

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Київський національний університет будівництва і архітектури

Л.І. Турчанінова, О.В. Доля

СТАТИСТИКА ЯКОСТІ
Конспект лекцій

*Для студентів, які навчаються за напрямом підготовки
0305 «Економіка та підприємництво»
всіх форм навчання
спеціальності
8.03051001 «Товарознавство і комерційна діяльність»*

Київ - 2014

УДК 311.1:65.018
ББК 60.6:65.9
Т 89

Укладачі: Л.І. Турчанінова, канд. техн. наук, доцент
О.В. Доля, канд. фіз.-мат. наук, доцент

Рецензенти:

Теренчук С.А., канд. фіз.-мат. наук, доцент кафедри інформаційних технологій
проектування та прикладної математики

Відповідальний за випуск В.М. Михайленко, доктор техн. наук, професор

*Затверджено на засіданні кафедри інформаційних технологій проектування та
прикладної математики, протокол №15 від 11 червня 2014 року.*

Видається в авторській редакції.

Л.І. Турчанінова, О.В. Доля
В55 Статистика якості: Конспект лекцій. – К.: КНУБА, 2014. – 36 с.

Конспект лекцій розроблено у відповідності з програмою курсу «Статистика якості» з урахуванням сучасного стану в цій галузі діяльності. Значна увага приділена проблемам забезпечення й підвищення якості продукції. Матеріал відповідає п'яти лекціям робочої програми з дисципліни.

Конспект лекцій розрахований на студентів вищих навчальних закладів напряму підготовки 0305 «Економіка та підприємництво» спеціальності «Товарознавство і комерційна діяльність» всіх форм навчання.

ЗМІСТ

Загальні положення.....	4
Лекція №1. Основні поняття статистики якості	4
Лекція №2. Прості статистичні методи	9
Інструменти контролю якості.....	9
Діаграма Парето	9
Діаграми Ісікави	11
Діаграма розсіювання	12
Гістограма	13
Контрольні листки	13
Діаграма послідовності дій	14
Лекція №3. Контрольні карти.....	14
Лекція №4. Статистичний контроль якості продукції	18
Стандарти статистичного приймального контролю товарів.....	18
Лекція №5. Статистичні моделі (Моделювання якості)	21
Побудова кореляційно-регресійних моделей.....	22
Метод найменших квадратів в лінійних і нелінійних моделях регресій.....	23
Параболічна регресія.....	26
Геометрична регресія	27
Множинний регресійний аналіз.....	27
Список використаної і рекомендованої літератури.....	33
Додатки.....	34

Загальні положення

Якість відноситься до числа найважливіших показників діяльності людини. Якість продукції значною мірою визначає конкурентоспроможність підприємства, ріст ефективності виробництва. Проблеми забезпечення й підвищення якості продукції актуальні для всіх країн і підприємств.

Конспект лекцій призначений для студентів з метою підготовки фахівців, які грамотно й усвідомлено володіють питаннями статистики якості й пропонує до вивчення:

- якість продукції;
- сучасні тенденції вдосконалення якості;
- питання щодо вирішення проблем якості на сучасному етапі;
- взаємозв'язок якості та ефективності розвитку національної економіки;
- теоретичні основи статистики якості.

Зміст конспекту лекцій відповідає робочій програмі з дисципліни «Статистика якості» з урахуванням сучасного стану в цій галузі діяльності.

Лекція №1. Основні поняття статистики якості

Статистика – це наука, яка здійснює кількісну оцінку масових явищ та процесів у зв'язку з їх якісним змістом. Виходячи із визначення статистики як науки і враховуючи її спрямування на якість як головну категорію дослідження, можна дати такі визначення.

Статистика якості – розділ статистики, що розробляє принципи та методи кількісного оцінювання та аналізу якості товарів і послуг.

Об'єкт статистики якості – якість процесів, товарів та послуг.

Предмет статистики якості – застосування статистичних методів у процесі управління якістю товарів, робіт і послуг.

Статистика якості складається з таких основних розділів:

- статистичний контроль;
- статистичне управління процесами;
- планування експерименту.

За впливом на якість кінцевого продукту (товару або послуги) ці розділи умовно можна охарактеризувати так:

- *статистичний приймальний контроль* відіграє пасивну роль, оскільки лише фіксує відповідність створеного продукту встановленим стандартам або вимогам споживача;

- *статистичне управління процесами* є напівактивним засобом управління якістю, тому що дає можливість оцінювати і аналізувати не створену продукцію, а процеси її виробництва, отже, в такий спосіб, попереджує і зменшує витрати від створення продукції, що не відповідає певним вимогам. Вдосконалення якості здійснюється лише після впровадження рішень, прийнятих за результатами виконаного аналізу;

- *методи планування експерименту* є активним способом вдосконалення якості, адже вони дають змогу визначити оптимальні значення параметрів якості та процесів створення продукту.

З метою забезпечення виготовлення якісної продукції підприємство повинно використовувати якісну сировину або інші ресурси. В свою чергу, для реалізації споживачам має передаватися тільки якісна продукція, що відповідає вимогам. На практиці забезпечити 100-відсоткову перевірку всього обсягу ресурсів або готової продукції нераціонально або неможливо. Тому доцільно виконувати контроль продукції, який би ґрунтувався на принципах вибіркової перевірки та статистичного аналізу результатів, - **статистичний приймальний контроль**. Об'єктом дослідження у статистичному приймальному контролі є не перебіг створення певного продукту – процес, а його результат – готова продукція.

Партія продукту, що підлягає контролю, - це сукупність одиниць продукції одного найменування, яка виготовлена протягом певного періоду часу в однорідних умовах. Однорідність умов виготовлення є основною ознакою формування партії продукту, що не виключає відмінності між собою її окремих одиниць. Під **одиницею продукції** розуміють окремих екземпляр штучної продукції або визначену в установленому порядку кількість не штучної або штучної продукції. Окрему одиницю можна характеризувати за кількісними або якісними ознаками. Залежно від типу ознак, за якими здійснюється характеристика одиниць продукції, розрізняють статистичний приймальний контроль за кількісними або якісними ознаками.

Контроль за якісними ознаками – це вид контролю, в ході якого визначають значення параметра, що вимірюється за інтервальною шкалою ознак, і рішення стосовно контрольованої сукупності приймають залежно від результатів порівняння отриманих значень параметра з контрольними нормативами.

Контроль за кількісними ознаками – це вид контролю, за якого кожен перевірену одиницю продукції відносять або до категорії тих, що відповідають певним вимогам, або до бракованих, і рішення стосовно контрольованої партії приймається залежно від результатів порівняння кількості визначених у вибірці дефектних одиниць або числа дефектів, що припадає на певну кількість одиниць продукції, із контрольними нормативними величинами.

Статистичний приймальний контроль здійснюється визначеними планами контролю.

План контролю – це сукупність вимог і правил, за якими приймається рішення про прийом чи бракування партії продукції.

Під **контрольним нормативом** розуміють мінімальне або максимальне значення, яке встановлене у відповідній нормативній документації і є критерієм для прийняття рішення щодо відповідності продукції встановленим вимогам за результатами вибіркового контролю.

До основних контрольних нормативів належать такі: рівень дефектності, приймальне число, бракувальне число, ризик постачальника, ризик споживача.

Рівень дефектності – це частка дефектних (таких, що не відповідають вимогам) одиниць партії продукції або кількість дефектів одиниць на сто одиниць. Рівень дефектності може бути охарактеризований залежно від місця аналізу або часу. Так, залежно від місця контролю розрізняють вхідний та вихідний рівні дефектності.

Вхідний рівень дефектності – це кількість дефектів у партії, яка надійшла на контроль, або в потоці продукції за певний період часу.

Вихідний рівень дефектності – це кількість дефектів у партії або потоці продукції за певний період часу після здійснення контролю (відбору певної кількості одиниць на контроль).

Значення середньої величини з часток дефектних одиниць у кількох партіях продукції має назву середнього рівня дефектності.

Залежно від впливу на результати прийняття рішення про відповідність всієї партії продукції встановленим вимогам розрізняють приймальний і бракувальний рівні дефектності.

Приймальний рівень дефектності (AQL) – це найбільший рівень дефектності для окремих партій або середній рівень для послідовності партій, який є задовільним для визначення продукції такого, що відповідає встановленим вимогам. Він визначається як відсоток дефектних одиниць у загальній сукупності (окремій партії): $AQL = \frac{d}{n} \cdot 100$, де d - кількість дефектних одиниць; n – кількість перевірених одиниць.

Бракувальний рівень дефектності – це мінімальний рівень дефектності для окремої партії, який є допустимим для визначення продукції такою, що не відповідає певним вимогам. Для послідовності партій він не встановлюється.

Контрольний норматив «**приймальне число**» - це максимальна кількість дефектних одиниць у вибірці, яка є критерієм для прийняття партії продукції.

Протилежним до нього є норматив «**бракувальне число**» - мінімальна кількість дефектних одиниць у вибірці, яка є критерієм для відхилення приймання партії продукції.

Застосування статистичного приймального контролю подібне до перевірки статистичних гіпотез щодо відповідності параметрів генеральної сукупності визначеним значенням. Таким чином, при застосуванні процедур статистичного приймального контролю виникають певні ризики. Основними з них є ризик постачальника і ризик споживача.

Ризик постачальника – це ймовірність бракування партії продукції, яка має приймальний рівень дефектності, тобто переважно складається з виробів, що відповідають певним вимогам.

Ризик споживача – це ймовірність приймання партії продукції з бракувальним рівнем дефектності, тобто такої, яка складається з одиниць продукції, що не відповідають певним вимогам.

Статистичний приймальний контроль здійснюється за певною схемою. Схема СПК – це комплект планів вибіркового контролю, який охоплює переважну частину можливих випадків, і сукупність правил використання цих планів на практиці.

Залежно від рівня вимог, які висуваються до надійності отриманих результатів контролю, розрізняють різні **види контролю**: нормальний, послаблений та посилений.

