UPPSALA UNIVERSITY
Department of Mathematics
Rolf Larsson

Exam in Mathematical Statistics
Inference Theory II, 1MS037
2021–06–08
Solutions

1. Suppose $X_1, ..., X_n$ are independent, distributed as X which has density function

$$f(x; \alpha) = \frac{2\alpha x}{(1 + x^2)^{\alpha+1}},$$

for $x > 0$, and 0 otherwise.

(a) Does this distribution belong to an exponential family?
If so, which is the natural parameter? (2p)

Solution: We may write the density function as

$$f(x; \alpha) = 2\alpha \exp\left\{-\alpha \log(1 + x^2)\right\} \frac{x}{1 + x^2}.$$

This is on the exponential family form, and so, we see that the distribution belongs to the exponential family. The natural parameter is $-\alpha$.

(b) Find a sufficient statistic for α. (2p)

Solution: From the expression in (a), the likelihood is

$$L(\alpha) = \prod_{i=1}^{n} f(x_i; \alpha) = (2\alpha)^n \exp\left\{-\alpha \sum_{i=1}^{n} \log(1 + x_i^2)\right\} \prod_{i=1}^{n} \frac{x_i}{1 + x_i^2}.$$

Hence, we find that the sufficient statistic is $T(\mathbf{x}) = \sum_{i=1}^{n} \log(1 + x_i^2)$, where $\mathbf{x} = (x_1, ..., x_n)$, the vector of observed values.

(c) Which is the smallest variance that an unbiased estimator of α can attain? (2p)

Solution: From the Cramér-Rao inequality, this variance is given by $1/I(\alpha)$, where $I(\alpha)$ is the Fisher information.
To calculate $I(\alpha)$, we at first write down the log likelihood:

$$l(\alpha) = \log L(\alpha) = C + n \log \alpha - \alpha T(\mathbf{x}),$$

where C is a constant and $T(\mathbf{x})$ is as above. This gives the derivatives

$$\begin{aligned} l'(\alpha) &= n\alpha^{-1} - T(\mathbf{x}), \\ l''(\alpha) &= -n\alpha^{-2}. \end{aligned}$$

The Fisher information is given by $I(\alpha) = -E(l''(\alpha)) = n\alpha^{-2}$, and so, the smallest possible variance that an unbiased estimator can attain is $1/I(\alpha) = \alpha^2/n$.

2. Suppose $X_1, ..., X_n$ are independent and normally distributed with expectation μ and variance σ^2. Let $\theta = (\mu, \sigma^2)$.

   Which of the following statistics are sufficient for θ, and which are also minimal sufficient for θ? Motivate your answer. (6p)

   (a) $\sum_{i=1}^n X_i$

   (b) $(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$

   (c) $(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^3)$

   (d) $(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2, \sum_{i=1}^n X_i^3)$

   Solution: Let the observations be $x_1, ..., x_n$. The likelihood is

$$L(\mu, \sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x_i-\mu)^2}{2\sigma^2}\right\} = (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{\sum_{i=1}^n (x_i-\mu)^2}{2\sigma^2}\right\}$$
$$= (2\pi\sigma^2)^{-n/2} \exp\left(-n\frac{\mu^2}{2\sigma^2}\right) \exp\left(\frac{\mu}{\sigma^2}\sum_{i=1}^n x_i - \frac{1}{\sigma^2}\sum_{i=1}^n x_i^2\right).$$

   Hence, by the factorization criterion, we find that $(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ is a sufficient statistic for $\theta = (\mu, \sigma^2)$.

   We also find that $\sum_{i=1}^n X_i$ does not contain enough information to be sufficient for θ. Nor does $(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^3)$, because we can not get $(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ as a function of this statistic. This is, however, possible for $(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2, \sum_{i=1}^n X_i^3)$, so this statistic is sufficient.

   Because we have a strictly 2-parametric exponential family (this can be seen from the likelihood), we find by theorem 3.9 that $(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ is also minimal sufficient. (Alternatively, see special case 3.13.) However, $(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2, \sum_{i=1}^n X_i^3)$ is not a function of $(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$, hence by definition, it can not be minimal sufficient.

   Conclusively, the answer is (a)- not sufficient, (b)- sufficient and minimal sufficient, (c)- not sufficient and (d)- sufficient but not minimal sufficient.

3. Suppose $X_1, ..., X_n$ are independent, distributed as X which is discrete with probability function

$$p(x;\theta) = \begin{cases} \frac{1}{4+\theta} & \text{if } x = 0, \\ \frac{3}{4+\theta} & \text{if } x = 1, \\ \frac{\theta}{4+\theta} & \text{if } x = 2, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Calculate the maximum likelihood estimate (MLE) of θ. (2p)

