THEORY PROBAB. APPL.
Vol. 35, No. 1

Translated from Russian Journal

STRUCTURE OF INVARIANT CONFIDENCE INTERVALS CONTAINING THE MAIN DISTRIBUTED MASS*

I. G. BAIRAMOV AND YU. I. PETUNIN

(*Translated by A. E. Shemyakin*)

Let $\mathbf{G} = \{G_\alpha, \alpha \in \mathfrak{A}\}$ be some class of general populations with distribution functions $F_\alpha(u)$ $(\alpha \in \mathfrak{A})$, $x_1^\alpha, x_2^\alpha, \cdots, x_n^\alpha$ being a sample from the general population G_α, obtained by a simple sampling and x^αan observation from G_α, independent of the sample $x_1^\alpha, \cdots, x_n^\alpha$. Consider two measurable functions of n variables, $f_1(u_1, u_2, \cdots, u_n)$ and $f_2(u_1, u_2, \cdots, u_n)$, satisfying the inequality $f_1(u_1, u_2, \cdots, u_n) \leqq f_2(u_1, u_2, \cdots, u_n)$ (for all $(u_1, u_2, \cdots, u_n) \in \mathbf{R}^n$).

With the help of the functions $f_1(u_1, u_2, \cdots, u_n)$, $f_2(u_1, u_2, \cdots, u_n)$ and the sample $x_1^\alpha, \cdots, x_n^\alpha$ the random confidence interval $J_\alpha = \big(f_1(x_1^\alpha, \cdots, x_n^\alpha), f_2(x_1^\alpha, \cdots, x_n^\alpha)\big)$ $(\alpha \in \mathfrak{A})$ may be constructed for the bulk of the distribution of the general population G_α.

DEFINITION 1. A random interval $J_\alpha = \big(f_1(x_1^\alpha, \cdots, x_n^\alpha), f_2(x_1^\alpha, \cdots, x_n^\alpha)\big)$ is called a confidence interval for the bulk of the distribution G_α with invariant significance level $1 - \theta$ in the class $\mathbf{G}$ (or simply an invariant significance interval for the class $\mathbf{G}$) if

$$P(x^\alpha \in J_\alpha) = \mathbf{P}\big(f_1(x_1^\alpha, \cdots, x_n^\alpha) < x^\alpha < f_2(x_1^\alpha, \cdots, x_n^\alpha)\big) = \theta \qquad (\forall \alpha \in \mathfrak{A}).$$

Note. The definition of an invariant confidence interval admits the cases when $f_1(u_1, \cdots, u_n) \equiv -\infty$ or $f_2(u_1, \cdots, u_n) \equiv \infty$; such invariant confidence intervals are called one-sided in contrast to two-sided, when neither $f_1(u_1, \cdots, u_n) \not\equiv -\infty$ nor $f_2(u_1, \cdots, u_n) \not\equiv \infty$.

Thus invariant confidence intervals in the class $\mathbf{G}$ have the significance level $\gamma = 1 - \theta$, which does not depend on the population $G_\alpha \in \mathbf{G}$. The main objective of the paper is to clarify the construction of invariant confidence intervals for the class $\mathbf{G}_C$ of general populations with continuous distribution $F_\alpha(u)$, and also to find the set of significance levels corresponding to all possible invariant confidence levels for the class $\mathbf{G}_C$.

Invariant confidence intervals for the class $\mathbf{G}_C$ were found for the first time in [1]; to describe these intervals we denote by $x_{(1)} \leqq x_{(2)} \leqq \cdots \leqq x_{(n)}$ the variational series [2], constructed according to the sample $x_1, x_2, \cdots, x_n$; then the functions $f_1(x_1, \cdots, x_n) = x_{(i)}$ and $f_2(x_1, \cdots, x_n) = x_{(j)}$ $(i < j)$, where $x_{(i)}$ and $x_{(j)}$ are the ith and the jth order statistics, respectively, define an invariant confidence interval J_α with confidence level $\theta = (j - i)/(n + 1)$:

$$\mathbf{P}\Big(x^\alpha \in \big(x_{(i)}^\alpha, x_{(j)}^\alpha\big)\Big) = \frac{j - i}{n + 1} \qquad (\forall G_\alpha \in \mathbf{G}_C).$$

It turns out that there is no other invariant confidence interval for the class $\mathbf{G}_C$; more precisely, the following theorem holds.

THEOREM 1. *Let $f_1(x_1, \cdots, x_n)$ and $f_2(x_1, \cdots, x_n)$ be two continuous symmetric functions satisfying the inequality*

$$f_1(x_1, \cdots, x_n) \leqq f_2(x_1, \cdots, x_n),$$

*Received by the editors June 10, 1987.

which coincide only on a set in $\mathbf{R}^n$ *of Lebesgue measure zero. A two-sided confidence interval* $J_\alpha = \big(f_1(x_1^\alpha, \cdots, x_n^\alpha), f_2(x_1^\alpha, \cdots, x_n^\alpha)\big)$, *containing the bulk of the distribution of the general population* G_α *with confidence level* θ *is invariant for the class* $\mathbf{G}_C$ *of the general populations with continuous distributions if and only if* $f_1(x_1, \cdots, x_n) = x_{(i)}$, $f_2(x_1, \cdots, x_n) = x_{(j)}$, *where* $i < j$ *and* $x_{(i)}$, $x_{(j)}$ *are some order statistics constructed from the sample* $x_1, \cdots, x_n$.

For the sake of brevity the proof of the theorem will be given for $n = 2$; in the case of arbitrary natural $n > 2$ the proof is quite similar. So let x_1, x_2, x_3 be a sample from a general population $G \in \mathbf{G}_C$ with continuous distribution function $F(u)$ and let $f_1(u_1, u_2)$, $f_2(u_1, u_2)$ be some functions satisfying the conditions of the theorem. Define the functional $g(F)$ on the set $\mathcal{F}_C$ of all continuous distributions $F(u)$ in the following way:

$$\begin{aligned} g(F) &= \mathbf{P}\Big\{x_3 \in \big(f_1(x_1, x_2), f_2(x_1, x_2)\big)\Big\} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Big[F\big(f_2(u_1, u_2)\big) - F\big(f_1(u_1, u_2)\big)\Big]\, dF(u_1)\, dF(u_2). \end{aligned} \tag{1}$$

Denote by $C(-\infty, \infty)$ the vector space of all continuous bounded functions, defined on the real line $\mathbf{R}^1$. It is obvious that the set $\mathcal{F}_C \subset C(-\infty, \infty)$ is convex.

DEFINITION 2. The derivative of a functional $h(F)$ defined on a convex set M of a vector space E at the point $F_0 \in M$ in the direction of $F_1 \in M$ is the value

$$h'_{F_1}(F_0) = \lim_{\substack{\alpha \to 0 \\ 0 < \alpha < 1}} \frac{h\big(F_0 + \alpha(F_1 - F_0)\big) - h(F_0)}{\alpha}.$$

The following obvious assertion is valid.

