University of BRISTOL

Limit laws and continuous random variables

Henry W J Reeve

henry.reeve@bristol.ac.uk

Statistical Computing & Empirical Methods (EMATM0061)

MSc in Data Science, Teaching block 1, 2021.

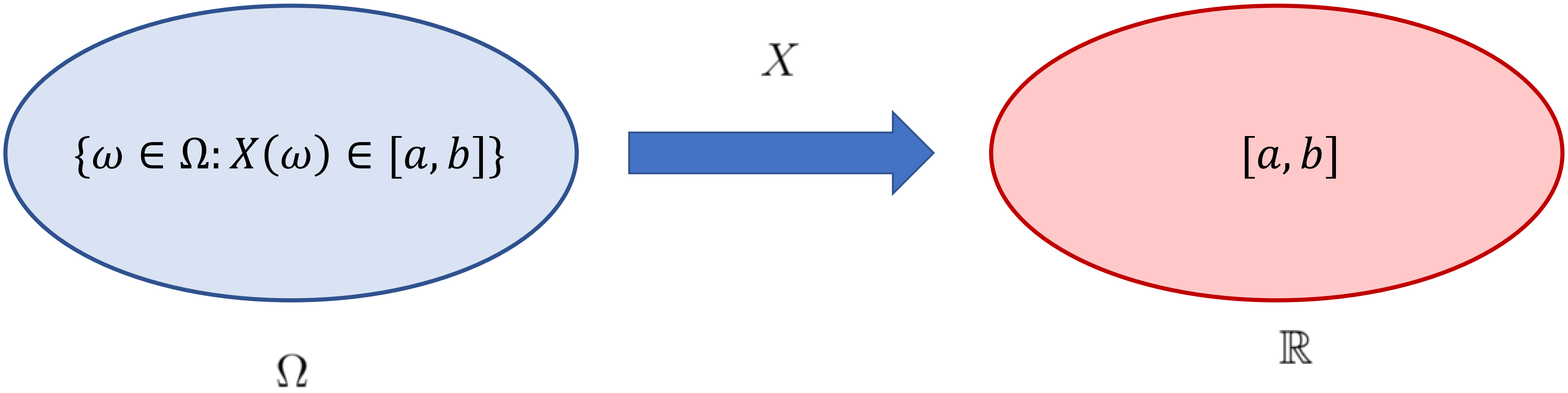
What will we cover today?

- We will introduce the concept of a continuous random variable.
- We will see continuous random variables can be understood via the probability density function.
- We will discuss expectation, variance, standard deviation, covariance and correlation in this context.
- We introduce the important example of Gaussian random variable.
- We will discuss the law of large numbers and the central limit theorem.

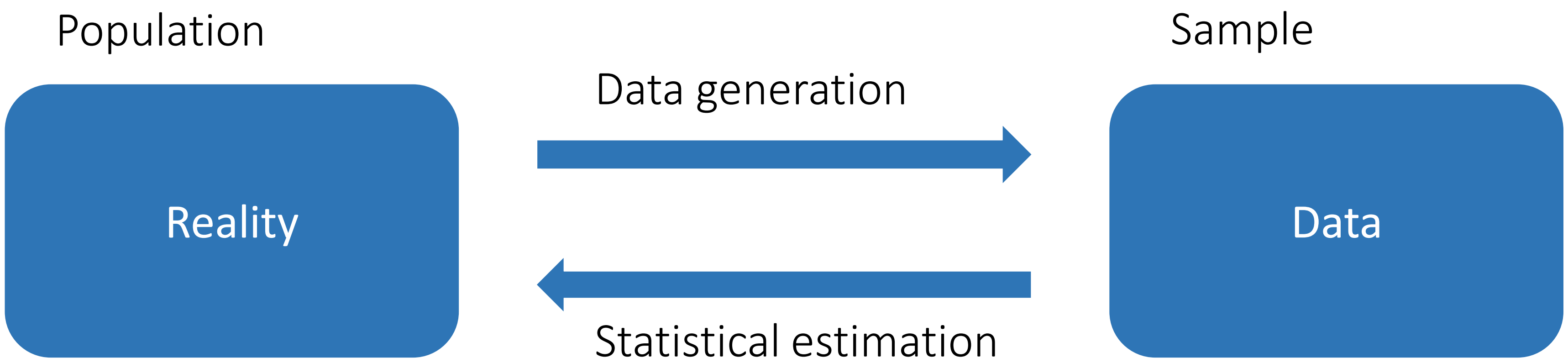
Random variables

Suppose we have a probability space $(\Omega, \mathcal{E}, \mathbb{P})$.

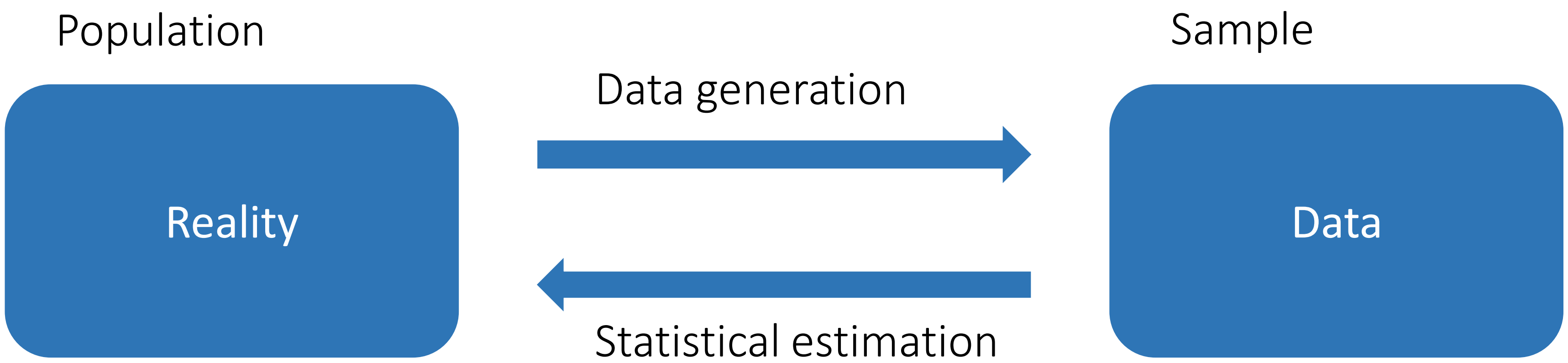
A **random variable** is a mapping $X : \Omega \to \mathbb{R}$ such that for every a, $b \in \mathbb{R}$, $\{\omega \in \Omega \;:\; X(\omega) \in [a, b]\} \in \mathcal{E}$ is an event.



Statistical estimation and probability



Statistical estimation and probability



A distribution P_X for a random variable X

A sample of independent copies $X_1, \ldots, X_n \sim P_X$.

Independent and identically distributed random variables

We model reality via the distribution P_X of a random variable $X : \Omega \to \mathbb{R}$.

Independent and identically distributed random variables

We model reality via the distribution P_X of a random variable $X : \Omega \to \mathbb{R}$.

Recall that $P_X(S) = \mathbb{P}(X \in S)$ for $S \in \mathfrak{B}(\mathbb{R})$ and $F_X(x) = \mathbb{P}(X \leq x)$ for $x \in \mathbb{R}$.

Independent and identically distributed random variables

We model reality via the distribution P_X of a random variable $X : \Omega \to \mathbb{R}$.

Recall that $P_X(S) = \mathbb{P}(X \in S)$ for $S \in \mathfrak{B}(\mathbb{R})$ and $F_X(x) = \mathbb{P}(X \leq x)$ for $x \in \mathbb{R}$.

We model our samples as a sequence of random variables $X_1, \ldots, X_n$.

Independent and identically distributed random variables

We model reality via the distribution P_X of a random variable $X : \Omega \to \mathbb{R}$.

Recall that $P_X(S) = \mathbb{P}(X \in S)$ for $S \in \mathfrak{B}(\mathbb{R})$ and $F_X(x) = \mathbb{P}(X \leq x)$ for $x \in \mathbb{R}$.

We model our samples as a sequence of random variables $X_1, \ldots, X_n$.

We say that $X_1, \ldots, X_n$ are **independent and identically distributed** **(i.i.d.)** if:

1. The sequence $X_1, \ldots, X_n$ consists of mutually independent random variables;
2. For all $x \in \mathbb{R}$ we have $F_{X_1}(x) = F_{X_2}(x) = \ldots = F_{X_n}(x)$.

Independent and identically distributed random variables

We model reality via the distribution P_X of a random variable $X : \Omega \to \mathbb{R}$.

Recall that $P_X(S) = \mathbb{P}(X \in S)$ for $S \in \mathfrak{B}(\mathbb{R})$ and $F_X(x) = \mathbb{P}(X \leq x)$ for $x \in \mathbb{R}$.

We model our samples as a sequence of random variables $X_1, \ldots, X_n$.

We say that $X_1, \ldots, X_n$ are **independent and identically distributed** **(i.i.d.)** if:

1. The sequence $X_1, \ldots, X_n$ consists of mutually independent random variables;
2. For all $x \in \mathbb{R}$ we have $F_{X_1}(x) = F_{X_2}(x) = \ldots = F_{X_n}(x)$.

Note that 2. is equivalent to $P_{X_1} = P_{X_2} = \ldots = P_{X_n}$.

Independent and identically distributed random variables

We model reality via the distribution P_X of a random variable $X : \Omega \to \mathbb{R}$.

Recall that $P_X(S) = \mathbb{P}(X \in S)$ for $S \in \mathfrak{B}(\mathbb{R})$ and $F_X(x) = \mathbb{P}(X \leq x)$ for $x \in \mathbb{R}$.

We model our samples as a sequence of random variables $X_1, \ldots, X_n$.

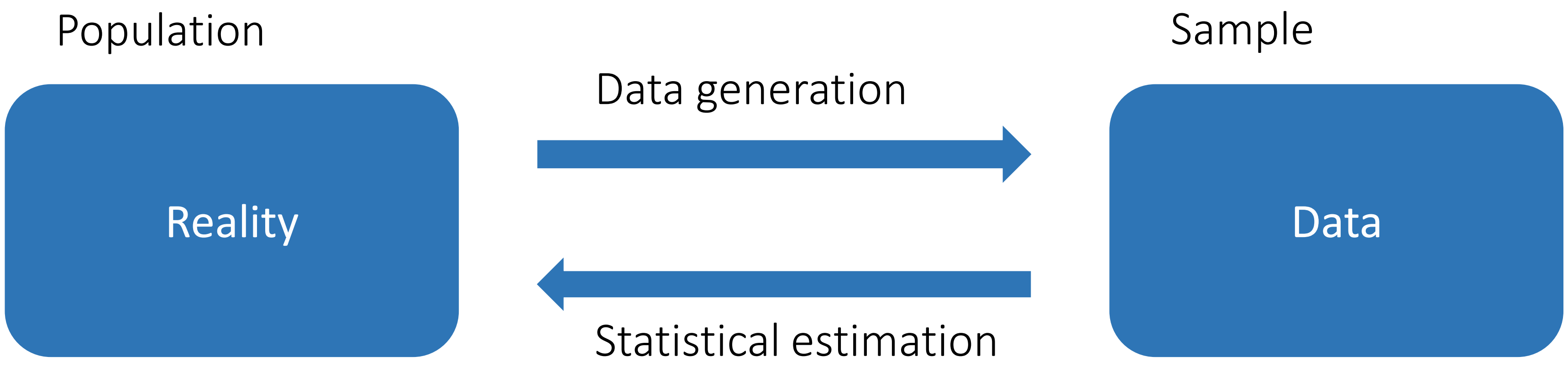
We say that $X_1, \ldots, X_n$ are **independent and identically distributed (i.i.d.)** if:

1. The sequence $X_1, \ldots, X_n$ consists of mutually independent random variables;
2. For all $x \in \mathbb{R}$ we have $F_{X_1}(x) = F_{X_2}(x) = \ldots = F_{X_n}(x)$.

Note that 2. is equivalent to $P_{X_1} = P_{X_2} = \ldots = P_{X_n}$.

If $P_{X_1} = P_{X_2} = \ldots = P_{X_n} = P_X$ then we refer to $X_1, \ldots, X_n$ as **independent copies** of X.

Statistical estimation and probability



A distribution P_X for a random variable X

A sample of independent copies $X_1, \ldots, X_n \sim P_X$.

The law of large numbers

Sequences of i.i.d. random variables are surprisingly well-behaved.

The law of large numbers

Sequences of i.i.d. random variables are surprisingly well-behaved.

Given the roll of a single fair dice the outcome X is highly unpredictable.

The law of large numbers

Sequences of i.i.d. random variables are surprisingly well-behaved.

Given the roll of a single fair dice the outcome X is highly unpredictable.

However, suppose we take a very large number of independent dolls of a fair dice rolls $X_1, \ldots, X_n$, and then compute the sample average of the sequence $\frac{1}{n}\sum_{i=1} X_i$.

The law of large numbers

Sequences of i.i.d. random variables are surprisingly well-behaved.

Given the roll of a single fair dice the outcome X is highly unpredictable.

However, suppose we take a very large number of independent dolls of a fair dice rolls $X_1, \ldots, X_n$, and then compute the sample average of the sequence $\frac{1}{n}\sum_{i=1} X_i$.

Then we expect the sample average $\frac{1}{n}\sum_{i=1} X_i$ to be close to $\mathbb{E}(X) = \frac{7}{2}$.

The law of large numbers

Sequences of i.i.d. random variables are surprisingly well-behaved.

Given the roll of a single fair dice the outcome X is highly unpredictable.

However, suppose we take a very large number of independent dolls of a fair dice rolls $X_1, \ldots, X_n$, and then compute the sample average of the sequence $\frac{1}{n}\sum_{i=1} X_i$.

Then we expect the sample average $\frac{1}{n}\sum_{i=1} X_i$ to be close to $\mathbb{E}(X) = \frac{7}{2}$.

This is an example of a law of large numbers.

The sample average of dice rolls

```r
num_trials<-1000000 # set the number of trials
set.seed(0) # set the random seed
sample_size<-5 # set the sample size
```

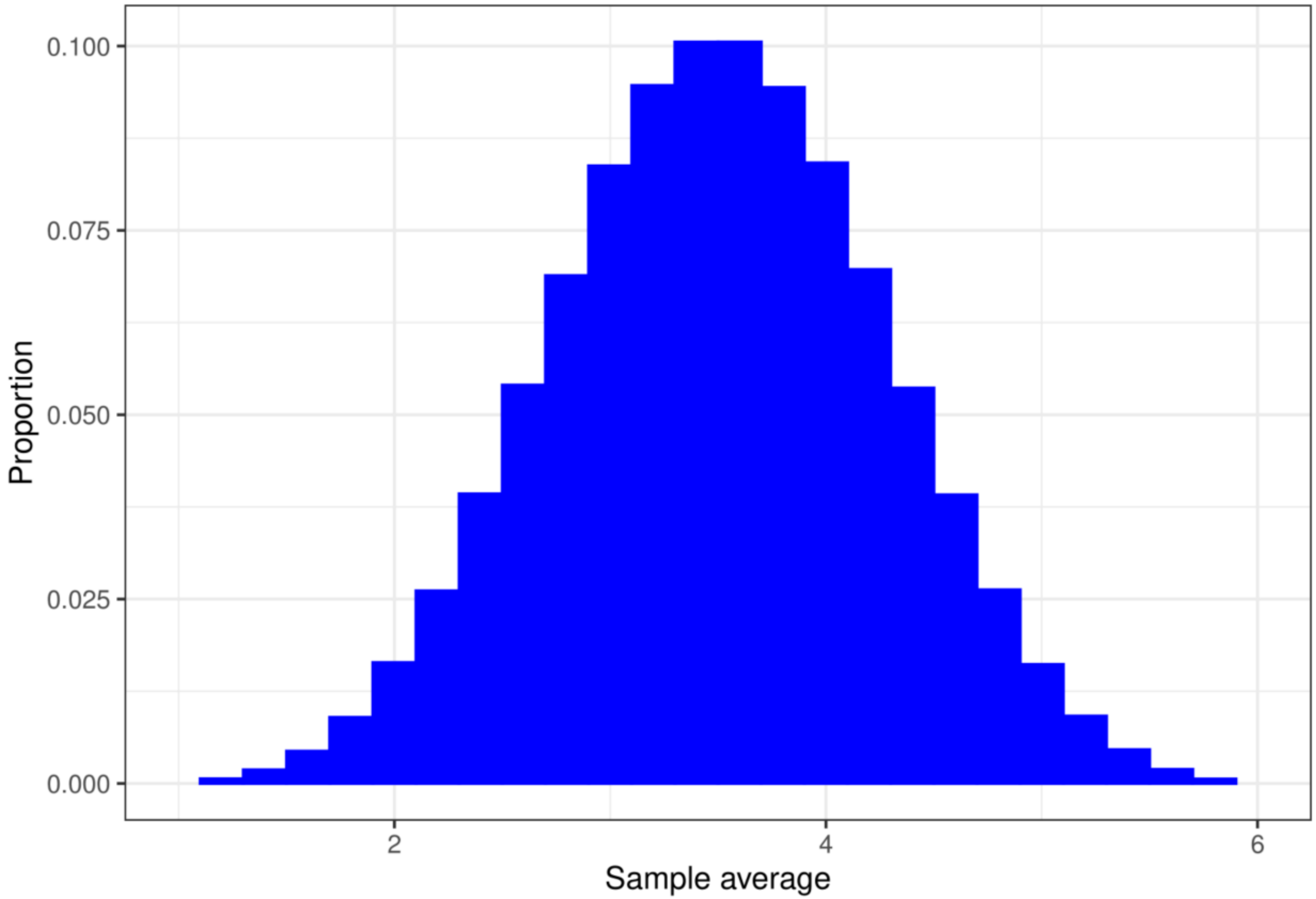
The sample average of dice rolls

```r
num_trials<-1000000 # set the number of trials
set.seed(0) # set the random seed
sample_size<-5 # set the sample size

# simulate the sample average of a dice roll
dice_sample_average_simulation<-data.frame(trial=1:num_trials)%>%
  mutate(dice_sample=map(.x=trial,~sample(6,sample_size,replace=TRUE)))%>%
  mutate(sample_avg=map_dbl(.x=dice_sample,~mean(.x)))
```

