Solutions of exercises Chapter 5

Table for exercise 1: (1 = "Tail" and 0 = "Head"**:** $4^2 = 16$ possible results in 4 tosses)

result	***X***	***Y***	**result**	***X***	***Y***	**result**	***X***	***Y***	**result**	***X***	***Y***
0000	0	0	1000	1	0	0110	2	1	1101	3	1
0001	1	1	0011	2	2	1010	2	1	1011	3	2
0010	1	1	0101	2	1	1100	2	0	0111	3	2
0100	1	0	1001	2	1	1110	3	1	1111	4	2

Table of joint probabilities
$P(X = x \text{ and } Y = y)$

Note: X *is* $B\left(4, \frac{1}{2}\right)$ *–distributed and*
Y *is* $B\left(2, \frac{1}{2}\right)$ *– distributed*

y / x	**0**	**1**	**2**	$P(X = x)$
0	1/16	0	0	**1/16**
1	2/16	2/16	0	**4/16**
2	1/16	4/16	1/16	**6/16**
3	0	2/16	2/16	**4/16**
4	0	0	1/16	**1/16**
$P(Y = y)$	¼	½	¼	**1**

1. a. *see the table above*

b. $P(X = 1|Y = 1) = \frac{P(X=1 \text{ en } Y=1)}{P(Y=1)} = \frac{2/16}{1/2} = \frac{1}{4}$,

Likewise: $P(X = 2|Y = 1) = \frac{1}{2}$ and $P(X = 3|Y = 1) = \frac{1}{4}$.

So $E(X|Y = 1) = \sum_x x \cdot P(X = x|Y = 1) = 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{2} + 3 \cdot \frac{1}{4} = 2$. (or directly because of symmetry)

c. $P(Y = 1|X = 3) = \frac{P(X=3 \text{ en } Y=1)}{P(X=3)} = \frac{2/16}{4/16} = \frac{1}{2}$

2. b. $P(X = i\,) = P(X = i \text{ and } Y = 0) + P(X = i \text{ and } Y = 1)$

$$= \left(\frac{1}{3}\right)^i \cdot \left(\frac{2}{3}\right) + \left(\frac{1}{3}\right)^{i-1} \left(\frac{2}{3}\right)^2 = \left(\frac{1}{3}\right)^{i-1} \cdot \left[\frac{1}{3} \cdot \frac{2}{3} + \left(\frac{2}{3}\right)^2\right] = \left(\frac{1}{3}\right)^{i-1} \left(\frac{2}{3}\right), \text{ for } i = 1, 2, 3, \ldots,$$

so $X \sim$geometric $\left(p = \frac{2}{3}\right)$ and $E(X) = \frac{1}{p} = \frac{3}{2}$

c. $P(Y = 0) = \sum_{i=1}^{\infty} \left(\frac{1}{3}\right)^i \cdot \left(\frac{2}{3}\right) = \cdots$ *geometric series ...* $= \left(\frac{1}{1-\frac{1}{3}} - 1\right) \cdot \left(\frac{2}{3}\right) = \frac{1}{3}$ and

(Note that $\sum_{i=1}^{\infty} \left(\frac{1}{3}\right)^i = \sum_{i=0}^{\infty} \left(\frac{1}{3}\right)^i - \left(\frac{1}{3}\right)^0$)

$P(Y = 1) = \sum_{i=1}^{\infty} \left(\frac{1}{3}\right)^{i-1} \cdot \left(\frac{2}{3}\right)^2 = \frac{2}{3} \cdot \sum_{i=1}^{\infty} \left(\frac{1}{3}\right)^{i-1} \left(\frac{2}{3}\right) = \frac{2}{3}$

($\sum_{i=1}^{\infty} \left(\frac{1}{3}\right)^{i-1} \left(\frac{2}{3}\right) = 1$, since this is the sum of geometric probabilities)

