# 管理数学的统计应用

田竞和

东北财经大学出版社

# 前　　言

现今数理统计方法已广泛应用于生产经营管理领域，如管理中的市场调查、产品需求预测、生产控制、产品质量控制、劳动力分配、资金管理、成本控制、盈亏分析等预测和决策，统计方法已成为经营管理者必需掌握的基本工具之一。

由于在统计理论和方法中渗入了管理思想，并与之结合起来处理管理中各种问题，从而逐渐形成一门研究管理活动的理论和方法的新兴科学——管理统计学。

管理统计学同样包括描述统计和推断统计。有关描述统计内容，在一般统计学书籍中已有较详的论述，本书着重介绍统计推断方面的一些基础知识，如统计预测、控制和决策，并侧重于应用。

推断可以理解为由一些陈述前提到达一个结论的过程。统计推断是在一个样本的基础上对总体作出结论。概率论在统计推断中具有重要意义。企业管理在预测、决策中每逢遇到不确定因素时，总是运用概率论的理论和方法去研究、探索，寻出答案。

运筹学特别是线性规划在第二次世界大战后，在制订经济建设规划和管理决策中得到广泛应用，并取得满意的结果。本书将概率论和线性规划的原理作简单扼要、通俗易懂的论述，以弥补读者有关这方面知识的不足，其他用得普遍的如盈亏分析法、马尔柯夫链等等也作必要的介绍。

本书分上、下两篇，上篇介绍数学方法，帮助读者解决学习中的一些具体困难，以便有助于下篇统计方法的学习。

本书是一种自学用管理统计的辅助性参考书，内容力求深入浅

点，多举例题，方法详为演算，使读者学习后，初步掌握一些这方面的基本知识。运用这些知识，必然对现代科学管理和改革产生积极作用。

全书共七章，约30万字。参加本书编写的除田竞和外，第一、二章和第四章由胡旭微编写，第三章由张明、张君编写。全书由田竞和主编，并承佟哲晖、贾宏宇等教授审阅，在此致谢。

本书的编写是一种新的尝试，不当之处请读者指出，并帮助我们改进。

编　者

1992年秋于上海

# 目　　录

## 上篇　数学方法指导

## 下篇　统计方法介绍

# 上篇　数学方法指导

## 第一章　概率论浅说

### 第一节　随机现象

**一、随机现象和随机事件**

人类在实践活动中会遇到各种各样的现象（或事件），这些现象从它的性质来说，可以归纳成两类：一类是确定性现象（或称必然现象）；另一类是不确定性现象（或称随机现象）。确定性现象可以根据已知的事实，推论它将要发生的结果，不确定性现象具有多种可能的结果，事先不能确切断定它将发生哪种结果的现象，即各次试验的结果具有不确定性。但是随机现象却是大量存在的，而且也不是完全不可捉摸的。如在相同条件下，进行大量观察（试验），将各次观测的结果，通过归纳分析，也会发现各种结果出现的可能性是有确定的规律性的。这种在大量重复试验观测中所显示出来的固有规律性为随机现象的统计规律性，可以用数学方法进行研究。

随机现象所显现出来的各种结果可用数量表示，或是可以设法用数量表示，这就可用变量来描述随机现象的性质，这种变量叫做

随机变量。

在一定条件下，试验结果的不确定性，或者说是随机性，也称为偶然性。随机现象偶然性的背后，都具有某种必然性和规律性。概率论就是认识现象的规律性的一种数学方法。

对随机现象的观测，称为随机试验。随机试验的每一种表现或结果，称为随机事件。每次试验中都出现的事件，称为必然事件。在每次试验中都不出现的事件，称为不可能事件。随机事件如在产品质量抽样检验时，从一批产品中抽取若干件产品检查其质量情况，是一个随机现象。而“全部合格”、“有一件或二件不合格”就是两个随机事件。测量某个零件长度的误差大小，是一个随机现象，误差不超过±0.5毫米就是一个随机事件。

