СӘТБАЕВ УНИВЕРСИТЕТІ

SATBAYEV UNIVERSITY

Статистика және деректерді талдау

Ержан Асел Ануаркызы,Ph.D., «Ақпараттық жүйелер» кафедрасының қауымдастырылған профессоры

Assel.yerzhan@satbayev.university



Орталық тенденция көрсеткіштері: Деректердің жүрегін тану

Статистиканың іргелі ұғымдары — орташа мән, медиана және мода — деректер жиынының мәнін ашып, экономикалық және әлеуметтік шешімдер қабылдаудың нақты бағытын белгілейді. Бұл презентация орталық тенденция көрсеткіштерін терең зерттеп, Қазақстан экономикасы мысалдарымен практикалық мазмұнын ашады.

КІРІСПЕ

1-бөлім: Статистикалық талдау негіздері

Статистикалық талдау — деректер дүниесінің тілі. Бұл бөлімде орталық тенденция көрсеткіштерінің теориялық негіздерін, олардың пайда болу тарихын және қазіргі заманғы ғылым мен практикадағы орнын қарастырамыз. Статистика — тек сандар жиынтығы емес, ол нақты шындықты сипаттайтын аналитикалық қуат.



Неге орталық тенденция маңызды?

Деректер — бізді қоршаған дүниенің санды бейнесі. Алайда жүздеген немесе мыңдаған сандардан тұратын жиынды тікелей түсіну мүмкін емес. Орталық тенденция көрсеткіштері осы мәселені шешеді — олар күрделі деректер жиынын бір немесе бірнеше мағыналы санмен сипаттауға мүмкіндік береді.

Деректерді жеңілдету

Жүздеген мәннен тұратын үлкен деректер жиынын бір ғана сипаттамалық санмен қорытындылау — талдаушыға уақытты үнемдетіп, шешімді жеделдетеді. Мысалы, 10 000 жалақы деңгейін орташа мән арқылы бірден бағалауға болады.

Экономикалық процестерді модельдеу

Күрделі макроэкономикалық үдерістерді — инфляция, жұмыссыздық, ЖІӨ өсімі — орталық тенденция арқылы жеңілдету экономистерге салыстырмалы талдау жүргізуге мүмкіндік береді. Бұл жеңілдету нақтылықты жоғалтпай, мағыналы қорытынды шығаруға септігін тигізеді.

Шешім қабылдаудың бағдары

Мемлекеттік саясаттан бастап бизнес стратегиясына дейін — барлық деңгейдегі шешімдер деректерге негізделеді. Орталық тенденция көрсеткіштері осы деректердің ең сенімді бағдары болып табылады: олар жалпы бағытты, үрдісті және ауытқуды анықтауға көмектеседі.

Орталық тенденцияның үш киелі көрсеткіші

Статистика ғылымында орталық тенденцияны сипаттайтын үш негізгі көрсеткіш бар. Олардың әрқайсысы деректер жиынының түрліше қырын ашады және белгілі бір жағдайларда артықшылыққа ие болады. Дұрыс көрсеткішті таңдай білу — талдаушының негізгі шеберлігі.



Орташа мән (Арифметикалық орта)

Барлық мәндерді қосып, санына бөлу арқылы есептелетін классикалық көрсеткіш. Деректер симметриялы таралған кезде ең тиімді нәтиже береді. Экономикалық есептеулерде кеңінен қолданылады.



Медиана (Орталық нүкте)

Деректерді ретімен орналастырғандағы дәл ортасындағы мән. Аномалиялық мәндерге (аутлайерлерге) сезімтал емес. Табыс пен байлықты бөлу талдауында алтын стандарт ретінде қолданылады.



Мода (Ең жиі кездесетін мән)

Деректер жиынында ең жиі қайталанатын мән. Сапалық деректерді талдауда және нарықтық зерттеулерде кеңінен қолданылады. Бірнеше мода болуы мүмкін — бұл таралудың көпмодалы сипатын білдіреді.

Орташа мән: Барлығының теңгерімі

Арифметикалық орта — статистиканың ең танымал және кеңінен қолданылатын көрсеткіші. Оның математикалық формуласы қарапайым: барлық мәндердің қосындысын олардың санына бөлу. Алайда оның мүмкіндіктері мен шектеуліліктерін терең түсіну маңызды.

Артықшылықтары

- Деректер жиынының барлық мәндерін есепке алады — ешбір ақпарат жоғалмайды
- Математикалық операцияларға қолайлы — дисперсия, стандартты ауытқу есептеулерінің негізі
- Кеңінен танымал және интуитивті түсінуге оңай
- Үлкен іріктемелерде аса тиімді жұмыс істейді

Кемшіліктері мен шектеулері

- Шектен тыс үлкен немесе кіші мәндерге (аутлайерлерге) өте сезімтал — бір ғана экстремалды мән нәтижені бұрмалай алады
- Ассиметриялы таралуларда нақты бейнені бермейді
- Сапалық деректерге қолдануға болмайды
- Кей жағдайда нақты мән ретінде кездеспейтін мән шығуы мүмкін

Мысалы: 10 адамның жалақысы 50 000–100 000 теңге аралығында болса, ал бір адамның жалақысы 5 000 000 теңге болса, орташа мән нақты картинаны бермейді.

