STA 302f13 Assignment Five[1]

These problems are preparation for the quiz in tutorial on Friday October 18th, and are not to be handed in.

For reference, the general linear model in matrix form is $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\mathbf{X}$ is an $n \times (k+1)$ matrix of observable constants, the columns of $\mathbf{X}$ are linearly independent, $\boldsymbol{\beta}$ is a $(k+1) \times 1$ vector of unknown constants (parameters), and $\boldsymbol{\epsilon}$ is an $n \times 1$ vector of unobservable random variables with $E(\boldsymbol{\epsilon}) = \mathbf{0}$ and $cov(\boldsymbol{\epsilon}) = \sigma^2 \mathbf{I}_n$, where $\sigma^2 > 0$ is an unknown constant parameter. The least squares estimator of $\boldsymbol{\beta}$ is $\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$.

1. Let $Y_i = \beta x_i + \epsilon_i$ for $i = 1, \ldots, n$, where $\epsilon_1, \ldots, \epsilon_n$ are a random sample from a distribution with expected value zero and variance σ^2, and β and σ^2 are unknown constants. The numbers $x_1, \ldots, x_n$ are known, observed constants. This is a special case of the general linear model, which is given above in matrix form.
   (a) What is $\mathbf{X}'\mathbf{X}$?
   (b) What is $\mathbf{X}'\mathbf{Y}$?
   (c) $(\mathbf{X}'\mathbf{X})^{-1}$?
   (d) What is $\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$?
   (e) You already know that $\widehat{\beta}$ is unbiased for β. It is also a *linear* estimator of the form $\mathbf{c}'\mathbf{Y} = \sum_{i=1}^n c_i Y_i$, so it is a *linear unbiased estimator*. What is c_i?

2. If we want to estimate $\mathbf{a}'\boldsymbol{\beta}$ based on sample data, the Gauss-Markov Theorem tells us that the most natural choice is also (in a sense) the best choice. This question leads you through the proof of the Gauss-Markov Theorem. Your class notes should help.
   (a) What is the most natural choice for estimating $\mathbf{a}'\boldsymbol{\beta}$?
   (b) Show that it's unbiased.
   (c) The natural estimator is a *linear* unbiased estimator of the form $\mathbf{c}_0'\mathbf{Y}$. What is the $n \times 1$ vector $\mathbf{c}_0$?
   (d) Of course there are lots of other possible linear unbiased estimators of $\mathbf{a}'\boldsymbol{\beta}$. They are all of the form $\mathbf{c}'\mathbf{Y}$; the natural estimator $\mathbf{c}_0'\mathbf{Y}$ is just one of these. The best one is the one with the smallest variance, because its distribution is the most concentrated around the right answer. What is $Var(\mathbf{c}'\mathbf{Y})$? Show your work.
   (e) We insist that $\mathbf{c}'\mathbf{Y}$ be unbiased. Show that if $E(\mathbf{c}'\mathbf{Y}) = \mathbf{a}'\boldsymbol{\beta}$ for *all* $\boldsymbol{\beta} \in \mathbb{R}^{k+1}$, we must have $\mathbf{X}'\mathbf{c} = \mathbf{a}$.
   (f) So, the task is to minimize $Var(\mathbf{c}'\mathbf{Y})$ by minimizing $\mathbf{c}'\mathbf{c}$ over all $\mathbf{c}$ subject to the constraint $\mathbf{X}'\mathbf{c} = \mathbf{a}$. As preparation for this, show $(\mathbf{c} - \mathbf{c}_0)'\mathbf{c}_0 = 0$.

[1]Copyright information is at the end of the last page.

(g) Using the result of the preceding question, show

$$\mathbf{c}'\mathbf{c} = (\mathbf{c} - \mathbf{c}_0)'(\mathbf{c} - \mathbf{c}_0) + \mathbf{c}_0'\mathbf{c}_0.$$

(h) Since the formula for $\mathbf{c}_0$ has no $\mathbf{c}$ in it, what choice of $\mathbf{c}$ minimizes the preceding expression? How do you know that the minimum is unique?

The conclusion is that $\mathbf{c}_0'\mathbf{Y} = \mathbf{a}'\widehat{\boldsymbol{\beta}}$ is the Best Linear Unbiased Estimator (BLUE) of $\mathbf{a}'\boldsymbol{\beta}$.

3. For the model of Question 1, let $\widehat{\beta}_2 = \frac{\overline{Y}_n}{\overline{x}_n}$.

   (a) Show that $\widehat{\beta}_2$ is unbiased for β.