Основним видом контролю є нормальний, він використовується в усіх випадках доти, доки не виникає потреба перейти на інші види контролю. Перехід від одного виду контролю до іншого здійснюється за такими правилами:

- від нормального до посиленого, якщо дві з п'яти послідовних партій буде визнано такими, що не відповідають встановленим вимогам;
- від нормального до послабленого, коли за умови стабільності процесу виробництва було прийнято не менше 10 послідовних партій, при цьому загальна кількість одиниць продукції, що не відповідають вимогам, не перевищує граничну кількість, встановлену для прийнятого рівня дефектності;
- від послабленого до нормального, якщо процеси не є статистично стабільними або відбулися зміни в технології, або наступна партія визнана такою, що не відповідає вимогам, при першій перевірці;
- від посиленого до нормального, коли під час посиленого контролю приймається п'ять послідовних партій.

Залежно від обсягів партій продукції та можливо допустимих вибірок розрізняють **рівні контролю**. Найпоширенішим є 7 рівнів контролю – 3 загальні (I-III) і 4 спеціальні (S-1 – S-4).

Основним є II рівень контролю. Рівень III використовують тоді, коли застосування рівнів I і II є недоцільним з точки зору вартості контролю. Спеціальні рівні застосовуються, якщо необхідно використати вибірки малого обсягу (наприклад, у разі руйнівних випробувань).

Ймовірнісною характеристикою ефективності застосування плану контролю є **операційна характеристика**. Це виражена у певний спосіб залежність ймовірності прийняття партії продукції від рівня якості цієї партії. Операційна характеристика може бути подана у табличній, аналітичній або графічній формі.

Закон розподілу Пуассона використовується для одержання операційних характеристик планів контролю, в яких AQL перевищує 10 і визначається кількість дефектів на 100 одиниць продукції.

Біноміальний закон розподілу використовується для визначення операційних характеристик планів контролю, в яких AQL менше 10 і обсяг вибірок не перевищує 80 одиниць. За їх допомогою контролюється частка дефектних одиниць продукції.

Кожний план контролю за альтернативною ознакою визначається низкою величин. До них належать такі:

- приймальний рівень (AQL);
- бракувальний рівень (LQ);
- середній вихідний рівень дефектності (AOQ);
- граничне значення середнього вихідного рівня дефектності (AOQL);
- приймальне число A_c ;
- бракувальне число R_e .

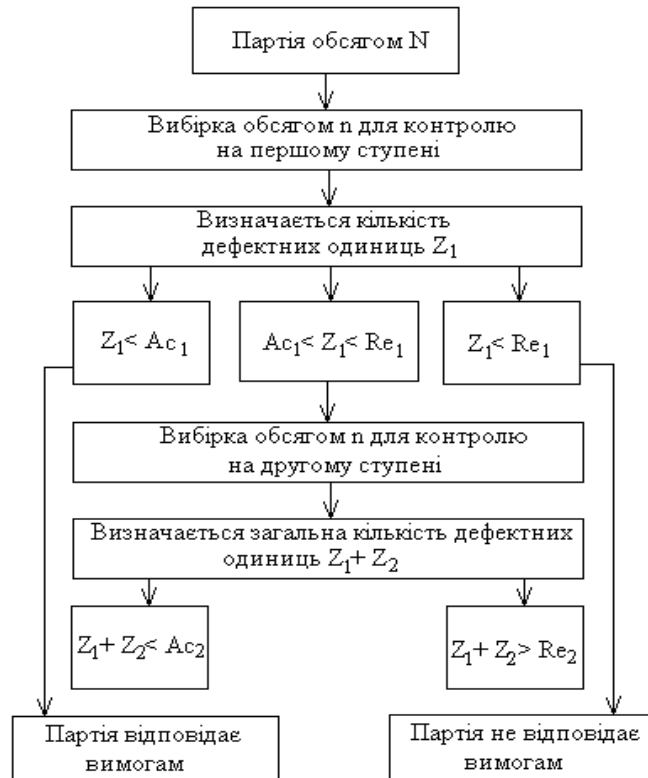


Рис. 1. Схема двоступеневого приймального контролю

Питання для самоконтролю

1. Дайте визначення ймовірності певної події.
2. Розкрийте особливості біноміального закону розподілу.
3. Розкрийте особливості закону розподілу Пуассона.
4. Які основні властивості нормального закону розподілу і в чому полягає його значення для практичного застосування статистики якості?
5. Які закони розподілу є похідними від нормального і де вони використовуються?
6. Розкрийте сутність статистичного приймального контролю. Поясніть його значення для забезпечення якості продукції.
7. Що таке план контролю? Які його основні властивості?
8. Розкрийте поняття видів контролю.
9. Яким є порядок застосування планів контролю?

Лекція №2. Прості статистичні методи

Інструменти контролю якості

Статистичні методи управління якістю направлені на аналіз, контроль та регулювання якості продукції та технологічних процесів.

При виробництві продукції виробник документує дані процедур запобігання невідповідності. Це необхідно йому для вирішення спірних питань щодо якості продукції та в разі спричинення шкоди споживачу.

Для ефективного виявлення і аналізу проблем, які виникають при виробництві невідповідної продукції використовують сім інструментів контролю якості:

- діаграма Парето;
- діаграма Ісікави;
- контрольні карти;
- гістограма;
- діаграма розсіювання;
- контрольний листок;
- діаграма послідовності дій.

Ці методи становлять основу проведення контролю якості на будь-якому промисловому підприємстві.

Діаграма Парето

Якщо поставлено завдання знизити брак виробу, то в кожному цеху будуть вести роботу зі зниження браку цього виробу й неминуче встане питання, як можна знизити брак у даній бригаді і які роботи треба вести в першу чергу. Якщо виникає питання про зниження числа операцій у процесі виробництва виробу, то кожна бригада на своїй виробничій ділянці або бюро контролю якості даної ділянки насамперед повинні вирішити, які саме операції варто раціоналізувати в першу чергу. Щоб виявити найбільш важливе проблемне питання, треба зрівняти безліч факторів. Діаграма Парето в цьому випадку є найбільш придатним методом. Вони дають можливість зрівняти фактори і побачити їх послідовність по ступені важливості.

Діаграми італійського економіста Парето (1845-1923) є найпоширенішим у Японії методом контролю якості в процесі виробництва, що вважається простим і ефективним.

Нижче перераховані деякі із проблем, які японські фахівці й рядові виконавці аналізують за допомогою діаграм Парето.

Якість:

- число випадків браку за його видами, суми втрат від браку;
- витрати часу й матеріальних засобів на виправлення браку;
- зміст рекламацій, що надходять від споживачів;
- кількість випадків поломок;
- витрати, пов'язані з рекламаціями.

Тимчасові фактори:

- причини затримки виготовлення виробів, число пристроїв, серйозність і масштаби аварій, поломок і т. ін.;
- розділ часу й операції;
- час простою через поломки й аварії;
- втрати у зв'язку зі скороченням обсягу виробництва.

Собівартість:

- суму по окремих статтях виробничого кошторису;
- витрати на виробництво (сировина й матеріали, витрати праці, допоміжні матеріали, грошові витрати, витрати на управління і т.д.);
- аналіз звітно-фінансової документації.

Безпека праці:

- аналіз готових причин пожеж і інших нещасних випадків (на японських підприємствах відомості про нещасні випадки по цехах подаються у вигляді діаграм Парето).

Попит:

- аналіз попиту на різні види продукції.

Всі види даних виражають у формі діаграм Парето й математичних таблиць. Оскільки дані збирають на виробництві з метою використання їх для контролю, поліпшення робіт і оцінки їхнього стану, то важливо, щоб контролери, керуючі, керівники виробничих ділянок могли на основі цих даних оперативно винести правильне рішення й вжити відповідних заходів.

На основі даних, наведених у табл. 2.1, креслять діаграму Парето. Для цього на осі абсцис записують види браку, на осі ординат – число випадків браку в числовому або відсотковому значенні (рис. 2.1) і креслять стовпчикову діаграму. Потім креслять кумулятивну криву, як це показано на рис. 2.2.

Таблиця 2.1

Види браку

Види браку	Число випадків браку	Сумарне число випадків браку за розглянутий період	Відсоткове співвідношення	Кумулятивний відсоток числа випадків браку
Брак сировини й матеріалів	48	48	41,7	41,7
Відхилення розмірів	32	80	27,8	69,5
Брак у формі	23	103	20,0	89,5
Брак в обробці	4	107	3,5	93,0
Інші види браку	8	115	7,0	100,
Разом:	115		100	

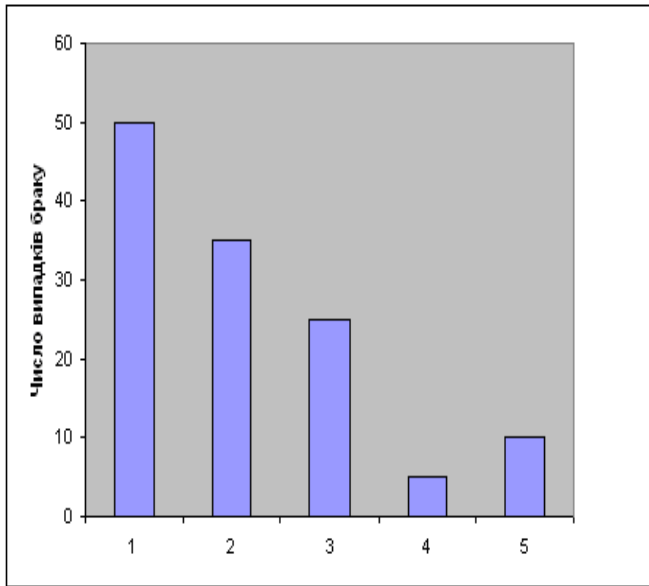


Рис. 2.1 Діаграма Парето за видами браку 1-брак матеріалів; 2-брак у розмірах; 3-брак у формі; 4-брак в обробці; 5-інші види браку.

