BUEC 333, Answers to problem set exercises
Part 1: Probability and Statistics

1

a. $P(X=1)=½$, $P(X=2)=P(X=3)=\ \ldots\ =P(X=6)\ =\frac{1}{10}$
b. $P(A)=P(X=4)+P(X=5)+P(X=6)=\frac{3}{10}$
c. $P(B)=P(X=1)+P(X=3)+P(X=5)=\frac{1}{2}+\frac{1}{10}+\frac{1}{10}=\frac{7}{10}$
d. $\mu_x=\sum_i X_i * P(X_i)=\ 2.5$

$$Unfair\ die:\quad Var(X)=\sum_i (X_i-\mu_x)^2 * P(X_i)$$
$$=\tfrac{1}{2}*(1-2.5)^2+\tfrac{1}{10}*\{(2-2.5)^2+(3-2.5)^2+(4-2.5)^2+(5-2.5)^2+(6-2.5)^2\}=\tfrac{1}{2}*2.25+\tfrac{1}{10}*21.25=2.25$$
$$\mu_{x'}=\sum_i X'_i * P(X'_i)=\ 3.5$$
$$Fair\ die:\ Var(X')=\sum_i (X'_i-\mu_{x'})^2 * P(X'_i)$$
$$=\tfrac{1}{6}*\{(1-3.5)^2+(2-3.5)^2+(3-3.5)^2+(4-3.5)^2+(5-3.5)^2+(6-3.5)^2\}=2.91$$

2

	X1	X2
HHH	3	3
HHT	2	1
HTH	2	1
THH	2	1
HTT	1	-1
TTH	1	-1
THT	1	-1
TTT	0	-3

a.

X1	3	2	1	0
Prob	⅛	⅜	⅜	⅛

b.

X2	3	1	-1	-3
Prob	⅛	⅜	⅜	⅛

3

a. Given $P(X=0)=p$, we have $P(X=1)=1-p$
b. $E(X)=p(0)+(1-p)(1)=1-p$
c. $var(X)=E(X^2)-(E(X))^2=(p(0^2)-(1-p)(1^2))-(1-p)^2=p(1-p)$

4. The variance question.

- Table 3. E(X)=½*(-1)+½*2=0.5. So the variance is (-1-0.5)^2*0.5+(2-0.5)^2*0.5=2.25. Alternatively, call the RV in Table 1 “X”, and the one in Table 3 “Y”. Note that Y=1.5*X+0.5. Therefore, Var(Y)=1.5^2 Var(X)

- Table 4. E(X)=8/10. Then Var(X)=1/10*(-1-8/10)^2 + 9/10*(1-8/10)^2 = 1/10*(-18/10)*(-18/10)+9/10*2/10*2/10 = 360/1000=0.36

7
Let X1 be the first shot, X2 be the second shot.
X1 = 1 if has a bullet, 0 otherwise
X2 = 1 if has a bullet, 0 otherwise

Spin again:
Pr(X2=0) = ½
Pull the trigger without spinning:
Pr(X2=0|X1=0) = ⅔

Better not to spin.

8
a. Recall $P(X,Y) = P(X|Y)P(Y)$ and $P(X|Y) = P(X)$ if X and Y are independent. Thus $P(X = 10, Y = 20) = 0.3 \times 0.6 = 0.18$
b. $P(X = 10|Y = 20) = P(X = 10, Y = 20)/P(Y = 20) = 0.18/0.6 = 0.3$, which is just the same as $P(X = 10)$ unconditionally.
c. No.

9
a.

$$E(V) = E(20 - 7Y) = 20 - 7E(Y) = 20 - 7 * 0.78 = 14.54$$
$$E(W) = E(3 + 6X) = 3 + 6E(X) = 3 + 6 * 0.7 = 7.2$$

b.

$$\sigma_V^2 = Var(20 - 7Y) = (-7)^2\sigma_Y^2 = 8.4084\ \sigma_W^2 = Var(3 + 6X) = 6^2\sigma_X^2 = 7.56$$

c.

$$\sigma_{WV} = cov(3 + 6X,\ 20 - 7Y) = 6(-7)con(X,Y) =- 42 * 0.084 =- 3.52$$
$$corr(W,V) = \frac{\sigma_{WV}}{\sigma_V * \sigma_W} =- \frac{3.528}{\sqrt{7.56 * 8.4084}} =- 0.4425$$

10 (a) P(X=6|Y=1) = P(X=6,Y=1) / P(Y=1) = ¼ / ½ =½.
(b) P(Y=0|X=4) = P(X=4,Y=0) / P(X=4) = 1/12 / (1/12+1/20) = ⅝
(c) P(X=3) = P(X=3,Y=0) + P(X=3,Y=1) = 1/12 + 1/20 = 2/15.
(d) E(X|Y=0) = 3.5, because the conditional prob P(X=x|Y=0)=⅙
P(X=6|Y=1) = ½, P(X=x|Y=1)=1/10 for x=1,...,5, so

E(X|Y=0) = ½*6 + 1/10*(1+...+5) = 3 + 15/10 = 4.5

(e) 3.5*½ + 4.5*½ = 4

(f) 1/10*((1-4.5)^2 + (2-4.5)^2 + (3-4.5)^2 + (4-4.5)^2 + (5-4.5)^2)) + 0.5*(6-4.5)^2 = 3.25.

