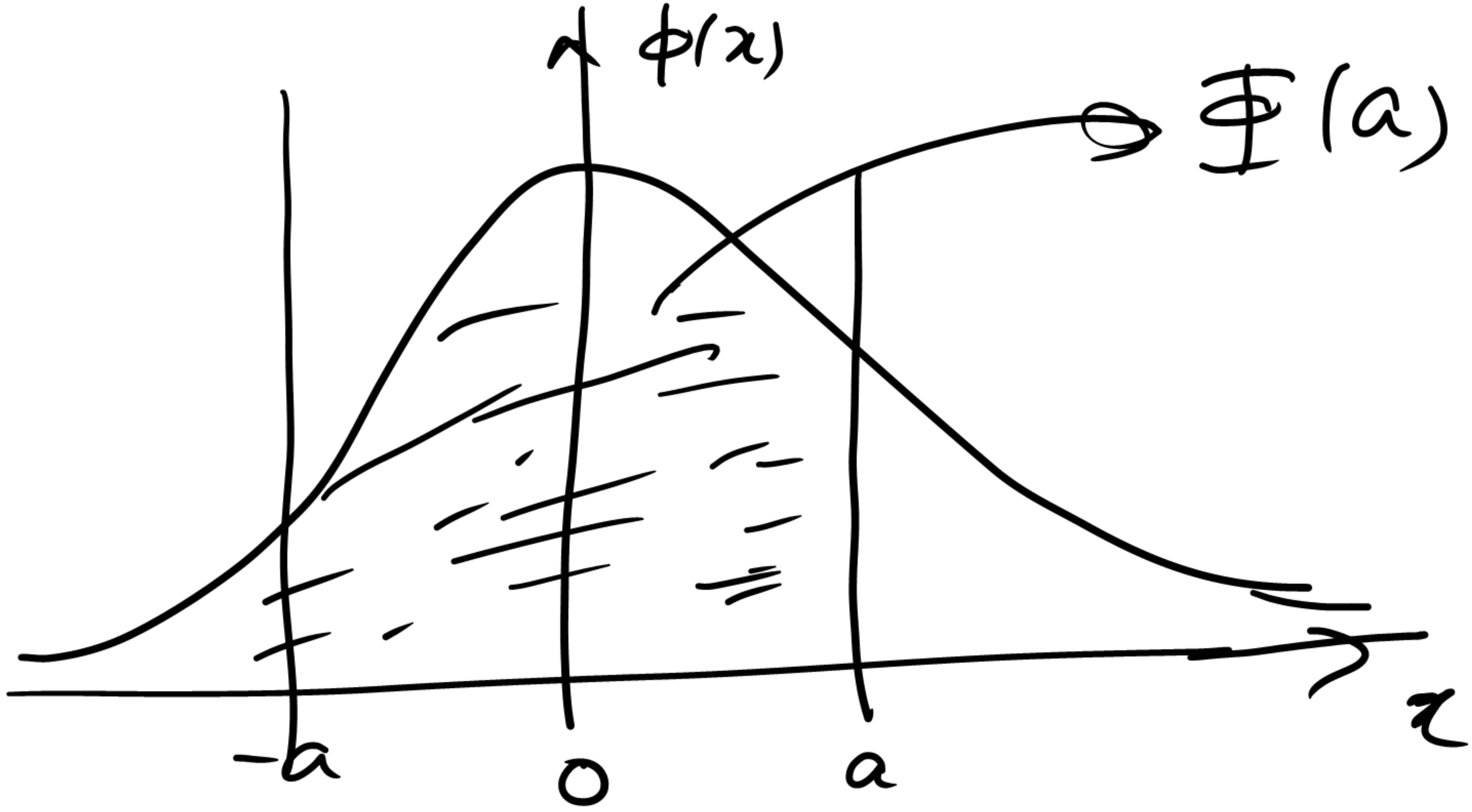
The pdf of a standard normal r.v. Z is $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, $x \in \mathbb{R}$

