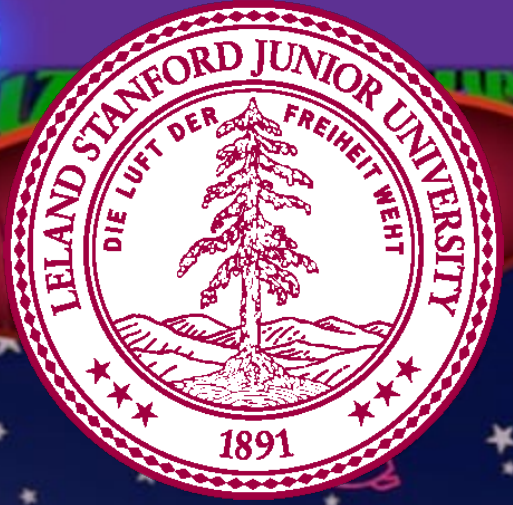


THE CLAW
CHALLENGE



TIMER: 29



POINTS: 01250



CLT and Sampling

Chris Piech

CS109, Stanford University

MAR
7TH



Review

The Insight to Convolution Proofs

What is the
probability that $X +$
 $Y = n$?

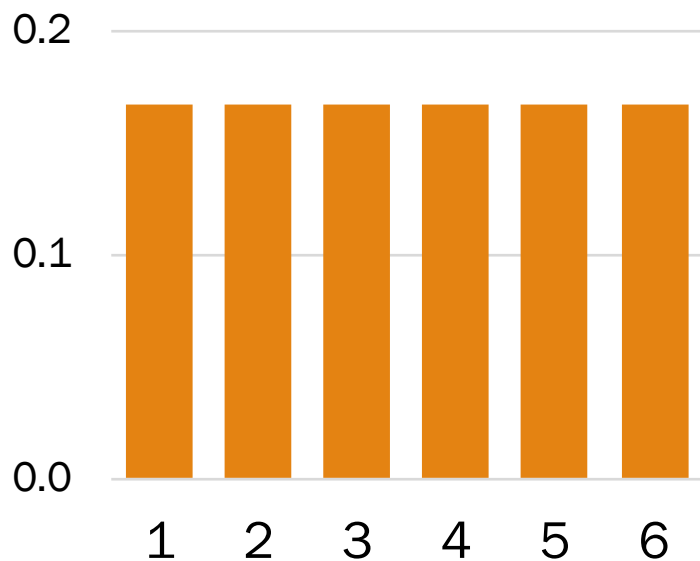
$$P(X + Y = n)?$$

$$P(X + Y = n) = \sum_{i=0}^n P(X = i, Y = n - i)$$

X	Y	i	
0	n	0	$P(X = 0, Y = n)$
1	$n - 1$	1	$P(X = 1, Y = n - 1)$
2	$n - 2$	2	$P(X = 2, Y = n - 2)$
	...		
n	0	n	$P(X = n, Y = 0)$

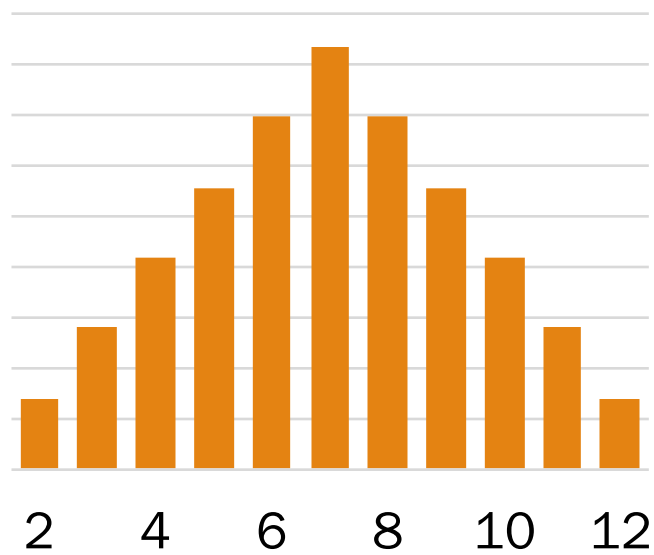
Sum of dice rolls

Roll n independent dice. Let X_i be the outcome of roll i . X_i are i.i.d.



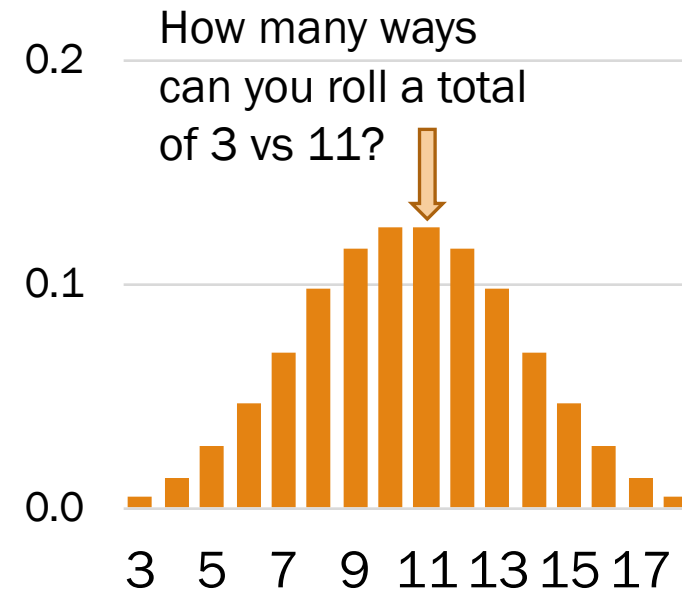
$$\sum_{i=1}^1 X_i$$

Sum of 1 die roll



$$\sum_{i=1}^2 X_i$$

Sum of 2 dice rolls



$$\sum_{i=1}^3 X_i$$

Sum of 3 dice rolls

Sum of 50 dice?

Central Limit Theorem

Consider n **independent and identically distributed (i.i.d)** variables X_1, X_2, \dots, X_n with $E[X_i] = \mu$ and $\text{Var}(X_i) = \sigma^2$.

$$\sum_{i=1}^n X_i \sim \mathcal{N}(n\mu, n\sigma^2) \quad \text{As } n \rightarrow \infty$$

The sum of n **i.i.d.** random variables is normally distributed with mean $n\mu$ and variance $n\sigma^2$.

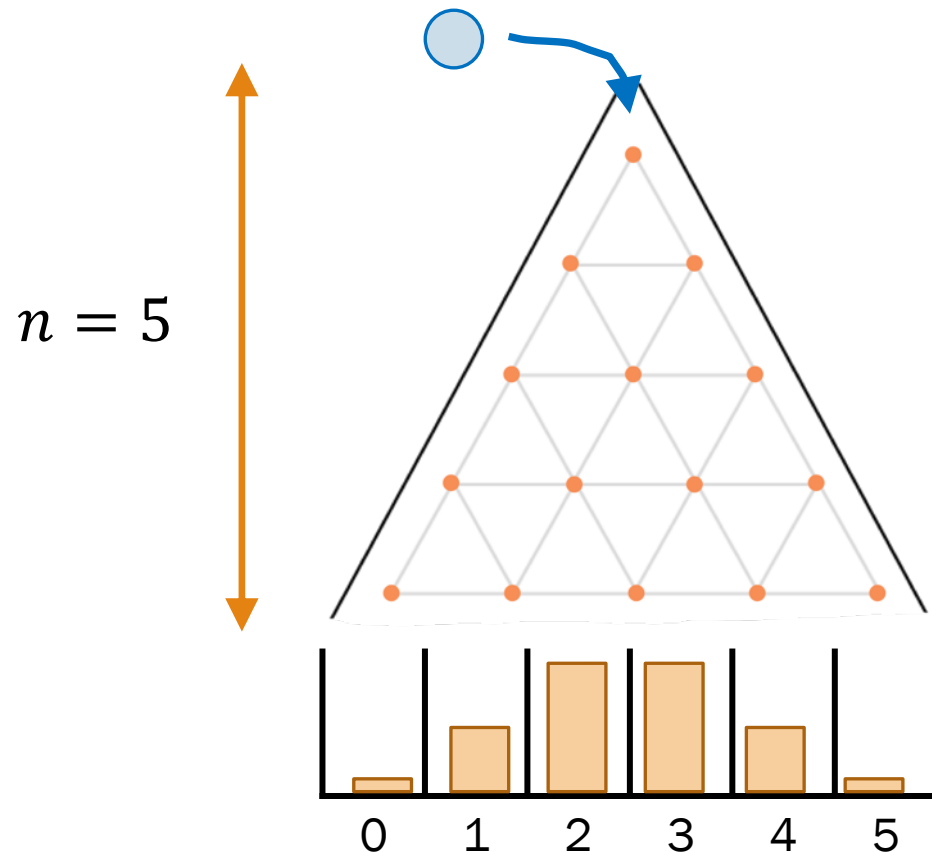
True happiness



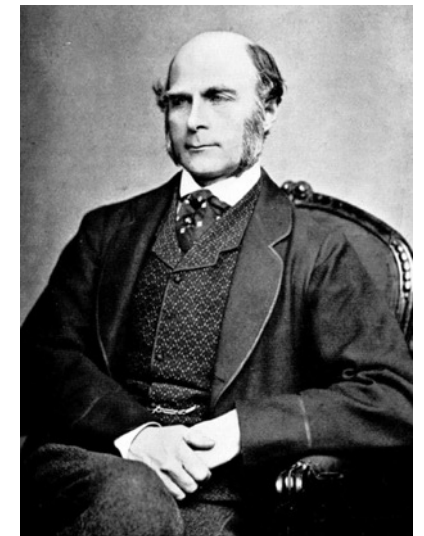
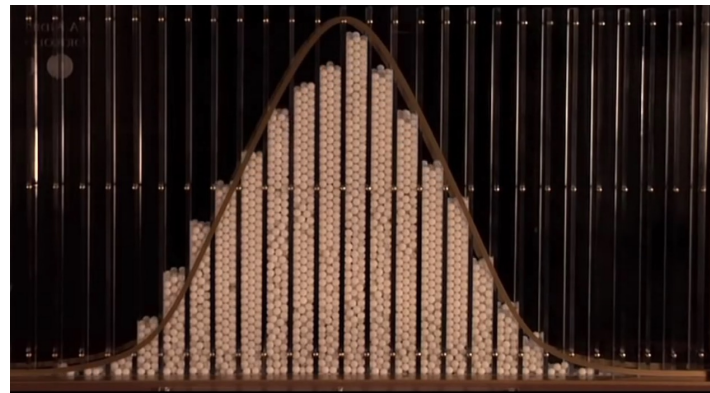
CLT explains a lot

$$\sum_{i=1}^n X_i \sim \mathcal{N}(n\mu, n\sigma^2) \quad \text{As } n \rightarrow \infty$$

The sum of n **i.i.d.** random variables is normally distributed with mean $n\mu$ and variance $n\sigma^2$.



