陈希孺统计文选

Chen Xiru tongji wenxuan

陈希孺 / 著

中国科学技术大学出版社
2003 · 合肥

图书在版编目（CIP）数据

陈希孺统计文选 / 陈希孺著. —合肥：中国科学技术大学出版社，2003.10
ISBN 7-312-01365-1

Ⅰ. 陈… Ⅱ. 陈… Ⅲ. ①陈希孺—文集 ②数理统计—文集 Ⅳ. O212-53

中国版本图书馆 CIP 数据核字（2003）第 071749 号

责任编辑：李攀峰
装帧设计：黄　彦

出版发行
中国科学技术大学出版社
（安徽省合肥市金寨路 96 号，230026）

照排
南京理工排版校对有限公司

印刷
合肥远东印务有限责任公司

经销
全国新华书店

开本
787×1092/16

印张
18

字数
461 千

版次
2003 年 10 月第 1 版

印次
2003 年 10 月第 1 次印刷

印数
1–1 000 册

定价
68.00 元

出版说明

本书是陈希孺院士的自选集，包含了他自20世纪60年代以来所发表的百余篇论文中的25篇，内容涉及参数估计、线性统计模型（LS估计、M估计等）、非参数统计（U-统计量、线性秩统计量、非参数回归和密度估计等）等领域，可以反映他在这些领域工作的基本成果.

本书的出版得到了中国科学技术大学科技处、出版社和统计与金融系等有关方面的大力支持和协助，具体工作主要是由目前在中国科学技术大学工作的陈希孺院士的学生们完成的. 这件事自2001年开始酝酿，历时两年多，在上述各方面通力合作之下，得以顺利完成，特书此以志其事.

目　　次

1

On Minimax Invariant Estimation of Scale and Location Parameters

1 Introduction

Let X be a random variable whose distribution function $F(x)$ is known. By changing the origin and scale we modify linearly the results of measurements. This leads us to consider a family of distributions $\left\{F\left(\frac{x-b}{a}\right)\right\}$, where $a > 0$ and b are parameters — the so-called scale and location parameters.

The problem of estimating these parameters has been investigated by many authors since 1937. Roughly Speaking, up to now the efforts have been centred on two points: the seeking of minimax invariant estimators (MIE) and the admissibility of such estimators. The second point has been investigated more successfully in the case when only a location parameter is present (see [4],[5],[6],[10],[13], [14], and [15]). As for the problem of existence of MIE, the present situation is that although a number of results in various particular cases were obtained (see [3],[4],[6],[7],[8], and [12]), the general case that two parameters are present simultaneously has not yet been worked out.

In this section an important paper by J. Kiefer should be here mentioned. In [11] Kiefer developed the general theory of invariance principle in the nonsequential case. His result is so general that the problem of existence of MIE may be considered as a solved one, at least partially. However, the solution by his theory, it seems, cannot be considered a rigorous and complete one, for, first, the main theorem of [11] states only that the restriction to invariant estimators does not enlarge the maximum risk and the problem of existence of MIE is still left open; second, little attention was paid in the paper to the measurability questions. So, although it may happen that in some cases MIE can be formally constructed, a verification is needed to insure the measurability of the estimator obtained. It goes without saying that from the rigorous mathematical viewpoint, it is very desirable to have these problems clarified. Third, Kiefer's theory does not apply to discrete distribution case. Owing to these

本文 1964 年发表于《*Scientia Sinica*》第 13 卷第 10 期.

reasons, it seems worth while to give a rigorous treatment of this problem.

The purpose of this paper is to prove the existence of MIE of scale and location parameters. For simplifying the wording, we shall discuss in detail only the one-dimensional case. The extension to multidimensional case will be indicated in the last section.

The case where the location parameter is the only one present may be treated in a somewhat similar manner (see in this connexion an article published by the author in *Acta Mathematica Sinica*, Vol. 14, No. 2, pp. 276～290).

2 The Main Results

Let $X = (X_1, \cdots, X_N)$ be a random vector with a known distribution F, and let $\varepsilon = (1, \cdots, 1)$. Then for any $a > 0$ and b, $aX + b\varepsilon$ has a distribution function

$$F(x \mid a, b) = F\left(\frac{x - b\varepsilon}{a}\right).$$

Now we draw a sample x from the population $F(x|a, b)$, and want to make some estimation of the parameter vector (a, b).

Let $V(a, b; d_1, d_2)$ be the loss function. We impose the natural condition $d_1 \geqslant 0$, and suppose that V is invariant, i. e., for any $c_1 > 0$ and c_2, we have

$$V(ac_1, bc_1 + c_2; d_1c_1, d_2c_1 + c_2) = V(a, b; d_1, d_2).$$

This is equivalent to the condition that there exists a function $W = W(t_1, t_2)$ defined on set $\Sigma = \{(t_1, t_2) \mid t_1 \geqslant 0\}$ such that

$$V(a, b; d_1, d_2) = W\left(\frac{d_1}{a}, \frac{d_2 - b}{a}\right).$$

In what follows the term "loss function" will be understood as the function W, and not V itself.

We call "estimator" any Borel measurable function defined on R_N with range Σ. The class of all estimators will be denoted by μ. An estimator (u, v) is called invariant if for $(c, d) \in \Sigma^0$ (Σ^0— interior of Σ) we have

$$u(cx + d\varepsilon) = cu(x), \quad v(cx + d\varepsilon) = cv(x) + d.$$

Let (u_0, v_0), where $u_0(x) > 0$, be a fixed invariant estimator. For any $(u, v) \in \mu$, we define

$$s(x) = \frac{u(x)}{u_0(x)}, \quad r(x) = \frac{v(x) - v_0(x)}{u_0(x)};$$

then $(s, r) \in \mu$, and the correspondence between (u, v) and (s, r) is one-to-one. A necessary and sufficient condition that (u, v) is invariant is that

$$s(cx + d\varepsilon) = s(x), \quad r(cx + d\varepsilon) = r(x) \tag{1}$$

for any $(c, d) \in \Sigma^0$. In what follows we shall often use (s, r) instead of (u, v), and the subclass of all functions in μ satisfying (1) will be denoted by ϑ.

