第十章 离散随机信号的统计分析基础

引 言

信号可以分为确定性信号和随机信号。

确定性信号 —— 信号的幅度随时间的变化有一定的规律性，可以用一个明确的数学关系进行描述，是可以再现的。

随机信号 —— 信号随时间的变化没有明确的变化规律，既不是有限能量又不是周期的，不能用明确的时间函数进行描述，也不能准确地重现。

随机信号的分类：

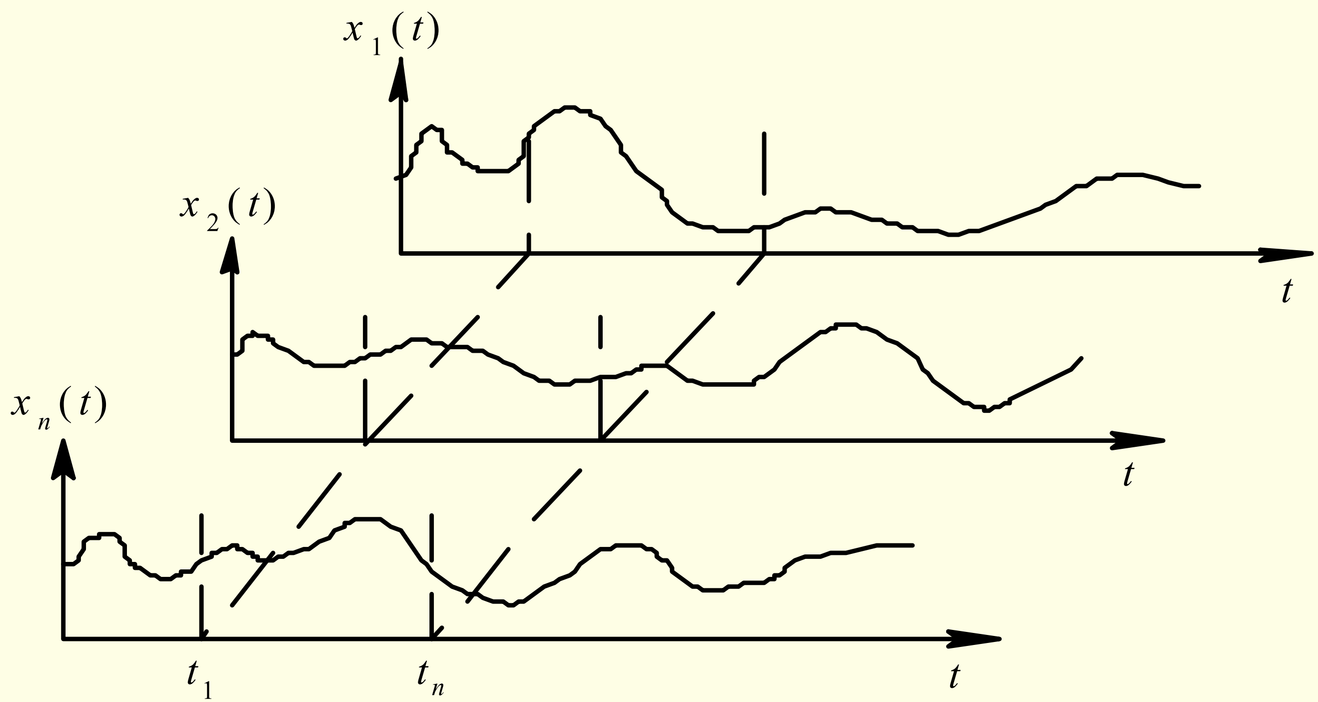
		时间变量	
		连续值	离散值
幅度	连续值	连续随机信号	时域离散随机信号(简称随机序列)
	离散值	幅度离散随机信号（例如随机脉冲信号， 其取值只有两个电平)	离散随机序列(也称为随机数字信号)

