Chapter 8: Limiting Theorems and Inequalities

①

Central Limit Theorem

$X \sim N(\mu, \sigma^2) \Rightarrow \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

Recall: If $X_1, X_2, \dots, X_n$ is a random sample from normal distribution with mean μ and variance σ^2, then $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

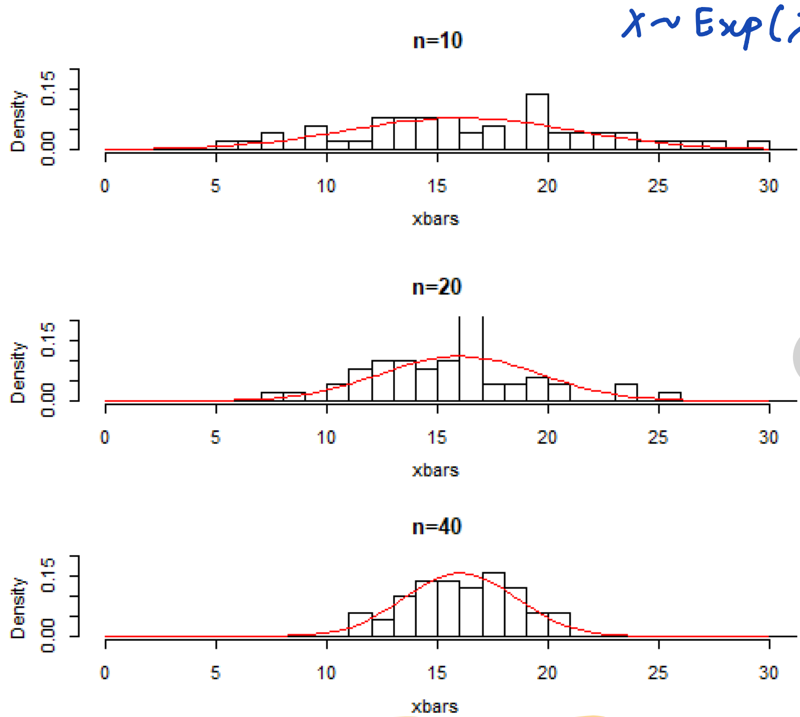
iid

What if $X_1, X_2, \dots, X_n$ is a random sample from something other than a normal distribution?

$\bar{X}$ can be a rv as well, but what if its not a std. one?

Example: We are interested in X, the time it takes PSU students to get to campus (in min). Suppose the population distribution of times is exponential with mean 16 min.

$X \sim Exp(\lambda)$ | $E(X) = 16$ (mean) $\Rightarrow \lambda = \frac{1}{16}$, $\sigma_X = 16$ (SD)



Time values are simulated for different sample sizes n.

Black histogram: the sampling distribution of $\bar{X}$

Red curve: the population distribution of $\bar{X}$

Website ⇒ NFL contracts

$X_1, \dots, X_n \longrightarrow \bar{X}_n$ (X bar sub n)

not completely normal only approx

Sample size	Sampling distribution of $\bar{X}$			Theoretical distribution of $\bar{X}$	
n	Mean	Standard deviation	Distribution	Mean μ	Standard deviation $\frac{\sigma}{\sqrt{n}}$
10	16.61	5.65	≈ normal *	16	$\frac{16}{\sqrt{10}} = 5.06$
20	15.42	3.64	≈ normal *	16	$\frac{16}{\sqrt{20}} = 3.58$
40	16.18	2.37	≈ normal	16	$\frac{16}{\sqrt{40}} = 2.53$

As n increases ⇒ mean value is coming closer to theoretical value 16

Central Limit Theorem (CLT):

Theoretical given

Let $X_1, X_2, \ldots, X_n$ be a *random sample* of size n from a population with mean μ and variance σ^2. Then as $n \to \infty$,

iid

$$\boxed{\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} Z \sim N(0,1)}$$

