·通用管理系列教材·

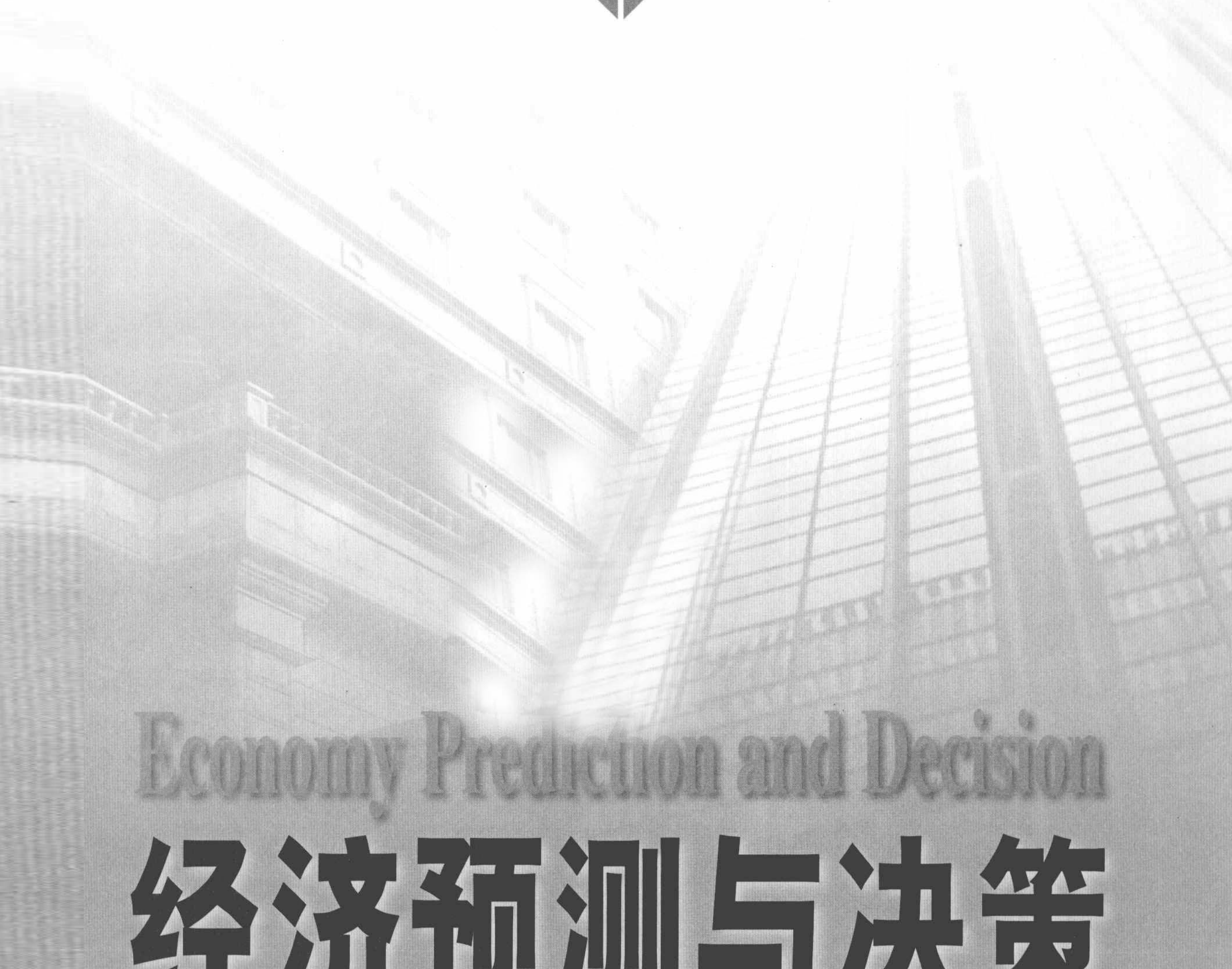
Economy Prediction and Decision

经济预测与决策

主 编　吴仁群

中国人民大学出版社

·通用管理系列教材·



Economy Prediction and Decision

经济预测与决策

图书在版编目(CIP)数据

经济预测与决策/吴仁群主编．—北京：中国人民大学出版社，2011.5
通用管理系列教材
ISBN 978-7-300-13670-7

Ⅰ.①经… Ⅱ.①吴… Ⅲ.①经济预测-高等学校-教材②经济决策-高等学校-教材 Ⅳ.①F20

中国版本图书馆 CIP 数据核字（2011）第 076668 号

通用管理系列教材
经济预测与决策
主　编　吴仁群
Jingji Yuce yu Juece

出版发行	中国人民大学出版社		
社　　址	北京中关村大街 31 号	邮政编码	100080
电　　话	010－62511242（总编室）		010－62511398（质管部）
	010－82501766（邮购部）		010－62514148（门市部）
	010－62515195（发行公司）		010－62515275（盗版举报）
网　　址	http://www.crup.com.cn		
	http://www.ttrnet.com(人大教研网)		
经　　销	新华书店		
印　　刷	北京密兴印刷有限公司		
规　　格	185 mm×260 mm　16 开本	版　　次	2011 年 9 月第 1 版
印　　张	13.75 插页 1	印　　次	2011 年 9 月第 1 次印刷
字　　数	278 000	定　　价	25.00 元

前言

PREFACE

