



STATISTIK TAHLILNING DATA SCIENCEDAGI O‘RNI

Marifjonova Maxliyoxon Xursanjon qizi

FarDU Axborot tizimlari va texnologiyalari yo‘nalishi 3-kurs talabasi.

marifjonovamaxliyoxon@gmail.com

Abdukadirov Baxtiyor Abduvaxitovich

FarDU Axborot texnologiyalari kafedrasi dotsenti

bakxtiyor.uz@gmail.com

Anotatsiya: Ushbu matnda statistik tahlil va matematik statistikaning ahamiyati, ularning o‘zaro bog‘liqligi hamda ma’lumotlarni tahlil qilish jarayonidagi o‘rni kengroq yoritilgan. Shuningdek, zamonaviy axborot texnologiyalari va Data Mining usullari yordamida katta hajmdagi ma’lumotlardan foydali bilimlar olish, ularni qayta ishlash va qaror qabul qilishdagi roli ham tushuntirilgan.

Kalit so‘zlar: statistik tahlil, matematik statistika, Data Mining, ma’lumotlar tahlili, qaror qabul qilish, axborot texnologiyalari.

THE ROLE OF STATISTICAL ANALYSIS IN DATA SCIENCE

Marifjonova Makhliyoxon Khursanjon qizi


is a 3rd-year student in the Information Systems and Technologies Department at FarDU

marifjonovamaxliyoxon@gmail.com

Abdukadirov Bakhtiyor Abduvakhitovich

Associate Professor of the Department of Information Technologies at FarDU

bakxtiyor.uz@gmail.com



Annotation: This paper explores the significance of statistical analysis and mathematical statistics, highlighting their interrelationships and their critical role in the data analysis process. Furthermore, it examines the use of modern information technologies and Data Mining techniques to extract valuable knowledge from large datasets. The role of these methods in data processing and enhancing decision-making processes is also thoroughly discussed.

Keywords: Statistical analysis, mathematical statistics, Data Mining, data analysis, decision-making, information technologies.



РОЛЬ СТАТИСТИЧЕСКОГО АНАЛИЗА В DATA SCIENCE

Марифжонова Махлиёхон Хурсанжон кизи

Студентка 3-го курса направления "Информационные системы и технологии"

ФерГУ

marifjonovataxliyoixon@gmail.com

Абдукадиров Бахтиёр Абдувахитович

Доцент кафедры информационных технологий ФерГУ

bakxtiyor.uz@gmail.com

Аннотация: В данной статье рассматривается значимость статистического анализа и математической статистики, их взаимосвязь и роль в процессе анализа данных. Также освещается роль современных информационных технологий и методов Data Mining в извлечении полезных знаний из больших объемов данных, их обработке и принятии управленческих решений.

Ключевые слова: Статистический анализ, математическая статистика, Data Mining, анализ данных, принятие решений, информационные технологии.

Kirish: Hozirgi zamonda axborot hajmining keskin ortib borishi natijasida ma'lumotlarni to'g'ri yig'ish, saqlash va tahlil qilish masalalari dolzarb ahamiyat kasb etmoqda. Turli sohalarda — iqtisodiyot, tibbiyot, sanoat va boshqa yo'nalishlarda — samarali qaror qabul qilish bevosita mavjud ma'lumotlarning sifatli tahliliga bog'liqdir. Shu sababli statistik tahlil va matematik statistika usullari muhim ilmiy va amaliy vosita sifatida keng qo'llanilmoqda.

Statistik tahlil nafaqat ma'lumotlarni umumlashtirish va tasvirlash, balki ular asosida xulosalar chiqarish, qonuniyatlarni aniqlash va kelajakdagi jarayonlarni bashorat qilish imkonini beradi. Ayniqsa, zamonaviy axborot texnologiyalarining rivojlanishi natijasida katta hajmdagi ma'lumotlarni qayta ishlash va ulardan yangi bilimlar olish imkoniyatlari kengaydi. Bu jarayonda Data Mining kabi usullar alohida o'rin egallaydi.

