第 6 讲 年龄与剩余寿命，泊松过程的汇合与分流

引理 1.1

对于强度为 λ 的泊松过程 $\{N(t)\}$，用 $N(t-)$ 表示区间 $[0,t)$ 内发生的事件数，则

(1) $N(t)-N(t-)=0$ a.s.;

(2) $N[s,t]=N(t)-N(s-)$ 是闭区间 $[s,t]$ 内发生的事件数;

(3) $N[s,t]=N(s,t]$ a.s..

一个使用寿命服从指数分布 $\mathcal{E}(\lambda)$ 的部件一旦失效后马上被换上同型号的备用部件继续工作. 用 $N(t)$ 表示 $(0,t]$ 中更换的部件数，则 $\{N(t)\}$ 是强度为 λ 的泊松过程. 通常称更换部件的时刻为呼叫时刻. 用 $\{S_n\}$ 表示相应的泊松流.

当 $N(t)=n$ 时，S_n 是时间 t 之前最后一次呼叫时刻，S_{n+1} 是时间 t 之后第一次呼叫时刻. 于是我们知道，$S_{N(t)}$ 是时间 t 前的最后一次呼叫时刻，$S_{N(t)+1}$ 是时间 t 之后的第一个呼叫时刻. 于是

$$S_{N(t)} \le t < S_{N(t)+1}.$$

定义

$$A(t)=t-S_{N(t)},\ R(t)=S_{N(t)+1}-t.$$

$A(t)$ 是 t 时服役的部件的使用年龄，$R(t)$ 是 t 时服役的部件的剩余寿命. 以后将 $A(t)$ 简称为 年龄，将 $R(t)$ 简称为 剩余寿命.

定理 1.2

在上面的定义下，有如下的结果：

(1) $R(t) \sim \mathcal{E}(\lambda)$;

(2) $P(A(t) \leq u) = \begin{cases} 1 - e^{-\lambda u}, & u \in [0, t), \\ 1, & u \geq t; \end{cases}$

(3) $A(t)$ 与 $R(t)$ 独立.

证明.

(1) 对 $v \geq 0$ 有 $\{R(t) > v\} = \{N(t, t+v] = 0\}$. 于是 $P(R(t) > v) = e^{-\lambda v}$, 即 $R(t) \sim \mathcal{E}(\lambda)$.

(2) 由于 $A(t) \leq t$, 所以对 $u \geq t$, $P(A(t) \leq u) = 1$. 对于 $u < t$, 由 $\{A(t) > u\} = \{N[t-u, t] = 0\}$ 和引理 1.1 (1), 我们有

$$P(A(t) > u) = P(N(t-u, t] = 0) = P(N[t-u, t] = 0) = e^{-\lambda u},$$

于是 (2) 成立.

(3) 对于 $u, v \geq 0$, 从 $N[t-u, t]$ 和 $N(t, t+v])$ 独立得到 $A(t)$ 和 $R(t)$ 独立. $\square$

定理 1.2 (1) 说明时间 t 时服役的部件剩余寿命 $R(t)$ 也服从指数分布 $\mathcal{E}(\lambda)$. 该性质是由指数分布的无记忆性所决定的.
利用上面结论和关系式

$$A(t) = t - S_{N(t)}, \ R(t) = S_{N(t)+1} - t.$$

我们有

推论 1.3

$S_{N(t)}$ 和 $S_{N(t)+1}$ 的分布函数分别为

$$P(S_{N(t)} \leq s) = \begin{cases} e^{-\lambda(t-s)} & 0 \leq s \leq t, \\ 1, & s > t, \end{cases}$$

$$P(S_{N(t)+1} \leq s) = \begin{cases} 1 - e^{-\lambda(s-t)} & 0 \leq s > t, \\ 0, & s \leq t. \end{cases}$$

指数分布的数学期望是 λ^{-1}. 如果用 $X(t)$ 表示时刻 t 时服役的部件的使用寿命，则有

$$X(t) = A(t) + R(t),\ E[X(t)] = E[A(t)] + E[R(t)] > \lambda^{-t}.$$

于是，t 时刻服役的部件的平均寿命比同型号的备用部件的平均使用寿命 $E[X_1] = \lambda^{-1}$ 要长.

