Problems: 3, 4, 5, 9, 12, 14, 15.

4. (a)

X \ Y	1	2	3	X
1	0.1	0.3	0.1	0.5
2	0.2	0	0.1	0.3
3	0	0.1	0.1	0.2
Y	0.3	0.4	0.3	

XY	p
1	0.1
2	0.5
3	0.1
4	0
6	0.2
9	0.1

$$\mathrm{cov}(X,Y) = E[X\cdot Y] - E[X]\cdot E[Y].$$

$$E[X] = 0.5\cdot(1) + 0.3\cdot(2) + 0.2\cdot(2)$$
$$= 1.7$$

$$E[Y] = 0.3\cdot(1) + 0.4\cdot(2) + 0.3\,(3)$$
$$= 2$$

$$E[XY] = 0.1\cdot(1) + 0.5\cdot(2) + 0.1\cdot(3)$$
$$+ 0.2\cdot(6) + 0.1\cdot(9)$$
$$= 3.5$$

$$\mathrm{cov}(X,Y) = 3.5 - (1.7)\cdot(2)$$
$$= 0.1$$

(b) $Var(X) = E[X^2] - (E[X])^2$

$(var(X) = cov(X, X))$

$E[X^2] = 0.5 \cdot (1^2) + 0.3 \cdot (2^2) + 0.2 \cdot (3^2)$

$= 3.5$

$var(X) = E[X^2] - (E[X])^2$

$= 3.5 - (1.7)^2$

$= 0.61$

(c) Very similar. (leave it as an exercise)

(d) A. $var(X+Y) = var(X) + var(Y) + 2 \cdot cov(X, Y)$.

(Useful when you know the three values)

B. Directly calculate the distribution of $X+Y$.

$X+Y$	p
2	0.1
3	0.5
4	0.1
5	0.2
6	0.1

$var(X+Y)$

$= E[(X+Y)^2] - (E[X+Y])^2$

3. $$\operatorname{corr}(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sqrt{\operatorname{var}(X)\cdot\operatorname{var}(Y)}}$$ $\rightarrow SD(X)\cdot SD(Y)$

$-1 \le \operatorname{corr} \le 1.$

- In this problem, $X+Y=0$.
- General case: $aX+bY=c$,

$a,b>0.$ $\operatorname{corr}(X,Y) = ?$

Prop. $\operatorname{var}(aX) = a^2\operatorname{var}(X)$.

$\operatorname{cov}(aX,Y) = a\cdot\operatorname{cov}(X,Y)$

$$\begin{aligned}
\operatorname{cov}(X,Y) &= \frac{1}{ab}\cdot\operatorname{cov}(aX,bY) \\
&= \frac{1}{ab}\cdot\operatorname{cov}(aX,\ c-aX) \\
&= \frac{1}{ab}\cdot\operatorname{cov}(aX,-aX) \\
&= (-a^2)\cdot\frac{1}{ab}\cdot\operatorname{cov}(X,X) \\
&= -\frac{a}{b}\cdot\operatorname{var}(X).
\end{aligned}$$

$$\operatorname{var}(X) = \operatorname{var}(X)$$

$$\begin{aligned}
\operatorname{var}(Y) &= \frac{1}{b^2}\cdot\operatorname{var}(bY) \\
&= \frac{1}{b^2}\operatorname{var}(-aX) \\
&= (-a)^2\cdot\frac{1}{b^2}\cdot\operatorname{var}(X) \\
&= \frac{a^2}{b^2}\cdot\operatorname{var}(X)
\end{aligned}$$

$$\begin{aligned}
\operatorname{corr}(X,Y) &= \frac{\operatorname{cov}(X,Y)}{\sqrt{\operatorname{var}(X)\cdot\operatorname{var}(Y)}} \\
&= \frac{-\cancel{\frac{a}{b}}\ \cancel{\operatorname{var}(X)}}{\sqrt{\cancel{\operatorname{var}(X)}\cdot\left(\cancel{\frac{a^2}{b^2}}\cancel{\operatorname{var}(X)}\right)}} = -1.
\end{aligned}$$

- Exercise: when $aX - bY = c$, $a, b > 0$,

  then $\mathrm{corr}(X, Y) = 1$.

