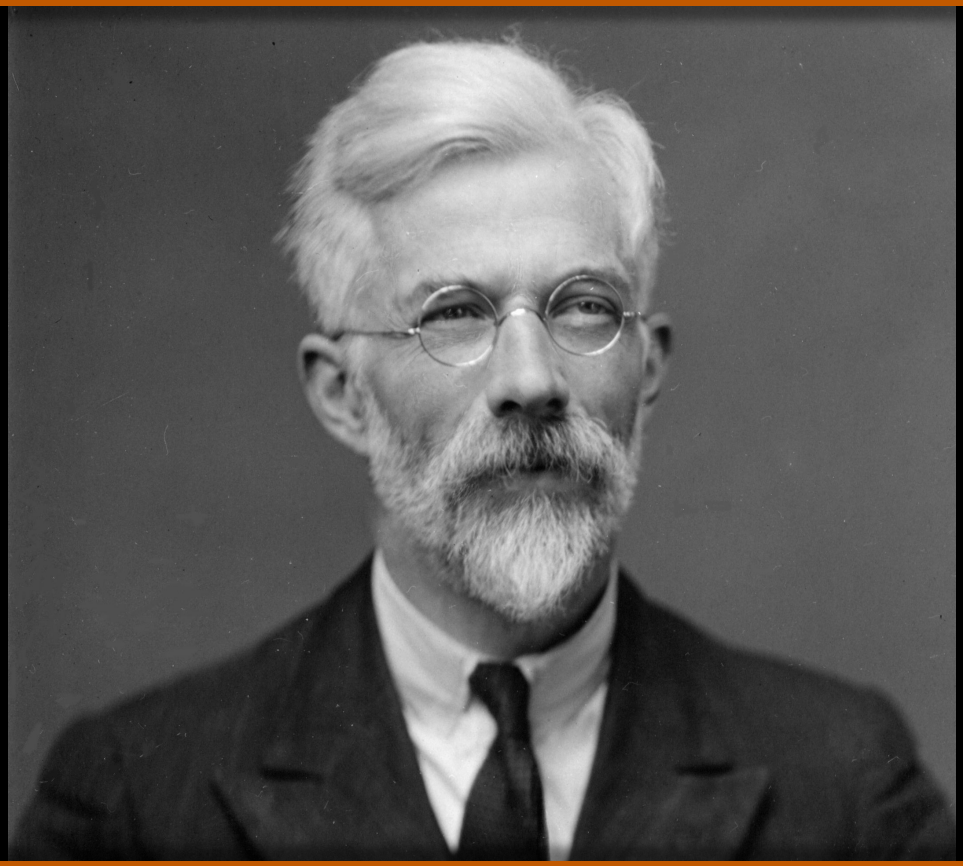
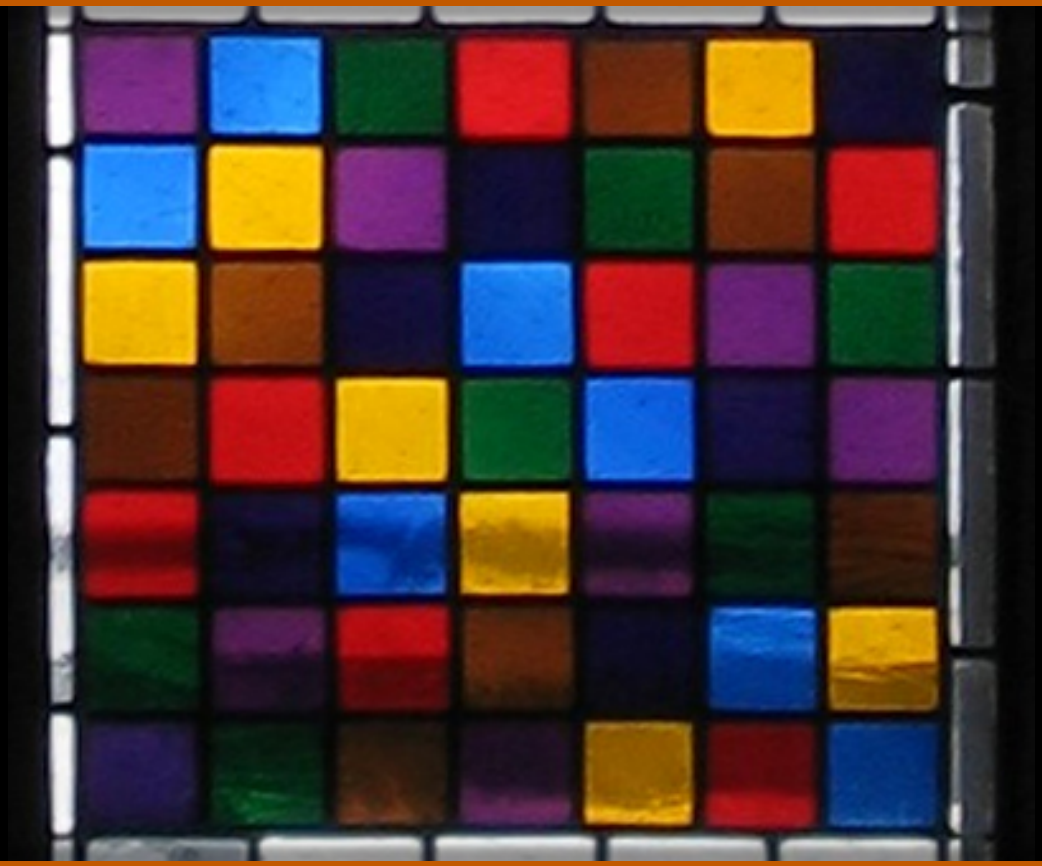
SOEE1475 Statistics and Data Analysis
Lecture 4: Null hypothesis statistical testing

