Processos de Poisson

Ricardo Ehlers
ehlers@icmc.usp.br

Departamento de Matemática Aplicada e Estatística
Universidade de São Paulo

Capitulo 5 Sheldon Ross

Distribuição de Poisson

Definição

Seja a variável aleatória X o número de ocorrências por intervalo fixo (de tempo ou espaço). Dizemos que X tem distribuição de Poisson com parâmetro λ,

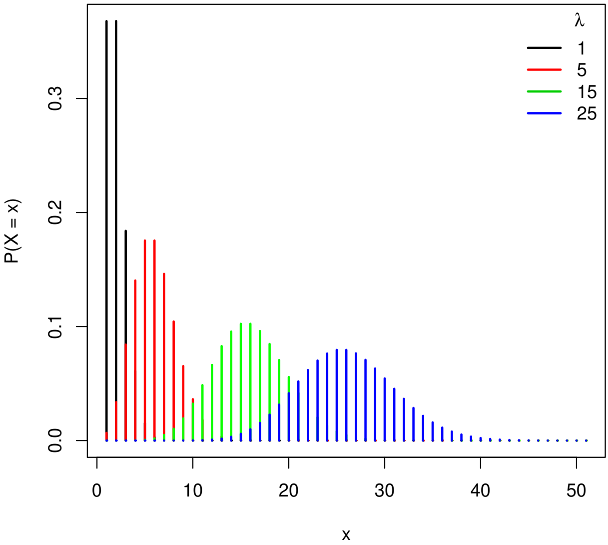
$$X \sim \text{Poisson}(\lambda),\ \lambda > 0,$$

com função de probabilidade,

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!},\ \lambda > 0, \quad k = 0, 1, \ldots.$$

$$
\begin{aligned}
E(X) &= \sum_{k=0}^{\infty} k \, \frac{\lambda^k e^{-\lambda}}{k!} = \lambda \\
Var(X) &= \sum_{k=0}^{\infty} (k - E(X))^2 \, \frac{\lambda^k e^{-\lambda}}{k!} = \lambda
\end{aligned}
$$

Probabilidades Poisson com $\lambda \in \{1, 5, 15, 25\}$



Teorema

Sejam X e Y variáveis aleatórias independentes tais que $X \sim \text{Poisson}(\lambda)$ e $X \sim \text{Poisson}(\nu)$. Então

$$X + Y \sim \text{Poisson}(\lambda + \nu).$$

Portanto,

$$P(X + Y = n) = \frac{(\lambda + \nu)^n \ e^{-(\lambda+\nu)}}{n!}, \ n = 0, 1, \ldots$$

Teorema

Sejam as variáveis aleatórias $N \sim \text{Poisson}(\lambda)$ e $M|N \sim \text{Binomial}(N, p)$. Então

$$M \sim \text{Poisson}(\lambda p).$$

Portanto,

$$P(M = k) = \frac{(\lambda p)^k \ e^{-\lambda p}}{k!}, \ k = 0, 1, \ldots$$

Distribuição Exponencial

Frequentemente usada para modelar o tempo entre eventos que ocorrem a uma taxa média constante.

Definição

Se X é uma variável aleatória com distribuição exponencial sua função de densidade de probabilidade tem a forma,

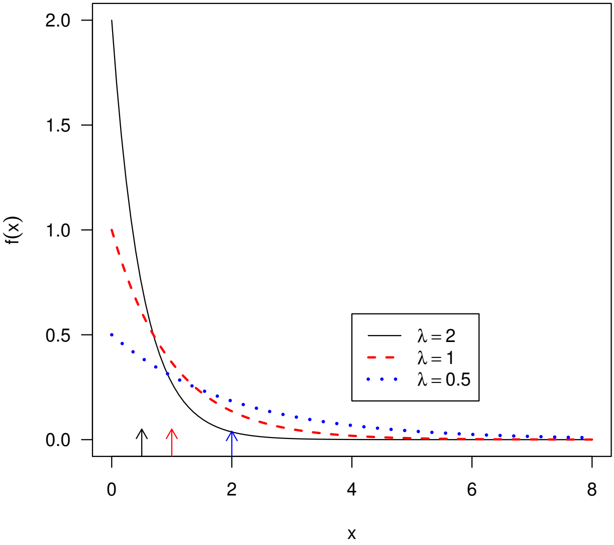
$$f(x) = \lambda e^{-\lambda x}, \ x > 0, \ \lambda > 0.$$

Notação: $X \sim \text{Exponencial}(\lambda)$ ou $X \sim \text{Exp}(\lambda)$.

