Department of Statistics, The Chinese University of Hong Kong
STAT5010 Advanced Statistical Inference (Term 1, 2022–23)

Assignment 4 · due on 5 December 2022
Please submit your answers in `.pdf` format via Blackboard.

1. (Hint for Problem 3 in Asg 5) Based on Theorem 3.5 in TPE, we first show the risk of the Bayes estimator η is given by

$$R[\eta, E(\eta|X)] = R[\eta, -\bigtriangledown \log h(X)] - \frac{2p}{\sigma^2+\tau^2} + \sum_{i=1}^{p} E\left(\frac{X_i-\mu}{\sigma^2+\tau^2}\right)^2 ,$$

since $m(x)$ in the normal hierarchical model is given by

$$m(x) = C(\sigma,\tau)\exp\left(-\frac{1}{2(\sigma^2+\tau^2)}\sum_{i=1}(X_i-\mu)^2\right) \;\Rightarrow\; \frac{\partial}{\partial x_i}\log m(x) = -\frac{(x_i-\mu)}{\sigma^2+\tau^2} \,.$$

As the normal hierarchical model has the following properties:

$$h(x) = \prod_{i=1}^{p}\left[\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{x_i^2}{\sigma^2}\right)\right] \;\Rightarrow\; \frac{\partial}{\partial x_i}\log h(x) = -\frac{x_i}{\sigma^2} \,,$$

$$R[\eta, -\bigtriangledown \log h(X)] = \ldots = \frac{p}{\sigma^2} \,,$$

which leads to

$$\begin{aligned} R[\eta, E(\eta|X)] &= \ldots \\ &= \frac{p\tau^4}{\sigma^2(\sigma^2+\tau^2)^2} + \left(\frac{\sigma^2}{\sigma^2+\tau^2}\right)^2 \sum a_i^2 \,, \end{aligned}$$

where $a_i = \eta_i - \mu/\sigma^2$. Finally, since the risk is given by

$$R[\eta, E(\eta|X)] = \frac{p\tau^4}{\sigma^2(\sigma^2+\tau^2)^2} + \left(\frac{\sigma^2}{\sigma^2+\tau^2}\right)^2 k$$

when $\sum a_i^2 = k$. Try to find the expression of η_i that solves $\sum a_i^2 = k$.

2. (K12.11) Laplace's law of succession gives a distribution for Bernoulli variables $X_1, X_2, \ldots$ in which $\Pr(X_1 = 1) = 1/2$, and

$$\Pr(X_{j+1} = 1 \mid X_1 = x_1, \ldots, X_j = x_j) = \frac{1 + x_1 + \ldots + x_j}{j+2}, \quad j \geq 1.$$

Consider testing the hypothesis H_1 that $X_1, \ldots, X_n$ have this distribution against the null hypothesis H_0 that the variables are iid with $\Pr(X_i = 1) = 1/2$. If $n = 10$, find the best test with size $\alpha = 0.05$. What is the power of this test?

3. (K12.17, p-values) Suppose we have a family of tests φ_α, $\alpha \in (0,1)$ indexed by level (so φ_α has level α), and that these tests are "nested" in the sense that $\varphi_\alpha(x)$ is nondecreasing as a function of α. We can then define the "p-value" or "attained significance" for observed data x as $\inf\{\alpha : \varphi_\alpha(x) = 1\}$, thought of as the smallest value for α where test φ_α rejects H_0. Suppose we are testing $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$ and that the densities for data X have monotone likelihood ratios in T. Further suppose T has continuous distribution.

   (a) Show that the family of uniformly most powerful tests are nested in the sense described.

   (b) Show that if $X = x$ is observed, the p-values $P(x)$ is

   $$\Pr{}_{\theta_0}(T(X) > t),$$

   where $t = T(x)$ is the observed value of T.

   (c) Determine the distribution of the p-value $P(X)$ when $\theta = \theta_0$.

4. (K12.19) Suppose X has a Poisson distribution with parameter λ. Determine the uniformly most powerful test of $H_0 : \lambda \leq 1$ versus $H_1 : \lambda > 1$ with level $\alpha = 0.05$.

