Statistics and your thesis:
How to do it right and enjoy the process

Jeromy Anglim

Department of Psychology,

The University of Melbourne

Correspondence: Jeromy Anglim
Email: jkanglim@unimelb.edu.au
For a complete copy of the presentation with
annotated slide comments:
Just email me or visit:
http://jeromyanglim.googlepages.com/

Statistics and your thesis: How to do it right and enjoy the process

For some students analysing their thesis data does not prove to be the most exciting experience in their academic life. This does not need to be the case. How can the results section be made the most important, interesting, and stimulating section in a thesis? What general frameworks can help us understand our results? What's the best way to approach the task of data analysis? How can we integrate statistics with our everyday understanding of the world?

In the talk, Jeromy will provide Advice on how to approach doing statistical analyses within the context of a thesis. Several big picture frameworks will be discussed through which we can try to understand the role and purpose of data analysis. In particular the importance of making reasoned decisions, telling an interesting story, becoming a scientist-practitioner, and the role of data analysis in improving democracy will be discussed. The importance of several contemporary issues relevant to anyone doing data analysis in psychology will also be discussed including, Confidence Intervals, Effect Size, Null Hypothesis Significance Testing, Power Analysis, Accuracy in Parameter Estimation and Meta Analytic thinking. The talk aims to provide some concrete strategies for assisting in the analysis of your thesis data, as well as showing why analysing your thesis data can be a motivating and passionate experience. Jeromy was a statistics consultant in the department for four years, where he spoke with hundreds of postgraduate psychology students about their thesis data analysis needs. He has also lectured 2nd and 3rd year statistics in psychology, lectured multivariate statistics in management, and regularly consults in relation to statistics in areas of climate surveys, market research, and test construction.

Overview

- The Big Picture
  - Answering a Research Question
  - Statistics & Democracy
  - Statistics as reasoned decision making
  - Staying open minded
- General orientation
  - Confidence Intervals
  - Effect Size
  - Null Hypothesis Significance Testing
  - Power Analysis
  - Accuracy in Parameter Estimation
  - Meta Analytic thinking
  - Sharing data with the world
- Getting the job done
  - Software
  - Learning Statistics
  - Taking your time
  - Telling a story

Jeromy Anglim - 2007

What does statistics mean to you?

Jeromy Anglim - 2007

Big Picture

- Statistics
  - Tool for **UNDERSTANDING** the world and developing theory
  - About explaining **VARIABILITY**
  - About **GENERALISING** from a set of observations to the broader population
- Tools for thinking about the world
  - Critically evaluate existing research
  - Conduct and report your own research

Jeromy Anglim - 2007

How can we refine our understanding of this process

- Ask some research questions
- Access Data
  - Take an existing dataset; or
  - Design and collect your own data
- Assess your research questions using the right statistical tools
- Answer your research questions
  - By writing up and drawing some conclusions about your research questions

"Can you create knowledge from empirical observations?
Or must you rely on others to digest it for you?"
provocative words put into Fisher's mouth by Anglim, 2007

Jeromy Anglim - 2007

There is something good for democracy about getting good at data analysis.
Good data analysis skills combined with an open mind to the answers that empirical data will bring represent a way of moving beyond ideology.
We can look at many debates at the moment and see how they could benefit from improved data and improved data analysis.
If we ourselves want to fully participate in these debates in our area of research or in the wider public domain, it is very powerful to be able to intelligently critique the conclusions that others have reached based on analysis of empirical observations.
Of course this requires much more than knowledge of data analysis. It requires knowledge of the substantive theory, research design, and principles of measurement, just to name a few.
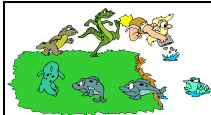
Staying open minded

- Pressures
  - Desire to support hypotheses
  - External pressure
- Advice
  - Listen to the data
  - Be ethical & Stay true to yourself
  - Don't turn psychology into infomercial-style empiricism

Jeromy Anglim - 2007

Statistics as reasoned argument

- Some popular *misconceptions*:
  - In statistics, there is always only one right answer
  - Statistical decision making is very different from everyday common sense
- Reality
  - Choices and decisions are everywhere
  - Need to make reasoned decisions based on data, theory & statistical literature

Making this shift requires that you know the "basics"… Because there are still plenty of wrong answers… and there are plenty of poorly reasoned arguments

Jeromy Anglim - 2007

I have noticed a tendency at least within psychology to make statistical decisions based on simple black and white rules. Examples of such rules: You need 100 participants to do multiple regression; p must be less than .05 to be statistically significant; if normality is violated, you must do a non-parametric test.

