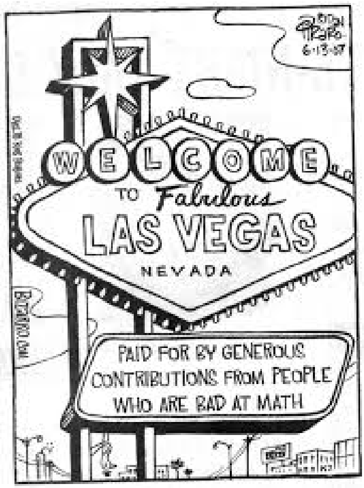

WELCOME
TO Fabulous
LAS VEGAS
NEVADA
PAID FOR BY GENEROUS CONTRIBUTIONS FROM PEOPLE WHO ARE BAD AT MATH

Probabilistic thinking

Erik Štrumbelj

2019

Would you agree or disagree?

Probability is fundamental to statistical practice.

Understanding probability is key to understanding statistical models and how to fit and interpret them.

What is probability?

I think that...

- probability is a language for expressing uncertainty.
- Like any language it has rules, but the rules of probability are very simple.
- Non-mathematicians are too often scared away by the mathematics of computing with probabilistic objects.
- Good news: the calculus and algebra of mathematical manipulations are practically almost irrelevant.

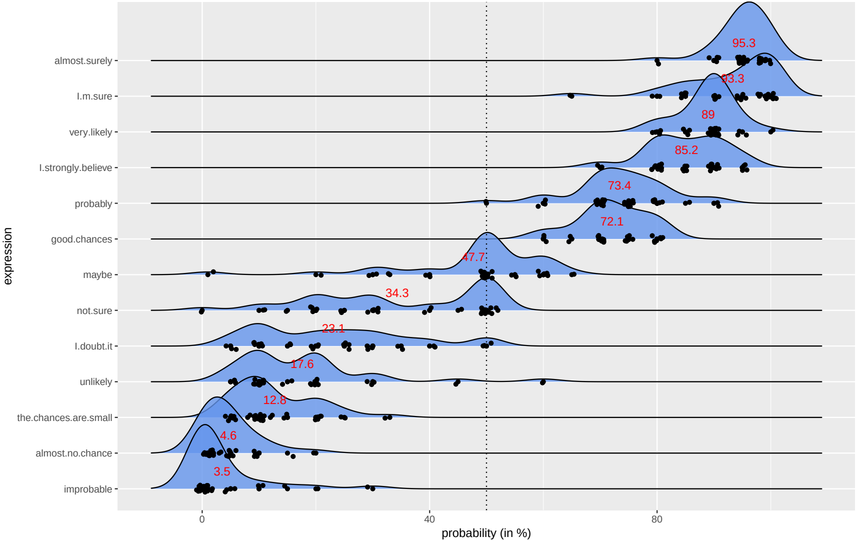
Probabilistic thinking

The goal of this lecture are to think about and discuss the following:

- The main task of statistics (and life) is making decisions in the presence of uncertainty.
- Probability is a formal language for expressing uncertainty.
- Probability is actually quite natural and the bare minimum we need to understand each other.
- Probability distributions are like expressions in a language - the more you know, the more expressive you can be.

Q: Will it rain in Moscow next Tuesday?

Natural language has support for expressing uncertainty



Q: What will the temperature ($^\circ$C) be in Moscow at midnight?

Important observations

- Uncertainty is part of everyday life...
- we've even developed natural language to express it!
- However, natural language is imprecise, incoherent and lacks expressiveness...
- which makes it useless for serious quantitative work.

Conclusion: **We need to agree on a more formal language for expressing uncertainty!**

Formal definition of probability

Set of outcomes Ω (set of all possible truths, sample space).

Set of events S (all subsets of Ω that we will assign probability to[i]).

A probability function P a function $P : S \to [0, 1]$ that satisfies the axioms of probability:

A1 $P(A) \geq 0$, for all $A \in S$.