The risk function $R(a, b; u, v)$ may conveniently be expressed as a functional ρ of (a, b) and (s, r). Thus

$$\begin{aligned}\rho(s, r; a, b) &= \rho(s, r; a, b; F, W)\\ &=\int W\left(\frac{u(x)}{a}, \frac{v(x)-b}{a}\right)\mathrm{d}F\left(\frac{x-b}{a}\right)\\ &=\int W[s(ay+b\varepsilon)u_0(y), r(ay+b\varepsilon)u_0(y)+v_0(y)]\mathrm{d}F(y).\end{aligned}$$

Notice that if $(s, r)\in\vartheta$, $\rho(s, r; a, b)$ would be independent of (a, b) and might justifiably be written as $\rho(s, r)$. Let

$$V = V(F, W) = \inf_{(s, r)\in\vartheta}\sup_{(a, b)\in\Sigma^0}\rho(s, r; a, b; F, W), \qquad (*)$$

$$\overline{V} = \overline{V}(F, W) = \inf_{(s, r)\in\mu}\sup_{(a, b)\in\Sigma^0}\rho(s, r; a, b; F, W),$$

$$\underline{V} = \underline{V}(F, W) = \sup_{G}\inf_{(s, r)\in\vartheta}\int\rho(s, r; a, b; F, W)\mathrm{d}G,$$

where G is an *a priori* distribution on Σ^0. We remark that in the foregoing expressions it is unnecessary to require that F is a probability measure.

It is easily seen that $V\geqslant\overline{V}\geqslant\underline{V}$. What we want to prove is $V=\overline{V}$ and the inf standing on the right hand of ($*$) is attainable. This will be proved in this paper under mild restrictions. We now formulate the conditions imposed on the loss function W:

1°. W is continuous on Σ^0.

2°. $W(1, 0)=0$.

3°. For any $t_1\geqslant 0$ fixed, W is nonincreasing in t_2 for $t_2\leqslant 0$ and nondecreasing in t_2 for $t_2\geqslant 0$. For any t_2 fixed, $W(t_1, t_2)$ viewed as a function of t_1 nonincreases for $0\leqslant t_1\leqslant 1$ and nondecreases for $t_1\geqslant 1$. Notice that from 2°, 3° it follows that W is nonnegative on Σ.

4°_1. The relation

$$W(t_1, t_2)\to\infty$$

holds uniformly as $\frac{1}{t_1}+t_1+|t_2|\to\infty$. This in conjunction with 3° implies that $W(t_1, t_2)=\infty$ when $t_1=0$.

In many cases it is reasonable to assume that the loss remains bound as the error of estimation remains bound. Accordingly we impose the following conditions.

4°_2. W is continuous on Σ and $W(t_1, t_2)\to\infty$ holds uniformly as $t_1+|t_2|\to\infty$. The important quadratic loss function

$$W(t_1, t_2) = (t_1-1)^2+t_2^2$$

satisfies this condition.

4°_3. W is continuous on Σ, and

$$W(t_1, t_2) \to \alpha < \infty$$

uniformly as $t_1 + |t_2| \to \infty$. An example of such a loss function is furnished by $W(t_1, t_2) = \frac{|1-t_1|+|t_2|}{1+t_1+|t_2|}$ with $\alpha = 1$.

Now we turn to the distribution F. We assume that it satisfies the condition:

$$F(A) = 0, \tag{2}$$

where

$$A = \{x \mid x_1 = \cdots = x_N\}. \tag{3}$$

It is easily seen that $aA + b\varepsilon = \{(ax_1 + b, \cdots, ax_N + b) \mid (x_1, \cdots, x_N) \in A\} = A$ for any $(a, b) \in \Sigma^0$, and so $P\{(ax + b\varepsilon) \in A\} = F(A) = 0$. This being the case, the invariant estimator (u_0, v_0) with $u_0 > 0$ exists and may be taken as

$$u_0(x) = \max_{1 \leqslant i < j \leqslant N} |x_i - x_j|, \quad v_0(x) = x_1.$$

The following may be said about the condition (2). Suppose that (u, v) is an invariant estimator and $x = (x_1, \cdots, x_1)$, then it is easily seen that

$$u(x) = 0, \; v(x) = x_1.$$

But under condition 4_1°, the risk of any such estimator is ∞ if $F(A) > 0$. Therefore, in this case, the condition (2) may be considered as necessary. The situation becomes more complicated when 4_2° or 4_3° holds. Fortunately the most important case where the condition (2) fails is that F is a discrete distribution, and we shall prove that in this case, the MIE exists under 4_2°.

The main results of this paper may be summarized in the following two theorems:

Theorem 1 If the loss function W satisfies conditions 1°, 2°, 3° and one of 4_1°, 4_2°, 4_3°, and if the distribution F satisfies (2), then there exists an MIE of the scale and location parameters.

Theorem 2 If W satisfies 1°, 2°, 3°, 4_2° and if F is discrete on A, i. e., there exists a denumerable set $A_1 \subset A$ such that $F(A_1) = F(A)$, then the conclusion of Theorem 1 still holds true.

In practical applications, the case most commonly dealt with is that F is the direct product of N one-dimensional distributions, all of which belong to the continuous of discrete type. In this case, the condition of F in the above theorems is satisfied.

