ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
IX, 1 (1966)

S. ZUBRZYCKI (Wrocław)

EXPLICIT FORMULAS FOR MINIMAX ADMISSIBLE ESTIMATORS IN SOME CASES OF RESTRICTIONS IMPOSED ON THE PARAMETER *

1. Introduction. It has been known for a long t me that an estimator which is minimax and/or admissible when the parameter ranges over the whole natural parameter space, can lose one or both of these properties if the variability of the parameter is restricted to a proper subset of the natural parameter space. The oldest example illustrating the situation might be this: In the case of quadratic loss, the sample mean is a minimax and admissible estimator of the expected value μ of a normal distribution with known constant variance when μ varies over the whole line $-\infty < \mu < \infty$. It is still minimax, but no longer admissible, if we know that $\mu \leqslant \mu_0$ or $\mu_0 \leqslant \mu$, where μ_0 is finite. It is neither minimax nor admissible if we know that $\mu_1 \leqslant \mu \leqslant \mu_2$ with both μ_1 and μ_2 finite (comp. Lehmann [9], section 4.2). Admissibility and minimaxity of estimators of location parameter have been since studied by many authors. Farell [5] presents an extensive systematization of what is known in this respect. Also the estimation of the expected value in exponential families of distributions has been studied from this point of view. Recently a new proof was given by Cheng Ping [1] for a result of Karlin [8] concerning admissibility of linear estimators of the expected value in an exponential family of distributions. His proof is based on the use of the Cramér-Rao inequality, an approach originated by Girshick & Savage [6] and Hodges & Lehmann [7]. He also states a fairly general sufficient condition for a linear estimator to be a minimax estimator of the expected value in an exponential family of distributions, applicable in the case of one-sided restriction of the variability of the parameter.

From the practical point of view, the restriction of the variability of the parameter to a proper subset of the natural parameter space ex-

* This research was partly supported by the Office of Naval Research under contract N-onr-477 (38), when the author served as Visiting Assistant Professor at the University of Washington.

presses an additional available information about the parameter. Therefore it is of some interest to study the vulnerability of minimaxity and admissibility of estimators with respect to such restrictions. Also it is interesting to know if and how such additional knowledge about the parameter can be built in into the estimator in order to improve its behavior over the restricted range of the parameter. Our subsequent considerations will be focused on the presentation of two examples where some one-sided restriction imposed on the parameter does not change the minimax risk of the statistician and does not affect the minimaxity of the estimators, and where, on the other hand, it is possible to improve the behavior of the minimax estimators over the restricted range of the parameter in a simple explicit manner. The first of our examples falls into the scheme of estimating the expectation in an exponential family of distributions, so that one obtains the conservation of minimax risk and minimaxity of the considered estimators as a corollary to the theorem of Cheng Ping. We discuss this in section 5. The other of our examples, however, does not fall into this scheme. Since our original derivations follow in both cases the same line, namely to consider Bayes risk corresponding to a conveniently chosen prior distribution of the parameter, we shall present them first, in sections 3 and 4. We begin with the formulation of two known results which served as our starting point.

2. Known results. Let me first recall a result of Dzan Dzo-i [4]. He investigated the minimax estimators of binomial probability p when it is known to lie in an interval of the form $a < p < 1-a$, where a is a nonnegative constant. He has shown among others that Rubin's estimator

$$\frac{m+\frac{1}{2}\sqrt{n}}{n+\sqrt{n}},$$

where m is the number of successes in n independent trials, remains minimax as long as $a \leqslant 1/[2(1+\sqrt{n})]$. Another example of a situation where a restriction of a certain type imposed on the parameter does not affect the minimaxity of the estimator was encountered by Czen Pin [2], and his result has been proved in another way and completed by the discussion of admissibility by Zubrzycki [10]. The problem is that of estimating the size of a finite population by inverse sampling. More precisely, the question is: what are the minimax estimators of the population size, based on the number of selections needed to select a marked element for the s-th time, if 1) the selections are made one by one with replacement, 2) we know the number of marked elements in the population at the beginning of the selections and 3) we know that the size of the population is greater than or equal to a certain number known in advance. The following has been proved (see Zubrzycki [10]).

If the probability distribution of m is given by

$$(2.1)\quad \Pr\{m = k \mid N\} = \binom{k-1}{s-1}\left(\frac{t}{N}\right)^{s}\left(1-\frac{t}{N}\right)^{k-s}, \qquad k = s, s+1, \ldots,$$

where s is a positive integer and N ranges continuously over $N_0 \leqslant N < \infty$, and

$$(2.2)\qquad L(N, N') = (N-N')^2/N^2$$

is the loss resulting if N is the true value of the parameter and we estimate it as N', then estimators of N given by

$$(2.3)\qquad \hat{N}(m) = b+mt/(s+1)$$

with

$$(2.4)\qquad \frac{N_0}{s+1}\left(1-\sqrt{1+\frac{st}{N_0}}\right) \leqslant b \leqslant \frac{N_0}{s+1}\left(1+\sqrt{1+\frac{st}{N_0}}\right),$$

are minimax with minimax risk equal to $1/(s+1)$, they are the only minimax estimators linear in m, and among them only those with

$$(2.5)\qquad \frac{N_0}{s+1} \leqslant b \leqslant \frac{N_0}{s+1}\left(1+\sqrt{1+\frac{st}{N_0}}\right)$$

are admissible in the class of estimators linear in m.

It is to be noted that both Czen Pin and Zubrzycki in their proofs regard N as a continuous parameter, where in the original problem N is integer-valued. Moreover, if we note that in the above result, which refers to sampling with replacement, $p = t/N$ may be interpreted as the probability of a success in a series of independent trials and m as the waiting time for the s-th success, and that $1/p$ is the expected value of this waiting time, we can give the above result the form revealing the connection with estimating the expectation of a negative binomial distribution (see Zubrzycki [10]):

If m has negative binomial distribution

$$(2.5)\qquad \Pr\{m = k \mid p\} = \binom{k-1}{s-1} p^{s}(1-p)^{k-s}, \qquad k = s, s+1, \ldots,$$

where p ranges over $0 < p \leqslant p_0 \leqslant 1$, and

$$(2.6)\qquad L(p, p') = p^2\left(\frac{1}{p}-p'\right)^2 = (pp'-1)^2$$

is the loss resulting if we estimate $1/p$ as p', then estimators of $1/p$ given by

$$\hat{p}(m) = d+m/(s+1) \tag{2.7}$$

with

$$\frac{1}{p_0(s+1)}(1-\sqrt{1+sp_0}) \leqslant d \leqslant \frac{1}{p_0(s+1)}(1+\sqrt{1+sp_0}) \tag{2.8}$$

are minimax with minimax risk $1/(s+1)$, they are the only minimax estimators linear in m, and among them only those with

$$\frac{1}{p_0(s+1)} \leqslant d \leqslant \frac{1}{p_0(s+1)}(1+\sqrt{1+sp_0}) \tag{2.9}$$

are admissible in the class of estimators linear in m.

