Math 224 **Spring 2022**
Homework 4 **Drew Armstrong**

---

1. Bilinearity of Covariance. Let X and Y be random variables on the same experiment with the following moments:

$$E[X] = 1, \quad E[X^2] = 2, \quad E[Y] = 2, \quad E[Y^2] = 6, \quad E[XY] = 5.$$

(a) Compute $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$ and $\mathrm{Cov}(X,Y)$.
(b) Use part (a) to compute $\mathrm{Cov}(2X - Y, 3X + 7Y)$.

Remark: Oops, the moments are supposed to satisfy $E[X^2] \cdot E[Y^2] \geq E[XY]^2$, which is called the *Cauchy-Schwarz inequality*. I'll fix this next time I teach the course.

(a): We use the algebraic formulas for variance and covariance to obtain

$$\begin{aligned}
\mathrm{Var}(X) &= E[X^2] - E[X]^2 = 2 - 1^2 = 1, \\
\mathrm{Var}(Y) &= E[Y^2] - E[Y]^2 = 6 - 2^2 = 2, \\
\mathrm{Cov}(X,Y) &= E[XY] - E[X] \cdot E[Y] = 5 - 1 \cdot 2 = 3.
\end{aligned}$$

(b): We use part (a) and the bilinearity of covariance to obtain

$$\begin{aligned}
\mathrm{Cov}(2X - Y, 3X + 7Y) &= \mathrm{Cov}(2X, 3X) + \mathrm{Cov}(2X, 7Y) + \mathrm{Cov}(-Y, 3X) + \mathrm{Cov}(-Y, 7Y) \\
&= 6 \cdot \mathrm{Cov}(X,X) + 14 \cdot \mathrm{Cov}(X,Y) - 3 \cdot \mathrm{Cov}(Y,X) - 7 \cdot \mathrm{Cov}(Y,Y) \\
&= 6 \cdot \mathrm{Var}(X) - 7 \cdot \mathrm{Var}(Y) + 11 \cdot \mathrm{Cov}(X,Y) \\
&= 6 \cdot 1 - 7 \cdot 2 + 11 \cdot 3 \\
&= 25.
\end{aligned}$$

2. Standardization. Let X be a random variable with $E[X] = \mu$ and $\mathrm{Var}(X) = \sigma^2$. Consider the random variable[1]

$$X' = \frac{X - \mu}{\sigma}.$$

(a) Use linearity of expectation to compute $E[X']$
(b) Use properties of variance to compute $\mathrm{Var}(X')$.

(a): We use the linearity of expectation to obtain

$$E[X'] = E\left[\frac{1}{\sigma} \cdot X - \frac{\mu}{\sigma}\right] = \frac{1}{\sigma} \cdot E[X] - \frac{\mu}{\sigma} = \frac{1}{\sigma} \cdot \mu - \frac{\mu}{\sigma} = 0.$$

(b): We use the algebraic properties of variance to obtain

$$\mathrm{Var}(X') = \mathrm{Var}\left(\frac{1}{\sigma} \cdot X - \frac{\mu}{\sigma}\right) = \left(\frac{1}{\sigma}\right)^2 \cdot \mathrm{Var}(X) + 0 = \left(\frac{1}{\sigma}\right)^2 \cdot \sigma^2 + 0 = 1.$$

Remark: Standardization is a very important trick for working with normal random variables. (See the next chapter.) If X is a normal random variable with mean μ and variance σ^2 then $Z = (X - \mu)/\sigma$ is a normal random variable with mean 0 and variance 1, called a *standard normal* random variable.

[1] This is not a derivative. I just didn't want to waste another letter of the alphabet.

3. Joint Distributions. Let X and Y be random variables on the same experiment. Suppose that X and Y have the following joint pmf table:[2]

$X \setminus Y$	0	1	3	
-1	$1/12$	$1/12$	$2/12$	$4/12$
1	$2/12$	$3/12$	$3/12$	$8/12$
	$3/12$	$4/12$	$5/12$	

(a) Compute $E[X]$ and $E[Y]$.
(b) Compute $\text{Var}(X)$ and $\text{Var}(Y)$.
(c) Compute $E[XY]$ and $\text{Cov}(X, Y)$.

