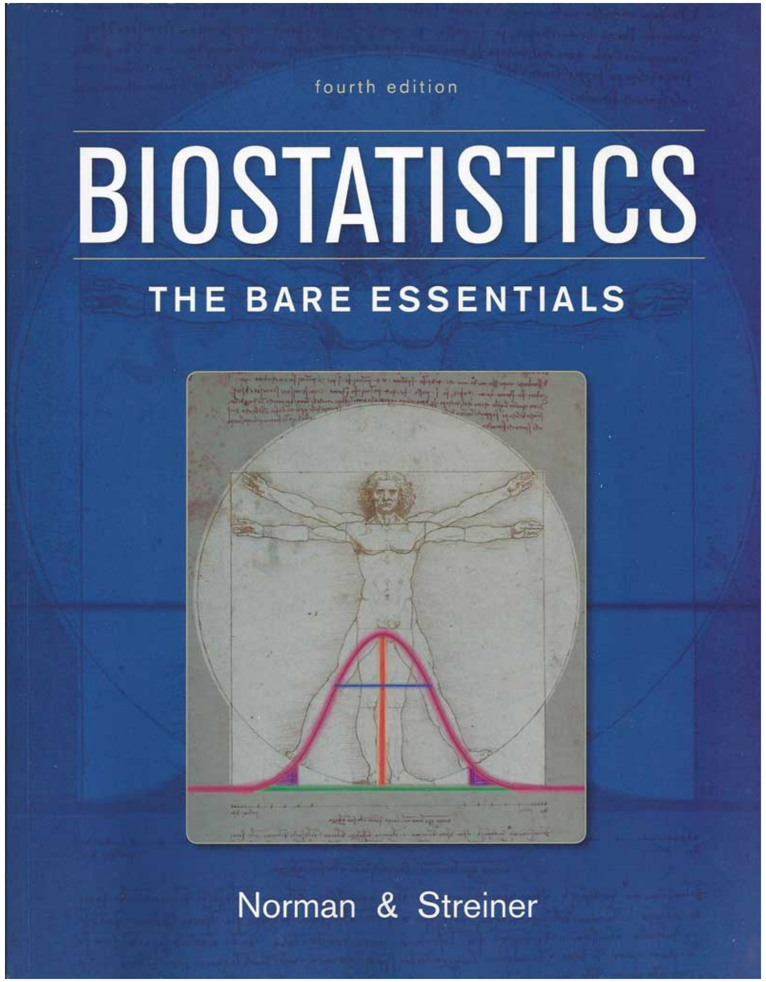
Data analysis and presentations: Examples from basic statistics

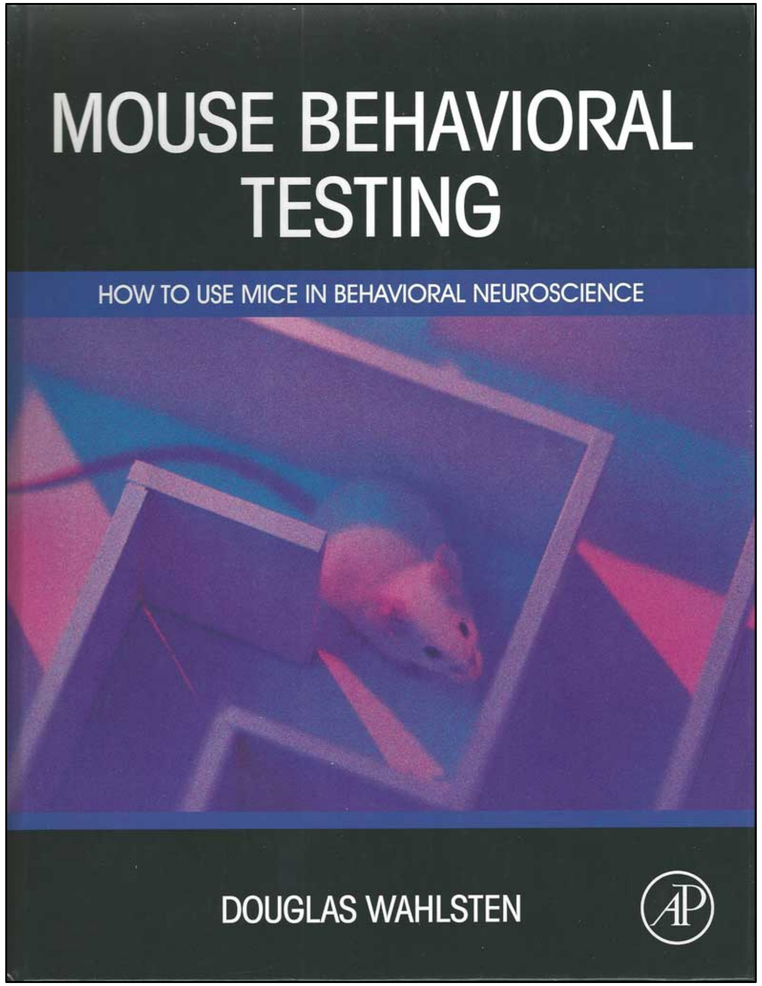
Fri 31.05.2019 13:00-14:30
ZNZ MD/PhD Neuroscience Course, Module BIO628
Room Y55 H12

