国外电子与电气工程技术丛书

HZ BOOKS
华章教育

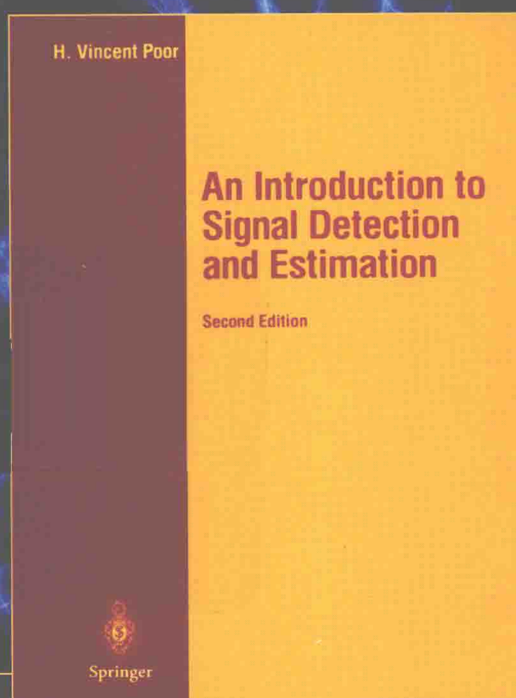
Springer

信号检测与估计

（原书第2版）

[美] H. Vincent Poor 著
廖桂生 杨志伟 等译

An Introduction to Signal Detection and Estimation
Second Edition



机械工业出版社
China Machine Press

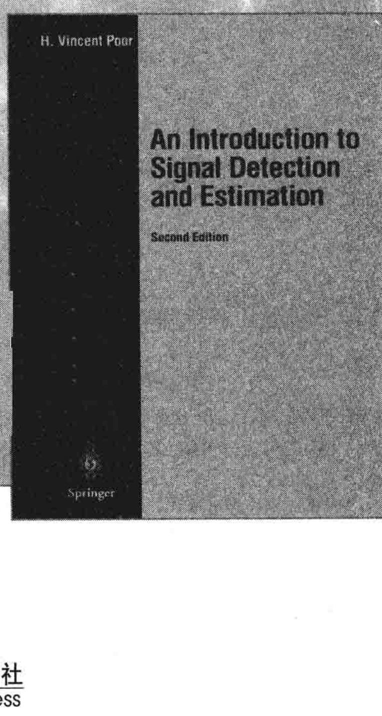
国外电子与电气工程技术丛书

信号检测与估计

（原书第2版）

[美] H. Vincent Poor 著
廖桂生 杨志伟 等译

An Introduction to Signal Detection and Estimation
Second Edition



机械工业出版社
China Machine Press

图书在版编目（CIP）数据

信号检测与估计（原书第 2 版）/（美）珀尔（Poor，H. V）著；廖桂生等译．—北京：机械工业出版社，2014.11

（国外电子与电气工程技术丛书）

书名原文：An Introduction to Signal Detection and Estimation, Second Edition

ISBN 978-7-111-48339-7

I. 信… II. ①珀… ②廖… III. ①信号检测 ②参数估计 IV. TN911.23

中国版本图书馆 CIP 数据核字（2014）第 244876 号

本书版权登记号：图字：01-2013-5973

本书介绍信号检测与估计的基本理论，主要包括三方面的内容。第 2、3、6 章讲述信号检测理论的基本概念，包括二元假设检验、复合假设检验、确知信号检测、部分参数确知信号检测、随机信号检测等内容。第 4 章讲述参数估计方法，包括贝叶斯参数估计方法、最大似然估计方法、向量参数估计、递归参数估计等内容。第 5、7 章讲述波形估计理论，包括 Kalman-Bucy 滤波、Wiener-Kolmogorov 滤波、线性 / 高斯估计、非线性滤波等内容。每章都有大量例题和习题，帮助读者理解抽象的理论知识。

本书可作为高等院校通信类、信息类和控制类等专业研究生或高年级本科生的专业教材或教学参考书，也可作为相关科研人员的参考用书。

出版发行：机械工业出版社（北京市西城区百万庄大街 22 号 邮政编码：100037）

责任编辑：曲 熠	责任校对：殷 虹
印 刷：北京瑞德印刷有限公司	版 次：2015 年 1 月第 1 版第 1 次印刷
开 本：185mm×260mm 1/16	印 张：16
书 号：ISBN 978-7-111-48339-7	定 价：49.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

客服热线：（010）88378991 88361066　　投稿热线：（010）88379604

购书热线：（010）68326294 88379649 68995259　　读者信箱：hzjsj@hzbook.com

译者序

本书英文版自出版以来，被全球众多著名大学选为研究生“统计信号处理”课程的教材或教学参考书。Poor 教授以多年的研究积淀和教学经验为基础撰写了本书，通过易于理解的方式，深入阐述信号检测和估计理论的基本思想。全书主要包括三方面的内容，第 2、3、6 章讲述信号检测理论的基本概念，包括离散域的二元假设检验、复合假设检验、序贯检测、非参数检测、稳健检测，连续域的确知信号检测、部分参数确知信号检测、随机信号检测等内容。第 4 章讲述参数估计方法，包括贝叶斯参数估计方法、最大似然估计方法、非随机参数估计的理论结构、向量参数估计、稳健参数估计、递归参数估计等内容。第 5、7 章讲述波形估计理论，包括离散域的 Kalman-Bucy 滤波、线性因果/非因果 Wiener-Kolmogorov 滤波，连续域的线性/高斯估计、非线性滤波等内容。本书最为突出的特点是提供了大量的实例，无论是在离散域还是连续域，都以简单实用的例子来描述检测与估计的基本理论与方法，并通过应用实例来比较各种方法的异同。此外，每章都有大量习题，这些习题与该章内容密切相关，很多习题是书中结论的证明或进一步阐述，以加深读者对抽象理论的理解。

