هوالحق

Stochastic Processes

Poisson Process

Note 3

مهدی جعفری سیاوشانی

دانشگاه صنعتی شریف

* مطالب این درس گفتار:

- معرفی فرایندهای arrival و renewal
- فرایند متعارفی پواسن
  - خاصیت بدون حافظه بودن یک متغیر تصادفی
  - خواص Stationary increment و independent increment توزیع پواسن
  - تابع چگالی احتمال S_n و تابع چگالی احتمال مشترک S_1 و ... و S_n
  - تابع جرمی احتمال متغیر تصادفی $N(t)$
  - چند نکته درباره توزیع پواسن
  - تعاریف دیگری برای فرایند پواسن
  - ترکیب کردن و انشعاب کردن فرایند پواسن

* فرایند تصادفی پواسن (Poisson Process)

فرایند تصادفی پواسن یک فرایند ساده و با کاربرد زیاد برای مدل کردن زمان هایی است که پدیده‌ای وارد یک سیستم می‌شود. این فرایند را می‌توان به عنوان نسخه پیوسته شده فرایند برنولی در نظر گرفت.

در فصل قبل اشاره کردیم که فرایند برنولی دنباله‌ای از متغیرهای تصادفی $\{Y_n, n\geq 1\}$ است به طوریکه $Y_i \in \{0,1\}$ و Y_i ها iid هستند. در این فرایند ورودها (arrivals) در زمان‌های گسسته‌ای اتفاق می‌افتند. قبلا اشاره کردیم که این فرایند را می‌توانیم توسط دنباله‌ای از متغیرهای تصادفی که فواصل هر دو ورود متوالی را مشخص می‌کنند، توصیف کنیم. این متغیرهای تصادفی جدید نیز iid هستند و توزیع هندسی دارند.

در فرایند پواسن، ورودها در هر زمان مثبتی ممکن است اتفاق بیافتند و احتمال ورود در یک لحظه خاص 0 است. به همین دلیل برای توصیف این فرایند راحت‌تر هستیم از دنباله متغیرهای تصادفی $X_1, X_2, \dots$ که فواصل بین ورودهای متوالی را مشخص می‌کنند، استفاده کنیم.

- Arrival Processes (فرایندهای ورود):

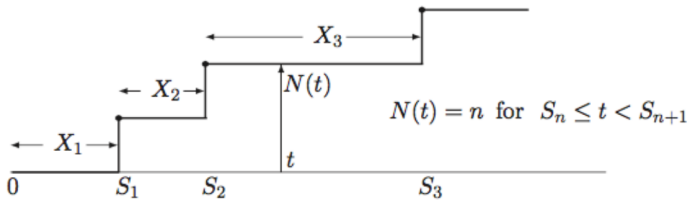
یک فرایند arrival دنباله‌ای افزایشی از متغیرهای تصادفی

$$0 < S_1 < S_2 < \cdots$$

است، که $S_i < S_{i+1}$ به معنی این است که $S_{i+1} - S_i$ یک متغیر تصادفی مثبت است. اگر تعریف کنیم $Z = S_{i+1} - S_i$ آنگاه $F_X(0) = 0$.

متغیرهای S_i را زمان ورود می‌نامند، که نشان دهنده زمان‌های متوالی ورود پدیده‌ای به سیستم هستند. توجه کنید که فرایند در زمان صفر شروع می‌شود و دو ورود همزمان اتفاق نمی‌افتند (گسستگی به حالتی که ممکن است چند ورود همزمان داشته باشیم مختص نیست. برای اینکار می‌توانیم به هر ورود یک متغیر تصادفی مثبت صحیح نسبت دهیم). برای اینکه این فرایند را به صورت کامل توسط دنباله $S_1, S_2, \dots$ مشخص کنیم باید توزیع مشترک $S_1, \dots, S_n$ را برای هر $n \geq 1$ داشته باشیم.

همانطور که در شکل بعد مشاهده می‌کنید، هر فرایند arrival $\{S_n; n\geq 1\}$ را می‌توان توسط دو فرایند دیگر نیز توصیف کرد. گزینه اول استفاده از دنباله زمان‌های بین هر دو ورود متوالی است: $\{X_i; i\geq 1\}$.



A sample function of an arrival process and its arrival epochs $\{S_1, S_2, \ldots\}$, its interarrival times $\{X_1, X_2, \ldots\}$, and its counting process $\{N(t);\ t > 0\}$

X_i ها متغیرهای تصادفی مثبتی هستند که بر اساس S_n ها به صورت زیر نوشته می‌شوند:

$$X_1 = S_1, \quad X_i = S_i - S_{i-1} \quad \forall\ i > 1$$

همچنین S_n توسط $\{X_i;\ i \geq 1\}$ به صورت زیر مشخص می‌شود:

$$S_n = \sum_{i=1}^{n} X_i$$

بنابراین توزیع مشترک $X_1, \ldots, X_n$ برای هر $n \geq 1$ هم قاعدتاً برای توصیف یک فرایند arrival کافی است.

