# Лекція Тема 5.

# СТАТИСТИЧНІ ГІПОТЕЗИ І СТАТИСТИЧНІ ВИСНОВКИ

**5.1. Сутність і види статистичних гіпотез**

***Статистичною гіпотезою*** *називається будь-яке припущення відносно виду або параметрів невідомого закону розподілу. В конкретній ситуації статистичну гіпотезу* формулюють як припущення, з певним рівнем значущості, про властивості *генеральної сукупності* на основі оцінок вибіркової сукупності.

Статистичну гіпотезу прийнято позначати латинською літерою *Н (від лат. Hypothesis).* Так, може бути висунута статистична гіпотеза відносно того, що середнє значення в генеральній сукупності дорівнює деякій величині *а* або є меншим за деяку величину *b*. У вигляді умовних позначень це може бути записано наступним чином:

*Н:*   *а* або *Н:*   *b* .

Це гіпотетичне твердження є або справедливим (істинним), або помилковим (хибним), що потребує його перевірки.

Розрізняють ***прості*** та ***складні*** статистичні гіпотези. Статистична гіпотеза вважається **простою**, якщо вона повністю визначає теоретичну функцію розподілу випадкової величини.

Наприклад, наведена раніше гіпотеза *Н:*   *а* є саме простою,

оскільки значення середньої може бути описано лише певним значенням *а*.

**Складна** гіпотеза містить скінченну чи безкінечну кількість простих гіпотез, причому певна область значень параметру повинна

бути вказаною. Наприклад, гіпотеза *Н:*   *b* складається з безлічі

простих гіпотез Н:   *с* , в яких *с* – будь-яке значення, що є меншим, ніж *b*.

Якщо гіпотеза описує параметри генеральної сукупності, то вона

називається **параметричною**; якщо гіпотеза описує закон розподілу –

## непараметричною.

Також статистичні гіпотези поділяються на нульові та альтернативні. Гіпотезу про те, що дві сукупності, які порівнюють між собою за однією чи декількома ознаками, є однаковими між собою, називають **нульовою гіпотезою** або **нуль-гіпотезою**. Ця гіпотеза позначається як *Н0* і є таким твердженням, що приймається, коли немає переконливих аргументів для його спростовування.

Наприклад, гіпотеза «*H0 : fi1 - fi2 = 0*» читається так: «висунута нульова гіпотеза про відсутність значущої відмінності між значеннями *fi1* і *fi2*»*.* Зазвичай нульова гіпотеза – це те, що ми хочемо спростувати, якщо перед нами стоїть завдання довести значущість відмінностей.

Нульова гіпотеза спростовується тоді, коли за вибіркою отримується результат, який при істинності нульової гіпотези є малоймовірним. Межею малоймовірного чи неймовірного вважають

  0,05 тобто 5% (інколи 0,01 або 0,001). Якщо орієнтуватися на

правило «трьох сигм», то ймовірність помилки α повинна дорівнювати 0,0027. Однак, для цього рівня ймовірності критерії рідко табулюються: зазвичай значення критеріїв у статистико- математичних таблицях розраховані для ймовірності помилки 0,05; 0,01; 0,001 [6, c. 272] .

**Альтернативна** або **дослідницька гіпотеза** є логічним запереченням нульової гіпотези та позначається як *H1*. Природно, що це гіпотеза про *існування* відмінностей. Наприклад, гіпотеза «*H1: fi1 - fi2 ≠ 0»* читається так: «висунута альтернативна гіпотеза про наявність значущої відмінності між значеннями *fi1* і *fi2*». Зазвичай альтернативна гіпотеза – це те, що ми хочемо довести. Проте існують завдання, коли бажано підтвердити нульову гіпотезу і переконатися, наприклад, що вибірки не відрізняються між собою за певними показниками.

Наведемо приклади нульових та альтернативних статистичних гіпотез, які були сформульовані відносно генеральної сукупності. Необхідно звернути увагу на те, що в кожному випадку обидві гіпотези не можуть бути істинними одночасно і для того, щоб вибрати одну з них, необхідно використовувати належні дані.

1. *Ситуація.* Випадково обрана група із 100 людей взяла участь у дегустації молочних продуктів підприємства у ході рекламної акції. Після цього реєструється кількість людей із групи, які купували дегустовані продукти протягом наступного тижня.