(a): Using the definition of expected value gives

$$\begin{aligned} E[X] &= (-1)(4/12) + (1)(8/12) = 4/12, \\ E[Y] &= (0)(3/12) + (1)(4/12) + (3)(5/12) = 19/12. \end{aligned}$$

(b): To compute the variances we also need to know $E[X^2]$ and $E[Y^2]$, which we compute using the formula for the expected value of a function of a random variable:

$$\begin{aligned} E[X^2] &= (-1)^2(4/12) + (1)^2(8/12) = 12/12 = 1, \\ E[Y^2] &= (0)^2(3/12) + (1)^2(4/12) + (3)^2(5/12) = 49/12. \end{aligned}$$

Then we use the algebraic formula for variance to obtain

$$\begin{aligned} \text{Var}(X) &= E[X^2] - E[X]^2 = 1 - (4/12)^2 = 128/144 = 8/9, \\ \text{Var}(Y) &= E[Y^2] - E[Y]^2 = (49/12) - (19/12)^2 = 227/144. \end{aligned}$$

(c): First we compute $E[XY]$ using the formula for the expected value of a function of two random variables:

$$\begin{aligned} E[XY] &= \sum_{k,\ell} k\ell P(X = k, Y = \ell) \\ &= (-1)(0)(1/12) + (-1)(1)(1/12) + (-1)(3)(2/12) \\ &\quad + (1)(0)(2/12) + (1)(1)(3/12) + (1)(3)(3/12) \\ &= -1/12 - 6/12 + 3/12 + 9/12 \\ &= 5/12. \end{aligned}$$

Then we use the algebraic formula for covariance to obtain

$$\text{Cov}(X, Y) = E[XY] - E[X] \cdot E[Y] = \frac{5}{12} - \left(\frac{4}{12}\right)\left(\frac{19}{12}\right) = -\frac{16}{144} = -\frac{1}{9}.$$

4. Multinomial Covariance. Consider a fair 3-sided die with sides labeled $\{a, b, c\}$. Roll the die 3 times and consider the following random variables:

$$\begin{aligned} A &= \text{the number of times that } a \text{ shows up}, \\ B &= \text{the number of times that } b \text{ shows up}. \end{aligned}$$

[2]For example, the table says that $P(X = -1, Y = 3) = 2/12$ and $P(Y = 3) = 5/12$.

(a) Write out the joint pmf table of A and B. [Hint: Recall the formula

$$P(A=k, B=\ell) = \frac{3!}{k!\ell!(3-k-\ell)!}(1/3)^k(1/3)^\ell(1/3)^{3-k-\ell}.]$$

(b) Use the joint pmf table to compute $\mathrm{Cov}(A,B)$. Observe that it is negative. Indeed, if the number of a's goes up then the number of b's has a tendency to go down (and vice versa) because the total number of rolls is fixed.

(a): Every number in the joint pmf table will have denominator 27 because

$$P(A=k, B=\ell) = \frac{3!}{k!\ell!(3-k-\ell)!}(1/3)^k(1/3)^\ell(1/3)^{3-k-\ell} = \frac{3!}{k!\ell!(3-k-\ell)!} \Big/ 27.$$

Filling in the numerators give the following table:

$A \setminus B$	0	1	2	3	
0	$\frac{1}{27}$	$\frac{3}{27}$	$\frac{3}{27}$	$\frac{1}{27}$	$\frac{8}{27}$
1	$\frac{3}{27}$	$\frac{6}{27}$	$\frac{3}{27}$	0	$\frac{12}{27}$
2	$\frac{3}{216}$	$\frac{3}{216}$	0	0	$\frac{6}{27}$
3	$\frac{1}{27}$	0	0	0	$\frac{1}{27}$
	$\frac{8}{27}$	$\frac{12}{27}$	$\frac{6}{27}$	$\frac{1}{27}$	

(b): Now we use the table to compute the covariance as we did in Problem 3. First we compute $E[A]$ and $E[B]$. Note that A and B have the same expected value:

$$\begin{aligned} E[A] &= (0)(8/27) + (1)(12/27) + (2)(6/27) + (3)(1/27) = 27/27 = 1, \\ E[B] &= (0)(8/27) + (1)(12/27) + (2)(6/27) + (3)(1/27) = 27/27 = 1. \end{aligned}$$

Then we compute the mixed moment:

$$\begin{aligned} E[AB] &= (0)(0)(1/27) + (0)(1)(3/27) + (0)(2)(3/27) + (0)(3)(1/27) \\ &\quad + (1)(0)(3/27) + (1)(1)(6/27) + (1)(2)(3/27) + (1)(3)(0) \\ &\quad + (2)(0)(3/27) + (2)(1)(3/27) + (2)(2)(0) + (2)(3)(0) \\ &\quad + (3)(0)(1/27) + (3)(1)(0) + (3)(2)(0) + (3)(3)(0) \\ &= 6/27 + 6/27 + 6/27 \\ &= 18/27 \\ &= 2/3. \end{aligned}$$

