知道了随机过程的有限维分布族就知道了过程 $\{X(t), t \in T\}$ 中任意 n 个随机变量的联合分布, 也就完全了解了这些变量之间的相互依赖关系. 有限维分布也有对称性, 它与变量 $X(t_1), \cdots, X(t_n)$ 的排序无关. 对 $\{1, \cdots, n\}$ 的任一置换 $(i_1, \cdots, i_n)$ 有

$$F_{t_{i_1}, \cdots, t_{i_n}}(x_{i_1}, \cdots, x_{i_n}) = F_{t_1, \cdots, t_n}(x_1, \cdots, x_n).$$

有限维分布还有相容性, 意思是当某些 $x \to \infty$ 时高维分布的边缘分布与相应的低维分布是一致的. 也即对 $m < n$ 有

$$F_{t_1, \cdots, t_m, t_{m+1}, \cdots t_n}(x_1, \cdots, x_m, \infty, \cdots \infty) = F_{t_1, \cdots, t_m}(x_1, \cdots, x_m).$$

有趣的是, 若一族给定的分布函数有上述对称性和相容性则保证了存在一个随机过程 $\{X(t), t \in T\}$, 使它的有限维分布族正好就是给定的分布函数族. 这是 1931 年由 Kolmogorov 证明的相容性基本定理.

定义 1.2 如果随机过程 $X(t)$ 对任意的 $t_1, \cdots, t_n \in T$ 和任何 h 有

$$(X(t_1+h), \cdots, X(t_n+h)) \overset{d}{=} (X(t_1), \cdots, X(t_n)), \tag{1.3}$$

则称为严格平稳的.

条件 (1.3) 很强也不易验证, 所以退而求其次有所谓宽平稳或二阶平稳过程. 引入如下定义.

定义 1.3 如果随机过程的所有二阶矩存在并有 $EX(t) = m$ 及协方差函数 $R_X(t, s)$ 只与时间差 $t - s$ 有关, 则称为宽平稳的或二阶矩平稳的.

对于宽平稳过程, 由于对 $-\infty < s, t < \infty, R_X(s, t) = R_X(0, t-s)$, 所以可以记为 $R_X(t-s)$. 显然对所有 $t, R_X(-t) = R_X(t)$, 即为偶函数. 所以 $R_X(t)$ 的图形是关于坐标轴对称的. 其在 0 点的值 $R_X(0)$ 就是过程的方差函数, 即 $R_X(0) = \mathrm{Var}X(t)$.

尽管 $X(t)$ 之间常常不是相互独立的, 但人们可以假定过程的增量之间是相互独立的. 这就是如下定义.

定义 1.4 如果对任意的 $t_1 < t_2 \cdots < t_n, t_1, \cdots, t_n \in T$, 随机变量 $X(t_2) - X(t_1), X(t_3) - X(t_2), \cdots, X(t_n) - X(t_{n-1})$ 是相互独立的, 则 $X(t)$ 称为独立增量过程. 如果进一步有对任意的 $t_1, t_2, X(t_1+h) - X(t_1) \overset{d}{=} X(t_2+h) - X(t_2)$, 则过程称为有平稳独立增量的过程.

可以证明平稳独立增量过程的均值函数一定是 t 的线性函数. 我们以后要介取极限即知给定 $S = 1$, T 是 $[0,1]$ 上的均匀分布. 习题 1 第 10 题提供另一解法, 不妨一试.

条件期望有些重要的性质我们总结为下面的命题.

命题 1.1 (a) 若 X 与 Y 独立, 则 $E(X \mid Y = y) = EX$.

(b) 条件期望有所谓的平滑性:

$$EX = \int E(X \mid Y = y) dF_Y(y) = E[E(X \mid Y)]. \tag{1.14}$$

(c) 对随机变量 X, Y 的函数 $\phi(X, Y)$ 恒有

$$E[\phi(X, Y) \mid Y = y] = E[\phi(X, y) \mid Y = y]. \tag{1.15}$$

1.2.2 矩母函数及生成函数

定义 1.5 随机变量 X 的矩母函数定义为随机变量 $\exp\{tX\}$ 的期望, 记作 $g(t)$, 即

$$g(t) = E(\exp\{tX\}) = \int \exp\{tx\} dF(x). \tag{1.17}$$

矩母函数刻画了随机变量的许多特征, 是研究它们特性的重要工具. 当矩母函数存在时, 它唯一地确定了 X 的分布. 通过 $g(t)$ 可以求出 X 的各阶矩, 即有

$$E[X^n] = g^{(n)}(0), \quad n \geqslant 1, \tag{1.18}$$

其中 $g^{(n)}(t)$ 是 $g(t)$ 的 n 阶导数在 t 的取值. 通过在积分号下求导数容易得到 (1.18) 式的证明. 对相互独立的随机变量 X 和 Y, 它们和的矩母函数就等于其矩母函数的积:

$$g_{X+Y}(t) = g_X(t) g_Y(t). \tag{1.19}$$

熟记一些常用随机变量分布的矩母函数有助于辨别和刻画不同的随机变量. 书末附表中列出了常用随机变量的分布, 矩母函数, 均值和方差.

如果随机变量 Y 等于一串独立随机变量 $X_1, X_2, \cdots, X_N$ 的和, 且下标 N 也是一个整数值随机变量, 则 Y 是随机变量的随机和. 这在实际问题中常常会碰到. 比如, N 是在某一时段到达服务台的顾客数, 而 X_i 代表第 i 个顾客所需要的服务时间, 则 $Y = \sum_{i=1}^{N} X_i$ 代表总服务时间. 这是排队论问题中的重要变量. 又比如在讨论人口和生物群体生长模型时, 在某一地区某类生物的总数为 N, X_i 为第 i 个体的后代数, 则 Y 就是该群体的总数. 下面讨论随机和 Y 的矩母函数与数字特征的计算.

例 1.12 随机和的矩母函数. 记 $X_1, X_2, \cdots$ 为一串独立同分布的随机变量, N 为非负整数值随机变量且与 X 序列相独立. Y 为随机和 $\sum_{i=1}^{N} X_i$. 求 Y 的矩母函数 $g_Y(t)$.

