

贝 叶 斯 统 计

姓名：韦 来 生 编 著

高 等 教 育 出 版 社

第一版，2016

贝 叶 斯 统 计

姓名：韦 来 生 编 著

高 等 教 育 出 版 社

第一版，2016

第5章 贝叶斯统计决策

1 5.1 引言

2 5.2 后验风险最小原则

- 5.2.1 后验风险的定义
- 5.2.2 后验风险与贝叶斯风险的关系
- 5.2.3 后验风险最小的原则

3 5.3 一般损失函数下的贝叶斯估计

- 5.3.1 平方损失下的贝叶斯估计
- 5.3.2 加权平方损失下的贝叶斯估计
- 5.3.3 绝对损失下的贝叶斯估计
- 5.3.4 线性损失函数下的贝叶斯估计

4 5.4 假设检验和有限行动（分类）问题

- 5.4.1 假设检验问题
- 5.4.2 多行动（分类）问题
- 5.4.3 统计决策中的区间估计问题*

5 5.5 Minimax 准则*

- 5.5.1 引言及定义
- 5.5.2 Minimax 解的求法

1 5.1 引言

2 5.2 后验风险最小原则

- 5.2.1 后验风险的定义
- 5.2.2 后验风险与贝叶斯风险的关系
- 5.2.3 后验风险最小的原则

3 5.3 一般损失函数下的贝叶斯估计

- 5.3.1 平方损失下的贝叶斯估计
- 5.3.2 加权平方损失下的贝叶斯估计
- 5.3.3 绝对损失下的贝叶斯估计
- 5.3.4 线性损失函数下的贝叶斯估计

4 5.4 假设检验和有限行动（分类）问题

- 5.4.1 假设检验问题
- 5.4.2 多行动 (分类) 问题
- 5.4.3 统计决策中的区间估计问题*

5 5.5 Minimax 准则*

- 5.5.1 引言及定义
- 5.5.2 Minimax 解的求法

1. 贝叶斯统计决策四要素

- **样本空间和样本分布族 $\{\mathcal{X}, F_\theta\}$:** 取值于样本空间 \mathcal{X} 内的随机变量 X 及其分布族 $\{F_\theta(x), \theta \in \Theta\}$ 是构成统计决策问题的第一个要素.
- **行动空间 \mathcal{D} :** 某个统计决策问题可能采取的行动所构成的非空集合, 被称为行动空间, 记为 \mathcal{D} . 估计问题中, $\mathcal{D} = \Theta$. 检验问题中 $\mathcal{D} = \{d_0, d_1\}$ 只有两个行动组成, d_0 表示接受原假设 H_0 , d_1 表示拒绝 H_0 .
- **损失函数 $L(\theta, d)$:** 损失函数是定义于 $\Theta \times \mathcal{D}$ 上的非负函数, 记为 $L(\theta, d)$. 它表示参数为 θ 时采取行动 $d \in \mathcal{D}$ 所蒙受的损失. 常用的有平方损失、绝对值损失和线性损失等.
- **先验分布 $\pi(\theta)$:** 在贝叶斯统计决策问题中还要求有一个定义于参数空间 Θ 上的先验分布 $\pi(\theta)$.
- **统计决策问题就是研究如何根据样本 X 的值恰当地选取行动 d 使得引起的平均损失最小.**

1. 贝叶斯统计决策四要素

- **样本空间和样本分布族 $\{\mathcal{X}, F_\theta\}$:** 取值于样本空间 \mathcal{X} 内的随机变量 X 及其分布族 $\{F_\theta(x), \theta \in \Theta\}$ 是构成统计决策问题的第一个要素.
- **行动空间 \mathcal{D} :** 某个统计决策问题可能采取的行动所构成的非空集合, 被称为行动空间, 记为 \mathcal{D} . 估计问题中, $\mathcal{D} = \Theta$. 检验问题中 $\mathcal{D} = \{d_0, d_1\}$ 只有两个行动组成, d_0 表示接受原假设 H_0 , d_1 表示拒绝 H_0 .
- **损失函数 $L(\theta, d)$:** 损失函数是定义于 $\Theta \times \mathcal{D}$ 上的非负函数, 记为 $L(\theta, d)$. 它表示参数为 θ 时采取行动 $d \in \mathcal{D}$ 所蒙受的损失. 常用的有平方损失、绝对值损失和线性损失等.
- **先验分布 $\pi(\theta)$:** 在贝叶斯统计决策问题中还要求有一个定义于参数空间 Θ 上的先验分布 $\pi(\theta)$.
- **统计决策问题就是研究如何根据样本 X 的值恰当地选取行动 d 使得引起的平均损失最小.**

1. 贝叶斯统计决策四要素

- **样本空间和样本分布族 $\{\mathcal{X}, F_\theta\}$:** 取值于样本空间 \mathcal{X} 内的随机变量 X 及其分布族 $\{F_\theta(x), \theta \in \Theta\}$ 是构成统计决策问题的第一个要素.
- **行动空间 \mathcal{D} :** 某个统计决策问题可能采取的行动所构成的非空集合, 被称为行动空间, 记为 \mathcal{D} . 估计问题中, $\mathcal{D} = \Theta$. 检验问题中 $\mathcal{D} = \{d_0, d_1\}$ 只有两个行动组成, d_0 表示接受原假设 H_0 , d_1 表示拒绝 H_0 .
- **损失函数 $L(\theta, d)$:** 损失函数是定义于 $\Theta \times \mathcal{D}$ 上的非负函数, 记为 $L(\theta, d)$. 它表示参数为 θ 时采取行动 $d \in \mathcal{D}$ 所蒙受的损失. 常用的有平方损失、绝对值损失和线性损失等.
- **先验分布 $\pi(\theta)$:** 在贝叶斯统计决策问题中还要求有一个定义于参数空间 Θ 上的先验分布 $\pi(\theta)$.
- **统计决策问题就是研究如何根据样本 X 的值恰当地选取行动 d 使得引起的平均损失最小.**

1. 贝叶斯统计决策四要素

- **样本空间和样本分布族 $\{\mathcal{X}, F_\theta\}$:** 取值于样本空间 \mathcal{X} 内的随机变量 X 及其分布族 $\{F_\theta(x), \theta \in \Theta\}$ 是构成统计决策问题的第一个要素.
- **行动空间 \mathcal{D} :** 某个统计决策问题可能采取的行动所构成的非空集合, 被称为行动空间, 记为 \mathcal{D} . 估计问题中, $\mathcal{D} = \Theta$. 检验问题中 $\mathcal{D} = \{d_0, d_1\}$ 只有两个行动组成, d_0 表示接受原假设 H_0 , d_1 表示拒绝 H_0 .
- **损失函数 $L(\theta, d)$:** 损失函数是定义于 $\Theta \times \mathcal{D}$ 上的非负函数, 记为 $L(\theta, d)$. 它表示参数为 θ 时采取行动 $d \in \mathcal{D}$ 所蒙受的损失. 常用的有平方损失、绝对值损失和线性损失等.
- **先验分布 $\pi(\theta)$:** 在贝叶斯统计决策问题中还要求有一个定义于参数空间 Θ 上的先验分布 $\pi(\theta)$.
- 统计决策问题就是研究如何根据样本 X 的值恰当地选取行动 d 使得引起的平均损失最小.

1. 贝叶斯统计决策四要素

- **样本空间和样本分布族 $\{\mathcal{X}, F_\theta\}$:** 取值于样本空间 \mathcal{X} 内的随机变量 X 及其分布族 $\{F_\theta(x), \theta \in \Theta\}$ 是构成统计决策问题的第一个要素.
- **行动空间 \mathcal{D} :** 某个统计决策问题可能采取的行动所构成的非空集合, 被称为行动空间, 记为 \mathcal{D} . 估计问题中, $\mathcal{D} = \Theta$. 检验问题中 $\mathcal{D} = \{d_0, d_1\}$ 只有两个行动组成, d_0 表示接受原假设 H_0 , d_1 表示拒绝 H_0 .
- **损失函数 $L(\theta, d)$:** 损失函数是定义于 $\Theta \times \mathcal{D}$ 上的非负函数, 记为 $L(\theta, d)$. 它表示参数为 θ 时采取行动 $d \in \mathcal{D}$ 所蒙受的损失. 常用的有平方损失、绝对值损失和线性损失等.
- **先验分布 $\pi(\theta)$:** 在贝叶斯统计决策问题中还要求有一个定义于参数空间 Θ 上的先验分布 $\pi(\theta)$.
- **统计决策问题就是研究如何根据样本 X 的值恰当地选取行动 d 使得引起的平均损失最小.**

2. 统计决策问题的一个例子

- 为了说明什么是决策 (判决) 问题, 请看下例:
- 例 5.1.1 一位投资者有一笔资金要进行投资, 有如下几个投资方案供他选择: a_1 : 买股票, 根据市场情况可净赚 5000元, 但也可能亏损 10000元. a_2 : 购买基金, 根据市场情况可净赚 3000元, 但也可能亏损 8000元. a_3 : 存入银行, 不管市场情况如何, 总可净赚 1000元. 他应如何决策?
- 这位投资者在与金融市场博弈. 未来的金融市场也有两种情况: 看涨 (θ_1) 与看跌 (θ_2). 根据上述情况, 可写出投资者的收益矩阵如下:

行动	a_1	a_2	a_3
θ_1	5000	3000	1000
θ_2	-10000	-8000	1000

2. 统计决策问题的一个例子

- 为了说明什么是决策 (判决) 问题, 请看下例:
- 例 5.1.1 一位投资者有一笔资金要进行投资, 有如下几个投资方案供他选择: a_1 : 买股票, 根据市场情况可净赚 5000元, 但也可能亏损 10000元. a_2 : 购买基金, 根据市场情况可净赚 3000元, 但也可能亏损 8000元. a_3 : 存入银行, 不管市场情况如何, 总可净赚 1000元. 他应如何决策?
- 这位投资者在与金融市场博弈. 未来的金融市场也有两种情况: 看涨 (θ_1) 与看跌 (θ_2). 根据上述情况, 可写出投资者的收益矩阵如下:

行动	a_1	a_2	a_3
θ_1	5000	3000	1000
θ_2	-10000	-8000	1000

2. 统计决策问题的一个例子

- 为了说明什么是决策 (判决) 问题, 请看下例:
- 例 5.1.1 一位投资者有一笔资金要进行投资, 有如下几个投资方案供他选择: a_1 : 买股票, 根据市场情况可净赚 5000元, 但也可能亏损 10000元. a_2 : 购买基金, 根据市场情况可净赚 3000元, 但也可能亏损 8000元. a_3 : 存入银行, 不管市场情况如何, 总可净赚 1000元. 他应如何决策?
- 这位投资者在与金融市场博弈. 未来的金融市场也有两种情况: 看涨 (θ_1) 与看跌 (θ_2). 根据上述情况, 可写出投资者的收益矩阵如下:

行动	a_1	a_2	a_3
θ_1	5000	3000	1000
θ_2	-10000	-8000	1000

3. 统计决策问题过程

- 投资者将依据收益矩阵决定资金投向何方. 这种人与自然界 (或社会) 的博弈问题称为决策问题. 在决策问题中, 主要寻求人对自然界 (或社会) 的最优策略. 决策问题也不一定要涉及统计方法. 如果它满足以下条件, 那就必然与统计方法有关, 因而就可以称为统计决策问题. 这条件是: 在作出决策时所依据的事实中, 至少有一个部分是受到随机性影响的观察值 (或试验数据) X .
- 决策实际上是一个过程, 它可分为两部分: 第一部分是把决策问题描述清楚; 第二部分是如何做决策使得收益最大 (或损失最小). 显然第二部分是我们研究的重点, 但首先要把第一部分搞清楚, 这就需要下面一些基本概念.

3. 统计决策问题过程

- 投资者将依据收益矩阵决定资金投向何方. 这种人与自然界 (或社会) 的博弈问题称为决策问题. 在决策问题中, 主要寻求人对自然界 (或社会) 的最优策略. 决策问题也不一定要涉及统计方法. 如果它满足以下条件, 那就必然与统计方法有关, 因而就可以称为统计决策问题. 这条件是: 在作出决策时所依据的事实中, 至少有一个部分是受到随机性影响的观察值 (或试验数据) X .
- 决策实际上是一个过程, 它可分为两部分: 第一部分是把决策问题描述清楚; 第二部分是如何做决策使得收益最大 (或损失最小). 显然第二部分是我们研究的重点, 但首先要把第一部分搞清楚, 这就需要下面一些基本概念.

1 5.1 引言

2 5.2 后验风险最小原则

- 5.2.1 后验风险的定义
- 5.2.2 后验风险与贝叶斯风险的关系
- 5.2.3 后验风险最小的原则

3 5.3 一般损失函数下的贝叶斯估计

- 5.3.1 平方损失下的贝叶斯估计
- 5.3.2 加权平方损失下的贝叶斯估计
- 5.3.3 绝对损失下的贝叶斯估计
- 5.3.4 线性损失函数下的贝叶斯估计

4 5.4 假设检验和有限行动（分类）问题

- 5.4.1 假设检验问题
- 5.4.2 多行动（分类）问题
- 5.4.3 统计决策中的区间估计问题*

5 5.5 Minimax 准则*

- 5.5.1 引言及定义
- 5.5.2 Minimax 解的求法

后验风险的定义

- 设 $L(\theta, \delta(x))$ 为损失函数，我们知道将损失函数按样本分布求平均就得到风险函数，若将损失函数按后验分布 $\pi(\theta|x)$ 求期望就得到后验风险，其定义如下：

- 定义 5.2.1 设 δ 是一个决策函数，称平均损失

$$R(\delta(x)|x) = E^{\theta|x}[L(\theta, \delta(x))]$$

$$= \begin{cases} \int_{\Theta} L(\theta, \delta(x)) \pi(\theta|x) d\theta & \text{当 } \theta \text{ 为连续型随机变量} \\ \sum_i L(\theta_i, \delta(x)) \pi(\theta_i|x) & \text{当 } \theta \text{ 为离散型随机变量} \end{cases}$$

为决策函数 $\delta(x)$ 的后验风险。

- 若存在决策函数 $\delta^*(x)$ ，使得

$$R(\delta^*|x) = \min_{\delta} R(\delta(x)|x), \quad \text{对 } \forall \text{ 判决函数 } \delta(x), \quad (5.2.1)$$

则称 $\delta^*(x)$ 为后验风险最小准则下的最优贝叶斯决策函数。

后验风险的定义

- 设 $L(\theta, \delta(x))$ 为损失函数，我们知道将损失函数按样本分布求平均就得到风险函数，若将损失函数按后验分布 $\pi(\theta|x)$ 求期望就得到后验风险，其定义如下：

- **定义 5.2.1** 设 δ 是一个决策函数，称平均损失

$$R(\delta(x)|x) = E^{\theta|x}[L(\theta, \delta(x))]$$

$$= \begin{cases} \int_{\Theta} L(\theta, \delta(x)) \pi(\theta|x) d\theta & \text{当 } \theta \text{ 为连续型随机变量} \\ \sum_i L(\theta_i, \delta(x)) \pi(\theta_i|x) & \text{当 } \theta \text{ 为离散型随机变量} \end{cases}$$

为决策函数 $\delta(x)$ 的后验风险。

- 若存在决策函数 $\delta^*(x)$ ，使得

$$R(\delta^*|x) = \min_{\delta} R(\delta(x)|x), \quad \text{对 } \forall \text{ 判决函数 } \delta(x), \quad (5.2.1)$$

则称 $\delta^*(x)$ 为后验风险最小准则下的最优贝叶斯决策函数。

后验风险的定义

- 设 $L(\theta, \delta(x))$ 为损失函数，我们知道将损失函数按样本分布求平均就得到风险函数，若将损失函数按后验分布 $\pi(\theta|x)$ 求期望就得到后验风险，其定义如下：

- **定义 5.2.1** 设 δ 是一个决策函数，称平均损失

$$R(\delta(x)|x) = E^{\theta|x}[L(\theta, \delta(x))]$$

$$= \begin{cases} \int_{\Theta} L(\theta, \delta(x)) \pi(\theta|x) d\theta & \text{当 } \theta \text{ 为连续型随机变量} \\ \sum_i L(\theta_i, \delta(x)) \pi(\theta_i|x) & \text{当 } \theta \text{ 为离散型随机变量} \end{cases}$$

为决策函数 $\delta(x)$ 的后验风险。

- 若存在决策函数 $\delta^*(x)$ ，使得

$$R(\delta^*|x) = \min_{\delta} R(\delta(x)|x), \quad \text{对 } \forall \text{ 判决函数 } \delta(x), \quad (5.2.1)$$

则称 $\delta^*(x)$ 为后验风险最小准则下的最优贝叶斯决策函数。

后验风险与贝叶斯风险的关系

- 利用下列事实: $f(x, \theta) = f(x|\theta)\pi(\theta) = \pi(\theta|x)m(x)$, 将 §1.3 中贝叶斯风险 $R_\pi(\delta)$ 表达式改写如下:

$$\begin{aligned} R_\pi(\delta) &= E^\theta[R(\theta, \delta(x))] = \int_{\Theta} R(\theta, \delta(x))\pi(\theta)d\theta \\ &= \int_{\Theta} \left[\int_{\mathcal{X}} L(\theta, \delta(x))f(x|\theta)dx \right] \pi(\theta)d\theta \\ &= \int_{\mathcal{X}} \left[\int_{\Theta} L(\theta, \delta(x))\pi(\theta|x)d\theta \right] m(x)dx \\ &= \int_{\mathcal{X}} R(\delta(x)|x)m(x)dx = E^X[R(\delta(x)|x)] \end{aligned}$$

- 可见贝叶斯风险有两种表达式

$$R_\pi(\delta(x)) = E^\theta[R(\theta, \delta(x))] = E^X[R(\delta(x)|x)],$$

即将风险函数 $R(\theta, \delta(x))$ 按 θ 的先验分布 $\pi(\theta)$ 求均值, 或者将后验风险按 X 的绝对分布 (边缘分布) $m(x)$ 求均值。

后验风险与贝叶斯风险的关系

- 利用下列事实: $f(x, \theta) = f(x|\theta)\pi(\theta) = \pi(\theta|x)m(x)$, 将 §1.3 中贝叶斯风险 $R_\pi(\delta)$ 表达式改写如下:

$$\begin{aligned} R_\pi(\delta) &= E^\theta[R(\theta, \delta(x))] = \int_{\Theta} R(\theta, \delta(x))\pi(\theta)d\theta \\ &= \int_{\Theta} \left[\int_{\mathcal{X}} L(\theta, \delta(x))f(x|\theta)dx \right] \pi(\theta)d\theta \\ &= \int_{\mathcal{X}} \left[\int_{\Theta} L(\theta, \delta(x))\pi(\theta|x)d\theta \right] m(x)dx \\ &= \int_{\mathcal{X}} R(\delta(x)|x)m(x)dx = E^X[R(\delta(x)|x)] \end{aligned}$$

- 可见贝叶斯风险有两种表达式

$$R_\pi(\delta(x)) = E^\theta[R(\theta, \delta(x))] = E^X[R(\delta(x)|x)],$$

即将风险函数 $R(\theta, \delta(x))$ 按 θ 的先验分布 $\pi(\theta)$ 求均值, 或者将后验风险按 X 的绝对分布 (边缘分布) $m(x)$ 求均值。

后验风险最小的原则

- 我们将证明由 (5.2.1) 式定义的后验风险最小准则下的决策函数就是贝叶斯解. 贝叶斯解在第一章中给过定义, 它是使贝叶斯风险 $R_{\pi}(\delta)$ 达到最小的决策函数, 即如果存在 δ^* , 使得 $R_{\pi}(\delta^*) = \min_{\delta} R_{\pi}(\delta)$, 对一切决策函数 $\delta(x)$ 成立.

