

Critical Appraisal Course for Emergency Medicine Trainees

Module 2

Statistics

Statistics

- Hypothesis testing and estimation
- P-values and confidence intervals
- Type I and type II errors
- Statistical significance
- Statistical power
- Multiple hypothesis testing

The role of chance

- Two common approaches to to estimate the potential effect of random error
- Depends upon how the research question is phrased
- Hypothesis testing: the P-value
- Estimation: the confidence interval

Hypothesis testing

- Research question phrased as a hypothesis
- Analysis aims to determine whether the hypothesis is true
- Stated hypothesis and null hypothesis
- Crude definition: P-value is the probability that the null hypothesis is true

Estimation

- Research question phrased as a measurement
- Analysis is used to estimate the measurement
- Confidence interval shows the precision of the estimate
- 95% CI is the range of values that has a 95% probability of encompassing the “true” value
- Narrower CI = more precise estimate

P-values v confidence interval

- Hypothesis testing requires a priori statement of what is considered clinically significant
- Both identify statistical significance
- CI shows the range of potential values
- CI shows potential for type II errors
- Too many p-values raise possibility of type I errors

Type I and II errors

	Alternative hypothesis is true	Null hypothesis is true
Experiment shows significant result	True positive No error	False positive Type I error
Experiment shows no significant result	False negative Type II error	True negative No error

Type I and II errors

Probability of statistical errors depends upon:

- Sample size
- Level at which statistical significance is set
- Number of tests performed

Statistical significance

- Set before undertaking the study
- Known as alpha
- P value below which the results are so improbable (given the null hypothesis) that we will reject the null hypothesis
- Alpha is therefore the maximum probability of a type I (false positive) error we are will to accept for each test performed

Statistical power

- Beta is the probability of a type II (false negative) error
- Having set alpha, beta is determined by the sample size
- The larger the sample size, the smaller beta
- Study power = $1 - \text{beta}$
- Conventionally power should be 80-90%

Determinants of study power

1. The level at which alpha is set
2. The sample size
3. The variability of the outcome measure, as defined by its standard deviation
4. The minimum clinically significant difference

Type I (false positive) errors

- If only one test is performed, probability of type I error is reflected by p value
- This is no longer true if more than one test is performed
- If you do enough tests you will eventually get a false positive result
- Isolated positive tests among a large number of tests are likely to be false positive

Multiple hypothesis testing

- Common mistake
- Sometimes known as “data dredging”
- Often seen in observational studies or routine data studies
- May reflect poorly planned (data driven) study
- May reflect attempt to extract a positive result from a negative trial (sub-group analysis)

Spotting type I errors

- Are there multiple hypothesis tests?
- Is there a clear rationale for the hypothesis tested?
- Does the hypothesis test flow from the objectives or only appear in the results?
- Does the methods section include a plan of analysis that specifies the positive test?
- Does the reported positive test suggest that other, more obvious, tests should have been performed?

Type II (false negative) errors

- How wide is the confidence interval?
- Does the confidence interval encompass potentially important differences?
- Is there a power calculation?
- What values were specified for alpha and beta in the power calculation?
- Does the minimum clinical significance seem reasonable?

Summary

- Hypothesis testing gives p-value
- Estimation gives confidence interval
- Type I = false positive error
- Look for multiple hypothesis tests
- Type II = false negative error
- Check power calculation and look at CI

Any questions or comments?