I am not saying that any of these rules are necessarily right or wrong. The point is that they are based on reasons. The reasons often get lost when following the rule. The idea about sample size selection relates to power and accuracy in parameter estimation. A sample size of 110 is not much different to a sample size of 90 when it comes to statistical power. The idea about normality relates to the accuracy of resulting p-values. The point is that good decisions have reasons. To make good statistical decisions theoretical understanding of the phenomena under study must be combined with the pros and cons of the options present.

- Know the options
- Understand at least some of the pros and cons of each
- Getting similar results from both approaches relaxes the burden of the decision a little

- Make a decision

Statistics as reasoned argument

- Decisions requiring reasoned argument:
    - What the research questions are
    - Which rule of thumb to apply
    - Which analysis to run
    - Within each analysis which options to choose
    - Interpretation and generalisation of results
- Some sources of reasons
    - Data
    - Theory
    - Statistical literature: Textbooks and primary journals
    - Purpose of research
    - Convention

Jeromy Anglim - 2007

What reasoned decisions do you foresee having to make?

What's your global orientation to statistics?

What are you thoughts on p values, effect size, power analysis and all that stuff?

Jeromy Anglim - 2007

Inferential Statistics

- Population Parameters
- Sample Statistics
- Hypothesis Testing
  - H0: Null Hypothesis
  - H1: Alternative Hypothesis
- P-value
  - The probability of obtaining a result as large as that observed in the sample if the null hypothesis were true.
- Alpha

Jeromy Anglim - 2007

Imagine we wanted to draw some conclusions about the nature of employees in a particular country. On average how many hours a week do they work? How much money do they earn on average? How many weeks holiday a year do they get? If we were going to research these questions, it is rarely feasible to obtain data from every person in the population of interest. Thus, we draw a smaller sample of people and assess them on how many hours they work, how much they earn and how many weeks holiday they get each year. We then may attempt to infer the characteristics of the broader population from our sample.

Population (parameters): A hypothetical or actual target population. We are trying to draw inferences about population parameters

Sample (statistics): A selection of individuals drawn from the population that provide sample statistics to estimate population parameters

Hypothesis testing: H0: null hypothesis; H1: alternative hypothesis

p value: The probability of obtaining a result as large as that observed in the sample if the null hypothesis were true.

Alpha: The probability of falsely rejecting the null hypothesis. Typically, we talk about alpha being .05 or .01

Hypothesis testing logic: if the p value is less than alpha (e.g., .05), the probability of the null hypothesis being true is low, we reject the null hypothesis, and accept the alternative hypothesis

Effect Size, Confidence Intervals, Power and Null hypothesis significance testing

- Old School
  - It's all about the p value
- In comes Effect Size, Power Analysis, and Confidence intervals
- Perhaps they can all work together as long as we remember their respective roles

Jeromy Anglim - 2007

Effect Size

- What are we trying to do?
  - Refine our estimates of population parameters
- Contrast with tests of statistical significance
  - Statistical significance does not mean practical importance
- Standardised vs Unstandardised
- Most common
  - Cohen's d
  - Pearson's R
  - Odd's ratio

Jeromy Anglim - 2007

Research enterprise: The main aim of most empirical research is to refine our estimates of population parameters. Whether we realise it or not, many of our research questions can be reduced to parameter estimates. If our research question concerns the effectiveness of treatment intervention versus a control group, then the parameter is the difference between the two group means. If we are interested in how extraversion relates to job performance, the parameter estimate is the correlation.

Contrast with tests of statistical significance: Effect size measures show the degree of relationship or extent of relationship. With a sufficiently large sample size, even small effects can be statistically significant. When evaluating the practical significance of a research finding, whether we are concerned with group differences or the relationship between variables, it is desirable to report a measure of effect size.

Standardised vs Unstandardised: Typically when people talk about effect size measures, they are referring to standardised effect size measures such as Cohen's d. However, there are many important unstandardised measures of effect. These include the difference between

two means and the standard error of the estimate. Unstandardised effect size measure are often preferable when the metrics are inherently meaningful. For example, if you were interested in knowing the difference in income levels of males and females, a good indicator would be difference in mean annual income in dollars.

Most common standardised effect size measures:

Cohen's d is the difference between two group means divided by the standard deviation. This can be the pooled within group standard deviation or the standard deviation of one of the groups.

Pearson's r is a correlation coefficient which ranges from -1 to 1 and expresses the degree of linear association between two variables.

