泊松过程

- 问题：例如：Internet WWW 服务器的访问次数与文件吞吐量
  - 模型、分布特征、统计特性？

- 讨论内容：
  - 计数过程
  - 泊松过程的基本概念
  - 泊松过程的统计特征
  - 泊松分布的几个问题
  - 非齐次泊松过程
  - 泊松过程的和、差
  - 复合泊松过程
  - 泊松过程举例

1 基本概念

1.1 独立增量过程和正交增量过程

定义：独立增量过程，

随机过程 $\{\xi(t), t \in T\}$，在 T 上任选 $t_1 < t_2 < \cdots < t_n$ n 个点，随机过程的增量

$\xi(t_2)-\xi(t_1)$，$\xi(t_4)-\xi(t_3)$，$\cdots, \xi(t_n)-\xi(t_{n-1})$ 是相互统计独立的，则称这类随机过程是独立增量过程。

定理：

独立增量过程一定是马尔可夫过程。

定义：正交增量过程，

设有二阶矩过程 $\{\xi(t), t \in T\}$，并设 $t_1 \le t_2 \le t_3 \le t_4$，$t_1, t_2, t_3, t_4 \in T$ 有

$E\{[\xi(t_2)-\xi(t_1)][\xi(t_4)-\xi(t_3)]^*\} = 0$，则称该过程为正交增量过程。

定理：

独立增量过程，若满足 $E\{\xi(t)\} = 0$，$E\{|\xi(t)|^2\} \le \infty$，则该过程也是一个正交增量过程。

定义：齐次独立增量过程，

如果独立增量过程的增量仅与时间差有关，而与时间本身无关，则称它为齐次独立增量过程。

1.2 计数过程

定义，在（0, t）内出现事件 A 的总数所组成的过程{N(t), t>0}称为计数过程。

计数过程{N(t), t>0}满足下列条件：

1. N(t)>0；
2. N(t)是一个正整数；
3. 如果有两个时刻 s, t，且 s<t，则 N(s)< N(t)；
4. 如果有两个时刻 s, t，且 s<t，则 N(t)-N(s)代表时间间隔（s, t）内事件 A 出现的次数。

独立增量计数过程

在不相交的时间间隔内出现事件 A 的次数是相互统计独立的。

平稳（齐次）增量计数过程

在时间间隔（t, t+s）内出现事件 A 的次数[N(t+s)-N(t)]仅与 s 有关而与 t 无关。

1.3 泊松过程的基本概念

定义，设有一个计数过程{N(t), t>0}满足下列假设，称为泊松过程，

1. 在 t=0 时，N(t)=0；
2. 该过程是独立增量计数过程；
3. 该过程是平稳增量计数过程；
4. 在（t, t+Δt）内出现一个事件的概率为 λΔt + 0(Δt)，λ为一常数，在（t, t+Δt）内出现两个或两个以上事件的概率为 0(Δt)，即 P{ N(t+Δt) - N(t)>1}=0(Δt)

2 泊松过程的分布特征

令 $P_n(t)=P\left[N(t)-N(0)=n\right]$ 表示在间隔 t 的时间内发生 n 次事件的概率。

对于泊松过程，在(0,t+Δt)内不出现事件，等价于在(0,t)间隔内不出现事件以及在(t,t+Δt)间隔内不出现事件；根据泊松过程的特征 4，在(t,t+Δt)内不出现事件的概率为：1-λΔt + 0(Δt)；根据泊松过程的独立增量特性，在(0,t) 与(t,t+Δt)间隔内出现事件次数是相互统计独立的，因而：

$$P_0(t+\Delta t)=P_0(t)P_0(\,\Delta t)$$

即：$P_0(t+\Delta t)=P_0(t)(1-\lambda\,\Delta t)+0(\Delta t)$

同理，在(0,t+Δt)内出现事件 n 次事件，等价于在(0,t)间隔内出现 n 次事件、在(t,t+Δt)