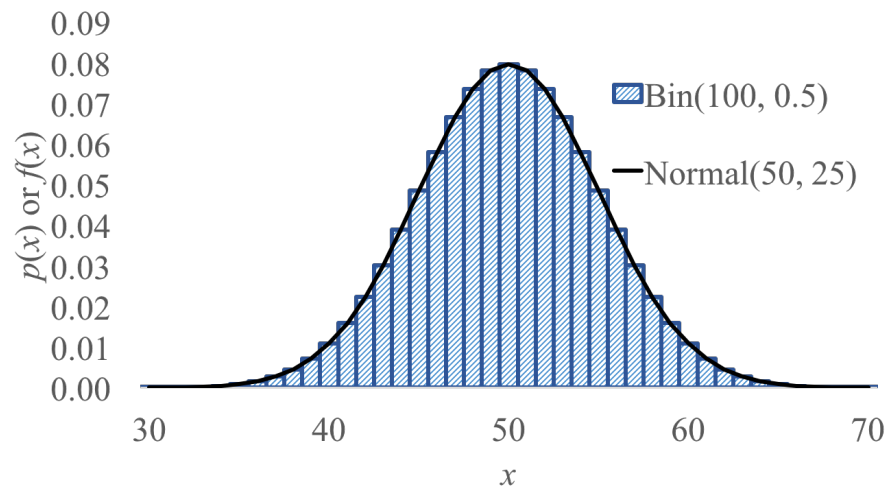
Galton Board, by Sir Francis Galton
(1822-1911)



CLT explains a lot

$$\sum_{i=1}^n X_i \sim \mathcal{N}(n\mu, n\sigma^2) \quad \text{As } n \rightarrow \infty$$

The sum of n **i.i.d.** random variables is normally distributed with mean $n\mu$ and variance $n\sigma^2$.



New Explanation:

Let $X_i \sim \text{Ber}(p)$ for $i = 1, \dots, n$, where X_i are i.i.d.
 $E[X_i] = p$, $\text{Var}(X_i) = p(1 - p)$

$$X = \sum_{i=1}^n X_i \quad (X \sim \text{Bin}(n, p))$$

$$X \sim \mathcal{N}(n\mu, n\sigma^2) \quad (\text{CLT, as } n \rightarrow \infty)$$

$$X \sim \mathcal{N}(np, np(1 - p)) \quad (\text{substitute mean, variance of Bernoulli})$$

Normal approximation of Binomial
Sum of i.i.d. Bernoulli RVs \approx Normal

Example CLT problem

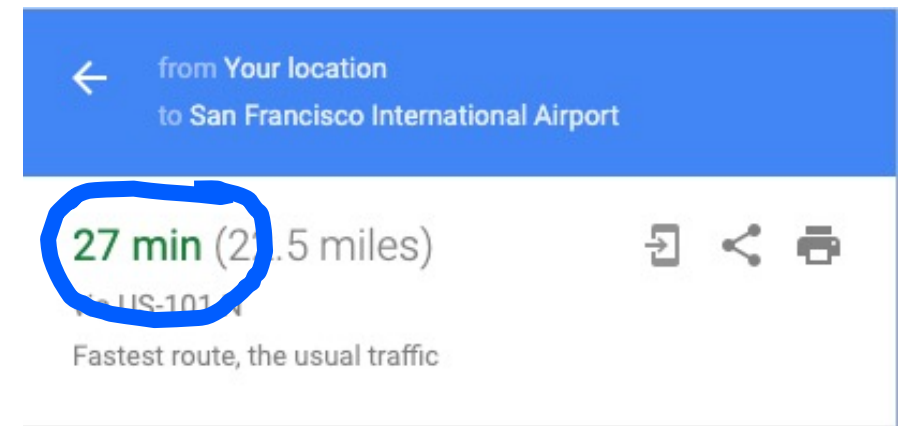
You hit 10 traffic lights on your way to work. You don't know the full distribution of the wait time, but for each you observe the average wait time is 45 seconds and the standard deviation is 5 seconds. You will be on time if your total wait time is less than 8 mins. What is the probability that you are on time? Assume the wait times are IID.

Answer: Let T be the total wait time. It is the sum of the 10 IID wait times. By the CLT

$$T \sim \mathcal{N}(n\mu, n\sigma^2)$$

$$T \sim \mathcal{N}(450, 250)$$

$$P(T \leq 480) = \Phi\left(\frac{480 - 450}{15.8}\right) \approx 0.97$$



Proof of CLT

$$\sum_{i=1}^n X_i \sim \mathcal{N}(n\mu, n\sigma^2) \quad \text{As } n \rightarrow \infty$$

The sum of n **i.i.d.** random variables is normally distributed with mean $n\mu$ and variance $n\sigma^2$.

Proof:

- The Fourier Transform of a PDF is called a **characteristic function**.
- Take the characteristic function of the probability mass of the sample, normalized
- Show that this approaches an exponential function in the limit as $n \rightarrow \infty$: $f(x) = e^{-\frac{x^2}{2}}$
- This function is in turn the characteristic function of a Normal Distribution

(this proof is beyond the scope of CS109)

For Details, See Video

CLT Proof Video

and $Y_n = \frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}}$, then:

$$\phi_{Y_n}(t) = \left(1 + \frac{it^2}{2n} + O\left(\frac{t^3}{n^{3/2}}\right) \right)^n$$

Lemma 1: for any z_j

$$\phi_{z_j}(t) = 1 - \frac{t^2}{2} + O(t^3)$$

Proof of lemma 1:

$$\begin{aligned} \phi_{z_j}(t) &= E[e^{itZ_j}] \leftarrow \text{by definition} \\ &= E\left[\sum_{k=0}^{\infty} \frac{(itZ_j)^k}{k!}\right] \leftarrow \text{Taylor expansion} \\ &= E\left[1 + itZ_j + \frac{(itZ_j)^2}{2} + O(t^3)\right] \leftarrow \text{algebra and approximating the last} \\ &= (\text{next column}) \end{aligned}$$

$E[1] + E[itZ_j] + E\left[\frac{(itZ_j)^2}{2}\right] + E[O(t^3)]$
 $= 1 + itE[Z_j] - \frac{t^2}{2}E[Z_j^2] + O(t^3)$ \leftarrow linearity of expectation
 $= 1 + it(0) - \frac{t^2}{2}(1) + O(t^3)$ \leftarrow plug in

Watch later Share

Recall Z_j

Watch on YouTube

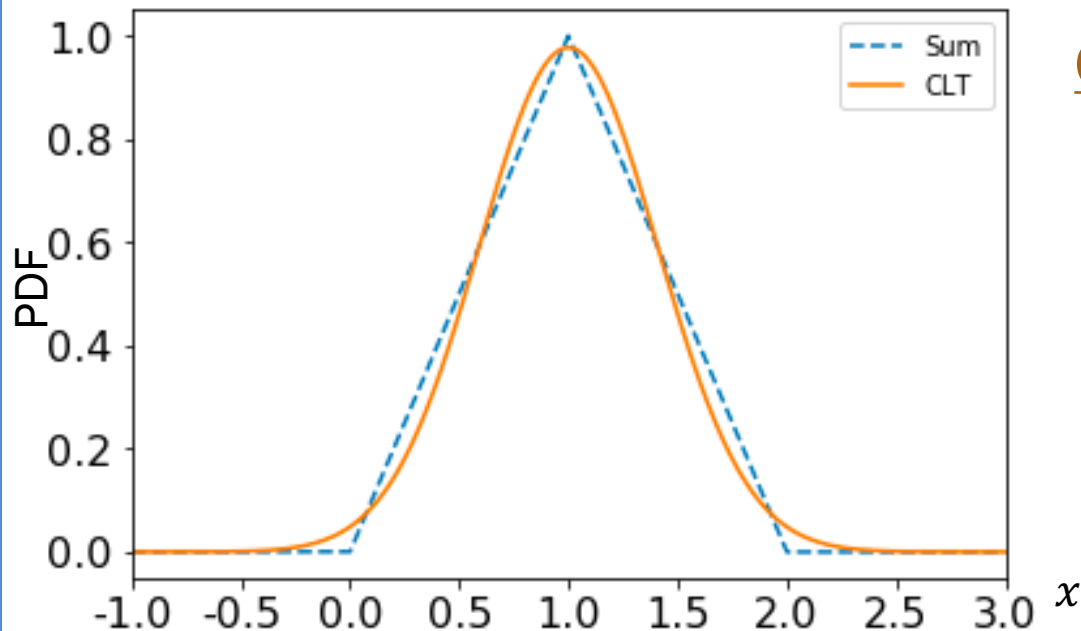
Testing CLT

Sum of n independent Uniform RVs

Let $X = \sum_{i=1}^n X_i$ be sum of i.i.d. RVs, where $X_i \sim \text{Uni}(0,1)$. $\mu = E[X_i] = 1/2$
 $\sigma^2 = \text{Var}(X_i) = 1/12$

For different n , how close is the CLT approximation of $P(X \leq n/3)$?

$n = 2$:



Exact

$$P(X \leq 2/3) \approx 0.2222$$

CLT approximation

$$X \approx Y \sim \mathcal{N}(n\mu, n\sigma^2) \implies Y \sim \mathcal{N}(1, 1/6)$$

$$P(X \leq 2/3) \approx P(Y \leq 2/3)$$

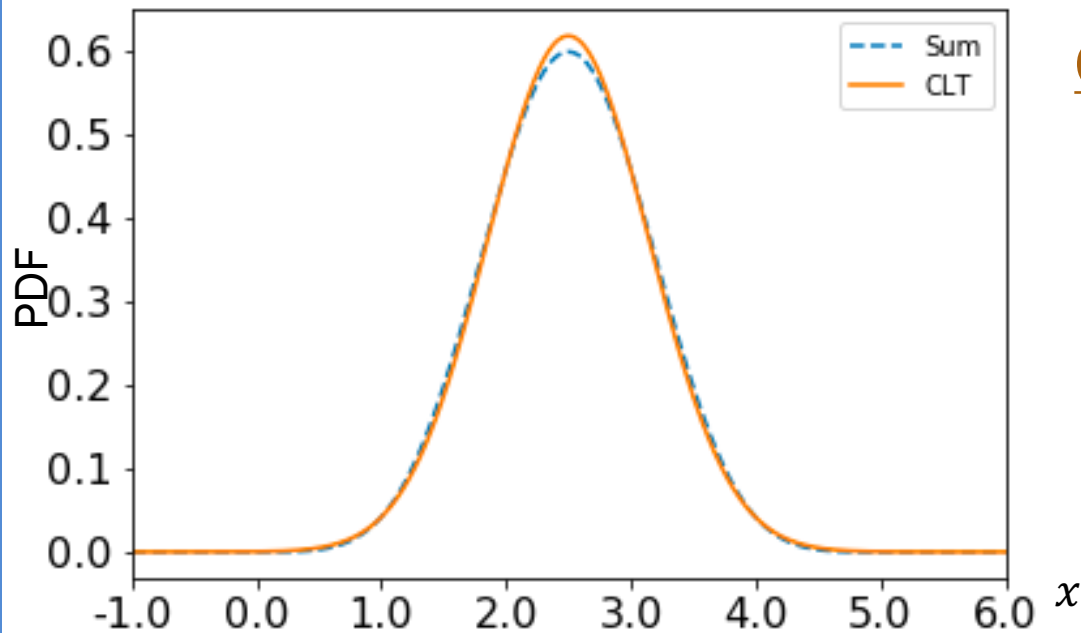
$$= \Phi\left(\frac{2/3 - 1}{\sqrt{1/6}}\right) \approx 0.2071$$

Sum of n independent Uniform RVs

Let $X = \sum_{i=1}^n X_i$ be sum of i.i.d. RVs, where $X_i \sim \text{Uni}(0,1)$. $\mu = E[X_i] = 1/2$
 $\sigma^2 = \text{Var}(X_i) = 1/12$

For different n , how close is the CLT approximation of $P(X \leq n/3)$?