Odd's ratio: This is used commonly used in logistic regression, when we are predicting a binary dependent variable. Odds represent the probability of an event occurring divided by something not occurring. The odds ratio involves dividing two separate odds. This is particularly common in medical research where you might for example compare the odds of getting a disease for those who have or have not had a particular vaccination (see the seminar on logistic regression for more detail).

The following table shows the relationship between Cohen's d and r. Cohen provided some rules of thumb that might guide the practical understanding of obtained effect sizes. These are also displayed on the table. It should be noted that the actual importance of an effect size will vary across contexts.

Slide 14

Connections:
Power, Effect size,
& Statistical Significance

Really understanding this slide will make you a better human being…
Okay… Maybe not… But I think it's important

- **POPULATION EFFECT SIZE**
  is **NOT** affected by:
  1) **SAMPLE SIZE**
  2) **ALPHA**

Aim:
To know the degree of relationship and the practical Importance of findings

- **POWER**
  is a function of:
  1) **SAMPLE SIZE**
  2) **POPULATION EFFECT SIZE**
  3) **ALPHA**

Aim:
To have a reasonable chance of testing the research question

Alpha: is by convention .05 or .01 unless some form of post-hoc adjustment is being applied
Sample size: is a function of: availability, money, time, & other resources

- **STATISTICAL SIGNIFICANCE**
  in a particular study is a function of:
  1) **POWER**
  2) **CHANCE**

Aim:
To rule out chance as a plausible explanation of observed relationships

Jeromy Anglim - 2007

Slide 15

Effect Size vs Significance Testing

Null Hypothesis Significance Testing (NHST)

- Key Question
  - What is the probability of obtaining the study's results if the null hypothesis were true ?
    - e.g., no difference between groups; no relationship between two variables
- Answering Question
  - Select appropriate significance test and assumptions
  - Set your alpha
  - SPSS: Compare alpha to obtained p value
  - Hand calculations: compare obtained test statistic to critical value of test statistic at the alpha you set

Effect Size

- Key Questions
  - What is the size of the effect?
  - What does the size of the effect mean in a practical sense?
- Answering Question
  - Select appropriate effect size
  - Contextualise obtained effect size:
    - Use clinical or professional judgement
    - Compare to obtained effect sizes in similar studies in research area
    - Compare to typical effect sizes in studies in the broader discipline
    - Use rules of thumb of small, medium and large effect sizes (e.g., Cohen's)

Jeromy Anglim - 2007

Confidence Intervals

- Meaning
  - 95% confident that *population parameter* lies in the specified interval
- In theory,
  - Confidence intervals are available for any standardised or unstandardised effect size measure
  - Any estimate of a population parameter from a sample statistic can have a confidence interval
- Major push in psychology
  - Think in terms of confidence intervals around effect sizes
    - Forces you to think about uncertainty
    - Focuses thinking on population estimation, which is the basis of theory
    - Encourages *meta-analytical thinking*

Jeromy Anglim - 2007

Meta Analysis

- Purpose
  - Estimate true population effect
- Composite effect size measure
  - typically d or r
- Adjustments
  - Reliability
  - Weighting by sample size
  - Publication bias / file-drawer effect
- Potential to examine moderators
- Implications for literature reviews

Jeromy Anglim - 2007

Purpose: Meta analysis is a method for systematically integrating the results of many studies.

> The emphasis is on getting a better estimate of the 'true' population relationship.
> While a complete understanding of the nature of meta-analyses is beyond the scope of this course. Some important elements include:

Composite effect size measure (typically d or r)

> Because different studies use different measurement scales, meta-analysis typically creates a standard effect size measure from the results from each study sampled.
> Cohen's d: difference between two groups in terms of standard deviation (e.g., effect of training on job performance)
> r: Correlation, typically used when looking at the relationship between interval level variables (e.g., commitment with job satisfaction)
> The Odds ratio is also commonly in medical settings

Adjustments

Meta analysis also adjusts for reliability problems. If measurement is less than perfectly reliable, observed correlations will be attenuated. Thus, an estimate of the true correlation can be obtained.

Frequently also samples are weighted by sample size so that studies with greater sample size have greater influence.

Issue of publication bias / file-drawer effect: This is the problem that studies with non-significant findings tend to be harder to publish. Thus, published articles may be over-optimistic about the population relationship.

