Foundations of Probability

Random experiments

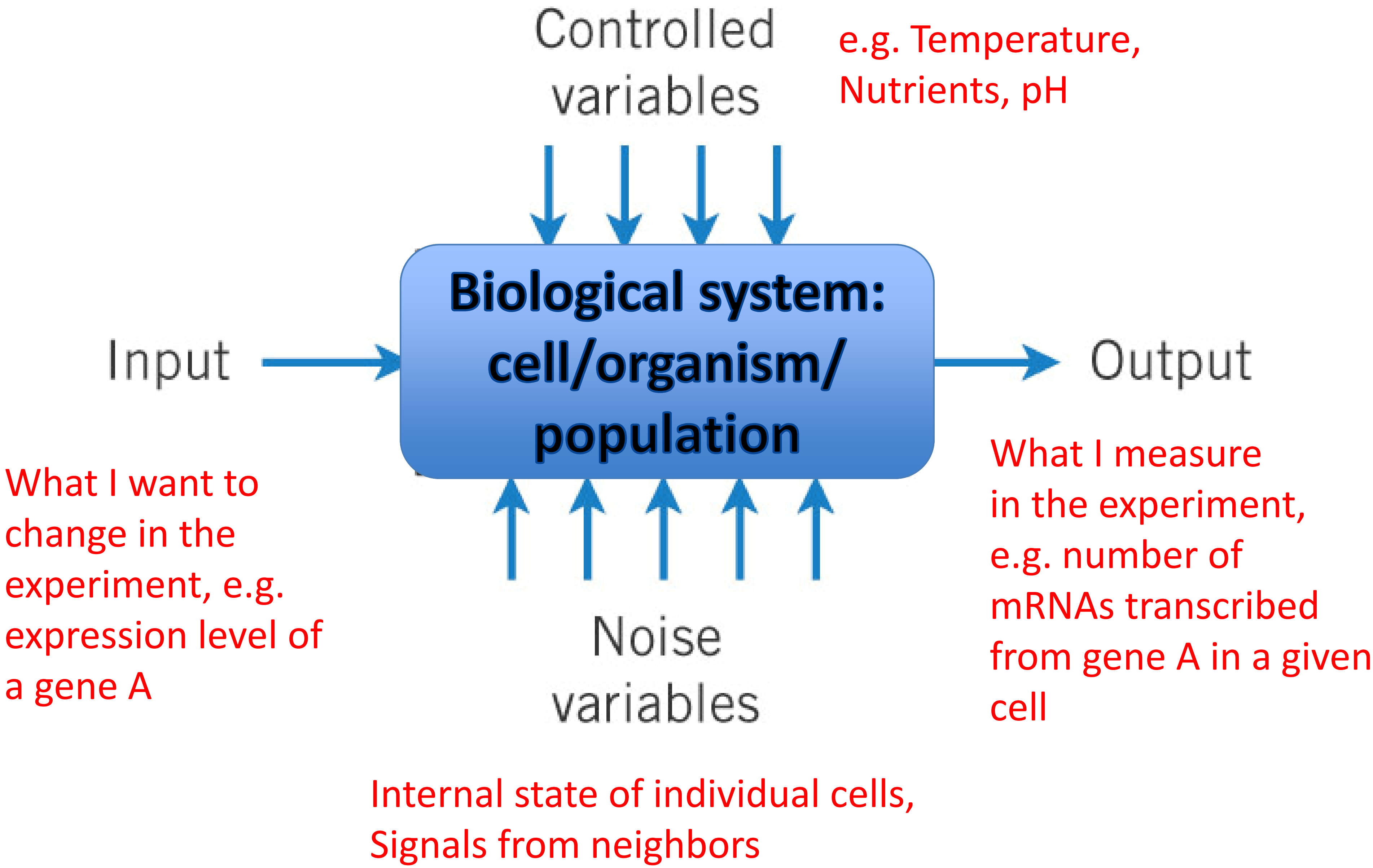
Sample spaces

Venn diagrams of random events

Random Experiments

- An experiment is an operation or procedure, carried out under controlled conditions
  - Example: measure the number of mRNA transcribed from gene A in an individual cell
- An experiment that can result in different outcomes, even if repeated in the same manner every time, is called a random experiment
  - Cell-to-cell variability due to history/genome variants
  - Noise in external parameters such as temperature, nutrients, pH, etc.
- Evolution offers ready-made random experiments
  - Genomes of different species
  - Genomes of different individuals within a species
  - Individual cancer cells isolated from the same patient

Variability/Noise Produce Output Variation



Are you in class?

A. Yes
B. No
C. I am not sure, I am still asleep

Get your i-clickers

Sample Spaces

- Random experiments have unique outcomes.
- The set of all possible outcomes of a random experiment is called the sample space, S.
- S is discrete if it consists of a finite or countable infinite set of outcomes.
- S is continuous if it contains an interval (either a finite or infinite width) of real numbers.

Examples of a Sample Space

- Experiment measuring the abundance of mRNA transcribed from gene A
  $S = \{x | x >= 0\}$: continuous.
- Bin it into four groups
  $S = \{\textit{below 10, 10-30, 30-100, above 100}\}$: discrete.
- Is gene "on" (mRNA above 30)?
  $S = \{\textit{true, false}\}$: logical/Boolean/discrete.

Event

An event (*E*) is a subset of the sample space of a random experiment, i.e., one or more outcomes of the sample space.

- The union of two events is the event that consists of all outcomes that are contained in either of the two events. We denote the union as $E_1 \cup E_2$

- The intersection of two events is the event that consists of all outcomes that are contained in both of the two events. We denote the intersection as $E_1 \cap E_2$

- The complement of an event in a sample space is the set of outcomes in the sample space that are not in the event. We denote the complement of the event E as E' (sometimes E^c or $\bar{E}$)

Examples

Discrete

1. Assume you toss a coin once. The sample space is $S = \{H, T\}$, where H = head and T = tail and the event of a head is $\{H\}$.

2. Assume you toss a coin twice. The sample space is $S = \{(H, H), (H, T), (T, H), (T, T)\}$, and the event of obtaining exactly one head is $\{(H, T), (T, H)\}$.

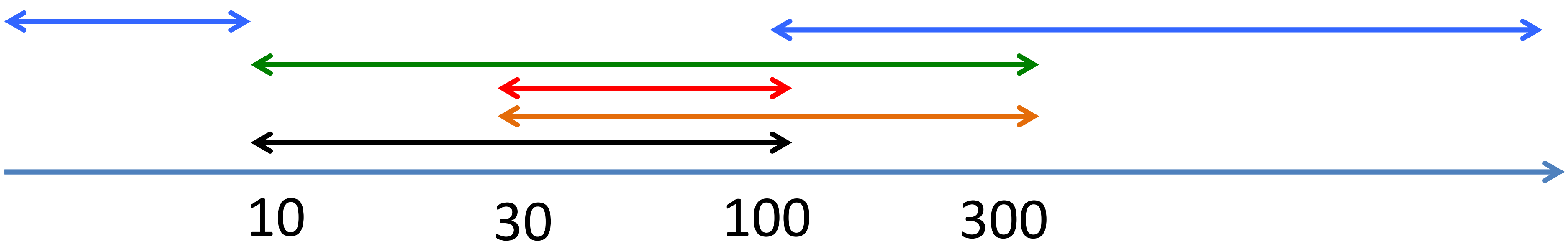
Continuous

Sample space for the expression level of a gene: $S = \{x | x \geq 0\}$

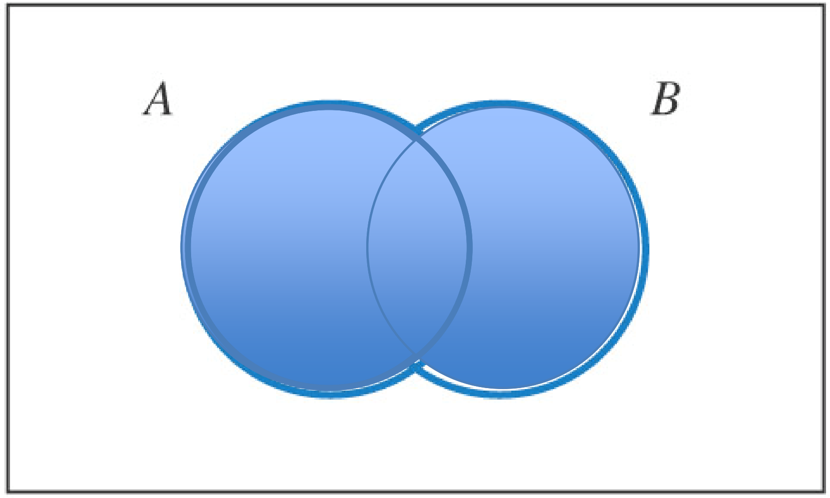
Two events:

