



Statistic traps and paradoxes

Limits of statistical methods

Petr Hála, Jan Hučín

29.10.2025

Outline

1. Simpson paradox
2. Survival bias
3. Publication bias
4. Benford law
5. Where did your data come from?
6. Beware the average
7. When normality fails
8. Log-Log scale
9. Pitfalls of measures and models
10. Justice system doesn't understand conditional probability



Simpson paradox

UC Berkeley gender bias

> admission figures for the fall of 1973

– Total

	All		Men		Women	
	Applicants	Admitted	Applicants	Admitted	Applicants	Admitted
Total	12,763	41%	8,442	44%	4,321	35%

UC Berkeley gender bias

> admission figures for the fall of 1973

– Total

	All		Men		Women	
	Applicants	Admitted	Applicants	Admitted	Applicants	Admitted
Total	12,763	41%	8,442	44%	4,321	35%

– TOP 6 Departments

Department	All		Men		Women	
	Applicants	Admitted	Applicants	Admitted	Applicants	Admitted
A	933	64%	825	62%	108	82%
B	585	63%	560	63%	25	68%
C	918	35%	325	37%	593	34%
D	792	34%	417	33%	375	35%
E	584	25%	191	28%	393	24%
F	714	6%	373	6%	341	7%
Total	4526	39%	2691	45%	1835	30%

Kidney stone treatment

> Compare 2 treatments

- Treatment A
 - Expensive
 - Invasive
 - **78% success rate**
- Treatment B
 - Cheaper
 - Less invasive
 - **83% success rate**

Kidney stone treatment

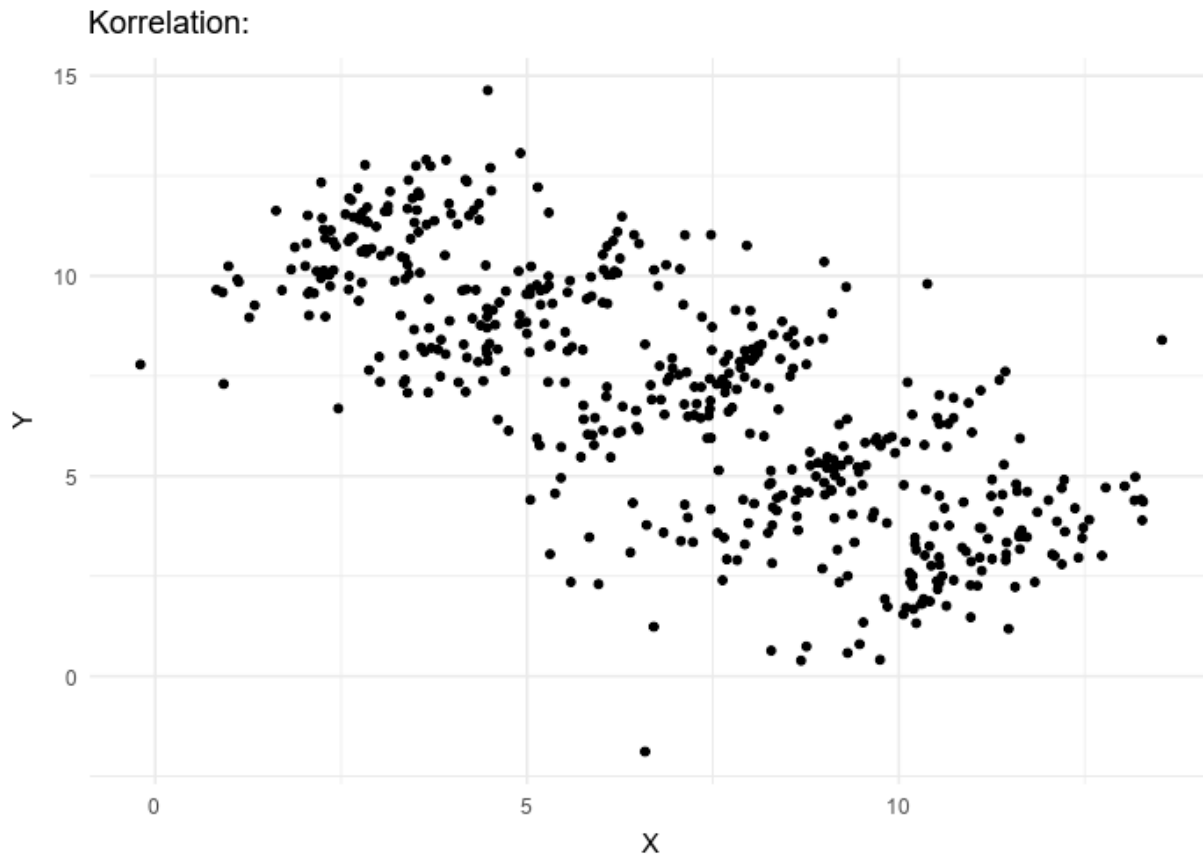
> Compare 2 treatments

- Treatment A
 - Expensive
 - Invasive
 - **78% success rate**
- Treatment B
 - Cheaper
 - Less invasive
 - **83% success rate**

Treatment Stone size	Treatment A	Treatment B
Small stones	Group 1 93% (81/87)	Group 2 87% (234/270)
Large stones	Group 3 73% (192/263)	Group 4 69% (55/80)
Both	78% (273/350)	83% (289/350)