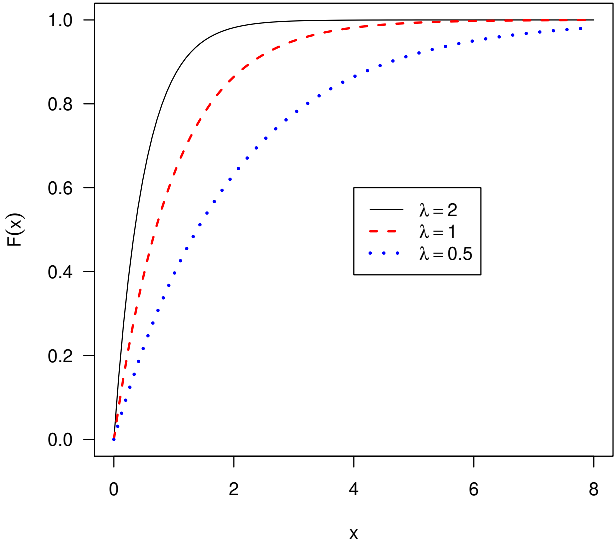
Probabilidades são facilmente calculadas,

$$
\begin{aligned}
P(a < X < b) &= \int_a^b \lambda e^{-\lambda x} dx = e^{-\lambda a} - e^{-\lambda b} \\
F(x) &= \int_0^x \lambda e^{-\lambda u} du = 1 - e^{-\lambda x} \\
E(X) &= \int_0^\infty x\ \lambda e^{-\lambda x} dx = \frac{1}{\lambda} \\
Var(X) &= \int_0^\infty (x - 1/\lambda)^2\ \lambda e^{-\lambda x} dx = \frac{1}{\lambda^2} = \frac{E(X)}{\lambda}
\end{aligned}
$$

Densidades exponenciais.



Funções de distribuição exponenciais.



Falta de Memória

Seja $X \sim \text{Exponencial}(\lambda)$. Então,

$$\begin{aligned} P(X > t + s | X > t) &= \frac{P(X > t + s, X > t)}{P(X > t)} = \frac{P(X > t + s)}{P(X > t)} \\ &= \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = e^{-\lambda s} = P(X > s). \end{aligned}$$

Esta é a única distribuição continua com tal propriedade.

Exemplo. Seja X uma variável aleatória que representa o intervalo de tempo (em minutos) entre emissões consecutivas de uma fonte radioativa. Assume-se que $X \sim$ Exponencial(λ) sendo $\lambda = 0.2$.

Calcule a probabilidade de haver 1 emissão em um intervalo maior do que 7 minutos sabendo que ele é maior do que 5 minutos.

Pela propriedade de falta de memória,

$$\begin{aligned} P(X > 7 | X > 5) &= P(X > 5 + 2 | X > 5) = P(X > 2) \\ &= 1 - P(X < 2) = e^{-2\lambda}. \end{aligned}$$

Taxa de falha

Definição

Seja X uma variável aleatória continua positiva com função de distribuição F e função de densidade f. A função *taxa de falha* é definida como,

$$h(t) = \frac{f(t)}{1 - F(t)}.$$

Em particular, se $X \sim \text{Exponencial}(\lambda)$,

$$h(t) = \frac{\lambda e^{-\lambda x}}{e^{-\lambda x}} = \lambda \text{ (taxa de falha constante).}$$

Sejam $X_1, \ldots, X_n$ variáveis aleatórias independentes. Então,

- Se $X_i \sim \text{Gama}(\alpha_i, \lambda)$,

$$Y = \sum_{i=1}^{n} X_i \sim \text{Gama}\left(\sum_{i=1}^{n} \alpha_i,\ \lambda\right).$$

- Se $X_i \sim \textit{Exponencial}(\lambda)$,

$$Y = \sum_{i=1}^{n} X_i \sim \text{Gama}(n, \lambda).$$

Prova por indução no livro texto.

Proposição

Sejam $X_1, \ldots, X_n$ variáveis aleatórias independentes tais que $X_i \sim$ Exponencial(λ_i), $i = 1, \ldots, n$. Então,

$$\min\{X_i\} \sim \text{Exponencial}\left(\sum_{i=1}^{n} \lambda_i\right).$$

Estude os exemplos da Seção 5.2.