(or using the formula: $P(Y = j) = \sum_{i=1}^{\infty} \left(\frac{1}{3}\right)^{i-j} \cdot \left(\frac{2}{3}\right)^{1+j} = \cdots = \left(\frac{1}{3}\right)^{1-j} \left(\frac{2}{3}\right)^j$, for $j = 0,1$)

d. Yes, $P(X = i \text{ and } Y = j) = \left(\frac{1}{3}\right)^{i-j} \left(\frac{2}{3}\right)^{1+j} = \left(\frac{1}{3}\right)^{i-1} \left(\frac{2}{3}\right) \cdot \left(\frac{1}{3}\right)^{1-j} \left(\frac{2}{3}\right)^j = P(X = i) \cdot P(Y = j)$

for all i, j combinations. You can also check the equality for $j = 0$ and $j = 1$ separately.

3.

a. Distribution of X: add the probabilities in each row:

i	0	1	2	Sum
$P(X = i)$	0.2	0.5	0.3	1

Distribution of Y: add the probabilities in each column:

j	0	1	2	3	sum
$P(Y = j)$	0.10	0.30	0.45	0.15	1

b. $E(X) = 1.1, E(X^2) = 1.7$ and $var(X) = 0.49$, $E(Y) = 1.65, E(Y^2) = 3.45$ and $var(Y) = 0.7275$

c. For $Z = 8Y$ we have $E(Z) = 8 \cdot E(Y) = 8 \cdot 1.65 = 13.20$ and

$$var(8Y) = 8^2 var(Y) = 64 \cdot 0.7275 = 46.56.$$

d. The values of $T = X + Y$ are integers 0 to 5. (*add the probabilities "diagonally"*)

t	0	1	2	3	4	5	Sum
$P(T = t)$	0.05	0.10	0.20	0.40	0.20	0.05	1

$E(T) = 2.75$ and $var(T) = 1.3875$

e. $E(T) = E(X + Y)$ is in agreement with the sum of expectations $E(X) = 1.10$ and $E(Y) = 1.65$.

f. $var(T) = var(X + Y)$ does not correspond with $var(X) + var(Y) = 0.490 + 0.7275 = 1.2175$, which is caused by the dependence of X and Y.

4. a. N is geometrically distributed with $p = \frac{1}{2}$, so

$P(N = 10) = \left(\frac{1}{2}\right)^{10}, P(X = 4|N = 10) = \binom{10}{4}\left(\frac{1}{2}\right)^{10}$ and

$P(X = 4 \text{ and } N = 10) = P(X = 4|N = 10) \cdot P(N = 10)$ (*Product rule* $P(A \cap B) = P(A|B) \cdot P(B)$)

$= \binom{10}{4}\left(\frac{1}{4}\right)^{10}$ (And $E(N) = \frac{1}{p} = 2$)

b. $P(N = n) = \left(\frac{1}{2}\right)^n, \ n = 1, 2, \ldots$ *as stated in a.!*

$P(X = x|N = n) = \binom{n}{x}\left(\frac{1}{2}\right)^n, \ x = 0, 1, 2, \ldots, n$, so X is, given $N = n$, $B\left(n, \frac{1}{2}\right)$-distributed.

$P(X = x \text{ and } N = n) = \binom{n}{x}\left(\frac{1}{4}\right)^n, \ x = 0, 1, 2, \ldots, n$ and $n = 1, 2, \ldots$

c. X is, given $N = n$, $B\left(n, \frac{1}{2}\right)$-distributed, so:

$E(X|N = 10) = 10 \cdot \frac{1}{2} = 5$,

$E(X|N = n) = n \cdot \frac{1}{2} = \frac{1}{2}n$, (*so* $E(X|N) = \frac{1}{2}N$ *is a* ***random variable***!)