注意 $A(t)$ 的分布函数在 t 处有一个跳跃高度 $e^{-\lambda t}$，用定理 1.2 (2) 得到，当 $t \to \infty$ 时，

$$\begin{aligned} E[A(t)] &= \int_0^\infty u dP(A(t) \le u) \\ &= \int_0^t \lambda u e^{-\lambda u} du + t e^{-\lambda t} \\ &\to \int_0^\infty \lambda u e^{-\lambda u} du = \frac{1}{\lambda}, \end{aligned}$$

所以有

$$\lim_{t\to\infty} E[X(t)] = \lim_{t\to\infty} (E[A(t)] + E[R(t)]) = 2\lambda^{-1} = 2E[X_1].$$

上式说明了，如果泊松过程在无穷远之前就开始了，则 t 时服役的部件的平均使用寿命是同型号备用部件的平均寿命的两倍.

注记 1.4

发生上述现象并不足为怪. 因为实际使用寿命长的部件在 t 时被遇到的概率比实际寿命短的部件被遇到的概率要大.

例 1.5

北京前门的公交车站，每 6 分钟发出一辆开往颐和园的公交车. 由于随机因素的干扰，汽车到达圆明园站时，两车之间的间隔时间成为独立同分布，服从指数分布的随机变量. 设某乘客等可能地到达车站候车，计算

(1) 他在前门候车时的平均候车时间;

(2) 他在圆明园站候车时的平均候车时间.

解.

(1) 用 T 表示甲到达前门站的时间. 对于任意长度为 6 分钟的发车间隔 $(0,6]$，已知 $T \in (0,6]$ 时，T 在 $(0,6]$ 中均匀分布. 所以平均候车时间是 3 分钟.

(2) 根据题意，公交车按照强度为 λ 的泊松过程 $\{N(t)\}$ 到达圆明园站. 由于这路公交车都要经过圆明园站，所以平均每 6 分钟停靠一辆，即有 $E[N(6)] = 6\lambda = 1$. 于是 $\lambda = 1/6$. 按照剩余寿命的定义，在 t 时到达的乘客的候车时间为 $R(t)$. 由 $R(t) \sim \mathcal{E}(\lambda)$ 知道平均候车时间为 $E[R(t)] = 6$(分钟). □

注记 1.6

上例的结果告诉我们，市内公交车的始发站和终点站距离太远时，会延长乘客的平均候车时间和降低公交车的平均利用率.

回顾：Poisson 过程

定义 2.1

设 $\lambda > 0$ 是常数. 如果计数过程 $\{N(t)\}$ 满足以下条件，则称它是强度为 λ 的泊松过程:

(1) $N(0) = 0$;

(2) $N(t)$ 是独立增量过程，有平稳增量性;

(3) 对任意 $t \geq 0$，当正数 $h \to 0$ 时有

$$\begin{cases} P(N(h) = 1) = \lambda h + o(h), \\ P(N(h) \geq 2) = o(h). \end{cases}$$

定理 2.2

设 $\{N_1(t)\}$ 和 $\{N_2(t)\}$ 是相互独立的、强度分别为 λ_1 和 λ_2 的泊松过程，则

$$N(t) = N_1(t) + N_2(t),\ \ t \geq 0$$

是强度为 $\lambda = \lambda_1 + \lambda_2$ 的泊松过程.

证明.

只需验证定义 2.1 的 (1)，(2) 和 (3). 因为 $\{N_1(t)\}$ 和 $\{N_2(t)\}$ 满足定义 2.1 的 (1)，(2) 和 (3)，所以有

$$N(0) = N_1(0) + N_2(0) = 0 + 0 = 0.$$

于是 (1) 成立. 对于任何正整数 n 以及 $0 = t_0 < t_1 < \cdots < t_n$，由 $\{N_1(t)\}$ 和 $\{N_2(t)\}$ 的独立增量性以及 $\{N_1(t)\}$ 和 $\{N_2(t)\}$ 是相互独立的我们知

$$N_1(t_{j-1}, t_j],\ N_2(t_{j-1}, t_j],\ j = 1, 2, \cdots, n$$

相互独立. 于是

$$N(t_{j-1}, t_j] = N_1(t_{j-1}, t_j] + N_2(t_{j-1}, t_j],\ j = 1, 2, \cdots, n$$

相互独立，所以 $\{N(t)\}$ 是独立增量过程.

