Lecture 26: Multiple Random Variables

11 March, 2019

Sunil Kumar Gauttam

Department of Mathematics, LNMIIT

Probabilistic models often involve several random variables. For example, in a medical diagnosis context, the results of several tests may be significant (These different tests would be observations on different random variables.), or in a networking context, the workloads of several routers may be of interest (These different workloads would be observations on different random variables, one for each router measured). Thus, we need to know how to describe and use probability models that deals with more than one random variable at a time.

All of these random variables are associated with the same experiment, sample space, and probability law, and their values may relate in interesting ways. This motivates us to consider simultaneously several random variables, i.e., the notion of random vector.

Definition 26.1 *Let $(\Omega, \mathcal{F}, P)$ be a probability space. A map $X : \Omega \to \mathbb{R}^n, X(\omega) = (X_1(\omega), X_2(\omega), \cdots, X_n(\omega))$ is called a n-dimensional random vector on $(\Omega, \mathcal{F}, P)$ if each X_i is a random variable on $(\Omega, \mathcal{F}, P)$.*

We will discuss mainly bivariate random vectors, i.e., two random variables. The extension to higher dimension is analogue.

Example 26.2 *Consider the experiment of tossing two fair dice. Then we know sample space would be $\Omega = \{(i,j) : i, j = 1, \cdots, 6\}$ with all outcomes be equally likely and σ-algebra is power set. Let us consider the following random variable*

$$X((i,j)) = i + j, Y((i,j)) = |i - j|$$

Then (X, Y) is a random vector.

Having define a random vector (X, Y), we can now discuss probabilities of events that are defined in terms of (X, Y). For example we can ask what is $P(X = 5 \text{ and } Y = 3)$?, Henceforth, we will write $P(X = 5, Y = 3)$ for $P(X = 5 \text{ and } Y = 3)$. Read the comma as "and". It is easy to see that there are only two sample points $(1, 4)$ and $(4, 1)$ that yields $X = 5$ and $Y = 3$. Please note that we are interested in those sample points (i, j) such that $X((i,j)) = 5, Y((i,j)) = 3$ simultaneously. Hence

$$P(X = 5, Y = 3) = \frac{2}{36} = \frac{1}{18}$$

Definition 26.3 (Discrete random vector) *We say that a random vector $X = (X_1, X_2)$ is a discrete random vector, if X_1 and X_2 both are discrete random variables.*

It follows from the definition that range of a discrete random vector is either finite or countable infinite.

1. If range of X_1 and X_2 both are finite, then X has finite range.
2. If range of X_1 is finite and X_2 is countable infinite, then X has countable infinite range. Similarly, If range of X_2 is finite and X_1 is countable infinite, then X has countable infinite range.
3. If range of X_1 and X_2 both are countable infinite, then X has countable infinite range.

Definition 26.4 (joint pmf) *Let $X = (X_1, X_2)$ be a discrete random vector. Define $f : \mathbb{R}^2 \to \mathbb{R}$ by*

$$f(x_1, x_2) = P\{X_1 = x_1, X_2 = x_2\}, \forall\, x_1, x_2 \in \mathbb{R}$$

Then f is called joint pmf of X.

If it is necessary to stress the fact that f is the joint pmf of (X_1, X_2) rather than some other random vector, the notation $f_{X_1,X_2}(x_1, x_2)$ will be used.

The joint pmf of (X, Y) completely determines the probability distribution of the discrete random vector (X, Y), just as the pmf of a discrete random variable completely defines its distribution.

For the discrete random vector defined in Example 26.2, there are 21 possible values of (X, Y).

Y \ X	2	3	4	5	6	7	8	9	10	11	12
0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$
1	0	$\frac{1}{18}$	0	$\frac{1}{18}$	0	$\frac{1}{18}$	0	$\frac{1}{18}$	0	$\frac{1}{18}$	0
2	0	0	$\frac{1}{18}$	0	$\frac{1}{18}$	0	$\frac{1}{18}$	0	$\frac{1}{18}$	0	0
3	0	0	0	$\frac{1}{18}$	0	$\frac{1}{18}$	0	$\frac{1}{18}$	0	0	0
4	0	0	0	0	$\frac{1}{18}$	0	$\frac{1}{18}$	0	0	0	0
5	0	0	0	0	0	$\frac{1}{18}$	0	0	0	0	0

As in the one-dimensional case, this function f has the following three properties:

1. $f(x, y) \geq 0$, for all $(x, y) \in \mathbb{R}^2$.
2. The set $\{(x, y) : f(x, y) \neq 0\}$ is at most countably infinite subset of $\mathbb{R}^2$.

3. $\sum_{(x,y)\in R(X,Y)} f(x,y) = 1.$

Any real-valued function f defined on $\mathbb{R}^2$ having these three properties will be called a two dimensional joint pmf.

One can easily verify all three properties for the joint pmf of Example 26.2

The joint PMF determines the probability of any event that can be specified in terms of the discrete random variables X_1 and X_2.

Theorem 26.5 *Let $X = (X_1, X_2)$ be a discrete random vector, with the joint pmf f. Then for any $A \subseteq \mathbb{R}^2$,*

$$P\{(X_1, X_2) \in A\} = \sum_{(x,y)\in A} f(x,y)$$

Exercise 26.6 *Suppose X be a random variable taking three values $-2, 1$ and 3, and let Y be a random variable that assume four values $-1, 0, 4, 6$. Their joint probabilities are given by the following table.*

X \ Y	-1	0	4	6
-2	$\frac{1}{9}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{9}$
1	$\frac{2}{9}$	0	$\frac{1}{9}$	$\frac{1}{9}$
3	0	0	$\frac{1}{9}$	$\frac{4}{27}$

Compute the probability of the event that XY is odd.

Solution: $\{XY \text{ is odd}\} = \{X = 1, Y = -1\} \bigcup \{X = 3, Y = -1\}$. Therefore

$$P(XY \text{ is odd}) = f(1,-1) + f(3,-1) = \frac{2}{9}$$

■

26.1 Marginal PMF

Even if we are considering a random vector (X, Y) there may be probabilities of interest that involve only one of the random variables in the random vector. We may wish to know $P(X = 2)$, for instance in Example 26.2. The random variable X is discrete and its probability distribution if described by it's pmf f_X (we now use the subscript to distinguish

$f_X(x)$ from $f_{X,Y}(x,y)$). We now call $f_X(x)$ the marginal pmf of X to emphasize the fact that it is the pmf of X but in the context of the joint pmf of the random vector (X, Y).

The marginal pmf of X or Y is easily calculated from the joint pmf of (X, Y).

Proposition 26.7 *If f is the joint pmf of X and Y, then*

$$f_X(x) = \sum_{y \in R(Y)} f(x,y), \quad f_Y(y) = \sum_{x \in R(X)} f(x,y),$$

Proof: Note that $\Omega = \bigcup_{y \in R(Y)} \{Y = y\}$. For each $x \in \mathbb{R}$, the events $\{X = x, Y = y\}, y \in R(Y)$ are disjoint and their union is the event $\{X = x\}$. Thus

$$\begin{aligned}
f_X(x) &= P(X = x) \\
&= P\left(\bigcup_{y \in R(Y)} \{X = x, Y = y\}\right) \\
&= \sum_{y \in R(Y)} P\left(X = x, Y = y\right) \\
&= \sum_{y \in R(Y)} f(x,y)
\end{aligned}$$

■