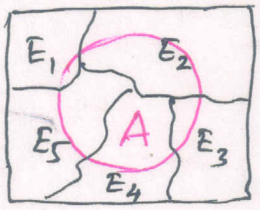
Bayesian Estimation

Classical Bayes' Theorem:

If $E_1, E_2, \cdots, E_n$ are mutually exclusive events with $P(E_i) \neq 0$, $i = 1(1)n$, Then for any arbitrary event A, $A \subset \bigcup_{i=1}^{n} E_i$ such That $P(A) > 0$, we have



$$P(E_i/A) = \frac{P(E_i) \cdot P(A/E_i)}{\sum_{i=1}^{n} P(E_i) \cdot P(A/E_i)}, \quad i = 1(1)n.$$

Proof: Since $A \subset \bigcup_{i=1}^{n} E_i$, we have $A = A \cap \left(\bigcup_{i=1}^{n} E_i\right) = \bigcup_{i=1}^{n} (A \cap E_i)$

Since $(A \cap E_i) \subset E_i$, $i = 1(1)n$ are mutually exclusive events, we have

$$P(A) = P\left\{\bigcup_{i=1}^{n} (A \cap E_i)\right\} = \sum_{i=1}^{n} P(A \cap E_i) = \sum_{i=1}^{n} P(E_i) \cdot P(A/E_i) \quad [\because P(A \cap E_i) = P(E_i) \cdot P(A/E_i)]$$

Therefore, $$P(E_i/A) = \frac{P(A \cap E_i)}{P(A)} = \frac{P(E_i) \cdot P(A/E_i)}{\sum_{i=1}^{n} P(E_i) \cdot P(A/E_i)}.$$

Remarks: 1. The probabilities $P(E_1), P(E_2), \cdots, P(E_n)$ are termed as the 'a prior probabilities' because they exist before we gain any information from the experiment itself.

2. The probabilities $P(A/E_i)$, $i = 1(1)n$ are called 'likelihoods' because they indicate how likely the event A under consideration is to occur, given each and every a prior probability.

3. The probability $P(E_i/A)$, $i = 1(1)n$ are called 'posterior probabilities' because they are determined after the results of the experiment are known. Sometimes, it is called revised estimate of $P(E_i)$, since it is ~~re calcul~~ recalculated after incorporating information contained in $P(A/E_i)$; $i = 1(1)n$.

Let X: A random variable, $X \in \mathcal{X}$

$F_\theta(x)$: $P(X \leq x)$, the cdf

$p_\theta(x)$: pmf or pdf of X given θ, $\theta \in \Theta$

$d(x)$: An estimator of θ or $\gamma(\theta)$, a function of θ. $d(x) \in \mathcal{D}$.

$L(\theta, d)$: Loss function such that $\Theta \times \mathcal{D} \longrightarrow \mathbb{R}$, ~~such that~~ when θ is true d is preferred to d' if $L(\theta, d) < L(\theta, d')$.

Before experiment is performed X is a r.v., $X \in \mathcal{X}$.
$d(X)$ is a r.v., so $L(\theta, d(X))$ is a real-valued r.v.
Now $E\{L(\theta, d(X))\} = R_d(\theta)$, $\theta \in \Theta$. This $R_d(\theta)$ is called the risk function of d.

Let $L(\theta, d) = (d-\theta)^2$

$$\therefore R_d(\theta) = E_\theta[d(X) - \theta]^2, \text{ the risk function}$$
$$= Var_\theta(d(X)) + \{b_\theta(d(X))\}^2$$
$$= Var_\theta(d(X)) + [E_\theta(d(X)) - \theta]^2$$

One can compare two different estimators in terms of this risk function. But it may or may not be possible to find an estimator for which this risk function is minimized for all θ generally.