- **Confounder** – seriousness of the condition

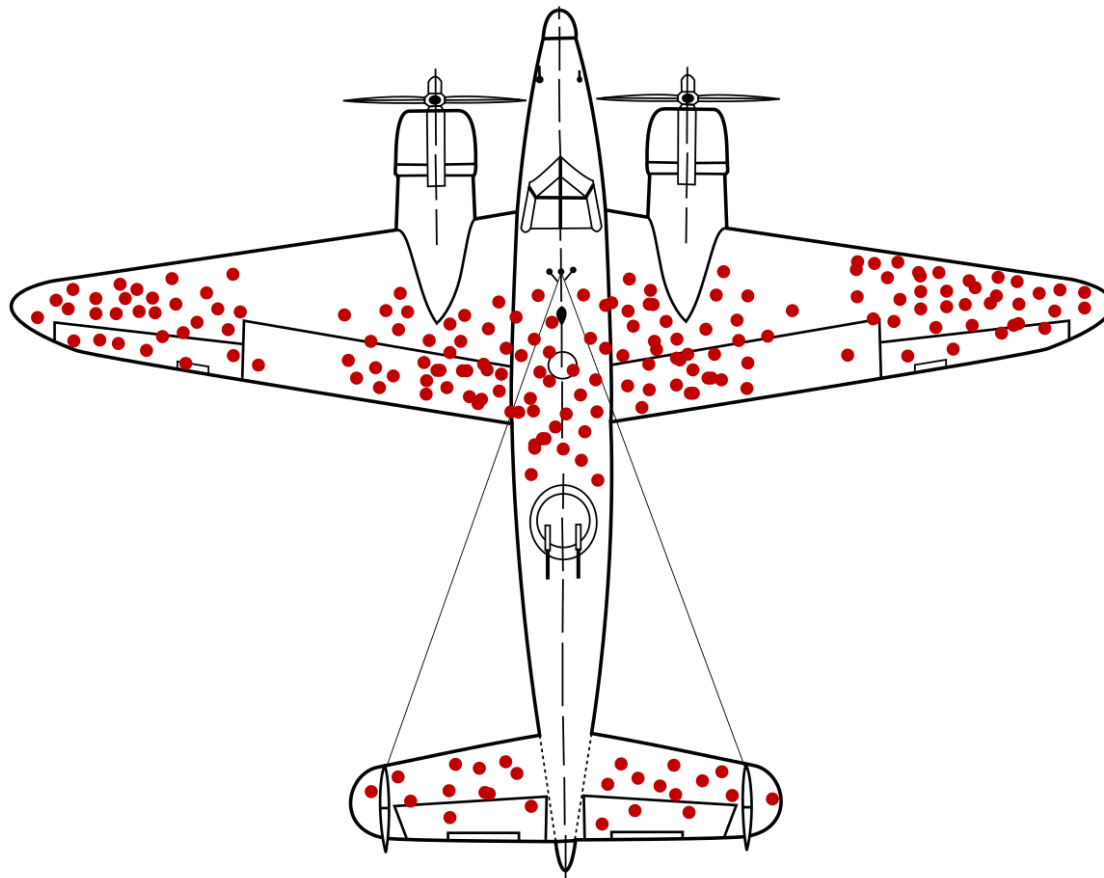
Illustrace



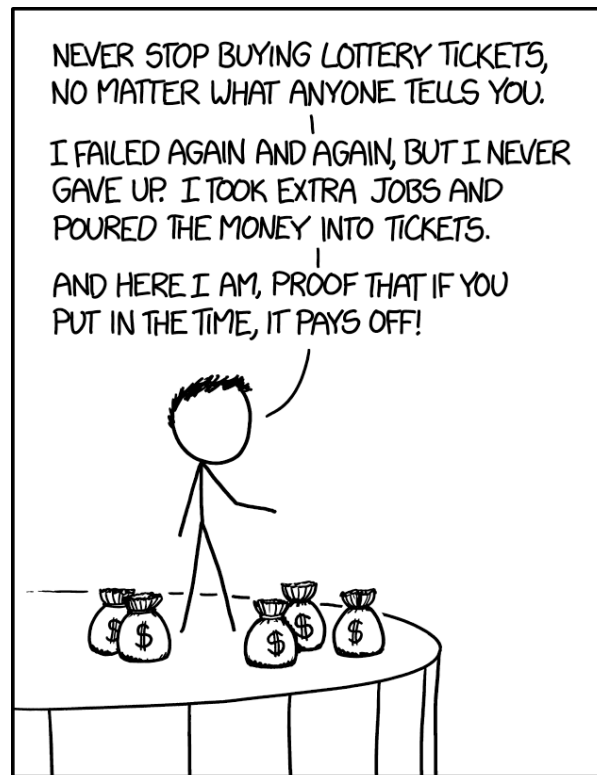
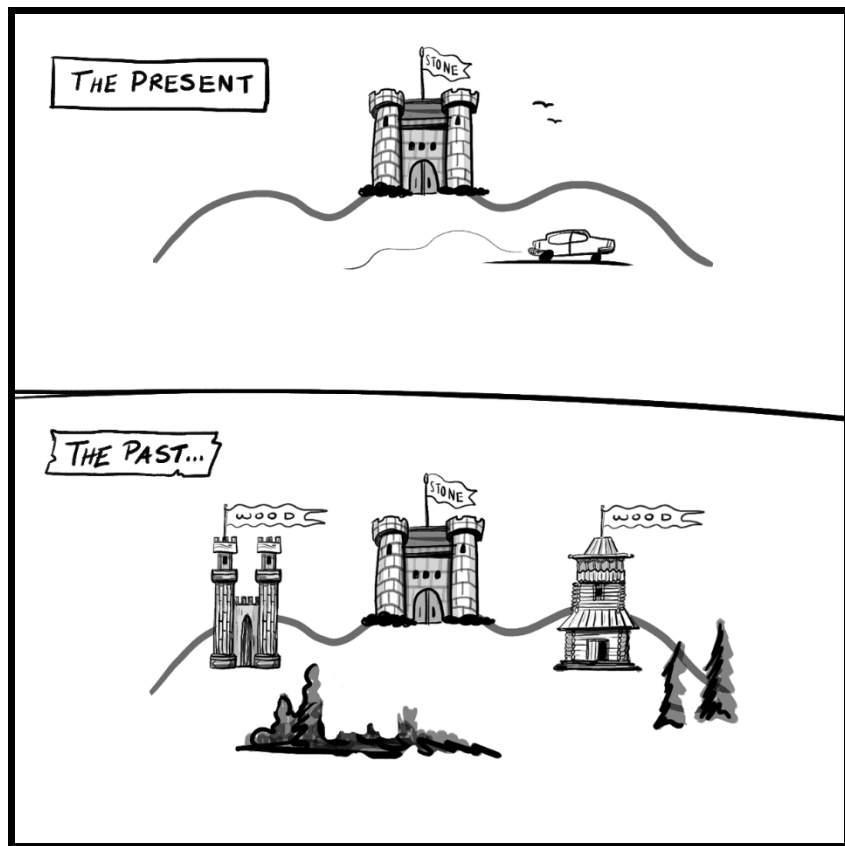


Survival bias

WW2 airplane damage



Other examples



EVERY INSPIRATIONAL SPEECH BY SOMEONE
SUCCESSFUL SHOULD HAVE TO START WITH
A DISCLAIMER ABOUT SURVIVORSHIP BIAS.

