STA 2212S Mar 19, 2021

Recall Example 11.28 in SM Empirical Bayes

$\rightarrow x_i|\theta_i \sim N(\theta_i, v_i)$ ind't v_i known

$\rightarrow \theta_i|\mu \sim N(\mu, \tau^2)$ i.i.d. τ^2 known

no hyperprior use data to est μ

$\Rightarrow x_i \sim N(\mu, v_i+\tau^2)$ ind't

$E(x_i) = E\{E(x_i|\theta_i)\} = E(\theta_i) = \mu$

$Var(x_i) = E\,var(x_i|\theta_i) + var\{E(x_i|\theta_i)\}$

$L_m(\mu) \propto \prod_{i=1}^{n} f(x_i; \mu, v_i+\tau^2) \Rightarrow \mu$

$$\hat{\mu}_{EB} = \sum_{i=1}^{n}\{x_i/(v_i+\tau^2)\} \Big/ \sum\{1/v_i+\tau^2\}$$

$$\hat{\theta}_{i,EB} = \hat{\mu}_{EB} + \frac{\tau^2}{v_i+\tau^2}(x_i - \hat{\mu}_{EB})$$ shrinkage est'r

$$E(\theta_i|x_i) = \mu + \frac{\tau^2}{v_i+\tau^1}(x_i - \mu)$$

$v_i \equiv 1$ ~~$x_i \sim N(\mu, 1)$~~ $x_i|\theta_i \sim N(\theta_i, 1)$

$\theta_i \sim N(\mu, \tau^2)$

$\Rightarrow x_i \underset{marg}{\sim} N(\mu, \tau^2+1)$ $i=1,\dots,n$ i.i.d.

n obs = n par. unkn. "nonparametric"

$\hat{\mu}_{EB} = \bar{x}$

$\hat{\theta}_{i,EB} = \bar{x} + \frac{\tau^2}{\tau^2+1}(x_i - \bar{x})$

Going further, τ^2 unknown $\Rightarrow \frac{1}{n}\sum(x_i-\bar{x})^2 = \hat{\tau}^2 + 1$

max. lik. est. of τ^2 [unbiased est. ...]

$$\hat{\tau}^2 = \left(\frac{1}{n}W - 1\right)_+ \qquad W = \sum (x_i - \bar{x})^2$$

$$\hat{\theta}_{i,EB} = \bar{x} + \left(\frac{\frac{1}{n}W - 1}{\frac{1}{n}W}\right)(x_i - \bar{x})$$

$$= \bar{x} + \left\{1 - \left(\tfrac{1}{n}W\right)^{-1}\right\}(x_i - \bar{x})$$

$$= \bar{x} + \left(1 - \frac{n}{W}\right)(x_i - \bar{x})$$

Instead, I'll use

$$\rightarrow \hat{\theta}_{i,EB} = \bar{x} + \left(1 - \frac{b}{W}\right)(x_i - \bar{x}) \qquad b \text{ to be chosen}$$

b to minimize $E\sum_{i=1}^{n}\left(\hat{\theta}_{i,EB} - \theta_i\right)^2$

$$R_b(\underline{\theta}, \underline{\hat{\theta}}) = \int L(\theta, \hat{\theta}(x)) f(x;\theta)\,dx$$

risk = exp'd loss

$$\min_b E \sum_{i=1}^{n} \left\{\bar{x} + \left(1 - \frac{b}{W}\right)(x_i - \bar{x}) - \theta_i\right\}^2$$

$$\min_b E \sum_{i=1}^{n} \left\{x_i - \theta_i - \frac{b}{W}(x_i - \bar{x})\right\}^2$$

$$= \min_b \sum \left\{E(x_i - \theta_i)^2 - \sum \frac{2b}{W}(x_i - \theta_i)(x_i - \bar{x}) + \frac{b^2}{W^2}\sum (x_i - \bar{x})^2\right\}$$

$$= \min_b \left\{-E\sum \frac{2b}{W}(x_i - \theta_i)(x_i - \bar{x}) + E\frac{b^2}{W}\right\}$$