The sample average of dice rolls

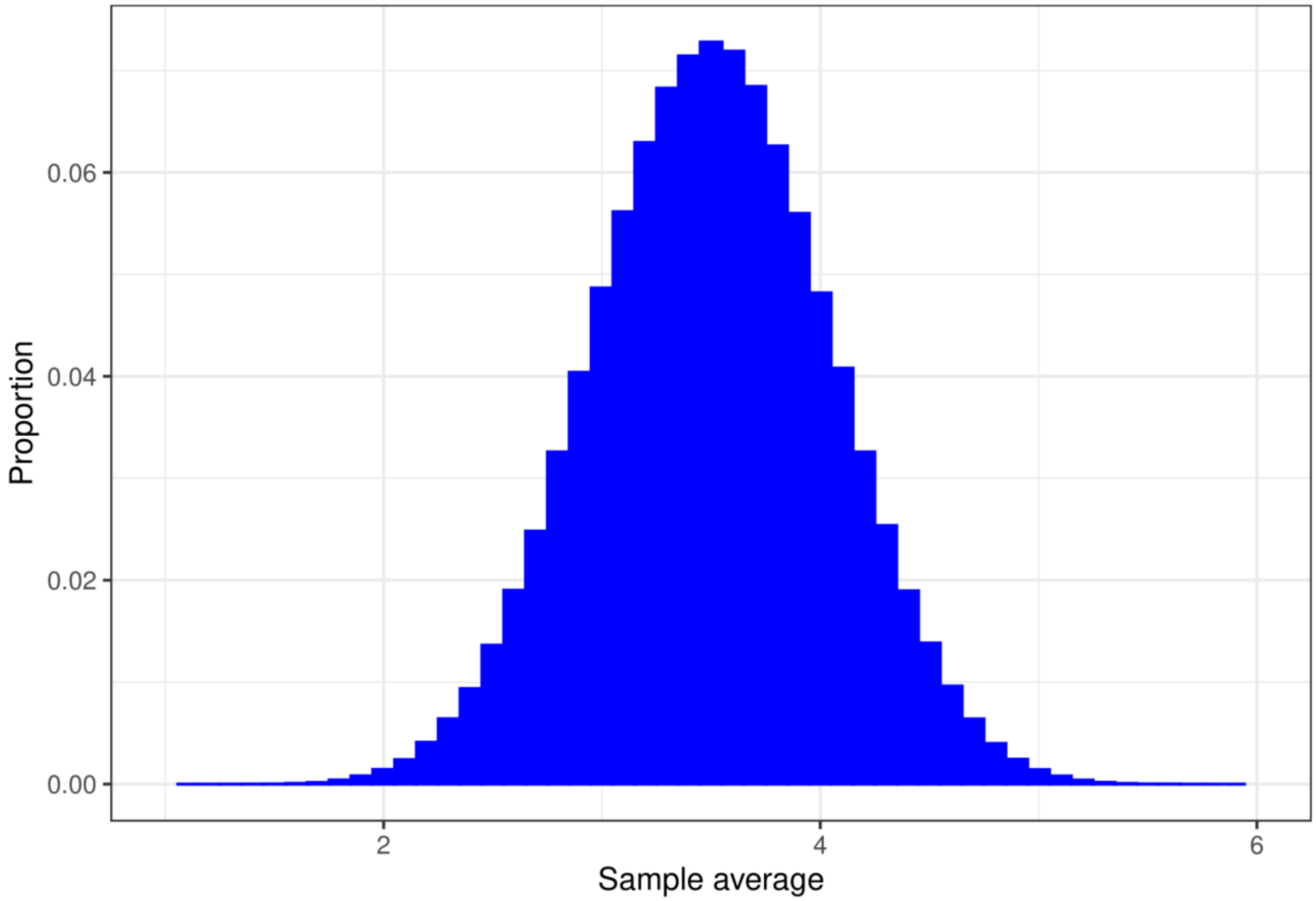
```
num_trials<-1000000 # set the number of trials
set.seed(0) # set the random seed
sample_size<-5 # set the sample size

# simulate the sample average of a dice roll
dice_sample_average_simulation<-data.frame(trial=1:num_trials)%>%
  mutate(dice_sample=map(.x=trial,~sample(6,sample_size,replace=TRUE)))%>%
  mutate(sample_avg=map_dbl(.x=dice_sample,~mean(.x)))

# plot a histogram to display the results
dice_sample_average_simulation%>%
  ggplot(aes(x=sample_avg))+
  geom_histogram(aes(y=..count../sum(..count..)),
                 binwidth=1/sample_size,fill="blue",color="blue")+
  theme_bw()+
  xlim(c(1,6))+
  xlab("Sample average")+ylab("Proportion")
```

