Bayesian inference and hypothesis testinig

David J. H. Shih

Intended learning outcomes

- Describe Bayesian inference and hypothesis testing
- Apply Bayesian inference to data and interpret the results

p-value

Assuming the null hypothesis H_0,

$$X \sim F_X,$$

then we can caculate the test statistic $T = g(X)$ from the data X, and derive the cumulative distribution function of T under H_0.

$$F_T(t) = P\left(T \leq t \mid H_0\right)$$

The (right-sided) *p*-value is

$$p = P\left(T > t \mid H_0\right) = 1 - F_T(t).$$

If we want a left-sided hypothesis, we can calculate

$$p = P\left(T \leq t \mid H_0\right) = F_T(t).$$

If we want a double-sided hypothesis, provided that f_T is even, we can calculate

$$p = P\left(|T| > |t| \mid H_0\right) = 2(1 - F_T(t)).$$

Under classical hypothesis testing, we cannot calculate the probability that the hypothesis is true!

Bayesian hypothesis testing

Given a discrete hypothesis H and observed data x,

$$p(H \mid x) = \frac{p(x \mid H)\, p(H)}{p(x)}.$$

Example: Medical testing

After observe test result x,

$$P(H = 1 \mid X = x) = \frac{P(X = x \mid H = 1)P(H = 1)}{P(X = x)},$$

where $H = 1$ indicates disease and $H = 0$ no disease.

Short-hand:

$$P(H_1 \mid x) = \frac{P(x \mid H_1)P(H_1)}{P(x)},$$

where H_1 is the event (hypothesis) that $H = 1$.

Posterior odds ratio

Definition

The posterior odds ratio for a hypothesis H is

$$\frac{P(H = h \mid x)}{P(H \neq h \mid x)} = \frac{P(x \mid H = h)}{P(x \mid H \neq h)} \, \frac{P(H = h)}{P(H \neq h)},$$

which follows from Bayes' theorem. First term is Bayes' factor. Second term is prior odds ratio.

Shorthand:
If $H \in \{0, 1\}$, we can write

$$\frac{P(H_1 \mid x)}{P(H_0 \mid x)} = \frac{P(x \mid H_1)}{P(x \mid H_0)} \, \frac{P(H_1)}{P(H_0)},$$

where H_h is the event (hypothesis) that $H = h$.

Direct inference on posterior of a parameter

For a parameter θ with posterior $p(\theta \mid x)$, we can simply calculate

$$P(\theta > c \mid x) \quad \text{or} \quad P(\theta < c \mid x),$$

for some threshold c.

We can also calculate

$$P(c_1 < \theta < c_2 \mid x).$$

If θ is discrete, we can calculate

$$P(\theta = c \mid x).$$

Bayesian hypothesis testing on a parameter

Greater-than hypothesis

$$H_1 : \theta > c, \qquad H_0 : \theta \leq c.$$

Less-than hypothesis

$$H_1 : \theta < c, \qquad H_0 : \theta \geq c.$$

Equality hypothesis

$$H_1 : \theta = c, \qquad H_0 :\neq c,$$

provided that θ is discrete.

Within-interval hypothesis

$$H_1 : \theta \in (c_1, c_2), \qquad H_0 : \theta \notin (c_1, c_2).$$

Beyond-interval hypothesis

$$H_1 : \theta \notin (c_1, c_2), \qquad H_0 : \theta \in (c_1, c_2).$$

Bayesian hypothesis testing on a model

For a model $m \in \mathcal{M}$ applied to observed data x,

$$p(m \mid x) = \frac{p(x \mid m)\, p(m)}{p(x)}.$$

Example

m_0 is a model with no unknown parameter.
m_1 is a model with unknown parameter θ.
m_2 is a model with unknown parameters θ and ϕ.

Defining the prior

The definition of prior affects the posterior.

We can use

- prior knowledge
- prior information
- objective prior

Or, avoid the defining the prior by using Bayes' factor for inference:

$$\frac{P(x \mid H_1)}{P(x \mid H_0)} > c,$$

for some threshold c.

Summary

“The value for which P=0.05, or 1 in 20, is 1.96 or nearly 2; it is convenient to take this point as a limit in judging whether a deviation ought to be considered significant or not.” - Ronald Fisher

Blais 2014, chapter 8. *Remark*: Classical hypothesis testing is described in this chapter, which compliments the Bayesian hypothesis testing described in the lecture.

Intended learning outcomes

- Describe Bayesian inference and hypothesis testing
- Apply Bayesian inference to data and interpret the results