随机过程
Stochastic Processes

讲义10：泊松过程-3

目录

（课本第5章部分3）

10.1 到达间隔与等待时间

（课本5.3.3）

10.1.1. 泊松过程的到达间隔时间列

- 泊松过程的到达间隔时间列$\{T_n, n = 1,2,\ldots\}$:
  - 第一个事件到达时间为T_1
  - 对$n > 1$，T_n记事件$n-1$和n之间的耗时

- 例：$N(t)$等于在或早于时刻t进入商店的人数
  - 9点开门，第一个人9点15来，第二个人9点20，第三个人9点30.
  - 有：$T_1 = 15$分钟, $T_2 = 5$分钟，$T_3 = 10$分钟

10.1.2. T_n的分布

- 命题5.2 $T_n(n=1,2,\dots)$是$\boldsymbol{i.i.d.}$的均值$1/\lambda$指数随机变量
  - 泊松过程的无记忆性
    - 泊松过程在任意时间点，概率意义下重新开始

10.1.2. T_n的分布

- 命题5.2 $T_n(n=1,2,\dots)$是$\boldsymbol{i.i.d.}$的均值$1/\lambda$指数随机变量

证明：

先推导T_1的分布：均值为$1/\lambda$的指数随机变量

- $P\{T_1>t\}=P\{N(t)=0\}=e^{-\lambda t}$
- 定理5.1：$N(t)\sim Poisson(\lambda t)$
- $\{T_1>t\}\leftrightarrow[0,t]$无事件发生

再推导T_2的分布：均值为$1/\lambda$的指数随机变量

- $P\{T_2>t\}=\int_0^\infty P\{T_2>t|T_1=s\}f_{T_1}(s)ds$
- $P\{T_2>t|T_1=s\}=P\{(s,s+t]中0个事件|T_1=s\}$
  $=P\{(s,s+t]中0个事件\}$
  $=\mathrm{e}^{-\lambda t}$
- $P\{T_2>t\}=\mathrm{e}^{-\lambda t}$

10.1.2. T_n的分布

- 命题5.2 $T_n(n=1,2,\ldots)$是$\boldsymbol{i.i.d.}$的均值$1/\lambda$指数随机变量

说明T_1与T_2的独立性：

- $F_{T_2|T_1=s}(t)=1-P\{T_2>t|T_1=s\}=1-e^{-\lambda t}$
- $f_{T_2|T_1=s}(t)=\lambda e^{-\lambda t}=f_{T_2}(t)$
- $f_{T_2,T_1}(t,s)=f_{T_2|T_1=s}(t)f_{T_1}(s)=f_{T_2}(t)f_{T_1}(s)$
- 符合密度函数定理

- 延续上述推导可证明结论

10.1.3. 泊松过程的等待时间S_n（到达时间）

- 等待时间列$\{S_n, n = 1,2,\dots\}$:
  - S_n记从开始到事件n的等待时间
    - 例：$N(t)$等于在或早于时刻t进入一个商店的人数，九点开门，第一个人9点15来，第二个人9点20，第三个人9点30.
    - 有：$S_1 = 15$分钟, $S_2 = 20$分钟，$S_3 = 30$分钟

- S_n的分布：具有参数n和λ的伽马分布
  - 即$f_{S_n}(t) = \lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}$， $t \geq 0$

- S_n的分布求解1：利用间隔时间
  - 由$S_n = \sum_{i=1}^{n} T_i$，$n \geq 1$；+ 讲义8-9中命题5.2

10.1.3. 泊松过程的等待时间S_n（到达时间）

- S_n的分布求解2：利用泊松过程第2种定义
  - 注意到：$N(t) \geq n \leftrightarrow S_n \leq t$
  - 所以：$F_{S_n}(t) = P\{S_n \leq t\} = P\{N(t) \geq n\} = \sum_{j=n}^{\infty} e^{-\lambda t} \frac{(\lambda t)^j}{j!}$
  - 微分可求得结果

10.1.3. 泊松过程的等待时间S_n（到达时间）

- S_n的分布求解3：利用泊松过程第1种定义，$o(h)$的解法
  - 回忆密度函数的近似表达式：$P\{x < X \leq x + \Delta x\} = f(x)\Delta x + o(\Delta x)$,
  - 希望给$P\{t < S_n < t + h\}$凑形式：

$$P\{t < S_n < t + h\}$$
$$= P\{N(t) = n - 1, 在(t, t + h)中有一个事件\} + o(h)$$

$$= P\{N(t) = n - 1\}P\{在(t, t + h)中有一个事件\} + o(h)$$
$$= e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}[\lambda h + o(h)] + o(h)$$
$$= e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}\lambda h + o(h)$$