Tasodifiy va oldindan aytib bo'lmaydigan ta'sirlar mavjud bo'lgan holatlarda ma'lumotlarni tahlil qilish bilan bog'liq vazifalarni yechish uchun matematiklar va boshqa tadqiqotchilar tomonidan so'nggi ikki yuz yil ichida matematik statistika deb ataluvchi metodlarning qudratli va moslashuvchan majmui yaratilgan.

O'tgan vaqt mobaynida bu metodlarni inson faoliyatining turli sohalarida (iqtisodiyotdan kosmik tadqiqotlargacha) samarali qo'llash borasida katta tajriba to'plangan. Muayyan sharoitlarda bu metodlar optimal yechimlarni olish imkonini beradi.

Umuman olganda, statistik tahlil metodik yo'nalish bo'lib, statistika sohasining markazi yoki "yuragi" hisoblanadi. Matematik statistikadan farqli o'laroq, statistik tahlil haqiqiy voqelikdagi masalalarni yechishga qaratilgan bo'lib, ekonometrika nuqtai nazaridan u iqtisodiyotdagi statistik usullarni anglatadi.

Biroq matematik statistika statistik tahlil uchun nazariy asos sifatida xizmat qiladi. Ya'ni statistik tahlil kursida asosiy e'tibor ma'lumotlarni tahlil qilish metodologiyasi va hisob-kitob algoritmlariga qaratiladi, teoremlardan esa ushbu algoritmlarni asoslashda foydalaniladi.

Bundan tashqari, statistik tahlil ikkita muhim yo'nalishni o'z ichiga oladi:

1. statistik tadqiqotni tashkil etish metodologiyasi (tadqiqotni rejalashtirish, ma'lumotlarni yig'ish va tahlilga tayyorlash, natijalarni talqin etish);
2. kompyuterda ma'lumotlarni qayta ishlashni tashkil etish (elektron jadvallar va ma'lumotlar bazalarini yaratish, maxsus dasturlardan foydalanish va h.k.).

Statistik tahlil usullarining turli sohalarda qo'llanishi natijasida sanoat statistikasi, tibbiyot statistikasi kabi amaliy yo'nalishlar shakllangan.

Statistik tahlil paketlari va dasturlariga kiritilgan usullar ko'pincha matematik statistika kursida to'liq o'rganilmaydi va natijada bu ikki yo'nalish o'rtasida uzilish yuzaga keladi. Bu uzilish qanday oqibatlarga olib kelishi mumkin?

- Birinchidan, matematik statistika bo'yicha tayyorlangan mutaxassislar haqiqiy ma'lumotlar bilan ishlashda qiyinchiliklarga duch keladi va amaliy masalalarni yechishda ojiz bo'lib qoladi.
- Ikkinchidan, tahlil jarayonida yo'l qo'yilgan xatolar muhim davlat ahamiyatiga ega bo'lgan hujjatlar, xususan statistika sohasidagi davlat standartlari sifatiga salbiy ta'sir ko'rsatishi mumkin.

Axborotni tahlil qilish atamasini har kim o'zicha tushunadi va natijada ko'p hollarda muammoning umumiy manzarasi aniq tasavvur qilinmaydi. Mazkur yondashuvdagi noizchilik nima yuz berayotganini va nima qilish kerakligini tushunmaslikka sabab bo'ladi. Hamma narsa bir-biri bilan yaxshi bog'lanmagan va umumiy asosga ega bo'lmagan bo'laklardan tashkil topadi.

Inson qarorlarni qanday qabul qilishidan kelib chiqib, tahlilning har bir mexanizmining o'rnini va vazifasini ko'rib chiqamiz. Biz fikr qanday tug'ilishini to'liq tushuntirib bera olmaganimiz uchun, asosiy e'tiborni bu jarayonda axborot texnologiyalaridan qanday foydalanish mumkinligiga qaratamiz.

