STAT 425: Introduction to Bayesian Analysis

Marina Vannucci

Rice University, USA

Fall 2017
Lecture 8: Bayes Factors

- Bayes factors for hypothesis testing
Frequentist hypothesis testing

In classical hypothesis testing, we proceed as follows:

1. state a null hypothesis, $H_0$, and an alternative hypothesis, $H_1$
2. determine an appropriate test statistic, $T(X)$
3. compute the $p$-value of the test as

\[
 p\text{-value} = P(T(X) \text{ more “extreme” than } T(x_{obs}) | \theta, H_0) 
\]

where “extremeness” in in the direction of $H_1$

4. if the $p$-value is less than the prespecified Type I error rate, $\alpha$, $H_0$ is rejected

Straightforward only when the two hypotheses are nested.
In Bayesian hypothesis testing, we proceed as follows:

1. state the two hypotheses, \( M_1 \) and \( M_2 \)
2. assign priors to \( \theta \) under \( M_1 \) and \( M_2 \), and specify \( p(M_1) \) and \( p(M_2) \)
3. evaluate \( P(M_1|x) \) and \( p(M_2|x) \) via Bayes’ theorem
4. compute the Bayes factor to assess the evidence in favor of \( M_1 \) as the ratio of the posterior odds vs the prior odds:

\[
BF = \frac{P(M_1|x)/P(M_2|x)}{P(M_1)/P(M_2)} = \frac{p(x|M_1)}{p(x|M_2)}.
\]

Does not require the two models to be nested. Can also be written as the ratio of the marginal densities for the two models.
The Bayes factor can also be written as the ratio of the observed marginal densities for the two models (via Bayes theorem)

$$BF = \frac{p(M_1|x)/p(M_2|x)}{p(M_1)/p(M_2)} = \frac{\left[ \frac{p(x|M_1)p(M_1)}{p(x)} \right]}{\left[ \frac{p(x|M_2)p(M_2)}{p(x)} \right]} / \frac{p(M_1)/p(M_2)}{p(M_1)/p(M_2)}$$

$$= \frac{p(x|M_1)}{p(x|M_2)}$$

The marginal distribution of $x$ under each model $M_i$ is

$$p(x|M_i) = \int f(x|\theta_i, M_i) \pi_i(\theta_i) \, d\theta_i, \quad i = 1, 2$$

In essence, how likely the data are, based on each model and integrating over the uncertainty in the parameters as represented by the prior.

The Bayes factor is only defined when the marginal density of $x$ under each model is proper. If $\pi_i(\theta_i)$ is improper, then $p(x|M_i)$ will necessarily be improper, and the Bayes factor is not defined.
Interpretation of Bayes factor

### Jeffreys’ – scale of evidence in favor of $M_1$

<table>
<thead>
<tr>
<th>$\log_{10} BF$</th>
<th>Bayes factor</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 – 0.5</td>
<td>$1 \leq BF \leq 3.2$</td>
<td>weak</td>
</tr>
<tr>
<td>0.5 – 1.0</td>
<td>$3.2 &lt; BF \leq 10$</td>
<td>substantial</td>
</tr>
<tr>
<td>1.0 – 2.0</td>
<td>$10 &lt; BF \leq 100$</td>
<td>strong</td>
</tr>
<tr>
<td>$&gt; 2$</td>
<td>$BF &gt; 100$</td>
<td>decisive</td>
</tr>
</tbody>
</table>

### Kass & Raftery – scale of evidence in favor of $M_1$

<table>
<thead>
<tr>
<th>$2 \ln BF$</th>
<th>Bayes factor</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 – 2</td>
<td>$1 \leq BF \leq 3$</td>
<td>weak</td>
</tr>
<tr>
<td>2 – 6</td>
<td>$3 &lt; BF \leq 20$</td>
<td>positive</td>
</tr>
<tr>
<td>6 – 10</td>
<td>$20 &lt; BF \leq 150$</td>
<td>strong</td>
</tr>
<tr>
<td>$&gt; 10$</td>
<td>$BF &gt; 150$</td>
<td>very strong</td>
</tr>
</tbody>
</table>
Example: Test of proportion

Suppose 16 customers have been recruited by a fast-food chain to compare two types of ground beef patty on the basis of flavor. All of the patties to be evaluated have been kept frozen for eight months.

- One set of 16 has been stored in a high-quality freezer that maintains a temperature that is consistently within $\pm 1^\circ F$.
- The other set of 16 has been stored in a freezer with temperature that varies anywhere between 0 and $15^\circ F$.

The food chain executives are interested in whether storage in the higher-quality freezer translates into a substantial improvement in taste, thus justifying the extra effort and cost associated with equipping all of their stores with these freezers.

Suppose that to be regarded as “substantial” improvement more than 60% of consumers must prefer the more expensive option. 13 of the 16 consumers state a preference for the more expensive patty.
Let $Y_i = 1$ if consumer $i$ states a preference for the more expensive patty and $Y_i = 0$ otherwise.

$$X = \sum_{i=1}^{16} Y_i \sim \text{Binomial}(16, \theta)$$

We want to test:

$$M_1 : \theta > 0.6 \quad \text{vs} \quad M_2 : \theta \leq 0.6$$

Suppose we consider “minimally informative” priors, $\pi(\theta)$:

- Jeffreys’ prior, Beta(.5, .5)
- a prior that we think of as “noninformative”, Beta(1, 1)
- Beta(2, 2) prior
The posterior distribution for $\theta$ is given by

$$\theta|x \sim \text{Beta}(\alpha + x, \beta + n - x)$$

| Prior     | Posterior quantile | $p(\theta > 0.6|x)$ | BF    |
|-----------|--------------------|---------------------|-------|
| Beta(.5, .5) | 0.579   0.806  0.944 | 0.964              | 34.432|
| Beta(1, 1)  | 0.566   0.788  0.932 | 0.954              | 30.812|
| Beta(2, 2)  | 0.544   0.758  0.909 | 0.930              | 24.604|

Strong evidence in favor of $M_1 : \theta > 0.6$. 