Tests

Macartan Humphreys

Feb 2017

Tests

In the classical approach to testing a hypothesis we ask:

How likely are we to see data like this if the hypothesis is true?

- If the answer is "not very likely" then we treat the hypothesis as suspect.
- If the answer is *not* "not very likely" then the hypothesis is maintained (some say "accepted" but this is tricky as you may want to "maintain" multiple incompatible hypotheses)

How unlikely is "not very likely"

Weighing Evidence

When we test a hypothesis we decide first on what sort of evidence we need to see in order to decide that the hypothesis is not reliable.

- **Othello** has a hypothesis that Desdemona is innocent.
- **Iago** confronts him with evidence:
- See how she looks at him: would she look a him like that if she were innocent?
- See how she defends him: would she defend him like that if she were innocent?
- See he carries her handkerchief: would he have her handkerchief if she were innocent?
- Othello, the chances of all of these things arising if she were innocent is surely less than 5%

Hypotheses are often rejected, sometimes maintained, but rarely accepted

Note that Othello is focused on the probability of the events if she were innocent but not the probability of the events if Iago were trying to trick him.

He is not directly assessing his belief in whether she is faithful, but rather how likely the data would be if she were faithful.

That is, he assesses:

$$\Pr(\text{Data}|\text{Hypothesis is TRUE})$$

In contrast a "Bayesian" would try to assess:

$$\Pr(\text{Hypothesis is TRUE}|\text{Data})$$

Not Bayes

Note: $Pr(\text{Data}|\text{Hypothesis is TRUE})$ and $Pr(\text{Hypothesis is TRUE}|\text{Data})$ are connected but in a slightly complex way (Bayes Rule):

$$\Pr(\text{H}|\text{D}) = \frac{\Pr(\text{D}|\text{H})\Pr(\text{H})}{\Pr(\text{D}|\text{H})\Pr(\text{H}) + \Pr(\text{D}|\text{NOT H})\Pr(\text{NOT H})}$$

So your belief about the hypothesis should depend not just on the likelihood of seeing the data given the hypothesis but also on your prior belief about how plausible the hypothesis is. But this second part is ignored in classical tests.

Bayes and Frequentist Inference: Example

Imagine a great test:

- If you have a disease then there is a 99% probability that the test will say you have the disease.
- If you do not have a disease then there is a 99% probability that the test will say you do not have the disease.
- 1% of people like you have this disease

The test says you have the disease:

- Will the frequentist reject the null that you have the diseaese?
- Does a Bayesian think you have the disease?

Bayes and Frequentist Inference: Example

Answer with a figure:

```
## Warning: package 'DeclareDesign' was built under R version

## Loading required package: randomizr

## Loading required package: fabricatr

## Loading required package: estimatr

## -- Attaching packages ------------------------------------

## v ggplot2 3.3.3     v purrr   0.3.4
## v tibble  3.1.2     v dplyr   1.0.6
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1

## -- Conflicts --------------------------------------- tid
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

Calculate a p value in your head

- Illustrating p values via "randomization inference"...
- Say you randomized assignment to treatment and your data looked like this.

Unit	1	2	3	4	5	6	7	8	9	10
Treatment	0	0	0	0	0	0	0	1	0	0
Healthy?	3	2	4	6	7	2	4	9	8	2

- Does the treatment improve your health?
- $p =?$

Calculate a p value in your head

- Illustrating p values via "randomization inference"...
- Say you randomized assignment to treatment and your data looked like this.

Unit	1	2	3	4	5	6	7	8	9	10
Treatment	0	0	0	0	0	0	0	0	1	0
Healthy?	3	2	4	6	7	2	4	9	8	2

- Does the treatment improve your health?
- $p =?$