3. Estimating the expected value of a gamma distribution. As we have seen, the result presented in section 2 can be viewed as referring to estimation of the expected value of the waiting time for the first success in a series of independent trials on the basis of s independent observations of that waiting time. The waiting time for the first success has, as it is well known, a geometric distribution. The question suggests itself, what can be said in the case of the continuous counterpart of geometric distribution, namely in the case of an exponential distribution. It is known that if we have s independent random variables having the same exponential distribution with probability density function

$$f(x \mid \lambda) = \lambda e^{-\lambda x}, \quad x > 0, \tag{3.1}$$

then their sum X, which is a sufficient statistic for λ, has the gamma distribution with the probability density

$$g(x \mid \lambda) = \frac{\lambda^s}{\Gamma(s)} x^{s-1} e^{-\lambda x}, \quad x > 0, \tag{3.2}$$

and expected value and variance of X are given by

$$E(X \mid \lambda) = \frac{s}{\lambda}, \quad D^2(X \mid \lambda) = \frac{s}{\lambda^2}. \tag{3.3}$$

Now formulas (3.2) and (3.3) make sense for any real $s > 0$. As s is fixed in our considerations, we can forget for a moment that s is an integer representing the sample size and look upon $g(x \mid \lambda)$ as the probability density of a continuous-time gamma process at moment s. We then can ask what are the mimimax extimators of $\beta = 1/\lambda$ and the corresponding minimax risk. The following will be proved:

THEOREM 1. *If X is a random variable with probability density $g(x \mid \lambda)$ given by* (3.2), *the parameter λ ranges over $0 < \lambda \leqslant \lambda_0$, and*

$$L(\beta, \beta') = (\beta - \beta')^2/\beta^2 = \lambda^2 \left(\frac{1}{\lambda} - \beta'\right)^2 = (1 - \lambda\beta')^2 \tag{3.4}$$

is the loss incurred if $\beta = 1/\lambda$ is estimated as β', then estimators of $\beta = 1/\lambda$ given by

$$\hat{\beta}(X) = \frac{X}{s+1} + b \tag{3.5}$$

with

$$0 \leqslant b \leqslant \frac{2}{\lambda_0(s+1)} \tag{3.6}$$

are minimax with minimax risk equal to $1/(s+1)$, these are the only minimax estimators linear in X, and among them only those estimators with

$$\frac{1}{\lambda_0(s+1)} \leqslant b \leqslant \frac{2}{\lambda_0(s+1)} \tag{3.7}$$

are admissible in the class of estimators linear in X.

We will divide the proof in a series of lemmas which may have some independent interest.

(a) *Estimators of $\beta = 1/\lambda$ given by* (3.5) *and* (3.6) *are the only minimax estimators among linear estimators of the form*

$$\hat{\beta}(X) = aX + b. \tag{3.8}$$

Proof. Indeed, for estimators of the form (3.8) the risk is given by

$$r(\lambda; aX+b) = [(1-as)^2 + a^2 s] - 2b(1-as)\lambda + b^2\lambda^2. \tag{3.9}$$

As this is a quadratic polynomial in λ with a non-negative coefficient at λ^2, its supremum with respect to λ is attained at one of the extremes of the interval $0 < \lambda \leqslant \lambda_0$. It is easily checked that the minimum of $(1-as)^2 + a^2 s$ equals $1/(s+1)$ and is attained at $a = 1/(s+1)$. With this choice of a the risk becomes

$$r\left(\lambda; \frac{X}{s+1} + b\right) = \frac{1}{s+1} - 2\frac{b}{s+1}\lambda + b^2\lambda^2.$$

In order that $1/(s+1)$ be the supremum of this risk with respect to λ we have only to impose on b the condition

$$b^2\lambda_0^2 - 2b\frac{\lambda_0}{s+1} \leqslant 0,$$

which is equivalent to (3.6).

(b) *Among the minimax estimators given by* (3.5) *and* (3.6), *only, those with b within the limits* (3.7) *are admissible in the class of estimators linear in* X.

Proof. Let us compare two estimators $\hat{\beta}_1$ and $\hat{\beta}_2$ with $0 \leqslant b_1 \leqslant b_2 \leqslant \leqslant 2/\lambda_0(s+1)$. We then have

$$r(\lambda; \hat{\beta}_1) = \frac{1}{s+1} - 2\lambda b_1 \frac{1}{s+1} + \lambda^2 b_1^2, \qquad r(\lambda; \hat{\beta}_2) = \frac{1}{s+1} - 2\lambda b_2 \frac{1}{s+1} + \lambda^2 b_2^2,$$

whence by subtraction

$$r(\lambda; \hat{\beta}_2) - r(\lambda; \hat{\beta}_1) = \lambda^2(b_2^2 - b_1^2) - 2\frac{b_2 - b_1}{s+1}\lambda = (b_2 - b_1)\lambda\left[\lambda(b_2 + b_1) - \frac{2}{s+1}\right].$$

This difference is a quadratic polynomial in λ with a positive coefficient at λ^2, so it is negative between its roots $\lambda_1 = 0$ and $\lambda_2 = 2/(b_2+b_1)(s+1)$. If now $\lambda_0 \leqslant \lambda_2$, then $\hat{\beta}_2$ is uniformly better than $\hat{\beta}_1$ in $0 < \lambda \leqslant \lambda_0$, that is, $\hat{\beta}_1$ is inadmissible. Now as long as $0 \leqslant b_1 < 1/\lambda_0(s+1)$, we can choose b_2 with $b_1 < b_2 < 1/\lambda_0(s+1)$ so as to have $0 \leqslant b_1 + b_2 < 2/\lambda_0(s+1)$ and thus $\lambda_2 = 2/(b_1+b_2)(s+1) > \lambda_0$. This proves that if $b_1 < 1/\lambda_0(s+1)$, then $\hat{\beta}_1$ is inadmissible.

On the other hand, if $1/\lambda_0(s+1) \leqslant b_1 < 2/\lambda_0(s+1)$, then with $b_1 < b_2 \leqslant 2/\lambda_0(s+1)$ we have $\lambda_2 < \lambda_0$, and thus $\hat{\beta}_1$ and $\hat{\beta}_2$ are incomparable. This proves that $\hat{\beta}_1$ with $1/\lambda_0(s+1) \leqslant b_1 \leqslant 2/\lambda_0(s+1)$ is admissible in the class of estimators linear in X.

(c) *If λ has a priori the gamma distribution with probability density*

$$h(\lambda) = h_{A,c}(\lambda) = \frac{A^c}{\Gamma(c)}\lambda^{c-1}e^{-A\lambda}, \quad c > 0, \quad A > 0, \quad \lambda > 0, \tag{3.10}$$

then the Bayes estimator of $\beta = 1/\lambda$ is given by

$$\hat{\beta}_{A,c}(X) = \frac{X+A}{s+c+1}, \tag{3.11}$$

and the conditional risk

$$\int_0^\infty L(\lambda, \hat{\beta}_{A,c})h(\lambda \mid x)d\lambda = \frac{1}{s+c+1} \tag{3.12}$$

is independent of x; here $h(\lambda \mid x)$ is the probability density of the posterior probability distribution of λ given $X = x$.

We would like to recall here that this last independence was used as an argument in [3] to prove that among all sequential procedures for estimating β the minimax one is a fixed time procedure.

