UiO : Department of informatics
University of Oslo

Adaptive Methods for Data-based Decision Making *Lecture 9*

IN-STK 5000 / 9000

Autumn 2022

Slides by Dr. Anne-Marie George, UiO

Please let me know what you think…

Leave your feedback *(for Anne-Marie's lectures)* on Flinga:

https://flinga.fi/s/FDN2DC3

- Were the contents understandable? Relevant? Interesting?
- How was the lecturing style?
- How was the teaching material?
- What could be improved?
- Any other comments?

IFI søker gruppelærere til våren 2023…

Arbeidsoppgaver

- Planlegging og fasilitering av gruppearbeid
- Svare på henvendelser fra studenter
- Retting av obligatoriske oppgaver
- Deltagelse på ukentlig gruppelærermøte

Kvalifikasjoner

- Du har tatt minst ett emne ved institutt for informatikk
- Ambisjoner om å bli en trygg og dyktig formidler
- Engasjement for faget ditt som du ønsker å dele med medstudenter

Vil tilbyr

- Opplæring og oppfølging
- En relevant og spennende deltidsjobb
- Kort vei mellom jobb og studiested
- Fleksibel arbeidstid

Lønn

- Bachelorstudenter u/gruppelærerkurs: 196,30 kr
- Bachelorstudenter m/gruppelærerkurs: 198,50 kr - 201,00 kr
- Masterstudenter u/gruppelærerkurs: 203,40 kr
- Masterstudenter m/gruppelærerkurs: 205,80 kr - 208,50 kr

What we talk about today

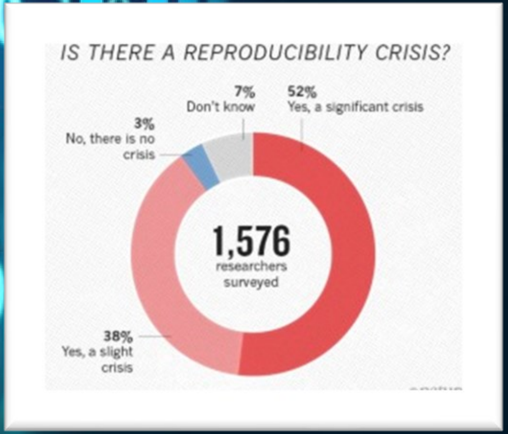
Reproducibility +

- Confidence Intervals
- Bandit Strategy: Upper Confidence Bound (UCB)
- Concentration Inequalities
- Confidence Intervals for ML Models
- Bootstrapping
- P-Hacking

Reproducibility

"Any **results should be documented** by making all **data and code available** in such a way that the computations **can be executed again** with identical results."

Source:
"The Ethical Algorithm" (Chapter 4) by Kearns & Roth



Source:
Monya Baker, *Nature*, 2016:
"1,500 scientists lift the lid on reproducibility"

The Reproducibility Crisis

Definition:

"The replication crisis (also called the replicability crisis and the reproducibility crisis) is an **ongoing methodological crisis** in which it has been found that the results of many scientific **studies are difficult or impossible to reproduce**. […] such failures undermine the credibility of theories building on them and potentially **call into question substantial parts of scientific knowledge**." - Wikipedia

Scope:

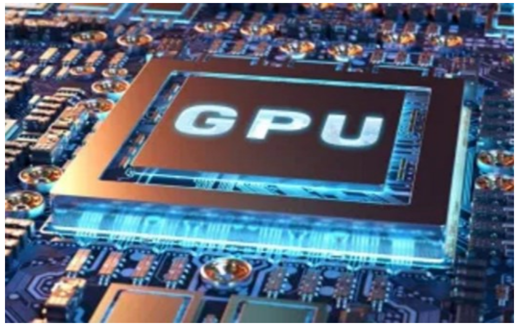
Monya Baker, *Nature*, 2016: "1,500 scientists lift the lid on reproducibility"
"More than 70% of researchers have tried and failed to reproduce another scientist's experiments, and more than half have failed to reproduce their own experiments.**"**

Causes:
p-hacking, poor study design /experimental technique, fraud and deception, "publish or parish" culture, …

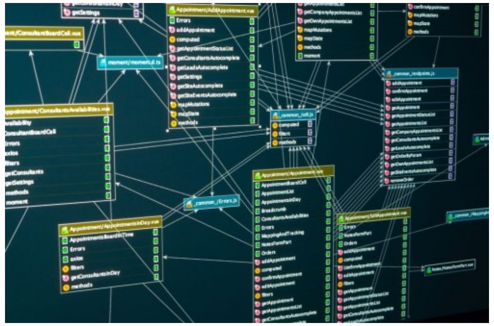
Things to keep in mind…

For someone else to be able to reproduce your results, you must publish:

Code + Documentation (e.g., code comments, parameter choices, random seed, …)



Publish specifications and versions of machines, packages / libraries etc. that were used



Publish data or data generation method and experiment design

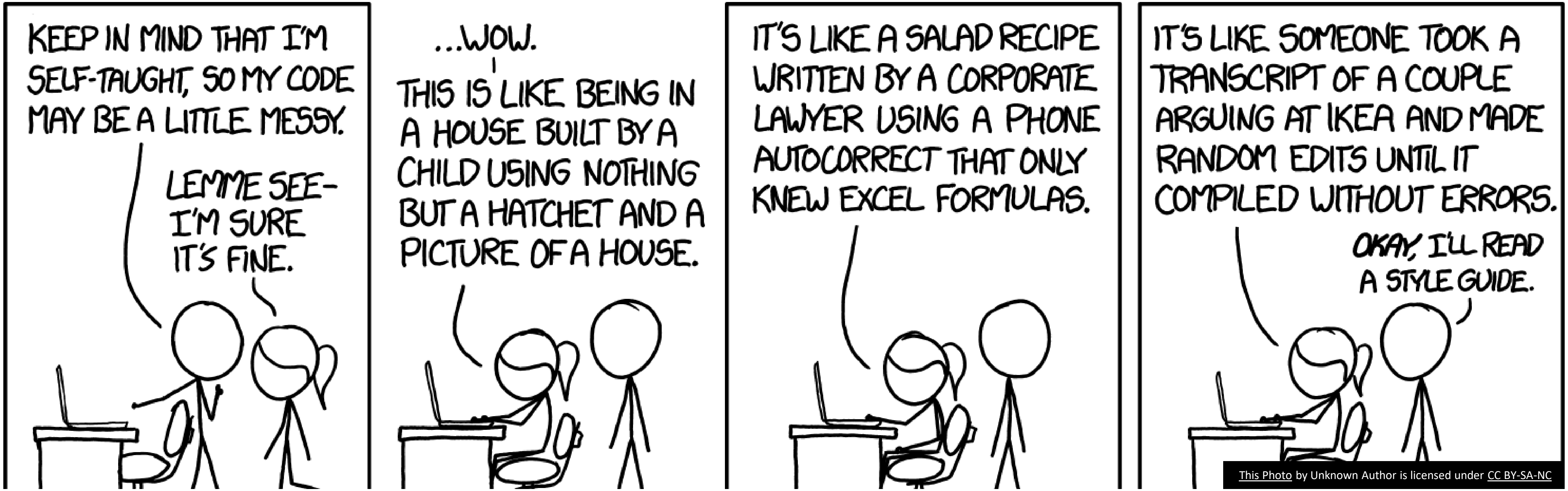
If randomness is involved:

Report the confidence / significance of your results! (cross-fold validation, averages, …)

Reproducibility

"Any results should be documented by making all data and code available in such a way that the computations can be executed again with identical results."

