STATS 331

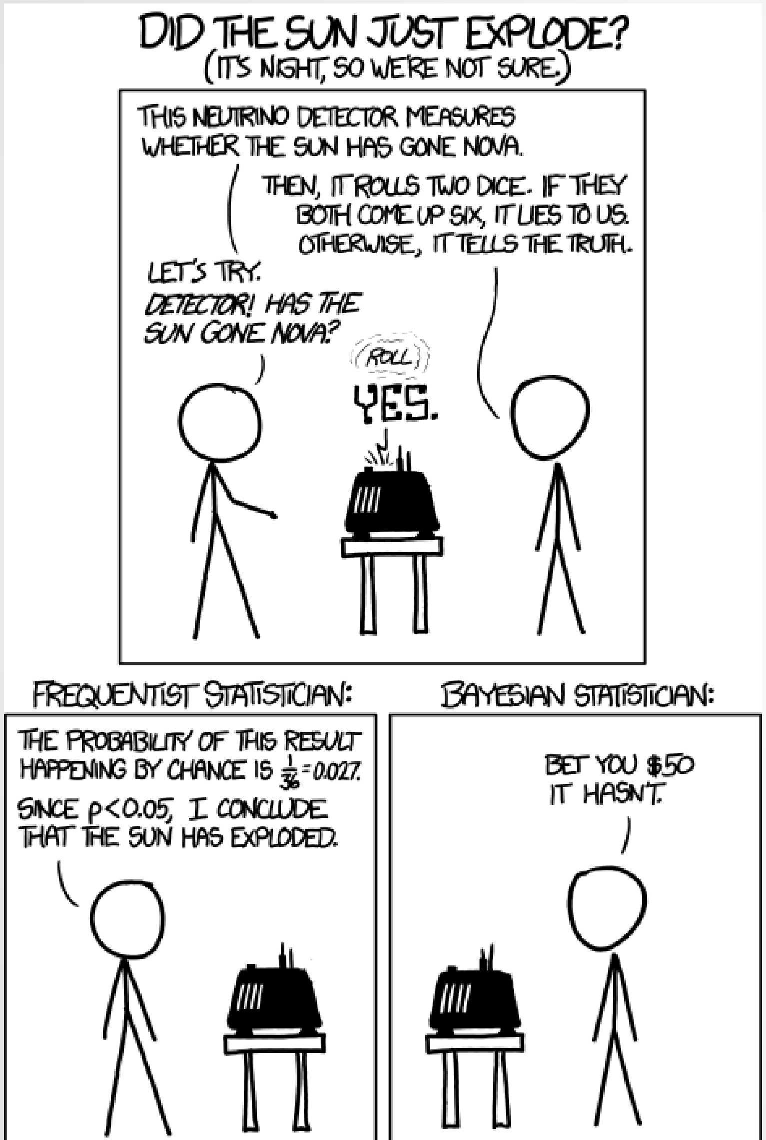


Image Credit: Randall Monroe, xkcd.com

Introduction to Bayesian Statistics

Semester 2, 2016

Hypothesis Testing 331 Style


I CAN'T BELIEVE SCHOOLS ARE STILL TEACHING KIDS ABOUT THE NULL HYPOTHESIS.
I REMEMBER READING A BIG STUDY THAT CONCLUSIVELY DISPROVED IT *YEARS* AGO.

xkcd.com

A Hypothesis Testing Scenario

- What if we considered the two hypotheses

$$\begin{aligned} H_0 &: \quad \theta = 0.5 \\ H_1 &: \quad \theta \neq 0.5 \end{aligned}$$

- There might be a good reason to suspect H_0 might be true. **A precise value**

Bayesian Hypothesis Testing

When we write down a list of hypotheses and get their posterior probabilities, we have tested them.

- In classical statistics, *estimation* and *hypothesis testing* are quite distinct
- For us, they're basically the same thing.

Ed Jaynes

“Any hypothesis testing procedure can be replaced by a parameter estimation procedure that is simpler and more informative”

Frequentist Hypothesis Test

$$H_0 : \quad \theta = 0.5$$
$$H_1 : \quad \theta \neq 0.5$$

- e.g. bus problem from notes. Observed 2 successes out of 5 trials
- P-value = P(x <= 2 or x >= 3 | H_0)

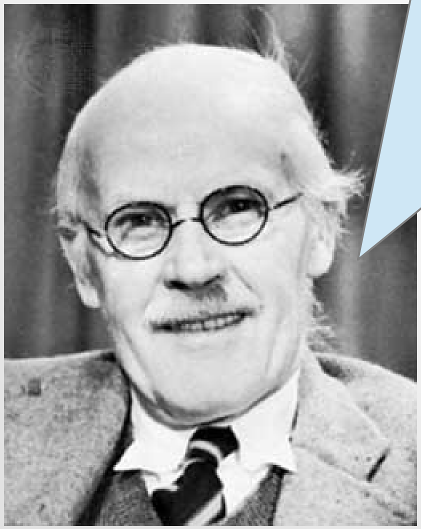
= 1 !!!!

