Statistical Inference: Examination

30 January 2024

Instructions: The time allotted for the examination is 180 minutes. You may answer in either English or French. No written material may be brought into the examination, but a simple calculator may be used if necessary. Full marks may be obtained with complete answers to four questions. The final mark will be based on the best four solutions.

First name:

Last name:

SCIPER number:

Exercise	Points	Indicative marks
1		/10 points
2		/10 points
3		/10 points
4		/10 points
5		/10 points
Total:		/40 points

Some formulae

Definition 1 *The moment-generating and cumulant-generating functions of a real-valued random variable X are*

$$M_X(t) = \mathrm{E}\left(e^{tX}\right), \quad K_X(t) = \log M_X(t), \quad t \in \mathcal{T},$$

where $\mathcal{T} = \{t \in \mathbb{R} : M_X(t) < \infty\}$.

Definition 2 *A Bernoulli random variable with parameter $p \in (0,1)$ has probability mass function*

$$f(x; p) = p^x (1-p)^{1-x}, \quad x \in \{0, 1\}.$$

Definition 3 *A geometric random variable with parameter $p \in (0,1)$ has probability mass function*

$$f(x; p) = (1-p)^{x-1} p, \quad x \in \{1, 2, \ldots\}.$$

Definition 4 *A Poisson variable with parameter $\lambda > 0$ has probability mass function*

$$f(x; \lambda) = \frac{\lambda^x}{x!} e^{-\lambda}, \quad x \in \{0, 1, \ldots\}.$$

Definition 5 *A normal (or Gaussian) random variable $X \sim \mathcal{N}(\mu, \sigma^2)$ has probability density function*

$$f(x; \mu, \sigma^2) = \frac{1}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right), \quad x \in \mathbb{R}, \quad \mu \in \mathbb{R}, \sigma^2 > 0,$$

where $\phi(u) = (2\pi)^{-1/2} e^{-u^2/2}$ for $u \in \mathbb{R}$, and we also define $\Phi(x) = \int_{-\infty}^{x} \phi(u)\,\mathrm{d}u$.

Definition 6 *A gamma random variable with shape parameter $\alpha > 0$ and rate parameter $\beta > 0$, $X \sim \mathrm{Gamma}(\alpha, \beta)$, has probability density function*

$$f(x; \alpha, \beta) = \begin{cases} \dfrac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x \geq 0, \\ 0, & x < 0, \end{cases}$$

where $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$, $\Gamma(\alpha) = (\alpha-1)!$ when α is a positive integer, and $\Gamma(1/2) = \sqrt{\pi}$.

Definition 7 *An exponential random variable X with rate parameter β, $X \sim \exp(\beta)$, has the gamma distribution with $\alpha = 1$.*

Definition 8 *A chi-squared random variable V with ν degrees of freedom, $V \sim \chi^2_\nu$, has the gamma distribution with $\alpha = \nu/2$ and $\beta = 1/2$, and can be expressed as $V \stackrel{\mathrm{D}}{=} Z_1^2 + \cdots + Z_\nu^2$, where $Z_1, \ldots, Z_\nu \stackrel{\mathrm{iid}}{\sim} \mathcal{N}(0,1)$.*

Solution 1

(a) [3, seen] Slides 28, 79, 80

(b) [3, unseen] We are told that the X_j are independent and identically distributed, so $\mathrm{E}(X_j) = p$ and $\mathrm{var}(X_j) = p(1-p)$, and

$$\mathrm{E}(T) = \frac{n}{n}\mathrm{E}\{X_{2j-1}(1-X_{2j})\} = \mathrm{E}(X_{2j-1})\mathrm{E}(1-X_{2j}) = p(1-p),$$

and as $X_j^2 = X_j$ and $(1-X_j)^2 = 1-X_j$,

$$\begin{aligned}
\mathrm{var}(T) &= n^{-1}\mathrm{var}\{X_{2j-1}(1-X_{2j})\} \\
&= n^{-1}\left[\mathrm{E}\left\{X_{2j-1}^2(1-X_{2j})^2\right\} - \mathrm{E}\{X_{2j-1}(1-X_{2j})\}^2\right] \\
&= n^{-1}\left[\mathrm{E}\left\{X_{2j-1}(1-X_{2j})\right\} - \mathrm{E}\{X_{2j-1}(1-X_{2j})\}^2\right],
\end{aligned}$$

and this equals $p(1-p)\{1-p(1-p)\}/n$.

Alternatively we could note that $I_j = X_{2j-1}(1-X_{2j}) = I(X_{2j-1} = 1, X_{2j} = 0)$ are independent indicator variables with success probability $p(1-p)$, giving the same result.

