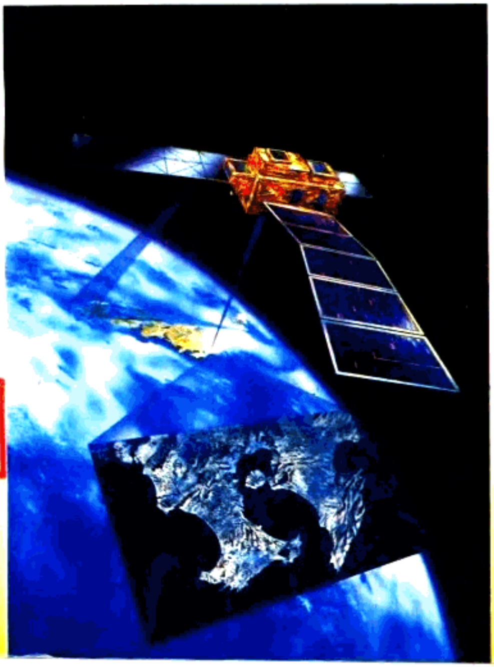
飞行器试验统计学



★张金槐 蔡 洪 编著★

技大学出版社★

前　　言

在学习了一般的概率论和数理统计的理论之后，我们又进入了飞行器试验统计学课程的学习。这是由于飞行器从研制、试验直至定型的整个过程中，统计学方法已越来越显示出它的重要性，而且在处理的方法上，飞行器试验统计分析有它的特殊性，涌现出不少需要研究的新问题。因此为了适应当前的需要，在专业基础理论上打下一个坚实的基础，我们新编了这本教科书。

我们将对在飞行器试验中常被运用的统计问题作比较系统、深入的论述。所论问题具有实际背景，能适应部队设计、研究和试验单位的需要。同时，注意思考方法，培养学员的思维能力和独立工作的能力。

这个教材是在原《飞行器试验统计学》(上、下册) 讲义的基础上重新编写的。我们力求做到内容精，反映近代成果且注重应用。书中不少理论和方法来源于我们的科研成果。

阅读本书时，必须具备高等数学、线性代数、概率论与数理统计、弹道导弹的基本知识等。本书适用于试验分析专门化及邻近专业的本科生。也可作为从事飞行器试验结果分析的工程技术人员的参考书。

编著者

1995 年 1 月

目　录

7 滤波的工程实现与自适应技术

8 Bayes 统计分析方法

9 仿真试验及其分析

10 飞行试验结果分析

附录 A1 向量代数

附录 A2 矩阵基础

1 引　　论

1.1 飞行器试验统计学研究的对象、特点及主要问题

顾名思义，飞行器试验统计学研究的对象是飞行器在试验（地面试验或飞行试验）中的随机现象。研究的目的则是要揭示出这种随机现象的统计规律性。这种统计规律性建立在观测信息的基础之上。读者已经学习了经典统计中的抽样分布理论、点估计及置信估计方法、假设检验方法、回归分析等内容。它们对于飞行器试验结果分析来说，都是基础性的知识，是非常有用的。然而，对飞行器试验来说，也还有它的特点。比如，飞行器的各系统技术复杂，而且价值昂贵，研制与试验的周期较长，且每批次的试验数少。而飞行试验中的弹道（轨道）特性分析必须结合飞行弹道（轨道）进行。因此，本课程的学习不同于一般的数理统计。除了必须掌握基本的理论以外，必须注重实际，结合专业中工程实践的问题来学习。就统计推断的思想方法来说，与经典统计也是不尽相同的。例如，在某些试验场合，不能如常规的统计处理那样，可以任意抽样。因此，人们习惯运用的大样本理论（中心极限定理、大数定理等）就受到限制。又如在飞行试验中，我

