CORRELATION & REGRESSION

Correlation

(1) $Co-variance$

$$Cov(x,y) = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)\left(\frac{1}{n}\sum_{i=1}^{n} y_i\right)$$

(2) ${\sigma_x}^2 = \frac{1}{n}\sum {x_i}^2 - \left(\frac{1}{n}\sum x_i\right)^2$

(3) ${\sigma_y}^2 = \frac{1}{n}\sum {y_i}^2 - \left(\frac{1}{n}\sum y_i\right)^2 \quad [where,\ \sigma = s.d.]$

(4) $Correlation\ co-efficient,$

$r_{xy} = \dfrac{Cov(x,y)}{\sigma_x \sigma_y}$

(5) $For\ any\ bi-variate\ data,\ -1 \le r_{xy} \le 1$

(6) $r_{xx} = 1\ \&\ r_{x(-x)} = -1$

(7) $If\ u = ax + b\ \&\ v = cy + d$

$then, \qquad r_{uv} = \dfrac{ac}{|a||c|} r_{xy}$

(8) $Var(x+y) = {\sigma_x}^2 + {\sigma_y}^2 + 2\sigma_x\sigma_y r_{xy}$

(9) $Var(x-y) = {\sigma_x}^2 + {\sigma_y}^2 - 2\sigma_x\sigma_y r_{xy}$

(10) $Var(x+y) = \frac{1}{n}\sum_{i=1}^{n}\{(x_i + y_i) - \overline{x+y}\}^2$

$\qquad = \frac{1}{n}\sum_{i=1}^{n}\{(x_i - \bar{x}) + (y_i - \bar{y})\}^2$

(11) $If\ u = \frac{x-\bar{x}}{\sigma_x}, v = \frac{y-\bar{y}}{\sigma_y},$

$then, \qquad r_{xy} = Cov(x,y)$

(12) $If\ r_{xy} = \pm 1,\ then\ y\ is\ a\ linear\ function\ of\ x\ \&\ vice-versa.$

(13) $If\ x, y\ are\ independent\ then\ they\ are\ un-correlated.$

(14) $If\ u = a_1 x + b_1 y + c_1$

$\qquad v = a_2 x + b_2 y + c_2$

$then,$

$$Cov(u,v) = a_1 a_2 Var(x) + (a_1 b_2 + a_2 b_1) Cov(x,y) + b_1 b_2 Var(y)$$

Regression

(1) $Regression\ line\ of\ x\ on\ y\ is$

$(x - \bar{x}) = b_{xy}(y - \bar{y})$