古人云：凡事预则立，不预则废。所谓预测是指通过对事物的过去和现在的情况进行分析、研究，找出其发展变化的规律，从而利用一定的方法或技术来预计和推测未来的情况，简言之，就是根据过去和现在来预计未来，根据已知的信息来推测未知的信息。经济预测作为预测的一个分支，是预测理论和方法在经济领域中的应用。因此，可以认为经济预测是根据经济活动的历史和现实，运用定性和定量方法，揭示出经济活动的客观规律，分析经济现象之间的联系以及作用机制，指出各类经济现象和经济过程未来发展的可能途径以及结果。预测是为决策提供信息的重要手段，是科学决策的基础。在经济活动中，预测有助于克服决策的盲目性，减少不确定性。经济领域中的许多问题都需要作预测。例如，预测未来一段时间内的生产发展趋势、商品需求情况。企业中典型的财务预测、成本预测以及发展前景预测等都属经济预测范畴。

本书共 10 章，还有一个附录。第 1 章介绍了经济预测与决策中涉及的统计知识；第 2 章介绍了经济预测的含义、类型、基本原理以及预测数据的处理；第 3 章介绍了定性预测的含义以及常见的定性预测方法：头脑风暴法、德尔菲法、主观概率法和对比类推法等；第 4 章介绍了趋势外推预测法的基本原理和应用举例；第 5 章介绍了回归分析预测法的基本原理和应用举例；第 6 章介绍了常见的确定型时间序列预测法，并举例进行说明；第 7 章介绍了随机型时间序列预测模型、自相关函数和偏相关函数、模型识别、参数估计以及预测等；第 8 章介绍了马尔科夫预测法的基本原理和应用举例；第 9 章介绍了决策的含义、基本过程，并对确定型决策、非确定型决策和风险型决策进行了介绍及举例说明；第 10 章介绍了主成分分析法、层次分析法和数据包络法等常用决策方法，并举例说明这些方法的应用。附录收集了作者编写的用于预测和决策的宏函数的使用说明，有关宏函数的源代码见作者提供的 Excel 电子计算模版，读者可登录 www. rdjg. com. cn 获取。

本书由北京印刷学院的吴仁群主编，系“北京市属市管高等学校人才强教计划资助项目”。在编写过程中，参考了本书参考文献所列举的资料，并得到了中国人民大学出版社的大力支持，在此对参考文献中图书的作者及中国人民大学出版社表示深深的感谢。