در بسیاری از کاربردها زمان‌های بین ورودها iid هستند. بنابراین در اینگونه موارد یا فقط توزیع مشترک X_i ها از S_i ها ساده‌تر است.

Counting process:

روش دیگر استفاده از فرایند شمارشی $\{N(t);\ t > 0\}$ است که در شکل نیز نمایش داده شده است که مجموعه ناشمارایی از متغیرهای تصادفی است به طوریکه متغیر تصادفی $N(t)$ (برای هر $t > 0$) مجموع تعداد دریافت‌ها در بازه زمانی $(0, t]$ است. برای این فرایند داریم:

$$N(\tau) \geq N(t), \quad \forall\ \tau \geq t > 0$$

برای هر $n \geq 1$ و $t > 0$، زمان ورود nام یعنی S_n و متغیر شمارشی $N(t)$ به صورت زیر به هم مربوط هستند:

$$\{S_n \leq t\} = \{N(t) \geq n\}$$

برای دیدن این نکته توجه کنید که واقعه $\{S_n \leq t\}$ به این معنی است nامین ورود در زمان $\tau \leq t$ رخ داده است. پس با توجه به تعاریف داریم: $N(\tau) = n$. پس

$$\{S_n \leq t\} \subseteq \{N(t) \geq n\}$$

از طرف دیگر واقعه $\{N(t)\geq n\}$ به این معنی است که $N(t)=m$ برای یک $m\geq n$. پس $S_m \leq t$

و بنابراین نتیجه می‌گیریم:

$$\{N(t)\geq n\} \subseteq \{S_n \leq t\}$$

در نتیجه برابری این دو واقعه اثبات می‌شود.

این رابطه به ما اجازه می‌دهد تابع توزیع تجمعی (CDF) مشترک S_n ها و $N(t)$ را بتوانیم برحسب هم بیان کنیم.

به طور خلاصه یک فرایند arrival را می‌توانیم با توجه به مشترک زمان‌های ورود (S_n ها) یا زمان‌های بین ورودهای متوالی (X_i ها) یا متغیرهای شمارش ($N(t)$ ها) مشخص کنیم.

- تعریف و خواص فرایند پواسن:

یکی از ساده‌ترین فرایندهای arrival فرایندهای renewal هستند که به صورت زیر تعریف می‌شوند.

- فرایندهای renewal:

یک فرایند renewal یک فرایند arrival است که زمان‌های بین دو ورود متوالی متغیرهای مثبت iid باشند.

- فرایند پواسن (Poisson):

فرایند پواسن یک فرایند renewal است که زمان‌های بین ورودهای متوالی یک متغیر تصادفی با توزیع نمایی است. یعنی برای یک عدد حقیقی $\lambda>0$، هر X_i تابع چگالی احتمال زیر را داشته باشد:

$$f_X(x)=\lambda e^{-\lambda x}, \quad \forall x\geq 0$$

به پارامتر λ نرخ این فرایند گویند. بعداً نشان خواهیم داد که برای هر بازه به طول t، λt امید ریاضی تعداد ورودها در این بازه است و به همین دلیل اگر گذاری نرخ برای λ معنی پیدا می‌کند.

- خاصیت بدون حافظه بودن (Memoryless Property):

یکی از نکاتی که فرایند پواسن را در میان فرایندهای renewal ویژه می‌سازد، خاصیت بدون حافظه بودن توزیع نمایی است.

تعریف (بدون حافظه بودن یک متغیر تصادفی): متغیر تصادفی X خاصیت بدون حافظه بودن را دارد اگر X یک متغیر تصادفی مثبت باشد که:

$$\mathbb{P}[X>t+s] = \mathbb{P}[X>t].\mathbb{P}[X>s] \quad \forall s,t\geq 0$$

برای متغیر تصادفی نمایی داریم:

$$F_X(x) = \mathbb{P}[X \leq x] = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

در نتیجه: $\mathbb{E}[X] = \frac{1}{\lambda}$, $Var[X] = \frac{1}{\lambda^2}$

پس داریم: $\mathbb{P}[X > t] = e^{-\lambda t}$ که در تعریف بدون حافظه بودن صدق می‌کند. از طرف دیگر تنها توزیعی که خاصیت بدون حافظه بودن را دارد توزیع نمایی است.

برای مشاهده این نکته فرض کنید $h(x) = \ln(\mathbb{P}[X > x])$ و توجه کنید که $h(x)$ یک تابع غیر صعودی است. همچنین شرط بدون حافظه بودن برحسب $h(x)$ به صورت زیر در می‌آید:

$$h(s+t) = h(s) + h(t)$$

می‌توان بررسی کرد که فقط تابع خطی این دو خاصیت را دارد. پس $\mathbb{P}[X > x]$ باید نسبت به x نمایی باشد.