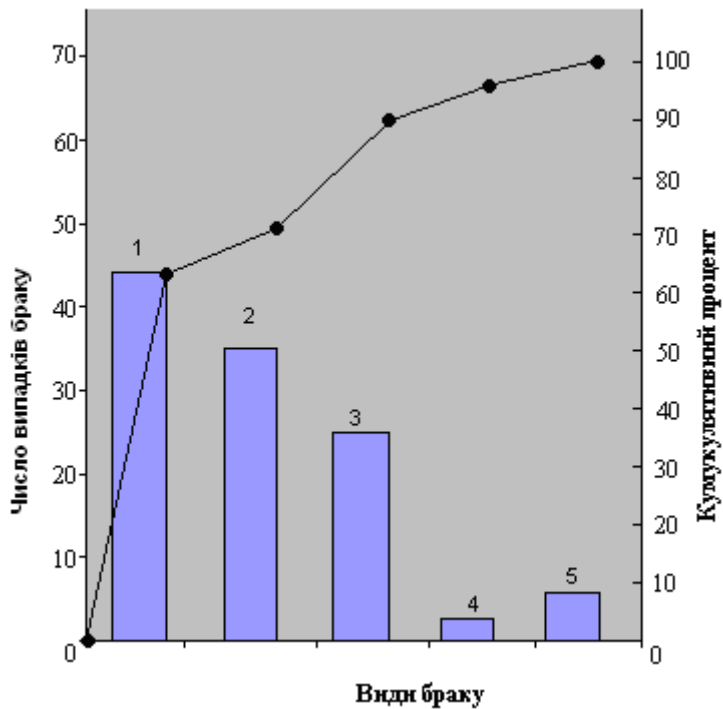


Рис. 2.2 Кумулятивна крива по видам браку 1-брак матеріалів; 2-брак у розмірах; 3-брак у формі; 4-брак в обробці; 5-інші види браку.

Якщо розглянути тепер складену діаграму Парето, легко зрозуміти розподіл основних видів браку. Очевидно, що брак матеріалів і розмірів становить 70%, а разом зі браком форми це 90% усього браку. При проведенні робіт зі зниження браку в цьому випадку раціонально поставити завдання зниження браку матеріалів на 20%, а браку розмірів і форми на 5%.

Діаграми Ісікави

При розгляді й аналізі питань забезпечення якості продукції й інших поточних питань виробництва на японських підприємствах широке

поширення одержав метод Ісікави. На діаграми часто посилаються також у доповідях і виступах на з'їздах і конференціях з питань контролю якості.

Професор Токійського університету, один із провідних японських спеціалістів в області забезпечення якості, Ісікава Каору в 1950 р. запропонував метод об'єднання різних факторів, що роблять вплив на кінцевий результат якогось процесу, і їхньої систематизації на діаграмі. Ці діаграми називають також схемами «причини-результат», а також через їхню специфічну форму – «риб'ячий кістяк», «риб'яча кістка» (рис. 2.3).

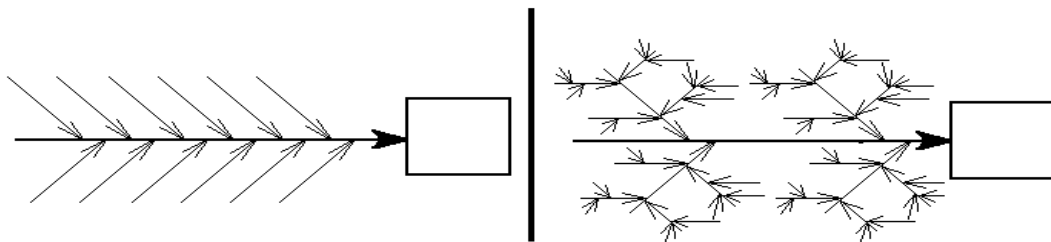


Рис. 2.3. Види схем Ісікави

Метод знайшов визначення не тільки в Японії. Зацікавленість до нього проявляють фахівці США й країн Західної Європи.

Як видно із рис. 2.4, за допомогою схеми Ісікави можна наочно показати зв'язок характеристики й факторів, які впливають на неї, і взаємозв'язок між самими факторами. Діаграми Ісікави дають можливість виявити такі одночасно діючі на якість продукції фактори, як робота виконавця, матеріали, верстати й устаткування, технологічні методи і т. ін.

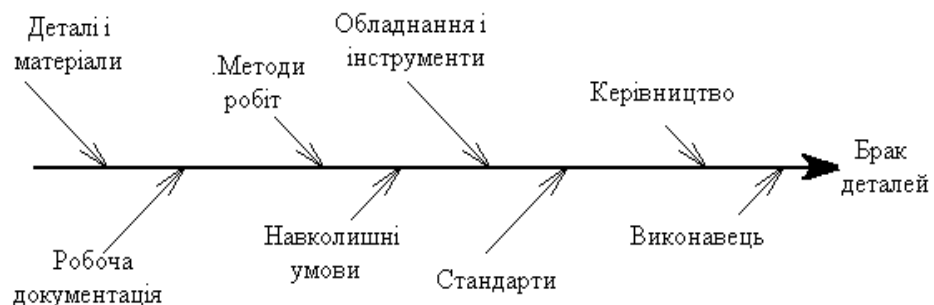


Рис. 2.4. Діаграма Ісікава, що розкриває різні фактори

Діаграма розсіювання

Для вивчення залежності між двома змінними показниками можна скористатися так званою, діаграмою розсіювання. По ній можна, використовуючи кореляційний і регресивний аналізи, виявити кількісний зв'язок між двома параметрами. Діаграма дозволяє наочно показати характер змін, параметрів якості в часі з урахуванням впливу різних факторів.

Якщо y – показник якості, x – фактор, що впливає на якість, то рис. 2.11 чітко простежується пряма кореляція (залежність). Зворотна кореляція спостерігається в тому випадку, коли при збільшенні значення x показник y зменшується.

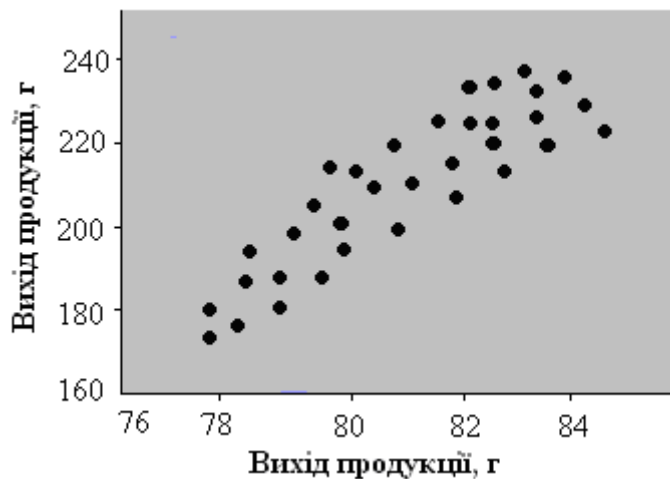


Рис. 2.11 Залежність виходу продукції від чистоти сировини (по 29 вимірам)

Однак на графіку розкидання точок значне і нагадує «пряму», то виражена залежність між параметрами x і y відсутня.

Кореляція – це метод аналізу зв'язків шляхом порівняння двох взаємозалежних величин x і y . У тих випадках, фактор x є безперервною випадковою й відповідально до зміни x безупинно міняється y , розташування по x неможливо, і тоді важливо безупинно вивчати взаємодія x і y . Із цією метою застосовують графіки розсіювання.

Гістограма

Гістограма – це метод спостереження згрупованих за частотами даних, який при простій обробці дає багато корисної інформації про розкид якості, середніх значеннях різних характеристик, про точність ходу технологічного процесу й точність роботи верстатів і устаткування.

Гістограма являє собою один з видів стовпчикової діаграми.

Гістограми знаходять широке застосування при складанні місячних звітів цехів і заводів про якість виробленої продукції, звітів про результати технічного контролю, при показі зміни рівня якості по місяцях і інших випадках.

Гістограми застосовують при навчанні виконавців методам контролю якості.

Контрольні листки

Контрольні листки – аркуші збору даних, використовуються при проведенні поточного контролю з метою розуміння виду розподілу показників якості, перевірки браку й обстеження порушень і неполадок у технологічному процесі й т.д. Форма листків розробляється відповідно до ситуації на місцях і з метою збору даних з урахуванням конкретної мети і завдання, що поставлені перед підприємством.

Як відомо, на виробництві часто буває необхідно знати вид розподілу таких показників якості, як твердість матеріалів, розміри деталей, склад хімічних продуктів. Для розуміння розподілу цих показників застосовують гістограми. Це звичайний спосіб, за допомогою якого одержують форму

розподілу, середні показники, розкид. Для складання гістограми збирають дані й обробляють їх. Використання контрольних листків спрощує збір і обробку даних, і складання гістограм.

Діаграма послідовності дій

Діаграма потоку (дій) дає можливість зобразити складні процеси з різними компетентністю й завданнями таким чином, щоб їхня структура й логіка були прозорі й зрозуміли. Діаграма дозволяє виконавцям повніше й наочніше довідатися перелік своїх завдань, у той час як письмове формулювання всього ходу подій охопити складніше.

Зразок заповнення діаграми послідовності дій:

1. Встановити невідповідну продукцію.
2. Позначити продукцію.
3. Покласти на ізолюваний склад.
4. Провести випробування.
5. Брак?
6. Зняти турботу.
7. Доробка можлива?
8. Провести доробку.
9. Провести повторне випробування.
10. Результати випробувань позитивні?
11. Спеціальний дозвіл.
12. Передати продукцію на відвантаження.
13. Включити витрати.

Питання для самоконтролю

1. Які методи становлять основу проведення контролю якості на підприємстві?
2. Які проблеми якості аналізують за допомогою діаграм Парето?
3. Який порядок складання схеми Ісікави?
4. Як будується діаграма розсіювання?
5. Як застосовується метод гістограм в роботі підприємства?
6. Де використовують контрольні листки?
7. Як заповнюють діаграму послідовності дій?