The sample average of dice rolls



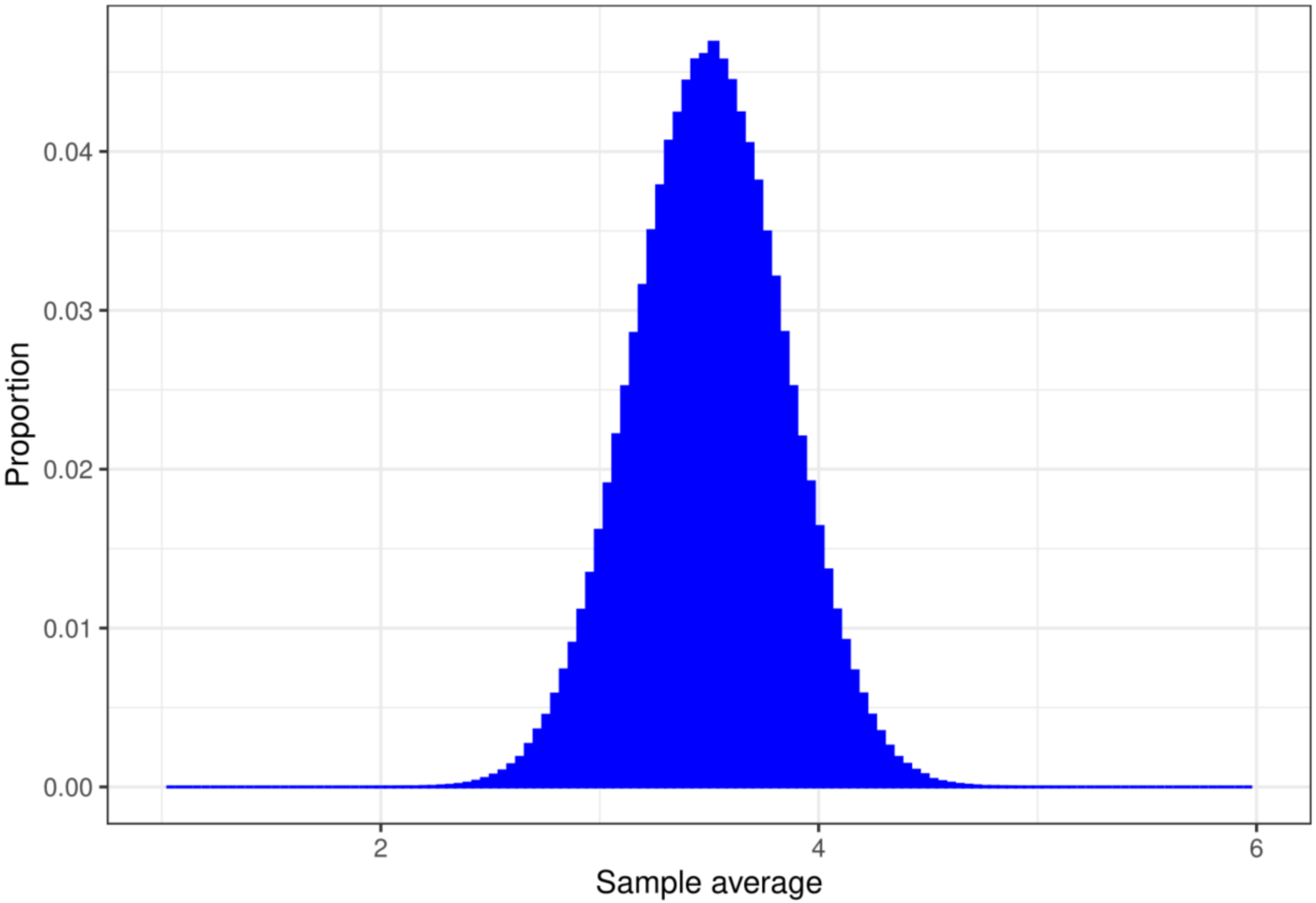
$$n = 5$$

The sample average of dice rolls



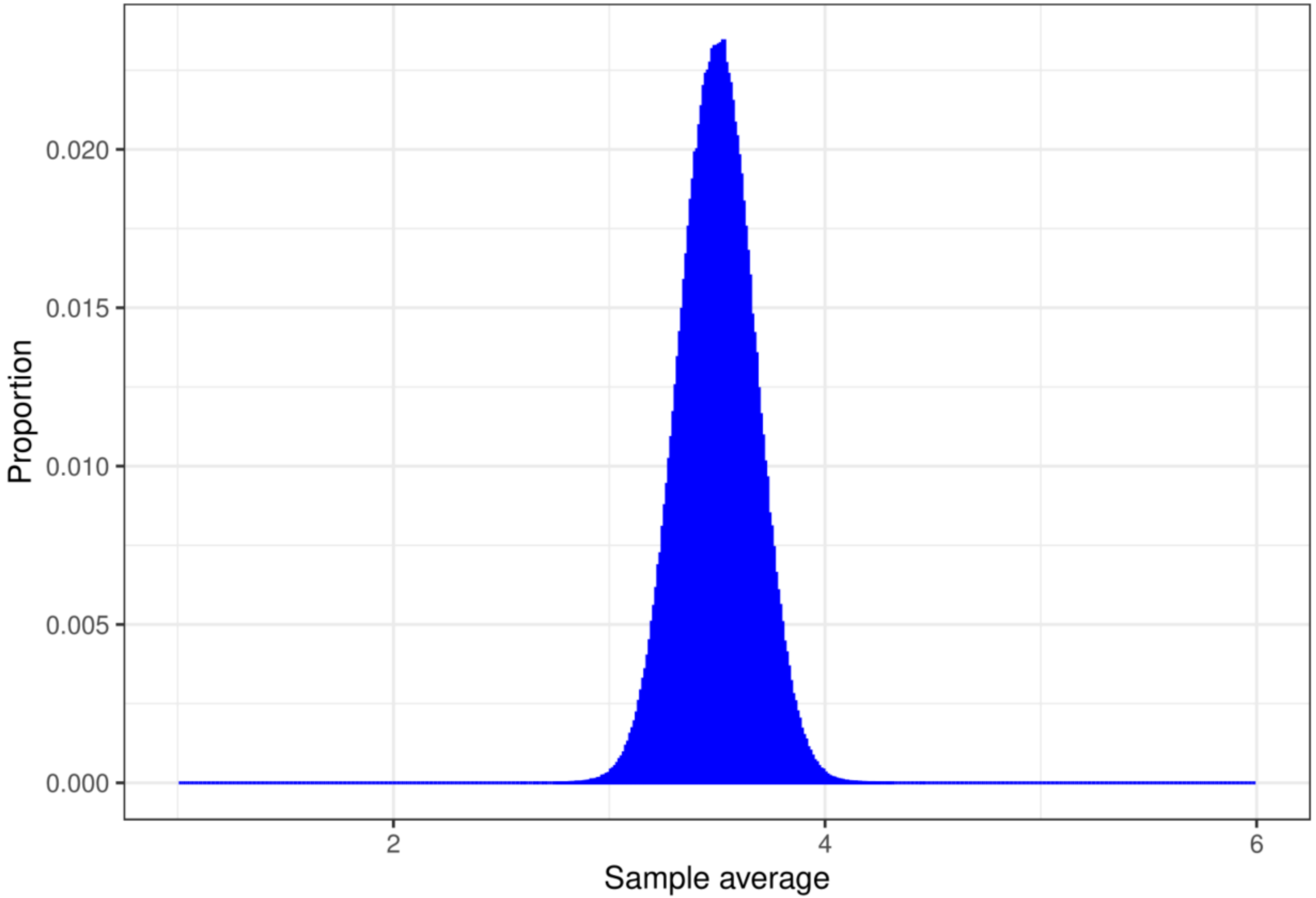
$n = 10$

The sample average of dice rolls



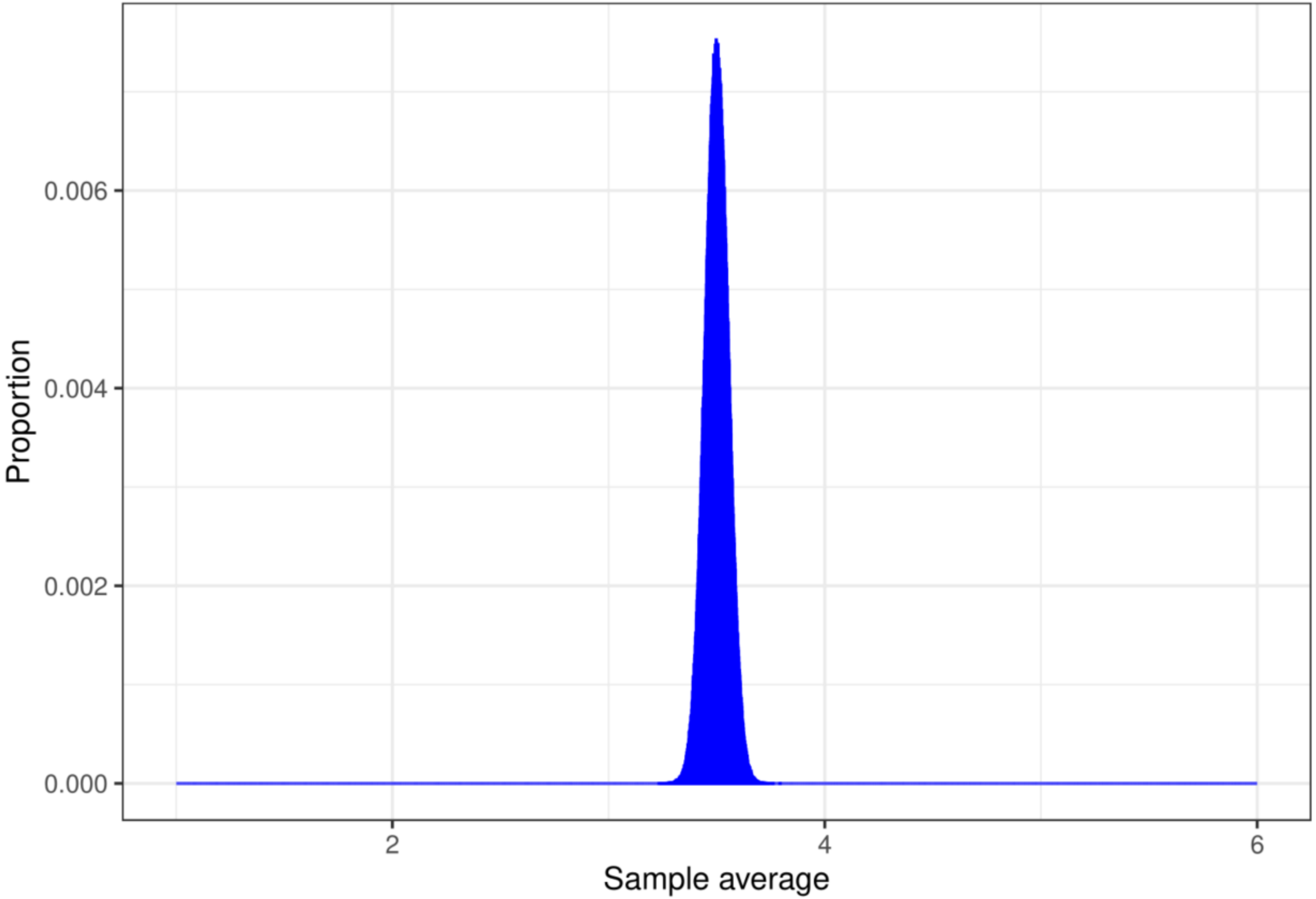
$$n = 25$$

The sample average of dice rolls



$n = 100$

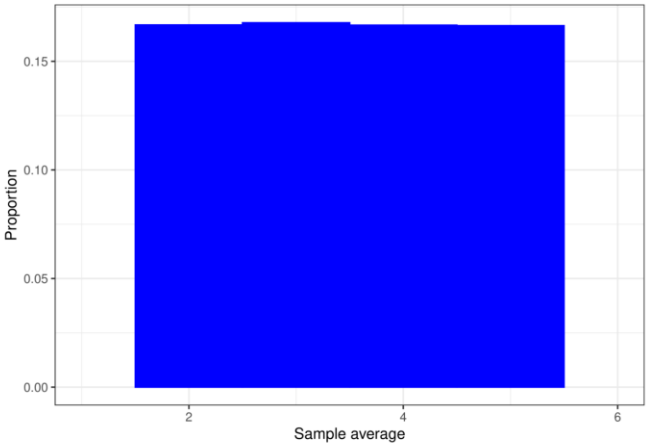
The sample average of dice rolls



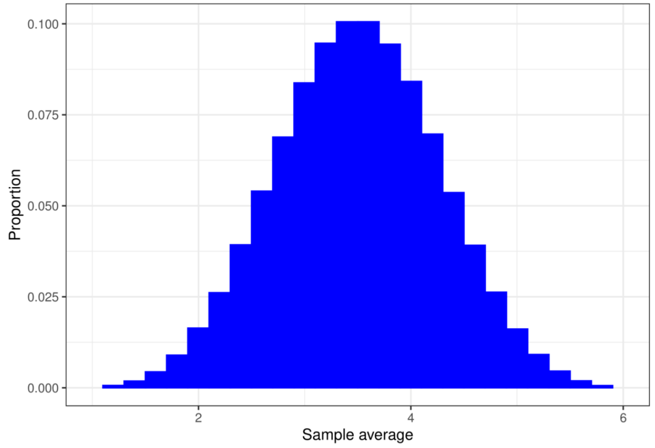
$$n = 1000$$

The sample average of dice rolls

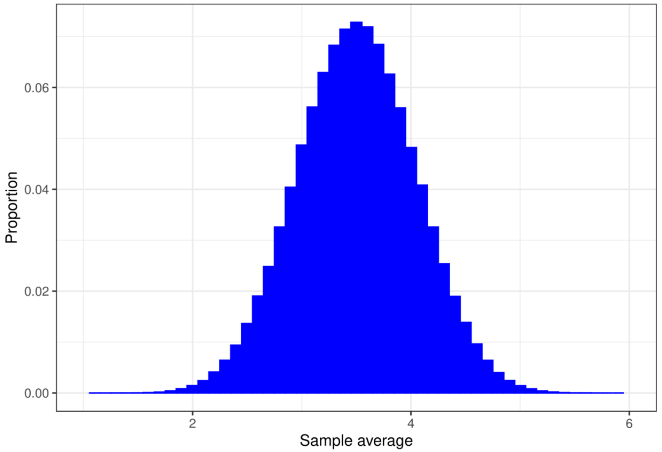
$n = 1$



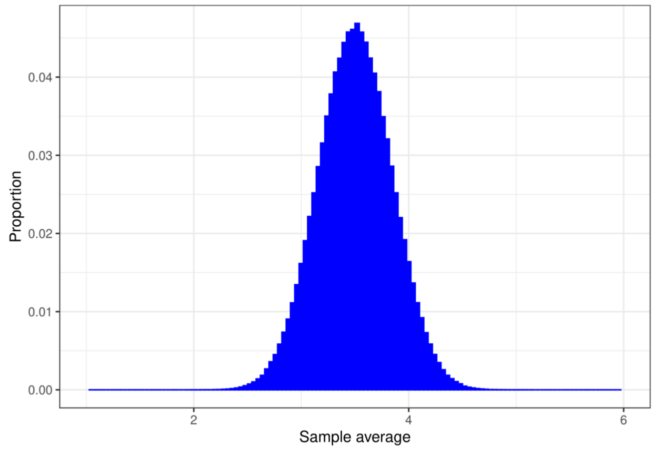
$n = 5$



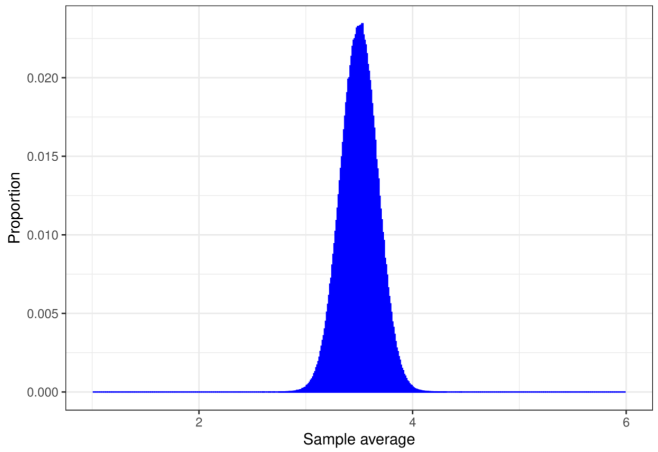
$n = 10$



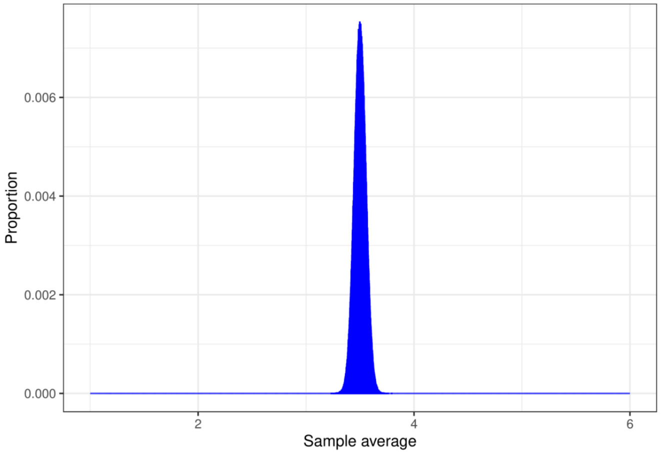
$n = 25$



$n = 100$



$n = 1000$



The law of large numbers

The law of large numbers tells us that the sample average converges towards the expectation, for sequences of independent and identically distributed random variables.

The law of large numbers

The law of large numbers tells us that the sample average converges towards the expectation, for sequences of independent and identically distributed random variables.

Theorem (Bernoulli, circa. 1700). *Let* $X : \Omega \to \mathbb{R}$ *be a random variable with a well-defined expectation* $\mu = \mathbb{E}(X)$. *Let* $X_1, \ldots, X_n : \Omega \to \mathbb{R}$ *be a sequence of independent copies of* X. *Then for all* $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left(\left| \frac{1}{n} \sum_{i=1}^{n} X_i - \mu \right| \geq \epsilon \right) = 0.$$

The law of large numbers

The law of large numbers tells us that the sample average converges towards the expectation, for sequences of independent and identically distributed random variables.

Theorem (Bernoulli, circa. 1700). *Let* $X : \Omega \to \mathbb{R}$ *be a random variable with a well-defined expectation* $\mu = \mathbb{E}(X)$. *Let* $X_1, \ldots, X_n : \Omega \to \mathbb{R}$ *be a sequence of independent copies of* X. *Then for all* $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left(\left| \frac{1}{n} \sum_{i=1}^{n} X_i - \mu \right| \geq \epsilon \right) = 0.$$

This result is known as **the weak law of large numbers**.

The law of large numbers

Theorem (Bernoulli, circa. 1700). *Let* $X : \Omega \to \mathbb{R}$ *be a random variable with a well-defined expectation* $\mu = \mathbb{E}(X)$. *Let* $X_1, \ldots, X_n : \Omega \to \mathbb{R}$ *be a sequence of independent copies of* X. *Then for all* $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \epsilon\right) = 0.$$

Intuition

Suppose X has expectation $\mathbb{E}(X) = \mu$ and variance $\mathrm{Var}(X) = \sigma^2$,

The law of large numbers

Theorem (Bernoulli, circa. 1700). *Let* $X : \Omega \to \mathbb{R}$ *be a random variable with a well-defined expectation* $\mu = \mathbb{E}(X)$. *Let* $X_1, \ldots, X_n : \Omega \to \mathbb{R}$ *be a sequence of independent copies of* X. *Then for all* $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \epsilon\right) = 0.$$

Intuition

Suppose X has expectation $\mathbb{E}(X) = \mu$ and variance $\mathrm{Var}(X) = \sigma^2$,

Since $X_1, \ldots, X_n$ are independent copies of X we have $\mathbb{E}(X_i) = \mu$ and $\mathrm{Var}(X_i) = \sigma^2$.

The law of large numbers

Theorem (Bernoulli, circa. 1700). *Let* $X : \Omega \to \mathbb{R}$ *be a random variable with a well-defined expectation* $\mu = \mathbb{E}(X)$. *Let* $X_1, \ldots, X_n : \Omega \to \mathbb{R}$ *be a sequence of independent copies of* X. *Then for all* $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \epsilon\right) = 0.$$

Intuition

Suppose X has expectation $\mathbb{E}(X) = \mu$ and variance $\mathrm{Var}(X) = \sigma^2$,

Since $X_1, \ldots, X_n$ are independent copies of X we have $\mathbb{E}(X_i) = \mu$ and $\mathrm{Var}(X_i) = \sigma^2$.

1. $\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(X_i) = \frac{1}{n}\sum_{i=1}^{n} \mu = \mu$;

The law of large numbers

Theorem (Bernoulli, circa. 1700). *Let $X : \Omega \to \mathbb{R}$ be a random variable with a well-defined expectation $\mu = \mathbb{E}(X)$. Let $X_1, \ldots, X_n : \Omega \to \mathbb{R}$ be a sequence of independent copies of X. Then for all $\epsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \epsilon\right) = 0.$$

Intuition

Suppose X has expectation $\mathbb{E}(X) = \mu$ and variance $\mathrm{Var}(X) = \sigma^2$,

Since $X_1, \ldots, X_n$ are independent copies of X we have $\mathbb{E}(X_i) = \mu$ and $\mathrm{Var}(X_i) = \sigma^2$.

1. $\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(X_i) = \frac{1}{n}\sum_{i=1}^{n} \mu = \mu$;
2. $\mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \left(\frac{1}{n}\right)^2 \cdot \sum_{i=1}^{n} \mathrm{Var}(X_i) = \left(\frac{1}{n}\right)^2 \cdot (n \cdot \sigma^2) = \frac{\sigma^2}{n}$.

The law of large numbers

Theorem (Bernoulli, circa. 1700). *Let $X : \Omega \to \mathbb{R}$ be a random variable with a well-defined expectation $\mu = \mathbb{E}(X)$. Let $X_1, \ldots, X_n : \Omega \to \mathbb{R}$ be a sequence of independent copies of X. Then for all $\epsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}\left(\left| \frac{1}{n} \sum_{i=1}^{n} X_i - \mu \right| \geq \epsilon \right) = 0.$$

Intuition

Suppose X has expectation $\mathbb{E}(X) = \mu$ and variance $\mathrm{Var}(X) = \sigma^2$,

Since $X_1, \ldots, X_n$ are independent copies of X we have $\mathbb{E}(X_i) = \mu$ and $\mathrm{Var}(X_i) = \sigma^2$.

1. $\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(X_i) = \frac{1}{n}\sum_{i=1}^{n} \mu = \mu$;
2. $\mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \left(\frac{1}{n}\right)^2 \cdot \sum_{i=1}^{n} \mathrm{Var}(X_i) = \left(\frac{1}{n}\right)^2 \cdot (n \cdot \sigma^2) = \frac{\sigma^2}{n}$.

This suggests $\frac{1}{n}\sum_{i=1}^{n} X_i - \mu$ is a random variable which shrinks to zero as n goes to infinity.

The law of large numbers

Suppose X has expectation $\mathbb{E}(X) = \mu$ and variance $\mathrm{Var}(X) = \sigma^2$,

Since $X_1, \ldots, X_n$ are independent copies of X we have $\mathbb{E}(X_i) = \mu$ and $\mathrm{Var}(X_i) = \sigma^2$.

1. $\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(X_i) = \frac{1}{n}\sum_{i=1}^{n} \mu = \mu$;
2. $\mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \left(\frac{1}{n}\right)^2 \cdot \sum_{i=1}^{n} \mathrm{Var}(X_i) = \left(\frac{1}{n}\right)^2 \cdot (n \cdot \sigma^2) = \frac{\sigma^2}{n}$.

The law of large numbers implies that $\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right) \to 0$ as $n \to \infty$.

The law of large numbers

Suppose X has expectation $\mathbb{E}(X) = \mu$ and variance $\mathrm{Var}(X) = \sigma^2$,

Since $X_1, \ldots, X_n$ are independent copies of X we have $\mathbb{E}(X_i) = \mu$ and $\mathrm{Var}(X_i) = \sigma^2$.

1. $\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n X_i\right) = \frac{1}{n}\sum_{i=1}^n \mathbb{E}(X_i) = \frac{1}{n}\sum_{i=1}^n \mu = \mu$;
2. $\mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^n X_i\right) = \left(\frac{1}{n}\right)^2 \cdot \sum_{i=1}^n \mathrm{Var}(X_i) = \left(\frac{1}{n}\right)^2 \cdot (n \cdot \sigma^2) = \frac{\sigma^2}{n}$.

The law of large numbers implies that $\left(\frac{1}{n}\sum_{i=1}^n X_i - \mu\right) \to 0$ as $n \to \infty$.

To gain further insight we can zoom in and consider the renormalised deviations

$$G_n := \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n X_i - \mu\right),$$

The law of large numbers

Suppose X has expectation $\mathbb{E}(X) = \mu$ and variance $\mathrm{Var}(X) = \sigma^2$,

Since $X_1, \ldots, X_n$ are independent copies of X we have $\mathbb{E}(X_i) = \mu$ and $\mathrm{Var}(X_i) = \sigma^2$.

1. $\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n X_i\right) = \frac{1}{n}\sum_{i=1}^n \mathbb{E}(X_i) = \frac{1}{n}\sum_{i=1}^n \mu = \mu$;
2. $\mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^n X_i\right) = \left(\frac{1}{n}\right)^2 \cdot \sum_{i=1}^n \mathrm{Var}(X_i) = \left(\frac{1}{n}\right)^2 \cdot (n \cdot \sigma^2) = \frac{\sigma^2}{n}$.

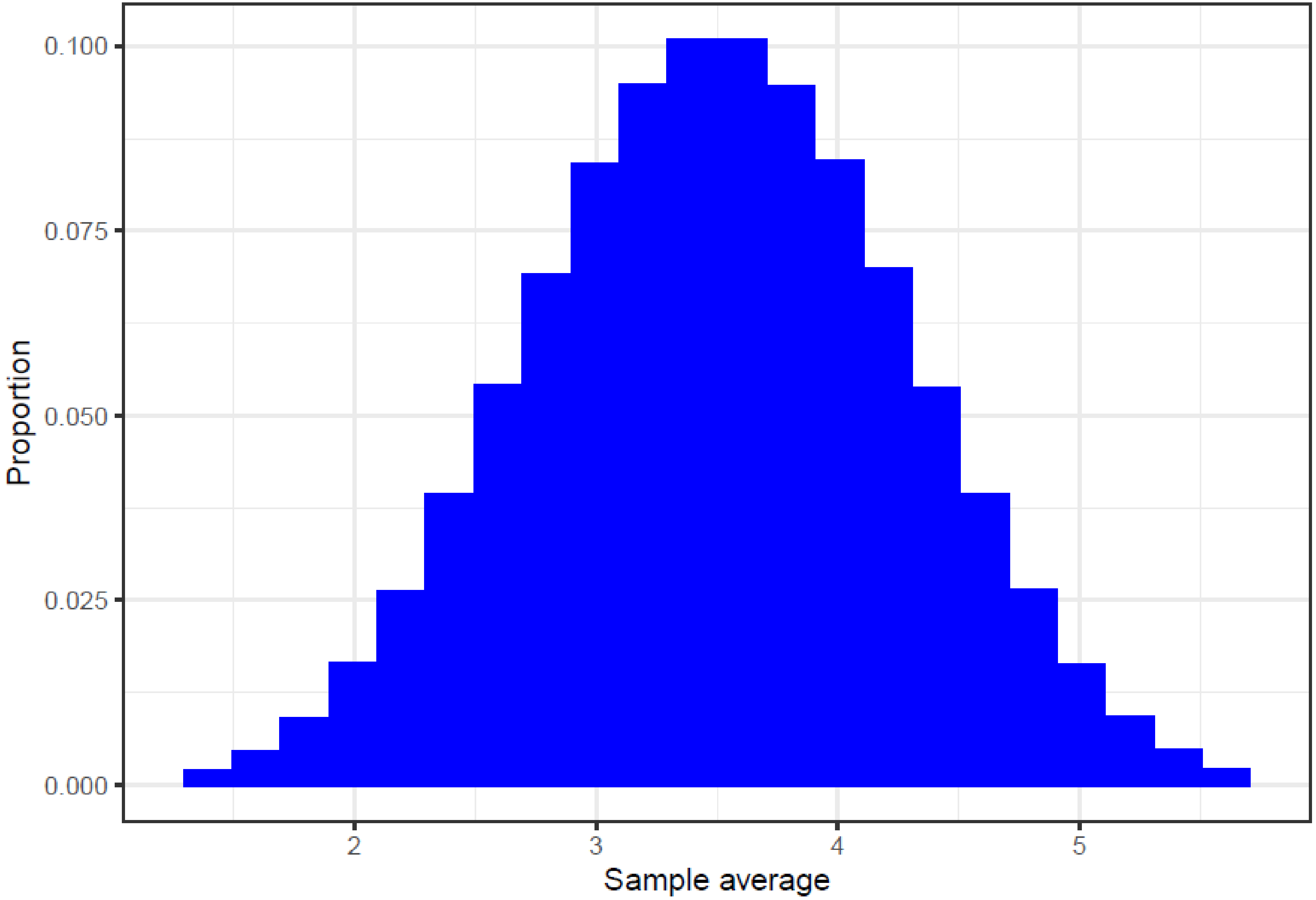
The law of large numbers implies that $\left(\frac{1}{n}\sum_{i=1}^n X_i - \mu\right) \to 0$ as $n \to \infty$.

