MTH 224, Spring 2024

Instructor: Bahman Angoshtari

Lecture 11

Section 4.1: discrete joint distribution, expectation of a function of multiple r.v.

11.1. Poisson distribution (continued)

- Recall that $X \sim \text{Pois}(\lambda)$ if $p_X(k) = \frac{\lambda^k}{k!}e^{-\lambda}$, $k = 0, 1, \ldots$.
- The Poisson distribution is obtained as the limit of the $\text{Bin}(n, p)$ distribution as n becomes large, p becomes small, and $np = \lambda$ is constant. It is also the number of events in a fixed interval if events occur according to the Poisson process.

Example 11.1. Births of twins in a certain city are described by a Poisson process with the constant rate of 1.2 births per year.

(a) What is the probability that more than two twin births will occur during the year 2021?

(b) What is the probability that no twin births will occur during the next five years?

(c) If we learn that there was at least one birth of twins during the year 2020, what is the conditional probability that there were no twin births during the first half of that year?

Solution.
(a) Let X be the number of twin births during 2021. Then $X \sim \text{Pois}(1.2)$, and hence

$$\mathbb{P}(X > 2) = 1 - \mathbb{P}(X \le 2) = 1 - e^{-1.2}\left(1 + \frac{1.2^1}{1} + \frac{1.2^2}{2}\right) \approx 0.12.$$

(b) Let Y be the number of twin births in the next 5 years. Then $Y \sim \text{Pois}(6)$, and hence

$$\mathbb{P}(Y = 0) = e^{-6} \approx 0.0025.$$

(c) Let $N_{[0,\frac{1}{2}]}, N_{[\frac{1}{2},1]}, N_{[0,1]}$ respectively be the number of twin births during the 1^{st} half, of 2020, 2^{nd} half of 2020, and the entire year of 2020. Then

$$\begin{aligned}\mathbb{P}\left(N_{[0,\frac{1}{2}]} = 0 \,|\, N_{[0,1]} \ge 1\right) &= \frac{\mathbb{P}\left(N_{[0,\frac{1}{2}]} = 0 \text{ and } N_{[0,1]} \ge 1\right)}{\mathbb{P}\left(N_{[0,1]} \ge 1\right)} = \frac{\mathbb{P}\left(N_{[0,\frac{1}{2}]} = 0 \text{ and } N_{[1,\frac{1}{2}]} \ge 1\right)}{\mathbb{P}\left(N_{[0,1]} \ge 1\right)} \\ &= \frac{\mathbb{P}\left(N_{[0,\frac{1}{2}]} = 0\right) \cdot \mathbb{P}\left(N_{[1,\frac{1}{2}]} \ge 1\right)}{\mathbb{P}\left(N_{[0,1]} \ge 1\right)} = \frac{e^{-\frac{1.2}{2}}\left(1 - e^{-\frac{1.2}{2}}\right)}{1 - e^{-1.2}} \approx 0.3543.\end{aligned}$$

11.2. Discrete joint distributions

- So far, we only considered one r.v. at a time. On many occasions, we encounter two or more interconnected random variables.

$$\omega \mapsto X_1(\omega), X_2(\omega), \ldots, X_N(\omega)$$

- We can interpret this scenario as having N different random variables. We can also see it as having one vector-valued random variable

$$\mathbf{X} : \Omega \to \mathbb{R}^N, \qquad \mathbf{X}(\omega) = \Big(X_1(\omega), \ldots, X_N(\omega)\Big).$$

- Examples of jointly distributed r.v.:
  - X = number of children in a family, Y = monthly income of a family.
  - X = person's height, Y = person's weight.
  - X = temperature, Y = precipitation .

Definition 11.2. Let $X_1, \ldots, X_N$ be discrete r.v.s. The joint pmf of $X_1, \ldots, X_N$ is

$$p_{X_1,\ldots,X_N}(x_1, \ldots, x_N) = \mathbb{P}(X_1 = x_1 \cap \cdots \cap X_N = x_N).$$

The individual pmf's $p_{X_1}(x) = \mathbb{P}(X_1 = x)$, ..., $p_{X_N}(x) = \mathbb{P}(X_N = x)$ are called the marginal pmfs.

