Statistical Theory - Solutions 5

Spring 2023. Email corrections to Kolyan Ray: kolyan.ray@ic.ac.uk

1. By the Neyman-Pearson lemma, this is given by the likelihood ratio test. The likelihood ratio equals

$$\Lambda(x) = \frac{f_{\sigma_1}(x)}{f_{\sigma_0}(x)} = \frac{\prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-x_i^2/(2\sigma_1^2)}}{\prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-x_i^2/(2\sigma_0^2)}} = (\sigma_0^2/\sigma_1^2)^{n/2} \exp\left(\left(\frac{1}{2\sigma_0^2} - \frac{1}{2\sigma_1^2}\right)\sum_{i=1}^n x_i^2\right),$$

and we reject H_0 if and only if $\Lambda(x) \geq k$, where $P_0(\Lambda(X) \geq k) = \alpha$. Rearranging and using that $\sigma_1^2 > \sigma_0^2$,

$$\begin{aligned}\{\Lambda(x) \geq k\} &= \left\{\exp\left(\frac{\sigma_1^2 - \sigma_0^2}{2\sigma_0^2\sigma_1^2}\sum_{i=1}^n x_i^2\right) \geq \left(\frac{\sigma_1^2}{\sigma_0^2}\right)^{n/2} k\right\} \\ &= \left\{\sum_{i=1}^n x_i^2 \geq \frac{2\sigma_0^2\sigma_1^2}{\sigma_1^2 - \sigma_0^2}\left(\frac{n}{2}\log(\sigma_1^2/\sigma_0^2) + \log k\right)\right\}.\end{aligned}$$

But under H_0, $\frac{1}{\sigma_0^2}\sum_{i=1}^n X_i^2 \sim \chi_n^2$, so we know $P_0(\sum_i X_i^2 \geq \sigma_0^2 \chi_n^2(\alpha)) = \alpha$ for $\chi_n^2(\alpha)$ the upper α-tail of the χ_n^2 distribution, i.e. satisfying $P(\chi_n^2 \geq \chi_n^2(\alpha)) = \alpha$. Thus we reject H_0 if and only if $\sum_i x_i^2 \geq \sigma_0^2 \chi_n^2(\alpha)$. [You can also rearrange to find k in terms of $\chi_n^2(\alpha)$].

2. We have likelihood ratio

$$\Lambda(x) = \frac{f_2(x)}{f_1(x)} = \frac{2}{(x+2)^2}\frac{(x+1)^2}{1} = 2\left(1 - \frac{1}{x+2}\right)^2,$$

and the LRT rejects H_0 if and only if $\Lambda(x) \geq k$. But since $x \mapsto (1 - \frac{1}{x+2})^2$ is increasing for $x > 0$, $\Lambda(x) \geq k \iff x \geq c$ for some $c = c_\alpha$ determined by the size of the test. For the type I error,

$$P_{H_0}(X \geq c) = \int_c^\infty \frac{1}{(x+1)^2} dx = \frac{1}{c+1} = \alpha,$$

i.e. $c = 1/\alpha - 1 = \frac{1-\alpha}{\alpha}$. Thus the likelihood ratio test rejects H_0 if and only if $X \geq \frac{1-\alpha}{\alpha}$. The type II error is the probability of accepting H_0 when H_1 is true:

$$P_{H_1}(X < c_\alpha) = \int_0^{c_\alpha} \frac{2}{(x+2)^2} dx = \left[-\frac{2}{x+2}\right]_0^{c_\alpha} = 1 - \frac{2}{c_\alpha + 2} = \frac{1-\alpha}{1+\alpha}.$$

Hence for $\alpha = 0.05$, the probability of a type II error is $\frac{1-0.05}{1+0.05} = 19/21$.

