Philosophy of Science and Statistical Reasoning

Introduction

But first, …

📢 Message Erwin van Vliet (programme director Bachelor Psychobiology).

Statistical reasoning module:

‘How can we use statistical inference to learn about the world?’

Remember?

- Methoden van Onderzoek en Statistiek (Lourens Waldorp)
- The Analysis of Biological Data (Whitlock & Schluter)
- R / RStudio (An Introduction to R)
- SOWISO

Expand foundation for follow-up courses and independent learning.

Expand knowledge and skills in statistics and statistical reasoning, develop statistical intuition, avoid common pitfalls and fallacies, and so on.

Module organization*

(*study guide is leading)

Lectures

1. Introduction
2. Frequentist inference
3. Linear regression & *t*-test
4. Moderation / interaction
5. ANOVA I & nonparametric inference
6. ANOVA II (complex models)
7. Bayesian inference

Team

Jonas van Nijnatten (co-coordinator), Anna Ansems and Guusje Schokker (teaching assistants), Alexander Savi (lecturer)

Assignments (formative)

- Weekly, sufficient/insufficient, deadline every Sunday 23:55 (correct answers available after deadline)
- Pass with 5/7 sufficient assignments
- 3 attempts per assignment, 2 prior checks per attempt
- ~~Exemption for recidivists (if previously sufficient)~~

Exam (summative)

- ⅔ Statistical Reasoning (⅓ Ph. of Sc.)
- 80% of final grade

Take control

> No one can teach you to play the guitar.
>
> But they can help you learn.
>
> — Dan Morgan, 1965, Guitar

During the course

- Read the literature
- Ask questions at the lectures
- Use the Canvas discussion forum
- Put effort in assignments
- Discuss & collaborate with peers
- Join think along sessions
- Give (anonymous) feedback
- *Play* with the offered material

After the course

- Fill in course evaluation form

Take control

- 💡 Get inspired
- 💭 Think it through
- 🛠️ Do it yourself
- 🚁 Helicopter view
- 💩 Uh?

