 – inc. Conditional

Example 3.7-2. Let X and Y have the joint p.d.f.

$$f(x, y) = 2, \qquad 0 \le x \le y \le 1.$$

Then $R = \{(x, y): \ 0 \le x \le y \le 1\}$ is the support and, for illustration,

$$P\left(0 \le X \le \frac{1}{2}, 0 \le Y \le \frac{1}{2}\right) = P\left(0 \le X \le Y, 0 \le Y \le \frac{1}{2}\right)$$

$$= \int_0^{1/2} \int_0^y 2\,dx\,dy = \int_0^{1/2} 2y\,dy = \frac{1}{4}.$$

The shaded region in Figure 3.7-2 is the region of integration that is a subset of R, and the given probability is the volume above that region under the surface $z = 2$. The marginal p.d.f.'s are given by

$$f_1(x) = \int_x^1 2\,dy = 2(1 - x), \qquad 0 \le x \le 1,$$

and

$$f_2(y) = \int_0^y 2\,dx = 2y, \qquad 0 \le y \le 1.$$

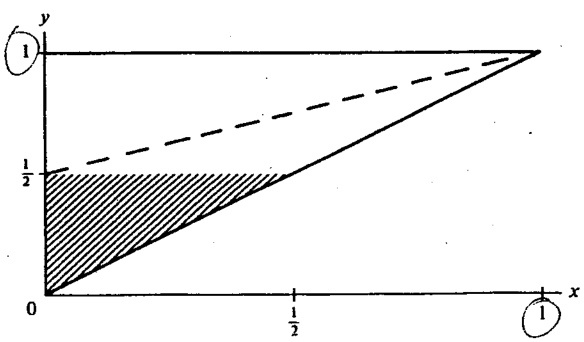

FIGURE 3.7-2

Four illustrations of expected values are

$$E(X) = \int_0^1 \int_x^1 2x\,dy\,dx = \int_0^1 2x(1 - x)\,dx = \frac{1}{3},$$

$$E(Y) = \int_0^1 \int_0^y 2y\,dx\,dy = \int_0^1 2y^2\,dy = \frac{2}{3},$$

$$E(Y^2) = \int_0^1 \int_0^y 2y^2\,dx\,dy = \int_0^1 2y^3\,dy = \frac{1}{2},$$

and

$$\mathrm{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] = E(XY) - \mu_X\mu_Y$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y)\,dx\,dy - \left(\frac{1}{3}\right)\left(\frac{2}{3}\right)$$

$$= \int_0^1 \int_0^y 2xy\,dx\,dy - \frac{2}{9} = \frac{1}{4} - \frac{2}{9} = \frac{1}{36}.$$

From these calculations it is obvious that $E(X)$, $E(Y)$, and $E(Y^2)$ could be calculated using the marginal p.d.f.'s as well as the joint one.

Conditional distributions:

Let X and Y have a distribution of the continuous type with joint p.d.f. $f(x, y)$ and marginal p.d.f.'s $f_1(x)$ and $f_2(y)$, respectively. So in accord with our policy of transition from the discrete to the continuous case, we have that the conditional p.d.f., mean, and variance of Y, given $X = x$, are, respectively,

$$h(y|x) = \frac{f(x, y)}{f_1(x)}, \qquad \text{provided } f_1(x) > 0,$$

$$E(Y|x) = \int_{-\infty}^{\infty} yh(y|x)\,dy,$$

and

$$\mathrm{Var}(Y|x) = E\{[Y - E(Y|x)]^2|x\}$$

$$= \int_{-\infty}^{\infty} [y - E(Y|x)]^2 h(y|x)\,dy$$

$$= E[Y^2|x] - [E(Y|x)]^2.$$

Similar expressions are associated with the conditional distribution of X, given $Y = y$.

