Lecture 6: May 12th, 2017

Lecturer: Suryapratim Banerjee *Notes By: Harsh Mistry*

6.1 Relative Risk

$$R.R = \frac{\frac{y_{11}}{(y_{11}+y_{12})}}{\frac{y_{21}}{(y_{21}+y_{22})}}$$

6.2 Correlation Coefficient

The correlation coefficient measures the degree of linear association between X and Y.

$$\text{Sample Correlation Coefficient } = r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left[\sum_{i=1}^{n}(x_i - \bar{x})^2\right]\left[\sum_{i=1}^{n}(y_i - \bar{y})^2\right]}}$$

6.2.1 Properties of the Correlation Coefficient

- $| r_{xy} | \leq 1$
- If $y_i = a + bx_i, \forall i = 1, \ldots, n$

$$r_x y = \begin{cases} 1 \text{ if } b > 0 \\ -1 \text{ if } b < 0 \end{cases}$$

- The sign of the value determines direction
- the value represents the strength of the linear relationship
- Higher values of $| r |$ represent a stronger linear association
- $| r | \approx 0$ indicates evidence of no linear association
- Ever if $| r |$ is really high, we cannot conclude Causation without further analysis