- 定理 5.2.1 对任何样本 x , 若存在决策函数 $\delta_{\pi}(x) \in \mathcal{D}$, 满足

$$R(\delta_{\pi}|x) = \inf_{\delta \in \mathcal{D}} R(\delta(x)|x) = \inf_{\delta \in \mathcal{D}} \int_{\Theta} L(\theta, \delta(x)) \pi(\theta|x) d\theta,$$

则 δ_{π} 为先验分布 $\pi(\theta)$ 之下的贝叶斯解.

定义 5.2.2 若 $\pi(\theta)$ 为广义先验分布, 且 $\delta_{\pi}(x)$ 是按 (5.2.1) 求得的最优决策函数, 则称 $\delta_{\pi}(x)$ 为广义贝叶斯解.

- 注 5.2.1 当 θ 的先验为广义先验分布, 定理的结果仍是对的. 此时后验风险最小准则下的决策函数, 称为广义贝叶斯解.

后验风险最小的原则

- 我们将证明由 (5.2.1) 式定义的后验风险最小准则下的决策函数就是贝叶斯解. 贝叶斯解在第一章中给过定义, 它是使贝叶斯风险 $R_\pi(\delta)$ 达到最小的决策函数, 即如果存在 δ^* , 使得 $R_\pi(\delta^*) = \min_{\delta} R_\pi(\delta)$, 对一切决策函数 $\delta(x)$ 成立.
- **定理 5.2.1** 对任何样本 x , 若存在决策函数 $\delta_\pi(x) \in \mathcal{D}$, 满足

$$R(\delta_\pi|x) = \inf_{\delta \in \mathcal{D}} R(\delta(x)|x) = \inf_{\delta \in \mathcal{D}} \int_{\Theta} L(\theta, \delta(x)) \pi(\theta|x) d\theta,$$

则 δ_π 为先验分布 $\pi(\theta)$ 之下的贝叶斯解.

定义 5.2.2 若 $\pi(\theta)$ 为广义先验分布, 且 $\delta_\pi(x)$ 是按 (5.2.1) 求得的最优决策函数, 则称 $\delta_\pi(x)$ 为广义贝叶斯解.

- **注 5.2.1** 当 θ 的先验为广义先验分布, 定理的结果仍是对的. 此时后验风险最小准则下的决策函数, 称为广义贝叶斯解.

后验风险最小的原则

- 我们将证明由 (5.2.1) 式定义的后验风险最小准则下的决策函数就是贝叶斯解. 贝叶斯解在第一章中给过定义, 它是使贝叶斯风险 $R_\pi(\delta)$ 达到最小的决策函数, 即如果存在 δ^* , 使得 $R_\pi(\delta^*) = \min_{\delta} R_\pi(\delta)$, 对一切决策函数 $\delta(x)$ 成立.
- **定理 5.2.1** 对任何样本 x , 若存在决策函数 $\delta_\pi(x) \in \mathcal{D}$, 满足

$$R(\delta_\pi|x) = \inf_{\delta \in \mathcal{D}} R(\delta(x)|x) = \inf_{\delta \in \mathcal{D}} \int_{\Theta} L(\theta, \delta(x)) \pi(\theta|x) d\theta,$$

则 δ_π 为先验分布 $\pi(\theta)$ 之下的贝叶斯解.

定义 5.2.2 若 $\pi(\theta)$ 为广义先验分布, 且 $\delta_\pi(x)$ 是按 (5.2.1) 求得的最优决策函数, 则称 $\delta_\pi(x)$ 为广义贝叶斯解.

- **注 5.2.1** 当 θ 的先验为广义先验分布, 定理的结果仍是对的. 此时后验风险最小准则下的决策函数, 称为广义贝叶斯解.

后验风险最小的原则

- 我们将证明由 (5.2.1) 式定义的后验风险最小准则下的决策函数就是贝叶斯解. 贝叶斯解在第一章中给过定义, 它是使贝叶斯风险 $R_\pi(\delta)$ 达到最小的决策函数, 即如果存在 δ^* , 使得 $R_\pi(\delta^*) = \min_{\delta} R_\pi(\delta)$, 对一切决策函数 $\delta(x)$ 成立.

- **定理 5.2.1** 对任何样本 x , 若存在决策函数 $\delta_\pi(x) \in \mathcal{D}$, 满足

$$R(\delta_\pi|x) = \inf_{\delta \in \mathcal{D}} R(\delta(x)|x) = \inf_{\delta \in \mathcal{D}} \int_{\Theta} L(\theta, \delta(x)) \pi(\theta|x) d\theta,$$

则 δ_π 为先验分布 $\pi(\theta)$ 之下的贝叶斯解.

定义 5.2.2 若 $\pi(\theta)$ 为广义先验分布, 且 $\delta_\pi(x)$ 是按 (5.2.1) 求得的最优决策函数, 则称 $\delta_\pi(x)$ 为广义贝叶斯解.

- **注 5.2.1** 当 θ 的先验为广义先验分布, 定理的结果仍是对的. 此时后验风险最小准则下的决策函数, 称为广义贝叶斯解.

一个例子

例 5.2.1 (续例 3.1.1) 通过血液可以帮助说明一个人是否患有某种疾病, 化验结果为阳性 (以 $X=1$ 表示) 或者为阴性 (以 $X=0$ 表示). 令 θ_1 表示有病, θ_2 表示无病, 记 $P(X=x|\theta) = p(x|\theta)$, 则

$$p(1|\theta_1) = 0.8, p(0|\theta_1) = 0.2, p(1|\theta_2) = 0.1, p(0|\theta_2) = 0.9,$$

设先验信息为 $\pi(\theta_1) = 0.05$, $\pi(\theta_2) = 0.95$, 此即该地区患病和不患病的比例. 知道化验结果后可能的决策行为是 a_1 , a_2 和 a_3 , 其中 a_1 表示: 治疗, a_2 表示: 不治疗, a_3 表示: 继续观察, 损失函数 $L(\theta, a)$ 如下表:

$\theta \backslash a$	a_1	a_2	a_3
θ_1	0	10	6
θ_2	4	0	2

求决策函数的后验风险和最优决策函.

解例 5.2.1

- 由例 3.1.1 可知参数 θ (只取 θ_1 和 θ_2 两个值) 的后验分布如下:

$$\begin{aligned}\pi(\theta_1|x=0) &= 0.012, & \pi(\theta_2|x=0) &= 0.988, \\ \pi(\theta_1|x=1) &= 0.296, & \pi(\theta_2|x=1) &= 0.704.\end{aligned}$$

- 故决策行为的后验风险分别为

$$\begin{aligned}R(a_1|x=0) &= E^{\theta|x}[L(a_1, \theta)] \\ &= L(a_1, \theta_1) \times \pi(\theta_1|x=0) + L(a_1, \theta_2) \times \pi(\theta_2|x=0) \\ &= 0 \times 0.012 + 4 \times 0.988 = 3.952;\end{aligned}$$

解例 5.2.1

- 由例 3.1.1 可知参数 θ (只取 θ_1 和 θ_2 两个值) 的后验分布如下:

$$\begin{aligned}\pi(\theta_1|x=0) &= 0.012, & \pi(\theta_2|x=0) &= 0.988, \\ \pi(\theta_1|x=1) &= 0.296, & \pi(\theta_2|x=1) &= 0.704.\end{aligned}$$

- 故决策行为的后验风险分别为

$$\begin{aligned}R(a_1|x=0) &= E^{\theta|x}[L(a_1, \theta)] \\ &= L(a_1, \theta_1) \times \pi(\theta_1|x=0) + L(a_1, \theta_2) \times \pi(\theta_2|x=0) \\ &= 0 \times 0.012 + 4 \times 0.988 = 3.952;\end{aligned}$$

解例 5.2.1

- 类似地有

$$R(a_2|x=0) = 10 \times 0.012 + 0 \times 0.988 = 0.12;$$

$$R(a_3|x=0) = 6 \times 0.012 + 2 \times 0.988 = 2.048;$$

因此按后验风险最小原则，当 $x = 0$ 时取最优决策函数为 $\delta^*(0) = a_2$.

- 同理，当 $x = 1$ 时可算得

$$R(a_1|x=1) = 0 \times 0.296 + 4 \times 0.704 = 2.816,$$

$$R(a_2|x=1) = 10 \times 0.296 + 0 \times 0.704 = 2.96,$$

$$R(a_3|x=1) = 6 \times 0.296 + 2 \times 0.704 = 3.184 .$$

按后验风险最小原则，当 $x = 1$ 最优决策函数为 $\delta^*(1) = a_1$.

- 因此，最优决策函数为 $\delta^*(x) = \begin{cases} a_2 & \text{当 } x = 0, \\ a_1 & \text{当 } x = 1 \end{cases}$

解例 5.2.1

- 类似地有

$$R(a_2|x=0) = 10 \times 0.012 + 0 \times 0.988 = 0.12;$$

$$R(a_3|x=0) = 6 \times 0.012 + 2 \times 0.988 = 2.048;$$

因此按后验风险最小原则，当 $x = 0$ 时取最优决策函数为 $\delta^*(0) = a_2$.

- 同理，当 $x = 1$ 时可算得

$$R(a_1|x=1) = 0 \times 0.296 + 4 \times 0.704 = 2.816,$$

$$R(a_2|x=1) = 10 \times 0.296 + 0 \times 0.704 = 2.96,$$

$$R(a_3|x=1) = 6 \times 0.296 + 2 \times 0.704 = 3.184 .$$

按后验风险最小原则，当 $x = 1$ 最优决策函数为 $\delta^*(1) = a_1$.