از آنجایی که یک متغیر تصادفی بدون حافظه باید توزیع نمایی داشته باشد پس برای آن $\lambda > 0$ وجود دارد که $\mathbb{P}[X > t] = e^{-\lambda t} > 0$. در این صورت تعریف بدون حافظه بودن را می‌توانیم به شکل زیر هم بیان کنیم:

$$\mathbb{P}[X > t + x \mid X > t] = \mathbb{P}[X > x]$$

این رابطه مفهوم بدون حافظه بودن متغیر تصادفی X را واضح‌تر بیان می‌کند. اگر X زمان انتظار برای وقوع رخدادی باشد (مثلا طول عمر وسیله‌ای یا زمان انتظار بیرون رستوران برای رسیدن نوبت ما) در این صورت اگر مدت زمان t صبر کرده باشیم و سوال بپرسیم که با چه احتمالی باید حداقل مقدار x بیشتر صبر کنیم، این احتمال برابر است با احتمال اینکه از ابتدا بخواهیم حداقل به میزان x منتظر شویم. یعنی زمان باقیمانده به یاد نمی‌آورد که تا قبلا چقدر منتظر مانده‌ایم.

نکته: اگر تعریف بدون حافظه بودن فقط به زمان‌های گسسته با مقدار صحیح (و مثبت) محدود شود، در این حالت به جای توزیع نمایی، توزیع هندسی بدون حافظه خواهد بود.

با این توصیف، فرایند برنولی (با توزیع هندسی بین زمان های ورود متوالی) مشابه نسخه زمان-گسسته فرایند پواسن (با توزیع نمایی بین زمان های ورود متوالی) است.

با استفاده از خاصیت بدون حافظه بودن متغیر نمایی می توانیم قضیه زیر را بیان کنیم.

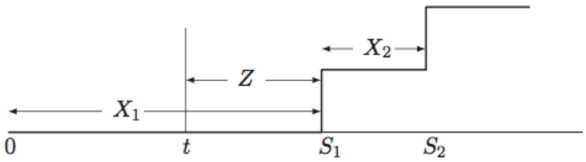
- قضیه: برای یک فرایند تصادفی پواسن با نرخ λ، و برای هر $t>0$، طول بازه از t تا اولین ورود بعد از زمان t یک متغیر تصادفی مثبت Z با تابع توزیع تجمعی (CDF) $1-e^{-\lambda z}$ برای $z\geq 0$ است. این متغیر تصادفی مستقل از

① $N(t)$ و زمان ورود قبل از t و

② مستقل از متغیر های تصادفی $\{N(\tau);\ \tau\leq t\}$ است.

ایده اصلی این قضیه این است که Z، مشروط به زمان τ آخرین ورود قبل از t، به وضوح زمان باقیمانده تا ورود بعدی است. از آنجایی که زمان های بین ورود نمایی هستند، چه کلی و چه با شروع از لحظه τ و بنابراین بدون حافظه هستند، پس Z مستقل از ورود لحظه $\tau\leq t$ و ورود های قبلی است.

اثبات زیر با جزییات بیشتری این مسئله را نشان می دهد.



For arbitrary fixed $t > 0$, consider the event $\{N(t) = 0\}$. Conditional on this event, Z is the distance from t to S_1, i.e., $Z = X_1 - t$.

اثبات: فرض کنید که Z فاصله بین t تا اولین ورود بعد از t باشد. ابتدا احتمال $\{Z>z\}$ را مشروط به واقعه $\{N(t)=0\}$ می کنیم. با شرط $N(t)=0$ داریم $X_1>t$ و $Z=X_1-t$. بنابراین:

$$\mathbb{P}[Z>z \mid N(t)=0] = \mathbb{P}[X_1>z+t \mid N(t)=0]$$
$$= \mathbb{P}[X_1>z+t \mid X_1>t]$$
$$= \mathbb{P}[X_1>z] = e^{-\lambda z}$$

بدون حافظه بودن X_1

$$\boxed{\begin{aligned} &\{S_n>t\}=\{N(t)<n\} \\ &\Rightarrow \{N(t)<1\}=\{S_1>t\} \\ &\Rightarrow \{N(t)=0\}=\{X_1>t\} \end{aligned}}$$

حال با مشروط کردن به وقایع $N(t)=n$ (برای هر $n\geq 1$) و $S_n=\tau$ (برای هر $\tau\leq t$)،

اولین ورود پس از زمان t اولین ورود پس از ورود لحظه S_n است. پس واقعه $\{Z=z\}$ معادل

است با واقعه $\{X_{n+1}=z+(t-\tau)\}$. بنابراین:

$$\mathbb{P}[Z>z \mid N(t)=n,\ S_n=\tau] = \mathbb{P}[X_{n+1}>z+t-\tau \mid N(t)=n,\ S_n=\tau]$$

$$\overset{a}{=} \mathbb{P}[X_{n+1}>z+t-\tau \mid X_{n+1}>t-\tau,\ S_n=\tau]$$

$$\overset{b}{=} \mathbb{P}[X_{n+1}>z+t-\tau \mid X_{n+1}>t-\tau]$$

$$\overset{c}{=} \mathbb{P}[X_{n+1}>z] = e^{-\lambda z}$$

رابطه a درست است چون با شرط $S_n=\tau\leq t$ داریم: $\{N(t)=n\}=\{X_{n+1}>t-\tau\}$

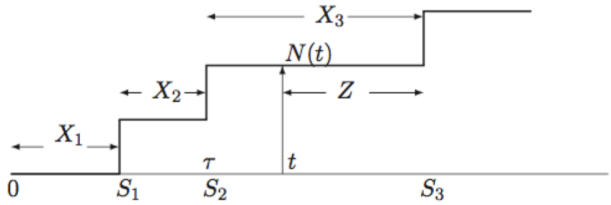
(شکل زیر را نیز نگاه کنید).