LEMMA 1. *The derivative of the functional* $g(F)$ *defined by the equality* (1) *on the convex set* $M = \mathcal{F}_C \subset E = C(-\infty, \infty)$ *at the point* F_0 *in the direction of* F *can be calculated using the formula*

$$\begin{aligned} g'_F(F_0) &= 2\iint_{-\infty}^{\infty} \Big[F_0\big(f_2(u_1, u_2)\big) - F_0\big(f_1(u_1, u_2)\big)\Big]\, dF(u_1)\, dF_0(u_2) \\ &\quad + \iint_{-\infty}^{\infty} \Big[F\big(f_2(u_1, u_2)\big) - F\big(f_1(u_1, u_2)\big)\Big]\, dF_0(u_1)\, dF_0(u_2) - 3g(F_0). \end{aligned} \tag{2}$$

LEMMA 2. *Let the functional* $g(F_0, F_1)$ *on the Cartesian product* $\mathcal{F}_C \times \mathcal{F}_C$ *of the set* $\mathcal{F}_C$ *of all continuous distribution functions be defined by the formula*

$$g(F_0, F_1) = \iint_{-\infty}^{\infty} \Big[F_0\big(f_2(u_1, u_2)\big) - F_0\big(f_1(u_1, u_2)\big)\Big]\, dF_1(u_1)\, dF_0(u_2);$$

moreover, let its restriction to the diagonal of the Cartesian product $g(F) = g(F, F)$ *be a constant* $g(F) = A^*$, $(F \in \mathcal{F}_C)$. *If there exists a pair* $(\overline{F}, \overline{\overline{F}})$ *in* $\mathcal{F}_C \times \mathcal{F}_C$ *such that* $g(\overline{F}, \overline{\overline{F}}) = 0$, *then, for each* $(F_0, F_1) \in \mathcal{F}_C \times \mathcal{F}_C$, *the inequality* $g(F_0, F_1) \leqq \frac{1}{2}$ *holds.*

Proof. By virtue of Lemma 1, we obtain

$$\begin{aligned} g'_F(F_0) =& 2\iint_{-\infty}^{\infty} \Big[F_0\big(f_2(u_1, u_2)\big) - F_0\big(f_1(u_1, u_2)\big)\Big]\, dF(u_1)\, dF_0(u_2) \\ &+ \iint_{-\infty}^{\infty} \Big[F\big(f_2(u_1, u_2)\big) - F\big(f_1(u_1, u_2)\big)\Big]\, dF_0(u_1)\, dF_0(u_2) - 3A^* = 0; \end{aligned}$$

whence

$$(3)\quad 2g(F_0, F_1) + \iint_{-\infty}^{\infty} \Big[F_1\big(f_2(u_1, u_2)\big) - F_1\big(f_1(u_1, u_2)\big)\Big]\, dF_0(u_1)\, dF_0(u_2) = 3A^*.$$

According to the conditions of the theorem there exist distributions $\overline{F}, \overline{\overline{F}}$ such that $g(\overline{F}, \overline{\overline{F}}) = 0$; thus, in view of (3) we have

$$3A^* = \iint_{-\infty}^{\infty} \Big[\overline{\overline{F}}\,\big(f_2(u_1, u_2)\big) - \overline{\overline{F}}\,\big(f_1(u_1, u_2)\big)\Big]\, d\overline{F}(u_1)\, d\overline{F}(u_2) \leqq 1,$$

so that $A^* \leqq \frac{1}{3}$. Assume

$$\delta = \sup\big\{g(F_0, F_1)\colon\ (F_0, F_1) \in \mathcal{F}_C \times \mathcal{F}_C\big\},$$

and let ε be an arbitrary positive number. Choose the distributions $\overline{F}_\varepsilon, \overline{\overline{F}}_\varepsilon$ so that $\delta < (\overline{F}_\varepsilon, \overline{\overline{F}}_\varepsilon) + \varepsilon$; then, according to (3),

$$2(\delta - \varepsilon) + \iint_{-\infty}^{\infty} \Big[\overline{F}_\varepsilon\big(f_2(u_1, u_2)\big) - \overline{F}_\varepsilon\big(f_1(u_1, u_2)\big)\Big]\, d\,\overline{\overline{F}}_\varepsilon(u_1)\, d\,\overline{\overline{F}}_\varepsilon(u_2) < 3A^*,$$

and hence $2(\delta - \varepsilon) \leqq 3A^*$. This inequality and $A^* \leqq \frac{1}{3}$ yield $\delta \leqq \frac{1}{2} + \varepsilon$ and, due to the arbitrary choice of $\varepsilon > 0$, $\delta \leqq \frac{1}{2}$, which completes the proof.

Let $u = (u_1, u_2, \cdots, u_n)$ be a random point in the n-dimensional Euclidean space $\mathbf{R}^n$ and $u_{(1)} \leqq \cdots \leqq u_{(n)}$ a permutation of the coordinates of u according to increasing order.

DEFINITION 3. A function of n variables $(n \geqq i)$ having the form

$$\varphi_i(u_1, u_2, \cdots, u_n) = u_{(i)}$$

is called a fundamental ith symmetric function.

The proof of the theorem will now be carried out by contradiction: assume that at least one of the functions $f_1(u_1, u_2)$ and $f_2(u_1, u_2)$ is not a fundamental symmetric function. Then, by virtue of results of [3], it follows that neither $f_1(u_1, u_2)$ nor $f_2(u_1, u_2)$ coincides with any of the fundamental functions $\varphi_1(u_1, u_2)$, $\varphi_2(u_1, u_2)$:

$$(4)\qquad f_i(u_1, u_2) \neq \varphi_1(u_1, u_2), \qquad f_i(u_1, u_2) \neq \varphi_2(u_1, u_2).$$

As a matter of fact, if at least one of these functions (for example, $f_1(u_1, u_2)$) does coincide with some fundamental symmetric function (for example, with $\varphi_1(u_1, u_2)$), then the one-sided confidence interval

$$\big\{x_3 < f_2(x_1, x_2)\big\} = \big\{f_1(x_1, x_2) < x_3 < f_2(x_1, x_2)\big\} \cup \big\{x_3 \leqq f_1(x_1, x_2)\big\}$$

being the union of two invariant confidence intervals, is also an invariant interval; thus, in view of a theorem in [3], $f_2(u_1, u_2) = u_{(2)}$, which contradicts the inequality $f_2(u_1, u_2) \neq \varphi_2(u_1, u_2)$. Hence, by the assumption, $f_1(u_1, u_2) \neq u_{(i)}$, $f_2(u_1, u_2) \neq u_{(i)}$ $(i = 1, 2)$. Assume

$$M_1(f_1) = \big\{(u_1, u_2)\colon\ f_1(u_1, u_2) \neq \varphi_1(u_1, u_2)\big\},$$

$$M_2(f_1) = \big\{(u_1, u_2)\colon\ f_1(u_1, u_2) \neq \varphi_2(u_1, u_2)\big\}.$$

It is easy to see that $M_1(f_1)$ and $M_2(f_1)$ are open sets, their intersection $M(f_1) = M_1(f_1) \cap M_2(f_1)$ being also a nonempty open set. Denote

$$T = \{(u_1, u_2)\colon\ f_1(u_1, u_2) = f_2(u_1, u_2)\}.$$

According to conditions of the theorem, the set T has Lebesgue measure equal to zero; thus, $M(f_1)\backslash T$ is not empty. Indeed, suppose that $M(f_1) \subset T$; since $M(f_1)$ is a nonempty open set, there exists a ball $S(x, \varepsilon) \subset M(f_1) \subset T$, which contradicts the fact that the Lebesgue measure of T is equal to zero. Thus there exists a point $(a_1, a_2) \in M(f_1)\backslash T(a_{(1)} < a_{(2)})$:

$$f_1(a_1, a_2) \neq \varphi_i(a_1, a_2), \quad (i = 1, 2), \quad f_1(a_1, a_2) < f_2(a_1, a_2).$$