们运用光学、雷达等测量设备，对飞行器进行观测，得到的观测量是随时间而变化的量，而这种观测，在一次飞行试验中是不能再现的，即是说不能进行重复测量。这样一来，常规的统计处理，遇到了困难。由此可知，在飞行器试验统计学中，常用的统计方法需要进一步发展，而飞行器试验中提出的新问题，又要求我们作深入一步的研究。我们编写的这本书就是为了这个目的。

就我国飞行器试验统计分析来说，大致有下列十个方面的工作：

1、飞行器各分系统试验结果分析。这种分析包括地面试验、飞行试验中特性参数的分析。例如，对于弹道式导弹来说，包括动力装置系统试验结果分析，如发动机推力、液体火箭发动机的秒耗量、比推力、燃烧室压力等的试验结果分析；制导系统误差分析，如惯性制导系统中的仪表误差分离；弹体结构强度分析；头部（烧腐）试验；分离机构试验中的结果分析等等。如果将地面试验结果应用于高空，那末必须考虑相关分析。

2、飞行过程中，弹道（轨道）运动参数的估计。在导弹飞行试验中，弹道运动参数（如位置、速度）的估计及其分析是必需的。这种评估，对于导弹飞行性能的分析是至关重要的。飞行试验中对导弹的跟踪、预报、飞行中的安全控制等就是属于这种工作。而飞行试验后的重建弹道也是一项经常性的工作。

3、导航和全球卫星定位系统（GPS）中的运动参数的实时估计（所谓滤波问题）。

4、故障检测。主要是运用统计学中的检验理论，发现试验中的故障，以便于研究设计部门改进飞行器的性能。如试验中安全控制故障的发现等，是试验场地的一项重要的工作。

5、飞行器定型试射中的精确度和密集度评估、最大射程评定、可靠性评定等工作。

武器系统的定型，包括各种不同类型的飞行器，从常规飞行

器至战略武器的定型（鉴定）都是必经之途。目前已成为一个相当热门的课题。

6、*弹道（轨道）仿真研究*。仿真技术，是国内外广泛重视的问题。不论是数字仿真，或是半实物、实物仿真，对于飞行器的研制和性能分析，仿真是一种重要的工具。例如进行弹体和控制回路的设计分析、优选参数、进行全弹道仿真试验等等。一般的仿真过程包括建立模型、录取仿真参数及干扰源的统计特性和计算。仿真结果分析，如模型的检测、参数精度分析、仿真的可信性分析等，统计分析方法是一个基本的有力的工具。

7、*测控设备的精度鉴定*。对于测控设备来说，在出厂时都有产品合格的履历表，其中记载着设备的精度指标。但是飞行器试验时的外场条件与工厂不同，因此测控设备的精度未必与记载的精度一致。为此，必须在现场试验条件下作出合理的精度评定。这个工作，是作好试验结果分析的重要前提。

8、*多种测量信息下的数据融合问题*。在飞行器现场试验过程中，运用了多种观测设备对飞行器进行观测。例如光学测量、雷达测量、遥测等。如何综合利用这些观测信息，以便对于飞行器的性能参数进行检测、估计，这是一项重要的工作。国外靶场试验中运用的所谓 EMBET（Error Modeling Ballastic Estimation Technique，运用误差模型的最佳弹道估算）方法，就是属于这类研究工作。

9、*小子样统计推断方法研究*。这是由于飞行器试验的特点所决定的。我们已经指出，飞行器的试验不能大量地进行。因此，统计推断方法必须与这种特点相适应。我们在本书中论述的 Bayes 方法，就是一种考虑小子样场合的统计分析的有效途径。

10、*试验法研究*。这里指的是制订飞行试验大纲中的试验法研究。它包括试验中达到的要求、项目、试验程序、批次发射数、定型状态下的最佳发射数以及定型验收方案。对于试验分析的技

术人员来说，研究最节省的试验途径，节省试验数，又能满足评估试验项目中的有关精度的要求。这个工作可以说是当务之急。对于试验决策人员和总体研究、设计部门都是迫切需要的。