11

a. $E(M|A=0) = \sum_i M_i Pr(M = M_i|A=0)$

$= 0 * 0.35/0.5 + 1 * 0.065/0.5 + 2 * 0.05/0.5 + 3 * 0.025/0.5 + 4 * 0.01/0.5 = 0.56$

$E(M|A=1) = \sum_i M_i Pr(M = M_i|A=1)$

$= 0 * 0.45/0.5 + 1 * 0.035/0.5 + 2 * 0.01/0.5 + 3 * 0.005/0.5 + 4 * 0.00/0.5 = 0.14$

b. Old computers are more likely to crash

c.

$$Var(M|A=0) = \sum_i (M_i - E(M|A=0))^2 - Pr(M = M_i|A=0)$$

$$= (0-0.56)^2 * 0.7 + \ ... \ + (4-0.56)^2 * 0.02 = 0.99$$

12

Cov(X,Y) = E(XY)-E(X)E(Y) = 171.7 - 5.33 x 30.15 = 11

Corr(X,Y) = Cov(X, Y)/[SD(X)SD(Y)] = 11/(2.6 x 14.77) = 0.286

13

a. Recall that if $Y \sim N(\mu, \sigma^2)$, then $\frac{Y-\mu}{\sigma} \sim N(0,1)$. Thus, given $Y \sim N(1,4)$,
$\Pr(Y \le 3) = \Pr(\frac{Y-1}{2} \le \frac{3-1}{2}) = \Phi(1) = 0.8413$.

b. Given $Y \sim (3,9)$, $\Pr(Y > 0) = 1 - \Pr(Y \le 0) = 1 - \Pr(\frac{Y-3}{3} \le \frac{0-3}{3}) = 1 - \Phi(-1) = \Phi(1) = 0.8413$

c. Given $Y \sim (50,25)$,
$\Pr(40 \le Y \le 52) = \Pr(\frac{40-50}{5} \le \frac{Y-50}{5} \le \frac{52-50}{5}) = \Phi(0.4) - \Phi(-2) = \Phi(0.4) - [1 - \Phi(2)] = 0.6326$

d. Given $Y \sim (5,2)$, $\Pr(6 \le Y \le 8) = \Pr(\frac{6-5}{\sqrt{2}} \le \frac{Y-5}{\sqrt{2}} \le \frac{8-5}{\sqrt{2}}) = \Phi(2.1213) - \Phi(0.7071) = 0.2229$.

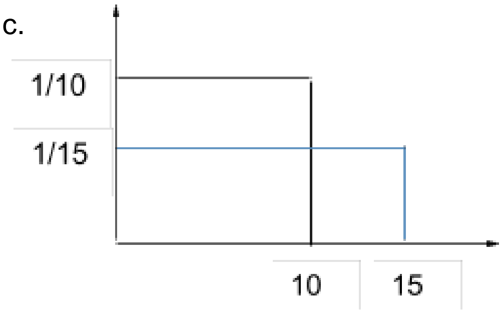
14 P(X^2 > 4) = P(X< -2) + P(X>2)
= P((X-3) / 2 < -5/2) + P((X-3) / 2 > -½)
= P(Z< -5/2) + P(Z> -½) where Z is a standard normal.

16

a. Because X_1 has a uniform probability distribution. X_1 is time and as a result is continuous.

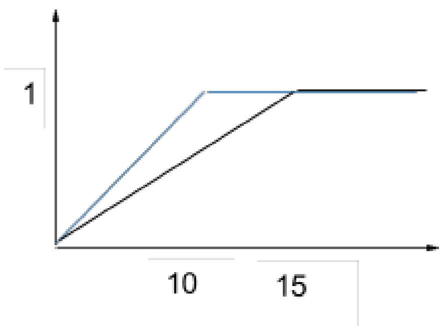
b. 0, 15

c.



e. $P(1 \le X_1 \le 3) = \frac{2}{15}$, $P(1 \le X_2 \le 3) = 2/10$

f.



