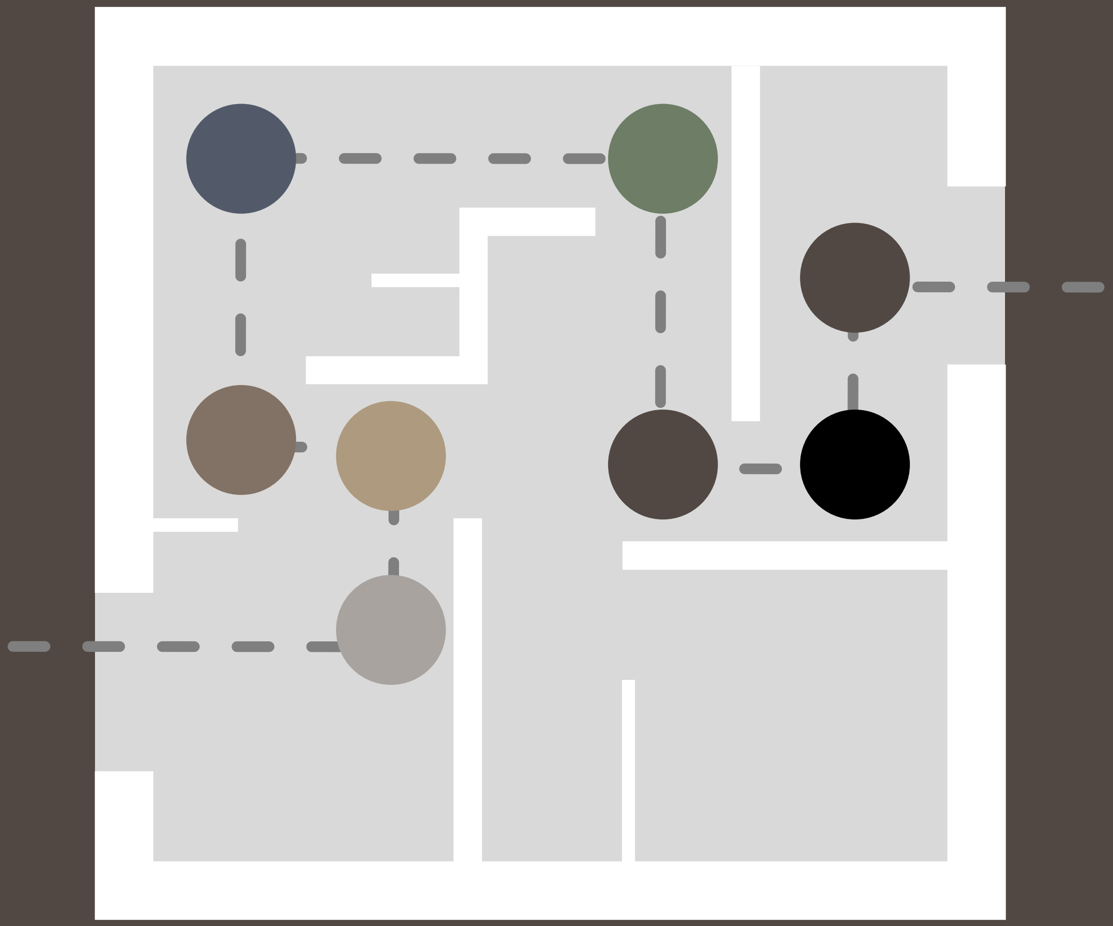
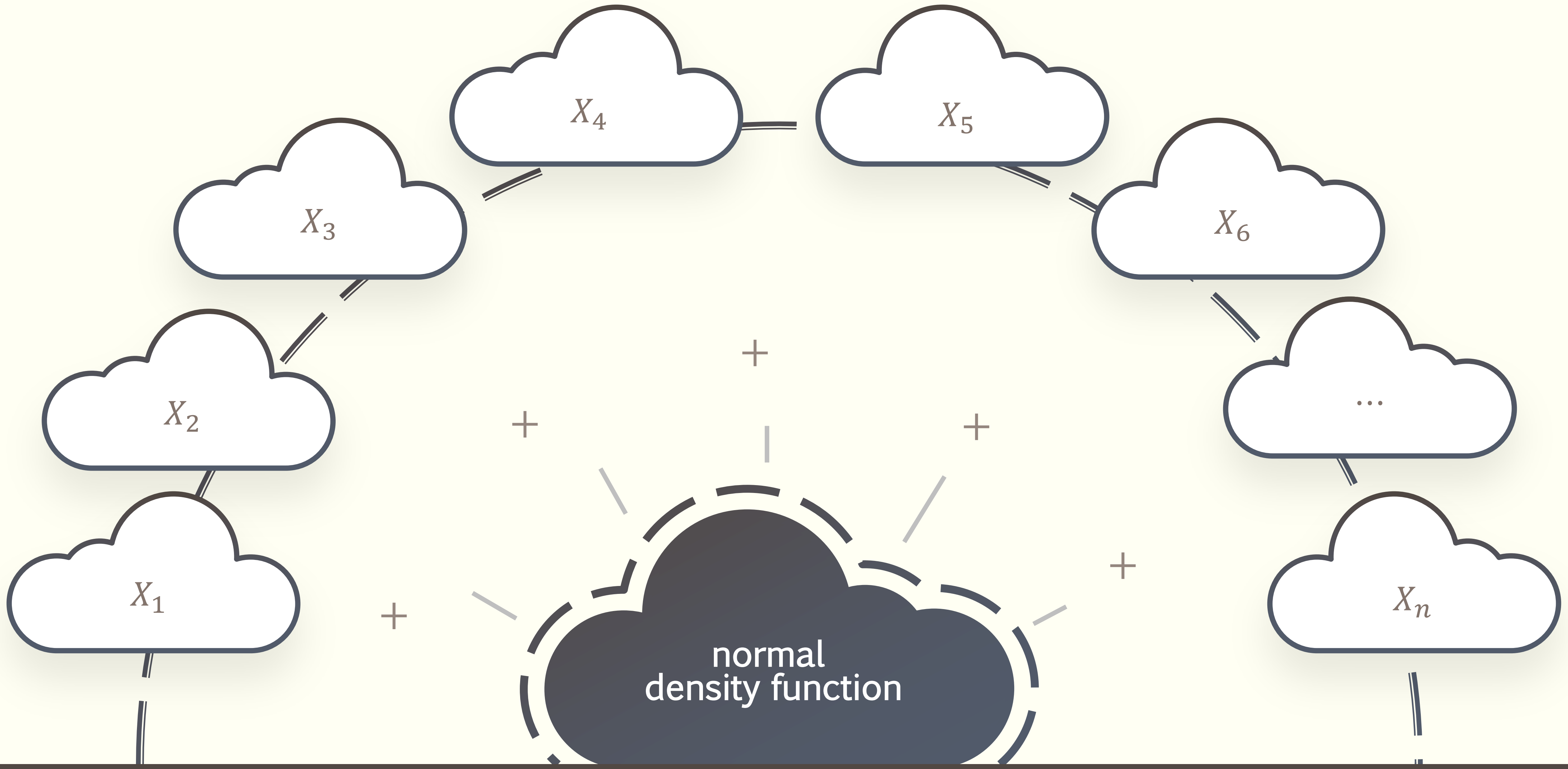


Central Limit Theorem

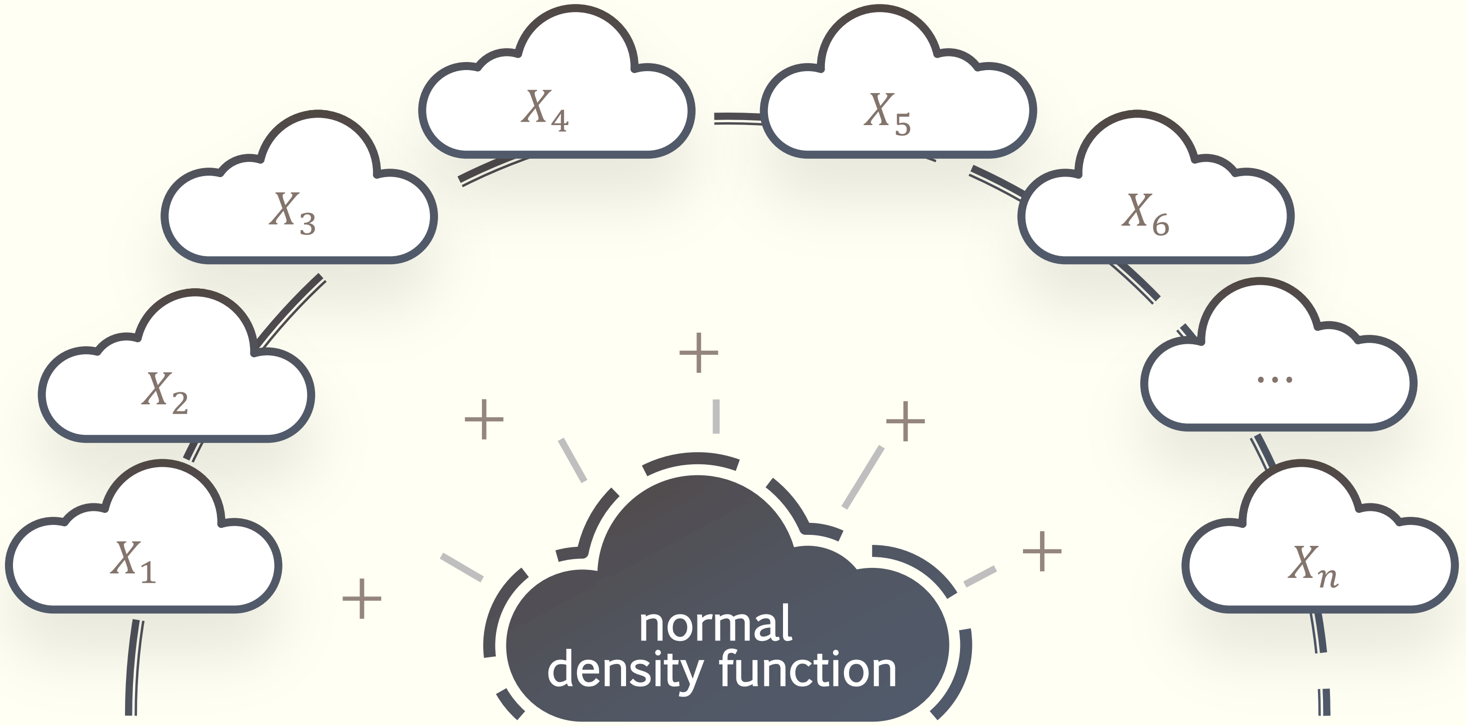
- number of measurements needed to estimate the mean
- convergence of the **estimated mean**
- a better approximation than the Chebyshev Inequality (LLN).

Xingru Chen
xingru.chen.gr@dartmouth.edu

Central Limit Theorem



Central Limit Theorem



If S_n is the sum of n mutually independent random variables, then the distribution function of S_n is well-approximated by a normal density function.

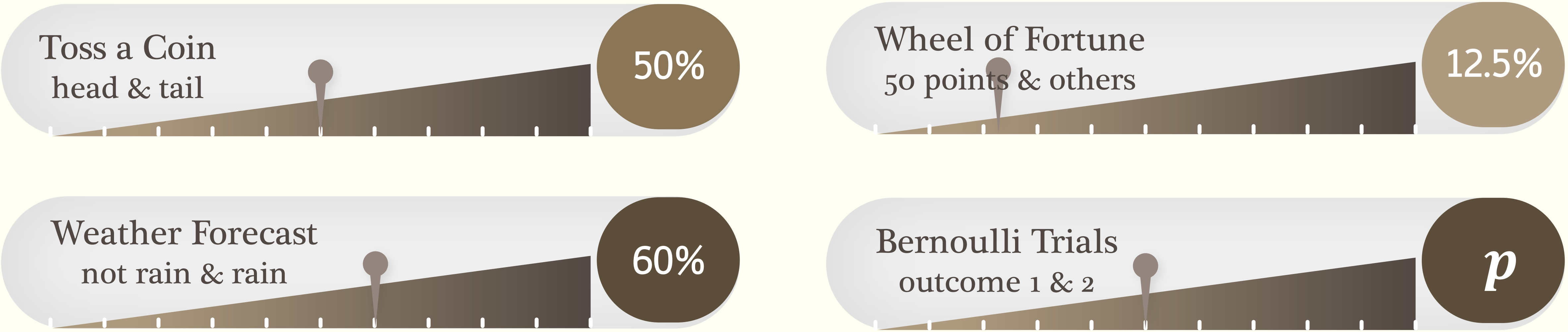
CENTRAL LIMIT THEOREM FOR BERNOULLI TRIALS

binomial distribution

Bernoulli Trials

A **Bernoulli trials process** is a sequence of n chance experiments such that

- Each experiment has **two possible outcomes**, which we may call success and failure.
- The probability p of success on each experiment is the same for each experiment, and this probability is not affected by any knowledge of previous outcomes. The probability q of failure is given by $q = 1 - p$.