*Нульова гіпотеза.* Рекламна акція підприємства не мала жодного ефекту. Інакше кажучи, відсоток покупців продукції в генеральній сукупності серед тих, хто брав участь у дегустації, дорівнює відсотку покупців продукції в генеральній сукупності, які не брали участь у дегустації. Рівень продажів залишається на минулому рівні, який, з досвіду, дорівнює 21,5 %.

*Альтернативна гіпотеза.* Проведена рекламна акція мала значний ефект. Тобто відсоток покупців продукції підприємства в генеральній сукупності, що приймали участь у дегустації, відрізняється від звичайного рівня продажів (21,5 %), який відповідає тим покупцям генеральної сукупності, що не брали участь у дегустації.

*Обговорення.* Зверніть увагу на те, що обидві гіпотези представляють собою ствердження відносно генеральної сукупності покупців, а не вибірки із 100 осіб, що приймали участь в дегустації. Вибіркові дані, що зібрані в результаті спостереження поведінки 100 випадково відібраних осіб, допоможе вирішити, яку з гіпотез необхідно прийняти. Оскільки нульова гіпотеза містить у собі точне значення відсотка (21,5 %), вона є більш визначеною, ніж альтернативна гіпотеза, яка говорить про більш широкий діапазон значень (про будь-яке значення, окрім 21,5 %). Якщо фахівці підприємства приймуть рішення, що рекламна акція з дегустацією була ефективною, то вони зроблять більш суворе ствердження, яке необхідно довести.

1. *Ситуація.* Підприємство протягом тижня оцінює нові верстати, які теоретично повинні прискорити виготовлення деталі *А*, що суттєво підвищить обсяги її виробництва.

*Нульова гіпотеза.* Нові верстати в довгостроковій перспективі ніяк не впливають на швидкість виготовлення та, відповідно, денні обсяги виробництва, які з минулого досвіду складають 121 деталь.

*Альтернативна гіпотеза.* У довгостроковій перспективі нове обладнання якимось чином впливає на швидкість виготовлення і обсяги виробництва, тобто значення випадкової величини може бути більшим чи меншим ніж 121 одиниця деталі *А* за день.

*Обговорення.* Нульова гіпотеза є більш певною. Обидві гіпотези сформульовані відносно генеральної сукупності (кількість виготовлених деталей у довгостроковій перспективі) замість оцінювання результатів роботи за минулий тиждень (вибірка). Таким чином, постачальнику нового устаткування необхідно довести, що воно не є неефективним. Це ні в якому разі не є завданням самого підприємства, що тестувало обладнання.

1. *Ситуація.* Власнику підприємства, який навчався у Європі та полюбляє все європейське, пред’явлено позов у несправедливому фаворитизмі співробітників, що закінчили іноземні вищі заклади

освіти, адже їх заробітна плата завжди є вищою, ніж у тих, хто навчався в національному закладі.

*Нульова гіпотеза.* Розміри заробітної плати працівників, які навчалися за кордоном, і тих, які не навчалися там, є рівними, без урахування випадкових відхилень у значеннях. Інакше кажучи, заробітна плата робітників не дуже б змінилася, якби власник знову сформував зарплатний фонд і розподілив саму зарплати між працівниками у випадковому порядку.

*Альтернативна гіпотеза.* Відмінність між заробітною платою тих, хто навчався за кордоном і тих, хто навчався в національних вищих навчальних закладах, є надто великою, щоб бути випадковою.

*Обговорення.* Перш за все, зверніть увагу на те, що розглянута сукупність є ідеалізованою, оскільки співробітники підприємства не можуть бути розглянуті як випадкова вибірка. Гіпотези відносяться до деякої ідеалізованої сукупності, яка представляє собою співробітників рівних з точки зору заробітної плати, а самі можливі відмінності у рівнях заробітної плати можуть бути пояснені випадковістю її розподілу між ними. Якщо нульова гіпотеза буде спростована, то власник підприємства буде мати значні проблеми.

