信息与系统

下 册

陈鸿彬 马世雄 程佩青 张澄波 宫兆祥

合 编

国 防 工 业 出 版 社

内 容 简 介

本书是工科无线电技术专业的专业基础理论课《信息与系统》的试用教材。内容包括信号理论（信号分析与综合），信号检测，离散系统，香农理论，及调制理论。本书部分内容可作研究生教材。本书也可作为无线电工程技术人员参考书。本书由陈鸿彬主编。赵华孟，肖国镇，梁传甲主审。由西安交通大学无线电通信教研室闻懋生等担任编辑。

目　　录

第六章　信号检测导论

第七章　卡尔曼滤波

第八章 离散线性变换

第九章　数字滤波器

第十章 通信系统

附录 B 线性差分方程

第六章　信号检测导论

§6.1　绪　　言

提要　由于存在随机干扰，所以在接收一测量到观测数据后，要进行信号提取的处理工作，这广义称为信号的检测。信号检测问题包括三个方面：①狭义的信号检测，即，对于信号的有限的几个可能状态的分辨，或在干扰背景之中信号是否存在的判断；②连续取态的信号参量的估计；③连续时间信号波形的最优复现。

在任何科学和技术领域中，人们总要测量一些系统中某组有意义的综合状态量，然后将它们传输，并且最后用之来控制某些事物。但是，由于热骚动干扰是客观存在的、器件机理固有的噪声或仪器的种种不完善性造成的随机干扰是客观存在的，以及有时存在人为的干扰因素，等等，在对于人们所关心的各种状态量的量测、传输和利用过程中，总难免发生待测状态量(称之为信号)被干扰沾污（迭加）的问题。在存在干扰的情况下要使得接收一测量到的观测值能正确执行控制，首先要在信号与干扰的迭加中提取信号或提取它所携带的消息，这是处理学科的重要方面之一。信号检测理论研究在存在干扰的情况下怎样以最优、最合理的方式提取信号或由它所携带的消息。

信号检测是一个广义的术语。它包括几个方面：①狭义的信号检测：在有干扰的情况下从接收一测量到的输入中判断其中所含信号系属于几个可能状态中的哪一个；它的最狭义理解是判断输入中是否存在信号。②参量估计：在有干扰的情况下对于接收-测量到的输入按某种准则，以最小的误差和损失将所含信号或信号所携带的消息的某些参量予以估计。③信号过滤：在有干扰的情况下从接收-测量到的时间连续波形中将信号波形以最小的误差连续地过滤出来。信号的连续过滤也称为信号波形的最优复现。

信号检测问题的解决方式归之于建立一个检测-处理系统。向这个系统输入的是受到干扰的信号，也可能是在一段时间内的连续观测波形记录。检测系统以这些输入作为数据基础，予以合理的加工组合，然后输出最优的结果。构成一个检测系统，需要对于信号及干扰的统计特性有事先的知识，以及对于当输出最后起了控制作用后的经济等方面后果的统计数据知识。广义地说，信号处理系统可泛称为过滤器。这种广义的过滤器并不一定具有通常称之为滤波器的形式。例如，我们对某一系统的某一状态量进行了二次在0均值干扰迭加下的量测，得到了二个数据。在二次观测时间的间距够大的情况下，一种良好的关于该待测状态量的估计值是将二次观测值各乘以1/2的权，然后相加。这个简单的加权加法装置就是一个过滤器。可见，它并不具有通常称之为滤波器的构成形式。当然，这个目的也可以用一个通常称之为滤波器的形式的电路（例如一个积分器）来实现。

本章将沿着前面提到的狭义信号检测→信号估计→信号过滤的发展顺序讨论这些问题。

§6.2 二种取态的信号的检测

提要 二种取态的信号集合是最广泛使用的信号集合。信号受到干扰后，就不能从观测值完全正确地判断信号的真实状态。根据一个接收-测量值，只能作出信号真实状态是二者之一或二者之另一这样的假设。一种假设成立而测得某一测量值的概率与另一种假设成立而测得同一个测量值的概率之比称为该接收-测量值的似然比，它定量地说明该接收-测量值更象(比较"似然")是由哪一个信号状态所产生。人们根据接收-测量值的似然比来合理地推断信号的真实状态。当似然比大于某边界值，判断信号属某一状态，当似然比小于该边界值，判断信号属于另一状态。

在很多情况下，信号是量化取态的，且在不少情况下取态数为2。根据接收到的信号究竟属于这二个状态中的哪个状态，是此种工作情况下信号检测的总任务。有的作者称这种二元检测问题为双择一检测问题。

