Joint distributions (part 1)

Lecture 5b (STAT 24400 F24)

Joint distributions

Recall that a random variable is defined as a function of the outcome of some random experiment or random process.

If we define multiple functions from the same random process, then this gives us two or more random variables whose distributions and probabilities are linked, governed by the same probability on the common sample space.

$$\begin{aligned} X: \quad \Omega &\to \mathbb{R} \\ Y: \quad \Omega &\to \mathbb{R} \\ &\cdots \end{aligned}$$

The **joint distribution** of X and Y, or of a list of random variables $X_1, \ldots, X_n$, refers to the characterization of the joint probabilities.

Joint distribution for a pair of r.v.'s

In general we are interested in calculating probabilities of the form $\mathbb{P}((X,Y) \in A)$, where A is any "reasonable" (measurable) region in $\mathbb{R}^2$.

The **joint** CDF F or $F_{X,Y}$:

$$F(x,y) = \mathbb{P}(X \le x, Y \le y)$$

read the "," as "and"

We can also ask questions about one variable on its own:
e.g., $\mathbb{P}(X \le 3)$ implicitly means $\mathbb{P}(X \le 3,$ and Y takes any value).

This is called the **marginal** distribution of X. Its CDF is:

$$F_X(x) = \mathbb{P}(X \le x) = \lim_{y \to +\infty} \mathbb{P}(X \le x, Y \le y) = \lim_{y \to +\infty} F(x,y).$$

Later, we will also study **conditional** distributions,

e.g., $\mathbb{P}(X \ge 3 \mid Y \le 7)$ or $\mathbb{P}(X = 1 \mid Y = 5)$.

Joint distribution for more than 2 r.v.'s

For joint distributions of more than two variables, everything is analogous, e.g. for a triple (X_1, X_2, X_3), the joint CDF is

$$F(x_1, x_2, x_3) = \mathbb{P}(X_1 \le x_1, X_2 \le x_2, X_3 \le x_3)$$

We can ask about the **marginal** distribution of any subset of r.v.'s, e.g.,

$$F_{X_1}(x) = \mathbb{P}(X_1 \le x), \qquad F_{X_2,X_3}(x,y) = \mathbb{P}(X_2 \le x, X_3 \le y)$$

We can also ask about **conditional** probabilities & distributions, e.g.,

$$\mathbb{P}(X_1 \le 3 \mid X_2 \ge 2), \qquad \mathbb{P}(X_1 \le 3 \mid X_2 > 1, X_3 \le 4)$$

Discrete joint distributions

If the pair (X, Y) takes only finitely many or countably infinitely many values, then we can characterize its distribution by the joint probability mass function p (or $p_{X,Y}$):

$$p(x, y) = \mathbb{P}(X = x, Y = y)$$

for every possible value (x, y) for the pair.

The marginal distribution for X is then calculated as

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

Similarly, the marginal distribution for Y can be calculated as

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

Example (from joint pmf to marginal pmf)

Example You play the following game: at each round, you toss a coin.

- If it's Heads, you roll a die and win \$1 if you rolled a 6
- If it's Tails, the game ends

What is the distribution of X = total \$ won?

First let's calculate joint distribution of X and $Y = \#$ of rounds played.

- Possible values are integer pairs (x, y) with $y \geq 1$ and $0 \leq x \leq y - 1$.
- Joint PMF: for (x, y) in the support,

$$\begin{aligned} p(x, y) = \mathbb{P}(X = x, Y = y) &= \overbrace{\mathbb{P}(Y = y)}^{\text{Geometric}(\frac{1}{2})} \cdot \overbrace{\mathbb{P}(X = x \mid Y = y)}^{\text{Binomial}(y-1, \frac{1}{6})} \\ &= \left(\tfrac{1}{2}\right)^{y-1} \cdot \tfrac{1}{2} \cdot \binom{y-1}{x} \cdot \left(\tfrac{1}{6}\right)^x \cdot \left(\tfrac{5}{6}\right)^{y-1-x} \end{aligned}$$

- Marginal PMF: for integer $x \geq 0$,

$$p_X(x) = \sum_y p(x, y) = \sum_{y=x+1}^{\infty} \left(\tfrac{1}{2}\right)^{y-1} \cdot \tfrac{1}{2} \cdot \binom{y-1}{x} \cdot \left(\tfrac{1}{6}\right)^x \cdot \left(\tfrac{5}{6}\right)^{y-1-x}.$$

Example (from joint pmf to marginal probability)

Suppose that, out of 80 people commuting to work, each one independently has a 75% chance of choosing to drive, and the chance of being over the speed limit is 45% for each drive.

