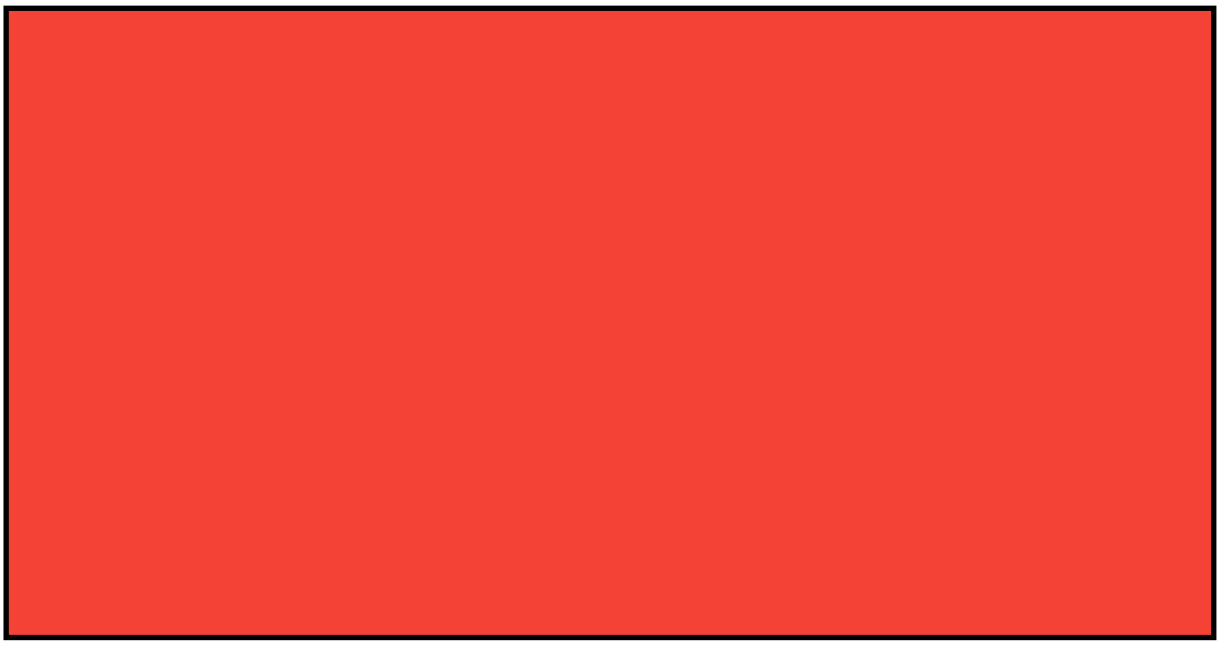

1. Suppose that we fit the following model to the n observations $(y_1, x_{11}, x_{21}), ..., (y_n, x_{1n}, x_{2n})$:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i,$$

for $i = 1, ..., n$, where all ϵ_i are identically and independently distributed as a normal random variable with mean zero and variance σ^2 and every x_{ji} is fixed.

   (a) [25 Points] Suppose the above model is the true model. Show that at any observation y_i, the point estimator of the mean response and its residual are two statistically independent normal random variables.

   (b) [25 Points] Suppose that the above model is the true model, but we fit the data to the following model (i.e., ignore the variable x_2):

$$y_i = \beta_0 + \beta_1 x_{1i} + \epsilon_i,$$

   for $i = 1, ..., n$. Assume that $\bar{x}_1 = 0$, $\bar{x}_2 = 0$, and $\sum_{i=1}^{n} x_{1i} x_{2i} = 0$. Derive the least-squares estimator of β_1 obtained from fitting this new model. Is this least squares estimator biased for β_1 in the original model? Why or why not?

2. Ten observations on the response variable y, associated with two regressor variables x_1 and x_2, are given in the following table.

Observation No.	y	x_1	x_2
1	7	9	1
2	8	6	1
3	5	7	1
4	3	8	1
5	2	5	1
6	10	7	-1
7	9	6	-1
8	10	3	-1
9	9	4	-1
10	8	4	-1

The model fitted to these observations is

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i,$$

for $i = 1, ..., n$, where all ϵ_i are identically and independently distributed as a normal random variable with mean zero and a known variance of $\sigma^2 = 3$.

(a) [25 Points] Test the null hypothesis, that there is no difference between the y-intercept for $x_2 = 1$ and the y-intercept for $x_2 = -1$, at a statistical significance level of 0.05.

(b) [25 Points] Now fit the following model to the above ten observations:

$$y_i = \beta_0 + \beta_2 x_{2i} + \epsilon_i.$$

Calculate the variance of the residual for observation #6. Make sure to state any assumption(s) used!