- "The Ethical Algorithm" (Chapter 4) by Kearns & Roth

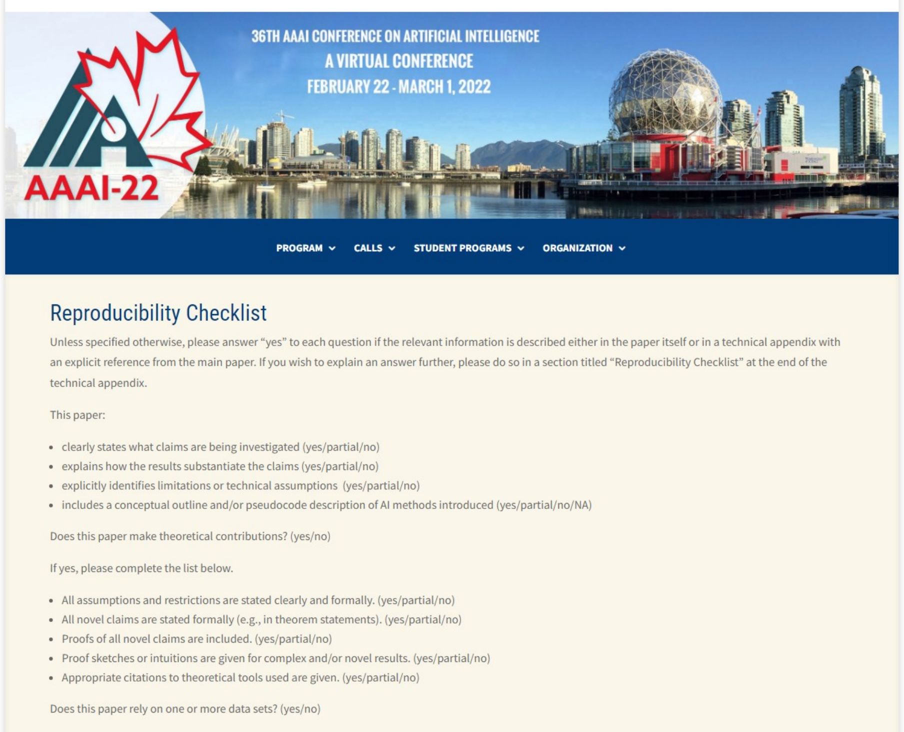


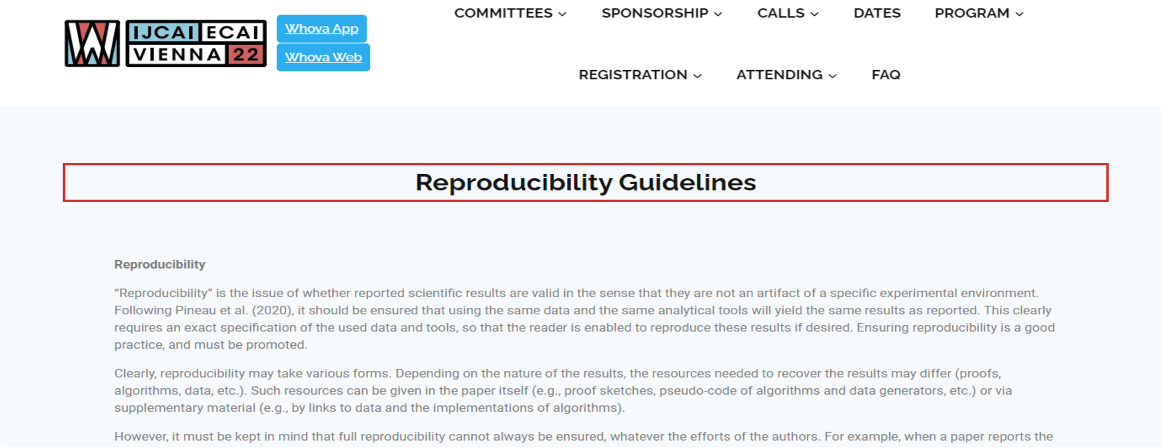
Source: https://xkcd.com/1513

Consequences for Research

Several high ranked conferences:

- Adopt reproducibility as base criterium in review process
- Set standards by comprehensive guidelines

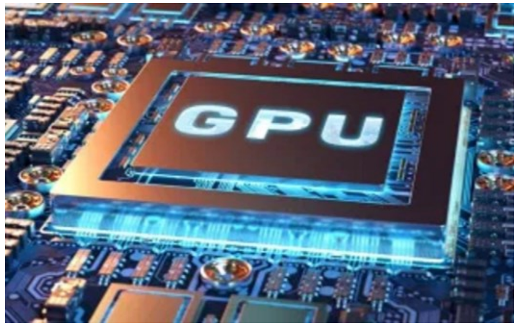




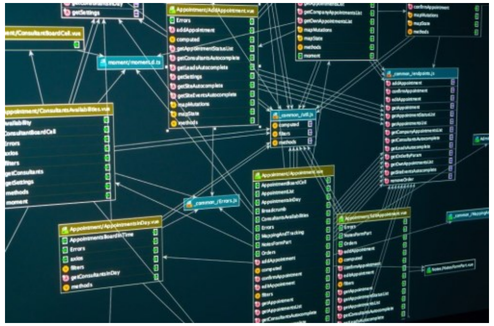
Things to keep in mind...

For someone else to be able to reproduce your results, you must publish:

Code + Documentation (e.g., code comments, parameter choices, random seed, …)



Publish specifications and versions of machines, packages / libraries etc. that were used



Publish data or data generation method and experiment design

If randomness is involved:

Report the confidence / significance of your results! (cross-fold validation, averages, …)

Reporting Uncertainty of Results

1. Average performance together with variance or standard deviation
2. Confidence Intervals
   - Over different train/test splits (cross validation)
   - Over bootstrapped sets
3. P-values

①

Reporting Uncertainty

Average & Standard Deviation

Reminder

Sample Versions of Standard Notions

	General: $x \sim p,\ p: R \longrightarrow [0,1],\ R \subseteq \mathbb{R}$	**Samples** $x_1, \dots, x_N \in \mathbb{R}$
Expectation / Mean	$\mu_x = \mathbb{E}_p[x] = \int_{r \in R} r \, dp(r)$ (R continuous) or $= \sum_{r \in R} p(r) \cdot r$ (R discrete)	$\bar{x} = \frac{1}{N} \sum_{i=1,\dots,N} x_i$
Variance	$\mathbb{V}[x] = var(x) = \mathbb{E}_p[x^2] - \left(\mathbb{E}_p[x]\right)^2 = \sigma^2$	$\bar{\sigma}^2 = \frac{1}{N-1} \sum_{i=1,\dots,N} (x_i - \bar{x})^2$
Standard Deviation	$\sigma = \sqrt{\mathbb{V}[x]}$	$\bar{\sigma} = \sqrt{sample\ variance}$