随机现象具有下列特点：

(1)随机性。每次试验究竟出现什么结果是不确定的，在试验前是无法确切断定的。

(2)试验的结果必须是可以观测的，也就是说它的结果是很明确的。

(3)随机试验必须假设可以在相同条件下重复进行。

## 二、随机变量

随机变量有两种性质（类型）：

(1)离散变量

离散变量之间都是以整数断开的，如人数、工厂数、机器台数等。它只能以计量的方法取得。仅能以0,1,2,…等数值表示。

例如：取100个标本，观测棉织坯布上的疵点数，则有100个观测值，是正整数。这100个观测值有相同的数值，也有不相同的数值。把它们归类，数值相同的归在一起，那末就可以得到经过整理后的有规则的资料，这种资料叫做分配数。

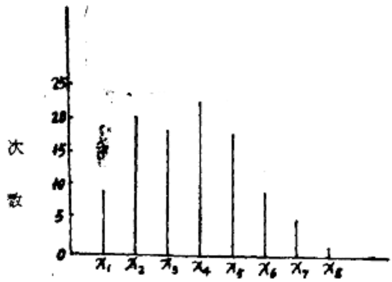
| 合计 | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | $x_8$ |
|---|---|---|---|---|---|---|---|---|
| 100 | 9 | 20 | 18 | 22 | 16 | 8 | 6 | 1 |

上述 $x_1$，$x_2$，…，$x_8$ 是按疵点数多少顺序排列，下一行是

各疵点数出现的次数（即重复的次数）。如$x_1$出现9次，$x_2$出现20次，等等。这些出现的次数叫做频数（或次数），记作$x_i$。所有变值的频数之和，等于样本容量（总数），用n表示。

各变量$x_i$的频数与总频数的比值，叫做频率。记作$f_i$。上例的频率分别是0.09，0.20，0.18，0.22，0.16，0.08，0.06，0.01。可见，$\sum f_i = 1$。

离散变量是各个孤立的数值，作图时为若干垂直线段，如图1—1所示。




点数

图 1—1

图中横轴代表疵点数，纵轴代表疵点数出现的次数。如第一条线疵点数$x_1$出现9次，线的长度为9，第二条线为20，等等。

（2）连续变量

连续变量是连续不断的，相邻两值之间可作无限分割。如棉纱强力（在一定长度和温湿度条件下，取样试验）为多少克。

例：100个样本试验的数据如下：

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 206 | 205 | 204 | 200 | 206 | 212 | 204 | 212 |
| 222 | 220 | 216 | 216 | 220 | 222 | 218 | 222 |
| 228 | 228 | 228 | 228 | 228 | 229 | 228 | 230 |
| 236 | 232 | 230 | 230 | 233 | 236 | 232 | 236 |
| 240 | 240 | 238 | 237 | 240 | 240 | 238 | 240 |
| 245 | 243 | 242 | 240 | 243 | 246 | 242 | 246 |
| 250 | 250 | 248 | 246 | 250 | 253 | 250 | 253 |
| 258 | 256 | 256 | 256 | 258 | 258 | 256 | 258 |
| 264 | 262 | 262 | 262 | 264 | 265 | 262 | 265 |
| 280 | 270 | 266 | 265 | 273 | 282 | 267 | |
| (176) | 230 | 246 | 265 | 222 | 240 | 260 | |
| 214 | 236 | 256 | 183 | 230 | 246 | 265 | |
| 222 | 240 | 260 | 214 | 237 | 256 | (292) | |

