Norwegian University of Science and Technology
Department of Mathematical Sciences

ENGLISH

Contact during exam:
Arvid Naess 73 59 70 53/ 99 53 83 50

SOLUTIONS EXAM IN COURSE TMA4295 STATISTICAL INFERENCE

December 7, 2009
Time: 09:00–13:00

Permitted aids: *Tabeller og formler i statistikk*, Tapir Forlag
K. Rottmann: *Matematisk formelsamling*
Calculator HP30S / CITIZEN SR-270X
Yellow, stamped A5-sheet with your own handwritten notes.

Problem 1

In this problem we model the lifetime (time to first failure) of a specific type of equipment as a random variable $X \sim \text{Exp}(1/\theta)$, that is, exponentially distributed, where θ is unknown ($0 < \theta < \infty$). The probability density function (pdf) of X: $f_X(x) = \theta \exp(-\theta x)$ for $x \geq 0$; $=$ 0 otherwise.

Suppose that n "identical" pieces of equipment are tested and the failure times $x_1, \ldots, x_n$ are observed. These data are considered as an outcome of the random sample $X_1, \ldots, X_n$, which are then iid $\text{Exp}(1/\theta)$. We want to estimate the probability of early failure of this type of equipment, that is, $P_\theta(X_1 \leq x) = 1 - e^{-\theta x}$, for some fixed x.

a) Show that $\text{Exp}(1/\theta)$ is an exponential family, and that $T = \sum_{i=1}^n X_i$ is a complete, sufficient estimator of θ.

SOLUTION:

The pdf of $X \sim \text{Exp}(1/\theta)$ is given as (for $x \geq 0$) $f_X(x|\theta) = \theta e^{-\theta x} = h(x)c(\theta)\exp\big(w(\theta)t(x)\big)$ with $h(x) = 1$, $c(\theta) = \theta$, $w(\theta) = -\theta$ and $t(x) = x$. Hence, $f_X(x|\theta)$ is a member of an exponetial family.

The joint sample pdf equals $f_{\mathbf{X}}(\mathbf{x}|\theta) = \prod_{i=1}^n f_X(x_i) = \theta^n \exp\big(-\theta\sum_{i=1}^n x_i\big)$. By the Factorization Theorem it follow that $T(\mathbf{X}) = \sum_{i=1}^n X_i$ is a sufficient statistic for θ. Since

$w(\theta) = -\theta$, and the parameter set of θ is the open, infinite interval $(0, \infty)$, the set of values of $w(\theta)$ is the open set $(-\infty, 0)$. Hence, T is also a complete statistic by the criterion for complete statistics in exponential families.

b) Show that $S(X_1) = 1_{[X_1 \leq x]}$ is an unbiased estimator of $\tau(\theta) = 1 - e^{-\theta x}$. ($1_A$ denotes the indicator function of the event A.)

Use $S(X_1)$ and T from part a) to construct a uniform minimum variance unbiased estimator (UMVUE) Q^* of $\tau(\theta)$.

SOLUTION:

$$\mathrm{E}_\theta[S(X_1)] = 1 \cdot P_\theta(X_1 \leq x) + 0 \cdot P_\theta(X_1 > x) = P_\theta(X_1 \leq x) = 1 - e^{-\theta x}\,,$$

which makes $S(X_1)$ an unbiased estimator of $\tau(\theta)$.

Since $S(X_1)$ is an unbiased estimator of $\tau(\theta)$ and T is a sufficient estimator of θ, then $Q^* = \mathrm{E}[S(X_1)|T]$ becomes an UMVUE of $\tau(\theta)$ by Rao-Blackwell's theorem. Hence, for an observed value t of T, the corresponding observed value q^* of Q^* is given as $q^* = \mathrm{E}[S(X_1)|T = t] = P(X_1 \leq x|T = t)$.

To calculate the explicit expression for Q^* in terms of $X_1, \ldots, X_n$, you may proceed as follows (or use another method, if you prefer).

c) First show that $V = X_1/T$ is independent of T, and that the probability density function of V is given as,

$$f_V(v) = (n-1)\,(1-v)^{n-2} \text{ for } 0 < v < 1\,,$$

and $f_V(v) = 0$ elsewhere.

Hint: From the two random variables X_1 and $\tilde{X}_2 = \sum_{j=2}^{n} X_j$, (T, V) is obtained by transformation of $(X_1, \tilde{X}_2)$.