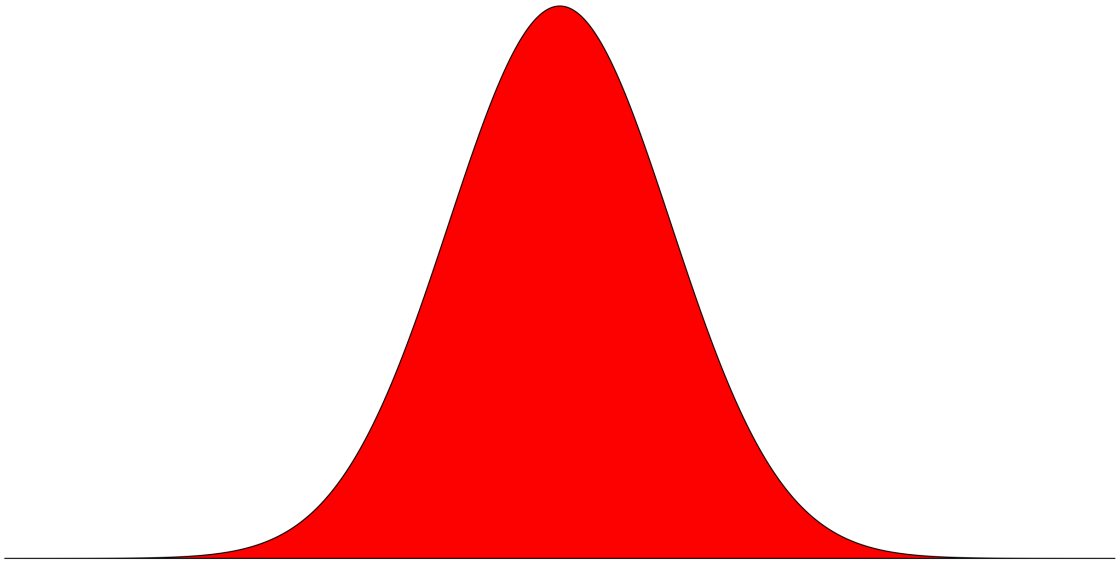




Graeme T. Lloyd

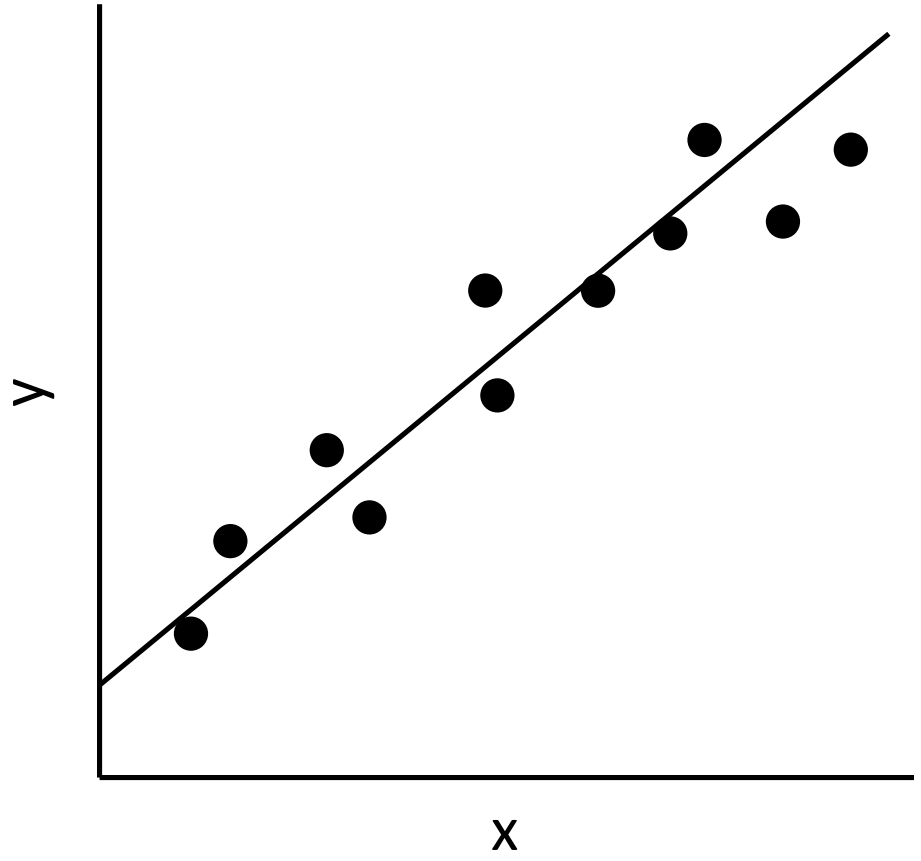
Univariate data

Population



Sample

Bivariate data



Future

- Multivariate data
- Temporal, spatial, and directional data

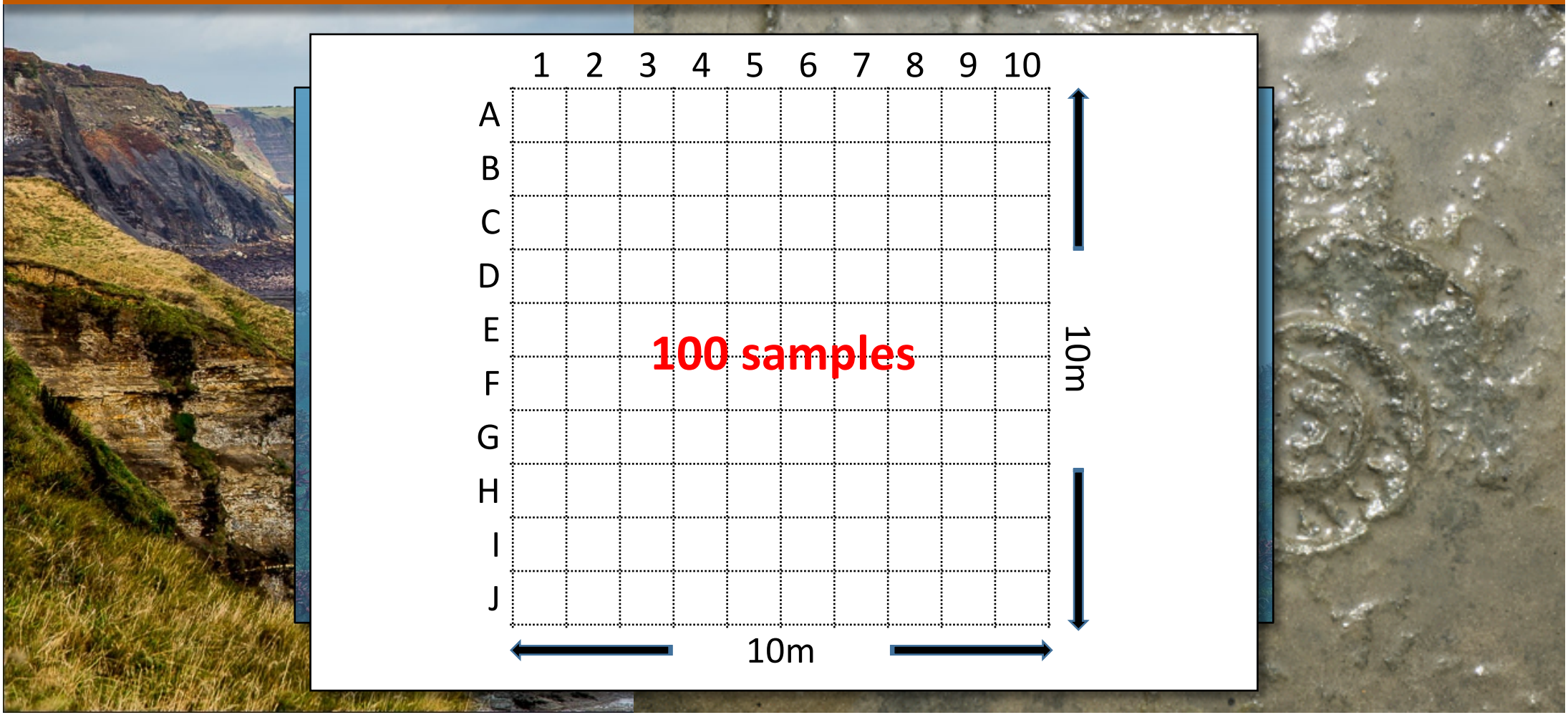
Today

- Null hypotheses
- Confidence and significance levels
- Type I and Type II errors

