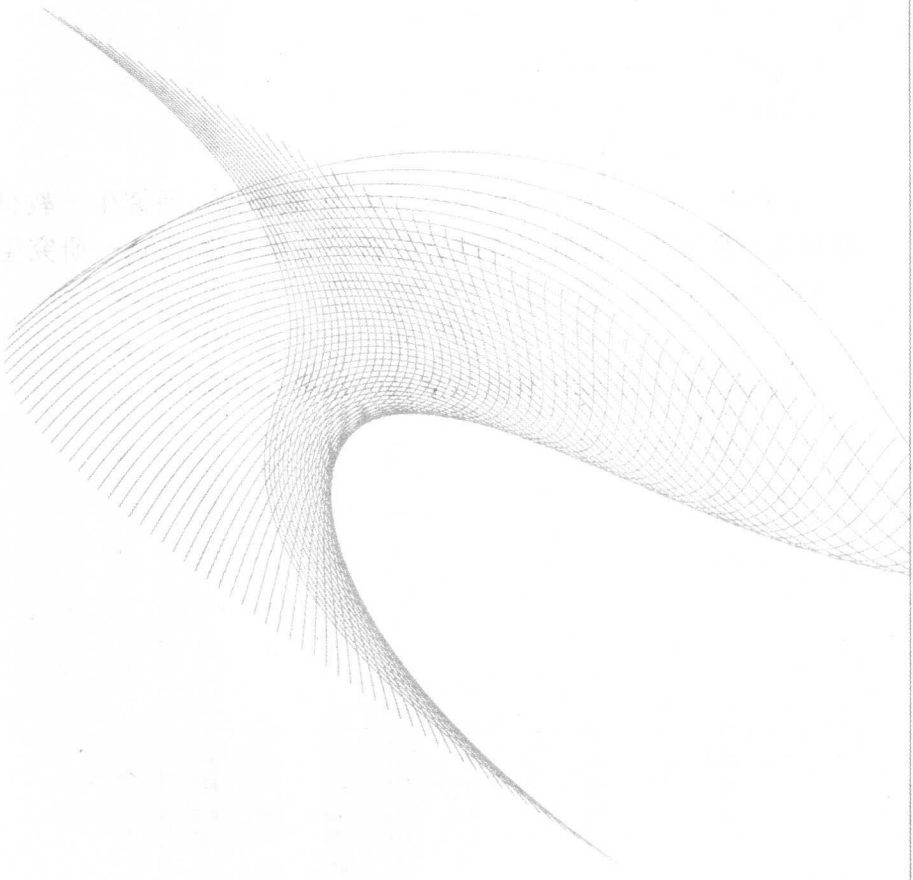
硕士研究生公共课教材

概率统计及SAS应用

余家林　肖枝洪　编著

WUHAN UNIVERSITY PRESS
武汉大学出版社

硕士研究生公共课教材



概率统计及SAS应用

余家林　肖枝洪　编著

WUHAN UNIVERSITY PRESS
武汉大学出版社

图书在版编目(CIP)数据

概率统计及SAS应用/余家林,肖枝洪编著.—武汉:武汉大学出版社,2007.7

硕士研究生公共课教材

ISBN 978-7-307-05608-4

Ⅰ.概… Ⅱ.①余… ②肖… Ⅲ.①概率论—研究生—教材 ②数理统计—研究生—教材 ③统计分析—应用软件,SAS—研究生—教材

Ⅳ.O21 C812

中国版本图书馆CIP数据核字(2007)第069410号

责任编辑:杨 华 责任校对:黄添生 版式设计:詹锦玲

出版发行:**武汉大学出版社** (430072 武昌 珞珈山)

(电子邮件:wdp4@whu.edu.cn 网址:www.wdp.com.cn)

印刷:湖北恒泰印务有限公司

开本:720×1000 1/16 印张:18.375 字数:316千字

版次:2007年7月第1版 2007年7月第1次印刷

ISBN 978-7-307-05608-4/O·362 定价:22.00元

内 容 提 要

统计学以概率论为理论基础，根据试验与观测所得到的数据资料，对研究对象的特征及内在规律进行估计与推断，具有十分广泛的应用。本教材包括概率论基础、数理统计的基本概念、点估计与区间估计、参数假设检验、正态性检验、试验设计与方差分析、回归分析与协方差分析、非参数检验、随机过程及 SAS 的应用等内容。本书既可作为非数学专业硕士研究生“概率统计”课程的教材，也可作为科技工作者的参考文献。

前　言

“概率统计”是非数学专业硕士研究生教学计划中普遍开设的一门公共基础课，各学校各专业讲授的内容大体一致。随着硕士研究生入学水平与课题研究水平的提高，亟需一本相适应的教材，既能加强理论基础，帮助研究生熟悉统计学原理，又能介绍近代流行的统计分析软件，使研究生在处理试验数据的过程中摆脱复杂计算的困扰。

