Knowledge point.

(• Descriptive Statistic

• Probability
- conditional probability & PC, multiplication principle

• Random variables
-expected value
-variance
-pdf, CDF, pgf, Mgf)

• Inequality
(- Markov’s inequality
- Chebyshev’s inequality)
+ Jensen’s inequality

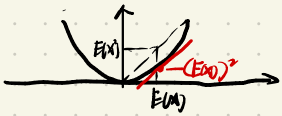
• Discrete distribution
(- uniform distribution
- Bernoulli distribution
- binomial distribution
- hypergeometric distribution)
+ geometric distribution
+ negative binomial distribution
+ Poisson distribution

• Continuous Distribution
+ The uniform distribution
+ The exponential distribution
+ Two-parameter exponential distribution
+ The gamma distribution
+ The Normal distribution and Normal approximation
+ The Chi-square distribution
+ The Weibull Distribution

lecture 4A:

Jensen's inequality: if $g(\cdot)$ is a convex function, then $g(E(X)) \le E(g(X))$ $g''(x) \ge 0$

if $g(\cdot)$ is strictly convex, then equality $g(E(X)) = E(g(X))$ will occur only if X can only assume one single value. $g''(x) > 0$



Geometric Distribution: X = number of trials needed to get the 1st success. $X \sim GEO(p)$

$$P(X=x) = p\cdot(1-p)^{x-1} \text{ for } x=1,2,\ldots$$

$G_X(t) = \frac{tp}{1-q}$ $E(X) = \frac{1}{p}$ $Var(X) = \frac{q}{p^2}$

$\boxed{\text{memoryless property}}$: $P(X>x) = q^x$ $F(x) = P(X \le x) = 1-q^x$

$\Rightarrow \boxed{P(X=k+j \mid X>k) = P(X=j)}$ - has waited k - trails for 1st success, 1st success has not come one step closer.

Negative Binomial Distribution: X = number of trials needed for the rth success $X \sim NEGBIN(r,p)$

$X=x$ only if first $x-1$ trials resulted in $r-1$ success, followed by a success at the x-th trial.

$$P(X=x) = \binom{x-1}{r-1} p^{r-1} q^{x-r} p = \binom{x-1}{r-1} p^r q^{x-r}$$

$G_X(t) = \frac{(pt)^r}{(1-qt)^r}$ $E(X) = \frac{r}{p}$ $Var(X) = \frac{qr}{p^2}$

Relation btw binomial & negative-binomial: $\underbrace{1\ 2\ \ldots\ n}_{<r}$ Y the number of successes

$$P(X>n) = P(Y<r) = P(Y \le r-1 \mid Y \sim Bin(n,p))$$
$$P(X \le n) = P(Y \ge r) = 1-P(Y \le r-1 \mid Y \sim Bin(n,p))$$

lecture 4B

Poisson Distribution: X = number of hits in a certain interval $X \sim POI(\mu)$:

$$P(X=x) = \frac{e^{-\mu}\mu^x}{x!} \quad x=0,1,2\ldots$$

$G = e^{\mu(t-1)}$ $E(X) = \mu$ $Var(X) = \mu$

$\mu = \lambda t$ $\boxed{\lambda \text{ is the intensity of the Poisson process}}$ 每分钟发生多少次.

Relation between Poisson & Binomial Distribution: $POI(\mu) \approx BIN(n,p)$ if $n \to \infty$, $np = \mu$ $p \to 0$

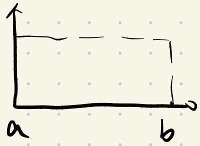
A binomial probability can be approximated by a Poisson probability if $n \ge 20$, $p \le 0.05$

$\to$ when $\mu > 15$ approximation.

lecture 5A:

Uniform Distribution (continuous case)

X = "waiting time for the first bus to arrive".



$f(x) = \frac{1}{b-a}$ for $a \le x \le b$

$$F(x) = P(X<x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{a} 0\,dt = 0 & x<a \\ \int_{-\infty}^{x} f(t)dt = \int_a^x \frac{1}{b-a}dt = \frac{x-a}{b-a}, & a \le x \le b \\ \int_{-\infty}^{x} f(t)dt = \int_a^b \frac{1}{b-a}dt = 1 & x>b. \end{cases}$$

$M_X(t) = \frac{e^{tb}-e^{ta}}{t(b-a)}$ $E(X) = \frac{1}{2}(b+a)$ $Var(X) = \frac{1}{12}(b-a)^2$