رابطه b برقرار است زیرا X_{n+1} از S_n مستقل است (X_{n+1} از همه $X_1,\dots,X_n$ مستقل است).

در نهایت رابطه c از بدون حافظه بودن X_{n+1} نتیجه می‌شود.

(توجه کنید که در روابط بالا چون $\mathbb{P}[Z>z\mid B_i]=f(z)$ و $\sum_i \mathbb{P}[B_i]=1$ نتیجه می‌گیریم

$\mathbb{P}[Z>z]=f(z)$).



Given $N(t)=2$, and $S_2=\tau$, X_3 is equal to $Z+(t-\tau)$. Also, the event $\{N(t)=2,\ S_2=\tau\}$ is the same as the event $\{S_2=\tau,\ X_3>t-\tau\}$.

در روابط بالا اگر به جای S_n همه $S_1,\dots,S_n$ ها را قرار دهیم باز استدلال بالا برقرار است. یعنی داریم:

$$\mathbb{P}[Z>z \mid N(t)=n,\ S_n=\tau_n,\dots,S_1=\tau_1] = e^{-\lambda z}$$

برای هر $0<\tau_1<\dots<\tau_n\leq t$.

واقعه‌ای که در عبارت بالا نسبت به آن مشروط کرده‌ایم معادل است با واقعه $\{N(\tau);\ 0<\tau\leq t\}$.

بنابراین داریم:

$$\mathbb{P}[Z>z \mid \{N(\tau);\ 0<\tau\leq t\}] = e^{-\lambda z}$$

■

حال دنباله زمان‌های بین ورودهای متوالی بعد از زمان t را مورد بررسی قرار می‌دهیم. اولین متغیر را زمان بین t تا اولین ورود بعد از آن در نظر می‌گیریم، یعنی $Z_1 = Z$. برای $m \geq 2$، Z_m را فاصله زمانی بین $(m-1)$ امین ورود پس از t و m امین ورود پس از t تعریف می‌کنیم. بنابراین با فرض اینکه بدانیم $N(t)=n$ و $S_n = \tau$ داریم $Z_m = X_{m+n}$ برای $m \geq 2$.

در تمرین (2.8 کتاب Gallager) اثبات می‌شود که مشروط به $N(t)=n$ و $S_n=\tau$، متغیرهای Z_i به صورت نمایی و iid توزیع شده‌اند. چون این از مقدار $N(t)$ و S_n مستقل است، نتیجه می‌گیریم که $Z_1, Z_2, \dots$ به صورت غیرمشروط هم iid هستند و همچنین مستقل از $N(t)$ و S_n هستند.

همچنین واضح است که $Z_1, Z_2, \dots$ مستقل از $\{N(\tau); 0 < \tau \leq t\}$ هستند.

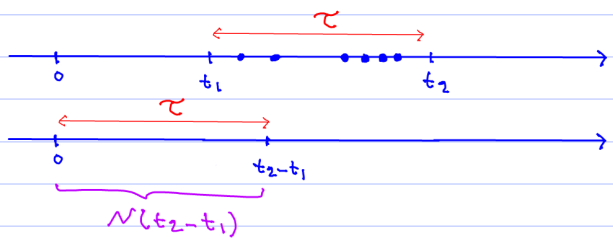
توضیحات بالا نشان می‌دهد که بخشی از فرایند پواسن که از زمان دلخواه $t>0$ شروع شود همان خواص آماری فرایند پواسنی را دارد که از زمان 0 شروع می‌شود.

فرایند شمارشی $N(t)$ را در نظر می‌گیریم و برای هر $t_2 > t_1 > 0$ تعریف می‌کنیم:

$$\tilde{N}(t_1, t_2) = N(t_2) - N(t_1)$$

• تعریف (Stationary increment property):

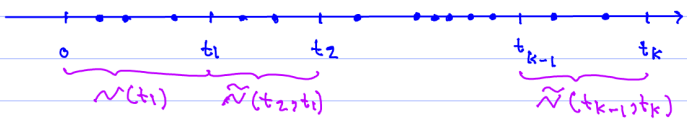
یک فرایند شمارشی $\{N(t); t>0\}$ دارای خاصیت increment property است اگر که برای هر $t_2 > t_1 > 0$ متغیر تصادفی $\tilde{N}(t_1, t_2)$ و $N(t_2 - t_1)$ توزیع یکسانی داشته باشند.