Calculated from graph

here we don't know if normal or not but if conditions are met we approximate them to normal

$$\boxed{\bar{X}_n = \frac{X_1 + X_2 + \cdots X_n}{n} = \bar{X} \approx N\left(\mu, \frac{\sigma^2}{n}\right)}$$

$\xrightarrow{d}$ means convergence in distribution i.e. $F_{\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}}(a) \to \phi(a)$ as $n \to \infty$

$\boxed{8.3 \Rightarrow \text{Pg } 398/399 \Rightarrow \text{Proof}}$

⊛ In practice, if $X_1, \ldots X_n$ are iid (random sample) with mean μ & Variance σ^2 and n is large enough then

$n \geq 30 \Rightarrow$ for good approx.

(2)

① $\boxed{\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \approx N(0,1)}$ — $0 \to \mu$, $1 \to \sigma^2$ — Std. Mean

② $\bar{X}_n \approx N\left(\mu, \frac{\sigma^2}{n}\right)$ → Mean

③ $T = \underbrace{\sum_{i=1}^{n} X_i}_{n\bar{X}_n} \approx N(n\mu, n\sigma^2)$ → Sum of X_i

Example 1: Suppose the distribution of an IQ test has a distribution with a mean of 100 and standard deviation of 16. A random sample of 37 students is taken from the population of all students who have taken this exam. What is the (approximate) probability that the average of the sample will be between 95 and 105?

↳ mean

$\boxed{X_1, X_2 \ldots X_{37}}$ → iid

X_i = The IQ score of the i^{th} student, $i = 1, 2 \ldots 37$

$\mu = 100, \quad \sigma = 16$

$n = 37 \geq 30 \Rightarrow$ CLT applies ✓

$\frac{\sigma^2}{n} = \sigma'^2$

$\bar{X}_{37} \approx N\left(\underset{\mu}{100}, \underset{\sigma^2/n}{\frac{16^2}{37}}\right) \Rightarrow P(95 \leq \bar{X} \leq 105) = P\left(\frac{95-100}{\sqrt{16^2/37}} \leq \bar{X} \leq \frac{105-100}{\sqrt{16^2/37}}\right)$

$\boxed{\frac{X - \text{mean}}{SD}}$ → $\frac{X-\mu}{\sigma'}$

$= P(-1.9 \leq Z \leq 1.9) = \phi(1.9) - \phi(-1.9) = 2(0.9713) - 1 =$

$\boxed{0.9426}$

Example 2: A random sample of size 55 is selected from Negative Binomial distribution with parameters $r = 5$ and $p = 0.2$. Approximate the probability that the sample average is at least 23.

→ as we don't know

$P(\bar{X} > 23)$; $X_1 \ldots X_{55}$ → iid $\quad X_i \sim \text{NegBinom}(r = 5, p = 0.2)$

$E(X_i) = \mu = \frac{r}{p} = \frac{5}{0.2} = 2.5$

$\sigma^2 = Var(X_i) = \frac{r(1-p)}{p^2} = \frac{5(1-0.2)}{0.2^2} = 100$ & $n = 55 \geq 30$ ✓ CLT

$\bar{X}_{55} \approx N\left(\underset{\mu}{25}, \underset{\sigma^2/n}{\frac{100}{55}}\right) \Rightarrow P(\bar{X} > 23) \approx P\left(Z \geq \frac{23-25}{\sqrt{100/55}}\right)$

∵ X_i are neg binom but $\bar{X}_n$ may be anything so we use CLT ∵ iid & ≥30 ✓

$= P(Z \geq -1.48) = 1 - \phi(-1.48) = \phi(1.48) =$

$\boxed{0.9306}$

True value = 0.93537 ⇒ if calculated by software

Example 3: A manager of a local bank claims that average waiting time of the entire population of their customers is 2 minutes. The waiting times are exponentially distributed. Since there were many complains through customer satisfaction surveys, headquarters send someone for inspection. He observed a random sample of 43 customers and found the average waiting time to be 2.9 minutes. Based on inspector's findings, what can we say about the manager's claim?