Statistics

1 SW 2.17

$\mu_{Y=0.4}$ and $\sigma_Y^2 = 0.4 * 0.6 = 0.24$

a. i. $P(\bar{Y} \geq 0.43) = Pr(\frac{\bar{Y} - 0.4}{\sqrt{0.24/n}} \geq \frac{0.43 - 0.4}{\sqrt{0.24/n}}) = Pr(\frac{\bar{Y} - 0.4}{\sqrt{0.24/n}} \geq 0.6124) = 0.27$

ii. $P(\bar{Y} \leq 0.37) = Pr(\frac{\bar{Y} - 0.4}{\sqrt{0.24/n}} \leq \frac{0.37 - 0.4}{\sqrt{0.24/n}}) = Pr(\frac{\bar{Y} - 0.4}{\sqrt{0.24/n}} \leq -1.22) = 0.27$

b. Pr(-1.96≤Z≤1.96)=0.95

therefore we need to find *n* such that:

$0.41 = \frac{0.41 - 0.4}{\sqrt{0.24/n}} > -1.96$ *and* $\frac{0.39 - 0.4}{\sqrt{0.24/n}} < -1.96.$ solving this gives us: $n \geq 9220$

SW 3.1

The CLT suggests that when the sample size (n) is large, the distribution of the sample average ($\bar{Y}$) is approximately $N(\mu_Y, \sigma_{\bar{Y}}^2)$ with $\sigma_{\bar{Y}}^2 = \sigma_Y^2/n$.

Given population $\mu_Y = 100$, σ_Y^2 =43 , we have

a. n=100, $\sigma_{\bar{Y}}^2 = \sigma_Y^2/n = 43/100 = 0.43$ and

$P(\bar{Y} < 101) = Pr(\frac{\bar{Y} - 100}{\sqrt{0.43}} < \frac{101 - 100}{\sqrt{0.43}}) \approx \Phi(1.525) = 0.9364$

b. n=64, $\sigma_{\bar{Y}}^2 = \sigma_Y^2/n = 43/64 = 0.6719,$ and

$P(101 < \bar{Y} < 103) = Pr(\frac{101 - 100}{\sqrt{0.6719}} < \frac{\bar{Y} - 100}{\sqrt{0.6719}} < \frac{103 - 100}{\sqrt{0.6719}}) \approx \Phi(3.6599) - \Phi(1.2200) = 0.9999 - 0.8888 = 0.1111$

c. n=165, $\sigma_{\bar{Y}}^2 = \sigma_Y^2/n = 43/165 = 0.2606$ and

$P(\bar{Y} > 98) = 1 - P(\bar{Y} \leq 98) = 1 - Pr(\frac{\bar{Y} - 100}{\sqrt{0.2606}} < \frac{98 - 100}{\sqrt{0.2606}}) \approx 1 - \Phi(-3.9178) = \Phi(3.9178) \approx 1$

2

3

<u>S&W 3.8</u>

Given that n = 1000, $\bar{X}$ = 1110, s_X= 123, a 95% confidence interval for the population mean is $\left\{ \bar{X} \pm 1.96\left(\frac{s_X}{\sqrt{n}}\right)\right\} = \left\{1110 \pm 1.96\left(\frac{123}{\sqrt{1000}}\right)\right\} = (1102.38,\ 1117.62)$

<u>S&W 3.13 (a)</u>

Given that n=420, $\bar{Y} = 646.2,\ s_Y = 19.5$, a 95% confidence interval for the population mean is $\left\{ \bar{Y} \pm 1.96\left(\frac{s_Y}{\sqrt{n}}\right)\right\} = \left\{646.2 \pm 1.96\left(\frac{19.5}{\sqrt{420}}\right)\right\} = (644.34,\ 648.06)$

4 SW 3.11

Assume that n is an even number. Then $\widehat{Y}$ is constructed by applying a weight of ½ to the n/2 "odd" observations and a weight of 3/2 to the remaining n/2 observations.

$E(\widehat{Y}) = 1/n\ [1/2\ E(Y_1) + 3/2E(Y_2) + ... + 1/2\ E(Y_{n-1}) + 3/2E(Y_n)]$

$= 1/n\ [1/2\ * n/2 * \mu_Y + 3/2\ * n/2 * \mu_Y] = \mu_Y$

$Var(\widehat{Y}) = 1/n^2\ [1/4\ Var(Y_1) + 9/4\ Var(Y_2) + ... + 1/4\ VarE(Y_{n-1}) + 9/4\ Var(Y_n)]$

$= 1/n^2[1/4 \, * n/2 * \sigma_Y{}^2 + 9/4 \, * n/2 * \sigma_Y{}^2] = \sigma_Y{}^2 = 1.25\, \sigma_Y{}^2/n$

5.