Лекція №3. Контрольні карти

Методологія контрольних карт, яка розвивається з 30-х років ХХ ст., стала основним засобом статистичного управління процесами. Завдяки можливості високого ступеня формалізації контрольні карти стали популярним засобом управління якістю в усьому світі. Так, в Японії, разом із діаграмою процесів, діаграмою «причин і наслідків», Парето-діаграмою, контрольним листком, гістограмою і діаграмою розсіювання вони становлять *«7 простих інструментів вдосконалення якості»*. Як зазначає японський

науковець Х. Куме, із застосування контрольних карт починається і ними ж завершується цикл із вдосконалення якості. Це означає, що контрольні карти є як засобом, з використанням якого починається вдосконалення якості, так і засобом підтвердження досягнутих результатів.

Контрольна карта являє собою графічний засіб оцінки певної ознаки якості, виміряні значення якої наносяться на графік відповідно до порядку отримання у часі (рис. 3.1). Графік контрольної карти складається з вертикальної осі, на яку наносять масштаб досліджуваної ознаки якості, та горизонтальної осі, яка характеризує послідовність отриманих даних. На графіку проводять центральну лінію, яка відповідає середньому значенню ознаки якості, і дві лінії, що називається контрольними межами – верхня (ВКМ) та нижня (НКМ). У зарубіжній літературі найчастіше вживається позначення UCL для верхньої та LCL для нижньої межі. Ці скорочення походять від термінів *Upper Control Limit* та *Lower Control Limit* відповідно. Можливі варіанти перекладів – верхня або нижня межа регулювання або межа управління.

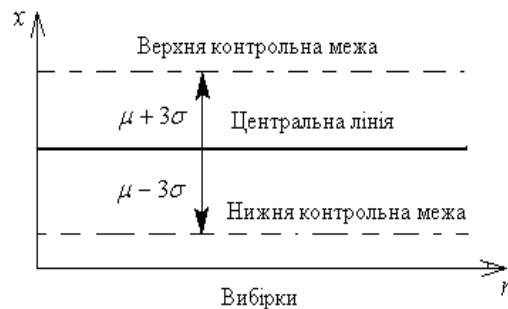


Рис. 3.1 Схема контрольної карти Шугарта

Потрібно особливо зазначити, що контрольні межі в жодній мірі не співвідносяться з межами допуску, які встановлюються при проектуванні на виготовлення продукції. Це зовсім інші характеристики, і недоцільно наносити їх на один графік. У найгіршому випадку така помилка спричинить руйнування системи статичного управління процесами, в кращому – ускладнить інтерпретацію контрольних карт. Процес може перебувати під впливом особливих причин, тобто мати непередбачувану поведінку, і при цьому бути в межах допуску. Рано чи пізно такий процес призведе до створення продукції, що не відповідатиме встановленим вимогам, якщо дію особливих причин не буде усунуто. Коли ж процес стабільний, можна очікувати, що він матиме такі самі параметри варіації і в майбутньому, і навіть якщо в поточний час створюється продукція, яка не відповідає вимогам допусків, завдяки певним діям можна зробити так, щоб виготовлена продукція відповідала вимогам.

Контрольні межі вибираються у такий спосіб: поки на процес впливають лише випадкові причини, майже всі вибіркові значення характеристики ознаки якості, які наносять на карту, потрапляють в інтервал між ними. поки точки вибіркового спостереження розміщуються в контрольних межах, вважається, що процес перебуває в стані статистичного управління або є стабільним. У тому разі, якщо точка яка відповідає певному

вибірковому спостереженню, розміщується поза контрольними межами, є підстави вважати, що процес вийшов зі стану стабільності. У цьому разі слід вжити коригувальні дії і здійснити поглиблений аналіз ситуації з метою визначення причин та їх подальшого усунення.

Окремі вчені розглядають контрольну карту як послідовність статистичних перевірок гіпотез. При цьому послідовно (кожним наступним значенням ознаки якості) перевіряється гіпотеза про рівність вибіркової статистики певному гіпотетичному значенню, тобто:

$$H_0 : \mu_1 = \mu_0,$$

$$H_1 : \mu_1 \neq \mu_0.$$

У цьому разі існування точки в контрольних межах не дає підстав відкинути гіпотезу про стабільність процесу, тоді як наявність точки за контрольними межами дає підстави прийняти альтернативну гіпотезу про вихід процесу зі стану керованості. Але між методологією контрольних карт і перевіркою гіпотез існують певні відмінності. Основна полягає в тому, що перевірка гіпотез прийнята лише для ситуації, за якої середнє змінюється один раз і після такої зміни залишиться стабільним на новому рівні, тоді як у реальних випадках можливі складніші ситуації.

Загальну модель контрольної карти можна описати так. Припустимо, що W – вибіркова величина, яка вимірює певну ознаку якості, і середнім цієї величини $W \in M_W$, а середньоквадратичним відхиленням - δ_W . Тоді центральну лінію, верхню та нижню контрольні межі контрольної карти можна подати так:

$$\text{ВКМ} = W + L \delta_W,$$

$$\text{ЦЛ} = M_W,$$

$$\text{НКМ} = W - L \delta_W.$$

де L – відстань між контрольними межами і центральною лінією.

В. Шугарт запропонував вимірювати відстань від контрольних меж до центральної лінії контрольної карти у середньоквадратичних відхиленнях вибіркової статистики W . Виходячи з емпіричних припущень, він встановив значення $L=3$ (рис. 3.2).

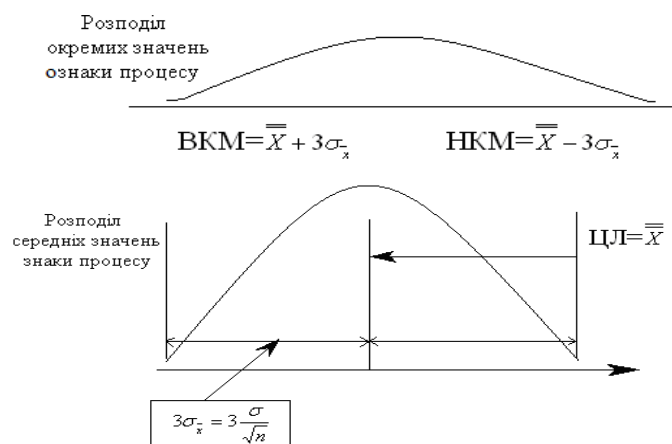


Рис. 3.2. Співвідношення між розподілом окремих значень, середніх значень ознаки якості і межами контрольної карти

Контрольні карти розроблені відповідно до цього принципу, на честь їх винахідника називають контрольними картами Шугарта.

Незважаючи на закон розподілу досліджуваної ознаки якості, в Сполучених Штатах Америки та Японії використовують множник, що дорівнює 3, тоді як у Німеччині та Великобританії найчастіше використовують ймовірнісні контрольні межі і вибирають їх у такий спосіб, щоб ймовірність появи точки за контрольними межами за стабільного процесу дорівнювала 0,001 або 0,027. Іноді застосовують контрольні межі за рівня ймовірності 0,005. Прихильники останнього підходу вважають, що контрольні карти Шугарта працюють лише за двох основних припущень:

- спостереження, які отримані у часі, незалежне одне від одного;
- досліджувана ознака має нормальний розподіл.

Інша група дослідників вважає, що контрольні карти Шугарта довели свою ефективність і при автокореляції даних, тобто коли результати спостережень впливають на наступні дані, а також при таких характеристиках розподілу ознаки якості, що не відповідають нормальному закону розподілу.

Контрольні карти Шугарта є здебільшого картами двох типів: контрольні карти кількісних змінних і контрольні карти альтернативних ознак. Для кожної з контрольних карт існують дві відмінні ситуації:

- коли стандартні значення не задано;
- коли стандартні значення задано.

Стандартними значеннями є певні спеціальні вимоги або цільові значення (рис. 3.3).



Рис. 3.3. Форми контрольних карт

Існують три основні типи контрольних карт:

- контрольні карти Шугарта;
- контрольні карти приймального контролю;
- адаптивні контрольні карти.

Контрольні карти Шугарта здебільшого застосовують для того, щоб оцінити наявність стану статистичного контролю, хоча карти цієї категорії часто використовують як засіб приймального контролю процесу, навіть якщо вони не розроблені саме щодо використання критерію або меж допусків технологічного процесу.

Приймальні карти призначені для встановлення відповідності виготовленої продукції існуючим вимогам.

Адаптивні контрольні карти використовуються для того, щоб регулювати технологічний процес за допомогою передбачення тенденцій його розвитку і вироблення заздалегідь налагоджень на основі передбачень.

Залежно від статистичних показників, на основі яких контролюють контрольні карти Шугарта, розглядають такі їх види:

1. Контрольні карти для кількісних змінних:

- карта середніх \bar{X} та розмахів варіації R ;
- карта середніх \bar{X} та стандартних відхилень s ;
- карта окремих значень X та ковзних розмахів MR ;
- карта медіан Me і розмахів R .

2. Контрольні карти для альтернативних змінних:

- карта частки дефектів (p -карта) або карта кількості дефектних виробів (np -карта);
- карта кількості дефектів на виріб (c -карта) або карта середньої кількості дефектів на одиницю (u -карта).

Питання для самоконтролю

1. На які типи поділяють контрольні карти?
2. Що таке контрольні карти Шугарта? Які їх види вам відомі?
3. З якою метою застосовують контрольні карти на виробництві?

Лекція №4. Статистичний контроль якості продукції

Якість продукції тлумачиться як «сукупність характеристик продукції (процесу, послуги), які стосуються її здатності задовольняти встановлені і передбачені потреби».

На практиці неможливо забезпечити перевірку характеристик кожної одиниці готової для реалізації продукції. Тому роблять **вибірковий контроль** та статистичний аналіз результатів, так званий – **статистичний приймальний контроль (СПК)**.