证明.

又因为 $N_1(t_1,t_2]$ 和 $N_1(t_1+s,t_2+s]$ 同分布，$N_2(t_1,t_2]$ 和 $N_2(t_1+s,t_2+s]$ 同分布，所以 $N(t_1,t_2]=N_1(t_1,t_2]+N_2(t_1,t_2]$ 与

$$N(t_1+s,t_2+s]=N_1(t_1+s,t_2+s]+N_2(t_1+s,t_2+s]$$

同分布. 所以 $\{N(t)\}$ 是平稳增量过程. 于是 (2) 成立.

证明.

下面验证普通性：设 $\lambda = \lambda_1 + \lambda_2$. 我们有当 $h \to 0$ 时，

$$\begin{aligned}
P(N(h) = 1) &= P(N_1(h) = 1, N_2(h) = 0) \\
&\quad + P(N_1(h) = 0, N_2(h) = 1) \\
&= P(N_1(h) = 1)P(N_2(h) = 0) \\
&\quad + P(N_1(h) = 0)P(N_2(h) = 1) \\
&= (\lambda_1 h + o(h))(1 - \lambda_2 h + o(h)) \\
&\quad + (\lambda_2 h + o(h))(1 - \lambda_1 h + o(h)) \\
&= (\lambda_1 + \lambda_2)h + o(h) = \lambda h + o(h).
\end{aligned}$$

证明.

$$
\begin{aligned}
P(N(h)=0) &= P(N_1(h)=0, N_2(h)=0) \\
&= (1-\lambda_1 h + o(h))(1-\lambda_2 h + o(h)) \\
&= 1-(\lambda_1+\lambda_2)h + o(h) = 1-\lambda h + o(h).
\end{aligned}
$$

于是得到

$$
\begin{aligned}
P(N(h)\geq 2) &= 1 - P(N(h)=0) - P(N(h)=1) \\
&= 1 - [1-\lambda h + o(h)] - [\lambda h + o(h)] = o(h).
\end{aligned}
$$

于是 (3) 成立. $\square$

应用归纳法我们知

推论 2.3

设 $\{N_j(t)\}(j=1,2,\cdots,m)$ 是相互独立的，强度分别为 λ_j 的泊松过程，则

$$N(t)=N_1(t)+N_2(t)+\cdots+N_m(t),\ \ t\geq 0$$

是强度为 $\lambda=\lambda_1+\lambda_2+\cdots+\lambda_m$ 的泊松过程.

上面推论所描述的性质称为泊松过程的可加性.

1. 年龄与剩余寿命
2. 泊松过程的汇合
3. **泊松过程的分流**

- 设旅客按照强度为 λ 的泊松过程到达长途汽车站，每次到达的旅客乘 A 线的概率是 p，乘 B 线的概率为 $q = 1 - p$，且与到达时间独立，也与其他的到达行为独立.
- 用 $N_1(t)$ 表示时间 $[0, t]$ 内乘 A 线的旅客到达次数，用 $N_2(t)$ 表示时间 $[0, t]$ 内乘 B 线的旅客到达次数.
- 因为旅客选择 A 线还是 B 线仅由他自己决定，与其他旅客的行为无关. 所以前往 A 线的旅客流与前往 B 线的旅客流是独立的，也就是说计数过程 $\{N_1(t)\}$ 和 $\{N_2(t)\}$ 应当是相互独立的.
- 用 λ_1 和 λ_2 分别表示他们的强度，于是 $\lambda = \lambda_1 + \lambda_2$.