  - 这里的h就是上述Δx，两边同除h，令$h \to 0$
  - 就得出了$f_{S_n}(t) = \lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}$

10.1.3. 间隔时间和等待时间的例子

- 例5.13：以小时为单位，杰伦奶茶店的顾客到达过程是速率$\lambda = 10$的泊松过程。奶茶店9点钟开门。
  - 1：本泊松过程的事件是什么？
    - 顾客的到达
  - 2：直到第10个顾客到达的时间的期望时间是几点？
    - $E[S_{10}] = E\left[\sum_{i=1}^{10} T_i\right] = \frac{10}{\lambda} = 1$小时
  - 3：第10个顾客和第11个顾客到达之间的时间超过10分钟的概率是多少？
    - $P\{T_{11} > 1/6\} = e^{-1/6\lambda} = e^{-10/6}$

10.1.4. 利用间隔时间定义泊松过程

- **泊松过程等价定义2(泊松过程第3种定义)**: 假设有均值为$1/\lambda$的$i.i.d.$指数随机变量$\{T_n, n \geq 1\}$.
- 现在定义一个计数过程，称过程的第n个事件在时间

$$S_0 \equiv 0, \qquad S_n \equiv T_1 + \cdots + T_n$$

发生。

- 现在这个定义的计数过程$\{N(t), t \geq 0\} \equiv max\{n: S_n \leq t\}$是速率$\lambda$的泊松过程。
  - ➢ 这个定义与前2种定义的等价性的严格证明超纲
  - ➢ 但部分结论是较容易得到的，例如，从第3种定义推第2种定义的核心条件：$N(t) \overset{D}{=} \text{Poi}(\lambda t)$
  - ➢ 要点: $P(N(t) = n) = P\{S_n \leq t < S_{n+1} = S_n + T_{n+1}\}$
  - ➢ $S_n \overset{D}{=} \Gamma(n, \lambda), T_{n+1} \overset{D}{=} \exp(\lambda)$; 可说明$N(t) \overset{D}{=} \text{Poi}(\lambda t)$

10.2 泊松过程的进一步性质
(课本5.3.4 部分)

10.2.1. 泊松过程的分流

- 考虑一个速率为λ的泊松过程$\{N(t), t \geq 0\}$，而且假设每次发生的事件分为I型和II型事件。

- 进一步假设每个事件独立于所有其他事件，以概率p为I型事件，以概率$1-p$为II型事件。
  - 例如顾客到达，顾客的性别可认为是I型和II型事件。

- 以$N_1(t)$和$N_2(t)$分别记在$[0, t]$发生的I型和II型事件的个数。
  - 注意到：$N(t) = N_1(t) + N_2(t)$

- 命题5.3：$\{N_1(t), t \geq 0\}$和$\{N_2(t), t \geq 0\}$两者分别是速率为λp和$\lambda(1-p)$的泊松过程。且，这两个泊松过程彼此独立。

10.2.1. 泊松过程的分流

- 命题5.3：$\{N_1(t), t \geq 0\}$和$\{N_2(t), t \geq 0\}$两者分别是速率为λp和$\lambda(1-p)$的泊松过程。且，这两个泊松过程彼此独立。

10.2.1. 泊松过程的分流

- 命题5.3理解I：$\{N_1(t), t \geq 0\}$ 和$\{N_2(t), t \geq 0\}$是泊松过程
  - 以$\{N_1(t), t \geq 0\}$ 为例
  - 条件1： $N_1(0) = 0$ 得自事实 $N(0) = 0$
  - 条件2：根据描述，$\{N_1(t), t \geq 0\}$和$\{N_2(t), t \geq 0\}$直观继承了过程$\{N(t), t \geq 0\}$的平稳和独立增量性。
    - 因为在一个区间中的I型事件的个数的分布可以由取条件于这个区间中的事件个数得到。而这个区间中事件个数的分布仅依赖于区间的长度。并且与任意与它不相交的区间中发生的事件是独立的。
  - 条件3：$P\{N_1(h) = 1\} =$ $P\{N_1(h) = 1|N(h) = 1\}P\{N(h) = 1\} +$ $P\{N_1(h) = 1|N(h) \geq 2\}P\{N(h) \geq 2\} = p\big(\lambda h + o(h)\big) +$ $o(h) = \lambda ph + o(h)$
  - 条件4：$P\{N_1(h) \geq 2\} \leq P\{N(h) \geq 2\} = o(h)$

10.2.1. 泊松过程的分流

- 命题5.3理解II：独立性
  - 独立增量性说明$(t, t+h)$区间中的事件独立于与$(t, t+h)$没有重叠的区间的事件
  - 因此，理解$(t, t+h)$区间中，$N(t)$ 分流出的两个泊松随机变量独立即可，即$N_1(t, t+h)$与$N_2(t, t+h)$独立
  - 回归到了第2讲例子3.23的情景

