

EX 5

Q1

Question 1.

Let $Y \sim N(\mu, \sigma^2)$, where $|\mu| \leq a$ and σ is known.

1. Find the MLE for μ .
2. Consider the family of linear estimators $\hat{\mu}_c = cY$.
 1. Find the minimax estimator for μ within this family w.r.t. the quadratic loss. What is the corresponding minimax risk?
 2. Consider the following two-point prior on μ : $\pi(a) = \pi(-a) = 0.5$ and zero otherwise. Show that this is the least favorable prior. Is the corresponding Bayesian estimator admissible?

a

$$L(\mu, \sigma, y) = \sum_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2} \Rightarrow$$

$$l = C - \sum_{i=1}^n (y_i - \mu)^2 \Rightarrow \frac{\partial l}{\partial \mu} = 2 \sum_{i=1}^n (y_i - \mu) = 0 \iff \mu = \bar{y} \Rightarrow \mu_{MLE} = \bar{y}$$

b

i

$$L(\mu, \delta) = (\mu - \delta)^2$$

$$R(\mu, \delta) = E_Y((\mu - \delta)^2) = Bias(\delta, \mu)^2 + V(\delta) = (c-1)^2\mu^2 + c^2\sigma^2$$

$$\delta^* = \underset{\delta}{\operatorname{argmin}} \underset{\mu}{\sup} R(\mu, \delta) = \underset{\delta}{\operatorname{argmin}} \underset{\mu}{\sup} E_Y((\mu - \delta)^2)$$

in order to maximize the risk function we want μ to be as far as possible

$$\mu^* = a \Rightarrow c^* = \underset{c}{\operatorname{argmin}} (c-1)^2 a^2 + c^2 \sigma^2$$

$$\frac{\partial (c-1)^2 a^2 + c^2 \sigma^2}{\partial c} = ca - a + c\sigma^2 = 0 \iff c = \frac{a^2}{a^2 + \sigma^2} \Rightarrow c^* = \frac{a^2}{a^2 + \sigma^2}$$

thus we get: $\hat{\mu}_c = \frac{a^2}{a^2 + \sigma^2} y$

ii

$$\rho(\pi, \delta) = \sum_{j=1}^2 R(\mu_j, \delta) \pi(\mu_j) = 0.5R(a, \delta) + 0.5R(-a, \delta) = 0.5((c-1)^2 a^2 + c^2 \sigma^2 + (c-1)^2 a^2 + c^2 \sigma^2) = (c-1)^2 a^2 + c^2 \sigma^2$$

as we saw the minmax estimator minimizes this function thus δ_π^* is a minmax rule ,therefore it is a least favorable prior

δ_π^* is unique and $\delta_\pi^* < \infty$ thus it is admissible

Q2

Question 2.

1. Show that the minimax estimator for the probability of success p from $Y \sim B(n, p)$ w.r.t. the quadratic loss $L(p, a) = (p - a)^2$ is $\frac{Y + 5\sqrt{n}}{n + \sqrt{n}}$.

(hint: consider the $Beta(\alpha, \beta)$ family of priors on p and show that for a specific choice of α and β this estimator is a Bayes rule whose risk does not depend on p).

2. Is the above estimator admissible? What is the corresponding least favorable prior?
3. Whether a "usual" (MLE, UMVUE, etc.) estimator Y/n is admissible w.r.t the quadratic loss?

(hint: use the fact that Y/n is a Bayes rule w.r.t. to the loss function $L(p, a) = \frac{(p-a)^2}{p(1-p)}$)

a

$$L(p, a) = (p - a)^2$$

using the hint and taking a beta prior we can calculate the posterior:

$$\pi(p|y) \propto \binom{n}{y} p^y (1-p)^{n-y} p^{\alpha-1} (1-p)^{\beta-1} \propto p^{y+\alpha-1} (y+1-p)^{n-y+\beta-1}$$

thus we get: $\pi(p|y) \sim Beta(y + \alpha, n - y + \beta)$

we can calculate the bayes rule estimator:

$$\hat{p} = E(p|y) = \frac{y + \alpha}{y + \alpha + n - y + \beta} = \frac{y + \alpha}{\alpha + n + \beta}$$

