
Solution 1 We can write $T \stackrel{\mathrm{D}}{=} Z/\sqrt{W/\nu}$, where $Z \sim \mathcal{N}(0,1)$ and $W \sim \chi^2_\nu$ are independent. Hence

$$T^2 \stackrel{\mathrm{D}}{=} \frac{Z^2}{W/\nu} \sim F_{1,n-p},$$

because $Z^2 \sim \chi^2_1$.

Solution 2 The MGF of ε is

$$M_\varepsilon(t) = \mathrm{E}\{\exp(t\varepsilon)\} = \mathrm{E}\{\exp(tX_1 - tX_2)\} = \mathrm{E}\{\exp(tX_1)\}\mathrm{E}\{\exp(-tX_2)\} = M_X(t)M_X(-t),$$

where

$$M_X(t) = \int_0^\infty e^{tx}\lambda e^{-tx}\,\mathrm{d}x = \frac{\lambda}{\lambda - t}, \quad t < \lambda,$$

so

$$M_\varepsilon(t) = \frac{\lambda^2}{\lambda^2 - t^2}, \quad |t| < \lambda.$$

The given density has MGF

$$\frac{\lambda}{2}\int_{-\infty}^\infty e^{tx}e^{-\lambda|x|}\,\mathrm{d}x = \frac{\lambda}{2}\int_0^\infty (e^{-tx-\lambda x} + e^{tx-\lambda x})\,\mathrm{d}x = \frac{\lambda}{2}\left(\frac{1}{\lambda+t} + \frac{1}{\lambda-t}\right) = \frac{\lambda^2}{\lambda^2-t^2}, \quad |t| < \lambda,$$

so it is the MGF of ε.

Clearly $\mathrm{E}(\varepsilon) = \mathrm{E}(X_1) - \mathrm{E}(X_2) = 0$ and $\mathrm{var}(\varepsilon) = 2\mathrm{var}(X_1) = 2/\lambda^2$. Thus we obtain variance σ^2 by setting $\lambda = \sqrt{2}/\sigma$.

This density has heavier tails than the normal, so it might be useful for dealing with data with symmetric errors but large tails than the normal.

Solution 3

(a) If the x_j are all equal then the matrix $X_{n\times 2}$ is not full-rank, so the parameters cannot be identified. If the x_j are not all equal, then

$$X = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}, \quad X^{\mathrm{T}}X = \begin{pmatrix} n & n\overline{x} \\ n\overline{x} & \sum x_i^2 \end{pmatrix}, \quad (X^{\mathrm{T}}X)^{-1} = \frac{1}{n\sum(x_i - \overline{x})^2}\begin{pmatrix} \sum x_i^2 & -n\overline{x} \\ -n\overline{x} & n \end{pmatrix},$$

and some algebra using the formulae $\widehat{\beta} = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}y$ and $H = X(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}$, so $h_{jj} = (1, x_j)(X^{\mathrm{T}}X)^{-1}(1, x_j)^{\mathrm{T}}$, leads to the given expressions.

(b) Standard formulae for sums of integers and their squares give

$$\sum x_j = c\sum j = cn(n+1)/2, \quad \sum x_j^2 = c^2\sum j^2 = c^2 n(n+1)(2n+1)/6,$$

so $\overline{x} = c(n+1)/2$, $\sum(x_j - \overline{x})^2 = c^2 n(n+1)(n-1)/12$, clearly h_{jj} is maximised for $j = 1, n$, and $x_n - \overline{x} = c(n-1)/2$, giving the stated formula.

(c) This uses the formula for summing a geometric series, i.e., $\sum_{j=1}^n p^j = p(p^n - 1)/(p-1)$ for $p \neq 1$, followed by some algebra.

(d) The sketch is easy. There is limiting normality in (b), but not in (c) (at least in general), because the response at x_n will dominate the limiting distribution. Of course if the errors in (c) were all normal, then there would be limiting normality ...

Solution 4

(a) If $\beta = \beta'$, then $\widehat{\beta} \sim \mathcal{N}_p\{\beta', \sigma^2(X^{\mathrm{T}}X)^{-1}\}$ and therefore $(\widehat{\beta} - \beta')^{\mathrm{T}}X^{\mathrm{T}}X(\widehat{\beta} - \beta')/\sigma^2 \sim \chi^2_p$, independent of the residual sum of squares. If σ^2 is unknown, then it can be estimated by $s^2 = y^{\mathrm{T}}(I - H)y/(n - p)$, and then under the null hypothesis we have

$$F = \frac{(\widehat{\beta} - \beta')^{\mathrm{T}}X^{\mathrm{T}}X(\widehat{\beta} - \beta')/p}{s^2} \sim F_{p.n-p}.$$

