Jointly Distributed Random Variables

COMP 245 STATISTICS

Dr N A Heard

Contents

1 Jointly Distributed Random Variables 1
  1.1 Definition ......................................................... 1
  1.2 Joint cdfs ....................................................... 2
  1.3 Joint Probability Mass Functions ................................. 2
  1.4 Joint Probability Density Functions ............................... 3

2 Independence and Expectation 4
  2.1 Independence ..................................................... 4
  2.2 Expectation ....................................................... 5
  2.3 Conditional Expectation .......................................... 5
  2.4 Covariance and Correlation ...................................... 6

3 Examples 6

1 Jointly Distributed Random Variables

1.1 Definition

Joint Probability Distribution

Suppose we have two random variables $X$ and $Y$ defined on a sample space $S$ with probability measure $P(E), E \subseteq S$. Jointly, they form a map $(X, Y) : S \rightarrow \mathbb{R}^2$ where $s \mapsto (X(s), Y(s))$.

From before, we know to define the marginal probability distributions $P_X$ and $P_Y$ by, for example,

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

We now define the joint probability distribution $P_{XY}$ by

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

So $P_{XY}(B_X, B_Y)$, the probability that $X \in B_X$ and $Y \in B_Y$, is given by the probability $P$ 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

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

and define

$$P_{XY}(B_{XY}) = P(S_{B_{XY}}).$$
1.2 Joint cdfs

Joint Cumulative Distribution Function

We thus define the joint cumulative distribution function as

\[ F_{XY}(x, y) = \mathbb{P}_{XY}((-\infty, x], (-\infty, y]), \quad x, y \in \mathbb{R}. \]

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.

Properties of a Joint cdf

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

Interval Probabilities

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 \(\mathbb{P}_{XY}([x_1, x_2], (-\infty, y]) = F_{XY}(x_2, y) - F_{XY}(x_1, y).\)

It is then easy to see that \(\mathbb{P}_{XY}((x_1, x_2], (y_1, y_2])\) is given by

\[ F_{XY}(x_2, y_2) - F_{XY}(x_1, y_2) - F_{XY}(x_2, y_1) + F_{XY}(x_1, y_1). \]

1.3 Joint Probability Mass Functions

If \(X\) and \(Y\) are both discrete random variables, then we can define the joint probability mass function as

\[ p_{XY}(x, y) = \mathbb{P}_{XY}({x}, {y}), \quad x, y \in \mathbb{R}. \]

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 a Joint pmf

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. \( \sum_y \sum_x p_{XY}(x, y) = 1. \)

1.4 Joint Probability Density Functions

On the other hand, if \( \exists f_{XY} : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} \) s.t.

\[
P_{XY}(B_{XY}) = \int_{(x,y) \in B_{XY}} f_{XY}(x, y) \, dx \, dy, \quad B_{XY} \subseteq \mathbb{R} \times \mathbb{R},
\]

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

In this case, we have

\[
F_{XY}(x, y) = \int_{t=-\infty}^{y} \int_{s=-\infty}^{x} f_{XY}(s, t) \, ds \, dt, \quad x, y \in \mathbb{R},
\]

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 \( f_Y \):

\[
f_X(x) = \frac{d}{dx} F_X(x) = \frac{d}{dx} F_{XY}(x, \infty)
\]

\[
= \frac{d}{dx} \int_{y=-\infty}^{\infty} \int_{s=-\infty}^{x} f_{XY}(s, y) \, ds \, dy.
\]

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 a Joint pdf

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. \( \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{\infty} f_{XY}(x, y) \, dx \, dy = 1. \)
2 Independence and Expectation

2.1 Independence

Independence of Random Variables

Two random variables \(X\) and \(Y\) are independent iff \(\forall B_X, B_Y \subseteq \mathbb{R},\)

\[
P_{XY}(B_X, B_Y) = P_X(B_X)P_Y(B_Y).
\]

More specifically, two discrete random variables \(X\) and \(Y\) are independent iff

\[
p_{XY}(x, y) = p_X(x)p_Y(y), \quad \forall x, y \in \mathbb{R};
\]

and two continuous random variables \(X\) and \(Y\) are independent iff

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

Conditional Distributions

For two r.v.s \(X, Y\) we define the conditional probability distribution \(P_{Y|X}\) by

\[
P_{Y|X}(B_Y|B_X) = \frac{P_{XY}(B_X, B_Y)}{P_X(B_X)}, \quad B_X, B_Y \subseteq \mathbb{R}, P_X(B_X) \neq 0.
\]

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 \(\iff\) \(P_{Y|X}(B_Y|B_X) = P_Y(B_Y), \forall B_X, B_Y \subseteq \mathbb{R}.