The cdf of standard normal r.v. is $\Phi(x) = \int_{-\infty}^{x} \phi(t)\,dt$



$$\Phi(-a) + \Phi(a) = 1 \quad , \quad \Phi(0) = \frac{1}{2}$$

$$\Phi(1) = 0.8413 , \quad \Phi(2) = 0.9772$$

$$\Phi(3) = 0.9987$$

$$P(-1 < Z < 1) = \Phi(1) - \Phi(-1)$$

$$= 2\Phi(1) - 1$$

$$= 0.6826$$

$$X \sim N(\mu, \sigma^2), \quad Z = \frac{X-\mu}{\sigma} \sim N(0,1)$$

$$\Rightarrow P\left(-1 < \frac{X-\mu}{\sigma} < 1\right) = 0.6826$$

$$\Rightarrow P(\mu - \sigma < X < \mu + \sigma) = 0.6826$$

So in a normal distⁿ 68.26% of the observations lie within $\mu \pm \sigma$ limits

$$P(-2 < Z < 2) = \Phi(2) - \Phi(-2)$$

$$= 2\Phi(2) - 1$$

$$= 0.9544$$

$$P\left(-2 < \frac{X-\mu}{\sigma} < 2\right) = 0.9544$$

$$P(\mu - 2\sigma < X < \mu + 2\sigma) = 0.9544$$

So in a normal distⁿ 95.44% observations lie with $\mu \pm 2\sigma$ limits

$$P(-3 < Z < 3) = \Phi(3) - \Phi(-3)$$

$$= 2\Phi(3) - 1 = 0.9974$$

$$P\left(-3 < \frac{X-\mu}{\sigma} < 3\right) = 0.9974$$

$$\Rightarrow P(\mu - 3\sigma < X < \mu + 3\sigma) = 0.9974$$

So in a normal dist^n 99.74% of the observations lie with $\mu \pm 3\sigma$ limits.

Exercise: Suppose the price of an item is a normal r.v. with $\mu = 10$ and

$\sigma = 0.25$. What is the shortest interval which has prob. 0.95 of including the price of the item?

$$P(10 - a < X < 10 + a) = 0.95$$

$$\Rightarrow \quad P\left(-4a < \overset{Z}{\frac{X-10}{0.25}} < 4a\right) = 0.95$$

$$\Rightarrow \quad \Phi(4a) - \Phi(-4a) = 0.95$$

$$\Rightarrow \quad 2\Phi(4a) - 1 = 0.95$$

$$\Rightarrow \quad \Phi(4a) = 0.975$$

$$\Rightarrow \quad 4a = 1.96 \Rightarrow a = 0.49$$

So the shortest interval is $(9.51, 10.49)$.

<u>Theorem</u> : Let $X \sim \text{Bin}(n, p)$.

As $n \to \infty$, the distⁿ of

$$Z = \frac{X - np}{\sqrt{npq}}$$ converges to $N(0,1)$.

<u>Proof</u> : $M_Z(t) = E\left(e^{tZ}\right)$

$$= E\left[e^{t\left(\frac{X-np}{\sqrt{npq}}\right)}\right] = e^{-\frac{npt}{\sqrt{npq}}} M_X\left(\frac{t}{\sqrt{npq}}\right)$$

$$= e^{-\frac{npt}{\sqrt{npq}}} \left[q + p e^{t/\sqrt{npq}} \right]^n$$

$$\log M_Z(t) = -\frac{npt}{\sqrt{npq}} + n \log \left(\qquad \right)$$

$$= -\frac{npt}{\sqrt{npq}} + n \log \left[1 + p \left(e^{t/\sqrt{npq}} - 1 \right) \right]$$

$$= -\frac{npt}{\sqrt{npq}} + n \log\left[1 + p\left(1 + \frac{t}{\sqrt{npq}} + \frac{t^2}{2npq}\right.\right.$$

$$\left.\left. + \frac{t^3}{3!\,(npq)^{3/2}} + \cdots - 1\right)\right]$$

$$= -\frac{npt}{\sqrt{npq}} + n \log\left[1 + \left(\frac{pt}{\sqrt{npq}} + \frac{pt^2}{2npq} + \cdots\right)\right]$$

$$= -\frac{npt}{\sqrt{npq}} + n\left[\left(\frac{pt}{\sqrt{npq}} + \frac{pt^2}{2npq} + \cdots\right)\right.$$

$$-\frac{1}{2}\left(\frac{pt}{\sqrt{npq}}+\frac{pt^2}{2npq}+\cdots\right)^2+\cdots\Bigg]$$

$$=\left(\frac{t^2}{2q}-\frac{p\,t^2}{2q}\right)+\text{terms containing powers of }\left(\frac{1}{n}\right)$$

$$\longrightarrow +\frac{t^2}{2}$$

So $\quad M_Z(t)\longrightarrow e^{t^2/2}$

which is mgf of $N(0,1)$.

