在噪声中信号的检测

A. D. 华伦（Whalen）

（上）

山　东　大　学

1978年

目　录(上册)

简单说明

本书是 Anthony D. Whalen（华伦）著的《在噪声中信号的检测》（Detection Signals in Noise）. 1971 年在美国出版（..cademic Press. Inc. New YorK）.

早在 50年代，信息论学科有了较快速的发展。当时形成了互相独立程度较大的两个学派。一派是以研究编码通讯而发展起来的，即所谓的狭义信息论的范畴，这一学派因此较注重于信息论的公理体系与基本原理，所以在后来的廿年实践中指明了除去完善信息论的基础理论和对编码方案方面有所发展以外，联系实际的进展是比较少的。但是这些内容是过去大学里"信息论"课的主要内容。另一个学派主要的是研究信号的检测、滤过、提取以及信号参量的估值等等，或者称为广义信息论，这一方面的研究是比较能联系实际的，为近代的通讯与雷达等方面的需要而有飞速的发展。

早期，人们对于通讯系统能够在实践中用要求信噪比最大为根据来进行设计。这当然是十分直观的，处于感性认识阶段，各种各样通讯系统的出现以及对有关系统所作的改善提出了要求在理论上总结提高的任务，因此发展了以概率论为基础的信息网络的理论与技术（如最佳滤波回与匹配滤波回以及相关技术）。我国在 1965年出版了"无线电技术基础资料汇编（第一集）"（高等教育出版社出版，常迥等编译），其中包含了上述的基础内容。

在 50年代末期，人们已经提出了对于以前发展起来的各种理论与见解有一个统一的认识，就是用数理统计学中的一套套理论（检验论、估值论以及统计决策理论等）来分析通讯系统与雷达系统，1960年在伦敦出版的 C. W. Helstrom："Statistical Theory of Signal Detection”一书是比较好的系统的一个总结。我国大约在1966年由国防出版社翻译出版了该书。（“信号检测的统计理论”）。

在上述基础上，在60年代，广泛地应用统计学的观点具体地联系了实际中很重要的问题，例如雷达中的多脉冲检测，有色噪声的问题和各种信号参量的估值问题，华伦（Whalen）的“在噪声中信号的检测”一书可以被认为是在此以前关于广义信息论方面基础理论的比较系统的总结。而且它还是一本可以被采用的入门性的参考书。譬如，最近 Harry L. Van Trees 从1968年到1972年写了四大厚本书（Detection, Estmation and Modulation Theory），其中总结的都是近期文献上的成果，有许多实际的雷达与通讯以及声纳系统的例子，但是没有入门阶段的予备知识往往不易看懂。又例如1969年出版的 D. K. Barton 与 H. R. Ward 著的“雷达测量手册”（Handbook of Radar Measurement）和 J. V. Difranco 与 W. L. Rubin 著的“雷达检测”（Radar Detection）等都是使用了信息论的概念，统计的分析以及有关的结果的。有了入门性的知识来参考这些资料就能比较顺利地了解了。最后，“在噪声中信号的检测”一书还结合大量实际通讯与雷达系统的最佳化方案的例子，有助于我们设计与鉴别各种通讯与雷达系统。

该书共分十一章，前四章介绍概率论的基础知识，第五章介绍了统计决策理论中的假设检验理论，第六章到第十章处理信号的检测，信号参量的估值问题。第十一章是比较抽象一些，那是为了便于参阅近代文献而加入的。

第一章 概率论*

§1－1 概率的简述

关于一事件（取出）的概率是直观地与事件发生的相对频率有联系的，例如，我们处理有A、B、C等等可能取出的一个实验（譬如一颗骰子的滚动）。我们重复做N次实验并记录每个事件发生的次数。分别记作n_A、n_B和n_C。事件发生的相对频率由比值n_A/N、n_B/N、n_C/N给定。当N趋于无穷时的极限情况下，我们说，这些比值趋于事件发生的真实概率。亦即

$$\lim_{N\to\infty} n_A/N \longrightarrow P(A)\ ,\quad \lim_{N\to\infty} n_B/N \longrightarrow P(B)\ ,\ 等等。$$

