smfprob1.tex

SMF PROBLEMS 1. 21.1.2011

Q1. Show that the regression (= least-squares) line for the data $(x_1, y_1), \ldots, (x_n, y_n)$ is

$$y - \overline{y} = r_{xy} \frac{s_y}{s_x} (x - \overline{x}),$$

with $r = r_{xy}$ the sample correlation coefficient, s_x, s_y the sample standard deviations.

Q2. With data y and two predictor variables (regressors) u and v, show that the regression (= least-squares) plane is $y - \overline{y} = a(u - \overline{u} + b(v - \overline{v})$, where a, b satisfy

$$\begin{aligned} a s_{uu} + b s_{uv} &= s_{yu}, \\ a s_{uv} + b s_{vv} &= s_{yv}. \end{aligned}$$

Q3. The *chi-square distribution* with n *degrees of freedom* (df), $\chi^2(n)$, is defined to be that of $X_1^2 + \ldots + X_n^2$ with the X_i independent $N(0, 1)$. Show that $\chi^2(n)$ has
(i) mean n and variance $2n$;
(ii) moment-generating function (MGF) $M(t) = 1/(1 - 2t)^{\frac{1}{2}n}$ $(t < \frac{1}{2})$;
(iii) density

$$f(x) = \frac{1}{2^{\frac{1}{2}n} \Gamma(\frac{1}{2}n)} .x^{\frac{1}{2}n - 1} \exp\{-\frac{1}{2}x\} \quad (x > 0).$$

Q4. For A the design matrix, show that
(i) $P := A(A^T A)^{-1} A^T$ is an $n \times n$ symmetric idempotent matrix $(P^2 = P)$ – a *projection*;
(ii) $I - P$ is a projection.

The *trace* $tr(A)$ of a matrix A is the sum of its diagonal elements. Show that
(a) $tr(A + B) = tr(A) + trB)$, $tr(AB) = tr(BA)$;
(b) $tr(P) = p$, $tr(I - P) = n - p$.

NHB