Під готовою продукцією можна розуміти сировину, матеріал, напівфабрикати, тобто будь-які результати діяльності на попередніх ланках створення продукції, все те, що являє собою певну сукупність і відносно чого потрібно прийняти рішення про відповідність визначеним вимогам.

Стандарти статистичного приймального контролю товарів

Для успішного застосування статистичних методів контролю якості продукції необхідні стандарти, доступні широкому колу інженерно-технічних працівників. Стандарти та статистичний приймальний контроль забезпечують можливість об'єктивно порівнювати рівні якості партій однотипної продукції як у часі, так і по різних підприємствах. Сформулюємо деякі вимоги до стандартів по статистичному контролю.

Стандарт повинен містити досить велику кількість планів, що мають різні оперативні характеристики. Це дозволить вибирати плани контролю з урахуванням особливостей виробництва й вимог споживача до якості продукції. Бажано, щоб у стандарті були зазначені різні типи планів: одноступінчасті, двоступінчасті, плани послідовного контролю й т.д.

При приймальному контролі якості партій готові продукції, сировини й напівфабрикатів важливі не стільки результати контролю окремої партії, скільки результат контролю послідовності партій виробів. Тому в стандартах повинна бути представлена система правил, що вказує, який конкретно план контролю з безлічі наявних у стандарті варто використовувати для контролю контрольної партії. Бажано враховувати результати контролю попередніх партій. При розкладанні технологічного процесу, коли на контроль надходять партії продукції з підвищеним утримуванням дефектних виробів, споживач може наполягати на використанні планів контролю, що забезпечують задане значення межі середнього рівня вихідної якості. У цьому зв'язку стандарт повинен містити правила переходу з нормального контролю на посилений і навпаки. У тих випадках, коли якість продукції досить висока, можна використовувати плани полегшеного контролю.

Таким чином, основними елементами стандартів по приймальному контролі є:

- 1) таблиці планів вибіркового контролю, використовувані в умовах нормального ходу виробництва, а також планів для посиленого контролю в умовах розладнань і для полегшеного контролю при досягненні високої якості;
- 2) правила вибору планів з урахуванням особливостей контролю;
- 3) правила переходу з нормального контролю на посилений або полегшений і зворотний перехід при нормальному ході виробництва;
- 4) методи обчислення наступних оцінок показників якості контрольованого процесу.

У ряді сучасних стандартів плани посиленого й полегшеного контролю пов'язані з таблицями планів нормального контролю простим співвідношенням. Наприклад, при одноступінчастому контролі перехід від нормального до посиленого зводиться до зменшення приймального числа з без зміни обсягу вибірки n . Якщо ж $z=0$, то збільшується тільки обсяг вибірки.

У тих випадках, коли відхилені партії піддаються суцільному контролю, план вибирають із урахуванням середнього числа виробів n_{cp} , контрольованих у партії при нормальному ході виробництва. Розрахунок n_{cp} провадять у припущенні, що частка дефектних виробів у партіях постійна й дорівнює \bar{q} .

Залежно від гарантій, забезпечуваних планами приймального контролю, існують різні методи побудови планів:

- установлюють значення ризику постачальника α і ризику споживача β і висувають вимогу, щоб оперативна характеристика $P(q)$ пройшла приблизно через дві точки: q_0, α і q_m, β , де q_0 і q_m – відповідно прийнятний і бракувальний рівні якості. Цей план можна назвати компромісним, тому що він забезпечує захист інтересів як споживача, так і постачальника;
- вибирають одну крапку оперативної характеристики й приймають одне або кілька додаткових незалежних умов (наприклад, q_L або мінімальний обсяг контролю n_{cp} при заданому значенні q).

Перша система планів статистичного приймального контролю, що знайшла широке застосування в промисловості, була розроблена Доджем і Ромінгом. Плани цієї системи передбачають суцільний контроль виробів із забракованих партій і заміну дефектних виробів придатними. Запропоновано таблиці одноступінчастих і двоступінчастих планів вибіркового контролю двох видів. Одні плани забезпечують ризик споживача $\beta=0,1$ при вісьмох значеннях бракувального рівня якості q_m (0,5; 1,0; 2,0; 3,0; 4,0; 5,0; 7,0; 10,0 %), а інші гарантують значення q_L межі середнього рівня вихідної якості. Крім того, плани Доджа-Ромінга забезпечують мінімальний обсяг контрольних операцій при середній частці дефектних виробів у продукції, що надходять на контроль m . Значення m визначається за результатами попереднього контролю. Умова мінімуму обсягу при $q_m = q_0$ є другим критерієм вибору конкурентного плану.

В Україні застосовуються наступні стандарти по статистичному контролю:

ДСТУ 3021-95. Випробування та контроль якості продукції. Терміни та визначення.

ГОСТ 16504-81. Випробування і контроль якості продукції. Основні терміни і визначення.

ГОСТ 18321-73. Статистичний контроль якості. Методи випадкового відбору вибірок поштучної продукції.

ГОСТ 24297-87. Вхідний контроль продукції. Основні положення.

Інструкція. «Приймання продукції по кількості і якості».

Тести для перевірки засвоєння навчального матеріалу

1. Статистична оцінка рівня якості продукції здійснюється на стадії:
 - а) експлуатація або споживання продукції;
 - б) *виробництва продукції*;
 - в) маркетингу та вивчення ринку;
 - г) проектування і розроблення продукції.
2. Оцінювання рівня якості виготовлення продукції за показниками ефективності здійснюється на стадії:
 - а) експлуатації або споживання продукції;
 - б) *виробництва продукції*;

- в) маркетингу та вивчення ринку;
- г) проектування і розроблення продукції.

3. Оцінювання рівня виготовлення продукції за показниками ефективності здійснюється на стадії:

- а) експлуатації або споживання продукції;
- б) *виробництва продукції*;
- в) маркетингу та вивчення ринку;
- г) проектування і розроблення продукції.

4. Встановлення способу збирання і отримання інформації про рівень якості здійснюється на стадії:

- а) *експлуатації або споживання продукції*;
- б) виробництва продукції;
- в) маркетингу та вивчення ринку;
- г) проектування і розроблення продукції.

5. Статистичний контроль якості на підприємстві передбачає:

- а) розподіл функцій і відповідальність за якість як між окремими працівниками, так і цеховим керівником або майстром;
- б) *використання контрольних карт*;
- в) наявність відділу управління якістю у складі організаційної структури підприємства;
- г) передбачає участь у роботах з якості всього персоналу підприємства.

Лекція №5. Статистичні моделі (Моделювання якості)

Під моделюванням якості продукції розуміють «такі напрямки виконання функції загального управління, які визначають політику, цілі і відповідальність у сфері якості, а також здійснюють їх за допомогою таких засобів, як планування якості, оперативне управління якістю, забезпечення якості та поліпшення якості в межах системи якості».

Основні принципи загального менеджменту якості (TQM) закладено у новій версії стандартів ISO 9001-2000. Відповідно до цих стандартів для ефективного досягнення результатів роботи організації слід управляти ресурсами і діяльністю як процесами. Система менеджменту якості орієнтована на безперервне вдосконалення і є метою організації. Цей підхід проілюстровано на рис. 5.1.

З цієї концептуальної моделі випливає висновок, що споживач відіграє основну роль у системі менеджменту якості, він формує вимоги до продукту, що створюється організацією, і визначає ступінь відповідності продукту своїм вимогам.

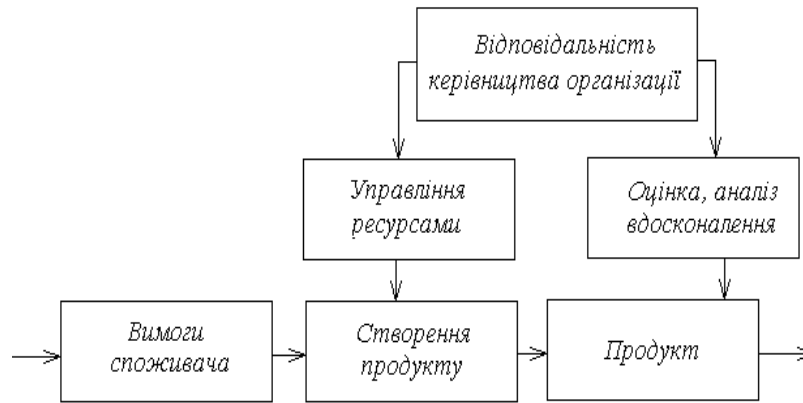


Рис. 5.1.

Побудова кореляційно-регресійних моделей

Для ефективного вдосконалення якості важливе значення має визначення наявності залежності певної ознаки якості від впливу різноманітних чинників, а також кількісна характеристика ступеня і виду цієї залежності.

Така залежність може бути **визначеною** (функціональною), але частіше на результативну ознаку різноманітні чинники впливають випадковим чином. В цьому випадку має місце стохастичний зв'язок, який визначається і характеризується за допомогою статистичних методів.

В математичній статистиці регресійний аналіз є одним з найбільш поширених методів обробки економічних даних при вивченні характеру залежностей між величинами. Термін «регресія» вперше був застосований для визначення залежності між двома величинами у останній треті XIX ст. англійським дослідником Френсісом Гальтоном. Він вивчав залежність зросту сина від зросту батька і встановив, що зріст сина завжди «регресує», тобто прагне повернутися до середньої величини.

В статистичній моделі припускається, що спостереження (або дані) є реалізацією (експериментом) деякої випадкової величини, що залежить від кількох інших не випадкових величин, значення яких відомі. Величини, або змінні, значення яких відомі і (або) визначаються умовами експерименту, називаються **факторами**. Величини, які вимірюються під час експерименту, називається **відгуками**. Наприклад, розглядається питання як якість продукції залежить від рівня освіти працівника, його статі, віку, місця проживання, сфери діяльності. Факторами в цій задачі будуть: рівень освіти (середня, середня спеціальна, вища), стать (чоловік чи жінка), вік (у роках), місце проживання (у місті чи сільській місцевості), галузь господарства (промисловість, сфера торгівлі і т. ін.). А відгуком буде якість кінцевого продукту.

