

# **Introduction to Econometrics (4th Edition)**

by

James H. Stock and Mark W. Watson

## **Solutions to End-of-Chapter Exercises: Chapter 2\***

(This version September 14, 2018)

\*Limited distribution: **For Instructors Only**. Answers to all odd-numbered questions are provided to students on the textbook website. If you find errors in the solutions, please pass them along to us at [mwatson@princeton.edu](mailto:mwatson@princeton.edu).

2.1. (a) Probability distribution function for  $Y$

Outcome (number of heads)	$Y = 0$	$Y = 1$	$Y = 2$
Probability	0.25	0.50	0.25

(b) Cumulative probability distribution function for  $Y$

Outcome (number of heads)	$Y < 0$	$0 \leq Y < 1$	$1 \leq Y < 2$	$Y \geq 2$
Probability	0	0.25	0.75	1.0

(c)  $\mu_Y = E(Y) = (0 \times 0.25) + (1 \times 0.50) + (2 \times 0.25) = 1.00$

Using Key Concept 2.3:  $\text{var}(Y) = E(Y^2) - [E(Y)]^2$ ,

and

$$E(Y^2) = (0^2 \times 0.25) + (1^2 \times 0.50) + (2^2 \times 0.25) = 1.50$$

so that

$$\text{var}(Y) = E(Y^2) - [E(Y)]^2 = 1.50 - (1.00)^2 = 0.50.$$

2.2. We know from Table 2.2 that  $\Pr(Y = 0) = 0.22$ ,  $\Pr(Y = 1) = 0.78$ ,  $\Pr(X = 0) = 0.30$ ,  $\Pr(X = 1) = 0.70$ . So

(a)

$$\begin{aligned}\mu_Y &= E(Y) = 0 \times \Pr(Y = 0) + 1 \times \Pr(Y = 1) \\ &= 0 \times 0.22 + 1 \times 0.78 = 0.78,\end{aligned}$$

$$\begin{aligned}\mu_X &= E(X) = 0 \times \Pr(X = 0) + 1 \times \Pr(X = 1) \\ &= 0 \times 0.30 + 1 \times 0.70 = 0.70.\end{aligned}$$

(b)

$$\begin{aligned}\sigma_X^2 &= E[(X - \mu_X)^2] \\ &= (0 - 0.70)^2 \times \Pr(X = 0) + (1 - 0.70)^2 \times \Pr(X = 1) \\ &= (-0.70)^2 \times 0.30 + 0.30^2 \times 0.70 = 0.21, \\ \sigma_Y^2 &= E[(Y - \mu_Y)^2] \\ &= (0 - 0.78)^2 \times \Pr(Y = 0) + (1 - 0.78)^2 \times \Pr(Y = 1) \\ &= (-0.78)^2 \times 0.22 + 0.22^2 \times 0.78 = 0.1716.\end{aligned}$$

(c)

$$\begin{aligned}\sigma_{XY} &= \text{cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] \\ &= (0 - 0.70)(0 - 0.78)\Pr(X = 0, Y = 0) + (0 - 0.70)(1 - 0.78)\Pr(X = 0, Y = 1) \\ &\quad + (1 - 0.70)(0 - 0.78)\Pr(X = 1, Y = 0) + (1 - 0.70)(1 - 0.78)\Pr(X = 1, Y = 1) \\ &= (-0.70) \times (-0.78) \times 0.15 + (-0.70) \times 0.22 \times 0.15 + 0.30 \times (-0.78) \times 0.07 + 0.30 \times 0.22 \times 0.63 \\ &= 0.084,\end{aligned}$$

$$\text{corr}(X, Y) = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{0.084}{\sqrt{0.21 \times 0.1716}} = 0.4425.$$

2.3. For the two new random variables  $W = 3 + 6X$  and  $V = 20 - 7Y$ , we have:

(a)

$$\begin{aligned}E(V) &= E(20 - 7Y) = 20 - 7E(Y) = 20 - 7 \times 0.78 = 14.54, \\E(W) &= E(3 + 6X) = 3 + 6E(X) = 3 + 6 \times 0.70 = 7.2.\end{aligned}$$

(b)

$$\begin{aligned}\sigma_W^2 &= \text{var}(3 + 6X) = 6^2 \sigma_X^2 = 36 \times 0.21 = 7.56, \\ \sigma_V^2 &= \text{var}(20 - 7Y) = (-7)^2 \times \sigma_Y^2 = 49 \times 0.1716 = 8.4084.\end{aligned}$$

(c)

$$\sigma_{WV} = \text{cov}(3 + 6X, 20 - 7Y) = 6(-7)\text{cov}(X, Y) = -42 \times 0.084 = -3.52$$

$$\text{corr}(W, V) = \frac{\sigma_{WV}}{\sigma_W \sigma_V} = \frac{-3.528}{\sqrt{7.56 \times 8.4084}} = -0.4425.$$

2.4. (a)  $E(X^3) = 0^3 \times (1-p) + 1^3 \times p = p$

(b)  $E(X^k) = 0^k \times (1-p) + 1^k \times p = p$

(c)  $E(X) = 0.3$

$$\text{var}(X) = E(X^2) - [E(X)]^2 = 0.3 - 0.09 = 0.21$$

Thus,  $\sigma = \sqrt{0.21} = 0.46$ .

