概率论.

1. Permutations

eg. Q.W.E.R.T.Y How many passwords?

sampling r items.

A_n^r $\frac{n!}{(n-r)!}$

2. combination.

C_n^r $\binom{n}{r}$ $\frac{n!}{(n-r)!\,r!}$

3. Binomial Theory

$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$

pf: $(x+y)^n = (x+y)\cdots(x+y)$

$\sum_{k=0}^{n} \binom{n}{k} x^n y^{n-k}$

4. multinomial coff

If we have n items and we want to allocate them to r subsets of sizes $n_1, n_2, \ldots n_r$ with $n_1+n_2+\ldots n_r = n$

then the total way, $\frac{n!}{n_1!n_2!\ldots n_r!} = \binom{n}{n_1,\ldots n_r}$

9 employee 2 mor, 3 aft, 4 night $\frac{9!}{2!3!4!}$

5. disjoint.

$A, B \in F$. $A \cap B = \emptyset$.

6. De Morgan's laws:

$(\bigcup_{i=1}^{n} A_i)^c = \bigcap_{i=1}^{n} A_i^c$ $(\bigcap_{i=1}^{n} A_i)^c = \bigcup_{i=1}^{n} A_i^c$

7. Union bound.

$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$

For any events $A_1, \ldots A_i$ $P(\bigcup A_i) \le \sum_{i=1}^{n} P(A_i)$

8. law of total prob.

let $\Omega = \bigcup_{i=1}^{n} B_i$. $B_i \cap B_j = \emptyset$

For any A, we have $P(A) = \sum_{i=1}^{n} P(A|B_i) P(B_i)$

pf: $C_i = A \cap B_i$ $= \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(C_i) = P(\bigcup_{i=1}^{n} C_i)$

9. Bayes Rule.

For any A, B, $P(A) > 0$, $P(B) > 0$

$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$

suppose $\Omega = \bigcup_{i=1}^{n} B_i$. $B_i \cap B_j = \emptyset$, $i \ne j$

$P(A) > 0$. $P(B_i) > 0$

$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{i=1} P(A|B_i) P(B_i)}$

10. Ind Events

$P(A|B) = P(A)$ $P(B|A) = P(B)$

$P(AB) = P(A)P(B)$

11. EFG are said to be independent.

if $P(EFG) = P(E)P(F)P(G)$

$P(EF) = P(E)P(F)$; $P(EG) = P(E)P(G)$. $P(FG) = P(F)P(G)$.

If $A \perp B$, $B \perp C$, $A \perp C$.

but. $A \cap B$ is not indep with C.

EX. consider tossing a coin twice.

$A = \{\text{heads on 1st}\}$ $\frac{1}{2}$

$B = \{\text{heads on 2nd}\}$ $\frac{1}{2}$

$C = \{\text{exactly one head}\}$ $\frac{1}{2}$

$P(AC) = \frac{1}{2} \times \frac{1}{2} = P(A)P(C)$

$P(AB) = P(A)P(B)$

$P(AC) = P(A)P(C)$

$P(C|A \cap B) \ne P(C)$

$A_1 \cdots A_n$ are indep

We say For any sub-collection. $A_{i_1} \to A_{i_k}$. satisfy.

$P(\bigcap_{l=1}^{k} A_{i_l}) = \prod_{l=1}^{k} P(A_{i_l})$.

12. convexity.

Def A is a convex set. if $x, y \in A$, $\lambda \in [0,1]$.

$\lambda x + (1-\lambda) y \in A$.

Def. f is function defined on A. A is convex set

f is convex if $\forall x, y \in A$, $\lambda \in [0,1]$

$\lambda f(x) + (1-\lambda) f(y) \ge f(\lambda x + (1-\lambda) y)$.

opposite: concave.

13. Convergent Sequences / series.

$a_n > 0$. $S_n = \sum_{i=1}^{n} a_i$ $\lim_{n \to \infty} S_n < \infty$ converge.

If S_n converges, then $a_n \to 0$ when $n \to \infty$.

14. $P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \le i < j \le n} P(A_i A_j) + \sum_{1 \le i < j < k \le n} P(A_i A_j A_k)$

$+ \ldots (-1)^{n-1} P(A_1 A_2 \cdots A_n)$

15. Discrete random Variable.

We say a rv is discrete if it can only take a finite or countable # of values.