间隔内不出现事件，或者在(0,t)间隔内出现 n-1 次事件、在(t,t+Δt)间隔内出现 1 次事件，或者在(t,t+Δt)间隔内不出现事件，或者在(0,t)间隔内出现 k<(n-1)次事件、在(t,t+Δt)间隔内出现 2 次或两次以上的事件；利用泊松过程的特征 4、泊松过陈的独立增量特性、以及全概率公式，得：

$$P_n(t+\Delta t)=P_n(t)P_0(\Delta t)+P_{n-1}(t)P_1(\Delta t)+\sum_{k=0}^{n-2}P_k(t)P_{n-k}(\Delta t),\quad n\ge 1$$

即：$P_n(t+\Delta t)=P_n(t)(1-\lambda\Delta t)+P_{n-1}(t)\lambda\Delta t+0(\Delta t),\quad n\ge 1$

整理上述两式，可得泊松过程递推微分方程：

$$\frac{d}{dt}P_0(t)=-\lambda P_0(t)$$
$$\frac{d}{dt}P_n(t)=-\lambda P_n(t)+\lambda\ P_{n-1}(t),\quad n\ge 1$$

泊松过程递推微分方程的解，

$$P_0(t)=e^{-\lambda t}$$

$$P_1(t)=\lambda\ t\cdot e^{-\lambda t}$$

$$P_n(t)=\frac{(\lambda\ \ t)^n}{n!}e^{-\lambda t}$$

因此：泊松过程的概率分布律为：

$$P\{N(t_0+t)-N(t_0)=n\}=P_n(t)=\frac{(\lambda\ \ t)^n}{n!}e^{-\lambda t}$$

泊松分布的母函数为：

$$\Psi(z)=\sum_{n=0}^{\infty}P_n z^n=\sum_{k=0}^{\infty}\frac{(\lambda t)^n}{n!}e^{-\lambda t}z^n=e^{-\lambda t(1-z)}$$

对于给定的时刻 s、t，且 s<t，$t=s+\Delta t$，以及相应的 N(s)、N(t)，转移概率分布为，

$$\begin{aligned}&P\{N(t)=n+k/N(s)=k\}\\&\quad=P\{N(t)-N(s)=n/N(s)=k\}\\&\quad=P\{N(s+\Delta t)-N(s)=n\}\\&\quad=P_n(\Delta t)=\frac{(\lambda\cdot\Delta t)^n}{n!}e^{-\lambda\cdot\Delta t}\end{aligned}$$

3. 泊松分布的统计特征

均值：

$$E\{N(t)\}=\sum_{n=0}^{\infty}n\cdot P\{N(t)=n\}$$

$$E\{N(t)\}=\left[\frac{d}{dz}\Phi(z)\right]_{z=1}$$

$$E\{N(t)\}=\sum_{n=0}^{\infty}nP\{N(t)=n\}=\sum_{n=0}^{\infty}nP\{N(t)=n\}\cdot z^{n}\Big|_{z=1}$$

$$=\frac{d}{dz}\left[\sum_{n=0}^{\infty}P\{N(t)=n\}\cdot z^{n}\right]_{z=1}=\left[\frac{d}{dz}\Phi(z)\right]_{z=1}$$

$$=\left[\frac{d}{dz}e^{-\lambda t(1-z)}\right]_{z=1}=\lambda t$$

$$E\{N(t)\}=\lambda t$$

方差：

$$E\left\{[N(t)]^{2}\right\}=\sum_{k=0}^{\infty}k^{2}P\{N(t)=k\}$$

$$=\sum_{k=0}^{\infty}k^{2}P(N(t)=k)\cdot z^{k}\Big|_{z=1}$$

$$=z\frac{d}{dz}z\frac{d}{dz}\left[\sum_{k=0}^{\infty}P(N(t)=k)\cdot z^{k}\right]_{z=1}$$

$$=\left[z\frac{d}{dz}z\frac{d}{dz}\Phi(z)\right]_{z=1}=\left[z\frac{d}{dz}z\frac{d}{dz}e^{-\lambda t(1-z)}\right]_{z=1}$$

$$=z\frac{d}{dz}\left[z\lambda te^{-\lambda t(1-z)}\right]_{z=1}=\left[z\lambda te^{-\lambda t(1-z)}+z^{2}(\lambda t)^{2}e^{-\lambda t(1-z)}\right]_{z=1}$$