由我们合编的《概率统计及 SAS 应用》是近几年来硕士研究生优质课程立项研究的一项成果。作为非数学专业硕士研究生的教材，编入了概率论基础、数理统计的基本概念、点估计与区间估计、参数假设检验，正态性检验、试验设计与方差分析、回归分析与协方差分析、非参数检验、随机过程及 SAS 的应用等内容，讲课及上机实习可控制在 72 课时以内。

在编写中，我们特别注意说明统计方法的实际背景，详细讲述用统计方法解决实际问题的思路，对于应用 Statistical Analysis System(简称为 SAS)所得到统计分析结果，则尽可能地与实际计算步骤一一对照，使初学者能够知其所以然。考虑到专业与课时设置的不同，本教材力求简明扼要、重点突出，通俗易懂、便于自学，例题与习题都在常识的范围之内。

本教材中第一章、第二章、第五章由余家林编写，第三章、第四章、第六章由肖枝洪编写。本教材的出版得到华中农业大学研究生处及武汉大学出版社的大力支持，在此表示衷心的感谢。由于编者的水平所限，编写不妥之处难以避免，敬请读者和使用本教材的同行学友批评指正。

编　者

2007 年 2 月 6 日

目　录

第一章　概率论基础 …… (1)
1.1　随机变量的分布与相互独立 …… (1)
1.2　随机变量的数字特征 …… (13)
1.3　多项分布与多维正态分布 …… (23)
1.4　连续型随机变量的变换及变换后的分布 …… (40)
1.5　卡方分布、t 分布及 F 分布 …… (50)

第二章　统计学导论 …… (57)
2.1　总体与样本 …… (57)
2.2　常用的统计量及其分布 …… (72)
2.3　总体分布参数的估计 …… (82)
2.4　总体分布参数的假设检验 …… (97)
2.5　正态性检验 …… (109)

第三章　试验设计与方差分析 …… (116)
3.1　单因素试验的方差分析 …… (116)
3.2　多重比较与数据转换 …… (129)
3.3　双因素试验的方差分析(一) …… (142)
3.4　双因素试验的方差分析(二) …… (148)
3.5　二级系统分组试验的方差分析 …… (153)
3.6　多因素试验的方差分析 …… (159)

第四章　回归分析与协方差分析 …… (173)
4.1　一元线性回归 …… (173)
4.2　一元非线性回归 …… (186)
4.3　统计控制与协方差分析(一) …… (191)

4.4 统计控制与协方差分析(二) …… (199)

第五章 非参数检验 …… (214)
5.1 总体分布的卡方检验 …… (214)
5.2 二项分布总体率的估计与假设检验 …… (221)
5.3 两组样本数据的检验 …… (230)
5.4 多组样本数据的检验 …… (242)
5.5 相关性指标与检验 …… (251)

第六章 Markov 链 …… (262)
6.1 随机过程的概念 …… (262)
6.2 Markov 链 …… (263)
6.3 一步转移矩阵的估计 …… (265)
6.4 Markov 链的状态分类 …… (266)
6.5 Markov 链的极限定理 …… (268)

附录一 标准正态分布的分布函数值表 …… (273)
附录二 t 分布的分位数表 …… (274)
附录三 卡方分布的分位数表 …… (275)
附录四 F 分布的分位数表 …… (276)
附录五 二项分布表 …… (278)
附录六 二项分布参数 p 的置信区间表 …… (281)
参考文献 …… (286)

第一章　概率论基础

概率论是统计学的基础. 本章先讲述随机变量的分布与相互独立，随机变量的数字特征，多项分布与多维正态分布，连续型随机变量的变换及变换后的分布，然后讲述卡方分布、t分布及F分布，同时介绍用SAS计算标准正态分布分布函数的值，计算标准正态分布、卡方分布、t分布及F分布的分位数. 在讲述正态随机变量的非奇线性变换与标准正态随机变量的正交变换后，将导出一系列与统计方法有关的重要结论. 通过本章的学习，理解这两个变换的实质，熟悉它们的应用.

1.1　随机变量的分布与相互独立

1.1.1　分布函数及边缘分布函数

首先回顾一下一维和二维随机变量的分布函数的定义.

若X为一维随机变量，x为任意实数，随机事件$\{X \leqslant x\}$出现的概率为$P\{X \leqslant x\}$，则称$P\{X \leqslant x\}$为X的**分布函数**. 它是关于x的一元函数，记作

$$F(x) = P\{X \leqslant x\}.$$

由定义可知，$0 \leqslant F(x) \leqslant 1$，

$$F(-\infty) = 0, \quad F(+\infty) = 1.$$

若(X,Y)为二维随机变量，x与y为任意实数，随机事件$\{X \leqslant x, Y \leqslant y\}$出现的概率为$P\{X \leqslant x, Y \leqslant y\}$，则称$P\{X \leqslant x, Y \leqslant y\}$为$(X,Y)$的**分布函数**. 它是关于$x$与$y$的二元函数，记作

$$F(x,y) = P\{X \leqslant x, Y \leqslant y\}.$$

由定义可知，$0 \leqslant F(x,y) \leqslant 1$，

$$F(-\infty,y) = 0, \quad F(x,-\infty) = 0, \quad F(+\infty,+\infty) = 1.$$

若$F_X(x), F_Y(y)$分别是一维随机变量X与Y的分布函数，则

$$F_X(x) = P\{X \leqslant x\} = P\{X \leqslant x, Y < +\infty\} = F(x, +\infty),$$

$$F_Y(y) = P\{Y \leqslant y\} = P\{X < +\infty, Y \leqslant y\} = F(+\infty, y),$$