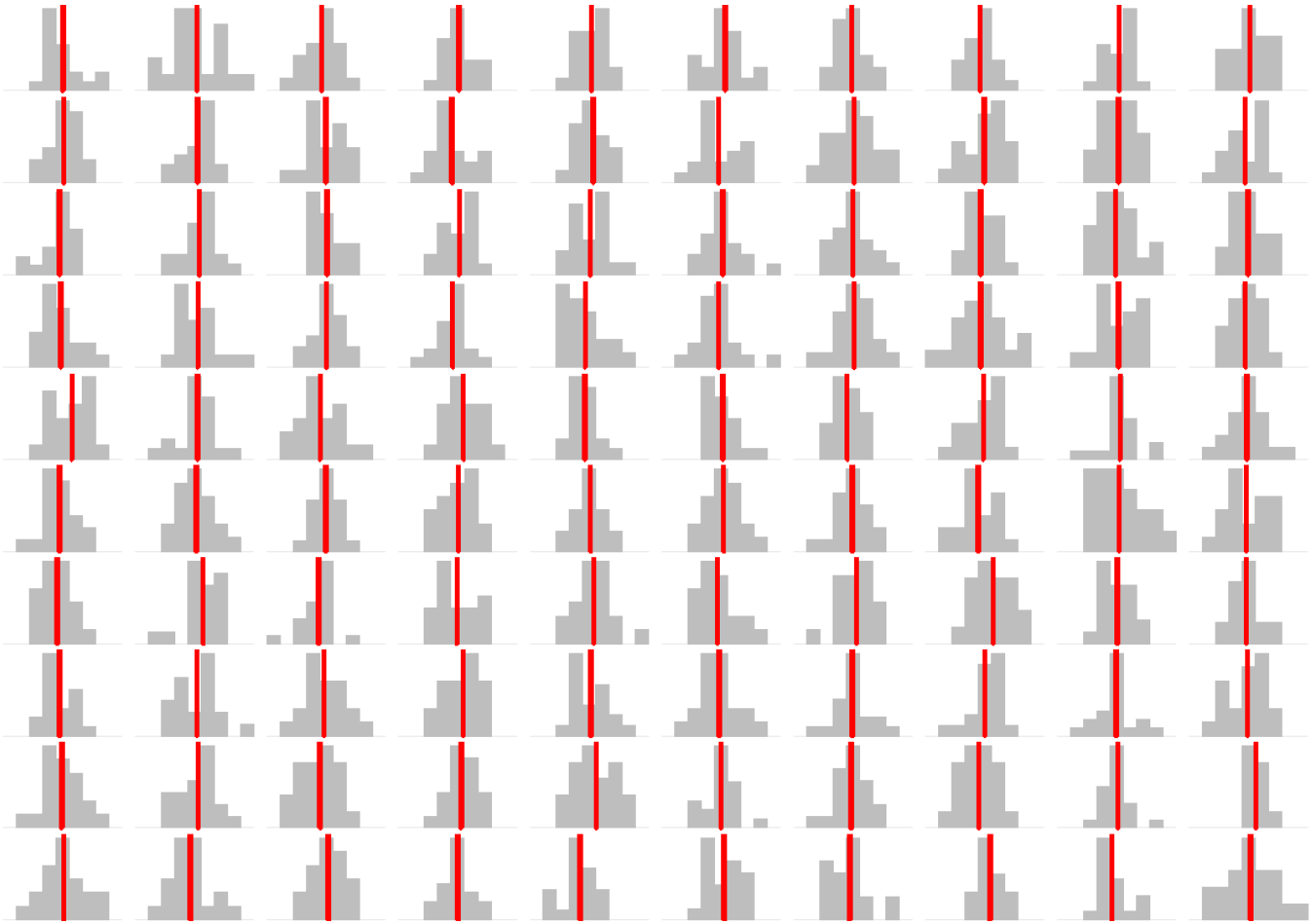
Populations and samples

	Population	Sample
Mean	μ	$\overline{x}$
Standard deviation	σ	s

Populations and samples

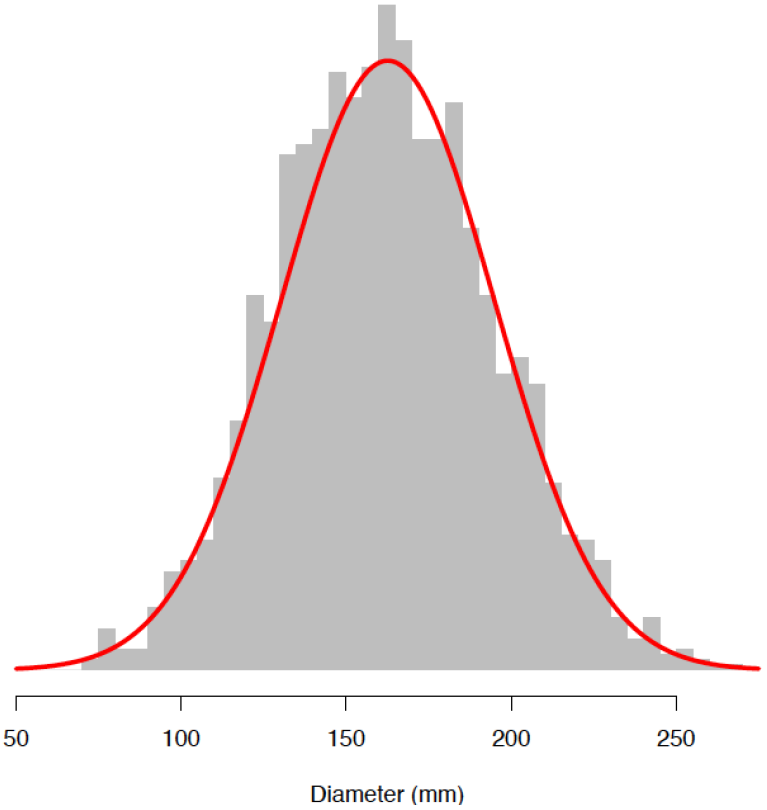


Populations and samples

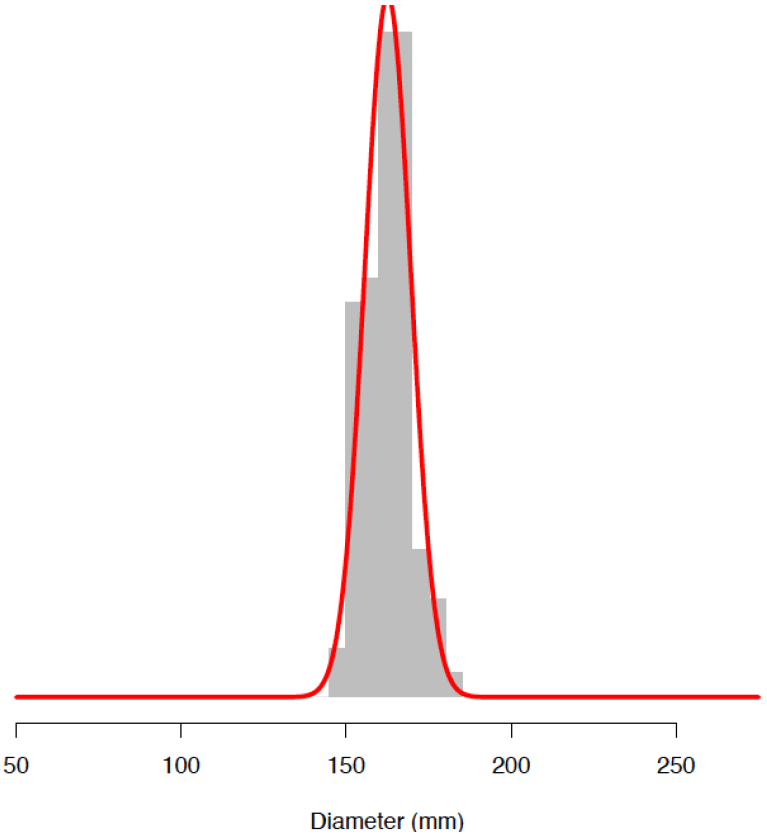


Populations and samples

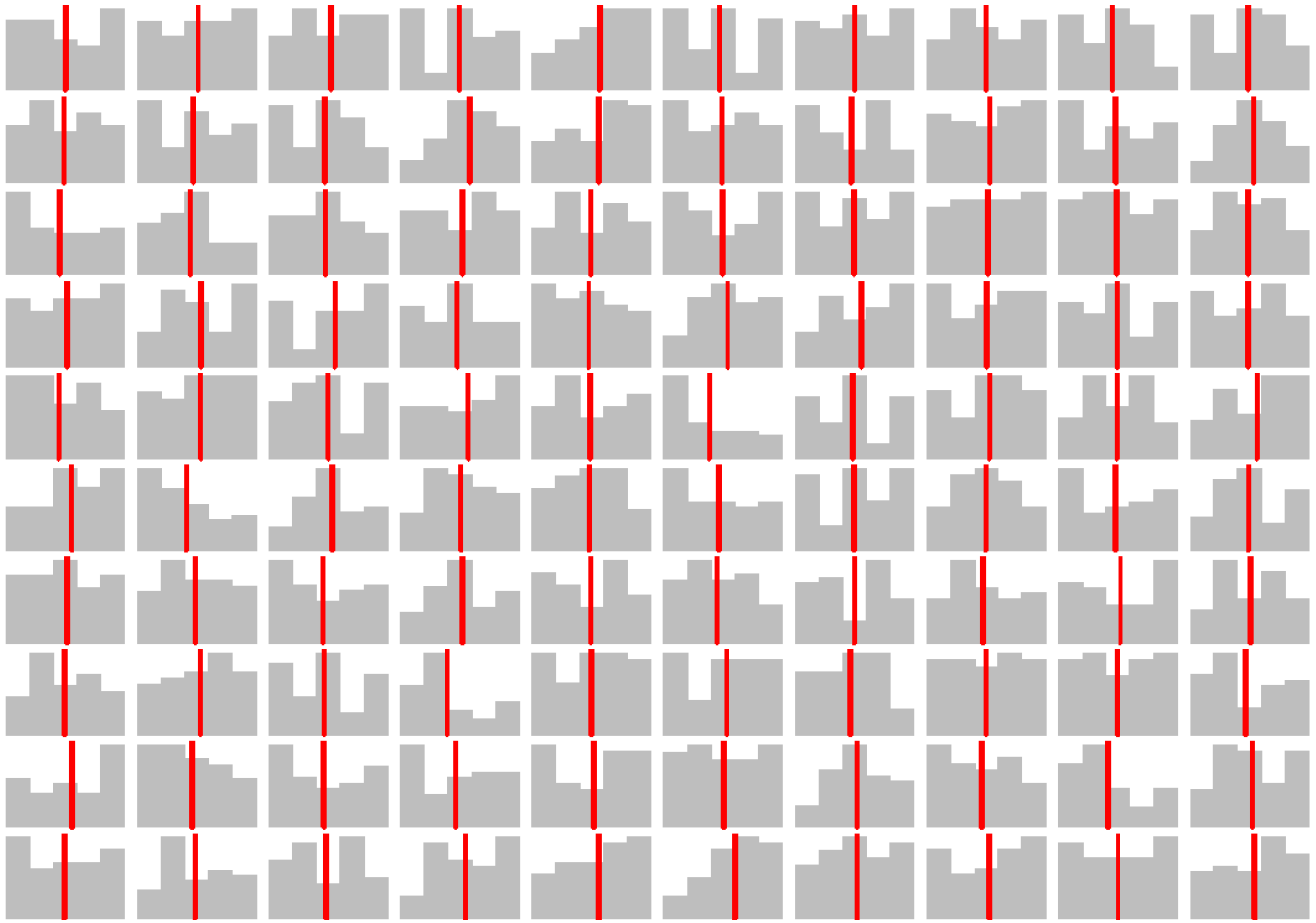
Pooled sample



Sample means

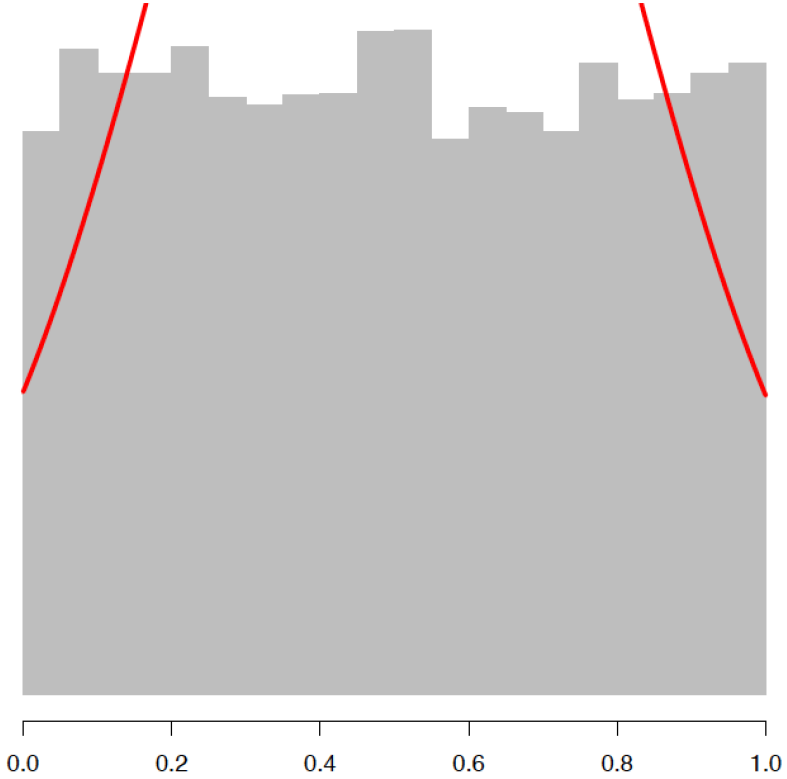


Populations and samples

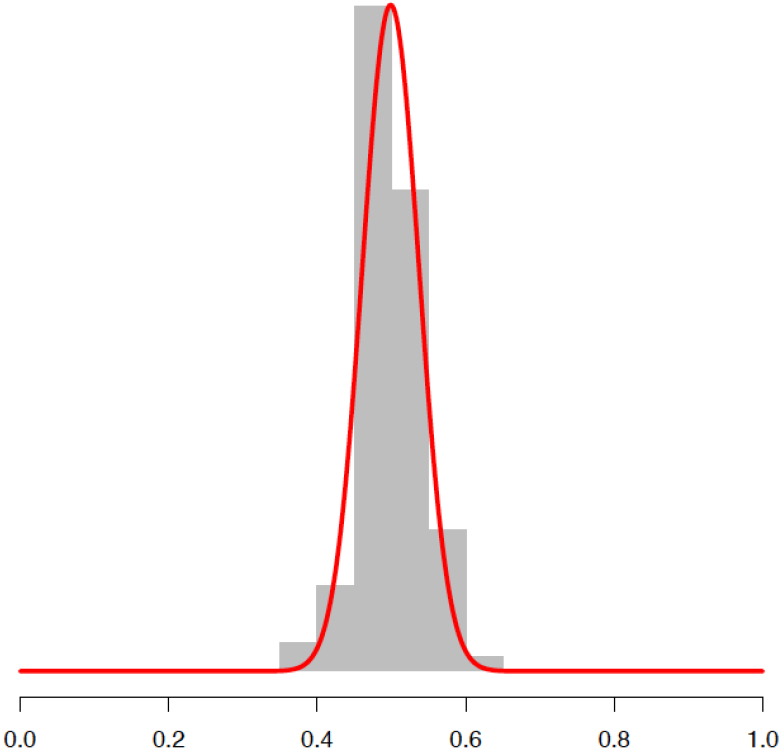