由于时间仓促，书中难免存在一些不足之处，敬请读者批评指正。

吴仁群

教师教学服务说明

中国人民大学出版社工商管理分社以出版经典、高品质的工商管理、财务会计、统计、市场营销、人力资源管理、运营管理、物流管理、旅游管理等领域的各层次教材为宗旨。同时，为了更好地服务于一线教师教学，工商管理分社近年来着力建设数字化、立体化的网络教学资源。老师们可以通过以下方式获得免费下载教学资源的权限：

（1）在“人大经管图书在线”（www. rdjg. com. cn）注册并下载“教师服务登记表”，或者直接填写下面的“教师服务登记表”后，加盖院系公章，然后邮寄或者传真给我们。我们收到表格后将在一个工作日内为您开通相关资源的下载权限。

（2）如果您有“人大出版社教研服务网络”（http://www. ttrnet. com）会员卡，可以将卡号发到我们的公共邮箱，无须重复注册，我们将直接为您开通相关专业领域教学资源的下载权限。

如果您需要帮助，请随时联系我们：

联系人：刘玉仙（010－62515735）　　李文重（010－82501704）

传真：010－62514775　　邮箱：rdcbsjg@crup. com. cn

邮寄地址：北京市海淀区中关村大街甲 59 号文化大厦 1501 室

中国人民大学出版社工商管理分社　邮编：100872

教师服务登记表

姓　　名		□先生　□女士	职　　称		
座机/手机			电子邮箱		
通信地址			邮　　编		
任教学校			所在院系		
所授课程	课程名称	现用教材名称	出版社	对象（本科生/研究生/MBA/其他）	学生人数
需要哪本教材的配套资源					
人大经管图书在线用户名					
院/系领导（签字）： 院/系办公室盖章					

目录

CONTENTS

第1章 统计基础

1.1 随机变量

1.1.1 随机变量及分布

1. 随机变量的含义

一个随机试验的可能结果（称为基本事件）的全体组成一个样本空间 Ω。

随机变量 X 是定义在样本空间 Ω 上的取值为实数的函数，即对每一个随机试验 $e\in\Omega$，有一个实数 $X(e)$ 与之对应。则称定义在 Ω 上的实值单值函数 $X=X(e)$ 为随机变量。

当一个随机变量的取值范围仅为有限个或可列无限个实数时，该随机变量称为离散型随机变量。例如，记 X 为一天内某证券交易市场的股民数，X 可能取值为 $\{0, 1, 2, 3, \cdots\}$，X 为离散型随机变量。

当一个随机变量的取值范围为数轴上的一个区间 (a, b) 时，该随机变量称为连续型随机变量。这里 a 可以为 $-\infty$，b 可以为 $+\infty$。例如，记 Y 为某产品直径，Y 可能取区间 $[1.2, 2.3]$ 之间的任何数，Y 为连续型随机变量。

2. 分布函数

设 X 是随机变量，x 为任意实数，函数 $F(x)=P\{X\leqslant x\}$ 为 X 的分布函数。对于任意实数 x_1，$x_2(x_1<x_2)$，有

$$P(x_1<X\leqslant x_2)=P(X\leqslant x_2)-P(X\leqslant x_1)=F(x_2)-F(x_1)$$

由上可知，若已知 X 的分布函数，就可以知道 X 落在任一区间 $(x_1, x_2]$ 上的概率，从这个意义上讲，分布函数完整地描述了随机变量的统计规律性。

3. 离散型随机变量的分布律

对离散型随机变量而言，常用分布律来表示其分布。

离散型随机变量 X 的分布律就是 X 所有可能取值及其概率。如果 X 的所有可能取值为 x_1，x_2，…，x_k，…，则 X 的分布律用公式表示为：

$$P\{X=x_i\}=p_i,\quad i=1,2,\cdots$$

分布律也可用表格方式表示，如表1—1所示。

表1—1　　随机变量 X 的分布律

X	x_1	x_2	$\cdots$	x_k	$\cdots$
p_i	p_1	p_2	$\cdots$	p_k	$\cdots$

以下是常见的离散型随机变量的分布律。

(1) 0—1分布。如果随机变量 X 只能取0或1两个值，其分布律为：

$$P\{X=k\}=p^k(1-p)^{1-k},\quad k=0,1;0<p<1$$

则称 X 服从0—1分布（见表1—2）。

表1—2　　0—1分布

X	0	1
p_k	p	$1-p$

(2) 二项分布。设随机实验的结果只有 A 或 $\overline{A}$，$p(A)=p$，$p(\overline{A})=1-p=q$。将实验独立地重复进行 n 次，则称这一串重复实验为 n 重伯努利试验，简称伯努利试验。其分布律为：