We also define the joint cdf:

$$F_{X_1,\ldots,X_N}(x_1, \ldots, x_N) = \mathbb{P}(X_1 \leq x_1 \cap \cdots \cap X_N \leq x_N). = \sum_{y_1 \leq x_1} \cdots \sum_{y_N \leq x_N} p_{X_1,\ldots,X_N}(y_1, \ldots, y_N).$$

- We can always compute the marginal pmfs from the joint distribution:

$$p_X(x) = \sum_y \mathbb{P}(X = x \cap Y = y) = \sum_y p_{X,Y}(x, y)$$

$$p_Y(y) = \sum_x \mathbb{P}(X = x \cap Y = y) = \sum_x p_{X,Y}(x, y).$$

  In general, to find the marginal pmf of a r.v., you sum the joint pmf over all the values of the remaining r.v.:

$$\begin{aligned} p_{X_1}(x) &= \sum_{x_2, x_3, \ldots, x_N} \cdots \sum \mathbb{P}(X_1 = x \cap X_2 = x_2 \cap \cdots \cap X_N = x_N) \\ &= \sum_{x_2, x_3, \ldots, x_N} \cdots \sum p_{X_1,\ldots,X_N}(x, x_2, \ldots, x_N) \end{aligned}$$

- The reverse is not always possible. That is, it may be impossible to retrieve the joint distribution from only its marginals without making additional assumptions. One such assumption is independence of r.v. which we will learn soon.

Example 11.3. We roll a pair of dice. Let X be the largest number rolled, and Y the smallest number. For example, if we roll the numbers 3 and 4 then $X = 4$ and $Y = 3$. Compute the joint and marginal distributions.

Solution. We first compute $p_{X,Y}(x, y) = \mathbb{P}(X = x, Y = y)$. We split the solution into three cases:
Case 1: $x = y$. In this case

$$p_{X,Y}(1, 1) = p_{X,Y}(2, 2) = \cdots = p_{X,Y}(6, 6) = \frac{1}{36}.$$

Case 2: $x > y$. In this case

$$p_{X,Y}(x, y) = \mathbb{P}(x \text{ and } y\text{, or } y \text{ and } x\text{, are rolled}) = \frac{1}{36} + \frac{1}{36} = \frac{1}{18}.$$

Case 3: $x < y$. Clearly, in this case, $p_{X,Y}(x, y) = 0$.

In a table:

	$X = 1$	$X = 2$	$X = 3$	$X = 4$	$X = 5$	$X = 6$	$p_Y(y)$
$Y = 1$	1/36	1/18	1/18	1/18	1/18	1/18	11/36
$Y = 2$	0	1/36	1/18	1/18	1/18	1/18	9/36
$Y = 3$	0	0	1/36	1/18	1/18	1/18	7/36
$Y = 4$	0	0	0	1/36	1/18	1/18	5/36
$Y = 5$	0	0	0	0	1/36	1/18	3/36
$Y = 6$	0	0	0	0	0	1/36	1/36
$p_X(x)$	1/36	3/36	5/36	7/36	9/36	11/36	

To find the marginal pmf of one r.v., we sum over all the values of the other r.v. For example, to find $p_X(2)$, we sum the second column. Similarly, we sum over the second row to find $p_Y(2)$.

11.3. Expected value of a function of two or more random variables

- Recall that for a discrete r.v. we have defined $\mathbb{E}[g(X)] = \sum_k g(x_k) p_X(x_k)$
- Now, we define the expected value for a function of two or more jointly distributed random variables.

Definition 11.4. Let $X_1, \dots, X_N$ be jointly distributed random variables with a joint pmf $p_{X_1,\dots,X_N}(x_1, \dots, x_N)$. For a multivariable function $g(x_1, \dots, x_N)$,we define:

$$\mathbb{E}[g(X_1, \dots, X_N)] = \sum_{x_1,\dots,x_N} \cdots \sum g(x_1, \dots, x_N) p_{X_1,\dots,X_N}(x_1, \dots, x_N).$$

Example 11.5. Let X and Y be two jointly distributed r.v. and a and b be two constants. Show that $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$.

Solution.

$$\begin{aligned}
\mathbb{E}(aX + bY) &= \sum_x \sum_y (ax + by) p_{X,Y}(x, y) \\
&= a \sum_x \sum_y x p_{X,Y}(x, y) + b \sum_x \sum_y y p_{X,Y}(x, y) \\
&= a \sum_x x \sum_y p_{X,Y}(x, y) + b \sum_y y \sum_x p_{X,Y}(x, y) \\
&= a \sum_x x p_X(x) + b \sum_y y p_Y(y) \\
&= a\mathbb{E}[X] + b\mathbb{E}[Y].
\end{aligned}$$

- By using a similar argument, we can prove the following theorem. $\mathbb{E}\left[b + \sum_{i=1}^N a_i X_i\right] = b + \sum_{i=1}^N a_i \mathbb{E}[X_i]$.