بحث قبل از تعریف بالا نشان می‌دهد که برای فرایند پواسن خاصیت بالا برقرار است. بنابراین تعداد ورودها (arrivals) در یک بازه فقط به طول بازه بستگی دارد و به نقطه شروع بازه بستگی ندارد.

• تعریف (Independent increment property):

یک فرایند شمارشی $\{N(t);\ t>0\}$ خاصیت فوق را دارد اگر که برای هر عدد صحیح $k>0$ و هم k تایی $0<t_1<t_2<\dots<t_k$ ، k متغیر تصادفی $N(t_1)$ ، $\tilde{N}(t_2,t_1)$ و ... و $\tilde{N}(t_{k-1},t_k)$ مستقل از یکدیگر باشند.

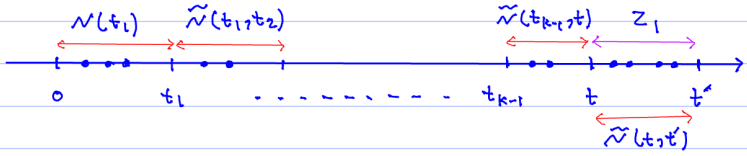


برای فرایند پواسن قضیه قبل بیان می‌کند که برای هر t ، زمان Z_1 تا اولین ورود بعد از t مستقل از $N(\tau)$ برای هر $\tau \le t$ است. حال فرض کنید:

$$t_1<t_2<\dots<t_{k-1}<t$$

در این صورت Z_1 مستقل از $N(t_1)$ و $\tilde{N}(t_1,t_2)$ و ... و $\tilde{N}(t_{k-1},t)$ است.

(شکل زیر را نیز نگاه کنید)



$$Z_1 \perp N(t_1),\ \tilde{N}(t_1,t_2),\ \dots,\ \tilde{N}(t_{k-1},t)$$

$$\tilde{N}(t,t') \perp N(t_1),\ \tilde{N}(t_1,t_2),\ \dots,\ \tilde{N}(t_{k-1},t)$$

همچنین دیدیم که زمان های بین ورود بعد از Z_1 و بنابراین $\tilde{N}(t,t')$ از $N(t_1)$ و $\tilde{N}(t_1,t_2)$ و ... و $\tilde{N}(t_{k-1},t)$ مستقل هستند.

اگر در روابط بالا قرار دهیم $t=t_k$ و $t'=t_{k+1}$ در این صورت $\tilde{N}(t_k,t_{k+1})$ مستقل از $N(t_1)$ و $\tilde{N}(t_1,t_2)$ و ... و $\tilde{N}(t_{k-1},t_k)$ است. از آنجایی که این برای هر k درست است پس فرایند پواسن خاصیت independent increment را دارد.

(اگر فقط زمان های گسسته را در نظر بگیریم، فرایند برنولی نیز دو خاصیت بالا را دارد).

- تابع چگالی احتمال S_n و تابع چگالی احتمال مشترک S_1 و ... و S_n :

یادآوری می‌کنیم که برای یک فرایند پواسن، S_n مجموع n متغیر iid تصادفی است که هرکدام توزیع نمایی با تابع چگالی احتمال زیر را دارند:

$$f_X(x) = \lambda e^{-\lambda x}, \qquad x \geq 0$$

همچنین به یاد بیاوریم که تابع چگالی احتمال مجموع دو متغیر تصادفی مستقل به صورت کانولوشن تابع چگالی هرکدام نوشته می‌شود. پس تابع چگالی S_2 به صورت زیر بدست می‌آید:

$$f_{S_2}(s) = (f_X * f_X)(s)$$
$$= \int_{-\infty}^{\infty} f_X(\tau) f_X(s-\tau)\, d\tau$$

و برای تابع چگالی احتمال S_3 داریم:

$$f_{S_3}(s) = (f_{S_2} * f_X)(s)$$

اگر این کار را n بار انجام دهیم برای تابع چگالی احتمال S_n به توزیع معروف به Erlang یا Gamma می‌رسیم که به صورت زیر است:

$$f_{S_n}(s) = \frac{\lambda^n s^{n-1} e^{-\lambda s}}{(n-1)!}, \qquad s \geq 0$$

برای اینکه این توزیع را بهتر درک کنیم، می‌توانیم به جای روش مکانیکی بالا از روش دیگری که شامل محاسبه توزیع مشترک $S_1, \dots, S_n$ است، همین نتیجه را بدست آوریم.

برای $n=2$، توزیع مشترک S_1 و S_2 (یا معادلاً X_1 و S_2) برابر است با:

$$f_{X_1 S_2}(x_1, s_2) = f_{X_1}(x_1) \cdot f_{S_2|X_1}(s_2|x_1) = \lambda e^{-\lambda x_1} \cdot \lambda e^{-\lambda(s_2 - x_1)}, \qquad 0 \leq x_1 \leq s_2$$

$$= \lambda^2 e^{-\lambda s_2}, \qquad 0 \leq x_1 \leq s_2$$

همانطور که مشاهده می‌کنید تابع چگالی مشترک فوق غیر از نامساوی $0 \leq x_1 \leq s_2$ وابستگی دیگری به x_1 ندارد. در این حالت برای یک مقدار ثابت s_2، تابع چگالی مشترک در بازه $0 \leq x_1 \leq s_2$ ثابت است.