3 Some Lemmas

The proof of the above theorems are preceeded by several lemmas, which are discussed in the present section. Some of them are of general interest.

Lemma 1 Suppose that the probability distribution F satisfies (2) and the condition

$$F(\{x \mid (u_0(x), v_0(x)) \in L_m\}) = 1, \tag{4}$$

where (u_0, v_0) is a fixed invariant estimation with $u_0(x) > 0$ for all x, and

$$L_m = \left\{(t_1, t_2) \left| \frac{1}{m} \leqslant t_1 \leqslant m, \mid t_2 \mid \leqslant m \right.\right\}.$$

Furthermore, we assume that one of the following conditions is satisfied by the loss function W:

(Ⅰ) W satisfies 2°, 3°, 4_1° and is finite on Σ^0;

(Ⅱ) W satisfies 2°, 3°, finite on Σ, and $\lim W(t_1, t_2) = \infty$ uniformly, as $t_1 + \mid t_2 \mid \to \infty$;

(Ⅲ) W is a bound and Borel measurable function on Σ.

Then we have

$$V(F, W) = \underline{V}(F, W).$$

This lemma① may be considered as an extension of the one proved by Girshick and Savage (see lemma 3.2 in [3]).

Proof For any $q > 1$, $(a, b) \in \Sigma - L_q$, we define

$$(a, b)^* = (a^*, b^*) \tag{5}$$

as the point on L_q with shortest distance to (a, b). For any $(a, b) \overline{\in} L_{2m^2}$ and $(c, d) \in L_m$, we have

$$W(ac, bc + d) \geqslant W(a^* c, b^* c + d),$$

where (a^*, b^*) is determined as above with $q = 2m^2$.

Now define

$$W^*(t_1, t_2) = \begin{cases} W(t_1, t_2), & \text{for } (t_1, t_2) \in L_{m(2m^2+1)}, \\ W(t_1^*, t_2^*), & \text{for } (t_1, t_2) \in \Sigma - L_{m(2m^2+1)}, \end{cases}$$

where (t_1^*, t_2^*) is determined by (5) with $q = m(2m^2+1)$. From the conditions imposed on W, it follows easily that W^* is a Borel measurable bound function, and $W^* \leqslant W$.

By the definition of $\underline{V}$, for any $\varepsilon > 0$ and $T > 1$, there exists $(s, r) \in \mu$ such that

$$\begin{aligned} &\underline{V}(F, W^*) + \varepsilon \\ &\geqslant \frac{1}{\varphi(T)} \int_{L_T} \left[\int W^*[s(ay + b\varepsilon)u_0(y), r(ay + b\varepsilon)u_0(y) + v_0(y)]\mathrm{d}F(y)\right] \frac{1}{a} \mathrm{d}a\mathrm{d}b \\ &= \frac{1}{\varphi(T)} \int \mathrm{d}F(y) \int_{L_T} W^*[s(ay + b\varepsilon)u_0(y), r(ay + b\varepsilon)u_0(y) + v_0(y)] \frac{1}{a} \mathrm{d}a\mathrm{d}b. \end{aligned} \tag{6}$$

① The author is indebted to Comrades Cheng Ping, Chien T T, and Tao Po for their assistance in the proof of this lemma.

Here $\varphi(T)=\int_{L_T}\frac{1}{a}\mathrm{d}a\mathrm{d}b=4T\log T$. But

$$\int_{L_T}W^*\left[s(ay+b\varepsilon)u_0(y),\ r(ay+b\varepsilon)u_0(y)+v_0(y)\right]\frac{\mathrm{d}a\mathrm{d}b}{a}$$
$$\geqslant\int_{1/T}^{T}\frac{1}{a}\mathrm{d}a\int_{-T-av_0(y)}^{T-av_0(y)}W^*\left[s(ay+b\varepsilon)u_0(y),\ r(ay+b\varepsilon)u_0(y)+v_0(y)\right]\mathrm{d}b$$
$$-2\mid v_0(y)\mid TM, \tag{7}$$

where

$$M=\sup_{(t_1,\,t_2)\in\Sigma}W^*(t_1,\,t_2)=\sup\{W(t_1,\,t_2)\mid(t_1,\,t_2)\in L_{m(2m^2+1)}\}<\infty.$$

Also, because of condition (4), we may restrict ourselves to those values of y, for which $(u_0(y),\ v_0(y))\in L_m$. Thus from (7) we have

$$\int_{L_T}W^*\left[s(ay+b\varepsilon)u_0(y),\ r(ay+b\varepsilon)u_0(y)+v_0(y)\right]\frac{\mathrm{d}a\mathrm{d}b}{a}$$
$$\geqslant\int_{[u_0(y)T]^{-1}}^{u_0^{-1}(y)T}\frac{\mathrm{d}a}{a}\int_{-T-av_0(y)}^{T-av_0(y)}W^*\left[s(ay+b\varepsilon)u_0(y),\ r(ay+b\varepsilon)u_0(y)+v_0(y)\right]\mathrm{d}b$$
$$-2mTM-2TM\int_{T}^{mT}\frac{\mathrm{d}a}{a}-2TM\int_{1/mT}^{1/T}\frac{\mathrm{d}a}{a}$$
$$=\int_{E}W^*\left[s(ay+b\varepsilon)u_0(y),\ r(ay+b\varepsilon)u_0(y)+v_0(y)\right]\frac{\mathrm{d}a\mathrm{d}b}{a}$$
$$-2mTM(1+2\log m), \tag{8}$$