$E(X) = E[E(X|N)] = E\left(\frac{1}{2}N\right) = \frac{1}{2}E(N) = \frac{1}{2} \cdot 2 = 1$

d. $P(X = 0) = P(X = 0 \text{ and } N = 1) + P(X = 0 \text{ and } N = 2) + \cdots$

$= \sum_{n=1}^{\infty} P(X = 0 \text{ and } N = n) = \sum_{n=1}^{\infty} P(X = 0 \mid N = n)P(N = n)$

$= \sum_{n=1}^{\infty}\left(\frac{1}{2}\right)^n \cdot \left(\frac{1}{2}\right)^n = \sum_{n=1}^{\infty}\left(\frac{1}{4}\right)^n = \frac{1}{4} \cdot \frac{1}{1-\frac{1}{4}} = \frac{1}{3}$.

e. $P(N = n|X = 0) = \frac{P(X=0 \text{ and } N=n)}{P(X=0)} = \frac{\left(\frac{1}{4}\right)^n}{\frac{1}{3}} = \frac{\left(\frac{1}{4}\right)^{n-1} \cdot \left(\frac{1}{4}\right)}{\frac{1}{3}} = \left(\frac{1}{4}\right)^{n-1} \cdot \left(\frac{3}{4}\right)$, for $n = 1, 2, 3, \ldots$

N is, given $X = 0$, geometrically distributed with parameter $p = \frac{3}{4}$, so $E(N|X = 0) = \frac{4}{3}$.

5. a. if $N = n$, then $S = X_1 + \cdots + X_n$

b. $E(X_i) = \sum_x x \cdot P(X_i = x) = 1000 \cdot \frac{1}{10} + 2000 \cdot \frac{3}{10} + 3000 \cdot \frac{4}{10} + 4000 \cdot \frac{2}{10} = 2700 \ (i = 1, \ldots, n)$

c. $E(S|N = n) = E(X_1 + \cdots + X_n) = E(X_1) + \cdots + E(X_n) = n \cdot 2700 = 2700n$

d. Using c. we find $E(S|N) = 2700N$ ($E(S|N)$ *is a random variable*!)

According to property 5.2.6 we have: $E(X) = E[E(X|Y)]$, so in this case:

$E(S) = E[\,E(S|N)] = E(2700N) = 2700 \cdot E(N) = 2700\mu$

Intuitively this is correct: we expect a number of damages (μ is not necessarily integer!), and for each damage one would expect an amount of 2700 Euro.

d. $P(N = n|X = 0) = \frac{P(X=0 \text{ en } N=n)}{P(X=0)} = \frac{\left(\frac{1}{4}\right)^n}{\frac{1}{3}} = \frac{\left(\frac{1}{4}\right)^{n-1} \cdot \left(\frac{1}{4}\right)}{\frac{1}{3}} = \left(\frac{1}{4}\right)^{n-1} \cdot \left(\frac{3}{4}\right)$, for $n = 1, 2, 3, \ldots$

N is, given $X = 0$, geometrically distributed with parameter $p = \frac{3}{4}$, so $E(N|X = 0) = \frac{4}{3}$.

6. a. $P(X = 8 \text{ and } Y = 2) = P(Y = 2|X = 8)P(X = 8)$ (*product rule* $P(AB) = P(A|B)P(B)$)

$= \binom{8}{2}0.3^2 0.7^6 \cdot \frac{10^8 \cdot e^{-10}}{8!}$ ($B(8,0.3)$- and Poisson (μ = 10)-distr., resp.)

$= 0.033.$

b. $E(Y|X=8) = np = 8 \cdot 0.3 = 2.4$ and $E(Y|X=x) = x \cdot 0.3 = 0.3x$. (*so* $E(Y|X) = 0.3X$)
c. $E(Y) = E[\,E(Y|X)] = E[\,0.3X] = 0.3 \cdot E(X) = 3$
Intuitively this result is correct: if X customers enter ("average 10"), we would expect that about 30% need a service time of more than 3 minutes.