2

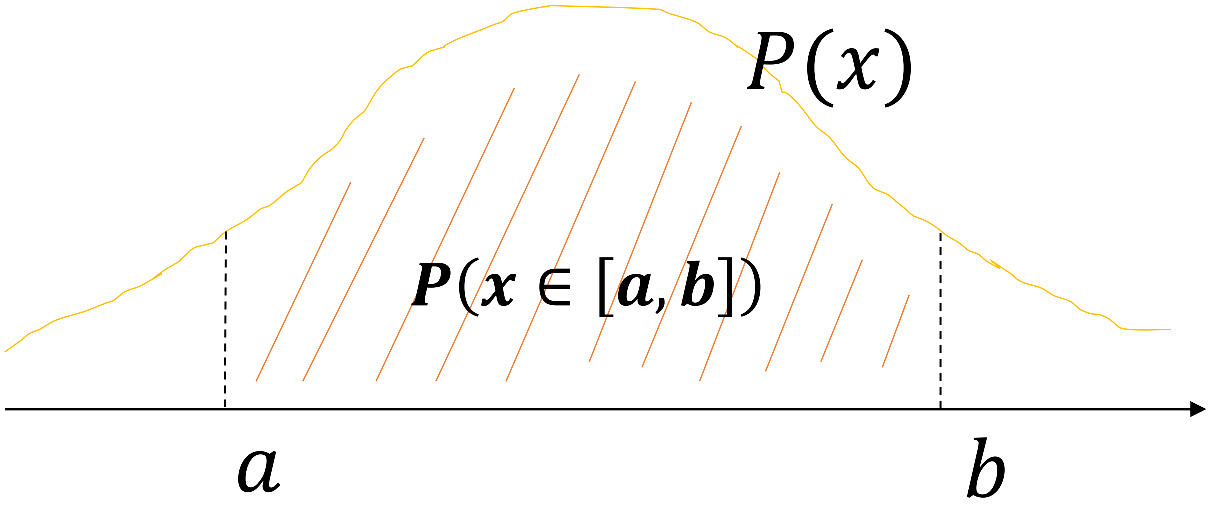
Reporting Uncertainty

Confidence Intervals

Confidence Intervals

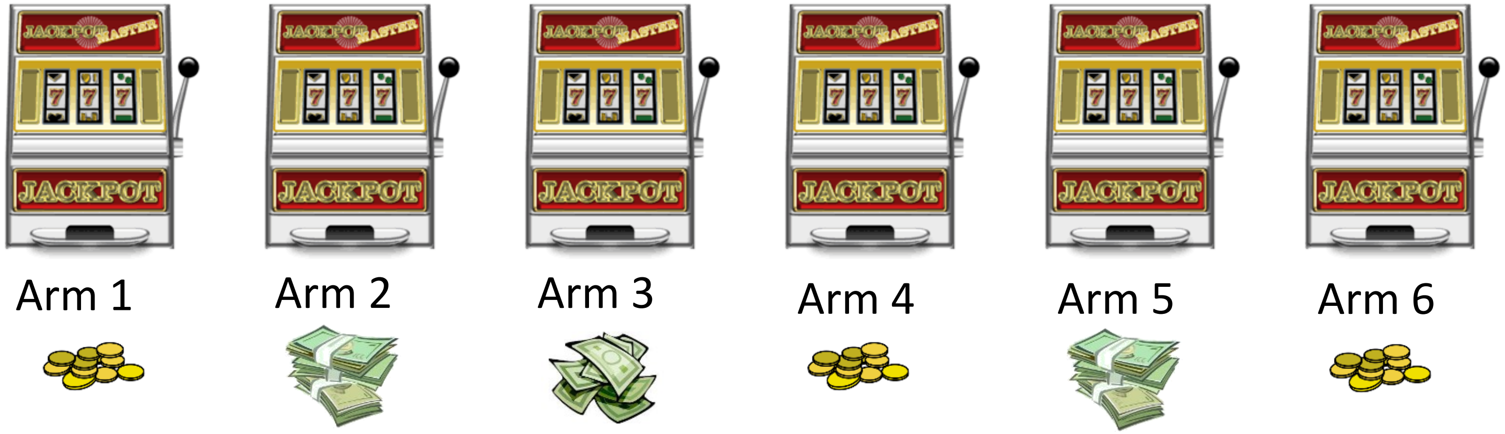
- Let x be a random variable and $x \sim P$ for probability distribution P.
- <u>Definition</u>: $[a, b]$ is a **$\boldsymbol{\gamma}$-confidence interval** for x if

$$P(a \leq x \leq b) = P(x \in [a, b]) = \gamma$$

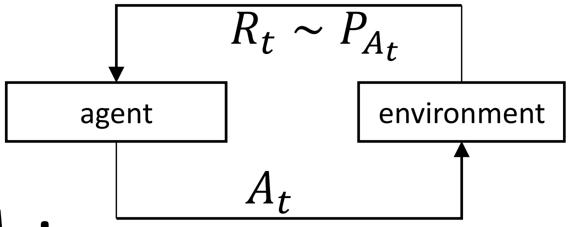


Multi-armed Bandits: Setting

- The Bandits:



- Actions: At any time step choose one arm to pull.
- Loop: Select action A_t, observe feedback (reward) R_t from unknown distribution P_{A_t}.
- Goal: Maximise rewards over time.



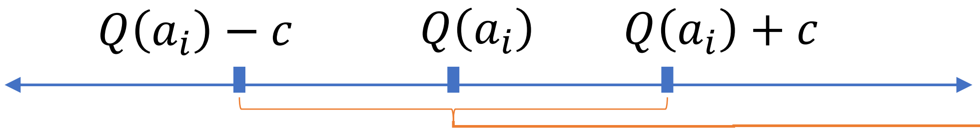
What is a confidence bound?

Arm a_i

- Suppose we have the following rewards from pulling arm a_i:
  3, 5, 6, -1, 2, -3

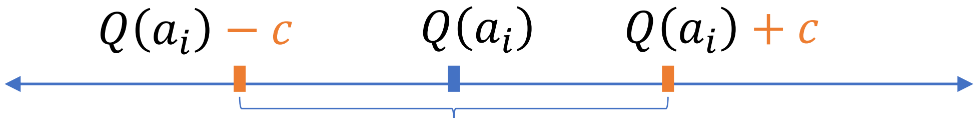
→ The sample average (=est. arm value) is: $Q^6(a_i) = 2$ (if $Q^0(a_i) = 0$)

- How confident are we that this is the correct mean of P_{a_i}?
  - With which probability is the true value $q^*(a_i) \in [Q(a_i) - c, Q(a_i) + c]$?



$\mathbb{P}\left[q^*(a_i) \in \left[Q(a_i) - c,\ Q_{a_i} + c\right]\right]$ = ?

  - In which interval does $q^*(a_i)$ lie with 95% certainty?

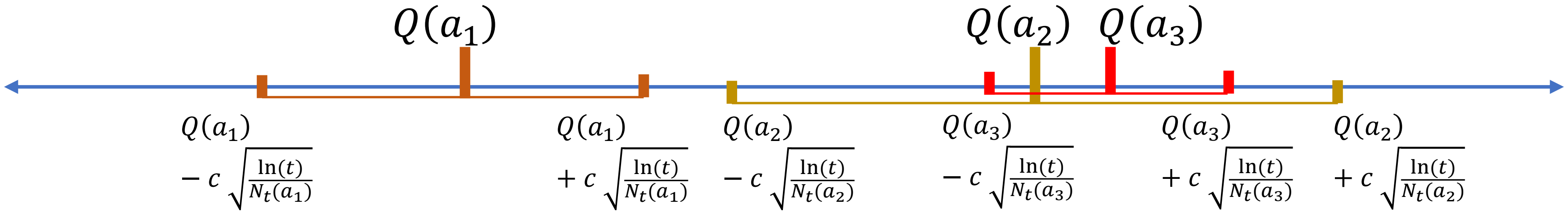


$\mathbb{P}\left[q^*(a_i) \in \left[Q(a_i) - c,\ Q_{a_i} + c\right]\right]$ = 0.95

Upper-Confidence-Bound Action Selection

- Idea: Select actions that are "uncertain", but "promising".
- Principle: Optimism in the face of uncertainty!
  → Find upper confidence bounds (by Hoeffding Inequality) of value estimates and choose the arm with the best bound:

$$A_t = \arg\max_a Q^t(a) + c\sqrt{\frac{\ln(t)}{N_t(a)}}.$$

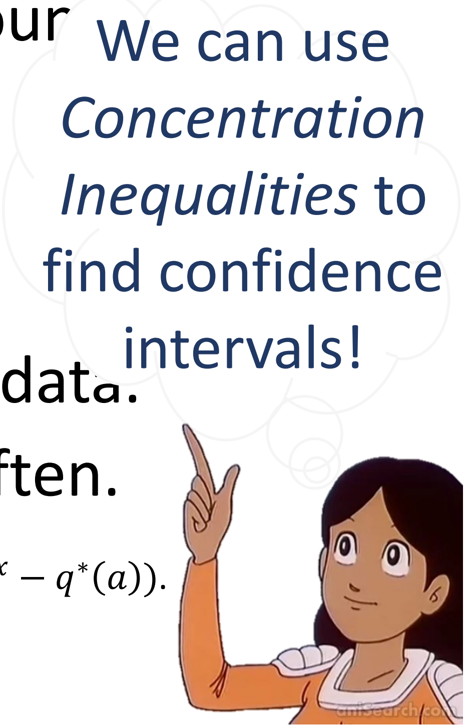


Upper-Confidence-Bound Action Selection

- Idea: Select actions that are "uncertain", but "promising".
- Principle: Optimism in the face of uncertainty!
  → Find upper confidence bounds (by Hoeffding Inequality) of value estimates and choose the arm with the best bour

$$A_t = \arg\max_a Q^t(a) + c\sqrt{\frac{\ln(t)}{N_t(a)}}.$$

  → Upper confidence bounds get tighter provided more data.
  → Intuitively, we will not select a suboptimal arm too often.
- Can implement this with almost optimal regret $O(\sqrt{k \cdot T \cdot \log(T)} + \sum_{a\in[k]} q^{max} - q^*(a))$.



