高等学校教材

工业技术应用数理統計学

上册

（修訂本）

周华章编

人民教育出版社

高等学校教材

工业技术应用数理統計学

上　冊

（修訂本）

周华章編

人民教育出版社

本书对数理統計学在工业技术上的应用作了比較詳細的闡述，分上下两册出版。此次出版的上册是第二版，比之第一版有較大修改。若干章节已完全改写，特别是第四、七、八、九、十、十一等章。第二版先从实驗数据的分組制图入手，引入常用的統計特征数的定义及其計算方法（第一至三章），再介紹概率論的基本概念、几种基本的分布及推导各种随机变量之函数的分布的方法（第四至七章），然后在此基础上介紹大子样与小子样檢驗統計假設及估計母体参数的方法（第八至十二章）；叙述时尽量举例說明。本书可供高等工业学校作教学参考用书，亦可供工业部門研究人員、工厂技术檢驗人員作参考。

工业技术应用数理統計学

上　册

（修訂本）

周　华　章　編

北京市书刊出版业营业許可证出字第2号

人民教育出版社出版（北京景山东街）

人民教育印刷厂印装

新华书店北京发行所发行

各地新华书店經售

統一书号K15010・901　开本 $850\times1168\ ^{1}/_{32}$　印张 $8^{10}/_{16}$　插頁1

字数 200,000　印数 9,001—16,000　定价（6）¥0.85

1960年7月第1版　1963年11月第2版　1964年9月北京第4次印刷

說　　明

本书上册第二版，比第一版已有多处經过修改。书中所用符号自第八章引入統計推断理論以后，比第一版亦有多处更动。为便于讀者查閱，特作統一說明如下。今后下册出版，亦将遵用以下符号，希已讀上册第一版的讀者注意。

(1)凡遇只有一个随机变量的場合，概用 ξ 表示随机变量。ξ 的观测值用 x 表示。譬如 $x_1, x_2, \cdots$ 分别表示 ξ 的第 $1, 2, \cdots$ 个观测值。

(2) ξ 的期望值 $E(\xi)$ 用 $\bar{u}$ 表示；方根差用 σ_ξ（或簡写作 σ）表示，即 ξ 的方差 $\mathrm{Var}(\xi)=\sigma_\xi^2$。

(3) ξ 的子样平均数作为一个随机变量，用 $\bar{\xi}$ 表示。$\bar{\xi}$ 的观测值用 $\bar{x}$ 表示。譬如 $\bar{x}_1, \bar{x}_2, \cdots$ 分别表示第 $1, 2, \cdots$ 个子样观测值的平均数。

(4) $\bar{\xi}$ 的期望值 $E(\bar{\xi})$ 因等于 $E(\xi)$，故仍用 $\bar{u}$ 表示。$\bar{\xi}$ 的方根差用 $\sigma_{\bar{\xi}}$，即 $\bar{\xi}$ 的方差 $\mathrm{Var}(\bar{\xi})=\sigma_{\bar{\xi}}^2$。

(5) ξ 的子样方根差作为一个随机变量，用 σ_x 表示。σ_x 的观测值，亦即由 ξ 的子样观测值 $x_1, x_2, \cdots, x_n$ 算出的方根差，用 s 表示。譬如 $s_1, s_2, \cdots$ 分别表示第 $1, 2, \cdots$ 个子样观测值的方根差。

(6) σ_x 的期望值为 $E(\sigma_x)$。σ_x 的方根差用 σ_{σ_x} 表示，即 σ_x 的方差 $\mathrm{Var}(\sigma_x)=\sigma_{\sigma_x}^2$。

(7)遇有二个以上随机变量出現的場合，用 $\xi, \eta, \zeta\cdots$ 等表示随机变量。其观测值分别用 $x, y, z, \cdots$ 表示。

(8) $\xi, \eta, \zeta, \cdots$ 的期望值 $E(\xi)$, $E(\eta)$, $E(\zeta)$ $\cdots$ 分别用 $\bar{u}, \bar{v}, \bar{w}, \cdots$ 表示。$\xi, \eta, \zeta, \cdots$ 的方根差分别以 $\sigma_\xi, \sigma_\eta, \sigma_\zeta \cdots$ 表示。

編 者

内容目录

第一章 引論

§1.1 数理統計学与工业技术

在金工車間里，有一名工人，在一台車床上，用同一种原材料，按照同一种設計，继續不断加工同一种机器零件，例如机器上的圓軸。一位质量檢查員对每一根軸測量直徑一次，发現測量結果并不完全相同，而有一定的差异。也就是說，測量結果表明軸的直徑并不必然等于某一数字，而表現出偶然的性质。不仅如此，这位檢查員对同一根軸，反复測量直徑多次，发現結果也不完全相同，而仍有一定的差异，虽然，后者的差异比之不同的軸之間的差异是較小些的。換句話說，同一根軸的測量結果仍不必然相同，而还是表現出偶然性质。工厂对于这两种差异的性质需要深入研究。譬如說，要弄清楚这差异是制造过程中造成的，还是质量檢查員測量直徑时的測量誤差造成的？工程人員必須找出这种差异发生的原因，或是說这种偶然性的来源，揭示其中的規律，才能較精确地測定軸的直徑。

