经典统计/频率学派

1. 样本统计量

2. 中心极限定理（Central Limit Theorem）

3. 最大似然估计（Maximum Likelihood Estimation）

贝叶斯统计/贝叶斯学派

1. 贝叶斯推断

单观测

$$P(\theta|x)=\frac{P(x|\theta)P(\theta)}{P(x)}\propto P(x|\theta)P(\theta)$$

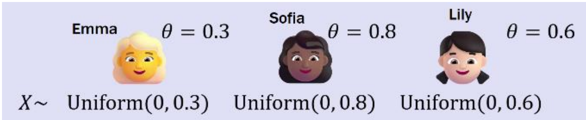
多观测（假设各个观测数据条件独立）

$$\begin{aligned}&f_{\Theta|X_1,\dots,X_n}(\theta|x_1,\dots,x_n)\\&=\frac{f_{X_1,\dots,X_n|\Theta}(x_1,\dots,x_n|\theta)f_\Theta(\theta)}{Z(x_1,\dots,x_n)}\\&\propto f_{X_1,\dots,X_n|\Theta}(x_1,\dots,x_n|\theta)f_\Theta(\theta)\\&=f_{X_1|\Theta}(x_1|\theta)\cdots f_{X_n|\Theta}(x_n|\theta)f_\Theta(\theta)\end{aligned}$$

注意条件独立≠独立. 也可以逐步更新.

① 约会大作战

John 想对 Apple 的迟到时间建模. 根据以往经验，前女友们迟到时间服从均匀分布：



先验：$\Theta\sim Uniform(0,1)$

似然：$X|\Theta\sim Uniform(0,\theta)$

单观测：Apple arrives $\frac{1}{2}$ hour late.

推断：

$$\begin{aligned}f_{\Theta|X}\left(\theta\middle|\frac{1}{2}\right)&\propto f_\Theta(\theta)f_{X|\Theta}\left(\frac{1}{2}\middle|\theta\right)\\&=\frac{1}{\theta}\quad\text{if }\frac{1}{2}\le\theta\le 1\end{aligned}$$

归一化常数（注意积分范围）：

$$Z(x)=\int_{\frac{1}{2}}^{1}\frac{1}{\theta}d\theta=\ln 2$$

后验：$f_{\Theta|X}\left(\theta\middle|\frac{1}{2}\right)=\frac{1}{\theta\ln 2}\quad\text{if }\frac{1}{2}\le\theta\le 1$

多观测：Apple is late by $\frac{1}{2},\frac{1}{4},\frac{1}{4}$ hours.

推断：

$$\begin{aligned}&f_{\Theta|X_1,X_2,X_3}\left(\theta\middle|\frac{1}{2},\frac{1}{4},\frac{1}{4}\right)\\&\propto f_{X_1|\Theta}\left(\frac{1}{2}\middle|\theta\right)f_{X_2|\Theta}\left(\frac{1}{4}\middle|\theta\right)f_{X_3|\Theta}\left(\frac{1}{4}\middle|\theta\right)f_\Theta(\theta)\\&=\frac{1}{\theta^3}\quad\text{if }\frac{1}{2}\le\theta\le 1\end{aligned}$$

归一化常数：$Z(x_1,x_2,x_3)=\int_{\frac{1}{2}}^{1}\frac{1}{\theta^3}d\theta=\frac{3}{2}$

后验：$f_{\Theta|X_1,X_2,X_3}\left(\theta\middle|\frac{1}{2},\frac{1}{4},\frac{1}{4}\right)=\frac{2}{3\theta^3}$ if $\frac{1}{2}\le\theta\le 1$

② 三类硬币模型

	H T	H H	T T
Prior	90%	5%	5%
	$\theta=1$	$\theta=2$	$\theta=3$

扔 2 个 Head，逐步更新.

