

Chapter 2

Review of Probability

- 2.6. The table shows that $\Pr(X = 0, Y = 0) = 0.037$, $\Pr(X = 0, Y = 1) = 0.622$,
 $\Pr(X = 1, Y = 0) = 0.009$, $\Pr(X = 1, Y = 1) = 0.332$, $\Pr(X = 0) = 0.659$, $\Pr(X = 1) = 0.341$,
 $\Pr(Y = 0) = 0.046$, $\Pr(Y = 1) = 0.954$.

(a) $E(Y) = \mu_Y = 0 \times \Pr(Y = 0) + 1 \times \Pr(Y = 1)$
 $= 0 \times 0.046 + 1 \times 0.954 = 0.954$.

(b) Unemployment Rate $= \frac{\#(\text{unemployed})}{\#(\text{labor force})}$
 $= \Pr(Y = 0) = 1 - \Pr(Y = 1) = 1 - E(Y) = 1 - 0.954 = 0.046$.

(c) Calculate the conditional probabilities first:

$$\Pr(Y = 0|X = 0) = \frac{\Pr(X = 0, Y = 0)}{\Pr(X = 0)} = \frac{0.037}{0.659} = 0.056,$$

$$\Pr(Y = 1|X = 0) = \frac{\Pr(X = 0, Y = 1)}{\Pr(X = 0)} = \frac{0.622}{0.659} = 0.944,$$

$$\Pr(Y = 0|X = 1) = \frac{\Pr(X = 1, Y = 0)}{\Pr(X = 1)} = \frac{0.009}{0.341} = 0.026,$$

$$\Pr(Y = 1|X = 1) = \frac{\Pr(X = 1, Y = 1)}{\Pr(X = 1)} = \frac{0.332}{0.341} = 0.974.$$

The conditional expectations are

$$E(Y|X = 1) = 0 \times \Pr(Y = 0|X = 1) + 1 \times \Pr(Y = 1|X = 1)$$
$$= 0 \times 0.026 + 1 \times 0.974 = 0.974,$$

$$E(Y|X = 0) = 0 \times \Pr(Y = 0|X = 0) + 1 \times \Pr(Y = 1|X = 0)$$
$$= 0 \times 0.056 + 1 \times 0.944 = 0.944.$$

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

Unemployment rate for college graduates $= 1 - E(Y|X = 1) = 1 - 0.974 = 0.026$

Unemployment rate for non-college graduates $= 1 - E(Y|X = 0) = 1 - 0.944 = 0.056$

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

$$\Pr(X = 1|Y = 0) = \frac{\Pr(X = 1, Y = 0)}{\Pr(Y = 0)} = \frac{0.009}{0.046} = 0.196.$$

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

$$\Pr(X = 0|Y = 0) = 1 - \Pr(X = 1|Y = 0) = 1 - 0.196 = 0.804.$$

- (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) $\Pr(X = 0|Y = 0) = 0.804$, while from the table $\Pr(X = 0) = 0.659$.

- 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 $e = 0.80$ Euros per dollar); each 1 dollar is therefore worth 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.

2.9.

		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

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

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

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

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

- 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 $N(\mu_Y, \sigma_{\bar{Y}}^2)$ with $\sigma_{\bar{Y}}^2 = \frac{\sigma_Y^2}{n}$. Given $\mu_Y = 100$, $\sigma_Y^2 = 43.0$,

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

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

(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

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

2.17. $\mu_Y = 0.4$ 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 $0.41 = \frac{0.41 - 0.40}{\sqrt{24/n}} > -1.96$
and $\frac{0.39 - 0.40}{\sqrt{24/n}} < -1.96$. 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,

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