信号的量化状态(取量化值 s_1 或 s_2)迭加上连续取态的干扰 n，构成检测系统的输入 x。x 叫作接收-测量值。当 n 的取值有一个可以根据其形成物理机理及条件计算得到的或者可以从过去的经验统计得到的分布特性时，就可分别求得 $x=s_1+n$ 或 $x=s_2+n$ 的条件分布特性。在二元(二态)检测时，$s(=s_1$ 或 $s_2)$ 固然是量化取值的随机变量，但是当 n 是连续取值的随机变量时，x 就成为连续取值的随机变量。接收-测量者在读得一个 x 值后，并不能肯定这个 x 所含的信号成分究竟是 s_1 还是 s_2，因为在 n 的作用下，不论 $s=s_1$ 或 $s=s_2$，都有可能被 n 沾污而成为同一个 x。接收-测量者只能作出一种假设：认为信号取二种可能状态之一，并根据信号取不同态时 x 的取值的统计特性来检验这种假设是否最合理或者最有利。以下简称认为 s 等于 s_1 的假设为假设 H_1，称认为 s 等于 s_2 的假设为假设 H_2。

在这个对于所作的假设检验中，要涉及当 $s=s_1$ 时 x 的取值条件概率密度函数 $w_1(x)=w(x|s_1)$ 及当 $s=s_2$ 时 x 的取值条件概率密度函数 $w_2(x)=w(x|s_2)$。这二个概率密度函数的比，记为

$$\Lambda(x)\triangleq\frac{w_2(x)}{w_1(x)} \tag{6-2-1}$$

称为随机变量 x 究竟起因于 $s=s_2$ 还是 $s=s_1$ 的似然比函数。对于一个观测值 x，有分别的条件概率密度 $w_1(x)$ 和 $w_2(x)$，所以，一个 x 的对应 $\Lambda(x)$ 代表着该 x 究竟起源于 $s=s_1$ 还是 $s=s_2$ 的条件概率密度之比。如果对某一 x 有 $\Lambda(x)=1$，则造成这个 x 的 s 是 s_2 或是 s_1 的概率(密度)是相等的；如果对某一个 x 有 $\Lambda(x)>1$，则造成这个 x 的 s 是 s_2 的概率(密度)大于 s 是 s_1 的概率(密度)，代表着这个 x 比较可能是(比较象是)由 $s=s_2$ 所引起。反之亦反。由此，$\Lambda(x)$ 被称为似然比函数，条件概率密度 $w_1(x)$ 和 $w_2(x)$ 分别又称为在测量得到 x 的基础上 s 等于 s_1 或 s_2 的似然函数。就是说，在发信号者的立场看，$w_1(x)$ 及 $w_2(x)$ 是条件概率密度函数，而在接收者的立场上，$w_1(x)$ 及 $w_2(x)$ 是 x 起因于 $s=s_1$ 或 $s=s_2$ 的似然函数。

当然，当某个 x 的 $\Lambda(x)>1$，并不意味着这个 x 肯定起源于 $s=s_2$。$\Lambda(x)>1$ 只代表：这个 x 比较象是 H_2 成立而已，反之也是如此。

由上面讨论，我们可以以 $\Lambda=1$ 为边界，作为在接收-测量到 x 时，判断究竟 H_1 还是 H_2 成立的依据。

如果 $\Lambda(x)$ 是 x 的单调函数，那么对于一个 Λ 值，只有一个 x 值。与边界 $\Lambda(x)=1$ 相对应的 x 值，称为对 x 进行判断的门槛值，记之为 x_0。如果 $\Lambda(x)$ 随 x 增大而增大，那么 $\Lambda(x)>1$ 与 $x>x_0$ 相当，$\Lambda(x)<1$ 与 $x<x_0$ 相当，当接收-测量到的 $x>x_0$，H_2 较可能成立，反之亦反。

由于中心极限定理，干扰取 n 值的概率密度常作正态分布：

$$w(n)=\frac{1}{\sqrt{2\pi}\,\sigma_n}\exp\left[-\frac{n^2}{2\sigma_n^{\,2}}\right] \tag{6-2-2}$$

这里 $\sigma_n^{\,2}$ 是 n 的均方值。因为 $w(n)$ 不依赖于 s 的取值，所以当 $s=s_1$，观测值取值 x 的概率密度为

$$w_1(x)=\frac{1}{\sqrt{2\pi}\,\sigma_n}\exp\left[-\frac{(x-s_1)^2}{2\sigma_n^{\,2}}\right] \tag{6-2-3}$$

当 $s=s_2$，观测值取 x 的概率密度为

$$w_2(x)=\frac{1}{\sqrt{2\pi}\,\sigma_n}\exp\left[-\frac{(x-s_2)^2}{2\sigma_n^{\,2}}\right] \tag{6-2-4}$$

因而，在正态分布的干扰下，有似然比函数

$$\begin{aligned}\Lambda(x)=\frac{w_2(x)}{w_1(x)}&=\exp\left[\frac{2x(s_2-s_1)-(s_2^{\,2}-s_1^{\,2})}{2\sigma_n^{\,2}}\right]\\&=\exp\left[-\frac{(s_2^{\,2}-s_1^{\,2})}{2\sigma_n^{\,2}}\right]\exp\left[\frac{x(s_2-s_1)}{\sigma_n^{\,2}}\right]\end{aligned} \tag{6-2-5}$$

