Chapter 8. Jointly Distributed Random Variables

Suppose we have two random variables X and Y defined on a sample space S with probability measure $\mathrm{P}(E), E \subseteq S$.

Note that S could be the set of outcomes from two 'experiments', and the sample space points be two-dimensional; then perhaps X could relate to the first experiment, and Y to the second.

Then from before we know to define the *marginal* probability distributions P_X and P_Y by, for example,

$$\mathrm{P}_X(B) = \mathrm{P}(X^{-1}(B)), \quad B \subseteq \mathbb{R}.$$

We now define the **joint probability distribution**:

Definition 8.0.1. *Given a pair of random variables, X and Y, we define the* **joint probability distribution** P_{XY} *as follows:*

$$\mathrm{P}_{XY}(B_X, B_Y) = \mathrm{P}\{X^{-1}(B_X) \cap Y^{-1}(B_Y)\}, \quad B_X, B_Y \subseteq \mathbb{R}.$$

So $\mathrm{P}_{XY}(B_X, B_Y)$, the probability that $X \in B_X$ **and** $Y \in B_Y$, is given by the probability of the set of all points in the sample space that get mapped **both** into B_X by X **and** into B_Y by Y.

More generally, for a single region $B_{XY} \subseteq \mathbb{R}^2$, find the collection of sample space elements

$$S_{XY} = \{s \in S | (X(s), Y(s)) \in B_{XY}\}$$

and define

$$\mathrm{P}_{XY}(B_{XY}) = \mathrm{P}(S_{XY}).$$

8.0.1 Joint Cumulative Distribution Function

We define the joint cumulative distribution as follows:

Definition 8.0.2. *Given a pair of random variables, X and Y, the joint cumulative distribution function is defined as*

$$F_{XY}(x, y) = \mathrm{P}_{XY}(X \leq x, Y \leq y), \quad x, y \in \mathbb{R}.$$

It is easy to check that the marginal cdfs for X and Y can be recovered by

$$F_X(x) = F_{XY}(x, \infty), \quad x \in \mathbb{R},$$
$$F_Y(y) = F_{XY}(\infty, y), \quad y \in \mathbb{R},$$

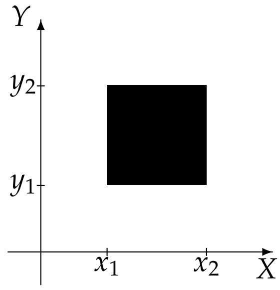
and that the two definitions will agree.

8.0.2 Properties of Joint CDF F_{XY}

For F_{XY} to be a valid cdf, we need to make sure the following conditions hold.

1. $0 \leq F_{XY}(x,y) \leq 1, \forall x,y \in \mathbb{R}$;
2. Monotonicity: $\forall x_1, x_2, y_1, y_2 \in \mathbb{R}$,
   $x_1 < x_2 \Rightarrow F_{XY}(x_1,y_1) \leq F_{XY}(x_2,y_1)$ and $y_1 < y_2 \Rightarrow F_{XY}(x_1,y_1) \leq F_{XY}(x_1,y_2)$;
3. $\forall x,y \in \mathbb{R}$,
   $F_{XY}(x,-\infty) = 0, F_{XY}(-\infty,y) = 0$ and $F_{XY}(\infty,\infty) = 1$.

Suppose we are interested in whether the random variable pair (X,Y) lie in the interval cross product $(x_1,x_2] \times (y_1,y_2]$; that is, if $x_1 < X \leq x_2$ *and* $y_1 < Y \leq y_2$.



First note that $P_{XY}(x_1 < X \leq x_2, Y \leq y) = F_{XY}(x_2,y) - F_{XY}(x_1,y)$.

It is then easy to see that $P_{XY}(x_1 < X \leq x_2, y_1 < Y \leq y_2)$ is given by

$$F_{XY}(x_2,y_2) - F_{XY}(x_1,y_2) - F_{XY}(x_2,y_1) + F_{XY}(x_1,y_1).$$

8.0.3 Joint Probability Mass Functions

Definition 8.0.3. *If X and Y are both discrete random variables, then we can define the* **joint probability mass function** *as*

$$p_{XY}(x,y) = P_{XY}(X = x, Y = y), \quad x,y \in \mathbb{R}.$$

We can recover the marginal pmfs p_X and p_Y since, by the law of total probability, $\forall x,y \in \mathbb{R}$

$$p_X(x) = \sum_y p_{XY}(x,y), \quad p_Y(y) = \sum_x p_{XY}(x,y).$$

Properties of Joint PMFs

For p_{XY} to be a valid pmf, we need to make sure the following conditions hold.

