

## KO'P O'ZGARUVCHILI STATISTIK MODELLAR VA ULARNING EHTIMOLLIK ASOSLARI

**Ikromjonova Malikabonu Shuxratbek qizi**

Andijon Davlat pedagogika instituti Matematika va informatika 4-bosqich talabasi.

**Xaydarxo'jayev Saydazimjon Nabixo'ja o'g'li**

Andijon Davlat pedagogika instituti Matematika va informatika 4-bosqich talabasi.

<https://doi.org/10.5281/zenodo.20031114>

***Annotatsiya.** Ushbu maqola ko'p o'zgaruvchili statistik modellarni ehtimollik nazariyasining asosiy tushunchalari bilan uzviy bog'lab, ularning nazariy asoslari va amaliy qo'llanish mantiqini tizimli yoritishga qaratilgan. Tadqiqotning maqsadi birgalikdagi taqsimotlar, shartli ehtimolliklar, kovariatsiya tuzilmalari hamda yashirin omillar orqali quriladigan modellar sinfining yagona probabilistik izohini berish va tanlash mezonlarini asoslashdan iborat. Metodologiya sifatida vektor tasodifiy kattaliklar nazariyasi, maksimal ehtimollik va Bayes yondashuvlari, identifikatsiya va regularizatsiya tamoyillari, shuningdek asimptotik yaqinlashuvlar qo'llanildi. Ilmiy yangilik ko'p o'zgaruvchili regressiya, kovariatsiya modellashtirish, aralashmalar va latent tuzilmalarni birlashtiruvchi ehtimollik asosidagi konseptual ramka taklif qilinishida, shuningdek yuqori o'lchamli sharoitda barqaror baholash va model mosligini tekshirish bo'yicha izchil mezonlar bayon etilishida namoyon bo'ladi.*

***Kalit so'zlar.** Ko'p o'zgaruvchili model; birgalikdagi taqsimot; kovariatsiya matritsasi; maksimal ehtimollik; Bayes inferensiyasi; identifikatsiya; regularizatsiya.*

Ko'p o'zgaruvchili statistik modellar zamonaviy fan va amaliyotda bog'liqliklarning murakkab tabiatini tushuntirish, prognozlash va qaror qabul qilishni asoslash uchun tayanch vositaga aylandi. Bitta natija o'zgaruvchisi bilan ishlaydigan klassik yondashuvlar ko'pincha real tizimlarda uchraydigan o'zaro ta'sirlar, yashirin omillar va korrelyatsiyalashgan xatoliklarni yetarli ifodalay olmaydi. Ijtimoiy-iqtisodiy jarayonlar, biotibbiyot, muhandislik diagnostikasi va moliyaviy risk boshqaruvida kuzatuvlar ko'pincha vektor ko'rinishida bo'ladi, ularning komponentlari esa bir-birini shartli ravishda ham, bevosita ham ta'sir ostiga oladi.

Shu sababli ko'p o'zgaruvchili modellashtirishning asosiy muammosi faqat parametrlarni baholash emas, balki ehtimollik tuzilmasini, ya'ni birgalikdagi taqsimotni adekvat tanlash va uni tekshirishdan iborat. Bu nuqtada ehtimollik nazariyasi nafaqat formal asos, balki modelning mazmuniy interpretatsiyasi va cheklovlarini ham belgilovchi metodologik karkas sifatida namoyon bo'ladi.

Adabiyotlarda ko'p o'zgaruvchili regressiya, kovariatsiya strukturalari, faktor tahlili va latent o'zgaruvchili modellar alohida yo'nalishlar sifatida rivojlangan bo'lsa-da, ularni yagona probabilistik til orqali uyg'unlashtirish ko'pincha yetarlicha izchil berilmaydi. Masalan, normal taqsimotga tayangan klassik kovariatsiya tahlili kuchli analitik qulayliklar beradi, biroq og'ir dumli taqsimotlar, aralashmalar yoki nolinearliklar mavjud bo'lganda model mosligi muammosi keskinlashadi. Yana bir tomondan, yuqori o'lchamli sharoitda kovariatsiya matritsasini barqaror baholash, multikollinearlik va identifikatsiya masalalari markaziy o'rin egallaydi; bu esa regularizatsiya va Bayes priorlari orqali ehtimollik asosida yechim topishni talab qiladi.