Example 3.7-3. Let X and Y be the random variables of Example 3.7-2. Thus

$$f(x, y) = 2, \qquad 0 \le x \le y \le 1,$$
$$f_1(x) = 2(1 - x), \qquad 0 \le x \le 1,$$

and

$$f_2(y) = 2y, \qquad 0 \le y \le 1.$$

Before we actually find the conditional p.d.f. of Y, given $X = x$, we shall give an intuitive argument. The joint p.d.f. is constant over the triangular region shown in Figure 3.7-2. If the value of X is known, say $X = x$, then the possible values of Y are between x and 1. Furthermore we would expect Y to be uniformly distributed on the interval $[x, 1]$. That is, we would anticipate that $h(y|x) = 1/(1-x)$, $x \le y \le 1$. More formally now, we have by definition that

$$h(y|x) = \frac{f(x, y)}{f_1(x)} = \frac{2}{2(1-x)}, \qquad x \le y \le 1, \quad 0 \le x \le 1.$$

The conditional mean of Y, given $X = x$, is

$$E(Y|x) = \int_x^1 y \frac{1}{1-x}\, dy = \left[\frac{y^2}{2(1-x)}\right]_x^1 = \frac{1+x}{2}, \qquad 0 \le x \le 1.$$

Note that, for a given x, the conditional mean of Y given x lies on the dotted line in Figure 3.7-2, a result that also agrees with our intuition.

In general, $\boxed{\text{if}}$ $E(Y|x)$ is linear, it is equal to

$$E(Y|x) = \mu_Y + \rho\left(\frac{\sigma_Y}{\sigma_X}\right)(x - \mu_X).$$

We ~~could have~~ Can calculate the **correlation coefficient** directly from the definition

$$\rho = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}.$$

In Example 3.7-2 we showed that $\text{Cov}(X, Y) = 1/36$. We also found $E(Y)$ and $E(Y^2)$ so that $\sigma_Y^2 = E(Y^2) - [E(Y)]^2 = 1/2 - (2/3)^2 = 1/18$. ~~Since~~ It turns out that $\sigma_X^2 = \sigma_Y^2$, so

$$\rho = \frac{\frac{1}{36}}{\sqrt{\frac{1}{18}}\sqrt{\frac{1}{18}}} = \frac{1}{2}.$$

An illustration of a computation of a conditional probability is

$$P\left(\frac{3}{4} \le Y \le \frac{7}{8} \,\middle|\, X = \frac{1}{4}\right) = \int_{3/4}^{7/8} h\left(y \,\middle|\, \frac{1}{4}\right) dy$$

$$= \int_{3/4}^{7/8} \frac{1}{3/4}\, dy = \frac{1}{6}.$$

Exercises

3.7-6. Let $f(x, y) = 1/20$, $x \le y \le x + 2$, $0 \le x \le 10$, be the joint p.d.f. of X and Y.

(a) Sketch the region for which $f(x, y) > 0$, that is, the support.

(b) Find $f_1(x)$, the marginal p.d.f. of X.

(c) Find $h(y|x)$, the conditional p.d.f. of Y, given $X = x$.

(d) Find the conditional mean ~~and variance~~ of Y, given $X = x$.

(e) Find $f_2(y)$, the marginal p.d.f. of Y. — Hint - this is something like Ash, p249 #3a.

(f) Find $P(Y>6|X=5)$, $P(Y>6|X=2)$, $P(Y>6|X=10)$

B. For the example 3.7-2,

a) Find σ_X^2

b) Plug in $\mu_X, \mu_Y, \rho, \sigma_X, \sigma_Y$ into $E(Y|x) = \mu_Y + \rho\left(\frac{\sigma_Y}{\sigma_X}\right)(x-\mu_X)$ and show it $= \frac{1+x}{2}$, $0 \le x \le 1$

C. $f(x,y) = (x+y)$, $0 \le x \le 1$, $0 \le y \le 1$ (the cardboard model

Find $f(y|x)$. Find $E(Y|x)$. (It will not be linear!)

Ash, p249, #2, #5a, #3