Potential to examine moderators

A moderator is a variable that alters the relationship between two other variables. When there are many studies, moderators can be examined.
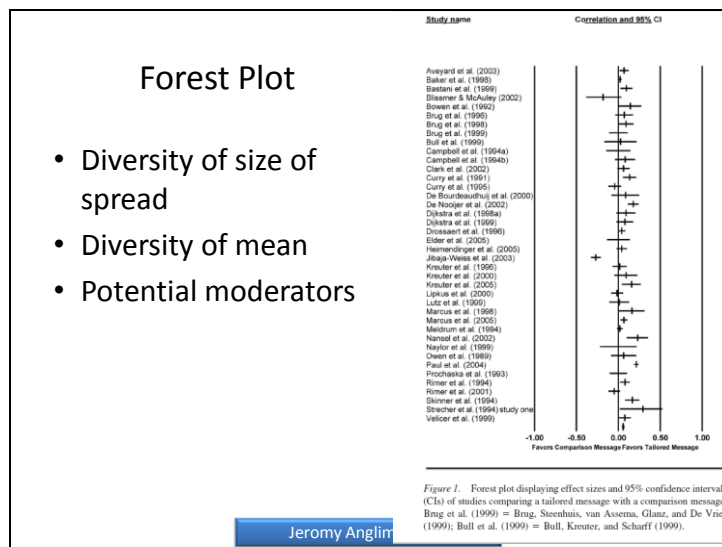
E.g., perhaps there is a stronger relationship between satisfaction and performance in studies where people have control over their work environment (i.e., control is a moderator of the satisfaction-performance relationship)

Implications for literature reviews: When summarising a literature, it is often useful to start with a meta-analysis where available rather than individual studies.

Slide 18
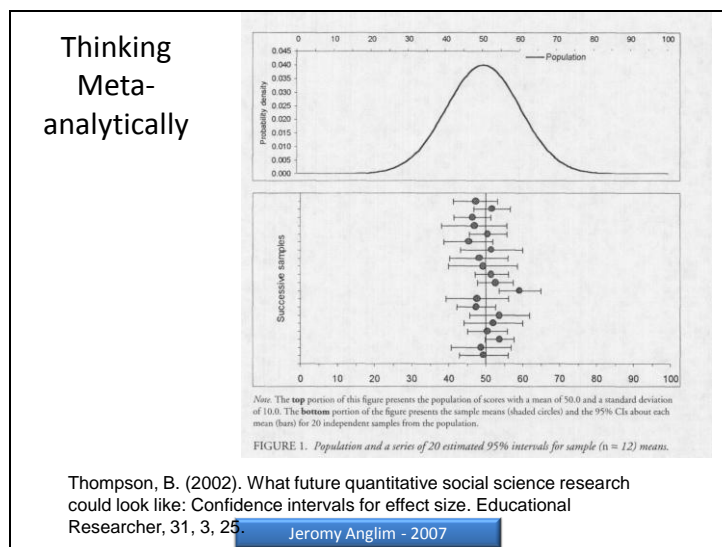
This diagram aims to show the meaning of a 95% confidence interval. The idea is that 95% of the time the true population parameter is contained with the confidence intervals.

Figure 1. Forest plot displaying effect sizes and 95% confidence intervals (CIs) of studies comparing a tailored message with a comparison message. Brug et al. (1999) = Brug, Steenhuis, van Assema, Glanz, and De Vries (1999); Bull et al. (1999) = Bull, Kreuter, and Scharff (1999).

Noar, benac, harris, 2007

The Forest plot shows cohen's d and associated confidence intervals for a series of studies looking at the difference in communication effectiveness of tailored messages with comparison messages.

Confidence intervals that do not cross zero represent statistically significant differences. Imagine if one of those studies was your own. Imagine also that you were not aware of the range of results that had proceeded. How would you interpret your result in isolation? How would you interpret your study if you interpreted it within this meta-analytic context?

Given the diversity of the results, it is quite probable that different studies are estimating different population effect sizes. Even studies that are looking at the same relationship (e.g., job satisfaction and performance) differ in important ways. Samples differ (e.g., one sample might be white collar and another blue collar); measurement tools differ (e.g., performance might be measured very differently); quality of study varies. All these factors can lead to variation in results.

The idea is that some of the variation in means we see is due to principles of random sampling and some of the variation is due to systematic differences in estimated population effects.

