Necessary and Sufficient Conditions of ε-Separability

Maria A. Ivanchuk[1], Igor V. Malyk[2]

[1]Department of Biological Physics and Medical Informatics, Bucovinian State Medical University, Chernivtsi, Ukraine

[2]Department of the System and Analysis and Insurance Financial Mathematics, Yuriy Fedkovych Chernivtsi National University, Chernivtsi, Ukraine

Email address:

mgracia@ukr.net (M. A. Ivanchuk), malyk.igor.v@gmail.com (I. V. Malyk)

Abstract: In this paper we propose a new approach for solving the classification problem, which is based on the using ε-nets theory. It is shown that for ε-separating of two sets one can use their ε-nets in the range space w.r.t. halfspaces, which considerably reduce the complexity of the separating algorithm for large sets' sizes. The separation space which contains the possible values of ε for ε-nets of both sets is considered. The separation space is quasi-convex in general case. To check necessary and sufficient conditions of ε-separability of two sets one can solve an optimisation problem, using the separation space as constraints. The lower bound of the separation space is convex for the exponential distribution and linear for the uniform distribution. So, we have convex and linear optimisation problems in these cases.

1. Introduction

In 1987 D. Haussler and E. Welzl [1] introduced ε-nets. Since that time ε-nets are used in computational and combinatorics geometry.

Let X be a set (possibly infinite) and $\mathcal{R} \subseteq 2^X$. The pair $(X, \mathcal{R})$ is called a range space with X its points and the elements of $\mathcal{R}$ its ranges. Typical examples of X range spaces in geometric context are set of halfspaces in R^d, set of balls in R^d, set of convex hulls in R^d, axis parallel rectangles in the plane [2], set of d-dimensional simplexes in R^d [3], triangles in the plane [4], set of α-fat wedges in the plain [5].

Numerous works study ε-nets of one set [2-6]. B. Aronov et al. shown the existence of ε-nets of size $O\left(\frac{1}{\varepsilon} \log \log \frac{1}{\varepsilon}\right)$ for planar point sets and axis-parallel rectangular ranges [2]. J. Kulkarni et al. get an ε-net of size $O\left(\frac{\pi}{\alpha\varepsilon}\right)$ for α-fat wedges in R^2 [5]. B .Gärtner proved the existing of ε-nets of size $\left\lceil \frac{d \ln n}{\varepsilon} \right\rceil$, where d is Vapnik-Chervonenkis dimension

[3]. Matousek J. et al. shown that disks and pseudo-disks in the plane as well as halfspaces in R^3 allow ε-nets of size only $O\left(\frac{1}{\varepsilon}\right)$, which is best possible up to a multiplicative constant [6]. But, unfortunately, the analogous questions for higher-dimensional spaces remain open.

In this paper we will build ε-nets of two sets for solving the classification problem.

Consider two sets A and B in Euclidean space R^d. Consider the set A contains n_A points, set B contains n_B points. Let's denote by $conv_A$ the convex hull of the set A. Let $conv_A \cap conv_B \neq \varnothing$.

Definition 1. Sets A and B are called *ε-separable* if there exist subsets $A_1 \subset A$, $B_1 \subset B$, such that

$$conv(A \setminus A_1) \bigcap conv(B \setminus B_1) = \varnothing \qquad (1)$$

and

$$n_{A_1} + n_{B_1} \leq \varepsilon(n_A + n_B). \qquad (2)$$

The classification problem consists of finding the hyperplane L, which separates R^d into two halfspaces L_ε^+

and L_ε^-, such that $\dfrac{n_{A\cap L_\varepsilon^+}+n_{B\cap L_\varepsilon^-}}{n_A+n_B}\geq 1-\varepsilon$.

Definition 2. Hyperplane L_ε is called *ε-separating* for the sets A and B if

$$\frac{n_{A\cap L_\varepsilon^+}+n_{B\cap L_\varepsilon^-}}{n_A+n_B}\geq 1-\varepsilon.$$

Consider an infinite range space (R^d, H^d), where H^d is the closed halfspaces in R^d bounded by hyperplanes. In [7] the following theorem was proved.

Theorem 1. A necessary and sufficient condition that two sets of points A and B are ε-separable is there exist $\varepsilon_A, \varepsilon_B$ and corresponding ε-nets $N_{\varepsilon_A}^A$, $N_{\varepsilon_B}^B$ in (R^d, H^d) such that

$$\varepsilon_A n_A + \varepsilon_B n_B < \varepsilon(n_A + n_B) \qquad (3)$$

and

$$conv N_{\varepsilon_A}^A \bigcap conv N_{\varepsilon_B}^B = \varnothing. \qquad (4)$$

To find $\varepsilon_A, \varepsilon_B$, which satisfy the inequality (3), let's denote the separation space.