它是 x 的单调函数。这是常常见到的情况。不难看出，当 $s_2>s_1$，此 $\Lambda(x)$ 随 x 而单调增加。当然，如果干扰的分布并不是高斯分布(6-2-2)，则 $\Lambda(x)$ 就会取其他形式，也有可能 $\Lambda(x)$ 是 x 的非单调函数。

事实上，二元检测问题并不总是可以简单地以 $\Lambda(x)=1$ 为判决的界线。因为，如果事实上发方根本不发 $s=s_1$ 而永远发 $s=s_2$，那么，接收-测量者同样既有可能接收-测量到 $x>x_0$，也有可能接收-测量到 $x<x_0$（这是因为干扰的作用）。在这种情形，如果接收-测量者以 x_0（它与 $\Lambda(x)=1$ 对应）为判决的门槛，一方面，当然将有 $x>x_0$ 的机会较多，但是也有出现 $x<x_0$ 的可能。因此，如果接收-测量者已知发方根本不可能发 $s=s_1$，而还要根据这样的门槛来进行判决，那就十分不合理的了。根本不可能出现 $s=s_1$ 这种情况代表先验概率特性 $p(s=s_1)=0$，同时 $p(s=s_2)=1$。

因此，从接收-测量值 x 判决 H_1 还是 H_2 成立，应该考虑到发方的先验概率特性。因为 $p(s_1)+p(s_2)=p_1+p_2=1$，所以这个先验概率特性可以用一个 p_1 来概括地描述。考虑到先验概率特性后，显然将导致对于门槛的修正。例如，如果 $p_1=p(s=s_1)=0$，则应该选门槛为 $-\infty$，以使得在任何 x 时，总有 x 大于门槛值，从而可认为 H_2 总是成立。

事实上，门槛值的决定，还要基于更多的综合的因素和前提。先验概率特性只是因素和前提的一个方面。试考虑以下情况：如果将 $x=s_1+n$ 错判为 $s=s_2$，要对出一定的错判代价，而将 $x=s_2+n$ 错判为 $s=s_1$ 也要对出另一种错判代价（在多数情况下，这二种错判代价并不一样），则门槛的最优选定，应该考虑到使得各次判决的平均代价最小。举例说，如果

$s_2>s_1$，而且 $\Lambda(x)$ 是 x 的单调函数，那么，当 $x=s_1+n$ 错判成 $s=s_2$ 的代价低于把 $x=s_2+n$ 错判为 $s=s_1$ 的代价，我们应该选门槛较低，以使得判决倾向于 $s=s_2$，从而使得统计平均代价最小。

总上来说，接收-测量者在策略上的任务是，选定一个最合理的 $\Lambda(x)$ 临界值，并相应地选定 x 门槛值。从以上显然可见，当综合因素和前提不同，将有不同的边界 $\Lambda(x)$ 和不同的门槛值。根据不同的综合因素和前提而选定不同的 $\Lambda(x)$ 边界以及不同的 x 门槛值，是因为在不同的因素和前提下，有不同的最优策略因而有不同的判断准则，下面将分节予以讨论。

§ 6.3 贝叶斯(Bayes)准则

提要 当信源(发方)吐出信号二个状态的先验概率特性是已知并且是稳定的，并且当判决正确或错误的统计平均后果(代价)是已知的，那么确定似然比边界的策略原则是应该使得平均代价最低。相应的准则称之为最小平均风险准则，或贝叶斯准则。按贝叶斯准则而确定的边界似然比，正就是使得二种假设的分别条件平均风险相等时的似然比。

使作出判决的平均代价最小的判决准则称为最小平均风险准则，或称为贝叶斯准则。根据最小平均风险的策略原则所定的似然比函数边界称为贝叶斯准则边界，相应的 x 门槛称为贝叶斯准则门槛。

所谓风险，就是对于一个 x 作出某种判决后因为有错误而要付出的代价，代价大称为风险大，代价小称为风险小。从一个 x，可以判 s 为 s_1 或 s 为 s_2。对于 $s=s_j$ 而判为 s_i，则相应的代价记为 c_{ij}。对二态检测一共有四种 c_{ij}：c_{11}、c_{12}、c_{21} 和 c_{22}。不妨可以用一个 2×2 方阵$[c]$表示：

$$[c] \triangleq \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \tag{6-3-1}$$

c_{11} 是发生 $s=s_1$ 而判决为 H_1 成立的代价。这既然是正确的判决，故常有

$$c_{11}=0$$

同理，也常有

$$c_{22}=0$$

c_{12} 是 $s=s_2$ 而错判为 H_1 成立的代价，c_{21} 是 $s=s_1$ 而错判为 H_2 成立的代价。c_{12} 和 c_{21} 通常都不为零。于是，常有