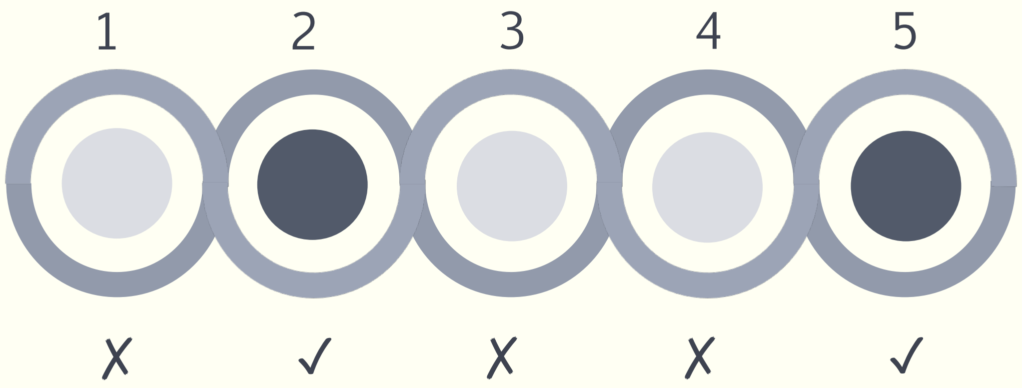


Binomial Distribution

- Let n be a positive integer and let p be a real number between 0 and 1.
- Let B be the random variable which counts the number of successes in a **Bernoulli trials process** with parameters n and p.
- Then the distribution $b(n, p, k)$ of B is called the binomial distribution.

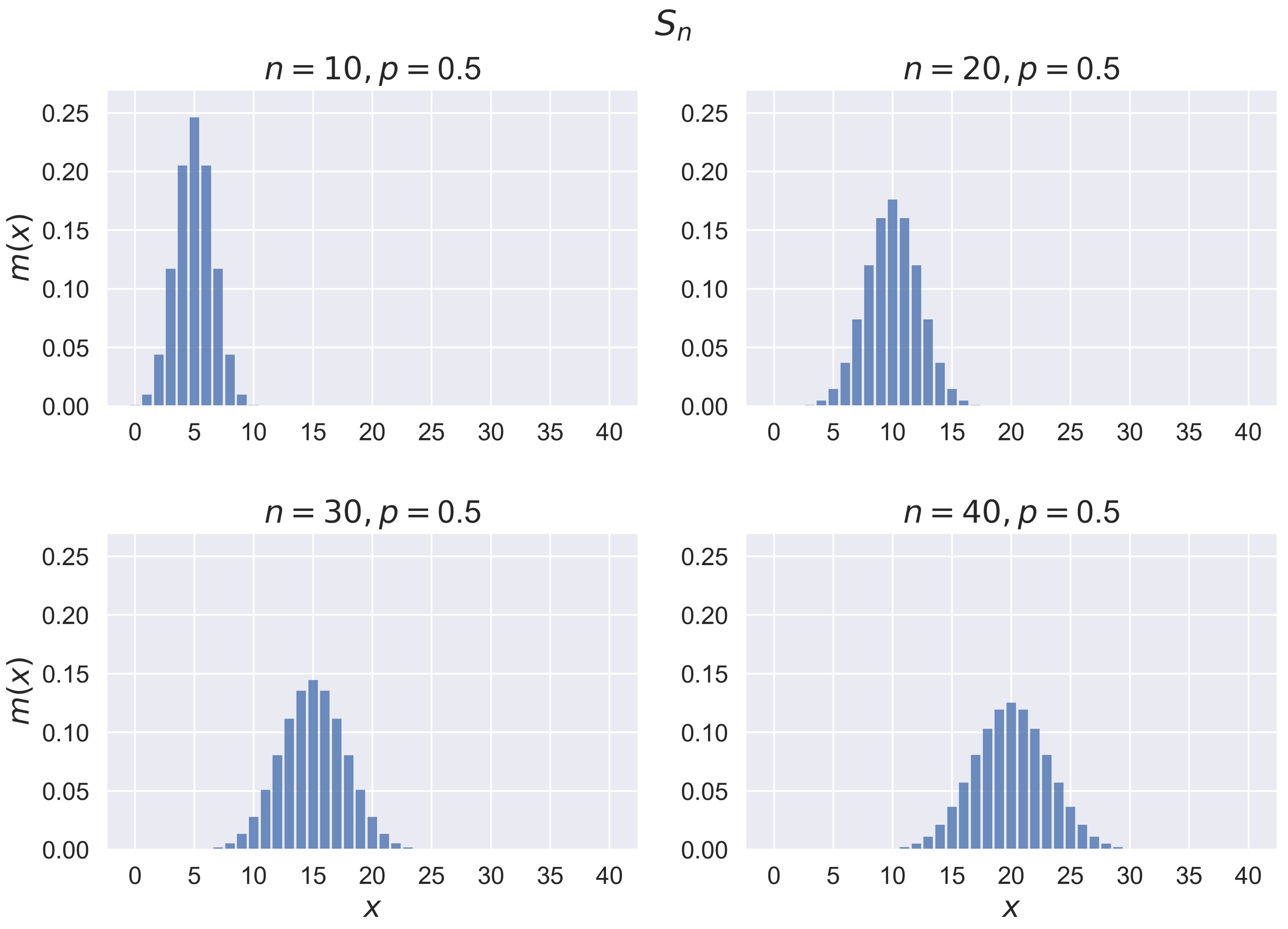
Binomial Distribution

$$b(n, p, k) = \binom{n}{k} p^k q^{n-k}$$



Binomial Distribution

- Consider a Bernoulli trials process with probability p for success on each trial.
- Let $X_i = 1$ or 0 according as the ith outcome is a success or failure.
- Let $S_n = X_1 + X_2 + \cdots + X_n$. Then S_n is the number of successes in n trials.
- We know that S_n has as its distribution the binomial probabilities $b(n, p, k)$.



n increases

drifts off to the right

flatten out

XC 2020

The Standardized Sum of S_n

- Consider a Bernoulli trials process with probability p for success on each trial.
- Let $X_i = 1$ or 0 according as the ith outcome is a success or failure.
- Let $S_n = X_1 + X_2 + \cdots + X_n$. Then S_n is the number of successes in n trials.
- We know that S_n has as its distribution the binomial probabilities $b(n, p, k)$.
- The standardized sum of S_n is given by

$$S_n^* = \frac{S_n - np}{\sqrt{npq}}.$$

$\mu = \cdots$

$\sigma^2 = \cdots$

The Standardized Sum of S_n

- Consider a Bernoulli trials process with probability p for success on each trial.
- Let $X_i = 1$ or 0 according as the ith outcome is a success or failure.
- Let $S_n = X_1 + X_2 + \cdots + X_n$. Then S_n is the number of successes in n trials.
- We know that S_n has as its distribution the binomial probabilities $b(n, p, k)$.
- The standardized sum of S_n is given by

$$S_n^* = \frac{S_n - np}{\sqrt{npq}}.$$

- S_n^* always has expected value 0 and variance 1.

$\mu = 0$

$\sigma^2 = 1$

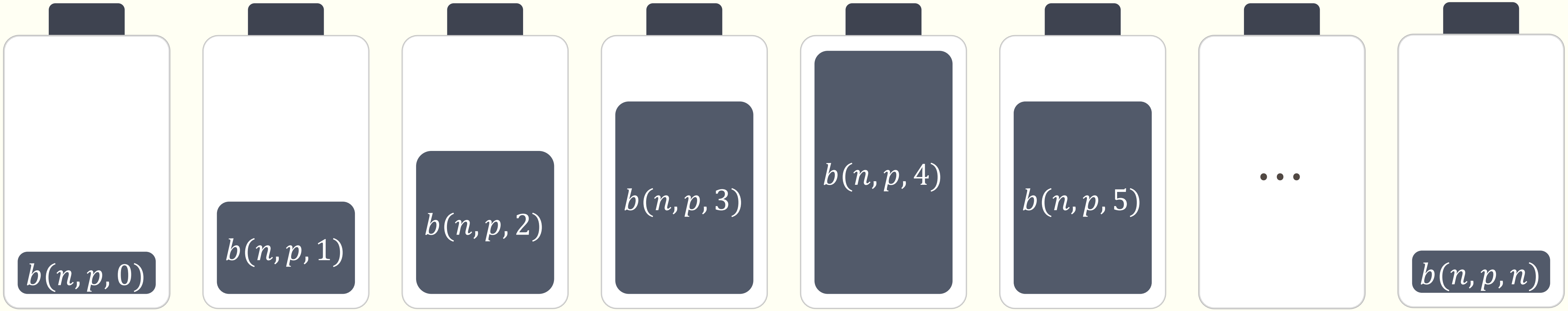
The Standardized Sum of S_n

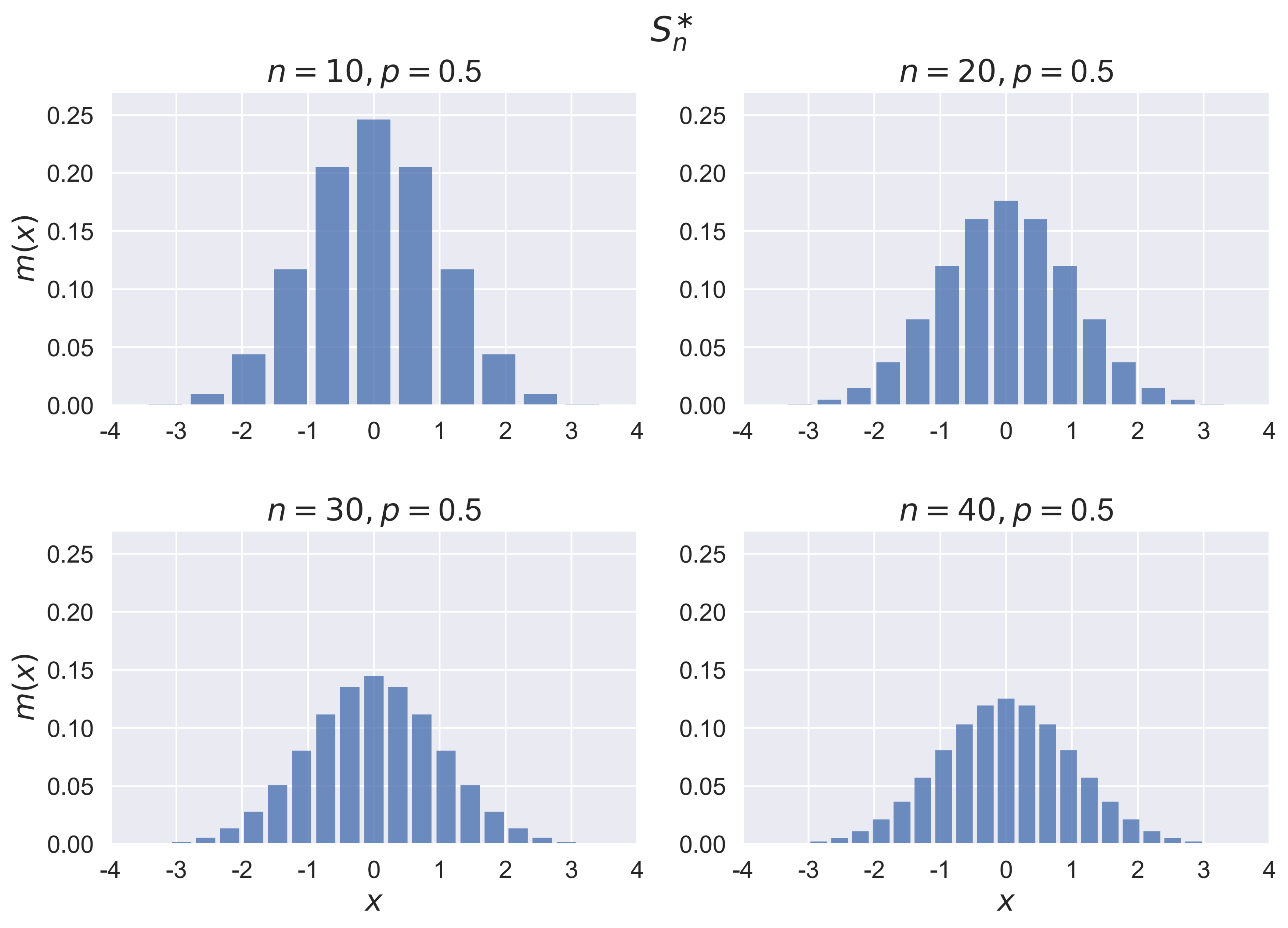
$$S_n^* = \frac{S_n - np}{\sqrt{npq}}$$

=

$b(n, p, k)$

k	0	1	2	3	4	5	...	n
x	$\frac{0-np}{\sqrt{npq}}$	$\frac{1-np}{\sqrt{npq}}$	$\frac{2-np}{\sqrt{npq}}$	$\frac{3-np}{\sqrt{npq}}$	$\frac{4-np}{\sqrt{npq}}$	$\frac{5-np}{\sqrt{npq}}$	...	$\frac{n-np}{\sqrt{npq}}$



