PPLS Summer Training: Intro to Bayesian Estimation

University of Edinburgh

27th June
Things we need to talk about...

- There is a lot to cover when it comes to Bayes
- We will try to touch on the main basics
- What’s different about Bayesian statistics?
- Why you may think you may need it?
- Probability of the event (Bayes vs frequentist)
- Hypothesis Testing
- Bayes Factor
Things we need to talk about...
Want to know the chance of observing head in fair coin toss
Have no idea what it will be
Toss a coin 4 times
Report the results of heads
Repeat
Analyse distribution of successes in four tosses over number of trials (trust Law of Large Numbers)
Results will tell us about the chances of observing the heads in a coin toss
Easy, right? (try here)
How do you know if something is probable?

Figure: Could be a joke, could be not
In Bayesian view, we start with some belief about the process first and compared to frequentist approach we have multiple components to consider:

- We often start with a belief about the event - prior distribution
- We also have a likelihood that our observation can be true given the model set up - likelihood
- Belief can be changed in light of evidence - data
- Combine belief and likelihood - posterior probability distribution
When we should prioritise Bayes

- In cases where the cost of making a Type 2 error is really high!
- Common examples often found in health research but can equally be observed in any discipline
- Bayes Analysis on small samples may provide a better picture with respect to uncertainty in your results
In the center of the setting we have a rule of conditional probability (Bayes Rule):

\[ P(A|B) = \frac{P(B|A)P(A)}{P(B)} \]  

(1)
Imagine that we have created a new test to check whether someone is at risk of getting a flu in the coming year - so can be vaccinated earlier for prevention.

- We need to specify possible outcomes to consider:
  - $P(\text{tested to have a flu—actually will have one}) = 0.9 = 90\%$
  - $P(\text{tested to not have a flu—will not have one}) = 0.95 = 95\%$
  - leaves us with complementary estimates for false positives and false negatives:
    - $P(\text{tested to have a flu—actually not have one}) = 0.1 = 10\%$
    - $P(\text{tested to not have a flu—will have one}) = 0.05 = 5\%$
We can also check what is the average chance of getting a flu at least once a year in our population, say we know that 1 in 30 is likely to get a flu:

$$P(\text{flu}) = \frac{1}{30} = 0.33$$

(2)
Let's work with an example

Using Bayes rule we can now calculate the probability of person having a flu given that the test indicated that they will have one.

\[
P(\text{have a flu|test positive}) = \frac{P(\text{Having a flu})(\text{Test positive})}{P(\text{Test positive})}
\]

(3)

Start with the numerator where joint probability:
\[P(\text{Having a flu and Test positive}) = 0.33 \times 0.9 = 0.27\]

denominator:
\[P(\text{Test positive|actually will have one}) \text{and}(\text{Test positive|will not have one}) = 0.27 + 0.034 = 0.30\]

We need quickly to show that:
\[P(\text{Test positive|will note have a flu}) = (1 - P(\text{tested to have a flu}) \times (1 - P(\text{Test Negative|will not have flu}) = (1 - 0.33) \times (1 - 0.95) = 0.034\]
We can now put everything together:

\[
P(\text{have a flu} | \text{test positive}) = \frac{P(\text{Having a flu})(\text{Test positive})}{P(\text{Test positive})} = \frac{0.27}{0.34} = 0.9
\] (4)

What we find is that even if the test returned positive, the actual probability of the flu without doing anything (0.9) or 90%, while still high number it includes more information about false negative rates. So, we can see that even if tested positive, the result is not a guarantee of actually having a flu. The result is dependent on our prior information about flu in population. Varying that number will provide you with different conclusions.
Updating your priors - beliefs

We can update the probability of the event to happen by retaking the test. For instance I can take my patients twice and only if the test shown the same results on the 2nd time I will then have more certainty about the process. Remember that we started:

\[ P(flu) = \frac{1}{30} = 0.33 \]  \hspace{1cm} (5)

Our results in the previous calculations lead us to:

\[
P(\text{have a flu|test positive}) = \frac{P(\text{Having a flu})(\text{Test positive})}{P(\text{Test positive})} = \frac{0.27}{0.34} = 0.9
\]  \hspace{1cm} (6)

We can use Bayes rule again to test what is the probability of the person having a flu and test coming positive in the 2nd round.
We can use Bayes rule again to test what is the probability of the person having a flu and test coming positive in the 2nd round:

$$P(\text{Having a flu} | \text{2nd test positive}) =$$

$$\frac{P(\text{Tested to have flu}) \times P(\text{2nd test positive} | \text{Will have a flu})}{P(\text{2nd test positive})}$$
Study: A total of 60 females came to the GP to ask for pain relief for headache. They were randomly assigned to ibuprofen (treatment) or placebo (control), 30 in each group and were asked to report how they feel two hours later. In the treatment group, 8 out of 30 reported no change in how they felt with their headache still being there within next two hours. In the control group, no change was reported for 20 out of 30 women.
Therefore, we can form the hypotheses as below:

