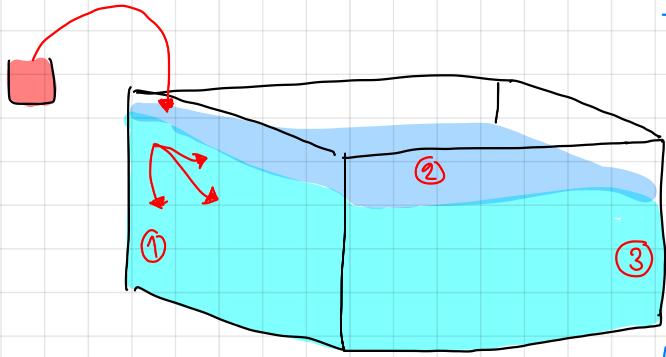
What can we learn about the population of world by taking a small sample and what are the limitation on taking the sample

↳ Sampling distribution

↳ how can we relate, sampling result, from the small sample we have back to the population from which it came

↳ เหมือนเรามีสารเคมีที่จะเอามาผสมกัน

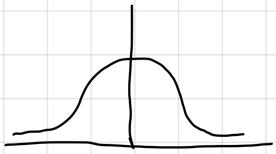


→ ลองคิดดูถ้าเทลงจากฝั่งซ้าย
↳ ถ้าเราตักขึ้นมาวัดความเข้มข้นของสารแต่ละจุด ①,②,③
↳ ตอนแรกเริ่ม
↳ ① เข้มข้น
② จางๆ
③ จางสุดๆ
↳ เมื่อมีการคนปล่อยไว้ให้เข้ากัน
↳ ①, ②, ③ อาจจะเหมือนหรือต่างกันได้

↳ การหยิบตัวอย่างในบ่อสารนั้นมาตรวจสอบ เราได้ส่วนที่ผสมเข้ากันดีแล้ว
↳ คำตอบของเราจะขึ้นอยู่กับ sample ที่เราเอามา
↳ แล้วเราจะรู้ได้ไงว่าอันไหน ok เพราะว่าต่าง sample ต่าง answer

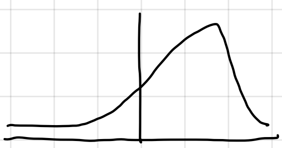
↳ เมื่อเราเอามาหลายๆ sample ที่มีค่าต่างกันมันคือ sampling distribution

↳ การกระจายข้อมูลในโลกนี้
↳ normal distribution



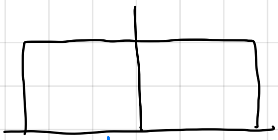
เกือบทุกอย่างบนโลกเป็นแบบนี้

↳ skewed distribution (เบ้)



เหมือนไปวัด IQ ของเด็ก MIT

↳ uniform distribution



ทุกๆอย่างมีโอกาสเกิดขึ้นเท่าๆกัน

population อาจจะหน้าตาเป็น normal แต่ sample เราอาจจะเป็น skewed ก็ได้

↳ ตามที่กล่าวไป → เราไม่รู้หน้าตาของประชากรหรอก ถ้าเรารู้เราไม่ต้องเรียนสถิติกันแล้ว

↳ population แต่ละชุดใหญ่เกินจนเราไม่สามารถเก็บข้อมูลได้หมด เราเลยทำการเก็บสำรวจแบบทุกๆ x คน

↳ in this lecture sample size "n"
↳ แต่ละ sample เราสามารถหาค่า mean, variance, std deviation ...
↳ เมื่อเราทำการ sample เยอะๆ ก็จะมีค่า ⤴ ของแต่ละ sample → ที่มีขนาด n
↳ เราก็จะได้ "sampling distribution"

↳ ถ้านำ mean มาต่อกันจะเป็น "sampling distribution of sample means"

มีของแต่ละ sample ที่สุ่มมา size n

↳ สรุป sampling distribution is when you get information and distribution of answer

↳ usually when you do that you calculate something $\{\bar{x}, \sigma, \sigma^2\}$, called "sampling distribution of sample mean"

Central Limit Theorem

↳ given: a population has mean (μ) and standard deviation (σ)

