概率论与数理统计

第十二章 随机过程及其统计描述

§ 3 泊松过程及维纳过程

- ◆独立增量过程
- ◆泊松过程的数学模型
- ◆维纳过程的数学模型

一、独立增量过程

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

 $X(t_1) - X(t_0), X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$ 相互独立,则称 $\{X(t), t \ge 0\}$ 为独立增量过程.

特征: 在互不重叠的区间上,状态的增量是相互独立的.

在 X(0) = 0 的条件下,独立增量过程的有限维分布函数族可以由增量 X(t) - X(s) ($0 \le s < t$)的分布确定.

如果对任意的实数 h和 $0 \le s + h < t + h$, X(t+h) - X(s+h)和 X(t) - X(s)具有相同的分布,则称增量具有平稳性.

如果增量具有平稳性,那么增量 *X(t)-X(s)* 的分布函数只依赖于时间差 *t-s*,而不依赖于 *t* 和 *s* 本身. 当增量具有平稳性时,称相应的独立增量过程是齐次的或时齐的.

独立增量过程的协方差函数 $C_x(s,t)$.

设 X(0) = 0, 方差函数 $D_X(t)$ 已知.

记 $Y(t) = X(t) - \mu_X(t)$.

当X(t)具有独立增量时,Y(t)也具有独立增

量; $Y(0) = 0, E[Y(t)] = 0, D_Y(t) = E[Y^2(t)] = D_X(t).$

因此,当 $0 \le s < t$ 时,有

$$C_X(s,t) = E[Y(s)Y(t)]$$

$$= E\{[Y(s)-Y(0)][(Y(t)-Y(s))+Y(s)]\}$$

$$= E[Y(s)-Y(0)]E[Y(t)-Y(s)]+E[Y^2(s)]$$

$$= D_X(s).$$

因此,对任意 $s,t \ge 0$,协方差函数可用方差函数表示为

$$C_X(s,t) = D_X(\min(s,t)).$$

二、泊松过程的数学模型

1.问题的提出

考虑下列随时间的推移迟早会重复出现的事件:

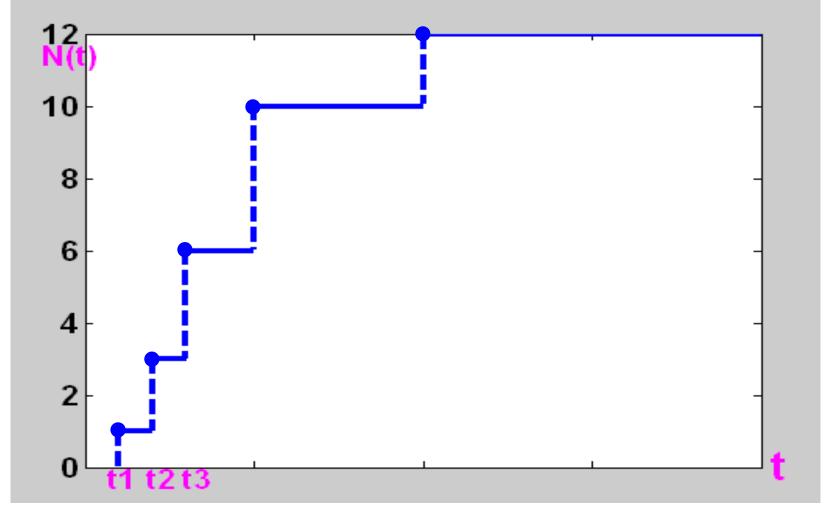
- (1) 自电子管阴极发射的电子到达阳极;
- (2) 意外事故或意外差错的发生;
- (3) 要求服务的顾客到达服务站.

2.问题的分析与求解

将电子、顾客等看作时间轴上的质点,电子到 达阳极、顾客到达服务站等事件的发生相当于质 点出现.因此研究的对象可以认为是随时间推移, 陆续地出现在时间轴上的许多质点所构成的随机 的质点流.

用 $N(t),t \geq 0$ 表示在时间间隔 (0,t]内时间轴上出现的质点数.