To gain further insight we can zoom in and consider the renormalised deviations

$$G_n := \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n X_i - \mu\right),$$

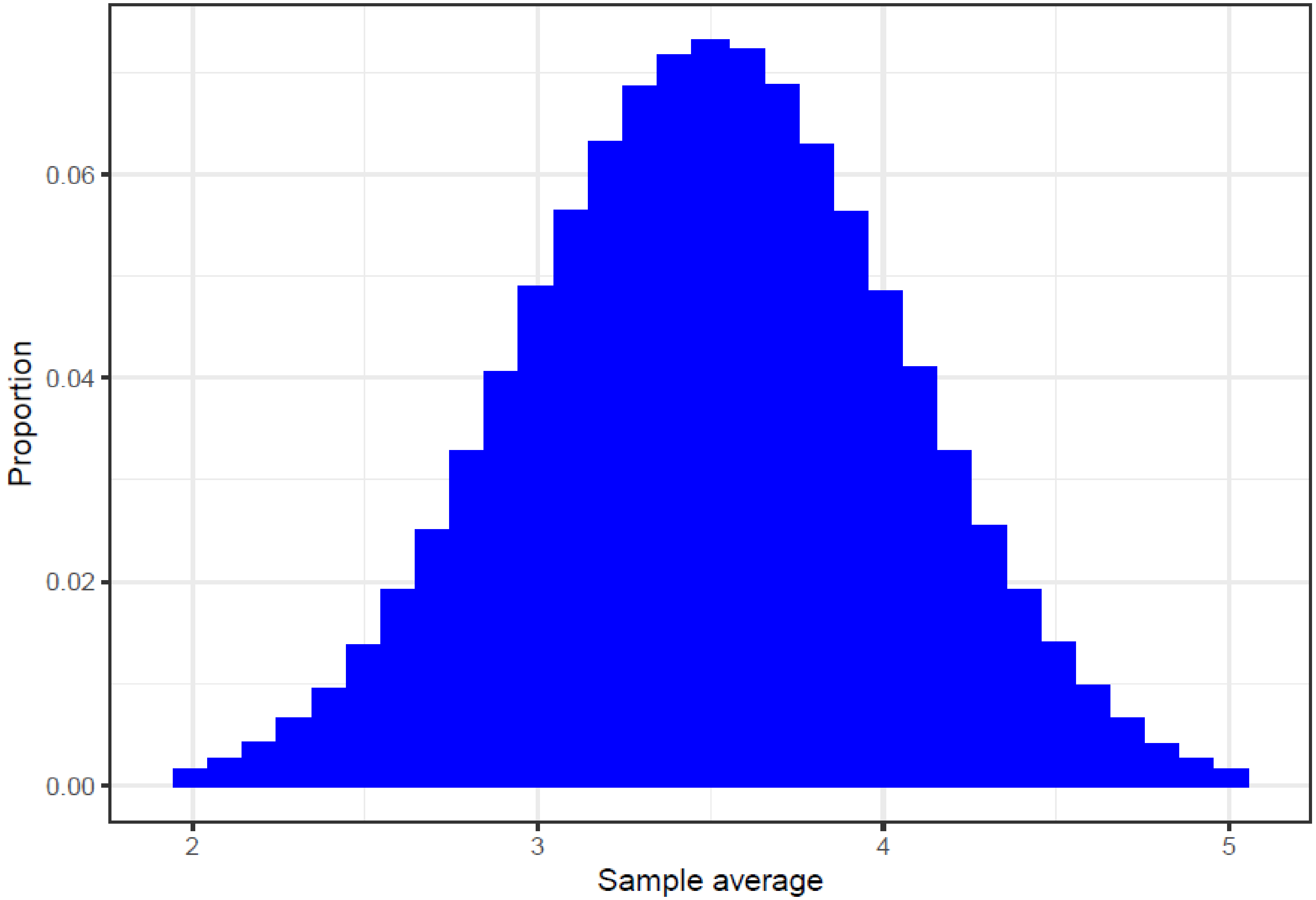
Note that $\mathrm{Var}(G_n) = \sigma^2$ for all $n \in \mathbb{N}$.

The sample average of dice rolls (renormalized)



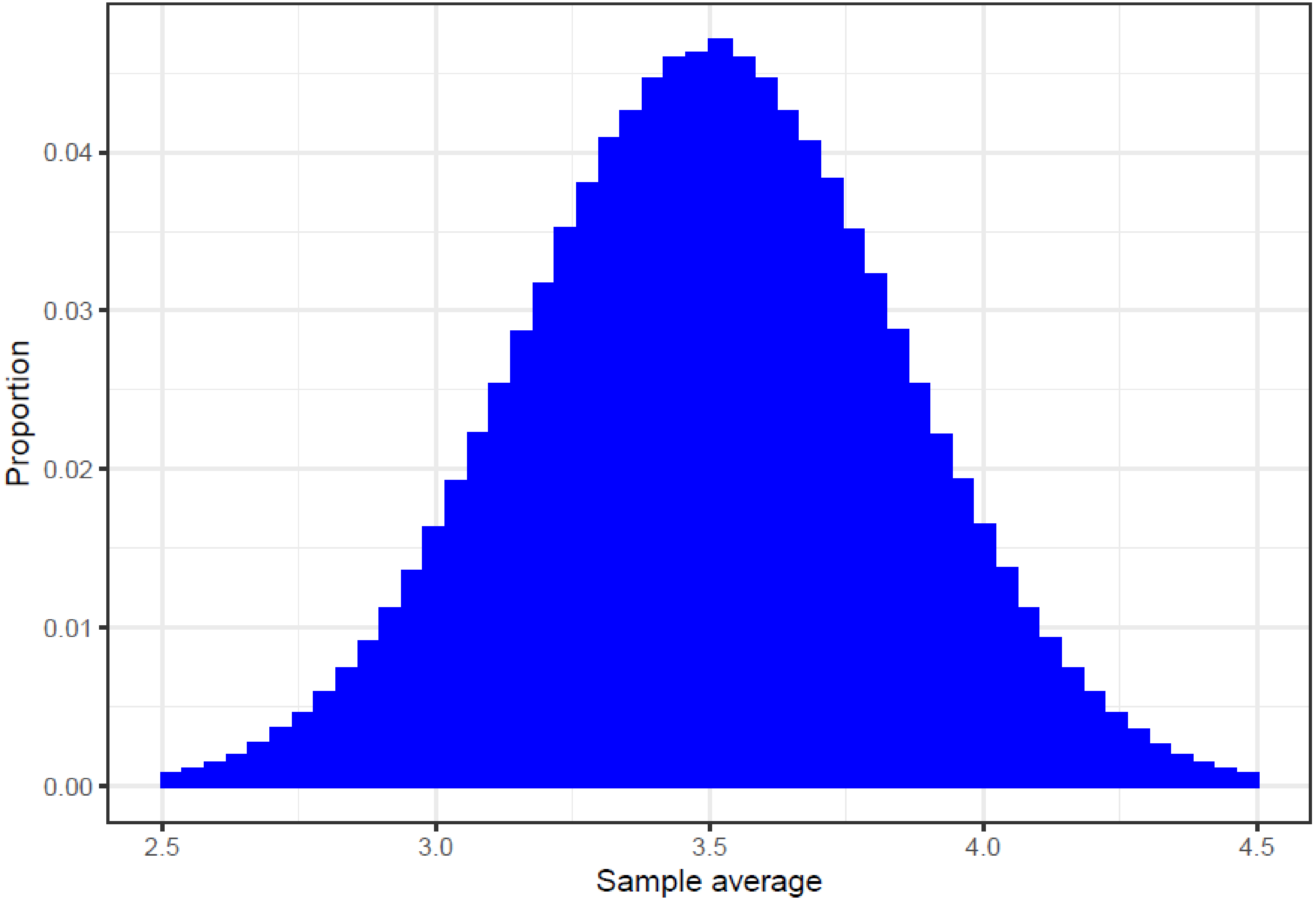
$n = 5$

The sample average of dice rolls (renormalized)



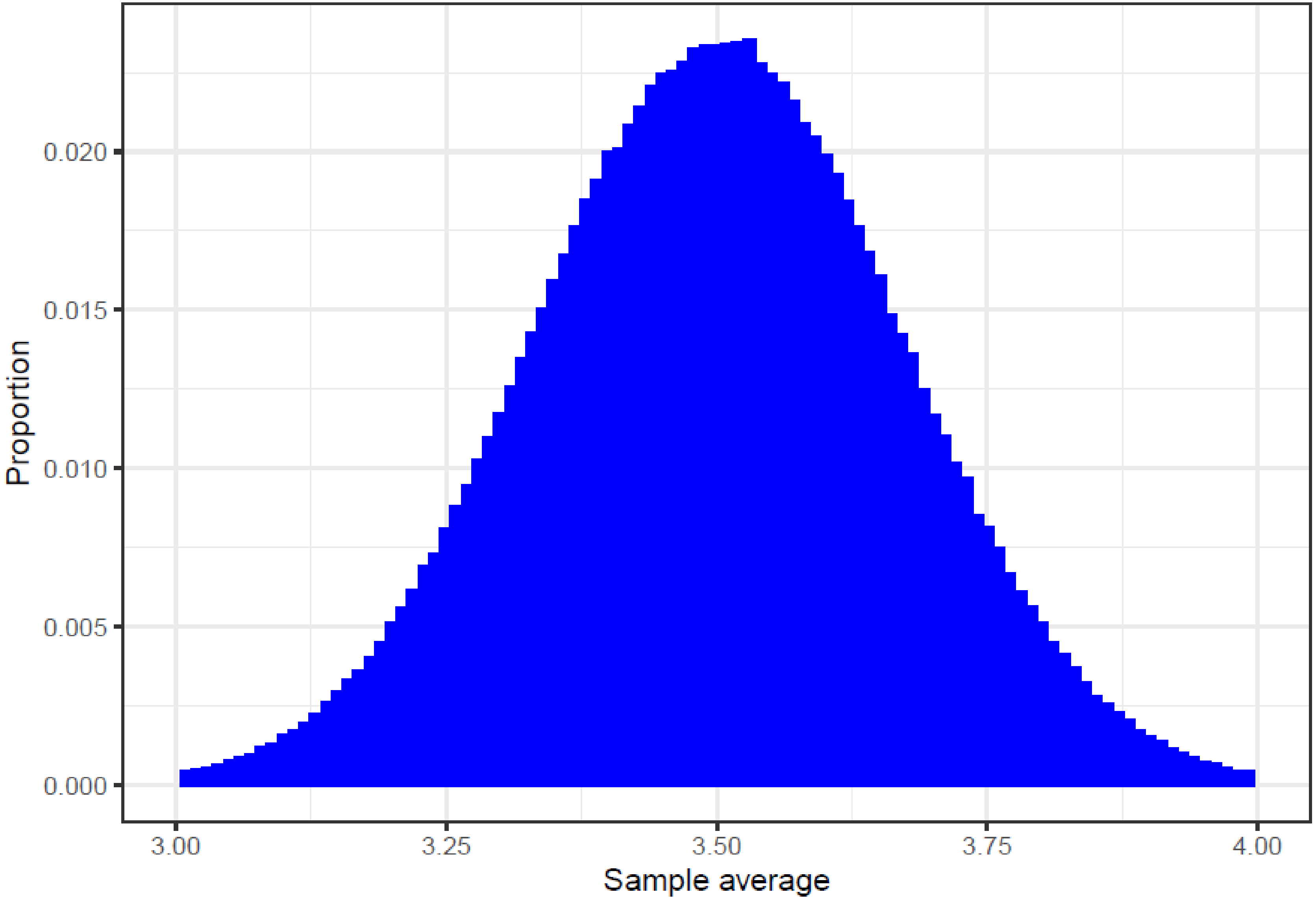
$n = 10$

The sample average of dice rolls (renormalized)



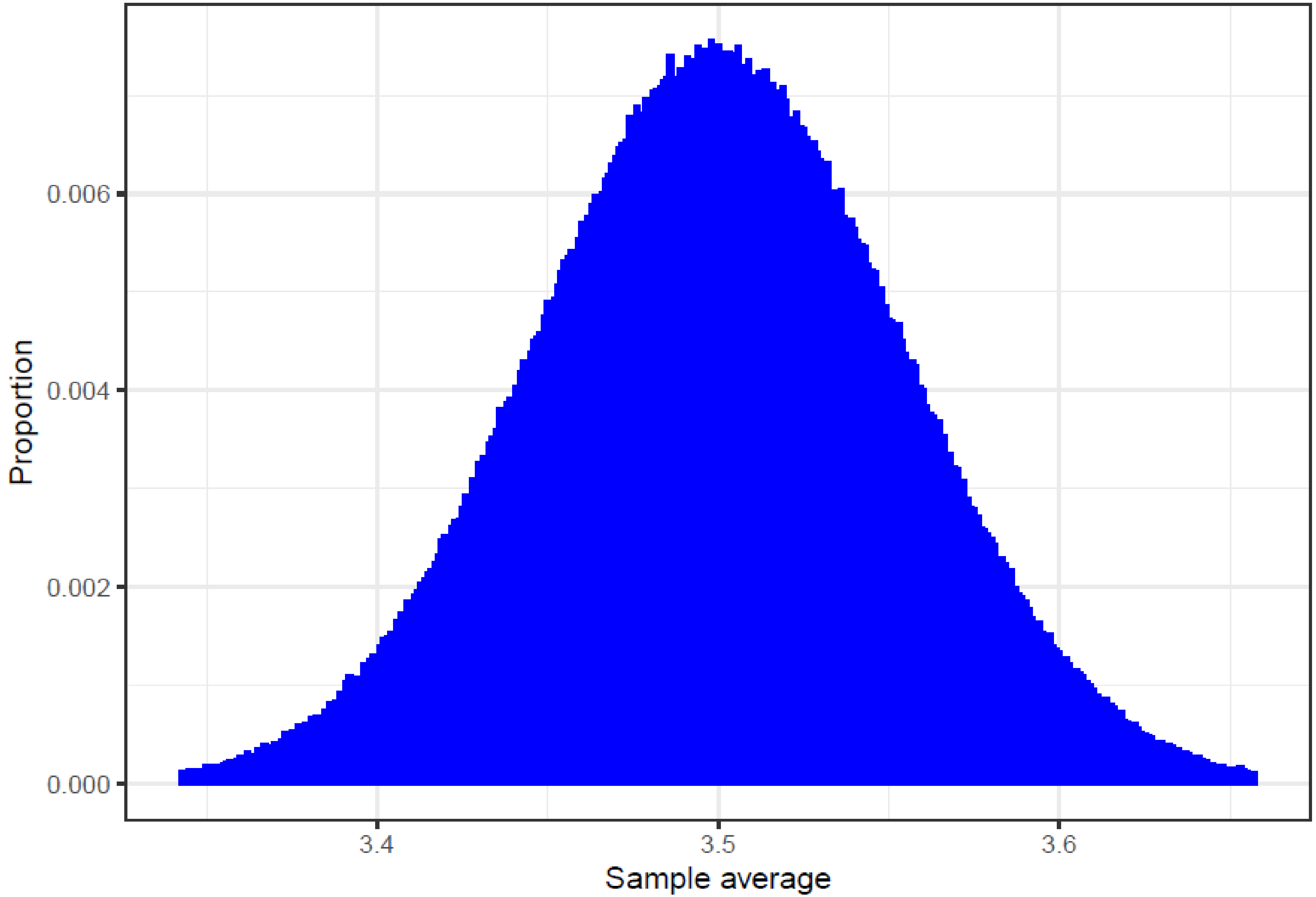
$n = 25$

The sample average of dice rolls (renormalized)



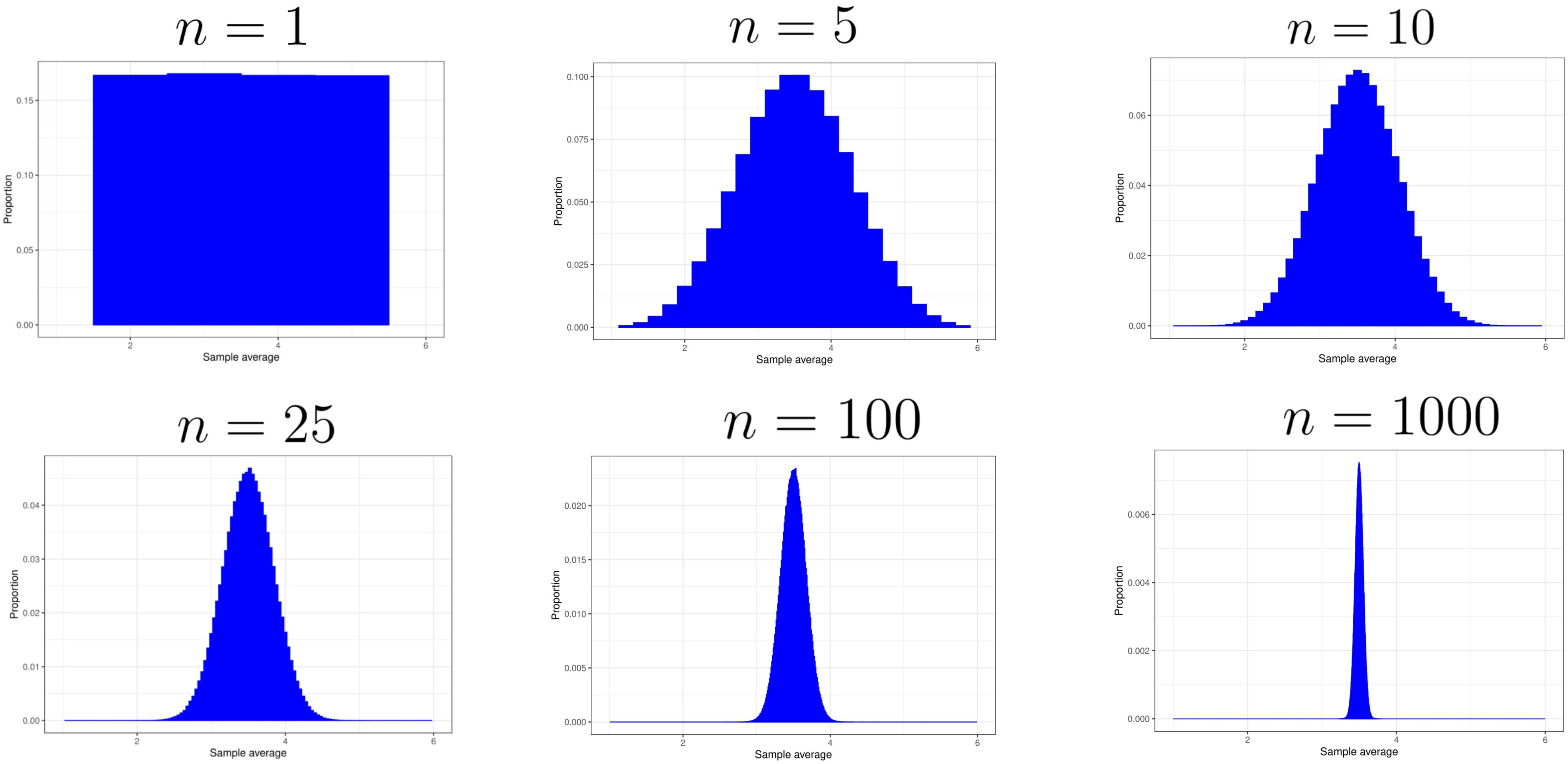
$$n = 100$$

The sample average of dice rolls (renormalized)



