Multivariate random variables

Definition 1 *Function $X = (X_1, \dots, X_d) : \Omega \to \mathbb{R}^d$ is a* ***multivariate random variable****, if $\{\omega : X(\omega) \in B\} \in \mathcal{A}$ for all d dimensional Borel measurable sets. Equivalently, if all coordinates are random variables. The distribution function of X is $F_X : \mathbb{R}^d \to \mathbb{R}$, and $F_X(z) = P(X < z)$. The inequality in the argument means coordinate-wise inequalities ($z \in \mathbb{R}^d$), i.e., $X < z$ if and only if $X_i < z_i$ for $i = 1, \dots, d$. F_X is the* ***cumulative distribution function*** *of multivariate random variable X.*

Properties:

1. $0 \le F_X(z) \le 1$
2. $F_X(z)$ is monotonically increasing in each coordinate
3. $\lim F_X(z) = 1$, if $\forall i\ z_i \to \infty$
4. $\lim F_X(z) = 0$, if for at least one i $z_i \to -\infty$
5. $F_X(z)$ is left continuous in each coordinate

Definition 2 *Let (X, Y) be a 2-dimensional random variable with cumulative distribution function $F_{(X,Y)}$. $\lim\limits_{x\to\infty} F_{(X,Y)}(x, y) = F_Y(y)$ and $\lim\limits_{y\to\infty} F_{(X,Y)}(x, y) = F_X(x)$ are* ***marginal distributions****.*

			Y	
		0	1	2
	0	0	1/30	2/30
	1	1/30	2/30	3/30
X	2	2/30	3/30	4/30
	3	3/30	4/30	5/30

Definition 3 *If F is the integral function of $f : \mathbb{R}^n \to \mathbb{R}$, i.e., $F(z) = \int_{-\infty}^{z_1} \dots \int_{-\infty}^{z_d} f(t_1, \dots, t_d) dt_d \dots dt_1$, then F is absolutely continuous with density function f.*

Proposition 1 (densities of marginal distributions) *If (X, Y) is a 2-dimensional absolutely continuous random variable with density function $f(x, y)$, then the density function of random variable X is $g_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dy$, and similarly for Y.*

Proof: $\int\limits_{-\infty}^{z} \int\limits_{-\infty}^{\infty} f_{X,Y}(x, y)dydx = F_{X,Y}(z, \infty) = P(X < z)$. □

Proposition 2 *The coordinates of multivariate random variable X are independent if and only if $F_X(z) = F_{X_1}(z_1) \cdot \ldots \cdot F_{X_d}(z_d)$ for all $z \in \mathbb{R}^d$. If X is absolutely continuous, it is also equivalent to $f_X(z) = f_{X_1}(z_1) \cdot \ldots \cdot f_{X_d}(z_d)$ for all $z \in \mathbb{R}^d$ (density functions).*

Estimation methods

Continuing the discussion of maximum likelihood method, remark that usually the estimator is not unbiased, unfortunately. Moreover, in some cases we can not calculate using differentiation, e.g. if the sample is from the Uniform$(0, \vartheta)$ distribution, and our aim is to estimate ϑ.

Fortunately, if $T(x)$ is the maximum likelihood estimator of ϑ, then for some g function, $g(T(x))$ is the ML estimator of $g(\vartheta)$.

Proposition 3 *If the likelihood function meets certain regularity conditions, then the maximum likelihood estimator exists, asymptotically unbiased, asymptotically MVUE and asymptotically normally distributed.*

Definition 4 (method of moments) *Suppose that the problem is to estimate k parameters $\theta_1, \ldots, \theta_k$ characterising the distribution of the RV X. Suppose further that the first k moments of the distribution can be expressed as functions of θ's, i.e.*

$$\begin{aligned}
\mu_1 = EX &= g_1(\theta_1, \ldots, \theta_k) \\
\mu_2 = EX^2 &= g_2(\theta_1, \ldots, \theta_k) \\
&\vdots \\
\mu_k = EX^k &= g_k(\theta_1, \ldots, \theta_k)
\end{aligned}$$

Let $\hat{\mu}_j = \frac{1}{n}\sum_{i=1}^n X_i^j$ the j^{th} empirical moment of a sample $X_1, \ldots, X_n$. Then the estimates $\hat{\theta}_1, \ldots, \hat{\theta}_k$ are to be determined as the solution of

$$\begin{aligned}
\hat{\mu}_1 &= g_1(\hat{\theta}_1, \ldots, \hat{\theta}_k) \\
\hat{\mu}_2 &= g_2(\hat{\theta}_1, \ldots, \hat{\theta}_k) \\
&\vdots \\
\hat{\mu}_k &= g_k(\hat{\theta}_1, \ldots, \hat{\theta}_k)
\end{aligned}$$

Remark 1 *The MoM (Method of Moments) is consistent under very weak assumptions, but it is often biased.*

Example 1 *For $N(\mu, \sigma^2)$ one has the following relations:*

$$\begin{aligned}
&\mu_1 = \mu \\
&\mu_2 = \sigma^2 + \mu^2
\end{aligned}$$

Hence

$$\begin{aligned}
\hat{\mu} &= \frac{1}{n}\sum_{i=1}^n X_i \\
\hat{\sigma^2} &= \frac{1}{n}\sum_{i=1}^n X_i^2 - (\frac{1}{n}\sum_{i=1}^n X_i)^2
\end{aligned}$$

Definition 5 (confidence interval) *An interval, which contains the (unknown) parameter of the distribution with at least probability $1-\alpha$, where $\alpha > 0$ is a real number (usually 0.01 or 0.05). The confidence interval is defined by statistics of sample, i.e., $P_\vartheta(T_1(X) < \vartheta < T_2(X)) \geq 1-\alpha$.*

Example 2 *Suppose an n-element sample from normal distribution, where the expected value, m is unknown, but the variance, σ is known. In this case, the confidence interval for m is $\left[\overline{x} - \frac{z_{1-\frac{\alpha}{2}}\sigma}{\sqrt{n}}; \overline{x} + \frac{z_{1-\frac{\alpha}{2}}\sigma}{\sqrt{n}}\right]$. Remark that the probability that m falls to this interval exactly equals to $1-\alpha$.*

Remark 2 *In the fractional in Example 2, $z_{1-\frac{\alpha}{2}}$ is the $1-\frac{\alpha}{2}$ quantile of standard normal distribution. If $\Phi(x)$ denotes the distribution function of standard normal distribution, then $z_p = \Phi^{-1}(p)$, i.e., z_p is the solution of the equation $\Phi(z_p) = p$.*