Однак, варто пам’ятати, що статистичні методи зазвичай описують тільки самі числа, але не пояснюють, чому саме числа є такими, як вони є. Відмінність у рівнях заробітної плати може бути пояснена дискримінаційним фактором залежності міста навчання, але може бути обґрунтована й іншими факторами, як то особистими навичками та вміннями, якістю отриманої освіти, рівнем володіння іноземними мовами, отриманим досвідом тощо. Перевірка статистичної гіпотези, що розглядає лише розмір заробітної плати та місце попереднього навчання, не може показати, які саме фактори вплинули на ці відмінності.

## 5.2.Критерій як інструмент перевірки гіпотези. Потужність критерію

Статистичні висновки робляться на підставі прийняття однієї гіпотези та відхилення іншої, а саме рішення приймається з певною достовірністю. Перевірка гіпотез здійснюється на основі так званих ***статистичних критеріїв***.

**Статистичним критерієм,** або просто критерієм, називають випадкову величину *К*, що слугує для перевірки нульової гіпотези.

Для різних гіпотез статистичний критерій є неоднаковим, він дозволяє визначити, чи суперечить нульова гіпотеза фактичним даним, чи ні.

У цілому перевірку статистичної гіпотези на основі статистичного критерію можна представити послідовністю наступних етапів:

1. Формування завдань дослідження у вигляді статистичної гіпотези.
2. Вибір статистичної характеристики гіпотези.
3. Вибір нульової та альтернативної гіпотез на основі аналізу помилкових рішень та їх наслідків.
4. Визначення області допустимих значень, критичної області, а також значення статистичного критерію на основі відповідної таблиці.
5. Обчислення фактичного значення статистичного критерію.
6. Перевірка висунутої гіпотези на основі порівняння фактичних значень і критичного значення критерію.
7. Прийняття або відхилення гіпотези залежно від результатів перевірки.

Під час перевірки гіпотез згідно одного з критеріїв можливими є прийняття 2-х помилкових рішень:

* + неправильне відхилення нульової гіпотези – помилка І роду. Імовірність появи цієї помилки зазвичай складає близько 5 %, а її саму можна контролювати, оскільки нульова гіпотеза є повністю визначеною, що передбачає існування точного значення ймовірності;
  + неправильне прийняття нульової гіпотези – помилка ІІ роду.

Коли нульова гіпотеза є вірною (І) або невірною (ІІ), можуть бути прийняті два помилкових рішення: нульова гіпотеза відхиляється і приймається альтернативна; нульова гіпотеза приймається.

Можливі рішення цієї ситуації наведені в табл. 4.1.

*Таблиця 4.1*

## Можливі рішення під час перевірки гіпотез

|  |  |  |
| --- | --- | --- |
| Рішення згідно критерію | Фактично | |
| *Н0* вірна | *Н0* невірна |
| *Н0* не відхиляється | помилка І роду | правильне рішення |
| *Н0* відхиляється | правильне рішення | помилка ІІ роду |

Наприклад, якщо встановлено, що нова харчова добавка підвищує імунітет краще за інші, хоча насправді це не так – то допущена помилка І роду. Якщо було вирішено, що нова і стара харчові добавки є однаковими, а насправді одна є кращою (гіршою) за іншу, то була допущена помилка ІІ роду.

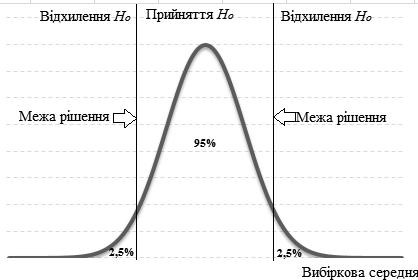
Імовірності, що відповідають невірним рішенням, називаються **ризиками** (І та ІІ). Ризик І дорівнює ймовірності помилки α (рівень значущості), ризик ІІ дорівнює ймовірності помилки β. Оскільки значення α завжди є більшим нуля, то ймовірність виникнення помилки β є завжди присутньою. За умови заданого значення α та вибірки розміру *n* значення β тим більше, чим менше значення α. Якщо значення *п* є великим, то відповідні значення α і β можуть бути скільки завгодно малими, тобто рішення будуть більш обґрунтованими [6, c. 273].

Зазвичай на практиці значення α є заданими, а значення β намагаються зробити якомога меншими. Імовірність 1 – β називається **потужністю критерію**, чим більшим є її значення, тим меншою є ймовірність появи помилки ІІ роду.