- Birinchi holat — qaror qabul qiluvchi shaxs (QQQSh) kompyuterdan faqat ma'lumot olish vositasi sifatida foydalanadi, xulosalarni esa mustaqil chiqaradi. Bunday vazifalarni yechishda hisobot tizimlaridan, ma'lumotlarni ko'p darajali tahlil qilishda diagrammalar va boshqa vizuallashtirish usullaridan foydalaniladi.
- Ikkinchi holatda esa dastur nafaqat ma'lumotlarni oladi, balki ularni dastlabki qayta ishlaydi, ya'ni tozalash, silliqlash va hokazo ishlarni bajaradi. Shundan so'ng qayta ishlangan ma'lumotlarga tahlilning matematik metodlari — klasterlash, tasniflash, regressiya va boshqalar qo'llaniladi. Bu holda QQQSh xom ma'lumotlarni emas, balki qayta ishlangan ma'lumotlarni oladi, ya'ni inson kompyuter tayyorlagan modellar bilan ishlaydi.

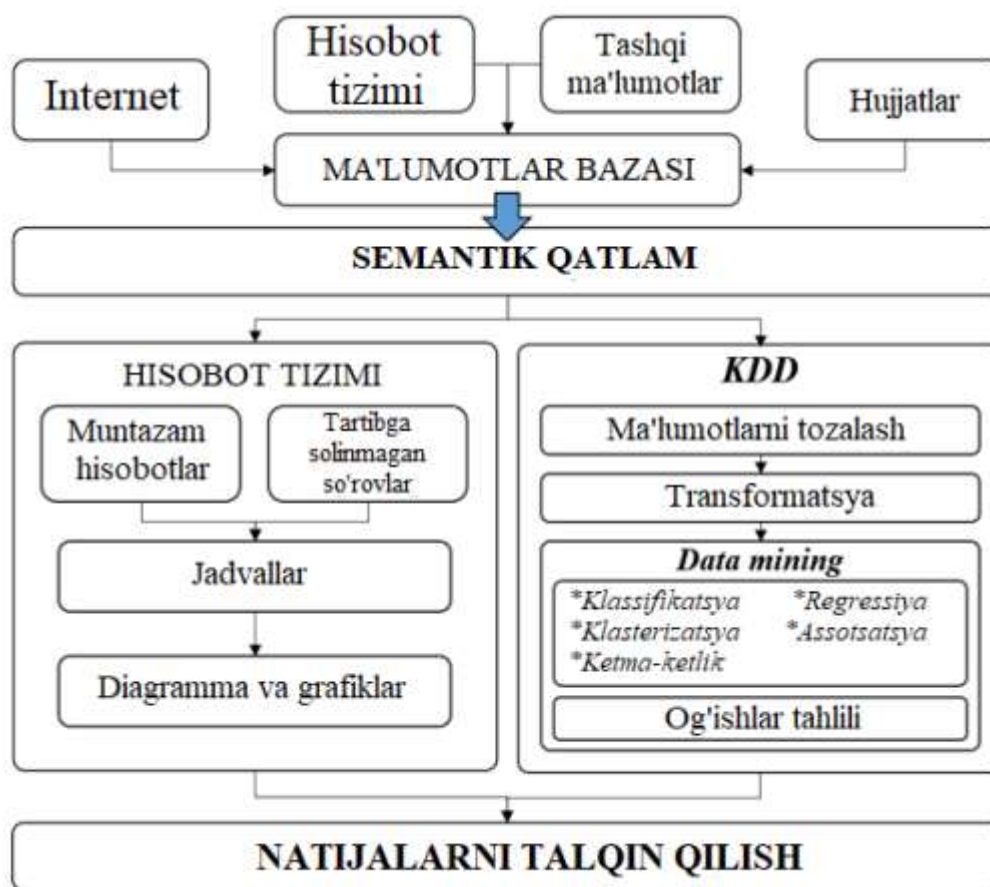
Birinchi holatda qaror qabul qilish bilan bog'liq barcha ish insonga yuklanadi. Mos modelni va qayta ishlash usullarini tanlash muammosi tahlil jarayonidan tashqarida qoladi.

