Solutions to Assignment 3

1. Suppose that a discrete random variable X takes different values in the set $\{-2, -1, 0, 1, 2\}$ with the probabilities given as below:

Value	-1	0	1	2
Probability	1/10	2/5	1/5	1/10

Answer the following questions:

1. What is the value of $P(X = -2)$?

Solution:

$$P(X = -2) = 1 - \sum_{r=-1}^{2} P(X = r) = 1/5$$

2. What is the value of $E(X)$?

Solution:

$$E(X) = \sum_{r=-2}^{2} rP(X = r) = (-2/5 - 1/10 + 0 + 1/5 + 2/10) = -1/10$$

3. Suppose Y is another random variable with distribution identical to X. What is $E(X + 1/(Y + 3))$

Solution: Since expectations are linear $E(X+1/(Y+3)) = E(X)+E(1/(Y+3))$. Since $1/(Y + 3)$ is a function of Y we have

$$E(1/(Y+3)) = \sum_{r=-2}^{2} (1/r+3)P(Y = r) = (1/5+1/20+2/15+1/20+1/50) = 116/300$$

4. What is the variance $\sigma^2(X)$?

Solution:

$$E(X^2) = \sum_{r=-2}^{2} r^2P(X = r) = (4/5 + 1/10 + 0 + 1/5 + 4/10) = 15/10$$

We then calculate $\sigma^2(X) = E(X^2) - E(X)^2 = 149/100$.

2. Let X and Y be distinct real valued random variables.

   1. For all real numbers a and b check that $a^2E(X^2) + 2abE(XY) + b^2E(Y^2)$ is a non-negative number. (Hint: What is its relation with $E((aX + bY)^2)$?)

      Solution: We note that

      $$E((aX+bY)^2) = E(a^2X^2+2abXY+b^2Y^2) = a^2E(X^2)+2abE(XY)+b^2E(Y^2)$$

      the last equality is due to linearity of expectation.

   2. Suppose A, B and C are real numbers such that $A + 2Bt + Ct^2 \geq 0$ for all real numbers t show that $AC \geq B^2$.

      Solution: If $C = 0$, then $A+2Bt$ cannot be non-negative for all t unless $B = 0$. In that case $B^2 = 0 = AC$.
      Let is assume that $C \neq 0$. We then get

      $$A + 2Bt + Ct^2 = (AC - B^2)/C + (B + Ct)^2/C$$

      Since this is positive for all t, we note that $AC - B^2/C \geq 0$. Now, for t large, the second term dominates, so $C > 0$ as well. It follows that $AC - B^2 \geq 0$.

   3. Use the above two steps to conclude that $E(XY)^2 \leq E(X^2)E(Y^2)$.

      Solution: Combining these two results above we see easily that $E(X^2)E(Y^2) \geq E(XY)^2$.

   4. Check that $E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y)$. (Hint: Expand!)

      Solution:

      $$E((X-a)(Y-b)) = E(XY - aY - bX + ab) = E(XY) - aE(Y) - bE(X) + ab$$

      Now putting $a = E(X)$ and $b = E(Y)$, we get

      $$\begin{aligned} &E((X - E(X))(Y - E(Y))) = \\ &\qquad E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y) = \\ &\qquad\qquad E(XY) - E(X)E(Y) \end{aligned}$$

5. Let us define $c(X, Y) = E(XY) - E(X)E(Y)$. Show that $c(X, Y)^2 \leq \sigma^2(X)\sigma^2(Y)$. (Hint: Apply steps 1-3 to $X - E(X)$ and $Y - E(Y)$.)

Solution:

$$E((X - E(X))^2)E((Y - E(Y))^2) \geq E((X - E(X)(Y - E(Y)))^2$$

This is just a restatement of

$$\sigma^2(X)\sigma^2(Y) \geq c(X, Y)^2$$

6. Show that $c(X, Y) = 0$ if X and Y are independent.

Solution: We have seen already that if X and Y are independent then $E(XY) = E(X)E(Y)$. Thus $c(X, Y) = 0$. Note that $c(X, Y) = 0$, then it need not be the case that X and Y are independent!

The number $c(X, Y)$ is called the covariance of X and Y. Assuming that $\sigma^2(X)\sigma^2(Y)$ is non-zero, the ratio $\rho(X, Y) = c(X, Y)/\sigma(X)\sigma(Y)$ is called the correlation of X and Y. It lies between -1 and 1.