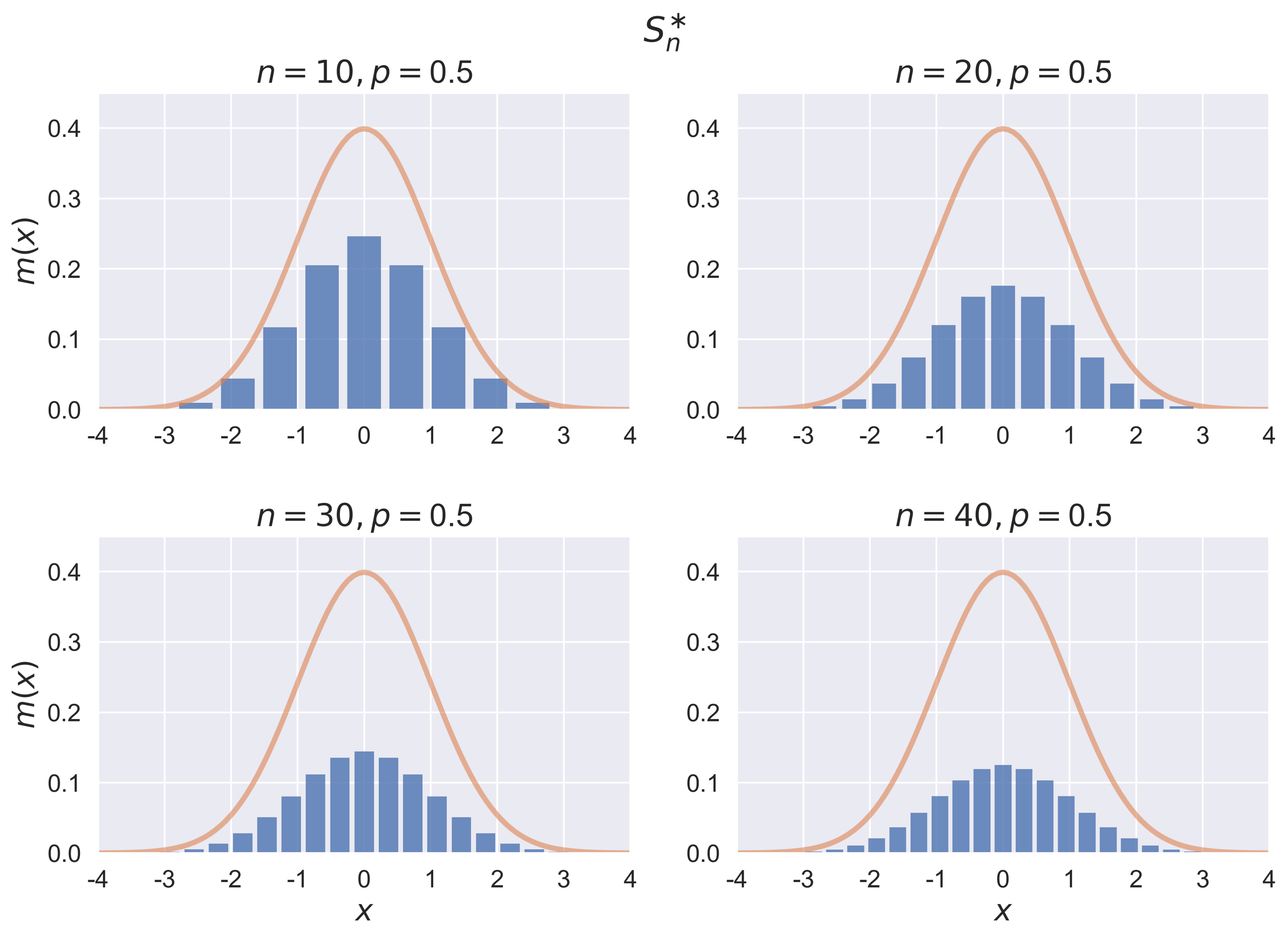


$$S_n^* = \frac{S_n - np}{\sqrt{npq}}$$

$$\mu = 0$$

$$\sigma^2 = 1$$

XC 2020

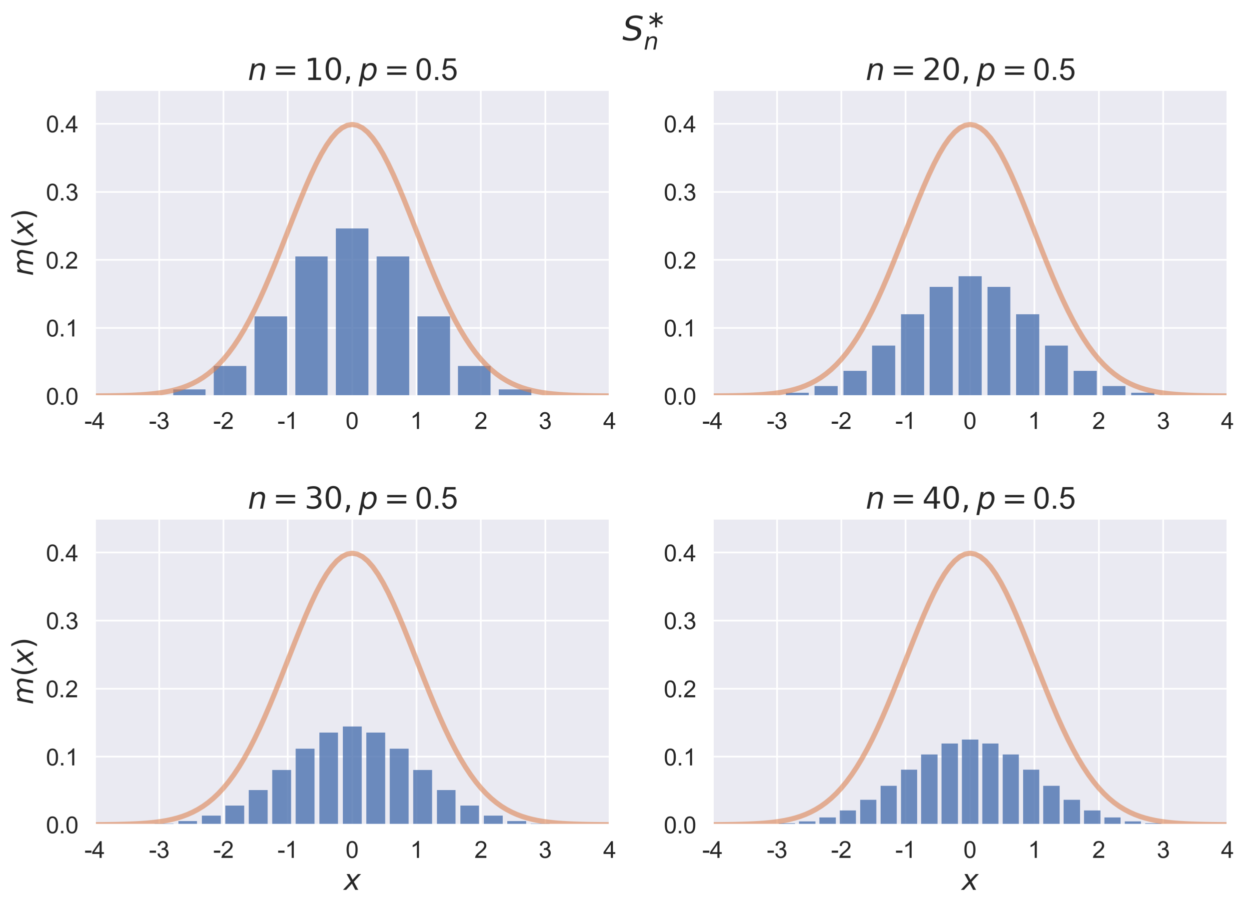


standardized sum

standard normal distribution

shapes: same
heights: different

why?



standardized sum

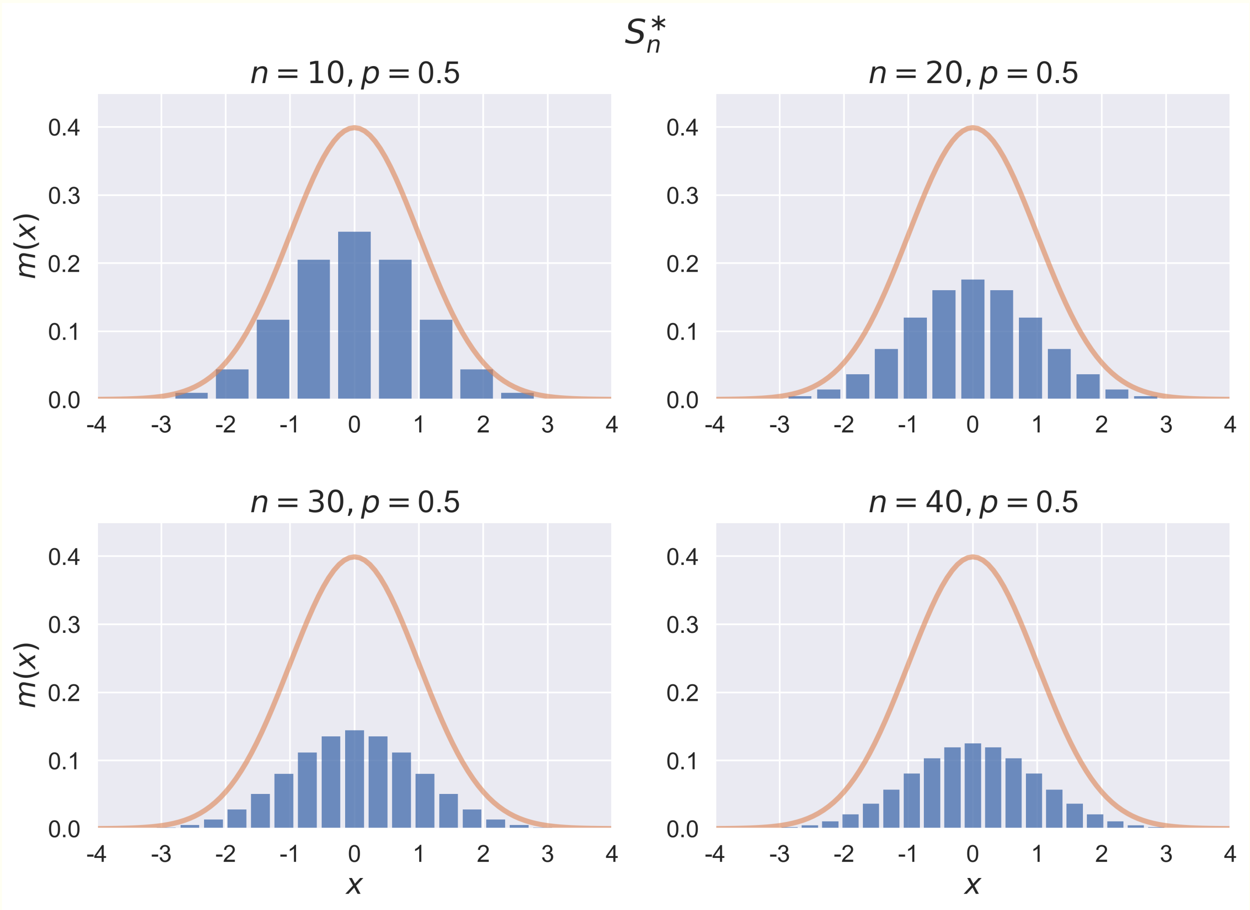
standard normal distribution

shapes: same
heights: different

why?

standardized sum: $\sum_{k=0}^{n} h_i = 1$

standard normal distribution
: $\int_{-\infty}^{+\infty} f(x)dx = 1$



standardized sum: $\sum_{k=0}^{n} h_i = 1$

standard normal distribution
: $\int_{-\infty}^{+\infty} f(x)dx = 1 \approx \sum_{k=0}^{n} f(x_i)dx$

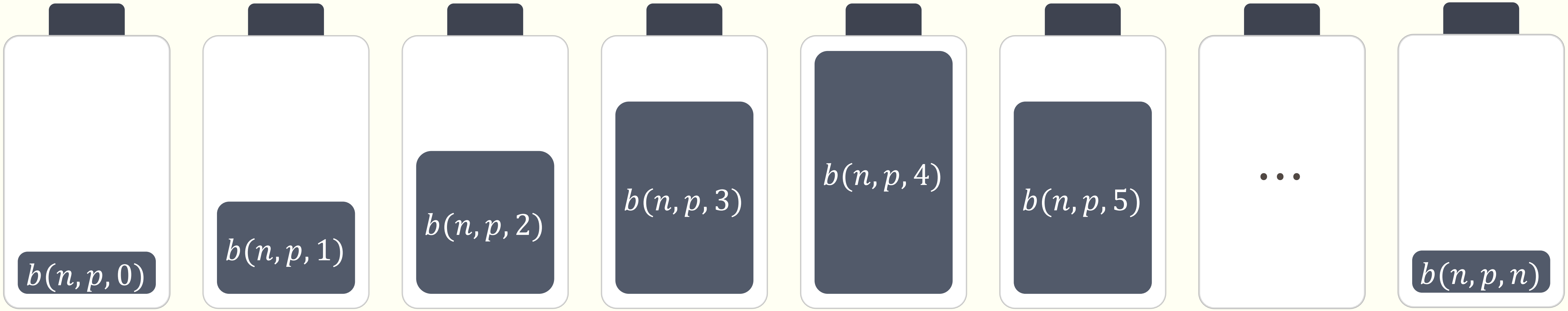
$$\sum_{k=0}^{n} h_i = 1$$

$$\sum_{k=0}^{n} f(x_i)dx \approx 1$$

$$h_k = f(x_k)dx$$

$$f(x_k) = \frac{h_k}{dx}$$

0	1	2	3	4	5	...	n



$\frac{0-np}{\sqrt{npq}}$	$\frac{1-np}{\sqrt{npq}}$	$\frac{2-np}{\sqrt{npq}}$	$\frac{3-np}{\sqrt{npq}}$	$\frac{4-np}{\sqrt{npq}}$	$\frac{5-np}{\sqrt{npq}}$	...	$\frac{n-np}{\sqrt{npq}}$

