Chapter 6

Normal errors

6 Normal errors **1**

1 The multivariate normal and related distributions 1

2 Rotation of axes 2

3 Facts about the multivariate normal 4

4 Least squares 5

1 The multivariate normal and related distributions

Let $Z_1, Z_2, \dots, Z_n$ be independent $N(0,1)$ random variables. When treated as the coordinates of a point in $\mathbb{R}^n$ they define a random vector $\mathbf{Z}$, whose (joint) density function is

$$f(\mathbf{z}) = (2\pi)^{-n/2} \exp\left(-\tfrac{1}{2}\sum\nolimits_i z_i^2\right) = (2\pi)^{-n/2} \exp\left(-\tfrac{1}{2}\|\mathbf{z}\|^2\right).$$

Such a random vector is said to have a ***spherical normal distribution***. That is, $\mathbf{Z} \sim N(0, I_n)$.

(i) The ***chi-square***, χ^2_n, is defined as the distribution of the sum of squares $Z_1^2 + \dots + Z_n^2$ of independent $N(0,1)$ random variables. The ***noncentral chi-square***, $\chi^2_n(\gamma)$, with noncentrality parameter $\gamma \ge 0$ is defined as the distribution of the sum of squares $(Z_1+\gamma)^2 + Z_2 \cdots + Z_n^2$.

(ii) If $Z \sim N(0,1)$ is independent of $S_k^2 \sim \chi^2_k$ then

$$\frac{Z}{\sqrt{S_k^2/k}} \text{ has a } t\text{-distribution on } k \text{ degrees of freedom } (t_k)$$

(iii) If $S_k^2 \sim \chi_k^2$ is independent of $S_\ell^2 \sim \chi_\ell^2$ then

$$\frac{S_\ell^2/\ell}{S_k^2/k} \text{ has an } F\text{-distribution on } \ell \text{ and } k \text{ degrees of freedom } (F_{\ell,k})$$

2 Rotation of axes

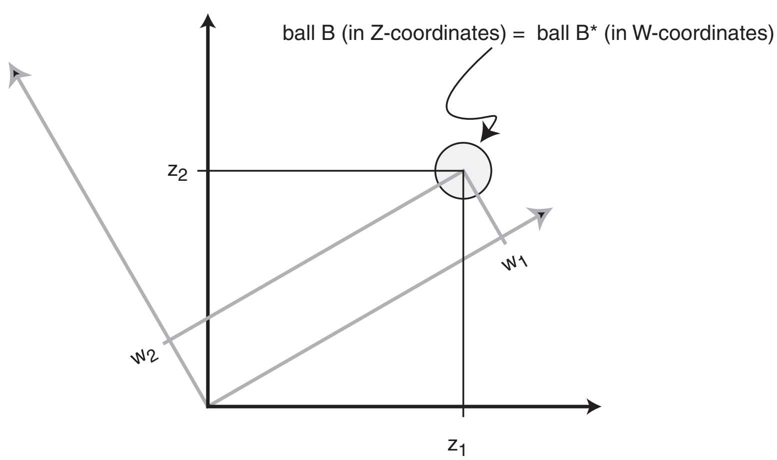
The spherical symmetry of the density $f(\cdot)$ is responsible for an important property of multivariate normals. Let $\mathbf{q}_1, \dots, \mathbf{q}_n$ be a new orthonormal basis for $\mathbb{R}^n$, and let

$$\mathbf{Z} = W_1\mathbf{q}_1 + \cdots + W_n\mathbf{q}_n$$

be the representation for $\mathbf{Z}$ in the new basis.

<6.1> **Theorem.** *The $W_1, \dots, W_n$ are also independent $N(0,1)$ distributed random variables.*

If you know about multivariate moment generating functions this is easy to establish using the matrix representation $\mathbf{Z} = Q\mathbf{W}$, where Q is the orthogonal matrix with columns $\mathbf{q}_1, \dots, \mathbf{q}_n$.