* 利用计算机只能处理随机数字信号。

10.1 随机过程的定义

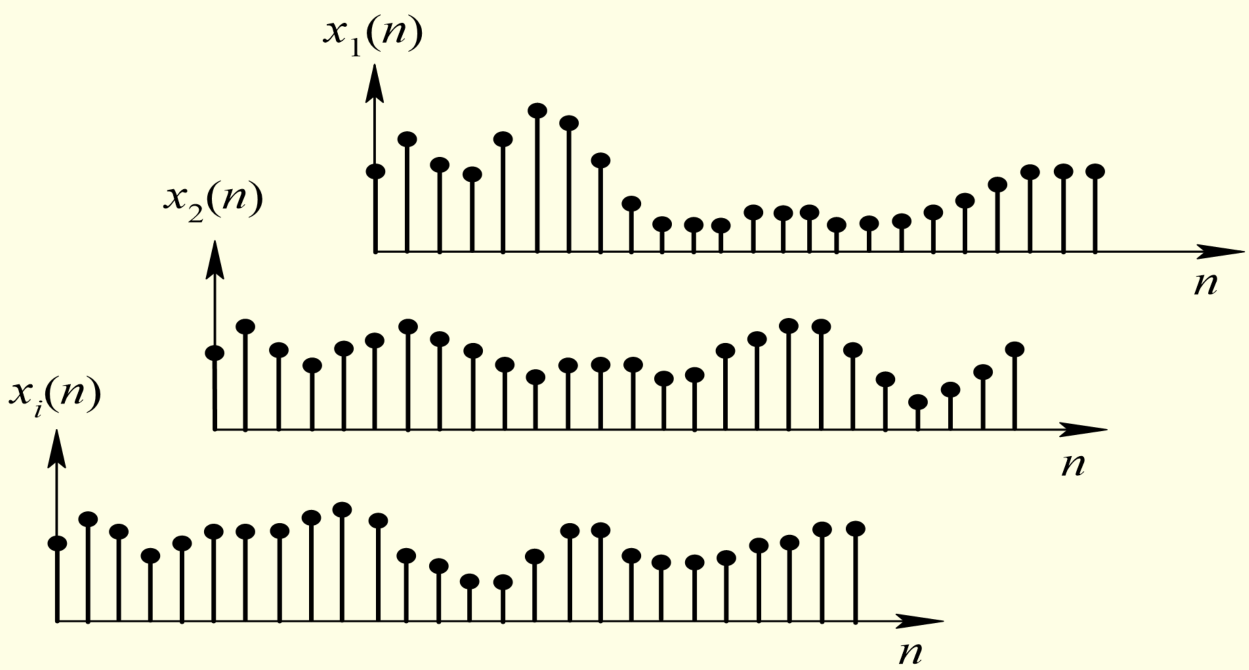
随机过程$X(t)$ 是由它所有可能的样本函数集合而成的，样本函数用$x_i(t)$，i=1，2，3,…表示。

如下图表示n部接收机的噪声电压，$x_i(t)$ 称为第i条样本曲线。



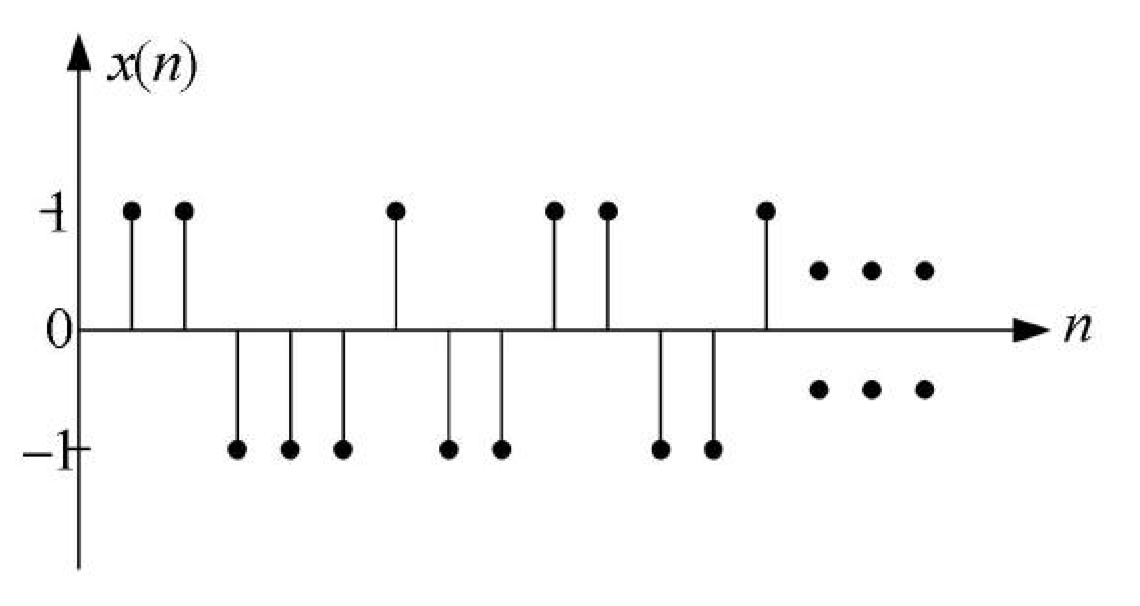
连续随机过程$X(t)$

- 对$X(t)$进行等间隔采样，得到$X(t_1)$，$X(t_2)$，$X(t_3)$，…，所构成的集合称为时域离散随机信号（随机序列），用$X(n)$表示。相应的$x_i(n)$，i=1，2，3，…，称为样本序列，它们是n的确定性函数。
- 而X(t1)，X(t2)，X(t3)则都是随机变量。
- 因此随机序列兼有随机变量和函数的特点。



