

ГТФХР 83.77.23:27.43.17

Ұ.М. Дүйсенбай

«Астана» университеті, Нұр-Сұлтан, Қазақстан

(E-mail: [Ulzhanka\\_dm@mail.ru](mailto:Ulzhanka_dm@mail.ru))**Статистикалық мәліметтерді есептеудің кейбір мәселелері**

**Аннотация:** Статистикалық ақпарат қолданушыларға зерттеуге қажетті өнімді жылдам алу үшін ең тиімді және қолайлы құрал болып табылады. Мақалада тиімді әрі түсінікті ақпараттар мен мәліметтерді алу үшін статистикалық талдаулар мен тәжірибелердің маңыздылығы, және оларды жасау жолдары көрсетілген.

**Кілт сөздер:** статистикалық мәлімет, математикалық статистика, талау, көпөлшемді, рангілік корреляция.

DOI: <https://doi.org/10.32523/2220-685X-2019-54-3-37-43>

Статистика мәліметтерді жинау тапсырмасымен және оның негізінде қандай да бір тұжырым алудың қиындығымен беттесетін кез келген ғылыми сала үшін маңызы зор. Осы қарапайым мазмұндалған материалдардың басында негізгі түсініктер және ықтималдылық теориясы, математикалық статистика жиналған. Статистика-ғылымның бір тармағы, практикалық қызмет саласы, білім бағыты, таным құралы. Статистика дегенде мәліметтерді жинақтау процесінде, оларды қайта өңдеп, бейтарап әрі анық талдау ережелерін де түсінеміз. Осы тұрғыдан алып қарағанда, статистика - әрі ғылым, әрі қызмет саласы, әрі мамандық түрі. Ол арнайы ұйымдастырылған білім тармағы да, белгілі бір мақсатты жүзеге асыруға қажетті пәрменді құрал да, кәсіпкерлік қызмет саласы да болып табылады.

Статистика сөзі ғылыми әдебиетте XVIII ғасырда қолданысқа енді және алғашқыда «мемлекеттану» деген мағынада ұғынылды. Бірақ статистика ғылымы XVII ғасырдың ортасында «саяси\_арифметика» нысанында дами бастады. Ғылымдағы бұл бағыттың бастауында ағылшын ғалымы - Уильям Петти тұрды. Статистика терминін ғылымға 1746 жылы неміс ғалымы Готфрид Ахенваль өзі сабақ беретін Германия университетінде «Мемлекет ісін жүргізу» курсының орнына «Статистиканы» қолдануды

ұсынды. Осыған қарамастан статистиалық санаулар бұдан да ерте жүргізілген болатын: Көне Қытайдахалық санағы, Көне Римде елдердің әскери күшін салыстыру, азаматтардың мүліктерін бағалау. Статистика XIX ғасырдың басында А. Кетленің және оның мектебі өкілдерінің еңбектерінде қоғамдық құбылыстардың заңдылықтарын зерделейтін ғылым мен құралға айналды. Статистика қоғамдық ғылым ретінде қоғамдық құбылыстарды олардың сапалық мазмұнымен тығыз байланыста қарастырылатын сандық қатынастарының қалыптасу және өзгеру заңдылықтарын зерделейді. Қоғамдық құбылыстарды белгілі бір уақытпен және орынмен байланысқан нақты тарихи жағдайларда сандық тұрғыда зерделеу статистика ғылымының айрықша белгісі болып табылады. Заңдылықтарды тану және құбылыстардың дамуын айқындайтын факторларды анықтау үшін пайдаланылатын статистикалық сипаттамалардың өзіндік ерекшелігі олардың сандық және сапалық қасиеттерінің ажырағысыз бірлігінде. Сондықтан статистика ғылымы әлеуметтік-экономикалық құбылыстарды талдауды өзіндік (ерекше) әдістер, тәсілдер, ережелер жиынтығы, яғни статистикалық әдіснама арқылы жасайды. Оны басқа да ғылымдар өз зерттеулерінде, тәжірибесінде қолданады. Статистикалық әдіснаманың негізгі бөліктеріне статистикалық байқау, топтастыру, орта шамалар, индекстер, корреляция талдау, т.б. әдістер жатады.