🏕️ Slow down: other approaches to the same material

👩‍💻 R code

🚄 Speed up: related, more advanced material, getting a bigger picture

🧠 Learning theory: strategies and insights

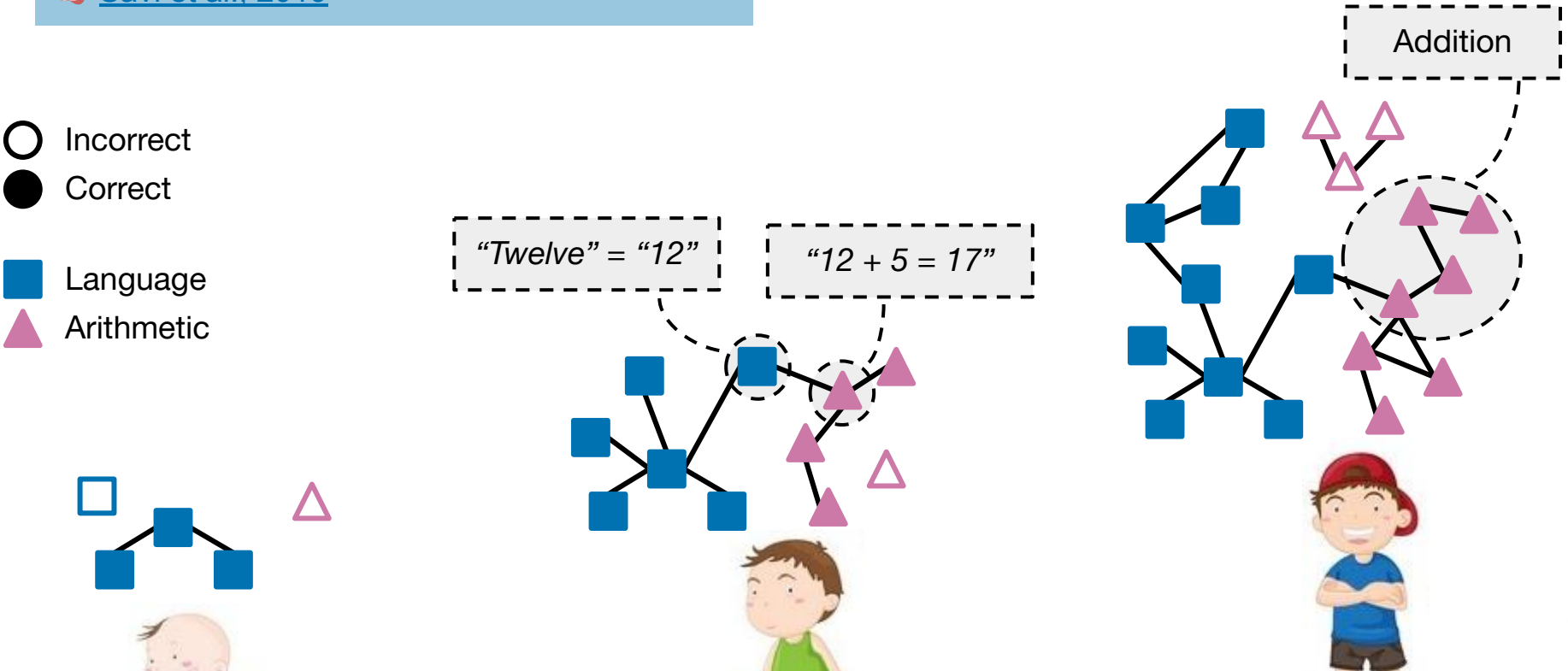
💩 Emoji science

The wiring of intelligence

🧠 Savi et al., 2019

Incorrect
Correct

Language
Arithmetic



What will we learn today?

Topics

Statistical reasoning
Empirical cycle
Probability distributions
Frequentist inference
Linear regression
t-Test
Moderation
ANOVA
Nonparametric inference
Bayesian inference

Questions

‘What is statistical reasoning?’ ‘Why is it important?’

‘Can we distinguish different types of inference?’ ‘How does statistical inference relate to those?’

‘What are probability distributions?’

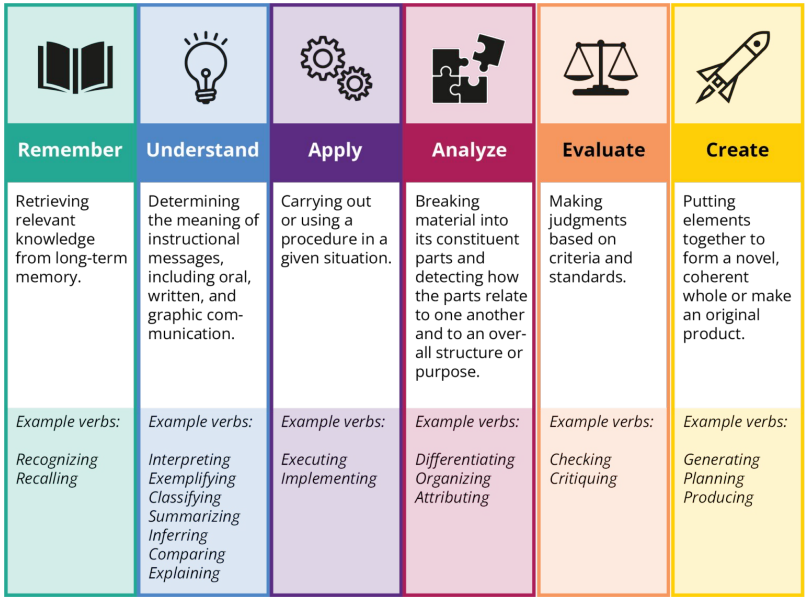
Statistical reasoning

Statistical [reasoning, literacy, thinking].

