随机过程

June 20, 2018

设 T 是一无限实数集，称为参数集；依赖于参数 $t \in T$ 的一族 (无限多个) 随机变量，称为随机过程，记为 $\{X(t), t \in T\}$；$X(t)$ 是一随机变量；

常把 t 看作时间，称 $X(t)$ 为时刻 t 时过程的状态；

随机过程的状态空间：对于一切 $t \in T$，$X(t)$ 所有可能取的一切值的全体；

随机过程的一个样本函数或样本曲线：对随机过程 $\{X(t), t \in T\}$，进行一次试验 (即在 T 上进行一次全程观测)，结果是 t 的函数，记为 $X(t), t \in T$；

所有不同的试验结果构成一族样本函数；

伯努利过程、伯努利随机序列

随机过程在任一时刻的状态是随机变量；

随机过程的分布函数族

给定随机过程 $\{X(t), t \in T\}$；对于每一个固定的 $t \in T$，随机变量 $X(t)$ 的分布函数一般与 t 有关，记为

$$F_X(x,t) = P\{X(t) \le x\}, x \in R \tag{1}$$

称为随机过程 $\{X(t), t \in T\}$ 的一维分布函数；$\{F_X(x,t), t \in T\}$ 称为一维分布函数族；

一维分布函数族刻画了随机过程在各个个别时刻的统计特征；

为了描述随机过程在不同时刻状态之间的统计联系，对任意 $n(n=2,3,\cdots)$ 个不同的时刻 $t_1,t_2,\cdots,t_n\in T$，引入 n 维随机变量 $(X(t_1),X(t_2),\cdots,X(t_n)\ )$，它的分布函数

$$F_X(x_1,x_2,\cdots,x_n;t_1,t_2,\cdots,t_n)=P\{X(t_1)\le x_1,X(t_2)\le x_2,\cdots,X(t_n)\le x_n\}, \tag{2}$$

$$x_i\in R,i=1,2,\cdots,n \tag{3}$$

对固定的 n，$\{F_X(x_1,x_2,\cdots,x_n;t_1,t_2,\cdots,t_n),t_i\in T\}$ 为随机过程 $\{X(t),t\in T\}$ 的n 维分布函数族；

科尔莫戈罗夫定理：

有限维分布函数族，即 $\{F_x(x_1,x_2,\cdots,x_n;t_1,t_2,\cdots,t_n),n=1,2,\cdots,t_i\in T\}$，完全确定了随机过程的统计特性。

1 数字特征

均值函数

给定随机过程 $\{X(t),t\in T\}$，固定 $t\in T$，$X(t)$ 是一随机变量，它的均值与 t 有关，

$$\mu_X(t)=E[X(t)] \tag{4}$$

$\mu_X(t)$ 是随机过程的所有样本函数在 t 时刻函数值的平均值；

集平均或统计平均；

均方值函数

二阶原点矩

$$\Psi_X^2(t)=E[X^2(t)] \tag{5}$$

方差函数

二阶中心矩

$$\sigma_X^2(t) = D_X(t) = \mathrm{Var}[X(t)] = E\{[X(t) - \mu_X(t)]^2\} \tag{6}$$

标准差函数$\sigma_X(t)$：方差函数的算术平方根；

表示随机过程 $X(t)$ 在时刻 t 对于均值 $\mu_X(t)$ 的平均偏离程度；

自相关函数(相关函数)

设任意 $t_1, t_2 \in T$，随机变量 $X(t_1), X(t_2)$ 的二阶原点混合矩

$$R_{XX}(t_1, t_2) = E[X(t_1)X(t_2)] \tag{7}$$

自协方差函数(协方差函数)

$X(t_1), X(t_2)$ 的二阶混合中心矩

$$\begin{aligned} C_{XX}(t_1, t_2) &= \mathrm{Cov}[X(t_1), X(t_2)] \\ &= E\{[X(t_1) - \mu_X(t_1)][X(t_2) - \mu_X(t_2)]\} \end{aligned} \tag{8}$$