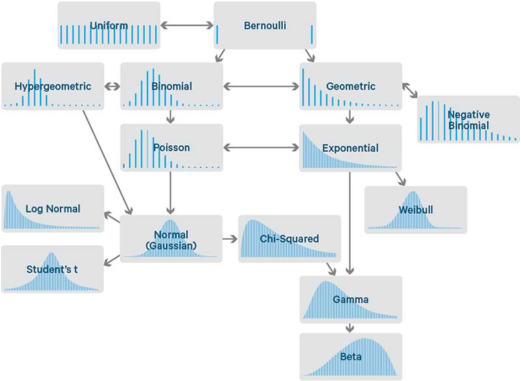
A2 $P(\Omega) = 1$.

A3 $P(A_1 \cup A_2 \cup A_3 \cup ...) = \sum_{i=1}^{\infty} P(A_i)$,
for any sequence of disjunct events.

Why is probability defined like this and not something else?

[i] The formalism used here is a $\sigma-$algebra.

Distributions

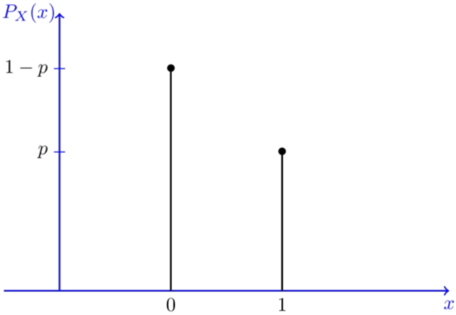


Probability distributions are the elementary expressions of probabilistic thinking and the basic building blocks of statistical models.

Distributions follow the axioms of probability and are therefore coherent and precise statements.

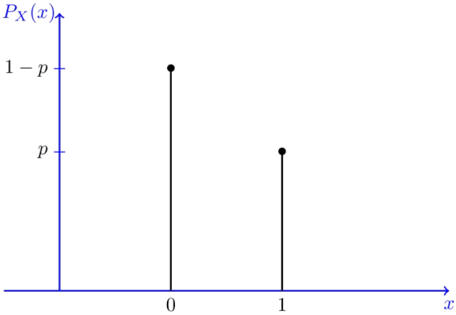
Bernoulli distribution

$$X \sim Bernoulli(p)$$



Bernoulli distribution

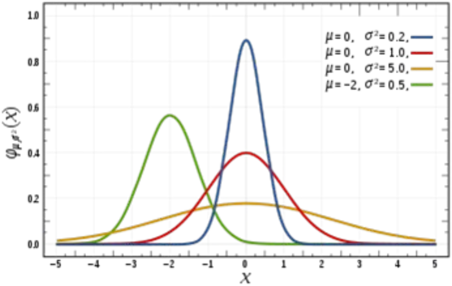
$$X \sim Bernoulli(p)$$



Q: Will it rain in Moscow next Tuesday?

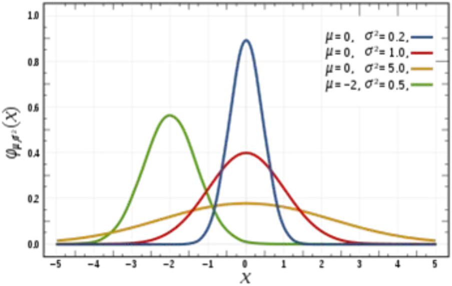
Normal distribution

$$X \sim N(\mu, \sigma^2)$$



Normal distribution

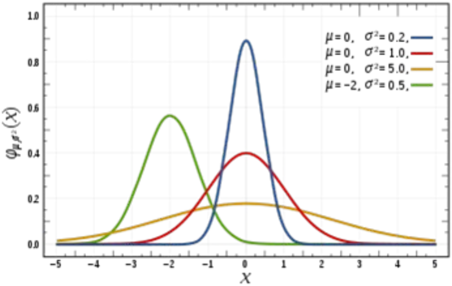
$$X \sim N(\mu, \sigma^2)$$



Q: What will the temperature (°C) be in Moscow at midnight?

Normal distribution

$$X \sim N(\mu, \sigma^2)$$

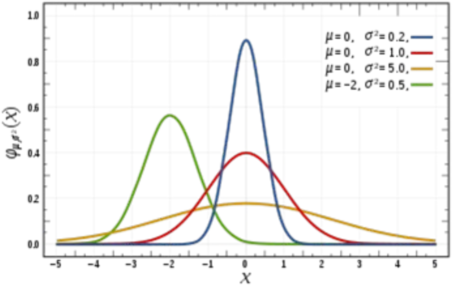


Q: What will the temperature (°C) be in Moscow at midnight?