1. E(1/2(Y1+Y2))=1/2 E(Y1+Y2) = 1/2 * (E(Y1) + E(Y2)) = 1/2 * 2 mu = mu
2. 1/8 var(Y), which is smaller than var(Y-bar) , so you prefer Y-tilde
3. It is biased
4. The estimator is based on the unknown quantity that you are trying to estimate!

6 Stock and Watson 3.2.

a. Let s be the number of successes in the trial. Then the fraction of success in n trials is

$$\hat{p} = \tfrac{s}{n} = \tfrac{1}{n}\sum_i Y_i = \overline{Y}$$

b. $E(\hat{p}) = E(\tfrac{1}{n}\sum_i Y_i) = \tfrac{1}{n}\sum_i E(Y_i) = \tfrac{1}{n}\sum_i p = p$

c. $var(\hat{p}) = var(\tfrac{1}{n}\sum_i Y_i) = \tfrac{1}{n^2}\sum_i var(Y_i) = \tfrac{1}{n^2}\sum_i p(1-p) = \tfrac{1}{n}p(1-p)$

7

8

9 Stock and Watson 2.7. Let M and F denote the randomly selected male and female earnings, respectively. Then $\mu_C = \mu_M + \mu_F$ and $\sigma_C^2 = \sigma_M^2 + \sigma_F^2 + 2cov(M,F)$.

a. $\mu_C = 40000 + 45000 = 85000$
b. From the correlation formula we get $cov(M,F) = \sigma_M\sigma_F corr(M,F) = 12 \times 18 \times 0.8 = 172.8$ where the units are squared thousands of dollars per year.
c. $\sigma_C = \sqrt{\sigma_M^2 + \sigma_F^2 + 2cov(M,F)} = \sqrt{12^2 + 18^2 + 2 \times 172.8} = \sqrt{813.6} = 28.524$
d. Let the exchange rate be E Euros per dollar. Then the mean combined income would be $E \times \mu_C$ Euros per year, with standard deviation $E \times \sigma_C \times 1000$ Euros per year. The correlation is unit free, and thus unchanged after conversion.

10

1. 3
2. (2-3)^2+(3-3)^2+(4-3)^2 / (n-1) = 1.
3. [3-1.64*SE(Xbar), 3+1.64*SE(Xbar)] = [3-1.64/sqrt(3), 3+1.64/sqrt(3)]
4. Incorrect answer: “contains μ with 90% probability”. Correct: “If we would repeat this procedure many, many times, it would contain μ ~90% of the time.”.

11 No. Note X-tilde = (n-2) /n X-bar, so that E[X-tilde] = (n-2)/n * mu

12

1 Mechanics and Fit

1. Let GPA be Y and ACT be X.

a) $\bar{Y} = 3.21$, $\bar{X} = 25.88$, $s_{XY} = 5.8125$, $s_X^2 = 56.875$. Recall that

$$\hat{\beta}_1 = \frac{s_{XY}}{s_X^2} = \frac{\sum_i (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_i (X_i - \bar{X})^2} \qquad \text{and} \qquad \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} \tag{1}$$

Thus $\hat{\beta}_1 = 0.1022$ and $\hat{\beta}_0 = 0.5681$

b) No

c) A one unit increase in ACT is, on average, associated with a $\hat{\beta}_1$ unit increase in GPA

d) Observation specific values are in the table below

Student	$\hat{Y}_i$	$\hat{u}_i$
1	2.7143	0.0857
2	3.0209	0.3791
3	3.2253	-0.2252
4	3.3275	0.1725
5	3.5319	0.0681
6	3.1231	-0.1231
7	3.1231	-0.4231
8	3.6341	0.0659
		$\sum_i \hat{u}_i = -0.0001$

e) $\hat{Y}(X = 20) = 0.5681 + 0.1022(20) = 2.6121$

f) $ESS = 0.5940$, $TSS = 1.0288$. Recall that

$$R^2 = \frac{ESS}{TSS} = \frac{\sum_i (\hat{Y}_i - \bar{Y})^2}{\sum_i (Y_i - \bar{Y})^2}$$

Thus $R^2 = 0.5774$.

