Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

LØSNINGSFORSLAG
EXAM IN TMA4295 STATISTICAL INFERENCE
Friday 19 May 2006
Time: 09:00–13:00

Oppgave 1

Suppose that $X_1, ..., X_n$ are iid $Poisson(\theta)$.

a) Find MLE of $(1+\theta)e^{-\theta}$.

Solution. MLE of θ is $\bar{X}$, therefore, due to the invariance principle

$$T_{MLE} = (1+\bar{X})e^{-\bar{X}}.$$

b) Find the best unbiased estimator of $(1+\theta)e^{-\theta}$.

Solution. $S = \sum_{i=1}^n X_i$ is a complete sufficient statistic. Set

$$T = \begin{cases} 1 & \text{if } X_1 = 0 \text{ or } X_1 = 1, \\ 0 & \text{otherwise.} \end{cases}$$

T is an unbiased estimator of $(1+\theta)e^{-\theta}$ therefore $E(T|S)$ is the best unbiased. For any $m = 0, 1, ...$

$$E(T|S=m) = P(T=1|S=m) = P(X_1=0|S=m) + P(X_1=1|S=m) =$$
$$= \frac{P(X_1=0, S=m)}{P(S=m)} + \frac{P(X_1=1, S=m)}{P(S=m)} = \frac{P(X_1=0, \sum_2^n X_i = m)}{P(S=m)} +$$
$$+ \frac{P(X_1=0, \sum_2^n X_i = m-1)}{P(S=m)} = \left(\frac{n-1}{n}\right)^m \left(1 + \frac{m}{n-1}\right).$$

Thus

$$T_{BUE} = \left(\frac{n-1}{n}\right)^S \left(1 + \frac{S}{n-1}\right).$$

c) Using a comparison of these two estimators show that MLE is biased. (*Hint:* note that both estimators are functions of a complete sufficient statistic.)

Solution. It follows from the Rao-Blackwell theorem and the uniqueness of the best unbiased estimator that for any function of the parameter there can be only one unbiased estimator which is a function of a complete sufficient statistic (If S is a complete sufficient statistic, and $T_1 = f_1(S), T_2 = f_2(S), ET_1 = \tau(\theta), ET_2 = \tau(\theta)$, then $0 = E(T_1 - T_2) = E(f_1(S) - f_2(S))$ and therefore $f_1(S) = f_2(S)$ a.s.). Both T_{MLE} and T_{BUE} are functions of $S = \sum_{i=1}^n X_i$, a complete sufficient statistic. It is easy to see that the two estimator do not coincide, therefore, since T_{BUE} is unbiased, T_{MLE} is biased.

Oppgave 2

Let $X_1, ..., X_n$ be iid from a distribution with pmf

$$\left(\frac{\theta}{2}\right)^{|x|}(1-\theta)^{1-|x|}, \quad x = -1, 0, 1, \quad 0 < \theta < 1.$$

Suppose that n is large enough so that the Central Limit Theorem can be used.

a) For testing $H_0 : \theta \le \theta_0$ versus $H_1 : \theta > \theta_0$ find an (approximate) level α UMP test.

Solution. The likelihood function is

$$L(\theta; X) = \left(\frac{\theta}{2}\right)^{\sum|X_i|}(1-\theta)^{n-\sum|X_i|},$$

therefore, if $\theta' < \theta''$, then the ratio

$$\frac{L(\theta'; X)}{L(\theta''; X)} = \left(\frac{1-\theta'}{1-\theta''}\right)^n \left[\frac{\theta'(1-\theta'')}{\theta''(1-\theta')}\right]^{\sum|X_i|}$$

is a monotone (decreasing) function of $T(X) = \sum|X_i|$. Therefore the UMP test has form

$$\sum_{i=1}^n |X_i| > c \Longrightarrow H_1$$

where c is determined from condition

$$P_{\theta_0}(\sum|X_i| > c) = \alpha.$$

To find c let us use CLT. We have $E|X_i| = \theta$, $Var(|X_i|) = \theta(1-\theta)$ therefore