Specialized experts

> Experts

- tend to come in when things go **WRONG**
- get paid to **DO SOMETHING**



Even well-respected experts may have their opinion preset by the nature of their position



Publication bias

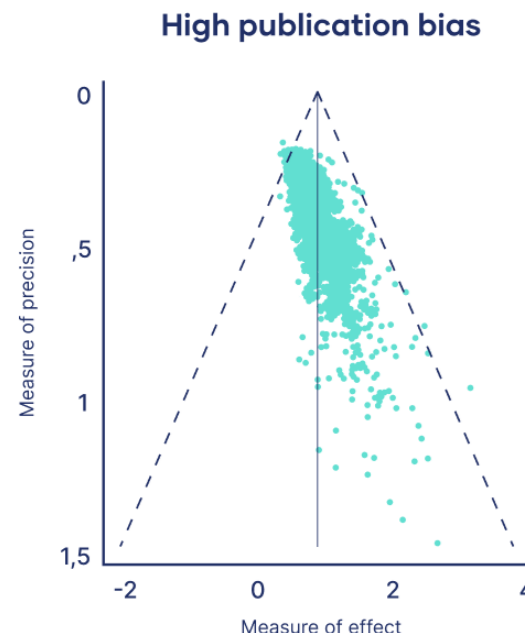
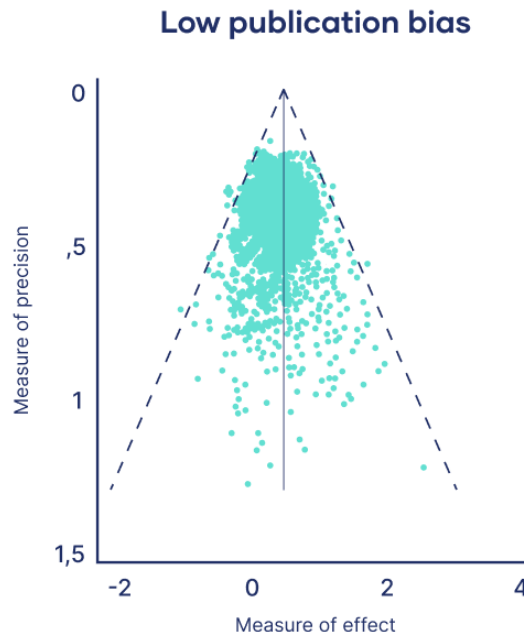
Publication bias

- Which scientific results are published is **dependent on the result itself**
 - Statistical significance
 - “Unfavourable” outcome

Publication bias

> Which scientific results are published is **dependent on the result itself**

- Statistical significance
- “Unfavourable” outcome

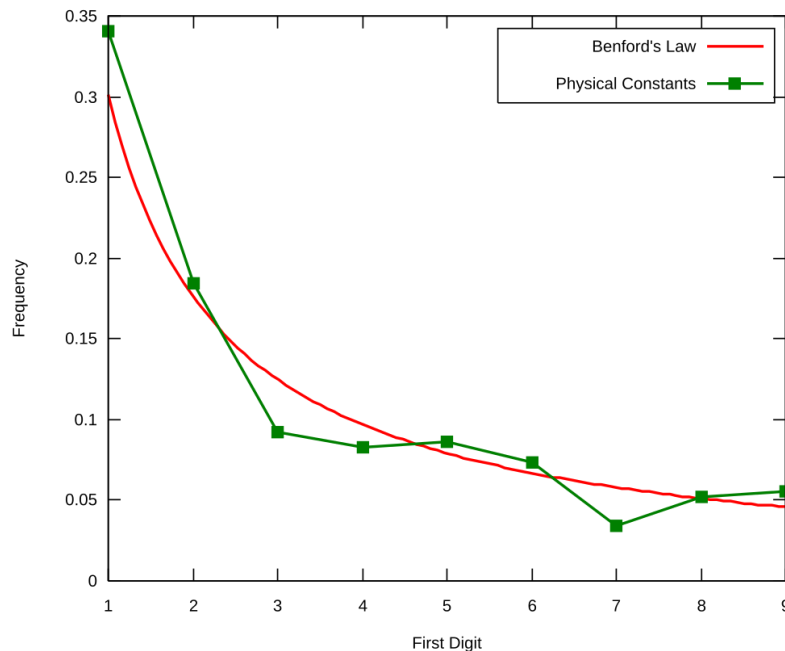
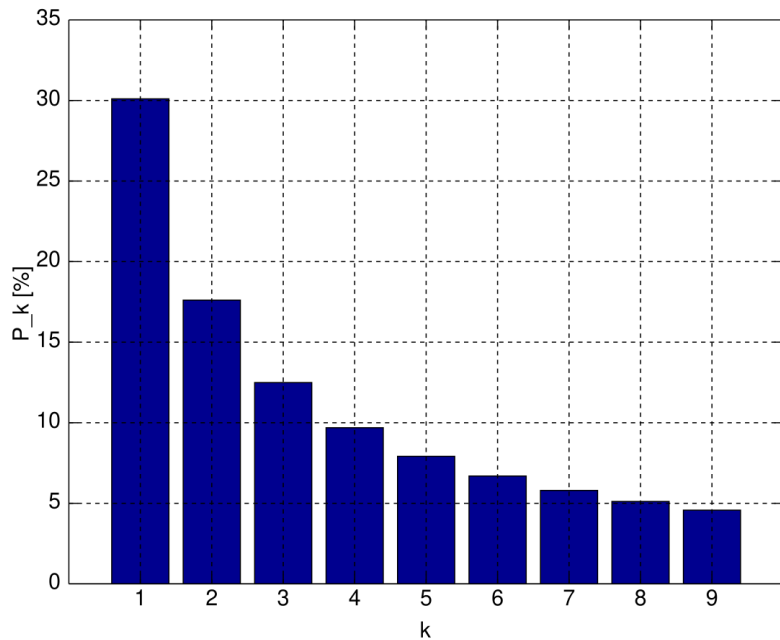




Benford law

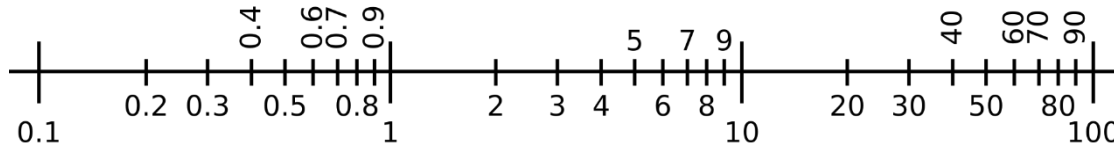
Benford law

- **law of anomalous numbers, or the first-digit law,**
 - in many real-life sets of numerical data, the leading digit is likely to be small.