因此，概率是在0到1之间所含有的一个数。实际上我们不能绝对准确地确定这种概率。

一个实验的所有可能取出的集合称为实验的取样空间。一个事件可以是取样空间的一个单一组元或者可以是可能取出的一个集合体。在每一种情况下，一事件的发生或不发生都可以用实验的取出来确定。我们将用花括号内的符号来表示事件，例如，$\{A\}$是具有特性A的组元在取样空间内的子集合。

对于任何事件$\{A\}$，存在$\{非A\}$的事件，我们用$\{\overline{A}\}$来表示它。事件$\{A或\overline{A}\}$**为所有可能取出的集合（肯定事件）。事件

---

* 在Davenport与Root的书[1]和Papoulis的书[2]中可以找到很易读的概率论方法，在本章的结尾还引入了另外的参考书。

** 事件$\{A或B\}$意味着或者$\{A\}$，或者$\{B\}$或者两者都发生，事件$\{A和B\}$意味着$\{A\}$和$\{B\}$一起发生。

{A和$\bar{A}$}为无组元的集合（零事件），根据相对频率的定义，一个直接结果是肯定事件与零事件分别有一与零的概率。如果事件{A}与{B}互相排斥（亦即，如果一个发生，则另一个就不能发生），则对于事件{A或B}来说我们有P(A或B)$=P(A)+P(B)$。

假设做两个实验，它们的可能联出分别表示成A_i，$i=1,2,\cdots$与B_j，$j=1,2,\cdots$。那么，联合事件{A_i和B_j}是确定的。如同单一事件的情形一样，与这个联合事件相应的有一个概率，用$P(A_i,B_j)$表示这个联合事件的概率。如果A_i事件与B_j事件是可穷尽的，亦即，如果不可能有其它的事件，则[1]

$$\sum_i\sum_j P(A_i,B_j)=1$$

$$\sum_i P(A_i,B_j)=P(B_j),\qquad \sum_j P(A_i,B_j)=P(A_i).$$

§1—2　条件概率与统计独立

条件概率与另一事件已发生而给予某事件的知识有关，在事件{A}的发生已给定时，事件{B}发生的概率表示成$P(B|A)$。定义"已给{A}时{B}的"条件概率为

$$P(B|A)=P(A,B)/P(A)\qquad 只要P(A)\neq 0 \qquad\cdots\cdots(1-1)$$

类似地，已给{B}时{A}发生的概率为

---

1. 和式含有对指标的所有值取和的意义。

$$P(A|B)=P(A,B)/P(B). \quad [P(B)\neq 0]$$

$$\cdots\cdots\cdots\cdots (1-2)$$

如果实验是由互相排斥且可穷尽的取出B_i（$i=1,2,\cdots$）所组成的，则

$$\sum_i P(B_i|A)=1.$$

如果以事件$\{A\}$与$\{B\}$而言，有$P(A|B)=P(A)$，那么从条件概率的定义可得到

$$P(A,B)=P(A)P(B), \qquad (1-3)$$

还可得到$P(B|A)=P(B)$。因此，这些事件之一发生的知识并没有为其它事件发生的概率提供信息。这样一些事件就称为是<u>统计独立的</u>。

如果三个事件$\{A_1\}$、$\{A_2\}$与$\{A_3\}$是统计独立的，则必须满足关系式：

$$P(A_1,A_2)=P(A_1)P(A_2), \quad P(A_1,A_3)=P(A_1)P(A_3).$$

$$P(A_2,A_3)=P(A_2)P(A_3), \quad P(A_1,A_2,A_3)=P(A_1)P(A_2)P(A_3).$$

$$\cdots\cdots\cdots\cdots (1-4)$$

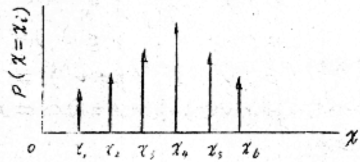
要使三个以上的事件是独立的，则须使每次取两个、三个、四个等等事件的概率等于这些单独事件的概率之乘积，