Markov's Inequality

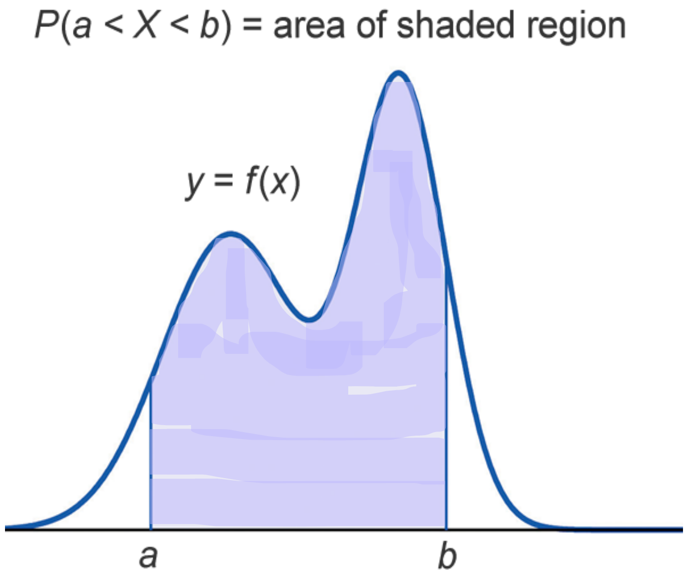
Andrey Andreyevich Markov
Source: Wikipedia

- Markov's Inequality: Let $\omega \in \mathbb{R}_{\geq 0}$ be a random variable with distribution $P(\omega)$ and $t > 0$. Then $P(\omega \geq t) \leq \mathbb{E}[\omega]/t$.
- Here, we interpret $P(\omega \geq t)$ as $P(\omega \in [t, \infty]) = \int_t^\infty p(\omega)\, d\omega$.
- Proof:

$$
\begin{aligned}
\mathbb{E}[\omega] &:= \int_0^\infty \omega \cdot p(\omega)\ d\omega \\
&= \int_0^t \omega \cdot p(\omega)\ d\omega + \int_t^\infty \omega \cdot p(\omega)\ d\omega \\
&\geq \int_t^\infty \omega \cdot p(\omega)\ d\omega \\
&\geq \int_t^\infty t \cdot p(\omega)\ d\omega \\
&= t \cdot \int_t^\infty p(\omega)\ d\omega \\
&= t \cdot P(\omega \geq t)
\end{aligned}
$$

Because $\omega \in \mathbb{R}_{\geq 0}$ and thus $\int_0^t \omega \cdot p(\omega)\ d\omega \geq 0$

Probability density function $f(x)$

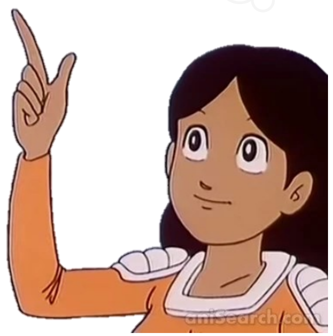


Markov's Inequality: Application Example

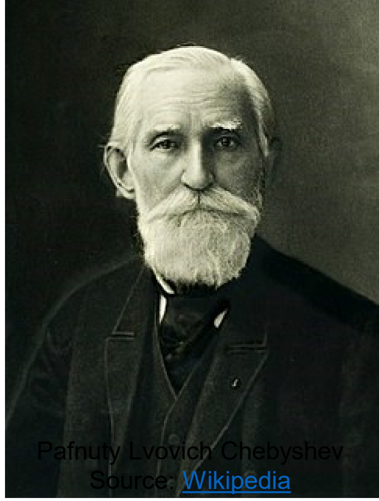
- <u>Markov's Inequality</u>: Let $\omega \in \mathbb{R}_{\geq 0}$ be a random variable with distribution $P(\omega)$ and $t > 0$. Then $P(\omega \geq t) \leq \mathbb{E}[\omega]/t$.

- <u>Example</u>:
  - Let $\xi = Beta(\alpha, \beta)$ be the Beta distr. with $\alpha = 1, \beta = 3$ and expectation $\mathbb{E}_{\omega\sim\xi}[\omega] = \frac{\alpha}{\alpha+\beta}$.
  - Q: If we want to be 90% certain, what upper bound can we put on ω?
  - A: The probability that $\omega \geq t$ is
    $\xi(\omega \geq t) \leq 1/t \cdot \mathbb{E}_{\omega\sim\xi}[\omega] = \frac{1 \cdot 1}{t \cdot 4} = 0.1 \quad \Rightarrow \quad t = \frac{1}{0.1 \cdot 4} = 2.5$
    $\Rightarrow \quad \omega \leq 2.5$ with probability at least 0.90

What if we want to lower-bound the probability? /
Have a closed interval?



Chebyshev Inequality



- Chebyshev Inequality:

  Let ω be a random variable with distribution $P(\omega)$ and $k > 0$.
  Then $P(|\omega - \mu| \geq k \cdot \sigma) \leq \frac{1}{\mathrm{k}^2}$, with expectation $\mu = \mathbb{E}[\omega] < \infty$
  and variance $\sigma^2 = \mathbb{E}[(\omega - \mu)^2] < \infty$.

- $|\omega - \mu| \geq k \cdot \sigma \quad \Leftrightarrow \omega \notin (\mu - k \cdot \sigma, \mu + k \cdot \sigma)$
  - if $\omega > \mu$: $\quad \begin{aligned} \omega - \mu &\geq k \cdot \sigma \\ \omega &\geq k \cdot \sigma + \mu \end{aligned}$
  - if $\omega < \mu$: $\quad \begin{aligned} \mu - \omega &\geq k \cdot \sigma \\ \omega &\leq -k \cdot \sigma + \mu \end{aligned}$

→ $[\mu - k \cdot \sigma, \mu + k \cdot \sigma]$ is a $(1 - \frac{1}{k^2})$ - confidence interval for ω

Chebyshev Inequality

- <u>Chebyshev Inequality</u>:

  Let $\omega \in \mathbb{R}_{\geq 0}$ be a random variable with distribution $P(\omega)$ and $k > 0$.
  Then $P(\omega \notin (\mu - k \cdot \sigma, \mu + k \cdot \sigma)\ ) \leq \frac{1}{\mathrm{k}^2}$, with expectation $\mu \ \ = \mathbb{E}[\omega]$
  and variance $\sigma^2 = \mathbb{V}(\omega) := \mathbb{E}[(\omega - \mu)^2]$.

- <u>Example</u>:
  - Let $P = \mathcal{N}(\mu = 1.5, \sigma = 0.5)$ be a normal distribution.
  - Q: What is the probability that $\omega \in [0.5{,}2.5]$ $(k = 2)$?
  - A: We have $P(\omega \notin (0.5{,}2.5)) \leq \frac{1}{2^2} = 0.25$
    $\Rightarrow P(\omega \in [0.5{,}2.5]) \geq P\big(\omega \in (0.5{,}2.5)\big) = 1 - P(\omega \notin (0.5{,}2.5)) \geq 0.75.$

Hoeffding Inequality

- <u>Hoeffding Inequality</u>: Let $x_1, \ldots, x_n \in [0,1]$ be random variables, $\omega = \sum x_i$ with distr. $P(\omega)$ and $\mu = \mathbb{E}[\omega]$.
  Then $\mathrm{P}(|\omega - \mu| \geq \epsilon) \leq e^{-\frac{2\epsilon^2}{n}} \triangleq \delta$.

- <u>Remark</u>: We have $\ln\left(\frac{1}{\delta}\right) = 2\epsilon^2/n \Leftrightarrow \epsilon = \sqrt{\frac{n \cdot \ln\left(\frac{1}{\delta}\right)}{2}}$.