Proof. With loss function (3.4) the conditional risk, given $X = x$, is equal to

$$E_{\lambda|x}\left(1-\lambda\hat{\beta}(x)\right)^2 = \hat{\beta}^2(x)\cdot E_{\lambda|x}(\lambda^2) - 2\hat{\beta}(x)\cdot E_{\lambda|x}(\lambda)+1,$$

and, as is easily seen, it is minimized to

$$D^2_{\lambda|x}(\lambda)/E_{\lambda|x}(\lambda^2) \tag{3.13}$$

by the choice

$$\hat{\beta}(x) = E_{\lambda|x}(\lambda)/E_{\lambda|x}(\lambda^2), \tag{3.14}$$

which is the form of the Bayes estimator in this case. Here and in the sequel $E_{\lambda|x}(\ )$ and $D^2\ (\ )$ denote the expected value and the variance of the random variable in parantheses with respect to the conditional distribution of λ given $X \epsilon x$.

If the prior probability distribution of λ is given by (3.10), then the posterior probability distribution has the probability density

$$h_{A,c}(\lambda \mid x) = \frac{(x+A)^{s+c}}{\Gamma(s+c)}\,\lambda^{s+c-1}e^{-(x+A)\lambda}.$$

Expected value, second moment and variance of this posterior distribution are

$$\frac{s+c}{x+A},\qquad \frac{(s+c)(s+c+1)}{(x+A)^2},\qquad \frac{s+c}{(x+A)^2},$$

which, when substituted into (3.14) and (3.13), yield (3.11) as the Bayes estimator and (3.12) as the conditional risk. This proves (c).

Note that this conditional risk, being independent of x, is equal to the Bayes risk corresponding to the prior distribution (3.10). Moreover, (a) with $\lambda_0 = \infty$ and (c) together prove that in the unrestricted case $1/(s+1)$ is the minimax risk and $X/(s+1)$ is the minimax estimator of $\beta = 1/\lambda$.

To complete the proof of our theorem we are now going to show that statistician's minimax risk remains unchanged if we impose the restriction $0 < \lambda \leqslant \lambda_0$ on parameter λ. To do so let us consider the truncated gamma distribution with probability density function

$$h(\lambda) = h_{A,c,\lambda_0}(\lambda) = \frac{\lambda^{c-1}e^{-A\lambda}}{\int_0^{\lambda_0}\lambda^{c-1}e^{-A\lambda}d\lambda},\qquad A>0,\qquad c>0,\qquad 0<\lambda\leqslant\lambda_0, \tag{3.15}$$

as the prior probability distribution of λ. For the posterior probability density of λ we have then the expression

$$h_{A,c,\lambda_0}(\lambda \mid x) = \frac{h(\lambda)g(x \mid \lambda)}{\int_0^{\lambda_0} h(\lambda)g(x \mid \lambda)d\lambda} = \frac{\lambda^{s+c-1}e^{-(x+A)\lambda}}{\int_0^{\lambda_0} \lambda^{s+c-1}e^{-(x+A)\lambda}d\lambda} \quad \text{for} \quad 0 < \lambda \leqslant \lambda_0,$$

and the expressions for the first and second moment of the posterior distribution of λ, given $X = x$, become

$$E_{A,c,\lambda_0}(\lambda \mid x) = \frac{\int_0^{\lambda_0} \lambda\,\lambda^{s+c-1}e^{-(x+A)\lambda}d\lambda}{\int_0^{\lambda_0} \lambda^{s+c-1}e^{-(x+A)\lambda}d\lambda}, \tag{3.16}$$

$$E_{A,c,\lambda_0}(\lambda^2 \mid x) = \frac{\int_0^{\lambda_0} \lambda^2\lambda^{s+c-1}e^{-(x+A)\lambda}d\lambda}{\int_0^{\lambda_0} \lambda^{s+c-1}e^{-(x+A)\lambda}d\lambda}, \tag{3.17}$$

hence, in view of (3.14), the Bayes estimator is

$$\beta_{A,c,\lambda_0}(x) = \frac{\int_0^{\lambda_0} \lambda\,\lambda^{s+c-1}e^{-(x+A)\lambda}d\lambda}{\int_0^{\lambda_0} \lambda^2\lambda^{s+c-1}e^{-(x+A)\lambda}d\lambda}. \tag{3.18}$$

We will prove the following:

(d) *For fixed $A > 0$ and $\lambda_0 > 0$ we have for the Bayes risk the relation*

$$\lim_{c\to 0} r(H_{A,c,\lambda_0}, \hat{\beta}_{A,c,\lambda_0}) = \frac{1}{s+1}; \tag{3.19}$$

here H_{A,c,λ_0} represents the prior distribution of λ with *density $h_{A,c,\lambda_0}(\lambda)$*.

Proof. Observe first that, for every $k = 0, 1, 2, \ldots$, we have

$$\lim_{x\to\infty} \frac{\int_0^{\lambda_0} \lambda^k\lambda^{s+c-1}e^{-A\lambda}e^{-x\lambda}d\lambda}{\int_0^{\infty} \lambda^k\lambda^{s+c-1}e^{-A\lambda}e^{-x\lambda}d\lambda} = 1 \tag{3.20}$$

and convergence is uniform in c for $0 < c \leqslant c_0$.

Instead of proving (3.20) directly we will look for the behavior of $1-(\int_0^{\lambda_0}/\int_0^{\infty}) = \int_{\lambda_0}^{\infty}/\int_0^{\infty}$ and we will prove that this complement converges to 0 uniformly in c. As a matter of fact, we can write

$$0 \leqslant \frac{\int_{\lambda_0}^{\infty} \lambda^k \lambda^{s+c-1} e^{-A\lambda} e^{-x\lambda} d\lambda}{\int_0^{\infty} \lambda^k \lambda^{s+c-1} e^{-Ax} e^{-x\lambda} d\lambda} \leqslant \frac{\int_{\lambda_0}^{\infty} \lambda^k \lambda^{s+c-1} e^{-A\lambda} e^{-x\lambda} d\lambda}{\int_0^{\lambda_0} \lambda^k \lambda^{s+c-1} e^{-A\lambda} e^{-x\lambda} d\lambda}$$

$$\leqslant \frac{\int_{\lambda_0}^{\infty} \lambda^k \lambda^{s+c_0-1} e^{-A\lambda} e^{-x\lambda} d\lambda}{\int_0^{\lambda_0} \lambda^k \lambda^{s+c_0-1} e^{-A\lambda} e^{-x\lambda} d\lambda} \leqslant \frac{\int_{\lambda_0}^{\infty} M e^{-x\lambda} d\lambda}{\int_0^{\lambda_0} \lambda^k \lambda^{s+c_0-1} e^{-A\lambda} e^{-x\lambda_0} d\lambda}$$

$$= \frac{M \int_{\lambda_0}^{\infty} e^{-x\lambda} d\lambda}{e^{-x\lambda_0} \int_0^{\lambda_0} \lambda^k \lambda^{s+c_0-1} e^{-A\lambda} d\lambda} = \frac{\frac{1}{x} M e^{-x\lambda_0}}{e^{-x\lambda_0} C} = \frac{1}{x} \frac{M}{C};$$

here M is a positive constant such that

$$\lambda^k \lambda^{s+c_0-1} e^{-A\lambda} \leqslant M \quad \text{for} \quad \lambda_0 \leqslant \lambda < \infty,$$
$$C = \int_0^{\lambda_0} \lambda^k \lambda^{s+c_0-1} e^{-A\lambda} d\lambda.$$

This bound of order $1/x$ being independent of c, we have (3.20) uniformly in c.