Populations and samples

Pooled sample



Sample means



Populations and samples – implications

Sample means are normally distributed

Holds regardless of the distribution of samples; Central Limit Theorem

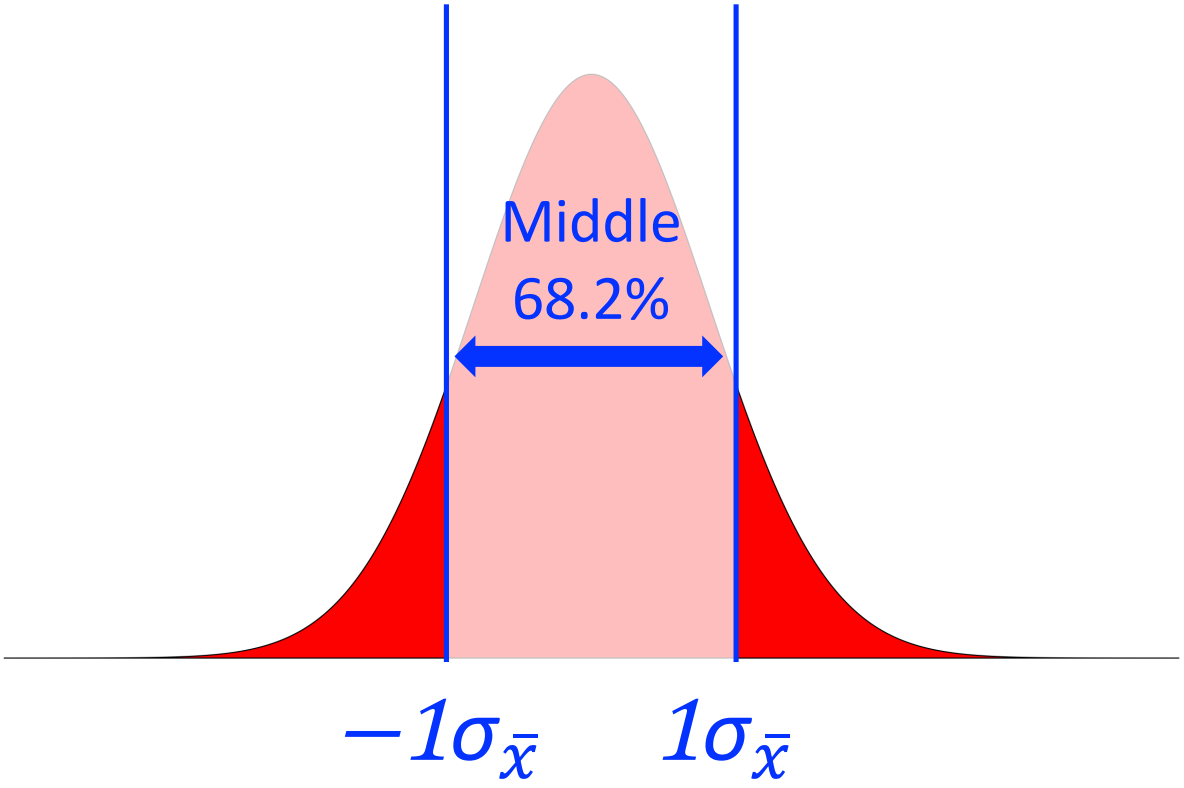
Mean of sample means ≈ population mean

If we have ∞ samples, mean of sample means = population mean

Distribution of sample means has its' own standard deviation

A dispersion metric, standard error of the mean (SEM)