称 $F_X(x)$ 为(X,Y) **关于 X 的边缘分布函数**，称 $F_Y(y)$ 为(X,Y) **关于 Y 的边缘分布函数**，而 $F(x,y)$ 又称为 X **与 Y 的联合分布函数**.

类似地，可以定义多维随机变量$(X_1, X_2, \cdots, X_n)$的分布函数及边缘分布函数.

若$(X_1, X_2, \cdots, X_n)$为 n 维随机变量，$x_1, x_2, \cdots, x_n$ 为任意实数，则称随机事件$\{X_1 \leqslant x_1, X_2 \leqslant x_2, \cdots, X_n \leqslant x_n\}$ 出现的概率

$$P\{X_1 \leqslant x_1, X_2 \leqslant x_2, \cdots, X_n \leqslant x_n\}$$

为$(X_1, X_2, \cdots, X_n)$ **的分布函数**. 它是关于 $x_1, x_2, \cdots, x_n$ 的 n 元函数，记作

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

由定义可知，$0 \leqslant F(x_1, x_2, \cdots, x_n) \leqslant 1$，

$$F(-\infty, x_2, \cdots, x_n) = 0,$$

$$F(x_1, -\infty, x_3, \cdots, x_n) = 0,$$

$$\cdots,$$

$$F(x_1, x_2, \cdots, x_{n-1}, -\infty) = 0,$$

$$F(+\infty, +\infty, \cdots, +\infty) = 1.$$

为表达方便起见，多维随机变量又称为**随机向量**. 若记

$$\boldsymbol{X} = (X_1, X_2, \cdots, X_n)', \quad \boldsymbol{x} = (x_1, x_2, \cdots, x_n)',$$

则$(X_1, X_2, \cdots, X_n)$ 的分布函数又可记作 $F(\boldsymbol{x})$.

从 $X_1, X_2, \cdots, X_n$ 中任意确定 r 个随机变量，不妨假设是 $X_1, X_2, \cdots, X_r$，若 $F_{X_1, X_2, \cdots, X_r}(x_1, x_2, \cdots, x_r)$ 是 r 维随机变量$(X_1, X_2, \cdots, X_r)$ 的分布函数，则

$$\begin{aligned} & F_{X_1, X_2, \cdots, X_r}(x_1, x_2, \cdots, x_r) \\ &= P\{X_1 \leqslant x_1, X_2 \leqslant x_2, \cdots, X_r \leqslant x_r\} \\ &= P\{X_1 \leqslant x_1, X_2 \leqslant x_2, \cdots, X_r \leqslant x_r, X_{r+1} < +\infty, \cdots, X_n < +\infty\} \\ &= F(x_1, x_2, \cdots, x_r, +\infty, \cdots, +\infty), \end{aligned}$$

称 $F_{X_1, X_2, \cdots, X_r}(x_1, x_2, \cdots, x_r)$ 为$(X_1, X_2, \cdots, X_n)$ **关于 $X_1, X_2, \cdots, X_r$ 的边缘分布函数**，式中的 $r = 1, 2, \cdots, n-1$.

下面根据随机变量取值的特征，先讲述离散型随机变量的概率函数及边缘概率函数，再讲述连续型随机变量的分布密度及边缘分布密度.

1.1.2 离散型随机变量的概率函数及边缘概率函数

研究随机变量的分布函数有时会感到不甚方便，于是对离散型随机变量就

着重地研究它的概率函数. 明确了概率函数，可以进一步求出它的分布函数.

若 X 在集合$\{a_1,a_2,\cdots,a_i,\cdots\}$ 中取值且

$$P\{X=a_i\}=p_i,\quad 0\leqslant p_i\leqslant 1,\quad \sum_i p_i=1,$$

则称 X 为**一维离散型随机变量**，称 $P\{X=a_i\}=p_i$ 为 X **的概率函数或分布律**.

若 X 在集合$\{a_1,a_2,\cdots,a_i,\cdots\}$ 中取值，Y 在集合$\{b_1,b_2,\cdots,b_j,\cdots\}$ 中取值且

$$P\{X=a_i,\ Y=b_j\}=p_{ij},\quad 0\leqslant p_{ij}\leqslant 1,\quad \sum_i\sum_j p_{ij}=1,$$

则称(X,Y) 为**二维离散型随机变量**，称 $P\{X=a_i,\ Y=b_j\}=p_{ij}$ 为(X,Y) **的概率函数或分布律**.

