Abstract. Let $Y$ be a random vector taking its values in a measurable space and let $z$ be a vector-valued function defined on that space. We consider gamma minimax estimation of the unknown expected value $p$ of the random vector $z(Y)$. We assume a weighted squared error loss function.

1. Introduction. Let $Y$ be a random variable (or vector) taking its values in a measurable space $(\mathcal{Y}, \mathcal{B})$, whose unknown distribution $P$ is assumed to be an element of the set

$$\mathcal{P} = \{\text{all probability measures on } (\mathcal{Y}, \mathcal{B})\}.$$ 

Further, let $Y^n = (Y_1, \ldots, Y_n)$ be a random sample of size $n$ from $P$ and let $z = (z_1, \ldots, z_k)^T$ be a bounded, measurable function on $(\mathcal{Y}, \mathcal{B})$ with values in $(\mathbb{R}^k, \mathcal{B}_{\mathbb{R}^k})$. Consider estimation of the unknown vector $p$ defined as the expected value of the random vector $Z = z(Y)$, i.e.

$$p = E_P Z.$$ 

We assume that the loss function, which describes the loss to the statistician if he estimates $p$ by $d$, has the form

$$L(d, P) = (d - p)^T C (d - p),$$ 

where the $k \times k$ matrix $C = [c_{ij}]$ is symmetric and nonnegative definite. To choose a reasonable decision rule $d \in \mathcal{D}$, where

$$\mathcal{D} = \{\text{all estimators } d = d(Y^n) \text{ of the unknown vector } p\},$$ 

we can use different principles. If we have no prior information on the unknown probability $P$ then we can use the minimax principle. Let $R(d, P)$ be the risk function of an estimator $d \in \mathcal{D}$, i.e.

$$R(d, P) = E_P [L(d(Y^n), P)].$$

2000 Mathematics Subject Classification: 62C20, 62G05.

Key words and phrases: gamma minimax estimation, nonparametric estimation, Dirichlet process.
Then a decision rule \( d_0 \) is said to be minimax if it minimizes the maximum expected loss, i.e.

\[
\sup_{P \in \mathcal{P}} R(d_0, P) = \inf_{d \in \mathcal{D}} \sup_{P \in \mathcal{P}} R(d, P).
\]

In Wilczyński (1992) it was proved that the minimax estimator of \( p \) under the loss function (1) has the form

\[
d_0(Y^n) = \frac{X^n + \sqrt{n} p_0}{n + \sqrt{n}}, \quad \text{where} \quad X^n = \sum_{j=1}^{n} z(Y_j).
\]

For the definition of the vector \( p_0 \) see Wilczyński (1992). Note that \( d_0 \) is affine (inhomogeneous linear) with respect to \( X^n \) and thus easy to evaluate and handle analytically.

If we know, on the other hand, that \( P \) is a random probability measure chosen according to a known prior distribution \( \pi \in \Pi \), where

\[
\Pi = \{ \text{all priors on the space of all probability measures on } (\mathcal{Y}, \mathcal{B}) \},
\]

we can use the (nonparametric) Bayes principle. Let \( r(d, \pi) \) be the \( \pi \)-Bayes risk of an estimator \( d \), i.e. the expected value of the risk function \( R(d, P) \) with respect to the prior \( \pi \):

\[
r(d, \pi) = E_\pi R(d, P).
\]

Then a decision rule \( d_\pi \) is said to be \( \pi \)-Bayes if it minimizes the \( \pi \)-Bayes risk, i.e.

\[
r(d_\pi, \pi) = \inf_{d \in \mathcal{D}} r(d, \pi).
\]

Unfortunately, finding a workable prior distribution \( \pi \) defined on the space of all probability measures on a given sample space is not an easy task. Ferguson (1973) stated that there are two desirable, but antagonistic, properties of a prior distribution for nonparametric problems: its support should be large and the posterior distribution given a sample of observations from the true probability distribution should be manageable analytically. The simplest priors which have the latter property are the Dirichlet processes introduced by Ferguson (1973). There are a large number of such processes, one for each finite nonnull measure on \( (\mathcal{Y}, \mathcal{B}) \). Suppose that \( \pi \) is the Dirichlet process corresponding to a measure \( A Q \), where \( A \) is a positive number and \( Q \in \mathcal{P} \). Then, from Ferguson (1973) (example b), the \( \pi \)-Bayes nonparametric estimator of \( p \) has the form

\[
d_\pi(Y^n) = \frac{X^n + A q}{n + A},
\]

where \( q = E_Q Z \).