Standard error of the mean, $\sigma_{\bar{x}}$



$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

$$\sigma_{\bar{x}} \approx \frac{s}{\sqrt{n}}$$

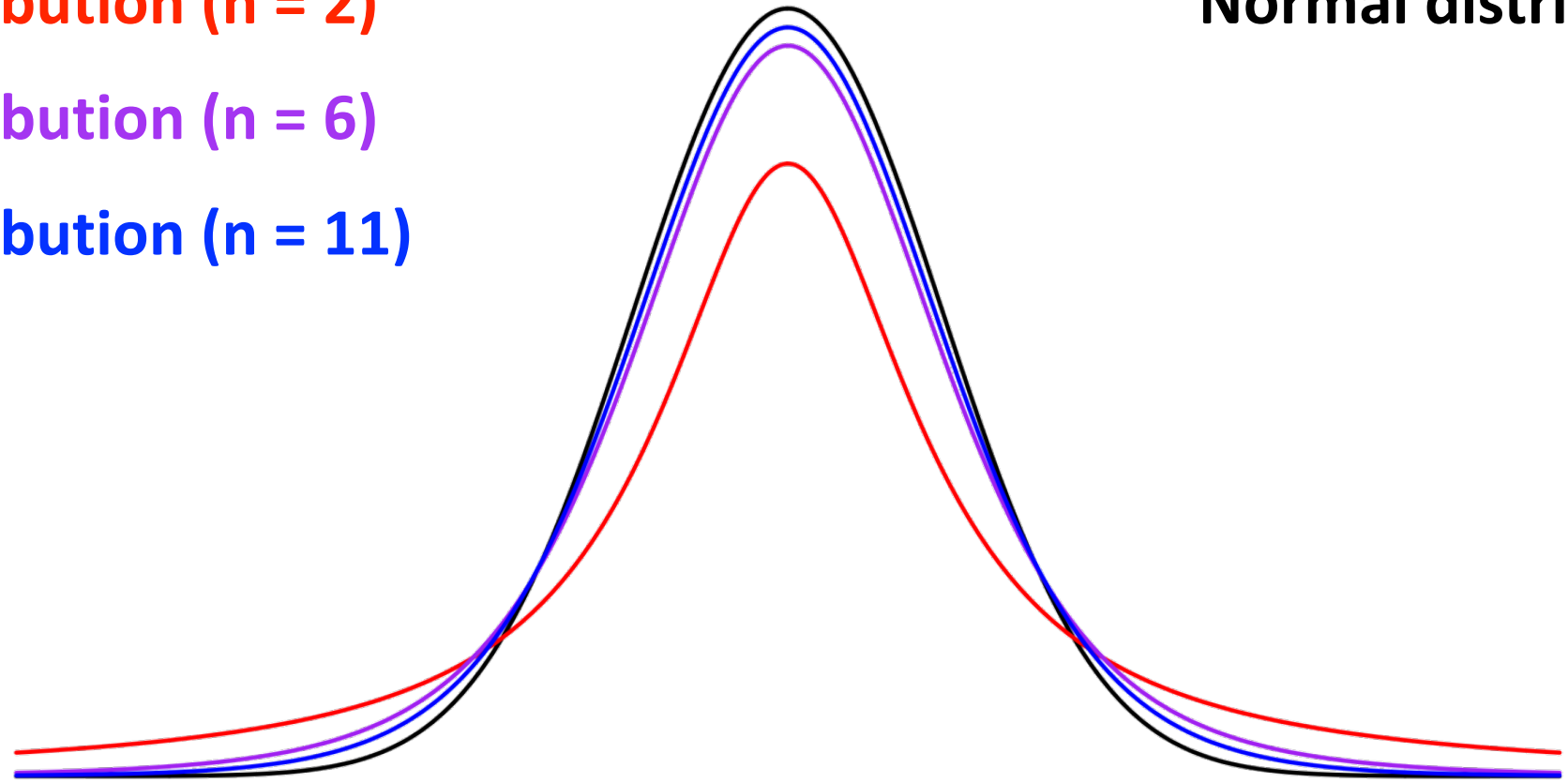
Standard error of the mean, $\sigma_{\bar{x}}$