Deals with inference (or prediction) when there is variability, probability, randomness, uncertainty, …

- Remember: vocabulary
- Understand: e.g., confidence intervals
- Apply: statistical procedures, R skills
- Analyze: e.g., frequentist inference
- Evaluate: fallacies, generalizability, other's work
- Create: e.g., simulations, new knowledge

🚁 Bloom's taxonomy can be applied to any domain, e.g., systems thinking.



Remember	Understand	Apply	Analyze	Evaluate	Create
Retrieving relevant knowledge from long-term memory.	Determining the meaning of instructional messages, including oral, written, and graphic com-munication.	Carrying out or using a procedure in a given situation.	Breaking material into its constituent parts and detecting how the parts relate to one another and to an over-all structure or purpose.	Making judgments based on criteria and standards.	Putting elements together to form a novel, coherent whole or make an original product.
Example verbs: *Recognizing* *Recalling*	*Example verbs:* *Interpreting* *Exemplifying* *Classifying* *Summarizing* *Inferring* *Comparing* *Explaining*	*Example verbs:* *Executing* *Implementing*	*Example verbs:* *Differentiating* *Organizing* *Attributing*	*Example verbs:* *Checking* *Critiquing*	*Example verbs:* *Generating* *Planning* *Producing*

Bloom's taxonomy (cognitive domain). Illustration by Utrecht University

Why?

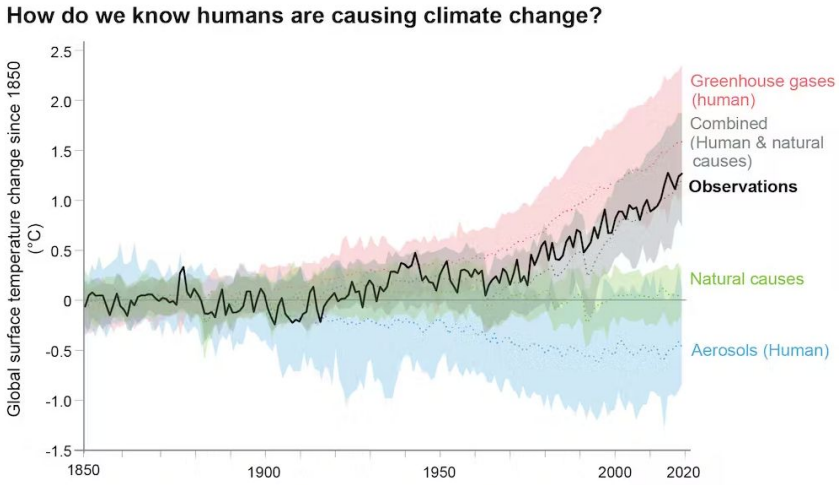


Illustration by IPCC (in The Conversation)

Illustration by Daniel Hertzberg (in Quanta)

More in Nautilus

Why?

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Which is more probable?

1. Linda is a bank teller.
2. Linda is a bank teller and is active in the feminist movement.

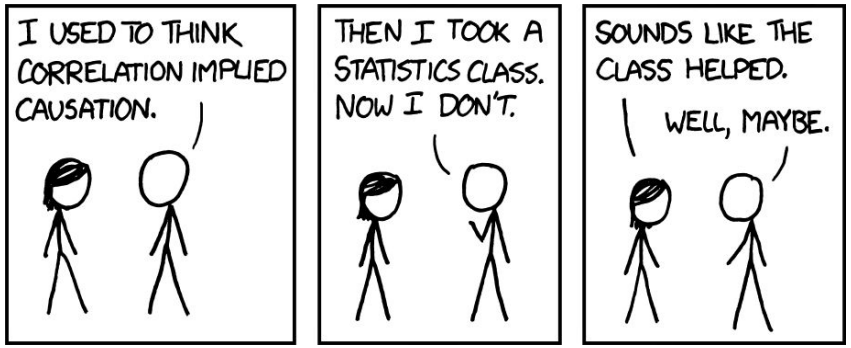


Illustration by Randall Munroe

Conjunction fallacy

Correlation does not imply causation

XKCD explained

💩 Why?