Processos de Poisson

Processo de contagem

Um processo estocástico $\{N(t), t \geq 0\}$ é um processo de contagem se $N(t)$ representa o total de ocorrências de um evento até o tempo t.

Consequentemente,

- $N(t) \geq 0$.
- $N(t)$ assume valores inteiros positivos.
- Se $s < t$ então $N(s) \leq N(t)$.
- Para $s < t$, $N(t) - N(s)$ é o número de ocorrências do evento no intervalo $(s, t]$.

Um processo de contagem tem *incrementos independentes* se os números de ocorrências em intervalos disjuntos são independentes.

Ou seja,

$$N(t_1) - N(t_0), N(t_2) - N(t_1), \ldots, N(t_n) - N(t_{n-1})$$

são variáveis aleatórias independentes para quaisquer $t_0 = 0 < t_1 < \cdots < t_n$, Um processo de contagem tem *incrementos estacionários* se o número de ocorrências do evento em qualquer intervalo de tempo depende somente do seu comprimento.

Ou seja, $N(s+t) - N(s)$ tem a mesma distribuição $\forall s$.

Definição

Um Processo de Poisson com taxa $\lambda > 0$ é um processo de contagem $\{N(t), t \geq 0\}$ tal que,

- os incrementos

$$N(t_1) - N(t_0), N(t_2) - N(t_1), \ldots, N(t_n) - N(t_{n-1})$$

  são variáveis aleatórias independentes para quaisquer $t_0 = 0 < t_1 < \cdots < t_n$,
- $N(0) = 0$.
- $P(N(t+h) - N(t) = 1) = \lambda h + o(h)$
- $P(N(t+h) - N(t) \geq 2) = o(h)$

Definição

Uma função $f(\cdot)$ é dita ser $o(h)$ se,

$$\lim_{h \to 0} \frac{f(h)}{h} = 0.$$

- Se $f(\cdot)$ é $o(h)$ e $g(\cdot)$ é $o(h)$, então $f(\cdot) + g(\cdot)$ é $o(h)$.
- Se $f(\cdot)$ é $o(h)$, então $g(\cdot) = cf(\cdot)$ é $o(h)$.

Qualquer combinação linear finita de funções $o(h)$ é $o(h)$.

Teorema

Seja $\{N(t), t \geq 0\}$ com taxa λ. Então,

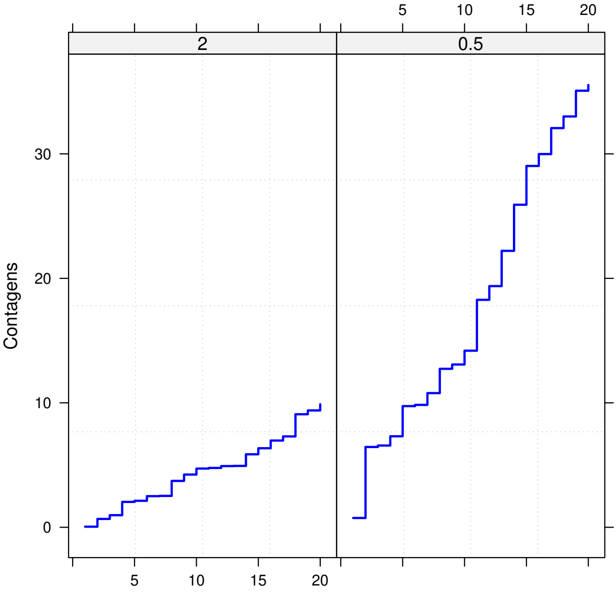
$$N(s+t) - N(s) \sim \text{Poisson}(\lambda t), \; s \geq 0, \; t > 0,$$

Ou seja, o número de eventos em qualquer intervalo de comprimento t tem distribuição de Poisson com taxa λt.

Propriedades

- Em particular, $N(t) \sim \text{Poisson}(\lambda t)$.
- Portanto, $E[N(t)] = Var[N(t)] = \lambda t$.
- Os incrementos são estacionários, i.e. a distribuição de $N(s+t) - N(s)$ só depende da amplitude t.

Processos de Poisson simulados com taxas 2 e 0.5.



