Multivariate Distribution

The random vector $\boldsymbol{X} = (X_1, \ldots, X_n)$ has a sample space that is a subset of $\mathbb{R}^n$. If $\boldsymbol{X}$ is discrete random vector, then the joint pmf of $\boldsymbol{x}$ is the function defined by $f(\boldsymbol{x}) = f(x_1, \ldots, x_n) = P(X_1 = x_1, \ldots, X_n - x_n)$ for each $(x_1, \ldots, x_n) \in \mathbb{R}^n$. Then for any $A \subset \mathbb{R}^n$,

$$P(\boldsymbol{X} \in A) = \sum_{\boldsymbol{x} \in A} f(\boldsymbol{x}).$$

If $\boldsymbol{X}$ is a continuous random vector, the joint pdf of $\boldsymbol{X}$ is a function $f(x_1, \ldots, x_n)$ that satisfies

$$P(\boldsymbol{X} \in A) = \int \cdots \int_A f(\boldsymbol{x}) d\boldsymbol{x} = \int \cdots \int_A f(x_1, \ldots, x_n) dx_1 \cdots dx_n.$$

Let $g(\boldsymbol{x}) = g(x_1, \ldots, x_n)$ be a real-valued function defined on the sample space of $\boldsymbol{X}$. Then $g(\boldsymbol{X})$ is a random variable and the expected value of $g(\boldsymbol{X})$ is

$$Eg(\boldsymbol{X}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(\boldsymbol{x}) f(\boldsymbol{x}) d\boldsymbol{x}$$

and

$$Eg(\boldsymbol{X}) = \sum_{\boldsymbol{x} \in \mathbb{R}^n} g(\boldsymbol{x}) f(\boldsymbol{x})$$

in the continuous and discrete cases, respectively.

The marginal distribution of $(X_1, \ldots, X_n)$, the first k coordinates of $(X_1, \ldots, X_n)$, is given by the pdf or pmf

$$f(x_1, \ldots, x_k) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n) dx_{k+1} \cdots dx_n$$

or

$$f(x_1, \ldots, x_k) = \sum_{(x_{k+1}, \ldots, x_n) \in \mathbb{R}^{n-k}} f(x_1, \ldots, x_n)$$

for every $(x_1, \ldots, x_k) \in \mathbb{R}^k$.

If $f(x_1, \ldots, x_k) > 0$, the conditional pdf or pmf of $(X_{k+1}, \ldots, X_n)$ given $X_1 = x_1, \ldots, X_k = x_k$ is the function of $(x_{k+1}, \ldots, x_n)$ defined by

$$f(x_{k=1}, \ldots, x_n | x_1, \ldots, x_k) = \frac{f(x_1, \ldots, x_n)}{f(x_1, \ldots, x_k)}.$$

Example 4.6.1 (Multivariate pdfs)

Let $n = 4$ and

$$f(x_1, x_2, x_3, x_4) = \begin{cases} \frac{3}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2) & 0 < x_i < 1, i = 1, 2, 3, 4 \\ 0 & \text{otherwise} \end{cases}$$

The joint pdf can be used to compute probabilities such as

$$P(X_1 < \frac{1}{2}, X_2 < \frac{3}{4}, X_4 > \frac{1}{2})$$
$$= \int_{\frac{1}{2}}^{1} \int_{0}^{1} \int_{0}^{\frac{3}{4}} \int_{0}^{\frac{1}{2}} \frac{3}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2) dx_1 dx_2 dx_3 dx_4 = \frac{151}{1024}.$$

The marginal pdf of (X_1, X_2) is

$$f(x_1, x_2) = \int_0^1 \int_0^1 \frac{3}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2) dx_2 dx_4 = \frac{3}{4}(x_1^2 + x_2^2) + \frac{1}{2}$$

for $0 < x_1 < 1$ and $0 < x_2 < 1$.