$$=\lambda t+(\lambda t)^{2}$$

$$D\{N(t)\}=E\left\{[N(t)]^{2}\right\}-\left[E\{N(t)\}\right]^{2}=\lambda t$$

自相关函数

$$R(t_1,t_2)=E\{N(t_1)N(t_2)\}$$

假设$t_1<t_2$，有

$$
\begin{aligned}
E\{N(t_1)N(t_2)\} &= E\left\{N(t_1)N(t_1)+N(t_1)\left[N(t_2)-N(t_1)\right]\right\} \\
&= E\{N(t_1)N(t_1)\}+E\left\{N(t_1)\right\}\cdot E\{\left[N(t_2)-N(t_1)\right]\} \\
&= \lambda t_1+\left(\lambda t_1\right)^2+\lambda t_1\cdot\lambda\left(t_2-t_1\right) \\
&= \lambda t_1+\lambda t_1\cdot\lambda t_2
\end{aligned}
$$

若$t_1 > t_2$，则有：

$$E\{N(t_1)N(t_2)\}=\lambda t_2+\lambda t_1\cdot\lambda t_2$$

总结起来有

$$R(t_1,t_2)=\begin{cases}\lambda t_2+\lambda^2 t_1 t_2 & t_1\ge t_2\\ \lambda t_1+\lambda^2 t_1 t_2 & t_1\le t_2\end{cases}$$

或：$R(t_1,t_2)=\lambda\cdot\min\left[t_1,t_2\right]+\lambda t_1\cdot\lambda t_2$

自协方差函数

$$
\begin{aligned}
C(t_1,t_2) &= E\left\{\left[N(t_1)-\right]\mu(t_1)\left[N(t_2)-\mu(t_2)\right]\right\} \\
&= E\left\{N(t_1)N(t_2)\right\}-\mu(t_1)\mu(t_2) \\
&= \lambda\min(t_1,t_2) \\
&= \lambda t_1 U(t_2-t_1)+\lambda t_2 U(t_1-t_2)
\end{aligned}
$$

4. 泊松分布的几个问题

4.1 第一个事件到达时间的概率密度

泊松过程{N(t), t>0}的第一个事件到达时间 T，T 落在$(\tau,d\tau)$内的概率：

泊松过程第 1 个事件在时刻$(\tau,d\tau)$发生，而在$(0,\tau)$时间内不发生事件的概率：

$$
\begin{aligned}
&P\{\tau<T<(\tau+d\tau)\}=P\{[N(\tau+d\tau)-N(\tau)]=1\}\cdot P\{[N(\tau)-N(0)]=0\} \\
&\quad=\lambda\cdot d\tau\cdot e^{-\lambda\tau}
\end{aligned}
$$

故，泊松过程{N(t), t>0}的第一个事件到达时间 T 的概率密度

$$f(\tau)=\lambda\cdot e^{-\lambda\tau}$$

4.2 第 n 个事件到达时间 t 的概率密度

泊松过程{N(t), t>0}的第 n 个事件到达时间T_n，T_n落在$(\tau,d\tau)$内的概率：

泊松过程第 n 个事件在时刻$(\tau, d\tau)$发生，而在$(0,\tau)$时间内发生(n-1)事件的概率：

$$P\{\tau < T_n < (\tau + d\tau)\} = P\{[N(\tau + d\tau) - N(\tau)] = 1\} \cdot P\{[N(\tau) - N(0)] = n-1\}$$
$$= \lambda \cdot d\tau \cdot \frac{(\lambda\tau)^{n-1}}{(n-1)!} e^{-\lambda\tau}$$

泊松过程{N(t), t>0}的第 n 个事件到达时间T_n的概率密度分布：

$$f(\tau) = \lambda \cdot \frac{(\lambda\tau)^{n-1}}{(n-1)!} e^{-\lambda\tau}$$

4.3 时间间隔（0，t）内发生一个事件，事件发生时间的概率密度

泊松过程{N(t), t>0}在（0,t）内有一个事件出现，事件到达时间 S 的概率分布

$$P\{S < s / N(t) = 1\} = P\{N(s) = 1 / N(t) = 1\}, \qquad (s < t)$$
$$= \frac{P\{N(s) = 1\} \cdot P\{N(t) = N(s)\}}{P\{N(t) = 1\}}$$
$$= \frac{\lambda s e^{-\lambda s} \cdot e^{-\lambda(t-s)}}{\lambda t e^{-\lambda t}} = \frac{s}{t}$$

