近代统计方法

Modern Statistical Methods

师义民　许　勇　周丙常　编著

高等教育出版社
HIGHER EDUCATION PRESS

近代统计方法

JINDAI TONGJI FANGFA

Modern Statistical Methods

师义民　许　勇　周丙常　编著

内容提要

本书介绍了近代统计的基本概念、基本方法和基本理论，充分反映近代统计的新发展与应用，力求做到理论与实际相结合，叙述严谨，由浅入深。

全书共分九章，主要包括：概率论基础，抽样分布，参数估计，稳健统计初步，贝叶斯估计，假设检验，回归分析，多元分析初步，可靠性统计简介。书中配有大量的例题与习题，强调近代统计的思想、理论、方法及应用。

本书可作为高等学校工科类、管理学类以及数学类相关专业研究生的“近代统计方法”课程教材，亦可供工程技术人员参考使用。

图书在版编目（CIP）数据

近代统计方法/师义民，许勇，周丙常编著.—北京：高等教育出版社，2011.7

ISBN 978-7-04-032291-0

Ⅰ.①近… Ⅱ.①师… ②许… ③周… Ⅲ.①统计方法-研究生-教材 Ⅳ.①C81

中国版本图书馆CIP数据核字(2011)第112483号

策划编辑 张长虹　责任编辑 田 玲　封面设计 李卫青　版式设计 马敬茹
插图绘制 杜晓丹　责任校对 金 辉　责任印制 张福涛

出版发行	高等教育出版社	咨询电话	400-810-0598
社　　址	北京市西城区德外大街4号	网　　址	http://www.hep.edu.cn
邮政编码	100120		http://www.hep.com.cn
印　　刷	北京七色印务有限公司	网上订购	http://www.landraco.com
开　　本	787mm×960mm 1/16		http://www.landraco.com.cn
印　　张	19.5	版　　次	2011年7月第1版
字　　数	360千字	印　　次	2011年7月第1次印刷
购书热线	010-58581118	定　　价	36.40元

物 料 号 32291-00

前　　言

为适应研究生数理统计课程教学改革的需要，我们编写了这本《近代统计方法》教科书。本书的前身是“近代统计方法”讲义，该讲义曾在我校研究生教学中使用6届。这次编写是在保留原讲义内容优点的基础上，补充和修改而成的。本书着力于近代统计的基本概念、基本方法和基本理论，充分反映近代统计的新发展与应用，力求做到理论与实际相结合，为读者进入理论研究领域和实际应用领域打下扎实的基础。

全书共分九章，依次为概率论基础、抽样分布、参数估计、稳健统计初步、贝叶斯估计、假设检验、回归分析、多元分析初步、可靠性统计简介。书中各章配有适量的习题。本书对近代统计的基本概念与核心内容，力求做到循序渐进、由浅入深、叙述严谨，而对应用方法部分，通过典型实例来介绍，培养学生应用近代统计的理论与方法解决实际问题的能力。

本书可作为高等学校工科类、管理学类以及数学类相关专业研究生的“近代统计方法”课程教材，亦可供工程技术人员参考使用。

本书的第一章至第四章由许勇编写，第五章、第八章和第九章由师义民编写，第六章和第七章由周丙常编写，全书由师义民统稿、定稿。

西北工业大学教务处、应用数学系概率统计教研室对本书的编写和出版给予了大力的支持和帮助，研究生顾昕、孙玉东、谭伟、王琳、毛松等为本书的部分章节打印、校对做了很多工作，在此一并致谢。

由于编者水平有限，错谬之处难免，恳请读者批评指正。

编者

2011年3月

目　录

第一章　概率论基础

§1.1　概率空间

1.1.1　概率的公理化定义

概率论是研究随机现象及其统计规律的学科.为了对随机现象及其统计规律性进行深入研究,就必须在一定的条件下对它进行多次观察.若把一次观察视为一次试验,观测到的结果就是试验结果.概率论中把满足下列条件的试验称为随机试验,通常我们用记号 E 表示随机试验.

