Jiapeng Wang.

Problem 5.1. Let $X \sim \text{Binomial}(m, p)$ and $Y \sim \text{Binomial}(n, \pi)$ be independent random variables. Let a and b be constants satisfying $0 \leq b \leq 1, 0 \leq a + b \leq 1$. Prove or disprove that estimators of the form $a\frac{Y}{n} + b$ are admissible for estimating p.

Suppose it's under square loss, we know that $-1 \leq a \leq 1$

$$R(p, a\tfrac{Y}{n}+b) = E(a\tfrac{Y}{n}+b-p)^2 = E^2(a\tfrac{Y}{n}+b-p) + Var(\tfrac{aY}{n})$$

$$= (a\pi+b-p)^2 + \frac{a^2\pi}{n}(1-\pi) \geq (a\pi+b-p)^2 = R(p, a\pi+b)$$

So, by definition, $a\frac{Y}{n}+b$ is not admissible.

12

Problem 5.2. Consider a random sample $X_1, \dots, X_n$ where each observable is from a common, unknown distribution with CDF denoted by F. Let $\mathbb{I}\{\cdot\}$ denote the indicator function. Define the decision procedure

$$\delta(\boldsymbol{X}) = \frac{\sum_{i=1}^{n} \mathbb{I}\{X_i \leq 0\}}{\sqrt{n}} \times \frac{1}{1+\sqrt{n}} + \frac{1}{2(1+\sqrt{n})}.$$

Prove that δ is minimax for estimating $F(0) = \mathbb{P}[X_i \leq 0]$ under squared error loss.

Problem 5.3. Let $X \sim \text{Poisson}(\lambda)$. For which constants a and b is $aX + b$ admissible for estimating $\mathbb{E}[X]$

We can transfer the problem like: Consider a random sample Y from Binomial $[n, \theta]$, $\theta = P[X_i \geq 0] = F(0)$, Y is equivalent to $\sum_{i=1}^{n} I(X_i \leq 0)$

Then $\delta(Y) = \frac{Y}{\sqrt{n}} \cdot \frac{1}{1+\sqrt{n}} + \frac{1}{2(1+\sqrt{n})}$, We need to prove δ is minimax for estimating θ

$$R(\theta, \delta(y)) = E_\theta(\theta - \delta(y))^2 = (\theta - E\delta(y))^2 + \mathrm{Var}\,\delta(y)$$
$$= \left(\theta - \frac{\sqrt{n}\,\theta}{1+\sqrt{n}} - \frac{1}{2(1+\sqrt{n})}\right)^2 + \frac{n\theta \cdot (1-\theta)}{n} \cdot \frac{1}{(1+\sqrt{n})^2}$$
$$= \frac{1}{4(1+\sqrt{n})^2}, \text{ which is constant}$$

And obviously $\delta(y)$ is sufficient for θ. Then we consider $r(a|y=k) = E[(\theta-a)^2 | y=k] = a^2 - 2aE(\theta|y=k) + E(\theta|y=k)^2$, in order to minimize it, we get $\delta^*(y=k) = E(\theta|y=k)$. We already know that Beta distribution and Binomial distribution are conjugate. So we can assume the prior distribution of θ: $\theta \sim \mathrm{Beta}(\alpha, \beta)$

then $\theta | y=k \sim \mathrm{Beta}(\alpha+k, \beta+n-k) \Rightarrow \delta^*(y=k) = \frac{\alpha+k}{\alpha+\beta+n}$

Let $\alpha = \beta = \frac{\sqrt{n}}{2}$, then $\delta^*(y=k) = \delta(y=k) = \frac{k}{\sqrt{n}} \cdot \frac{1}{1+\sqrt{n}} + \frac{1}{2(1+\sqrt{n})}$

So we find a prior distribution π^* s.t ① δ is Bayes for π^*

② $R(\theta, \delta)$ is constant for all θ. Then by Thm E.1.1,

$\delta(Y)$ is minimax, i.e $\delta(\vec{X})$ is minimax

$\square$

Problem 5.3. Let $X \sim \text{Poisson}(\lambda)$. For which constants a and b is $aX + b$ admissible for estimating $\mathbb{E}[X]$ under squared error loss?

