Universitetet i Oslo

Matematisk institutt

Continuation exam in: **STK 4011 – Statistical inference theory**

Day of examination: **Monday 20/1/2020**

Examination hours: **09:00–13:00**

Permitted aids: **One single sheet of paper with the candidate's own handwritten notes**

This exam set contains 3 exercises and comprises 3 pages.

Exercise 1

Consider the density given by

$$f(x;\theta,b)=\frac{x^{1/\theta-1}}{\theta b^{1/\theta}},\quad \text{for } 0\le x\le b,$$

and zero otherwise, where $\theta, b>0$. Below you may use that if Y is a Gamma distributed random variable with positive parameters (a,b), it has density $g(y)=\{b^a/\Gamma(a)\}y^{a-1}\exp(-by)$, $y>0$.

(a) Show that the expectation of $X\sim f(x;\theta,b)$ is

$$\mathrm{E}\,X=\frac{b}{\theta+1},$$

and find an expression for its variance.

(b) Let $X_1,\dots,X_n$ be independent draws from $f(x;\theta,b)$. Find a two-dimensional statistic $T=(T_1,T_2)$ that is sufficient for (θ,b). Is your T a minimal sufficient statistic?

(c) Provide expressions for the expectation and variance of both T_1 and T_2.

(d) Find expressions for the maximum likelihood estimators, say $\widehat{\theta}_{\rm ml}$ and $\widehat{b}_{\rm ml}$. Show that $\widehat{\theta}_{\rm ml}$ and $\widehat{b}_{\rm ml}$ are consistent for θ and b, respectively. *Hint:* Start with $\widehat{b}_{\rm ml}$.

We now assume that b is known and equals 1, and turn our attention to the median of the distribution, namely

$$\mu=(1/2)^{\theta}.$$

We will look at two different estimators of this quantity.

(e) Recall that if $U_1,\dots,U_n$ are independent uniforms on $(0,1)$, and $M_n=\mathrm{median}(U_1,\dots,U_n)$, then $\sqrt{n}(M_n-1/2)\to_d \mathrm{N}(0,1/4)$. Let $\widetilde{\mu}_n$ be the median of $X_1,\dots,X_n$. We will use $\widetilde{\mu}_n$ as an estimator of μ. Find an expression for its limiting variance.

(f) Find the maximum likelihood estimator of μ, say $\widehat{\mu}_{\rm ml}$, and find an expression for the mean of this estimator in terms of μ. Use this expression to show that $\widehat{\mu}_{\rm ml}$ is asymptotically unbiased, i.e,

$$\mathrm{E}\,\widehat{\mu}_{\rm ml}\to\mu,\quad \text{as } n\to\infty.$$

(g) Find the Cramér–Rao lower bound for unbiased estimators of μ. *Hint*: If $f(x;\theta)$ and $f(x;\mu)$ are two ways of parametrising a model, then the expected Fisher informations are related $I_\mu(\mu)=I_\theta(\theta)(\mathrm{d}\theta/\mathrm{d}\mu)^2$, given that $\theta=\theta(\mu)$ is continuously differentiable.

(h) Find the limiting distribution of $\sqrt{n}(\widehat{\mu}_{\rm ml}-\mu)$. Which estimator do you prefer, $\widetilde{\mu}_n$ or $\widehat{\mu}_{\rm ml}$, and why? Comment on this in view of what you found in (b).

Exercise 2

How does the long run become relevant to a particular set of data?, Sir David Cox asked and came up with an answer, but not all agreed (Reid, N. (1994) *A conversation with Sir David Cox*, Statistical Science, 9:439–455).

(a) We observe Y_1, Y_2 given by

$$Y_i = X_i\theta + (1 - X_i)U_i, \quad i = 1, 2,$$

where X_1, X_2 are independent Bernoulli(1/2), independent of U_1, U_2, which are independent uniforms on $(0, 1)$, and $0 < \theta < 1$ is an unknown parameter. Find an unbiased estimator of θ.

(b) Propose an estimator of θ, say $\delta(X)$, that allows you to say that $\delta(x) = \theta$ with a certain probability. Which of the two estimators of θ do you prefer, and why?

Suppose now that we are given n independent and unbiased measurements $Y_1, \ldots, Y_n$ of a quantity θ from one of two instruments, along with indicators $(a_1, \ldots, a_n)$ of which of the two instruments were used for the i'th measurement. Specifically, let $Y_i \mid a_i$, $i = 1, \ldots, n$ be independent $\mathrm{N}(\theta, \sigma^2_{a_i})$, where $a_1, \ldots, a_n$ are independent random variables with distribution $\Pr(a_i = 1) = \Pr(a_i = 0) = 1/2$. The variances σ_0^2 and σ_1^2 are known and unequal. Use $a = \sum_{i=1}^n a_i$ to denote the number of Y_i's in the sample with variance σ_1^2.