For testing against $H_1 : \theta > 1$, we check that the family has a monotone likelihood ratio. For $\theta_2 > \theta_1$,

$$\frac{f_{\theta_2}(x)}{f_{\theta_1}(x)} = \frac{\theta_2}{\theta_1}\left(\frac{x+\theta_1}{x+\theta_2}\right)^2 = \frac{\theta_2}{\theta_1}\left(1 - \frac{\theta_2 - \theta_1}{x+\theta_2}\right)^2,$$

which is a strictly increasing function of $x > 0$. Thus the likelihood ratio test is also uniformly most powerful in this one-sided setting and we do not change the test (see the Karlin-Rubin theorem).

3. The likelihood ratio test statistics are given by

$$\Lambda_n(x) = \frac{\sup_{\theta\in\Theta}\prod_{i=1}^n f_\theta(X_i)}{\prod_{i=1}^n f_{\theta_0}(X_i)} = \frac{\prod_{i=1}^n f_{\hat\theta_{MLE}}(X_i)}{\prod_{i=1}^n f_{\theta_0}(X_i)}$$

and can be found in the following table (see Q1 on PS2 for the MLEs):

Distribution	$\prod_{i=1}^n f_\theta(x_i)$	$\Lambda_n(x)$
(a) Bernoulli(θ)	$\theta^{n\bar X_n}(1-\theta)^{n-n\bar X_n}$	$(\frac{\bar X_n}{\theta_0})^{n\bar X_n}(\frac{1-\bar X_n}{1-\theta_0})^{n-n\bar X_n}$
(b) N(θ,1)	$(2\pi)^{-n/2}e^{-\frac{1}{2}\sum(x_i-\theta)^2}$	$e^{\frac{1}{2}n(\bar X_n-\theta_0)^2}$
(c) N(0,θ)	$(2\pi\theta)^{-n/2}e^{-\frac{1}{2\theta}\sum x_i^2}$	$(\frac{\theta_0}{\overline{X^2}})^{n/2}e^{\frac{n}{2}(\frac{\overline{X^2}}{\theta_0}-1)}$
(d) N(μ,σ^2)	$(2\pi\sigma^2)^{-n/2}e^{-\frac{1}{2\sigma^2}\sum(x_i-\mu)^2}$	$(\frac{\sigma_0^2}{\hat\sigma^2})^{n/2}\exp(\frac{n}{2}(\frac{\hat\sigma^2}{\sigma_0^2}-1))$
(e) Poisson(θ)	$\theta^{n\bar X_n}e^{-n\theta}/(\prod x_i!)$	$(\frac{\bar X_n}{\theta_0})^{n\bar X_n}e^{n(\theta_0-\bar X_n)}$
(f) $(1/\theta)e^{-x/\theta}$	$\frac{1}{\theta^n}e^{-n\bar X_n/\theta}$	$(\frac{\theta_0}{\bar X_n})^n\exp(n(\frac{\bar X_n}{\theta_0}-1))$
(g) $\theta e^{-\theta x}$	$\theta^n e^{-n\theta\bar X_n}$	$(\frac{1}{\theta_0\bar X_n})^n e^{n(\theta_0\bar X_n-1)}$

4. (a) For two simple hypotheses, the Neyman-Pearson lemma says the likelihood ratio test is UMP. Since

$$x \mapsto \frac{f_1(x)}{f_0(x)} = \frac{x}{1-x}$$

is an increasing function of $x\in[0,1]$ (it has derivative $(1-x)^{-2}>0$ on $[0,1]$),

$$\frac{x}{1-x}\ge k \iff x\ge B$$

for some $B=B(k)$. Thus the UMP test has critical region of the desired form.