There is an intermediate approach between the Bayes and the minimax principles, the \( \Gamma \)-minimax principle, which is appropriate in the following
situation. Suppose that $P$ is a random probability measure chosen according to an unknown prior distribution $\pi$ which belongs to a given subset $\Gamma$ of $\Pi$. Then a decision rule $d_{\Gamma}$ is said to be $\Gamma$-minimax if it minimizes the maximum Bayes risk with respect to the elements of $\Gamma$, i.e.

$$\sup_{\pi \in \Gamma} r(d_{\Gamma}, \pi) = \inf_{d \in D} \sup_{\pi \in \Gamma} r(d, \pi).$$

(2)

In this paper we consider $\Gamma$-minimax estimation of an unknown vector $p$ under the loss function (1). We assume that the set $\Gamma$ has the form

$$\Gamma = \{\pi \in \Pi : (\nu_1(\pi), \nu_2(\pi)) \in G\},$$

where $G$ is a given convex subset of $\mathbb{R}^{k+1}$ and $\nu_1(\pi)$ and $\nu_2(\pi)$ denote the first moments of $E_PZ$ and $E_PZ^TCZ$ with respect to the prior $\pi \in \Pi$, i.e.

$$\nu_1(\pi) = E_\pi(E_PZ), \quad \nu_2(\pi) = E_\pi(E_PZ^TCZ).$$

We prove that the $\Gamma$-minimax estimator $d_{\Gamma}$ is an affine transformation of the random vector $X^n$.

As is well known, a decision rule which is $\Gamma$-minimax for $\Gamma = \Pi$ has another optimal property—it is also minimax. Thus, we generalize our previous result concerning minimax nonparametric estimation (cf. Wilczyński (1992)), because $\Gamma = \Pi$ when $G = \mathbb{R}^{k+1}$.

The problem of $\Gamma$-minimax estimation has been considered by many authors. In particular, the case where $\Gamma$ consists of all distributions whose first two moments are within some given bounds has been described by Jackson et al. (1970), Robbins (1964), Eichenauer et al. (1988) and Chen and Eichenauer-Hermann (1990). A similar set of priors has been chosen by Chen et al. (1991) who have explicitly determined the $\Gamma$-minimax estimator for the unknown parameter $\theta$ of a one-parameter exponential family. This general result has been obtained under the assumption that there exists an unbiased statistic for $\theta$ with variance which is quadratic in the parameter. Next, this result has been strengthened by Eichenauer (1991) who has assumed that the set $\Gamma$ of priors consists of all distributions whose first two moments are within some given convex compact set $G$. A set $\Gamma$ determined by certain moment-type conditions has also been considered in the paper of Magiera (2001), where the aim is to estimate unknown parameters of Markov-additive processes from the data observed up to a random stopping time.

In all the references given above the problem of estimation is parametric and the observed random variable (or vector) $Y$ has a distribution which depends on the unknown parameter $\theta$, which takes its values in a finite-dimensional Euclidean space. In contrast, we consider the nonparametric version of the problem of $\Gamma$-minimax estimation. We assume that the unknown distribution of $Y$ can be described by any probability measure $P$ defined on the measurable space $(\mathcal{Y}, \mathcal{B})$ in which $Y$ takes its values.
2. Gamma minimax estimate. Since the vector-valued function \( z \) is assumed to be bounded there exists a positive number \( M \) such that 
\[
\sup_{y \in Y} \| z(y) \| \leq M,
\]
where \( \| \cdot \| \) denotes the standard norm in \( \mathbb{R}^k \). This implies that the random vector \( Z := z(Y) \) is bounded and takes its values in the convex compact subset \( M \) of \( \mathbb{R}^k \) defined by
\[
M = \{ x \in \mathbb{R}^k : \| x \| \leq M \}.
\]