In view of (3.11), (3.18) and (3.20) we can conclude that

$$\lim_{x\to\infty} \frac{\hat{\beta}_{A,c,\lambda_0}(x)}{\hat{\beta}_{A,c}(x)} = 1 \quad \text{uniformly in } c. \tag{3.21}$$

Now let us make some observations about $\hat{\beta}_{A,c}(x)$. It is a non-decreasing function of x, so for $x < x_0$ and $\lambda < \lambda_0$ the loss

$$\left(1-\hat{\beta}_{A,c}(x)\lambda\right)^2$$

is bounded, and hence

$$\int_0^{x_0} \left(1-\hat{\beta}_{A,c}(x)\lambda\right)^2 g(x \mid \lambda) dx$$

can be made as small as we wish by taking λ sufficiently small, because for every $x_0 > 0$

$$\lim_{\lambda\to 0} \int_0^{x_0} g(x \mid \lambda) dx = 0. \tag{3.22}$$

Now, in view of (3.9) and (3.11), we have

$$(3.23)\qquad r(\lambda,\hat{\beta}_{A,c}) = \int_0^\infty \big(1-\lambda\hat{\beta}_{A,c}(x)\big)^2 g(x\mid\lambda)\,dx$$
$$= \frac{s+(c+1)^2}{(s+c+1)^2} - 2\frac{A(c+1)}{(s+c+1)^2}\lambda + \frac{A^2}{(s+c+1)^2}\lambda^2.$$

Therefore $r(\lambda,\hat{\beta}_{A,c})$ approaches $1/(s+1)$ as $\lambda\to 0$ and $c\to 0$, and so does, for any fixed $x_0>0$, the integral

$$\int_{x_0}^\infty \big(1-\lambda\hat{\beta}_{A,c}(x)\big)^2 g(x\mid\lambda)\,dx.$$

Now it is easily seen that if we choose $x_0>0$ so as to have, for a given small $\eta>0$,

$$(1-\eta)\hat{\beta}_{A,c}(x) \leqslant \sqrt{(1-\eta)}\,\hat{\beta}_{A,c}(x) \leqslant \hat{\beta}_{A,c,\lambda_0}(x) \leqslant \sqrt{(1+\eta)}\,\hat{\beta}_{A,c}(x)$$
$$\leqslant (1+\eta)\hat{\beta}_{A,c}(x),$$

then we will have

$$(3.24)\qquad \int_{x_0}^\infty \big(1-\lambda\hat{\beta}_{A,c,\lambda_0}(x)\big)^2 g(x\mid\lambda)\,dx \geqslant \int_{x_0}^\infty \big(1-\lambda\hat{\beta}_{A,c}(x)\big)^2 g(x\mid\lambda)\,dx\,(1-\eta),$$

because of the inequalities

$$-(1+\eta)\cdot 2\lambda\hat{\beta}_{A,c}(x) \leqslant -2\lambda\hat{\beta}_{A,c,\lambda_0}(x),$$
$$(1-\eta)\lambda^2\hat{\beta}^2_{A,c}(x) \leqslant \lambda^2\hat{\beta}^2_{A,c,\lambda_0}(x),$$
$$(1-\eta)\cdot 1 \leqslant 1$$

which, when summed on both sides and integrated with respect to $g(x\mid\lambda)\,dx$, yield (3.24).

We thus see that for any $\varepsilon>0$ there are x_0 and λ' such that for $\lambda<\lambda'$ the left-hand side of (3.24) is greater than or equal to $(1-\varepsilon)/(s+1)$ uniformly in $c\leqslant c_0$. Now we can find a $c'\leqslant c_0$ such that for $c\leqslant c'$

$$\int_0^{\lambda'} h_{A,c,\lambda_0}(\lambda)\,d\lambda \geqslant (1-\varepsilon).$$

Then we have

$$r(H_{A,c,\lambda_0},\hat{\beta}_{A,c,\lambda_0}) \geqslant (1-\varepsilon)^2/(s+1).$$

Because $\varepsilon>0$ is arbitrary, (d) is proved.

Lemmas (a), (b), (c) and (d) together prove Theorem 1.

4. Estimating λ in the gamma distribution of the form $f(x|\lambda) = \lambda^s x^{s-1} e^{-\lambda x}/\Gamma(s)$. In the preceding section we have been interested in the estimation of $\beta = 1/\lambda$, that is, the rate of growth of the underlying gamma process as measured by the expected value of the increase per unit of time. If we go back to the case of integer s and exponential distribution serving as the distribution of waiting time for the next signal in the homogeneous continuous-time Poisson process, we can say that we wanted to estimate the expected value of this waiting time. On the other hand, one might be as well interested in the estimation of λ, the intensity of the Poisson process. In our setting this is the estimation of intensity λ on the basis of the observed lapse of time to the appearance of the s-th signal, as opposed to the possibility of observing the number of signals in a time-interval of fixed length. Again we want to show that in this case the restriction of the variability of λ from below does not affect statistician's minimax risk and we want to present an explicit family of minimax estimators of λ with some additional knowledge about the parameter built into them.

Let us start with the following observation:

If a random variable X has the probability density

$$f(x \mid \lambda) = \frac{\lambda^s}{\Gamma(s)} x^{s-1} e^{-\lambda x}, \quad x > 0, \quad \lambda > 0, \quad s > 2, \tag{4.1}$$

then we have

$$E\left(\frac{1}{X} \mid \lambda\right) = \int_0^\infty \frac{1}{x} f(x \mid \lambda)\, dx = \frac{\lambda}{s-1},$$

$$E\left(\frac{1}{X^2} \mid \lambda\right) = \int_0^\infty \frac{1}{x^2} f(x \mid \lambda)\, dx = \frac{\lambda^2}{(s-1)(s-2)}, \tag{4.2}$$

$$D^2\left(\frac{1}{X} \mid \lambda\right) = \frac{\lambda^2}{(s-1)^2(s-2)}.$$

We want to prove the following:

THEOREM 2. *Let X be a random variable with probability density $f(x \mid \lambda)$ given by* (4.1), *where $s > 2$ and*

$$0 \leqslant \lambda_0 \leqslant \lambda < \infty. \tag{4.3}$$

Let the loss of estimating λ as λ' be given by

$$L(\lambda, \lambda') = (\lambda - \lambda')^2/\lambda^2. \tag{4.4}$$

Then estimators of λ given by

$$\frac{s-2}{X} + b \tag{4.5}$$

with

$$0 \leqslant b \leqslant \frac{2\lambda_0}{s-1} \tag{4.6}$$

are minimax in the class of all estimators, they are the only minimax estimators linear in $1/X$, *the minimax risk equals* $1/(s-1)$, *and among these estimators only those with*

$$\frac{\lambda_0}{s-1} \leqslant b \leqslant \frac{2\lambda_0}{s-1} \tag{4.7}$$

are admissible in the class of estimators linear in $1/X$.

As in the preceding section, we will prove several lemmas, which might be of some independent interest, and from which Theorem 2 follows immediately.

