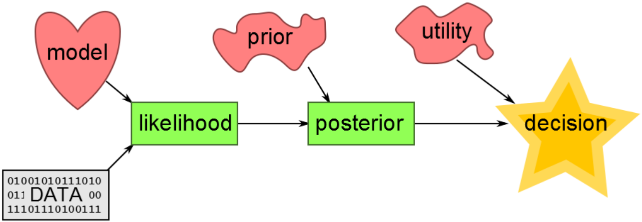
第 5 讲：贝叶斯决策

张伟平

目录

5.1 贝叶斯统计决策



- 模型 *model* 是贝叶斯统计决策的核心. *model* 一般基于刻画现实世界的科学理论, 其为一个随机模型, 具有可调节的参数, 可以用来计算观测到特定结果的概率.(似然函数)

- 先验 *prior* 表示了没有看到数据前对参数分布的认识 (先验分布)

- 效用函数 *utility function* 或损失函数 *loss function* 用来评估给定参数值后, 采取一个行动的后果. 效用和损失互为相反数. 商业领域常用效用, 因为要最大化利润. 科学领域则常用损失, 因为最小化误差是感兴趣所在. (损失函数)

- 贝叶斯统计决策综合了模型, 先验和损失. 先验和损失是主观化的, 不同人可能使用不同的假设.

统计决策理论

Abraham Wald

Wald 的统计决策理论为点估计，区间估计和假设检验等各种问题提供了一个统一的理论框架.

根据统计决策过程，常把**样本空间和样本分布族**、**行动空间**和**损失函数**称为统计决策三要素.

- **样本空间和样本分布族:** 取值于样本空间 $\mathscr{X}$ 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. 这里 $F_\theta(x)$ 是 X 的分布函数, θ 是未知参数, Θ 为参数空间, $\mathbf{X}=(X_1,\cdots,X_n)$ 是总体 X 的简单样本.

- **行动空间:** 决策者或统计工作者对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 $\mathscr{A}$. 在估计问题中, $\mathscr{A}$ 由一切估计量 $\delta(x)$ 构成, 常取 $\mathscr{A}=\Theta$. 在检验问题中 $\mathscr{A}$ 只有两个行动组成, 即 $\mathscr{A}=\{a_0,a_1\}$, 其中 a_0 表示接受原假设 H_0, a_1 表示拒绝 H_0.

- **损失函数:** 损失函数是定义于 $\Theta\times\mathscr{A}$ 上的非负函数, 记为 $L(\theta,a)$. 它表示参数为 θ 时采取行动 $a\in\mathscr{A}$ 所蒙受的损失. 常用的有平方损失 $L(\theta,a)=(a-\theta)^2$、绝对值损失 $L(\theta,a)=|\theta-a|$ 和

线性损失

$$L(\theta,a)=\begin{cases} k_1(\theta-a) & \theta\ge a\\ k_2(a-\theta) & \theta<a\end{cases}$$

■ **例 4** LINEX 损失函数定义为

$$L(\theta,a)=\exp\{c(a-\theta)\}-c(a-\theta)-1, c\in R$$

对 $c>0$, 损失函数 $L(\theta,a)$ 关于 0 非常不对称, 高估的成本比低估要高很多. 当 $|a-\theta|\to\infty$, 损失 $L(\theta,a)$ 在 $a-\theta>0$ 时指数增加, 在 $a-\theta<0$ 时几乎线性增加; 对 $c<0$, 线性-指数增加现象刚好反过来. 而且当 $|a-\theta|$ 非常小时候, $L(\theta,a)$ 接近 $c(a-\theta)^2/2$.

- 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 a 使得引起的平均损失最小.

5.2 风险函数和一致最优决策函数

风险函数

定义于样本空间 $\mathscr{X}$ 内而取值于行动空间 $\mathscr{A}$ 内的函数 $\delta = \delta(\mathbf{x})$ 称为**决策函数**或**判决函数**.

Definition

若 δ 是决策行动, 参数为 θ, 损失函数是 $L(\theta, \delta)$, 这个量与样本 $\mathbf{X}$ 有关, 因而是随机的. 故采取行动 δ 的效果用平均损失去度量是相对合理的. 这就引入如下风险函数的概念.