***t*-distribution (n = 2)**

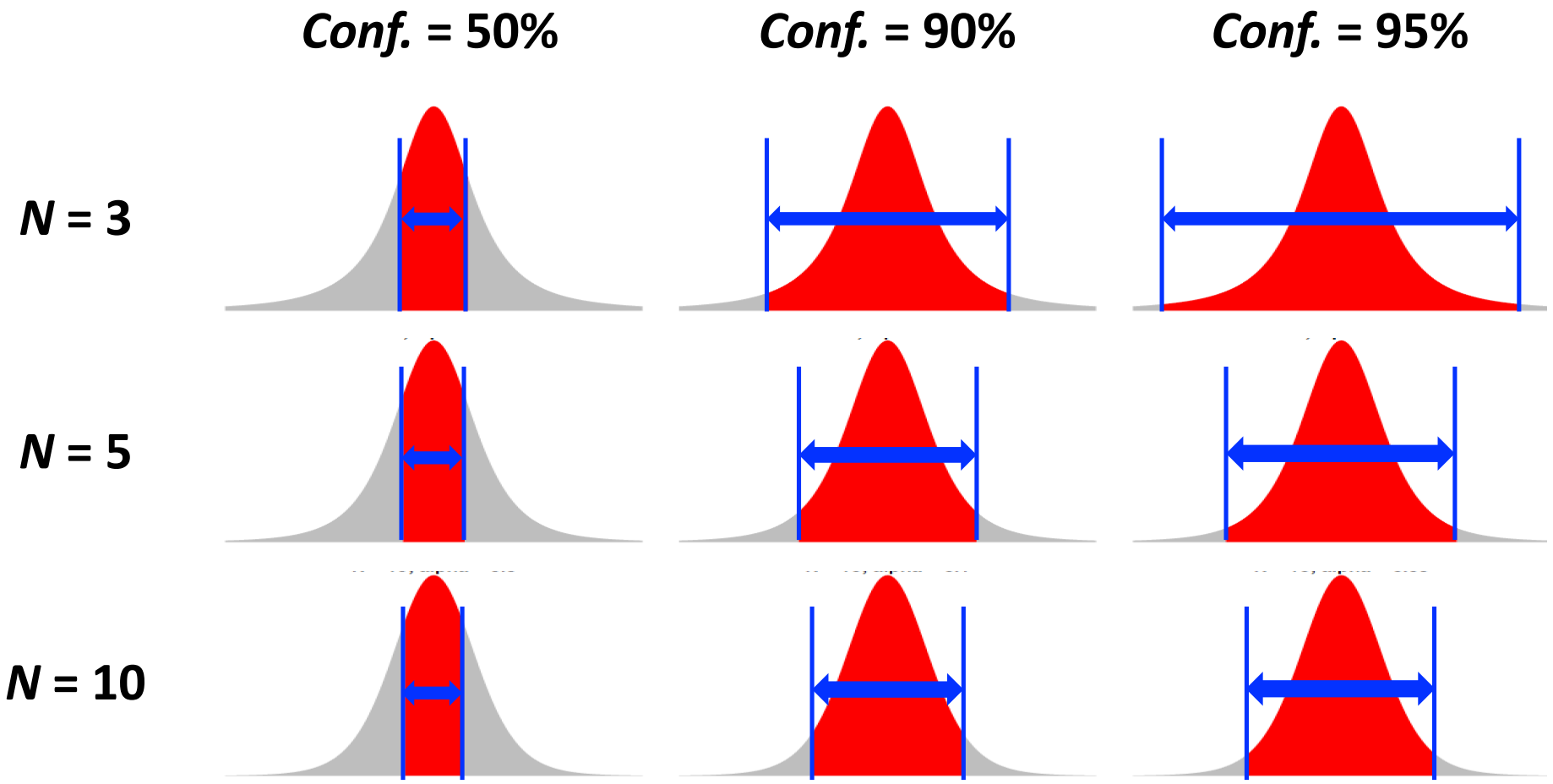
***t*-distribution (n = 6)**

***t*-distribution (n = 11)**

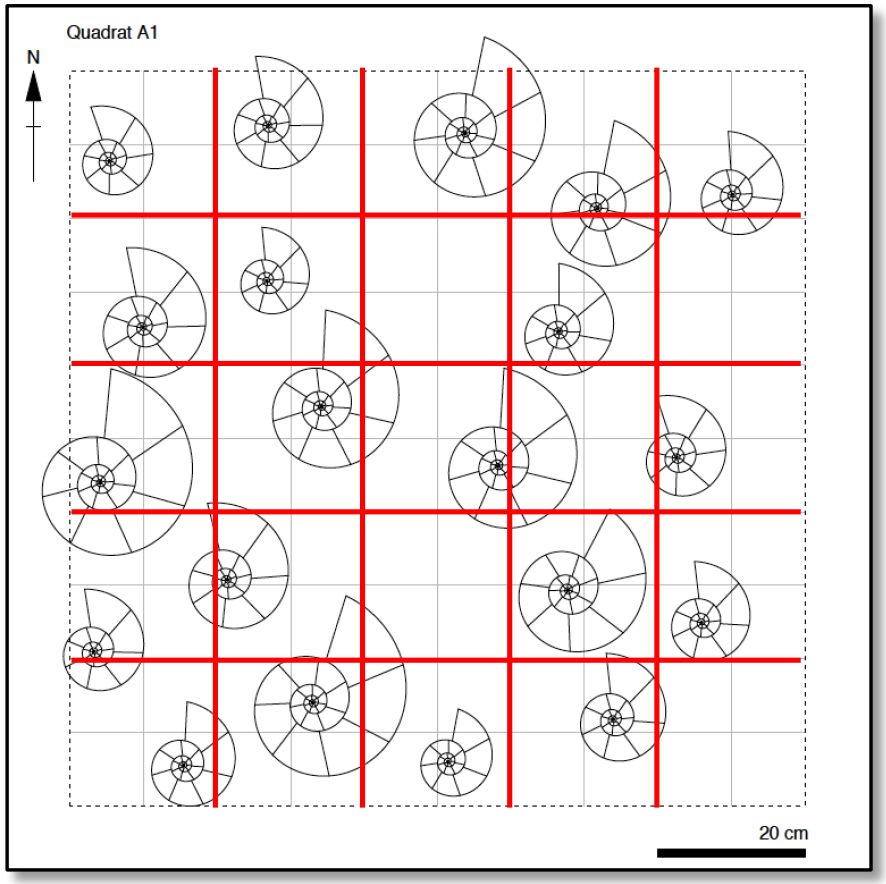
Normal distribution

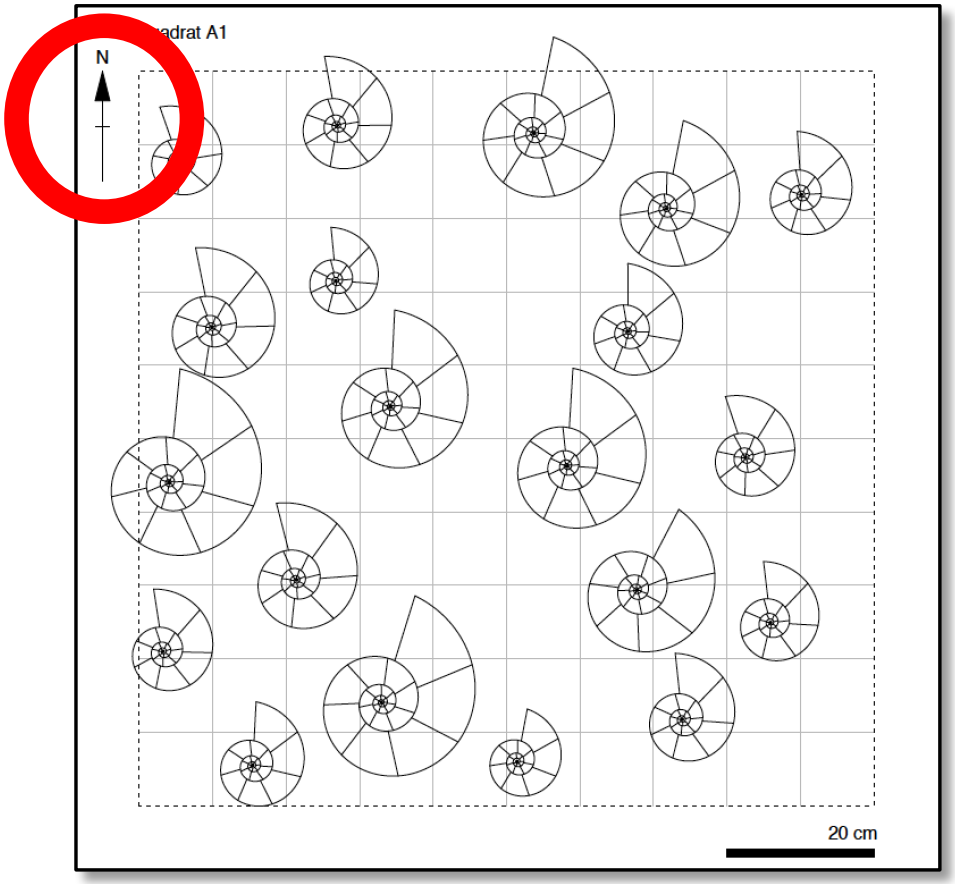


Confidence intervals and $\sigma_{\bar{x}}$



Exercise





Exercise

p(at least two people in the room share a birthday)

Exercise

p(2+ shared birthdays) =

Option A	0.0-0.2
Option B	0.2-0.4
Option C	0.4-0.6
Option D	0.6-0.8
Option E	0.8-1.0

A blueprint for science

Pick interesting question

Framed in form of a hypothesis

Collect data

Measure at appropriate precision; avoid biased samples

Perform statistics

Choose appropriate test; accept or reject hypothesis

Null and alternative hypotheses

Null hypothesis – H_{null} / H_0

Simplest, boring, often stochastic explanation

Alternative hypothesis – H_{alt} / H_1

Not the null, interesting, may incorporate multiple outcomes

Null and alternative hypotheses

	H_{null} / H_0	H_{alt} / H_1
One-sample t-test	$CI_{low} < \mu_{hyp.} < CI_{upp}$	$\mu_{hyp.} > CI_{upp}$ or $\mu_{hyp.} < CI_{low}$
Two-sample t-test	$\mu_A = \mu_B$	$\mu_A < \mu_B$ or $\mu_A > \mu_B$
ANOVA	$\mu_A = \mu_B = \mu_C ...$	$\mu_A > \mu_B = \mu_C$ or...
Correlation	$r = 0$	$r > 0$ or $r < 0$