where $E=\{(a,\,b)\mid(au_0(y),\ b+av_0(y))\in L_T\}$. Making the transform $a'=au_0(y)$, $b'=av_0(y)+b$ and rewriting $(a,\,b)$ instead of $(a',\,b')$, we obtain

$$\int_{E}W^*\left[s(ay+b\varepsilon)u_0(y),\ r(ay+b\varepsilon)u_0(y)+v_0(y)\right]\frac{1}{a}\mathrm{d}a\mathrm{d}b$$
$$=\int_{L_T}W^*\left[s\left\{\frac{a}{u_0(y)}y+\left(b-\frac{av_0(y)}{u_0(y)}\right)\varepsilon\right\}u_0(y),\right.$$
$$\left.r\left\{\frac{a}{u_0(y)}y+\left(b-\frac{av_0(y)}{u_0(y)}\right)\varepsilon\right\}u_0(y)+v_0(y)\right]\frac{1}{a}\mathrm{d}a\mathrm{d}b. \tag{9}$$

Define

$$\begin{aligned}\bar{s}_{ab}(x)&=s\left\{\frac{a}{u_0(x)}x+\left(b-\frac{av_0(x)}{u_0(x)}\right)\varepsilon\right\},\\ \bar{r}_{ab}(x)&=r\left\{\frac{a}{u_0(x)}x+\left(b-\frac{av_0(x)}{u_0(x)}\right)\varepsilon\right\}.\end{aligned} \tag{10}$$

Then it is easy to verify that $(\bar{s}_{ab},\ \bar{r}_{ab})\in\vartheta$. Therefore, from (6), (7), (8), (9), and (10), it follows that

$$\underline{V}(F, W^*) + \varepsilon \geq \frac{1}{\varphi(T)} \inf_{(a,b)\in\Sigma''} \int W^*[\bar{s}_{ab}(y)u_0(y), \bar{r}_{ab}(y)u_0(y) + v_0(y)]\mathrm{d}F(y)\int_{L_T} \frac{1}{a}\mathrm{d}a\mathrm{d}b$$
$$- \frac{1}{\varphi(T)} 2mTM(1 + 2\log m)$$
$$\geq \inf_{(s,r)\in\vartheta} \int W^*[s(y)u_0(y), r(y)u_0(y) + v_0(y)]\mathrm{d}F(y) - \varepsilon(T). \tag{11}$$

where

$$\lim_{T\to\infty} \varepsilon(T) = \lim_{T\to\infty} \frac{mM}{2\log T}(1 + 2\log m) = 0. \tag{12}$$

But it is not difficult to show that

$$\inf_{(s,r)\in\vartheta} \int W^*[s(y)u_0(y), r(y)u_0(y) + v_0(y)]\mathrm{d}F(y)$$
$$= \inf\Big\{ \int W^*[s(y)u_0(y), r(y)u_0(y) + v_0(y)]\mathrm{d}F(y) \mid (s, r) \in \vartheta;$$
$$\text{for any } y, (s(y), r(y)) \in L_{2m^2} \Big\}. \tag{13}$$

In fact, for any $(s, r)\in\vartheta$, let us define a pair of functions s_1, r_1 in the following manner:

$$(s_1(x), r_1(x)) = \begin{cases} (s(x), r(x)), & \text{for } (s(x), r(x)) \in L_{2m^2}; \\ ((s(x))^*, (r(x))^*), & \text{for } (s(x), r(x)) \overline{\in} L_{2m^2}. \end{cases} \tag{14}$$

Here $((s(x))^*, (r(x))^*)$ is determined from $(s(x), r(x))$ by means of (5), with $q = 2m^2$. Clearly we have $(s_1, r_1)\in\vartheta$, $(s_1(x), r_1(x))\in L_{2m^2}$ for every x, and

$$W^*[s_1(x)u_0(x), r_1(x)u_0(x) + v_0(x)] \leq W^*[s(x)u_0(x), r(x)u_0(x) + v_0(x)].$$

Therefore

$$\int W^*[s_1(y)u_0(y), r_1(y)u_0(y) + v_0(y)]\mathrm{d}F(y)$$
$$\leq \int W^*[s(y)u_0(y), r(y)u_0(y) + v_0(y)]\mathrm{d}F(y),$$

and (13) is proved. But we have $(s_1(x)u_0(x), r_1(x)u_0(x) + v_0(x)) \in L_{m(2m^2+1)}$ when $(s_1(x), r_1(x))\in L_{2m^2}$, and $(u_0(x), v_0(x))\in L_m$. So, from the definition of W it follows at once

"The right hand side of (11)"

$$\geq \inf_{(s,r)\in\vartheta} \int W[s(y)u_0(y), r(y)u_0(y) + v_0(y)]\mathrm{d}F(y) - \varepsilon(T)$$
$$= V(F, W) - \varepsilon(T). \tag{15}$$

From (11) and (15) it follows that

$$\underline{V}(F, W^*) + \varepsilon \geq V(F, W) - \varepsilon(T).$$

But obviously $\underline{V}(F, W) \geq \underline{V}(F, W^*)$. Inserting this into the above inequality and letting

$T\to\infty$ and then $\varepsilon\to 0$, we obtain $\underline{V}(F, W) \geqslant V(F, W)$. As the reverse inequality holds true in any case $V(F, W) = \underline{V}(F, W)$ follows and the proof is terminated.

In case W is bound, it follows by the same argument as used in [3] that $V(F, W) = \underline{V}(F, W)$, provided F satisfies (4). Nevertheles, it seems impossible to prove the existence of MIE without some additional assumptions.