Медиана: Ортадағы бақылаушы

Медиана — деректер жиынын өсу немесе кему тәртібімен орналастырғанда дәл ортада тұратын мән. Жұп санды деректерде екі ортаңғы мәннің арифметикалық ортасы алынады. Медианаңың ең басты ерекшелігі — ол шеткі мәндердің ықпалына бой алдырмайды.

1 Есептеу алгоритмі

Деректерді өсу тәртібімен реттеңіз. Тақ санды жиынта ортадағы элемент медиана болады. Жұп санды жиында екі ортаңғы элементтің орташасы есептеледі. Мысалы: {3, 7, 12, 20, 45} → медиана = 12; {4, 8, 15, 22} → медиана = (8+15)/2 = 11,5.

2 Аномалияларға төзімділік

Бір немесе бірнеше экстремалды мән медианаңы өзгертпейді. Мысалы, 5 адамның жалақысы: {80 000, 90 000, 100 000, 110 000, 2 000 000} теңге. Орташа мән: 476 000 теңге (шындыққа сай емес), ал медиана: 100 000 теңге — нақты бейне береді. Осыған байланысты табыс статистикасында медиана кеңінен қолданылады.

3 Практикалық қолданыс аясы

Табыс деңгейі мен байлықты бөлу статистикасы, жылжымайтын мүлік бағасын талдау, медициналық зерттеулердегі өмір сүру ұзақтығы деректері — барлығы медианаңы негізгі көрсеткіш ретінде пайдаланады. Қазақстан Ұлттық банкі де инфляция деңгейін бақылауда медианалық бағаларды қолданады.

Мода: Көпшіліктің таңдауы

Мода (французша *mode* — сән, үрдіс) — деректер жиынында ең жиі кездесетін мән. Ол жиіліктер кестесінде ең жоғарғы шыңға сәйкес келеді. Мода — орташа мән мен медианадан өзгеше, ол сандық та, сапалық та деректерге қолданыла алады.

Мода түрлері

- **Унимодалды:** тек бір басым мән бар — таралу симметриялы немесе бір шыңды
- **Бимодалды:** екі тең жиілікті мән бар — екі топтың болуын көрсетеді
- **Полимодалды:** үштен артық мода — деректердің күрделі гетерогенді табиғатын сипаттайды
- **Модасыз:** барлық мән бірдей жиілікте — мода анықталмайды

Қолдану салалары

- Маркетинг: ең танымал өнім өлшемін, түсін, брендін анықтау
- Сайлау социологиясы: партиялар мен кандидаттардың рейтингі
- Медицина: ең жиі диагноз, дозировка немесе жанама эффект
- Білім: ең жиі қойылатын баға, тест нәтижесінің ең жиі кездесетін мәні
- Қазақстандық нарық: ең сатылымды тауар санаттарын анықтау



2-БӨЛІМ

Қазақстан экономикасы деректеріндегі мысалдар

Теорияны практикамен бекіту — статистикалық білімнің негізгі принципі. Бұл бөлімде Қазақстанның нақты экономикалық деректері негізінде орталық тенденция көрсеткіштерінің қолданылуын талдаймыз. ЖІӨ, халық саны, жұмыссыздық және инфляция деректері — статистикалық талдаудың тамаша нысандары.

ЖІӨ динамикасы: 2001–2016 жылдар

Қазақстанның Жалпы Ішкі Өнімі (ЖІӨ) 2001–2016 жылдар аралығындағы деректер орталық тенденция көрсеткіштерінің практикалық мазмұнын айқын бейнелейді. 16 жылдық ұзақ уақыт кезеңіндегі деректер жиыны экономикалық циклдерді, өсу және құлдырау кезеңдерін айна-қатесіз көрсетеді.

122.53

Орташа ЖІӨ

Миллиард АҚШ долларымен есептелген 16 жылдық кезеңдегі орташа арифметикалық мән. Бұл Қазақстан экономикасының базалық күшін сипаттайды.

16

Талданған жыл

2001 жылдан 2016 жылға дейінгі жылдық ЖІӨ деректері негізінде орташа мән, медиана және мода есептелді.

~3x

Өсу қарқыны

Кезең барысында ЖІӨ шамамен үш есеге өсті — орташа мәннің жылдан жылға жоғарылауы бұл өсімді айқын бейнелейді.

Максималды және минималды мәндер арасындағы алшақтық орталық тенденцияның шеткі деректерге сезімталдығын тексеруге мүмкіндік берді. ЖІӨ мәліметтері асимметриялы таралған кезде медиана орташа мәнге қарағанда нақтырақ орталық бағаны береді — бұл Қазақстанның мұнай бағасына тәуелді экономикасы үшін ерекше өзекті.