Example: Let $X \sim N(0, \theta)$. Let $x_1, x_2, \dots, x_n$ are n sample observations. Objective is to estimate θ with squared error loss, $0 < \theta < \infty$.
We know MVUE of $\theta = \frac{1}{n}\sum_{i=1}^{n} x_i^2$

Assume that $c\sum x_i^2$ is the best estimator in terms of risk.

$$R_d(\theta) = E_\theta[c.\sum x_i^2 - \theta]^2$$
$$= E_\theta[c.\theta\chi^2_n - \theta]^2$$
$$= \theta^2 E[c\chi^2_n - 1]^2$$
$$= \theta^2[c^2 E(\chi^2_n)^2 - 2c.E(\chi^2_n) + 1]$$
$$= \theta^2\left[c^2 \frac{\Gamma(\frac{n}{2}+2).2^2}{\Gamma(\frac{n}{2})} - 2.c.\frac{\Gamma(\frac{n}{2}+1)}{\Gamma(\frac{n}{2})}.2 + 1\right].$$

$= \theta^2 [c^2 n(n+2) - 2cn + 1]$

$$\frac{d R_d(\theta)}{dc} = 0 \Rightarrow c = \frac{1}{n+2}$$

Now $d_1(\underline{x}) = \frac{1}{n} \sum x_i^2$

$d_2(\underline{x}) = \frac{1}{n+2} \sum x_i^2$

$R_{d_2}(\theta) < R_{d_1}(\theta) \ \forall \theta$

Here we are not sure whether $d_2(x)$ is best.

Take $d_3(\underline{x}) = 1$.

Then $R_{d_3}(\theta) = 0$, if $\theta = 1$

> 0 , if $\theta \neq 1$

So, for $\theta = 1$, d_3 is best.

It is very difficult to find an estimator which is universally b

Bayes Principle-

Let Θ is either countable or subset of a Euclidian space.
θ is assumed to be a r.v. with a known probability distribution (prior distribution θ).

(i) If Θ is countable, prior represented by pmf $\tau(\theta)$; $\tau(\theta) \geq 0$, $\theta \in \Theta$, $\sum_{\theta \in \Theta} \tau(\theta) = 1$ and $\mathcal{X}$ is countable with pmf $p(x/\theta)$.

(ii) If Θ is subset of Euclidian Space, prior represented by pdf $\tau(\theta)$; $\tau(\theta) \geq 0$, $\int_{\Theta} \tau(\theta) d\theta = 1$.

Now $r_d(\tau)$ = Bayes risk of d under the prior τ

= Average risk when θ follows the prior suggested by $\tau(\theta)$.

$$= \begin{cases} \sum_{\theta \in \Theta} R_d(\theta) \tau(\theta) \\ \int_{\Theta} R_d(\theta) \tau(\theta) d\theta \end{cases}$$

Now d_0 is Bayes rule under τ if $r_{d_0}(\tau) = \min_{d \in \mathcal{D}} r_d(\tau)$.

Method of determination of Bayes Rule when it exists

Case I: $\mathcal{X}$ is countable and for each θ, distribution of x is discrete with pmf $p(x/\theta)$. Also Θ is discrete. Prior represent by pmf $\tau(\theta)$.

Our problem is to determine $d = d_0 \ni$

$$r_{d_0}(\tau) = \min_{d \in \mathcal{D}} r_d(\tau)$$

i.e. $$\sum_{\Theta} \tau(\theta) R_{d_0}(\theta) = \min_{d \in \mathcal{D}} \sum_{\Theta} \tau(\theta) R_d(\theta)$$

i.e. $$\sum_{\Theta} \tau(\theta) \sum_{\mathcal{X}} L(\theta, d_0(x)) p(x/\theta) = \min_{d \in \mathcal{D}} \sum_{\Theta} \tau(\theta) \sum_{\mathcal{X}} L(\theta, d(x)) p(x/\theta)$$

i.e. $$\sum_{\mathcal{X}} \sum_{\Theta} L(\theta, d_0(x)) \tau(\theta). p(x/\theta) = \min_{d \in \mathcal{D}} \sum_{\mathcal{X}} \sum_{\Theta} L(\theta, d(x)). \tau(\theta) p(x/\theta) \quad \text{---- (A)}$$

