Two random variables

Joint density function and joint distribution function

The behavior of two random variables X and Y is completely described by their joint (bivariate) density function $f_{XY}(x,y)$

$$P\left(\{a < X \leq b\} \cap \{c < Y \leq d\}\right) = \int_{x=a}^{x=b} \int_{y=c}^{y=d} f_{XY}(x,y)\,dx\,dy$$

It follows from the above definion that the joint density function $f_{XY}(x,y)$ will satisfy

$$f_{XY}(x,y) \geq 0 \quad \forall\,(x,y)$$

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

The joint distribution function $F_{XY}(x,y)$ is the indefinite integral of the joint density function $f_{XY}(x,y)$

$$F_{XY}(x,y) = P\left(\{X \leq x\} \cap \{Y \leq y\}\right)$$

$$= \int_{\xi=-\infty}^{\xi=x} \int_{\eta=-\infty}^{\eta=y} f_{XY}(\xi,\eta)\,d\xi\,d\eta$$

Thus, we have that

$$f_{XY}(x,y) = \frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y}$$

Marginal density functions

If we wish to consider only one of the two random variables X or Y, we can integrate out the independent variable in the joint density function corresponding to the random variable that is not of interest. So we have

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

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)dx$$

It follows immediately that

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

Expectation for two random variables

Similar to what we did for a single random variable, we can define the expectation $E\{g(X,Y)\}$ of any function $g(\cdot,\cdot)$ of the two random variables X and Y

$$E\{g(X,Y)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{XY}(x,y)dxdy$$

In terms of two random variables X and Y, we can express the property of linearity of expectation as

$$E\{ag(X,Y)+bh(X,Y)\} = aE\{g(X,Y)\}+bE\{h(X,Y)\}$$

Now, let us consider some important special cases

Correlation

Here we let $g(x,y) = xy$. We obtain

$$\overline{XY} = E\{XY\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{XY}(x,y)dxdy$$

Covariance

Here we let $g(x,y) = \left(x - \bar{X}\right)\left(y - \bar{Y}\right)$. We obtain

$$\sigma^2{}_{XY} = E\left\{\left(X - \bar{X}\right)\left(Y - \bar{Y}\right)\right\}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(x - \bar{X}\right)\left(y - \bar{Y}\right) f_{XY}(x,y)dxdy$$

By multiplying out the term $\left(X - \bar{X}\right)\left(Y - \bar{Y}\right)$ in $E\left\{\left(X - \bar{X}\right)\left(Y - \bar{Y}\right)\right\}$ and exploiting the property of linearity, we can easily show that

$$\sigma^2{}_{XY} = \overline{XY} - \bar{X}\bar{Y}$$

Correlation Coefficient

The correlation coefficient ρ_{XY} for two random variables X and Y is defined to be the normalized covariance

$$\rho_{XY} = \frac{\sigma_{XY}^2}{\sigma_X \sigma_Y}$$

The correlation coefficient is bounded in magnitude between 0 and 1

$$0 \le \left|\rho_{XY}\right| \le 1$$

Independence of two random variables

Two random variables X and Y are said to be independent if and only if their joint density function $f_{XY}(x,y)$ satisfies

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

It follows directly from the above definitions and relationships that if X and Y are independent, then we have that

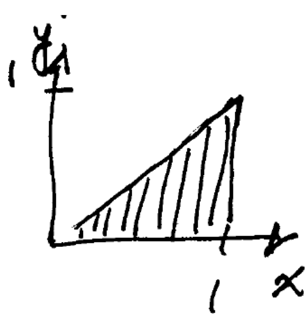
$$\overline{XY} = \overline{X}\,\overline{Y}$$

$$\sigma_{XY}^2 = 0$$

$$\rho_{XY} = 0$$

Note that the converse is not true, i.e. two random variable X and Y can be uncorrelated, i.e. $\overline{XY} = 0$, yet not be independent, i.e. $f_{XY}(x,y) \neq f_X(x) f_Y(y)$.