离散随机过程X(n)

一个例子：



抛掷硬币得到的 *离散随机样本序列* $x_1(n)$

10.2 离散随机过程的时域统计描述

10.2.1 概率分布函数和密度函数

（1）概率分布函数

对于随机变量X$_n$, 其概率分布函数用下式描述：

$$F_{X_n}(x_1,n)=\mathrm{P}(x_n\le x_1)$$

式中P（）表示概率，x1为任一实数。

（2） 概率质量函数

对于值域离散的随机变量Xn，定义概率质量函数如下：

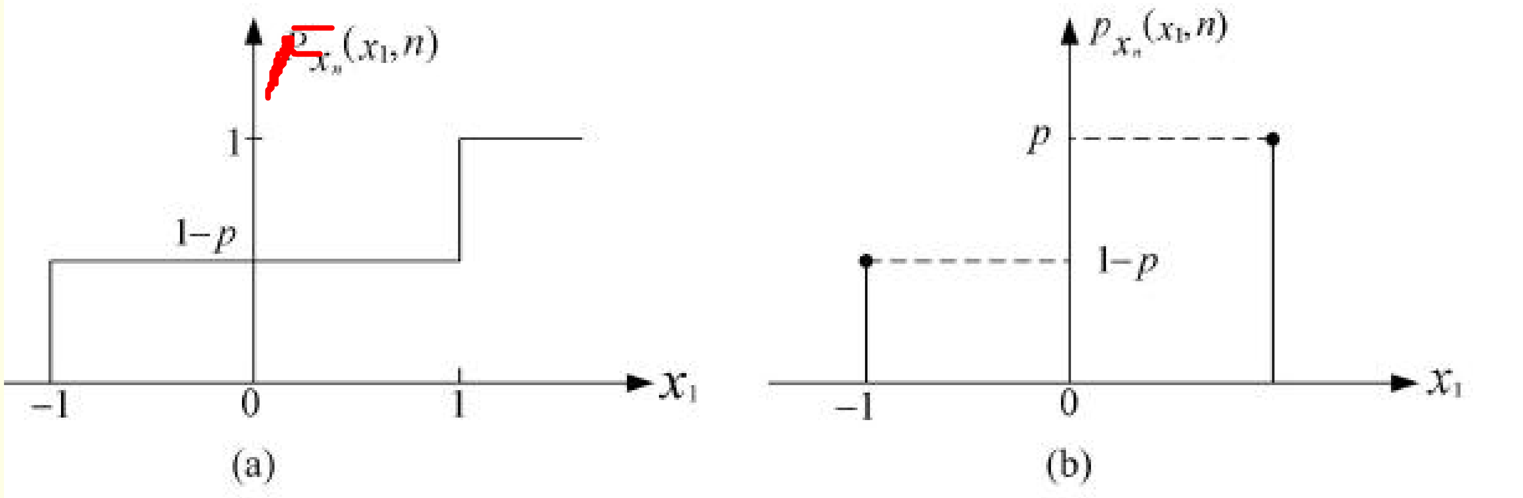
$$p_{X_n}(x_1,n)=\mathrm{P}(x_n=x_1)$$

如果x$_n$取连续值，则称之为概率密度函数。

概率分布函数与概率质量函数的关系是

$$F_{X_n}(x_1,n)=\sum_{x_n\le x_1}p_{X_n}(x_1,n)$$

- 抛硬币的随机过程就是xn*取值是离散的例子*。设*xn = +1*（正面）的概率为*p*，则*xn = -1*（反面）的概率为*(1-p)*。下图给出了离散型随机变量xn的概率分布函数和概率质量函数。