Халық саны мен ЖІӨ корреляциясы

Орталық тенденция көрсеткіштері жеке-дара зерттелумен шектелмейді — олар екі және одан да көп айнымалылар арасындағы байланысты зерттейтін корреляциялық талдаумен тығыз байланысты. Қазақстанның халық санының өсімі мен ЖІӨ динамикасы арасындағы корреляция осы байланыстың нақты мысалы болып табылады.

Корреляция коэффициенті: r = 0,85

Оң және күшті байланыс. Бұл коэффициент халық санының өсуі ЖІӨ өсімімен қатар жүретінін статистикалық тұрғыдан растайды. Пирсон корреляция коэффициентінің 0,8-ден жоғары болуы өте күшті сызықтық байланыстың болуын білдіреді.

Интерпретация және шектеулер

Корреляция себеп-салдар байланысын білдірмейді — халық санының өсуі ЖІӨ өсімінің себебі емес, екі үдеріс бір-бірімен қатар жүреді. Алайда осы байланыс негізінде болжамдық модельдер жасауға болады: халық санының болжамы арқылы ЖІӨ-нің күтілетін деңгейін бағалауға мүмкіндік туады. Орталық тенденция көрсеткіштері мұнда референтті нүктені белгілейді.

- Регрессиялық модель арқылы болашақ ЖІӨ деңгейін болжау
- Орташа мән мен медианаңы пайдалана отырып деректер сапасын тексеру
- Аутлайерлерді анықтап, олардың корреляцияға әсерін бағалау

Жұмыссыздық пен инфляция: Кері байланыс

Экономикалық теорияда жұмыссыздық пен инфляция арасындағы кері байланыс Филлипс қисығы арқылы белгілі. Алайда Қазақстанның нақты деректері бойынша бұл байланыс күткендей айқын емес.

Корреляция коэффициенті: r = –0,059

Коэффициенттің нөлге жақын болуы екі айнымалы арасында статистикалық тұрғыдан маңызды байланыстың жоқтығын білдіреді. Бұл Қазақстан экономикасының ерекше табиғатымен — мұнай экспортына тәуелділігімен — байланысты болуы мүмкін. Мұнай бағасының өзгерісі жұмыссыздық пен инфляцияны бір мезгілде қозғайды, бұл классикалық Филлипс теңдіктен ауытқуға алып келеді.

Орталық тенденция арқылы жеке талдау

Жұмыссыздықтың орташа мәні мен медианасы — экономикалық тұрақтылықтың индикаторы. Инфляцияның орташа мәні жылдар бойынша мұнай бағасының тікелей ықпалында болды. Орталық тенденция арқылы қандай жылдары инфляция «қалыпты» деңгейден жоғары болғанын нақты анықтауға болады.

Тұрақтылықты бағалаудың қиындығы

Байланыс коэффициентінің төмен болуы саясаткерлер мен экономистерге жұмыссыздық пен инфляцияны бір-бірімен «сауда-саттық» жасамай дербес реттеу мүмкіндігін береді — бірақ бұл сонымен бірге бір инструментпен екі проблеманы шешу қиын екенін де білдіреді.