(c) [2, unseen] We have

$$\mathrm{E}(\overline{X}) = p, \quad \mathrm{var}(\overline{X}) = \mathrm{E}(\overline{X}^2) - \mathrm{E}(\overline{X})^2 = p(1-p)/(2n),$$

so

$$\mathrm{E}\left\{\overline{X}(1-\overline{X})\right\} = \mathrm{E}(\overline{X}) - \mathrm{E}(\overline{X}^2) = p - \left\{p(1-p)/(2n) + p^2\right\} = (2n-1)p(1-p)/(2n),$$

so $S = 2n\overline{X}(1-\overline{X})/(2n-1)$ is unbiased for θ.

(d) [2, unseen] $\overline{X}$ is minimal sufficient for p, and therefore for θ, and it is complete because the Bernoulli model is an exponential family. Therefore S is the unique minimum variance unbiased estimator of θ, using the Rao–Blackwell theorem, so it has the smallest possible variance among T and any other unbiased estimators.

Solution 2

(a) [4, unseen] The likelihood is defined for $\beta, \gamma > 0$ and $\psi \geq 0$, and is

$$L(\psi,\beta,\gamma) = f(y_1;\psi,\beta,\gamma)f(y_2;\beta)f(y_3;\gamma) = \frac{1}{y_1!y_2!y_3!}(\beta+\gamma\psi)^{y_1}e^{-(\beta+\gamma\psi)}(\beta u)^{y_2}e^{-\beta u}(\gamma t)^{y_3}e^{-\gamma t},$$

which reduces to

$$m(y)\exp\left\{y_1\varphi_1 + y_2\varphi_2 + y_3\varphi_3 - k(\varphi)\right\}$$

with canonical statistic $s(y) = (y_1, y_2, y_3)$ and canonical parameter $\varphi = (\varphi_1, \varphi_2, \varphi_3)$, where

$$\varphi_1 = \log(\beta+\gamma\psi), \quad \varphi_2 = \log\beta, \quad \varphi_3 = \log\gamma, \quad k(\varphi) = e^{\varphi_1} + ue^{\varphi_2} + te^{\varphi_3},$$

and $m(y)$ absorbs the multiplicative constants.

(b) [4, seen/unseen] With $\gamma = \alpha\beta$ we have an exponential family of the form above with

$$\varphi_1' = \log(1+\alpha\psi), \quad \varphi_2' = \log\beta, \quad \varphi_3' = \log\alpha, \quad y_1' = y_1, \quad y_2' = y_1+y_2+y_3, \quad y_3' = y_3,$$

say, so the distribution of (y_1, y_2, y_3) conditional on $w = y_1 + y_2 + y_3$ does not depend on $\log\beta$ (or equivalently on β). Now w has a Poisson distribution with mean $\beta(1+\alpha\psi+u+t\alpha)$ (i.e., the sum of the means of y_1, y_2, y_3), so the conditional density is

$$\frac{(y_1+y_2+y_3)!}{y_1!y_2!y_3!}\frac{(1+\alpha\psi)^{y_1}u^{y_2}(t\alpha)^{y_3}}{(1+\alpha\psi+u+t\alpha)^{y_1+y_2+y_3}},$$

corresponding to observing a trinomial variable (y_1, y_2, y_3) with probabilities

$$\pi_1 = \frac{1+\alpha\psi}{1+\alpha\psi+u+t\alpha}, \quad \pi_2 = \frac{u}{1+\alpha\psi+u+t\alpha}, \quad \pi_3 = \frac{t\alpha}{1+\alpha\psi+u+t\alpha}.$$

The transformation is interest-respecting, because ψ is unchanged.

(c) [2, unseen] No; it is a $(3K, 2K+1)$ curved exponential family.

Solution 3

(a) [3, seen] Slides 91–98, 107–108

(b) [3, seen] Slides 109–110

(c) [4, unseen] The density function is

$$f(y;\alpha,\lambda) = \frac{\alpha y^{\alpha-1}}{\lambda^\alpha}\exp\left\{-(y/\lambda)^\alpha\right\}, \quad y > 0,$$

so the log likelihood based on a random sample $y_1, \ldots, y_n$ is

$$\ell(\alpha,\lambda) = n\log\alpha + (\alpha-1)\sum_{j=1}^n \log y_j - n\alpha\log\lambda - \lambda^{-\alpha}\sum_{j=1}^n y_j^\alpha,$$

which thus gives $S(\alpha) = \sum_{j=1}^n y_j^\alpha$. Differentiation with respect to λ gives

$$\frac{\partial\ell(\alpha,\lambda)}{\partial\lambda} = -\frac{n\alpha}{\lambda} + \alpha\lambda^{-\alpha-1}S(\lambda), \qquad \frac{\partial^2\ell(\alpha,\lambda)}{\partial\lambda^2} = \frac{n\alpha}{\lambda^2} - \alpha(\alpha+1)\lambda^{-\alpha-2}S(\lambda),$$

so $\widehat{\lambda}_\alpha = \{n^{-1}S(\lambda)\}^{1/\alpha}$ is the root of the first equation, and inserting this into the second derivative gives

$$\frac{n\alpha}{\widehat{\lambda}_\alpha} - \alpha(\alpha+1)\widehat{\lambda}_\alpha^{-\alpha-2}S(\lambda) = -n\alpha^2/\widehat{\lambda}_\alpha^2 < 0,$$

so $\widehat{\lambda}_\alpha$ is the MLE of λ for fixed α. Substitution of this into $\ell(\alpha,\lambda)$ gives the stated expression after a little algebra; call this $\ell_{\mathrm{p}}(\alpha)$.