Definition 4.6.2 Let n and m be positive integers and let $p_1, \ldots, p_n$ be numbers satisfying $0 \leq p_i \leq 1$, $i = 1, \ldots, n$, and $\sum_{i=1}^n p_i = 1$. Then the random vector $(X_1, \ldots, X_n)$ has a multinomial distribution with m trials and cell proabilities $p_1, \ldots, p_n$ if the joint pmf of $(X_1, \ldots, X_n)$ is

$$f(x_1, \ldots, x_n) = \frac{m!}{x_1! \cdots x_n!} p_1^{x_1} \cdots p_n^{x_n} = m! \prod_{i=1}^n \frac{p_i^{x_i}}{x_i!}$$

on the set of $(x_1, \ldots, x_n)$ such that each x_i is a nonnegative integer and $\sum_{i=1}^n x_i = m$.

Example 4.6.3 (Multivariate pmf) Consider tossing a six-sided die 10 times. Suppose the die is unbalanced so that the probability of observing an i is $i/21$. Now consider the vector $(X_1, \ldots, X_6)$, where X_i counts the number of times i comes up in the 10 tosses. Then $(X_1, \ldots, X_6)$ has a multinomial distribution with $m = 10$ and cell probabilities $p_1 = \frac{1}{21}, \ldots, p_6 = \frac{6}{21}$. For example, the probability of the vector $(0, 0, 1, 2, 3, 4)$ is

$$f(0, 0, 1, 2, 3, 4) = \frac{10!}{0!0!1!2!3!4!}(\frac{1}{21})^0(\frac{2}{21})^0(\frac{3}{21})^1(\frac{4}{21})^2(\frac{5}{21})^3(\frac{6}{21})^4 = 0.0059.$$

The factor $\frac{m!}{x_1! \cdots x_n!}$ is called a multinomial coefficient. It is the number of ways that m objects can be divided into n groups with x_1 in the first group, x_2 in the second group, ..., and x_n in the nth group.

Theorem 4.6.4 (Multinomial Theorem)

Let m and n be positive integers. Let A be the set of vectors $\boldsymbol{x} = (x_1, \ldots, x_n)$ such that each x_i is a nonnegative integer and $\sum_{i=1}^{n} x_i = m$. Then, for any real numbers $p_1, \ldots, p_n$,

$$(p_1 + \ldots + p_n)^m = \sum_{\boldsymbol{x} \in A} \frac{m!}{x_1! \cdots x_n!} p_1^{x_1} \ldots p_n^{x_n}.$$

Definition 4.6.5 Let $X_1, \ldots, X_n$ be random vectors with joint pdf or pmf $f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$. Let $f_{\boldsymbol{X}_i}(x_i)$ denote the marginal pdf or pmf of $\boldsymbol{X}_i$. Then $X_1, \ldots, X_n$ are called mutually independent random vectors if, for every $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$,

$$f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = f_{\boldsymbol{X}_1}(\boldsymbol{x}_1) \ldots f_{\boldsymbol{X}_n}(\boldsymbol{x}_n) = \prod_{i=1}^{n} f_{\boldsymbol{X}_i}(\boldsymbol{x}_i).$$

If the X_i's are all one dimensional, then $X_1, \ldots, X_n$ are called mutually independent random variables.

Mutually independent random variables have many nice properties. The proofs of the following theorems are analogous to the proofs of their counterparts in Sections 4.2 and 4.3.

Theorem 4.6.6 (Generalization of Theorem 4.2.10)

Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be mutually independent random variables. Let $g_1, \ldots, g_n$ be real-valued functions such that $g_i(x_i)$ is a function only of x_i, $i = 1, \ldots, n$. Then

$$E(g_1(X_1) \cdots g(X_n)) = (Eg_1(X_1)) \cdots (Eg_n(X_n)).$$

Theorem 4.6.7 (Generalization of Theorem 4.2.12)

Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be mutually independent random variables with mgfs $M_{X_1}(t), \ldots, M_{X_n}(t)$. Let $Z = X_1 + \cdots + X_n$. Then the mgf of Z is

$$M_Z(t) = M_{X_1}(t) \cdots M_{X_n}(t).$$

In particular, if $X_1, \ldots, X_n$ all have the same distribution with mgf $M_X(t)$, then

$$M_Z(t) = (M_X(t))^n.$$

Example 4.6.8 (Mgf of a sum of gamma variables)

Suppose $X_1, \ldots, X_n$ are mutually independent random variables, and the distribution of X_i is gamma(α_i, β). Thus, if $Z = X_1 + \ldots + X_n$, the mgf of Z is

$$M_Z(t) = M_{X_1}(t) \cdots M_{X_n}(t) = (1-\beta t)^{-\alpha_1} \cdots (1-\beta t)^{-\alpha_n} = (1-\beta t)^{-(\alpha_1+\cdots+\alpha_n)}.$$

This is the mgf of a gamma$(\alpha_1 + \cdots + \alpha_n, \beta)$ distribution. Thus, the sum of a independent gamma random variables that have a common scale parameter β also has a gamma distribution.