SOLUTION:

From the properties of $X_1, \ldots, X_n$ it follows that X_1 and $\tilde{X}_2$ are independent and $\tilde{X}_2 \sim$ Gamma$(n-1, 1/\theta)$. The joint pdf of X_1 and $\tilde{X}_2$ is then given as,

$$f_{X_1\tilde{X}_2}(x, y) = \theta e^{-\theta x}\,\frac{\theta^{n-1} y^{n-2}}{\Gamma(n-1)} e^{-\theta y} = \frac{\theta^n y^{n-2}}{\Gamma(n-1)} e^{-\theta(x+y)}\,, \text{ for } x > 0, y > 0\,.$$

The transformation $(X_1, \tilde{X}_2) \to (T, V)$ is obtained by $T = X_1 + \tilde{X}_2$ and $V = X_1/(X_1 + \tilde{X}_2)$. This leads to the following relation for the pdf of (T, V),

$$f_{TV}(t, v) = |J| f_{X_1\tilde{X}_2}(tv, t - tv) = t\,\frac{\theta^n (t - tv)^{n-2}}{\Gamma(n-1)} e^{-\theta t}\,, \text{ for } t > 0, 0 < v < 1\,,$$

where the Jacobian J of the transformation is calculated to be t. Rewriting this gives,

$$f_{TV}(t,v) = \frac{\theta^n t^{n-1}}{\Gamma(n)} e^{-\theta t} \cdot (n-1)(1-v)^{n-2}\,, \text{ for } t > 0, 0 < v < 1\,.$$

The first factor on the right hand side is recognized as the pdf of T, which is a Gamma$(n, 1/\theta)$ variable. Since the second factor does not depend on t, it follows that T and V are independent random variables, and the pdf of V is given as,

$$f_V(v) = (n-1)(1-v)^{n-2}\,, \text{ for } 0 < v < 1\,.$$

d) Prove that,

$$P(X_1 \le x | T = t) = P(V \le \frac{x}{t})\,.$$

Then show that,

$$P(V \le \frac{x}{t}) = 1 - \left(1 - \frac{x}{t}\right)^{n-1}, \quad t > x\,;$$
$$= 1\,, \quad t \le x\,.$$

Finally, write down the expression for Q^* in terms of $X_1, \ldots, X_n$.

SOLUTION:

Given that $T = t$, we may write

$$P(X_1 \le x | T = t) = P(\frac{X_1}{T} \le \frac{x}{t} | T = t) = P(V \le \frac{x}{t} | T = t) = P(V \le \frac{x}{t})\,.$$

The last equation follows since V is independent of T.

From the expression for $f_V(v)$ it follows that $F_V(v) = P(V \le v) = 1 - (1-v)^{n-1}$ for $0 < v < 1$ and $F_V(v) = P(V \le v) = 1$ for $v \ge 1$, which leads to the desired formula $P(V \le \frac{x}{t}) = 1 - \left(1 - \frac{x}{t}\right)^{n-1}$ for $t > x$ and $P(V \le \frac{x}{t}) = 1$ for $t \le x$. The expression for Q^* is then found to be gives as,

$$Q^* = 1 - \left(1 - \frac{x}{T}\right)^{n-1} = 1 - \left(1 - \frac{x}{\sum_{i=1}^n X_i}\right)^{n-1} \text{ for } \sum_{i=1}^n X_i > x\,,$$

and $Q^* = 1$ for $\sum_{i=1}^n X_i \le x$.

e) Show that the maximum likelihood estimator (MLE) $\hat{Q}$ of $\tau(\theta)$ based on the random sample $X_1, \ldots, X_n$ is given as $\hat{Q} = 1 - \exp(-x/\overline{X})$, where $\overline{X} = \sum_{i=1}^n X_i/n$.

SOLUTION:

By the invariance principle for MLE, it follows that $\hat{Q} = \tau(\hat{\Theta})$, where $\hat{\Theta}$ is the MLE of θ. The log-likelihood function $\ell(\theta|\mathbf{x})$ for θ is given as

$$\ell(\theta|\mathbf{x}) = n \ln \theta - \theta \sum_{i=1}^n x_i$$

The MLE $\hat{\theta}$ is found by solving the equation,

$$0 = \partial \ell(\theta|\mathbf{x})/\partial\theta = \frac{n}{\theta} - \sum_{i=1}^n x_i \,,$$

which gives the solution

$$\hat{\theta} = \frac{n}{\sum_{i=1}^n x_i} \,.$$

Hence,

$$\hat{Q} = 1 - \exp(-nx/\sum_{i=1}^n X_i) = 1 - \exp(-x/\overline{X}) \,.$$

f) Verify that $\mathrm{E}[\hat{Q}] > \mathrm{E}[Q^*]$ by using Jensen's inequality, or any other suitable method. What can you say about the asymptotic behaviour of $\hat{Q}$ relative to Q^*? (You may assume that the requisite regularity conditions are satisfied.)