Статистикалық мәліметтер дегеніміз жүргізілген байқаулардың, тәжірибелердің қорытындысы. Математикалық статистиканың негізгі міндеттерінің бірі статистикалық мәліметтердің заңдылықтарын айқындап, оларға сәйкес экономикалық модель құру және шешімдер қабылдау. Математикалық статистиканың бірінші міндеті байқаулар, тәжірибелер арқылы алынған мәліметтерді жинастырып, топтастыру әдістерін көрсету. Ал екінші міндеті статистикалық мәліметтерді алға қойған мақсатқа байланысты талдау жасау әдістерін табу.

Математикалық статистика-ол өңдеу және бақылау немесе өлшеудің нәтиже анализінің кездейсоқ көрінісін зерттейтін ғылым. Математикалық статистиканың бірінші міндеті-арнайы қойылған тәжірибелер немесе жасалған өлшеулерді бақылау нәтижесінде жиналған статистикалық мәліметтерді топтау және өңдеуді алу жолдарын көрсетеді.

Математикалық статистиканың екінші міндеті-зерттеудің мақсатына байланысты статистикалық анықтамалардың анализ әдістерін құру. Мысалы, зерттеудің мақсаты мынадай болуы мүмкін:

- оқиғаның белгісіз ықтималдығының бағасы;
- кездейсоқ шаманың үлестірілу параметрінің бағасы;
- үлестірілу параметрлері немесе белгісіз үлестірілу түрі туралы гипотездерді тексеру;
- бір кездейсоқ шамаға немесе бірнеше кездейсоқ шамаға байланысты және тағы басқа баға.

*Бас жиынтық және таңдама.* Бас жиынтық деп белгілі қасиеттерімен бірігетін барлық қарастырылып отырған объектілер жиынын атайды. Жеке объект осы жиынның элементі болады. Егер бір кәсіпорынды алатын болсақ, онда жеке жұмысшы объект болады да, ал кәсіпорынның барлық жұмысшылары бас жиынтық болады. Бас жиынтықтағы объектілер саны ақырлы немесе ақырсыз көп сан болуына байланысты бас жиынтық шектелген және шектелмеген болып екіге бөлінеді. Таңдама дегеніміз бас жиынтықтан кездейсоқ таңдап алынған объектілер жиынтығы. Бас және таңдама жиынтықтағы объектілер саны олардың көлемі деп аталады. Бас жиынтықтың көлемі үлкен болса, түгелдей тексеру мүмкін емес. Егер зерттелініп отырған объектілер жиыны жойылуға жататын (өзіндік құнын жоғалтуына байланысты), немесе практикада өте көп материалдық шығындарды қажет ететін болса, онда толық талдаудың мәні болмайды. Бас жиынтық жөніндегі тұжырым дұрыс болуы үшін бөлініп алынған таңдама бас жиынтық туралы жан-жақты сипаттама беруге тиіс. Бас жиынтық алдын-ала өзара қиылыспайтын бірнеше серияға бөлінеді де, ол сериялардың біреуі кездейсоқ таңдап алынса, онда осы серияның барлық элементтері таңдама жиынтық құрады.

*Таңдаманың статистикалық үлестіруі.* Бас жиынтықтан көлемі  $n$  болатын таңдама бөлініп алынды делік. Оларды  $x$ -тің мәндерінің өсу ретіне байланысты  $x_1, x_2, \dots, x_k$  деп жазамыз. Бақылау бойынша  $x_1$  элементі  $n_1$  рет,  $x_2$  элементі  $n_2$  рет, тағы сол сияқты  $x_k$  элементі  $n_k$  рет пайда болсын, мұндағы  $n_1 + n_2 + \dots + n_k = n$ , онда тәжірибелер нәтижесінде алынған  $x_i$ -варианталар, ал  $x_1, x_2, \dots, x_k$  тізбегі вариациялық қатар деп

аталады.  $n_1, n_2, \dots, n_k$ -сәйкес варианттардың жиіліктері,  $\omega_i$ -варианттардың салыстырмалы жиіліктері, мұндағы:

$$\omega_1 = \frac{n_1}{n}, \omega_1 + \omega_2 + \dots + \omega_k = 1.$$

Таңдаманың статистикалық үлестіруі деп варианттар мен олардың жиіліктерінің немесе салыстырмалы жиіліктерінің арасындағы өзара бірмәнді сәйкестікті атайды. Таңдаманың құлашы деп оның ең үлкен және ең кіші варианттарының айырымын айтады. Вариациялық қатар дискретті және үздіксіз деп екіге бөлінеді. Егер шаманың мәндері бір-бірінен кемінде кейбір тұрақты саннан өзгеше болса, онда вариациялық қатар дискретті деп аталады.