16. Bernoulli rV.

$X=1$ w/p 1 / $X=0$ w/p 0.

$P_X(x)=p^x(1-p)^{1-x}$. $x\in\{0,1\}$.

Define indicator function.

$I_A(w)=1$ if $w\in A$, $=0$ if $w\notin A$.

Sum of Bernoulli. If $X_1\to X_n$ are indep Bernoulli(p)

Binomial (n, p)

$P(A)=\binom{n}{k}p^k(1-p)^{n-k}$.

For Bernoulli. $E(X)=p$ $Var(X)=EX^2-(EX)^2$

$=p-p^2$

$=p(1-p)$

For Binomial. $E(X)=E\sum E(X_i)=np$.

or $E(X)=\sum_{k=0}^{n}k\binom{n}{k}p^k(1-p)^{n-k}=np\sum_{k=0}^{n}\frac{(n-1)!}{(k-1)!(n-k)!}p^{k-1}(1-p)^{n-k}=n$

$E(X^2)=\sum_{k=0}^{n}k^2\binom{n}{k}p^k(1-p)^{n-k}=\sum_{k=0}^{n}[k(k-1)+k]\binom{n}{k}p^k(1-p)^{n-k}$.

$=n(n-1)p^2+np$.

$Var(X)=EX^2-(EX)^2=np(1-p)$

---

17. Geometric distribution. (无记忆性) $P(X>m+n|X>m)=P(X>n)$

记每次试验中事件A发生的概率为P，如果X为事件A首次出现时的试验次数

$P_X(k)=(1-p)^{k-1}p$.

$E(X)=\sum_{k=1}^{\infty}(1-p)^{k-1}pk=p\sum_{k=1}^{\infty}k(1-p)^{k-1}\overset{def}{\underset{1-p=q}{=}}p\sum_{k=1}^{\infty}kq^{k-1}=p\sum_{k=1}^{\infty}\frac{dq^k}{dq}$

$=\frac{1}{p}$

$E(X^2)=\sum_{k=1}^{\infty}(1-p)^{k-1}pk^2=\sum[k(k-1)+k](1-p)^{k-1}p$, ... 二阶导.

$Var(X)=\frac{1-p}{p^2}$

---

18. negative. binomial (r, p) (帕斯卡分布)

记每次试验中事件A发生的概率为P. X为事件A第r次出现时的次数

$P(X=k)=\binom{k-1}{r-1}p^{r-1}(1-p)^{k-r}$, $k=r, r+1, \ldots$

$E(X)=\frac{r}{p}$ $Var(X)=r\frac{(1-p)}{p^2}$

如果将第1个A出现时的试验次数记为X_1，第二个出现的试验次数(从第一次A出现后算起)记为X_2 .. X_r.

$X=X_1+\ldots X_r$ $X_i\sim$ 几何分布. $\boxed{E.\ Var\text{好算}}$.

19. Hypergeometric

设有N件产品，其中有M件不合格，不放回随机抽取n件 则其中含有次品的件数X服从超几何分布. $X\sim(n, N, M)$

$P(X=k)=\frac{\binom{M}{k}\binom{n-M}{n-k}}{\binom{N}{n}}$

$E(X)=n\frac{M}{N}$

$Var(X)=\frac{nM(N-M)(N-n)}{N^2(N-1)}$

---

19. Poisson (λ). story: Many opportunities for something bad to happen but each oppor is unlikely.

$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$ 泰勒展开. $e^x=1+x+\frac{x^2}{2!}+\ldots$

$k=0,1,2\ldots$ $=\sum_{k=0}\frac{\lambda^k}{k!}$

$E(X)=\sum_{k=0}k\frac{\lambda^k}{k!}e^{-\lambda}=\lambda\sum_{k=1}\frac{\lambda^{k-1}}{(k-1)!}e^{-\lambda}=\lambda$

$E(X^2)=\sum_{k=0}k^2\frac{\lambda^k}{k!}e^{-\lambda}=\sum_{k=0}[k(k-1)+k]\frac{\lambda^k}{k!}e^{-\lambda}$

$Var(X)=\lambda$ $=\lambda^2+\lambda$

二项分布的泊松近似.

