ISyE 3770, Spring 2024
Statistics and Applications

Sampling Distribution

Instructor: Jie Wang
H. Milton Stewart School of
Industrial and Systems Engineering
Georgia Tech

jwang3163@gatech.edu
Office: ISyE Main 447

Model for Samples: Random sampling

Random Sample

The random variables $X_1, X_2, \ldots, X_n$ are a **random sample** of size n if (a) the X_i's are independent random variables, and (b) every X_i has the same probability distribution.

$X_1, \ldots, X_n$ are independent

$X_1, \ldots, X_n$ same distribution.

Observations in a random sample are also known as

independent and identically distributed (***i.i.d.***)

i.i.d.

random variables

Statistic

A **statistic** is any function of the observations in a random sample.

$$S^2 = \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{n-1}$$

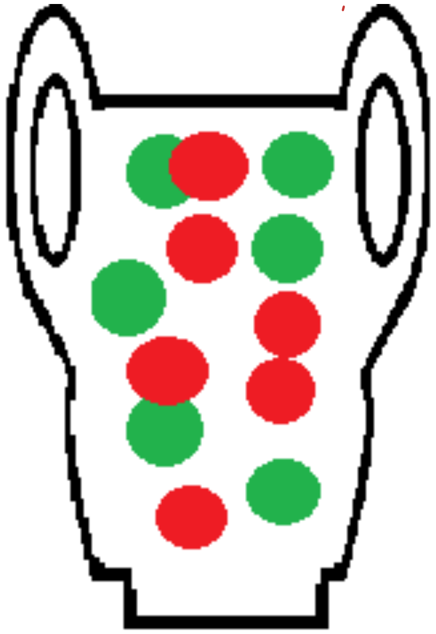
$$\text{e.g., } X_1, X_2, \cdots, X_n \rightarrow \bar{X}, S^2$$

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

The probability distribution of a statistic is called a **sampling distribution.**

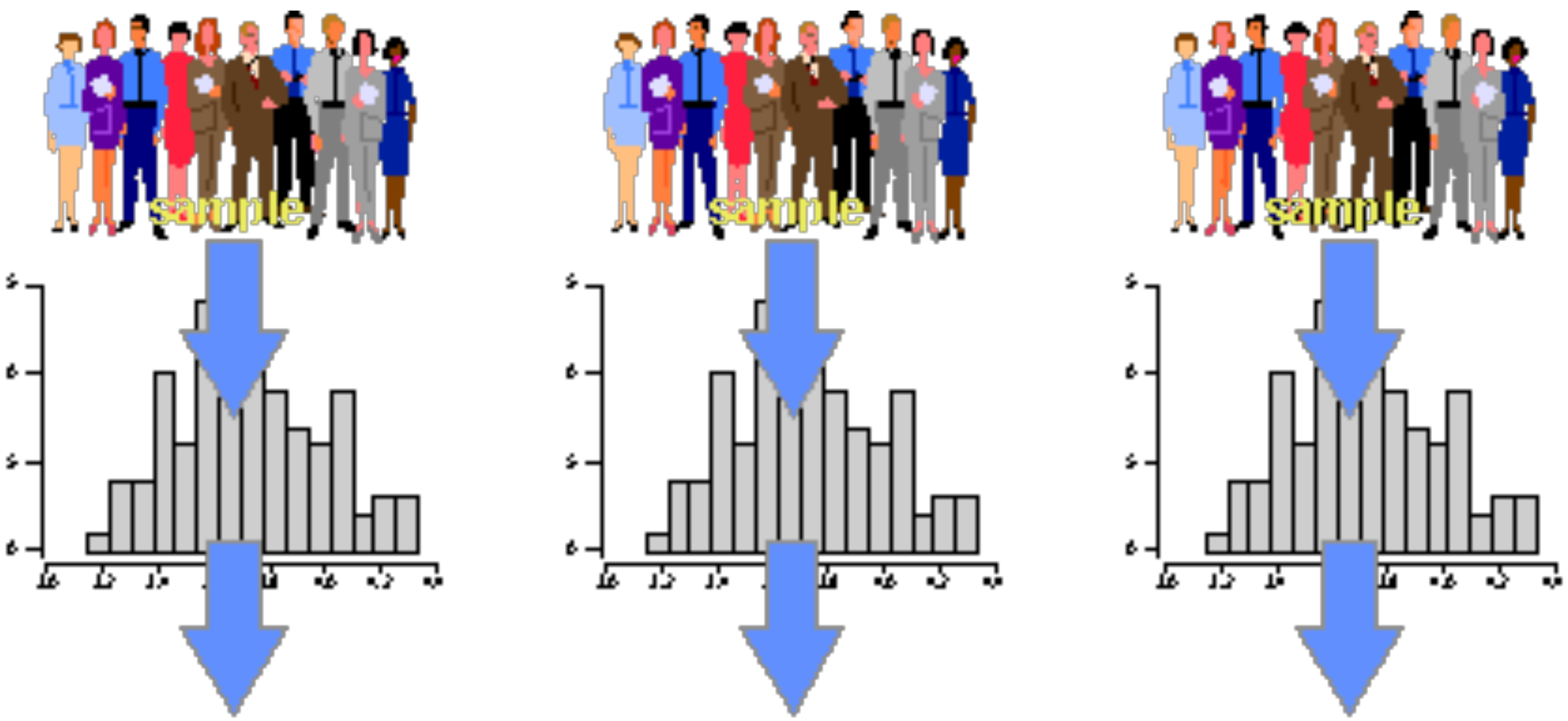
Goal: study distribution of $\bar{X}$ or S^2

Class activity: urn model



Sampling distribution

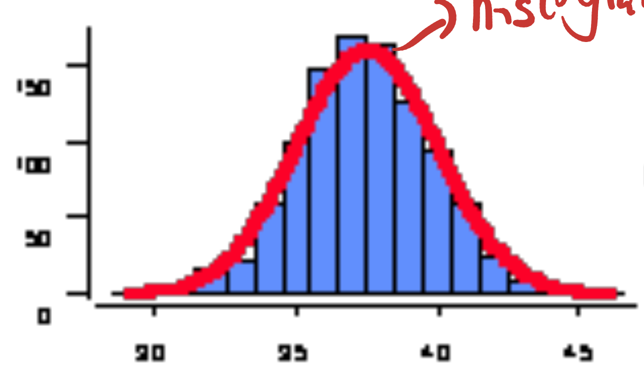
population



$\bar{X}^{(1)}$ **Average** | **Average** | $\bar{X}^{(2)}$ **Average** $\bar{X}^{(3)}$ ..... $\bar{X}^{(N)}$

histogram of $\{\bar{X}^{(n)}\}_{n\geq 1}$.

The Sampling Distribution...



...is the distribution of a statistic across an infinite number of samples

Why sampling distribution?