10.2.1. 泊松过程的分流

- 例 5.14：移民以每周10人的泊松速率到达A地区，若每个移民是英格兰后裔的概率为1/12，那么在二月份没有英格兰后裔移民到A地区的概率是多少？
  - 原泊松过程：$\lambda = 40$ ？
  - 分流泊松过程：
    - 英格兰后裔：$\lambda p = 10/3$
    - 非英格兰后裔：$\lambda(1-p) = 110/3$
  - $P\{N(1) = 0\} = e^{-\frac{10}{3}} \approx 0.036$
    - $N(1) \sim Pois(10/3)$

10.2.1. 泊松过程的分流

- 例 5.15：你作为老板想出售某件物品，外界的出价以速率为λ的泊松过程到达。假定每次出价是具有密度函数$f(x)$的随机变量的值。一旦出价提供给你，你必须接受或者拒绝并等待下一个出价。
- 假设部件卖出去以前，你以每个单位时间c的速率需要花费，而你的目标是使得你的期望总回报最大，其中总回报等于收到的钱的数目减去总花费。
- 假设你的策略是，接受第一个超过某个特定值y的出价。(这种类型的策略称为y策略)y的最优值是多少？

$$p = f(y)$$

10.2.1. 泊松过程的分流

第一步：确定子泊松过程：

- 以X记一个随机出价的值，而以$\bar{F}(x)=P(X>x)=\int_x^\infty f(u)du$记它的分布函数的尾部概率
- 每次出价，以概率$\bar{F}(y)$大于y，这就推出了这种出价按照速率为$\lambda\bar{F}(y)$的泊松过程发生
- 满足要求的出价事件的过程，服从速率为$\lambda\bar{F}(y)$的泊松过程

10.2.1. 泊松过程的分流

第二步：确定期望回报：

- 因此，直到一次出价被接受的时间是一个速率为$\lambda\bar{F}(y)$的指数随机变量。
- 以$R(y)$记从y策略得到的总回报，有

$$
\begin{aligned}
E[R(y)] &= E[\text{接受的出价}] - cE[\text{到接受的时间}] \\
&= E[X|X > y] - \frac{c}{\lambda\bar{F}(y)} \\
&= \int_0^\infty x f_{X|X>y}(x)dx - \frac{c}{\lambda\bar{F}(y)} \\
&= \int_y^\infty x \frac{f(x)}{\bar{F}(y)} dx - \frac{c}{\lambda\bar{F}(y)} \\
&= \frac{\int_y^\infty x f(x)dx - c/\lambda}{\bar{F}(y)}
\end{aligned}
$$

10.2.1. 泊松过程的分流

- 积分求导（也称费曼技巧）的定理：

若a(t)和$b(t)$皆可导，则

$$\frac{d}{dt}\int_{a(t)}^{b(t)} f(x,t)dx =$$

$$\int_{a(t)}^{b(t)} \frac{\partial}{\partial t} f(x,t)dx + f(b(t),t)b'(t) - f(a(t),t)a'(t)$$

当然，要满足下列条件：

1. 存在$\alpha < \beta, c_1 < c_2$ 使得 $f(x,t)$ 和 $\frac{\partial}{\partial t} f(x,t)$ 在$[\alpha,\beta] \times (c_1,c_2)$上是连续的
2. 对所有$t \in (c_1,c_2)$，有 $a(t) \in [\alpha,\beta], b(t) \in [\alpha,\beta]$
3. 对$(x,t) \in [\alpha,\beta] \times (c_1,c_2)$，$|f(x,t)|$ 和 $|\frac{\partial}{\partial t} f(x,t)|$有上界，具体的对$(x,t) \in [\alpha,\beta] \times (c_1,c_2)$, 有$|f(x,t)| \leq A(x), |\frac{\partial}{\partial t} f(x,t)| \leq B(x)$, 且$\int_{\alpha}^{\beta} A(x)dx$和 $\int_{\alpha}^{\beta} B(x)dx$存在。

10.2.1. 泊松过程的分流

第三步：求解最佳临界值：

$$\frac{d}{dy}E[R(y)]=0 \leftrightarrow \left(\int_y^{\infty} xf(x)dx-\frac{c}{\lambda}\right)f(y)-yf(y)\bar{F}(y)=0$$