- Processos de Poisson são em geral bons modelos para ocorrências aleatórias de um evento em tempo continuo.
- Alguns exemplos: chegadas de clientes em uma fila, emissão de particulas radiotivas, ocorrência de uma doença rara, fótons que chegam a um telescópio espacial.

Exemplo. Clientes chegam a uma loja segundo um processo de Poisson com taxa $\lambda = 4$ clientes por hora. Sabendo que a loja abre as 9h,

- Qual a probabilidade de exatamente 1 cliente ter chegado até as 9:30?
- Qual a probabilidade de chegarem 5 clientes até as 11:30?

Exemplo. Defeitos ocorrem em um cabo submarino segundo um processo de Poisson com taxa $\lambda = 0.1$ por milha.

- Qual a probabilidade de nenhum defeito ocorrer nas primeiras 2 milhas?

  Seja $N(t)$ o número de defeitos no tempo t.

$$\begin{aligned} N(t) &\sim \text{Poisson}(\lambda t) \\ N(2) &\sim \text{Poisson}(0.2) \\ P(N(2) = 0) &= e^{-0.2} \end{aligned}$$

- Qual a probabilidade de nenhum defeito ocorrer entre 2 e 3 milhas dado que não ocorreu defeito nas primeiras 2 milhas?

  Pede-se $P(N(3) - N(2) = 0 | N(2) = 0)$.

  Como $N(3) - N(2)$ e $N(2) - N(0)$ são independentes e $N(3) - N(2) \sim \text{Poisson}(0.1)$, ento

$$\begin{aligned} P(N(3) - N(2) = 0 | N(2) = 0) &= P(N(3) - N(2) = 0) \\ &= e^{-0.1} \end{aligned}$$

Tempos entre ocorrências

Seja $\{N(t), t \geq 0\}$ um processo de Poisson com taxa λ.

- Denota-se por T_n o tempo entre a $(n-1)$ e n-ésima ocorrência de eventos, sendo T_1 o tempo até a primeira ocorrência.
- A sequência $\{T_n, n = 1, 2, \ldots\}$ é chamada de *sequência de tempos entre ocorrências ou entre chegadas*.

Proposição

$T_1, T_2, \ldots$ são variáveis aleatórias indenpendentes com distribuição Exponencial(λ).

Tempos de Espera

Seja S_n o tempo até a n-ésima ocorrência (tempo de espera).

Então,

$$S_n = \sum_{i=1}^{n} T_i, \;\; n = 1, 2, \ldots$$

e portanto $S_n \sim \text{Gama}(n, \lambda)$ com função de densidade,

$$f(t) = \frac{\lambda^n}{\Gamma(n)} t^{n-1} e^{-\lambda t} = \frac{(\lambda t)^{n-1}}{(n-1)!} \; \lambda e^{-\lambda t}, \;\; t \geq 0.$$

Definição alternativa

Seja uma sequência $\{T_n, n \geq 1\}$ de variáveis aleatórias independentes tais que $T_n \sim \text{Exp}(\lambda), \;\; n = 1, 2, \ldots$. Defina um processo de contagem tal que o n-ésimo evento deste processo ocorre no tempo,

$$S_n = \sum_{i=1}^{n} T_i.$$

Então $\{N(t), t \geq 0\}$ tal que $N(t) = \max\{n : S_n \leq t\}$ sendo $S_0 = 0$ é um processo de Poisson com taxa λ.

Exemplo. Pessoas imigram para um território segundo um processo de Poisson com taxa $\lambda = 1$ por dia.

- Qual o tempo esperado até que o 10o migrante chegue?
- Qual a probabilidade de que o tempo entre as chegadas do 10o e do 11o imigrante exceda 2 dias?

Na nossa notação deseja-se calcular $E(S_{10})$ e $P(T_{11} > 2)$,

sendo que,

$$\begin{aligned} T_{11} &\sim Exp(1), \\ S_{10} &\sim Gama(10, 1). \end{aligned}$$

Teorema

Seja $\{N(t), t \geq 0\}$ um processo de Poisson com taxa λ. Para $0 < u < t$ e $k = 0, 1, \ldots, n$,

$$P(N(u) = k | N(t) = n) = \binom{n}{k} \left(\frac{u}{t}\right)^k \left(1 - \frac{u}{t}\right)^{n-k},$$

ou seja $N(u)|N(t) = n \sim \text{Binomial}\left(n, \frac{u}{t}\right)$.