 $\{N(t), t \geq 0\}$ 是一个状态取非负整数、时间连续的随机过程,称为计数过程。



计数过程的一个典型样本函数

记 $N(t_0,t) = N(t) - N(t_0), 0 \le t_0 < t$, 表示在时间间隔 $(t_0,t]$ 内出现的质点数.

随机事件 $\{N(t_0,t)=k\}$ 的概率为

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

对N(t)的假设

- (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 ,

 $P_1(t,t+\Delta t) = P\{N(t,t+\Delta t) = 1\} = \lambda \Delta t + o(\Delta t),$ 常数 $\lambda > 0$ 称为过程 N(t) 的强度.

(3)对于充分小的△t,

$$\sum_{j=2}^{\infty} P_j(t,t+\Delta t) = \sum_{j=2}^{\infty} P\{N(t,t+\Delta t) = j\} = o(\Delta t)$$
(4) $N(0) = 0$.

满足条件(1)(2)(3)(4)的计数过程 $\{N(t), t \geq 0\}$ 称作强度为 λ 的泊松过程.

相应的质点流或质 点出现的随机时刻 t_1,t_2,\cdots 称作强度为 λ 的泊松流.

增量的分布律

由于
$$\sum_{k=0}^{\infty} P_k(t_0,t) = 1$$
, 所以根据假设有
$$P_0(t,t+\Delta t) = 1 - P_1(t,t+\Delta t) - \sum_{k=2}^{\infty} P_k(t,t+\Delta t)$$

$$= 1 - \lambda \Delta t + o(\Delta t).$$

概率的计算

当 $\Delta t > 0$ 时, 先计算 $P_0(t_0,t)$.

$$\begin{split} P_0(t_0, t + \Delta t) &= P\{N(t_0, t + \Delta t) = 0\} \\ &= P\{N(t_0, t) + N(t, t + \Delta t) = 0\} \\ &= P\{N(t_0, t) = 0, N(t, t + \Delta t) = 0\}, \end{split}$$

$$\begin{split} P_0(t_0, t + \Delta t) &= P\{N(t_0, t) = 0, N(t, t + \Delta t) = 0\} \\ &= P\{N(t_0, t) = 0\}P\{N(t, t + \Delta t) = 0\} \\ &= P_0(t_0, t)[1 - \lambda \Delta t + o(\Delta t)], \\ &\Rightarrow P_0(t_0, t + \Delta t) - P_0(t_0, t) = -\lambda P_0(t_0, t) \Delta t + o(\Delta t), \\ &\frac{P_0(t_0, t + \Delta t) - P_0(t_0, t)}{\Delta t} = -\lambda P_0(t_0, t) + \frac{o(\Delta t)}{\Delta t}, \end{split}$$

令 $\Delta t \rightarrow 0$, 取极限得微分方程

$$\frac{\mathrm{d}P_0(t_0,t)}{\mathrm{d}t} = -\lambda P_0(t_0,t).$$

因为 $N(t_0,t_0)=0$, 所以 $P_0(t_0,t_0)=1$.

利用初值条件求解微分方程可得

$$P_0(t_0,t) = e^{-\lambda(t-t_0)}, \quad t > t_0.$$

再计算 $P_k(t_0,t), k \geq 1$.

$$P\{N(t_0, t + \Delta t) = k\} = P\{N(t_0, t) + N(t, t + \Delta t) = k\}$$

$$= \sum_{j=0}^{k} P\{N(t, t + \Delta t) = j\}P\{N(t_0, t) = k - j\}$$

$$= \sum_{j=0}^{k} P_j(t, t + \Delta t)P_{k-j}(t_0, t)$$

因为当 $k \ge 2$ 时,

$$\sum_{j=2}^k P_j(t,t+\Delta t)P_{k-j}(t_0,t) \leq \sum_{j=2}^\infty P_j(t,t+\Delta t) = o(\Delta t),$$

所以
$$P_k(t_0,t+\Delta t) = \sum_{j=0}^k P_j(t,t+\Delta t)P_{k-j}(t_0,t)$$

$$= [1 - \lambda \Delta t + o(\Delta t)]P_k(t_0, t)$$

$$+ [\lambda \Delta t + o(\Delta t)]P_{k-1}(t_0, t) + o(\Delta t), \quad (k \ge 1).$$