- 因此，最优决策函数为 $\delta^*(x) = \begin{cases} a_2 & \text{当 } x = 0, \\ a_1 & \text{当 } x = 1 \end{cases}$

例 5.2.2

某工厂的陶瓷产品每 100 件装成一箱运交运客户. 在向客户交货前面临如下两个行动的选择: a_1 : 一箱中逐一检查, a_2 : 一箱中一件也不检查. 若工厂选择行动 a_1 , 则可保证交货时每件产品都是合格品. 但因每件产品的检查费为 0.8 元, 为此工厂要支付检查费 80 元/箱, 若工厂选择行动 a_2 , 可免付每箱 80 元检查费. 但客户发现不合格品时, 按合同不允许更换, 而且每件要支付 12.5 元的赔偿费. 用 θ 表示一箱中的产品不合格率. 假如工厂决定先在每箱中随机抽取两件进行检查, 得到 $X = (X_1, X_2)$, 其中 $X_i = 0, 1$ 分别表示第 i 件产品是否合格, 显然 $X_i \sim B(1, \theta)$, $i = 1, 2$. 且 $X = X_1 + X_2 \sim B(2, \theta)$, 另外从历史资料得知, 该厂产品的不合格率 θ 不超过 0.12. 且取均匀分布 $U(0, 0.12)$ 作为 θ 的先验分布.

例 5.2.2

- (1) 求 θ 的后验分布 $\pi(\theta|x)$, $x = 0, 1, 2$.
- (2) 取损失函数

$$L(\theta, a_1) = \begin{cases} 78.4 - 1250\theta, & \text{当 } \theta \leq \theta_0 \\ 0, & \text{当 } \theta > \theta_0 \end{cases}$$
$$L(\theta, a_2) = \begin{cases} 0, & \text{当 } \theta \leq \theta_0 \\ -78.4 + 1250\theta, & \text{当 } \theta > \theta_0. \end{cases}$$

此处 $\theta_0 = 0.062$. 求决策行为 a_i ($i = 1, 2$) 的后验风险.

例 5.2.2

- (1) 求 θ 的后验分布 $\pi(\theta|x)$, $x = 0, 1, 2$.
- (2) 取损失函数

$$L(\theta, a_1) = \begin{cases} 78.4 - 1250\theta, & \text{当 } \theta \leq \theta_0 \\ 0, & \text{当 } \theta > \theta_0 \end{cases}$$
$$L(\theta, a_2) = \begin{cases} 0, & \text{当 } \theta \leq \theta_0 \\ -78.4 + 1250\theta, & \text{当 } \theta > \theta_0. \end{cases}$$

此处 $\theta_0 = 0.062$. 求决策行为 a_i ($i = 1, 2$) 的后验风险.

例 5.2.2

- (1) 求 θ 的后验分布 $\pi(\theta|x)$, $x = 0, 1, 2$.
- (2) 取损失函数

$$L(\theta, a_1) = \begin{cases} 78.4 - 1250\theta, & \text{当 } \theta \leq \theta_0 \\ 0, & \text{当 } \theta > \theta_0 \end{cases}$$
$$L(\theta, a_2) = \begin{cases} 0, & \text{当 } \theta \leq \theta_0 \\ -78.4 + 1250\theta, & \text{当 } \theta > \theta_0. \end{cases}$$

此处 $\theta_0 = 0.062$. 求决策行为 a_i ($i = 1, 2$) 的后验风险.

1 5.1 引言

2 5.2 后验风险最小原则

- 5.2.1 后验风险的定义
- 5.2.2 后验风险与贝叶斯风险的关系
- 5.2.3 后验风险最小的原则

3 5.3 一般损失函数下的贝叶斯估计

- 5.3.1 平方损失下的贝叶斯估计
- 5.3.2 加权平方损失下的贝叶斯估计
- 5.3.3 绝对损失下的贝叶斯估计
- 5.3.4 线性损失函数下的贝叶斯估计

4 5.4 假设检验和有限行动（分类）问题

- 5.4.1 假设检验问题
- 5.4.2 多行动（分类）问题
- 5.4.3 统计决策中的区间估计问题*

5 5.5 Minimax 准则*

- 5.5.1 引言及定义
- 5.5.2 Minimax 解的求法

平方损失下的贝叶斯估计

- **定理 5.3.1** 在平方损失 $L(\theta, a) = (\theta - a)^2$ 下, θ 的贝叶斯估计为后验期望值, 即

$$\hat{\theta}_B(x) = E(\theta|x).$$

- **例 5.3.1** 设 $X \sim N(\theta, \sigma^2)$, 其中 σ^2 已知. 先验分布 $\pi(\theta)$ 为 $N(\mu, \tau^2)$, μ 和 τ^2 已知, 求平方损失下 θ 的贝叶斯估计.
- **例 5.3.2** 设 $\mathbf{X} = (X_1, \dots, X_n)$ 是从 Poisson 分布 $P(\theta)$ 中抽取的简单样本. 取 θ 的先验分布为 $\Gamma(\alpha, \lambda)$, 即

$$\pi(\theta) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} \exp\{-\lambda\theta\} I_{(0,\infty)}(\theta),$$

其中 $\lambda > 0$ 已知, 求平方损失下 θ 的贝叶斯估计.

平方损失下的贝叶斯估计

- **定理 5.3.1** 在平方损失 $L(\theta, a) = (\theta - a)^2$ 下, θ 的贝叶斯估计为后验期望值, 即

$$\hat{\theta}_B(x) = E(\theta|x).$$

- **例 5.3.1** 设 $X \sim N(\theta, \sigma^2)$, 其中 σ^2 已知. 先验分布 $\pi(\theta)$ 为 $N(\mu, \tau^2)$, μ 和 τ^2 已知, 求平方损失下 θ 的贝叶斯估计.
- **例 5.3.2** 设 $\mathbf{X} = (X_1, \dots, X_n)$ 是从 Poisson 分布 $P(\theta)$ 中抽取的简单样本. 取 θ 的先验分布为 $\Gamma(\alpha, \lambda)$, 即

$$\pi(\theta) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} \exp\{-\lambda\theta\} I_{(0,\infty)}(\theta),$$

其中 $\lambda > 0$ 已知, 求平方损失下 θ 的贝叶斯估计.

平方损失下的贝叶斯估计

- **定理 5.3.1** 在平方损失 $L(\theta, a) = (\theta - a)^2$ 下, θ 的贝叶斯估计为后验期望值, 即

$$\hat{\theta}_B(x) = E(\theta|x).$$

- **例 5.3.1** 设 $X \sim N(\theta, \sigma^2)$, 其中 σ^2 已知. 先验分布 $\pi(\theta)$ 为 $N(\mu, \tau^2)$, μ 和 τ^2 已知, 求平方损失下 θ 的贝叶斯估计.
- **例 5.3.2** 设 $\mathbf{X} = (X_1, \dots, X_n)$ 是从 Poisson 分布 $P(\theta)$ 中抽取的简单样本. 取 θ 的先验分布为 $\Gamma(\alpha, \lambda)$, 即

$$\pi(\theta) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} \exp\{-\lambda\theta\} I_{(0,\infty)}(\theta),$$

其中 $\lambda > 0$ 已知, 求平方损失下 θ 的贝叶斯估计.

加权平方损失下的贝叶斯估计

- **定理 5.3.2** 在加权平方损失 $L(\theta, a) = w(\theta)(\theta - a)^2$ 下, θ 的贝叶斯估计为

$$\hat{\theta}_B = \frac{E(\theta w(\theta)|x)}{E(w(\theta)|x)},$$

其中 $w(\theta)$ 为参数空间上的正值函数.

- **例 5.3.3** 设 $\mathbf{X} = (X_1, \dots, X_n)$ 是从下列指数分布中抽取的简单样本,

$$f(x|\theta) = \theta^{-1} \exp\{-x/\theta\} I_{(0,\infty)}(x),$$

此处 $\theta > 0$. 设 θ 的先验分布服从逆伽玛分布 $\Gamma^{-1}(\alpha, \lambda)$, 即先验密度是

$$\pi(\theta) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \theta^{-(\alpha+1)} \exp\{-\lambda/\theta\} I_{(0,\infty)}(\theta).$$

求 θ 在加权平方损失 $L(\theta, \delta) = (\theta - \delta)^2/\theta^2$ 下的贝叶斯估计.

加权平方损失下的贝叶斯估计

- **定理 5.3.2** 在加权平方损失 $L(\theta, a) = w(\theta)(\theta - a)^2$ 下, θ 的贝叶斯估计为

$$\hat{\theta}_B = \frac{E(\theta w(\theta)|x)}{E(w(\theta)|x)},$$

其中 $w(\theta)$ 为参数空间上的正值函数.

- **例 5.3.3** 设 $\mathbf{X} = (X_1, \dots, X_n)$ 是从下列指数分布中抽取的简单样本,

$$f(x|\theta) = \theta^{-1} \exp\{-x/\theta\} I_{(0,\infty)}(x),$$

此处 $\theta > 0$. 设 θ 的先验分布服从逆伽玛分布 $\Gamma^{-1}(\alpha, \lambda)$, 即先验密度是

$$\pi(\theta) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \theta^{-(\alpha+1)} \exp\{-\lambda/\theta\} I_{(0,\infty)}(\theta).$$

求 θ 在加权平方损失 $L(\theta, \delta) = (\theta - \delta)^2/\theta^2$ 下的贝叶斯估计.

绝对值损失下的贝叶斯解

- **定理 5.3.3** 在绝对损失 $L(\delta, \theta) = |\delta - \theta|$ 下, θ 的贝叶斯估计为后验中位数.
- **例 5.3.4** 设 $\mathbf{X} = (X_1, \dots, X_n)$ 是从均匀分布 $U(0, \theta)$ 中抽取的简单样本, θ 的先验分布是 Pareto 分布, 其密度函数为

$$H(\theta) = 1 - (\theta_0/\theta)^\alpha, \quad \pi(\theta) = \frac{\alpha\theta_0^\alpha}{\theta^{\alpha+1}}, \quad \theta > \theta_0$$

求 θ 的在绝对值损失下的贝叶斯估计.

