Review Session 7 – Solutions

1 Solutions

1.1 Previous Core Competency Problems

Problem 1 (September 2018, # 5)**.** We obtain observations $Y_1, \ldots, Y_n$ which can be described by the relationship

$$Y_i = i \times \theta + \epsilon_i,$$

where $\epsilon_1, \ldots, \epsilon_n$ are i.i.d. $N(0, \sigma^2)$; $\sigma^2 > 0$. Here θ and σ^2 are unknown.

(i) Find the least squares estimator $\hat{\theta}$ of θ; i.e., $\hat{\theta} = \operatorname{argmin}_{\theta \in \mathbb{R}} \sum_{i=1}^n (Y_i - i\theta)^2$.

(ii) Is $\hat{\theta}$ unbiased?

(iii) Find the exact distribution of $\hat{\theta}$.

(iv) Find the asymptotic (non-degenerate) distribution of $\hat{\theta}$ (properly normalized).

(v) How would you test the hypothesis $H_0 : \theta = 0$ versus $H_1 : \theta \neq 0$ (at level $(\alpha \in (0,1))$)? Describe the test statistic and the critical value.

Solution

(i) Taking the derivative of the objective with respect to θ, it's straightforward to compute $\hat{\theta} = \frac{\sum_{i=1}^n i \cdot Y_i}{\sum_{i=1}^n i^2}$.

(ii) $\hat{\theta}$ is unbiased by linearity of expectation.

(iii) $\hat{\theta} \sim \mathcal{N}(\theta, \sigma^2 / \sum_{i=1}^n i^2)$.

(iv) From (iii), we have $\sqrt{\sum_{i=1}^n i^2} \cdot (\hat{\theta} - \theta) \xrightarrow{\text{d}} \mathcal{N}(0, \sigma^2)$.

(v) To find a T-statistic for this problem (from which we can derive a level α t-test), it suffices to find an appropriate estimator of the variance σ^2. We'll consider a variant of the sample variance which takes into account the different means of the Y_i's. Let

$$S^2 := \frac{1}{n-1} \sum_{i=1}^n (Y_i - i\hat{\theta})^2.$$

We can then observe $\forall i \in [n] : \operatorname{Cov}(Y_i - i\hat{\theta}, \hat{\theta}) = 0$. Thus, S^2 and $\hat{\theta}$ are independent and next we claim

$$(n-1) \cdot S^2 \sim \sigma^2 \cdot \chi^2_{n-1}.$$

This follows from standard facts about normal linear regression where $X_i := i$ is our covariate. For completeness, I give a full proof of the above claim. First, for ease of notation, define $\mathbf{Y} = (Y_1, \ldots, Y_n)^T$, $\mathbf{X} = (1, 2, \ldots, n)^T$, and $\epsilon = (\epsilon_1, \ldots, \epsilon_n)^T$. Then, we have

$$\begin{aligned}
\hat{\theta} &= (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y} \\
&= (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T(\mathbf{X}\theta + \epsilon) \\
&= \theta + (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\epsilon. \qquad (1)
\end{aligned}$$

Thus,

$$\begin{aligned}(n-1)\cdot S^2 &= (\mathbf{Y}-\mathbf{X}\hat{\theta})^T(\mathbf{Y}-\mathbf{X}\theta)\\ &= (\mathbf{X}(\theta-\hat{\theta})+\epsilon)^T(\mathbf{X}(\theta-\hat{\theta})+\epsilon)\\ &= (-\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\epsilon+\epsilon)^T(-\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\epsilon+\epsilon) \text{ (using (1))}\\ &= \epsilon^T(\mathrm{Id}-\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T)^T(\mathrm{Id}-\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T)\epsilon.\end{aligned}$$

We have $\epsilon/\sigma \sim \mathcal{N}(\mathbf{0}_n, \mathrm{Id})$. Then, we claim that the quadratic form $(\epsilon/\sigma)^T A(\epsilon/\sigma)$ with matrix $A = (\mathrm{Id}-\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T)^T(\mathrm{Id}-\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T)$ is a χ^2_{n-1} distribution. This follows from verifying A is symmetric, idempotent, has trace $n-1$, and then using Problem 10 in the 2018 Summer Practice exam. Thus, our claim $(n-1)\cdot S^2 \sim \sigma^2\cdot\chi^2_{n-1}$ is proven. Then, it follows tha

$$\frac{\hat{\theta}-\theta}{S/\sqrt{\sum_{i=1}^n i^2}} \sim T_{n-1}$$

from which it is straightforward to derive a level α hypothesis test.

