Известия Саратовского университета. Новая серия. Серия: Математика. Механика. Информатика. 2022. Т. 22, вып. 2. С. 233–240
Izvestiya of Saratov University. Mathematics. Mechanics. Informatics, 2022, vol. 22, iss. 2, pp. 233–240
https://mmi.sgu.ru https://doi.org/10.18500/1816-9791-2022-22-2-233-240

Article

What scientific folklore knows about the distances between the most popular distributions

M. Y. Kelbert[1], **Yu. M. Suhov**[2,3,4]✉

[1]Higher School of Economics – National Research University, 20 Myasnitskaya St., Moscow 101000, Russia
[2]University of Cambridge, The Old Schools, Trinity Ln, Cambridge CB2 1TN, UK
[3]University of Pennsylvania, 201 Old Main, State College, PA 16802, USA
[4]Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute), 19 build. 1 Bolshoy Karetny per., Moscow 127051, Russia
Mark Y. Kelbert, mkelbert@hse.ru, https://orcid.org/0000-0002-3952-2012, AuthorID: 1137288
Yurii M. Suhov, yms@statslab.cam.ac.uk, AuthorID: 1131362

Abstract. We present a number of upper and low bounds for the total variation distances between the most popular probability distributions. In particular, some estimates of the total variation distances between one-dimensional Gaussian distributions, between two Poisson distributions, between two binomial distributions, between a binomial and a Poisson distribution, and also between two negative binomial distributions are given. The Kolmogorov – Smirnov distance is also presented.
Keywords: probability distribution, variation distance, Pinsker's inequality, Le Cam's inequalities, distances between distributions
For citation: Kelbert M. Y., Suhov Yu. M. What scientific folklore knows about the distances between the most popular distributions. *Izvestiya of Saratov University. Mathematics. Mechanics. Informatics*, 2022, vol. 22, iss. 2, pp. 233–240. https://doi.org/10.18500/1816-9791-2022-22-2-233-240

Научная статья
УДК 519.85

Что научный фольклор знает о расстояниях между наиболее популярными распределениями

М. Я. Кельберт[1], **Ю. М. Сухов**[2,3,4]✉

[1]Национальный исследовательский университет «Высшая школа экономики», Россия, 101000, г. Москва, ул. Мясницкая, д. 20
[2]Кембриджский университет, Великобритания, The Old Schools, Trinity Ln, Cambridge CB2 1TN
[3]Университет штата Пенсильвания, Соединенные Штаты Америки, Пенсильвания, 16802, г. Стейт-Колледж, кампус Юниверсити-Парк, ул. Олд Мейн, д. 201
[4]Институт проблем передачи информации имени А. А. Харкевича Российской академии наук, Россия, 127051, г. Москва, Б. Каретный пер., д. 19, стр. 1

Кельберт Марк Яковлевич, кандидат физико-математических наук, профессор-исследователь департамента статистики и анализа данных факультета экономических наук, mkelbert@hse.ru, https://orcid.org/0000-0002-3952-2012, AuthorID: 1137288

Сухов Юрий Михайлович, кандидат физико-математических наук, [2]профессор кафедры чистой математики и математической статистики; [3]профессор математического факультета; [4]научный сотрудник, yms@statslab.cam.ac.uk, AuthorID: 1131362

Аннотация. Представлен ряд верхних и нижних оценок для расстояний по вариации между наиболее популярными распределениями вероятностей. В частности, приводятся оценки расстояний по вариации между одномерными гауссовскими, между двумя пуассоновскими, между двумя биномиальными распределениями, между биномиальным и пуассоновским распределениями и между двумя негативными биномиальными распределениями. Также исследуется расстояние Колмогорова – Смирнова.

Ключевые слова: распределение вероятностей, расстояние вариации, неравенство Пинскера, неравенства Ле Кама, расстояния между распределениями

Для цитирования: *Kelbert M. Y., Suhov Yu. M.* What scientific folklore knows about the distances between the most popular distributions [*Кельберт М. Я., Сухов Ю. М.* Что научный фольклор знает о расстояниях между наиболее популярными распределениями] // Известия Саратовского университета. Новая серия. Серия: Математика. Механика. Информатика. 2022. Т. 22, вып. 2. С. 233–240. https://doi.org/10.18500/1816-9791-2022-22-2-233-240

Introduction

A tale that becomes folklore is one that is passed down and whispered around. The second half of the word, lore, comes from Old English lār, i.e. 'instruction'. Different bounds for the distances between the most popular probability distributions (see [1]) appear in many problems of applied probability. Unfortunately, the available textbooks and reference books do not present them in a systematic way. In this short note, we make an attempt to fill this gap.