泊松定理. 在n重伯努利试验中，记事件A在一次试验中发生的概率为P_n. $n\to\infty$ $np_n\to\lambda$ n large p small,

则 $\lim_{n\to\infty}\binom{n}{k}p_n^k(1-p_n)^{n-k}=\frac{\lambda^k}{k!}e^{-\lambda}$.

Let's see how Poisson (λ) comes from Binomial 用 (n, p).

Binomial (n, p). $P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$ | $\lambda=np$

$=\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}$. | $p=\frac{\lambda}{n}$

$=\frac{n!}{k!(n-k)!}\left(\frac{\lambda}{n}\right)^k\left(\frac{n-\lambda}{n}\right)^{n-k}$

$=\frac{\lambda^k}{k!}\underbrace{\frac{n!}{(n-k)!n^k}}_{1}\underbrace{\left(\frac{n-\lambda}{n}\right)^n}_{e^{-\lambda}}\underbrace{\left(\frac{n-\lambda}{n}\right)^{-k}}_{1}$ as $n\to\infty$.

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

差存在 → 期望存在 ✓
← ×

$(<X\le b) = F_X(b) - F_X(a)$

$= P_X(X\le b) - P_X(X\le a)$

iform distribution

$\sim \text{Uniform}[a,b]$ $\qquad E(X) = \frac{b+a}{2}$

$(x) = \begin{cases} \frac{1}{b-a}, & x\in[a,b] \\ 0, & \end{cases}$ $\qquad Var(X) = \frac{(b-a)^2}{12}$

$(x) = \int_a^x \frac{1}{b-a}dx = \frac{x-a}{b-a}$

ponential distribution (无记忆性) $P(T>t+s \mid T>t) = P(T>s)$

$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x\ge 0 \\ 0, & x<0 \end{cases}$ $\qquad E(X) = \frac{1}{\lambda}$

$(x) = \begin{cases} 1-e^{-\lambda x}, & x\ge 0 \\ 0, & x<0 \end{cases}$ $\qquad var(X) = \frac{1}{\lambda^2}$

ussian or Normal rV

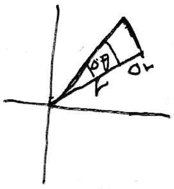
$(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$ $\qquad \mu.$

$\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}dx = \sqrt{2\pi}$ $\qquad \sigma^2.$

$\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}dx \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}dy = 2\pi$

$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}}dxdy = 2\pi$

$r^2 = x^2+y^2$
$r\cos\theta = x$ $\qquad dxdy = rdrd\theta$
$r\sin\theta = y$



$\int_0^{2\pi}\int_0^{\infty} e^{-\frac{r^2}{2}} rdrd\theta.$

$\frac{1}{2}\int_0^{\infty} e^{-\frac{r^2}{2}}dr^2 \int_0^{2\pi} d\theta.$

2π

g be a strictly increasing and differentiable fcn.

$Y=g(x)$. $f_Y(y) = f_X(g^{-1}(y)) \times \left|\frac{dg^{-1}(y)}{dy}\right|$.

ppose X has cts strictly increasing F_X, generate new $\tilde{X}\sim X$.

$\tilde{X} = F_X^{-1}(u)$. $\qquad u\sim \text{uniform}[0,1]$

$F_{\tilde{X}}(x) = P(\tilde{X}\le x) = P(u\le F(x)) = F_X(x)$.

Joint distributions for cts r.v.

$f_{X,Y}(x,y) = P(x\le X\le x+dx,\ y\le Y\le y+dy)$.

$F_{X,Y}(x,y) = P(X\le x, Y\le y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(u,v)dudv$.

边际分布.

$F_X(x) = P(X\le x, -\infty\le Y\le\infty)$
$= \int_{-\infty}^{x}\int_{-\infty}^{\infty} f_{X,Y}(u,v)dvdu$

$f_X(x) = F_X'(x) = \int_{-\infty}^{\infty} f_{X,Y}(u,v)dv$.

独立.

$F(x_1,\dots,x_n) = \prod_{i=1}^{n} F(x_i)$ $\qquad P(x_1\dots x_n) = \prod_{i=1}^{n} P(x_i)$

卷积公式.