(b) The risks equal

$$R(\delta_B,H_0) = E_{H_0}1\{\delta_B\neq 0\} = P_{H_0}(X\ge B) = \int_B^1 2(x-1)dx = (1-B)^2,$$
$$R(\delta_B,H_1) = E_{H_1}1\{\delta_B\neq 1\} = P_{H_1}(X< B) = \int_0^B 2xdx = B^2.$$

The minimax test procedure minimizes the worst case risk over the whole parameter space $\{H_0,H_1\}$, i.e. we want to minimize

$$\sup_{\theta\in\{0,1\}} R(\delta_B,H_\theta) = \max\{R(\delta_B,H_0),R(\delta_B,H_1)\} = \max\{(1-B)^2,B^2\}$$
$$= \begin{cases}(1-B)^2 & \text{if } B\le 1/2,\\ B^2 & \text{if } B\ge 1/2.\end{cases}$$

This is minimized at $B=1/2$, so $\delta_{1/2}$ is the minimax test procedure.

(c) We compute the Bayes risk directly:

$$R_\pi(\delta_B) = E_{\theta\sim\pi}R(\delta_B,H_\theta) = \nu(1-B)^2+(1-\nu)B^2.$$

Minimizing this quadratic in B yields $B(\nu)=\nu$, i.e. δ_ν is the Bayes test procedure. [Note that if we give more prior weight to H_0 by taking ν closer to 1, then we take $B(\nu)=\nu$ larger to reduce the chance of incorrectly rejecting H_0].

Equating the risks for $\delta_{B(\nu)}=\delta_\nu$,

$$(1-\nu)^2 = R(\delta_{B(\nu)},H_0) = R(\delta_{B(\nu)},H_1) = \nu^2,$$

which implies $\nu = 1/2$. Thus for the prior with $\nu = 1/2$, the Bayes rule $\delta_{1/2}$ has constant risk and hence is minimax. Note this is the same estimator we computed directly in (b).

5. By the CLT, $\frac{\sqrt{n}\,\overline{X}_n}{\sigma} \to^d N(0,1)$. Furthermore,

$$S_n^2 = \frac{n}{n-1}\left(\frac{1}{n}\sum_{j=1}^n (X_j - \overline{X}_n)^2\right) = \frac{n}{n-1}\left(\frac{1}{n}\sum_{j=1}^n X_j^2 - (\overline{X}_n)^2\right) \to^p \sigma^2$$

since $\frac{1}{n}\sum_{j=1}^n X_j^2 \to^p \sigma^2$ and $\overline{X}_n \to^p 0$, both by the WLLN. Slutsky's lemma yields

$$t_n = \frac{\sqrt{n}\,\overline{X}_n}{S_n} \to^d N(0,1).$$

This implies that $\frac{\sqrt{n}(\bar{X}_n - EX_1)}{S_n}$ is an asymptotically pivotal quantity, i.e. its asymptotic distribution does not depend on EX_1. Let $q_{1-\alpha/2}$ be the $(1-\alpha/2)$-quantile of the normal distribution. Then we have

$$\lim_{n\to\infty} P(EX_1 \in [\overline{X}_n - n^{-1/2}S_n q_{1-\alpha/2}, \overline{X}_n + n^{-1/2}S_n q_{1-\alpha/2}]) = 1-\alpha.$$

6. For $\theta_2 > \theta_1$,

$$\frac{f_{\theta_2}(x)}{f_{\theta_1}(x)} = \frac{\prod_i \theta_2 e^{-\theta_2 x_i}}{\prod_i \theta_1 e^{-\theta_1 x_i}} = (\theta_2/\theta_1)^n e^{(\theta_1-\theta_2)\sum_i x_i}$$

is a *decreasing* function of $T(x) = \sum_i x_i$. Thus we have a monotone likelihood ratio (note the same proofs hold as long as the function is either increasing or decreasing). So by the Karlin-Rubin theorem, the UMP test is to reject H_0 if and only if $T(X) \le k$, where $P_0(T(X) \le k) = \alpha$. Since $T(X) \sim \Gamma(n, \theta_0)$ under H_0, the hint gives $2\theta_0 T(X) \sim \chi^2_{2n}$ under H_0. Therefore, take $k = \frac{1}{2\theta_0} q_{2n}(\alpha)$ the α-quantile of the χ^2_{2n} distribution, i.e. such that $P(\chi^2_{2n} \le q_{2n}(\alpha)) = \alpha$.

