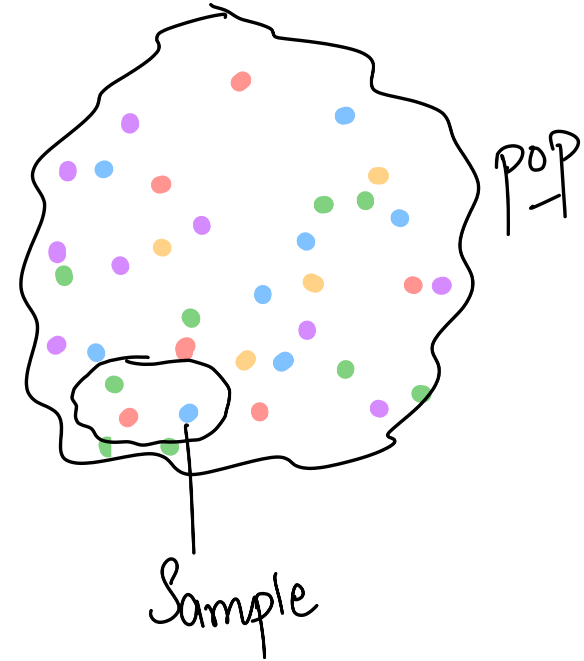
CENTRAL LIMIT THEOREM

① CLT → (Today)

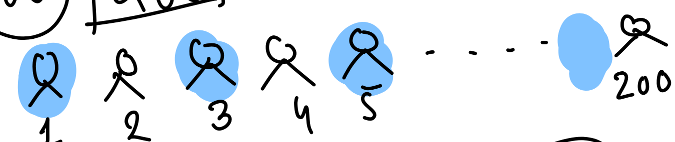
② CI → Confidence Interval (Next)



