Key concepts: ① random variables ② probability distributions

③ independence
④ Expectation
⑤ Variance

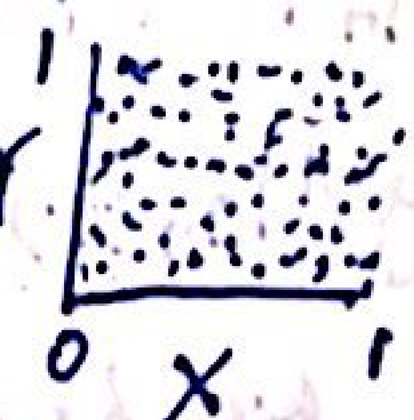
$\boxed{\begin{array}{l}\text{discrete: prob. mass fun. (pmf)} \\ \text{cont: prob. density fun (pdf)}\end{array}}$

cdf: $\mathbb{P}(X \le x) = F(x)$

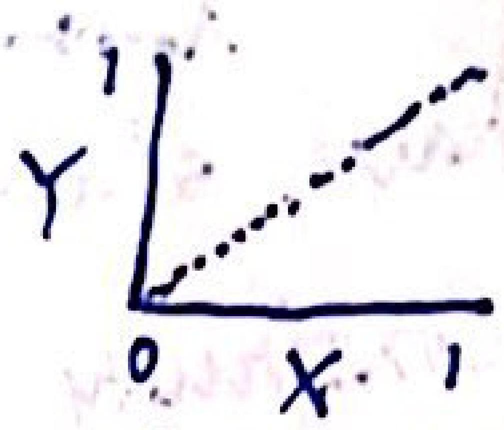
MGF $M_X(t) = E[e^{tX}]$

bivariate: a pair of r.v.s (X, Y)

$X \sim U[0,1]$
$Y \sim U[0,1]$



if $X \perp Y$
$X|Y \sim U[0,1]$



if $X = Y$.
$X|Y$ = point mass at Y

joint MGF: $M_{XY}(s,t) = E[e^{sX+tY}]$

marginal MGFs: $M_X(s) = E[e^{sX}]$, $M_Y(t) = E[e^{tY}]$

e.g. $M_X(0) = E[e^{0X}]$; $M_X(1) = E[e^{1X}]$

$X \perp Y \iff M_{X,Y}(s,t) = M_X(s) \cdot M_Y(t)$

eg. $X_1, \cdots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$
$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ is indep of $X_i - \bar{X}_n$ for all $i = 1, \cdots, n$

Check combined lecture notes Pages 13-14: Binomial(n,p) can be approx. by $N(np, np(1-p))$ when n is large

Example: a sample of $n=100$ values from a population w/ mean $\mu = 500$ and standard deviation $\sigma = 80$.

a. What is the probability that the sample mean is in the interval $(490, 510)$?

$X_1, \cdots, X_n \overset{iid}{\sim}$ $\mu = 500$, $\sigma = 80$, $n = 100$

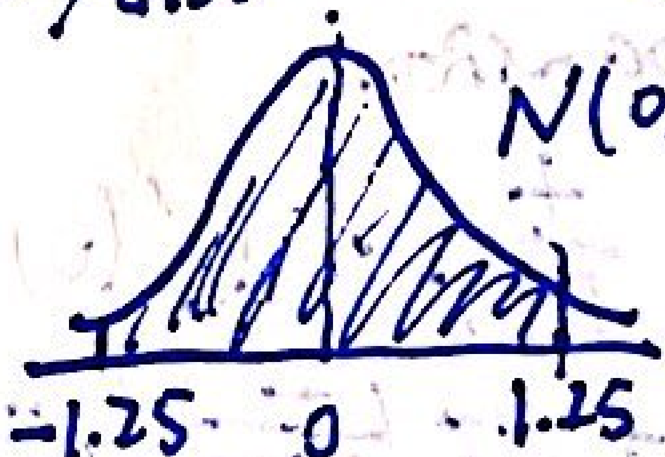
CLT: $\bar{X}_n \overset{approx}{\sim} N(E[\bar{X}_n], var(\bar{X}_n))$
$= N(\mu, \frac{\sigma^2}{n})$
$= N(500, \frac{80^2}{100})$

$E[\bar{X}_n] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n}\sum_{i=1}^{n}\mu = \mu$

$Var[\bar{X}_n] = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \overset{ind}{=} \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} var(X_i) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{1}{n}\sigma^2$

Ans: $\mathbb{P}(490 < \bar{X}_n < 510)$
$= \mathbb{P}\left(\frac{490-500}{8} < \frac{\bar{X}_n - 500}{8} < \frac{510-500}{8}\right)$
$= \mathbb{P}\left(-1.25 < \frac{\bar{X}_n - 500}{8} < 1.25\right)$
$= \cancel{////} = 0.7887$

$\frac{\bar{X}_n - 500}{80/\sqrt{100}} = \frac{\bar{X}_n - 500}{8} \overset{approx}{\sim} N(0,1)$



n iid w/ mean μ, sd σ ; $P(a < \overline{X}_n < b)$ when n is large

Ans: $\overline{X}_n \overset{\text{approx}}{\sim} N(\mu, \frac{\sigma^2}{n}) \Leftrightarrow \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \overset{\text{approx}}{\sim} N(0,1)$

$$P(a < \overline{X}_n < b) = P\left(\frac{a-\mu}{\sigma/\sqrt{n}} < \frac{\overline{X}_n-\mu}{\sigma/\sqrt{n}} < \frac{b-\mu}{\sigma/\sqrt{n}}\right) \approx \Phi\left(\frac{b-\mu}{\sigma/\sqrt{n}}\right) - \Phi\left(\frac{a-\mu}{\sigma/\sqrt{n}}\right)$$

when Φ is the CDF of $N(0,1)$:

$\Phi(x) = P(Z \le x)$ where $Z \sim N(0,1)$

Φ = "pnorm()"

Gamma distribution : $X \sim \text{Gamma}(\alpha, \beta)$ cont. $X \geq 0$

$\alpha, \beta > 0$

pdf : $f(x) = \frac{x^{\alpha-1} \cdot e^{-x/\beta}}{\beta^\alpha \cdot \Gamma(\alpha)}$

$$\int_0^\infty f(x)\,dx = 1 \Leftrightarrow \int_0^\infty \frac{x^{\alpha-1} e^{-x/\beta}}{\beta^\alpha \Gamma(\alpha)}\,dx = 1 \Leftrightarrow \Gamma(\alpha) = \int_0^\infty \frac{x^{\alpha-1} e^{-x/\beta}}{\beta^\alpha}\,dx$$

Special cases: ① $\alpha = 1$, $\beta = \frac{1}{\lambda}$, $\lambda > 0$; then $f(x) = \frac{e^{-\lambda x}}{(\frac{1}{\lambda}) \cdot \Gamma(1)} = \lambda e^{-\lambda x}$

— exponential dist

② $\alpha = \frac{1}{2}$, $\beta = 2$. $f(x) = \frac{x^{-\frac{1}{2}} e^{-x/2}}{2^{\frac{1}{2}} \Gamma(\frac{1}{2})}$ — χ_1^2 dist

Theorem 1: Let $Z \sim N(0,1)$. Then $Z^2 \sim \chi_1^2$.

Pf : idea: CDF of Z^2: $\underset{F(x)}{\underbrace{P(Z^2 \le x)}} \Rightarrow$ pdf of Z^2 $f(x) = F'(x)$

$\Rightarrow$ show that $f(x) = \frac{x^{-\frac{1}{2}} e^{-x/2}}{2^{\frac{1}{2}} \Gamma(\frac{1}{2})}$