Solution: Let n_0, n_1, n_2 be the number of zeros, ones and twos in the sample, respectively. Then, the likelihood is

$$L(\theta) = \left(\frac{1}{4+\theta}\right)^{n_0} \left(\frac{3}{4+\theta}\right)^{n_1} \left(\frac{\theta}{4+\theta}\right)^{n_2} = 3^{n_1}\theta^{n_2}(4+\theta)^{-n},$$

where $n = n_0 + n_1 + n_2$. This gives us the log likelihood

$$l(\theta) = \log L(\theta) = n_1 \log 3 + n_2 \log\theta - n\log(4+\theta),$$

with derivatives

$$l'(\theta) = \frac{n_2}{\theta} - \frac{n}{4+\theta},$$
$$l''(\theta) = -\frac{n_2}{\theta^2} + \frac{n}{(4+\theta)^2}.$$

Solving $l'(\hat{\theta}) = 0$ yields

$$\hat{\theta} = \frac{4n_2}{n-n_2} = \frac{4n_2}{n_0+n_1},$$

and moreover, it follows that

$$l''(\hat{\theta}) = -\frac{(n_0+n_1)^3}{16nn_2} < 0,$$

so we have a maximum.

(b) Assume that the observations are $2, 0, 2, 2$. Consider testing H_0: $\theta = 4$ vs H_1: $\theta > 4$, using the MLE as test statistic.
Calculate the p value. (3p)

Solution: Since the MLE is monotonely increasing in n_2, it is equivalent to use n_2 as test statistic. With $n = 4$ and $\theta = 4$, $n_2 \sim \text{Bin}(4, 1/2)$. The observed value of n_2 is 3, and we should reject for large n_2. Hence, the p value is

$$P_{H_0}(n_2 \geq 3) = P_{H_0}(n_2 = 3) + P_{H_0}(n_2 = 4) = \frac{5}{16} = 0.3125.$$

4. Suppose X_1 and X_2 are independent and Poisson with parameter λ.

(a) Show that $S = X_1 + X_2$ is a sufficient statistic for λ. (2p)

Solution: Let the observations be x_1 and x_2. The likelihood is

$$L(\lambda) = \frac{\lambda^{x_1}}{x_1!}e^{-\lambda}\frac{\lambda^{x_2}}{x_2!}e^{-\lambda} = \frac{1}{x_1!x_2!}\lambda^{x_1+x_2}e^{-2\lambda},$$

and it follows from the factorization theorem that $S = X_1+X_2$ is sufficient for λ.

(b) Show that the statistic $T = X_1$ is unbiased for λ. (1p)

Solution: This follows because $E(T) = E(X_1) = \lambda$.

(c) Use the Rao-Blackwell theorem to find an unbiased estimator of λ with smaller variance than the variance of T. (3p)

Solution: By the Rao-Blackwell theorem, the statistic $E(T|S)$ has smaller variance. Since $S \sim \text{Po}(2\lambda)$, we get

$$\begin{aligned} P(T = k|S = s) &= \frac{P(X_1 = k, X_2 = s-k)}{P(S = s)} = \frac{\frac{\lambda^k}{k!}e^{-\lambda}\frac{\lambda^{s-k}}{(s-k)!}e^{-\lambda}}{\frac{(2\lambda)^s}{s!}e^{-2\lambda}} \\ &= 2^{-s}\frac{s!}{k!(s-k)!} = \binom{s}{k}2^{-s}, \end{aligned}$$

and it follows that $(T|S = s) \sim \text{Bin}(s, 1/2)$. Hence,

$$E(T|S) = \frac{S}{2} = \frac{X_1 + X_2}{2} = \bar{X}$$

is a statistic with smaller variance than the variance of T.

5. Consider testing that the observation x comes from a discrete distribution with probability function $p_0(x)$ vs the alternative that it comes from a discrete distribution with probability function $p_1(x)$, where these two probability functions are given in the following table:

x	1	2	3	4	5
$p_0(x)$	0.3	0.2	0.2	0.2	0.1
$p_1(x)$	0.1	0.1	0.2	0.3	0.3

(a) Which is the most powerful (MP) test at level $\alpha = 0.2$? (2p)

Solution: We use the Neyman-Pearson lemma. The most powrful test rejects for the smallest values of $p_0(x)/p_1(x)$. These values are given in the following extended table:

x	1	2	3	4	5
$p_0(x)$	0.3	0.2	0.2	0.2	0.1
$p_1(x)$	0.1	0.1	0.2	0.3	0.3
$p_0(x)/p_1(x)$	3	2	1	0.67	0.33

The smallest value occurs for $x = 5$, with probability 0.1 under H_0. Hence, at level 0.2, we should reject for $x = 5$. The next to smallest value occurs for $x = 4$. Now, the probability to get $x \in \{4, 5\}$ is $0.1 + 0.2 = 0.3 > 0.2$, so we need to randomize. If we reject for $x = 4$ with probability 0.5, we get the correct size of 0.2. Hence, the MP test has test function

$$\varphi(x) = \begin{cases} 1 & \text{if } x = 5, \\ 0.5 & \text{if } x = 4, \\ 0 & \text{otherwise.} \end{cases}$$

(b) Calculate the size of the type II error and the power for the MP test.(2p)

Solution: The power $1 - \beta$, where β is the type II error, equals the expectation of the test function under H_1. In our case, this is

$$1 - \beta = 1 * p_1(5) + 0.5 * p_1(4) = 1 * 0.3 + 0.5 * 0.3 = 0.45.$$

Hence, the size of the type II error is $\beta = 1 - 0.45 = 0.55$.