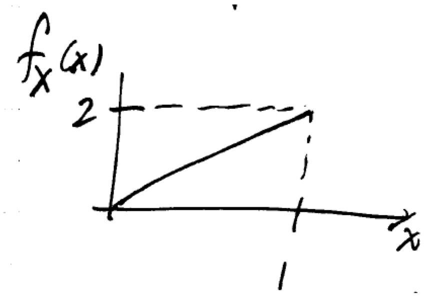
Example



$$f_{XY}(x,y) = \begin{cases} 2, & 0 \le y < x \le 1 \\ 0, & \text{else} \end{cases}$$

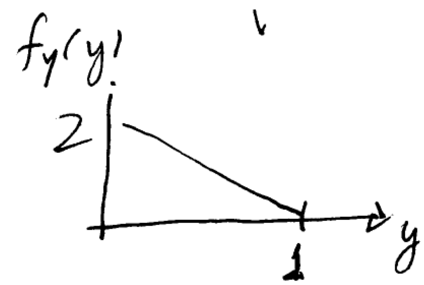
The density function $f_{XY}(x,y)$ satisfies $f_{XY}(x,y) \ge 0 \quad \forall\, (x,y)$ and $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,dx\,dy = 1$ by inspection.

Marginal density functions



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

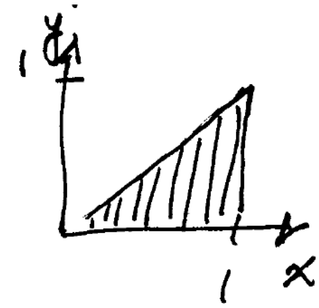
$$= \begin{cases} \int_0^x 2dy = 2x, & 0 \le x \le 1 \\ 0, & \text{else} \end{cases}$$



$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)dx$$

$$= \begin{cases} \int_y^1 2dx = 2(1-y), & 0 \le y \le 1 \\ 0, & \text{else} \end{cases}$$

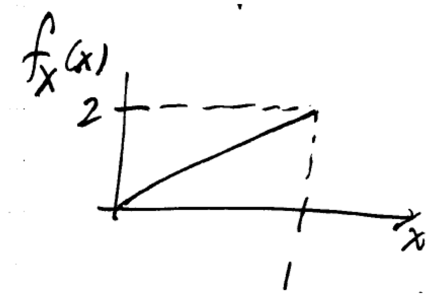
Correlation, Covariance, and Correlation Coefficient



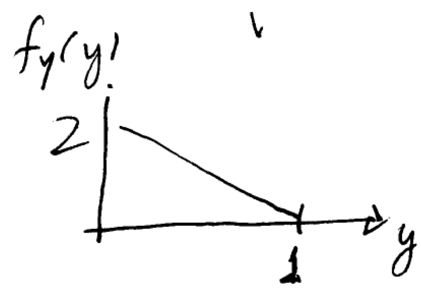
$$\overline{XY} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{XY}(x,y)\,dx\,dy$$

$$= \int_{x=0}^{1}\int_{y=0}^{x} xy2\,dx\,dy = \int_{x=0}^{1} x\left\{\int_{y=0}^{x} 2y\,dy\right\}dx$$

$$= \int_{x=0}^{1} x\left\{2\left.\frac{y^2}{2}\right|_0^x\right\}dx = \int_{x=0}^{1} x^3\,dx = \left.\frac{x^4}{4}\right|_0^1 = \frac{1}{4}$$



$$\bar{X} = \int_{-\infty}^{\infty} x f_X(x)\,dx = \int_0^1 x2x\,dx = 2\left.\frac{x^3}{3}\right|_0^1 = \frac{2}{3}$$



By symmetry, we conclude that $\bar{Y} = \frac{1}{3}$.

$$\overline{X^2} = \int_{-\infty}^{\infty} x^2 f_X(x)dx = \int_0^1 x^2 2x dx = 2\frac{x^4}{4}\bigg|_0^1 = \frac{1}{2}$$

Therefore, $\sigma_X^2 = \overline{X^2} - \left(\overline{X}\right)^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \frac{4}{9} = \frac{9-8}{18} = \frac{1}{18}$.

Again, by symmetry, we conclude that $\sigma_Y^2 = \frac{1}{18}$, also.

Thus, $\sigma_{XY}^2 = \overline{XY} - \overline{X}\overline{Y} = \frac{1}{4} - \left(\frac{2}{3}\right)\left(\frac{1}{3}\right) = \frac{9-8}{36} = \frac{1}{36}$

And, **finally**, $\rho_{XY} = \frac{\sigma_{XY}^2}{\sigma_X \sigma_Y} = \frac{1/36}{(1/18)^{1/2}(1/18)^{1/2}} = \frac{1}{2}$.