Lecture 20 - 10/7

Sunday, October 09, 2022 11:51 AM

Review:

① exponential distribution.

T has exponential distribution with parameter $\lambda > 0$, $T \sim exp(\lambda)$

$$f_T(t) = \begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & o.w \end{cases}$$

- cdf: $P(T \leq t) = F_T(t) = \begin{cases} 1 - e^{-\lambda t} & t \geq 0 \\ 0 & o.w. \end{cases}$

- complimentary cdf: $P(T > t) = F_T^c(t) = \begin{cases} e^{-\lambda t} & t \geq 0 \\ 0 & o.w. \end{cases}$

- Mean & Variance: $E[T] = \frac{1}{\lambda}$, $Var(T) = \frac{1}{\lambda^2}$

- Memoryless property:

$$P(T > t+s \mid T > s) = P(T > t)$$

② relation between geometric distribution and exponential distribution:

Let X_h denote a geometric random variable with parameter $p_h = \lambda h$ for all $h > 0$.

Let $T_h = h X_h$. We have

$$P(T_h > t) \longrightarrow P(T > t) \quad \text{as } h \longrightarrow 0$$

where T is a random variable with parameter $\lambda > 0$.

Today:

① Important remark about continuous random variable

② Poisson process as a limit of time-scaled Bernoulli process

③ Poisson process

④ Erlang distribution

① Important remark about continuous random variable:

Recall that a continuous random variable X, is defined by the relation between cdf and pdf.

$$P(X \leq a) = F_X(a) = \int_{-\infty}^{a} f_X(u)\, du$$

As we discussed before

$$P(a < X \leq b) = F_X(b) - F_X(a) = \int_b^a f_X(u)\, du$$

$$P(a < X \leq b) = F_X(b) - F_X(a) = \int_b^a f_X(u)\, du$$

Hence,

$$P(X > a) = F_X^c(a) = 1 - F_X(a) = \int_a^\infty f_X(u)\, du$$

and in generall, for any $A \subseteq \mathbb{R}$:

$$P(X \in A) = \int_A f_X(u)\, du$$

---

② Poisson process as a limit of time-scaled Bernoulli process.

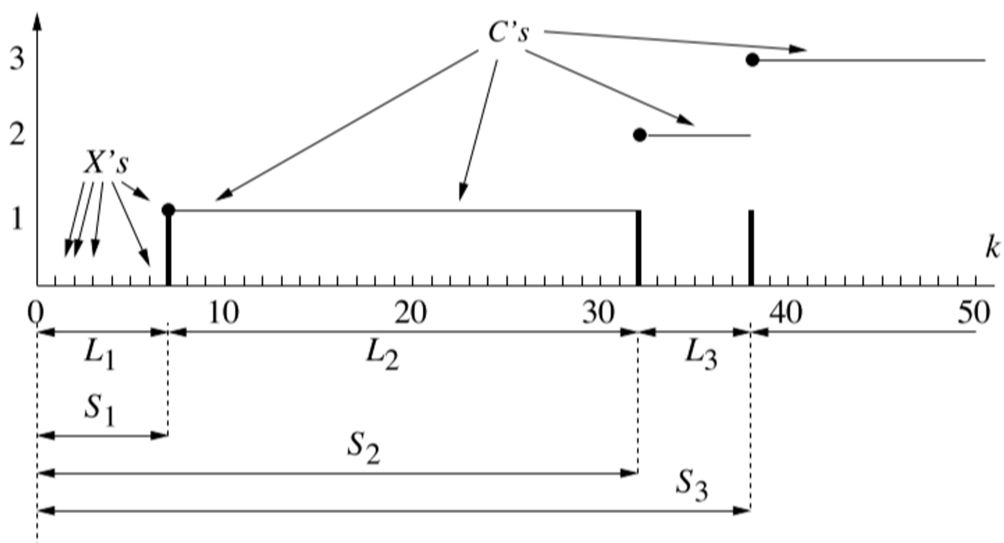
Recall that a Bernoulli process is given by $(X_1, X_2, \dots)$ where for each n, X_n is a Bernoulli random variable with parameter p. Also, recall that it is a random process, i.e.,

(i) for each $n \in \{1,2,3,\dots\}$. X_n is a random variable. More specifically, $X_n \sim Ber(p)$

(ii) for each $\omega \in \Omega$, $(X_1(\omega), X_2(\omega), X_3(\omega), \dots)$ is a sequence of zeros and ones.