<u>Example</u>: The prob that a person recovers from a serious disease is 0.4. If 100 persons are getting treated what is the prob that not more than 30 will recover?

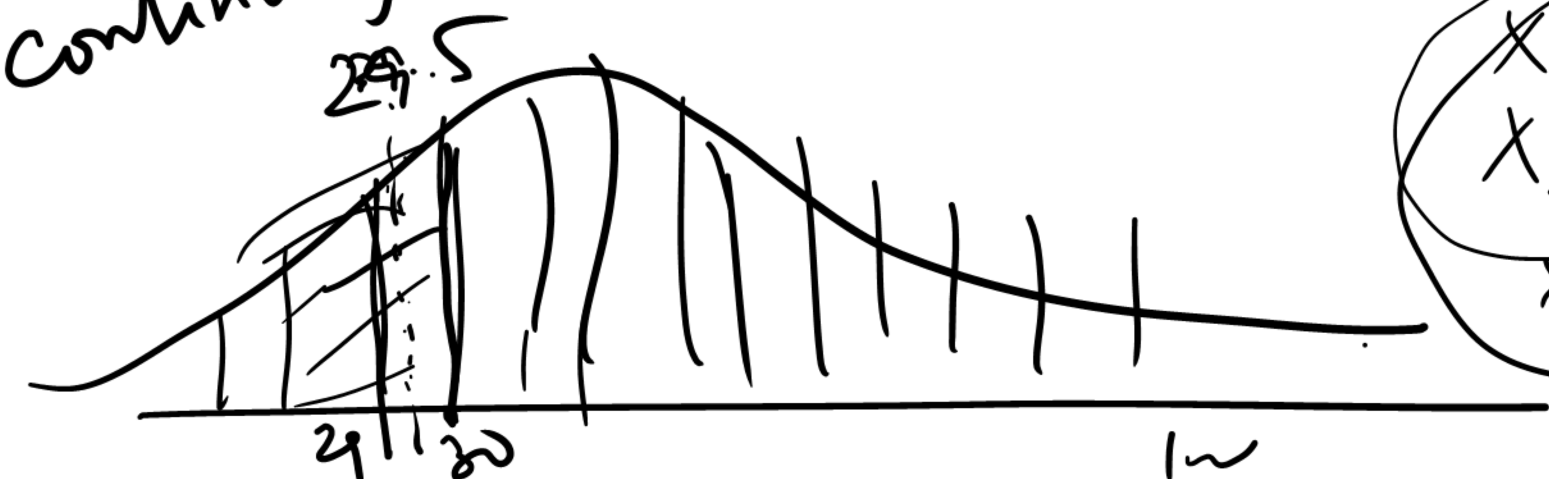
Here n is large but p is not small. So we cannot use

Poisson approximation

$$P(X \leq 30) = \sum_{x=0}^{29} \binom{100}{x} (0.4)^{x} (0.6)^{n-x}$$

$$np = 40, \quad npq = 24, \quad \sqrt{npq} \simeq 4.899$$

continuity correction $\quad X < 30$



$X < 29$

$X \leq 29$

$X < 30$

$$P(X < 30) \cong P(X \leq 30.5)$$

$$\cong P\left(\frac{X-40}{4.899} \leq \frac{30.5-40}{4.899}\right)$$

$$= P(Z \leq -2.14) = \Phi(-1.94)$$

$$= 0.0262$$

<u>Theorem</u>: Let $X \sim P(\lambda)$ and $Z = \frac{X-\lambda}{\sqrt{\lambda}}$. As $\lambda \to \infty$, the

dist^n of Z converges to $N(0,1)$.