工业技术数理統計学是数学的一个分枝。它的內容就是分析上述一类問題的。上述問題中，假定零件是由同一名工人，在同一台車床上，用同一种原材料，在同一天里加工而成的。如果工人，車床，原材料，日期再不固定，那么所加工的軸的直徑，可能有更大的差异。此时分析差异造成的原因，正确地了解零件直徑的情况，就更需要数理統計学的知識。因此，数理統計学在近代工业技术应用中占有很重要的地位。但讀者不可誤认为分析差异的原因是数理統計学所研究的唯一內容，共实它不过是数理統計

学所研究的一个問題。关于数理統計学的主要內容下一节将加以闡述。

数理統計学旣应用数学的方法，又应用統計学的方法，所以它涉及的領域非常广泛。数理統計学应用数学方法的地方，在本书中随处可見，这里毋須申述；至于它是怎样应用統計学方法的，現略述如下：

（一）分組法：——分組法是統計学上的特有方法，数理統計学上也用。例如有这么一家工厂：这家工厂的車床每天加工了很多只鉚釘，但每加工一只鉚釘任意換一次材料，或每加工一只鉚釘任意換一名車工。在这种情况下，虽然鉚釘仍旧大量出現，可是它們每一只与另一只的生产条件不同，如果运用数理統計方法于这台車床所加工的鉚釘，那就难于发現甚么有意义的規律，所以在整理、計算或分析之前一定要按照事物特点将所收集的資料加以分組，再来說明。

（二）大量观察法：統計学主要是研究大量現象，数理統計方法也是研究大量偶然事件的某一类客观規律的科学方法。关于它的性质在下面将有所討論。譬如一台車床一天里加工的鉚釘，一个細紗車間一天里紡成的細紗，都是大量观察的对象。假如一台車床一天里只加工了一只鉚釘，那么要了解这台車床一天里加工的鉚釘的情况，就不需要数理統計方法。

（三）綜合指标法：統計学上分析事物、現象或发展过程时，常应用指标，数理統計学亦然。譬如在鉚釘口徑，皮带拉力的例子里，常計算平均口徑和平均拉力；再如一家紡織厂生产的棉紗，按其粗細均勻度可以分为优級，一級，二級等，生产优級紗天数所占的百分比，生产一級紗天数所占百分比等，和平均口徑或拉力等都是綜合指标。

§1.2 两个概念

根据上节的討論，可以总結起来說，数理統計学是借助統計学和数学中的方法，研究偶然事件的某一类客观規律的科学。这里有两个概念須交代清楚。第一个是偶然事件的性质，第二个是"某一类客观規律"究竟指怎样的一类客观規律。

第一，在概率論和数理統計学中，称一次独立試驗的結果为一"事件"。譬如測量軸的直徑一次，得到一个測量数字，这个数字是測量的結果，我們称之为一个"事件"。如果在一定条件下，某事件必然出現，这事件称为必然事件。如果在一定条件下，某事件有时出現，有时不出現，那么这事件便称为偶然事件。譬如說有一名工人在一台車床上，用同一种原材料，按照同一种設計，继續不断地加工机器上用的圓軸，如果每次測量这些軸的直徑，都不变地得到同一数字，那么說这測量結果是在上述工人，車床，原材料，設計及測量工作本身等条件下出現的必然事件。反之，如果在以上条件下，測量結果幷非不变的同一数字，則称这种結果为**偶然事件**。

概率論和数理統計学既然研究偶然事件，那么弄清楚偶然事件的性质应該是十分必要的。为了說明方便，举一个比加工圓軸更簡单的例子。学过普通物理学的人都知道一条粗細均匀的金屬导綫的电阻值 R 与其长度 l 成正比，而与其截面面积 S 成反比，比例常数 k 称为电阻系数，是与导綫溫度有关的一个数，即 $R=k\cdot\frac{l}{S}$。但在实驗室中进行电阻測量时，根据經驗，往往是在不同的日子，不同的环境里，即使溫度相同，同样一条导綫的电阻的观測值幷不是固定不变，而是表現出一定差异的。譬如按公式計算，电阻值等于1欧姆，但观測值有时大于1欧姆，有时小于1欧姆。这样，好像按公式算出的电阻值幷非必然出現的必然事件，实际出現的电阻的观測值是带有偶然性的偶然事件。这样說来，岂不是电阻值