本书不可能对上述问题一一进行详细讨论，特别是具体的技术细节问题。我们将着重讨论上述问题中广为运用的试验统计分析的理论和方法。一些具体应用，将以示例的形式给出。

下面，对于本课程所论述的内容作一点说明，使读者能在总体上有个一般的了解，并在学习中了解彼此的关系及其地位。

1.2 内容梗概

1.2.1 线性模型的参数估计问题

先从最小二乘方（Least Square，简记 LS）方法说起。这个方法在实验结果曲线拟合问题中常被应用。不仿看一个例。设飞行器在飞行过程中测得的沿某个坐标轴的位置可表示为

$$z(t_i) = \theta_1 + \theta_2 t_i + \theta_3 t_i^2 + \cdots + \theta_m t_i^{m-1} + \varepsilon(t_i)$$

$$i = 1, \cdots, n \tag{1.2.1}$$

其中 t_i 为时刻，$\varepsilon(t_i)$ 为 t_i 时刻的观测误差，其均值为 0，方差为 σ^2，$\theta_1, \cdots, \theta_m$ 为待估参数。要求我们由 n 个时刻的观测量 $z(t_1), \cdots, z(t_n)$ 去估计参数 $\theta_1, \cdots, \theta_m$.（1.2.1）式表示飞行器的运动轨迹是一条关于时间 t 的 $m-1$ 次多项式曲线。如果记

$$s(t_i) = z(t_i) - (\theta_1 + \theta_2 t_i + \cdots + \theta_m t_i^{m-1})$$

它表示了在 t_i 时刻的"残差"（观测值与多项式之差），而 $\sum_{i=1}^{n} s^2(t_i)$ 就是在观测点处的残差平方和。在这个和式中，$\theta_1, \cdots, \theta_m$ 是待定的。如果取 $\hat{\theta}_1, \cdots, \hat{\theta}_m$，它们使

$$S^2 \triangleq \sum_{i=1}^{n}[z(t_i) - (\hat{\theta}_1 + \hat{\theta}_2 t_i + \cdots + \hat{\theta}_m t_i^{m-1})]^2$$
$$= \text{Min}$$

这样所作出的 θ_i 的估计 $\hat{\theta}_i(i=1,\cdots,m)$ 称为 LS 估计。要具体去确定出 $\hat{\theta}_i$，只要用求极小值的方法就可以了。为此，只要令

$$\frac{\partial S^2}{\partial \hat{\theta}_i} = 0, \quad i = 1,\cdots,m$$

解出 $\hat{\theta}_i$，就是我们所需要的 LS 估计。现在我们用向量的形式来描述上述 LS 估计。记

$$Z = \begin{bmatrix} z(t_1) \\ \vdots \\ z(t_n) \end{bmatrix}$$

$$H = \begin{bmatrix} 1 & t_1 & t_1^2 & \cdots & t_1^{m-1} \\ 1 & t_2 & t_2^2 & \cdots & t_2^{m-1} \\ \vdots & & & & \\ 1 & t_n & t_n^2 & \cdots & t_n^{m-1} \end{bmatrix}_{n \times m}$$

$$X = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_m \end{bmatrix}, \qquad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

于是，$\text{Var}[\varepsilon] = E[\varepsilon\varepsilon^T] = \sigma^2 I$，而（1.2.1）式可以表示为

$$Z = HX + \varepsilon \tag{1.2.1'}$$

我们将会看到（见第 3 章），X 的 LS 估计是

$$\hat{X}_{LS} = (X^T X)^{-1} X^T Z \tag{1.2.2}$$

且它是 X 的无偏估计。估计误差的方差阵为

$$\text{Var}(\hat{X}_{LS} - X) = E[(\hat{X}_{LS} - X)(\hat{X}_{LS} - X)^T]$$
$$= \sigma^2 (X^T X)^{-1} \tag{1.2.3}$$

