Exercises Lecture 6: Joint distributions

1. We will consider the example of Lecture 3. There, 150 patients with episodes of depression were randomized to 4 treatment groups: Placebo and treatments I, L and C. Relapse may happen within two to three years.

	Treatment group (Y)			
Response (X)	Imipramine (1)	Lithium (2)	Combination (3)	Placebo (4)
Relapse (0)	0.120	0.087	0.146	0.160
No relapse (1)	0.147	0.166	0.107	0.067

Suppose that a patient is selected at random from the 150 patients in that study and we record Y, an indicator of the treatment group for that patient, and X, an indicator of whether or not the patient relapsed. The Table above contains the joint pmf of X and Y.

(a) Calculate the probability that a patient selected at random from this study used either treatment (2) or treatment (3) and did not relapse.

(b) Calculate the probability that the patient had a relapse (without regard to the treatment group).

<u>Solution:</u>

(a) Let $f(x, y)$, the joint pmf given on the Table above. $P(\{Y \in \{2, 3\}\} \cap \{X = 1\}) = f(1, 2) + f(1, 3) = 0.166 + 0.107$.

(b) $P(X = 0) = f(0, 1) + f(0, 2) + f(0, 3) + f(0, 4) = 0.513$.

2. The joint pmf of two discrete random variables X and Y is given in the following table:

	x			
y	1	2	3	4
1	0.1	0.05	0.02	0.02
2	0.05	0.20	0.05	0.02
3	0.02	0.05	0.20	0.04
4	0.02	0.02	0.04	0.10

(*a*) Find the marginal pmf of X and Y.

(*b*) $P(X < Y)$.

Solution:

(*a*) They can be found on the table:

		x			
y	1	2	3	4	$f(Y)$
1	0.1	0.05	0.02	0.02	0.19
2	0.05	0.20	0.05	0.02	0.32
3	0.02	0.05	0.20	0.04	0.31
4	0.02	0.02	0.04	0.10	0.18
$f(X)$	0.19	0.32	0.31	0.18	1

(*b*) Sum thus probabilities in the table for which $x < y$. $P(X < Y) = 0.05 + 0.02 + 0.02 + 0.05 + 0.02 + 0.04 = 0.2$.

3. Two random variables are independent and have the following pmf:

j	1	3	5	7	9
$P_X(j)$	0.1	0.2	0.4	0.2	0.1

k	2	4	6	8
$P_Y(k)$	0.1	0.2	0.3	0.4

(*a*) Find $P(X = 3, Y = 6)$.

(*b*) $P(X \leq 3, Y \leq 6)$.

Solution:

(*a*) $P(X = 3, Y = 6) = P(X = 3) \cdot P(Y = 6) = 0.2 \cdot 0.3 = 0.06$.

(*b*) $P(X \leq 3, Y \leq 6) = P(X \leq 3) \cdot P(Y \leq 6) = (0.1 + 0.2) \cdot (0.1 + 0.2 + 0.3) = 0.18$.

4. Consider two random variables Y and X. In the following table a part of their joint probabilities is given:

	x			
y	0	1	2	
0	0.03	0.15	?	?
1	0.04	?	?	?
2	0.03	?	?	?
	?	?	?	?

Assume that X and Y are independent. Find the missing probabilities i.e. joint and marginal pmfs.

Solution:

	x			
y	0	1	2	$f(Y)$
0	0.03	0.15	0.12	0.3
1	0.04	0.2	0.16	0.4
2	0.03	0.15	0.12	0.3
$f(X)$	0.1	0.5	0.4	1

5. Let us assume that the random variable X follows the uniform distribution on the interval $[0, 1]$, the random variable Y follows the uniform distribution on the interval $[5, 9]$, and that X and Y are independent. Let also that a rectangle is to be constructed for which the lengths of two adjacent sides are X and Y. What is the expected value of the area of the rectangle?
Hint: The mean of a Uniform random variable in $[a, b]$ is $\frac{a+b}{2}$ and the area of a rectangle is computed by length times width.

   Solution:

   The area is computed by $X \cdot Y$. Since X and Y are independent $E(XY) = E(X) \cdot E(Y)$. We have $E(X) = 1/2$ and $E(Y) = 7$, Thus $E(XY) = 7/2$.

6. Let X and Y be independent random variables with equal variances. What is the value of $Cov(X + Y, X - Y)$?

   Solution:
   We use use the properties of the covariance: $Cov(X+Y, X-Y) = Cov(X, X)+Cov(Y, X)-Cov(Y, X) - Cov(Y, Y) = Var(X) - Var(Y) = 0$.

7. The joint distribution of the number of B alleles that two siblings carry is given below:

	0	1	2
0	0.3	0.1	0.01
1	0.1	0.3	0.06
2	0.01	0.05	0.07

What is the value of the covariance of X and Y?

Solution:

We have $Cov(X, Y) = E(X \cdot Y) - E(X) \cdot E(Y)$. So we need to compute each term separately.

$$\begin{aligned} E(X \cdot Y) &= \sum_{x=0}^{2} \sum_{y=0}^{2} x \cdot y \cdot p_{XY}(x, y) \\ &= 0 \cdot 0 \cdot p_{X,Y}(0,0) + 0 \cdot 1 \cdot p_{X,Y}(0,1) + 0 \cdot 2 \cdot p_{X,Y}(0,2) + \ldots \\ &= 0 \cdot 0 \cdot 0.3 + 0 \cdot 1 \cdot 0.1 + 0 \cdot 2 \cdot 0.01 + \ldots = 0.8 \end{aligned}$$

To compute: $E(X)$ and $E(Y)$ we need first the marginals pmf $p_X(x)$ and $p_Y(y)$.

$p_X(x) = \sum_{y=0}^{2} p_{XY}(x, y)$.

So $p_X(X = 0) = 0.3 + 0.1 + 0.01 = 0.41$.

Then $p_X(X = 1) = 0.1 + 0.3 + 0.05 = 0.45$.

And finally $p_X(X = 2) = 0.01 + 0.06 + 0.07 = 0.14$.

$E(X) = \sum_{x=0}^{2} x p_X(x) = 0 \cdot 0.41 + 1 \cdot 0.45 + 2 \cdot 0.14 = 0.73$.

Similarly you may compute $E(Y) = 0.72$.
You may also use the R code in the slides of this lecture.

```
PXY <- rbind(c(0.3, 0.1, 0.01), c(0.1, 0.3, 0.05), c(0.01, 0.06, 0.07))
EXY <- PXY[2, 2] + 2 * PXY[2, 3] + 2 * PXY[3, 2] + 4 * PXY[3, 3]
EXY
[1] 0.8
EX <- apply(PXY, 1, sum) %*% 0:2
EY <- apply(PXY, 2, sum) %*% 0:2
COVXY <- EXY - EX * EY
COVXY
    [,1]
[1,] 0.2744
```