1. $0 \leq p_{XY}(x,y) \leq 1, \forall x,y \in \mathbb{R}$;
2. $\sum_y \sum_x p_{XY}(x,y) = 1$.

8.0.4 Joint Probability Density Functions

On the other hand, if $\exists f_{XY} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ s.t.

$$\mathrm{P}_{XY}(B_{XY}) = \int_{(x,y)\in B_{XY}} f_{XY}(x,y)dxdy, \quad B_{XY} \subseteq \mathbb{R} \times \mathbb{R},$$

then we say X and Y are **jointly continuous** and we refer to f_{XY} as the **joint probability density function** of X and Y.

In this case, we have

$$F_{XY}(x,y) = \int_{t=-\infty}^{y} \int_{s=-\infty}^{x} f_{XY}(s,t)dsdt, \quad x,y \in \mathbb{R},$$

Definition 8.0.4. *By the Fundamental Theorem of Calculus we can identify the joint pdf as*

$$f_{XY}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{XY}(x,y).$$

Furthermore, we can recover the marginal densities f_X and f_Y:

$$\begin{aligned} f_X(x) &= \frac{d}{dx} F_X(x) = \frac{d}{dx} F_{XY}(x,\infty) \\ &= \frac{d}{dx} \int_{y=-\infty}^{\infty} \int_{s=-\infty}^{x} f_{XY}(s,y)dsdy. \end{aligned}$$

By the Fundamental Theorem of Calculus, and through a symmetric argument for Y, we thus get

$$f_X(x) = \int_{y=-\infty}^{\infty} f_{XY}(x,y)dy, \quad f_Y(y) = \int_{x=-\infty}^{\infty} f_{XY}(x,y)dx.$$

Properties of Joint PDFs

For f_{XY} to be a valid pdf, we need to make sure the following conditions hold.

1. $f_{XY}(x,y) \geq 0, \forall x,y \in \mathbb{R}$;
2. $\int_{y=-\infty}^{\infty} \int_{x=-\infty}^{\infty} f_{XY}(x,y)dxdy = 1$.

8.1 Independence and Expectation

8.1.1 Independence

Two random variables X and Y are **independent if and only if** $\forall B_X, B_Y \subseteq \mathbb{R}.$,

$$\mathrm{P}_{XY}(B_X, B_Y) = \mathrm{P}_X(B_X)\mathrm{P}_Y(B_Y).$$

More specifically,

Definition 8.1.1. *Two random variables X and Y are independent* **if and only if**

$$f_{XY}(x,y) = f_X(x) f_Y(y), \quad \forall x, y \in \mathbb{R}.$$

Definition 8.1.2. *For two random variables X, Y we define the* **conditional probability distribution** $\mathrm{P}_{Y|X}$ *by*

$$\mathrm{P}_{Y|X}(B_Y|B_X) = \frac{\mathrm{P}_{XY}(B_X, B_Y)}{\mathrm{P}_X(B_X)}, \quad B_X, B_Y \subseteq \mathbb{R}.$$

This is the revised probability of Y falling inside B_Y *given* that we now know $X \in B_X$.

Then we have X and Y are independent $\iff \mathrm{P}_{Y|X}(B_Y|B_X) = \mathrm{P}_Y(B_Y), \forall B_X, B_Y \subseteq \mathbb{R}$.

Definition 8.1.3. *For random variables X, Y we define the* **conditional probability density function** $f_{Y|X}$ *by*

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)}, \quad x, y \in \mathbb{R}.$$

Note The random variables X and Y are independent $\iff f_{Y|X}(y|x) = f_Y(y), \forall x, y \in \mathbb{R}$.

8.1.2 Expectation

Suppose we have a (measurable) bivariate function of interest of the random variables X and Y, $g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$.