由上述观测值的曲线拟合方法可以看出，观测向量 Z 是未知参数向量 X 的线性函数。常称模型（1.2.1′）为线性模型，且记

作 $(Z,HX,\sigma^2 I)$．X 的 LS 估计及精度分别用（1.2.2）和（1.2.3）表示。这种 LS 估计方法是工程技术人员乐意采用的。因为应用方便，而且也完全不用给出误差向量的分布。但是，我们也应该指出，这种最小二乘方估计方法并不是在任何场合下都能适用的。例如，如果 X^TX 是退化阵（它在 H 不是满列秩的情况下就会发生这种情况），那末 X^TX 的逆矩阵（$X^TX)^{-1}$ 不存在，于是 $\hat{X}_{LS}$ 是没有意义的。此外，在工程应用中，还会发生 X^TX 近似退化的情况（它在 H 的列间具有近似线性相关性时将会发生这种情况）。在后续的讨论中，将会看到此时估计误差的方差将会变得十分大，以致估值 $\hat{X}$ 是不可信赖的。这种情况称之为 H 矩阵具有复共线性。

因此，当 H 具有复共线性时，最小二乘方估计必须进行改进。这种情况在专业实践中是必须注意的。我们在第 4 章中将专门讨论这个问题。

关于线性模型参数估计问题，我们还将同时讨论估计的优良性指标问题。例如，那种无偏估计的要求是否总是必要的？在应用中，人们关心的往往是估计量和未知参数之间的差距是否很小，而无偏性倒不一定十分重要。因此，我们还将引入估计优良性的其他指标。在不同的指标之下，改进估计的方法是不同的。

1.2.2 线性系统状态估计问题

对于总体分布的参数进行估计或检验，一般数理统计课程中已有讨论。但是对于一类随机过程的统计估值问题，我们还必须予以充分的注意。例如，飞行器的跟踪定轨、预报、拦截、导航等等就属于这类问题。需要研究的问题是随时间而变化的状态估计。例如导弹运动中的三个位置参数和三个速度参数。中东战争中，以色列以“爱国者”导弹拦截来袭的“飞毛腿”导弹，就要运用状态估计技术。一方面，通过雷达观测知道“飞毛腿”的飞行位置及速度（带有误差），经过统计处理之后给出来袭目标的状

态估计，并预报未来状态。另一方面，对提供的状态信息，由雷达导引“爱国者”导弹去击中目标。这个过程，观测信息的获取、统计估计、信息传递、快速的计算，都是十分重要的。关于状态估计的一般方法将在第5、6章中讨论。至于估计方法的工程实现，我们还将作为重点问题来讨论，这就是第7章的内容。

1.2.3 Bayes 统计推断方法

我们过去学习的数理统计是属于经典学派或称之为频率学派的理论和方法。而对于另一种学派——Bayes 学派的统计理论和方法则了解甚少。由于 Bayes 统计有它自身的特点，且越来越受到人们的关注。因此，我们在第8章将专门论述这种方法。这里，先作一点必要的介绍。设随机变量 X 的分布参数为 θ，它是我们要研究的对象。当获得了子样 $\boldsymbol{X}=(X_1,\cdots,X_n)$ 之后，要对于 θ 作统计推断。在经典统计中，θ 是未知的常量。但在 Bayes 统计中，将 θ 看作是随机的，而且在获取子样 $(X_1,\cdots,X_n)$ 之前，有关于 θ 的验前信息，且以验前分布的形式给出。记 θ 的验前分布密度为 $\pi(\theta)$. 当获得子样 $\boldsymbol{X}$ 之后，又可以计算 θ 的验后密度，记作 $\pi(\theta/\boldsymbol{X})$. 由 Bayes 公式

$$\pi(\theta/\boldsymbol{X})=\frac{\pi(\theta)p(\boldsymbol{X}/\theta)}{\int_{\Theta}\pi(\theta)p(\boldsymbol{X}/\theta)\mathrm{d}\theta}$$