$n = 5$:



Exact

$$P(X \leq 5/3) \approx 0.1017$$

CLT approximation

$$X \approx Y \sim \mathcal{N}(n\mu, n\sigma^2) \implies Y \sim \mathcal{N}(5/2, 5/12)$$

$$P(X \leq 5/3) \approx P(Y \leq 5/3)$$

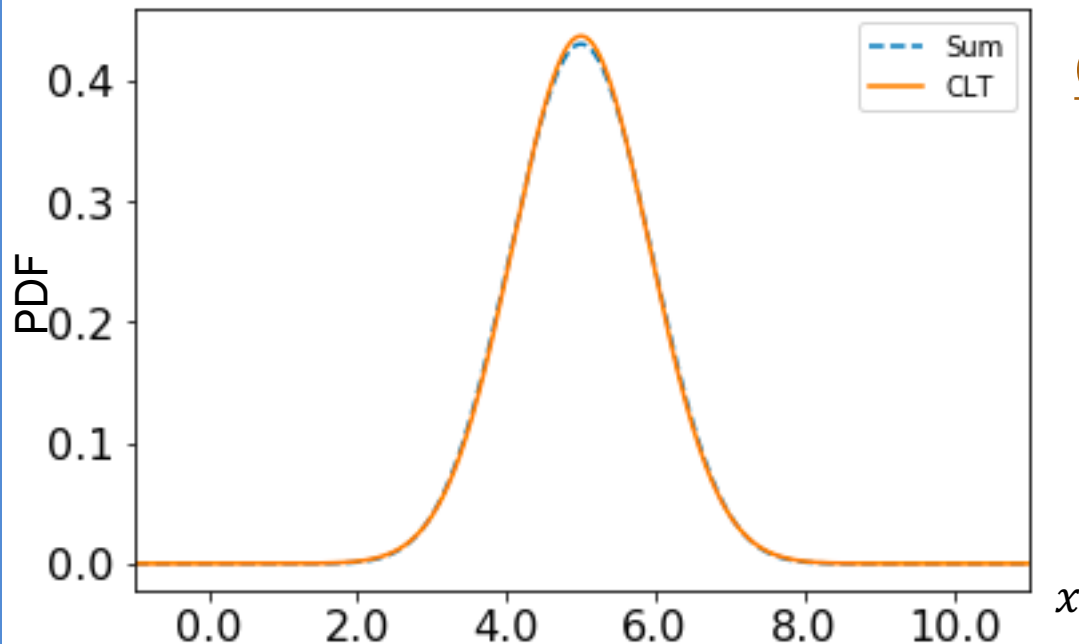
$$= \Phi\left(\frac{5/3 - 5/2}{\sqrt{5/12}}\right) \approx 0.0984$$

Sum of n independent Uniform RVs

Let $X = \sum_{i=1}^n X_i$ be sum of i.i.d. RVs, where $X_i \sim \text{Uni}(0,1)$. $\mu = E[X_i] = 1/2$
 $\sigma^2 = \text{Var}(X_i) = 1/12$

For different n , how close is the CLT approximation of $P(X \leq n/3)$?

$n = 10$:



Exact

$$P(X \leq 10/3) \approx 0.0337$$

CLT approximation

$$X \approx Y \sim \mathcal{N}(n\mu, n\sigma^2) \implies Y \sim \mathcal{N}(5, 5/6)$$

$$P(X \leq 10/3) \approx P(Y \leq 10/3)$$

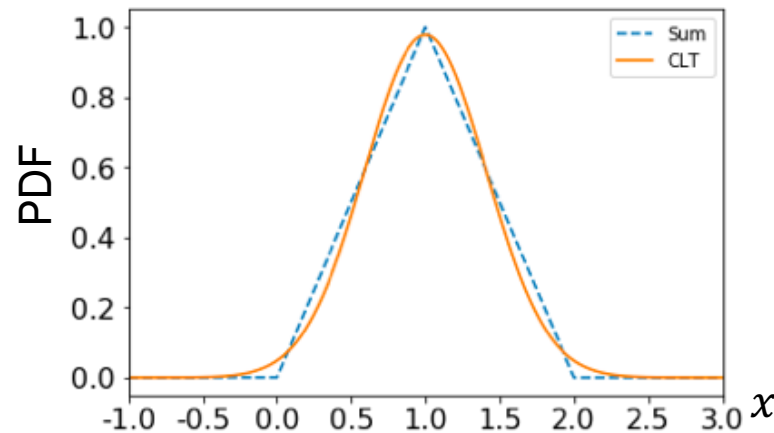
$$= \Phi\left(\frac{10/3 - 5}{\sqrt{5/6}}\right) \approx 0.0339$$

Sum of n independent Uniform RVs

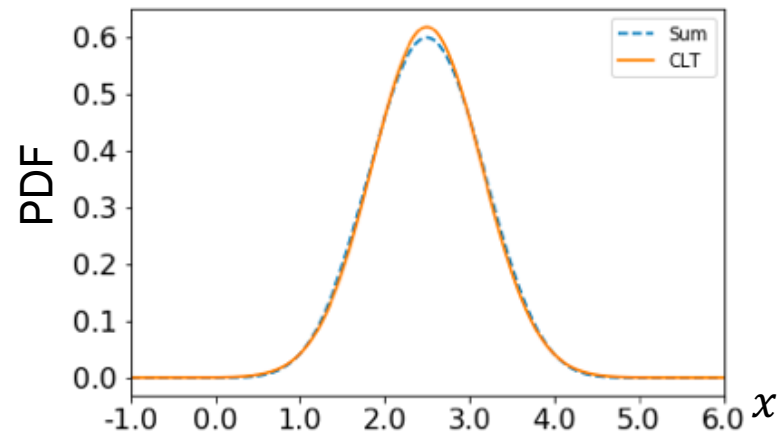
Let $X = \sum_{i=1}^n X_i$ be sum of i.i.d. RVs, where $X_i \sim \text{Uni}(0,1)$. $\mu = E[X_i] = 1/2$
 $\sigma^2 = \text{Var}(X_i) = 1/12$

For different n , how close is the CLT approximation of $P(X \leq n/3)$?

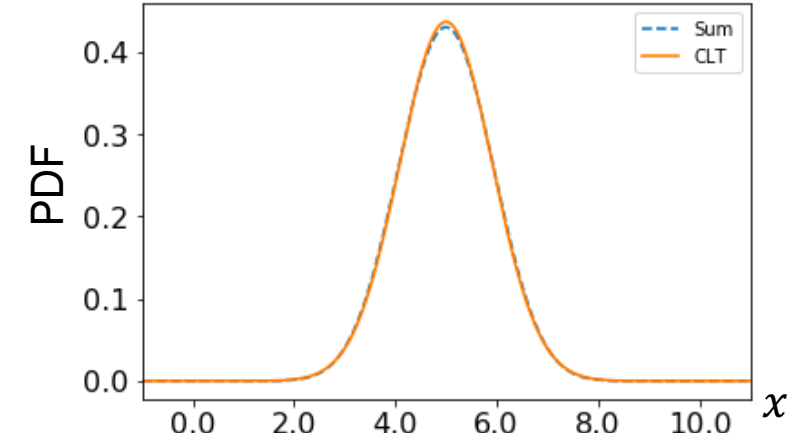
$n = 2$:



$n = 5$:



$n = 10$:



Most books will tell you that CLT holds if $n \geq 30$, but it can hold for smaller n depending on the distribution of your i.i.d. X_i 's.

The sum of independent, identically distributed variables:

$$Y = \sum_{i=0}^n X_i$$



Is normally distributed:

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

where $\mu = E[X_i]$

$$\sigma^2 = \text{Var}(X_i)$$

What about other functions?

Sum of iid? **Normal**

Average of iid?

Max of iid?

Average of IID Variables?

Let X_i be i.i.d. variables. There are n . Let \bar{X} be the average

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

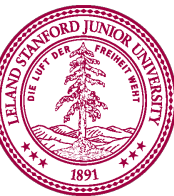
Gaussian by CLT

$$N(n\mu, n\sigma^2)$$



By the Central Limit Theorem, the mean of IID variables are distributed normally. As $n \rightarrow \infty$

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



Average of IID Variables Demo



Demo

http://onlinestatbook.com/stat_sim/sampling_dist/

What about other functions?

Sum of iid? Normal

Average of iid? Normal

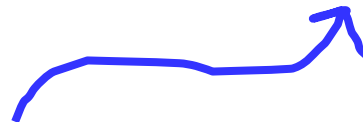
Max of iid?

What about other functions?

Sum of iid? Normal

Average of iid? Normal

Max of iid? Gumbel



See Fisher Trippett Gnedenko Theorem

Once Upon a Time...

Abraham De Moivre

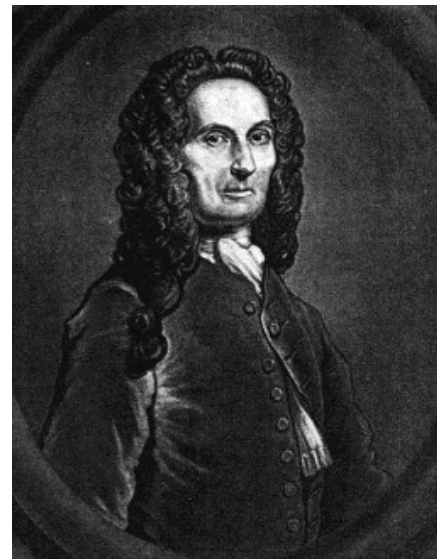
THE
DOCTRINE
OF
CHANCES:

OR,
A Method of Calculating the Probability
of Events in Play.



By *A. De Moivre*. F. R. S.

L O N D O N:
Printed by *W. Pearson*, for the Author. MDCCLXVIII.



1733

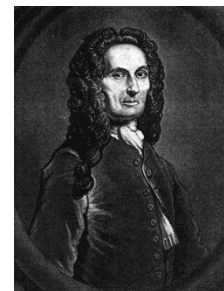
Piech, CS109, Stanford University



Once Upon a Time...