(a) *Under assumptions of Theorem* 2, *estimators described by* (4.5) *and* (4.6) *are the only minimax estimators among estimators linear in* $1/X$.

Proof. For an estimator of the form

$$\hat{\lambda}(X) = \frac{a}{X} + b$$

the risk is given by

$$r\left(\lambda, \frac{a}{X} + b\right) = E\left(\frac{1}{\lambda^2}\left(\lambda - \frac{a}{X} - b\right)^2 \middle| \lambda\right)$$

$$= \left(\frac{a^2}{(s-1)(s-2)} - 2\frac{a}{s-1} + 1\right) + 2b\frac{1}{\lambda}\left(\frac{a}{s-1} - 1\right) + b^2\frac{1}{\lambda^2}.$$

The term free of λ is minimized to $1/(s-1)$ by the choice of $a = s-2$. The risk then becomes

$$r\left(\lambda, \frac{s-2}{X} + b\right) = \frac{1}{s-1} - 2\frac{b}{s-1}\frac{1}{\lambda} + b^2\frac{1}{\lambda^2}. \tag{4.8}$$

Now it is easily seen that

$$b^2\frac{1}{\lambda^2} - 2\frac{b}{s-1}\frac{1}{\lambda} \leqslant 0 \quad \text{for} \quad \lambda_0 \leqslant \lambda < \infty$$

if only b is within the limits (4.6). Thus (a) is proved.

(b) *Among estimators described by* (4.5) *and* (4.6) *only those with* b *within the limits* (4.7) *are admissible in the class of estimators linear in* $1/X$.

Proof. In view of (4.8) the difference of risks corresponding to the two estimators

$$\hat{\lambda}_1(X) = \frac{s-2}{X} + b_1 \quad \text{and} \quad \hat{\lambda}_2(X) = \frac{s-2}{X} + b_2$$

with $0 \leqslant b_1 < b_2 \leqslant 2\lambda_0/(s-1)$ is given by

$$r(\lambda, \hat{\lambda}_1) - r(\lambda, \hat{\lambda}_2) = (b_1^2 - b_2^2)\frac{1}{\lambda^2} - 2\frac{b_1 - b_2}{s-1}\frac{1}{\lambda}.$$

This quadratic polynomial in $1/\lambda$ is positive between its roots which are 0 and $2/(s-1)(b_1+b_2)$. In order to have

$$r(\lambda, \hat{\lambda}_1) - r(\lambda, \hat{\lambda}_2) > 0 \quad \text{for} \quad \lambda_0 \leqslant \lambda < \infty,$$

in which case $\hat{\lambda}_1$ is inadmissible because of being uniformly worse than $\hat{\lambda}_2$, it is enough to choose b_2 so as to have

$$\frac{1}{\lambda_0} < \frac{2}{(s-1)(b_1+b_2)}$$

or

$$b_1 + b_2 < \frac{2\lambda_0}{s-1}.$$

We see that this is possible as long as $b_1 < \lambda_0/(s-1)$. If $b_1 \geqslant \lambda_0/(s-1)$, the estimators $\hat{\lambda}_1$ and $\hat{\lambda}_2$ are incomparable: for some values of λ the estimator $\hat{\lambda}_1$ has smaller risk, for some other $\hat{\lambda}_2$ has smaller risk. So (b) is proved.

Let $h(\lambda)$ be the probability density and $H(\lambda)$ the cumulative distribution function of a prior probability distribution of λ. Let further $h(\lambda \mid x)$ denote the probability density of the corresponding conditional distribution of λ with x fixed. With these notations and the loss function given by (4.4), the Bayes estimator corresponding to a prior distribution $H(\lambda)$, that is, the estimator $\hat{\lambda}_H(x)$ minimizing

$$r(H, \hat{\lambda}) = \int_0^\infty r(\lambda, \hat{\lambda})\,dH(\lambda) = \int_0^\infty \Big(\int_0^\infty L\big(\lambda, \hat{\lambda}(x)\big) f(x \mid \lambda)\,dx\Big)\,dH(\lambda),$$

is given by

$$\hat{\lambda}_H(x) = \frac{E_{\lambda|x}\left(\frac{1}{\lambda}\right)}{E_{\lambda|x}\left(\frac{1}{\lambda^2}\right)}, \tag{4.9}$$

provided both expectations exist.

Actually, the conditional risk

$$\int_0^\infty L(\lambda, \hat{\lambda}(x))h(\lambda \mid x)d\lambda = \int_0^\infty (\lambda-\hat{\lambda}(x))^2 \frac{1}{\lambda^2} h(\lambda \mid x)d\lambda$$
$$= 1-2\hat{\lambda}(x)E_{\lambda|x}\left(\frac{1}{\lambda}\right)+\hat{\lambda}^2(x)E_{\lambda|x}\left(\frac{1}{\lambda^2}\right)$$

is minimized to

$$D^2_{\lambda|x}\left(\frac{1}{\lambda}\right)\Big/E_{\lambda|x}\left(\frac{1}{\lambda^2}\right) \tag{4.10}$$

by choosing $\hat{\lambda}(x) = \hat{\lambda}_H(x)$ as given by (4.9), if all moments involved are finite.

(c) *If λ has a priori the gamma distribution with probability density*

$$h(\lambda) = h_{A,c}(\lambda) = \frac{A^c}{\Gamma(c)}\lambda^{c-1}e^{-A\lambda}, \quad A>0, \quad c>0, \quad \lambda>0, \tag{4.11}$$

then

$$h(\lambda \mid x) = h_{A,c}(\lambda \mid x) = \frac{(x+A)^{s+c}}{\Gamma(s+c)}\lambda^{s+c-1}e^{-(x+A)\lambda}, \tag{4.12}$$

$$\hat{\lambda}_{A,c}(x) = \frac{\int_0^\infty \lambda^{s+c-2}e^{-(x+A)\lambda}d\lambda}{\int_0^\infty \lambda^{s+c-3}e^{-(x+A)\lambda}d\lambda} = \frac{s+c-2}{x+A}, \tag{4.13}$$

$$r(H_{A,c}, \hat{\lambda}_{A,c}) = \frac{1}{s+c-1}. \tag{4.14}$$

Proof. Straightforward computation yields (4.12) and (4.13). Then we easily see that the conditional risk, expressed by (4.10), equals $1/(s+c-1)$ independently of x; thus (4.14) follows.

Suppose now that λ has a priori the truncated gamma distribution with probability density given by

$$h(\lambda) = h_{A,c,\lambda_0}(\lambda) = \frac{\lambda^{c-1}e^{-A\lambda}}{\int_{\lambda_0}^\infty \lambda^{c-1}e^{-A\lambda}d\lambda} \quad \text{for} \quad \lambda_0 \leqslant \lambda < \infty. \tag{4.15}$$

We then easily conclude that the Bayes estimator corresponding to this prior distribution is, according to (4.9), given by

$$\lambda_{A,c,\lambda_0}(x) = \frac{\int_{\lambda_0}^\infty \lambda^{s+c-2}e^{-(x+A)\lambda}d\lambda}{\int_{\lambda_0}^\infty \lambda^{s+c-3}e^{-(x+A)\lambda}d\lambda}. \tag{4.16}$$

(Because of $s>2$ this is well defined.)

