

**DVORETSKIY–KIEFER–VOLFOVITS TENGISIZLIGINING
NAZARIY ASOSLARI VA OPTIMAL O‘ZGARMASLARNI
ANIQLASH MASALASI**

Yuldosheva Durdona

Navoiy davlat universiteti

Matematika mutaxassisligi

1-kurs magistranti

Annotatsiya. UShbu tezisda Dvoretzkiy–Kiefer–Volfovits tengisizligining nazariy asoslari hamda empirik taqsimot funksiyasining haqiqiy taqsimot funksiyasiga bir jinsli yaqinlashuvi masalasi kompleks tarzda tahlil qilinadi. Tadqiqot jarayonida tengisizlikning matematik mohiyati, uning ehtimollar nazariyasidagi o‘rni va konsentratsiya tengisizliklari tizimidagi ahamiyati ilmiy jihatdan asoslab berilgan. Shuningdek, tengisizlikdagi o‘zgarmaslarning optimal qiymatini aniqlash muammosi katta og‘ishlar nazariyasi, empirik jarayonlar va asimptotik tahlil yondashuvlari asosida yoritilgan. Empirik taqsimot funksiyasining maksimal og‘ishi uchun beriladigan eksponensial baholarning keskinligi tahlil qilinib, optimal konstantaning nazariy va amaliy ahamiyati ochib berilgan.

Kalit so‘zlar. Dvoretzkiy–Kiefer–Volfovits tengisizligi, empirik taqsimot funksiyasi, optimal o‘zgarmaslar, empirik jarayonlar, ehtimollar nazariyasi, konsentratsiya tengisizliklari, katta og‘ishlar nazariyasi, bir jinsli yaqinlashuv, asimptotik tahlil, statistik baholash.

Abstract. This thesis comprehensively analyzes the theoretical foundations of the Dvoretzky–Kiefer–Wolfowitz inequality and the issue of homogeneous approximation of the empirical distribution function to the true distribution function. In the course of the research, the mathematical essence of the inequality, its role in probability theory, and its significance in the system

of concentration inequalities are scientifically substantiated. Also, the problem of determining the optimal value of the constants in the inequality is addressed based on the theory of large deviations, empirical processes, and asymptotic analysis approaches. The sharpness of the exponential estimates given for the maximum deviation of the empirical distribution function is analyzed, and the theoretical and practical significance of the optimal constant is revealed.

Keywords. Dvoretzky–Kiefer–Wolfowitz inequality, empirical distribution function, optimal constants, empirical processes, probability theory, concentration inequalities, large deviations theory, homogeneous approximation, asymptotic analysis, statistical evaluation.

Ehtimollar nazariyasi va matematik statistikaning zamonaviy rivojlanish bosqichida empirik jarayonlarni o‘rganish muhim nazariy va amaliy ahamiyat kasb etmoqda. Xususan, empirik taqsimot funksiyasining haqiqiy taqsimot funksiyasidan og‘ishi masalasini aniqlik bilan baholash statistik xulosalarning ishonchliligi va barqarorligini ta’minlashda asosiy o‘rin tutadi. Shu nuqtai nazardan, Dvoretzkiy–Kiefer–Volfovits tengsizligi empirik taqsimot funksiyasining bir jinsli (uniform) yaqinlashuv tezligini baholovchi fundamental natijalardan biri hisoblanadi.

Mazkur tengsizlik empirik taqsimot funksiyasi bilan nazariy taqsimot funksiyasi orasidagi maksimal og‘ish ehtimolini eksponensial tipdagi baho orqali ifodalaydi. Uning ustun jihati shundaki, baho taqsimot turiga bog‘liq emas, ya’ni u taqsimotdan mustaqil (distribution-free) xarakterga ega. Bu xususiyat tengsizlikni nazariy tadqiqotlarda ham, amaliy statistik masalalarda ham keng qo‘llash imkonini beradi.

Tengsizlikdagi o‘zgarmaslarning optimal qiymatlarini aniqlash esa alohida ilmiy qiziqish uyg‘otadi. Sababi, baholashdagi o‘zgarmaslar aniqligi statistik intervallar kengligi, testlar kuchi va ehtimollik baholarining ishonchliligiga bevosita ta’sir ko‘rsatadi. Optimal o‘zgarmaslarni aniqlash

masalasi nafaqat tengisizlikning kuchaytirilgan shakllarini ishlab chiqish, balki empirik jarayonlar nazariyasini yanada chuqurlashtirishga ham xizmat qiladi.