- **注 5.3.1** 当后验分布是单峰对称时, 后验均值也是后验中位数, 二者相同. 如在例 4.2.1 儿童智商测验的例子中, 后验分布仍为正态, 后验均值也是后验中位数, 故绝对值损失下的贝叶斯解也为 $E(\theta|x) = \mu(x)$.

绝对值损失下的贝叶斯解

- **定理 5.3.3** 在绝对损失 $L(\delta, \theta) = |\delta - \theta|$ 下, θ 的贝叶斯估计为后验中位数.
- **例 5.3.4** 设 $\mathbf{X} = (X_1, \dots, X_n)$ 是从均匀分布 $U(0, \theta)$ 中抽取的简单样本, θ 的先验分布是 Pareto 分布, 其密度函数为

$$H(\theta) = 1 - (\theta_0/\theta)^\alpha, \quad \pi(\theta) = \frac{\alpha\theta_0^\alpha}{\theta^{\alpha+1}}, \quad \theta > \theta_0$$

求 θ 的在绝对值损失下的贝叶斯估计.

- **注 5.3.1** 当后验分布是单峰对称时, 后验均值也是后验中位数, 二者相同. 如在例 4.2.1 儿童智商测验的例子中, 后验分布仍为正态, 后验均值也是后验中位数, 故绝对值损失下的贝叶斯解也为 $E(\theta|x) = \mu(x)$.

绝对值损失下的贝叶斯解

- **定理 5.3.3** 在绝对损失 $L(\delta, \theta) = |\delta - \theta|$ 下, θ 的贝叶斯估计为后验中位数.
- **例 5.3.4** 设 $\mathbf{X} = (X_1, \dots, X_n)$ 是从均匀分布 $U(0, \theta)$ 中抽取的简单样本, θ 的先验分布是 Pareto 分布, 其密度函数为

$$H(\theta) = 1 - (\theta_0/\theta)^\alpha, \quad \pi(\theta) = \frac{\alpha\theta_0^\alpha}{\theta^{\alpha+1}}, \quad \theta > \theta_0$$

求 θ 的在绝对值损失下的贝叶斯估计.

- **注 5.3.1** 当后验分布是单峰对称时, 后验均值也是后验中位数, 二者相同. 如在例 4.2.1 儿童智商测验的例子中, 后验分布仍为正态, 后验均值也是后验中位数, 故绝对值损失下的贝叶斯解也为 $E(\theta|x) = \mu(x)$.

线性损失函数下的贝叶斯估计

- **定理 5.3.4** 在线性损失函数

$$L(\theta, a) = \begin{cases} k_0(\theta - a) & \text{当 } \theta - a \geq 0 \\ k_1(a - \theta) & \text{当 } \theta - a < 0 \end{cases}$$

下, 后验风险最小准则下的贝叶斯估计为后验分布的 $\frac{k_0}{k_0 + k_1}$ 分位数.

- **例 5.3.5** (续例 4.2.1) 在估计那个儿童智商 IQ 时, 若认为低估比高估的损失高两倍, 则使用线性损失是合理的. 其损失函数为

$$L(\theta, a) = \begin{cases} 2(\theta - a) & \text{当 } \theta - a \geq 0 \\ a - \theta & \text{当 } \theta - a < 0 \end{cases}$$

求儿童智商 θ 的贝叶斯估计.

线性损失函数下的贝叶斯估计

- **定理 5.3.4** 在线性损失函数

$$L(\theta, a) = \begin{cases} k_0(\theta - a) & \text{当 } \theta - a \geq 0 \\ k_1(a - \theta) & \text{当 } \theta - a < 0 \end{cases}$$

下, 后验风险最小准则下的贝叶斯估计为后验分布的 $\frac{k_0}{k_0 + k_1}$ 分位数.

- **例 5.3.5** (续例 4.2.1) 在估计那个儿童智商 IQ 时, 若认为低估比高估的损失高两倍, 则使用线性损失是合理的. 其损失函数为

$$L(\theta, a) = \begin{cases} 2(\theta - a) & \text{当 } \theta - a \geq 0 \\ a - \theta & \text{当 } \theta - a < 0 \end{cases}$$

求儿童智商 θ 的贝叶斯估计.

1 5.1 引言

2 5.2 后验风险最小原则

- 5.2.1 后验风险的定义
- 5.2.2 后验风险与贝叶斯风险的关系
- 5.2.3 后验风险最小的原则

3 5.3 一般损失函数下的贝叶斯估计

- 5.3.1 平方损失下的贝叶斯估计
- 5.3.2 加权平方损失下的贝叶斯估计
- 5.3.3 绝对损失下的贝叶斯估计
- 5.3.4 线性损失函数下的贝叶斯估计

4 5.4 假设检验和有限行动（分类）问题

- 5.4.1 假设检验问题
- 5.4.2 多行动 (分类) 问题)
- 5.4.3 统计决策中的区间估计问题*

5 5.5 Minimax 准则*

- 5.5.1 引言及定义
- 5.5.2 Minimax 解的求法

引言

- 在估计问题中，一般有无穷多个行动可供选择. 然而有不少统计决策问题只能在有限个行动中选择. 最重要的有限行动问题是假设检验.
- 对这类问题使用贝叶斯统计决策方法是很容易解决的. 例如行动空间有 r 个行动组成，即行动空间 $\mathcal{D} = \{a_1, \dots, a_r\}$. 设在采取行动 a_i 下的损失为 $L(\theta, a_i)$, $i = 1, \dots, r$, 则贝叶斯决策就是选择使后验风险 $R(a_i|\mathbf{x}) = E^{\theta|\mathbf{x}}[L(\theta, a_i)]$ 达到最小的那个行动.
- 以下我们将分别讨论两行动（假设检验）问题和多行动（分类）问题.

引言

- 在估计问题中，一般有无穷多个行动可供选择. 然而有不少统计决策问题只能在有限个行动中选择. 最重要的有限行动问题是假设检验.
- 对这类问题使用贝叶斯统计决策方法是很容易解决的. 例如行动空间有 r 个行动组成，即行动空间 $\mathcal{D} = \{a_1, \dots, a_r\}$. 设在采取行动 a_i 下的损失为 $L(\theta, a_i)$, $i = 1, \dots, r$, 则贝叶斯决策就是选择使后验风险 $R(a_i|\mathbf{x}) = E^{\theta|\mathbf{x}}[L(\theta, a_i)]$ 达到最小的那个行动.
- 以下我们将分别讨论两行动（假设检验）问题和多行动（分类）问题.

引言

- 在估计问题中，一般有无穷多个行动可供选择. 然而有不少统计决策问题只能在有限个行动中选择. 最重要的有限行动问题是假设检验.
- 对这类问题使用贝叶斯统计决策方法是很容易解决的. 例如行动空间有 r 个行动组成，即行动空间 $\mathcal{D} = \{a_1, \dots, a_r\}$. 设在采取行动 a_i 下的损失为 $L(\theta, a_i)$, $i = 1, \dots, r$, 则贝叶斯决策就是选择使后验风险 $R(a_i|\mathbf{x}) = E^{\theta|\mathbf{x}}[L(\theta, a_i)]$ 达到最小的那个行动.
- 以下我们将分别讨论两行动（假设检验）问题和多行动（分类）问题.

检验问题：0-1 损失情形

- 设有如下的两个假设

$$H_0: \theta \in \Theta_0 \longleftrightarrow H_1: \theta \in \Theta_1, \quad \text{且 } \Theta_0 \cup \Theta_1 = \Theta.$$

决策行动有两个: a_i ($i = 0, 1$) 表示接受 H_i 的行动.

- 若为 0-1 损失

$$L(\theta, a_i) = \begin{cases} 0, & \text{若 } \theta \in \Theta_i, \\ 1, & \text{若 } \theta \notin \Theta_i, \end{cases} \quad i = 0, 1.$$

- 其后验风险为

$$R(a_0|x) = E_{\theta|x}[L(\theta, a_0)] = P(\Theta_1|x),$$

$$R(a_1|x) = E_{\theta|x}[L(\theta, a_1)] = P(\Theta_0|x).$$

- 按后验风险最小准则, 若

$$P(\Theta_1|x) \geq P(\Theta_0|x) \text{ 时否定 } H_0, \text{ 接受 } H_1.$$

因此, 贝叶斯决策就是接受具有较大后验概率的假设. 这与贝叶斯统计推断中的结论是一致的.

检验问题：0-1 损失情形

- 设有如下的两个假设

$$H_0: \theta \in \Theta_0 \longleftrightarrow H_1: \theta \in \Theta_1, \quad \text{且 } \Theta_0 \cup \Theta_1 = \Theta.$$

决策行动有两个: a_i ($i = 0, 1$) 表示接受 H_i 的行动.

- 若为 0-1 损失

$$L(\theta, a_i) = \begin{cases} 0, & \text{若 } \theta \in \Theta_i, \\ 1, & \text{若 } \theta \notin \Theta_i, \end{cases} \quad i = 0, 1.$$

- 其后验风险为

$$R(a_0|x) = E_{\theta|x}[L(\theta, a_0)] = P(\Theta_1|x),$$

$$R(a_1|x) = E_{\theta|x}[L(\theta, a_1)] = P(\Theta_0|x).$$

- 按后验风险最小准则, 若

$$P(\Theta_1|x) \geq P(\Theta_0|x) \text{ 时否定 } H_0, \text{ 接受 } H_1.$$

因此, 贝叶斯决策就是接受具有较大后验概率的假设. 这与贝叶斯统计推断中的结论是一致的.