↳ sample: choose all sample of size "n" from population

↳ เลือกมากลุ่มละ n ทั้งหมดเท่าที่เป็นไปได้

↳ ตัว central limit theorem จะเอาไปใช้งานได้กับทุก distribution

n=2 เหมือนเดิม

sample#1, n=2
↳ val1, val2 $\} \rightarrow \bar{x}_1$ (sample mean)

sample#2, n=2
↳ val1, val2 $\} \rightarrow \bar{x}_2$

sample#3, n=2
↳ val1, val2 $\} \rightarrow \bar{x}_3$

* ไม่ได้เลือกออกมาเป็นกลุ่มๆ เป็นๆ แต่เป็นการ random ออกมา

เลือกไปจนกว่าจะหมด หรือเหนื่อย

first part

↳ theorem: the mean of the sample means is equal to the mean of the population

$$\mu_{\bar{x}} = \mu$$

↳ ไม่สนว่า sample size จะมีขนาดเท่าไหร่

↳ ถ้าตามตัวอย่าง $\mu_{\bar{x}} = \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3}{3}$ เพื่อมีค่าเท่ากับ μ

เหมือนตาม sense ทั่วไปคือเอาค่าเฉลี่ยของกลุ่มเล็กๆ มาหาค่าเฉลี่ยรวมอีกที

↳ ก็เหมือนกับไปสำรวจจากทุกคนอะ "every posible sample"

↳ theorem: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

standard deviation of sample mean

↳ เมื่อเอาค่าของ σ จาก $\bar{x}$ มาแล้ว และรู้ n เราสามารถหา standard deviation of population ได้

↳ ต่อให้เราไม่ได้ดูตัวจริงของ sample ทั้ง population เราก็สามารถเข้าถึง σ ของ population ได้

second part

↳ theorem: if population is normal, then the sample mean will have a normal distribution. (independent of sample size) → เมื่อมันเป็น normal ใช้ z ได้

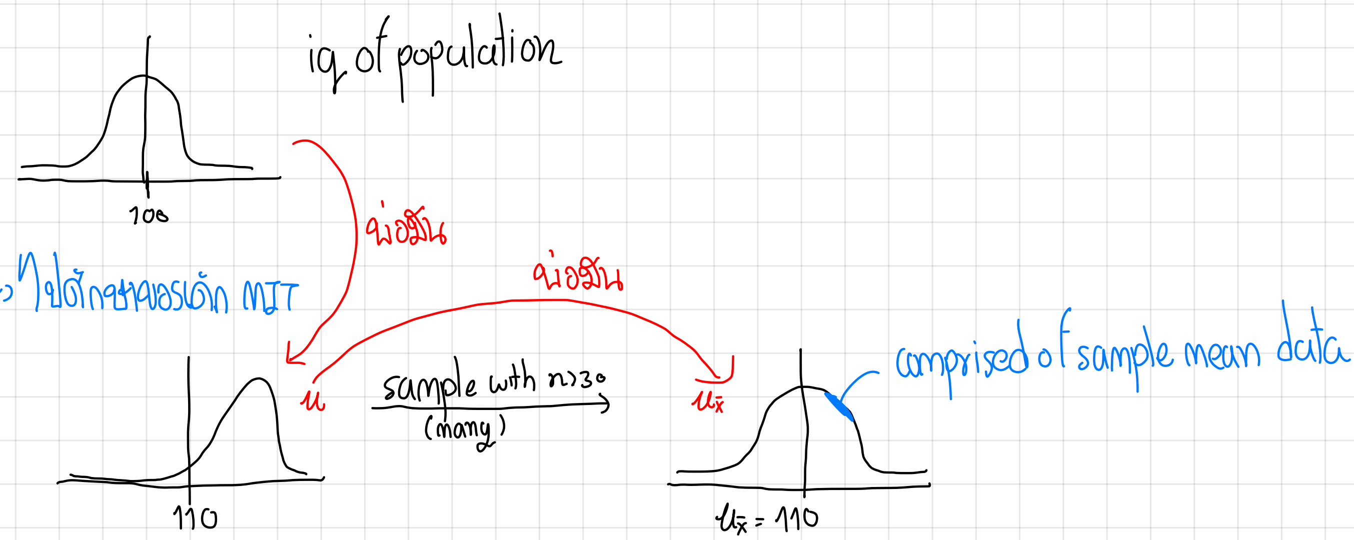
↳ theorem: if population is not normal, but the sample size "n" > 30, then sampling distribution of sample mean approximates a normal distribution for any population distribution shape.

↳ สมมติ population มีการกระจายแบบ skewed → ทำการ sample ($n=30$) ทุกๆตัวในประชากร คำนวณหา mean ของแต่ละอันจะได้ sampling distribution of sample mean เหมือน normal

↳ I don't need to know what my population look like, as long as I sample it with a greater than 30 then the sampling distribution of sample mean that I get is going to look normal → normal เลยใช้ z

last part

↳ iq of population



↳ **ex1.** Given $\mu = 35$, $\sigma = 9$ ↳ sample $n = 64$ } what is $\mu_{\bar{x}}$ and $\sigma_{\bar{x}}$?

