# Appendix 2

## Gaussian (Normal) and $\chi^2$ Distributions

### A2.1 Gaussian probability distribution

Given a vector $\mathbf{X}$ of random variables $x_i$ for $i = 1, \ldots, N$, with mean $\overline{\mathbf{X}} = E[\mathbf{X}]$, where $E[\cdot]$ represents the expected value, and $\Delta\mathbf{X} = \mathbf{X} - \overline{\mathbf{X}}$, the covariance matrix $\Sigma$ is an $N \times N$ matrix given by

$$\Sigma = E[\Delta\mathbf{X}\,\Delta\mathbf{X}^{\mathsf{T}}]$$

so that $\Sigma_{ij} = E[\Delta x_i \Delta x_j]$. The diagonal entries of the matrix $\Sigma$ are the variances of the individual variables $x_i$, whereas the off-diagonal entries are the cross-covariance values.

The variables $x_i$ are said to conform to a joint Gaussian distribution, if the probability distribution of $\mathbf{X}$ is of the form

$$P(\overline{\mathbf{X}} + \Delta\mathbf{X}) = (2\pi)^{-N/2} \det(\Sigma^{-1})^{1/2} \exp\left(-(\Delta\mathbf{X})^{\mathsf{T}}\Sigma^{-1}(\Delta\mathbf{X})/2\right) \qquad \text{(A2.1)}$$

for some positive-semidefinite matrix $\Sigma^{-1}$. It may be verified that $\overline{\mathbf{X}}$ and $\Sigma$ are the mean and covariance of the distribution. A Gaussian distribution is uniquely determined by its mean and covariance. The factor $(2\pi)^{-N/2}\det(\Sigma^{-1})^{1/2}$ is just the normalizing factor necessary to make the total integral of the distribution equal to 1.

In the special case where $\Sigma$ is a scalar matrix $\Sigma = \sigma^2 \mathtt{I}$ the Gaussian PDF takes a simple form

$$P(\mathbf{X}) = (\sqrt{2\pi}\sigma)^{-N} \exp\left(-\sum_{i=1}^{N}(x_i - \bar{x}_i)^2/2\sigma^2\right)$$

where $\mathbf{X} = (x_1, x_2, \ldots, x_N)^{\mathsf{T}}$. This distribution is called an *isotropic Gaussian distribution*.

**Mahalanobis distance.** Note that in this case the value of the PDF at a point $\mathbf{X}$ is simply a function of the Euclidean distance $\left(\sum_{i=1}^{N}(x_i - \bar{x}_i)^2\right)^{1/2}$ of the point $\mathbf{X}$ from the mean $\overline{\mathbf{X}} = (\bar{x}_1, \ldots, \bar{x}_N)^{\mathsf{T}}$. By analogy with this one may define the *Mahalanobis distance* between two vectors $\mathbf{X}$ and $\mathbf{Y}$ to be

$$\|\mathbf{X} - \mathbf{Y}\|_{\Sigma} = \left((\mathbf{X} - \mathbf{Y})^{\mathsf{T}}\Sigma^{-1}(\mathbf{X} - \mathbf{Y})\right)^{1/2} \quad .$$

One verifies that for a positive-definite matrix $\Sigma$, this defines a metric on $\mathbb{R}^N$. Using this notation, the general form of the Gaussian PDF may be written as

$$P(\mathbf{X}) \approx \exp\left(-\|\mathbf{X}-\overline{\mathbf{X}}\|_{\Sigma}^2/2\right)$$

where the normalizing factor has been omitted. Thus, the value of the Gaussian PDF is a function of the Mahalanobis distance of the point $\mathbf{X}$ from the mean.

**Change of coordinates.** Since $\Sigma$ is symmetric and positive-definite, it may be written as $\Sigma = \mathtt{U}^\mathsf{T}\mathtt{D}\mathtt{U}$, where $\mathtt{U}$ is an orthogonal matrix and $\mathtt{D} = (\sigma_1^2, \sigma_2^2, \ldots, \sigma_N^2)$ is diagonal. Writing $\mathbf{X}' = \mathtt{U}\mathbf{X}$ and $\overline{\mathbf{X}}' = \mathtt{U}\overline{\mathbf{X}}$, and substituting in (A2.1), leads to

$$\begin{aligned}\exp\left(-(\mathbf{X}-\overline{\mathbf{X}})^\mathsf{T}\Sigma^{-1}(\mathbf{X}-\overline{\mathbf{X}})/2\right) &= \exp\left(-(\mathbf{X}'-\overline{\mathbf{X}}')^\mathsf{T}\mathtt{U}\Sigma^{-1}\mathtt{U}^\mathsf{T}(\mathbf{X}'-\overline{\mathbf{X}}')/2\right)\\ &= \exp\left(-(\mathbf{X}'-\overline{\mathbf{X}}')^\mathsf{T}\mathtt{D}^{-1}(\mathbf{X}'-\overline{\mathbf{X}}')/2\right)\end{aligned}$$

