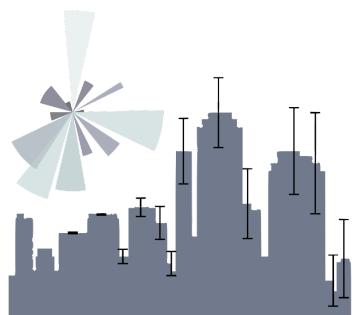


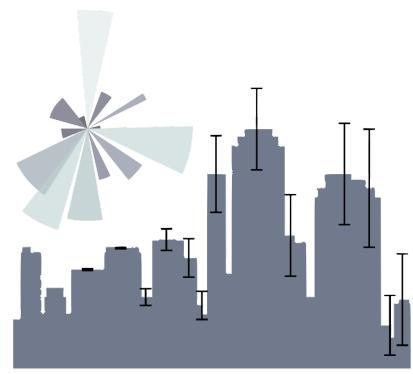
principles of Urban Science



dr.federica bianco | fbb.space |  fedhere |  fedhere

1. Workflow of data-driven projects
 2. The Demarcation problem: what is science
 3. Going from idea to hypothesis
-
1. Reading in data
 2. Descriptive statistics (central tendency, spread...)
 3. Extracting descriptive statistics from data

this slide deck: https://slides.com/federicabianco/pus2020_2



1 data science in public policy

The workflow of a data driven project

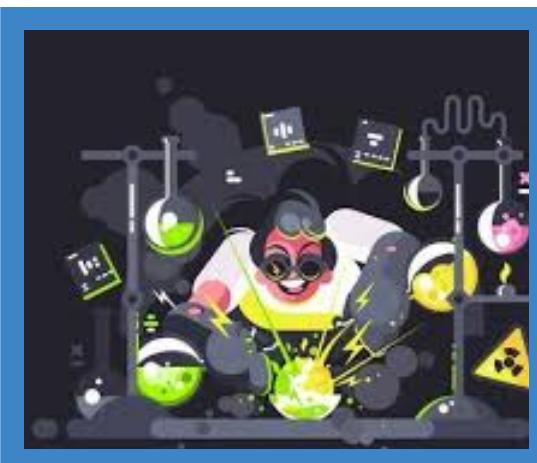
- IDEA
 - formulate a hypothesis
- DATA IDENTIFICATION
 - define ideal data
 - find a suitable dataset
 - figure out best data available
 - figure out if you can get new data
 - obtain data (including policy issues + technical issues)
 - create new data
 - survey
 - plan a survey
 - run a survey (including getting permissions, IRB, figuring out a distribution strategy...)
 - set up a remote sensing facility
 - buy instruments
 - deploy instruments (including getting permissions...)
 - connect instruments
 - run instruments
- DATA WRANGLING
 - cleaning datasets
 - data aggregation (joining datasets)
- DATA EXPLORATION
 - understanding structure in the dat
 - understanding correlation
 - machine learning (clustering, dimensionality reduction?)
- PREDICTING
 - models (e.g. regression)
 - prediction
 - validation (simulations)
 - transferability
- INTERPRETATION
 - feature importance
 - mechanistic models
- RESULT COMMUNICATION
 - visualizations
 - write a report
 - write a paper

The workflow of a data driven project

- IDEA
 - formulate a hypothesis
- DATA IDENTIFICATION
 - define ideal data
 - find a suitable dataset
 - figure out best data available
 - figure out if you can get new data
 - obtain data (including policy issues + technical issues)
 - create new data
 - survey
 - plan a survey
 - run a survey (including getting permissions, IRB, figuring out a distribution strategy...)
 - set up a remote sensing facility
 - buy instruments
 - deploy instruments (including getting permissions...)
 - connect instruments
 - run instruments



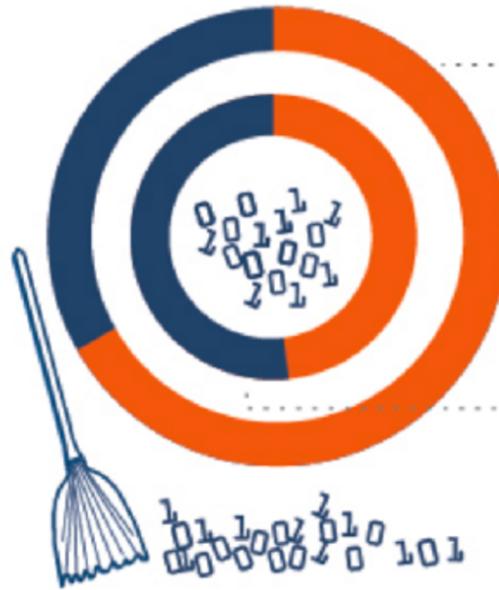
- DATA WRANGLING
 - cleaning datasets
 - data aggregation (joining datasets)
- DATA EXPLORATION
 - understanding structure in the dat
 - understanding correlation
 - machine learning (clustering, dimensionality reduction?)
- PREDICTING
 - models (e.g. regression)
 - prediction
 - validation (simulations)
 - transferability
- INTERPRETATION
 - feature importance
 - mechanistic models
- RESULT COMMUNICATION
 - visualizations
 - write a report
 - write a paper



The workflow of a data driven project

WHAT ARE THEIR BIGGEST CHALLENGES?

Dirty data is the #1 hurdle for data scientists.



66.7%

of data scientists say cleaning and organizing data is their most time-consuming task.

52.3%

of data scientists cite poor quality data as their biggest daily obstacle.

60% of data scientists spend most of their time cleaning and labeling data.

57% said it was the least enjoyable thing they do

<https://visit.figure-eight.com/2015-data-scientist-report>

- DATA WRANGLING
 - cleaning datasets
 - data aggregation (joining datasets)
- DATA EXPLORATION
 - understanding structure in the data
 - understanding correlation
 - machine learning (clustering, dimensionality reduction?)

The workflow of a data driven project



60% of data scientists spend most of their time cleaning and labeling data.

57% said it was the least enjoyable thing they do

<https://visit.figure-eight.com/2015-data-scientist-report>

- DATA WRANGLING
 - cleaning datasets
 - data aggregation (joining datasets)
- DATA EXPLORATION
 - understanding structure in the data
 - understanding correlation
 - machine learning (clustering, dimensionality reduction?)

google is your friend!

how do i remove values smaller than something in a dataframe

All Images Videos Shopping News More Settings

About 880,000 results (0.58 seconds)

stackoverflow.com › questions › remove-rows-in-pytho... ▾

[Remove rows in python less than a certain value - Stack ...](#)

Jan 23, 2017 - I feel like this question must have been answered by someone before, but I can't find an answer on stack overflow! I have a **dataframe** result that ...

2 answers

The workflow of a data driven project

WHAT ARE THEIR BIGGEST CHALLENGES?

Dirty data is the #1 hurdle for data scientists.



66.7%

of data scientists say cleaning and organizing data is their most time-consuming task.

52.3%

of data scientists cite poor quality data as their biggest daily obstacle.

60% of data scientists spend most of their time cleaning and labeling data.

57% said it was the least enjoyable thing they do

<https://visit.figure-eight.com/2015-data-scientist-report>

- DATA WRANGLING
 - cleaning datasets
 - data aggregation (joining datasets)
- DATA EXPLORATION
 - understanding structure in the data
 - understanding correlation
 - machine learning (clustering, dimensionality reduction?)

stackoverflow is your friend!

certain value

Asked 3 years, 7 months ago Active 2 years, 3 months ago Viewed 54k times

I feel like this question must have been answered by someone before, but I can't find an answer on stack overflow!

25

I have a dataframe `result` that looks like this and I want to remove all the values *less than or equal to* 10

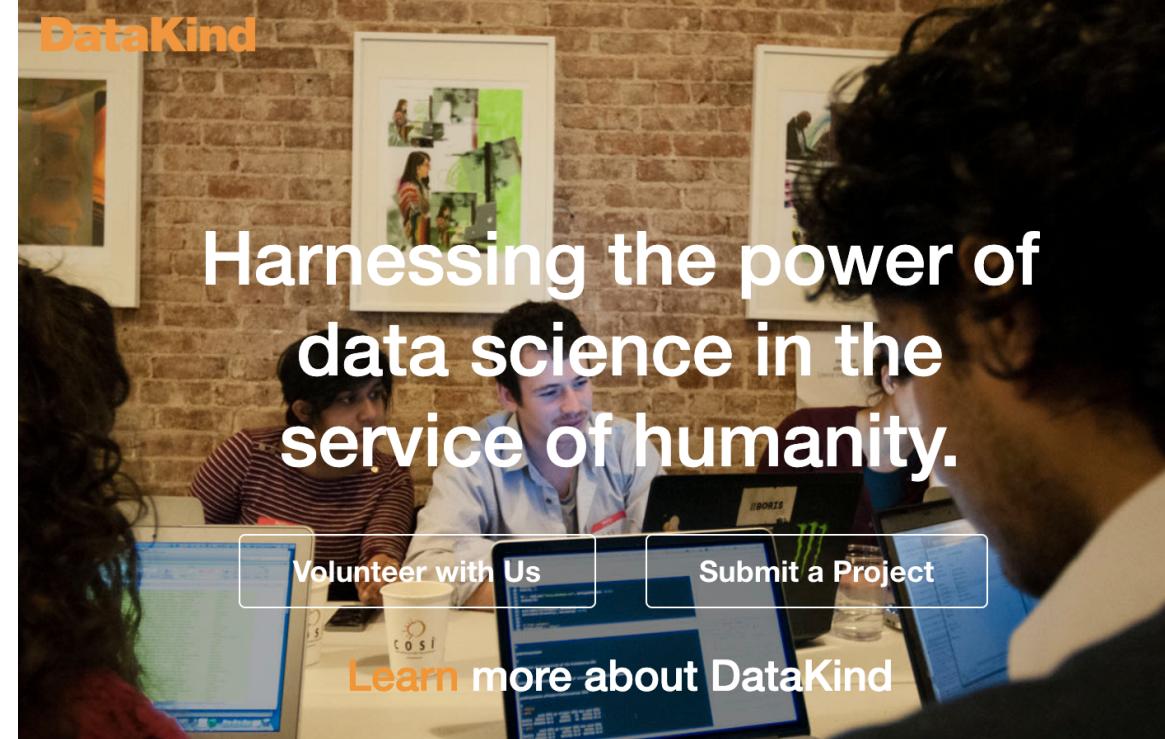
7

1

>>> result	Name	Value	Date
189	Sall	19.0	11/14/15
191	Sam	10.0	11/14/15
192	Richard	21.0	11/14/15
193	Ingrid	4.0	11/14/15

(sometime it may not seem like it is... its a forum. it may be toxic)

Advantage of bringing DS into public policy



Who Are We?

Data For Good is a collective of do gooders, who want to use their powers for good, and not evil, to help make our communities better through data. We are a national not for profit organization, with chapters across the country, that help other not for profit, and non-governmental, organizations harness the power of their data to make more informed and better decisions in their quest to make their communities flourish.

[DOWNLOAD OUR INFORMATION SHEET.](#)

Advantage of bringing DS into public policy

Kinder Institute



The Kinder Institute for Urban Research

Building Better Cities

Building Better Lives

The Kinder Institute for Urban Research builds better cities and improves people's lives by bringing together data, research, engagement and action.