Thus, $P\left(|\omega - \mu| \geq \sqrt{\frac{n \cdot \ln\left(\frac{1}{\delta}\right)}{2}}\right) \leq \delta$, i.e., $P\left(\left|\frac{\omega}{n} - \frac{\mu}{n}\right| \geq \sqrt{\frac{\ln\left(\frac{1}{\delta}\right)}{2 \cdot n}}\right) \leq \delta$

Hoeffding Inequality

- <u>Hoeffding Inequality</u>: Let $x_1, \ldots, x_n \in [0,1]$ be random variables, $\omega = \sum x_i$ with distr. $P(\omega)$ and $\mu = \mathbb{E}[\omega]$.

  Then $P\left(\left|\frac{\omega}{n} - \frac{\mu}{n}\right| \geq \sqrt{\frac{\ln\left(\frac{1}{\delta}\right)}{2 \cdot n}}\right) \leq \delta.$

- <u>Example</u>: Samples $0, 0.5, 1, 0, 1, 0, 0.5, 1$ from pulls of bandit arm a.
  - Sample average is $\mathrm{Q(a)} = 1/2$. What is the 95% confidence interval?
  - $P\left(|Q(a) - q^*(a)| \geq \sqrt{\frac{\ln\left(\frac{1}{0.05}\right)}{2 \cdot 8}}\right) \leq 0.05$ and $\sqrt{\frac{\ln\left(\frac{1}{0.05}\right)}{2 \cdot 8}} < 0.44$
  - $q^*(a) \in [Q(a) - 0.44, Q(a) + 0.44]$ with (at least) 95% confidence

Let's take a Quiz…

… go to Mentimeter!

Let’s take a break…

Back on in 5 min!

Hoeffding Inequality

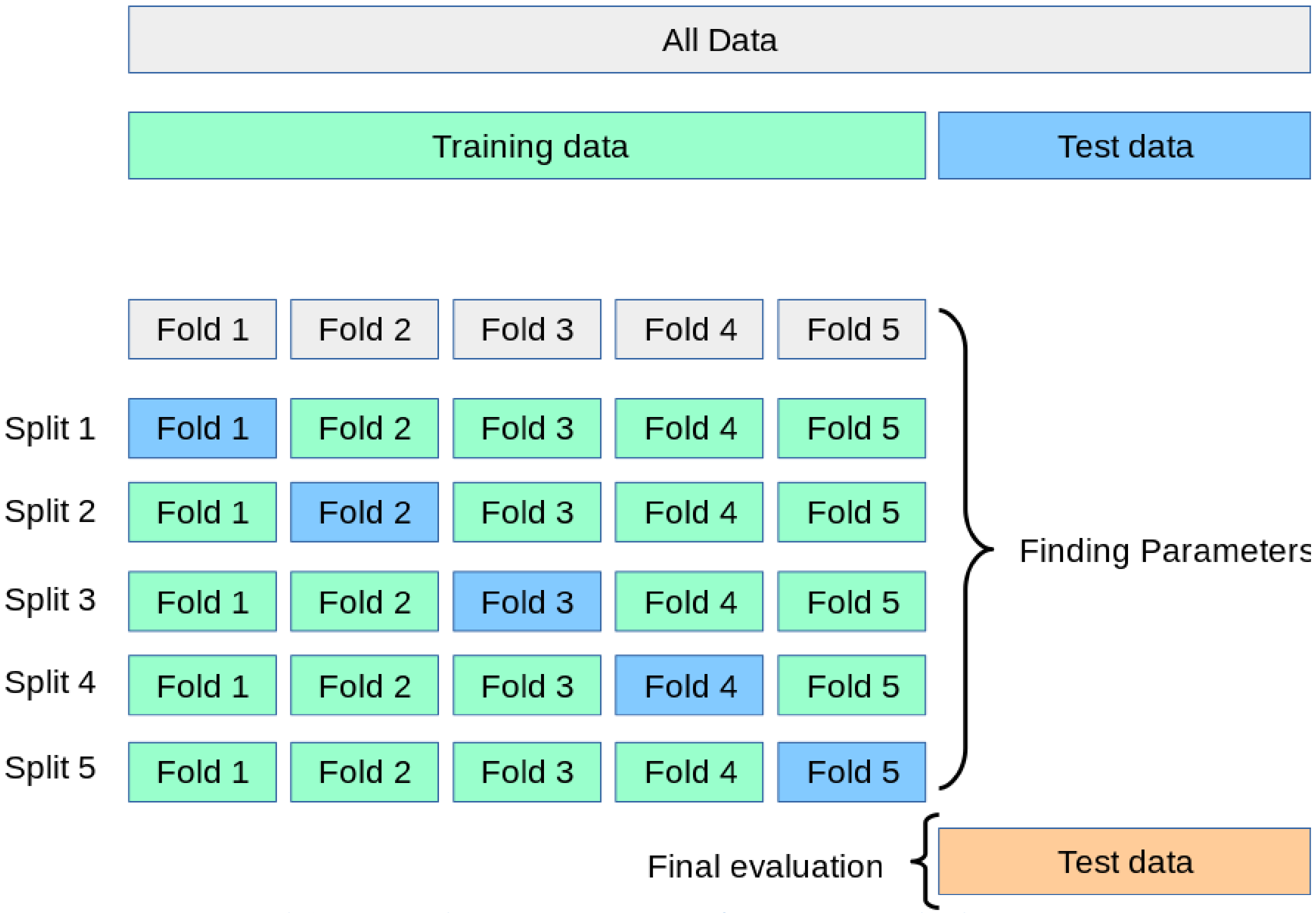
The larger the test set, the surer we can be of the accuracy on the test data.
What if we have more train-test splits?

- Hoeffding Inequality: Let $x_1, \dots, x_n \in [0,1]$ be random variables, $\omega = \sum x_i$ with distr. $P(\omega)$ and $\mu = \mathbb{E}[\omega]$.

  Then $P\left(\left|\frac{\omega}{n} - \frac{\mu}{n}\right| \geq \sqrt{\frac{\ln\left(\frac{1}{\delta}\right)}{2 \cdot n}}\right) \leq \delta.$

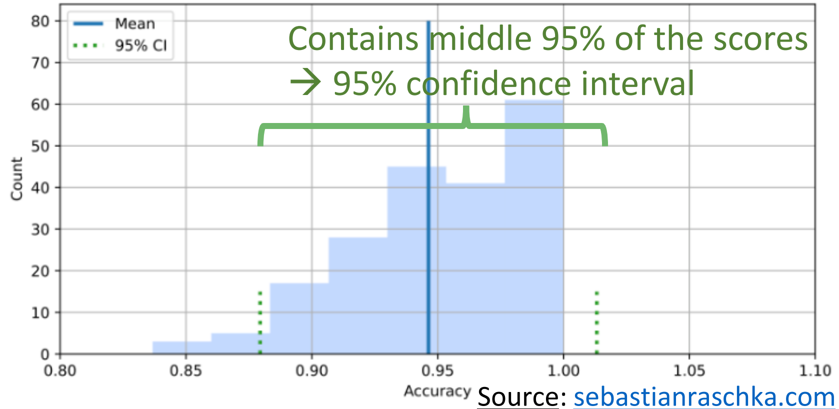
- Example: Classifier has $acc_{test} = \frac{True\ Predictions}{Total\ Predictions} = \frac{89}{100}$ **on test data**.
  - What is the 95% confidence interval?
  - $P\left(|acc_{test} - acc_{true}| \geq \sqrt{\frac{\ln\left(\frac{1}{0.05}\right)}{2 \cdot 100}}\right) \leq 0.05$ and $\sqrt{\frac{\ln\left(\frac{1}{0.05}\right)}{2 \cdot 100}} < 0.13$
  - $acc_{true} \in [0.89 - 0.13, 0.89 + 0.13] = [0.76, 1.0]$ with (at least) 95% confidence

Multiple Train-Test Splits: k-Fold Cross Validation



Source: [3.1. Cross-validation: evaluating estimator performance — scikit-learn 1.1.2 documentation]

- Get performance scores *scores* (e.g. accuracy) for every fold
- Report mean and standard dev.: *scores.mean(), scores.std()* **or…**
- Report (clipped) confidence interval:



Source: [sebastianraschka.com]

- Works if enough data for cross validation is available.