From Thm E.1.4, we get that $aX+b$ is inadmissible under squared error loss when any of the following condition holds, (i) $a>1$ (ii) $a<1$, (iii) $a=1$ and $b\neq 0$, so $aX+b$ is admissible only if $a=1$, $b=0$, i.e. $aX+b=X$. Then from Lemma E.6.2, for Poisson, we get that X is admissible for estimating EX.

$\square$

Problem 5.4. Suppose that $X \mid \theta \sim \mathcal{N}(\theta, 1^2)$ and θ has the (improper) prior density $\pi(\theta) = \exp(\theta)$, where $-\infty < \theta < \infty$. Obtain the Bayes estimator of θ under squared error loss.
Bonus: Show that this estimator is neither minimax nor admissible.

$$\pi(\theta \mid \vec{x}) \propto f(\vec{x} \mid \theta) \cdot \pi(\theta) \propto \exp\left(-\frac{1}{2}\sum_{i=1}^{n}(x_i - \theta)^2\right) \cdot \exp\theta$$

$$\propto \exp\left(-\frac{n}{2}\theta^2 + n\bar{x}\theta + \theta\right) = \exp\left(-\frac{n}{2}\theta^2 + (n\bar{x}+1)\theta\right) \propto \exp\left(-\frac{n}{2}\left(\theta - \frac{n\bar{x}+1}{n}\right)^2\right)$$

$$\Rightarrow \quad \hat{\theta}_{Bayes} = E[\theta \mid \vec{x}] = \bar{x} + \frac{1}{n}$$

$$R(\theta, \hat{\theta}_{Bayes}) = E\left[(\theta - \hat{\theta}_{Bayes})^2\right] = E\left[(\theta - \bar{x} - \tfrac{1}{n})^2\right]$$

$$= E[(\theta - \bar{x})]^2 + \frac{1}{n^2} - \frac{2}{n} E_X(\theta - \bar{x})$$

$$= E[(\theta - \bar{x})]^2 + \frac{1}{n^2} > E[(\theta - \bar{x})]^2 = R(\theta, \bar{x})$$

Thus this estimator is not admissible

$$\sup R(p, \hat{\theta}) = \frac{1}{n^2} + \frac{1}{n}$$

But since $\pi(\theta) = \exp(\theta)$ is improper, $\int_{-\infty}^{\infty} \exp(\theta)\, d\theta = \infty$

the lower bound $E[R(P, \hat{\theta})]$ will be ∞. Thus $\hat{\theta} = \bar{X} + \frac{1}{n}$

is also not minimax

$\square$

dropped?

8.3 Here, the LRT alluded to in Example 8.2.9 will be derived. Suppose that we observe m iid Bernoulli(θ) random variables, denoted by $Y_1, \ldots, Y_m$. Show that the LRT of $H_0: \theta \le \theta_0$ versus $H_1: \theta > \theta_0$ will reject H_0 if $\sum_{i=1}^m Y_i > b$.

8.4 Prove the assertion made in the text after Definition 8.2.1. If $f(x|\theta)$ is the pmf of a dis-

Calculating the likelihood ratio gives $\Lambda(\vec{y}) = \dfrac{\sup_{\theta \le \theta_0} f_m(\vec{y}|\theta)}{\sup f_m(\vec{y}|\theta)}$

$$= \frac{\sup_{\theta\le\theta_0} \theta^{\sum y_i}(1-\theta)^{m-\sum y_i}}{\sup \theta^{\sum y_i}(1-\theta)^{m-\sum y_i}} = \begin{cases} 1, & \text{if } \frac{\sum y_i}{m} \le \theta_0 \\ \dfrac{(\theta_0)^{\sum y_i} \cdot (1-\theta_0)^{m-\sum y_i}}{\left(\frac{\sum y_i}{m}\right)^{\sum y_i} \cdot \left(1-\frac{\sum y_i}{m}\right)^{m-\sum y_i}} & \text{if } \frac{\sum y_i}{m} > \theta_0 \end{cases}$$