Finally, we use the algebraic formula for covariance:

$$\mathrm{Cov}(A,B) = E[AB] - E[A] \cdot E[B] = \frac{2}{3} - 1^2 = -\frac{1}{3}.$$

Remark: We could have predicted this answer with a general formula. Suppose that a fair s-sided die is rolled n times. Let p_i be the probability that side i shows up and let X_i be the number of times that side i shows up. Then we have

$$\mathrm{Var}(X_i) = np_i(1-p_i) \quad \text{and} \quad \mathrm{Cov}(X_i, X_j) = -np_ip_j \text{ for } i \neq j.$$

In our case we had $s = 3$, $n = 3$, $p_1 = p_2 = p_3 = 1/3$, $A = X_1$ and $B = X_2$, hence

$$\mathrm{Cov}(A, B) = \mathrm{Cov}(X_1, X_2) = -np_1p_2 = -3(1/3)(1/3) = -1/3.$$

You do not need to memorize this general formula.

5. The Hat Check Problem. Suppose that n people go to a party and leave their hats with the hat check person.[3] At the end of the party the hat check person returns the hats randomly. Consider the following Bernoulli variables:

$$X_i = \begin{cases} 1 & \text{if the } i\text{th person gets their own hat back,} \\ 0 & \text{otherwise.} \end{cases}$$

Let $X = X_1 + \cdots + X_n$ be the total number of people who get their own hat back.

(a) Compute $E[X_i]$ and $\mathrm{Var}(X_i)$ for any i. [Hint: Compute $P(X_i = 1)$.]

(b) Use linearity to compute the expected value $E[X]$.

(c) Compute the mixed moment $E[X_iX_j]$ for $i \neq j$. [Hint: Note that

$$X_iX_j = \begin{cases} 1 & \text{if the } i\text{th and } j\text{th persons both get their own hat back,} \\ 0 & \text{otherwise.} \end{cases}$$

This implies that $P(X_iX_j = 1) = P(X_i = 1, X_j = 1) = P(X_i = 1)P(X_j = 1|X_i = 1)$.]

(d) Use parts (a) and (c) to compute the covariance $\mathrm{Cov}(X_i, X_j) = E[X_iX_i] - E[X_i]E[X_j]$ and the variance $\mathrm{Var}(X)$. [Hint: Bilinearity and symmetry of covariance gives

$$\begin{aligned} \mathrm{Var}(X) &= \mathrm{Cov}(X_1 + \cdots + X_n, X_1 + \cdots + X_n) \\ &= \sum_{i,j} \mathrm{Cov}(X_i, X_j) \\ &= \sum_i \mathrm{Cov}(X_i, X_i) + \sum_{i \neq j} \mathrm{Cov}(X_i, X_j) \\ &= \sum_i \mathrm{Cov}(X_i, X_i) + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j) \\ &= \sum_i \mathrm{Var}(X_i) + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j). \end{aligned}$$

The number of pairs in the second sum is $\binom{n}{2} = n(n-1)/2$.]

(a): Intuition suggests that any individual person has a $1/n$ chance of getting their own hat back. In other words, we have $P(X_i = 1) = 1/n$ for any i.[4] Then since X_i is Bernoulli with $P(X_i = 1) = 1/n$ and $P(X_i = 0) = (n-1)/n)$ we have

$$\begin{aligned} E[X_i] &= 0 \cdot P(X_i = 0) + 1 \cdot P(X_i = 1) = 0 \cdot (n-1)/n + 1 \cdot 1/n = 1/n, \\ E[X_i^2] &= 0^2 \cdot P(X_i = 0) + 1^2 \cdot P(X_i = 1) = 0^2 \cdot (n-1)/n + 1^2 \cdot 1/n = 1/n, \\ \mathrm{Var}(X_i) &= E[X_i^2] - E[X_i]^2 = (1/n) - (1/n)^2 = (n-1)/n^2. \end{aligned}$$

---

[3] Long ago people used to wear hats, but not indoors.

[4] We can make this more precise as follows. There are $n!$ ways to return all the hats, and we assume that each of these ways is equally likely. If person i gets their own hat back then there are $(n-1)!$ ways to return the hats to the other $n-1$ people. Hence

$$P(X_i = 1) = P(\text{person } i \text{ gets their own hat back}) = (n-1)!/n! = 1/n.$$

(b): Let $X = X_1 + X_2 + \cdots + X_n$ be the total number of people who get their own hat back. Then part (a) and linearity of expectation gives

$$\begin{aligned} E[X] &= E[X_1] + E[X_2] + \cdots + E[X_n] \\ &= \underbrace{1/n + 1/n + \cdots + 1/n}_{n \text{ times}} = n(1/n) = 1. \end{aligned}$$

In other words, if the hats are returned randomly then, on average, exactly one person will get their own hat back. And this is independent of the number n. That's interesting.