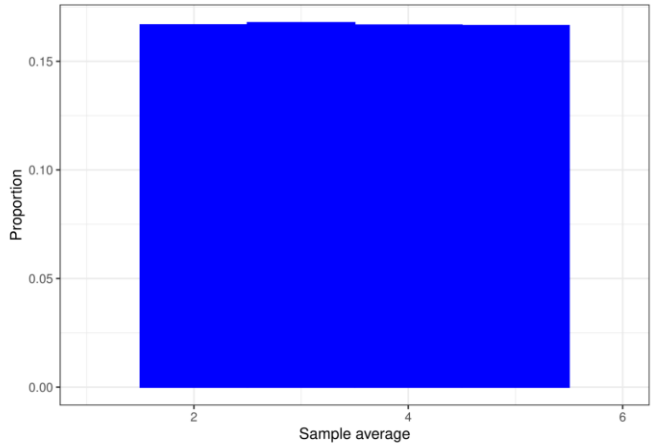
$n = 1000$

The sample average of dice rolls

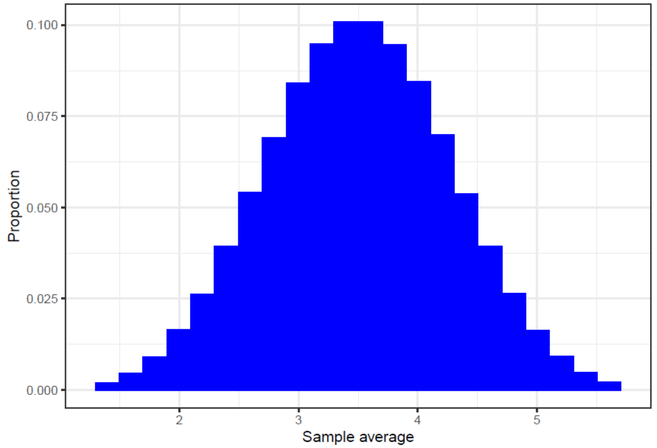


The sample average of dice rolls (renormalized)

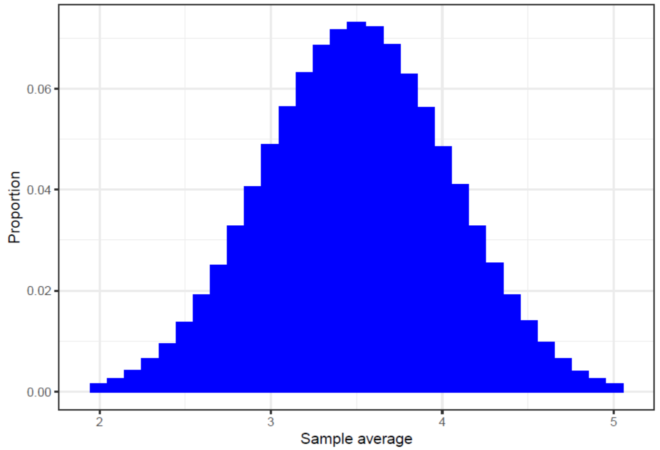
$n = 1$



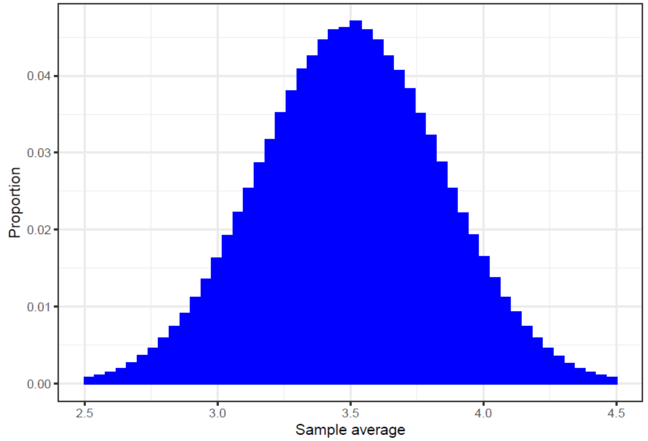
$n = 5$



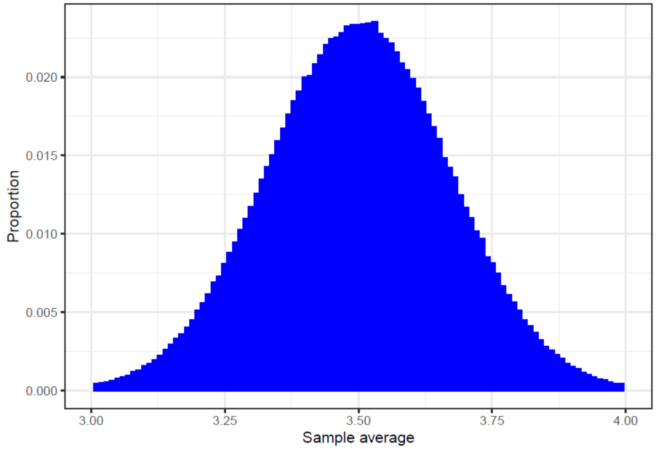
$n = 10$



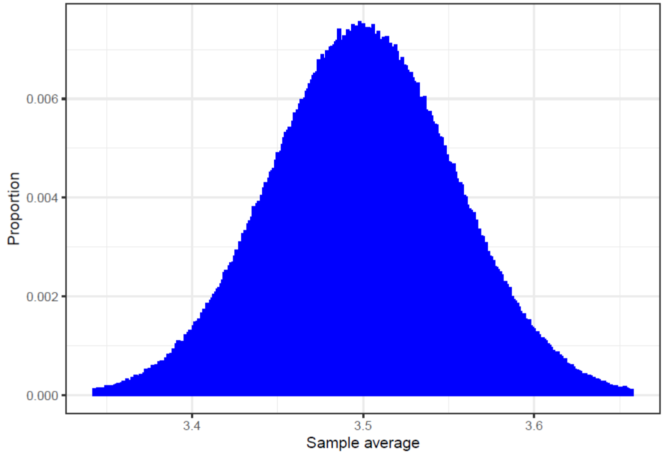
$n = 25$



$n = 100$



$n = 1000$



Now take a break!

Continuous random variables

We often want to model stochastic quantities which can take on a continuum of possible values.

Continuous random variables

We often want to model stochastic quantities which can take on a continuum of possible values.

Examples

- The wing span of a penguin selected at random from a population of penguins on an island;

Continuous random variables

We often want to model stochastic quantities which can take on a continuum of possible values.

Examples

- The wing span of a penguin selected at random from a population of penguins on an island;

- The time taken to by an athlete to run the London marathon;

Continuous random variables

We often want to model stochastic quantities which can take on a continuum of possible values.

Examples

- The wing span of a penguin selected at random from a population of penguins on an island;

- The time taken to by an athlete to run the London marathon;

- The level of rainfall in a given location on a particular day of the year.

Probability density functions

Suppose we have a probability space $(\Omega, \mathcal{E}, \mathbb{P})$.

A **random variable** is a mapping $X : \Omega \to \mathbb{R}$ such that for every a, $b \in \mathbb{R}$, $\{\omega \in \Omega \ : \ X(\omega) \in [a, b]\} \in \mathcal{E}$ is an event.

Probability density functions

Suppose we have a probability space $(\Omega, \mathcal{E}, \mathbb{P})$.

A **random variable** is a mapping $X : \Omega \to \mathbb{R}$ such that for every a, $b \in \mathbb{R}$, $\{\omega \in \Omega \ : \ X(\omega) \in [a, b]\} \in \mathcal{E}$ is an event.

Discrete random variable can be characterised by is probability mass function $x \mapsto \mathbb{P}\,(X = x)$.

Probability density functions

Suppose we have a probability space $(\Omega, \mathcal{E}, \mathbb{P})$.

A **random variable** is a mapping $X : \Omega \to \mathbb{R}$ such that for every a, $b \in \mathbb{R}$, $\{\omega \in \Omega \;:\; X(\omega) \in [a, b]\} \in \mathcal{E}$ is an event.

Discrete random variable can be characterised by is probability mass function $x \mapsto \mathbb{P}\left(X = x\right)$.

However, for a continuous random variable we have $\mathbb{P}\left(X = x\right) = 0$ for all $x \in \mathbb{R}$.

Probability density functions

Suppose we have a probability space $(\Omega, \mathcal{E}, \mathbb{P})$.

A **random variable** is a mapping $X : \Omega \to \mathbb{R}$ such that for every a, $b \in \mathbb{R}$, $\{\omega \in \Omega \ : \ X(\omega) \in [a, b]\} \in \mathcal{E}$ is an event.

Discrete random variable can be characterised by is probability mass function $x \mapsto \mathbb{P}(X = x)$.

However, for a continuous random variable we have $\mathbb{P}(X = x) = 0$ for all $x \in \mathbb{R}$.

To understand continuous random variables $X : \Omega \to \mathbb{R}$ we require a probability density function.

Probability density functions

Suppose we have a probability space $(\Omega, \mathcal{E}, \mathbb{P})$.

A **random variable** is a mapping $X : \Omega \to \mathbb{R}$ such that for every a, $b \in \mathbb{R}$, $\{\omega \in \Omega \ : \ X(\omega) \in [a, b]\} \in \mathcal{E}$ is an event.

Discrete random variable can be characterised by is probability mass function $x \mapsto \mathbb{P}(X = x)$.

However, for a continuous random variable we have $\mathbb{P}(X = x) = 0$ for all $x \in \mathbb{R}$.

To understand continuous random variables $X : \Omega \to \mathbb{R}$ we require a probability density function.

A **probability density function** is a function $f : \mathbb{R} \to [0, \infty)$ such that $\int_{-\infty}^{\infty} f(x)dx = 1$.

Continuous random variables and probability density functions

Suppose we have a probability space $(\Omega, \mathcal{E}, \mathbb{P})$.

A **random variable** is a mapping $X : \Omega \to \mathbb{R}$ such that for every a, $b \in \mathbb{R}$, $\{\omega \in \Omega \ : \ X(\omega) \in [a, b]\} \in \mathcal{E}$ is an event.

A **probability density function** (p.d.f.) is a function $f : \mathbb{R} \to [0, \infty)$ such that $\int_{-\infty}^{\infty} f(x)dx = 1$.

Continuous random variables and probability density functions

Suppose we have a probability space $(\Omega, \mathcal{E}, \mathbb{P})$.

A **random variable** is a mapping $X : \Omega \to \mathbb{R}$ such that for every a, $b \in \mathbb{R}$, $\{\omega \in \Omega \ : \ X(\omega) \in [a, b]\} \in \mathcal{E}$ is an event.

A **probability density function** (p.d.f.) is a function $f : \mathbb{R} \to [0, \infty)$ such that $\int_{-\infty}^{\infty} f(x)dx = 1$.

A **continuous random variable** is a random variable $X : \Omega \to \mathbb{R}$ with a probability density function $f_X : \mathbb{R} \to [0, \infty)$ such that for all $a, b \in \mathbb{R}$ we have

$$\mathbb{P}(X \in [a, b]) = \int_a^b f_X(x)dx.$$

Continuous random variables and probability density functions

Suppose we have a probability space $(\Omega, \mathcal{E}, \mathbb{P})$.

A **random variable** is a mapping $X : \Omega \to \mathbb{R}$ such that for every a, $b \in \mathbb{R}$, $\{\omega \in \Omega \ : \ X(\omega) \in [a, b]\} \in \mathcal{E}$ is an event.

A **probability density function** (p.d.f.) is a function $f : \mathbb{R} \to [0, \infty)$ such that $\int_{-\infty}^{\infty} f(x)dx = 1$.

A **continuous random variable** is a random variable $X : \Omega \to \mathbb{R}$ with a probability density function $f_X : \mathbb{R} \to [0, \infty)$ such that for all $a, b \in \mathbb{R}$ we have

$$\mathbb{P}(X \in [a, b]) = \int_a^b f_X(x)dx.$$

If this holds, we say that $f_X : \mathbb{R} \to [0, \infty)$ is the p.d.f. of the random variable $X : \Omega \to \mathbb{R}$.

It follows that the for all $x \in \mathbb{R}$, $F_X(x) := \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(z)dz$.

Gaussian random variables

A classical example of continuous random variable X is a Gaussian with parameters (μ, σ)

Gaussian random variables

A classical example of continuous random variable X is a Gaussian with parameters (μ, σ)

The associated density $f_{\mu,\sigma} : \mathbb{R} \to [0, \infty)$ given by

$$f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) \quad \text{for all} \quad x \in \mathbb{R}$$

Gaussian random variables

A classical example of continuous random variable X is a Gaussian with parameters (μ, σ)

The associated density $f_{\mu,\sigma} : \mathbb{R} \to [0, \infty)$ given by

$$f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) \quad \text{for all} \quad x \in \mathbb{R}$$

A Gaussian random variable is often referred to as a normal random variable.

Gaussian random variables

A classical example of continuous random variable X is a Gaussian with parameters (μ, σ)

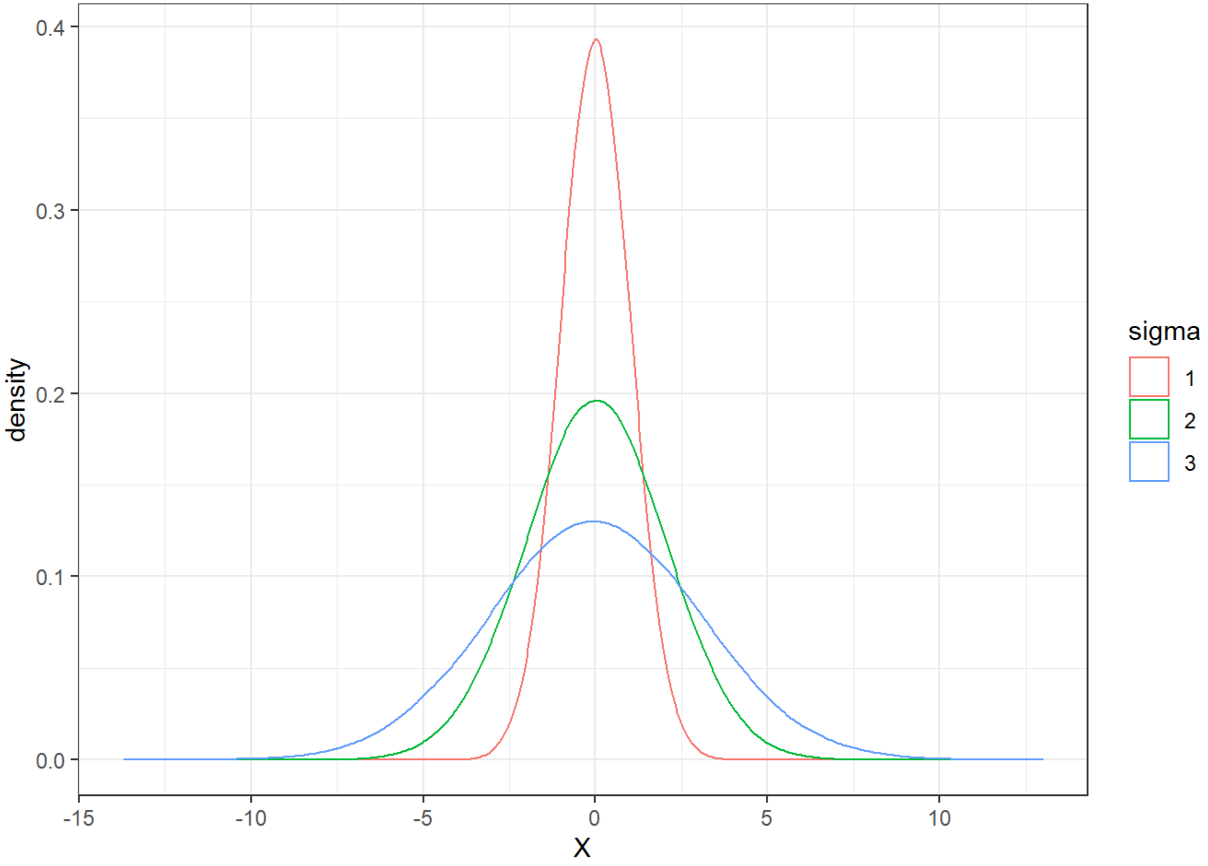
The associated density $f_{\mu,\sigma} : \mathbb{R} \to [0, \infty)$ given by

$$f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) \quad \text{for all} \quad x \in \mathbb{R}$$

A Gaussian random variable is often referred to as a normal random variable.

