Computing Probability by Conditioning

$P(X \in A) = E(I_A(X))$ $\quad I_A(X) = \begin{cases} +1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$

$E(I_A(X) \mid Y=y) = P(X \in A \mid Y=y)$

$P(A) = P(X \in A) = E(I_A(X)) = E_Y E_{X|Y}(I_A(X) \mid Y)$

$$= \begin{cases} \sum_{y \in S_Y} P(A \mid Y=y)\, P(Y=y) & \text{for } Y \text{ is discrete.} \\ \int_{-\infty}^{\infty} P(A \mid Y=y)\, f_Y(y)\,dy & \text{for } Y \text{ continuous.} \end{cases}$$

ex: X, Y - independent CRV's f_X, f_Y $\cancel{P(X<Y)}$ $\cancel{P(X>Y)}$ $P(X<Y)$

$$= \cancel{\int_{-\infty}^{\infty} P(X>Y \mid Y)} = \int_{-\infty}^{\infty} P(X<Y \mid Y=y)\, f_Y(y)\,dy$$

$$= \int_{-\infty}^{\infty} P(X<y)\, f_Y(y)\,dy \quad \left(\text{Independent } P(X|Y) = P(X)\right)$$

$$\boxed{P(X<Y) = \int_{-\infty}^{\infty} F_X(y) f_Y(y)\,dy}$$

ex: X, Y iid CRV's common PDF f, F

$$P(X<Y) = P(X>Y) = \int_{-\infty}^{\infty} F_Y(y) f_Y(y)\,dy$$

$F_Y(y) = t$

$dF_Y(y) = dt$

$\Rightarrow f_Y(y)\,dy = dt$

$$= \cancel{\int_{-\infty}^{\infty}}$$

$$= \int_0^1 t\,dt = 0.5$$

$$P(X=Y) = \int_{-\infty}^{\infty} P(X=y)\, f_Y(y)\,dy = 0$$

ex: X, Y are two independent RV's (discrete or contin)

$X+Y$

→ X, Y are DRV. Then X+Y will be DRV.

→ So the PMF of Z=X+Y

$$f_Z(z) = P(X+Y=z)$$

$$= \sum_{y \in S_Y} P(X+Y=z \mid Y=y)\, P(Y=y)$$

$$= \sum_{y \in S_Y} P(X=z-y)\, P(Y=y)$$

X, Y are CRV. Then X+Y will be

CDF of z

$$F_Z(z) = P(X+Y \leq z)$$

$$= \cancel{\sum_{y \in S_Y}} \int_{-\infty}^{\infty} P(X+Y \leq z \mid Y=y)\, f_Y(y)\, dy$$

$$= \int_{-\infty}^{\infty} P(X \leq z-y)\, f_Y(y)\, dy \quad \text{(Independent)}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{z-y} f_X(x)\, f_Y(y)\, dx\, dy$$

$x' = x+y$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{z} f_X(x'-y)\, f_Y(y)\, dx'\, dy$$

$$= \int_{-\infty}^{z} \left[\int_{-\infty}^{\infty} f_X(x'-y)\, f_Y(y)\, dy \right] dx'$$

Hence X+Y is CRV with PDF,

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(z-y)\, f_Y(y)\, dy \quad \forall z \in \mathbb{R}$$