To compute the skewness, use the formula from exercise 2.21:

$$\begin{aligned} E(X - \mu)^3 &= E(X^3) - 3[E(X^2)][E(X)] + 2[E(X)]^3 \\ &= 0.3 - 3 \times 0.3^2 + 2 \times 0.3^3 = 0.084 \end{aligned}$$

Alternatively,  $E(X - \mu)^3 = [(1-0.3)^3 \times 0.3] + [(0-0.3)^3 \times 0.7] = 0.084$

Thus, skewness  $= E(X - \mu)^3 / \sigma^3 = .084 / 0.46^3 = 0.87$ .

To compute the kurtosis, use the formula from exercise 2.21:

$$\begin{aligned} E(X - \mu)^4 &= E(X^4) - 4[E(X)][E(X^3)] + 6[E(X)]^2[E(X^2)] - 3[E(X)]^4 \\ &= 0.3 - 4 \times 0.3^2 + 6 \times 0.3^3 - 3 \times 0.3^4 = 0.0777 \end{aligned}$$

Alternatively,  $E(X - \mu)^4 = [(1-0.3)^4 \times 0.3] + [(0-0.3)^4 \times 0.7] = 0.0777$

Thus, kurtosis is  $E(X - \mu)^4 / \sigma^4 = .0777 / 0.46^4 = 1.76$

- 2.5. Let  $X$  denote temperature in °F and  $Y$  denote temperature in °C. Recall that  $Y = 0$  when  $X = 32$  and  $Y = 100$  when  $X = 212$ .

This implies  $Y = (100/180) \times (X - 32)$  or  $Y = -17.78 + (5/9) \times X$ .

Using Key Concept 2.3,  $\mu_X = 70^\circ\text{F}$  implies that  $\mu_Y = -17.78 + (5/9) \times 70 = 21.11^\circ\text{C}$ , and  $\sigma_X = 7^\circ\text{F}$  implies  $\sigma_Y = (5/9) \times 7 = 3.89^\circ\text{C}$ .

2.6. The table shows that  $\Pr(X=0, Y=0) = 0.026$ ,  $\Pr(X=0, Y=1) = 0.576$ ,  
 $\Pr(X=1, Y=0) = 0.009$ ,  $\Pr(X=1, Y=1) = 0.389$ ,  $\Pr(X=0) = 0.602$ ,  
 $\Pr(X=1) = 0.398$ ,  $\Pr(Y=0) = 0.035$ ,  $\Pr(Y=1) = 0.965$ .

(a)

$$\begin{aligned} E(Y) &= \mu_Y = 0 \times \Pr(Y=0) + 1 \times \Pr(Y=1) \\ &= 0 \times 0.035 + 1 \times 0.965 = 0.965. \end{aligned}$$

(b)

$$\begin{aligned} \text{Unemployment Rate} &= \frac{\#(\text{unemployed})}{\#(\text{labor force})} \\ &= \Pr(Y=0) = 1 - \Pr(Y=1) = 1 - E(Y) = 1 - 0.965 = 0.035. \end{aligned}$$

(c) Calculate the conditional probabilities first:

$$\Pr(Y=0|X=0) = \frac{\Pr(X=0, Y=0)}{\Pr(X=0)} = \frac{0.026}{0.602} = 0.043,$$

$$\Pr(Y=1|X=0) = \frac{\Pr(X=0, Y=1)}{\Pr(X=0)} = \frac{0.576}{0.602} = 0.957,$$

$$\Pr(Y=0|X=1) = \frac{\Pr(X=1, Y=0)}{\Pr(X=1)} = \frac{0.009}{0.398} = 0.023,$$

$$\Pr(Y=1|X=1) = \frac{\Pr(X=1, Y=1)}{\Pr(X=1)} = \frac{0.389}{0.398} = 0.978.$$

The conditional expectations are

$$\begin{aligned} E(Y|X=1) &= 0 \times \Pr(Y=0|X=1) + 1 \times \Pr(Y=1|X=1) \\ &= 0 \times 0.023 + 1 \times 0.978 = 0.978, \end{aligned}$$

$$\begin{aligned} E(Y|X=0) &= 0 \times \Pr(Y=0|X=0) + 1 \times \Pr(Y=1|X=0) \\ &= 0 \times 0.043 + 1 \times 0.957 = 0.957. \end{aligned}$$

(d) Use the solution to part (b),

$$\text{Unemployment rate for college graduates} = 1 - E(Y|X=1) = 1 - 0.978 = 0.023.$$

$$\text{Unemployment rate for non-college graduates} = 1 - E(Y|X=0) = 1 - 0.957 = 0.043$$

(e) The probability that a randomly selected worker who is reported being unemployed is a college graduate is

$$\Pr(X = 1|Y = 0) = \frac{\Pr(X = 1, Y = 0)}{\Pr(Y = 0)} = \frac{0.009}{0.035} = 0.257.$$

The probability that this worker is a non-college graduate is

$$\Pr(X = 0|Y = 0) = 1 - \Pr(X = 1|Y = 0) = 1 - 0.257 = 0.743.$$

(f) Educational achievement and employment status are not independent because they do not satisfy that, for all values of  $x$  and  $y$ ,

$$\Pr(X = x|Y = y) = \Pr(X = x).$$

For example, from part (e)  $\Pr(X = 0|Y = 0) = 0.743$ , while from the table  $\Pr(X = 0) = 0.602$ .