解 为求 g_Y, 先算条件期望

$$\begin{aligned} E[e^{tY} \mid N = n] &= E\left[\exp\left\{t \sum_{i=1}^{N} X_i\right\} \mid N = n\right] \\ &= E\left[\exp\left\{t \sum_{i=1}^{n} X_i\right\} \mid N = n\right] \\ &= E\left[\exp\left\{t \sum_{i=1}^{n} X_i\right\}\right] = [g_X(t)]^n. \end{aligned}$$

于是有 $g_Y(t) = E[\exp\{tY\}] = E\{E[\exp\{tY\}|N]\} = E[(g_X(t))^N]$. 对 $g_Y(t)$ 关于 t 求导即有

$$\begin{aligned} g'_Y(t) &= E[N(g_X(t))^{N-1} g'_X(t)], \\ g''_Y(t) &= E[N(N-1)(g_X(t))^{N-2}(g'_X(t))^2 + N(g_X(t))^{N-1} g''_X(t)]. \end{aligned}$$

将 $t = 0$ 代入上面两式得

$$\begin{aligned} EY &= E[NE(X)] = EN \cdot EX, \\ EY^2 &= EN \cdot \mathrm{Var}X + EN^2 \cdot E^2 X, \\ \mathrm{Var}Y &= EN \cdot \mathrm{Var}X + E^2 X \cdot \mathrm{Var}N. \end{aligned} \tag{1.20}$$

定义 1.6 若 X 为离散随机变量, 则期望 $E(s^X)$ 为其概率生成函数, 记作 $\phi_X(s)$. 特别地, 若 $P(X = k) = p_k, k = 0, 1, 2, \cdots$, 则 $\phi_X(s) = \sum_{k=0}^{\infty} p_k s^k$.

由定义知, $\phi_X(s)$ 是以概率 p_k 为系数的幂级数. $\phi_X(s)$ 与 X 的概率分布也是一一对应的, 而且有 $p_0 = \phi_X(0)$,

$$p_k = \frac{1}{k!} \frac{d^k}{ds^k} \phi_X(s)|_{s=1}, \quad k = 1, 2, \cdots. \tag{1.21}$$

与矩母函数的性质类似, 若 X, Y 为独立随机变量, 则它们和的概率生成函数为

$$\phi_{X+Y}(s) = \phi_X(s) \phi_Y(s), \tag{1.22}$$

而 X 的期望及有关高阶矩也可由 $\phi_X(s)$ 关于 s 的导函数在 $s = 1$ 时的值来确定:

$$\begin{aligned} EX &= \phi'_X(s)|_{s=1}, \\ E\{X(X-1)\cdots(X-r+1)\} &= \frac{d^r}{ds^r} \phi_X(s)|_{s=1}. \end{aligned} \tag{1.23}$$

在随机和的情形, 若 X_i 为离散随机变量, 其概率生成函数为 $\phi_X(s)$, 则对 $Y = \sum_{i=1}^{N} X_i$ 有

$$\begin{aligned} \phi_Y(s) &= E(s^Y) = E\{E[s^Y \mid N]\} \\ &= E\{(\phi_X(s))^N\} = \phi_N(\phi_X(\phi)). \end{aligned} \tag{1.24}$$

定义 1.7 设 $\{X_n, n \geqslant 1\}$ 是一列随机变量, 若存在随机变量 X, 使对 $\forall \varepsilon > 0$, 有

$$\lim_{n \to \infty} P(|X_n - X| \geqslant \varepsilon) = 0, \tag{1.25}$$

则称随机变量序列 $\{X_n, n \geqslant 1\}$ 依概率收敛于 X, 记为 $X_n \xrightarrow{p} X$.

如果事件 $\{\omega : \lim_{n \to \infty} (X_n(\omega) - X(\omega)) = 0\}$ 的概率为 1, 即

$$P(\lim_{n \to \infty} (X_n - X) = 0) = 1, \tag{1.26}$$

则称随机变量序列 $\{X_n, n \geqslant 1\}$ 几乎必然 (almost sure) 收敛于 X, 记为 $X_n \to X$, a.s., 也称随机变量序列以概率 1 收敛于 X.

定义 1.8 设随机变量 X 和 $X_n, n \geqslant 1$, 都有有限的二阶矩, 如果

$$\lim_{n \to \infty} E(X_n - X)^2 = 0, \tag{1.27}$$

则称 X_n 均方收敛于 X, 记为 $X_n \xrightarrow{L_2} X$. 这里 "均" 是平均, "方" 是平方.

在例 1.13 中, 由于 $E\left(\frac{S_n}{n} - p\right)^2 = \frac{p(1-p)}{n}$, 故 $S_n/n \xrightarrow{L_2} p$.

这三种收敛性之间的关系是: 均方收敛和几乎必然收敛都蕴涵依概率收敛, 反过来不必成立; 其次均方收敛和几乎必然收敛互不包含.

定义 2.1 一个整数值随机过程 $\{N(t), t \geqslant 0\}$ 满足下述三个条件就称作强度为 $\lambda > 0$ 的 Poisson 过程:

(i) $N(0) = 0$;

(ii) $N(t)$ 是独立增量过程;

(iii) 对任何 $t > 0$, $s \geqslant 0$, 增量 $N(s+t) - N(t)$ 服从参数为 λt 的 Poisson 分布, 即

$$P\{N(s+t) - N(t) = k\} = \frac{(\lambda t)^k \exp\{-\lambda t\}}{k!}, \quad k = 0, 1, \cdots. \tag{2.1}$$

定义的条件 (i) 告诉我们随机事件从时刻 0 开始计数, 条件 (iii) 是过程称为 Poisson 过程的直接理由. 由 Poisson 分布的性质我们马上可以求出 $EN(t) = \mathrm{Var}[N(t)] = \lambda t$. 增量 $N(s+t) - N(t)$ 代表时间区间 $(s, s+t]$ 中发生的随机事件数, 由条件 (iii), 这一增量的分布与 s 无关因而增量具有平稳性. 因此条件 (ii) 和 (iii) 充分刻画了模型所要求的两个性质: 前后的独立性与时间上的均匀性. 强度 λ 有时也称为速率, 它描绘随机事件发生的频繁程度.

稀有事件的概率常服从 Poisson 分布. 这是由于当试验次数很多而每次试验成功的概率很小时, Poisson 分布可以逼近二项分布. 若记 N 为试验次数, p 为成功概率, 而当 N 很大时, Np 趋于一个常量 λ, 则 N 次试验中的总成功次数近似地服从参数为 λ 的 Poisson 分布. 这一想法很自然地可以推广到随机过程的情况. 记 $[0, \infty)$ 为观察过程的时间轴, 0 代表起始时刻, $N(b) - N(a)$ 代表时间区间 $(a, b]$ 上发生的事件数. 我们特作如下假定:

(1) 在不相交区间中事件发生的数目相互独立, 也即对任何整数 $n = 1, 2, \cdots$, 设时刻 $t_0 = 0 < t_1 < t_2 < \cdots < t_n$, 增量 $N(t_1) - N(t_0), N(t_2) - N(t_1), \cdots N(t_n) - N(t_{n-1})$ 相互独立;

(2) 对任何时刻 t 和正数 h, 随机变量 (增量)$N(t+h) - N(t)$ 的分布只依赖于区间长度 h 而不依赖时刻 t;