Perspectives on Psychological Science
Volume 6, Issue 3, May 2011, Page 313

https://doi.org/10.1177/1745691611406928

SAGE journals

Bayesians Caught Smuggling Priors Into Rotterdam Harbor

Rogier A. Kievit

Department of Psychology, University of Amsterdam, The Netherlands

Corresponding author(s):
Roger A. Kievit, Department of Psychology, University of Amsterdam, Roetersstraat 15, 1018WB Amsterdam, The Netherlands E-mail: r.a.kievit@uva.nl

Amsterdam, April 13, 2011. A group of international Bayesians was arrested today in the Rotterdam harbor. According to Dutch customs, they were attempting to smuggle over 1.5 million priors into the country, hidden between electronic equipment. The arrest represents the largest capture of priors in history.

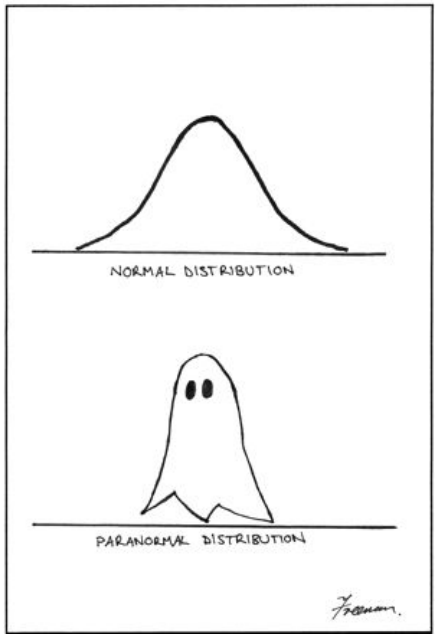
"This is our biggest catch yet. Uniform priors, Gaussian priors, Dirichlet priors, even informative priors, it's all here," says customs officers Benjamin Roosken, responsible for the arrest. "There are priors for memory experiments, intelligence tests, flanker tasks, meta-analyses, political preference, everything! God only knows what would have happened if this had gotten through. We're pretty lucky to catch them too. The chance of being in the right place, given the right time, if you take into account the number of arrests, divided by the number of successful arrests every year, it's pretty slim. We're very glad indeed."

Sources suggest that the shipment of priors was going to be introduced into the Dutch scientific community by "white-washing" them. "They are getting very good at it. They found ghost-journals with fake articles, refer to the papers where the priors are allegedly based on empirical data, and before you know it, they're out in the open. Of course, when you look up the reference, everything is long gone," says Roosken.

Until recently, the Dutch government adopted a lenient, pragmatic approach toward priors. As an anonymous source states, "It was quite simple. Scientists were allowed to *use* priors, but not to create them at home. It may sound a bit counterintuitive, but it worked quite well, for a while at least." However, according to critics, this policy created an uncontrollable backdoor industry.

The discovery of international smuggling rings has caused the government to revise its strategy and crack down hard on illegal trade. The capture of the smuggling ring symbolizes a new, tough stance on priors. "We will not stand for this unjustified and illegal use of priors any longer," says Roosken. If found guilty, the defendants may face 12 years in prison (95% CI [10.2, 13.8], $p < .01$).