根据随机事件$\{X=a_i\}$ 与$\sum\limits_j\{X=a_i,\ Y=b_j\}$ 等价，$\{Y=b_j\}$ 与$\sum\limits_i\{X=a_i,\ Y=b_j\}$ 等价，可以导出 X 的概率函数 $P\{X=a_i\}=\sum\limits_j p_{ij}$，$Y$ 的概率函数 $P\{Y=b_j\}=\sum\limits_i p_{ij}$. 若记

$$\sum_j p_{ij}=p_{i\cdot},\quad \sum_i p_{ij}=p_{\cdot j},$$

则称 X 的概率函数$P\{X=a_i\}=p_{i\cdot}$ 为(X,Y) **关于** X **的边缘概率函数**，称Y 的概率函数 $P\{Y=b_j\}=p_{\cdot j}$ 为(X,Y) **关于** Y **的边缘概率函数**，而(X,Y) 的概率函数又称为 X **与** Y **的联合概率函数**.

类似地，可以定义多维离散型随机变量$(X_1,X_2,\cdots,X_n)$ 的概率函数及边缘概率函数.

若$(X_1,X_2,\cdots,X_n)$ 在集合$\{(a_{j_1},a_{j_2},\cdots,a_{j_i},\cdots,a_{j_n})\colon j_i=1,2,\cdots,\ i=1,2,\cdots,n\}$ 中取值且

$$P\{X_1=a_{j_1},\ X_2=a_{j_2},\ \cdots,\ X_n=a_{j_n}\}=p_{j_1,j_2,\cdots,j_n},$$

$$0\leqslant p_{j_1,j_2,\cdots,j_n}\leqslant 1,\quad \sum_{j_1}\sum_{j_2}\cdots\sum_{j_n}p_{j_1,j_2,\cdots,j_n}=1,$$

则称$(X_1,X_2,\cdots,X_n)$ 为**多维离散型随机变量**，称

$$P\{X_1=a_{j_1},\ X_2=a_{j_2},\ \cdots,\ X_n=a_{j_n}\}=p_{j_1,j_2,\cdots,j_n}$$

为$(X_1,X_2,\cdots,X_n)$ **的概率函数**.

从 $X_1,X_2,\cdots,X_n$ 中任意确定r 个随机变量，不妨假设是 $X_1,X_2,\cdots,X_r$，根据随机事件$\{X_1=a_{j_1},\ X_2=a_{j_2},\ \cdots,\ X_r=a_{j_r}\}$ 与

$$\sum_{j_{r+1}}\cdots\sum_{j_n}\{X_1=a_{j_1},\ X_2=a_{j_2},\ \cdots,\ X_r=a_{j_r},X_{r+1}=a_{j_{r+1}},\cdots,X_n=a_{j_n}\}$$

等价，可以导出$(X_1,X_2,\cdots,X_r)$ 的概率函数为

$$P\{X_1 = a_{j_1},\ X_2 = a_{j_2},\ \cdots,\ X_r = a_{j_r}\}$$
$$= \sum_{j_{r+1}} \cdots \sum_{j_n} P\{X_1 = a_{j_1},\ X_2 = a_{j_2},\ \cdots,\ X_r = a_{j_r},$$
$$X_{r+1} = a_{j_{r+1}}, \cdots, X_n = a_{j_n}\},$$

称$(X_1, X_2, \cdots, X_r)$的概率函数为$(X_1, X_2, \cdots, X_n)$**关于** $X_1, X_2, \cdots, X_r$ **的边缘概率函数**.

1.1.3 连续型随机变量的分布密度及边缘分布密度

若 X 在实数轴 x 上的某个区间内取值，$p(x)$ 在 x 轴上有定义且

(1) $p(x) \geqslant 0$,

(2) $\int_{-\infty}^{+\infty} p(x)\mathrm{d}x = 1$,

(3) 对 x 轴上的任一区间 D_x 都有 $P\{X \in D_x\} = \int_{D_x} p(x)\mathrm{d}x$,

则称 X 为**一维连续型随机变量**，称 $p(x)$ 为 X **的分布密度**.

若(X,Y)在直角坐标平面 xOy 上的某个区域内取值，$p(x,y)$ 在平面 xOy 上有定义且

(1) $p(x,y) \geqslant 0$,

(2) $\iint_{xOy} p(x,y)\mathrm{d}x\,\mathrm{d}y = 1$,

(3) 对平面 xOy 上的任一区域 D_{xy} 都有

$$P\{(X,Y) \in D_{xy}\} = \iint_{D_{xy}} p(x,y)\mathrm{d}x\,\mathrm{d}y,$$

则称(X,Y)为**二维连续型随机变量**，称 $p(x,y)$ 为(X,Y) **的分布密度**.

若

$$p_X(x) = \int_{-\infty}^{+\infty} p(x,y)\mathrm{d}y, \quad p_Y(y) = \int_{-\infty}^{+\infty} p(x,y)\mathrm{d}x,$$

则称 $p_X(x)$ 为(X,Y) **关于** X **的边缘分布密度**，称 $p_Y(y)$ 为(X,Y) **关于** Y **的边缘分布密度**，它们分别是一维随机变量 X 与 Y 的分布密度，而 $p(x,y)$ 又称为 X **与** Y **的联合分布密度**.