2.7. Using obvious notation,  $C = M + F$ ; thus  $\mu_C = \mu_M + \mu_F$  and  $\sigma_C^2 = \sigma_M^2 + \sigma_F^2 + 2\text{cov}(M, F)$ . This implies

(a)  $\mu_C = 40 + 45 = \$85,000$  per year.

(b)  $\text{corr}(M, F) = \frac{\text{cov}(M, F)}{\sigma_M \sigma_F}$ , so that  $\text{cov}(M, F) = \sigma_M \sigma_F \text{corr}(M, F)$ . Thus  $\text{cov}(M, F) = 12 \times 18 \times 0.80 = 172.80$ , where the units are squared thousands of dollars per year.

(c)  $\sigma_C^2 = \sigma_M^2 + \sigma_F^2 + 2\text{cov}(M, F)$ , so that  $\sigma_C^2 = 12^2 + 18^2 + 2 \times 172.80 = 813.60$ , and  $\sigma_C = \sqrt{813.60} = 28.524$  thousand dollars per year.

(d) First you need to look up the current Euro/dollar exchange rate in the Wall Street Journal, the Federal Reserve web page, or other financial data outlet. Suppose that this exchange rate is  $e$  (say  $e = 0.85$  Euros per Dollar or  $1/e = 1.18$  Dollars per Euro); each 1 Dollar is therefore with  $e$  Euros. The mean is therefore  $e \times \mu_C$  (in units of thousands of euros per year), and the standard deviation is  $e \times \sigma_C$  (in units of thousands of euros per year). The correlation is unit-free, and is unchanged.

2.8.  $\mu_Y = E(Y) = 1$ ,  $\sigma_Y^2 = \text{var}(Y) = 4$ . With  $Z = \frac{1}{2}(Y - 1)$ ,

$$\mu_Z = E\left(\frac{1}{2}(Y - 1)\right) = \frac{1}{2}(\mu_Y - 1) = \frac{1}{2}(1 - 1) = 0,$$

$$\sigma_Z^2 = \text{var}\left(\frac{1}{2}(Y - 1)\right) = \frac{1}{4}\sigma_Y^2 = \frac{1}{4} \times 4 = 1.$$

2.9.

		Value of $Y$					Probability Distribution of $X$
		14	22	30	40	65	
Value of $X$	1	0.02	0.05	0.10	0.03	0.01	0.21
	5	0.17	0.15	0.05	0.02	0.01	0.40
	8	0.02	0.03	0.15	0.10	0.09	0.39
Probability distribution of $Y$		0.21	0.23	0.30	0.15	0.11	1.00

(a) The probability distribution is given in the table above.

$$E(Y) = 14 \times 0.21 + 22 \times 0.23 + 30 \times 0.30 + 40 \times 0.15 + 65 \times 0.11 = 30.15$$

$$E(Y^2) = 14^2 \times 0.21 + 22^2 \times 0.23 + 30^2 \times 0.30 + 40^2 \times 0.15 + 65^2 \times 0.11 = 1127.23$$

$$\text{var}(Y) = E(Y^2) - [E(Y)]^2 = 218.21$$

$$\sigma_Y = 14.77$$

(b) The conditional probability of  $Y|X = 8$  is given in the table below

Value of $Y$				
14	22	30	40	65
0.02/0.39	0.03/0.39	0.15/0.39	0.10/0.39	0.09/0.39

$$E(Y|X = 8) = 14 \times (0.02/0.39) + 22 \times (0.03/0.39) + 30 \times (0.15/0.39) + 40 \times (0.10/0.39) + 65 \times (0.09/0.39) = 39.21$$

$$E(Y^2|X = 8) = 14^2 \times (0.02/0.39) + 22^2 \times (0.03/0.39) + 30^2 \times (0.15/0.39) + 40^2 \times (0.10/0.39) + 65^2 \times (0.09/0.39) = 1778.7$$

$$\text{var}(Y) = 1778.7 - 39.21^2 = 241.65$$

$$\sigma_{Y|X=8} = 15.54$$

(c)

$$E(XY) = (1 \times 14 \times 0.02) + (1 \times 22 \times 0.05) + \dots + (8 \times 65 \times 0.09) = 171.7$$

$$\text{cov}(X, Y) = E(XY) - E(X)E(Y) = 171.7 - 5.33 \times 30.15 = 11.0$$

$$\text{corr}(X, Y) = \text{cov}(X, Y) / (\sigma_X \sigma_Y) = 11.0 / (2.60 \times 14.77) = 0.286$$

2.10. Using the fact that if  $Y \sim N(\mu_Y, \sigma_Y^2)$  then  $\frac{Y - \mu_Y}{\sigma_Y} \sim N(0, 1)$  and Appendix Table 1, we have

$$(a) \quad \Pr(Y \leq 3) = \Pr\left(\frac{Y - 1}{2} \leq \frac{3 - 1}{2}\right) = \Phi(1) = 0.8413.$$

