Notes and Sample Questions
Chapter 10

1 Useful equations for Chapter 10

If we are given the joint distribution of two random variables X,Y, we are able to calculate the expectation of any function $g(X,Y)$ w.r.t X,Y:

Definition: $E[g(X,Y)] = \begin{cases} \sum_x \sum_y g(x,y)P(X=x,Y=y), & X \text{ is discrete} \\ \int \int_x yg(x,y)f(x,y)dx, & X \text{is continuous} \end{cases}$

Definition: $Var[g(X,Y)] = E[(g(X,Y) - E[g(X,Y)])^2] = E[g(X,Y)^2] - E[g(X,Y)]^2$

Definition: $Cov(X,Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y], \rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var[X]Var[Y]}}$

Definition: $Cov(g(X), g(Y)) = E[g(X)g(Y)] - E[g(X)]E[g(Y)]$

Lemma: $Cov(rX+s, tY+u) = rtCov(X,Y); \rho(rX+s, tY+u) = \begin{cases} -\rho(X,Y), if\ rt < 0 \\ \rho(X,Y), if\ rt > 0 \end{cases}$

Puzzle: Why does joint distribution exist? (Weight, Height can also be measured jointly.)

2 Sample Questions

10.1 Random variables X and Y have the following joint probabilities:

$$P(X=0,Y=0) = P(X=0,Y=8) = P(X=16,Y=0) = P(X=16,Y=16) = \frac{1}{4}$$

a. Make a table of the joint distribution of X and Y and add their marginal distributions to the table.

	b			
a	0	8	16	P(X=a)
0	$\frac{1}{4}$	$\frac{1}{4}$	0	$\frac{1}{2}$
16	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{2}$
P(Y=b)	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	1

b. Compute $Cov(X,Y)$. Are X and Y positively correlated, negative corrected, or uncorrelated ?
Solution 1

Definition Let X and Y be two random variables. The *covariance* between X and Y is defined by

$$Cov(X,Y) = E\left[(X - E[X])\left(Y - E[Y]\right)\right]$$

First,

$$E[X] = 0 \times \frac{1}{2} + 16 \times \frac{1}{2} = 8$$
$$E[Y] = 0 \times \frac{1}{2} + 8 \times \frac{1}{4} + 16 \times \frac{1}{4} = 6$$

Let $X' = X - E[X], Y' = Y - E[Y]$, then, we construct a table of the joint distribution of X' and Y' here,

	b			
a	-6=0-6	2=8-6	10=16-6	P(X'=a)
-8=0-8	$\frac{1}{4}$	$\frac{1}{4}$	0	$\frac{1}{2}$
8=16-8	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{2}$
P(Y'=b)	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	1

From the definition, we know that $Cov(X,Y) = E\left[(X - E[X])\left(Y - E[Y]\right)\right]$. By replacing $X - E[X]$ with X' and $Y - E[Y]$ with Y', we obtain $Cov(X,Y) = E[X'Y']$.

TWO-DIMENSIONAL CHANGE-OF-VARIABLE FORMULA Let X and Y be random variables, and let $g : \mathbb{R}^2 \to \mathbb{R}$ be a function. If X and Y are *discrete* random variables with values $a_1, a_2, \ldots$ and $b_1, b_2, \ldots$ respectively, then

$$E\left[g\left(X,Y\right)\right] = \sum_i \sum_j g(a_i, b_j) P(X = a_i, Y = b_j).$$

If X and Y are *continuous* random variables with joint probability density function f, then

$$E\left[g\left(X,Y\right)\right] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f(x,y) dx dy.$$

In order to compute the $E[X'Y']$, we can use the two-dimensional change-of-variable formula by setting $g(X', Y') = X'Y'$. Thus,

$$\begin{aligned}
E\left[X'Y'\right] &= \sum_{a \in \{-8,8\}} \sum_{b \in \{-6,2,10\}} g(a,b) P(X' = a, Y' = b) \\
&= \sum_{a \in \{-8,8\}} \sum_{b \in \{-6,2,10\}} ab P(X' = a, Y' = b) \\
&= (-8)(-6)(\frac{1}{4}) + (-8)(2)(\frac{1}{4}) + (8)(-6)(\frac{1}{4}) + (8)(10)(\frac{1}{4}) \\
&= 12 + (-4) + (-12) + (20) \\
&= 16
\end{aligned}$$

Thus, $Cov(X,Y) = 16$.

