统计学笔记

Ruijun Shi[1]

2021 年 11 月 18 日

[1]GitHub：https://github.com/RuijunShi

摘要

简单的统计学笔记，主要是在天文学，特别是引力波和pulsar timing中遇到的统计学，现在没写多少内容。缓慢更新中[1]。同时这也是我第一次用latex编写书籍，学习过程艰难啊。有错误请大家指出！latex资料参考https://github.com/Ali-loner

Contents

Chapter 1

代数基础

未完待续。。。图片测试 SVD分解 1.1

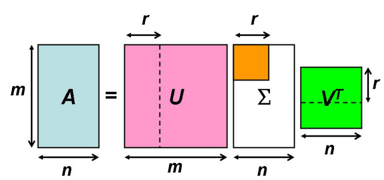

图 1.1: tSVD分解

Chapter 2

数理统计基础

2.1 概率

1. 概率的定义（略）概率满足：非负性，规范性，可列可加性

2. 概率的性质：

   重点：逆事件概率；加法公式；有限可加性

3. 条件概率：

$$P(A|B) = \frac{P(AB)}{P(B)} \tag{2.1}$$

4. 乘法定理：

$$P(AB) = P(A|B)P(A) \tag{2.2}$$

5. 全概率公式：

$$P(A) = P\left(A \mid B_1\right) P\left(B_1\right) + P\left(A \mid B_2\right) P\left(B_2\right) + \ldots + P\left(A \mid B_n\right) P\left(B_n\right) \tag{2.3}$$

6. 独立性：满足

$$P(AB) = P(A)P(B) \tag{2.4a}$$
$$P(B|A) = P(B) \tag{2.4b}$$
$$\tag{2.4c}$$

2.2 单变量分布

1. 随机变量的概念（略）

2. 分布函数的概念（略）和性质：不减函数；$0 \leq F(x) \leq 1$; $F(x+0) = F(x)$

3. 概率密度函数

$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{dt} \tag{2.5}$$

性质：

$$f(x) \leq 0 \tag{2.6a}$$

$$\int_{-\infty}^{\infty} f(x)dx = 1 \tag{2.6b}$$

$$P\{x_1 < X < x_2\} = \int_{x_1}^{x_2} f(x)dx \tag{2.6c}$$

$$F'(x) = f(x) \tag{2.6d}$$

2.3 常见分布

1. (0-1)分布:

$$P(X=k) = p^k(1-p)^{1-k}, 0 < p < 1, k = 0, 1 \tag{2.7}$$

2. 二项分布:

$$P(X=k) = \left(\begin{array}{c} n \\ k \end{array} \right) p^k(1-p)^{n-k} \tag{2.8}$$

3. 泊松分布:

$$P(X=k) = \frac{\lambda^k e^{-\lambda}}{k!}, k = 0, 1, 2, \ldots \tag{2.9}$$

4. Beta分布:

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \tag{2.10}$$

其中：$0 \leq x \leq 1,\ \alpha > 0,\ \beta > 0,\ \Gamma(z) = \int_0^{+\infty} t^{z-1}e^{-t}dt$

5. 均匀分布:

$$f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{otherwise} \end{cases} \tag{2.11}$$

6. 指数分布:

$$f(x) = \begin{cases} \frac{1}{\theta}e^{-x/\theta} & x > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.12}$$

7. 正态分布:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{2.13}$$

$f(x)$关于μ对称；$f(\mu) = max(f(x)) = \frac{1}{\sqrt{2\pi}\sigma}$。

8. $Gamma$分布:

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \tag{2.14}$$

其中：$x > 0,\ \alpha > 0,\ \beta > 0$

9. Inv-$Gamma$分布:

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-(\alpha+1)} e^{-\frac{\beta}{x}} \tag{2.15}$$

其中：$x > 0,\ \alpha > 0,\ \beta > 0$

10. χ^2分布:

$$f_k(x) = \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} x^{\frac{k}{2}-1} e^{-\frac{x}{2}} \tag{2.16}$$