→ For less data use bootstrapping!

Bootstrapping

“To pull oneself up by one's bootstraps”
Meaning: Improve one's situation without outside help.

A. George

Sample → Compute Average → Approximation for Population Average

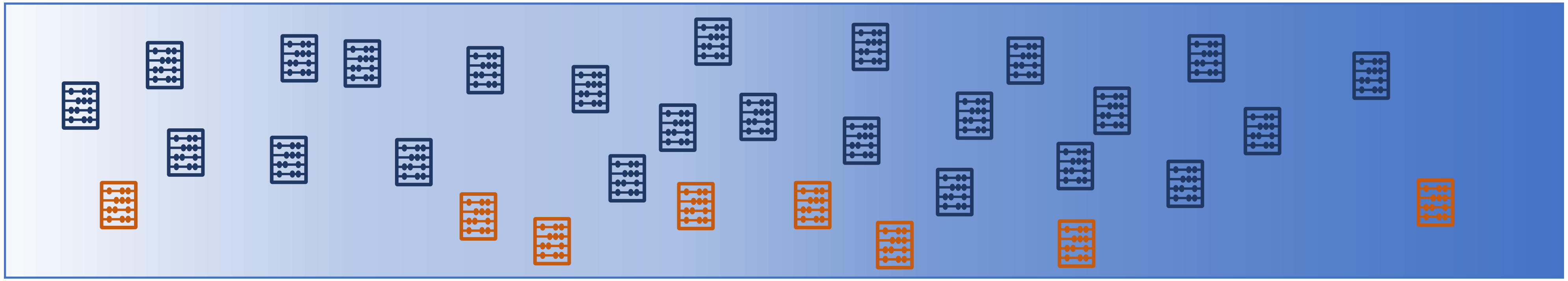
Problem: We don't know the true population
→ We don't know how good our estimate is!

Bootstrapping idea: **Resample with replacement** (several times) from the given sample population uniformly at random. → Get a distribution of estimates!

Resample → Compute Average → Approximation for Sample Population Average

Example

Determine the average IQ of all students in UiO



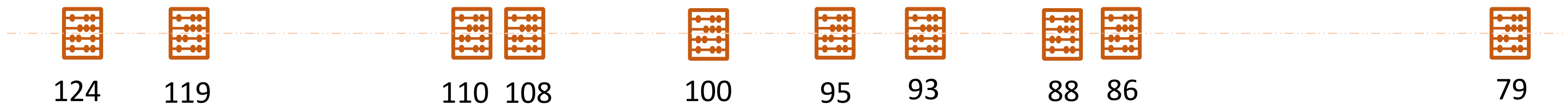
... While only knowing part of the IQs!

→ Could take the average only over the these...
... but we don't know how close this average is to the true one?!?

Resampling

- Sample of 10 student's IQs:

Average:



= 100.2

- Resample 10 students uniformly at random (with replacement)

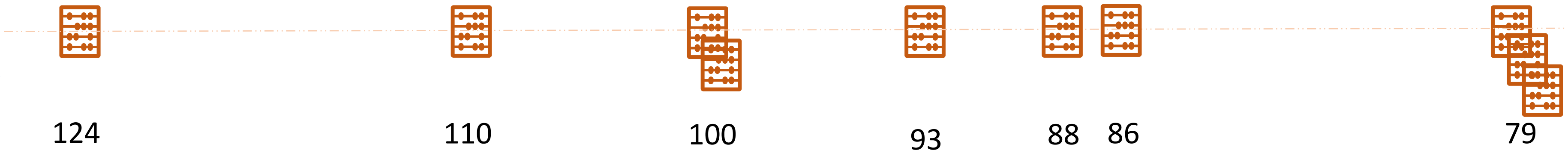
= 104.6

… (say ~1000 times) …

…

Bootstrap samples

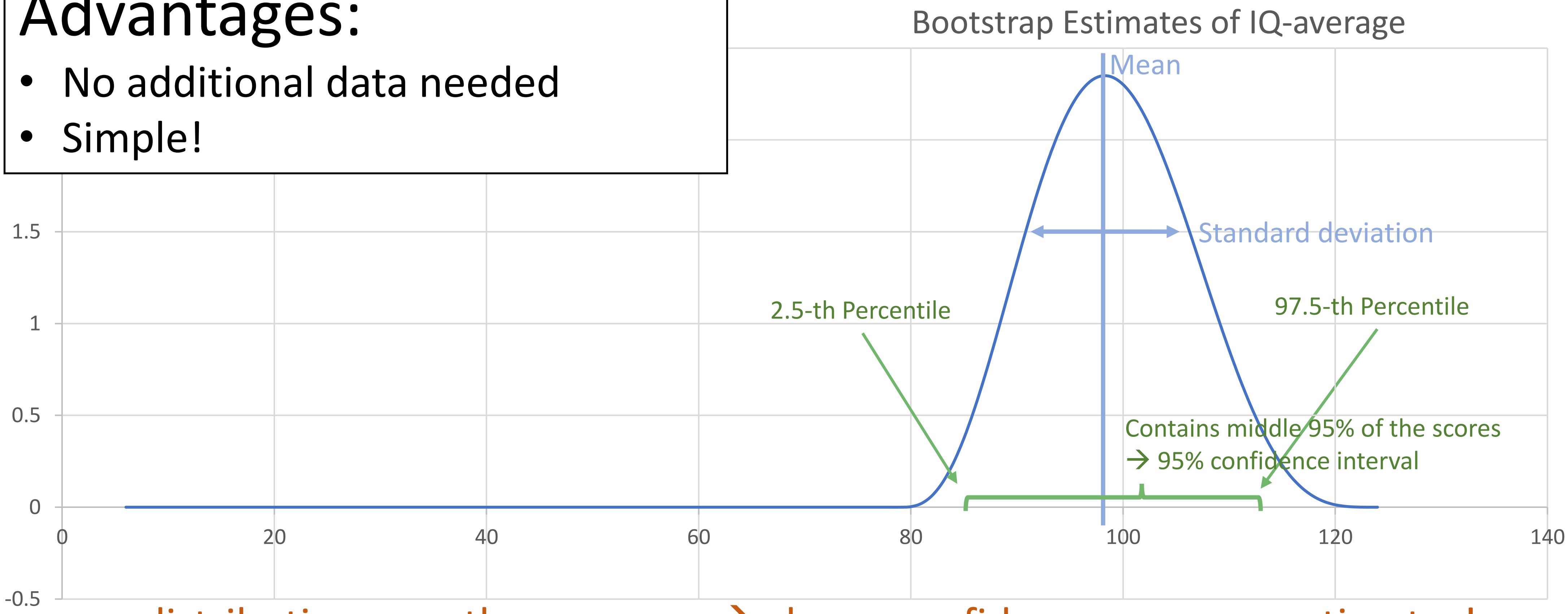
Bootstrap estimates



= 93.8

Analysing Bootstrap Estimates

Advantages:
- No additional data needed
- Simple!



… a distribution over the averages → shows confidence over our estimates!

Binomial Proportion Confidence Interval

Confidence interval for probability of success calculated from the outcome of a series of success–failure experiments (Bernoulli trials):

- Bernoulli samples $x_1, \dots x_n$, e.g., true/false prediction on test set
- Estimated success probabilities, e.g., from k-fold CV / bootstr.:
  A_i proportion of true predictions in i-th test
- **Assumption**: $A_1, \dots, A_k$ are normally distributed with mean $\bar{A}$
- True success probability:

$$A \in \bar{A} \pm z\sqrt{\frac{\bar{A}(1-\bar{A})}{k}}$$, with z = z-value (incl. confidence)

A Word of Caution… Bias in Data Collection

- Any data set is only a sample from the real population.
- Bias in data collection:
  - Intentional bias
  - Faulty or inaccurate measurement tools
  - High variance in minorities of population
  - Over- or under-sampling
  - Labelling bias
- Even with bootstrapping or other sampling techniques…
- … we cannot expect a model to perform as in the tests when it is employed on the true population if the data is biased!