- $H_0 = 0.5$ (equal chance - no difference to the state which is equally likely to come from the treatment or control group)
- $H_1 < 0.5$ (in treatment group there is less chance to observe no change, there was an effect of treatment)

Using known to us approach, we will be looking at p values, lower the p-value - we will conclude that what we observe is unlikely to be found under the null.
Inference for Proportions - simple analysis

- I can use Bernoulli distribution to find out.
- If we know that the probability of observing change/no change in women state is 0.5, we can calculate the p-value from 30 independent Bernoulli trials where the probability of success is 0.5.
- Remember that we observed 8 out of 30. The probability of observing 8 (no change) cases in 30 trials with probability of (no change ) = 0.5, will be: 0.008
Inference for Proportions - simple analysis

THIS IS WHAT P < .05 FEELS LIKE
In Bayesian setting we will need to think a bit more. We will use:

- Prior
- Likelihood - $P(data|model)$
- Posterior (Bayes Rule) - $\frac{P(model) \times P(data|model)}{P(data)}$
Updating your priors - beliefs

Figure: Illustration of prior, likelihood and posterior distribution for $\beta(10, 10)$ distributed prior
Elicitation

- If you are a Bayesian, you express your beliefs and uncertainty through probability distributions.
- You can self-elicit one based on what you know.
- You then should update your beliefs in the light of new data.
- Importantly, no matter which distribution you choose with enough data you will converge to an accurate posterior distribution. That means that if you and your friend started with a very different priors with more evidence observed - both of you will get to the same distribution.
You should now by now pretty well the structure of hypothesis testing in frequentist approach.

In Bayesian setting, when dealing with two competing hypothesis, we first want to find the posterior probability of $H_1$ given the data and then compare it with $H_2$ given the data.

Once you know the posterior of each you can just pick one with higher probability

There is however an extra thing to consider

Decision you made and loss associated with that

You can incorporate these in Bayesian setting too
Posterior odds and Bayes factor

If there are two competing hypotheses being considered, then the prior odds of $H_1$ to $H_2$ can be defined as:

$$O[H_1 : H_2] = \frac{P(H_1)}{P(H_2)} \quad (9)$$

The posterior odds should then be defined as:

$$PO[H_1 : H_2] = \frac{P(H_1|Data)}{P(H_2|Data)} \quad (10)$$

this is equivalent to:

$$PO[H_1 : H_2] = \frac{P(Data|H_1)}{P(Data|H_2)} \cdot \frac{P(H_1)}{P(H_2)} \quad (11)$$

This is also known as:

$$PO[H_1 : H_2] = Bayesfactor \cdot PriorOdds \quad (12)$$

Note that Bayes factor is the ratio of two likelihoods.
Posterior odds and Bayes factor (Quick Example)

- $H_1$: Patient has a flu
- $H_2$: Patient doesn’t have a flu

We know that probability of a flu is 1 in 30 = 0.33, hence we have the following priors:

- $H_1$: 0.33
- $H_2$: 0.67
The prior odds are:

\[
O[H_1 : H_2] = \frac{P(H_1)}{P(H_2)} = \frac{0.33}{0.67} = 0.49
\]  

(13)

We have tested a sample of patients and found that probability to get a positive test results and have a flu was 0.9 with a complement of 0.1. So our posteriors are:

- \(H_1|+: 0.9\)
- \(H_2|+: 0.1\)

Which gives us posterior odds:

\[
PO[H_1 : H_2] = \frac{P(H_1|Data)}{P(H_2|Data)} = \frac{0.9}{0.1} = 9
\]

(14)
The Bayes factor:

\[ BF[H_1 : H_2] = \frac{PO[H_1 : H_2]}{O[H_1 : H_2]} = \frac{9}{0.49} = 18.36 \] \hspace{1cm} (15)

The general rule for interpretation is:

<table>
<thead>
<tr>
<th>BF[H_1 : H_2]</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-3</td>
<td>Nor worth a mention</td>
</tr>
<tr>
<td>3-20</td>
<td>Positive support for ( H_2 )</td>
</tr>
<tr>
<td>20-150</td>
<td>Strong support for ( H_2 )</td>
</tr>
<tr>
<td>&gt; 150</td>
<td>A very strong support for ( H_2 )</td>
</tr>
</tbody>
</table>
An Overview and Conclusions

- **p values** - how to be pragmatic about them?
- Can we do better with Bayes?
- The answer depends on the sample and uncertainty too
- In large enough sample we can minimise uncertainty in both frequentist and Bayesian setting
Stay pragmatic

Figure: The balance between complexity and interpretability
An Overview and Conclusions

- If everything is fine as it is - no need for Bayesian analysis
- It can be computationally demanding
- If you are interested in uncertainty - go for it
- Otherwise, always think what was the whole point to start
Examples time (let’s get back to R)

But first...

Any questions?

Let’s take a break now and come back in 10 min!