



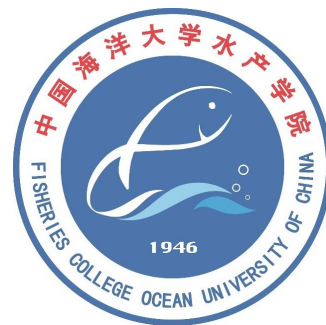
贝叶斯数据分析

贝叶斯数据分析笔记

作者：韩方成

组织：中国海洋大学水产学院

时间：Dec 5, 2021



别回头，跟着光

目录

1	概率和推理	1
1.1	贝叶斯数据分析的三大步骤	1
1.2	贝叶斯推理	1
1.2.1	概率符号	1
1.2.2	贝叶斯法则	1
1.2.3	预测	1
1.2.4	似然和比率	1
1.3	概率论中的一些有用的结果	2
1.3.1	变量的变换	2
A	Gamma,Beta 分布	3
A.1	Gamma 函数	3

第 1 章 概率和推理

1.1 贝叶斯数据分析的三大步骤

贝叶斯数据分析的过程可以被分为一下三步：

1. 建立全概率模型，即一个问题中所有可观测和不可观测的量的联合概率密度。
2. 计算后验概率：在给定已知数据下未知参数的分布。
3. 评估模型的拟合程度。

1.2 贝叶斯推理

1.2.1 概率符号

我们定义变异系数为 $\text{sd}(\theta)/E(\theta)$ ，几何平均数为 $\exp(E[\log(\theta)])$ ，并且几何标准差为 $\exp(\text{sd}[\log(\theta)])$ 。

1.2.2 贝叶斯法则

贝叶斯法则可以写为：

$$p(\theta|y) = \frac{p(\theta, y)}{p(y)} = \frac{p(\theta)p(y|\theta)}{p(y)} = \frac{p(\theta)p(y|\theta)}{\int_{\theta} p(\theta)p(y|\theta)d\theta} \quad (1.1)$$

因为在给定 y 的条件下 $p(y)$ 为常数，与 θ 无关，所以我们常将1.1写为式1.2所示的非标准化后验分布：

$$p(\theta|y) \propto p(\theta)p(y|\theta) \quad (1.2)$$

1.2.3 预测

为了对一个未知的可观测数据做预测，即预测推断。在数据 y 被考虑之前，未知但是可观测的 y 的分布为

$$p(y) = \int p(y, \theta)d\theta = \int p(\theta)p(y|\theta)d\theta \quad (1.3)$$

这被称为先验预测分布：先验是因为它没有以任何先前的观测为条件，预测是因为他是一个可观测数据的分布。

当数据 y 被观测到后，我们可以用同样的方式来预测一个未知可观测的 \tilde{y} 。 \tilde{y} 的分布被称为后验预测分布，后验是因为它以观测 y 为条件，预测是因为他是可观测数据 \tilde{y} 的预测：

$$\begin{aligned} p(\tilde{y}|y) &= \int p(\tilde{y}, \theta|y)d\theta \\ &= \int p(\tilde{y}|\theta, y)p(\theta|y)d\theta \\ &= \int p(\tilde{y}|\theta)p(\theta|y)d\theta \end{aligned} \quad (1.4)$$

1.2.4 似然和比率

在给定模型下在点 θ_1 和 θ_2 的后验密度 $p(\theta|y)$ 比值称为 θ_1 相对于 θ_2 的后验几率：

$$\frac{p(\theta_1|y)}{p(\theta_2|y)} = \frac{p(\theta_1)p(y|\theta_1)/p(y)}{p(\theta_2)p(y|\theta_2)/p(y)} = \frac{p(\theta_1)}{p(\theta_2)} \frac{p(y|\theta_1)}{p(y|\theta_2)} \quad (1.5)$$

由此可以看出后验比率是先验比率和似然比例的乘积。

1.3 概率论中的一些有用的结果

我们经常用下列方式表示随机变量 u 的期望:

$$E(u) = E(E(u|v)) \quad (1.6)$$

公式1.6很好证明:

$$\iint up(u, v)dudv = \iint up(u|v)dup(v)dv = \int E(u|v)p(v)dv \quad (1.7)$$

方差包括两项:

$$\text{var}(u) = E(\text{var}(u|v)) + \text{var}(E(u|v)) \quad (1.8)$$

式1.8的证明为:

$$\begin{aligned} E(\text{var}(u|v)) + \text{var}(E(u|v)) &= E(E(u^2|v) - (E(u|v))^2) + E((E(u|v))^2 - (E(E(u|v)))^2) \\ &= E(u^2) - E((E(u|v))^2) + E((E(u|v))^2) - (E(u))^2 \\ &= E(u^2) - (E(u))^2 \\ &= \text{var}(u) \end{aligned}$$