7. a. $P(X_1 = 10) = 0.9^9 \cdot 0.1 = 3.87\%$
b. *It is convenient to use the property* $P(X > x) = (1-p)^x$ *of the geometric distribution:*
$P(20 \le X_1 \le 30) = P(X_1 > 19) - P(X_1 > 30) = 0.9^{19} - 0.9^{30} \approx 9.27\%$
c. $P(X_1 = X_2) = P(X_1 = 1 \text{ and } X_2 = 1) + P(X_1 = 2 \text{ and } X_2 = 2) + \cdots$

$$\overset{\text{ind.}}{=} \sum_{k=1}^{\infty} P(X_1 = k) \cdot P(X_2 = k)$$
$$= \sum_{k=1}^{\infty} 0.9^{k-1} 0.1 \cdot\ 0.9^{k-1} 0.1$$
$$= \sum_{k=1}^{\infty} 0.81^{k-1} 0.01$$
$$= \frac{1}{1-0.81} \cdot 0.01 \approx 5.26\%$$

d. $P(X_1 + X_2 = 20) = \sum_{k=1}^{19} P(X_1 = k) \cdot P(X_2 = 20-k)$
$$= \sum_{k=1}^{19} 0.9^{k-1} 0.1 \cdot\ 0.9^{20-k-1} 0.1$$
$$= \sum_{k=1}^{19} 0.9^{18} 0.1^2$$
$$= 19 \cdot 0.9^{18} 0.1^2 \approx 2.85\%$$

8. a. $P(X > i \text{ and } Y > i) \overset{\text{ind.}}{=} P(X > i)P(Y > i) = (1-p)^i \cdot (1-p)^i = [(1-p)^2]^i$
b. $P(min(X,Y) > i) = P(X > i \text{ and } Y > i) = [(1-p)^2]^i$, for $i = 0, 1, 2, \ldots$
c. $P(min(X,Y) = i) = P(min(X,Y) > i-1) - P(min(X,Y) > i)$
$$= [(1-p)^2]^{i-1} - [(1-p)^2]^i = [(1-p)^2]^{i-1}[1-(1-p)^2]$$
for $i = 1, 2, \ldots$
d. So $min(X,Y)$ is geometrically distributed with parameter $1-(1-p)^2 = 2p - p^2$ and
$E[min(X,Y)] = \frac{1}{2p-p^2}$

9. a. $E(X+Y) = E(X) + E(Y) = \frac{2}{p}$ and $var(X+Y) \overset{\text{ind.}}{=} var(X) + var(Y) = 2 \cdot \frac{1-p}{p^2}$
b. $P(X = i \text{ and } Y = j) \overset{\text{ind.}}{=} P(X = i) \cdot P(Y = j) = (1-p)^{i-1}p \cdot (1-p)^{j-1}p$
$$= (1-p)^{i+j-2}p^2 \qquad (i,j = 1, 2, \ldots)$$
c. $P(X+Y = n) = \sum_{i=1}^{n-1} P(X = i) \cdot P(Y = n-i)$
$$= \sum_{i=1}^{n-1} (1-p)^{i-1}p \cdot (1-p)^{n-i-1}p = \sum_{i=1}^{n-1} (1-p)^{n-2}p^2$$
$$= (n-1) \cdot (1-p)^{n-2}p^2, n = 1, 2, 3, \ldots$$

10. a. X and Y have the same distribution (alongside):
$E(X) = 1$ (symmetry) and
$var(X) = E(X^2) - \mu^2 = 1.6 - 1 = 0.6$
(so $\sigma_X = \sigma_Y = \sqrt{0.6}$.)