بنابراین برای محاسبه $f_{S_2}(s)$ به جای اینکه مجبور باشیم انتگرال کانولوشن بگیریم، کافی است طول بازه یعنی s_2 را در $f_{X_1 S_2}(x_1, s_2)$ ضرب کنیم. داریم:

$$f_{S_2}(s_2) = \lambda^2 s_2 e^{-\lambda s_2}, \qquad s_2 \geq 0$$

در حالت کلی قضیه زیر را داریم:

• قضیه: فرض کنید X_1 و X_2 و ... متغیرهای تصادفی iid با تابع چگالی احتمال

$$f_X(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$

باشند. برای هر $n \geq 1$ تعریف می‌کنیم: $S_n = X_1 + \cdots + X_n$ • در نتیجه برای هر $n \geq 2$ داریم:

$$f_{S_1 \cdots S_n}(s_1, \ldots, s_n) = \lambda^n e^{-\lambda s_n}, \quad 0 \leq s_1 \leq s_2 \cdots \leq s_n \tag{2.15}$$

اثبات: در قسمت قبل اگر به جای X_1 قرار دهیم S_1، می‌بینیم که قضیه فوق را برای $n=2$ ثابت کرده‌ایم. این را به عنوان پایه استدلال استقرایی که در ادامه می‌آید قرار می‌دهیم.

فرض کنید حکم مسئله برای یک n داده شده برقرار باشد. پس می‌توانیم بنویسیم:

$$f_{S_1 \cdots S_{n+1}}(s_1, \ldots, s_{n+1}) = f_{S_1 \cdots S_n}(s_1, \ldots, s_n)\, f_{S_{n+1} | S_1 \cdots S_n}(s_{n+1} | s_1, \ldots, s_n)$$

$$= \lambda^n e^{-\lambda s_n} \cdot f_{S_{n+1} | S_1 \cdots S_n}(s_{n+1} | s_1, \ldots, s_n)$$

حال می‌دانیم که $S_{n+1} = S_n + X_{n+1}$. چون X_{n+1} مستقل از S_1 و ... و S_n است، بنابراین:

$$f_{S_{n+1} | S_1 \cdots S_n}(s_{n+1} | s_1, \ldots, s_n) = \lambda e^{-\lambda(s_{n+1} - s_n)}$$

با جایگذاری در عبارت قبلی، حکم قضیه ثابت می‌شود.

■

در قضیه بالا نیز تابع چگالی احتمال مشترک S_1 و ... و S_n به غیر از نامساوی $0 \leq S_1 \leq \cdots \leq S_n$ به هیچ متغیری غیر از S_n بستگی ندارد. با انتگرال گرفتن بر روی S_1 و S_2 و ... همان توزیع Erlang که به آن اشاره شده به دست می‌آید:

$$f_{S_n}(s_n) = \int_0^{s_n} \int_0^{s_{n-1}} \cdots \int_0^{s_2} f_{S_1 \cdots S_n}(s_1, \ldots, s_n)\, ds_1\, ds_2 \cdots ds_{n-1}$$

$$= \lambda^n e^{-\lambda s_n} \int_0^{s_n} \int_0^{s_{n-1}} \cdots \int_0^{s_2} ds_1\, ds_2 \cdots ds_{n-1}$$

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

توجه کنید که رابطه بالا برای هر n، توزیع مشترک زمان‌های ورود را بیان می‌کند و بنابراین فرایند پواسن را به طور کامل توصیف می‌کند.

- تابع جرمی احتمال (PMF) متغیر تصادفی $N(t)$

فرایند شمارش پواسن $\{N(t);\ t>0\}$ شامل متغیرهای صحیح غیر منفی $N(t)$ برای هر $t>0$ است.

در ادامه نشان می‌دهیم که PMF این متغیر تصادفی تابع توزیع مشهور پواسن است، همانطور که در قضیه بعد نشان داده می‌شود.

• قضیه: برای یک فرایند پواسن با نرخ λ، و برای هر $t>0$، تابع جرمی احتمال متغیر تصادفی $N(t)$ برابر است با PMF یک توزیع پواسن به صورت زیر:

$$P_{N(t)}(n) = \mathbb{P}[N(t)=n] = \frac{(\lambda t)^n e^{-\lambda t}}{n!}$$

(دقت کنید $N(t)$ برابر است با تعداد ورودها در بازه $(0,t]$).

اثبات 1: این اثبات، برای هر n و t داده شده، به مبنای محاسبه احتمال $\mathbb{P}[t < S_{n+1} < t+\delta]$ به دو طریق برای مقدار کوچک δ است. در راه اول با استفاده از تابع چگالی احتمال S_{n+1} که قبلا محاسبه کرده‌ایم، داریم:

$$\mathbb{P}[t < S_{n+1} < t+\delta] = \int_t^{t+\delta} f_{S_{n+1}}(\tau)\,d\tau = f_{S_{n+1}}(t)\,[\delta + o(\delta)]$$

(یادآوری: می‌گوییم تابع $g(\delta)$ از مرتبه $o(\delta)$ است اگر $\lim_{\delta\to 0} \frac{g(\delta)}{\delta} = 0$.)