plug it into the risk function: $R(p, \hat{p}) = Bias(p - \hat{p}) + V(\hat{p})$

$$Bias(\hat{p} - p) = \frac{np + \alpha}{\alpha + n + \beta} - p = \frac{\alpha - (\alpha + \beta)p}{\alpha + n + \beta}$$

$$V(\hat{p}) = \frac{V(y)}{(\alpha + n + \beta)^2} = \frac{np(1-p)}{(\alpha + n + \beta)^2}$$

plug both equations back in $R(p, \hat{p})$:

$$\begin{aligned} R(p, \hat{p}) &= \left(\frac{\alpha - (\alpha + \beta)p}{\alpha + n + \beta} \right)^2 + \frac{np(1-p)}{(\alpha + n + \beta)^2} = \frac{np(1-p) + (\alpha - (\alpha + \beta)p)^2}{(\alpha + n + \beta)^2} = \frac{np(1-p) + \alpha^2 - 2\alpha(\alpha + \beta)p + ((\alpha + \beta)p)^2}{(\alpha + n + \beta)^2} = \\ &= \frac{\alpha^2 + (-2\alpha(\alpha + \beta) + np)p + ((\alpha + \beta)^2 - n)p^2}{(\alpha + n + \beta)^2} \end{aligned}$$

thus if $(-2\alpha(\alpha + \beta) + np) = 0$ and $(\alpha + \beta)^2 - n = 0$

thus if $(-2\alpha(\alpha + \beta) + np) = 0$ and $(\alpha + \beta)^2 - n = 0$

we get a risk function which does not depend on p and we have a minmax estimator

if we take $\alpha = \beta = 0.5\sqrt{n}$:

$$(1) : (-2\alpha(\alpha + \beta) + np) = 0 \quad (2) : (\alpha + \beta)^2 - n = 0$$

thus we get that the minmax estimator is: $\hat{p} = E(p|y) = \frac{y + 0.5\sqrt{n}}{n + \sqrt{n}}$

b

the above estimator is admissible since it is a unique bayes rule
 $Beta(\alpha, \beta)$ is a least favorable prior (the risk function is constant for all p)

$$\begin{aligned} \rho(\pi, a) &= \int L(p, a) Beta(\alpha, \beta) dp = \int (p - a)^2 \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha, \beta)} dp \\ \int (p - a)^2 \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha, \beta)} dp &= \int \frac{p^{\alpha+1}(1-p)^{\beta-1}}{B(\alpha, \beta)} - 2a \frac{p^\alpha(1-p)^{\beta-1}}{B(\alpha, \beta)} - a^2 \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha, \beta)} dp = \\ \int \frac{p^{\alpha+1}(1-p)^{\beta-1}}{B(\alpha, \beta)} dp &= \frac{\alpha(\alpha+1)}{(\alpha+\beta+1)(\alpha+\beta)} B(\alpha, \beta) \int \frac{p^{\alpha+1}(1-p)^{\beta-1}}{B(\alpha, \beta)B(\alpha+2, \beta)} = \frac{\alpha(\alpha+1)}{(\alpha+\beta+1)(\alpha+\beta)} \\ \int 2a \frac{p^\alpha(1-p)^{\beta-1}}{B(\alpha, \beta)} dp &= \dots = 2a \frac{\alpha}{(\alpha+\beta)} \\ \int a^2 \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha, \beta)} dp &= a^2 \end{aligned}$$

plugging all the calculations back in we get:

$$\begin{aligned} \int (p - a)^2 \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha, \beta)} dp &= \frac{\alpha(\alpha+1)}{(\alpha+\beta+1)(\alpha+\beta)} - 2a \frac{\alpha}{(\alpha+\beta)} + a^2 \\ \frac{\alpha(\alpha+1)}{(\alpha+\beta+1)(\alpha+\beta)} - 2 \frac{y+\alpha}{\alpha+n+\beta} \frac{\alpha}{(\alpha+\beta)} + \left(\frac{y+\alpha}{\alpha+n+\beta} \right)^2 & \end{aligned}$$