Meta analytic thinking

- Thompson (2002, p.28) defines meta analytic thinking as:
  - "a) the ***prospective formulation*** of study expectations and design by explicitly invoking prior effect size measures and
  - b) the ***retrospective interpretation*** of new results, once they are in hand, via explicit, direct comparison with the prior effect sizes in the related literature" [italics added]

Jeromy Anglim - 2007

Power Analysis

- Power:
  - Probability of correctly rejecting H_0 when H_0 is false
- Type I error rate (alpha)
- Type II Error rate (Beta)
- Power (1- Beta)

- Increasing Power
  - Increase Effect Size
  - Increase Sample size
  - Reduce Alpha
- Power is a property of a hypothesis not a study

Objective Status	Decision	
	Retain H0	Reject Ho
H0 True	$(1-\alpha)$	Type 1 error (α)
H0 False	Type 2 error (β)	Power $(1-\beta)$

Jeromy Anglim - 2007

In the logic of hypothesis testing, there are two possible states in the world and we can draw two possible conclusions. If the null hypothesis is true and conclude that this is the case, then we have made a correct decision.

Power is the probability of correctly rejecting the null hypothesis when it is false in the population.

To state the case strongly, a study that with insufficient power is not worth doing.

From a practical perspective, it is important to know what power is and what increases it.

Power increases with bigger samples.

Power increases with bigger effect sizes.

Power is a property of a hypothesis not a study.

Typical opening question regarding power analysis with a statistics consultant:

Student: "My supervisor/committee/grant form/ ethics form requires me to do a power analysis to work out what sample size I need to have reasonable power. What sample size do I need for my study to have 80% power?"

What's wrong with this question? Power analysis is not a property of a study, it is a property of a hypothesis.
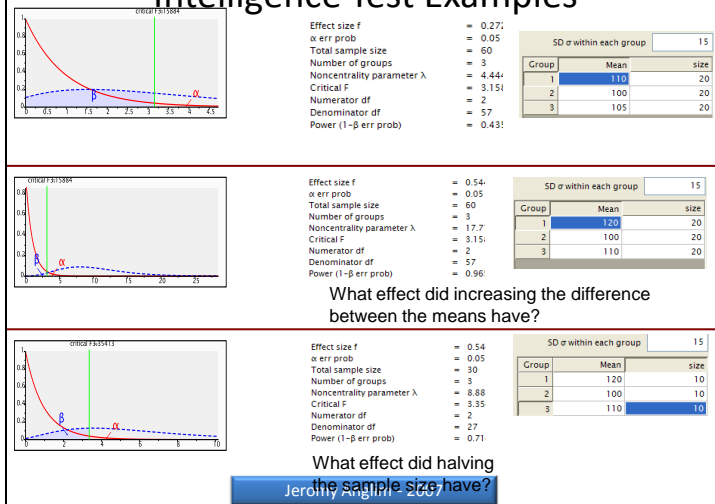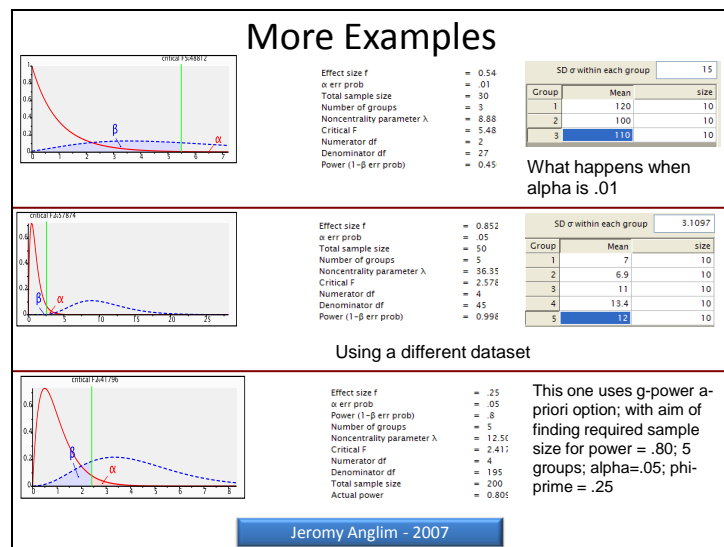
Slide 22

Slide 23

One of the nice aspects of G-power is that you can play around with different results and train your intuition about the relationship between power, sample size, effect size and alpha.

Slide 24

Slide 25

A-priori power analysis is typically used at the design phase to determine what sample size is required to have a reasonable chance of obtaining statistically significant results.

To work out the required sample size, we need to specify the population effect size, alpha and power.

Alpha is typically set at .05 by convention.

Population effect size is usually not known exactly. Meta analyses and large scale studies are typically the best indicators of what to expect. Any prior research in the area or in related areas may be useful. You might also fall back on Cohen's rules of thumbs combined with you intuition about what the effect size is hypothesised to be.