$$P\{X=k\}=C_n^k p^k q^{n-k},\quad k=0,1,\cdots,n$$

如果随机变量 X 的分布律满足上式，则称随机变量 X 服从参数为 n，p 的二项分布，记为 $X\sim b(n,\ p)$。

(3) 泊松分布。设随机变量 X 所有可能取值为0，1，2，…，取各个值的概率为：

$$P\{X=k\}=\frac{\lambda^k e^{1-\lambda}}{k!},\quad k=0,1,2,\cdots$$

式中，$\lambda>0$，是常数，则称随机变量 X 服从参数为 λ 的泊松分布，记为 $X\sim\pi(\lambda)$。

4. 连续型随机变量的概率密度函数

对随机变量 X 的分布函数 $F(X)$，若存在非负函数 $f(x)$ 使得对于任意实数 x，有

$$F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t$$

则称 $f(x)$ 为 X 的概率密度函数。

以下是常见的连续型随机变量的概率密度函数。

(1) 均匀分布。若连续型随机变量 X 的概率密度函数为：

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\ 0, & 其他\end{cases}$$

则称 X 在区间 $(a,\ b)$ 上服从均匀分布。

(2) 正态分布。若连续型随机变量 X 的概率密度函数为：

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty<x<\infty$$

则称 X 服从参数为 μ，σ 的正态分布，记为 $X\sim N(\mu, \sigma^2)$。其中 μ，$\sigma(\sigma>0)$ 为常数。

若 $\mu=0$，$\sigma=1$，则称 X 服从标准正态分布，记为 $X\sim N(0, 1)$。

对一般正态分布可通过下面的变换转换成标准正态分布。

若 $X\sim N(\mu, \sigma^2)$，则 $Z=\frac{X-\mu}{\sigma}\sim N(0, 1)$。

若 $X\sim N(0, 1)$，z_α 满足以下条件：

$$P(X>z_\alpha)=\alpha, \quad 0<\alpha<1$$

则称 z_α 为标准正态分布的上 α 分位点，如图 1—1 所示。

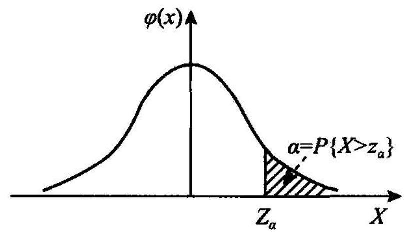


图 1—1 正态分布的上 α 分位点

1.1.2 随机变量的数字特征

1. 数学期望及方差

对离散型随机变量 X 而言，数学期望是 X 的各种取值与其概率的乘积。若离散型随机变量 X 的分布律为：

$$P\{X=x_k\}=p_k, \quad k=0,1,\cdots$$

X 的数学期望 $E(X)$ 为：

$$E(X)=\sum_{k=1}^{\infty}x_k p_k$$

对连续型随机变量 X 而言，其数学期望可通过概率密度函数来计算。

设连续型随机变量 X 的概率密度函数为 $f(x)$，如果积分 $\int_{-\infty}^{\infty}xf(x)\mathrm{d}x$ 绝对收敛，则称其为 X 的数学期望，记为 $E(X)$，即