classmate

Atleast one is cont. X+Y is CRV

Let X be CRV & Y be DRV

CDF of $Z=X+Y$

$$F_Z(z) = P(X+Y \le z)$$

$$= \sum_{y\in S_Y} P(X+Y \le z \mid Y=y)\, f_Y(y)$$

$$= \sum_{y\in S_Y} P(X \le z-y)\, f_Y(y)$$

$$= \sum_{y\in S_Y} \int_{-\infty}^{z-y} f_X(x) f_Y(y)\,dx \qquad x' = x+y$$

$$= \sum_{y\in S_Y} \int_{-\infty}^{z} f_X(x'-y)\, f_Y(y)\, dx'$$

$$= \int_{-\infty}^{z} \Big(\sum_{y\in S_Y} f_X(x'-y) f_Y(y) \Big) dx'$$

X+Y is CRV with PDF $\sum_{y\in S_Y} f_X(x'-y)\, f_Y(y)$

$$= \sum_{y\in S_Y} f_X(z-y) f_Y(y) \quad \forall z \in \mathbb{R}$$

<u>Def</u>: (X,Y) be Random vector. Then

$$E\big(h(X,Y) \mid (X,Y)\in A\big) = \frac{E\big(h(X,Y)\, I_A(X,Y)\big)}{P((X,Y)\in A)}$$

<u>ex</u>: $X \sim Exp(1)$, $E(X \mid X \ge 2)$.

$$= \frac{E(X\, I_A(X\ge 2))}{P(X\ge 2)} = \frac{\int_2^\infty x e^{-x}\,dx}{1-(1-e^{-2})} = \frac{3e^{-2}}{e^{-2}} = 3$$

classmate

Ex: (X,Y) is uniform on unit sq.

$$E(X \mid X+Y>1) = \frac{E(X\, I_{(1,\infty)}(X+Y))}{P(X+Y>1)}$$

$$= \frac{\int_0^1\int_0^1 x\, I_{(1,\infty)}(x+y)\,dx\,dy}{\int_0^1\int_{1-y}^1 1\,dx\,dy}$$

$$= \frac{\int_0^1\int_{1-y}^1 x\,dy\,dx}{1/2} = \frac{\int_0^1 \left(\frac{1-(1-y)^2}{\cancel{2}}\right)dy}{\cancel{1/2}}$$

$$= \int_0^1 1-(1+y^2-2y)\,dy$$

$$= \left[y^2 - y^3/3\right]_0^1$$

$$= 2/3$$

$X, \{X_n\}$ be a sequence of rv's in $(S, \mathcal{F}, P)$

Def: (Almost sure convergence):

X_n converges almost surely or with probability 1 to rv X, if

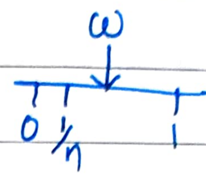
$$P(\omega \in S \mid X_n(\omega) \to X(\omega)) = 1$$

(There ma~~ny~~ be ω' s.t $X_n(\omega') \not\to X(\omega)$, but then $P(\omega')=0$)

$$\{\omega \in S \mid X_n(\omega) \to X(\omega)\} \leftarrow \left\{\begin{array}{l} X_n(\omega_1) \to X(\omega_1) \text{ as } n\to\infty \\ \vdots \end{array}\right.$$

classmate

$S = [0\ 1]$ P uniform. Define $X_n = 1_{[0, \frac{1}{n}]}$ (Indicator)
zero RV $\Rightarrow$ $X(\omega) = 0$ $\forall \omega \in S$.
for fixed $\omega \in [0,1]$
we can find n_0 s.t. $\frac{1}{n} < \omega$ $\forall n \geq n_0$



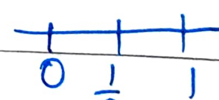
So,
$X_n(\omega) \to 0 = X(\omega)$ as $n \to \infty$.
Hence $X_n \to X$ almost surely.

Thm: Let $\{X_n\}$ and $\{Y_n\}$ be sequence of rv's on $(S, \mathcal{F}, P)$.
Suppose $X_n \to X$ w.p. 1 & $Y_n \to Y$ w.p. 1.
$\to$ $X_n + Y_n \to X + Y$ w.p. 1
$\to$ $X_n Y_n \to XY$ w.p. 1
$\to$ $f(X_n) \to f(X)$ w.p. 1

Def: (convergence in probability) We say that X_n converges in probability to X if $\forall \varepsilon > 0$

$P(|X_n - X| > \varepsilon) \to 0$ as $n \to \infty$. (limit outside)
$\lim_{n \to \infty} P(|X_n - X| > \varepsilon) = 0$