Seja $\{N(t), t \geq 0\}$ um processo de Poisson com taxa λ.

- Cada vez que um evento ocorre ele é classificado como tipo I ou tipo II com probabilidades p e $1-p$ independente dos demais.
- Sejam $\{N_1(t), t \geq 0\}$ e $\{N_2(t), t \geq 0\}$ os números de eventos do tipo I e tipo II que ocorrem no intervalo $(0, t]$.
- Portanto,
$$N(t) = N_1(t) + N_2(t).$$

Proposição

$\{N_1(t), t \geq 0\}$ e $\{N_2(t), t \geq 0\}$ são processos de Poisson com taxas λp e $\lambda(1-p)$. Além disso, os 2 processos são independentes.

Sejam,

$$Y_i = \begin{cases} 1, & \text{ocorre evento tipo I,} \\ 0, & \text{ocorre evento tipo II} \end{cases}$$

sendo $P(Y_i = 1) = p$.

Considere separadamente os processos marcados por 1's e 0's,

$$\begin{aligned} N_1(t) &= \sum_{k=1}^{N(t)} Y_k \\ N_2(t) &= N(t) - N_1(t) \end{aligned}$$

Note que,

- $N_1(t)$ tem incrementos independentes,
- $N_1(0) = 0$,
- $N_1(t) \sim \text{Poisson}(\lambda pt)$, pois $E(Y_k) = p$.

Em resumo,

- $N_1(t)$ é um processo de Poisson com taxa λp e,
- por um argumento análogo $N_2(t)$ é um processo de Poisson com taxa $\lambda(1-p)$.

Exemplo. Clientes entram em uma loja de acordo com um processo de Poisson com taxa $\lambda = 10$ por hora. De forma independente cada cliente compra alguma coisa com probabilidade 0.3 ou sai da loja sem comprar nada com probabilidade 0.7.

Sejam,

- $N_0(t)$: o número de pessoas que não compram nada até o tempo t e,
- $N_1(t)$: o número de pessoas que compram algo até o tempo t,

Então, $N_0(t)$ e $N_1(t)$ são processos de Poisson independentes com taxas $(1-p)\lambda$ e $p\lambda$.

A probabilidade de que durante a primeira hora 9 pessoas entrem na loja, das quais 3 compram alguma coisa e 6 vão embora sem comprar é

$$P(N_0(1) = 6, N_1(1) = 3) = P(N_0(1) = 6)\ P(N_1(1) = 3),$$

sendo que $N_0(1) \sim \text{Poisson}(7)$ e $N_1(1) \sim \text{Poisson}(3)$.

Portanto,

$$P(N_0(1) = 6, N_1(1) = 3) = \frac{7^6 e^{-7}}{6!}\ \frac{3^3 e^{-3}}{3!} = 0.0334.$$

Considere agora variáveis aleatórias discretas $Y_1, Y_2, \ldots$ independentes e cada Y_n pode assumir valores $0, 1, 2, \ldots$ com probabilidades,

$$P(Y_n = k) = p_k, \;\; k = 0, 1, \ldots, \text{ sendo } \sum_{k=0}^{\infty} p_k = 1.$$

- Cada um dos processos resultantes $N_k(t)$, $k = 0, 1, 2, \ldots$ é um processo de Poisson com taxa λp_k, e
- $N_0(t), N_1(t), \ldots,$ são processos independentes.

Exemplo. Considere um sistema no qual equipamentos podem ser classificados em estados $\{0, 1, \ldots, r\}$. Assume-se que,

- Cada equipamento muda de estado segundo uma cadeia de Markov com probabilidades de transição P_{ij}, $i, j = 0, 1, \ldots, r$.
- Os equipamentos mudam de estado independentemente uns dos outros.
- Os números de equipamentos inicialmente nos estados $\{0, 1, \ldots, r\}$ são independentes com distribuição de Poisson e parâmetros $\{\lambda_0, \lambda_1, \ldots, \lambda_r\}$.

Qual a distribuição conjunta dos números de equipamentos nos estados $\{0, 1, \ldots, r\}$ em um tempo n.

Seja $N_j(i)$: o número de equipamentos inicialmente em i que estão em j no tempo n.