Ya'ni qaror qabul qilish uchun asos sifatida yo yo'riqnoma (masalan, og'ishlarga qanday ta'sir qilish mumkinligi), yoki intuisiya xizmat qiladi. Ayrim hollarda bu yetarli bo'lishi mumkin, lekin agar QQQSh yanada chuqurroq bilim olishni istasa, bu yerda faqat ma'lumot olish mexanizmlari yetarli bo'lmaydi.

Qo'llaniluvchi dastlabki qayta ishlash va tahlil qilish mexanizmlarining barchasi QQQShga yuqoriroq darajada ishlash imkonini beradi. Birinchi holat taktik va operativ vazifalarni yechishga, ikkinchi holat esa bilimlarni ko'paytirish va strategik muammolarni hal etishga mos keladi.

Tahlilga nisbatan ikkala yondashuvni birgalikda qo'llash ideal holat hisoblanadi. Ular tashkilotning biznes-axborotni tahlil qilishga bo'lgan barcha ehtiyojlarini qondirish imkonini beradi. Vazifalarga qarab turli metodikalardan foydalanish orqali har qanday holatda mavjud axborotdan imkon qadar ko'proq bilim olish mumkin bo'ladi.

Ma'lumotlarning birlamchi manbai sifatida korxonani boshqarish tizimlarining ma'lumotlar bazalari, ofis hujjatlari va Internet xizmat qiladi. Chunki qaror qabul qilishda foydali bo'lishi mumkin bo'lgan barcha ma'lumotlardan foydalanish zarur.



Ma'lumotlarni qayd etish va saqlash metodlarining rivojlanishi yig'iladigan va tahlil qilinadigan axborot hajmlarining tez sur'atlarda o'sishiga olib keldi. Ma'lumotlar hajmi shunchalik kattaki, inson ularni mustaqil ravishda tahlil qilishga qodir emas. Ammo bunday tahlilni amalga oshirish zarurligi aniq, chunki aynan xom ma'lumotlarda qaror qabul qilishda

foydalanish mumkin bo'lgan bilimlar mavjud bo'ladi. Ma'lumotlarni avtomatik tahlil qilish uchun Data Mining usullaridan foydalaniladi.

Data Mining — bu inson faoliyatining turli sohalarida qaror qabul qilish uchun zarur bo'lgan, ilgari ma'lum bo'lmagan, murakkab (notrivial), amalda foydali va talqin qilish mumkin bo'lgan bilimlarni xom ma'lumotlardan aniqlash jarayonidir. Data Mining Knowledge Discovery in Databases (KDD) jarayonining muhim bosqichlaridan biri hisoblanadi.

Data Mining metodlarini qo'llash natijasida olingan axborot yangi va ilgari noma'lum bo'lishi kerak. Masalan, sotuvlarning o'rtacha qiymati bunday axborot hisoblanmaydi. Olingan bilimlar xususiyatlar o'rtasidagi yangi bog'lanishlarni ko'rsatishi, shuningdek ayrim belgilar asosida boshqa belgilar qiymatlarini bashorat qilish imkonini berishi lozim.

Xulosa: Tadqiqot davomida to'plangan ma'lumotlar va o'rganilgan adabiyotlar tahlili shuni ko'rsatadiki, bugungi shiddat bilan rivojlanayotgan raqamli iqtisodiyot davrida statistik tahlil shunchaki raqamlar yig'indisi emas, balki ulkan ma'lumotlar oqimidan foydali bilimlarni ajratib oluvchi asosiy intellektual filtrdir. James T. McClave va Terry Sincich kabi olimlarning ilmiy qarashlarida ta'kidlanganidek, statistik metodlar tasodifiylik ichidagi qonuniyatlarni aniqlash imkonini beradi, bu esa o'z navbatida Data Science sohasining fundamental poydevorini tashkil etadi.