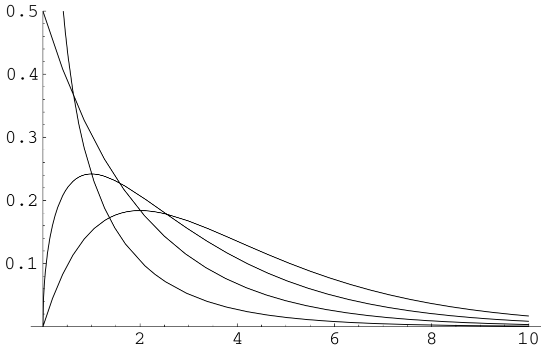
Thus, the orthogonal change of coordinates from $\mathbf{X}$ to $\mathbf{X}' = \mathtt{U}\mathbf{X}$ transforms a general Gaussian PDF into one with diagonal covariance matrix. A further scaling by $\sigma_i$ in each coordinate direction may be applied to transform it to an isotropic Gaussian distribution. Equivalently stated, a change of coordinates may be applied to transform Mahalanobis distance to ordinary Euclidean distance.

## A2.2 $\chi^2$ distribution

The $\chi_n^2$ distribution is the distribution of the sum of squares of $n$ independent Gaussian random variables. As applied to a Gaussian random vector $\mathbf{v}$ with non-singular covariance matrix $\Sigma$, the value of $(\mathbf{v}-\bar{\mathbf{v}})^\mathsf{T}\Sigma^{-1}(\mathbf{v}-\bar{\mathbf{v}})$ satisfies a $\chi_n^2$ distribution, where $n$ is the dimension of $\mathbf{v}$. If the covariance matrix $\Sigma$ is singular, then we must replace $\Sigma^{-1}$ with the pseudo-inverse $\Sigma^+$. In this case

- *If $\mathbf{v}$ is a Gaussian random vector with mean $\bar{\mathbf{v}}$ and covariance matrix $\Sigma$, then the value of $(\mathbf{v}-\bar{\mathbf{v}})^\mathsf{T}\Sigma^+(\mathbf{v}-\bar{\mathbf{v}})$ satisfies a $\chi_r^2$ distribution, where $r = \mathrm{rank}\Sigma$.*

The cumulative chi-squared distribution is defined as $F_n(k^2) = \int_0^{k^2} \chi_n^2(\xi)d\xi$. This represents the probability that the value of the $\chi_n^2$ random variable is less than $k^2$. Graphs of the $\chi_n^2$ distribution and inverse cumulative $\chi_n^2$ distributions for $n = 1, \ldots, 4$ are shown in figure A2.1 A program for computing the cumulative chi-squared distribution $F_n(k^2)$ is given in [Press-88]. Since it is a monotonically increasing function, one may compute the inverse function by any simple technique such as subdivision, and values are tabulated in table A2.1 (compare with figure A2.1).



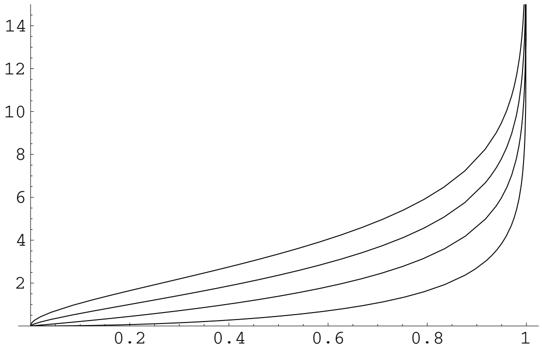




Fig. A2.1. *The $\chi_n^2$ distribution (left) and inverse cumulative $\chi_n^2$ distribution $F_n^{-1}$ (right) for $n = 1, \ldots, 4$. In both cases, graphs are for $n = 1, \ldots, 4$ bottom to top (at middle point of horizontal axis).*

| $n$ | $\alpha = 0.95$ | $\alpha = 0.99$ |
|---|---|---|
| 1 | 3.84 | 6.63 |
| 2 | 5.99 | 9.21 |
| 3 | 7.81 | 11.34 |
| 4 | 9.49 | 13.28 |

Table A2.1. *Values of $k^2$ for which $F_n(k^2)$, the cumulative $\chi^2$ distribution with $n$ degrees of freedom, equals $\alpha$, i.e. $k^2 = F_n^{-1}(\alpha)$, where $\alpha$ is the probability.*