We reject H_0 if $\left(\frac{m\theta_0}{\sum y_i}\right)^{\sum y_i} \cdot \left(\frac{m(1-\theta_0)}{m-\sum y_i}\right)^{m-\sum y_i} < c$

and we know $\frac{m\theta_0}{\sum y_i} < 1$, $\frac{m(1-\theta_0)}{m-\sum y_i} > 1$ at this time

So when $\sum y_i$ increases $\left(\frac{m\theta_0}{\sum y_i}\right)^{\sum y_i}$ decreases

and $\left(\frac{m(1-\theta_0)}{m-\sum y_i}\right)^{m-\sum y_i}$ decreases as well.

Thus, there exists b s.t when $\sum_{i=1}^m Y_i > b$,

$\left(\frac{m\theta_0}{\sum y_i}\right)^{\sum y_i} \cdot \left(\frac{m(1-\theta_0)}{m-\sum y_i}\right)^{m-\sum y_i} < c$, i.e. we reject H_0 if $\sum_{i=1}^m Y_i > b$

$\square$

8.5 A random sample, $X_1, \ldots, X_n$, is drawn from a Pareto population with pdf

$$f(x|\theta,\nu) = \frac{\theta\nu^\theta}{x^{\theta+1}} I_{[\nu,\infty)}(x), \quad \theta > 0, \quad \nu > 0.$$

(a) Find the MLEs of θ and ν.

(b) Show that the LRT of

$$H_0: \theta = 1,\ \nu \text{ unknown}, \qquad \text{versus} \qquad H_1: \theta \neq 1,\ \nu \text{ unknown},$$

has critical region of the form $\{\mathbf{x}: T(\mathbf{x}) \le c_1 \text{ or } T(\mathbf{x}) \ge c_2\}$, where $0 < c_1 < c_2$ and

$$T = \log\left[\frac{\prod_{i=1}^n X_i}{(\min_i X_i)^n}\right].$$

(c) Show that, under H_0, $2T$ has a chi squared distribution, and find the number of degrees of freedom. (*Hint:* Obtain the joint distribution of the $n-1$ nontrivial terms $X_i/(\min_i X_i)$ conditional on $\min_i X_i$. Put these $n-1$ terms together, and notice that the distribution of T given $\min_i X_i$ does not depend on $\min_i X_i$, so it is the unconditional distribution of T.)

(a). $L(\theta,\nu|\vec{x}) = \frac{\theta^n \nu^{n\theta}}{(\prod x_i)^{\theta+1}} I_{[\nu,\infty)}(x_{(1)})$

$l(\theta,\nu|\vec{x}) = n\log\theta + n\theta\log\nu - (\theta+1)(\sum_i \log x_i)$ $\quad \nu \le x_{(1)}$

As this is an increasing function of ν, so $\hat{\nu}_{MLE} = x_{(1)}$

Then $\frac{\partial l}{\partial\theta} = \frac{n}{\theta} + n\log\nu - \sum_i \log x_i \overset{set}{=} 0$, $\Rightarrow \hat\theta_{PLE} = \frac{n}{\log\left(\frac{\prod X_i}{(X_{(1)})^n}\right)}$

And $\frac{\partial^2 l}{\partial\theta^2} = \frac{-n}{\theta^2} < 0$, so $\hat\theta_{MLE} = \hat\theta_{PLE} = \frac{n}{\log\left(\frac{\prod X_i}{(X_{(1)})^n}\right)} = \frac{n}{T}$