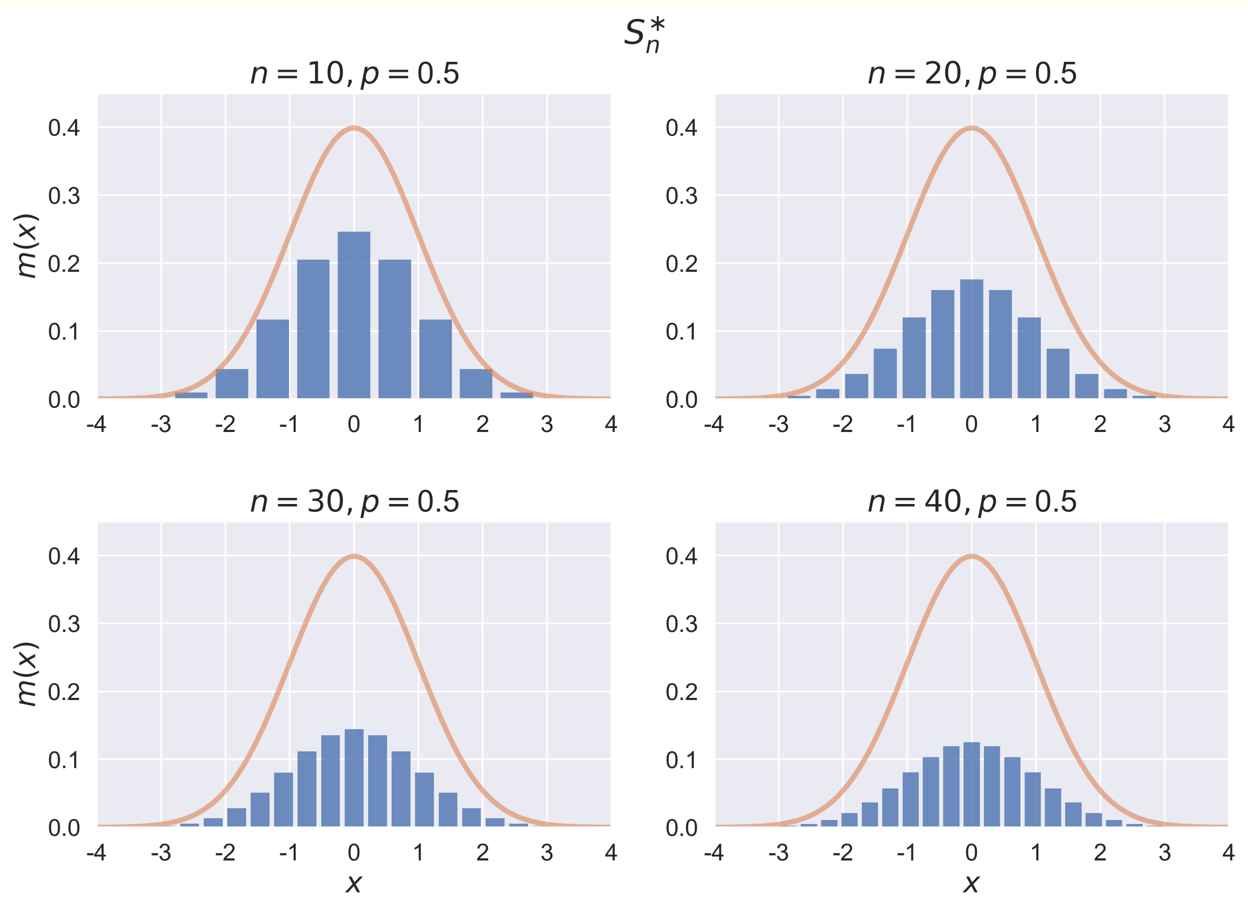
$$S_n^* = \frac{S_n - np}{\sqrt{npq}}$$ =

$$h_k = b(n,p,k)$$ =

$$dx = x_{k+1} - x_k = \frac{1}{\sqrt{npq}}$$ =

$$k = \langle \sqrt{npq}\, x_k + np \rangle$$ =

$\langle a \rangle$: the integer nearest to a



$$h_k = b(n, p, k)$$

$$dx = x_{k+1} - x_k = \frac{1}{\sqrt{npq}}$$

$$k = \langle \sqrt{npq}\, x_k + np \rangle$$

$$f(x_k) = \frac{h_k}{dx} = \sqrt{npq}\;\; b(n, p, \langle \sqrt{npq}\, x_k + np \rangle)$$

Central Limit Theorem for Binomial Distributions

- For the binomial distribution $b(n,p,k)$ we have

$$\lim_{n\to\infty} \sqrt{npq}\, b\big(n,p,\langle np + x\sqrt{npq}\rangle\big) = \phi(x),$$

  where $\phi(x)$ is the standard normal density.

- The proof of this theorem can be carried out using Stirling's approximation.

Stirling's Formula
The sequence $n!$ is asymptotically equal to
$$n^n e^{-n}\sqrt{2\pi n}.$$

Approximating Binomial Distributions

$$\lim_{n\to\infty} \sqrt{npq}\, b(n, p, \langle np + x\sqrt{npq}\rangle) = \phi(x)$$

=

binomial → normal

$$b(n, p, k) \approx \cdots$$

?

normal → binomial

Approximating Binomial Distributions

$$\lim_{n\to\infty} \sqrt{npq}\, b(n, p, \langle np + x\sqrt{npq}\rangle) = \phi(x)$$

=

$$b(n, p, k) \approx \cdots$$

?

$$k = np + x\sqrt{npq}$$

$$x = \frac{k - np}{\sqrt{npq}}$$

$$b(n, p, k) \approx \frac{\phi(x)}{\sqrt{npq}}$$

$$b(n, p, k) \approx \frac{1}{\sqrt{npq}} \phi\left(\frac{k - np}{\sqrt{npq}}\right)$$

Example

Toss a coin

Find the probability of exactly 55 heads in 100 tosses of a coin.

$$\lim_{n\to\infty}\sqrt{npq}\,b(n,p,\langle np+x\sqrt{npq}\rangle)=\phi(x)$$

=

$$b(n,p,k)\approx\frac{1}{\sqrt{npq}}\phi\left(\frac{k-np}{\sqrt{npq}}\right)$$

Example

Toss a coin

Find the probability of exactly 55 heads in 100 tosses of a coin.

$n = 100$

$p = \frac{1}{2}$

$np = 50$

$\sqrt{npq} = 5$

$k = 55$

$x = \frac{k - np}{\sqrt{npq}} = 1$

$b(n, p, k) \approx \frac{1}{\sqrt{npq}} \phi(\frac{k - np}{\sqrt{npq}})$

$b(100, \frac{1}{2}, 55) \approx \frac{1}{5} \phi(1)$

Example

Toss a coin

Find the probability of exactly 55 heads in 100 tosses of a coin.

Standard normal distribution Z

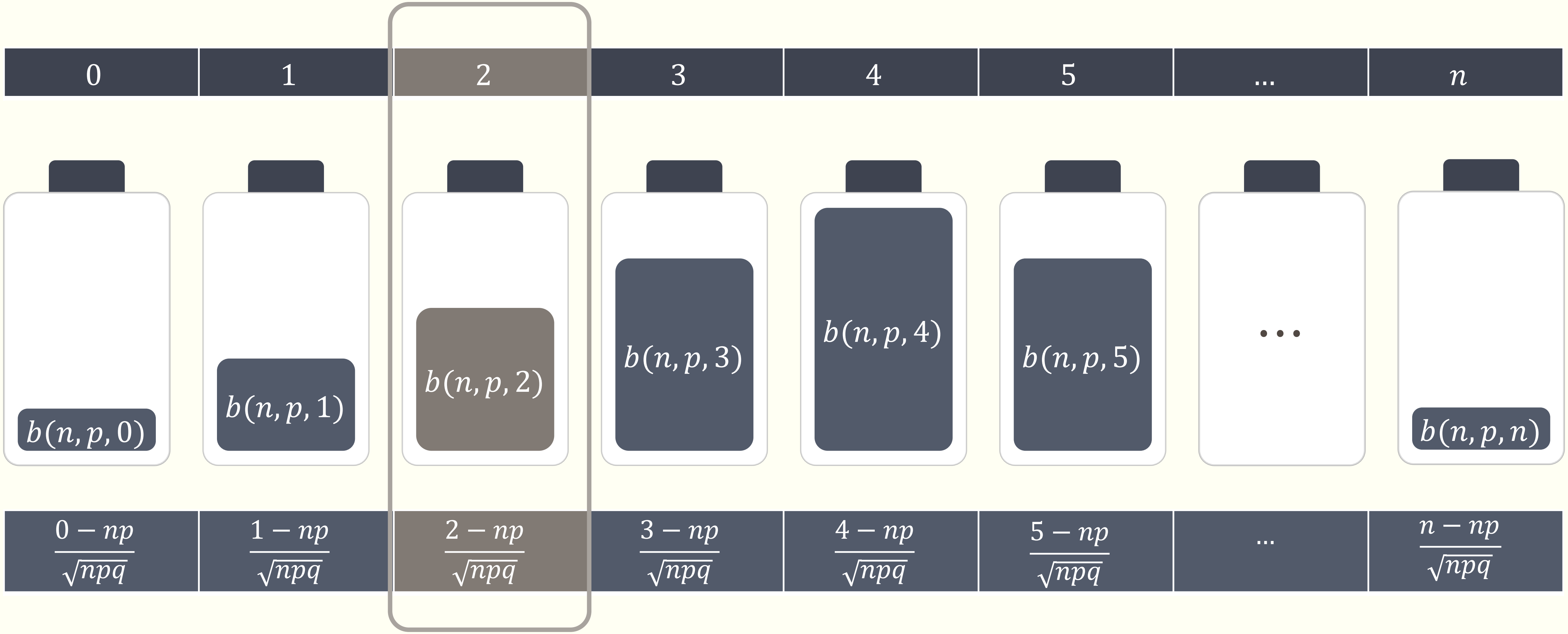
$\mu = 0$ and $\sigma = 1$

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

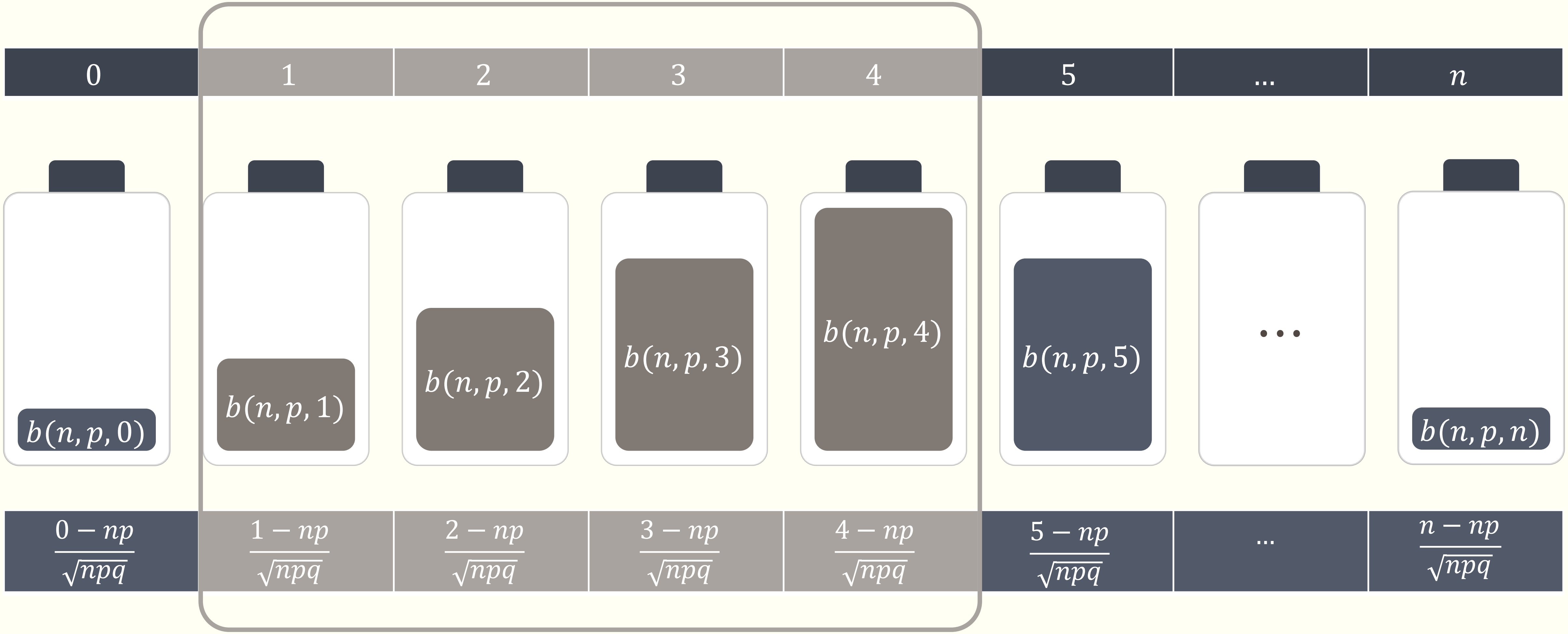
$$b(n,p,k) \approx \frac{1}{\sqrt{npq}} \phi(\frac{k-np}{\sqrt{npq}})$$

$$b\left(100, \frac{1}{2}, 55\right) \approx \frac{1}{5}\phi(1) = \frac{1}{5}(\frac{1}{\sqrt{2\pi}} e^{-1/2})$$