What's the probability that more than half of the drivers are speeding?

Let $X = \#$ drivers, $Y = \#$ speeding drivers. Joint PMF:

$$p(k, \ell) = \mathbb{P}(X = k)\mathbb{P}(Y = \ell | X = k) = \binom{80}{k} 0.75^k 0.25^{80-k} \cdot \binom{k}{\ell} 0.45^\ell 0.55^{k-\ell},$$

with possible values = integer pairs (k, ℓ) with $0 \leq \ell \leq k \leq 80$.

$$\mathbb{P}\left(\begin{array}{c}\text{more than half}\\ \text{are speeding}\end{array}\right) = \sum_{k=0}^{80} \sum_{\ell=\lfloor \frac{k}{2}+1 \rfloor}^{k} \binom{80}{k} 0.75^k 0.25^{80-k} \cdot \binom{k}{\ell} 0.45^\ell 0.55^{k-\ell}.$$

Independence of random variables

Implied: defined on the same sample space,
i.e. (X, Y) has a joint distribution

Random variables X and Y are **independent** if

$$F_{X,Y}(x, y) = F_X(x) F_Y(y)$$

for all values (x, y). We write this as $X \perp\!\!\!\perp Y$.

Independence is equivalent to:

- $F_{X,Y}(x, y) =$ (some function of x) $\cdot$ (some function of y) for all x, y
- $\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$ for all subsets $A \subseteq \mathbb{R}$, $B \subseteq \mathbb{R}$

Independence of discrete r.v.'s

For a discrete joint distribution on (X, Y), independence is equivalent to:

- It holds for all x, y that
$$p_{X,Y}(x, y) = p_X(x) \cdot p_Y(y)$$
- It holds for all x, y that
$$p_{X,Y}(x, y) = (\text{some function of } x) \cdot (\text{some function of } y)$$

Example (independent or not)

Draw a hand of 10 cards. Let $X = \#$ of red cards, $Y = \#$ of Kings.
Does it hold that $X \perp\!\!\!\perp Y$?

- Possible values for X: 0, 1, ..., 10
- Possible values for Y: 0, 1, 2, 3, 4
- But not all combinations are possible —
  for example if $X = 10$ (all red cards) then $Y \leq 2$
- We can see that $X \not\perp\!\!\!\perp Y$ because, for instance,
$$\underbrace{p_{X,Y}(10, 4)}_{=0} \neq \underbrace{p_X(10) \cdot p_Y(4)}_{\text{both } > 0}$$

Remarks (domain and independence)

General principle:

Under independence, the following condition must hold:

$$\{\text{possible } (X, Y) \text{ values}\} = \{\text{possible } X \text{ values}\} \times \{\text{possible } Y \text{ values}\}$$

If this condition is violated, for example, when we check the supports of the marginal distributions of X and Y, such as in the previous example, then $X \not\perp\!\!\!\perp Y$.

On the other hand, if this condition holds, all we know is that X and Y are possibly independent, but not necessarily.

Recap and look ahead

- Expected values and variance of random variables
- Joint distributions of two (or more) r.v.'s and probability of events
  $\mathbb{P}((X, Y) \in A)$, $F_{X,Y}$, $p_{X,Y}(x, y)$, $f_{X,Y}(x, y)$ (next)
- Marginal distributions $F_X(x)$, $F_Y(y)$
- Independent r.v.'s:
  joint CDF: $F_{X,Y} = F_X(x)\, F_Y(y)$,
  discrete joint PMF: $p_{X,Y}(x, y) = p_X(x) p_Y(y)$,
  continuous joint PDF: $f_{X,Y}(x, y) = f_X(x) f_Y(y)$ (next)
- Conditional distributions, conditional PMF and PDF (next)
$$p_{X|Y}(x|y) = \mathbb{P}(X = x \mid Y = y) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)} = \frac{p_{X,Y}(x, y)}{p_Y(y)}$$
$$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$