则　S 的概率密度是 1/t。

结论：如果已知在(0,t)内发生 n 次事件，则 n 次事件的发生时间是 n 个独立同分布的随机变量的顺序序列，每一随机变量均匀分布于(0,t)内。

4.4 时间间隔$(0,t_2)$内发生 n 个事件时，$(0,t_1 < t_2)$内发生 k 次事件的概率

考虑条件概率，$P\left[N(t_1) = k \mid N(t_2) = n\right]$，$t_1 \le t_2, k = 0,1,\cdots,n$

$$P\left\{N(t_1) = k \mid N(t_2) = n\right\} = \frac{P\left\{N(t_1) = k, N(t_2) = n\right\}}{P\left\{N(t_2) = n\right\}}$$
$$= \frac{P\left\{N(t_1) = k, N(t_2 - t_1) = n - k\right\}}{P\left\{N(t_2) = n\right\}}$$
$$= \frac{e^{-\lambda t_1}\left(\lambda t_1\right)^k}{k!} \frac{e^{-\lambda(t_2 - t_1)}\left(\lambda\left(t_2 - t_1\right)\right)^{n-k}}{(n-k)!} \frac{n!}{e^{-\lambda t_2}\left(\lambda t_2\right)^n}$$
$$= \binom{n}{k}\left(\frac{t_1}{t_2}\right)^k \left(1 - \frac{t_1}{t_2}\right)^{n-k}$$

结论：条件概率服从二项式分布

4.5 第一个过程的第一个事件先于第二个过程的第一个事件的概率

有两个相互独立的泊松过程$\{N_1(t), t>0)$及$\{N_2(t), t>0)$，它们在单位时间内出现事件的平均数分别是λ_1, λ_2，设 x,y,分别是两个过程出现第一次事件的时刻，求第一个过程的第一个事件先于第二个过程的第一个事件的概率，即 P{ x<y},

解：首先考虑第一个过程第 1 个事件在时刻(x+dx)发生，而第二个过程在（0,x）时间内不发生任何事件的概率密度，再考虑 x 从 0 到∞的积分。

$$\int_0^\infty dx\lambda_1 \cdot e^{-\lambda_1 x} \cdot e^{-\lambda_2 x}$$
$$= \frac{\lambda_1 \cdot}{(\lambda_1 + \lambda_2)} \cdot \int_0^\infty dx \cdot (\lambda_1 + \lambda_2) e^{-(\lambda_1 + \lambda_2)x}$$
$$= \frac{\lambda_1 \cdot}{(\lambda_1 + \lambda_2)} \cdot$$

4.6 第一个过程的第 k 个事件先于第二个过程的第一个事件的概率

有两个相互独立的泊松过程{N1(t), t>0}及{N2(t), t>0}，它们在单位时间内出现事件的平均数分别是λ1，λ2，设 t1k,t21,分别是两个过程出现第一次事件的时刻，求第一个过程的第 k 个事件先于第二个过程的第一个事件的概率，即 Pr{ t1k < t21}。

首先考虑第一个过程第 k 个事件在时刻(x+dx)发生，而第二个过程在（0,x）时间内不发生任何事件的概率密度，再考虑 x 从 0 到∞的积分。

$$\int_0^\infty dx\lambda_1 \cdot \frac{(\lambda_1 x)^{k-1}}{(k-1)!} e^{-\lambda_1 x} \cdot e^{-\lambda_2 x}$$
$$= \left(\frac{\lambda_1 \cdot}{\lambda_1 + \lambda_2}\right)^k (\lambda_1 + \lambda_2) \cdot \int_0^\infty dx \cdot \frac{[(\lambda_1 + \lambda_2)x]^{k-1}}{(k-1)!} e^{-(\lambda_1 + \lambda_2)x}$$
$$= \left(\frac{\lambda_1 \cdot}{\lambda_1 + \lambda_2}\right)^k$$