(1) 试验可以在相同条件下重复进行;

(2) 每次试验的可能结果不止一个,但试验的所有可能结果事先是明确的;

(3) 在试验结果揭晓前,无法预言会出现哪一个可能结果.

若无特别声明,本书以后所指的试验均指随机试验.例如

(1) 在一定的条件下进行射击练习,考虑中靶的环数;

(2) 掷一颗均匀的骰子,考虑出现的点数;

(3) 记录某汽车站某时段内候车的人数.

定义 1.1　对于随机试验 E,我们把每一个可能出现的结果称为样本点,把某些样本点构成的集合称为随机事件,简称事件.把单个样本点构成的集合称为基本事件,把所有样本点构成的集合称为必然事件或样本空间,记为 Ω.

通常我们用 A,B,C 等大写英文字母表示随机事件.我们规定不含任何元素的空集也为事件,称为不可能事件,用$\varnothing$表示.

1.1.2　概率空间

定义 1.2　样本空间 Ω 的子集集合 $\mathscr{F}$,如果满足如下性质:

(1) $\Omega\in\mathscr{F}$;

(2) 如果 $A\in\mathscr{F}$,则 $\bar{A}\in\mathscr{F}$;

(3) 如果 $A_i\in\mathscr{F},i=1,2,3\cdots$,则 $\bigcup\limits_{i=1}^{+\infty}A_i\in\mathscr{F}$,

则称$\mathscr{F}$为σ－域. 其中$\bar{A}$为A在Ω中的补集：$\bar{A}=\Omega-A$.

例 1.1 $\mathscr{F}=\{\varnothing,\Omega\}$. 不难验证$\mathscr{F}$是一个$\sigma$－域.

例 1.2 $\mathscr{F}=\{\varnothing,A,\bar{A},\Omega\}$. 可以验证$\mathscr{F}$也是一个$\sigma$－域.

定理 1.1 设$\mathscr{F}$为一个σ－域，则

(1) $\varnothing\in\mathscr{F}$;

(2) 如果$A,B\in\mathscr{F}$，则$A\cup B,AB,A-B\in\mathscr{F}$;

(3) 如果$A_i\in\mathscr{F},i=1,2,3,\cdots,n$，则$\bigcup\limits_{i=1}^{n}A_i\in\mathscr{F}$.

证明略.

定义 1.3（概率的公理化定义） 设$\mathscr{F}$是一个σ－域，在$\mathscr{F}$上定义实值集函数P，对每一事件$A\in\mathscr{F}$，函数值为$P(A)$，如果它满足如下三个条件：

(1) 非负性：对任意的$A\in\mathscr{F}$，有$P(A)\geqslant0$;

(2) 规范性：$P(\Omega)=1$;

(3) 可列可加性：若$A_i\in\mathscr{F},i=1,2,3\cdots$，且$A_iA_j=\varnothing(i\neq j)$，即所有随机事件两两互不相容，则有$P\left(\bigcup\limits_{i=1}^{+\infty}A_i\right)=\sum\limits_{i=1}^{+\infty}P(A_i)$,

则称P为$\mathscr{F}$上的概率，称$P(A)$为事件A的概率.

通常，我们将$\Omega,\mathscr{F},P$联系在一起，称三元整体$(\Omega,\mathscr{F},P)$为概率空间. 在实际问题中，$\Omega,\mathscr{F},P$都认为是预先给定的，并以此作为出发点讨论种种问题. 至于实际问题中，如何选定Ω，怎样构造$\mathscr{F}$，怎样给定P，则要视具体情况而定.

例 1.3 掷硬币试验：

$$\Omega=\{(\omega_1)=\text{“正面向上”},(\omega_2)=\text{“反面向上”}\},$$

$$\mathscr{F}=\{(\omega_1),(\omega_2),\Omega,\varnothing\},$$

$$P=\left\{P(\omega_1)=P(\omega_2)=\frac{1}{2},P(\Omega)=1,P(\varnothing)=0\right\},$$

$(\Omega,\mathscr{F},P)$就是掷硬币试验的概率空间.

