Bayesian Estimation

- Bayesian approach. $\int R(\theta, \delta)\, d\Lambda(\theta)$.

  Λ —— a specific probability measure on Ω.

- $\Theta \sim \Lambda$. prior distribution

  $X \mid \Theta = \theta \sim P_\theta$.

$$E[L(\theta, \delta(X)) \mid \Theta = \theta] = \int L(\theta, \delta(x))\, dP_\theta(x) = R(\theta, \delta)$$

$$E L(\Theta, \delta) = E\, E[L(\theta, \delta) \mid \Theta] = E R(\theta, \delta) = \int R(\theta, \delta)\, d\Lambda(\theta)$$

- The choice of the prior distribution Λ is critical.

  subjective opinion.

- minimize $\int R(\theta, \delta)\, d\Lambda(\theta) \longrightarrow$ called Bayes.

Thm: Suppose $\Theta \sim \Lambda$. $X \mid \Theta = \theta \sim P_\theta$ $L(\theta, d) \geq 0$. if

a) $EL(\Theta, \delta_0) < \infty$ for some δ_0

b) for a.e. x. $\exists\ \delta_\Lambda(x)$ minimize $E[L(\Theta, d) \mid X = x]$ w.r.t. d.

$\Longrightarrow$ δ_Λ is a Bayes estimator.

$$E(L(\theta, \delta(X)) \mid X) \geq E(L(\theta, \delta_\Lambda(X)) \mid X)$$

ex. Weighted L_2-loss.

$$L(\theta, d) = \omega(\theta)(d - g(\theta))^2.$$

minimize.

$$E[\omega(\theta)\{(d-g(\theta))\}^2 \mid X=x]$$

$$= d^2 E[\omega(\theta) \mid X=x] - 2d\, E[\omega(\theta) g(\theta) \mid X=x] + E[\omega(\theta) g^2(\theta) \mid X=x]$$

$$\delta_\Lambda(x) = \frac{E[\omega(\theta) g(\theta) \mid X=x]}{E[\omega(\theta) \mid X=x]}.$$

If $\omega = 1$. $\delta_\Lambda(x) = E[g(\theta) \mid x]$. — posterior mean of $g(\theta)$

Λ is absolutely continuous with Lebesgue density λ.

joint (X, θ): $p_\theta(x)\lambda(\theta)$

marginal X: $g(x) = \int p_\theta(x)\lambda(\theta)\,d\theta$

$\theta \mid X=x$. $\lambda(\theta \mid x) = \dfrac{p_\theta(x)\lambda(\theta)}{g(x)}$

$$\delta_\Lambda(x) = \frac{\int \omega(\theta) g(\theta) p_\theta(x) \lambda(\theta)\, d\theta}{\int \omega(\theta) p_\theta(x) \lambda(\theta)\, d\theta}$$

Ex. Binomial

$X \sim Bin(n, \theta)$.

prior. $\theta \sim Beta(\alpha, \beta)$. $\lambda(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1}(1-\theta)^{\beta-1}$ $0 < \theta < 1$.

$$\int_0^1 \theta^{\alpha-1}(1-\theta)^{\beta-1} d\theta = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

$$E\theta = \frac{\alpha}{\alpha+\beta}.$$

$$q(x) = \int p_\theta(x)\lambda(\theta)d\theta = \int_0^1 \binom{n}{x} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{x+\alpha-1}(1-\theta)^{n-x+\beta-1} d\theta$$

$$= \binom{n}{x} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(x+\alpha)\Gamma(n-x+\beta)}{\Gamma(n+\alpha+\beta)}$$

$$\lambda(\theta|x) = \frac{\Gamma(n+\alpha+\beta)}{\Gamma(\alpha+x)\Gamma(\beta+n-x)} \theta^{x+\alpha-1}(1-\theta)^{n-x+\beta-1} \quad \theta \in (0,1)$$

$$\theta | X=x \sim Beta(x+\alpha, n-x+\beta)$$

$$\delta_\lambda(x) = E(\theta|x) = \underline{\frac{x+\alpha}{n+\alpha+\beta}} = \frac{n}{n+\alpha+\beta}\frac{x}{n} + \left[1 - \frac{n}{n+\alpha+\beta}\right]\frac{\alpha}{\alpha+\beta}.$$

Prior distributions that ensures a posterior from the same class are called conjugate.

ex. Negative Binomial.

X — number of failures before the second success.

$$p_\theta(x) = P_\theta(X=x) = (x+1)\,\theta^2(1-\theta)^x. \qquad x = 0, 1, \dots$$

$$g(\theta) = \frac{1}{\theta}. \qquad \Theta \sim \text{Unif}(0,1).$$

$$\lambda(\theta|x) \propto_\theta p_\theta(x)\lambda(\theta) \propto_\theta \theta^2(1-\theta)^x.$$

$$\theta | X=x \sim \text{Beta}(3, x+1)$$

$$\delta_0(x) = E(\Theta^{-1} | X=x) = \frac{\Gamma(x+4)}{\Gamma(3)\Gamma(x+1)} \int_0^1 \theta(1-\theta)^x \, d\theta$$

$$= \frac{\Gamma(x+4)\Gamma(2)\Gamma(x+1)}{\Gamma(3)\Gamma(x+1)\Gamma(x+3)} = \frac{x+3}{2}.$$

ex 5.3. UMVUE of $\frac{1}{\theta}$: $\delta_1(x) = \frac{x+2}{2}$

$$\delta_0(x) = \delta_1(x) + \frac{1}{2}.$$

bias: $b(\theta, \delta_0) = E_\theta \delta_0(X) - \frac{1}{\theta} = \frac{1}{2}.$

$$R(\theta, \delta_0) = \text{Var}(\delta_0) + \frac{1}{4} = \text{Var}_\theta(\delta_1) + \frac{1}{4} = R(\theta, \delta_1) + \frac{1}{4}$$

UMVUE δ_1 has uniformly smaller risk than δ_0!

An estimator is call <u>inadmissible</u> if a competing estimator has a better risk function.

Condition (a) of thm fails!