Jensen's inequality: If $g(\cdot)$ is a strictly convex function on the value range of a non-constant random variable $T \geq 0$, $0 < \mathrm{E}[T] < \infty$, then $\mathrm{E}[g(T)] > g(\mathrm{E}[T])$. (Note: $g(t) = 1 - \exp(-nx/t)$ is a strictly convex function on $(0, \infty)$.)

SOLUTION:

Invoking Jensen's inequality for the strictly convex function $g(t) = 1 - \exp(-nx/t)$ and the random variable $T = \sum_{i=1}^n X_i$, it follows that $\mathrm{E}[\hat{Q}] = \mathrm{E}[1 - \exp(-nx/T)] > 1 - \exp(-nx/\mathrm{E}[T])$. Now, $\mathrm{E}[T] = n\mathrm{E}[X_1] = n/\theta$, which gives $\mathrm{E}[\hat{Q}] > 1 - \exp(-\theta x) = \mathrm{E}[Q^*]$.

From the asymptotic properties of the MLE, it follows that $\mathrm{E}[\hat{Q}] \to \mathrm{E}[Q^*]$ and $\mathrm{Var}[\hat{Q}] \to \mathrm{Var}[Q^*]$ as $n \to \infty$, that is, $\hat{Q}$ is asymptotically unbiased and efficient.

Problem 2

Let $X_1, \ldots, X_n$ be a random sample from the uniform distribution Unif$(\theta, \theta+1)$, where θ is unknown. If $X \sim$ Unif$(\theta, \theta+1)$, then the pdf $f_0(x|\theta)$ of X is $f_0(x|\theta) = 1$ for $\theta \le x \le 1+\theta$; $=$ 0 otherwise.

a) Show that the joint pdf of $\mathbf{X} = (X_1, \ldots, X_n)$ can be written as,

$$f(\mathbf{x}|\theta) = 1\,, \max(x_1, \ldots, x_n) - 1 \le \theta \le \min(x_1, \ldots, x_n)\,;$$
$$= 0\,, \text{ otherwise}\,.$$

Then prove that $T(\mathbf{X}) = (Y_1, Y_n)$, where $Y_1 = \min(X_1, \ldots, X_n)$ and $Y_n = \max(X_1, \ldots, X_n)$, is a minimal sufficient statistic for θ.

What is the MLE of θ?

SOLUTION:

Let $f_0(x|\theta)$ denote the pdf of a Unif$(\theta, \theta+1)$ variable. Then $f_0(x|\theta) = 1$ for $\theta \le x \le 1+\theta$, $f_0(x|\theta) = 0$ otherwise. Hence, $f(\mathbf{x}|\theta) = 1$ for $\theta \le x_i \le 1+\theta$ for every $i = 1, \ldots, n$, and $f(\mathbf{x}|\theta) = 0$ otherwise. This gives that $f(\mathbf{x}|\theta) = 1$ for $\theta \le \min(x_1, \ldots, x_n)$ and $\max(x_1, \ldots, x_n) \le 1+\theta$, and $f(\mathbf{x}|\theta) = 0$ otherwise. Summing up:

$$f(\mathbf{x}|\theta) = 1\,, \max(x_1, \ldots, x_n) - 1 \le \theta \le \min(x_1, \ldots, x_n)\,;$$
$$= 0\,, \text{ otherwise}\,.$$

To verify that $T(\mathbf{X}) = (Y_1, Y_n)$ is a minimal sufficient statistic, we shall verify that the ratio $f(\mathbf{x}|\theta)/f(\mathbf{y}|\theta)$ is constant as a function of θ if and only if $T(\mathbf{x}) = T(\mathbf{y})$. It is seen that $f(\mathbf{x}|\theta)$ and $f(\mathbf{y}|\theta)$ will both be positive if and only if $\min(x_1, \ldots, x_n) = \min(y_1, \ldots, y_n)$ and $\max(x_1, \ldots, x_n) = \max(y_1, \ldots, y_n)$. But then they are both equal to 1, so the ratio is constant. Hence, $T(\mathbf{X}) = (Y_1, Y_n)$ is indeed a minimal sufficient statistic for θ.

