

STA 2101/442 Assignment Six¹

Please bring two *separate* R printouts to the quiz. *Your printouts must not contain answers to the non-computer parts of this assignment*, The non-computer questions are just practice for the quiz, and will not be handed in.

Note that the [formula sheet](#) will be displayed with Quiz 6. See the link on the course home page in case the one in this document does not work.

1. You probably should do one Fisher Information exercise by hand. This example is very standard, but informative. Among other things, it shows when you do and when you do *not* divide by n to get asymptotic variances.

Let Y_1, \dots, Y_n be a random sample from a $N(\mu, \sigma^2)$ distribution. You may use the well-known maximum likelihood estimators $\hat{\mu} = \bar{Y}$ and $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2$ without deriving them.

- (a) What is $\text{Var}(\bar{Y})$? Just write down the answer.
- (b) Calculate the Fisher Information matrix $\mathcal{I}(\theta)$. It is a function of μ and σ^2 .
- (c) What is $\mathcal{I}(\mu, \sigma^2)^{-1}$?
- (d) Recalling $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} \mathbf{T} \sim N_k(\mathbf{0}, \mathcal{I}(\theta)^{-1})$,
 - i. What is the asymptotic covariance matrix of $(\bar{Y}, \hat{\sigma}^2)$?
 - ii. What part of your answer agrees with your answer to Question 1a?
 - iii. Why does it make sense that the asymptotic covariance matrix is diagonal?
- (e) As a cross-check on your asymptotic variance of $\hat{\sigma}^2$, recall that if $W \sim \chi^2(\nu)$, then $\text{Var}(W) = 2\nu$, and for a normal random sample, $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$. Use this to calculate the *exact* (not asymptotic) variance of $\hat{\sigma}^2$.
- (f) Based on what you have done so far, give the *estimated* asymptotic covariance matrix of $(\bar{Y}, \hat{\sigma}^2)$. Your answer is a 2×2 matrix of handwritten symbols.
- (g) Another formula for the asymptotic covariance matrix is based on the observed Fisher information. Look at the formula sheet, and calculate it for this problem. Compare it to your answer to Question 1f.
- (h) The file [normal.data](#) contains some data that are rounded, but otherwise closer to normal than almost anything in nature. There is a link from the course website in case the one in this document does not work.
 - i. Give the estimated asymptotic covariance matrix of $(\bar{Y}, \hat{\sigma}^2)$ based on your work so far. It is a 2×2 matrix of numbers. Produce it with R.
 - ii. Do the job numerically with R's `nlm` function. Obtain an estimated asymptotic covariance matrix of $(\bar{Y}, \hat{\sigma}^2)$ by inverting the Hessian. Again, your answer is a 2×2 matrix of numbers. Are you shocked? Why or why not?
 - iii. Now do the job with the bootstrap, using $B = 1,000$ as usual. Again, your answer is a 2×2 matrix of numbers. Are you even more shocked? Which estimated asymptotic covariance matrix do you like most?

Bring your printout to the quiz. Your printout must *not* contain answers to the non-computer parts of this question. That is, it must contain only numerical answers.

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

2. Men and women are calling a technical support line according to independent Poisson processes with rates λ_1 and λ_2 per hour. Data for 144 hours are available, but unfortunately the sex of the caller was not recorded. All we have is the number of callers for each hour, which is distributed $\text{Poisson}(\lambda_1 + \lambda_2)$. Here are the data, which are also available in the file [poisson.data](#) on the course website:

```

12  9 25 11 22 16 17  8 14 17 14 17  8 14 18 16 13 17 13 11  9 15 18
14 16 17 16 19 13 18 12 12 13 10  8 15 13 11 15 15  6 10 13 13 11 13
 9 15 16  9  5 10  8 18 13 17  7 13 13 17 12 17 14 16  6 12 17 10  9
14 11 19 13 17 15 20 14 10 13 14 17  9 13 14  7 16 16  9 25 10 10  9
17  7 15 12 14 21 14 18 14 12 13 15 12 11 16 14 15 16  8 19 13 17 15
11 18 13 12 11 19 14 16 17 13 13 19 19 11 19 10 12  9 18 11 14  9 14
14 14 13  9 13 18