A more intuitive explanation is based on the approximation

$$\mathbb{P}\{\mathbf{Z} \in B\} \approx f(\mathbf{z})(\text{volume of } B)$$

for a small ball B centered at $\mathbf{z}$. The transformation from $\mathbf{Z}$ to $\mathbf{W}$ corresponds to a rotation, so

$$\mathbb{P}\{\mathbf{Z} \in B\} = \mathbb{P}\{\mathbf{W} \in B^*\},$$

where B^* is a ball of the same radius, but centered at the point $\mathbf{w} = (w_1, \dots, w_n)$ for which $w_1\mathbf{q}_1 + \cdots + w_n\mathbf{q}_n = \mathbf{z}$. The last equality implies $\|\mathbf{w}\| = \|\mathbf{z}\|$, from which we get

$$\mathbb{P}\{\mathbf{W} \in B^*\} \approx (2\pi)^{-n/2} \exp(-\tfrac{1}{2} \|\mathbf{w}\|^2)(\text{volume of } B^*).$$

That is, $\mathbf{W}$ has the asserted spherical normal density.

To prove results about the spherical normal it is often merely a matter of transforming to an appropriate orthonormal basis.

<6.2> **Theorem.** *Let $\mathcal{X}$ be an m-dimensional subspace of $\mathbb{R}^n$. Let $\mathbf{Z}$ be a vector of independent $N(0,1)$ random variables, and $\boldsymbol{\mu}$ be a vector of constants. Then*

(i) the projection $\widehat{\mathbf{Z}}$ of $\mathbf{Z}$ onto $\mathcal{X}$ is independent of the projection $\mathbf{Z} - \widehat{\mathbf{Z}}$ of $\mathbf{Z}$ onto $\mathcal{X}^\perp$, the orthogonal complement of $\mathcal{X}$.

(ii) $\left\|\widehat{\mathbf{Z}}\right\|^2$ has a χ^2_m distribution.

(iii) $\|\mathbf{Z} + \boldsymbol{\mu}\|^2$ has a noncentral $\chi^2_n(\gamma)$ distribution, with $\gamma = \|\boldsymbol{\mu}\|$.

(iv) $\left\|\widehat{\mathbf{Z}} + \boldsymbol{\mu}\right\|^2$ has a noncentral $\chi^2_m(\gamma)$ distribution, with $\gamma = \|\boldsymbol{\mu}\|$.

Proof Let $\mathbf{q}_1, \dots \mathbf{q}_n$ be an orthonormal basis of $\mathbb{R}^n$ such that $\mathbf{q}_1, \dots, \mathbf{q}_m$ span the space $\mathcal{X}$ and $\mathbf{q}_{m+1}, \dots, \mathbf{q}_n$ span $\mathcal{X}^\perp$. If $\mathbf{Z} = W_1\mathbf{q}_1 + \cdots + W_n\mathbf{q}_n$ then

$$\begin{aligned}
\widehat{\mathbf{Z}} &= W_1\mathbf{q}_1 + \cdots + W_m\mathbf{q}_m, \\
\mathbf{Z} - \widehat{\mathbf{Z}} &= W_{m+1}\mathbf{q}_{m+1} + \cdots + W_n\mathbf{q}_n, \\
\|\mathbf{Z}\|^2 &= W_1^2 + \cdots + W_m^2,
\end{aligned}$$

from which the first two asserted properties follow.

For the third and fourth assertions, choose the basis so that $\boldsymbol{\mu} = \gamma\mathbf{q}_1$. Then

$$\begin{aligned}
\mathbf{Z} + \boldsymbol{\mu} &= (W_1 + \gamma)\mathbf{q}_1 + W_2\mathbf{q}_2 + \dots \qquad\qquad + W_n\mathbf{q}_n \\
\widehat{\mathbf{Z}} + \boldsymbol{\mu} &= (W_1 + \gamma)\mathbf{q}_1 + W_2\mathbf{q}_2 + \cdots + W_m\mathbf{q}_m
\end{aligned}$$

from which we get the noncentral chi-squares.

