Simple Linear Regression – Bollywood Movie Revenues and Budgets

Data: Sample of n = 190 Bollywood films released between 2013-2017

Goal: Observe the relationship between Revenues and Budgets. After looking at several plots (see plots from program), let **Y= log(Revenues)** and **X = log(Budget)**, which shows a **linear relation** and **constant variance**.

Q.1. Give the simple linear regression model: _____

Q.2. Give the 3 (unknown) parameters and their interpretations:

Parameter 1: _____ Parameter 2: _____ Parameter 3: _____

Q.3. After fitting the regression model, give point estimates of the 3 parameters and the fitted equation.

$a =$ _____ $b =$ _____ $\hat{\sigma} =$ _____ $\hat{y} =$ _____

Q.4. "Race 2" had Budget=65 (log(Budget) = x = 4.17) and a Revenue=96.34 (log(Rev) = y =4. 57), give its predicted y, its residual and its predicted Revenue (exponentiate predicted y).

$\hat{y} =$ _____ $e = y - \hat{y} =$ _____ $\exp\left(\hat{y}\right) =$ _____

Q.5. Give the correlation between:

i) Budget,Revenue _____ ii) Y,X _____ iii) Spearman r _____

Q.6. Give the sums of squares:

Total: $TSS =$ _____ Regression: $SSR =$ _____ Error (Residual): $SSE =$ _____

Q.7. Give the coefficient of determination r^2 and its interpretation:

$r^2 =$ _____ Interpretation: _____

Q.8. Test whether there is an association between Y and X (t-test): $b =$ _____ $se\{b\} =$ _____

$H_0:$ _____ $H_A:$ _____ Test Stat: _____ Rej Reg. _____ $P =$ _____

Q.9.Give the 95% Confidence Interval for β:

Estimate $\pm$ $t^* se =$ _____ 95% CI: _____

Q.10. Complete the following Analysis of Variance table:

Source	df	Sum Sq	Mean Sq	F	F(.05)	P-value
Regression						
Error (Residual)				#N/A	#N/A	#N/A
Total			#N/A	#N/A	#N/A	#N/A