等价$\alpha = k/2, \beta = 1/2$的Gamma分布

11. Inv-χ^2分布:

$$f(x) = \frac{2^{-\frac{k}{2}}}{\Gamma\left(\frac{k}{2}\right)} x^{-\left(\frac{k}{2}+1\right)} e^{-\frac{1}{2x}} \tag{2.17}$$

等价$\alpha = k/2, \beta = 1/2$的Inv-Gamma分布

12. Scaled Inv-χ^2分布:

$$f(x) = \frac{\frac{k}{2}^{-\frac{k}{2}} s^k}{\Gamma\left(\frac{k}{2}\right)} x^{-\left(\frac{k}{2}+1\right)} e^{-\frac{ks^2}{2x}} \tag{2.18}$$

等价$\alpha = k/2, \beta = ks^2/2$的Inv-Gamma分布。

2.4 多元随机变量

2.5 随机变量的数学特征

2.6 频率学派的参数估计

2.7 频率学派的假设检验

Chapter 3

贝叶斯统计

3.1 贝叶斯定理推导

1. 贝叶斯公式

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^{n} P(A|B_i)P(B_j)} = \frac{P(A|B_i)P(B_i)}{P(A)} \tag{3.1}$$

先验：$P(B)$

似然: $P(A|B)$

后验: $P(B|A)$

证据（归一化）: $P(A)$

2. 贝叶斯公式含义：通过数据推算模型参数的概率。即：

$$P(\text{Model}(\theta)|\text{Data}) = P(\text{Data}|\text{Model}(\theta))P(\theta) \tag{3.2}$$

3. 贝叶斯统计的优势：将这个某种程度上是主观性的信息明确表达在先验概率中，而不是隐藏在没有明确指出的假设中; 让数据说话，减少主观性的先验概率。

3.2 贝叶斯单参数估计

一、二项分布估计

1. 无信息先验

- 似然：

$$p(y|\theta) = \begin{pmatrix} n \\ y \end{pmatrix} \theta^y (1-\theta)^{n-y} \tag{3.3}$$

- 先验：均匀分布
- 后验：

$$p(\theta|y) \propto \theta^y (1-\theta)^{n-y} \sim Beta(y+1, n-y+1)$$

- 预测：

$$Pr(\widetilde{y}=1|y) = \int_0^1 Pr(\widetilde{y}=1|\theta, y) = \int_0^1 \theta p(\theta|y) d\theta = E(\theta|y)$$

1. **有信息先验**

- 先验：$p(\theta) \propto \theta^{\alpha-1}(1-\theta)^{\beta-1}$
- 似然：

$$p(y|\theta) = \begin{pmatrix} n \\ y \end{pmatrix} \theta^y (1-\theta)^{n-y} \tag{3.4}$$

- 后验：

$$p(\theta|y) \propto \theta^{y+\alpha-1}(1-\theta)^{n-y+\beta-1} \sim Beta(\alpha+y, \beta+n-y)$$

- 后验期望：$E(\theta|y) = \frac{\alpha+y}{\alpha+\beta+n}$
- 先验期望：$E(y) = \frac{\alpha}{\alpha+\beta}$

当$n \to \infty$: $E(\theta|y) \to y/n$

数据很大的时候可以用正态分布近似后验分布

二、正态分布参数估计

1. **已知方差求均值**

- 似然：

$$p(y|\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(y-\theta)^2\right] \sim N(\theta, \sigma^2)$$

- 先验：

$$p(\theta) \propto \exp\left(-\frac{1}{2\tau_0^2}(\theta-\mu_0)^2\right) \sim N(\mu_0, \tau_0^2)$$

- 后验：

$$p(\theta|y) \propto \exp\left(-\frac{1}{2\tau_1^2}(\theta-\mu_1)^2\right) \sim N(\mu_1, \tau_1)$$