(b) Although there are three angles, with angles α, β, γ, say, their sum is the constant $\alpha+\beta+\gamma = \pi$, and so just two angles can vary independently. In terms of α and β, we have $y_A = \alpha + \varepsilon_A$, $y_B = \beta + \varepsilon_B$, and $y_C = \pi - \alpha - \beta + \varepsilon_C$, and this gives the linear model

$$\begin{pmatrix} y_A \\ y_B \\ y_C - \pi \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \begin{pmatrix} \varepsilon_A \\ \varepsilon_B \\ \varepsilon_C \end{pmatrix}.$$

Hence
$$\begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix} = \tfrac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} \pi + y_A - y_C \\ \pi + y_B - y_C \end{pmatrix} = \tfrac{1}{3}\begin{pmatrix} 2y_A + \pi - y_B - y_C \\ 2y_B + \pi - y_A - y_C \end{pmatrix}$$

It is straightforward to show that $s^2 = (y_A + y_B + y_C - \pi)^2/3$.

The triangle is equilateral if $\alpha = \beta = \pi/3$, which corresponds to the setup in (a) with $\beta' = (\pi/3, \pi/3)^{\mathrm{T}}$, and would lead to a test based on an $F_{2,1}$ statistic.

Solution 5

(a) If we write

$$\begin{pmatrix} X_1^{\mathrm{T}}X_1 & X_1^{\mathrm{T}}X_2 \\ X_2^{\mathrm{T}}X_1 & X_2^{\mathrm{T}}X_2 \end{pmatrix}^{-1} = \begin{pmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{pmatrix},$$

then by symmetry we must have $A^{21} = (A^{12})^{\mathrm{T}}$, and therefore

$$\widehat{\beta}_2 = A^{21}X_1^{\mathrm{T}}y + A^{22}X_2^{\mathrm{T}}y = A^{22}X_2^{\mathrm{T}}y - A^{22}X_2^{\mathrm{T}}X_1(X_1^{\mathrm{T}}X_1)^{-1}X_1^{\mathrm{T}}y = A^{22}By,$$

where

$$A^{22} = \left\{X_2^{\mathrm{T}}X_2 - X_2^{\mathrm{T}}X_1(X_1^{\mathrm{T}}X_1)^{-1}X_1^{\mathrm{T}}X_2\right\}^{-1} = (X_2^{\mathrm{T}}P_1X_2)^{-1}, \quad B = X_2^{\mathrm{T}}P_1,$$

and this gives the required expression. Also the fact that $P_1P_1^{\mathrm{T}} = P_1$ gives that $\mathrm{var}(\widehat{\beta}_2)$ equals

$$\mathrm{var}\left\{(X_2^{\mathrm{T}}P_1X_2)^{-1}X_2^{\mathrm{T}}P_1y\right\} = (X_2^{\mathrm{T}}P_1X_2)^{-1}X_2^{\mathrm{T}}P_1\mathrm{var}(y)\left\{(X_2^{\mathrm{T}}P_1X_2)^{-1}X_2^{\mathrm{T}}P_1\right\}^{\mathrm{T}} = \sigma^2(X_2^{\mathrm{T}}P_1X_2)^{-1}.$$

If we consider the estimate resulting from regressing P_1y on the columns of P_1X_2, we get $\widehat{\beta}_2$. Since P_1y is the residual from regressing y on the columns of X_1, we see that $\widehat{\beta}_2$ can be seen as coming from a two-stage procedure: first, we regress both y and X_2 on the columns of X_1; second, we regress the residuals for y from this first regression on the residuals for X_2 from this first regression; the result is $\widehat{\beta}_2$.

By symmetry we must have $\widehat{\beta}_1 = (X_1^{\mathrm{T}}P_2X_1)^{-1}X_1^{\mathrm{T}}P_2y$.

(b) Conditional on $H_1 y$, we see that $X_2 = f(H_1 y)$ is constant, so the results for the usual linear model imply that if X_1 is $n \times p$, then in the usual notation

$$\widehat{\beta}_2 = (X_2^{\mathrm{T}} P_1 X_2)^{-1} X_2^{\mathrm{T}} P_1 y \sim \mathcal{N}\left\{\beta_2, \sigma^2 (X_2^{\mathrm{T}} P_1 X_2)^{-1}\right\} \quad \Longrightarrow \quad \frac{\widehat{\beta}_2 - \beta_2}{S(X_2^{\mathrm{T}} P_1 X_2)^{-1/2}} \sim t_{n-p-1}.$$

As this distribution does not depend on $H_1 y$, this result is also true unconditionally.