引入独立同分布的随机变量

$$Y_j = \begin{cases} 1, & \text{当第 } j \text{ 次到达乘 A 线}, \\ 0, & \text{当第 } j \text{ 次到达乘 B 线}, \end{cases} \tag{3.1}$$

则有

$$P(Y_j = 1) = p, \ P(Y_j = 0) = q, \ j = 1, 2, \cdots .$$

如果已知在时刻 $[0,t]$ 内有 $N(t)=n$ 次到达，则 $[0,t]$ 内有 $\sum_{j=1}^{n} Y_j$ 次到达是乘 A 线的. 当在时刻 $[0,t]$ 内有 $N(t)$ 次到达，则有

$$N_1(t)=\sum_{j=1}^{N(t)} Y_j \tag{3.2}$$

次到达是乘 A 线的. 同理有

$$N_2(t)=\sum_{j=1}^{N(t)} (1-Y_j) \tag{3.3}$$

次到达是乘 B 线的.
这里和以后对 $a<b$，我们总是规定 $\sum_{j=b}^{a}(\cdot)=0$. 对于 $k<0$ 或 $k>n$，总规定 $C_n^k=0$.

定理 3.1

设 $\{N(t)\}$ 是强度为 λ 的泊松过程，$\{Y_j\}$ 是独立同分布的随机序列，服从两点分布 (3.1). 计数过程 $\{N_1(t)\}$ 和 $\{N_2(t)\}$ 分别由 (3.2) 和 (3.3) 定义. 如果 $\{Y_j\}$ 与 $\{N(t)\}$ 独立，则 $\{N_1(t)\}$ 与 $\{N_2(t)\}$ 相互独立，分别是强度为 $\lambda_1 = \lambda p$ 和 $\lambda_2 = \lambda q$ 的泊松过程.

在上面定理中，$\{N_1(t)\}$ 和 $\{N_2(t)\}$ 被称为 $N(t)$ 的 分流过程.

证明.

对于 $0 \le s < t$,

$$N_1(s,t] = N_1(t) - N_1(s) = \sum_{j=N(s)+1}^{N(t)} Y_j \tag{3.4}$$

是时刻 $(s,t]$ 内乘 A 线的乘客数. 因为 $\{Y_j\}$ 与 $\{N(t)\}$ 独立，对于 $k \ge l$，容易计算出

$$\begin{aligned} P(N_1(s,t] = n | N(s) = l, N(t) = k) &= P(\sum_{j=l+1}^{k} Y_j = n | N(s) = l, N(t) = k) \\ &= P(\sum_{j=l+1}^{k} Y_j = n) = P(\sum_{j=1}^{k-l} Y_j = n) \\ &= g(k-l, n), \end{aligned}$$

其中 $g(k-l,n) = P(\sum_{j=l}^{k} Y_j = n)$ 满足 $g(1,1) = p$ 以及对任意 $n \ge 1$，$g(0.n) = 0$.

证明.

由条件概率的定义我们知

$$P(N_1(s,t]=n|N(s),N(t))=g(N(s,t],n).$$

对上式求期望得到

$$P(N_1(s,t]=n)=E[g(N(s,t],n)].$$

我们证明 $\{N_1(t)\}$ 为强度 $\lambda_1=\lambda p_1$ 的泊松过程，我们验证定义 2.1 中的条件 (1)，(2)，(3).

(1) $N_1(0)=\sum_{j=1}^{0} Y_j=0$.

证明.

(2) 独立增量性: 对任意正整数 m, $0 = t_0 < t_1 < t_2 < \cdots < t_m$ 以及整数 $0 = n_0 \le n_1 \le n_2 \le \cdots \le n_m$, 定义

$$\boldsymbol{N} = (N(t_1), N(t_2), \cdots, N(t_m)), \ \boldsymbol{n} = (n_1, n_2, \cdots, n_m).$$

从 (3.4) 知道, 在条件 $\boldsymbol{N} = \boldsymbol{n}$ 下, 随机变量

$$N_1(t_{j-1}, t_j] = \sum_{i=n_{j-1}+1}^{n_j} Y_i, \ j = 1, 2, \cdots, m$$

相互独立.

证明.