频率风险函数

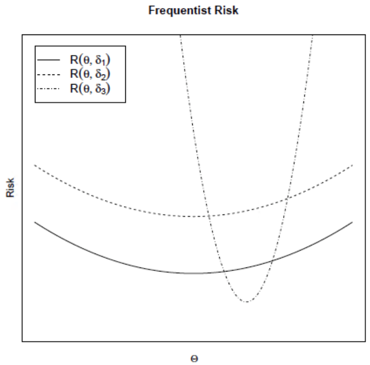
设 $\delta(\mathbf{x})$ 是一个决策函数，称平均损失

$$R(\theta,\delta)=E[L(\theta,\delta(\mathbf{X}))]=\int_{\mathscr{X}}L(\theta,\delta(\mathbf{x}))dF(\mathbf{x}|\theta)$$

$$=\begin{cases}\int_{\mathscr{X}}L(\theta,\delta(\mathbf{x}))f(\mathbf{x}|\theta)d\mathbf{x}, & \text{当 } X \text{ 为连续随机变量,}\\ \sum\limits_{\mathbf{x}\in\mathscr{X}}L(\theta,\delta(\mathbf{x}))f(\mathbf{x}|\theta), & \text{当 } X \text{ 为离散随机变量.}\end{cases}$$

为 δ 的风险函数，此处 $F(\mathbf{x}|\theta)$ 是给定 θ 时 $\mathbf{X}$ 的分布函数.

Definition



三种不同决策函数的风险函数. 那个估计能够控制其他估计?

一致最优决策函数

按照 Wald 的统计决策理论，评价一个决策函数的唯一依据，就是其风险函数. 风险函数愈小愈好.

设 $\delta_1(\mathbf{x})$ 和 $\delta_2(\mathbf{x})$ 为 θ 的两个不同的决策函数，若 $R(\theta,\delta_1(\mathbf{x})) \leq R(\theta,\delta_2(\mathbf{x}))$，对一切 $\theta \in \Theta$，且至少存在一个 $\theta_0 \in \Theta$ 使严格不等号成立，则称 $\delta_1(\mathbf{x})$ 优于 $\delta_2(\mathbf{x})$. 若存在 $\delta^*(\mathbf{x})$，使得对任一决策函数 $\delta(\mathbf{x})$ 有

$$R(\theta,\delta^*(\mathbf{x})) \leq R(\theta,\delta(\mathbf{x})), \quad 对一切\ \theta \in \Theta,$$

则称 $\delta^*(\mathbf{x})$ 为一致最优解或一致最优决策函数.

Definition

对决策函数 $\delta(\mathbf{x})$, 若不存在一致优于它的决策函数, 则称 $\delta(\mathbf{x})$ 为可容许的决策函数.

Definition

由于要求一致最优性, 可容许的决策往往并不存在. 人们转而讨论 minimax 决策和贝叶斯决策.

如果存在 $\delta_0(\mathbf{x})$ 满足 $R(\theta, \delta_0(\mathbf{x})) = \inf_\delta \sup_\theta R(\theta, \delta)$, 则称 $\delta_0(\mathbf{x})$ 为一 minimax 决策.

Definition

5.2.1 贝叶斯风险和贝叶斯解

贝叶斯风险

设 $R(\theta,\delta(\mathbf{x}))$ 为风险函数，$\pi(\theta)$ 为 θ 的先验分布，则称

$$R_\pi(\delta(\mathbf{x})) = \int_\Theta R(\theta,\delta(\mathbf{x}))dF^\pi(\theta) = E^\pi[R(\theta,\delta(\mathbf{X}))]$$

$$= \int_\Theta \int_{\mathscr{X}} L(\theta,\delta(\mathbf{x}))dF(\mathbf{x}|\theta)dF^\pi(\theta)$$

是 $\delta(\mathbf{x})$ 的贝叶斯风险，即它是将风险函数对 θ 的先验分布再求一次均值.

