University of California, Los Angeles
Department of Statistics

Statistics 100C **Instructor: Nicolas Christou**

Week 9 - practice problems

Exercise 1
Consider the model $\mathbf{y} = \mu\mathbf{1} + \boldsymbol{\epsilon}$, where $E(\boldsymbol{\epsilon}) = \mathbf{0}$ and $var(\boldsymbol{\epsilon}) = \boldsymbol{\Sigma}$. The entries of $\boldsymbol{\Sigma}$ are computed using $var(Y_i) = \sigma^2, i = 1, \ldots, n$ and $cov(Y_i, Y_j) = \rho\sigma^2, i \neq j$, with σ^2 and ρ known constants. Suppose the prediction of a new observation Y_0 is of interest. Assume that $\hat{Y}_0 = \mathbf{w}'\mathbf{Y}$ and $E(\hat{Y}_0) = \mu$. By minimizing the mean-squared error, $E(Y_0 - \hat{Y}_0)^2$ show that $\hat{Y}_0 = \hat{\mu} + \mathbf{c}'\boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \hat{\mu}\mathbf{1})$. where $\hat{\mu}$ is the generalized least squares estimate of μ. Note: To find the entries of $\mathbf{c}$ it is reasonable to assume that the new observation Y_0 is also correlated with $\mathbf{y}$, i.e., $cov(Y_i, Y_0) = \rho\sigma^2, i = 1, \ldots, n$.

Exercise 2
Researchers 1 and 2 were working on similar problems.

Using n_1 data points, researcher 1 formed the model $\mathbf{y_1} = \mathbf{X_1}\boldsymbol{\beta} + \boldsymbol{\epsilon_1}$, where $\mathbf{y_1}$ is $n_1 \times 1$, $\mathbf{X_1}$ is $n_1 \times (k+1)$, $\boldsymbol{\beta}$ is $(k+1) \times 1$, and $\boldsymbol{\epsilon_1}$ is $n_1 \times 1$, with $E(\boldsymbol{\epsilon_1}) = \mathbf{0}$ and $\text{cov}(\boldsymbol{\epsilon_1}) = \sigma^2\mathbf{I}$.

Using n_2 data points, researcher 2 formed the model $\mathbf{y_2} = \mathbf{X_2}\boldsymbol{\beta} + \boldsymbol{\epsilon_2}$, where $\mathbf{y_2}$ is $n_2 \times 1$, $\mathbf{X_2}$ is $n_2 \times (k+1)$, $\boldsymbol{\beta}$ is $(k+1) \times 1$, and $\boldsymbol{\epsilon_2}$ is $n_2 \times 1$, with $E(\boldsymbol{\epsilon_2}) = \mathbf{0}$ and $\text{cov}(\boldsymbol{\epsilon_2}) = \sigma^2\mathbf{I}$.

Note: Each researcher is trying to estimate the same coefficient vector $\boldsymbol{\beta}$.

Answer the following questions:

a. Suppose that the researchers worked independently. Give $\hat{\boldsymbol{\beta}}_\mathbf{1}$ and $\hat{\boldsymbol{\beta}}_\mathbf{2}$, their separate least squares estimators.

b. Suppose that they cooperate and pool their data. So now the model will be:

$$\begin{pmatrix} \mathbf{y_1} \\ \mathbf{y_2} \end{pmatrix} = \begin{pmatrix} \mathbf{x_1} \\ \mathbf{x_2} \end{pmatrix} \boldsymbol{\beta} + \begin{pmatrix} \boldsymbol{\epsilon_1} \\ \boldsymbol{\epsilon_2} \end{pmatrix}.$$

Find the combined least squares estimator.

c. Consider the case $n_1 = n_2$ and $\mathbf{X_1} = \mathbf{X_2}$, but $\mathbf{y_1}$ not necessarily equals to $\mathbf{y_2}$. Now answer question (b) again.

d. Consider the situation in (c). Find the variance covariance matrix of $\hat{\boldsymbol{\beta}}$.

e. Consider again the situation in (c). Now assume that for researcher 2 we have $\text{var}(\boldsymbol{\epsilon_2}) = c\sigma^2\mathbf{I}$, while $\text{var}(\boldsymbol{\epsilon_1}) = \sigma^2\mathbf{I}$. Write the matrix $\mathbf{W}$ to be used in weighted least squares.

Exercise 3
Consider the centered and scaled model with two predictors: $y_i = \gamma_0 + \delta_1 Zs_{i1} + \delta_2 Zs_{i2} + \epsilon_i$. Suppose, $\epsilon_i \sim N(0, 1)$. Show that the test statistic for testing $H_0 : \delta_2 = 0$ follows $N(0, 1)$. Now consider the model $y_i = \gamma_0 + \delta_1 Zs_{i1} + \epsilon_i$. Does the test statistic for testing $H_0 : \delta_1 = 0$ in this model follow $N(0, 1)$ too? Please explain.