3-БӨЛІМ

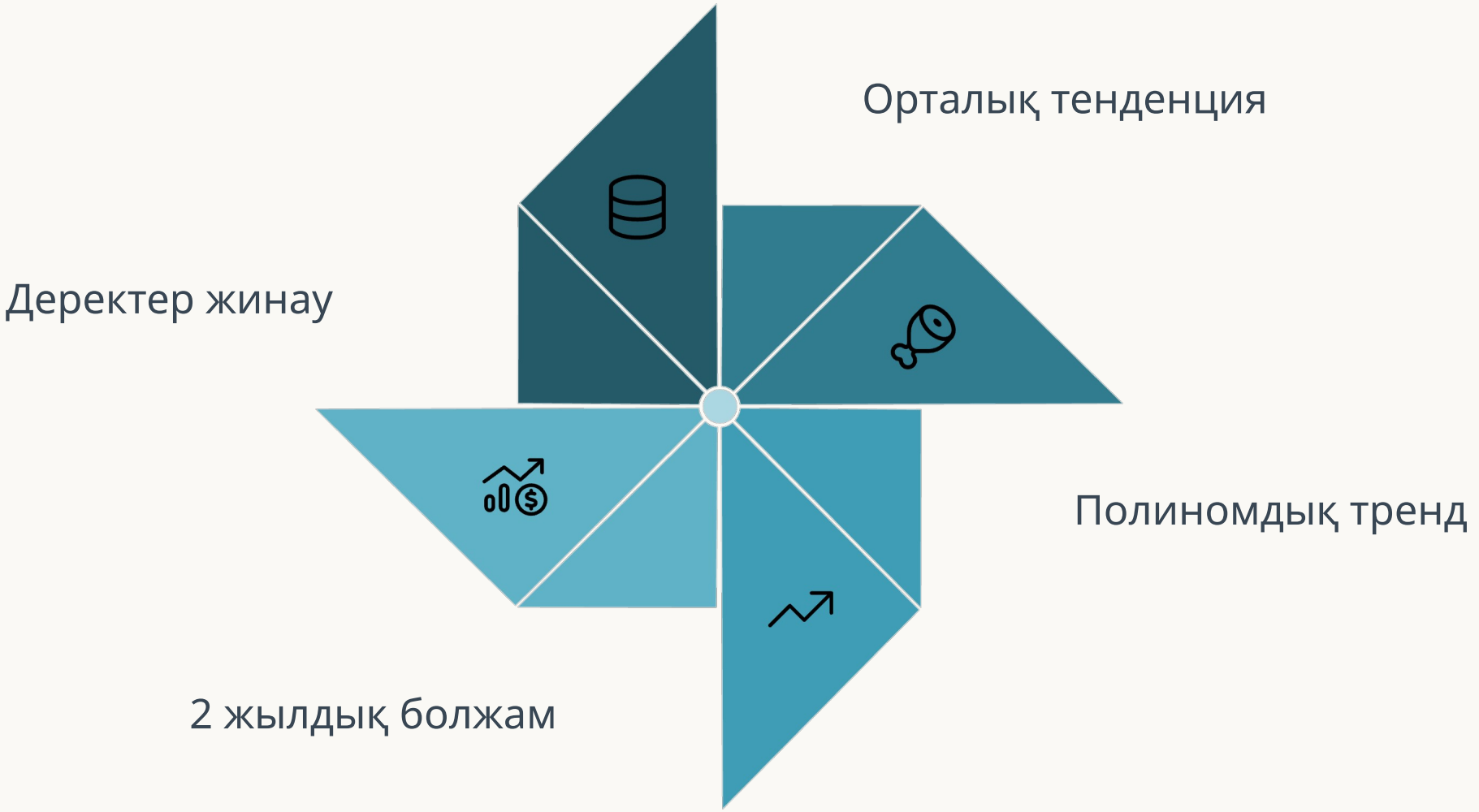
Практикалық қолдану және болжау

Орталық тенденция көрсеткіштері — тек өткенді сипаттайтын дескриптивті статистика ғана емес. Олар болжамдық модельдердің де іргетасы болып табылады. Бұл бөлімде полиноминалды трендтер, салалық мониторинг және болжамдық талдаудың практикалық тәсілдері қарастырылады.



Полиноминалды кестелер арқылы болжам

Уақыт қатарларын талдауда орташа мән жеткіліксіз — трендті, яғни уақыт өткен сайынғы жүйелі өзгерісті анықтау керек. Полиноминалды регрессия осы мақсатта орталық тенденция идеясын динамикалық кеңістікке ауыстырады.



Болжамның негіздері

Өткен деректер негізінде болашақты болжау үшін трендтің математикалық моделін құру қажет. Полиноминалды модель — сызықтық тренд жеткіліксіз болған жағдайда, яғни деректер қисық сызықты үрдіс көрсеткенде қолданылады. Орташа мән мен медиана деректердің «нормалды» деңгейін анықтай отырып, болжам дәлдігін тексеруге мүмкіндік береді.

Excel арқылы практикалық іске асыру

Microsoft Excel-дің Trend, Forecast және Polynomial Trendline функциялары орталық тенденцияны болжамдық модельге біріктіреді. 2 жылдық тұрақтылық болжамы экономикалық жоспарлауда кеңінен қолданылатын орта мерзімді горизонт болып табылады — бұл кезеңде деректердің орташа мәні болашақтың ең сенімді жуық бағасы ретінде алынады.

Салалық құрылым: 2026 жылғы мониторинг

Қазақстанның 2026 жылғы салалық мониторингі 3 492 кәсіпорынның пікіртерімі негізінде жүргізілді. Бұл деректер жиыны орталық тенденция көрсеткіштерінің салалық талдаудағы қолданылу мысалын жарқын бейнелейді.

Сауда саласы

Іскерлік белсенділік агрегатты көрсеткіші (ААК) = **0,27** — ең жоғары көрсеткіш. Бұл саудагерлердің жалпы жағдайды орташадан жақсы деп бағалайтынын білдіреді. Сауда кәсіпорындарының медианалық табысы да оң динамика көрсетті.

Көлік және қойма

Рентабельділік деңгейі: **34,8%** — мониторинг нысанындағы салалардың ішіндегі ең жоғарғы рентабельділік. Орташа мән бойынша бұл сала инвестициялық тартымдылық рейтингінде жоғары орын алады.

Өнеркәсіп

3 492 кәсіпорынның деректері бойынша өнеркәсіп саласының модалық жалақы деңгейі орта республикалық деңгейден жоғары. Медианалық табыстылық өнеркәсіпте тұрақты орта деңгейді сақтайды.

Пікіртерім деректерін өңдеуде орталық тенденция көрсеткіштері шешуші рөл атқарды: ААК орташа мән негізінде есептелді, ал жеке секторлардың рентабельділігін салыстыруда медиана қолданылды. Бұл тәсіл бірнеше аутлайер кәсіпорынның жалпы картинаны бұрмалауына жол бермеді.