Definition 3. The set

$$D_{A,B}=\left\{(\varepsilon_1,\varepsilon_2)\in(0,1)^2:\exists N_A^{\varepsilon_1},N_B^{\varepsilon_2},conv N_A^{\varepsilon_1}\cap conv N_B^{\varepsilon_2}=\varnothing\right\} \quad (5)$$

is called the separation space for A, B.

Let sets A and B are generated by the general populations ξ and η with distributions F_ξ and F_η.

Definition 4. The set

$$D_{\xi,\eta}:=\left\{(x,y)\in(0,1)^2:\exists L\in R^d, P\{\xi\in L^+\}\leq x, P\{\eta\in L^-\}\leq y\right\} \quad (6)$$

is called the separation space for ξ, η.

Consider the Euclidian space R^1. In [8] the following lemma was proved.

Lemma 1. Let the inverse function F_ξ exist. Then the sets $D_{\xi,\eta}$ and $\overline{D}_{\xi,\eta}:=(0,1)^2\setminus D_{\xi,\eta}$ are separated by the curve

$$y(x)=\min\left(F_\eta\left(F_\xi^{-1}(1-x)\right),1-F_\eta\left(F_\xi^{-1}(x)\right)\right). \qquad (7)$$

Let's consider the general case, when distribution functions don't have the inverse functions in some points. We will use the generalized inverse [9].

Definition 5. For an increasing function $T:R\to R$ with $T(-\infty)=\lim_{x\downarrow-\infty}T(x)$ and $T(\infty)=\lim_{x\uparrow\infty}T(x)$, the *generalized inverse*

$$T^-(y)=\inf\{x\in R:T(x)\geq y\}, y\in R \qquad (8)$$

with the convention that $\inf\varnothing=\infty$. If $T:R\to[0,1]$ is a

distribution function, $T^-:[0,1]\to\overline{R}$ is also called the *quantile function* of T.

In [8] the following lemma was proved.

Lemma 2. Sets $D_{\xi,\eta}$ and $\overline{D}_{\xi,\eta}:=(0,1)^2\setminus D_{\xi,\eta}$ are separated by the line

$$y(x)=\min\left(F_\eta\left(\left(F_\xi\right)^-(1-x)\right),1-F_\eta\left(\left(F_\xi\right)^-(x)\right)\right). \qquad (9)$$

For the line $y_{A,B}(x)$, which separates the sets $D_{A,B}$ and $\overline{D_{A,B}}$ the following theorem was proved in [8].

Theorem 2. Let the following conditions
(1) The sets A, B of size n_A, n_B are generated by the independent continuous random variables ξ, η.
(2) The sets $D_{A,B}$ and $\overline{D_{A,B}}$ are separated by the line $y_{A,B}(x)$.
are satisfied. Then there exist the following equality

$$\lim_{n_A, n_B\to\infty} y_{A,B}(x)=y(x),$$

where

$$y(x)=\min\left(F_\eta\left(\left(F_\xi\right)^{-1}(1-x)\right),1-F_\eta\left(\left(F_\xi\right)^{-1}(x)\right)\right).$$

In this paper we will consider separation curve $y(x)$ for the sets $D_{\xi,\eta}$, $\overline{D_{\xi,\eta}}$ and it's properties.

2. Results and Discussion

2.1. Quasiconvexity of the Separation Curve

Consider the function $f=y(x)$ where $y(x)$ is separation curve for the sets $D_{\xi,\eta}$, $\overline{D_{\xi,\eta}}$. Let's show that in the general case function f is quasiconvex [10].

Definition 6. Let X be the convex subset of R^d. Function $f:X\to R$ is called *quasiconvex*, if for any $x,y\in X$ and $\lambda\in[0,1]$

$$f(\lambda x+(1-\lambda)y)\leq \max(f(x),f(y)).$$

Theorem 3. Function

$$y(x)=\min\left(F_\eta\left(\left(F_\xi\right)_G^{-1}(1-x)\right),1-F_\eta\left(\left(F_\xi\right)_G^{-1}(x)\right)\right)$$

is quasiconvex.

Proof. A continuous function $f:X\to R$ is quasiconvex if and only if at least one of the following conditions holds: f is nondecreasing; f is nonincreasing; there is a point $c\in X$ such that for $t\in X, t\leq c$, f is nonincreasing, and for $t\in X, t>c$, f is nondecreasing [10]. Let's show that

function $y(x)$ is nonincreasing.