(3) 存在正常数 λ, 当 $h \downarrow 0$ 时, 使在长度为 h 的小区间中事件至少发生一次的概率

$$P\{N(t+h) - N(t) \geqslant 1\} = \lambda h + o(h);$$

(4) 在小区间 $(t, t+h]$ 发生两个或两个以上事件的概率为 $o(h)$ (可以忽略不计), 即当 $h \downarrow 0$,

$$P\{N(t+h) - N(t) \geqslant 2\} = o(h).$$

从逼近的观点看, (1) 说明试验是独立的; (2) 说明在每个长度相同的小区间上事件发生有相同的概率 p; (3) 告诉我们成功 (事件发生) 概率 $p \doteq \lambda h$, 而且 p 很小; (4) 是说明事件不发生的概率为 $1 - \lambda h \doteq 1 - p$. 这正好是独立 Bernoulli 试验的模型. 若观察区间为 $[0, 1]$, 则 $N \doteq \frac{t}{h}$, 所以当 $h \downarrow 0$ 时, $N \to \infty$, 而 $Np \doteq \lambda t$. 从而二项分布的极限是参数为 λt 的 Poisson 分布. 从直观意义看, 如前所述 (1) 为前后的独立性, (2) 为时间上的均匀性或齐次性, (3) 表明事件是稀有的, 而 (4) 则称为相继性 (orderliness), 意思指事件是一件一件地发生的, 在同一瞬间同时发生多个事件的可能性很小很小. 基于这些假定我们可以证明如下结论.

命题 2.1 满足假定 (1) ~ (4) 的随机过程 $N(t)$ 为 Poisson 过程.

图 2.1 中第 $n - 1$ 次与第 n 次事件间的间隔时间记作 X_n, 而 $W_n = \sum_{i=1}^{n} X_i$ 为第 n 次事件的到达或等待时间. 我们来求 X_n 与 W_n 的分布.

命题 2.2 $X_n, n = 1, 2, \cdots$ 是均值为 $\frac{1}{\lambda}$ 的独立同分布的指数随机变量, W_n 服从参数为 n 和 λ 的 Γ 分布.

证 事件 $\{X_n > t\}$ 表示第一次事件发生在时刻 t 之后, 其发生当且仅当在时间区间 $(0, t]$ 中 Poisson 过程不曾有事件发生过. 所以,

$$P\{X_1 > t\} = P\{N(t) = 0\} = \mathrm{e}^{-\lambda t},$$

而

$$\begin{aligned} P\{X_2 > t \mid X_1 = s\} &= P\{(s, s+t] \text{ 中事件不发生 } \mid X_1 = s\} \\ &= P\{(s, s+t] \text{ 中事件不发生}\} (\text{独立增量性}) \\ &= P\{(0, t] \text{ 中事件不发生}\} (\text{平稳增量性}) \\ &= \mathrm{e}^{-\lambda t}. \end{aligned}$$

这就可以得出结论: X_2 也服从以 λ 为参数的指数分布而且 X_1 与 X_2 是独立的. 类似地, 可对其他 X_i 证明命题的结论. 利用习题 1 第 13 题可以知道 W_n 服从 Γ 分布, 参数为 n, λ. 但我们宁愿在这里再给一个直接的证明. 因事件 $\{N(t) \geqslant n\}$ 是与 $W_n \leqslant t$ 等价的, 它们都表明第 n 次事件发生在时刻 t 之前, 或者换言之, 到时刻 t 已经至少发生了 n 件事. 于是

$$P\{W_n \leqslant t\} = P\{N(t) \geqslant n\} = \sum_{j=n}^{\infty} \mathrm{e}^{-\lambda t} \frac{(\lambda t)^j}{j!}, \tag{2.5}$$

对 W_n 的分布函数 (2.5) 式关于 t 求导即可求出 W_n 的密度函数

$$\begin{aligned} f_{W_n}(t) &= -\sum_{j=n}^{\infty} \lambda \mathrm{e}^{-\lambda t} \frac{(\lambda t)^j}{j!} + \sum_{j=n}^{\infty} \lambda \mathrm{e}^{-\lambda t} \frac{(\lambda t)^{j-1}}{(j-1)!} \\ &= \lambda \mathrm{e}^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}. \end{aligned}$$

的 Poisson 过程 $\{N(t), t \geqslant 0\}$. λ 越大表明事件平均间隔时间 $\frac{1}{\lambda}$ 越短, 事件的发生就越频繁, 强度也就越大.

给定了到时刻 t 总计发生了 n 件事, 在过去某时刻 $u < t$ 发生了 k 件事的条件概率可由习题 2 第 1 题知道是二项分布. 但人们常常对给定 $N(t) = n$ 事件发生后 $W_1, \cdots, W_n$ 的联合分布更感兴趣. 关于这一条件联合分布, 我们有如下定理.

定理 2.1 若 $N(t), t \geqslant 0$ 为 Poisson 过程, 则给定 $N(t) = n$ 下等待时间 $W_1, \cdots, W_n$ 的联合密度为

$$f_{W_1, \cdots, W_n|N(t)=n}(w_1, \cdots, w_n|n) = \frac{n!}{t^n}, \quad 0 < w_1 < \cdots < w_n \leqslant t. \tag{2.6}$$

例 2.2 顾客依速率为 λ 的 Poisson 过程到达车站. 若火车在时刻 t 离站, 问在 $(0, t]$ 区间里顾客的平均总等待时间是多少?

解 作为依 Poisson 过程到达的第一位顾客, 他的到达时间为 W_1, 等到时刻 t 发车需等待 $t - W_1$. 而第 i 位旅客的等待时间为 $t - W_i$. 在 $(0, t]$ 区段总共来了 $N(t)$ 位客人, 所以总等待时间为 $\sum_{i=1}^{N(t)} (t - W_i)$. 而所要求的平均总等待时间就是 $E\left[\sum_{i=1}^{N(t)} (t - W_i)\right]$. 为求出它可以先求条件期望

$$\begin{aligned} E\left[\sum_{i=1}^{N(t)} (t - W_i) | N(t) = n\right] &= E\left[\sum_{i=1}^{n} (t - W_i) | N(t) = n\right] \\ &= nt - E\left[\sum_{i=1}^{n} W_i | N(t) = n\right]. \end{aligned}$$

注意到给定 $N(t) = n, W_i, i = 1, \cdots, n$ 的联合密度是与 $(0, t]$ 上均匀分布中随机样本 $U_i, i = 1, \cdots, n$ 的次序统计量 $U_{(i)}, i = 1, \cdots, n$ 的联合密度是一样的. 于是,

$$E\left[\sum_{i=1}^{n} W_i | N(t) = n\right] = E\left[\sum_{i=1}^{n} U_{(i)}\right] = E\left[\sum_{i=1}^{n} U_i\right] = \frac{nt}{2}.$$

因此,

$$E\left[\sum_{i=1}^{N(t)} (t - W_i) | N(t) = n\right] = nt - \frac{nt}{2} = \frac{nt}{2}.$$

最后得到

$$E\left[\sum_{i=1}^{N(t)} (t - W_i)\right] = \frac{t}{2} E[N(t)] = \frac{\lambda t^2}{2}.$$