$$[c] = \begin{bmatrix} 0 & c_{12} \\ c_{21} & 0 \end{bmatrix} \tag{6-3-2}$$

一般来说，c_{12} 并不一定等于 c_{21}。

令 $s=s_1$ 而正确判为 H_1 的概率记为 Q_{11}。当 $s_2>s_1$，它可以根据 $w_1(x)$ 计算如下：

$$Q_{11}=\int_{-\infty}^{x_0} w_1(x)dx \tag{6-3-3}$$

x_0 是判决的门槛。当 $x>x_0$，判 H_2 成立；当 $x<x_0$，判 H_1 成立。这个 x_0 是有待选定的量，它应这样选定，使得平均风险最小。

令 $s=s_2$ 而错判为 H_1 成立的概率记为 Q_{12}，在上述 $s_2>s_1$ 的情况下，它可以根据 $w_2(x)$ 计算如下：

$$Q_{12}=\int_{-\infty}^{x_0} w_2(x)dx \tag{6-3-4}$$

$s=s_1$ 而错判为 H_2 成立的概率记为 Q_{21}，它可以根据 $w_1(x)$ 计算如下：

$$Q_{21}=\int_{x_0}^{\infty} w_1(x)dx \tag{6-3-5}$$

同样，$s=s_2$ 而正确判为 H_2 成立的概率记为 Q_{22}，它可以根据 $w_2(x)$ 计算如下：

$$Q_{22}=\int_{x_0}^{\infty} w_2(x)dx \tag{6-3-6}$$

容易理解，随着 x_0 增大，

$$\left.\begin{array}{lll} Q_{11} & \text{增大:} & \left(\dfrac{dQ_{11}(x_0)}{dx_0}>0\right) \\ Q_{12} & \text{增大:} & \left(\dfrac{dQ_{12}(x_0)}{dx_0}>0\right) \\ Q_{21} & \text{减小:} & \left(\dfrac{dQ_{21}(x_0)}{dx_0}<0\right) \\ Q_{22} & \text{减小:} & \left(\dfrac{dQ_{22}(x_0)}{dx_0}<0\right) \end{array}\right\} \tag{6-3-7}$$

但

如果信源给出（发送）s_1，有可能被正确判决也有可能被错误判决，所以，发送 s_1 的统计平均代价（称之为发状态 s_1 的平均风险）记为 $\overline{C}_1$，为

$$\overline{C}_1=C_{11}Q_{11}+C_{21}Q_{21} \tag{6-3-8}$$

送 s_2 同样有可能被正确或错误地判决，送 s_2 的统计平均代价（称之为发状态 s_2 的平均风险），记为 $\overline{C}_2$，为

$$\overline{C}_2=C_{12}Q_{12}+C_{22}Q_{22} \tag{6-3-9}$$

$\overline{C}_1$ 和 $\overline{C}_2$ 都决定于同一 x_0。

如果 $s=s_1$ 的先验概率为 P_1，则 $s=s_2$ 的先验概率为

$$P_2=1-P_1 \tag{6-3-10}$$

于是，选 x_0 作为门槛的每次判决的统计平均风险为

$$\begin{aligned}\overline{C}&=P_1\overline{C}_1+P_2\overline{C}_2=P_1\overline{C}_1+(1-P_1)\overline{C}_2\\ &=P_1\left[C_{11}\int_{-\infty}^{x_0} w_1(x)dx+C_{21}\int_{x_0}^{\infty} w_1(x)dx\right]\\ &\quad+(1-P_1)\left[C_{12}\int_{-\infty}^{x_0} w_2(x)dx+C_{22}\int_{x_0}^{\infty} w_2(x)dx\right]\end{aligned} \tag{6-3-11}$$

上式中的 P_1，$[C]$，$w_1(x)$ 及 $w_2(x)$ 有时都是可以予先知道的。显然，$\overline{C}$ 是 x_0 的函数。当 x_0 取某值 x_{00} 时，可有 $\overline{C}$ 取一个最小值，这个 x_{00} 就是按照最小平均风险策略原则的贝叶斯准则门槛似然比值。

因为当 x_0 增加一个 $dx_0>0$ 时，有

$$\begin{aligned}d\overline{C}(x_0)=&P_1[C_{11}w_1(x_0)dx_0-C_{21}w_1(x_0)dx_0]+\\ &+(1-P_1)[C_{12}w_2(x_0)dx_0-C_{22}w_2(x_0)dx_0]\end{aligned}$$

因为当 $x_0=x_{00}$ 时应有

$$\left.\frac{d\overline{C}}{dx_0}\right|_{(x_0=x_{00})}=0$$

得：

$$P_1[C_{11}w_1(x_{00})-C_{21}w_1(x_{00})]+(1-P_1)[C_{12}w_2(x_{00})-C_{22}w_2(x_{00})]=0$$