↳ ตาม theorem: $\mu_{\bar{x}} = \mu$

$\mu_{\bar{x}} = 35$

และ $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$ → ดูการกระจายตัวแค่ไหนของ sample mean

→ ถ้าขนาด sample น้อยลง อันนี้จะแผ่ออกเพราะ $\sigma_{\bar{x}}$ เยอะขึ้น

$= \dfrac{9}{\sqrt{64}} = \dfrac{9}{8}$

↳ ตามหลักแล้วข้อนี้จะ → มันแปลกๆในเมื่อเรามีข้อมูลประชากรอยู่แล้วจะหาตัวอย่างทำไม แต่เขาสมมติให้ทำเฉยๆ

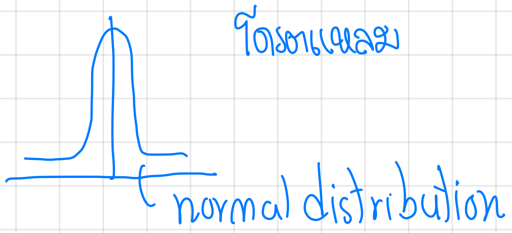
↳ **ex2** given $\mu = 12$, $\sigma = 2.3$, $n = 36$

อยากได้ $\mu_{\bar{x}}$, $\sigma_{\bar{x}}$

↳ $\mu_{\bar{x}} = \mu$

$= 12$

$\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{2.3}{\sqrt{36}}$ | $= \dfrac{2.3}{6} \approx 0.383$



Applying Central Limit Theorem to Population Means

- เงื่อนไขตัวเลขบนกล่องว่า ถ้าเกิดมี sample size > 30, sampling distribution of sample mean will be normal

  อ้วนข้างผอมข้าง

- z-score
  - $z = \frac{x-\mu}{\sigma}$ ⟹ ค่าการเปลี่ยนรูปแบบของตัวแปร

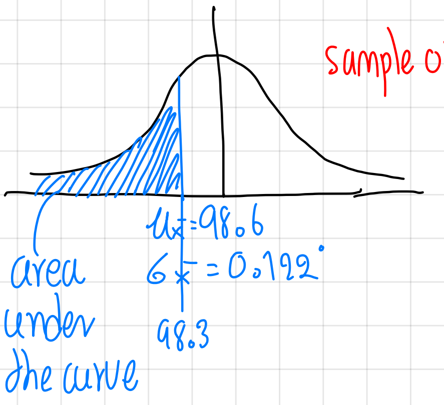
    $z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$ → เกิดแต่ละครั้งของการ sample

    $z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$

- ex body temperature of adults are normally distributed with a mean of 98.6 deg F and standard deviation of 0.73 deg F. Find the probability that a sample of 36 adults have average body temp less than 98.3 deg F?
  - $\mu = 98.6$, $\sigma = 0.73$

    $n = 36$,
    - ตอนนี้เกิด sample mean ก้อนหนึ่ง → จริงๆ มี sample distribution of sample mean
    - เหมือนเป็นการถามว่า จะมีโอกาสที่ sample means would be less than 98.3
    - $n > 30$ → sampling distribution เหมือนเป็น normal distribution

      $\mu_{\bar{x}} = \mu = 98.6$ $\qquad$ $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.73}{\sqrt{36}} = 0.1217$



sample of 36 people

$\frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$

z-score $= \frac{98.3 - 98.6}{0.122} = -2.47 \xrightarrow{\text{เปิดตาราง}} 0.0068$ ← นี่คือ prob ที่เราต้องการ

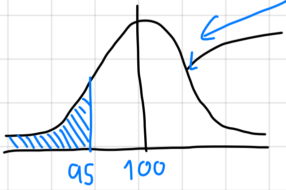
$P(z < -2.47)$

- ex IQ score for the USA have a mean of 100 and standard diviation of 15. If a sample of 50 people are given an IQ tests, what is the probability that their mean IQ will be less than 95?
  - $\mu = 100$, $\sigma = 15$

$n = 50 \rightarrow n \geq 30 \rightarrow$ sampling distribution เป็นแบบ normal

$\mu_{\bar{x}} = \mu$
$= 100$

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$
$= \frac{15}{\sqrt{50}}$
$= 2.121$