We are now going to prove

(d) *For any $\varepsilon > 0$ there are λ' and A' such that for $\lambda \geqslant \lambda'$ and $A \leqslant A'$ we have*

$$r(H_{A,c,\lambda_0}, \hat{\lambda}_{A,c,\lambda_0}) \geqslant \frac{1}{s+c-1} - \varepsilon.$$

Proof. Observe first that

$$\lim_{A+x\to 0} \frac{\hat{\lambda}_{A,c,\lambda_0}(x)}{\hat{\lambda}_{A,c}(x)} = 1 \text{ uniformly in } c \leqslant c_0. \tag{4.17}$$

Indeed, by substitution of $u = (x+A)\lambda$ into (4.13) and (4.16) we get the following representations of $\hat{\lambda}_{A,c}$ and $\hat{\lambda}_{A,c,\lambda_0}$:

$$\hat{\lambda}_{A,c}(x) = \frac{1}{x+A} \frac{\int_0^\infty u^{s+c-2} e^{-u} du}{\int_0^\infty u^{s+c-3} e^{-u} du},$$

$$\hat{\lambda}_{A,c,\lambda_0}(x) = \frac{1}{x-A} \frac{\int_{(x+A)\lambda_0}^\infty u^{s+c-2} e^{-u} du}{\int_{(x+A)\lambda_0}^\infty u^{s+c-3} e^{-u} du}.$$

With this (4.17) is evident.

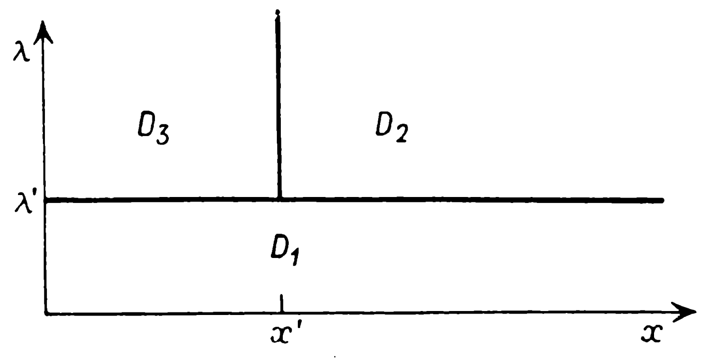

Fig. 1

Consider now the three domains in the (x, λ)-plane given by

$$D_1 = \{(x, \lambda):\ 0 < x < \infty,\ 0 < \lambda \leqslant \lambda'\},$$
$$D_2 = \{(x, \lambda):\ x' \leqslant x < \infty,\ \lambda' \leqslant \lambda < \infty\},$$
$$D_1 = \{(x, \lambda):\ 0 < x < x',\ \lambda' \leqslant \lambda < \infty\},$$

indicated on Fig. 1. Also, to shorten formulas, denote $f(x \mid \lambda) h_{A,c}(\lambda) dx d\lambda$ by $d\mu$.

By (4.17) for any small $\varepsilon > 0$ we can choose A', λ' and x' so as to have

$$(1-\eta)\hat{\lambda}_{A,c}(x) \leqslant \hat{\lambda}_{A,c,\lambda_0}(x) \leqslant (1+\eta)\hat{\lambda}_{A,c}(x),$$

and

$$(1-\eta)\hat{\lambda}^2_{A,c}(x) \leqslant \hat{\lambda}^2_{A,c,\lambda_0}(x) \leqslant (1+\eta)\hat{\lambda}^2_{A,c}(x).$$

for $A < A'$ and $(x, \lambda) \epsilon D_3$.

Because

$$L(\lambda, \hat{\lambda}) = \left(1-2\frac{1}{\lambda}\hat{\lambda}+\frac{1}{\lambda^2}\hat{\lambda}^2\right),$$

we can conclude, by estimating term by term, that

$$\begin{aligned}(4.18) \qquad \iint_{D_3} L(\lambda, \hat{\lambda}_{A,c,\lambda_0})\,d\mu &\geqslant \iint_{D_3} L(\lambda, \hat{\lambda}_{A,c})\,d\mu - \iint_{D_3}\left(2\frac{1}{\lambda}\hat{\lambda}_{A,c}+\frac{1}{\lambda^2}\hat{\lambda}^2_{A,c}\right)d\mu \\ &\geqslant \iint_{D_3} L(\lambda, \hat{\lambda}_{A,c})\,d\mu - \eta M,\end{aligned}$$

where M is an upper bound of

$$\iint_{D_3}\left(2\frac{1}{\lambda}\hat{\lambda}_{A,c}+\frac{1}{\lambda^2}\hat{\lambda}^2_{A,c}\right)d\mu$$

independent of λ', x' and A. Estimation (4.18) remains valid if we enlarge λ'.

Observe next that the probability distribution with density (4.1) concentrates for large values of λ on small values of x. To be more specific, for every $x' > 0$ we have

$$\int_{x'}^{\infty} f(x \mid \lambda)\,dx = \frac{\int_{x'}^{\infty} x^{s-1}e^{-\lambda x}\,dx}{\int_0^{\infty} x^{s-1}e^{-\lambda x}\,dx} = \frac{\int_{x'\lambda}^{\infty} u^{s-1}e^{-u}\,du}{\int_0^{\infty} u^{s-1}e^{-u}\,su} \xrightarrow[\lambda\to\infty]{} 0.$$

Because the loss $L(\lambda, \lambda_{A,c}(x))$ is bounded in D_2, we see that by choosing λ' large enough we can make

$$\int_{x'}^{\infty} L(\lambda, \lambda_{A,c}(x))f(x \mid \lambda)\,dx$$

smaller than a given $\eta > 0$ for all $\lambda \geqslant \lambda'$. Consequently we will have

$$(4.19) \qquad \iint_{D_2} L(\lambda, \hat{\lambda}_{A,c}(x))\,d\mu < \eta,$$

with (4.18) still holding.

Now it is easy to see that $\int_0^\infty L(\lambda, \hat{\lambda}_{A,c}(x))f(x \mid \lambda)dx$ is bounded from above independently of λ, for we have

$$\int_0^\infty L(\lambda, \hat{\lambda}_{A,c}(x))f(x \mid \lambda)dx = \int_0^\infty \left(1-2\frac{1}{\lambda}\hat{\lambda}_{A,c}(x)+\frac{1}{\lambda^2}\hat{\lambda}_{A,c}^2(x)\right)f(x \mid \lambda)dx$$

$$\leqslant \int_0^\infty \left(1+2\frac{1}{\lambda}\frac{s+c-2}{x}+\frac{1}{\lambda^2}\left(\frac{s+c-2}{x}\right)^2\right)f(x \mid \lambda)dx,$$

and this last expression, as we conclude from (4.2), is independent of λ. This, when combined with the fact that the probability distribution with density (4.11) concentrates for small values of A on large values of λ, that is, for every $\lambda' > 0$ we have

$$\int_{\lambda'}^\infty h_{A,c}(\lambda)d\lambda = \frac{\int_{\lambda'}^\infty \lambda^{c-1}e^{-A\lambda}d\lambda}{\int_0^\infty \lambda^{c-1}e^{-A\lambda}d\lambda} = \frac{\int_{A\lambda'}^\infty u^{c-1}e^{-u}du}{\int_0^\infty u^{c-1}e^{-u}du} \xrightarrow[A\to 0]{} 1,$$

enables us to adjust A' introduced above so as to have, beside (4.18), also

$$\iint_{D_1} L(\lambda, \hat{\lambda}_{A,c}(x))d\mu < \eta \quad \text{for} \quad A < A'. \tag{4.20}$$

