$q = 1-p$

$$\begin{cases} \text{Bernulli: } M_1(t) = (pe^t + q) \\ \text{Binomial MGF } E(e^{tX}) = M_1(t) M_2(t) \cdots M_n(t) = (pe^t + q)^n \end{cases}$$

$P(X_i = 1) = E(X_i) = p$ T_1 (# of trials untill 1st success) (可能是 infinity)

PMF for T_1 $P(T_1 = k) = P(X_1 = 0, \cdots X_{k-1} = 0, X_k = 1) = P(F \cdots FS) = (1-p)^{k-1} p$ $k = 1, 2, 3, \cdots$

→ k次实验 1次成功，k-1次失败 第k次成功

Check $\sum p_k = \sum_{k=1}^{\infty} P(T_1 = k) = \sum_{k=1}^{\infty} (1-p)^{k-1} p = p \sum_{k=1}^{\infty} (1-p)^{k-1} = p \sum_{k=0}^{\infty} (1-p)^k$

$= 1$

$$\sum_{k=n}^{\infty} q^k = \frac{q^n}{1-q} \quad |q| < 1$$

考试前

It's similar to T_r $r = 2 \cdots$ rth success

$\Rightarrow T_r = k$ [# trials till you get r successes] $\Rightarrow$ ____ $\underset{k\text{th trail}}{\overset{r\text{th}}{S(\text{success})}}$

(k-1) trilals $\begin{cases} (r-1) \text{ successes} \\ k-1-(r-1) \text{ failures} \end{cases}$

$P\{T_r = k\} = \binom{k-1}{r-1} p^{r-1} (1-p)^{k-r} \cdot p$ ⇒ negative binomial

$= \binom{k-1}{r-1} p^r (1-p)^{k-r}$

r=1时，是 geometric distribution

r次成功要k次实验的概率

$\{k = r, r+1 \cdots\}$

→ 至少试r次才有r次成功

一直在用 Binomial

$T_r = k \Rightarrow k-r$ failures till rth successes $\Rightarrow$ $\tilde{T}_r = T_r - r$ (negative distribution)

k次实验 r次成功中 失败的次数

$$\begin{cases} E(\tilde{T}_r) = E(T_r) - r \\ Var(\tilde{T}_r) = Var(T_r) \end{cases}$$

不和 textbook 上的符号一样 failure 的次数 $\tilde{T}_1 = T_1 - 1$ $T_1 = \tilde{T}_1 + 1$

((单次成功概率 p 越小) ⇒ 等待总时间越长 (negative Bernulli)

标准化后 ⇒ P越小 ⇒ 实验次数越多 ⇒ 越接近正态分布. (p=0.9时，成功来得很快且分布不对称

P13)

Binomial:

$$1=(p+q)^n=\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}$$

~~negative $w=1-p=q$ $1-w=p$ (失败概率 w)~~

$$\cancel{1=\sum_{k=r}^{\infty}\underbrace{\binom{k-1}{r-1}(1-p)^{k-r}\cdot p^r}_{P_r(T_r=k)}}$$

?

$$T_r=T_1+(T_2-T_1)+\cdots+(T_r-T_{r-1})$$

独立 (各项的独立性)

$$P_r(T_1=k,T_2-T_1=j)=P_r((\underbrace{F\cdots F}_{k-1}\underset{k\text{th}}{S})(\underbrace{F\cdots F}_{j-1}\underset{j\text{th}}{S}))=P_r(\underbrace{F\cdots F}_{k-1}S)P_r(\underbrace{F\cdots F}_{j-1}S)$$

$$=P_r(T_1=k)P_r(T_2-T_1=j)$$

T_2-T_1 inter-arrival time (time between 2nd and 1st success)

$\Downarrow$

T_k-T_{k-1}

Geometric series $M_1(t)=\dfrac{pe^t}{1-qe^t}$ (r=1时的试验次数)

$E(T_1)=\dfrac{1}{p}$ (等待时间与实验成功概率成反比)

单位

$Var(T_1)=\dfrac{1-p}{p^2}$

Negative Binomial $M(t)=\left(\dfrac{pe^t}{1-qe^t}\right)^r$ $E(T_r)=\dfrac{r}{p}$ $Var(T_r)=\dfrac{r(1-p)}{p^2}$

Example 2.3 Accidents in a workplace.