Көпөлшемді статистикалық талдауды дәстүрлі жалпы статистика курсының логикалық жалғасы ретінде қарастыруға болады. Мұндағы түбегейлі айырмашылық нысандар мен құбылыстардың өзгерістері бір немесе екі белгінің емес, көптеген белгінің жиынының әсері тұрғысынан зерттеледі. Ол зерттелетін нысандар мен құбылыстарды толықтай сипаттауға және қорытындының шынайы болуына мүмкіндік береді.

Мысалы, адамды тек еңбекақысымен немесе еңбекақы мен білім деңгейімен сипаттасақ, қорытынды шектеулі, әрі жаңсақтау болар еді. Ал егер адамды денсаулығы, әлеуметтік жағдайы, кәсібилігі т.с.с. тұрғысынан сипаттасақ, субъектінің бейнесін барынша толық модельдей алар едік.

Кәсіпорында еңбек өндірімділігі жоғары болса, ол кәсіпорын тұрақты жұмыс істейтін сенімді серік бола алады деген тұжырым туғызбайды. Ол үшін кәсіпорынның өтімділігі, қаржылық тұрақтылық, инвестицияның тиімділігі т.с.с. мәліметтері қажет. Сондықтан нақты зерттелетін нысандар мен құбылыстар әрдайым көпөлшемді немесе көпбелгілі болып келеді. Көпөлшемді статистикалық талдау - алдын ала берілген ақпараттың көпөлшемді геометриялық кеңістікте беріліп және жасырын түрде ықпал беретін заңдылықтарды анықтау үшін қолданылатын терең статистикалық тәсілдердің жиынтығы.

Оның ерекшеліктері мынада:

- көпөлшемді статистикалық талдау-салыстырмалы жас ғылым;

- көпөлшемді статистикалық талдау-кіретін факторлық талдау - XIX-XX ғасырлар тоғысында, көпөлшемді талқылау - XX ғасырдың 30-шы жылдарында, ал кластерлік талдау XX ғасырдың

10-20-шы жылдарында пайда болды. Тәсілдердің негізгі бөлігі әлі де жетілдірілу үстінде және қолданылатын салалары әлі нақтылана қойған жоқ;

- көпөлшемді статистикалық талдаудың тәсілдері күрделі логикалық-математикалық құрылыммен және формулалармен жүзеге асырылады. Ол экономикалық теорияны және математиканы терең білуді қажет етеді;

- көпөлшемді статистикалық талдау тәсілдерін қолдану белгілі бір дәрежеде шығармашылықты талап етеді. Мысалы, бір типті есепті шығару үшін ондаған әдіс-амал, ал кластерлік талдауда 200 түрлі әдіс керек. Сондықтан есепті дұрыс шешу үшін кәсібилік пен алдын ала көрегендік қажет;

- көпөлшемді статистикалық талдауда көпбелгілі мәліметтер қолданылады. Бірнысанды сипаттайтын белгінің саны 100-ге дейін жетуі мүмкін. Нақты нысанды сипаттау үшін 33 белгі жеткілікті деген көзқарас бар. Бұл көзқарас дәлелденген. Бірақ белгінің саны зерттеу мақсаты мен міндетіне байланысты болады;

- көпөлшемді статистикалық талдаудың барлық тәсілдерін қолдану есептегіш техниканы қажет етеді, себебі есептеу ауқымы мен күрделілігінің деңгейі өте жоғары болып келеді.

Рангілік корреляция.

Екі айнымалының айнымалының арасындағы корреляция коэффициентінің мәнін есептеу айнымалылардың мәнінің дәл болуын талап етеді. Көп жағдайда мәндер дәл болмай жатады. Мұндайда рангілік корреляция коэффициенті қолданылады. Оның мәнісі әрбір мәнге рангінің берілуінде және сол рангілердің арасындағы байланыстың тығыздылығының анықталуында.

Дисперсиялық талдаудың негізгі мақсаты орташа мәндерді өзара салыстыра отырып, айырмашылықтарының елеулі немесе елеусіз екендігін анықтау болып табылады.