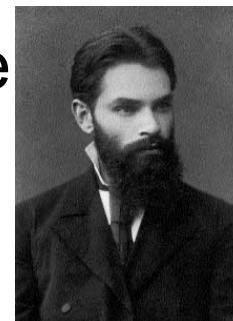
- History of the Central Limit Theorem

- 1733: CLT for $X \sim \text{Ber}(1/2)$ postulated by Abraham de Moivre



- 1823: Pierre-Simon Laplace extends de Moivre's work to approximating $\text{Bin}(n, p)$ with Normal

- 1901: Aleksandr Lyapunov provides precise definition and rigorous proof of CLT



- 2016: Beyonce releases Lemonade

- It was her 6th album, bringing her total number of songs to 214

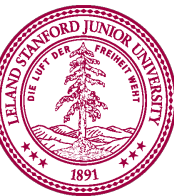
- Mean quality of subsamples of songs is Normally distributed (thanks to the Central Limit Theorem)



Estimating Clock Running Time

- Have new algorithm to test for running time
 - Mean (clock) running time: $\mu = t$ sec.
 - Variance of running time: $\sigma^2 = 4$ sec².
 - Run algorithm repeatedly (I.I.D. trials), measure time
 - How many trials s.t. estimated time = $t \pm 0.5$ with 95% certainty?
 - X_i = running time of i -th run (for $1 \leq i \leq n$), \bar{X} is the mean
-

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \sim N\left(t, \frac{4}{n}\right)$$

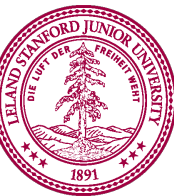


$$0.95 = P(-0.5 < \bar{X} - t < 0.5) \quad \bar{X} - t \sim N\left(0, \frac{4}{n}\right)$$

$$0.95 = F_{\bar{X}-t}(0.5) - F_{\bar{X}-t}(-0.5)$$

$$= \Phi\left(\frac{0.5 - 0}{\sqrt{4/n}}\right) - \Phi\left(\frac{-0.5 - 0}{\sqrt{4/n}}\right)$$

$$= 2\phi\left(\frac{\sqrt{n}}{4}\right) - 1$$



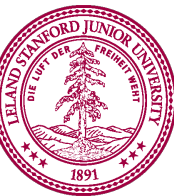
$$0.95 = 2\phi\left(\frac{\sqrt{n}}{4}\right) - 1$$

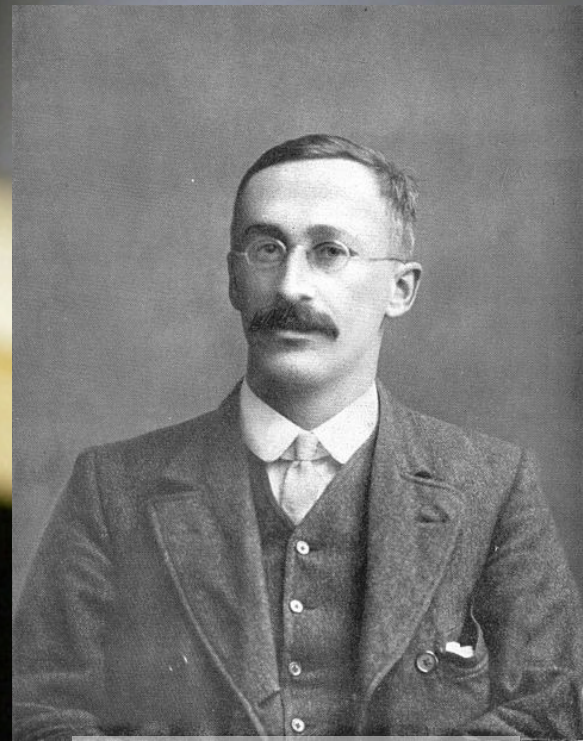
$$0.975 = \phi\left(\frac{\sqrt{n}}{4}\right)$$

$$\phi^{-1}(0.975) = \frac{\sqrt{n}}{4}$$

$$1.96 = \frac{\sqrt{n}}{4}$$

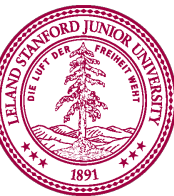
$$n = 61.4$$





William Sealy Gosset
(aka Student)

It's play time!



Sum of Dice

- You will roll 10 6-sided dice (X_1, X_2, \dots, X_{10})
 - $X =$ total value of all 10 dice $= X_1 + X_2 + \dots + X_{10}$
 - Win if: $X \leq 25$ or $X \geq 45$
 - Roll!
- And now the truth (according to the CLT)...



Sum of Dice

- You will roll 10 6-sided dice (X_1, X_2, \dots, X_{10})
 - X = total value of all 10 dice = $X_1 + X_2 + \dots + X_{10}$
 - Win if: $X \leq 25$ or $X \geq 45$

-
- Recall CLT: $X = \sum_{i=1}^n X_i \rightarrow N(n\mu, n\sigma^2)$ As $n \rightarrow \infty$

- Determine $P(X \leq 25 \text{ or } X \geq 45)$ using CLT:

$$\mu = E[X_i] = 3.5 \qquad \sigma^2 = \text{Var}(X_i) = \frac{35}{12} \qquad X \approx N(35, 29.2)$$

$$1 - P(25.5 < X < 44.5) = 1 - P\left(\frac{25.5 - 35}{\sqrt{29.2}} < Z < \frac{44.5 - 35}{\sqrt{29.2}}\right)$$

$$\approx 1 - (2\Phi(1.76) - 1) \approx 2(1 - 0.9608) = 0.0784$$

Sampling Statistics

Three Types of Sampling

1 Known distribution,
Parametric
Eg Sample from a normal

```
>>> import numpy as np
>>> np.random.normal(10, 1)
10.070856788485646
>>> np.random.normal(10, 1)
9.95160925474276
```

2 Known distribution,
Non-Parametric
Eg Sample from a histogram

```
>>> np.random.choice([0,1,2], p=[0.8,0.1,.1])
0
>>> np.random.choice([0,1,2], p=[0.8,0.1,.1])
2
```

3 Unknown,
Non-Parametric
Eg Get data from a trial

```
>>> test_algorithm_runtime()
12.1232394737634875
>>> test_algorithm_runtime()
18.2098234987956728
```

We know how to sample from a
parametric distribution

Now let us think about sampling from a population (unknown distribution)

Allow us to think about probabilities of
scientific claims!

Motivating example

You want to know the true mean and variance of happiness in Bhutan.

- But you can't ask everyone.
- You poll 200 random people.
- Your data looks like this:

Happiness = {72, 85, 79, 91, 68, ..., 71}

- The mean of all these numbers is 83.

Is this the **true mean happiness** of Bhutanese people?