本书要求读者具备理工科研究生基本课程中关于应用概率论及随机过程的相关知识。此外，在教学过程中，可根据需要进行内容取舍。

参加本书翻译工作的主要有廖桂生（前言，第 1、2、4 章）和杨志伟（第 3、5、6、7 章）。参加翻译、校对和译稿资料整理工作的还有刘志凌、郭小路、谢坚、党博、李东、粟嘉、何嘉懿、张学攀、许京伟、高永婵、杨东、束宇翔、杜文韬。全书由廖桂生进行审校。

由于译者水平有限，书中难免有不当之处，敬请读者批评指正。

前　言

本书旨在介绍信号检测与估计的基本理论。读者在阅读本书时，应当具备理工科研究生第一学期基本课程中关于应用概率论及随机过程的相关知识。这些预备知识可以参考 Wong（1983）的著作。信号检测与估计理论所涉及的更复杂的概念将主要在第 6 章和第 7 章中介绍，包括连续时间信号检测和估计问题。

本书改编自伊利诺伊大学和普林斯顿大学一个学期的第二阶段研究生课程。如果将授课范围限制在第 1～5 章，本书也可用于短期课程或者第一阶段课程，此时仅要求读者具备应用概率论的基本知识，包括随机向量和条件期望等概念。为了获取足够的背景知识，读者可以参考 Thomas（1986）的著作。

此外，还可以以其他教学模式使用本书。例如，第 2、3、6 章可用于介绍信号检测的短期课程，第 4、5、7 章可用于介绍信号估计的短期课程。类似地，还可以将第 1 章和第 4 章用于入门级课程，而将第 5 章和第 7 章用于高级课程。

在第 2 版的筹备过程中，听取了很多读者的意见和建议。由于人数众多，在此不便一一列出，但要对这些读者致以深深的谢意。

目　录

第1章 引言

信号检测与估计是信号处理的一个研究领域，其目的是从信号中提取需要的信息。信号检测与估计的相关理论已在通信、自动控制等领域得到广泛应用。例如，在雷达和数据传输等通信领域，信号检测和估计是设计高效通信接收机的理论基础和分析工具。此外，在自动控制领域，检测与估计理论是精确推断进程状态和进行系统控制的基础。

雷达是检测与估计技术的一个应用实例。雷达的基本工作方式是发射一个电磁脉冲信号，然后等待目标可能反射的回波信号。受接收机噪声、大气扰动、地面和其他目标寄生反射以及信号失真等因素的影响，往往无法绝对准确地判断目标是否存在。因此，必须根据接收天线输出的(非理想的)观测结果来推断目标是否存在，此时，基于检测理论可设计出最优化的检测方法。进而，在判断出目标以某种概率存在之后，还希望估计目标的一些特性，比如位置和速度等，而这在广义上属于估计理论的范畴。上述估计结果可用于控制天线以跟踪目标或者远程控制目标以实现特定航迹飞行。检测估计技术的其他具体应用还包括地震学、射电天文学、声呐、语音和图像处理、医学信号处理及光通信等。

在检测与估计技术的实际应用中，通常会涉及基于观测值的推断处理，并且这些观测值往往存在不明原因导致的失真和缺损，甚至从这些观测值中提取的信息也是未知的。因此，将检测和估计问题置于概率框架下处理是非常有效的，此时可以将未知因素作为随机量来处理。就此而论，在统计推断领域内研究检测估计理论非常适合，并且在本书的论述过程中都将以这种方式进行阐述。

信号检测与估计理论学习中的一个基本概念是随机观测量 Y，它取值于观测值集合 Γ，其中 Γ 可以是向量的集合、波形的集合、实数的集合或者其他任意集合。我们期望从观测量 Y 中提取有关 Y 的一些特征信息。本书主要关注两类问题：检测问题和估计问题。在检测问题中，我们期望基于一些有限数量的可能情况或“自然状态”做出判决；在估计问题中，我们期望对一些无法直接观测的量进行估值。在上述两类问题中，观测值和待提取信息之间的关系均是概率性的，而不是直接必然的，在此意义上而言，Y 的统计特性受其自然状态或待估计量的真实值影响。因此，在建立检测/估计模型时，必须涉及基于 Γ 的一族概率分布，其中各元素均对应不同自然状态或在不同待估计量值下所呈现的统计状态。检测/估计模型一旦建立，则相应的检测/估计问题的目标即转化为找到一种处理观测量 Y 的最优方式，以便提取所需信息。区分这类问题的基本特征包括：待估计信息的性质(离散或连续)，关于待估计量或自然状态的先验信息的多少，以及用于比较不同检测估计过程的性能指标。

本书旨在向读者介绍检测与估计理论的基本原理。第 2 章、第 3 章和第 6 章主要讨论信号估计的相关问题。第 2 章介绍二元假设检验的基础知识，这部分知识是大多数信号检

测问题的基础。第 3 章将应用上述基础知识针对特定的信号检测模型推导最优检测器并分析其检测性能。此外，第 3 章还将讨论用于处理非标准情况下信号检测问题的几种特殊信号检测方法。第 2 章和第 3 章主要讨论观测值为向量时的情况，它与基于离散时间观测值(采样观测值)的信号检测问题相对应。基于连续时间观测值的信号检测问题将在第 6 章中进行讨论。连续时间模型的信号检测问题虽然与离散时间模型在本质上相同，但是由于对于这类问题的分析中涉及了更复杂的分析方法，因此将对其进行单独讨论。第 4 章、第 5 章和第 7 章讨论估计问题。第 4 章中给出了参数估计问题的基本原理和结构。与信号检测类似，由于分析难度不同，所以对离散时间估计(第 4 章、第 5 章)和连续时间估计(第 7 章)分别进行了讨论。