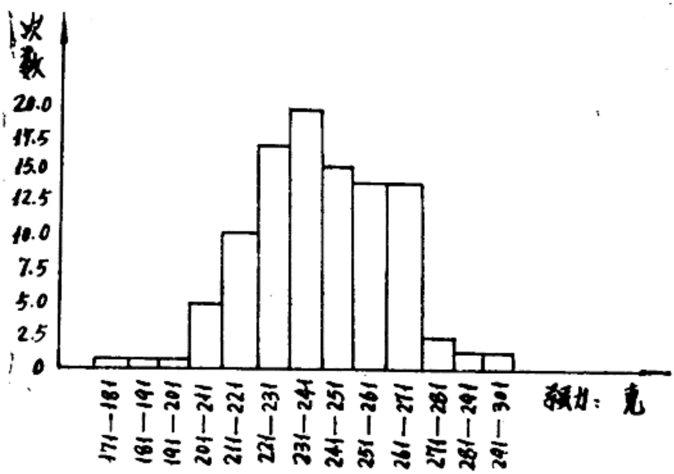



图 7—2

连续变量较分散，范围广，用分组方法整理。

上述数据最小为176，最大为292，如果分13组，每组组距为10，整理如表1—1，并作图如图1—2所示。

表1—1

| 分　组 | 划　记 | 频　数 | 频　率 |
|---|---|---|---|
| 171～181 | 一 | 1 | 0.01 |
| 181～191 | 一 | 1 | 0.01 |
| 191～201 | 一 | 1 | 0.01 |
| 201～211 | 正 | 5 | 0.05 |
| 211～221 | 正𠄌 | 9 | 0.09 |
| 221～231 | 正正正丅 | 17 | 0.17 |
| 231～241 | 正正正𠄌 | 19 | 0.19 |
| 241～251 | 正正正 | 15 | 0.15 |
| 251～261 | 正正正 | 14 | 0.14 |
| 261～271 | 正正𠄌 | 14 | 0.14 |
| 271～281 | 丅 | 2 | 0.02 |
| 281～291 | 一 | 1 | 0.01 |
| 291～301 | 一 | 1 | 0.01 |
| | | 100 | 1.00 |

### 三、随机变量的数学特征

随机变量 x 的概率分布，完整地描述了随机变量的统计特征，但在实际情况中，有时并不要求我们全面考察随机变量的变化情况，而是只需要知道随机变量的某些数字特征就可以了。在根据概率做决策时，是根据随机变量的平均值选择方案，有时还要考虑随机变量的分布与平均值的偏离程度来考虑方案的风险程度。随机变量的平均值及偏离程度描述了随机变量的特征。

例如，我们要知道灯泡的质量，知道平均寿命以及各灯泡的寿命与平均寿命的偏离程度就可以了。平均寿命高，且各灯泡的寿命与平均寿命的偏离程度小，灯泡质量就好。不必细致地了解其详细的概率特征。这样工作就省事多了。

又如，某企业甲车间有同种类型的机床4台，这4台机床不完全是

同时开动。因此，任一时刻工作的机床数是一个随机变量。为了确切知道车间的电力负荷，需要知道同时工作着的机床平均数。为此，进行了25次观察，结果如表1—2，表中资料说明一台机床工作只有一次；2台机床工作有4次；3台机床同时工作有12次，等等。

表1—2

| 工作机床数 | 频　数 | 频　率 |
|---|---|---|
| 0 | 0 | 0 |
| 1 | 1 | 0.0417 |
| 2 | 4 | 0.1667 |
| 3 | 12 | 0.5000 |
| 4 | 7 | 0.2916 |

