Norwegian University of Science and Technology
Department of Mathematical Sciences

TMA4295 Statistical inference Fall 2023

Exercise set 7

Read the questions carefully and make your own assumptions if needed.

1 Bernoulli

Let the data $x = (x_1, \dots, x_n)$ be a random sample from the Bernoulli $\mathrm{B}(p)$.

a) Calculate the Fisher information in a single Bernoulli variable, in the sample, and in $y = n\overline{x}$. Hint: Use independence and sufficiency.

b) Determine the UMVU estimators of the parameters $\operatorname{Var} X_1$ and p^2. Compare the variances with the corresponding Cramer-Rao lower bounds. Calculate the Fisher information in the estimators and the resulting Cramer-Rao lower bounds.

c) Let T be a statistic. Show that $\tau = \mathrm{E}\, T$ is a polynomial in p, and that any n'th order polynomial can be obtained like this. Can you find UMVU estimators for τ, $1/p$, and $\eta = \ln p - \ln(1-p)$? What about UMRU estimators using the loss $l = (\hat{\eta} - \eta)^2$?

d) When is it possible for a statistic T to obtain the lower bound in the Cramer-Rao inequality? Hint: Equality in the Cauchy-Schwarz inequality.

e) Let f be the $\mathrm{B}(p)$ density. Explain that $\sqrt{f}$ is a unit vector in the plane and illustrate with a drawing. Identify p with the corresponding point in the plane. Find a formula for the direct distance between $\hat{p}$ and p, and also the distance along the unit circle. Explain that this gives two different metrics on the space of Bernoulli densities. Is there a relation between the resulting two metric spaces?

f) What is the Jeffreys' prior for p? Show that the Haldane prior $1/[p(1-p)]$ gives a uniform prior on $\eta = \ln p - \ln(1-p)$. Show that both priors determines a metric by

$$\mathsf{d}(\hat{p}, p) = \left| \int_{\hat{p}}^{p} \pi(p)\, dp \right| \tag{1}$$

and calculate explicit formulas for the distances. Find a formula for the distance given by the Fisher information metric.

g) Find a formula for the Kullbach-Leibler divergence $D(\hat{p} \,\|\, p)$. Is this given by a prior density? Decide if the previous distances and divergence define convex loss functions.

h) How would you calculate Bayes estimators for p using the previous priors and with loss given by absolute distance or distance squared?

2 Fisher information for the multinomial

Let f be the density a random point X with values $\mathrm{R}(X) = \{x_1, \ldots, x_m\}$.

a) Explain that $\sqrt{f}$ is a unit vector in the vector space $\mathbb{R}^m$. Assume P_X is known when the time t is known. Explain that this defines a statistical model, and also a path on the unit sphere in $\mathbb{R}^m$. Let $\iota(t) = \mathrm{E}\, S^2$ with $S = \partial_t \ln f(X)$. Prove that the length of a part of the assumed smooth path is

$$\mathsf{d}(t_1, t_2) = \left| \int_{t_0}^{t_1} \frac{\sqrt{\iota(t)}}{2} \, dt \right| \tag{2}$$

Hint: Distance is given by speed and time.

b) Let $\tau = \mathrm{E}\, T$. Prove that the Cauchy inequality $|\mathrm{E}[(T-\tau)S]| \leq \|T-\tau\| \, \|S\|$ implies the Cramer-Rao inequality $\mathrm{Var}\, T \geq \dot{\tau}^2/\iota$.

c) Assume P_X is known when θ is known and assume θ is known when t is known. Prove $\iota(t) = \dot{\theta}\iota\dot{\theta}^T$ where ι is the Fisher information. Explain that ι is a differential metric on the model space Ω_Θ. The length element and the volume element $\sqrt{|\det(\iota)|}\, d\theta$ are both coordinate independent. What does this mean? Hint: Equation (2) and the chain rule for differentiation.

d) Prove the multivariate Cramer-Rao inequality. Hint: Reduce to a component of T and use the chain rule on the one-dimensional Cramer-Rao inequality.

e) Are the previous arguments valid if it is only assumed that f is a density? Find $\mathsf{d}(\beta_1, \beta_2)$ for exponential distributions with scales β_1 and β_2.

3 Standard uncertainty and Cramer-Rao

Let $X \sim \mathrm{N}(\theta, \Sigma)$ where the variance matrix $\mathrm{Var}\, X = \Sigma$ is known.

a) Let $\tau = A + B\theta$ where A and B are matrices. Find an UMVU T for τ and determine its distribution. Compare $\mathrm{Var}\, T$ with the Cramer-Rao lower bound.

b) Let τ be a parameter. Determine the MLE of τ. Justify that the MLE has a distribution close to the normal distribution if Σ is small. Compare the variance of this normal distribution with the Cramer-Rao lower bound.

c) Explain that the previous gives a recipe for calculating a standard uncertainty when estimating τ, but that the actual uncertainty is larger. Explain how you could use simulation on a computer to find an improved standard uncertainty. Explain how you could use a finite difference approximation to calculate a standard uncertainty if calculation of τ is costly.

d) The length λ of a pendulum with period τ is

$$\lambda = \left(\frac{\tau}{2\pi}\right)^2 g \tag{3}$$

Assume that T estimates τ and that S estimates g. Find an estimator of λ. Assume that T and S are unbiased with normal distribution and known variances. Find an UMVU of λ. Determine standard uncertainties for the two competing estimates, and compare with the resulting Cramer-Rao lower bounds.