observed value

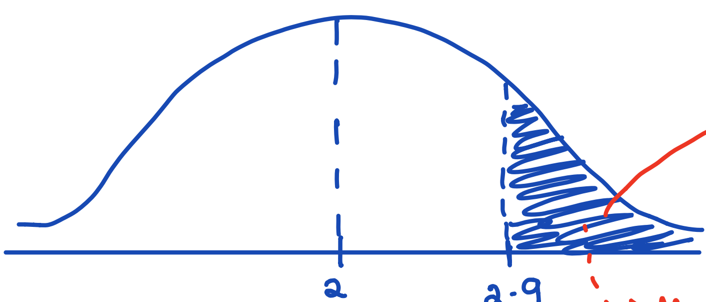
X_i = the waiting time of i^{th} customer, $i = 1, \dots 43$

$$\boxed{X_i \sim Exp(\lambda)}$$

If the manager's claim is true $\Rightarrow$ $E(X_i) = 2 = \frac{1}{\lambda}$ $\Rightarrow$ $\boxed{\lambda = \frac{1}{2}}$

$X_1, \dots X_{43}$ iid w/ mean $(\mu) = 2$, $\sigma^2 = 4$ & $n = 43 \geq 30$

$\Rightarrow$ By CLT $\Rightarrow$ $\bar{X} \approx N(\mu, \frac{\sigma^2}{n}) = N(2, \frac{4}{43})$



Mark by observed or even more extreme values

$\circledast$ $P(\bar{X} \geq 2.9) \approx P\left(Z \geq \frac{2.9 - 2}{\sqrt{4/43}}\right)$

$= P(Z \geq 2.95)$

$= 1 - \Phi(2.95) = 1 - 0.9984$

$= \boxed{0.0016}$ < 0.05

But is it small or big?

$\alpha = 0.05$

If Probability is $> \alpha$ $\Rightarrow$ large prob

If probability is $< \alpha$ $\Rightarrow$ small prob

If prob $= \alpha$

get more data

$\Rightarrow$ Manager's claim that waiting time is 2 min is not believable, because $P(X \geq 2.9)$ is very small which we calculated using values of $\mu = 2$, $\sigma^2 = 4$ as per manager's claim.

Example 4: The effective life of a component used in jet engine is a random variable with mean 5000 h and standard deviation 40 h. The engine manufacturer introduces an improvement to the manufacturing process that increases the mean life to 5050 h and decreases standard deviation to 30 h. A random sample of 32 components is selected from "old" process, and a random sample of 35 components is selected from the "new" process. Assume the samples are independent. What is the (approximate) probability that the difference between the average life for the new process exceeds that of the old one by at least 25 hours?

- X_i = life of ith component from old process, $i = 1 \ldots 32$

$X_1, \ldots X_{32}$ iid w/ $\mu_X = 5000$, $\sigma_X = 40$ & $n = 32 \geq 30$

By CLT $\Rightarrow$ $\bar{X} \approx N\left(\mu_X, \frac{\sigma_X^2}{n}\right) = N\left(5000, \frac{40^2}{32}\right)$

- Y_j = life of jth component from new process, $j = 1, \ldots 35$

$Y_1, \ldots Y_{35}$ iid with $\mu_Y = 5050$, $\sigma_Y = 30$ & $m = 35 \geq 30$

By CLT $\Rightarrow$ $\bar{Y} \approx N\left(\mu_Y, \frac{\sigma_Y^2}{m}\right) = N\left(5050, \frac{30^2}{35}\right)$