We denote by \( (\pi_j) \) a sequence of priors from \( \Gamma \) for which
\[
\lim_{j \to \infty} (\nu_2(\pi_j) - \nu_1^T(\pi_j)C\nu_1(\pi_j)) = \sup_{\pi \in \Gamma} (\nu_2(\pi) - \nu_1^T(\pi)C\nu_1(\pi)).
\]

Since \( Z \in M \), the corresponding sequence
\[
(\nu_1(\pi_j)) = (E_{\pi_j}(E_PZ))
\]
takes its values in \( M \) and therefore has a cluster point \( p_\Gamma \in M \). The following theorem is the main result of the paper:

**Theorem 1.** The \( \Gamma \)-minimax estimator of the unknown vector \( p \) under the loss function (1) has the form
\[
d_\Gamma(Y^n) = \frac{X^n + \sqrt{n}p_\Gamma}{n + \sqrt{n}},
\]
and its \( \Gamma \)-minimax risk equals
\[
\sup_{\pi \in \Gamma} r(d_\Gamma, \pi) = \sup_{\pi \in \Gamma} \frac{\nu_2(\pi) - \nu_1^T(\pi)C\nu_1(\pi)}{(\sqrt{n} + 1)^2}.
\]

**Proof.** We will use a method analogous to that in Wilczyński (1992). First we will show that \( d_\Gamma(Y^n) \) is \( \Gamma \)-minimax if the class of estimators is restricted to a subset \( D_0 \subset D \) defined by
\[
D_0 = \left\{ d^b \in D : d^b(Y^n) = \frac{X^n + \sqrt{n}b}{n + \sqrt{n}}, b \in \mathcal{M} \right\}.
\]

Then, using some implications of this fact, we will find the least upper bound for the Bayes risk of \( d_\Gamma(Y^n) \). Finally, we will construct a sequence \( (\pi^*_j) \) of priors from \( \Gamma \) and a sequence of \( \pi^*_j \)-Bayes estimators \( (d_{\pi^*_j}) \) for which the corresponding sequence of Bayes risks \( (r(d_{\pi^*_j}, \pi^*_j)) \) approaches that upper bound. This will complete the proof of the theorem.

We first calculate the risk function for an estimator \( d^b \in D_0 \). We note that \( z(Y_1), \ldots, z(Y_n) \) are i.i.d. random vectors distributed as \( Z = z(Y) \). Therefore,
\[
E_P(X^n - np)^T C(X^n - np) = \sum_{j=1}^n E_P(z(Y_j) - p)^T C(z(Y_j) - p) = nE_P(Z - p)^T C(Z - p) = n(E_PZ^TCZ - p^TCp).
\]
This implies that the risk function of an estimator $d^b$ from $D_0$ has the form

$$R(d^b, P) = \frac{n(E_PZ^TCZ - p^TCp) + (\sqrt{n})^2(b - p)^TC(b - p)}{(n + \sqrt{n})^2}$$

$$= \frac{(E_PZ^TCZ - p^TCp) + (b - p)^TC(b - p)}{(\sqrt{n} + 1)^2}$$

$$= E_PZ^TCZ - 2b^TCp + b^TCb$$

Moreover, for any prior $\pi \in \Pi$ the $\pi$-Bayes risk of $d^b$ is

$$r(d^b, \pi) = \frac{E_PZ^TCZ - 2b^TC\pi p + b^TCb}{(\sqrt{n} + 1)^2}$$

$$= \nu_2(\pi) - 2b^TC\nu_1(\pi) + b^TCb$$

Let the function $r_1 : M \times \Gamma \to [0, \infty]$ be defined by

$$r_1(b, \pi) := r(d^b, \pi).$$

Note that $M$ and $\Gamma$ are convex sets and $M$ is compact. Moreover, for each fixed $\pi \in \Gamma$, $r_1(b, \pi)$ is convex, continuous with respect to $b \in M$, and for each fixed $b \in M$, $r_1(b, \pi)$ is concave (linear) with respect $\pi \in \Gamma$. This means that all the assumptions of the Nikaido theorem (see Aubin (1980), p. 217) are fulfilled and thus there exists a point $b^*$ for which

