Several random variables

Considering systems of only one random variable may not be good enough. We will extend now the concepts we have discussed so far to two random variables. This extension will prove to be enough to carry out the most usual types of analysis in systems that involve an infinite number of random variables.

Definition. Joint probability distribution. Let X and Y be two random variables, we define the joint probability distribution as:

$F(x,y) = Pr(X \leq x, Y \leq y)$

Properties

1) $0 \leq F(x,y) \leq 1$, $-\infty < x < \infty$ $-\infty < y < \infty$

2) $\lim_{x \to -\infty} F(x,y) = \lim_{y \to -\infty} F(x,y) = \lim_{x,y \to -\infty} F(x,y) = 0$

3) $\lim_{x,y \to \infty} F(x,y) = 1$

4) $F(x,y)$ is a non-decreasing function as either x, y or both increase.

5) $\lim_{x \to \infty} F(x,y) = F_Y(y)$, $\lim_{y \to \infty} F(x,y) = F_X(x)$

Definition Joint probability density function. For X and Y continuous random variables $f(x,y) = \dfrac{\partial^2 F(x,y)}{\partial x \partial y}$

Definition. Marginal probability distribution function. $F_X(x) = \lim_{y \to \infty} F(x,y)$

Properties

Continuous	Discrete
1) $f(x,y) \geq 0 \quad -\infty < x < \infty, -\infty < y < \infty$	
2) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$	$\sum_i \sum_j f(x_i, y_j) = 1$
3) $F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v)\,du\,dv$	$F(x,y) = \sum_{u_i \in S}^{x} \sum_{v_j \in S}^{y} f(u_i, v_j)$
4) $f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$	$f_X(x) = \sum_{y_j \in S} f(x, y_j)$
$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$	$f_Y(y) = \sum_{x_i \in S} f(x_i, y)$
5) $P_r(x_1 \leq X \leq x_2,\ y_1 \leq Y \leq y_2) = \int_{x_1}^{x_2} \int_{y_1}^{y_2} f(x,y)\,dx\,dy$	$P_r(X = x, Y = y) = f(x,y)$

Definition. Marginal probability density function. $f_X(x) = \dfrac{\partial F(x,y)}{\partial x} = \int_{-\infty}^{\infty} f(x,y)\,dy$

Example. A fair coin is tossed 3 times. X: # of heads on the first toss

Y: total number of heads $\quad S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$

X \ Y	0	1	2	3	
0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	0	$\frac{1}{2}$
1	0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{1}{2}$
	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	

$f(0,0) = \frac{1}{8}$ $\quad f_X(0) = \frac{1}{2}$ $\quad f_Y(0) = \frac{1}{8}$

$f(0,1) = \frac{2}{8}$ $\quad f_X(1) = \frac{1}{2}$ $\quad f_Y(1) = \frac{3}{8}$

$\vdots$ $\quad f_Y(2) = \frac{2}{8}$

$f(1,3) = \frac{1}{8}$ $\quad f_Y(3) = \frac{1}{8}$

Example. Consider the manufacture of rectangular semiconductor substrates. The values of each one of the two dimensions X and Y might be random variables uniformly distributed between certain values (x_1, x_2) and (y_1, y_2) thus:

$$f(x,y) = \frac{1}{(x_2-x_1)(y_2-y_1)}$$

$$f_X(x) = \int_{y_1}^{y_2} \frac{1}{(x_2-x_1)(y_2-y_1)} dy = \frac{1}{(x_2-x_1)(y_1-y_2)} y \Big|_{y_1}^{y_2} = \frac{1}{x_2-x_1}$$

$$f_Y(y) = \int_{x_1}^{x_2} \frac{1}{(x_2-x_1)(y_2-y_1)} dx = \frac{1}{y_2-y_1}$$

Similarly to the one dimensional case, expected values can be computed as:

$$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f(x,y)\, dx\, dy$$

<u>Multinomial</u> n independent trials of an experiment can result in r types of outcomes with probabilities $p_1, p_2, \ldots, p_r$ respectively. Let N_i: total number of outcomes of type $i = 1, 2, \ldots, r$, then

$$Pr(N_1 = n_1, N_2 = n_2, \ldots, N_r = n_r) = f(n_1, n_2, \ldots, n_r) = \binom{n}{n_1\, n_2 \ldots n_r} p_1^{n_1} p_2^{n_2} \ldots p_r^{n_r}$$

$$f_{N_i}(n_i) = \binom{n}{n_i} p_i^{n_i}(1-p_i)^{n-n_i}$$

$$\frac{n!}{n_1!\, n_2! \cdots n_r!}$$

<u>Bivariate normal</u>

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right]\right\}$$

$-\infty < \mu_X, \mu_Y < \infty$, $\sigma_X, \sigma_Y > 0$, $-1 < \rho < 1$