离散: $Z=X+Y$.
$P(Z=k) = \sum_{i=0}^{k} P(X=i)P(Y=k-i)$

$X\sim P(\lambda_1), Y\sim P(\lambda_2)$ X与Y独立.
$X+Y\sim P(\lambda_1+\lambda_2)$

连续: X与Y是两个相互独立的连续型随机变量.
$Z=X+Y$
$P_Z(z) = \int_{-\infty}^{\infty} P_X(z-y)P_Y(y)dy = \int_{-\infty}^{\infty} P_X(x)P_Y(z-x)dx$.

正态分布的可加性.

$X\sim N(\mu_1,\sigma_1^2), Y\sim N(\mu_2,\sigma_2^2)$ X与Y独立.
$Z=X+Y\sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$

条件Pmf.

$P_{Y|X}(y|x) = \frac{P_{X,Y}(x,y)}{P_X(x)}$

条件Pdf.

$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$

最大值分布. 最小值分布.

$F_{X_{(n)}}(x) = P(X_{(n)}\le x) = P(X_1\le x,\dots X_n\le x)$
$= (F(x))^n$

$P_{X_{(n)}}(x) = n[F(x)]^{n-1}f(x)$.

$F_{X_{(1)}}(x) = P(X_{(1)}\le x) = 1-P(X_{(1)}>x)$
$= 1-P(X_1>x,\dots X_n>x)$
$= 1-(1-F(x))^n$.

$P_{X_{(1)}}(x) = n(1-F(x))^{n-1}f(x)$

变量变换法

设二维随机变量 (X,Y) 的联合密度函数为 $P(x,y)$

如果函数 $\begin{cases} u=g_1(x,y) \\ v=g_2(x,y) \end{cases}$

有连续偏导数，且存在唯一反函数. $\begin{cases} x=x(u,v) \\ y=y(u,v) \end{cases}$

且变换的雅可比行列式.

$$J=\frac{\partial(x,y)}{\partial(u,v)}=\begin{vmatrix} \frac{\partial x}{\partial u}, & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{vmatrix}=\left(\frac{\partial(u,v)}{\partial(x,y)}\right)^{-1}$$

$$P(u,v)=P(x(u,v),y(u,v))\,|J|.$$

例子: $U=X+Y$ $\quad X,Y\sim N(\mu,\sigma^2)$

$V=X-Y$.

$x=\frac{u+v}{2}$

$y=\frac{u-v}{2}$.

$$J=\begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{vmatrix}=\begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix}=-\frac{1}{2}$$

$$P(u,v)=P_X\left(\frac{u+v}{2}\right)P_Y\left(\frac{u-v}{2}\right)\left|-\frac{1}{2}\right|$$

$$V(y)=\int_{a(y)}^{B(y)} f(x,y)\,dx.$$

$$V'(y)=\int_{a(y)}^{B(y)}\frac{\partial f(x,y)}{\partial y}dx+f(B(y),y)B'(y)$$

$$-f(a(y),y)a'(y)$$

Indicator Variable.

Recall if A is an event, then we def $1_A(w)=\begin{cases} 1, & \text{if } w\in A \\ 0, & \text{if } w\notin A \end{cases}$

$E[1_A]=P(A)$

Note: 1) $1_{A^c}=1-1_A$

2) $1_{A_1\cap A_2}=1_{A_1}\cdot 1_{A_2}$

3) $(A_1\cup A_2)^c=A_1^c\cap A_2^c$

$1_{A_1\cup A_2}=1-1_{A_1^c\cap A_2^c}$

4) $P(\bigcup_{i=1}^{n} A_i)\le\sum_{i=1}^{n}P(A_i)$

Pf: $1_{A_1\cup A_2}\le 1_{A_1}+1_{A_2}$

$E(\ )\le E(\ )$

$P(A_1\cup A_2)\le P(A_1)+P(A_2)$

covariance 描述两个变量之间的相关联程度.

$\mathrm{cov}(X,Y)=E[(X-EX)(Y-EY)]=E(XY)-E(X)E(Y)$

cts: $\mathrm{cov}(X,Y)=\iint(x-\mu_x)(y-\mu_y)f_{X,Y}(x,y)\,dx\,dy$

Discrete: $\mathrm{cov}(X,Y)=\sum_x\sum_y(x-\mu_x)(y-\mu_y)P_{X,Y}(x,y)$

$$\mathrm{Var}(\sum X_i)=\sum \mathrm{Var}(X_i)+2\sum_{i<j}\mathrm{cov}(X_i,X_j)$$

E.g. $X\sim N(0,\sigma^2)$ $Y=X^2$. Y和X不独立.

