Regression Methods: Problems MATH-408

Anthony Davison

Problem 1 If $T \sim t_{n-p}$, find the distribution of T^2.

Problem 2 If X_1, X_2 are independent with density $\lambda \exp(-\lambda x)$ for $x > 0$ and $\lambda > 0$, use moment-generating functions to show that $\varepsilon = X_1 - X_2$ has density $(\lambda/2)\exp(-\lambda|\varepsilon|)$ for $\varepsilon \in \mathbb{R}$, and find the moments of ε. Hence say what value of λ ensures that $\mathrm{var}(\varepsilon) = \sigma^2$. Under what circumstances do you think this might be a useful density for the errors in a regression model?

Problem 3 The straight-line regression model has uncorrelated observations $y_j \sim (\beta_0 + \beta_1 x_j, \sigma^2)$, where $j = 1, \dots, n$ and the x_j are real scalars.

(a) What problem arises if $x_1, \dots, x_n$ are all equal? If this is not the case, show that

$$\widehat{\beta}_1 = \frac{\sum(x_j - \overline{x})y_j}{\sum(x_j - \overline{x})^2}, \quad \widehat{\beta}_0 = \overline{y} - \widehat{\beta}_1\overline{x}, \quad h_{jj} = \frac{1}{n} + \frac{(x_j - \overline{x})^2}{\sum_i(x_i - \overline{x})^2}, \quad j = 1, \dots, n,$$

where $\overline{x} = n^{-1}\sum x_j$ and $\overline{y} = n^{-1}\sum y_j$.

(b) If the design points are equally-spaced, i.e., $x_j = cj$ for some non-zero c, show that $\overline{x} = c(n+1)/2$ and $\sum(x_j - \overline{x})^2 = c^2 n(n+1)(n-1)/12$, and deduce that

$$\max_j h_{jj} = h_{nn} = \frac{1}{n} + \frac{3(n-1)}{n(n+1)}.$$

(c) If $x_j = c2^j$ for some non-zero c, show that

$$\overline{x} = c2(2^n - 1)/n, \quad \sum(x_j - \overline{x})^2 = c^2\left\{\frac{4}{3}(4^n - 1) - \frac{4^{n+1} + 4 - 2^{n+3}}{n}\right\},$$

and deduce that $\max_j h_{jj} = h_{nn} \to 3/4$ as $n \to \infty$.

(d) Sketch the designs in (b) and (c), and discuss the limiting distribution of $\widehat{\beta}$ in each case.

Problem 4

(a) In the usual linear model notation, show that a test of the hypothesis $\beta = \beta'$ can be based on $(\widehat{\beta} - \beta')^{\mathrm{T}} X^{\mathrm{T}} X(\widehat{\beta} - \beta')/\sigma^2$, and give its null distribution in a normal model. Give a suitable test statistic and its null distribution when σ^2 is unknown.

(b) The angles of a triangle ABC are each measured once independently and unbiasedly with common error variance σ^2. Find unbiased estimators of the angles at A and B and of σ^2. If the errors are normally distributed, say how you would test whether the triangle is equilateral.

Problem 5 Consider a linear model $y \sim (X\beta, \sigma^2 I_n)$, where $X\beta = X_1\beta_1 + X_2\beta_2$ with full-rank matrices X_1 and X_2, so that

$$\widehat{\beta} = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}y = \begin{pmatrix}\widehat{\beta}_1\\ \widehat{\beta}_2\end{pmatrix} = \begin{pmatrix}X_1^{\mathrm{T}}X_1 & X_1^{\mathrm{T}}X_2\\ X_2^{\mathrm{T}}X_1 & X_2^{\mathrm{T}}X_2\end{pmatrix}^{-1}\begin{pmatrix}X_1^{\mathrm{T}}y\\ X_2^{\mathrm{T}}y\end{pmatrix}.$$

Let $H_1 = X_1(X_1^{\mathrm{T}}X_1)^{-1}X_1^{\mathrm{T}}$, $P_1 = I_n - H_1$, and define H_2 and P_2 similarly; notice that these matrices are symmetric and idempotent.

(a) Use the hint to show that we can write $\widehat{\beta}_2 = A^{22}By$, for suitable matrices A^{22} and B, and deduce that $\widehat{\beta}_2 = (X_2^{\mathrm{T}} P_1 X_2)^{-1} X_2^{\mathrm{T}} P_1 y$, with variance matrix $\sigma^2 (X_2^{\mathrm{T}} P_1 X_2)^{-1}$. Hence interpret $\widehat{\beta}_2$ as the parameter estimate from the regression of $P_1 y$ on the columns of $P_1 X_2$. What is the corresponding expression for $\widehat{\beta}_1$?

(b) Suppose that in a normal linear model, X_2 is a single column that depends on y only through the fitted values from regression of y on X_1, so that X_2 is itself random. Noting that the residuals $P_1 y$ are independent of the fitted values, $H_1 y$, and arguing conditionally on $H_1 y$, show that the t statistic for $\widehat{\beta}_2$ has a distribution that is independent of X_2.

Hint: If the matrix A is partitioned as

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

and the necessary inverses exist, then the elements of the corresponding partition of A^{-1} are

$$\begin{aligned} A^{11} &= (A_{11} - A_{12} A_{22}^{-1} A_{21})^{-1}, \quad A^{22} = (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1}, \\ A^{12} &= -A_{11}^{-1} A_{12} A^{22}, \quad A^{21} = -A_{22}^{-1} A_{21} A^{11}. \end{aligned}$$