[Kievit, 2011](#) 🔒



[Freeman, 2006](#)

🥼 Statistical immersion

Skunk

Five rounds. Each round:

1. You stand up
2. I roll two dice
3. If I roll a 1, the game is over and your score for that round is lost
4. If I don't roll a 1, you add the number on the dice to your score (write it down)
5. If you sit down before I roll the dice, you forfeit that round and carry your score over to the next round

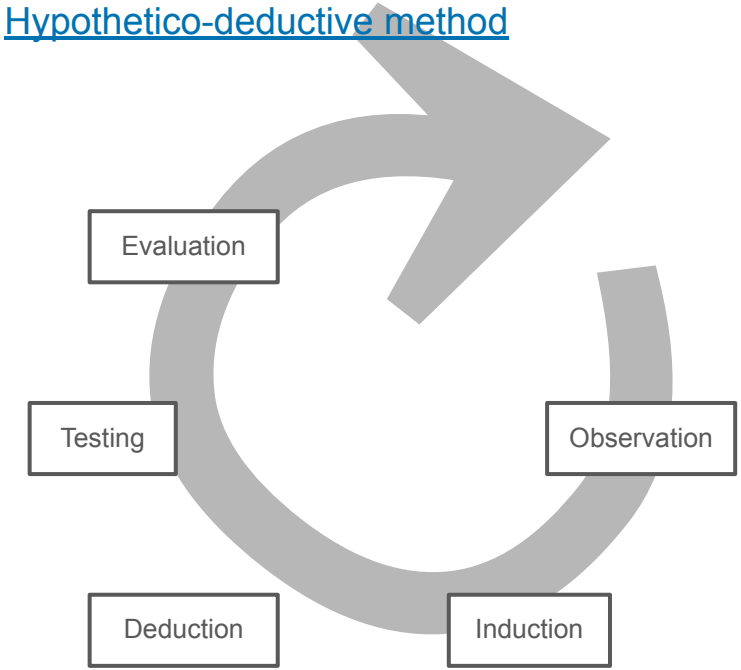
The winner collects the most coins in all rounds.

👨‍💻 ?sample

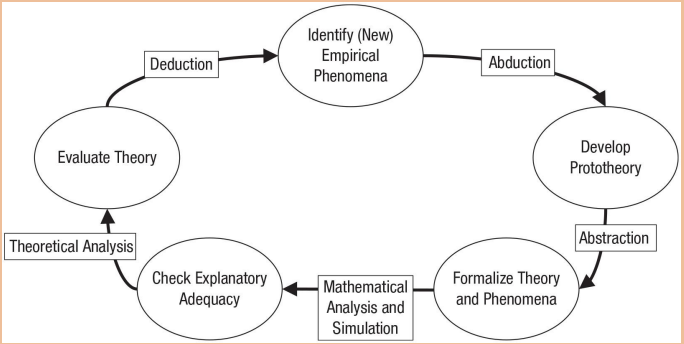
15:00

Empirical cycle

[Hypothetico-deductive method](#)



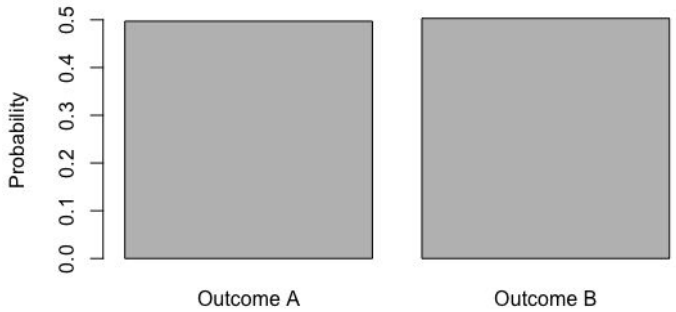
🚁 Multiple, complementary scientific methods; [abductive method](#), [generative method](#)



[Borsboom et al., 2021](#)

Probability distributions

> " A probability distribution is the mathematical function that gives the probabilities of occurrence of different possible outcomes for an experiment.
> — Wikipedia



🤦‍♂️

```
p_heads <- .5
n_rep <- 1000
toss <- sample(
  x = c("Heads", "Tails"),
  size = n_rep,
  replace = TRUE,
  prob = c(p_heads, 1 - p_heads))
prob <- table(toss) / length(toss)
barplot(prob, ylab = "Probability")
```