$\hookrightarrow E(e^{tx}) = \int e^{tx}\frac{1}{b-a}dx$

$G_X(t) = E(t^X)$

Exponential Distribution:

X = the time elapses between two subsequent events

$X \sim EXP(\lambda)$

$f(x) = \lambda e^{-\lambda x}$ for $x \ge 0$ $E(X) = \int_0^\infty x(\lambda e^{-\lambda x})dx = \frac{1}{\lambda}$

$F(x) = \int_0^x \lambda e^{-\lambda t}dt = 1-e^{-\lambda x}$ for $x \ge 0$ $Var(X) = E(X^2)-(E(X))^2 = \frac{1}{\lambda^2}$

$P(X>t) = \alpha \rightarrow e^{-\lambda t} = \alpha$ $M_X(t) = E(e^{tX}) = \underline{\frac{\lambda}{\lambda-t} = \lambda(\lambda-t)^{-1}}$

$\boxed{\text{memoryless property: } P(X>a+b \mid X>a) = P(X>b)}$ $\hookrightarrow \frac{\lambda}{\lambda-t}\int_0^\infty (\lambda-t)e^{-(\lambda-t)x}dx$

$F(x) = P(X \le x) = 1-e^{-\lambda x}$

$P(X>x) = P(Y_x=0) \Rightarrow$ Poisson Exponential. $P(X \ge x) = e^{-\lambda x}$ $\int \mu e^{-\mu x}dx = 1$

Two parameter Distribution

You certainly know that you have to $\boxed{\text{wait a time } \eta}$ until the next event occurs.

X = the time elapses. $X \sim Exp(\lambda, \eta)$

$f(x) = \lambda e^{-\lambda(x-\eta)}$ for $x-\eta$

$M_X(t) = \frac{\lambda e^{\eta t}}{\lambda-t}$ $E(X) = \eta+\frac{1}{\lambda}$ $Var(X) = \frac{1}{\lambda^2}$

Gamma distribution:

X = waiting time until the rth event $X \sim GAM(\theta, r)$.

$X>t$ only at most $r-1$ in $[0,t] \rightarrow Y_t \sim POI(\lambda t)$

$P(X>t) = P(Y_t \le r-1) = \sum_{k=0}^{r-1} e^{-\lambda t}\frac{(\lambda t)^k}{k!} \Rightarrow f_X(t) = \frac{dF_X(t)}{dt} = \frac{d}{dt}(1-P(X>t)) = -\frac{d}{dt}\left(\sum_{k=0}^{r-1} e^{-\lambda t}\frac{(\lambda t)^k}{k!}\right)$

$= -\frac{d}{dt}\left(e^{-\lambda t}+\sum_{k=1}^{r-1} e^{-\lambda t}\frac{(\lambda t)^k}{k!}\right) = \lambda e^{-\lambda t}+\sum_{k=0}^{r-1}\left(\lambda e^{-\lambda t}\frac{(\lambda t)^k}{k!} - e^{-\lambda t}\frac{k(\lambda t)^{k-1}\lambda}{k!}\right)$

$\lambda = \frac{1}{\theta}$

$= \lambda e^{-\lambda t}+\lambda e^{-\lambda t}\sum_{k=1}^{r-1}\left(\frac{(\lambda t)^k}{k!} - \frac{(\lambda t)^{k-1}}{(k-1)!}\right)$

$f(t) = \frac{1}{\theta^r(r-1)!}t^{r-1}e^{-\frac{t}{\theta}}$ for $t>0$

$= \lambda e^{-\lambda t} + \lambda e^{-\lambda t}[(\cancel{\frac{\lambda t}{1}} - 1) + (\frac{\lambda^2 t^2}{2!} - \cancel{\frac{\lambda t}{1}}) + \dots + (\frac{(\lambda t)^{r-1}}{(r-1)!} - \cancel{\frac{(\lambda t)^{r-2}}{(r-2)!}})]$

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

$\Gamma(r) = (r-1)!$ $\Gamma(x) = (x-1)\Gamma(x-1)$

$\Gamma(n) = (n-1)!$ $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

$M_X(t) = \left(\frac{1}{1-\theta t}\right)^r$ $M_{Y_i}(t) = \frac{1}{1-\theta t} = \frac{\lambda}{\lambda-t}$. $E(X) = r\theta$ $Var(X) = r\theta^2$

Lecture 5B.

normal distribution: $X \sim N(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad F(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} du$$

$M_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$

$M'_X(t) = \mu$

$M''_X(t) = \mu^2 + \sigma^2$. $Var(X) = \sigma^2$.