第 1 个 Head：

$$P(\theta=1|H_1)=\frac{P(H_1|\theta=1)P(\theta=1)}{Z(H_1)}=\frac{0.45}{Z(H_1)}$$
$$P(\theta=2|H_1)=\frac{P(H_1|\theta=2)P(\theta=2)}{Z(H_1)}=\frac{0.05}{Z(H_1)}$$
$$P(\theta=3|H_1)=\frac{P(H_1|\theta=3)P(\theta=3)}{Z(H_1)}=\frac{0}{Z(H_1)}$$
$$Z(H_1)=0.45+0.05+0=0.5$$

后验： $P(\theta=1|H_1)=0.9$ $P(\theta=2|H_1)=0.1$ $P(\theta=3|H_1)=0$

第 2 个 Head：

$$P(\theta=1|H_2H_1)=\frac{0.45}{0.55}\approx 0.82$$
$$P(\theta=2|H_2H_1)=\frac{0.1}{0.55}\approx 0.18$$
$$P(\theta=3|H_2H_1)=\frac{0}{0.55}=0$$

③ 连续硬币模型（Beta-Bernoulli）

先验： $\Theta\sim Uniform(0,1)$，即 $Beta(1,1)$

似然（抛硬币）： $X|\Theta\sim Bernoulli(\theta)$

多观测：观测到 H,T,T.

推断：

$$f_{\Theta|X_1,X_2,X_3}(\theta|H,T,T)$$
$$\propto P_{X_1|\Theta}(H|\theta)P_{X_2|\Theta}(T|\theta)P_{X_3|\Theta}(T|\theta)f_\Theta(\theta)$$
$$=\theta(1-\theta)^2 \qquad \text{if } 0\le\theta\le 1$$

归一化常数： $Z(x_1,x_2,x_3)=\int_0^1\theta(1-\theta)^2\,d\theta=\frac{1}{12}$

可以看出是B(2,3)，但最好用计算器算.

因为 $Beta$ 函数要求积分范围为[0,1].

后验： $f_{\Theta|X_1,X_2,X_3}(\theta|H,T,T)=12\theta(1-\theta)^2$ if $0\le\theta\le 1$

④ iPhone 销售模型（Gamma-Poisson/Exp）

Gamma-Poisson 型，观测每天售出台数.

先验： $\Theta\sim Gamma(3,2)$，2 天卖 3 台

$$f_\Theta(\theta)=\frac{2^3}{\Gamma(3)}\theta^{3-1}e^{-2\theta}\ \text{ if } \theta>0$$

似然： $X|\Theta\sim Poisson(\theta)$

$$P_{X|\Theta}(x|\theta)=\frac{e^{-\theta}\theta^x}{x!}\quad \text{for } x=0,1,2..$$

单观测：观测到 1 天卖了 3 台.

推断： $f_{\Theta|X}(\theta|3)\propto P_{X|\Theta}(3|\theta)f_\Theta(\theta)\propto\theta^5e^{-3\theta}$ if $\theta>0$

归一化常数： $Z(x)=\int_0^\infty\theta^{6-1}e^{-3\theta}\,d\theta=\frac{\Gamma(6)}{3^6}$

后验： $f_{\Theta|X}(\theta|3)=\frac{3^6}{\Gamma(6)}\theta^{6-1}e^{-3\theta}$ if $\theta>0$

即Gamma(6,3).

Gamma-Exponential 型,观测每台售出时间.

先验： $\Theta\sim Gamma(1,2)$，2 小时卖 1 台

似然： $X|\Theta\sim Exp(\theta)$

$$f_{X|\Theta}(x|\theta)=\theta e^{-\theta x}\quad \text{for } x\ge 0$$

多观测：隔 1.5h, 2h, 2.5h，共卖出 3 台.

推断：

$$f_{\Theta|X_1,X_2,X_3}(\theta|1.5,2,2.5)$$
$$\propto f_{X_1|\Theta}(1.5|\theta)f_{X_2|\Theta}(2|\theta)f_{X_3|\Theta}(2.5|\theta)f_\Theta(\theta)$$
$$\propto\theta^3e^{-8\theta}\qquad \text{if } \theta>0$$

归一化常数： $Z(x_1,x_2,x_3)=\int_0^\infty\theta^{4-1}e^{-8\theta}\,d\theta=\frac{\Gamma(4)}{8^4}$

后验： $f_{\Theta|X_1,X_2,X_3}(\theta|1.5,2,2.5)=\frac{8^4}{\Gamma(4)}\theta^{4-1}e^{-8\theta}$ if $\theta>0$

即 $Gamma(4,8)$.

