AMS 131: Introduction to Probability Theory (Spring 2009)

Chapter 3: Sections 3.4–3.6

A bivariate distribution refers to the joint probability distribution of two random variables defined on a common sample space and each taking values on (a subset of) the real line $\mathbb{R}$.

Discrete joint distributions: A bivariate r.v. $(X, Y)$ is discrete if both r.v.s $X$ and $Y$ have either a finite number or countable number of possible values. The probability distribution of a discrete bivariate r.v. $(X, Y)$ can be defined through the joint probability function (p.f.),

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

for any possible value $(x, y)$ of $(X, Y)$ (the definition can be extended for any $(x, y) \in \mathbb{R}^2$, since $f(x, y) = 0$ for all $(x, y)$ that are not possible values of $(X, Y)$). A joint p.f. must satisfy two conditions: (1) $f(x, y) \geq 0$, for all $(x, y)$; and (2) $\sum_x \sum_y f(x, y) = 1$. The probability of any subset $A$ of $\mathbb{R}^2$ can be obtained from $\Pr((X, Y) \in A) = \sum \sum_{(x,y) \in A} f(x, y)$.

The marginal p.f.s of $X$ and $Y$ are given by

$$f_1(x) = \Pr(X = x) = \sum_y f(x, y), \ x \in \mathbb{R} \quad \text{and} \quad f_2(y) = \Pr(Y = y) = \sum_x f(x, y), \ y \in \mathbb{R}.$$

For any $y$ such that $f_2(y) > 0$, the conditional p.f. of $X$ given that $Y = y$ arises directly from the definition of the conditional probability of event $\{X = x\}$ given event $\{Y = y\}$. In particular,

$$g_1(x \mid y) = \Pr(X = x \mid Y = y) = \frac{f(x, y)}{f_2(y)}, \ x \in \mathbb{R}.$$

Analogously, for any $x$ such that $f_1(x) > 0$, the conditional p.f. of $Y$ given $X = x$ is given by $g_2(y \mid x) = \Pr(Y = y \mid X = x) = f(x, y)/f_1(x), \ y \in \mathbb{R}$.

Continuous joint distributions: A bivariate r.v. $(X, Y)$ is continuous if both r.v.s $X$ and $Y$ take an uncountable number of possible values (say, values in a bounded interval, in $\mathbb{R}^+$, or in $\mathbb{R}$). The probability distribution of a continuous bivariate r.v. $(X, Y)$ is determined by its joint probability density function (p.d.f.), $f(x, y)$, so that for any subset $A$ of $\mathbb{R}^2$,

$$\Pr((X, Y) \in A) = \int \int_A f(x, y) \, dx \, dy.$$

(Note that under a continuous bivariate distribution, points in $\mathbb{R}^2$ and one-dimensional subsets of $\mathbb{R}^2$ have probability 0.) Any joint p.d.f. must satisfy two conditions: (1) $f(x, y) \geq 0$, for all $(x, y) \in \mathbb{R}^2$; and (2) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \, dx \, dy = 1$.

The marginal p.d.f.s of $X$ and $Y$ are given by

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) \, dy, \ x \in \mathbb{R} \quad \text{and} \quad f_2(y) = \int_{-\infty}^{\infty} f(x, y) \, dx, \ y \in \mathbb{R}.$$

For any $y$ such that $f_2(y) > 0$, the conditional p.d.f. of $X$ given that $Y = y$ is defined by

$$g_1(x \mid y) = \frac{f(x, y)}{f_2(y)}, \ x \in \mathbb{R}.$$

Analogously, for any $x$ such that $f_1(x) > 0$, the conditional p.d.f. of $Y$ given $X = x$ is defined by $g_2(y \mid x) = f(x, y)/f_1(x), \ y \in \mathbb{R}$.
Mixed bivariate distributions: In certain applications, we may wish to work with pairs of r.v.s where one is discrete and the other is continuous.