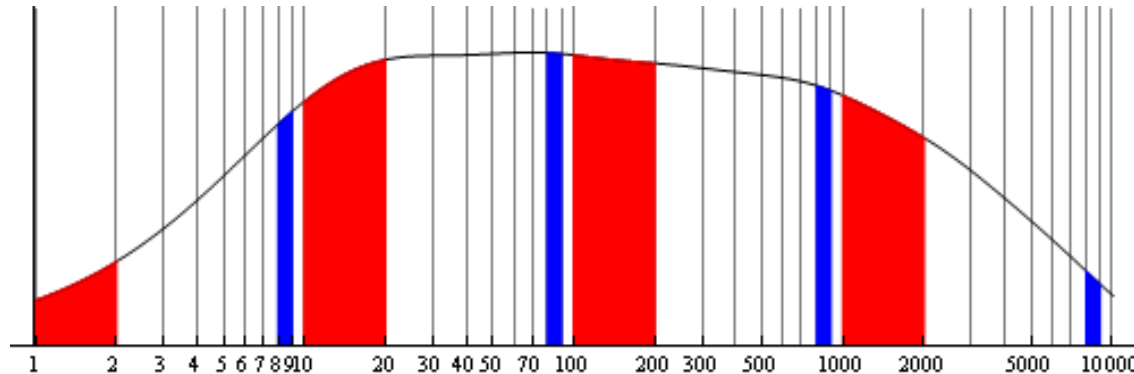


Benford law

- > The 1st digit **distribution is invariant to the units of measurement**
 - Physical constants, Length of world rivers, Area of countries, ...
- > Theoretically - data **uniformly distributed on a logarithmic scale**



- > Practically - data spanning over **several orders of magnitude**



Fraud detection

- People's brains are bad at generating randomness
- People think linearly
- Financial data
- Elections
 - If enough variation in electoral district sizes

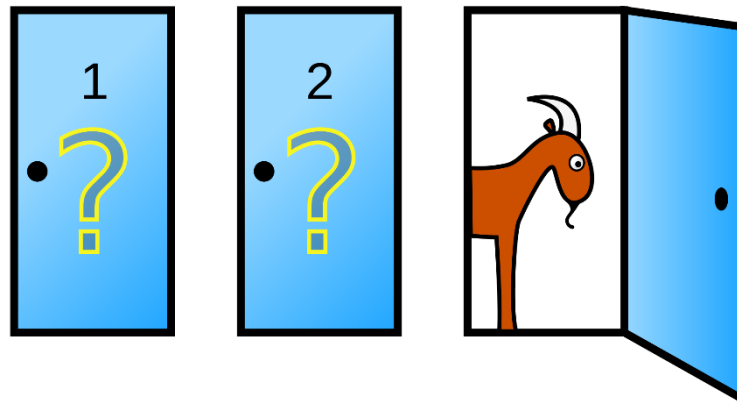




Where did your data come from?

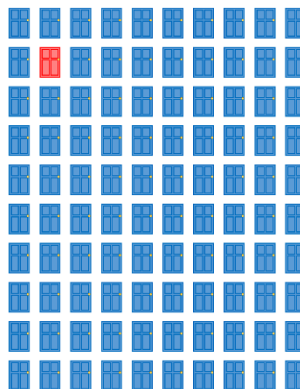
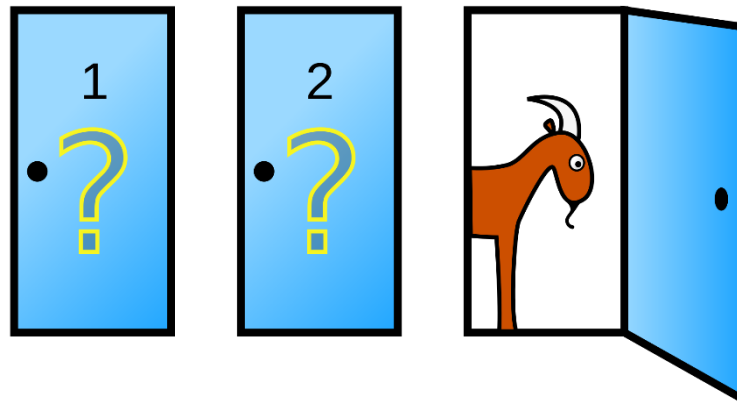
Monty Hall problem

- American TV game show
- 3 doors
 - 1 car
 - 2 goats
- Once you choose, the host opens one of the other door – with a goat.
- Should you change your pick?

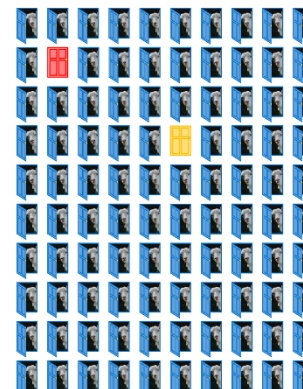


Monty Hall problem

- > American TV game show
- > 3 doors
 - 1 car
 - 2 goats
- > Once you choose, the host opens one of the other door – with a goat.
- > Should you change your pick?
- > Imagine 100 doors



At first, you choose 1 in 100 doors: 0.01 probability (1% chances) of getting the winning door!



The TV host opens all but the winning door and the one you chose: 0.99 probability (99% chances) of winning if you change door!

Jackie's improbable sister

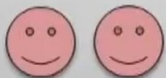
- Jackie is a girl in a family with two children.
- What is the probability that Jackie has a sister?

Jackie's improbable sister

- Jackie is a girl in a family with two children.
- What is the probability that Jackie has a sister?
- **How we found Jackie?**

Option A: 1/2

- 1) Pick a two child family at random.
- 2) Pick a child from the family at random.



Option B: 1/3

- 1) Pick a two child family with *at least one girl* at random.
- 2) Report one girl's name for each family.