Problem 2 (2019 May, # 1)**.** Let $X \sim$ Poisson(λ) and $Y \sim$ Poisson(μ), where $\lambda, \mu > 0$ and assume that X and Y are independent.

(i) Find the conditional distribution of X given that $X + Y = n$.

(ii) Use the above, or otherwise, to test the hypothesis (at level $\alpha \in (0,1)$)

$$H_0 : \lambda = \mu \qquad \text{versus} \qquad \lambda > \mu.$$

Solution

(i) A quick mgf calculation reveals $X + Y \sim$ Poisson$(\lambda+\mu)$. Next, we have for $k \in \mathbb{N}\cap[0,n]$:

$$\mathbb{P}(X=k|X+Y=n) = \frac{\mathbb{P}(X=k, Y=n-k)}{\mathbb{P}(X+Y=n)} = \frac{\left(\frac{e^{-\lambda}}{k!}\lambda^k\right)\left(\frac{e^{-\mu}}{(n-k)!}\mu^{n-k}\right)}{\frac{e^{-(\lambda+\mu)}(\lambda+\mu)^n}{n!}} = \binom{n}{k}\frac{\lambda^k\mu^{n-k}}{(\lambda+\mu)^n}.$$

We can further factor the above RHS as $\binom{n}{k}\left(\frac{\lambda}{\lambda+\mu}\right)^k\left(\frac{\mu}{\lambda+\mu}\right)^{n-k}$. Thus, $X|X+Y=n \sim$ Binomial$(n, \lambda/(\lambda+\mu))$.

(ii) We can then form a level α test with rejection region $\{X > c_\alpha\}$ where c_α is chosen such that for $X+Y=n$:

$$\sum_{k=c_\alpha}^{n}\binom{n}{k}\cdot\frac{1}{2^n} \leq \alpha.$$

Problem 3 (2020 May, # 5)**.** Suppose that $X_1, \ldots, X_n$ are i.i.d. observations from the exponential distribution with parameter λ (recall that $\mathbb{E}(X_1) = \lambda^{-1}$). Consider the following testing problem:

$$H_0 : \lambda = \lambda_0 \qquad \text{versus} \qquad H_1 : \lambda = \lambda_1,$$

where $0 < \lambda_1 < \lambda_0$. Let $f_0(X_1, \ldots, X_n)$ be the likelihood of the data under H_0 and $f_1(X_1, \ldots, X_n)$ that under H_1.

(i) Show that $\log\frac{f_1(X_1,\ldots,X_n)}{f_0(X_1,\ldots,X_n)}$ is an increasing function of $\overline{X}_n := \frac{1}{n}\sum_{i=1}^n X_i$.

(ii) Suppose that $c_{\alpha,n}$ is such that $\mathbb{P}_{\lambda_0}(\overline{X}_n \geq c_{\alpha,n}) = \alpha$, for $\alpha \in (0,1)$. Relate $c_{\alpha,n}$ to $q_k(\beta)$ – the β-th quantile of the χ^2_k distribution (for some k).

(iii) How would you test the hypothesis

$$H_0 : \lambda = \lambda_0 \qquad \text{versus} \qquad H_1 : \lambda < \lambda_1,$$

Derive an expression for the power function of the test.

Solution

(i)

$$\log \frac{f_1(X_1, \ldots, X_n)}{f_0(X_1, \ldots, X_n)} = n \log(\lambda_1/\lambda_0) - (\lambda_1 - \lambda_0) \sum_{i=1}^{n} X_i,$$

which is increasing in $\overline{X}_n$.

(ii) An mgf computation reveals $\exp(\lambda) \sim \frac{1}{2\lambda}\chi_2^2$ and thus an i.i.d. sum of $\exp(\lambda)$ R.V.'s is $\frac{1}{2\lambda}\chi_{2n}^2$. Thus,

$$\mathbb{P}_{\lambda_0}(\overline{X}_n < c_{\alpha,n}) = 1 - \alpha = \mathbb{P}_{\lambda_0}(\chi_{2n}^2 < 2\lambda_0 n \cdot c_{\alpha,n}).$$

Thus, $q_{2n}(1-\alpha) = 2\lambda_0 n \cdot c_{\alpha,n}$.

