Lecture 18 : Pairs of Continuous Random Variables

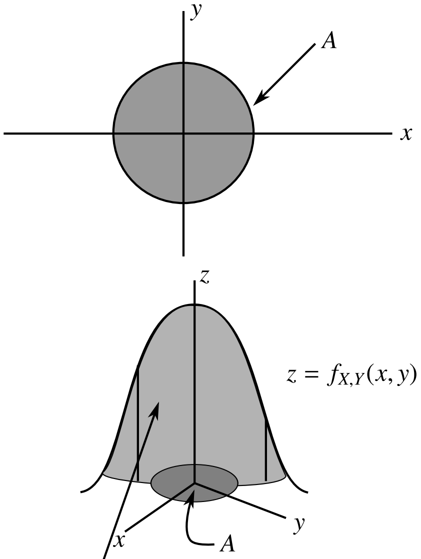
Definition

Let X and Y be continuous random variables defined on the same sample space S. Then the joint probability density function, joint pdf, $f_{X,Y}(x,y)$ is the function such that

$$P((X,Y) \in A) = \underbrace{\iint\limits_A f_{X,Y}(x,y)dx\,dy}_{\text{double integral}} \qquad (*)$$

for any region A in the plane.

Again the geometric interpretation of (*) is very important



$P((X, Y) \in A)$ = the <u>volume</u> under the graph of f and above the region A.

For $f(x, y)$ to be a joint *pdf* for some pair of random variables X and Y it is necessary and sufficient that

$$f(x, y) \geq 0, \quad \text{all} \quad x, y$$

and

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx\, dy = 1$$

or geometrically, the total volume under the graph of f has to be 1.

Example 5.3 (from text)

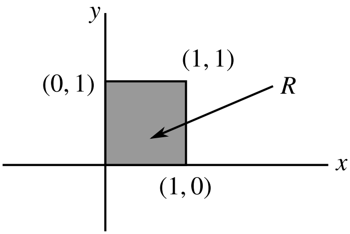
A bank operates a drive-up window and *a* walkup window. On a randomly selected day, let

X = proportion of time the drive-up facilty is in use.

Y = proportion of time the walk-up facilty is in use.

The set of possible outcomes for the pair (X, Y) is the square

$$R = \{(x, y),\ 0 \le x \le 1,\ 0 \le y \le 1\}$$



Suppose the joint *pdf* of (X, Y) is given by

$$f_{x,y}(x,y) = \begin{cases} 6/5(x+y^2), & \begin{array}{l} 0 \le x \le 1 \\ 0 \le y \le 1 \end{array} \\ 0, & \text{otherwise} \end{cases}$$

Find the probability that neither facilty is in use more than 1/4 of the time.

Solution

Neither facilty is in use more than $\frac{1}{4}$ of the time when re-expressed in terms of X and Y is

$$X \le \frac{1}{4} \left(\textit{the drive-up facilty is in use} \le \frac{1}{4} \textit{ of the time}\right)$$

and

$$Y \le \frac{1}{4} \left(\textit{the walk-up facilty is in use} \le \frac{1}{4} \textit{ of the time}\right)$$

Solution (Cont.)

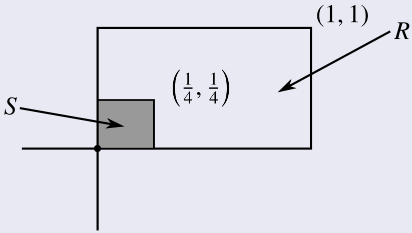
The author formulated the problem in a confusing fashion, don't worry about it. So we want