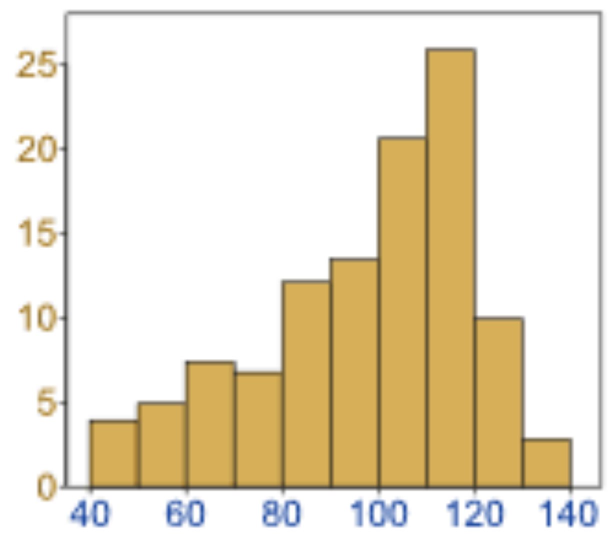
The probability distribution of a statistic is called a **sampling distribution.**

Statistical inference is concerned with making decisions about a population based on the information contained in a *random sample* from that population.

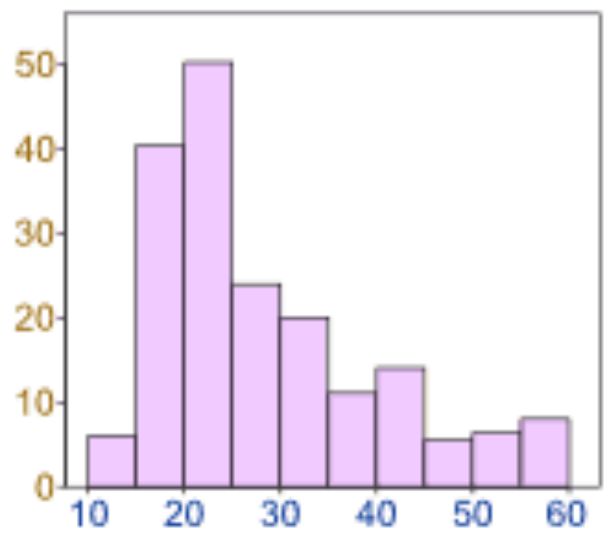
Sampling distribution is the link between probability and statistics.

Empirical distribution

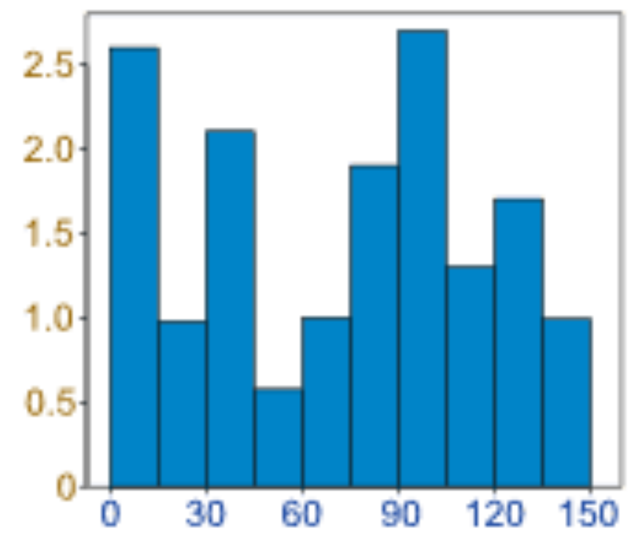
- **Data can be "distributed" (spread out) in different ways**



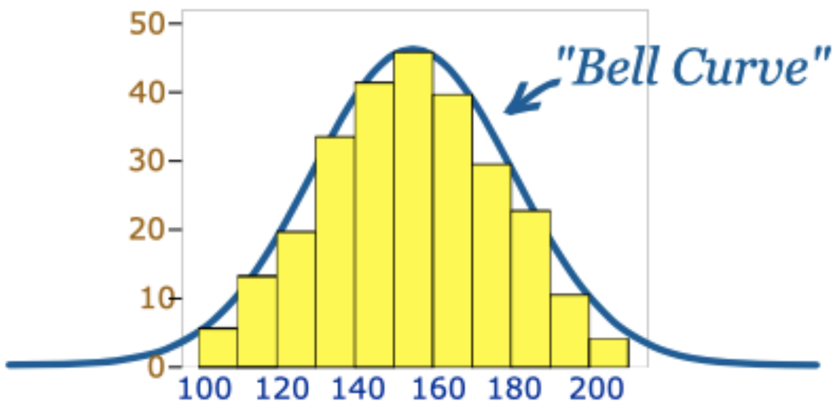
It can be spread out more on the left



Or more on the right



Or it can be all jumbled up



A Normal Distribution

Model sampling distribution

- **Relationships between Bernoulli and Binomial distributions**

$$X_i \sim BERN(p), i = 1, 2, \dots, n$$

random sample

$$\bar{X} = \sum_{i=1}^{n} X_i \sim BIN(n, p)$$

statistic

$\bar{X} \sim$ Binomial(n, p)

sampling distribution

- **In this setting**
- **Each time X_i is the outcome of each draw:**

  = 1, if black, otherwise = 0
- **$\bar{X}$ is the number of black stones**
- **Multiple experiments $\bar{X}$ is different and has variability**

Alternative view

- **Sample proportion is the percentage of black stones**

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$X_i \sim Ber(p)$

$E[X_i] = p$

$Var(X_i) = p(1-p)$

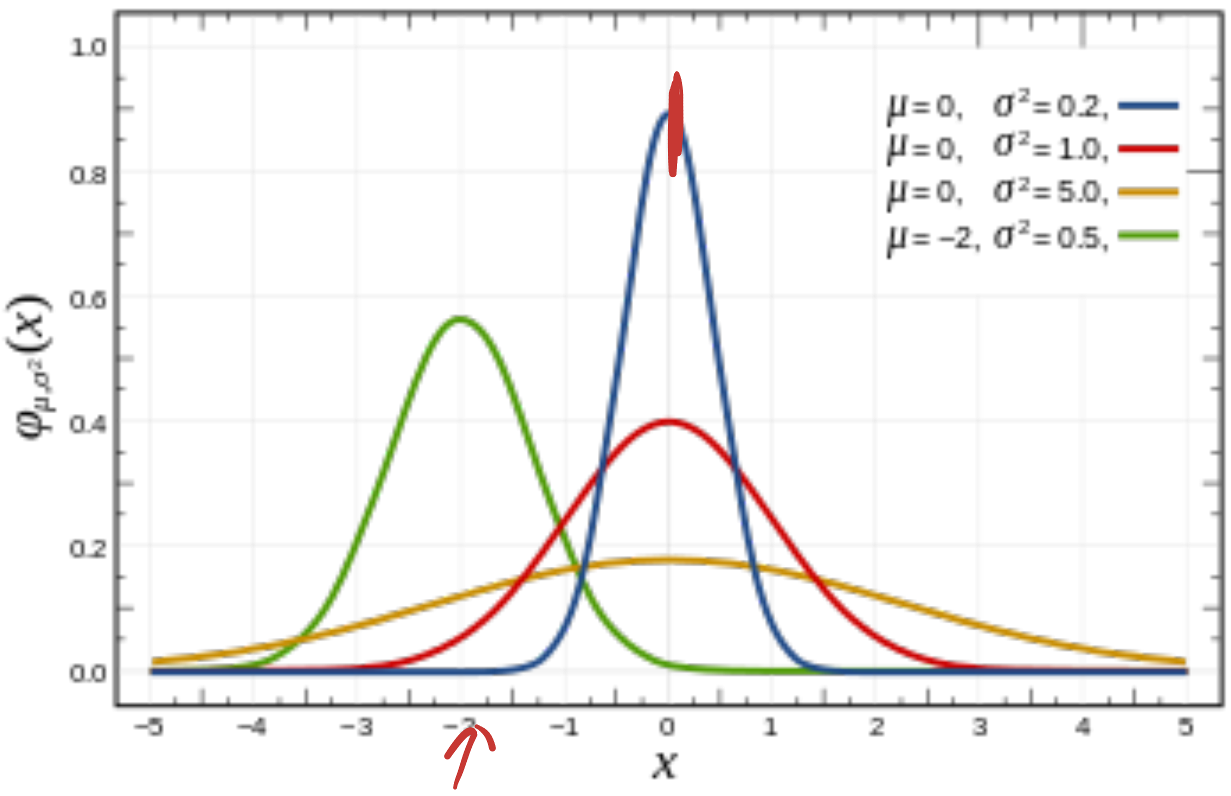
- **Claim: $\bar{X}$ is approximately normal distributed with mean p and variance = $\frac{p(1-p)}{n}$**