$\to$ $S = [0,1]$, P is uniform. $X_n = n 1_{[0, \frac{1}{n}]}$ $\begin{cases} n & \text{if } \omega \in [0, \frac{1}{n}] \\ 0 & \text{o.w} \end{cases}$
zero RV; $X(\omega) = 0$ $\forall \omega \in S$
for any fixed $\varepsilon > 0$,
$|X_n - X| > \varepsilon$ only in $[0, \frac{1}{n}] \ni \omega$



So, $P(|X_n - X| > \varepsilon) = \frac{1}{n}$
$\lim_{n \to \infty} P(|X_n - X| > \varepsilon) = 0$

Thm: $\{X_n\}$ $\{Y_n\}$ sequences of RV. $X_n \to X$ in Probability
$Y_n \to Y$ in Probability

$\rightarrow$ $X_n + Y_n \to X + Y$

$\rightarrow$ $X_n Y_n \to XY$

$\rightarrow$ $f(X_n) \to f(X)$ | X is continuous

$X_n = 1_{[0,\frac{1}{n}]}$ X zero RV

$X_n \to X$ almost surely (strong)

$\leftarrow$ $X_n \to X$ in probability (Weak)

$X_n = n_{[0,\frac{1}{n}]}$

Def: (Convergence in rth mean) we say X_n convergence in rth mean rv X if

$$E\,|X_n - X|^r \to 0 \text{ as } n \to \infty$$

$$\lim_{n\to\infty} E\,|X_n - X|^r = 0$$

$S = [0,1]$

$X_n = 1_{[0,\frac{1}{n}]}$ claim: $X_n \to X$ in rth mean
(zero RV)

$$|X_n - X| = \begin{cases} 1 & \text{w.p. } \frac{1}{n} \\ 0 & \text{w.p } \left(1-\frac{1}{n}\right) \end{cases}$$

$$E\left(|X_n - X|^r\right) = 1^r \frac{1}{n} + 0^r\left(1 - \frac{1}{n}\right)$$
$$= \frac{1}{n}$$

$$\lim_{n\to\infty} E\left(|X_n - X|^r\right) = 0$$

classmate

Thm: ~~$\{x_n\}, \{y_n\}$~~ $\{X_n\}, \{Y_n\} \to$ Two seq. of rv's

$\to$ If $X_n \to X$ in rth mean, $Y_n \to Y \Rightarrow X_n + Y_n \to X+Y$

$\to$ $f(X_n) \to f(X)$ in rth mean, if f is bounded-continuous.

Def: (Convergence in Distribution)

if $F_n(X) \to F(X)$ as $n \to \infty$.

$\forall x$, F is continuous. F_ns are distribution f^n of X_ns, F is distribution function of X. (CDF's)

$\to$ Here X_ns can be defined on different probability spaces. We are only interested in distribution functions.

ex: $P(X_n = 1/n) = 1$ X – 0 RV

$$F_n(X) = \begin{cases} 0 & \text{if } x < \frac{1}{n} \\ 1 & x \geq \frac{1}{n} \end{cases}$$

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$$

Clearly, $F_n(X) \to F(X)$ as $n \to \infty$ $\forall x \neq 0$ (but $x=0$, F is not continuous) $\forall x$ where F is cont.

So, $X_n \to X$ in distribution

Thm: $\{X_n\}, \{Y_n\}$

$X_n \to X$ in distri, $Y_n \to c$ in probability for some constant.

$\Rightarrow X_n + Y_n \to X + c$ in distri $\Rightarrow X_n Y_n \to cX$ in distri

$\Rightarrow f(X_n) \to f(X)$ in distr f is continuous

$X_n+Y_n \to X+Y$ may not be in distribution
(both in distribution)

Let $X \sim N(0,1)$.
$Y \sim N(0,1)$

Define $X_n = X$, $Y_n = Y$ for $n=1,2,\ldots$

$X_n \to X$ in distribution
$Y_n \to Y$
$Y_n \to X$ in distribution, both X, Y have same