与长度成正比，与截面面积成反比的規律，就沒有眞实性了嗎？事实究竟是怎样？对于一条导綫的电阻来說，当溫度固定时，导綫的长度和截面面积是决定电阻大小的决定性因素，但除此之外，对电阻有影响的还有其他許多因素，如空气中的放电現象，空气的湿度，观测电阻时的观测誤差，电阻仪本身的測量誤差等等。由于这許多因素的存在，使每一次观测得到的电阻值并不一定等于按規律所算出之値 1 欧姆，而有时大于，有时小于 1 欧姆，也就是在 1 欧姆左右摆动不定，因而表現出某种偶然性。这种偶然性的存在，并不說明計算电阻的規律不正确，因为这条規律正确地反映了导綫电阻怎样受它决定性因素的制約的客观情况，所以是一条有价值的客观的动态規律。但是观測值搖摆于按动态規律所算出的值左右这件事实，也提供了一个新問題，让概率論和数理統計学来处理，那就是概率規律性或統計規律性的問題。因为世界上任何現象都和其他現象有多种多样的联系，所以动态規律性总是和概率規律性或統計規律性交織在一起。因此研究概率論和数理統計学就有它的意义了。

类似地，加工机器用的圓軸时，工人的技术水平，車床的性能，原材料的性能，一定的設計等等是决定軸的直徑精密度的决定性因素。这些因素即使能够完全固定不变，軸的直徑的观测值并不能完全相等，而仍有一定的差异，因为除了以上列举的諸因素外，还有很多其他因素对它发生影响。

現在我們要回过头来看一看，为甚么概率論和数理統計学的研究必須借助分組法，大量观察法和綜合指标法。这是不难理解的。首先，概率論和数理統計学所要研究的概率規律性是一种数量規律性。所以必須借助統計学里的数量指标。其次，概率規律性是指大量偶然現象，或偶然事件的規律性。如果只观測电阻一次，那么这一次的电阻值可能大于 1 欧姆，也可能小于 1 欧姆，找

不出甚么規律，但在百次，千次……的观测中，观测出有多少次大于1欧姆，多少次小于1欧姆，則通过这样的观察就可寻出某一規律，所以被研究的对象必須是大量出現的事件。再次，分組法是为保证主要条件具有某种程度的不变性所必需的性质。如果导綫的长度和截面面积一經改变，譬如測量电阻几次后，換上了另一根导綫，那么这批电阻的观測值就不是圍繞着1欧姆而摆动的数字，从而寻求規律的工作就应該从这两批观測值，而不是从一批观測值着手。对于一家工厂加工鉚釘来說，如果更換工人、原材料，等等，那么如果不从这些主要因素变化的影响分批研究，只从其加工的鉚釘口徑考虑，是难于看出甚么規律来的。

以上討論了偶然事件的性质。以下还要談談第二个概念，即“某一类客观規律”究竟指甚么。經过以上的討論，現在可以說概率論是研究大量偶然事件的概率規律性的科学。但数理統計学和概率論有何不同呢？正如上面提到，它們之間很难找出截然的分界綫。但大体說来，数理統計学是以概率論为基础，对統計資料进行分析、研究，导出其概率規律性的科学，更具体些說，是研究从一定母体中所抽子样的某些特征数字所表現的概率規律性的科学，（关于母体和子样这些名詞下面有較詳細的解釋，現在只要理解为全体与部分就够了。）譬如一台車床加工的鉚釘为数极多，我們从其中抽取10只，測量其口徑，这10只鉚釘的平均口徑可能大于，也可能小于这种鉚釘的标准口徑。这里可以发現一定的概率規律性。正是这一类概率規律性基本上是数理統計学所要研究的。我們强調基本两字，因为在少数情形中数理統計学也可以研究母体本身。

§1.3 数理統計学的內容

数理統計学应用范圍很广，它可以涉及各門科学及各种工业技术的領域，已如上述。它所研究的問題的性质也是多种多样的。

現代工业技术和各門科学的領域是在不断地扩展着，新的数理統計問題也随时涌現出来，要求解决，所以数理統計学的面貌也在日新月异地变化着。本书并不能遍及数理統計学的各个部門，只是結合工业技术方面的应用，討論若干基本原理和几种常用的方法。本书所說的工业技术应用，并不专門对某一种工业，而是对机器制造，土木建筑，化学，紡織，冶金等各种工业的应用。