可以看到顾客平均总等待时间是和 t^2 成正比的, 比例因子的大小取决于 Poisson 过程的强度 λ.

$$P\{N(t+h) - N(t) = 1\} = \lambda h \cdot \mathrm{e}^{-\lambda h} = \lambda h + o(h),$$

如果允许这一比例因子依赖于时刻 t, 所得的过程就是非齐次 Poisson 过程. 其参数为 t 的函数 $\lambda(t)$. 这也提供了过程增量不平稳的例子. 此时,

$$\begin{aligned} &P\{N(t+h) - N(t) = k\} \\ &= \frac{\left(\int_t^{t+h} \lambda(u) du\right)^k \exp\left(-\int_t^{t+h} \lambda(u) du\right)}{k!}, \quad k = 0, 1, \cdots. \end{aligned} \tag{2.8}$$

人们只需将 2.1 节中的假定 (3) 改为 (3)$'$:

$$P\{N(t+h) - N(t) \geqslant 1\} = \lambda(t) h + o(h)$$

就可以完全类似地证明.

2.3.5 更新过程

由命题 2.2 知事件的间隔时间 X_i 相互独立且有相同的均值为 $\frac{1}{\lambda}$ 的指数分布. 这是 Poisson 过程的重要的特征. 如果把时间间隔 X_i 服从的指数分布改为一般的分布函数 $F(x)$, 那么所得的将会是什么过程呢? 这就是所谓的更新过程.

定义 2.3 如果 X_i, $i = 1, 2, \cdots$ 为一串非负的随机变量, 它们独立同分布, 分布函数为 $F(x)$. 记 $W_0 = 0, W_n = \sum_{i=1}^{n} X_i$, W_n 表示第 n 次事件发生的时刻, 则称

$$N(t) = \max\{n : W_n \leqslant t\}$$

为更新过程.

定义中 $N(t)$ 代表了到时刻 t 时事件的总数. W_i, $i = 1, 2, \cdots$, 也常常称为是更新点, 在这些更新点上过程又重新开始. 在更新过程中事件平均发生的次数称为是更新函数, 记作 $m(t)$, 即 $m(t) = E[N(t)]$. 更新理论的主体是研究更新函数的性质, 我们仅给出最基本的.

命题 2.4 更新过程 $N(t)$ 的分布

$$P\{N(t) = n\} = F^{(n)}(t) - F^{(n+1)}(t),$$

而更新函数 $m(t) = \sum_{n=1}^{\infty} F^{(n)}(t)$. 其中 $F^{(n)}(t)$ 为 $F(t)$ 的 n 重卷积, $F(t)$ 即为 X_i 的分布函数.

定理 3.1 Markov 链的 n 步转移概率矩阵满足

$$P_{ij}^{(n)} = \sum_{k=0}^{\infty} P_{ik} P_{kj}^{(n-1)}, \tag{3.5}$$

在上式中我们约定 $P_{ii}^{(0)} = 1$, 当 $j \neq i$ 时 $P_{ij}^{(0)} = 0$.

定义 3.3 (可达与互达) 如果对某一 $n \geqslant 0$, 有 $P_{ij}^{(n)} > 0$, 则称状态 j 是从状态 i 可达的 (accessible), 记作 $i \to j$. 它表示从状态 i 经过有限步的转移可以到达状态 j. 两个互相可达的状态 i 和 j 则称为互达的 (communicate), 记作 $i \leftrightarrow j$.

如果两个状态 i 和 j 不是互达的, 那就有对所有 $n \geqslant 0, P_{ij}^{(n)} = 0$ 或者对所有 $n \geqslant 0$, $P_{ji}^{(n)} = 0$, 或者两者都成立. 三种情况必居其一. 互达性是一种数学上的等价关系, 也就是说它满足自反性、对称性和传递性.

命题 3.1 互达性是等价关系, 即

(i) $i \leftrightarrow i$, 自反性;

(ii) 若 $i \leftrightarrow j$, 则 $j \leftrightarrow i$, 对称性;

(iii) 若 $i \leftrightarrow j$, 且 $j \leftrightarrow k$, 则 $i \leftrightarrow k$, 传递性.

定义 3.4 状态 i 的周期. 设 i 为 Markov 链的一个状态, 使 $P_{ii}^{(n)} > 0$ 的所有正整数 n $(n \geqslant 1)$ 的最大公约数称作是状态 i 的周期, 记作 $d(i)$. 如果对所有 $n \geqslant 1$, 都有 $P_{ii}^{(n)} = 0$ 则约定周期为 ∞; $d(i) = 1$ 的状态 i 则称为是非周期的.

由定义立即可知, 如果 n 不能被周期 $d(i)$ 整除, 则必有 $P_{ii}^{(n)} = 0$.

命题 3.2 如果 $i \leftrightarrow j$ 则 $d(i) = d(j)$.

推论 3.1 如果 $P_{ji}^{(m)} > 0$, 则存在正整数 N 使得对 $n \geqslant N$ 恒有 $P_{ji}^{(m+nd(i))} > 0$.

讨论这些性质的意义何在呢? 它可以帮助我们研究当 n 很大时 $P_{ji}^{(n)}$ 的极限是否存在? 存在的条件是什么? 如果存在又如何简便地求出它们? 周期性是一种等价类中全体状态共有的性质. 当周期为 1 时, Markov 链称为是非周期的. 对非周期不可约的 Markov 链我们有如下结论.

命题 3.4 令 $\boldsymbol{P}$ 为不可约、非周期、有限状态 Markov 链的转移概率阵, 则必存在 N, 使得当 $n \geqslant N$ 时, n 步转移概率阵 $\boldsymbol{P}^{(n)}$ 的所有元素都非零.

证 由于 Markov 链是不可约的, 过程的任两个状态 i 和 j 都是可达的, 于是存在 m (与 i, j 有关) 使 $P_{ij}^{(m)} > 0$. 由推论 3.1 及非周期性存在 N, 使得当 $n \geqslant N$ 时有 $P_{ij}^{(m+n \cdot 1)} > 0$. 因状态空间有限, 对全部的状态对 (i, j) 求出 $N(i, j)$. 并取 $N = \max_{(i,j)} (m(i, j) + N(i, j))$, 则显然对所有状态 i, j, 当 $n > N$ 时有 $P_{ij}^{(n)} > 0$.

这种存在 n 使 P^n 的元素全部非零的 k 个状态 (记为 $0, 1, \cdots, k-1$) 的 Markov 链称为是正则的. 而最重要的事实是对一个正则的有限状态 Markov 链, 极限 $\lim_{n \to \infty} P_{ij}^{(n)}$ 总是存在的, 记为 π_j (注意极限与初始状态 i 无关). 并且这些极限构成一概率分布, 即对 $\pi = (\pi_0, \pi_1, \cdots, \pi_{k-1})$, 满足 $\pi_j > 0$, $j = 0, 1, \cdots, k-1$, 及 $\sum_j \pi_j = 1$. 这些是下一节极限定理和平稳分布所要讨论的内容.