$\mathrm{cov}(Y,X)=E(X^3)-E(X)E(X^2)=0$.

coefficient.

$$\mathrm{corr}(X,Y)=\frac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$$

Cauchy-Schwarz inequality.

$$[\mathrm{cov}(X,Y)]^2\le \mathrm{Var}(X)\,\mathrm{Var}(Y)$$

Pf: 设 $\mathrm{Var}(X)>0$. 因为当 $\mathrm{Var}(X)=0$ 成立.

$$g(t)=E[t(X-EX)+(Y-EY)]^2=\mathrm{Var}(X)t^2+2\,\mathrm{cov}(X,Y)t+\mathrm{Var}(Y)$$

$\Delta\le 0$ $\quad 4\,\mathrm{cov}(X,Y)^2-4\,\mathrm{Var}(X)\mathrm{Var}(Y)\le 0.$

$\Rightarrow \checkmark$

重期望公式.

$E(X)=E(E(X|Y))$

$\mathrm{Var}(Y)=\mathrm{Var}(E(Y|X))+E(\mathrm{Var}(Y|X))$

where $\mathrm{Var}(Y|X)=E[Y^2|X]-[E[Y|X]]^2$

Moment generating fcn.

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

Theorem. If two rv's X & Y satisfy $M_X(t)=M_Y(t)$ for all $t\in(-\varepsilon,\varepsilon)$ for some $\varepsilon>0$ then $F_X(u)=F_Y(u)$ for all u

other fact. If X has an mgf, then for any $n\ge 1$,

$E[X^n]=M_X^{(n)}(0)$

Note: $M_X^{(1)}(t)=M_X'(t)$ / $M_X^{(2)}(t)=M_X''(t)$

e^{tx} 在0处展开 $1+tx+\frac{(tx)^2}{2!}\ldots$

$E[e^{tX}]=1+tE(X)+t^2\frac{EX^2}{2!}+\ldots$

$Y=a+bX$ $\quad M_Y(t)=E(e^{t(a+bX)})=e^{ta}M_X(bt)$

If X, Y indep. $\quad M_{X+Y}(t)=E[e^{t(X+Y)}]=M_X(t)M_Y(t)$

Law of Large Numbers

LLN.

Informal statement: If $x_1, x_2 \cdots$ are indep rv w/ $E[X_i] = u$ for all i.
then $\bar{X} = \frac{1}{n}\sum X_i$ will be close to u with prob as n becomes large.

Tools for proving LLN.

① Let Y be non-neg cts r.v.

$$E(Y) = \int_0^\infty P(Y>y)\,dy$$

Pf:

$$\int_0^\infty P(Y>y)dy = \int_0^\infty \int_y^\infty f_Y(u)\,du\,dy = \int_0^\infty \int_0^\infty I(y<u) f_Y(u)\,du\,dy$$

$$= \int_0^\infty \int_0^u dy\, f_Y(u)\,du$$

$$= \int_0^\infty u f_Y(u)\,du$$

$$= E(Y)$$

① Markov's ineq.

Let $W \geq 0$ be a non-neg rv. Then

$$P(W \geq t) \leq \frac{E(W)}{t} \quad (\text{for any } t>0)$$

$$\int_t^{+\infty} f_W(w)\,dw \leq \int_t^{+\infty} \frac{w}{t} f_W(w)\,dw = \frac{E(W)}{t}.$$

② chebyshev 不等式.

设随机变量 X 的数学期望和方差都存在

$\forall \varepsilon > 0$.

$$P(|X-EX| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}$$

$$P(|X-EX| < \varepsilon) \geq 1 - \frac{Var(X)}{\varepsilon^2}.$$

$$P(|X-EX| \geq \varepsilon) = P(|X-EX|^k \geq \varepsilon^k)$$

$$= \int_{\{x:\, |x-EX|^k \geq \varepsilon^k\}} f_X(x)\,dx \leq \int_{\{x:\, |x-EX|^k \geq \varepsilon^k\}} \frac{|x-EX|^k}{\varepsilon^k} f_X(x)\,dx$$

$$\leq \frac{E|X-EX|^k}{\varepsilon^k}$$

$k=2$

$$= \frac{Var(X)}{\varepsilon^2}$$

Formal statement of LLN.