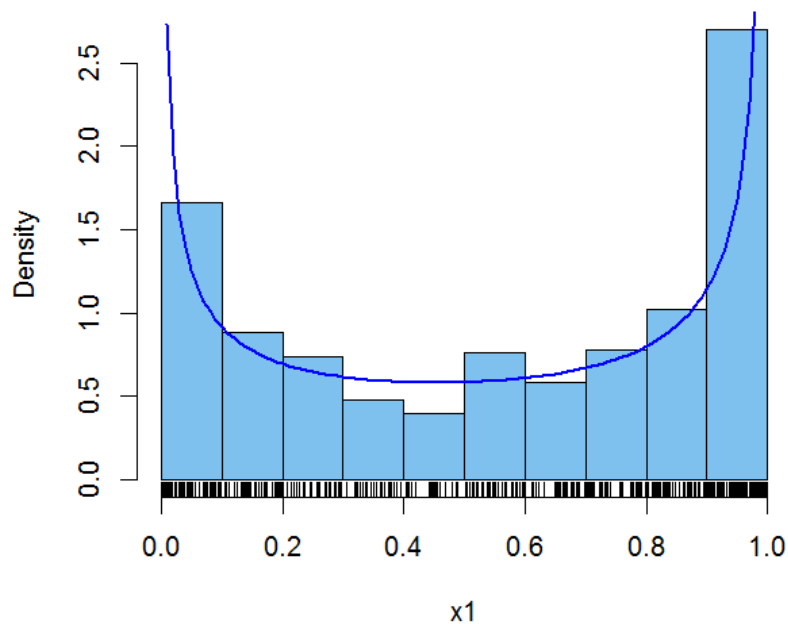




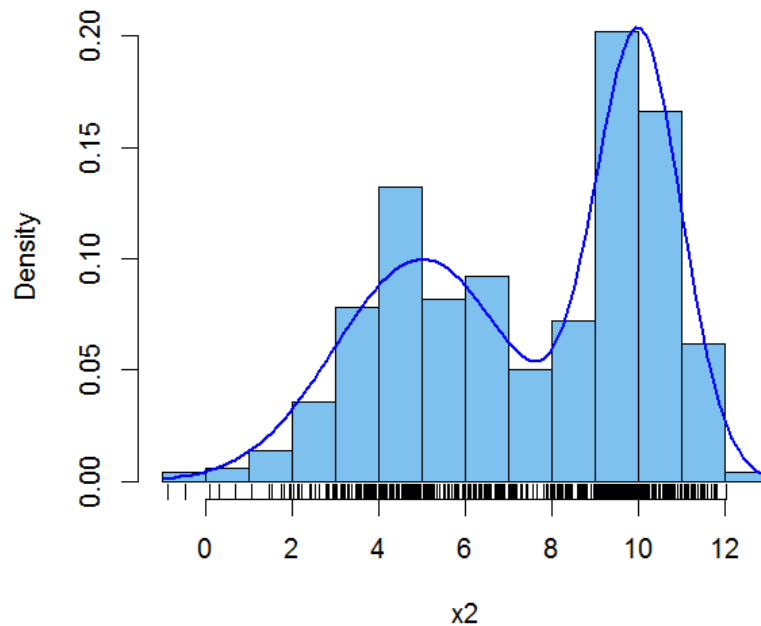
Beware the average

Beware the average

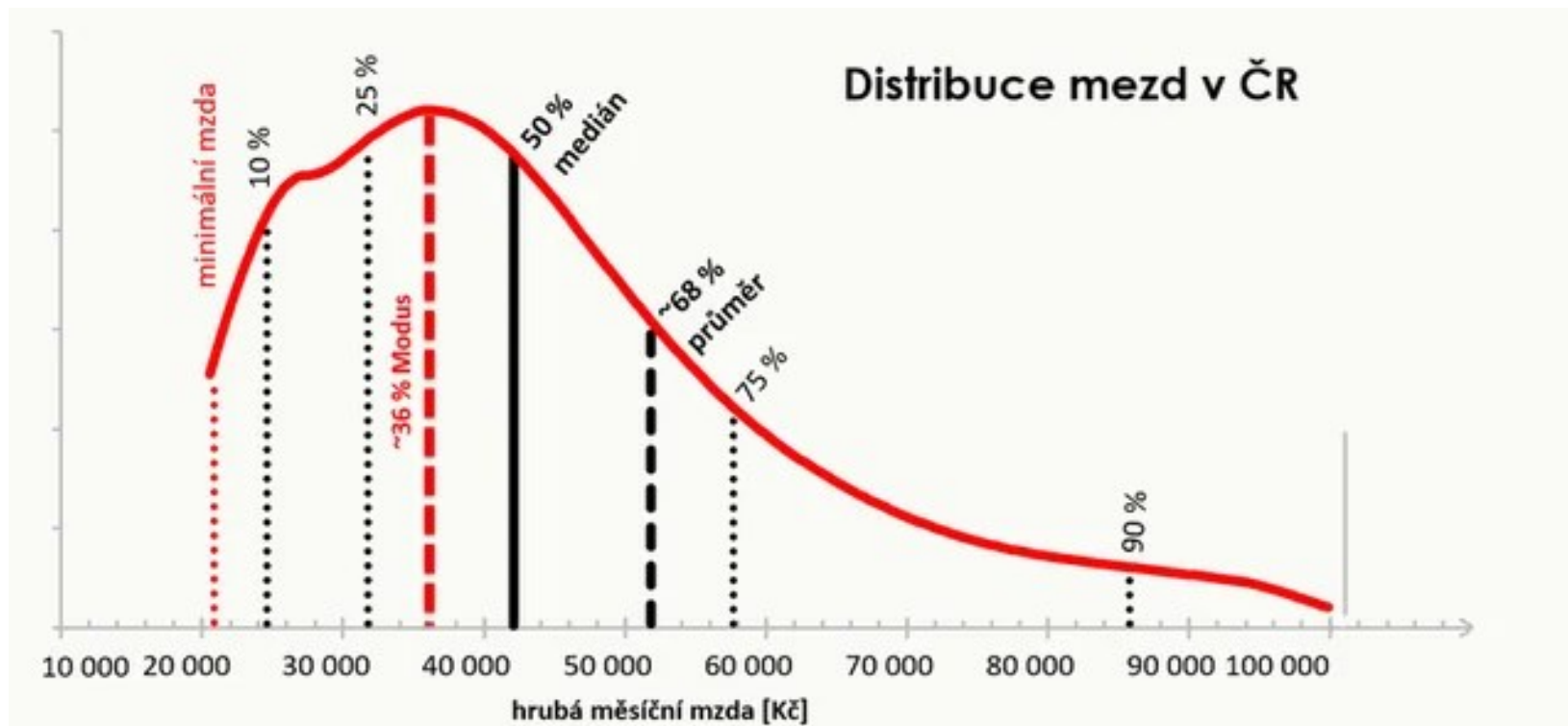
BETA(.5, .4)