Let us remind that for probability measures $\mathbf{P}, \mathbf{Q}$ with densities p, q

$$\mathrm{TV}(\mathbf{P}, \mathbf{Q}) = \sup_{A \subset \mathbf{R}^d} |\mathbf{P}(A) - \mathbf{Q}(A)| = \frac{1}{2} \int_{\mathbf{R}^d} |p(u) - q(u)| \mathrm{d}u.$$

Let us remind the coupling characterization of the total variation distance. For two distributions $\mathbf{P}$ and $\mathbf{Q}$, a pair (X, Y) of random variables defined on the same probability space is called a *coupling* for $\mathbf{P}$ and $\mathbf{Q}$ if $X \sim \mathbf{P}$ and $Y \sim \mathbf{Q}$.

One of the useful facts is that there exists a coupling (X, Y) such that $\mathbf{P}(X \neq Y) = \mathrm{TV}(\mathbf{P}, \mathbf{Q})$. Therefore, for any function f, we have $\mathbf{P}(f(X) \neq f(Y)) \leqslant \mathrm{TV}(X, Y)$ with equality iff f is reversible.

1. Gaussian distributions

The total variation distance between one-dimensional Gaussian distributions is equal to

$$\tau = \tau(X_1, X_2) = \mathrm{TV}(N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2))$$

and it depends on the parameters $\Delta = |\delta|$, with $\delta = \mu_1 - \mu_2$, and σ_1^2, σ_2^2:

$$\frac{1}{200}\min\left[1, \max[\frac{|\sigma_1^2 - \sigma_2^2|}{\min[\sigma_1^2, \sigma_2^2]}, \frac{40\Delta}{\min[\sigma_1, \sigma_2]}]\right] \leqslant \tau \leqslant \frac{3|\sigma_1^2 - \sigma_2^2|}{2\max[\sigma_1^2, \sigma_2^2]} + \frac{\Delta}{2\max[\sigma_1, \sigma_2]}.$$

In the case $\sigma_1^2 = \sigma_2^2$ the following identity holds: $\tau = \Phi(\frac{\Delta}{2\sigma}) - \frac{1}{2}$.

1.1. Pinsker's inequality

In the general case, the upper bound is a version of *Pinsker's inequality* [2] for $\tau(X_1, X_2) = \mathrm{TV}(X_1, X_2)$:

$$\tau(X_1, X_2) \leqslant \min\{1, \sqrt{\mathrm{KL}(\mathbf{P}_{X_1}||\mathbf{P}_{X_2})/2}\}, \tag{1}$$

where

$$\mathrm{KL}(\mathbf{P}_{X_1}||\mathbf{P}_{X_2}) = \frac{1}{2}\left(\frac{\sigma_2^2}{\sigma_1^2} - 1 + \frac{\Delta^2}{\sigma_1^2} - \ln\frac{\sigma_2^2}{\sigma_1^2}\right).$$

For multidimensional Gaussian case

$$\mathrm{KL}(\mathbf{P}_{X_1}||\mathbf{P}_{X_2}) = \frac{1}{2}\left(\mathrm{tr}(\Sigma_1^{-1}\Sigma_2 - \mathbf{I}) + \delta^T\Sigma_1^{-1}\delta - \ln\det(\Sigma_2\Sigma_1^{-1})\right).$$

Let us prove the Pinsker's inequality (1).

We need the following bound

$$|x-1| \leqslant \sqrt{\left(\frac{4}{3} + \frac{2x}{3}\right)\phi(x)}, \quad \phi(x) = x\ln x - x + 1. \tag{2}$$

If $\mathbf{P}$ and $\mathbf{Q}$ are singular, then $\mathrm{KL} = \infty$ and Pinsker's inequality holds true. Assume $\mathbf{P}$ and $\mathbf{Q}$ are absolutely continuous. In view of (2) and Cauchy – Schwarz inequality

$$\tau(X,Y) = \frac{1}{2}\int |p-q| = \frac{1}{2}\int q|\frac{p}{q} - 1|\mathbf{1}_{\{q>0\}} \leqslant$$
$$\leqslant \frac{1}{2}\left(\int\left(\frac{4q}{3} + \frac{2p}{3}\right)\mathbf{1}_{\{q>0\}}\right)^{1/2}\left(\int q\phi(\frac{p}{q})\mathbf{1}_{\{q>0\}}\right)^{1/2} =$$
$$= \left(\frac{1}{2}\int p\ln(\frac{p}{q})\mathbf{1}_{\{q>0\}\}}\right)^{1/2} = (\mathrm{KL}(\mathbf{P}||\mathbf{Q})/2)^{1/2}.$$