5. (K12.22) Suppose we observe a single observation X from $N(\theta, \theta^2)$.

   (a) Do the densities for X have monotone likelihood ratios?

   (b) Let ϕ^* be the best level $alpha$ test of $H_0 : \theta = 1$ versus $H_1 : \theta = 2$. Is ϕ^* also the best level α test of $H_0 : \theta = 1$ versus $H_1 : \theta = 4$?

6. (K12.29) Suppose Y_1 and Y_2 are independent variables, both uniformly distributed on $(0, \theta)$, but our observation is $X = Y_1 + Y_2$.

   (a) Show that the densities for X have monotone likelihood ratios.

   (b) Find the UMP level α test of $H_0 : \theta = \theta_0$ versus $H_1 : \theta > \theta_0$ based on X.

7. Let the variables X_i, $1 \leq i \leq n$ be independently distributed with distribution $Poisson(\lambda_i)$, $1 \leq i \leq n$ respectively. For testing the hypothesis

   $$H_0 \ : \sum_{i=1}^{n} \lambda_i \leq a \quad v.s. \quad H_1 \ : \sum_{i=1}^{n} \lambda_i > a \ .$$

   (for example, that the combined radioactivity of a number of pieces of radioactive material does not exceed a), show that there exists a UMP test, which rejects when $\sum_{i=1}^{n} X_i > C$.

8. (Optional) Suppose we wish to test n hypotheses $H_1, H_2, \cdots, H_n$. And we assume that the null p-values are uniformly distributed. In this problem, we are interested in procedures which operate in two steps:

- **Step 1**. Select a set $\mathcal{S} \subset \{1, 2, \cdots, n\}$ of 'promising' hypotheses.
- **Step 2**. Apply a multiple testing procedure to test those hypotheses in $\mathcal{S}$, namely, $\{H_i\}_{i \in \mathcal{S}}$.

Below we shall assume that the selection step is monotone in the following sense: if $\mathcal{S}(p)$ is the set of selected hypotheses on the basis of the n p-values $(p_1, \cdots, p_n)$, then $p_i \leq p'_i$ for all i ($p \leq p'$ for short) implies that $\mathcal{S}(p') \subset \mathcal{S}(p)$.

(a) Suppose we apply the Benjamini-Hochberg (BH) procedure to the selected set of hypotheses with an FDR target level set to q (this means that the critical thresholds would be equal to $q_i/|\mathcal{S}|$ for $i = 1, 2, \cdots, |\mathcal{S}|$). Under independence of all n p-values, would you expect FDR control at level q? Explain why or why not. Similarly, imagine you were to apply the Bonferroni correction at level $\alpha/|\mathcal{S}|$, would you expect FWER control at level α?

(b) Suppose now that you apply the BH procedure to the selected hypotheses with an FDR target set to $q|\mathcal{S}|/n$. Under independence between all the p-values, show that this two-step procedure would control the FDR at level q.
Hint: You may use the following claim: whenever a function $f : (p_1, \cdots, p_n) \to [0, 1]$ is nonincreasing (recall that this means that $p \leq p'$ implies $f(p) \geq f(p')$), we have

$$\mathbb{E}\left[\frac{I_{\{p_i < f(p)\}}}{f(p)}\right] \leq 1,$$

provided the p-values obey the PRDS property.

(c) Suppose then that the n p-values actually obey the PRDS property, would FDR control at level q continue to hold? Explain why or why not.

(d) Under independence between the p-values, can I set a nominal threshold higher than $q|\mathcal{S}|/n$ and expect FDR control in general? Explain why or why not.

(e) Describe an application where it might make sense to use the two-step procedure we have just described.

(f) Prove the claim from the hint.