$$E(X)=\int_{-\infty}^{\infty}xf(x)\mathrm{d}x$$

设 X 是随机变量，若 $E\{[X-E(X)]^2\}$ 存在，则称其为 X 的方差，记为 $D(X)$，即

$$D(X)=E\{[X-E(X)]^2\}$$

方差用于度量随机变量 X 与其均值 $E(X)$ 即期望的偏离程度。

$\sqrt{D(X)}$为 X 的标准差或均方差，记为 $\sigma(X)$。

表 1—3 中是常见随机变量的数学期望及方差。

表 1—3　　常见随机变量的数学期望及方差

分布	参数	分布律或概率密度函数	数学期望 $E(X)$	方差 $D(X)$
0−1 分布	$0<p<1$	$P\{X=k\}=p^k(1-p)^{1-k}(k=0,1)$	p	$p(1-p)$
二项分布	$n\geqslant1$ $0<p<1$	$P\{X=k\}=C_n^kp^kq^{1-k}(k=0,1,\cdots,n)$	np	$np(1-p)$
泊松分布	$\lambda>0$	$P\{X=k\}=\frac{\lambda^k e^{1-\lambda}}{k!}(k=0,1,2,\cdots)$	λ	λ
均匀分布	$a<b$	$f(x)=\begin{cases}\frac{1}{b-a}, & a<x<b\\ 0, & \text{其他}\end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
正态分布	μ，σ $\sigma>0$	$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2

2. 协方差及相关系数

$E\{[X-E(X)][Y-E(Y)]\}$ 称为随机变量 X 与 Y 的协方差，记为 $\mathrm{Cov}(X,Y)$，即

$$\mathrm{Cov}(X,Y)=E\{[X-E(X)][Y-E(Y)]\}$$

协方差作为描述随机变量 X 和 Y 相关程度的量，在同一物理量纲之下有一定的作用，但同样的两个量采用不同的量纲使它们的协方差在数值上表现出很大的差异。此时可使用相关系数来表示。

随机变量 X 与 Y 的相关系数为：

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

当 $\rho_{XY}=0$ 时，称 X 与 Y 不相关；当 $\rho_{XY}>0$ 时，称 X 与 Y 正相关；当 $\rho_{XY}<0$ 时，称 X 与 Y 负相关。

3. 矩、协方差矩阵

设 X 和 Y 是随机变量，若 $E(X^k)$（$k=1,2,\cdots$）存在，则称其为 X 的 k 阶原点矩，简称 k 阶矩。

若 $E\{[X-E(X)]^k\}(k=1,2,\cdots)$ 存在，则称其为 X 的 k 阶中心矩。

若 $E(X^kY^l)(k,l=1,2,\cdots)$ 存在，则称其为 X 和 Y 的 $k+l$ 阶混合矩。

若 $E\{[X-E(X)]^k[Y-E(Y)]^l\}(k,l=1,2,\cdots)$ 存在，则称其为 X 和 Y 的 $k+l$ 阶混合中心矩。

设 n 维随机变量（$X_1,X_2,\cdots,X_n$）的二阶混合中心矩

$$c_{ij}=\mathrm{Cov}(X_i, X_j)=E\{[X_i-E(X_i)][X_j-E(X_j)]\}, \quad i, j=1, 2, \cdots, n$$

都存在，则称矩阵

$$C=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

为 n 维随机变量（X_1，X_2，…，X_n）的协方差矩阵。

1.1.3 应用举例

例 1—1 已知随机变量 X 的概率分布如表 1—4 所示。请计算 X 的数学期望和方差。

表 1—4 **随机变量 X 的概率分布**

随机变量 X	7	3	5	6	4
概率	0.1	0.2	0.3	0.3	0.1

解 随机变量 X 的数学期望值 $E(X)$ 为：

$$E(X)=\sum_{i=1}^{n}X_ip_i=7\times0.1+3\times0.2+5\times0.3+6\times0.3+4\times0.1=5$$

随机变量 X 的方差 $\mathrm{Var}(X)$ 为：

$$\begin{aligned}\mathrm{Var}(X) &= E\{[X-E(X)]^2\}=\sum_{i=1}^{n}X_i[-E(X)]^2p_i \\ &=(7-5)^2\times0.1+(3-5)^2\times0.2+(5-5)^2\times0.3 \\ &\quad+(6-5)^2\times0.3+(4-5)^2\times0.1 \\ &=1.6\end{aligned}$$