检验问题：0-1 损失情形

- 设有如下的两个假设

$$H_0: \theta \in \Theta_0 \longleftrightarrow H_1: \theta \in \Theta_1, \quad \text{且 } \Theta_0 \cup \Theta_1 = \Theta.$$

决策行动有两个: a_i ($i = 0, 1$) 表示接受 H_i 的行动.

- 若为 0-1 损失

$$L(\theta, a_i) = \begin{cases} 0, & \text{若 } \theta \in \Theta_i, \\ 1, & \text{若 } \theta \notin \Theta_i, \end{cases} \quad i = 0, 1.$$

- 其后验风险为

$$R(a_0|x) = E_{\theta|x}[L(\theta, a_0)] = P(\Theta_1|x),$$

$$R(a_1|x) = E_{\theta|x}[L(\theta, a_1)] = P(\Theta_0|x).$$

- 按后验风险最小准则, 若

$$P(\Theta_1|x) \geq P(\Theta_0|x) \text{ 时否定 } H_0, \text{ 接受 } H_1.$$

因此, 贝叶斯决策就是接受具有较大后验概率的假设. 这与贝叶斯统计推断中的结论是一致的.

检验问题：0-1 损失情形

- 设有如下的两个假设

$$H_0: \theta \in \Theta_0 \longleftrightarrow H_1: \theta \in \Theta_1, \quad \text{且 } \Theta_0 \cup \Theta_1 = \Theta.$$

决策行动有两个: a_i ($i = 0, 1$) 表示接受 H_i 的行动.

- 若为 0-1 损失

$$L(\theta, a_i) = \begin{cases} 0, & \text{若 } \theta \in \Theta_i, \\ 1, & \text{若 } \theta \notin \Theta_i, \end{cases} \quad i = 0, 1.$$

- 其后验风险为

$$R(a_0|x) = E_{\theta|x}[L(\theta, a_0)] = P(\Theta_1|x),$$

$$R(a_1|x) = E_{\theta|x}[L(\theta, a_1)] = P(\Theta_0|x).$$

- 按后验风险最小准则, 若

$$P(\Theta_1|x) \geq P(\Theta_0|x) \text{ 时否定 } H_0, \text{ 接受 } H_1.$$

因此, 贝叶斯决策就是接受具有较大后验概率的假设. 这与贝叶斯统计推断中的结论是一致的.

检验问题： $0 - k_i$ 损失情形

- 若为 $0 - k_i$ 损失

$$L(\theta, a_i) = \begin{cases} 0, & \text{若 } \theta \in \Theta_i ; \\ k_i, & \text{若 } \theta \notin \Theta_i ; \end{cases} \quad i = 0, 1.$$

- 此时后验风险分别为

$$\begin{aligned} R(a_0|x) &= E_{\theta|x}[L(\theta, a_0)] = k_0 P(\Theta_1|x), \\ R(a_1|x) &= E_{\theta|x}[L(\theta, a_1)] = k_1 P(\Theta_0|x). \end{aligned}$$

- 按后验风险最小准则, 若

$$P(\Theta_1|x) \geq \frac{k_1}{k_0 + k_1} \quad \text{时否定 } H_0, \text{ 否则接受 } H_0.$$

检验问题： $0 - k_i$ 损失情形

- 若为 $0 - k_i$ 损失

$$L(\theta, a_i) = \begin{cases} 0, & \text{若 } \theta \in \Theta_i ; \\ k_i, & \text{若 } \theta \notin \Theta_i ; \end{cases} \quad i = 0, 1.$$

- 此时后验风险分别为

$$R(a_0|x) = E_{\theta|x}[L(\theta, a_0)] = k_0 P(\Theta_1|x),$$

$$R(a_1|x) = E_{\theta|x}[L(\theta, a_1)] = k_1 P(\Theta_0|x).$$

- 按后验风险最小准则, 若

$$P(\Theta_1|x) \geq \frac{k_1}{k_0 + k_1} \quad \text{时否定 } H_0, \text{ 否则接受 } H_0.$$

检验问题： $0 - k_i$ 损失情形

- 若为 $0 - k_i$ 损失

$$L(\theta, a_i) = \begin{cases} 0, & \text{若 } \theta \in \Theta_i ; \\ k_i, & \text{若 } \theta \notin \Theta_i ; \end{cases} \quad i = 0, 1.$$

- 此时后验风险分别为

$$\begin{aligned} R(a_0|x) &= E_{\theta|x}[L(\theta, a_0)] = k_0 P(\Theta_1|x), \\ R(a_1|x) &= E_{\theta|x}[L(\theta, a_1)] = k_1 P(\Theta_0|x). \end{aligned}$$

- 按后验风险最小准则, 若

$$P(\Theta_1|x) \geq \frac{k_1}{k_0 + k_1} \quad \text{时否定 } H_0, \text{ 否则接受 } H_0.$$

例子

- **例 5.4.1** (续例 4.2.1) 在儿童智商问题的例子中, 设测验结果 $X \sim N(\theta, 100)$, 其中 θ 为这个孩子测验中的智商 IQ 真值, θ 的先验分布为 $N(100, 225)$. 该儿童测验得分 $x = 115$. 取损失函数为 $0-1$ 损失, 求下列检验问题:

$$H_0 : \theta \leq 105 \leftrightarrow H_1 : \theta > 105.$$

- **例 5.4.2** 设 $X \sim N(\theta, \sigma^2)$, 其中 σ^2 已知. θ 的先验分布为 $N(\mu, \tau^2)$, 损失函数为 " $0-k_i$ " 损失. 令 X_1, \dots, X_n 为从总体 X 中抽取的简单样本. 求检验问题:

$$H_0 : \theta \geq \theta_0 \longleftrightarrow H_1 : \theta < \theta_0.$$

例子

- **例 5.4.1** (续例 4.2.1) 在儿童智商问题的例子中, 设测验结果 $X \sim N(\theta, 100)$, 其中 θ 为这个孩子测验中的智商 IQ 真值, θ 的先验分布为 $N(100, 225)$. 该儿童测验得分 $x = 115$. 取损失函数为 $0-1$ 损失, 求下列检验问题:

$$H_0 : \theta \leq 105 \leftrightarrow H_1 : \theta > 105.$$

- **例 5.4.2** 设 $X \sim N(\theta, \sigma^2)$, 其中 σ^2 已知. θ 的先验分布为 $N(\mu, \tau^2)$, 损失函数为 " $0-k_i$ " 损失. 令 X_1, \dots, X_n 为从总体 X 中抽取的简单样本. 求检验问题:

$$H_0 : \theta \geq \theta_0 \longleftrightarrow H_1 : \theta < \theta_0.$$

多行动(分类)问题

- 很多决策问题可能采取的行动多于两个. 例如在假设检验问题中常常存在两者皆可的区域. 即除了 $\theta \in \Theta_0$ 及 $\theta \in \Theta_1$ 分别采取行动 a_0 和 a_1 之外, 还存在第三个行动 a_2 , 它表示当 $\theta \in \Theta_2$ 时采取两者皆可的行动. 例如, 若要求检验两种药物的治愈率, 合理方法是检验下列三个假设:

$H_0: \theta_1 - \theta_2 < -\varepsilon$, $H_1: \theta_1 - \theta_2 > \varepsilon$, $H_2: |\theta_1 - \theta_2| \leq \varepsilon$,
其中 $\varepsilon > 0$ 选择使得当 $|\theta_1 - \theta_2| \leq \varepsilon$ 时两种药物被认为等效.

- 即使在经典的假设检验中, 也有三个行动可供选择: a_0 表示接受 H_0 , a_1 表示拒绝 H_0 , 而 a_2 表示接受 H_0 或 H_1 都无足够的证据. 经典方法是通过犯错误概率来做选择. 下面将利用贝叶斯统计决策方法来研究, 采用后验风险最小的原则.
- 常见的有限行动问题的另一个类型是分类问题. 获得观测值后, 将未知参数分到几个可能的区域中去, 这与前面的多行动检验类似, 采用的准则仍是后验风险最小的原则.

多行动(分类)问题

- 很多决策问题可能采取的行动多于两个. 例如在假设检验问题中常常存在两者皆可的区域. 即除了 $\theta \in \Theta_0$ 及 $\theta \in \Theta_1$ 分别采取行动 a_0 和 a_1 之外, 还存在第三个行动 a_2 , 它表示当 $\theta \in \Theta_2$ 时采取两者皆可的行动. 例如, 若要求检验两种药物的治愈率, 合理方法是检验下列三个假设:

$H_0: \theta_1 - \theta_2 < -\varepsilon$, $H_1: \theta_1 - \theta_2 > \varepsilon$, $H_2: |\theta_1 - \theta_2| \leq \varepsilon$,
其中 $\varepsilon > 0$ 选择使得当 $|\theta_1 - \theta_2| \leq \varepsilon$ 时两种药物被认为等效.

- 即使在经典的假设检验中, 也有三个行动可供选择: a_0 表示接受 H_0 , a_1 表示拒绝 H_0 , 而 a_2 表示接受 H_0 或 H_1 都无足够的证据. 经典方法是通过犯错误概率来做选择. 下面将利用贝叶斯统计决策方法来研究, 采用后验风险最小的原则.
- 常见的有限行动问题的另一个类型是分类问题. 获得观测值后, 将未知参数分到几个可能的区域中去, 这与前面的多行动检验类似, 采用的准则仍是后验风险最小的原则.

多行动(分类)问题

- 很多决策问题可能采取的行动多于两个. 例如在假设检验问题中常常存在两者皆可的区域. 即除了 $\theta \in \Theta_0$ 及 $\theta \in \Theta_1$ 分别采取行动 a_0 和 a_1 之外, 还存在第三个行动 a_2 , 它表示当 $\theta \in \Theta_2$ 时采取两者皆可的行动. 例如, 若要求检验两种药物的治愈率, 合理方法是检验下列三个假设:

$H_0: \theta_1 - \theta_2 < -\varepsilon$, $H_1: \theta_1 - \theta_2 > \varepsilon$, $H_2: |\theta_1 - \theta_2| \leq \varepsilon$,
其中 $\varepsilon > 0$ 选择使得当 $|\theta_1 - \theta_2| \leq \varepsilon$ 时两种药物被认为等效.

