2021. Sec 1. Q1.

$f(T_i) = \frac{1}{\lambda} \exp(-\frac{T_i}{\lambda})$. $F(T_i) = 1 - e^{-\frac{1}{\lambda} T_i}$ $E(T_i) = \lambda$. → missing

$U_i \overset{iid}{\sim} U(0,1)$. $R_i = I(T_i \in [0, U_i])$ → observed.

(a). $p(u_i, R_i) = p(R_i | u_i) \cdot p(u_i) = p(T_i \in [0, u_i])^{R_i} \cdot [1 - p(T_i \in [0, u_i])]^{1-R_i}$

$p(T_i \in [0, u_i]) = F_{T_i}(u_i) = 1 - \exp(-\frac{1}{\lambda} u_i)$

Thus:

$$L_n(\lambda | u_1, \cdots u_n, R_1, \cdots, R_n) = \prod_{i=1}^{n} [1 - \exp(-\frac{u_i}{\lambda})]^{R_i} \cdot [\exp(-\frac{u_i}{\lambda})]^{1-R_i}$$

$$l_n(\lambda | u_1, \cdots u_n, R_1 \cdots R_n) = \sum_{i=1}^{n} \left[R_i \cdot \log(1 - \exp(-\frac{u_i}{\lambda})) + (1-R_i) \cdot (-\frac{u_i}{\lambda}) \right]$$

$$SC_n = \frac{\partial l_n}{\partial \lambda} = \sum_{i=1}^{n} \left\{ R_i \cdot \frac{-e^{-\frac{u_i}{\lambda}}}{1 - e^{-\frac{u_i}{\lambda}}} \cdot \frac{u_i}{\lambda^2} + (1-R_i) \cdot \frac{u_i}{\lambda^2} \right\}$$

$$= \frac{1}{\lambda^2} \cdot \sum_{i=1}^{n} \left\{ \frac{-R_i u_i \cdot e^{-\frac{u_i}{\lambda}}}{1 - e^{-\frac{u_i}{\lambda}}} + (1-R_i) u_i \right\}.$$

$$SC_n \overset{set}{=} 0 : \quad \sum_{i=1}^{n} (1-R_i) u_i = \sum_{i=1}^{n} \frac{R_i u_i \cdot e^{-\frac{u_i}{\lambda}}}{1 - e^{-\frac{u_i}{\lambda}}}.$$ solve for $\hat{\lambda}_{MLE}$.

(b). $\sqrt{n}(\hat{\lambda} - \lambda) \xrightarrow{d} N(0, I_1(\lambda)^{-1})$.

$I_1(\lambda) = \lim_{n \to \infty} \frac{1}{n} I_n(\lambda)$.

$I_n(\lambda) = E\left(-\frac{\partial^2 l_n}{\partial \lambda^2}\right)$

$\theta = \lambda^{-1}$.

$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, V_1)$.

$g(x) = x^{-1}$.

$\sqrt{n}(g(\hat{\theta}) - g(\theta)) \xrightarrow{d} g'(\theta) \cdot N(0, V_1)$.

$\Rightarrow \sqrt{n}(\hat{\lambda} - \lambda) \xrightarrow{d} N(0, V_2)$.

$V_2 = (-\theta^{-2})^2 \cdot V_1 = \lambda^4 \cdot V_1$.

(c).

$$f(T_i, R_i, U_i) = f(T_i) \cdot f(U_i) \cdot f(R_i \mid T_i, U_i) = f(T_i).$$

$$l_n(\lambda) = \sum_{i=1}^{n} \log\left(\frac{1}{\lambda} \exp(-\frac{T_i}{\lambda})\right) = -n \cdot \log\lambda - \frac{\sum_{i=1}^{n} T_i}{\lambda}.$$

M-step: We would like to minimize $E\left(-n \cdot \log\lambda - \frac{\sum_{i=1}^{n} T_i}{\lambda} \,\middle|\, \tilde{R}, \tilde{U}, \lambda^{(K)}\right)$ w.r.t. λ.

$$\Rightarrow \text{minimize } -n \cdot \log\lambda - \frac{1}{\lambda}\sum_{i=1}^{n} E(T_i \mid R_i, U_i, \lambda^{(K)})$$

$$\Rightarrow \frac{-n}{\lambda} - \sum_{i=1}^{n} E(T_i \mid R_i, U_i, \lambda^{(K)}) \cdot \frac{-1}{\lambda^2} \overset{\text{set}}{=} 0.$$

$$\Rightarrow \lambda^{(K+1)} = \frac{1}{n}\sum_{i=1}^{n} E(T_i \mid R_i, U_i, \lambda^{(K)}).$$