Suppose $X_1, \cdots, X_n \overset{i.i.d.}{\sim}$ distribution with mean μ, variance σ^2

$\bar{X} \overset{\text{approximately}}{\sim} N(\mu, \frac{\sigma^2}{n})$

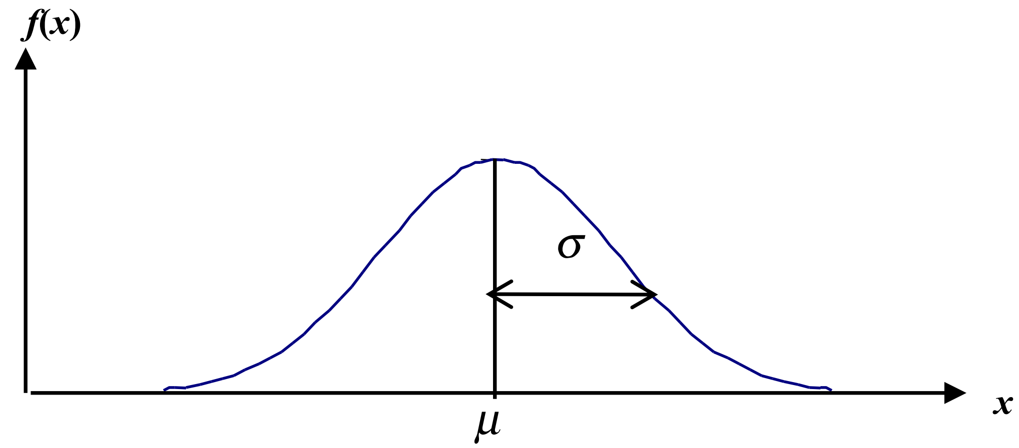
Sampling distribution describes the distribution of sample mean

Normal Distribution



Normal Distribution

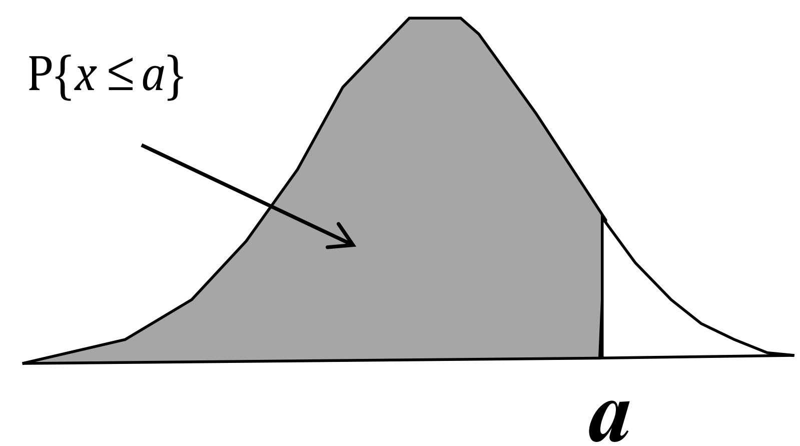
$$X \sim N(\mu, \sigma^2); \qquad -\infty < x < +\infty$$



$$f(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$E(x) = \mu \qquad Var(x) = \sigma^2$$

$$P\{x \le a\} = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} dx$$



Important Fact

$x_1 \sim N(\mu_1, \sigma_1^2)$ $x_2 \sim N(\mu_2, \sigma_2^2)$

- **Fact:** If x_1, x_2 are independently normally distributed variables, then

$$y = x_1 + x_2$$

also follows the normal distribution:

$$y \sim N(\mu_1+\mu_2, \sigma_1^2 + \sigma_2^2)$$

It can be shown by deriving cdf of y.

Special case

- **Making normal assumption about samples:**

  Xi' s are **normally** independently distributed (a random sample from a Normal distribution with the known variance)

$$X_1, X_2, \cdots, X_n \sim N(\mu, \sigma^2) \rightarrow \bar{X} \sim N(\mu, \frac{\sigma^2}{n}) \rightarrow \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

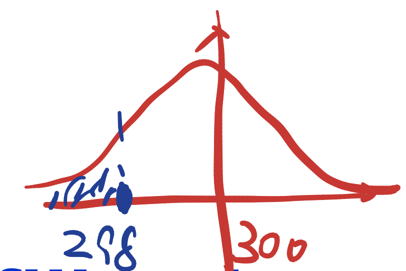
- **Proof?**

Sampling distribution of <u>sample mean</u> is normal, when samples are normal

• Remark: If X_i are i.i.d, not normally distributed, $\bar{X}$ is approximately $N(\mu, \frac{\sigma^2}{n})$. typically $n \geq 30$.

Example

1. The design of the machine has fill volume 300 mls, and variance 9ml. An engineer takes a random sample of 25 cans, what's the sampling distribution of mean filling volume of a can of soft drink?



2. The engineer finds the sample mean of fill volume to be 298 mls. Is this considered to be normal?

$X_1, \dots, X_{25} \sim N(300, 9)$

$\bar{X} = \frac{\sum_{i=1}^{25} X_i}{25} \sim N(300, \frac{9}{25})$

Pnorm (-3.3)

$P(\bar{X} \le 298)$

$= P(N(300, \frac{9}{25}) \le 298)$

$= P(N(0,1) \le \frac{298-300}{3/5})$

$= P(N(0,1) \le -3.3) = 0.0004$

Review.

- Give a R.V. X
- Consider n i.i.d observations, $X_1, \cdots, X_n$.

  i.i.d: independent identically distributed.

  Assume $X_1, \cdots, X_n$ have same distribution as X.

- We say $X_1, \cdots, X_n$ is a random sample, with observation size n.

- A statistic is a function of random sample.

$$\bar{X} = \frac{1}{n} \sum_i X_i \qquad S^2 = \frac{1}{n-1} \sum_i (X_i - \bar{X})^2$$

- Sampling distribution describes distribution of statistic.

Prop. $X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2) \Rightarrow \bar{X} = \frac{1}{n}\sum_i X_i \sim N(\mu, \frac{\sigma^2}{n})$

Extension: Suppose $X_1, \ldots, X_n$ are i.i.d., not necessarily normal, with mean μ, variance σ^2.

$\bar{X} = \frac{1}{n} \sum_i X_i$ approximately follows $N(\mu, \frac{\sigma^2}{n})$.

