

Activity submission remainder

- You need to use the template from the **isdas** package.
- You need to submit knitted **PDFs**.
- You should finish **all** the coding parts in the lab activities submission.
- You need to answer **all** the questions in the activity sections (highlights and threshold concepts are optional) in the in-class activity submission.
- We have **two** activities each week, and you need to submit both in each submission.
- You need to submit both PDFs in the **same** submission.

Line breaks in markdown

Simply changing a line in Markdown won't create a line break.

Unlike *L^AT_EX* (`\newline`) or HTML (`
`), Markdown does not have explicit line break syntax.

The following example would produce a single line sentence after knitting:

```
1 first line of text
2 still the first line of text
```

In Markdown, line breaks are created implicitly using empty lines.

The following example would produce two lines after knitting:

```
1 first line of text
2
3 second line of text
```

Packages we use today

Load the following three packages.

```
1 library(isdas)
2 library(sf)
3 library(tidyverse)
4 library(spatstat)
```

If you have trouble restoring the reproducible environment, you need to manually install the packages first.

```
1 install.packages("remotes")
2 remotes::install_github("paezha/isdas")
3
4 install.packages("sf")
5
6 install.packages("tidyverse")
7
8 install.packages("spatstat")
```

New package for today: **spatstat**

spatstat is an R package for spatial statistics with a strong focus on analyzing spatial point patterns in 2D.

You can find documentation for this package at:

- <https://spatstat.org/>
- <https://spatstat.org/resources/spatstatQuickref.pdf>

Random vs. deterministic process

Let's denote the probability of an event occurring at point (x, y) as $Prob(Event_{(x,y)})$. The points are located within a unit square ranging from 0 to 1.

- A random process:

$$Prob(Event_{(x,y)}) = B(n = 1, p = 0.5)$$

- A deterministic process:

$$Prob(Event_{(x,y)}) = x$$

- A stochastic process:

$$Prob(Event_{(x,y)}) = x - x \cdot B(n = 1, p = 0.5)$$

Generate random values from a distribution

There is a family of functions that start with `r*` capable of generating random values from a given distribution.

From a normal distribution:

```
1 rnorm(n = 5, mean = 0, sd = 1)
```

```
[1] 0.5185718 -0.5801848 1.7541242 -1.4321465 -1.3573983
```

From a binomial distribution:

```
1 rbinom(n = 5, size = 1, prob = 0.5)
```

```
[1] 0 1 0 1 0
```

These functions are very useful for simulations.

Random number generator and seed

In computers, there is no true random number generator; they all use pseudo-random number generators. This means that the random numbers produced by computers are just the output of a very complex function based on an input, or seed.

As a result, if we know the seed, we can perfectly predict the random numbers generated by a computer's random number generator.

Typically, computers use the current time as the seed (R uses this approach).

Setting a seed in R

You can control the random number generation algorithm by setting a seed value in R.

```
1 set.seed(437988)
2 rbinom(n = 5, size = 1, prob = 0.5)
```

```
[1] 1 0 0 0 1
```

```
1 set.seed(437988)
2 rbinom(n = 5, size = 1, prob = 0.5)
```

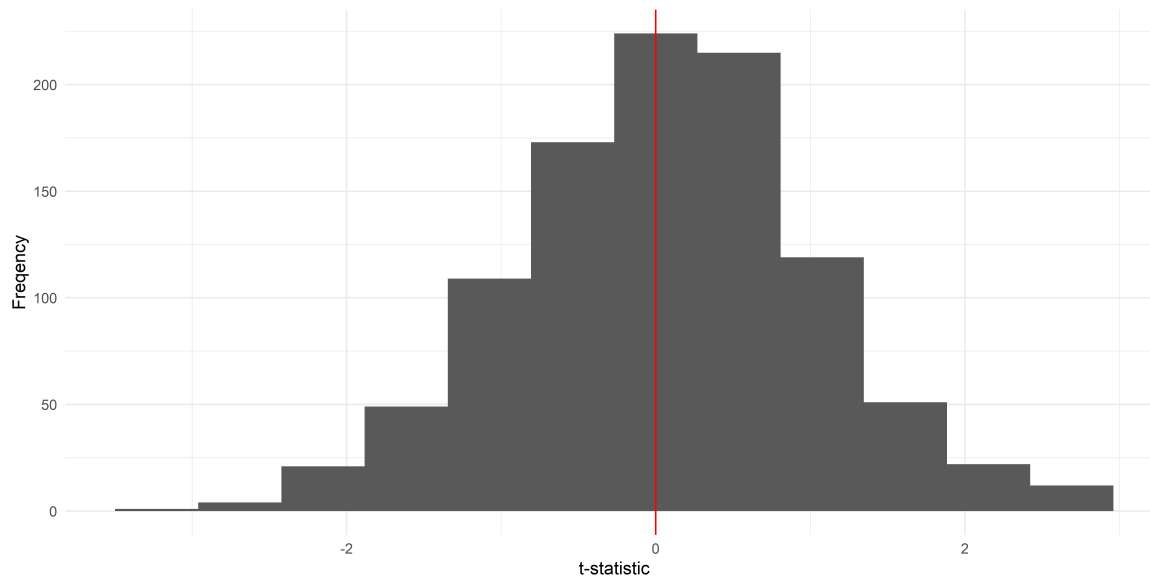
```
[1] 1 0 0 0 1
```

Note that the seed in R is actually a vector, a sequence of values, and `set.seed` changes this vector to a particular state.

If you do not set the seed again before rerunning the second `rbinom`, it will produce different results.

What is p -value?

Let's say we have two samples from two different distributions. We want to perform a statistical test to compare their means ($H_1 : \overline{x}_1 \neq \overline{x}_2$). The p -value is the area under the curve that is more extreme than the test statistic, multiplied by 2.



Online Office Hours

If you need additional help, especially while working on your course project, you're welcome to book a 15- or 30-minute online office hour with me.

Please email me at yinz39@mcmaster.ca with a brief **description of your question** and several **possible meeting times**.

I'll follow up to confirm the appointment and share a Zoom link.

Activities for today

- We will work on the following chapter from the textbook:
 - Chapter 8: Activity 3: Maps as Processes
 - Chapter 10: Activity 4: Point Pattern Analysis I
- The hard deadline is **Tuesday, January 27 (8:00 am)**.