Maqolada ko'rib chiqilganidek, matematik statistika va amaliy statistik tahlil o'rtasidagi uzviy bog'liqlikni tushunish strategik ahamiyatga ega. Agar matematik statistika nazariy asos va isbotlarni taqdim etsa, statistik tahlil bu nazariyalarni hayotiy masalalarga, xususan, David M. Levine va uning hammualliflari ta'kidlaganidek, boshqaruvda qaror qabul qilish jarayonlariga tatbiq etadi. Zamonaviy mutaxassis nafaqat matematik formulalarni bilishi, balki ushbu algoritmlarni Microsoft Excel, RAD Studio yoki boshqa maxsus dasturiy paketlar orqali real vaqt rejimida qo'llay olish ko'nikmasiga ega bo'lishi lozim.

O'zbekistonlik olimlar N.M. Soatov, X. Shodiyev va I. Xabibullayevlarning darsliklarida bayon etilgan uslubiy yondashuvlar shuni tasdiqlaydiki, mamlakatimizda ham statistika fani iqtisodiy-ijtimoiy jarayonlarni modellashtirishning asosi sifatida shakllangan. Biroq, zamonaviy axborot texnologiyalari va Data Mining usullari ushbu an'anaviy metodlarni yangi sifat bosqichiga olib chiqdi. Endilikda nafaqat o'tmishdagi ma'lumotlarni tahlil qilish, balki kelajakdagi tendensiyalarni yuqori aniqlikda bashorat qilish (forecasting) imkoniyati kengaydi.

Tadqiqot natijalariga ko'ra, Data Science sohasida muvaffaqiyatga erishish uchun faqatgina dasturlash tillarini bilish yetarli emas degan xulosaga kelindi. Haqiqiy bilimlar manbai — bu ma'lumotlarning statistik tabiatini chuqur anglashdir. Tahlilning operativ va strategik darajalarini birlashtirish orqali qaror qabul qilish tizimini avtomatlashtirish va inson omili bilan bog'liq xatolar ehtimolini minimal darajaga tushirish mumkin. Shunday qilib, statistik tahlil metodologiyasini mukammal egallash — har bir zamonaviy axborot tizimlari mutaxassisi uchun nafaqat akademik talab, balki uning mehnat bozoridagi kasbiy raqobatbardoshligini ta'minlovchi eng muhim omil hisoblanadi.



ADABIYOTLAR RO‘YXATI:

1. James T. McClave, Terry Sincich. Statistics. – 12th ed. – Pearson, 2013. – 814 p.
2. David M. Levine, David F. Stephan, Timothy C. Krehbiel, Mark L. Berenson. Statistics for Managers: Using Microsoft Excel. – 5th Edition. – Pearson Education, Inc, 2008. – 538 p.
3. Soatov N.M. Statistika: darslik. – T.: Ibn Sino, 2003.
4. Shodiev X., Xabibullayev I. Statistika: darslik. – T.: “IQTISOD-MOLIYA”, 2019. – 454 b.
5. Shodiyev X., Xabibullayev I. tahriri ostida. Statistika bo'yicha praktikum: o'quv qo'llanma. – T.: Tafakkur bo'stoni, 2015. – 336 b.
6. Aybjonov A.H., Mamatqulov B.X., Sayfullayev S.N. “Statistika” fanidan amaliy mashg'ulotlarni o'tkazish uchun o'quv qo'llanma. – T.: TDIU, 2016. – 180 b.
7. Елисеева И.И. Статистика: Учебник. – М.: Юрайт, 2012. – 559 с.
8. Shodiev X. Moliya statistikasi: darslik. – T.: “IQTISOD-MOLIYA”, 2010. – 320
9. Nabixo'jaev A. va boshqalar. Moliyaviy dasturlash: o'quv qo'llanma. – T.: “Iqtisodiyot”, 2010. – 208 b.
10. www.gov.uz - O‘zbekiston Respublikasi Hukumat portal.
11. www.lex.uz - O‘zbekiston Respublikasi Qonun hujjatlari ma'lumotlari milliy bazasi.
12. www.stat.uz - O‘zbekiston Respublikasi Statistika agentligi (sobiq Qo‘mitasi).