其中 $p(\boldsymbol{X}/\theta)$ 为 θ 固定时子样 $\boldsymbol{X}$ 的条件密度函数，$\boldsymbol{\Theta}$ 为 θ 的变化区域。关于 θ 的统计推断将以 θ 的验后密度 $\pi(\theta/\boldsymbol{X})$ 作为出发点。这就是 Bayes 统计的基本思想。由此可以看出，关于 θ 的统计推断，不但与获得的子样 $\boldsymbol{X}$ 有关，且与 θ 的验前分布有关。这样，Bayes 统计较之经典统计，具有较多的信息。因此，如果验前的信息比较充分，虽然现场试验的次数较少，也可以作出合理的关于 θ 的统计推断。这种观点，容易被工程技术界所接受。特别是昂贵产品

(如导弹、卫星等)试验后的统计分析，就更是如此。我们讨论 Bayes 统计方法，也是出于这种思考。可以预期，Bayes 统计的运用，将可减少武器系统定型时的试射发数，缩短研制周期，使部队及时获得新的装备。因此经济和军事意义是明显的。

1.2.4 仿真试验的分析

仿真技术，已广泛地应用于各种不同的领域。对于飞行器的特性分析，无论是研制、试验或定型、使用等阶段，仿真是一种常用的手段。从统计分析的角度看，既然仿真试验可以大量地进行，因此，运用经典统计方法，仿真结果分析似乎是不困难的。然而，问题在于仿真结果和真实的试验（现场试验）结果是否一致。特别是对于动态参数的辨识，这是目前广为关注的问题。需要指出的是，由于昂贵产品或试验具有破坏性的产品，总不能大量抽试，因此，依赖于仿真提供的信息以及少量（有时甚至是极少量）的现场试验信息，对仿真系统的输出特性进行评估，这是十分有意义的工作。我们将会看到（第 9 章），运用 Bayes 统计推断方法，将上述两方面的信息融合，对于飞行器试验的分析工作将有重大的作用。

上述几个方面，只是本课程所论内容的一个基本索引。此外，为了更好地掌握本门课程的理论，我们编写了一章（第 2 章）随机过程导论。对于那些仅仅学习了一般的概率统计而没有学习随机过程理论的读者，学习这一章内容是完全必要的。

由于本书的论述中运用较多的矩阵知识，我们特意编写了一个附录，作为读者学习中的查考。

2 随机过程导论

飞行器试验统计学这门学科的建立，是基于对飞行器试验过程中的随机现象进行分析和统计处理，因而概率论与随机过程的一些概念和方法是飞行器试验统计学的重要基础。这一章在假定读者已具备工科概率论基础，但对随机过程知识不足的情况下，对随机过程的知识作一些必要的介绍。

2.1 随机过程的概念

自然界中事物变化的过程可以广泛地分为两类：一类是变化过程具有确定的形式，或者说有必然的变化规律；另一类是变化过程没有确定的变化形式，没有必然的变化规律。用数学语言来说，确定性过程可以用一个（或多个）关于时间 t 的确定的函数来描绘；而非确定性过程不能用一个（或多个）关于时间 t 的确定函数来描绘。严格说来，实际存在的变化过程都是非确定性的，确定性过程只是实际过程中忽略不确定性的因素而得到的一种理想化描述。例如，飞行器的飞行具有一定的规律可循，但由于各种不确定性因素（如风）的干扰，使得飞行器的运动具有不确定性。因此飞行器的飞行轨道是试验前不可准确预知的，进行多次试验也不可能有任何两次完全相同的结果出现。单次试验的结果仅是无限次可能结果中的一次，一次飞行试验得到的飞行器飞行轨道是时间的函数，它是随机现象的一次表现，称为样本函数，而所有可能产生的样本函数的集合就称为随机过程。由此，我们引出