CDF $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$

$X_n + Y_n \to X+Y \sim N(0,2)$ $\quad E(X+Y)=0$

$2X \sim N(0,4)$ $\quad V(X+Y)=2$

$X_n+Y_n \to 2X$ $\quad V(2X)=4$

So, X_n+Y_n doesnot converge to $X+Y$

almost sure $\quad$ mean

$\searrow \quad \swarrow$

Probability

$\downarrow$

Dist

Counter ex. $\quad X_n = \begin{cases} n & \text{if } \omega \in [0,\frac{1}{n}] \\ 0 & \text{ow} \end{cases}$

ex: $S=[0,1]$, P uniform. $X_n = n \mathbb{1}_{[0,\frac{1}{n}]}$. $X_n \to 0$ in Probability & almost surely but not in rth mean

for any fixed $\omega \in S$. We can find n_0
s.t $\frac{1}{n} < \omega$. $\forall n \geq n_0$

Then, $X_n(\omega) \to 0 = X(\omega)$ (zero RV) $\forall \omega \in S$.
$P\{\omega \in S \mid X_n(\omega) \to X(\omega)\} = 1$
So, $X_n \to X$ almost surely.

$0 \quad \frac{1}{n} \quad \omega \quad 1$

classmate

$$E|X_n - X|^r = n^r \frac{1}{n} + 0^r \left(\frac{n-1}{n}\right)$$

$$= n^{r-1}$$

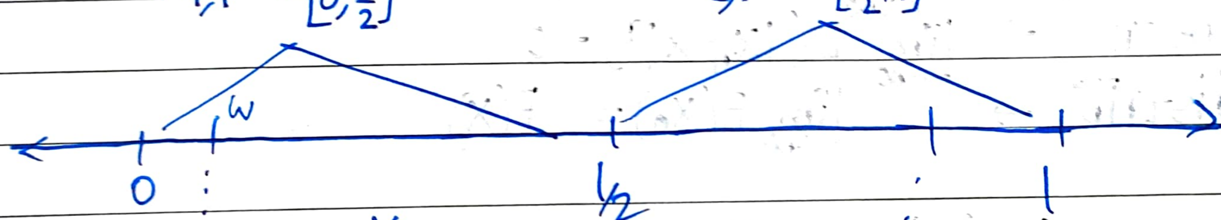
$$\to \begin{cases} 1 & \text{if } r=1 \\ \infty & \text{if } r>1 \end{cases} \text{ as } n\to\infty.$$

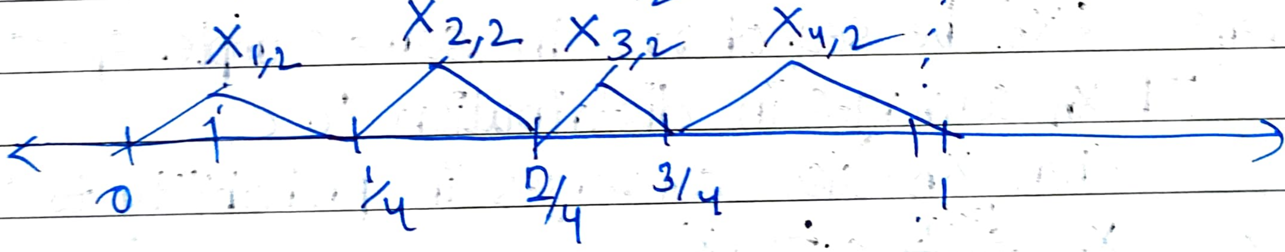
$X_n \not\to X$ in rth mean for $r \geq 1$.

ex: $X_{1,1} = 1_{[0,\frac{1}{2}]}$, $X_{2,1} = 1_{[\frac{1}{2},1]}$, $X_{1,2} = 1_{[0,\frac{1}{4}]}$, $X_{2,2} = 1_{[\frac{1}{4},\frac{1}{2}]}$,