Sampling Techniques

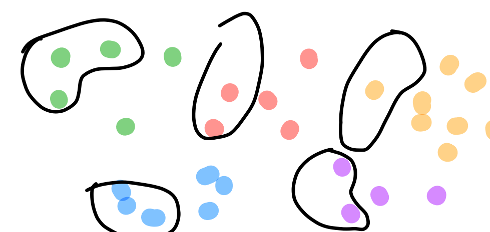
① Random Sampling

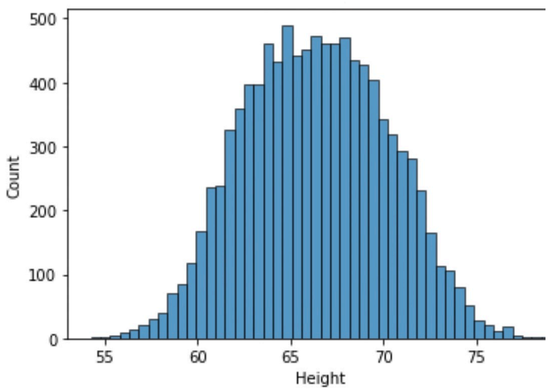
② Systematic Sampling 100 $\boxed{200}$

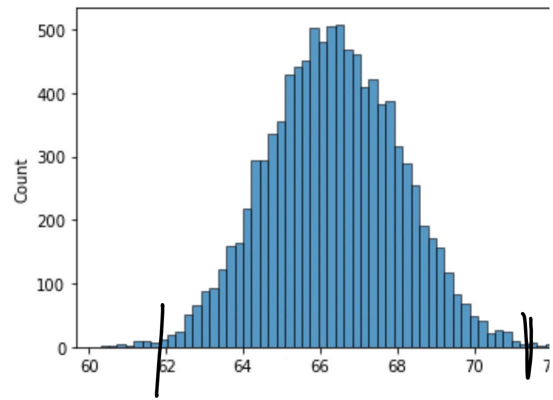


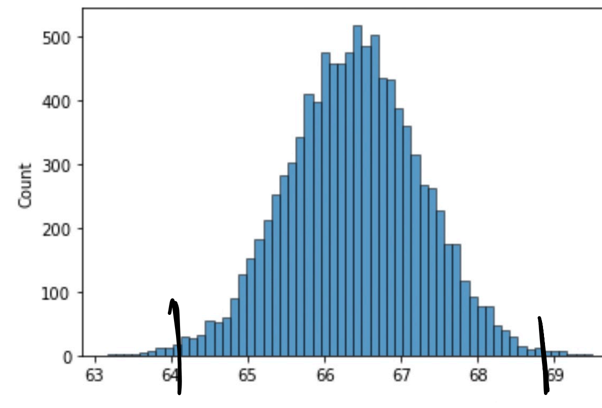
③ Stratified Sampling

Cluster based Sampling









Normal Distribution

Sample size = 5

Sample size = 20

$\mu = 66.36$

$\sigma = 3.8475$

$\mu = 66.34$

$\sigma_5 = 1.709$

$\mu = 66.38$

$\sigma_{20} = 0.849$

$\sigma > \sigma_5 > \sigma_{20}$

Central Limit Theorem

If "n" denotes sample size & "σ" denotes population standard deviation

$$\left\{ \begin{array}{l} n=5 \\ n=10 \\ 20 \\ 30 \\ \vdots \end{array} \right\}$$

then standard deviation of distribution of sampling means

$$\boxed{s = \frac{\sigma}{\sqrt{n}}}$$

AKA "Standard Error"

$$\bar{X} = \frac{X_1 + X_2 + X_3 + X_4 + X_5 + \dots X_n}{n}$$

$\bar{X}$ is Random Variable.

↳ denoting Sample means (RV)

$\bar{X}$ follow Gaussian Distribution
Normal Distribution

- Mean / Expectation $E[\bar{X}] = \mu$
- Std of $\bar{X} = \frac{\sigma}{\sqrt{n}}$

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

$\boxed{n > 30}$

* If our original pop^n is normally distributed.

n doesn't matter

$n = 1$ $\quad \bar{X} \sim N$

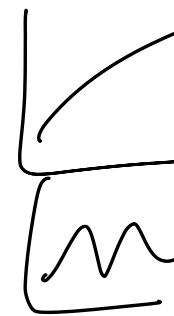
$n = 2$ $\quad \bar{X} \sim N$

Z test $\rightarrow n > 30$

t test $\rightarrow n < 30$

↳ t student's distribution

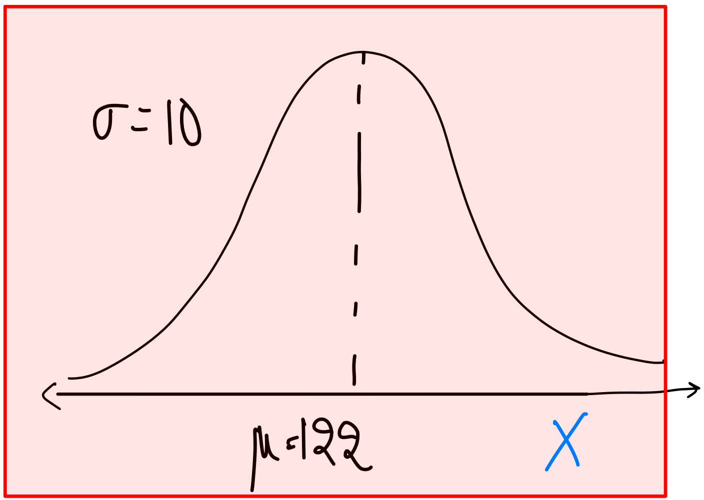
* If our original pop^n is not normally distributed.



$\bar{X} \sim N$ $\quad n > 30$

Systolic blood pressure of a group of people is known to have an average of 122 mmHg and a standard deviation of 10 mmHg

Calculate the probability that the average blood pressure of 16 people will be greater than 125 mmHg.



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

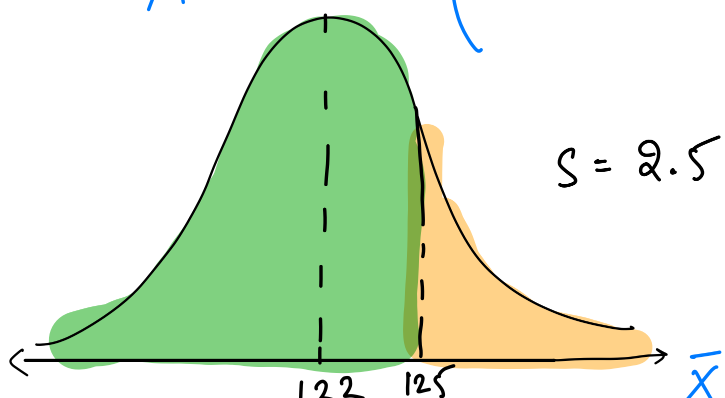
Sample mean

$n=16$

$$\bar{X} \sim N(122, 2.5)$$

$$P[X>125] = 1 - P[X<125]$$

norm.cdf(z score)