Definition

贝叶斯解

设 $\delta_1(\mathbf{x})$ 和 $\delta_2(\mathbf{x})$ 为 θ 的两个决策函数, $\pi(\theta)$ 为 θ 的先验分布,若 $R_\pi(\delta_1(\mathbf{x})) \le R_\pi(\delta_2(\mathbf{x}))$,则称 $\delta_1(\mathbf{x})$ 在贝叶斯风险下优于 $\delta_2(\mathbf{x})$. 若存在 $\delta^*(\mathbf{x})$, 使得对任一决策函数 $\delta(\mathbf{x})$ 有

$$R_\pi(\delta^*(\mathbf{x})) \le R_\pi(\delta(\mathbf{x})),$$

则称 $\delta^*(\mathbf{x})$ 为所考虑的统计判决问题的贝叶斯解.

Definition

贝叶斯后验风险

设 δ 是一个决策函数，称平均损失

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

为决策函数 $\delta(\mathbf{x})$ 的后验风险.

Definition

若存在决策函数 $\delta^*(x)$, 使得

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

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：

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

可见贝叶斯风险有两种表达式

$$R_\pi(\delta(\mathbf{x}))=E^\theta\left[R(\theta,\delta(\mathbf{x}))\right]=E^X\left[R(\delta(\mathbf{x})|\mathbf{x})\right],$$

即将风险函数 $R(\theta,\delta(\mathbf{x}))$ 按 θ 的先验分布 $\pi(\theta)$ 求均值，或者将后验风险按 X 的边缘分布 $m(x)$ 求均值.

定理 1. 对任何样本 $\mathbf{x}$, 若存在非随机化决策函数 $\delta_\pi(\mathbf{x})\in\mathscr{D}$, 满足

$$R(\delta_\pi|\mathbf{x})=\inf_{\delta\in\mathscr{D}}R(\delta(\mathbf{x})|\mathbf{x})$$

则 δ_π 为先验分布 $\pi(\theta)$ 之下的贝叶斯解.

证明. 设 $\delta(\mathbf{x})$ 为任一非随机化决策函数. 由已知条件, 可知

$$\begin{aligned}R(\delta(x)|x)&=\int_\Theta L(\theta,\delta(x))\pi(\mathrm{d}\theta|x)\\&\geqslant\int_\Theta L\left(\theta,\delta_\pi\right)\pi(\mathrm{d}\theta|x)=R\left(\delta_\pi(x)|x\right)\end{aligned}$$

对一切 $\mathbf{x} \in \mathcal{X}$ 成立, 从而将上式两边按 X 的边际分布求积分, 可得到

$$
\begin{aligned}
R_\pi(\delta(x)) &= \int_{\mathscr{X}} R(\delta(x)|x)m(x)\mathrm{d}x \\
&\geqslant \int_{\mathscr{X}} R\left(\delta_\pi(x)|x\right) m(x)\mathrm{d}x = R_\pi\left(\delta_\pi(x)\right)
\end{aligned}
$$

因此, 使贝叶斯风险达到最小的决策函数 $\delta_\pi(\mathbf{x})$ 是贝叶斯解. □

注 若 $\pi(\theta)$ 为广义先验分布, 且 $\delta_\pi(\mathbf{x})$ 是最小化后验贝叶斯风险得的最优决策函数, 则称 $\delta_\pi(\mathbf{x})$ 为广义贝叶斯解. 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.

例 P158 例 5.4.4 (续例 3.1.1) 通过血液检验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:

a \ θ	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数.

- 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:

$$\pi(\theta_1|x=0)=0.012, \quad \pi(\theta_2|x=0)=0.988,$$
$$\pi(\theta_1|x=1)=0.296, \quad \pi(\theta_2|x=1)=0.704\,.$$

- 故决策行为的后验风险分别为

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

- 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 $\delta^*(x)=\begin{cases} a_2 & \text{当 } x=0, \\ a_1 & \text{当 } x=1 \end{cases}$

5.3 一般损失函数下的贝叶斯解

5.3.1 平方损失下的贝叶斯估计

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

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

例 5.3.1;P147 设 $X \sim N(\theta, \sigma^2)$, 其中 σ^2 已知. θ 的先验分布是 $N(\mu, \tau^2)$, μ 和 τ^2 已知，求平方损失下 θ 的贝叶斯估计.