(c) Give an expression for the log-likelihood function based on having observed $(Y_1, a_1), \ldots, (Y_n, a_n)$. Find the maximum likelihood estimator $\widehat{\theta}$, and show that it is unbiased for θ.

(d) Find an expression for the variance of $\widehat{\theta}$ conditionally on the $(a_1, \ldots, a_n)$ you observed.

(e) Find the Cramér–Rao lower bound for unbiased estimators of θ. Comment on this in view of what you found in (d).

In the remainder of this exercise we assume that $a = \sum_{i=1}^n a_i = n/2 = 50$, and all we do will be conditional on the sequence $(a_1, \ldots, a_n)$ that was observed.

(f) It has come to your attention that instrument 0, instead of measuring θ, might be overshooting the target a bit and is measuring $\theta + \xi$, for some $\xi > 0$. Device a test to check if this might indeed be the case.

(g) It might be even worse. You suspect that the amount of overshooting may not be the same from measurement to measurement on instrument 0. That is, perhaps the data from instrument 0 are measurements of $\theta + \xi_1, \ldots, \theta + \xi_{50}$. Device a test to detect if the one or more of the ξ_i's differ from the others.

Exercise 3

For this exercise you will need the following prior to posterior result: If we observe data that are $X \mid \theta \sim \mathrm{N}(\theta, \sigma^2)$ with a mean that is $\theta \sim \mathrm{N}(0, \tau^2)$, then

$$\theta \mid (X = x) \sim \mathrm{N}(m, v^2), \quad \text{with} \quad m = \frac{\tau^2 x}{\sigma^2 + \tau^2}, \quad \text{and} \quad v^2 = \frac{\tau^2\sigma^2}{\sigma^2 + \tau^2}.$$

We also recall that $\mathrm{E}\,X = \mathrm{E}\,\mathrm{E}\,(X \mid Y)$ and that $\mathrm{Var}\,X = \mathrm{E}\,\mathrm{Var}\,(X \mid Y) + \mathrm{Var}\,\mathrm{E}\,(X \mid Y)$.

Consider the regression model

$$Y_i = \beta x_i + \varepsilon_i, \quad i = 1, \ldots, n, \tag{1}$$

where $\varepsilon_1, \ldots, \varepsilon_n$ independent standard normal random variables, and $x_1, \ldots, x_n$ are fixed and known constants. We are to estimate β under squared error loss $L(\delta, \beta) = (\delta - \beta)^2$.

(a) Find an expression for the maximum likelihood estimator, say $\widehat{\beta}_{\rm ml}$, and write down its distribution. Provide also an expression for its risk function.

(b) Find the Cramér–Rao lower bound for unbiased estimators of β, and comment on what you find.

(c) Then place a normal prior with mean 0 and variance σ_0^2 over β. First explain why the posteriors $\beta \mid (Y_1 = y_1, \ldots, Y_n = y_n)$ and $\beta \mid \widehat{\beta}_{\rm ml}$ must be the same, then find the posterior.

(d) Give an expression for the Bayes estimator, $\widetilde{\beta}_{\rm b}$ say. Indicate how you would use $\widetilde{\beta}_{\rm b}$ to show that the maximum likelihood estimator $\widehat{\beta}_{\rm ml}$ is admissible. Is it also minimax; why or why not?

(e) Find the risk function of the Bayes estimator, and compare it to the risk function of the maximum likelihood estimator in the point $\beta = 0$.

It has come to your attention that the model in display (1) has been found unsatisfactory due to the omission of one unobserved covariate inducing a correlation between the x_i's and the ε_i's; a so called confounder. For such a thing to make sense statistically modelling wise, the covariates $x_1, \ldots, x_n$ need to be random variables. Assume therefore that

$$\begin{pmatrix} x_i \\ \varepsilon_i \end{pmatrix} \sim {\rm N}_2\big(0, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\big). \tag{2}$$

independently for $i = 1, \ldots, n$. Using results from the curriculum one might show that ${\rm E}\,(\varepsilon_i \mid x_i) = \rho x_i$ and ${\rm E}\,(x_i \mid \varepsilon_i) = \rho \varepsilon_i$, and that ${\rm Var}(\varepsilon_i \mid x_i) = {\rm Var}(x_i \mid \varepsilon_i) = 1 - \rho^2$.

(f) A colleague tells you that the ρ is known to be $1/2$. Find an estimator that is unbiased for β, given that this supposition is correct.

(g) Still under the model in display (1) and the relation between the covariates and the noise described in display (2), show that the estimator you found in (e) is consistent, and provide a 95% confidence interval for β.