2. $\bar{X} = \bar{Y} = 0$

a) If $k = 2$ for $i = 1, \ldots, 6$, $Y_i = X_i$ (i.e. $\hat{\beta}_0 = 0$, $\hat{\beta}_1 = 1$), $\hat{u}_i = Y_i - \hat{Y}_i = Y_i - X_i = 0$, $SSR = \sum_i \hat{u}_i = 0$, $R^2 = 1 - \frac{SSR}{TSS} = 1 - 0 = 1$

b) Given the formulae in equation (1) above,

$$\hat{\beta}_1 = \frac{5+k}{7}, \qquad \text{and} \qquad \hat{\beta}_0 = 0$$

c) If $k = 1$, $\hat{\beta}_0 = 0, \hat{\beta}_1 = \frac{6}{7}, \hat{Y}_i = \frac{6}{7}X_i$. Thus $R^2 = \frac{ESS}{TSS} = \frac{144/7}{22} = \frac{72}{77}$

d) If $k = 3$, $\hat{\beta}_0 = 0, \hat{\beta}_1 = \frac{8}{7}, \hat{Y}_i = \frac{8}{7}X_i$. $R^2 = \frac{256/7}{38} = \frac{128}{133}$

e) If $k = 10$, $\hat{\beta}_0 = 0, \hat{\beta}_1 = \frac{15}{7}, \hat{Y}_i = \frac{15}{7}X_i$. $R^2 = \frac{900/7}{220} = \frac{45}{77}$

f) Given that $\hat{\beta}_0 = 0$ and $\hat{\beta}_1 = \frac{5+k}{7}$, $\hat{\beta}_1 \to \infty$ as $k \to \infty$. That is, the fitted line converges to a vertical line going through the origin. And since $\hat{Y}_i = \frac{5+k}{7}X_i$, $ESS = \frac{1}{7}(100 + 4k^2 + 40k)$, $TSS = 20 + 2k^2$, and

$$R^2 = \frac{50 + 20k + 2k^2}{70 + 7k^2}$$

$R^2 \to \frac{2}{7}$ as $k \to \infty$

3. Stock & Watson 4.1

a) The predicted average test score is

$$\widehat{TestScore} = 520.4 - 5.82 \times 22 = 392.36$$

b) The predicted change in the classroom average test score is

$$\Delta\widehat{TestScore} = (-5.82 \times 19) - (-5.82 \times 23) = 23.28$$

c) Using the formula for $\hat{\beta}_0$ in equation (4.8), we know the sample average of the test score across the 100 classroom is

$$\overline{TestScore} = \hat{\beta}_0 + \hat{\beta}_1\overline{CS} = 520.4 - 5.82 \times 21.4 = 395.85$$

2 Estimation and Inference

1. Stock & Watson 4.4

2. Stock & Watson 4.6

$$\begin{aligned} E(Y_i \mid X_i) &= E(\beta_0 + \beta_1 X_i + u_i \mid X_i) \\ &= E(\beta_0 \mid X_i) + E(\beta_1 X_i \mid X_i) + E(u_i \mid X_i) \\ &= \beta_0 + \beta_1 X_i \end{aligned}$$

3. Stock & Watson 4.7

$$\begin{aligned} E(\hat{\beta}_0) &= E(\bar{Y} - \hat{\beta}_1 \bar{X}) \\ &= E\left[\left(\beta_0 + \beta_1 \bar{X} + \frac{1}{n}\sum_i u_i\right) - \hat{\beta}_1 \bar{X}\right] \\ &= \beta_0 + \underbrace{E(\beta_1 - \hat{\beta}_1)}_{\substack{=0 \text{ since } \hat{\beta}_1 \\ \text{is unbiased}}} \bar{X} + \frac{1}{n}\sum_i \underbrace{E(u_i)}_{\substack{=0 \\ \text{by assumption}}} \\ &= \beta_0 \end{aligned}$$

4. Stock & Watson 5.2 (a)-(d)

a)

$$\begin{aligned} \beta_1 &= E(Wage_i \mid Male_i = 1) - E(Wage_i \mid Male_i = 0) \\ &= \text{difference in mean earnings between men and women} \\ &= \text{wage gender gap} \end{aligned}$$

The estimated gender gap $= \hat{\beta}_1 = \$2.12$ /hour

b) $H_0 : \beta_1 = 0$, $H_1 : \beta \neq 0$. $t = \frac{2.12-0}{0.36} = 5.89$, p-value $= \Pr(Z \leq -5.89) \times 2 \approx 0 < 0.01$. Thus reject H_0 at $\alpha = 1\%$. That is, the estimated gender gap is significantly different from 0.

c) $\{2.12 \pm 1.96 \times 0.36\} = (1.4144, 2.8256)$

d) $\overline{Wage} = \hat{\beta}_0 + \hat{\beta}_1 \overline{Male}$.
For women:

$$Male_i = 0 \Rightarrow \overline{Male} = 0 \Rightarrow \overline{Wage} = \hat{\beta}_0 + \hat{\beta}_1 \times 0 = \$12.52/\text{hour}$$

For men:

$$Male_i = 1 \Rightarrow \overline{Male} = 1 \Rightarrow \overline{Wage} = \hat{\beta}_0 + \hat{\beta}_1 = \$14.64/\text{hour}$$

5. Stock & Watson 5.3

The 99% confidence interval is $1.5 \times \{3.94 \pm 2.58 \times 0.31\}$ or

$$4.71\text{lbs} \leq WeightGain \leq 7.11\text{lbs}.$$

6.

a. 130

b. 124

c. The R^2 never decreases. This question relies on Chapter 6 knowledge.

d. No, the estimator is a random variable.