Zamonaviy yondashuvlarda model tanlashning informatsion mezonlari, kross-valitatsiya va posterior bashorat tekshiruvlari birgalikda qo'llanmoqda, ammo ularning ehtimollik mohiyati va qo'llanish shartlari doimo to'liq anglab yetilmaydi.

Shu kontekstda maqola ko'p o'zgaruvchili statistik modellarni ehtimollik asoslari bilan bir tizimda ko'rib, asosiy tushunchalar orasidagi mantiqiy bog'lanishlarni ochib berishni dolzarb deb biladi.

Tadqiqotning maqsadi ko'p o'zgaruvchili model sinflarini birgalikdagi taqsimot, shartli mustaqillik, kovariatsiya parametrizatsiyasi va latent tuzilmalar nuqtai nazaridan birlashtiruvchi konseptual ramkani ishlab chiqishdan iborat. Vazifalar sifatida quyidagilar ko'zda tutiladi: birinchidan, vektor tasodifiy kattaliklar uchun ehtimollik modeli va likelihood konstruktsiyasini umumiy shaklda bayon qilish; ikkinchidan, kovariatsiya va korrelyatsiya tuzilmasini tanlashning identifikatsion shartlarini ko'rsatish; uchinchidan, maksimal ehtimollik va Bayes inferensiyasi doirasida baholash va noaniqlikni kvantlashning asosiy mexanizmlarini taqqoslash; to'rtinchidan, yuqori o'lchamli va cheklangan tanlanma sharoitida regularizatsiya hamda model tekshiruvining izchil mezonlarini asoslash.

Ilmiy yangilik taklif qilinayotgan ramkaning turli model oilalarini yagona ehtimollik mantiqida tutashtirishi va u orqali amaliy tahlilda uchraydigan barqarorlik hamda moslik muammolariga yo'naltirilgan tavsifiy mezonlarning izchillik bilan bayon qilinishida ifodalanadi.

Amaliy ahamiyat esa ko'p o'zgaruvchili ma'lumotlarda sababiy talqinga da'vo qilmasdan turib, bog'liqliklarni ishonchli baholash, prognoz sifatini nazorat qilish va model xatoliklarini tizimli aniqlash imkoniyatlarining kengayishidir.

Materials and Methods. Tadqiqot nazariy xarakterga ega bo'lib, asosiy material sifatida ehtimollik nazariyasining vektor tasodifiy kattaliklar bo'limi, statistik inferensiyaning likelihood va Bayes paradigmatlari hamda ko'p o'zgaruvchili modellashtirishda qo'llanadigan standart parametrizatsiyalar olindi. Ma'lumotlar bazasi sifatida empirik kuzatuvlar jamlanmasi emas, balki modellar sinfini aniqlovchi umumiy probabilistik konstruktsiyalar, ularning identifikatsiya shartlari va asimptotik xossalari tahlil qilindi. Metodologik yondashuv quyidagi ketma-ketlikka asoslandi: avvalo,  $Y$  vektor javob va  $X$  dizayn matritsasi sharoitida birgalikdagi taqsimot  $p(Y, X)$  yoki shartli taqsimot  $p(Y|X, \theta)$  ko'rinishida model berilishi, keyin esa parametr  $\theta$  uchun likelihood  $L(\theta) = p(y|x, \theta)$  orqali baholash masalasi qo'yilishi. Shundan so'ng kovariatsiya matritsasi  $\Sigma$  ning ijobiy aniqlanganlik sharti, uning strukturaviy cheklovlari va parametrlar fazosining geometriyasi muhokama qilindi. Yashirin o'zgaruvchili holatlarda  $Z$  latent vektor kiritilib,  $p(Y|X) = \int p(Y|X, Z)p(Z) dZ$  parchalanishi orqali aralashmalar va faktor modellarining umumiy ko'rinishi berildi. Baholashda maksimal ehtimollik, kutilmalarni maksimalashtirishga olib keluvchi algoritmik konstruktsiyalar va Bayesda posterior  $p(\theta|y)$  proporsional  $p(y|\theta)p(\theta)$  tamoyili tayanilgan.