类似地，可以定义多维连续型随机变量$(X_1, X_2, \cdots, X_n)$的分布密度及边缘分布密度.

若$(X_1, X_2, \cdots, X_n)$在 n 维直角坐标空间 $\mathbf{R}^n$ 中的某个部分空间内取值，$p(x_1, x_2, \cdots, x_n)$ 在 $\mathbf{R}^n$ 中有定义，且

(1) $p(x_1,x_2,\cdots,x_n)\geqslant 0$,

(2) $\int\cdots\int_{\mathbf{R}^n} p(x_1,x_2,\cdots,x_n)\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n=1$,

(3) 对 $\mathbf{R}^n$ 中的任一区域 $D_{x_1,x_2,\cdots,x_n}$ 都有

$$P\{(X_1,X_2,\cdots,X_n)\in D_{x_1,x_2,\cdots,x_n}\}$$
$$=\int\cdots\int_{D_{x_1,x_2,\cdots,x_n}} p(x_1,x_2,\cdots,x_n)\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n,$$

则称$(X_1,X_2,\cdots,X_n)$为**多维连续型随机变量**，称 $p(x_1,x_2,\cdots,x_n)$ 为$(X_1,X_2,\cdots,X_n)$ **的分布密度**.

从 $X_1,X_2,\cdots,X_n$ 中任意确定 r 个随机变量，不妨假设是 $X_1,X_2,\cdots,X_r$，若

$$p_{X_1,X_2,\cdots,X_r}(x_1,x_2,\cdots,x_r)$$
$$=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} p(x_1,x_2,\cdots,x_r,x_{r+1},\cdots,x_n)\mathrm{d}x_{r+1}\cdots\mathrm{d}x_n,$$

则称 $p_{X_1,X_2,\cdots,X_r}(x_1,x_2,\cdots,x_r)$ 为$(X_1,X_2,\cdots,X_n)$ **关于** $X_1,X_2,\cdots,X_r$ **的边缘分布密度**.

1.1.4　条件概率函数、条件分布密度与条件分布函数

设(X,Y) 为离散型随机变量，对固定的 j，若 $P\{Y=b_j\}=p_{\cdot j}>0$，根据条件概率的计算方法，有

$$P\{X=a_i\mid Y=b_j\}=\frac{P\{X=a_i,\ Y=b_j\}}{P\{Y=b_j\}}=\frac{p_{ij}}{p_{\cdot j}},$$

称 $P\{X=a_i\mid Y=b_j\}$ 为**在条件** $Y=b_j$ **下随机变量** X **的条件概率函数**，称 $F(x\mid b_j)=P\{X\leqslant x\mid Y=b_j\}$ 为**在条件** $Y=b_j$ **下随机变量** X **的条件分布函数**.

同理，对固定的 i，若 $P\{X=a_i\}=p_{i\cdot}>0$，则称 $P\{Y=b_j\mid X=a_i\}$ 为**在条件** $X=a_i$ **下随机变量** X **的条件概率函数**，称 $F(y\mid a_i)=P\{Y\leqslant y\mid X=a_i\}$ 为**在条件** $X=a_i$ **下随机变量** Y **的条件分布函数**.

类似地，可以定义多维离散型随机变量$(X_1,X_2,\cdots,X_n)$的条件概率函数及条件分布函数.

从 $X_1,X_2,\cdots,X_n$ 中任意确定 r 个随机变量，不妨假设是 $X_1,X_2,\cdots,X_r$，若 $P\{X_1=a_{j_1},\ X_2=a_{j_2},\ \cdots,\ X_r=a_{j_r}\}>0$，则称

$$P\{X_{r+1}=a_{j_{r+1}},\ \cdots,\ X_n=a_{j_n}\mid X_1=a_{j_1},\ \cdots,\ X_r=a_{j_r}\}$$
$$=\frac{P\{X_1=a_{j_1},\ X_2=a_{j_2},\ \cdots,\ X_n=a_{j_n}\}}{P\{X_1=a_{j_1},\ X_2=a_{j_2},\ \cdots,\ X_r=a_{j_r}\}}$$

为在条件 $X_1 = a_{j_1}$，$X_2 = a_{j_2}$，…，$X_r = a_{j_r}$ 下随机变量$(X_{r+1}, X_{r+2}, \cdots, X_n)$的条件概率函数，称

$$F(x_{r+1}, x_{r+2}, \cdots, x_n \mid a_{j_1}, a_{j_2}, \cdots, a_{j_r})$$
$$= P\{X_{r+1} \leqslant x_{r+1}, \cdots, X_n \leqslant x_n \mid X_1 = a_{j_1}, \cdots, X_r = a_{j_r}\}$$

为在条件 $X_1 = a_{j_1}$，$X_2 = a_{j_2}$，…，$X_r = a_{j_r}$ 下随机变量$(X_{r+1}, X_{r+2}, \cdots, X_n)$的条件分布函数.