การsample หลายๆครั้งของ $n=50$

$z\text{-score} = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$
$= \frac{95 - 100}{2.121}$
$= -2.36 \longrightarrow$ เปิดตาราง $P(Z < -2.36) = 0.0091$

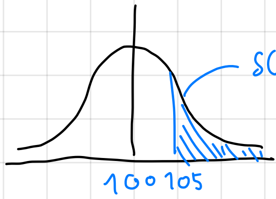
↳ เป็นเหตุผลว่าสิ่งนี้เป็นประโยชน์ของโลก central limit theorem

↳ ตามหลักโจทย์ที่ข้อมูลจะไม่มีรูปแบบ distribution ใดก็ใช้ได้เลยนะ แค่บอก μ, σ

↳ ex the population IQ mean is 100 with standard deviation of 15. If we sample 50 people, what is probability that the average IQ of the sample is greater than 105?

↳ $\mu = 100$, $\sigma = 1.5$
$n = 50$,
$\mu_{\bar{x}} = \mu$
$= 100$

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$
$= \frac{1.5}{\sqrt{50}}$
$= 2.121$



sample mean

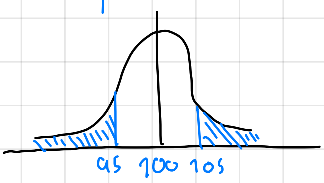
$z\text{-score} = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$
$= \frac{105 - 100}{2.121}$
$= 2.36$

$P(Z > 2.36) = P(Z < -2.36)$

เปิดตาราง

ปกติตารางจากซ้าย
ไปขวา

↳ โอกาสที่ sample มาแล้วค่าเฉลี่ยจะออกจากค่าเฉลี่ยมากกว่า 5 หน่วยนะ



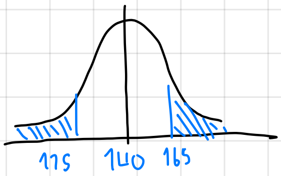
↳ ex average fetal heartrate is 140 bpm, with standard deviation of 12 bpm. What is the probability that randomly chosen fetal heartrate differs from the mean by more than 25 bpm?

↳ $\mu = 140$, $\sigma = 12$

↳ เมื่อคุณเจอปัญหา central limit theorem ต้องรู้ก่อนเลย sample size

↳ ทุกคนคิดว่าเลือกยากมากว่าเป็นเรื่อง central limit theorem แต่คุณไม่สามารถทำเป็นได้หรอก

↳ ทำแบบปกติเลย → ไม่เกี่ยวกับเรื่องนี้



$z_1 = \frac{115-140}{12} = -2.08$ → เปิดตาราง ซึ่งนำที่ได้เป็นบวกคูณ 2

$z_2 = \frac{165-140}{12} = 2.08$

↳ ไม่ต้องทำหลายรอบ

↳ ex the average person play video game 7.5 hours per week with standard deviation of 3 hours if I choose 110 people randomly, what the probability that their mean game play is more than 8 hour per week?

↳ $\mu = 7.5,\ \sigma = 3$

$n = 110$

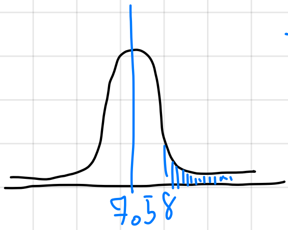
$\mu_{\bar{x}} = \mu$

$= 7.5$

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

$= \frac{3}{\sqrt{110}}$

$= 0.286$



$z = \frac{8-7.5}{0.28}$

$= 1.79$

$P(z > 1.79) = P(z < -1.79)$

$= 0.0401$

Applying Central Limit Theorem to Population Proportions

↳ propotion in statistic is a bit difference than in basic mathematics
  ↳ is very specific thing
  ↳ population propotion is % of a population that has a characteristic
    ↳ 97% of all people carry mobile phone
      97% → %
      carry mobile phone → characteristic
  ↳ represent as "p"

↳ sample propotion
  ↳ is % of sample that has a characteristic
  ↳ represent as "$\hat{p}$" called "p-hat"

↳ ตัวอย่างรอบเลือกตั้ง
  ↳ จะมี poll station หลายๆที่เพื่อเก็บข้อมูล → แต่ละอันก็สามารถคำนวณหา propotion ของมันได้
    ↳ และจะได้สิ่งที่เรียกว่า sampling distribution of sample propotions
      ↳ อันเก่า sample mean → เราคำนวณเกี่ยวกับ mean
    ↳ อันเก่าแค่ sample size > 30 ก็สามารถใช้สูตรได้แล้ว แต่กับ propotion มันจะมีเงื่อนไขแปลกๆ → นอกจากกรอบ เหมือนเดิม แต่มีเงื่อนไขกับสูตรเพิ่ม
      ↳ ensure : $np \geq 5$ and $n(1-p) \geq 5$ → อย่าเพิ่งสงสัยมันเลย 555+
        ↳ then sampling distribution of sample propotion is normal distribution
        ↳ then $z\text{-score} = \dfrac{\hat{p} - p}{\sqrt{\dfrac{p(1-p)}{n}}}$

↳ ex 79% of voters in a city are Democrats. If we sample 100 people, what is the probability that more than 68 will vote Democrats?

