Chapter 10 – Two or More Random Variables

Often, variables (ambition and income) are related, and this must be taken into account and quantified.

Example A – 5 Coin Tosses: Let X be the number of Heads in the 5 tosses, and Y be the number of changes in the sequence. For example, in the sequence 'THHTT', there is Y = 2 changes in the sequence, and in the sequence 'HTHTH', Y = 4. Then the separate (or so-called *marginal*) distributions of X and Y are the following:

x	0	1	2	3	4	5
f(x)	1/32	5/32	10/32	10/32	5/32	1/32

y	0	1	2	3	4
f(y)	2/32	8/32	12/32	8/32	2/32

Deriving the distribution for X is easy since X ~ Binomial($n=5, \pi=\frac{1}{2}$); for Y, we list the possibilities (p.269) and count so as to obtain the probabilities. Since X is Binomial, we know $\mu_X = E(X) = 2.5$ and $\sigma_X^2 = Var(X) = 1.25$. Also,

$$\mu_Y = E(Y) = (1/32)*\{1*8 + 2*12 + 3*8 + 4*2\} = 64/32 = 2$$

$$E(Y^2) = (1/32)*\{1^2*8 + 2^2*12 + 3^2*8 + 4^2*2\} = 160/32 = 5, \text{ so } \sigma_Y^2 = 1$$

For this example, the table of *joint probabilities* is the following:

		Y					Marginal
		0	1	2	3	4	Prob's.
X	0	1/32	0	0	0	0	1/32
	1	0	2/32	3/32	0	0	5/32
	2	0	2/32	3/32	4/32	1/32	10/32
	3	0	2/32	3/32	4/32	1/32	10/32
	4	0	2/32	3/32	0	0	5/32
	5	1/32	0	0	0	0	1/32
Marginal	Prob's.	2/32	8/32	12/32	8/32	2/32	1

These *joint probabilities* are again obtained by counting from the complete list of the 32 sequences on p.269. We can express the joint probabilities as fractions here (Table 10.1, p.271), or as decimals as in Table 10.2 on p.272. In contrast with the *joint* and *marginal* probabilities above we can also consider *conditional probabilities*: suppose we are told that X = 2, what is the *conditional* probability distribution of Y? We get it from the X = 2 row of the above joint table after dividing through by 10/32:

y	0	1	2	3	4
f(y/X=2)	0	2/10	3/10	4/10	1/10

Clearly, *this* distribution is not the same as the *marginal* distribution for Y, so we say that X and Y are not *<u>independent</u>* (they are therefore *<u>dependent</u>*). More on this later.

<u>Example B</u> – X and Y are two discrete RVs with probabilities:

		Y				Marginal
		1	2	3	4	Probabilities
X	0	0.10	0.20	0.14	0.06	0.50
	1	0.04	0.10	0.22	0.14	0.50
Marginal	Probabilities	0.14	0.30	0.36	0.20	1

Here, X is Bernoulli with $\pi = ½$, so $\mu_X = 0.5$ and $\sigma_X^2 = 0.25$. Also,

- $\mu_Y = \{1*0.14 + 2*0.30 + 3*0.36 + 4*0.20\} = 2.62$
- $E(Y^2) = \{1^2*0.14 + 2^2*0.30 + 3^2*0.36 + 4^2*0.20\} = 7.78$,
- so $\sigma_Y^2 = 7.78 - 2.62^2 = 0.9156$

Clearly, here too the *conditional* probability distributions of Y given X and the *marginal* distribution of Y are not the same.

It goes without saying that the *joint* probabilities sum to one, and that the *marginal* distributions are obtained by summation (see p.271). Also, cdf's (cumulative) are obtained by summing too (p.272).

Departing from theory, Tables 10.4 (p.273) and 10.5 (p.274) provide actual (empirical) probabilities for 15162 families with at least 5 kids with X = # of Girls and Y = # of changes in gender sequence. This is analogous to X = # of H's and Y = # of changes in sequence.

Previously, we had that events E and F are independent IFF

- **Pr(E/F) = Pr(E), or equivalently,**
- **$Pr(E \cap F) = Pr(E)*Pr(F)$**

Here, we have the same results for RVs X and Y:

Random variables X and Y are *independent* IFF • **f(x,y) = f(x)*f(y) for all x and y, or equivalently,** • **f(x/y) = f(x) for all x and y**

This means that all the joint probabilities in the table must be the same as those obtained by multiplying the corresponding row and column marginal probabilities. It follows that X and Y in both above examples are dependent. An 'independent' example is given on p.277.

<u>Definition</u>: For RVs X and Y with joint pmf f(x,y),

$$E\{g(X,Y)\} = \Sigma_x \Sigma_y \, g(x,y)*f(x,y)$$

Therefore, $E(XY) = \Sigma_x \Sigma_y \, xyf(x,y)$

- **For Example B above,**
  E(XY) = (1)(0)(0.10) + (2)(0)(0.20) + (3)(0)(0.14) + (4)(0)(0.06) + (1)(1)(0.04) + (2)(1)(0.10) + (3)(1)(0.22) + (4)(1)(0.14) = 1.46
- **For Example A above, verify that E(XY) = 5**

<u>Definition</u>: For RVs X and Y with joint pmf f(x,y), the *covariance* between X and Y, denoted σ_{XY}, is equal to

$$\sigma_{XY} = Cov(X,Y) = E\{(X - \mu_X)*(Y - \mu_Y)\}$$

By Theorem 10.5, $\sigma_{XY} = E(XY) - \mu_X*\mu_Y$ (Shortcut formula)

It follows that for Example B, σ_{XY} = 1.46 – (0.5)*(2.62) = 0.15 and for Example A, σ_{XY} = 5 – (2)*(2.5) = 0.

By Theorem 10.5, the *correlation* between X and Y, denoted ρ_{XY}, is equal to

$$\rho_{XY} = \text{Corr}(X,Y) = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

So, to find the correlation (coefficient), we first find the covariance, then divide by the square root of the product of the variances. If the covariance is zero, then so is the correlation. So, for Example A, ρ = 0 (we can omit the XY subscript since it is clear here). For Example B, ρ = 0.15 / sqrt{0.25*0.9156} = 0.313522.

<u>Miscellany</u>:

- **Theorem 10.1: E{*} is a linear operator**
- **This is applied in the Corollary, E(K) = nπ**
- **Theorem 10.2: X and Y are independent IFF E(XY) = E(X)*E(Y)**
- **Theorem 10.3: X and Y Independent ➔ they're uncorrelated**
- **NOT vice versa! (So 'Independence' is "stronger")**
- **Correlation is only assessing LINEAR relationships**
- **Linear combinations are discussed in Section 10.4**
- **Var(X + Y) = Var(X) + 2Cov(X,Y) + Var(Y)**
- **Var(X - Y) = Var(X) - 2Cov(X,Y) + Var(Y)**
- **When X and Y are uncorrelated (which could result from their being independent), Var(X + Y) = Var(X - Y) = Var(X) + Var(Y)**