4-БӨЛІМ

Статистикалық ауытқулар және олардың мәні

Орталық тенденция — бастапқы нүкте, бірақ деректердің толық бейнесі үшін ол жеткіліксіз. Деректердің шашыраңқылығы мен ауытқуы — статистикалық сурет толық болуы үшін міндетті түрде зерттелуі тиіс өлшемдер. Дисперсия, стандартты ауытқу және аутлайерлер туралы білім — кәсіби аналитиктің арсеналы.



Деректердің шашыраңқылығы: Орталықтан ауытқу

Екі деректер жиыны бірдей орташа мәнге ие бола отырып, мүлдем өзгеше табиғатта болуы мүмкін. Мысалы, {50, 50, 50, 50} және {0, 20, 80, 100} жиындарының орташасы бірдей — 50. Алайда бірінші жиын өте тұрақты, екіншісі — өте шашыраңқы. Осы айырмашылықты анықтауға шашыраңқылық өлшемдері қызмет етеді.

Дисперсия (Variance)

Орталық мәннен ауытқулардың квадраттарының орташасы. Дисперсия неғұрлым үлкен болса, деректер орташадан соғұрлым алшақ орналасқан. Есептеу формуласы: $\sigma^2 = \Sigma(x_i - \bar{x})^2 / n$. Дисперсияның бірлігі — бастапқы деректер бірлігінің квадраты, бұл интерпретацияны қиындатады.

Стандартты ауытқу

Дисперсияның квадрат түбірі — σ. Бастапқы деректермен бір бірліктегі болғандықтан, интерпретациялауға ыңғайлы. Стандартты ауытқудың кіші болуы деректердің орталыққа жиналғанын, үлкен болуы — шашыраңқы екенін білдіреді. Қаржылық тәуекелді бағалауда кеңінен қолданылады.

Өңірлік дисперсия

Қазақстан өңірлерінің табыс деңгейіндегі дисперсия жоғары. Астана мен Алматы қалаларының орташа жалақысы ауылдық өңірлерден едәуір жоғары. Бұл жоғары дисперсия республикалық орташа мәнді нақты картина ретінде қабылдауға болмайтынын ескертеді — медиана мен өңірлік жіктелген талдау міндетті.

Аутлайерлер: Статистикалық қателік пе, әлде жаңалық па?

Аутлайерлер (шеткі мәндер) — деректер жиынының басым бөлігінен мәнді түрде ерекшеленетін бақылаулар. Олар кейде өлшеу қателіктерінің белгісі, ал кейде — жаңа ашылымның ізі. Аутлайерлерді дұрыс анықтау және олармен дұрыс жұмыс жасау — аналитиктің жоғары шеберлігін талап ететін міндет.

Орташа мәнге әсері — айтарлықтай

Бір ғана экстремалды мән орташа арифметиканы елеулі мөлшерге жылжытады. Мысалы, 9 адамның айлығы 100 000 теңге, бір адамның — 10 000 000 теңге болса, орташа мән 1 081 818 теңгені көрсетеді — бұл 9 адамның нақты жағдайын мүлдем бейнелемейді. Сондықтан аутлайер бар деректерде орташаны негізгі көрсеткіш ретінде пайдалану қауіпті.

Медианаңың артықшылығы

Медиана аутлайерлерге иммунды. Деректер жиынына қаншалықты үлкен немесе кіші мән қосылса да, егер ол шеттерге орналасса, медиана өзгермейді. Бұл медианаңы экономикалық дағдарыс кезеңдерінде — бағалар мен табыстардың күрт секіруі байқалғанда — аса бағалы құрал етеді.

Аномалияларды анықтау

Экономикалық дағдарыс кезіндегі аномалдық деректерді анықтауда **интерквартильдік аралық (IQR)** кеңінен қолданылады: IQR = Q3 – Q1. Q1 – 1,5·IQR-дан төмен немесе Q3 + 1,5·IQR-дан жоғары мәндер аутлайер деп саналады. 2008–2009 жылдардағы қаржылық дағдарыс кезіндегі Қазақстанның кейбір макроэкономикалық деректері осылай анықталды.

5-БӨЛІМ

Деректерді визуализациялау өнері

Деректер тек кестелерде өмір сүрмейді — олар диаграммалар мен графиктер арқылы тірі болады. Орталық тенденция көрсеткіштерін визуализациялау — оларды аудиторияға тиімді жеткізудің ең сенімді жолы. Гистограммалар, қораптық диаграммалар және тығыздық қисықтары — аналитиктің визуалды тілі.



Гистограммалар: Жиіліктердің көрінісі

Гистограмма — деректердің жиілік таралуын баған диаграммасы түрінде бейнелейтін классикалық статистикалық график. Ол орталық тенденцияны — әсіресе моданы — визуалды түрде дереу анықтауға мүмкіндік береді.