To check (2) define $g(x) = (x-1)^2 - \left(\frac{4}{3} + \frac{2x}{3}\right)\phi(x)$. Then $g(1) = g'(1) = 0$, $g''(x) = -\frac{4\phi(x)}{3x} < 0$. Hence,

$$g(x) = g(1) + g'(1)(x-1) + \frac{1}{2}g''(\xi)(x-1)^2 = -\frac{4\phi(\xi)}{6\xi}(x-1)^2 \leqslant 0.$$

Remark. Mark S. Pinsker was invited to be the Shannon Lecturer at the 1979 IEEE International Symposium on Information Theory, but could not obtain permission at that time to travel to the symposium. However, he was officially recognized by the IEEE Information Theory Society as the 1979 Shannon Award recipient.

1.2. Le Cam's inequalities

Le Cam's inequalities were presented in [3] for Hellinger distance defined by

$$\eta(X,Y) = \frac{1}{\sqrt{2}}\left(\int (\sqrt{p_X(u)} - \sqrt{p_Y(u)})^2 \mathrm{d}u\right)^{1/2}$$

as follows:

$$\eta(X,Y)^2 \leqslant \tau(X,Y) \leqslant \eta(X,Y)\left(2 - \eta(X,Y)^2\right)^{1/2}. \quad (3)$$

For one-dimensional Gaussian distributions we get

$$\eta(X,Y)^2 = 1 - \frac{\sqrt{2\sigma_1\sigma_2}}{\sqrt{\sigma_1^2+\sigma_2^2}} e^{-\frac{\Delta^2}{4(\sigma_1^2+\sigma_2^2)}}.$$

Let us present the proof of Le Cam's inequalities (3).

From $\tau(X,Y) = \frac{1}{2}\int |p-q| = 1 - \int \min[p,q]$ and $\min[p,q] \leqslant \sqrt{pq}$, it follows $\tau(X,Y) \geqslant 1 - \int\sqrt{pq} = \eta^2(X,Y)$. Next, $\int \min[p,q] + \int \max[p,q] = 2$. Therefore, by Cauchy – Schwarz inequality we get

$$\left(\int \sqrt{pq}\right)^2 = \left(\int \sqrt{\min[p,q]\max[p,q]}\right)^2 \leqslant \int \min[p,q] \int \max[p,q] =$$
$$= \int \min[p,q]\left(2 - \int \min[p,q]\right).$$

Hence, it follows from

$$\left(1 - \eta(X,Y)^2\right)^2 \leqslant (1-\tau(X,Y))(1+\tau(X,Y))$$

that

$$\tau(X,Y) \leqslant \eta(X,Y)\left(2 - \eta(X,Y)^2\right)^{1/2}.$$

2. Poisson and binomial distributions

2.1. Two Poisson distributions

Let X_i are Poisson distributed random variables, i.e. $X_i \sim \mathrm{Po}(\lambda_i)$, where $0 < \lambda_1 < \lambda_2$. Then the distance between two Poisson distributions is

$$\tau(X_1,X_2) = \int_{\lambda_1}^{\lambda_2} \mathbf{P}(N(u) = l-1)\mathrm{d}u \leqslant \min\left[\lambda_2 - \lambda_1, \sqrt{\frac{2}{e}}(\sqrt{\lambda_2} - \sqrt{\lambda_1})\right],$$

where $N(u) \sim \mathrm{Po}(u)$. Here $\lceil\lambda_1\rceil \leqslant l \leqslant \lceil\lambda_2\rceil$, and

$$l = l(\lambda_1,\lambda_2) = \lceil(\lambda_2-\lambda_1)\left(\ln(\lambda_2/\lambda_1)\right)^{-1}\rceil.$$