For discrete r.v.s \(X, Y\) we define the conditional probability mass function \(p_{Y|X}\) by

\[
p_{Y|X}(y|x) = \frac{p_{XY}(x,y)}{p_X(x)}, \quad x, y \in \mathbb{R}, p_X(x) \neq 0,
\]

and for continuous r.v.s \(X, Y\) we define the conditional probability density function \(f_{Y|X}\) by

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

[Justification:]

\[
P(Y \leq y|X \in [x, x + dx)) = \frac{P_{XY}([x, x + dx), (-\infty, y])}{P_X([x, x + dx))} = \frac{F(y, x + dx) - F(y, x)}{F(x + dx) - F(x)} = \frac{\{F(y, x + dx) - F(y, x)\}/dx}{\{F(x + dx) - F(x)\}/dx} \Rightarrow f(y|X \in [x, x + dx)) = \frac{d\{F(y, x + dx) - F(y, x)\}/dxdy}{\{F(x + dx) - F(x)\}/dx} \Rightarrow f(y|x) = \lim_{dx \to 0} f(y|X \in [x, x + dx)) = \frac{f(y, x)}{f(x)}dfrac{dx}{dx}.
\]

In either case, \(X\) and \(Y\) are independent \(\iff\) \(p_{Y|X}(y|x) = p_Y(y)\) or \(f_{Y|X}(y|x) = f_Y(y), \forall x, y \in \mathbb{R}.\)
2.2 Expectation

\( E\{g(X, Y)\} \)

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

If \( X \) and \( Y \) are discrete, we define \( E\{g(X, Y)\} \) by

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

If \( X \) and \( Y \) are jointly continuous, we define \( E\{g(X, Y)\} \) by

\[
E_{XY}\{g(X, Y)\} = \int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} g(x, y)f_{XY}(x, y)dxdy.
\]

Immediately from these definitions we have the following:

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

\[
E_{XY}\{g_1(X) + g_2(Y)\} = E_X\{g_1(X)\} + E_Y\{g_2(Y)\}.
\]

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

\[
E_{XY}\{g_1(X)g_2(Y)\} = E_X\{g_1(X)\}E_Y\{g_2(Y)\}.
\]

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

\[
E_{XY}(XY) = E_X(X)E_Y(Y).
\]

2.3 Conditional Expectation

\( E_{Y|X}(Y|X = x) \)

In general \( E_{XY}(XY) \neq E_X(X)E_Y(Y) \).

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

The **conditional expectation** of \( Y \) given \( X = x \) is therefore

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

Similarly, 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 \( Y \).
2.4 Covariance and Correlation

Covariance

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

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

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

Correlation

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

We define the correlation of $X$ and $Y$ by

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

Unlike the covariance, the correlation is invariant to the scale of the r.v.s $X$ and $Y$. It is easily shown that if $X$ and $Y$ are independent random variables, then $\sigma_{XY} = \rho_{XY} = 0$.

3 Examples

Example 1

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

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

Are lifetime and brightness independent?

The marginal pdf for $X$ is

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

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

Example 2

Suppose continuous r.v.s $(X, Y) \in \mathbb{R}^2$ have joint pdf

$$f(x, y) = \begin{cases} \frac{1}{\pi}, & x^2 + y^2 \leq 1 \\ 0, & \text{otherwise}. \end{cases}$$

Determine the marginal pdfs for $X$ and $Y$.

Well $x^2 + y^2 \leq 1 \iff |y| \leq \sqrt{1 - x^2}$. So

$$f(x) = \int_{y=-\infty}^{\infty} f(x, y) \, dy = \int_{y=0}^{\sqrt{1-x^2}} \frac{1}{\pi} \, dy = \frac{2}{\pi} \sqrt{1-x^2}$$

Similarly $f(y) = \frac{2}{\pi} \sqrt{1-y^2}$. Hence $f(x, y) \neq f(x)f(y)$ and $X$ and $Y$ are not independent.