其中：

$$\mu_1 = \frac{\frac{1}{\tau_0^2}\mu_0 + \frac{1}{\sigma^2}y}{\frac{1}{\tau_0^2} + \frac{1}{\sigma^2}}$$

$$\frac{1}{\tau_1^2} = \frac{1}{\tau_0^2} + \frac{1}{\sigma^2}$$

精度：方差的倒数

- 预测：

$$\begin{aligned}
p(\tilde{y}|y) &= \int p(\tilde{y}|\theta)p(\theta|y)d\theta \\
&\propto \int \exp\left(-\frac{(\tilde{y}-\theta)^2}{2\sigma^2}\right)\exp\left(-\frac{(\theta-\mu_1)^2}{2\tau_1^2}\right)d\theta \\
E(\tilde{y}|\theta) &= \theta \\
D(\tilde{y}|\theta) &= \sigma^2 + \tau_1^2
\end{aligned}$$

- 多个相互独立数据：

$$\begin{aligned}
p(\theta \mid y) &\propto p(\theta)p(y \mid \theta) \\
&= p(\theta)\prod_{i=1}^{n} p\left(y_i \mid \theta\right) \\
&\propto \exp\left(-\frac{1}{2\tau_0^2}\left(\theta-\mu_0\right)^2\right)\prod_{i=1}^{n}\exp\left(-\frac{1}{2\sigma^2}\left(y_i-\theta\right)^2\right) \\
&\propto \exp\left(-\frac{1}{2}\left[\frac{1}{\tau_0^2}\left(\theta-\mu_0\right)^2 + \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(y_i-\theta\right)^2\right]\right)
\end{aligned} \tag{3.5}$$

可得：

$$p(\theta|y_1,\cdots,y_n)=p(\theta|\bar{y})=N(\theta|\mu_n,\tau_n^2)$$

其中：

$$\mu_n=\frac{\frac{1}{\tau_0^2}\mu_0+\frac{n}{\sigma^2}\bar{y}}{\frac{1}{\tau_0^2}+\frac{n}{\sigma^2}}$$

$$\frac{1}{\tau_n^2}=\frac{1}{\tau_0^2}+\frac{n}{\sigma^2}$$

若$n\to+\infty$，τ_0不变，则：$\theta|y\sim N(\bar{y},\sigma^2/n)$

若$\tau\to+\infty$，n 不变，则：$\theta|y\sim N(\bar{y},\sigma^2/n)$

1. **已知均值求方差**

由于有n个服从$\sim N(\theta,\sigma^2)$的分布，因此：

- 似然：

$$p(y|\sigma^2)\propto\sigma^{-n}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\theta)^2\right)=(\sigma^2)^{-n/2}e^{-\frac{n}{2\sigma^2}v}$$

其中：

$$v=\frac{1}{n}\sum_{i=1}^{n}(y_i-\theta)^2$$

- 共轭先验：$p(\sigma^2)\propto(\sigma^2)^{-\alpha+1}e^{\beta/\sigma^2}\sim \mathrm{Inv}-\chi^2(v_0,\sigma_0^2)$
- 后验：

$$\begin{aligned}p(\sigma^2|y)&\propto p(\sigma^2)p(y|\sigma^2)\\&\propto\left(\frac{\sigma_0^2}{\sigma^2}\right)^{v_0/2+1}\exp\left(-\frac{v_0\sigma_0^2}{2\sigma^2}\right)\cdot(\sigma^2)^{-n/2}\exp\left(-\frac{n}{2}\frac{v}{\sigma^2}\right)\\&\propto(\sigma^2)^{-((n+v_0)/2+1)}\exp\left(-\frac{1}{2\sigma^2}(v_0\sigma_0^2+nv)\right)\\&\sim Inv-\chi^2(v_0+n,\frac{v_0\sigma_0^2+nv}{v_0+n})\end{aligned}\tag{3.6}$$