David P. Wolfer, MD

Institute of Human Movement Sciences and Sport, ETH Zurich
Institute of Anatomy, University of Zurich

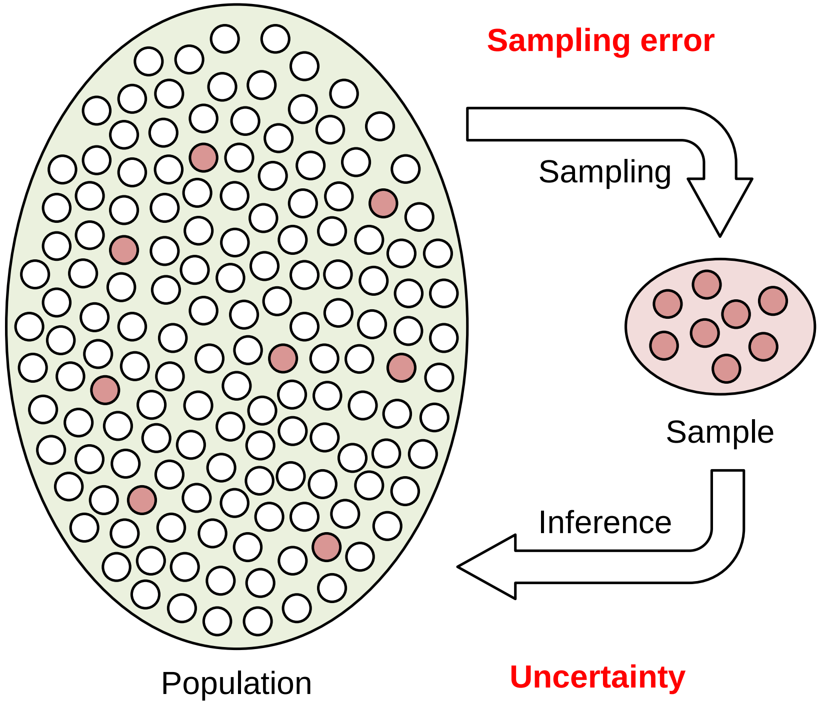


Norman GR, Streiner DL
Biostatistics, the bare essentials
BC Decker, 4. edition, 2014



Wahlsten D
Mouse Behavioral Testing
Academic Press, 1. edition, 2011

Population, sample and sampling error



Population	Sample	
$\mu = \frac{\Sigma\,(X)}{N}$	$M = \frac{\Sigma\,(X)}{n}$	mean
$\sigma^2 = \frac{\Sigma\,(X-\mu)^2}{N}$	$S^2 = \frac{\Sigma\,(X-M)^2}{n-1}$	variance
σ	S	standard deviation