$$\sup_{\pi \in \Gamma} r_1(b, \pi) = \inf_{b \in M} \sup_{\pi \in \Gamma} r_1(b, \pi) = \sup_{\pi \in \Gamma} \inf_{b \in M} r_1(b, \pi).$$

The last equality implies that the $\Gamma$-minimax risk in $D_0$ equals

$$\inf_{b \in M} \sup_{\pi \in \Gamma} r_1(b, \pi) = \sup_{\pi \in \Gamma} \inf_{b \in M} r_1(b, \pi) = \sup_{\pi \in \Gamma} \frac{\nu_2(\pi) - \nu_1^T(\pi)C\nu_1(\pi)}{(\sqrt{n} + 1)^2},$$

because, for a fixed distribution $\pi \in \Gamma$, the convex function $r_1(b, \pi)$ of the variable $b$ attains its global minimum over $M$ at the point $b(\pi) = \nu_1(\pi)$. Now it remains to prove that $b = p_\Gamma$. Set, for simplicity,

$$r_2(\pi) = \nu_2(\pi) - \nu_1^T(\pi)C\nu_1(\pi).$$

Let $(\pi_j)$ be a sequence of priors satisfying (3). Since the functions $\nu_1(\pi)$ and $\nu_2(\pi)$ are linear in $\pi$, an easy calculation shows that for any $\pi \in \Gamma$ and $0 < \beta < 1$,

$$\sup_{\pi \in \Gamma} r_2(\pi) \geq r_2(\beta \pi + (1 - \beta)\pi_j) = \beta r_2(\pi) + (1 - \beta)r_2(\pi_j)$$

$$+ \beta(1 - \beta)(\nu_1(\pi_j) - \nu_1(\pi))^T C(\nu_1(\pi_j) - \nu_1(\pi)).$$

This implies that

$$\sup_{\pi \in \Gamma} r_2(\pi) \geq \beta r_2(\pi) + (1 - \beta) \sup_{\pi \in \Gamma} r_2(\pi) + \beta(1 - \beta)(p_\Gamma - \nu_1(\pi))^T C(p_\Gamma - \nu_1(\pi)), $$

$$= \infty.$$
because $p_\Gamma$ is a cluster point of the sequence $(\nu_1(\pi_j))$, and $\lim_{j \to \infty} r_2(\pi_j) = \sup_{\pi \in \Gamma} r(\pi)$ by (3). Therefore,

$$\beta \sup_{\pi \in \Gamma} r_2(\pi) \geq \beta r_2(\pi) + \beta(1 - \beta)(p_\Gamma - \nu_1(\pi))^T C(p_\Gamma - \nu_1(\pi)),$$

and since $\beta$ is positive,

$$\sup_{\pi \in \Gamma} r_2(\pi) \geq r_2(\pi) + (1 - \beta)(p_\Gamma - \nu_1(\pi))^T C(p_\Gamma - \nu_1(\pi)).$$

Letting $\beta \to 0^+$, we can see by (9) that

$$\sup_{\pi \in \Gamma} r_2(\pi) \geq r_2(\pi) + (1 - \beta)(p_\Gamma - \nu_1(\pi))^T C(p_\Gamma - \nu_1(\pi)) = \nu_2(\pi) - 2p_\Gamma^T C\nu_1(\pi) + p_\Gamma^T C p_\Gamma,$$

which implies by (7) that

$$\sup_{\pi \in \Gamma} \frac{\nu_2(\pi) - \nu_1^T(\pi) C \nu_1(\pi)}{(\sqrt{n} + 1)^2} \geq \frac{\nu_2(\pi) - 2p_\Gamma^T C\nu_1(\pi) + p_\Gamma^T C p_\Gamma}{(\sqrt{n} + 1)^2} = r(d_\Gamma, \pi).$$

