Regression Analysis Midterm Examination 07 November 2016

1. Consider the simple linear regression model

$$y_i = \beta_0 + \beta_1 x_{i1} + \varepsilon_i,$$

with $\mathrm{E}(\varepsilon_i) = 0$, $\mathrm{Var}(\varepsilon_i) = \sigma^2$, $i = 1, \ldots, n$, and ε_i's uncorrelated.

(a) (15 pts) Suppose β_0 is known. Find the least squares estimator of β_1 and find the bias and variance of this estimator.

(b) (15 pts) Suppose we have fit the simple linear regression model $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1$ with β_0 and β_1 both unknown, but the true regression function is

$$\mathrm{E}(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 .$$

Find the bias of $\hat{\beta}_1$. Compare the variance of $\hat{\beta}_1$ with the variance of the least squares estimator of β_1 under the true model.

2. Consider two independent sets of observations

$$y_{1i} = \beta_0 + \beta_1 x_{1i} + \varepsilon_{1i},\ i = 1, \cdots, n,$$

where the ε_{1i}'s are iid Normal$(0, \sigma^2)$, and

$$y_{2i} = \beta_0 + \beta_2 x_{2i} + \varepsilon_{2i},\ i = 1, \cdots, n$$

where the ε_{2i}'s are iid Normal$(0, \sigma^2)$. Here, y_{1i}'s and y_{2i}'s are observations of the same response variable and x_{1i}'s and x_{2i}'s are observations of the same covariate.

(a) (5 pts) Find the least squares estimator of β_0 based on the first set of observations $(x_{11}, y_{11}), \cdots, (x_{1n}, y_{1n})$.

(b) (5 pts) Find the least squares estimator of $\beta = (\beta_0, \beta_1, \beta_2)^T$ based on the $2n$ observations, and find its mean vector and covariance matrix.

(c) (10 pts) Compare biases and variances of the two estimators of β_0 in (a) and (b).

(d) (5 pts) Find an estimator for σ^2.

(e) (10 pts) Derive a level-α test for the hypotheses $H_0 : \beta_1 = \beta_2$ versus $H_1 : \beta_1 \neq \beta_2$.

TURN OVER

3. Consider the multiple linear regression model

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + \varepsilon_i,\ i = 1, \cdots, n,$$

where the ε_i's are uncorrelated each with zero mean and constant variance σ^2. Let $\beta = (\beta_0, \beta_1, \cdots, \beta_k)^T$, $y = (y_1, \cdots, y_n)^T$, $\varepsilon = (\varepsilon_1, \cdots, \varepsilon_n)^T$ and $X = (x_1, \cdots, x_n)^T$ with $x_i = (1, x_{i1}, \cdots, x_{ik})^T$, $i = 1, \cdots, n$. Then the above model can be written as

$$y = X\beta + \varepsilon\,.$$

40 Assume that the design matrix X is of rank $k+1$.

(a) (5 pts) Write down the least squares problem for estimation of β and find the solution, denoted by $\hat{\beta}$.

(b) (10 pts) Let $s^2 = (y - X\hat{\beta})^T(y - X\hat{\beta})/(n-k-1)$. Show that s^2 is an unbiased estimator of σ^2.

(c) (10) Let $r^2_{y\hat{y}}$ be the correlation coefficient between the y_i's and $\hat{y}_i's$. Show that $r^2_{y\hat{y}} = R^2$.

(d) (5 pts) Let $H = (h_{ij})$ be the hat matrix. Show that $0 \le h_{ii} \le 1$, $i = 1, \cdots, n$.

(e) (10) Show that, for any constant $(k+1)$-vector l, $l'\hat{\beta}$ is a best linear unbiased estimator of $l'\beta$ under the Gauss-Markov conditions.

4. Consider the reparameterization $Y = W\alpha + \varepsilon$ of the model $Y = X\beta + \varepsilon$, where $W = XC$ and C is nonsingular.