Normal Mixture



Beware the average





When normality fails

When normality fails

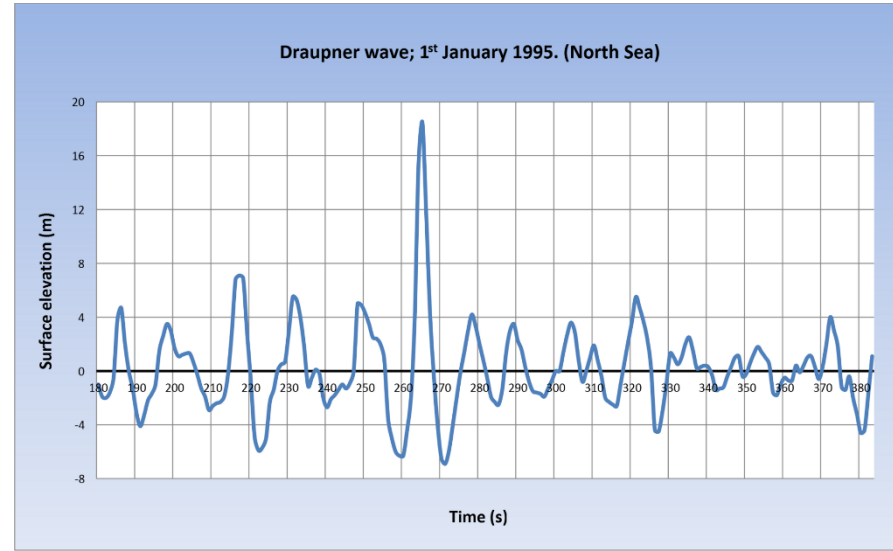
- Black Monday (Oct 19, 1987)
 - Dow Jones dropped 22.6% in single day
 - **25 sigma event** (5 sigma daily event happens once in 10 000 years)

When normality fails

- Black Monday (Oct 19, 1987)
 - Dow Jones dropped 22.6% in single day
 - **25 sigma event** (5 sigma daily event happens once in 10 000 years)
- 2008 global financial crisis
 - Mortgage defaults **highly correlated**

When normality fails

- Black Monday (Oct 19, 1987)
 - Dow Jones dropped 22.6% in single day
 - **25 sigma event** (5 sigma daily event happens once in 10 000 years)
- 2008 global financial crisis
 - Mortgage defaults **highly correlated**
- Rogue waves
 - Gaussian model:
 $\approx 12 \text{ m waves} \Rightarrow \text{max. } 15 \text{ m waves}$
 - Draupner wave (1995)
 $\approx 12 \text{ m waves} \Rightarrow 25.6 \text{ m}$
 - Waves add up together
 - Schrödinger equation may apply





Log-Log scale

Multiplicative nature

- > Many natural processes are **multiplicative**, not additive
 - whole **economy**...

GDP per capita, 1870 to 2016

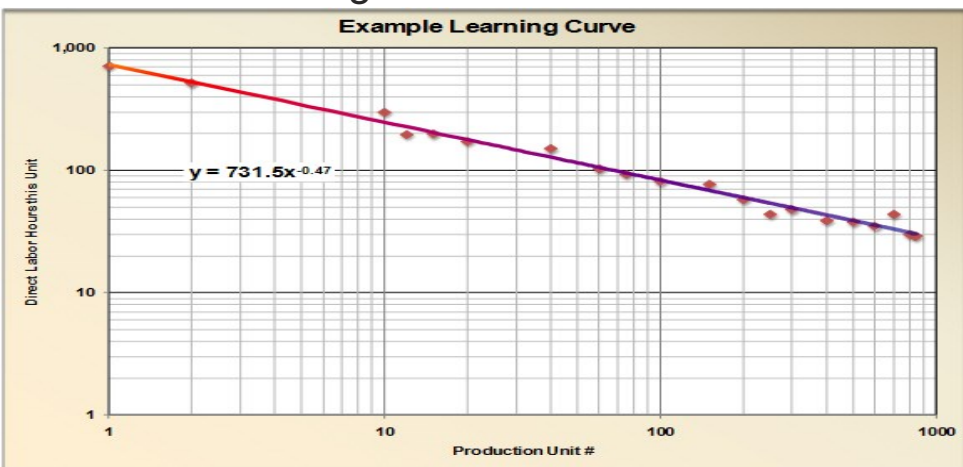
GDP per capita adjusted for price changes over time (inflation) and price differences between countries – it is measured in international-\$ in 2011 prices.

LINEAR LOG



Multiplicative nature

- Many natural processes are **multiplicative**, not additive
 - whole **economy**...
- Sometimes, the **other axis is also not linear**
 - Learning curve...



GDP per capita, 1870 to 2016

GDP per capita adjusted for price changes over time (inflation) and price differences between countries – it is measured in international-\$ in 2011 prices.

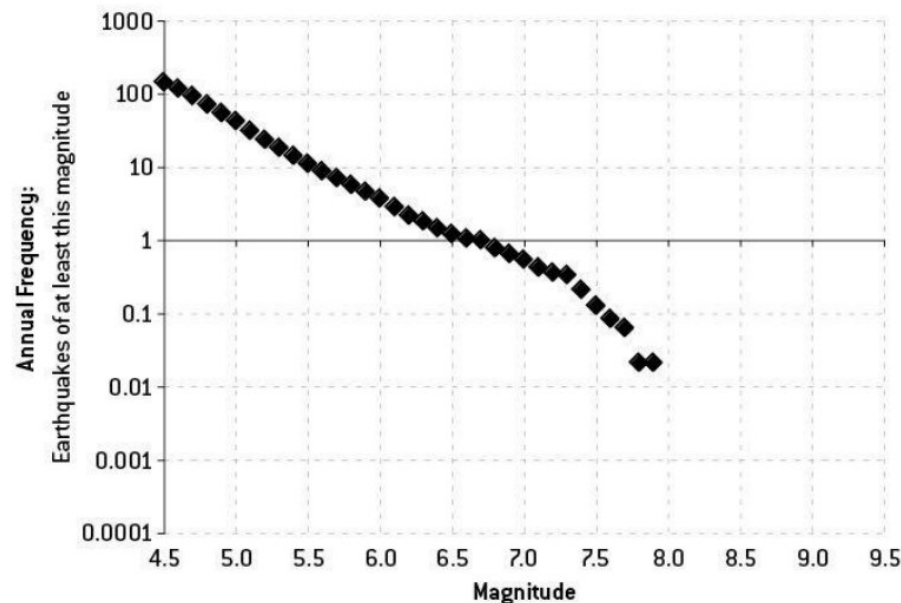
LINEAR LOG



Earthquakes & Fukushima

- Archeological evidence of ≈ 40 m tsunamis disregarded
- Reactor build to withstand 8.6
- Earthquake in 2011 was 9.1

FIGURE 5-7A: TŌHOKU, JAPAN EARTHQUAKE FREQUENCIES
JANUARY 1, 1964–MARCH 10, 2011



NATO terrorism

FIGURE 13-4: TERROR ATTACK FREQUENCY BY DEATH TOLL IN NATO COUNTRIES, 1979–2009 (LINEAR SCALE)