$P(\bar{Y} - \bar{X} \geq 25) = P(\bar{Y} \geq \bar{X} + 25)$

$a_1\mu_1 + a_2\mu_2$ $\qquad$ $a_1^2\sigma_1^2 + a_2^2\sigma_2^2$

$\bar{X}$ & $\bar{Y}$ are independent $\Rightarrow$ $\bar{Y} - \bar{X} \approx N\left(5050 - 5000, \frac{30^2}{35} + \frac{40^2}{32}\right)$

$= N(50, 75.7)$

$\Rightarrow P(\bar{Y} - \bar{X} \geq 25) \approx P\left(Z \geq \frac{25 - 50}{\sqrt{75.7}}\right) = P(Z \geq -2.87)$

$= \Phi(2.87) = \boxed{0.9979}$

$$\boxed{\begin{aligned} &\mathrm{Var}(\bar{Y} - \bar{X}) = \\ &1^2(\mathrm{Var}\,\bar{Y}) + (-1)^2\mathrm{Var}(\bar{X}) \end{aligned}}$$

Normal Approximation to Binomial

actual distribution given

$X \sim \text{Binom}(n, p)$. We want $P(a \le X \le b)$.

iid RVs

Recall: X can be represented as $X = Y_1 + Y_2 + \cdots + Y_n$, where

$$Y_i = \begin{cases} 1 & \text{if } i\text{th trial is success} \\ 0 & \text{if } i\text{th trial is failure} \end{cases}, \qquad \text{i.e. } Y_i \sim \text{Bernoulli}(p),\ i = 1, \dots, n$$

Check conditions:

for using this topic

① i) $n \geq 30$

ii) $np \geq 5$

iii) $n(1-p) \geq 5$

$\rightarrow$ $\boxed{X \approx N(np,\ np(1-p))}$

↳ To check the shape, so make sure to check always explicitly

Also use use continuity correction of $\boxed{0.5!}$

can't use Poisson ∵ p is big

Example 5: Let $X \sim \text{Binom}(100, 0.6)$. Approximate $P(50 \le X < 66)$.

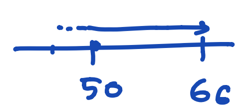
i) Check: $n = 100 \geq 30$ ✓

$np = 100(0.6) = 60$ ✓ ≥ 5

$n(1-p) = 100(0.4) = 40 \geq 5$ ✓

$\Rightarrow X \approx N(60,\ 100(0.6)(0.4)) \approx \boxed{N(\underset{\mu}{60}, \underset{\sigma^2}{24})}$

ii) Continuity Correction of 0.5



② $P(50 \le X < 66) = P(49.5 \le X < 65.5)$ (allowed val shouldn't change)

50: 49.5 or 50.5

which one doesn't change our original ∴ 49.5 & not 50.5

66: 66.5 ✗ or 65.5 ✓ (Similarly)

$$P(49.5 < X < 65.5) \approx P\left(\frac{49.5-60}{\sqrt{24}} < Z < \frac{65.5-60}{\sqrt{24}}\right)$$

$$= P(-2.14 < Z < 1.12) = \Phi(1.12) - \Phi(-2.14)$$

$$= \Phi(1.12) - (1 - \Phi(2.14)) = 0.8686 - (1 - 0.9838)$$

$$= \boxed{0.8524}$$

Without continuity correction $\Rightarrow$ answer = $\boxed{0.8681}$

Exact answer is $\boxed{0.8529}$

$\therefore$ Continuity correction helps us get a closer value to exact answer than without it

Law of Large Numbers

LLN

Recall: If $X_1, X_2, \ldots, X_n$ are iid with mean μ and variance σ^2, then $E(\bar{X}_n) = \mu$ and $Var(\bar{X}_n) = \frac{\sigma^2}{n}$.

