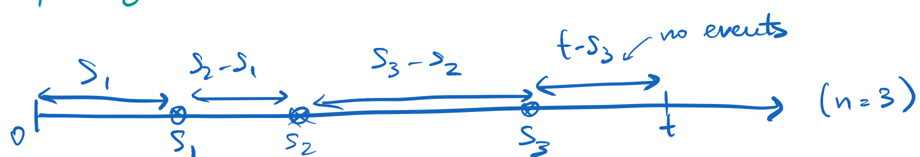
Theorem: Given that $N(t)=n$, the n arrival times $S_1, \dots, S_n$ have the same distribution as the order statistics corresponding to n independent random variables uniformly distributed on the interval $[0,t]$.

Definition: Let $Y_1, \dots, Y_n$ be r.v.'s. Their order statistics $Y_{(1)}, Y_{(2)}, \dots, Y_{(n)}$ are the values of $Y_1, \dots, Y_n$ in increasing order.

Last time we showed that if $Y_1, \dots, Y_n$ are i.i.d. Uniform$([0,t])$, the joint pdf of $Y_{(1)}, \dots, Y_{(n)}$ is $f(y_1, \dots, y_n) = \frac{n!}{t^n}$ for any $0 \le y_1 \le \dots \le y_n \le t$.

Proof of Theorem: We will show that the joint density of $S_1, \dots, S_n$ given $N(t)=n$ is $f(s_1, \dots, s_n | N(t)=n) = \frac{n!}{t^n}$.



The event $\{S_1 = s_1, \dots, S_n = s_n, N(t)=n\}$ for $s_1 < \dots < s_n \le t$ is the same as the event that the first $n+1$ interarrival times $T_1, \dots, T_{n+1}$ satisfy $T_1 = s_1$, $T_2 = s_2 - s_1, \dots, T_n = s_n - s_{n-1}$, $T_{n+1} > t - s_n$.

Thus, $f(s_1, \dots, s_n | N(t)=n) = \dfrac{f(s_1, \dots, s_n, T_{n+1} > t - s_n)}{P(N(t)=n)}$

$$= \frac{f(T_1 = s_1, T_2 = s_2 - s_1, \dots, T_n = s_n - s_{n-1}, T_{n+1} > t - s_n)}{P(N(t)=n)}$$

$$= \frac{\lambda e^{-\lambda s_1} \lambda e^{-\lambda(s_2 - s_1)} \cdots \lambda e^{-\lambda(s_n - s_{n-1})} e^{-\lambda(t - s_n)}}{(\lambda t)^n e^{-\lambda t}/n!} = \frac{\lambda^n e^{-\lambda t} n!}{\lambda^n t^n e^{-\lambda t}} = \frac{n!}{t^n}$$

Now suppose that we label events of a Poisson process into k possible types s.t. at each time s, there is a distribution $P_i(s)$ $(i=1,\dots,k)$ s.t. P(event occurring at time s is type i) $= P_i(s)$

(the type selection is indep. of what previously occurred)

<u>Proposition</u>: Let $N_i(t) = \#$ Type i events by time t. Then, $N_1(t),\dots,N_k(t)$ are independent Poisson r.v.'s with means $\mathbb{E}[N_i(t)] = \lambda\int_0^t P_i(s)\,ds$.

<u>Remark</u>: We recover part of the Poisson thinning proposition: if $p_i = p$,

then $\mathbb{E}[N_i(t)] = \lambda\int_0^t p\,ds = p\cdot\lambda t$.

<u>Proof</u>: $P(N_1(t)=n_1, \dots, N_k(t)=n_k)$

$$= P\left(N_1(t)=n_1,\dots,N_k(t)=n_k \,\middle|\, N(t)=\sum_{i=1}^k n_i\right) P\left(N(t)=\sum_{i=1}^k n_i\right)$$

Conditional on n events, arrival times are indep. and uniform on $[0,t]$, so $P(\text{event is type } i) = \int_0^t P(\text{event is type } i \mid \text{event occurs at } s)\frac{1}{t}ds$

$$= \frac{1}{t}\int_0^t P_i(s)\,ds =: \bar{p}_i$$

prob of arrival at s

Each of the n events has prob $\bar{p}_i$ of being type i and they are independent.