- 所以，y的最优值满足

$$y\bar{F}(y)=\int_y^{\infty} xf(x)dx-c/\lambda$$

- 即，　$\text{y}\int_y^{\infty} f(x)dx=\int_y^{\infty} xf(x)dx-c/\lambda$
- 也即，　$\int_y^{\infty}(x-y)f(x)dx=c/\lambda$
- 注意到：$\int_y^{\infty}(x-y)f(x)dx$ 是关于y单调递减的，所以存在唯一的y满足上述方程，令该值为y^*, 有

$$\int_{y^*}^{\infty}(x-y^*)f(x)dx=c/\lambda$$

- 将$y=y^*$带入，由$\int_{y^*}^{\infty} xf(x)dx=\frac{c}{\lambda}+y^*\int_{y^*}^{\infty} f(x)dx$
- 得$E[R(y)]=\dfrac{y^*\int_{y^*}^{\infty} f(x)dx}{\bar{F}(y^*)}=y^*$，即，最佳临界值也是最佳期望回报值！

10.2.2. 泊松过程的分流的推广1

泊松过程的多类型分流:

- 考虑一个速率为λ的泊松过程$\{N(t), t \geq 0\}$，而且假设每次发生的事件分为I、II、…型事件。
- 假设每个事件独立于所有其他事件，以概率p_1为I型事件，以概率p_2为II型事件, 等，以概率p_n为n型事件，$\sum_{i=1}^{n} p_i = 1$
- 注意到，这里的概率是常数
- 以$N_1(t)$，$N_2(t)$，…分别记在$[0, t]$发生的各型事件的个数。
  - ➢ 注意到：$N(t) = N_1(t) + N_2(t) + \cdots$

10.2.2. 泊松过程的分流的推广1

泊松过程的多类型分流:

- 命题：依前述情境，$\{N_1(t), t \geq 0\}$、$\{N_2(t), t \geq 0\}$、…分别是速率为 $\lambda p_1, \lambda p_2, \dots$ 的泊松过程。且这些泊松过程是彼此独立的。
  - 理解过程与命题5.3类似
  - 独立性的理解需用到例3.23的例子的拓展例子

10.2.3. 独立泊松过程的汇合

独立泊松过程的汇合（聚流）：

●定理：$\{N_1(t), t \geq 0\}$、$\{N_2(t), t \geq 0\}$、 ...分别是速率为$\lambda_1, \lambda_2, \ldots$的泊松过程。且这些泊松过程是彼此独立的。

●则

$$N(t) = N_1(t) + N_2(t) + \cdots$$

是速率为$\lambda_1 + \lambda_2 + \cdots$ 的泊松过程。

10.2.3. 独立泊松过程的汇合

独立泊松过程的汇合（聚流）：

●定理理解:说明两个泊松过程的聚流即可

- 条件1：$N(0) = N_1(0) + N_2(0) = 0$
- 条件2：独立和平稳增量性，易得
  - 对任何正整数n和$0 = t_0 < t_1 < \cdots < t_n$，由$\{N_1(t)\}$和$\{N_2(t)\}$的独立增量性可知随机变量

$$N_1\left(t_{j-1}, t_j\right],\ N_2(t_{j-1}, t_j],\quad j = 1,2,\dots,n$$

均相互独立，可得

$$N\left(t_{j-1}, t_j\right] = N_1\left(t_{j-1}, t_j\right] + N_2(t_{j-1}, t_j],\quad j = 1,2,\dots,n$$

相互独立。

  - 由平稳增量性，知

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

与

$$N(t_1, t_2] = N_1(t_1, t_2] + N_2(t_1, t_s],\text{同分布}$$

10.2.3. 独立泊松过程的汇合

独立泊松过程的汇合（聚流）：

- 条件3:

$$P\{N(t+h)-N(t)=1\}$$
$$=P(N(h)=1)$$
$$=P(N_1(h)=1,N_2(h)=0)+P(N_1(h)=0,N_2(h)=1)$$
$$=P(N_1(h)=1\,)P(N_2(h)=0)+P(N_1(h)=0)P(N_2(h)=1)$$
$$=\big(\lambda_1 h+o(h)\big)\big(1-\lambda_2 h+o(h)\big)+\big(\lambda_2 h+o(h)\big)\big(1-\lambda_1 h+o(h)\big)$$
$$=(\lambda_1+\lambda_2)h+o(h)$$

- 条件4:

$$P\{N(t+h)-N(t)=0\}$$
$$=P(N(h)=0)=P(N_1(h)=0\,)P(N_2(h)=0)$$
$$=1-(\lambda_1+\lambda_2)h+o(h)$$
$$P(N(h)\geq 2)=1-P(N(h)=0)-P(N(h)=1)=o(h)$$