- E1 = {x|10 < x < 100}
- E2 = {x|30 < x < 300}

- E1 ∩ E2 = {x|30 < x < 100 }
- E1 U E2 = {x|10 < x < 300}
- E1’ = {x|x ≤ 10 or x ≥ 100}



Venn diagrams

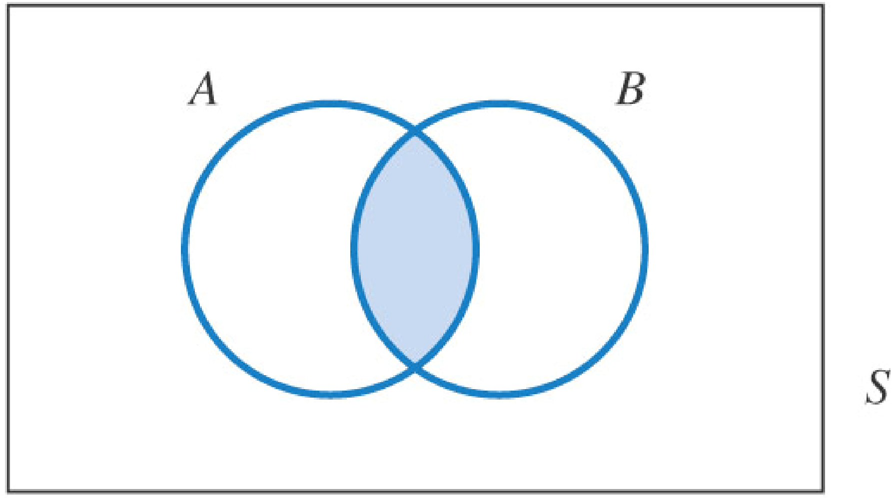


A U B

S

John Venn (1843-1923)
British logician

Venn diagrams

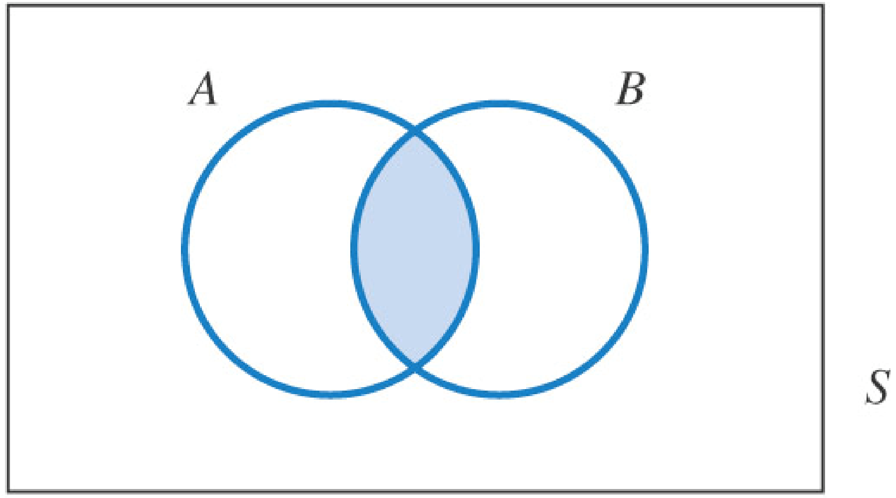


Which formula describes the blue region?

A. A ∪ B
B. A ∩ B
C. A’
D. B’

Get your i-clickers

Venn diagrams



Which formula describes the blue region?

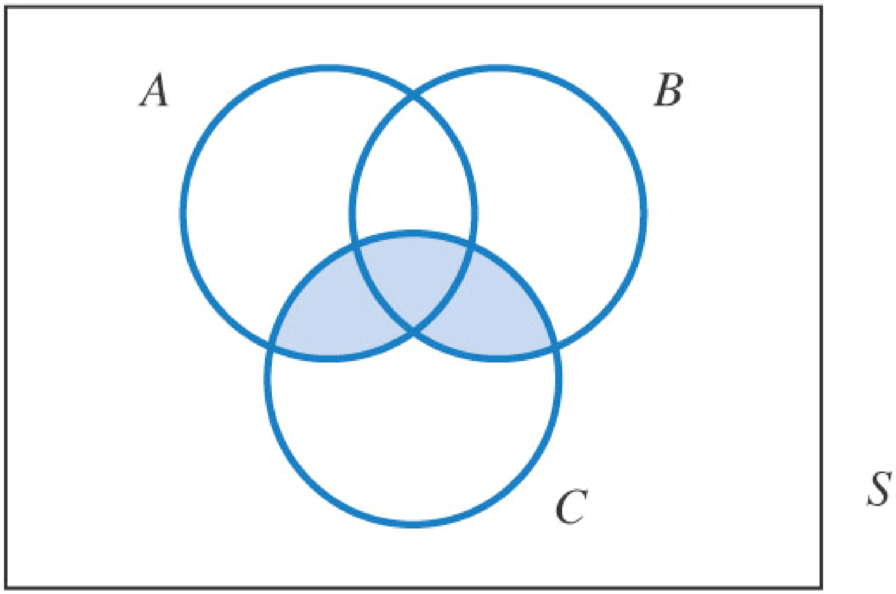
A. A ∪ B

B. A ∩ B

C. A’

D. B’

Venn diagrams



Which formula describes the blue region?

A. (A U B) ∩ C

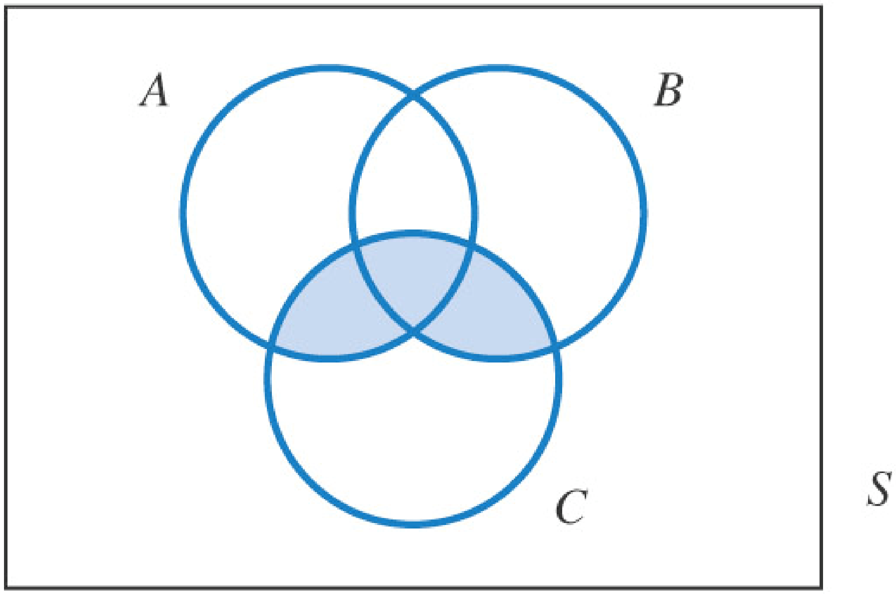
B. (A ∩ B) ∩ C

C. (A U B) U C

D. (A ∩ B) U C

Get your i-clickers

Venn diagrams



Which formula describes the blue region?