Binomial distribution

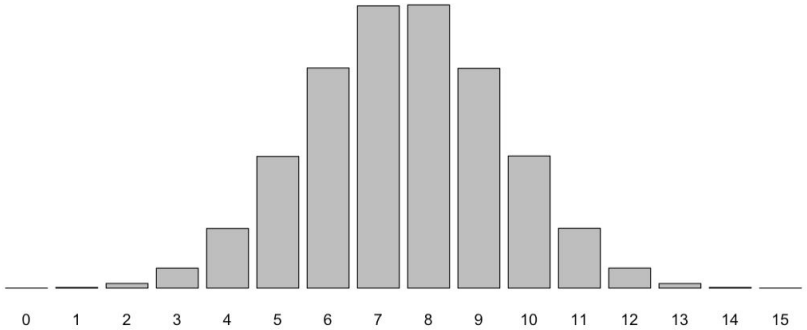
Bi = two
Nomial = terms
Discrete

- 🪙 Head / Tail (/ Edge)
- 🪙 Correct / Incorrect
- 🪙 Success / Failure
- 🪙 True / False
- 🪙 Yes / No (/ Maybe)
- 🪙 Female / Male (/ Intersex)
- 🪙 … / …

👩‍💻

```
p_heads <- .5
n_rep <- 1000
toss <- sample(
  x = c("Heads", "Tails"),
  size = n_rep,
  replace = TRUE,
  prob = c(p_heads, 1 - p_heads))
prob <- table(toss) / length(toss)
barplot(prob, ylab = "Probability")
```

Binomial distribution



👩‍💻

```
n_rep = 1000000  # e.g., persons
n = 15  # e.g., items
dat <- rbinom(n = n_rep, size = n, prob = .5)
tab <- table(dat)
barplot(tab, xlab = "Sum", ylab = "Frequency")
```

💭 If we change the probabilities,

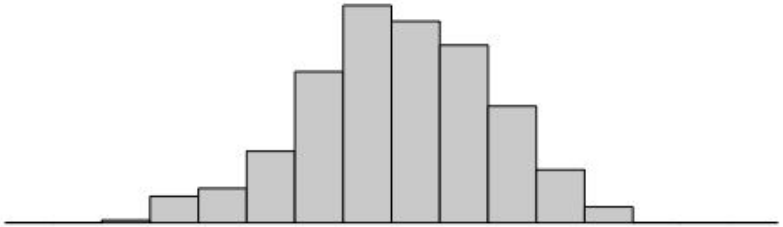
- will the frequency distribution change?
- will the probability distribution change?

💭 If we change the number of repetitions,

- will the frequency distribution change?
- will the probability distribution change?

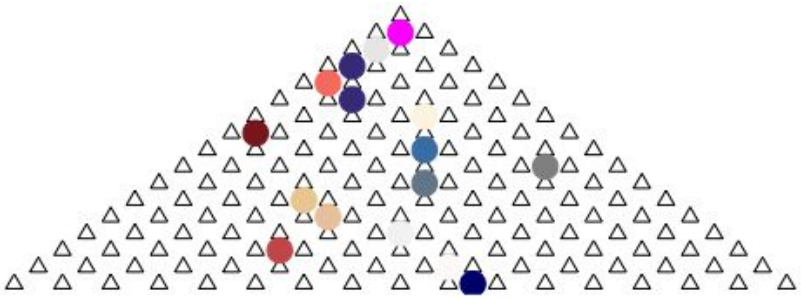
Galton board

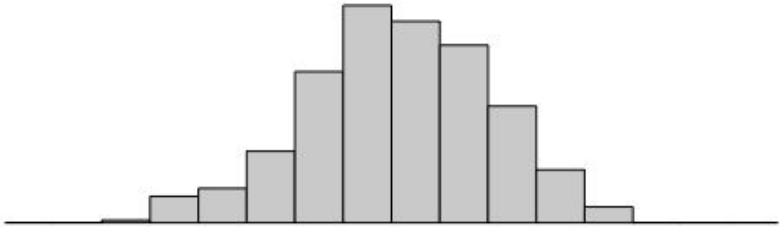
Video by Wikipedia



👩‍💻

```
library("animation")
n_rep <- 500
n <- 15
n_layers <- n + 2
animation::ani.options(interval = 0.05, nmax =
n_rep + n_layers)
animation::quincunx(balls = n_rep, layers =
n_layers)
```