c

$$\hat{p} = \frac{Y}{n} \text{ is admissible:}$$

$\frac{Y}{n}$ is a bayes rule w.r.t $\frac{(p-a)^2}{p(1-p)}$ (thus admissible)

assume there exists some estimator \tilde{p} such that: $MSE(p - \tilde{p}) < MSE(p - \hat{p})$ we get the following:

$$L^*(p, \tilde{p}) = \frac{E[(p - \tilde{p})^2]}{p(1-p)} \leq \frac{E[(p - \hat{p})^2]}{p(1-p)} = L^*(p, \hat{p})$$

which is a contradiction to the admissibility of \hat{p} w.r.t L^*

Q3

Question 3.

Assume we want to test two simple hypotheses $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$ based on a random sample $Y_1, \dots, Y_n \sim f(y|\theta)$, where the prior probabilities of the hypotheses are π_0 and $\pi_1 = 1 - \pi_0$ respectively. The loss is L_0 for erroneous rejection of H_0 , L_1 for erroneous rejection of H_1 and zero for a correct decision ("0 - L_i " loss).

1. Derive the resulting Bayes testing rule.
2. Interpretate the Bayes rule from the previous paragraph in terms of the frequentist approach. Is it a most powerful test? What is the critical value for the test statistic?
3. Let α and β be the probabilities of the I and II Type Errors respectively for the above test. Show that if $L_0\alpha = L_1\beta$, then it is also a minimax test.
4. Let Y_1, \dots, Y_{100} be a random sample from a $N(\mu, 25)$ distribution. Obtain the minimax test for testing $H_0 : \mu = 0$ vs. $H_1 : \mu = 2$ under "0 - L_i " loss, where $L_0 = 25$ and $L_1 = 10$.

a

$$\begin{aligned}
 & \text{define : } a = 0 \Rightarrow \text{accept } H_0, \quad a = 1 \Rightarrow \text{accept } H_1 \\
 & p_{yi} = P(\theta = \theta_i | y) \\
 & L(\theta, a) = aL_1 I(\theta = \theta_0) + (1 - a)L_0 I(\theta = \theta_1) \\
 & \rho(\theta, \delta) = E_{\theta|y}(L(\theta, a)) = \delta L_1 p_{y0} + (1 - \delta)L_0 p_{y1} = \delta(L_1 p_{y0} - L_0 p_{y1}) + L_0 p_{y1} \\
 & \delta^* = \operatorname{argmin}_{\delta} \delta(L_1 p_{y0} - L_0 p_{y1}) + L_0 p_{y1} \\
 & \text{thus if: } \delta^* = I(L_1 p_{y0} < L_0 p_{y1}) \\
 & p_{y0} = P(\theta = \theta_0 | y) = \frac{f(y|\theta_0)p_0}{f(y)} \\
 & p_{y1} = P(\theta = \theta_1 | y) = \frac{f(y|\theta_1)p_1}{f(y)} = \frac{f(y|\theta)(1 - p_0)}{f(y)} \\
 & L_1 p_{y0} < L_0 p_{y1} \iff \frac{p_{y1}}{p_{y0}} < \frac{L_0}{L_1} \iff \frac{\frac{f(y)}{f(y|\theta_1)(1-p_0)}}{\frac{f(y)}{f(y|\theta_0)p_0}} < \frac{L_0}{L_1} \iff \frac{f(y|\theta_0)}{f(y|\theta_1)} < \frac{(1 - p_0)L_0}{p_0 L_1} \\
 & \text{thus the bayes testing rule is: } \frac{p_0 L_1}{(1 - p_0)L_0} < \frac{f(y|\theta_1)}{f(y|\theta_0)}
 \end{aligned}$$

b

this is a likelihood ratio test and thus a most powerful test. the critical values of the test is: $\frac{p_0 L_1}{(1 - p_0)L_0}$

c

$$\begin{aligned}
 & \text{the risk function is:} \\
 & R(\theta, \delta^*) = \begin{cases} P(\delta^* = 1)L_0 & \theta = \theta_0 \\ P(\delta^* = 0)L_1 & \theta = \theta_1 \end{cases} = \begin{cases} \alpha L_0 & \theta = \theta_0 \\ \beta L_1 & \theta = \theta_1 \end{cases} = \begin{cases} \alpha L_0 & \theta = \theta_0 \\ \alpha L_0 & \theta = \theta_1 \end{cases} \\
 & \text{which does not depend on } \theta \text{ thus the test is also minmax}
 \end{aligned}$$