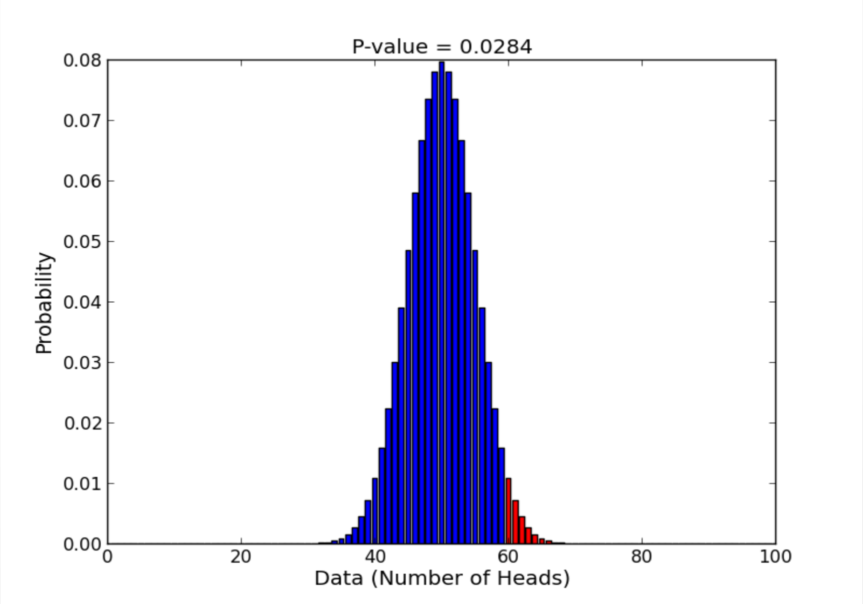
Criticisms of p-values

- P-values are not the probability of the null hypothesis
- There is not even a general relationship between the two!
- What data did you expect to see if the *alternative* was true? P-values ignore this

Another criticism of p-values

A hypothesis may be rejected because it has not predicted observable results that *did not occur!*





Testing Prior

- "Hypothesis testing" singles out a specific value as being special
- **Just give it more prior probability than everything else**

$$
\begin{aligned}
H_0 &: \quad \theta = 0.5 \\
H_1 &: \quad \theta \neq 0.5
\end{aligned}
$$

Bayes Box: “Parameter Estimation”

Possible Values	Prior	Likelihood	Prior x Likelihood	Posterior
0	0.0909			
0.1	0.0909			
0.2	0.0909			
0.3	0.0909			
0.4	0.0909			
0.5	0.0909			
0.6	0.0909			
0.7	0.0909			
0.8	0.0909			
0.9	0.0909			
1	0.0909			
TOTALS	1			

Bayes Box: “Testing Prior”

Possible Values	Prior	Likelihood	Prior x Likelihood	Posterior
0	0.05	0	0	0
0.1	0.05	0.0729	0.003645	0.0163
0.2	0.05	0.2048	0.01024	0.0457
0.3	0.05	0.3087	0.01544	0.0689
0.4	0.05	0.3456	0.01728	0.0772
0.5	**0.5**	**0.3125**	**0.1563**	**0.6977**
0.6	0.05	0.2304	0.0115	0.0514
0.7	0.05	0.1323	0.0662	0.0295
0.8	0.05	0.0512	0.00256	0.0114
0.9	0.05	0.0081	0.000405	0.0018
1	0.05	0	0	0
TOTALS	1		0.224	1

Testing Prior in R

```
theta = seq(0, 1, by=0.1)
prior = rep(0.5/(length(theta) - 1),
                  length(theta))
prior[abs(theta - 0.5) < 1E-6] = 0.5
```

- Make the uniform stuff first, then “fix up” the special hypothesis afterwards

Key Differences from Frequentist

- No p-value. Just get the posterior probability
- Clearer interpretation, and what we really want to know

- Can get evidence **for** the null hypothesis!

Terminology

- Lots of words are in use and we need to know them so we can communicate with people who use them!

Evidence

Odds Ratio

Bayes Factor

Two Hypotheses

- Here is Bayes' Rule for Two Hypotheses
- Can manipulate these and name the parts

$$P(H_0|x) = \frac{P(H_0)p(x|H_0)}{p(x)}$$

$$P(H_1|x) = \frac{P(H_1)p(x|H_1)}{p(x)}$$

Odds Ratio

- One probability divided by another – denominator cancels out
- Prior odds, posterior odds

$$\underbrace{\frac{P(H_0|x)}{P(H_1|x)}}_{\text{Posterior Odds}} = \underbrace{\frac{P(H_0)}{P(H_1)}}_{\text{Prior Odds}} \times \frac{p(x|H_0)}{p(x|H_1)}$$

Odds

- If the odds are “3:1” in favour of H_0
- What does this mean?

$$P(H_0)/P(H_1) = 3$$

$$P(H_0|H_0 \vee H_1) = 3/4$$

Evidence

- Normalising constant
- Prior predictive
- Marginal likelihood
- Evidence

$p(x)$

- All of these names make sense!
- This quantity is related to “model selection” and we will discuss how tomorrow

Bayes Factor

- Likelihood ratio

$$\frac{P(H_0|x)}{P(H_1|x)} = \frac{P(H_0)}{P(H_1)} \times \frac{p(x|H_0)}{p(x|H_1)}$$

Posterior Odds

Prior Odds

Bayes Factor
Likelihood Ratio
Evidence Ratio

First method

- If all of your hypotheses (including “null” and “alternative” if you like to label things that way) are included in your analysis, there is nothing special you have to do.

Example from MacKay

▷ Exercise 3.15.[2, p.63] A statistical statement appeared in *The Guardian* on Friday January 4, 2002:

> When spun on edge 250 times, a Belgian one-euro coin came up heads 140 times and tails 110. 'It looks very suspicious to me', said Barry Blight, a statistics lecturer at the London School of Economics. 'If the coin were unbiased the chance of getting a result as extreme as that would be less than 7%'.

But *do* these data give evidence that the coin is biased rather than fair? [Hint: see equation (3.22).]

Second method

- Can study two different “models” separately.
- Use marginal likelihood within each model as the likelihood for the model as a whole

- Example question from 2014 exam

Warning!

- Bayesian model selection (as this is often called) results can depend strongly on priors.
- I’ll put some examples of this in Assignment 2.