We have to choose $d_0(x)$, $x \in \mathcal{X}$ so that (A) holds.
i.e. We have to choose for each x a $d_0(x)$ so that (A) holds.
Suppose for any $x \in \mathcal{X}$, $\exists$ an $d_0(x) \ni$

$$\sum_{\Theta} L(\theta, d_0(x)) \tau(\theta). p(x/\theta) = \min_{d \in \mathcal{D}} \underbrace{\sum_{\Theta} L(\theta, d(x)) \tau(\theta). p(x/\theta)}_{L^*(d/x),\ \text{say}} \quad \text{---(B)}$$

Hence our aim is to minimize $\sum_{x \in \mathcal{X}} L^*(d/x)$, by choosing $d(x)$.

Then, if for each x, $d_0(x) \ni$

$$L^*(d_0(x)/x) = \min_{d \in \mathcal{D}} L^*(d/x)$$, the rule $d_0(x)$ will minimize $r_d(\tau)$

Alternatively, define $\bar{p}(x) = \sum_{\Theta} p(x/\theta). \tau(\theta)$

Now $\dfrac{p(x/\theta). \tau(\theta)}{\bar{p}(x)} = q(\theta/x)$ whenever $\bar{p}(x) > 0$, is called the aposteriori distribution of θ given x.

Now $\bar{p}(x) = 0 \Rightarrow p(x|\theta) \cdot \tau(\theta) = 0 \ \forall \theta$

Assuming WLG that the outer sum in (A) is taken only over those x for which $\bar{p}(x) > 0$.

$$r_d(\tau) = \sum_{x \in \mathcal{X}} \bar{p}(x) \underbrace{\sum_{\theta \in \Theta} L(\theta, d(x)) \cdot q(\theta|x)}_{L^{**}(d|x)}$$

$$= \sum_{x \in \mathcal{X}} \bar{p}(x) \cdot L^{**}(d(x)|x),$$

where $L^{**}(d|x) = \dfrac{L^{*}(d|x)}{\bar{p}(x)}$, $\bar{p}(x) > 0$.

$L^{**}(d|x)$ can be considered the conditional mean loss for d (or posterior mean loss for d).

Bayes rule is determined by taking for each x that d ~~action~~ ~~for~~ which minimizes the posterior mean loss.

Case-II: Θ is measurable subset of a Euclidian space. $\tau(\theta)$ = pdf of θ and $p(x|\theta)$ = pdf of x.

Then $L^{*}(d|x) = \int_{\Theta} L(\theta, d(x))\, p(x|\theta)\, \tau(\theta)\, d\theta$ and for all

must all x, $d_0(x) \ni L^{*}(d_0(x)|x) = \min_{d \in \mathcal{D}} L^{*}(d(x)|x)$.

Let $\int_{\Theta} p(x|\theta)\, \tau(\theta)\, d\theta = \bar{p}(x)$, the marginal pdf of x.

~~Now~~ $\dfrac{p(x|\theta) \cdot \tau(\theta)}{\bar{p}(x)} = q(\theta|x)$, the posterior pdf of $\theta | x$. $(\bar{p}(x) > 0)$.

~~Now~~ $r_d(\tau) = \int_{\mathcal{X}} \bar{p}(x) \left[\int_{\Theta} L(\theta, d(x))\, q(\theta|x)\, d\theta \right] dx.$

Now $L^{**}(d|x) = \int_{\Theta} L(\theta, d(x))\, q(\theta|x)\, d\theta$ for all $x \ni \bar{p}(x) > 0$

$\therefore L^{**}(d_0(x)|x) = \min_{d \in \mathcal{D}} L^{**}(d(x)|x)$.