随机过程的定义：

定义 2.1.1 随机过程是依赖于时间 t 的一族随机变量，常记为

$$\{X(t), t \in T\}$$

其中 T 是时间 t 的变化范围，对于每一个固定的时刻 $t_1 \in T$，$X(t_1)$ 是随机变量。$X(t)$ 的每一次表现（或称为实现）$x(t)$，为时间的确定性函数，称为样本函数。

例 2.1.1 在测量运动目标的距离时存在有误差，若以 $\varepsilon(t)$ 表示在时刻 t 的测量误差，则它是一个随机变量。当目标随时间 t 的变化按一定规律运动时，测量误差也随时间 t 而变化，换句话说，$\varepsilon(t)$ 是依赖于时间 t 的一族随机变量，亦即 $\varepsilon(t)$ 是一随机过程。

例 2.1.2 设随机相位余弦波 $X(t) = a\cos(\omega t + \Theta)$，其中 a，ω 为常数，Θ 是在 $[0,2\pi]$ 上服从均匀分布的随机变量。显然 $X(t)$ 是一个随机过程。对于每一个固定的时刻 $t = t_1$，$X(t_1) = a\cos(\omega t_1 + \Theta)$ 是一个随机变量；在 $[0,2\pi]$ 上随机地取一数 θ_i，相应地得到这个随机过程的一个样本函数：

$$x(t) = a\cos(\omega t + \theta_i), \quad \theta_i \in [0,2\pi]$$

图 2.1 画出了这个随机过程的三条样本曲线。

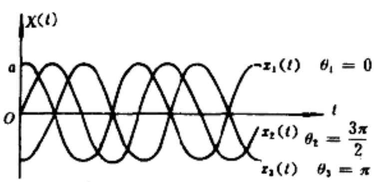


图 2.1 随机相位余弦波的样本函数

随机现象在工程技术中几乎随处可见。例如地震波幅、结构物承受的风载荷、船舶甲板“上浪”的次数以及通讯系统、自动控制系统中的各种噪声和干扰等等的变化过程都可以用随机过程这一数学模型来描绘，不过实际中绝大多数的随机过程不能像随机相位余弦波那样可具体地用时间和随机变量的关系式来表示，其主要原因在于自然界产生随机因素的机理是极为复杂的，甚至是不可能被观测到的。因而对于这样的随机过程，只有通过分析由测量所得到的样本函数集合才能掌握它们随时间变化的统计规律性。

随机过程 $\{X(t), t \in T\}$ 可按状态分为连续型随机过程和离散型随机过程。连续型随机过程在任意给定的时刻均为连续型随机变量；离散型随机过程在任意给定的时刻均为离散型随机变量。

另一方面，随机过程 $\{X(t), t \in T\}$ 按其时间参数集 T 的性质分为连续时间随机过程和离散时间随机过程。连续时间随机过程常简称为随机过程，其时间参数集 T 为连续时间集合，如 $T = \{t, t_0 \leqslant t \leqslant t_1\}$ 或 $T = \{t, t \geqslant t_0\}$；离散时间随机过程亦称为随机序列，其时间参数集为离散时间集合，如 $T = \{t_k, k = 0,1,\cdots,N\}$ 或 $T = \{t_k, k = 0,1,2,\cdots\}$.

下面我们介绍随机过程方面的知识。

2.2 随机过程的概率分布

随机过程是依赖于时间的一族随机变量。类似于随机变量，可以定义随机过程的概率分布函数和概率密度函数。

设 $X(t)$ 是一随机过程，对于每一个固定的时刻 $t_1 \in T$，$X(t_1)$ 是一个随机变量，它的分布函数一般与 t_1 有关，记为

$$F_1(x_1, t_1) = \mathrm{P}\{X(t_1) \leqslant x_1\} \tag{2.2.1}$$

称 $F_1(x_1, t_1)$ 为随机过程 $X(t)$ 的一维分布函数。