STA 623 – Fall 2009 – Dr. Charnigo

Section 4.2: Conditional Distributions and Independence

Conditional probability mass function and expectation. Let $\mathbf{X}$ be a discrete random vector with components X and Y. Let $f_{X,Y}(x,y)$ denote the joint probability mass function of X and Y, $f_X(x)$ the marginal probability mass function of X, and $f_Y(y)$ the marginal probability mass function of Y.

For any $x \in \mathbb{R}$ at which $f_X(x) > 0$, we define

$$f_{Y|X}(y|x) := f_{X,Y}(x,y)/f_X(x) = P(Y=y, X=x)/P(X=x) = P(Y=y|X=x)$$

to be the conditional probability mass function of Y given that $X = x$.

The interpretation of $f_{Y|X}(y|x)$ is that, for any (appropriately measurable) set $A \subset \mathbb{R}$,

$$P(Y \in A|X=x) = \sum_{y \in A \cap S_{Y|x}} f_{Y|X}(y|x),$$

where $S_{Y|x} := \{y \in \mathbb{R} : f_{Y|X}(y|x) > 0\}$.

For any (appropriately measurable) function $g(y)$, we define the conditional expectation of $g(Y)$ given that $X = x$ as

$$E[g(Y)|X=x] = \sum_{y \in \mathcal{S}_{Y|x}} g(y) f_{Y|X}(y|x).$$

Example (conditional probability mass function and expectation). Suppose that for $\{(x,y)' \in \mathbb{R}^2 : x \in \{0,1\}, y \in \{0,1\}\}$ we have

$$f_{X,Y}(x,y) = (x+2y+1)/10.$$

We have $f_X(x) = (4+2x)/10$ for $x \in \{0,1\}$, so

$$f_{Y|X}(y|0) =$$

$$f_{Y|X}(y|1) =$$

$$E[Y^2|X=0] =$$

Conditional probability density function and expectation. Let $\mathbf{X}$ be a continuous random vector with components X and Y. Let $f_{X,Y}(x,y)$ denote the joint probability density function of X and Y, $f_X(x)$ the marginal probability density function of X, and $f_Y(y)$ the marginal probability density function of Y.

For any $x \in \mathbb{R}$ at which $f_X(x) > 0$, we define

$$f_{Y|X}(y|x) := f_{X,Y}(x,y)/f_X(x)$$

to be the conditional probability density function of Y given that $X = x$.

The interpretation of $f_{Y|X}(y|x)$ is that, for any (appropriately measurable) set $A \subset \mathbb{R}$,

$$P(Y \in A|X = x) = \int_A f_{Y|X}(y|x)\ dy.$$

For any (appropriately measurable) function $g(y)$, we define the conditional expectation of $g(Y)$ given that $X = x$ as

$$E[g(Y)|X = x] = \int_{\mathbb{R}} g(y) f_{Y|X}(y|x)\ dy.$$

Example (conditional probability density function and expectation). Suppose that

$$f_{X,Y}(x,y) = 8xy1_{\{0<x<y<1\}}.$$

We have $f_X(x) = 4(x - x^3)1_{\{0<x<1\}}$, so

$$f_{Y|X}(y|x) =$$

$$P(Y \le 3/4|X = 1/2) =$$

$$E[Y^2|X = 1/2] =$$

Independence. We say that X and Y are independent if

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$$

for any (appropriately measurable) sets $A, B \subset \mathbb{R}$. If $\mathbf{X}$ is a discrete random vector, then as a special case we may take $A := \{x\}$ and $B := \{y\}$ to obtain the probability mass decomposition

$$f_{X,Y}(x, y) = f_X(x)f_Y(y)$$

for any $(x, y)' \in \mathbb{R}^2$. Likewise, the probability density decomposition

$$f_{X,Y}(x, y) = f_X(x)f_Y(y)$$

may be used to characterize independence if $\mathbf{X}$ is a continuous random vector, with the technical caveat that the probability density decomposition is not required for all $(x, y)' \in \mathbb{R}^2$ but only for $(x, y)' \in C \subset \mathbb{R}^2$ with $P(\mathbf{X} \in C) = 1$.

Example (independence). Suppose that X and Y are independent (and continuous, for simplicity in the calculations to follow). Then, whenever all quantities below exist as finite numbers, we have

$$\begin{aligned}
M_{X+Y}(t) &= E[\exp\{t(X+Y)\}] \\
&= E[\exp\{tX\}\exp\{tY\}] \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} \exp[tx]\exp[ty] f_{X,Y}(x, y)\ dx\ dy \\
&= \int_{\mathbb{R}} \exp[tx] f_X(x)\ dx \int_{\mathbb{R}} \exp[ty] f_Y(y)\ dy \\
&= E[\exp\{tX\}]E[\exp\{tY\}] \\
&= M_X(t)M_Y(t).
\end{aligned}$$

This result can be used to prove that the sum of two independent normal random variables is normal, among some other useful relations that you will establish in Written Assignment 5.