NEW YORK CITY'S LEADER IN URBAN INFORMATICS

A UNIQUE RESEARCH CENTER THAT USES NEW YORK CITY AS ITS LABORATORY TO HELP CITIES

Carnegie Mellon University

Search

Metro21: Smart Cities Institute



both roles need a solid understanding of DS

we know that these
risks exist

what we do not fully
understand is how to
mitigate them

US House of Representatives
Science, Space, and Tech
committee chairperson

Eddie Bernice Johnson (D-TX)

https://www.youtube.com/embed/rXyR0_gFOM?start=403&enablejsapi=1

both roles need a solid understanding of DS

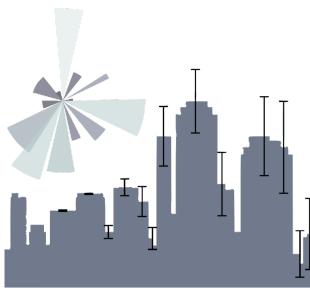
we know that these
risks exist

what we do not fully
understand is how to
mitigate them

US House of Representatives
Science, Space, and Tech
committee chairperson

Eddie Bernice Johnson (D-TX)

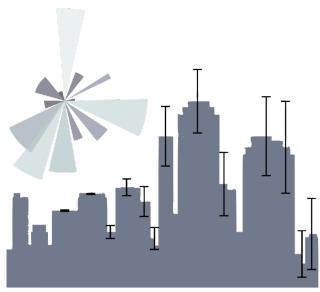
<https://www.youtube.com/embed/BcgDvEdGEXg?start=613&enablejsapi=1>



2 what is a scientific theory?

- # The workflow of a data driven project
- IDEA
 - formulate a hypothesis
 - DATA IDENTIFICATION
 - define ideal data
 - find a suitable dataset
 - figure out best data available
 - figure out if you can get new data
 - obtain data (including policy issues + technical issues)
 - create new data
 - survey
 - plan a survey
 - run a survey (including getting permissions, IRB, figuring out a distribution strategy...)
 - set up a remote sensing facility
 - buy instruments
 - deploy instruments (including getting permissions...)
 - connect instruments
 - run instruments
 - DATA WRANGLING
 - cleaning datasets
 - data aggregation (joining datasets)
 - DATA EXPLORATION
 - understanding structure in the data
 - understanding correlation
 - machine learning (clustering, dimensionality reduction?)
 - PREDICTING
 - models (e.g. regression)
 - prediction
 - validation (simulations)
 - transferability
 - INTERPRETATION
 - feature importance
 - mechanistic models
 - RESULT COMMUNICATION
 - visualizations
 - write a report
 - write a paper

the "demarcation" problem?

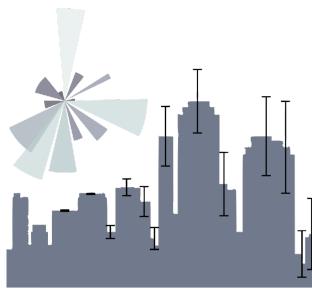


A scientific theory must be *falsifiable*

My proposal is based upon an **asymmetry** between verifiability and **falsifiability**; an asymmetry which results from the logical form of universal statements. For these are never derivable from singular statements, but can be contradicted by singular statements.

—Karl Popper, *The Logic of Scientific Discovery*

the "demarcation" problem?

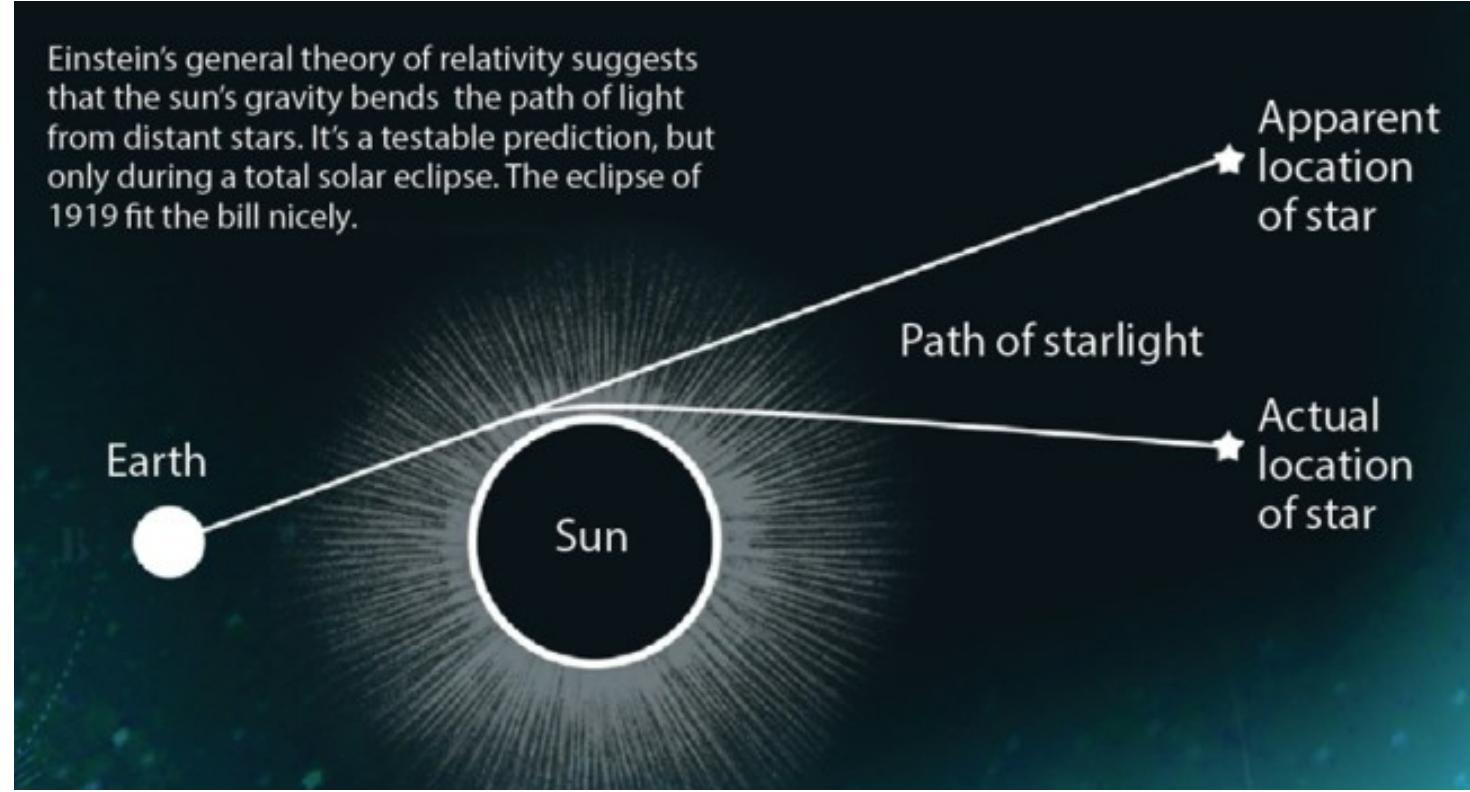


Einstein's General Relativity

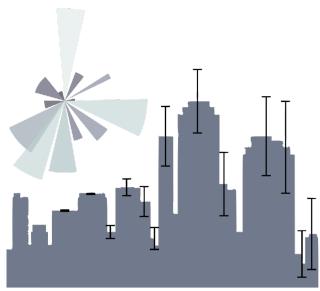
theory is falsifiable:

hypothesis: if GR is true the (apparent) position of a star would change during an eclipse

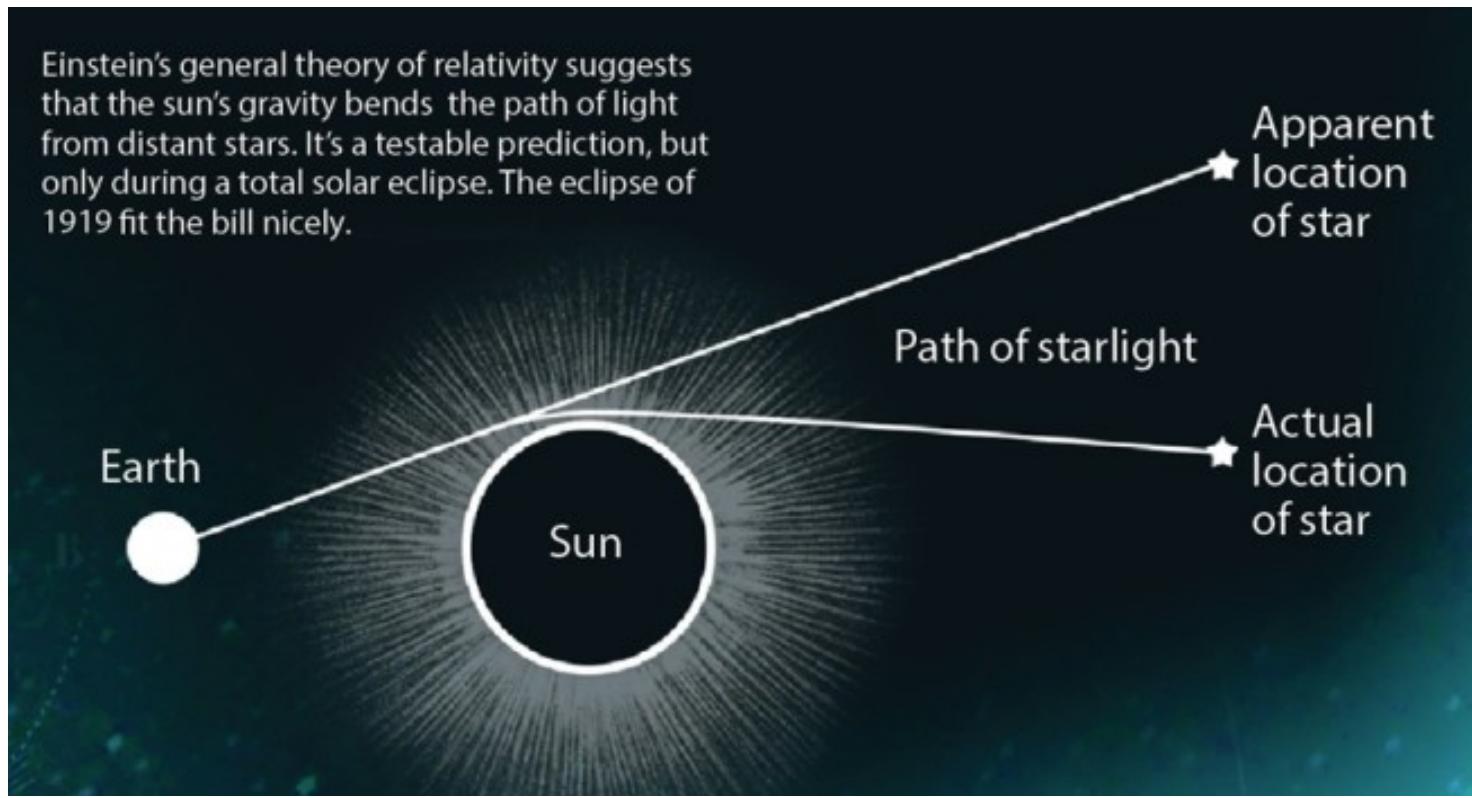
if the position of the star during the eclipse did not change the theory would be falsified.



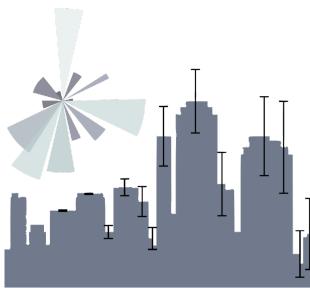
the "demarcation" problem?



*Can you think of a theory
or even a field
that is not falsifiable?*

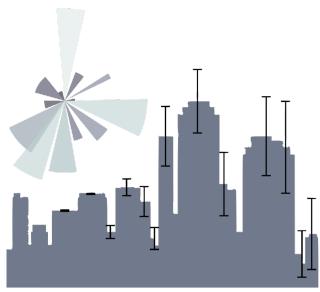


Einstein's general theory of relativity suggests that the sun's gravity bends the path of light from distant stars. It's a testable prediction, but only during a total solar eclipse. The eclipse of 1919 fit the bill nicely.



3 probabilistic induction, falsifiability

the "demarcation" problem?

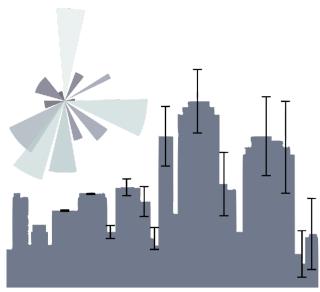


A scientific theory must be *falsifiable*

While limits and accuracy claims are not necessarily part of the theory proper, they are part of the theory as actually taught and applied. Indeed, although people try to extrapolate, one can never be sure whether a theory is correct outside the domain where the data were collected.