Sampling error

- *Due to sampling error, mean and standard deviation of samples always differ from the true population values, they are only estimates of the mean and standard deviation of the population*

- *If multiple samples are drawn from the same population, due to sampling error their means and standard deviations will always differ from each other*

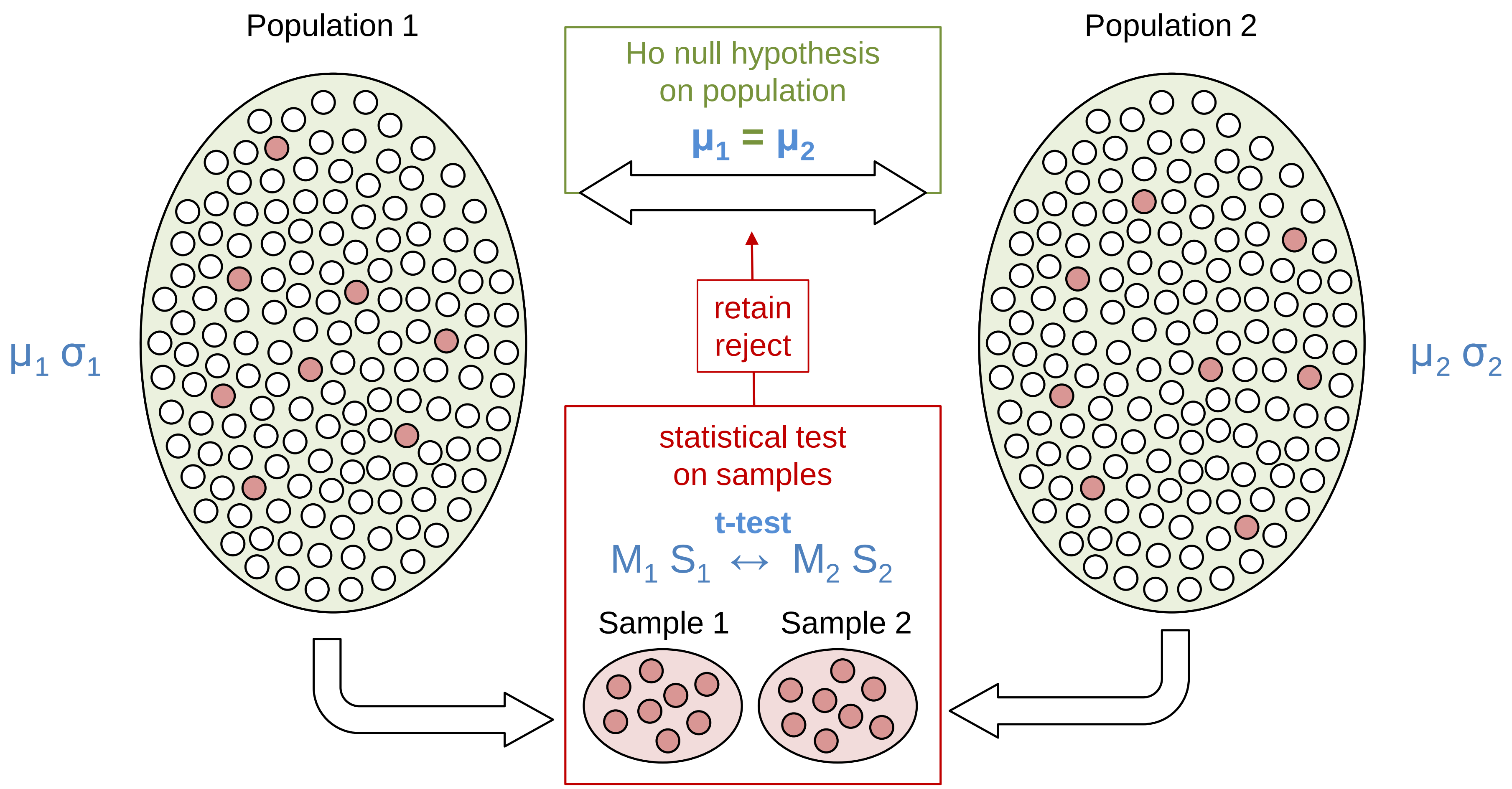
- *Repeated estimates of mean and standard deviations converge on the true population values, provided that*
  - *population data are normally distributed*
  - *sampling is unbiased*