因而有：

$$\boxed{\Lambda(x_{00})=\frac{w_2(x_{00})}{w_1(x_{00})}=\frac{P_1(C_{21}-C_{11})}{(1-P_1)(C_{12}-C_{22})}} \tag{6-3-12}$$

于是，在已知 P_1 及[c]情况下，可以方便地求得 $\Lambda(x_{00})$，并且在已知 $w_1(x)$ 及 $w_2(x)$的情况下，可以从而求得 x_{00}，如果 $w(n)$作正态分布，均方值为 σ_n^2，则按(6-2-5)有

$$\Lambda(x_{00})=\exp\left[-\frac{s_2^2-s_1^2}{2\sigma_n^2}\right]\exp\left[\frac{x_{00}(s_2-s_1)}{\sigma_n^2}\right] \tag{6-3-13}$$

因而，

$$\boxed{x_{00}=\sigma_n^2\ln\left\{\frac{\Lambda(x_{00})}{\exp\left[-\frac{s_2^2-s_1^2}{2\sigma_n^2}\right]}\right\}\Big/(s_2-s_1)} \tag{6-3-14}$$

或，

$$\boxed{x_{00}=\frac{\sigma_n^2}{s_2-s_1}\ln\left[\frac{P_1(C_{21}-C_{11})}{(1-P_1)(C_{12}-C_{22})}\right]+\frac{s_2+s_1}{2}} \tag{6-3-15}$$

举例说，若有一个二进制数字通信系统，s的二个可能状态为 $s_1=0$ 及 $s_2=1$，$P_1=\frac{1}{2}$（出现0与1等概率）且

$$[c]=\begin{bmatrix}0 & k\\ k & 0\end{bmatrix}$$

（二种错判的代价相同），可计算得贝叶斯门槛 x_{00} 应为 $\frac{1}{2}$。

在正确选用了门槛 x_{00} 后，$\overline{C}$ 成为最小可能值，记为 $\overline{C}_{\min}$：

$$\begin{aligned}\overline{C}_{\min}=&P_1[C_{11}Q_{11}(x_{00})+C_{21}Q_{21}(x_{00})]\\&+(1-P_1)[C_{12}Q_{12}(x_{00})+C_{22}Q_{22}(x_{00})]\end{aligned} \tag{6-3-16}$$

当 $w(n)$作高斯分布均方值为 σ_n^2，且 $s_2>s_1$，则

$$Q_{11}(x_{00})=\int_{-\infty}^{x_{00}}w_1(x)dx=1-\mathrm{erfc}\ (x_{00}-s_1) \tag{6-3-17}$$

$$Q_{21}(x_{00})=\int_{x_{00}}^{\infty}w_1(x)dx=\mathrm{erfc}\ (x_{00}-s_1) \tag{6-3-18}$$

$$Q_{12}(x_{00})=\int_{-\infty}^{x_{00}}w_2(x)dx=1-\mathrm{erfc}\ (x_{00}-s_2) \tag{6-3-19}$$

$$Q_{22}(x_{00})=\int_{x_{00}}^{\infty}w_2(x)dx=\mathrm{erfc}\ (x_{00}-s_2) \tag{6-3-20}$$

各式中，$\mathrm{erfc}(y)=\frac{1}{\sqrt{2\pi}}\int_y^{\infty}e^{-x^2/2}dx$。于是，在高斯干扰下，在正确选用了门槛 x_{00} 后，相应的最小平均风险为

$$\overline{C}_{min}=P_1\{C_{11}[1-\mathrm{erfc}(x_{00}-s_1)]+C_{21}\mathrm{erfc}(x_{00}-s_1)\}$$
$$+(1-P_1)\{C_{12}[1-\mathrm{erfc}(x_{00}-s_2)]+C_{22}\mathrm{erfc}(x_{00}-s_2)]\} \qquad (6\text{-}3\text{-}21)$$

此式中 x_{00} 按式(6-3-15)选取其值。由此，$\overline{C}_{min}$ 归根结底是 P_1，$[C]$，σ_n^2 及 s_1、s_2 的函数。如将(6-3-15)代入(6-3-21)可见 $\overline{C}_{min}$ 是 P_1，$[C]$ 及 $\frac{s_2-s_1}{\sigma_n}$ 的函数。当$[C]$及 $\frac{s_2-s_1}{\sigma_n}$ 给定后，$\overline{C}_{min}$ 就只决定于 P_1。举例说，当

$$[C]=\begin{bmatrix} 0 & 2k \\ k & 0 \end{bmatrix}$$

$$\left.\begin{aligned} s_1&=-1 \\ s_2&=+1 \\ \sigma_n^2&=4 \end{aligned}\right\} \qquad \frac{s_2-s_1}{\sigma_n}=1;$$