2.2. Distances between binomial distributions

Let X_i are drawn from binomial distributions, i.e. $X_i \sim \mathrm{Bin}(n,p_i)$, $0 < p_1 < p_2 < 1$. Then the distance between two binomial distributions is equal to

$$\tau(X_1,X_2) = n\int_{p_1}^{p_2} \mathbf{P}(S_{n-1}(u) = l-1)\mathrm{d}u \leqslant \frac{\sqrt{e}}{2}\frac{\psi(p_2-p_1)}{(1-\psi(p_2-p_1))^2},$$

where $S_{n-1}(u)$ ~Bin$(n-1,u)$ and $\psi(x) = x\sqrt{\frac{n+2}{2p_1(1-p_1)}}$. Finally, define

$$l = \left\lceil \frac{-n\ln(1-\frac{p_2-p_1}{1-p_1})}{\ln(1+\frac{p_2-p_1}{p_1}) - \ln(1-\frac{p_2-p_1}{1-p_1}))} \right\rceil$$

with $\lceil np_1 \rceil \leqslant l \leqslant \lceil np_2 \rceil$.

2.3. Distance between binomial and Poisson distributions

Let X ~Bin(n,p) and Y ~Po(np), $0 < np < 2-\sqrt{2}$, then

$$\tau(X,Y) = np[(1-p)^{n-1} - e^{-np}].$$

For the sum of Bernoulli r.v. $S_n = \sum\limits_{j=1}^{n} X_j$ with $\mathbf{P}(X_i = 1) = p_i$ we have

$$\tau(S_n, Y_n) = \frac{1}{2}\sum_{k=1}^{\infty} |\mathbf{P}(S_n = k) - \frac{\lambda_n^k}{k!}e^{-\lambda_n}| < \sum_{i=1}^{n} p_i^2,$$

where Y_n ~Po(λ_n), $\lambda_n = p_1 + p_2 + \ldots + p_n$ [4]. A stronger result: for $X_i \sim$ Bernoulli(p_i) and $Y_i \sim$ Po$(\lambda_i = p_i)$ there exists a coupling such that

$$\tau(X_i, Y_i) = \mathbf{P}(X_i \neq Y_i) = p_i(1 - e^{-p_i}).$$

2.4. Distance between negative binomial distributions

Let X_i be drawn from negative binomial distributions, i.e. $X_i \sim$ NegBin(m, p_i), $0 < p_1 < p_2 < 1$. Then

$$\tau(X_1, X_2) = (m + l - 1)\int_{p_1}^{p_2} \mathbf{P}(S_{m+l-2}(u) = m-1)\mathrm{d}u,$$

where $S_n(u)$ ~Bin(n,u) and

$$l = \lceil -m\frac{\ln(1+\frac{p_2-p_1}{p_1})}{\ln(1-\frac{p_2-p_1}{1-p_1})} \rceil$$

with $\lceil m\frac{1-p_2}{p_2} \rceil \leqslant l \leqslant \lceil m\frac{1-p_1}{p_1} \rceil$.

3. Multidimensional Gaussian distributions

In the case of multidimensional Gaussian distributions the distance is

$$\tau = \mathrm{TV}(N(\mu_1, \Sigma_1), N(\mu_2, \Sigma_2)),$$

where Σ_1, Σ_2 are positive-definite.

Let $\delta = \mu_1 - \mu_2$ and Π be a $d \times (d-1)$ matrix whose columns form a basis for subspace orthogonal to δ. Let $\lambda_1, \ldots, \lambda_{d-1}$ denote the eigenvalues of the matrix $(\Pi^T\Sigma_1\Pi)^{-1}\Pi^T\Sigma_2\Pi - \mathbf{I}_{d-1}$ and $\lambda = \sqrt{\sum\limits_{i=1}^{d-1}\lambda_i^2}$. If $\mu_1 \neq \mu_2$ then

$$\frac{1}{200}\min[1, \varphi(\delta, \Sigma_1, \Sigma_2)] \leqslant \tau \leqslant \frac{9}{2}\min[1, \varphi(\delta, \Sigma_1, \Sigma_2)], \qquad (4)$$

where
$$\varphi(\delta, \Sigma_1, \Sigma_2) = \max\left[\frac{\delta^T(\Sigma_1 - \Sigma_2)\delta}{\delta^T\Sigma_1\delta}, \frac{\sqrt{\delta^T\delta}}{\sqrt{\delta^T\Sigma_1\delta}}, \lambda\right].$$

In the case of equal means $\mu_1 = \mu_2$ the bound (4) is simplified as follows:
$$\frac{1}{100}\min[1, \lambda] \leqslant \tau \leqslant \frac{3}{2}\min[1, \lambda].$$

Here $\lambda = \sqrt{\sum\limits_{j=1}^{d}\lambda_j^2}$, $\lambda_1, \ldots, \lambda_d$ are the eigenvalues of $\Sigma_1^{-1}\Sigma_2 - \mathbf{I}_d$ for positive-definite Σ_1, Σ_2. In the case $\Sigma_1 = \Sigma_2$ the following equality holds: $\tau = \Phi\left(||\Sigma^{-1/2}\delta||/2\right) - \frac{1}{2}$.