Desired power is often set at .80 as a reasonable minimum. However, this is only a convention.
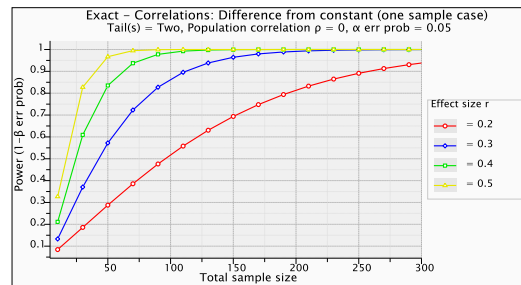
G-Power – Post hoc power analysis

- Did I have a hope of finding a statistically significant result?
  - Effect size
  - Alpha
  - Sample size

Jeromy Anglim - 2007

Post hoc power analysis aims to determine the power of the study on the basis of alpha, the sample size and the obtained sample effect size.

Power Curves

Exact – Correlations: Difference from constant (one sample case)
Tail(s) = Two, Population correlation ρ = 0, α err prob = 0.05

Effect size r
= 0.2
= 0.3
= 0.4
= 0.5

Jeromy Anglim - 2007

G Power 3 lets you produce some very nice power curves that show the relationship between the variables.

This graph shows that if you are expecting a small correlation (i.e., r=.2) you will need 180 participants in order to have 80% power.

The Core Message:

– Sample as many people as you can

– Appreciate that in most studies there are fairly large 95% confidence intervals around results

– Don't replace p-value fundamentalism with effect size or power analysis fundamentalism

– See results in terms of a meta-analytic context

– Sample size estimation is more than just power analysis

Jeromy Anglim - 2007

Accuracy in Parameter Estimation (AIPE)
Going beyond Power Analysis

- Better to think about sufficiently small confidence intervals on parameter estimates
- Tasks
  – **PLANNING:** What sample size is required to have a defined confidence interval around an effect size?
  – **INTERPRETATION:** What confidence interval is around the obtained effect size

Jeromy Anglim - 2007

The aim of a priori power analysis is typically to determine what sample size is required to infer with reasonable probability (e.g., 80%) that a population parameter is significantly different from zero. While this has the potential to improve study design, in many areas of research, the interest is not about whether there is or is not an effect, but rather the size of the effect.

If my interest is on the relationship between intelligence and job performance, I am already pretty confident that there is a relationship from the thousands of studies and meta-analyses that have established this in the past. A conclusion that the correlation is greater than 0 would be of little interest. My interest is centred on what the strength of the correlation. Is it small (r=.1), medium (r=.3), Large (r=.5) or very large (r=.7)? How can I make sure that the confidence interval of the correlation that I obtain is small enough so that I can get a reasonable estimate of the true population correlation?

On a practical level, this tends to mean that even larger sample sizes are required than would be suggested based on power analysis.

If you are looking for practical rules of thumb, I tend to find N=100 to be reasonable and N=200 to be quite good. At these levels 95% confidence intervals around correlations become reasonably accurate in the sense that we have a reasonable approximation of the population effect size relative to rules of thumb.
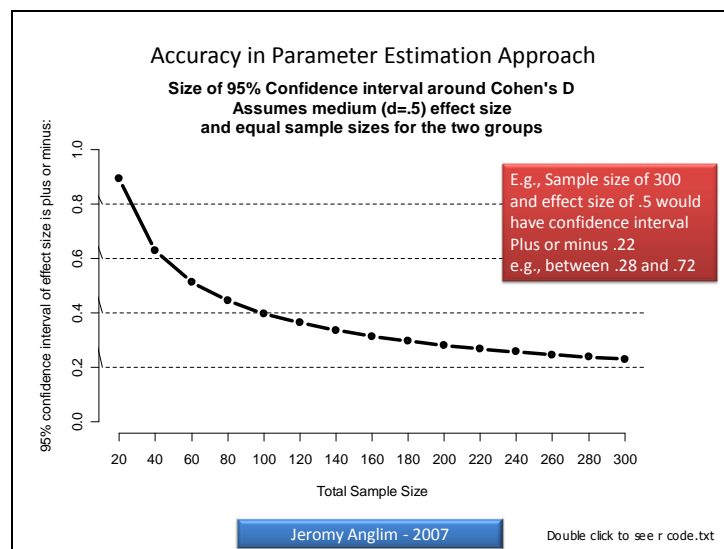
I have seen otherwise good textbooks recommend not having too large a sample size as this may lead to unimportant relationships becoming statistically significant. I find this advice misguided. If you are trained in an appreciation of effect sizes, you know that statistical significance does not mean practical significance. All else being equal bigger samples will always provide better population estimates.