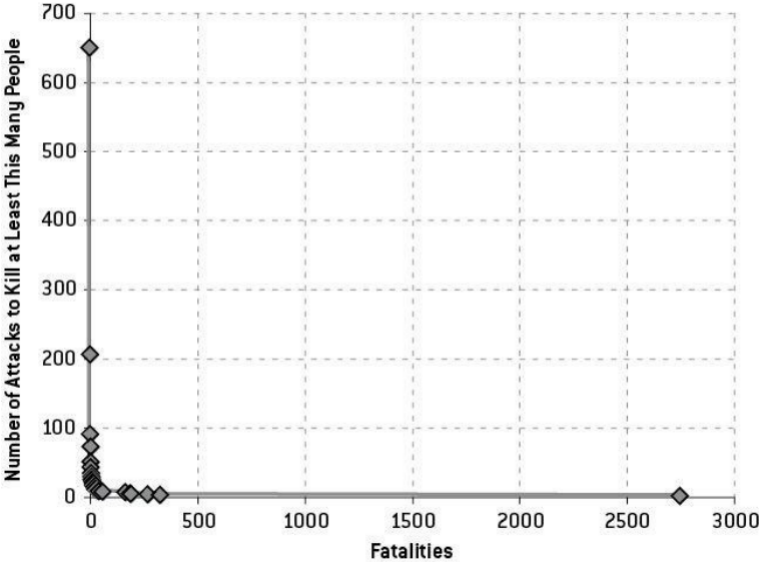
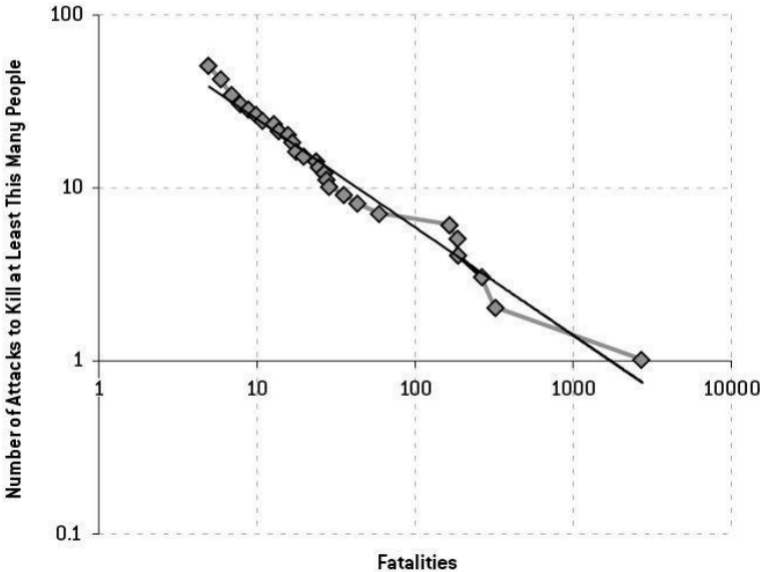


FIGURE 13-5: TERROR ATTACK FREQUENCY BY DEATH TOLL, NATO COUNTRIES, 1979–2009 (LOGARITHMIC SCALE)



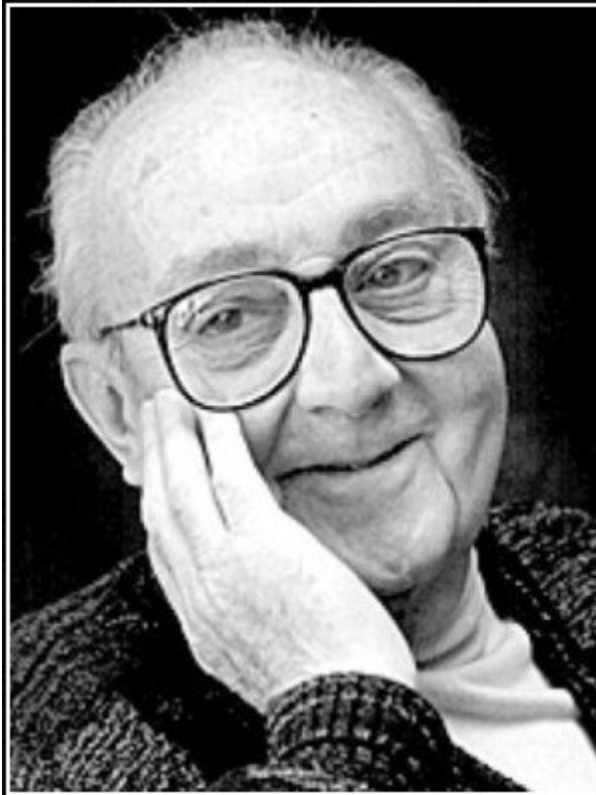
NATO terrorism

- What's the biggest threat?
- Nuclear attack, killing $\approx 1\,000\,000$ people
 $\approx 1 : 1\,500$ years
 ≈ 667 deaths / year
- Airplane attack, killing $\approx 3\,000$ people
 $\approx 1 : 30$ years
 ≈ 100 deaths / year
- Where should the security efforts aim?





Pitfalls of measures and models



All models are approximations.
Essentially, all models are wrong, but
some are useful. However, the
approximate nature of the model
must always be borne in mind.

— *George E. P. Box* —

Goodhart's law

GOODHART'S LAW

WHEN A MEASURE BECOMES A TARGET,
IT CEASES TO BE A GOOD MEASURE

IF YOU
MEASURE
PEOPLE ON...

NUMBER OF
NAILS MADE

WEIGHT OF
NAILS MADE

THEN YOU
MIGHT GET

1000'S OF
TINY NAILS

A FEW GIANT,
HEAVY NAILS



Goodhart's law

GOODHART'S LAW

WHEN A MEASURE BECOMES A TARGET,
IT CEASES TO BE A GOOD MEASURE

IF YOU
MEASURE
PEOPLE ON...

NUMBER OF
NAILS MADE

WEIGHT OF
NAILS MADE

THEN YOU
MIGHT GET

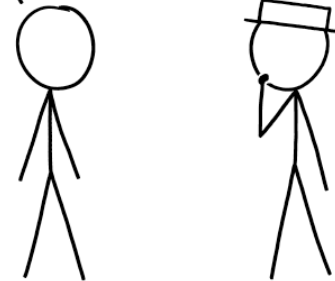
1000'S OF
TINY NAILS

A FEW GIANT,
HEAVY NAILS



WHEN A METRIC BECOMES A TARGET,
IT CEASES TO BE A GOOD METRIC.

SOUNDS BAD. LET'S OFFER
A BONUS TO ANYONE WHO
IDENTIFIES A METRIC THAT
HAS BECOME A TARGET.