3.3 贝叶斯多参数模型

一、多参数模型处理

1. **置之不理**
2. **边缘化**

$$p(\theta_1|y) = \int p(\theta_1, \theta_2|y) d\theta_2 \tag{3.7}$$

   用贝叶斯公式展开：

$$p(\theta_1, \theta_2|y) \propto p(y|\theta_1, \theta_2)p(\theta_1, \theta_2)$$

   将θ_2边缘化积分，得到θ_1的后验分布。

3. **平均化**

$$p(\theta_1|y) = \int p(\theta_1|\theta_2, y)p(\theta_2|y)d\theta_2$$

二、无信息先验的正态分布

- 先验：

$$p(\mu, \ln \sigma^2) \sim U(\mu, \ln \sigma^2)$$

  或者先验写为：

$$p(\mu, \sigma^2) \sim \frac{1}{\sigma^2}$$

- 似然（**有n次观测**）：

$$p(y|\mu, \sigma^2) = \sigma^{-n} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \mu)^2\right)$$

- 联合后验：

$$p(\mu, \sigma^2|y) \propto \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}[(n-1)s^2 + n(\bar{y} - \mu)^2]\right)$$

  其中：

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$$

- 边缘后验$p(\sigma^2|y)$：

$$\begin{aligned}
p(\sigma^2|y) &\propto \int p(\mu,\sigma^2|y)d\mu \\
&\propto (\sigma^2)^{-\frac{n+1}{2}} \exp\left(-\frac{(n-1)s^2}{2\sigma^2}\right) \\
&\sim Inv-\chi^2(n-1,s^2)
\end{aligned}$$

- 边缘后验$p(\mu|y)$：

$$\begin{aligned}
p(\mu|y) &= \int_0^{\infty} p(\mu,\sigma^2|y)d\sigma^2 \\
&\propto \left[1+\frac{n(\mu-\bar{y})^2}{(n-1)s^2}^{-n/2}\right] \\
&\sim t_{n-1}(\bar{y},s^2/n)
\end{aligned}$$

- 预测后验分布

$$p(\widetilde{y}|y) = \iint p(\widetilde{y}|\mu,\sigma^2,y)p(\mu,\sigma^2|y)d\mu d\sigma^2$$

三、共轭先验分布

- 先验分布：

  $\mu|\sigma^2 \sim N(\mu_0,\sigma^2/\kappa_0)$

  $\sigma^2 \sim Inv-\chi^2(\nu_0,\sigma_0^2)$

- 联合先验分布

$$\begin{aligned}
p(\mu,\sigma^2) &= p(\mu|\sigma^2)p(\sigma^2) \\
&\propto N-Inv-\chi^2(\mu_0,\sigma_0^2/\kappa_0\ ;\ \nu_0,\sigma_0^2)
\end{aligned}$$

- 似然分布

$$p(y|\mu,\sigma^2)=\sigma^{-n}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2\right)$$

- 联合后验分布

$$\begin{aligned} p\left(\mu, \sigma^2 \mid y\right) &\propto \sigma^{-1}\left(\sigma^2\right)^{\left(\nu_0/2+1\right)} e^{-\frac{1}{2\sigma^2}\left[\nu_0\sigma_0^2+\kappa_0\left(\mu-\mu_0\right)^2\right]}\left(\sigma^2\right)^{-n/2} e^{-\frac{1}{2\sigma^2}\left[(n-1)s^2+n(\bar{y}-\mu)^2\right]} \\ &\propto \mathrm{N}-\operatorname{Inv}-\chi^2\left(\mu_n, \sigma_n^2/\kappa_n ; \nu_n, \sigma_n^2\right) \end{aligned} \tag{3.8}$$

其中参数为:

$$\begin{gathered} \mu_n=\tfrac{\kappa_0}{\kappa_0+n}\mu_0+\tfrac{n}{\kappa_0+n}\bar{y} \\ \kappa_n=\kappa_0+n \\ \nu_n=\nu_0+n \\ \nu_n\sigma_n^2=\nu_0\sigma_0^2+(n-1)s^2+\tfrac{\kappa_0 n}{\kappa_0+n}\left(\bar{y}-\mu_0\right)^2 \end{gathered} \tag{3.9}$$

