Assessment and student self-assessment of exit level statistics modules in a computer based world

Sean van der Merwe

University of the Free State, South Africa

2 Aug 2019

DON'T PANIC

Overview

What is statistics? Like even really?

- "Statistics is the grammar of science." — Karl Pearson
- Statisticians facilitate the following process (specifically the arrows):
- Raw measurements $\longrightarrow$ Data $\longrightarrow$ Information $\longrightarrow$ Knowledge
- So the main skills required are: manipulating data and understanding uncertainty
- In practical terms, in 2019, this means high level (*i.e.* easy) programming skills and decent maths skills (with a focus on randomness and probability)

UNIVERSITY OF THE FREE STATE
UNIVERSITEIT VAN DIE VRYSTAAT
YUNIVESITHI YA FREISTATA
UFS UV

What does a statistics test look like?

- For the most part they are rather boring
- They generally have a lot of application of established formulas
- But some modules allow for making things interesting
- In the third year and fourth year modules that I taught in recent years I've done some unorthodox things, particularly with assessments
- That's what this talk will focus on
  - I will try my best not to preach or tell you what to do, only throw out ideas that I tried that weren't a complete flop
    - (to the surprise of most of my colleagues when they hear about them for the first time)

UNIVERSITY OF THE FREE STATE
UNIVERSITEIT VAN DIE VRYSTAAT
YUNIVESITHI YA FREISTATA
UFS UV

Overview

Unorthodox?

- Assignments and semester tests are done on computer
  - Instructions and resources are obtained on Blackboard
  - Students work on lab computers, laptops, or home computers
  - They submit on Blackboard using the 'Assignment' assessment type
- Tests are open-book and open-internet
  - Only restriction is communication tools — semester tests are individual assessments
  - I don't have exams — traditional exams don't reflect the working environment these students will face
- This idea is scary for many colleagues because I'm giving students control
- I constantly get the question: "How do you stop them cheating?"

Cheating

- I avoid asking questions with a definitive solution
- Shift the focus to solving complex problems
- Award marks for the thinking process and logical steps taken, not for giving the answer
- Force students to discuss the results they get, not just paste them raw
- And keep the pressure on so that they don't have time to scheme
- Of course I still have to watch them in tests so that they don't pass around flash disks or email each other, but when the vast majority aren't trying to cheat then a cheater is easier to spot
- Simply put: a student will not look at their neighbour if they know that their neighbour also doesn't know the answer

Examples

- In this example we are trying to test whether students can apply basic hypothesis tests on given data.
- In this example we are trying to test whether students can work with data measured over time.

Overview

Reasoning for peer assessment

- For the first semester test in one module, which is the most formative, I was inspired by my colleague (Zani) to implement peer assessment
- The goal was to force students to engage with the test, and particularly their shortcomings, after the test
- My students love to lie to themselves and think they understand the work when they don't, and this helped them to realise this
- It was not less work in total, but far more engaging and interesting for everyone
- But since the test counted a lot of marks I had to be careful. I had to...

The lecturer does oversight

- anonymise all the scripts thoroughly
- redistribute them randomly (through email) ensuring no student gets their own test
- get all the tests assessed and returned to me in good time (through Blackboard)
- record and anonymise the marking
- redistribute them randomly again for review (email)
- https://seanvdm.co.za/shinyapps/listshuffler/
- get them back again in good time (Blackboard)
- spot check the marking
- Give them back to the students (Blackboard)
- Take student complaints about how their test was marked and evaluate them

Overview

Why 'Assignment'?

- To enable good assessment, students must be able to answer with code, maths, graphs, and written explanations interspersed
- Thus we need submissions in a modern shareable and mark-able format such as .pdf or .docx
- Until recently that meant that students would:
  - Make a Word document
  - Read a blog post or watch a YouTube video to see how to do the thing they were supposed to study but didn't
  - Copy their code to the document
  - Run their code and realise there's a problem
  - Fix the problem and run again
  - Copy the output (text and graphs) to the document
  - Forget to copy the changed code, creating inconsistencies
  - Misinterpret or fail to interpret the output, and type nonsense in their document

Markdown

- The pattern on the previous slide is still the status quo
- But there is a new way that is becoming very popular in research and some industries: markdown
- It was used to good effect by the least lazy students in my classes over the last year
- Markdown is based on the WYSIWYM document editing principle (like LateX)
  - Instead of micro-managing your document, you use codes to indicate headings, links, or emphasis
  - You type your document and code in a single editor
  - In the end you press a button and it compiles your document for you, running the code and embedding the output along the way
  - Cannot be a disconnect between code, output, and discussion if used properly

Markdown interface

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function | Addins

MarkdownExample.Rmd

Knit | Insert | Run

```
---
title: "Student results analysis for Eleanor Bernard"
author: "Sean van der Merwe"
date: "2019/06/10"
output:
  word_document: default
---

### Links

[Consultation Unit](https://www.ufs.ac.za/natagri/departments-and-divisions/mathematical-statistics-and-actuarial-science-home/statistical-consultation-unit/statistical-consultation-unit)

[Consultant](https://seanvdm.co.za)

[The computer software used is the free program R.](https://en.wikipedia.org/wiki/R_(programming_language))

### Load Data

```{r readdata}
library(openxlsx)
rawdata <- read.xlsx(file.choose())
attach(rawdata)
```

t-tests

In this section we will analyse each set of marks to determine whether the treatment did or did not appear to have an effect. More explicitly, we assume that the treatment had no effect, and then calculate the p-value: the probability of seeing a difference as or more extreme than the one observed, under the null hypothesis of no difference.

For an introduction to this topic see [the Wikipedia page on the t-test](https://en.wikipedia.org/wiki/Student%27s_t-test) and [the corresponding page on the p-value](https://en.wikipedia.org/wiki/P-value) by clicking on these links.

```{r ttests}
t.test(Sem1~Group)
```

Statistical discussion of results

* We assume that the treatment (watching the films) had no effect on student performance. If we see a small p-value we treat that as evidence against our assumption. The weight of that evidence and the practical interpretation is the domain of the principle investigator.
* The average semester 1 mark of the treatment group (`r round(mean(Sem1[Group=='Treatment'],na.rm=T),3)`) is significantly higher (p-value=`r round(t.test(Sem1~Group)$p.value,3)`) than the average semester 1 mark of the control group (`r round(mean(Sem1[Group=='Control'],na.rm=T),3)`).

Please contact me should any further clarity be required.
```

UFS UV

Overview

What about practising simpler problems?

- What if a student just wants to practice the basics?
- Students don't want to go to the library and take out books and figure out which problems are relevant to what they are studying
- And you should see the faces they pull when I suggest they make up their own problems based on the examples I give them, do them, and bring them to me to mark!
- They want to press a button on their phone and get a hit of validation
- So here and there I try to indulge them:
  - https://seanvdm.co.za/shinyapps/hypothesistest/
  - https://seanvdm.co.za/shinyapps/timeseriespractice1/
  - https://seanvdm.co.za/shinyapps/exploredists1/
  - https://seanvdm.co.za/shinyapps/arithmeticgame/

Overview

The point

- At exit level we should try to prepare students for the work environment
- At work our students are going to sit in front of a computer with internet access
- So my assessments take place in front of a computer with internet access
- This lets me raise the cognitive level of the tasks I assign, and shift the focus to problem solving