Гистограмманы оқу

- **X осі:** деректер мәндерінің диапазондары (интервалдар)
- **Y осі:** әр интервалдағы бақылаулар саны (жиілік)
- **Ең биік баған:** мода орналасқан интервал
- **Симметриялы форма:** орташа мән ≈ медиана ≈ мода
- **Оңға қиғаш:** орташа мән > медиана (табыс таралуына тән)
- **Солға қиғаш:** орташа мән < медиана

Қазақстанның демографиялық мысалы

Қазақстанның жас құрылымының гистограммасы оңға қиғаш таралуды көрсетеді: жас топтары арасында 25–34 жас тобы ең жиі кездесетін мода болып табылады. Бұл Қазақстан халқының демографиялық дивидендтен (жас жұмыс күшінің мол болуынан) пайда алу мүмкіндігін статистикалық тұрғыдан растайды.

Нормалды (Гаусс) үлестірімде орташа мән, медиана және мода дәл бір нүктеде сәйкес келеді — бұл «жетілген» симметриялы таралудың белгісі. Нақты экономикалық деректер мұндай идеалды үлестірімнен жиі ауытқиды.

Қораптық диаграмма (Box Plot): Статистикалық портрет

Қораптық диаграмма бір суретте бес негізгі статистикалық сипаттаманы бейнелейді: минималды мән, бірінші квартиль (Q1), медиана, үшінші квартиль (Q3) және максималды мән. Бұл «бес санның қорытындысы» (five-number summary) деп аталады және деректер жиынының толық портретін шыншыл береді.

1

Минимум мен Максимум

Деректердің толық ауқымын — диапазонын — белгілейді. Мұсафат (Range) = Max – Min. Аутлайерлер болмаған жағдайда бұл деректердің шеткі нүктелері болады.

2

Квартильдер Q1 және Q3

Q1 — деректердің төменгі 25%-ын, Q3 — жоғарғы 25%-ын шектейді. Қорап (box) осы екі квартиль арасындағы орта 50% деректерді қамтиды. IQR = Q3 – Q1 деректердің негізгі шоғырлану аймағын көрсетеді.

3

Медиана — қораптың ішінде

Медианаңы білдіретін сызық қораптың ішінде орналасады. Егер ол қораптың ортасынан жоғары немесе төмен болса — таралудың асимметриялылығы туралы хабарлайды.

4

Өңірлік ЖІӨ салыстыруы

Қазақстанның өңірлік ЖІӨ деректерін қораптық диаграммамен салыстыру өңірлер арасындағы теңсіздіктің нақты бейнесін береді. Астана мен Атырау өңірлерінің ЖІӨ мәндері жоғарғы аутлайер ретінде айқын ерекшеленеді.



6-БӨЛІМ

Болжау және стратегиялық жоспарлау

Орталық тенденция — өткенді сипаттаудан болашақты болжауға дейін созылатын аналитикалық жол. Трендті анықтау, жылжымалы орташаны есептеу және медианалық болжам — стратегиялық жоспарлаудың статистикалық арқауы. Бұл бөлімде осы тәсілдердің практикалық механизмдері ашылады.

Орталық тенденция арқылы трендті анықтау

Уақыт қатарларындағы тренд — деректердің ұзақ мерзімді жүйелі бағыты. Орталық тенденция оны анықтаудың және тегістеудің негізгі құралы болып табылады.

Жылжымалы орташа (Moving Average)

Уақыт қатарындағы орташа мән «тереземен» жылжытылады — мысалы, 3 жылдық немесе 5 жылдық жылжымалы орташа. Бұл тәсіл қысқа мерзімді тербелістерді тегістеп, ұзақ мерзімді тренді айқын ашады. ЖІӨ динамикасын талдауда кеңінен қолданылады.

1

Инфляцияны бақылаудағы медиана

Орталық банктер инфляцияны бақылауда медианалық инфляцияны қолданады — жеке тауарлар мен қызметтер бағасының өзгерісінің медианасы. Бұл тәсіл бірнеше тауарда болатын күрт баға секірісінің жалпы инфляция бейнесін бұрмалауынан қорғайды.

2

2026–2030 жылдарға болжам

Қазақстан Ұлттық экономика министрлігінің болжамдары орташа мән мен трендтік регрессияны үйлестіреді. ЖІӨ өсімінің орташа болжамы жылына 4–5%, инфляцияның медианалық болжамы — 7–9% деңгейінде. Бұл болжамдар деректер жиынынан есептелген орталық тенденцияға негізделеді.

3

Салалық талдау: Қай көрсеткіш қайда тиімді?

Орташа мән, медиана және мода — бірін-бірі алмастыратын баламалар емес, бір-бірін толықтырып, бірге пайдаланылатын аналитикалық триада. Саланың табиғатына қарай тиімді көрсеткішті таңдай білу — статистиктің кәсіби шеберлігінің белгісі.

Жалақы деңгейі

Медиана — таптырмас таңдау. Жоғары жалақылы топ орташа мәнді жоғары тартады, ал медиана нақты «орталық жұмысшының» жалақысын бейнелейді. Қазақстанда медианалық жалақы орташа жалақыдан 15–25%-ға төмен — теңсіздіктің айқын белгісі.