Proof: $M_Z(t) = E\left(e^{tZ}\right)$

$$= E\left[e^{t\left(\frac{X-\lambda}{\sqrt{\lambda}}\right)}\right] = e^{-t\sqrt{\lambda}} M_X\left(\frac{t}{\sqrt{\lambda}}\right)$$

$$= e^{-t\sqrt{\lambda}} \; e^{\lambda\left(e^{t/\sqrt{\lambda}} - 1\right)}$$

$$= e^{-t\sqrt{\lambda}} \cdot e^{\lambda\left(\frac{t}{\sqrt{\lambda}} + \frac{t^2}{2\lambda} + \frac{t^3}{6\lambda^{3/2}} + \cdots\right)}$$

$$= e^{t^2/2 + \text{terms with negative powers of } \lambda}$$

$$\longrightarrow e^{t^2/2} \quad \text{which is mgf of } N(0,1)$$

<u>Example</u>: $X \sim \mathcal{P}(15)$

$$P(X \le 10) \simeq P(X \le 10.5)$$

$$\simeq P\left(\frac{X-15}{\sqrt{15}} \le \frac{10.5-15}{\sqrt{15}}\right)$$

$$= P(Z \le -1.16) = \Phi(-1.16)$$
$$= 0.1230$$

Examples. 1. A vaccine for desensitizing patients to bee stings is to be packed with 3 vials in each box. Each vial is checked for strength before packing. The prob. that a vial meets the specifications is 0.9. Let X denote the number of vials that must be checked to fill a box. Find the prob that out of 10 boxes to be filled only three boxes need exactly 3 testings each.

$$X \sim NB(3, 0.9)$$

$$p_X(k) = \binom{k-1}{2} (0.9)^3 (0.1)^{k-3}, \quad k = 3, 4, \ldots$$

$$E(X) = \frac{r}{p} = \frac{3}{0.9} = \frac{10}{3}$$

$$V(X) = \frac{rq}{p^2} = \frac{3 \times 0.1}{(0.9)^2} = \frac{10}{27}.$$

Y denote the number of boxes which need exactly 3 testings

Then $Y \sim \text{Bin}(10, p^*)$

$p^* = P($ a box needs exactly 3 testing$)$

$$= (0.9)^3 = 0.729$$

$$P(Y=3) = \binom{10}{3}(0.729)^3(.271)^7$$

$$\cong 0.00499$$

2. In a battle, missiles are launched to destroy an enemy target. The prob. that a missile will hit the target is 0.5. How many missiles should be

launched so that the prob of destroying the target is at least 0.99? (2 successful hits are required to destroy the target)
Suppose n missiles are launched and X is number of missiles hitting the target.

$$P(X \geq 2) \geq 0.99$$

$$\Rightarrow 1 - P(X=0) - P(X=1) \geq 0.99$$

$$\Rightarrow \left(\frac{1}{2}\right)^n + \binom{n}{1}\left(\frac{1}{2}\right)^n \leq 0.01$$

$$\Rightarrow \quad 2^n \geqslant 100(n+1).$$

The smallest value of n for which this inequality is satisfied is

$$n = 11.$$

3.. A purchaser of electronic components buys them in lots of size 10. The policy is to inspect 3 components randomly from a lot, If all 3 are nondefective the lot is accepted. It is known that 30% of lots have 4 defective components and 70% of lots have 1 defective. What is the proportion of lots which will be rejected.

P(lot is accepted) =

P(lot is accepted | lot has 4 def) P(lot has 4 def)

$+ P(\text{lot is accepted} \mid \text{lot has 1 def}) \, P(\text{lot has 1 def})$

$$= \frac{\binom{4}{0}\binom{6}{3}}{\binom{10}{3}} \times \frac{3}{10} + \frac{\binom{1}{0}\binom{9}{3}}{\binom{10}{3}} \times \frac{7}{10}$$

$$= 0.54$$

$P(\text{Rejecting the lot}) = 0.46$

46% of lots are rejected.