1. what is chance that there are no accidents in a year?

$$\begin{cases} T_r \text{ (\# months untill the rth accident)} \\ X_t \text{ (\# accidents in t years) } B(12 \times t, \; p = 0.05) \end{cases}$$

$12 \times t$ → 实验次数; $p=0.05$ → 每次成功概率

assumption: In a given month, we will have 0/1 accident only

Success: accident occurs. $p=0.05$

$$P(T_1 > 12) = P_r(X_1 = 0) \text{ (in one year we have zero accident}$$

$$= (1-0.05)^{12} \approx 0.54$$

② prob of 4th accident occurs more than 10 years from now?

$$P(T_4 > 120) = P(T_3 = 120) + P(T_2 = 120) + P(T_1 = 120)$$

$$= P(X_{10} \le 3)$$

$$= \binom{120}{3}(0.05)^3(0.95)^{117} + \binom{120}{2}(0.05)^2(0.95)^{118} + \binom{120}{1}(0.05)(0.95)^{119} + (0.95)^{120}$$

$$= 0.1444076$$

③ the mean and standard deviation of time in years (单位不同,需转换) to the first/fourth accident.

$$T_r \sim NegBinom(r, 0.05) \text{ in years}$$

$r=4$ ⇒ 月数 (发生事故的月数)

事故发生的次数.

$$E(T_4) = \frac{4}{0.05} \div 12$$

$$\sqrt{Var(T_4)} = \sqrt{\frac{r(1-p)}{p^2}} \div 12 =$$

$$E(T_1) = \frac{\frac{1}{0.05}}{12} = \frac{20}{12}$$

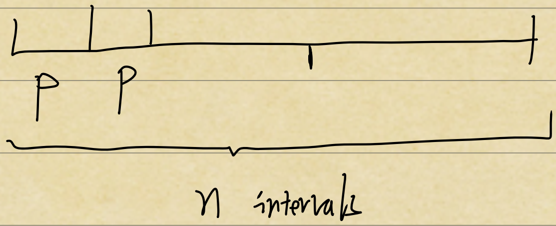
$$\sqrt{Var(T_1)} = \frac{\sqrt{\frac{1-p}{p^2}}}{12} = \frac{\sqrt{\frac{0.95}{0.05^2}}}{12} \Rightarrow \text{geometric distribution}$$

Variance in years? $\div 12^2$ $\qquad Var(\underset{\text{in years}}{Y}) = Var(\frac{\overset{\text{in months}}{T}}{12}) = \frac{1}{12^2} Var(T)$

possion distribution → accident can occur one time but in very small time interval

一个月30天，每天都可能有事故发生

total n periods $\qquad X$ # of accidents



$$\left| Pr(X=k) - e^{-np} \times \frac{(np)^k}{k!} \right| < p$$

Binomil $\binom{n}{k} p^k (1-p)^{n-k}$ $\qquad$ $Pr(Y=k)$ possion $(0,1,\dots,n)$

当n足够大时，p足够小，posion与二项分布 (近似)

主要取决于p

$$\boxed{Pr(Y=k) = e^{-\lambda}\frac{\lambda^k}{k!}}$$

($\lambda = np$, parameter of approximate poisson prob)

Binom$(n,p) \approx$ Poisson$(\lambda = n \cdot p)$ if p is small, n个 small interval

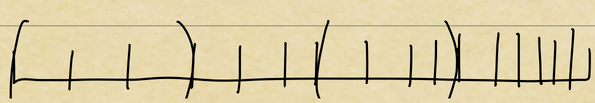
$$\boxed{e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}} \Rightarrow \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = e^{-\lambda}e^{\lambda} = 1$$

$$M(t) = E(e^{tX}) = \sum_{k=0}^{\infty} e^{tk}\frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{(e^t\lambda)^k}{k!} = e^{-\lambda}e^{\lambda e^t} = e^{\lambda(e^t-1)}$$

$E(X) = M'(0) = \lambda \qquad Var(X) = \lambda$

Event occur in continous time with no double occurence of event's at any specific time.