i	0	1	2	som
$P(X=i)$	0.3	0.4	0.3	1
$i \cdot P(X=i)$	0	0.4	0.6	$1 = E(X)$
$i^2 \cdot P(X=i)$	0	0.4	1.2	$1.6 = E(X^2)$

b. $+\frac{2}{3}$, 0, 0 and -1, respectively. Explanation:
Distribution 1: If you graph the points in the XY-plane, where i is on the X-axis and j on the Y-axis, these points lie "around the line $y = x$": the correlation coefficient must be positive so 2/3, 1 or 2: 2 is impossible since $\rho \le 1$ and 1 (strictly linear relation) is impossible since not all points are on the line $y = x$. So $\rho(X,Y) = \frac{2}{3}$, which is the result if we really compute it.
Distributions 2 and 3 do not show some linear relation: $\rho = 0$.
Distribution 4: all points lie on a line with a negative slope: a strictly linear, negative relation: $\rho = -1$
c. Only in distribution 3 X and Y are independent:
$P(X = i \text{ and } = j) = P(X = i) \cdot P(Y = j)$, for all (i,j).
The dependence of X and Y is easily proven in the other 3 cases: if the joint probability is 0 in a cell (i,j) and the accompanying row and column totals are not 0.
Note: for distribution 3: X and Y are ***independent and not correlated*** ($\rho = 0$), *but for*

*distribution 2: X and Y are **dependent and not correlated**.*
*These examples illustrate: **independence implies non-correlation, but not reverse**!.*

d. $E(XY) = \sum_i \sum_j \ i \cdot j \cdot P(X = i \text{ en } Y = j) = 1 \cdot 0.2 + 2 \cdot 0.1 + 2 \cdot 0.1 + 4 \cdot 0.2 = 1.4$

$cov(X,Y) = E(XY) - E(X)E(Y) = 1.4 - 1 \cdot 1 = 0.4$

$\rho(X,Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{0.4}{\sqrt{0.6}\sqrt{0.6}} = +\frac{2}{3}$

e. $cov(3X, 2-Y) = 3 \cdot -1 \cdot cov(X,Y) = -3 \cdot 0.4 = -1.2$,

dus $\rho(3X, 2-Y) = -\rho(X,Y) = -\frac{2}{3}$

f. $E(XY) = 1$ (d.) and $EX = EY = 1$, so:

$cov(X,Y) = E(XY) - E(X)E(Y) = 1 - 1 \cdot 1 = 0$, and $\rho(X,Y) = 0$

g. The results are in the table, e.g. for $Y = 0$: $P(X = 0|Y = 0) = \frac{P(X=0 \text{ and } Y=0)}{P(Y=0)} = \frac{0.2}{0.3} = \frac{2}{3}$

x	0	1	2	total	
$P(X = x\|Y = 0)$	$\frac{2}{3}$	$\frac{1}{3}$	0	1	So $E(X\|Y = 0) = 0 \cdot \frac{2}{3} + 1 \cdot \frac{1}{3} = \frac{1}{3}$
$P(X = x\|Y = 1)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	1	So $E(X\|Y = 1) = 1\ (symmetry)$
$P(X = x\|Y = 2)$	0	$\frac{1}{3}$	$\frac{2}{3}$	1	So $E(X\|Y = 2) = 1 \cdot \frac{1}{3} + 2 \cdot \frac{2}{3} = \frac{5}{3}$

From a. we know that $E(X) = 1$.

And using the results in the table: $\sum_y E(X|Y = y) \cdot P(Y = y) = \frac{1}{3} \cdot 0.3 + 1 \cdot 0.4 + \frac{5}{3} \cdot 0.3 = 1.0$

(*Note: the last expression determines the expected value of the variable $E(X|Y)$, that can take on values $E(X|Y = 0)$, $E(X|Y = 1)$ and $E(X|Y = 2)$ with probabilities $P(Y = 0), P(Y = 1)$ and $P(Y = 2)$, respectively.* **)**

11. $\rho(X,Y) = \rho(X, -3X + 4) = \frac{cov(X, -3X+4)}{\sigma_X \sigma_{-3X+4}}$