$p = 0.79$
$n = 100$

① $np \geq 5$
$100(0.79) \geq 5$
$79 \geq 5$ ✓

② $n(1-p) \geq 5$
$100(1-0.79) \geq 5$
$100(0.21) \geq 5$
$21 \geq 5$ ✓

} sampling distribution of sample propotion look normal

$\hat{p} = \dfrac{68}{100} = 0.68$

$z = \dfrac{\hat{p}-p}{\sqrt{\dfrac{p(1-p)}{n}}} = \dfrac{0.68 - 0.79}{\sqrt{\dfrac{0.79(1-0.79)}{100}}} = -2.68$

$\therefore p(z > -2.68) = p(z < 2.68)$
$= 0.9963$
↳ make sense

↳ ex 47% of people in a city want to buy a bike. If we survey 61 people randomly, what is the probability that less than 40% want to buy a bike?

↳ สงสัยไหมเขาเอาข้อมูลประชากรมาจากไหน → อาจจะมาจากข้อมูลในอดีตก็ได้

↳ $n = 61$, $p = 0.47$, $\hat{p} = 0.4$ (จุดที่เราสนใจ)

$np \geq 5$ $\qquad$ $n(1-p) \geq 5$

$61(0.47) \geq 5$ $\qquad$ $61(1-0.47) \geq 5$

$28.7 \geq 5$ ✓ $\qquad$ $32.3 \geq 5$ ✓

ถ้าทั้ง 2 เป็นจริงแสดงว่าเรา sample จำนวนมากเพียงพอที่จะ =

ทำ central limit theorem applying

↳ it will look normal

$$Z = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.4-0.47}{\sqrt{\frac{0.47(1-0.47)}{61}}} = \frac{-0.07}{0.0639} = -1.10$$

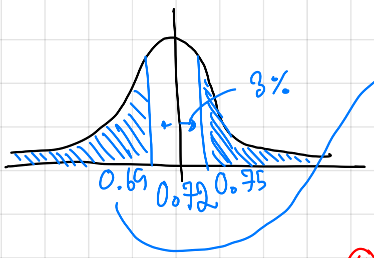
$P(Z < -1.10) = 0.1357$; เปิดตารางหา

↳ ex 72% of shoppers at a store are women. 40 shoppers are chosen randomly. What is probability that the propotion of women sampled differs form the mean by more than 3%?

↳ $n = 40$, $p = 0.72$ $\qquad$ $np = 40(0.72) = 28.8 \geq 5$ ✓

1 $\qquad$ 2 $\quad = 40(1-0.72) = 11.2 \geq 5$ ✓

step เหมือนเดิม



3 $$Z = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.69-0.72}{\sqrt{\frac{0.72(1-0.72)}{40}}} = \frac{-0.03}{0.071} = -0.423$$

$P(Z < -0.423) = 0.3372$ → สองข้างเท่ากันแน่นอน

4

$P(\text{ช่วงกลาง}) = 0.6744$

Confidence interval for Population Mean

↳ the purpose of statistics in most case is really that you would like to learn something about the large population but it's very impossible to do that

↳ so you would take some sample and you would like to see if that sample is somehow representative of the population.

↳ สุ่มก็ต้องขึ้นอยู่กับขนาดของ sample size อีก ว่าเท่าไหร่ถึงดีน้อ

↳ แล้วจะรู้ได้ไงว่ามันใกล้เคียงกับประชากรแล้ว

↳ Point estimate : single num estimate of a population parameter

↳ the best point estimate for population mean μ is the sample mean $\bar{x}$

↳ at the starting, the population mean is probably pretty close to that (เป็นจุดเริ่มต้นที่ดี) clearly we know it's not, but it's great starting point.

↳ then we're gonna figure out what is the error associated with that? what is level of confidence about we've done?

↳ ถ้าใช้ตาม central limit theorem → เราต้องทำการ sample all possible sample groups → sample whole pop.