A specific outcome



Outcomes in an Interval



Central Limit Theorem for Binomial Distributions

- Let S_n be the number of successes in n Bernoulli trials with probability p for success, and let a and b be two fixed real numbers.
- Then

$$\lim_{n\to+\infty} P\left(a \le \frac{S_n - np}{\sqrt{npq}} \le b\right) = \int_a^b \phi(x)dx = \mathrm{NA}(a, b).$$

- We denote this area by $\mathrm{NA}(a, b)$.

$$a \le \frac{S_n - np}{\sqrt{npq}} = S_n^* \le b$$

$$a\sqrt{npq} + np \le S_n \le b\sqrt{npq} + np$$

Approximating Binomial Distributions

$$a \leq \frac{S_n - np}{\sqrt{npq}} = S_n^* \leq b$$

$$a\sqrt{npq} + np \leq S_n \leq b\sqrt{npq} + np$$

$$\lim_{n \to +\infty} P\left(a \leq \frac{S_n - np}{\sqrt{npq}} \leq b\right) = \mathrm{NA}(a, b).$$

$$i \leq S_n \leq j$$

$$a = \frac{i - np}{\sqrt{npq}} \leq S_n^* \leq \frac{j - np}{\sqrt{npq}} = b$$

$$\lim_{n \to +\infty} P(i \leq S_n \leq j) = \lim_{n \to +\infty} P\left(\frac{i - np}{\sqrt{npq}} \leq S_n^* \leq \frac{j - np}{\sqrt{npq}}\right) = \mathrm{NA}\left(\frac{i - np}{\sqrt{npq}}, \frac{j - np}{\sqrt{npq}}\right).$$

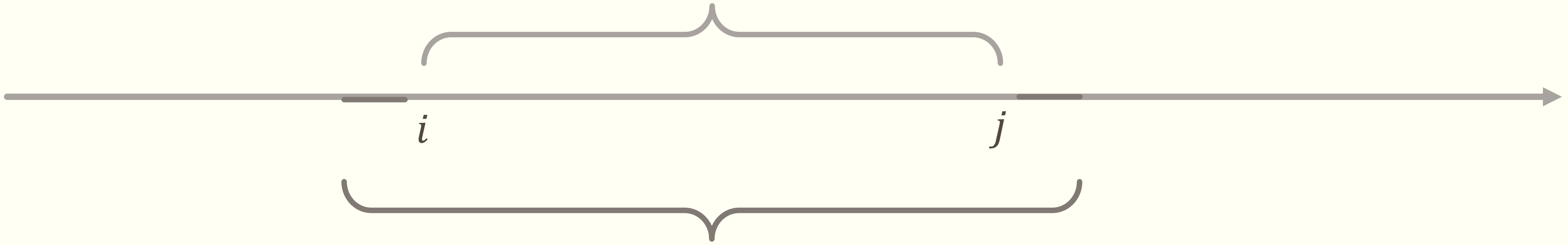
Approximating Binomial Distributions (more accurate)

$i \le S_n \le j$

$$a = \frac{i - np}{\sqrt{npq}} \le S_n^* \le \frac{j - np}{\sqrt{npq}} = b$$

$$\lim_{n\to+\infty} P(i \le S_n \le j) = \lim_{n\to+\infty} P\left(\frac{i-np}{\sqrt{npq}} \le S_n^* \le \frac{j-np}{\sqrt{npq}}\right) = \mathrm{NA}\left(\frac{i-np}{\sqrt{npq}}, \frac{j-np}{\sqrt{npq}}\right).$$

$$\lim_{n\to+\infty} P(i \le S_n \le j) = \mathrm{NA}\left(\frac{i-\frac{1}{2}-np}{\sqrt{npq}}, \frac{j+\frac{1}{2}-np}{\sqrt{npq}}\right).$$



Example

Toss a coin

A coin is tossed 100 times. Estimate the probability that the number of heads lies between 40 and 60 (including the end points).

$$\lim_{n\to+\infty} P(i \le S_n \le j) = \mathrm{NA}(\frac{i-\frac{1}{2}-np}{\sqrt{npq}}, \frac{j+\frac{1}{2}-np}{\sqrt{npq}})$$

=

Toss a coin

A coin is tossed 100 times. Estimate the probability that the number of heads lies between 40 and 60 (including the end points).

the outcome will not deviate by more than two standard deviations from the expected value

$n = 100$ | $p = \frac{1}{2}$

$np = 50$ | $\sqrt{npq} = 5$

$i = 40 \rightarrow \dfrac{i - \frac{1}{2} - np}{\sqrt{npq}} = -2.1$

$j = 60 \rightarrow \dfrac{j + \frac{1}{2} - np}{\sqrt{npq}} = 2.1$

Toss a coin

A coin is tossed 100 times. Estimate the probability that the number of heads lies between 40 and 60 (including the end points).

$i = 40$

$j = 60$

$$\frac{i - \frac{1}{2} - np}{\sqrt{npq}} = -2.1$$

$$\frac{j + \frac{1}{2} - np}{\sqrt{npq}} = 2.1$$

$$\lim_{n\to+\infty} P(i \le S_n \le j) = \mathrm{NA}\left(\frac{i - \frac{1}{2} - np}{\sqrt{npq}}, \frac{j + \frac{1}{2} - np}{\sqrt{npq}}\right)$$

=

$$\lim_{n\to+\infty} P(40 \le S_n \le 60) = \mathrm{NA}(-2.1,2.1)$$
$$= 2\mathrm{NA}(0, 2.1)$$

=

$2\sigma \longrightarrow 2.1\sigma$

Approximating Binomial Distribution

$\boldsymbol{S_n^* = x}$

$$\lim_{n\to\infty} \sqrt{npq}\, b(n, p, \langle np + x\sqrt{npq}\rangle) = \phi(x)$$

$\boldsymbol{a \le S_n^* \le b}$

$$\lim_{n\to+\infty} P\left(a \le \frac{S_n - np}{\sqrt{npq}} \le b\right) = \mathrm{NA}(a, b).$$

CLT

$\boldsymbol{S_n = k}$

$$b(n, p, k) \approx \frac{1}{\sqrt{npq}} \phi\left(\frac{k - np}{\sqrt{npq}}\right)$$

$\boldsymbol{i \le S_n \le j}$

$$\lim_{n\to+\infty} P(i \le S_n \le j) = \mathrm{NA}\left(\frac{i - \frac{1}{2} - np}{\sqrt{npq}}, \frac{j + \frac{1}{2} - np}{\sqrt{npq}}\right).$$

CENTRAL LIMIT THEOREM FOR DISCRETE INDEPENDENT TRIALS

any independent trials process such that the individual trials have finite variance

The Standardized Sum of S_n

- Consider an independent trials process with common distribution function $m(x)$ defined on the integers, with expected value μ and variance σ^2.
- Let $S_n = X_1 + X_2 + \cdots + X_n$ be the sum of n independent discrete random variables of the process.
- The standardized sum of S_n is given by

$$S_n^* = \frac{S_n - n\mu}{\sqrt{n\sigma^2}}.$$

- S_n^* always has expected value 0 and variance 1.

$E(S_n) = \cdots$

$V(S_n) = \cdots$

$$S_n^* = \frac{S_n - n\mu}{\sqrt{n\sigma^2}}$$

=

k

$$x_k = \frac{k - n\mu}{\sqrt{n\sigma^2}}$$

independent trials process

$P(S_n = k)$

normal

$\sqrt{n\sigma^2}P(S_n = k)$

Approximation Theorem

- Let $X_1, X_2, \cdots, X_n$ be an independent trials process and let $S_n = X_1 + X_2 + \cdots + X_n$.
- Assume that the greatest common divisor of the differences of all the values that the X_k can take on is 1.
- Let $E(X_k) = \mu$ and $V(X_k) = \sigma^2$.
- Then for n large,

$$P(S_n = k) \approx \frac{1}{\sqrt{n\sigma^2}} \phi(\frac{k-n\mu}{\sqrt{n\sigma^2}}).$$

- Here $\phi(x)$ is the standard normal density.

Central Limit Theorem for a Discrete Independent Trials Process

- Let $S_n = X_1 + X_2 + \cdots + X_n$ be the sum of n discrete independent random variables with common distribution having expected value μ and variance σ^2.
- Then, for $a < b$,

$$\lim_{n\to+\infty} P\left(a \le \frac{S_n - n\mu}{\sqrt{n\sigma^2}} \le b\right) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-x^2/2}dx = \mathrm{NA}(a,b).$$

Approximating a Discrete Independent Trials Process

$\boldsymbol{S_n = k}$

$$P(S_n = k) \approx \frac{1}{\sqrt{n\sigma^2}}\phi\left(\frac{k - n\mu}{\sqrt{n\sigma^2}}\right)$$

CLT

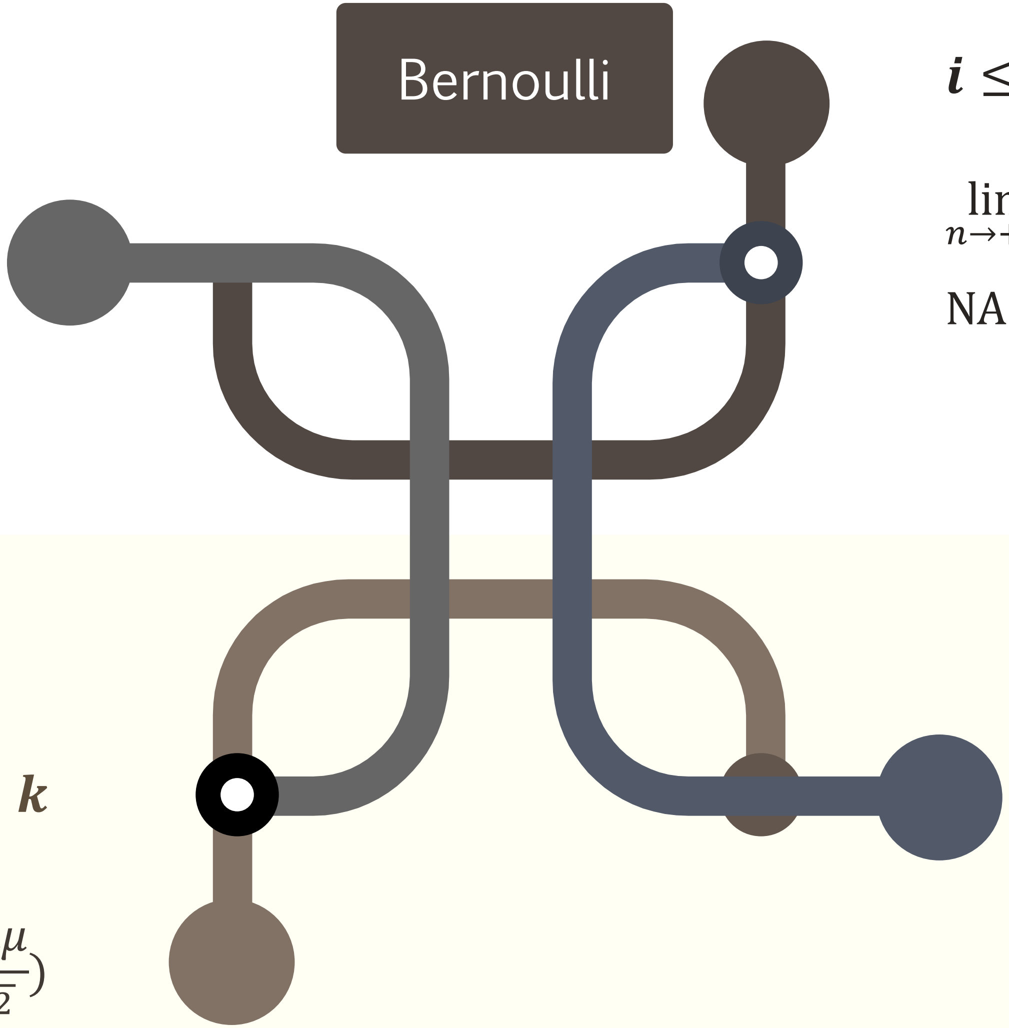
$\boldsymbol{c \leq S_n \leq d}$

$$\lim_{n\to+\infty} P(c \leq S_n \leq d) = \mathrm{NA}\left(\frac{c-n\mu}{\sqrt{n\sigma^2}}, \frac{d-n\mu}{\sqrt{n\sigma^2}}\right).$$

Approximating a Discrete Independent Trials Process

$\boldsymbol{S_n = k}$

$$b(n, p, k) \approx \frac{1}{\sqrt{npq}} \phi\left(\frac{k - np}{\sqrt{npq}}\right)$$



$\boldsymbol{i \leq S_n \leq j}$

$$\lim_{n \to +\infty} P(i \leq S_n \leq j) = \mathrm{NA}\left(\frac{i - \frac{1}{2} - np}{\sqrt{npq}}, \frac{j + \frac{1}{2} - np}{\sqrt{npq}}\right).$$

$\boldsymbol{S_n = k}$

$$P(S_n = k) \approx \frac{1}{\sqrt{n\sigma^2}} \phi\left(\frac{k - n\mu}{\sqrt{n\sigma^2}}\right)$$

$\boldsymbol{c \leq S_n \leq d}$

$$\lim_{n \to +\infty} P(c \leq S_n \leq d) = \mathrm{NA}\left(\frac{c - n\mu}{\sqrt{n\sigma^2}}, \frac{d - n\mu}{\sqrt{n\sigma^2}}\right).$$

Example

A surveying instrument makes an error of -2, -1, 0, 1, or 2 feet with equal probabilities when measuring the height of a 200-foot tower.