设(X,Y)为连续型随机变量，联合分布密度为 $p(x,y)$，边缘分布密度为 $p_X(x)$ 和 $p_Y(y)$，对固定的 y，若 $p_Y(y) > 0$，则称

$$p(x \mid y) = \frac{p(x,y)}{p_Y(y)}$$

为在条件 $Y = y$ 下随机变量 X 的条件分布密度，称

$$F(x \mid y) = \frac{F(x,y)}{p_Y(y)}$$

为在条件 $Y = y$ 下随机变量 X 的条件分布函数.

同理，对固定的 x，若 $p_X(x) > 0$，称

$$p(y \mid x) = \frac{p(x,y)}{p_X(x)}$$

为在条件 $X = x$ 下随机变量 Y 的条件分布密度，称

$$F(y \mid x) = \frac{F(x,y)}{p_X(x)}$$

为在条件 $X = x$ 下随机变量 Y 的条件分布函数.

类似地，可以定义多维连续型随机变量$(X_1, X_2, \cdots, X_n)$的条件分布密度及条件分布函数.

从 $X_1, X_2, \cdots, X_n$ 中任意确定 r 个随机变量，不妨假设是 $X_1, X_2, \cdots, X_r$，对固定的 $x_1, x_2, \cdots, x_r$，若 $p_{X_1, X_2, \cdots, X_r}(x_1, x_2, \cdots, x_r) > 0$，则称

$$p(x_{r+1}, x_{r+2}, \cdots, x_n \mid x_1, x_2, \cdots, x_r) = \frac{p(x_1, x_2, \cdots, x_r, x_{r+1}, x_{r+2}, \cdots, x_n)}{p_{X_1, X_2, \cdots, X_r}(x_1, x_2, \cdots, x_r)}$$

为在条件 $X_1 = x_1, X_2 = x_2, \cdots, X_r = x_r$ 下随机变量$(X_{r+1}, X_{r+2}, \cdots, X_n)$的条件分布密度，称

$$F(x_{r+1}, x_{r+2}, \cdots, x_n \mid x_1, x_2, \cdots, x_r) = \frac{F(x_1, x_2, \cdots, x_r, x_{r+1}, x_{r+2}, \cdots, x_n)}{p_{X_1, X_2, \cdots, X_r}(x_1, x_2, \cdots, x_r)}$$

为在条件 $X_1 = x_1$，$X_2 = x_2$，…，$X_r = x_r$ 下随机变量$(X_{r+1}, X_{r+2}, \cdots, X_n)$的条件分布函数.

1.1.5　两个随机变量相互独立

对任意实数 x 与 y，设 $F(x,y)$ 为二维随机变量(X,Y) 的联合分布函数，$F_X(x),F_Y(y)$ 分别为 X 与 Y 的分布函数. 若 $F(x,y)=F_X(x)F_Y(y)$，则称**随机变量 X 与 Y 相互独立**.

定理 1.1.1　(1) 对于离散型二维随机变量(X,Y)，若 X 在集合$\{a_1,a_2,\cdots,a_i,\cdots\}$ 中取值，Y 在集合$\{b_1,b_2,\cdots,b_j,\cdots\}$ 中取值，且

$$P\{X=a_i,\ Y=b_j\}=p_{ij},\quad P\{X=a_i\}=p_{i\cdot},\quad P\{Y=b_j\}=p_{\cdot j}$$

分别为(X,Y) 及 X 与 Y 的概率函数时，则 X 与 Y 相互独立的充要条件是对a_i 与 b_j 的一切组合，

$$P\{X=a_i,\ Y=b_j\}=P\{X=a_i\}P\{Y=b_j\}$$

或 $p_{ij}=p_{i\cdot}p_{\cdot j}$.

(2) 对于连续型二维随机变量(X,Y)，若 $p(x,y)$ 及 $p_X(x),p_Y(y)$ 分别是(X,Y) 及 X 与 Y 的分布密度，则 X 与 Y 相互独立的充要条件是在 $p(x,y)$ 及$p_X(x),p_Y(y)$ 都连续的点(x,y) 处，

$$p(x,y)=p_X(x)p_Y(y).$$

以下对(1) 给出证明，(2) 的证明与(1) 相类似.

证　充分性　对 a_i 与b_j 的一切组合，当 $P\{X=a_i,\ Y=b_j\}=P\{X=a_i\}P\{Y=b_j\}$ 时，

$$\begin{aligned}F(x,y)&=\sum_{a_i\leqslant x}\sum_{b_j\leqslant y}P\{X=a_i,\ Y=b_j\}\\&=\sum_{a_i\leqslant x}\sum_{b_j\leqslant y}P\{X=a_i\}P\{Y=b_j\}\\&=\sum_{a_i\leqslant x}P\{X=a_i\}\sum_{b_j\leqslant y}P\{Y=b_j\}\\&=F_X(x)F_Y(y),\end{aligned}$$

因此，X 与 Y 相互独立.