Since $var(-3X + 4) = (-3)^2 \cdot var(X) = 9 \cdot var(X)$, so $\sigma_{-3X+4} = 3\sigma_X$, and

$cov(X, -3X + 4) = -3 \cdot cov(X,X) = -3 \cdot var(X) = -3\sigma_X^2$,

we find: $\rho(X,Y) = \frac{-3\sigma_X^2}{\sigma_X \cdot 3\sigma_X} = -1$

12. a. $cov(X_1, X_1 + X_2) = cov(X_1, X_1) + cov(X_1, X_2) \overset{\text{ind.}}{=} var(X_1) + 0 = 2$

$$\rho(X_1, X_1 + X_2) = \frac{cov(X_1, X_1 + X_2)}{\sigma_{X_1}\sigma_{X_1+X_2}} = \frac{var(X_1) + cov(X_1, X_2)}{\sigma_{X_1}\sqrt{var(X_1) + var(X_2)}} = \frac{var(X_1)}{\sigma_{X_1} \cdot \sqrt{2}\sigma_{X_1}} = \frac{1}{\sqrt{2}}$$

b. $\rho(X_1, X_1 + X_2 + \ldots + X_n) = \frac{cov((X_1, X_1 + X_2 + \ldots + X_n)}{\sigma_{X_1}\sigma_{X_1+\cdots+X_n}} = \frac{var(X_1)}{\sqrt{n}var(X_1)} = \frac{1}{\sqrt{n}} < \frac{1}{3}$, if $n > 9$

13. a. $E(X_1) = 1 \cdot \frac{1}{10} + 0 \cdot \frac{9}{10} = \frac{1}{10}$, $E(X_1{}^2) = \frac{1}{10}$ and $var(X_1) = \frac{1}{10} - \left(\frac{1}{10}\right)^2 = \frac{9}{100}$

$cov(X_1, X_2) = E(X_1X_2) - E(X_1)E(X_2) = 1 \cdot 1 \cdot \frac{1}{10} \cdot \frac{1}{9} - \left(\frac{1}{10}\right)^2 = \frac{1}{900}$

b. $E(S) = E\left(\sum_{i=1}^{10} X_i\right) = \sum_{i=1}^{10} E(X_i) = 10 \cdot E(X_1) = 1$

In the computation of $var(S)$ we will use that $cov(X_i, X_j)$ are the same for all $i \neq j$:

$var(S) = var\left(\sum_{i=1}^{10} X_i\right) = \sum_{i=1}^{10} var(X_i) + \sum \ \sum_{i \neq j} cov(X_i, X_j)$

$= 10 \cdot var(X_1) + 90 \cdot cov(X_1, X_2) = 10 \cdot \frac{9}{100} + 90 \cdot \frac{1}{900} = 1$

14. a. $X + Y \sim Poisson(\mu_1 + \mu_2)$, with $\mu_1 + \mu_2 = 2 + 3$

b. $P(X = k|X + Y = n) = \frac{P(X=x \text{ and } Y=n-k)}{P(X+Y=n)} = \frac{\frac{2^k e^{-2}}{k!} \cdot \frac{3^{n-k} e^{-3}}{(n-k)!}}{\frac{5^n e^{-5}}{n!}} = \frac{n!}{k!(n-k)!} \left(\frac{2}{5}\right)^k \left(\frac{3}{5}\right)^{n-k}$

(In this case $\mu_1 = 2$ and $\mu_2 = 3$).

This is for $k = 0, 1, 2, .., n$ the binomial probability function with $p = \frac{\mu_1}{\mu_1 + \mu_2} = \frac{2}{5}$.

c. $E(X|X+Y=7) = 7 \cdot \frac{2}{2+3} = 2.8$, since X is, given $X+Y=n$, $B\left(n, \frac{2}{5}\right)$– distributed.

Assumptions: X and Y, the numbers of appendicitis and kidney stones are independent and Poisson distributed with parameters 2 and 3, respectively.