Let us present below the sketch of proof, cf. [5].

Let X_i ∼$\mathrm{N}(\mu_i, \Sigma_i), i = 1, 2$. Without the loss of generality we can assume that Σ_1, Σ_2 are positively definite as
$$\mathrm{TV}(N(0, \Sigma_1), N(0, \Sigma_2)) = \mathrm{TV}(N(0, \Pi^T\Sigma_1\Pi), N(0, \Pi^T\Sigma_2\Pi)),$$
where Π is $d \times r$ matrix whose columns form orthogonal bases for range($\Sigma_{1,2}$). Denote $u = (\mu_1 + \mu_2)/2, \delta = \mu_1 - \mu_2$ and decompose $\forall w \in \mathbf{R}^d$ as
$$w = u + f_1(w)\delta + f_2(w), f_2(w)^T\delta = 0.$$

Then
$$\max[\mathrm{TV}(f_1(X_1), f_1(X_2)), \mathrm{TV}(f_2(X_1), f_2(X_2))] \leqslant \mathrm{TV}(X_1, X_2) \leqslant$$
$$\leqslant \mathrm{TV}(f_1(X_1), f_1(X_2)) + \mathrm{TV}(f_2(X_1), f_2(X_2)).$$

All the components are Gaussian and $f_1(X_1)$ ∼$\mathrm{N}\left(\frac{1}{2}, \frac{\delta^T\Sigma_1\delta}{\delta^T\delta}\right)$, $f_1(X_2)$ ∼$\mathrm{N}\left(-\frac{1}{2}, \frac{\delta^T\Sigma_2\delta}{\delta^T\delta}\right)$, $f_2(X_1)$ ∼$\mathrm{N}(0, \mathbf{P}\Sigma_1\mathbf{P})$, $f_2(X_2)$ ∼$\mathrm{N}(0, \mathbf{P}\Sigma_2\mathbf{P})$, $\mathbf{P} = \mathbf{I}_d - \frac{\delta\delta^T}{\delta^T\delta}$. We claim that
$$\frac{1}{200}\min\left[1, \max\left[\frac{\delta^T(\Sigma_1 - \Sigma_2)\delta}{2\delta^T\Sigma_1\delta}, \frac{40\sqrt{\delta^T\delta}}{\sqrt{\delta^T\Sigma_1\delta}}\right]\right] \leqslant \mathrm{TV}(f_1(X_1), f_1(X_2)) \leqslant$$
$$\leqslant \frac{3\delta^T(\Sigma_1 - \Sigma_2)\delta}{2\delta^T\Sigma_1\delta} + \frac{\sqrt{\delta^T\delta}}{2\sqrt{\delta^T\Sigma_1\delta}}.$$

Then
$$\frac{1}{100}\min[1, \lambda] \leqslant \mathrm{TV}(f_2(X_1), f_2(X_2)) \leqslant \frac{3}{2}\lambda,$$
where $\lambda = \left(\sum\limits_{j=1}^{d}\lambda_j\right)^{1/2}$ and λ_i are the eigenvalues of $\Sigma_1^{-1}\Sigma_2 - \mathbf{I}_d$.

Here we present only the proof of the upper bound. Let $d = 1$ and $\sigma_2 \geqslant \sigma_1$. Then for $x = \frac{\sigma_2^2}{\sigma_1^2}$ we have $x - 1 - \ln x \leqslant (x - 1)^2$ and, by Pinsker's inequality,
$$\mathrm{TV}\left(\mathrm{N}(\mu_1, \sigma_1^2), \mathrm{N}(\mu_2, \sigma_2^2)\right) \leqslant \frac{1}{2}\sqrt{\frac{\sigma_2^2}{\sigma_1^2} - 1 - \ln\frac{\sigma_2^2}{\sigma_1^2} + \frac{\Delta^2}{\sigma_1^2}} \leqslant$$

$$\leqslant \frac{1}{2}\sqrt{\frac{\sigma_2^2}{\sigma_1^2} - 1 - \ln\frac{\sigma_2^2}{\sigma_1^2}} + \frac{1}{2}\sqrt{\frac{\Delta^2}{\sigma_1^2}} \leqslant \frac{1}{2}\frac{|\sigma_2^2 - \sigma_1^2|}{\sigma_1^2} + \frac{1}{2}\frac{\Delta}{\sigma_1}.$$