(nearly holds when $n \geq 30$).

Review:

- Given a random variable X
- Consider n i.i.d. observations

  $X_1, \dots, X_n$, following same distribution as in X

- $X_1, \dots, X_n$ are called a random sample of size n.
- A statistic is a function of random sample,

  i.e., $\bar{X} = \frac{1}{n} \sum_i X_i$

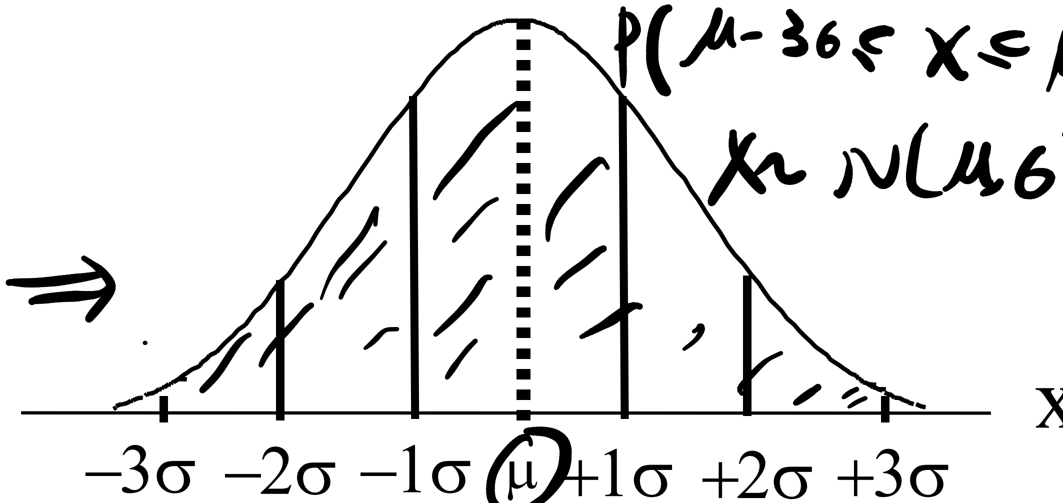
Prop. Let $X_1, \dots, X_n \sim N(\mu, \sigma^2)$. $\Rightarrow$ $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

Standard Normal Distribution

$X \sim N(\mu, \sigma^2)$; $\quad -\infty < x < +\infty$

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}$$

$\Rightarrow$



$P(\mu - 3\sigma \le X \le \mu + 3\sigma)$

$X \sim N(\mu, \sigma^2)$

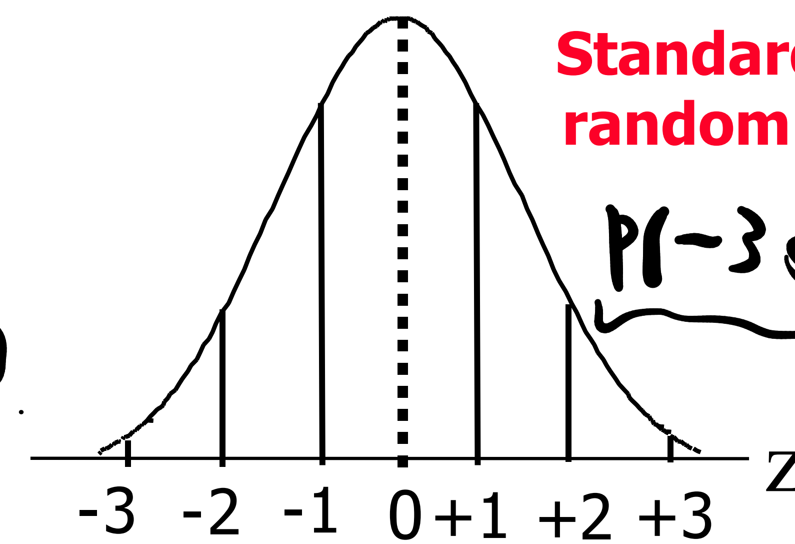
Map any X into Z

$$\boxed{Z = \frac{X-\mu}{\sigma}}$$

$Z \sim N(0, 1^2)$; $\quad -\infty < z < +\infty$

$$f(z;\mu=0,\sigma^2=1) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z^2}{2}\right)$$

$\Rightarrow$



Standard normal random variable

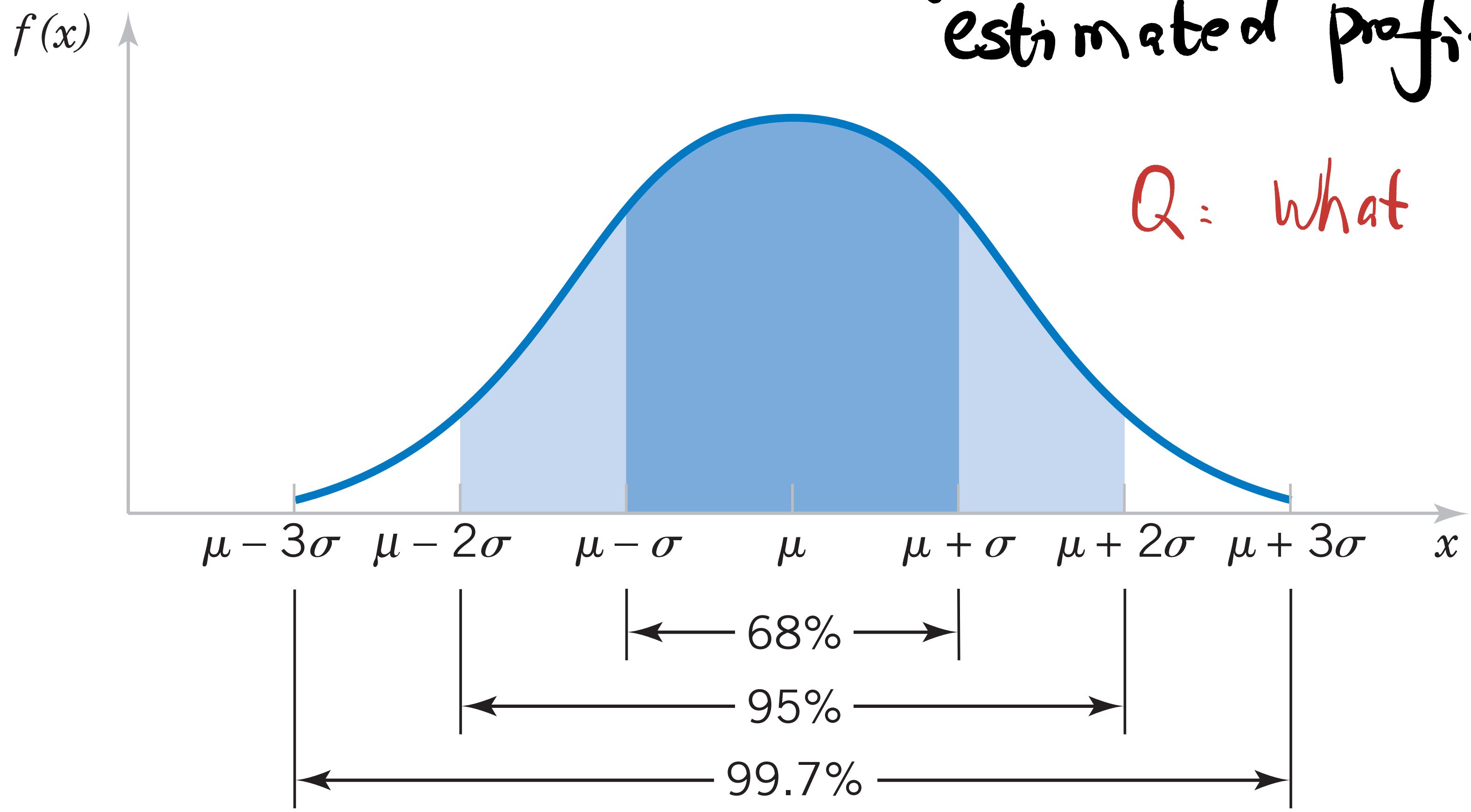
$P(-3 \le Z \le +3)$

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

↓

estimated profit.