A. Neumaier

the "demarcation" problem?



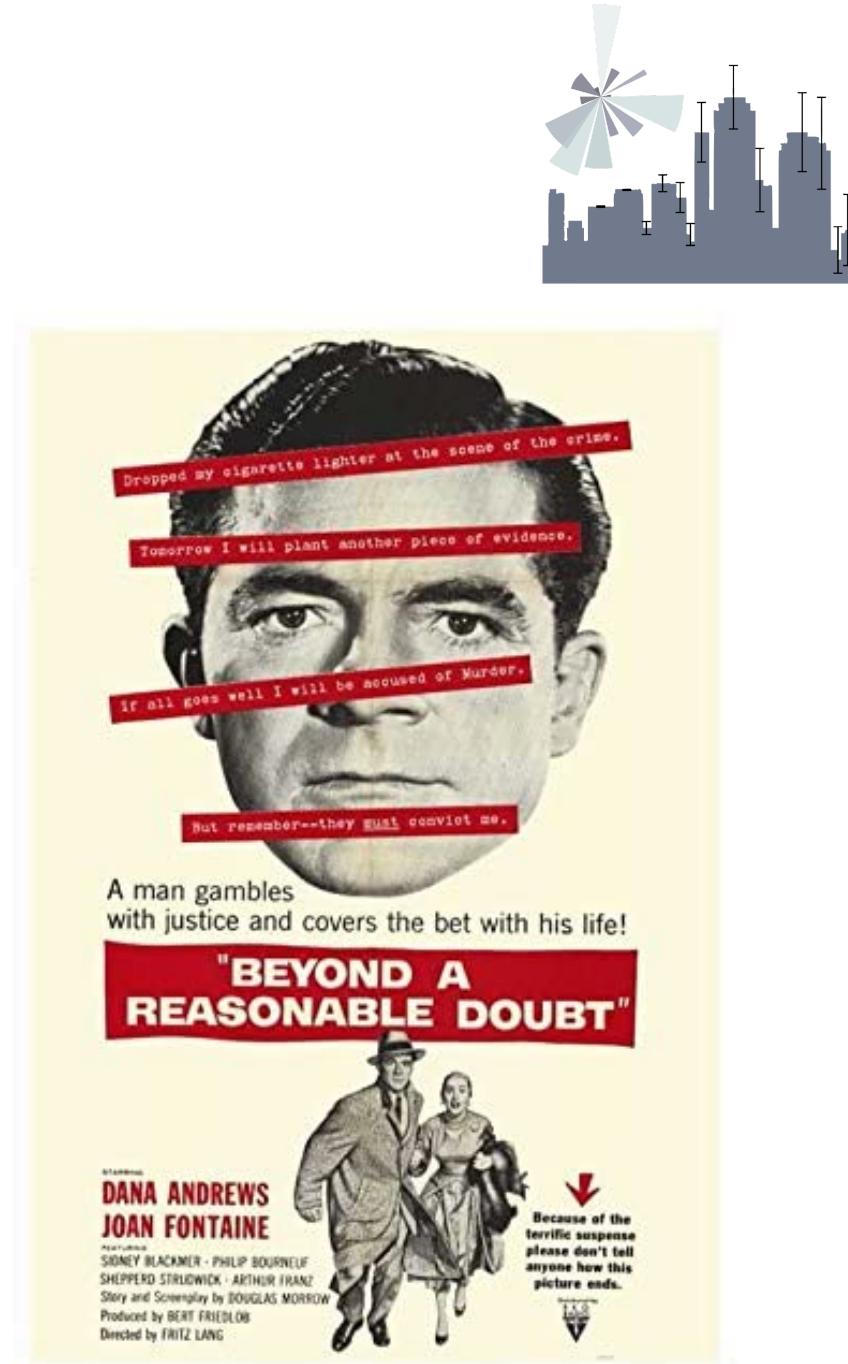
A scientific theory must be *falsifiable*

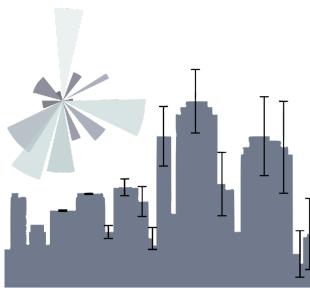
by most scientific theories are actually based largely on
probabilistic induction and
modern inductive inference (Solomonoff, frequentist vs
Bayesian methods...)

the "demarcation" problem?

The Falsifiability principle is common our own society!

proving beyond a reasonable doubt means falsifying (to a reasonable extent) any alternative explanation



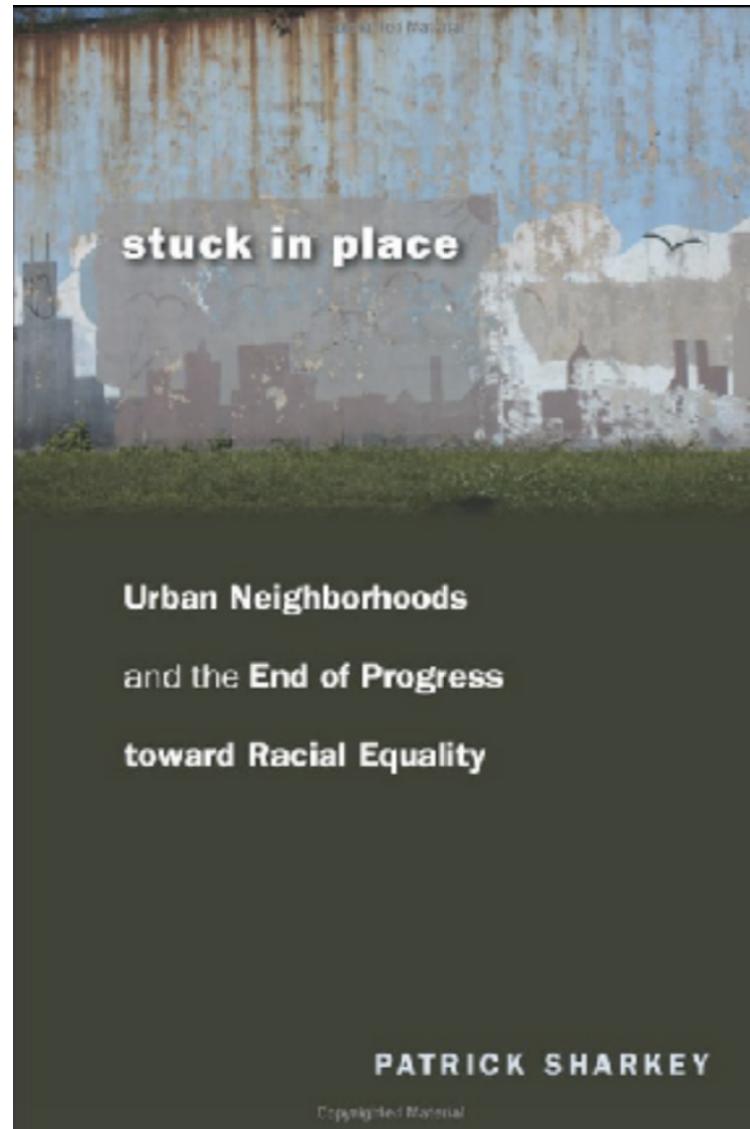


4

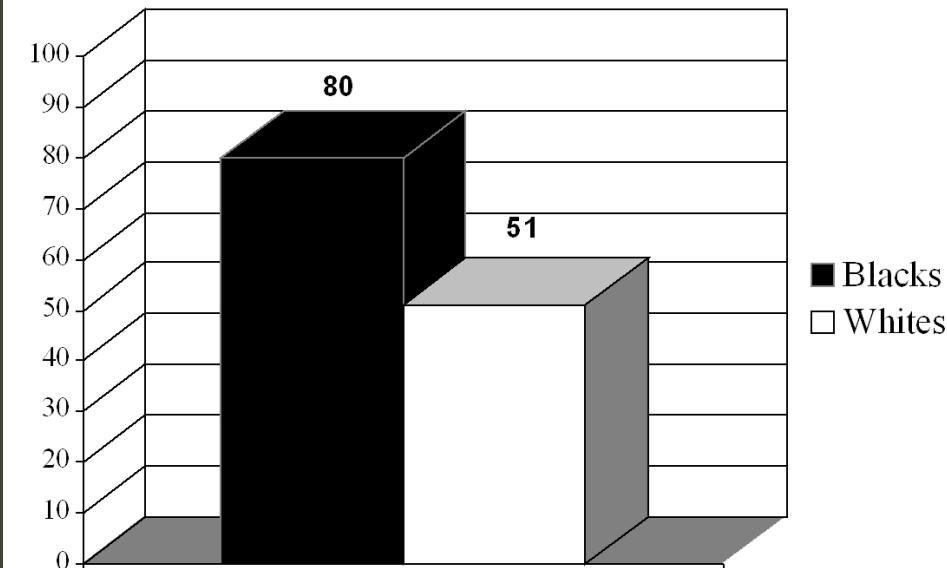
from idea to hypothesis

data-drive measures of social mobility

OBSERVATION:
there a
persistent
disparity in the
distribution of
wealth by race



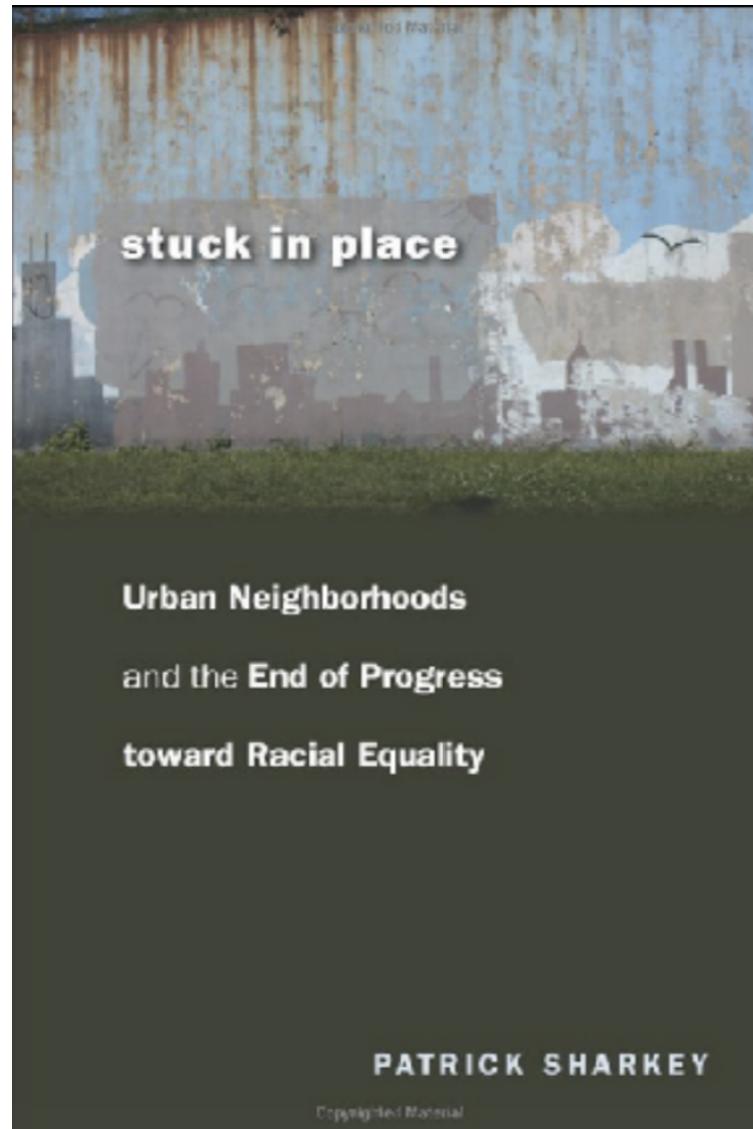
% of individuals in poor neighborhoods whose parent was also raised in a poor neighborhood