必要性　当 X 与 Y 相互独立时，$F(x,y)=F_X(x)F_Y(y)$，即

$$P\{X\leqslant x,\ Y\leqslant y\}=P\{X\leqslant x\}P\{Y\leqslant y\},$$

于是对 a_i 与b_j 的一切组合，有

$$P\{X\leqslant a_i,\ Y\leqslant b_j\}=P\{X\leqslant a_i\}P\{Y\leqslant b_j\},$$

$$P\{X\leqslant a_i,\ Y<b_j\}=P\{X\leqslant a_i\}P\{Y<b_j\}.$$

两式相减得到

$$P\{X \leqslant a_i, Y = b_j\} = P\{X \leqslant a_i\}P\{Y = b_j\}.$$

同样地

$$P\{X < a_i, Y = b_j\} = P\{X < a_i\}P\{Y = b_j\}.$$

两式相减得到

$$P\{X = a_i, Y = b_j\} = P\{X = a_i\}P\{Y = b_j\}.$$ □

1.1.6 多个随机变量相互独立

对任意实数 $x_1, x_2, \cdots, x_n$，设 $F(x_1, x_2, \cdots, x_n)$ 为 n 维随机变量 $(X_1, X_2, \cdots, X_n)$ 的分布函数，$F_1(x_1), F_2(x_2), \cdots, F_n(x_n)$ 分别为 $X_1, X_2, \cdots, X_n$ 的分布函数. 若

$$F(x_1, x_2, \cdots, x_n) = F_1(x_1)F_2(x_2)\cdots F_n(x_n),$$

则称**随机变量** $X_1, X_2, \cdots, X_n$ **相互独立**. 其充要条件与两个随机变量的情形相类似.

注 (1) 当随机变量 $X_1, X_2, \cdots, X_n$ 相互独立时，上述随机变量中的任意 k 个随机变量也相互独立，其中 $k = 2, 3, \cdots, n-1$，即全体相互独立时其部分也相互独立.

(2) 当随机变量 $X_1, X_2, \cdots, X_n$ 相互独立时，它们各自的连续函数 $f_1(X_1), f_2(X_2), \cdots, f_n(X_n)$ 也相互独立.

1.1.7 两组随机变量相互独立

对任意实数 $x_1, x_2, \cdots, x_n$ 与 $y_1, y_2, \cdots, y_m$，设 n 维随机变量 $(X_1, X_2, \cdots, X_n)$ 的分布函数为 $F_1(x_1, x_2, \cdots, x_n)$，$m$ 维随机变量 $(Y_1, Y_2, \cdots, Y_m)$ 的分布函数为 $F_2(y_1, y_2, \cdots, y_m)$，$n+m$ 维随机变量 $(X_1, X_2, \cdots, X_n, Y_1, Y_2, \cdots, Y_m)$ 的分布函数为 $F(x_1, x_2, \cdots, x_n, y_1, y_2, \cdots, y_m)$. 若

$$F(x_1, x_2, \cdots, x_n, y_1, y_2, \cdots, y_m) = F_1(x_1, x_2, \cdots, x_n)F_2(y_1, y_2, \cdots, y_m),$$

则称 $(X_1, X_2, \cdots, X_n)$ 与 $(Y_1, Y_2, \cdots, Y_m)$ **相互独立**. 其充要条件与两个随机变量的情形相类似.

注 当随机变量 $(X_1, X_2, \cdots, X_n)$ 与 $(Y_1, Y_2, \cdots, Y_m)$ 相互独立时，

(1) 任一 $X_i (i = 1, 2, \cdots, n)$ 与 $Y_j (j = 1, 2, \cdots, m)$ 相互独立；

(2) $(X_1, X_2, \cdots, X_n)$ 的连续函数 $h(X_1, X_2, \cdots, X_n)$ 与 $(Y_1, Y_2, \cdots, Y_m)$ 的连续函数 $g(Y_1, Y_2, \cdots, Y_m)$ 也相互独立. 例如，$\overline{X} = \frac{1}{n}\sum_i X_i$ 与 $\overline{Y} = \frac{1}{m}\sum_j Y_j$ 相

互独立，$S_X^2=\frac{1}{n}\sum_i(X_i-\overline{X})^2$ 与 $S_Y^2=\frac{1}{m}\sum_j(Y_j-\overline{Y})^2$ 也相互独立.

要特别注意的是：随机变量的相互独立性在统计学中是一个十分重要的基本概念，有关的一些结论将成为研究后续若干内容的依据. 但是，随机变量的相互独立性通常都不是根据定义而是根据专业知识或通过抽样方法与试验设计来确定的. 在统计学中，只要随机变量是相互独立的，便可以根据上述充要条件写出它们所构成的二维或多维随机变量的分布函数、概率函数或分布密度. 也有少数情形，随机变量的相互独立性仍然需要通过证明来加以确认，由例 1.1.1 ～ 例 1.1.3 可见证明的依据及步骤. 在 1.4 节中还要证明 $X_1,X_2,\cdots,X_n$ 相互独立且都服从正态分布时，$\overline{X}=\frac{1}{n}\sum_i X_i$ 与 $S_X^2=\frac{1}{n}\sum_i(X_i-\overline{X})^2$ 相互独立，证明时将用到标准正态分布作正交变换的结论.