(b)

$$\begin{aligned} \Pr(Y > 0) &= 1 - \Pr(Y \leq 0) \\ &= 1 - \Pr\left(\frac{Y - 3}{3} \leq \frac{0 - 3}{3}\right) = 1 - \Phi(-1) = \Phi(1) = 0.8413. \end{aligned}$$

(c)

$$\begin{aligned} \Pr(40 \leq Y \leq 52) &= \Pr\left(\frac{40 - 50}{5} \leq \frac{Y - 50}{5} \leq \frac{52 - 50}{5}\right) \\ &= \Phi(0.4) - \Phi(-2) = \Phi(0.4) - [1 - \Phi(2)] \\ &= 0.6554 - 1 + 0.9772 = 0.6326. \end{aligned}$$

(d)

$$\begin{aligned} \Pr(6 \leq Y \leq 8) &= \Pr\left(\frac{6 - 5}{\sqrt{2}} \leq \frac{Y - 5}{\sqrt{2}} \leq \frac{8 - 5}{\sqrt{2}}\right) \\ &= \Phi(2.1213) - \Phi(0.7071) \\ &= 0.9831 - 0.7602 = 0.2229. \end{aligned}$$

2.11. (a) 0.90

(b) 0.05

(c) 0.05

(d) When  $Y \sim \chi^2_{10}$ , then  $Y/10 \sim F_{10,\infty}$ .

(e)  $Y = Z^2$ , where  $Z \sim N(0,1)$ , thus  $\Pr(Y \leq 1) = \Pr(-1 \leq Z \leq 1) = 0.32$ .

---

2.12. (a) 0.05

(b) 0.950

(c) 0.953

(d) The  $t_{df}$  distribution and  $N(0, 1)$  are approximately the same when  $df$  is large.

(e) 0.10

(f) 0.01

2.13. (a)  $E(Y^2) = \text{Var}(Y) + \mu_Y^2 = 1 + 0 = 1$ ;  $E(W^2) = \text{Var}(W) + \mu_W^2 = 100 + 0 = 100$ .

(b)  $Y$  and  $W$  are symmetric around 0, thus skewness is equal to 0; because their mean is zero, this means that the third moment is zero.

(c) The kurtosis of the normal is 3, so  $3 = \frac{E(Y - \mu_Y)^4}{\sigma_Y^4}$ ; solving yields  $E(Y^4) = 3$ ; a similar calculation yields the results for  $W$ .

(d) First, condition on  $X = 0$ , so that  $S = W$ :

$$E(S|X = 0) = 0; E(S^2|X = 0) = 100, E(S^3|X = 0) = 0, E(S^4|X = 0) = 3 \times 100^2$$

Similarly,

$$E(S|X = 1) = 0; E(S^2|X = 1) = 1, E(S^3|X = 1) = 0, E(S^4|X = 1) = 3.$$

From the large of iterated expectations

$$E(S) = E(S|X = 0) \times \Pr(X = 0) + E(S|X = 1) \times \Pr(X = 1) = 0$$

$$E(S^2) = E(S^2|X = 0) \times \Pr(X = 0) + E(S^2|X = 1) \times \Pr(X = 1) = 100 \times 0.01 + 1 \times 0.99 = 1.99$$

$$E(S^3) = E(S^3|X = 0) \times \Pr(X = 0) + E(S^3|X = 1) \times \Pr(X = 1) = 0$$

$$\begin{aligned} E(S^4) &= E(S^4|X = 0) \times \Pr(X = 0) + E(S^4|X = 1) \times \Pr(X = 1) \\ &= 3 \times 100^2 \times 0.01 + 3 \times 1 \times 0.99 = 302.97 \end{aligned}$$

(e)  $\mu_S = E(S) = 0$ , thus  $E(S - \mu_S)^3 = E(S^3) = 0$  from part (d). Thus skewness = 0.

Similarly,  $\sigma_S^2 = E(S - \mu_S)^2 = E(S^2) = 1.99$ , and  $E(S - \mu_S)^4 = E(S^4) = 302.97$ .

Thus, kurtosis =  $302.97 / (1.99^2) = 76.5$

2.14. The central limit theorem suggests that when the sample size ( $n$ ) is large, the distribution of the sample average ( $\bar{Y}$ ) is approximately  $N(\mu_Y, \sigma_Y^2)$  with  $\sigma_Y^2 = \frac{\sigma_Y^2}{n}$ .

Given  $\mu_Y = 100$ ,  $\sigma_Y^2 = 43.0$ ,

(a)  $n = 100$ ,  $\sigma_{\bar{Y}}^2 = \frac{\sigma_Y^2}{n} = \frac{43}{100} = 0.43$ , and

$$\Pr(\bar{Y} \leq 101) = \Pr\left(\frac{\bar{Y} - 100}{\sqrt{0.43}} \leq \frac{101 - 100}{\sqrt{0.43}}\right) \approx \Phi(1.525) = 0.9364.$$