10.2.4. 独立泊松过程间事件发生的先后

- 例子：令$\{N_1(t), t \geq 0\}$和$\{N_2(t), t \geq 0\}$是分别具有速率λ_1和λ_2的独立泊松过程。
- 问题：以S_n^1记$\{N_1(t)\}$的n个事件发生的时间，S_m^2记$\{N_2(t)\}$的m个事件发生的时间：求 $P\{S_n^1 < S_m^2\}$
  - $\{N_1(t)\}$中n个事件发生先于$\{N_2(t)\}$中m个事件发生的概率

10.2.4. 独立泊松过程间事件发生的先后

•问题：以S_n^1记$\{N_1(t)\}$的n个事件发生的时间， S_m^2记$\{N_2(t)\}$的m个事件发生的时间：求 $P\{S_n^1 < S_m^2\}$

•课本解法:

1. $n = m = 1$时， $S_n^1 = S_1^1$和$S_m^2 = S_1^2$是均值$1/\lambda_1$和$1/\lambda_2$的独立指数随机变量；可知 $P\{S_1^1 < S_1^2\} = \frac{\lambda_1}{\lambda_1+\lambda_2}$
2. $n = 2, m = 1$ 时，首先发生的初始事件是$N_1(t)$ 过程的事件，这个事件发生的概率是$P\{S_1^1 < S_1^2\} = \frac{\lambda_1}{\lambda_1+\lambda_2}$

➢ 第一个事件后，两过程均重新开始

➢ 因此，第二个事件是$N_1(t)$过程的事件的概率仍为 $P\{S_1^1 < S_1^2\} = \frac{\lambda_1}{\lambda_1+\lambda_2}$

➢ 所以要求的概率为$P\{S_2^1 < S_1^2\} = \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^2$

10.2.4. 独立泊松过程间事件发生的先后

得到如下结论：

- 由泊松过程的无记忆性，可知，独立于以前发生的所有情况，每个事件以概率$\frac{\lambda_1}{\lambda_1+\lambda_2}$是$N_1(t)$过程的事件，以概率$\frac{\lambda_2}{\lambda_1+\lambda_2}$是$N_2(t)$过程的事件

继续考虑：

3. $n=2, m=2$时，如何保证事件$\{S_2^1 < S_2^2\}$呢？
   - 要求，前3次事件，至少有2次为$N_1(t)$过程的事件
   - 即$P\{S_2^1 < S_2^2\} = \sum_{k=2}^{3} \binom{3}{k} \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k \left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{3-k}$

4. $n=2, m=3$时，如何保证事件$\{S_2^1 < S_3^2\}$呢？
   - 要求，前4次事件，至少有2次为$N_1(t)$过程的事件
   - 即$P\{S_2^1 < S_3^2\} = \sum_{k=2}^{4} \binom{4}{k} \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k \left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{4-k}$

10.2.4. 独立泊松过程间事件发生的先后

5. $n=3, m=2$时，如何保证事件$\{S_3^1 < S_2^2\}$呢？

- 要求，前4次事件，至少有3次为$N_1(t)$过程的事件
- 即$P\{S_3^1 < S_2^2\} = \sum_{k=3}^{4} \binom{4}{k} \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k \left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{4-k}$

最终结论：

- 如何保证事件$\{S_n^1 < S_m^2\}$呢？
  - 要求，前$n+m-1$次事件，至少有n次为$N_1(t)$过程的事件
  - 即$P\{S_n^1 < S_m^2\} = \sum_{k=n}^{n+m-1} \binom{n+m-1}{k} \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k \left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n+m-1-k}$
  - 经典问题: $N_1(t)$过程到达n先于$N_2(t)$过程到达m的概率，正是在抛掷一枚以概率$p=\frac{\lambda_1}{\lambda_1+\lambda_2}$出现正面的硬币时出现$n$次正面先于$m$次反面的概率。

10.2.4. 独立泊松过程间事件发生的先后

●问题：以S_n^1记$\{N_1(t)\}$的n个事件发生的时间，S_m^2记$\{N_2(t)\}$的m个事件发生的时间：求 $P\{S_n^1 < S_m^2\}$

后续思考：

➢ 本题可以理解为泊松过程汇合后的分流

$\{N_1(t)\}$和$\{N_2(t)\}$的汇合：

➢ $\{N_1(t)\}$和$\{N_2(t)\}$汇合成的泊松过程为：$\{N(t)\}$速率$\lambda_1 + \lambda_2$

$\{N(t)\}$的分流：

➢ $\{N(t)\}$各事件独立地以概率$\frac{\lambda_1}{\lambda_1+\lambda_2}$和$\frac{\lambda_2}{\lambda_1+\lambda_2}$分为I、II类事件