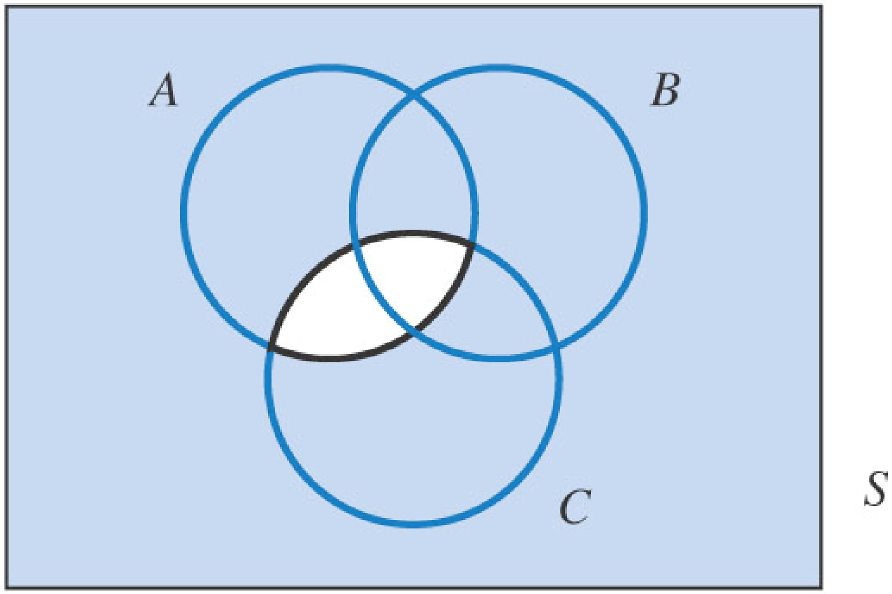
A. $(A \cup B) \cap C$

B. $(A \cap B) \cap C$

C. $(A \cup B) \cup C$

D. $(A \cap B) \cup C$

Venn diagrams



Which formula describes the blue region?

A. A ∩ C

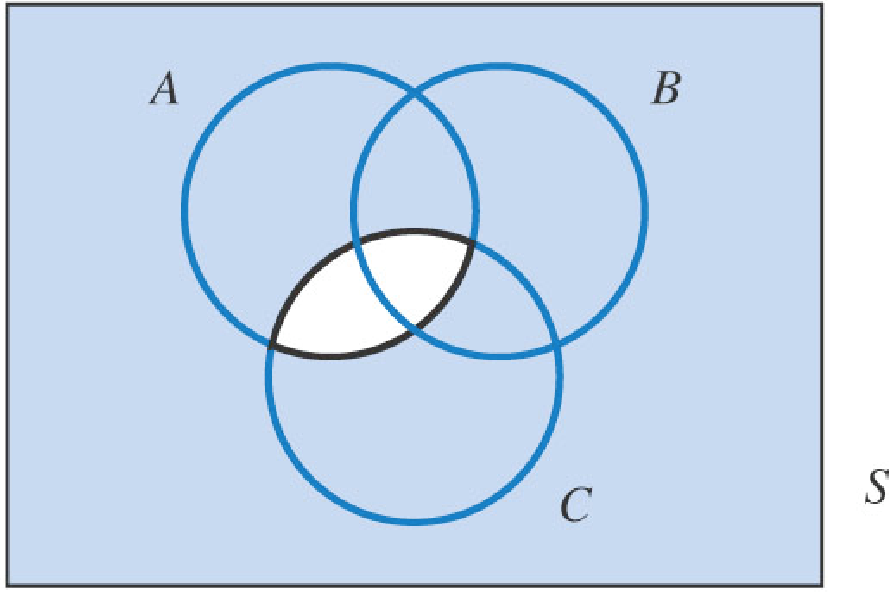
B. A’ U C’

C. (A ∩ B ∩ C)’

D. (A ∩ B) ∩ C

Get your i-clickers

Venn diagrams



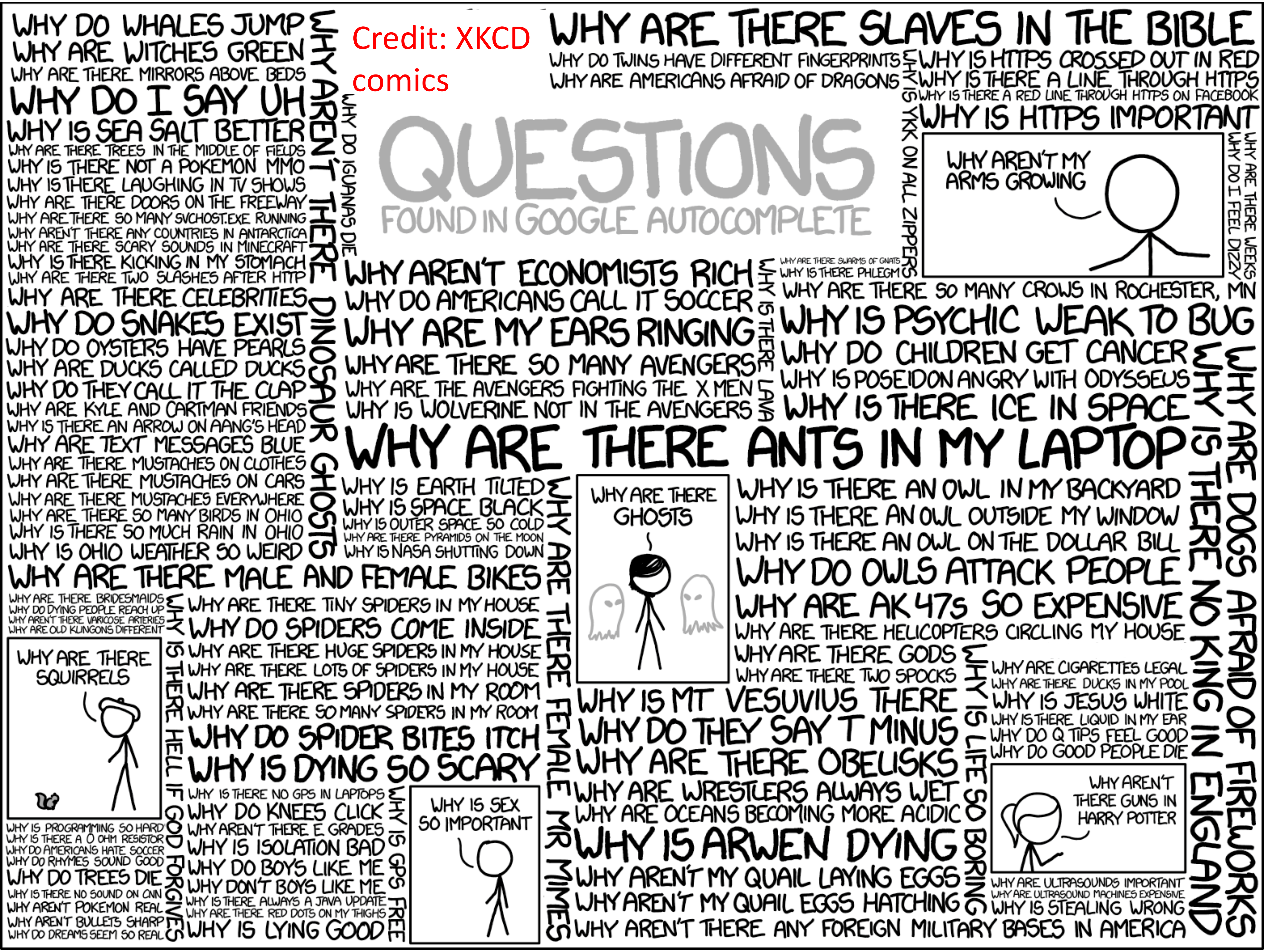
Which formula describes the blue region?

A. $A \cap C$

B. $A' \cup C'$

C. $(A \cap B \cap C)'$

D. $(A \cap B) \cap C$



Credit: XKCD comics

Definitions of Probability

Two definitions of probability

- (1) STATISTICAL PROBABILITY: the relative frequency with which an event occurs in the long run

- (2) INDUCTIVE PROBABILITY: the degree of belief which it is reasonable to place in a proposition on given evidence

Bulmer, M. G.. Principles of Statistics (Dover Books on Mathematics)