3.2.2 常返 (recurrent) 与瞬过 (transient)

为了以下讨论的方便, 我们引入一个重要的概率 $f_{ij}^{(n)}$, 它表示从 i 出发在 n 步转移时首次到达 j 的概率. 用式子表达即是

$$\begin{aligned} f_{ij}^{(0)} &= 0, \\ f_{ij}^{(n)} &= P\{X_n = j, X_k \neq j, k = 1, \cdots, n-1 | X_0 = i\}, \end{aligned}$$

定义 3.5 如果 $f_{ii}=1$, 我们称状态 i 是常返 (recurrent) 的. 一个非常返状态就称为瞬过 (transient) 的.

从定义知道如果 i 是常返的状态, 那么从 i 出发经过有限步转移后最终又回到 i 的概率为 1. 那么如何判断一个状态 i 是否是常返的呢? 我们给出用 n 步转移概率 $P_{ii}^{(n)}$ 表示的一个判别准则.

定理 3.2 状态 i 常返的充分必要条件是

$$\sum_{n=1}^{\infty} P_{ii}^{(n)}=\infty. \tag{3.8}$$

当然与此等价地有, 状态 i 是瞬过的当且仅当

$$\sum_{n=1}^{\infty} P_{ii}^{(n)}<\infty. \tag{3.9}$$

推论 3.2 如果 i 是常返的, 且 $i\leftrightarrow j$, 则 j 也是常返的.

对常返状态 i, 我们定义 T_i 为首次返回状态 i 的时刻, 称作常返时. 记 $\mu_i=ET_i$, 则有

$$\mu_i=\sum_{n=1}^{\infty} n f_{ii}^{(n)}. \tag{3.10}$$

回想 $f_{ii}^{(n)}$ 代表从 i 出发在第 n 步转移时首次回到 i 的概率, 所以 μ_i 是首次返回 i 的期望步数, 也叫状态 i 的平均常返时. 利用 μ_i 可以对常返状态作进一步的分类, 分为零常返和正常返.

定义 3.6 一个常返状态 i 当且仅当 $\mu_i=\infty$ 时称为零常返的, 而当且仅当 $\mu_i<\infty$ 时称为正常返的.

对于只有有限多个状态的 Markov 链, μ_i 总是有限的, 所以只有在有可列无穷多个状态时才可能出现零常返的状态 (参看习题 3 第 15 题). 当状态数目不大时直

在上一节有关 Markov 链状态分类的讨论中已经引进了常返状态 i 的常返时 T_i. 它还可记为

$$T_i=\min\{n\geqslant 1 : X_n=i\},$$

而

$$\begin{aligned}f_{ii}^{(n)}&=P\{X_n=i,X_k\neq i,k=1,\cdots,n-1|X_0=i\}\\&=P\{T_i=n|X_0=i\},\quad n=1,2,\cdots\end{aligned}$$

表示了 T_i 的条件概率分布, $\mu_i=\sum_{n=1}^{\infty} nf_{ii}^{(n)}$ 则为 T_i 在给定 $X_0=i$ 时的条件期望.

在进行更深入的讨论之前先看如下的例子.

长则过程就长期而言处于状态 0 的概率越小. π_0 有两重含义, 它既可以反映在时间长河中过程处于状态 0 的份额或机会, 又同时代表当整个过程处于平衡状态 (极限情形下) 时过程处于状态 0 的机会.

定理 3.3 Markov 链的基本极限定理

(a) 若状态 i 是瞬过的或者是零常返的, 则

$$\lim_{n\to\infty} P_{ii}^{(n)}=0. \tag{3.11}$$

(b) 若状态 i 是周期为 d 的常返状态, 则

$$\lim_{n\to\infty} P_{ii}^{(nd)}=\frac{d}{\mu_i}. \tag{3.12}$$

(c) 当状态 i 是非周期的正常返状态, 则

$$\lim_{n\to\infty} P_{ii}^{(n)}=\frac{1}{\mu_i}. \tag{3.13}$$

一个正常返非周期的状态也称作是遍历的 (ergodic). 定理 3.3 (c) 告诉我们对遍历状态 i 有 $\lim_{n\to\infty} P_{ii}^{(n)}=\frac{1}{\mu_i}$, 其中 μ_i 是常返时的期望.

至于不同状态间的 n 步转移概率 $P_{ji}^{(n)}$ 有如下推论.

推论 3.3 如果状态 i 是遍历的, 则对所有 $i\to j$ 有

$$\lim_{n\to\infty} P_{ji}^{(n)}=\lim_{n\to\infty} P_{ii}^{(n)}=\frac{1}{\mu_i}.$$

证明留给读者作为练习.

直接求 $\boldsymbol{P}^n$ 的每一元素 $P_{ij}^{(n)}$ 的极限并不是一件容易的事. 而求条件期望 ET_i 也并不方便. 它涉及求一系列条件概率 $f_{ii}^{(n)}$. 因而人们寻求更简捷的方法来处理极限分布.

定义 3.7 Markov 链有转移概率阵 $\boldsymbol{P}=(P_{ij})$. 一个概率分布 $\{\pi_i,i\geqslant 0\}$ 如果满足 $\pi_j=\sum_{i=0}^{\infty}\pi_i P_{ij}$, 则称为这一 Markov 链的平稳分布.

定理 3.4 若一个不可约 Markov 链中的所有状态都是遍历的, 则对所有 i,j, 极限 $\lim_{n\to\infty} P_{ij}^{(n)}=\pi_j$ 存在且 $\pi=\{\pi_j,j\geqslant 0\}$ 为平稳分布. 也即

$$\sum_j \pi_j=1,\quad \pi_j>0, \tag{3.15}$$

$$\sum_i \pi_i P_{ij}=\pi_j. \tag{3.16}$$

反之, 若一个不可约 Markov 链只存在一个平稳分布, 即满足 (3.15) 式及 (3.16) 式, 且这个 Markov 链的所有状态都是遍历的. 则该平稳分布就是这一 Markov 链的极限分布, 即对任何 i 有

$$\lim_{n\to\infty} P_{ij}^{(n)}=\pi_j. \tag{3.17}$$

在实际应用中 $\{\pi_j\}$ 有两种解释: 一是作为 $P_{ij}^{(n)}$ 的极限分布, 它告诉我们在过程的长期运行中不论初始状态 i 是什么, 经过一段时期后发现过程处于状态 j 的概率就是 π_j. 另一解释是 π_j 也代表了就长期而言过程访问 j 的次数在总时间中的平均份额或比例. 这可从下面的推理中看出. 设