Shuningdek, ushbu masala katta og‘ishlar nazariyasi, konsentratsiya tengisizliklari hamda zamonaviy ehtimollik metodlari bilan uzviy bog‘liqdir. So‘nggi yillarda o‘zgarmaslarni aniqlashtirish va baholarni keskinlashtirishga qaratilgan tadqiqotlar mazkur yo‘nalishning dolzarbligini yana bir bor tasdiqlamoqda.

Empirik taqsimot funksiyasining nazariy taqsimot funksiyasidan og‘ishini baholash masalasi XX asrning o‘rtalaridan boshlab ehtimollar nazariyasi va matematik statistikaning markaziy yo‘nalishlaridan biriga aylandi. Bu boradagi eng muhim natijalardan biri 1956-yilda Aryeh Dvoretzky, Jack Kiefer va Jacob Wolfowitz tomonidan isbotlangan tengisizlik hisoblanadi. Ular empirik taqsimot funksiyasi $F_n(x)$ va haqiqiy taqsimot funksiyasi $F(x)$ orasidagi maksimal og‘ish ehtimoli uchun eksponensial tipdagi universal baho olishga muvaffaq bo‘ldilar. Ushbu natija keyinchalik Dvoretzkiy–Kiefer–Volfovits (DKW) tengisizligi nomi bilan ilmiy adabiyotga kirdi.

Dastlabki shaklda tengisizlikdagi o‘zgarmaslar optimal emas edi. Mazkur masalani yanada chuqurlashtirish 1990-yillarda Pascal Massart tomonidan amalga oshirildi. U 1990-yilda tengisizlikdagi eksponensial bahodagi koeffitsientning optimal qiymatini aniqlab, klassik natijani keskinlashtirdi. Massartning ishi DKW tengisizligining “aniq” (sharp) shaklini asoslab berdi va optimal o‘zgarmas 2 ga teng ekanligini ko‘rsatdi. Bu natija empirik jarayonlar nazariyasida muhim burilish yasadi [2].

Empirik jarayonlar nazariyasining umumiy asoslari Patrick Billingsley, Michel Ledoux va Michel Talagrand ishlarida keng yoritilgan. Xususan, Talagrand konsentratsiya tengisizliklari nazariyasini rivojlantirib, empirik jarayonlar uchun yanada nozik baholar olish imkonini berdi. Bu yondashuv DKW tengisizligining umumlashtirilgan shakllarini ishlab chiqishda nazariy asos bo‘lib xizmat qildi [2, 5].

Bundan tashqari, Vladimir Vapnik va Alexey Chervonenkis tomonidan ishlab chiqilgan VC-o'lcham nazariyasi empirik jarayonlarning bir jinsli yaqinlashuvini o'rganishda yangi bosqichni boshlab berdi. Ularning natijalari DKW tengisizligining ko'p o'lchamli va funksional sinflarga nisbatan umumlashtirilgan variantlarini shakllantirishga zamin yaratdi [6].

Katta og'ishlar nazariyasi doirasida Amir Dembo va Ofer Zeitouni ishlari muhim ahamiyatga ega. Ularning tadqiqotlari eksponensial tipdagi baholarning asimptotik xususiyatlarini chuqur tahlil qilish imkonini berdi. Bu esa DKW tengisizligidagi o'zgarmaslarning optimal chegaralarini aniqlash masalasini katta og'ishlar printsiplari bilan bog'lashga yordam berdi [7].

Shuningdek, zamonaviy konsentratsiya tengisizliklari yo'nalishida Roman Vershynin va Olivier Boucheron tomonidan olib borilgan tadqiqotlar ehtimollik baholarini aniqlashtirishning yangi usullarini taklif etdi. Ularning ishlari DKW tengisizligini boshqa konsentratsiya tengisizliklari (masalan, Hoeffding yoki Bernstein tipidagi baholar) bilan taqqoslash imkonini beradi [8; 9].