$$\alpha = P_{\theta_0}(\sum|X_i| > c) = P_{\theta_0}\left(\frac{\sum|X_i| - n\theta_0}{\sqrt{n\theta_0(1-\theta_0)}} > \frac{c - n\theta_0}{\sqrt{n\theta_0(1-\theta_0)}}\right) \approx$$

$$\approx 1 - \Phi\left(\frac{c - n\theta_0}{\sqrt{n\theta_0(1-\theta_0)}}\right)$$

and

$$c = n\theta_0 + \sqrt{n\theta_0(1-\theta_0)}z_{1-\alpha}.$$

b) For the specific case $\theta_0 = 1/3$, $\alpha = 0.05$ determine the sample size n for which the probability of the Type II error for $\theta = 2/3$ is no greater than 0.0001.

5**Solution.** The power function is

$$\pi(\theta) = P_\theta(\sum |X_i| > c) =$$

$$= P_\theta\left(\frac{\sum |X_i| - n\theta}{\sqrt{n\theta(1-\theta)}} > \frac{n(\theta_0 - \theta) + \sqrt{n\theta_0(1-\theta_0)}z_{1-\alpha}}{\sqrt{n\theta(1-\theta)}}\right) \approx$$

$$\approx 1 - \Phi\left(\frac{\sqrt{n}(\theta_0 - \theta) + \sqrt{\theta_0(1-\theta_0)}z_{1-\alpha}}{\sqrt{\theta(1-\theta)}}\right),$$

therefore condition $1 - \pi(1/3) < 0.0001$ is equivalent to

$$n > 55.$$

c) Prove that there does not exist a level α UMP test of $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$, $0 < \alpha < 1$.

Solution. Suppose it exists. Denote C its critical region. Consider two values θ_1 and θ_2 such that $\theta_1 < \theta_0 < \theta_2$. Then

$$P_{\theta_0}(X \in C) = \alpha$$

and

$$P_{\theta_1}(X \in C) \geq P_{\theta_1}(X \in C'), \quad P_{\theta_2}(X \in C) \geq P_{\theta_2}(X \in C')$$

for any C' such that

$$P_{\theta_0}(X \in C') \leq \alpha$$

i.e. C is the most powerful level α test for both problems (a) $H_0 : \theta = \theta_0,\ H_1 : \theta = \theta_1$ and (b) $H_0 : \theta = \theta_0,\ H_1 : \theta = \theta_2$. Due to the Neyman-Pearson Lemma, this means that C is NPT (Neyman-Pearson test) for problem (a) and for problem (b). But NPT for (a) has form $\bar{X} < t' \Rightarrow H_1$ while for (b) it has form $\bar{X} > t'' \Rightarrow H_1$. Contradiction.

Oppgave 3

Let X be one observation from a distribution with pdf $\theta x^{\theta-1}$, $0 < x < 1$, $\theta > 0$.

a) Prove that X^θ is a pivotal quantity. Find its distribution.

Solution. X^θ has the uniform $(0,1)$ distribution (this is found either directly or using theory – distribtions of functions of random variables).

b) Using this pivot construct a $(1-\alpha)$ confidence interval for θ, $0 < \alpha < 1$.

Solution. Let $\alpha_1 + \alpha_2 = \alpha$, $(\alpha_1 > 0, \alpha_2 > 0)$. Then

$$\alpha = P\left(\alpha_1 \le X^\theta \le 1 - \alpha_2\right) = P\left(\frac{\ln(1/(1-\alpha_2))}{\ln(1/X)} \le \theta \le \frac{\ln(1/\alpha_1)}{\ln(1/X)}\right)$$

therefore each interval

$$\left[\frac{\ln(1/(1-\alpha_2))}{\ln(1/X)}, \frac{\ln(1/\alpha_1)}{\ln(1/X)}\right]$$

is a $(1-\alpha)$ confidence interval for θ.