Lemma 2 (A) Suppose that $W \geqslant 0$ is continuous on Σ^0, $W(0, t_2) = \infty$ and $W(t_1, t_2) \to\infty$ uniformly as $\frac{1}{t_1} + t_1 + |t_2| \to \infty$. Let μ be a measure defined on the σ-field $\mathscr{F}$ of all Borel sets in Σ^0 such that $\mu(B) < \infty$ for any closed bound set B, then

$$\lim_{m\to\infty} \inf_{(a, b)\in\Sigma^0} \int_{L_m} W(at_1, bt_1 + t_2)\mathrm{d}\mu = \inf_{(a, b)\in\Sigma^0} \int_{\Sigma^0} W(at_1, bt_1 + t_2)\mathrm{d}\mu. \tag{16}$$

(B) Suppose that $W \geqslant 0$ is continuous on Σ, $W(t_1, t_2) \to \alpha \leqslant \infty$ uniformly as $t_1 + |t_2| \to \infty$, and μ is a measure on $\mathscr{F}$ such that $\mu(C) < \infty$ for any bound Borel set C, then (16) holds true, but the symbol " $\inf\limits_{(a, b)\in\Sigma^0}$ " should be replaced by " $\inf\limits_{(a, b)\in\Sigma}$ ".

Proof We shall prove only (A), for the proof of (B) is essentially the same. Without loss of generality we can assume that m takes only positive integer value. Write

$$h_m(a, b) = \int_{L_m} W(at_1, bt_1 + t_2)\mathrm{d}\mu.$$

It is easily seen that h_m is continuous on Σ^0. Clearly we may assume that $\mu(\Sigma^0) > 0$, so $\mu(L_m) > 0$ for the sufficiently large m. For such an m, $W(t_1, t_2)\to\infty$ uniformly as $\frac{1}{t_1} + t_1 + |t_2| \to\infty$ and h_m attains its minimum at some finite point $(a_m, b_m) \in \Sigma^0$. Writing $c_m = h_m(a_m, b_m)$, we have $c_1 \leqslant c_2 \leqslant \cdots$ and $\lim\limits_{m\to\infty} c_m = c \leqslant \infty$ exists. Two cases may occur: (Ⅰ) $h(a, b) = \int_{\Sigma^0} W(at_1, bt_1 + t_2)\mathrm{d}\mu \equiv \infty$, and (Ⅱ) $h \not\equiv \infty$. We consider only case (Ⅰ) as case (Ⅱ) may be discussed in a similar way.

We have to show that $c = \infty$. Suppose, on the contrary, that $c < \infty$; then the sequence $\left\{\left(\frac{1}{a_m} + a_m + |b_m|\right)\right\}$ must be bound, for otherwise we may assume that $\lim\limits_{m\to\infty}\left(\frac{1}{a_m} + a_m + |b_m|\right) = \infty$. Then for T sufficiently large (so as to make $\mu(L_T) > 0$), one has

$$\begin{aligned}\infty > c = \lim_{m\to\infty} c_m &= \lim_{m\to\infty}\int_{L_T} W(a_m t_1, b_m t_1 + t_2)\mathrm{d}\mu \\ &\geqslant \int_{L_T} \liminf_{m\to\infty} W(a_m t_1, b_m t_1 + t_2)\mathrm{d}\mu = \infty.\end{aligned}$$

This contradiction proves the boundness of $\left\{\left(\frac{1}{a_m} + a_m + |b_m|\right)\right\}$. Therefore without loss of generality we may assume that $(a_m, b_m)\to(a_0, b_0) \in \Sigma^0$. For any fixed T and $m \geqslant T$, we have

$$\int_{L_T} W(a_m t_1, b_m t_1 + t_2) d\mu \leqslant \int_{L_m} W(a_m t_1, b_m t_1 + t_2) d\mu = c_m \leqslant c < \infty.$$

Let $m \to \infty$ and then $T \to \infty$. Then it follows that $h(a_0, b_0) \leqslant c < \infty$, Which contradicts $h \equiv \infty$, and the lemma is proved.

Remark 1 Suppose that W is a strictly convex function on Σ and μ is a nonzero measure on $\mathscr{F}$. Then it is easily seen that $h(a, b) = \int_{\Sigma^0} W(at_1, bt_1 + t_2) d\mu$ is strictly convex on the convex set $H = \{(a, b) \mid h(a, b) < \infty\}$. In case H is not empty, it follows that the point at which h attains its minimum is unique.

Before formulating Lemma 3, it is convenient to introduce the following notion: Let $A \subset R_k$ be a bound set. Define

$$\bar{a}_1 = \sup\{a_1 \mid a = (a_1, \cdots, a_k) \in A\}.$$

The intersection of $\bar{A}$ (closure of A) and the hyperplane $x_1 = a_1$ is a bound set A_1. Let

$$\bar{a}_2 = \sup\{a_2 \mid a = (a_1, \cdots, a_k) \in A_1\}.$$

Suppose now that the numbers $\bar{a}_1, \cdots, \bar{a}_p (p \leqslant k-1)$ have already been determined. Let $A_p = \bar{A} \cap \{(a_1, \cdots, a_k) \mid a_i = \bar{a}_i, i \leqslant p\}$ and $\bar{a}_{p+1} = \sup\{a_{p+1} \mid a = (a_1, \cdots, a_k) \in A_p\}$. The point $\bar{a} = (\bar{a}_1, \cdots, \bar{a}_k)$ thus obtained clearly belongs to $\bar{A}$ and will be denoted by

$$\bar{a} = \sup a\{A\}.$$

Lemma 3 Let $\varphi^{(i)}(y) = (\varphi_1^{(i)}(y), \cdots, \varphi_k^{(i)}(y))$, $i = 1, 2, \cdots$ be a sequence of Borel measurable functions defined on a Borel set $E \subset R_m$. Suppose that for every $y \in E$ the set $A_y = \{\varphi^{(i)}(y), i = 1, 2, \cdots\}$ is bound; then the function f defined on E by the formula $f(y) = (f_1(y), \cdots, f_k(y)) = \sup a\{A_y\}$ is Borel measurable.

