STA 623 – Fall 2010 – Dr. Charnigo

Section 4.6: Multivariate Distributions

Probability mass functions. Let $\mathbf{X}$ be a random vector with components $X_1, \ldots, X_n$, where n is a positive integer greater than 2. If $\mathbf{X}$ realizes only finitely or countably many values, then we say that $\mathbf{X}$ is discrete. In this case, $\mathbf{X}$ may be characterized by a joint probability mass function

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) := P(X_1 = x_1, \ldots, X_n = x_n)$$

such that

$$P(\mathbf{X} \in A) = \sum_{(x_1,\ldots,x_n)' \in A \cap S_{\mathbf{X}}} f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$$

for any (appropriately measurable) set $A \subset \mathbb{R}^n$ and $S_{\mathbf{X}}$ denotes the support of the joint probability mass function.

The marginal probability mass function of X_1 is obtained by summation,

$$f_{X_1}(u) = \sum_{(x_1,\ldots,x_n)' \in S_{\mathbf{X}} \;:\; x_1 = u} f_{X_1,\ldots,X_n}(x_1, \ldots, x_n).$$

What if we want the joint probability mass function of X_1 and X_2 only (i.e., not also X_3 through X_n)? Again we can employ summation,

$$f_{X_1,X_2}(u, v) = \sum_{(x_1,\ldots,x_n)' \in S_{\mathbf{X}} \;:\; x_1 = u, x_2 = v} f_{X_1,\ldots,X_n}(x_1, \ldots, x_n).$$

The conditional probability mass function of X_1 given that $X_2 = x_2, \ldots, X_n = x_n$ (assumed to be an event with positive probability) is obtained by division,

$$f_{X_1|X_2,\ldots,X_n}(x_1|x_2, \ldots, x_n) = f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)/f_{X_2,\ldots,X_n}(x_2, \ldots, x_n).$$

If the joint probability mass function of $X_1, \ldots, X_n$ factors into their marginal probability mass functions, then we say that $X_1, \ldots, X_n$ are independent and in this case the conditional probability mass functions coincide with the marginal probability mass functions.

Example (probability mass functions). Let $\mathbf{X} = (X_1, X_2, X_3)'$ have the joint probability mass function $(1 + x_1 + 2x_2 + 3x_3)/32$ for $x_1, x_2, x_3 \in \{0, 1\}$. The marginal probability mass function of X_1 is

The joint probability mass function of X_2 and X_3 is

The conditional probability mass function of X_1 given that $X_2 = X_3 = 0$ is

Probability density functions. There are parallel developments for probability density functions pertaining to the components of a continuous random vector $\mathbf{X}$. First try formulating these developments yourself, without consulting the textbook. Then use the textbook as a check.

Expectations. Let g be a(n appropriately measurable) function from $\mathbb{R}^n$ to $\mathbb{R}$ and $\mathbf{X}$ a discrete or continuous random vector. Assuming absolute convergence of the appropriate expression below, we define the expected value of $g(X_1, \ldots, X_n)$ as

$$\sum_{(x_1,\ldots,x_n)' \in S_{\mathbf{X}}} g(x_1, \ldots, x_n) f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$$

if $\mathbf{X}$ is discrete and as

$$\int_{\mathbb{R}^n} g(x_1, \ldots, x_n) f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)\; dx_1 \ldots dx_n$$

is $\mathbf{X}$ is continuous. For an expectation conditional on $X_2 = x_2, \ldots, X_n = x_n$, replace $f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$ by $f_{X_1|X_2,\ldots,X_n}(x_1|x_2, \ldots, x_n)$, summation over $(x_1, \ldots, x_n)' \in S_{\mathbf{X}}$ by summation over $x_1 \in S_{X_1|x_2,\ldots,x_n}$, and integration in $dx_1 \ldots dx_n$ over $\mathbb{R}^n$ by integration in dx_1 over $\mathbb{R}$.

Useful results on expectations. If $X_1, \ldots, X_n$ are independent, then for (appropriately measurable) functions $g_1, \ldots, g_n$ from $\mathbb{R}$ to $\mathbb{R}$ we have

$$E[g_1(X_1) \times \cdots \times g_n(X_n)] = E[g_1(X_1)] \times \cdots \times E[g_n(X_n)],$$

assuming all expectations exist as finite numbers. A special case occurs when $g_1(x) = \cdots = g_n(x) = \exp[tx]$,

$$M_{X_1 + \ldots + X_n}(t) = M_{X_1}(t) \times \cdots \times M_{X_n}(t).$$

The latter result can be used to show that some families of distributions are closed under convolution (i.e., summing independent random variables from the family yields another random variable from the same family).

One-to-one multivariate transformation formula. Let $\mathbf{X}$ be a continuous random vector with components $X_1, \ldots, X_n$ whose joint probability density function is $f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$. Suppose that $g_1(x_1, \ldots, x_n), \ldots, g_n(x_1, \ldots, x_n)$ are (appropriately measurable) functions that define a one-to-one multivariate transformation so that the equations $u_1 = g_1(x_1, \ldots, x_n), \ldots, u_n = g_n(x_1, \ldots, x_n)$ can be solved for $x_1, \ldots, x_n$, say $x_1 = h_1(u_1, \ldots, u_n), \ldots, x_n = h_n(u_1, \ldots, u_n)$. Put $U_1 := g_1(X_1, \ldots, X_n), \ldots, U_n := g_n(X_1, \ldots, X_n)$. Assuming the existence of all partial derivatives referenced below, the joint probability density function of $U_1, \ldots, U_n$ is

$$f_{U_1,\ldots,U_n}(u_1, \ldots, u_n) = f_{X_1,\ldots,X_n}(h_1(u_1, \ldots, u_n), \ldots, h_n(u_1, \ldots, u_n)) \times$$
$$\left| Det \left[\frac{\partial \mathbf{h}}{\partial \mathbf{u}} \right] \right| 1_{\{u_1,\ldots,u_n)' \in S_{U_1,\ldots,U_n}\}},$$

where $Det[\cdot]$ returns the determinant of a matrix (assumed nonzero here), $\frac{\partial \mathbf{h}}{\partial \mathbf{u}}$ is the $n \times n$ matrix of partial derivatives of $h_1(u_1, \ldots, u_n), \ldots, h_n(u_1, \ldots, u_n)$ in $u_1, \ldots, u_n$, and

$$S_{U_1,\ldots,U_n} := \{(u_1, \ldots, u_n)' \in \mathbb{R}^n : \exists (x_1, \ldots, x_n)' \in \mathbb{R}^n \text{ with}$$
$$u_1 = g_1(x_1, \ldots, x_n), \ldots, u_n = g_n(x_1, \ldots, x_n), f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) > 0\}.$$