(a) (5) Show that the hat matrices H_W and H_X for the two models are the same.

(b) (5) Suppose $\hat{\alpha}$ and $\hat{\beta}$ are least squares estimators of α and β. Express $\hat{\alpha}$ in terms of $\hat{\beta}$.

Regression Analysis Final Examination January 9, 2017

18 1. Consider the linear regression model $Y = X\beta + \varepsilon$.

(a) (6 pts) Give the standardized residuals d_i, $i = 1, \cdots, n$, and explain a use of them in model diagnostics.

(b) (6 pts) Give the studentized residuals r_i, $i = 1, \cdots, n$, and explain a use of them in model diagnostics.

(c) (6 pts) Give the PRESS residuals $e_{(i)}$, $i = 1, \cdots, n$, and explain a use of them in model diagnostics.

8 2. (8 pts) Consider the linear model $Y = X\beta + \varepsilon$. Suppose the ε_i's are uncorrelated with zero mean and variances $\mathrm{Var}(\varepsilon_i) = c_i^2\sigma^2$ where $c_1, \cdots, c_n$ are known positive constants Give the BLUE of β and an unbiased estimator of σ^2.

18 3. Consider the model $y_i = x_i^T\beta + \varepsilon_i$, $i = 1, \cdots, n$, where $\varepsilon_i = \rho\,\varepsilon_{i-1} + u_i$ with $|\rho| < 1$ and u_i's being uncorrelated with mean zero and variance σ^2. Assume that $\mathrm{E}(\varepsilon_1) = 0$, $\mathrm{Var}(\varepsilon_1) = \sigma^2/(1-\rho^2)$ and ε_1 is independent of $u_2, \cdots, u_n$.

(a) (5 pts) Find the generalized least squares estimator of β.

(b) (5 pts) Show that $\hat{\rho} = \sum_{i=2}^n e_i e_{i-1} / \sum_{i=1}^n e_i^2$ is asymptotically unbiased for ρ.

(c) (8 pts) Find the estimated generalized least squares estimator of β and an asymptotically unbiased estimator of σ^2.

17 4. Consider the simple linear regression model $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, where $\mathrm{Var}(\varepsilon_i) = \sigma^2 x_i^2$, $i = 1, \ldots, n$.

(a) (8 pts) Suppose that we use the transformations $y' = y/x$ and $x' = 1/x$. Is this a variance-stabilizing transformation? What are the final estimators of β_0, β_1, and σ^2?

(b) (5 pts) Suppose we use the weighted least squares method with weights $w_i = 1/x_i^2$, $i = 1, \ldots, n$. Is this equivalent to the transformations introduced in part (a)?

15 5. Consider the model $y_i = x_i'\beta + \varepsilon_i$, $i = 1, \cdots, n$, where the ε_i's are uncorrelated with zero mean.

(a) (6 pts) Describe problems caused by influential points, and how we can cope with the problems.

(b) (9 pts) Define three measures to detect influential points.

23

6. Consider the one-way analysis of variance model $y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$, $i = 1, \ldots, 3$, $j = 1, \ldots, n$, where y_{ij} is the jth observation for the ith treatment level, μ is the grand mean, α_i is effect of the ith treatment, and ε_{ij}'s are independent $\mathrm{N}(0, \sigma^2)$ errors.

(a) (5 pts) Write down a linear regression model of the form $Y = X\beta + \varepsilon$ for the data, using indicator variables.

(b) (8 pts) Find the ordinary least squares estimators of μ, α_1, α_2, and α_3.

(c) (10 pts) Write down the One-Way Analysis of Variance table.

7. Consider the regression model $Y = X\beta + \varepsilon$. Describe the following three possible remedies to multicollinearity.

(a) (5 pts) Incomplete principal component regression.

(b) (5 pts) Ridge regression.

(c) (5 pts) Add additional observations.