以下说明如何使用 Excel 中的函数来计算数学期望和方差。

在 Excel 工作表中建立计算模版如表 1—5 所示。

表 1—5 **计算模版**

	A	B	C
1	随机变量 X	概率	$[X-E(X)]^2$
2	7.00	0.10	4.00
3	3.00	0.20	4.00
4	5.00	0.30	0.00
5	6.00	0.30	1.00
6	4.00	0.10	1.00
7			
8	数学期望 E(X)	5	
9	方差	1.6	
10	标准差	1.26	

有关表 1—5 各单元格中公式及其功能的描述详见表 1—6。

表 1—6　　表 1—5 各单元格中公式及其功能

单元格	公式	功能	公式复制操作
B8	=sumproduct(a2:a6, b2:b6)	计算 $E(X)$	
C2	=(a2-b8)*(a2-b8)	计算 $[X-E(X)]^2$	复制到区域 c3:c6
B9	=sumproduct(c2:c6, b2:b6)	计算 $D(X)$	
B10	=sqrt(b9)	计算 $\sigma(X)$	

说明：公式复制操作按照如下步骤进行，首选选择被复制对象并按组合键 Ctrl+C，其次选择复制区域并按组合键 Ctrl+V 完成复制。本书所有公式复制操作均如此实现，故以后不再介绍。

表 1—6 中函数 sumproduct(arr1，arr2) 的功能是计算单元格区域 arr1 和 arr2 对应单元格的乘积和。sqrt(　) 的功能是求平方根。这两个函数为 Excel 内置函数，详细说明见 Excel 中帮助项。

因此，只要在有关单元格输入计算公式后，便可自动计算出数学期望和方差，同时当随机变量 X 的取值及其概率发生变化时，计算结果也随之发生变化。这也是利用 Excel 进行计算的好处，即对同一类型问题，只要设计一个公共计算模版后，并在需要输入数据的地方输入对应的数据，便可实现不同计算。

1.2 样本及抽样分布

1.2.1 基础知识

1. 总体、个体、样本

研究对象的某项数量指标的值的全体称为总体，构成总体的每个成员称为个体。

设 X 是具有分布函数 F 的随机变量，若 X_1，X_2，…，X_n 是具有同一分布函数 F 且相互独立的随机变量，则称 X_1，X_2，…，X_n 为从分布函数 F 得到的容量为 n 的简单随机样本，简称样本，它们的观察值 x_1，x_2，…，x_n 称为样本值，又称为 X 的 n 个独立的观察值。

2. 统计量及观察值

设 X_1，X_2，…，X_n 是来自总体 X 的一个样本，$g(X_1, X_2, \cdots, X_n)$ 是 X_1，X_2，…，X_n 的函数，若 g 是连续函数且 g 中不含任何未知参数，则称 $g(X_1, X_2, \cdots, X_n)$ 为统计量。

若 x_1，x_2，…，x_n 为样本 X_1，X_2，…，X_n 的观察值，则 $g(x_1, x_2, \cdots, x_n)$ 为统计量 $g(X_1, X_2, \cdots, X_n)$ 的观察值。

表 1—7 给出了常见统计量及观察值。

表 1—7 常见统计量及观察值

统计量名称	统计量	观察值
样本平均值	$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$	$\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$
样本方差	$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$	$s^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2$
样本标准差	$S=\sqrt{S^2}$	$s=\sqrt{s^2}$
样本 k 阶矩	$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k(k=1,2,\cdots)$	$a_k=\frac{1}{n}\sum_{i=1}^{n}x_i^k(k=1,2,\cdots)$
样本 k 阶中心矩	$B_k=\frac{1}{n}\sum_{i=1}^{n}(X_k-\overline{X})^k(k=1,2,\cdots)$	$b_k=\frac{1}{n}\sum_{i=1}^{n}(x_k-\overline{x})^k(k=1,2,\cdots)$