Q: what profit you want to report?



$$\mathbb{P}(\mu - \sigma \le X \le \mu + \sigma) = \mathbb{P}(-1 \le Z \le 1) = 0.6827,$$

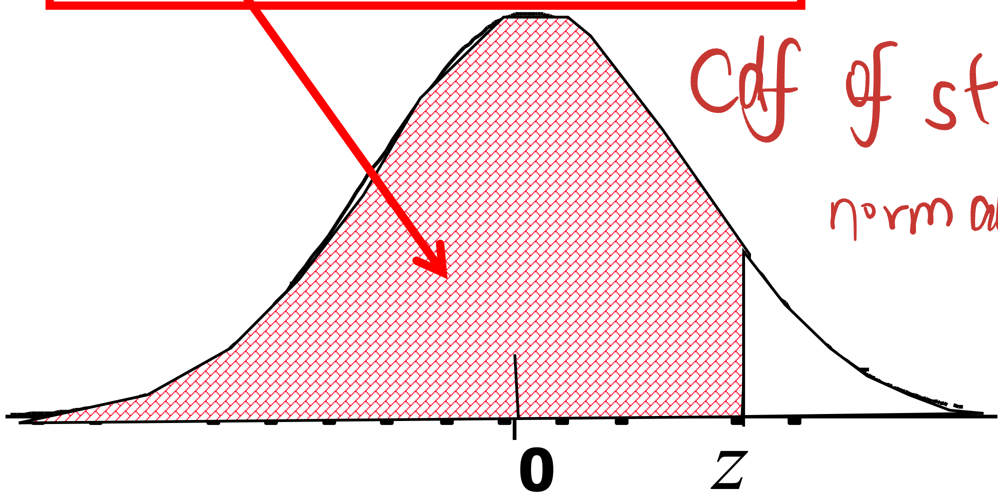
$$\boxed{\mathbb{P}(\mu - 2\sigma \le X \le \mu + 2\sigma) = \mathbb{P}(-2 \le Z \le 2) = 0.9545,} \; \star$$

$$\mathbb{P}(\mu - 3\sigma \le X \le \mu + 3\sigma) = \mathbb{P}(-3 \le Z \le 3) = 0.9973. \; \star$$

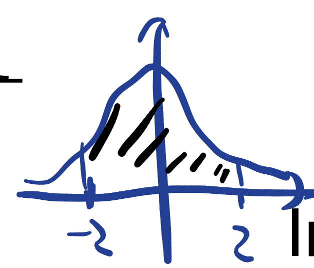
Z-Values

$$\boxed{\Phi(z) = P(Z \le z)}$$

cdf of standard normal distribution

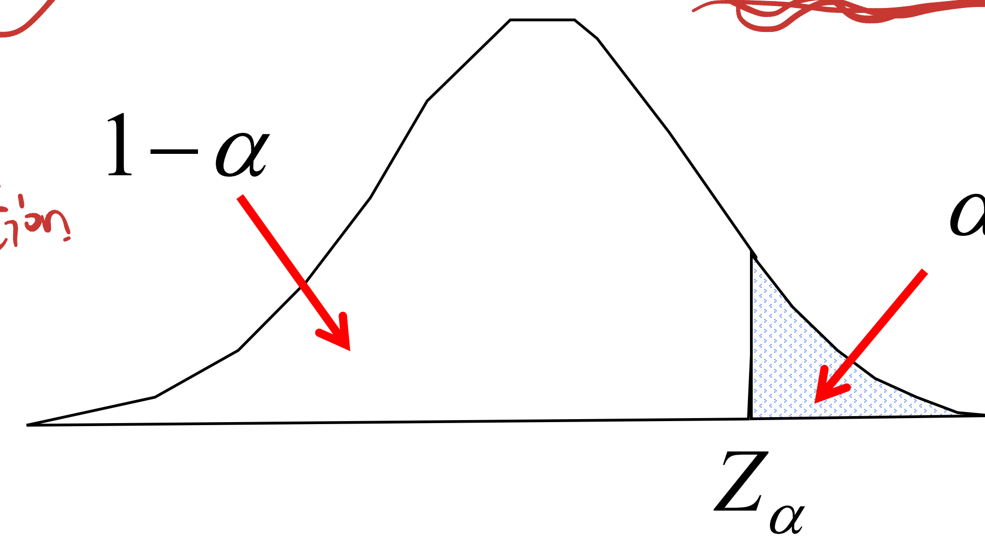


Input: real number (z)
Output: probability Φ(z)



$\phi(2) \approx 0.975$

$$\boxed{Z_\alpha = \Phi^{-1}(1-\alpha)} \Leftrightarrow P(Z > Z_\alpha) = \alpha$$



Input: probability α
Output: real number Z_α **Upper percentage point**

$$\Phi(1.645) = P(Z < 1.645) = 0.95 \Leftrightarrow z_{0.05} = \Phi^{-1}(0.95) = 1.645$$

Useful equations:

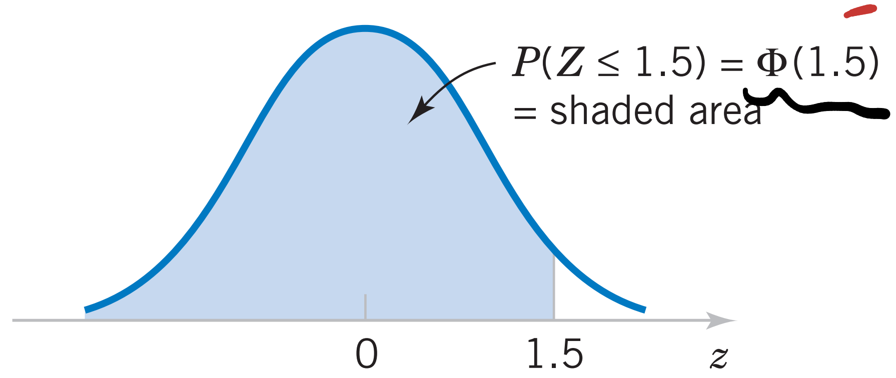
$$\Phi(Z_\alpha) = P(Z < Z_\alpha) = 1-\alpha \Rightarrow \alpha = 1 - \Phi(Z_\alpha)$$

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

$$Z_{1-\alpha} = -Z_\alpha$$

Normal table

Figure 4-13 Standard normal probability density function.