Assume first that $f_2(a_1, a_2) \neq \varphi_i(a_1, a_2)$ $(i = 1, 2)$. Let $A = f_1(a_1, a_2)$, $B = f_2(a_1, a_2)$, where $A < B$; then, for some $\varepsilon > 0$, the intervals $(a_{(1)} - \varepsilon, a_{(1)} + \varepsilon)$, $(a_{(2)} - \varepsilon, a_{(2)} + \varepsilon)$, $(A - \varepsilon, A + \varepsilon)$, $(B - \varepsilon, B + \varepsilon)$ are mutually disjoint. According to the conditions of the theorem, the functions $f_1(u_1, u_2)$ and $f_2(u_1, u_2)$ are continuous and symmetric; thus, it is possible to choose an $\varepsilon_1 \in (0, \varepsilon)$ such that

$$A - \varepsilon \leqq f_1(u_1, u_2) \leqq A + \varepsilon \quad \text{and} \quad B - \varepsilon \leqq f_2(u_1, u_2) \leqq B + \varepsilon$$

for any (u_1, u_2) from the square $K = \{(u_1, u_2)\colon\ a_{(1)} - \varepsilon_1 \leqq u_1 \leqq a_{(1)} + \varepsilon_1,\ a_{(2)} - \varepsilon_1 \leqq u_2 \leqq a_{(2)} + \varepsilon_1\}$. By construction, all the intervals $(A - \varepsilon, A + \varepsilon)$, $(B - \varepsilon, B + \varepsilon)$, $(a_{(1)} - \varepsilon_1, a_{(1)} + \varepsilon_1)$, $(a_{(2)} - \varepsilon_1, a_{(2)} + \varepsilon_1)$ will be mutually disjoint. The following arrangements of the midpoints of these intervals may occur:

1. $A < a_{(1)} < B < a_{(2)}$,	4. $A < a_{(1)} < a_{(2)} < B$,
2. $a_{(1)} < A < a_{(2)} < B$,	5. $A < B < a_{(1)} < a_{(2)}$,
3. $a_{(1)} < A < B < a_{(2)}$,	6. $a_{(1)} < a_{(2)} < A < B$.

Consider the alternatives 1 and 2 first. Let us show that in these cases two pairs of distributions may be constructed: F_0, F_1 and F_2, F_3, for which $g(F_0, F_1) = 0$, $g(F_2, F_3) = 1$. This contradicts Lemma 2. In case 1, one can take as $F_0(u)$ the uniform distribution in the interval $(a_{(2)} - \varepsilon_1, a_{(2)} + \varepsilon_1)$ and as $F_1(u)$ the uniform distribution in the interval $(a_{(1)} - \varepsilon_1, a_{(1)} + \varepsilon_1)$; then

$$\begin{aligned} g(F_0, F_1) &= \iint_{-\infty}^{\infty} \left[F_0\big(f_2(u_1, u_2)\big) - F_0\big(f_1(u_1, u_2)\big)\right] dF_0(u_1)\, dF_1(u_1) \\ &= \frac{1}{4\varepsilon_1^2} \iint_K \left[F_0\big(f_2(u_1, u_2)\big) - F_0\big(f_1(u_1, u_2)\big)\right] du_1\, du_2 = 0, \end{aligned}$$

since $f_1(u_1, u_2) \in (A - \varepsilon, A + \varepsilon)$, $f_2(u_1, u_2) \in (B - \varepsilon, B + \varepsilon)$ for any $(u_1, u_2) \in K$.

Assuming now that $F_2(u) = F_1(u)$, $F_3(u) = F_0(u)$, we obtain

$$g(F_2, F_3) = \frac{1}{4\varepsilon_1^2} \iint_K \left[F_1\big(f_2(u_1, u_2)\big) - F_1\big(f_1(u_1, u_2)\big)\right] du_1\, du_2 = 1,$$

since $F_1\big(f_1(u_1, u_2)\big) = 0$ for any $(u_1, u_2) \in K$, and $F_1\big(f_2(u_1, u_2)\big) = 1$.

In case 2, we select as $F_0(u)$ the uniform distribution in the interval $(a_{(1)} - \varepsilon_1, a_{(1)} + \varepsilon_1)$, and as $F_1(u)$ the uniform distribution in the interval $(a_{(2)} - \varepsilon_1, a_{(2)} + \varepsilon_1)$; moreover, $g(F_0, F_1) = 0$. Substituting now $F_1(u)$ for $F_2(u)$ and $F_0(u)$ for $F_3(u)$, we obtain $g(F_2, F_3) = 1$.

The subsequent reasoning will be based on the following lemma.

LEMMA 3. *Let (x_1, x_2, x_3) be a sample from a general population with continuous distribution and $f_1(u_1, u_2)$, $f_2(u_1, u_2)$ continuous symmetric functions, forming* *an invariant confidence interval* $J = \big(f_1(x_1,x_2), f_2(x_1,x_2)\big)$ *for the class* $\mathcal{F}_C$ *of all continuous distributions. If there exists a point* $(a_1,a_2) \in \mathbf{R}^2$ *such that the intervals* $(A-\varepsilon, A+\varepsilon)$, $(B-\varepsilon, B+\varepsilon)$, $(a_{(1)}-\varepsilon_1, a_{(1)}+\varepsilon_1)$, $(a_{(2)}-\varepsilon_1, a_{(2)}+\varepsilon_1)$ *are mutually disjoint for some* $\varepsilon, \varepsilon_1 > 0$, *where* $A = f_1(a_1,a_2)$, $B = f_2(a_1,a_2)$, $a_{(1)} = \min(a_1,a_2)$, $a_{(2)} = \max(a_1,a_2)$, *then* $0 < A^* = \mathbf{P}(x_3 \in J) < 1$.

Proof. Denote by $[a,b]$ the minimal segment containing all the intervals $(A-\varepsilon, A+\varepsilon)$, $(B-\varepsilon, B+\varepsilon)$, $(a_{(1)}-\varepsilon_1, a_{(1)}+\varepsilon_1)$, $(a_{(2)}-\varepsilon_1, a_{(2)}+\varepsilon_1)$. Consider the distribution $F(u)$ with the following properties:

1. $F(u) = 0$ if $u \leqq a$; $F(u) = 1$ if $u \geqq b$.
2. The function $F(u)$ is linear on the segments $[a_{(1)}-\varepsilon_1, a_{(1)}+\varepsilon_1]$, $[a_{(2)}-\varepsilon_1, a_{(2)}+\varepsilon_1]$, and $F(a_{(1)}+\varepsilon_1) - F(a_{(1)}-\varepsilon_1) = F(a_{(2)}+\varepsilon_1) - F(a_{(2)}-\varepsilon_1) = \frac{1}{3}$.
3. The function $F(u)$ takes on the constant value c_0 on the segment $[A-\varepsilon, A+\varepsilon]$ and the constant value c_1 on the segment $[B-\varepsilon, B+\varepsilon]$, where $\frac{1}{3} \leqq c_1 - c_0 \leqq \frac{2}{3}$.
4. To the other points of $[a,b]$ the function $F(u)$ is extended continuously (for example, linearly).

One may observe that such a distribution function exists for any arrangement of mutually disjoint intervals $(A-\varepsilon, A+\varepsilon)$, $(B-\varepsilon, B+\varepsilon)$, $(a_{(1)}-\varepsilon_1, a_{(1)}+\varepsilon_1)$, $(a_{(2)}-\varepsilon_1, a_{(2)}+\varepsilon_1)$ on the segment $[a,b]$.