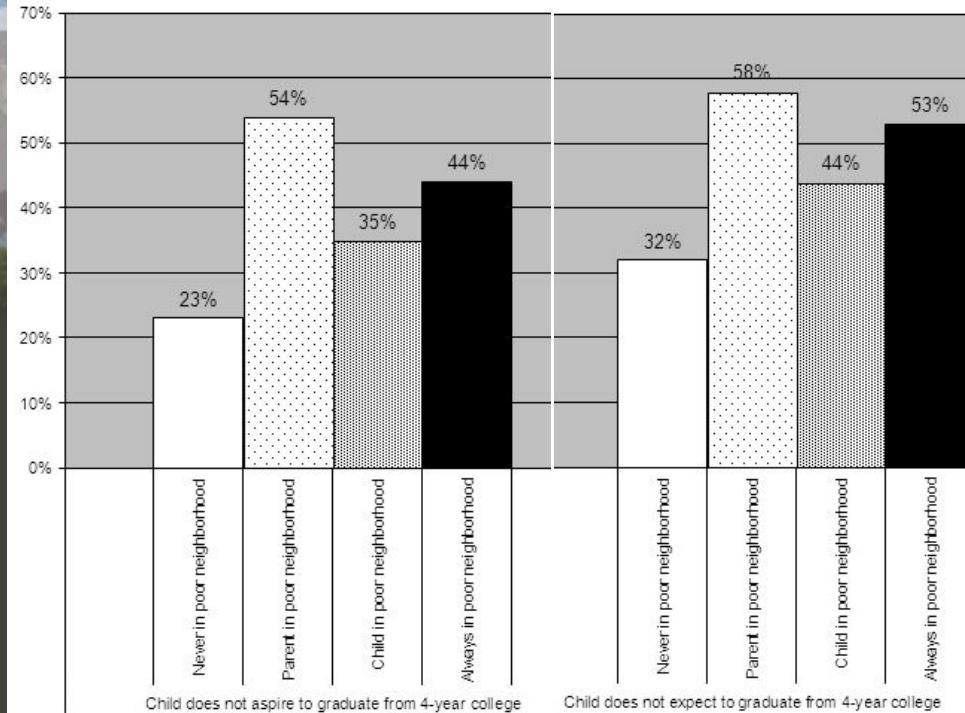
data-drive measures of social mobility

QUESTION: why is the disparity persistent?

the goal is to formulate a *hypothesis* that can be *falsified*



The consequences of multigenerational neighborhood disadvantage: Educational aspirations



formulation of a scientific hypothesis

QUESTION: does
proximity to violence
affect children's
development ?

the goal is to formulate a
hypothesis
that can be
falsified

- 1) identify a *measurable* variable related to the effect you propose exists
- 2) identify which quantity (*statistics*) of the measurable variable is appropriate to identify an effect (e.g. the average, the spread, the distribution shape?)
- 3) identify to what *significance* you expect to measure the effect in a probabilistic sense

formulation of a scientific hypothesis

QUESTION: does
proximity to violence
affect children's
development ?

the goal is to formulate a
hypothesis
that can be
falsified

reading test
score

- 1) identify a *measurable* variable related to the effect you propose exists
- 2) identify which quantity (*statistics*) of the measurable variable is appropriate to identify an effect (e.g. the average, the spread, the distribution shape?)
- 3) identify to what *significance* you expect to measure the effect in a probabilistic sense

formulation of a scientific hypothesis

QUESTION: does
proximity to violence
affect children's
development ?

the goal is to formulate a
hypothesis
that can be
falsified

- 1) identify a *measurable* variable related to the effect you propose exists
- 2) identify which quantity (*statistics*) of the measurable variable is appropriate to identify an effect (e.g. the average, the spread, the distribution shape?)
- 3) identify to what *significance* you expect to measure the effect in a probabilistic sense

reading test
score

the average
reading test score
of children **within**
2 blocks of a crime

formulation of a scientific hypothesis

QUESTION: does
proximity to violence
affect children's
development ?

the goal is to formulate a
hypothesis
that can be
falsified

- 1) identify a *measurable* variable related to the effect you propose exists
- 2) identify which quantity (statistics) of the measurable variable is appropriate to identify an effect (e.g. the average, the spread, the distribution shape?)
- 3) identify to what *significance* you expect to measure the effect in a probabilistic sense

reading test
score

the average
reading test score
of children **within**
2 blocks of a crime

p-value
(we'll talk about it
next week)

formulation of a scientific hypothesis

QUESTION: does proximity to violence affect children's development ?

the goal is to formulate a *hypothesis* that can be *falsified*

FALSIFIABILITY:

Instead of directly testing our hypothesis, we test the opposite (*complement*) of it.

We call that the *Null hypothesis (H₀)*

We call the hypothesis originating from our original idea *Alternative hypothesis (H_a)*

If we can reject the Null hypothesis, the Alternative Hypothesis holds (for now).

formulation of a scientific hypothesis

QUESTION: does proximity to violence affect children's development ?

the goal is to formulate a *hypothesis* that can be *falsified*

IDEA:

violence may affect learning outcomes

ALTERNATIVE HYPOTHESIS H_a :

the ***average test score*** of children who live within 2 blocks of the site of a violent crime is ***significantly lower*** up to 30 days after the crime

NULL HYPOTHESIS:

the ***average reading test score*** of children who live *within a block of the site of a violent crime* is the ***same or higher*** than the average score for the control group within 30 days following the crime, ***significance level $p=0.05$***

formulation of a scientific hypothesis

QUESTION: does proximity to violence affect children's development ?

the goal is to formulate a *hypothesis* that can be *falsified* a measurable quantity

IDEA:

violence may affect learning outcomes

ALTERNATIVE HYPOTHESIS H_a :

the **average test score** of children who live within 2 blocks of the site of a violent crime is **significantly lower** up to 30 days after the crime

NULL HYPOTHESIS:

the *average reading test score* of children who live *within a block of the site of a violent crime* is the **same or higher** than the average score for the control group within 30 days following the crime, *significance level p=0.05*

formulation of a scientific hypothesis

QUESTION: does proximity to violence affect children's development ?

the goal is to formulate a *hypothesis* that can be *falsified*
a statistics of the measurable quantity

IDEA:

violence may affect learning outcomes

ALTERNATIVE HYPOTHESIS H_a :

the **average test score** of children who live within 2 blocks of the site of a violent crime is **significantly lower** up to 30 days after the crime

NULL HYPOTHESIS:

the *average reading test score* of children who live *within a block of the site of a violent crime* is the **same or higher** than the average score for the control group within 30 days following the crime, *significance level p=0.05*

formulation of a scientific hypothesis

QUESTION: does proximity to violence affect children's development ?

the goal is to formulate a *hypothesis* that can be *falsified*

what we would expect to measure under the Alternative hypothesis *Ha* is

IDEA:

violence may affect learning outcomes

ALTERNATIVE HYPOTHESIS *Ha*:

the *average test score* of children who live within 2 blocks of the site of a violent crime is *significantly lower* up to 30 days after the crime

NULL HYPOTHESIS:

the *average reading test score* of children who live *within a block of the site of a violent crime* is the *same or higher* than the average score for the control group within 30 days following the crime, *significance level p=0.05*

formulation of a scientific hypothesis

QUESTION: does proximity to violence affect children's development ?

the goal is to formulate a *hypothesis* that can be *falsified*

what we are going to test is if the complement of our expectation is falsified

IDEA:

violence may affect learning outcomes

ALTERNATIVE HYPOTHESIS *Ha*:

the *average test score* of children who live within 2 blocks of the site of a violent crime is *significantly lower* up to 30 days after the crime

NULL HYPOTHESIS:

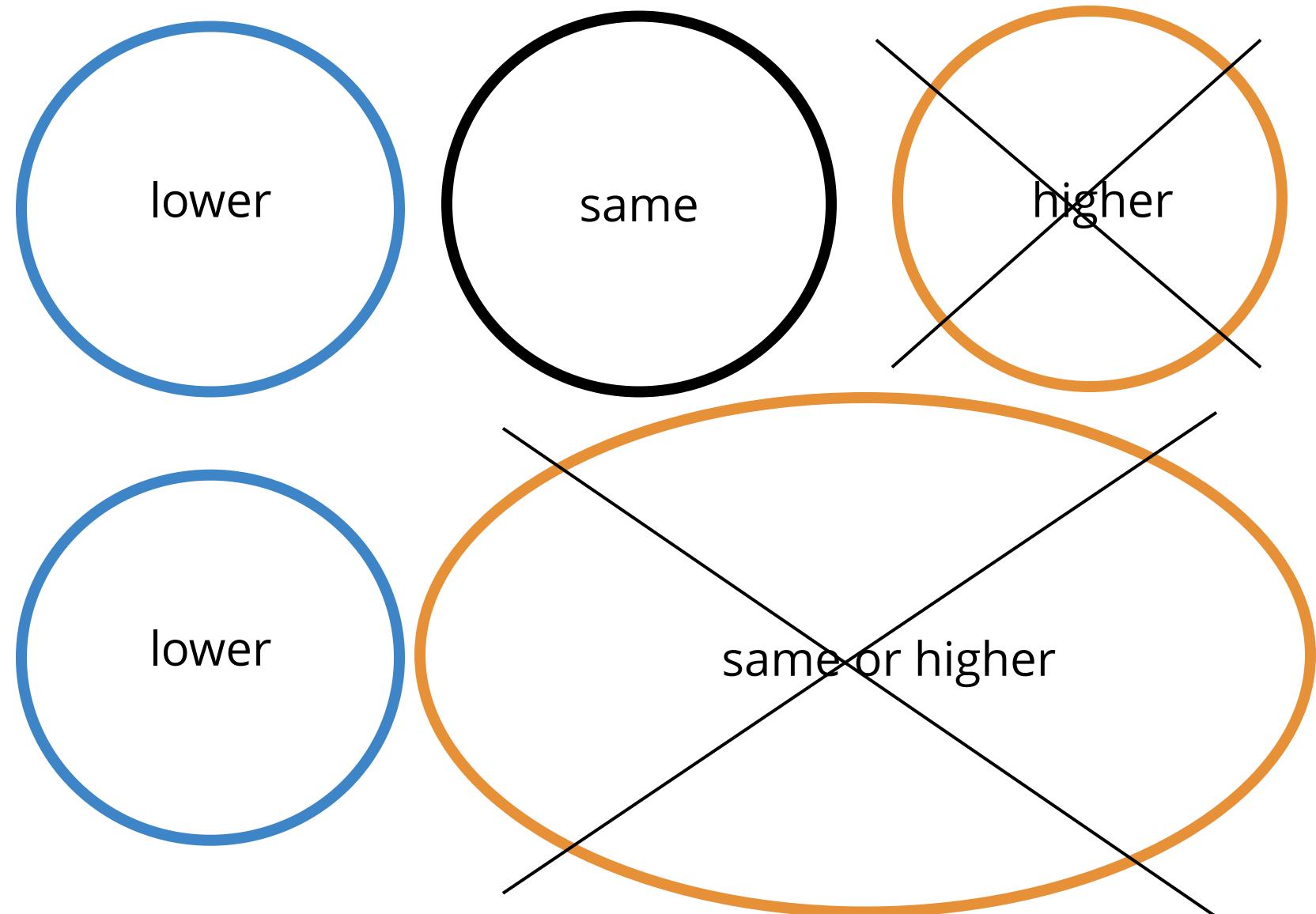
the *average reading test score* of children who live *within a block of the site of a violent crime* is the *same or higher* than the average score for the control group within 30 days following the crime, *significance level p=0.05*

formulation of a scientific hypothesis

QUESTION: does proximity to violence affect children's development ?

the goal is to formulate a *hypothesis* that can be *falsified*

what we are going to test is if the *complement* of our expectation is falsified



formulation of a scientific hypothesis

QUESTION: does proximity to violence affect children's development ?

the goal is to formulate a *hypothesis* that can be *falsified* **be specific!**

IDEA:

violence may affect learning outcomes

ALTERNATIVE HYPOTHESIS H_a :

the ***average test score*** of children who live within 2 blocks of the site of a violent crime is ***significantly lower*** up to 30 days after the crime

NULL HYPOTHESIS:

the ***average reading test score*** of children who live *within a block of the site of a violent crime* is the ***same or higher*** than the average score for the control group within 30 days following the crime, ***significance level $p=0.05$***

formulation of a scientific hypothesis

QUESTION: does proximity to violence affect children's development ?

the goal is to formulate a *hypothesis* that can be *falsified*

IDEA:

violence may affect learning outcomes

ALTERNATIVE HYPOTHESIS H_a :

the ***average test score*** of children who live within 2 blocks of the site of a violent crime is ***significantly lower*** up to 30 days after the crime

NULL HYPOTHESIS:

the ***average reading test score*** of children who live *within a block of the site of a violent crime* is the ***same or higher*** than the average score for the control group within 30 days following the crime, ***significance level $p=0.05$***

formulation of a scientific hypothesis

example: food deserts

<https://flowingdata.com/2013/08/27/in-search-of-food-deserts/>

practicing formulating hypothesis

general idea

1 response

food deserts are responsible for poor health outcomes in the US. where there is less availability of healthy food store outcomes may be worse.