В тому випадку, коли факторів два і більше задача називається **багатофакторною**, якщо ж фактор один – то експеримент називають **парним**.

Нехай вибіркові спостереження складаються з n пар вимірювань (x_i, y_i) , $i=1 \div n$. Мірою щільності зв'язку між фактором X і результатом Y буде статистика

$$r = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{n}{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \frac{n}{\sum_{i=1}^n (y_i - \bar{y})^2}}} \quad (5.1)$$

Статистика r називається **коефіцієнтом кореляції**. Чисельник у формулі (4.1) є середній добуток відповідних відхилень фактора x та результату y від їх вибіркових середніх значень. Цю величину називають **вибірковою коваріацією** X та Y і позначають S_{xy} . Знаменник представляє собою добуток стандартних відхилень чисельних значень відповідно x та y , які зазвичай позначають відповідно S_x і S_y . Тоді формулу (4.1) можна представити у вигляді

$$r = \frac{S_{xy}}{S_x \cdot S_y} \quad (5.2)$$

Величина коефіцієнта кореляції r не має одиниць вимірювання, вона не залежить від одиниць вимірювання X та Y та від вибору початку координат. Значення r належать інтервалу: $-1 \leq r \leq 1$ і що більше його абсолютне значення, то сильніший лінійний зв'язок між фактором і результатом. Крім того коефіцієнт r вказує і на напрямок зв'язку: при позитивних значеннях r зв'язок прямий, при від'ємних – зв'язок зворотній. Обчислення r спрощується, якщо скористатись формулами:

$$S_{xy} = \frac{\sum_{i=1}^n x_i y_i}{n} - \bar{x} \bar{y} \quad (5.3)$$

$$S_x^2 = \frac{\sum_{i=1}^n x_i^2}{n} - \bar{x}^2 \quad (5.4)$$

$$S_y^2 = \frac{\sum_{i=1}^n y_i^2}{n} - \bar{y}^2 \quad (5.5)$$

Метод найменших квадратів (МНК) в лінійних і нелінійних моделях однофакторної регресії

Для визначення залежності відгуку від фактора будується **регресійна модель**. В моделі регресійного аналізу виходячи з вигляду діаграми розсіювання підбирають теоретичну **лінію регресії**, яка задається функцією

$\hat{y} = f(x, \theta)$ від невідомого параметра θ . Ця функція називається **рівнянням регресії** або емпіричною формулою. Зазвичай функцію \hat{y} обирають з числа простих за виглядом аналітичних функцій (табл. 5.1.) Зрозуміло, що строге співпадання значень \hat{y} з експериментальними точками (x_i, y_i) спостерігається дуже рідко, але емпірична функція дозволяє прогнозувати значення реальної залежності для нетабличних значень x_i , тим самим «згладжувати» результати вимірювань величини відгуку. Треба зауважити, що множину точок (x_i, y_i) можна моделювати різними емпіричними формулами і обирати з них найкращу модель.

Таблиця 5.1

Вид регресії	№ п/п	Назва регресії	Рівняння регресії	Невідомий параметр θ
Лінійна	1	Лінійна	$\hat{y} = ax + b$	$\{a, b\}$
	2	Параболічна	$\hat{y} = ax^2 + bx + c$	$\{a, b, c\}$
	3	Степенева (геометрична)	$\hat{y} = ax^b$	$\{a, b\}$
	4	Гіперболічна	$\hat{y} = \frac{a}{x} + b$	$\{a, b\}$
	5	Показникова	$\hat{y} = a \cdot e^{bx}, a > 0$	$\{a, b\}$
	6	Логарифмічна	$\hat{y} = a \ln x + b$	$\{a, b\}$
	7	Дробово-лінійна	$\hat{y} = \frac{1}{ax + b}$	$\{a, b\}$
	8	Дробово-раціональна	$\hat{y} = \frac{x}{ax + b}$	$\{a, b\}$

Для оцінювання невідомого параметру рівняння регресії в математичній статистиці використовується декілька методів, наприклад метод максимальної вірогідності. Найбільш важливим серед всіх методів є метод найменших квадратів (МНК), який запропонував Лежандр у 1805 р. Основні досягнення цього методу пов'язані з іменами К. Гауса та О. Маркова. Цей метод відрізняється від методу максимальної вірогідності, але має власні оптимальні властивості. Питання про доцільність використання МНК зводиться до питання про властивості оцінок найменших квадратів. Цей метод дає незсунені оцінки, дисперсія яких мінімальна у класі лінійних оцінок від спостережень.

Розглянемо задачу **лінійної регресії**. Припустимо, що є n вибіркових спостережень (x_i, y_i) і ми хочемо апроксимувати їх («підігнати») лінійним рівнянням регресії $\hat{y} = ax + b$, що задає пряму лінію. Основою МНК є вибір зі всієї множини прямих на площині такої, якій відповідає найменше значення суми квадратів відхилень від неї до точок графіка розсіювання, тобто

$$F = \sum_{i=1}^n [y_i - \hat{y}_i]^2 = \sum_{i=1}^n [y_i - (ax_i + b)]^2 \quad (5.6)$$

Величини x_i і y_i задачі фіксовані числа, невідомий параметр $\theta = \theta(a, b)$, тому функціонал F залежить від a і b :

$$F(a,b) = \sum_{i=1}^n [y_i - \hat{y}_i]^2 \rightarrow \min_{i=1 \div n} \quad (5.7)$$

Параметр a – **коефіцієнт регресії** показує на скільки одиниць у середньому зміниться y із зміною x на одиницю. Він має такі самі одиниці виміру, що і відгук. У разі прямого зв'язку a – величина додатня, а при зворотному – від'ємна. Параметр b – вільний член рівняння регресії, тобто це значення \hat{y} при $x=0$.

Значення параметрів a і b , що мінімізують функціонал $F(a,b)$ (рис. 5.3), є рішенням системи рівнянь:

$$\begin{cases} \frac{\partial F}{\partial a} = 0 \\ \frac{\partial F}{\partial b} = 0 \end{cases} \quad (5.8)$$

або

$$\begin{cases} \sum_{i=1}^n y_i = a \sum_{i=1}^n x_i + bn \\ \sum_{i=1}^n x_i y_i = a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i \end{cases} \quad (5.9)$$

Рівняння (4.9) називаються **нормальними рівняннями регресії**. Остаточні формули для обчислення значень параметрів a і b мають вигляд:

$$a = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad (5.10)$$

$$b = \frac{\left(\sum_{i=1}^n y_i \right) \left(\sum_{i=1}^n x_i^2 \right) - \left(\sum_{i=1}^n x_i \right) \sum x_i y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad (5.11)$$

Пряма $\hat{y} = ax + b$, де a та b є рішенням системи нормальних рівнянь (4.9) називається **регресією Y на X** . Аналогічно можна одержати рівняння **регресії X на Y** :

$$x = c\hat{y} + d.$$

Взагалі, якщо треба завбачити значення y по даному значенню x , необхідно користуватися регресією Y на X ; для знаходження x по даному y – регресією X на Y .

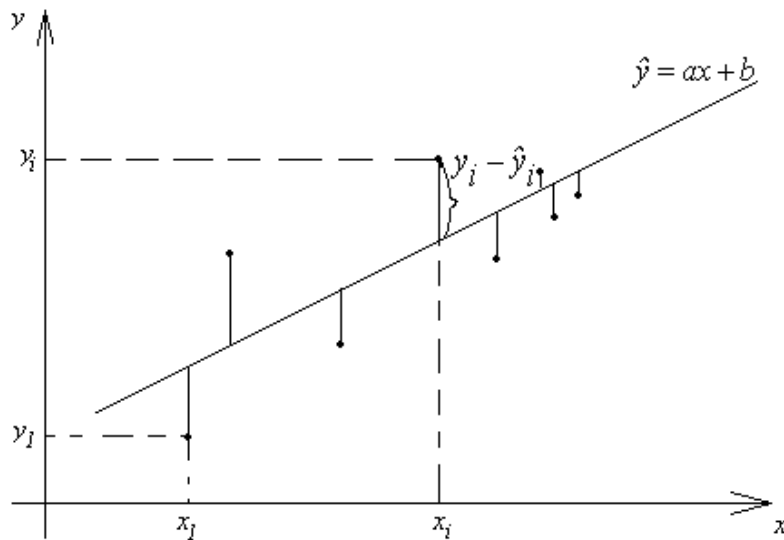


Рис.4.3. Лінійна модель регресії

Рівняння прямої регресії може бути записано й за допомогою коефіцієнту кореляції. Регресія Y на X :

$$y' = x'r, \quad (5.12)$$

регресія X на Y :

$$x' = y'r$$

де x' та y' - стандартизовані змінні:

$$x' = \frac{x - \bar{x}}{S_x}, \quad y' = \frac{y - \bar{y}}{S_y} \quad (5.13)$$

МНК узагальнюється й на випадок, коли лінія регресії парного експерименту довільна, тобто не є прямою.

Параболічна регресія

Рівняння регресії має вигляд (табл. 4.1):

$$\hat{y} = ax^2 + bx + c.$$

Треба підібрати параметр $\theta = \{a, b, c\}$ цієї моделі мінімізуючи функціонал

$$F(a, b, c) = \sum_{i=1}^n \left[y_i - (ax_i^2 + bx_i + c) \right]^2.$$

Необхідна умова мінімуму функції $F(a, b, c)$:

$$\frac{\partial F}{\partial a} = 0, \quad \frac{\partial F}{\partial b} = 0, \quad \frac{\partial F}{\partial c} = 0, \quad (5.14)$$

приймає вигляд:

$$\begin{cases} a \sum_{i=1}^n x_i^4 + b \sum_{i=1}^n x_i^3 + c \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i^2 y_i \\ a \sum_{i=1}^n x_i^3 + b \sum_{i=1}^n x_i^2 + c \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i \\ a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i + c \cdot n = \sum_{i=1}^n y_i \end{cases} \quad (5.15)$$

Геометрична регресія

В цьому випадку рівняння регресії моделюється функцією

$$\hat{y} = ax^m. \quad (5.16)$$

Треба оцінити невідомий параметр $\theta = \{a, m\}$.