⑤ Normal-Normal 模型

Suppose $X_1, \ldots, X_n$ form a random sample from Normal distribution with an unknown mean μ and a known variance σ^2 $(\sigma^2 > 0)$. If the prior distribution $f_{\mathrm{M}}(\mu)$ is the Normal distribution $N(\mu_0, \sigma_0^2)$, then the posterior distribution $f_{\mathrm{M}|X_1,\ldots,X_n}(\mu|x_1,\ldots,x_n)$ given observed data $\{x_i\}_{i=1}^n$, $\sum_{i=1}^n x_i = k$ is also the Normal distribution $N(\mu', \sigma'^2)$, where

$$\mu' = \frac{\sigma^2\mu_0 + \sigma_0^2 k}{\sigma^2 + n\sigma_0^2} \qquad \sigma'^2 = \frac{\sigma^2\sigma_0^2}{\sigma^2 + n\sigma_0^2}$$

Special case: both σ_0^2 and σ^2 are 1

$$\mu' = \frac{\mu_0 + k}{n+1} \qquad \sigma'^2 = \frac{1}{n+1}$$

Example: A $N(\theta, 1)$ random variable takes value 3.97. θ is a value of the random variable Θ. Θ follows a standard normal. What is the posterior of Θ?

似然：$X|\Theta \sim N(\theta, 1)$

$$f_{X|\Theta}(x|\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\theta)^2}$$

先验：$\Theta \sim N(0,1)$

$$f_\Theta(\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\theta^2}$$

贝叶斯推断：

$$\begin{aligned} f_{\Theta|X}(\theta|x) &\propto f_{X|\Theta}(x|\theta) f_\Theta(\theta) \\ &\propto e^{-\frac{1}{2}[(x-\theta)^2+\theta^2]} \\ &\propto e^{-\frac{1}{2}(\sqrt{2}\theta - \frac{\sqrt{2}}{2}x)^2} \\ &= e^{-\frac{1}{2}(\frac{\theta - \frac{x}{2}}{\frac{1}{\sqrt{2}}})^2} \end{aligned}$$

> 即 $N(\frac{x}{2}, (\frac{1}{\sqrt{2}})^2)$.

归一化常数：

$$Z(x) = \int_{-\infty}^{\infty} e^{-\frac{1}{2}(\frac{\theta - \frac{x}{2}}{\frac{1}{\sqrt{2}}})^2} d\theta = \sqrt{2\pi\sigma^2} = \sqrt{\pi}$$

后验：$\Theta|X \sim N(\frac{x}{2}, (\frac{1}{\sqrt{2}})^2)$

$$f_{\Theta|X}(\theta|3.97) = \frac{1}{\sqrt{\pi}} e^{-\frac{1}{2}(\frac{\theta - 1.985}{\frac{1}{\sqrt{2}}})^2}$$

Example: Three independent $N(\theta, 1)$ random variables take values $3.97, 4.09, 3.11$. What is Θ?

似然：$X_i|\Theta \sim N(\theta, 1)$

$$f_{X_i|\Theta}(x_i|\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_i-\theta)^2}$$

先验：$\Theta \sim N(0,1)$

$$f_\Theta(\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\theta^2}$$

贝叶斯推断：$x_1 = 3.97$，$x_2 = 4.09$，$x_3 = 3.11$

$$\begin{aligned} f_{\Theta|X_1,X_2,X_3}(\theta|x_1,x_2,x_3) &\propto f_{X_1|\Theta}(x_1|\theta) f_{X_2|\Theta}(x_2|\theta) f_{X_3|\Theta}(x_3|\theta) f_\Theta(\theta) \\ &\propto e^{-\frac{1}{2}[(x_1-\theta)^2+(x_2-\theta)^2+(x_3-\theta)^2+\theta^2]} \\ &\propto e^{-\frac{1}{2}(2\theta - \frac{x_1+x_2+x_3}{2})^2} \\ &= e^{-\frac{1}{2}(\frac{\theta - \frac{x_1+x_2+x_3}{4}}{\frac{1}{2}})^2} \end{aligned}$$

> 即 $N(\frac{x_1+x_2+x_3}{4}, \frac{1}{4})$.