The power function equals

$$\pi_\phi(\theta) = P_\theta(\text{reject } H_0) = P_\theta(T(X) \le k) = P_\theta(2\theta T(X) \le 2\theta k) = P\left(\chi^2_{2n} \le \frac{\theta}{\theta_0} q_{2n}(\alpha)\right).$$

We now invert the hypothesis test. Let

$$A(\theta') = \{(x_1, \ldots, x_n) : \sum x_i \ge \tfrac{1}{2\theta'} q_{2n}(\alpha)\}$$

be the non-rejection region for $H_0' : \theta = \theta'$. Then a $(1-\alpha)100\%$ confidence region is

$$C(X) = \{\theta : (X_1, \ldots, X_n) \in A(\theta)\} = \{\theta : \sum X_i \ge \tfrac{1}{2\theta} q_{2n}(\alpha)\} = \left[\frac{q_{2n}(\alpha)}{2\sum_i X_i}, \infty\right).$$

7. (a) The log-likelihood equals

$$\ell_n(\theta, \nu) = \log\left(\prod_{i=1}^n \frac{\theta\nu^\theta}{x_i^{\theta+1}} 1_{[\nu,\infty)}(x_i)\right) = n\log\theta + n\theta\log\nu - (\theta+1)\log\left(\prod x_i\right) - \infty 1\{\min_i x_i < \nu\},$$

where $\infty \times 0 = 0$. For any $\theta > 0$, ℓ_n is an increasing function of ν until the indicator function becomes 1, so the log-likelihood is maximized at $\hat{\nu} = \min_i x_i$. For the MLE of θ,

$$\frac{\partial \ell_n(\theta, \hat{\nu})}{\partial\theta} = \frac{n}{\theta} + n\log\hat{\nu} - \log\left(\prod x_i\right) = 0$$
$$\frac{\partial^2 \ell_n(\theta, \hat{\nu})}{\partial\theta^2} = -\frac{n}{\theta^2} < 0,$$

which is solved by

$$\hat{\theta} = \frac{n}{\log(\prod_i x_i) - n\log\hat{\nu}} = \frac{n}{\log(\prod_i x_i/(\min_i x_i)^n)} = \frac{n}{T}.$$

(b) Under H_0, the MLE for θ is $\hat{\theta}_0 = 1$ and the MLE for ν is unchanged. Since $T = \log((\prod_i x_i)/\hat{\nu}^n)$, writing the likelihood ratio statistic as a function of T,

$$\Lambda(x) = \frac{L_n(\hat{\theta},\hat{\nu})}{L_n(\hat{\theta}_0,\hat{\nu})} = \frac{\prod_i \frac{\hat{\theta}\hat{\nu}^{\hat{\theta}}}{x_i^{\hat{\theta}+1}}}{\prod_i \frac{\hat{\nu}}{x_i^2}} = \hat{\theta}^n \left(\frac{\hat{\nu}^n}{\prod_i x_i}\right)^{\hat{\theta}-1} = (n/T)^n (e^{-T})^{n/T-1} = \left(\frac{n}{T}\right)^n e^{T-n}.$$

Differentiating,

$$\frac{\partial}{\partial T}\log\Lambda(x) = \frac{\partial}{\partial T}(n\log n - n\log T + T - n) = 1 - n/T.$$

Hence $\Lambda(x)$ is increasing if $T \geq n$ and decreasing if $T \leq n$. So $\Lambda(x) \geq k$ is equivalent to $T \leq c_1$ or $T \geq c_2$ for suitably chosen constants c_1, c_2.