§1-3 概率分布函数

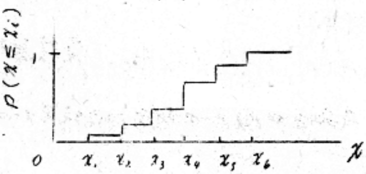
我们定义一<u>随机变量</u>或<u>随机的变化</u>为取样空间的一个实值函数，亦即一个实验的取出的实值函数。例如，掷一颗骰子，出现的点数是一个随机变量或随机地变化的。点数的任何一个函数也

是一个随机变量。如果随机变量假定可以取的数值（取样空间）是有限的或可数无穷的*，则就称它为离散随机变量。否则，它就是连续随机变量。

假设有一个离散随机变量**，它在任何一次可取，譬如说六个可能的数值 x_i，其中 $x_6 > x_5 > x_4 > x_3 > x_2 > x_1$，相应的概率表示为 $P(x_i)$ 或 $P(X=x_i)$，示于图1—1。对于这个例子来说，譬如提出随机变量取小于或等于 x_3 之值的概率是多少的问题是合理的，在这种情况下，$P(X \le x_3) = P(x_1) + P(x_2) + P(x_3)$。由 $P(X \le x)$ 定义的 x 的函数就称为随机变量X的概率分布函数。（它也称为分布函数与累积的分布函数）。图1—1还指明了上面例子的累积分布，取出 $\{X \le x\}$ 在一般概率意义上是一个事件，所以累积分布函数必须满足上



（A）



（B）

图1—1 离散随机变量的概率函数
（A）表示离散概率的脉冲函数
（B）累积分布函数

---

* 一个数的集合是可数的，只要它与正整数的集合建立一一对应关系。

** 在本章中，随机变量用大写字母表示，取样空间中一个元素用对应的小写字母表示。但是，使用不同符号来区分它们变得很麻烦，所以在后面几章中只用一个符号，只要在用时有清楚的解释。

已讨论过的特性。特别是，$P(X<-\infty)=0$ 与 $P(X<\infty)=1$ 还有，X落入区间 $x_i<X<x_j$ 的概率为 $P(X\leq x_j)-P(X\leq x_i)=P(x_i<X\leq x_j)$。

上述结果容易外推到两个随机变量（双变量分布）和多个随机变量（多变量分布）的情形，对于两个随机变量X与Y，它们可以是连续的或离散的，有

$$\left.\begin{aligned}
&P(X\leq-\infty,\ Y\leq y)=0,\\
&P(X\leq x,\ Y\leq-\infty)=0,\\
&P(X\leq\infty,\ Y\leq\infty)=1,\\
&P(X\leq x,\ Y\leq\infty)=P(X\leq x),\\
&P(X\leq\infty,\ Y\leq y)=P(Y\leq y),
\end{aligned}\right\}\qquad(1-5)$$

§1-4 连续随机变量

考虑具有如图1—2所示的连续累积分布函数的随机变量X．

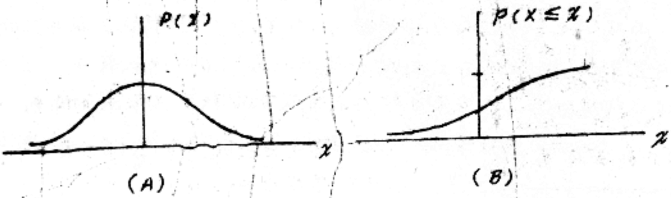


图1—2 连续随机变量的概率函数

(A) 密度函数

(B) 累积分布函数

这是连续随机变量的一个例子。这样一个随机变量可取不可数的值。例如取样空间可以是整个实数轴。对于连续随机变量而言，我们定义概率密度函数（或简称密度函数）为累积分布函数的导数，只要这个导数是存在的。用 $p(x)$ 表示随机变量 X 的概率密度函数，于是有