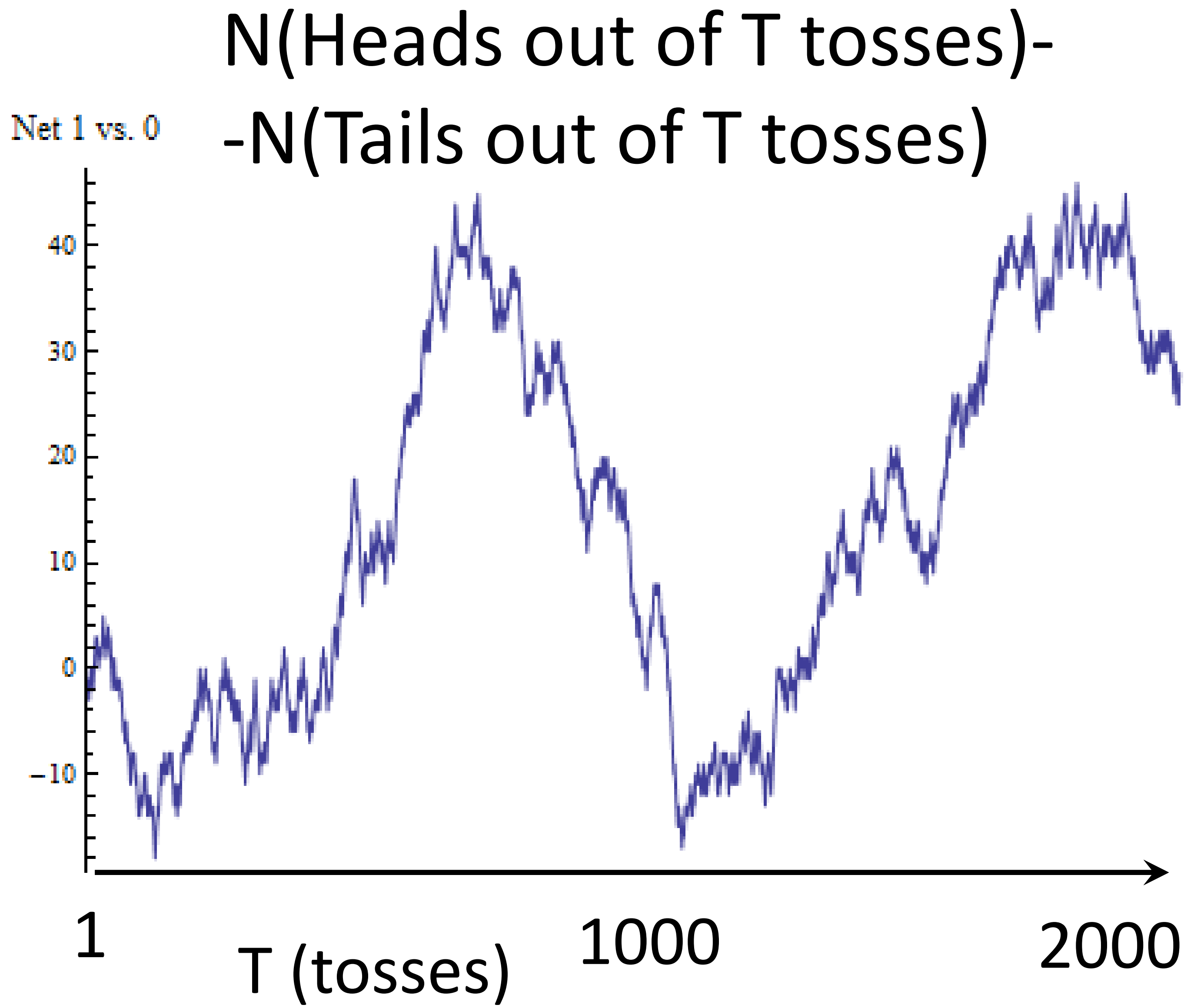
Statistical Probability

A statistical probability of an event is the limiting value of the relative frequency with it occurs in a very large number of independent trials

Empirical

Statistical Probability of a Coin Toss

John Edmund Kerrich
(1903–1985)
British/South African mathematician

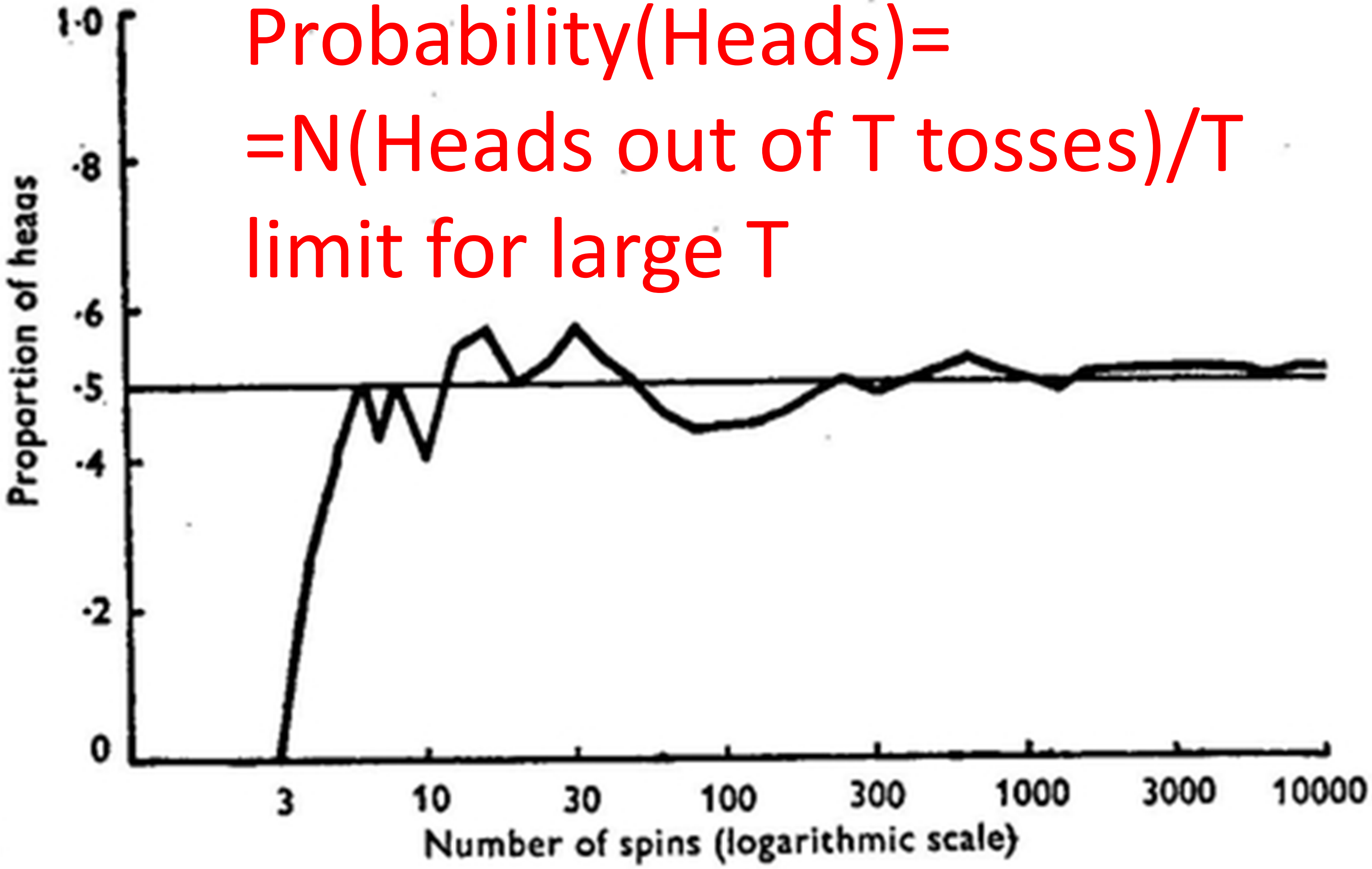


Excess of heads among the first 2,000 out of 10,000 tosses performed (Kerrich 1946)