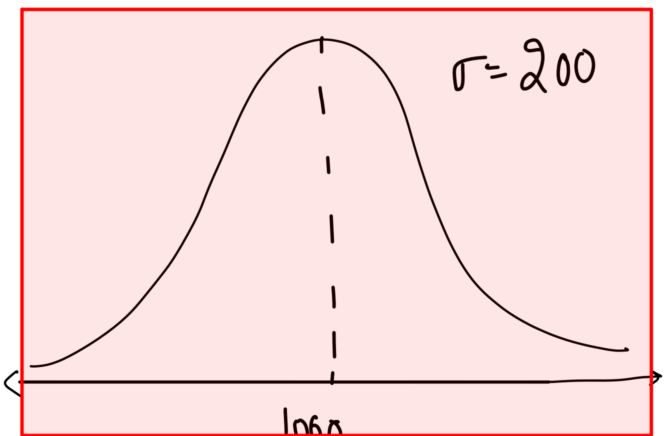
Weekly toothpaste sales have a mean 1000 and std dev 200. What is the probability that the average weekly sales next month is more than 1110?

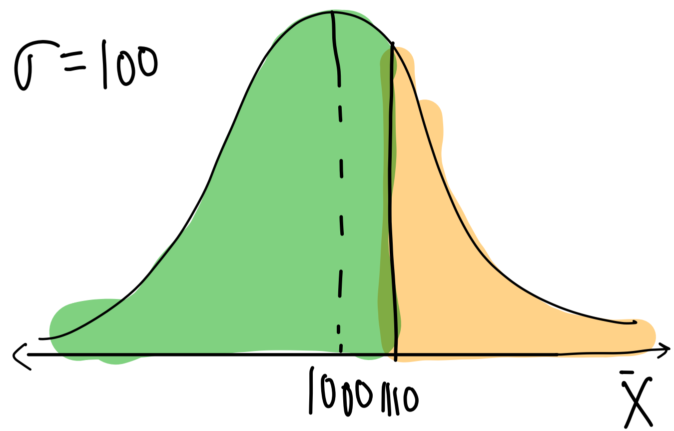
32 users have participated

A	0.29	28%
B	0.13	41%
C	0.11	22%
D	0.08	9%

$$\bar{X} = \frac{X_1 + X_2 + X_3 + X_4}{4} \quad \rightarrow \quad \underline{n = 4}$$

$$\bar{X} \sim N\left(1000, \frac{200}{\sqrt{4}}\right)$$





In an e-commerce website, the average purchase amount per customer is \$80 with a standard deviation of \$15. If we randomly select a sample of 50 customers, what is the probability that the average purchase amount in the sample will be less than \$75?

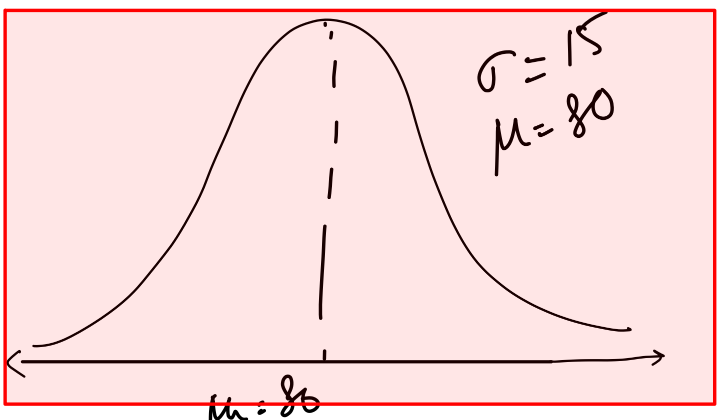
23 users have participated

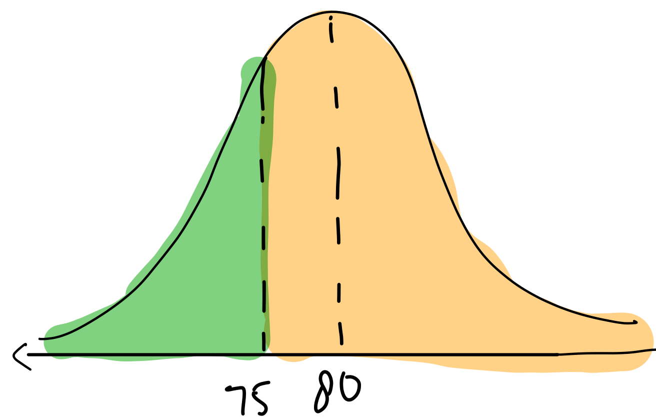
A	0.36	13%
B	0.18	30%
C	0.01	13%
D	0.009	43%