Standard normal distribution: $Z \sim N(0,1)$

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \quad F(z) = \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du.$$

mgf $= M_Z(t) = e^{\frac{1}{2}t^2}$

$$M_Z(t) = E(e^{tZ}) = \int_{-\infty}^{\infty} e^{tz} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = e^{\frac{1}{2}t^2} \underbrace{\int \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-t)^2}}_{1} dz = e^{\frac{1}{2}t^2}$$

pdf of $N(t,1)$

$E(Z) = 0$

$E(Z^2) = M_Z^{(2)}(0) = 1$.

Standardisation of normal r.v.: if $X \sim N(\mu, \sigma^2)$ and we define $Z = \frac{X-\mu}{\sigma}$ then $Z \sim N(0,1)$

reversed: $Z \sim N(0,1)$ $X = \sigma Z + \mu$ $X \sim N(\mu, \sigma^2)$

$$F_Z(z) = P(Z \le z) = P\left(\frac{X-\mu}{\sigma} \le z\right) = P(X \le \mu + z\sigma) = F_X(\mu + z\sigma)$$

$$\Rightarrow f_Z(z) = \frac{dF_Z(z)}{dz} = f_X(\mu + z\sigma)\,\sigma = \frac{\sigma}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\mu+z\sigma-\mu}{\sigma}\right)^2} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

① $M_{X_1+X_2}(t) = M_{X_1}(t) M_{X_2}(t)$; $X_1 \sim N(\mu_1, \sigma_1^2)$ $X_2 \sim N(\mu_2, \sigma_2^2) \Rightarrow X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

② $S_n \sim N(n\mu, n\sigma^2)$ $\bar{X} = \frac{S_n}{n}$ $M_{\bar{X}}(t) = E(e^{t\bar{X}}) = e^{\mu t + \frac{\sigma^2 t^2}{2n}}$ $\rightarrow \bar{X} \sim N(\mu, \frac{\sigma^2}{n})$.

$X \sim N(\mu, \sigma^2)$. sample mean $\sigma = \frac{\sigma}{\sqrt{n}}$.

Central limit theorem: $S_n = X_1 + X_2 + \dots + X_n$. $(n \ge 30)$

$S_n \sim N(n\mu, n\sigma^2)$ / $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$.

Normal approximation to the binomial distribution: $np \ge 5$ $nq \ge 5$

continuity correction.

Normal approximation to the poisson distribution: $\mu \ge 15$.

Lecture 6A.

Chi-square distribution: $X \sim GAM(\theta = 2, r = \frac{v}{2}) \rightarrow X \sim \chi^2(v)$.

$$f(x) = \frac{1}{\theta^r \Gamma(r)} x^{r-1} e^{-\frac{x}{\theta}} \Rightarrow f(x) = \frac{e^{-\frac{1}{2}x} x^{\frac{1}{2}v - 1}}{2^{\frac{1}{2}v} \Gamma(\frac{1}{2}v)}$$

$$M_X(t) = \left(\frac{1}{1-\theta t}\right)^r \Rightarrow M_X(t) = (1-2t)^{-\frac{v}{2}} \rightarrow M_X(t) = \left(\frac{1}{1-2t}\right)^{\frac{v}{2}}$$

$E(X) = r\theta$ $\Rightarrow E(X) = \frac{v}{2} \cdot 2 = v$

$Var(X) = r\theta^2$ $\Rightarrow Var(X) = 2v$

property: ① if $Z \sim N(0,1)$, $X = Z^2$ then $X \sim \chi^2(1)$

② $X \sim GAM(\theta, r)$ $Y := \frac{2X}{\theta} \sim \chi^2(2r)$

Weibull Distribution: $X \sim WEI(\theta, \beta)$

$$f(x) = \frac{\beta x^{\beta-1}}{\theta^\beta} e^{-(\frac{x}{\theta})^\beta}$$

Transformation Method:

① CDF method: 记得画图 当不是one to one function时.

② Transformation method: $f_Y(y) = f_X(u^{-1}(y)) \left|\frac{d}{dy} u^{-1}(y)\right|$

③ MGF method:
- Start with the definition of mgf of Y
- write the mgf as the expected value of a function of X
- rewrite the integral such that the result can be recognized as the mgf of a distribution

$Z \sim N(0,1)$ $Y = Z^2$ $Y \sim \chi^2(1)$

$M_Y(t) = E(e^{tY}) = E(e^{tZ^2}) = (1-2t)^{-\frac{1}{2}}$

Simulation:

Probability Integral Transformation: X is continuous. $Y = F_X(x)$ has a uniform distribution on $(0,1)$
CDF $F_X(\cdot)$

inverse PIT: $F(\cdot)$ be CDF. $U \sim UNIF(0,1)$. CDF of the random variable $X = F^{-1}(U)$ is equal to $F(\cdot)$

Type of parameters:

① location parameter μ
② Scale parameter θ
③ Intensity parameter $\lambda = \frac{1}{\theta}$
④ Shape parameter β, r

Hazard rate:

T life time CDF $F(\cdot)$

- Reliability of a system on time t: $R(t) = P(T > t) = 1 - F(t)$.

