence to indicate that Design 2 is superior to Design 1.

9.72 a. To determine if the variance for population 2 is greater than that for population 1, we test:

H_0: $\sigma_1^2 = \sigma_2^2$
H_a: $\sigma_1^2 < \sigma_2^2$

The test statistic is $F = \frac{s_2^2}{s_1^2} = \frac{2.9729^2}{1.4359^2} = 4.29$

The rejection region requires $\alpha = .05$ in the upper tail of the F-distribution with $\nu_1 = n_2 - 1 = 5 - 1 = 4$ and $\nu_2 = n_1 - 1 = 6 - 1 = 5$. From Table IX, Appendix B, $F_{.05} = 5.19$. The rejection region is $F > 5.19$.

Since the observed value of the test statistic does not fall in the rejection region ($F = 4.29 \not> 5.19$), H_0 is not rejected. There is insufficient evidence to indicate the variance for population 2 is greater than that for population 1 at $\alpha = .05$.

b. The p-value is $P(F \geq 4.29)$. From Tables VIII and IX, with $\nu_1 = 4$ and $\nu_2 = 5$,

$.05 < P(F \geq 4.29) < .10$

There is no evidence to reject H_0 for $\alpha < .05$ but there is evidence to reject H_0 for $\alpha = .10$.

10.10 a.

x_i	y_i	x_i^2	x_iy_i
7	2	$7^2 = 49$	$7(2) = 14$
4	4	$4^2 = 16$	$4(4) = 16$
6	2	$6^2 = 36$	$6(2) = 12$
2	5	$2^2 = 4$	$2(5) = 10$
1	7	$1^2 = 1$	$1(7) = 7$
1	6	$1^2 = 1$	$1(6) = 6$
3	5	$3^2 = 9$	$3(5) = 15$

Totals:
$\sum x_i = 7 + 4 + 6 + 2 + 1 + 1 + 3 = 24$
$\sum y_i = 2 + 4 + 2 + 5 + 7 + 6 + 5 = 31$
$\sum x_i^2 = 49 + 16 + 36 + 4 + 1 + 1 + 9 = 116$
$\sum x_iy_i = 14 + 16 + 12 + 10 + 7 + 6 + 15 = 80$

b. $SS_{xy} = \sum x_iy_i - \frac{\left(\sum x_i\right)\left(\sum y_i\right)}{n} = 80 - \frac{(24)(31)}{7} = 80 - 106.2857143 = -26.2857143$

c. $SS_{xx} = \sum x_i^2 - \frac{\left(\sum x_i\right)^2}{7} = 116 - \frac{(24)^2}{7} = 116 - 82.28571429 = 33.71428571$

d. $\hat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}} = \frac{-26.2857143}{33.71428571} = -.779661017 \approx -.7797$

e. $\bar{x} = \frac{\sum x_i}{n} = \frac{24}{7} = 3.428571429$ $\qquad \bar{y} = \frac{\sum y_i}{n} = \frac{31}{7} = 4.428571429$

f. $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} = 4.428571429 - (-.779661017)(3.428571429)$
$= 4.428571429 - (-2.673123487) = 7.101694916 \approx 7.102$

g. The least squares line is $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x = 7.102 - .7797x$.

10.22 a. $SSE = SS_{yy} - \hat{\beta}_1 SS_{xy} = 95 - .75(50) = 57.5$

$$s^2 = \frac{\sum x}{n} = \frac{57.5}{20 - 2} = 3.19444$$

b. $$SS_{yy} = \sum y^2 - \frac{(\sum y)^2}{n} = 860 - \frac{50^2}{40} = 797.5$$

$SSE = SS_{yy} - \hat{\beta}_1 SS_{xy} = 797.5 - .2(2700) = 257.5$

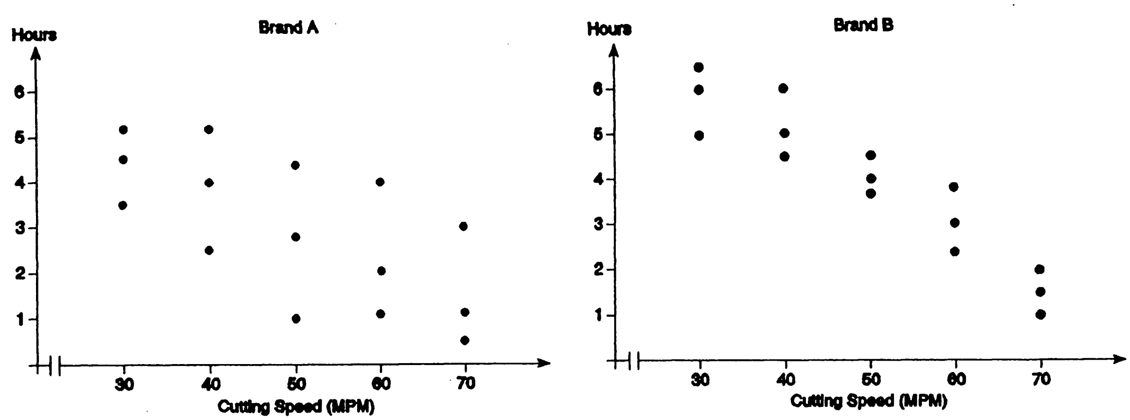
$$s^2 = \frac{SSE}{n - 2} = \frac{257.5}{40 - 2} = 6.776315789 \approx 6.7763$$

c. $SS_{yy} = \sum (y_i - \bar{y})^2 = 58$ $\quad \hat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}} = \frac{91}{170} = .535294117$