$= 0.93$

z	0.00	0.01	0.02	0.03
0	0.50000	0.50399	0.50398	0.51197
⋮		⋮		
1.5	0.93319	0.93448	0.93574	0.93699

1.5 : row
0.00 : column

$\Rightarrow 0.93$

$\Phi(2.5)$?

R studio

pnorm(2.5, 0, 1) = 0.9937

$$\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}\, du$$

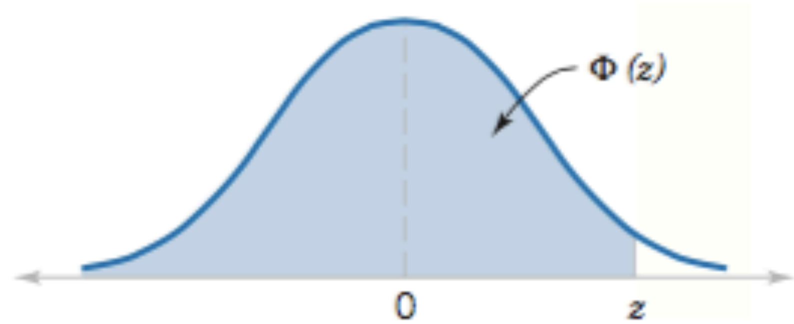


Table III Cumulative Standard Normal Distribution (*continued*)

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.500000	0.503989	0.507978	0.511967	0.515953	0.519939	0.532922	0.527903	0.531881	0.535856
0.1	0.539828	0.543795	0.547758	0.551717	0.555760	0.559618	0.563559	0.567495	0.571424	0.575345
0.2	0.579260	0.583166	0.587064	0.590954	0.594835	0.598706	0.602568	0.606420	0.610261	0.614092
0.3	0.617911	0.621719	0.625516	0.629300	0.633072	0.636831	0.640576	0.644309	0.648027	0.651732
0.4	0.655422	0.659097	0.662757	0.666402	0.670031	0.673645	0.677242	0.680822	0.684386	0.687933
0.5	0.691462	0.694974	0.698468	0.701944	0.705401	0.708840	0.712260	0.715661	0.719043	0.722405
0.6	0.725747	0.729069	0.732371	0.735653	0.738914	0.742154	0.745373	0.748571	0.751748	0.754903
0.7	0.758036	0.761148	0.764238	0.767305	0.770350	0.773373	0.776373	0.779350	0.782305	0.785236
0.8	0.788145	0.791030	0.793892	0.796731	0.799546	0.802338	0.805106	0.807850	0.810570	0.813267
0.9	0.815940	0.818589	0.821214	0.823815	0.826391	0.828944	0.831472	0.833977	0.836457	0.838913
1.0	0.841345	0.843752	0.846136	0.848495	0.850830	0.853141	0.855428	0.857690	0.859929	0.862143
1.1	0.864334	0.866500	0.868643	0.870762	0.872857	0.874928	0.876976	0.878999	0.881000	0.882977
1.2	0.884930	0.886860	0.888767	0.890651	0.892512	0.894350	0.896165	0.897958	0.899727	0.901475
1.3	0.903199	0.904902	0.906582	0.908241	0.909877	0.911492	0.913085	0.914657	0.916207	0.917736
1.4	0.919243	0.920730	0.922196	0.923641	0.925066	0.926471	0.927855	0.929219	0.930563	0.931888
1.5	0.933193	0.934478	0.935744	0.936992	0.938220	0.939429	0.940620	0.941792	0.942947	0.944083
1.6	0.945201	0.946301	0.947384	0.948449	0.949497	0.950529	0.951543	0.952540	0.953521	0.954486

Normal table

$$\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}\, du$$

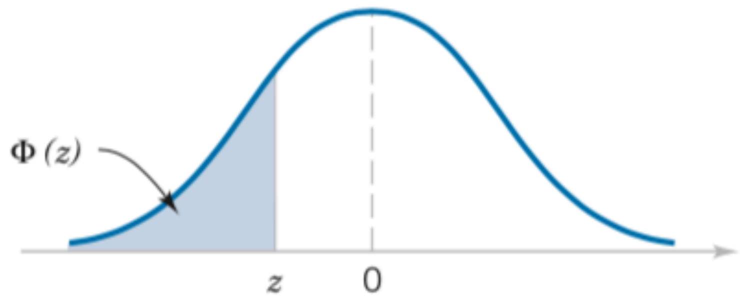


Table III Cumulative Standard Normal Distribution

z	−0.09	−0.08	−0.07	−0.06	−0.05	−0.04	−0.03	−0.02	−0.01	−0.00
−3.9	0.000033	0.000034	0.000036	0.000037	0.000039	0.000041	0.000042	0.000044	0.000046	0.000048
−3.8	0.000050	0.000052	0.000054	0.000057	0.000059	0.000062	0.000064	0.000067	0.000069	0.000072
−3.7	0.000075	0.000078	0.000082	0.000085	0.000088	0.000092	0.000096	0.000100	0.000104	0.000108
−3.6	0.000112	0.000117	0.000121	0.000126	0.000131	0.000136	0.000142	0.000147	0.000153	0.000159
−3.5	0.000165	0.000172	0.000179	0.000185	0.000193	0.000200	0.000208	0.000216	0.000224	0.000233
−3.4	0.000242	0.000251	0.000260	0.000270	0.000280	0.000291	0.000302	0.000313	0.000325	0.000337
−3.3	0.000350	0.000362	0.000376	0.000390	0.000404	0.000419	0.000434	0.000450	0.000467	0.000483
−3.2	0.000501	0.000519	0.000538	0.000557	0.000577	0.000598	0.000619	0.000641	0.000664	0.000687

Exercise

$1 - \text{pnorm}(0.3, 0, 1)$

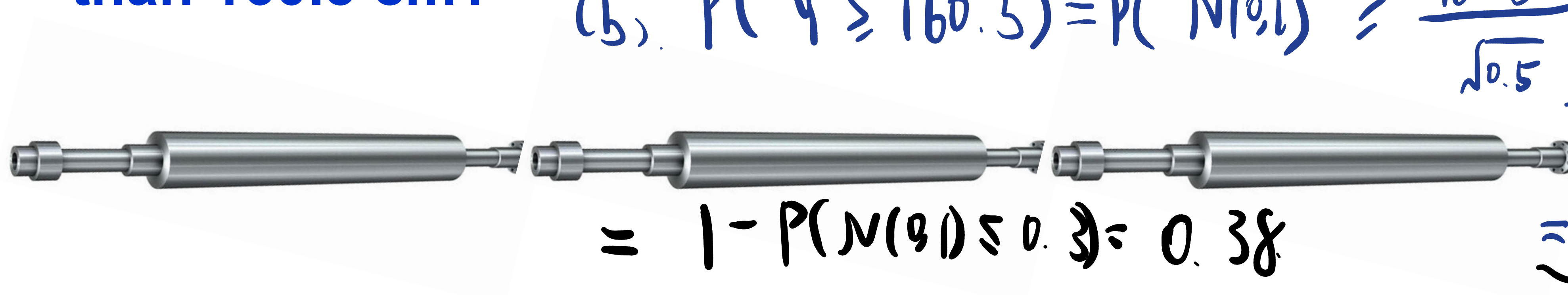
- **Example: Three shafts are made and assembled in a machine. The length of each shaft, in centimeters, is distributed as follows:**

$(a)\ Y = X_1 + X_2 + X_3 \sim N(75+60+25,\ 0.09+0.16+0.25) = N(160, 0.5).$

Shaft 1: ~ N (75, 0.09) $\to X_1$
Shaft 2: ~ N (60, 0.16) $\to X_2$
Shaft 3: ~ N (25, 0.25) $\to X_3$

Assume the shafts' length are independent to each other:
(a) What is the distribution of the linkage?
(b) What is the probability that the linkage will be longer than 160.5 cm?