于是得到

$$
\begin{aligned}
& P(N_1(t_{j-1},t_j]=k_j, 1\le j\le m|\boldsymbol{N}=\boldsymbol{n}) \\
=& P(\sum_{i=n_{j-1}+1}^{n_j} Y_i=k_i, 1\le j\le m|\boldsymbol{N}=\boldsymbol{n}) \\
=& P(\sum_{i=1}^{n_1} Y_i=k_1)P(\sum_{i=n_1+1}^{n_2} Y_i=k_2)\cdots P(\sum_{i=n_{m-1}+1}^{n_m} Y_i=k_m) \\
=& g(n_1,k_1)g(n_2-n_1,k_2)\cdots g(n_m-n_{m-1},k_m).
\end{aligned}
$$

按条件概率的定义我们知

$$
\begin{aligned}
& P(N_1(t_{j-1},t_j]=k_j, 1\le j\le m|\boldsymbol{N}) \\
=& g(N(t_1),k_1)g(N(t_1,t_2],k_2)\cdots g(N(t_{m-1},t_m],k_m).
\end{aligned}
$$

证明.

对于上式求数学期望，利用 $\{N(t)\}$ 的独立增量性我们知

$$
\begin{aligned}
&P(N_1(t_{j-1},t_j]=k_j, 1\le j\le m)\\
=&E[g(N(t_1),k_1)]E[g(N(t_1,t_2],k_2)]\cdots E[g(N(t_{m-1},t_m],k_m)]\\
=&P(N_1(t_0,t_1]=k_1)P(N_1(t_1,t_2]=k_2)\cdots P(N_1(t_{m-1},t_m]=k_m).
\end{aligned}
$$

这说明随机变量 $N_1(t_{j-1},t_j](1\le j\le n)$ 相互独立. 于是 $\{N_1(t)\}$ 是独立增量过程.

证明.

平稳增量性：因为 $N(t_1, t_2]$ 和 $N(t_1+s, t_2+s]$ 同分布，所以我们知对任意 $n \in \mathbb{N}$，有

$$
\begin{aligned}
P(N(t_1, t_2] = n) &= Eg(N(t_1, t_2], n) \\
&= E(N(t_1+s, t_2+s], n) \\
&= P(N(t_1+s, t_2+s] = n)
\end{aligned}
$$

于是 $N_1(t_1, t_2]$ 和 $N_1(t_1+s, t_2+s]$ 同分布.

证明.

(3) 普通性: 注意到 $g(0,1)=0$, $g(1,1)=p$ 以及

$$g(N(t,t+h],1)I[N(t,t+h]=1]=\begin{cases}0 & N(t,t+h]\neq 1\text{ 时},\\ g(1,1) & N(t,t+h]=1\text{ 时},\end{cases}$$

我们有

$$\begin{aligned}P(N_1(t,t+h]=1) &= E[g(N(t,t+h],1)]\\ &= E[g(N(t,t+h],1)(I[N(t,t+h]=1]+I[N(t,t+h]\geq 2])]\\ &= g(1,1)P(N(t,t+h]=1)+o(h)\\ &= p(\lambda h+o(h))+o(h)=p\lambda h+o(h).\end{aligned}$$

证明.

因为 $N_1(s,t] \le N(s,t]$ 对 $s<t$ 成立，故由

$$P(N_1(t,t+h] \ge 2) \le P(N(t,t+h] \ge 2) = o(h)$$

我们有 $P(N_1(t,t+h] \ge 2) = o(h)$.

综上我们知 $\{N_1(t)\}$ 是强度为 λp 的泊松过程. 完全对称地可以证明 $\{N_2\}$ 是强度为 λq 的泊松过程.

证明.

下面证明 $\{N_1(t)\}$ 和 $\{N_2(t)\}$ 相互独立. 我们只需要证明对于任何 $n \geq 1$, $0 \leq t_1 < t_2 < \cdots < t_n$, 随机变量

$$(N_1(t_1), N_1(t_2), \cdots, N_1(t_n)) \text{ 和 } (N_2(t_1), N_2(t_2), \cdots, N_2(t_n))$$

独立. 对于整数

$$0 = k_0 \leq k_1 \leq k_2 \leq \cdots \leq k_n,\ 0 = m_0 \leq m_1 \leq m_2 \leq \cdots \leq m_n,$$

引入 $n_j = k_j + m_j$, $\boldsymbol{n} = (k_1 + m_1, k_2 + m_2, \cdots, \cdots, k_n + m_n)$. 随机变量

$$\xi_j = \sum_{i=n_{j-1}+1}^{n_j} Y_i,\ j = 1, 2, \cdots, n$$

相互独立.