Let x_1, x_2, x_3 be a sample obtained by a simple sampling from the general population G with distribution $F(u)$. Then

$$\begin{aligned}
&\mathbf{P}\{x_3 \in \big(f_1(x_1,x_2), f_2(x_1,x_2)\big)\}\\
&\quad= \mathbf{P}\{f_1(x_1,x_2) < x_3 < f_2(x_1,x_2),\ (x_1,x_2) \in K\}\\
&\qquad+ \mathbf{P}\{f_1(x_1,x_2) < x_3 < f_2(x_1,x_2),\ (x_1,x_2) \in CK\}\\
&\quad\geqq \mathbf{P}\{f_1(x_1,x_2) < x_3 < f_2(x_1,x_2),\ (x_1,x_2) \in K\}\\
&\qquad+ \mathbf{P}\{A+\varepsilon < x_3 < B-\varepsilon,\ (x_1,x_2) \in K\}\\
&\quad= \mathbf{P}\{A+\varepsilon < x_3 < B-\varepsilon\}\mathbf{P}\{(x_1,x_2) \in K\} \geqq 1/27,
\end{aligned}$$

where $K = \{(x_1,x_2)\colon a_{(1)}-\varepsilon_1 \leqq x_1 \leqq a_{(1)}+\varepsilon_1,\ a_{(2)}-\varepsilon_1 \leqq x_2 \leqq a_{(2)}+\varepsilon_1\}$ and CK is the complement of K. On the other hand,

$$\begin{aligned}
&\mathbf{P}\Big\{x_3 \in \big(f_1(x_1,x_2), f_2(x_1,x_2)\big)\Big\}\\
&\quad= \mathbf{P}\{f_1(x_1,x_2) < x_3 < f_2(x_1,x_2),\ (x_1,x_2) \in K\}\\
&\qquad+ \mathbf{P}\{f_1(x_1,x_2) < x_3 < f_2(x_1,x_2),\ (x_1,x_2) \in CK\}\\
&\quad\leqq \mathbf{P}(A-\varepsilon < x_3 < B+\varepsilon)\mathbf{P}\{(x_1,x_2) \in K\}\\
&\qquad+ \mathbf{P}\{(x_1,x_2) \in CK\} \leqq \frac{2}{27} + \frac{8}{9} = \frac{26}{27} < 1.
\end{aligned}$$

The lemma is proved.

Let us now go over to the study of alternative 3. Consider the distribution $F_0(u)$, which is uniform in $(a_{(1)}-\varepsilon_1, a_{(1)}+\varepsilon_1)$, and the distribution $F_1(u)$, which is uniform in $(a_{(2)}-\varepsilon_1, a_{(2)}+\varepsilon_1)$. Then

$$\begin{aligned}
&\int_{a_{(1)}-\varepsilon_1}^{a_{(1)}+\varepsilon_1} \int_{a_{(2)}-\varepsilon_1}^{a_{(2)}+\varepsilon_1} \Big[F_0\big(f_2(u_1,u_2)\big) - F_0\big(f_1(u_1,u_2)\big)\Big]\, dF_0(u_1)\, dF_1(u_1)\\
&\qquad= \frac{1}{4\varepsilon_1^2} \int_{a_{(1)}-\varepsilon_1}^{a_{(1)}+\varepsilon_1} \int_{a_{(2)}-\varepsilon_1}^{a_{(2)}+\varepsilon_1} \Big[F_1\big(f_2(u_1,u_2)\big) - F_1\big(f_1(u_1,u_2)\big)\Big]\, du_1\, du_2 = 0.
\end{aligned}$$

This and equality (3) yield

$$(5)\qquad J = \frac{1}{4\varepsilon_1^2}\int_{a_{(1)}-\varepsilon_1}^{a_{(1)}+\varepsilon_1}\int_{a_{(2)}-\varepsilon_1}^{a_{(2)}+\varepsilon_1}\Big[F_1\big(f_2(u_1,u_2)\big) - F_1\big(f_1(u_1,u_2)\big)\Big]\,du_1\,du_2 = 3A^*.$$

By Lemma 3, the probability A^* lies in $(0,1)$ provided that $(f_1(x_1,x_2), f_2(x_1,x_2))$ is an invariant confidence interval; thus, the equalities $A^* = 0$ or $A^* = 1$ contradict the conditions of the theorem.

Denote $A_1 = f_1(a_{(1)}, a_{(1)})$, $B_1 = f_2(a_{(1)}, a_{(1)})$. If $A_1 = B_1$, then for the uniform distribution in the interval $(a_{(1)} - \varepsilon_1, a_{(1)} + \varepsilon_1)$ the probability A^* equals to zero. Assume $A_1 < B_1$. Suppose first that at least one of the inequalities $a_{(1)} \neq A_1$, $a_{(2)} \neq B_1$ holds. The following arrangements of $a_{(1)}$, $a_{(2)}$, B_1, A_1 are possible:

1. $A_1 < B_1 < a_{(2)}$; then

$$\begin{aligned} J &= \int_{a_{(1)}-\varepsilon_1}^{a_{(1)}+\varepsilon_1}\int_{a_{(1)}-\varepsilon_1}^{a_{(1)}+\varepsilon_1}\Big[F_1\big(f_2(u_1,u_2)\big) - F_1\big(f_1(u_1,u_2)\big)\Big]\,dF_0(u_1)\,dF_0(u_2) \\ &= \frac{1}{4\varepsilon_1^2}\int_{a_{(1)}-\varepsilon_1}^{a_{(1)}+\varepsilon_1}\int_{a_{(1)}-\varepsilon_1}^{a_{(1)}+\varepsilon_1}\Big[F_1\big(f_2(u_1,u_2)\big) - F_1\big(f_1(u_1,u_2)\big)\Big]\,du_1\,du_2 = 0, \end{aligned}$$

since $F_1\big(f_2(u_1,u_2)\big) = 0$ and $F_1\big(f_1(u_1,u_2)\big) = 0$ if $(u_1,u_2) \in K_1 = \big\{(u_1,u_2)\colon u_1 \in (a_{(1)} - \varepsilon_1, a_{(1)} + \varepsilon_1),\ u_2 \in (a_{(1)} - \varepsilon_1, a_{(1)} + \varepsilon_1)\big\}$. This and formula (5) yield $A^* = 0$.

2. $a_{(2)} < A_1 < B$; moreover, $J = 0$, since $F_1\big(f_2(u_1,u_2)\big) = 1$ and $F_1\big(f_1(u_1,u_2)\big) = 1$ if $(u_1,u_2) \in K_1$; thus, due to (5), $A^* = 0$.