(1) Consider function $f_1(x) = F_\eta\left(\left(F_\xi\right)_G^{-1}(1-x)\right)$. Function $F_\xi(x)$ is nondecreasing, so, $\left(F_\xi\right)_G^{-1}(x)$ is also nondecreasing. Then $\left(F_\xi\right)_G^{-1}(1-x)$ is nonincreasing. Since $F_\eta(x)$ is nondecreasing, $f_1(x) = F_\eta\left(\left(F_\xi\right)_G^{-1}(1-x)\right)$ is nonincreasing.

(2) Consider function $f_2(x) = 1 - F_\eta\left(\left(F_\xi\right)_G^{-1}(x)\right)$. Function $\left(F_\xi\right)_G^{-1}(x)$ is nondecreasing, so $F_\eta\left(\left(F_\xi\right)_G^{-1}(x)\right)$ is also nondecreasing. Then $f_2(x) = 1 - F_\eta\left(\left(F_\xi\right)_G^{-1}(x)\right)$ is nonincreasing.

Since function $f_1(x)$ and $f_2(x)$ are nonincreasing, $y = \min\{f_1(x), f_2(x)\}$ is nonincreasing. Thus, Theorem 3 is proved.

In particular cases function $y(x)$ may be convex or linear. Consider corresponding examples.

2.2. Exponential Distribution

Let ξ, η two random variables which is exponentially distributed with parameters μ_ξ, μ_η. Then according to Lemma 1 we have

$$y(x) = \min\left(1 - x^{\frac{\mu_\xi}{\mu_\eta}}, (1-x)^{\frac{\mu_\xi}{\mu_\eta}}\right), x \in (0,1).$$

Lemma 3. Let two random variables ξ, η exponentially distributed with parameters μ_ξ, μ_η, then function $y(x)$, which separates sets $D_{\xi,\eta}$ and $\overline{D_{\xi,\eta}}$ is convex.

Proof. Let's consider the following two possible cases.

(1) Let $\mu_\xi > \mu_\eta$, then

$$y(x) = F_\eta\left(F_\xi^{-1}(1-x)\right) = (1-x)^{\frac{\mu_\xi}{\mu_\eta}}, x \in (0,1)$$

We need to show that the function $y(x)$ is convex. Since

$$y''(x) = \frac{\mu_\xi}{\mu_\eta}\left(\frac{\mu_\xi}{\mu_\eta} - 1\right)(1-x)^{\frac{\mu_\xi}{\mu_\eta}-2} > 0 \text{ as } x \in (0,1).$$

So, function $y(x)$ is convex as $\mu_\xi > \mu_\eta$.

(2) Let $\mu_\xi < \mu_\eta$, then

$$y(x) = 1 - F_\eta\left(F_\xi^{-1}(x)\right) = 1 - x^{\frac{\mu_\xi}{\mu_\eta}}, x \in (0,1)$$

Also, need to show that the function $y(x)$ is convex. Since

$$y''(x) = -\frac{\mu_\xi}{\mu_\eta}\left(\frac{\mu_\xi}{\mu_\eta} - 1\right)x^{\frac{\mu_\xi}{\mu_\eta}-2} > 0 \text{ as } x \in (0,1).$$

So, function $y(x)$ is convex as $\mu_\xi < \mu_\eta$. Thus, Lemma 3 is proved.

Let $\mu_\xi = 1$ and $\mu_\eta = 5$. Function $y(x)$ is illustrated in the figure 1.
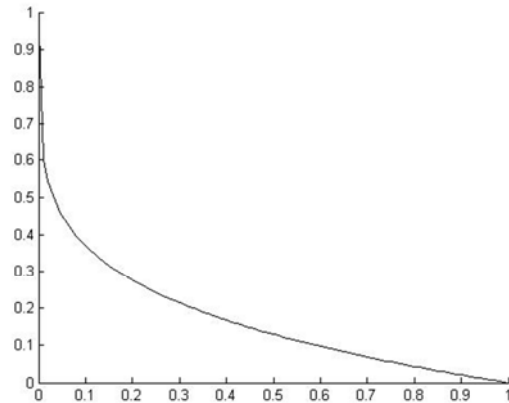