Solution 2:

AN ALTERNATIVE EXPRESSION FOR THE COVARIANCE. Let X and Y be two random variables, then

$$Cov(X,Y) = E[XY] - E[X]E[Y].$$

From solution 1, we know that $E[X] = 8, E[Y] = 6$, we then compute $E[XY]$. We can compute it by using the two-dimensional change-of-variable formula by setting $g(x,y) = xy$. Then,

$$E[XY] = \sum a \in \{0,16\} \sum_{b \in \{0,8,16\}} g(a,b) P(X = a, Y = b) \tag{1}$$
$$= (0)(0)(\frac{1}{4}) + (0)(8)(\frac{1}{4}) + (16)(0)(\frac{1}{4}) + (16)(16)(\frac{1}{4}) \tag{2}$$
$$= 64 \tag{3}$$

Thus,

$$\begin{aligned} Cov(X,Y) &= E[XY] - E[X]E[Y] \\ &= 64 - (8)(6) \\ &= 16 \end{aligned}$$

If $Cov(X,Y) > 0$, X and Y are positively correlated.
If $Cov(X,Y) < 0$, X and Y are negatively correlated.
If $Cov(X,Y) = 0$, X and Y are uncorrelated.
Since $Cov(X,Y) = 16 > 0$, then X and Y are positively correlated.

c. Compute the correlation coefficient between X and Y. **Definition** Let X and Y be two random variable. The *correlation coefficient* $\rho(X,Y)$ is defined to be 0 if $Var(X) = 0$ or $Var(Y) = 0$, and otherwise

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}.$$

We compute $Var(X) = E[X^2] - (E[X])^2$, while

$$\begin{aligned} E[X^2] &= (0^2)(\frac{1}{2}) + (16^2)(\frac{1}{2}) = 128 \\ E[X] &= 8 \\ Var(X) &= E[X^2] - (E[X])^2 = 128 - 8^2 = 64 \end{aligned}$$

Similarly, we compute $Var(Y) = E[Y^2] - (E[Y])^2$ such that

$$\begin{aligned} E[Y^2] &= (0^2)(\frac{1}{2}) + (8^2)(\frac{1}{4}) + (16^2)(\frac{1}{4}) = 0 + 16 + 64 = 80 \\ E[Y] &= 6 \\ Var(Y) &= E[Y^2] - (E[Y])^2 = 80 - 6^2 = 44 \end{aligned}$$

We know that $Cov(X,Y) = 16$. Thus,

$$\begin{aligned} \rho(X,Y) &= \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} \\ &= \frac{16}{\sqrt{(64)(44)}} \\ &= \frac{\sqrt{11}}{11} \\ &\approx 0.3015 \end{aligned}$$

10.5 Suppose X and Y are discrete random variables taking values 0, 1, and 2. The following is given about the joint and marginal distributions:

	a			
b	0	1	2	P(Y=b)
0	$\frac{8}{72}$	...	$\frac{10}{72}$	$\frac{1}{3}$
1	$\frac{12}{72}$	$\frac{9}{72}$	...	$\frac{1}{2}$
2	...	$\frac{3}{72}$	...	...
P(X=a)	$\frac{1}{3}$	...	...	1

a. Complete the table.

$$
\begin{aligned}
P(X=1,Y=0) &= P(Y=0) - P(X=0,Y=0) - P(X=2,Y=0) = \frac{1}{3} - \frac{8}{72} - \frac{10}{72} \\
&= \frac{6}{72} \\
P(X=0,Y=2) &= P(X=0) - P(X=0,Y=0) - P(X=0,Y=1) = \frac{1}{3} - \frac{8}{72} - \frac{12}{72} \\
&= \frac{4}{72} \\
P(X=2,Y=1) &= P(Y=1) - P(X=0,Y=1) - P(X=1,Y=1) = \frac{1}{2} - frac1272 - \frac{9}{72} \\
&= \frac{15}{72} \\
P(X=1) &= P(X=1,Y=0) + P(X=1,Y=1) + P(X=1,Y=2) = \frac{6}{72} + \frac{9}{72} + \frac{3}{72} \\
&= \frac{1}{4} \\
P(Y=2) &= 1 - P(Y=0) - P(Y=1) = 1 - \frac{1}{3} - \frac{1}{2} \\
&= \frac{1}{6} \\
P(X=2,Y=2) &= P(Y=2) - P(X=0,Y=2) - P(X=1,Y=2) = \frac{1}{6} - \frac{4}{72} - \frac{3}{72} \\
&= \frac{5}{72} \\
P(X=2) &= P(X=2,Y=0) + P(X=2,Y=1) + P(X=2,Y=2) = \frac{10}{72} + \frac{15}{72} + \frac{5}{72} \\
&= \frac{5}{12}
\end{aligned}
$$