(b)  $n = 165$ ,  $\sigma_{\bar{Y}}^2 = \frac{\sigma_Y^2}{n} = \frac{43}{165} = 0.2606$ , and

$$\begin{aligned} \Pr(\bar{Y} > 98) &= 1 - \Pr(\bar{Y} \leq 98) = 1 - \Pr\left(\frac{\bar{Y} - 100}{\sqrt{0.2606}} \leq \frac{98 - 100}{\sqrt{0.2606}}\right) \\ &\approx 1 - \Phi(-3.9178) = \Phi(3.9178) = 1.000 \text{ (rounded to four decimal places).} \end{aligned}$$

(c)  $n = 64$ ,  $\sigma_{\bar{Y}}^2 = \frac{\sigma_Y^2}{n} = \frac{43}{64} = 0.6719$ , and

$$\begin{aligned} \Pr(101 \leq \bar{Y} \leq 103) &= \Pr\left(\frac{101 - 100}{\sqrt{0.6719}} \leq \frac{\bar{Y} - 100}{\sqrt{0.6719}} \leq \frac{103 - 100}{\sqrt{0.6719}}\right) \\ &\approx \Phi(3.6599) - \Phi(1.2200) = 0.9999 - 0.8888 = 0.1111. \end{aligned}$$



2.15. (a)

$$\begin{aligned}\Pr(9.6 \leq \bar{Y} \leq 10.4) &= \Pr\left(\frac{9.6-10}{\sqrt{4/n}} \leq \frac{\bar{Y}-10}{\sqrt{4/n}} \leq \frac{10.4-10}{\sqrt{4/n}}\right) \\ &= \Pr\left(\frac{9.6-10}{\sqrt{4/n}} \leq Z \leq \frac{10.4-10}{\sqrt{4/n}}\right)\end{aligned}$$

where  $Z \sim N(0, 1)$ . Thus,

$$(i) \ n = 20; \Pr\left(\frac{9.6-10}{\sqrt{4/n}} \leq Z \leq \frac{10.4-10}{\sqrt{4/n}}\right) = \Pr(-0.89 \leq Z \leq 0.89) = 0.63$$

$$(ii) \ n = 100; \Pr\left(\frac{9.6-10}{\sqrt{4/n}} \leq Z \leq \frac{10.4-10}{\sqrt{4/n}}\right) = \Pr(-2.00 \leq Z \leq 2.00) = 0.954$$

$$(iii) \ n = 1000; \Pr\left(\frac{9.6-10}{\sqrt{4/n}} \leq Z \leq \frac{10.4-10}{\sqrt{4/n}}\right) = \Pr(-6.32 \leq Z \leq 6.32) = 1.000$$

(b)

$$\begin{aligned}\Pr(10-c \leq \bar{Y} \leq 10+c) &= \Pr\left(\frac{-c}{\sqrt{4/n}} \leq \frac{\bar{Y}-10}{\sqrt{4/n}} \leq \frac{c}{\sqrt{4/n}}\right) \\ &= \Pr\left(\frac{-c}{\sqrt{4/n}} \leq Z \leq \frac{c}{\sqrt{4/n}}\right).\end{aligned}$$

As  $n$  get large  $\frac{c}{\sqrt{4/n}}$  gets large, and the probability converges to 1.

(c) This follows from (b) and the definition of convergence in probability given in Key Concept 2.6.

- 2.16. There are several ways to do this. Here is one way. Generate  $n$  draws of  $Y$ ,  $Y_1, Y_2, \dots, Y_n$ . Let  $X_i = 1$  if  $Y_i < 3.6$ , otherwise set  $X_i = 0$ . Notice that  $X_i$  is a Bernoulli random variable with  $\mu_X = \Pr(X = 1) = \Pr(Y < 3.6)$ . Compute  $\bar{X}$ . Because  $\bar{X}$  converges in probability to  $\mu_X = \Pr(X = 1) = \Pr(Y < 3.6)$ ,  $\bar{X}$  will be an accurate approximation if  $n$  is large.

2.17.  $\mu_Y = 0.4$  and  $\sigma_Y^2 = 0.4 \times 0.6 = 0.24$

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

$$(ii) P(\bar{Y} \leq 0.37) = \Pr\left(\frac{\bar{Y} - 0.4}{\sqrt{0.24/n}} \leq \frac{0.37 - 0.4}{\sqrt{0.24/n}}\right) = \Pr\left(\frac{\bar{Y} - 0.4}{\sqrt{0.24/n}} \leq -1.22\right) = 0.11$$

b) We know  $\Pr(-1.96 \leq Z \leq 1.96) = 0.95$ , thus we want  $n$  to satisfy

$$0.41 = \frac{0.41 - 0.4}{\sqrt{0.24/n}} > -1.96 \quad \text{and} \quad \frac{0.39 - 0.4}{\sqrt{0.24/n}} < -1.96. \quad \text{Solving these inequalities yields } n \geq$$

9220.

2.18.  $\Pr(Y = \$0) = 0.95$ ,  $\Pr(Y = \$20000) = 0.05$ .

(a) The mean of  $Y$  is

$$\mu_Y = 0 \times \Pr(Y = \$0) + 20,000 \times \Pr(Y = \$20000) = \$1000.$$

The variance of  $Y$  is

$$\begin{aligned}\sigma_Y^2 &= E[(Y - \mu_Y)^2] \\ &= (0 - 1000)^2 \times \Pr(Y = 0) + (20000 - 1000)^2 \times \Pr(Y = 20000) \\ &= (-1000)^2 \times 0.95 + 19000^2 \times 0.05 = 1.9 \times 10^7,\end{aligned}$$

so the standard deviation of  $Y$  is  $\sigma_Y = (1.9 \times 10^7)^{\frac{1}{2}} = \$4359$ .