法二：由 special case, when both σ_0^2 and σ^2 are 1,

$$\mu' = \frac{\mu_0 + k}{n+1} = \frac{0 + x_1 + x_2 + x_3}{3+1} \qquad \sigma'^2 = \frac{1}{n+1} = \frac{1}{3+1}$$

归一化常数：

$$Z(x) = \int_{-\infty}^{\infty} e^{-\frac{1}{2}(\frac{\theta - \frac{x_1+x_2+x_3}{4}}{\frac{1}{2}})^2} d\theta = \sqrt{2\pi\sigma^2} = \sqrt{\frac{\pi}{2}}$$

后验：$\Theta|X_1, X_2, X_3 \sim N(\frac{x_1+x_2+x_3}{4}, \frac{1}{4})$

$$f_{\Theta|X_1,X_2,X_3}(\theta|3.97, 4.09, 3.11) = \frac{1}{\sqrt{\frac{\pi}{2}}} e^{-\frac{1}{2}(\frac{\theta - 2.7925}{\frac{1}{2}})^2}$$

A more general case

Suppose $X_1, \ldots, X_n$ form a random sample from Normal distributions with a common unknown mean μ and the known variance σ_i^2 $(\sigma_i^2 > 0)$. If the prior distribution $f_{\mathrm{M}}(\mu)$ is the Normal distribution $N(\mu_0, \sigma_0^2)$, then the posterior distribution $f_{\mathrm{M}|X_1,\ldots,X_n}(\mu|x_1,\ldots,x_n)$ given observed data $\{x_i\}_{i=1}^n$, $\sum_{i=1}^n x_i = k$ is also the Normal distribution $N(\mu', \sigma'^2)$, where

$$\frac{\mu'}{\sigma'^2} = \frac{\mu_0}{\sigma_0^2} + \frac{x_1}{\sigma_1^2} + \cdots + \frac{x_n}{\sigma_n^2} \qquad \frac{1}{\sigma'^2} = \frac{1}{\sigma_0^2} + \frac{1}{\sigma_1^2} + \cdots + \frac{1}{\sigma_n^2}$$

2. 预测/估计/假设检验

贝叶斯预测

① 离散变量

$$P(x^*|X = x) = \int_{-\infty}^{+\infty} P(x^*|\theta) f_{\Theta|X}(\theta|x) d\theta$$

② 连续变量

$$\begin{aligned} P(x^* \in [a,b]|X = x) &= \int_{-\infty}^{+\infty} P(x^* \in [a,b]|\theta) f_{\Theta|X}(\theta|x) d\theta \\ &= \int_{-\infty}^{+\infty} \int_a^b f_{X|\Theta}(x^*|\theta) dx^* \, f_{\Theta|X}(\theta|x) d\theta \end{aligned}$$

MAP 估计

MAP 决策准则

似然比检验

分布表

$$X \sim \text{Exponential}(\theta), f(x) = \begin{cases} \theta e^{-\theta x} & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases}$$

$$X \sim \mathcal{N}(\mu, \sigma^2), f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

$$\mathrm{B}(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

$$X \sim \text{Beta}(\alpha, \beta), f(x) = \begin{cases} \frac{1}{\mathrm{B}(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} & \text{for } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

$\Gamma(\alpha) = (\alpha-1)!$ for positive integer α

$$X \sim \text{Gamma}(\alpha, \beta), f(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} & \text{for } x > 0 \\ 0 & \text{for } x \leq 0 \end{cases}$$

$$X \sim \text{Uniform}(a, b), f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

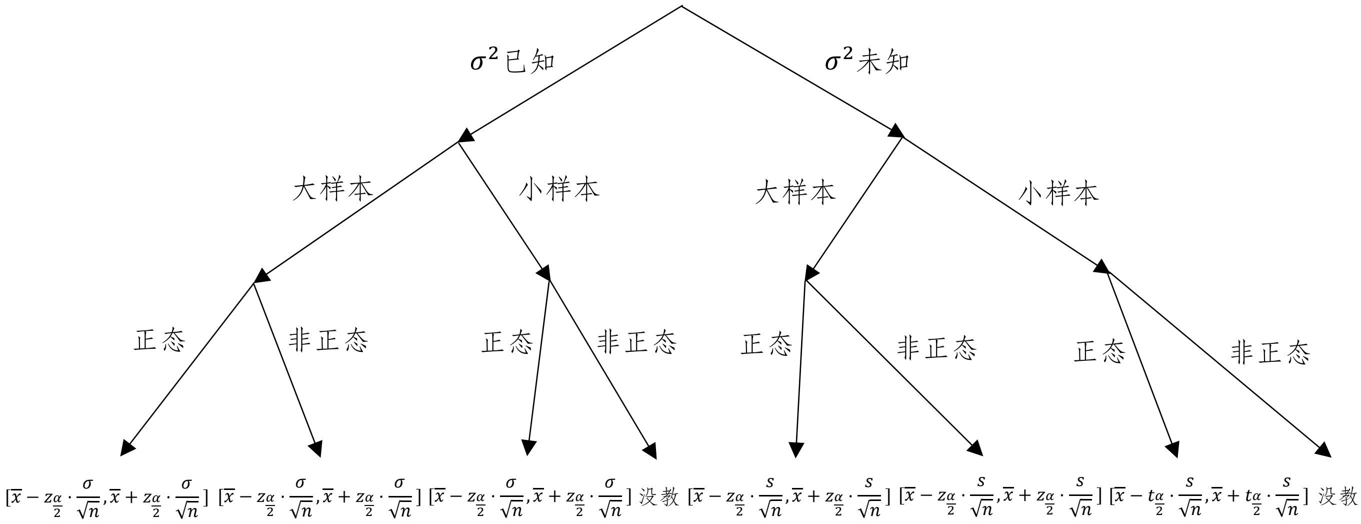
$$X \sim \text{Poisson}(\theta), P(x) = \begin{cases} \frac{e^{-\theta}\theta^x}{x!} & \text{for } x = 0,1,2 \ldots \\ 0 & \text{otherwise} \end{cases}$$