$X_{3,2} = 1_{[\frac{1}{2},\frac{3}{4}]}$, $X_{4,2} = 1_{[\frac{3}{4},1]}$ ..., $X_{m,n} = 1_{[\frac{m-1}{2^n}, \frac{m}{2^n}]}$

$X_{1,1} = 1_{[0,\frac{1}{2}]}$ $X_{2,1} = 1_{[\frac{1}{2},1]}$





X-Zero RV, $P(|X_{m,n} - X| > \varepsilon) = \frac{1}{2^n}$

$\lim_{n\to\infty} P(|X_{m,n} - X| > \varepsilon) = 0 \to$ converge in Probability

$E|X_{m,n} - X|^r = 1^r \cdot \frac{1}{2^n} \Rightarrow \lim_{n\to\infty} E|X_{m,n} - X|^r = 0 \to$ converges in rth mean

for fixed $w \in S$, Then $\exists$ sub seq of seq. of real nos $\{X_{m,n}(w)\}$ converges to 1. And $\exists$ sub seq. that converges to 0. So, $\{X_{m,n}(w)\}$ doesnot converge for all w, $P\{w \in S \mid X_{m,n}$ converges$\} = 0$

doesnot converge almost surely.

classmate

ex: $X \sim N(0,1)$ rv

$-X \sim N(0,1)$

$X_n = X \quad X_n \not\to X$ in distribution.

Now, $P(|X_n - -X| \leq \varepsilon) = P(|X_n + X| \leq \varepsilon)$

$= P(|2X| \leq \varepsilon)$

$= P(-\varepsilon/2 \leq x \leq \varepsilon/2)$

$= 2\Phi(\varepsilon/2) - 1 \neq 1$

So, not converges in Probability

as $\rightarrow P(\omega \mid X_n(\omega) \rightarrow X(\omega)) = 1$

Probability $\rightarrow \lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0$

rth mean $\rightarrow \lim_{n\to\infty} E(|X_n - X|^r) = 0$

distrib^n $\rightarrow \lim_{n\to\infty} F(X_n) = F(X)$

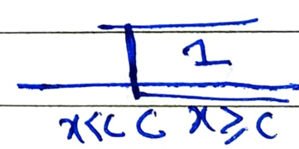
<u>Theorem:</u> $\{X_n\}$ be seq. of rv's in same sample space.

If X_n converges in distribution to c,

then X_n also converges in probability to c.

$\rightarrow \quad X_n \xrightarrow{distri} c \rightarrow (X)$

$$F(X_n) = F(X) = \begin{cases} 0 & \text{if } x < c \\ 1 & \text{if } x \geq c \end{cases}$$



(degenerate distrib^n)

fix $\varepsilon > 0$

To show $P(|X_n - X| > \varepsilon) \rightarrow 0$

$P(|X_n - X| > \varepsilon) \geq 0$

$\Rightarrow P(X_n - c > \varepsilon) + P(X_n - c < -\varepsilon) \geq 0$

$\Rightarrow 1 - P(X_n \leq c + \varepsilon) + P(X_n < c - \varepsilon)$

$\Rightarrow 0 \leq 1 - F_n(c+\varepsilon) + F_n(c-\varepsilon)$

$\Rightarrow \quad \leq 0 \longrightarrow$ So, converges in probability

Theorem: $\{X_n\}$ be seq. of rv's with MGF's $M_n(t)$.
X be rv with MGF $M(t)$. If $M_n(t) \to M(t)$
$\forall t$ in an open interval containing 0, then
$X_n \xrightarrow{distr^n} X$

ex: $X_n \sim Bin(n, P_n)$, $P_n \to 0$ & $nP_n \to \lambda\ (>0)$
$X \sim Poi(\lambda)$

MGF of X_n is $E(e^{tX_n}) = \sum_{x_n=0}^{n} e^{tx_n} \binom{n}{x_n} P_n^{x_n} (1-P_n)^{n-x_n}$

$= (1-P_n + P_n e^t)^n$

MGF of X is $E(e^{tX}) = \sum_{x_n=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} e^{tx}$