$\bar{X} \sim N\left(80, \frac{15}{\sqrt{50}}\right)$

$n = 50$

$\boxed{0.009}$





Aw pr

Manual

ML

1

DL

NN

$P[A/B]$

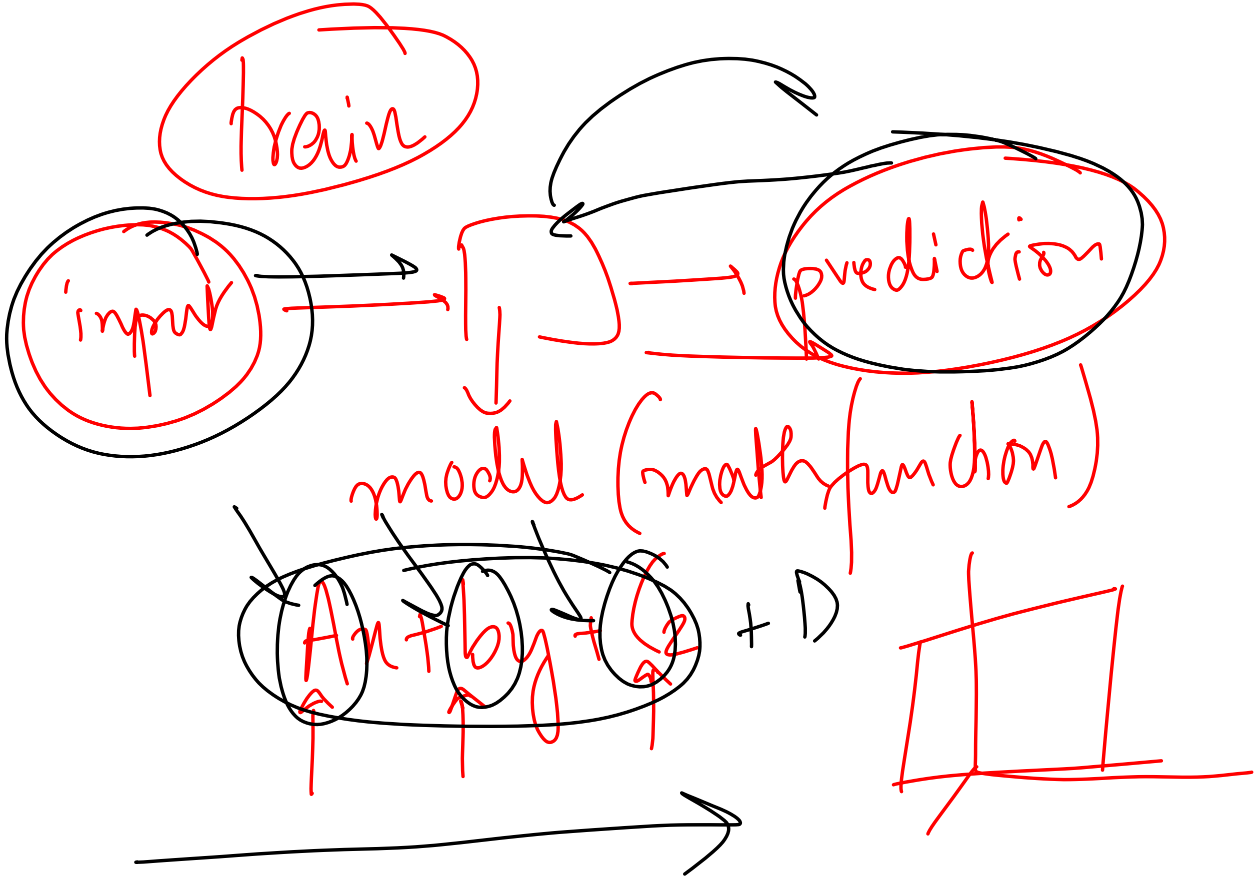
$P[A/B \cap C]$

$P[A/B \cap C \cap D]$

A B C

Income Gender Age

Gradient Descent


train
input
prediction
model (math function)
Ax+by+Cz + D