硬币抛掷问题(a)抛掷硬币的随机变量x_n的概率分布函数；(b) 相应的概率质量函数

- 一些离散随机变量的常用分布。

(1) 伯努利分布

该分布的概率质量函数是

$p(-1) = 1 - q; p(1) = q; q \in [0, 1]$；抛掷单个硬币就属于这种分布。

(2) 二项分布

该分布的概率质量函数是 $p(k) = C_n^k p^k (1-p)^{n-k}, k \in \{0,1,\cdots,n\}$

多次抛掷硬币时，k次正面或反面的次数可以这个分布描述。

(3) 均匀分布

该分布的概率质量函数是 $p(k) = 1/n, \quad k \in \{0,1,\cdots,n\}$

许多传感信号和社会计算的模型均可用这种分布。

(4) 泊松分布

该分布的概率质量函数是 $p(k) = \lambda^k e^{-1}/k!, \quad k \in \{0,1,\cdots,n\}$

该分布可以描述数据传输、电子辐射、电话呼叫中的数据到达时间。

（3）二维概率分布函数:

$$F_{X_n,X_m}(x_1,n,x_2,m)=P(x_n\le x_1,x_m\le x_2)$$

二维概率质量函数:

$$p_{X_n,X_m}(x_1,n,x_2,m)=P(x_n=x_1,x_m=x_2)$$

以此类推, *N*维概率分布函数为

$$F_{X_1,X_2\cdots,X_M}(x_1,n_1,x_2,n_2,\cdots,x_M,n_M)=P(x_{n_1}\le x_1,x_{n_2}\le x_2,\cdots,x_{n_M}\le x_M)$$

10.2.2 平稳随机过程

定义：

平稳随机序列指它的N维概率分布函数或N维概率质量（密度）函数与时间n的起始位置无关，其统计特性不随时间的平移而发生变化。如果将随机序列在时间上平移k，其统计特性满足下式：

$$
\begin{aligned}
&F_{X_{1+k},X_{2+k},\cdots,X_{N+k}}(x_1,n_1+k,x_2,n_2+k,\cdots,x_M,n_M+k) \\
&=F_{X_1,X_2,\cdots,X_N}(x_1,n_1,x_2,n_2,\cdots,x_M,n_M)
\end{aligned}
$$

- 上述这类序列称为狭义平稳随机序列，其条件在实际情况下很难满足。
- 对于随机序列的均值和方差不随时间改变，则其相关函数仅是时间差的函数，将这一类随机序列称为广义平稳随机序列。

平稳随机序列的一维概率质量（密度）函数与时间无关，因此均值、方差和均方值均与时间无关，它们可分别用下式表示：

$$m_x = E[X_n] = E[X_{n+m}] = E[X]$$

$$E[X_n^2] = E[X_{n+m}^2] = E[X^2]$$

$$\sigma_x^2 = E[(X_n - m_x)^2] = E[(X_{n+m} - m_x)^2] = E[(X - m_x)^2]$$

二维概率密度函数仅决定于时间差，与起始时间无关，因此自相关函数与自协方差函数是时间差的函数：

$$\varphi_{xx}(m) = E[X_n X_{n+m}]$$

$$\gamma_{xx}(m) = E[(X_n - m_x)(X_{n+m} - m_x)]$$

对于两个各自平稳且联合平稳的随机过程X、Y，其互相关函数为

$$\varphi_{xy}(m)=\varphi_{xy}(n,n+m)=E[X_n Y_{n+m}]$$

如果对于所有的m，满足公式：$\phi_{xy}(m)=0$，则称两个随机序列互为正交。

如果对于所有的m，满足公式：$\varphi_{xy}(m)=m_x m_y$，$r_{xy}(m)=0$，则称两个随机序列互不相关。

实平稳随机序列的相关函数、协方差函数具有以下性质：

（1）自相关函数和自协方差函数是m的偶函数，用下式表示：

$$\varphi_{xx}(m)=\varphi_{xx}(-m),\gamma_{xx}(m)=\gamma_{xx}(-m)$$

$$\varphi_{xy}(m)=\varphi_{yx}(-m),\gamma_{xy}(m)=\gamma_{yx}(-m)$$

（2） $$\varphi_{xx}(0)=E[X_n^2]$$

（3） $$\varphi_{xx}(0) \geq |\varphi_{xx}(m)|$$

（4） $$\lim_{m\to\infty} \varphi_{xx}(m) = m_x^2$$

$$\lim_{m\to\infty} \varphi_{xy}(m) = m_x m_y$$

（5） $$\gamma_{xx}(m) = \varphi_{xx}(m) - m_x^2, \qquad \gamma_{xx}(0) = \sigma_x^2$$