7.

a-c. See book.

d. No. Yes. Yes. Last one: yes, because it depends on $\hat{\beta}_1$.

3 Prediction and Forecasting

1. Stock & Watson 14.1

a) Y_t is stationary so its probability distribution does not change over time. As a results expected value of Y_t and Y_{t-1} are the same.

b)

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + u_t \Rightarrow E(Y_t) = \beta_0 + \beta_1 E(Y_{t-1})$$

From (a) we know $E(Y_t) = E(Y_{t-1})$

$$\Rightarrow E(Y_t) = \beta_0 + \beta_1 E(Y_t)$$

$$\Rightarrow (1 - \beta_1) E(Y_t) = \beta_0$$

$$\Rightarrow E(Y_t) = \beta_0/(1-\beta_1)$$

2.

a) Using the result from 1(b), $E(Y_t) = \frac{\beta_0}{1-\beta_1} = \frac{2.5}{1-0.7} = 8.33.$

$Var(Y_t) = Var(0.7Y_{t-1} + u_t) = 0.7^2 Var(Y_{t-1}) + Var(u_t) = 0.49Var(Y_t) + Var(u_t)$
(Note that $Var(Y_t) = Var(Y_{t-1})$ because Y_t is stationary.)

Thus $(1-0.49)Var(Y_t) = 9,$ and therefore $Var(Y_t) = 9/(1-0.49) = 17.647.$

b) $Cov(Y_t, Y_{t-1}) = Cov(2.5 + 0.7Y_{t-1} + u_t, Y_{t-1}) = 0.7Var(Y_{t-1}) + Cov(u_t, Y_{t-1})$
$= 0.7Var(Y_t) = 0.7 \times 17.647 = 12.3529$

$Cov(Y_t, Y_{t-2}) = Cov(2.5 + 0.7(2.5 + 0.7Y_{t-2} + u_{t-1}) + u_t, Y_{t-2})$
$= Cov(0.49Y_{t-2} + 0.7u_{t-1} + u_t, Y_{t-2})$
$= 0.49Cov(Y_{t-2}, Y_{t-2}) + 0.7Cov(u_{t-1}, Y_{t-2}) + Cov(u_t, Y_{t-2})$
$= 0.49Var(Y_{t-2}) = 0.49Var(Y_t) = 0.49 \times 17.647 = 8.64703$

c) (not assigned)

d) $Y_{T+1|T} = E(Y_{T+1}|Y_T, Y_{T-1}, ...) = E(2.5 + 0.7Y_T + u_t|Y_T, Y_{T-1,...}) = 2.5 + 0.7Y_T$
$= 2.5 + 0.7 \times 102.3 = 74.11$

3. Let $(Y_1, Y_2, Y_3) = (-1, 1, -1)$.

a) Recall that

$$\widehat{\text{cov}(Y_t, Y_{t-j})} = \frac{1}{T}\sum_{t=j+1}^{T}(Y_t - \bar{Y}_{j+1:T})(Y_{t-j} - \bar{Y}_{1:T-j})$$

where $T = 3$ and $j = 1$ for this question (and note that both $\bar{Y}$s equal zero). Applying this formula we get

$$\widehat{\text{cov}(Y_t, Y_{t-1})} = \frac{1}{3}\Big[(Y_2 - 0)(Y_1 - 0) + (Y_3 - 0)(Y_2 - 0)\Big] = -\frac{2}{3}.$$

b) Recall that

$$\hat{\rho}_j = \frac{\widehat{\text{cov}(Y_t, Y_{t-j})}}{\widehat{\text{var}(Y_t)}}$$

where $j = 1$ for this question. Since $\bar{Y} = \frac{1}{3}(-1 + 1 - 1) = -\frac{1}{3}$,

$$\widehat{\text{var}(Y_t)} = \frac{1}{2}\left[\left(-1 + \frac{1}{3}\right)^2 + \left(1 + \frac{1}{3}\right)^2 + \left(-1 + \frac{1}{3}\right)^2\right] = \frac{4}{3}$$

Thus,

$$\hat{\rho}_1 = \frac{-2/3}{4/3} = -\frac{1}{2}.$$

c) Yes, consider $(0, 0, 0, 0, 0)$.