$$p(x)=dP(X\leq x)/dx \qquad (1-6)$$

详细确定一个密度函数必须包含密度函数可应用的范围。

如果函数 $P(X\leq x)$ 取定积分的形式，则能够进行微分的莱布尼茨法则[3]是：

如果 $\phi(x)=\int_{a(x)}^{b(x)} f(t,x)\,dt$，其中 $a(x)$ 与 $b(x)$ 是 x 的可微函数，并且 $f(t,x)$ 与 $\partial f(t,x)/\partial x$ 是 x 与 t 的连续函数，则

$$\frac{d\phi}{dx}=\int_{a(x)}^{b(x)}\frac{\partial f(t,x)}{\partial x}dt+f[b(x),x]\frac{db(x)}{dx}-$$

$$-f[a(x),x]\frac{da(x)}{dx}$$

因为累积分布函数是非减的，$p(x)\geq 0$，图 1-2 指出了概率密度函数的一个例子。用狄拉克 δ 函数⑴（脉冲函数），也可以对离散随机变量的概率密度函数定义为累积分布函数的导数，在图 1-1 所示的不连续点引入 δ 函数。例如在图 1-1 中，密度函

⑴ 狄拉克 δ 函数 $\delta(x)$ 是当它的变量为零时成无穷大并在其它地方处处为零，它有这样的面积 $\int_{-\infty}^{\infty}\delta(x)dx=1$，而且 $\int_{-\infty}^{\infty}f(x)\delta(x-a)dx=f(a)$。见 Lighthill[4] 对广义函数的讨论。

数可表示成

$$P(X)=\sum_{i=1}^{n} P(x)\delta(x-x_i)$$

一般地，随机变量可以是混合型的，并且累积分布函数中一部分具有跳跃的不连续性，而另一部分则是处处都连续的。这种随机变量的例子如图1-3所示。

根据概率密度函数的定义，立即得到

$$P(X\leq a)=\int_{-\infty}^{a} P(x)dx$$

量 $P(x)dx$ 可以解释为随机变量落入 x 与 $x+dx$ 之间的概率。随机变量落入间隔 $a<x<b$ 内的概率为

$$P(a<X\leq b)=\int_{a}^{b} P(x)dx$$

对于连续随机变量，变量落入一个趋于零的间隔内的概率随间隔而下降。这一点在上面表达式中用 $a+\varepsilon$ 代替 b，并使 ε 趋于零可以看出。

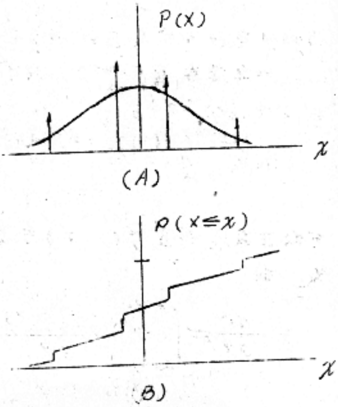


图1-3　混合随机变量的概率函数
(A). 密度函数
(B). 累积分布函数

对于两个随机变量 X 与 Y. 联合概率密度函数 $P(x,y)$ 定义为

$$P(x,y)=\frac{\partial^2}{\partial x\partial y}P(X\leq x,\ Y\leq y). \quad (1-7)$$