(b). Calculating the likelihood ratio get:

$$\Lambda(\vec x) = \frac{\sup_{\theta=1} L(\theta,\nu|\vec x)}{\sup L(\theta,\nu|\vec x)} = \frac{x_{(1)}^n / (\prod x_i)^2}{(\frac{n}{T})^n \cdot (x_{(1)})^{\frac{n^2}{T}} / (\prod x_i)^{\frac{n}{T}+1}}$$

$$= \left(\frac{T}{n}\right)^n x_{(1)}^{n-\frac{n^2}{T}} \cdot \left(\prod x_i\right)^{\frac{n}{T}-1} = \left(\frac{T}{n}\right)^n \cdot \left(\prod x_i/(x_{(1)})^n\right)^{\frac{n}{T}-1}$$

$$= \left(\frac{T}{n}\right)^n \cdot (e^T)^{\frac{n}{T}-1} = \left(\frac{T}{n}\right)^n \cdot e^{n-T}, \quad \frac{\partial}{\partial T}\log\Lambda(\vec x) = \frac{n}{T} - 1, \text{ hence}$$

$\Lambda(x)$ is increasing if $T \le n$ and decreasing if $T > n$

Thus, $\Lambda(\vec{X}) \leq c$ is equivalent to $T \leq c_1$ or $T \geq c_2$ for chosen c_1 and c_2, i.e the LRT has the critical region $\{X : T(X) \leq c_1 \text{ or } T(X) \geq c_2\}$

(c). $T = \log\left(\frac{X_1 \cdots X_n}{X_{(1)}^n}\right) = \sum \log X_i - n \log X_{(1)}$, Let $Y_i = \log X_i$

$\Rightarrow T = \sum Y_i - n Y_{(1)} = (Y_1 - Y_{(1)}) + \cdots + (Y_n - Y_{(1)})$

Consider $Z_i = Y_i - Y_{(1)}$ and suppose $Y_i \neq Y_{(1)}$

$$f_X(x|\theta,\nu) = \frac{\theta \nu^\theta}{x^{\theta+1}} \quad (x \geq \nu), \quad f_{Y_i}(y_i|\theta,\nu) = \left|\frac{\partial e^{y_i}}{\partial y_i}\right| \cdot f_X(e^{y_i}|\theta,\nu)$$

$$= \theta \cdot e^{-\theta y_i} \nu^\theta$$

$$f_{Y_{(1)}}(y_{(1)}|\theta,\nu) = n\nu^n \cdot e^{-n y_{(1)}} \quad \Rightarrow Z_i \sim \exp(1)$$

$$\Rightarrow \sum Z_i \sim \Gamma(n-1, 1), \text{ i.e } \sum \log X_i - n \log X_{(1)} \sim \Gamma(n-1, 1)$$

which means $T \sim \Gamma(n-1, 1) \Rightarrow 2T \sim \Gamma(n-1, \frac{1}{2}) = \chi^2(2n-2)$

$\square$

8.8 A special case of a normal family is one in which the mean and the variance are related, the $n(\theta, a\theta)$ family. If we are interested in testing this relationship, regardless of the value of θ, we are again faced with a nuisance parameter problem.

(a) Find the LRT of $H_0: a = 1$ versus $H_1: a \neq 1$ based on a sample $X_1, \ldots, X_n$ from a $n(\theta, a\theta)$ family, where θ is unknown.

(b) A similar question can be asked about a related family, the $n(\theta, a\theta^2)$ family. Thus, if $X_1, \ldots, X_n$ are iid $n(\theta, a\theta^2)$, where θ is unknown, find the LRT of $H_0: a = 1$ versus $H_1: a \neq 1$.