4. Buses arrive at a specified stop at 15 minute intervals starting at 7:00 a.m. A passenger comes to the stop at a random time between 7:00 - 7:30 a.m. Find the prob that

(i) he/she waits less than 5 minutes

(ii) at least 12 minutes

$X \to$ time past 7:00 a.m. that the

passenger arrives at the stop

Then $X \sim U(0, 30)$ $f(x) = \frac{1}{30}$,
$0 < x < 30$

P(waiting time is less than 5 min)

$= P(10 < X < 15) + P(25 < X < 30)$

$$= \int_{10}^{15} \frac{1}{30}\,dx + \int_{25}^{30} \frac{1}{30}\,dx = \frac{1}{3}$$

P(waiting time is at least 12 min)

$$= P(0 < X < 3) + P(15 < X < 18)$$

$$= \int_0^3 \frac{1}{30}\,dx + \int_{15}^{18} \frac{1}{30}\,dx = \frac{1}{5}$$

5. A large IC has multiple copies of circuits. If the circuit fails, the IC knows how to repair itself. The average no of defects per IC is 300. What is the prob that at

most 4 defects will be found in a randomly selected area comprising 2% of the total surface area of IC.?

$$X \sim P(300)$$

$$\lambda t \rightarrow 300 \times 0.02 = 6$$

$$P(X \leq 4) = \sum_{k=0}^{4} \frac{e^{-6} 6^{k}}{k!} \approx 0.285$$

8. The lifetime (in years) of a component follows a Weibull distⁿ with failure rate $z(t) = t^3$. Find the prob that the item survives to age 2? What is the prob that the life is between 0.4 to 1.4? What is the prob that a 1 year item survives to 2 years?

$$R(t) = e^{-\int t^3 dt} = e^{-t^4/4}$$

$$R(2) = e^{-(2)^4/4} = e^{-4}$$

$$\approx 0.0183$$

$$P(0.4 \leq X \leq 1.4) = R(0.4) - R(1.4)$$

$$= e^{-\frac{(0.4)^4}{4}} - e^{-\frac{(1.4)^4}{4}} \approx 0.61$$

$$P(X \geq 2 \mid X \geq 1) = P(X \geq 2) / P(X \geq 1)$$

$$= \frac{R(2)}{R(1)} = \frac{e^{-(2)^4/4}}{e^{-1/4}} = e^{-15/4}$$

$$\approx 0.0235$$

9. X has Gamma distn with mean 20 & s.d. 10

$P(X > 15)$??

$$\frac{r}{\lambda} = 20, \qquad \frac{r}{\lambda^2} = 100$$

$r = 4, \quad \lambda = 1/5$

$$P(X > 15) = \int_{15}^{\infty} \frac{1}{5^4 \Gamma 4} e^{-x/5} x^3 \, dx$$

$$\frac{x}{5} = t, \quad \frac{1}{5} dx = dt$$

$$= \frac{1}{6} \int_3^{\infty} t^3 e^{-t} \, dt$$

∞

$$= \frac{1}{6}\left[-t^3 - 3t^2 - 6t - 6\right]e^{-t}\Big|_3$$

$$= 27 + 27 + 18 + 6$$

$$= 13e^{-3} \simeq 0.6472$$

Lognormal Distribution :

Let $Y \sim N(\mu, \sigma^2)$.

Then $X = e^Y$ is said to have a lognormal (μ, σ^2) distn.

The pdf of X is

$$f_X(x) = \frac{1}{\sigma x \sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(\log x - \mu)^2}, \quad x > 0$$

$$-\infty < \mu < \infty, \sigma > 0$$

$$E(X^k) = E(e^{kY}) = M_Y(k)$$

$$= e^{k\mu + \frac{1}{2}k^2\sigma^2}$$

$$\mu_1' = E(X) = e^{(\mu + \sigma^2/2)}$$

$$\mu_2' = E(X^2) = e^{2\mu + 2\sigma^2}$$

$$V(X) = e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right)$$

$X \sim$ lognormal (mean 7.43
variance 0.56)