3. $a_{(1)} < A_1 < a_{(2)} < B_1$ and $A_1 < a_{(1)} < a_{(2)} < B_1$. For these cases consider a sample (x_1, x_2, x_3) from the general population with distribution $F_0(u)$; then

$$\begin{aligned} A^* &= \mathbf{P}\Big\{x_3 \in \big(f_1(x_1,x_2),\ f_2(x_1,x_2)\big)\Big\} \\ &= \iint_{K_1}\Big[F_0\big(f_2(u_1,u_2)\big) - F_0\big(f_1(u_1,u_2)\big)\Big]\,dF_0(u_1)\,dF_0(u_2) \\ &= \frac{1}{4\varepsilon_1^2}\iint_{K_1}\Big[F_0\big(f_2(u_1,u_2)\big) - F_0\big(f_1(u_1,u_2)\big)\Big]\,du_1\,du_2 \\ &= \begin{cases} 0 & \text{if} \quad a_{(1)} < A_1 < a_{(2)} < B_1, \\ 1 & \text{if} \quad A_1 < a_{(1)} < a_{(2)} < B_1. \end{cases} \end{aligned}$$

4. $A_1 \neq a_{(1)}$, $B_1 = a_{(2)}$. Under these conditions, a sample from a general population with distribution $F_0(u)$ satisfies

$$\begin{aligned} A^* &= \mathbf{P}\Big\{x_3 \in \big(f_1(x_1,x_2),\ f_2(x_1,x_2)\big)\Big\} \\ &= \frac{1}{4\varepsilon_1^2}\iint_{K_1}\Big[F_0\big(f_2(u_1,u_2)\big) - F_0\big(f_1(u_1,u_2)\big)\Big]\,du_1\,du_2 \\ &= \begin{cases} 0 & \text{if} \quad a_{(1)} < A_1, \\ 1 & \text{if} \quad a_{(1)} > A_1. \end{cases} \end{aligned}$$

Assume now that $A_1 = a_{(1)}$, $B_1 = a_{(2)}$. Set $A_2 = f_1(a_{(2)}, a_{(2)})$ and $B_2 = f_2(a_{(2)}, a_{(2)})$. If $A_2 \neq a_{(2)}$ and $B_2 \neq a_{(2)}$, then a uniform distribution in the interval $(a_{(2)} - \varepsilon_1, a_{(2)} + \varepsilon_1)$ satisfies the equality

$$A^* = \mathbf{P}\Big\{x_3 \in \big(f_1(x_1,x_2), f_2(x_1,x_2)\big)\Big\} = \begin{cases} 0 & \text{if} \quad a_{(2)} \notin (A_2, B_2), \\ 1 & \text{if} \quad a_{(2)} \in (A_2, B_2). \end{cases}$$

Suppose now that $A_2 = a_{(2)}$, $B_2 = a_{(2)}$ or $A_2 \neq a_{(2)}$, $B_2 = a_{(2)}$, or else $A_2 = a_{(2)}$, $B_2 \neq a_{(2)}$. Let $\widetilde{\mathbf{G}} = \{G_\alpha,\ \alpha \in [0,1]\}$ be the class of general populations with distribution functions $F_\alpha(u)$, $(\alpha \in [0,1])$ of the form

$$F_\alpha(u) = \begin{cases} \alpha\frac{u-(a_{(1)}-\varepsilon_1)}{2\varepsilon_1} & \text{if } \ a_{(1)} - \varepsilon_1 \leqq u \leqq a_{(1)} + \varepsilon_1, \\ (1-\alpha)\frac{u-(a_{(2)}-\varepsilon_1)}{2\varepsilon_1} + \alpha & \text{if } \ a_{(2)} - \varepsilon_1 \leqq u \leqq a_{(2)} + \varepsilon_1, \\ 0 & \text{if } \ u < a_{(1)} - \varepsilon_1, \\ \alpha & \text{if } \ a_{(1)} - \varepsilon_1 < u < a_{(2)} - \varepsilon_1, \\ 1 & \text{if } \ u > a_{(2)} - \varepsilon_1. \end{cases} \tag{6}$$

Denote

$$K_2 = \{(u_1, u_2)\colon\ u_1 \in (a_{(2)} - \varepsilon_1, a_{(2)} + \varepsilon_1),\ u_2 \in (a_{(2)} - \varepsilon_1, a_{(2)} + \varepsilon_1)\},$$

$$K_3 = \{(u_1, u_2)\colon\ u_1 \in (a_{(2)} - \varepsilon_1, a_{(2)} + \varepsilon_1),\ u_2 \in (a_{(1)} - \varepsilon_1, a_{(1)} + \varepsilon_1)\};$$

recall that

$$K = \{(u_1, u_2)\colon\ u_1 \in (a_{(1)} - \varepsilon_1, a_{(1)} + \varepsilon_1),\ u_2 \in (a_{(2)} - \varepsilon_1, a_{(2)} + \varepsilon_1)\}.$$

Let $x_1^\alpha, x_2^\alpha, x_3^\alpha$ be a sample from a general population with distribution $F_\alpha(\alpha \in [0,1])$. Then the following equality holds:

$$\begin{aligned} &\mathbf{P}\{x_3^\alpha \in (f_1(x_1^\alpha, x_2^\alpha), f_2(x_1^\alpha, x_2^\alpha))\} \\ &\quad = \mathbf{P}\{x_3^\alpha \in (f_1(x_1^\alpha, x_2^\alpha), f_2(x_1^\alpha, x_2^\alpha)),\ (x_1^\alpha, x_2^\alpha) \in K_1\} \\ &\qquad + \mathbf{P}\{x_3^\alpha \in (f_1(x_1^\alpha, x_2^\alpha), f_2(x_1^\alpha, x_2^\alpha)),\ (x_1^\alpha, x_2^\alpha) \in K_2\} \\ &\qquad + \mathbf{P}\{x_3^\alpha \in (f_1(x_1^\alpha, x_2^\alpha), f_2(x_1^\alpha, x_2^\alpha)),\ (x_1^\alpha, x_2^\alpha) \in K_3\} \\ &\qquad + \mathbf{P}\{x_3^\alpha \in (f_1(x_1^\alpha, x_2^\alpha), f_2(x_1^\alpha, x_2^\alpha)),\ (x_1^\alpha, x_2^\alpha) \in K\} \\ &\quad = \mathbf{P}\{x_3^\alpha \in (f_1(x_1^\alpha, x_2^\alpha), f_2(x_1^\alpha, x_2^\alpha)),\ (x_1^\alpha, x_2^\alpha) \in K_1\} \\ &\qquad + \mathbf{P}\{x_3^\alpha \in (f_1(x_1^\alpha, x_2^\alpha), f_2(x_1^\alpha, x_2^\alpha)),\ (x_1^\alpha, x_2^\alpha) \in K_2\}. \end{aligned} \tag{7}$$

Furthermore,

$$\begin{aligned} &\mathbf{P}\Big\{x_3^\alpha \in (f_1(x_1^\alpha, x_2^\alpha),\ f_2(x_1^\alpha, x_2^\alpha)),\ (x_1^\alpha, x_2^\alpha) \in K_1\Big\} \\ &\quad = \iint_{K_1} \Big[F_\alpha(f_2(u_1, u_2)) - F_\alpha(f_1(u_1, u_2))\Big]\, dF_\alpha(u_1)\, dF_\alpha(u_2) \\ &\quad = \frac{\alpha^2}{4\varepsilon_1^2} \iint_{K_1} \Big[F_\alpha(f_2(u_1, u_2)) - F_\alpha(f_1(u_1, u_2))\Big]\, du_1\, du_2 = \alpha^2 B_1^*, \end{aligned} \tag{8}$$

where

$$B_1^* = \frac{1}{4\varepsilon_1^2} \int_{a_{(1)}-\varepsilon_1}^{a_{(1)}+\varepsilon_1} \int_{a_{(1)}-\varepsilon_1}^{a_{(1)}+\varepsilon_1} \Big[F_\alpha(f_2(u_1, u_2)) - F_\alpha(f_1(u_1, u_2))\Big]\, du_1\, du_2.$$

Similarly,

$$\begin{aligned} &\mathbf{P}\Big\{x_3^\alpha \in (f_1(x_1^\alpha, x_2^\alpha), f_2(x_1^\alpha, x_2^\alpha)),\ (x_1^\alpha, x_2^\alpha) \in K_2\Big\} \\ &\quad = \frac{(1-\alpha)^2}{4\varepsilon_1^2} \iint_{K_2} \Big[F_\alpha(f_2(u_1, u_2)) - F_\alpha(f_1(u_1, u_2))\Big]\, du_1\, du_2 = (1-\alpha)^2 B_2^*, \end{aligned} \tag{9}$$

where

$$B_2^* = \frac{1}{4\varepsilon_1^2} \int_{a_{(2)}-\varepsilon_1}^{a_{(2)}+\varepsilon_1} \int_{a_{(2)}-\varepsilon_1}^{a_{(2)}+\varepsilon_1} \Big[F_\alpha\big(f_2(u_1,u_2)\big) - F_\alpha\big(f_1(u_1,u_2)\big)\Big]\, du_1\, du_2.$$