Өндіріс көлемі

Мода — өндіріс кәсіпорындарының ең жиі кездесетін өндіріс деңгейін анықтайда. Модалық кәсіпорын — орташа емес, ең типтік кәсіпорын. Нарықтық стратегия жасауда мода нақтырақ бағдар береді.

Инвестициялық жобалар

Орташа мән мен дисперсия — инвестициялық тиімділікті бағалаудың жұбы. Орташа табыстылық жобаның күтілетін нәтижесін, ал дисперсия (немесе стандартты ауытқу) — тәуекел деңгейін өлшейді. Екеуі бірге «тәуекел-табыстылық» профилін толық сипаттайды.

7-БӨЛІМ

Заманауи цифрлық талдау

Цифрлық революция статистикалық талдаудың масштабын, жылдамдығын және тереңдігін мүлдем жаңа деңгейге көтерді. Үлкен деректер (Big Data), бұлтты есептеу және жасанды интеллект — орталық тенденция көрсеткіштерін нақты уақыт режимінде миллиондаған бақылаулар бойынша есептеуге мүмкіндік береді.



Үлкен деректер (Big Data) және орталық тенденция

Big Data дәуірінде орталық тенденция көрсеткіштерін есептеу — секундтар мәселесі. Алайда масштаб ұлғайғанда жаңа сын-қатерлер де туындайды: деректер сапасы, жиіліктің дұрыс бейнеленуі, аутлайерлердің ықпалы — бәрі де жаңаша шешімдерді талап етеді.

Транзакциялық деректерді талдау

Қазақстан банктері күн сайын миллиондаған транзакцияны өңдейді. Орташа транзакция мөлшері, медианалық шығын, модалық сатып алу санаты — бұл мәліметтер тұтынушылық мінез-құлықты нақты бейнелейді. Нақты уақыт режиміндегі орташа есептеу алаяқтықты (фрод) дереу анықтауға мүмкіндік береді: транзакция орташадан 3 стандартты ауытқуға асып кетсе — жүйе автоматты түрде ескертеді.

Нақты уақыт режиміндегі есептеу

Apache Spark, Apache Kafka сияқты заманауи технологиялар секундына мыңдаған жаңа бақылаулар қосылып жатқан кезде де орташа мән мен медианаңы үздіксіз жаңартып отырады. Бұл мемлекеттік статистикалық органдардың мониторинг жүйелерін нақты уақыт режиміне көшіруіне мүмкіндік береді.

Цифрлық Қазақстан бағдарламасы

«Цифрлық Қазақстан» мемлекеттік бағдарламасы шеңберінде мемлекеттік қызметтердің тиімділігін бағалауда орталық тенденция кеңінен қолданылады. Қызмет көрсету уақытының медианасы, азаматтардың қанағаттану деңгейінің орташа мәні, шағымдар санының модасы — бәрі де e-government жүйесін жетілдірудің деректерге негізделген жолын белгілейді.

Жасанды интеллект және статистикалық болжам

Жасанды интеллект алгоритмдерінің негізінде — статистика. Машиналық оқыту модельдерінің абсолютті көпшілігі орталық тенденция концепцияларын — орташа мән, дисперсия, нормализация — кеңінен пайдаланады. ЖИ — статистикалық талдаудың жаңа, масштабталған нысаны.

Алгоритмдердегі орталық тенденция

- **Деректерді нормализациялау:** орташа мән мен стандартты ауытқу арқылы деректерді стандартты шкалаға келтіру — нейрондық желілердің алғашқы қадамы
- **k-means кластерлеу:** кластердің орталығы — орташа мәннің тікелей қолданысы
- **Аномалияларды анықтау:** орташадан N стандартты ауытқуға асып кеткен мәндерді бөлу
- **Регрессиялық болжам:** орташа квадраттық қате (MSE) минимизациясы арқылы үйрену

Басқарушылық шешімдер

Деректерге негізделген басқару (Data-Driven Governance) — Қазақстандың мемлекеттік басқарудағы жаңа бағыты. Орталық тенденция арқылы есептелген индикаторлар:

- Мемлекеттік бюджет шығыстарының медианалық тиімділігі
- Мектеп нәтижелерінің өңірлік орташа мәні
- Ауруханалардың хизмет уақытының модасы
- Азаматтардың цифрлық белсенділігінің тренді

ЖИ моделдері осы индикаторлар негізінде саясат нәтижелерін болжап, ресурстарды оңтайлы бөлуге ұсыныстар береді.

8-БӨЛІМ

Қорытынды және болашаққа көзқарас

Орталық тенденция — деректердің жүрегі. Бұл бөлімде біз статистикалық сауаттылықтың қоғамдық мәнін, деректерді манипуляциядан қорғаудың жолдарын және болашақтың деректерге негізделген дүниесін зерттейміз. Қазақстан үшін статистикалық мәдениет — бәсекеге қабілеттіліктің іргетасы.