$h(t) = \frac{f(t)}{R(t)} = -\frac{d \ln R(t)}{dt}$

$f(t \mid T > t) = \frac{d}{dt} F(t \mid T \geq t) = h(t)$. conditional density is a pdf. $\int_t^\infty f(x \mid T > t)\, dx = 1$.

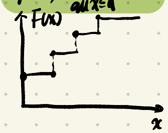
求 median: $P(X=m)=50\%$

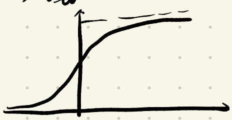
$F(m)=0.5$

求 mode: $f'(x)=0 \Leftrightarrow x=m-2$

Discrete & Continuous

Discrete	Continuous
$f(x)=P(X=x)$	$f(x)\neq P(X=x)$ $\quad P(X=x)=0$
$0\le f(x)\le 1$	$f(x)\ge 0$
$\sum_{\text{all } x} f(x)=1$	$\int_{-\infty}^{\infty} f(x)\,dx=1$
$F(a)=\sum_{\text{all } x\le a} f(x)$	$F(a)=\int_{-\infty}^{a} f(x)\,dx$





Markov's inequality: $P(|X|\ge c)\le \frac{E(X)}{c}$

Chebyshev's inequality $P(|X-\mu|\ge k\sigma)\le \frac{1}{k^2}$

Jensen's inequality $g(E(X))\le E g(X))$

Expected Value:

Discrete

$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$

$\mu = E(X) = \sum_{\text{all } x} x_i f(x)$

$E(g(x)) = \sum_{\text{all } x} g(x) f(x)$

quality: ① $E(c) = c$

② $E(c\,g(x)) = cE(g(x))$

③ $E(g_1(x) + g_2(x)) = E(g_1(x)) + g_2(x))$

④ $E(aX+b) = aE(X)+b$

Continuous

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

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

property $E(c) = c$

$E(c\,g(x)) = cE(g(x))$

$E(g_1(x) + g_2(x)) = E(g_1(x)) + E(g_2(x))$

$E(aX+b) = aE(x)+b$.

Variance:

$\sigma^2 = Var(X) = E[(X-\mu)^2] = \sum_{\text{all } x} (x-\mu)^2 f(x)$

$Var(X) = E(X^2) - (E(X))^2 \rightarrow E(X^2) - \mu^2$.

$Var(aX+b) = a^2 Var(X)$

s.d: $\sigma = \sqrt{\sigma^2} = \sqrt{Var(X)}$

$\sigma_{aX+b} = |a|\sigma_X$

Same

$\sigma^2 = Var(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx$

$= \int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2 = \int_{-\infty}^{\infty} x^2 f(x)dx - [E(X)]^2$

$\sigma^2 = \int_{-\infty}^{\infty} (x^2 - 2\mu x + \mu^2) f(x)\,dx$.

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

$= E(x^2) - \mu^2$

$\sigma = \sqrt{Var(X)}$ $\quad$ $Var(Y) = Var(aX+b) = a^2 Var(X)$

moment & central moment:

$E(X^k)$ - k th moment $\quad$ k=1 - Expected

$E((X-\mu)^k)$ - k th central moment $\quad$ k=2 variance

Probability Generating Function: for discrete r.v

$G_X(t) = E(t^X) = \sum_{x=0}^{\infty} t^x P(X=x) = \sum_{x=0}^{\infty} t^x f_X(x)$.