(c) Calculate sizes of the errors of type I and II as well as the power for the test with critical region $\{x = 4\}$. Compare to the power for the MP test. (2p)

Solution: The size of the error of type I for this test is $\alpha = p_0(4) = 0.2$, hence the same size as for the test discussed above.

The power is $1 - \beta = p_1(4) = 0.3$, implying that the probability of a type II error is $\beta = 0.7$. Hence, as expected, the power is lower for this test than for the NP test (which is the test with the highest possible power), and β is higher.

6. Suppose $X_1, ..., X_n$ are independent, distributed as X which has density function

$$f(x; \beta) = \frac{2x}{\beta} \exp\left(-\frac{x^2}{\beta}\right), \quad x > 0,$$

and 0 otherwise. Let $x_1, ..., x_n$ be the observations.

(a) Show that this distribution belongs to a one-parameter exponential family. (1p)

Solution: Clearly, the density is on exponential family form with one natural parameter (see also (b)).

(b) Give the natural parameter and the sufficient statistic. (1p)

Solution: Let $x_1, ..., x_n$ be the observations. The likelihood is

$$L(\beta) = 2^n \beta^{-n} \exp\left(-\frac{1}{\beta} \sum_{i=1}^{n} x_i^2\right) \prod_{i=1}^{n} x_i,$$

which gives the natural parameter $\eta = -1/\beta$ and the sufficient statistic $T = \sum_{i=1}^{n} x_i^2$.

(c) Consider testing H_0: $\beta \leq \beta_0$ vs H_1: $\beta > \beta_0$. Show that the uniformly most powerful (UMP) test has critical region $\sum_{i=1}^{n} x_i^2/n > C$ where C is some constant. (3p)

Solution: The likelihood is monotonely increasing with η. Hence, we have an MLR family, and by Blackwell's theorem (5.4), the test of $H_0 : \eta \leq \eta_0$ vs $H_0 : \eta > \eta_0$ with critical region $T > k$ for some k is UMP. This is equivalent to testing H_0: $\beta \leq \beta_0 = -1/\eta_0$ vs H_1: $\beta > \beta_0$ with critical region $\sum_{i=1}^{n} x_i^2/n > C$ for $C = nk$.

7. Suppose $X_1, ..., X_n$ are independent and normally distributed with expectation μ and variance 1. We want to test H_0: $\mu = 0$ vs H_1: $\mu \neq 0$ at level α.

   Define z_α through $P(Z < z_\alpha) = \alpha$ where Z is standard normal.
   Let $\bar{X} = n^{-1} \sum_{i=1}^{n} X_i$, and let T_{obs} be the observed value of $T = \sqrt{n}\bar{X}$.

   (a) Which of the following tests (if any) are unbiased size α tests, and why? (3p)

      i. The test that rejects if.f. $T_{obs} > z_{1-\alpha}$.
      ii. The test that rejects if.f. $|T_{obs}| > z_{1-\frac{\alpha}{2}}$.
      iii. The test that rejects if.f. $T_{obs} < z_{\frac{\alpha}{4}}$ or $T_{obs} > z_{1-\frac{3\alpha}{4}}$.

      Solution: An unbiased test should attain minimum rejection probability when the parameter is in the null region. This is not true for the test in i., where the rejection probability goes to zero as $\mu \to -\infty$. But the rejection probability for $\mu = 0$ equals $\alpha > 0$.

      In fact, if we test $H_0 : \mu = -\delta$ vs $H_1 : \mu \neq -\delta$ for any $\delta > 0$, then the probability under H_0 of getting $T_{obs} > z_{1-\alpha}$ would be smaller than α. This shows that the rejection probability for all negative μ is less than α, which contradicts unbiasedness.

      The test in ii, however, is unbiased. If fact, by corollary 5.3 (see also special case 5.5), it is even UMPU.

      As for the test in iii, it is not unbiased. This is because the rejection region is not symmetric. For example, we reject for all observed T between $z_{1-(3/4)\alpha}$ and $z_{1-(1/4)\alpha}$, but not for any observed T between $z_{(3/4)\alpha}$ and $z_{(1/4)\alpha}$. Now, say that we test $H_0 : \mu = -\delta$ vs $H_1 : \mu \neq -\delta$ for some small $\delta > 0$, with the same rejection region as in iii. Then, T would tend to be a little smaller under H_0, and the probability for T to attain a value in the upper part of the rejection region would decrease while the probability for the lower part increases. However, because the density function is steeper at the boundary of the upper part ($z_{1-\frac{3\alpha}{4}}$) than at the boundary of the lower part ($z_{\frac{\alpha}{4}}$), the decrease in the upper part would be larger than the increase in the lower part. Hence, the probability of rejection would be slightly smaller than α. This shows that for some negative μ close to zero, the probability of rejection is smaller than α. Hence, the test is not unbiased.

   (b) Is any of the tests in (a) uniformly most powerful unbiased (UMPU), and in that case, which one? Explain why! (3p)

      Solution: Yes, the test in ii, see the solution of (a).