符号说明

为了详细描述观测集合 Γ 的概率分布，需要对 Γ 的子集进行概率分配。对于某些观测空间而言，无法为 Γ 的所有子集分配一致的概率值，因此，我们总是将 Γ 与它的一类子集 $\mathcal{G}$ 联系在一起，并期望对这类子集进行概率分配。$\mathcal{G}$ 中的集合称为**观测结果**，而组合 $(\Gamma, \mathcal{G})$ 则称为**观测空间**。为了便于分析，始终假设 $\mathcal{G}$ 满足 σ 代数，即假设 $\mathcal{G}$ 包含所有补集(关于 Γ)和其元素的可数并集。[⊖]

本书主要关注 $(\Gamma, \mathcal{G})$ 的两种情况。第一种情况是 $\Gamma=\mathbf{R}^n$，即 Γ 为 n 维实向量的集合；第二种情况是 Γ 为离散(可数)集合，且 $\Gamma=\{\gamma_1, \gamma_2, \cdots\}$。在第一种情况中，很自然地期望为形如 $\{y=(y_1, \cdots, y_n)^{\mathrm{T}}\in\mathbf{R}^n \mid a_1\leqslant y_1\leqslant b_1, \cdots, a_n\leqslant y_n\leqslant b_n\}$ 的集合分配概率，其中，a_i 和 b_i 均为任意实数。因此，对于 $\Gamma=\mathbf{R}^n$，把 $\mathcal{G}$ 看作包含 a_i 和 b_i 所确定的全部子集的最小 σ 代数，其中，a_i 和 b_i 在整个实数范围内取值。这类 σ 代数一般用 $\mathcal{B}^n$ 表示，称为 $\mathbf{R}^n$ 中的 Borel 集合。对于第二种情况，可以将 $\mathcal{G}$ 定义为 Γ 的所有子集。这类 σ 代数一般用 2^Γ 表示，称为 Γ 的幂集合。在上述两种观测空间内足以描述第 2 章～第 5 章中所讨论的大部分离散时间检测与估计问题。如无特别说明，均假设 $(\Gamma, \mathcal{G})$ 属于上述两种情况之一。在讨论连续时间的检测估计问题时会涉及更为抽象的观测空间，相关内容会在第 6 章和第 7 章加以介绍。

对于离散观测空间 $(\Gamma, 2^\Gamma)$，以**概率质量函数**的形式为 Γ 中的子集分配概率，即 p: $\Gamma\to[0, 1]$，且

$$P(A)=\sum_{\gamma_i\in A}p(\gamma_i),\quad A\in 2^\Gamma \tag{1.1}$$

其中，$P(A)$ 表示观测量 Y 的取值位于集合 A 中的概率。如果满足条件 $\sum_{i=1}^{\infty}p(\gamma_i)=1$，则将 Γ 映射到 $[0, 1]$ 的任意函数均是合理的概率质量函数。对于观测空间 $(\mathbf{R}^n, \mathcal{B}^n)$，则主要关注所谓的**连续随机向量**，且可按照**概率密度函数**的形式对其进行概率分配，即 p: $\mathbf{R}^n\to[0, \infty)$，且[注意：式(1.2)中的积分为 n 重积分]

$$P(A)=\int_A p(y)\mathrm{d}y,\quad A\in\mathcal{B}^n \tag{1.2}$$

⊖ 也就是说，$\mathcal{G}$ 具有下述性质：若 $A\in\mathcal{G}$，则有 $A^c\in\mathcal{G}$(本书中上标 c 表示取补集)；若 $A_1, A_2, \cdots\in\mathcal{G}$，则有 $\bigcup_{i=1}^{\infty}A_i\in\mathcal{G}$。

如果满足条件$\int_{\mathbf{R}^n} p(y)\mathrm{d}y = 1$，则将$\mathbf{R}^n$映射到$[0, \infty)$的任意可积函数均是合理的概率密度函数。为了简化术语和公式表述，对概率质量函数和概率密度函数均使用密度一词，且将式(1.1)中的求和公式和式(1.2)中的积分公式统一用下述公式表示

$$P(A) = \int_A p(y)\mu(\mathrm{d}y) \tag{1.3}$$

其中，也可以省略积分变量，将上式进一步简写为

$$P(A) = \int_A p\,\mathrm{d}\mu \tag{1.4}$$

对于以随机观测量Y为自变量的实值函数g，经常需要求解$g(Y)$的期望值，即$E\{g(Y)\}$。对于离散观测空间$(\Gamma, 2^{\Gamma})$，上述期望值为

$$E\{g(Y)\} = \sum_{i=1}^{\infty} g(\gamma_i)p(\gamma_i) \tag{1.5}$$

而对于$(\mathbf{R}^n, \mathcal{B}^n)$中的连续随机向量，有

$$E\{g(Y)\} = \int_{\mathbf{R}^n} g(y)p(y)\mathrm{d}y \tag{1.6}$$

其中，对于每一种情况，均假设相应的求和或积分是存在的。同样地，为了简化公式，采用下述公式对式(1.5)和式(1.6)进行统一表述

$$E\{g(Y)\} = \int_{\Gamma} g(y)p(y)\mu(\mathrm{d}y) = \int_{\Gamma} gp\,\mathrm{d}\mu \tag{1.7}$$

对上述公式的进一步阐述将在第6章中给出。易知，式(1.3)和式(1.4)是式(1.7)中的函数g按下式定义时的特殊情况

$$g(y) = \begin{cases} 1, & y \in A \\ 0, & y \in A^c \end{cases} \tag{1.8}$$

在后续的讨论中，均采用大写字母表示随机量，并采用小写字母表示这些随机量的特定取值。因此，本书的表述方式为：随机观测量Y的可能取值为y。

第2章

假设检验基础

2.1 引言

大多数信号检测问题都可以归结为 M 元假设检验问题，这一假设以观测数据(可用向量或函数表示)为基础，通过判决 M 元统计信号来描述观测数据。例如，在 M 元通信接收系统中，观测数据的波形为 M 元信号之一，该信号会受到随机信道或接收机噪声的干扰，因此我们期望通过统计判决来确定可表示观测数据的具体信号形式。显然，对于任意给定的判决问题，都有一系列可行的判决策略或准则可以应用，然而，本书希望选择一种在某种意义上是最优的判决准则。尽管就最优性而言，存在多种定义，但本章仅考虑三种最常见的准则——贝叶斯准则、极小化极大准则和尼曼-皮尔逊准则，并推导这些准则的最优解。为了不失一般性，虽然本章中的大多数结论可直接推广到 M 元假设检验(在习题中会具体探讨)，但是本章只考虑二元假设检验问题。第 3 章和第 6 章将详细讨论这些理论在信号检测中的具体应用。