因此得到[1]

$$P(X\leq a\quad Y\leq b)=\int_{-\infty}^{b}dy\int_{-\infty}^{a}dx\,p(x,y).$$

$$P(X\leq a)=\int_{-\infty}^{\infty}dy\int_{-\infty}^{a}dx\;p(x,y).$$

$$P(Y\leq b)=\int_{-\infty}^{\infty}dx\int_{-\infty}^{b}dy\,p(x,y).$$

类似地，可以证明，

$$p(x)=\int_{-\infty}^{\infty}p(x,y)\,dy$$

与

$$p(y)=\int_{-\infty}^{\infty}p(x,y)\,dx$$

用这种方法产生的概率密度函数有时就称为单边密度函数。

在已给随机变量 X 时变量 Y 的条件概率密度函数定义为

$$p(y|x)=p(x,y)/p(x)\qquad(p(x)\neq 0)$$

……………………（1—8）

因此，
$$P(Y\leq b|x)=\int_{-\infty}^{b}p(y|x)\,dy.$$

它可解释为在已给 $\{X=x\}$ 时，事件 $\{Y\leq y\}$ 的概率。条件概率密度函数的定义可以推广到在另一个随机变量的集合 $\{X_j,\ j=1,\cdots\cdots,m\}$ 已给的条件下，某一随机变量的集合 $\{Y_i,\ i=1,\cdots,n\}$

---

[1] 下面共同规定，用符号 $p(\cdot)$ 来表示 X 的和 Y 的概率密度函数以及 X 与 Y 的联合概率密度函数，而不管它们是否有不同的函数形式。变量含有区分这些函数的意义，一旦这个规定引起前后含糊不清时，就用脚标来区分不同的函数形式，

中去。因此

$$p(y_1,\cdots,y_n\mid x_1,\cdots,x_m)=\frac{p(y_1,\cdots,y_n,x_1,\cdots,x_m)}{p(x_1,\cdots,x_m)}$$

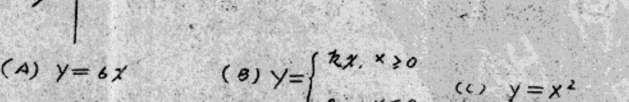
统计独立的条件也可以用概率密度函数来叙述：当且仅当

$$p(x,y,\cdots,z)=p(x)p(y)\cdots p(z) \quad (1-9)$$

对 x，y，…… z 的所有值都成立时，则随机变量 X，Y，……，Z 是统计独立的。

§1—5 随机变量的函数

在检测理论以及其它包括概率论与统计学的学科中常常出现一个或多个随机变量的函数。一个随机变量的函数 $y=f(x)$ 可以作如下确定：规定从进行的实验中得出实数 x，然后由定义 $y=f(x)$ 来作数学运算。典型的例子显示于图1—4，它可以推



(A) $y=bx$ $-\infty<x<\infty$

(B) $y=\begin{cases} bx, & x\ge 0 \\ 0, & x<0 \end{cases}$

(C) $y=x^2$ $-\infty<x<\infty$

图1—4 随机变量函数的例子

(A) 线性运算，(B) 半波检流，(3) 平方律

广到多于一个随机变量的函数，例如，$y=f(x, z)$，通过观测成对的实数x与z，并作出$y=f(x, z)$的数学运算来详细确定这种函数运算。求和$y=x+z$就是一个例子。

为了阐明如何确定一随机变量函数的统计学的直接方法，考虑图1—4(A)所指明的情形，这是一种线性函数$y=bx$，只含有一个标量的变化，假设X的概率密度函数是已知的，而Y的密度函数是待定的。$\{Y\le y\}$的概率就等于$\{X\le y/b\}$的概率。这就是说小，$P_Y\{Y\le y\}=P_X\{X\le y/b\}$由概率密度函数的定义直接得到

$$p_Y(y)=\frac{d}{dy}P_X\{X\le y/b\}.$$

因为分布函数是非减的，导数就不能是负数，并用莱布尼茨法则可以证明

$$p_Y(y)=\frac{1}{|b|}p_X(x=y/b),$$

其中$|\cdot|$表示绝对值，Y的区域是X的区域乘上b。

[例1-5-1]．如果在上述情形中，X的密度函数是指数的（换句话说，x为指数分布的），则对$x>0$，$p_X(x)=e^{-x}$，立即得到

$$p_Y(y)=\frac{1}{|b|}e^{-y/b}\qquad\begin{cases}\text{如果 } b>0,\ y\ge 0;\\ \text{如果 } b<0,\ y\le 0.\end{cases}$$