Fig. 1. *Function $y(x)$ for exponential distribution.*

2.3. Uniform Distribution

Let ξ, η two random variables which is uniformly distributed with parameters a_ξ, b_ξ and a_η, b_η. According to the Lemma 1, the set $D_{\xi,\eta}$ is lower bounded by the function

$$y(x) = \min\left(\frac{a_\xi - b_\xi}{b_\eta - a_\eta}x + \frac{b_\xi - a_\eta}{b_\eta - a_\eta}, \frac{a_\xi - b_\xi}{b_\eta - a_\eta}x + \frac{a_\eta - a_\xi}{b_\eta - a_\eta}\right), x \in (0,1).$$

Lemma 4. Let two random variables ξ, η uniformly distributed with parameters a_ξ, b_ξ and a_η, b_η, then function $y(x)$, which separates sets $D_{\xi,\eta}$ and $\overline{D_{\xi,\eta}}$ is linear decreasing function.

Proof. For this reason consider the following two possible cases:

(1) Let for $h \in (-\infty, \infty)$: $F_\xi(h) > F_\eta(h)$, then

$$y(x) = F_\eta\left(F_\xi^{-1}(1-x)\right) = \frac{a_\xi - b_\xi}{b_\eta - a_\eta}x + \frac{b_\xi - a_\eta}{b_\eta - a_\eta};$$

So, $y(x)$ is linear function. Since $a_\xi - b_\xi < 0$, function $y(x)$ is decreasing.

(2) Let for $h \in (-\infty, \infty)$: $F_\xi(h) < F_\eta(h)$, then

$$y(x) = 1 - F_\eta\left(F_\xi^{-1}(x)\right) = \frac{a_\xi - b_\xi}{b_\eta - a_\eta}x + \frac{a_\eta - a_\xi}{b_\eta - a_\eta}$$

So, $y(x)$ is linear function. Since $a_\xi - b_\xi < 0$, function $y(x)$ is decreasing. Thus, Lemma 4 is proved.

Let $a_\xi = 1, b_\xi = 5$ and $a_\eta = 4, b_\eta = 7$. Function $y(x)$ is illustrated in the figure 2.
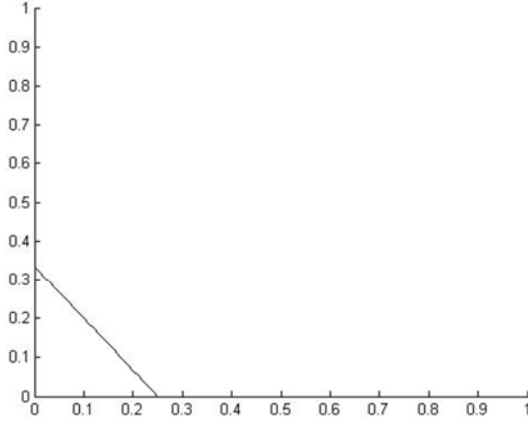

Fig. 2. Function $y(x)$ for uniform distribution.

2.4. Normal Distribution

Let ξ, η two random variables which is normally distributed with parameters μ_ξ, σ_ξ and μ_η, σ_η. According to lemma 1 the set $D_{\xi,\eta}$ is lower bounded by the function

$$y(x) = \min(N_1, N_2), x \in (0,1),$$

where

$$N_1 = \Phi\left(\frac{m_\xi + \sigma_\xi \Phi^{-1}(1-x) - \mu_\eta}{\sigma_\eta}\right)$$

$$N_2 = 1 - \Phi\left(\frac{m_\xi + \sigma_\xi \Phi^{-1}(x) - \mu_\eta}{\sigma_\eta}\right).$$

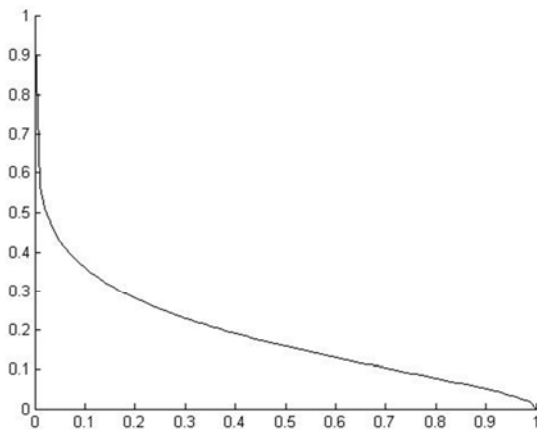Let $\mu_\xi = 3, \sigma_\xi = 1$ and $\mu_\eta = 5, \sigma_\eta = 2$. Function $y(x)$ is illustrated in the figure 3.

Fig.3. Function $y(x)$ for normal distribution.