工作机床总数 $= 1\times1+2\times4+3\times12+4\times7=73$

同时工作的机床的平均数：$\dfrac{73}{24}=3.04$（3台）

根据随机变量的可能值与其相应的频率的乘积之和，也可得到同样结果：

$1\times0.0417+2\times0.3334+3\times0.5000+4\times0.2916=3.04$

此数即为随机变量的数学期望值，记为：

$$E(x)=\Sigma xP$$

式中：$E(x)$为随机变量的数学期望值，$x$为频数，$P$为频率。

随机变量的数学期望值表示随机变量的“平均值”，但在实际中，往往需要研究分布的集中或分散程度。如考查一批零件的长度，只知道它的平均值达到标准还是不够的，如果长度参差不齐，长短不一，我们不能认为这批零件是适用的。因为与标准短的尺度就不适用了。故还要取其数学期望的偏离程度，即方差。随机变量的方差为：

$$D(x)=E\{[x-E(x)]^2\}$$ D(x)为方差

也可按下式计算：$D(x)=E(x^2)-[E(x)]^2$

方差是用来描述随机变量取值的分散程度，方差数值大，说明随机变量分布较为分散；方差数值小，说明随机变量分布较为集中，且集中在平均数的周围。

## 第二节　概率和概率的运算

**一、概率**

在解释概率之前，先解释频率。因为频率与概率有相同的特点，但两者却是两个不同的概念。

（1）频率

频率是指在n次试验中，设随机事件a出现了$n_a$次，这个比值$\frac{n_a}{n}$为n次试验中事件a出现的频率。记作$f_n$（a），即：$f_n(a)=\frac{n_a}{n}$。

它有下列一些性质：

①随机事件a在n次试验中，出现的频率总是介于零与1之间的一个数，即$0\leqslant f_n(a)\leqslant 1$。

②如果a是必然事件，则事件a在n次试验中出现的次数$n_a=n$，所以必然事件出现的频率等于1，即$f_n(U)=1$（必然事件常用U表示）。如果a是不可能事件，它的出现的频率等于0，即$f_n(V)=0$（不可能事件常用V表示）。

经验表明：只要试验在相同条件下进行，随机事件出现的频率逐渐稳定在某个常数P，这个P是事件本身的一种属性。

如果试验次数n的增大，事件a的频率$\{f_n(a)=\frac{n_a}{n}\}$在区间（0，1）上的某个数值p附近摆动，那么事件a的概率为P（a）＝p。

（2）概率

概率是某种现象（事物）发生某一状态或结果的可能性大小的度量，以0到1之间的分数或小数表示。概率为零表示这一现象不可能出现概率为1，表示这种现象必然出现，概率为0.5表示这种现

象出现或不出现的可能性各半。

例如：抛一枚硬币，一面是国徽，一面是币值，只有两面。要么出现国徽面，要么出现币值面。概率各为$\frac{1}{2}$即0.5，用数学公式表示为：

出现国徽面的概率记作P(H)

$$P(H)=\frac{1}{1+1}=\frac{1}{2}\text{或}0.5$$

这个公式的子项表示出现国徽的结果是1（种）；母项表示在一次抛掷中，可能出现一个国徽面，也可能出现一个币值面，全部可能结果为1＋1＝2。

在概率论的书中，往往一开始就用掷骰子的例子。一个骰子有6面，从1到6计六个数字。如掷一次出现“3”点的概率为$P(3)=\frac{1}{1+1+1+1+1+1}=\frac{1}{6}$，无论那一个点子，均为1/6。

又如：在一批100个零件中，合格品有98个，不合格品有2个，那末从100个零件中抽取一个，可能抽到的是合格品，也可能抽到的是不合格品，但两者的概率是不相同的。抽到合格品的概率为$\frac{98}{100}$，抽到不合格品的概率为$\frac{2}{100}$。

随机事件是在一次试验中，可能发生也可能不发生，即具有偶然性。但是在大量重复试验中，可以发现它具有内在的规律性，即它发生的可能性大小是可以度量的。随机事件的概率就是用来说明随机事件发生的可能性大小的一个概念，它是概率论中最基本的概念之一。