رابطه بالا به این علت برقرار است چرا که تابع چگالی احتمال S_{n+1} یک تابع پیوسته در بازه $[t, t+\delta]$ است.

راه دیگر محاسبه $\mathbb{P}[t < S_{n+1} < t+\delta]$ دقت به این نکته است که احتمال اینکه بیشتر از یک ورود در بازه $[t, t+\delta]$ داشته باشیم از مرتبه $o(\delta)$ است. بنابراین با صرف نظر کردن از این حالت، واقعه $\{t < S_{n+1} < t+\delta\}$ در صورتی اتفاق می‌افتد که دقیقا n ورود در بازه $(0,t]$ و یک ورود هم در بازه $[t, t+\delta]$ داشته باشیم. به علت خاصیت independent increment که فرایند پواسن دارد، برای احتمال این واقعه داریم:

$$\mathbb{P}[t < S_{n+1} < t+\delta] = P_{N(t)}(n)\cdot \mathbb{P}[\tilde{N}(t, t+\delta) = 1] + o(\delta)$$

$$= P_{N(t)}(n)\cdot \mathbb{P}[N(\delta) = 1] + o(\delta) \quad \text{: due to Stationary inc. Prop.}$$

$$= P_{N(t)}(n)\cdot \mathbb{P}[X_1 \in (0,\delta],\ X_2 \in (\delta,\infty)] + o(\delta)$$

$$= P_{N(t)}(n)\cdot \mathbb{P}[X_1 \in (0,\delta)]\cdot \mathbb{P}[X_2 \in (\delta - X_1, \infty) \mid X_1 \in (0,\delta)] + o(\delta)$$

داریم:

$$\mathbb{P}[X_1 \in (0,\delta]] = \int_0^\delta \lambda e^{-\lambda t}\, dt = -e^{-\lambda t}\Big]_0^\delta$$

$$= 1 - e^{-\lambda\delta} = 1 - (1 - \lambda\delta + o(\delta)) = \lambda\delta + o(\delta)$$

به طریق مشابه برای جمله شرطی دوم در عبارت قبل می‌توان اثبات کرد که:

$$\mathbb{P}[X_2 \in (\delta - X_1, \infty) \mid X_1 \in (0,\delta)] = 1 - o(\delta)$$

بنابراین:

$$\mathbb{P}[t < S_{n+1} < t+\delta] = P_{N(t)}(n)\,[\lambda\delta + o(\delta)]$$

در نهایت با مساوی قرار دادن عباراتی که از این دو روش بدست آمده اند داریم:

$$P_{N(t)}(n)\,[\lambda\delta + o(\delta)] = f_{S_{n+1}}(t)\,[\delta + o(\delta)]$$

با تقسیم کردن بر δ و محاسبه حد $\delta \to 0$ داریم:

$$P_{N(t)}(n) = \frac{f_{S_{n+1}}(t)}{\lambda} = \frac{(\lambda t)^n e^{-\lambda t}}{n!}$$

■

اثبات 2: به کتاب Gallager مراجعه شود.

چون فرایند پواسن خاصیت Stationary increment دارد پس تعداد ورود بین زمان t_1 و t_2 یعنی $\tilde{N}(t_1,t_2) = N(t_2) - N(t_1)$ همان توزیع $N(t_2-t_1)$ را دارد که طبق قضیه ۱ توزیع پواسن است. پس تعداد ورود یک فرایند پواسن بین زمان های t_1 و t_2 یک متغیر تصادفی پواسن با میانگین $\lambda(t_2-t_1)$ است.

$$\tilde{N}(t_1,t_2) \sim \text{Poisson}(\lambda(t_2-t_1))$$

- چند نکته درباره توزیع پواسن

متغیر تصادفی ای که این توزیع را دارد مقادیر $\{0,1,2,\dots\}$ را به خود می‌گیرد و تابع جرمی احتمال زیر را دارد:

$$\mathbb{P}[X=k] = \frac{\lambda^k}{k!} e^{-\lambda}$$

همچنین داریم:

$$\mathbb{E}[X] = \text{Var}[X] = \lambda$$

تابع مولد گشتاور توزیع پواسن برابر است با:

$$\phi_X(t) = \mathbb{E}[e^{tX}] = e^{\lambda(e^t - 1)}$$

بنابراین جمع دو متغیر پواسن مستقل با پارامترهای λ_1 و λ_2 یک متغیر پواسن با پارامتر $\lambda_1+\lambda_2$ است.

• توزیع پواسن به عنوان حد توزیع دو جمله‌ای:

در بسیاری از مسایل با نوعی توزیع دو جمله‌ای سروکار داریم که n (تعداد آزمایش‌ها) خیلی زیاد است و p (احتمال رخداد واقعه مورد نظر در هر آزمایش) خیلی کوچک، به طوری که همچنان ضریب np ثابت است.