3

As n increases, the sample mean $\bar{X}_n$ should get *closer* to μ.

as we studied in the table

still quite strong, the term is comparative that's it

Weak Law of Large Numbers:

WLLN

Let $X_1, X_2, \ldots, X_n$ be a random sample from a population with a finite mean μ. Then for any $\varepsilon > 0$,

$$\lim_{n\to\infty} P(\,|\bar{X}_n - \mu| \geq \varepsilon\,) = 0$$

Probability that the distance b/w $\bar{X}$ & μ decreases to 0, as $n \to \infty$ is 1 $\Leftrightarrow$

$$\boxed{\lim_{n\to\infty} P(\,|\bar{X}_n - \mu| < \varepsilon) = 1}$$

distance b/w $\bar{X}$ & μ

any small number

$\boxed{\bar{X}_n \xrightarrow{P} \mu}$ $\Rightarrow$ convergence in probability

Strong Law of Large Numbers:

SLLN

Let $X_1, X_2, \ldots, X_n$ be a random sample from a population with a finite mean μ. Then for any $\varepsilon > 0$,

$$P\left(\lim_{n\to\infty} (\bar{X}_n) = \mu\right) = 1$$

Limit of $\bar{X}_n$ is equal to μ $\Rightarrow$ probability of this statement $= 1$

- $\bar{X}_n \xrightarrow{w.p.1} \mu$ w.p.1 $\Rightarrow$ convergence w/ prob $= 1$ (weak)
- $\bar{X}_n \xrightarrow{a.s.} \mu$ a.s $\Rightarrow$ convergence almost surely (strong)

Visit again Apr 22

Example: You are given a coin and asked whether the coin is fair. You flipped the coin 35 times and observed

$$H, H, H, H, H, T, H, H, T, T, H, T, T, T, H, T, H, H, T, T, H, H, H, H, H, H, H, T, T, T, T, H, H, T, T$$

$\boxed{H \to 20 \quad T \to 15}$

p = True probability of heads

Let y denote # of heads among n flips of a coin

$n = 35 \to$ Bernoulli trials

$y \sim \text{Binom}(n, p)$

$\frac{y}{n}$ = relative frequency of heads $= \bar{X}_n \to$ Bernoulli

$$X_i = \begin{cases} 1 & \text{if the } i\text{th flip is H} \\ 0 & \text{if the } i\text{th flip is T} \end{cases} \qquad i = 1, 2 \ldots n$$

$X_1, X_2 \ldots X_n$ iid Bernoulli(p)

$\boxed{\mu_x = p \quad \sigma_x^2 = p(1-p)}$

$$\boxed{y = X_1 + X_2 + \cdots X_n \Rightarrow \frac{y}{n} = \frac{X_1 + \cdots X_n}{n} = \bar{X}_n}$$

Mean of $\frac{y}{n}$ or $\bar{X} = p = E\left(\frac{y}{n}\right) = E(\bar{X}_n) \Rightarrow$ Use $\frac{y}{n}$ to estimate p

Variance $\Rightarrow \text{Var}(\bar{X}_n) = \text{Var}\left(\frac{y}{n}\right) = \frac{p(1-p)}{n}$

By WLLN $\Rightarrow P\left(\left|\frac{y}{n} - p\right| < \varepsilon\right) \to 1$ as $n \to \infty$

However, this is sufficiently close to true val

$$\boxed{y = 20 \quad n = 35 \Rightarrow \frac{y}{n} = \frac{20}{35} = 0.57 = \hat{p}}$$

$\Rightarrow$ This is an estimate of p & not actual p

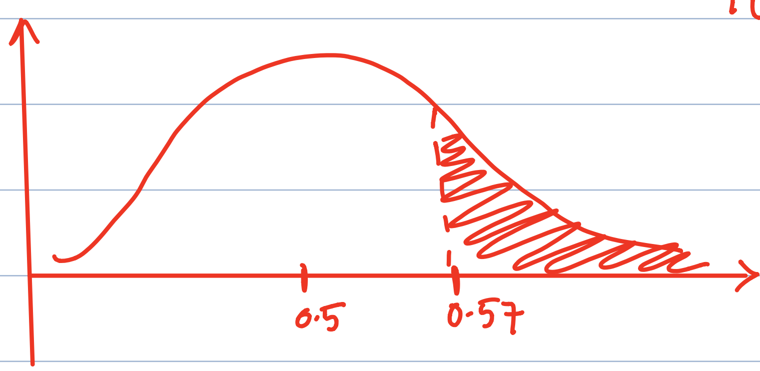
But for different samples $\Rightarrow \hat{p}$ will change $\therefore$ only an estimate