Using equalities (7), (8), and (9), we obtain

$$\begin{aligned} A^* &= \mathbf{P}\Big\{x_3^\alpha \in \big(f_1(x_1^\alpha, x_2^\alpha), f_2(x_1^\alpha, x_2^\alpha)\big)\Big\} \\ &= \frac{\alpha^2}{4\varepsilon_1^2} \iint_{K_1} \Big[F_\alpha\big(f_2(u_1,u_2)\big) - F_\alpha\big(f_1(u_1,u_2)\big)\Big]\, du_1\, du_2 \\ &\quad + \frac{(1-\alpha)^2}{4\varepsilon_1^2} \iint_{K_1} \Big[F_\alpha\big(f_2(u_1,u_2)\big) - F_\alpha\big(f_1(u_1,u_2)\big)\Big]\, du_1\, du_2 \\ &= \alpha^2 B_1^* + (1-\alpha)^2 B_2^* = \varphi(\alpha), \end{aligned}$$

where $B_1^* > 0$, $B_2^* > 0$. It is obvious that the expression $\varphi(\alpha) = \alpha^2 B_1^* + (1-\alpha)^2 B_2^*$ depends on α contradicting the conditions of the theorem. Thus, the theorem is proved for case 3. Its proof for cases 4, 5, and 6 is carried out analogously.

Let now $A = f_1(a_1,a_2) \neq a_{(1)} \neq a_{(2)}$, $B = f_2(a_1,a_2) = a_{(2)}$; then one can assume, without loss of generality, that in some neighbourhood $K = \big\{(u_1,u_2)\colon u_1 \in (a_{(1)} - \varepsilon_1, a_{(1)} + \varepsilon_1),\ u_2 \in (a_{(2)} - \varepsilon_1, a_{(2)} + \varepsilon_1)\big\}$ of the point (a_1,a_2) the function $f_2(u_1,u_2)$ coincides with the fundamental symmetric function: $f_2(u_1,u_2) = u_2\ \forall (u_1,u_2) \in K$. Choose $\varepsilon > 0$ and $\varepsilon_1 < \varepsilon$ in such a way that the intervals $(A-\varepsilon, A+\varepsilon)$, $(a_{(1)} - \varepsilon_1, a_{(1)} + \varepsilon_1)$, $(a_{(2)} - \varepsilon_1, a_{(2)} + \varepsilon_1)$ do not intersect and $A - \varepsilon \leqq f_1(u_1,u_2) \leqq A + \varepsilon$, $a_{(2)} - \varepsilon \leqq f_2(u_1,u_2) \leqq a_{(2)} + \varepsilon$ if $a_{(1)} - \varepsilon_1 \leqq u_1 \leqq a_{(1)} + \varepsilon_1$, $a_{(2)} - \varepsilon_1 \leqq u_2 \leqq a_{(2)} + \varepsilon_1$.

It is easy to see that Lemma 3 remains correct if the intervals $(a_{(1)} - \varepsilon_1, a_{(1)} + \varepsilon_1), (A-\varepsilon, A+\varepsilon)$, $(a_{(2)} - \varepsilon_1, a_{(2)} + \varepsilon_1)$ are mutually disjoint; thus, in this case

$$0 < A^* < 1. \tag{10}$$

Suppose first that there exists a point (a,a) lying on the diagonal such that

$$\bar{A}_1 = f_1(a,a) \neq a, \qquad \bar{B}_1 = f_2(a,a) \neq a. \tag{11}$$

Reducing if necessary the numbers ε, ε_1, one can achieve it that the intervals $(\bar{A}_1 - \varepsilon, \bar{A}_1 + \varepsilon)$, $(\bar{B}_1 - \varepsilon, \bar{B}_1 + \varepsilon)$ do not intersect $(a - \varepsilon_1, a + \varepsilon_1)$, and that $f_1(u_1,u_2) \in (\bar{A}_1 - \varepsilon, \bar{A}_1 + \varepsilon)$, $f_2(u_1,u_2) \in (\bar{B}_1 - \varepsilon, \bar{B}_1 + \varepsilon)$ if $a - \varepsilon_1 \leqq u_1 \leqq a + \varepsilon_1$, $a - \varepsilon_1 \leqq u_2 \leqq a + \varepsilon_1$.

Consider the distribution $F(u)$, which is uniform on the segment $[a - \varepsilon_1, a + \varepsilon_1]$; then the sample x_1, x_2, x_3 from the general population with distribution $F(u)$ satisfies the equality

$$\begin{aligned} A^* &= \mathbf{P}\Big\{x_3 \in \big(f_1(x_1,x_2), f_2(x_1,x_2)\big)\Big\} \\ &= \frac{1}{4\varepsilon_1^2} \int_{a-\varepsilon_1}^{a+\varepsilon_1} \int_{a-\varepsilon_1}^{a+\varepsilon_1} \Big[F\big(f_2(u_1,u_2)\big) - F\big(f_1(u_1,u_2)\big)\Big]\, du_1\, du_2 \\ &= \begin{cases} 0 & \text{if } \bar{A}_1 < a < \bar{B}_1, \\ 1 & \text{if } a \notin (\bar{A}_1, \bar{B}_1); \end{cases} \end{aligned}$$

and that contradicts the inequality (10).

If the conditions (11) are not satisfied the two following alternatives occur:

1. $f_1(u,u) = u,\ f_2(u,u) = u\ \forall u \in \mathbf{R}^1$,

2. $f_2(u,u) \neq u$ at the points (u,u), where $f_1(u,u) = u$.

Consider first case 1. Choose $\varepsilon_1 > 0$ in such a way that all the previous relations in which ε_1 occurs are satisfied and, moreover, $f_1(u_1,u_2) \in (a_{(i)} - \varepsilon, a_{(i)} + \varepsilon)$, $f_2(u_1,u_2) \in (a_{(i)} - \varepsilon, a_{(i)} + \varepsilon)$ if $(u_1,u_2) \in \{(u_1,u_2)\colon u_1 \in (a_{(i)} - \varepsilon_1, a_{(i)} + \varepsilon_1),\ u_2 \in (a_{(i)} - \varepsilon_1, a_{(i)} + \varepsilon_1)\}$ $(i = 1,2)$. Let $F_1(u)$ be a uniform distribution on the interval $(a_{(2)} - \varepsilon_1, a_{(2)} + \varepsilon_1)$ and $F_2(u)$ on the interval $(a_{(1)} - \varepsilon_1, a_{(1)} + \varepsilon_1)$. Then, by formula (2),