(b) (i)  $E(\bar{Y}) = \mu_Y = \$1000$ ,  $\sigma_{\bar{Y}}^2 = \frac{\sigma_Y^2}{n} = \frac{1.9 \times 10^7}{100} = 1.9 \times 10^5$ .

(ii) Using the central limit theorem,

$$\begin{aligned}\Pr(\bar{Y} > 2000) &= 1 - \Pr(\bar{Y} \leq 2000) \\ &= 1 - \Pr\left(\frac{\bar{Y} - 1000}{\sqrt{1.9 \times 10^5}} \leq \frac{2,000 - 1,000}{\sqrt{1.9 \times 10^5}}\right) \\ &\approx 1 - \Phi(2.2942) = 1 - 0.9891 = 0.0109.\end{aligned}$$

2.19. (a)

$$\begin{aligned}\Pr(Y = y_j) &= \sum_{i=1}^l \Pr(X = x_i, Y = y_j) \\ &= \sum_{i=1}^l \Pr(Y = y_j | X = x_i) \Pr(X = x_i)\end{aligned}$$

(b)

$$\begin{aligned}E(Y) &= \sum_{j=1}^k y_j \Pr(Y = y_j) = \sum_{j=1}^k y_j \sum_{i=1}^l \Pr(Y = y_j | X = x_i) \Pr(X = x_i) \\ &= \sum_{i=1}^l \left( \sum_{j=1}^k y_j \Pr(Y = y_j | X = x_i) \right) \Pr(X = x_i) \\ &= \sum_{i=1}^l E(Y | X = x_i) \Pr(X = x_i).\end{aligned}$$

(c) When  $X$  and  $Y$  are independent,

$$\Pr(X = x_i, Y = y_j) = \Pr(X = x_i) \Pr(Y = y_j),$$

so

$$\begin{aligned}\sigma_{XY} &= E[(X - \mu_X)(Y - \mu_Y)] \\ &= \sum_{i=1}^l \sum_{j=1}^k (x_i - \mu_X)(y_j - \mu_Y) \Pr(X = x_i, Y = y_j) \\ &= \sum_{i=1}^l \sum_{j=1}^k (x_i - \mu_X)(y_j - \mu_Y) \Pr(X = x_i) \Pr(Y = y_j) \\ &= \left( \sum_{i=1}^l (x_i - \mu_X) \Pr(X = x_i) \right) \left( \sum_{j=1}^k (y_j - \mu_Y) \Pr(Y = y_j) \right) \\ &= E(X - \mu_X) E(Y - \mu_Y) = 0 \times 0 = 0,\end{aligned}$$

$$\text{cor}(X, Y) = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{0}{\sigma_X \sigma_Y} = 0.$$

$$2.20. (a) \Pr(Y = y_i) = \sum_{j=1}^l \sum_{h=1}^m \Pr(Y = y_i | X = x_j, Z = z_h) \Pr(X = x_j, Z = z_h)$$

(b)

$$\begin{aligned} E(Y) &= \sum_{i=1}^k y_i \Pr(Y = y_i) \Pr(Y = y_i) \\ &= \sum_{i=1}^k y_i \sum_{j=1}^l \sum_{h=1}^m \Pr(Y = y_i | X = x_j, Z = z_h) \Pr(X = x_j, Z = z_h) \\ &= \sum_{j=1}^l \sum_{h=1}^m \left[ \sum_{i=1}^k y_i \Pr(Y = y_i | X = x_j, Z = z_h) \right] \Pr(X = x_j, Z = z_h) \\ &= \sum_{j=1}^l \sum_{h=1}^m E(Y | X = x_j, Z = z_h) \Pr(X = x_j, Z = z_h) \end{aligned}$$

where the first line in the definition of the mean, the second uses (a), the third is a rearrangement, and the final line uses the definition of the conditional expectation.

2. 21.

(a)

$$\begin{aligned} E(X - \mu)^3 &= E[(X - \mu)^2(X - \mu)] = E[X^3 - 2X^2\mu + X\mu^2 - X^2\mu + 2X\mu^2 - \mu^3] \\ &= E(X^3) - 3E(X^2)\mu + 3E(X)\mu^2 - \mu^3 = E(X^3) - 3E(X^2)E(X) + 3E(X)[E(X)]^2 - [E(X)]^3 \\ &= E(X^3) - 3E(X^2)E(X) + 2E(X)^3 \end{aligned}$$

(b)

$$\begin{aligned} E(X - \mu)^4 &= E[(X^3 - 3X^2\mu + 3X\mu^2 - \mu^3)(X - \mu)] \\ &= E[X^4 - 3X^3\mu + 3X^2\mu^2 - X\mu^3 - X^3\mu + 3X^2\mu^2 - 3X\mu^3 + \mu^4] \\ &= E(X^4) - 4E(X^3)E(X) + 6E(X^2)E(X)^2 - 4E(X)E(X)^3 + E(X)^4 \\ &= E(X^4) - 4[E(X)][E(X^3)] + 6[E(X)]^2[E(X^2)] - 3[E(X)]^4 \end{aligned}$$

2. 22. The mean and variance of  $R$  are given by

$$\begin{aligned}\mu &= w \times 0.08 + (1 - w) \times 0.05 \\ \sigma^2 &= w^2 \times 0.07^2 + (1 - w)^2 \times 0.04^2 + 2 \times w \times (1 - w) \times [0.07 \times 0.04 \times 0.25]\end{aligned}$$

where  $0.07 \times 0.04 \times 0.25 = \text{Cov}(R_s, R_b)$  follows from the definition of the correlation between  $R_s$  and  $R_b$ .