d

$$\begin{aligned}
 & * \text{LR- Likelihood ratio, LLR- Log Likelihood ratio} \\
 & LR = \frac{f(y|\mu_1)}{f(y|\mu_0)} = \exp \left\{ -\frac{1}{50} \sum_{i=1}^{100} (y_i - 2)^2 + \frac{1}{50} \sum_{i=1}^{100} y_i^2 \right\} \\
 & \frac{10p_0}{25(1 - p_0)} < LR \iff LLR > \ln \left(\frac{10p_0}{25(1 - p_0)} \right) \iff \left\{ -\frac{1}{50} \sum_{i=1}^{100} (y_i - 2)^2 + \frac{1}{50} \sum_{i=1}^{100} y_i^2 \right\} > \ln \left(\frac{10p_0}{25(1 - p_0)} \right) \iff \\
 & \iff -\frac{1}{50} \sum_{i=1}^{100} (y_i^2 - 4y_i + 4 - y_i^2) > \ln \left(\frac{10p_0}{25(1 - p_0)} \right) \iff 8\bar{y} - 8 > \ln \left(\frac{10p_0}{25(1 - p_0)} \right) \iff \\
 & \iff \bar{y} > 1 + \frac{1}{8} \ln \left(\frac{10p_0}{25(1 - p_0)} \right)
 \end{aligned}$$

just need to add plug in p_0 and we have a valid test

Q4

Question 4.

A device has been created to classify type of blood: A, B, AB or O. The device measures a certain quantity X, which has a density $f(x|\theta) = e^{-(x-\theta)}$, $x \geq \theta$. If $0 < \theta < 1$, the blood is of type AB; if $1 < \theta < 2$, the blood is of type A; if $2 < \theta < 3$, the blood is of type B; and if $\theta > 3$, the blood is of type O. It is known that in the population as a whole, $\theta \sim \exp(1)$.

The loss in misclassifying the blood is given in the following table:

		Classified As			
		AB	A	B	O
True Blood Type	AB	0	1	1	2
	A	1	0	2	2
	B	1	2	0	2
		O	3	3	0

A patient has been tested and $x=4$ is observed. What is the Bayes action?

calculate the posterior:

$$\pi(\theta|x) \propto e^{-(x-\theta)}e^{-\theta} I(0 < \theta < x) = e^{-x} I(0 < \theta < x) \text{ which is a constant function of } \theta$$

thus $\theta|x \sim U[0, x]$

$$\rho(\theta, \delta) = E_{\theta|x} L(\theta, \delta) \begin{cases} P(1 < \theta < 2) + P(2 < \theta < 3) + 3P(3 < \theta < 4) & \delta = AB \\ 1P(0 < \theta < 1) + 2P(2 < \theta < 3) + 3P(3 < \theta < 4) & \delta = A \\ 1P(0 < \theta < 1) + 2P(1 < \theta < 2) + 3P(3 < \theta < 4) & \delta = B \\ 2P(0 < \theta < 1) + 2P(1 < \theta < 2) + 2P(2 < \theta < 3) & \delta = O \end{cases}$$

for the case where $x=4$: $P(1 < \theta < 2) = 0.25$ plugging it back in yields:

$$\rho(\theta, \delta) = E_{\theta|x} L(\theta, \delta) \begin{cases} 1.25 & \delta = AB \\ 1.5 & \delta = A \\ 1.5 & \delta = B \\ 1.5 & \delta = O \end{cases}$$

the bayes action that minimizes $\rho(\theta, \delta)$ is $\delta = AB$

Q5

Question 5.

Children are given an intelligence test. The test result $X \sim N(\mu, 100)$, where μ is the true IQ (intelligence) level of child. It is known that in the population as a whole, μ is distributed $N(100, 225)$. A young Genius got 115 in the test.

