

PAST CAS AND SoA EXAMINATION QUESTIONS

1. Given the 5 pairs of points (x, y) shown below,

x	4	0	-2	3	1
y	5	0	0	6	3

what line of the form $y = x + a$ best fits the data by the method of least squares?

- A. $y = x + 6/5$ (791-2-11) **B.** $y = x + 8/5$ C. $y = x + 14/5$ D. $y = x + 11/6$ E. $y = x + 13/6$

2. Suppose the line $y = bx + 1$ is fit by the method of least squares to the three data points (x, y) : $(1, 2)$, $(2, 2)$, $(4, 0)$. What is the value of the constant b ?

- A. $-5/7$ **B.** $-1/21$ C. $1/21$ D. $1/7$ E. $5/7$ (792-2-48)

3. Suppose the straight line of the form $y = bx + 1$ is fitted to the four data points $(0, 2)$, $(1, 2)$, $(2, 0)$, and $(3, -1)$ by the method of least squares. What is the value of b ?

- A. $-1/4$ B. $-5/14$ **C.** $-1/2$ D. -1 E. $-11/10$ (81F-2-7)

4. Suppose 3 observations on the ordered pair (x, y) yield $(0, -1)$, $(1, 0)$, and $(1, 1)$. What would be the least-squares estimate of b in fitting a line of the form $y = bx$?

- A. $-1/2$ B. 0 C. $1/2$ D. $2/3$ E. 1 (82S-2-49)

5. You wish to fit the Gompertz form by unweighted simple linear regression to the following data;

x	1	2	3	4	5
$\ln \mu_x$	-3.10	-3.07	-3.06	-3.04	-3.03

Determine $\ln B$.

- A. -3.162 B. -3.145 C. -3.128 D. -3.111 E. -3.094 (84F-5B-5)

6. For the data $(x_1, y_1) = (1, 2)$ and $(x_2, y_2) = (5, 3)$ and the model $E(Y) = bx$, the least-squares estimate of b is:

- A. $1/4$ B. $17/26$ C. $17/13$ D. $17/6$ E. 4 (85S-2-19)

7. Let Y_1, \dots, Y_n be random variables such that $E[Y_i] = 2 + bx_i$, $i = 1, \dots, n$. Given observations $(x_1, y_1), \dots, (x_n, y_n)$, what is the least-squares estimate for b ?

- A.** $\left[\sum_{i=1}^n y_i x_i - 2 \sum_{i=1}^n x_i \right] / \sum_{i=1}^n x_i^2$ B. $\left[\sum_{i=1}^n y_i x_i - 2n \right] / \sum_{i=1}^n x_i^2$ C. $\left[\sum_{i=1}^n y_i x_i \right] / \sum_{i=1}^n x_i(2 - x_i)$
 D. $\left[\sum_{i=1}^n y_i x_i \right] / \left[2 - \sum_{i=1}^n x_i^2 \right]$ E. $\sum_{i=1}^n y_i / \left[2 - \sum_{i=1}^n x_i^2 \right]$ (86F-110-25)

8. You are using a linear regression model to represent the following data:

Independent variable (X)	1	2	3	4	5	6	7	8	9
Dependent variable (Y)	2	4	1	3	6	8	7	5	9

If $\sum_{i=1}^9 (Y_i - \hat{Y}_i)^2 = 21.6$, then $a = .8$ to the nearest .1. (87S-120-8-MC)

9. You are to use simple linear regression to assess the impact of the unemployment rate on the duration of disability for injured workers. You have collected data on the unemployment rate (X_i) and the duration of disability expressed in days (Y_i) for 1967 through 1987. You have determined:

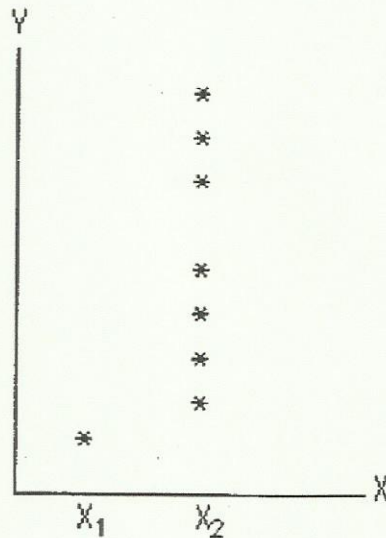
$$\sum_{i=1967}^{1987} X_i = 1.5095 \quad \sum_{i=1967}^{1987} Y_i = 1.599.84 \quad \sum_{i=1967}^{1987} (X_i - \bar{X})^2 = .01491$$

$$\sum_{i=1967}^{1987} (Y_i - \bar{Y})^2 = 761.43 \quad \sum_{i=1967}^{1987} (X_i - \bar{X})(Y_i - \bar{Y}) = 2.8396$$

If the unemployment rate is forecast to be .06 in 1988 and the cost of a day of disability is expected to be 100, estimate the expected cost of an injury in 1988.

- A. 7,000 B. 7,200 **C. 7,400** D. 7,600 E. 7,800 (88S-120-7)

10. Eight pairs of observations (X_i, Y_i) are represented by the following scatter diagram:



The least-squares line passes through the point $([X_1 + X_2]/2, k)$. Determine k .

- A. $3Y_1/8 + 5\bar{Y}/8$ B. $3Y_1/7 + 4\bar{Y}/7$ C. $Y_1/2 + \bar{Y}/2$ D. $4Y_1/7 + 3\bar{Y}/7$ E. $5Y_1/8 + 3\bar{Y}/8$ (90S-120-7)

11. Which of the following are valid expressions for b , the slope coefficient in the simple linear regression of Y on X .

1. $\frac{\sum X_i Y_i - \bar{Y} \sum X_i}{\sum X_i^2 - \bar{X} \sum X_i}$ 2. $\frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum X_i^2 - (\bar{X})^2}$ 3. $\frac{\sum (X_i)(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$

- A. 1, 2 **B. 1, 3** C. 2, 3 D. 1, 2, 3 E. None of these answers is correct. (90S-120-11)

12. You use the method of least squares to fit the model $Y_i = 10 + bX_i$ to the following data:

X	0	1	2	3
Y	11	12	15	16

Determine b.

- A. 10/7 **B.** 15/7 C. 30/7 D. 40/7 E. 45/7 (93S-120-7)
13. You are given the following information about a simple regression model fit to 10 observations:

$$\sum_{i=1}^{10} X_i = 20 \quad \sum_{i=1}^{10} Y_i = 100 \quad \sqrt{\frac{1}{9} \sum_{i=1}^{10} (X_i - \bar{X})^2} = 2 \quad \sqrt{\frac{1}{9} \sum_{i=1}^{10} (Y_i - \bar{Y})^2} = 8$$

You are also given that the sample correlation coefficient $r = -.98$. Determine the predicted value of Y when $X = 5$.

- A. -10 **B.** -2 C. 11 D. 30 E. 37 (96S-120-7)
14. The method of least squares is applied to ten pairs of values (X_i, Y_i) and yields 10 and 20, respectively, as estimates of the intercept and slope of the least-squares line. Determine the least-squares estimates of the intercept and slope if every X_i is doubled and every Y_i remains the same.
- A. Intercept = 5, slope = 10 **B.** Intercept = 10, slope = 10 C. Intercept = 10, slope = 20
 D. Intercept = 20, slope = 20 E. Intercept = 20, slope = 40 (98F-120-1)

15. You are given:

- i) Z_1 and Z_2 are independent $N(0, 1)$ random variables.
 ii) a, b, c, d, e, f are constants.
 iii) $Y = a + bZ_1 + cZ_2$ and $X = d + eZ_1 + fZ_2$

Determine $E(Y|X)$.

- A.** a B. $a + (b + c)(X - d)$ C. $a + (be + cf)(X - d)$ D. $a + [(be + cf)/(e^2 + f^2)]X$
 E. $a + [(be + cf)/(e^2 + f^2)](X - d)$ (03F-4-13)