Lucas critique

- Traditional forecasting or modeling often assumes that the **relationships observed in the past will hold in the future**
- But **when a major change occurs** (e.g., a new rule, technology, incentive, or policy), **people adjust their behavior**
- Those adjustments can **undermine or completely alter the original relationships** on which the predictions were based

Lucas critique - examples

> Phillips curve

- Expectations are the driving factor
- Lower unemployment $\Leftrightarrow$ unexpectedly high inflation

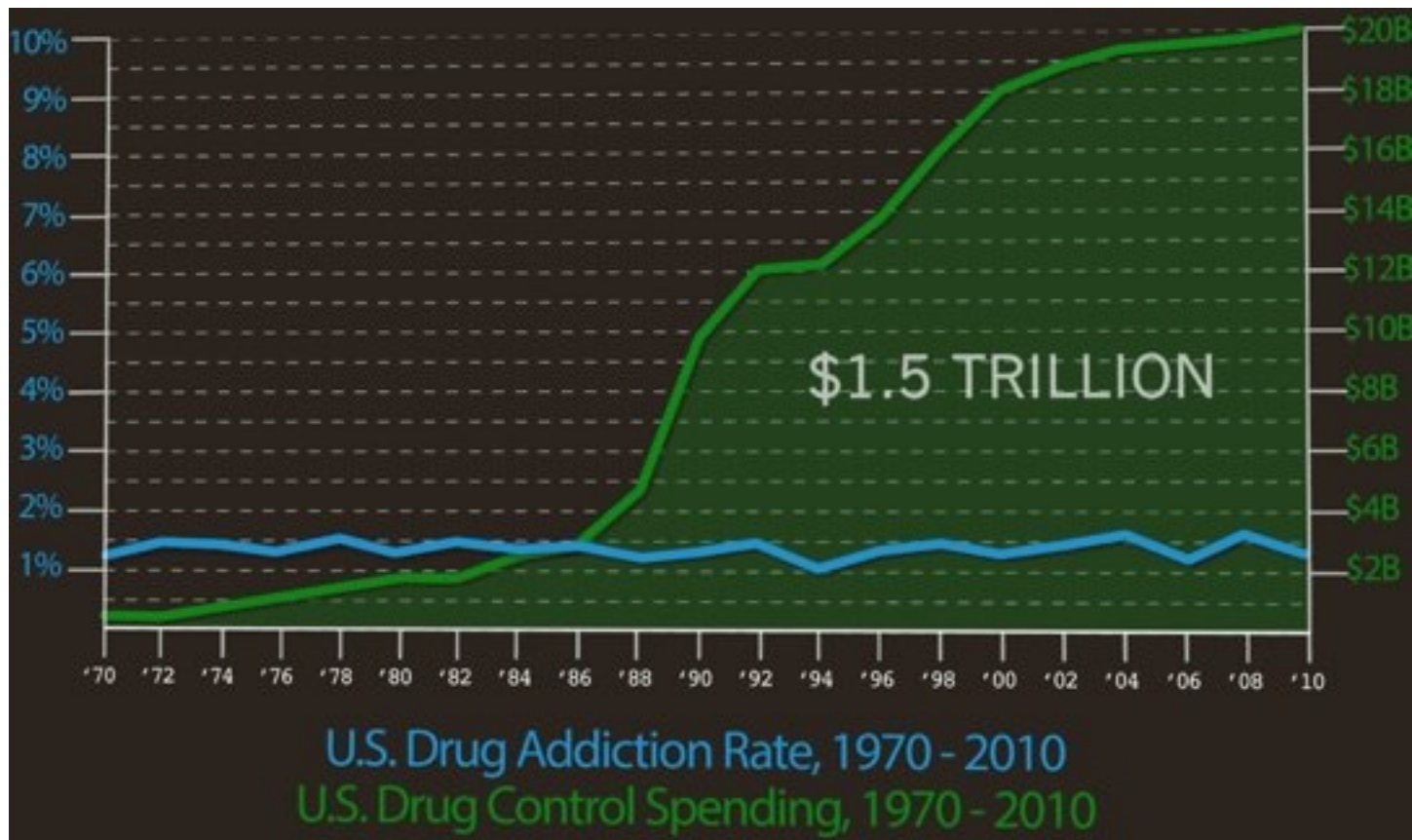


Lucas critique - examples

- > Phillips curve
 - Expectations are the driving factor
 - Lower unemployment $\Leftrightarrow$ unexpectedly high inflation
- > Ban guns $\Rightarrow$ people will use other weapons
- > Add security checks $\Rightarrow$ terrorists attack the que before the checkpoint
- > New highway $\Rightarrow$ shorter travel times $\Rightarrow$ more cars...



War on drugs





**Justice system doesn't understand
conditional probability**

People v. Collins

- 1968 American robbery trial in California noted for its misuse of probability and as an example of the **prosecutor's fallacy**



Black man with beard	1 in 10
Man with mustache	1 in 4
White woman with pony tail	1 in 10
White woman with blond hair	1 in 3
Yellow motor car	1 in 10
Interracial couple in car	1 in 1,000

People v. Collins

- Ignoring dependencies between the characteristics
- $P(A|B) \neq P(B|A)$

People v. Collins

- Ignoring dependencies between the characteristics
- $P(A|B) \neq P(B|A)$
- $P(4 \text{ legs} \mid \text{dog}) \neq P(\text{dog} \mid 4 \text{ legs})$

People v. Collins

- Ignoring dependencies between the characteristics
- $P(A|B) \neq P(B|A)$
- $P(4 \text{ legs} | \text{dog}) \neq P(\text{dog} | 4 \text{ legs})$
- $P(\text{evidence} | \text{innocent}) \neq P(\text{innocent} | \text{evidence})$

$$P(I|E) = \frac{P(E|I) * P(I)}{P(E|I) * P(I) + P(E|G) * P(G)}$$

Sally Clark

- Her 2 sons suffered SIDS in 1996 and 1998
- Probability of SIDS said to be 1 : 8 543
- Prosecution claimed 1 : 73 000 000 chance
- Convicted of double murder and sentenced to life imprisonment in 1999
- At least one of the sons actually died of Staphylococcus infection
- Released in 2003
- Botched, nonsensical statistics never addressed
- Never mentally recovered, died in 2007



Discussion



Thank You for Your attention

Profinit EU, s.r.o.
Tychonova 2, 160 00 Praha 6

Tel.: + 420 224 316 016, web: www.profinit.eu

 **LINKEDIN**
linkedin.com/company/profinit

 **TWITTER**
[@profinit_EU](https://twitter.com/profinit_EU)

 **FACEBOOK**
facebook.com/Profinit.EU

 **YOUTUBE**
Profinit EU, s.r.o.