Proof The case $k = 1$ being a well-known result, we proceed to prove the general case by the method of induction. We first show that for any real u

$$\{y \mid y \in E; f_k(y) > u\}$$
$$= \sum_{n=1}^{\infty} \prod_{m=1}^{\infty} \sum_{i=1}^{\infty} \left\{ y \middle| y \in E;\ \varphi_d^{(i)}(y) > f_d(y) - \frac{1}{m},\ d = 1, \cdots, k-1;\ \varphi_k^{(i)}(y) \geqslant u + \frac{1}{n} \right\}. \tag{17}$$

In fact, assume that y_0 belongs to the left hand of (17); then $f_k(y_0) - u = \delta > \frac{1}{n}$ for n_0 sufficiently large. Let m be arbitrarily fixed. Since $f(y_0) \in \bar{A}_{y_0}$, an integer i_m may be found such that $\varphi^{(i_m)}(y_0) \in V_\varepsilon$ where V_ε is a sphere with a centre $f(y_0)$ and a radius $\varepsilon < \min\left(\frac{1}{m}, \delta - \frac{1}{n_0}\right)$. Consequently we have shown that there exists an integer n such that to every m there corresponds an integer i_m such that

$$\varphi_d^{(i_m)}(y_0) > f_d(y_0) - \frac{1}{m},\ d = 1, \cdots, k-1,\ \varphi_k^{(i_m)}(y_0) \geqslant u + \frac{1}{n_0}, \tag{18}$$

where y_0 belongs to the right hand of (17). Conversely, if y_0 belongs to the right hand of (17), then an integer n_0 exists such that for any m, an integer i_m satisfying (18) may be found. The sequence $\{\varphi^{(i_m)}(y_0)\}$ being bound, there exists a subsequence $\{\varphi^{(i_{m_r})}(y)\}$ converging to a point $a = (a_1, \cdots, a_k)$. Clearly

$$a_d \geqslant f_d(y_0),\ d = 1, \cdots, k-1;\ a_k \geqslant u + \frac{1}{n_0}.$$

As the reverse inequality $a_d \leqslant f_d(y_0)$, $d = 1, \cdots, k-1$ follows easily from the definition of "sup a", we have $a_d = f_d(y_0)$ for $d = 1, \cdots, k-1$. Therefore $f_k(y_0) \geqslant a_k \geqslant u + 1/n_0 > u$ by the very definition of "sup a". This shows that y_0 belongs to the left hand of (17) and concludes the proof of (17).

Suppose now that Lemma 3 holds true for $k-1$ instead of k; then the right hand of (17) is a Borel set, and this proves the measurability of f_k. The truth of Lemma 3 for k follows, and the proof is terminated.

Lemma 4 Suppose that a sequence of functions $\{\varphi^{(i)}(y) = (\varphi_1^{(i)}(y), \cdots, \varphi_k^{(i)}(y)), i = 1, 2, \cdots\}$ defined on a Borel set $E \subset R_m$ satisfies the condition of Lemma 3. Then there exists a Borel measurable function $\varphi(y) = (\varphi_1(y), \cdots, \varphi_k(y))$ defined on E such that for any $y \in E$, a subsequence $\{i_r\}$ (depending, in general, on y) of positive integers may be found such that $\varphi(y) = \lim_{r\to\infty}\varphi^{(i_r)}(y)$.

Proof For $k = 1$ the lemma follows by taking $\varphi(y) = \limsup_{n\to\infty}\varphi^{(n)}(y)$. The general case may again be treated by the method of induction.

For any $y \in E$, let $A_{y_n} = \{\varphi^{(i)}(y), i = n, n+1, \cdots\}$, $n = 1, 2, \cdots$. By Lemma 3, the functions $f^{(n)}(y) = (f_1^{(n)}(y), \cdots, f_k^{(n)}(y)) = \sup a\{A_{y_n}\}$, $n = 1, 2, \cdots$ are Borel measurable on E. Clearly, for any $y \in E$, $\lim_{n\to\infty} f_1^{(n)}(y)$ exists. Consider the sequence of functions defined on E by $g^{(n)}(y) = (f_2^{(n)}(y), \cdots, f_k^{(n)}(y))$, $n = 1, 2, \cdots$. This sequence satisfies all conditions of Lemma 4. Therefore by induction hypothesis there exists on E a Borel measurable function $g(y) = (g_2(y), \cdots, g_k(y))$ such that for any $y \in E$, a subsequence $\{g^{(n_s)}(y)\}$ tending to $g(y)$ exists. Now define

$$\varphi(y) = (\varphi_1(y), \cdots, \varphi_k(y)) = (\lim_{n\to\infty} f_1^{(n)}(y), g_2(y), \cdots, g_k(y))$$

and proceed to show that it satisfies all conditions of Lemma 4. Since the measurability of φ follows from g, we have only to show that for any $y \in E$, there exists a subsequence $\{\varphi^{(i_r)}(y)\}$ tending to $\varphi(y)$. Suppose on the contrary that for some $\varepsilon > 0$, $\varphi^{(i)}(y) \overline{\in} V$ as $i \geqslant n_0$, where V is a sphere with a radius ε and a centre $\varphi(y)$. From the definition of "sup a" it then follows that $f^{(n)}(y) = \sup a\{A_{y_n}\} \overline{\in} V$ for $n \geqslant n_0$. This is a contradiction since for any y we have $g^{(n_s)}(y) \to g(y)$, and thus $f^{(n_s)}(y) \to \varphi(y)$ as $s \to \infty$. The proof of Lemma 4 is concluded.