2.2 贝叶斯假设检验

本章所考虑的首要问题是简单的二元假设检验问题，它有两种可能的假设，称为“状态或类型”，即 H_0 和 H_1，分别对应观测空间$(\Gamma, \mathcal{G})$中的两种可能的概率分布 P_0 和 P_1。可将问题描述如下

$$\begin{aligned} H_0: Y &\sim P_0 \\ H_1: Y &\sim P_1 \end{aligned} \tag{2.2.1}$$

式中，记号“$Y \sim P$”表示“Y 满足 P 分布”。假设 H_0 和假设 H_1 有时也称为零假设和非零假设。关于假设 H_0 和假设 H_1 的判决准则 δ 将观测集 Γ 分成 $\Gamma_1 \in \mathcal{G}$ 和 $\Gamma_0 = \Gamma_1^c$ 两个子集，如果 $j=0$ 或 1 时存在 $y \in \Gamma_j$，则判决假设 H_j 成立。集合 Γ_1 称为拒绝域(或关键域)，集合 Γ_0 称为接受域。也可将判决准则 δ 看成是由下式给出的在集合 Γ 上的函数，即

$$\delta(y) = \begin{cases} 1, & y \in \Gamma_1 \\ 0, & y \in \Gamma_1^c \end{cases} \tag{2.2.2}$$

对给定的 $y \in \Gamma$，δ 的取值等于判决规则 δ 所接受假设的下标值。

本章旨在以某种最优的方法选择 Γ_1，在这种思路的指引下，需要对不同判决赋予不同的代价。特别地，假设正数 C_{ij}($i=0$，1，$j=0$，1)是在假设 H_j 为真的情况下，判决假设 H_i 成立所需付出的代价。对于每个假设，可将条件风险定义为：当该假设为真时，由判决准则所引发的平均代价或期望代价，即

$$R_j(\delta) = C_{1j}P_j(\Gamma_1) + C_{0j}P_j(\Gamma_0), \quad j = 0,1 \tag{2.2.3}$$

注意，$R_j(\delta)$是当 H_j 为真时选择 H_1 的代价与选择 H_1 的概率的乘积加上选择 H_0 的代价与选择 H_0 的概率的乘积。

假设 H_0 和 H_1 的概率是 π_0 和 $\pi_1=1-\pi_0$，即 π_j 是假设 H_j 为真的概率，它与 Y 的取值无关。概率 π_0 和 π_1 也称为两种假设的先验概率。给定先验概率后，平均风险或贝叶斯风险可用来描述判决准则 δ 的总平均代价，表达式由下式给出

$$r(\delta)=\pi_0 R_0(\delta)+\pi_1 R_1(\delta) \tag{2.2.4}$$

根据贝叶斯风险的定义，可设计关于假设 H_0 和假设 H_1 的最优判决准则，该判决准则在所有判决准则中具有最小的贝叶斯风险，称为关于 H_0 和 H_1 的贝叶斯准则。

综合式(2.2.3)和式(2.2.4)并利用 $P_j(\Gamma_1^c)=1-P_j(\Gamma_1)$后，可以得到

$$\begin{aligned} r(\delta) &= \sum_{j=0}^{1}\pi_j[C_{0j}(1-P_j(\Gamma_1))+C_{1j}P_j(\Gamma_1)] \\ &= \sum_{j=0}^{1}\pi_j C_{0j}+\sum_{j=0}^{1}\pi_j(C_{1j}-C_{0j})P_j(\Gamma_1) \end{aligned} \tag{2.2.5}$$

假设 P_j 的概率密度是 $p_j(j=0，1)$，根据第 1 章的符号说明，可将式(2.2.5)表示为

$$r(\delta)=\sum_{j=0}^{1}\pi_j C_{0j}+\int_{\Gamma_1}\Big[\sum_{j=0}^{1}\pi_j(C_{1j}-C_{0j})p_j(y)\Big]\mu(\mathrm{d}y) \tag{2.2.6}$$

如果按照下式选取 Γ_1，则 $r(\delta)$可取得极小值，即

$$\begin{aligned} \Gamma_1 &= \Big\{y\in\Gamma\,\Big|\,\sum_{j=0}^{1}\pi_j(C_{1j}-C_{0j})p_j(y)\leqslant 0\Big\} \\ &= \{y\in\Gamma\,|\,\pi_1(C_{11}-C_{01})p_1(y)\leqslant\pi_0(C_{00}-C_{10})p_0(y)\} \end{aligned} \tag{2.2.7}$$