Case III: It is possible that x is discrete and θ is continuous (vice versa) and in that case, there arises one sum and one integral.

Particular Case

Case of point estimation of a real parameteric function with squared error loss.

~~Here τ: Θ~~
Now problem is to get a point estimate of $\gamma(\theta)$ with squared error.
Let $L(\theta, a) = (a - \gamma(\theta))^2$
Here if θ is discrete

$$L^*(a|x) = \sum_{\theta \in \Theta} (a - \gamma(\theta))^2 \, p(x|\theta) \tau(\theta)$$

If θ is continuous

$$L^*(a|x) = \int_{\Theta} (a - \gamma(\theta))^2 \, p(x|\theta) \tau(\theta) \, d\theta$$

Now minimizing $L^*(a|x)$ w.r.t. a, we have

$$a_0 = \sum_{\Theta} \gamma(\theta) \, p(x|\theta) \tau(\theta) \Big/ \underbrace{\sum_{\Theta} p(x|\theta) \cdot \tau(\theta)}_{(= \bar{p}(x))}$$

$$= \sum_{\Theta} \gamma(\theta) q(\theta|x) \quad (\Leftrightarrow E_x \gamma(\theta))$$

(or $\int_{\Theta} \gamma(\theta) p(x|\theta) \tau(\theta) d\theta / \bar{p}(x) = \int_{\Theta} \gamma(\theta) q(\theta|x) d\theta$) has the corresponding value of $d_0(x)$.

Here the posterior mean of $\gamma(\theta)$, for each x, has the corresponding value of $d_0(x)$.

Example 1: Let $X \sim R(0, \theta)$ and $\tau(\theta) = \theta e^{-\theta}$
We want to estimate $\gamma(\theta) = \theta$, $\theta \in (0, \infty)$

Here $a_0(x) = \int_0^\infty \cancel{\theta} \cdot \frac{1}{\cancel{\theta}} I_{(0,\theta)}(x) \cdot \theta \cdot e^{-\theta} d\theta \Big/ \int_0^\infty \frac{1}{\cancel{\theta}} I_{(0,\theta)}(x) \cdot \cancel{\theta} \cdot e^{-\theta} d\theta$

$$= \int_x^\infty \theta e^{-\theta} d\theta \Big/ \int_x^\infty e^{-\theta} d\theta = x+1 \quad \text{(check)}$$

Let $\tau(\theta) = c^p \theta^{p-1} e^{-c\theta} / \Gamma(p)$; $p > 0$

Hence $a_0(x) = \int_x^\infty \cancel{\theta} \cdot \frac{1}{\cancel{\theta}} \cancel{c^p} e^{-c\theta} \theta^{p-1} d\theta \Big/ \int_x^\infty \frac{1}{\theta} \cdot \cancel{c^p} e^{-c\theta} \theta^{p-1} d\theta$

$$= \int_{cx}^\infty e^{-u} u^{p-1} du \Big/ c \cdot \int_{cx}^\infty e^{-u} u^{p-2} du$$

Example 2: Let $X \sim Bin(n, \theta)$; n is known, $\theta \in (0,1)$

$$\tau(\theta) = \frac{1}{B(\alpha, \beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1}, \quad \alpha, \beta > 0 \quad \text{i.e. Beta prior.}$$

Bayes rule

$$d_0(x) = \frac{\cancel{\frac{1}{B(\alpha,\beta)}} \int_0^1 \theta \cancel{\binom{n}{x}} \theta^x (1-\theta)^{n-x} \theta^{\alpha-1} (1-\theta)^{\beta-1} d\theta}{\cancel{\frac{1}{B(\alpha,\beta)}} \int_0^1 \cancel{\binom{n}{x}} \theta^x (1-\theta)^{n-x} \theta^{\alpha-1} (1-\theta)^{\beta-1} d\theta}$$

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

$$= B(x+\alpha+1, n-x+\beta) / B(x+\alpha, n-x+\beta) = \frac{x+\alpha}{n+\alpha+\beta}.$$

So with Beta priors we can not get standard MVUE as a Bayes estimator as α, β may not be zero.