Thus, the completed table is

	a			
b	0	1	2	P(Y=b)
0	$\frac{8}{72}$	$\frac{6}{72}$	$\frac{10}{72}$	$\frac{1}{3}$
1	$\frac{12}{72}$	$\frac{9}{72}$	$\frac{15}{72}$	$\frac{1}{2}$
2	$\frac{4}{72}$	$\frac{3}{72}$	$\frac{5}{72}$	$\frac{1}{6}$
P(X=a)	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{5}{12}$	1

b. Compute the expectation of X and of Y and the covariance between X and Y.

$$\begin{aligned} E[X] &= \sum_{a \in \{0,1,2\}} aP(X=a) \\ &= (0)(\frac{1}{3}) + (1)(\frac{1}{4}) + (2)(\frac{5}{12}) \\ &= \frac{13}{12} \\ E[Y] &= \sum_{b \in \{0,1,2\}} bP(Y=b) \\ &= (0)(\frac{1}{3}) + (1)(\frac{1}{2}) + (2)(\frac{1}{6}) \\ &= \frac{5}{6} \end{aligned}$$

In order to compute the $Cov(X,Y)$, we can use $Cov(X,Y) = E[XY] - E[X]E[Y]$. We then have to compute $E[XY]$. We can use the two-dimensional change-of-variable formula by setting $g(x,y) = xy$. Thus,

$$\begin{aligned} E[XY] &= \sum_{a \in \{0,1,2\}} \sum_{b \in \{0,1,2\}} g(a,b)P(X=a, Y=b) \\ &= (1)(1)(\frac{9}{72}) + (1)(2)(\frac{15}{72}) + (2)(1)(\frac{3}{72}) + (2)(2)(\frac{5}{72}) \\ &= \frac{65}{72} \end{aligned}$$

Thus,

$$\begin{aligned} Cov(X,Y) &= E[XY] - E[X]E[Y] = \frac{65}{72} - \frac{13}{12}\frac{5}{6} \\ &= 0 \end{aligned}$$

c. Are X and Y are independent ?

Question Here ? Can we say that X and Y are independent because X and Y are uncorrelated ? **NO!!!**

INDEPENDENCE VERSUS UNCORRELATED. If two random variable X and Y are independent, then X and Y are uncorrelated.
Note that the reverse is not necessarily true. If X and Y are uncorrelated, they need not be independent.

Remember the definition: **DEFINITION** An event A is called *independent* of B if $P(A|B) = P(A)$.
INDEPENDENCE. To show that A and B are independent it suffice to prove *just one* of the following:

$$\begin{aligned} P(A|B) &= P(A) \\ P(B|A) &= P(B) \\ P(A \cap B) &= P(A)P(B), \end{aligned}$$

where A may be replaced by A^c and B replaced by B^c, or both. IF one of these statements holds, *all* of them are true. If two events are not independent, they are called *dependent*.

To show Independent If you want to show that A and B are independent, you can justified it by mathematically showing that one of those statements holds. Or, you can check one of those statements. It holds for *all* possible outcomes of A and *all* possible outcomes of B. Then you can say that A and B are independent.

To show Dependent You can just show for one outcome of A (e.g., a) and one outcome of B (e.g., b) that one of those statements doesn't hold. **Anyone of the pairs with specific value of a and specific value of b**.

For this question, these two variables are actually independent. Yes, X and Y are independent. We can check for any a and b, $P(X = a, Y = b) = P(X = a)P(Y = b)$.