刻画随机过程自身在两个不同时刻的状态之间统计依赖关系

$$\Psi_X^2(t) = R_{XX}(t, t) \tag{9}$$

$$C_{XX}(t_1, t_2) = R_{XX}(t_1, t_2) - \mu_X(t_1)\mu_X(t_2) \tag{10}$$

当 $t_1 = t_2 = t$ 时，

$$\sigma_X^2(t) = C_{XX}(t, t) = R_{XX}(t, t) - \mu^2(t) \tag{11}$$

二维随机过程

设 $X(t), Y(t)$ 是依赖于同一参数 $t \in T$ 的随机过程，对于不同的 $t \in T$，$(X(t), Y(t)\ )$ 是不同的二维随机变量，称 $\{(X(t), Y(t)\ ), t \in T\}$ 为二维随机过程；

给定二维随机过程 $\{(X(t),Y(t)\ ),t\in T\}$，$t_1,t_2,\cdots,t_n;t_1',t_2',\cdots,t_m'$ 是 T 中任意两组实数，称 $n+m$ 维随机变量

$$(X(t_1),X(t_2),\cdots,X(t_n);Y(t_1'),Y(t_2'),\cdots,Y(t_n')\ ) \tag{12}$$

的分布函数

$$\begin{aligned}&F_X(x_1,x_2,\cdots,x_n;t_1,t_2,\cdots,t_n:y_1,y_2,\cdots,y_m;t_1',t_2',\cdots,t_m'),\\&x_i,y_j\in\mathbf{R},\ i=1,2,\cdots,n,\ j=1,2,\cdots,m\end{aligned} \tag{13}$$

为这个二维随机过程的 $n+m$ 维分布函数或随机过程 $X(t)$ 与 $Y(t)$ 的 $n+m$ 维联合分布函数;

二维随机过程的 $n+m$ 维分布函数族

有限维分布函数族

相互独立

对任意的正整数 n、m，任意的数组 $t_1,t_2,\cdots,t_n\in T$，$t_1',t_2',\cdots,t_m'\in T$，$n$ 维随机变量 $(X(t_1),X(t_2),\cdots,X(t_n)\ )$ 与 m 维随机变量 $(Y(t_1'),Y(t_2'),\cdots,Y(t_n')\ )$ 相互独立;

互相关函数

$X(t)$ 和 $Y(t)$ 的二阶混合原点矩

$$R_{XY}(t_1,t_2)=E[X(t_1)Y(t_2)],\ t_1,t_2\in T \tag{14}$$

互协方差函数

$$\begin{aligned}C_{XY}(t_1,t_2)&=E\{[X(t_1)-\mu_X(t_1)][Y(t_2)-\mu_Y(t_2)]\}\\&=R_{XY}(t_1,t_2)-\mu_X(t_1)\mu_Y(t_2),\ t_1,t_2\in T\end{aligned} \tag{15}$$

不相关

若二维随机过程 $(X(t), Y(t))$ 对任意的 $t_1, t_2 \in T$ 恒有

$$C_{XY}(t_1, t_2) = 0, \tag{16}$$

随机过程 $(X(t), Y(t))$ 是不相关的；

两个随机过程若是相互独立的，且二阶矩存在，则它们必然不相关；但从不相关不能推断它们是相互独立的；

考虑三个随机过程之和：

$$W(t) = X(t) + Y(t) + Z(t), \tag{17}$$

均值函数为

$$\mu_W(t) = \mu_X(t) + \mu_Y(t) + \mu_Z(t), \tag{18}$$

$W(t)$ 的自相关函数

$$\begin{aligned} R_{WW}(t_1, t_2) &= E[W(t_1)W(t_2)] \\ &= R_{XX}(t_1, t_2) + R_{XY}(t_1, t_2) + R_{XZ}(t_1, t_2) \\ &\quad + R_{YX}(t_1, t_2) + R_{YY}(t_1, t_2) + R_{YZ}(t_1, t_2) \\ &\quad + R_{ZX}(t_1, t_2) + R_{ZY}(t_1, t_2) + R_{ZZ}(t_1, t_2) \end{aligned} \tag{19}$$

几个随机过程之和的自相关函数 = 各个随机过程的自相关函数 + 各对随机过程的互相关函数；

若随机过程两两不相关，且各自的均值函数都为 0，则诸互相关函数均为 0，$\rightarrow W(t)$ 的自相关函数 = 各个过程的自相关函数之和，即

$$R_{WW}(t_1, t_2) = R_{XX}(t_1, t_2) + R_{YY}(t_1, t_2) + R_{ZZ}(t_1, t_2), \tag{20}$$

令 $t_1 = t_2 = t$，则

$$\sigma_W^2(t) = \Psi_W^2(t) = \Psi_X^2(t) + \Psi_Y^2(t) + \Psi_Z^2(t), \tag{21}$$