We often write $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ to mean X is Gaussian with parameters (μ, σ)

Gaussian random variables



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

Gaussian random variables

A Gaussian random variable $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ has density $f_{\mu,\sigma} : \mathbb{R} \to [0, \infty)$

defined by $f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$

Gaussian random variables

A Gaussian random variable $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ has density $f_{\mu,\sigma} : \mathbb{R} \to [0, \infty)$

defined by $f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$

To be a well-defined density we must have $\int_{-\infty}^{\infty} f_{\mu,\sigma}(x)dx = 1$

Gaussian random variables

A Gaussian random variable $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ has density $f_{\mu,\sigma} : \mathbb{R} \to [0, \infty)$

defined by $f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$

To be a well-defined density we must have $\int_{-\infty}^{\infty} f_{\mu,\sigma}(x)dx = 1$

To see this start with the **Gaussian integral** $\int_{-\infty}^{\infty} e^{-z^2} dz = \sqrt{\pi}$

Gaussian random variables

A Gaussian random variable $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ has density $f_{\mu,\sigma} : \mathbb{R} \to [0, \infty)$

defined by $f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$

To be a well-defined density we must have $\int_{-\infty}^{\infty} f_{\mu,\sigma}(x)dx = 1$

To see this start with the **Gaussian integral** $\int_{-\infty}^{\infty} e^{-z^2} dz = \sqrt{\pi}$

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

Gaussian random variables

A Gaussian random variable $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ has density $f_{\mu,\sigma} : \mathbb{R} \to [0, \infty)$

defined by $f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$

To be a well-defined density we must have $\int_{-\infty}^{\infty} f_{\mu,\sigma}(x)dx = 1$

To see this start with the **Gaussian integral** $\int_{-\infty}^{\infty} e^{-z^2} dz = \sqrt{\pi}$ and then

apply a change of variables $z = \frac{x-\mu}{\sigma\sqrt{2}}$ with derivative $\frac{\partial z}{\partial x} = \frac{1}{\sigma\sqrt{2}}$,

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

Gaussian random variables

A Gaussian random variable $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ has density $f_{\mu,\sigma} : \mathbb{R} \to [0, \infty)$

defined by $f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$

To be a well-defined density we must have $\int_{-\infty}^{\infty} f_{\mu,\sigma}(x)dx = 1$

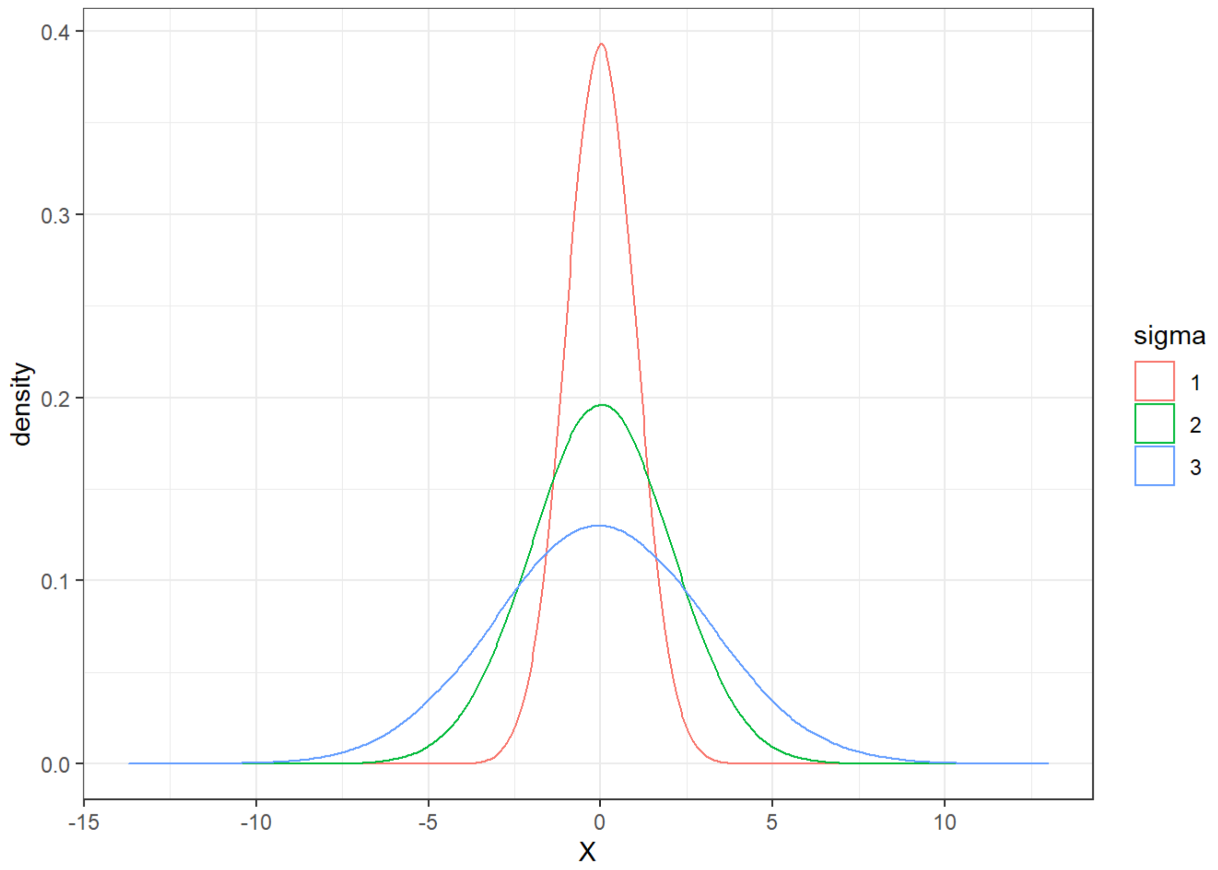
To see this start with the **Gaussian integral** $\int_{-\infty}^{\infty} e^{-z^2} dz = \sqrt{\pi}$ and then

apply a change of variables $z = \frac{x-\mu}{\sigma\sqrt{2}}$ with derivative $\frac{\partial z}{\partial x} = \frac{1}{\sigma\sqrt{2}}$,

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

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2} \left(\frac{\partial z}{\partial x}\right)^{-1} dz = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-z^2} dz = 1.$$

Gaussian random variables



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

Now take a break!

Expectation for continuous random variables

A **continuous random variable** is a random variable $X : \Omega \to \mathbb{R}$ with a probability density function $f_X : \mathbb{R} \to [0, \infty)$ such that for all $a, b \in \mathbb{R}$ we have

$$\mathbb{P}(X \in [a, b]) = \int_a^b f_X(x) dx.$$

Expectation for continuous random variables

A **continuous random variable** is a random variable $X : \Omega \to \mathbb{R}$ with a probability density function $f_X : \mathbb{R} \to [0, \infty)$ such that for all $a, b \in \mathbb{R}$ we have

$$\mathbb{P}(X \in [a, b]) = \int_a^b f_X(x)dx.$$

The **expectation** of a continuous random variable is defined by $\mathbb{E}(X) := \int_{-\infty}^{\infty} x f_X(x)dx$.

Expectation for continuous random variables

A **continuous random variable** is a random variable $X : \Omega \to \mathbb{R}$ with a probability density function $f_X : \mathbb{R} \to [0, \infty)$ such that for all $a, b \in \mathbb{R}$ we have

$$\mathbb{P}\left(X \in [a, b]\right) = \int_a^b f_X(x)dx.$$

The **expectation** of a continuous random variable is defined by $\mathbb{E}(X) := \int_{-\infty}^{\infty} x f_X(x)dx$.

Expectation is *linear* in the following sense:

Given random variables $X_1, \ldots, X_K : \Omega \to \mathbb{R}$ and numbers $\alpha_1, \ldots, \alpha_K \in \mathbb{R}$ we have

$$\mathbb{E}\left(\sum_{i=1}^{K} \alpha_i X_i\right) = \sum_{i=1}^{K} \alpha_i \mathbb{E}\left(X_i\right).$$

The expectation of Gaussian random variables

A Gaussian random variable $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ has density $f_{\mu,\sigma} : \mathbb{R} \to [0, \infty)$

defined by $f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$

The expectation of Gaussian random variables

A Gaussian random variable $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ has density $f_{\mu,\sigma} : \mathbb{R} \to [0, \infty)$

defined by $f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$

We compute $\mathbb{E}[X]$ by combining $\int_{-\infty}^{\infty} e^{-z^2} dz = \sqrt{\pi}$ with the change of variables $z = \frac{x-\mu}{\sigma\sqrt{2}}$

The expectation of Gaussian random variables

A Gaussian random variable $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ has density $f_{\mu,\sigma} : \mathbb{R} \to [0, \infty)$

defined by $f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$

We compute $\mathbb{E}[X]$ by combining $\int_{-\infty}^{\infty} e^{-z^2} dz = \sqrt{\pi}$ with the change of variables $z = \frac{x-\mu}{\sigma\sqrt{2}}$

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_{\mu,\sigma}(x) dx = \int_{-\infty}^{\infty} \frac{x}{\sigma\sqrt{2\pi}} \exp\left(-\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)^2\right) dx$$

The expectation of Gaussian random variables

A Gaussian random variable $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ has density $f_{\mu,\sigma} : \mathbb{R} \to [0, \infty)$

defined by $f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$

We compute $\mathbb{E}[X]$ by combining $\int_{-\infty}^{\infty} e^{-z^2} dz = \sqrt{\pi}$ with the change of variables $z = \frac{x-\mu}{\sigma\sqrt{2}}$

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_{\mu,\sigma}(x) dx = \int_{-\infty}^{\infty} \frac{x}{\sigma\sqrt{2\pi}} \exp\left(-\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)^2\right) dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (\mu + \sigma\sqrt{2} \cdot z) \cdot e^{-z^2} \left(\frac{\partial z}{\partial x}\right)^{-1} dz = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} (\mu + \sigma\sqrt{2} \cdot z) \cdot e^{-z^2} dz$$

The expectation of Gaussian random variables

A Gaussian random variable $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ has density $f_{\mu,\sigma} : \mathbb{R} \to [0, \infty)$

defined by $f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$

We compute $\mathbb{E}[X]$ by combining $\int_{-\infty}^{\infty} e^{-z^2} dz = \sqrt{\pi}$ with the change of variables $z = \frac{x-\mu}{\sigma\sqrt{2}}$

$$\begin{aligned}
\mathbb{E}[X] &= \int_{-\infty}^{\infty} x f_{\mu,\sigma}(x) dx = \int_{-\infty}^{\infty} \frac{x}{\sigma\sqrt{2\pi}} \exp\left(-\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)^2\right) dx \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (\mu + \sigma\sqrt{2} \cdot z) \cdot e^{-z^2} \left(\frac{\partial z}{\partial x}\right)^{-1} dz = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} (\mu + \sigma\sqrt{2} \cdot z) \cdot e^{-z^2} dz \\
&= \frac{\mu}{\sqrt{\pi}} \cdot \int_{-\infty}^{\infty} e^{-z^2} dz + \frac{\sigma\sqrt{2}}{\sqrt{\pi}} \int_{-\infty}^{\infty} z \cdot e^{-z^2} dz = \mu.
\end{aligned}$$

Variance for continuous random variables

A **continuous random variable** is a random variable $X : \Omega \to \mathbb{R}$ with a probability density function $f_X : \mathbb{R} \to [0, \infty)$ such that for all $a, b \in \mathbb{R}$ we have

$$\mathbb{P}\left(X \in [a, b]\right) = \int_a^b f_X(x)dx.$$

The **expectation** of a continuous random variable is defined by $\mathbb{E}(X) := \int_{-\infty}^{\infty} x f_X(x)dx$.

Variance for continuous random variables

A **continuous random variable** is a random variable $X : \Omega \to \mathbb{R}$ with a probability density function $f_X : \mathbb{R} \to [0, \infty)$ such that for all $a, b \in \mathbb{R}$ we have

$$\mathbb{P}(X \in [a,b]) = \int_a^b f_X(x)dx.$$

The **expectation** of a continuous random variable is defined by $\mathbb{E}(X) := \int_{-\infty}^{\infty} x f_X(x)dx$.

We define the **variance** of a continuous random variable $\mathrm{Var}(X) := \mathbb{E}\left[\{X - \mathbb{E}(X)\}^2\right]$.

Standard deviation for continuous random variables

A **continuous random variable** is a random variable $X : \Omega \to \mathbb{R}$ with a probability density function $f_X : \mathbb{R} \to [0, \infty)$ such that for all $a, b \in \mathbb{R}$ we have

$$\mathbb{P}(X \in [a, b]) = \int_a^b f_X(x)dx.$$

The **expectation** of a continuous random variable is defined by $\mathbb{E}(X) := \int_{-\infty}^{\infty} x f_X(x)dx$.

We define the **variance** of a continuous random variable $\mathrm{Var}(X) := \mathbb{E}\left[\{X - \mathbb{E}(X)\}^2\right]$.

In addition, the standard deviation of a random variable is defined by $\sigma(X) := \sqrt{\mathrm{Var}(X)}$.

Variance and covariance for continuous random variables

A **continuous random variable** is a random variable $X : \Omega \to \mathbb{R}$ with a probability density function $f_X : \mathbb{R} \to [0, \infty)$ such that for all $a, b \in \mathbb{R}$ we have

$$\mathbb{P}\left(X \in [a,b]\right) = \int_a^b f_X(x)dx.$$

The **expectation** of a continuous random variable is defined by $\mathbb{E}(X) := \int_{-\infty}^{\infty} x f_X(x)dx$.

We define the **variance** of a continuous random variable $\mathrm{Var}(X) := \mathbb{E}\left[\{X - \mathbb{E}(X)\}^2\right]$.

Given a pair of random variables $X, Y : \Omega \to \mathbb{R}$ we define the covariance and correlation by

$$\mathrm{Cov}(X,Y) := \mathbb{E}\left[\{X - \mathbb{E}(X)\}\{Y - \mathbb{E}(Y)\}\right] = \mathbb{E}(X \cdot Y) - \mathbb{E}(X) \cdot \mathbb{E}(Y)$$

$$\mathrm{Corr}(X,Y) := \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X) \cdot \mathrm{Var}(Y)}}.$$

Variance and covariance for continuous random variables

The **expectation** of a continuous random variable is defined by $\mathbb{E}(X) := \int_{-\infty}^{\infty} x f_X(x) dx$.

We define the **variance** of a continuous random variable $\mathrm{Var}(X) := \mathbb{E}\left[\{X - \mathbb{E}(X)\}^2\right]$.

Given a pair of random variables $X,Y : \Omega \to \mathbb{R}$ we define the covariance and correlation by

$$\mathrm{Cov}(X, Y) := \mathbb{E}\left[\{X - \mathbb{E}(X)\}\{Y - \mathbb{E}(Y)\}\right] = \mathbb{E}(X \cdot Y) - \mathbb{E}(X) \cdot \mathbb{E}(Y).$$

Lemma *Given random variables* $X_1, \dots, X_K : \Omega \to \mathbb{R}$ *and* $\alpha_1, \dots, \alpha_K \in \mathbb{R}$ *we have*

$$\mathrm{Var}\left(\sum_{i=1}^{K} \alpha_i X_i\right) = \sum_{i=1}^{K} \alpha_i^2 \mathrm{Var}(X_i) + 2 \sum_{1 \le i < j \le k} \alpha_i \cdot \alpha_j \mathrm{Cov}(X_i, X_j).$$

In particular, if $X_1, \dots, X_k$ *are independent then* $\mathrm{Var}\left(\sum_{i=1}^{K} \alpha_i X_i\right) = \sum_{i=1}^{K} \alpha_i^2 \mathrm{Var}(X_i)$.

The variance of Gaussian random variables

A Gaussian random variable $X \sim \mathcal{N}(\mu, \sigma^2)$ has density $f_{\mu,\sigma} : \mathbb{R} \to [0, \infty)$

defined by $f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$

We want to compute $\mathrm{Var}(X)$

The variance of Gaussian random variables

A Gaussian random variable $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ has density $f_{\mu,\sigma} : \mathbb{R} \to [0, \infty)$

defined by $f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$

We compute $\mathrm{Var}(X)$ by first using $\int_{-\infty}^{\infty} e^{-z^2} dz = \sqrt{\pi}$ to show $2\int_{-\infty}^{\infty} z^2 \cdot e^{-z^2} dz = \sqrt{\pi}.$

The variance of Gaussian random variables

A Gaussian random variable $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ has density $f_{\mu,\sigma} : \mathbb{R} \to [0, \infty)$

defined by $f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$

We compute $\operatorname{Var}(X)$ by first using $\int_{-\infty}^{\infty} e^{-z^2} dz = \sqrt{\pi}$ to show $2\int_{-\infty}^{\infty} z^2 \cdot e^{-z^2} dz = \sqrt{\pi}.$

$$u = e^{-z^2} \qquad v = -z$$

$$\frac{du}{dz} = -2ze^{-z^2} \qquad \frac{dv}{dz} = -1$$

By integration by parts, $2\int_{-\infty}^{\infty} z^2 \cdot e^{-z^2} dz = \int_{-\infty}^{\infty} v\frac{du}{dz} dz = [uv]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} u\frac{dv}{dz} dz$

$$= \left[-z \cdot e^{-z^2}\right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-z^2} dz = \sqrt{\pi}.$$

The variance of Gaussian random variables

A Gaussian random variable $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ has density $f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$

We compute $\mathrm{Var}(X)$ by applying $2\int_{-\infty}^{\infty} z^2 \cdot e^{-z^2} dz = \sqrt{\pi}.$

with the change of variables $z = \frac{x-\mu}{\sigma\sqrt{2}}$ so that $\frac{\partial z}{\partial x} = \frac{1}{\sigma\sqrt{2}}$

The variance of Gaussian random variables

A Gaussian random variable $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ has density $f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$

We compute $\mathrm{Var}(X)$ by applying $2\int_{-\infty}^{\infty} z^2 \cdot e^{-z^2} dz = \sqrt{\pi}.$

with the change of variables $z = \frac{x-\mu}{\sigma\sqrt{2}}$ so that $\frac{\partial z}{\partial x} = \frac{1}{\sigma\sqrt{2}}$

$$\mathrm{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}\left[(X-\mu)^2\right]$$

The variance of Gaussian random variables

A Gaussian random variable $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ has density $f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$

We compute $\mathrm{Var}(X)$ by applying $2\int_{-\infty}^{\infty} z^2 \cdot e^{-z^2} dz = \sqrt{\pi}.$

with the change of variables $z = \frac{x-\mu}{\sigma\sqrt{2}}$ so that $\frac{\partial z}{\partial x} = \frac{1}{\sigma\sqrt{2}}$

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}\left[(X-\mu)^2\right] \\ &= \int_{-\infty}^{\infty} (x-\mu)^2 f_{\mu,\sigma}(x) dx = \int_{-\infty}^{\infty} \frac{(x-\mu)^2}{\sigma\sqrt{2\pi}} \exp\left(-\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)^2\right) dx\end{aligned}$$