經过上面的簡短說明以后，我們来列举几个工业技术中主要的数理統計問題。

(I)表达一件事物的性質

正如§1.1中所說，由同一名工人在同一台車床上加工的許多件同品种机器零件，如机器用的軸或汽缸上的螺栓，它們的直徑并不完全相同，而有一定的差异。由若干名工人在不同的車床上加工的同一种零件，它們的直徑更不能完全相同，而常表現更大的差异。一台或几台車床加工的零件为数甚多，工厂檢查員了解它們的直徑的情况时，如果一只一只地列举出它們的直徑，則工作旣极繁瑣，而对整个情况仍得不到一个概括的全面的了解，因此需要用少数几个数字来表达出整体的情况。如何去找少数几个数字来表达整体的情况，是数理統計学中經常遇到的第一类問題。这种少数的几个数字在数理統計学中称为特征数。

(II)比較两件事物的差异

例如比較两台机床的加工精度时，即使是同一型号的两台机床，其加工精度也会有所不同。而且，同一台机床在不同时間內加工出来的零件的精度也会有所差异，譬如昨天的加工情况和今天的加工情况就可能不一样。如何用最簡便的方法来进行分析、研究及比較昨天和今天的加工精度，是一个很重要的問題。

比較两件事物有沒有差异，当事物的性質很单純时，原为很簡单的事。例如知道了两只螺釘的直徑，即可判定那一只尺寸較大。

但在比較两批为数众多的螺釘的平均直徑时，若各批中螺釘直徑的变化較大，同时在測量直徑时还有測量誤差，那么問題就变得比較复杂了。数理統計学中分析这类問題的方法是对两件事物的性质先給以某种假設，例如假設两批螺釘的平均直徑是相等的，然后从每一批中抽取若干只螺釘，測量其直徑。如果被測的螺釘直徑的某一特征数(例如被測螺釘直徑的平均数，中位数，极差等，这些名詞下面都有定义)的差异超过一定程度时，即认为所作假設不可信。反之，如果差异不超过一定程度，則认为所作假設是可信的，即可相信两批螺釘的平均直徑是相等的。

(Ⅲ)分析影响事物变化的因素

如果两批螺釘的加工精度是有差异的，那么我們自然要問，是那些因素引起了这种差异。引起差异的因素可能是原材料的不同，可能是加工处理方法的不同，也有可能是測量螺釘直徑时引起的測量誤差。这些因素可以单独引起差异，也可以两个或三个因素合在一起，引起差异。用数理統計学的分析方法，可以对这个問題作出結論，并能計算出各个因素所产生的影响的大小。

(Ⅳ)一件事物的两种性质之間的相互关系，或环境对于事物性质的影响

在工业技术中常遇到下面的一类問題：工程师需要确定机器零件的加工精度与刀具磨損之間，青磚的抗压强度与抗折强度之間，紡紗車間的溫、湿度与細紗断头率之間，是否存在着某种关系。总起来說，这一类問題就是要研究一件工业事物的两种性质之間可能存在着的相互关系，或这件事物的周圍环境对它的性质的影响。这一类問題在数理統計学中称为“相关”問題。談到相关二字，我們很自然地联想到一些在中学时代就熟悉的問題，譬如一定质量的气体，当溫度固定不变时，其容积与压力之間存在着的相互关系。实际上物理学中气体容积与压力之間的“相关”与数理統計

学中的“相关”，仅有准确程度上的差别，并沒有本质上的差别。这就是說，物理学中气体的容积与压力的关系比較确定，其关系图为一单純光滑的曲綫，但在通常数理統計学处理的相关問題中，各个变量之間的关系，不如物理問題中之光滑准确，但其基本性质却沒有甚么不同。

(V)研究取样与試驗的方法

要求知道一件事物的性质，譬如一批螺釘的直徑，我們不能对全部螺釘的直徑一一进行測量，而只能从其中抽取一部分来进行測量和分析。但所抽取的这一部分螺釘的直徑，須能代表全部螺釘的直徑。現在問題就是要决定，这一部分螺釘应該怎样从全部中抽取出来？应該抽取多少只？这就是抽样方法的問題。再如，一名质量檢查員用一只螺旋測徑仪測量同一件机器零件的直徑，每測一次，所得数値与前一次所得的不能完全相同。那么究竟对同一件零件，应測量多少次，才能对它的眞正口徑有較准确的了解呢？这就是試驗方法中的一个問題，其他有关試驗方法的問題还有很多，下面将有机会談到。

上面列举的五类問題，是工业技术应用中最常遇到的問題，但并不包括数理統計学的全部內容。正如上文說过，現代工业日新月异，新的数理統計問題不断地涌現出来，无有止境。然而人們自然要問，問題虽多，有沒有一个綫索把它們貫串起来呢？这也等于說，以上列举的五类，以及其他許多問題，是不是許多孤立的問題？抑或有一定的联系，而构成一个有机的整体？要回答这問題，我們先要在§1.4里闡明甚么是母体、个体和子样，然后在§1.5里再回过头来討論它。

§1.4 母体、个体与子样

母体、个体与子样，是数理統計学中常用的名詞。初学的人必