Statistical Probability of a Coin Toss

John Edmund Kerrich
(1903–1985)
British/South African mathematician

Probability(Heads)=
=N(Heads out of T tosses)/T
limit for large T

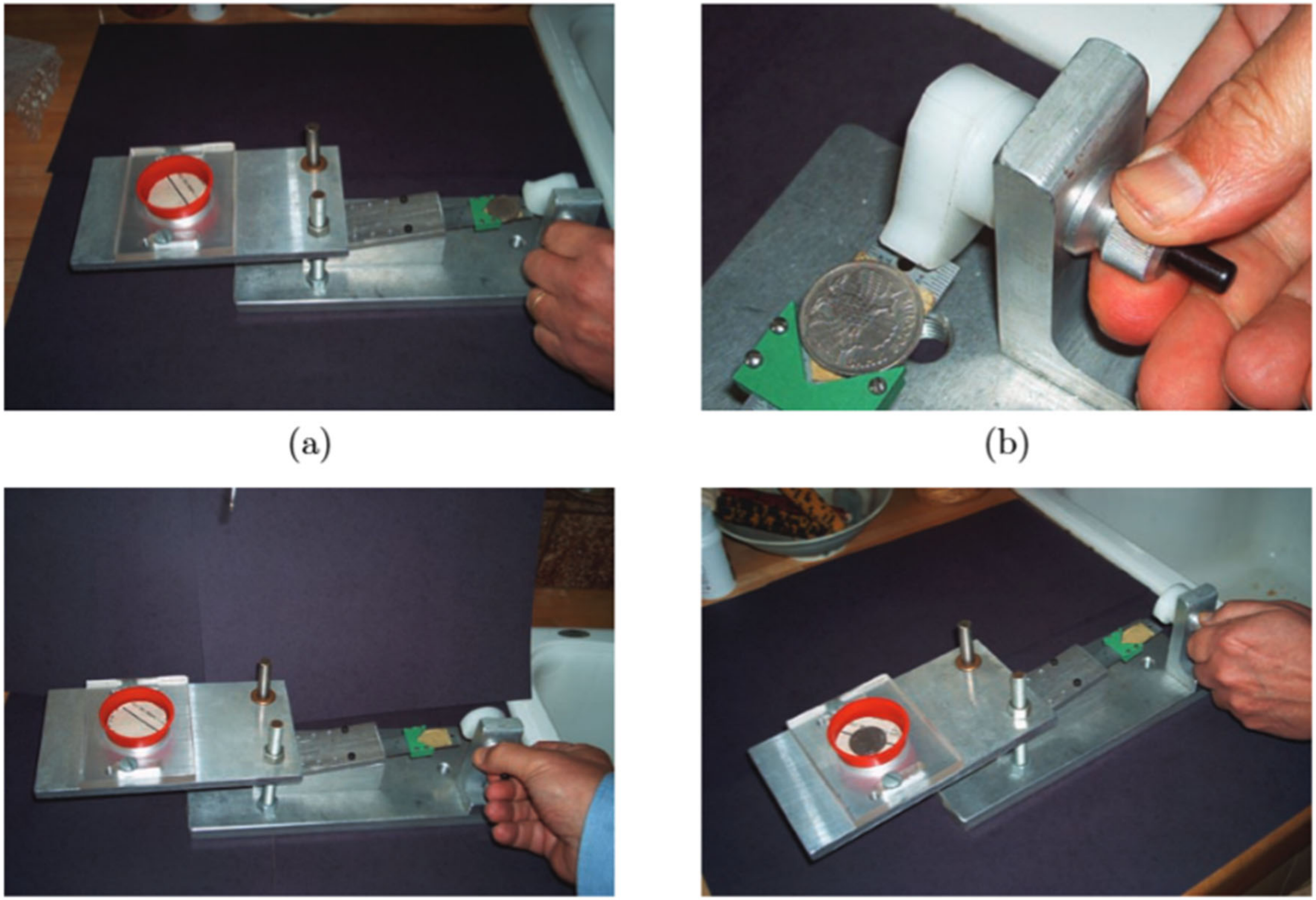


Proportion of heads among 10,000 coin tosses (Kerrich 1946)

Kerrich's predecessors and followers

- In 1777 French naturalist Count de Buffon collected a series of 2048 uninterrupted coin flips
  - possibly the first statistical experiment ever conducted
  - de Buffon: out of Africa hypothesis, theory of evolution (struggle for existence, heredity)
- In 1897 the statistician Karl Pearson flipped a coin 24,000 times to obtain 12,012 tails
- In 1946 John Kerrich flipped a coin 10,000 times for a total of 5067 heads
- In 2007 Diaconis, Holmes, and Montgomery from Stanford and UC Santa Cruiz predicted that a coin tends to land on the side that was facing up when it was tossed

Information from Bartoš F, Sarafoglou A, Godmann HR, Sahrani A, Leunk DK, Gui PY, et al. J Am Stat Assoc. 2025;ahead-of-print: 1–10. doi:10.1080/01621459.2025.2516210

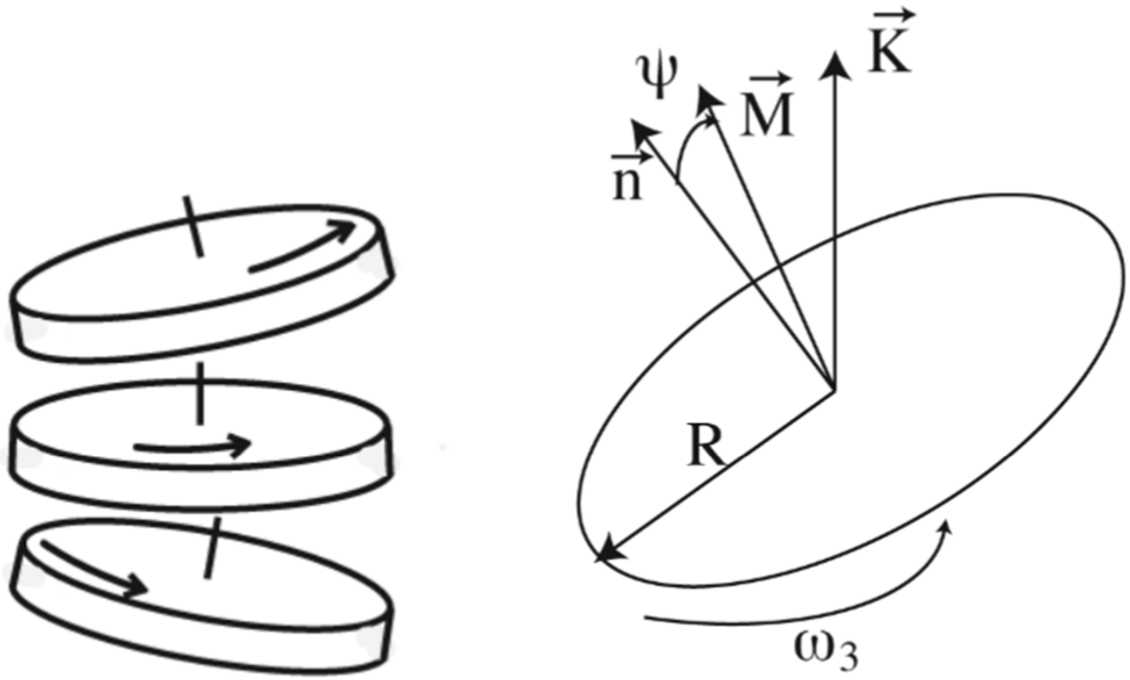


(a) (b)

(c) (d)

Fig. 1

Diaconis P, Holmes S, Montgomery R. Dynamical Bias in the Coin Toss. SIAM Rev. 2007;49: 211–235. doi:10.1137/s0036144504446436



(a) (b)

Fig. 2 (a) *Diagram of a precessing coin.* (b) *Coordinates of precessing coin: $\vec{K}$ is the upward direction, $\vec{n}$ is the normal to the coin, $\vec{M}$ is the angular momentum vector, and ω_3 is the rate of rotation around the normal $\vec{n}$.*