E-step: Solve $E(T_i \mid R_i, U_i, \lambda^{(K)})$.

$$\begin{cases} R_i = 1 \Rightarrow T_i \in [0, U_i] \\ R_i = 0 \Rightarrow T_i \in (U_i, \infty). \end{cases}$$

Thus, $E(T_i \mid R_i = 1, U_i, \lambda^{(K)}) = \int_0^{U_i} t \cdot \frac{1}{\lambda} \exp(-\frac{t}{\lambda})\, dt$

$$= \left(-t \cdot e^{-\frac{t}{\lambda}} - \lambda e^{-\frac{t}{\lambda}}\right)\Big|_0^{U_i} = \lambda^{(K)} - U_i \cdot e^{-\frac{U_i}{\lambda^{(K)}}} - \lambda^{(K)} \cdot e^{-\frac{U_i}{\lambda^{(K)}}}$$

$$E(T_i \mid R_i = 0, U_i, \lambda^{(K)}) = \int_{U_i}^{\infty} t \cdot \frac{1}{\lambda} \exp(-\frac{t}{\lambda})\, dt = U_i \cdot e^{-\frac{U_i}{\lambda^{(K)}}} + \lambda^{(K)} \cdot e^{-\frac{U_i}{\lambda^{(K)}}}$$

(d). (i)

$$l_n^{pseudo} = \sum_{i: T_i \in [0, U_i]} \left(-\log\lambda - \frac{U_i}{2\lambda}\right) = -n_0 \cdot \log\lambda - \frac{\sum_{i: T_i \in [\,]} U_i}{2\lambda}$$

Not MLE !
Can only use CLT !

$$\frac{\partial l_n^{pseudo}}{\partial \lambda} \overset{\text{set}}{=} 0 \Rightarrow \tilde{\lambda} = \frac{\sum_{i: T_i \in [0, U_i]} U_i}{2 n_0}, \quad \text{where } n_0 = \sum_{i=1}^{n} I(T_i \in [0, U_i]).$$

$$\cancel{\sqrt{n}(\tilde{\lambda} - \lambda) \xrightarrow{d} N(0, I_1(\vec{\lambda}))}, \quad \cancel{\text{where } I_1(\lambda) = \lim_{n \to \infty} \frac{1}{n} \cdot E\left(-\frac{\partial^2 l_n^{pseudo}}{\partial \lambda^2}\right)}$$

Let $R_i = I(T_i \in [0, u_i])$. $\quad \tilde{\lambda} = \dfrac{\sum_{i=1}^{n} R_i U_i}{2 \cdot \sum R_i}$.

By CLT, $\sqrt{n}\left(\begin{pmatrix} \overline{R_i U_i} \\ \overline{R_i} \end{pmatrix} - \begin{pmatrix} E(R_1 U_1) \\ E(R_1) \end{pmatrix} \right) \xrightarrow{d} N\left(0, \begin{pmatrix} Var(R_1U_1) & cov(R_1U_1, R_1) \\ cov(R_1, R_1U_1) & Var(R_1) \end{pmatrix}\right)$.

Double Exp formula.

$$E(R_1U_1) = E_{U_1} E(R_1 U_1 | U_1) = E_{U_1}\left(U_1 \cdot E(R_1|U_1)\right) = E\left[U_1\left(1 - \exp\left(-\tfrac{1}{\lambda} U_1\right)\right)\right]$$

$$= E(U_1) - E\left(U_1 \exp\left(-\tfrac{U_1}{\lambda}\right)\right) = \tfrac{1}{2} - \left(-\lambda \cdot x e^{-\frac{x}{\lambda}} - \lambda^2 \cdot e^{-\frac{x}{\lambda}}\right)\Big|_0^1$$

$$= \tfrac{1}{2} + \lambda \cdot e^{-\frac{1}{\lambda}} + \lambda^2 \cdot e^{-\frac{1}{\lambda}} - \lambda^2$$

$E(R_1) = E_{U_1} E(R_1|U_1)$

$= 1 - E_U(e^{-\frac{U}{\lambda}})$

$= 1 + \lambda e^{-\frac{1}{\lambda}} - \lambda$.

$= E(R_1|U_1)$

$E(R_1) \cancel{=} P(T_i \in [0, u_i]) = F_{T_1}(u_1) = 1 - \exp\left(-\tfrac{1}{\lambda} u_1\right)$.