- 即使在经典的假设检验中, 也有三个行动可供选择: a_0 表示接受 H_0 , a_1 表示拒绝 H_0 , 而 a_2 表示接受 H_0 或 H_1 都无足够的证据. 经典方法是通过犯错误概率来做选择. 下面将利用贝叶斯统计决策方法来研究, 采用后验风险最小的原则.
- 常见的有限行动问题的另一个类型是分类问题. 获得观测值后, 将未知参数分到几个可能的区域中去, 这与前面的多行动检验类似, 采用的准则仍是后验风险最小的原则.

例子

例 5.4.3 (续例 4.2.1) 在儿童智商 IQ 测试问题的例子中, 对那个孩子的智商作出如下三个假设:

$$H_1: \theta < 90, \quad H_2: 90 \leq \theta \leq 110, \quad H_3: \theta > 110. \quad (5.4.6)$$

设有三个行动: a_i ($i = 1, 2, 3$) 表示接受 H_i , 取下列损失函数:

$$L(\theta, a_1) = \begin{cases} 0 & \text{当 } \theta < 90 \\ \theta - 90 & \text{当 } 90 \leq \theta \leq 110 \\ 2(\theta - 90) & \text{当 } \theta > 110 \end{cases}$$

$$L(\theta, a_2) = \begin{cases} 90 - \theta & \text{当 } \theta < 90 \\ 0 & \text{当 } 90 \leq \theta \leq 110 \\ \theta - 110 & \text{当 } \theta > 110 \end{cases}$$

$$L(\theta, a_3) = \begin{cases} 2(110 - \theta) & \text{当 } \theta < 90 \\ 110 - \theta & \text{当 } 90 \leq \theta \leq 110 \\ 0 & \text{当 } \theta > 110. \end{cases}$$

已知后验分布 $\pi(\theta|x)$ 是 $N(110.38, 8.32^2)$, 求检验问题 (5.4.6).

例子

例 5.4.3 (续例 4.2.1) 在儿童智商 IQ 测试问题的例子中, 对那个孩子的智商作出如下三个假设:

$$H_1: \theta < 90, \quad H_2: 90 \leq \theta \leq 110, \quad H_3: \theta > 110. \quad (5.4.6)$$

设有三个行动: a_i ($i = 1, 2, 3$) 表示接受 H_i , 取下列损失函数:

$$L(\theta, a_1) = \begin{cases} 0 & \text{当 } \theta < 90 \\ \theta - 90 & \text{当 } 90 \leq \theta \leq 110 \\ 2(\theta - 90) & \text{当 } \theta > 110 \end{cases}$$

$$L(\theta, a_2) = \begin{cases} 90 - \theta & \text{当 } \theta < 90 \\ 0 & \text{当 } 90 \leq \theta \leq 110 \\ \theta - 110 & \text{当 } \theta > 110 \end{cases}$$

$$L(\theta, a_3) = \begin{cases} 2(110 - \theta) & \text{当 } \theta < 90 \\ 110 - \theta & \text{当 } 90 \leq \theta \leq 110 \\ 0 & \text{当 } \theta > 110. \end{cases}$$

已知后验分布 $\pi(\theta|x)$ 是 $N(110.38, 8.32^2)$, 求检验问题 (5.4.6).

区间估计问题

- 区间估计问题也可以用统计决策的方法去考虑. 为简单计, 设 $C(x) = (d_1(x), d_2(x))$ 为 θ 的一个区间估计.
- 损失函数的一种取法为

$$L(\theta, C(x)) = m_1[d_2(x) - d_1(x)] + m_2[1 - I_{C(x)}(\theta)],$$

此处 $m_1 > 0$, $m_2 > 0$ 为给定常数. 显见, 第一部分表示区间长度引起的损失, 第二部分表示当 θ 不属于 $C(x)$ 引起的损失.

- 按后验风险最小的原则, 应使区间估计的后验风险

$$\begin{aligned} R(C(x)|x) &= E_{\theta|x}[L(\theta, C(x))] \\ &= m_1(d_2(x) - d_1(x)) + m_2 P_{\theta|x}(\theta \notin C(x)) \end{aligned}$$

越小越好. 但是要找出最优解, 并非易事. 优化问题能够得以解决的不多.

区间估计问题

- 区间估计问题也可以用统计决策的方法去考虑. 为简单计, 设 $C(x) = (d_1(x), d_2(x))$ 为 θ 的一个区间估计.
- 损失函数的一种取法为

$$L(\theta, C(x)) = m_1[d_2(x) - d_1(x)] + m_2[1 - I_{C(x)}(\theta)],$$

此处 $m_1 > 0$, $m_2 > 0$ 为给定常数. 显见, 第一部分表示区间长度引起的损失, 第二部分表示当 θ 不属于 $C(x)$ 引起的损失.

- 按后验风险最小的原则, 应使区间估计的后验风险

$$\begin{aligned} R(C(x)|x) &= E_{\theta|x}[L(\theta, C(x))] \\ &= m_1(d_2(x) - d_1(x)) + m_2 P_{\theta|x}(\theta \notin C(x)) \end{aligned}$$

越小越好. 但是要找出最优解, 并非易事. 优化问题能够得以解决的不多.

区间估计问题

- 区间估计问题也可以用统计决策的方法去考虑. 为简单计, 设 $C(x) = (d_1(x), d_2(x))$ 为 θ 的一个区间估计.
- 损失函数的一种取法为

$$L(\theta, C(x)) = m_1[d_2(x) - d_1(x)] + m_2[1 - I_{C(x)}(\theta)],$$

此处 $m_1 > 0$, $m_2 > 0$ 为给定常数. 显见, 第一部分表示区间长度引起的损失, 第二部分表示当 θ 不属于 $C(x)$ 引起的损失.

- 按后验风险最小的原则, 应使区间估计的后验风险

$$\begin{aligned} R(C(x)|x) &= E_{\theta|x}[L(\theta, C(x))] \\ &= m_1(d_2(x) - d_1(x)) + m_2 P_{\theta|x}(\theta \notin C(x)) \end{aligned}$$

越小越好. 但是要找出最优解, 并非易事. 优化问题能够得以解决的不多.

1 5.1 引言

2 5.2 后验风险最小原则

- 5.2.1 后验风险的定义
- 5.2.2 后验风险与贝叶斯风险的关系
- 5.2.3 后验风险最小的原则

3 5.3 一般损失函数下的贝叶斯估计

- 5.3.1 平方损失下的贝叶斯估计
- 5.3.2 加权平方损失下的贝叶斯估计
- 5.3.3 绝对损失下的贝叶斯估计
- 5.3.4 线性损失函数下的贝叶斯估计

4 5.4 假设检验和有限行动（分类）问题

- 5.4.1 假设检验问题
- 5.4.2 多行动（分类）问题
- 5.4.3 统计决策中的区间估计问题*

5 5.5 Minimax 准则*

- 5.5.1 引言及定义
- 5.5.2 Minimax 解的求法

引言

- 前一节已经说过, 一致最优的决策函数通常不存在, 因此我们必须把标准放宽些, 引进一些比一致最优准则更弱的优良性准则.
- 途径之一, 是用某种方法制定优良性的综合指标, 以之作为比较的标准. 贝叶斯准则属于这一类. 和贝叶斯风险 $R_{\pi}(\delta)$ 一样, 下列定义的最大风险 $M(\delta)$ 也是一种优良性的综合指标, 用它作为比较决策函数的标准, 称为 Minimax 准则. 因此 Minimax 准则是从综合指标考虑的另一种优良性准则.
- 设 δ 为一决策函数, $R(\theta, \delta)$ 为其风险函数, 令

$$M(\delta) = \sup_{\theta \in \Theta} R(\theta, \delta).$$

易见 $M(\delta)$ 表示采用 δ 时所遭受的最大风险. 如果在某项应用中, 使这个最大风险尽可能小是很重要的话, 我们就可以制定如下准则, 通常称为 Minimax 准则或极小极大准则.

引言

- 前一节已经说过, 一致最优的决策函数通常不存在, 因此我们必须把标准放宽些, 引进一些比一致最优准则更弱的优良性准则.
- 途径之一, 是用某种方法制定优良性的综合指标, 以之作为比较的标准. 贝叶斯准则属于这一类. 和贝叶斯风险 $R_{\pi}(\delta)$ 一样, 下列定义的最大风险 $M(\delta)$ 也是一种优良性的综合指标, 用它作为比较决策函数的标准, 称为 **Minimax 准则**. 因此 **Minimax 准则**是从综合指标考虑的另一种优良性准则.
- 设 δ 为一决策函数, $R(\theta, \delta)$ 为其风险函数, 令

$$M(\delta) = \sup_{\theta \in \Theta} R(\theta, \delta).$$

易见 $M(\delta)$ 表示采用 δ 时所遭受的最大风险. 如果在某项应用中, 使这个最大风险尽可能小是很重要的话, 我们就可以制定如下准则, 通常称为 **Minimax 准则**或**极小极大准则**.

引言

- 前一节已经说过, 一致最优的决策函数通常不存在, 因此我们必须把标准放宽些, 引进一些比一致最优准则更弱的优良性准则.
- 途径之一, 是用某种方法制定优良性的综合指标, 以之作为比较的标准. 贝叶斯准则属于这一类. 和贝叶斯风险 $R_{\pi}(\delta)$ 一样, 下列定义的最大风险 $M(\delta)$ 也是一种优良性的综合指标, 用它作为比较决策函数的标准, 称为 **Minimax 准则**. 因此 **Minimax 准则**是从综合指标考虑的另一种优良性准则.
- 设 δ 为一决策函数, $R(\theta, \delta)$ 为其风险函数, 令

$$M(\delta) = \sup_{\theta \in \Theta} R(\theta, \delta).$$

易见 $M(\delta)$ 表示采用 δ 时所遭受的最大风险. 如果在某项应用中, 使这个最大风险尽可能小是很重要的话, 我们就可以制定如下准则, 通常称为 **Minimax 准则**或**极小极大准则**.