(iii) A natural level α test then has rejection region $\{\overline{X}_n \geq c_{\alpha,n}\}$ with power

$$\mathbb{P}_{\lambda}(\overline{X}_n \geq c_{\alpha,n}) = \mathbb{P}(\chi_{2n}^2 \geq 2\lambda n \cdot c_{\alpha,n}).$$

Problem 4 (2020 May, # 7)**.** Consider the random variable $X = \mu + \sigma Z$, where $\mu \in \mathbb{R}$, $\sigma > 0$ and Z is a random variable with density f. Suppose that μ and σ are unknown parameters and that the density f is known (completely specified). We have a random i.i.d sample $X_1, \ldots, X_n$ with the same distribution as X. *You may assume for this problem that* $\mathbb{E}[Z] = 0$*,* $\mathbb{E}[|Z|] < \infty$*, and* $\operatorname{Var}(Z) \in (0, \infty)$.

(i) Propose unbiased estimators, $\hat{\mu}$ and $\hat{\sigma}^2$, of μ and σ^2.

(ii) Does the joint distribution of $(X_i - \hat{\mu})/\hat{\sigma}$ ($i = 1, \ldots, n$) depend on μ and σ? Explain your answer.

(iii) For a given level $\alpha \in (0, 1)$, describe a way to construct a confidence interval for μ with exact coverage probability $1 - \alpha$.

Note: I added some extra assumptions to this problem (in *italics*) since I don't think the problem is solvable without them in general.

Solution

(i) Let $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$. Then, the estimators $\hat{\mu} := \overline{X}_n$ and $\hat{\sigma}^2 := S^2/\operatorname{Var}(Z)$ are unbiased estimators of μ and σ^2, respectively.

> **Remark.** Here are my justifications for the extra assumptions I added.
>
> First, if $\operatorname{Var}(Z) = 0$, then Z is a.s. constant which means we cannot form an unbiased estimator of either μ or σ since we only observe $\mu + \sigma \cdot Z$.
>
> If $\mathbb{E}[Z] \neq 0$, then I claim no unbiased estimator of the location μ can exist. For if $\hat{\mu}$ was an unbiased estimator of μ, then $\frac{\overline{X}_n - \hat{\mu}}{\mathbb{E}[Z]}$ would be an unbiased estimator of σ. However, an unbiased estimator of σ cannot exist in general (cf. Theorem 2.3 of here).
>
> If $\mathbb{E}[|Z|], \operatorname{Var}(Z)$ are undefined, then unbiased estimators need not exist for either parameter (e.g., for a Cauchy distribution; see Proposition 4 of the same paper). I think the problem-writers meant to impose more conditions here.

(ii) Let $X_i = \mu + \sigma \cdot Z_i$ for $Z_i \overset{\text{i.i.d.}}{\sim} f$ and define $\overline{Z}_n = n^{-1}\sum_{i=1}^n Z_i$. Then, we have

$$\frac{X_i - \hat{\mu}}{\hat{\sigma}} = \frac{X_i - \overline{X}_n}{\sqrt{\hat{\sigma}^2}} = \frac{Z_i - \overline{Z}_n}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n (Z_i - \overline{Z}_n)^2}}.$$

Thus, the distribution of $\frac{X_i - \hat{\mu}}{\hat{\sigma}}$ only depends on the distribution of Z and not on μ, σ.

(iii) We claim the distribution of $T := \frac{\hat{\mu} - \mu}{\sqrt{\hat{\sigma}^2}}$ is free of μ, σ^2. Let $S_Z^2 := \frac{1}{n-1}\sum_{i=1}^n (Z_i - \overline{Z}_n)^2$. Then,

$$\overline{X}_n = \mu + \sigma \cdot \overline{Z}_n$$
$$S^2 = \sigma^2 \cdot S_Z^2.$$

So, we have

$$T := \frac{\hat{\mu} - \mu}{\sqrt{\hat{\sigma}^2}} = \frac{\sigma \cdot \overline{Z}_n}{\sigma\sqrt{S_Z^2 / \operatorname{Var}(Z)}} = \frac{\overline{Z}_n}{\sqrt{S_Z^2 / \operatorname{Var}(Z)}}$$