Нехай $x_i > 0$ і $y_i > 0$. Прологарифмуємо рівняння (5.16):

$$\ln \hat{y} = \ln a + m \ln x. \quad (5.17)$$

Введемо нову змінну $u = \ln x$, тоді з (4.17) матимемо:

$$\ln \hat{y} = \ln a + mu = g(u).$$

Позначимо:

$$m = A, \quad \ln a = B \quad (5.18)$$

Тоді рівняння (5.17) приймає вигляд:

$$\hat{g}(u, A, B) = Au + B, \quad (5.19)$$

Тобто задача звелась до задачі лінійної регресії. Практично для знаходження теоретичної функції (5.16) необхідно зробити наступне:

- за таблицю вибірових значень (x_i, y_i) скласти нову таблицю, в якій прологарифмувати значення x_i і y_i ;
- по новій таблиці знайти параметри A і B для функції $\hat{g}(u, A, B)$ за допомогою формул (5.10) і (5.11);
- використавши позначення (5.18), знайти значення невідомих параметрів a і m і підставити їх у (5.16).

Множинний регресійний аналіз

У практичному аналізі якості здебільшого використовують методи багатофакторної регресії, тобто досліджується зв'язок між відгуком y і факторами x_1, x_2, \dots, x_n за допомогою функції регресії, що має невідомі параметри. Рівняння такої моделі можна записати у вигляді

$$\hat{y} = f(x_1, x_2, \dots, x_p; \beta_1, \beta_2, \dots, \beta_k) + e,$$

де $\beta_1, \beta_2, \dots, \beta_k$ - невідомі параметри;

e – помилка апроксимації Y за допомогою функції регресії.

Зокрема, якщо $k = p + 1$ і

$$f(x_1, x_2, \dots, x_p; \beta_0, \beta_1, \dots, \beta_p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p,$$

модель називається **лінійною множинною**. Тоді

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + e \quad (5.22)$$

В цьому рівнянні деякі факторні змінні можуть бути функціями інших змінних або залежати одна від іншої. Наприклад,

$$\hat{y} = \beta_0 + \beta_1 \sin z_1 + \beta_2 \cos z_1 + e$$

є модель множинної лінійної регресії з $x_1 = \sin z_1$ і $x_2 = \cos z_1$. В частковому випадку, якщо $x_i = x^i, i = 1 \div p$ отримується модель **поліноміальної регресії**.

$$\hat{y} = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_p x^p + e.$$

Нарешті, треба розуміти, що слово «лінійна» розуміє лінійність відносно параметрів, але не відносно факторних змінних. Так

$$\hat{y} = \beta_0 + \sin(\beta x_1) + \beta_2 x_2$$

не є лінійною функцією параметрів.

В цьому розділі будемо розглядати модель множинної лінійної регресії у вигляді (5.22). В цій моделі коефіцієнти регресії $\beta_1, \beta_2, \dots, \beta_n$ показують, на скільки одиниць зміниться відгук \hat{y} у разі зміни відповідної факторної ознаки на одиницю за умови, що значення інших факторних ознак x , які входять до моделі, є фіксованими.

Параметри моделі оцінюються за вибіркою об'єму n , що отримана таким чином: фіксуємо факторні змінні x_1, x_2, \dots, x_p і спостерігаємо значення відгуку y_1 ; знову фіксуємо факторні змінні x_1, x_2, \dots, x_p і спостерігаємо значення відгуку y_2 і т.д. Таким чином отримуємо вибірку з n спостережень $(y_1; x_{11}; x_{21}, \dots, x_{p1}), \dots, (y_n; x_{1n}, \dots, x_{pn})$. Для моделі множинної лінійної регресії маємо

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} + e_i, \quad i = 1 \div n \quad (5.23)$$

де $\beta_1, \beta_2, \dots, \beta_k$ - невідомі параметри;

e_1, \dots, e_n - незалежні випадкові помилки, які розподілені за законом

$$N(0, \sigma^2).$$

Крім того, будемо вважати, що $x_{1i}, x_{2i}, \dots, x_{pi}, \quad i = 1 \div n$ є фіксовані значення незалежних змінних x_1, \dots, x_p .

МНК – оцінки b_0, b_1, \dots, b_p параметрів $\beta_0, \beta_1, \dots, \beta_p$ мінімізують суму квадратів відхилень.

$$F = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{1i} - \dots - \beta_p x_{pi})^2 \quad (5.24)$$

Зазвичай їх називають **частинними коефіцієнтами регресії**. Іноді оцінку b_0 називають **вільним членом, константою або зсуном по у**. Оцінка рівняння множинної лінійної регресії або **площина найменших квадратів** може бути записаною у вигляді

$$\hat{y} = b_0 + b_1 x_1 + \dots + b_p x_p. \quad (5.25)$$

Модель регресії у загальному вигляді в матричній формі може бути записана як:

$$Y = X\beta + e,$$

де β - матриця невідомих параметрів розмірності $(p + 1) \cdot 1$;

Y – матриця з n спостережень розмірності $n \cdot 1$;

e – матриця з n помилок розмірності $n \cdot 1$;

X^T – так звана **матриця плану** розмірності $n \cdot (p + 1)$.

$$X^T = \begin{pmatrix} 1 & x_{11} & x_{21} & \dots & x_{p1} \\ 1 & x_{12} & x_{22} & \dots & x_{p2} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{1n} & x_{2n} & \dots & x_{pn} \end{pmatrix}$$

Рівняння (4.23) можуть бути записані у вигляді

$$Y = X^T \beta + e,$$

а вираз (4.24) приймає вид

$$F = (Y - X^T \beta)^T (Y - X^T \beta).$$

Матриця МНК – оцінок $B = (b_0, b_1, \dots, b_p)^T$ невідомих параметрів знаходиться як розв'язок системи **нормальних рівнянь**

$$X \cdot X^T \cdot \beta = XY$$

і має вигляд

$$B = (X \cdot X^T)^{-1} (X \cdot Y) \quad (5.26)$$

Для отримання оцінок коефіцієнтів багатofакторної моделі регресії доцільно використовувати пакети статистичних програм (ПСП). Наприклад, використання функції пакета EXCEL «ЛИНЕЙН», наведено в [16].

У вихідних даних програм множинної лінійної регресії зазвичай містяться ще декілька величин. Наприклад – **залишкова сума квадратів** (або **помилка**) SF_R , що представляє собою значення F (5.24), яке отримується при підстановці МНК - оцінок замість невідомих параметрів, тобто

$$SF_R = \sum_{i=1}^n (y_i - b_0 - b_1 x_{1i} - \dots - b_p x_{pi})^2 \quad (5.27)$$

Якщо цю величину поділити на **число степенів свободи залишків або помилок**

$$v_R = n - p - 1, \quad (5.28)$$

то отримується незсунена оцінка дисперсії помилок, яка називається **залишковим середнім квадратом помилки** і позначається

$$MF_R = \frac{SF_R}{v_R} \quad (5.29)$$

Квадратний корінь з MF_R називається **стандартною похибкою оцінки**

$$S = \sqrt{MF_R} \quad (5.30)$$

Зауваження. У регресійній моделі (5.22) коефіцієнт β_i вимірює степінь зміни відгуку y в залежності від фактора x_i , коли значення інших факторів фіксовані. На практиці часто виникає ситуація, коли факторні змінні вимірюються в різних одиницях (на відміну від ситуації приклада 5.2)

і коефіцієнти не можливо порівняти за величиною. Ця проблема вирішується завдяки застосування **стандартизованих факторних змінних**

$$z_j = \frac{x_j}{S_j}, \quad j = 1 \div p, \quad (5.31)$$

де

$$S_j^2 = \frac{1}{n-1} \sum_{i=1}^n (x_{ji} - \bar{x}_j)^2.$$

Тоді модель множинної лінійної регресії в термінах z_j буде задаватись рівняннями, аналогічними (5.23):

$$y_i = \gamma_0 + \gamma_1 z_{1i} + \dots + \gamma_p z_{pi} + e_i, \quad i = 1 \div n, \quad (5.32)$$

де γ_k - невідомі параметри, $k = 0 \div p$;

e_i - незалежні випадкові помилки, які розподілені за законом $N(0, \sigma^2)$.

Перевага стандартизації (5.31) полягає в тому, що $\gamma_1, \dots, \gamma_p$ вимірюють степінь зміни в одній шкалі, що дозволяє робити висновки про вплив факторних змінних на відгук (згадаємо наведений на початку розділу приклад аналізу якості продукції в залежності від таких факторів як рівень освіти (x_1), вік (x_2), стать (x_3), місце проживання (x_4) і т.д. працівника).

Після знаходження оцінок невідомих параметрів (5.25) визначають їх значущість. Для цього при заданому рівні значущості α найчастіше використовують t -критерій Стюдента для двосторонньої області з ν_R степенями вільності і обчислюють статистичну t :

$$H_0 : b_j = 0; \quad H_1 : b_j \neq 0; \quad t = \frac{|b_j|}{\sqrt{MF_R \cdot c_{jj}}},$$

де c_{jj} - елемент головної діагоналі оберненої матриці $(X \cdot X^T)^{-1}$.

Критичне значення $t_{кр}$ знаходять за таблицями критичних точок t -розподілу Стюдента для ν_R степенів вільності при заданому рівні значущості α (додаток 2). Якщо гіпотеза $H_0 : b_j = 0$ приймається для всіх j , то на цьому регресійний аналіз закінчується. Для значущих факторів рівняння регресії будують інтервальні оцінки коефіцієнтів. Довірчий інтервал для β_j :

$$b_j - t_{кр} \sqrt{MF_R \cdot c_{jj}} \leq \beta_j \leq b_j + t_{кр} \sqrt{MF_R \cdot c_{jj}} \quad (5.32)$$

Деякі програми з ПСП друкують значення статистики t для кожного коефіцієнта b_j . Іноді це значення називають величиною **t -включення**.