Recall that there are three other random variables that are related to Bernoulli process.

- L_n = number of trials between $(n-1)$'th success and n'th success. Recall that $L_i \sim Geo(p)$
- S_r = number of trials till the r'th success. Recall that $S_r \sim NBi(r,p)$
- C_n = number of successes in the first n trials. Recall that $C_n \sim Bi(n,p)$



Consider the following setting: customer are entering bank, we have a clock that ticks every h seconds, and define

$$X_n^{(h)} = \begin{cases} 1 & \text{if at least one customer entered the bank during time } ((n-1)h, nh] \\ 0 & \text{otherwise} \end{cases}$$

Let us make the following assumptions:

$\lfloor 0$ otherwise

Let us make the following assumptions:

(*) For any $h>0$, $(X_1^{(h)}, X_2^{(h)}, \dots)$ is a Bernoulli process with parameter p_h.

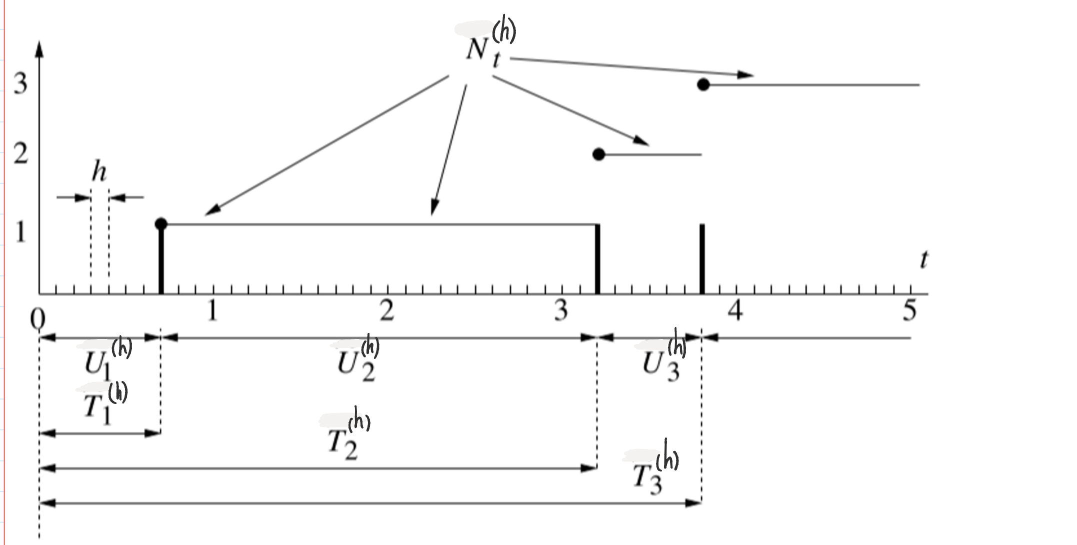
(**) For small $h>0$, number of customers that can enter the bank during interval $((n-1)h, nh]$ is at most 1.

(***) $p_h = \lambda h$ for all small $h>0$. This can be justified the same way as we did in the previous session, i.e., by showing that $p_{2h} \approx p_h$ for all small $h>0$. and using the same argumet to show $p_{ah} = a p_h$.

Assume that $h>0$ is small. Let us consider the timed counterpart of $L_n^{(h)}, S_r^{(h)}$ and $C_n^{(h)}$:

- $U_n^{(h)}$ = time it takes till the n'th arrival. Notice that $U_n^{(h)} = h\,L_n^{(h)}$.
- $T_r^{(h)}$ = time it takes till r'th arrival. Notice that $T_r^{(h)} = h\,S_r^{(h)}$.
- $N_t^{(h)}$ = number of times $X_i^{(h)} = 1$ for $i < \frac{t}{h}$. Notice that $N_t^{(h)} = C_{\lfloor t/h \rfloor}$.

Intuitively speaking, $N_t^{(h)}$ is very close the number arrivals till time t, i.e., we are just ignoring arrivals during time interval $\left(h\lfloor \frac{t}{h} \rfloor, th\right]$



As $h \to 0$, we have:

- distribution of $U_n^{(h)}$ converges to exponential distribution with parameter λ.

$$P(U_n^{(h)} > t) = P\left(L_n^{(h)} > \frac{t}{h}\right) = (1-\lambda h)^{\lfloor \frac{t}{h} \rfloor} \longrightarrow e^{-\lambda t} \quad \text{as } h \to 0.$$

- distribution of $N_t^{(h)}$ converges to poisson distribution with parameter λt. Notice that $N_t^{(h)} \sim Bi\left(\lfloor \frac{t}{h} \rfloor, p_h\right)$ and $p_h \cdot \lfloor \frac{t}{h} \rfloor = \lambda h \cdot \lfloor \frac{t}{h} \rfloor \approx \lambda t$. Hence, by poisson approximation

$$P(N_t^{(h)} = k) \longrightarrow e^{-\lambda t} \frac{(\lambda t)^k}{k!}$$

Hence, the limit of time-scaled Bernoulli process is a continuous process which is called poisson process.