2.5. Checking the Necessary and Sufficient Conditions for ε-Separability

Obviously, for any $\varepsilon_A, \varepsilon_B \in D_{A,B}$ the equality (4) holds. So, to check the ε-separability of the sets A and B it's enough to check the existing such $(\varepsilon_A, \varepsilon_B) \in D_{A,B}$ that the inequality (3) holds. Consider the optimisation problem

$$\frac{\varepsilon_A n_A + \varepsilon_B n_B}{n_A + n_B} \to \min \qquad (10)$$

$$(\varepsilon_A, \varepsilon_B) \in D_{A,B} \qquad (11)$$

Let $\left(\varepsilon_A^{\min}, \varepsilon_B^{\min}\right)$ be the solution of the task (10), (11). If the equality (3) doesn't hold for $\left(\varepsilon_A^{\min}, \varepsilon_B^{\min}\right)$, it doesn't hold for all $(\varepsilon_A, \varepsilon_B) \in D_{A,B}$, it means the sets A and B are not ε-separable. According to the Theorem 2, the condition (10) can be changed into

$$(\varepsilon_A, \varepsilon_B) \in D_{\xi,\eta}. \qquad (12)$$

According to the Theorem 3, in the general case, the set $D_{\xi,\eta}$ is quasiconvex. According to the Lemma 4, the set $D_{\xi,\eta}$ is convex if the distribution is exponential. In this case the problem (10), (12) is the problem of convex programming. In case of uniform distribution, according to the Lemma 5, the task (10), (12) is the task of linear programming.

Let the point (x_0, y_0) is the solution of the optimisation problem (10), (12).

In the paper [8] the following theorem was proved.

Theorem 4. Let

(1) $\xi_i, i \geq 1$ have distribution F_ξ, $\eta_j, j \geq 1$ have distribution F_η,

(2) $\xi_i, \eta_j, i, j \geq 1$ be independent random variables,

(3) Functions F_ξ, F_η have inverse and bounded densities f_ξ, f_η exist.

Then we obtain the weak convergence

$$\zeta_{n,m} = \zeta_{n,m}(y) = \sqrt{n}\left(F_n\left(F_m^{-1}(y)\right) - F_\eta\left(F_\xi^{-1}(y)\right)\right) \to N(0, \sigma^2),$$

where

$$\sigma^2 = F_\eta\left(F_\xi^{-1}(y)\right)\left(1 - F_\eta\left(F_\xi^{-1}(y)\right)\right), \text{ as } n \to \infty,$$

$$m = O\left(n^\alpha\right), \alpha > 0$$

According to the theorem 4, the point $\left(\varepsilon_A^{\min}, \varepsilon_B^{\min}\right)$ belongs to the neighbourhood

$$\left(\left[x_0 - \frac{z_p\sigma}{\sqrt{n_A}}, x_0 + \frac{z_p\sigma}{\sqrt{n_A}}\right], \left[y_0 - \frac{z_p\sigma}{\sqrt{n_B}}, y_0 + \frac{z_p\sigma}{\sqrt{n_B}}\right]\right),$$

where $\sigma = \sqrt{F_\eta\left(F_\xi^{-1}(y)\right)\left(1 - F_\eta\left(F_\xi^{-1}(y)\right)\right)}$, p is probability, z_p is t-distribution quantile for the corresponding probabilities in p (fig.4).

If in the boundary point $\left(x_0 - \frac{z_p\sigma}{\sqrt{n_A}}, y_0 - \frac{z_p\sigma}{\sqrt{n_B}}\right)$ the condition (4) of the theorem 1 doesn't hold, namely:

$$\frac{\left(x_0 - \frac{z_p\sigma}{\sqrt{n_A}}\right)n_A + \left(y_0 - \frac{z_p\sigma}{\sqrt{n_B}}\right)n_B}{n_A + n_B} > \varepsilon,$$
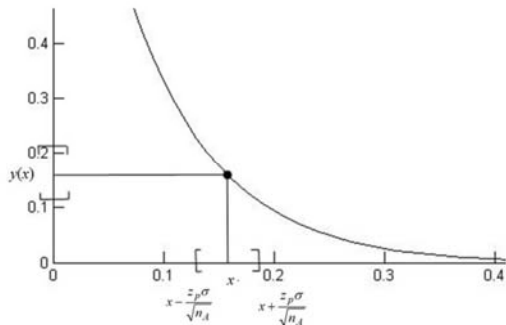
then the sets A and B are not ε-separable.