将此式进行整理后可得

$$P_{k}(t_{0},t+\Delta t) - P_{k}(t_{0},t) = -\lambda P_{k}(t_{0},t)\Delta t + \lambda P_{k-1}(t_{0},t)\Delta t + [P_{k}(t_{0},t) + P_{k-1}(t_{0},t) + 1]o(\Delta t),$$

两边除以 Δt , 令 $\Delta t \rightarrow 0$, 取极限得微分 – 差分方程 $\frac{dP_k(t_0,t)}{dt} = -\lambda P_k(t_0,t) + \lambda P_{k-1}(t_0,t), \quad t > t_0.$

因为 $N(t_0,t_0)=0$,所以 $P_k(t_0,t_0)=0$, $k\geq 1$.

令 k = 1,利用初值条件和 $P_0(t_0, t)$ 的表达式可得 $P_1(t_0, t) = \lambda(t - t_0)e^{-\lambda(t - t_0)}, \quad t > t_0.$

令 k = 2, 利用初始条件和 $P_0(t_0,t)$, $P_1(t_0,t)$ 可得

$$P_2(t_0,t) = \frac{\left[\lambda(t-t_0)\right]^2}{2} e^{-\lambda(t-t_0)}, \quad t > t_0.$$

如此重复,一般地可得到

$$P_{k}(t_{0},t) = P\{N(t_{0},t) = k\}$$

$$= \frac{[\lambda(t-t_{0})]^{k}}{k!} e^{-\lambda(t-t_{0})}, \quad t > t_{0}, k = 0,1,\dots.$$

结论 增量 $N(t_0,t) = N(t) - N(t_0)$ 的概率分布是参数为 $\lambda(t-t_0)$ 的泊松分布,且只与时间差 $t-t_0$ 有关;强度为 λ 的泊松过程是齐次的独 立增量过程.

泊松过程的数字特征

$$N(t) - N(t_0) \sim \pi(\lambda(t - t_0)), \quad t > t_0 \geq 0.$$
 $E[N(t) - N(t_0)] = \operatorname{Var}[N(t) - N(t_0)] = \lambda(t - t_0).$
令 $t_0 = 0$,根据假设 $N(0) = 0$ 可得 $E[N(t)] = \lambda t$, 均值函数 $D_N(t) = \operatorname{Var}[N(t)] = \lambda t$, 方差函数 $\lambda = E[\frac{N(t)}{t}].$

泊松过程的强度等于单位长时间间隔内出现的质点数目的期望值.

$$C_N(s,t) = \lambda \min\{s,t\}, \quad s,t \ge 0.$$
 协方差函数
$$R_N(s,t) = E[N(s)N(t)] = \lambda^2 st + \lambda \min\{s,t\}, s,t \ge 0.$$
 相关函数

若 λ 是时间t的函数 $\lambda = \lambda(t), t \geq 0$, 则称泊松过程是非齐次的.

3.与泊松过程有关的随机变量

(1)等待时间

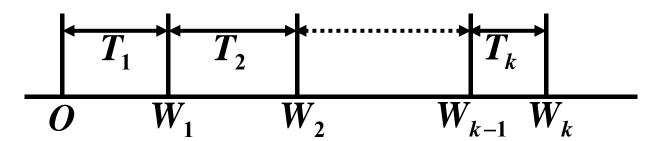
设质点(或事件)依次重复出现的时刻

$$t_1, t_2, \cdots, t_n, \cdots$$

是强度为 λ 的泊松流, $\{N(t), t \geq 0\}$ 为相应泊松过程.

记
$$W_0 = 0$$
, $W_n = t_n$, $n = 1, 2, \cdots$

则 W_n 是随机变量,表示第n个质点(或事件第n次)出现的等待时间.



$$W_n$$
的分布函数 $F_{W_n}(t) = P\{W_n \le t\}$

因为
$$\{W_n > t\} = \{N(t) < n\},$$

所以
$$F_{W_n}(t) = P\{W_n \le t\} = 1 - P\{W_n > t\}$$

$$= 1 - P\{N(t) < n\} = P\{N(t) \ge n\}$$

$$= \sum_{k=n}^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \quad t \ge 0,$$

$$F_{W_n}(t) = 0, \quad t < 0.$$

对时间 t 求导,可得 W_n 的概率密度函数为

$$f_{W_n}(t) = rac{dF_{W_n}(t)}{dt} = egin{cases} rac{\lambda(\lambda t)^{n-1}}{(n-1)!} \mathrm{e}^{-\lambda t}, & t > 0, \ 0, & 其他. \end{cases}$$

泊松流(泊松过程)的等待时间 W_n 服从 Γ 分布.