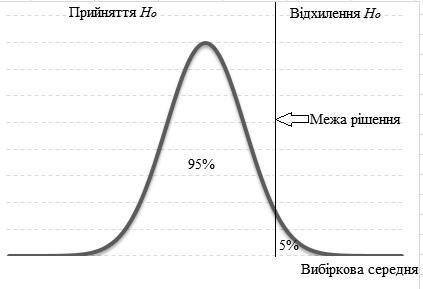
Область, в яку потрапляння значення статистичного критерію приводить до відхилення *Н0*, називається **критичною областю**. Імовірність попадання в цю область дорівнює прийнятому рівню значущості. Доповнює критичну область ще одна – **область допустимих значень**. Якщо значення критерію потрапляє в область допустимих значень, то нульова гіпотеза не суперечить фактичним даним.

Точки, що розділяють область допустимих значень і критичну область, називають **межею критичної області** або **граничними точками**. Залежно від формулювання альтернативної гіпотези критична область може бути **двосторонньою** (рис. 5.1) або **односторонньою** (рис. 5.2) – **лівостороння** або **правостороння**. Якщо обчислене значення критерію потрапляє в критичну область, нульова гіпотеза відхиляється, оскільки вона суперечить фактичним даним.

Для перевірки істинності чи хибності гіпотез зазвичай використовують *t*-критерій Стьюдента, критерій Колмагорова- Смірнова та критерій хі-квадрат тощо.



### Рис. 5.1. 5% двостороння критична область

******

***Рис. 5.2. 5% одностороння критична область***

Зокрема, для визначення достовірності відмінностей між середніми двох вибірок застосовують метод Стьюдента, а для того щоб робити висновки щодо відмінності між трьома або більшою кількістю вибірок – *F*-тест або дисперсійний аналіз (ANOVA). Якщо дослідник має справу з даними, що отримані в неметричних (номінальних або порядкових) шкалах, або вибірки надто малі для впевненості в тому, що генеральні сукупності, з яких вони взяті, підкоряються нормальному розподілу, використовують непараметричні методи – критерій хі-квадрат, критерій Манна-Уїтні, критерій Уілкоксона тощо.

## 5.3. Двостороння перевірка гіпотези. Перевірка гіпотези про рівність середньої генеральної сукупності певному заданому значенню

Одним із найпростіших випадків перевірки гіпотези є перевірка рівності між певним числом (µ0) і середнім значенням генеральної сукупності (µ). Так, нульова гіпотеза буде стверджувати, що значення є рівними між собою (Н0: µ = µ0), а альтернативна гіпотеза буде свідчити навпаки, що значення між собою є нерівними (Н1: µ ≠ µ0).

Означена перевірка гіпотези буде двосторонньою, оскільки альтернативна гіпотеза передбачає будь-яке значення середньої в генеральній сукупності: воно може бути як більшим за обране значення µ0, так і меншим.

Звертаємо вашу увагу на те, що в цій теоретичній задачі будуть присутніми цілих три числа: невідоме середнє значення генеральної сукупності (µ), задане значення (µ0), з яким порівнюють середнє по сукупності, та *х* – відоме вибіркове середнє, яке буде використовуватися для прийняття рішення щодо істинності чи хибності гіпотези. Зі всіх трьох перелічених чисел лише значення *х* є випадковою величиною, оскільки воно буде розраховано відповідно до даних вибірки з генеральної сукупності. Ця величина є оціночним значенням, яке представляє µ.

Таким чином, сутність перевірки гіпотези буде полягати у порівнянні значень µ0 та *х* , тобто у порівнянні значень уже відомих величин. Якщо ці значення відрізняються надто сильно, що не можна пояснити простою випадковістю, то нульову гіпотезу відхиляють. Якщо ж відмінність між значеннями є не надто великою, то нульову гіпотезу приймають. Критерієм «незначущості» відмінностей між

значеннями виступає стандартна похибка *x* , яка вимірює ступінь

*S*

випадковості *х* . Таким чином, коли значення *х* та µ0 знаходяться

один від одного на відстані достатньої кількості *Sx* , то це є достатнім

доказом того, що µ та µ0 є нерівними. Розглянемо два методи перевірки істинності гіпотези й отримання результату: використання довірчого інтервалу та використання методу *t*-статистики (критерію Стьюдента).