   (b) Which has the smaller variance, $\widehat{\beta}_2$ or your estimator from Question 1? How do you know? This is quick if you see it.

4. For the model of Question 1, let $\widehat{\beta}_3 = \frac{1}{n}\sum_{i=1}^n \frac{Y_i}{x_i}$.

   (a) Show that $\widehat{\beta}_3$ is unbiased for β.

   (b) Which has the smaller variance, $\widehat{\beta}_3$ or your estimator from Question 1? How do you know? Again, this is quick.

5. The first parts of this question were in Assignment One. Let $Y_1, \ldots, Y_n$ be independent random variables with $E(Y_i) = \mu$ and $Var(Y_i) = \sigma^2$ for $i = 1, \ldots, n$.

   (a) Write down $E(\overline{Y})$ and $Var(\overline{Y})$.

   (b) Let $c_1, \ldots, c_n$ be constants and define the linear combination L by $L = \sum_{i=1}^n c_i Y_i$. What condition on the c_i values makes L an unbiased estimator of μ?

   (c) Is $\overline{Y}$ a special case of L? If so, what are the c_i values?

   (d) What is $Var(L)$?

   (e) Now show that $Var(\overline{Y}) < Var(L)$ for every unbiased $L \neq \overline{Y}$. Hint: Add and subtract $\frac{1}{n}$ as in Question 2.

6. Another way to express the model of Question 5 is to say $Y_i = \mu + \epsilon_i$ for $i = 1, \ldots, n$, where $\epsilon_1, \ldots, \epsilon_n$ are independent random variables from a distribution with expected value zero and variance σ^2. This is a regression with *no independent variables* (weird), and $\beta_0 = \mu$.

   (a) What is the $\mathbf{X}$ matrix?

   (b) What is $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$?

   (c) Now how do you know $Var(\overline{Y}) < Var(L)$ in Question 5, without any calculations?

7. For this question, you may use the following, as well as standard properties of moment-generating functions.

$$Y \sim N(\mu, \sigma^2) \text{ means } M_Y(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$$

$$M_{\mathbf{Y}}(\mathbf{t}) = E(e^{\mathbf{t}'\mathbf{Y}}) \qquad M_{\mathbf{AY}}(\mathbf{t}) = M_{\mathbf{Y}}(\mathbf{A}'\mathbf{t}) \qquad M_{\mathbf{Y}+\mathbf{c}}(\mathbf{t}) = e^{\mathbf{t}'\mathbf{c}} M_{\mathbf{Y}}(\mathbf{t})$$

(a) Let $Z_1, \ldots, Z_p$ be independent standard normal random variables, and $\mathbf{Z} = (Z_1, \ldots, Z_p)'$.

i. What is $E(\mathbf{Z})$?

ii. What is $cov(\mathbf{Z})$?

iii. What is $M_{\mathbf{Z}}(\mathbf{t})$?

iv. Let $\boldsymbol{\Sigma}$ be a $p \times p$ symmetric non-negative definite (real) matrix and $\boldsymbol{\mu} \in \mathbb{R}^p$. Letting $\mathbf{Y} = \boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{Z} + \boldsymbol{\mu}$, show that $M_{\mathbf{Y}}(\mathbf{t}) = e^{\mathbf{t}'\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}}$. A random vector with this moment-generating function will be called *multivariate normal* with parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, and we will write $\mathbf{Y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

(b) Let $\mathbf{Y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and let $\mathbf{A}$ be an $r \times p$ matrix of real constants. Show that the random vector $\mathbf{AY}$ has a multivariate normal distribution. Give the mean vector and covariance matrix.

(c) Let $\mathbf{X} = (X_1, X_2, X_3)'$ be multivariate normal with

$$\boldsymbol{\mu} = \begin{bmatrix} 1 \\ 0 \\ 6 \end{bmatrix} \text{ and } \boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Let $Y_1 = X_1 + X_2$ and $Y_2 = X_2 + X_3$. Find the joint distribution of Y_1 and Y_2.

(d) Let X_1 be Normal(μ_1, σ_1^2), and X_2 be Normal(μ_2, σ_2^2), independent of X_1. What is the joint distribution of $Y_1 = X_1 + X_2$ and $Y_2 = X_1 - X_2$? What is required for Y_1 and Y_2 to be independent? Hint: Use matrices.

(e) Show that if $\mathbf{Y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where the covariance matrix $\boldsymbol{\Sigma}$ is strictly positive definite, $W = (\mathbf{Y} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \boldsymbol{\mu})$ has a chi-squared distribution with p degrees of freedom.