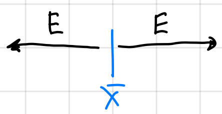
↳ ในกรณีนี้เราไม่ทำแบบนั้น เราจะทำการ sample แค่ครั้งเดียว

↳ จุดเริ่มต้นที่ดีที่สุดคือ sample mean

↳ มี error เรียกว่า "margin of error" (E)

↳ margin of error (E)

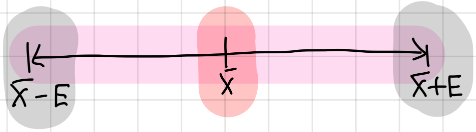
↳ คือ largest distance from the point estimate that will contain population mean



↳ คือ construct interval ดีๆ นี่เอง

↳ interval estimate

↳ range of value that contain the population parameter.



95% ว่า μ อยู่ในนี้

↳ level of confidence

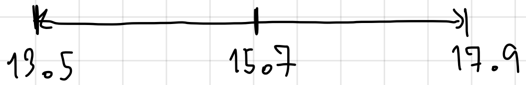
↳ how certain we are that the interval estimate contain the mean

↳ Confidence Interval

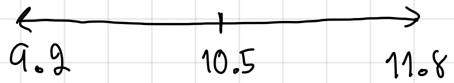
↳ interval estimate that goes with a specific level of confidence

↳ ex a survey of 200 male show that they read on avg 15.7 hour per week. If margin of error is 2.2 hrs at 95% confidence, construct their confidence interval.

$13.5 < \mu < 17.9$ ← 95% CI



↳ ex a survey of 600 people finds they sleep an avg of 10.5 hours per night. If margin of error at 98% confidence level is 1.3 hours, construct CI.



98% CL

$9.2 < \mu < 17.9$

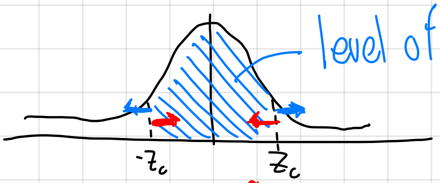
↳ ต่อให้ไม่ได้กำหนด margin of error เองแล้ว also depend on CL and sample size.

Estimate Population Means (Large # of Sample) (n>30)

↳ assume → random are random.
→ n>30 "large"
→ population standard deviation (σ) is unknow → ในหนังสือเล่มอะ = มึงรู้ได้ไงเล่าเว้ย

then → use t-distribution to find E (margin of error)
↳ but when n>30 "large" we can use normal distribution to approximate t-dis

↳ Procedure → how much confidence level
→ find the critical value of z (z_c)



level of confidence = Area ex = 0.95

CL น้อยลง
CL มากขึ้น

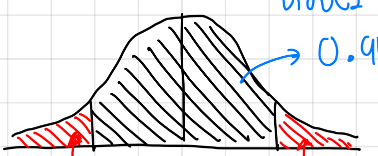
→ the margin error is $\boxed{E = z_c \frac{s}{\sqrt{n}}}$

standard deviation
sample size
critical value

↳ สามารถเอาไปหา choice of sample size ได้

↳ หา critical value กัน

ถ้าขอ CI 95%



0.95

$\frac{0.05}{2}$ $\frac{0.05}{2}$

$= 0.025$ $= 0.025$

ไปหา z ที่มีพื้นที่ = 0.025 จะได้ z = -1.96

↳ ไม่จำเป็นต้องรู้ค่าทุกรอบ ส่วนใหญ่ใช้ 95%

↳ จะเขียนแบบไหนก็ได้ ข้อสอบส่วนใหญ่จะเป็น $Z_{\frac{\alpha}{2}}$

$\boxed{Z_c = Z_{\frac{\alpha}{2}}}$ → เปิดตาราง ตอนนี้มีเตรียมไว้ในมือตอนแล้ว #

โดย $\alpha = 1-c$

ถ้า C	Z_c
0.80	1.28
0.85	1.44
0.90	1.645
0.95	1.96
0.98	2.33
0.99	2.575

↳ **ex** We asked 100 people how much a supreme pizza costs. The average answer is \$25.99 with standard deviation of \$15.50 Construct 99% CI that contain the average price of supreme pizza?