证明.

注意 ξ_j 服从二项分布 $\mathcal{B}(n_j-n_{j-1},p)$，我们可以得到

$$
\begin{aligned}
&P(N_1(t_j)=k_j,N_2(t_j)=m_j;1\le j\le n)\\
=&P(N_1(t_j)=k_j,N(t_j)=n_j;1\le j\le n)=P(N_1(t_{j-1},t_j]=k_j-k_{j-1},N(t_j)=n_j;1\le j\le n)\\
=&P(\xi_j=k_j-k_{j-1},N(t_{j-1},t_j]=n_j-n_{j-1};1\le j\le n)\\
=&\prod_{j=1}^{n}[P(\xi_j=k_j-k_{j-1})P(N(t_{j-1},t_j]=n_j-n_{j-1})]\\
=&\prod_{i=1}^{n}\left[\frac{(n_j-n_{j-1})!}{(k_j-k_{j-1})!(m_j-m_{j-1})!}p^{k_j-k_{j-1}}q^{m_j-m_{j-1}}\cdot\frac{[\lambda(t_j-t_{j-1})]^{n_j-n_{j-1}}}{(n_j-n_{j-1})!}e^{-\lambda(t_j-t_{j-1})}\right]\\
=&\prod_{i=1}^{n}\left[\frac{[\lambda p(t_j-t_{j-1})]^{k_j-k_{j-1}}}{(k_j-k_{j-1})!}e^{-\lambda p(t_j-t_{j-1})}\frac{[\lambda q(t_j-t_{j-1})]^{m_j-m_{j-1}}}{(m_j-m_{j-1})!}e^{-\lambda q(t_j-t_{j-1})}\right]\\
=&P(N_1(t_j)=k_j,1\le j\le n)P(N_2(t_j)=m_j,1\le j\le n).
\end{aligned}
$$

这说明 $(N_1(t_1),N_1(t_2),\cdots,N_1(t_n))$ 和 $(N_2(t_1),N_2(t_2),\cdots,N_2(t_n))$ 独立. $\square$

定理 3.2

设旅客按强度为 λ 的泊松过程到达某长途汽车站，每次到达的旅客以概率 p_i 前往 A_i 线，且前往哪个路线与到达时间独立，也与其他的到达行为独立. 用 $N_i(t)$ 表示 $[0,t]$ 内前往 A_i 线的到达次数时，$\{N_i(t)\}$ 是强度 $\lambda_i = p_i\lambda$ 的泊松过程. 当 $p_1 + p_2 + \cdots + p_n = 1$ 时，这 n 个泊松过程是相互独立的.

例 3.3

汽车按 Poisson 流驶向立体立交桥. 经过调查知道，由东面每分钟平均驶入 6 辆汽车，由南面每分钟平均驶入 6.5 辆汽车，由西面平均每分钟平均驶入 9 辆汽车，由北面平均每分钟平均驶入 8.5 辆汽车. 在桥 A 上，每辆车向左或向右转向行驶的概率是 0.3，直行的概率是 0.35，掉头行驶的概率是 0.05. 计算各个方向上，离开立交桥的汽车流的车流强度.

证明.

用 $\{N_1(t)\}$ 表示由东面驶入的汽车流. 根据题意，每分钟平均驶入 $E[N_1(t,t+1]]=6$ 辆汽车. 所以 Poisson 过程 $\{N_1(t)\}$ 的强度是 $\lambda_1=6$ (辆/分钟). 根据 Poisson 过程的可分解性，东面驶入的车流 $\{N_1(t)\}$ 分流给东、南、西、北的分流强度分别是 $0.05\lambda_1$，$0.3\lambda_1$，$0.35\lambda_1$，$0.3\lambda_1$. 完全类似地可以列出其他方向的汽车流的分流情况，列入下面的分流表:

证明.