$$\begin{aligned}
3A^* &= 2\iint_{-\infty}^{\infty} \Big[F_1\big(f_2(u_1,u_2)\big) - F_1\big(f_1(u_1,u_2)\big)\Big]\, dF_1(u_1)\, dF_2(u_2) \\
&\quad + \iint_{-\infty}^{\infty} \Big[F_2\big(f_2(u_1,u_2)\big) - F_2\big(f_1(u_1,u_2)\big)\Big]\, dF_1(u_1)\, dF_1(u_2) \\
&= \frac{2}{4\varepsilon_1^2} \int_{a_{(1)}-\varepsilon_1}^{a_{(1)}+\varepsilon_1} \int_{a_{(2)}-\varepsilon_1}^{a_{(2)}+\varepsilon_1} F_1\big(f_2(u_1,u_2)\big)\, du_1\, du_2 \\
&\quad + \frac{1}{4\varepsilon_1^2} \int_{a_{(2)}-\varepsilon_1}^{a_{(2)}+\varepsilon_1} \int_{a_{(2)}-\varepsilon_1}^{a_{(2)}+\varepsilon_1} \Big[F_2\big(f_2(u_1,u_2)\big) - F_2\big(f_1(u_1,u_2)\big)\Big]\, du_1\, du_2 \\
&= \frac{2}{4\varepsilon_1^2} \int_{a_{(1)}-\varepsilon_1}^{a_{(1)}+\varepsilon_1} \left[\int_{a_{(2)}-\varepsilon_1}^{a_{(2)}+\varepsilon_1} \frac{u_2 - (a_{(2)} - \varepsilon_1)}{2\varepsilon_1}\, du_2\right] du_1 = 1;
\end{aligned}$$

thus,

$$A^* = \frac{1}{3}. \tag{12}$$

Let now $F_i(u)$ be a uniform distribution in the interval $(a_{(i)} - \varepsilon_1, a_{(i)} + \varepsilon_1)$ $(i = 1,2)$. Using formula (2) as in the previous case, we obtain

$$\begin{aligned}
3A^* &= \frac{2}{4\varepsilon_1^2} \int_{a_{(1)}-\varepsilon_1}^{a_{(1)}+\varepsilon_1} \int_{a_{(2)}-\varepsilon_1}^{a_{(2)}+\varepsilon_1} \Big[F_1\big(f_2(u_1,u_2)\big) - F_1\big(f_1(u_1,u_2)\big)\Big]\, du_1\, du_2 \\
&\quad + \frac{1}{4\varepsilon_1^2} \int_{a_{(1)}-\varepsilon_1}^{a_{(1)}+\varepsilon_1} \int_{a_{(1)}-\varepsilon_1}^{a_{(1)}+\varepsilon_1} \Big[F_2\big(f_2(u_1,u_2)\big) - F_2\big(f_1(u_1,u_2)\big)\Big]\, du_1\, du_2 \\
&= \frac{2}{4\varepsilon_1^2} \int_{a_{(1)}-\varepsilon_1}^{a_{(1)}+\varepsilon_1} \int_{a_{(2)}-\varepsilon_1}^{a_{(2)}+\varepsilon_1} \Big[1 - F_1\big(f_1(u_1,u_2)\big)\Big]\, du_1\, du_2 = \begin{cases} 0 & \text{if } A > a_{(1)}, \\ 2 & \text{if } A < a_{(1)}. \end{cases}
\end{aligned}$$

Hence, $A^* = 0$ if $A > a_{(1)}$, and $A^* = \frac{2}{3}$ if $A < a_{(1)}$, which contradicts the equality (12).

We pass now to the study of case 2. Condition 2 implies that one of the following relations holds at the point $(a_{(2)}, a_{(2)})$:

(a) $f_1(a_{(2)}, a_{(2)}) = a_{(2)}$, $f_2(a_{(2)}, a_{(2)}) > a_{(2)}$,

(b) $f_2(a_{(2)}, a_{(2)}) = a_{(2)}$, $f_1(a_{(2)}, a_{(2)}) < a_{(2)}$.

Since conditions (4) do not hold in this case, the function $f_1(u_1,u_2)$ is equal to $u_{(2)}$ in the points where $f_1(u_1,u_2) \neq u_{(1)}$, $f_1(u_1,u_2) \neq u_{(2)}$. Consider case (a). We have $f_1(a_{(2)}, a_{(2)}) \neq a_{(2)}$, $f_2(a_{(2)}, a_{(2)}) = a_{(2)}$, where the last equality may be assumed to hold in some neighbourhood $K_2 = \{(u_1,u_2)\colon a_{(2)} - \varepsilon_1 \leqq u_1 \leqq a_{(2)} + \varepsilon_1,\ a_{(2)} - \varepsilon_1 \leqq u_2 \leqq a_{(2)} + \varepsilon_1\}$ of the point $(a_{(2)}, a_{(2)})$. Let $\mathbf{G}_{[a_{(2)}-\varepsilon_1, a_{(2)}+\varepsilon_1]}$ be the class of general populations with continuous distribution functions which are concentrated on the segment $[a_{(2)} - \varepsilon_1, a_{(2)} + \varepsilon_1]$. It is clear that, for this class of distributions,

$$\begin{aligned} A^* &= \mathbf{P}\{x_3 < f_2(x_1, x_2)\} - \mathbf{P}\{x_3 < f_1(x_1, x_2)\} \\ &= \int_{a_{(2)}-\varepsilon_1}^{a_{(2)}+\varepsilon_1} \int_{a_{(2)}-\varepsilon_1}^{a_{(2)}+\varepsilon_1} F\big(f_2(u_1, u_2)\big)\, dF(u_1)\, dF(u_2) \\ &\quad - \int_{a_{(2)}-\varepsilon_1}^{a_{(2)}+\varepsilon_1} \int_{a_{(2)}-\varepsilon_1}^{a_{(2)}+\varepsilon_1} F\big(f_1(u_1, u_2)\big)\, dF(u_1)\, dF(u_2), \end{aligned}$$

where the first integral, as it was shown in [3], does not depend on the distribution $F(u)$ of the general population $G \in \mathbf{G}_{[a_{(2)}-\varepsilon_1, a_{(2)}+\varepsilon_1]}$, while the second one varies depending on F. Hence, we arrive at a contradiction, since $(f_1(x_1, x_2), f_2(x_1, x_2))$ is an invariant confidence interval.

Case (b) is analyzed similarly.

For an arbitrary natural number $n > 2$ the proof of the theorem is carried out using the procedure introduced above by analyzing various arrangements of the points $A, B, a_{(1)}, \cdots, a_{(n)}$ which are analogous to alternatives 1–6. For example, consider one of the most general cases, when there exists a point $(a, a, \cdots, a) \in \mathbf{R}^n$ such that

$$\begin{aligned} D_1 &= f_1(a, a, \cdots, a) \neq a, \\ D_2 &= f_2(a, a, \cdots, a) \neq a. \end{aligned} \tag{13}$$

Then, for $D_1 \neq D_2$ one can choose an $\varepsilon > 0$, for which the intervals $(a - \varepsilon, a + \varepsilon)$, $(D_1 - \varepsilon, D_1 + \varepsilon)$, $(D_2 - \varepsilon, D_2 + \varepsilon)$ do not intersect (or, if $D_1 = D_2 = D$, the intervals $(a - \varepsilon, a + \varepsilon)$, $(D - \varepsilon, D + \varepsilon)$ do not intersect). Since the functions $f_1(u_1, u_2, \cdots, u_n)$, $f_2(u_1, u_2, \cdots, u_n)$ are continuous according to the conditions of the theorem, there exists, for any $\varepsilon > 0$, an $\varepsilon_1 \in (0, \varepsilon)$ such that

$$\begin{aligned} D_1 - \varepsilon \leqq f_1(u_1, u_2, \cdots, u_n) \leqq D_1 + \varepsilon, \\ D_2 - \varepsilon \leqq f_2(u_1, u_2, \cdots, u_n) \leqq D_2 + \varepsilon, \end{aligned} \tag{14}$$

for all $(u_1, u_2, \cdots, u_n) \in \Pi = \big\{(u_1, u_2, \cdots, u_n)\colon\ a - \varepsilon_1 \leqq u_1 \leqq a + \varepsilon_1, \cdots, a - \varepsilon_1 \leqq u_n \leqq a + \varepsilon_1\big\}$.