Alternative Hypothesis

1 response

the incidence (number per person) of diabetes or obesity in areas where the closest food store is $d \geq 3$ miles (food desert) is higher than in non food deserts

Null Hypothesis

1 response

the incidence (number per person) of diabetes or obesity in areas where the closest food store is $d \geq 3$ miles (food desert) is lower or the same than in non food deserts

Are food deserts responsible for health issues?

count and category of the food stores by neighborhood and health diagnoses for citizens in each neighborhood

math formulation

1 response

$$H_0 : (N_d + N_o - N_{do})_{(d \geq 3 \text{ miles})} \leq (N_d + N_o - N_{do})_{(d < 3 \text{ miles})}$$

$$H_1 : (N_d + N_o - N_{do})_{(d \geq 3 \text{ miles})} > (N_d + N_o - N_{do})_{(d < 3 \text{ miles})}$$

comments on experimental design

1 response

food desert needs to be defined:

with the given data: number of health store per neighborhood 100 residents < 1
better: nearest health store is > 3 miles

health outcomes need to be defined: obesity and diabetes are related to diet.
target variable: N cases of obesity or diabetes in the neighborhood

issues:

- obesity and diabetes may be correlated : correct for covariance (hard)
- other health outcomes may be relevant
- there may be genetic predisposition : common solution split by race
- transportation may matter (3 miles may not be a lot in rural areas)
- diagnoses rate may be correlated to availability of resources including food stores

formulation of a scientific hypothesis

practicing formulating
hypothesis

the goal is to formulate a
hypothesis
that can be
falsified

pick 2

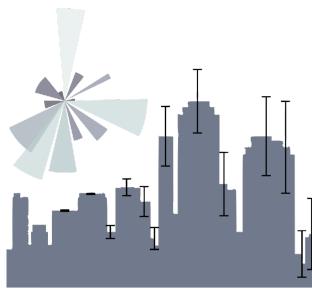
Does exercise help lose more weight than diet alone?
<http://bit.ly/htexercize>

Does smoking while pregnant affect children's
intellectual development? <http://bit.ly/htsmoking>

Do teens use cell phones to access the internet
more than adults? <http://bit.ly/htinternetaccess>

Are citibikes a threat to road safety 1: geographic?
<http://bit.ly/htcitibike1>

Are citibikes a threat to road safety 1: longitudinal?
<http://bit.ly/htcitibike2>

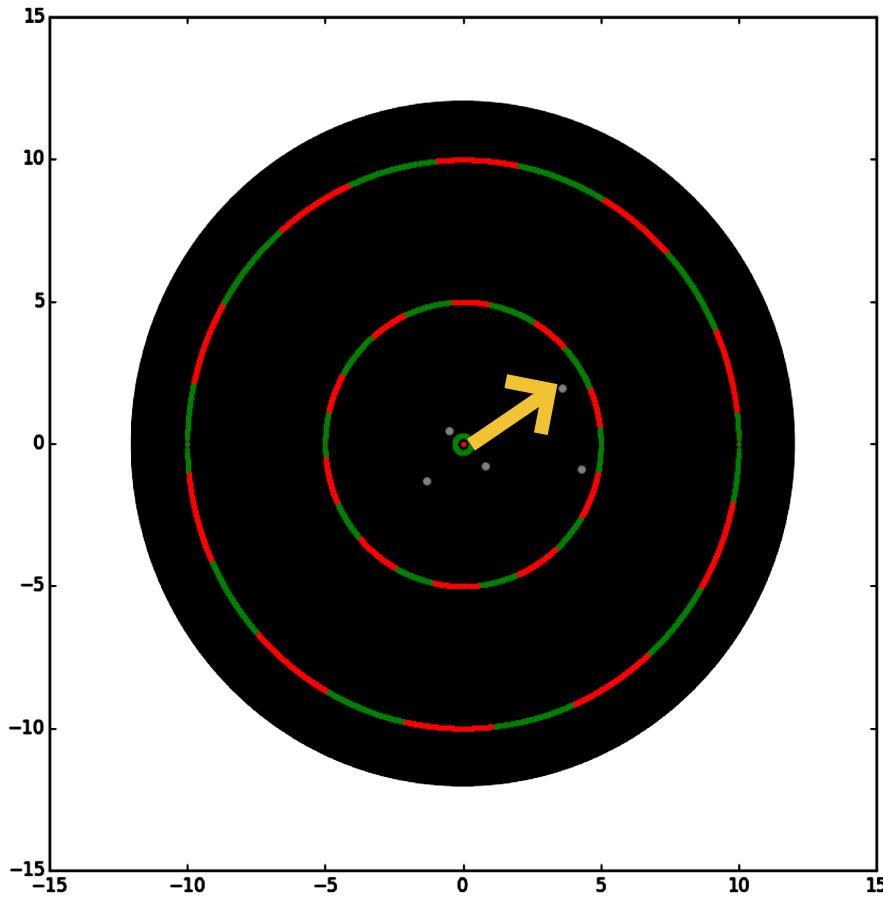


6

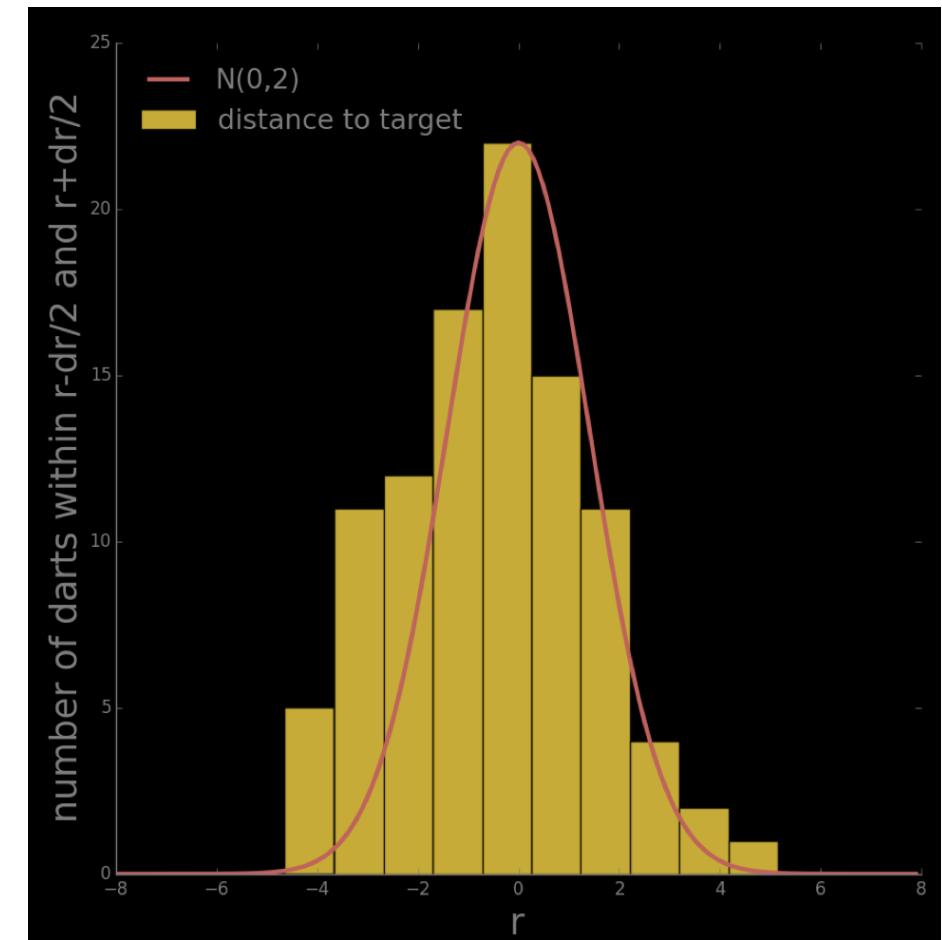
descriptive statistics

distribution

Consider the distance from the bull's eye:
its never 0 but, if you are a good dart
player, its more likely to be small than big

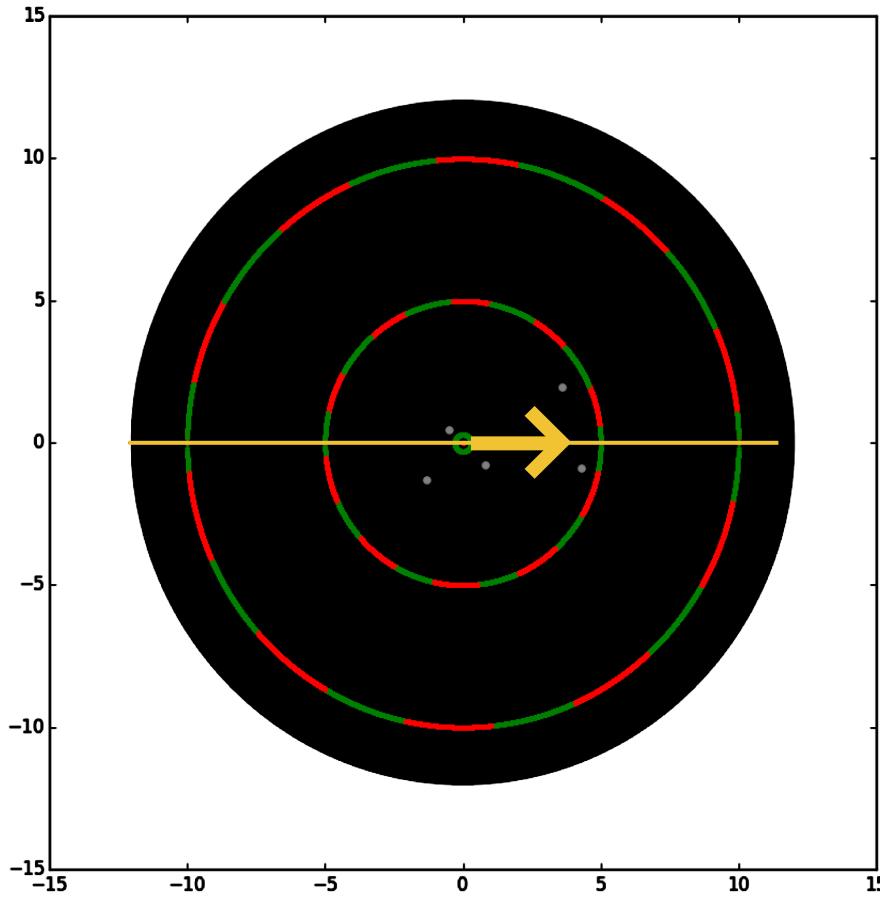


$$N(r|\mu, \sigma) \sim \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(r-\mu)^2}{2\sigma^2}}$$