一般说来，若样本空间Ω内的元素只有有限个，譬如n个，$\mathscr{F}$是由Ω的所有子集所组成的σ－域，则$\mathscr{F}$内的事件总个数是2^n.

§1.2 多维随机变量及其分布

1.2.1 多维随机变量及其分布

定义 1.4 若随机变量$X_1,X_2,\cdots,X_n$定义在同一概率空间$(\Omega,\mathscr{F},P)$上，则

称 $X=(X_1,X_2,\cdots,X_n)$ 构成一个 n 维(或 n 元)随机向量或随机变量.

定义 1.5 设 $x_1,x_2,\cdots,x_n$ 是 n 个实数,称 n 元函数

$$F(x_1,x_2,\cdots,x_n)=P(X_1\leqslant x_1,X_2\leqslant x_2,\cdots,X_n\leqslant x_n)$$

为随机向量 $X=(X_1,X_2,\cdots,X_n)$ 的联合分布函数.

随机向量的分布函数也有离散型与连续型的区别. 在离散型场合,概率分布集中在有限个或可列个点上. 在连续型场合,存在着非负可积函数 $f(x_1,x_2,\cdots,x_n)$,使

$$F(x_1,x_2,\cdots,x_n)=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n}f(y_1,\cdots,y_n)\mathrm{d}y_1\cdots\mathrm{d}y_n,$$

其中 $f(x_1,x_2,\cdots,x_n)$ 称为联合密度函数,满足条件 $f(x_1,x_2,\cdots,x_n)\geqslant 0$,且

$$\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}f(x_1,x_2,\cdots,x_n)\mathrm{d}x_1\cdots\mathrm{d}x_n=1.$$

例 1.4 现做 n 次独立重复试验,每次试验的结果有 $A_1,A_2,\cdots,A_m$, $P(A_i)=p_i$, $i=1,2,\cdots,m$,且 $p_i\geqslant 0$, $\sum\limits_{i=1}^{m}p_i=1$. 记 X_i 表示在 n 次试验中事件 A_i 出现的次数,则 $(X_1,X_2,\cdots,X_m)$ 是一 m 维随机向量,其分布律为

$$P(X_1=n_1,X_2=n_2,\cdots,X_m=n_m)=\frac{n!}{n_1!n_2!\cdots n_m!}p_1^{n_1}p_2^{n_2}\cdots p_m^{n_m},\quad(1.1)$$

这里 $n_i\geqslant 0$, $\sum\limits_{i=1}^{m}n_i=n$. 公式(1.1)称为多项分布,记为 $M(n,p_1,\cdots,p_m)$. 它是 $(p_1+p_2+\cdots+p_m)^n$ 的展开式的一般项,且

$$\sum_{n_1+n_2+\cdots+n_m=n}\frac{n!}{n_1!n_2!\cdots n_m!}p_1^{n_1}p_2^{n_2}\cdots p_m^{n_m}=1.$$

$m=2$ 时,式(1.1)即为二项分布. 显然多项分布是二项分布的推广.

例 1.5 若随机变量 $X=(X_1,X_2,\cdots,X_n)$ 的概率密度函数为

$$f(x_1,x_2,\cdots,x_n)=\begin{cases}\dfrac{1}{S}, & (x_1,x_2,\cdots,x_n)\in G,\\ 0, & (x_1,x_2,\cdots,x_n)\notin G,\end{cases}\quad(1.2)$$

其中 G 是 $\mathbf{R}^n$ 中的有限区域, G 的体积 $S>0$, 则称随机变量 X 服从 G 上的均匀分布.

例 1.6 设 $\Sigma=(\sigma_{ij})$ 是 n 阶正定矩阵, $|\Sigma|$ 表示 Σ 的行列式的值, $x=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 为任意实值列向量,则由密度函数

$$f(x_1,x_2,\cdots,x_n)=\frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}(x-\mu)^{\mathrm{T}}\Sigma^{-1}(x-\mu)\right\}\quad(1.3)$$

定义的分布称为 n 维(或 n 元)正态分布,简记为 $N(\mu,\Sigma)$.