- 条件后验分布$p(\mu|\sigma^2,y)$

$$\left(\mu \mid \sigma^2, y\right) \sim \mathrm{N}\left(\mu_n \frac{\sigma^2}{\kappa_n}\right)$$

- 方差边缘后验分布$p(\sigma^2|y)$

$$\left(\sigma^2 \mid y\right) \sim \operatorname{Inv}-\chi^2\left(\nu_n, \sigma_n^2\right)$$

- 均值边缘后验分布$p(\mu|y)$

$$\begin{aligned} p(\mu \mid y) &\propto\left[1+\frac{\kappa_n\left(\mu-\mu_n\right)^2}{\nu_n\sigma_n^2}\right]^{-\left(\nu_n+1\right)/2} \\ &=t_{\nu_n}\left(\mu \mid mu_n, \sigma_n^2/\kappa_n\right) \end{aligned} \tag{3.10}$$

3.4 Chapter 5层次化模型

一、参数化先验分布

1. 先验分布

   先验分布是由某个未知参数的分布ϕ给出:

$$p(\theta|\phi)=\prod_{j=1}^{J}p(\theta_j|\phi)\qquad\theta=(\theta_1,\theta_2,\cdots)$$

边缘化：

$$p(\theta) = \int \left(\prod_{j=1}^{J} p(\theta_j|\phi) \right) p(\phi) d\phi$$

2. 联合先验分布 $p(\phi, \theta) = p(\theta|\phi)p(\phi)$

3. 超先验分布：$p(\phi)$

4. 后验分布

   - 联合后验：$p(\phi, \theta|y) = p(y|\theta)p(\theta|\phi)p(\phi)$
   - 条件后验：$p(\theta|\phi, y)$
   - 边缘后验：$p(\phi|y)$

$$p(\phi|y) = \frac{p(\theta, \phi|y)}{p(\theta|\phi, y)}$$

5. 层次化贝叶斯完整表述

$$\begin{aligned} p(\phi, \theta|y) &\propto p(y|\phi, \theta)p(\phi, \theta) \\ &= p(y|\theta)p(\phi, \theta) \\ &= p(y|\theta)p(\theta|\phi)p(\phi) \end{aligned} \tag{3.11}$$

1. 层次化贝叶斯计算步骤

- 写出联合后验分布$p(\theta, \phi|y)$：即超先验分布，总体分布和似然分布的乘积
- 确定条件后验分布$p(\theta|\phi, y)$：

$$p(\theta|\phi, y) = \prod_{j=1}^{J} p(\theta_j|\phi, y)$$

- 边缘化给出ϕ的贝叶斯估计

二、二项分布的分层贝叶斯模型

1. y_i先验（组内模型） $y_j \sim Bin(n_j, \theta_j)$

2. θ_j先验（组间模型）：$\theta_j \sim Beta(\alpha, \beta)$

3. 联合先验：$p(\alpha, \beta, \theta) = p(\alpha, \beta)p(\theta|\alpha, \beta)$

4. 似然：$p(y|\theta,\alpha,\beta)$

5. 联合后验：$p(\theta,\alpha,\beta|y)$

$$
\begin{aligned}
p(\theta,\alpha,\beta|y) &\propto p(\alpha,\beta)p(\theta|\alpha,\beta)p(y|\theta,\alpha,\beta) \\
&= p(\alpha,\beta)\prod_{j=1}^{J}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta_j^{\alpha-1}(1-\theta_j)^{\beta-1}\prod_{j=1}^{J}\theta_j^{y_j}(1-\theta_j)^{n_j-y_j}
\end{aligned}
\tag{3.12}
$$

6. 条件后验： $p(\theta|\alpha,\beta,y)$ ：单参数模型给定的后验

$$
\begin{aligned}
p(\theta \mid \alpha,\beta,y) &= \prod_{j=1}^{J}\frac{\Gamma(\alpha+\beta+n_j)}{\Gamma(\alpha+y_j)\Gamma(\beta+n_j-y_j)}\theta_j^{\alpha+y_j-1}(1-\theta_j)^{\beta+n_j-y_j-1} \\
&\sim \prod_{j=1}^{J} Beta(\alpha+y_j,\beta+n_j-y_j)
\end{aligned}
\tag{3.13}
$$