概率实质上是频率的数学抽象，频率所具有的特性，概率同样有。

**二、概率的加法定理和乘法定理**

不相容事件的加法法则

如果事件A与事件B不能在一次试验中同时发生，则事件A和事件B互不相容。则：

$P\{A+B\}=P\{A\}+P\{B\}$

例：某工厂产品的质量分为一等品、二等品和等外品。在正常条件下，出现二等品的概率为5%，出现等外品的概率为2%，其余都是一等品，则出现非一等品的概率为：

$P\{A+B\}=P\{A\}+P\{B\}=5\%+2\%=7\%$

如果有n个互不相容事件$A_i$（$i=1,2,\cdots,n$）则有：

$$P\left\{\sum_{i=1}^{n}A_i\right\}=\sum_{i=1}^{n}P\{A_i\}$$

概率的乘法法则

两个或两个以上独立事件同时发生或连续发生的概率等于它们无条件概率的乘积。

所谓独立事件例如掷一枚硬币，P（正）=0.5，P（反）=0.5，出现正面与反面的概率都是0.5，对某一次抛掷来说，都是正确的，每次抛掷都是独立的事件。

$P\{AB\}=P\{A\}\times P\{B\}$

这个公式是以事件A和事件B相互独立作为条件的。

$P\{AB\}$为事件A和事件B同时发生的概率，也称联合概率。$P\{A\}$与$P\{B\}$为事件A（或B）出现的无条件概率。如连续抛掷硬币二次，都出现正面(A)的概率为：

$P\{H_1H_2\}=P\{H_1\}\times P\{H_2\}=0.5\times0.5=0.25$

上面是两个独立事件的无条件概率。此外还有条件概率，如，$P\{B|A\}$，读作在事件A发生的条件下，事件B的概率，对于独立事件是一事件的概率不受其他事件发生与否的影响，所以在事件A已经出现的条件下，事件B的概率就是B的单一概率。如：

$P\{B|A\}=P\{B\}$；$P\{AB\}=P\{A\}P\{B|A\}$；

$P\{AB\}=P\{B\}P\{A|B\}$

A，B是相互独立的，事件B的出现不受事件A出现与否的影响。

同样：$P\{A|B\}=P\{A\}$。

上述独立事件相当于抽样中的返回抽样，如果说是不返回抽样，情况就不是这样。

例如：一个袋中有20个产品，其中18个是合格品，2个是不合格品。甲从中抽取一个产品，恰好是不合格品，为事件A；乙从中抽取一个产品，也恰好为不合格品，为事件B。

$$P\{A\}=\frac{2}{20}=\frac{1}{10}$$

当甲取后不放回，乙再取用。此时产品只有19个。不合格品只有一个，故$P\{A|B\}=\frac{1}{19}\left(\neq P\{A\}\right)$，$P\{B|A\}=\frac{1}{19}\left(\neq P\{B\}\right)$，它说明A和B互相影响。

$$P\{AB\}=P\{B\}P\{A|B\};P\{AB\}=P\{A\}P\{B|A\}$$

例：某一随机事件，如机床中的某一零件，它的质量要求有二个，零件的厚度和零件的直径。我们可以把零件厚度合格记为事件A，直径合格记为B。可应用联合概率求解；即$P(AB)=P(A)\cdot P(B|A)$或$P(AB)=P(B)P(A|B)$。

例：某进出口公司拟对外推销一种新产品，预测市场销售情况有三种：畅销、中等、滞销。其概率及利润额如表1—3所示。

表1—3

| 预测结果Q | 概率P(A) | 利润（万元） |
|---|---|---|
| 畅销 $Q_1$ | 0.26 | 45 |
| 中等 $Q_2$ | 0.34 | 3 |
| 滞销 $Q_3$ | 0.40 | −18 |

由于畅销概率仅26%，该公司要承担的风险较大。公司考虑欲委托一咨询公司作一次市场调查，费用需要2.5万元。从历史资料看，同类产品的销售概率见下页表。

表中$S_i$是指几种实际销售情况，$P_i$是几种预测销售情况，表中数字为预测结果为$P_i$，实际销售为$S_i$时的概率。例如，$P_iS_i$是指预