$$X \sim \text{Geometric}(p), P(x) = \begin{cases} (1-p)^{x-1} p & \text{for } x = 1,2 \ldots \\ 0 & \text{otherwise} \end{cases}$$

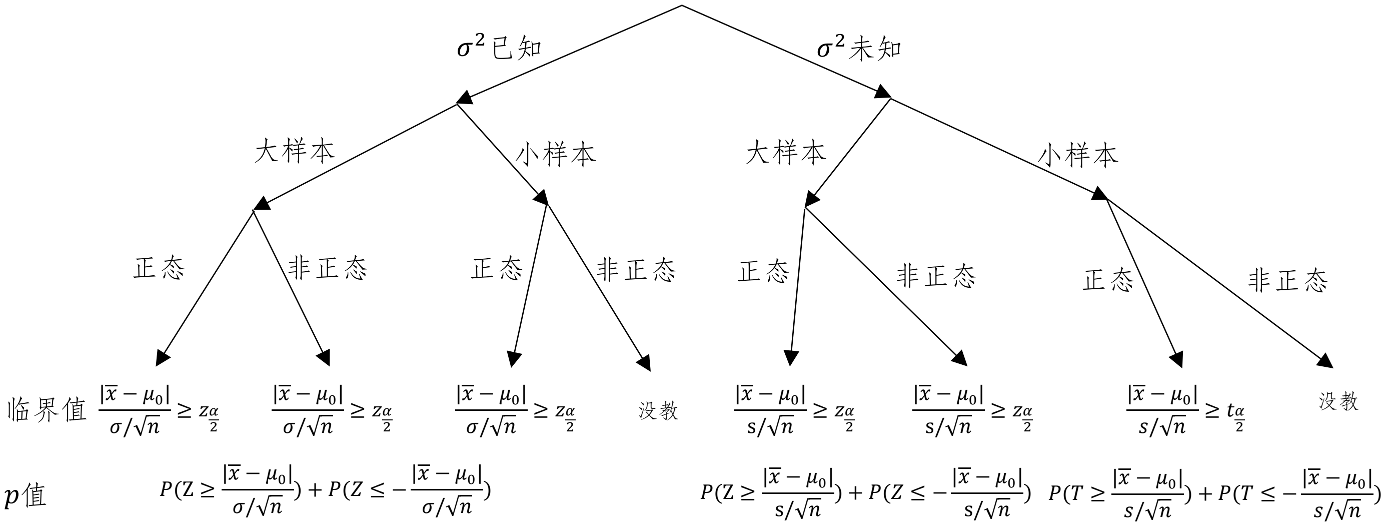
$$X \sim \text{Bernoulli}(p), P(x) = p^x (1-p)^{1-x} \quad \text{for } x \in \{0,1\}$$

$$X \sim \text{Binomial}(n, p), P(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{for } x = 0,1,\ldots,n \\ 0 & \text{otherwise} \end{cases}$$

置信区间二叉树



假设检验二叉树 1：单个总体均值的假设检验 $H_0{:}\mu=\mu_0$



假设检验二叉树 2：比较两个总体均值 $H_0{:}\mu_x-\mu_y=0$