根据(6-3-15)有

$$x_{00}=\frac{4}{2}\ln\left[\frac{P_1k}{(1-P_1)2k}\right]+0=2\ln\left[\frac{P_1}{2(1-P_1)}\right]$$

从而可以计算得 $\overline{C}_{min}(P_1)$。此例中，$\overline{C}_{min}(P_1)$ 作为 P_1 的函数，如图 6-3-1 所示。$\overline{C}_{min}$ 在

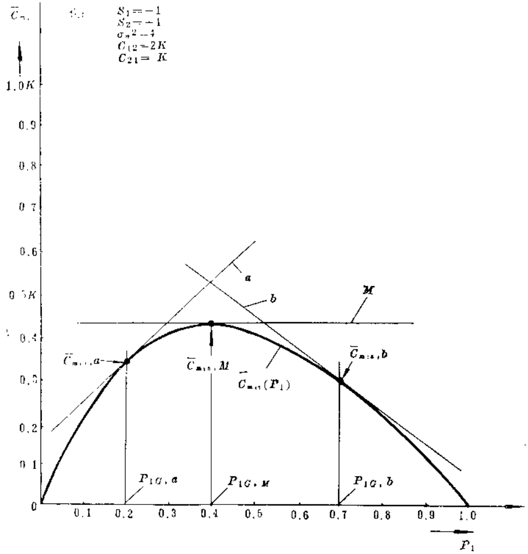

图 6-3-1 说明贝叶斯准则系统及最小最大准则系统

$P_1=0$ 及 $P_1=1$ 时都为 0，而在某一 P_1 值时有其最大值。此例中，$\overline{C}_{\min}$ 的最大值发生在 $P_1=0.4$，相应的 $\overline{C}_{\min}(P_1=0.4)=[\overline{C}_{\min}]_{\max}=0.42K$。

在有些情况下，例如，在博奕或战争中，P_1 被掌握在斗争的对方，可以由对方主观地被改变。

最小平均风险策略准则（贝叶斯准则）只能在 P_1 已知的情况下有意义。

贝叶斯准则的似然比边界以及判决门槛，也可以从下面的讨论中得到。

当我们观察到某 x 的基础上而作假设 H_1，由于此假设可能符合也可能不符合真正的 S 的状态，所以有一个平均风险，称为作假设 H_1 的条件平均风险，记为 $\overline{C}(H_1/x)$。同样，在观察到 x 的基础上而作假设 H_2，有一个作假设 H_2 的条件平均风险，记为 $\overline{C}(H_2/x)$。

根据贝叶斯定理，出现 x 并且 H_1 符合事实的概率密度为

$$w(x)P(H_1/x)=P_1w_1(x) \tag{6-3-22}$$

出现 x 并且 H_2 符合事实的概率密度为

$$w(x)P(H_2/x)=P_2w_2(x)=(1-P_1)w_2(x) \tag{6-3-23}$$

因而，在出现 x 的基础上 H_1 成立的概率是

$$P(H_1/x)=\frac{P_1w_1(x)}{w(x)} \tag{6-3-24}$$

在出现 x 的基础上 H_2 成立的概率是

$$P(H_2/x)=\frac{(1-P_1)w_2(x)}{w(x)} \tag{6-3-25}$$

当然，

$$P(H_1/x)+P(H_2/x)=1 \tag{6-3-26}$$

$P(H_1/x)$ 和 $P(H_2/x)$ 是在发生 x 的基础上 H_1 或 H_2 成立的后验概率。

由此，

$$\begin{aligned}\overline{C}(H_1/x)&=C_{11}P(H_1/x)+C_{12}P(H_2/x)\\&=C_{11}\frac{P_1w_1(x)}{w(x)}+C_{12}\frac{(1-P_1)w_2(x)}{w(x)}\end{aligned} \tag{6-3-27}$$

同理，

$$\begin{aligned}\overline{C}(H_2/x)&=C_{22}P(H_2/x)+C_{21}P(H_1/x)\\&=C_{22}\frac{(1-P_1)w_2(x)}{w(x)}+C_{21}\frac{P_1w_1(x)}{w(x)}\end{aligned} \tag{6-3-28}$$

$\overline{C}(H_1/x)$ 和 $\overline{C}(H_2/x)$ 分别称为在观测结果为 x 的条件下，H_1 成立的条件平均风险及 H_2 成立的条件平均风险。

接收-测量者在一个观测值 x 下作出判决时，希望平均风险最小，所以，他在 $\overline{C}(H_1/x)$ 及 $\overline{C}(H_2/x)$ 二者中看哪个小而选相应的假设。因此，当正好发生 $x=x_{00}$ 时，应有 $\overline{C}(H_1/x=x_{00})$ 与 $\overline{C}(H_2/x=x_{00})$ 相等，于是，对于 $x=x_{00}$，成立等式：

$$\begin{aligned}&C_{11}\frac{P_1w_1(x_{00})}{w(x_{00})}+C_{12}\frac{(1-P_1)w_2(x_{00})}{w(x_{00})}\\&=C_{21}\frac{P_1w_1(x_{00})}{w(x_{00})}+C_{22}\frac{(1-P_1)w_2(x_{00})}{w(x_{00})}\end{aligned}$$