5. X : wage income

   Y : asset income.

   $E[X] = 20{,}000$, $SD[X] = 5{,}000$.

   $E[Y] = 3{,}000$, $SD[Y] = 2{,}500$

   $\mathrm{corr}(X, Y) = 0.4$.

   $E[X+Y] = ?$ $SD[X+Y] = ?$

$$E[X+Y] = E[X] + E[Y] = 20{,}000 + 3{,}000 = 23000.$$

$$\begin{aligned}
\mathrm{Var}(X+Y) &= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2 \cdot \mathrm{Cov}(X, Y) \\
&= (SD[X])^2 + (SD[Y])^2 + 2 \cdot \mathrm{corr}(X, Y) \cdot SD[X] \cdot SD[Y]. \\
&= (5000)^2 + (2500)^2 + 2 \cdot (0.4) \cdot (2500) \cdot (5000) \\
&= 2500^2 (2^2 + 1^2 + 2 \cdot (0.4) \cdot 1 \cdot 2) \\
&= 2500^2 (6.6).
\end{aligned}$$

$$SD(X+Y) = 2500 \cdot \sqrt{6.6} = 6422.62.$$

9. prediction of X is always denoted by $\hat{X}$.

squared error cost: $\min_{\hat{X}} E(\hat{x}-x)^2$.

Scalar r.v.

$\Rightarrow \hat{x} = E[x]$.

absolute error cost: $\min_{\hat{x}} E|\hat{x}-x|$.

$\Rightarrow \hat{x} = \text{median}(x)$.

(i) $\hat{x} = E[x] = 7.5$

(ii) $\hat{x} = \text{median}(x)$, by the definition (for continuous r.v.), $F(\hat{x}) = 0.5$

$\Rightarrow (\hat{x})^3 / 1000 = 0.5$

$\Rightarrow \hat{x} = (500)^{1/3} = 7.937$.

12. predicted value of Y given $X=1$, is 0
(minimizing squared error cost.)

- The predicted value should be $E[Y \mid X=1]$

$$= \sum_{y} \mathbb{P}(Y=y \mid X=1) \cdot y$$

- the sum of all probabilities is 1.

$$\Rightarrow \quad a+b = 1 - 0.1 - 0.2 - 0.2 - 0.1 - 0.1$$

$$= 0.3$$

$$= \frac{0.1}{0.1+a+b} \cdot (-1) + \frac{b}{0.1+a+b} \cdot (1)$$

$$= \frac{0.1}{\cancel{0.4}} \cdot (-1) + \frac{b}{\cancel{0.4}} \cdot (1)$$

It equals to 0, So we know that $b=0.1$,

and hence $a = 0.3 - b = 0.2$

14. $\mathrm{var}(\bar{X}_n - \mu)$

$= \mathrm{var}(\bar{X}_n)$

$= \mathrm{var}\left(\frac{1}{n}(X_1 + X_2 + \cdots + X_n)\right)$

$= \frac{1}{n^2}\mathrm{var}(X_1 + X_2 + \cdots + X_n)$

Because $X_1, X_2, \cdots, X_n$ are independent.

$= \frac{1}{n^2}\left(\mathrm{var}(X_1) + \mathrm{var}(X_2) + \cdots + \mathrm{var}(X_n)\right)$

$= \frac{1}{n^2}\left(\sigma^2 + \sigma^2 + \cdots + \sigma^2\right)$

$= \frac{1}{n}\sigma^2$

This would get smaller when n is large.

• Note: $E[\bar{X}_n - \mu] = E[\bar{X}_n] - \mu$

$= 0.$

15. When $X_1, \cdots, X_n$ i.i.d with mean μ, variance σ^2,

we know that

$$\bar{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right), \text{ for large } n.$$

(Central Limit Theorem)

• $\bar{X}_{50} \sim N(80, 2)$, $\mathbb{P}(\bar{X}_{50} \leq 79.5)$

R code: pnorm(79.5, 80, sqrt(2))

(0.3618)

• $\bar{X}_{100} \sim N(80, 1)$, $\mathbb{P}(\bar{X}_{100} \leq 79.5)$

R code: pnorm(79.5, 80, 1)

(0.3085)