Since by (c) we have

$$r(H_{A,c}, \hat{\lambda}_{A,c}) = \iint_{D_1\cup D_2\cup D_3} L(\lambda, \hat{\lambda}_{A,c}(x))d\mu = \frac{1}{s+c-1},$$

we conclude from (4.19) and (4.20) that

$$\iint_{D_3} L(\lambda, \hat{\lambda}_{A,c}(x))d\mu \geqslant \frac{1}{s+c-1}-2\eta,$$

which, when combined with (4.18), yields

$$\iint_{D_3} L(\lambda, \hat{\lambda}_{A,c,\lambda_0}(x))d\mu \geqslant \frac{1}{s+c-1}-\eta(M+2). \tag{4.21}$$

If we now realize that for $\lambda > \lambda_0$ the density $h_{A,c,\lambda_0}(\lambda)$ is proportional to $h_{A,c}(\lambda)$ with a coefficient of proportionality greater than 1, we can write

$$r(H_{A,c,\lambda_0}, \hat{\lambda}_{A,c,\lambda_0}) \geqslant \iint_{D_3} L(\lambda, \hat{\lambda}_{A,c,\lambda_0})d\mu \geqslant \frac{1}{s+c-1}-\eta(M+2).$$

Because $\eta > 0$ is arbitrary, we have the conclusion of (d).

Lemma (d) implies in a straightforward manner that the statistician's minimax risk remains greater than or equal to $1/(s-1)$ if we impose the restriction $\lambda_0 \leqslant \lambda < \infty$ on parameter λ. This together with (a) and (b) proves Theorem 2.

5. An alternative proof of minimaxity of estimators indicated in Theorem 1. Consider a one-parameter exponential family of distributions with probability density functions of the form

$$f(x \mid \omega) = C(\omega)e^{\omega x} \tag{5.1}$$

with respect to a fixed measure μ over the real line X of points x. Call the set $\Omega = \{\omega\colon \int_X e^{\omega x} d\mu(x) < \infty\}$ the natural parameter space. It is known that Ω is a connected set (a finite or infinite interval), that $\int_X e^{\omega x} d\mu(x)$ is an analytic function of ω inside Ω and that moments of this distribution can be expressed by $C(\omega)$ and its derivatives. In particular, we have the following expressions for expected value and variance:

$$\begin{gathered}\theta(\omega) = E(x \mid \omega) = \int_X xf(x \mid \omega)d\mu(x) = -\frac{C'(\omega)}{C(\omega)}, \\ D^2(x \mid \omega) = \int_X \big(x-\theta(\omega)\big)^2 f(x \mid \omega)d\mu(x) = -\frac{C''(\omega)C(\omega)-C'^2(\omega)}{C^2(\omega)} = \theta'(\omega).\end{gathered} \tag{5.2}$$

Suppose now we want to estimate $\theta(\omega)$ on the basis of observing x. It is immediate that x is an unbiased estimator of $\theta(\omega)$. Is it also minimax and/or admissible? The answer depends on the choice of the loss function. If we define it by

$$L\big(\theta(\omega), \hat{\theta}(x)\big) = \big(\theta(\omega)-\hat{\theta}(x)\big)^2/D^2(x \mid \omega), \tag{5.3}$$

then $\hat{\theta}(x) = x$ is evidently an estimator with constant risk. Girshick and Savage [6] proved the minimaxity of this estimator by showing that it is admissible, provided $\Omega = \{-\infty < \omega < \infty\}$.

The most general result concerning the admissibility of estimators of expected value in exponential families of distributions is seemingly due to Karlin [8], whose main theorem has been proved in another way by Cheng Ping [1]. It is this:

THEOREM (Karlin). *Consider an exponential family of distributions of the form* (5.1) *with the interior of Ω equal to $\{\omega\colon a < \omega < b\}$ (one or both of the a and b might be infinite), and let the loss be given by* (5.3). *Let $a < \omega_0 < b$.*

Then

$$\hat{\theta}(x) = \frac{x+ka}{1+a}, \tag{5.4}$$

where $a \neq -1$ and k are constants, is an admissible estimator of $\theta(\omega)$, if

$$\lim_{\omega_1 \to a} \int_{\omega_1}^{\omega_0} C^{-a}(\omega) e^{-ka\omega} d\omega = \infty \tag{5.5}$$

and

$$\lim_{\omega_2 \to 0} \int_{\omega_0}^{\omega_2} C^{-a}(\omega) e^{-ka\omega} d\omega = \infty. \tag{5.6}$$

Now, as shown by the example quoted in section 1, it might well happen that if we restrict the variability of the parameter to a proper subset of Ω, a formerly minimax and admissible estimator is still minimax without being admissible any more; and some other estimators may be minimax then. So one might want to find instead a similarly general condition ensuring minimaxity of an estimator. We find a result of that sort in Cheng Ping [1]:

THEOREM (Cheng Ping). *Consider an exponential family of distribution of the form* (5.1) *and suppose ω ranges over a subset $\{\omega: a < \omega < b\}$ of the natural parameter space. Let the loss function be given by* (5.3). *Then $\hat{\theta}(x)$ given by* (5.4) *is a minimax estimator of $\theta(\omega)$ if* (5.5) *holds and we have*

$$\sup_{a<\omega<b} \frac{(\theta(\omega)-k)^2}{D^2(x \mid \omega)} = \lim_{\omega \to a} \frac{(\theta(\omega)-k)^2}{D^2(x \mid \omega)} = \frac{1}{a}, \qquad a \geqslant 0, \tag{5.7}$$

and

$$\underline{\lim}_{\omega \to a} C^a(\omega) e^{ka\omega} D(x \mid \omega) = c > 0, \tag{5.8}$$

or else, if (5.6) *holds and we have*

$$\sup_{a<\omega<b} \frac{(\theta(\omega)-k)^2}{D^2(x \mid \omega)} = \lim_{\omega \to b} \frac{(\theta(\omega)-k)^2}{D^2(x \mid \omega)} = \frac{1}{a}, \qquad a \geqslant 0, \tag{5.9}$$

and

$$\underline{\lim}_{\omega \to b} C^a(\omega) e^{ka\omega} D(x \mid \omega) = c > 0. \tag{5.10}$$

Now let us return to our Theorem 1. Densities

$$g(x \mid \lambda) = \frac{\lambda^s}{\Gamma(s)} e^{-\lambda x} x^{s-1}$$

are evidently of the form (5.1) with $\omega = \lambda$, $C(\lambda) = \lambda^s/\Gamma(s)$, $d\mu(x) = x^{s-1}dx$ and $\Omega = \{\lambda\colon 0 < \lambda < \infty\}$. In Theorem 1, we were actually interested not in estimation of the expected value $\theta(\lambda) = s/\lambda$, but of $1/\lambda$, and so to prove the minimaxity of estimators of $1/\lambda$ given by (3.5) and (3.6) we have to show that if λ ranges over $0 < \lambda \leqslant \lambda_0$, then

$$\hat{\theta}(x) = \frac{s}{s+1}x+b = \frac{x+b\dfrac{s+1}{s}}{\dfrac{s+1}{s}} = \frac{x+b(s+1)\dfrac{1}{s}}{1+\dfrac{1}{s}}$$

with

$$0 \leqslant b \leqslant \frac{2s}{(s+1)} \tag{5.11}$$

are minimax estimators of $\theta(\lambda)$.