If n is large enough $\Rightarrow$ $\frac{y}{n} = \bar{X}_n$ $\Rightarrow$

$$\boxed{\frac{y}{n} \approx N\left(p, \frac{p(1-p)}{n}\right)} \quad (\text{By CLT})$$

Under fair coin $\Rightarrow$ $p = 0.5$

$$\frac{y}{35} \approx N\left(0.5, \frac{0.5(0.5)}{35}\right)$$



$$P\left(\frac{y}{35} \geq 0.57\right) \approx P\left(z \geq \frac{0.57 - 0.5}{\sqrt{\frac{0.5(0.5)}{35}}}\right)$$

$$= P(z \geq 0.84)$$

$$= 1 - \phi(0.84) = 1 - 0.7881 = \boxed{0.2119} > \alpha = 0.05$$

⊛ ∴ There is not enough evidence that the coin is not fair. $\Rightarrow$ Unable to disprove, not proving that it is fair / unfair. We are just failing to reach a conclusion.

Inequalities

Suppose the distribution of random variable X is unknown, except for $E(X) = \mu$ and $Var(X) = \sigma^2$.

when we don't know much, we use inequalities

like the distribution of RV X, then only resort to this

① **Markov's Inequality:** If X is a non-negative random variable, then for any constant $a > 0$,

$$\boxed{P(X \geq a) \leq \frac{\mu}{a}}$$

⇒ upper bound on probability

If $a > \mu$ ⇒ informative but if $\mu > a$ ⇒ doesn't make sense

② **Chebyshev's Inequality:** For any constant $c > 0$,

Cheby Shev

$$\boxed{P(\,|X-\mu| \geq c) \leq \frac{\sigma^2}{c^2}}$$

Again if $\frac{\sigma^2}{c^2} > 1$ ⇒ no sense ⇒ Else informative bound

$\boxed{\text{let } k = \frac{c}{\sigma} \Rightarrow c = k\sigma}$ ⇒ $P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$

- If $k=2$ ⇒ $P(|X-\mu| \geq 2\sigma) \leq \frac{1}{4}$ ⇒ X deviating from mean by at least 2 S.D ⇒ $\boxed{k=2}$
- If $k=3$ ⇒ $P(|X-\mu| \geq 3\sigma) \leq \frac{1}{9}$ ⇒ - - - - at least 3.SD

3

One-sided Chebyshev's Inequality: For any constant $a > 0$,

1. $P(X \geq \mu + a) \leq \dfrac{\sigma^2}{\sigma^2 + a^2}$
2. $P(X \leq \mu - a) \leq \dfrac{\sigma^2}{\sigma^2 + a^2}$

$X \geq \mu + a$
$\Rightarrow X - \mu \geq a$

always informative
$\because$ won't exceed 1

$X - \mu \leq -a \Rightarrow -(X-\mu) \geq a$

Example 6: Suppose X has mean 48 and variance 28.8.

a) Use at least two appropriate inequalities to provide a bound on $P(X \geq 50)$.
b) What can be said about $P(X \leq 40)$?
c) A random sample of 120 people is selected. A person favors the new local tax proposition with probability 0.4, independently from the others. We are interested in the probability that at least 50 of the selected are in favor of the new tax proposition.

