

A DISTANCE BETWEEN DISTRIBUTIONS CONNECTED WITH THEIR VALUES ON CONVEX SETS

S. M. Sadikova

1967

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-196701.94401>

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.

Abstract

Full Text

UDC 519.214

MATHEMATICS

S. M. SADIKOVA

A DISTANCE BETWEEN DISTRIBUTIONS CONNECTED WITH THEIR VALUES ON CONVEX SETS

(Presented by Academician A. N. Kolmogorov, December 17, 1966)

In the paper ⁽¹⁾ the author obtained an inequality (a two-dimensional analogue of Esseen's inequality ⁽²⁾, § 39) that makes it possible to estimate, in terms of the difference of characteristic functions, the difference of the values $P(A) - Q(A)$ of the corresponding distributions for all rectangles A with sides parallel to the coordinate axes. The purpose of the present note is to derive inequality (4), which gives an upper estimate for the distance

$$\rho(P, Q) = \sup_A |P(A) - Q(A)|$$

(where the supremum is taken over all convex sets of the space R^s) in terms of characteristic functions. Let us note that a method for estimating the difference $P(A) - Q(A)$ for balls was indicated in ⁽³⁾ (see also ⁽⁴⁾).

The starting point in this and similar cases may be the following relation, easily derived from the inversion formula for Fourier integrals. Let ξ, η , and ζ be s -dimensional random vectors, ζ independent of ξ and η and having an absolutely integrable characteristic function $h(t)$. Then

$$\Delta_A = Pr\{\xi + \zeta \in A\} - Pr\{\eta + \zeta \in A\} = \frac{1}{(2\pi)^s} \int_{R^s} \overline{\tau_A(t)} (f(t) - g(t)) h(t) dt.$$

Here A is an arbitrary bounded Borel set;

$$\tau_A(t) = \int_A e^{i(t,x)} dx;$$

f and g are the characteristic functions of the vectors ξ and η , respectively.

Representing the quantity $(2\pi)^s \Delta_A$ as the sum $I_1 + I_2 + I_3$, where

$$I_1 = \int_{|t| \leq 1}, \quad I_2 = \int_{1 < |t| < T}, \quad I_3 = \int_{|t| \geq T},$$

and applying the Cauchy–Bunyakovsky inequality to each integral, we obtain

$$(2\pi)^s |\Delta_A| \leq \left(\int_{|t| \leq 1} |t|^2 |\tau_A|^2 \right)^{1/2} J_1 + \left(\int_{1 < |t| < T} |\tau_A|^2 \right) J_2 + 2 \left(\int_{|t| \geq T} |\tau_A|^2 \right)^{1/2} J_3, \quad (1)$$

where

$$J_1 = \left(\int_{|t| \leq 1} \frac{|f-g|^2}{|t|^2} \right)^{1/2}, \quad J_2 = \left(\int_{1 < |t| < T} |f-g|^2 \right)^{1/2}, \quad J_3 = \left(\int_{|t| \geq T} |h|^2 \right)^{1/2}.$$

Let O_r denote the ball of radius r centered at the origin, and let

$$\Delta_r = \sup_{A \subset O_r} |\Delta_A|,$$

where the supremum is taken over all convex $A \subset O_r$. As is easy to see,

$$\gamma_1 = \sup_A |\Delta_A| \leq 2\Delta_r + 2Pr\{\eta + \zeta \in \overline{O_r}\}.$$

The quantity Δ_r can be estimated with the aid of inequality (1). Indeed, if A is a convex set lying in the ball O_r , and $S(A)$ is the magnitude of its surface, then

$$S(A) \leq S(O_r) \leq 2\pi^{s/2} \Gamma^{-1}(s/2) r^{s-1}.$$

Therefore, taking into account the inequalities for the integrals of $\tau_A(t)$ (formulas of the paper ⁽¹⁾, p. 379), we obtain, for $A \subset O_r$,

$$|\Delta_A| \leq \frac{\sqrt{\lambda_s S(O_r)}}{(2\pi)^s} \left[J_1 + 2\sqrt{2} J_2 + \frac{4\sqrt{2}}{\sqrt{T}} J_3 \right],$$

where

$$\lambda_s = \frac{(2\pi)^{s+1}}{4} \left(\int_0^\pi \sin^s \alpha \, d\alpha \right)^{-1}. \quad (2)$$

Denote the right-hand side of (2) by Λ_r^* . Then

$$\gamma_1 \leq 2\Lambda_r^* + 2Pr\{\eta + \zeta \in \overline{O_r}\}.$$

Here the constants r and T and the random variable ζ are at our disposal. The lemma stated below serves for passing from γ_1 to $\gamma = \rho(P, Q)$. Before formulating it, introduce the following notation: A is a convex set; $\dot{A}$ is its boundary; $(\dot{A})^\delta$ is the δ -neighborhood of A ; $\omega_\eta(\delta) = \sup_A \Pr\{\eta \in (\dot{A})^\delta\}$. Let us note that, for any A , the inequalities

$$\Pr\{\xi \in A\} \leq \Pr\{\xi + \zeta \in A \cup (\dot{A})^\delta\} + \Pr\{|\zeta| \geq \delta\}, \quad (3)$$

$$\Pr\{\xi \in A\} \geq \Pr\{\xi + \zeta \in A \setminus (\dot{A})^\delta\} - \Pr\{|\zeta| \geq \delta\}.$$

It is also easy to prove that $\omega_{\eta+\zeta}(\delta) \leq \omega_\eta(\delta)$.