可以看出$\varphi_{xx}(m)$是描述随机过程X最主要的统计量，它不仅说明了随机变量间的相关性，而且蕴含了均值、均方值，方差等主要特征量。

10.2.3　概率分布特性的特征量

概率分布函数能对随机序列进行完整的描述，但实际中往往无法得到它。而随机序列的一些数字特征则比较容易进行测量和计算，在实际中，知道这些数字特征就足够用了。

（1）数学期望(统计平均值)

n时刻对应的离散随机变量Xn的数学期望定义为

$$m_{x_n} = E[X_n] = \sum_{x} x p_{X_n}(x,n)$$

（2） 均方值与方差

n时刻的离散随机变量Xn的均方值定义为：

$$E[X_n^2]=\sum_x x^2 p_{X_n}(x,n)$$

方差定义为：

$$\sigma_{x_n}^2=E[\left(X_n-m_{x_n}\right)^2]$$

可以推出以下等价的形式:

$$\sigma_{x_n}^2=E[X_n^2]-m_{x_n}^2$$

10.2.4 相关函数和协方差函数

随机序列不同时刻的状态之间存在着一定关联性，即互相有影响，包括随机序列本身或者不同随机序列之间。

n、m时刻分别对应的随机变量Xn、Xm，其自相关函数定义为：

$$\varphi_{xx}(n,m)=E[X_nX_m]=\sum_{x_n}\sum_{x_m}x_nx_mp_{X_n,X_m}(x_n,n,x_m,m)$$

其自协方差函数定义为

$$\gamma_{xx}(n,m)=\mathrm{cov}[X_n,X_m]=E[(X_n-m_{x_n})(X_m-m_{x_m})]$$

可以推出上式的另一种等价形式：

$$\gamma_{xx}(n,m)=\varphi_{xx}(n,m)-m_{x_n}m_{x_m}$$

对于零均值随机序列，$m_{x_n}=m_{x_m}=0$，则 $\gamma_{xx}(n,m)=\varphi_{xx}(n,m)$

这种情况下，自相关函数和自协方差函数相等。

n、m时刻分别对应的随机变量Xn、Ym，其互相关函数定义为：

$$\varphi_{xy}(n,m) = E[X_n Y_m] = \sum_{x_n}\sum_{y_m} x_n y_m p_{X_n,Y_m}(x_n, n, y_m, m)$$

式中$p_{X_n,Y_m}(x_n, n, y_m, m)$表示$X_n$和$Y_m$的联合概率质量函数。对于两个不同的随机序列之间的关联性，可以用互相关函数和下面的互协方差函数来描述。

互协方差函数定义为：

$$\gamma_{xy}(n,m) = \mathrm{cov}(X_n, Y_m) = E[(X_n - m_{x_n})(Y_m - m_{y_m})]$$

$$= \varphi_{xy}(n,m) - m_{x_n} m_{y_m}$$

同样，当$m_{X_n}=m_{Y_m}=0$时，

$$\gamma_{xy}(n,m) = \varphi_{xy}(n,m)$$

10.2.5 各态历经性与时间平均

上述数字特征都要求对大量的样本进行平均，即集合平均。但实际中这种做法是不现实的，在很多情况下，用一条样本曲线来描述随机序列更加容易。

设$x(n)$是平稳随机序列$X(n)$的一条样本曲线，其时间平均为:

$$<x(n)>=\lim_{N\to\infty}\frac{1}{2N+1}\sum_{n=-N}^{N}x(n)$$

式中〈·〉表示时间平均算子。

类似地，其时间自相关函数为:

$$<x(n)x(n+m)>=\lim_{N\to\infty}\frac{1}{2N+1}\sum_{n=-N}^{N}x(n)x(n+m)$$

定义：

如果平稳随机序列的集合平均值与集合自相关函数值依概率趋于其样本函数的时间平均值与时间自相关函数，即满足下面两式：

$$\langle x(n)\rangle = m_x = E[X]$$

$$\langle x(n)x(n+m)\rangle = \varphi_{xx}(m) = E[X(n)X(n+m)]$$

则称该平稳随机序列具有各态历经性。

- 这样用研究平稳随机序列的一条样本曲线代替研究其集合，即用时间平均代替集合平均，给研究平稳随机序列带来极大的方便。

10.3 离散随机过程的频域统计描述

平稳随机序列是能量无限信号，故无法直接利用傅里叶变换进行分析。但自相关函数却随着时间差m的增大，趋近于随机序列的均值。如果随机序列的均值为0，即m_x=0，$\phi_{xx}(m)$是收敛序列，可以进行Z变换：

