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*

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

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

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

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 $Z_1,\dots,Z_\nu \stackrel{\mathrm{iid}}{\sim} \mathcal{N}(0,1)$.*

Question 1

(a) Explain the meaning of the italicised terms in the phrase 'the *statistical model* was formulated in order to find a *minimum variance unbiased estimator* of a *parameter*'.

(b) Independent Bernoulli variables $X_1, \ldots, X_{2n}$ with common success probability p are available. Show that

$$T = \frac{1}{n} \sum_{j=1}^{n} X_{2j-1}(1 - X_{2j})$$

is an unbiased estimator of $\theta = p(1-p)$ and find its variance.

(c) If $\overline{X} = (2n)^{-1} \sum_{j=1}^{2n} X_j$, find an unbiased estimator of θ based on $\overline{X}(1 - \overline{X})$.

(d) How does your estimator compare to T and any other unbiased estimators of θ?

Question 2 Many particle physics experiments result in noisy count data. A single-channel experiment results in three independent Poisson variables y_1, y_2, y_3 with respective means $\beta + \gamma\psi$, βu, γt, where u and t are positive and known, and the parameters β, γ and $\psi \geq 0$ are unknown. The parameter $\beta > 0$ represents a background rate for the event of interest and is estimated from an experiment that gives y_2 events over a time period u, and $\gamma \in (0, 1)$, which represents the efficiency of the detector, is estimated from another experiment that gives y_3 events over a time period t. The main experiment results in y_1 and is intended to assess whether $\psi > 0$, which would indicate the presence of more particles than could be explained by the background, and might therefore indicate a need for 'new physics'.

(a) Write the above model in exponential family form, giving the canonical parameters and statistics.

(b) Show that if we write $\gamma = \beta\alpha$ for some $\alpha > 0$, then β can be removed by a conditioning argument. Is this an interest-respecting transformation?

(c) K independent channels each give independent Poisson observations y_{1k}, y_{2k}, y_{3k} with corresponding means $\beta_k + \gamma_k\psi$, $\beta_k u_k$, $\gamma_k t_k$ $(k = 1, \ldots, K)$, where $u_1, \ldots, u_K$ and $t_1, \ldots, t_K$ are known and the parameters must be estimated. Is this model a full exponential family?

Question 3 Suppose that $Y_1, \ldots, Y_n$ are independent realisations of a random variable Y whose density function $f(y; \theta)$ satisfies suitable regularity conditions and where the vector parameter θ lies in an open subset of $\mathbb{R}^d$.

(a) State the limiting distribution of the maximum likelihood estimator $\widehat{\theta}$ as $n \to \infty$, and explain how this result can be used for inference on elements of θ.

(b) What is a *profile log likelihood*? Under what circumstances would you use one, and why?

(c) Show that the log likelihood for a Weibull random sample $y_1, \ldots, y_n$, for which

$$\mathrm{P}(Y > y) = \exp\left\{-(y/\lambda)^\alpha\right\}, \quad y > 0, \quad \alpha, \lambda > 0,$$

can be written in the form

$$\ell(\alpha, \lambda) = n \log \alpha - n\alpha \log \lambda + (\alpha - 1) \sum_{j=1}^{n} \log y_j - S(\alpha)/\lambda^\alpha, \quad \alpha, \lambda > 0,$$

where $S(\alpha)$ should be specified, and deduce that apart from additive constants,

$$\max_{\lambda} \ell(\alpha, \lambda) = n \log \alpha - n \log S(\alpha) + \alpha \sum_{j=1}^{n} \log y_j, \quad \alpha > 0.$$

How would you use this function to verify whether $\alpha = 1$?

Question 4

(a) Explain the terms *null hypothesis*, *test statistic* and *P-value*. A hypothesis test is performed and gives P-value 0.001. What do you conclude?

(b) Define the *Bayes factor* and say how it is used in Bayesian hypothesis testing.

(c) Show that Bayesian inference for a parameter θ depends on data Y only through a minimal sufficient statistic $S = s(Y)$, and deduce that the Bayes factor for comparing two models with the same minimal sufficient statistic S for θ can be expressed as

$$\mathrm{E}_1\{f(s \mid \theta)\}/\mathrm{E}_0\{f(s \mid \theta)\},$$

where the expectations are over the different prior distributions for θ under the models. Do you find this surprising? Explain.

Question 5 In current status data all that is known about individuals is their status at a single time. For example, at time zero n skiers are struck by an avalanche, and when rescuers locate skier j at a later time c_j they find that s/he is either alive (1) or dead (0).

(a) The likelihood for the survival time distribution F based on such a dataset is stated to be $\prod_{j=1}^{n} F(c_j)^{1-d_j}\{1 - F(c_j)\}^{d_j}$. On what assumptions does this depend?

(b) If $F(x) = 1 - \exp(-\lambda x)$, for $\lambda > 0$ and $x > 0$, and all the c_j are equal, then find the maximum likelihood estimator of λ and the corresponding Fisher information.

(c) Under right-censoring the observations are of the form $(Y, D) = (\min(T, c), I(T > c))$; i.e., the failure time is T observed exactly up to a non-random censoring time c, but otherwise c is recorded, and D is the indicator of survival beyond c. Explain why the likelihood for independent observations is then $\prod_{j=1}^{n} f(y_j)^{1-d_j}\{1 - F(c_j)\}^{d_j}$, where f is the density function corresponding to F.

(d) Find the asymptotic relative efficiency of the estimator in (b) relative to that in (c) when $T \sim \exp(\lambda)$ and $c_1 = \cdots = c_n = c$.

——————— END OF THE EXAM PAPER ———————