**Довірчим інтервалом** називають такий обчислений на існуючих даних інтервал, який з відомою ймовірністю може містити необхідний досліднику невідомий параметр генеральної сукупності.

Ця ймовірність визначається з урахуванням випадкового експерименту, який починається з формування випадкової вибірки. Також існує можливість вибирати ймовірність твердження, яку називають **довірчим рівнем**, **коефіцієнтом довіри** або **довірчою ймовірністю**. Традиційно цей показник установлюють рівним 95 %, але часто використовують значення 90, 99 і навіть 99,9 %. Межі довірчого інтервалу мають наступний вигляд:

*х*  *tтабл Sx* та *х*  *tтаблSx* , (4.1)

де значення *t* беруть із таблиць про двосторонній довірчий рівень, наприклад, 95%, який ми будемо використовувати у подальших розрахунках. Таблиця значень *t* наведена в *Додатку Б*.

Таким чином, для того, щоб прийняти чи спростувати гіпотези Н0: µ = µ0 та Н1: µ ≠ µ0 необхідно побудувати 95 % довірчий інтервал

за допомогою наведених формул і виходячи зі значень *х* і *Sx* . Після

цього потрібно впевнитись, що значення µ0 знаходиться в межах довірчого інтервалу. Якщо це дійсно так, то µ0 може бути розглянуто як допустиме значення середнього значення генеральної сукупності, отже необхідно прийняти нульову гіпотезу як істинну. У протилежному випадку приймають альтернативну гіпотезу. Сутність даного підходу проілюстрована на рис. 5.3.

Якщо задане опорне значення знаходиться в цьому інтервалі, то сере- днє вибірки значно від- різняється від нього. Фа- ктичне значення вибірки є значно більшим

Якщо наведене опорне зна- чення знаходиться в цьому інтервалі, то навіть якщо се- реднє вибірки відрізняється від нього, то ця відмінність незначна

Якщо задане опорне значення знаходиться в цьому інтервалі, то сере- днє вибірки значно від- різняється від нього. Фа- ктичне значення вибірки є значно меншим

95% довірчий інтервал

Стандартна похибка Середня вибірки

### Рис. 5.3. Схематичне зображення перевірки гіпотези за допомогою довірчого інтервалу

Характеристика можливих результатів дослідження за допомогою довірчого інтервалу наведена в табл. 5.2.

*Таблиця 5.2*

## Методика прийняття рішення відносно перевірки гіпотези стосовно середньої в генеральній сукупності

|  |  |
| --- | --- |
| Значення µ0 знаходиться в інтервалі  від *х*  *tтабл Sx* до *х*  *tтабл Sx* | Значення µ0 не знаходиться в інтервалі  від *х*  *tтабл Sx* до *х*  *tтабл Sx* |
| Нульова гіпотеза *Н0* приймається як допустима можливість | Альтернативна гіпотеза *Н1*  приймається як допустима можливість |
| Альтернативна гіпотеза *Н1*  спростовується | Нульова гіпотеза *Н0* спростовується |
| Відмінності між значеннями *х* та µ0 є випадковими | Відмінності між значеннями *х* та µ0 не можуть бути обумовлені лише випадковістю |
| Результат перевірки не є статистично значущим | Результат перевірки є статистично значущим |

Розглянемо використання довірчого інтервалу на прикладі, який ми згадували на початку розділу. Нехай існує підприємство, яке протягом тижня тестувало нове обладнання на своєму виробництві. Теоретично нове обладнання працює швидше, що означає збільшення обсягів виробництва деталей типу *А* протягом дня. Вихідні дані прикладу наведені в табл. 5.3.

*Таблиця 5.3*

## Вихідні дані прикладу

|  |  |  |  |
| --- | --- | --- | --- |
| Показник | Умовне позначення | Значення | Одиниця вимірювання |
| Середній денний обсяг виготовленої продукції протягом минулого тижня (з  використанням нового обладнання) | *х* | 135 | штуки |
| Стандартна похибка | *S*  *x* | 15 | штуки |
| Розмір вибірки | *п* | 7 | дні |
| Середній обсяг виготовленої продукції протягом довгострокового періоду (без використання нового  обладнання) | µ0 | 121 | штуки |

Наведені дані містять у собі 7 днів спостереження за обсягами виробництва деталі *А* за допомогою нового обладнання. Генеральна сукупність складається з усіх можливих денних обсягів виробництва

деталі з використанням нової техніки; зокрема середнє генеральної сукупності µ є середнім значенням обсягу виробництва, отримане за тривалий період використання нового обладнання (в таблиці це значення відсутнє, оскільки є невідомим). Вибіркове значення *х* є найкращою оцінкою µ.