7. 边缘后验：$p(\alpha,\beta|y)$

$$
\begin{aligned}
p(\alpha,\beta \mid y) =& \frac{p(\theta,\alpha,\beta \mid y)}{p(\theta \mid \alpha,\beta,y)} \\
\propto& \frac{p(\alpha,\beta)\prod_{j=1}^{J}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta_j^{\alpha-1}(1-\theta_j)^{\beta-1}\prod_{j=1}^{J}\theta_j^{y_j}(1-\theta_j)^{n_j-y_j}}{\prod_{j=1}^{J}\frac{\Gamma(\alpha+\beta+n_j)}{\Gamma(\alpha+y_j)\Gamma(\beta+n_j-y_j)}\theta_j^{\alpha+y_j-1}(1-\theta_j)^{\beta+n_j-y_j-1}} \\
=& p(\alpha,\beta)\prod_{j=1}^{J}\frac{\Gamma(\alpha+\beta)\Gamma(\alpha+y_j)\Gamma(\beta+n_j-y_j)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\alpha+\beta+n_j)}
\end{aligned}
\tag{3.14}
$$

三、正态分布的分层贝叶斯模型

1. 数据结构

假设J个独立试验，每个实验都由θ_j给出其参数估计，估计n_j个$i.i.d$正态分布的数据点y_{ij}，每个点方差为σ^2：

$$
y_{ij}|\theta_j \sim N\left(\theta_j,\sigma^2\right) \quad i=1,\ldots,n_j \quad j=1,\ldots,J
$$

样本均值（充分统计量）：

$$
\bar{y}_{\cdot j} = \frac{1}{n_j}\sum_{i=1}^{n_j} y_{ij}
$$

样本均值的分布

$$\bar{y}_{\cdot j} \sim N(\theta_j, \sigma_j^2)$$

样本方差：

$$\sigma_j^2 = \frac{\sigma^2}{n_j}$$

样本的均值是从θ中估计，样本方差是从σ中估计。

相当于 θ 的似然分布：

$$\bar{y}_{\cdot j}|\theta \sim N(\theta_j, \sigma_j^2)$$

由于 σ 是已知的，下面所有的分布都是在 σ 已知情况下成立。

1. **层次化模型：无信息先验**

θ是从参数(μ, τ)中抽取：

$$p\left(\theta_1, \ldots, \theta_J \mid \mu, \tau^2\right) = \prod_{j=1}^{J} p\left(\theta_j \mid \mu, \tau^2\right)$$

边缘化：

$$p\left(\theta_1, \ldots, \theta_J\right) = \iint \prod_{j=1}^{J} \left[p\left(\theta_j \mid \mu, \tau^2\right)\right] p\left(\mu, \tau^2\right) d\mu d\tau$$

- **先验和似然**

组内抽样：$\bar{y}_{\cdot j} \sim N(\theta_j, \sigma_j^2)$

θ的先验（组内模型）：$\theta|\mu, \tau \sim N(\mu, \tau^2)$

μ的先验： $p(\mu, \tau) = p(\mu|\tau)p(\tau) \propto p(\tau)$

θ_j似然分布：$p(y|\theta) \sim (\bar{y}_{\cdot j}|\theta) \sim N(\theta_j, \sigma_j^2)$

- **联合后验：**$p(\theta, \phi|y)$

$$\begin{aligned} p(\theta, \mu, \tau|y) &\propto p(\mu, \tau)p(\theta|\mu, \tau)p(y|\theta) \\ &\propto p(\mu, \tau) \prod_{j=1}^{J} p(\theta_j|\mu, \tau^2) \prod_{j=1}^{J} p(\bar{y}_{\cdot j}|\theta_j, \sigma_j^2) \end{aligned} \tag{3.15}$$