方向	向东分流	向南分流	向西分流	向北分流
东面驶入 $\lambda_1 = 6.0$	$0.05\lambda_1$	$0.30\lambda_1$	$0.35\lambda_1$	$0.30\lambda_1$
南面驶入 $\lambda_2 = 6.5$	$0.30\lambda_2$	$0.05\lambda_2$	$0.30\lambda_2$	$0.35\lambda_2$
西面驶入 $\lambda_3 = 9.0$	$0.35\lambda_3$	$0.30\lambda_3$	$0.05\lambda_3$	$0.30\lambda_3$
北面驶入 $\lambda_4 = 8.5$	$0.30\lambda_4$	$0.35\lambda_4$	$0.30\lambda_4$	$0.05\lambda_4$
驶出强度 (辆 / 分钟)	$\lambda_E = 7.95$	$\lambda_S = 7.80$	$\lambda_W = 7.05$	$\lambda_N = 7.20$

表中最后一行表示各个方向驶出立交桥的车流强度，是其所在列各分流强度之和.
所有的分流都是 Poisson 流. □

例 3.4

从时刻 $t=0$ 开始，客户按强度为 λ 的泊松流点击一个网站．每个客户点击后的浏览时间是相互独立的，有共同的分布函数 $G(t)$．用 $N_1(t)$ 表示 t 时刻已经离线的客户数，用 $N_2(t)$ 表示 t 时在线的客户数，则 $N_1(t)$ 和 $N_2(t)$ 是两个相互独立的泊松随机变量，分别有数学期望

$$E[N_1(t)] = \lambda \int_0^t G(s)ds, \ E[N_2(t)] = \lambda \int_0^t \overline{G}(s)ds,$$

其中 $\overline{G}(s) = 1 - G(s)$.

证明.

对于一个客户来讲，用 S 表示他进入网站的时间，用 A 表示他 t 时已经离线，用 Y 表示他的在线时间. 对于 $s \le t$，有

$$P(A|S=s) = P(Y \le t-s) = G(t-s).$$

因为在 $S \le t$ 条件下，S 在 $[0,t]$ 内均匀分布，且 $P_t(A) \equiv P(A|S \le t)$ 是概率，所以

$$\begin{aligned}
p = P_t(A) &= \int_0^t P_t(A|S=s)dP_t(S \le s) \\
&= \int_0^t P(A|S \le t, S=s)dP(S \le s|S \le t) \\
&= \frac{1}{t}\int_0^t P(A|S=s)ds = \frac{1}{t}\int_0^t G(t-s)ds = \frac{1}{t}\int_0^t G(s)ds, \\
q = P(A^c|S \le t) &= 1-p = \frac{1}{t}\int_0^t \bar{G}(s)ds.
\end{aligned}$$

证明.

每个在时刻 $[0,t]$ 内进入网站的人在 t 时离线的概率是 p，在线的概率是 q，与其他客户的行为独立. 用 $\{N(t)\}$ 表示所述的泊松过程，利用二项分布得到

$$P(N_1(t)=k, N_2(t)=j|N(t)=k+j)=C_{k+j}^k p^k q^j.$$

于是得到

$$\begin{aligned} P(N_1(t)=k, N_2(t)=j) &= P(N(t)=k+j)P(N_1(t)=k, N_2(t)=j|N(t)=k+j) \\ &= P(N(t)=k+j)C_{k+j}^k p^k q^j \\ &= \frac{(\lambda tp)^k}{k!}e^{-\lambda tp}\frac{(\lambda tq)^j}{j!}e^{-\lambda tq}. \end{aligned}$$

证明.