Інформація, яка наведена в табл. 5.3, свідчить, що використання нового обладнання можливо є ефективним. Середньоденний обсяг виробництва деталі *А*, досягнутий з використанням нової техніки ( *х* = 135 шт.), є на 16 одиниць більшим, ніж середній очікуваний обсяг виробництва, розрахований за минулий довгостроковий період (µ0 = 121). Постає питання, чи є ця відмінність випадковою, чи вона дійсно викликана використанням нового обладнання.

*Висунемо гіпотези.* Таким чином, нульова гіпотеза буде стверджувати, що під час використання нового обладнання невідоме середнє значення обсягу продукції за тривалий період часу µ ***дорівнює*** заданому значенню µ0 = 121 одиниць деталі *А*, що аналогічно обсягу виробництва продукції без використання нової техніки (*Н0*: µ0 = 121).

Альтернативна гіпотеза, навпаки, стверджує, що під час використання нового обладнання, невідоме середнє значення обсягу продукції за тривалий період часу µ ***не дорівнює*** заданому значенню µ0 = 121 одиниць деталі *А* – обсягу виробництва без використання нової техніки (*Н1*: µ0 ≠ 121).

Наступним кроком нашого дослідження є розрахунок довірчого інтервалу за формулою 5.1 та значення *t* із *Додатку Б*. З кількістю ступенів свободи 6 (*п*-1=7-1=6) табличне значення t = 2,447. Отже, з імовірністю 95 % можна стверджувати, що під час використання нового обладнання протягом довгострокового періоду середнє значення обсягу виробництва деталі *А* буде знаходитись в інтервалі від 98 до 172 одиниць.

Для того, щоб перевірити правдивість гіпотези, необхідно визначити, чи входить значення µ0 = 121 до довірчого інтервалу. Як ми бачимо – входить, адже 121 знаходиться між значеннями 98 та

172. Це свідчить, що ми можемо прийняти нульову гіпотезу *Н0*: µ0 = 121 як допустиму можливість і спростувати альтернативну гіпотезу *Н1*: µ0 ≠ 121. Відмінність між середнім вибірковим значенням обсягу виробництва продукції *х* = 135 і заданим значенням µ0 = 121 може бути обумовлена лише випадковістю. Результат перевірки не є статистично значущим.

Варто відмітити, що перевірка гіпотези не довела ефективність використання нового обладнання для виготовлення деталі *А*. Воно може бути як ефективним, так і неефективним – на даний час не має доказів ні для того, ні для іншого. Варіантом вирішення такого завдання на практиці може бути збільшення вибірки, тобто використання нового обладнання ще протягом деякого періоду часу, наприклад ще два тижні або місяць (за умови вигідних пропозицій з боку постачальника цієї техніки).

Наведена формула 4.1, яку ми використовували для вирішення цього завдання, є дійсною для нормального розподілу випадкової величини. Для інших варіантів розподілу, разом з іншими формулами, вона може мати інший вигляд. Наведемо формули і позначення для нормального та біноміального законів розподілу в табл. 5.4.

*Таблиця 5.4*

## Формули розрахунку показників та їх умовні позначення

|  |  |  |
| --- | --- | --- |
| Показник | Нормальний розподіл | Біноміальний розподіл |
| Середнє значення в сукупності | µ | Ω |
| Задане значення | µ0 | Ω0 |
| Нульова гіпотеза | *Н0*: µ = µ0 | *Н0*: Ω = Ω 0 |
| Альтернативна гіпотеза | *Н1*: µ ≠ µ0 | *Н1*: Ω ≠ Ω 0 |
| Дані | *х1, х2….хп* | *х* настання подій в *п*  випробуваннях |
| Оцінка | *х* | *р*  *х* , де *п* – обсяг  *п*  вибіркової сукупності |
| Стандартна похибка | *S*  *S n*  *x* | *S*  *р*(1  *р*)  *p п* |
| Довірчий інтервал | від *х*  *tтабл Sх* до *х*  *tтабл Sх* | від *х*  *tтабл Sр* до *х*  *tтабл Sр* |
| t-статистика | *t*  *x*  0  *статистика S*  *x* | *t*  *p*  0  *статистика S*  *p* |