$= \sum_{x_n=0}^{\infty} e^{-\lambda} \frac{(\lambda e^t)^x}{x!}$

$= e^{-\lambda} e^{\lambda e^t} \sum_{x=0}^{\infty} \cancel{e^{-\lambda e^t} \frac{(\lambda e^t)^x}{x!}}\ 1$

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

$\left(1+\frac{x}{n}\right)^n \to e^x$

$M_n(t) = (1+P_n(e^t-1))^n$
$= \left(1+\frac{\lambda}{n}(e^t-1)\right)^n$

$\Longrightarrow e^{\lambda(e^t-1)} = M(t)$ as $n \to \infty$ $\forall t$

So, $X_n \xrightarrow{dist} X$

classmate

Theorem: $\{X_n\}$ be seq. of DRV's with PMF's $f_n(\cdot)$.
X be DRV with PMF $f(\cdot)$.
If $f_n(x) \to f(x)\ \forall x$, then $X_n \xrightarrow{dist} X$.

ex: $f_n(x) = \binom{n}{x} p_n^x (1-p_n)^{n-x} = \frac{n!}{x!(n-x)!} p_n^x (1-p_n)^{n-x}$

$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$

$$= \frac{n!}{x!(n-x)!} \frac{\lambda^x}{n^x} \frac{(n-\lambda)^{n-x}}{n^{n-x}}$$

$$= \frac{n(n-1)\cdots(n-x+1)}{n^x}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-x}\frac{\lambda^x}{x!}$$

$$= (1)\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-x}\frac{\lambda^x}{x!}$$

$$\longrightarrow 1 \cdot e^{-\lambda} \cdot 1 \cdot \frac{\lambda^x}{x!}$$

$$= \frac{e^{-\lambda}\lambda^x}{x!} = f(x)$$

Theorem: $\{X_n\}$ be seq. of CRV with PDF $f_n(\cdot)$.
X be CRV with PDF $f(\cdot)$.
$f_n(x) \to f(x)\ \forall x$. Then $X_n \xrightarrow{dist} X$

Ex: $X_n \sim U(0, 1+1/n)$ $X \sim U(0,1)$.

$$f_n(x) = \begin{cases} \frac{1}{1+1/n} & \text{if } 0<x<1+\frac{1}{n} \\ 0 & \text{ow.} \end{cases}$$

$$f(x) = \begin{cases} 1 & \text{if } 0<x<1 \\ 0 & \text{ow.} \end{cases}$$

$f_n(x) \to f(x)\ \forall x$ as $n \to \infty$.

$X_n \xrightarrow{dist} X$

Limit Theorems

Theorem: (Strong law of large numbers)

$\{X_n\}$ seq. of iid RV's with mean μ.
$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Then $\{\bar{X}_n\}$ converges to μ almost-
-surely. (Then $\{\bar{X}_n\} \xrightarrow{Prob} \mu$)

$\frac{1}{n}\sum_{i=1}^{n} h_i \rightarrow$ size n (random)

ex: Bernoulli ~~probab~~ proportion converges to success probability.

$E \rightarrow$ fixed event.

$$X_i = \begin{cases} 1 & \text{if } E \text{ occurs in ith trial} \\ 0 & \text{ow.} \end{cases}$$

$$\bar{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n} \rightarrow \mu = E(X_1) = P(E).$$

$\rightarrow \frac{\text{no. of times } E \text{ occurs}}{n} = \text{proportion}$

ex: Monte Carlo Integ^n.

$$I = \int_a^b h(x)\,dx$$

we can write, $I = (b-a)\int_a^b \frac{h(x)}{b-a}\,dx$

$$= (b-a)\,E(Y).$$

where $Y = h(X)$
$X \sim U(a,b)$

$\{X_n\}$ be seq. of iid rv's with distribution $U(a,b)$
& assume $Y_n = h(X_n)$, $n = 1, 2, 3, \ldots$

using SLLN,

$$\overline{Y}_n = \frac{Y_1+Y_2+\cdots+Y_n}{n} \rightarrow E(Y_1) = \frac{I}{b-a}$$

$$= \frac{\sum_{i=1}^{n} h(X)}{n}$$

$$\Rightarrow \frac{b-a}{n} \sum_{i=1}^{n} h(X_i) \rightarrow I \text{ (almost surely)}$$