- *Dispersion of estimates of mean and standard deviation decreases with increasing sample size:*

$$SE = \frac{s}{\sqrt{n}}$$

standard error of the mean

Statistical comparison of two populations



Hypothesis testing 1

	Ho false	Ho true
test positive: reject Ho	true positive	false positive
test negative: retain Ho	false negative	true negative

Sampling error:
We may be correct or not when using the results of a statistical test as criterion to reject or retain the null hypothesis

Hypothesis testing 1

	Ho false	Ho true
test positive: reject Ho	true positive	**α** $P(\text{reject Ho} \mid \text{Ho true})$ Type I error false positive
test negative: retain Ho	false negative	**1-α** $P(\text{retain Ho} \mid \text{Ho true})$ specificity true negative

Type-I error

- *reject Ho when it is in fact true = false positive*
- *likelihood **α** estimated from experimental sample data by statistical tests*
- *Ho rejected if estimate ≤ threshold, typically **0.05***

Hypothesis testing 2

	Ho false	Ho true
test positive: reject Ho	**1-β** P(reject Ho | Ho false) power = sensitivity true positive	**α** P(reject Ho | Ho true) Type I error false positive
test negative: retain Ho	**β** P(retain Ho | Ho false) Type II error false negative	**1-α** P(retain Ho | Ho true) specificity true negative

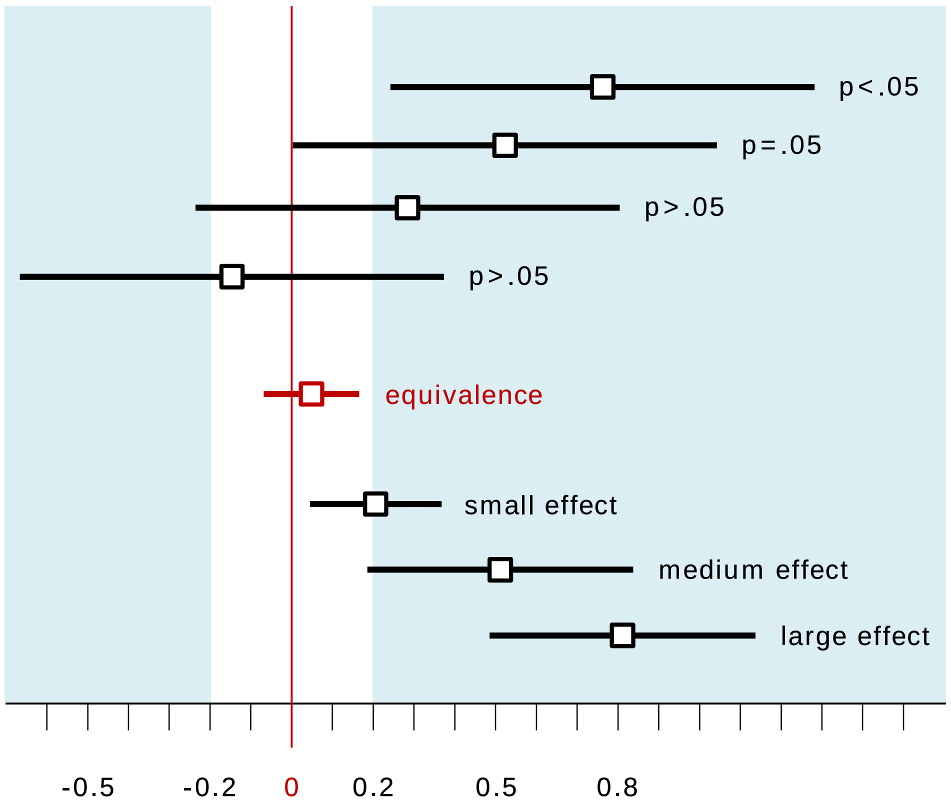
Type-II error

- *retain Ho when it is in fact false = false negative*
- *likelihood β determined by experimental design:*
  - *- sample size*
  - *- type-I error threshold*
  - *- effect size*
- *Typically accepted $\beta \leq$* **0.2**, *same as power 1-$\beta \geq$* **0.8**