Yuqoridagi tahlillar shuni ko'rsatadiki, Dvoretzkiy–Kiefer–Volfovits tengisizligi dastlab empirik taqsimot funksiyasi uchun umumiy baho sifatida shakllangan bo'lsa, keyinchalik optimal o'zgarmaslarni aniqlash, keskinlashtirish va umumlashtirish yo'nalishlarida keng rivojlangan. Shunga qaramay, tengisizlikdagi o'zgarmaslarning turli sharoitlarda optimal qiymatlarini aniqlash, ayniqsa cheklangan namunalar uchun baholar olish masalasi hanuz dolzarb bo'lib qolmoqda.

Dvoretzkiy–Kiefer–Volfovits tengisizligini tahlil qilish empirik taqsimot funksiyasining nazariy taqsimot funksiyasiga bir jinsli yaqinlashuvi masalasini chuqur o'rganishdan boshlanadi. Faraz qilaylik, $X_1, X_2, \dots, X_n$ — bir xil taqsimlangan va o'zaro mustaqil tasodifiy miqdorlar bo'lib, ularning taqsimot funksiyasi $F(x)$ bilan belgilanadi. Empirik taqsimot funksiyasi esa

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}\{X_i \leq x\}$$

ko‘rinishda aniqlanadi. Tadqiqotning markaziy masalasi $\sup_x |F_n(x) - F(x)|$ kattalikning ehtimollik taqsimotini baholashdan iborat.

Dvoretzkiy–Kiefer–Volfovits tengisizligi quyidagi ko‘rinishda ifodalanadi:

$$\mathbb{P} \left(\sup_x |F_n(x) - F(x)| > \varepsilon \right) \leq 2e^{-2n\varepsilon^2}.$$

Mazkur bahoning muhim jihati shundaki, u taqsimot turiga bog‘liq emas, ya’ni universal xarakterga ega. Bu xossa uni noaniq taqsimotli statistik modellar uchun ham qo‘llash imkonini beradi. Tengisizlikning isbotida kombinatorik usullar, eksponensial momentlar bahosi hamda binom taqsimotning simmetrik xususiyatlaridan foydalaniladi. Empirik jarayonni Bernulli tipidagi yig‘indilar orqali ifodalash va Hoeffding tipidagi baholardan foydalanish asosiy texnik vosita hisoblanadi.

Tengisizlikdagi asosiy e’tibor o‘zgarmaslarning optimal qiymatiga qaratiladi. Dastlabki natijalarda eksponensial bahodagi koeffitsient aniq emas edi, keyinchalik olib borilgan tadqiqotlar natijasida eksponent oldidagi 2 koeffitsienti optimal ekani ko‘rsatildi. Optimal o‘zgarmasni aniqlash jarayoni supremum funksionalning aniq ehtimollik bahosini hisoblash bilan chambarchas bog‘liq. Bu yerda empirik jarayonning Kolmogorov statistikasi bilan bog‘lanishi muhim rol o‘ynaydi, chunki $\sqrt{n}(F_n - F)$ jarayonning chegaraviy taqsimoti Braun ko‘prigi bilan tavsiflanadi. Asimptotik tahlil eksponensial baholarning aniq chegaralarini aniqlashga yordam beradi, biroq cheklangan n uchun optimal konstantani topish ancha murakkab vazifadir.

Optimal o‘zgarmaslarni aniqlashda katta og‘ishlar printsipi muhim metodologik asos bo‘lib xizmat qiladi. Supremum funksional uchun eksponensial tezlik funksiyasini aniqlash orqali baholarning keskinligi tekshiriladi. Agar bahoda berilgan koeffitsient kamaytirilsa, tengisizlik

umumiy holda buzilishi mumkinligi ko'rsatiladi. Shu sababli optimal konstantani aniqlash ehtimollik chegaralarining "eng keskin" shaklini topish deganidir.

Shuningdek, tengsizlikning bir tomonlama varianti ham muhim ahamiyatga ega:

$$\mathbb{P} \left(\sup_x (F_n(x) - F(x)) > \varepsilon \right) \leq e^{-2n\varepsilon^2}.$$

Bu ko'rinish amaliy statistikada, xususan ishonchli intervallarni qurishda qo'llanadi. Bir tomonlama va ikki tomonlama baholar orasidagi farq ham o'zgarmlarning optimal qiymatini aniqlash jarayonida e'tiborga olinadi.