Lemma. The inequalities hold

$$\gamma_1 \leq \gamma, \quad \gamma \leq \gamma_1 + \omega_\eta(\delta) + 2\Pr\{|\zeta| \geq \delta\}.$$

We omit the proof of the lemma.

Let us now note that

$$\begin{aligned} \Pr\{\eta + \zeta \in \overline{O}_r\} &\leq \Pr\{(\eta + \zeta \in \overline{O}_r) \cap (|\zeta| < \delta)\} + \Pr\{|\zeta| \geq \delta\} \leq \\ &\leq \Pr\{|\eta| \geq r\} + \omega_\eta(\delta) + \Pr\{|\zeta| \geq \delta\}. \end{aligned}$$

Therefore

$$\gamma = \rho(P, Q) \leq 2\Lambda_r^* + 2\Pr\{|\eta| \geq r\} + 3\omega_\eta(\delta) + 4\Pr\{|\zeta| \geq \delta\},$$

$$\begin{aligned} \rho(P, Q) &\leq C_s r^{(s-1)/2} \left[J_1 + 2\sqrt{2}J_2 + \frac{4\sqrt{2}}{\sqrt{T}}J_3 \right] + \\ &+ 3\omega_\eta(\delta) + 2\Pr\{|\eta| \geq r\} + 4\Pr\{|\zeta| \geq \delta\}, \end{aligned} \quad (4)$$

where $C_s = \sqrt{\lambda_s S(O_1)}/(2\pi)^s$.

From Stirling's formula it is not difficult to derive that

$$C_s = \sqrt[4]{2} [s(s+1)]^{1/4} \left(\frac{e}{2\pi s} \right)^{s/4} e^{\theta/s + \theta'/12s}, \quad |\theta|, |\theta'| \leq 1.$$

The function $h(t)$ and the constants r, δ , and T entering inequality (4) may be chosen depending on our aims.

We shall now show how the function $\omega_\eta(\delta)$ can be estimated in the example of the normal distribution in R^s with density

$$p_\eta(x) = \frac{1}{(2\pi)^{s/2}} e^{-|x|^2/2}.$$

Let A be an arbitrary convex set in R^s . Without loss of generality, one may assume that A does not lie in any subspace of R^s of dimension $s' < s$. Consider the system of concentric spheres O_n , $n = 0, 1, 2, \dots$, with centers at the origin and such that the radius of O_n is equal to n . The part of the surface A lying between the n -th and $(n+1)$ -st spheres does not exceed its part lying inside the $(n+1)$ -st sphere. Suppose that $\delta < 1$. Then the measure of the part of the δ -neighborhood of A lying between the n -th and $(n+1)$ -st spheres does not exceed the measure of the δ -neighborhood of the $(n+1)$ -st sphere, i.e., in any case, does not exceed $2\delta S(O_{n+2})$. The probability of hitting the mentioned part of the δ -neighborhood of A does not exceed

$$\frac{1}{(2\pi)^{s/2}} e^{-(n-1)^2/2} 2\delta \frac{2\pi^{s/2}}{\Gamma(s/2)} (n+2)^{s-1}$$

(for $n \geq 1$; for $n = 0$ the factor $e^{-(n-1)^2/2}$ must be replaced by 1). Therefore

$$Pr\{\eta \in (\dot{A})^\delta\} \leq \frac{4\delta}{2^{s/2}\Gamma(s/2)} \left[2^{s-1} + \sum_{n=1}^{\infty} e^{-(n-1)^2/2} (n+2)^{s-1} \right].$$

From this it is not difficult to derive that for $\delta < 1$

$$\omega_\eta(\delta) \leq \delta \frac{4}{2^{s/2}\Gamma(s/2)} \left[2^{s-1} + 3^{s-1} + 2^{3s/2-3}\Gamma(s/2) + 2^{3s-4} \frac{\sqrt{2\pi}}{2} \right].$$

From the last inequality and Stirling's formula it follows that

$$\omega_\eta(\delta) \leq \delta C_{\text{abs}} \cdot 2^s,$$

where C_{abs} is an absolute constant.

Moscow Engineering Physics
Institute

Received
13 XII 1966

REFERENCES

1. S. M. Sadikova, *Theory of Probability and Its Applications*, **11**, 3 (1966).
2. B. V. Gnedenko, A. N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables*, Moscow, 1949.
3. K. G. Esseen, *Acta Math.*, **77**, 1 (1944).
4. G. de Barra, *Proc. Cambridge Phil. Soc.*, **59**, 2 (1963).

Note: Figure translations are in progress. See original paper for figures.

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.