Probability and Statistics
for Final Year Engineering Students

By Yoni Nazarathy, Last Updated: May 1, 2011.

Lecture 4 – part B: Several Random Variables

Note Lecture 4 is divided into Part A: Joint Distributions, Covariance, Correlation and Part B: Multinomial Distribution, Conditional Distributions, Joint Gaussian Distribution

Brief review of part A:

In part A we saw that the joint distribution of X and Y is

$p(x, y) = P(X = x, Y = y)$ for discrete RVS, or

$f(x, y)dxdy \approx P(X \in [x, x + dx], Y \in [y, y + dy])$ for continuous RVs (f is called the density).

We also defined the covariance:
$Cov(X, Y) = E[(X - E[X])(Y - E[Y])] = \cdots = E[XY] - E[X]E[Y]$, noted that,

$E[XY] = E[X]E[Y]$ (when independent)

And that,

$$Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y).$$

We also defined the correlation coefficient:

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}}.$$

Marginal distributions:

If we know the joint distribution we can get the distribution of one of the coordinates by summing (integrating) over the other. This is called the marginal distribution.

$p(x) = \sum_y p(x, y),\ p(y) = \sum_x p(x, y).$

$f(x) = \int f(x, y)dy,\ f(y) = \int f(x, y)dx$

The Multinomial Distribution:

Recall first the binomial distribution. $X \sim Bin(n, p)$ let Y=n-X, what is the joint distribution of X and Y?

Answer: $p(x, y) = \frac{n!}{x!y!} p^x (1 - p)^y$ for $0 \le x, y$ and $x + y = n$.

This is a special case of the multinomial distribution where there are two outcomes for each experiment, with probabilities $p_1 = p$ and $p_2 = 1 - p$.

This can be generalized to a case of M possible outcomes in a set of n indepdendent experiments. Denote, $p_1 + \cdots + p_M = 1$. (Binomial is the case of M=2). Then the probability of k_1 events of type 1 (each occurring with probability p_1), k_2 events of type 2, etc… is given by,

$$\frac{n!}{k_1!\,k_2!\ldots k_M!} {p_1}^{k_1} {p_2}^{k_1} \ldots . {p_M}^{k_1}.$$

Note this is the joint distribution of M random variables.

What do you think is the marginal distribution?

Conditional distributions:

The conditional distribution of X given Y gives the conditional probabilities, P(X=x | Y=y) (in the case of discrete RVs). This is the probability of the random variable X getting the value x if we know that Y=y.

We have,

$$p_{X|Y}(x|y) = P(X = x|Y = y) = \frac{p(x,y)}{p(y)} = \frac{p(x,y)}{\sum_x p(x,y)}.$$

Similarly,

$$p_{Y|X}(y|x) = P(Y = y|Y = x) = \frac{p(x,y)}{p(x)} = \frac{p(x,y)}{\sum_y p(x,y)}.$$

For continuous random variables:

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f(y)} = \frac{f(x,y)}{\int f(x,y)dx}$$

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f(x)} = \frac{f(x,y)}{\int f(x,y)dy}$$

Example – a simple router:

Example needs to be filled in.

The Bivariate Gaussian Distribution:

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2(\sqrt{1-\rho^2}} \exp\{-\frac{1}{2(1-\rho^2)}\left(\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right)\}$$

The means are: μ_1, μ_2

The variances are σ_1^2, σ_2^2.

The parameter is ρ is the correlation coefficient (this can be shown).

The PDF is centered around (μ_1, μ_2) and is constant for values where

$$\left(\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right) = constant.$$

The Marginal PDFs are simply those of Normal(μ_i, σ_i^2) $i = 1,2$.

The conditional PDF is as follows:

$$f_{X_1|X_2}(x_1|x_2) = \frac{1}{\sqrt{2\pi\sigma_1(1-\rho^2)}} \exp\{-\frac{1}{2(1-\rho^2)\sigma_1^2}\left(x - \rho\frac{\sigma_1}{\sigma_2}(x_2-\mu_2) - \mu_1\right)^2\}$$

So this shows that if we know that $X_2 = x_2$ then $X_1 \sim Normal(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(x_2-\mu_2), \quad (1-\rho^2)\sigma_1^2)$.