1.3.1 变量的变换

如果 p_u 为一个离散分布, 并且 f 为一个一对一的函数, 而 $v = f(u)$, 则 v 的概率密度为

$$p_v(v) = p_u(f^{-1}(v))$$

如果 f 是一个多对一函数, 那么会对 $p_v(v)$ 有一个求和。

如果 p_u 为一个连续分布且 $v = f(u)$ 是一个一对一的函数, 则

$$p_v(v) = |J|p_u(f^{-1}(v))$$

其中 $|J|$ 为变换 $u = f^{-1}(v)$ 的雅可比行列式, 如果 f 为多对一函数, 则 $p_v(v)$ 上有积分或求和项。

当处理一维变量时, 我们经常利用对数函数来将参数空间从 $(0, \infty)$ 转换到 $(-\infty, \infty)$ 。当处理唯一开单位区间 $(0, 1)$ 上的参数时, 我们用 logistic 变换:

$$\text{logit}(u) = \log\left(\frac{u}{1-u}\right) \quad (1.9)$$

它的逆变换为:

$$\text{logit}^{-1}(v) = \frac{e^v}{1+e^v}$$

定理 1.1

如果 X 是一个连续随机变量, 并且它的累积概率密度函数为 F_X , 则 $U = F_X(X) \sim U(0, 1)$



证明 在证明之前我们先定义:

$$F_X^{-1}(u) = \inf\{x : F_X(x) = u\}, 0 < u < 1$$

如果随机变量 $U \sim U(0, 1)$, 则对于所有 $x \in \mathbb{R}$:

$$\begin{aligned} P(F_X^{-1}(U) \leq x) &= P(\inf\{t : F_X(t) = U\} \leq x) \\ &= P(U \leq F_X(x)) \\ &= F_U(F_X(x)) \\ &= F_X(x) \end{aligned}$$

得证。

第 A 章 Gamma, Beta 分布

以下内容基于 LDA 的数学八卦。

A.1 Gamma 函数

定义 A.1 (Gamma 函数)

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt \quad (\text{A.1})$$

通过分布积分的方法，可以推导出这个函数有如下的递归性质：

$$\Gamma(x+1) = x\Gamma(x)$$

于是很容易证明， $\Gamma(x)$ 函数可以当成是阶乘在实数集上的延拓，具有如下性质：

$$\Gamma(n) = (n-1)!$$



那么 Gamma 函数是怎么来的呢？这个来源于哥德巴赫在将阶乘推广到实数域的问题上，欧拉解决了这个问题，他发现 $n!$ 可以用如下的一个无穷乘积表达：

$$\left[\left(\frac{2}{1} \right)^n \frac{1}{n+1} \right] \left[\left(\frac{3}{2} \right)^n \frac{2}{n+2} \right] \left[\left(\frac{4}{3} \right)^n \frac{3}{n+3} \right] \cdots = n! \quad (\text{A.2})$$

用极限形式，这个式子整理后可以写为

$$\lim_{m \rightarrow \infty} \frac{1 \cdot 2 \cdot 3 \cdots m}{(1+n)(2+n) \cdots (m+n)} (m+1)^n = n! \quad (\text{A.3})$$

左边可以整理为：

$$\begin{aligned} & \frac{1 \cdot 2 \cdot 3 \cdots m}{(1+n)(2+n) \cdots (m+n)} (m+1)^n \\ &= 1 \cdot 2 \cdot 3 \cdots n \cdot \frac{(n+1)(n+2)m}{(1+n)(2+n) \cdots m} \cdot \frac{(m+1)^n}{(m+1)(m+2) \cdots (m+n)} \\ &= n! \frac{(m+1)^n}{(m+1)(m+2) \cdots (m+n)} \\ &= n! \prod_{k=1}^n \frac{m+1}{m+k} \rightarrow n! \quad (m \rightarrow \infty) \end{aligned}$$

所以上式成立。

之后欧拉对得到的式子进行了处理，得到了积分的形式：

$$n! = \int_0^1 (-\log t)^n dt$$

如果我们令 $t = e^{-u}$ ，就可得我们常见的 Gamma 函数形式：

$$n! = \int_0^{\infty} u^n e^{-u} du$$

于是，利用上式把阶乘延拓到实数集上，我们就得到了 Gamma 函数的一般形式：

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$$