5 非齐次泊松过程

定义，计数过程{N(t), t>0}满足下列条件，称它为非齐次泊松过程，

1. 在 t=0 时，N(t)=0；
2. 该过程{N1(t), t>0}是独立增量计数过程；
3. P{ N(t+Δt) - N(t)>1}=0(Δt)；

4. P{ N(t+Δt) - N(t)=1}=λ(t) Δt + 0(Δt)。

非齐次泊松过程的概率分布：

$$P\{N(t_0+t)-N(t_0)=n\}=\frac{[m(t_0+t)-m(t_0)]^n}{n!}e^{-[m(t_0+t)-m(t_0)]}$$

其中，

$$m(t)=\int_0^t \lambda(s)\,ds$$

6 泊松过程的和、差

泊松过程的和

如果 $N_1(t)$，$N_2(t)$ 是参数分别为 λ_1，λ_2 的独立泊松过程，则他们的和 $N_1(t)+N_2(t)$ 是参数为 $(\lambda_1+\lambda_2)$ 的泊松过程

泊松过程的差

如果 $N_1(t)$，$N_2(t)$ 是参数分别为 λ_1，λ_2 的独立泊松过程，则他们的差 $y(t)=N_1(t)-N_2(t)$ 取值为 n 的概率是，

$$\begin{aligned}P\{y(t)=n\}&=\sum_{k=0}^{\infty}P\{N_1(t)=n+k\}P\{N_2(t)=k\}\\&=\sum_{k=0}^{\infty}e^{-\lambda_1 t}\frac{(\lambda_1 t)^{n+k}}{(n+k)!}e^{-\lambda_2 t}\frac{(\lambda_2 t)^k}{k!}\\&=e^{-(\lambda_1+\lambda_2)t}\left(\frac{\lambda_1}{\lambda_2}\right)^{n/2}\sum_{k=0}^{\infty}\frac{\left(\sqrt{\lambda_1\lambda_2}t\right)^{n+2k}}{k!(n+k)!}\\&=e^{-(\lambda_1+\lambda_2)t}\left(\frac{\lambda_1}{\lambda_2}\right)^{n/2}I_{|n|}\left(2\sqrt{\lambda_1\lambda_2}t\right)\end{aligned}$$

其中 $I_{|n|}\left(2\sqrt{\lambda_1\lambda_2}t\right)$ 是 n 阶修正贝塞尔函数。泊松过程的差的均值和方差是，

$$E\{y(t)\}=(\lambda_1-\lambda_2)t \qquad Var\{y(t)\}=(\lambda_1+\lambda_2)t。$$

两个独立泊松过程的差不是泊松过程。

7 复合泊松过程

定义：

设有泊松过程{N(t), t>0}和一族独立同分布的随机变量{Y_n}，n=1，2，3，且{N(t), t>0}和{Y_n}也是相互统计独立的。随机过程 $x(t)=\sum_{n=0}^{N(t)} Y_n$ 称为复合泊松过程。若 Y_n=1，则复合泊松过程就是普通的泊松过程。

在任意时刻复合泊松过程 $x(t)$ 的矩生成函数是

$$
\begin{aligned}
\Phi_x(z) &= E\left[z^{x(t)}\right]=E\left\{E\left[z^{x(t)} \mid N(t)=k\right]\right\} \\
&= E\left\{E\left[z^{\sum_{i=1}^{k} y_i} \mid N(t)=k\right]\right\}=\sum_{k=0}^{\infty} E\left[z^{\sum_{i=1}^{k} y_i} \mid N(t)=k\right] P\left(N(t)=k\right) \\
&= \sum_{k=0}^{\infty} E\left[\prod_{i=1}^{k} z^{y_i}\right] P\left(N(t)=k\right)=\sum_{k=0}^{\infty}\left[\prod_{i=1}^{k} E\left[z^{y_i}\right]\right] P\left(N(t)=k\right) \\
&= \sum_{k=0}^{\infty}\left[\prod_{i=1}^{k} \phi_y(z)\right] \cdot e^{-\lambda t} \frac{(\lambda t)^k}{k!}=\sum_{k=0}^{\infty}\left[\phi_y(z)\right]^k e^{-\lambda t} \frac{(\lambda t)^k}{k!} \\
&= e^{-\lambda t\left[1-\phi_y(z)\right]}
\end{aligned}
$$