For articles about the approach:

http://www.indiana.edu/~kenkel/publications.shtml

http://cran.r-project.org/src/contrib/Descriptions/MBESS.html

Slide 30

The above graph was created in the free open source software package R. The code used to create the graph is available on the digital version of the slide, if you are interested. The code may appear complex at first but the more you use R and the more online tutorials you work through, the more familiar it all appears. In particular obtaining the basic information about confidence intervals for effect sizes is not too difficult once you know the basics. The help files are quite extensive.

In terms of interpreting the graph, the key point is that if we are trying to assess the population effect size for our study, we want to know what level of confidence we want to have. Thus, instead of choosing a sample size based on power analysis, we choose a sample based on what will give us a desired confidence of the population effect size.

We might decide that plus or minus .3 in terms of Cohen's D would be adequate. Reading off the graph

I have found the prospect of greater sharing of data quite exciting. Imagine a world where for any journal article you were reading, you could obtain the original data in a well documented form and quickly run the analyses that you wanted to run.

Advances in meta analysis: One of the critiques of meta-analysis is that it is based on summary data. If researchers systematically provided complete datasets as part of publication, datasets could be combined to allow for sophisticated analyses beyond the scope of a single study.

Educational Aid: Reading existing journal articles is an excellent way of learning how to report particular statistical techniques. However, imagine the benefits that could be obtained if doctoral students had the opportunity of recreating the results of the core journal articles they have read.

Increased acceptance, if the ideas are good. In many journal articles the authors decide to test a particular regression model or decide on a factor analysis with a particular number of factors. What if you think that they might not have made the right decisions. What if you would have run the analyses differently? What if there were some analyses which they could have done based on the variables they reported, but they didn't run? In all these situations, imagine the contribution that could be made if the community of scholars were able to access this data.

Further reading: The following authors have wrote passionately on this topic:

David H. Johnson: http://www.psychologicalscience.org/observer/0102/databases.html

Garry King: http://gking.harvard.edu/replication.shtml

Right to first publication: There is concern about sharing data that others may publish articles based on your research before you have had opportunity to publish yourself. Of course time limits can be placed on access.

Confidentiality: Data may be stripped of identifying information.

Fear of being caught making a mistake: While this is may be a rational fear particularly for those less skilled in data analysis, it hardly seems grounded in our desire to advance the profession. If anything, accountability tends to improve quality.

It is unfortunate that governments, granting bodies, and universities do not seem to determine merit in terms of contributing data to the research community.

Statistical Software

- Major packages
    - SPSS, SAS,STATA, S-Plus, R, and more
- Dedicated Packages
    - SEM: Amos, Lisrel, more
    - Multilevel Modelling: HLM, MLWin, more
    - More…
- Decision
    - What kind of user are you?

Jeromy Anglim - 2007

There are many statistics packages out there.
Most people I know in psychology, market research and management use SPSS.
Once you learn one tool, it is not too difficult to learn another.

Considering 'R'

- Open Source – i.e., Free
- http://www.r-project.org/
- Very powerful
- Encourages a better orientation to data analysis
- Takes longer to learn
    - Command line; not really GUI
- Many innovative features not found in other programs
- Improving rapidly

Jeromy Anglim - 2007

Open Source – i.e., Free: Being open source software, it is free to use. Thus, if after university, you work for an organisation and it does not have SPSS or SAS or any other commercial package, you will not be limited.
Takes longer to learn It runs on the command line (although GUIs do exist) and thus requires a different mindset to that of typical point-and-click programs such as SPSS.
There are extensive tutorials on the internet to get you up an running. I would think that about 100 hours of practice would be enough to get you up and running with R. But equally, once you do learn it, many tasks become a lot easier.
Who's it for? While anyone could get benefits from using R, the following people are likely to get the most benefit:
- You need to automate graphs and reports on large datasets
- You are quite good at statistics

- You see statistical analyses as an ongoing integral part of your career, academic, consulting, or otherwise
- You have done some form of computer programming in the past
- You want to take advantage of particular features that R offers
- You want to write your own custom statistical procedures or techniques and perhaps share them with others

Users: *Dabblers*: If you learn the basics of R, you can use it just for specific features that are not available in your normal commercial software. *Mainstream Users:* You can use R to do the things you do in normal statistics packages *Power Users:* You take your statistical analysis to another level of sophistication and are able to do many sophisticated tasks efficiently. You are able to produce large scale customised reports in minimal time.

