Econometrics, Birkbeck College

Walter Beckert, 2020/21

1. Multivariate Normal Distribution

Definition: Let $y \in \mathbb{R}^N$ be a realization of the vector valued random variable Y, with support[1] $\mathbb{R}^N$. Then, the N components of the random vector Y have a joint distribution which is multivariate normal if their joint density $f_Y(y)$ is of the form

$$f_Y(y) = \frac{1}{(2\pi)^{\frac{N}{2}}|\Omega|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(y-\mu)'\Omega^{-1}(y-\mu)\right),$$

where $\mu = \mathbb{E}[Y]$, $\Omega = \mathbb{E}[(Y-\mu)(Y-\mu)']$, $\mathrm{rk}(\Omega) = N$, and the diagonal elements of Ω are the variances of the components of Y, while the off-diagonal elements are the covariances. Note that the distribution depends solely on the parameters μ and Ω, (N1) so the multivariate normal distribution is completely characterized by the first two (centered) moments.

Linear transformations of normals are also normal. (N2) If Y is distributed $N(\mu, \Omega)$ and α and Γ are a commensurate vector and matrix, respectively, then $Z = \alpha + \Gamma Y$ is dtsributed $N(\alpha + \Gamma\mu, \Gamma\Omega\Gamma')$.

(N3) In the case of the multivariate normal, zero covariance is equivalent to independence[2]. Hence, if $\Omega_{nm} = 0$ for $n \neq m$, $n, m = 1, \cdots, N$, then the joint density factorizes into N marginal univariate normal densities,

$$f_Y(y) = \prod_{n=1}^{N} \frac{1}{(2\pi\Omega_{nn})^{\frac{1}{2}}} \exp\left(-\frac{1}{2\Omega_{nn}}(y_n-\mu_n)^2\right).$$

The case when the idiosyncratic variances $\mathrm{var}(y_n) = \Omega_{nn}$ differ across n is called heteroskedasticity. If, on the other hand, the idiosyncratic variances are homoskedastic, i.e. $\mathrm{var}(y_n) = \sigma^2$ for all n, then the joint density of Y simplifies further,

$$f_Y(y) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left(-\frac{1}{2\sigma^2}\sum_{n=1}^{N}(y_n-\mu_n)^2\right).$$

[1] The support of a random variable is the set of possible realizations of the random variable, i.e. the union of set of realizations that occur with positive probability.

[2] Recall that two events A and B are independent if, and only if, $\Pr(A \cap B) = Pr(A)Pr(B)$.

Let $Y' = [Y_1', Y_2']$, where $Y_1 \in \mathbb{R}^M$ and $Y_2 \in \mathbb{R}^K$, $N = M + K$, and partition μ and Ω accordingly as $\mu' = [\mu_1', \mu_2']$ and Ω into the block-diagonal matrix that has Ω_{11} as the upper-left $M \times M$ submatrix, Ω_{22} as the lower-right $K \times K$ submatrix, and $\Omega_{12} = \Omega_{21}'$ as the off-diagonal $K \times M$ and $M \times K$ blocks. In the case of the multivariate normal, the conditional distribution of of Y_1, given Y_2, is also multivariate normal, with conditional mean $\mu_{1|2}(Y_2) := \mathbb{E}[Y_1|Y_2] = \mu_1 + \Omega_{12}\Omega_{22}^{-1}(Y_2 - \mu_2)$ and conditional variance $\Omega_{1|2} := \text{var}(Y_1|Y_2) = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}$. Hence,

$$f_{Y_1|Y_2}(y_1|y_2) = \frac{1}{(2\pi)^{\frac{M}{2}}|\Omega_{1|2}|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(y_1 - \mu_{1|2}(y_2))'\Omega_{1|2}^{-1}(y_1 - \mu_{1|2}(y_2))\right).$$

2. Useful Results for Interval Estimation (Confidence Intervals) and Hypothesis Testing

Theorem 1: (χ^2 Distribution with m degrees of freedom)

Suppose $\epsilon \in \mathbb{R}^N$, with $\epsilon \sim N(\mathbf{0}, \sigma^2\mathbf{I}_N)$, $\sigma^2 > 0$, and $\mathbf{A}$ is a square, symmetric and idempotent matrix of dimension $N \times N$. Then,

$$\epsilon'\mathbf{A}\epsilon \sim \sigma^2\chi^2_m,$$

where $m = rk(\mathbf{A}) = tr(\mathbf{A})$.

(Non-central χ^2 distribution with m degrees of freedom)

Suppose $\epsilon \in \mathbb{R}^N$, with $\epsilon \sim N(\mu, \sigma^2\mathbf{I}_N)$, $\sigma^2 > 0$, and $\mathbf{A}$ is a square, symmetric and idempotent matrix of dimension $N \times N$. Then,

$$\epsilon'\mathbf{A}\epsilon \sim \sigma^2\chi^2_m(\lambda),$$

where $m = rk(\mathbf{A}) = tr(\mathbf{A})$ and non-centrality parameter $\lambda = \mu'A\mu \geq 0$.

Theorem 2: (F distribution with ν_1 numerator and ν_2 denominator degrees of freedom)

Suppose $Y_1 \sim \chi^2_{\nu_1}$, $Y_2 \sim \chi^2_{\nu_2}$, and Y_1 and Y_2 are independent. Then,

$$\frac{Y_1/\nu_1}{Y_2/\nu_2} \sim F_{\nu_1,\nu_2}.$$

(Non-central F distribution with ν_1 numerator and ν_2 denominator degrees of freedom)

Suppose $Y_1 \sim \chi^2_{\nu_1}(\lambda)$, $\lambda > 0$, $Y_2 \sim \chi^2_{\nu_2}$, *and* Y_1 *and* Y_2 *are independent. Then,*

$$\frac{Y_1/\nu_1}{Y_2/\nu_2} \sim F_{\nu_1,\nu_2}(\lambda),$$

with non-centrality parameter $\lambda > 0$.

Theorem 3: (t distribution with ν degrees of freedom)

Suppose $X_1 \sim N(0,1)$, $X_2 \sim \chi^2_\nu$, *and* X_1 *and* X_2 *are independent. Then,*

$$\frac{X_1}{\sqrt{X_2/\nu}} \sim t_\nu.$$

(Non-central t distribution with ν degrees of freedom)

Suppose $X_1 \sim N(\mu,1)$, $X_2 \sim \chi^2_\nu$, *and* X_1 *and* X_2 *are independent. Then,*

$$\frac{X_1}{\sqrt{X_2/\nu}} \sim t_\nu(\mu),$$

with non-centrality parameter μ.