分别对 j,k 求和，就得到边缘分布

$$P(N_1(t)=k)=\frac{(\lambda tp)^k}{k!}e^{-\lambda tp},\ P(N_2(t)=j)=\frac{(\lambda tq)^j}{j!}e^{-\lambda tq}.$$

这说明 $N_1(t),N_2(t)$ 是相互独立的泊松随机变量，分别有数学期望 λtp 和 λtq，即有

$$E[N_1(t)]=\lambda tp=\lambda\int_0^t G(s)ds,\ E[N_2(t)]=\lambda tq=\lambda\int_0^t \overline{G}(s)ds.$$

从而得证. □

用 $b_G=\inf\{s|G(s)=1\}$ 表示 $G(t)$ 的右端点，用 μ_G 表示 $G(t)$ 的数学期望. 当 $t\geq b_G$ 时，

$$E[N_2(t)]=\lambda\int_0^{b_G}\overline{G}(s)ds=\lambda\mu_G,$$

$$E[N_1(t)]=E[N(t)]-E[N_2(t)]=\lambda(t-\mu_G),$$

这说明 $t\geq b_G$ 后，在线的客户平均数稳定在 $e[N_2(t)]=\lambda\mu_G$. 在 $(b_G,b_G+t]$ 中离线的客户数

$$N_3(t)=N_1(t+b_G)-N_1(b_G),\ t\geq 0$$

有数学期望

$$E[N_3(t)]=E[N_1(t+b_G)]-E[N_1(b_G)]=\lambda t.$$

- 上述例子可以有如下一般的描述.
- 假设有 k 种可能类型的事件, 而一个事件被分类为类型 i $(i=1,\cdots,k)$ 的概率依赖于事件发生的时间, 但独立于之前发生的任何事件.
- 特别地, 假设一个在时间 y 发生事件被分类为类型 i $(i=1,\cdots,k)$ 的概率为 $P_i(y)$, 其中 $\sum_{i=1}^{k} P_i(y)=1$.

命题 3.5

如果 $N_i(t)$, $(i = 1, \cdots, k)$ 表示到时刻 t 为止类型 1 事件发生的个数, 那么 $N_i(t)$ $(i = 1, \cdots, k)$ 是均值

$$E[N_i(t)] = \lambda \int_0^t P_i(s)ds$$

的泊松随机变量.

应用: 追踪感染 HIV 的人数

- 从个体感染 HIV (人体免疫缺陷病毒) 到症状出现, 有相对较长的潜伏期.
- 因此, 对于卫生部门来说, 对任意给定时间确定总体中的感染人数很困难.
- 我们现在介绍一个描述这一现象的初级近似模型, 以大致估计受感染的人数.
- 我们假设感染 HIV 的人数构成一个强度为 λ 的泊松过程.
- 假设个体从感染到出现症状的时间是服从已知分布的随机变量. 同时假设不同的感染个体的潜伏期是独立的.

令 $N_1(t)$ 为到时刻 t 为止已经出现症状的人数, $N_2(t)$ 为到时刻 t 为止 HIV 阳性但是还没有出现任何症状的人数. 由于在时刻 s 受到感染的个体以概率 $G(t-s)$ 在时刻 t 出现症状, 而以概率 $\overline{G}(t-s)$ 在时刻 t 不出现症状. 因此由命题 **??** 知, $N_1(t)$ 和 $N_2(t)$ 是独立的泊松随机变量, 其均值分别为

$$E[N_1] = \lambda \int_0^t G(t-s)ds = \lambda \int_0^t G(y)dy$$

和

$$E[N_2] = \lambda \int_0^t \overline{G}(t-s)ds = \lambda \int_0^t \overline{G}(y)dy.$$

如果我们知道 λ, 那么我们可以通过均值 $E[N_2(t)]$ 来估计受到感染但是在时刻 t 没有出现症状的人数 $N_2(t)$. 但是由于 λ 未知, 我们需要先得到它的一个估计.

我们现在知道 $N_1(t)$ 的值，从而可以将它作为均值 $E[N_1(t)]$ 的估计，即如果到时刻 t 为止出现症状的人数为 n_1，那么我们可以估计

$$n_1 \approx E[N_1(t)] = \lambda \int_0^t G(y)dy.$$

从而我们可以用

$$\hat{\lambda} = n_1 / \int_0^t G(y)dy$$

来估计 λ. 利用 λ 的这个估计来估计 $N_2(t)$:

$$N_2(t) \text{ 的估计} = \hat{\lambda} \int_0^t \overline{G}(y)dy = \frac{n_1 \int_0^t \overline{G}(y)dy}{\int_0^t G(y)dy}$$

来估计受到感染但是在时间 t 没有出现症状的人数.