① $n > 30$
② σ unknow

↳ $\bar{x} = 25.99$, $c = 0.99$, $s = 15.50$, $n = 100$

หา critical value $z_c = z_{\frac{1-0.99}{2}}$

$= 2.575$

หา margin of error $= z_c \frac{s}{\sqrt{n}}$

$= \frac{2.575(15.50)}{\sqrt{100}}$

$= 3.99$

$\bar{x} - E$	25.99	$\bar{x} + E$
22		29.98

99% that the mean of population will fall in

↳ $E = z_c \frac{s}{\sqrt{n}}$ ถ้าเกิดเพิ่ม sample size เยอะๆ → ค่าในช่วงจะยิ่งแคบลง $E \propto \frac{1}{n}$

↳ **ex** 78 student surveyed said they study an average of 15 hours per week with standard deviation of 2.3 hrs. What is the margin of error for 90% CI for the student population?

↳ $n > 78$ "large sample size", not know σ

$n = 78$, $\bar{x} = 15$, $s = 2.3$

$E = z_c \frac{s}{\sqrt{n}}$

$= 1.645\left(\frac{2.3}{\sqrt{78}}\right)$

$= 0.428$

หา $z_c = z_{\frac{1-\alpha}{2}}$

$= 1.645$

จะหา CI ก็คือ $\bar{x} \pm E$

↳ **ex** $n = 89$, $s = 2.01$, $c = 0.95$, $\bar{x} = 45$, 95% CI

$E = z_c \frac{s}{\sqrt{n}}$

$= 1.96\left(\frac{2.01}{\sqrt{89}}\right)$

$= 0.42$

44.58	45	45.42

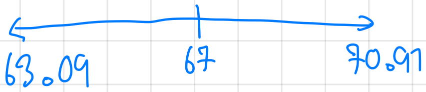
↳ **ex** 85 homeowners answered that they spent 67\$ per month on repair with standard deviation of 14\$ find the 99% CI how much money spent on repair/month by all homeowners.

↳ $n = 85$, $\bar{x} = 67$, $s = 14$, $c = 99\%$

$$E = z_c\left(\frac{s}{\sqrt{n}}\right)$$

$$= 2.575\left(\frac{14}{\sqrt{85}}\right)$$

$$= 3.91$$

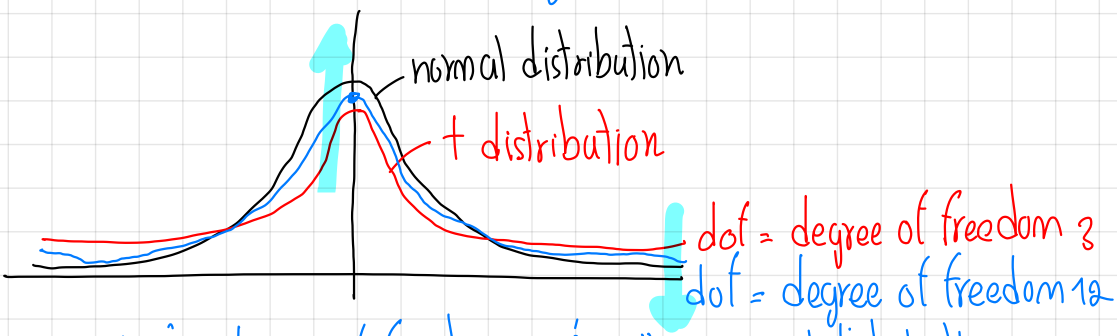


↳ ก่อนจะจบ ตั้งคำถาม ถ้า sample size < 30 หล่ะ ?

The Student t-Distribution

- เมื่อ n≥30 สามารถใช้ normal distribution แทน t distribution แทนได้ → t ใช้เมื่อไหร่ก็ได้
  - เนื่องจากความเกือบเหมือนกันในรูปร่าง
- เมื่อ n<30 เราก็จะสามารถใช้ t ได้
  - ผลกระทบของรูปกราฟ จากขนาดค่าเฉลี่ย, ขึ้นเสียง
- properties : → symmetrical and bell shaped ✓
  → centered about 0
  → area under curve = 1 ✓
  → x-axis horizontal asymtote ✓

  เหมือน normal distribution



  - เมื่อยิ่งเพิ่ม degree of freedom มันยิ่งเหมือน normal distribution
- dof → ∞ → shape is determine by dof
  - จะทำให้ t-distribution become normal distribution → shape is determine by $\mu, \bar{x}$
- so what the fuck is degree of freedom?
  - degree of freedom (dof) = n(sample size) − 1

$$dof = n-1$$

- ถ้าเกิด n<30 มันจะไม่ค่อยเหมือน normal distribution → ถ้าใช้การกระจายแบบ normal ไม่ค่อยสอดคล้องความเป็นจริง