从而有

$$A(x_{00}) = \frac{w_2(x_{00})}{w_1(x_{00})} = \frac{P_1(C_{21}-C_{11})}{(1-P)(C_{12}-C_{22})} \tag{6-3-29}$$

这就是公式(6-3-12)。因此，贝叶斯准则的等效叙述是，二种假设的条件平均风险相等，这是一个容易记住的结论。

贝叶斯准则有较广泛的实用意义。

§6.4 "安全"平均风险准则（极小极大准则）

提要 当信号的二个状态的先验概率特性不能知道时，贝叶斯策略就失去一个统计知识基础。这时，接收-测量者可以设法猜测一种先验概率特性。如果这种猜测不符合实际情况，可能接收-测量者要付出非常大的平均代价。为了使得这种平均代价被限制在不过分大的范围之内，使得接收-测量者可以"保险"不付出过分大的平均代价，相对地得到一种"安全稳定"的保证，接收-测量者应该合理地猜测某一个先验概率特性。这个合理的猜测，与在贝叶斯准则下所得的最小平均风险作为先验概率的函数有关。所得的结果，称为极小极大（最小最大）准则的结果。这个结果的似然比边界，正是使得信号分别取二种状态而分别的平均风险相等时的门槛值的相应似然比。

在不少的情况下，(出现 $s=s_1$ 的)先验概率 P_1 及 $P_2=1-P_1$ 是不知道的。在这种情况下，为了追求最小平均风险，接收-测量者不妨可以猜测或假定一个 P_1，然后用贝叶斯准则的方法选定 x_{00}（及 $A(x_{00})$），从而谋求侥倖地可以付出最小平均风险来进行选定假设。但是，事实上是不可能总是侥倖猜中 P_1 的。从(6-3-12)式，在一个被猜测的 P_1(记为 $P_{1,G}$)下，就有相应的 $A(x_{00})$，因而根据 s_1 及 s_2 的值及已知的 $w(n)$ 及$[c]$就可求出相应的贝叶斯 x_{00}。从而，进一步可以求得全部 Q_{11}，Q_{12}，Q_{21} 及 Q_{22}。按(6-3-11)式，有相应的平均风险 $\overline{C}$ 为

$$\begin{aligned}\overline{C} = P_1&\left[C_{11}\int_{-\infty}^{x_{00}} w_1(x)dx + C_{21}\int_{x_{00}}^{\infty} w_1(x)dx\right]\\ &+(1-P_1)\left[C_{12}\int_{-\infty}^{x_{00}} w_2(x)dx + C_{22}\int_{x_{00}}^{\infty} w_2(x)dx\right]\end{aligned} \tag{6-4-1}$$

此式中 x_{00} 是在给定 $[C]$ 及 $P_1=P_{1G}$ 下确定的。今 P_{1G} 是一个被猜测的值，当然相应的 x_{00} 是一个与此猜测有关的值，记为 x_{00G}。因此，上式中两个方括号都是根据猜测而定的值。于是对于一个猜测之中的 P_{1G}，$\overline{C}$ 是 P_1 的直线函数［式(6-4-1)］。直线 $\overline{C}(P_1)$ 的斜率显然是(6-4-1)右方第一个方括号

$$\left[C_{11}\int_{\infty}^{x_{00G}} w_1(x)dx + C_{21}\int_{x_{00G}}^{\infty} w_1(x)dx\right]$$

减去第二个方括号

$$\left[C_{12}\cdot\int_{-\infty}^{x_{00G}} w_2(x)\cdot dx + C_{22}\cdot\int_{x_{00G}}^{\infty} w_2(x)\cdot dx\right]。$$

当接收-测量者猜测的 P_{1G} 不同，有不同的 x_{00G}，有不同的二个方括号，也因而有不同的直线

斜率。使用图 6-3-1 的例子，已知$[C]=\begin{bmatrix}0 & 2k\\ k & 0\end{bmatrix}$，如接收-测量者猜测 P_1 为 $P_{1G,a}=0.2$，则有$\overline{C}(P_1)$直线如图 6-3-1 中直线 a 所示；如猜测 P_1 为 $P_{1G,b}=0.7$，则有 $\overline{C}(P_1)$ 直线如图中直线 b 所示。容易理解，当 P_1 正好等于所猜测的 P_{1G} 时，对应于该 P_1 的 $\overline{C}(P_1)$ 等于已知 P_1 为该 P_{1G} 时的贝叶斯风险，因此，$\overline{C}(P_1)$ 直线必定在 $P_1=P_{1G}$ 点上与贝叶斯风险 $\overline{C}_{\min}(P_1)$曲线相接触，可参见图 6-3-1。此外，由于$\overline{C}_{\min}(P_1)$曲线的斜率为