Metod tanlovining ilmiy asosi shundaki, ehtimollik modeli bir tomondan kuzatiladigan bog'liqliklarni formal ifodalasa, ikkinchi tomondan noaniqlikni boshqarish va taqqoslanadigan xulosalar olish imkonini beradi. Ayniqsa, yuqori o'lchamli holatlarda  $\Sigma$  ni bevosita baholashning beqarorligi regularizatsiyani zarur qiladi; bu jarayon penalizatsiyalangan likelihood yoki informativ priorlar orqali ehtimollik doirasida asoslanadi. Model tanlash uchun informatsion mezonlarning umumiy ko'rinishi likelihoodning maksimal qiymati va parametrlar soni o'rtasidagi muvozanat sifatida talqin qilindi, Bayesda esa marginal likelihood va posterior bashorat tekshiruvlari orqali moslik baholandi.

Metodologik baholashda ko'p o'zgaruvchili normal modelning nazariy qulayliklari, shuningdek elliptik taqsimotlar va aralashmalar orqali og'ir dumlilikni hisobga olish imkoniyatlari solishtirildi.

Bundan tashqari, kovariatsiya strukturalarini tahlil qilishda spektral parchalanish, Cholesky parametrizatsiyasi va grafli modellar orqali shartli mustaqillikni kodlash usullari umumlashtirildi.

**Natija.** Tadqiqot natijasida ko'p o'zgaruvchili statistik modellarni yagona ehtimollik asosida tasniflash va izohlash imkonini beruvchi konseptual sxema shakllantirildi. Birinchi natija shundan iboratki, ko'p o'zgaruvchili regressiya va kovariatsiya modellashtirish aslida bitta birgalikdagi taqsimot tanlashning ikki ko'rinishi bo'lib, regressiya shartli o'rtacha funksiyani, kovariatsiya esa shartli yoki shartsiz bog'liqlik tuzilmasini belgilaydi. Agar  $Y|X$  normal deb olinsa, shartli o'rtacha  $X$  Beta va kovariatsiya Sigma parametrlarining ajratilishi likelihoodni nisbatan soddalashtiradi; biroq bu ajratilish taqsimot og'ir dumli yoki aralashma bo'lganda buziladi va kovariatsiya bahosi shartli o'rtacha xatolariga sezgir bo'ladi.

Ikkinchi natija shuki, identifikatsiya masalasi ko'p o'zgaruvchili modellashtirishda markaziy o'ringa ega: parametrlar turli kombinatsiyalarda bir xil birgalikdagi taqsimotni hosil qilishi mumkin bo'lgan holatlarda, qo'shimcha cheklovlar yoki normalizatsiyalar talab etiladi.

Faktor modellarida yuklamalarning aylanish noaniqligi, aralashma modellarida komponentlarni almashtirish simmetriyasi va grafli modellarida ekvivalent strukturalar shular jumlasidandir. Uchinchi natija sifatida, yuqori o'lchamli sharoitda kovariatsiya matritsasini oddiy tanlanma kovariatsiyasi bilan baholashning beqarorligi ehtimollik nuqtai nazaridan ham izohlandi: kuzatuvlar soni  $p$  ga yaqinlashganda yoki undan kichik bo'lganda spektral baholar shovqinga juda sezgir bo'lib, likelihood yuzasi yassi va ko'p cho'qqili tus oladi. Regularizatsiya penalizatsiya orqali yoki priorlar orqali kiritilganda esa posterior yoki penalizatsiyalangan likelihood konsentratsiyasi kuchayadi, bu barqaror baholash va prognozni yaxshilaydi. To'rtinchi natija model mosligini tekshirishda ikki turdagi obyektivlashtirilgan ko'rsatkichlarni ajratishdan iborat bo'ldi: birinchisi, fit darajasini ifodalovchi likelihood asosli mezonlar; ikkinchisi, bashorat xatolarini o'lchovchi tekshiruvlar. Ularning ehtimollik talqini shuni ko'rsatdiki, birinchi tur parametrik farazlarga kuchli bog'liq, ikkinchisi esa ma'lum darajada robust bo'lsa-da, ma'lumotni bo'lish sxemalariga sezgir.