Using a Student t-Distribution Statistical Table

t_α

↳ area of t that give the area of α to the right → normal เป็น area to the left

↳ การอ่านตาราง

dof	← α listed as column →
↓	read t_α from the body of table

→ area to the right

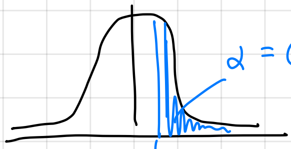
อันในเลขของตารางค่า score อันนอกเลขของตารางค่า area

t เขียน area , ดู จากรูปของโจทย์

↳ **ex** what is $t_{0.025}$ with 25 df

↳ ตาราง

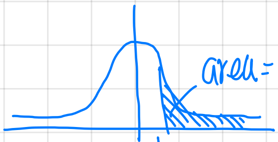
	0.05	0.025	0.01	0.005	0.0005
24					
25		2.05954			
26					



$t_{0.025}$

$\therefore t_{0.025} = 2.060$

↳ **ex**



$t_?$

df = 17

ถามว่าคือ t?

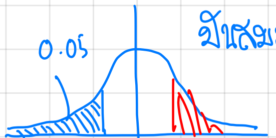
t ตามด้วย area

↓

ดูแค่ one tail หรือ 2 tail

↳ ตอบ $t_{area} \rightarrow t_{0.1} \rightarrow$ เปิดตาราง = 1.333379

↳ **ex**



$t_?$

ฝั่งสมมาตร

11 df

↳ พื้นที่แดง = ฟ้า

$t_? = -t_{0.05}$

$= -1.796$

Estimate Population means (Small Samples)

↳ กรณีที่ $n < 30$

↳ หลักการเหมือนเดิมแค่หา margin of error

↳ assume : all sample are random

→ $n < 30$

→ pop distribution is normal → ถ้าไม่ normal แต่ไม่ต่างจากนอร์มอลมากเกิน.

→ population standard deviation (σ) is unknow

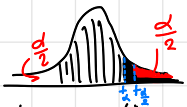
then → use student-t-distribution to find confidence interval

↳ หาจาก margin of error

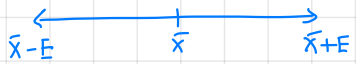
$$E = t_{\frac{\alpha}{2}}\left(\frac{s}{\sqrt{n}}\right)$$

คล้ายกับของเดิมเลย ; ก็เป็นค่า t_{area}

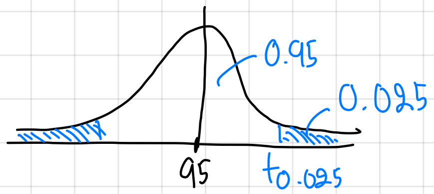
↳ พื้นที่ด้านซ้าย $\frac{\alpha}{2}$ ขวา $\frac{\alpha}{2}$



↳ ถ้าเกิดเราใช้ E ด้วย t_α → มันต้องพื้นที่ 2 ฝั่ง → ทำให้ E เราเพี้ยนได้



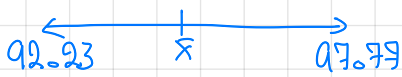
↳ ex $n=14$, $\bar{x}=95$, $s=4.8$, CI 95% หา confidence interval



$$E = t_{0.025}\left(\frac{s}{\sqrt{n}}\right)$$

$$= 2.1604\left(\frac{4.8}{\sqrt{14}}\right)$$

$$= 2.77$$



↳ ex 25 tigers were found to have average weight of 600 pounds with standard deviation of 90 pounds. What is the magin of error on a 98% confidence interval

↳ $\bar{x}=600$, $n=25$, $dof=24$, $s=90$, $c=0.98$

$$E = t_{\frac{\alpha}{2}}\left(\frac{s}{\sqrt{n}}\right)$$

$$= 2.492\left(\frac{90}{\sqrt{25}}\right)$$

$$= 44.86$$

$$t_{\frac{\alpha}{2}} = t_{\frac{1-c}{2}}$$

$$= t_{\frac{0.02}{2}}$$

$$= t_{0.01}$$

$$= 2.492$$

↳ ex a sample of 4 bus crashes show on average 49 people died with standard deviation of 15. Find 99% confidence interval for the number of fatalities in a bus crash nationwide.

$n = 4$, $dof = 3$, $\bar{x} = 49$, $s = 15$, $C = 0.99$

หา margin error $= t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$

$= t_{\frac{0.01}{2}} \frac{15}{\sqrt{4}}$

$= 5.84 \left(\frac{15}{\sqrt{4}}\right)$ ต้องรูทด้วย

$= 43.8$ จากตาราง

โคตรกว้าง → data ไม่ค่อยดี หรือไม่พอ