Thus, the distribution of T only depends on Z and is free of μ, σ. Then, if $T_{1-\alpha/2}, T_{\alpha/2}$ are the corresponding quantiles of T, we have $1-\alpha$ C.I.:

$$\hat{\mu} - T_{\alpha/2} \cdot \hat{\sigma} \le \mu \le \hat{\mu} + T_{1-\alpha/2} \cdot \hat{\sigma}.$$

Problem 5 (2020 September, # 1). Researchers notice that a mutation in a gene predisposes individuals to a kind of radiation-induced cancer. The researchers theorize that the gene is involved in repairing damage from radiation, and that the mutation disables the gene. To explore their theory, the researchers obtain cells growing in a laboratory that have the mutation. They take eight different clumps of the cells, and randomize the clumps to treatment with radiation or no radiation (four in each group). They then examine a marker of damage from radiation in each cell in each clump, recording whether or not there appears to be damage. The researchers run the same experiment in clumps of cells that do not have the mutation. They explain that the cells that do not have the mutation are a “control”. The researchers ask you to analyze the results.

(a) Propose a “reasonable” model to analyze the data.

(b) Propose how you plan to conclude whether mutation plays a role in repairing radiation damage.

Solution

Let $M := \mathbf{1}\{\text{clump is mutated}\}$ and let $R := \mathbf{1}\{\text{clump is irradiated}\}$. Then, our response vaiable is $Y := \mathbf{1}\{\text{clump exhibits damage}\}$, so we can consider a logistic regression model

$$\log\left(\frac{\mathbb{E}[Y|M,R]}{1-\mathbb{E}[Y|M,R]}\right) = \beta_0 + \beta_1 \cdot M + \beta_2 \cdot R,$$

and conduct a hypothesis test of significance for $H_0 : \beta_1 = 0$ vs. $H_1 : \beta_1 \neq 0$. This can be done, for example, using Wald’s test which treats $\hat{\beta}_1/\text{SE}(\hat{\beta}_1)$ as approximately $\mathcal{N}(0,1)$, where $\hat{\beta}_1$ is the MLE of β_1.

Problem 6 (2020 September, # 3). Suppose that $(X_{1i}, X_{2i}) \overset{i.i.d.}{\sim} N_2(\theta, I_2)$ for $1 \le i \le n$, where the parameter space is restricted to $\Theta := \{\theta = (\theta_1, \theta_2) : \theta_1, \theta_2 \ge 0\}$. Consider the following hypothesis testing problem:

$$H_0 : \theta = (0,0) \qquad \text{versus} \qquad H_1 : \theta \in \Theta \backslash \{(0,0)\}.$$

(a) Find the MLE of θ (when $\theta \in \Theta$).

(b) Find an expression for the likelihood ratio statistic $\Lambda_n \in (0,1]$ in this case.

(c) Find the asymptotic distribution of $-2\log\Lambda_n$, under H_0 [Hint: You may want to consider the cases where $(\overline{X}_1, \overline{X}_2)$ belongs to each of the four quadrants separately.]

Solution

(a) The MLE for θ_i is $\overline{X}_i \cdot \mathbf{1}\{\overline{X}_i \ge 0\}$ for $i = 1, 2$ where $\overline{X}_i = n^{-1}\sum_{j=1}^n X_{ij}$.

(b) The joint likelihood is proportional to

$$L(\theta_1, \theta_2) = \exp\left(-\frac{n(\overline{X}_1 - \theta_1)^2}{2} - \frac{n(\overline{X}_2 - \theta_2)^2}{2}\right),$$

which is maximized at the MLE or

$$\sup_{(\theta_1,\theta_2)\in\Theta} L(\theta_1, \theta_2) = \exp\left(-\frac{n}{2}\left(\overline{X}_1^2 \cdot \mathbf{1}\{\overline{X}_1 < 0\} + \overline{X}_2^2 \cdot \mathbf{1}\{\overline{X}_2 < 0\}\right)\right).$$

On the other hand,

$$L(0,0) = \exp\left(-\frac{n}{2}\left(\overline{X}_1^2 + \overline{X}_2^2\right)\right).$$