$$P\left(0 \le X \le \frac{1}{4},\ 0 \le Y \le \frac{1}{4}\right)$$

or

$$P((X, Y) \in S)$$

where S is the small square



This probability is given by

$$\int_0^{\frac{1}{4}} \int_0^{\frac{1}{4}} \frac{6}{5}(x + y^2)dx\ dy$$

$$\iint_S \frac{6}{5}(x + y^2)dx\ dy \qquad (\sharp)$$

Remark

For general (X, Y) we have

$$P(a \leq X \leq b,\ c \leq Y \leq d)$$

$$= \int_a^b \int_c^d f_{X,Y}(x, y) dx\, dy$$

Let's do the integral $(\sharp)$. We will do the x-integration first. So

$$P\left(0 \leq X \leq \frac{1}{4},\ 0 \leq Y \leq \frac{1}{4}\right)$$

$$= \int_0^{\frac{1}{4}} \left(\int_0^{\frac{1}{4}} \frac{6}{5}(x + y^2) dx \right) dy$$

$$= \frac{6}{5} \int_0^{\frac{1}{4}} \left(\frac{X^2}{2} + xy^2 \right) \Big|_{x=0}^{x=\frac{1}{4}} dy$$

Remark (Cont.)

$$= \frac{6}{5} \int_0^{\frac{1}{4}} \left(\frac{1}{32} + \frac{y}{4} \right) dy$$

$$= \frac{6}{5} \left[\left(\frac{y}{32} + \frac{y^3}{12} \right) \Big|_{y=0}^{y=\frac{1}{4}} \right]$$

$$= \frac{6}{5} \left[\frac{1}{128} + \frac{1}{(64)(12)} \right]$$

$$= \left(\frac{6}{5} \right) \left(\frac{1}{64} \right) \left(\frac{1}{2} + \frac{1}{12} \right)$$

$$= \left(\frac{\cancel{6}}{5} \right) \left(\frac{1}{64} \right) \left(\frac{7}{\frac{\cancel{12}}{2}} \right)$$

$$= \frac{7}{640}$$

An exercise in the forgotten art of fractions- more of the same later.

More Theory Marginal Distributions in the Continuous Case

Problem

Suppose you know the joint pdf $f_{X,Y}(x,y)$ of (X,Y). How do you find the individual pdf's $f_X(x)$ of X and $f_Y(y)$. The answer is

Proposition

$$\begin{aligned} &\text{(i)}\quad f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy \\ &\text{(ii)}\quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx \end{aligned} \qquad (*)$$

Proposition (Cont.)

The formula (*) is the continuous analogue of the formula for the discrete case. Namely

Discrete Case

$$f_X(x) = \sum_{\text{all}_y} f_{X,Y}(y)$$

Continuous Case

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$$

In the first case we sum away the "extra variable" y and in the second case we integrate it away.
By analogy once again we call $f_X(x)$ and $f_Y(y)$ (obtained via (*)) the marginal densities or marginal *pdf*'s.

Note the $f_X(x)$ and $f_Y(y)$ are the two partial definite integrals of $f_{X,Y}(x,y)$ - see Lecture 16.

Example 5.4

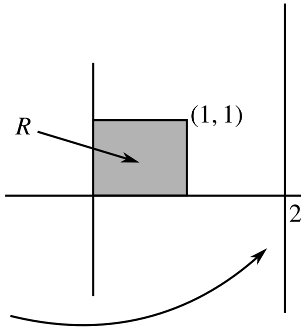
We compute the two marginal *pdf*'s for the bank problem, Example 5.3.

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$$

$$\underset{\text{this is a little tricky.}}{=} \begin{cases} \int_0^1 \frac{6}{5}(x+y^2)dy, & 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$$

The formula for $f_X(x)$ says you integrate $f_{X,Y}(x,y)$ over the vertical

line passing through x.
If x does not satisfy $0 \leq x \leq 1$ then the vertical line does not pass through the square R where $f_{X,Y}(x,y)$ is non zero



You get $f_X(2)$ by integrating over the line $x = 2$ above which $f_{X,Y}(x,y) = 0$.
Equivalently (without geometry)

$$f_X(2) = \int_{-\infty}^{\infty} f_{X,Y}(2,g)dy = \int_{-\infty}^{\infty} 0 \; dy = 0$$