```

- The parameter in this problem is $\boldsymbol{\theta} = (\lambda_1, \lambda_2)'$. Try to find the MLE analytically. Show your work. Are there any points in the parameter space where both partial derivatives are zero?
- Now try to find the MLE numerically with R's `nlm` function². The Hessian is interesting; ask for it. Try two different starting values. Compare the minus log likelihoods at your two answers. What seems to be happening here?
- Try inverting the Hessian to get the asymptotic covariance matrix. Any comments?
- To understand what happened in the last item, calculate the Fisher information in a single observation from the definition. That is, letting $\ell = \log f(Y; \boldsymbol{\theta})$, calculate the elements of the 2×2 matrix whose (i, j) element is

$$-E \left(\frac{\partial^2 \ell}{\partial \theta_i \partial \theta_j} \right)$$

- The Fisher information in the sample is just n times the Fisher information in a single observation. Using the numerical MLEs from one of your `nlm` runs, estimate this quantity (a 2×2 matrix). Compare it to the Hessian. Now do you see what happened when you tried to calculate the asymptotic covariance matrix?
- Why did estimation fail for this fairly realistic model?

Bring your printout to the quiz. Your printout must *not* contain answers to the non-computer parts of this question. That is, it must contain only numerical answers.

²When I did this, I got lots of warning messages with some starting values, when the search repeatedly left the parameter space and then bounced back in. For this problem, you don't need to worry about the warnings as long as the exit code is one.

3. Consider the usual fixed effects multiple regression model in scalar form:

$$Y_i = \beta_0 + \beta_1 x_{i,1} + \cdots + \beta_{p-1} x_{i,p-1} + \epsilon_i,$$

where the $x_{i,j}$ quantities are fixed constants, and the ϵ_i variables are independent $N(0, \sigma^2)$.

- (a) Write the likelihood function.
 (b) The least squares estimates of the regression coefficients are the β_j values that minimize

$$Q(\boldsymbol{\beta}) = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 x_{i,1} - \cdots - \beta_{p-1} x_{i,p-1})^2$$

Are the least squares and maximum likelihood estimates the same in this case (the normal case), or are they different? You don't need to derive $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ in order to see the answer, so don't do it.

- (c) Differentiate $Q(\boldsymbol{\beta})$ with respect to β_0 and set the derivative to zero, obtaining the first *normal equation*.
 (d) Noting that the quantities $\hat{\beta}_0, \dots, \hat{\beta}_{p-1}$ must satisfy the first normal equation, show that the least squares plane must pass through the point $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_{p-1}, \bar{Y})$.
 (e) Defining "predicted" Y_i as $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i,1} + \cdots + \hat{\beta}_{p-1} x_{i,p-1}$, show that $\sum_{i=1}^n \hat{Y}_i = \sum_{i=1}^n Y_i$.
 (f) The *residual* for observation i is defined by $e_i = Y_i - \hat{Y}_i$. Show that the sum of residuals equals exactly zero.
4. For the general fixed effects linear regression model in matrix form (see formula sheet), show that the $n \times p$ matrix of covariances $C(\mathbf{e}, \hat{\boldsymbol{\beta}}) = \mathbf{0}$. Why does this establish that *SSE* and $\hat{\boldsymbol{\beta}}$ are independent?
5. For the general fixed effects linear regression model, tests and confidence intervals for linear combinations of regression coefficients are very useful. Derive the appropriate t distribution and some applications by following these steps. Let \mathbf{a} be a $p \times 1$ vector of constants.
- (a) What is the distribution of $\mathbf{a}'\hat{\boldsymbol{\beta}}$? Show a little work. Your answer includes both the expected value and the variance.
 (b) Now standardize the difference (subtract off the mean and divide by the standard deviation) to obtain a standard normal.
 (c) Divide by the square root of a well-chosen chi-squared random variable, divided by its degrees of freedom, and simplify. Call the result T .
 (d) How do you know numerator and denominator are independent?
 (e) Suppose you wanted to test $H_0 : \mathbf{a}'\boldsymbol{\beta} = c$. Write down a formula for the test statistic.
 (f) Suppose you wanted to test $H_0 : \beta_2 = 0$. Give the vector \mathbf{a} .
 (g) Suppose you wanted to test $H_0 : \beta_1 = \beta_2$. Give the vector \mathbf{a} .
 (h) Letting $t_{\alpha/2}$ denote the point cutting off the top $\alpha/2$ of the t distribution with $n - p$ degrees of freedom, give a $(1 - \alpha) \times 100\%$ confidence interval for $\mathbf{a}'\boldsymbol{\beta}$.

This assignment was prepared by [Jerry Brunner](#), Department of Statistics, University of Toronto. It is licensed under a [Creative Commons Attribution - ShareAlike 3.0 Unported License](#). Use any part of it as you like and share the result freely. The L^AT_EX source code is available from the course website: <http://www.utstat.toronto.edu/~brunner/oldclass/appliedf13>