Problem 1. An analysis is performed to study the relationship between a response variable Y and a single explanatory variable X. The model guiding the analysis is

$$Yi = \beta_0 + \beta_1 X_i + \epsilon_i,$$

where the ϵ_i's are independently and identically distributed as $N(0, \sigma^2)$ and σ^2 is unknown. The data with a sample size of $n = 5$ were analyzed using R and produced the following output.

Observations	x	y	Residual
1	0.0	0.9	0.027
2	4.1	5.9	0.188
3	5.1	6.2	-0.692
4	6.1	8.7	0.628
5	$*$	9.1	$*$

Fill in the two missing values (denoted by "$*$") in the table above.

Problem 2. In fitting a SLR model, ten of the residuals for the regression of weight on age for eleven children are:

$$0.02, -0.02, 0.01, -0.01, -0.04, 0.03, 0.01, 0.02, 0.01, 0.05$$

What is the eleventh residual?

Problem 3. Consider the multiple regression model:

$$Y_i = \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i, i = 1 \ldots, n,$$

where the ϵ_i are uncorrelated, with $E(\epsilon_i) = 0$ and $\text{Var}(\epsilon_i) = \sigma^2$. Let $X = \begin{pmatrix} X_{11} & X_{21} & \ldots & X_{n1} \\ X_{12} & X_{22} & \ldots & X_{n2} \end{pmatrix}^T$ and $X^T X$ is invertible (i.e., full rank).

(a) Write the normal equations using the least square criterion.

(b) What are the least square estimators of β_1 and β_2?

(c) Let e_i denote the residual and $\hat{Y}_i$ the fitted value for $i = 1, \ldots, n$. Which of the following statements are true?

(c1) $\sum_{i=1}^n e_i = 0$ always.

(c2) $\sum_{i=1}^n \hat{Y}_i = 0$ always.

(c3) $\sum_{i=1}^n X_{ij} = 0$ always for $j = 1, 2$.

(c4) $\sum_{i=1}^n Y_i = 0$ always.