因而上一代撞击所产生中子的多少会直接影响下一代所能产生中子的数量. 记 X_0 为第 0 代祖先的数目, 一般假定为 1, 并记 X_n 为第 n 代后裔的大小. Markov 链 $\{X_n,n\geqslant 0\}$ 就称作分支过程. 记 Z_i 为第 n 代中第 i 个个体所繁衍的后代数, 则有 $X_{n+1}=\sum_{i=1}^{X_n} Z_i$ 也就是说 X_{n+1} 为 X_n 个随机变量 Z_i 的随机和. 我们可以简化地假定各代个体的繁衍能力是相同的, 而且同一代个体的繁衍是相互独立的. 这时 Z_i 独立同分布, $P(Z_1=k)=p_k$, 且记 $EZ_1=\mu$, $\mathrm{Var}\,Z_1=\sigma^2$. 这时就可以算出 Markov 链 $\{X_n,n\geqslant 0\}$ 的转移概率为

$$P_{ij}=P\{X_{n+1}=j \mid X_n=i\}=P\left\{\sum_{k=1}^{i} Z_k=j\right\}.$$

利用第 1 章的 (1.20) 式我们可求出 X_{n+1} 的均值与方差满足

$$EX_{n+1}=EX_n\cdot EZ_1. \tag{3.19}$$

$$\mathrm{Var}\,X_{n+1}=EX_n\cdot\mathrm{Var}\,Z_1+\mathrm{Var}\,X_n\cdot(EZ_1)^2. \tag{3.20}$$

由 (3.19) 式不断地迭代即有

$$EX_{n+1}=EX_n\cdot\mu=EX_{n-1}\cdot\mu^2=\cdots=EX_1\cdot\mu^n=\mu^{n+1}.$$

这里要注意到第 0 代只有一个个体, 所以 X_1 与 Z_1 同分布. 类似地通过 (3.20) 式可知

$$\begin{aligned}\mathrm{Var}\,X_{n+1}&=\mu^n\cdot\sigma^2+\mathrm{Var}\,X_n\cdot\mu^2\\&=\begin{cases}\sigma^2\mu^n\dfrac{(1-\mu^{n+1})}{1-\mu}, & \mu\neq 1,\\ (n+1)\sigma^2, & \mu=1.\end{cases}\end{aligned}$$

$$\boxed{\mathrm{Res}\,f(z_0)=\frac{1}{(m-1)!}\lim_{z\to z_0}\frac{\mathrm{d}^{m-1}}{\mathrm{d}z^{m-1}}[(z-z_0)^m f(z)]}$$

由 (3.19) 式, (3.20) 式可以看出 μ 的大小对于整个群体和家族的繁衍存亡至关重要. 当 $\mu<1$, 且 n 很大时, EX_n 与 $\mathrm{Var}\,X_n$ 都趋于 0. 由著名的 Chebyshev 不等式马上可以推出对任意给定的 $\varepsilon>0$ 有

$$P\{|X_n-EX_n|>\varepsilon\}\leqslant\frac{\mathrm{Var}\,X_n}{\varepsilon^2},$$

也即当 $n\to\infty$ 时可以证明 $P\{|X_n-0|>\varepsilon\}\to 0$, 或者说 X_n 会以概率趋于 0, 群体终将消亡. 而当 $\mu\geqslant 1$ 群体消亡的概率就是一个较为复杂的问题, 但这是分支过程中的令人感兴趣的量. 从 Markov 链状态分类的观点看群体消亡就是在某个时刻 n 过程 X_n 进入吸收状态, 而群体消亡的概率就是过程最终进入吸收状态的概率. 至于在中子轰击原子裂变的物理问题中, $p_0=P(Z_1=0)=0$, 即粒子不会消失, 而 p_1 则为原子没被击中或不分裂的概率. 这时人们最感兴趣的就不是消亡概率而是粒子数增加的速率了.

我们用生成函数来研究消亡概率. 记过程 X_n 的生成函数为 $\phi_n(s)$, X_1 的生成函数则直接记为 $\phi(s)$, 即 $\phi(s)=\sum_{k=0}^{\infty} p_k s^k$, 其中 $p_k=P(X_1=k)$, $k=0,1,2,\cdots$. 由第 1 章 (1.24) 式随机和 X_{n+1} 的生成函数可以表为 $\phi_{n+1}(s)=\phi_n(\phi(s))$ 或 $\phi_{n+1}(s)=\phi(\phi_n(s))$. 记 π_n 为由单个祖先开始家族将在第 n 代之前消亡的概率, 则 $\pi_n=P\{X_n=0\}=\phi_n(0)$. 我们所感兴趣的整个群体终将消亡的概率是当 $n\to\infty$ 时 π_n 的极限. 当 $p_0=1$, 家族不能发端; 而当 $p_0=0$, 它将永不会消亡. 所以一般假定 $0<p_0<1$. 这时对 $s\in(0,1)$,

$$\phi'(s)=\frac{d}{ds}\left(\sum_{j=0}^{\infty} p_j s^j\right)=\sum_{j=1}^{\infty} p_j\cdot j\cdot s^{j-1}>0,$$

$\phi(s)$ 是 (0,1) 区间上的单调函数, 而 $\phi(0)=p_0>0$, 所以

$$\pi_{n+1}=\phi_{n+1}(0)=\phi_n[\phi(0)]>\phi_n(0)=\pi_n.$$

因此 π_n 随 n 单调上升, 极限 $\lim_{n\to\infty}\pi_n$ 存在. 记为 π, π 就是群体消亡的概率. 由 $\phi_{n+1}=\phi(\phi_n(s))$ 知 $\pi_{n+1}=\phi(\pi_n)$, 再令 $n\to\infty$, 即有 $\pi=\phi(\pi)$. 于是通过求解函数方程 $\phi(s)=s$ 就可以求出 π. 事实上还可以严格证明下面的定理.

定理 3.5 对分支过程 X_n, 若 $p_0>0$, $p_0+p_1<1$, 则有

(a) 群体消亡概率 π 是方程 $\phi(s)=s$ 的最小正解, 其中 $\phi(s)=\sum_{j=0}^{\infty} p_j s^j$, $\{p_j\}$ 是 X_1 与 Z_1 的概率分布.

(b) $\pi=1$ 当且仅当 $\mu\leqslant 1$, 其中 $\mu=EZ_1$.

以下设 T 为具有如下性质的下标集合: 若 $t_1,t_2\in T$, 则 $t_1+t_2\in T$. 通常 T 取如下几种集合之一:

(i) $T=\{0,1,2,\cdots\}$;
(ii) $T=\{0,\pm1,\pm2,\cdots\}$;
(iii) $T=\{t: t\geqslant 0\}$;
(iv) $T=\{t: -\infty<t<\infty\}$.

直观上, 下标集合可以理解为时间 (当然, T 也可以表示空间位置或其他).