Q: What was the temperature (°C) in Moscow exactly 50 years ago?

Normal distribution

$$X \sim N(\mu, \sigma^2)$$



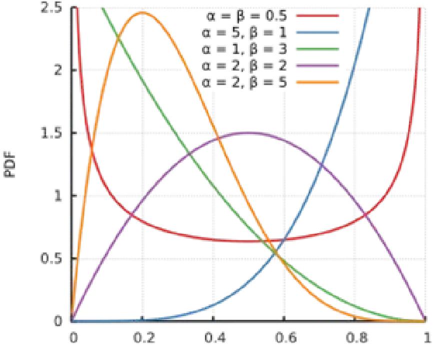
Q: What will the temperature (°C) be in Moscow at midnight?

Q: What was the temperature (°C) in Moscow exactly 50 years ago?

What is the cause of uncertainty in the second case?
What are some other common causes of uncertainty?

Beta distribution

$$X \sim Beta(\alpha, \beta)$$



Tricky questions

Assume that these are independent tosses of a (possibly unfair) coin:

T T H T T H T T T H (?)

- Q1: Will the 11th flip be heads or tails?

Tricky questions

Assume that these are independent tosses of a (possibly unfair) coin:

T T H T T H T T T H (?)

- Q1: Will the 11th flip be heads or tails?
- Q2: What is the coin's probability of landing heads?

Tricky questions

Assume that these are independent tosses of a (possibly unfair) coin:

T T H T T H T T T H (?)

- Q1: Will the 11th flip be heads or tails?
- Q2: What is the coin's probability of landing heads?
- Q3: Is this coin practically fair (is its probability between 49% and 51%)?

Summary

- Probability theory is a precise language for expressing uncertainty.
- If we don't follow the rules of probability, others won't understand what we are saying! However, probability doesn't require us to be objective or even sensible.[ii]
- And we find it natural to have probabilistic opinions about properties that are arguably not random.[iii]

[ii] For example, it's quite coherent to say that the world will end tomorrow with 0.99999 probability and there's a 0.00001 probability that it won't. But saying that there is a 0.99999 probability that it won't and 0.00002 probability that it will - that just doesn't make sense.

[iii] This is inherently Bayesian!

If I had just one more slide on probability, this is what I'd write on it

Definition of conditional probability:

$$P(A|B) = \frac{P(AB)}{P(B)}$$

Marginal probability and Bayes' theorem:

Let $\mathcal{H} = \{H_1, ..., H_n\}$ be a partition such that $P(\mathcal{H}) = 1$. For any event E:
$P(E) = \sum_{i=1}^{n} P(EH_i)$ and

$$P(H_i|E) = \frac{P(H_i E)}{P(E)} = \frac{P(E|H_i)P(H_i)}{P(E)} = \frac{P(E|H_i)P(H_i)}{\sum_{j=1}^{n} P(E|H_j)P(H_j)}.$$

Definition of independence and conditional independence:

A and B are independent $\Leftrightarrow P(AB) = P(A)P(B)$

A and B are cond. independent given C $\Leftrightarrow P(AB|C) = P(A|C)P(B|C)$

Note: Independence does not imply conditional independence (or vice versa).

For continuous distributions:

$p(x|y) = \frac{p(x,y)}{p(y)}$, $p(x) = \int p(x,y)dy$, independence: $p(x,y) = p(x)p(y)$...

Further reading

See [HOFF:Ch02], [BDA:Ch01] for more on Bayesian probability. For a more thorough understanding of Bayesian probability, in particular, how axioms of probability and conditional probability follow from rationally avoiding being a sure loser in a betting game, see [POU:Ch01].

BDA Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). Bayesian data analysis. CRC press.

HOFF Hoff, P. D. (2009). A first course in Bayesian statistical methods. New York: Springer.

POU Kadane, J. B. (2011). Principles of uncertainty. CRC Press.