HG 4Nov 2019

Supplement to the lecture 4/11 2019 on the multivariate normal distribution

The last example of the lecture seemed to cause some confusion.

If

(1) $$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left[\begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix}\right] = N\left[E\begin{pmatrix} X \\ Y \end{pmatrix}, \operatorname{cov}\begin{pmatrix} X \\ Y \end{pmatrix}\right]$$

then, according to a property of the multinormal distribution, the regression of Y with respect to X, is necessarily linear, $E(Y \mid x) = \alpha + \beta x$, and homoscedastic, $\operatorname{var}(Y \mid x) = \tau^2$ constant.

The parameters α, β, τ^2 are not new additional parameters, but are determined by the original parameters in (1) by

(2) $$\beta = \frac{\sigma_{xy}}{\sigma_x^2}, \quad \alpha = \mu_y - \beta\mu_x, \text{ and } \tau^2 = \sigma_y^2(1-\rho^2), \text{ where the correlation } \rho = \frac{\sigma_{xy}}{\sigma_x\sigma_y}$$

This follows from earlier exercises in the course (see, e.g., the week 41 exercises), but it may be good to repeat the algebra here. The easiest way to prove (2) is probably to use the theorem of total expectation:

Proof of (2):

We have $\mu_y = E(Y) = E\left[E(Y \mid X)\right] = E\left(\alpha + \beta X\right) = \alpha + \beta\mu_x$. Hence

(3) $$\alpha = \mu_y - \beta\mu_x$$

Next $\sigma_{xy} = \operatorname{cov}(X,Y) = E\left(XY\right) - \left(EX\right)\left(EY\right) = E\left(XY\right) - \mu_x\mu_y$, and

$$E\left(XY\right) = E\left[X \cdot E(Y \mid X)\right] = E\left[X(\alpha + \beta X)\right] = \alpha EX + \beta E\left(X^2\right) = \alpha\mu_x + \beta\left(\operatorname{var} X + \left(EX\right)^2\right) =$$

$$= \alpha\mu_x + \beta\left(\sigma_x^2 + \mu_x^2\right) = \mu_x(\alpha + \beta\mu_x) + \beta\sigma_x^2 = \mu_x\mu_y + \beta\sigma_x^2$$

Hence, $\sigma_{xy} = E\left(XY\right) - \mu_x\mu_y = \mu_x\mu_y + \beta\sigma_x^2 - \mu_x\mu_y = \beta\sigma_x^2$, giving

(4) $$\beta = \frac{\sigma_{xy}}{\sigma_x^2}$$

determining α and β from (1). We use the variance-formula to determine τ^2 :

$$\sigma_y^2 = \operatorname{var} Y = E\left[\operatorname{var}(Y \mid X)\right] + \operatorname{var}\left[E(Y \mid X)\right] = E\left(\tau^2\right) + \operatorname{var}\left(\alpha + \beta X\right) = \tau^2 + \beta^2\sigma_x^2$$

Hence, $\tau^2 = \sigma_y^2 - \beta^2\sigma_x^2 = \sigma_y^2 - \dfrac{\sigma_{xy}^2}{\sigma_x^4}\sigma_x^2 = \sigma_y^2 - \dfrac{\sigma_{xy}^2}{\sigma_x^2\sigma_y^2}\sigma_y^2 = \sigma_y^2 - \rho^2\sigma_y^2$, giving

(5) $$\tau^2 = \sigma_y^2(1-\rho^2)$$

End of proof