2 独立增量过程

二阶矩过程

设随机过程 $\{X(t), t \in T\}$，若对每一个 $t \in T$，二阶矩 $E[X^2(t)]$ 都存在；

二阶矩过程的相关函数总存在；

正态过程

设随机过程 $\{X(t), t \in T\}$，它的每一个有限维分布都是正态分布，亦即对任意整数 $n \geq 1$ 以及任意 $t_1, t_2, \cdots, t_n \in T$，$(X(t_1), X(t_2), \cdots, X(t_n)\ )$ 服从 n 维正态分布；

正态过程的全部统计特性完全由均值函数和自协方差函数 (自相关函数) 确定；

独立增量过程

给定二阶矩过程 $\{X(t), t \geq 0\}$，称随机变量 $X(t) - X(s)$，$0 \leq s < t$ 为随机过程在区间 $(s, t]$ 上的增量；对任意选定的正整数 n 和任意选定的 $0 \leq t_0 < t_1 < t_2 < \cdots < t_n$，$n$ 个增量

$$X(t_1) - X(t_0), X(t_2) - X(t_1), \cdots, X(t_n) - X(t_{n-1}), \tag{22}$$

相互独立，称 $\{X(t), t \geq 0\}$ 为独立增量过程；

在互不重叠的区间上，状态的增量是相互独立的；

可以证明：对独立增量过程，$X(0) = 0$ 时，它的有限维分布函数族可以由增量 $X(t) - X(s)(0 \leq s < t)$ 的分布所确定；

增量具有平稳性

对任意的实数 h 和 $0 \leq s + h < t + h$，$X(t+h) - X(s+h)$ 与 $X(t) - X(s)$ 具有相同的分布；

增量 $X(t) - X(s)$ 的分布函数只依赖于时间差 $t - s(0 \leq s < t)$，不依赖于 t 和 s 本身；

当增量具有平稳性时，相应的独立增量过程是齐次的或时齐的；

2.1 泊松过程

计数过程:

以 $N(t), t \geq 0$ 表示在时间间隔 $(0, t]$ 内出现的质点数；$\{N(t), t \geq 0\}$ 是一状态取非负整数、时间连续的随机过程；

强度为 λ 的泊松过程:

将增量 $N(t) - N(t_0)$ 记成 $N(t_0, t), 0 \leq t_0 < t$，表示时间间隔 $(t_0, t]$ 内出现的质点数；"在 $(t_0, t]$ 内出现 k 个质点"，即 $\{N(t_0, t) = k\}$，是一事件，其概率记为

$$P_k(t_0, t) = P\{N(t_0, t) = k\}, k = 0, 1, 2, \cdots \tag{23}$$

假设 $N(t)$ 满足如下条件:

1. 在不相重叠的区间上的增量具有独立性；

2. 对于充分小的 Δt

$$P_1(t, t + \Delta t) = P\{N(t, t + \Delta t) = 1\} = \lambda \Delta t + o(\Delta t), \tag{24}$$

$\lambda > 0$：过程 $N(t)$ 的强度，

$o(\Delta t)$：当 $\Delta t \to 0$ 时是关于 Δt 的高阶无穷小；

3. 对于充分小的 Δt

$$\sum_{j=2}^{\infty} P_j(t, t + \Delta t) = \sum_{j=2}^{\infty} P\{N(t, t + \Delta t) = j\} = o(\Delta t) \tag{25}$$

即对于充分小的 Δt，在 $(t, t + \Delta t]$ 内出现 2 个或 2 个以上质点的概率与出现一个质点的概率相比可以忽略不计；

4. $N(0) = 0$

强度为 λ 的泊松流:

相应的质点流或即质点出现的随机时刻 $t_1, t_2, \cdots$;

散粒噪声

2.2 维纳 (Wiener) 过程

Brown 运动

3 Random Processes

[1] A stochastic or random variable is a variable quantity with a definite range of values, each one of which, depending on chance, can be attained with a definite probability. A stochastic variable is defined (a) if the set of possible values is given, and (b) if the probability of attaining each particular value is also given.

The sum of a large number of independent stochastic variables is a stochastic variable. The central limit theorem says that under very general conditions the distribution of the sum tends toward a normal (Gaussian) distribution law as the number of terms is increased.