---

③ Poisson process.

Def: Let $\lambda > 0$. A poisson process with rate λ is a counting process $N = (N_t ; t \geq 0)$ such that

N.1: N has independent increament: for any $t_0 \leq t_1 \leq \cdots \leq t_n$, the increaments

$$N_{t_2} - N_{t_1}, \; N_{t_3} - N_{t_2}, \; \ldots, \; N_{t_n} - N_{t_{n-1}}$$

are independents.

N.2: The increament $N_t - N_s$ has poisson distribution with parameter $\lambda(t-s)$,

$$P(N_t - N_s = k) = e^{-\lambda(t-s)} \frac{(\lambda(t-s))^k}{k!}$$

N.3: $N_0 = 0$.

Interpretation:

- N is a random process, this means that
  - (i) if we fix $t > 0$, N_t is a random variable, i.e., $N_t : \Omega \to \{0,1,2,3,\ldots\}$
    In particular N_t is a discrete random variable, and more specifically $N_t \sim Poi(\lambda t)$
  - (ii) if we fix $\omega \in \Omega$, $N(\omega)$ is a sample path, i.e., it is a function of time:
    $N(\omega) : [0,+\infty) \to \{0,1,2,3,\ldots\}$, i.e., $N_t(\omega)$ is a non-negative integer.
- N is a counting process. Intuitively speaking, this means that N counts number of events that has happenned over time, i.e.,

  N_t: number of occurance of an event over time interval $[0,t]$

  $N_t - N_s$: number of occurance of an event over time interval $(s,t]$
- One example to visualize the above properties is the customers that appear over time in a bank. Let N denote this process and assume N is a poisson process:
  - (i) N_t: number of customers that showed up during time interval $[0,t]$
  - (ii) $N_t - N_s$: number of customers that showed up during time interval $(s,t]$
  - (iii) independent increament means number of customers that showed up during intervals $(t_1, t_2]$ and $(t_3, t_4]$ for any $0 \leq t_1 \leq t_2 \leq t_3 \leq t_4$ are independent.
  - (iv) $N_0 = 0$ means at the point of openning the bank, there is no customer inside the bank.

Related variables:

Associated with the poisson process $N = (N_t ; t \geq 0)$, there are important random variables:

① For any $n \in \{0,1,2,\ldots\}$, we define

U_n = the time between $(n-1)$'th occurance & n'th occurance,

i.e., inter-arrival time between the $(n-1)$'th customer and n'th customer.

② For any $n \in \{0,1,2,\ldots\}$, we define

T_n = time of n'th occurance

i.e., time at which n'th customer entered the bank.

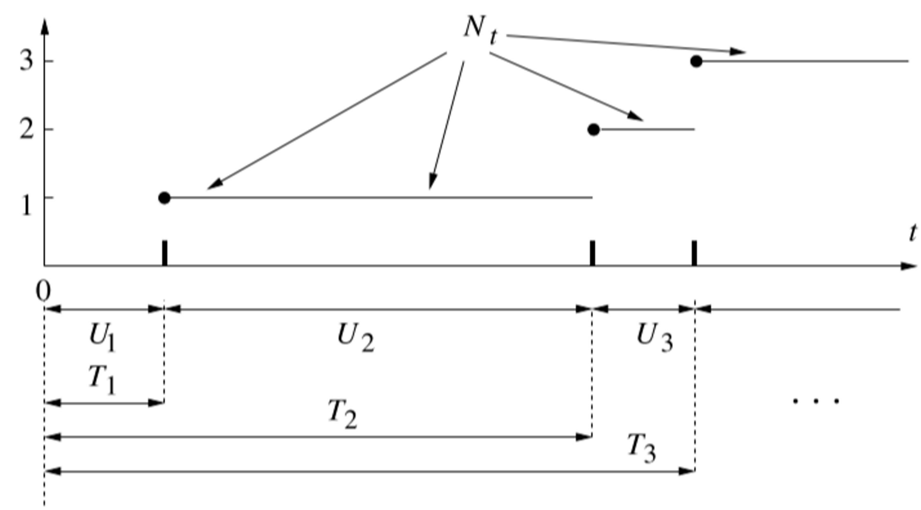


Figure 3.9: A sample path of a Poisson process.