1.2.2 抽样分布

统计量是样本的函数，它是一个随机变量，其分布称为抽样分布。

以下介绍来自正态总体的几种常见统计量的分布。

1. χ^2 分布

设 X_1，X_2，…，X_n 是来自总体 $N(0,1)$ 的样本，则称统计量 $\chi^2=X_1^2+X_2^2+\cdots+X_n^2$ 服从自由度为 n 的 χ^2 分布，记为 $\chi^2\sim\chi^2(n)$。

$\chi^2(n)$ 分布的概率密度函数为：

$$F(y)=\begin{cases}\dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\dfrac{n}{2}\right)}y^{\frac{n}{2}-1}e^{-\frac{y}{2}}, & y>0\\ 0, & \text{其他}\end{cases}$$

对于给定的正数 α（$0<\alpha<1$），称满足条件 $P\{\chi^2>\chi_\alpha^2(n)\}=\int_{\chi_\alpha^2(n)}^{\infty}f(y)\mathrm{d}y=\alpha$ 的点 $\chi_\alpha^2(n)$ 为 $\chi^2(n)$ 分布的上 α 分位点，如图 1—2 所示。

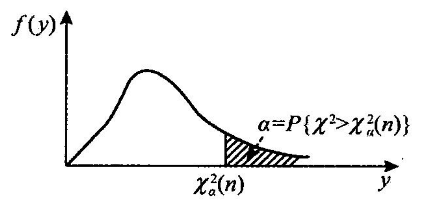


图 1—2 χ^2 分布的上 α 分位点

2. t 分布

设 $X\sim N(0,1)$，$Y\sim\chi^2(n)$，X 与 Y 相互独立，则称统计量 $t=\dfrac{X}{\sqrt{\dfrac{Y}{n}}}$ 服从自由度为 n 的 t 分布，记为 $t\sim t(n)$。

$t(n)$ 分布的概率密度函数为：

$$h(t)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}},\quad -\infty<t<\infty$$

对于给定的正数 α（$0<\alpha<1$），称满足条件 $P\{t>t_\alpha(n)\}=\int_{t_\alpha(n)}^{\infty}h(t)\mathrm{d}t=\alpha$ 的点 $t_\alpha(n)$为 $t(n)$ 分布的上 α 分位点，如图 1—3 所示。

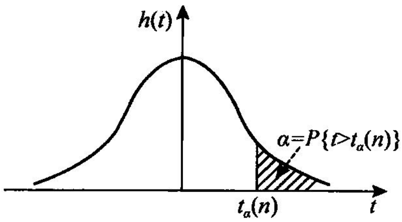


图 1—3　t 分布的上 α 分位点

3. F 分布

设 $U\sim\chi^2(n_1)$，$V\sim\chi^2(n_2)$，U 与 V 相互独立，则称统计量 $F=\dfrac{\frac{U}{n_1}}{\frac{V}{n_2}}$ 服从自由度为（n_1，n_2）的 F 分布，记为 $F\sim F(n_1,\ n_2)$。

$F(n_1,\ n_2)$ 分布的概率密度函数为：

$$\Psi(y)=\begin{cases}\dfrac{\Gamma\left(\frac{n_1+n_2}{2}\right)\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}}y^{\frac{n_1}{2}-1}}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)\left(1+\frac{n_1y}{n_2}\right)^{\frac{n_1+n_2}{2}}}, & y>0\\ 0, & \text{其他}\end{cases}$$

对于给定的正数 α，$0<\alpha<1$，称满足条件 $P\{F>F_\alpha(n_1,\ n_2)\}=\int_{F_\alpha(n_1,n_2)}^{\infty}\Psi(y)\mathrm{d}y=\alpha$的点 $F_\alpha(n_1,\ n_2)$ 为 $F(n_1,\ n_2)$ 分布的上 α 分位点，如图 1—4 所示。