Consider a distribution $F_a(u)$ which is uniform on the segment $[a - \varepsilon_1, a + \varepsilon_1]$. Let $x_1, x_2, \cdots, x_n$, x be a sample from the general population with distribution function $F_a(u)$; then conditions (13), (14), and the definition of $F_a(u)$ imply that

$$\begin{aligned} &\mathbf{P}\Big\{x \in \big(f_1(x_1, \cdots, x_n), f_2(x_1, \cdots, x_n)\big)\Big\} \\ &\quad = \frac{1}{(2\varepsilon_1)^n} \int_{a-\varepsilon_1}^{a+\varepsilon_1} \cdots \int_{a-\varepsilon_1}^{a+\varepsilon_1} \Big[F_a\big(f_2(u_1, \cdots, u_n)\big) - F_a\big(f_1(u_1, \cdots, u_n)\big)\Big]\, du_1 \cdots du_n \\ &\quad = \begin{cases} 0 & \text{if } D_1 < a,\ D_2 < a \text{ or } D_1 > a,\ D_2 > a \quad (D < a), \\ 1 & \text{if } D_1 < a,\ D_2 > a \quad (D > a), \end{cases} \end{aligned}$$

since either $F_a\big(f_2(u_1, \cdots, u_n)\big) - F_a\big(f_1(u_1, \cdots, u_n)\big) = 0$ or $F_a\big(f_2(u_1, \cdots, u_n)\big) - F_a\big(f_1(u_1, \cdots, u_n)\big) = 1$, depending on the arrangements of the points D_1 and D_2 (D) for $(u_1, u_2, \cdots, u_n) \in \Pi$, contradicting Lemma 3.

The theorem is proved.

Note. A result analogous to Theorem 1 has been obtained by Robbins for nonparametrical tolerance intervals ([4], § 2.6); however, these theorems are only slightly related and do not imply each other. Moreover, the proof of Robbins' theorem is much simpler than that of Theorem 1.

Let $\mathbf{G} = \{G_\alpha,\ \alpha \in \mathfrak{A}\}$ be an arbitrary class of general populations, $\{J_t(x_1, x_2, \cdots, x_n),\ t \in T\}$, $(n \in N)$ the set of all invariant confidence intervals for the class $\mathbf{G}$ and $\mathfrak{B}(\mathbf{G}) = \{\beta_{t,n} = \mathbf{P}\{x^\alpha \in J_t(x_1, \cdots, x_n),\ t \in T,\ x^\alpha \in G_\alpha\}\}$ the set of all confidence levels, corresponding to the confidence intervals $J_t(x_1, x_2, \cdots, x_n)$. Theorem 1 and the equality

$$\mathbf{P}\big(x \in (x_{(i)}, x_{(j)})\big) = \frac{j-i}{n+1} \qquad (i < j,\ i, j \in \{1, 2, \cdots, n\})$$

immediately yield the following corollary.

COROLLARY. *The totality $\mathfrak{B}(\mathbf{G}_C)$ of all confidence levels corresponding to all possible invariant confidence intervals, satisfying the conditions of Theorem 1 for the class $\mathbf{G}_C$ of general populations with continuous distribution functions is the set of all rationals from the interval $[0,1] : \mathfrak{B}(\mathbf{G}_C) = [0,1] \cap \mathbf{Q}$.*

THEOREM 2. *Let $(x_1, x_2, \cdots, x_n)$ be a sample from the general population G with continuous distribution, obtained by means of simple sampling, let $J = \big(f_1(x_1, \cdots, x_n), f_2(x_1, \cdots, x_n)\big)$ be an invariant confidence interval constructed from the sample values $x_1, \cdots, x_n$, satisfying the conditions of Theorem 1, and let $x_{n+1}, x_{n+2}, \cdots, x_{n+m}$ be subsequent sample values from the general population G. Then,*

$$\mathbf{P}(x_{n+1}, \cdots, x_{n+m} \in J) = \frac{n!(m+s-r-1)!}{(s-r-1)!(m+n)!},$$

where $r < s$; $r, s \in \{1, 2, \cdots, n\}$.

Proof. Indeed, by virtue of Theorem 1, the invariant confidence interval J for the class of continuous distributions has the form $J = (x_{(t)}, x_{(s)})$, where $r < s$; $r, s \in \{1, 2, \cdots, n\}$. Furthermore, the joint density function of the uniform order statistics $\tilde{x}_{(r)}$ and $\tilde{x}_{(s)}$ is determined by the formula

$$f_{r,s}(u,v) = \frac{n!}{(r-1)!(s-r-1)!(n-s)!}\, u^{r-1}(v-u)^{s-r-1}(1-v)^{n-s}$$

$(0 \leqq u < v \leqq 1)$; thus, the following equality holds:

$$\int_0^1 \int_u^1 u^{r-1}(v-u)^{s-r-1}(1-v)^{n-s}\, dv\, du = (r-1)!(s-r-1)!(n-s)!/n!.$$

Taking Smirnov's transformation into account, we obtain

$$\begin{aligned}
&\mathbf{P}(x_{n+1}, \cdots, x_{n+m} \in J)\\
&\quad = \mathbf{P}\big(x_{n+1} \in (x_{(r)}, x_{(s)}), \cdots, x_{n+m} \in (x_{(r)}, x_{(s)})\big)\\
&\quad = \mathbf{P}\big(\tilde{x}_{n+1} \in (\tilde{x}_{(r)}, \tilde{x}_{(s)}), \cdots, \tilde{x}_{n+m} \in (\tilde{x}_{(r)}, \tilde{x}_{(s)})\big)\\
&\quad = \frac{n!}{(r-1)!(s-r-1)!(n-s)!} \int_0^1 \int_u^1 (v-u)^m f_{r,s}(u,v)\, dv\, du\\
&\quad = \frac{n!}{(r-1)!(s-r-1)!(n-s)!} \int_0^1 \int_u^1 u^{r-1}(v-u)^{s+m-r-1}(1-v)^{n-s}\, dv\, du\\
&\quad = \frac{n!}{(r-1)!(s-r-1)!(n-s)!}\, \frac{(r-1)!(s+m-r-1)!(n-s)!}{(m+n)!}\\
&\quad = \frac{n!(s+m-r-1)!}{(n+m)!(s-r-1)!}.
\end{aligned}$$

The theorem is proved.

COROLLARY. *Under the conditions of Theorem 2, the equality*

$$\mathbf{P}(x_{n+1}, x_{n+2}, \cdots, x_{n+m}, \cdots \in J) = 0$$

holds.

REFERENCES

[1] I. MADREIMOV AND YU. I. PETUNIN, *Characterization of an uniform distribution with the help of order statistics,* Theory Prob. and Math. Statist., 27 (1982), pp. 96–102. (In Russian.)

[2] B. L. VAN DER WAERDEN, *Matematische Statistik,* Springer-Verlag, Berlin–Goettingen–Heidelberg, 1957.

[3] I. G. BAIRAMOV AND YU. I. PETUNIN, *Invariant confidence intervals for the bulk of the distribution,* Vychislit. Prikl. Mat., 65 (1988), pp. 112–117. (In Russian.)

[4] H. A. DAVID, *Order Statistics,* John Wiley, London–New York, 1970.