Перейдемо до другого методу перевірки гіпотез із тих, що ми згадували – методу t-статистики. Сутність цього методу полягає в

*t*  *x*  0

тому, щоб розрахувати значення t за формулою

*S*

*статистика*

*x*

, щоб

після цього з використанням t-таблиці (*Додаток Б*) вирішити, яку

саме гіпотезу необхідно прийняти. За умови правильності розрахунків, результат повинен бути той самий, що й при використанні методу довірчого інтервалу. Тому немає жодного значення, який із методів використовувати.

Відповідно до другого методу, перевірку статистичної гіпотези починають з того, що для визначення відмінності між двома гіпотезами на основі найкращих даних розраховують так звану тест- статистику. Отриману величину тест-статистики (як то, t-статистику) далі порівнюють з відповідним критичним значенням із таблиці критичних значень (t-таблиці); на основі цього порівняння й визнають, яку гіпотезу необхідно прийняти. Тест-статистика показує, скільки стандартних помилок знаходиться між значеннями *х* та µ0.

Для нормального закону розподілу t-статистика розраховується за наступною формулою:

*t*  *х*  0

(5.2)

*статистика*

*S*

*x*

Для біноміального закону вона прийме наступний вид:

*t*  *p*  0

(5.3)

*статистика*

*S*

*p*

Залежно від того, чи перевищує розрахована t-статистика табличне значення, можливі наступні варіанти розвитку подій (табл. 5.5).

*Таблиця 5.5*

## Використання t-статистики для перевірки гіпотез

|  |  |
| --- | --- |
| *t*-статистика за абсолютною величиною (за модулем) є меншою, ніж *t*-значення із *t*-таблиці | *t*-статистика за абсолютною величиною (за модулем) є більшою, ніж *t*-значення із *t*-таблиці |
| Нульова гіпотеза *Н0* приймається як допустима можливість | Альтернативна гіпотеза *Н1*  приймається як допустима можливість |
| Альтернативна гіпотеза *Н1*  спростовується | Нульова гіпотеза *Н0* спростовується |
| Відмінність між значеннями *х* та µ0 є випадковою | Відмінність між значеннями *х* та µ0 не може бути обумовлена лише  випадковістю |
| Результат перевірки не є статистично значущим | Результат перевірки є статистично значущим |

Звертаємо вашу увагу на те, що таблиці бракує ще одного варіанту порівняння – коли фактичне значення t-статистики *співпадає* з t-значенням із таблиці. У цьому випадку µ0 точно співпадає з межею довірчого інтервалу. Така ситуація буває вкрай рідко, тим не менше, все ж трапляється. Тоді необхідно провести розрахунок з більшим ступенем точності та залишити більше знаків після коми. Або, другий варіант, можна прийняти нульову гіпотезу і зробити зауваження, що значення є граничним.

Корисно також запам’ятати наступне правило: якщо значення t- статистики за абсолютною величною є меншим 2, то нульову гіпотезу спростовують, і навпаки. Однак, це правило можна використовувати тільки в тих випадках, коли вибіркова сукупність є більшою, ніж 40 одиниць, тоді використовують число 2 як апроксимацію t-значення 1,96.

Розрахуємо значення t-статистики для попереднього прикладу з використанням нового обладнання на підприємстві для виробництва

деталі *А*. Згідно з умовою, нам було відомо, що *п* = 7, *х* = 135, *x*  15 ,

*S*

а задане значення 0 = 121. Тоді, відповідно до формули 4.2, отримуємо:

*tстатистика*

 135  121  0,933 .

15

Оскільки отримане значення t-статистики за модулем є меншим, ніж табличне значення 2,447 (*Додаток Б*), то ми можемо зробити висновок, що необхідно прийняти нульову гіпотезу і спростувати альтернативну. Отже, за допомогою методу t-статистики ми отримали той самий результат, що й під час використання методу довірчого інтервалу.