Empirik jarayonni funksional fazoda ko'rib chiqish yondashuvi tengsizlikni yanada umumlashtirish imkonini beradi. Funksiyalar sinfi cheklangan VC-o'lchamga ega bo'lsa, bir jinsli yaqinlashuv tezligi saqlanadi, biroq bu holatda eksponensial bahodagi o'zgarmlar funksiyalar sinfining murakkabligiga bog'liq bo'ladi. Demak, klassik Dvoretzkiy–Kiefer–Volfovits tengsizligi empirik jarayonlar nazariyasining maxsus, ammo fundamental holi sifatida namoyon bo'ladi.

Yuqoridagi natijalari shuni tasdiqladiki, Dvoretzkiy–Kiefer–Volfovits tengsizligi taqsimotdan mustaqil universal baho berishi bilan alohida ilmiy qiymatga ega. Tengsizlikdagi eksponensial ko'rsatkich va uning oldidagi o'zgarmlarning optimal qiymati ehtimollik chegaralarining keskinligini belgilaydi. Optimal konstantaning aniqlanishi nazariy jihatdan bahoning eng aniq shaklini beradi, amaliy jihatdan esa statistik ishonchlilik intervallari va gipotezalarini tekshirish jarayonida yuqori aniqlikni ta'minlaydi.

Shuningdek, empirik jarayonning asimptotik xatti-harakati, katta og'ishlar printsipli va konsentratsiya tengsizliklari bilan o'zaro bog'liqlik tahlil qilinib, optimal o'zgarmlarni aniqlash muammosi kengroq nazariy kontekstda qaraldi. Bu esa mazkur tengsizlikni empirik jarayonlar nazariyasining muhim elementi sifatida baholash imkonini berdi.

Umuman olganda, olib borilgan tahlil Dvoretzkiy–Kiefer–Volfovits tengisizligining nazariy mustahkamligini, undagi o‘zgarmaslarning optimal qiymatini aniqlashning dolzarbligini hamda ushbu yo‘nalishning kelgusida chuqurroq tadqiq etilishi zarurligini ko‘rsatdi. Olingan natijalar empirik jarayonlar nazariyasini rivojlantirish, statistik baholash usullarini takomillashtirish va konsentratsiya tengisizliklarining keskin shakllarini ishlab chiqish uchun metodologik asos bo‘lib xizmat qilishi mumkin.

FOYDALANILGAN ADABIYOTLAR

1. Dvoretzky A., Kiefer J., Wolfowitz J. Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator // *Annals of Mathematical Statistics*. – 1956. – Vol. 27, № 3. – P. 642–669.
2. Massart P. The tight constant in the Dvoretzky–Kiefer–Wolfowitz inequality // *Annals of Probability*. – 1990. – Vol. 18, № 3. – P. 1269–1283.
3. Billingsley P. *Convergence of probability measures*. – 2nd ed. – New York: John Wiley & Sons, 1999. – 277 p.
4. Ledoux M. *The concentration of measure phenomenon*. – Providence: American Mathematical Society, 2001. – 181 p.
5. Talagrand M. New concentration inequalities in product spaces // *Inventiones Mathematicae*. – 1996. – Vol. 126, № 3. – P. 505–563.
6. Vapnik V. N., Chervonenkis A. Ya. On the uniform convergence of relative frequencies of events to their probabilities // *Theory of Probability and Its Applications*. – 1971. – Vol. 16, № 2. – P. 264–280.
7. Dembo A., Zeitouni O. *Large deviations techniques and applications*. – 2nd ed. – New York: Springer, 1998. – 396 p.
8. Boucheron O., Lugosi G., Massart P. *Concentration inequalities: A nonasymptotic theory of independence*. – Oxford: Oxford University Press, 2013. – 481 p.

9. Vershynin R. High-dimensional probability: An introduction with applications in data science. – Cambridge: Cambridge University Press, 2018. – 306 p.
10. Shorack G. R., Wellner J. A. Empirical processes with applications to statistics. – New York: John Wiley & Sons, 1986. – 938 p.