Find $P(X > 8)$

$$e^{\mu + \frac{\sigma^2}{2}} = 7.43$$

$$e^{2\mu + \sigma^2}\left(e^{\sigma^2} - 1\right) = 0.56$$

$$\mu = 2, \quad \sigma = 0.1$$

$$P(X > 8) = P(\log_e X > \ln_e 8)$$

$$= P\left(\frac{\ln_e X - 2}{0.1} > \frac{\ln_e 8 - 2}{0.1}\right)$$

$$= P(Z > 0.79) = \Phi(-0.79)$$

$$= 0.2148$$

<u>Beta Distribution</u> : A continuous r.v. X is said to have a beta distn with parameters (a, b) if

it has density

$$f_X(x) = \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1},$$

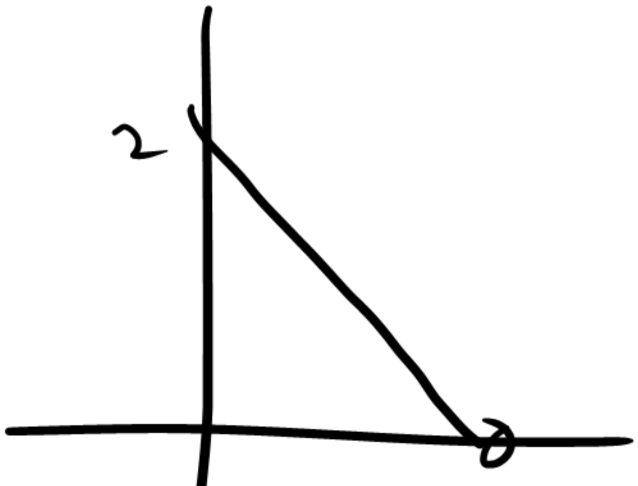
$$0 < x < 1$$

$$a > 0, \; b > 0$$

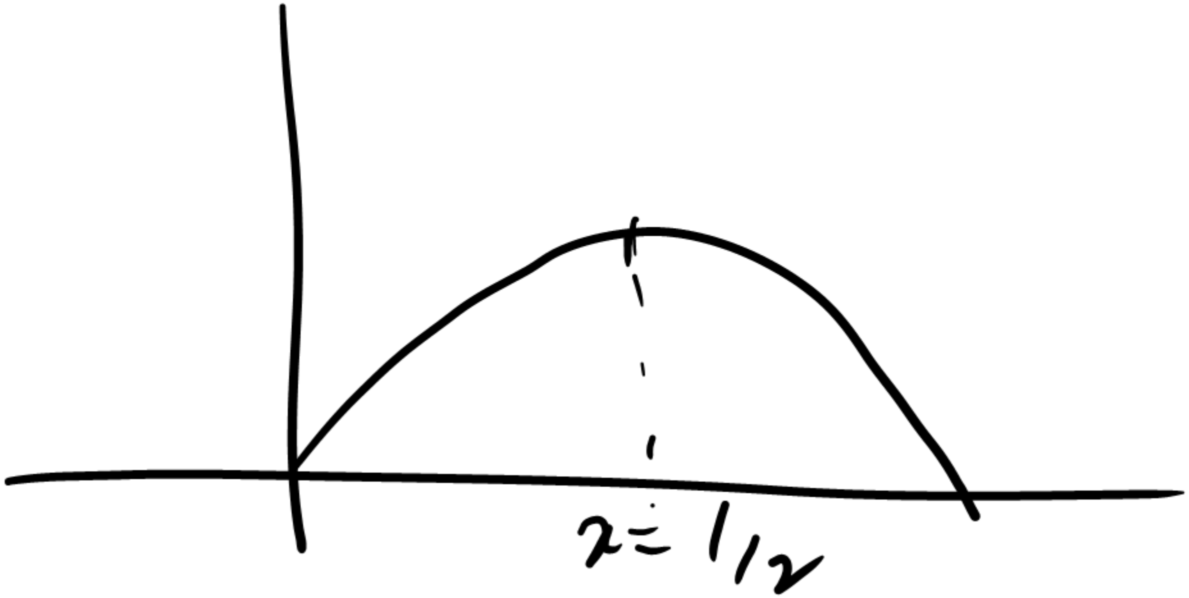
$\alpha = \beta = 1$: This is $U(0,1)$

$\alpha = 2, \; \beta = 1, \quad f(x) = 2x, \; 0 < x < 1$

$\alpha = 1, \beta = 2, \quad f(x) = 2(1-x), \quad 0 < x < 1$



$\alpha = 2, \ \beta = 2, \ f(x) = 6x(1-x), \quad 0 < x < 1$



