data don't speak for themselves

a quick introduction to Bayesian statistics

João P. Faria

Disclaimer

I'm not a statistician.

I never gave this course before.

I borrowed (heavily) from material presented in the references.

But, unless otherwise noted, all opinions are my own.

Outline

- main differences between frequentist and Bayesian statistics
    - problems with p-values, confidence intervals and null hypothesis testing
- the basic rules of probability theory
- how to assign probability distributions
    - the role of priors
    - the likelihood
- the simplest models in Bayesian statistics
    - linear regression
    - beta-binomial model
    - hierarchical models

Outline

- main differences between frequentist and Bayesian statistics
  - ~~problems with p-values, confidence intervals and null hypothesis testing~~
- the basic rules of probability theory
- how to assign probability distributions
  - the role of priors
  - the likelihood
- the simplest models in Bayesian statistics
  - linear regression
  - ~~beta-binomial model~~
  - hierarchical models
- MCMC

Q/A

Introduction

brief historical sketch

- statistical inference was invented because of astronomy
  (this claim is not based on an in-depth search)

brief historical sketch

- statistical inference was invented because of astronomy
  (this claim is not based on an in-depth search)
- Tycho Brahe, Galileo, Legendre, Laplace, Gauss
  used and developed statistical methods

brief historical sketch

- statistical inference was invented because of astronomy
  (this claim is not based on an in-depth search)
- Tycho Brahe, Galileo, Legendre, Laplace, Gauss
  used and developed statistical methods
- $\sim$ 20th century, astronomers focused on least-squares
  techniques, and heuristic procedures

brief historical sketch

- statistical inference was invented because of astronomy
  (this claim is not based on an in-depth search)
- Tycho Brahe, Galileo, Legendre, Laplace, Gauss
  used and developed statistical methods
- $\sim$ 20th century, astronomers focused on least-squares techniques, and heuristic procedures
- "astrostatistics" emerged in the late 1990s – collaborations between astronomers and statisticians
  names like Babu, Feigelson, Gregory, Hobson

brief historical sketch

- statistical inference was invented because of astronomy
  (this claim is not based on an in-depth search)
- Tycho Brahe, Galileo, Legendre, Laplace, Gauss
  used and developed statistical methods
- $\sim$ 20th century, astronomers focused on least-squares techniques, and heuristic procedures
- "astrostatistics" emerged in the late 1990s – collaborations between astronomers and statisticians
  names like Babu, Feigelson, Gregory, Hobson
- statistics today is mostly Bayesian
  (many) astronomers are (very) sceptical of (sophisticated) statistics

frequentist and Bayesian

Consider n independent measurements of the same quantity, under identical conditions

Calculate their mean $\bar{x}$ and standard deviation σ

frequentist and Bayesian

Consider n independent measurements of the same quantity, under identical conditions

Calculate their mean $\bar{x}$ and standard deviation σ

$$\mu = \bar{x} \pm \frac{\sigma}{\sqrt{n}}$$

frequentist and Bayesian

Consider n independent measurements of the same quantity, under identical conditions
Calculate their mean $\bar{x}$ and standard deviation σ

$$\mu = \bar{x} \pm \frac{\sigma}{\sqrt{n}}$$

what you think this means:

$$p(\bar{x} - \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + \frac{\sigma}{\sqrt{n}}) = 68\%$$

frequentist and Bayesian

Consider n independent measurements of the same quantity, under identical conditions

Calculate their mean $\bar{x}$ and standard deviation σ

$$\mu = \bar{x} \pm \frac{\sigma}{\sqrt{n}}$$

what you think this means:

$$p(\bar{x} - \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + \frac{\sigma}{\sqrt{n}}) = 68\%$$

what it actually means:

$$p(\mu - \frac{\sigma}{\sqrt{n}} \leq \bar{x} \leq \mu + \frac{\sigma}{\sqrt{n}}) = 68\%$$

frequentist and Bayesian

Consider n independent measurements of the same quantity, under identical conditions