- Cada equipamento inicialmente no estado i estará em j no tempo n com probabilidade P_{ij}^n, o elemento (i, j) da matriz P^n.
- Portanto, $N_j(i) \sim \text{Poisson}(\lambda_i P_{ij}^n)$ independentes, $j = 0, 1, \ldots, r$.
- Distribuição do número total de equipamentos no estado j no tempo n,

$$\sum_{i=0}^{r} N_j(i) \sim \text{Poisson}\left(\sum_{i=0}^{r} \lambda_i P_{ij}^n\right), \ j = 0, \ldots, r.$$

Distribuição condicional dos tempos de chegada

Seja $\{N(t), t \geq 0\}$ com taxa λ.

- Sabe-se que exatamente 1 evento ocorreu no intervalo $(0, t]$.
- Qual a distribuição do tempo até a ocorrência do evento?
- Equivalentemente, deseja-se obter

$$P(T_1 < s|N(t) = 1), \; s \leq t.$$

Pela hipótese de incrementos independentes e estacionários é razoável assumir que a distribuição é uniforme em $(0, t]$.

Pode-se verficar que,

$$P(T_1 < s|N(t) = 1) = \frac{s}{t}, \; s \leq t.$$

Estatisticas de ordem

Sejam n variáveis aleatórias $Y_1, \ldots, Y_n$. Então $Y_{(1)}, \ldots, Y_{(n)}$ são as *estatisticas de ordem* se $Y_{(k)}$ é o k-ésimo menor valor dentre $Y_1, \ldots, Y_n$, $k = 1, \ldots, n$.

Se os Y_i forem continuas, independentes e identicamente distribuidas,

$$f(y_1, \ldots, y_n) = n! \prod_{i=1}^{n} f(y_i), \;\; y_1 < y_2 < \cdots < y_n$$

é a densidade conjunta das estatisticas de ordem.

Teorema

Dado que $N(t) = n$, os n tempos de chegada $S_1, \ldots, S_n$ tem a mesma distribuição que as estatisticas de ordem de n variáveis aleatórias com distribuição uniforme em $(0, t]$.

Portanto,

$$
\begin{aligned}
f(s_1, \ldots, s_n | N(t) = n) &= n! \prod_{i=1}^{n} \frac{1}{t} \\
&= \frac{n!}{t^n}, \ 0 < s_1 < \cdots < s_n < t.
\end{aligned}
$$

Processos de Poisson não homogêneos

Em situações práticas pode ser mais razoável assumir que a taxa de ocorrência não é constante e depende do tempo (ou espaço).

Definição

Um processo de Poisson é não homogêneo ou não estacionário com *função de intensidade* $\lambda = \lambda(t)$, $t > 0$ se,

- os incrementos são independentes,
- $N(0) = 0$.
- $P(N(t+h) - N(t) = 1) = \lambda(t)h + o(h)$
- $P(N(t+h) - N(t) \geq 2) = o(h)$

Definição

Função valor médio do processo de Poisson não homogêneo,

$$m(t) = \int_0^t \lambda(u)du.$$

Teorema

Seja $\{N(t), t \geq 0\}$ um processo de Poisson não estacionário com intensidade $\lambda(t)$. Então,

$$N(t+s) - N(s) \sim Poisson(m(t+s) - m(s))$$

sendo,

$$m(t+s) - m(s) = \int_{s}^{t+s} \lambda(u)du.$$

Note que os incrementos não são mais estacionários. Ou seja, eventos podem ser mais prováveis de ocorrer em certos periodos de tempo.

Exemplo. A demanda por atendimento em um posto de saúde occorre segundo um processo de Poisson com taxa,

$$\lambda(t) = \begin{cases} 2t, \ 0 \leq t < 1 \\ 2, \ 1 \leq t < 2 \\ 4 - t, \ 2 \leq t < 4 \end{cases}$$

sendo t medido em horas.

- Qual a probabilidade de que 2 atendimentos ocorram nas primeiras 2 horas?

  Seja $N(t)$ o número de atendimentos no tempo t. Pede-se $P(N(2) = 2)$ sendo $N(2) \sim \text{Poisson}(\mu)$ e,

$$\mu = \int_0^1 2t dt + \int_1^2 2 dt.$$

- Qual a probabilidade de que mais 2 atendimentos ocorram nas 2 horas seguintes?