Remark: We could have tried to compute the expectation directly from the definition:

$$E[X] = \sum_k k \cdot P(X = k).$$

But in this case the probabilities $P(X = k)$ are very difficult to compute. For example, I claim that the probability that no one gets their own hat back is

$$P(X = 0) = \frac{1}{0!} - \frac{1}{1!} + \frac{1}{2!} - \cdots + (-1)^n \frac{1}{n!}.$$

Once again, this illustrates the usefulness of the linearity of expectation.

Parts (c) and (d) are tricky and you do not need to know them for the exam.

(c): Since X_i and X_j have values 0 or 1 then the product X_iX_j also has value 0 or 1. In particular we have $X_iX_j = 1$ precisely when **both** $X_i = 1$ and $X_j = 1$. In other words,

$$X_iX_j = \begin{cases} 1 & \text{if the } i\text{th and } j\text{th persons } \textbf{both} \text{ get their own hat back,} \\ 0 & \text{otherwise.} \end{cases}$$

This allows us to compute the probability $P(X_iX_j = 1)$ as follows:

$$\begin{aligned} P(X_iX_j = 1) &= P(X_i = 1 \text{ and } X_j = 1) \\ &= P(X_i = 1)P(X_j = 1 \,|\, X_i = 1). \end{aligned}$$

The last step comes from the definition of conditional probability. From part (a) we know that $P(X_i = 1) = 1/n$. And the conditional probability can also be computed by intuition. After the ith person gets their own hat, there are $n-1$ remaining hats, so there is a $1/(n-1)$ that the jth person gets their own hat:

$$P(X_j = 1 \,|\, X_i = 1) = 1/(n-1).$$

Putting these together gives

$$P(X_iX_j = 1) = \frac{1}{n} \cdot \frac{1}{n-1},$$

and hence the expected value is

$$E[X_iX_j] = 0 \cdot P(X_iX_j = 0) + 1 \cdot P(X_iX_j = 1) = P(X_iX_j = 1) = \frac{1}{n(n-1)}.$$

(c): It follows from parts (a) and (c) that

$$\text{Cov}(X_i, X_j) = E[X_iX_j] - E[X_i] \cdot E[X_j] = \frac{1}{n(n-1)} - \left(\frac{1}{n}\right)^2 = \frac{1}{n^2(n-1)}.$$

We will use this fact and the formula $\mathrm{Var}(X_i) = (n-1)/n^2$ from part (a) to compute $\mathrm{Var}(X)$. As in the hint, we can apply the bilinearity of covariance to obtain

$$\begin{aligned}
\mathrm{Var}(X) &= \mathrm{Cov}(X_1 + \cdots + X_n, X_1 + \cdots + X_n) \\
&= \sum_{i,j} \mathrm{Cov}(X_i, X_j) \\
&= \sum_{i} \mathrm{Cov}(X_i, X_i) + \sum_{i \neq j} \mathrm{Cov}(X_i, X_j) \\
&= \sum_{i} \mathrm{Cov}(X_i, X_i) + 2\sum_{i < j} \mathrm{Cov}(X_i, X_j) \\
&= \sum_{i} \mathrm{Var}(X_i) + 2\sum_{i < j} \mathrm{Cov}(X_i, X_j) \\
&= \sum_{i} \frac{n-1}{n^2} + 2\sum_{i<j} \frac{1}{n^2(n-1)}.
\end{aligned}$$

Since the summands don't depend on i or j we only need to determine how many summands there are.[5] The sum over i has n terms and the sum over $i < j$ has $\binom{n}{2} = n(n-1)/2$ terms because there are $\binom{n}{2}$ ways to choose the numbers i and j from the index set $\{1, 2, \ldots, n\}$. It follows that

$$\begin{aligned}
\mathrm{Var}(X) &= \sum_{i} \frac{n-1}{n^2} + 2\sum_{i<j} \frac{1}{n^2(n-1)} \\
&= n \cdot \frac{n-1}{n^2} + 2 \cdot \binom{n}{2} \cdot \frac{1}{n^2(n-1)} \\
&= n \cdot \frac{n-1}{n^2} + 2 \cdot \frac{n(n-1)}{2} \cdot \frac{1}{n^2(n-1)} \\
&= \frac{n-1}{n} + \frac{1}{n} \\
&= 1.
\end{aligned}$$

That's a surprise.

[5]For example, if a is constant then we have $\sum_{i=1}^{n} a = na$ and $\sum_{1 \leq i < j \leq n} a = [n(n-1)/2]a$.