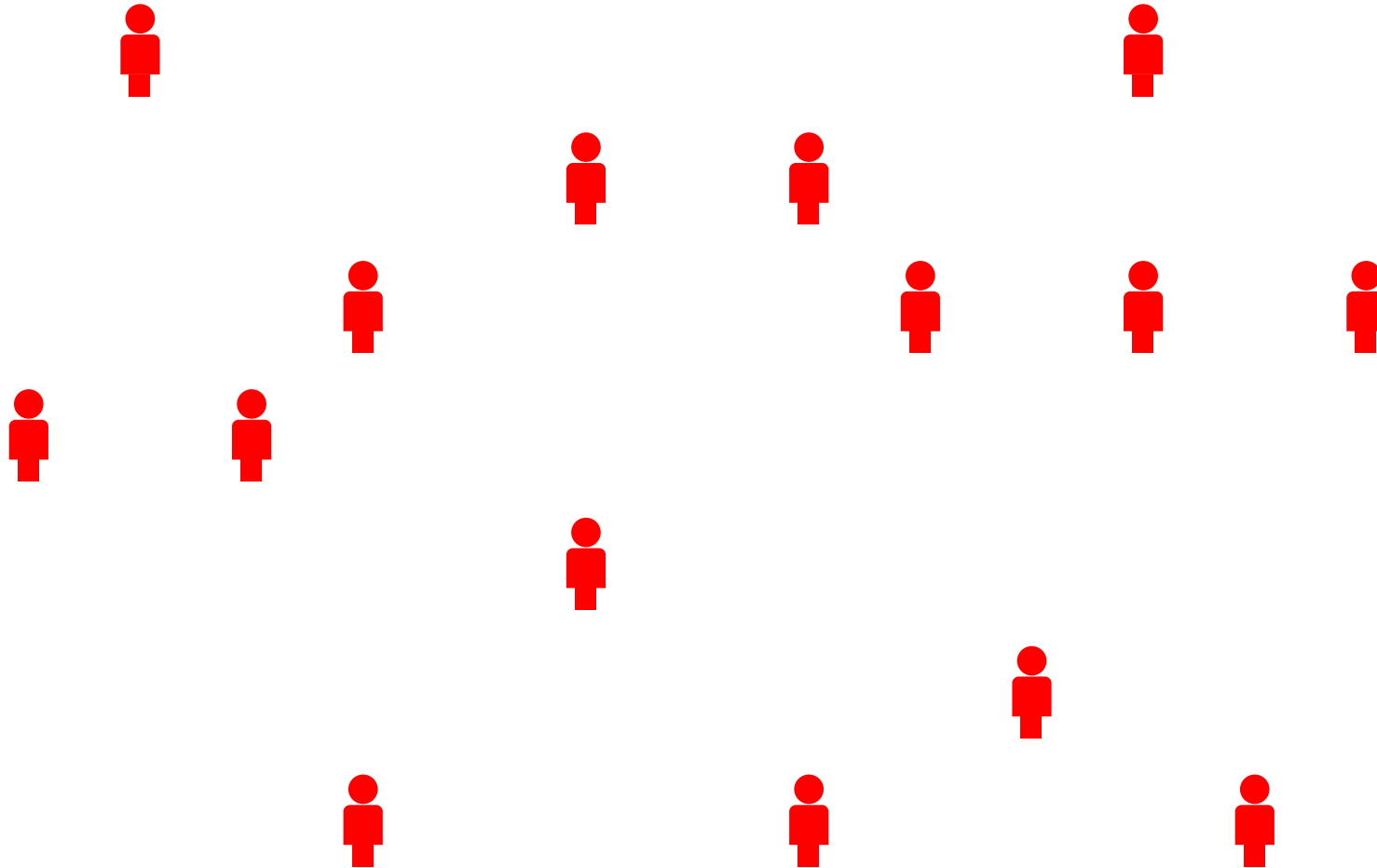
Population



Sample



Sample



Collect one (or more) numbers from each person



Sample





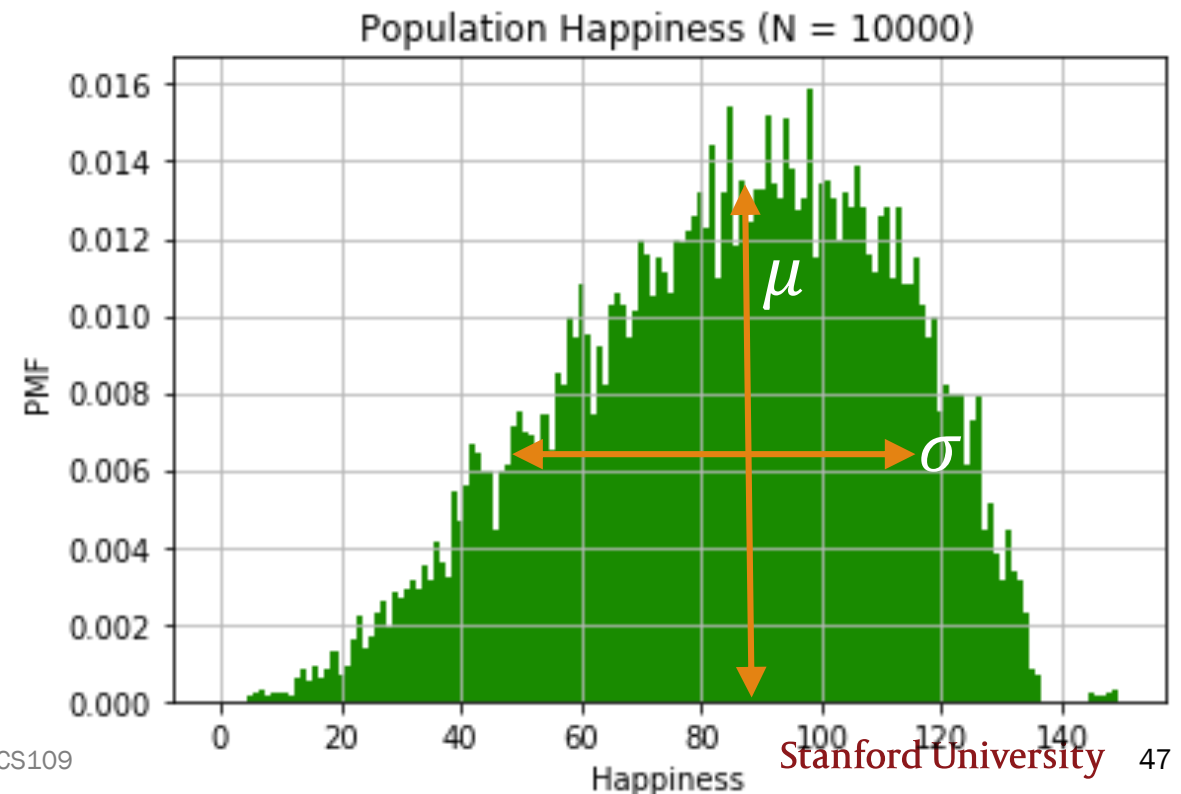
DATA SCIENCE

A sample, mathematically

Consider n random variables X_1, X_2, \dots, X_n .

The sequence X_1, X_2, \dots, X_n is a **sample** from distribution F if:

- X_i are all independent and identically distributed (i.i.d.)
- X_i all have same distribution function F (the **underlying distribution**), where $E[X_i] = \mu$, $\text{Var}(X_i) = \sigma^2$



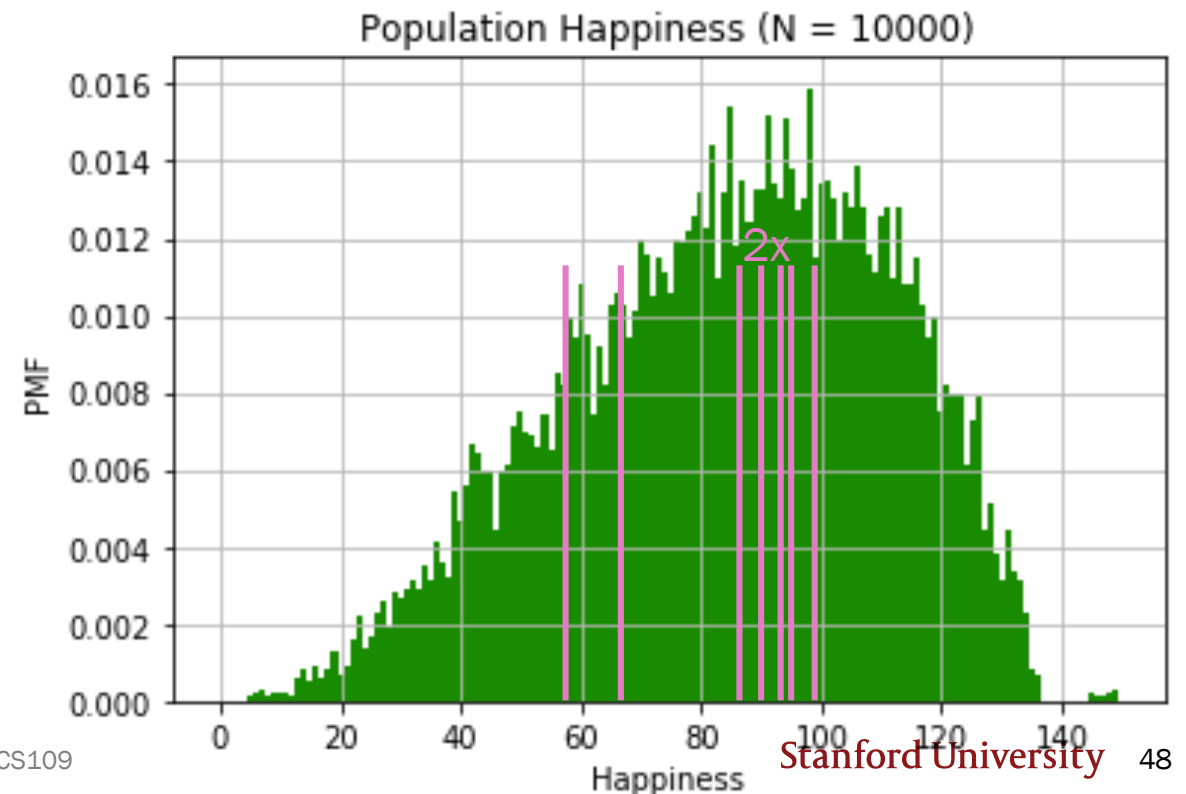
A sample, mathematically

A sample of **sample size 8**:

$(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8)$

A **realization** of a sample of size 8:

$(59, 87, 94, 99, 87, 78, 69, 91)$



A single sample



A happy
person

If we had a distribution F of our entire population, we could compute exact statistics about about happiness.

But we only have 200 people (a sample).

Today: If we only have a single sample,

- How do we report *estimated* statistics?
- How do we report estimated error of these estimates?

Tomorrow:

How do we perform hypothesis testing?

Estimating Core Statistics (Mean + Var)

A single sample

If we had a distribution F of our entire population, we could compute exact statistics about about happiness.



A happy person

But we only have 200 people (a sample).

So these population statistics are unknown:

- μ , the **population mean**
- σ^2 , the **population variance**

A single sample

If we had a distribution F of our entire population, we could compute exact statistics about about happiness.



A happy
person

But we only have 200 people (a sample).


- From these 200 people, what is our best estimate of **population mean** and **population variance**?
- How do we define best estimate?

Estimating the Mean

Consider n random variables X_1, X_2, \dots, X_n

- X_i are all independently and identically distributed (I.I.D.)
- Have same distribution function F and $E[X_i] = \mu$
- We call sequence of X_i a **sample** from distribution F
- *How would you estimate the population mean??*

$$\text{Estimate} = \frac{1}{n} \sum_{i=0}^n X_i$$

Sample Mean: This is a fancy way of saying "your estimate of the mean" 

$$\bar{X} = \frac{1}{n} \sum_{i=0}^n X_i$$

Say hello to our new friend: Sample mean

\bar{X}

You know I am mean because I have a really *scary* looking hat...

Celebrity look alike:



Al Capone

Is that estimate any good?

$$\bar{X} = \frac{1}{n} \sum_{i=0}^n X_i$$

Consider n random variables X_1, X_2, \dots, X_n

- Have same distribution function F and $E[X_i] = \mu$
- *Is our estimate of mean any good??*

$$\begin{aligned} E[\bar{X}] &= E\left[\sum_{i=1}^n \frac{X_i}{n}\right] = \frac{1}{n} E\left[\sum_{i=1}^n X_i\right] \\ &= \frac{1}{n} \sum_{i=1}^n E[X_i] = \frac{1}{n} \sum_{i=1}^n \mu = \frac{1}{n} n\mu = \mu \end{aligned}$$

Estimating the population mean



1. What is our best estimate of μ , the **mean happiness** of Bhutanese people?

If we only have a sample, (X_1, X_2, \dots, X_n) :

The best estimate of μ is the **sample mean**:

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

\bar{X} is an unbiased estimator of the population mean μ . $E[\bar{X}] = \mu$

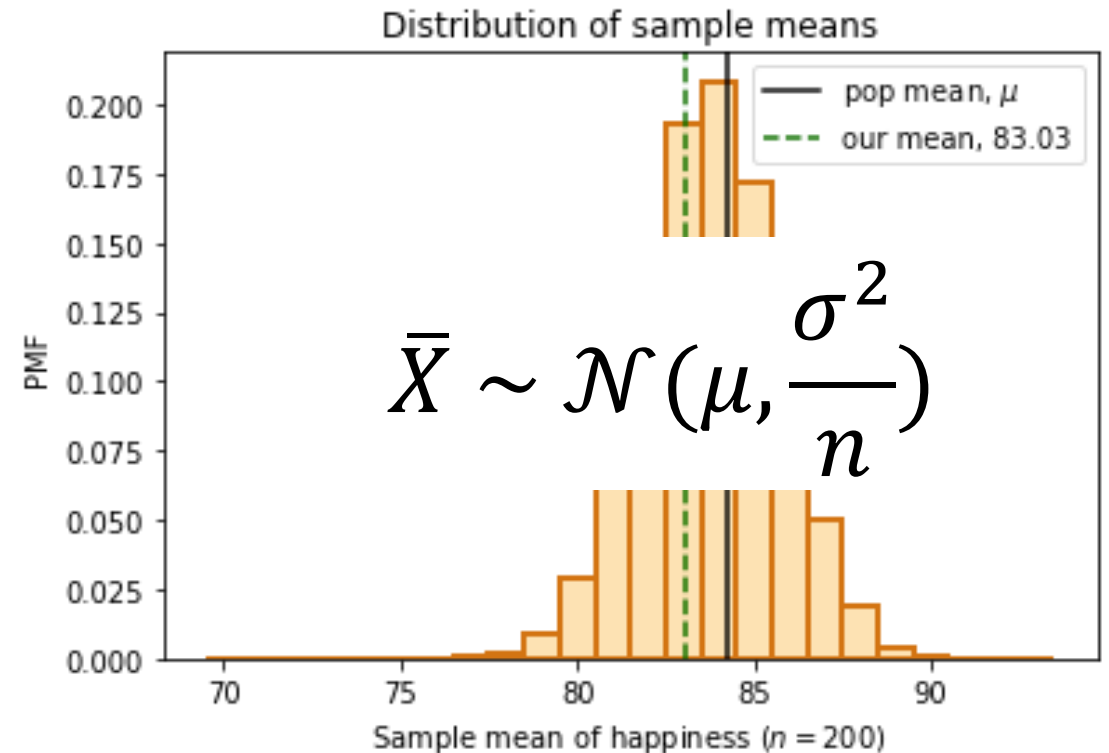
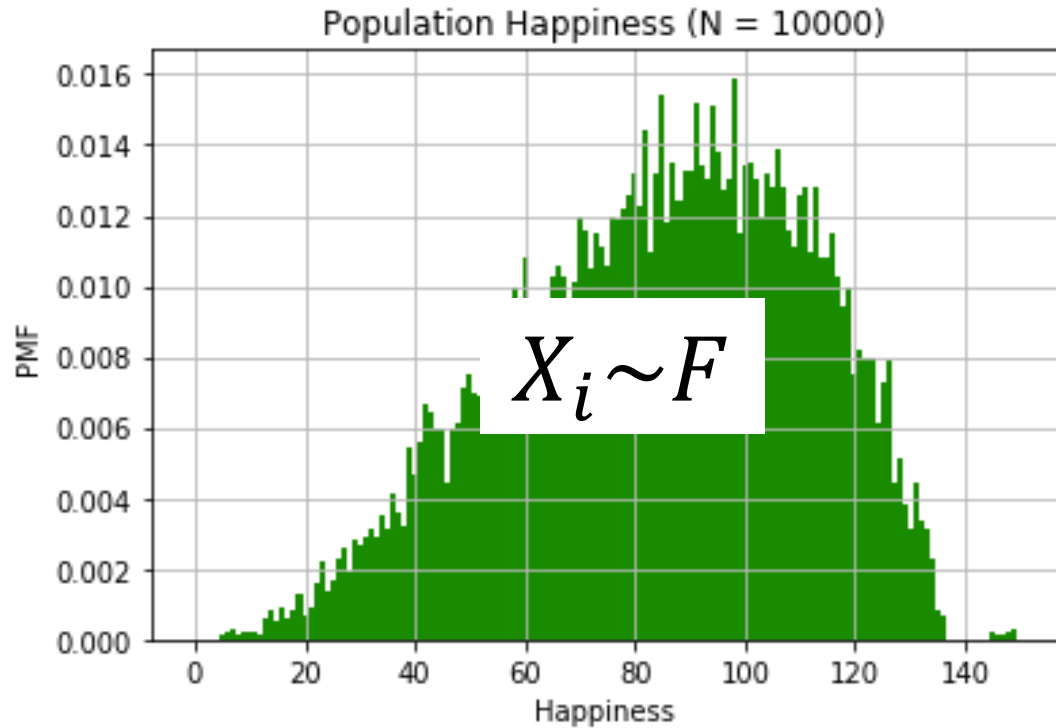
Intuition: By the CLT, $\bar{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$