定义 4.1 设 $X=\{X(t),t\in T\}$ 为一随机过程, 若对任意正整数 k 及 T 中任意 k 个时刻 $t_1<t_2<\cdots<t_k$, 及 T 中的 h, 有

$$\{X(t_1),\cdots,X(t_k)\}\overset{d}{=}\{X(t_1+h),\cdots,X(t_k+h)\}, \tag{4.1}$$

则随机过程 X 称为严平稳过程. 这里 "d" 表示等式两边 k 维随机向量的分布相同.

注 4.1 如果 $T=\{0,\pm1,\cdots\}$, 我们一般把 X 称为随机序列. 如果 X 还是严平稳的, 则称为严平稳序列, 以下不再一一说明.

设 $X=\{X(t),t\in T\}$ 为严平稳过程, 由定义如果均值函数 $m(t)=EX(t)$ 存在, 则必为常数, 即 $m(t)=m$, $t\in T$. 同样, 如果方差函数存在, 即 $\mathrm{Var}(X(t))=E(X(t)-m)^2$ 也是一个常数, 记为 σ^2. 设 $s,t\in T$ (不妨设 $s<t$), 由平稳性, 其协方差函数

$$E(X(t)-m)(X(s)-m)=E(X(t-s)-m)(X(0)-m).$$

等式右端只依赖于时间差 $t-s$. 若记

$$R(h)=E(X(h)-m)(X(0)-m),$$

则

$$E(X(t+h)-m)(X(t)-m)=R(h),$$

即协方差函数仅与时间差有关, 而与起点无关. 当然, 由定义 $\mathrm{Var}(X(t))=R(0)$. 此外, 易知 $r(\tau)=EX(t)X(t+\tau)$ 与起点 t 无关, 我们分别称 $r(\tau)$ 和

$$\rho(v)=R(v)/\sigma^2=R(v)/R(0)$$

为平稳过程 X 的自相关函数和标准自相关函数. 由概率论中相关系数性质易知 $\rho(0)=1$ 及 $|\rho(v)|\leqslant 1$.

定义 4.2 设 $X=\{X(t),t\in T\}$ 为一实值随机过程, 如果对 $\forall t\in T$, $EX^2(t)<\infty$, $EX(t)=m$ 以及协方差函数 $E(X(t)-m)(X(s)-m)$ 仅与 $t-s$ 有关, 则称 X 为宽平稳过程.

一般来说, 这两个过程是互不包含的, 即严平稳由于不一定有二阶矩而不必是宽平稳; 反之, 宽平稳由于其有限维联合分布可能不满足 (4.1) 式而不一定是严平稳. 但只要过程的二阶矩存在, 则严平稳过程一定是宽平稳过程.

一个自然的问题是是否存在一类过程, 使这两个平稳过程的定义等价呢? 答案是肯定的, 为此我们引入如下定义.

定义 4.3 设 $G=\{G(t),-\infty<t<\infty\}$ 为一随机过程, 如果对任一正整数 k 以及 k 个时刻 $t_1\leqslant t_2\leqslant\cdots\leqslant t_k$, $(G(t_1),G(t_2),\cdots,G(t_k))$ 的联合分布为 k 维正态分布, 则称 G 为高斯(Gauss) 过程.

我们知道, k 维正态分布完全由协方差矩阵和均值向量所唯一确定, 而这些量仅与它们的二阶矩有关, 所以对 Gauss 过程而言, 两个平稳的定义是等价的.

以下主要研究宽平稳过程. 为方便起见, 我们就称宽平稳过程为平稳过程.

定义 4.5 设 $X=\{X(t),-\infty<t<\infty\}$ 为一平稳过程(或序列), 若

$$\overline{X}=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} X(t)dt\overset{L_2}{=}m \tag{4.6}$$

或

$$\overline{X}=\lim_{N\to\infty}\frac{1}{2N+1}\sum_{k=-N}^{N} X(k)\overset{L_2}{=}m, \tag{4.7}$$

则称 X 的均值有遍历性. 如果

$$\hat{R}(\tau)=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}(X(t)-m)(X(t+\tau)-m)dt\overset{L_2}{=}R(\tau) \tag{4.8}$$

或

$$\hat{R}(\tau)=\lim_{N\to\infty}\frac{1}{2N+1}\sum_{k=1}^{n}(X(k+\tau)-\hat{m}_n)(X(k)-\hat{m}_n)\overset{L_2}{=}R(\tau), \tag{4.9}$$

则称 X 的协方差函数有遍历性. 若随机过程(或序列)的均值和协方差函数都有遍历性, 则称此随机过程有遍历性.

注意, 我们这里的极限是定义 1.8 中的均方极限, 以 (4.6) 式为例, 即

$$E\left(\frac{1}{2T}\int_{-T}^{T} X(t)dt-m\right)^2\to 0\quad (T\to\infty).$$

遍历性又称各态历经性. 直观上可以这样理解: 考虑只有有限个状态的平稳序列

先考虑平稳过程均值的遍历性问题.

定理 4.1(均值遍历性定理)

(i) 设 $X=\{X_n,n=0,\pm1,\cdots\}$ 为平稳序列, 其协方差函数为 $R(\tau)$, 则 X 有遍历性的充分必要条件是

$$\lim_{N\to\infty}\frac{1}{N}\sum_{\tau=0}^{N-1} R(\tau)=0.$$

(ii) 若 $X=\{X(t),-\infty<t<\infty\}$ 为平稳过程, 则 X 有遍历性的充分必要条件是

$$\lim_{T\to\infty}\frac{1}{T}\int_0^{2T}(1-\frac{\tau}{2T})R(\tau)d\tau=0.$$

由此定理, 可以推出一些判断平稳过程均值有遍历性的充分条件.

推论 4.1 若 $\int_{-\infty}^{\infty}|R(\tau)|<\infty$, 则均值遍历性成立.

这是由于当 $0\leqslant\tau\leqslant 2T$ 时, $|(1-\frac{\tau}{2T})R(\tau)|\leqslant|R(\tau)|$,

$$\frac{1}{T}\left|\int_0^{2T}(1-\tau/2T)R(\tau)d\tau\right|\leqslant\frac{1}{T}\int_0^{2T}|R(\tau)|d\tau\leqslant\frac{1}{T}\int_0^{\infty}|R(\tau)|d\tau\to 0.$$

推论 4.2 对平稳序列而言, 若 $R(\tau)\to 0(\tau\to\infty)$, 则均值遍历性成立.

定理 4.2 (协方差函数遍历性定理) 设 $X=\{X(t),-\infty<t<\infty\}$ 为平稳过程, $Y_\tau=\{Y_\tau(t),-\infty<t<\infty\}$ 其中 Y_τ 由上面所定义, 则对给定的 τ, X 的协方差函数 $R(\tau)$ 有遍历性的充分必要条件是

$$\lim_{T\to\infty}\frac{1}{T}\int_0^{2T}\left(1-\frac{\tau_1}{2T}\right)(B(\tau_1)-R^2(\tau))d\tau_1=0,$$

其中

$$B(\tau_1)=EX(t+\tau+\tau_1)X(t+\tau_1)X(t+\tau)X(t).$$

对于定义在 $[0,\infty)$ 上的平稳过程, 只要把遍历性理解为 (4.10),(4.11) 等式, 则定理 4.1 和定理 4.2 仍成立.