1.2.2 边缘分布

设 $F(x_1,x_2,\cdots,x_n)$ 是 n 元分布函数,任意保留 k 个 $x_i(1\leqslant k\leqslant n)$,例如 $x_1,x_2,\cdots,x_k$,而令其他的 x_j 趋向于 $+\infty$,即

$$F(x_1,x_2,\cdots,x_k,+\infty,\cdots,+\infty)=\lim_{\substack{x_{k+1}\to+\infty,\cdots,\\ x_n\to+\infty}}F(x_1,x_2,\cdots,x_n).\tag{1.4}$$

容易看出 $F(x_1,x_2,\cdots,x_k,+\infty,\cdots,+\infty)$ 是一个 k 元分布函数,称它为 $F(x_1,x_2,\cdots,x_n)$ 的 k 维(或 k 元)边缘分布函数. 由于自 $x_1,x_2,\cdots,x_n$ 中挑选 k 个 x_i 的方法共有 C_n^k 种,故共有 C_n^k 个 k 维边缘分布函数.

如果 $F(x_1,x_2,\cdots,x_n)$ 是连续型的,相应的密度函数为 $f(x_1,x_2,\cdots,x_n)$,则

$$\begin{aligned}&F(x_1,x_2,\cdots,x_k,+\infty,\cdots,+\infty)\\&=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_k}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}f(y_1,\cdots,y_n)\,\mathrm{d}y_n\cdots\mathrm{d}y_{k+1}\mathrm{d}y_k\cdots\mathrm{d}y_1.\end{aligned}$$

可见 $F(x_1,x_2,\cdots,x_k,+\infty,\cdots,+\infty)$ 也是连续型的,相应的密度函数为

$$f_{1,2,\cdots,k}(x_1,x_2,\cdots,x_k)=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}f(x_1,\cdots,x_n)\,\mathrm{d}x_n\cdots\mathrm{d}x_{k+1}.$$

如果 $F(x_1,x_2,\cdots,x_n)$ 是离散型的,那么 $F(x_1,x_2,\cdots,x_k,+\infty,\cdots,+\infty)$ 也是离散型的. 其边缘分布律可仿上式求得,只需把积分号改为求和号.

例 1.7 设二维随机变量 (X,Y) 的联合分布律为

$$\begin{pmatrix}(1,1) & (2,1) & (2,2) & (3,4)\\ \dfrac{1}{3} & \dfrac{3}{24} & \dfrac{1}{24} & \dfrac{1}{2}\end{pmatrix}$$

求 (X,Y) 关于 X 和关于 Y 的边缘分布律.

解 X 的可能取值构成点集 $(1,2,3)$,可得 X 的边缘分布律为

$$\begin{pmatrix}1 & 2 & 3\\ \dfrac{1}{3} & \dfrac{3}{24}+\dfrac{1}{24} & \dfrac{1}{2}\end{pmatrix}=\begin{pmatrix}1 & 2 & 3\\ \dfrac{1}{3} & \dfrac{1}{6} & \dfrac{1}{2}\end{pmatrix}$$

类似地,Y 的可能取值构成点集 $(1,2,4)$,Y 的边缘分布律为

$$\begin{pmatrix}1 & 2 & 4\\ \dfrac{1}{3}+\dfrac{3}{24} & \dfrac{1}{24} & \dfrac{1}{2}\end{pmatrix}=\begin{pmatrix}1 & 2 & 4\\ \dfrac{11}{24} & \dfrac{1}{24} & \dfrac{1}{2}\end{pmatrix}$$

以上我们对保留 $x_1,x_2,\cdots,x_k$ 的情况进行了讨论,如果保留的是 $(x_{i_1},x_{i_2},\cdots,x_{i_k})$,讨论完全类似.

边缘分布函数由分布函数 $F(x_1,x_2,\cdots,x_n)$ 唯一决定;但反之不然,也就是说,不同的分布函数可以有完全相同的边缘分布函数.