⊛ We don't know any thing about the distribution

a) By Markov's Identity $\Rightarrow P(X \geq 50) \leq \dfrac{E(X)}{a} = \dfrac{48}{50} = \boxed{0.96}$

$a = 50$

By one-sided Cheb. Ineq $\Rightarrow P(X \geq 50) \leq \dfrac{\sigma^2}{\sigma^2 + a^2} = \dfrac{28.8}{28.8 + 2^2}$

$\mu + a$ $(a = 2)$ $= \boxed{0.878}$

b) By One sided Cheb. Ineq $\Rightarrow P(X \leq 40) \leq \dfrac{\sigma^2}{\sigma^2 + a^2} = \dfrac{28.8}{28.8 + 8^2} =$

$\mu - a$ $a = 8$ $\boxed{0.31}$

If X was non negative $\Rightarrow$ Markov's inequality

$X \geq 0 \Rightarrow P(X \leq 40) = \begin{cases} \text{Markov's Ineq} \\ P(X > 40) \leq \dfrac{E(X)}{40} = \dfrac{48}{40} \end{cases}$

$P(X \leq 40) = 1 - P(X > 40) > 1 - \dfrac{48}{40} = \dfrac{-8}{40}$

Both are not informative

$\boxed{-P(X \leq 40) \leq 8/40}$

c) X = # people among 120 who favor the proposition

$\Rightarrow X \sim \text{Binom}(n=120,\ p=0.4)$

$E(X) = np = 48$; $\text{Var}(X) = np(1-p) = 28.8$

$P(X \geq 50) = ?$

⊛ Let's also practice normal distributions

Check : $n = 120 \geq 30$ ✓

$np = 48 \geq 5$ ✓ $\qquad X \approx N(48, 28.8)$

$np(1-p) = 72 \geq 5$ ✓

$$P(X \geq 50) = P(X > 49.5) \approx P\left(Z > \frac{49.5 - 48}{\sqrt{28.8}}\right)$$

$$= P(Z > 0.28) = 1 - \Phi(0.28) = 1 - 0.6103$$

$$= \boxed{0.3897}$$

4 **Recall:** A twice-differentiable real-valued function $g(x)$ is convex if $g''(x) \geq 0$ for all x. ∪
A twice-differentiable real-valued function $g(x)$ is concave if $g''(x) \leq 0$ for all x ∩

Jensen's Inequality: on S_X ⇒ that's it, no matter what happens elsewhere

- If $g(x)$ is a convex function, then $E[g(X)] \geq g(E[X])$.
- If $g(x)$ is a concave function, then $E[g(X)] \leq g(E[X])$.

Example $g(x) = x^2 \Rightarrow g'(x) = 2x \Rightarrow g''(x) = 2 > 0 \; \forall x$

$\therefore$ Strictly convex

$$E(g(X)) \geq g(E(X))$$

$$E(X^2) \geq (E(X))^2 \Rightarrow \boxed{E(X^2) - (E(X))^2 \geq 0}$$

$$\Rightarrow \boxed{Var(X) \geq 0}$$

Example let $g(x) = (x+2)^{1/7}, \; x > -2$

$$\Rightarrow g'(x) = \frac{1}{7}(x+2)^{-\frac{6}{7}}$$

$$\Rightarrow g''(x) = \frac{-6}{49}(x+2)^{-\frac{13}{7}} < 0 \text{ for all } x > -2$$

⇒ concave

$$\circledast \; E(g(X)) = E\left((X+2)^{1/7}\right) \leq g(E(X)) = (E(X)+2)^{1/7}$$

should always be > 0 $(x > 0$ & not $x > -2)$

Now suppose $X \sim$ Binom $(n = 120, p = 0.4)$ for calculations

$$\circledast \Rightarrow \boxed{E\left((X+2)^{1/7}\right) = 1.747 \leq (E(X)+2)^{1/7} = 1.919}$$

visit again

$\ln(1.1$ ✓ x

E

$.1^{5}$

$.9$

LB