Lemma 5 Let $P(Z, B)$ be a Borel measurable probability defined on $D \times \mathscr{F}$, where $D \subset R_n$ is a Borel set. Suppose further that the loss function W satisfies the conditions of

Lemma 2 (A); then for any m, $0 < m < \infty$, there exists a Borel measurable function (a_m, b_m) mapping D into Σ^0 such that

$$\int_{L_m} W(a_m(z)t_1, b_m(z)t_1 + t_2)P(z, d\omega) = \inf_{(a, b) \in \Sigma^0} \int_{L_m} W(at_1, bt_1 + t_2)P(x, d\omega), \quad (19)$$

where $d\omega = d(t_1, t_2)$. There exists also a Borel measurable function (a, b) mapping D into Σ^0 such that

$$\int_{\Sigma^0} W(a(z)t_1, b(z)t_1 + t_2)P(z, d\omega) = \inf_{(a, b) \in \Sigma^0} \int_{\Sigma^0} W(at_1, bt_1 + t_2)P(z, d\omega). \quad (20)$$

If the function W satisfies the conditions of Lemma 2 (B), then the above conclusion remains true, only that the symbol " $\inf\limits_{(a, b) \in \Sigma^0}$ " should be replaced by " $\inf\limits_{(a, b) \in \Sigma}$ ", and the functions (a_m, b_m), (a, b) will map D into Σ.

Proof We shall treat only the first case for in the second case the proof remains essentially unchanged.

The set $B_m = \{z \mid z \in D, P(z, L_m) = 0\}$ is a Borel set. If $z_0 \in B_m$, then any point in Σ^0 may be chosen for $(a_m(z_0), b_m(z_0))$ in (19). So without loss of generality we can assume $B_m = \varnothing$. Define

$$H_m(a, b; z) = \int_{L_m} W(at_1, bt_1 + t_2)P(z, d\omega).$$

By a well-known fact of measure theory, $H_m(a, b; z)$ is Borel measurable on D for any fixed $(a, b) \in \Sigma^0$. The continuity of W implies the continuity of $H_m(a, b; z)$ in (a, b) for any fixed $z \in D$, and we have $\inf\limits_{(a, b) \in \Sigma^0} H_m(a, b; z) = \inf\limits_n H_m(a_n, b_n, z)$, where $\{(a_n, b_n)\}$ is the sequence of all rational points in Σ^0. In this way, we have proved that the right hand of (19) is Borel measurable. Now let $\varepsilon > 0$ be given, and define

$$E_{n\varepsilon} = \{z \mid z \in D, H_m(a_n, b_n, z) \leqslant \inf_{(a, b) \in \Sigma^0} H_m(a, b, z) + \varepsilon\}, \ n = 1, 2, \cdots$$

Clearly, $\{E_{n\varepsilon}\}$ is a sequence of the Borel set and $D = \sum\limits_{n=1}^{\infty} E_{n\varepsilon}$. Write $E^0_{1\varepsilon} = E_{1\varepsilon}, \cdots, E^0_{n\varepsilon} = E_{n\varepsilon} - (E_{1\varepsilon} + \cdots + E_{n-1\varepsilon}), \cdots$; then the sets of $\{E^0_{n\varepsilon}\}$ are pairwise disjoint. Now define a sequence of functions on D as follows:

$$(a_{m\varepsilon}(z), b_{m\varepsilon}(z)) = (a_n, b_n), \text{ for } z \in E^0_{n\varepsilon}, \ n = 1, 2, \cdots$$

These functions are Borel measurable on D. If we choose a sequence $\{\varepsilon_s\}$, $\varepsilon_s \downarrow 0$, and if for every s we define a function $(a_{m\varepsilon_s}, b_{m\varepsilon_s})$ as above, then for any $z \in D$ we have

$$\inf_{(a, b) \in \Sigma^0} H_m(a, b; z) \leqslant H_m(a_{m\varepsilon_s}(z), b_{m\varepsilon_s}(z); z) \leqslant \inf_{(a, b) \in \Sigma^0} H_m(a, b; z) + \varepsilon_s. \quad (21)$$

Setting $s \to \infty$, we obtain

$$\lim_{s\to\infty} H_m(a_{mt_s}(z),\ b_{mt_s}(z);\ z) = \inf_{(a,\ b)\in\Sigma^0} H_m(a,\ b;\ z). \tag{22}$$

Since the right hand of (19) is finite and since by assumption $P(z,\ L_m) > 0$, by an argument similar to what was used in the proof of Lemma 2, for any fixed $z \in D$, the sequence $\{(a_{mt_s}^{-1}(z) + a_{mt_s}(z) + |b_{mt_s}(z)|)\}_{s=1}^{\infty}$ is bound, and therefore there exists a compact set $J_z \subset \Sigma^0$ such that $(a_{mt_s}(z),\ b_{mt_s}(z)) \in J_z$ for all s. By Lemma 4, there exists a function on D such that for any $z \in D$ a subsequence $\{s_i\}$ can be found to satisfy the relation

$$\lim_{i\to\infty}(a_{mt_{s_i}}(z),\ b_{mt_{s_i}}(z)) = (a_m(z),\ b_m(z)).$$

Clearly $a_m(z) > 0$ for all z. Owing to the continuity of $H_m(a,\ b;\ z)$ in $(a,\ b)$, from (21) and (22) it follows at once that $H_m(a_m(z),\ b_m(z);\ z) = \inf_{(a,\ b)\in\Sigma^0} H_m(a,\ b;\ z)$. This proves (19).

