Handout 12: Multivariate Models

While a single random variable might be the sole focus in a particular experiment, there are many occasions in which two or more variables of interest occur together and are recorded simultaneously. This handout investigates those cases where we have multiple random variables defined over a single probability space.

Basic definitions

For discrete random variables X and Y the joint probability distribution is given by the function $p_{X,Y}(x,y)$. This gives the probability mass of observing $X = x$ and $Y = y$. The marginal probability mass functions are given by:

$$p_X(x) = \sum_{\text{all} \quad y} p_{X,Y}(x,y)$$
$$p_Y(y) = \sum_{\text{all} \quad x} p_{X,Y}(x,y)$$

And, similarly to the definition of conditional probabilities, we have the concept of a conditional distribution:

$$p_{X|Y}(x|y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

Notice that this suggest a very useful functional form for defining $p_{X,Y}(x,y)$:

$$p_{X,Y}(x,y) = p_{X|Y}(x|y) \cdot p_Y(y)$$

So we can construct a joint density by first specifying the marginal density on Y and then describing X in relationship to Y. For example we could use something like this:

$$Y \sim Beta(\alpha, \beta)$$
$$X|Y \sim Binomial(1, Y)$$

This type of model is very useful in statistical applications. The conditional distribution also leads to the analogous definition of independence and of Bayes Theorem. For Bayes, we have:

$$f(X|Y) = \frac{f(Y|X) \cdot f(X)}{f(Y)}$$

Two random variables are said to be independent if, for all values of x and y:

$$p_{X,Y}(x,y) = p_X(x) \cdot p_Y(y).$$

The definitions and theorems for expected values apply over the joint, marginal, and conditional probabilities as these are in fact all valid random variables. Also, the exact same definitions holds for probability density functions by replacing summations with integral signs.

Expectation and Variance Rules

The expectation of a sum is always equal the sum of expectations:

Theorem 1 *Suppose that a random vector $X = (X_1, X_2, \ldots, X_n)$ has a density or mass function $f_{x_1,\ldots,x_n}$. For arbitrary real numbers $a_1, \ldots, a_n$:*

$$\mathbb{E}\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i \cdot \mathbb{E}X_i$$

Variances add only if the variables are independent.

Theorem 2 *Suppose that a random vector $X = (X_1, X_2, \ldots, X_n)$ has a density or mass function $f_{x_1,\ldots,x_n}$, and each X_i is independent of all other X_j's. For arbitrary real numbers $a_1, \ldots, a_n$:*

$$Var\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 \cdot Var(X_i)$$

In the bivariate case, the following will also be very useful:[1]

Theorem 3 *For any random variables X and Y we have:*

$$\mathbb{E}X = \mathbb{E}(\mathbb{E}(X|Y))$$

The proofs of all of these theorems boil down to statements about the linearity of the integral and summation signs.

[1] I have found this result sometimes troubles new probability students, but it is just essentially saying that we can compute a double integral but first computing the inner integral and then calculating the outer one; this is something you should have seen a lot of multivariate calculus.