Joint, Marginal, and Conditional Distributions

Problems involving the joint distribution of random variables X and Y use the pdf of the joint distribution, denoted $f_{X,Y}(x, y)$. This pdf is usually given, although some problems only give it up to a constant. The methods for solving problems involving joint distributions are similar to the methods for single random variables, except that we work with double integrals and 2-dimensional probability spaces instead of single integrals and 1-dimensional probability spaces. We illustrate these methods by example.

Discrete Case: Analogous to the discrete single random variable case, we have

$$0 \leq f_{X,Y}(x, y) = \Pr((X = x) \wedge (Y = y)) \leq 1$$

The Continuous Case is illustrated with examples.

The Mixed Case (one of the random variables is discrete, the other is continuous) is also illustrated with examples.

cdf of Joint Distribution – denoted $F_{X,Y}(x, y)$

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

Notice that we can get the cdf's for X and Y from the joint cdf as follows:

$$F_X(x) = F_{X,Y}(x, \infty)$$
$$F_Y(y) = F_{X,Y}(\infty, y)$$

cdf of Joint Distribution (continued)

Discrete Case: We have $F_{X,Y}(x,y) = \sum_{s=-\infty}^{x}\sum_{t=-\infty}^{y} f_{X,Y}(s,t)$

Continuous Case: We have
$$F_{X,Y}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(s,t)\,dt\,ds$$
$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y)$$

General Expectation using Joint Distribution

Discrete Case: We have $E[h(X,Y)] = \sum_{x}\sum_{y} h(x,y)\cdot f_{X,Y}(x,y)$

Continuous Case: We have $E[h(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)\cdot f_{X,Y}(x,y)\,dy\,dx$

The covariance of random variables X and Y is a generalization of variance of a single random variable

$$Cov(X,Y) = Cov(Y,X) = E[(X-\mu_X)\cdot(Y-\mu_Y)] = E[X\cdot Y] - E[X]\cdot E[Y]$$

Remarks:

1. $Var(X) = Cov(X,X)$
2. $Cov(aX+bY, cZ+d) = a\cdot c\cdot Cov(X,Z) + b\cdot c\cdot Cov(Y,Z)$ where a, b, c, and d are constants and X, Y, and Z are random variables
3. $Var(X+Y) = Cov(X+Y, X+Y) = Var(X) + Var(Y) + 2\cdot Cov(X,Y)$

The mgf for Joint Distribution is a generalization of mgf for a single random variable

$$M_{X,Y}(s,t) = E[e^{s \cdot X + t \cdot Y}]$$

Remarks:

1. $E[X^n \cdot Y^m] = \frac{\partial^{n+m}}{(\partial^n s)(\partial^m t)} M_{X,Y}(s,t)\,|_{(s,t)=(0,0)}$

   Important special case are

   $E[X \cdot Y] = \frac{\partial^2}{\partial s \partial t} M_{X,Y}(s,t)\,|_{(s,t)=(0,0)}$, $E[X] = \frac{\partial}{\partial s} M_{X,Y}(s,t)\,|_{(s,t)=(0,0)}$

   and $E[Y] = \frac{\partial}{\partial t} M_{X,Y}(s,t)\,|_{(s,t)=(0,0)}$

2. $M_X(t) = M_{X,Y}(t,0)$ and $M_Y(t) = M_{X,Y}(0,t)$

   So we also have $E[X] = \frac{d}{dt} M_{X,Y}(t,0)\,|_{t=0}$ and $E[Y] = \frac{d}{dt} M_{X,Y}(0,t)\,|_{t=0}$

Marginal Distributions (pdfs) of *X* and *Y*

Discrete Case: We have
$$f_X(x) = \sum_y f_{X,Y}(x,y)$$
$$f_Y(y) = \sum_x f_{X,Y}(x,y)$$

Continuous Case: We have
$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$$
$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx$$

Notice that in order to get the marginal distribution of *X*, we sum out (discrete case) or integrate out (continuous case) *Y*. Similarly, for getting the marginal distribution for *Y*, we sum out (discrete case) or integrate out (continuous case) *X*.

Conditional Distributions for *X*, given *Y* and for *Y*, given *X*

$$f_{X|Y=y}(x \mid y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$
$$f_{Y|X=x}(y \mid x) = \frac{f_{X,Y}(x, y)}{f_X(x)}$$

One way to remember these is by saying the words: the conditional distribution is the joint distribution divided by the marginal distribution. Also notice the probability interpretation when *X* and *Y* are discrete.

Independence of the jointly distributed random variables *X* and *Y*

If *X* and *Y* are independent, then each of the following will be true:

1. $f_{X|Y=y}(x \mid y) = f_X(x)$ and $f_{Y|X=x}(y \mid x) = f_Y(y)$
   and therefore we have $f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)$
   (Think of the probability interpretation in the discrete case.)

2. $E[X \cdot Y] = E[X] \cdot E[Y]$ and more generally,
   $E[h(X) \cdot g(Y)] = E[h(X)] \cdot E[g(Y)]$

3. $Cov(X, Y) = 0$

4. $Var(X + Y) = Var(X) + Var(Y)$

Double Expectation Theorem (Very Important and Useful)

$$E[X] = E[E[X \mid Y]]$$
$$Var(X) = E[Var(X \mid Y)] + Var(E[X \mid Y])$$