M1S TUTORIAL SHEET : WEEK 11

JOINT PROBABILITY DISTRIBUTIONS

Ideas developed to enable specfication of probability distributions for single random variables can be extended to model situations in which more than one random variable are present. We will only study the extension to two random variables, but the ideas generalize to higher dimensional problems.

DISCRETE CASE

If X and Y are two discrete random variables, then the **joint probability mass function** of X and Y is written $f_{X,Y}$, and defined by

$$f_{X,Y}(x,y) = \mathrm{P}[\ X = x,\ Y = y\ ] = \mathrm{P}[\ (X = x) \cap (Y = y)\ ]$$

for real-values x and y. The probability axioms dictate that $f_{X,Y}$ satisfies

$$\text{(i) } f_{X,Y}(x,y) \geq 0, \forall x, y \qquad \text{(ii) } \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} f_{X,Y}(x,y) = 1.$$

This joint specfication describes the probability of various *pairs* of outcomes (x,y). The **joint cumulative distribution function**, $F_{X,Y}$, is defined by

$$F_{X,Y}(x,y) = \mathrm{P}[\ X \leq x,\ Y \leq y\ ] = \mathrm{P}[\ (X \leq x) \cap (Y \leq y)\ ]$$

and calculated by double summation of the joint mass function over appropriate ranges.

This joint distribution *automatically* specifies the **marginal** probability distribution of each variable individually. The **marginal mass functions** of X and Y, f_X and f_Y, are defined respectively by

$$f_X(x) = \sum_{y=-\infty}^{\infty} f_{X,Y}(x,y) \qquad f_Y(y) = \sum_{x=-\infty}^{\infty} f_{X,Y}(x,y)$$

which follow directly from the Theorem of Total Probability.

Finally, we can also consider **conditional** probability distributions. Using the definition of conditional probability, it is possible to define the **conditional mass functions** of X given Y and Y given X, $f_{X|Y}$ and $f_{Y|X}$ by

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \qquad f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

whenever $f_X(x) > 0$ and $f_Y(y) > 0$. These distributions describe the probability distribution for one variable, given that the variable is fixed equal to a specfied value.

In the discrete case, joint distributions are commonly displayed in tables, with rows defining the possible values of X and columns defining the possible values of Y. In this context, the marginal mass functions are obtained as the row and column sums, and the conditional mass functions are obtained as individual rows or columns, re-scaled to sum to one.

CONTINUOUS CASE
For continuous random variables, the **joint probability density function** of X and Y $f_{X,Y}$, and the **joint cumulative distribution function**, $F_{X,Y}$, are defined by

$$F_{X,Y}(x,y) = \mathrm{P}[\ X \leq x,\ Y \leq y\ ] = \int_{-\infty}^{y}\int_{-\infty}^{x} f_{X,Y}(s,t)\ ds\ dt$$

for real-values x and y. The probability axioms dictate that $f_{X,Y}$ satisfies

$$\text{(i) } f_{X,Y}(x,y) \geq 0, \forall x,y \qquad \text{(ii) } \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\ dx\ dy = 1.$$

Again, the joint distribution specifies the marginal and conditional probability distributions of X and Y; the **marginal density functions** of X and Y, f_X and f_Y, are defined respectively by

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

and the **conditional density functions** of X given Y and Y given X, $f_{X|Y}$ and $f_{Y|X}$ by

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \qquad f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

whenever $f_X(x) > 0$ and $f_Y(y) > 0$.

INDEPENDENCE
X and Y are **independent** if, for all values of x and y we have

$$F_{X,Y}(x,y) = F_X(x)F_Y(y) \qquad \text{or, equivalently} \qquad f_{X,Y}(x,y) = f_X(x)f_Y(y)$$

EXPECTATIONS IN JOINT DISTRIBUTIONS : COVARIANCE
It is possible to consider expectations of functions of pairs of random variables, as an extension to the concept of expectations for single variables. For a bivariate, real-valued function g, the expectation of $g(X,Y)$ is defined (in the continuous case) by

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

A special case $g(s,t) = (s-\mu_X)(t-\mu_Y)$, where μ_X and μ_Y are the (marginal) expectations of X and Y respectively, defines the **covariance** of X and Y, that is

$$\begin{aligned}
\mathrm{Cov}_{f_{X,Y}}[X,Y] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_X)(y-\mu_Y) f_{X,Y}(x,y)\ dx\ dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{X,Y}(x,y)\ dx\ dy - \mu_X\mu_Y = \mathrm{E}_{f_{X,Y}}[XY] - \mathrm{E}_{f_X}[X]\mathrm{E}_{f_Y}[Y]
\end{aligned}$$

The covariance is a *measure of association* between X and Y, and is related to the **correlation** between X and Y by

$$\mathrm{Corr}_{f_{X,Y}}[X,Y] = \frac{\mathrm{Cov}_{f_{X,Y}}[X,Y]}{\sqrt{\mathrm{Var}_{f_X}[X]\mathrm{Var}_{f_Y}[Y]}}$$