Innovative features: Confidence intervals for effect sizes; cool plots; polychoric correlations; meta-analysis tools; item response theory tools; built in datasets to assist your statistical learning; many advanced regression modelling procedures

Improving rapidly: Many statisticians develop the latest techniques for R many years before they are ported to other statistics packages, if they are ever ported to other statistics packages.

Slide 34

The "R" Orientation

- Customise output for your needs
- Feed results of one analysis into the results of the next
- Use graphs extensively to explore data
- Document all analyses with syntax
- Integrate more with mainstream statistical community

Jeromy Anglim - 2007

Even if you decide R is not for you. There are many useful messages about an orientation towards statistics emphasised by R which you can take into any statistical package.
In particular, if you are using SPSS, it is essential that you start using syntax. Record your analyses using the paste command. It's important that you have a record of the analyses that you have run and the steps you have taken. This allows you to return to your analyses at a later date and re-run them easily. It also allows others to check your work.

Statistics Self-development plan

- Where do you want to be?
- What books, articles, websites to read?
- What courses, practice exercises are there to help you learn?

Jeromy Anglim - 2007

The right support materials – textbooks

- Textbooks
  - Hair, et al – Multivariate Data Analysis (6th Ed.)
  - Tabachnick & Fiddel – Using Multivariate Statistics
  - Pallant – SPSS Survival Manual
  - Field – Discovering Statistics Using SPSS
- The internet
  - Google the technique
  - http://www.ats.ucla.edu/stat/
  - http://www2.chass.ncsu.edu/garson/pa765/statnote.htm
- A statistical friend

Jeromy Anglim - 2007

Hair et al: A massive and comprehensive book on multivariate statistics written for people doing research in business; includes section on SEM and Confirmatory factor analysis; justifications and references to primary statistics literature are sometimes a little light-on

Tabachnick & Fiddel: Another massive multivariate book; greater coverage of underlying matrix algebra and formulas than Hair; Examples are drawn more from psychology.

Pallant: Designed specifically for people going through the process of using SPSS for their thesis; more like a cook-book; If you don't do statistics very often and want an easy to apply recipe – this is the book.

Field : This book is awesome. Its funny. It explains the statistical ideas clearly using everyday language without overly sacrificing sophistication of understanding. Covers regression, PCA, different types of ANOVA and more.

Howell: This book is good if you want to get a deeper understanding of statistics including formulas with particular strengths in various ANOVA designs.

Internet:

http://www2.chass.ncsu.edu/garson/pa765/statnote.htm
This site is comprehensive and also has a nice set of links to other key sites
Basically with a few simple searches you can usually get useful information about almost any technique you are likely to use. You can also usually get it explained in SPSS.
http://www.ats.ucla.edu/stat/
The UCLA site is excellent and comprehensive. It has many great tutorials including videos and worked examples from key textbooks. It also covers all major statistics packages and many niches ones as well.
A statistical friend: Perhaps your office buddy; your supervisor; a lecturer who knows their stats; a statistical consultant. The statistics department also runs a consulting service.
The key to making the most of such people: 1) show that you've thought about the problem and made some initial attempts at answering it by reading up; 2) be prepared, structured and specific with your questions; 3) understand that statistical consultants are there to assist you in understanding the pros and cons of different options . Many statistical decisions can not be outsourced and rely on reasoned decisions that should be made by the researcher.

Slide 37

Learning Strategies

- Practice with Feedback
- Explore formulas
- Try to see commonalities across topics
- Consolidate the basics if necessary
- Integrate statistical understanding with everyday awareness of the world
- Look at examples of how others have written up similar analyses

Jeromy Anglim - 2007

Taking your time

- "Conducting data analysis is like drinking a fine wine. It is important to swirl and sniff the wine, to unpack the complex bouquet and to appreciate the experience." (Wright, 2003)
- Graphical exploration
- Examine individual cases
- Assess assumptions
- Learn about your statistical models
- Ask questions of the data

Jeromy Anglim - 2007

Wright, D.B. (2003) Making friends with your data: Improving how statistics are conducted and reported. British Journal of Educational Psychology, 73(Mar), 123- 136.

Telling a story

- Link between empirical observations and knowledge
- Show what is important
- Explain reasons for decisions
- Explain why it is important

Jeromy Anglim - 2007

Self development plan: Acquiring statistical skills requires a strategy. Knowing about the resources available is important.

The Research Question: Know how to choose the right statistical tests given the research question. Know how to link statistical tools and statistical output back to the research question and your intuition. Keeping the research question as the central focus for

evaluating everything we do helps to keep us focused.