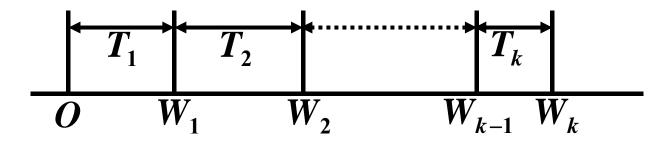
取 n = 1, 得质点(或事件)首次出现的等待时间 W_1 服从指数分布:

$$f_{W_1}(t) = \begin{cases} \lambda e^{-\lambda t}, & t > 0, \\ 0, & 其他. \end{cases}$$

(2)点间间距

记
$$T_i = W_i - W_{i-1}, i = 1, 2, \cdots$$

则 T_i 也是一个连续型随机变量,称为相继出现的第i-1个质点和第i个质点的点间间距。



因为 $T_1 = W_1$,所以 T_1 服从指数分布

$$f_{T_1}(t) = f_{W_1}(t) = \begin{cases} \lambda e^{-\lambda t}, & t > 0, \\ 0, & \text{ #...} \end{cases}$$

当 $i \ge 2$ 时,第i-1个质点出现在时刻 t_{i-1} 的条件下, T_i 的条件分布函数

$$\begin{split} F_{T_{i}|t_{i-1}}(t|t_{i-1}) &= P\{T_{i} \leq t|t_{i-1} = t_{i-1}\} \\ &= P\{N(t_{i-1} + t) - N(t_{i-1}) \geq 1|N(t_{i-1}) = i - 1\} \\ &= P\{N(t_{i-1} + t) - N(t_{i-1}) \geq 1\} \\ &= 1 - P\{N(t_{i-1} + t) - N(t_{i-1}) = 0\} \\ &= 1 - P\{N(t) = 0\} = 1 - e^{-\lambda t}, \quad t > 0, \\ F_{T_{i}|t_{i-1}}(t|t_{i-1}) &= 0, \quad t \leq 0. \end{split}$$

求导可得条件概率密度函数为

$$f_{T_i|t_{i-1}}(t|t_{i-1}) = \begin{cases} \lambda e^{-\lambda t}, & t > 0, \\ 0, & t \leq 0. \end{cases}$$

随机变量 T_i, t_{i-1} 的联合概率密度函数为

$$f(t,t_{i-1}) = \begin{cases} \lambda e^{-\lambda t} f_{t_{i-1}}(t_{i-1}), & t > 0, t_{i-1} > 0, \\ 0, & \text{ #.e.} \end{cases}$$

对上式关于 t_{i-1} 积分,得 $T_i(i=2,3,\cdots)$ 的概率密度为

$$f_{T_{i}}(t) = \int_{0}^{+\infty} \lambda e^{-\lambda t} f_{t_{i-1}}(t_{i-1}) dt_{i-1} = \lambda e^{-\lambda t} \int_{0}^{+\infty} f_{t_{i-1}}(t_{i-1}) dt_{i-1}$$
$$= \lambda e^{-\lambda t}, \quad t > 0,$$

$$f_{T_i}(t)=0, \qquad t\leq 0.$$

$$f_{T_i}(t) = \begin{cases} \lambda e^{-\lambda t}, & t > 0, \\ 0, & t \le 0. \end{cases} \quad i = 2, 3, \dots.$$

结论 点间间距序列 $\{T_i\}$ 服从相同的指数分布 . 理论上, $T_1,T_2,\cdots,T_i,\cdots$ 是相互独立的随机变量 .