(c) Expanding out T:

$$T(X) = \sum_{i=1}^n [\log(X_i) - \log(\min_i X_i)].$$

We first consider the distribution of $Y_i = \log X_i$: for $e^t \geq \nu$,

$$P(\log X_i \leq t) = P(X_i \leq e^t) = \int_\nu^{e^t} \frac{\theta\nu^\theta}{x^{\theta+1}}dx = \left[-v^\theta x^{-\theta}\right]_\nu^{e^t} = 1 - (\nu e^{-t})^\theta.$$

Therefore, setting $\nu e^{-t} = e^{-s}$ gives for all $s \geq 0$,

$$P(\log(X_i/\nu) \leq s) = P(X_i \leq \nu e^{-s}) = 1 - e^{-s\theta},$$

i.e. $Y_i = \log(X_i/\nu) \sim^{iid} \text{Exp}(\theta) = \text{Exp}(1)$ since $\theta = 1$ under H_0. We now establish the joint distribution of $Y_i - \min_j Y_j$ for $i = 1,\ldots,n$. Since the Y_i are i.i.d., the minimum is attained by each Y_j with equal probability $1/n$. Suppose $Y_1 = \min_j Y_j$ attains the minimum. Then $Y_1 - \min Y_j = Y_1 - Y_1 = 0$ and for $t_2,\ldots,t_n \geq 0$,

$$P(Y_i - \min_j Y_j \geq t_i,\ i = 2,\ldots,n | Y_1 = \min_j Y_j) = \frac{P(Y_i - Y_1 \geq t_i,\ i = 2,\ldots,n)}{P(Y_1 = \min_j Y_j)}.$$

Since $Y_1,\ldots,Y_n \sim \text{Exp}(1)$,

$$\begin{aligned}
P(Y_i - Y_1 \geq t_i,\ i = 2,\ldots,n) &= \int_{y_1=0}^\infty e^{-y_1} \prod_{i=2}^n \int_{y_i=y_1+t_i}^\infty e^{-y_i} dy_i dy_1 \\
&= \int_{y_1=0}^\infty e^{-y_1} \prod_{i=2}^n e^{-(y_1+t_i)} dy_1 \\
&= e^{-\sum_{i=2}^n t_i} \int_{y_1=0}^\infty e^{-ny_1} dy_1 \\
&= \frac{1}{n} e^{-\sum_{i=2}^n t_i} = \frac{1}{n}\prod_{i=2}^n e^{-t_i} = \frac{1}{n}\prod_{i=2}^n P(\text{Exp}(1) \geq t_i).
\end{aligned}$$

In particular,

$$P(Y_i - \min_j Y_j \geq t_i,\ i = 2,\ldots,n | Y_1 = \min_j Y_j) = \prod_{i=2}^n e^{-t_i}.$$

Thus conditional on $Y_1 = \min_j Y_j$, we have that $(Y_1 - \min_j Y_j, \ldots, Y_n - \min_j Y_j)$ has the first coordinate equal to zero and the other $i = 2, \ldots, n$ coordinates equal to independent Exp(1) distributions, so that

$$\sum_{i=1}^{n}[Y_i - \min_j Y_j]\Big|Y_1 = \min_j Y_j =^d \sum_{i=2}^{n}\mathrm{Exp}(1) \sim \Gamma(n-1, 1).$$

However, the same argument applies to any of $Y_2, \ldots, Y_n$ attaining the minimum, and so $\sum_{i=1}^{n} Y_i - \min_j Y_j \sim \Gamma(n-1, 1)$ irrespective of which Y_i attains the minimum.
Finally, note that if $Z \sim \chi^2_r$ and $c > 0$, then $cZ \sim \Gamma(r/2, 1/(2c))$. Setting $r = 2(n-1)$ and $c = 1/2$ yields if $Z \sim \chi^2_{2(n-1)}$ then $\frac{1}{2}Z \sim \Gamma(n-1, 1)$. Setting $Z = 2T = 2\sum_{i=1}^{n}[Y_i - \min_j Y_j] \sim \chi^2_{2(n-1)}$ gives the result.