The likelihood function for θ is given as $L(\theta|\mathbf{x}) = f(\mathbf{x}|\theta) = 1$ for $\max(x_1, \ldots, x_n) - 1 \le \theta \le \min(x_1, \ldots, x_n)$, and $= 0$ otherwise. Therefore, any $\hat{\theta}$ satisfying $\max(x_1, \ldots, x_n) - 1 \le \hat{\theta} \le \min(x_1, \ldots, x_n)$ will serve as a maximum likelihood estimate of θ. Hence, any random variable $\hat{\Theta}$ satisfying the inequality

$$\max(X_1, \ldots, X_n) - 1 \le \hat{\Theta} \le \min(X_1, \ldots, X_n)$$

will be a MLE. One example could be

$$\hat{\Theta} = \frac{1}{2}\big(\min(X_1, \ldots, X_n) + \max(X_1, \ldots, X_n) - 1\big)\,,$$

and for this choice it turns out that $\mathrm{E}[\hat{\Theta}] = \theta$.

b) Show that the joint cumulative distribution function (cdf) $F_{Y_1Y_n}(y_1, y_n)$ of Y_1 and Y_n is given (partly) as follows,

$$F_{Y_1Y_n}(y_1, y_n) = (y_n - \theta)^n - (y_n - y_1)^n\,,\ \ \theta \leq y_1 < y_n \leq 1 + \theta$$

<u>SOLUTION:</u>

The CDF of a Unif$(\theta, \theta + 1)$ variable is given as

$$\begin{aligned} &= 0\,,\ \ x < \theta \\ F_X(x) &= x - \theta\,,\ \ \theta \leq x \leq 1 + \theta \\ &= 1\,,\ \ x > 1 + \theta\,. \end{aligned}$$

For $\theta \leq y_1 < y_n \leq 1 + \theta$,

$$\begin{aligned} F_{Y_1Y_n}(y_1, y_n) &= P(Y_1 \leq y_1, Y_n \leq y_n) = P(Y_n \leq y_n) - P(Y_1 > y_1, Y_n \leq y_n) \\ &= P(X_1 \leq y_n, \ldots, X_n \leq y_n) - P(y_1 < X_1 \leq y_n, \ldots, y_1 < X_n \leq y_n) \\ &= \prod_{i=1}^{n} P(X_i \leq y_n) - \prod_{i=1}^{n} P(y_1 < X_i \leq y_n) \\ &= (y_n - \theta)^n - (y_n - y_1)^n \end{aligned}$$

The following hypothesis is to be tested:

$$H_0 : \theta = 0$$
$$H_1 : \theta > 0$$

With a rejection region given as,

$$R = \{\mathbf{x} | y_1 = \min(x_1, \ldots, x_n) \geq k \text{ or } y_n = \max(x_1, \ldots, x_n) \geq 1\}$$

for a suitable constant $k \geq 0$.

c) Show that the power function $\beta(\theta) = P_\theta(\mathbf{X} \in R)$ for this test can be expressed as,

$$\begin{aligned} \beta(\theta) &= 1 - (1 - \theta)^n + (1 - k)^n\,,\ \ \theta \leq k < 1\,, \\ &= 1\,,\ \ 0 \leq k < \theta\,. \end{aligned}$$

SOLUTION:

The power function is,

$$\begin{aligned}\beta(\theta) &= P(Y_1 \geq k \text{ or } Y_n \geq 1) = 1 - F_{Y_1 Y_n}(k, 1)\\ &= 1 - (1-\theta)^n + (1-k)^n\,.\end{aligned}$$

This holds for $\theta \leq k < 1$. Since $Y_1 \geq \theta$, it follows that for $k < \theta$, the event $Y_1 \geq k$ is the certain event, that is, $P(Y_1 \geq k) = 1$. But then also $\beta(\theta) = 1$.

d) Find values of n and k so that a size 0.10 test will have power at least 0.8 if $\theta \geq 0.2$.

SOLUTION:

A size 0.10 test requires that $\beta(0) = 0.10$. This leads to the equation $(1-k)^n = 0.10$.

The requirement of power at least 0.8 if $\theta \geq 0.2$ leads to $\beta(0.2) \geq 0.8$, that is $1 - (1 - 0.2)^n + (1-k)^n = 1 - 0.8^n + 0.10 \geq 0.8$. This leads to $0.8^n \leq 0.3$. The smallest integer satisfying this inequality is $n = 6$. Hence, $(1-k)^6 = 0.10$, which gives $k = 0.32$.