Example

error	-2	-1	0	1	2
probability	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

Find the expected value and the variance
for the height obtained using this instrument once.

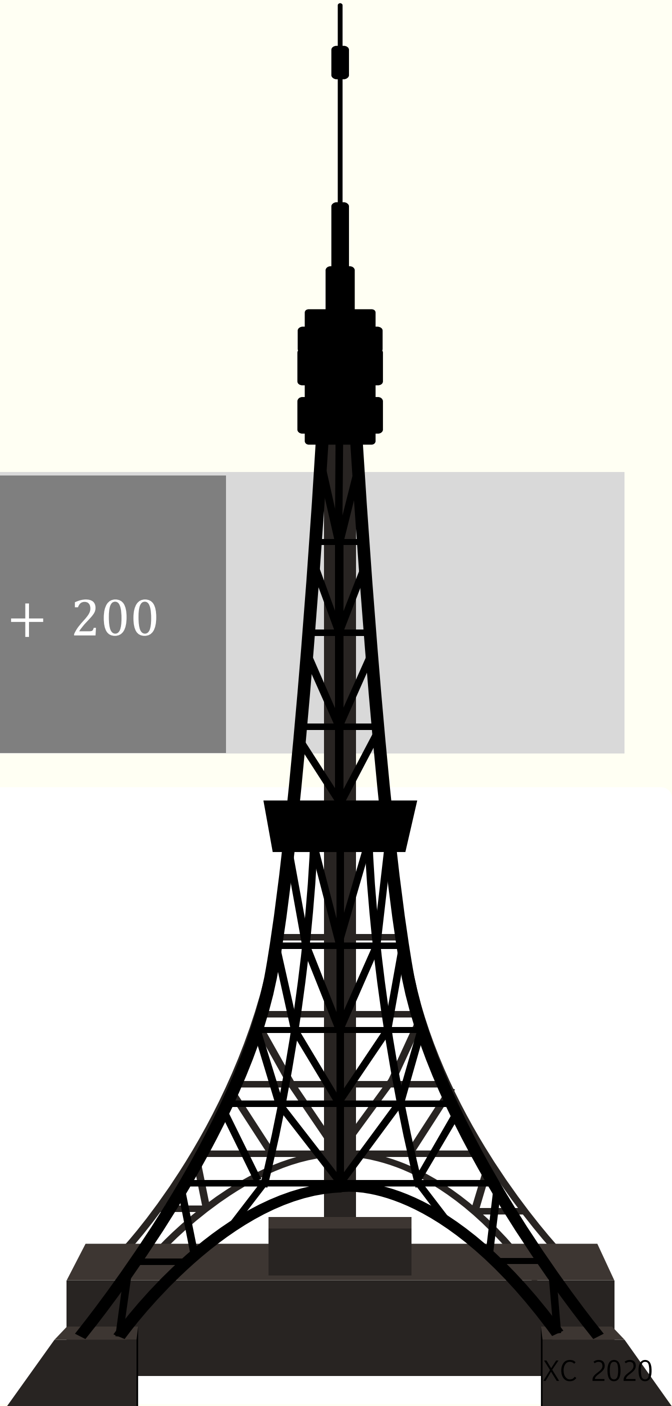
$$\text{height} = \text{error} + 200$$

$$E(\text{error}) = \frac{1}{5}(-2 - 1 + 0 + 1 + 2) = 0$$

$$E(\text{height}) \\ = 0 + 200 = 200$$

$$V(\text{error}) = \frac{1}{5}\left((-2)^2 + (-1)^2 + 0^2 + 1^2 + 2^2\right) = 2$$

$$V(\text{height}) = 2$$



Example

Estimate the probability that in 18 independent measurements of this tower, the average of the measurements is between 199 and 201, inclusive.

$\mu = 200$

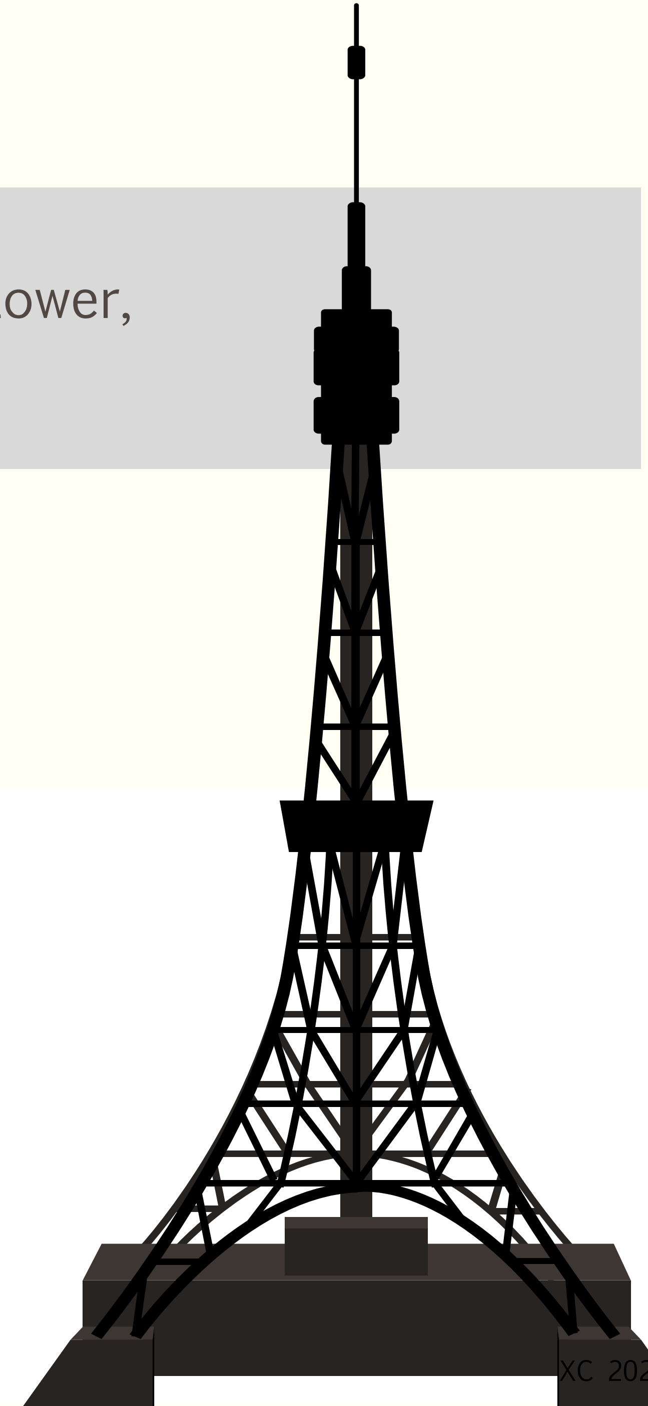
$\sigma^2 = 2$

$n = 18$

$199 \leq \frac{S_n}{n} = \frac{S_n}{18} \leq 201$

$$\lim_{n \to +\infty} P(c \leq S_n \leq d) = \mathrm{NA}\left(\frac{c - n\mu}{\sqrt{n\sigma^2}}, \frac{d - n\mu}{\sqrt{n\sigma^2}}\right).$$

=



XC 2020

Example

$\mu = 200$

$\sigma^2 = 2$

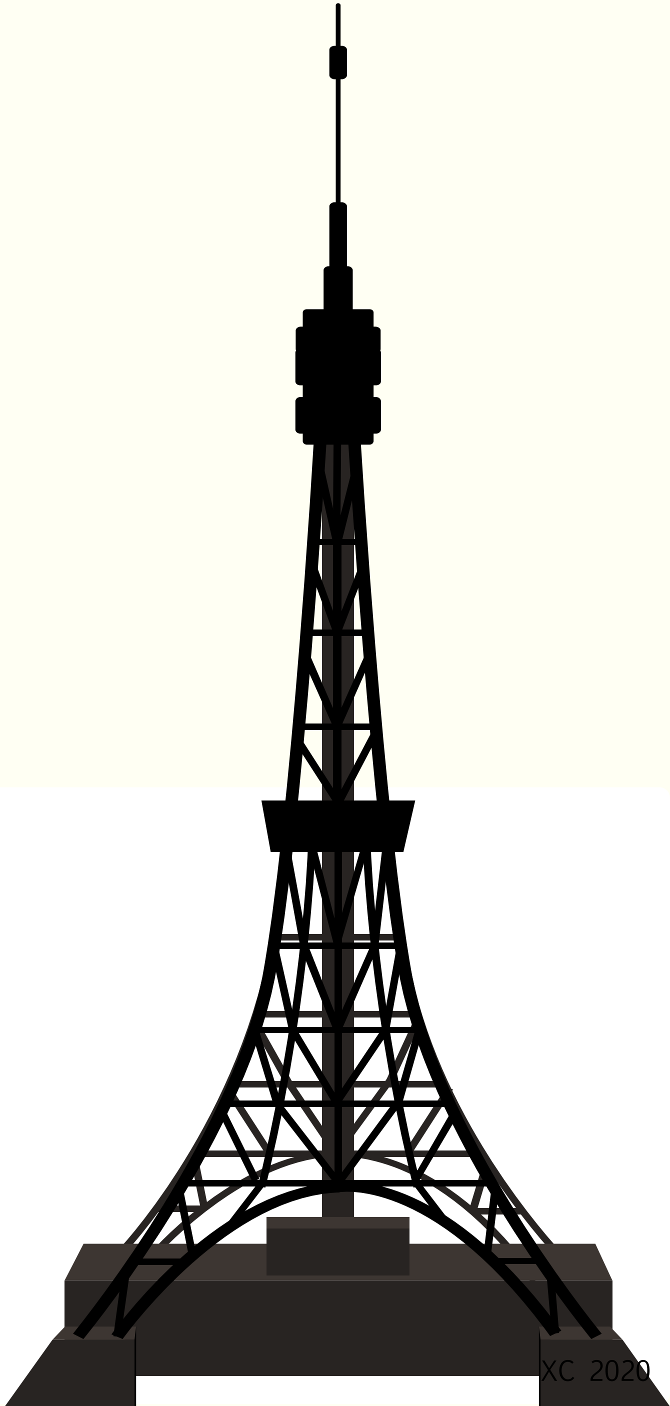
$n = 18$

$199 \le \frac{S_n}{n} = \frac{S_n}{18} \le 201$

$$\lim_{n\to+\infty} P(c \le S_n \le d) = \mathrm{NA}\left(\frac{c-n\mu}{\sqrt{n\sigma^2}}, \frac{d-n\mu}{\sqrt{n\sigma^2}}\right).$$

=

$$P\left(199 \le \frac{S_n}{18} \le 201\right) = P(3582 \le S_n \le 3618)$$
$$\approx \mathrm{NA}\left(\frac{3582-3600}{\sqrt{36}}, \frac{3618-3600}{\sqrt{36}}\right) = \mathrm{NA}(-3, 3).$$



Central Limit Theorem

01

Bernoulli Trials

- independent
- identical

02

Discrete Trials

- independent
  - identical

03

Discrete Random Variables

- independent

Central Limit Theorem

- Let $X_1, X_2, \cdots, X_n$ be a sequence of independent discrete random variables. There exist a constant A, such that $|X_i| \leq A$ for all i.
- Let $S_n = X_1 + X_2 + \cdots + X_n$ be the sum. Assume that $S_n \to \infty$.
- For each i, denote the expected value and variance of X_i by μ_i and σ_i^2, respectively.
- Define the expected value and variance of S_n to be m_n and s_n^2, respectively.
- For $a < b$,

$$\lim_{n \to +\infty} P\left(a \leq \frac{S_n - m_n}{s_n} \leq b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx = \mathrm{NA}(a, b).$$

CENTRAL LIMIT THEOREM FOR CONTINUOUS INDEPENDENT TRIALS

continuous random variables with a common density function

Central Limit Theorem

- Let $S_n = X_1 + X_2 + \cdots + X_n$ be the sum of n independent continuous random variables with common density function p having expected value μ and variance σ^2.
- Let