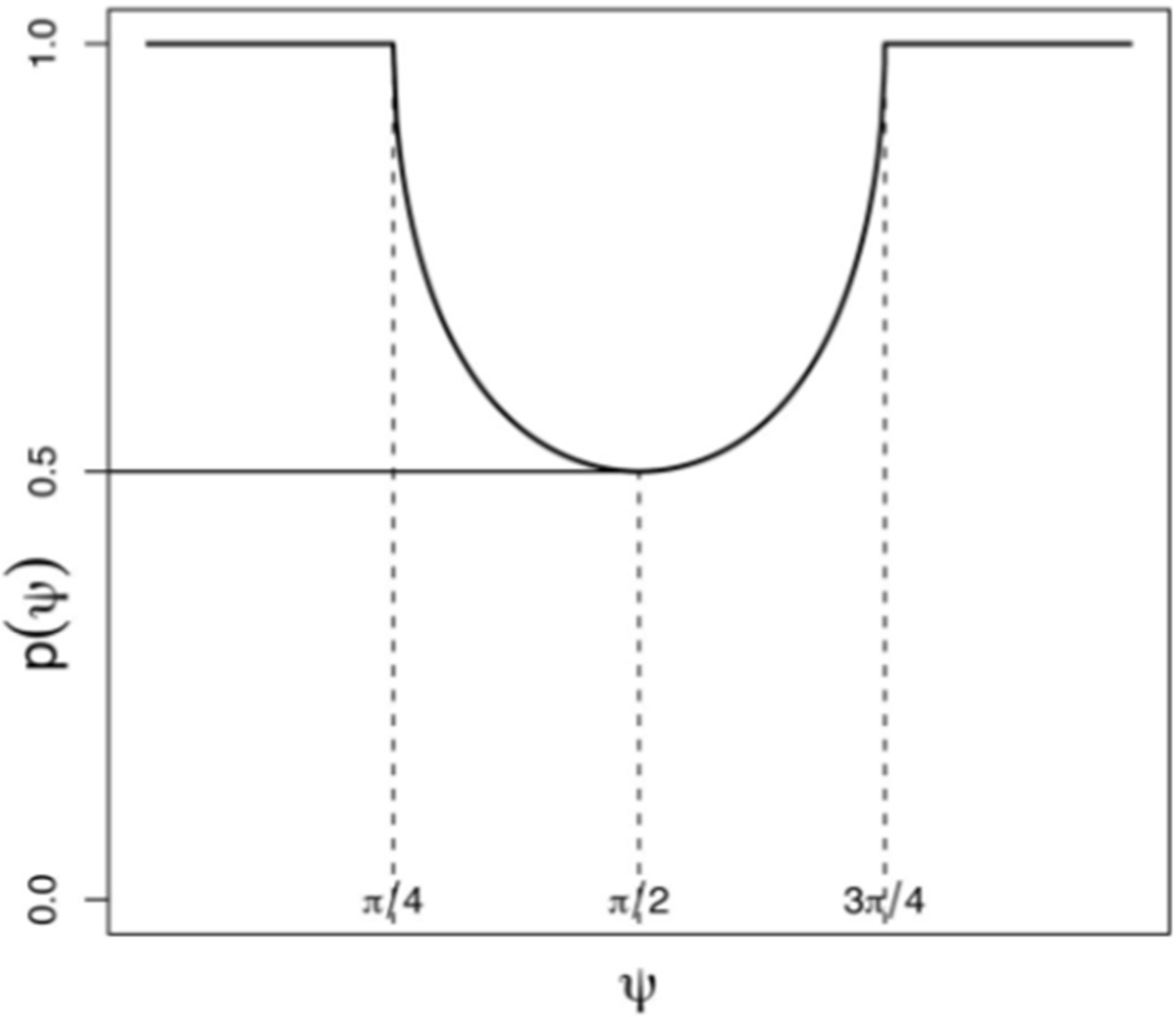
THEOREM 1. *For a coin tossed starting heads up at time 0, let $\tau(t) = \vec{N}(t) \cdot \vec{K}$ be the cosine of the angle between the normal at time t and the up direction $\vec{K}$. Then*

$$\tau(t) = A + B \ \cos(\omega_N t), \tag{1.1}$$

with $A = \cos^2 \psi$, $B = \sin^2 \psi$, $\omega_N = \|\vec{M}\|/I_1$, $I_1 = \frac{1}{4}(mR^2 + \frac{1}{3}mh^2)$ for coins with radius R, thickness h, and mass m. Here ψ is the angle between the angular momentum vector $\vec{M}$ and the normal at time $t = 0$, and $\|\cdot\|$ is the usual Euclidean norm.

Diaconis P, Holmes S, Montgomery R. Dynamical Bias in the Coin Toss. SIAM Rev. 2007;49: 211–235. doi:10.1137/s0036144504446436

PERSI DIACONIS, SUSAN HOLMES, AND RICHARD MONTGOMERY



Diaconis P, Holmes S, Montgomery R. Dynamical Bias in the Coin Toss.
SIAM Rev. 2007;49: 211–235. doi:10.1137/s0036144504446436

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION
2025, VOL. 00, NO. 0, 1–10: Applications and Case Studies
https://doi.org/10.1080/01621459.2025.2516210

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS

Check for updates

Fair Coins Tend to Land on the Same Side They Started: Evidence from 350,757 Flips

František Bartoš[a], Alexandra Sarafoglou[a], Henrik R. Godmann[a], Amir Sahrani[a], David Klein Leunk[a], Pierre Y. Gui[b], David Voss[b], Kaleem Ullah[b], Malte Zoubek[c,*], Franziska Nippold[b], Frederik Aust[a], Felipe Fontana Vieira[d,**], Chris-Gabriel Islam[e,f], Anton J. Zoubek[g], Sara Shabani[h], Jonas Petter[a], Ingeborg B. Roos[i], Adam Finnemann[a,j], Aaron B. Lob[b,k], Madlen F. Hoffstadt[a], Jason Nak[b], Jill de Ron[b], Koen Derks[l], Karoline Huth[a,m], Sjoerd Terpstra[n], Thomas Bastelica[o,p], Magda Matetovici[b,q], Vincent L. Ott[b], Andreea S. Zetea[b], Katharina Karnbach[b], Michelle C. Donzallaz[a], Arne John[a], Roy M. Moore[b], Franziska Assion[r], Riet van Bork[a], Theresa E. Leidinger[b], Xiaochang Zhao[b], Adrian Karami Motaghi[b], Ting Pan[b,i], Hannah Armstrong[b], Tianqi Peng[b,†], Mara Bialas[b,s], Joyce Y.-C. Pang[b], Bohan Fu[b], Shujun Yang[b], Xiaoyi Lin[b], Dana Sleiffer[b], Miklos Bognar[t], Balazs Aczel[u], and Eric-Jan Wagenmakers[a]

[a]Department of Psychological Methods, University of Amsterdam, Amsterdam, Netherlands; [b]Department of Psychology, University of Amsterdam, Amsterdam, Netherlands; [c]Institute of Economic Law, University of Kassel, Kassel, Germany; [d]Faculty of Psychology and Educational Sciences, KU Leuven, Leuven, Belgium; [e]Chairs of Statistics and Econometrics, Georg August University of Göttingen, Göttingen, Germany; [f]Centre for Environmental Sciences, Hasselt University, Hasselt, Belgium; [g]Gutenberg School Wiesbaden, Wiesbaden, Germany; [h]Department of Psychology, Justus Liebig University, Giessen, Germany; [i]Department of Adult Psychiatry, Amsterdam University Medical Centre, Amsterdam, Netherlands; [j]Centre for Urban Mental Health, University of Amsterdam, Amsterdam, Netherlands; [k]Cognitive and Behavioral Decision Research, University of Zurich, Zurich, Switzerland; [l]Center of Accounting, Auditing & Control, Nyenrode Business University, Breukelen, Netherlands; [m]Department of Psychiatry, Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands; [n]Institute for Marine and Atmospheric Research Utrecht, Utrecht University, Utrecht, Netherlands; [o]Centre Hospitalier Le Vinatier, Bron, France; [p]INSERM, U1028; CNRS, UMR5292, Lyon Neuroscience Research Center, Psychiatric Disorders: from Resistance to Response Team, Lyon, France; [q]Research Institute of Child Development and Education, University of Amsterdam, Amsterdam, Netherlands; [r]Faculty of Behavioral and Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam, Netherlands; [s]Behavioural Science Institute, Radboud University, Nijmegen, Netherlands; [t]Doctoral School of Psychology, ELTE Eotvos Lorand University, Budapest, Hungary; [u]Institute of Psychology, ELTE Eotvos Lorand University, Budapest, Hungary

[redacted]

How Bartos et al got 350,000 flips?

- Sequence of 40,000 coin flips collected by Janet Larwood and Priscilla Ku (Berkeley 2009)
  - Larwood always started the flips heads-up, and Ku always tails-up
- A group of five bachelor students collected at least 15,000 coin flips each as a part of their bachelor thesis project, contributing 75,036 coin flips in total
- A series of on-site "coin flipping marathons" where 35 people spent up to 12 hr coin-flipping contributing a total of 203,440 coin flips. (see e.g. https://www.youtube.com/watch?v=3xNg51mv-fk)
- They issued a call for collaboration via Twitter, which resulted in an additional seven people contributing a total of 72,281 coin flips. Verified via video recordings.

Conclusions of Bartos et al (2025)

- Probability (coin lands on the same side) = 0.508
- 95% credible interval (CI) [0.506, 0.509]
- Odds to get this with unbiased coin <10^{-17}

$$\mathrm{BF}_{10} = \frac{p(\text{data} \mid \mathcal{H}_1)}{p(\text{data} \mid \mathcal{H}_0)},$$

which contrasts the competing hypotheses in terms of their predictive performance for the observed data. The Bayes factor hypothesis test indicates extreme evidence in favor of the same-side bias predicted by the DHM model, $\mathrm{BF}_{\text{same-side bias}} = 1.76 \times 10^{17}$.