  Pede-se $P(N(4) - N(2) = 2)$ sendo

$$N(4) - N(2) \sim \text{Poisson}(\mu),$$

$$\mu = \int_2^4 (4 - t) dt.$$

Gerando um processo de Poisson não homogêneo

Seja $\{N(t), t \geq 0\}$ um processo de Poisson com taxa λ.

- Um evento que ocorre no tempo t é contado com probabilidade $p(t)$ independente do que ocorreu antes de t.
- Seja $N_c(t)$ o número de eventos contados até o tempo t.

Então, o processo de contagem $\{N_c(t), t \geq 0\}$ é um processo de Poisson não homogêneo com função de intensidade,

$$\lambda(t) = \lambda p(t).$$

Verificar que $\{N_c(t), t \geq 0\}$ satisfaz aos axiomas de processos de Poisson não homogêneos.

Fazendo uma mudança deterministica na escala temporal e definindo um novo processo $Y(s) = N(t)$ com,

$$s = \Lambda(t) = \int_0^t \lambda(u)du$$

pode-se mostrar que $Y(s)$ é um processo de Poisson homogêneo com taxa igual a 1.

Processos de Poisson compostos

Suponha que $\{N(t), t \geq 0\}$ é um processo de Poisson com taxa λ e cada evento está associado a uma variável aleatória.

Definição

Um processo estocástico $\{Z(t), t \geq 0\}$ é dito ser um *processo de Poisson composto* se puder ser representado como,

$$Z(t) = \sum_{i=1}^{N(t)} Y_i, \;\; t \geq 0,$$

sendo $Y_1, Y_2, \ldots$ variáveis aleatórias independentes e independentes de $N(t)$.

Note que se $Y_i = 1, \; i = 1, 2, \ldots, N(t)$ então $Z(t) = N(t)$.

Exemplo. Suponha que ônibus chegam a um evento de acordo com um processo de Poisson com taxa λ. Os números de pessoas em cada ônibus são independentes e identicamente distribuidos.

Se $Z(t)$ representa o número total de pessoas que chegam de ônibus ao evento e Y_i representa o número de pessoas no ônibus i, então $\{Z(t), t \geq 0\}$ é um processo de Poisson composto,

$$Z(t) = \sum_{i=1}^{N(t)} Y_i, \; t \geq 0,$$

Exemplo. Suponha que clientes saiam de um supermercado de acordo com um processo de Poisson com taxa λ. Sejam $\{N(t), t \geq 0\}$ o número de clientes que saem no tempo t e Y_i a quantia gasta pelo cliente i, independentes e independentes de $N(t)$. Então,

$$Z(t) = \sum_{i=1}^{N(t)} Y_i, \;\; t \geq 0,$$

é um processo de Poisson composto que representa o total gasto no tempo t.

Denotando $E(Y_i) = \mu$ e $Var(Y_i) = \sigma^2$, $i = 1, 2, \ldots,$

$$
\begin{aligned}
E[Z(t)] &= E\left[\sum_{i=1}^{N(t)} Y_i\right] \\
&= E\left[E\left(\sum_{i=1}^{N(t)} Y_i \middle| N(t) = n\right)\right] \\
&= \sum_{n=1}^{\infty}\left[E\left(\sum_{i=1}^{N(t)} Y_i \middle| N(t) = n\right)\right] P(N(t) = n) \\
&= \sum_{n=1}^{\infty} E(Y_1 + \cdots + Y_n) P(N(t) = n) \\
&= \mu \sum_{n=1}^{\infty} n P(N(t) = n) = \mu E[N(t)] = \mu\lambda t
\end{aligned}
$$

Analogamente,

$$Var[Z(t)] = (\mu^2 + \sigma^2)\lambda t.$$

Exemplo. Suponha que famílias migram para uma certa região segundo um processo de Poisson com taxa igual a 2 famílias por semana. Cada família pode ter 1, 2, 3 ou 4 pessoas com probabilidades 1/6,1/3,1/3,1/6. Obtenha o valor esperado e a variância do número de pessoas que migram para esta região num período de 4 semanas e meia.

Sejam Y_i o número de pessoas na familia i e $N(t)$ o número de familias que migram para esta região até o tempo t. Então,

$$Z(t) = \sum_{i=1}^{N(t)} Y_i, \ \ t \geq 0,$$

é um processo de Poisson composto que representa o número de migrantes no tempo t.