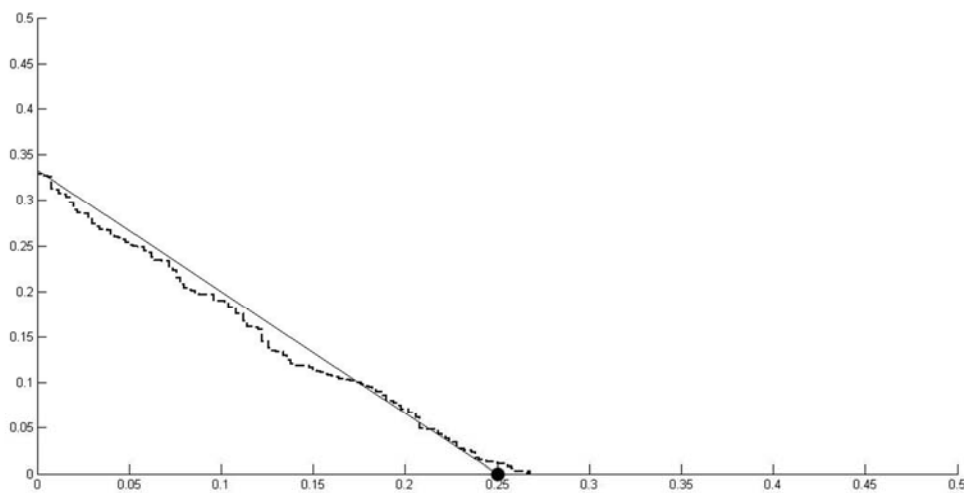**Fig. 4.** *The solution of the optimisation problem.*

Fig.5. *Lower bounds of the sets $D_{\xi,\eta}$ and $D_{A,B}$ and the linear programing problem's solution the point $(0,25;0)$*

Let's verify the condition (3) of the theorem 1

$$\frac{\varepsilon_A^{\min} n_A + \varepsilon_B^{\min} n_B}{n_A + n_B} = 0,0833 > 0,05.$$

Example 1. Consider two sets of points A and B generated by the uniformly distributed general populations ξ and η. Let $n_A = 500, n_B = 1000$, $a_\xi = 1$, $b_\xi = 5$, $a_\eta = 4$, $b_\eta = 7$. Let's verify the ε-separability of the sets A and B as $\varepsilon = 0,05$.

Build the set $D_{\xi,\eta}$ (fig.5) and consider the task of linear programming

$$\varepsilon_A + 2\varepsilon_B \to \min \tag{13}$$

$$-\frac{4}{3}\varepsilon_A - \varepsilon_B \leq \frac{1}{3} \tag{14}$$

$$0 \leq \varepsilon_A \leq 1, 0 \leq \varepsilon_B \leq 1. \tag{15}$$

The solution of the linear programming problem (13)-(15) is the point $(0,25;0)$ (fig.5). Let's find the confidential interval for the solution of the problem

$$\varepsilon_A + 2\varepsilon_B \to \min \tag{16}$$

$$(\varepsilon_A, \varepsilon_B) \in D_{A,B} \tag{17}$$

with probability $p = 0,95$.
According to the theorem 4

$$\sigma^2 = F_\eta\left(F_\xi^{-1}(y)\right)\left(1 - F_\eta\left(F_\xi^{-1}(y)\right)\right) = 0 \text{ as } y = 0,$$

so, the confidential interval for the problem (16), (17) solution degenerates to the point $(0,25;0)$.

The condition (3) doesn't hold for $\left(\varepsilon_A^{\min}, \varepsilon_B^{\min}\right)$, hence it doesn't hold for all $(\varepsilon_A, \varepsilon_B) \in D_{\xi,\eta}$. So, the sets A and B are not ε-separable as $\varepsilon = 0,05$.

Example 2. Consider two sets of points A and B generated by the exponentialy distributed general populations

ξ and η. Let $\mu_\xi = 1$, $\mu_\eta = 5$, $n_A = 500$, $n_B = 500$. Let's verify the ε-separability of the sets A and B as $\varepsilon = 0,05$.