Choosing a hypothesis

Significance-level

Decides threshold value between accepting or rejecting null; termed alpha

Confidence-level

One minus alpha; often expressed as a percentage

Choosing a hypothesis

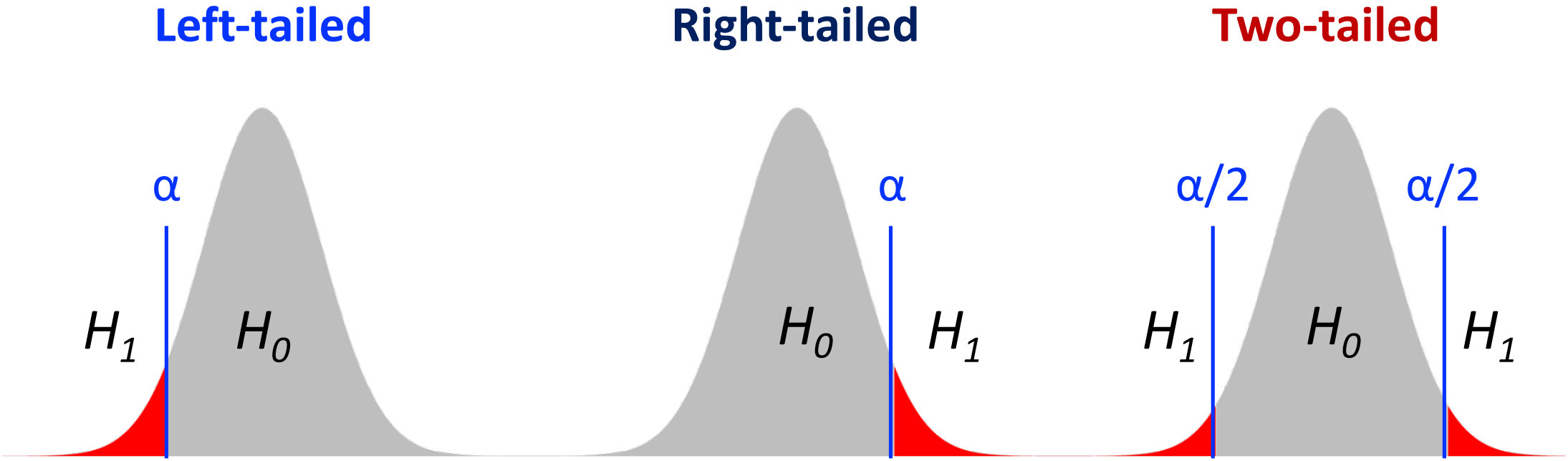
One-tailed

Only care whether value falls to left or right of alpha

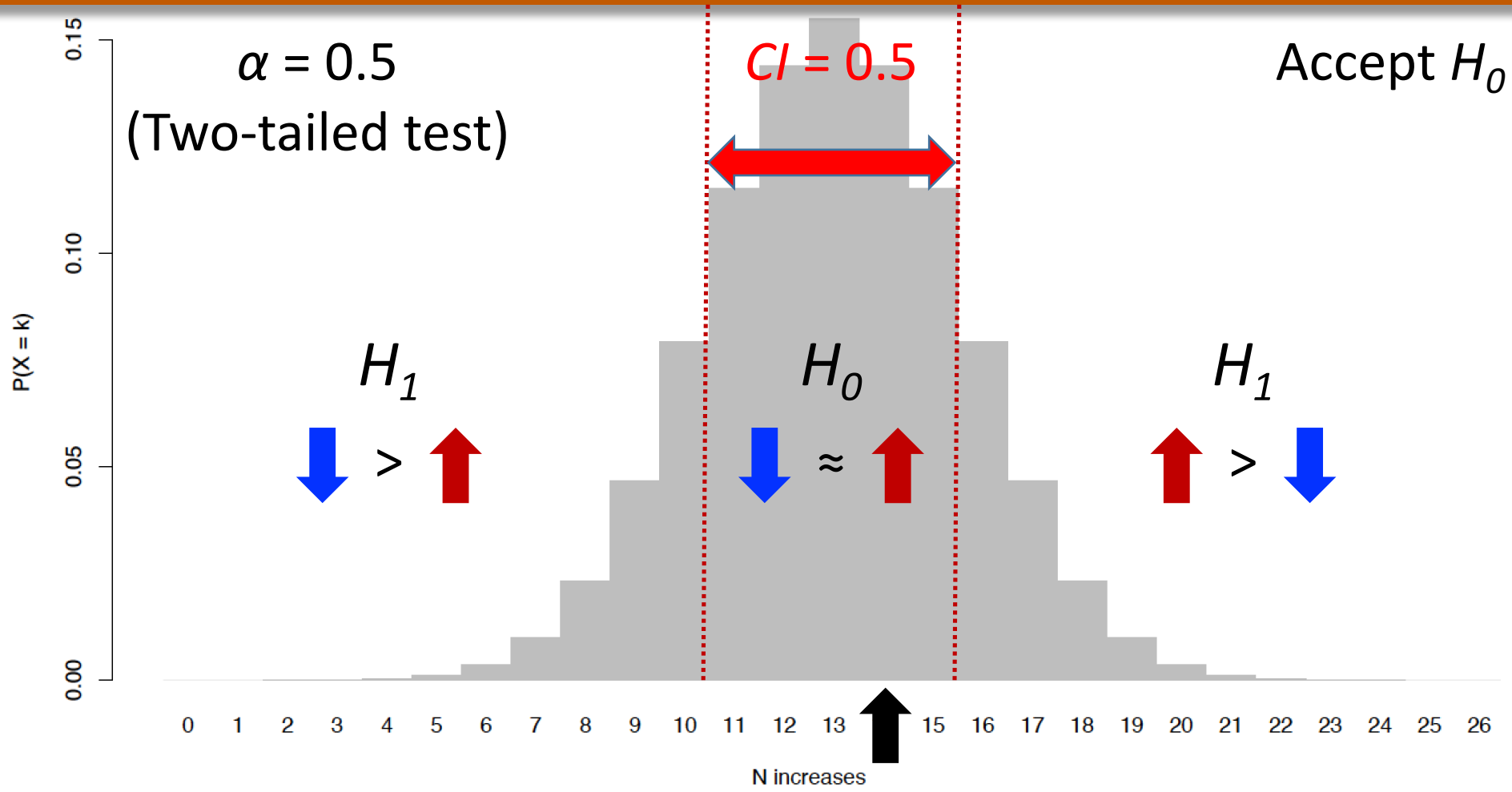
Two-tailed

Allow alpha to be split equally in two; accounts for higher and lower values

Choosing a hypothesis



Choosing a hypothesis



Choosing a hypothesis

Type I error

Rejecting the null when the null is correct (false positive)

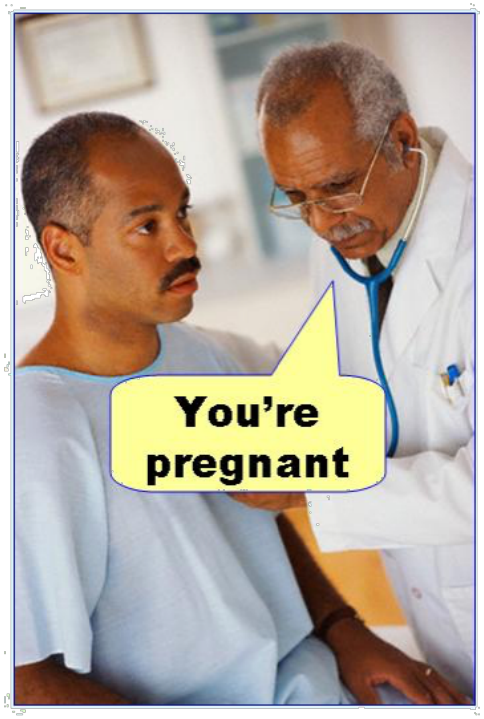
Type II error

Accepting the null when the null is incorrect (false negative)