$G_X(1) = E(1^X) = 1 = \sum_{x=0}^{\infty} f_X(x) = 1$

E(X) $\quad G'_X(t) = \frac{d}{dt}\sum_{x=0}^{\infty} t^x f_X(x) \Rightarrow E(X) = G'_X(1)$

$G''_X(t) \quad G''_X(1) = \sum_{x=2}^{\infty} x(x-1) f(x)$

$= E(X(X-1)) = E(X^2) - E(X)$

$Var(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2$

$P(X=n) = \frac{G_X^{(n)}(0)}{n!}$

Moment Generating Function:

$M_X(t) = E(e^{tX}) = \begin{cases} \sum e^{tx} f_X(x) \\ \int_{-\infty}^{\infty} e^{tx} f_X(x)dx \end{cases}$

$M_X(0) = E(e^{0X}) = 1$

$M'_X(t) = E(Xe^{tX}) \Rightarrow M'_X(0) = E(X)$

$M^{k}_X(t) = \left(\frac{d}{dt}\right)^k M_X(t) = E(X^k e^{tX}) \Rightarrow M_X^{(k)}(0) = E(X^k)$

① $M_Y(t) = e^{bt} M_X(at)$

$\rightarrow Y = aX+b$.

② $M_{X+Y}(t) = M_X(t) M_Y(t)$.

$M_X(t) = G_X(e^t) = E(e^{tX})$

$G_X(t) = E(t^X)$

$E(X+Y) = E(X) + E(Y)$

$Var(X+Y) = Var(X) + Var(Y)$ only independent.

$G_{X+Y}(t) = G_X(t) G_Y(t)$.

$M_X(t) = 1 + \sum_{k=1}^{\infty} \frac{E(X^k)}{k!} t^k = 1 + M'_X(0)t + \frac{M''_X(0)}{2}t^2 + \frac{M_X^{(3)}(0)}{3!}t^3$

$M_X(t) = E(e^{tX}) = E(1 + tX + \frac{t^2X^2}{2!} + \frac{t^3X^3}{3!} + \dots)$

$= 1 + \sum_{k=1}^{\infty} \frac{E(X^k)}{k!} t^k \qquad e^{tX} = \sum_{k=0}^{\infty} \frac{(tX)^k}{k!}$

Discrete Distribution:

⑤ Geometric Distribution:

$P(X=x)=pq^{x-1}$ for $x=1,2,3\ldots$

$$G_X(t)=\sum t^x P(X=x)=\sum t^x pq^{x-1}=\frac{pt}{1-qt}$$

$$E(X)=G'_X(1)=\frac{p(1-qt)-pt(-q)}{(1-qt)^2}=\frac{p}{(1-qt)^2}=\frac{p}{(1-\cancel{q})^2}=\frac{1}{p}$$

$$Var(X)=G''_X(1)+E(X)-(E(X))^2=\frac{-p\cdot 2(1-qt)(-q)}{(1-qt)^4}=\frac{2pq}{(1-qt)^3}+\frac{1}{p}-\frac{1}{p^2}=\frac{2q}{p^2}+\frac{1}{p}-\frac{1}{p^2}=\frac{2q}{p^2}$$

⑥ Negative Binomial Distribution

$$P(X=x)=\binom{x-1}{r-1}p^{r-1}q^{x-r}\cdot p=\binom{x-1}{r-1}p^r q^{x-r}$$

$$G_X(t)=E(t^X)=\sum_{k=r}^{\infty}t^k\binom{k-1}{r-1}p^r q^{k-r}=\sum_{k=r}^{\infty}\binom{k-1}{r-1}(pt)^r(qt)^{k-r}$$

$$=\frac{(pt)^r}{(1-qt)^r}\underbrace{\sum\binom{k-1}{r-1}(1-qt)^r(qt)^{k-r}}_{\sum_{k=r}^{\infty}\binom{k-1}{r-1}p^r q^{k-r}=1}=\frac{(pt)^r}{(1-qt)^r}$$

$$E(X)=G'_X(1)=\frac{r}{p}$$

$$Var(X)=G''_X(1)+E(X)-(E(X))^2=\frac{rp}{q^2}$$

⑦ Poisson Distribution:

$$G_X(t)=E(t^X)=\sum_{k=0}^{\infty}t^k\frac{e^{-\mu}\mu^k}{k!}=\sum_{k=0}^{\infty}e^{-\mu}\frac{(t\mu)^k}{k!}=e^{-\mu}\underbrace{\sum_{k=0}^{\infty}\frac{(t\mu)^k}{k!}}_{e^{t\mu}}=e^{-\mu+t\mu}=e^{\mu(t-1)}$$

$$e^y=\sum_{k=0}^{\infty}\frac{y^k}{k!}$$

$$E(X)=G'_X(1)=\mu e^{\mu(t-1)}=\mu$$

$$G''_X(t)=\mu^2e^{\mu(t-1)}\qquad Var(X)=G''_X(1)+E(X)-(E(X))^2=\mu^2+\mu-\mu^2=\mu.$$