STAT 5010 Homework 4 Dec 5th Xiaocheng Zhou 1155184323

1. Same with the result of HW 3.3(c)

$$\sum_{i=1}^{p} a_i^2 \geq \frac{1}{n}\left(\sum_{i=1}^{p} a_i\right)^2 = \frac{1}{n}\left(\sum_{i=1}^{p} \eta_i - \frac{p\mu}{\sigma^2}\right)^2$$ where the equality holds when $a_i = a_j$ for $\forall i, j$

then $p a_i^2 = k \Rightarrow a_i = \eta_i - \frac{\mu}{\sigma^2} = \sqrt{\frac{k}{p}}$

Thus we can solve $\eta_i = \sqrt{\frac{k}{p}} + \frac{\mu}{\sigma^2}$, $i = 1, \cdots, p$, and in this case $R(\eta, \delta')$ reaches minimum.

2. Construct likelihood ratio as follow, considering $H_0: P(X_i = 1) = \frac{1}{2}$ for $\forall i$
H_1: Laplace Law.

$$L(\vec{x}) = \frac{P_1(\vec{x})}{P_0(\vec{x})} = \frac{P(X_1 = x_1) \prod_{j=1}^{n-1} P(X_{j+1} = x_{j+1} \mid X_1 = x_1, \cdots, X_j = x_j)}{(\frac{1}{2})^n}$$

In this question, $n = 10$, then we can write that

$$L(\vec{x}) = \frac{T(\vec{x})}{(\frac{1}{2})^9} \quad \text{where} \quad T(\vec{x}) = \prod_{j=1}^{p} \left(\frac{1 + \sum_{i=1}^{j} x_i}{j+2}\right)^{x_{j+1}} \left(1 - \frac{1 + \sum_{i=1}^{j} x_i}{j+2}\right)^{x_{j+1}}$$

then L is monotone (non-decreasing) likelihood ratio in T

Let

$$\varphi(\vec{x}) = \begin{cases} 1 & \text{if } T(\vec{x}) > c \\ r & \text{if } T(\vec{x}) = c \\ 0 & \text{if } T(\vec{x}) < 0 \end{cases} \quad \text{with } r \,\&\, c \text{ to be determined.}$$

$$\alpha = \mathbb{E}_0 \varphi(\vec{X}) = P_0(T(\vec{X}) > c) + r P_0(T(\vec{X}) = c) = 0.05$$

then $c = \cdots$ $r = \cdots$. And with such c and r, we have

$$\beta = \mathbb{E}_1 \varphi(\vec{X}) = \cdots$$

3.

4. Construct likelihood ratio as follow.

$$L(x) = \frac{P_1(x)}{P_0(x)} = \frac{e^{-\lambda_1}\lambda_1^x/x!}{e^{-\lambda_0}\lambda_0^x/x!} = e^{-(\lambda_1-\lambda_0)}\left(\frac{\lambda_1}{\lambda_0}\right)^x$$

Since $\lambda_0 \leq 1 < \lambda_1$, then $\left(\frac{\lambda_1}{\lambda_0}\right) > 1$.
Take $T(X)=X$, then L is nondecreasing likelihood in T.

Let
$$\varphi(X) = \begin{cases} 1 & \text{if } x > c \\ r & \text{if } x = c \\ 0 & \text{if } x < c \end{cases}$$

$$\alpha = \max_{\lambda_0} \mathbb{E}(\varphi(X)|\lambda_0) = \mathbb{E}(\varphi(X)|\lambda=1) = P(X>c|\lambda=1) + rP(X=c|\lambda=1)$$

$\Rightarrow c=3$ (when $P(X>3|\lambda=1) = 0.01899$ is around and less than 0.05)
in order to let the equality hold, take $r = 0.5058$

Thus the UMP is $\varphi(X) = \mathbb{1}(X>3) + 0.5058 \times \mathbb{1}(X=3)$

5. (a) Assume $H_0: \theta = \theta_0$, $H_1: \theta=\theta_1$, then the likelihood ratio is

$$L(x) = \frac{P_1(x)}{P_0(x)} = \frac{\frac{1}{\sqrt{2\pi}\,\theta_1}\exp\left\{-\frac{(x-\theta_1)^2}{2\theta_1^2}\right\}}{\frac{1}{\sqrt{2\pi}\,\theta_0}\exp\left\{-\frac{(x-\theta_0)^2}{2\theta_0^2}\right\}} =$$

$$= \frac{\theta_0}{\theta_1}\exp\left\{-\left[\frac{(x-\theta_1)^2}{2\theta_1^2} - \frac{(x-\theta_0)^2}{2\theta_0^2}\right]\right\}$$

$$= \frac{\theta_0}{\theta_1}\exp\left\{-\frac{\theta_0^2-\theta_1^2}{2\theta_0^2\theta_1^2}\left(x-\frac{\theta_1\theta_0}{\theta_0+\theta_1}\right)^2\right\}\exp\left\{\frac{\theta_0-\theta_1}{2(\theta_0+\theta_1)}\right\}$$