t-test effect size and confidence interval

- *effect size for populations: $\delta = \Delta\mu/\sigma$*
- *effect size estimated from sample data:* **Cohen's d** $= \Delta M/S_{pooled}$
- *confidence intervals (CI) for mean difference or effect size can be estimated from data as measure of dispersion (larger with low power)*
- *alternative to α for* ***hypothesis testing****:* *p<.05 when 95% CI excludes zero*
- *criterion for* ***equivalence****:* *CI excludes ranges of relevant effects*

95% CI for Cohen's d



Hypothesis testing 2

	Ho false	Ho true
test positive: reject Ho	**1-β** P(reject Ho | Ho false) power = sensitivity true positive	**α** P(reject Ho | Ho true) Type I error false positive
test negative: retain Ho	**β** P(retain Ho | Ho false) Type II error false negative	**1-α** P(retain Ho | Ho true) specificity true negative

Determining sample size:

- *before experiment is done!*
- *depends on*
  - *type-I error threshold, typically* ***α=0.05***
  - *expected effect size, eg.* ***δ=1***
  - *desired power, typically* ***1-β=0.8***

Hypothesis testing 3

FPR ≠ α !

	Ho false	Ho true
test positive: reject Ho	**1-β** P(reject Ho | Ho false) power = sensitivity true positive	**α** P(reject Ho | Ho true) Type I error false positive
test negative: retain Ho	**β** P(retain Ho | Ho false) Type II error false negative	**1-α** P(retain Ho | Ho true) specificity true negative

FPR *(false positive risk)*

P(Ho true | test positive)

PPV = 1 - FPR (positive predictive value)

$$FPR = \frac{\alpha}{(1-\beta) \cdot \mathbf{R} + \alpha}$$

R = Ho false / Ho true
(pre-study odds)

prior probability = **R** / (**R** +1)

Hypothesis testing 4

NPV ≠ 1-α !

	Ho false	Ho true
test positive: reject Ho	**1-β** P(reject Ho | Ho false) power = sensitivity true positive	**α** P(reject Ho | Ho true) Type I error false positive
test negative: retain Ho	**β** P(retain Ho | Ho false) Type II error false negative	**1-α** P(retain Ho | Ho true) specificity true negative

***NPV** (negative predictive value)*

P(Ho true | test negative)

$$NPV = \frac{(1-\alpha)}{(1-\alpha) + \beta \cdot \mathbf{R}}$$

R = Ho false / Ho true
(pre-study odds)

Interpretation of test outcomes

- *positive test outcome:*
  type-I error p-value is not a measure of false positive risk
- *false positive risk is typically larger than type-I error p-value, especially with*
  - *underpowered studies*
  - *low pre-study odds*
- *Typically don't assume pre-study odds >1*
  - *pre-study odds <1 without prior data*
  - *<<1 for screening with poor or no hypotheses*
  - *~ 1 with prior data (replication, phase III)*
  - *> 1 for reproduction of established effect in rescue experiment*
- *false positive risk may be ≈0.05 when*
  - *p<.003 for hypothesis without prior evidence*
  - *p ≈.05 for hypothesis with prior evidence*
- *negative test outcome:*
  absence of evidence is not necessarily evidence of absence of an effect
- *negative predictive value is low and result inconclusive with*
  - *underpowered studies*
  - *high pre-study odds*
- *absence of effect should be demonstrated using dedicated tests for equivalence*