假设 $C_{11}<C_{01}$(正确选择 H_1 的代价小于错误拒绝 H_1 的代价)，式(2.2.7)可以重写成

$$\Gamma_1=\{y\in\Gamma\,|\,p_1(y)\geqslant\tau p_0(y)\} \tag{2.2.8}$$

式中

$$\tau\triangleq\frac{\pi_0(C_{10}-C_{00})}{\pi_1(C_{01}-C_{11})} \tag{2.2.9}$$

注意，区域$\{y\in\Gamma\,|\,p_1(y)=\tau p_0(y)\}$对平均误差没有影响，因此可以根据需要将其从 Γ_1 中全部或部分省略。

将式(2.2.8)中由拒绝域所描述的判决准则定义为似然比检验(或概率比检验)，该检验在假设检验理论中处于核心地位。注意，式(2.2.8)中的 Γ_1 可以重写成

$$\Gamma_1=\{y\in\Gamma\,|\,[p_1(y)/p_0(y)]\geqslant\tau\} \tag{2.2.10}$$

其中，定义对于任何 $k\geqslant 0$ 的情况，$k/0$ 的值为∞。H_0 和 H_1 之间的似然比定义为

$$L(y)=\frac{p_1(y)}{p_0(y)},\quad y\in\Gamma \tag{2.2.11}$$

将上式定义为 H_0 和 H_1 之间的似然比(或似然比统计)。显然，式(2.2.8)所描述的贝叶斯判决准则首先需要计算出基于 Y 的观测值的似然比，然后通过与门限值 τ 作对比来完成判决。换言之，贝叶斯判决准则可写成

$$\delta_B(y)=\begin{cases}1, & L(y)\geqslant\tau \\ 0, & L(y)<\tau\end{cases} \tag{2.2.12}$$

此外，均匀代价是常用的代价分配方法，即

$$C_{ij}=\begin{cases}0, & i=j\\ 1, & i\neq j\end{cases} \tag{2.2.13}$$

与之对应的贝叶斯风险为

$$r(\delta)=\pi_0 P_0(\Gamma_1)+\pi_1 P_1(\Gamma_0) \tag{2.2.14}$$

注意，$P_i(\Gamma_j)$是当H_i为真时，选择H_j的概率。因此，对于$i\neq j$的情况，$P_i(\Gamma_j)$是H_i为真时发生误判的概率，$r(\delta)$是由判决准则δ产生的平均错误概率。又因为似然比判决$\tau=\pi_0/\pi_1$可最小化式(2.2.13)对应的贝叶斯风险$r(\delta)$，所以它也是最小错误概率判决。

根据贝叶斯准则[参见 Thomas(1986)]，在随机观测值Y取值为y的条件下，H_j为真的条件概率可表示为

$$\pi_j(y)=P(H_j\ \text{为真}\,|\,Y=y)=\frac{p_j(y)\pi_j}{p(y)} \tag{2.2.15}$$

式中，$p(y)=\pi_0 p_0(y)+\pi_1 p_1(y)$表示$Y$的平均或全局概率密度。$\pi_0(y)$和$\pi_1(y)$称为两种假设的后验概率。利用式(2.2.15)，式(2.2.7)所描述的贝叶斯判决准则的关键域可重写成

$$\Gamma_1=\{y\in\Gamma\,|\,C_{10}\pi_0(y)+C_{11}\pi_1(y)\leqslant C_{00}\pi_0(y)+C_{01}\pi_1(y)\} \tag{2.2.16}$$

因此，在贝叶斯判决准则框架中，最优判决是基于后验概率的。即利用观测值将先验概率更新为后验概率。与之对应的后验代价为

$$C_{i0}\pi_0(y)+C_{i1}\pi_1(y) \tag{2.2.17}$$

式(2.2.17)是已知Y等于y时，选择假设H_i所产生的平均代价。因此贝叶斯准则通过选择具有最小后验代价的假设来实现判决。例如，对于由式(2.2.13)描述的均匀代价准则而言，贝叶斯准则可以写成

$$\delta_B(y)=\begin{cases}1, & \pi_1(y)\geqslant\pi_0(y)\\ 0, & \pi_1(y)<\pi_0(y)\end{cases} \tag{2.2.18}$$

因此，最小错误概率判决准则通过选取$Y=y$时具有最大后验概率的假设实现判决。该判决准则也称为二元假设检验的 MAP(maximum a posteriori)判决。

下面的例子将详细阐明贝叶斯判决准则。

例 2.2.1 二元信道

假设通信信道内数据以二进制(即“0”或“1”)进行传输，观测值Y的取值为 0 或 1。由于存在信道噪声、非理想调制或解调等原因，信道中传输的“0”可能以概率λ_0被接收成“1”，以概率$1-\lambda_0$被接收成“0”，其中$0<\lambda_0<1$。类似地，传输的“1”以概率λ_1被接收成“0”，以概率$1-\lambda_1$被接收成“1”，其中$0<\lambda_1<1$(具体描述见图 2.2.1)。因此，由观测值Y无法精确判定传输的数字究竟是“0”还是“1”，需要寻求某个优化准则以实现判决。

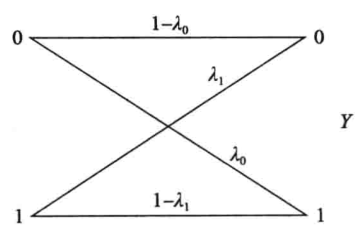


图 2.2.1 二元信道

该问题可建模成二元假设检验问题，其中，假设H_j表示“j”被传输($j=0, 1$)，观测集Γ是$\{0, 1\}$，观测值Y的概率密度可表示为

$$p_j(y)=\begin{cases}\lambda_j, & y\neq j\\ 1-\lambda_j, & y=j\end{cases} \tag{2.2.19}$$

其中，$j=0$ 和 1。因此，似然比可由下式给出

$$L(y)=\frac{p_1(y)}{p_0(y)}=\begin{cases}\dfrac{\lambda_1}{1-\lambda_0}, & y=0\\[2mm] \dfrac{1-\lambda_1}{\lambda_0}, & y=1\end{cases} \tag{2.2.20}$$

对于贝叶斯检验，根据式(2.2.9)中的代价和先验概率来确定门限值 τ。如果 λ_1、λ_0 和 τ 满足 $\lambda_1\geqslant\tau(1-\lambda_0)$，式(2.2.12)所表示的似然比检验可解释为将接收的“0”判决为传输的“1”；否则，接收的“0”对应传输的“0”。类似地，如果 $1-\lambda_1\geqslant\tau\lambda_0$，似然比检验就可解释为将传输的“1”接收成“1”；如果 $1-\lambda_1<\tau\lambda_0$，则接收到的“1”将被看成是“0”。对于本例，最小的贝叶斯风险 $r(\delta_B)$ 可以由式(2.2.5)直接计算出来(参见习题 1)。

