The multivariate normal distribution and related topics

math612.0 A1: From numbers through algebra to calculus and linear algebra

Gunnar Stefansson (editor) with contributions from very many students

March 7, 2022

Transformations of random variables

Recall that if X is a vector of continuous random variables with a joint probability density function and if $Y = h(X)$ such that h is a 1-1 function and continuously differentiable with inverse g so $X = g(Y)$, then the density of Y is given by

$$f_Y(y) = f(g(y))|J|$$

The multivariate normal distribution

Univariate normal transforms

The general univariate normal distribution with density

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

is a special case of the multivariate version.

Transforms to lower dimensions

If $Y \sim n\,(\boldsymbol{\mu}, \Sigma)$ is a random vector of length n and A is an $m \times n$ matrix of rank $m \leq n$, then $AY \sim n(A\mu, A\Sigma A')$.

The OLS estimator

Suppose $Y \sim n(X\beta, \sigma^2 I)$. The ordinary least squares estimator, when the $n \times p$ matrix is of full rank, p, where $p \leq n$, is:

$$\hat{\beta} = (X'X)^{-1}X'Y$$

The random variable which describes the process giving the data and estimate is:

$$b = (X'X)^{-1}X'Y$$

It follows that

$$\hat{\beta} \sim n(\beta, \sigma^2(X'X)^{-1})$$