To verify if $\alpha = 1$, we would find the MLE $\widehat{\alpha}$ and then compare $W_p(1) = 2\{\ell_{\mathrm{p}}(\widehat{\alpha}) - \ell_{\mathrm{p}}(1)\}$ with the χ_1^2 distribution. The results in (b) imply that large values of $W_p(1)$ relative to the χ_1^2 distribution would be suggestive that $\alpha \neq 1$.

Solution 4

(a) [4, seen] Slides 124–130

(b) [3, seen] Slides 181–182

(c) [3, seen] If S is minimal sufficient for θ, then the marginal density of Y is

$$f(y) = \int f(y \mid \theta)\pi(\theta)\,\mathrm{d}\theta = \int f(y \mid s)f(s \mid \theta)\pi(\theta)\,\mathrm{d}\theta = f(y \mid s)\mathrm{E}_\theta\left\{f(s \mid \theta)\right\}.$$

Hence the ratio of marginal densities for y under two different priors is

$$\frac{f_1(y)}{f_0(y)} = \frac{f(y \mid s)\mathrm{E}_1\left\{f(s \mid \theta)\right\}}{f(y \mid s)\mathrm{E}_0\left\{f(s \mid \theta)\right\}} = \frac{\mathrm{E}_1\left\{f(s \mid \theta)\right\}}{\mathrm{E}_0\left\{f(s \mid \theta)\right\}},$$

as required.

The comparison is of different prior densities for θ, so the conditional density of Y given S, which does not depend on θ, is irrelevant. So this is not surprising.

Solution 5

(a) [2, unseen] If we assume that the times to death have common distribution F, then the probability of death by time c is $F(c)$, and the probability of being alive is thus $1 - F(c)$. Hence if d is the indicator that the individual is alive, the corresponding likelihood contribution, $F(c)^{1-d}\{1 - F(c)\}^d$, yields the given likelihood if the outcomes are independent.

(b) [3, unseen] Writing $p(\lambda) = \exp(-\lambda c)$ and with $s = \sum_j d_j$ survivors, the log likelihood can be written as

$$\ell(\lambda) = (n - s)\log\{1 - p(\lambda)\} + s\log p(\lambda), \quad \lambda > 0,$$

so $p(\widehat{\lambda}) = s/n$, which yields $\widehat{\lambda} = c^{-1}\log(n/s)$. For the Fisher information we note that $S \sim B\{n, p(\lambda)\}$, and then after a little work obtain

$$\mathrm{E}\left\{-\frac{\partial^2\ell(\lambda)}{\partial\lambda^2}\right\} = -\mathrm{E}\left\{\left(\frac{\partial p}{\partial\lambda}\right)^2\frac{\partial^2\ell}{\partial p^2} + \frac{\partial^2 p}{\partial\lambda^2}\frac{\partial\ell}{\partial p}\right\} = \frac{nc^2p(\lambda)}{1 - p(\lambda)},$$

because $\partial p(\lambda)/\partial\lambda = -cp(\lambda)$, $\mathrm{E}(\partial\ell/\partial p) = 0$ and $\mathrm{E}(S) = np(\lambda)$.

(c) [2, unseen] If an individual failure time is observed exactly up to c and and right-censored, then the likelihood contribution will be $f(y)$ if $y \le c$, i.e., $D = 0$, and will be $1 - F(c)$ if $y > c$, i.e., $D = 1$. The likelihood contribution is then $f(y)^{1-d}\{1 - F(c)\}^d$, and the given formula is the product of these terms over the n independent individuals.

(d) [3, unseen] In this case the likelihood contribution for an individual is $(\lambda e^{-\lambda y})^{1-d}(e^{-\lambda c})^d$, so with $s = \sum_j d_j$ the overall log likelihood is

$$\sum_{j=1}^n (1 - d_j)(\log\lambda - \lambda y_j) - \lambda cs, \quad \lambda > 0.$$

This has second derivative $-(n - s)/\lambda^2$, leading to Fisher information $\{n - \mathrm{E}(S)\}/\lambda^2 = n\{1 - p(\lambda)\}/\lambda^2$ so the asymptotic relative efficiency of using current status data is

$$\frac{nc^2p(\lambda)}{1 - p(\lambda)} \div \frac{n\{1 - p(\lambda)\}}{\lambda^2} = \frac{\lambda^2c^2p(\lambda)}{\{1 - p(\lambda)\}^2} = \frac{p(\lambda)\{\log p(\lambda)\}^2}{\{1 - p(\lambda)\}^2}.$$

—————— END OF THE EXAM PAPER ——————