Type I and Type II errors

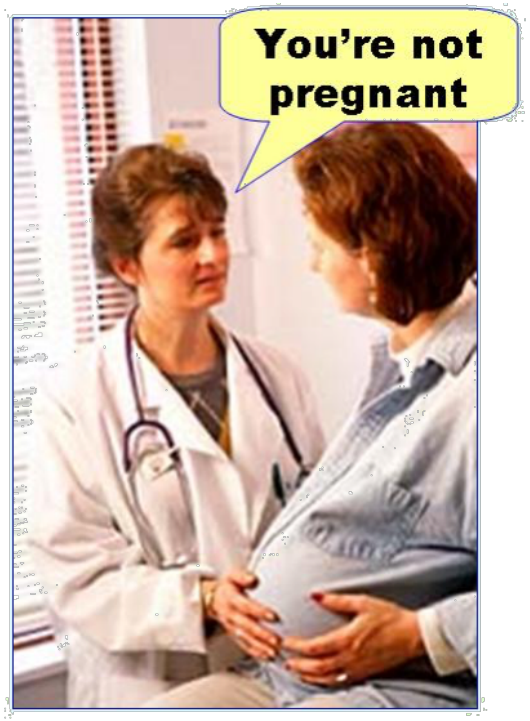
Type I error

False positive



Type II error

False negative



Problems with null hypothesis tests

Choosing alpha

No clear ideal value

Counterintuitive

Typically about low probability of model we believe is incorrect

Sample size dependency

Low sample sizes frequently lead to Type II errors

Alternatives to null hypotheses

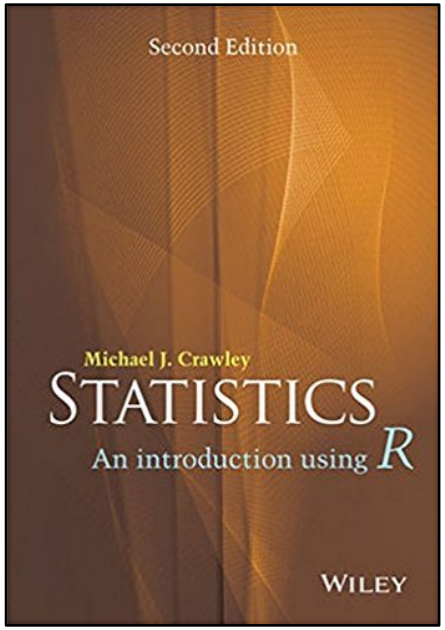
Maximum likelihood

Most likely value or model regardless of absolute probability

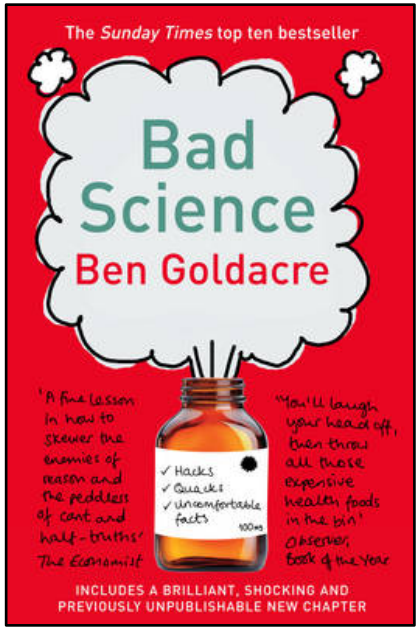
Bayesian statistics

Expresses probabilities as degrees of belief that update with new data

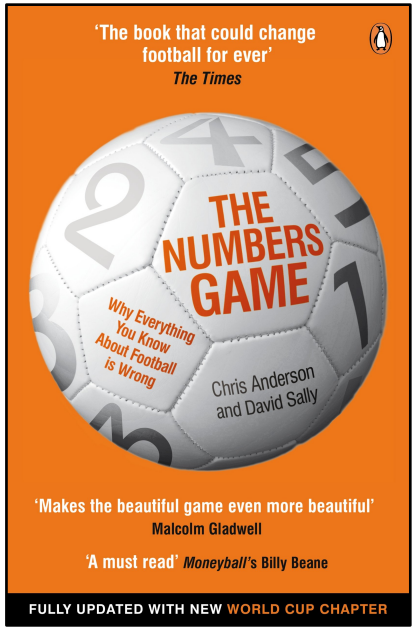
Reading



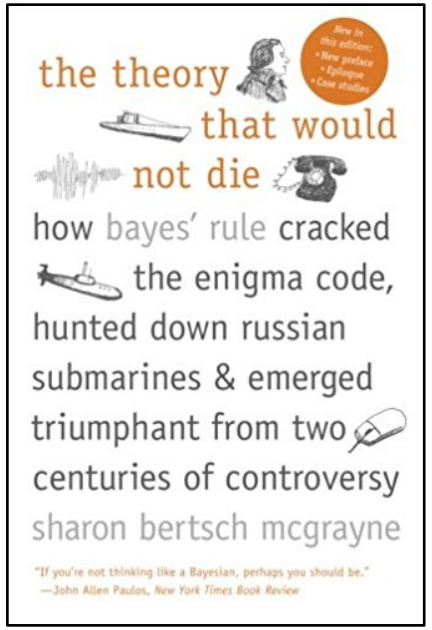
Statistics An Introduction Using R
Crawley



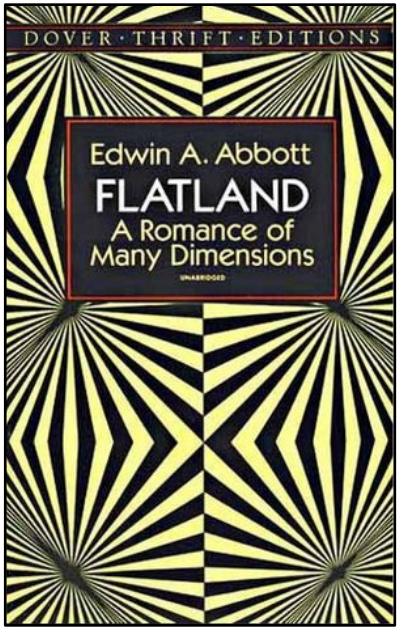
Bad Science
Goldacre



The Numbers Game
Anderson and Sally

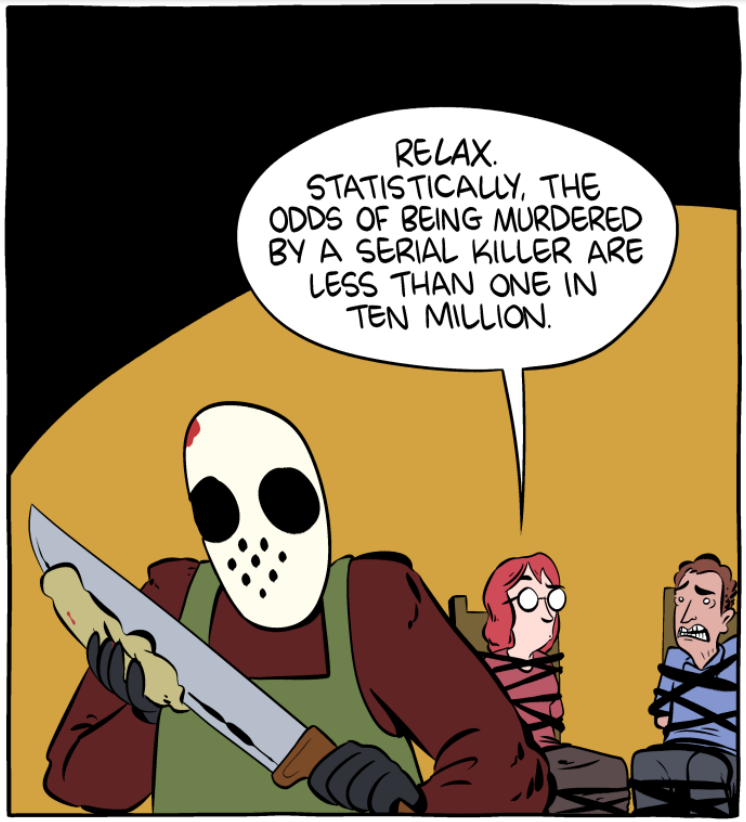


The Theory That Would Not Die
McGrayne

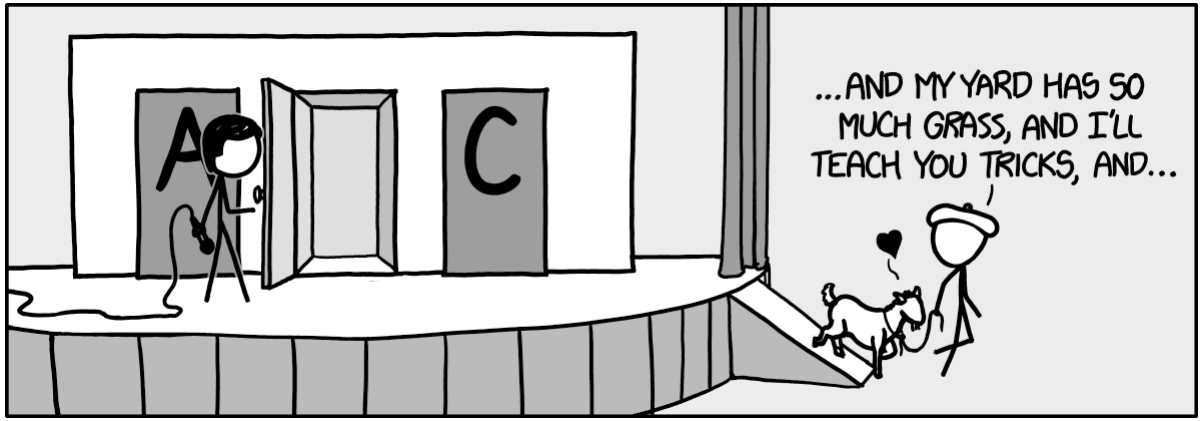


Flatland
Abbott

Web comics



Saturday Morning Breakfast Cereal



xkcd