You take an exam with 15 two-choice items. Where on the Galton board are the answers to the following questions captured?

- The sum score of the test is 8. How many possible ways can get you to that sum score? (1)
- What is the probability of sum score 8? (2)
- What is the probability of precisely this series (with sum score 8): 000010110011111?
- What is the probability of sum score 8 or less? (3)
- Which lowest sum scores have a probability of 10% or less? (4)
- What are the two factors that determine these probabilities? (binomial theorem!)

👩‍💻

```
choose(15, 8)  # (1) number of ways to score an 8
dbinom(8, 15, .5)  # (2) probability of sum score 8
pbinom(8, 15, .5)  # (3) probability of sum score 8 or less
qbinom(.1, 15, .5)  # (4) lowest sum scores with probability 10% or less
rbinom(n = 500, size = 15, prob = .5)  # sample 500 sum scores
```

🚁 rbinom(); rnorm(); rt(); rf()

You take an exam with 15 two-choice items. Where on the Galton board are the answers to the following questions captured?

- The sum score of the test is 8. How many possible ways can get you to that sum score? (1)
- What is the probability of sum score 8? (2)
- What is the probability of precisely this series (with sum score 8): 000010110011111?
- What is the probability of sum score 8 or less? (3)
- Which lowest sum scores have a probability of 10% or less? (4)
- What are the two factors that determine these probabilities? (binomial theorem!)

Probability distributions

Discrete probability distributions

- Bernoulli distribution
- Binomial distribution

Continuous probability distributions

- Normal/Gaussian distribution
- *t*-distribution
- Gamma distribution
- *F*-distribution

💡 The joy of stats (Hans Rosling)

🛠️ Probability distributions (Seeing Theory)

💩 Buy a probability distribution

🚁 List of probability distributions (Wikipedia)

Cooling down

What did we learn?

Assuming you know nothing more about Alice, which of 1-5 is most likely? Or does 6 hold?

1. Alice is a rock star or she works in a bank.
2. Alice is quiet and works in a bank.
3. Alice is a rock star.
4. Alice is honest and works in a bank.
5. Alice works in a bank.
6. There is no definite answer.

Think it through (1 minute) and write down your answer.

What did we learn?

Assuming you know nothing more about Alice, which of 1-5 is most likely? Or does 6 hold?

1. Alice is a rock star and works in a bank.
2. Alice is quiet and works in a bank.
3. Alice is quiet and reserved and works in a bank.
4. Alice is honest and works in a bank.
5. Alice works in a bank.
6. There is no definite answer.

Think it through (1 minute) and write down your answer.

Take-home assignments

📅 Weekly assignment

Q3: Sharon is the previous lecturer.

🍻 Pub quiz

Create an *informative* four-choice question about the content of today's lecture.

An informative question has a large spread in responses across answer options.

Clarify answer options (which are (in)correct and why).

Illustration adapted from Snippets.com

Statistical reasoning
Empirical cycle
Probability distributions
Frequentist inference
Sample / sampling distribution
Central limit theorem
Normal distribution
P value
Type I/II errors
Effect size
Confidence interval
Power
Test statistics
Linear regression
t-Test
Moderation
ANOVA
Nonparametric inference
Bayesian inference

Illustration by Jennifer Cheuk

Colophon

Slides
alexandersavi.nl/teaching/