According to Cheng Ping theorem, we have to verify the conditions (5.5), (5.7) and (5.8), which in our case become

$$\lim_{\lambda_1\to 0}\int_{\lambda_1}^{\lambda_0}\left(\frac{\lambda^s}{\Gamma(s)}\right)^{-1/s} e^{-b\lambda(s+1)/s}d\lambda, \tag{5.12}$$

$$\sup_{0<\lambda\leqslant\lambda_0}\frac{\big((s/\lambda)-b(s+1)\big)^2}{s/\lambda^2} = \lim_{\lambda\to 0}\frac{\big((s/\lambda)-b(s+1)\big)^2}{s/\lambda^2} = s, \tag{5.13}$$

$$\varliminf_{\lambda\to 0}\left(\frac{\lambda^s}{\Gamma(s)}\right)^{1/s} e^{b(s+1)\frac{1}{s}\lambda}\cdot\frac{\sqrt{s}}{\lambda} = c > 0, \tag{5.14}$$

and are easily checked to hold true. Especially (5.13) holds if and only if b is within the limits (5.11).

Thus the minimaxity of the estimators in question follows.

A little bit more generally, we can see that in our case the three Cheng Ping's conditions (5.5), (5.7) and (5.8) are fulfilled only by $a = 1/s$ and $k = b(s+1)$ with b within the limits (5.11).

Let us note, however, that Cheng Ping states formally only a sufficient condition for an estimator of the form (5.4) to be minimax, and it is not clear to what extent his conditions are also necessary. On the other hand, by our straighforward analysis in section 3 we were able to state that estimators described by (3.5) and (3.6) are *the only* minimax estimators *linear* in x.

As regards Theorem 2, λ is not the expected value of the considered exponential family of distributions, so Cheng Ping's theorem does not apply.

References

[1] Cheng Ping, *Minimax estimates of parameters of distributions belonging to an exponential family*, Chinese Mathematics 5 (1964), pp. 277-299.

[2] Czen Pin, *O minimaksowym estymatorze liczności populacji*, Zastosow. Mat. 6 (1962), pp. 137-148.

[3] A. Dvoretzky, J. Kiefer and J. Wolfowitz, *Sequential decision problems for processes with continuous time parameter. Testing hypotheses*, Annals of Mathematical Statistics 24 (1953), pp. 254-264; *Sequential decision problems for processes with continuous time parameter. Problems of estimation*, Annals of Mathematical Statistics 24 (1953), pp. 403-415.

[4] Dzan Dzo-i, *O minimaksowej estymacji parametru rozkladu dwumianowego*, Zastosow. Mat. 6 (1961), pp. 31-42.

[5] R. H. Farell, *Estimators of a location parameter in the absolutely continuous case*, Annals of Mathematical Statistics 35 (1964), pp. 949-998.

[6] M. A. Girshick and L. J. Savage, *Bayes and minimax estimates for quadratic loss functions*, Proc. 2nd Berkeley Symposium on Math. Stat. and Prob., Berkeley and Los Angeles 1951, pp. 53-73.

[7] J. L. Hodges, Jr., and E. L. Lehmann, *Some applications of the Cramer-Rao inequality*, Proc. 2nd Berkeley Symposium on Math. Stat. and Prob., Berkeley and Los Angeles 1951, pp. 13-22.

[8] S. Karlin, *Admissibility for estimation with quadratic loss*, Annals of Math. Stat. 29 (1958), pp. 406-436.

[9] E. L. Lehmann, *Notes of the theory of estimation* (*recorded by C. Blyth*), Associated Students Store, University of California, Berkeley 1950.

[10] S. Zubrzycki, *O minimaksowym szacowaniu liczności populacji*, Zastosow. Mat. 7 (1963), pp. 183-194.

THE MATHEMATICAL INSTITUTE OF THE POLISH ACADEMY OF SCIENCES
UNIVERSITY OF WASHINGTON

Received on 18. 11. 1965

S. ZUBRZYCKI (Wrocław)

JAWNE WZORY NA MINIMAKSOWE ESTYMATORY DOPUSZCZALNE W PRZYPADKU PEWNYCH OGRANICZEŃ NAŁOŻONYCH NA PARAMETR

STRESZCZENIE

Wiadomo, że estymator, który jest minimaksowy lub dopuszczalny, gdy parametr przebiega całą naturalną przestrzeń parametrów, może utracić jedną lub obie te własności, gdy zmienność parametru zostanie ograniczona do właściwego podzbioru naturalnej przestrzeni parametrów. W niniejszej pracy dyskutuje się szczegółowo dwa przypadki, kiedy parametr jest w taki sposób ograniczony, a jednak udaje się otrzymać jawne wzory na minimaksowe estymatory, które są ponadto dopuszczalne

w pewnej ograniczonej klasie estymatorów. Idzie mianowicie o estymowanie parametru $\beta = 1/\lambda$ w rozkładzie gamma o gęstości

$$g(x \mid \lambda) = \lambda^s x^{s-1} e^{-\lambda x}/\Gamma(s),$$

gdy wiadomo, że parametr λ spełnia nierówności $0 < \lambda \leqslant \lambda_0 < \infty$, i o estymowanie parametru λ, gdy wiadomo, że spełnia on nierówności $0 < \lambda_0 \leqslant \lambda < \infty$.

С. ЗУБЖИЦКИ (Вроцлав)

РАЗВЁРНУТЫЕ ФОРМУЛЫ МИНИМАКСНЫХ ДОПУСТИМЫХ ОЦЕНОК В СЛУЧАЕ НЕКОТОРЫХ ОГРАНИЧЕНИЙ НАЛОЖЕННЫХ НА ПАРАМЕТР

РЕЗЮМЕ

Известно, что оценка, являющаяся минимаксной или допустимой, когда параметр принимает значения из всего естественного пространства параметров, может потерять одно из этих свойств или даже оба, если ограничить область изменения параметра до некоторого собственного подмножества естественного пространства параметров. В этой работе рассматривается детально два случая, когда параметр ограничен вышеуказанным образом, но когда всё таки удаётся получить развёрнутые формулы минимаксных оценок, которые являются сверх того допустимыми в некотором ограниченном классе оценок. Дело касается именно оценки параметра $\beta = 1/\lambda$ в распределении гамма с плотностью

$$g(x \mid \lambda) = \lambda^s x^{s-1} e^{-\lambda x}/\Gamma(s),$$

когда известно, что параметр λ удовлетворяет неравенствам $0 < \lambda \leqslant \lambda_0 < \infty$, и оценки параметра λ, когда известно, что удовлетворяет он неравенствам $0 < \lambda_0 \leqslant \lambda < \infty$.
