Asymptotic Statistics

Lecture 3

The multivariate normal distribution and the multivariate CLT

Univariate normal distribution

For $\mu \in \mathbb{R}$, $\sigma > 0$, continuous distribution $N(\mu, \sigma^2)$ on $\mathbb{R}$ with density

$$x \mapsto \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}(x-\mu)^2/\sigma^2}.$$

We define $N(\mu, 0)$ to be the Dirac measure δ_μ which puts all mass at the point μ.

We want to generalize to $\mathbb{R}^k$-valued random vectors.

Mean and covariance of random vectors

Let X be a random vector in $\mathbb{R}^k$, $X = (X_1, \ldots, X_k)$. Define its expectation and covariance matrix by

$$\mathbb{E}X = \begin{pmatrix} \mathbb{E}X_1 \\ \mathbb{E}X_2 \\ \vdots \\ \mathbb{E}X_k \end{pmatrix}, \quad \mathbb{C}\mathrm{ov}X = \begin{pmatrix} \mathbb{C}\mathrm{ov}(X_1, X_1) & \cdots & \mathbb{C}\mathrm{ov}(X_1, X_k) \\ \mathbb{C}\mathrm{ov}(X_2, X_1) & \cdots & \mathbb{C}\mathrm{ov}(X_2, X_k) \\ \vdots & & \vdots \\ \mathbb{C}\mathrm{ov}(X_k, X_1) & \cdots & \mathbb{C}\mathrm{ov}(X_k, X_k) \end{pmatrix}.$$

Mean and covariance of random vectors

Lemma.

Let X be a random vector in $\mathbb{R}^k$, A a matrix, b a vector. Then

(i) $\mathbb{E}(AX + b) = A\mathbb{E}X + b$,

(ii) $\mathbb{C}\text{ov}(AX) = A(\mathbb{C}\text{ov}X)A^T$,

(iii) $\mathbb{C}\text{ov}X$ is symmetric and nonnegative definite,

(iv) $P(X \in \mathbb{E}X + \text{range}(\mathbb{C}\text{ov}X)) = 1$.

Sketch of proof.

(i)+(ii): do yourself! (iii): use that variances are nonnegative.
(iv): use that $\text{range}(M^T) = \text{kernel}(M)^\perp$. □

Square root of a covariance matrix

Lemma.
Every $k \times k$ covariance matrix Σ can be written as $\Sigma = LL^T$ for a $k \times k$ matrix L.

Proof.
Σ is a symmetric, nonnegative definite matrix. Hence, there exists an orthonormal basis of eigenvectors of Σ, with corresponding nonnegative eigenvalues. Let O be the orthogonal matrix of eigenvectors and let D be the diagonal matrix with the corresponding eigenvalues on the diagonal. Then $\Sigma = ODO^T$. Then we can take $L = O\sqrt{D}O^T$ (check!), where $\sqrt{D}$ is the matrix obtained from D by taking the square roots of the diagonal entries. □

The matrix L in the proof is called the positive square root of Σ.

Multivariate normal distribution

For random variables, we have $X \sim N(\mu, \sigma^2)$ if and only if X is distributed as $\mu + \sigma Z$, where Z is standard normal.

Definition.
A random vector X in $\mathbb{R}^k$ is said to have a multivariate normal distribution with parameters μ and Σ, if X is distributed as $\mu + LZ$, for $Z = (Z_1, \ldots, Z_k)$ a vector of independent, standard normal variables and L such that $\Sigma = LL^T$. Notation: $X \sim N_k(\mu, \Sigma)$.

Note that if $X \sim N_k(\mu, \Sigma)$, then $\mathbb{E}X = \mu$, $\mathbb{C}\text{ov}X = \Sigma$.

If $X \sim N_k(0, I)$, i.e. X is a vector of independent standard normal variables, then X is called a standard normal vector in $\mathbb{R}^k$.

Density of the multivariate normal distribution

Lemma.
If $X \sim N_k(\mu, \Sigma)$ for a nonsingular covariance matrix Σ, it has density

$$x \mapsto \frac{1}{(2\pi)^{k/2}\sqrt{\det \Sigma}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}.$$

Sketch of proof.
For $X = Z$ standard normal it follows from independence of the marginals. For general X, write $X = \mu + LZ$ and compute its distribution function using a change of variables. □

Properties of the normal vectors

Lemma.
We have $X \sim N_k(\mu, \Sigma)$ if and only if $a^T X \sim N_1(a^T\mu, a^T\Sigma a)$ for every $a \in \mathbb{R}^k$.

N.B. In particular, all marginals of a normal vector are normal. Not vice versa!

Corollary.
If $X \sim N_k(\mu, \Sigma)$ and A a $m \times k$ matrix, then $AX \sim N_m(A\mu, A\Sigma A^T)$.

Special property for normal vectors:

Lemma.
If $X \sim N_k(\mu, \Sigma)$ and Σ is diagonal, i.e. the X_i are uncorrelated, then the X_i are independent.

Multivariate CLT

Theorem.
Let $Y_1, Y_2, \ldots$ be i.i.d. random vectors in $\mathbb{R}^k$ with finite mean μ and covariance matrix Σ. Then

$$\sqrt{n}(\bar{Y}_n - \mu) \rightsquigarrow N_k(0, \Sigma)$$

as $n \to \infty$.

Sketch of proof.
We use the so-called Cramér-Wold argument. For all $a \in \mathbb{R}^k$, the ordinary CLT implies that

$$a^T \sqrt{n}(\bar{Y}_n - \mu) \rightsquigarrow N_k(0, a^T \Sigma a).$$

This implies the statement of the theorem. □

Quadratic forms

Chisquare distribution

Definition.
The chisquare distribution with k degrees of freedom is the distribution of $Z_1^2 + \ldots + Z_k^2$, for $Z_1, \ldots, Z_k$ independent, standard normal random variables. Notation: χ_k^2.

In other words: if $Z \sim N_k(0, I)$, then $\|Z\|^2 \sim \chi_k^2$.

Quadratic forms

Lemma.
If $X \sim N_k(0, \Sigma)$, then $\|X\|^2 \stackrel{\mathrm{d}}{=} \sum_{i=1}^k \lambda_i Z_i^2$, where $\lambda_1, \ldots, \lambda_k$ are the eigenvalues of Σ and $Z_1, \ldots, Z_k$ independent, standard normal random variables.

Sketch of proof.
Use that $\|X\|^2 = \|OX\|^2$ for O an orthogonal matrix of eigenvectors of Σ. $\square$