$$P_{xx}(z)=\sum_{m=-\infty}^{\infty}\varphi_{xx}(m)z^{-m}$$

$$\varphi_{xx}(m)=\frac{1}{2\pi j}\oint_{c}P_{xx}(z)z^{m-1}\mathrm{d}z$$

$P_{xx}(z)$的收敛域如下，故包含单位圆：

$$R_a<|z|<R_a^{-1}$$

0≤R_a≤1

由于$P_{xx}(z)$的收敛域包含单位圆，因此$\varphi_{xx}(m)$的DTFT存在。令$z=\exp(j\omega)$, 有

$$P_{xx}(\mathrm{e}^{\mathrm{j}\omega}) = \sum_{m=-\infty}^{\infty} \varphi_{xx}(m)\mathrm{e}^{-j\omega m}$$

$$\varphi_{xx}(m) = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_{xx}(\mathrm{e}^{j\omega})\mathrm{e}^{j\omega m}\mathrm{d}\omega$$

功率谱的性质：

（1） 功率谱是ω的偶函数：

$$P_{xx}(\omega) = P_{xx}(-\omega)$$

该结果可直接由自相关函数是时间差的偶函数来证明。

（2） 功率谱是实的非负函数， 即

$$P_{xx}(\omega)\geq 0$$

类似地，互相关函数的Z变换为：

$$P_{xy}(z)=\sum_{m=-\infty}^{\infty}\varphi_{xy}(m)z^{-m}$$

其相对应的DTFT称为互功率谱：

$$P_{xy}(\omega)=\sum_{m=-\infty}^{\infty}\varphi_{xy}(m)e^{-j\omega m}$$

$$\varphi_{xy}(m)=\frac{1}{2\pi}\int_{-\pi}^{\pi}P_{xy}(\omega)e^{j\omega m}d\omega$$

互功率谱的性质：

$$P_{xy}(\omega)=P_{yx}(-\omega)$$

白噪声序列

对于随机序列X(*n*)，如果其随机变量在不同时刻的取值无关联性，即两两不相关，表示为自协方差函数

$$\gamma_{xx}(n,m)=\mathrm{cov}(X_n,X_m)=\sigma_{x_n}^2\delta_{mn}$$

式中

$$\delta_{mn}=\begin{cases}1 & m=n\\ 0 & m\neq n\end{cases}$$

则称该序列为白噪声序列。如果白噪声序列还是平稳的， 则

$$\gamma_{xx}(n,m)=\mathrm{cov}(X_n,X_m)=\sigma^2\delta_{mn}$$

式中， σ^2是常数。设均值$m_{x_n}=0$，则其功率谱$P_{xx}(e^{j\omega})=\sigma^2$，即在整个频带上功率谱是一个常数。

10.4 离散线性系统对随机信号的响应

1. 相关卷积定理

假设线性时不变系统的单位脉冲响应为实序列$h(n)$，则当输入是平稳随机序列$x(n)$，其输出为

$$y(n)=\sum_{k=-\infty}^{\infty}h(k)x(n-k)$$

卷积

那么，可以推导出输出信号的自相关函数为：

$$\varphi_{yy}(m)=\sum_{l=-\infty}^{\infty}\varphi_{xx}(m-l)v(l)$$

$$=\varphi_{xx}(m)*v(m) \qquad (1)$$

式中 $v(l)=\sum_{k=-\infty}^{\infty}h(k)h(l+k)$ 为$h(n)$的自相关函数。 (2)

上述公式用语言叙述如下：$x(n)$ 与 $h(n)$ 卷积的自相关等于 $x(n)$ 的自相关和 $h(n)$ 的自相关的卷积。即卷积的相关等于相关的卷积。 用一般公式表示如下：

如果 $$\left.\begin{aligned} e(n) &= a(n) * b(n) \\ f(n) &= c(n) * d(n) \end{aligned}\right\}$$

那么 $$\varphi_{ef}(m) = \varphi_{ac}(m) * \varphi_{bd}(m)$$

这就是相关卷积定理，描述了随机信号通过线性系统的时域表达。

2. 输出响应的功率谱密度函数

将上述(1)、(2)式分别写成*Z*变换形式，表示如下：

$$V(z)=H(z)H\left(\frac{1}{z}\right)$$

$$P_{yy}(z)=P_{xx}(z)H(z)H\left(\frac{1}{z}\right)$$

将z=e$^{j\omega}$代入上式，得到输出功率谱：

$$P_{yy}(e^{j\omega})=P_{xx}(e^{j\omega})|H(e^{j\omega})|^2$$

3. 系统的输入、 输出互相关函数

线性系统输入与输出间的互相关函数满足下列关系：

$$\varphi_{xy}(m) = E[x(n)y(n+m)]$$
$$= \sum_{k=-\infty}^{\infty} h(k)\varphi_{xx}(m-k)$$
$$= h(m) * \varphi_{xx}(m)$$