به عنوان مثال تعداد افرادی را که در یک روز خاص وارد ایستگاه قطار شهر می‌شوند در نظر بگیرید. هر فرد خاص با احتمال خیلی کمی (و با تقریب خوبی مستقل از بقیه) وارد ایستگاه می‌شود. ولی چون تعداد افراد کل شهر خیلی زیاد است، بالاخره در هر روز به طور متوسط تعدادی از افراد وارد ایستگاه می‌شوند. در اینجا تعداد آزمایش‌ها (ورود افراد) خیلی زیاد و وارد شدن هر فرد خاص خیلی کم است.

این توزیع همان توزیع پواسن است. به بیان دقیق‌تر توزیع پواسن حالت حدی توزیع دوجمله‌ای است وقتی که:

$$n \longrightarrow \infty, \quad p \longrightarrow 0, \quad np = \lambda = cte$$

به این ترتیب با فرض $np=\lambda$ و $k=o(n)$ داریم:

$$\mathbb{P}[X=k] = \lim_{\substack{n\to\infty, p\to 0 \\ np=\lambda}} \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k}, \quad k \in \{0,1,2,\dots\}$$

$$= \lim_{n\to\infty} \frac{n!}{k!\,(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{\lambda^k}{k!} \lim_{n\to\infty} \frac{n(n-1)\cdots(n-k+1)}{n^k} \left(1-\frac{\lambda}{n}\right)^n \cdot \left(1-\frac{\lambda}{n}\right)^{-k}$$

$$= \frac{\lambda^k}{k!} e^{-\lambda}$$

- تعاریف دیگری برای فرایند پواسن

• تعریف 2 برای فرایند پواسن: یک فرایند شمارش پواسن $\{N(t);\ t>0\}$ یک فرایند شمارش است که تابع جرمی احتمال پواسن داشته باشد، یعنی:

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

و خواص stationary increment و independent increment را داشته باشد.

با استفاده از تعریف اول دیدیم که شرایط تعریف دوم برقرار است. تمرین 2.4 کتاب Gallager نشان می‌دهد که از تعریف دوم می‌توان نشان داد که زمان‌های بین ورود iid با توزیع نمایی هستند، که در نتیجه این دو تعریف معادل می‌شوند.

• تعریف 3 برای فرایند پواسن: یک فرایند شمارش پواسن یک فرایند شمارش است که روابط زیر برقرار باشند:

$$\mathbb{P}[\tilde{N}(t,t+\delta)=0] = e^{-\lambda\delta} \approx 1-\lambda\delta+o(\delta)$$

$$\mathbb{P}[\tilde{N}(t,t+\delta)=1] = \lambda\delta e^{-\lambda\delta} \approx \lambda\delta+o(\delta)$$

$$\mathbb{P}[\tilde{N}(t,t+\delta)\geq 2] = o(\delta)$$

و خواص Stationary increment و independent increment را داشته باشد.

با استفاده از تعریف اول دیدیم که شرایط تعریف دوم برقرار است. عکس این قضیه در تمرین 2.7 کتاب Gallager اثبات می‌شود.

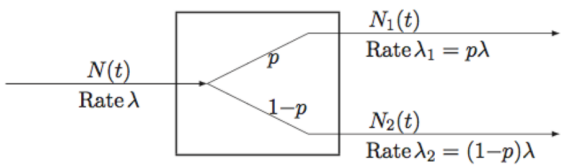
- ترکیب کردن و انشعاب کردن فرایندهای پواسن

• قضیه:

فرض کنید دو فرایند شمارش پواسن مستقل $\{N_1(t);\ t>0\}$ و $\{N_2(t);\ t>0\}$ با نرخ‌های λ_1 و λ_2 داشته باشیم. جمع این دو فرایند را به صورت $N(t)=N_1(t)+N_2(t)$ تعریف می‌کنیم. به این ترتیب فرایند شامل همه ورودهای هر دو فرایند اول و دوم است. $\{N(t);\ t>0\}$ یک فرایند پواسن با نرخ $\lambda=\lambda_1+\lambda_2$ است.

برای اثبات این قضیه به بخش 2.3 کتاب Gallager مراجعه کنید.

• قضیه: فرض کنید یک فرایند شمارش پواسن $\{N(t);\ t>0\}$ با نرخ λ داریم و می‌خواهیم آن را به دو فرایند $\{N_1(t);\ t>0\}$ و $\{N_2(t);\ t>0\}$ تقسیم کنیم. فرض کنید هر ورود در فرایند اولیه به صورت تصادفی و مستقل با احتمال p به فرایند اول و با احتمال $1-p$ به فرایند دوم ارسال شود (مطابق شکل زیر).



Each arrival is independently sent to process 1 with probability p and to process 2 otherwise.

در این صورت هرکدام این فرایندهای جدید، فرایند پواسن هستند با نرخ $\lambda_1=\lambda p$ و $\lambda_2=\lambda(1-p)$ و همچنین این دو فرایند مستقل از یکدیگر هستند.