(??) **Example 3:** Let $X_i; i = 1(1)n \sim N(\theta, \sigma^2)$; σ^2 is known
$\theta \sim N(\gamma, \delta^2)$; γ, δ^2 are both known.
We know $\bar{X}_n$ is sufficient for Θ.
Now $\bar{x}|\theta \sim N(\theta, \frac{\sigma^2}{n})$ $\quad \theta \in \mathbb{R} = \Theta$

$$d_0(\underline{x}) = d_0(\bar{x}) = E(\theta | \bar{x} = \bar{x})$$
$$= \gamma + \frac{\delta^2}{\delta^2 + \frac{\sigma^2}{n}} (\bar{x} - \gamma) = \frac{\frac{\sigma^2}{n} \cdot \gamma + \delta^2 \cdot \bar{x}}{\delta^2 + \frac{\sigma^2}{n}} = \frac{\frac{\gamma}{\delta^2} + \frac{n}{\sigma^2} \bar{x}}{\frac{1}{\delta^2} + \frac{n}{\sigma^2}}$$

(check), which is the weighted average of γ and $\bar{x}$ and the weights being $\frac{1}{\delta^2}$ and $\frac{n}{\sigma^2}$ respectively.

Risk function $R_{d_0}(\theta) = E_{\bar{x}/\theta}\{d_0(\bar{x}) - \theta\}^2$

$$= Var(d_0(\bar{x})|\theta) + [E\{d_0(\bar{x})|\theta\} - \theta]^2$$

$$= \frac{\frac{n}{\sigma^2}}{\left(\frac{1}{\delta^2}+\frac{n}{\sigma^2}\right)^2} + \left[\frac{\frac{1}{\delta^2}(\nu-\theta)}{\frac{1}{\delta^2}+\frac{n}{\sigma^2}}\right]^2$$

Bayes risk $r_{d_0}(\tau)$ = Bayes risk of d_0

$$= E_\theta R_{d_0}(\theta) = \frac{\left(\frac{n}{\sigma^2}+\frac{1}{\delta^2}\right)}{\left(\frac{n}{\sigma^2}+\frac{1}{\delta^2}\right)^2} = \frac{1}{\left(\frac{n}{\sigma^2}+\frac{1}{\delta^2}\right)}$$

If the value of δ is very small the $d_0(\bar{x})$ is dominated by ν and ~~apt~~ apriori distribution is more precise. On the other hand if δ is very ~~small~~ large then the apriori distribution will be vague and $d_0(\bar{x})$ is dominated by $\bar{x}$ only.

(??) Example 4: In excercise 3, let $\nu(\theta) = \theta^2$. (H.T.)

Here $d_0(\underline{x}) = d_0(\bar{x}_n) = E(\nu(\theta)|\bar{x}=\bar{x})$

$$= E(\theta^2|\bar{x}=\bar{x})$$

$$= Var(\theta|\bar{x}=\bar{x}) + E^2(\theta|\bar{x}=\bar{x})$$

$$= \sigma_1^2(1-\rho^2) + \left\{\frac{\frac{\nu}{\delta^2}+\frac{n}{\sigma^2}\cdot\bar{x}}{\frac{1}{\delta^2}+\frac{n}{\sigma^2}}\right\}^2$$

$$= \frac{\delta^2\cdot\sigma^2}{\sigma^2+n\delta^2} + \left\{\frac{\frac{\nu}{\delta^2}+\frac{n}{\sigma^2}\cdot\bar{x}}{\frac{1}{\delta^2}+\frac{n}{\sigma^2}}\right\}^2$$

$$= \frac{1}{\frac{1}{\delta^2}+\frac{n}{\sigma^2}} + \left\{\frac{\frac{\nu}{\delta^2}+\frac{n}{\sigma^2}\bar{x}}{\frac{1}{\delta^2}+\frac{n}{\sigma^2}}\right\}^2$$