Зауважимо, що гіпотеза $H_0 : b_j = 0, j = 1 \div p$ може розглядатися як гіпотеза про те, що факторна змінна x_j не покращує передбачення відгуку у в порівнянні з передбаченням, що отримане за допомогою регресії y за $(p-1)$

іншими змінними. Коефіцієнт b_0 можна вважати передбаченим значенням відгуку y при $x_1 = \dots = x_p = 0$.

Степінь лінійного зв'язку між факторами і результатом і факторами визначається за допомогою часткових і множинних коефіцієнтів кореляції (Нелінійні зв'язки не визначаються. В цьому випадку для оцінки використовують кореляційне відношення).

Оцінкою парного коефіцієнта кореляції є статистика

$$r_{jk} = \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{nS_x S_y}, \quad j, k = 1 \div p.$$

Оцінкою часткового коефіцієнта кореляції l -го порядку ($l = p - 2$) є **вибірковий частковий коефіцієнт кореляції l -го порядку**. Він обчислюється на базі кореляційної матриці, що складена з вибірових парних коефіцієнтів кореляції:

$$Q_p = \begin{pmatrix} 1 & r_{12} & \dots & r_{1p} \\ r_{21} & 1 & \dots & r_{2p} \\ & & r_{jk} & \\ r_{p1} & r_{p2} & \dots & 1 \end{pmatrix}, \quad (5.33)$$

за формулою

$$r_{jk \cdot 1, 2, \dots, p} = \frac{q_{jk}}{(q_{jj} \cdot q_{kk})^{1/2}} \quad (5.34)$$

де q_{jk}, q_{jj}, q_{kk} - алгебраїчні доповнення до відповідних елементів матриці (5.33).

Зв'язок одного з факторів (наприклад x_j) з усіма іншими можна оцінити за допомогою **вибіркового множинного коефіцієнта кореляції**

$$R_{j \cdot 1, 2, \dots, p} = \sqrt{1 - \frac{|Q_p|}{q_{jj}}}, \quad (5.35)$$

де $|Q_p|$ - визначник кореляційної матриці (4.33),

q_{jj} - алгебраїчне доповнення до елемента r_{jj} .

Квадрат вибірового множинного коефіцієнта кореляції

$$D = R_{j \cdot 1, 2, \dots, p}^2$$

називається **вибірковим множинним коефіцієнтом детермінації**. Коефіцієнти $R_{j \cdot 1, 2, \dots, p}$ і D - величини додатні і приймають значення від 0 до 1.

Розглянемо частковий випадок двофакторної (x_1, x_2) моделі

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + e \quad (5.36)$$

За міру лінійного зв'язку між відгуком y і факторами x_1, x_2 використовують статистику множинний коефіцієнт кореляції

$$R_{y \cdot 1,2} = \sqrt{\frac{r_{y1}^2 + r_{y2}^2 - 2 \cdot r_{y1} \cdot r_{y2} \cdot r_{12}}{1 - r_{12}^2}} \quad (5.37)$$

де r_{ij} - парні коефіцієнти кореляції.

Можна розглядати також $R_{1 \cdot y,2}$ і $R_{2 \cdot y,1}$. Якщо $R_{y \cdot 1,2} = 1$, то відгук y однозначно визначається функціональною залежністю (5.36).

Для встановлення впливу фактору x_1 (або x_2) на зміну відгуку y використовують частковий коефіцієнт кореляції

$$r_{yx_1 \cdot x_2} = \frac{r_{y1} - r_{y2} \cdot r_{12}}{\sqrt{(1 - r_{12}^2)(1 - r_{y2}^2)}} \quad (5.38)$$

Аналогічно визначається і $r_{yx_2 \cdot x_1}$.

Список використаної і рекомендованої літератури

1. ДСТУ 3230-95. Управління якістю та забезпечення якості. Терміни та визначення. – К.: Держстандарт України, 1995. – 38 с.
2. ДСТУ 8286-03. Контрольні карти Шугарта. – К.: Держспоживстандарт, 2003. – 32 с.
3. Афифи А., Ёйзен С. Статистический анализ: Подход с использованием ЭВМ. Пер. с англ. – М.: Мир, 1982. – 488 с.
4. Баран Р.Я., Жолоб Л.І., Кондур О.С., Крупа О.М. Статистика: Навчальний посібник для студ. Економічних спеціальностей вузів. – Чернівці: ТОВ «Видавництво «Наші книги», 2008. – 240 с.
5. Богатырёв А.А., Филинов Ю.Д. Стандартизация статистических методов управления качеством. – М.: Изд-во стандартов, 1989 г.
6. Гельфанд С.Ю., Дьяконова Э.В. Статистические методы контроля качества в конс. – М.: Лёгкая и пищевая пр-ть, 1984 г.
7. Єріна А.М. Статистичне моделювання та прогнозування: Навч. посіб. – К.: КНЕУ, 2001. – 170 с.
8. Заварыкин В.М. и др. Численные методы: Учебн. пособие. – М.: Посвещение, 1990. – 176 с.
9. Захожай В.Б., Чорний А.Ю. Статистика якості: Підр. для студ. вищ. навч. закл. – К.: МАУП, 2005. – 576 с.
10. Кендалл М. Дж., Стюарт А. Статистические выводы и связи. – М.: Наука, 1977. – 899 с.
11. Коуден. Статистические методы контроля качества. – 1961 г. Физмат.
12. Кулинич О.І., Кулинич Р.О. Теорія статистики: підручник. – 5-те вид., переробл. і допов. – К.: Знання, 2010. – 239 с.
13. Лойко Д.П., Вотченікова О.В., Удовіченко О.П., Котляр М.А. Управління якістю: навч. посібник. – Донецьк: ДонНУЕТ, 2008. – 230 с.
14. Мердок. Контрольные карты. Изд-во Фин. статистика, 1886 г.
15. Сошникова Л.А., Тамашевич В.Н., Уебе Г., Шефер М. Многомерный статистический анализ в экономике: Учебн. пособие для вузов / Под. ред. проф. В.Н. Тамашевича. – М.: ЮНИТИ-ДАНА, 1999. – 598 с.
16. Чорний А.Ю. Статистика якості. Практикум: навч. посіб. – К.: Київ. нац. торг.-екон. ун-т, 2011. – 264 с.
17. Штангерт А.М., Копилюк О.І. Статистика: Навчальний посібник. – Київ: Центр навчальної літератури, 2005. – 232 с.

Словник уживаних термінів

α (рівень значущості) у тому випадку, коли висунута нульова гіпотеза є правильною і вона відхиляється, припускається помилка I виду. Імовірність зробити помилку I виду позначається як грецькою літерою α і називається рівнем значущості.

Варіація – коливання ознаки, мінливість величини ознаки одиниць, що входять до складу сукупності. Існує декілька статистичних характеристик коливання ознаки: варіаційний розмах, середнє лінійне відхилення, середнє квадратичне відхилення, дисперсія, коефіцієнт.

Вибіркова сукупність – сукупність випадково відібраних об'єктів з генеральної сукупності.

Випадкова величина – величина, яка набуває залежно від деяких випадкових обставин одне зі значень x_1, x_2, \dots, x_n , що мають певні імовірності p_1, p_2, \dots, p_n .

Випадкова змінна – це будь-яка змінна, значення якої не може бути визначено.

Генеральна сукупність – сукупність об'єктів, з яких відбирається вибірка.

Гіпотеза – наукове припущення, що висувається для пояснення якогось явища та потребує перевірки досвідом і теоретичного обґрунтування для того, щоб стати науковою теорією.

Дисперсія – середній квадрат відхилення варіант x від середнього арифметичного значення \bar{x} .

Кількісні зміни – скалярне вимірювання у визначеній шкалі ступеня проявлення досліджуваних властивостей об'єкта.

Контрольна карта – карта, на якій для наочності відображення стану процесу відмічають значення відповідної вибіркової характеристики суміжних вибірок у часовій послідовності.

Кореляція – систематичний і обумовлений зв'язок між двома явищами чи процесами.

Кореляційна матриця – матриця, елементами якої є коефіцієнти кореляції між парами випадкових величин.

Моніторинг – 1) постійне спостереження за будь-яким процесом з метою виявлення його відповідності бажаному результату або первинними припущенням; 2) спостереження, оцінка і прогноз стану навколишнього середовища у зв'язку з господарською діяльністю людини.

Якість – сукупність характеристик об'єкта, які стосуються його здатності задовольнити встановлені й передбачені потреби.

Якість продукції – сукупність характеристик продукції (процесу, послуги), які стосуються її здатності задовольняти встановлені і передбачені потреби.

Навчальне видання

СТАТИСТИКА ЯКОСТІ

Конспект лекцій

Укладачі: Турчанінова Людмила Іванівна,
Доля Олена Вікторівна

Комп'ютерна верстка

Підписано до друку	Формат
Папір офсетний. Гарнітура Таймс. Друк на різнографі.	
Ум. друк. арк.	Обл. –вид. арк.
Ум. фарбовідб.	Тираж прим. Вид. № Зам. №

КНУБА, Повітрофлотський проспект, 31, Київ, Україна, 03680

E-mail: red_istad@ua.fm

Віддруковано в редакційному відділі
Київського національного університету будівництва і архітектури

Свідоцтво про внесення до Державного реєстру суб'єктів видавничої справи
ДК №

СТАТИСТИКА ЯКОСТІ

Конспект лекцій

*для студентів, які навчаються за напрямом підготовки
0305 «Економіка та підприємництво» всіх форм навчання
спеціальності 8.03051001
«Товарознавство і комерційна діяльність»*

Всі цитати, цифровий
та фактичний матеріал,
бібліографічні відомості
перевірені. Написання
одиниць вимірювання
відповідає стандартам

Підписи авторів _____

« _____ » _____ 20__ р.

Підпис голови методичної комісії факультету

« _____ » _____ 20__ р.