$$W = \sum_1 (x_i - \bar{x})^2 \qquad x_i \sim N(\mu, \sigma^2 + 1)$$

$$\sum (x_i - \bar{x})^2 \sim \sigma^2 \chi^2_{n-1}$$

$$\sim (\sigma^2 + 1)\chi^2_{n-1}$$

$E\left(\frac{1}{W}\right)$ — easy

$E\left\{\frac{(x_i - \theta_i)(x_i - \bar{x})}{W}\right\}$ — bit harder

$$= \min_b \left\{ n + b^2 E\left(\frac{1}{W}\right) - 2E\sum_{i=1}^{n} \frac{(x_i - \bar{x})(x_i - \theta_i)}{W} \right\}$$

$$= \ldots \min_b \left\{ n + b(b - 2(n-3)) E(W^{-1}) \right\}$$

$\frac{\partial}{\partial b}$: $\{2b - 2(n-3)\} E(W^{-1}) = 0$

$$b = n - 3$$

$$\hat{\theta}_{i,JS} = \bar{x} + \left(1 - \frac{n-3}{W}\right)(x_i - \bar{x}) \qquad n \geq 4 \text{ for this argument}$$

James-Stein estimator

this has smaller E(mse) than x_i

$$R(\underset{\sim}{\theta}, \hat{\underset{\sim}{\theta}}_{JS}) \leq R(\underset{\sim}{\theta}, \underset{\sim}{x})$$

if $X \sim N(\theta, 1)$ then X is best est. of θ

$X_1, \ldots, X_n$ iid $N(\theta, 1)$ then $\bar{X}$ is best est of θ

$\mathbb{R}^2$ if $(X_{1i}, X_{2i})^T \sim N_2\left(\begin{pmatrix}\theta_1\\ \theta_2\end{pmatrix}, \begin{matrix}1 & 0\\ 0 & 1\end{matrix}\right)$ $i = 1, \dots, n$

then $(\bar{X}_1, \bar{X}_2)^T$ has smallest mse

$\mathbb{R}^3$??

in $\mathbb{R}^4$ $\underset{\sim}{\bar{X}} = (\bar{X}_1, \dots, \bar{X}_4)$ does not have smallest mse

HW6/7 bonus

$X_i \sim N(\mu_i, 1)$

$\pi(\mu_i) \propto 1$ "flat"

$\pi(\mu_i | x_i) \propto N(x_i, 1) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\mu_i - x_i)^2}$

$\pi(\underset{\sim}{\mu} | \underset{\sim}{x}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\mu_i - x_i)^2}$

But if you're interested in $\sum \mu_i^2$

$E(\sum \mu_i^2 \mid \sum x_i^2) = n + \sum x_i^2$

A frequentist would note

$E(\sum x_i^2 \mid \underset{\sim}{\mu}) = n + \sum \mu_i^2$

$\sum X_i^2 - n$ is unbiased for $\sum \mu_i^2$

Bayes posterior mean $= \sum x_i^2 + n = \tilde{\mu}_{Bayes}$

$E(\tilde{\mu}_{Bayes} \mid \underset{\sim}{x})$ under model $= \sum \mu_i^2 + 2n$

$$\hat{\theta}^{PP}_{i,JS} = \bar{x} + \left(1 - \frac{n-3}{\Sigma(x_i-\bar{x})^2}\right)_+ (x_i - \bar{x})$$

$\hat{\underset{\sim}{\theta}}_{JS} = (\hat{\theta}_{1,JS}, \cdots, \hat{\theta}_{n,JS})$ smaller mse than

$\hat{\theta}_{MLE} = (x_1, \cdots, x_n)$

A.S (12.12)

$$\hat{\theta}_{i,JS_2} = \left(1 - \frac{n-2}{\Sigma x_i^2}\right)_+ x_i \quad ??$$

$x_i | \theta_i \sim N(\theta_i, 1) \qquad \theta_i \sim N(0, \tau^2)$ (pr)

JS Thm is $(\hat{\underset{\sim}{\theta}}_{1,JS_2}, \cdots, \hat{\theta}_{n,JS_2})$ has smaller mse than $(x_1, \cdots, x_n)$

"positive part J-S estimator"