Ол тексеру жалпы дисперсияны кездейсоқ қате мен орташа мәндердің айырмашылықтары арқылы сипаттаумен жүргізіледі. Дисперсияның соңғы компоненті орташа мәндердің айырмашылығының статистикалық елеулі болса, онда нөлдік жорамал теріске шығарылып, балама болжам қабылданады.

Сонымен қатар дисперсиялық талдау белгілі бір нәтижелік  $X$  факторына әсер ететін факторлардың ықпалын бағалау үшін қолданылады. Іс жүзінде дисперсиялық талдау  $F_1, F_2, \dots, F_n$

мәндері бар  $F$  сапалық фактордың  $X$  нәтижелік факторға елеулі әсер ететіндігін сипаттау үшін қолданылады. Дисперсиялық талдаудың алғашқы жұмыстары нақ осы егін шаруашылығында жүзеге асырылды. Дисперсиялық талдаудың негізгі идеясы - фактор туғызған «факторлық дисперсия» мен кездейсоқ себептерден пайда болатын «қалдық» дисперсияны өзара салыстыру. Егер аталған дисперсиялардың өзара айырмашылығы елеулі болса, онда  $F$  факторы  $X$  нәтижелік факторға елеулі түрде ықпал береді.

Егер қарастырылатын фактор біреу болса, онда оны бір факторлы дисперсиялық талдау деп атайды. Егер фактор бірнеше, әрі олардың бірнеше мән бар болса, онда оны көпфакторлы дисперсиялық талдау деп атайды.

Қоршаған орта әрдайым күрделі және көпөлшемді. Кейбір құбылыстардың тек бір айнымалымен сипатталуы өте сирек. Жалпы айтқанда ең басты қойылған мақсаттарға жету үшін көптеген анализ, бақылау жұмыстарын жүргізу қажет. «Статистика» қарапайым өмірде өте қажетті ұғым. Кейбір қиындық тудыратын жағдайларда статистикалық анализ жасау арқылы шешу әлдеқайда оңай болады.

### Әдебиет тізімі

1. Методы и алгоритмы финансовой математики : пер. с англ. / Ю.-Д. Люю . – М. : БИНОМ. Лаборатория знаний, 2007 – 751 бет.
2. Боровиков В.П., «Популярное введение в программу Statistica» - М.:Филинь, 2007-269 бет.
3. Айдынов З.П. «Мәліметтерді талдау және экономикадағы болжау»: оқу құралы - Алматы: «Бастау» баспасы, 2014-240 бет.
4. Сизова Т.М. Статистика: Учебное пособие. – СПб.: СПб ГУИТМО, 2005. – 80 с.

У.М. Дуйсенбай

*Университет «Астана», Нур-Султан, Қазақстан*

### Некоторые проблемы вычисления статистических данных

**Аннотация:** Статистическая информация является наиболее эффективным и удобным инструментом для пользователей, чтобы быстро получить продукты, которые они должны исследовать. В статье подчеркивается важность статистического анализа и практики для

получения эффективной и понятной информации и данных, а также способы их разработки.

**Ключевые слова:** статистическая информация, математическая статистика, анализ, ранговая корреляция.

U. M. Duisenbay

*University «Astana», Nur-Sultan, Kazakhstan*

### **Some problems in calculation of statistical data**

**Abstract:** Statistical information is the most effective and convenient tool for users to quickly get the products they need to research. The article emphasizes the importance of statistical analysis and practice for obtaining effective and understandable information and data, as well as ways to develop them.

**Key words:** statistical information, mathematical statistics, analysis, rank correlation.

### **References**

1. Liu Y.-D. *Metody i algoritmy finansovoy matematiki* [Methods and algorithms of financial mathematics] (Moscow, BINOM. Knowledge lab, 2007). [in Russian]
2. Borovikov V.P. *Populyarnoye vvedeniye v programmu Statistica* [Popular introduction to Statistica] (Moscow, Filin, 2007). [in Russian]
3. Aydinov Z.P. *Malimetterdi talday zhane ekonomikadagu bolzhay* [Data analysis and forecasting in the economy] (Almaty, Bastay, 2014). [in Kazakh]
4. Sizova T.M. *Statistika* [Statistics] (SPb, SPb GUITMO, 2005). [in Russian]