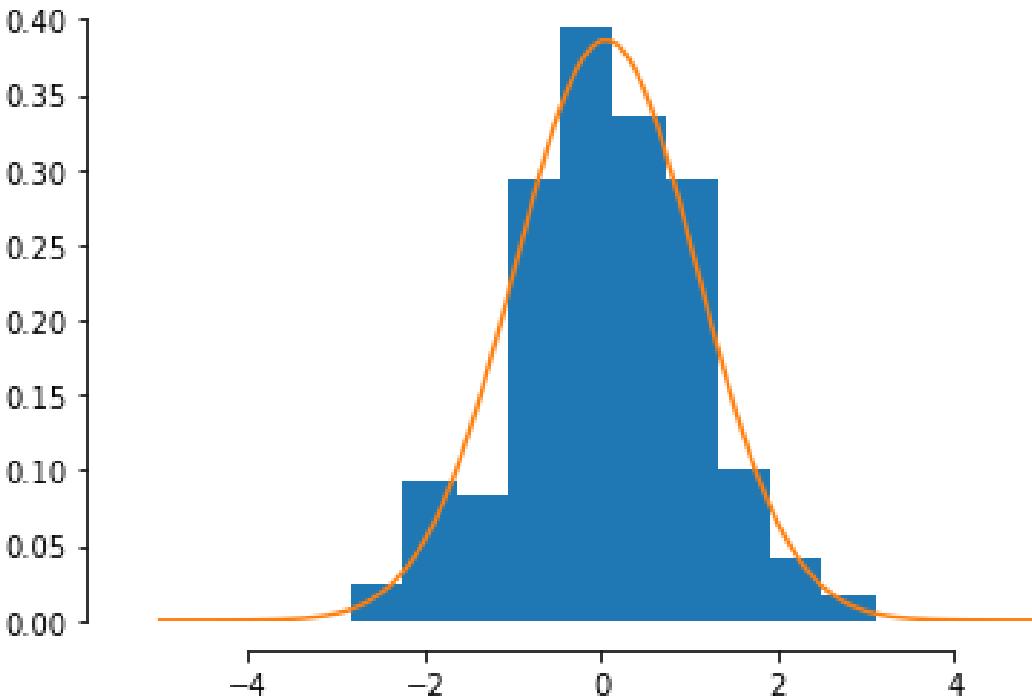
distributions

Consider the distance from the bull's eye:
its never 0 but, if you are a good dart
player, its more likely to be small than big



support parameters

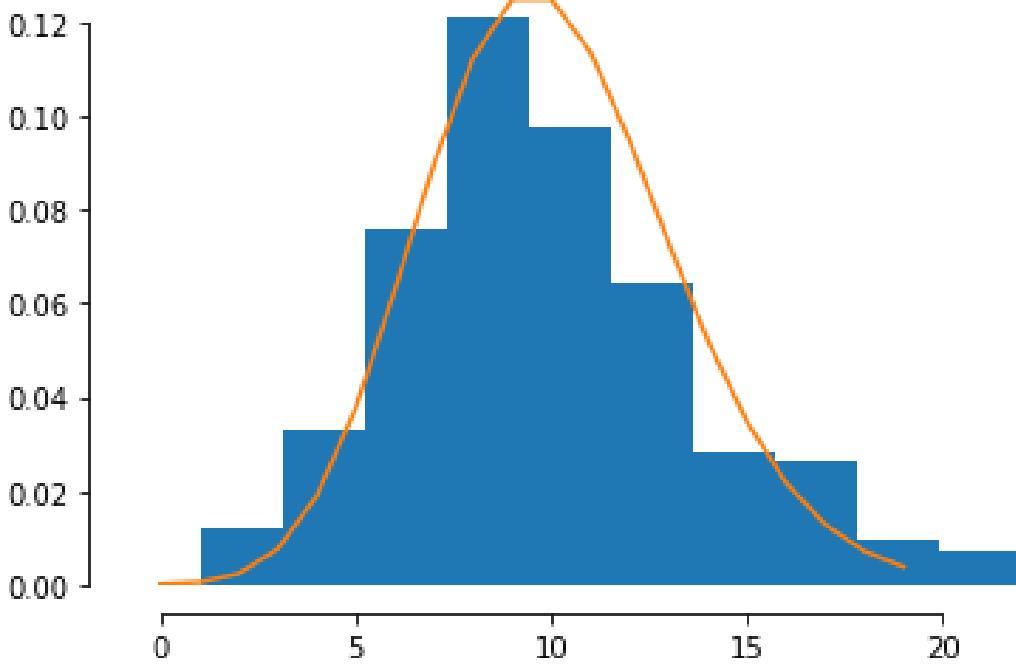
$$N(r|\mu, \sigma) \sim \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(r-\mu)^2}{2\sigma^2}}$$



distributions

parameters ($\lambda=10$)

$$P(k|\lambda) \sim \frac{\lambda^k e^{-\lambda}}{!k}$$

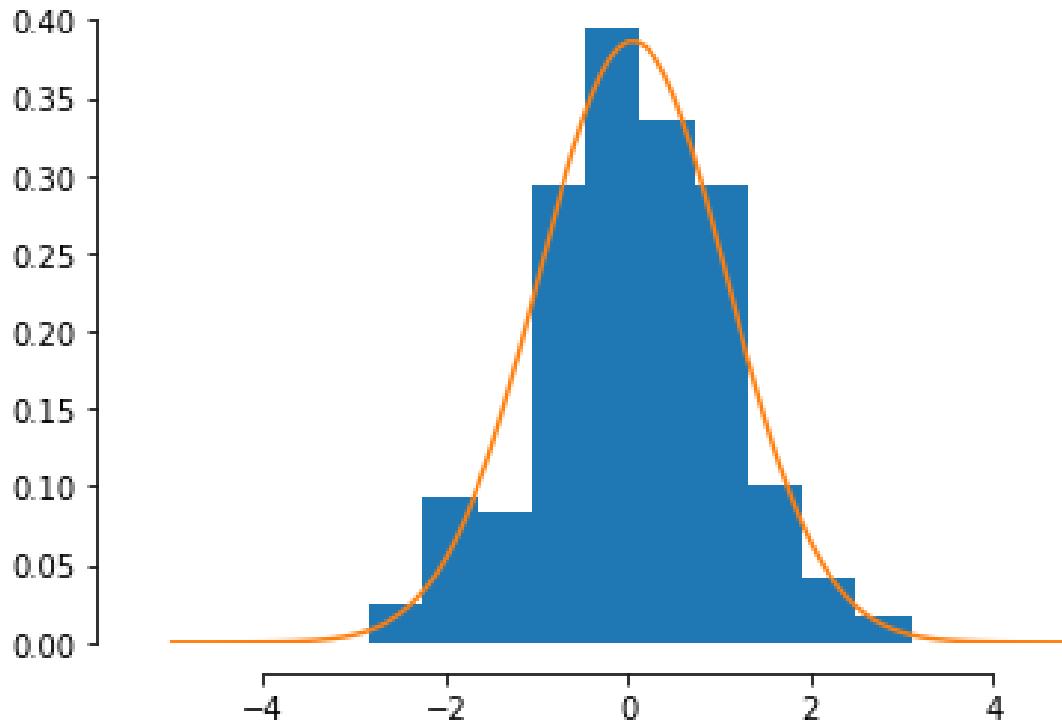


Poisson

discrete support $(1, +\infty]$

support parameters $(-0.1, 0.9)$

$$N(r|\mu, \sigma) \sim \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(r-\mu)^2}{2\sigma^2}}$$



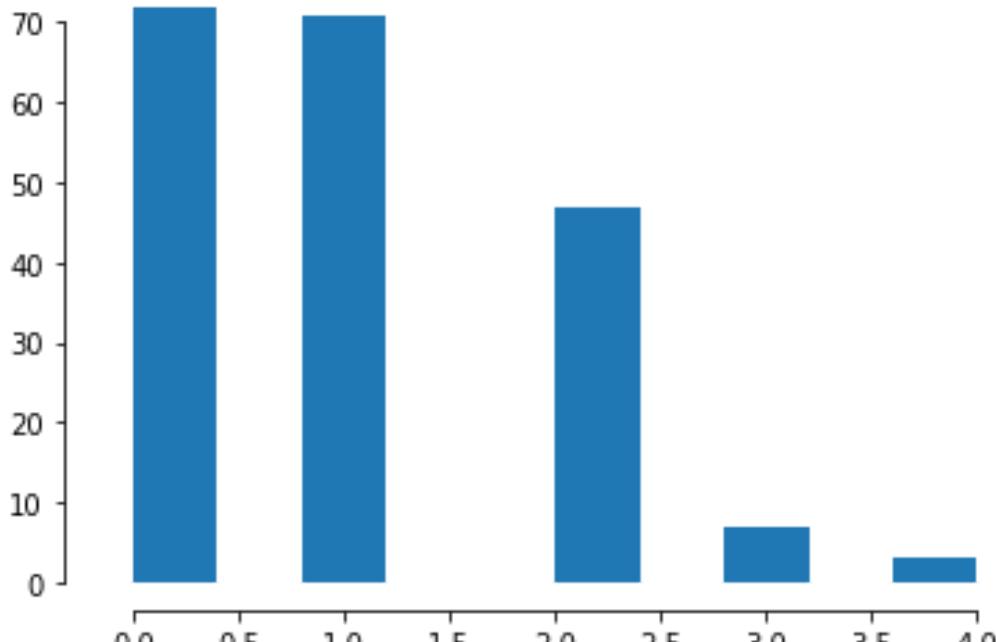
normal or Gaussian

continuous support $[-\infty, +\infty]$

distributions

parameters ($\lambda=1$)

$$P(k|\lambda) \sim \frac{\lambda^k e^{-\lambda}}{!k}$$

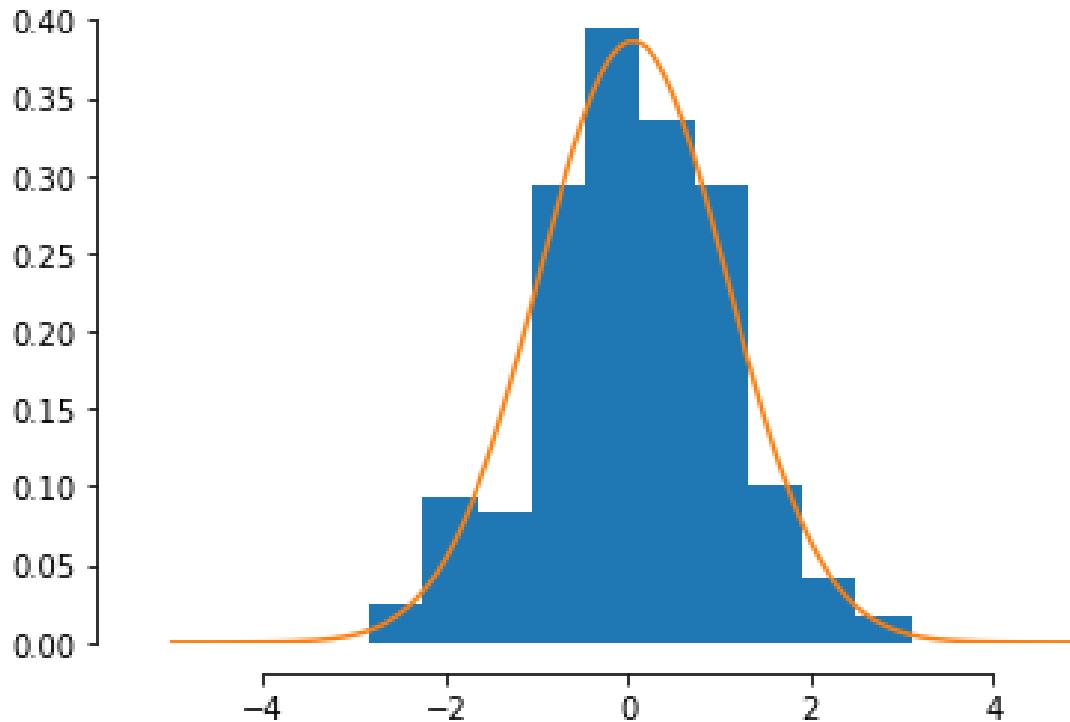


Poisson

discrete support $(1, +\infty]$

support parameters $(-0.1, 0.9)$

$$N(r|\mu, \sigma) \sim \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(r-\mu)^2}{2\sigma^2}}$$



normal or Gaussian

continuous support $[-\infty, +\infty]$

Moments and frequentist probability

a distribution's moments summarize its properties:

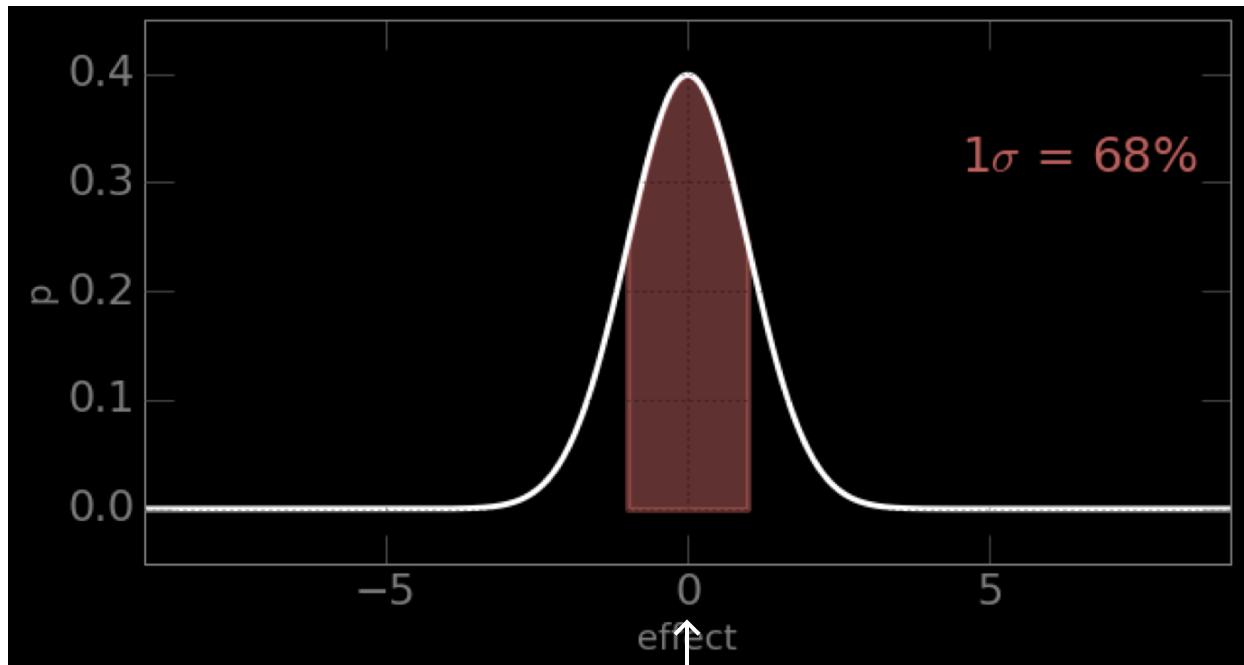
$$m_n = \int_{-\infty}^{\infty} (x - c)^n f(X) dx$$

central tendency: mean ($n=1$), median, mode

spread: standard deviation/variance ($n=2$), quartiles range

symmetry: skewness ($n=3$)

cuspiness: kurtosis ($n=4$)



Moments and frequentist probability

a distribution's moments summarize its properties:

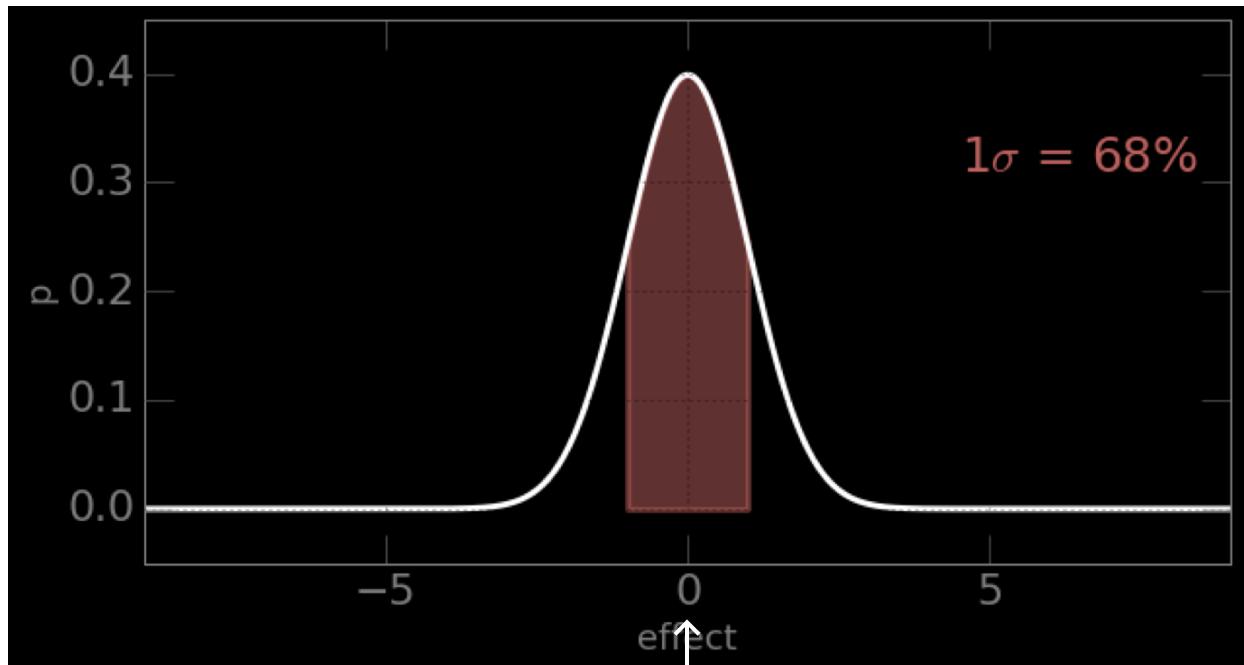
$$m_n = \int_{-\infty}^{\infty} (x - c)^n f(X) dx$$

central tendency: mean ($n=1$), median, mode

spread: standard deviation/variance ($n=2$), quartiles range

symmetry: skewness ($n=3$)

cuspiness: kurtosis ($n=4$)



Moments and frequentist probability

a distribution's moments summarize its properties:

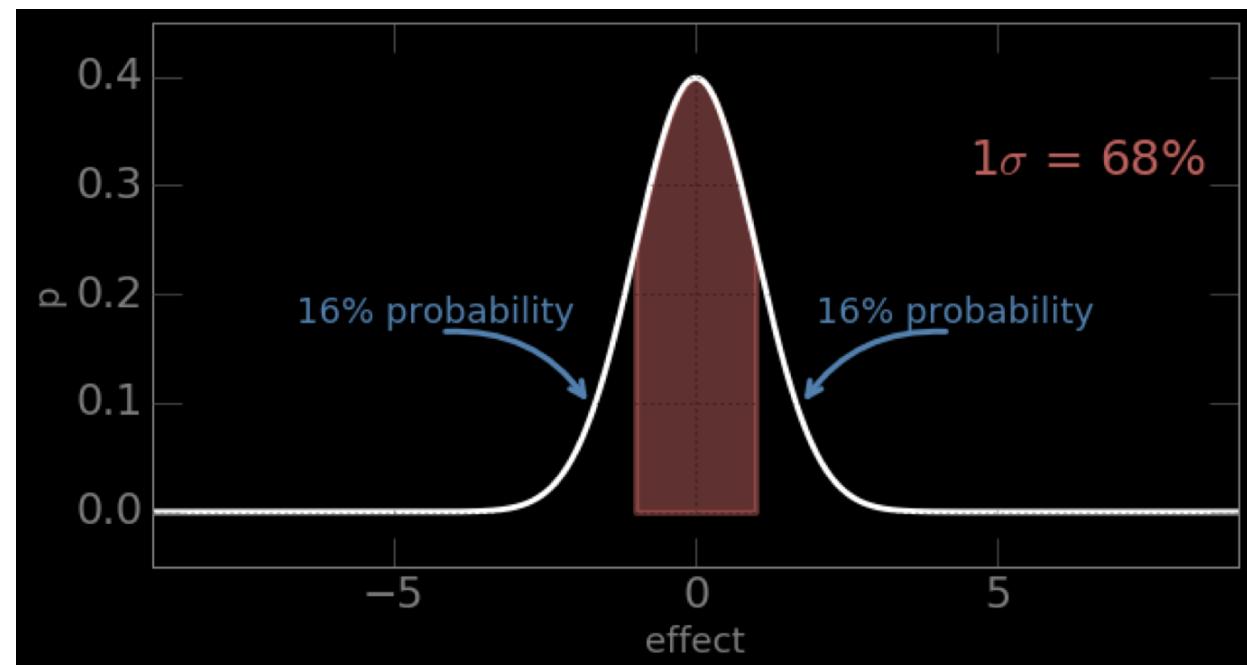
$$m_n = \int_{-\infty}^{\infty} (x - c)^n f(X) dx$$

central tendency: mean ($n=1$), median, mode

spread: standard deviation/variance ($n=2$), quartiles range

symmetry: skewness ($n=3$)

cuspiness: kurtosis ($n=4$)



Moments and frequentist probability

a distribution's moments summarize its properties:

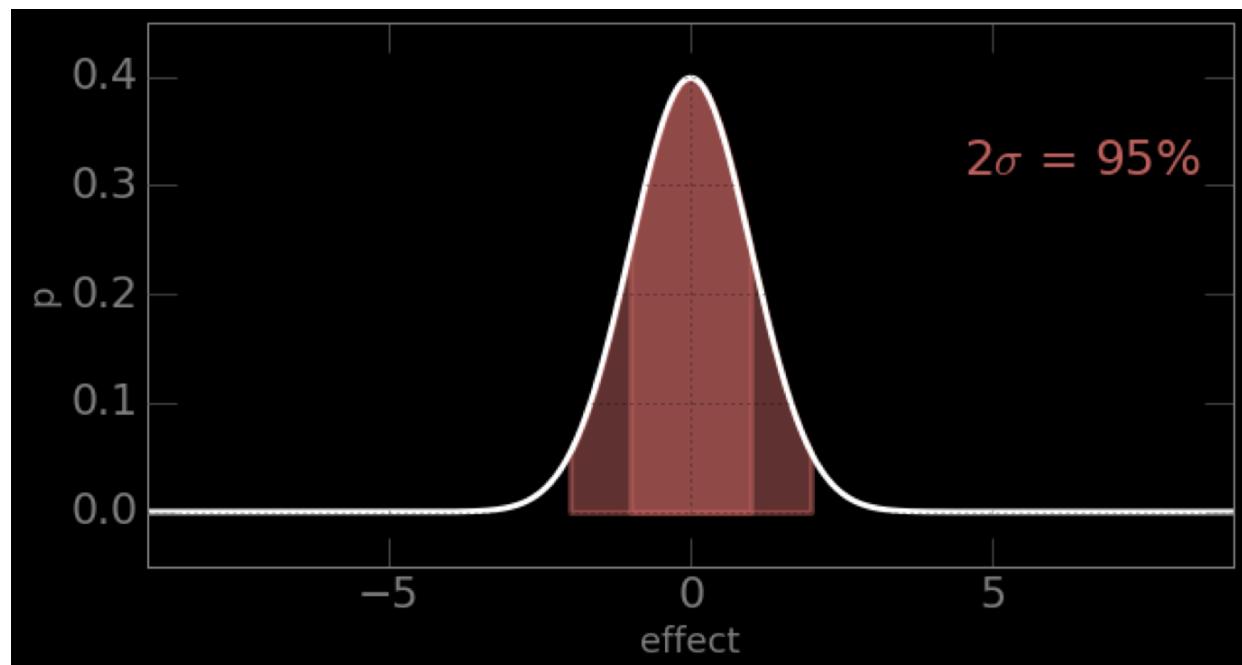
$$m_n = \int_{-\infty}^{\infty} (x - c)^n f(X) dx$$

central tendency: mean ($n=1$), median, mode

spread: standard deviation/variance ($n=2$), quartiles range

symmetry: skewness ($n=3$)

cuspiness: kurtosis ($n=4$)