Now generate N random numbers from $U(a,b)$.
(can be done by R, Python, Matlab etc)
Now I can be approximated using SLLN as

$$\frac{b-a}{N} \sum_{i=1}^{N} h(X_i)$$

higher order finite
$\Rightarrow$ lower also
$V = E(X^2) - E(X)$

<u>Theorem</u>: (Central Limit Theorem)

Let $\{X_n\}$ be seq. of iid RV's with mean μ & variance $\sigma^2 < \infty$. Then, as $n \rightarrow \infty$.

$$P\left(\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} \leq a\right) \rightarrow \Phi(a) = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \, dt.$$

given, $E(X_i) = \mu$, $V(X_i) = \sigma^2$

$$E(\overline{X}_n) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n} E\left(\sum_{i=1}^{n} X_i\right) = \mu$$

$$V(\overline{X}_n) = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} n\sigma^2$$

$$= \sigma^2/n$$

Standardised.

$$\frac{\overline{X}_n - E(\overline{X}_n)}{\sqrt{Var(\overline{X}_n)}} = \frac{\overline{X}_n - \mu}{\sqrt{\sigma^2/n}} = \frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma}$$

example 3:

$X_n \sim Bin(n,p)$. To show $P\left(\frac{X_n - np}{\sqrt{np(1-p)}} < a\right) \to \Phi(a)$

let $Y_i \sim Bernoulli(p)$ independent. $E(Y_i) = p$

$X_n = Y_1 + Y_2 + \dots + Y_n$. $V(Y_i) = p(1-p)$

$$P\left(\frac{\overline{Y}_n - p}{\sqrt{\frac{p(1-p)}{n}}} < a\right) \longrightarrow \Phi(a)$$

$$\Rightarrow \frac{\overline{Y}_n - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{\frac{X_n}{n} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{X_n - np}{\sqrt{np(1-p)}}$$

same, so converge in distribution.

$$\Rightarrow P\left(\frac{X_n - np}{\sqrt{np(1-p)}} < a\right) \to \Phi(a)$$

ex: mean $= 40$, var $= 400$. $n = 25$.

$X_i =$ ith Battery life.

$P(X_1 + \dots + X_{25} > 1100)$

$= P\left(\overline{X}_{25} > \frac{1100}{25}\right)$

$= P\left(\frac{\overline{X}_{25} - 40}{\frac{20}{\sqrt{25}}} > \frac{44 - 40}{\frac{20}{\sqrt{25}}}\right) = 1 - \Phi($

classmate

Bivariate normal

Def: A two dimensional random vector $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ is said to have a bivariate normal distribution if $aX_1 + bX_2$ is univariable normal $\forall\, (a,b) \in \mathbb{R}^2 \setminus (0,0)$

Theorem:

X has bivariate normal distribution $\rightarrow X_1$ & X_2 is Univariate Normal. All moments of X_1, X_2 exist

$$\mu = E(X) = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

$$\Sigma = Var(X) = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}.$$

$\sigma_{11} = Cov(X_1, X_1)$

$\sigma_{12} = Cov(X_1, X_2)$

$\sigma_{21} =$ $\sigma_{22} = Cov(X_2, X_2)$

$$|\rho| \leq 1 \Rightarrow \frac{|Cov(X_1, X_2)|}{\sqrt{V(X_1)V(X_2)}} < 1 \Rightarrow Cov(X_1, X_2) \leq \sqrt{V(X_1)V(X_2)}$$

also exists.

