Statistical Inference: Dummy Examination

1 April 1349

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. 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	Marks	Indicative marks
1		10
2		10
3		10
4		10
5		10
Total:		40

Notation

The material below **may or may not** be useful in some questions.

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

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

Definition 4 *A uniform random variable $X \sim U(a, b)$ with $a < b$ has probability density function*

$$f(x; a, b) = \begin{cases} \frac{1}{b-a}, & a < x \leq b, \\ 0, & \textit{otherwise}. \end{cases}$$

Definition 5 *A normal (or Gaussian) random variable $X \sim N(\mu, \sigma^2)$ 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 $Z_1, \ldots, Z_\nu \stackrel{\mathrm{iid}}{\sim} \mathcal{N}(0, 1)$.*

Question 1 Where does the probability come from? Write an essay on the probabilistic basis of statistics in which you describe, compare and contrast at least three main approaches to inference, giving examples as appropriate.

Question 2

(a) How does a *minimal sufficient statistic* differ from a *sufficient statistic*?

(b) State the factorisation theorem and explain how it is useful in finding a sufficient statistic for the parameter θ of a statistical model $f(y;\theta)$. How would you verify whether a sufficient statistic is minimal?

(c) The Poisson random variables Y_1 and Y_2 are independent with respective means $m\theta$ and $m(1-\theta)$, where $m>0$ is known and $\theta \in (0,1)$. Find a minimal sufficient statistic for θ. Show that the value of $Y_1 + Y_2$ is not informative about θ, and explain how this might be used in inference.

Question 3

(a) Large values of a continuous test statistic T are considered to give evidence against a null hypothesis H_0. Explain the terms *null distribution of* T and *P-value* in this context.

Express the P-value P in terms of T and its null distribution F and show that $P \sim U(0,1)$ when H_0 is true.

(b) Independent P-values $P_1, \ldots, P_m$ result from testing the null hypotheses $H_1, \ldots, H_m$. Show that $P^* = 1 - \{\max_{j=1}^m (1-P_j)\}^m$ has a $U(0,1)$ distribution under the global null hypothesis H_0 that $H_1, \ldots, H_m$ are all true.

Under what circumstances would P^* cast doubt on H_0?

(c) How does the Bonferroni procedure for testing H_0 compare to the use of P^*?

Question 4

(a) What problems might be posed by the presence of nuisance parameters when performing likelihood inference for a scalar parameter of interest? How might you attempt to solve such problems?

(b) What is a profile log likelihood? When might you use one?

(c) Independent observations arise as follows: $y_{j1}, y_{j2} \overset{\text{iid}}{\sim} \mathcal{N}(\mu_j, \sigma^2)$, for $j = 1, \ldots, m$. Find the maximum likelihood estimators of the μ_j and show that (apart from additive constants) the profile log likelihood for σ^2 is

$$-m \log \sigma^2 - \frac{1}{4\sigma^2} \sum_{j=1}^{m} (y_{j1} - y_{j2})^2, \quad \sigma^2 > 0.$$

Deduce that the maximum likelihood estimator of σ^2 converges to $\sigma^2/2$.

(d) Discuss how the incorrect convergence in (c) can be remedied.

Hint: $z_j = y_{j1} - y_{j2} \overset{\text{iid}}{\sim} \mathcal{N}(0, 2\sigma^2)$.

Question 5

(a) What is a hierarchical Bayesian model? In what circumstances would you think of using one?

(b) Observations $y_1, \ldots, y_n$ are available under the following setup:

$$\begin{aligned} y_j \mid \lambda_j &\overset{\text{ind}}{\sim} \exp(\lambda_j), \\ \lambda_1, \ldots, \lambda_n \mid \beta &\overset{\text{iid}}{\sim} \text{Gamma}(\alpha, \beta), \\ \beta &\sim \exp(\tau), \end{aligned}$$

where $\text{Gamma}(\alpha, \beta)$ denotes the gamma distribution with shape parameter α and rate parameter β, and α and τ are known. Deduce that

$$f(y_1, \ldots, y_n) = \int_0^\infty \tau \alpha^n \beta^{n\alpha} e^{-\tau\beta} \prod_{j=1}^n (y_j + \beta)^{-\alpha-1} \, \mathrm{d}\beta,$$

and outline the steps you would need to obtain a Laplace approximation to this integral.

(c) Express the marginal posterior density for λ_1 as a ratio of two integrals, and discuss how they might be approximated. Would you expect the resulting approximation to be more or less accurate than that in (b)?