□

<6.3> **Example.** Suppose $X_1, \dots, X_n$ are independent random variables, each distributed $N(\mu, \sigma^2)$. Define $\overline{X} = n^{-1}\sum_{i\le n} X_i$ and $S^2 = \sum_{i\le n}(X_i - \overline{X})^2$. Many textbooks prove the following assertion in a gruesome way:

$$\overline{X} \sim N(\mu, \sigma^2/n) \text{ independent of } S^2/\sigma^2 \sim \chi^2_{n-1}.$$

The clean proof uses the fact that the random variables $Z_i = (X_i - \mu)/\sigma$ are independent $N(0,1)$'s, so that $\mathbf{Z} = (Z_1, \dots, Z_n) \sim N(0, I_n)$. Define $q_1 = \mathbb{1}/\sqrt{n}$ then find $q_2, \dots, q_n$ so that $\{q_i : 1 \le i \le n\}$ is an onb for $\mathbb{R}^n$. (Actually it is not necessary to calculate $q_2, \dots, q_n$ explicitly. It suffices to know that such q_i's exist.

From Theorem <6.1>, if

$$\mathbf{Z} = W_1 q_1 + \dots + W_n q_n$$

then the W_i's are independent $N(0,1)$. In particular, $\overline{Z} = \mathbb{1}^T\mathbf{Z}/n = W_1 \sim N(0,1)$ so that

$$\overline{X} = \mu + \sigma\overline{Z} \sim N(\mu, \sigma^2).$$

Also $\mathbf{Z} - \overline{Z}\mathbb{1} = \sum_{i=2}^n W_i q_i$ so that

$$S^2 = \sigma^2 \sum\nolimits_{i\le n}(Z_i - \overline{Z})^2 = \sum\nolimits_{2\le i\le n} W_i^2 \sim \chi^2_{n-1}.$$

The independence comes from the fact that $\overline{X}$ is a function of W_1 and S^2 is a function of $W_2, \dots, W_n$.

□

3 Facts about the multivariate normal

Suppose $Z \sim N(0, I_n)$ and μ is an $m \times 1$ vector of constants. If A is an $m \times n$ matrix of constants then the random vector $X = \mu + AZ$ has expected value μ and variance matrix $V = AA'$, and moment generating function

$$\mathbb{E}\exp(t^T X) = \exp(t^T\mu + t^T AA^T t/2) == \exp(t^T\mu + t^T V t/2).$$

The distribution of X depends only on μ and V. The random vector X has a $N(\mu, V)$ distribution.

If γ is a $k \times 1$ vector of constants and B is a $k \times m$ matrix of constants then

$$\gamma + BX = (\gamma + B\mu) + BAZ \sim N(\gamma + B\mu, BVB').$$

4 Least squares

Much of the distribution theory for least squares has been worked out for the simple model where $y = \mu + \xi \sim N(\mu, \sigma^2 I_n)$, where μ is assumed to lie in some p-dimensional subspace $\mathcal{X}$ of $\mathbb{R}^n$. Both μ and σ^2 are unknown.

Let $q_1, \dots, q_n$ be an onb for $\mathbb{R}^n$ such that $q_1, \dots, q_p$ are an onb for $\mathcal{X}$ and $q_{p+1}, \dots, q_n$ are an onb for $\mathcal{X}^\perp$. Write ξ as $\sigma \mathbf{Z}$, where

$$\mathbf{Z} = [Z_1, \dots, Z_n] = \sum\nolimits_{i \le n} W_i q_i.$$

From Theorem <6.1>, $\mathbf{W} \sim N(0, I_n)$.

The matrix

$$H = \sum\nolimits_{i \le p} q_i q_i^T$$

projects vectors orthogonally onto $\mathcal{X}$. Thus

$$\widehat{y} = H(\mu + \sigma \mathbf{Z}) = \mu + \sigma H \mathbf{Z} = \mu + \sigma \sum\nolimits_{i \le p} W_i q_i.$$

To be continued.