复合泊松过程的均值：

$$
\begin{aligned}
E\{x(t)\} &= \Phi'_x(z)|_{z=1} \\
&= \left[e^{-\lambda t\left[1-\phi_y(z)\right]}\right]' |_{z=1} \\
&= \lambda t \phi_y'(z)|_{z=1} \\
&= \lambda t E\{Y\}
\end{aligned}
$$

复合泊松过程的方差：

$$
E\{x^2(t)\}=\Phi''_x(z)|_{z=1}+\Phi^1_x(z)|_{z=1}
$$

$$
\begin{aligned}
D\{x(t)\} &= E\{x^2(t)\}-[E\{x(t)\}]^2 \\
&= \Phi_x''(z)|_{z=1}+\Phi_x^1(z)|_{z=1}-[\Phi_x^1(z)|_{z=1}]^2 \\
&= \lambda t[\phi_y''(z)+\phi_y'(z)]|_{z=1} \\
&= \lambda t E\{Y^2\}
\end{aligned}
$$

复合泊松过程举例。

8 泊松过程举例

例 1 非齐次泊松过程，

一酒店，每日 8 点开始营业，从早 8:00 到 11:00 平均顾客到达率线性增加，在 8:00 顾客平均到达率为 5 人/小时，11:00 到达率最高峰 20 人/小时，从上午 11:00 到下午 1:00 平均顾客到达率不变，为 20 人/小时。从下午 1:00 到 5:00 顾客到达率线性下降，到下午 5:00 顾客到达率为 12 人/小时。假设不相交叠的时间间隔内到达商店的顾客数是相互统计独立的 ，问在上午 8:30—9:30 间无顾客到达商店的概率是多少？在这段时间内到达商店顾客数学期望是多少？

解：

到达率函数：

进行时间变换，8 点到下午 5 点对应 0-9

$$\lambda(t)=\begin{cases}5+5t, & 0\le t\le 3\\ 20, & 3\le t\le 5\\ 20-2(t-5), & 5\le t\le 9\end{cases}$$

该过程是一个非齐次泊松过程。

8 点半到 9 点半

$$m(1.5)-m(0.5)=\int_{0.5}^{1.5}\lambda(t)dt$$

$$=\int_{0.5}^{1.5}(5+5t)dt$$

$$=10$$

8 点半到 9 点半

无顾客到达的概率是 $e^{m(1.5)-m(0.5)}$

平均顾客到达数为 $m(1.5)-m(0.5)=10$

例 2 复合泊松过程，

设移民到某地区的户数是一泊松过程，平均每周有两户定居，即 λ =2。如果每户的人口是一个随机变量，一户 4 人的概率是 1/6，一户 3 人的概率是 1/3，一户 2 人的的概率是 1/3，一户 1 人的概率是 1/6。并设每户的人口数是相互独立的随机变量。求在五周内移民到该地区的人口数的数学期望及其方差。

解：

设Y_n代表第 n 户的人口数，$X(t)$代表总的人口数

$$X(t)=\sum_{n=1}^{N(t)}Y_n$$

考虑到

$$E\{Y_n\}=4\times\frac{1}{6}+3\times\frac{1}{3}+2\times\frac{1}{3}+1\times\frac{1}{6}=\frac{5}{2}$$

$$E\{{Y_n}^2\}=4^2\times\frac{1}{6}+3^2\times\frac{1}{3}+2^2\times\frac{1}{3}+1^2\times\frac{1}{6}=\frac{43}{6}$$

根据前面利用矩生成函数得到的特征值得表示：

X(t)的均值为：

$$E\{x(t)\}=\lambda tE\{Y\}=2\times5\times\frac{5}{2}=25$$

X(t)的方差为：

$$D\{x(t)\}=\lambda tE\{Y^2\}=2\times5\times\frac{43}{6}=\frac{215}{3}$$