例如，对于均匀代价且先验概率相等($\pi_0=\pi_1=1/2$)的情况，有 $\tau=1$，且式(2.2.12)的贝叶斯准则变为

$$\delta_B(0)=\begin{cases}1, & \lambda_1\geqslant(1-\lambda_0)\\ 0, & \lambda_1<(1-\lambda_0)\end{cases} \tag{2.2.21a}$$

$$\delta_B(1)=\begin{cases}1, & (1-\lambda_1)\geqslant\lambda_0\\ 0, & (1-\lambda_1)<\lambda_0\end{cases} \tag{2.2.21b}$$

根据前面的讨论，$L(y)$ 等于 τ 的边界点可任意判决为 Γ_1 或者 Γ_0。因此从风险角度出发，等价的贝叶斯检验可表示为

$$\delta_B(y)=\begin{cases}y, & (1-\lambda_1)\geqslant\lambda_0\\ 1-y, & (1-\lambda_1)<\lambda_0\end{cases} \tag{2.2.22}$$

如果进一步假设信道对称，即 $\lambda_1=\lambda_0=1$，则式(2.2.22)变成

$$\delta_B(y)=\begin{cases}y, & \lambda\leqslant 1/2\\ 1-y, & \lambda>1/2\end{cases} \tag{2.2.23}$$

式(2.2.23)的含义非常直观：如果信道发生比特翻转的概率较高($\lambda>1/2$)，则接收端需要进行比特翻转，即通过翻转所接收到的比特来作判决；反之，接收端不进行比特翻转。对于后一种情况，最小贝叶斯风险为

$$r(\delta_B)=\min\{\lambda,1-\lambda\} \tag{2.2.24}$$

因此，直接传输或翻转比特后再传输的信道越可靠，传输性能就越好。值得注意的是，在均匀代价和相等的先验概率条件下，即使缺乏观测值 y，随机猜测的风险也是 1/2。因此，就 $\lambda=1/2$ 而言，观测是没有价值的。◀

例 2.2.2　高斯误差下的位置检验

考虑关于实值观测值 Y 的两种假设

$$\begin{aligned}H_0&:Y=\varepsilon+\mu_0\\ H_1&:Y=\varepsilon+\mu_1\end{aligned} \tag{2.2.25}$$

其中，ε 是均值为 0、方差为 σ^2 的高斯随机变量，μ_0 和 μ_1 是两个固定值，并满足 $\mu_1>\mu_0$。注意，由于 μ_0 或 μ_1 与 ε 的和仅改变观测值的均值，因此需要对观测数据的两种分布或“位

置"进行判决。从观测空间的分布角度看，式(2.2.25)的两个假设可重写成

$$\begin{aligned} H_0: Y &\sim \mathcal{N}(\mu_0, \sigma^2) \\ H_1: Y &\sim \mathcal{N}(\mu_1, \sigma^2) \end{aligned} \tag{2.2.26}$$

其中，$\mathcal{N}(\mu, \sigma^2)$表示均值为 μ、方差为 σ^2 的高斯(正态)分布。服从$\mathcal{N}(\mu, \sigma^2)$分布的随机变量的概率密度可表示为

$$(1/\sqrt{2\pi}\sigma)\exp\{-(x-\mu)^2/2\sigma^2\}, \quad x \in \mathbf{R}$$

式(2.2.26)的似然比为

$$L(y) = \frac{p_1(y)}{p_0(y)} = \frac{\dfrac{1}{\sqrt{2\pi}\sigma}e^{-(y-\mu_1)^2/2\sigma^2}}{\dfrac{1}{\sqrt{2\pi}\sigma}e^{-(y-\mu_0)^2/2\sigma^2}} = \exp\left\{\frac{\mu_1-\mu_0}{\sigma^2}\left(y-\frac{\mu_0+\mu_1}{2}\right)\right\} \tag{2.2.27}$$

即式(2.2.26)的贝叶斯检验为

$$\delta_B(y) = \begin{cases} 1, & \exp\left\{\dfrac{\mu_1-\mu_0}{\sigma^2}\left(y-\dfrac{\mu_0+\mu_1}{2}\right)\right\} \geqslant \tau \\ 0, & \text{其他} \end{cases} \tag{2.2.28}$$

其中，τ 是判决门限。因为 $\mu_1 > \mu_0$，所以式(2.2.27)的似然比是关于观测值 y 的严格增函数$[\mathrm{d}L(y)/\mathrm{d}y=(\mu_1-\mu_0)L(y)/\sigma^2>0]$。所以，将 $L(y)$与门限值 τ 作比较等价于将 y 与另一个门限值 $\tau'=L^{-1}(\tau)$作比较，其中 L^{-1}是 L 的反函数。特别地，对式(2.2.28)中不等式分别取对数，并重新整理之后得到

$$\delta_B(y) = \begin{cases} 1, & y \geqslant \tau' \\ 0, & y < \tau' \end{cases} \tag{2.2.29}$$

其中

$$\tau' = \frac{\sigma^2}{\mu_1-\mu_0}\log(\tau) + \frac{\mu_0+\mu_1}{2} \tag{2.2.30}$$

例如，在均匀代价和先验概率相等的情况下，可以得到 $\tau=1$ 和 $\tau'=(\mu_0+\mu_1)/2$。此时，贝叶斯准则将观测值与 μ_0 和 μ_1 的平均值作比较。如果 y 大于或等于平均值，选择 H_1；如果 y 小于平均值，选择 H_0。图 2.2.2 是与之对应的判决示意图。