Thus,

$$\Lambda_n = \frac{L(0,0)}{\sup_{(\theta_1,\theta_2)\in\Theta} L(\theta_1,\theta_2)} = \exp\left(-\frac{n}{2}\left(\overline{X}_1^2 \cdot \mathbf{1}\{\overline{X}_1 \geq 0\} + \overline{X}_2^2 \cdot \mathbf{1}\{\overline{X}_2 \geq 0\}\right)\right).$$

(c)

$$-2\log\Lambda_n = n\left(\overline{X}_1^2 \cdot \mathbf{1}\{\overline{X}_1 \geq 0\} + \overline{X}_2^2 \cdot \mathbf{1}\{\overline{X}_2 \geq 0\}\right).$$

The function $g(x) = x^2 \cdot \mathbf{1}\{x \geq 0\}$ is continuous. Thus, since $\sqrt{n} \cdot \overline{X}_i \xrightarrow{\text{d}} \mathcal{N}(0,1)$ under H_0 by CLT, the continuous mapping theorem gives us:

$$-2\log\Lambda_n \xrightarrow{\text{d}} Z_1^2 \cdot \mathbf{1}\{Z_1 \geq 0\} + Z_2^2 \cdot \mathbf{1}\{Z_2 \geq 0\},$$

where $Z_1, Z_2 \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$.

Problem 7 (2021 May, # 2)**.** Let $X_1, X_2, \ldots, X_n$ be from an i.i.d. random sample Uniform$(0,\theta)$, where $\theta > 0$ is an unknown parameter. Suppose that we want to test the following hypothesis:

$$H_0 : 3 \leq \theta \leq 4 \qquad \text{versus} \qquad H_1 : \theta < 3 \text{ or } \theta > 4. \tag{2}$$

Let $Y_n = \max\{X_1, \ldots, X_n\}$. Consider the following two tests:

$$\delta_1 : \text{Reject } H_0 \text{ if } Y_n \leq 2.9 \text{ or } Y_n \geq 4$$

and

$$\delta_2 : \text{Reject } H_0 \text{ if } Y_n \leq 2.9 \text{ or } Y_n \geq 4.5.$$

(i) Find the power functions of δ_1 and δ_2, when $\theta \leq 4$

(ii) Find the power functions of δ_1 and δ_2, when $\theta > 4$.

(iii) Which of the two tests seems better for testing the hypothesis (2)?

Solution

The power functions of δ_1 and δ_2 are, respectively

$$\mathbb{P}(Y_n \leq 2.9 \cup Y_n \geq 4|\theta) = \begin{cases} \left(\frac{2.9}{\theta}\right)^n + 1 - \left(\frac{4}{\theta}\right)^n & \theta > 4 \\ \left(\frac{2.9}{\theta}\right)^n & \theta \in [2.9, 3) \\ 1 & \theta < 2.9 \end{cases}$$

$$\mathbb{P}(Y_n \leq 2.9 \cup Y_n \geq 4.5|\theta) = \begin{cases} \left(\frac{2.9}{\theta}\right)^n + 1 - \left(\frac{4.5}{\theta}\right)^n & \theta > 4.5 \\ \left(\frac{2.9}{\theta}\right)^n & \theta \in [2.9, 3) \cup (4, 4.5] \\ 1 & \theta < 2.9 \end{cases}.$$

From the above, we see that δ_1 has higher power than δ_2 for all values of θ.

1.2 Additional Practice

Problem 8 (Casella & Berger, Exercise 8.8)**.** A special case of a normal family is one in which the mean and variance are related, the $\mathcal{N}(\theta, a\cdot\theta)$ family.

(a) Find the LRT of $H_0 : a = 1$ versus $H_1 : a \neq 0$ based on a sample $X_1, \ldots, X_n \overset{\text{i.i.d}}{\sim} \mathcal{N}(\theta, a\cdot\theta)$, where θ is unknown.

(b) Now consider the $\mathcal{N}(\theta, a\cdot\theta^2)$ family. If $X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\theta, a\cdot\theta^2)$, where θ is unknown, find the LRT of $H_0 : a = 1$ versus $H_1 : a \neq 1$.