(a)  $\mu = 0.065; \sigma = 0.044$

(b)  $\mu = 0.0725; \sigma = 0.056$

(c)  $w = 1$  maximizes  $\mu; \sigma = 0.07$  for this value of  $w$ .

(d) The derivative of  $\sigma^2$  with respect to  $w$  is

$$\begin{aligned}\frac{d\sigma^2}{dw} &= 2w \times .07^2 - 2(1 - w) \times 0.04^2 + (2 - 4w) \times [0.07 \times 0.04 \times 0.25] \\ &= 0.0102w - 0.0018\end{aligned}$$

Solving for  $w$  yields  $w = 18 / 102 = 0.18$ . (Notice that the second derivative is positive, so that this is the global minimum.) With  $w = 0.18, \sigma_R = .038$ .



2. 23.  $X$  and  $Z$  are two independently distributed standard normal random variables, so

$$\mu_X = \mu_Z = 0, \sigma_X^2 = \sigma_Z^2 = 1, \sigma_{XZ} = 0.$$

(a) Because of the independence between  $X$  and  $Z$ ,  $\Pr(Z = z|X = x) = \Pr(Z = z)$ , and  $E(Z|X) = E(Z) = 0$ . Thus

$$E(Y|X) = E(X^2 + Z|X) = E(X^2|X) + E(Z|X) = X^2 + 0 = X^2.$$

(b)  $E(X^2) = \sigma_X^2 + \mu_X^2 = 1$ , and  $\mu_Y = E(X^2 + Z) = E(X^2) + \mu_Z = 1 + 0 = 1$ .

(c)  $E(XY) = E(X^3 + ZX) = E(X^3) + E(ZX)$ . Using the fact that the odd moments of a standard normal random variable are all zero, we have  $E(X^3) = 0$ . Using the independence between  $X$  and  $Z$ , we have  $E(ZX) = \mu_Z \mu_X = 0$ . Thus  $E(XY) = E(X^3) + E(ZX) = 0$ .

(d)

$$\begin{aligned} \text{cov}(XY) &= E[(X - \mu_X)(Y - \mu_Y)] = E[(X - 0)(Y - 1)] \\ &= E(XY - X) = E(XY) - E(X) \\ &= 0 - 0 = 0. \end{aligned}$$

$$\text{corr}(X, Y) = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{0}{\sigma_X \sigma_Y} = 0.$$

2.24. (a)  $E(Y_i^2) = \sigma^2 + \mu^2 = \sigma^2$  and the result follows directly.

(b)  $(Y_i/\sigma)$  is distributed i.i.d.  $N(0,1)$ ,  $W = \sum_{i=1}^n (Y_i/\sigma)^2$ , and the result follows from the definition of a  $\chi_n^2$  random variable.

$$(c) \quad E(W) = E \sum_{i=1}^n \frac{Y_i^2}{\sigma^2} = \sum_{i=1}^n E \frac{Y_i^2}{\sigma^2} = n.$$

(d) Write

$$V = \frac{Y_1}{\sqrt{\frac{\sum_{i=2}^n Y_i^2}{n-1}}} = \frac{Y_1/\sigma}{\sqrt{\frac{\sum_{i=2}^n (Y_i/\sigma)^2}{n-1}}}$$

which follows from dividing the numerator and denominator by  $\sigma$ .  $Y_1/\sigma \sim N(0,1)$ ,  $\sum_{i=2}^n (Y_i/\sigma)^2 \sim \chi_{n-1}^2$ , and  $Y_1/\sigma$  and  $\sum_{i=2}^n (Y_i/\sigma)^2$  are independent. The result then follows from the definition of the  $t$  distribution.

$$2.25. (a) \sum_{i=1}^n ax_i = (ax_1 + ax_2 + ax_3 + \cdots + ax_n) = a(x_1 + x_2 + x_3 + \cdots + x_n) = a \sum_{i=1}^n x_i$$