Information from Bartoš F, Sarafoglou A, Godmann HR, Sahrani A, Leunk DK, Gui PY, et al. J Am Stat Assoc. 2025;ahead-of-print: 1–10. doi:10.1080/01621459.2025.2516210

Who is ready to use Matlab?

A. I have Matlab installed on my laptop
B. I am ready to use Matlab on EWS
C. I don't have it ready but plan to install it
D. I am not ready but plan to use EWS
E. I plan to use other software (Python, R, etc.)

Get your i-clickers

Who is familiar with Matlab?

A. I used Matlab extensively
B. I used Matlab from time to time
C. I used Matlab a couple of times
D. I never used Matlab but want to learn it
E. I never used Matlab but don't want to learn it, Instead of Matlab I will use something else: Python, R, etc.

Get your i-clickers

Matlab is easy to learn

- Matlab is the lingua franca of all of **engineering**
- Use online tutorials e.g.: https://www.youtube.com/watch?v=82TGgQApFIQ
- Matlab is designed to work with Matrices → symbols * and / are understood as matrix multiplication and division
- Use .* and ./ for regular (non-matrix) multiplication
- Add ; in the end of the line to avoid displaying the output on the screen
- Loops: for i=1:100; f(i)=floor(2.*rand); end;
- Conditional statements: if rand>0.5; count=count+1; end;
- Plotting: plot(x,y,'ko-'); or semilogx(x,y,'ko-'); or loglog(x,y,'ko-'); .
  To keep adding plots onto the same axes use: hold on;
  To create a new axes use figure;
- Generating matrices: rand(100) – generates square matrix 100x100. **Confusing!** Use rand(100,1) or zeros(30,20), or randn(1,40) (Gaussian);
- If Matlab complains multiplying matrices check sizes using whos and if needed use transpose operation: x=x';

A Matlab Cheat-sheet (MIT 18.06, Fall 2007)

Basics:

`save 'file.mat'`	save variables to *file.mat*
`load 'file.mat'`	load variables from *file.mat*
`diary on`	record input/output to file *diary*
`diary off`	stop recording
`whos`	list all variables currenly defined
`clear`	delete/undefine all variables
`help command`	quick help on a given *command*
`doc command`	extensive help on a given *command*

Defining/changing variables:

`x = 3`	define variable x to be 3
`x = [1 2 3]`	set x to the 1×3 row-vector (1,2,3)
`x = [1 2 3];`	same, but don't echo x to output
`x = [1;2;3]`	set x to the 3×1 column-vector (1,2,3)
`A = [1 2 3 4;5 6 7 8;9 10 11 12];`	set A to the 3×4 matrix with rows 1,2,3,4 etc.
`x(2) = 7`	change x from (1,2,3) to (1,7,3)
`A(2,1) = 0`	change $A_{2,1}$ from 5 to 0

Arithmetic and functions of numbers:

`3*4, 7+4, 2-6 8/3`	multiply, add, subtract, and divide numbers
`3^7, 3^(8+2i)`	compute 3 to the 7th power, or 3 to the 8+2i power
`sqrt(-5)`	compute the square root of –5
`exp(12)`	compute e^{12}
`log(3), log10(100)`	compute the natural log (ln) and base-10 log ($\log_{10}$)
`abs(-5)`	compute the absolute value |–5|
`sin(5*pi/3)`	compute the sine of 5π/3
`besselj(2,6)`	compute the Bessel function $J_2(6)$

Arithmetic and functions of vectors and matrices:

`x * 3`	multiply every element of x by 3
`x + 2`	add 2 to every element of x
`x + y`	element-wise addition of two vectors x and y
`A * y`	product of a matrix A and a vector y
`A * B`	product of two matrices A and B
`x * y`	not allowed if x and y are two column vectors!
`x .* y`	element-wise product of vectors x and y
`A^3`	the square matrix A to the 3rd power
`x^3`	not allowed if x is not a square matrix!
`x.^3`	every element of x is taken to the 3rd power
`cos(x)`	the cosine of every element of x
`abs(A)`	the absolute value of every element of A
`exp(A)`	e to the power of every element of A
`sqrt(A)`	the square root of every element of A
`expm(A)`	the matrix exponential e^A
`sqrtm(A)`	the matrix whose square is A

Constructing a few simple matrices:

`rand(12,4)`	a 12×4 matrix with uniform random numbers in [0,1)
`randn(12,4)`	a 12×4 matrix with Gaussian random (center 0, variance 1)
`zeros(12,4)`	a 12×4 matrix of zeros
`ones(12,4)`	a 12×4 matrix of ones
`eye(5)`	a 5×5 identity matrix I ("eye")
`eye(12,4)`	a 12×4 matrix whose first 4 rows are the 4×4 identity
`linspace(1.2,4.7,100)`	row vector of 100 equally-spaced numbers from 1.2 to 4.7
`7:15`	row vector of 7,8,9,…,14,15
`diag(x)`	matrix whose diagonal is the entries of x (and other elements = 0)

Portions of matrices and vectors:

`x(2:12)`	the 2nd to the 12th elements of x
`x(2:end)`	the 2nd to the last elements of x
`x(1:3:end)`	every third element of x, from 1st to the last
`x(:)`	all the elements of x
`A(5,:)`	the row vector of every element in the 5th row of A
`A(5,1:3)`	the row vector of the first 3 elements in the 5th row of A
`A(:,2)`	the column vector of every element in the 2nd column of A
`diag(A)`	column vector of the diagonal elements of A

Solving linear equations:

`A \ b`	for A a matrix and b a column vector, the solution x to $Ax=b$
`inv(A)`	the inverse matrix A^{-1}
`[L,U,P] = lu(A)`	the LU factorization $PA=LU$
`eig(A)`	the eigenvalues of A
`[V,D] = eig(A)`	the columns of V are the eigenvectors of A, and the diagonals `diag(D)` are the eigenvalues of A

Plotting:

`plot(y)`	plot y as the y axis, with 1,2,3,… as the x axis
`plot(x,y)`	plot y versus x (must have same length)
`plot(x,A)`	plot columns of A versus x (must have same # rows)
`loglog(x,y)`	plot y versus x on a log-log scale
`semilogx(x,y)`	plot y versus x with x on a log scale
`semilogy(x,y)`	plot y versus x with y on a log scale
`fplot(@(x) …expression…,[a,b])`	plot some expression in x from $x=a$ to $x=b$
`axis equal`	force the x and y axes of the current plot to be scaled equally
`title('A Title')`	add a title *A Title* at the top of the plot
`xlabel('blah')`	label the x axis as *blah*
`ylabel('blah')`	label the y axis as *blah*
`legend('foo','bar')`	label 2 curves in the plot *foo* and *bar*
`grid`	include a grid in the plot
`figure`	open up a new figure window

Transposes and dot products:

`x.', A.'`	the transposes of x and A
`x', A'`	the complex-conjugate of the transposes of x and A
`x' * y`	the dot (inner) product of two *column* vectors x and y
`dot(x,y), sum(x.*y)`	…two other ways to write the dot product
`x * y'`	the *outer* product of two *column* vectors x and y

http://web.mit.edu/18.06/www/Spring09/matlab-cheatsheet.pdf

VIA app by Kramer needs to be updated

- Get the latest version app from https://k.kramerav.com/support/download.asp?f=61213

