Вариант 30.

Имеются данные за 10 лет по прибылям X и Y (в %) двух компаний:

год	1	2	3	4	5	6	7	8	9	10
X	19.2	15.8	12.5	10.3	5.7	5.8	3.5	5.2	7.3	6.7
Y	20.1	18.0	10.3	12.5	6.0	6.8	2.8	3.0	8.5	8.0

По данным таблицы: а) оценить коэффициент детерминации R^2 для линейного уравнения регрессии Y по X и дайте интерпретацию полученного результата; Известно, что: $\sum_{i=1}^{10} x_i=92$; $\sum_{i=1}^{10}(x_i)^2=1084.2$; $\sum_{i=1}^{10} y_i=96$; $\sum_{i=1}^{10}(y_i)^2=1225.7$; $\sum_{i=1}^{10} x_i y_i=1142.5$; б) рассчитать параметры обратной функции $y=1/(\beta_0+\beta_1 x+\varepsilon)$. Известно, что: $\sum_{i=1}^{10}(1/y_i)=1.5$; $\sum_{i=1}^{10}(1/y_i)\,x_i=10.4$; в) для временного ряда y_t найти коэффициент автокорреляции (для лага $\tau=2$). Известно, что: $\sum_{t=3}^{10} y_t=57,9$; $\sum_{i=3}^{10}(y_t)^2=497,7$; $\sum_{t=3}^{10} y_{t-2}=79,5$; $\sum_{i=3}^{10}(y_{t-2})^2=1089$; $\sum_{t=3}^{10} y_t y_{t-2}=663,8$.

Решение

<u>а) Коэффициент детерминации равен квадрату коэффициента корреляции: $R^2 = r^2$</u>

Коэффициент корреляции рассчитывается по формуле:

$$r_{xy}=\frac{cov(x,y)}{\sigma_x\cdot\sigma_y}=\frac{n\cdot\sum x_i y_i-\sum x_i\sum y_i}{\sqrt{n\cdot\sum x_i^2-(\sum x_i)^2}\cdot\sqrt{n\cdot\sum y_i^2-(\sum y_i)^2}}$$

$$cov(x,y)=n\cdot\sum x_i y_i-\sum x_i\sum y_i=10*1142,5-92*96=2593$$

$$\sigma_x=\sqrt{n\cdot\sum x_i^2-(\sum x_i)^2}=\sqrt{10*1084,2-92*92}=48,7647$$

$$\sigma_y=\sqrt{n\cdot\sum y_i^2-(\sum y_i)^2}=\sqrt{10*1225,7-96*96}=55,145$$

$$r_{xy}=\frac{cov(x,y)}{\sigma_x\cdot\sigma_y}=\frac{2593}{48,7647*55,145}=0,9642$$

Отсюда, коэффициент детерминации: $R^2 = r^2 = 0,9642^2 = 0,9298$

Его можно интерпретировать следующим образом: вариация прибыли Y на 92,98% объясняется вариацией прибыли X и на 7,02% объясняется прочими факторами.

<u>б) Рассчитаем параметры обратной функции $Y = 1 / (b_0+b_1x)$</u>

Для нахождения уравнения линейной регрессии используется метод наименьших квадратов. С использованием данного метода коэффициенты уравнения регрессии могут быть найдены по следующим формулам:

$$b1=\frac{n\cdot\sum x_i y_i-\sum x_i\sum y_i}{n\cdot\sum x_i^2-(\sum x_i)^2}$$

$$b0=\frac{\sum x_i^2\sum y_i-\sum x_i\sum x_i y_i}{n\cdot\sum x_i^2-(\sum x_i)^2}$$

Для обратной регрессии необходимо заменить $\sum y_i$ на $\sum 1/y_i$, а $\sum x_i y_i$на $\sum x_i/y_i$

Тогда получим:

$$b1=\frac{n\cdot\sum x_i/y_i-\sum x_i\sum 1/y_i}{n\cdot\sum x_i^2-(\sum x_i)^2}=\frac{10\cdot 10,4-92*1,5}{10*1084,2-92*92}=\frac{-34}{2378}=-0,0143$$

$$b0 = \frac{\sum x_i^2 \sum 1/y_i - \sum x_i \sum x_i/y_i}{n \cdot \sum x_i^2 - (\sum x_i)^2} = \frac{1084{,}2 * 1{,}5 - 92 * 10{,}4}{10 * 1084{,}2 - 92 * 92} = \frac{669{,}5}{2378} = 0{,}28154$$

Таким образом, обратная функция будет иметь следующий вид: Y = 1 / (0.28154–0.0143·X)

в) Определим коэффициент автокорреляции для временного ряда y_t для лага τ=2

Коэффициент автокорреляции определяется также, как и обычный коэффициент корреляции, только вычисляется корреляция между различными уровнями временного ряда. Коэффициент автокорреляции второго порядка рассчитывается по следующей формуле:

$$r(t)_2 = \frac{cov(y_t, y_{t-2})}{\sigma_y \cdot \sigma_{y(t-2)}} = \frac{n \cdot \sum y_t y_{t-2} - \sum y_t \sum y_{t-2}}{\sqrt{n \cdot \sum y_t^2 - (\sum y_t)^2} \cdot \sqrt{n \cdot \sum y_{t-2}^2 - (\sum y_{t-2})^2}}$$

$$cov(y_t, y_{t-2}) = n \cdot \sum y_t y_{t-2} - \sum y_t \sum y_{t-2} = 8 * 663.8 - 57.9 * 79.5 = 707.35$$

$$\sigma_y = \sqrt{n \cdot \sum y_t^2 - (\sum y_t)^2} = \sqrt{8 * 497.7 - 57.9 * 57.9} = 25.084$$

$$\sigma_{y(t-2)} = \sqrt{n \cdot \sum y_{t-2}^2 - (\sum y_{t-2})^2} = \sqrt{8 * 1089 - 79.5 * 79.5} = 48.9$$

$$r(t)_2 = \frac{cov(y_t, y_{t-2})}{\sigma_y \cdot \sigma_{y(t-2)}} = \frac{707.35}{25.084 \cdot 48.9} = 0.576$$

Коэффициент автокорреляции второго порядка равен 0,576.