If we could take *multiple* samples of size n :

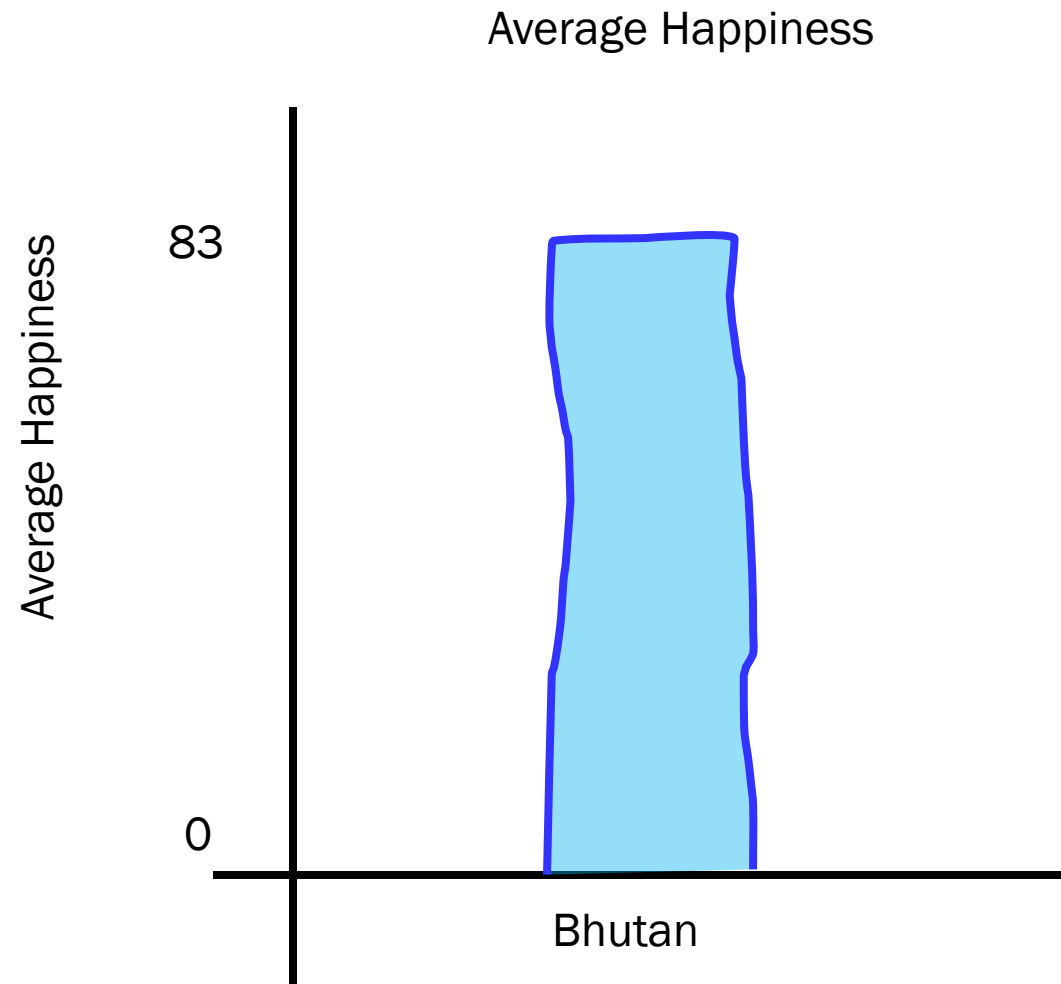
1. For each sample, compute sample mean
2. On average, we would get the population mean

Sample mean



Even if we can't report μ , we can report our sample mean 83.03, which is an unbiased estimate of μ .

Our Report to Bhutan Government





Sample Mean:

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

ith sample

Size of the sample

Estimating the population variance



2. What is σ^2 , the **variance of happiness** of Bhutanese people?

If we knew the entire population (x_1, x_2, \dots, x_N) :

population variance

$$\sigma^2 = E[(X - \mu)^2] = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2$$

population mean

If we only have a sample, (X_1, X_2, \dots, X_n) :

sample variance

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

sample mean

Intuition about the sample variance, S^2

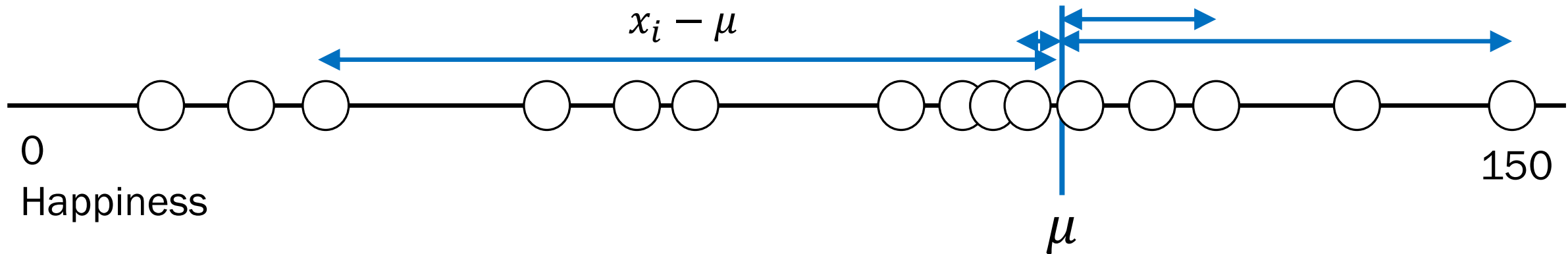
Actual, σ^2

population mean

population
variance

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2$$

$x_i - \mu$



Population size, N

Calculating population statistics exactly requires us knowing all N datapoints.

Intuition about the sample variance, S^2

Actual, σ^2

Estimate, S^2

population
variance

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2$$

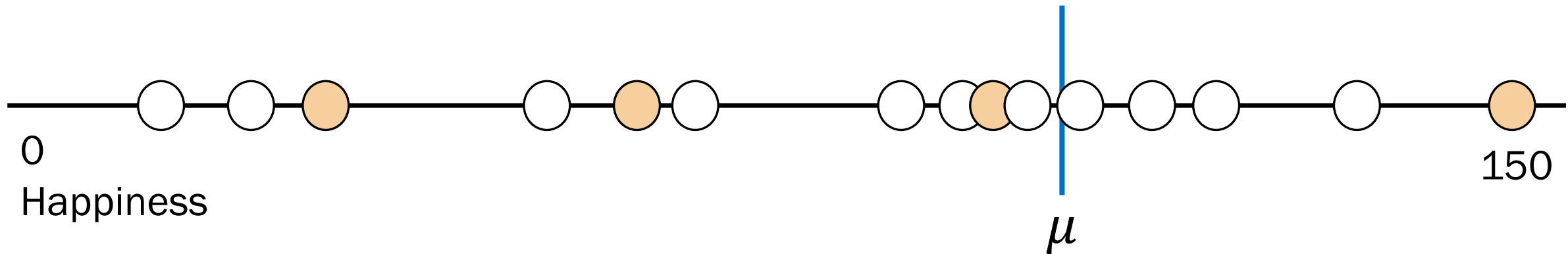
population mean



sample
variance

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

sample mean



Population size, N

Intuition about the sample variance, S^2

Actual, σ^2

Estimate, S^2

population
variance

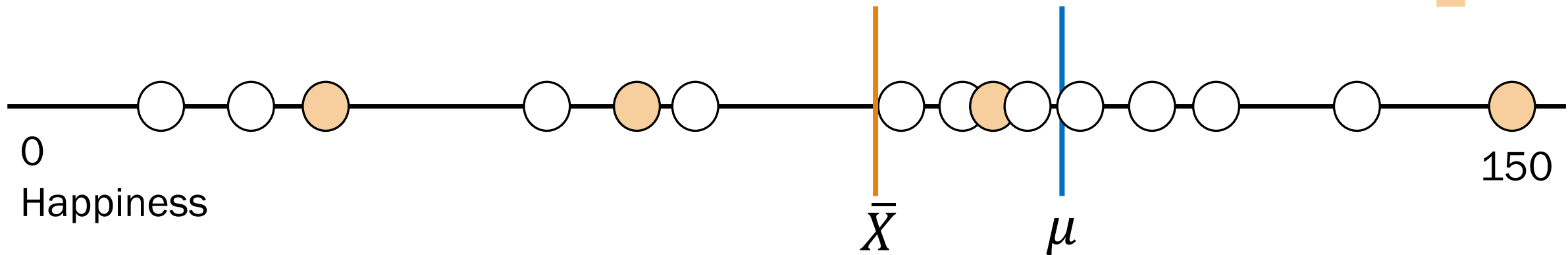
$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2$$

population mean

sample
variance

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

sample mean



Population size, N

Intuition about the sample variance, S^2

Actual, σ^2

Estimate, S^2

population
variance

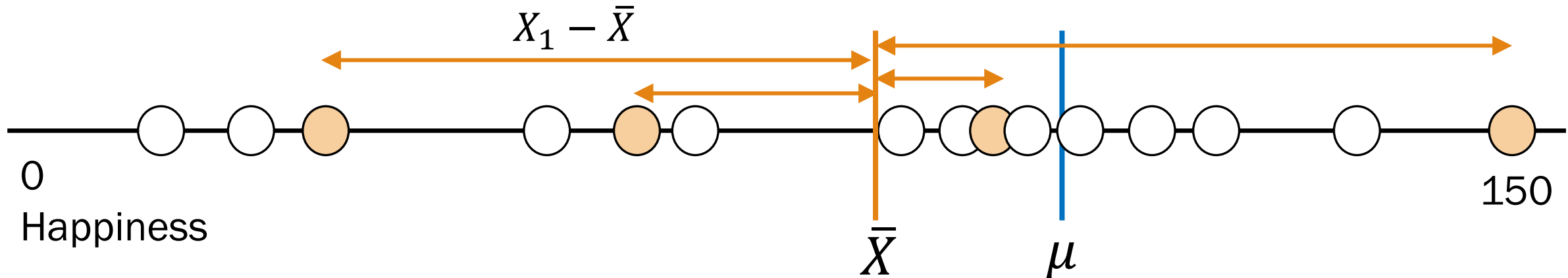
population mean

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2$$

sample
variance

sample mean

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



Population size, N

Sample variance is an estimate using an estimate, so it needs additional scaling.