The probability distribution of a mixed bivariate r.v. \((X, Y)\), where, for example, \(X\) is discrete and \(Y\) is continuous, is defined by a joint p.f./p.d.f., \(f(x, y)\), which provides the probability of any subset \(A\) of \(\mathbb{R}^2\) by summing the values of \(f(x, y)\) over \(x\) and integrating \(f(x, y)\) over \(y\), for all \((x, y)\) \(\in A\). Such a joint p.f./p.d.f. must satisfy two conditions: (1) \(f(x, y) \geq 0\), for all \((x, y) \in \mathbb{R}^2\); and (2) \(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \, dy \, dx = 1\).

Here, the marginal p.f. for \(X\) is given by \(f_1(x) = \Pr(X = x) = \int_{-\infty}^{\infty} f(x, y) \, dy\), \(x \in \mathbb{R}\), and the marginal p.d.f. for \(Y\) by \(f_2(y) = \int_{-\infty}^{\infty} f(x, y) \, dx\), \(y \in \mathbb{R}\).

Bivariate distribution functions: Although less convenient to work with than the univariate case, the definition of the distribution function can be extended to the joint distribution function (d.f.), which can again be used to characterize the distribution of a bivariate r.v. regardless of its type.

The joint d.f. of a bivariate r.v. \((X, Y)\) is a function on \(\mathbb{R}^2\) with values in [0, 1] defined by

\[
F(x, y) = \Pr(\{X \leq x\} \cap \{Y \leq y\}), \quad (x, y) \in \mathbb{R}^2.
\]

The marginal d.f. for r.v. \(X\) can be obtained from \(F_1(x) = \lim_{y \to -\infty} F(x, y)\), \(x \in \mathbb{R}\), and the marginal d.f. for r.v. \(Y\) can be obtained from \(F_2(y) = \lim_{x \to -\infty} F(x, y)\), \(y \in \mathbb{R}\).

For a continuous bivariate r.v. \((X, Y)\), we have \(F(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(r, s) \, dr \, ds\). Therefore, the joint p.d.f. can be obtained from the joint d.f. using

\[
f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}
\]

for all \((x, y)\) at which the second-order partial derivative above exists.

Independent random variables: Two r.v.s \(X\) and \(Y\) are independent if, by definition, for any subsets \(A\) and \(B\) of \(\mathbb{R}\), \(\Pr(\{X \in A\} \cap \{Y \in B\}) = \Pr(X \in A)\Pr(Y \in B)\).

It can be shown that r.v.s \(X\) and \(Y\) are independent if and only if

\[
F(x, y) = F_1(x)F_2(y), \quad \text{for all } (x, y) \in \mathbb{R}^2.
\]

Moreover, a similar factorization applies under independence to the joint p.f., joint p.d.f. or joint p.f./p.d.f., \(f(x, y)\), for discrete, continuous or mixed bivariate r.v.s, respectively. Specifically, r.v.s \(X\) and \(Y\) are independent if and only if

\[
f(x, y) = f_1(x)f_2(y), \quad \text{for all } (x, y) \in \mathbb{R}^2.
\]

Bayes theorem for random variables: Assume that the joint p.f., p.d.f or p.f./p.d.f., \(f(x, y)\), for r.v.s \(X\) and \(Y\) is built from the marginal p.f. or p.d.f., \(f_2(y)\), for \(Y\) and the conditional p.f. or p.d.f., \(g_1(x \mid y)\), of \(X\) given \(Y\). Then, the conditional distribution of \(Y\) given \(X\) can be obtained through its conditional p.f. or p.d.f.,

\[
g_2(y \mid x) = \frac{g_1(x \mid y)f_2(y)}{f_1(x)}
\]

where \(f_1(x) = \sum_y g_1(x \mid y)f_2(y)\) or \(f_1(x) = \int_{-\infty}^{\infty} g_1(x \mid y)f_2(y) \, dy\), if \(Y\) is a discrete or continuous r.v., respectively.