3. Suppose that X and Y are independent random variables.

1. Show that aX and bY are independent random variables for any non-zero constants a and b.

Solution: (We assume that the variables are discrete. The continuous case is similar.)

$$P((aX = r) \wedge (bY = s)) = P((X = r/a) \wedge (Y = s/b)) = \\ P(X = r/a) \cdot P(Y = s/b) = P(aX = r) \cdot P(bY = s)$$

Thus, these are independent.

2. Let $U = X + Y$ and $V = 2X - 3Y$. Calculate $E(U)$, $\sigma^2(U)$, $E(V)$ and $\sigma^2(V)$ in terms of $E(X)$, $\sigma^2(X)$, $E(Y)$ and $\sigma^2(Y)$.

Solution: By linearity of expectation

$$E(U) = E(X) + E(Y) \text{and} E(V) = 2E(X) - 3E(Y)$$

Since aX and bY are independent, we have

$$\sigma^2(U) = \sigma^2(X) + \sigma^2(Y) \text{and} \sigma^2(V) = 4\sigma^2(X) + 9\sigma^2(Y)$$

3. Let $W = 3X - 2Y = U + V$. Calculate $E(W)$ and $\sigma^2(W)$ in terms of $E(X)$, $E(Y)$, $\sigma^2(X)$ and $\sigma^2(Y)$.

Solution: As above we have

$$E(W) = 3E(X) - 2E(Y) \text{and} \sigma^2(W) = 9\sigma^2(X) + 4\sigma^2(Y)$$

4. Check whether U and V are independent.

Solution: Let us compare $\sigma^2(W) = \sigma^2(U + V)$ with $\sigma^2(U) + \sigma^2(V)$. The first one is $9\sigma^2(X) + 4\sigma^2(Y)$ and the second one is $5\sigma^2(X) + 10\sigma^2(Y)$. For most values of $\sigma^2(X)$ and $\sigma^2(Y)$ these numbers are different! Hence U and V are (in general) not independent.

This exercise shows that independence is quite a different notion from what one might think since X and Y (which are independent) can be written linearly in terms of U and V which are not independent!

4. Given any two positive numbers a and b such that $b < a$ and $a^2 = k(a - b)$ for some positive integer k. Show that there is a random variable X such that $E(X) = a$ and $\sigma^2(X) = b$. (Hint: Consider a suitable Binomial random variable.)

Solution: Assume that $p = 1 - b/a$ and consider the Binomial random variable X that counts the number of successes in k Bernoulli trials with probability p of success. Then $E(X) = k(1 - b/a) = k(a - b)/a = a$ and $\sigma^2(X) = k(b/a)(1 - b/a) = b$.

5. Given any two positive numbers a and b such that $a < b$ and $a^2 = k(b - a)$ for some positive integer k. Show that there is a random variable X such that $E(X) = a$ and $\sigma^2(X) = b$. (Hint: Consider a suitable Negative Binomial random variable.)

Solution: Assume that $p = 1 - a/b$ and consider the negative Binomial random variable X that counts the number of successes before finding k failures in a sequence of independent Bernoulli trials with probability p of success. Then $E(X) = k(1 - a/b)/(a/b) = k(b - a)/a = a$ and $\sigma^2(X) = k(1 - a/b)/(a/b)^2 = k(b - a)b/a^2 = b$

Note that, in fact, given any real number a and any positive number b, we can define a random variable with $E(X) = a$ and $\sigma^2(X) = b$. (The sign of a does not matter since $E(-X) = -a$. The tricky part is getting the case $|a| = b$!)

6. Suppose we flip a fair coin 100 times. Using Chebychev's inequality, estimate the probability that we see between 40 and 60 heads.

Solution: Let X be the random variable that counts the number of heads. Then $E(X) = 100/2 = 50$ and $\sigma^2(X) = 100/4 = 25$. Chebychev's inequality in this case becomes

$$P(|X - 50| > 10) \leq 25/10^2 = 1/4$$

So the probability that we have between 40 and 50 heads is $1 - 1/4 = 3/4$ by this inequality.

Note that this is actually a very weak result. If you calculate it directly in terms of the Binomial distribution you get a number like 0.965!