Proof that S^2 is unbiased (just for reference)

$$E[S^2] = \sigma^2$$

$$E[S^2] = E\left[\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2\right] \Rightarrow (n-1)E[S^2] = E\left[\sum_{i=1}^n (X_i - \bar{X})^2\right]$$

$$(n-1)E[S^2] = E\left[\sum_{i=1}^n ((X_i - \mu) + (\mu - \bar{X}))^2\right] \quad (\text{introduce } \mu - \mu)$$

$$= E\left[\sum_{i=1}^n (X_i - \mu)^2 + \sum_{i=1}^n (\mu - \bar{X})^2 + 2 \sum_{i=1}^n (X_i - \mu)(\mu - \bar{X})\right]$$

$$= E\left[\sum_{i=1}^n (X_i - \mu)^2 + n(\mu - \bar{X})^2 - 2n(\mu - \bar{X})^2\right]$$

$$= E\left[\sum_{i=1}^n (X_i - \mu)^2 - n(\mu - \bar{X})^2\right] = \sum_{i=1}^n E[(X_i - \mu)^2] - nE[(\bar{X} - \mu)^2]$$

$$= n\sigma^2 - n\text{Var}(\bar{X}) = n\sigma^2 - n\frac{\sigma^2}{n} = n\sigma^2 - \sigma^2 = (n-1)\sigma^2 \quad \text{Therefore } E[S^2] = \sigma^2$$

Estimating the population variance



2. What is σ^2 , the **variance of happiness** of Bhutanese people?

If we only have a sample, (X_1, X_2, \dots, X_n) :

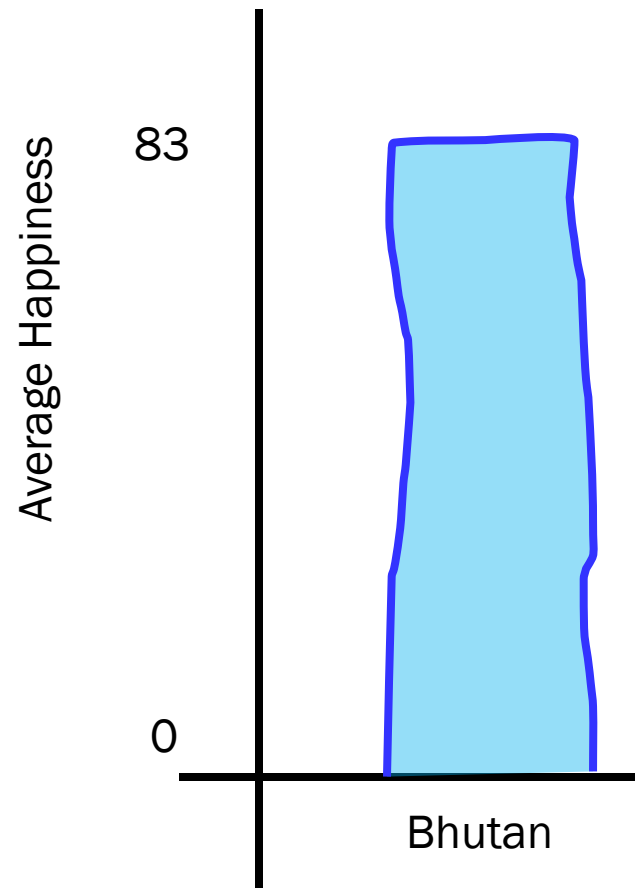
The best estimate of σ^2 is the **sample variance**:

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

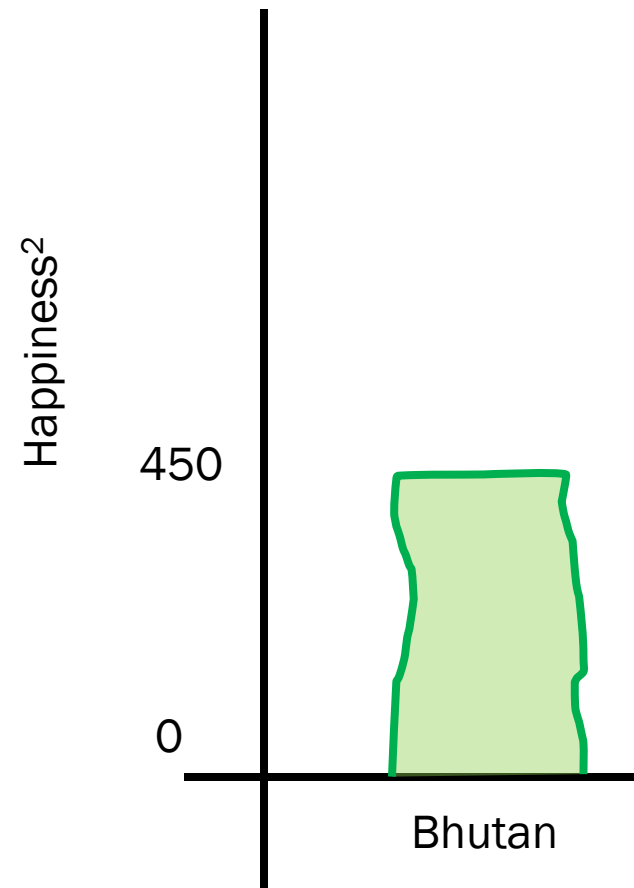
S^2 is an **unbiased estimator** of the population variance, σ^2 . $E[S^2] = \sigma^2$

Our Report to Bhutan Government

Average Happiness



Variance of Happiness





Sample Variance:

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

Sample mean

Makes it "unbiased"

No Error Bars ☹️

Quick check

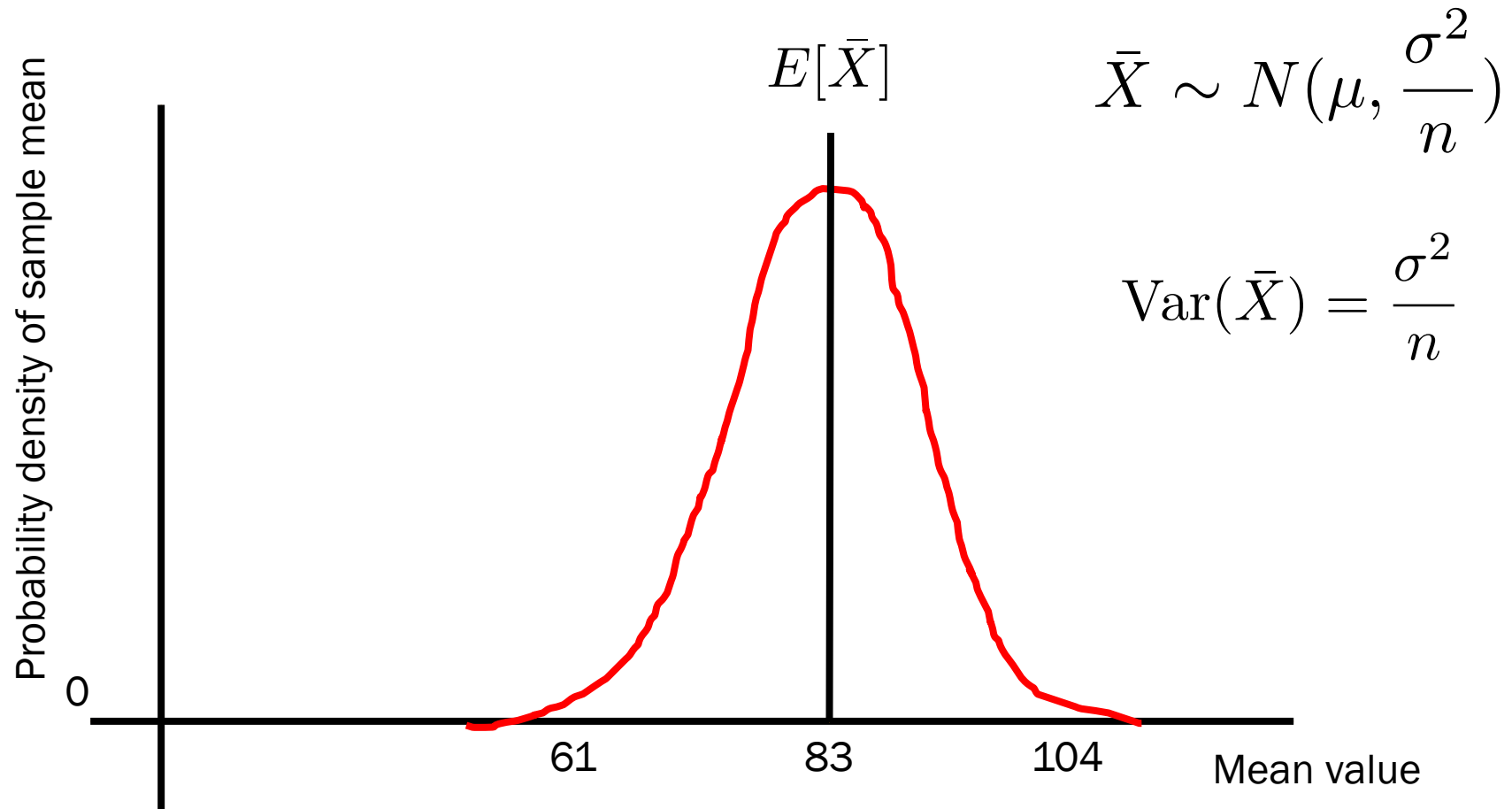
1. μ , the population mean
2. $(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8)$, a sample
3. σ^2 , the population variance
4. \bar{X} , the sample mean
5. $\bar{X} = 83$
6. $(X_1 = 59, X_2 = 87, X_3 = 94, X_4 = 99,$
 $X_5 = 87, X_6 = 78, X_7 = 69, X_8 = 91)$