$(\theta, \bar{x}) \sim N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \sigma_{12})$
$\mu_1 = \nu,\ \sigma_1^2 = \delta^2,\ \alpha = 0$

Now $R_{d_0}(\theta) = E_{\bar{x}/\theta}\{d_0(\bar{x}) - \nu(\theta)\}^2$

$$= V_{\bar{x}|\theta}(d_0(\bar{x})) + [E\{d_0(\bar{x})|\theta\} - \nu(\theta)]^2$$

$$= \frac{1}{\{\frac{1}{\delta^2}+\frac{n}{\sigma^2}\}^4}\left[\frac{4\nu^2n^2}{\delta^4\sigma^4}V(\bar{x}|\theta) + \frac{n^4}{\sigma^4}V(\bar{x}^2|\theta)\right]$$

$$+ \left[\frac{1}{\frac{1}{\delta^2}+\frac{n}{\sigma^2}} + \frac{1}{[\frac{1}{\delta^2}+\frac{n}{\sigma^2}]^2}\left[\frac{\nu^2}{\sigma^4} + \frac{2\nu n}{\delta^2\sigma^2}\cdot\theta + \frac{n^2}{\sigma^4}\left\{\frac{\sigma^2}{n}+\theta^2\right\}\right] - \theta^2\right]^2$$

$$= \frac{\frac{4\nu^2n^2}{\delta^4\sigma^4}\cdot\frac{\sigma^2}{n} + \frac{n^4}{\sigma^4}\left(2+\frac{4n\theta^2}{\sigma^2}\right)}{\{\frac{1}{\delta^2}+\frac{n}{\sigma^2}\}^4} + \frac{\left(\frac{1}{\delta^2}+\frac{n}{\sigma^2}\right) + \left\{\frac{\nu^2}{\delta^4} + \frac{2n\cdot\nu\theta}{\delta^2\sigma^2} + \frac{n}{\sigma^2} + \theta^2\left(\frac{n^2}{\sigma^4}-1\right)\right\}}{[\frac{1}{\delta^2}+\frac{n}{\sigma^2}]^4}$$

$\because \bar{x}/\theta \sim N(\theta, \frac{\sigma^2}{n})$
$\bar{x}^2/\theta \sim \chi_1'^2\left(\frac{n\theta^2}{\sigma^2}\right)$
$V(\bar{x}^2/\theta) = 2n + 4\delta'$
$= 2 + \frac{4n\theta^2}{\sigma^2}$

$\therefore$ Minimum Bayes ~~Rate~~ risk $= E_\theta R_{d_0}(\theta)$

$$= \frac{1}{\{\frac{1}{\delta^2}+\frac{n}{\sigma^2}\}^4}\left[\frac{4n\nu^2}{\delta^4\sigma^2} + \frac{2n^4}{\sigma^4}\left\{1+\frac{2n}{\sigma^2}(\delta^2+\nu^2)\right\} + \left(\frac{1}{\delta^2}+\frac{2n}{\sigma^2}\right) + \frac{\nu^2}{\delta^4} + \frac{2n\nu^2}{\delta^2\sigma^2} + \left(\frac{n^2}{\sigma^4}-1\right)(\delta^2+\nu^2)\right]$$

Example 5: $x \sim Bin(n, \theta_1)$ and $y \sim Bin(n, \theta_2)$ and they are independent

Here $\underline{\theta} = (\theta_1, \theta_2)$, $\nu = \theta_1 - \theta_2$

Apriori θ_1, θ_2 independently distributed as $R(0,1)$

$f(\theta_1, \theta_2) = 1;\ 0 < \theta_1, \theta_2 < 1$

$f(x, y) = \binom{n}{x}\theta_1^x(1-\theta_1)^{n-x}\binom{n}{y}\theta_2^y(1-\theta_2)^{n-y}$

$d_0(x, y) = E(\theta_1 - \theta_2/x, y) = E(\theta_1|x) - E(\theta_2|y) = \frac{x-y}{n+2}$ (Check)

(H.T.) Use β-prior.