【例 1.1.1】 证明：若随机变量 X 只取一个值 a，则 X 与任意的随机变量相互独立.

证 由 $P\{X=a\}=1$ 知，它的分布函数为

$$F_1(x)=\begin{cases}0, & x<a,\\ 1, & x\geqslant a.\end{cases}$$

设任意的随机变量 Y 的分布函数为 $F_2(y)$，考虑 (X,Y) 的分布函数

$$F(x,y)=P\{X\leqslant x,Y\leqslant y\}.$$

当 $x<a$ 时，$\{X\leqslant x\}$ 与 $\{X\leqslant x,Y\leqslant y\}$ 为不可能事件，$F_1(x)=0$，

$$F(x,y)=P\{X\leqslant x,Y\leqslant y\}=0,$$

从而 $F(x,y)=F_1(x)F_2(y)$.

当 $x\geqslant a$ 时，$\{X\leqslant x\}$ 为必然事件，$F_1(x)=1$，$\{X\leqslant x,Y\leqslant y\}=\{Y\leqslant y\}$，从而

$$F(x,y)=P\{X\leqslant x,Y\leqslant y\}=P\{Y\leqslant y\}=F_2(y),$$

故 $F(x,y)=F_1(x)F_2(y)$.

因此，X 与 Y 相互独立. □

注 由于随机变量 X 只取一个值 a，可以看做是一个常量，因此上述结论应该理解为常量与任一随机变量相互独立.

【例 1.1.2】 设三维随机变量 (X,Y,Z) 的分布密度为

$$p(x,y,z)=\begin{cases}(x+y)\mathrm{e}^{-z}, & 0<x<1,0<y<1,z>0,\\ 0, & \text{其他},\end{cases}$$

试证明：X 与 Z 相互独立，X 与 (Y,Z) 不相互独立.

证 当 $0<y<1$，$z>0$ 时，(Y,Z) 的分布密度

$$p_{23}(y,z)=\int_0^1(x+y)\mathrm{e}^{-z}\mathrm{d}x=\left(\frac{1}{2}+y\right)\mathrm{e}^{-z};$$

当 $0<x<1$，$z>0$ 时，(X,Z) 的分布密度

$$p_{13}(x,z)=\int_0^1(x+y)\mathrm{e}^{-z}\mathrm{d}y=\left(\frac{1}{2}+x\right)\mathrm{e}^{-z};$$

当 $0<x<1$ 时，X 的分布密度

$$p_1(x)=\int_0^{+\infty}\left(\frac{1}{2}+x\right)\mathrm{e}^{-z}\mathrm{d}z=\frac{1}{2}+x;$$

当 $z>0$ 时，Z 的分布密度

$$p_3(x)=\int_0^1\left(\frac{1}{2}+x\right)\mathrm{e}^{-z}\mathrm{d}x=\mathrm{e}^{-z}.$$

从而

$$p_{23}(y,z)=\begin{cases}\left(\frac{1}{2}+y\right)\mathrm{e}^{-z}, & 0<y<1,\ z>0,\\ 0, & \text{其他},\end{cases}$$

$$p_{13}(x,z)=\begin{cases}\left(\frac{1}{2}+x\right)\mathrm{e}^{-z}, & 0<x<1,\ z>0,\\ 0, & \text{其他},\end{cases}$$

$$p_1(x)=\begin{cases}\frac{1}{2}+x, & 0<x<1,\\ 0, & \text{其他},\end{cases}$$

$$p_3(z)=\begin{cases}\mathrm{e}^{-z} & z>0,\\ 0, & \text{其他}.\end{cases}$$

由此易知，对于任意的实数 x 与 z，$p_{13}(x,z)=p_1(x)p_3(z)$，故 X 与 Z 相互独立；对于 $0<x<1$，$0<y<1$，$z>0$，$p(x,y,z)\neq p_1(x)p_{23}(y,z)$，故 X 与 (Y,Z) 不相互独立. □

【例 1.1.3】 设三维随机变量 (X,Y,Z) 的分布密度为

$$p(x,y,z)=\begin{cases}\frac{1-\sin x\sin y\sin z}{8\pi^3}, & 0<x<2\pi,\ 0<y<2\pi,\ 0<z<2\pi,\\ 0, & \text{其他},\end{cases}$$

试证明：X,Y,Z 两两相互独立，但 X,Y,Z 不相互独立.

证 当 $0<x<2\pi$，$0<y<2\pi$ 时，(X,Y) 的分布密度为

$$p_{12}(x,y)=\int_0^{2\pi}\frac{1-\sin x\sin y\sin z}{8\pi^3}\mathrm{d}z=\frac{1}{4\pi^2},$$