The random variables associated with the poisson process have the following relations:

$$N_t = \sum_{n=1}^{\infty} I_{\{T_n \leq t\}}$$

$$T_n = \min\{t : N_t = n\} = \min\{t : N_t \geq n\}$$

$$T_n = U_1 + U_2 + \cdots + U_n$$

where

$$I\{T_n \leq t\} = \begin{cases} 1 & \text{if } T_n \leq t \\ 0 & \text{if } T_n > t \end{cases}$$

$$I\{T_n \leq t\} = \begin{cases} & \\ 0 & \text{if } T_n > t \end{cases}$$

Distribution of U_n:

Notice that by the argument of previous section, we expect U_n to be a exponentially distributed random variable with parameter λ. We can prove the same property by deffinition of Poisson process:

$$P(U_n > t) = P(\{\text{after } (n-1)\text{th arrival, there is no arrival over an interval of length } t\})$$
$$= e^{-\lambda t} \frac{(\lambda t)^0}{0!} = e^{-\lambda t},$$

where we used the fact that number of arrivals over any interval of length t is a poisson random variable with parameter λt.

---

④ Erlang distribution

Recall that $T_n = U_1 + U_2 + \dots + U_n$ is the sum of n exponential random variables. Notice that

$$\{T_n > t\} = \{\text{number of arrivals in } [0,t] \text{ is less than } n\} = \{N_t < n\} = \{N_t \leq n-1\}$$

Hence,

$$F^c_{T_n}(t) = \mathbb{P}(T_n > t) = \mathbb{P}(N_t \leq n-1)$$
$$= \sum_{k=0}^{n-1} e^{-\lambda t} \frac{(\lambda t)^k}{k!}$$

$$f_{T_n}(t) = -\frac{dF^c_{T_n}(t)}{dt} = \sum_{k=0}^{n-1} \left(\lambda e^{-\lambda t} \frac{(\lambda t)^k}{k!} - e^{-\lambda t} \frac{(\lambda t)^{k-1}}{k!} \cdot k\lambda \right)$$

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

$$= \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}$$

Def: We say a continuous random variable T has erlang distribution with parameters $r \in \{1,2,3,\dots\}$ and $p \in [0,1]$ if

$$f_T(t) = \begin{cases} \frac{\lambda e^{-\lambda t} (\lambda t)^{r-1}}{(r-1)!}, & t \geq 0 \\ \end{cases}$$

$$f_T(t) = \begin{cases} \dfrac{\ \ }{(r-1)!} & \\ 0 & , \ t<0 \end{cases}$$

CDF of T:

$$F_T(t) = \begin{cases} 1 - \sum_{k=0}^{r-1} e^{-\lambda t} \dfrac{(\lambda t)^k}{k!} & , \ t \geq 0 \\ 0 & , \ t<0 \end{cases}$$

Mean & variance:

$$E[T] = \int_{-\infty}^{+\infty} u\, f_T(u)\, du = \int_{-\infty}^{+\infty} u \cdot \frac{\lambda e^{-\lambda u} (\lambda u)^{r-1}}{(r-1)!} du = \frac{r}{\lambda} \int_{-\infty}^{+\infty} \frac{\lambda e^{-\lambda t} (\lambda u)^r}{r!} du = 1$$

since $\dfrac{\lambda e^{-\lambda t} (\lambda u)^r}{r!}$ is the pdf of an Erlang distributed random variable with parameters $(r+1, p)$

$$E[T^2] = \int_{-\infty}^{+\infty} u^2 \cdot \frac{\lambda e^{-\lambda u} (\lambda u)^{r-1}}{(r-1)!} du = \frac{(r+1)r}{\lambda^2} \int_{-\infty}^{+\infty} \frac{\lambda e^{-\lambda u} (\lambda u)^{r+1}}{(r+1)!} du = \frac{r(r+1)}{\lambda^2}$$

by a similar reasoning as before.

$$Var(T) = E[T^2] - (E[T])^2 = \frac{r(r+1)}{\lambda^2} - \frac{r^2}{\lambda^2} = \frac{r}{\lambda^2}$$