Theorem: Under squared error loss no unbiased estimate of a real parameter can be Bayes estimate except in trivial situations.

Proof: Let $\theta \in \Theta = \mathbb{R}$ and loss function, $L(\theta, a) = (a-\theta)^2$.

Let $d(x)$ be an unbiased estimator of θ: $E_\theta d(x) = \theta \quad \forall \theta$.

If possible, let $d(x)$ be the Bayes estimator w.r.t. some prior, $\tau(\theta)$

Now $d(x) = E_{\theta/x}(\theta)$

Bayes risk of $d = r_d(\tau) = E_{\theta,x}\{d(x)-\theta\}^2$

$$= E_{\theta,x} d^2(x) + E_{\theta,x}\theta^2 - 2E_{\theta,x}\{d(x).\theta\}.$$

Now $E_{\theta,x}\{d(x).\theta\} = E_\theta\{\theta . E_{x/\theta} d(x)\} = E_\theta \theta^2$

Again, $E_{\theta,x}\{d(x).\theta\} = E_x\{d(x) E_{\theta/x}(\theta)\}$

$$= E_x d^2(x).$$

So $r_d(\tau) = 0 = E_\theta E_{x/\theta}\{d(x)-\theta\}^2$

$\Rightarrow E_{x/\theta}\{d(x)-\theta\}^2 = 0$ for ~~all~~ almost all θ (regarding τ)

Again for a particular θ,

$E_{x/\theta}\{d(x)-\theta\}^2 = 0 \Rightarrow d(x) = \theta$ a.s.

Other loss functions generally used for point estimation of a real parameter θ.

a) Weighted squared error : $L(\theta, a) = w(\theta)(a-\theta)^2$; $w(\theta) > 0$

b) Absolute error : $L(\theta, a) = |a-\theta|$

Example: $X \sim Bin(n, \theta)$, n is known, $\theta \in (0,1)$

Let $L(\theta, a) = \dfrac{(a-\theta)^2}{\theta(1-\theta)}$, where $w(\theta) = \dfrac{1}{\theta(1-\theta)}$ minimum at $\theta = \frac{1}{2}$

$\rightarrow \infty$ as $\theta \rightarrow 0, 1$

In general $L^*(a|x) = \int_\Theta (\theta-a)^2 w(\theta)\tau(\theta) p(x|\theta)\, d\theta$

So, $a_x = \int_\Theta \theta . w(\theta)\tau(\theta)p(x|\theta)\,d\theta \Big/ \int_\Theta w(\theta)\tau(\theta)p(x|\theta)\,d\theta$

As if we are assuming, prior density $\propto w(\theta).\tau(\theta)$ and using unweighted squared error loss.

In this particular example, for prior $\tau(\theta)$,

Bayes rule $= d_0(x) = \int_0^1 \theta\, \theta^{-1}(1-\theta)^{-1} \tau(\theta)\, \theta^x (1-\theta)^{n-x} d\theta \Big/ \int_0^1 \theta^{-1}(1-\theta)^{-1}\tau(\theta)\theta^x(1-\theta)^{n-x} d\theta$

In particular, when $\tau(\theta) = 1$ i.e. $R(0,1)$

$$d_0(x) = \frac{B(x+1, n-x)}{B(x, n-x)} = \frac{x}{n}; \; 0 < x < n.$$

Example: $L(\theta, a) = |a-\theta|$

$L^*(a|x) = \int_\Theta |\theta - a|\tau(\theta)p(x|\theta)\,d\theta$

$d_0(x) =$ median of the distr. with density $\propto \tau(\theta).p(x|\theta)$, $\theta \in \Theta$.