$\sigma_{11} = V(X_1) > 0$

$\sigma_{22} = V(X_2) > 0$ (Symmetric positive definite)

$x'Ax \geq 0 \rightarrow$ Semi Positive definite

$x'Ax \leq 0 \rightarrow$ Semi Negative definite

$x'Ax > 0 \rightarrow$ Positive definite

$x'Ax < 0 \rightarrow$ Negative definite

Theorem:

X - Bivariate normal distribution

$\mu = E(X)$

$\Sigma = Var(X)$

for any fixed $u = (a,b) \in \mathbb{R}^2 \setminus (0,0)$

$$u'X \sim N(u'\mu, u'\Sigma u)$$

$$E(u'X) = a\mu_1 + b\mu_2 = (a,b)\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix} = u'\mu$$

$$V(u'X) = a^2 V(X_1) + b^2 V(X_2) + 2ab\, Cov(X_1, X_2)$$

$$= a^2\sigma_{11} + b^2\sigma_{11} + 2ab\,\sigma_{12}$$

$$= u'\Sigma u$$

Theorem:

MGF of X $\quad M_X(t) = E(e^{t'X})$ $\qquad t'X \sim N(t'\mu, t'\Sigma t)$

$$= M_{t'X}(1)$$

$$= e^{t'\mu + \frac{1}{2}t'\Sigma t}$$

Theorem: Thus bivariate normal distribution is completely defined by μ, Σ

$$X \sim N_2(\mu, \Sigma)$$

Def: $X \sim$ Bivariate Normal, if $X = \mu + AY$ $\quad A - 2\times 2$ real

$Y = (Y_1, Y_2)$, Y_1, Y_2 are iid $N(0,1)$

$E(X) = \mu$

$\Sigma = AA'$

Theorem

If $X \sim N_2(\mu, \Sigma)$

then $X_1 \sim N(\mu_1, \sigma_{11})$

$X_2 \sim N(\mu_2, \sigma_{22})$

Remark: Converse of above is not true

let $X \sim N(0,1)$

let Z be a DRV, independent of X such that

$P(Z=1) = \frac{1}{2} = P(Z=-1)$

Claim: $Y = ZX \sim N(0,1)$

CDF of Y for $y \in \mathbb{R}$

$$
\begin{aligned}
P(Y \le y) &= P(ZX \le y) \\
&= P(ZX \le y \mid Z=1)\, P(Z=1) + P(ZX \le y \mid Z=-1)\, P(Z=-1) \\
&= \frac{1}{2} P(X \le y) + \frac{1}{2} P(X \ge -y) \\
&= \frac{1}{2} \Phi(y) + \frac{1}{2} \Phi(y) \\
&= \Phi(y) \Rightarrow Y \sim N(0,1)
\end{aligned}
$$

Claim: $X \sim N(0,1)$, $Y \sim N(0,1)$ But $\binom{X}{Y}$ doesnot follow ~~Bino~~ variate Normal distribution

$$
\begin{aligned}
P(X+Y=0) &= P(X+ZX=0) \\
&= P(Z=-1) \\
&= \frac{1}{2}
\end{aligned}
$$

$X+Y$ doesnot follow univariate Normal distribution. So $\binom{X}{Y}$ is not Bivariate Normal.

Remark:

If $X \sim N_2(\mu, \Sigma)$ and $Cov(X_1, X_2) = 0$, then X_1 and X_2 are independent.

$Cov(X_1, X_2) = 0$

$$\Sigma = \begin{bmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{bmatrix}$$

So, $M_X(t) = e^{t'\mu + \frac{1}{2}t'\Sigma t}$

$= e^{t_1\mu_1 + \frac{1}{2}t_1^2\sigma_{11}} \, e^{t_2\mu_2 + \frac{1}{2}t_2^2\sigma_{22}}$

$= M_{X_1}(t_1) \, M_{X_2}(t_2)$

So, X_1, X_2 are independent.

$$t'\mu = (t_1 \; t_2)\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = t_1\mu_1 + t_2\mu_2$$

$$(t_1 \; t_2)\begin{pmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{pmatrix}\begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = t_1^2\sigma_{11} + t_2^2\sigma_{22}$$

classmate