4. See "forecasts and predicted values" on p.533 of Stock & Watson.

1 Introduction to Multiple Linear Regression

1. Stock & Watson 6.1–6.4

6.1 Recall that

$$\bar{R}^2 = 1 - \frac{n-1}{n-k-1}(1-R^2).$$

Thus, the values of $\bar{R}^2$ are 0.162, 0.180, 0.181 for columns (1)–(3).

6.2 (a) Workers with college degrees earn \$8.31/hour more, on average, than workers with only high school degrees.

(b) Men earn \$3.85/hour more, on average, than women.

6.3 (a) On average, a worker earns \$0.51/hour more for each year he ages.

(b) Sally's earnings prediction is $1.87 + 8.32 \times 1 - 3.81 \times 1 + 0.51 \times 29 = \$21.17/\text{hour}$. Betsy's earnings prediction is $1.87 + 8.32 \times 1 - 3.81 \times 1 + 0.51 \times 34 = \$23.72/\text{hour}$.

6.4 (a) Workers in the Northeast earn \$0.18 more per hour than workers in the West, on average, controlling for other variables in the regression. Workers in the Midwest earn \$1.23 less per hour than workers in the West, on average, controlling for other variables in the regression. Workers in the South earn \$0.43 less than workers in the West, controlling for other variables in the regression.

(b) The regressor *West* is omitted to avoid perfect multicollinearity. If *West* is included, then the intercept can be written as a perfect linear function of the four regional regressors.

(c) The expected difference in earnings between Juanita and Jennifer is

$$-0.43 - (-1.23) = \$0.80/\text{hour}$$

2. Stock & Watson 6.5

(a) The expected increase in the value of the house is 23,400 dollars (note that *Price* is in 1000 dollars).

(b) The expected increase in the value of the house is $23,400 + 15,600 = 39,000$ dollars.

(c) The loss in value is 48,800 dollars.

(d) $R^2 = 1 - \frac{n-k-1}{n-1}\left(1 - \bar{R}^2\right) = 0.728$

3. Stock & Watson 6.6

   (a) There are other important determinants of a country's crime rate, including demographic characteristics of the population.

   (b) Suppose that the crime rate is positively affected by the fraction of young males i nthe population, and that counties with high crime rates tend to hire more police. In this case, the size of the police force is likely to be positively correlated with the fraction of young males in the population leading to a positive value for the omitted variable bias, so that $\hat{\beta}_1 > \beta_1$.

6. Stock & Watson 6.9

   - For omitted variable bias to occur, two conditions must be true: X_1 (the included regressor) is correlated with the omitted variable, and the omitted variable is a determinant of the dependent variable. Since X_1 and X_2 are uncorrelated, the estimator of β_1 does not suffer from omitted variable bias.

7. Final, Summer 2014

   (a) Dummy variable trap / perfect multicollinearity

   (b) Exclude either F_i or M_i, or drop the constant β_0 from the regression equation.

8. Final Q5, Summer 2014

   (a) $100 + 6 \times 5 = 130$

   (b) 124

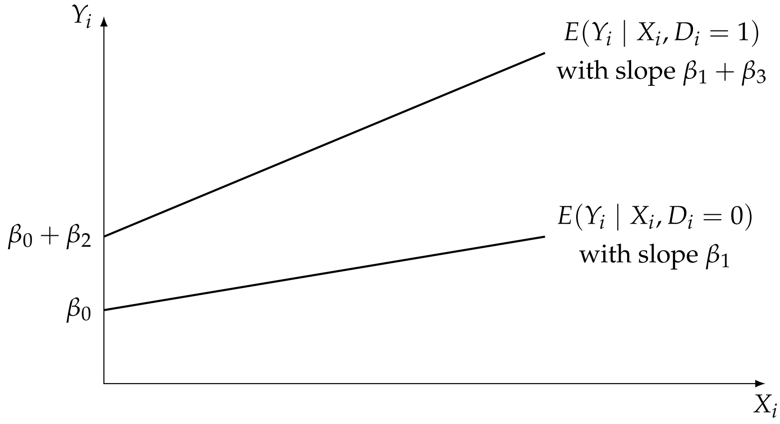
   (c) In multiple regression, the R^2 increases whenever a regressor is added, unless the estimated coefficient on the added regressor is exactly zero. To see this, think about starting with one regressor and then adding a second. When you use OLS to estimate the model with both regressors, OLS finds the values of the coefficients that minimize the sum of squared residuals. If OLS happens to choose the coefficient on the new regressor to be exactly zero, then the SSR will be the same whether or not the second variable is included in the regression. But if OLS chooses any value other than zero, then it must be that this value reduced the SSR relative to the regression that excludes this regressor. In practice, it is extremely unusual for an estimated coefficient to be exactly zero, so in general the SSR will decrease when a new regressor is added. But this means that the R^2 generally increases (and never decreases) when a new regressor is added.