- A. Random variable(s)
- B. Value
- C. Event



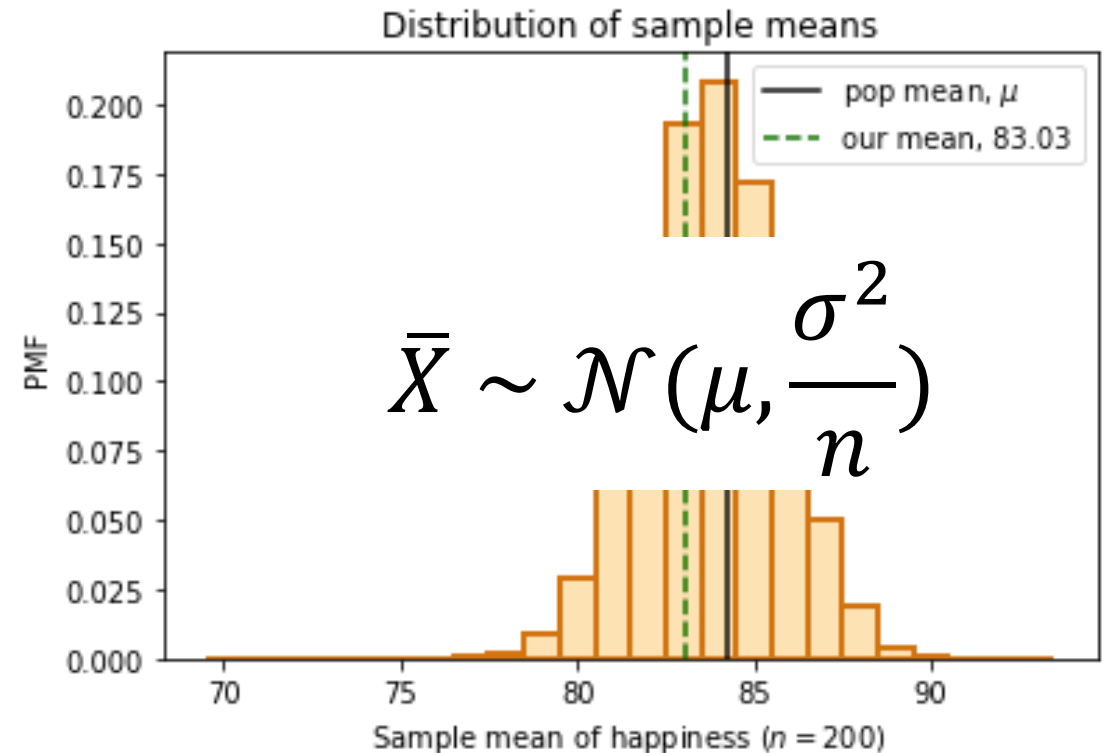
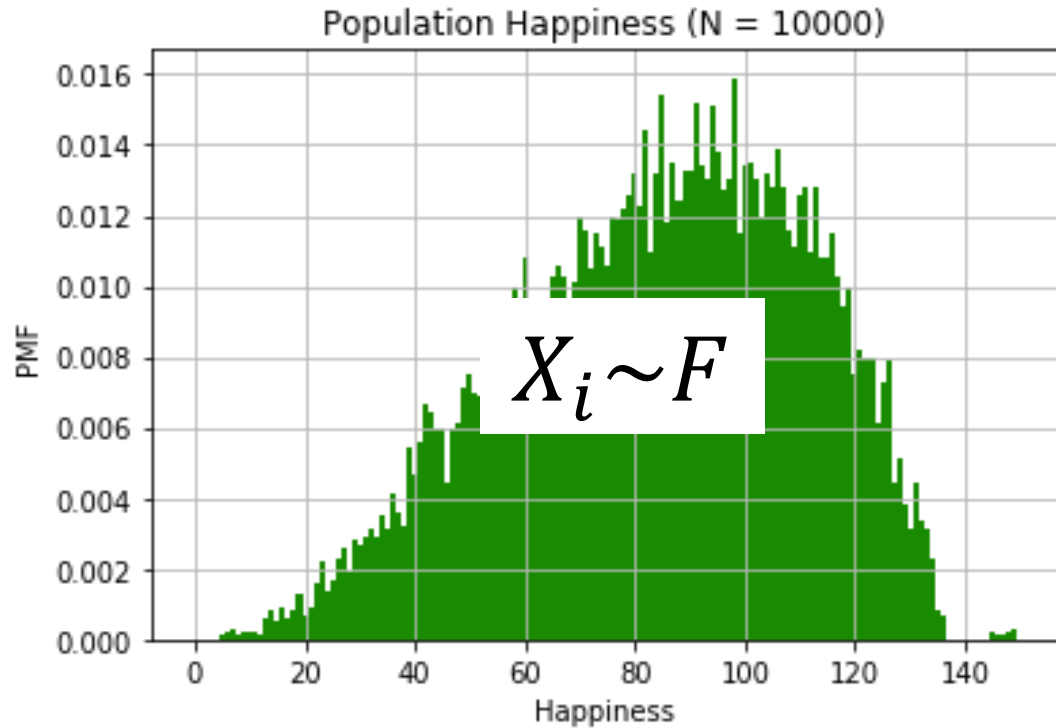
Insight: Sample Mean is an RV with known Var

By central limit theorem:



Standard error of the mean

Sample mean



- $\text{Var}(\bar{X})$ is a measure of how “close” \bar{X} is to μ .
- How do we estimate $\text{Var}(\bar{X})$?

Standard Error of the Mean

$$E[\bar{X}] = \mu$$

$$\text{Var}(\bar{X}) = \frac{\sigma^2}{n}$$

We want to estimate this

def The **standard error** of the mean is an estimate of the standard deviation of \bar{X} .

$$SE = \sqrt{\frac{S^2}{n}}$$

Intuition:

- S^2 is an unbiased estimate of σ^2
- S^2/n is an unbiased estimate of $\sigma^2/n = \text{Var}(\bar{X})$
- $\sqrt{S^2/n}$ can estimate $\sqrt{\text{Var}(\bar{X})}$

More info on bias of standard error: [wikipedia](#)

Standard Error of the Mean

$$\text{Var}(\bar{X}) = \text{Var}\left(\sum_{i=1}^n \frac{X_i}{n}\right) = \left(\frac{1}{n}\right)^2 \text{Var}\left(\sum_{i=1}^n X_i\right) = \frac{\sigma^2}{n}$$

$$\text{Var}(\bar{X}) = \frac{\sigma^2}{n}$$

$$= \frac{S^2}{n}$$

Since S^2 is an unbiased estimate

$$\text{Std}(\bar{X}) = \sqrt{\frac{S^2}{n}}$$

Change variance to standard deviation

$$= \sqrt{\frac{450}{200}}$$

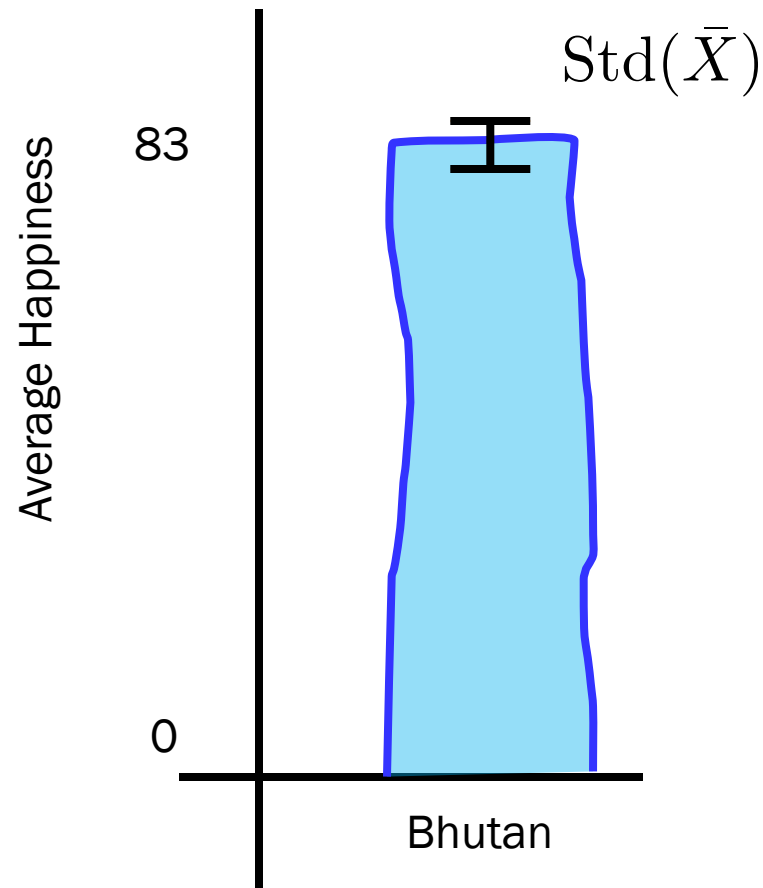
The numbers for our Bhutanesse poll

$$= 1.5$$

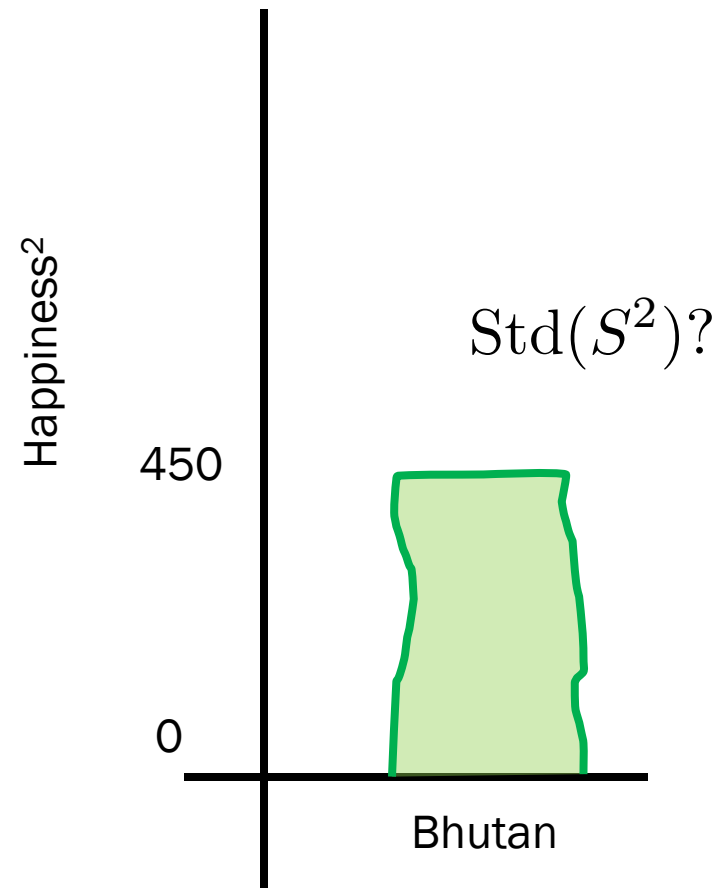
Bhutanesse standard error of the mean

Our Report to Bhutan Government

Average Happiness



Variance of Happiness



Claim: The average happiness of Bhutan is 83 ± 2

Bootstrapping