Minimax 准则的定义

- **定义 5.5.1** 设 δ_1 和 δ_2 为同一个统计决策问题的两个决策函数, 若 $M(\delta_1) < M(\delta_2)$, 则称决策函数 δ_1 优于 δ_2 . 如果存在某个决策函数 δ^* , 对任何决策函数 δ 都有

$$M(\delta^*) \leq M(\delta)$$

则称 δ^* 为该统计决策问题的 Minimax 解, 或称为 Minimax 决策函数. 当统计决策问题为估计或检验时, 称 δ^* 为 Minimax 估计或 Minimax 检验.

- **注 5.5.1** 以最大风险的大小作为评判决策函数的标准, 是考虑最不利的情形, 使最不利情形尽可能地好. 因此 Minimax 准则是一种偏保守的准则.
- 在实际中常使用 Minimax 准则这种策略思想作决策. 形象地说, 这一准则“不求得到很多, 但求不失去很多”. 例如, 在地震多发地区, 重要建筑物的建筑设计按 Minimax 准则, 要求在可抗八级地震的条件下, 尽量减少建造费用.

Minimax 准则的定义

- **定义 5.5.1** 设 δ_1 和 δ_2 为同一个统计决策问题的两个决策函数, 若 $M(\delta_1) < M(\delta_2)$, 则称决策函数 δ_1 优于 δ_2 . 如果存在某个决策函数 δ^* , 对任何决策函数 δ 都有

$$M(\delta^*) \leq M(\delta)$$

则称 δ^* 为该统计决策问题的 Minimax 解, 或称为 Minimax 决策函数. 当统计决策问题为估计或检验时, 称 δ^* 为 Minimax 估计或 Minimax 检验.

- **注 5.5.1** 以最大风险的大小作为评判决策函数的标准, 是考虑最不利的情形, 使最不利情形尽可能地好. 因此 Minimax 准则是一种偏保守的准则.
- 在实际中常使用 Minimax 准则这种策略思想作决策. 形象地说, 这一准则“不求得到很多, 但求不失去很多”. 例如, 在地震多发地区, 重要建筑物的建筑设计按 Minimax 准则, 要求在可抗八级地震的条件下, 尽量减少建造费用.

Minimax 准则的定义

- **定义 5.5.1** 设 δ_1 和 δ_2 为同一个统计决策问题的两个决策函数, 若 $M(\delta_1) < M(\delta_2)$, 则称决策函数 δ_1 优于 δ_2 . 如果存在某个决策函数 δ^* , 对任何决策函数 δ 都有

$$M(\delta^*) \leq M(\delta)$$

则称 δ^* 为该统计决策问题的 Minimax 解, 或称为 Minimax 决策函数. 当统计决策问题为估计或检验时, 称 δ^* 为 Minimax 估计或 Minimax 检验.

- **注 5.5.1** 以最大风险的大小作为评判决策函数的标准, 是考虑最不利的情形, 使最不利情形尽可能地好. 因此 Minimax 准则是一种偏保守的准则.
- 在实际中常使用 Minimax 准则这种策略思想作决策. 形象地说, 这一准则“不求得到很多, 但求不失去很多”. 例如, 在地震多发地区, 重要建筑物的建筑设计按 Minimax 准则, 要求在可抗八级地震的条件下, 尽量减少建造费用.

注释及图

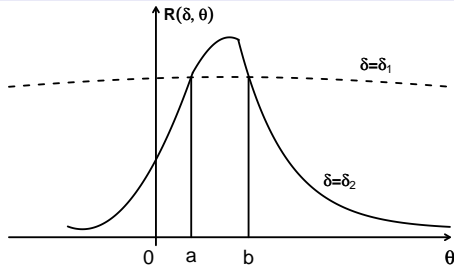


图 5.5.1 风险函数与 Minimax 准则

常有如图 5.5.1 的情形, 其中 $M(\delta_1) < M(\delta_2)$, 故按 Minimax 准则, δ_1 优于 δ_2 . 但细看二者的风险函数, 发现对大多数 θ 而言, δ_2 优于 δ_1 , 仅当 $a < \theta < b$ 时, δ_1 优于 δ_2 .

如果没有足够的先验信息说明 θ 处在 a, b 之间, 就很难说 δ_1 优于 δ_2 了. 故贝叶斯学派认为, 只在人们对 θ 的先验分布很没把握时, 作为一种替代, 才使用 Minimax 解. 只要对先验分布有一定把握, 则宁肯用贝叶斯准则.

Minimax 解的求解方法一

- 求 Minimax 解通常比较困难, 下列两个定理与其说是求 Minimax 解的方法, 不如说是验证某一特定的解为 Minimax 解的方法.
- 定理 5.5.1 设 δ^* 为在先验分布 $H(\theta)$ 之下的贝叶斯解, 且 δ^* 的风险函数为常数 c , 即 $R(\delta^*, \theta) = c$ 对任何 $\theta \in \Theta$, 则 δ^* 为一个 Minimax 解.
- 例 5.5.1 设 $X \sim B(n, \theta)$, θ 的先验分布是 $Be(a, b)$. 损失函数为 $L(\theta, d) = (\theta - d)^2$, 求 θ 的 Minimax 估计.

Minimax 解的求解方法一

- 求 Minimax 解通常比较困难, 下列两个定理与其说是求 Minimax 解的方法, 不如说是验证某一特定的解为 Minimax 解的方法.
- **定理 5.5.1** 设 δ^* 为在先验分布 $H(\theta)$ 之下的贝叶斯解, 且 δ^* 的风险函数为常数 c , 即 $R(\delta^*, \theta) = c$ 对任何 $\theta \in \Theta$, 则 δ^* 为一个 Minimax 解.
- **例 5.5.1** 设 $X \sim B(n, \theta)$, θ 的先验分布是 $Be(a, b)$. 损失函数为 $L(\theta, d) = (\theta - d)^2$, 求 θ 的 Minimax 估计.

Minimax 解的求解方法一

- 求 Minimax 解通常比较困难, 下列两个定理与其说是求 Minimax 解的方法, 不如说是验证某一特定的解为 Minimax 解的方法.
- **定理 5.5.1** 设 δ^* 为在先验分布 $H(\theta)$ 之下的贝叶斯解, 且 δ^* 的风险函数为常数 c , 即 $R(\delta^*, \theta) = c$ 对任何 $\theta \in \Theta$, 则 δ^* 为一个 Minimax 解.
- **例 5.5.1** 设 $X \sim B(n, \theta)$, θ 的先验分布是 $Be(a, b)$. 损失函数为 $L(\theta, d) = (\theta - d)^2$, 求 θ 的 Minimax 估计.

Minimax 解的求解方法二

- 定理 5.5.1 的使用面较窄, 因为一般很难找到一个其风险函数为常数的贝叶斯解. 下面定理的应用要广泛的多.
- 定理 5.5.2 设 δ_k 为一个统计决策问题在先验分布 H_k 之下的贝叶斯解, 假定 δ_k 的贝叶斯风险为 r_k , $k = 1, 2, \dots$, 且有

$$\lim_{k \rightarrow \infty} r_k = r < \infty.$$

又设 δ^* 为同一问题的一个决策函数, 满足条件 $M(\delta^*) \leq r$, 则 δ^* 为此决策问题的 Minimax 解.

- 例 5.5.2 设 $\mathbf{X} = (X_1, \dots, X_n)$ 是从正态总体 $N(\theta, 1)$ 中抽取的简单样本, 取损失函数为 $L(\theta, d) = (\theta - d)^2$, 求 θ 的 Minimax 估计.

Minimax 解的求解方法二

- 定理 5.5.1 的使用面较窄，因为一般很难找到一个其风险函数为常数的贝叶斯解。下面定理的应用要广泛的多。
- 定理 5.5.2 设 δ_k 为一个统计决策问题在先验分布 H_k 之下的贝叶斯解，假定 δ_k 的贝叶斯风险为 r_k , $k = 1, 2, \dots$ ，且有

$$\lim_{k \rightarrow \infty} r_k = r < \infty.$$

又设 δ^* 为同一问题的一个决策函数，满足条件 $M(\delta^*) \leq r$ ，则 δ^* 为此决策问题的 Minimax 解。

- 例 5.5.2 设 $\mathbf{X} = (X_1, \dots, X_n)$ 是从正态总体 $N(\theta, 1)$ 中抽取的简单样本，取损失函数为 $L(\theta, d) = (\theta - d)^2$ ，求 θ 的 Minimax 估计。

Minimax 解的求解方法二

- 定理 5.5.1 的使用面较窄, 因为一般很难找到一个其风险函数为常数的贝叶斯解. 下面定理的应用要广泛的多.
- 定理 5.5.2 设 δ_k 为一个统计决策问题在先验分布 H_k 之下的贝叶斯解, 假定 δ_k 的贝叶斯风险为 r_k , $k = 1, 2, \dots$, 且有

$$\lim_{k \rightarrow \infty} r_k = r < \infty.$$

又设 δ^* 为同一问题的一个决策函数, 满足条件 $M(\delta^*) \leq r$, 则 δ^* 为此决策问题的 Minimax 解.

- 例 5.5.2 设 $\mathbf{X} = (X_1, \dots, X_n)$ 是从正态总体 $N(\theta, 1)$ 中抽取的简单样本, 取损失函数为 $L(\theta, d) = (\theta - d)^2$, 求 θ 的 Minimax 估计.