Let's build the set $D_{\xi,\eta}$ (fig.6). Consider the problem of convex programming

$$\varepsilon_A + \varepsilon_B \to \min \tag{18}$$

$$-\varepsilon_A^{1/5} - \varepsilon_B \le -1 \tag{19}$$

$$0 \le \varepsilon_A \le 1, 0 \le \varepsilon_B \le 1. \tag{20}$$

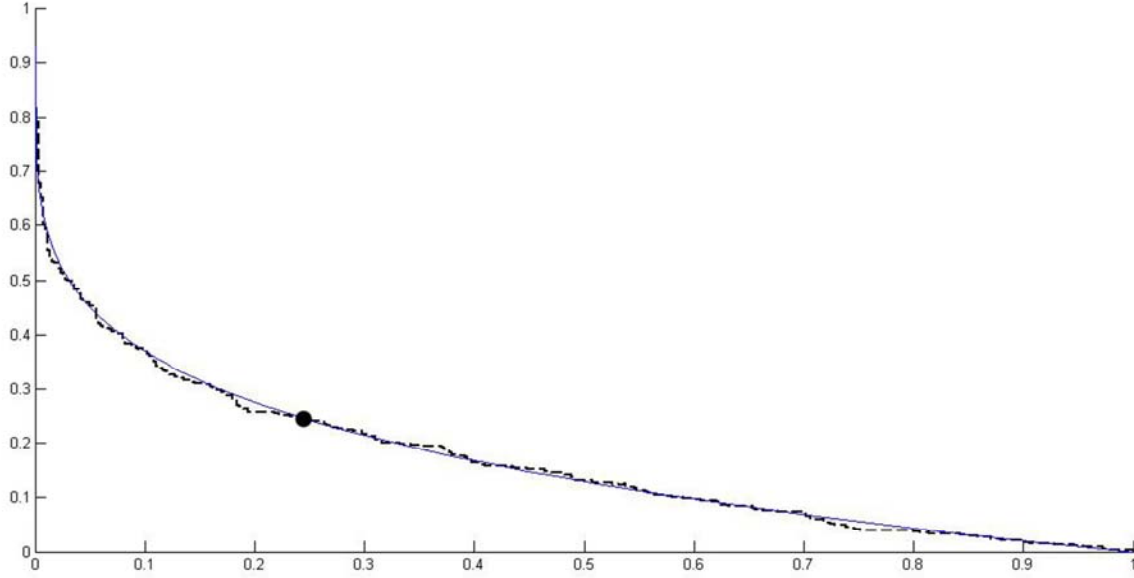The solution of the convex programming problem (18)-(20) is the point $(0,245; 0,245)$ (fig.6).

Fig.6. *Lower bounds of the sets $D_{\xi,\eta}$ and $D_{A,B}$ and the convex programing problem's solution the point* $(0,245; 0,245)$

Let's find the confidence interval for problem's solution with probability $p = 0,95$

$$\varepsilon_A + \varepsilon_B \to \min \tag{21}$$

$$(\varepsilon_A, \varepsilon_B) \in D_{A,B} \tag{22}$$

According to theorem 4

$$\sigma^2 = F_\eta\left(F_\xi^{-1}(y)\right)\left(1 - F_\eta\left(F_\xi^{-1}(y)\right)\right) = 0,0517 \text{ as } y = 0,245,$$

then

$$\frac{z_{0,95}\sigma}{\sqrt{n_A}} = 0,02.$$

So, the confidence interval for the problem's (18)-(20) solution is $([0,225; 0,265], [0,225; 0,265])$.

Let's verify the condition (3) in the boundary point $(0,225; 0,225)$

$$\frac{\varepsilon_A^{\min} n_A + \varepsilon_B^{\min} n_B}{n_A + n_B} = 0,225 > 0,05.$$

The condition (3) doesn't hold for $\left(\varepsilon_A^{\min}, \varepsilon_B^{\min}\right)$, hence it

doesn't hold for all $(\varepsilon_A, \varepsilon_B) \in D_{\xi,\eta}$. So, the sets A and B are not ε-separable as $\varepsilon = 0,05$.

Example 3. Consider two sets of points A and B generated by the general populations ξ, η. Let random variable ξ be uniformly distributed with the parameters $a_\xi = 1$, $b_\xi = 5$ and random variable η be normally distributed with the parameters $\mu_\eta = 7, \sigma_\eta = 1$. Let $n_A = 1000$, $n_B = 5000$. Let's verify the ε-separability of the sets A and B as $\varepsilon = 0,05$. Consider the problem of convex programming

$$\varepsilon_A + 5\varepsilon_B \ge \min \tag{23}$$

$$\varepsilon_B \ge \Phi\left(\frac{b_\xi(1 - \varepsilon_A) + a_\xi\varepsilon_A - \mu_\eta}{\sigma_\eta}\right) \tag{24}$$