Example

Let $X_1, \ldots, X_n$ be mutually independent random variables with $X_i \sim N(\mu_i, \sigma_i^2)$. Let $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ be fixed constants. Then

$$Z = \sum_{i=1}^{n} (a_i X_i + b_i) \sim N(\sum_{i=1}^{n} (a_i \mu_i + b_i), \sum_{i=1}^{n} a_i^2 \sigma_i^2).$$

Theorem 4.6.11 (Generalization of Lemma 4.2.7)

Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be random vectors. Then $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ are mutually independent random vectors if and only if there exist functions $g_i(\boldsymbol{x}_i)$, $i = 1, \ldots, n$, such that the joint pdf or pmf of $(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)$ can be written as

$$f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = g_1(\boldsymbol{x}_1) \cdots g_n(\boldsymbol{x}_n).$$

Theorem 4,6,12 (Generalization of Theorem 4.3.5)

Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be random vectors. Let $g_i(\boldsymbol{x}_i)$ be a function only of $\boldsymbol{x}_i$, $i = 1, \ldots, n$. Then the random vectors $U_i = g_i(\boldsymbol{X}_i)$, $i = 1, \ldots, n$, are mutually independent.

Let $(X_1, \ldots, X_n)$ be a random vector with pdf $f_X(x_1, \ldots, x_n)$. Let $\mathcal{A} = \{\boldsymbol{x} : f_X(\boldsymbol{x}) > 0\}$. Consider a new random vector $(U_1, \ldots, U_n)$, defined by $U_1 = g_1(X_1, \ldots, X_n)$, $\ldots$, $U_n = g_n(X_1, \ldots, X_n)$. Suppose that $A_0, A_1, \ldots, A_k$ form a partition of $\mathcal{A}$ with these properties. The set A_0, which may be empty, satisfies $P((X_1, \ldots, X_n) \in A_0) = 0$. The transformation $(U_1, \ldots, U_n) = (g_1(\boldsymbol{X}), \ldots, g_n(\boldsymbol{X}))$ is a one-to-one transformation from A_i onto B for each $i = 1, 2, \ldots, k$. Then for each i, the inverse functions from B to A_i can be found. Denote the

ith inverse by $x_1 = h_{1i}(u-1, \ldots, u_n), \ldots, x_n = h_{ni}(u_1, \ldots, u_n)$. Let J_i denote the Jacobian computed from the ith inverse. That is,

$$J_i = \begin{vmatrix} \frac{\partial h_{1i}(\boldsymbol{u})}{\partial u_1} & \frac{\partial h_{1i}(\boldsymbol{u})}{\partial u_2} & \cdots & \frac{\partial h_{1i}(\boldsymbol{u})}{\partial u_1} \\ \frac{\partial h_{2i}(\boldsymbol{u})}{\partial u_1} & \frac{\partial h_{2i}(\boldsymbol{u})}{\partial u_2} & \cdots & \frac{\partial h_{2i}(\boldsymbol{u})}{\partial u_1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_{ni}(\boldsymbol{u})}{\partial u_1} & \frac{\partial h_{ni}(\boldsymbol{u})}{\partial u_2} & \cdots & \frac{\partial h_{ni}(\boldsymbol{u})}{\partial u_1} \end{vmatrix}$$

the determinant of an $n \times n$ matrix. Assuming that these Jacobians do not vanish identically on B, we have the following representation of the joint pdf, $f_U(u_1, \ldots, u_n)$, for $\boldsymbol{u} \in B$:

$$f_{\boldsymbol{u}}(u_1, \ldots, u_n) = \sum_{i=1}^{k} f_X(h_{1i}(u_1, \ldots, u_n), \ldots, h_{ni}(u_1, \ldots, u_n))|J_i|.$$