③ Reporting Uncertainty

p-Values

Null Hypothesis Testing

- Null Hypothesis: H_0 The hypothesis we want to test.

  “A does not have a causal effect on B”, i.e., $P(B|A) = P(B)$

- Alternate Hypothesis: Negation of null-hypothesis

  “A has a causal effect on B”, i.e., $P(B|A) \neq P(B)$

Desired value for $P(rejecting\ H_0\ |\ H_0\ \ true)$. Typically, $\alpha = 0.05$.

- Perform a Hypothesis Test: t-test, Z-test, Chi-sq. … → get p-value ← Probability of obtaining data at least as extreme, given that the null hypothesis is true.
  - If $p < \alpha$: Reject the null hypothesis!

    (result significant) (Enough evidence to say that “A has a causal effect on B”!)
  - If $p \geq \alpha$: Cannot reject the null hypothesis!

    (result not significant) (Not enough evidence to say whether A has a causal effect on B, or not!)

The Reproducibility Crisis

Definition:

“The replication crisis (also called the replicability crisis and the reproducibility crisis) is an **ongoing methodological crisis** in which it has been found that the results of many scientific **studies are difficult or impossible to reproduce**. […] such failures undermine the credibility of theories building on them and potentially **call into question substantial parts of scientific knowledge**.” - Wikipedia

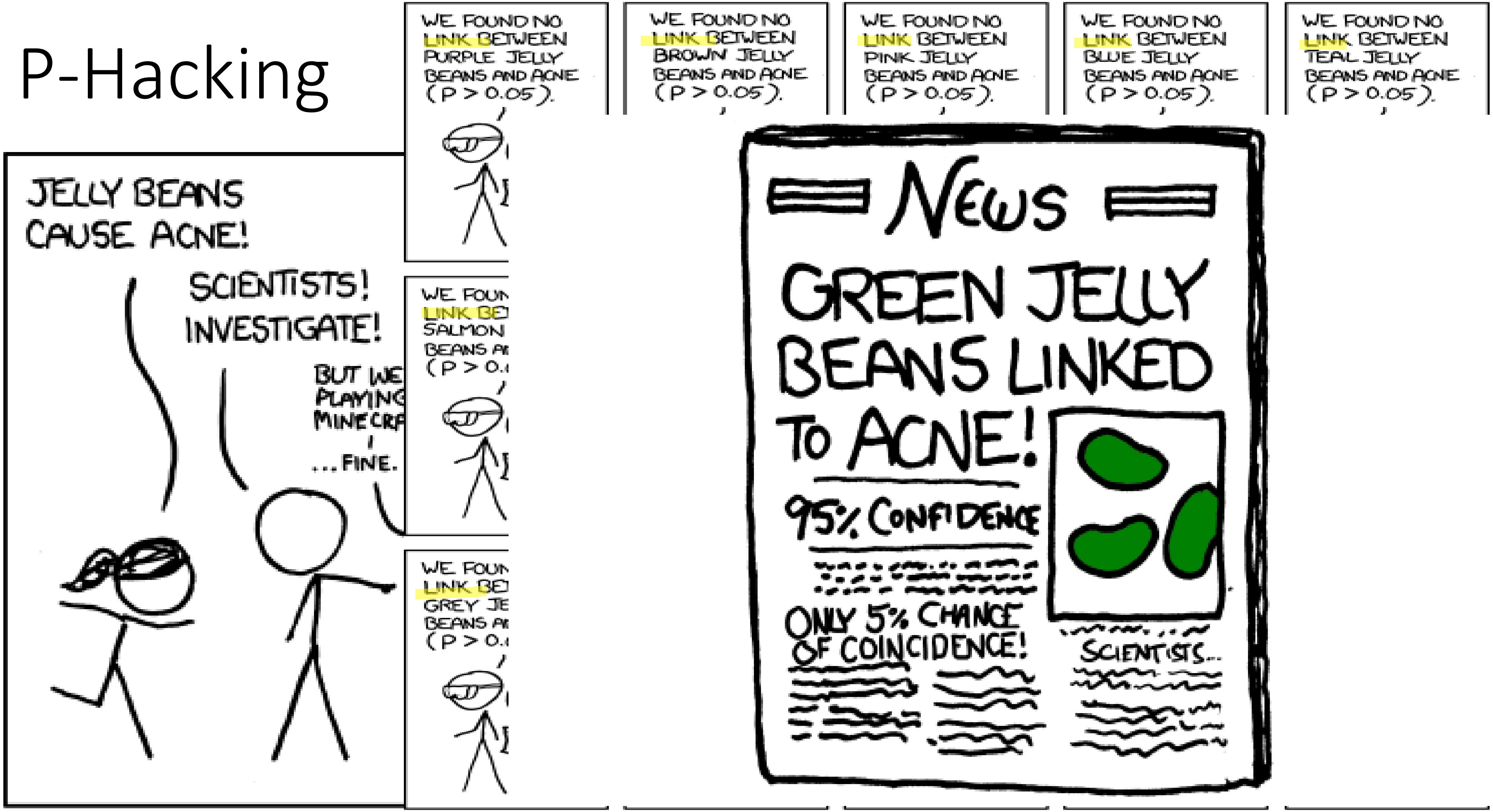
Scope:

Monya Baker, *Nature*, 2016: “1,500 scientists lift the lid on reproducibility”
“More than 70% of researchers have tried and failed to reproduce another scientist's experiments, and more than half have failed to reproduce their own experiments.**”**

Causes:

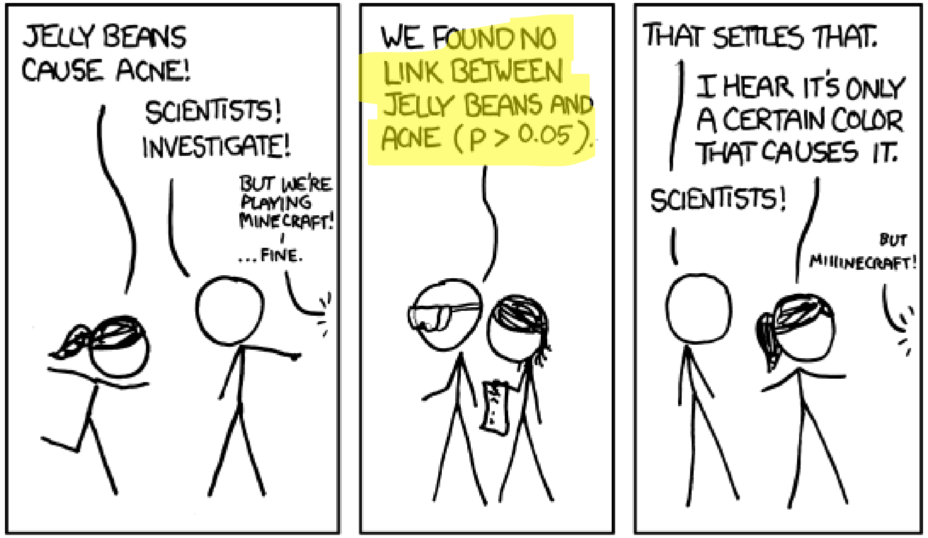
p-hacking, poor study design /experimental technique, fraud and deception, “publish or parish” culture, …

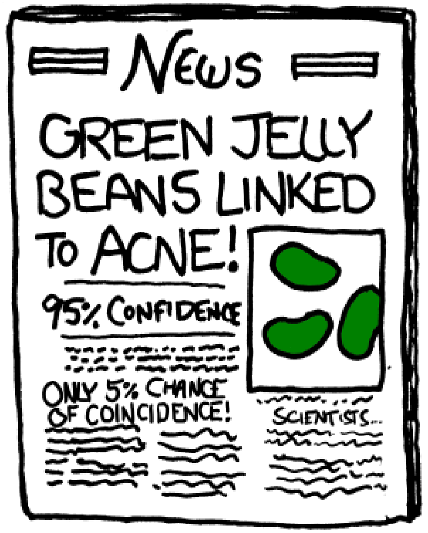
P-Hacking


JELLY BEANS CAUSE ACNE!
SCIENTISTS! INVESTIGATE!
...FINE.
WE FOUND NO LINK BETWEEN PURPLE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN BROWN JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN PINK JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN BLUE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN TEAL JELLY BEANS AND ACNE (P > 0.05).
NEWS
GREEN JELLY BEANS LINKED TO ACNE!
95% CONFIDENCE
ONLY 5% CHANCE OF COINCIDENCE!
SCIENTISTS...