定理一

强度为 2的泊松流 (泊松过程)的点间间距是相互独立的随机变量,且服从同一个指数分布

$$f_{T_i}(t) = \begin{cases} \lambda e^{-\lambda t}, & t > 0, \\ 0, & t \le 0. \end{cases} \quad i = 2, 3, \dots$$

定理二

如果任意相继出现的两个质点的点间间距是相互独立的,且服从同一个指数分布

$$f_{T_i}(t) = \begin{cases} \lambda e^{-\lambda t}, & t > 0, \\ 0, & t \le 0. \end{cases} \quad i = 2, 3, \dots$$

则质点流构成了一强度为2的泊松过程.

定理的意义 定理刻画出了泊松过程的特征.

要确定一个计数过程是不是泊松过程,只要用统计方法检验点间间距是否独立,并且服从同一个指数分布.

例1 设 X_t 为(0,t]内来到某一服务机构的顾客数,再设 $\{X_t, t \ge 0\}$ 是一个以 λ 为强度的泊松过程,令 $0 < \tau_1 < \tau_2 < \cdots < \tau_n < \cdots$,

 τ_n 是第n 个顾客到达时刻, $T_1 = \tau_1$, $T_2 = \tau_2 - \tau_1$,…,即 T_n 是第n-1 个至第n 个顾客到达时刻的间距,试求 τ_n 的分布函数,数学期望与方差.

解 由于 $P\{X_t \le n-1\} = P\{\tau_{n-1} \ge t\} = P\{\tau_n > t\}$, 所以 τ_n 的分布函数为

$$F_{\tau_n}(t) = P\{\tau_n \le t\} = 1 - P\{X_t \le n - 1\}$$

$$= 1 - e^{-\lambda t} \left[1 + \lambda t + \dots + \frac{(\lambda t)^{n-1}}{(n-1)!}\right] \quad (n \ge 1, t > 0),$$

两端对t求导,得 τ_n 的分布密度为

$$f_{\tau_n}(t) = \begin{cases} \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}, & t > 0, \\ 0, & t \leq 0, \end{cases}$$

即 $F_{\tau_n}(t)$ 是参数为n和 λ 的 Γ 分布.

$$E(\tau_n) = \int_0^{+\infty} t \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} dt = \frac{\Gamma(n+1)}{\Gamma(n)\lambda} = \frac{n}{\lambda},$$

$$E(\tau_n^2) = \int_0^{+\infty} t^2 \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} dt = \frac{\Gamma(n+2)}{\lambda^2 \Gamma(n)} = \frac{n(n+1)}{\lambda^2},$$

故
$$D(\tau_n) = E(\tau_n^2) - [E(\tau_n)]^2 = \frac{n(n+1)}{\lambda^2} - \frac{n^2}{\lambda^2} = \frac{n}{\lambda^2}.$$

例2 假定进入中国上空的流星的个数构成一个泊松过程,且平均每年进入10000个,每个流星在大气中未烧完而以陨石落于地面的概率为0.0001,设W是一个月内落入中国地面的陨石数,试求E(W)和 $P\{W \geq 2\}$.

解 设 X 是 t 年内进入中国上空的流星数,则

$$P\{X_t = k\} = e^{-10000t} \frac{(10000t)^k}{k!}, \quad k \ge 0,$$

由二项分布

$$P\{W=j \mid X_{\frac{1}{12}}=k\} = \binom{k}{j} (0.0001)^{j} (0.9999)^{k-j}, \\ j=0,1,\cdots,k.$$

$$P\{W \ge 2\} = \sum_{k=2}^{\infty} P\{W = k\}$$

$$= \sum_{k=2}^{\infty} \sum_{j=k}^{\infty} P\{X_{\frac{1}{12}} = j\} P\{W = k | X_{\frac{1}{12}} = j\}$$

$$= \sum_{k=2}^{\infty} \sum_{j=k}^{\infty} \frac{(10000/12)^{j}}{j!} e^{\frac{-10000}{12}} \cdot {k \choose j} \cdot (0.0001)^{j} (0.9999)^{j-k}$$

$$= e^{-\frac{1}{12}} \sum_{k=2}^{\infty} (\frac{1}{12})^{k} \frac{1}{k!} = e^{-\frac{1}{12}} [e^{\frac{1}{12}} - 1 - \frac{1}{12}]$$

$$= 1 - e^{-\frac{1}{12}} - \frac{1}{12} e^{-\frac{1}{12}} \approx 0.003285217485.$$