(b)

$$\begin{aligned} \sum_{i=1}^n (x_i + y_i) &= (x_1 + y_1 + x_2 + y_2 + \cdots + x_n + y_n) \\ &= (x_1 + x_2 + \cdots + x_n) + (y_1 + y_2 + \cdots + y_n) \\ &= \sum_{i=1}^n x_i + \sum_{i=1}^n y_i \end{aligned}$$

$$(c) \sum_{i=1}^n a = (a + a + a + \cdots + a) = na$$

(d)

$$\begin{aligned} \sum_{i=1}^n (a + bx_i + cy_i)^2 &= \sum_{i=1}^n (a^2 + b^2 x_i^2 + c^2 y_i^2 + 2abx_i + 2acy_i + 2bcx_i y_i) \\ &= na^2 + b^2 \sum_{i=1}^n x_i^2 + c^2 \sum_{i=1}^n y_i^2 + 2ab \sum_{i=1}^n x_i + 2ac \sum_{i=1}^n y_i + 2bc \sum_{i=1}^n x_i y_i \end{aligned}$$

$$2.26. \text{ (a) } \text{corr}(Y_i, Y_j) = \frac{\text{cov}(Y_i, Y_j)}{\sigma_{Y_i} \sigma_{Y_j}} = \frac{\text{cov}(Y_i, Y_j)}{\sigma_Y \sigma_Y} = \frac{\text{cov}(Y_i, Y_j)}{\sigma_Y^2} = \rho, \text{ where the first equality}$$

uses the definition of correlation, the second uses the fact that  $Y_i$  and  $Y_j$  have the same variance (and standard deviation), the third equality uses the definition of standard deviation, and the fourth uses the correlation given in the problem. Solving for  $\text{cov}(Y_i, Y_j)$  from the last equality gives the desired result.

$$\text{(b) } \bar{Y} = \frac{1}{2}Y_1 + \frac{1}{2}Y_2, \text{ so that } E(\bar{Y}) = \frac{1}{2}E(Y_1) + \frac{1}{2}E(Y_2) = \mu_Y$$

$$\text{var}(\bar{Y}) = \frac{1}{4}\text{var}(Y_1) + \frac{1}{4}\text{var}(Y_2) + \frac{2}{4}\text{cov}(Y_1, Y_2) = \frac{\sigma_Y^2}{2} + \frac{\rho\sigma_Y^2}{2}$$

$$\text{(c) } \bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i, \text{ so that } E(\bar{Y}) = \frac{1}{n} \sum_{i=1}^n E(Y_i) = \frac{1}{n} \sum_{i=1}^n \mu_Y = \mu_Y$$

$$\begin{aligned} \text{var}(\bar{Y}) &= \text{var}\left(\frac{1}{n} \sum_{i=1}^n Y_i\right) \\ &= \frac{1}{n^2} \sum_{i=1}^n \text{var}(Y_i) + \frac{2}{n^2} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{cov}(Y_i, Y_j) \\ &= \frac{1}{n^2} \sum_{i=1}^n \sigma_Y^2 + \frac{2}{n^2} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \rho\sigma_Y^2 \\ &= \frac{\sigma_Y^2}{n} + \frac{n(n-1)}{n^2} \rho\sigma_Y^2 \\ &= \frac{\sigma_Y^2}{n} + \left(1 - \frac{1}{n}\right) \rho\sigma_Y^2 \end{aligned}$$

where the fourth line uses  $\sum_{i=1}^{n-1} \sum_{j=i+1}^n a = a(1 + 2 + 3 + \cdots + n - 1) = \frac{an(n-1)}{2}$  for any variable  $a$ .

$$\text{(d) When } n \text{ is large } \frac{\sigma_Y^2}{n} \approx 0 \text{ and } \frac{1}{n} \approx 0, \text{ and the result follows from (c).}$$

2.27

$$(a) E(u) = E[E(u|X)] = E[E(Y - \hat{Y})|X] = E[E(Y|X) - E(Y|X)] = 0.$$

$$(b) E(uX) = E[E(uX|X)] = E[XE(u)|X] = E[X \times 0] = 0$$

(c) Using the hint:  $v = (Y - \hat{Y}) - h(X) = u - h(X)$ , so that  $E(v^2) = E[u^2] + E[h(X)^2] - 2 \times E[u \times h(X)]$ . Using an argument like that in (b),  $E[u \times h(X)] = 0$ . Thus,  $E(v^2) = E(u^2) + E[h(X)^2]$ , and the result follows by recognizing that  $E[h(X)^2] \geq 0$  because  $h(x)^2 \geq 0$  for any value of  $x$ .

2.28

$$\begin{aligned} \text{(a) } \Pr(M = 3) &= \Pr(M=3, A=0) + \Pr(M=3, A=1) \\ &= \Pr(M=3|A=0)\Pr(A=0) + \Pr(M=3|A=1)\Pr(A=1) \\ &= 0.05 \times 0.5 + 0.01 \times 0.5 = 0.03 \end{aligned}$$

(b)

$$\Pr(A = 0 | M = 3) = \Pr(M = 3 | A = 0) \times \frac{\Pr(A = 0)}{\Pr(M = 3)} = 0.05 \times \frac{0.5}{0.03} = 0.83$$

(c)

$$\begin{aligned} \Pr(M = 3) &= \Pr(M=3, A=0) + \Pr(M=3, A=1) \\ &= \Pr(M=3|A=0)\Pr(A=0) + \Pr(M=3|A=1)\Pr(A=1) \\ &= 0.05 \times 0.25 + 0.01 \times 0.75 = 0.02 \\ \Pr(A = 0 | M = 3) &= \Pr(M = 3 | A = 0) \times \frac{\Pr(A = 0)}{\Pr(M = 3)} = 0.05 \times \frac{0.25}{0.02} = 0.63 \end{aligned}$$