Definition 8.1.4. *If X and Y are discrete, we define $E\{g(X,Y)\}$ by*

$$E_{XY}\{g(X,Y)\} = \sum_y \sum_x g(x,y) p_{XY}(x,y).$$

Definition 8.1.5. *If X and Y are jointly continuous, we define $E\{g(X,Y)\}$ by*

$$E_{XY}\{g(X,Y)\} = \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{\infty} g(x,y) f_{XY}(x,y) dx dy.$$

Immediately from these definitions we have the following:

- If $g(X,Y) = g_1(X) + g_2(Y)$,

$$\mathrm{E}_{XY}\{g_1(X) + g_2(Y)\} = E_X\{g_1(X)\} + E_Y\{g_2(Y)\}.$$

- If $g(X,Y) = g_1(X)g_2(Y)$ and X and Y are **independent**,

$$\mathrm{E}_{XY}\{g_1(X)g_2(Y)\} = \mathrm{E}_X\{g_1(X)\}\mathrm{E}_Y\{g_2(Y)\}.$$

  In particular, considering $g(X,Y) = XY$ for independent X, Y we have

$$\mathrm{E}_{XY}(XY) = \mathrm{E}_X(X)\mathrm{E}_Y(Y).$$

8.1.3 Conditional Expectation

Warning! In general $\mathrm{E}_{XY}(XY) \neq \mathrm{E}_X(X)\mathrm{E}_Y(Y)$.

Suppose X and Y are discrete random variables with joint pmf $p(x,y)$. If we are given the value x of the random variable X, our revised pmf for Y is the conditional pmf $p(y|x)$, for $y \in \mathbb{R}$.

Definition 8.1.6. *The* **conditional expectation** *of Y given $X = x$ is therefore*

$$E_{Y|X}(Y|X=x) = \sum_y y\ p(y|x).$$

Similarly,

Definition 8.1.7. *If X and Y were continuous,*

$$E_{Y|X}(Y|X=x) = \int_{y=-\infty}^{\infty} y\ f(y|x)dy.$$

In either case, the conditional expectation is a function of x but not the unknown Y.

For a single variable X we considered the expectation of $g(X) = (X-\mu_X)(X-\mu_X)$, called the variance and denoted σ_X^2.

The bivariate extension of this is the expectation of $g(X,Y) = (X-\mu_X)(Y-\mu_Y)$. We define the **covariance** of X and Y by

$$\sigma_{XY} = \mathrm{Cov}(X,Y) = \mathrm{E}_{XY}[(X-\mu_X)(Y-\mu_Y)].$$

Covariance measures how the random variables move in tandem with one another, and so is closely related to the idea of correlation.

Definition 8.1.8. *We define the* **correlation** *of X and Y by*

$$\rho_{XY} = Cor(X,Y) = \frac{\sigma_{XY}}{\sigma_X\sigma_Y}.$$

Unlike the covariance, the correlation is invariant to the scale of the random variables X and Y.

It is easily shown that if X and Y are independent random variables, then $\sigma_{XY} = \rho_{XY} = 0$.

8.2 Examples

Example Suppose that the lifetime, X, and brightness, Y of a light bulb are modelled as continuous random variables. Let their joint pdf be given by

$$f(x,y) = \lambda_1\lambda_2 e^{-\lambda_1 x - \lambda_2 y}, \quad x,y > 0.$$

Question Are lifetime and brightness independent?

Solution If the lifetime and brightness are independent we would have

$$f(x,y) = f(x)f(y) \quad \text{for all } x,y \in \mathbb{R}.$$

The marginal pdf for X is

$$\begin{aligned} f(x) &= \int_{y=-\infty}^{\infty} f(x,y)dy = \int_{y=0}^{\infty} \lambda_1\lambda_2 e^{-\lambda_1 x - \lambda_2 y} dy \\ &= \lambda_1 e^{-\lambda_1 x}. \end{aligned}$$

Similarly $f(y) = \lambda_2 e^{-\lambda_2 y}$. Hence $f(x,y) = f(x)f(y)$ and X and Y are independent. ■

Example Suppose continuous random variables $(X,Y) \in \mathbb{R}^2$ have joint pdf

$$f(x,y) = \begin{cases} 1, & |x| + |y| < 1/\sqrt{2} \\ 0, & \text{otherwise.} \end{cases}$$

Question Determine the marginal pdfs for X and Y.

Solution
We have $|x| + |y| < 1/\sqrt{2} \iff |y| < 1/\sqrt{2} - |x|$. So

$$f(x) = \int_{y=-(\frac{1}{\sqrt{2}}-|x|)}^{\frac{1}{\sqrt{2}}-|x|} dy = \sqrt{2} - 2|x|.$$

Similarly $f(y) = \sqrt{2} - 2|y|$. Hence $f(x,y) \neq f(x)f(y)$ and X and Y are not independent. ■