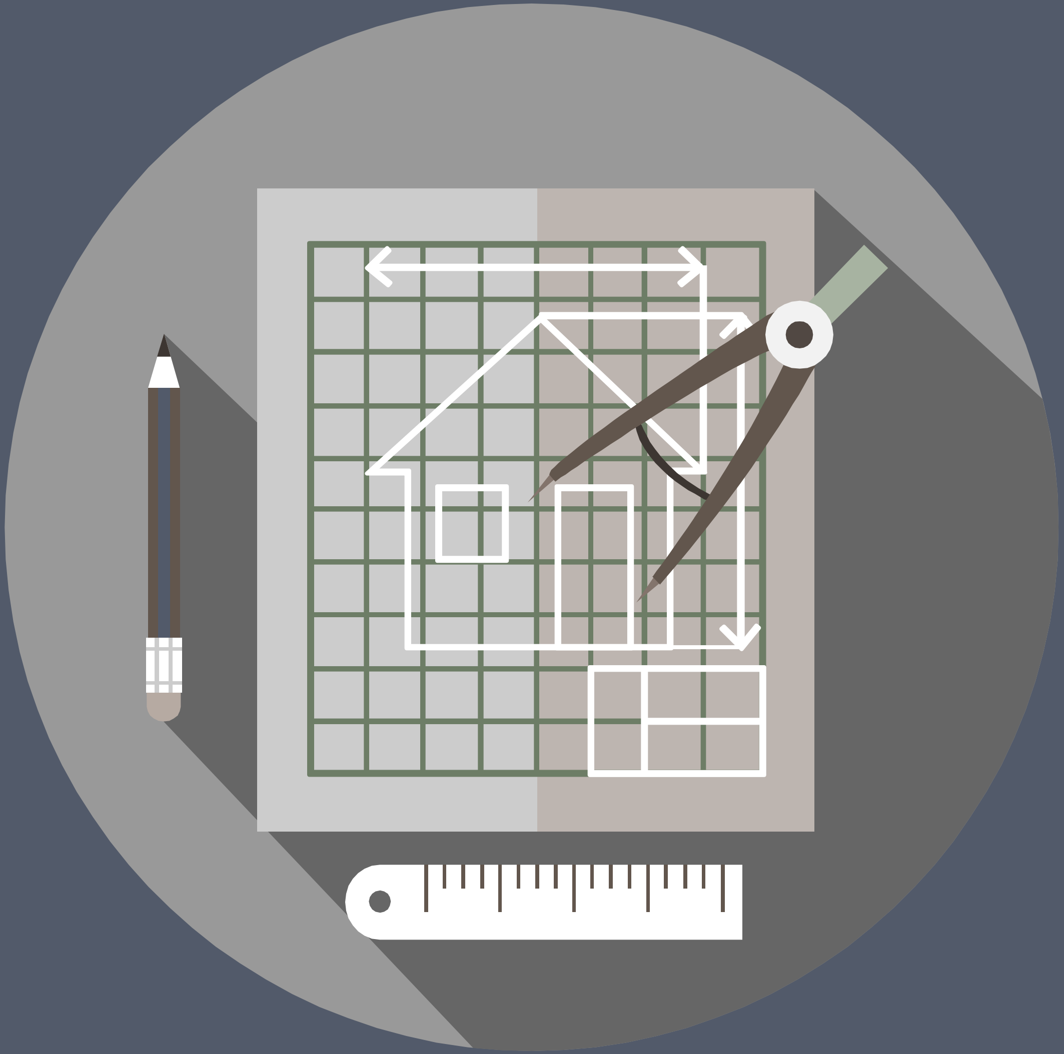
$$S_n^* = \frac{S_n - n\mu}{\sqrt{n\sigma^2}}.$$

- Then we have, for all $a < b$,

$$\lim_{n \to +\infty} P(a \le S_n^* \le b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx = \mathrm{NA}(a, b).$$

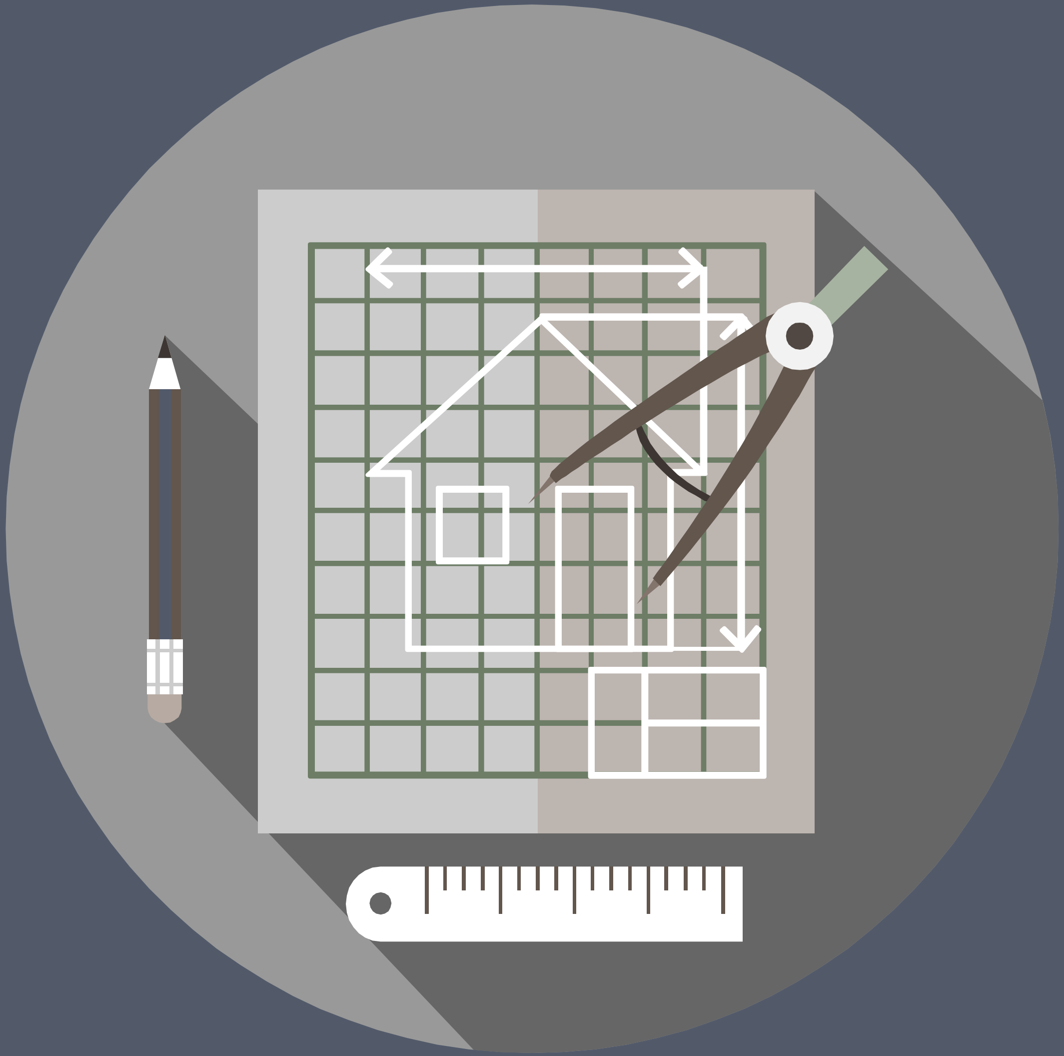
Example

- Suppose a surveyor wants to measure a known distance, say of 1 mile, using a transit and some method of triangulation.
- He knows that because of possible motion of the transit, atmospheric distortions, and human error, any one measurement is apt to slightly in error.
- He plans to make several measurements and take an average.
- He assumes that his measurements are independent random variables with common distribution of mean $\mu = 1$ and standard deviation $\sigma = 0.0002$.



Example

- He can say that if n is large, the average $\frac{S_n}{n}$ has a density function that is approximately normal, with mean 1 mile, and standard deviation $\frac{0.0002}{\sqrt{n}}$ miles.
- How many measurements should he make to be reasonably sure that his average lies within 0.0001 of the true value?



Example

- He can say that if n is large, the average $\frac{S_n}{n}$ has a density function that is approximately normal, with mean 1 mile, and standard deviation $\frac{0.0002}{\sqrt{n}}$ miles.

Expected Value

$$\mu = 1$$

Standard Deviation

$$\sigma = 0.0002$$

Variance

$$\sigma^2 = (0.0002)^2$$

Sum of n independent measurements

$$S_n = X_1 + X_2 + \cdots + X_n$$

Average of n independent measurements

$$A_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

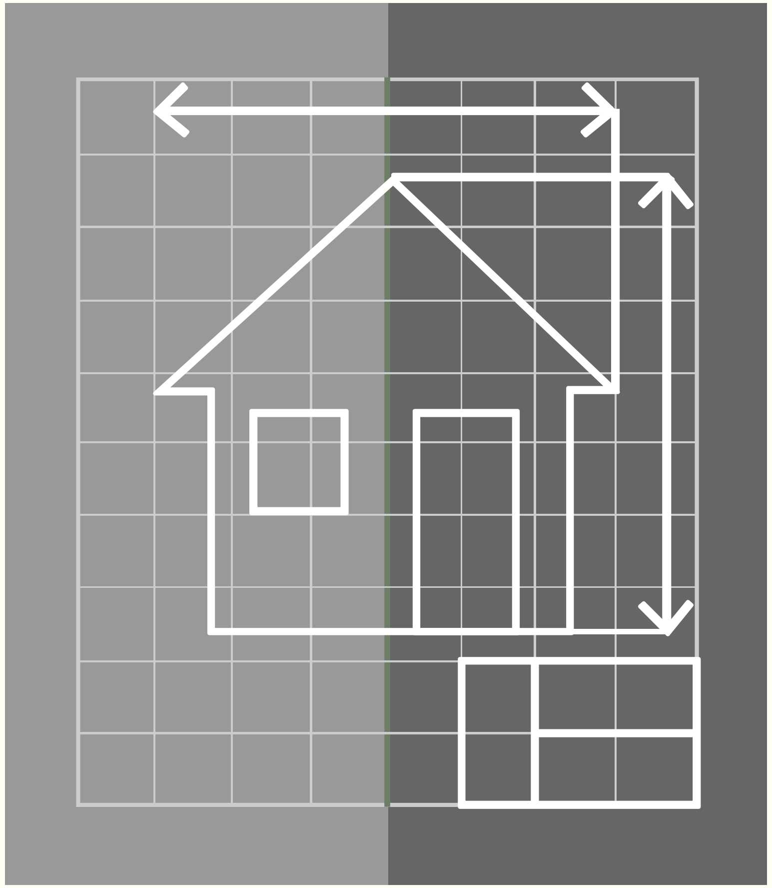
$$E(A_n) = \mu$$

$$D(A_n) = \frac{\sigma}{\sqrt{n}}$$

$$V(A_n) = \frac{\sigma^2}{n}$$

Example

- How many measurements should he make to be reasonably sure that his average lies within 0.0001 of the true value?



LLN

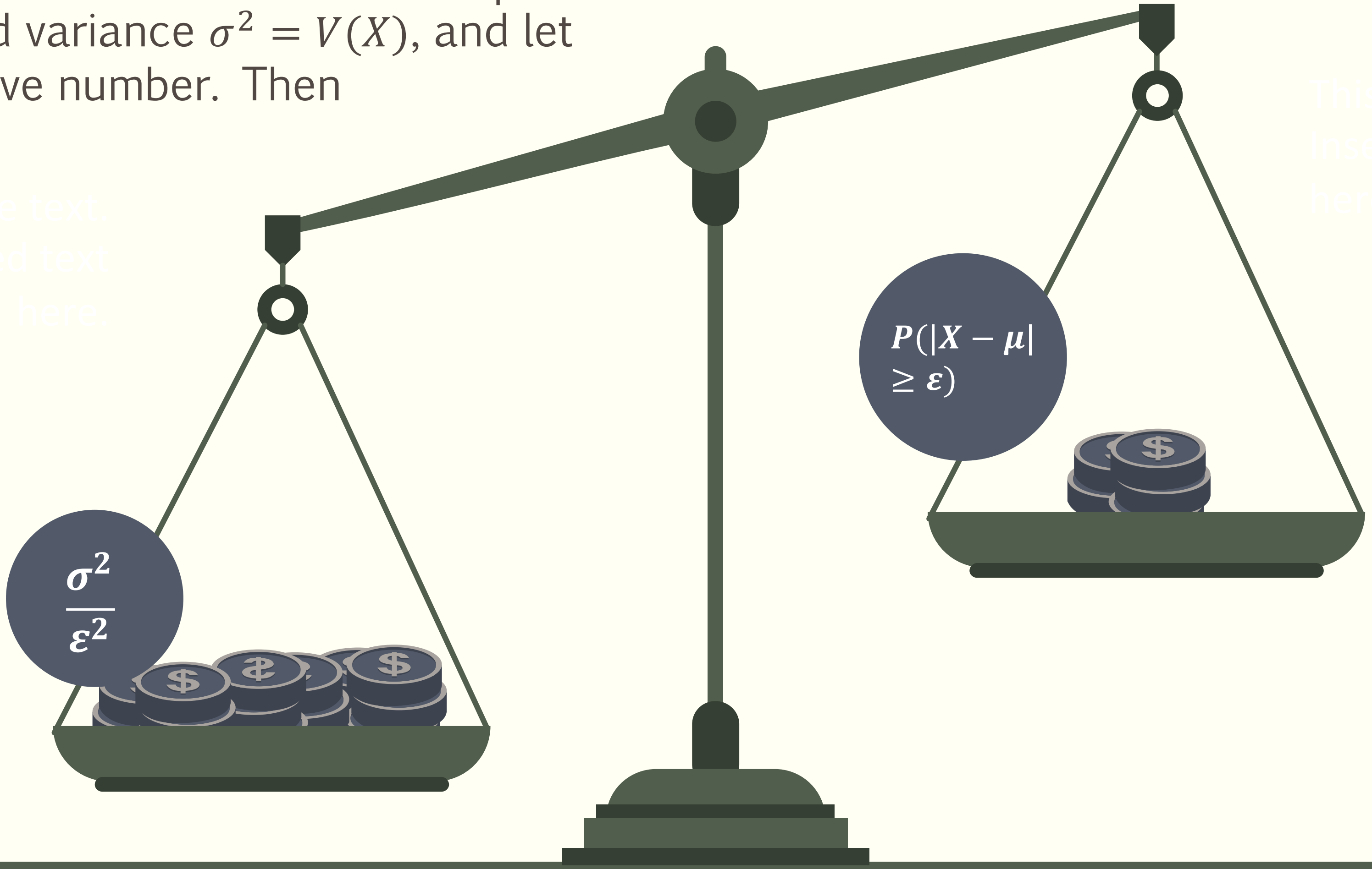
Chebyshev inequality

CLT

$$S_n^* = \frac{S_n - n\mu}{\sqrt{n\sigma^2}}$$

Let X be a discrete random variable with expected value $\mu = E(X)$ and variance $\sigma^2 = V(X)$, and let $\varepsilon > 0$ be any positive number. Then

not necessarily positive



Chebyshev inequality

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

Example

- How many measurements should he make to be reasonably sure that his average lies within 0.0001 of the true value?