L is a symmetrical function in x (symmetrical with $x = \frac{\theta_1\theta_0}{\theta_0+\theta_1}$)

so there is not monotone likelihood ratios for X.

(b) If ϕ^* is MP test for both testings

6. (a) $f_X(x) = \int f_{Y_1}(y_1) f_{Y_2}(x-y_1)\,dy_1 = \frac{1}{\theta}\int_0^\theta f_{Y_2}(x-y_1)\,dy_1$

if $0 \le x \le \theta$, $\quad \frac{1}{\theta}\int_0^\theta f_{Y_2}(x-y_1)\,dy_1 = \frac{1}{\theta}\int_0^x \frac{1}{\theta}\,dy_1 = \frac{x}{\theta^2}$

if $\theta < x < 2\theta$, $\quad \frac{1}{\theta}\int_0^\theta f_{Y_2}(x-y_1)\,dy_1 = \frac{1}{\theta}\int_{x-\theta}^\theta \frac{1}{\theta}\,dy_1 = \frac{1}{\theta^2}(2\theta - x)$

then the likelihood ratio is

$$L(x) = \frac{p_{\theta_1}(x)}{p_{\theta_0}(x)} = \left(\frac{\theta_0}{\theta_1}\right)^2 \quad \text{if } 0 \le x \le \theta_1 < \theta_2 \quad (\text{non-decreasing in } x)$$

$$L(x) = \frac{p_{\theta_1}(x)}{p_{\theta_0}(x)} = \left(\frac{\theta_0}{\theta_1}\right)^2\left(1 + \frac{2\theta_0}{2\theta_0 - x}\right) \text{ if } \theta_0 < x \le \min(\theta_1, 2\theta_0) \quad (\text{non-decreasing in } x)$$

$$L(x) = \frac{p_{\theta_1}(x)}{p_{\theta_0}(x)} = \left(\frac{\theta_0}{\theta_1}\right)^2\left(1 + \frac{2(\theta_1-\theta_0)}{2\theta_1 - x}\right) \text{ if } \theta_0 < \min(\theta_1, 2\theta_0) \quad (\text{non-decreasing in } x)$$

Thus it has MLR in $T(X) = X$

(b) Let UMP test has form like

$$\varphi(X) = \begin{cases} 1, & X > k \\ 0, & X < k \end{cases}$$

$$\alpha = \mathbb{E}_0 \varphi(X) = P_0(X > k) = 1 - F_0(k)$$

then $k = F_0^{-1}(1-\alpha) = \begin{cases} \theta_0\sqrt{2(1-\alpha)}, & \alpha \ge \frac{1}{2} \\ 2 - \sqrt{2\alpha}, & \alpha < \frac{1}{2}. \end{cases}$

7. Construct statistic $T(\vec{X}) = \sum_{i=1}^n X_i$, then $T \sim \text{Poisson}(\sum_{i=1}^n \lambda_i)$

then $L(\vec{x}) = \frac{P_1(\vec{x})}{P_0(\vec{x})} = \frac{e^{-\mu_1}\cdot \mu_1^t / t!}{e^{-\mu_0}\cdot \mu_0^t / t!}$

$= e^{-(\mu_1 - \mu_0)}\left(\frac{\mu_1}{\mu_0}\right)^t \quad$ where $\mu = \sum_{i=1}^n \lambda_i$. $t = \sum_{i=1}^n x_i$

is nondecreasing in t, so there exists a UMP test like.

$$\varphi(\vec{\ }) = \begin{cases} 1 & \sum X_i > c \\ r & \sum X_i = c \\ 0 & \sum X_i < c \end{cases}$$