Beshinchi natijada latent o'zgaruvchili konstruktsiyalar turli model oilalarini birlashtiruvchi mexanizm ekanligi ko'rsatildi: aralashma modellar latent sinf indikatorlari orqali, faktor tahlili latent omillar orqali, holat fazosi modellar esa vaqt bo'yicha evolyutsiyalanuvchi latent holat orqali ifodalanadi. Barcha holatlarda kuzatiladigan taqsimot latent o'zgaruvchi bo'yicha integrallash natijasida olinadi, bu esa ehtimollik hisobining markaziy amali sifatida paydo bo'ladi. Natijada, EM turidagi algoritmlar va Bayesda MCMC yoki variatsion yaqinlashuvlar bir xil ehtimollik tuzilmasidan kelib chiqadigan hisoblash strategiyalari sifatida ko'rindi.

**Muhokama.** Olingan natijalar mavjud ilmiy yondashuvlar bilan qiyoslanganda, ko'p o'zgaruvchili modellashtirishni faqat alohida usullar majmui sifatida emas, balki birgalikdagi taqsimot tanlash va tekshirish jarayoni sifatida talqin qilishning ustunligini ko'rsatadi. Klassik ko'p o'zgaruvchili tahlilda normal farazning keng qo'llanishi Anderson maktabi an'alarida chuqur asoslangan bo'lib, u identifikatsiya, asimptotika va testlar uchun yopiq shaklli natijalar beradi [3]. Biroq zamonaviy ma'lumotlar ko'pincha og'ir dumlilik, heteroskedastiklik va nolinearlikni namoyon qiladi; bu holatda aralashmalar va elliptik sinflar orqali moslashuvchanlik oshadi, lekin identifikatsiya va hisoblash murakkablashadi. Taklif qilingan ramka aynan shu ziddiyatni ehtimollik geometriyasi orqali tushuntiradi: moslashuvchanlik oshgani sari likelihood ko'p modali bo'lishi, posterior esa priorlarga sezgir bo'lib qolishi mumkin.

Shunday ekan, model tanlash faqat informatsion mezon bilan cheklanmasdan, moslikning diagnostik tekshiruvlari va prognoz barqarorligi bilan birga olib borilishi kerak. Bayes yondashuvi natijalari nuqtai nazaridan, regularizatsiya va identifikatsiya o'rtasidagi bog'liqlik alohida ahamiyatga ega. Priorlar faqat subyektiv qo'shimcha emas, balki yuqori o'lchamli muammolarda parametrlar fazosini tartibga soluvchi ehtimollik cheklovi sifatida ishlaydi; bu Gelman va hammualliflar ta'kidlaganidek, posterior bashorat tekshiruvlari bilan uyg'unlashganda model mosligini amaliy jihatdan kuchaytiradi. Shu bilan birga, Bayes usullarida hisoblash algoritmlarining yaqinlashuvi va tekshiruvchanligi muammosi mavjud bo'lib, noto'g'ri konvergentsiya diagnostikasi noto'g'ri xulosalarga olib kelishi mumkin.