- Find the Bayes estimate for the Genius' IQ w.r.t. the quadratic loss $L(\mu, a) = (\mu - a)^2$.
- In estimating IQ, it is deemed to be twice harmful to underestimate as to overestimate and the following loss is felt appropriate: $L(\mu, a) = 2(\mu - a)$, $\mu \geq a$ and $L(\mu, a) = (a - \mu)$, $\mu < a$. Find the Bayes estimate for the Genius' IQ w.r.t. this loss.
- Some people say that it is important to detect particularly high or low IQs and use the *weighted* quadratic loss
 $L(\mu, a) = (\mu - a)^2 e^{(\mu-100)^2/900}$ (note that this means that detecting an IQ of 145 or 55 is about nine times as important as detecting an IQ of 100). Find the Bayes estimate for Genius' IQ.
- Genius is to be classified as having below average IQ (less than 90), average (90 to 110), or above average IQ (over 110). Find the corresponding Bayes action and classify Genius to one of these three groups.

a

we saw that for: $X \sim N(\mu_x, \sigma_x^2)$, $\mu_x \sim N(\tilde{\mu}, \sigma^2)$

$$\mu_x|x \sim N\left(\frac{\sigma^2}{\sigma^2 + \sigma_x^2}x + \frac{\sigma_x^2}{\sigma^2 + \sigma_x^2}\tilde{\mu}, \left(\frac{1}{\sigma^2} + \frac{1}{\sigma_x^2}\right)^{-1}\right)$$

the bayes estimate w.r.t the quadratic loss is: $E_{\mu_x|x} = \frac{\sigma^2}{\sigma^2 + \sigma_x^2}x + \frac{\sigma_x^2}{\sigma^2 + \sigma_x^2}\tilde{\mu}$

pulging in our setup yields: $E_{\mu_x|x} = \frac{225}{325}115 + \frac{100}{325}100 = 110.3846$

b

$$\begin{aligned}\rho(x, \delta) &= E_{\mu|x} L(\mu, \delta) = \int_{-\infty}^{\delta} (\delta - \mu) f(\mu|x) d\mu + 2 \int_{\delta}^{\infty} (\mu - \delta) f(\mu|x) d\mu = \\ &= \delta \int_{-\infty}^{\delta} f(\mu|x) d\mu - \int_{-\infty}^{\delta} \mu f(\mu|x) d\mu + 2 \int_{\delta}^{\infty} \mu f(\mu|x) d\mu - 2\delta \int_{\delta}^{\infty} f(\mu|x) d\mu = \\ &= \delta F(\delta) - \int_{-\infty}^{\delta} \mu f(\mu|x) d\mu + 2 \int_{\delta}^{\infty} \mu f(\mu|x) d\mu - 2\delta(1 - F(\delta))\end{aligned}$$

$$\frac{\partial \rho(x, \delta)}{\partial \delta} = F(\delta) + \delta f(\delta|x) - \delta f(\delta|x) - 2\delta f(\delta|x) - 2 + 2F(\delta) + 2\delta f(\delta|x) = 3F(\delta|x) - 2 = 0 \iff \delta = F^{-1}(2/3) \Rightarrow \delta^* = F^{-1}(2/3)$$

using r qnorm function we can estimate the genius IQ w.r.t to this loss which yileds: 113.9685

$$* qnorm(2/3, 110.3846, sqrt((\frac{1}{100} + \frac{1}{225})^{-1}))$$

c

$$\begin{aligned}\rho(x, \delta) &= E_{\mu|x} L(\mu, \delta) = \int_{-\infty}^{\infty} (\delta - \mu)^2 e^{\frac{(\mu-100)^2}{900}} f(\mu|x) d\mu \\ \frac{\partial \rho(x, \delta)}{\partial \delta} &= 2 \int_{-\infty}^{\infty} (\delta - \mu) e^{\frac{(\mu-100)^2}{900}} f(\mu|x) d\mu\end{aligned}$$

setting the derivative to 0 yields:

$$\begin{aligned}0 &= \int_{-\infty}^{\infty} (\delta - \mu) e^{\frac{(\mu-100)^2}{900}} f(\mu|x) d\mu = \delta \int_{-\infty}^{\infty} e^{\frac{(\mu-100)^2}{900}} f(\mu|x) d\mu - \int_{-\infty}^{\infty} \mu e^{\frac{(\mu-100)^2}{900}} f(\mu|x) d\mu \Rightarrow \\ &\Rightarrow \delta^* = \frac{\int_{-\infty}^{\infty} \mu e^{\frac{(\mu-100)^2}{900}} f(\mu|x) d\mu}{\int_{-\infty}^{\infty} e^{\frac{(\mu-100)^2}{900}} f(\mu|x) d\mu}\end{aligned}$$

lets break down this expression, fisrt define $E(\mu|x) = \theta$, $V(\mu|x) = \tau^2$

also lets forget about the normalizing constant for the next calcultations as it will cancel out

$$e^{\frac{(\mu-100)^2}{900}} f(\mu|x) = e^{\frac{(\mu-100)^2}{900} - \frac{(\mu-\theta)^2}{2\tau^2}}$$

taking a closer look at the power:

$$\begin{aligned}\frac{(\mu-100)^2}{900} - \frac{(\mu-\theta)^2}{2\tau^2} &= \frac{1}{2} \frac{\tau^2(\mu-100)^2 - 450(\mu-\theta)^2}{450\tau^2} = \frac{1}{2} \frac{\tau^2(\mu^2 - 2 * 100\mu + 100^2) - 450(\mu^2 - 2\mu\theta + \theta^2)}{450\tau^2} = \\ &= \frac{1}{2} \frac{(\tau^2 - 450)\mu^2 + (900\theta - 200\tau^2)\mu - 450\theta^2 + 100^2\tau^2}{450\tau^2} = \frac{1}{2} \left[\left(\frac{1}{450} - \frac{1}{\tau^2} \right) \mu^2 - 2 \left(\frac{100}{450} - \frac{\theta}{\tau^2} \right) \mu + \frac{100^2}{450} - \frac{\theta^2}{\tau^2} \right] = \\ &= -\frac{1}{2} \left(\frac{1}{\tau^2} - \frac{1}{450} \right) [\mu - \frac{\frac{\theta}{\tau^2} - \frac{100}{450}}{\frac{1}{\tau^2} - \frac{1}{450}}]^2 + C\end{aligned}$$

where C does not depend on μ

plugging it back in we get:

$$\delta^* = \frac{\int_{-\infty}^{\infty} \mu e^{-\frac{1}{2}(\frac{1}{\tau^2} - \frac{1}{450})[\mu - \frac{\frac{\theta}{\tau^2} - \frac{100}{450}}{\frac{1}{\tau^2} - \frac{1}{450}}]^2 + C} d\mu}{\int_{-\infty}^{\infty} e^{-\frac{1}{2}(\frac{1}{\tau^2} - \frac{1}{450})[\mu - \frac{\frac{\theta}{\tau^2} - \frac{100}{450}}{\frac{1}{\tau^2} - \frac{1}{450}}]^2 + C} d\mu}$$

if we multiple and divide by the relevent normalizing consant the numerator is the expected value of a normally distributed

variable with an expected value of $\frac{\theta}{\tau^2} - \frac{100}{450}$, and the denominator is 1 (an integral over the density)

plugging our known θ, τ , yields:

$$(110.3846/69.23077 - 100/450)/(1/69.23077 - 1/450) = 112.2727$$

d

$$P(below|x = 115) = pnorm(90, 110.3846, sqrt(900/13)) = 0.007$$

$$P(avg|x = 115) = pnorm(90, 110.3846, sqrt(900/13)) = 0.475$$

$$P(above|x = 115) = 1 - pnorm(110, 110.3846, sqrt(900/13)) = 0.518$$

$$\rho(\mu, \delta) = E_{\mu|x} L(\mu, \delta) = \begin{cases} 1 - P(below|x = 115) & \delta = below \\ 1 - P(avg|x = 115) & \delta = avg \\ 1 - P(above|x = 115) & \delta = above \end{cases} = \begin{cases} 0.993 & \delta = below \\ 0.525 & \delta = avg \\ 0.482 & \delta = above \end{cases}$$

thus $\delta^* = above$