Because this is true for all $\pi \in \Gamma$, it follows from (8) that

$$\sup_{\pi \in \Gamma} r(d_\Gamma, \pi) \leq \sup_{\pi \in \Gamma} \frac{\nu_2(\pi) - \nu_1^T(\pi) C \nu_1(\pi)}{(\sqrt{n} + 1)^2} = \inf_{b \in M, \pi \in \Gamma} \sup_{\beta} r_1(b, \pi) = \inf_{b \in M, \pi \in \Gamma} \sup_{d \in D_0, \pi \in \Gamma} r(d, \pi).$$

This implies that $d_\Gamma(Y^n)$ is $\Gamma$-minimax if the class of estimators is restricted to $D_0$ of $D$.

To complete the proof we will construct a sequence $(\pi_j^*)$ of priors from $\Gamma$ and a sequence $(d_{\pi_j^*})$ of $\pi_j^*$-Bayes estimators for which

$$\lim_{j \to \infty} r(d_{\pi_j^*}, \pi_j^*) = \sup_{\pi \in \Gamma} \frac{\nu_2(\pi) - \nu_1^T(\pi) C \nu_1(\pi)}{(\sqrt{n} + 1)^2} = \sup_{\pi \in \Gamma} r(d_\Gamma, \pi).$$

Let $(\pi_j)$ be a sequence of priors from $\Gamma$ satisfying (3) and let $(P_j)$ be a sequence of probability measures on $(\mathcal{Y}, \mathcal{B})$ such that

$$\bigwedge_{A \in \mathcal{B}} P_j(A) = E_{\pi_j}(P(A)).$$

For each $j \geq 1$ we denote by $\pi_j^*$ a Dirichlet prior process on $(\mathcal{Y}, \mathcal{B})$ with parameter $\beta_j = \sqrt{n} P_j$. To prove that $\pi_j^* \in \Gamma$ we note first that by Ferguson (1973) (Theorems 3 and 4),

$$E_{\pi_j^*}[E_P Z^T C Z] = E_{P_j} Z^T C Z, \quad E_{\pi_j^*} p = E_{\pi_j^*}[E_P Z] = E_{P_j} Z.$$

Since by the definition of the probability measure $P_j$,

$$E_{P_j} Z = E_{\pi_j^*}[E_P Z] = \nu_1(\pi_j), \quad E_{P_j} Z^T C Z = E_{\pi_j^*}[E_P Z^T C Z] = \nu_2(\pi_j),$$
we deduce that

$$\nu_1(\pi^*_j) = E_{\pi^*_j}[E_P Z] = \nu_1(\pi_j), \quad \nu_2(\pi^*_j) = E_{\pi^*_j}[E_P Z^T C Z] = \nu_2(\pi_j).$$

This obviously implies that $\pi^*_j \in \Gamma$, because $\pi_j \in \Gamma$. Moreover, from Ferguson (1973) (example b), the $\pi^*_j$-Bayes nonparametric estimator of $p = E_P Z$ has the form

$$d_{\pi^*_j}(Y^n) = \frac{\sqrt{n}}{n + \sqrt{n}} E_P Z + \frac{n}{n + \sqrt{n}} \frac{1}{n} \sum_{j=1}^{n} z(Y_j) = \frac{X^n + \sqrt{n} \nu_1(\pi_j)}{n + \sqrt{n}},$$

because $E_P Z = \nu_1(\pi_j)$. To calculate the $\pi^*_j$-Bayes risk $r(d_{\pi^*_j}, \pi^*_j)$ we note that $d_{\pi^*_j}(Y^n) = d^b(Y^n) \in D_0$ with $b = \nu_1(\pi_j)$. Thus, by (7),