$\mu = 1$

$\sigma^2 = (0.0002)^2$

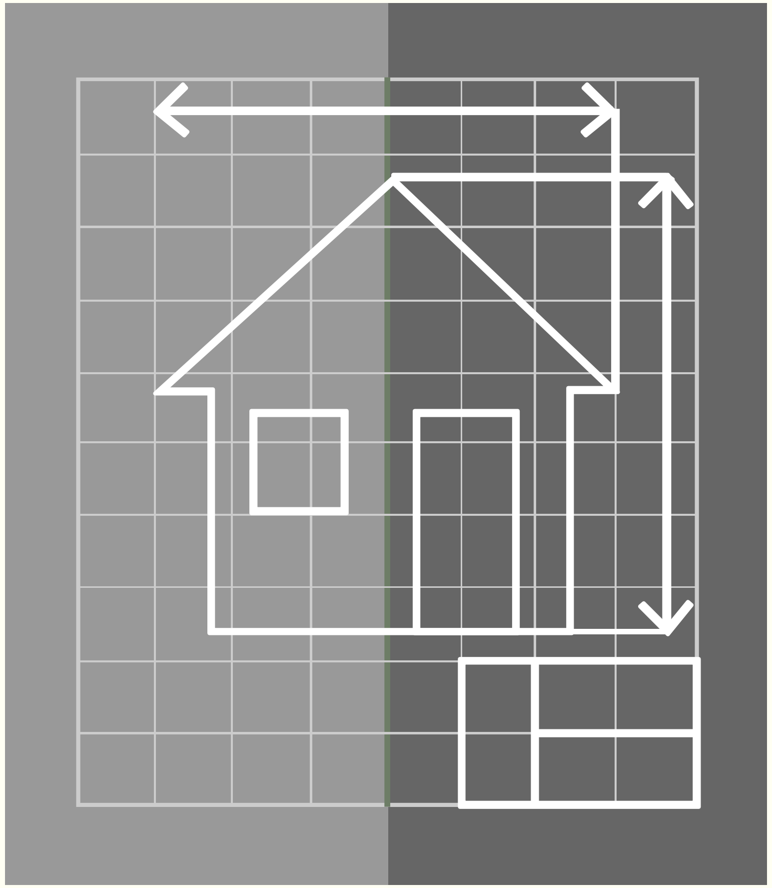
LLN

Chebyshev inequality

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

=

$$P(|A_n-\mu| \geq 0.0001) \leq \frac{\sigma^2}{n\varepsilon^2} = \frac{(0.0002)^2}{n(0.0001)^2} = \frac{4}{n}$$



Example

- How many measurements should he make to be reasonably sure that his average lies within 0.0001 of the true value?

LLN

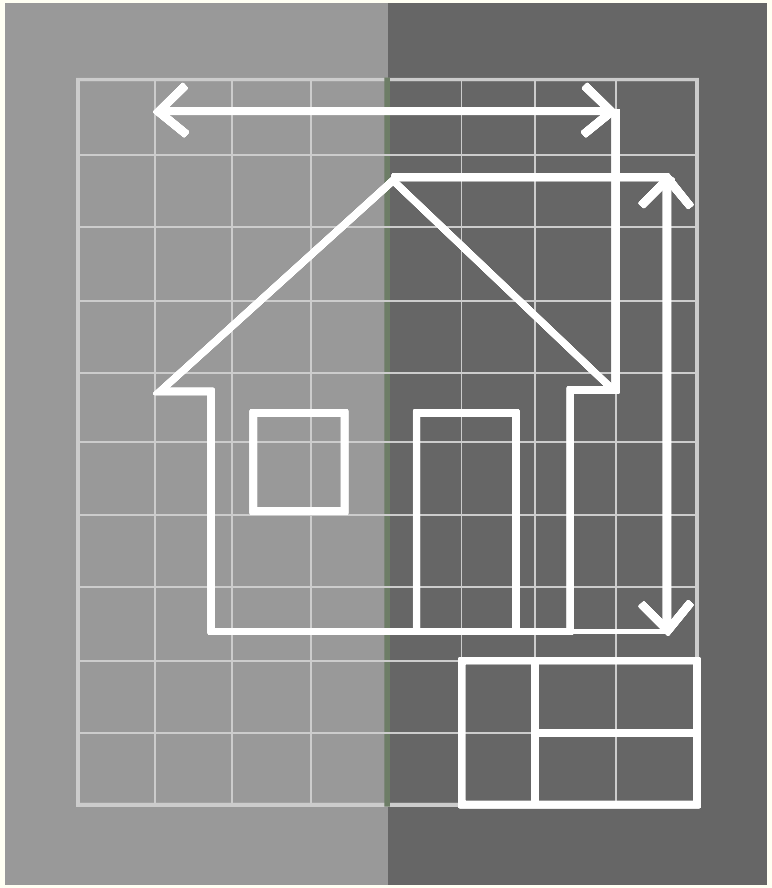
Chebyshev inequality

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

=

$$P(|A_n-\mu| \geq 0.0001) \leq \frac{\sigma^2}{n\varepsilon^2} = \frac{4}{n} = 0.05$$

$n = 80$



Example

$\mu = 1$

$\sigma^2 = (0.0002)^2$

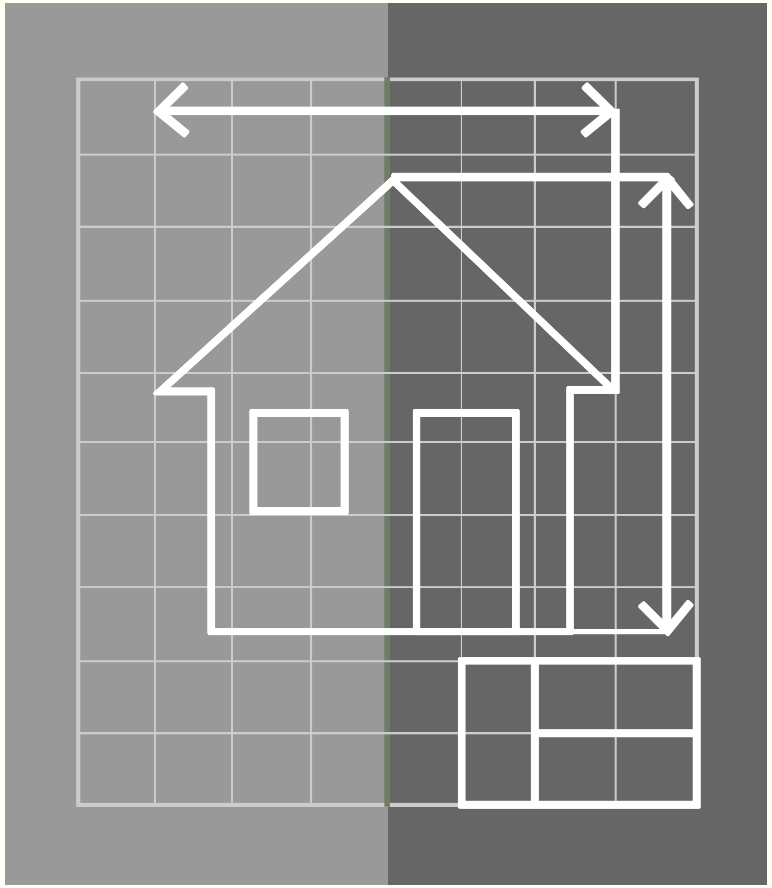
CLT

$$S_n^* = \frac{S_n - n\mu}{\sqrt{n\sigma^2}}$$

$$\lim_{n\to+\infty} P(a \le S_n^* \le b) = \mathrm{NA}(a,b)$$

=

$$P(|A_n - \mu| \le 0.0001) = P(|S_n - n\mu| \le 0.5n\sigma)$$
$$= P\left(\left|\frac{S_n - n\mu}{\sqrt{n\sigma^2}}\right| \le 0.5\sqrt{n}\right) = P(-0.5\sqrt{n} \le S_n^* \le 0.5\sqrt{n})$$
$$\approx \mathrm{NA}(-0.5\sqrt{n}, 0.5\sqrt{n})$$



Example

CLT

$$S_n^* = \frac{S_n - n\mu}{\sqrt{n\sigma^2}}$$

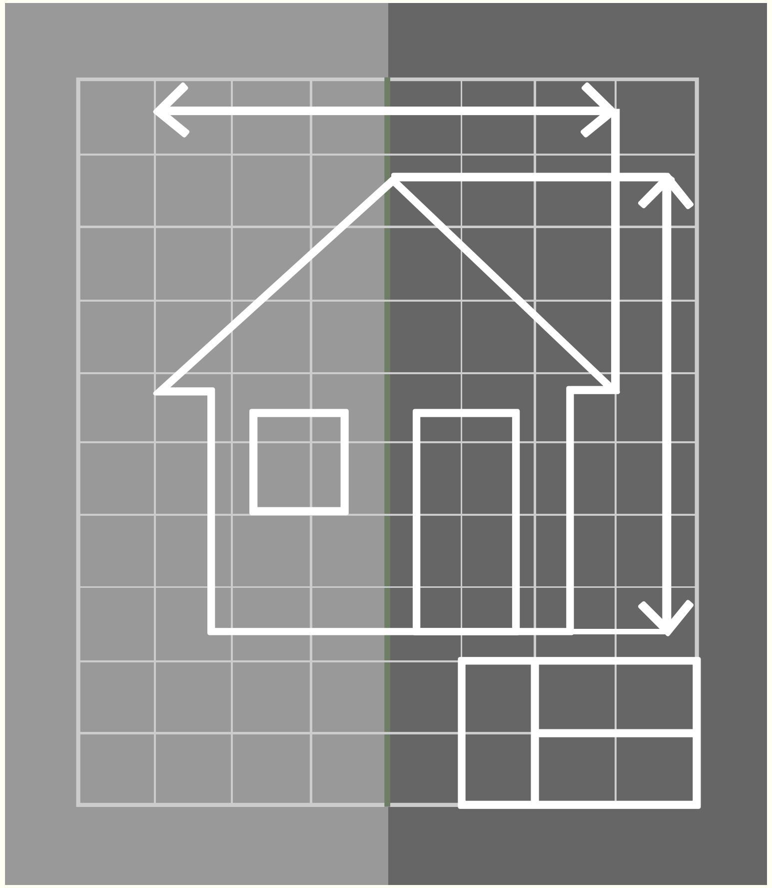
$$\lim_{n\to+\infty} P(a \le S_n^* \le b) = \mathrm{NA}(a,b)$$

=

$$P(|A_n - \mu| \le 0.0001) \approx \mathrm{NA}(-0.5\sqrt{n}, 0.5\sqrt{n}) = 0.95$$

$$0.5\sqrt{n} = 2$$

$$n = 16$$



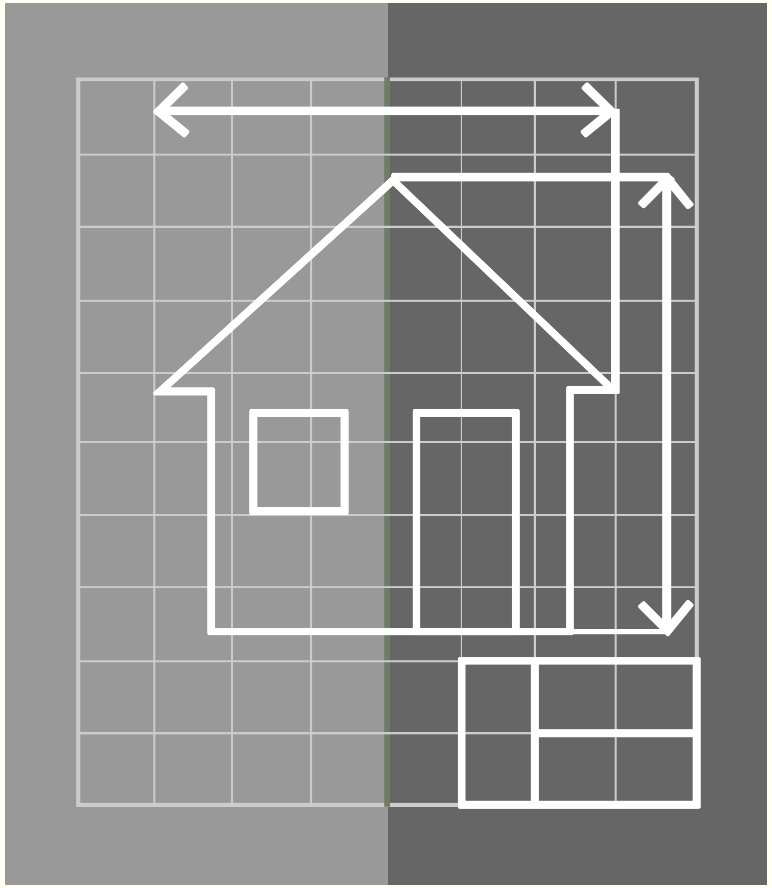
Example

- How many measurements should he make to be reasonably sure that his average lies within 0.0001 of the true value?

LLN

Chebyshev inequality

$n = 80$



CLT

$$S_n^* = \frac{S_n - n\mu}{\sqrt{n\sigma^2}}$$

$n = 16$