Probability

Wenxiao Yang[*]

[*]Department of Mathematics, University of Illinois at Urbana-Champaign

2021

目录

1 Poisson Distribution $Pois(\lambda)$: 单位时间发生 k 次事件的概率

1.1 λ: 单位时间发生该时间的平均次数

$$\Pr(X{=}k) = \frac{\lambda^k e^{-\lambda}}{k!},\ k = 0, 1, 2, 3...$$

1.2 $E(X) = Var(X) = \lambda$

1.3 推导

我们考虑一段时间 (讲单位时间微分成 n 等分, $n \to \infty$)，每一刻 (瞬间) 都有一个 event may occur, which follows binomial distribution $B(n, p)$. where $n \to \infty, p \to 0$; $\lambda = n \cdot p$ is the expected number of events in this period of time.
现在我们考虑发生 k 次 event 的概率:

$$\begin{aligned}
\Pr(X = k) &= \lim_{n\to\infty} \binom{n}{k} (\frac{\lambda}{n})^k (1 - \frac{\lambda}{n})^{n-k} \\
&= \lim_{n\to\infty} \frac{n!}{(n-k)!k!} (\frac{\lambda}{n})^k (1 - \frac{\lambda}{n})^n (1 - \frac{\lambda}{n})^{-k} \\
&= \lim_{n\to\infty} \frac{n!}{(n-k)!k!} (\frac{\lambda}{n})^k e^{-\lambda} \\
&= \frac{\lambda^k e^{-\lambda}}{k!} \lim_{n\to\infty} \frac{n!}{(n-k)!n^k} \\
&= \frac{\lambda^k e^{-\lambda}}{k!} \lim_{n\to\infty} \frac{n}{n} \frac{n-1}{n} \cdots \frac{n-k+1}{n} \\
&= \frac{\lambda^k e^{-\lambda}}{k!}
\end{aligned}$$

2 Exponential distribution $Exp(\lambda)$: 独立随机事件的发生间隔/第一次发生事件的时间

2.1 λ: 单位时间发生该时间的平均次数

随机变量 X 服从参数为 λ 或 β 的指数分布，则记作

$$X \sim \mathrm{Exp}(\lambda) \text{ or } X \sim \mathrm{Exp}(\beta)$$

两者意义相同，只是 λ 与 β 互为倒数关系.

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

$$f(x; \beta) = \begin{cases} \frac{1}{\beta} e^{-\frac{1}{\beta} x} & x \geq 0 \\ 0 & x < 0. \end{cases}$$

累积分布函数为:

$$F(x;\lambda)=\begin{cases}1-e^{-\lambda x} & x\geq 0\\ 0 & x<0.\end{cases}$$

其中 $\lambda>0$ 是分布的参数，即每单位时间发生该事件的次数；$\beta>0$ 为尺度参数，即该事件在每单位时间内的发生率。两者常被称为率参数（rate parameter）。指数分布的区间是 $[0,\infty)$。

2.2 $\mathbb{E}(X)=\frac{1}{\lambda}$: 预期事件的发生间隔；$Var(X)=\frac{1}{\lambda^2}$

$$\mathbb{E}(X)=\frac{1}{\lambda};\ Var(X)=\frac{1}{\lambda^2}$$

2.3 Memorylessness: $\Pr\left(T>s+t\mid T>s\right)=\Pr(T>t)$

$$\begin{aligned}\Pr(T>s+t\mid T>s)&=\frac{\Pr(T>s+t\text{ and }T>s)}{\Pr(T>s)}\\&=\frac{\Pr(T>s+t)}{\Pr(T>s)}\\&=\frac{e^{-\lambda(s+t)}}{e^{-\lambda s}}\\&=e^{-\lambda t}\\&=\Pr(T>t)\end{aligned}$$

2.4 推导

我们考虑一段时间 (讲单位时间微分成 n 等分, $n\to\infty$)，每一刻 (瞬间) 都有一个 event may occur, which follows binomial distribution $B(n,p)$. where $n\to\infty,p\to 0$; $\lambda=n\cdot p$ is the expected number of events in this period of time. （与 Poisson 设定相同）
CDF: 现在我们考虑第一次发生 event 的时间大于 x 的概率:

$$1-F(x;\lambda)=\lim_{n\to\infty}(1-\frac{\lambda}{n})^{nx}=e^{-\lambda x}\Rightarrow F(x;\lambda)=1-e^{-\lambda x}$$

PDF:

$$f(x;\lambda)=\frac{\partial F(x;\lambda)}{\partial x}=\lambda e^{-\lambda x}$$

3 Poisson process: A sequence of arrivals in continuous time with rate λ

3.1 Definition

3.1.1 $N(t) \sim Pois(\lambda t)$: Number of arrivals in length t follows Poisson distribution

$$N(t) \sim Pois(\lambda t)$$
$$\Pr(N(t) = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}$$

3.1.2 The number of arrivals in disjoint time invervals are independent

3.2 T_j: time of j^{th} arrival

$T_1 > t$ is same as $N(t) = 0$: $P(T_1 > t) = P(N(t) = 0) = e^{-\lambda t}$
$\Rightarrow T_1 \sim Expo(\lambda) \Rightarrow T_j - T_{j-1} \sim Expo(\lambda); T_j \sim Gamma(j, \lambda)$

3.3 Theorem (Conditional counts): $N(t_1)|N(t_2) = n \sim Bin(n, \frac{t_1}{t_2})$

可以理解为 n 个点散落在 $(0, t_2]$ 上的概率每处均等 $= \frac{1}{t_2}$; 所以散落在 $(0, t_1]$ 上的概率为 $\frac{t_1}{t_2}$