Correlation and Regression

Karl Pearson's Co-efficient of Correlation:

Covariance:

$Cov(x, y) = \frac{1}{n} \sum (x - \bar{x})(y - \bar{y}) = \frac{\sum d_x d_y}{n}$

Correlation Co-efficient:

$\rho(x, y)$ or $r = \frac{Cov(x, y)}{\sigma_x \sigma_y} = \frac{Cov(x, y)}{\sqrt{\sigma_x^2 \sigma_y^2}}$ (Product Moment Methos)

$\rho(x, y)$ or $r = \frac{\sum d_x d_y}{n \cdot \sigma_x \sigma_y} = \frac{\sum d_x d_y}{\sqrt{(\sum {d_x}^2)(\sum {d_y}^2)}}$

$r = \frac{n \sum xy - \sum x \cdot \sum y}{\sqrt{n \sum x^2 - (\sum x)^2} \sqrt{n \sum y^2 - (\sum y)^2}}$ (Direct Method)

$r = \frac{n \sum uv - \sum u \cdot \sum v}{\sqrt{n \sum u^2 - (\sum u)^2} \sqrt{n \sum v^2 - (\sum v)^2}}$ (Assumed Mean Method)
($u = x - A$ and $v = y - B$, A and B being assumed means)

Spearman's Rank Correlation:

$r = 1 - \frac{6 \left[\sum D^2 + \sum CF\right]}{n^3 - n}$, $CF = \frac{m^3 - m}{12}$

Regression Analysis:

Regression of Y on X: $y = b_{yx} \cdot x + c$
Regression of X on Y: $x = b_{xy} \cdot y + c$
The lines intersect at $(\bar{x}, \bar{y})$

When x is independant, use $y - \bar{y} = b_{yx}(x - \bar{x})$
When y is independant, use $x - \bar{x} = b_{xy}(y - \bar{y})$

$b_{yx} = \frac{r \cdot \sigma_y}{\sigma_x} = \frac{\sum d_x d_y}{\sum {d_x}^2} = \frac{n \sum xy - \sum x \cdot \sum y}{n \sum x^2 - (\sum x)^2} = \frac{n \sum uv - \sum u \cdot \sum v}{n \sum u^2 - (\sum u)^2}$

$b_{xy} = \frac{r \cdot \sigma_x}{\sigma_y} = \frac{\sum d_x d_y}{\sum {d_y}^2} = \frac{n \sum xy - \sum x \cdot \sum y}{n \sum y^2 - (\sum y)^2} = \frac{n \sum uv - \sum u \cdot \sum v}{n \sum v^2 - (\sum v)^2}$

$r = \sqrt{b_{xy} \cdot b_{yx}}$

If one correlation coefficient is more than unity, other is less than unity.
The correlation coefficient and regression coefficients have the same sign.