Lecture 5 – Bivariate Data & Least Squares Estimates

Probability Theory

- As we have **two random variables** (X and Y), the outcomes from an experiment could be descried by more than one variable, so we have a **multivariate probability distribution** (e.g. Y = printers sold, X = PCs sold)

Frequency Distribution

	# PCs (X)					
# Printers (Y)	**0**	**1**	**2**	**3**	**4**	***Total Pr(Y)***
0	6	6	4	4	2	22
1	4	10	12	4	2	32
2	2	4	20	10	10	46
3	2	2	10	20	20	54
4	2	2	2	10	30	46
Total Pr(X)	16	24	48	48	64	200

Relative Frequency Distribution

	# PCs (X)					
# Printers (Y)	**0**	**1**	**2**	**3**	**4**	***Total Pr(Y)***
0	0.03	0.03	0.02	0.02	0.01	0.11
1	0.02	0.05	0.06	0.02	0.01	0.16
2	0.01	0.02	0.10	0.05	0.05	0.23
3	0.01	0.01	0.05	0.10	0.10	0.27
4	0.01	0.01	0.01	0.05	0.15	0.23
Total Pr(X)	0.08	0.12	0.24	0.24	0.23	1

  - Their joint probability of occurrence is defined by the joint probability mass (density) function $\Pr[X = x, Y = y]$ (e.g. Pr [X=1, Y=2] = 0.2)
  - Their marginal probability is the probability that X (or Y) assumes a given value regardless of the values taken by Y (or X) (e.g. Pr [X = 1] = 0.12)
- **Conditional Probability Functions** – finding the probability that Y = y, conditional upon that X = x, and vice versa

$$\boldsymbol{Pr[Y = y \mid X = x]}$$

- **Conditional Distribution**
  - *Bayes Theorem* (note: ∩ means 'and'):

$$\boldsymbol{Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}}$$

$$\boldsymbol{Pr[Y = y \mid X = x] = \frac{Pr[Y = y,\ X = x]}{Pr[X = x]}}$$

  - In regression analysis, interest is in studying the behaviour of one variable (dependent) conditional upon the knowledge of another variable (independent) (e.g. the hourly wage conditional upon education)
- **Statistical Independence**: X and Y are statistically independent if the value of X is unrelated to the value of Y
  - *Statistically Independent if*:

$$\boldsymbol{Pr[Y = y \mid X = x] = Pr[Y = y]}$$

$$\boldsymbol{Pr[Y = y \cap X = x] = Pr[Y = y] * Pr[X = x]}$$