  ➢ 保证分流出的$\{N_1(t)\}$和$\{N_2(t)\}$分别具备速率λ_1、λ_2

➢ 事件$\{S_n^1 < S_m^2\}$自然转为了m个II类事件之前，发生n次I类事件

10.3到达时间的条件分布（课本5.3.5 部分）

10.3.0. 条件到达时间引入

$\{N(t)\}$是速率为λ的泊松过程

- 思考：若已知$N(t)=n$, n个事件到达时间$S_1,\ldots,S_n$的条件分布如何？

10.3.1. 到时间t为止仅发生1个事件的情况

- 假设已知到时间 t 为止仅发生1个事件，发生时间的条件分布：

事件发生的时间均匀分布在[0, t]上

直观理解：

- 泊松过程有平稳和独立增量，在$[0, t]$中每个相等长度的区间应该有相同的概率包含这个事件

10.3.1. 到时间t为止仅发生1个事件的情况

- 假设已知到时间 t 为止仅发生1个事件，发生时间的条件分布：
  事件发生的时间均匀分布在[0, t]上

数学推导：

- $P\{T_1 < s|N(t) = 1\} = \frac{P\{T_1<s,N(t)=1\}}{P\{N(t)=1\}}$
- $= \frac{P\{[0,s)\text{中}1\text{个事件，}[s,t]\text{中}0\text{个事件}\}}{P\{N(t)=1\}}$
- $= \frac{P\left\{[0,s)\text{中}1\text{个事件}\right\}P\left\{[s,t]\text{中}0\text{个事件}\right\}}{P\{N(t)=1\}}$
- $= \frac{\lambda s e^{-\lambda s} e^{-\lambda(t-s)}}{\lambda t e^{-\lambda t}}$
- $= \frac{s}{t}$

10.3.2. 条件到达时间进一步思考

- **思考**：给定$N(t)=n$，n个事件到达时间$S_1,\ldots,S_n$的条件分布如何？

现在再理解一下：

- 泊松过程有平稳和独立增量，在$[0,t]$中每个相等长度的区间应该有相同的概率包含任意一个事件

10.3.2. 次序统计量

•令$Y_1,\ldots,Y_n$是具有概率密度函数f的$i.i.d.$随机变量，且$Y_{(k)}$是在$Y_1,\ldots,Y_n$中第k小的值，则$Y_{(1)},\ldots,Y_{(n)}$是$Y_1,\ldots,Y_n$ 的次序统计量

•$Y_{(1)},\ldots,Y_{(n)}$的联合密度为$f_{Y_{(1)},\ldots,Y_{(n)}}(y_1,\ldots,y_n)=n!\prod_{i=1}^{n}f(y_i)$，其中$y_1<\cdots<y_n$

理解：

- 注意到，$(Y_1,\ldots,Y_n)$等于$(y_1,\ldots,y_n)$的$n!$种排列的任一种，均对应了$\left(Y_{(1)},\ldots,Y_{(n)}\right)=(y_1,\ldots,y_n)$
- 而$(Y_1,\ldots,Y_n)$等于$(y_1,\ldots,y_n)$的$n!$中排列的任意一种的概率密度都是$\prod_{i=1}^{n}f(y_i)$
- 证明不作要求，理解即可，这种用了一种近似的概率理解

•重要：若Y_i在$(0,t)$上均匀分布，则次序统计量联合密度函数：

- $f(y_1,y_2,\ldots,y_n)=\frac{n!}{t^n}$，$0<y_1<y_2<\cdots<y_n<t$

10.3.3. 条件到达时间

- 定理5.2：给定$N(t)=n$，n个到达时间$S_1,\dots,S_n$的条件分布为

$$f_{S_1,\dots S_n}(s_1,\dots,s_n|n)=\frac{n!}{t^n},0<s_1<\cdots<s_n<t$$

- 与n个$(0,t)$上独立均匀随机变量的次序统计量的分布相同

证明：

- 注意到，对$0<s_1<\cdots<s_n<t$，
事件$\{S_1=s_1,\dots,S_n=s_n,N(t)=n\}\leftrightarrow$事件$\{T_1=s_1,T_2=s_2-s_1,\dots,T_n=s_n-s_{n-1},T_{n+1}>t-s_n\}$
- 已知$T_n(n=1,2,\dots)$是$i.i.d.$的$Exp(\lambda)$（命题5.2）
- 所以，$f_{S_1,\dots S_n}(s_1,\dots,s_n|n)=\frac{f(s_1,\dots,s_n,n)}{P\{N(t)=n\}}$
- $=\dfrac{\lambda e^{-\lambda s_1}\lambda e^{-\lambda(s_2-s_1)}\dots\lambda e^{-\lambda(s_n-s_{n-1})}e^{-\lambda(t-s_n)}}{e^{-\lambda t}(\lambda t)^n/n!}$
- $=\frac{n!}{t^n}$，$0<s_1<\cdots<s_n<t$