Maksimal ehtimollik yondashuvi esa ko'pincha tezkor va aniq optimallashtirishga tayanadi, ammo kichik tanlanmalarda va murakkab latent tuzilmada baholashning chetga og'ishi hamda standart xatolarni to'g'ri baholash masalalari dolzarblashadi. Mahalliy adabiyotlarda ham ko'p o'zgaruvchili regressiya va kovariatsiya tahlilining amaliy tadbirlari ko'p uchrasa-da, yuqori o'lchamli regularizatsiya va model diagnostikasining ehtimollik talqini yetarlicha tizimlashtirilmaganligi seziladi. Shu jihatdan maqola nazariy birlikni kuchaytirish orqali o'quv va tadqiqot amaliyotiga metodologik aniqlik kiritadi.

Cheklovlar sifatida tadqiqotning empirik validatsiya emas, nazariy sintez xarakteriga egaligi qayd etiladi; ya'ni taqdim etilgan mezonlar umumiy qoida sifatida berilgan bo'lib, aniq sohalarida taqsimot farazlari va ma'lumot yig'ish mexanizmi alohida tekshiruvni talab qiladi.

Kelajakdagi izlanishlar uchun uch yo'nalish istiqbolli ko'rinadi: birinchisi, yarimparametrik va noan'anaviy bog'liqlik tuzilmalarini ko'p o'zgaruvchili model tanlash nazariyasiga integratsiya qilish; ikkinchisi, kauzal talqin bilan mos keluvchi shartli mustaqillik graflarini robust baholash; uchinchisi, hisoblash samaradorligini oshiradigan variatsion va skalyalanadigan Bayes algoritmlarining nazariy kafolatlarini kuchaytirish.

**Xulosa.** Maqolada ko'p o'zgaruvchili statistik modellar ehtimollik nazariyasining birgalikdagi taqsimot, shartli ehtimollik va bog'liqlik tuzilmalari haqidagi asosiy g'oyalari orqali yagona konseptual ramkada yoritildi. Asosiy xulosa shundan iboratki, regressiya, kovariatsiya modellashtirish, aralashmalar va latent o'zgaruvchili konstruktsiyalar turli ko'rinishda namoyon bo'lsa-da, ularning barchasi birgalikdagi taqsimotni aniqlash va uning parametrlarini identifikatsiya qilish muammosiga borib taqaladi.

Yuqori o'lchamli sharoitda baholashning beqarorligi ehtimollik yuzasining geometriyasi bilan bog'lanib, regularizatsiyaning zarurligini asoslaydi; regularizatsiya esa penalizatsiya yoki priorlar orqali barqaror inferensiyani ta'minlaydi. Model tanlash va moslikni tekshirishda likelihood asosli mezonlar bilan bir qatorda bashoratga yo'naltirilgan tekshiruvlar ham zarur ekani ko'rsatildi.

Natijada, ko'p o'zgaruvchili tahlilda ishonchli xulosa chiqarish uchun ehtimollik farazlari, identifikatsiya cheklovlari, baholash barqarorligi va diagnostika bir butun metodologik tizim sifatida qaralishi lozim.

#### **Adabiyotlar ro'yxati.**

1. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. New York, Springer, 2009. 745 p.
2. Vahobov A. A., Karimov S. B. Matematik statistika va ehtimollik nazariyasi asoslari. Toshkent, O'zbekiston, 2019. 384 b.

3. Anderson T. W. An Introduction to Multivariate Statistical Analysis. New York, Wiley, 2003. 752 p.
4. G'aniev B. T., Qodirov D. R. Ko'p o'zgaruvchili regressiya va korrelyatsion tahlil usullari. Samarqand, Samarqand universiteti nashriyoti, 2020. 256 b.
5. Bishop C. M. Pattern Recognition and Machine Learning. New York, Springer, 2006. 738 p.
6. Gelman A., Carlin J. B., Stern H. S., Dunson D. B., Vehtari A., Rubin D. B. Bayesian Data Analysis. Boca Raton, CRC Press, 2013. 675 p.
7. Lorana D., Loran M. Graficheskie modeli: teoriya i prilozheniya. Moskva, Fizmatlit, 2011. 412 s.