The variance of Gaussian random variables

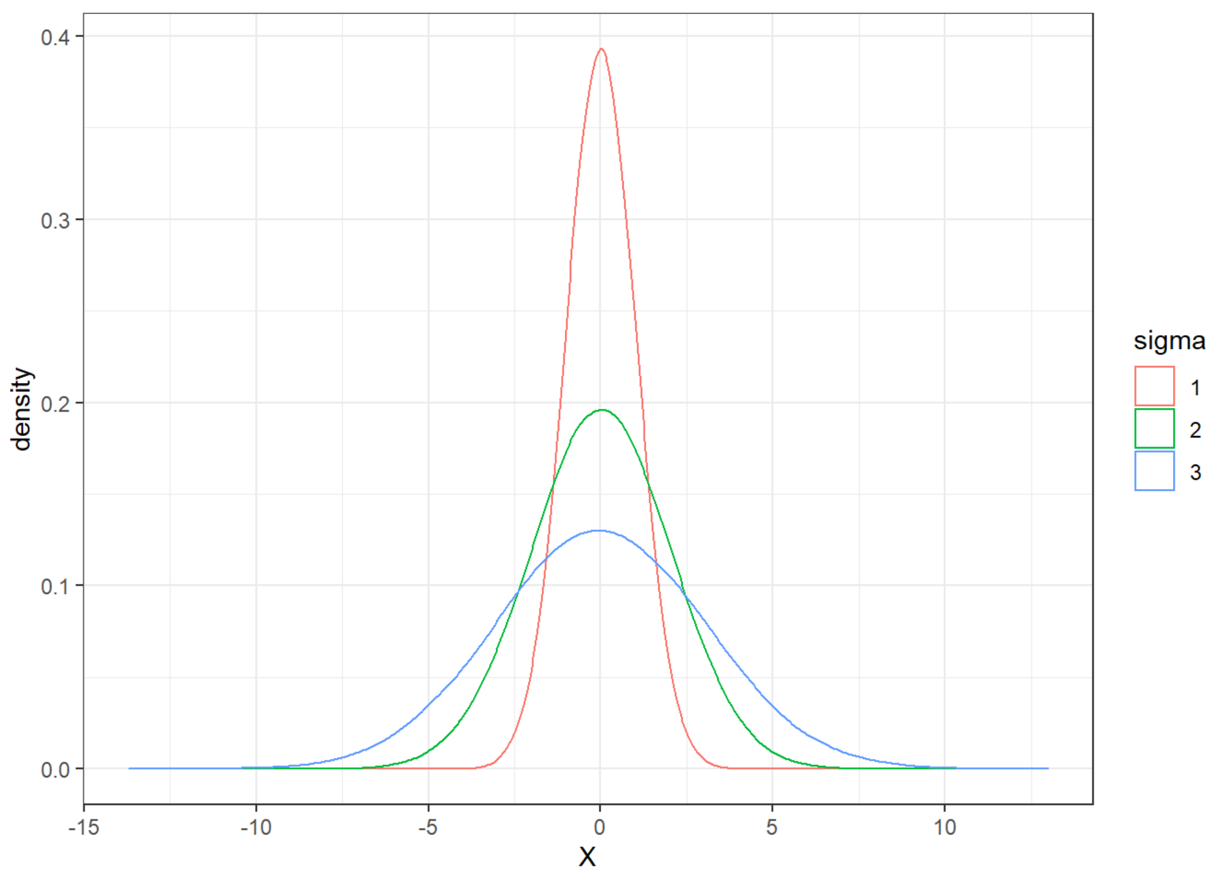
A Gaussian random variable $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ has density $f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$

We compute $\mathrm{Var}(X)$ by applying $2\int_{-\infty}^{\infty} z^2 \cdot e^{-z^2} dz = \sqrt{\pi}.$

with the change of variables $z = \frac{x-\mu}{\sigma\sqrt{2}}$ so that $\frac{\partial z}{\partial x} = \frac{1}{\sigma\sqrt{2}}$

$$
\begin{aligned}
\mathrm{Var}(X) &= \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}\left[(X-\mu)^2\right] \\
&= \int_{-\infty}^{\infty} (x-\mu)^2 f_{\mu,\sigma}(x) dx = \int_{-\infty}^{\infty} \frac{(x-\mu)^2}{\sigma\sqrt{2\pi}} \exp\left(-\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)^2\right) dx \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma\sqrt{2} \cdot z)^2 \cdot e^{-z^2} \left(\frac{\partial z}{\partial x}\right)^{-1} dz = \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} z^2 \cdot e^{-z^2} dz = \sigma^2.
\end{aligned}
$$

Gaussian random variables

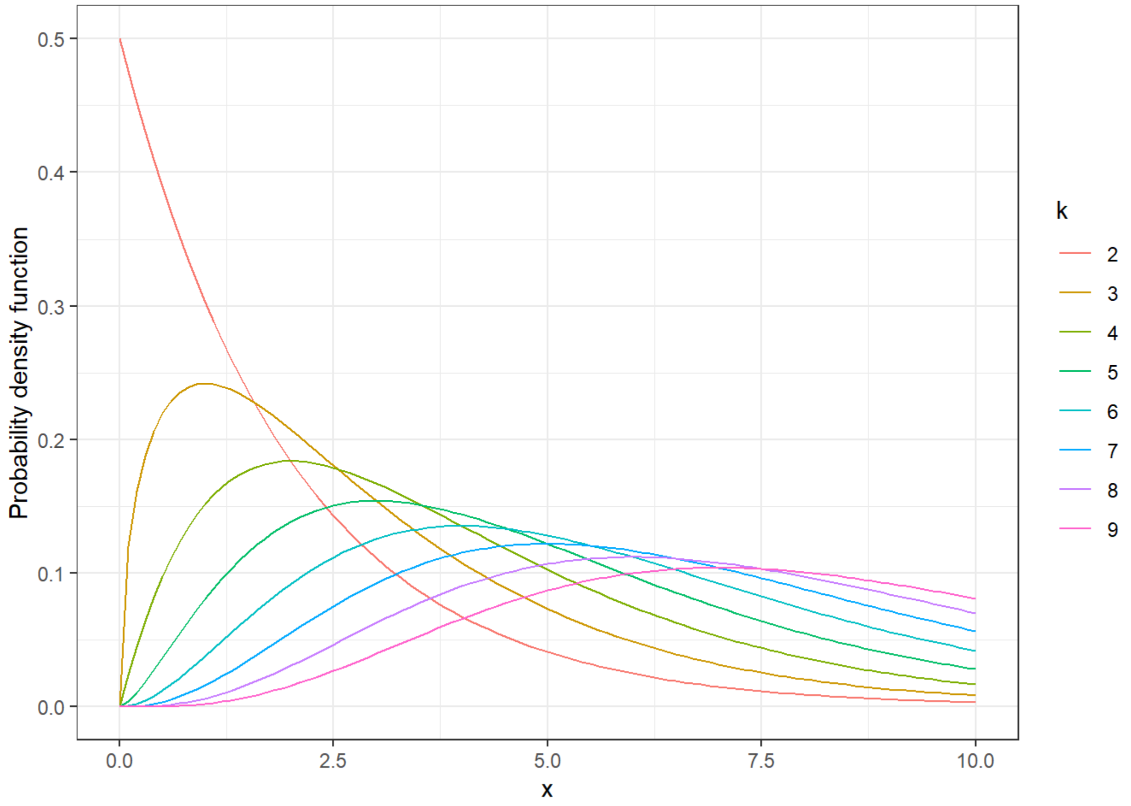


Given $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ we have $\mathbb{E}[X] = \mu$ and $\mathrm{Var}(X) = \sigma^2$

Chi squared distribution

A random variable Q is said to be chi-squared with k degrees of freedom $Q \sim \chi^2(k)$ if

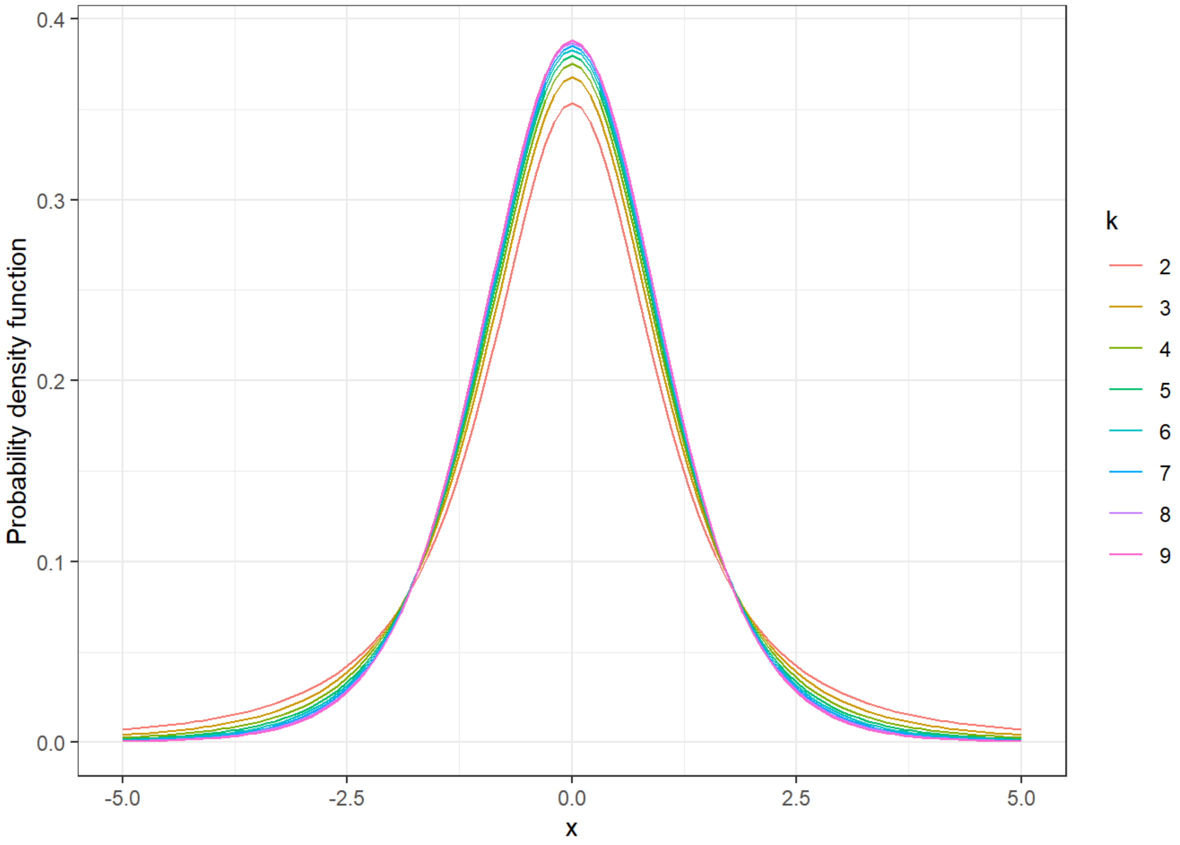
$Q = \sum_{i=1}^{k} Z_i^2$ with $Z_1, \cdots, Z_k \sim \mathcal{N}(0,1)$ independent and identically distributed.



Student's t distribution

A random variable T is said to be t distributed with k degrees of freedom if $T = \dfrac{Z}{\sqrt{Q/k}}$

for two independent random variables $Z \sim \mathcal{N}(0,1)$ and $Q \sim \chi^2(k)$.



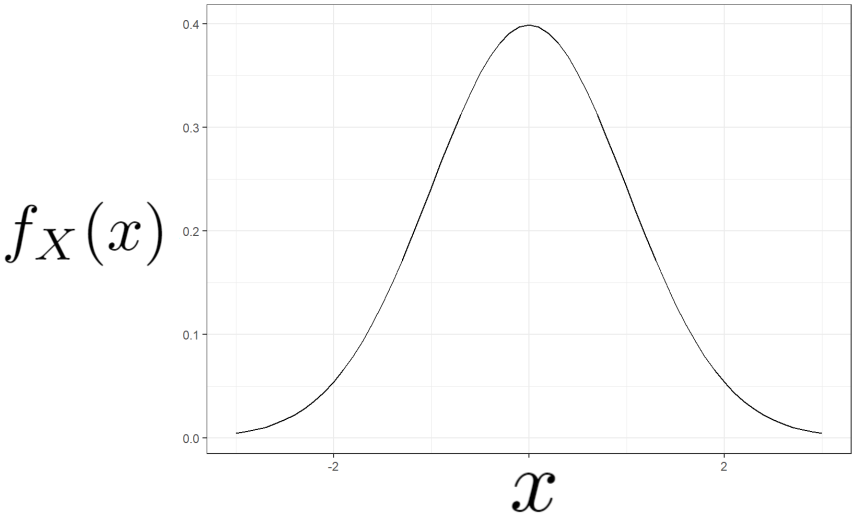
Now take a break!

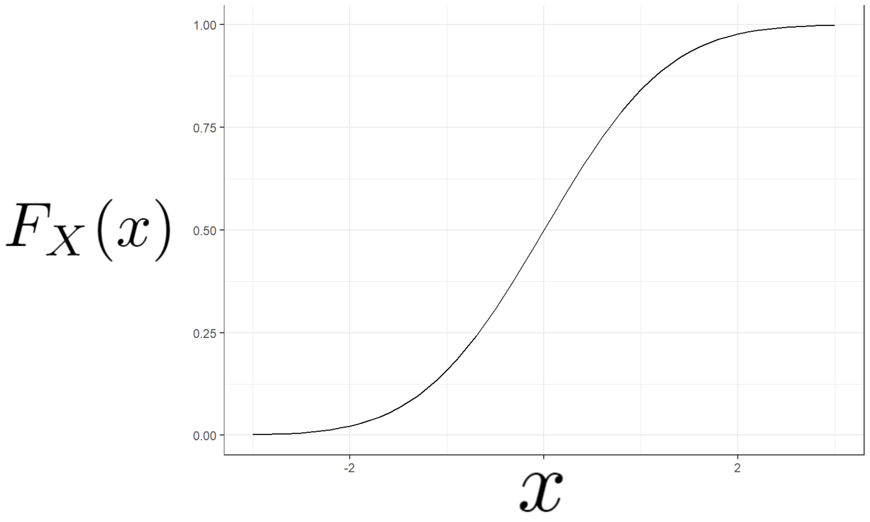
The cumulative distribution function and the density

A **continuous random variable** is a random variable $X : \Omega \to \mathbb{R}$ with a probability density function $f_X : \mathbb{R} \to [0, \infty)$ such that for all $a, b \in \mathbb{R}$ we have $\mathbb{P}\left(X \in [a, b]\right) = \int_a^b f_X(x)dx$.

The **cumulative distribution function** satisfies $F_X(x) := \mathbb{P}(X \leq x) = \int_{-\infty}^x f(z)dz$.

Equivalently, for all $z \in \mathbb{R}$ we have $f_X(z) = \frac{dF_X(x)}{dx}\Big|_z$.





The quantile function

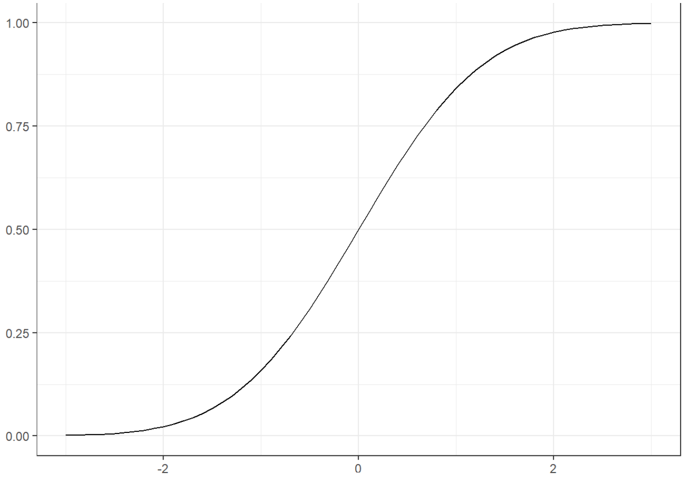
A **continuous random variable** is a random variable $X : \Omega \to \mathbb{R}$ with a probability density function $f_X : \mathbb{R} \to [0, \infty)$ such that for all $a, b \in \mathbb{R}$ we have $\mathbb{P}\left(X \in [a, b]\right) = \int_a^b f_X(x)dx$.

The distribution function $F_X : \mathbb{R} \to [0, 1]$ satisfies $F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f_X(z)dz$.

The **quantile function** $F_X^{-1} : [0, 1] \to \mathbb{R}$ is defined by

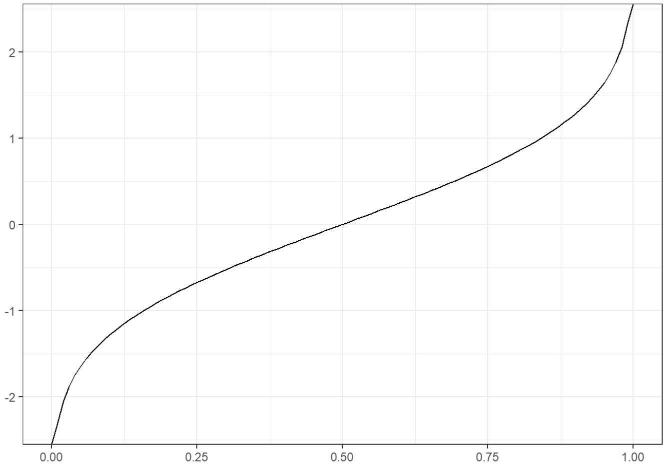
$$F_X^{-1}(p) := \inf\left\{x \in \mathbb{R} : F_X(x) = \mathbb{P}\left(X \leq x\right) \geq p\right\}.$$

$F_X(x)$



x

$F_X^{-1}(p)$



p

The quantile function

A **continuous random variable** is a random variable $X : \Omega \to \mathbb{R}$ with a probability density function $f_X : \mathbb{R} \to [0, \infty)$ such that for all $a, b \in \mathbb{R}$ we have $\mathbb{P}(X \in [a, b]) = \int_a^b f_X(x)dx$.

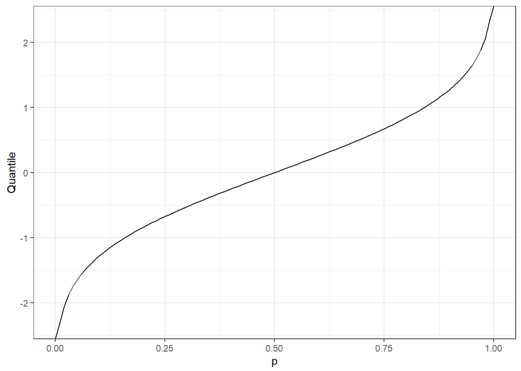
The distribution function $F_X : \mathbb{R} \to [0, 1]$ satisfies $F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^x f_X(z)dz$.

