(b) Now let's generalize the derivations from (a) to the case of general OLS regression

$$Q(\hat{\beta}) = (y - X\hat{\beta})'(y - X\hat{\beta})$$

1. Find $dQ(\hat{\beta})$ and $d^2Q(\hat{\beta})$
2. Write down FOC for OLS problem
3. Find $\hat{\beta}$ assuming that $X'X$ is invertible
4. Show that $\hat{\beta}$ is unbiased
5. Find $Var(\hat{\beta})$
6. Derive in matrix from TSS = ESS + RSS and state assumptions under which this equation is true

1. The random variable X takes three values: 1, 2 and 3 with probabilities p_1, p_2 and p_3.

   How the entropy of X will change if we split each value v into two new values, $v - 0.1$ and $v + 0.1$, with equal probabilities $p_v/2$?

2. The random variable X takes three values: 1, 2 and 3. There are two probability measuers, p and q, $\mathbb{P}(X = 1) = 0.2$, $\mathbb{P}(X = 2) = 0.3$, $\mathbb{P}(X = 3) = 0.5$.

   (a) Find the probabilities $Q(X = 1)$, $Q(X = 2)$, $Q(X = 3)$ that maximize cross-entropy $CE(p||q)$.

   (b) Find the probabilities $Q(X = 1)$, $Q(X = 2)$, $Q(X = 3)$ that maximize cross-entropy $CE(q||p)$.

   Hint: you may use python if you can't solve the first order conditions by hand.

3. Consider the 1$ lottery ticket that pays you either 5$ or nothing with equal probabilities.

   How much of your current wellfare should you invest in this lottery to maximize the long-term interest rate?

4. The response variable is binary. Elon Musk has split the node of a tree according to the new X-criterion into two non-empty child nodes.

   Can the Gini impurity index increase after this splitting?

5. I have a toy dataset of 5 observations. All values of all variables are pairwise different.

   Consider the random forest algorithm.

   (a) What is the probability that the first tree will use five identical observations?

   (b) What is the probability that the second tree will use all five initial observations?

6. Random variables y_1, y_2, ..., y_n is the initial random sample from uniform distribution on $[0; 1]$. Consider one of the bootstrap samples, y_1^*, y_2^*, ..., y_n^*.

   (a) What is the probability that y_5 will be included exactly 3 times in the bootstrap sample?

   (b) What is the limit of probability in the point (a) when $n \to \infty$?

   (c) Find the probability $\mathbb{P}(\max\{y_1, \ldots, y_n\} > \max\{y_1^*, \ldots, y_n^*\})$.