$SSE = SS_{yy} - \hat{\beta}_1 SS_{xy} = 58 - .535294117(91) = 9.2882353 \approx 9.288$

$$s^2 = \frac{SSE}{n - 2} = \frac{9.2882353}{10 - 2} = 1.161029413 \approx 1.1610$$

10.30 a.



b. For Brand A, from printout, $\hat{y} = 6.62 - .0727x$
For Brand B, $\hat{y} = 9.31 - .1077x$

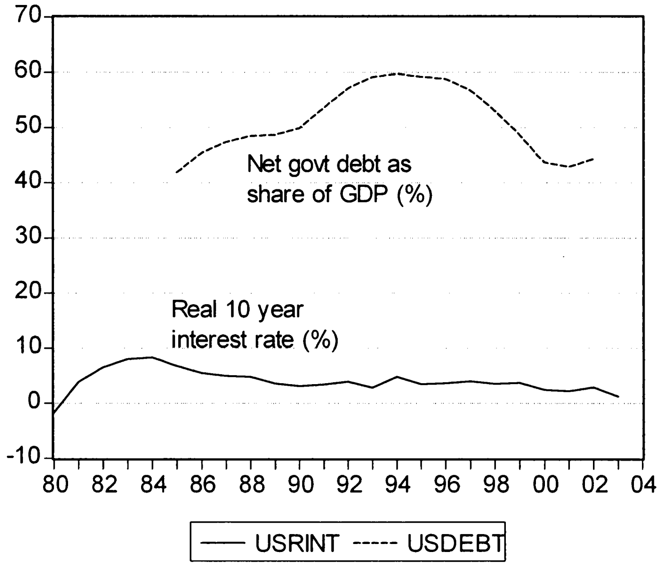
c. For Brand A, SSE = 19.056, s^2 = MSE = 1.465846154, and s = 1.210721336
For Brand B, SSE = 4.833, s^2 = MSE = .371769231, and s = .609728817

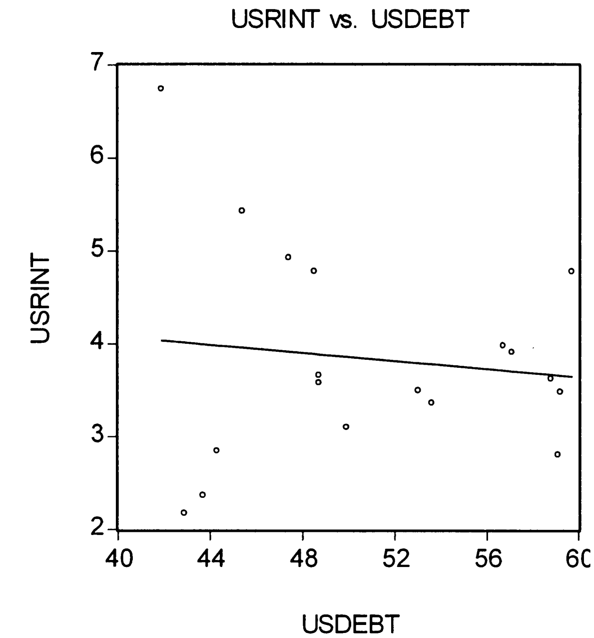
d. For Brand A, $\hat{y} = 6.62 - .0727x$. For $x = 70$, $\hat{y} = 6.62 - .0727(70) = 1.531$
$2s = 2(1.211) = 2.422$
Therefore, $\hat{y} \pm 2s \Rightarrow 1.531 \pm 2.422 \Rightarrow (-.891, 3.593)$

For Brand B, $\hat{y} = 9.31 - .1077x$. For $x = 70$, $\hat{y} = 9.31 - .1077(70) = 1.771$
$2s = 2(.61) = 1.22$
Therefore, $\hat{y} \pm 2s \Rightarrow 1.771 \pm 1.22 \Rightarrow (.551, 2.991)$

e. More confident with Brand B since there is less variation (s is smaller).

Problem W. Below are data for the United States over the 1980 to 2004 period. In graph 1, the US real interest rate, *USRINT*, (the nominal yield on 10 year constant maturity government bonds, subtracting off the lagged one year CPI inflation rate) and the net government debt as a share of gross domestic product, *USDEBT*, are plotted (in percent). In Figure 2, a scatterplot is presented, along with a bivariate regression line.