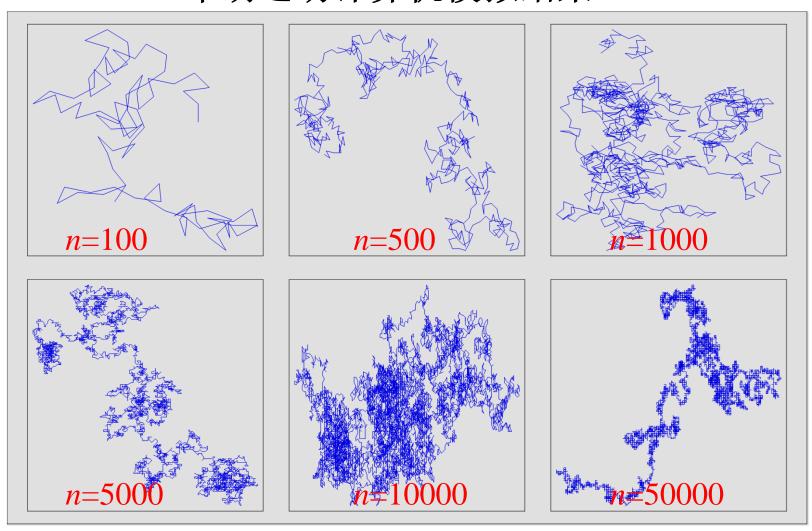
三、维纳过程的数学模型

1.布朗运动简介

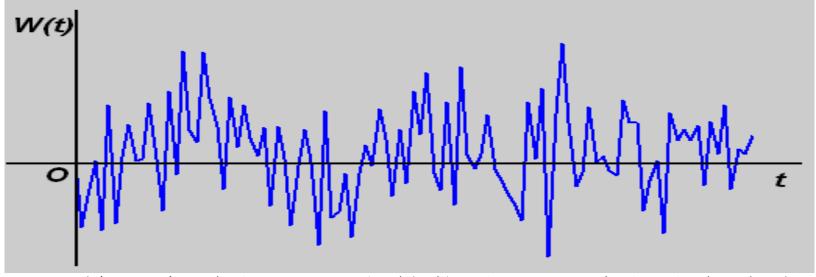
英国植物学家布朗(Brown)在显微镜下,观察 漂浮在平静的液面上的微小粒子,发现它们不断地 进行着杂乱无章的运动,这种现象称为布朗运动.

爱因斯坦(Enisten)1905年提出一种理论,认为微粒的这种运动是由于受到大量随机的、相互独立的分子碰撞的结果.

布朗运动计算机模拟结果

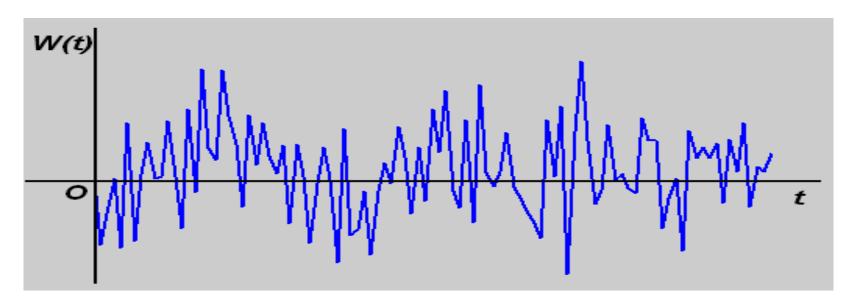


以W(t)表示运动中一微粒从时刻t=0到时刻t>0的位移的横坐标,且W(0)=0.



