Introduction to Econometrics (4th Edition)

by

James H. Stock and Mark W. Watson

Solutions to Odd-Numbered End-of-Chapter Exercises: Chapter 2*

(This version September 14, 2018)

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)	$Y < 0$	$0 \le Y < 1$	$1 \le Y < 2$	$Y \ge 2$
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

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

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

(a)

$$E(V) = E(20 - 7Y) = 20 - 7E(Y) = 20 - 7 \times 0.78 = 14.54,$$
$$E(W) = E(3 + 6X) = 3 + 6E(X) = 3 + 6 \times 0.70 = 7.2.$$

(b)

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

(c)

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

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

2.5. 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, $\mu_X = 70$°F implies that $\mu_Y = -17.78 + (5/9) \times 70 = 21.11$°C,

and $\sigma_X = 7$°F implies $\sigma_Y = (5/9) \times 7 = 3.89$°C.

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 Eollar 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.

2.9.

		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)

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

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

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

2.11. (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 $Z \sim \text{N}(0,1)$, thus $\Pr(Y \leq 1) = \Pr(-1 \leq Z \leq 1) = 0.32$.

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 $\text{E}(Y^4) = 3$; 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

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

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

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

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

(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, $\text{kurtosis} = 302.97 / (1.99^2) = 76.5$

2.15. (a)

$$\Pr(9.6 \le \overline{Y} \le 10.4) = \Pr\left(\frac{9.6-10}{\sqrt{4/n}} \le \frac{\overline{Y}-10}{\sqrt{4/n}} \le \frac{10.4-10}{\sqrt{4/n}}\right)$$
$$= \Pr\left(\frac{9.6-10}{\sqrt{4/n}} \le Z \le \frac{10.4-10}{\sqrt{4/n}}\right)$$

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

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

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

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

(b)

$$\Pr(10-c \le \overline{Y} \le 10+c) = \Pr\left(\frac{-c}{\sqrt{4/n}} \le \frac{\overline{Y}-10}{\sqrt{4/n}} \le \frac{c}{\sqrt{4/n}}\right)$$
$$= \Pr\left(\frac{-c}{\sqrt{4/n}} \le Z \le \frac{c}{\sqrt{4/n}}\right).$$

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.

2.17. $\mu_Y = 0.4$ 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 $0.41 = \frac{0.41-0.4}{\sqrt{0.24/n}} > -1.96$ and $\frac{0.39-0.4}{\sqrt{0.24/n}} < -1.96$. Solving these inequalities yields $n \geq$ 9220.

2.19. (a)

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

(b)

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

(c) When X and Y are independent,

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

so

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

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

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. 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)

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

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

2.25. (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)

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

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.