Let $X_1, X_2 \ldots$ be indep rv w/ $E[X_i] = u$ for all.
Then for any $\varepsilon > 0$, $P(|\bar{X}_n - u| > \varepsilon) \to 0$ as $n \to \infty$

Pf: (under extra assumption $Var(X_i) = \sigma^2$ for all i.

$$P(|\bar{X}_n - u| > \varepsilon) \leq \frac{Var(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \to 0 \text{ as } n \to \infty.$$

Notation: Let $W_1, W_2 \ldots$ be a seq of rv's. and Let c be a const. We say $W_n \xrightarrow{P} c$ if for any $\varepsilon > 0$, $P(|W_n - c| > \varepsilon) \to 0$ as $n \to \infty$ 依概率收敛

converge of sequence of #'s.

Let $a_1 \ldots a_n$ be a seq of real #'s
Let a be fixed then we say $a_n \to a$ as $n \to \infty$ If for any $\delta > 0$, there is $N_\delta > 0$ s.t $n \geq N_\delta$, $|a_n - a| \leq \delta$.

依分布收敛. $Z_n \to Z$ in distribution

$$\Leftrightarrow F_{Z_n}(t) \to F_Z(t). \ \forall t.$$

$$E((Z_n - Z)^2) \to 0.$$

quadratic $\to$ prob $\to$ distribution.

Jense 不等式.

convex fcn.

$\lambda \in [0,1]$. $g(\lambda x + (1-\lambda)y) \leq \lambda g(x) + (1-\lambda) g(y)$

For convex functions, local minimam are gloal minima.

二阶导 ≥ 0. $f(x), g(x)$ 都凸

max, min 都是凸函数. $f(x)$ 凸 \ $f(Ay+b) = g(y)$ 凸.

$f(x)+g(x)$. 凸.

$g[E(X)] \leq E[g(X)]$. Jenze ineq.

CLT.

Informal statement

if $X_1, \cdots X_n$ are iid rv and $\bar{X}_n = \frac{1}{n}(X_1 + \cdots X_n)$

$\bar{X}$ fluctuate $\mu = E[X_i]$ like a Gaussian rv

~~Def: If W_1, W_2 is a seq of rv. then we say W_n converges in dist to a rv. ($W_n \xrightarrow{d} W$)~~

If ~~$F_{W_n}(d)$~~ $\to F_W(d)$, ~~at every~~ where

F_W is 连续.

~~Note that if we~~ can approximate F_{Z_n} asymptotically $(n \to \infty)$

Formal Version

$$Z_n = \frac{S_n - n\mu}{\sqrt{n\sigma^2}}$$

Let $X_1, \cdots X_n \overset{iid}{\sim}$ r.v. w/ $E[X_1] = \mu$, $Var(X_i) = \sigma^2$

$\forall c$, then $F_{Z_n}(u) \to \Phi(u)$ where Φ is cdf $N(0,1)$

设$\{X_n\}$是独立同分布的随机变量序列，且$E(X_i) = \mu$, $Var(X_i) = \sigma^2 > 0$.

存在，若记 $Y_n^* = \frac{X_1 + X_2 + \ldots X_n - n\mu}{\sigma\sqrt{n}}$

则对任意实数y，有

$$\lim_{n \to \infty} P(Y_n^* \le y) = \Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{t^2}{2}} dt.$$

Let's prove CLT assuming MGF's exist.

Let X_1, X_2 iid $E[X_i] = 0$ & $Var[X_i] = 1$ for all i.

$Z_n = \frac{S_n}{\sqrt{n}}$ Want show $M_{Z_n}(t) \to M_Z(t)$, as $n \to \infty$.

$$M_{Z_n}(t) = E\left[e^{\frac{tS_n}{\sqrt{n}}}\right] \overset{(*)}{=} M_X^n\left(\frac{t}{\sqrt{n}}\right)$$

Let's expand $M_{X_1}(s)$ around 0.

$$M_{X_1}(s) = M_{X_1}(0) + sM'_{X}(0) + \frac{s^2}{2!} M''_{X_1}(0) + \cdots$$

$$= 1 + 0 + \frac{s^2}{2} + \cdots$$

$$(*) = \left[1 + \frac{t^2}{2n} \cdots\right]^n = e^{\frac{t^2}{2}} \to$$