Solution

(a) We first compute the unconstrained MLE of (a, θ). We have

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

Thus,

$$\frac{\partial \log(L)}{\partial a} = \sum_{i=1}^{n}\left(-\frac{1}{2a} + \frac{1}{2\theta a^2}(x_i-\theta)^2\right) = -\frac{n}{2a} + \frac{1}{2\theta a^2}\sum_{i=1}^{n}(x_i-\theta)^2$$
$$\frac{\partial \log(L)}{\partial \theta} = \sum_{i=1}^{n} -\frac{1}{2\theta} + \frac{1}{2a\theta^2}(x_i-\theta)^2 + \frac{1}{a\theta}(x_i-\theta)$$

Setting the first equation equal to 0, we see that $a = \frac{1}{n\theta}\sum_{i=1}^{n}(x_i-\theta)^2$. Substituting this into the second equation and setting it equal to 0 then gives us

$$-\frac{n}{2\theta} + \frac{n}{2\theta} + \frac{n(\overline{x}-\theta)}{a\theta} = 0 \implies \hat{\theta}_{\text{MLE}} = \overline{x}_n.$$

Thus, $\hat{a}_{\text{MLE}} = \frac{1}{n\cdot\overline{x}_n}\sum_{i=1}^{n}(x_i-\overline{x}_n)^2 = \frac{\hat{\sigma}^2}{\overline{x}_n}$, where $\hat{\sigma}^2$ is the usual MLE of the variance. For $a = 1$, the restricted MLE is found similarly by setting log-likelihood derivative (which is a quadratic in θ) equal to 0 when $a = 1$:

$$\hat{\theta}_{\text{MLE},a=1} = \frac{-1+\sqrt{1+4(\hat{\sigma}^2+\overline{x}^2)}}{2}.$$

Noting that $\hat{a}_{\text{MLE}}\cdot\hat{\theta}_{\text{MLE}} = \hat{\sigma}^2$, the LRT is found to be

$$\lambda(\mathbf{x}) = \frac{L(\hat{\theta}_{\text{MLE},a=1}|\mathbf{x})}{L(\hat{a}_{\text{MLE}},\hat{\theta}_{\text{MLE}}|\mathbf{x})}$$
$$= \frac{\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\hat{\theta}_{\text{MLE},a=1}}}e^{-(x_i-\hat{\theta}_{\text{MLE},a=1})^2/(2\hat{\theta}_{\text{MLE},a=1})}}{\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\hat{a}_{\text{MLE}}\cdot\hat{\theta}_{\text{MLE}}}}e^{-(x_i-\hat{\theta}_{\text{MLE}})^2/(2\hat{a}_{\text{MLE}}\cdot\hat{\theta}_{\text{MLE}})}}$$
$$= \frac{(1/(2\pi\hat{\theta}_{\text{MLE},a=1}))^{n/2}e^{-\sum_i(x_i-\hat{\theta}_{\text{MLE},a=1})^2/(2\hat{\theta}_{\text{MLE},a=1})}}{1/(2\pi\hat{\sigma}^2)^{n/2}e^{-\sum_i(x_i-\overline{x}_n)^2/(2\hat{\sigma}^2)}}$$
$$= (\hat{\sigma}^2/\hat{\theta}_{\text{MLE},a=1})^{n/2}e^{(n/2)-\sum_i(x_i-\hat{\theta}_{\text{MLE},a=1})^2/(2\hat{\theta}_{\text{MLE},a=1})}.$$

(b) In this case we have

$$\log L(a,\theta|\mathbf{x}) = \sum_{i=1}^{n} -\frac{1}{2}\log(2\pi a\theta^2) - \frac{1}{2a\theta^2}(x_i-\theta)^2.$$

Thus,

$$\frac{\partial \log(L)}{\partial a} = \sum_{i=1}^{n} -\frac{1}{2a} + \frac{1}{2a^2\theta^2}(x_i-\theta)^2 = -\frac{n}{2a} + \frac{1}{2a^2\theta^2}\sum_{i=1}^{n}(x_i-\theta)^2$$
$$\frac{\partial \log(L)}{\partial \theta} = \sum_{i=1}^{n} -\frac{1}{\theta} + \frac{1}{a\theta^3}(x_i-\theta)^2 + \frac{1}{a\theta^2}(x_i-\theta)$$
$$= -\frac{n}{\theta} + \frac{1}{a\theta^3}\sum_{i=1}^{n}(x_i-\theta)^2 + \frac{1}{a\theta^2}\sum_{i=1}^{n}(x_i-\theta).$$