$$0 \le \varepsilon_A \le 1, 0 \le \varepsilon_B \le 1 \tag{25}$$

The convex programming problem's solution $(0,0044; 0,022)$ is illustrated in the figure 7.
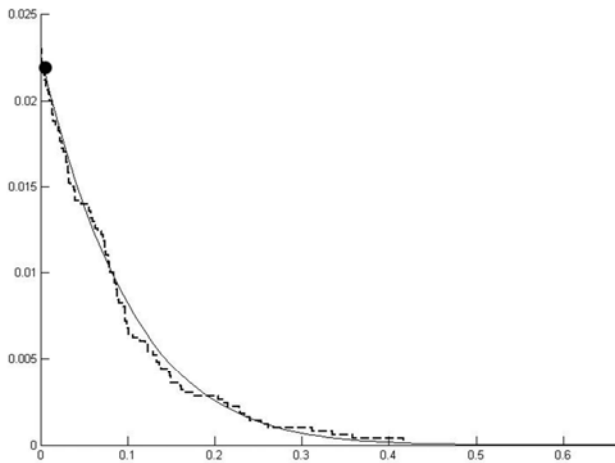
Fig. 7. *Lower bounds of the sets $D_{\xi,\eta}$ and $D_{A,B}$ and the convex programing problem's solution the point $(0,0044;0,022)$*

Let's find the confidence interval for problem's solution with probability $p = 0,95$

$$\varepsilon_A + 5\varepsilon_B \geq \min \qquad (26)$$

$$(\varepsilon_A, \varepsilon_B) \in D_{A,B} \qquad (27)$$

$$\sigma^2 = F_\eta\left(F_\xi^{-1}(y)\right)\left(1 - F_\eta\left(F_\xi^{-1}(y)\right)\right) \to 0 \quad \text{as} \quad y = 0,022 \quad,$$

hence the confidential interval degenerates to the point $(0,0044;0,022)$. Let's verify the condition (3)

$$\frac{\varepsilon_A^{\min} n_A + \varepsilon_B^{\min} n_B}{n_A + n_B} = 0,0191 < 0,05 .$$

The condition holds. So, the sets A and B are ε-separable as $\varepsilon = 0,05$.

3. Conclusions

Two sets A and B are ε-separable iff there exist their ε-nets $N_{\varepsilon_A}^A$, $N_{\varepsilon_B}^B$ in (R^d, H^d) such that conditions (3)-(4) hold. To verify the conditions (3)-(4) it is enough to solve the optimisation problem (10)-(11). If the solution of the problem (10)-(11) does not satisfy the condition (3), then the sets A and B are not ε-separable. According to the theorem 2, one can use the theoretical set $D_{\xi,\eta}$ as constraints for the optimisation problem (10)-(11). Three examples show verification of ε-separability for uniform and exponential distributions.

References

[1] Haussler D. and Welzl E. "Epsilon-nets and simplex range queries", *Discrete Comput. Geom.*, 1987, №2, p.p. 127–151.

[2] Aronov B., Ezra E., Sharir M. "Small-size epsilon-nets for axis-parallel rectangles and boxes", *Symposium on Theory of Computing*, 2009, p.p. 639–648.

[3] Gärtner B., Hoffmann M. *Computational Geometry*, http://www.ti.inf.ethz.ch/ew/lehre/CG12/lecture/CG%20lecture%20notes.pdf.

[4] Hausler S. *VC Dimension. A Tutorial for the Course Computational Intelligence*, http://www.igi.tugraz.at/lehre/CI.

[5] Kulkarni J., Govindarajan S. "New ε-Net Constructions", *Canadian Conference on Computational Geometry* (CCCG), 2010, P.P.159-162.

[6] Matousek J., Seidel R., Welzl E. "How to Net a Lot with Little:Small ε-Nets for Disks and Halfspaces" *In Proc. sixth annual symposium on Computational geometry*, 1990, P.P. 16–22.

[7] Ivanchuk M.A., Malyk I.V. "Solving the Classification Problem Using ε-nets", *Cybernetics and Systems Analysis*, Vol.52, Iss.4 (2016), p.p. 613-622.

[8] M. A. Ivanchuk, I. V. Malyk "Using ε-Nets for Linear Separation of Two Sets in a Euclidean Space R^d ", *Cybernetics and Systems Analysis*, Vol.51, Iss.6, 2015, P.P. 965-968. - DOI 10.1007/s10559-015-9789-7.

[9] Embrechts P., Hofert M. "A note on generalized inverses", *Mathematical Methods of Operations Research* , 2013, 77(3), 423-432.

[10] Boyd S., Vandenberghe L. *Convex Optimization*, Cambridge University Press, 2009, 716 p.