Central Limit Theorem

Let $x_1, x_2, \cdots, x_n$ be independent stochastic variables with their means equal to 0, possessing absolute moments $\mu_{2+\delta}^{(i)}$ of the order $2+\delta$, where δ is some number > 0. If, denoting by B_n the mean square fluctuation of the sum $x_1 + x_2 + \cdots + x_n$, the quotient

$$\omega_n = \frac{\sum_{i=1}^{n} \mu_{2+\delta}^{(i)}}{B_n^{1+\delta/2}} \tag{26}$$

tends to zero as $n \to 0$, the probability of the inequality

$$\frac{x_1 + x_2 + \cdots x_n}{\sqrt{B_n}} < t \tag{27}$$

tends uniformly to the limit

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \exp\left[-\frac{u^2}{2}\right] \mathrm{d}u \ . \tag{28}$$

For a distribution $f(x_i)$, the absolute moment of order α is defined as

$$\mu_{\alpha}^{(i)} = \int_{-\infty}^{\infty} |x_i|^{\alpha} f(x_i) \mathrm{d}x_i \ . \tag{29}$$

Almost all the probability distributions $f(x)$ of stochastic variables x of interest to us in physical problems will satisfy the requirements of the central limit theorem.

By a random process or stochastic process $x(t)$, we mean a process in which the variable x does not depend in a completely definite way on the independent variable t, which may denote the time. In observations on the different systems of a representative ensemble we find different functions $x(t)$. All we can do is to study certain probability distributions - we can not obtain the functions $x(t)$ themselves for the members of the ensemble.

We can determine, for example

$$p_1(x,t)\mathrm{d}x = \text{Probability of finding } x \text{ in the range } (x, x+\mathrm{d}x) \text{ at time } t \ , \tag{30}$$

$$p_2(x_1,t_1;x_2,t_2)\mathrm{d}x_1\mathrm{d}x_2 = \text{Probability of finding } x \text{ in}$$

$$(x_1, x_1+\mathrm{d}x_1) \text{ at time } t_1 \text{ and in the range } (x_2, x_2+\mathrm{d}x_2) \text{ at time } t_2 \ , \tag{31}$$

If we had an actual oscillogram record covering a long period of time we might construct an ensemble by cutting the record up into strips of equal length T and mounting them one over the other. The probabilities p_1 and p_2 will be found from the ensemble. Proceeding similarly we can form $p_3, p_4, \cdots$. The whole set of probability distributions $P_n(n = 1, 2, \cdots, \infty)$ may be necessary to describe the random process completely. In many important cases p_2 contains all the information we need. When this is true, the random process is called a Markoff process. A stationary random process is one for which the joint probability distributions p_n are invariant under a displacement of the origin of time.

It is useful to introduce the conditional probability $P_2(x_1, 0|x_2, t)\mathrm{d}x_2$ for the probability that given x_1 one finds x in $\mathrm{d}x_2$ at x_2 a time t later. Then it obvious that

$$p_2(x_1, 0; x_2, t) = p_1(x_1, 0)P_2(x_1, 0|x_2, t) \ . \tag{32}$$

4 Wiener-Khintchine Theorem

[1] The Wiener-Khintchine theorem states a relationship between two important characteristics of a random process: the power spectrum of the process and the correlation function of the process.

Suppose we develop one of the records of $x(t)$ for $0 < t < T$ in a Fourier series:

$$x(t) = \sum_{n=1}^{\infty} (a_n \cos 2\pi f_n t + b_n \sin 2\pi f_n t) \ , \tag{33}$$

where $f_n = n/T$. We assume that $\langle x(t) \rangle = 0$, where the angular parentheses $\langle \rangle$ denote time average; because the average is assumed zero there is no constant term in the Fourier series. The Fourier coefficients are highly variable from one record of duration T to another. For many types of noise that a_n, b_n have Gaussian distributions. When this is true, the process is said to be a Gaussian random process.