Consider $g: \mathbb{R}^2 \to \mathbb{R}^1$. $g\begin{pmatrix} x \\ y \end{pmatrix} = \dfrac{x}{2y}$. Then, $g\begin{pmatrix} \overline{R_i U_i} \\ \overline{R_i} \end{pmatrix} = \tilde{\lambda}$.

by Delta method, $\sqrt{n}(\tilde{\lambda} - \lambda_0) \xrightarrow{d} N(0, \Sigma)$. where:

$$\lambda_0 = \frac{E(R_1U_1)}{2 \cdot E(R_1)} = \frac{\frac{1}{2} + \lambda \cdot e^{-\frac{1}{\lambda}} + \lambda^2 \cdot e^{-\frac{1}{\lambda}} - \lambda^2}{2 \cdot \left(1 + \lambda e^{-\frac{1}{\lambda}} - \lambda\right)} \qquad \tilde{\lambda} \xrightarrow{P} \lambda_0.$$

(ii)

$$\nabla g\begin{pmatrix} x \\ y \end{pmatrix} = \left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right) = \left(\frac{1}{2y}, \frac{-x}{2y^2}\right).$$

$$\Sigma = \nabla g\left(\begin{pmatrix} E(R_1U_1) \\ E(R_1) \end{pmatrix}\right) \cdot \begin{pmatrix} Var(R_1U_1) & cov(R_1U_1, R_1) \\ cov(R_1, R_1U_1) & Var(R_1) \end{pmatrix} \cdot \nabla g\left(\begin{pmatrix} E(R_1U_1) \\ E(R_1) \end{pmatrix}\right)^T$$

$cov(R_1U_1, R_1) = E(R_1^2 U_1) - E(R_1^2)E(U_1) = E(R_1U_1) - E(R_1)E(U_1)$

$Var(R_1U_1) = E_{U_1} Var(R_1U_1 | U_1) + Var_{U_1} E(R_1U_1 | U_1)$.

2021. Sec 1. Q2.

See 761 mid 1.

BIOS 761

Spring 2022

Exam 1

There are 25 total points on this exam. Please show your solution in full details in order to get full credit.

1. Suppose $X_1, \ldots, X_n$ are i.i.d samples from $N(0, \sigma^2)$, where σ^2 is unknown.

   (i) [5 points] Consider the prior distribution of $\tau = 1/(2\sigma^2)$ to be a Gamma distribution, which has a density function of

   $$f(\tau) = \frac{1}{\Gamma(a)b^a}\tau^{a-1}e^{-\tau/b}, \text{ with } E(\tau) = ab \text{ and } Var(\tau) = ab^2.$$

   Under the loss function $L(\sigma^2, d) = (d - \sigma^2)^2/\sigma^4$, find the Bayes rule of σ^2.

   (ii) [5 points] Let $Y = n^{-1}\sum_{i=1}^n X_i^2$. Suppose we focus on the rule set $\mathcal{D} = \{\alpha Y + \beta\}$, where α and β are two real numbers. Prove that, under the loss function $L(\sigma^2, d) = (d - \sigma^2)^2$, any rule $Y + \beta$ with $\beta \neq 0$ in inadmissible in $\mathcal{D}$.

2. Suppose $X_1, \ldots, X_n$ are i.i.d samples from $N(\mu, 1)$ and the prior distribution of μ is $P(\mu = 0) = \lambda$ and $P(\mu = 1) = 1 - \lambda$ with $0 < \lambda < 1$. Consider the hypothesis test of $H_0 : \mu = 0$ versus $H_a : \mu = 1$ with the 0–1 loss function.

   (i) [5 points] Derive the Bayes rule for this test.

   (ii) [5 points] Derive the minimax rule for this test.

Solution given by Quefang.

Solution

1. (i) $$z \mid X \propto z^{a-1} e^{-z/b} \left(\frac{1}{\sigma}\right)^n \exp\left\{-\frac{\sum x_i^2}{2\sigma^2}\right\}$$

$$\propto z^{a-1} e^{-z/b} z^{\frac{n}{2}} \exp\left\{-z \sum x_i^2\right\}$$

$$= z^{a+\frac{n}{2}-1} \exp\left\{-z\left(\frac{1}{b} + \sum x_i^2\right)\right\}$$

Therefore $z \mid X \sim \text{Gamma}\left(a+\frac{n}{2}, \left(\frac{1}{b} + \sum x_i^2\right)^{-1}\right)$.