(a), We first find the MLE of a, θ,

$$L(a,\theta|\vec{x}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi a\theta}} \cdot e^{-(x_i-\theta)^2/(2a\theta)}$$

$$l(a,\theta|\vec{x}) = -\frac{n}{2}\log(2\pi a\theta) - \sum_{i=1}^{n}(x_i-\theta)^2/(2a\theta)$$

$$\begin{cases} \frac{\partial l}{\partial a} = 0 \\ \frac{\partial l}{\partial \theta} = 0 \end{cases} \Rightarrow \quad \hat{a}_{MLE} = \frac{1}{n\bar{x}}\sum_{i=1}^{n}(x_i-\bar{x})^2 = \frac{\sigma^2}{\bar{x}}, \quad \hat{\theta}_{MLE} = \bar{x}$$

And when $a=1$, $\hat{\theta}_R = -\frac{1}{2} + \frac{1}{2}\sqrt{1+4(\sigma^2+\bar{x}^2)}$

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

Now. Calculating the likelihood ratio:

$$\Lambda(\vec{x}) = \frac{\sup_{a=1} L(a,\theta|\vec{x})}{\sup L(a,\theta|\vec{x})} = \hat{a}^{\frac{n}{2}} \frac{\exp(-\sum(x_i-\hat{\theta}_R)^2/2\hat{\theta}_R)}{\exp(-\sum(x_i-\hat{\theta}_{MLE})^2/(2\hat{a}_{MLE}\cdot\hat{\theta}_{MLE}))}$$

$$= \left(\frac{\sigma^2}{\bar{x}}\right)^{\frac{n}{2}} \cdot \exp\left(-\sum(x_i-\hat{\theta}_R)^2/2\hat{\theta}_R + \frac{n}{2}\right)$$

(b). Similar we get $\hat{a}_{MLE} = \frac{\sigma^2}{\bar{x}^2}$, $\hat{\theta}_{MLE} = \bar{x}$

$$\hat{\theta}_R = \bar{x} + \sqrt{\bar{x}+4(\sigma^2+\bar{x})^2}/2$$

the LRT is $\left(\frac{\sigma}{\hat{\theta}_R}\right)^n \exp\left(\frac{n}{2} - \sum(x_i-\hat{\theta}_R)^2/(2\hat{\theta}_R)\right)$

$\square$

8.10 Let $X_1, \ldots, X_n$ be iid Poisson(λ), and let λ have a gamma(α, β) distribution, the conjugate family for the Poisson. In Exercise 7.24 the posterior distribution of λ was found, including the posterior mean and variance. Now consider a Bayesian test of $H_0: \lambda \le \lambda_0$ versus $H_1: \lambda > \lambda_0$.

(a) Calculate expressions for the posterior probabilities of H_0 and H_1.

(b) If $\alpha = \frac{5}{2}$ and $\beta = 2$, the prior distribution is a chi squared distribution with 5 degrees of freedom. Explain how a chi squared table could be used to perform a Bayesian test.

(a). $X_1, \cdots, X_n$ i.i.d Poisson, $\lambda \sim$ gamma(α, β) Thus from previous exercise, we get $\lambda | \vec{x} \sim$ gamma$(\sum X_i + \alpha, n+\beta)$

$$\hat{\lambda}_{Bayes} = \frac{\sum X_i + \alpha}{n+\beta}$$

H_0: $\sum X_i \le (n+\beta)\lambda_0 - \alpha$

H_1: $\sum X_i > (n+\beta)\lambda_0 - \alpha$

(b). $\lambda | \vec{x} \sim \Gamma(n\bar{x} + \frac{5}{2}, 2+n) \Rightarrow (2n+4)\lambda | \vec{x} \sim \Gamma(n\bar{x} + \frac{5}{2}, \frac{1}{2}) : \chi^2(2n\bar{x}+5)$

Thus $P(\lambda \le \lambda_0 | \vec{x}) = P((2n+4)\lambda \le (2n+4)\lambda_0 | \vec{x}) = P(\chi^2 \le 2(n+4)\lambda_0)$

$\chi^2 \sim \chi^2(2n\bar{x}+5)$

$\square$