Temos então que,

$$
\begin{aligned}
E(Y_i) &= \sum_{i=1}^{4} yP(Y=y) = 5/2 = 2.5 \\
E(Y_i^2) &= \sum_{i=1}^{4} y^2 P(Y=y) = 43/6 \\
Var(Y_i) &= 43/6 - (5/2)^2 = 11/12
\end{aligned}
$$

Portanto, como $N(t)$ tem taxa $\lambda = 2$,

$$
\begin{aligned}
E[Z(t)] &= \mu\lambda t = 2(5/2)t = 5t \\
Var[Z(t)] &= (\mu^2 + \sigma^2)\lambda t = 2[(5/2)^2 + 11/12]t
\end{aligned}
$$

Para $t = 4.5$ semanas o número médio de migrantes é $E[Z(t)] = 22.5$ com variância $Var[Z(t)] = 64.5$.

Processos de Cox

Definição

Um processo de Cox é um processo de Poisson não homogêneo cuja taxa é também um processo estocástico $\{\lambda(t), t \geq 0\}$.

Sejam $\{N'(t), t \geq 0\}$ um processo de Poisson com taxa constante $\lambda = 1$ e o processo $N(t) = N'(\Theta t)$ sendo Θ uma variável aleatória.

- $N(t)$ é um processo de Cox chamado processo de Poisson misto.
- $N(t)|\Theta$ é um processo de Poisson com taxa constante $\lambda = \Theta$.

Em particular, se Θ for continua com densidade $f(\theta)$ a distribuição marginal de $N(t)$ é,

$$\begin{aligned} P(N(t)=k) &= \int_0^\infty P(N(t)=k|\theta)\ f(\theta)d\theta \\ &= \int_0^\infty \frac{(\theta t)^k e^{-\theta t}}{k!}\ f(\theta)d\theta. \end{aligned}$$

Equivalentemente,

$$P(N(t+s)-N(s)=k) = \int_0^\infty \frac{(\theta t)^k e^{-\theta t}}{k!}\ f(\theta)d\theta.$$

- Um processo de Poisson misto tem incrementos estacionários.
- Um processo de Poisson misto em geral não tem incrementos independentes.
- Portanto, um processo de Poisson misto em geral não é um processo de Poisson.

Veja a Seção 5.4.3 do livro texto.

Exemplo. Considere um processo de Poisson misto com parâmetro $\Theta \sim$ Exponencial(1). Obtenha a distribuição marginal de $N(t)$.

$$
\begin{aligned}
P(N(t)=k) &= \int_0^{\infty} \frac{(\theta t)^k e^{-\theta t}}{k!} e^{-\theta} d\theta \\
&= \frac{t^k}{k!} \int_0^{\infty} \theta^k e^{-(t+1)\theta} d\theta \\
&= \frac{t^k}{k!} \frac{\Gamma(k+1)}{(t+1)^{k+1}} = \left(\frac{t}{t+1}\right)^k \left(\frac{1}{t+1}\right),
\end{aligned}
$$

pois,

$$
\int_0^{\infty} \frac{(t+1)^{k+1}}{\Gamma(k+1)} \theta^k e^{-(t+1)\theta} d\theta = 1.
$$

Portanto,

$$N(t) \sim Geometrica(p), \; p = \frac{1}{t+1}, \; t \geq 0.$$

Repita o exemplo para $\Theta \sim \text{Gama}(r, \lambda)$.

Como $N(t)|\Theta$ é um processo de Poisson com taxa Θ então,

$$E[N(t)|\Theta] = Var[N(t)|\Theta] = \Theta t.$$

Portanto,

$$Var[N(t)] = E(\Theta t) + Var(\Theta t) = tE(\Theta) + t^2 Var(\Theta).$$

Finalmente obtem-se que,

$$P(\Theta < x | N(t) = k) \quad = \quad \frac{P(\Theta < x, N(t) = k)}{P(N(t) = k)}$$

sendo,

$$\begin{aligned} P(\Theta < x, N(t) = k) &= \int_0^{\infty} P(\Theta < x, N(t) = k | \theta) g(\theta) d\theta \\ &= \int_0^{x} P(N(t) = k | \Theta = \theta) g(\theta) d\theta. \end{aligned}$$

A função de densidade condicional fica,

$$f(\theta | k) = C \; e^{-\theta t} \; \theta^k \; g(\theta),$$

e a constante C não depende de θ.