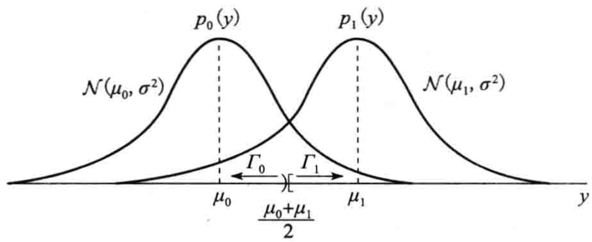


图 2.2.2 高斯误差、均匀代价和等先验概率下的位置检验

如果已知 $P_j(\Gamma_1)(j=0, 1)$，则最小贝叶斯风险 $r(\delta_B)$可由式(2.2.5)计算得到。又因为$\Gamma_1=\{y\in\mathbf{R}\,|\,y\geqslant\tau'\}$，进而可以得到

$$P_j(\Gamma_1) = \int_{\tau'}^{\infty} p_j(y)\mathrm{d}y = 1-\Phi\left(\frac{\tau'-\mu_j}{\sigma}\right) = \begin{cases} 1-\Phi\left(\dfrac{\log\tau}{d}+\dfrac{d}{2}\right), & j=0 \\ 1-\Phi\left(\dfrac{\log\tau}{d}-\dfrac{d}{2}\right), & j=1 \end{cases} \tag{2.2.31}$$

其中，Φ 表示满足$\mathcal{N}(0, 1)$分布的随机变量的累积概率分布函数，$d \triangleq (\mu_1-\mu_0)/\sigma$。对于均匀代价和先验概率相等的特殊情况，可以直接得到

$$r(\delta_B) = 1 - \Phi(d/2) \tag{2.2.32}$$

图 2.2.3 给出了贝叶斯风险与 d 的关系。注意，风险随均值($\mu_1 - \mu_0$)与误差标准差 σ 的比值的增大而单调递减。d 可理解为信噪比，在接下来的章节中将会深入介绍其含义。◀

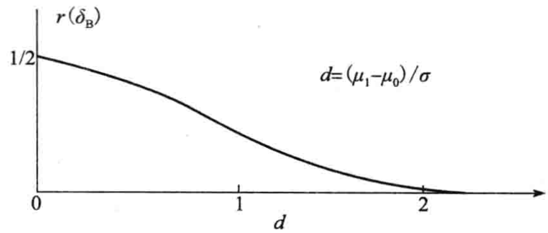


图 2.2.3　高斯误差下位置判决的贝叶斯风险

虽然例 2.2.1 和例 2.2.2 相当简单，但却说明了贝叶斯假设检验的基本原理。第 3 章和第 4 章将讨论更复杂的例子，本章末还有一系列相关习题。

本节最重要的结论是：给定代价 C_{ij} 和先验概率 π_i 后，式(2.2.12)所描述的似然比是贝叶斯意义上的最优判决准则。下一节将讨论在先验概率与(或)代价未知情况下的其他最优判决准则。

2.3　极小化极大假设检验

本节考虑先验概率 π_0 和 π_1 未知的判决准则设计问题，这是因为在设计判决准则时通常无法控制或精确知道系统的运行机制。此时，最小化平均或贝叶斯风险都不是可接受的设计准则，因为不太可能利用单一的判决准则来最小化每个可能的先验概率分布的平均风险。因此，有必要寻求新的判决准则，该准则相对所有的判决准则 δ 而言能使最大的条件风险 $R_0(\delta)$ 和 $R_1(\delta)$ 最小化，即判决准则最小化性能指标

$$\max\{R_0(\delta), R_1(\delta)\} \tag{2.3.1}$$

该判决准则也称为极小化极大准则。接下来的章节将会讨论该判决准则的具体实现方法。

为了寻求最小化式(2.3.1)的判决准则，首先考虑已知先验概率 $\pi_0 \in [0,1]$ 和判决准则 δ 的平均风险 $r(\pi_0, \delta)$，即

$$r(\pi_0, \delta) = \pi_0 R_0(\delta) + (1 - \pi_0) R_1(\delta) \tag{2.3.2}$$

在图 2.3.1 中，π_0 的函数 $r(\pi_0, \delta)$ 是一条从 $r(0, \delta) = R_1(\delta)$ 到 $r(1, \delta) = R_0(\delta)$ 的直线。因此，当 δ 值固定，且 π_0 在 0～1 之间取值时，$r(\pi_0, \delta)$ 的最大值不是在 $\pi_0 = 0$ 处，就是在 $\pi_0 = 1$ 处，并且最大值是 $\max\{R_0(\delta), R_1(\delta)\}$。所以最小化式(2.3.1)的问题等价于在整个 δ 上最小化下式

$$\max_{0 \leqslant \pi_0 \leqslant 1} r(\pi_0, \delta) \tag{2.3.3}$$

相比式(2.3.1)而言，式(2.3.3)更容易求解。

对于每个先验概率 $\pi_0 \in [0,1]$，假设 δ_{π_0} 表示对应该先验概率的贝叶斯判决准则。令 $V(\pi_0) = r(\pi_0, \delta_{\pi_0})$ 表示关于先验概率 π_0 的最小贝叶斯风险。易知 $V(\pi_0)$ 是关于 π_0 的连续凹函数，其中 $\pi_0 \in [0,1]$，$V(0) = C_{11}$，$V(1) = C_{00}$(参见习题 8)。图 2.3.1 画出了典型的 $V(\pi_0)$ 曲线。

为了讨论方便，假设 $V(\pi_0)$ 和 $r(\pi_0, \delta)$ 如图 2.3.1 所示。其中还有一条标有 $r(\pi_0, \delta_{\pi_0'})$ 的直线，该直线平行于 $r(\pi_0, \delta)$，并且相切于 $V(\pi_0)$。注意，对于这种情况，δ 不是极小化极大准则，因为 $r(\pi_0, \delta_{\pi_0'})$ 所表示的代价直线全部在直线 $r(\pi_0, \delta)$ 之下，即 $r(\pi_0, \delta_{\pi_0'})$ 的最大值小于 $r(\pi_0, \delta)$。又因为 $r(\pi_0, \delta_{\pi_0'})$ 和 $V(\pi_0)$ 在点 $\pi_0 = \pi_0'$ 处相切，所以 $\delta_{\pi_0'}$ 是给