Now we finish the job

$$\int_0^1 \frac{6}{5}(x+y^2)dy = \frac{6}{5}\int_0^1 (x+y^2)dy$$

$$= \frac{6}{5}\left(xy + \frac{y^3}{3}\right)\Big|_{y=0}^{y=1} = \frac{6}{5}\left(x + \frac{1}{3}\right)$$

Similarly

$$f_Y(y) = \begin{cases} \frac{6}{5}\int_0^1 \frac{6}{5}(x+y^2)dx, & 0 \le y \le 1 \\ 0, & \text{otherwise} \end{cases}$$

$$= \begin{cases} \frac{6}{5}y^2 + \frac{3}{5}, & 0 \le y \le 1 \\ 0, & \text{otherwise} \end{cases}$$

Independence of Two Continuous Random Variables

Definition

Two continuous random variables X and Y are independent of their joint pdf $f_{X,Y}(x,y)$ is the product of the two marginal pdf's $f_X(x)$ and $f_Y(y)$ so

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

This not true for the bank example pg. 5.

$$f_{X,Y}(x,y) = {}^{6}\!/\!{}_{5}\underbrace{(x+y^2)}_{\text{not a product}}$$

Covariance and Correlation of Pairs of Continuous Random Variables

We continue with a pair of continuous random variables X and Y as before. Again we define

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

and

$$\rho_{X,Y} = \mathrm{Corr}(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

But now

$$E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy \; f_{X,Y}(x, y) dx \; dy$$

We will now compute the $\mathrm{Cov}(X, Y)$ and $\mathrm{Corr}(X, Y)$ for the bank problem. So

$$f_{X,Y}(x,y) = \begin{cases} \frac{6}{5}(x+y^2), & \begin{array}{c} 0 \le x \le 1 \\ 0 \le y \le 1 \end{array} \\ 0, & \text{otherwise} \end{cases}$$

$$f_X(x) = \begin{cases} \frac{6}{5}\left(x+\frac{1}{3}\right), & 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{6}{5}y^2+\frac{3}{5}, & 0 \le y \le 1 \\ 0, & \text{otherwise} \end{cases}$$

Let's first do the calculations for X and Y - we need

$$E(X), E(Y), \sigma_X = \sqrt{V(X)} \quad \text{and} \quad \sigma_Y = \sqrt{V(Y)}$$

$$E(X) = \int_0^1 x \frac{6}{5}\left(x + \frac{1}{3}\right) dx$$

$$= \frac{6}{5} \int_0^1 \left(x^2 + \frac{x}{3}\right) dx = \frac{6}{5}\left(\frac{x^3}{3} + \frac{x^2}{6}\right)\Big|_{x=0}^{x=1}$$

$$= \frac{6}{5}\left(\frac{1}{3} + \frac{1}{6}\right) = \frac{6}{5}\left(\frac{3}{6}\right) = \frac{3}{5}$$

$$E(X^2) = \int_0^1 x^2 \frac{6}{5}\left(x + \frac{1}{3}\right) dx$$

$$= \frac{6}{5} \int_0^1 \left(x^3 + \frac{x^2}{3}\right) dx = \frac{6}{5}\left(\frac{x^4}{4} + \frac{x^3}{9}\right)\Big|_{x=0}^{x=1}$$

$$= \frac{6}{5}\left(\frac{1}{4} + \frac{1}{9}\right) = \frac{6}{5}\left(\frac{13}{36}\right) = \frac{13}{30}$$

$$V(X) = \frac{13}{30} - \left(\frac{3}{5}\right)^2 = \frac{13}{30} - \frac{9}{25} = \frac{65 - 54}{150} = \frac{11}{150}$$

