Lecture 30: Nov. 9

Last time

- Exponential families
- Location and Scale families
- Chebychev's Inequality

Today

- Multiple Random Variables (Chapter 4)

Joint and Marginal Distributions

In previous lectures, we have discussed probability models and computation of probability for events involving only one random variable. These are called *univariate models*.

In an experimental situation, it would be very unusual to observe only the value of one random variable. For example, in an experiment designed to gain information about some health characteristics of a population of people, the body weights of several people in the population might be measured. These different weights would be observations on difference random variables, one for each person measured. Multiple observations could also arise because several physical characteristics were measured on each person. Thus, we need to know how to describe and use probability models that deal with more than one random variable at a time.

Definition: An n-dimensional random vector $\boldsymbol{X} = (X_1, \ldots, X_n)$ is a function from a sample space S into $\mathbb{R}^n$.

- Each coordinate X_i is a random variable.
- The random vector is associated with a probability space $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), F)$.
- For each Borel set B,
$$\Pr\{\boldsymbol{X} \in B\} = \Pr\{\boldsymbol{X}^{-1}(B)\} \tag{1}$$
where
$$\boldsymbol{X}^{-1}(B) = \{w : \boldsymbol{X}(w) \in B\}$$

Example (Bivariate random variable) A fair coin is flipped 3 times. Define the random vector (X, Y) where X represents the number of heads on the last toss and Y the total number of heads. Then, the probabilities of various outcomes are given in the following table:

Outcome	(x, y)	Pr(*outcome*)
(H, H, H)	(1, 3)	1/8
(H, T, H), (T, H, H)	(1, 2)	2/8
(H, H, T)	(0, 2)	1/8
(T, T, H)	(1, 1)	1/8
(T, H, T), (H, T, T)	(0, 1)	2/8
(T, T, T)	(0, 0)	1/8

Definition Two random variables X and Y are said to be jointly *discrete* if there is an associated *joint probability mass function*,

$$f_{X,Y}(x, y) = \Pr\{X = x, Y = y\}$$

which sums to 1 over a finite or possibly countable combinations of x and y for which $f_{X,Y}(x, y) > 0$, i.e.,

$$\sum_{x,y} f_{X,Y}(x, y) = 1$$

From this, one can also obtain the marginal pmfs of X and Y as follows:

$$f_X(x) = \Pr(X = x) = \sum_y f_{X,Y}(x, y)$$
$$f_Y(y) = \Pr(Y = y) = \sum_x f_{X,Y}(x, y)$$

Example Back to the fair coin example again. From the definition, we can construct the joint pmf of X and Y:

		Y			
		0	1	2	3
X	0	1/8	1/4	1/8	0
	1	0	1/8	1/4	1/8

The marginal distributions of X and Y are also easy to find. Note: Marginals do not determine joint pmf.

Bivariate cdfs Whether they are discrete or continuous or some combination of the two, we can always define the *joint cdf*. For $n = 2$, the *bivariate cumulative distribution function* is

$$F_{X,Y}(x, y) = \Pr\{X \leqslant x, Y \leqslant y\}$$

Properties:

- $F_{X,Y}(x, y) \geqslant 0$
- $F_{X,Y}(\infty, \infty) = 1$
- $F_{X,Y}(-\infty, y) = F_{X,Y}(x, -\infty) = 0$
- $F_{X,Y}(-\infty, -\infty) = 0$
- F is non-decreasing and right-continuous in each variable separately.

Joint probabilities All joint probability statements about X and Y can be answered in terms of their joint cdf:

$$\Pr(x_1 < X \leqslant x_2, y_1 < Y \leqslant y_2) = \\ F_{X,Y}(x_2, y_2) + F_{X,Y}(x_1, y_1) - F_{X,Y}(x_1, y_2) - F_{X,Y}(x_2, y_1)$$

Example

$$\Pr(X > x, Y > y) = 1 - F_X(x) - F_Y(y) + F_{X,Y}(x, y)$$

Note: To ensure that a bivariate function $F(x, y)$ is a proper cdf, it must satisfy all the properties mentioned above and the rectangular property above.

Marginal distributions From $F_{X,Y}$, we can derive the univariate distribution functions for X and Y. These are generally called *marginal distributions*.

$$F_X(x) = \Pr\{X \leqslant x\} = \Pr\{X \leqslant x, Y \leqslant \infty\} = F_{X,Y}(x, \infty)$$
$$F_Y(y) = \Pr\{Y \leqslant y\} = \Pr\{X < \infty, Y \leqslant y\} = F_{X,Y}(\infty, y)$$

Note: Although we can obtain $F_X(x)$ and $F_Y(y)$ from the joint cdf, we cannot do the reverse.

Continuous Bivariate RVs The random variables X and Y are said to be *jointly continuous* if there exists a function $f_{X,Y}(x, y)$, such that for any Borel set B of 2-tuples in $\mathbb{R}^2$,

$$\Pr\{(X, Y) \in B\} = \int\int_{(x,y)\in B} f_{X,Y}(x, y)dxdy.$$

The function $f_{X,Y}(x, y)$ is called the *joint probability density function* for X and Y. It follows in this case that

$$F_{X,Y}(x, y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(s, t)dsdt,$$
$$f_{X,Y}(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}$$

Properties of the bivariate pdf

- $f_{X,Y}(x, y) \geqslant 0$
- $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x, y)dxdy = 1$
- $f_{X,Y}(x, y)$ is not a probability, but can be thought of as a relative probability of (X, Y) falling into a small rectangle located at (x, y):

$$\Pr\{x < X \leqslant x + dx, y < Y \leqslant y + dy\} \approx f(x, y)dxdy$$

- The *marginal probability density functions* for X and Y can be obtained as

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dy$$
$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dx$$

Example 1

$$
\begin{aligned}
F_{X,Y}(x,y) &= xy \qquad 0 < x \leqslant 1, 0 < y \leqslant 1 \\
f_{X,Y}(x,y) &= \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y} = \\
f_X(x) &= \\
f_Y(y) &=
\end{aligned}
$$

Example 2

$$
\begin{aligned}
F_{X,Y}(x,y) &= x - x \log \frac{x}{y} \qquad 0 < x \leqslant y \leqslant 1 \\
f_{X,Y}(x,y) &= \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y} = \\
f_X(x) &= \\
f_Y(y) &=
\end{aligned}
$$

Note: Once we have $f_X(x)$ and $f_Y(y)$, we can obtain $F_X(x)$ and $F_Y(y)$ directly. Double check: $F_X(x) = F_{X,Y}(x, \infty)$.