即输入输出间的互相关函数等于h(n)与输入自相关函数的卷积。称为输入、输出互相关定理。

证明：根据相关卷积定理，有

$$x(n)=x(n)*\delta(n)$$

$$y(n)=x(n)*h(n)$$

$$\varphi_{xy}(m)=\varphi_{xx}(m)*\varphi_{\delta h}(m)$$

其中

$$\varphi_{\delta h}(m)=\sum_{l=-\infty}^{\infty}\delta(l)h(m+l)=h(m)$$

代入上式，得到

$$\varphi_{xy}(m)=\varphi_{xx}(m)*h(m)$$

再由线性卷积的交换律，得到

$$\varphi_{xy}(m)=h(m)*\varphi_{xx}(m) \qquad \square$$

设$x(n)$是零均值平稳随机序列，对上述结果进行Z变换，得到

$$P_{xy}(z) = H(z)P_{xx}(z)$$

其功率谱表示为

$$P_{xy}(\mathrm{e}^{j\omega}) = H(\mathrm{e}^{\mathrm{j}\omega})P_{xx}(\mathrm{e}^{\mathrm{j}\omega})$$

称为互功率谱密度函数。

4. 小结

信号通过线性系统的主要分析工具对比：

	确定性信号	平稳随机信号
时域	冲激响应 $h(n)$	自相关函数 $\varphi_{xx}(m)$ $\varphi_{yy}(m)$ 互相关函数 $\varphi_{xy}(m)$
频域	系统函数 $H(z)$ 频率响应 $H(\mathrm{e}^{\mathrm{j}\omega})$	功率谱密度 $P_{xx}(\mathrm{e}^{\mathrm{j}\omega})$ $P_{yy}(\mathrm{e}^{\mathrm{j}\omega})$ 互谱密度 $P_{xy}(\mathrm{e}^{\mathrm{j}\omega})$

DTFT DTFT

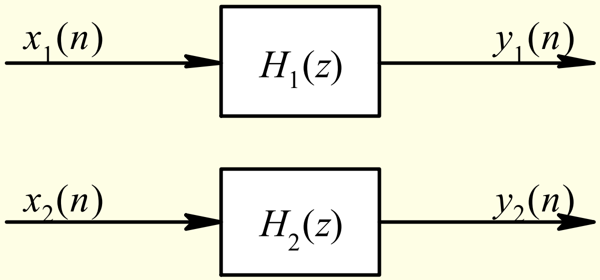
因此相关运算、功率谱估计成为随机信号处理中最基本的研究内容。

例 1. 已知下图两个系统的输入互谱密度函数 $P_{x_1x_2}$ 及系统函数，试求其输出互谱密度函数 $P_{y_1y_2}$ 。

解:

$y_1(n)=x_1(n)*h_1(n)$

$y_2(n)=x_2(n)*h_2(n)$



按照相关卷积定理, 有

$$\varphi_{y_1y_2}(m)=\varphi_{x_1x_2}(m)*\varphi_{h_1h_2}(m)$$

$$\varphi_{h_1h_2}(m)=\sum_{n=-\infty}^{\infty}h_1(n)h_2(n+m)=\sum_{n=-\infty}^{\infty}h_1(-n)h_2(m-n)=h_1(-m)*h_2(m)$$

$$\left(x(n)*h(n)=\sum_{m=-\infty}^{\infty}x(m)h(n-m)\right)$$

$$\therefore P_{y_1y_2}(z)=P_{x_1x_2}(z)H_1\left(z^{-1}\right)H_2(z)$$

例2 已知实平稳白噪声$x(n)$的功率谱是$\sigma^2{}_x$，使通过一个q阶的FIR网络，求输出自相关函数$\varphi_{yy}(m)$、功率谱$P_{yy}(e^{j\omega})$、互相关函数$\varphi_{xy}(m)$和互谱$P_{xy}(e^{j\omega})$。

解 设系统的传输函数用下式表示：

$$H(z)=\sum_{n=0}^{q}b_n z^{-n}$$

式中系数b_n是实数，网络输出的功率谱为

$$P_{yy}(\mathrm{e}^{\mathrm{j}\omega}) = P_{xx}(\mathrm{e}^{\mathrm{j}\omega})\,|H(\mathrm{e}^{\mathrm{j}\omega})|^2$$

$$= \sigma_x^2 \left| \sum_{n=0}^{q} b_n \mathrm{e}^{-\mathrm{j}\omega n} \right|^2$$

$$= \sigma_x^2 \left[\left(\sum_{n=0}^{q} b_n \cos(\omega n) \right)^2 + \left(\sum_{n=0}^{q} b_n \sin(\omega n) \right)^2 \right]$$

$$= \sigma_x^2 [\sum_{n=0}^{q} b_n^2 + \sum_{\substack{n=0 \\ n \neq j}}^{q} \sum_{j=0}^{q} b_n b_j \cos(n-j)\omega]$$