X与Y的密度函数在$b=2$时的情形如图1—5所示。

[例1-5-2]　设$P_X(X)$具有与上例相同的密度函数．要确定Y的密度函数，其中$Y=X+a$

---

*：在这里为了避免可能的混淆。X与Y的分布函数分别用$P_X(\cdot)$与$P_Y(\cdot)$来表示。

用直接方法，

$$P_Y(Y\le y)=P_X(X\le y-a)$$

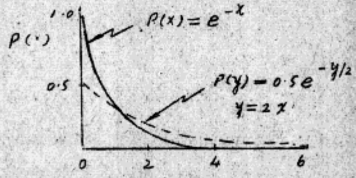


图 1-5　例 1-5-1 的密度函数

现在，

$$p_Y(y)=\frac{d}{dy}P_Y(Y\le y)$$

$$=\frac{d}{dy}P_X(X\le y-a)$$

所以，

$$p_Y(y)=p_X(x=y-a)=e^{-(y-a)},$$

$$y>a.$$

它由图 1-6 所示。

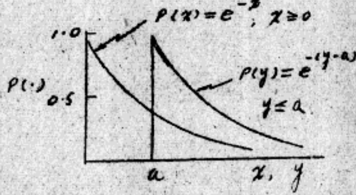


图 1-6　例 1-5-2 的密度函数

下面讨论更为复杂但含有与如上所讨论的简单情形相同原理的情况。一个或多个随机变量变换或投影到另一个随机变量的集合中去。多维随机变量落入取样空间的给定区域的概率与变换后的多维随机变量落入新取样空间的变换后区域的概率是相同的。

我们假设有一个随机变量的集合，例如说 $X_1, X_2, \cdots, X_N$. 对此，用 $p_X(x_1, x_2, \cdots, x_N)$ 表示联合概率密度函数，它是已知的。我们要求出新随机变量集合，例如说 $Y_1, Y_2, \cdots, Y_N$，的联合概率密度函数 $p_Y(y_1, y_2, \cdots, y_N)$，这两个随机变量的集合有如下的函数关系：

$$Y_1=g_1(X_1, X_2, \cdots, X_N),$$

$$\vdots$$

$$y_N = g_N(x_1, x_2, \cdots x_N) \quad \cdots\cdots\cdots \quad (1-10)$$

例如，

$$y_1 = x_1 + x_2, \qquad y_2 = x_1 - x_2 .$$

现在，新随机变另 y_i 的数目（N）假定是等于老随机变另 x_i 的数目，假设新随机变另是老随机变另的单值连续函数（处处有连续的偏导数），并更进一步地，老变另也可以表示为新变另的单值连续函数：

$$x_1 = f_1(y_1, y_2, \cdots y_N)$$

$$\vdots \qquad\qquad \cdots\cdots\cdots \quad (1-11)$$

$$x_N = f_N(y_1, y_2, \cdots, y_N)$$

联系到上面的例子，

$$x_1 = \frac{1}{2}(y_1 + y_2), \qquad x_2 = \frac{1}{2}(y_1 - y_2).$$

因此，对于 x_i 的取样空间中每一点，存在着相应 y_i 的取样空间中的一个且唯一的点，所以老变另到新变另具有一一对应的射形。

假设在 x 域内的一个区域 A 中含有取样点的一个特定的集合，

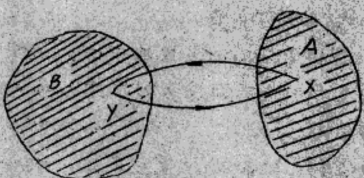


图 1—7　空间 A 到空间 B 的一一对应射形的举例，x 与 y 分别表示多维随机变量 $x_1, x_2, \cdots x_N$ 与 $y_1, y_2, \cdots y_N$。