$(b)\ P(Y \geq 160.5) = P\left(N(0,1) \geq \frac{160.5-160}{\sqrt{0.5}}\right) = P\left(N(0,1) \geq \frac{\sqrt{2}}{2}\right) = P(N(0,1) \geq 0.7)$



$= 1 - P(N(0,1) \leq 0.3) = 0.38$

Exercise

$X \sim N(40, 5^2)$

$p(x \le 37.9) =$ $\underline{\text{pnorm}(37.9, 40, 5)} = 0.33$

$p(x \ge a) = 0.3783$ $\Leftrightarrow$ Find a s.t. $P(X \le a) = 1 - 0.3783$

$= 0.62$

$\underline{\text{qnorm}(0.62, 40, 5)} = 41.52$

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$$

Appendix A: Table III

z	0.00	0.01	0.02	0.03	0.04	z
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.0
0.1	0.53983	0.54379	0.54776	0.55172	0.55567	0.1
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.2
0.3	0.61791	0.62172	0.62551	0.62930	0.63307	0.3
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.4
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.5
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.6
0.7	0.75803	0.76115	0.76424	0.76730	0.77035	0.7

Exercise: Airport Check-in

The amount of time that a customer spends waiting in the airport check-in counter is a normal random variable with mean 8.2 minutes and standard deviation 1.5 minutes. Suppose that a random sample of 49 customers is observed. $X_1, \dots, X_n \overset{i.i.d.}{\sim} N(8.2, 1.5^2)$ $n = 49$

$\bar{X} \sim N(8.2, \frac{1.5^2}{49})$ $std(\bar{x}) = 1.5/7$

1. Find the probability that the average waiting time for these customers is:

(a) Less than 10 minutes; $Pr(\bar{X} \le 10)$

(b) Between 5 and 10 minutes. $= \text{pnorm}(10, 8.2, 1.5/7) = 1$

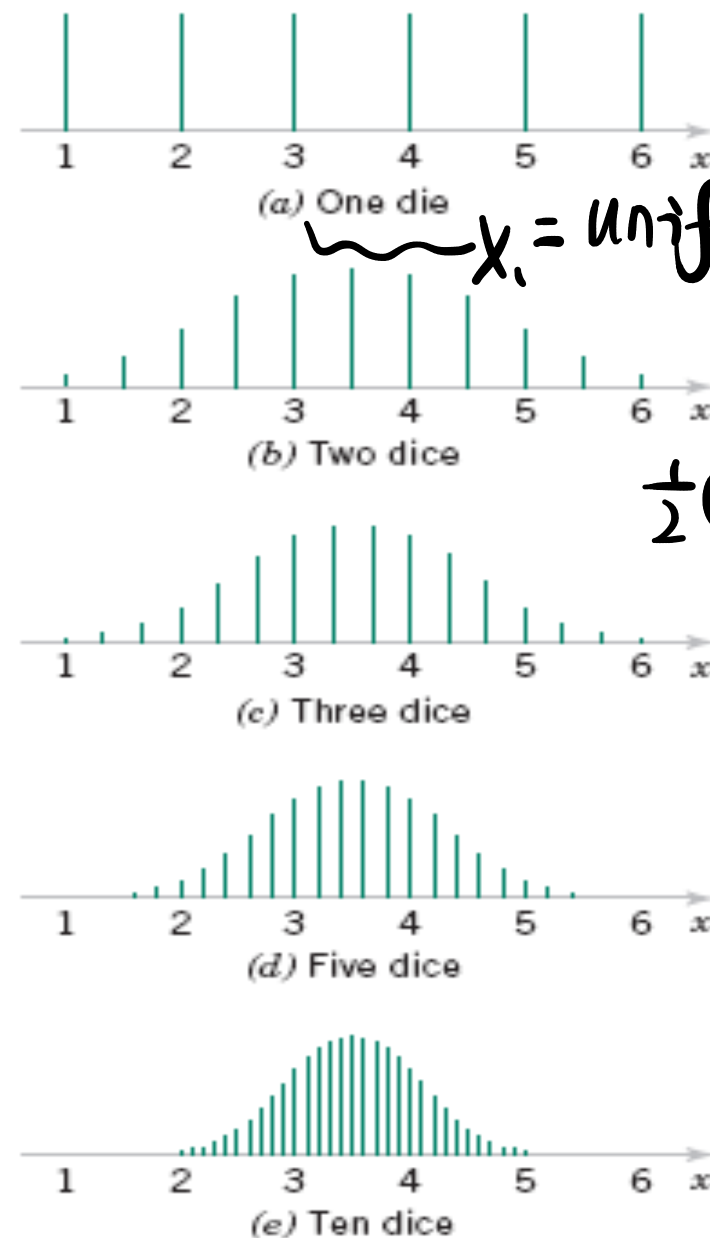
$Pr(5 \le \bar{X} \le 10) = 1$

2. What is a value such that 90% of chance, average wait time will wait shorter than that? Find a s.t. $P(\bar{X} \le a) = 0.9$

$\text{qnorm}(0.9, 8.2, 1.5/7) = 8.47$

General case: Central Limit Theorem (CLT)

Sampling distribution of <u>sample mean</u> is normal, even when samples are NOT normal



$X_1 = \text{uniform}(1, 2, \dots, 6)$

$\frac{1}{2}(X_1 + X_2)$

$\frac{1}{3}(X_1 + X_2 + X_3)$

Figure 7-1 Distributions of average scores from throwing dice. [Adapted with permission from Box, Hunter, and Hunter (1978).]

Rule of thumb: when n > 30 this works pretty well.

CENTRAL LIMIT THEOREM If all samples of a particular size are selected from any population, the sampling distribution of the sample mean is approximately a normal distribution. This approximation improves with larger samples.

Exercise: Coffee

For the "number of coffee drink / day" question, assume the number of coffee drink / day is a random variable with mean 1 and variance 0.5.

There are 58 responses of survey. What the sampling distribution of the sample mean?

- population X s.t. $E[X]=1$ $Var(X)=0.5$.
- $\bar{X} = \frac{1}{n}\sum_{i} X_n$ $n=58$,

$\sim N(1, \frac{0.5}{58})$ $X_1, \cdots, X_n \overset{i.i.d.}{\sim} X$.