例 1.8 设有两个二元分布函数 $F(x,y)$ 和 $G(x,y)$,密度函数分别为

$$f(x,y)=\begin{cases}x+y, & \text{若 } 0\leqslant x\leqslant 1,\ 0\leqslant y\leqslant 1,\\ 0, & \text{其他};\end{cases}$$

$$g(x,y)=\begin{cases}\left(\dfrac{1}{2}+x\right)\left(\dfrac{1}{2}+y\right), & \text{若 } 0\leqslant x\leqslant 1,\ 0\leqslant y\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

容易看出,方程$\left(\frac{1}{2}+x\right)\left(\frac{1}{2}+y\right)=x+y$只有根$x=\frac{1}{2}$,$y=\frac{1}{2}$,因而在正方形$\{(x,y)\in\mathbf{R}^2:0\leqslant x\leqslant 1,0\leqslant y\leqslant 1\}$中,只有在两直线$x=\frac{1}{2}$及$y=\frac{1}{2}$上,$f(x,y)=g(x,y)$,可见$F(x,y)$与$G(x,y)$不恒等.然而它们的边缘密度函数相等.事实上,

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\,\mathrm{d}y=\begin{cases}\displaystyle\int_0^1(x+y)\,\mathrm{d}y=x+\frac{1}{2}, & 0\leqslant x\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

$$g_X(x)=\int_{-\infty}^{+\infty}g(x,y)\,\mathrm{d}y=\begin{cases}\displaystyle\int_0^1\left(\frac{1}{2}+x\right)\left(\frac{1}{2}+y\right)\mathrm{d}y=x+\frac{1}{2}, & 0\leqslant x\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

所以$f_X(x)=g_X(x)$.

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\,\mathrm{d}x=\begin{cases}\displaystyle\int_0^1(x+y)\,\mathrm{d}x=y+\frac{1}{2}, & 0\leqslant y\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

$$g_Y(y)=\int_{-\infty}^{+\infty}g(x,y)\,\mathrm{d}x=\begin{cases}\displaystyle\int_0^1\left(\frac{1}{2}+x\right)\left(\frac{1}{2}+y\right)\mathrm{d}x=y+\frac{1}{2}, & 0\leqslant y\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

因此$f_Y(y)=g_Y(y)$.

例 1.9　设随机变量(X,Y)的密度函数为

$$f(x,y)=\begin{cases}x\mathrm{e}^{-y}, & 0<x<y<+\infty;\\ 0, & \text{其他}.\end{cases}$$

求边缘密度函数$f_X(x)$,$f_Y(y)$.

解　$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\,\mathrm{d}y=\begin{cases}\displaystyle\int_x^{+\infty}x\mathrm{e}^{-y}\mathrm{d}y, & x>0\\ 0, & x\leqslant 0\end{cases}=\begin{cases}x\mathrm{e}^{-x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\,\mathrm{d}x=\begin{cases}\displaystyle\int_0^y x\mathrm{e}^{-y}\mathrm{d}x, & y>0\\ 0, & y\leqslant 0\end{cases}=\begin{cases}\dfrac{1}{2}y^2\mathrm{e}^{-y}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

1.2.3 条件分布

设 X,Y 为两个随机变量,对任意 $C\in B_1$(B_1 为某个可测集),如果 $P(X\in C)>0$,可考虑 $y\in\mathbf{R}$ 的函数

$$P(Y\leqslant y\mid X\in C)=\frac{P(Y\leqslant y,X\in C)}{P(X\in C)}. \tag{1.5}$$

容易看出 $P(Y\leqslant y\mid X\in C)$是一维分布函数,称它为在条件 $X\in C$ 下 Y 的条件分布函数.