Dependent Variable: USRINT
Method: Least Squares
Date: 11/24/03 Time: 20:31
Sample(adjusted): 1985 2002
Included observations: 18 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.949735	2.336725	2.118236	0.0502
USDEBT	-0.021767	0.045470	-0.478715	0.6386

R-squared	0.014121	Mean dependent var	3.838888
Adjusted R-squared	-0.047497	S.D. dependent var	1.140295
S.E. of regression	1.167061	Akaike info criterion	3.251294
Sum squared resid	21.79251	Schwarz criterion	3.350224
Log likelihood	-27.26165	F-statistic	0.229168
Durbin-Watson stat	0.625861	Prob(F-statistic)	0.638618

a. In words, interpret the coefficient on *USDEBT*.

Generally, the coefficient is interpreted as the change in the left-hand side (dependent) variable for a given change in the right hand side (independent) variable. For the units that these variables are measured in, this coefficient is the percentage point change in US real interest rates for a one percentage point change in the US debt to GDP ratio. When the US debt to GDP ratio rises by one percentage point, then the US real interest rate falls by 0.022 percentage points.

b. Conduct a two-sided t-test using a 5% significance level.

The test in question is a two-sided t-test for the null hypothesis that the true value of the parameter is 0, i.e.:

H_0: $\beta_1=0$

H_a: $\beta_1 \neq 0$

Reject the null if $|t| > t_{\alpha/2}$. For 5% significance, $t_{\alpha/2} = t_{.025} = 2.12$ (for 16 d.f.).

Calculating t = (-0.021767-0)/ 0.045470 = -0.47871

Since the absolute value of this t-statistic is less 2.12, then one should fail to reject the null hypothesis.

c. Calculate the standard error of the regression, using the statistics reported in the output (show your work!).

The entry "Sum squared resid" is the same as what the textbook calls the Sum of Squared Errors (SSE).

$$s = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{21.79251}{16}} = 1.16706$$

which is the same as the entry under "S.E. of regression".

d. Calculate the value of the "S.E. of regression" using the "Sum of squared resid" (also termed the Sum of Squared Errors in the textbook).

Same as c – this question is just to check whether you understand the regression output and the concepts in the textbook.

e. Calculate the R-squared using the values for SSE and the "S.D. dependent var" (which is the standard deviation of the dependent variable).

$R^2 = 1 - (SSE/SS_{yy})$. While SS_{yy} is not reported in the regression output, notice that the standard deviation of the dependent variable y *is* reported. In fact $SS_{yy} = (SD^2)x(n-1) = (1.140295)^2x(17) = 22.10464$

Hence, $R^2 = 1 - (SSE/SS_{yy}) = 1 - (21.79251/22.10464) = 0.014120$ (which differs from the regression output entry for this measure slightly due to rounding errors).

Suppose you believe that not only does today's debt matter, but also additional debt the government is expected to incur over the next two years (*USDEBT_2-USDEBT_0*), as well as the state of the economy (*USGAP*). Then the following regression might be estimated:

Dependent Variable: USRINT
Method: Least Squares
Date: 11/24/03 Time: 20:45
Sample(adjusted): 1988 2002
Included observations: 15 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.205692	1.603793	0.128253	0.9003
USDEBT	0.065595	0.030105	2.178859	0.0520
USDEBT_2-USDEBT_0	0.151589	0.084782	1.787993	0.1013
USGAP	0.373826	0.167481	2.232059	0.0474

R-squared	0.491138	Mean dependent var	3.466149
Adjusted R-squared	0.352358	S.D. dependent var	0.748689
S.E. of regression	0.602517	Akaike info criterion	2.047776
Sum squared resid	3.993292	Schwarz criterion	2.236590
Log likelihood	-11.35832	F-statistic	3.538961
Durbin-Watson stat	2.175840	Prob(F-statistic)	0.051746

f. Interpret the coefficient on USDEBT.

This coefficient is the percentage point change in US real interest rates for a one percentage point change in the expected increase in the US debt to GDP ratio over the next two years, *holding all other variables constant*. When the expected change in the US debt to GDP ratio rises by one percentage point, *and all other variables remain constant*, then the US real interest rate *rises* by 0.152 percentage points.

g. Calculate the standard error of the regression, using the statistics reported in the output (again, show your work!).

$$s = \sqrt{\frac{SSE}{n - (k + 1)}} = \sqrt{\frac{3.993292}{11}} = 0.602517$$

which is the same as the entry under "S.E. of regression".

h. Form a 95% confidence interval around the coefficient on (*USDEBT_2-USDEBT_0*).

Note a 95% confidence interval implies α = 0.05, α /2 = 0.025; also there are 11 degrees of freedom.

The confidence interval is given by $\hat{\beta}_2 \pm t_{\alpha/2} \times s_{\hat{\beta}_2}$ = 0.151589 ± 2.201 x 0.084782

= (-0.035016, 0.338194)