10.3.3. 条件到达时间

进一步结论

- 重要：对一个泊松过程，在$(0,t)$中已经发生n个事件的条件下，各事件发生的时间是独立地在$(0,t)$上均匀分布的。
  - ➢ 各事件均可单独理解为独立均匀分布（详述如下）
  - ➢ 所以是，n个独立的均匀随机变量叠加，自然形成了一个次序统计量。

- 详述：已知$N(t)=n$，令$V_1,V_2,\ldots,V_n$为n个事件（无序的）的发生时间，则$V_1,\ldots,V_n$为独立同分布的均匀随机变量。也即，已知$N(t)=n$，考虑任意一个发生事件的到达时间，这个到达时间是均匀分布在$[0,t]$上的。

10.3.3. 条件到达时间

- 例5.21.：杰伦奶茶店的顾客按照速率为λ的泊松过程分布的时间来店内消费，相继的消费金额是具有均值为μ的独立随机变量，且独立于到达时间。
- 令S_i和C_i分别记第i个消费者到达的时间和消费金额。
- 主理人希望确定到时间t的所有消费金额的期望现时价值$D(t)$，考虑连续的无风险利率为α。

10.3.3. 条件到达时间

●解答：

- 注意到：$D(t)=\sum_{i=1}^{N(t)} e^{-\alpha S_i}C_i$
  - 其中$N(t)$是到时间t的顾客数
- $E[D(t)]=\sum_{n=0}^{\infty} E[D(t)|N(t)=n]\frac{e^{-\lambda t}(\lambda t)^n}{n!}=E[E[D(t)|N(t)]]$
- 应用进一步结论：$E[D(t)|N(t)=n]=E[\sum_{i=1}^{n} e^{-\alpha U_i}C_i]$

$=\sum_{i=1}^{n} E[C_i e^{-\alpha U_i}]=\sum_{i=1}^{n} E[C_i]E[e^{-\alpha U_i}]=\mu E[\sum_{i=1}^{n} e^{-\alpha U_i}]$

- $\mu E[\sum_{i=1}^{n} e^{-\alpha U_i}]=n\mu E[e^{-\alpha U}]=n\frac{\mu}{t}\int_0^t e^{-\alpha x}\,dx=n\frac{\mu}{\alpha t}(1-e^{\alpha t})$
- 所以:$E[D(t)|N(t)]=N(t)\frac{\mu}{\alpha t}(1-e^{\alpha t})$
- $E\big[E[D(t)|N(t)]\big]=\frac{\mu}{\alpha t}(1-e^{\alpha t})E[N(t)]=\frac{\lambda\mu}{\alpha}(1-e^{\alpha t})$

10.3.3. 条件到达时间

教材解法：不应用进一步结论

- 定理5.2: $E[D(t)|N(t)=n]=E\left[\sum_{i=1}^{n} e^{-\alpha U_{(i)}} C_i\right]$

$=\sum_{i=1}^{n} E[C_i e^{-\alpha U_{(i)}}]=\sum_{i=1}^{n} E[C_i]E[e^{-\alpha U_{(i)}}]=\mu E[\sum_{i=1}^{n} e^{-\alpha U_{(i)}}]$

- 注意到：$\mu E\left[\sum_{i=1}^{n} e^{-\alpha U_{(i)}}\right]=\mu E[\sum_{i=1}^{n} e^{-\alpha U_i}]$

10.3.3. 条件到达时间

- **命题5.5**：给定$S_n = t$，集合$S_1, \ldots, S_{n-1}$的分布：

$$f_{S_1, \ldots S_{n-1}}(s_1, \ldots, s_{n-1} | S_n = t) = \frac{(n-1)!}{t^{n-1}}, 0 < s_1 < \cdots < s_{n-1} < t$$

➢ 与$n-1$个独立的$(0, t)$均匀随机变量的次序统计量分布相同

10.3.3. 条件到达时间

- 命题5.5：给定$S_n=t$, 集合$S_1,\ldots,S_{n-1}$的分布：

$$f_{S_1,\ldots S_{n-1}}(s_1,\ldots,s_{n-1}|S_n=t)=\frac{(n-1)!}{t^{n-1}},0<s_1<\cdots<s_{n-1}<t$$

➢ 与$n-1$个独立的$(0,t)$均匀随机变量的次序统计量分布相同

证明：

➢ 注意到，对$0<s_1<\cdots<s_n=t$, 事件$\{S_1=s_1,\ldots,S_{n-1}=s_{n-1},S_n=t\}$等价于前$n$个间隔时间$T_1=s_1,T_2=s_2-s_1,\ldots,T_n=t-s_{n-1}$这个事件。

