

**University of Maryland, College Park**  
**Department of Civil and Environmental Engineering**

Quiz 5 Solution, closed book & notes, for 15 minutes  
 December 3, 2001

ENCE 302

Probability and Statistics for Civil Engineers

Name:       SAMPLE      **Problem 1**

In certain type of metal test specimen, the normal stress on a specimen is known to be functionally related to the shear resistance. The following is a set of experimental data on the two variables:

$\sigma$ (kg/cm <sup>2</sup> )	26.8	25.4	28.9	23.6	27.7
$\tau$ (kg/cm <sup>2</sup> )	26.5	27.3	24.2	27.1	23.6

- Estimate the regression linear model  $\hat{\tau} = b_0 + b_1\sigma$  using the principle of least squares.
- Estimate the shear resistance for a normal stress of 24.5 kg/cm<sup>2</sup>.
- Compute the correlation coefficient.
- Compute both the standard deviation  $S_\tau$  and the standard error of estimate  $S_e$ .
- Do you think that the regression model has improved the prediction? Why

## \*\*\* SOLUTION \*\*\*

	TV	EV	UV	EV + UV						
$i$	$\sigma_i$	$\tau_i$	$\sigma_i^2$	$\sigma_i \tau_i$	$\hat{\tau}_i$	$(\tau_i - \bar{\tau})^2$	$(\hat{\tau}_i - \bar{\tau})^2$	$(\tau_i - \hat{\tau}_i)^2$		$\tau_i^2$
1	26.8	26.5	718.24	710.2	25.51864	0.5776	0.048999	0.963063	1.012063	702.25
2	25.4	27.3	645.16	693.42	26.48708	2.4336	0.558133	0.660834	1.218967	745.29
3	28.9	24.2	835.21	699.38	24.06598	2.3716	2.802341	0.017961	2.820302	585.64
4	23.6	27.1	556.96	639.56	27.73222	1.8496	3.968946	0.399704	4.36865	734.41
5	27.7	23.6	767.29	653.72	24.89607	4.5796	0.712213	1.679805	2.392018	556.96
$\Sigma$	132.4	128.7	3522.86	3396.28	128.7	11.812	8.090633	3.721367	11.812	3324.55

- (a) Let
- $\sigma = x$
- and
- $\tau = y$

$$b_1 = \frac{\sum_{i=1}^n x_i y_i - \frac{1}{n} \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sum_{i=1}^n x_i^2 - \frac{1}{n} \left( \sum_{i=1}^n x_i \right)^2} = \frac{3396.28 - \frac{1}{5} (132.4)(128.7)}{3522.86 - \frac{(132.4)^2}{5}} = -0.6917$$

$$b_0 = \bar{Y} - b_1 \bar{X} = \frac{1}{n} \sum_{i=1}^n y_i - b_1 \frac{1}{n} \sum_{i=1}^n x_i = \frac{128.4}{5} - \frac{-0.6917(132.4)}{5} = 44.0573$$

Hence, the regression line is:

$$\hat{\tau} = 44.0573 - 0.6917\sigma$$

(b)

$$\hat{\tau}(24.5) = 44.0573 - 0.6917(24.5) = 27.11 \text{ kg/cm}^2$$

(c)

$$R^2 = \frac{\text{EV}}{\text{TV}} = \frac{8.090633}{11.812} = 0.68495 \quad \Rightarrow \quad R = 0.8276$$

Alternatively, the following equation can be used:

$$R = \frac{\sqrt{\sum_{i=1}^n (\hat{y}_i - \bar{Y})^2}}{\sqrt{\sum_{i=1}^n (y_i - \bar{Y})^2}} = \frac{\sum_{i=1}^n x_i y_i - \frac{1}{n} \left( \sum_{i=1}^n x_i \right) \left( \sum_{i=1}^n y_i \right)}{\sqrt{\sum_{i=1}^n x_i^2 - \frac{1}{n} \left( \sum_{i=1}^n x_i \right)^2} \sqrt{\sum_{i=1}^n y_i^2 - \frac{1}{n} \left( \sum_{i=1}^n y_i \right)^2}}$$

