Problem 1 (Least Squares Coefficients). Recall from lecture our least-squares estimates for the coefficients $\beta_0, \beta_1$ in linear regression, where we chose our estimates $\hat{\beta}_0, \hat{\beta}_1$ as the minimizers of

$$S(\beta_0, \beta_1) = \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i))^2,$$

where $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ are the observed pairs of predictors ($x$s) and responses ($y$s). In what follows, all sums are understood to be over the integers $1, 2, \ldots, n$.

(a) Compute the partial derivatives of $S$ with respect to $\beta_0$ and $\beta_1$ to obtain expressions for $\frac{\partial}{\partial \beta_0} S(\beta_0, \beta_1)$ and $\frac{\partial}{\partial \beta_1} S(\beta_0, \beta_1)$.

(b) Set these partial derivatives equal to 0 and solve for $\beta_0$ and $\beta_1$ to show that

$$\hat{\beta}_0 = \frac{\sum_i x_i^2 \sum_i y_i - (\sum_i x_i)(\sum_i x_i y_i)}{n \sum_i x_i^2 - (\sum_i x_i)^2} \quad (1)$$

and

$$\hat{\beta}_1 = \frac{n \sum_i x_i y_i - (\sum_i x_i)(\sum_i y_i)}{n \sum_i x_i^2 - (\sum_i x_i)^2}. \quad (2)$$

You may assume that the solution $(\hat{\beta}_0, \hat{\beta}_1)$ is indeed a minimizer of $S(\beta_0, \beta_1)$.

(c) Rearrange the sums in the numerator and denominator of $\hat{\beta}_1$ to show that

$$\hat{\beta}_1 = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2}.$$

(d) Rearrange the expression for $\hat{\beta}_0$ in Equation (1) or simply set $\frac{\partial}{\partial \beta_0} S(\beta_0, \beta_1) = 0$ to show that

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}.$$
(e) Use the above identities for $\hat{\beta}_0$ and $\hat{\beta}_1$ to show that the least-squares regression line $y = \hat{\beta}_0 + \hat{\beta}_1 x$ passes through the point $(\bar{x}, \bar{y})$.

**Problem 2** (Expectation of Estimated Coefficients Under Independent Errors). In lecture, we considered the setting where the observations $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ were drawn according to the model $y_i = \beta_0 + \beta_1 x_i + e_i$, where the coefficients $\beta_0, \beta_1$ and the predictors $x_1, x_2, \ldots, x_n$ were fixed, and the measurement errors $e_1, e_2, \ldots, e_n$ were independent with shared mean $\mathbb{E} e_1 = \mathbb{E} e_2 = \cdots = \mathbb{E} e_n = 0$. Let $\hat{\beta}_0$ and $\hat{\beta}_1$ be the estimators defined in Equations (1) and (2), respectively.

1. Show that under this model, $\mathbb{E} \hat{\beta}_0 = \beta_0$.
2. Show that under this model, $\mathbb{E} \hat{\beta}_1 = \beta_1$.
3. Now, suppose that we make the additional model assumptions that $\text{Var} e_i = \sigma^2$ for all $i = 1, 2, \ldots, n$ and that the error terms are uncorrelated, so that $\mathbb{E} e_i e_j = 0$ for $i \neq j$.

   (i) Show that
   $$\text{Var} \hat{\beta}_1 = \frac{n \sigma^2}{\sum_i x_i^2 - (\sum_i x_i)^2}.$$

   (ii) Show that
   $$\text{Var} \hat{\beta}_0 = \frac{\sigma^2 \sum_i x_i^2}{n \sum_i x_i^2 - (\sum_i x_i)^2}.$$

   (iii) Show that
   $$\text{Cov}(\hat{\beta}_0, \hat{\beta}_1) = \frac{-\sigma^2 \sum_i x_i}{n \sum_i x_i^2 - (\sum_i x_i)^2}.$$

**Problem 3** (Weighted Least Squares). Suppose that we make observations $y_i = \beta_0 + \beta_1 x_i + e_i$ for $i = 1, 2, \ldots, n$, where the errors $e_1, e_2, \ldots, e_n$ are independent with mean 0, but the measurement errors may differ in their variances. Suppose that the $i$-th error has variance $\text{Var} e_i = \rho_i^2 \sigma^2$, where $\sigma^2 > 0$ and $\rho_i > 0$ for all $i = 1, 2, \ldots, n$. Because the errors have different variances, the theory we discussed in lecture does not apply. Intuitively, we should make sure that the observations with larger variances (i.e., larger values of $\rho_i^2$) do not influence our estimates as much as the observations with smaller variances.

1. Suppose that we instead consider $z_i = \rho_i^{-1} y_i$, so that
   $$z_i = u_i \beta_0 + v_i \beta_1 + \delta_i,$$
   where $u_i = \rho_i^{-1}, \ v_i = \rho_i^{-1} x_i$ and $\delta_i = \rho_i^{-1} e_i$. Show that the $\delta_1, \delta_2, \ldots, \delta_n$ error terms are independent with mean 0 and shared variance $\sigma^2$.

2. Consider the least-squares objective
   $$\sum_{i=1}^n (z_i - u_i \beta_0 - v_i \beta_1)^2.$$
   Take derivatives and show that the least-squares estimates for $\beta_0$ and $\beta_1$ are
   $$\hat{\beta}_0 = \frac{\sum_i u_i z_i \sum_j v_j^2 - \sum_j u_i v_i \sum_j v_j z_j}{\sum_i u_i^2 \sum_j v_j^2 - (\sum_i u_i v_i)^2}.$$
and
\[ \hat{\beta}_1 = \frac{\sum_i u_i^2 \sum_j v_j z_j - \sum_i u_i v_i \sum_j u_j z_j}{\sum_i u_i^2 \sum_j v_j^2 - (\sum_i u_i v_i)^2}. \]

3. Verify that these estimates are equivalent to what we would obtain if we instead minimized the weighted least squares objective
\[ \sum_{i=1}^n \rho_i^{-2} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2. \]

**Problem 4** (Predictors and Variance). Suppose that we observe \( y_i = \beta_0 + \beta_1 x_i + e_i \) with independent, mean 0 errors \( e_1, e_2, \ldots, e_n \) with shared variance \( \sigma^2 \). Suppose that \( \bar{x} = 0 \). Show that the slope and intercept of the least-squares fitted line are uncorrelated. **Hint:** use Problem 2.