Moments and frequentist probability

a distribution's moments summarize its properties:

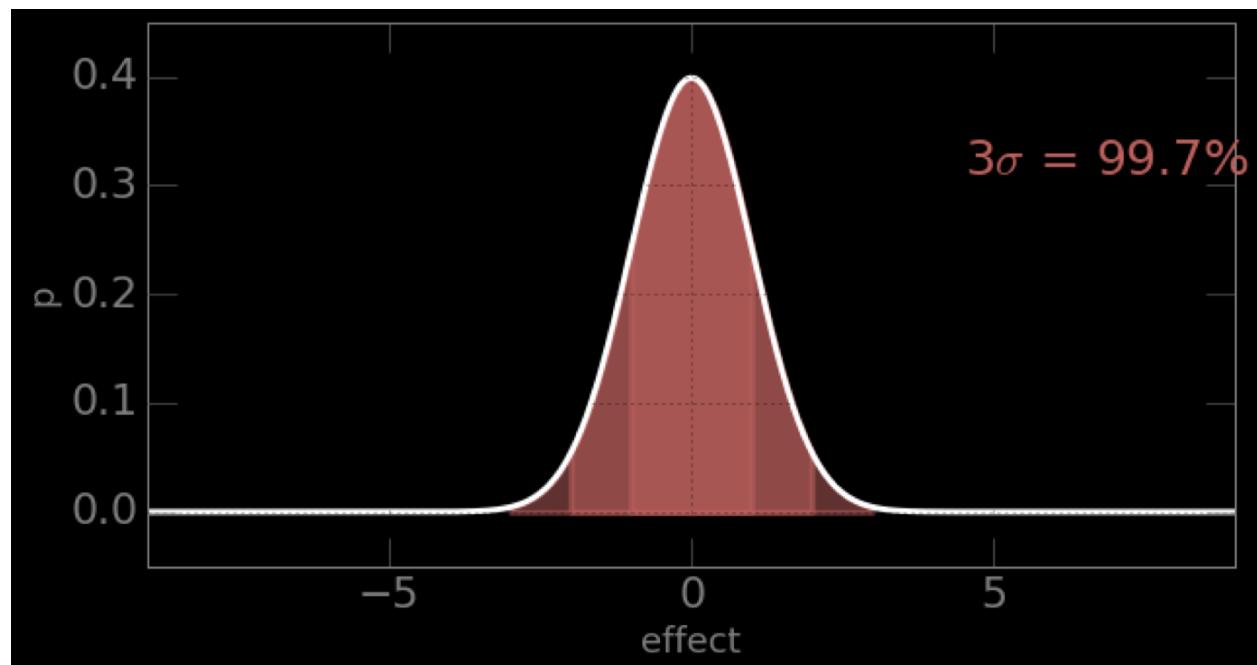
$$m_n = \int_{-\infty}^{\infty} (x - c)^n f(X) dx$$

central tendency: mean ($n=1$), median, mode

spread: standard deviation/variance ($n=2$), quartiles range

symmetry: skewness ($n=3$)

cuspiness: kurtosis ($n=4$)



Moments and frequentist probability

a distribution's moments summarize its properties:

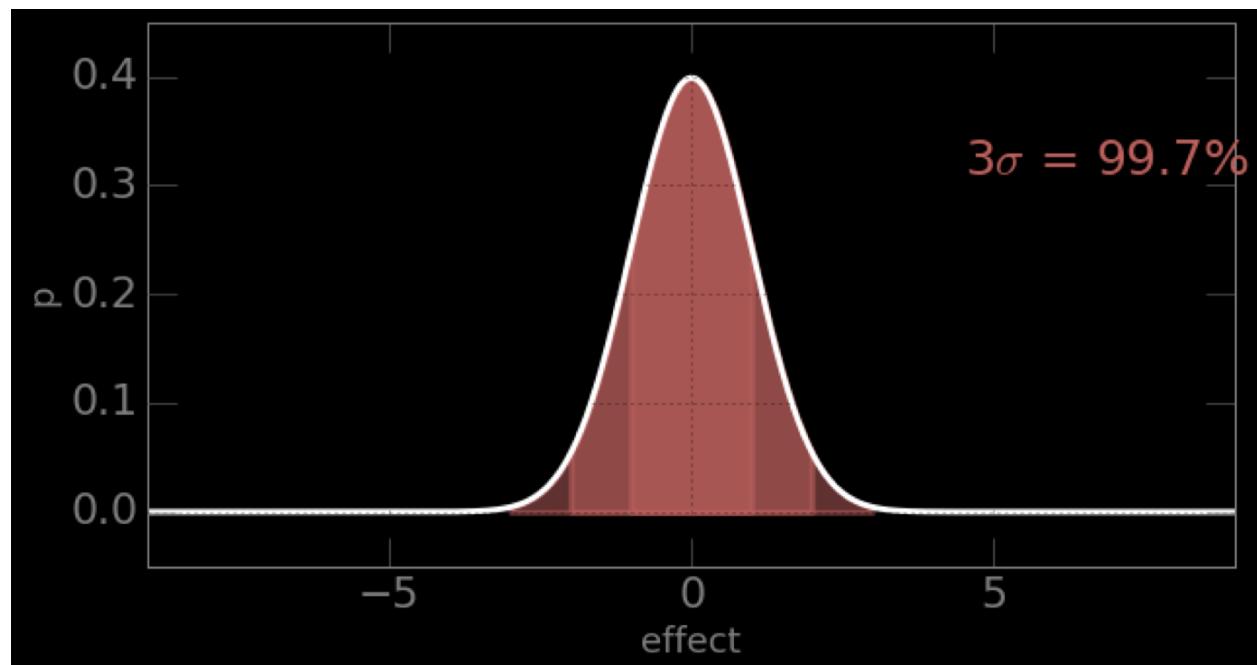
$$m_n = \int_{-\infty}^{\infty} (x - c)^n f(X) dx$$

central tendency: mean ($n=1$), median, mode

spread: standard deviation/variance ($n=2$), quartiles range

symmetry: skewness ($n=3$)

cuspiness: kurtosis ($n=4$)



key concepts

Applications of machine learning and data science to public policy enable the collection and interpretation of evidence for evidence based policy

AI and ML are pervasive technologies with that also pose ethical threats. An understanding of way machine learning works is necessary to sensibly regulate technology

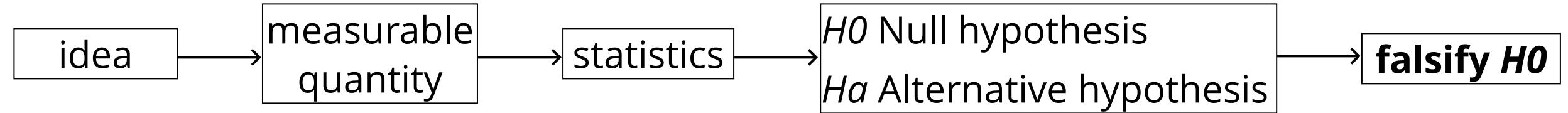
Data Science applications in public policy requires to import the scientific framework including the scientific way to ask questions, develop theories, validate results

Applications of machine learning and data science to public policy enable the collection and interpretation of evidence for evidence based policy

Falsification, rather than verification, is the hallmark of a scientific theory

key concepts

From idea to hypothesis



Gaussian and Poisson distribution

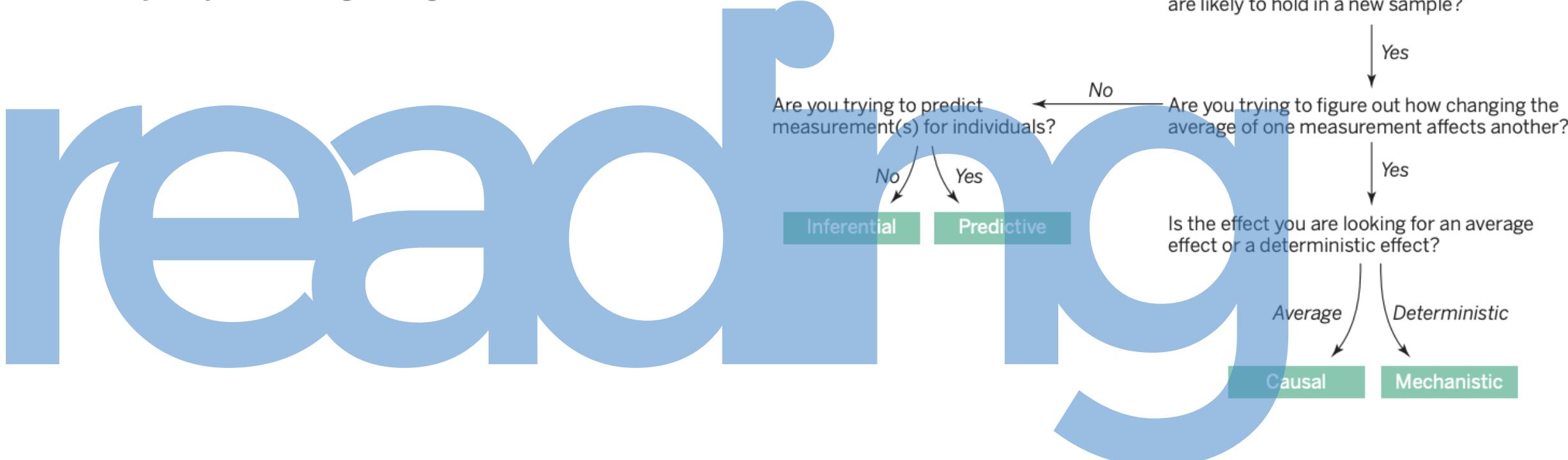
Moments of a distribution

STATISTICS

What is the question?

Mistaking the type of question being considered is the most common error in data analysis

By Jeffery T. Leek and Roger D. Peng





fedhere / phillyparks_instructions.ipynb

Secret

Created 26 seconds ago

Edit

Delete

Star

0

<> Code -o Revisions 2



Download ZIP

PhillyParks_instructions

phillyparks_instructions.ipynb



Raw

CO Open in Colab

HW2 : Summary statistics

this homework performs strictly an exploratory analysis of the data

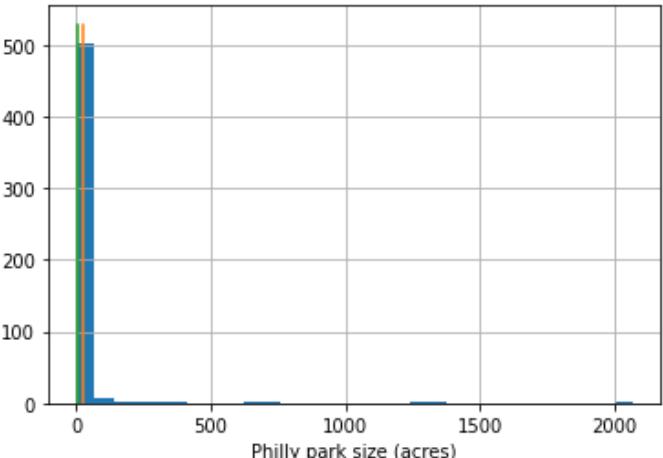
Consult the <https://www.cin.ufpe.br/~embat/Python%20for%20Data%20Analysis.pdf> Python for Data Analysis book here
(pandas book) chapter 7 or this digital rendition of its content here
https://github.com/fedhere/PUS2020_FBianco/blob/master/HW2/PandasCh7.ipynb.

Also use the code examples in the folder classdemo https://github.com/fedhere/PUS2020_FBianco/tree/master/classdemo

important turn in the notebook by uploading it to your github repository in a folder called HW2. Before you upload the notebook make sure you

- click above: Runtime -> Restart and run all

this will assure that your notebook has no bugs due to running the cells of code out of order.



https://github.com/fedhere/PUS2020_FBianco/blob/master/HW2/README.md