Статистикалық сауаттылық: Қоғамның даму кілті

Деректер — жаңа заманның мұнайы деп айтылады. Алайда мұнайды өндіру қалай дұрыс технология талап етсе, деректерді дұрыс оқу да арнайы білімді — статистикалық сауаттылықты талап етеді. Орталық тенденция көрсеткіштерін дұрыс түсіну азаматтарды манипуляциядан қорғайды.

Деректерді дұрыс оқудың маңызы

Орташа жалақы өссе — бірақ медиана өспесе — бұл байлықтың тек жоғарғы топта жинақталып жатқанын білдіреді. Статистикалық сауатты азамат бұл айырмашылықты дереу түсінеді. Мұндай сыни ойлау саяси шешімдерді бағалауда, тұтынушылық таңдауда және мансаптық стратегияда ажырамас дағды.

Манипуляциядан қорғану

Орташа мәнді қалай бұрмалайды? Жарнамашылар «орташа тұтынушы» деп өз мақсаттарына сай таңдалған сегментті ұсынады. Саясаткерлер «орташа жалақы өсті» дейді — бірақ медиана өспеген. БАҚ «ең жоғарғы» немесе «ең төменгі» мысалды жалпы тенденция ретінде көрсетеді. Бұл манипуляциялардың бәрін орталық тенденция туралы базалық білім ашып береді.

Азаматтық статистикалық мәдениет

Финляндия, Эстония, Сингапур сияқты елдерде статистикалық сауаттылық мектеп бағдарламасының міндетті бөлігі. Қазақстанда да мектептерде деректерді оқу, графиктерді интерпретациялау және орталық тенденция көрсеткіштерін қолдану дағдылары кеңейтілуде. Бұл — болашақтың бәсекеге қабілетті ұлтын қалыптастырудың іргетасы.

Статистикалық сауаттылық: Деректерге негізделген шешімдер

Орталық тенденция көрсеткіштерін терең түсіну — кез келген саладағы тиімді басқарудың іргетасы. Дәрігер, экономист, мұғалім, кәсіпкер немесе мемлекеттік қызметкер — барлығы деректермен жұмыс жасайды. Орташа мәнді, медианаңы және моданы дұрыс таңдай білу — дұрыс шешімнің бірінші қадамы.

Деректер — тек сандар емес

Деректерді тек сандар жиынтығы ретінде емес, нақты экономикалық және әлеуметтік процестердің индикаторы ретінде қабылдау — статистикалық ойлаудың жоғары деңгейі. Орташа жалақы артындағы нақты адамдарды, медиана артындағы экономикалық теңсіздікті, мода артындағы нарықтық үрдісті «оқи» білу — кәсіби аналитиктің белгісі.

Статистикалық сауаттылық — қоғамның кілті

Статистикалық сауаттылық — жеке мансаптың ғана емес, бүкіл қоғамның дамуының кілті. Деректерге негізделген шешімдер қабылдайтын үкімет, бизнес және азаматтық қоғам — тиімді, мөлдір және бәсекеге қабілетті мемлекеттің негізі. Қазақстанның 2050 жылға арналған стратегиялық мақсаттарына жетудің жолы — деректер арқылы өтеді.

Орталық тенденция — статистиканың іргетасы. Оны дұрыс қолдану: дұрыс сұрақ қою, дұрыс деректерді жинау, дұрыс көрсеткішті таңдау және нәтижені сыни бағалаудан басталады.

Болашаққа көзқарас: Сандардың күшімен жаңа белестерге

Болашақта орталық тенденция көрсеткіштері жасанды интеллект пен үлкен деректер (Big Data) арқылы одан да дәл, одан да жылдам және одан да кең ауқымда есептелетін болады. Нақты уақыт режиміндегі экономикалық индикаторлар, автоматтандырылған аномалия анықтау жүйелері және болжамдық аналитика — бәрі де орталық тенденция концепцияларының заманауи жалғасы.

ЖИ + Статистика = Жаңа аналитика

Жасанды интеллект орталық тенденция есептеулерін автоматтандырып, адам аналитигін интерпретация мен стратегиялық шешімге жіберуге мүмкіндік береді. Адам мен алгоритмнің ынтымақтастығы — болашақтың аналитикалық парадигмасы.

Қазақстан экономикасының тұрақтылығы

Деректерді дұрыс интерпретациялау — Қазақстан экономикасының тұрақтылығы мен инновациялық өсімін қамтамасыз ететін негізгі құрал. Орталық тенденция арқылы экономикалық циклдерді ертерек анықтап, дағдарысты алдын ала болғызбайтын саясат жасауға болады.

Өткеннен болашаққа

Статистикалық талдау — өткеннің тәжірибесін болашақтың табысына айналдырудың ең сенімді жолы. Орталық тенденция көрсеткіштері — осы жолдың ең сенімді бағдары, деректер дүниесінің жүрегі, сандар тіліндегі шындықтың дауысы.