其中：$\bar{y}_{\cdot j} \sim N(\theta_j, \sigma_j^2)$

可以忽略只依赖 y 和 σ_j 的参数，因为其已知。

- **θ条件后验**：$p(\theta_j \mid \mu, \tau, y_{\cdot,j})$

$$\theta_j \mid \mu, \tau, y_{\cdot,j} \sim N\left(\hat{\theta}_j, V_j\right)$$

其中：

$$\hat{\theta}_j = \frac{\frac{1}{\sigma_j^2}\bar{y}_{\cdot j} + \frac{1}{\tau^2}\mu}{\frac{1}{\sigma_j^2} + \frac{1}{\tau^2}}$$

$$V_j = \frac{1}{\frac{1}{\sigma_j^2} + \frac{1}{\tau^2}}$$

- **超参数边缘后验**：$p(\mu, \tau \mid y)$

$$p(\mu, \tau \mid y) \propto p(\mu, \tau) p(y \mid \mu, \tau)$$

对于正态分布：

$$\bar{y}_{\cdot j} \sim N(\mu, \tau^2 + \sigma^2)$$

因此：

$$p(\mu, \tau \mid y) \propto p(\mu, \tau) \prod_{j=1}^{J} p\left(\bar{y}_{\cdot j} \mid \mu, \tau^2 + \sigma_j^2\right)$$

$$\bar{y}_{\cdot j} \mid \mu, (\tau^2 + \sigma_j^2) \sim N\left(\mu, \tau^2 + \sigma_j^2\right)$$

- **给定τ下μ的边缘后验分布**： $p(\mu \mid \tau, y)$从单参数模型中得出的结论

$$\mu \mid \tau, y \sim N\left(\hat{\mu}, V_\mu\right) \tag{3.16}$$

$$\hat{\mu} = \frac{\sum_{j=1}^{J} \frac{1}{\sigma_j^2 + \tau^2} \bar{y}_{\cdot j}}{\sum_{j=1}^{J} \frac{1}{\sigma_j^2 + \tau^2}}$$

$$V_\mu^{-1} = \sum_{j=1}^{J} \frac{1}{\sigma_j^2 + \tau^2}$$

- τ的后验：$p(\tau \mid y)$

$$\begin{aligned} p(\tau \mid y) &= \frac{p(\mu, \tau \mid y)}{p(\mu \mid \tau, y)} \\ &\propto \frac{p(\tau) \prod_{j=1}^{J} N\left(\bar{y}_{\cdot j} \mid \mu, \sigma_j^2 + \tau^2\right)}{N\left(\mu \mid \hat{\mu}, V_\mu\right)} \\ &\propto p(\tau) V_\mu^{1/2} \prod_{j=1}^{J} \left(\sigma_j^2 + \tau^2\right)^{-1/2} \exp\left(-\frac{\left(\bar{y}_{\cdot j} - \hat{\mu}\right)^2}{2\left(\sigma_j^2 + \tau^2\right)}\right) \end{aligned} \tag{3.17}$$

- 从后往前采样，即先算出τ的后验，然后依次采样算出μ的后验，超参数联合后验，θ的条件后验，联合后验等。

3.5 贝叶斯回归

3.6 贝叶斯模型选择

Chapter 4

随机过程

4.1 随机过程及其统计描述

4.2 平稳随机过程

4.3 马尔科夫链

Chapter 5

MCMC

5.1 拒绝-接受采样

5.2 吉布斯采用

5.3 M-H采样

5.4 Nested采样

Chapter 6

高斯随机过程

6.1 高斯随机过程及其统计描述

6.2 核密度估计

6.3 高斯混合模型

6.4 高斯学习

参考文献

[1] Jaranowski, P., and Królak, A. *Analysis of gravitational-wave data.* No. 29 in Cambridge monographs on particle physics, nuclear physics and cosmology. Cambridge University Press, Cambridge ; New York, 2009.