- On 8/24/2023 the version that worked was 4.0.3.1344

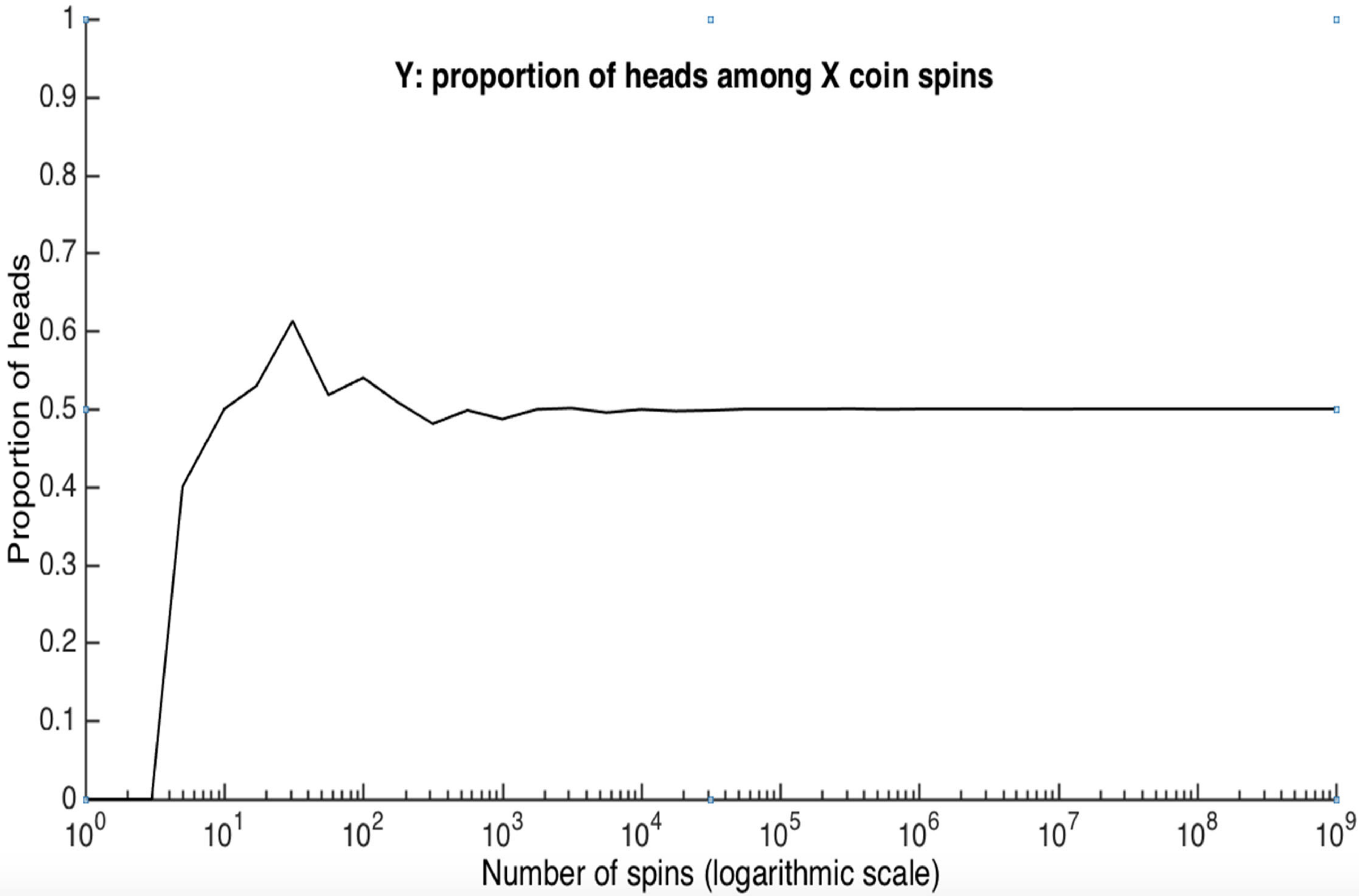
Matlab group exercise

Each table to edit the file coin_toss_template.m (replace all ?? with commands/variables/operations) or writes a new Matlab (Python, R, or anything else) script to:

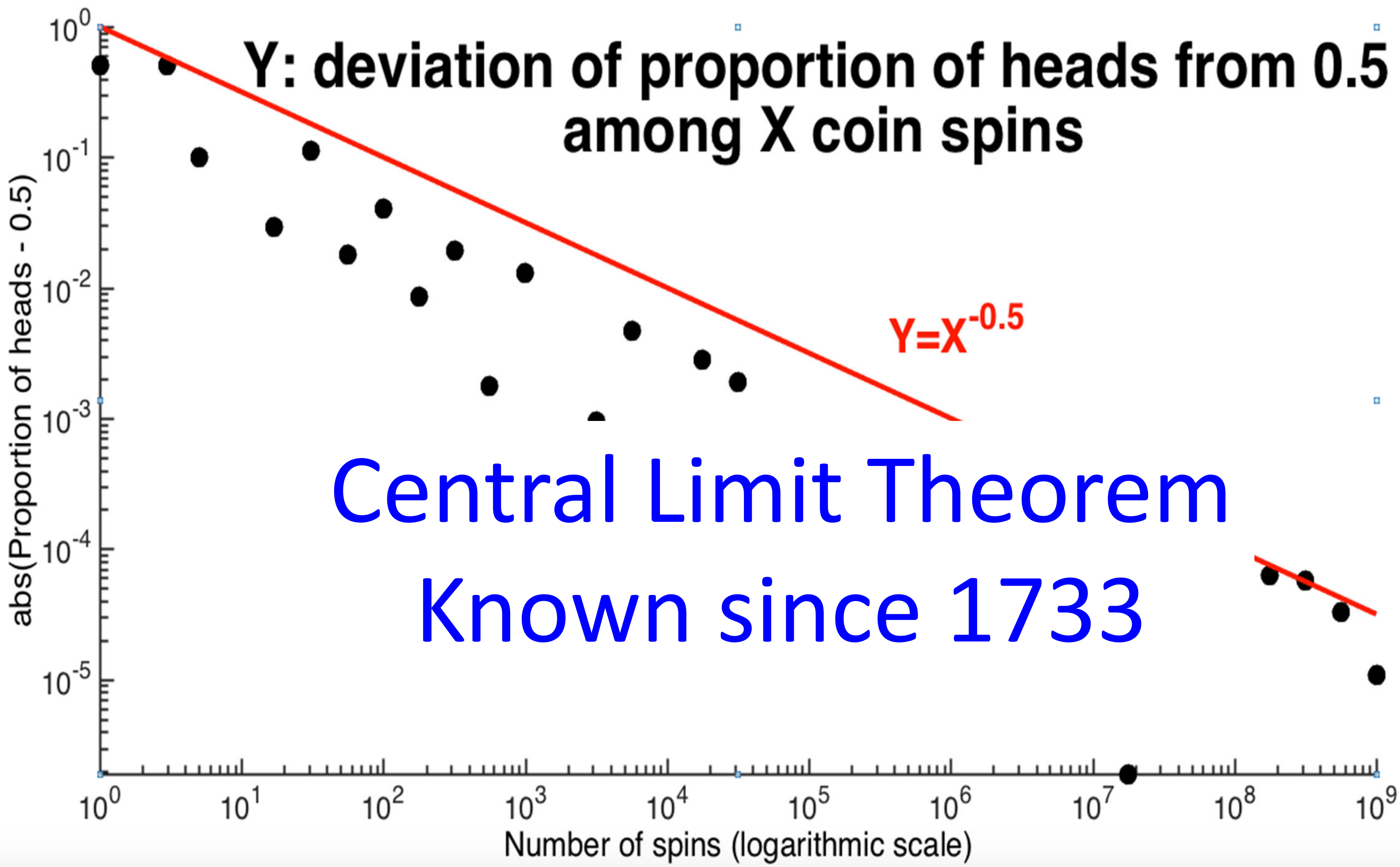
- Simulate a fair coin toss experiment
- Generate multiple tosses of a fair coin:
  1 – heads, 0 - tails
- Calculate the fraction of heads (f_heads(t)) at timepoints:
  t=10; 100; 1000; 10,000; 100,000; 1,000,000;10,000,000 coin tosses
- Plot fraction of heads f_heads(t) vs t with a logarithmic t-axis
- Plot abs(f_heads(t)-0.5) vs t on a log-log plot
  (both axes are logarithmic)

How I did it

- Stats=1e7;
- r0=rand(Stats,1); r1=floor(2.*r0);
- n_heads(1)=r1(1);
- for t=2:Stats; n_heads(t)=n_heads(t-1)+r1(t); end;
- tp=[1, 10,100,1000, 10000, 100000, 1000000, 10000000]
- np=n_heads(tp); fp=np./tp
- figure; semilogx(tp,fp,'ko-');
- hold on; semilogx([1,10000000],[0.5,0.5],'r--');
- figure; loglog(tp,abs(fp-0.5),'ko-');
- hold on; loglog(tp,0.5./sqrt(tp),'r--');



Proportion of heads among 1,000,000,000 coin tosses
(10^5 more than Kerrich) took me 33 seconds on my Surface Book



ABS(Proportion of heads-0.5)
among 100,000,000 coin tosses