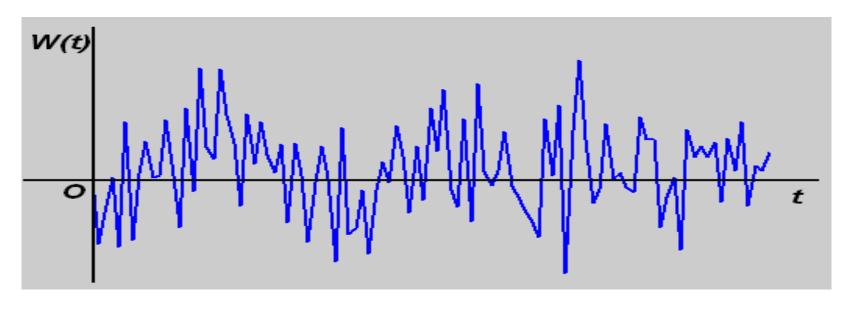
粒子在时段 (s,t]上的位移可以看成是许多微小位移的代数和.

根据中心极限定理,假设 位移W(t)-W(s) 服从正态分布.



由于粒子的运动完全是由液体分子的不规则碰撞而引起的,因此,在不相重叠的时间间隔内,碰撞的次数、大小和方向可假定是相互独立的.

位移W(t)具有独立的增量



液面处于平衡状态,这时粒子在一时段上位移的概率分布可以认为只依赖于这时段的长度,而与观察的起始时刻无关.

位移W(t)具有平稳的增量

2.维纳过程的数学模型

给定二阶矩过程 $\{W(t), t \geq 0\}$, 如果它满足

- (1)具有独立增量;
- (2)对任意的 $t > s \ge 0$,增量

则称此过程为维纳过程.

3.维纳过程的特征

维纳过程增量的分布只与时间差有关,所以 维纳过程是齐次的独立增量过程,也是正态过程. 其分布完全由均值函数和自协方差函数(或者自 相关函数)所确定.

$$E[W(t)] = 0,$$

$$D_W(t) = \sigma^2 t,$$

$$C_W(s,t) = R_W(s,t)$$

$$= \sigma^2 \min\{s,t\}, \quad s,t > 0.$$

例3 设 $\{W_t, t \in [0,\infty)\}$ 是以 σ^2 为参数的维纳过程.令 (a) $X_t = (1-t)W_{\frac{t}{1-t}}$, 0 < t < 1; (b) $X_t = W_t^2$, $0 \le t \le \infty$, 试求 $\mu_{X}(t)$, $C_{XX}(s,t)$, 并判断 $\{X_{t}\}$ 是否是正态过程. 解 $(a)\mu_X(t)=(1-t)E[W_{\underline{t}}]=0,$ $C_{XX}(s,t) = E[X_s, X_t]$ $= (1-s)(1-t)Cov[W_{\frac{t}{1-t}}, W_{\frac{s}{1-s}}]$ $= (1-t)(1-s) \cdot [\min(\frac{t}{1-t}, \frac{s}{1-s})]\sigma^{2}$ $= \begin{cases} t(1-s)\sigma^{2}, t \leq s, \\ s(1-t)\sigma^{2}, t > s. \end{cases} \{X_{t}, t \geq 0\}$ 是正态过程.

(b)
$$\mu_X(t) = E[W_t^2] = D[W_t] = \sigma^2 t$$
,
 $C_{XX}(s,t) = \text{Cov}(W_s^2, W_t^2)$

$$=2C_{XX}^{2}(s,t)=2[\min(s,t)\sigma^{2}]^{2}.$$

因为 $X_t = W_t^2 \ge 0$,所以 $\{X_t\}$ 不是正态过程.

小结

1. 独立增量过程

特征: 在互不重叠的区间上, 状态的增量是相互独立.

2. 泊松过程

数学模型 增量的概率分布

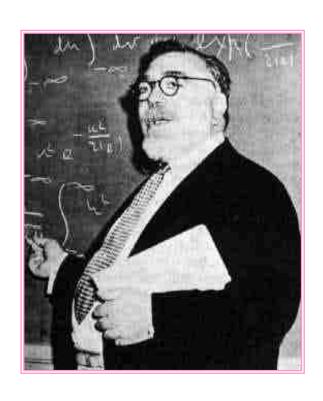
数字特征 有关的随机变量

3. 维纳过程

布朗运动 数学模型 数字特征

作业: 课后习题 10、11

维纳资料



Norbert Wiener

Born: 26 Nov. 1894 in Columbia, Missouri, USA Died: 18 Mar. 1964 in Stockholm, Sweden