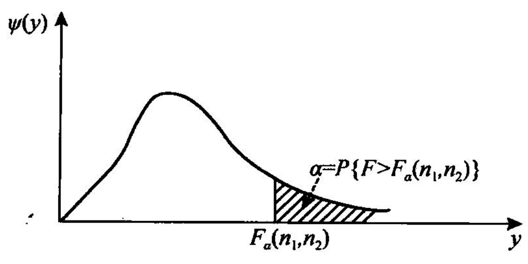


图 1—4　F 分布的上 α 分位点

4. 正态总体的样本均值和样本方差的分布

设 X_1，X_2，…，X_n 是来自总体 $N(\mu, \sigma^2)$ 的样本，$\overline{X}$，S^2 分别是样本均值和样本方差，则有

(1) $\overline{X}$ 与 S^2 相互独立。

(2) $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$。

(3) $\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$。

设 X_1，X_2，…，X_n 与 Y_1，Y_2，…，Y_n 分别是具有相同方差的正态总体 $N(\mu_1, \sigma^2)$，$N(\mu_2, \sigma^2)$的样本，且两样本相互独立。$\overline{X}$，$\overline{Y}$ 分别为两样本的样本均值，S_1^2，S_2^2 分别是两样本的样本方差，则有

$$\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t(n_1+n_2-2)$$

式中，$S_w=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$。

1.2.3 应用举例

例 1—2 已知随机变量 X 的一个样本及其观察值如表 1—8 所示。试计算样本均值、样本方差和样本标准差。

表 1—8　　随机变量 X 的观察值

序号	1	2	3	4	5	6	7	8	9	10
样本观察值	82.92	31.64	77.34	33.73	79.16	48.80	46.07	14.46	33.10	20.12

解 样本均值 $\overline{X}$ 为：

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i=46.73$$

样本方差 S^2 为：

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2=626.09$$

样本标准差 S 为：

$$S=\sqrt{S^2}=25.02$$

下面说明如何使用 Excel 中的函数来计算样本均值、样本方差和样本标准差。

在 Excel 工作表中建立计算模版如表 1—9 所示。

表 1—9 计算模版

	A	B	C	…
1	序号	样本观察值		
2	1	82.92		
3	2	31.64		
4	3	77.34		
5	4	33.73		
6	5	79.16		
7	6	48.80		
8	7	46.07		
9	8	14.46		
10	9	33.10		
11	10	20.12		
12				
13	样本均值	46.73		
14	样本方差	626.09		
15	样本标准差	25.02		

有关表 1—9 各单元格各公式及其功能的描述详见表 1—10。

表 1—10 有关表 1—9 各单元格中公式及其功能

单元格	公式	功能
B13	=average(b2:b11)	计算 $\overline{X}$
B14	=var(b2:b11)	计算 S^2
B15	=sqrt(b14)	计算 S

显然在数据量比较大时，使用人工方式计算均值、方差等计算量比较大，不如使用 Excel 计算来得快。

例 1—3 随机变量 X 服从正态分布 $N(40, 1.5^2)$，$\alpha=0.05$，计算 N 分布的上 α 分位点 z_α。

解 对标准正态分布 $N(0, 1)$ 而言，$z^0_{\alpha=0.05}=1.645$，因此对于一般正态分布 $N(40, 1.5^2)$，其上 α 分位点 z_α 为：

$$z_\alpha=z^0_{\alpha=0.05}\times 1.5+40=42.67$$

在 Excel 中内置函数 NORMINV 为计算一般正态分布函数的上 α 分位点 z_α。使用格式为：

$$z_\alpha=\text{NORMINV}(1-\alpha,\mu,\sigma)$$

对本例来说：

$$z_\alpha=\text{NORMINV}(1-0.05,40,1.5)=42.4673$$

函数 NORMINV(prob，μ，σ) 返回概率为 prob(=$1-\alpha$)、均值为 μ、方差为 σ^2 的正态分布的上 α 分位点 z_α。对应的相反函数为 NORMDIST，详见 Excel 中的帮助项。

表 1—11 给出了常见分布在 Excel 中对应的计算上 α 分位点的内置函数。