Sampling Distribution of Sample Mean With Known Variance

One Population:

Xi' s are normally independently distributed (a random sample from a Normal distribution with the known variance)

$$X_1, X_2, \cdots, X_n \sim N(\mu, \sigma^2) \rightarrow \bar{X} \sim N(\mu, \frac{\sigma^2}{n}) \rightarrow \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

$\bar X - \bar Y = \bar X + (-\bar Y) \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$

Two Populations:

$\bar X \sim N(\mu_1, \frac{\sigma_1^2}{n_1})$

Two independent random samples from two Normal distributions with the known variances

$\bar Y \sim N(\mu_2, \frac{\sigma_2^2}{n_2})$, $-\bar Y \sim N(-\mu_2, \frac{\sigma_2^2}{n_2})$

$$\left.\begin{array}{l} X_1, X_2, \cdots, X_{n_1} \sim N(\mu_1, \sigma_1^2) \\ \\ Y_1, Y_2, \cdots, Y_{n_2} \sim N(\mu_2, \sigma_2^2) \end{array}\right\} \rightarrow \bar{X} - \bar{Y} \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

$P(N(-50, 136) \geq 25) = 1 - P(N(-50, 136) \leq 25) = 1 - 0.7 = 0.3$

Aircraft Engine Life

- The effective life of a component, X1, used in a jet-turbine aircraft engine is a random variable with mean 5000 hours and standard deviation 40 hours. The engine manufacture design a new component X2, which increases the mean life to 5050 hours and decreases the standard deviation to 30 hours. Assume X1 and X2 is fairly close to a normal distribution. Suppose n1 = 16 samples of old components, and n2 = 25 samples from the new components, are selected. What is the probability that the difference in two sample means is at least 25 hours?

$X_1 \sim (5000, 40^2)$ $X_2 \sim (5050, 30^2)$

$P(\bar{X}_1 - \bar{X}_2 \geq 25)$

$\bar{X}_1 - \bar{X}_2 \sim N\left(-50, \frac{40^2}{n_1} + \frac{30^2}{n_2}\right) = N\left(-50, \frac{40^2}{16} + \frac{30^2}{25}\right)$

$= N(-50, 136)$

Central Limit Theorem (CLT) for two populations

Approximate Sampling Distribution of a Difference in Sample Means

If we have two independent populations with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2, and if $\bar{X}_1$ and $\bar{X}_2$ are the sample means of two independent random samples of sizes n_1 and n_2 from these populations, then the sampling distribution of

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \tag{7-4}$$

is approximately standard normal, if the conditions of the central limit theorem apply. If the two populations are normal, the sampling distribution of Z is exactly standard normal.

$$\bar{X}_1 - \bar{X}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right).$$

- Group 1: population mean/var (μ_1, σ_1^2). observation size n_1
- Group 2: population mean/var (μ_2, σ_2^2) observation size n_2

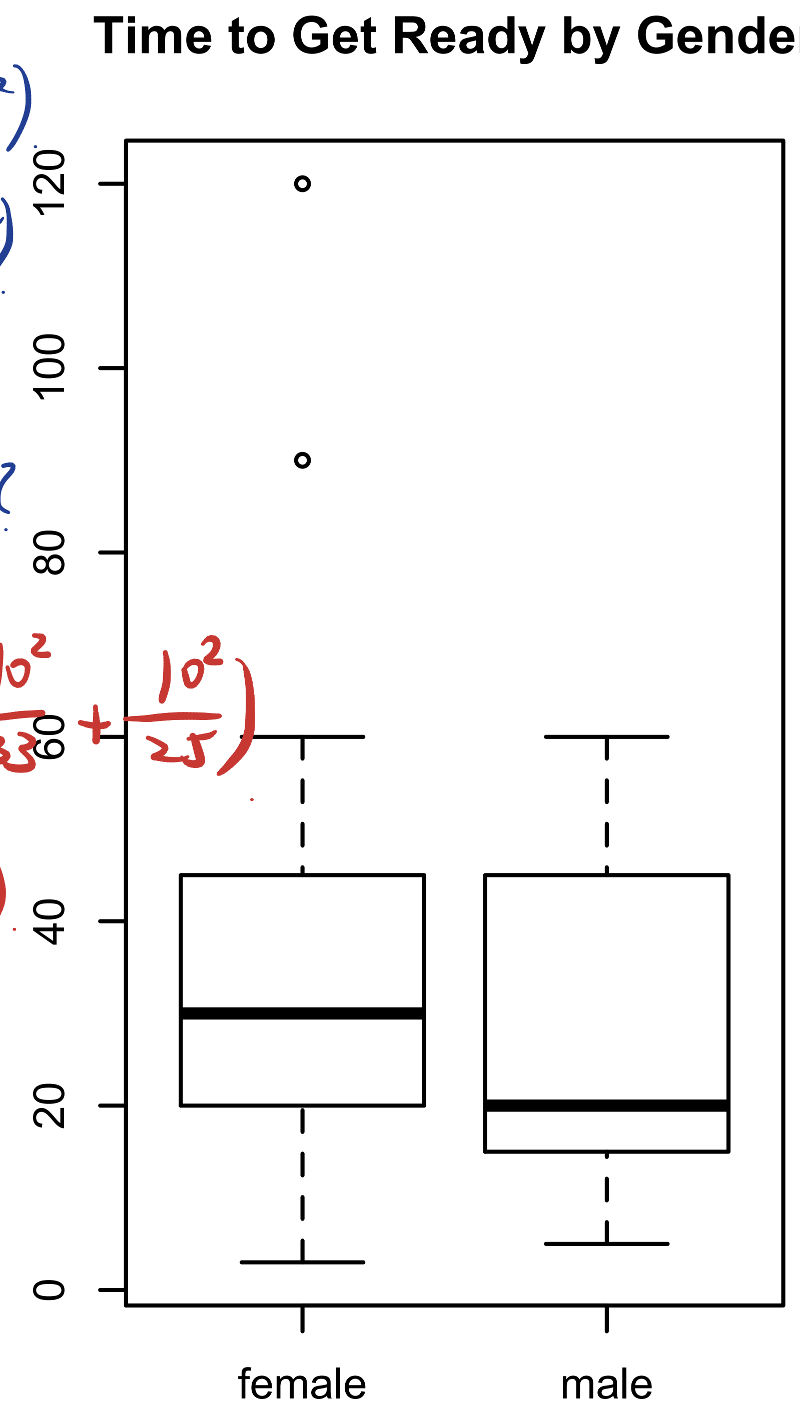
Example: Time in the morning

- The time for students to get ready in the morning, for male and female students.
- There are n1 = 33 girls and n2 = 25 guys who provides the answer
- Assume the time for girl is a random variable with mean 30 minutes, and the time for guy is a random variable with mean 20 minutes. The standard deviation for both of them is 10 minutes.
- What is the probability that the difference in two sample means is at least 10 minutes?

$X_1 \sim N(30, 10^2)$

$X_2 \sim N(20, 10^2)$

$P(\bar{X}_1 - \bar{X}_2 \geq 10)$?



$\bar{X}_1 - \bar{X}_2 \sim N\left(10, \frac{10^2}{33} + \frac{10^2}{25}\right)$

$= N(10, 7.03)$

$P(N(10, 7.03) \geq 10) = \frac{1}{2}$

Of course the variance is unknown…

Summary

For sample mean: (μ, σ^2) are known

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1) \qquad \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

For sample mean, μ is known
σ^2 is unknown

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1), \qquad \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t(n_1-1,\ n_2-2)$$

X X

For sample variance ? $\chi^2(n-1)$, $F(n_1-1, n_2-1)$.

X

Summary and Extension.

For sample mean,

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1), \quad \text{if } (\mu, \sigma^2) \text{ known.}$$

$$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1) \quad \text{if } (\mu, \sigma^2) \text{ known.}$$

For sample mean, but unknown variance,

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1), \qquad \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$$

For sample variance, $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$. $\quad \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1-1, n_2-1)$.