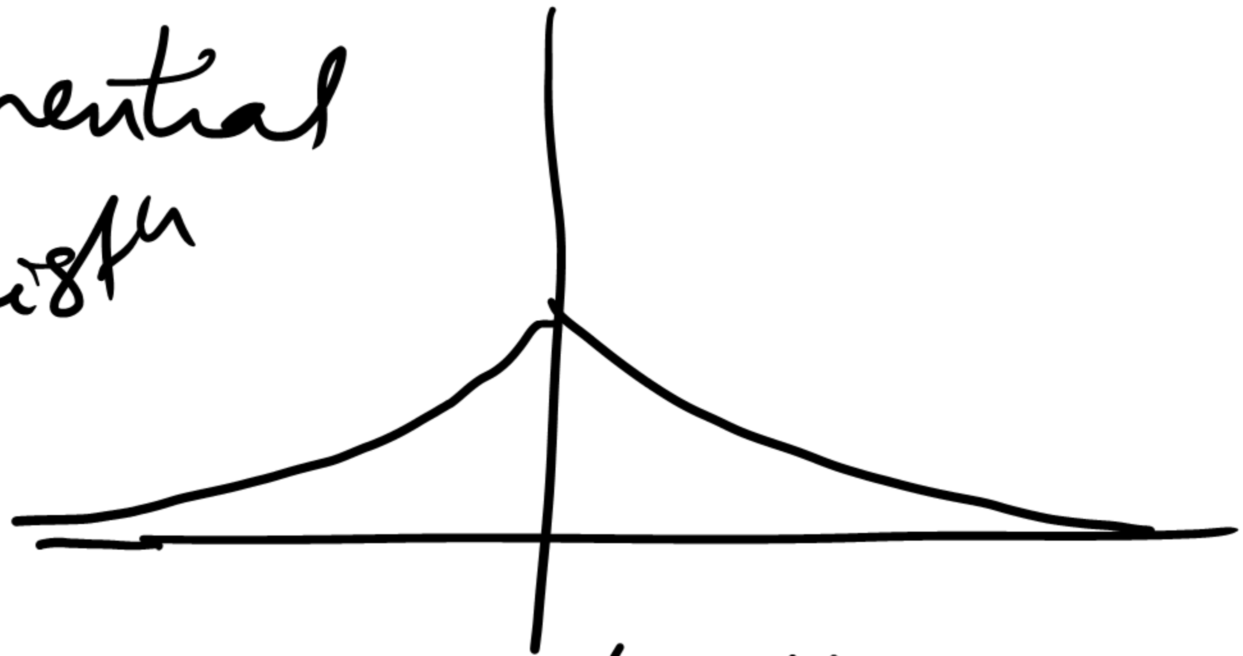
$$\mu_k' = E(X^k) = \int_0^1 \frac{x^k \cdot x^{a-1}(1-x)^{b-1}}{B(a,b)} \, dx$$

$$= \frac{B(a+k,\ b)}{B(a,b)} .$$

$$\mu_1' = \frac{a}{a+b} , \quad \mu_2' = \frac{a(a+1)}{(a+b)(a+b+1)}$$

$$\mu_2 = \frac{ab}{(a+b)^2(a+b+1)}$$

Double Exponential
or Laplace Dist^{n}



$$f_X(x) = \frac{1}{2\sigma} e^{-\frac{|x-\mu|}{\sigma}}, \quad x \in \mathbb{R}$$

$\mu \in \mathbb{R}, \ \sigma > 0$

$E(X) = \mu, \quad \text{Med}(X) = \mu, \quad \text{Mode}(X) = \mu$

$$Var(X) = 2\sigma^2.$$

<u>Cauchy Dist^n</u>:

$$f_X(x) = \frac{\sigma}{\pi\,(\sigma^2 + (x-\mu)^2)},$$

$$x \in \mathbb{R}, \quad \mu \in \mathbb{R}, \quad \sigma > 0$$

$E(X)$ does not exist

median $(X) =$ mode $(X) = \mu$

Pareto Dist^n $f(x) = \dfrac{\beta \alpha^{\beta}}{x^{\beta+1}}$,

$x > \alpha$.

Usually income / wealth dist^n.

+ve skewed