➢ 所以，$f_{S_1,\ldots S_{n-1}}(s_1,\ldots,s_{n-1}|t)=\frac{f(s_1,\ldots,t)}{f_{S_n}(t)}$

➢ $=\frac{\lambda e^{-\lambda s_1}\lambda e^{-\lambda(s_2-s_1)}\ldots\lambda e^{-\lambda(t-s_{n-1})}}{\frac{\lambda^n}{(n-1)!}t^{n-1}e^{-\lambda t}}$

➢ $=\frac{(n-1)!}{t^{n-1}},0<s_1<\cdots<s_{n-1}<t$

10.3.4. 泊松过程的分流的推广二

泊松过程的事件分类概率依赖于该类型事件发生时间：

- 泊松过程每个事件独立的以概率p_i分类为$i(i=1,\ldots,k)$型事件，我们假设这些概率依赖于事件发生的时间。
- 特别地，若一个事件在时刻s发生，则独立于以前发生的任何事件，它以概率$P_i(s)$被分为类型i事件，其中$\sum_{i=1}^{k} P_i(s)=1$。
- **命题5.4**：如果$N_i(t)(i=1,\ldots,k)$表示到时刻t为止，类型i事件发生的个数，那么$N_i(t)(i=1,\ldots,k)$是具有均值

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

的独立泊松随机变量。

10.3.4. 泊松过程的分流的推广二

10.3.4. 泊松过程的分流的推广二

- 例 5.20：一个人从感染某种病毒的时间，到症状出现，有相对长的潜伏期。其结果是，对于负责公众健康的部门来说，在任意给定的时间确定总体中受到感染的人数很困难。
- 这里给出这个现象的一个初级近似模型，用它可粗略估计感染人数。
- 假设感染病毒的人数服从速率λ的泊松过程。假设个体感染到出现症状的时间是分布函数G的随机变量，并假设不同个体之间的潜伏期独立。

第一步：确定考虑的事件类

- 这里考虑两类事件：1.显现疾病者和2.暂无症状者
  - 以$N_1(t)$记直到时刻t为止显现疾病的人数。
  - 以$N_2(t)$记直到时刻t为止感染病毒但暂无症状的人数。

10.3.4. 泊松过程的分流的推广二

第二步：确定随时间变化的分类概率$P_1(s)$和$P_2(s)$

- 注意到，在时刻s感染病毒的人以概率$G(t-s)$在时刻t已经出现症状，即，$P_1(s)=G(t-s)$
- 在时刻s感染病毒的人，以概率$\bar{G}(t-s)=1-G(t-s)$在时刻t仍无症状，即，$P_2(s)=\bar{G}(t-s)$

10.3.4. 泊松过程的分流的推广二

第三步：得到$N_1(t)$与$N_2(t)$的分布

- 由命题5.4，$N_1(t)$与$N_2(t)$是独立泊松随机变量，均值如下

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

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

第四步：预测无症状（潜伏期）人数

- 若λ已知，那么可以通过均值$E[N_2(t)]$估计受到感染但是在时间t无症状的人数$N_2(t)$.
- 若λ未知，怎么办呢？

10.3.4. 泊松过程的分流的推广二

第五步：根据有症状病人数，估计λ

- $N_1(t)$的值是已知的
- 当已知$N_1(t)=n_1$，可用此已知值，作为$N_1(t)$的均值$E[N_1(t)]$的估计
- 即，若直到时间t为止呈现症状的人数是n_1，则可以估计

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

- 所以λ的估计值$\hat{\lambda}$可得

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

- 利用λ的估计值$\hat{\lambda}$，可以估计$N_2(t)$

$$\widehat{N_2}(t) = \hat{\lambda} \int_0^t \bar{G}(y)dy = n_1 \int_0^t \bar{G}(y)dy \,/ \int_0^t G(y)dy$$

10.3.4. 泊松过程的分流的推广二

分布及数据计算示例

- 若G是均值μ的指数分布，那么$\bar{G}(y) = e^{-y/\mu}$, 求积分可得

$$\widehat{N_2}(t) = \frac{n_1\mu(1-e^{-t/\mu})}{t-\mu(1-e^{-t/\mu})}$$

- 若$t = 16$ 年，$\mu = 10$年，而 $n_1 = 22$ 万，那么受到感染但在16年暂无症状的人数的估计值为

$$\widehat{N_2}(t) = \frac{200\,(1-e^{-1.6})}{16-10(1-e^{-1.6})} = 21.896$$

- 潜伏期是均值10年的指数随机变量，且在传染的前16年有症状的人数为22万，则可近似估计无症状人数为21.9万。