$$r(d_{\pi^*_j}, \pi^*_j) = \frac{\nu_2(\pi^*_j) - 2\nu^T_1(\pi^*_j) C \nu_1(\pi^*_j) + \nu_1^T(\pi^*_j) C \nu_1(\pi^*_j)}{(\sqrt{n} + 1)^2}$$

$$= \frac{\nu_2(\pi_j) - \nu^T_1(\pi_j) C \nu_1(\pi_j)}{(\sqrt{n} + 1)^2},$$

because $\nu_1(\pi^*_j) = \nu_1(\pi_j)$ and $\nu_2(\pi^*_j) = \nu_2(\pi_j)$. Therefore, by (10),

$$\inf_{d \in D} \sup_{\pi \in \Gamma} r(d, \pi) = \lim_{j \to \infty} r(d_{\pi^*_j}, \pi^*_j) = \lim_{j \to \infty} \frac{\nu_2(\pi_j) - \nu^T_1(\pi_j) C \nu_1(\pi_j)}{(\sqrt{n} + 1)^2}$$

$$\geq \sup_{\pi \in \Gamma} \frac{\nu_2(\pi) - \nu^T_1(\pi) C \nu_1(\pi)}{(\sqrt{n} + 1)^2} \geq \sup_{d \in D} \sup_{\pi \in \Gamma} r(d, \pi),$$

which implies that the estimator $d_F(Y^n)$ is $\Gamma$-minimax and its $\Gamma$-minimax risk is given by (6). This completes the proof of Theorem 1. $

3. Generalization. In this section we present a slight generalization of Theorem 1. Instead of assuming that the function $z$ is bounded on $Y$, we suppose that a weaker condition is fulfilled: $\sup_{y \in Y} \|C^{1/2} z(y)\| < \infty$, where $C^{1/2}$ is the square root of the matrix $C$, i.e. $C^{1/2}C^{1/2} = C$. Then the random vector

$$Z^* := C^{1/2} z(Y) = C^{1/2} Z$$

is bounded, which implies that for each affine estimator $d^b \in D_0$ its risk function $R(d^b, P)$ is bounded for $P \in \mathcal{P}$. Let $\nu_1^*(\pi)$ and $\nu_2^*(\pi)$ denote the first moments of $E_P Z^*$ and $E_P (Z^*)^T Z^* = E_P \|Z^*\|^2$ with respect to a prior $\pi \in \Pi$, i.e.

$$\nu_1^*(\pi) = E_{\pi}(E_P Z^*), \quad \nu_2^*(\pi) = E_{\pi}(E_P \|Z^*\|^2),$$

and let $(\pi_j)$ be a sequence of priors from $\Gamma$ for which
\[
\lim_{j \to \infty} (\nu_j^2(\pi_j) - \|\nu_j^* (\pi_j)\|^2) = \sup_{\pi \in \Gamma} (\nu^*_2(\pi) - \|\nu^*_1(\pi)\|^2).
\]

Then, by the same arguments as in the previous section, the sequence \((\nu_j^*(\pi_j))\), where
\[
\nu_j^* (\pi_j) = E_{\pi_j}(E_{\Gamma} Z^*), \quad j \geq 1,
\]
has a cluster point \(p_j^*\). Since \(p_j^*\) belongs to the linear space generated by the columns of the matrix \(C^{1/2}\), there exists a vector \(p_\Gamma\) for which
\[
C^{1/2} p_\Gamma = p_j^*.
\]

The following theorem generalizes the results of the previous section.

**Theorem 2.** Suppose that \(\sup_{y \in Y} \|C^{1/2} z(y)\| < \infty\). Then the \(\Gamma\)-minimax estimator of the unknown vector \(p\) under the loss function (1) has the form
\[
d_\Gamma(Y^n) = \frac{X^n + \sqrt{n} p_\Gamma}{n + \sqrt{n}},
\]
where \(p_\Gamma\) is any solution of (12). Moreover, the \(\Gamma\)-minimax risk for \(d_\Gamma\) is
\[
\sup_{\pi \in \Gamma} r(d_\Gamma, \pi) = \sup_{\pi \in \Gamma} \frac{\nu_2^*(\pi) - \|\nu_1^*(\pi)\|^2}{(\sqrt{n} + 1)^2}.
\]