↑Example
↓Example

↑Example

例 5.3.2;P148 设 $\mathbf{X}=(X_1,\cdots,X_n)$ 是从泊松分布 $P(\theta)$ 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.

↓Example

5.3.2 加权平方损失下的贝叶斯估计

定理 3. 在加权平方损失 $L(\theta, a) = w(\theta)(\theta - a)^2$ 下, θ 的贝叶斯估计为

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

其中 $w(\theta)$ 为参数空间上的正值函数.

例 5.3.3;P149 设 $\mathbf{X} = (X_1, \cdots, X_n)$ 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.

↑Example
↓Example

5.3.3 绝对损失下的贝叶斯估计

绝对值损失下的贝叶斯解

定理 4. 在绝对损失 $L(\delta, a) = |\theta - a|$ 下，θ 的贝叶斯估计为后验中位数.

↑Example

例 5.3.4;P151 设 $\mathbf{X}=(X_1,\cdots,X_n)$ 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

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

求 θ 的在绝对值损失下的贝叶斯估计.

↓Example

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 5.3.5 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 $E(\theta|x) = 110.39$.

5.3.4 线性损失函数下的贝叶斯估计

定理 5. 在线性损失函数

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

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.

↑Example

例 5.3.6;P153 (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

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

求儿童智商 θ 的贝叶斯估计.

↓Example

5.4 假设检验和有限行动问题

- 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 $\mathscr{D}=\{a_1,\cdots,a_r\}$. 设在采取行动 a_i 下的损失为 $L(\theta,a_i)$，$i=1,\cdots,r$, 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．

5.4.1 假设检验问题

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

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

- 其后验风险为

$$R(a_0|\mathbf{x})=E_{\theta|\mathbf{x}}\big[L(\theta,a_0)\big]=P(\Theta_1|\mathbf{x}),$$
$$R(a_1|\mathbf{x})=E_{\theta|\mathbf{x}}\big[L(\theta,a_1)\big]=P(\Theta_0|\mathbf{x}).$$

- 按后验风险最小准则, 若

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

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.

↑Example

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:

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

↓Example

检验问题： $0 - k_i$ 损失情形

- 若为 $\mathbf{0-k_i}$ 损失

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

- 此时后验风险分别为

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

- 按后验风险最小准则, 若

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

↑Example

例 5.4.2;P155 设 $X \sim N(\theta, \sigma^2)$, 此处 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 $X_1, \cdots, X_n$ 为从总体 X 中抽取的简单样本. 求检验问题:

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

↓Example

5.4.2 多行动问题

- 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:

$$H_0 : \theta_1 - \theta_2 < -\varepsilon, \quad H_1 : \theta_1 - \theta_2 > \varepsilon, \quad H_2 : |\theta_1 - \theta_2| \leqslant \varepsilon,$$

  其中 $\varepsilon > 0$ 选择使得当 $|\theta_1 - \theta_2| \leqslant \varepsilon$ 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 5.4.3;P157 (续例 5.3.6) 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:

$$H_1: \ \theta < 90, \quad H_2: \ 90 \le \theta \le 110, \quad H_3: \ \theta > 110.$$

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:

$L(\theta, a)$	$\theta < 90$	$90 \le \theta \le 110$	$\theta > 110$
a_1	0	$\theta - 90$	$2(\theta - 90)$
a_2	$90 - \theta$	0	$\theta - 110$
a_3	$2(110 - \theta)$	$110 - \theta$	0

已知后验分布 $\pi(\theta|x)$ 是 $N(110.38, 8.32^2)$, 求多重检验问题的贝叶斯解.

5.4.3 统计决策中的区间估计问题

- 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 $C(\mathbf{x}) = (d_1(\mathbf{x}),\ d_2(\mathbf{x}))$ 为 θ 的一个区间估计.

- 损失函数的一种取法为

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

  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 $C(\mathbf{x})$ 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

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

  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.