Solving the first equation for a in terms of θ yields

$$a = \frac{1}{n\theta^2}\sum_{i=1}^{n}(x_i-\theta)^2.$$

Substituting this into the second equation, we get

$$-\frac{n}{\theta}+\frac{n}{\theta}+n\frac{\sum_i(x_i-\theta)}{\sum_i(x_i-\theta)^2}=0 \implies \hat{\theta}=\overline{x},$$

so that again $\hat{a}=\frac{\hat{\sigma}^2}{\hat{x}^2}$. In the restricted case, set $a=1$ in the second equation above to get

$$\frac{\partial\log(L)}{\partial\theta} = -\frac{n}{\theta}+\frac{1}{\theta^3}\sum_{i=1}^{n}(x_i-\theta)^2+\frac{1}{\theta^2}\sum_{i=1}^{n}(x_i-\theta)=0 \implies -\theta^2+\hat{\sigma}^2+(\overline{x}_n-\theta)^2+\theta(\overline{x}_n-\theta)=0.$$

This is a quadratic in θ with solution for the MLE

$$\hat{\theta}_{\text{MLE},a=1}=\overline{x}_n+\sqrt{\overline{x}_n+4(\hat{\sigma}^2+\overline{x}_n^2)}/2,$$

which yields the LRT statistic

$$\lambda(\mathbf{x})=\frac{L(\hat{\theta}_{\text{MLE},a=1}|\mathbf{x})}{L(\hat{a}_{\text{MLE}},\hat{\theta}_{\text{MLE}}|\mathbf{x})}=\left(\frac{\hat{\sigma}}{\hat{\theta}_{\text{MLE},a=1}}\right)^n e^{(n/2)-\sum_i(x_i-\hat{\theta}_{\text{MLE},a=1})^2/(2\hat{\theta}_{\text{MLE},a=1})}.$$

Problem 9 (Casella & Berger, Exercise 8.13). Let X_1, X_2 be i.i.d. Unif$[(\theta,\theta+1)]$. For testing $H_0:\theta=0$ versus $H_1:\theta>0$, consider two competing tests

$$\phi_1(X_1): \text{Reject } H_0 \text{ if } X_1>.95$$
$$\phi_2(X_1,X_2): \text{Reject } H_0 \text{ if } X_1+X_2>C.$$

(i) Find the value of C so that ϕ_2 has the same size (i.e., probability of committing a Type 1 error) as ϕ_1.

(ii) Calculate the power function of each test.

(iii) Which test ϕ_1 or ϕ_2 is more powerful (i.e., has larger power function)? Does it depend on the value of θ?

Solution

(i) The size of ϕ_1 is $\mathbb{P}(X_1>.95|\theta=0)=.05$. The size of ϕ_2 is

$$\mathbb{P}(X_1+X_2>C|\theta=0)=\int_{1-C}^{1}\int_{C-x_1}^{1}1\,dx_2\,dx_1=\frac{(2-C)^2}{2}.$$

Setting this equal to $.05$ and solving for C gives $C=2-\sqrt{.1}\approx 1.68$.

(ii) ϕ_1 has power function

$$\beta_1(\theta)=\mathbb{P}_\theta(X_1>.95)=\begin{cases}0 & \theta\le -.05\\ \theta+.05 & \theta\in(-.05,.95]\\ 1 & .95<\theta\end{cases}.$$

Using the distribution of $Y=X_1+X_2$, given by

$$f_Y(y|\theta)=\begin{cases}y-2\theta & 2\theta\le y<2\theta+1\\ 2\theta+2-y & 2\theta+1\le y<2\theta+2\\ 0 & \text{otherwise},\end{cases}$$

we obtain the power function for ϕ_2 as

$$\beta_2(\theta) = \mathbb{P}_\theta(Y > C) = \begin{cases} 0 & \theta \leq (C/2) - 1 \\ (2\theta + 2 - C)^2/2 & (C/2) - 1 < \theta \leq (C-1)/2 \\ 1 - (C - 2\theta)^2/2 & (C-1)/2 < \theta \leq C/2 \\ 1 & C/2 < \theta \end{cases}.$$

From looking at the graphs of the two power functions, we see that ϕ_1 is more powerful for θ near 0, but ϕ_2 is more powerful for larger θ's. Neither test is uniformly more powerful than the other.