Let us now imagine that $x(t)$ is an electric current flowing through unit resistance. The instantaneous power dissipation is $x^2(t)$. Each Fourier component will contribute to the total power dissipation. The power in the nth component is

$$P_n = (a_n \cos 2\pi f_n t + b_n \sin 2\pi f_n t)^2 \ . \tag{34}$$

We do not consider cross product terms in the power of the form

$$(a_n \cos 2\pi f_n t + b_n \sin 2\pi f_n t)(a_m \cos 2\pi f_m t + b_m \sin 2\pi f_m t) \tag{35}$$

because for $n \neq m$ the time average of such terms will be zero. The time average of P_n is

$$\langle P_n \rangle = \langle a_n^2 + b_n^2 \rangle / 2 \ , \tag{36}$$

because

$$\langle \cos^2 2\pi f_n t \rangle = \frac{1}{2} \ ,$$
$$\langle \sin^2 2\pi f_n t \rangle = \frac{1}{2} \ ,$$
$$\langle \cos 2\pi f_n t \sin 2\pi f_n t \rangle = 0 \ .$$

We turn to ensemble averages, denoted by a bar over the quantity. An ensemble average is an average over a large set of independent records, each record running in time from 0 to T. For a random process,

$$\overline{a_n} = 0 \;, \quad \overline{b_n} = 0 \;, \quad \overline{a_n b_m} = 0 \;, \tag{37}$$

$$\overline{a_n a_m} = \overline{b_n b_m} = \sigma_n^2 \delta_{nm} \;, \tag{38}$$

where for a Gaussian random process, σ_n is just the standard deviation.

$$\overline{(a_n \cos 2\pi f_n t + b_n \sin 2\pi f_n t)^2} = \sigma_n^2 (\cos^2 2\pi f_n t + \sin^2 2\pi f_n t) = \sigma_n^2 \;. \tag{39}$$

From Eq. (36), the ensemble average of the time average power dissipation associated with the nth component of $x(t)$ is

$$\overline{\langle P_n \rangle} = \sigma_n^2 \;. \tag{40}$$

5 Power Spectrum

[1] Define the power spectrum or spectral density $G(f)$ of the random process as the ensemble average of the time average of the power dissipation in unit resistance per unit frequency bandwidth. If we pick a frequency band width Δf_n equal to the separation between two adjacent frequencies

$$\Delta f_n = f_{n+1} - f_n = \frac{n+1}{T} - \frac{n}{T} = \frac{1}{T} \;, \tag{41}$$

we have

$$G(f_n) \Delta f_n = \overline{\langle P_n \rangle} = \sigma_n^2 \;. \tag{42}$$

$$\overline{x^2(t)} = \sum_n \sigma_n^2 = \sum_n G(f_n) \Delta f_n = \int_0^\infty G(f) \mathrm{d}f \;.$$

The integral of the power spectrum over all frequencies gives the ensemble average total power, which we assume is independent of time, so we speak of it simply as the average total power.

6 Correlation Function

[1] Consider the correlation function

$$C(\tau) = \langle x(t)x(t+\tau)\rangle \tag{43}$$

where the average is over the time t. The function is also called the autocorrelation function. We may take an ensemble average of the time average $\langle x(t)x(t+\tau)\rangle$, so that

$$\begin{aligned}
C(\tau) &= \overline{\langle x(t)x(t+\tau)\rangle} \\
&= \overline{\langle \sum_{n,m} [a_n \cos 2\pi f_n t + b_n \sin 2\pi f_n t][a_m \cos 2\pi f_m (t+\tau) + b_m \sin 2\pi f_m (t+\tau)]\rangle} \\
&= \frac{1}{2}\sum_n \overline{a_n^2 + b_n^2} \cos 2\pi f_n \tau \\
&= \sum_n \sigma_n^2 \cos 2\pi f_n \tau \ .
\end{aligned} \tag{44}$$

$$C(\tau) = \int_0^\infty G(f) \cos 2\pi f \tau \mathrm{d}f \ . \tag{45}$$

The correlation function is the Fourier cosine transform of the power spectrum. If we set in

$$f(x) = \sqrt{\frac{2}{\pi}} \int_0^\infty g(u) \cos ux \mathrm{d}u \ , \tag{46}$$

by

$$u = 2\pi f \ , \tag{47}$$

$$2\sqrt{2\pi}g(u) = G(f) \ , \tag{48}$$

then

$$g(u) = \sqrt{\frac{2}{\pi}} \int_0^\infty f(t) \cos ut \mathrm{d}t \ , \tag{49}$$

gives

$$G(f) = 4 \int_0^\infty C(\tau) \cos 2\pi f \tau \mathrm{d}\tau \ . \tag{50}$$

This, together with Eq. (45), is the Wiener-Khintchine theorem. It has an obvious physical content. The correlation function tells us essentially how rapidly the random process is changing.

References

[1] C. Kittel. *Elementary statistical physics.* Wiley, 1958.