设 X,Y 是离散的,(X,Y) 的联合分布律为 $P(X=x_i,Y=y_j)=p_{ij},i,j=1,2,\cdots$,于是 $p_{i\cdot}=\sum\limits_j p_{ij},i=1,2,\cdots,p_{\cdot j}=\sum\limits_i p_{ij},j=1,2,\cdots$; 从而

$$P(Y=y_j\mid X=x_i)=\frac{p_{ij}}{p_{i\cdot}}=\frac{p_{ij}}{\sum\limits_j p_{ij}}, \tag{1.6}$$

$$P(Y\leqslant y\mid X=x_i)=\frac{\sum\limits_{j:y_j\leqslant y}p_{ij}}{\sum\limits_j p_{ij}}. \tag{1.7}$$

设 X,Y 是连续型随机变量,联合密度函数为$f(x,y)$,如果在定点 x,

$$\int_{-\infty}^{+\infty}f(x,y)\,\mathrm{d}y>0, \tag{1.8}$$

定义

$$P(Y\leqslant y\mid X=x)=\frac{\int_{-\infty}^{y}f(x,z)\,\mathrm{d}z}{\int_{-\infty}^{+\infty}f(x,z)\,\mathrm{d}z} \tag{1.9}$$

为给定 $X=x$ 条件下 Y 的条件分布函数. 称函数

$$f(y\mid x)=\frac{f(x,y)}{\int_{-\infty}^{+\infty}f(x,z)\,\mathrm{d}z} \tag{1.10}$$

为在条件 $X=x$ 下 Y 的密度函数. 注意$f_X(x)=\int_{-\infty}^{+\infty}f(x,z)\,\mathrm{d}z$ 是(X,Y) 关于 X 的边缘密度函数,显然式(1.9) 可写为

$$P(Y\leqslant y\mid X=x)=\int_{-\infty}^{y}f(z\mid x)\,\mathrm{d}z. \tag{1.11}$$

同理可得

$$f(x\mid y)=\frac{f(x,y)}{\int_{-\infty}^{+\infty}f(z,y)\,\mathrm{d}z}=\frac{f(x,y)}{f_Y(y)}. \tag{1.12}$$

把式(1.10)和式(1.12)变形得

$$f(x,y)=f_X(x)f(y\mid x), \tag{1.13}$$

$$f(x,y)=f_Y(y)f(x\mid y). \tag{1.14}$$

这就表明用一个变量的分布与在这个变量的给定情况下另一个变量的条件分布,可求出联合分布. 比较式(1.12)与式(1.14)得

$$f(x\mid y)=\frac{f_X(x)f(y\mid x)}{f_Y(y)}=\frac{f_X(x)f(y\mid x)}{\int_{-\infty}^{+\infty}f(y\mid x)f_X(x)\,\mathrm{d}x}. \tag{1.15}$$

这就是贝叶斯公式的密度函数形式.

例 1.10　设 X 服从均匀分布 $U(0,1)$($U(a,b)$表示区间(a,b)上的均匀分布),x 是一个观测值,又设在 $X=x$ 条件下 Y 的条件分布为均匀分布 $U(x,1)$.

(1) 求(X,Y)的联合分布;

(2) 求 Y 的边缘分布及 $P(Y>0.5)$.

解　(1) X 的密度函数为

$$f(x)=\begin{cases}1, & 0<x<1,\\ 0, & \text{其他}.\end{cases}$$

在给定 $X=x$ 条件下,Y 的条件密度为

$$f(y\mid x)=\begin{cases}\dfrac{1}{1-x}, & 0<x<y<1,\\ 0, & \text{其他}.\end{cases}$$

由此得(X,Y)的联合密度为

$$f(x,y)=f_X(x)f(y\mid x)=\begin{cases}\dfrac{1}{1-x}, & 0<x<y<1,\\ 0, & \text{其他}.\end{cases}$$