multinomial distribution

$$\Rightarrow P\left(N_1(t)=n_1,\dots,N_k(t)=n_k \,\middle|\, N(t)=\sum n_i\right) = \frac{n!}{n_1!\cdots n_k!}\bar{p}_1^{n_1}\cdots\bar{p}_k^{n_k}$$

$$\Rightarrow P(N_1(t)=n_1,\dots,N_k(t)=n_k) = \frac{\cancel{n!}}{n_1!\cdots n_k!}\bar{p}_1^{n_1}\cdots\bar{p}_k^{n_k}\frac{(\lambda t)e^{-\lambda t}}{\cancel{n!}}$$

$$= \left(\frac{(\lambda t\bar{p}_1)^{n_1}}{n_1!}e^{-\lambda t\bar{p}_1}\right)\cdots\left(\frac{(\lambda t\bar{p}_k)^{n_k}e^{-\lambda t\bar{p}_k}}{n_k!}\right)$$

(we used that $n_1+\dots+n_k = n$ and $\sum\bar{p}_i = 1$)

$\Rightarrow N_1(t),\dots,N_k(t)$ are independent and $N_i(t)\sim \text{Poisson}\left(\lambda\int_0^t P_i(s)\,ds\right)$

<u>Example</u>: let's get back to our example: individuals independently get infected following a λ-Poisson process, s.t. symptoms appear after time T since infection where T has CDF G.

$N_1(t)$ = # individuals infected with symptoms

$N_2(t)$ = # individuals infected with no symptoms.

$N(t) = N_1(t) + N_2(t)$ = total # of individuals infected.

<u>Goal</u>: Find $\mathbb{E}[N_2(t)]$, which is an estimate of $N_2(t)$.

<u>Solution</u>: Fix t and use previous proposition with $k=2$.

$$P(N_1(t) = n_1, N_2(t) = n_2) = P(N_1(t) = n_1, N_2(t) = n_2 \mid N(t) = n_1 + n_2)\, P(N(t) = n_1 + n_2).$$

For $0 \le s \le t$, an individual infected at time s has symptoms by time t w.p. $\underbrace{G(t-s)}_{P_1(s)}$ and has no symptoms by time t w.p. $\underbrace{1-G(t-s)}_{P_2(s)}$.

$$\Rightarrow \quad N_1(t) \sim \text{Poisson}\left(\lambda \int_0^t G(t-s)\,ds\right)$$
$$N_2(t) \sim \text{Poisson}\left(\lambda \int_0^t 1 - G(t-s)\,ds\right)$$

$$\Rightarrow \quad \mathbb{E}[N_1(t)] = \lambda \int_0^t G(t-s)\,ds$$
$$\mathbb{E}[N_2(t)] = \underbrace{\lambda}_{\text{need to estimate.}} \int_0^t \overbrace{(1-G(t-s))}^{\text{known}}\,ds$$

To estimate λ, since we know $N_1(t)$, say $N_1(t) = n_1$ (number of people infected with symptoms by time t), then use

$$n_1 \approx \mathbb{E}[N_1(t)] = \lambda \int_0^t G(t-s)\,ds$$ to get an estimate

$$\hat{\lambda} = \frac{n_1}{\int_0^t G(t-s)\,ds} \Rightarrow \hat{n}_2 = \hat{\lambda} \int_0^t (1-G(t-s))\,ds = \frac{n_1 \int_0^t (1-G(t-s))\,ds}{\int_0^t G(t-s)\,ds}.$$

e.g., $t = 16$ days, $G \sim \text{Exp}\left(\frac{1}{10\text{ days}}\right)$, $n_1 = 2000$, then $\hat{n}_2 = 2000 \dfrac{\int_0^{16} e^{-\frac{(16-s)}{10}}\,ds}{\int_0^{16} (1-e^{-\frac{(16-s)}{10}})\,ds}$

$$= 2000 \cdot \frac{10\left(1 - e^{-16/10}\right)}{16 - 10\left(1 - e^{-16/10}\right)} , \quad \hat{n}_2 \approx 114 .$$