Calculate their mean $\bar{x}$ and standard deviation σ

$$\mu = \bar{x} \pm \frac{\sigma}{\sqrt{n}}$$

what you want it to mean:

$$p(\bar{x} - \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + \frac{\sigma}{\sqrt{n}}) = 68\%$$

what it actually means:

$$p(\mu - \frac{\sigma}{\sqrt{n}} \leq \bar{x} \leq \mu + \frac{\sigma}{\sqrt{n}}) = 68\%$$

frequentist and Bayesian

Here is the main difference:

- Frequentists assign a probability to what is *random*
- Bayesians assign a probability to what is *unknown*

Blocks

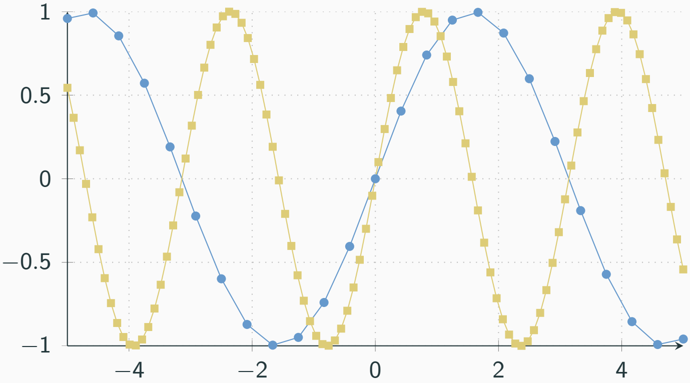
Three different block environments are pre-defined and may be styled with an optional background color.

Default	**Default**
Block content.	Block content.
Alert	**Alert**
Block content.	Block content.
Example	**Example**
Block content.	Block content.

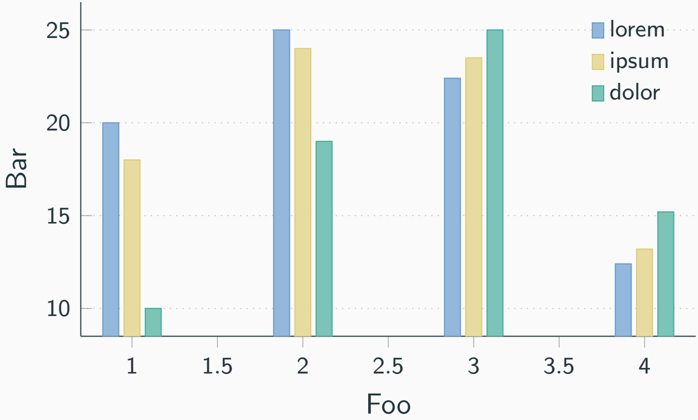
Math

$$e = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n$$

Line plots



Bar charts



Quotes

Veni, Vidi, Vici

References

Some references to showcase [allowframebreaks] [4, 3, 1, 2]

Conclusion

Summary

Get the source of this theme and the demo presentation from

`github.com/matze/mtheme`

Q/A

everything

References I

P. Figueira, J. P. Faria, E. Delgado-Mena, V. Z. Adibekyan, S. G. Sousa, N. C. Santos, and G. Israelian.
Exoplanet hosts reveal lithium depletion. Results from a homogeneous statistical analysis.
Astronomy and Astrophysics, 570:A21, Oct. 2014.

J. Lillo-Box, D. Barrado, N. C. Santos, L. Mancini, P. Figueira, S. Ciceri, and T. Henning.
Kepler-447b: a hot-Jupiter with an extremely grazing transit.
ArXiv150203267 Astro-Ph, Feb. 2015.

References II

N. C. Santos, G. Israelian, M. Mayor, R. Rebolo, and S. Udry.
Statistical properties of exoplanets: II. Metallicity, orbital parameters, and space velocities.
Astron. Astrophys., 398(1):363–376, Jan. 2003.

J. Schneider, V. Lainey, and J. Cabrera.
A next step in exoplanetology: exo-moons.
Int. J. Astrobiol., 14(Special Issue 02):191–199, Apr. 2015.