Proof. Let the random vector \(X^{*n}\) be defined by
\[
X^{*n} := C^{1/2} X^n = \sum_{j=1}^n z^*(Y_j).
\]

As can easily be seen, it suffices to show that the decision rule \(d_\Gamma^*(Y^n) = C^{1/2} d_\Gamma(Y^n)\), which by (13) and (12) has the form
\[
d_\Gamma^*(Y^n) = \frac{X^{*n} + \sqrt{n} p_\Gamma}{n + \sqrt{n}},
\]
is the \(\Gamma\)-minimax estimator of the vector \(p^* = C^{1/2} p = E_{\Gamma} Z^*\) under the loss function
\[
L^*(d^*, P) = (d^* - p^*)^T (d^* - p^*) = \|d^* - p^*\|^2.
\]

This, however, can be easily deduced from Theorem 1. Moreover, since
\[
\bigwedge_{P \in \mathcal{P}} L^*(d_\Gamma^*, P) = L(d_\Gamma, P),
\]
the estimators \(d_\Gamma^*\) and \(d_\Gamma\) have the same risk functions, and (6) yields (14).

**4. Example.** Finding analytically the cluster point \(p_\Gamma\) is not an easy task. However, in the following example this can easily be done.
Example. Suppose that the set $\mathcal{Y}$ is centrosymmetric about $0$ and that
\begin{equation}
    z(y) = -z(-y), \quad y \in \mathcal{Y}, \quad (\nu_1, \nu_2) \in \mathcal{G} \iff (-\nu_1, \nu_2) \in \mathcal{G}.
\end{equation}
Let $P^-$ stand for the distribution of the random vector $-Y$, whenever $Y$ is distributed according to $P$. For any prior $\pi \in \Pi$ we denote by $\pi^-$ its modified version in which each probability distribution $P$ chosen by $\pi$ is replaced by $P^-$. The assumption (15) implies that $\pi \in \Gamma \iff \pi^- \in \Gamma$, because
\begin{align*}
    \nu_2(\pi^-) &= \nu_2(\pi), \quad \nu_1(\pi^-) = -\nu_1(\pi).
\end{align*}
Now, let $(\pi_j)$ be a sequence of priors from $\Gamma$ satisfying (3). Then for each $j \geq 1$, the prior $\bar{\pi}_j = \frac{1}{2}(\pi_j + \pi_j^-)$ belongs to $\Gamma$, because $\bar{\pi}_j \in \Gamma$, $\pi_j^- \in \Gamma$ and the set $\Gamma$ is convex. Moreover, since $\nu_2(\bar{\pi}_j) = \nu_2(\pi_j)$ and $\nu_1(\bar{\pi}_j) = 0$, we conclude that
\begin{align*}
    \nu_2(\bar{\pi}_j) - \nu_2^T(\bar{\pi}_j) C \nu_1(\bar{\pi}_j) &= \nu_2(\pi_j) - \nu_2^T(\pi_j) C \nu_1(\pi_j).
\end{align*}
This implies that the sequence $(\bar{\pi}_j)$ also satisfies (3). Therefore, the estimator
\begin{align*}
    d_\Gamma(Y^n) = \frac{X^n}{n + \sqrt{n}}
\end{align*}
is $\Gamma$-minimax, because $p_\Gamma = \lim_{j \to \infty} \nu_1(\bar{\pi}_j) = 0$.

Acknowledgements. The author wishes to thank the referee for his helpful comments which improved the original version of the paper.

References

L. Chen and J. Eichenauer-Herrmann (1990), Gamma-minimax estimation of a multivariate normal mean, Metrika 37, 1–6.

Institute of Mathematics
Wrocław University of Technology
Wybrzeże Wyspiańskiego 27
50-370 Wrocław, Poland
E-mail: wilczyn@im.pwr.wroc.pl

Received on 21.9.2001;
revised version on 8.7.2002