关于协方差函数的遍历性, 由于牵涉到过程的四阶矩, 一般很难验证. 但对于 Gauss 过程来说, 问题要简单得多, 比如我们有如下的结果.

定理 4.3 设 $X=\{X_n,n=0,\pm1,\cdots\}$ 是均值为 0 的 Gauss 平稳过程, 如果

$$\lim_{T\to\infty}\frac{1}{N}\sum_{k=0}^{N-1} R^2(k)=0,$$

则 Gauss 过程的协方差函数有遍历性.

4.3.1 协方差函数

对平稳过程 X 的协方差函数 $R(\tau)$, 容易由定义得到如下性质:

1. 对称性, 即 $R(-\tau)=R(\tau)$.
2. 有界性, 即 $|R(\tau)|\leqslant R(0)$.
3. 非负定性. 即对任意的时刻 t_n 及实数 a_n, $n=1,2,\cdots,N$, 有

$$\sum_{n=1}^{N}\sum_{m=1}^{N} a_n a_m R(t_n-t_m)\geqslant 0.$$

这是很重要的一条性质, 证明见 (1.2) 式.

关于平稳过程导数, 我们作如下定义: 如果存在 $Y(t),t\in T$ 使

$$\lim_{h\to 0}E\left|\frac{X(t+h)-X(t)}{h}-Y(t)\right|^2=0,$$

则称 $Y=\{Y(t)\}$ 为过程 X 在 t 点的均方导数, 简称导数, 并记为 $X'(t)$ 或 $\frac{dX(t)}{dt}$. 可以验证, 均方导数存在的充分必要条件是

$$\lim_{h\to 0,k\to 0}\frac{R(0)-R(h)-R(k)+R(h-k)}{hk}$$

存在有限. 由此我们可以推出如下的有趣性质: 只要下面所涉及的导数都存在, 则性质成立.

4. 平稳过程 n 阶导数的协方差函数为

$$\mathrm{Cov}(X^n(t),X^n(t+\tau))=(-1)^n R^{(2n)}(\tau). \tag{4.16}$$

由 $S(\omega)$ 的定义可知

$$\bar{S}(\omega)=S(\omega)\geqslant 0,\qquad S(-\omega)=S(\omega).$$

这是因为 $|F(\omega,T)|^2=F(\omega,T)F(-\omega,T)$ 为实的, 非负偶函数. 其次, 由 (4.27) 定义的平均功率谱密度 $S(\omega)$ 和协方差函数 $R(\tau)$(假定平稳过程的均值为零) 是一对 Fourier 变换. 一般由 (4.27) 定义的平均功率谱密度 $S(\omega)$ 和自相关函数 $r(\tau)$ 也是一对 Fourier 变换. 更具体地, 我们有如下定理.

定理 4.4 (Wiener-Khintchine公式) 假定 $EX(t)=0$, 且 $\int|R(\tau)|d\tau<\infty$, 则

$$S(\omega)=\int R(\tau)\mathrm{e}^{-j\omega\tau}d\tau, \tag{4.33}$$

$$R(\tau)=\frac{1}{2\pi}\int S(\omega)\mathrm{e}^{j\omega\tau}d\omega. \tag{4.34}$$

$$S(\omega)=\sum_{\tau=-\infty}^{\infty}\mathrm{e}^{-j\omega\tau}R(\tau), \tag{4.37}$$

$$R(\tau)=\frac{1}{2\pi}\int_{-\pi}^{\pi}S(\omega)\cos\omega\tau d\omega. \tag{4.38}$$

Fourier 分析中知道, 如果允许谱密度和协方差函数中含有 δ 函数, 则在广义 Fourier 变换下, 仍成立 Wiener-Khintchine 公式. 这主要是利用 δ 函数的如下基本性质: 对任一连续函数 $f(\tau)$,

$$\int\delta(\tau-\tau_0)f(\tau)d\tau=f(\tau_0).$$

由此可得

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{j\omega\tau}d\omega=\delta(\tau), \tag{4.39}$$

$$\int\delta(\tau)\mathrm{e}^{-j\omega\tau}d\tau=1. \tag{4.40}$$

由 (4.39) 式, 当协方差函数为常数 1 时, 对应的谱密度函数为 $2\pi\delta(\omega)$. 反之, 当谱密度为 1 时, 由 (4.39) 知协方差函数为 $\delta(\tau)$. 再由

$$\cos\omega\tau=\frac{1}{2}(\mathrm{e}^{j\omega\tau}+\mathrm{e}^{-j\omega\tau}),\qquad \sin\omega\tau=\frac{1}{2j}(\mathrm{e}^{j\omega\tau}-\mathrm{e}^{-j\omega\tau}).$$

可以得到当谱密度为 $a\cos\omega\tau_0$ 时, 其对应的协方差函数为

$$\begin{aligned}R(\tau)&=\frac{1}{2\pi}\int a\cos\omega\tau_0\mathrm{e}^{j\omega\tau}d\omega\\&=\frac{a}{4\pi}\left[\int\mathrm{e}^{j\omega(\tau+\tau_0)}d\omega+\int\mathrm{e}^{j\omega(\tau-\tau_0)}d\omega\right]\\&=\frac{a}{2}(\delta(\tau+\tau_0)+\delta(\tau-\tau_0)).\end{aligned}$$

分布	$P(X=x)/f(x)$	矩母	EX	Var(X)
$B(n,p)$	$C_n^x p^x(1-p)^{n-x}$	$\left(pe^t+(1-p)\right)^n$	np	$np(1-p)$
$poisson$	$\frac{e^{-\lambda}\lambda^x}{x!}$	$exp\{\lambda(e^t-1)\}$	λ	λ
几何	$p(1-p)^{x-1}$	$\frac{pe^t}{1-(1-p)e^t}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
负二项	$\mathrm{C}_{\mathrm{x}-1}^{r-1}p^r(1-p)^{x-r}$	$\left(\frac{pe^t}{1-(1-p)e^t}\right)^r$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$
$U(a,b)$	$\frac{1}{b-a}$	$\frac{e^{ta}-e^{tb}}{t(a-b)}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
指数	$\lambda e^{-\lambda x}$	$\frac{\lambda}{\lambda-t}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$\Gamma(n,\lambda)$	$\frac{\lambda e^{-\lambda x}(\lambda x)^{n-1}}{(n-1)!}$	$\left(\frac{\lambda}{\lambda-t}\right)^n$	$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$
$N(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$\exp\left\{\mu t+\frac{\sigma^2t^2}{2}\right\}$	μ	σ^2

made by Lyncien