(d)

$$S_{\tau} = \sqrt{\frac{1}{5-1} \left( 3324.55 - \frac{1}{5} (128.7)^2 \right)} = 1.718$$

$$S_e = \sqrt{\left( \frac{n-1}{n-p-1} \right) S_y (1-R^2)} = \sqrt{\left( \frac{5-1}{5-1-1} \right) (1.718)^2 (1-0.68495)} = 1.1138$$

(e)

Because  $S_e (= 1.114)$  is not considerably smaller than  $S_{\tau} (= 1.718)$ , the regression line (or model) has slightly improved the prediction.

**Formulas**

■ Approximate Methods (Random Vector)

– First-order (approximate) Mean

$$E(Y) = \mu_Y = g[E(X_1), E(X_2), \dots, E(X_n)]$$

– First-order (approximate) Variance

$$\text{Var}(Y) = \sigma_Y^2 = \sum_{i=1}^n \sum_{j=1}^n \left. \frac{\partial g(\mathbf{X})}{\partial X_i} \right|_{E(X_i)} \left. \frac{\partial g(\mathbf{X})}{\partial X_j} \right|_{E(X_j)} \text{Cov}(X_i, X_j)$$

■ Approximate Methods (Random Vector)

– First-order (approximate) Variance

If the  $X_i$ 's are uncorrelated (statistically independent), then

$$\text{Var}(Y) = \sigma_Y^2 \approx \sum_{i=1}^n \left( \left. \frac{\partial g(\mathbf{X})}{\partial X_i} \right|_{E(X_i)} \right)^2 \text{Var}(X_i)$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad P(B) \neq 0$$

■ Marginal Distributions

The marginal mass function for  $X_2$  that is not equal to zero is

$$P_{X_2}(x_2) = \sum_{\text{all } x_1} P_{X_1 X_2}(x_1, x_2)$$

The marginal mass function for  $X_1$  that is not equal to zero is

$$P_{X_1}(x_1) = \sum_{\text{all } x_2} P_{X_1 X_2}(x_1, x_2)$$

■ Conditional Probability Mass Function

The conditional probability mass function for two random variables  $X_1$  and  $X_2$  is given by

$$P_{X_1|X_2}(x_1 | x_2) = \frac{P_{X_1 X_2}(x_1, x_2)}{P_{X_2}(x_2)}$$

where  $P_{X_1|X_2}(x_1 | x_2)$  results in the probability of  $X_1 = x_1$  given that  $X_2 = x_2$ .

$P_{X_2}(x_2)$  = marginal mass function for  $X_2$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2$$

$$S^2 = \frac{1}{n-1} \left[ \sum_{i=1}^n x_i^2 - \frac{1}{n} \left( \sum_{i=1}^n x_i \right)^2 \right]$$

OR

$$S^2 = \frac{1}{n-1} \left[ \sum_{i=1}^n x_i^2 - n\bar{X}^2 \right]$$

■ Multiple Random Variables

– If the function  $Y = g(\mathbf{X})$  is given by

$$Y = g(\mathbf{X}) = a_0 + a_1 X_1 + a_2 X_2 + \dots + a_n X_n$$

Then

$$E(Y) = a_0 + a_1 E(X_1) + a_2 E(X_2) + \dots + a_n E(X_n)$$

and

$$\text{Var}(Y) = \sum_{i=1}^n \sum_{j=1}^n a_i a_j \text{Cov}(X_i, X_j) = \sum_{i=1}^n \sum_{j=1}^n a_i a_j \rho_{X_i X_j} \sigma_{X_i} \sigma_{X_j}$$

If the random variables of  $\mathbf{X}$  are uncorrelated, then

$$\text{Var}(Y) = \sum_{i=1}^n a_i^2 \text{Var}(X_i)$$