If you torture the data …

Ronald Coase, Nobel Prize–winning British economist:
"If you torture the data for long enough, it will confess to anything."





Source: https://xkcd.com/882

P-hacking

P-hacking = Testing many null hypothesis / statistics on same data and only reporting the ones that are significant

Cause: Non-significant results less interesting / not publishable

Example: "Green Jelly Beans cause acne"

→ Splitting data up by colors of jelly beans and finding correlations to acne

→ The more colors, the more likely to find a correlation

Multiple Comparisons Problem

- The Problem: Run an experiment repeatedly and only report the "interesting" findings.
- Example:
  - Experiment: Flipping a coin 100 times (bias=b).
  - Observing 100× Heads is unlikely: b^{100}.
  - Repeating the experiment k times
    → more likely to observe 100× Heads in one of them: $k \cdot b^{100}$
    → reporting that 100× Heads occurred (in one of the experiments) paints a very incomplete picture!

Bonferroni-Correction

- Testing m null-hypothesis $H_1, \ldots H_m$
- We can observe:

	Null Hypothesis True	Alternative True	Total
Test significant $p < \alpha$	V	S	R
Test non-significant	U	T	$m - R$
Total	m_0	$m - m_0$	m

- Significance for independent tests: $\bar{\alpha} = 1 - (1 - \alpha)^m$ (increasing with m)
  for non-independence: $\bar{\alpha} \leq \mathrm{m} \cdot \alpha$
- **Bonferroni-correction**: Report significance α / m

Bonferroni-Correction

- **Bonferroni-correction**: Report significance $\frac{\alpha}{m}$
  → Counteracts multiple-comparisons
- Conservative method: Only reject null-hypothesis when very certain!
- Non-Adaptive: We assume that the experiment is set beforehand and not adjusted depending on the findings!

Bonferroni-Correction & Adaptivity

Example: Building a binary classifier based on d binary features.

1. Identify (anti-) correlations of features with target → relevant features
   Age 60+, Tattoos, blue eyes, … : appear *slightly* correlated with "likes Spaghetti" in dataset

2. Classify as Yes if at least half of the relevant features are positive

Problem:

→ Features can be randomly distr. & actually uncorrelated with target

→ The classifier can't perform better than random

→ But we could find slight correlations and classifier performs well on data!

Bonferroni-Correction & Adaptivity

- Target: “Likes spaghetti”
  Correlations: Blue eyes, owns cat
  Anti-correlations: Age 60+, Tattoos
- Bonferroni-Correction only corrects the multiple correlation tests considered
- … does not account for **every possible sets** of correlated features, i.e., all classifiers!

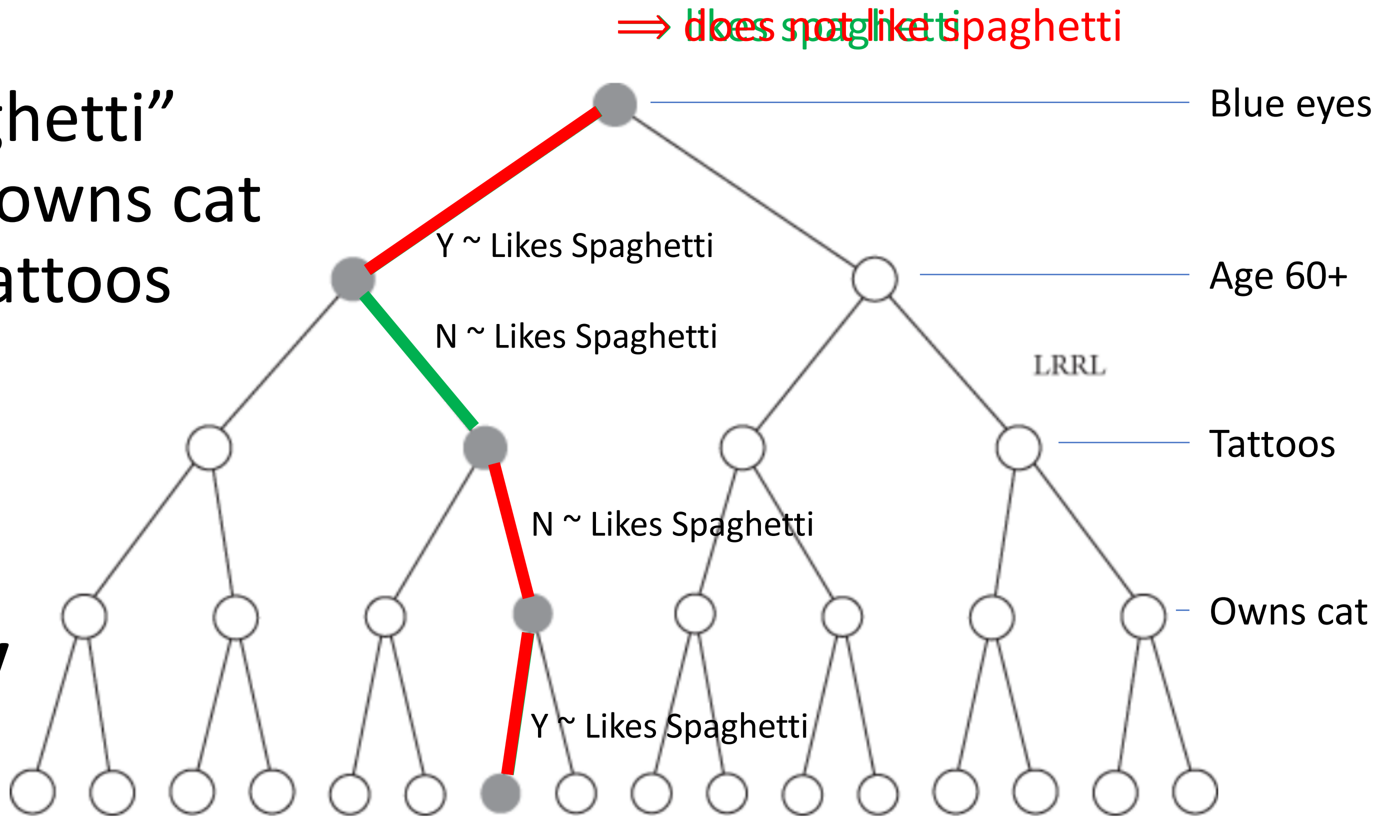


Fig. 26. Tree illustrating the dangers of adaptive data analysis and *p*-hacking. Each level of the tree corresponds to a feature that could be correlated (left) or anti-correlated (right) with the label. The gray path (LRRL) represents the outcomes of the correlation tests. Each leaf corresponds to a classifier that results from a sequence of correlation tests.

Source: “The Ethical Algorithm” (Chapter 4) by Kearns & Roth

Further reading

- Chapter 3&4 - The Ethical Algorithm [Video] (oreilly.com)
- 3. Model selection and evaluation — scikit-learn 1.1.2 documentation
- Chapter 8 Bootstrapping and Confidence Intervals | Statistical Inference via Data Science (moderndive.com)
- The resources of last years' courses (see Canvas links)
- Lecture 4 from IN-STK 5100 in 2022

What did we talk about today?

Discuss with your neighbor!

Please let me know what you think…

Leave your feedback *(for Anne-Marie's lectures)* on Flinga:

https://flinga.fi/s/FDN2DC3

- Were the contents understandable? Relevant? Interesting?
- How was the lecturing style?
- How was the teaching material?
- What could be improved?
- Any other comments?