定先验概率 π'_0 的贝叶斯准则。显然，对于任意判决方法 δ，都可以画出类似的正切线(低于两个条件风险)，因此只有贝叶斯准则才有可能是极小化极大准则。此外，通过分析图 2.3.2，在 V 有内部最大值时，极小化极大准则就是在 $\pi_0\in[0,1]$上最大化 $V(\pi_0)$的先验概率 π_L 所对应的贝叶斯准则。注意，对于这个先验概率，易知 $r(\pi_0,\delta_{\pi_L})$是关于 π_0 的常量，即 $\max\{R_0(\delta_{\pi_L}),R_1(\delta_{\pi_L})\}=R_0(\delta_{\pi_L})=R_1(\delta_{\pi_L})$(条件风险相等的判决方法称为均衡准则)。事实上，从图 2.3.2 易知，δ_{π_L} 是极小化极大值。因为如果 $\pi'_0<\pi_L$，有 $\max\{R_0(\delta_{\pi'_0}),R_1(\delta_{\pi'_0})\}=R_0(\delta_{\pi'_0})>R_0(\delta_{\pi_L})$；如果 $\pi''_0>\pi_L$，有 $\max\{R_0(\delta_{\pi''_0}),R_1(\delta_{\pi''_0})\}=R_1(\delta_{\pi''_0})>R_1(\delta_{\pi_L})$。

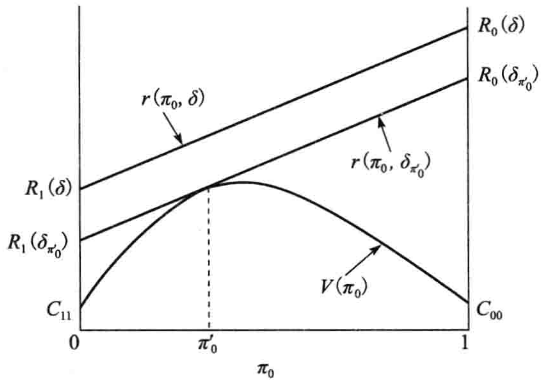


图 2.3.1 函数 $r(\pi_0,\delta)$和 $V(\pi_0)$

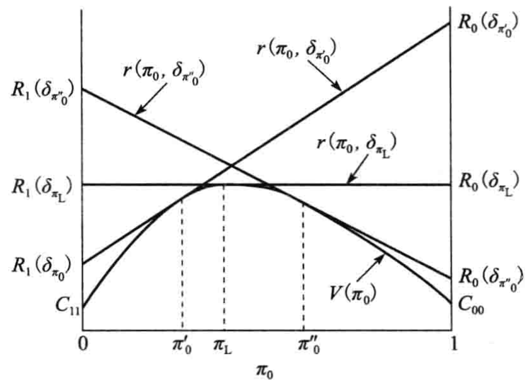


图 2.3.2 当 V 有内部最大值时的极小化极大值

图 2.3.2 中，π_L 对应极大化极小贝叶斯风险，因此也称为最不利先验概率。此时，极小化极大判决方法是关于最不利先验概率的贝叶斯准则。虽然在以上讨论中，没有考虑 $\max V(\pi_0)(0\leqslant\pi_0\leqslant1)$可能会出现在 $\pi_0=0$ 或 $\pi_0=1$ 的情况，也不能保证 $V(\pi_0)$在每一点上都有直线与之相切(即该函数不是在所有点上可微)。但是，极小化极大准则几乎总是关于最不利先验概率的贝叶斯准则。下面推导极小化极大假设检验问题的一般解。首先讨论如下命题，其中 $V(\pi_0)$如图 2.3.1 和图 2.3.2 所示，即 $\pi_L=0$ 或 $\pi_L=1$。

命题 2.3.1 极小化极大检验

假设 π_L 是式 $V(\pi_L)=\max V(\pi_0)(0\leqslant\pi_0\leqslant1)$的解，假设 $\pi_L=0$ 或 $\pi_L=1$ 或 $R_1(\delta_{\pi_L})=R_0(\delta_{\pi_L})$，则 δ_{π_L} 是极小化极大准则。

证明 首先，考虑 $R_1(\delta_{\pi_L})=R_0(\delta_{\pi_L})$的情况。对于任何先验概率 π_0，有

$$\max_{0\leqslant\pi_0\leqslant1}\min_{\delta} r(\pi_0,\delta)=r(\pi_L,\delta_{\pi_L})=r(\pi_0,\delta_{\pi_L}) \tag{2.3.4}$$

其中，第一个等式利用了 V 和 π_L 的定义，第二个等式利用了 $r(\pi_0,\delta_{\pi_L})$在 π_0 处是常数这一事实。进而可以得到

$$\max_{0\leqslant\pi_0\leqslant1}\min_{\delta} r(\pi_0,\delta)=\max_{0\leqslant\pi_0\leqslant1} r(\pi_0,\delta_{\pi_L})\geqslant\min_{\delta}\max_{0\leqslant\pi_0\leqslant1} r(\pi_0,\delta) \tag{2.3.5}$$

对于每个 δ 而言，都有

$$\max_{0\leqslant\pi_0\leqslant1} r(\pi_0,\delta)\geqslant\max_{0\leqslant\pi_0\leqslant1}\min_{\delta} r(\pi_0,\delta) \tag{2.3.6}$$

这意味着

$$\min_{\delta}\max_{0\leqslant\pi_0\leqslant1} r(\pi_0,\delta)\geqslant\max_{0\leqslant\pi_0\leqslant1}\min_{\delta} r(\pi_0,\delta) \tag{2.3.7}$$