STAT201B Lecture 16 Bayes Factor

Logic

在 Bayesian analysis 中, hypothesis 也可以像 parameters 一样被 probability distributions 描述

1 Baye Factor

Logic

我们首先考虑最简单的情况: hypothesis 用于描述 θ 所落入的区间, 即

$$H_i : \theta \in \Theta_i$$

例如:

- θ 的 prior 为: $\mathrm{Unif}[0,1]$
- 检验 $H_0 : \theta \leq \frac{1}{2} \quad v.s. \quad H_1 : \theta > \frac{1}{2}$

1.1 Hypothesis 的 posterior (simple case)

若

- hypothesis 为 $H_i : \theta \in \Theta_i$,
- θ 的 prior 为 $f(\theta)$

则:

- Prior probability 为:

$$\mathbb{P}(H_i) = \mathbb{P}(\theta \in \Theta_i) = \int_{\Theta_i} f(\theta)\, d\theta$$

- Posterior probability 为:

$$\mathbb{P}(H_i|x^n) = \mathbb{P}(\theta \in \Theta_i|x^n) = \int_{\Theta_i} f(\theta|x^n)\, d\theta$$

Logic

接下来我们考虑更复杂的情况, 即

- 存在 K 个 hypotheses
- Hypotheses 不再仅用于描述 θ 落入的区间 (且 θ 在不同 hypothesis 下的 prior 可能不同)
- Hypotheses 也存在 prior, 且可能相互间不相等

例如:

- Independent samples X 和 Y 满足: $X|p_1 \sim Bin(n, p_1)$, $Y|p_2 \sim Bin(n, p_2)$
- Null hypothesis 和 alternative hypothesis 分别为:

- $H_0 : p_1 = p_2 = p,\ p \sim \text{Unif}(0,1)$
- $H_1 : p_1 \neq p_2,\ p_1 \sim \text{Unif}(0,1),\ p_2 \sim \text{Unif}(0,1)$

1.2 Hypothesis 的 posterior (general cases)

若:

- 存在 K 个 hypotheses $H_0, \dots, H_{K-1}$
- H_k 的 prior 为 $\mathbb{P}(H_k)$ (通常各个 prior 的先验概率都相等)

则 hypothesis H_k 的 posterior 为:

$$\mathbb{P}(H_k|x^n) = \frac{f(x^n|H_k)\,\mathbb{P}(H_k)}{\sum_{k=1}^{K} f(x^n|H_k)\,\mathbb{P}(H_k)}$$

Remark: H_k 的 priors 均相等的情况

若 H_k 的 priors 均相等, 则 hypothesis H_k 的 posterior 为:

$$\mathbb{P}(H_k|x^n) = \frac{f(x^n|H_k)}{\sum_{k=1}^{K} f(x^n|H_k)}$$

Logic

通常我们会比较两个 hypotheses H_i 和 H_j, 此时可以考虑使用 posterior odds

1.3 Definition: Posterior odds 和 Bayes factor

若:

- 存在 K 个 hypotheses $H_0, \dots, H_{K-1}$
- H_k 的 prior 为 $\mathbb{P}(H_k)$ (通常各个 prior 的先验概率都相等)

则 the **posterior odds** of H_i relative to H_j 为:

$$\frac{\mathbb{P}(H_i|x^n)}{\mathbb{P}(H_j|x^n)} = \frac{f(x^n|H_i)}{f(x^n|H_j)} \times \frac{\mathbb{P}(H_i)}{\mathbb{P}(H_j)}$$

其中, **Bayes factor** for comparing H_i to H_j 为:

$$BF_{ij} = \frac{f(x^n|H_i)}{f(x^n|H_j)}$$

Remark: H_k 的 priors 均相等的情况

若 H_k 的 priors 均相等, 则 posterior odds 等于 Bayes factor, 即

$$\frac{\mathbb{P}(H_i|x^n)}{\mathbb{P}(H_j|x^n)} = \frac{f(x^n|H_i)}{f(x^n|H_j)} = BF_{ij}$$

Logic

可以看到, Bayes factor 可以用于辅助在两个 hypotheses 之间做选择, 接下来我们研究如何计算 Bayes factor

1.4 Bayes factor 的计算

情况一: Simple hypothesis testing

若 hypothesis 为两点检验, 即:

$$H_i : \theta = \theta_i, \quad v.s. \quad H_j : \theta = \theta_j$$

则 Bayes factor 为 likelihood ratio, 即:

$$BF_{ij} = \frac{f(x^n|\theta_i)}{f(x^n|\theta_j)}$$

情况二: Hypothesis 被用于描述 region of the parameter space

若 Hypothesis 被用于描述 region of the parameter space, 即:

$$H_i : \theta \in \Theta_i, \quad v.s. \quad H_j : \theta \in \Theta_j$$

则 Bayes factor 可以通过计算 prior 和 posterior odds 来推出, 即:

$$BF_{ij} = \frac{\mathbb{P}(H_i|x^n)}{\mathbb{P}(H_j|x^n)} \Big/ \frac{\mathbb{P}(H_i)}{\mathbb{P}(H_j)} = \frac{\int_{\Theta_i} f(\theta|x^n)\, d\theta}{\int_{\Theta_j} f(\theta|x^n)\, d\theta} \Big/ \frac{\mathbb{P}(H_i)}{\mathbb{P}(H_j)}$$

情况三: General cases

对于 general cases, Bayes factor 可以通过定义式计算, 即:

$$BF_{ij} = \frac{f(x^n|H_i)}{f(x^n|H_j)}$$

其中

$$f(x^n|H_i) = \int_\Theta f(x^n|\theta, H_i) f(\theta|H_i)\, d\theta$$

被称为 marginal likelihood

Remark

若 $f(\theta|H_i)$ 为 conjugate, 则可以计算出 $f(x^n|H_i)$ 的 closed form, 否则可能需要使用 sampling (如 MC integration, with sampling from $f(\theta|H_i)$) 进行 approximation

Example

情境:

Example: Albert Pujols (St. Louis Cardinals) was voted the "most feared hitter in baseball" in 2010. However, Ichiro Suzuki (Seattle Mariners) has a very similar batting average. Here are their career statistics from 2001 to 2010, when they both played major league baseball.