(2) Y 的边缘密度为

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\,\mathrm{d}x=\int_0^y\frac{1}{1-x}\mathrm{d}x=-\ln(1-y),\quad 0<y<1.$$

$$P(Y>0.5)=-\int_{0.5}^{1}\ln(1-y)\,\mathrm{d}y=-\int_0^{0.5}\ln u\,\mathrm{d}u$$

$$=-(u\ln u-u)\Big|_0^{0.5}=0.5\ln 2+0.5=0.8466.$$

1.2.4　随机变量的独立性

定义 1.6　设 $X_1,X_2,\cdots,X_n$ 为 n 个随机变量,若对于任意的实数 $x_1,x_2,\cdots,x_n$ 成立

$$P(X_1\leqslant x_1,\cdots,X_n\leqslant x_n)=P(X_1\leqslant x_1)\cdots P(X_n\leqslant x_n), \tag{1.16}$$

则称 $X_1,X_2,\cdots,X_n$ 是相互独立的.

若 X_i 的分布函数为 $F_i(x_i)$,$i=1,2,\cdots,n$,$X_1,X_2,\cdots,X_n$ 的联合分布函数为

$F(x_1,x_2,\cdots,x_n)$,则式(1.16)等价于对任何一组可能取的值 $x_1,x_2,\cdots,x_n$ 都有

$$F(x_1,\cdots,x_n) = F_1(x_1)\cdots F_n(x_n). \tag{1.17}$$

在这种场合,由随机变量的边缘分布函数可唯一地决定联合分布函数.

对于离散型随机变量,式(1.16)等价于对任何一组可能取的值 $(x_1,x_2,\cdots,x_n)$ 都有

$$P(X_1 = x_1, X_2 = x_2,\cdots,X_n = x_n) = P(X_1 = x_1)P(X_2 = x_2)\cdots P(X_n = x_n). \tag{1.18}$$

对于连续型随机变量,$X_1,X_2,\cdots,X_n$ 相互独立的充分必要条件是

$$f(x_1,x_2,\cdots,x_n) = f_1(x_1)f_2(x_2)\cdots f_n(x_n). \tag{1.19}$$

这里 $f(x_1,x_2,\cdots,x_n)$ 是联合分布密度函数,而 $f_i(x_i)$ 是各随机变量的密度函数.

例 1.11　设 (X,Y) 服从二维正态分布 $N(\mu,\Sigma)$,其中

$$\mu = \begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix},\ \Sigma = \begin{pmatrix}\sigma_1^2 & r\sigma_1\sigma_2\\ r\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix},$$

证明 X 与 Y 相互独立的充分必要条件是 $r=0$.

证　我们先来求 X 和 Y 的边缘密度函数. 令 $\dfrac{x-\mu_1}{\sigma_1}=u$, $\dfrac{y-\mu_2}{\sigma_2}=v$, 则

$$\begin{aligned}
f_1(x) &= \int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}y\\
&= \frac{1}{2\pi\sigma_1\sqrt{1-r^2}}\int_{-\infty}^{+\infty}\exp\left\{-\frac{1}{2(1-r^2)}[u^2-2ruv+v^2]\right\}\mathrm{d}v\\
&= \frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-\frac{u^2}{2}}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi(1-r^2)}}\exp\left\{-\frac{r^2u^2-2ruv+v^2}{2(1-r^2)}\right\}\mathrm{d}v\\
&= \frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-\frac{u^2}{2}}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi(1-r^2)}}\mathrm{e}^{-\frac{(v-ru)^2}{2(1-r^2)}}\mathrm{d}v\\
&= \frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-\frac{u^2}{2}} = \frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}},
\end{aligned}$$

即 $f_1(x)$ 是正态分布 $N(\mu_1,\sigma_1^2)$ 的密度函数,同理可得 $f_2(y)=\dfrac{1}{\sqrt{2\pi}\sigma_2}\mathrm{e}^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$.

如果 X 与 Y 相互独立,则 $f(x,y)=f_1(x)f_2(y)$,从而在 $x=\mu_1, y=\mu_2$ 处,上式成立,即

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} = \frac{1}{2\pi\sigma_1\sigma_2},$$

亦即 $\sqrt{1-r^2}=1$,所以 $r=0$. 反之若 $r=0$,则

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2}\exp\left\{-\frac{1}{2}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$$