因为$\delta(n)$的傅里叶变换是1，$\delta(n+m)+\delta(n-m)$的傅里叶变换是$2\cos(\omega m)$，那么对上式进行反变换，得到

$$\varphi_{yy}(m)=\sigma_x^2\left[\sum_{n=0}^{q}b_n^2\right]\delta(m)+\sigma_x^2\sum_{\substack{n=0\\n\neq j}}^{q}\sum_{j=0}^{q}\frac{b_n b_j}{2}\left[\delta(m-j+n)+\delta(m+j-n)\right]$$

上式表明输出信号的自相关函数$\varphi_{yy}(m)$有限长，存在于$\pm q$之间。 类似地，可求出互谱和互相关函数为

$$P_{xy}(\mathrm{e}^{\mathrm{j}\omega})=\sigma_x^2\left[b_0+\sum_{n=1}^{q}b_n\cos(\omega n)-j\sum_{n=1}^{q}b_n\sin(\omega n)\right]$$

$$\varphi_{xy}(m)=E[x(n)y(n+m)]$$

$$=E[x(n)\left(\sum_{k=0}^{q}bk\right)$$

例3 设实平稳白噪声$x(n)$的方差是σ^2_x, 均值$m_x=0$，让$x(n)$通过一个网络，网络的差分方程为

$$y(n)=x(n)+ay(n-1)$$

式中a是实数。求网络输出的功率谱和自相关函数。

解 先用归纳法求网络输出的自相关函数

$$\varphi_{yy}(m)=E[y(n)y(n+m)]$$

令$m=0$, 则

$$\varphi_{yy}(0)=E[y^2(n)]=E[(x(n)+ay(n-1))^2]$$

$$\varphi_{yy}(0)=E[x^2(n)]+a^2E[y^2(n-1)]+2aE[x(n)y(n-1)]$$

上式中$y(n-1)$发生在$x(n)$之前，它只和$x(n-1)$, $x(n-2)$, …有关，而且$x(n)$是白噪声，$x(n)$和$x(n-1)x(n-2)$, …无关，因此上式中的第三项等于0, 那么

$$\varphi_{yy}(0)=\sigma_x^2+a^2\varphi_{yy}(0)$$

$$\varphi_{yy}(0)=\frac{\sigma_x^2}{1-a^2}$$

令$m=1$，则

$$\varphi_{yy}(1)=E[y(n)y(n+1)]$$

$$\varphi_{yy}(1)=E[y(n)(ay(n)+x(n+1))]=a\varphi_{yy}(0)$$

令m=2, 则

$$\begin{aligned}\varphi_{yy}(2)&=E[y(n)y(n+2)]\\&=E[y(n)(ay(n+1)+x(n+2))]\\&=a\varphi_{yy}(1)=a^2\varphi_{yy}(0)\end{aligned}$$

总结规律，因此有

$$\varphi_{yy}(m)=a^m\varphi_{yy}(0)=\frac{a^m}{1-a^2}\sigma_x^2$$

下面再求网络输出的功率谱, 由给定的网络差分方程, 得到网络系统函数

$$H(z)=\frac{1}{1-az^{-1}}$$

网络输出功率谱为

$$P_{yy}(\mathrm{e}^{\mathrm{j}\omega}) = P_{xx}(\mathrm{e}^{\mathrm{j}\omega})\,|\,H(\mathrm{e}^{\mathrm{j}\omega})\,|^2 = \sigma_x^2 \left| \frac{1}{1-a\mathrm{e}^{-\mathrm{j}\omega}} \right|^2 = \frac{\sigma_x^2}{1+a^2-2a\cos\omega}$$

式中，a是网络的极点，为了稳定，要求$|a|<1$。a愈接近于单位圆，功率谱峰愈尖锐，带宽愈窄，但相关函数衰减愈慢；反过来，a愈小，功率谱下降愈慢，自相关函数衰减愈加快。

本章小结

- 随机过程的定义
- 离散随机过程的时域统计描述
- 各态历经与时间平均
- 离散随机过程的频域统计描述
- 离散线性系统对随机信号的响应