Pujols: 5146 at bats; 1717 hits
Suzuki: 6099 at bats; 2030 hits

If we consider that each player has a "true" batting average p, around which their actual batting average fluctuates, we might be interested in looking at evidence for/against the hypothesis $p_{Pujols} = p_{Suzuki}$.

问题:

对于上述情境, 考虑以下问题:

令:

- $X|p_1 \sim Bin(n, p_1)$, $Y|p_2 \sim Bin(m, p_2)$

- $H_0: p_1 = p_2 = p,\ p \sim \text{Unif}(0,1)$
- $H_1: p_1 \neq p_2,\ p_1 \sim \text{Unif}(0,1),\ p_2 \sim \text{Unif}(0,1)$

求:

1. $f(x, y|H_0)$
2. $f(x, y|H_1)$
3. Bayes factor BF_{10}

解答:

Question 1:

$$
\begin{aligned}
f(x,y|H_0) &= \int_0^1 f(x,y,p|H_0)\,dp \\
&= \int_0^1 f(x|p,H_0)\cdot f(y|p,H_0)\cdot f(p|H_0)\,dp \\
&= \int_0^1 \frac{n!}{x!(n-x)!}\cdot p^x\cdot(1-p)^{n-x}\cdot\frac{m!}{y!(m-y)!}\cdot p^y\cdot(1-p)^{m-y}\cdot 1\,dp \\
&= \int_0^1 \frac{n!}{x!(n-x)!}\cdot\frac{m!}{y!(m-y)!}\cdot p^{x+y}\cdot(1-p)^{n+m-x-y}\,dp \\
&= \frac{n!}{x!(n-x)!}\cdot\frac{m!}{y!(m-y)!}\cdot\frac{\Gamma(x+y+1)\Gamma(n+m-x-y+1)}{\Gamma(n+m+2)}\cdot\int_0^1\frac{\Gamma(n+m+2)}{\Gamma(x+y+1)\Gamma(n+m-x-y+1)}\cdot p^x \\
&= \frac{n!}{x!(n-x)!}\cdot\frac{m!}{y!(m-y)!}\cdot\frac{\Gamma(x+y+1)\Gamma(n+m-x-y+1)}{\Gamma(n+m+2)} \\
&= \frac{\binom{n}{x}\binom{m}{y}}{\binom{n+m}{x+y}}\cdot\frac{1}{m+n+1}
\end{aligned}
$$

Question 2:

$$
\begin{aligned}
f(x,y|H_1) &= \int_0^1\int_0^1 f(x,y,p_1,p_2|H_1)\,dp_1\,dp_2 \\
&= \int_0^1\int_0^1 f(x|p_1,H_1)\cdot f(p_1|H_1)\cdot f(y|p_2,H_1)\cdot f(p_2|H_1)\,dp_1\,dp_2 \\
&= \int_0^1 f(x|p_1,H_1)\cdot f(p_1|H_1)\,dp_1\cdot\int_0^1 f(y|p_2,H_1)\cdot f(p_2|H_1)\,dp_2
\end{aligned}
$$

其中,

$$
\begin{aligned}
\int_0^1 f(x|p_1,H_1)\cdot f(p_1|H_1)\,dp_1 &= \int_0^1\frac{n!}{x!(n-x)!}\cdot p_1^x\cdot(1-p_1)^{n-x}\,dp_1 \\
&= \frac{n!}{x!(n-x)!}\frac{\Gamma(x+1)\Gamma(n-x+1)}{\Gamma(n+2)}\int_0^1\frac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n-x+1)}\cdot p_1^x\cdot(1-p_1)^{n-x}\,dp_1 \\
&= \frac{n!}{x!(n-x)!}\frac{\Gamma(x+1)\Gamma(n-x+1)}{\Gamma(n+2)} \\
&= \frac{1}{n+1}
\end{aligned}
$$

类似地,

$$
\int_0^1 f(y|p_2,H_1)\cdot f(p_2|H_1)\,dp_2 = \frac{1}{m+1}
$$

因此,

$$
f(x,y|H_1) = \frac{1}{(n+1)(m+1)}
$$

Question 3:

$$
BF_{10} = \frac{f(x,y|H_1)}{f(x,y|H_0)} = \frac{(n+m+1)\binom{n+m}{x+y}}{(n+1)(m+1)\binom{n}{x}\binom{m}{y}}
$$

1.5 对 Bayes factor 的 interpretation

Hypotheses 没有 prior 的情况:

此时 Bayes factor 反映了相较于 H_0, 数据对于 H_1 的支撑强度:

$\log_{10}(BF_{10})$	BF_{10}	**证据强度 (Evidence)**
0–0.5	1–3.2	Weak（弱）
0.5–1	3.2–10	Moderate（中等）
1–2	10–100	Strong（强）
> 2	> 100	Decisive（决定性）

Hypotheses 有 prior 的情况:

令:

- H_1 的 prior 为 $p = \mathbb{P}(H_1)$
- H_1 的 posterior 表示为 $p^* = \mathbb{P}(H_1|Data)$

则可以求出 posterior 为:

$$p^* = \frac{\frac{p}{1-p}BF_{10}}{1 + \frac{p}{1-p}BF_{10}}$$