For $d > 1$, by Pinsker's inequality, one gets the upper bound in the case $\mu_1 = \mu_2 = 0$: if $\lambda_i > -\frac{2}{3}\ \forall i$

$$4\mathrm{TV}\left(\mathrm{N}(0,\Sigma_1),\mathrm{N}(0,\Sigma_2)\right)^2 \leqslant \sum_{i=1}^{d}\lambda_i - \ln(1+\lambda_i) \leqslant \sum_{i=1}^{d}\lambda_i^2 = \lambda^2.$$

4. Kolmogorov – Smirnov distance

Kolmogorov – Smirnov distance (only for probability measures on $\mathbf{R}$) is defined by

$$\mathrm{Kolm}(\mathbf{P},\mathbf{Q}) := \sup_{x\in\mathbf{R}}|\mathbf{P}(-\infty,x) - \mathbf{Q}(-\infty,x)|.$$

We have

$$\mathrm{Kolm}(\mathbf{P},\mathbf{Q}) \leqslant \mathrm{TV}(\mathbf{P},\mathbf{Q}).$$

Suppose $X \sim \mathbf{P}, Y \sim \mathbf{Q}$ are two random variables and Y has a density with respect to a Lebesgue measure bounded by a constant C. Then

$$\mathrm{Kolm}(\mathbf{P},\mathbf{Q}) \leqslant 2\sqrt{C\mathrm{Wass}_1(\mathbf{P},\mathbf{Q})},$$

where $\mathrm{Wass}_1(\mathbf{P},\mathbf{Q}) = \inf[\mathbf{E}|X-Y| : X \sim \mathbf{P}, Y \sim \mathbf{Q}]$.

Let $N(t) \sim \mathrm{Po}(t)$ then, via integration by part,

$$\mathbf{P}(N(t) \leqslant n) = \sum_{k=0}^{n} e^{-t}\frac{t^k}{k!} = \int_t^\infty e^{-u}\frac{u^n}{n!}\mathrm{d}u = \int_t^\infty \mathbf{P}(N(u) = n)\mathrm{d}u.$$

Hence,

$$\mathrm{Kolm}(X_1,X_2) = \tau(X_1,X_2) = \mathbf{P}(X_2 \geqslant l) - \mathbf{P}(X_1 \geqslant l) =$$
$$= \mathbf{P}(X_1 \leqslant l-1) - \mathbf{P}(X_2 \leqslant l-1) = \int_{\lambda_1}^{\lambda_2}\mathbf{P}(N(u) = l-1)\mathrm{d}u,$$

where $l = \min[k \in \mathbf{Z}_+ : f(k) \geqslant 1]$ and $f(k) = \frac{\mathbf{P}(N(\lambda_2)=k)}{\mathbf{P}(N(\lambda_1)=k)}$.

Conclusion

This short review discusses only the most popular and well-known inequalities. Another interesting cases, i.e. the total variation distance between Binomial distribution and Gaussian with equal parameters, deserve special attention. Also, applications of these bounds in different problems of mathematical statistics, including classification theory and machine learning algorithms, are a rich field in the state of extensive development.

References

1. Suhov Yu., Kelbert M. *Probability and Statistics by Example. Vol. I. Basic Probability and Statistics*. 2nd ed. Cambridge, UK, Cambridge University Press, 2014. 470 p. https://doi.org/10.1017/CBO9781139087773
2. Pinsker M. *Information and Information Stability of Random Variables and Processes*. San Francisco, USA, Holden-Day Inc., 1964. 243 p.

3. Le Cam L. *Asymptotic Methods in Statistical Decision Theory*. Springer Series in Statistics. New York, NY, Springer, 1986. 742 p. https://doi.org/10.1007/978-1-4612-4946-7
4. Le Cam L. An approximation theorem for the Poisson binomial distribution. *Pacific Journal of Mathematics*, 1960, vol. 10, no. 4, pp. 1181–1197. https://doi.org/10.2140/pjm.1960.10.1181
5. Devroye L., Mehrabian A., Reddad T. The total variation distance between high-dimensional Gaussians. *ArXiv*, 2020, ArXiv:1810.08693v5, pp. 1–12.

Поступила в редакцию / Received 25.11.2021
Принята к публикации / Accepted 27.12.2021
Опубликована / Published 31.05.2022