Assignment 1: Part 1 (20%)

Airfreight breakage. A substance used in biological and medical research is shipped by air-freight to users in cartons of 1,000 ampules. The data below, involving 10 shipments, were collected on the number of times the carton was transferred from one aircraft to another over the shipment route (X) and the number of ampules found to be broken upon arrival (Y). Assume that first-order regression model (1.1) is appropriate.

i:	1	2	3	4	5	6	7	8	9	10
X_i:	1	0	2	0	3	1	0	1	2	0
Y_i:	16	9	17	12	22	13	8	15	19	11

a. Obtain the estimated regression function. Plot the estimated regression function and the data. Does a linear regression function appear to give a good fit here?

b. Obtain a point estimate of the expected number of broken ampules when $X = 1$ transfer is made.

c. Estimate the increase in the expected number of ampules broken when there are 2 transfers as compared to 1 transfer.

d. Verify that your fitted regression line goes through the point $(\bar{X}, \bar{Y})$.

Assignment 1: Part 2 (20%)

1.19. **Grade point average.** The director of admissions of a small college selected 120 students at random from the new freshman class in a study to determine whether a student's grade point average (GPA) at the end of the freshman year (Y) can be predicted from the ACT test score (X). The results of the study follow. Assume that first-order regression model (1.1) is appropriate.

i:	1	2	3	...	118	119	120
X_i:	21	14	28	...	28	16	28
Y_i:	3.897	3.885	3.778	...	3.914	1.860	2.948

a. Obtain the least squares estimates of β_0 and β_1, and state the estimated regression function.

b. Plot the estimated regression function and the data. Does the estimated regression function appear to fit the data well?

c. Obtain a point estimate of the mean freshman GPA for students with ACT test score $X = 30$.

d. What is the point estimate of the change in the mean response when the entrance test score increases by one point?

e. Obtain the residuals e_i. Do they sum to zero in accord with (1.17)?

f. Estimate σ^2 and σ. In what units is σ expressed?

Assignment 1: Part 3 (20%)

1.27. **Muscle mass.** A person's muscle mass is expected to decrease with age. To explore this relationship in women, a nutritionist randomly selected 15 women from each 10-year age group, beginning with age 40 and ending with age 79. The results follow; X is age, and Y is a measure of muscle mass. Assume that first-order regression model (1.1) is appropriate.

i:	1	2	3	...	58	59	60
X_i:	43	41	47	...	76	72	76
Y_i:	106	106	97	...	56	70	74

a. Obtain the estimated regression function. Plot the estimated regression function and the data. Does a linear regression function appear to give a good fit here? Does your plot support the anticipation that muscle mass decreases with age?

b. Obtain the following: (1) a point estimate of the difference in the mean muscle mass for women differing in age by one year, (2) a point estimate of the mean muscle mass for women aged $X = 60$ years, (3) the value of the residual for the eighth case, (4) a point estimate of σ^2.

Assignment 1: Part 4 (20%)

Typographical errors. Shown below are the number of galleys for a manuscript (X) and the dollar cost of correcting typographical errors (Y) in a random sample of recent orders handled by a firm specializing in technical manuscripts. Assume that the regression model $Y_i = \beta_1 X_i + \varepsilon_i$ is appropriate, with normally distributed independent error terms whose variance is $\sigma^2 = 16$.

i:	1	2	3	4	5	6
X_i:	7	12	4	14	25	30
Y_i:	128	213	75	250	446	540

a. State the likelihood function for the six Y observations, for $\sigma^2 = 16$.

b. Evaluate the likelihood function for $\beta_1 = 17$, 18, and 19. For which of these β_1 values is the likelihood function largest?

c. The maximum likelihood estimator is $b_1 = \sum X_i Y_i / \sum X_i^2$. Find the maximum likelihood estimate. Are your results in part (b) consistent with this estimate?

d. Using a statistics package, evaluate the likelihood function for values of β_1 between $\beta_1 = 17$ and $\beta_1 = 19$ and plot the function. Does the point at which the likelihood function is maximized correspond to the maximum likelihood estimate found in part (c)?

Assignment 1: Part 5 (20%)

Refer to **Typographical errors** Problem . Assume that first-order regression model (1.1) is appropriate, with normally distributed independent error terms whose variance is $\sigma^2 = 16$.

a. State the likelihood function for the six observations, for $\sigma^2 = 16$.

b. Obtain the maximum likelihood estimates of β_0 and β_1, using (1.27).

c. Using a statistics package, obtain a three-dimensional plot of the likelihood function for values of β_0 between $\beta_0 = -10$ and $\beta_0 = 10$ and for values of β_1 between $\beta_1 = 17$ and $\beta_1 = 19$. Does the likelihood appear to be maximized by the maximum likelihood estimates found in part (b)?

Assignment 1

- Submit your responses in Blackboard in a **single** pdf file by Monday September 11, midnight.
- Use the following file name: LASTNAME_FIRSTNAME_ASUID_ASSIGNMENTNUMBER
- Prepare your pdfs carefully; each week some of you will present their work.
- Include the `R` commands you used.