The **quantile function** $F_X^{-1} : [0, 1] \to \mathbb{R}$ is defined by

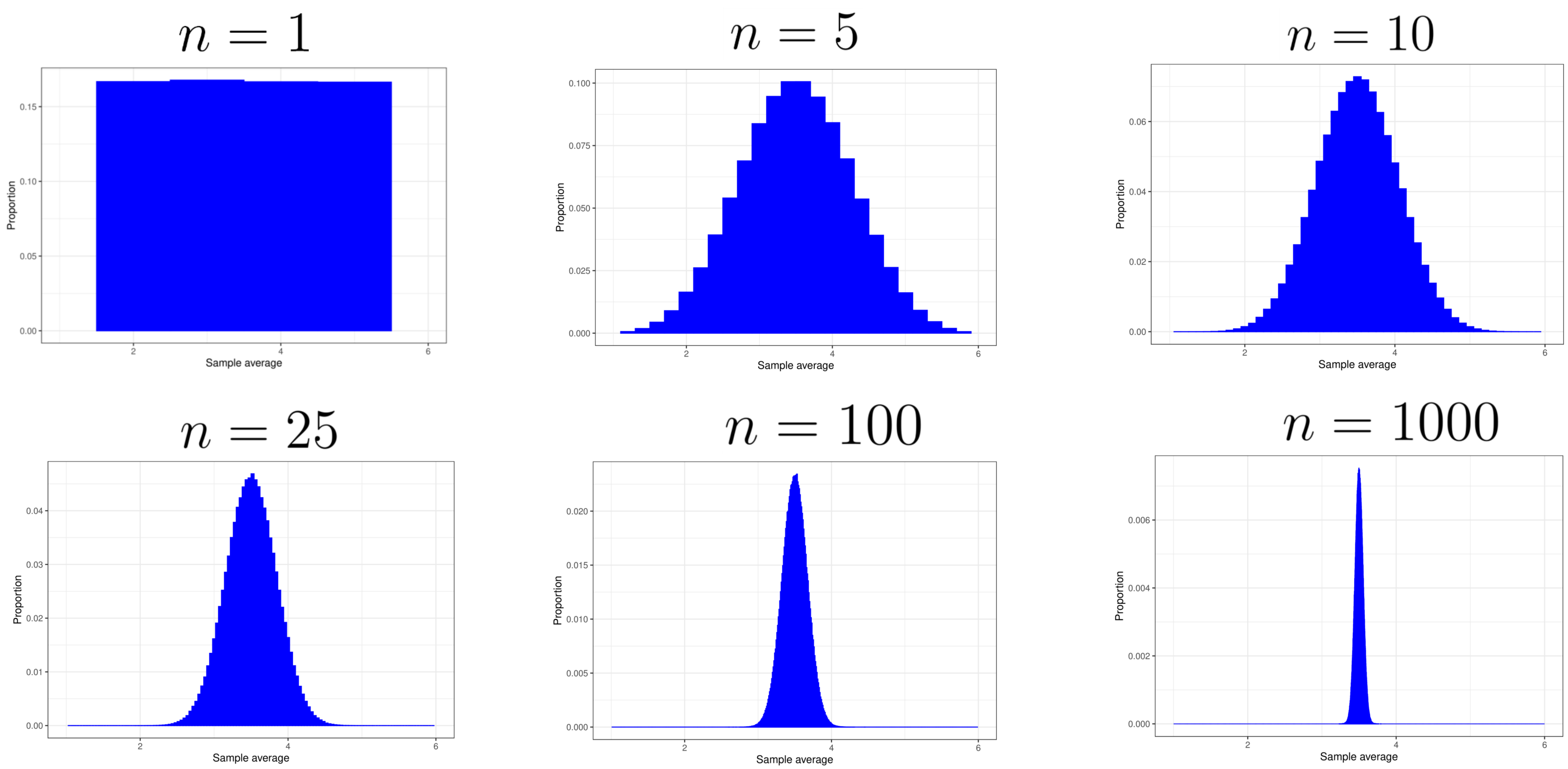
$$F_X^{-1}(p) := \inf \{x \in \mathbb{R} : F_X(x) = \mathbb{P}(X \leq x) \geq p\}.$$

We refer to the values $F_X^{-1}(p)$ for given p as population quantiles.

In particular, $F_X^{-1}(0.5)$ is the population median of the distribution of X.



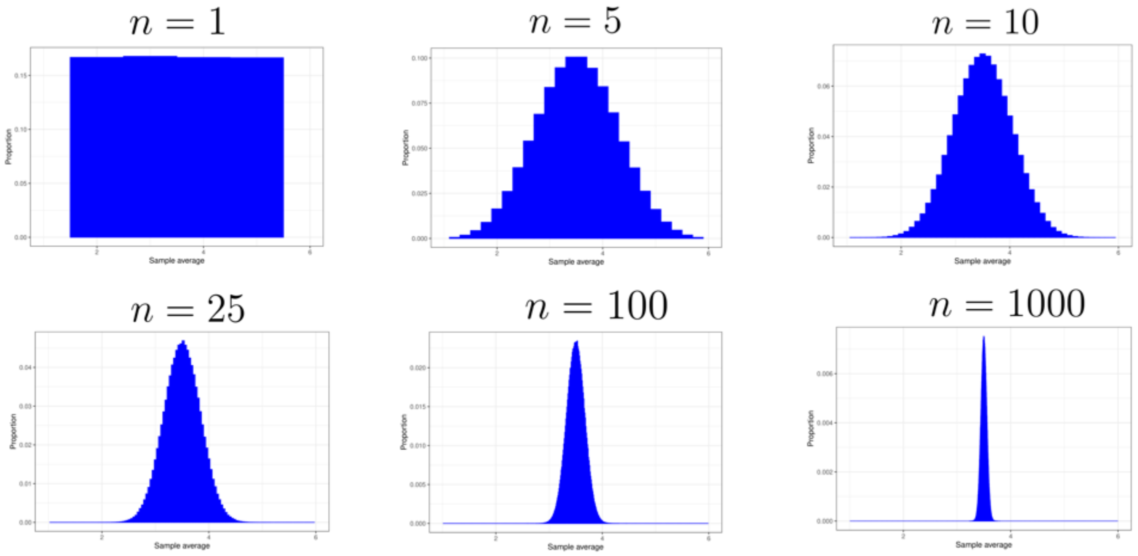
The law of large numbers



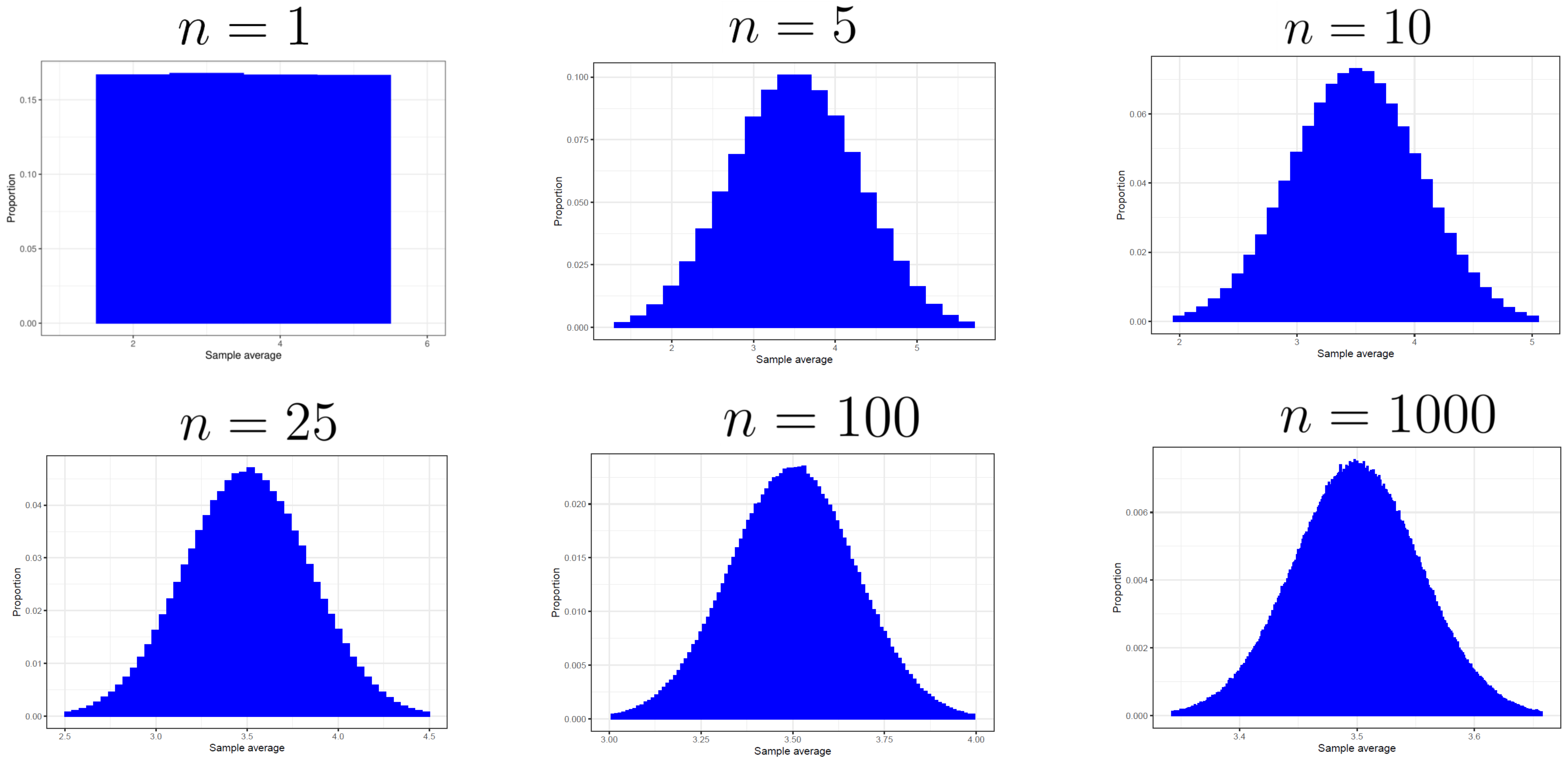
The law of large numbers

Theorem (Bernoulli, circa. 1700). *Let* $X : \Omega \to \mathbb{R}$ *be a random variable with a well-defined expectation* $\mu = \mathbb{E}(X)$. *Let* $X_1, \ldots, X_n : \Omega \to \mathbb{R}$ *be a sequence of independent copies of* X. *Then for all* $\epsilon > 0$,

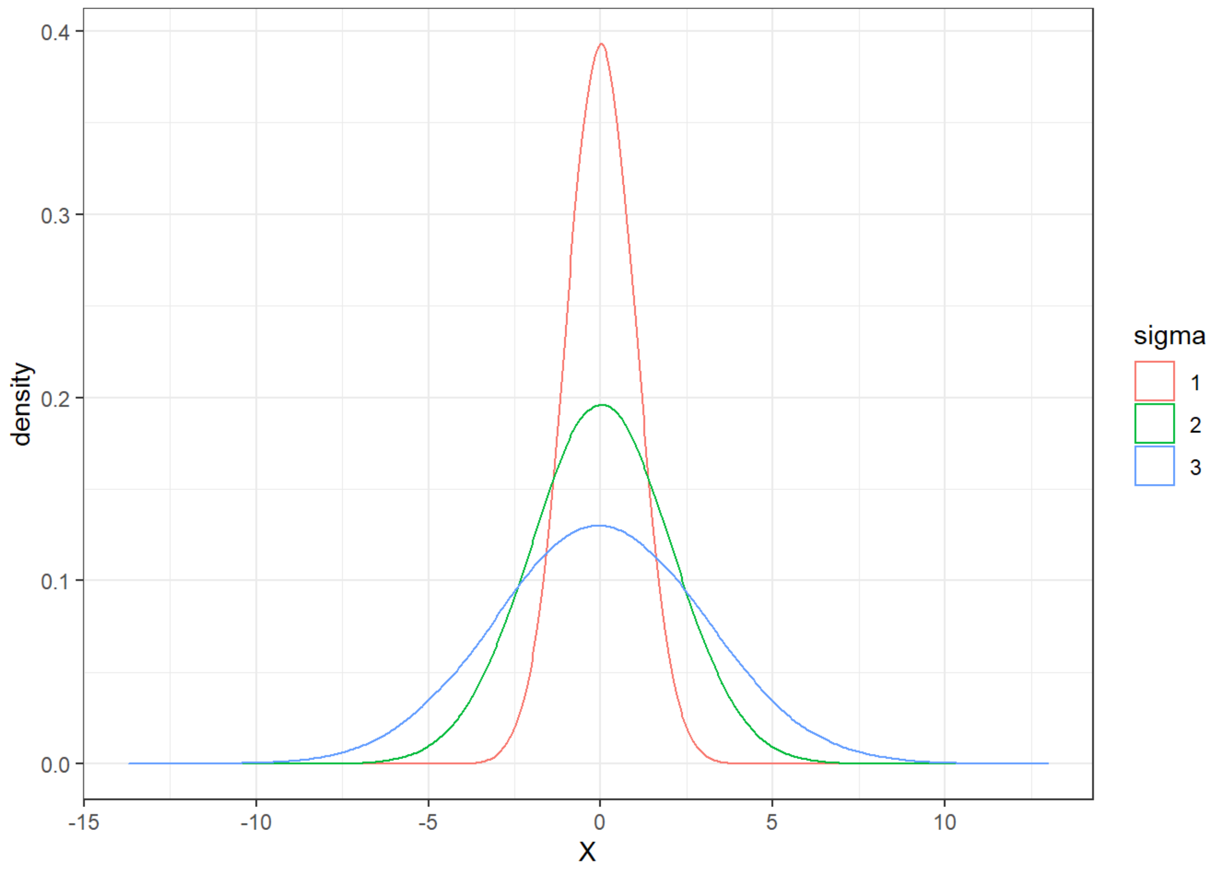
$$\lim_{n \to \infty} \mathbb{P}\left(\left| \frac{1}{n} \sum_{i=1}^{n} X_i - \mu \right| \geq \epsilon \right) = 0.$$



The central limit theorem



Gaussian random variables

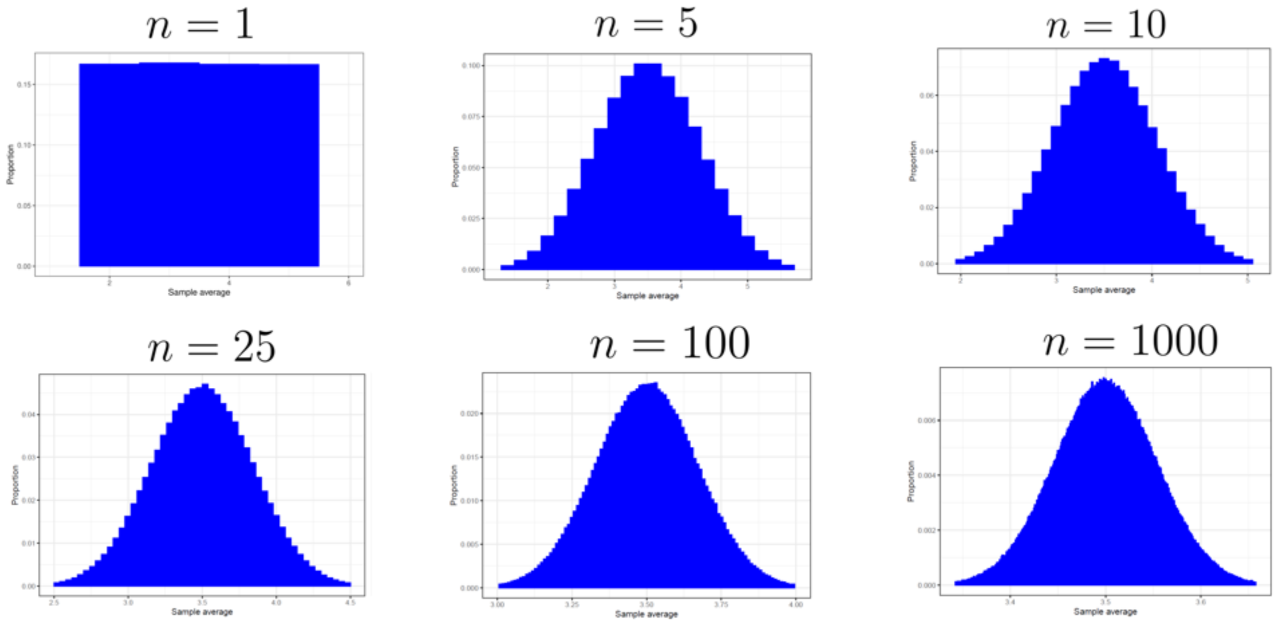


Given $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ we have $\mathbb{E}[X] = \mu$ and $\mathrm{Var}(X) = \sigma^2$

The central limit theorem

Theorem (Lindeberg–Lévy). *Let $X : \Omega \to \mathbb{R}$ be a random variable with expectation $\mu = \mathbb{E}(X)$ and variance $\sigma^2 = \mathrm{Var}(X)$. Let $X_1, \ldots, X_n : \Omega \to \mathbb{R}$ be a sequence of independent copies of X. Let $Z \sim \mathcal{N}(0,1)$ be a standard Gaussian random variable. Then for all $x \in \mathbb{R}$,*

$$\lim_{n\to\infty} \mathbb{P}\left\{\sqrt{\frac{n}{\sigma^2}} \cdot \left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right) \le x\right\} = \mathbb{P}\left(Z \le x\right).$$



What have we covered?

- We introduced the concept of a continuous random variable.
- We saw how continuous random variables can be understood via the probability density function.
- We looked at expectation, variance, standard deviation, covariance and correlation in this context.
- We discussed Gaussian random variables and briefly looked at chi-square & Student's t distributions.
- We saw how the law of large numbers describes the limiting behavior of the average.
- We saw how the central limit theorem gives us greater insight as we zoom in around the mean.

University of BRISTOL

Thanks for listening!

Henry W J Reeve

henry.reeve@bristol.ac.uk

Include EMATM0061 in the subject of your email.

Statistical Computing & Empirical Methods (EMATM0061)

MSc in Data Science, Teaching block 1, 2021.