Counts of events in disjoint time intervals are indepdent random variables



long time interval

independent

$\frac{t}{n} \times n$

$\#\text{events} \sim \text{Binom}(n, p) \approx \text{Pois}(np)$, $p = \frac{\lambda}{n}$, $np = \lambda$

when every interval is small enough

0 ———————— t

Poisson Process: $P_r(X=k) = e^{-\lambda t}\frac{(\lambda t)^k}{k!}$ 在t时间内事件发生k次的概率

$X \sim \text{Pois}(\lambda t)$ $t\uparrow \rightarrow$ 总时间越长 $\rightarrow$ X的发生数量越大

t=1时是 poisson(λ) $E(X_{(0,t)}) = \lambda t$

\# event expected to arrive is proportional to the length of Interval

interval $\uparrow \longrightarrow$ more events expected to arrive

$X_{0,t}$ and X_{t+2} X_{t+3} 不同interval的泊松分布相互独立

$\frac{1}{n}\ \frac{1}{n}\ \cdots$

t_1 $\text{Pois}(\lambda t_1)$ t_2 $\text{Pois}(\lambda t_2)$

泊松分布：单位时间事件发生次数的概率

泊松过程：多个泊松分布累加，各分布之间独立；多个独立过程的累积增量

$E(T_{(0,t_1)}) = \lambda t_1$

$E(T_{(t_1,t_2)}) = \lambda t_2$

$P(X_1 = 0) = e^{-0.62}$ $P(X=k) = e^{-\lambda}\frac{\lambda^k}{k!}$

$\lambda = 10 \times 0.62 = 6.2$

$P(T_4 > 10) = P(X_{10} \le 3)$

$= P(X_{10}=1) + P(X_{10}=0) + P(X_{10}=2) + P(X_{10}=3)$

mean: var : 0.62

PDF prob density function

CDF Cumulative prob $\quad F'(x) = f(x)$

T_1 — Poison process (time until 1st event arrives in with rate of λ)

$Pr(t_1 = t) = 0$

$Pr(T_1 \le t) =$ CDF

$$Pr(X(0,t) \ge 1) = 1 - Pr(X(0,t) = 0) = \begin{cases} 1 - e^{-\lambda t}, & t \ge 0 \\ 0 & t < 0 \end{cases}$$

$\sim$ pois(λt)

$T_1 \sim Expon(\lambda)$ (第一次发生的等待时间遵循指数分布)

类似连续情况下的几何分布

differentiable 可导

PMF for discrete

$(CDF)' = PDF$ (Density function) → for continuous

In what condition (Density 可能大于1)?

$$\int_{-\infty}^{x} f(x)\,dx = F(x)$$

$$Pr(1 < T_1 \le 2) = P(T_1 \le 2) - P(T_1 \le 1)$$

$$= (1 - e^{-0.62 \times 2}) - (1 - e^{-0.62 \times 1})$$

$$= 0.249$$

$\mathcal{E}exp(\lambda)$

$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx$ $\qquad$ $E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)\, dx$

MGF: $\int_{-\infty}^{+\infty} e^{tx} f(x)\, dx$

连续时间第4次事件发生耗费时间的概率 $\Gamma(x) = \left[M_{Exp}(t)\right]^4$ (假设求T_4)

Sum(exp) $\int_0^{\infty} \lambda e^{-\lambda x} dx = 1$

每次事件发生的等待时间概率 $\sim Exp(\lambda)$ (Exp的单个期望求和 $\Rightarrow$ $\Gamma(x)$期望)

$$\begin{cases} 4E(T_1) = E(T_4) \\ \sqrt{4Var(T_1)} = \sqrt{Var(T_4)} \end{cases}$$

连续时间事件发生的概率

$P(X_3 < t) \Rightarrow [Poison(\lambda t)]$

poisson MGF $= \left(\frac{\lambda}{\lambda - t}\right)^n$ $\qquad$ (T_n to the nth event of Poisson process of rate λ is M_n given)

Mean : $E(T_n) = \frac{n}{\lambda}$ $\qquad$ $SD(T_n) = \frac{\sqrt{n}}{\lambda}$

$\Gamma(x)$ $\quad f(x) = \begin{cases} 0 & x < 0 \\ \ldots \end{cases}$

$\frac{\lambda^{\alpha} x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}$ $x \geq 0$ waitingtime untill nth event arrive

$\begin{cases} T_n & \text{continous} \\ X_n & \text{discrete} \end{cases}$

poisson : how many events would come in a time interval
Gama : how many time would be spent if nthevent come $\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1} e^{-y}$