under the weighted loss function, the Bayes rule is

$$\delta(X) = \frac{E\left(\frac{1}{\sigma^2} \mid X\right)}{E\left(\frac{1}{\sigma^4} \mid X\right)}$$

For the numerator, $\frac{1}{\sigma^2} = 2z$. Let $S = \sum_{i=1}^{n} x_i^2$

numerator $= E(2z \mid X)$

$$= \int_0^\infty 2z \cdot \frac{1}{\Gamma(a+\frac{n}{2})\left(\frac{1}{b^{-1}+S}\right)^{a+\frac{n}{2}}} z^{a+\frac{n}{2}-1} \exp\left\{-z(b^{-1}+S)\right\} dz$$

$$= \frac{\frac{2}{b^{-1}+S}\Gamma(a+1+\frac{n}{2})}{\Gamma(a+\frac{n}{2})} \int_0^\infty \frac{1}{\Gamma(a+1+\frac{n}{2})\left(\frac{1}{b^{-1}+S}\right)^{a+1+\frac{n}{2}}} z^{a+\frac{n}{2}} \exp\left\{-z(b^{-1}+S)\right\} dz$$

$$= \frac{2a+n}{b^{-1}+S}$$

For the denominator, $\frac{1}{\sigma^4} = 4z^2$. Therefore,

$$E\left(\frac{1}{\sigma^4} \mid X\right) = E(4z^2 \mid X) = 4E(z^2 \mid X)$$

Since $z \mid X \sim \text{Gamma}\left(a+\frac{n}{2}, (b^{-1}+S)^{-1}\right)$

$$E(\tau^2|X) = \{E(\tau|X)\}^2 + \text{var}(\tau|X)$$

$$= \left(\frac{a+\frac{n}{2}}{b^{-1}+S}\right)^2 + \frac{a+n/2}{(b^{-1}+S)^2}$$

$$= \frac{(a+n/2)(a+n/2+1)}{(b^{-1}+S)^2}$$

Therefore $\delta(X) = \dfrac{2(a+n/2)/(b^{-1}+S)}{4(a+n/2)(a+n/2+1)/(b^{-1}+S)^2}$

$$= \frac{(b^{-1}+S)}{2(a+n/2+1)} = \frac{b^{-1}+\Sigma x_i^2}{n+2a+2}$$

(iii) Let $d = \alpha Y + \beta$. Then, the risk of d

$$R(\sigma^2, d) = E(\alpha Y + \beta - \sigma^2)^2$$

$$= E\{\alpha Y - E(\alpha Y) + E(\alpha Y) + \beta - \sigma^2\}^2$$

$$= \text{var}(\alpha Y) + \{E(\alpha Y) + \beta - \sigma^2\}^2$$

$$= \alpha^2 \frac{1}{n} \text{var}(X_1^2) + \{\alpha\sigma^2 + \beta - \sigma^2\}^2$$

$$= \frac{2\alpha^2}{n}\sigma^2 + \{(\alpha-1)\sigma^2 + \beta\}^2$$

When $\alpha = 1$, the risk of d is

$$2\sigma^2/n + \beta^2 > 2\sigma^2/n,$$

which is the risk of Y. Therefore any rule $Y+\beta$ with $\beta \neq 0$ is inadmissible.

2. (i) The posterior distribution of $\mu / x_1, \ldots x_n$

$$\propto \lambda \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left\{-\frac{\Sigma x_i^2}{2}\right\} I(\mu=0)$$

$$+ (1-\lambda)\left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left\{-\frac{\Sigma(x_i-1)^2}{2}\right\} I(\mu=1)$$

Under 0-1 loss, the Bayes rule is the posterior mode. Therefore, the Bayes rule is

$$\phi(x) = \begin{cases} 1 & \text{when } \lambda \exp\left\{-\frac{\Sigma x_i^2}{2}\right\} < (1-\lambda)\exp\left\{-\frac{\Sigma(x_i-1)^2}{2}\right\} \\ 0 & \text{otherwise.} \end{cases}$$

(ii) We know that the Bayes rule with a constant risk is also the minimax rule. In order for the above Bayes rule to have constant risk, we only need to choose $\lambda = 1/2$. Hence, the minimax rule is

$$\phi^*(x) = \begin{cases} 1 & \text{when } \exp\left\{-\frac{\Sigma x_i^2}{2}\right\} < \exp\left\{-\frac{\Sigma(x_i-1)^2}{2}\right\} \\ 0 & \text{otherwise.} \end{cases}$$