$$
\begin{aligned}
\frac{d\overline{C}_{\min}}{dP_1}=&\{P_1[C_{11}w_1(x_{00})-C_{21}w_1(x_{00})]+(1-P_1)[C_{12}w_2(x_{00})\\
&-C_{22}w_2(x_{00})]\}\frac{dx_{00}}{dP_1}+\left[C_{11}\int_{-\infty}^{x_{00}}w_1(x)dx\right.\\
&\left.+C_{21}\int_{x_{00}}^{\infty}w_1(x)dx\right]-\left[C_{12}\int_{-\infty}^{x_{00}}w_2(x)dx+C_{22}\int_{x_{00}}^{\infty}w_2(x)dx\right]
\end{aligned}
$$

而贝叶斯风险的得来条件为

$$
\begin{aligned}
&P_1[C_{11}w_1(x_{00})-C_{21}w_1(x_{00})]+(1-P_1)[C_{12}w_2(x_{00})\\
&-C_{22}w_2(x_{00})]=0
\end{aligned}
$$

[见(6-3-12)式的导出过程]，因而，在$\overline{C}_{\min}(P_1)$曲线上必有

$$
\begin{aligned}
\frac{d\overline{C}_{\min}(P_1)}{dP_1}=&\left[C_{11}\int_{-\infty}^{x_{00}}w_1(x)dx+C_{21}\int_{x_{00}}^{\infty}w_1(x)dx\right]\\
&-\left[C_{12}\int_{-\infty}^{x_{00}}w_2(x)dx+C_{22}\int_{x_{00}}^{\infty}w_2(x)dx\right]
\end{aligned}
$$

这正就是在猜测一个 P_{1G} 时，直线$\overline{C}(P_1)$的斜率。因此，直线$\overline{C}(P_1)$不仅在 $P_1=P_{1G}$ 点与贝叶斯风险曲线$\overline{C}_{\min}(P_1)$相迁，而且二者在 $P_1=P_{1G}$ 点相切（参见图 6-3-1）。

从图(6-3-1)的直线 a 可见，如果接收-测量者猜测了 $P_{1G}=P_{1G,a}$，而它的对立面却用了 $P_1>P_{1G,a}$ 的话，将可能使得平均风险 $\overline{C}(P_1)$非常大；同样，如果接收-测量者猜测了 $P_{1G}=P_{1G,b}$，而它的对立面却用了 $P_1<P_{1G,b}$ 的话，也将可能有平均风险 $\overline{C}(P_1)$ 非常大。因此，在不能知道 P_1 的情况下，接收-测量者如果猜测 $P_{1G}=P_{1G,a}$ 或 $P_{1G}=P_{1G,b}$，将会冒相当大的风险。但是，如果接收-测量者所猜测的 P_{1G} 为对应于贝叶斯平均风险曲线极值点的 P_1 值——图 6-3-1 中的 $P_{1G,M}$，则对应于这个 $P_{1G,M}$ 猜测值，将有直线$\overline{C}(P_1)$为水平线。在这个情况，不论对立面选用什么 P_1，都恒有$\overline{C}$为对应于 $P_{1G,M}$ 点的$\overline{C}_{\min}$ 值。因此，接收-测量者的安全对策是选择贝叶斯风险$\overline{C}_{\min}$ 取最大值：$\overline{C}_{\min,M}$ 的 P_{1G} 值。在作其它对于猜测的选择时（例如猜测 $P_{1G,a}$ 或 $P_{1G,b}$），一方面固然有可能 $\overline{C}(P_1)$ 十分剧烈，但另一方面也有可能 $\overline{C}(P_1)$小于$\overline{C}_{\min,M}$ 并且还可能甚小于 $\overline{C}_{\min,M}$，然而，存在付出很大风险的危险。与之相比，接收-测量者选择 $P_{1G,M}$ 作为对策使得总有$\overline{C}(P)$等于且不大于贝叶斯风险的最大值$\overline{C}_{\min,M}$。这种猜测的风险对于接收-测量者来说，心中有数充其量为多少，因而是相对地最“安全”（保险）的。接收-测量者付出不贪图偶然可以得到很小的平均风险的代价以及对出最大的最小平均风险（贝叶斯风险）这种保守做法的“代价”，可以换取免于可能使得平均风险过分大的“安全”。

事实上，在对立面的立场上，对于 P_1 的选取策略的做法，也是选 $P_1=P_{1G,M}$，因为，如果他选取其它的 P_1（例如 $P_{1G,a}$），则当接收-测量者正好猜测了一个接近这个 P_1 的 P_{1G} 值时（按贝叶斯准则，他选了 x_{00}），接收方面付出的 $\overline{C}(P_{1G})$ 将比 $\overline{C}_{\min,M}$ 小，这是对立面