$$\sigma_X = \sqrt{\frac{11}{150}} = \frac{1}{5}\sqrt{\frac{11}{6}}$$

$$E(Y) = \int_{y}^{1} \left(\frac{6}{5}y^2 + \frac{3}{5}\right) dy$$

$$= \frac{6}{5}\int_{0}^{1} y^3 dy + \frac{3}{5}\int_{0}^{1} y dy$$

$$= \left(\frac{6}{5}\right)\left(\frac{1}{4}\right) + \left(\frac{3}{5}\right)\left(\frac{1}{2}\right) = \frac{6}{20} + \frac{3}{10} = \frac{12}{20}$$

$$E(Y^2) = \int_{0}^{1} y^2 \left(\frac{6}{5}y^2 + \frac{3}{5}\right) dy$$

$$= \frac{6}{5}\int_{0}^{1} y^4 dy + \frac{3}{5}\int_{0}^{1} y^2 dy$$

$$= \left(\frac{6}{5}\right)\left(\frac{1}{5}\right) + \left(\frac{\cancel{3}}{5}\right)\left(\frac{1}{\cancel{3}}\right) = \frac{6}{25} + \frac{??}{5} = \frac{11}{25}$$

$$V(Y) = \frac{11}{25} - \frac{144}{400} = \frac{176}{400} - \frac{144}{400} = \frac{32}{400} = \frac{2}{25}$$

$$\sigma_Y = \sqrt{\frac{2}{25}} = \frac{1}{5}\sqrt{2}$$

Finally we need

$$E(XY) = \int_0^1 \int_0^1 (xy)\frac{6}{5}(x+y^2)dx\,dy$$

$$= \int_0^1 \int_0^1 \underbrace{xy\frac{6}{5}x}_{\text{product function}} dx\,dy + \int_0^1 \int_0^1 \underbrace{xy\frac{6}{5}y^2}_{\text{product function}} dx\,dy$$

$$= \frac{6}{5}\left(\int_0^1 x^2 dx\right)\left(\int_0^1 y\,dy\right) + \frac{6}{5}\left(\int_0^1 x\,dx\right)\left(\int_0^1 y^3 dy\right)$$

$$= \left(\frac{6}{5}\right)\left(\frac{1}{3}\right)\left(\frac{1}{2}\right) + \left(\frac{6}{5}\right)\left(\frac{1}{2}\right)\left(\frac{1}{4}\right)$$

$$= \left(\frac{6}{5}\right)\left(\frac{1}{2}\right)\left(\frac{1}{3}+\frac{1}{4}\right) = \left(\frac{\cancel{6}}{5}\right)\left(\frac{1}{2}\right)\left(\frac{7}{\underset{2}{\cancel{12}}}\right) = \frac{7}{20}$$

Now we can mop the fruits of our labours.

$$
\begin{aligned}
\mathrm{Cov}(X, Y) &= E(XY) - E(X)E(Y) \\
&= \frac{7}{20} - \left(\frac{3}{5}\right)\left(\frac{12}{20}\right) \\
&= \frac{7}{20} - \frac{36}{100} = \frac{35}{100} - \frac{36}{100} \\
\mathrm{Cov}(X, Y) &= \frac{-1}{100} \\
\mathrm{Cov}(X, Y) &= \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\frac{-1}{100}}{\left(\frac{1}{5}\sqrt{\frac{11}{6}}\right)\left(\frac{1}{5}\sqrt{2}\right)} \\
&= \left(\frac{-1}{\cancel{100}}\right)\left(\frac{\cancel{5}}{\sqrt{\frac{11}{\underset{3}{\cancel{6}}}}}\right)\left(\frac{\cancel{5}}{\sqrt{\cancel{2}}}\right) = -\frac{1}{4}\left(\frac{1}{\sqrt{\frac{11}{3}}}\right) = -\frac{\sqrt{3}}{4\sqrt{11}}
\end{aligned}
$$

Independence of Continuous Random Variables

Definition

Two continuous random variables X and Y are independent if the joint pdf is the product of the two marginal pdf's

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

(so $\Longleftrightarrow$ the joint pdf is a product function)
So in Example 5.3, page 4, X and Y are NOT independent.