   (d) Sampling variability. Even if $\beta_{ID} = 0$, $\hat{\beta}_{ID} \sim \mathcal{N}(0, \sigma^2_{\hat{\beta}})$

2 Topics in Linear Regression

5. Final Q6, Summer 2014

(a) Assuming $\beta_0, \beta_1, \beta_2, \beta_3$ are all positive, then



(b) Difference in slopes $\frac{\partial E(Y_i|X_i,D_i)}{\partial X_i}$ between $E(Y_i \mid X_i, D_i = 1)$ and $E(Y_i \mid X_i, D_i = 0)$.

7. Stock & Watson 7.2

(a) The t-statistic is $8.31/0.23 = 36.1 > 1.96$, so the coefficient is statistically significant at the 5% level. The 95% confidence interval is $8.31 \pm (1.96 \times 0.23)$.

(b) The t-statistic is $-3.85/0.23 = -16.7 > 1.96$, so the coefficient is statistically significant at the 5% level. The 95% confidence interval is $-3.85 \pm (1.96 \times 0.23)$.

9. β_1 measures the expected unit change in Y_i given a one-percent change in X_i.

10. Stock & Watson 8.7

(a) (i) $ln(Earnings)$ for females are, on average, 0.44 lower for men than for women.
(ii) The error term has a standard deviation of 2.65 (measured in log-points).
(iii) Yes. But the regression does not control for many factors (size of firm, industry, profitability, experience and so forth).
(iv) No. In isolation, these results do not imply gender discrimination. Gender discrimination means that two workers, identical in every way but gender, are paid different wages. Thus, it is also important to control for characteristics of the workers that may affect their productivity (education, years of experience, etc.) If these characteristics are systematically different between men and women, then they may

be responsible for the difference in mean wages. (If this were true, it would raise an interesting and important question of why women tend to have less education or less experience than men, but that is a question about something other than gender discrimination.) These are potentially important omitted variables in the regression that will lead to bias in the OLS coefficient estimator for Female.
Since these characteristics were not controlled for in the statistical analysis, it is premature to reach a conclusion about gender discrimination.

(b) (i) If MarketValue increases by 1%, earnings increase by 0.37%.
(ii) Female is correlated with the two new included variables and at least one of the variables is important for explaining $ln(Earnings)$. Thus the regression in part (a) suffered from omitted variable bias.

(c) Forgetting about the effect or Return, whose effects seems small and statistically insignificant, the omitted variable bias formula (see equation (6.1)) suggests that Female is negatively correlated with $ln(MarketValue)$.

3 Intrumental Variables

1. Stock & Watson 12.5

(a) Instrument relevance. Z_i does not enter the population regression for X_i.

(b) $\hat{X}_i$ will be perfectly collinear with W_i. Alternatively, the first stage regression suffers from perfect multicollinearity.

(c) W_i is perfectly multicollinear with the constant term.

(d) Z_i violates instrument exogeneity, because it's correlated with the error term.

2. Stock & Watson 12.9

(a) There are other factors that could affect both the choice to serve in the military and annual earnings. One example could be education, although this could be included in the regression as a control variable. Another variable is "ability" which is difficult to measure, and thus difficult to control for in the regression.

(b) The draft was determined by a national lottery so the choice of serving in the military was random. Because it was randomly selected, the lottery number is uncorrelated with individual characteristics that may affect earning and hence the instrument is exogenous. Because it affected the probability of serving in the military, the lottery number is relevant.

3. Stock & Watson 12.10

$$\begin{aligned}
\hat{\beta}_1^{TSLS} &= \frac{s_{ZY}}{s_{ZX}} \\
&\xrightarrow{p} \frac{cov(Z_i, Y_i)}{cov(Z_i, X_i)} \\
&= \frac{cov(Z_i, \beta_0 + \beta_1 X_i + \beta_2 W_i + u_i)}{cov(Z_i, X_i)} \\
&= \frac{\beta_1 cov(Z_i, X_i) + \beta_2 cov(Z_i, W_i)}{cov(Z_i, X_i)}
\end{aligned}$$

(a) If $cov(Z_i, W_i) = 0$, $\hat{\beta}_1^{TSLS} \xrightarrow{p} \beta_1$, the IV estimator is consistent.

(b) If $cov(Z_i, W_i) \neq 0$, $\hat{\beta}_1^{TSLS} \xrightarrow{p} \frac{\beta_1 cov(Z_i,X_i)+\beta_2 cov(Z_i,W_i)}{cov(Z_i,X_i)}$, the IV estimator is not consistent.