We now proceed to prove (20). Write

$$H(a,\ b;\ z) = \int_{\Sigma^0} W(at_1,\ bt_1 + t_2)P(z,\ d\omega).$$

By (16) and the Borel measurability of $\inf_{(a,\ b)\in\Sigma^0} H_m(a,\ b;\ z)$ it follows that $\inf_{(a,\ b)\in\Sigma^0}\int_{\Sigma^0} W(at_1,\ bt_1 + t_2)P(z,\ d\omega)$ is Borel measurable. As before we may assume that $B = \{z \mid z \in D,\ H(a,\ b;\ z) \equiv \infty\}$ is empty.

Now let $(a_m,\ b_m)$, $m = 1,\ 2,\ \cdots$ be the functions satisfying (19) found above. For any fixed b' and $m \geqslant b'$,

$$\begin{aligned}
&\int_{L_{b'}} W(a_m(z)t_1,\ b_m(z)t_1 + t_2)P(z,\ d\omega) \\
&\quad\leqslant \int_{L_m} W(a_m(z)t_1,\ b_m(z)t_1 + t_2)P(z,\ d\omega) \\
&\quad\leqslant \inf_{(a,\ b)\in\Sigma^0} H(a,\ b;\ z) < \infty,
\end{aligned} \tag{23}$$

and as before for any fixed $z \in D$ there exists a compact set $J_z \subset \Sigma^0$ such that $(a_m(z),\ b_m(z)) \in J_z$ for all m. By Lemma 4 there exists a Borel measurable function $(a,\ b)$ such that for any $z \in D$ a subsequence $\{m_i\}$ can be found to satisfy

$$\lim_{i\to\infty}(a_{m_i}(z),\ b_{m_i}(z)) = (a(z),\ b(z));$$

clearly, $a(z) > 0$. From (23) it follows that

$$\begin{aligned}
&\int_{L_{b'}} W(a(z)t_1,\ b(z)t_1 + t_2)P(z,\ d\omega) \\
&\quad= \lim_{i\to\infty}\int_{L_{b'}} W(a_{m_i}(z)t_1,\ b_{m_i}(z)t_1 + t_2)P(z,\ d\omega) \\
&\quad\leqslant \inf_{(a,\ b)\in\Sigma^0} H(a,\ b;\ z).
\end{aligned} \tag{24}$$

Setting $b' \to \infty$, we see that (20) follows from (24). Lemma 5 is thus proved.

Remark 2 If W is strictly convex and $B = \varnothing$, then it follows from Remark 1 and

Lemma 5 that the function (a, b) found above to satisfy (20) is unique.

4 Proof of Theorem 1

Define a function z on $R_N - A$:

$$z(x) = \left(i_r, \operatorname{sign}(x_{i_r} - x_1), \frac{x_{i_r} - x_j}{x_{i_r} - x_1}, j = 2, \cdots, N\right),$$

where $i_r = \min\{j \mid x_j \neq x_1, j = 2, \cdots, N\}$. It is readily seen that z is a maximal invariant, and therefore the condition $(s, r) \in \vartheta$ is equivalent to the condition that (s, r) depends only on z. In such a case, we shall write $s(x) = s(z(x))$ for convenience. It is easily seen that z is a Borel measurable function.

Let $P(z, B)$ be the mixed conditional distribution of (u_0, v_0) (see Section 2) given $z(X)$, the underlying distribution being F (see, for example, p. 361 in [16]). Suppose that $(s, r) \in \vartheta$; we have

$$\begin{aligned}
&\mathrm{E}_F\{W[s(X)u_0(X), r(X)u_0(X) + v_0(X)] \mid Z\} \\
&= \mathrm{E}_F\{W[s(Z)u_0(X), r(Z)u_0(X) + v_0(X)] \mid Z\} \\
&= \int_{\Sigma^0} W[s(Z)t_1, r(Z)t_1 + t_2]P(Z, \mathrm{d}\omega).
\end{aligned}$$

By Lemma 5, there exists a Borel measurable function (s_0, r_0) such that

$$\begin{aligned}
w(Z) &= \mathrm{E}_F\{W[s_0(Z)u_0(X), r_0(Z)u_0(X) + v_0(X)] \mid Z\} \\
&= \int_{\Sigma^0} W[s_0(Z)t_1, r_0(Z)t_1 + t_2]P(Z, \mathrm{d}\omega) \\
&= \inf_{(a, b) \in \Sigma^0} \int_{\Sigma^0} W[at_1, bt_1 + t_2]P(Z, \mathrm{d}\omega).
\end{aligned}$$

Therefore, for any $(s, r) \in \vartheta$ we have

$$\begin{aligned}
\rho(s, r) &= \mathrm{E}_F\left\{\int_{\Sigma^0} W[s(Z)t_1, r(Z)t_1 + t_2]P(Z, \mathrm{d}\omega)\right\} \\
&\geqslant \mathrm{E}_F\left\{\int_{\Sigma^0} W[s_0(Z)t_1, r_0(Z)t_1 + t_2]P(Z, \mathrm{d}\omega)\right\} \\
&= \rho(s_0, r_0). \qquad (25)
\end{aligned}$$

This leads to

$$V = V(F, W) = \inf_{(s, r) \in \vartheta} \rho(s, r) = \rho(s_0, r_0) = \mathrm{E}_F\{w(Z)\}.$$

Now let m be a positive integer, define a measure F_m:

$$F_m(C) = F(C \cap \{x : (u_0(x), v_0(x)) \in L_m\}),$$

and let $h(x, m)$ be the characteristic function of the set $\{x \mid (u_0(x), v_0(x)) \in L_m\}$; then for any $(s, r) \in \vartheta$, we have