## Одностороння перевірка гіпотези

Варто повторити, що перевірка гіпотези, про яку йшла мова до цього моменту, була двосторонньою, оскільки розглядалася нульова гіпотеза Н0: µ = µ0 проти альтернативної Н1: µ ≠ µ0. Ця альтернативна гіпотеза є двосторонньою, адже середнє в сукупності може виявитися як більшим, так і меншим, ніж задане опорне значення.

Однак, в аналізі економічних процесів часто трапляються більш конкретні ситуації, коли дослідника цікавить точніший результат, наприклад, чи є середнє в сукупності більшим заданого значення (чи меншим, залежно від ситуації). Скажімо, підприємство закупить нове обладнання тільки в тому випадку, якщо відсоток бракованих виробів виявиться меншим деякого конкретного числа, або покупці придбають нову харчову добавку тільки в тому випадку, коли кількість корисних речовин в ній виявиться більшою, ніж певне значення.

У той же час відмітимо, що використання односторонньої перевірки не є обов’язковим. Для подібних випадків можна скористатися і вже відомою двосторонньою перевіркою. Якщо результат двосторонньої перевірки свідчить, що нульова гіпотеза спростовується, а приймається альтернативна, то з урахуванням того, меншим чи більшим є середнє значення вибірки *х* , ніж задане опорне значення, можна зробити наступні висновки:

* в ситуаціях, коли *х* > µ0, середнє значення вибірки *х* є значно більшим заданого значення µ0;
* в ситуаціях, коли *х* < µ0, середнє значення вибірки *х* є значно меншим заданого значення µ0.

Однак, бувають ситуації, коли необхідно використовувати саме односторонню перевірку, наприклад, коли її результати свідчать про те, що необхідно прийняти нульову гіпотезу, а результати двосторонньої – альтернативну. Одностороння перевірка приділяє увагу тільки одній стороні, тому може краще знайти відмінність між середнім і певним заданим значенням. У той же час, існує недолік, адже ця перевірка нічого не знає про іншу сторону значень від заданого числа.

Під час процедури односторонньої перевірки гіпотези нульова гіпотеза стверджує, що значення µ знаходиться по одну сторону від µ0, а альтернативна гіпотеза – по іншу. Випадок, коли µ = µ0, завжди включають до нульової гіпотези, яку необхідно спростувати, оскільки це гарантує, що прийняття альтернативної гіпотези дозволяє зробити більш точний висновок «набагато більше, ніж …» або «набагато менше, ніж…».

Одностороння гіпотеза про те, що µ є меншим, ніж µ0, формулюються наступним чином:

*Н0: µ*  *µ0* , (4.4)

*Н1: µ < µ0* (4.5)

Отже, нульова гіпотеза стверджує, що невідоме середнє значення генеральної сукупності є щонайменше таким же великим як і значення µ0. Альтернативна гіпотеза говорить, що невідоме середнє значення генеральної сукупності є меншим за значення µ0.

Одностороння гіпотеза про те, що µ є більшим, ніж µ0, формулюється так:

*Н0: µ*  *µ0,* (4.6)

*Н1: µ > µ0* (4.7)

Таким чином, нульова гіпотеза стверджує, що невідоме середнє значення генеральної сукупності є не більшим, ніж значення µ0. Альтернативна гіпотеза говорить, що невідоме середнє значення генеральної сукупності є більшим за задане опорне значення µ0.

Саме використання односторонньої перевірки йде з деяким обмеженням: дослідник повинен бути впевнений, що незалежно від того, як поводять себе дані, ця перевірка буде продовжуватись використовуватися лише на цьому самому боці. Якщо внаслідок зміни характеру даних виникає необхідність або сумнів щодо можливості використання перевірки іншого боку, то необхідно залишити односторонню перевірку і перейти до двосторонньої.

Також відмітимо, що одностороння перевірка залишає деякі сумніви, адже містить суб’єктивну думку дослідника про те, чи є значення тільки більшим, чи тільки меншим за певне число. У ситуаціях, коли необхідно переконати людей із протилежною точкою зору, краще використовувати двосторонню перевірку. У той же час, якщо дослідник представляє результати своїх досліджень колу однодумців, які підтримують його точку зору, можна використовувати односторонню перевірку.

Як і для двосторонньої перевірки, для односторонньої існують два способи перевірки гіпотези: метод довірчого інтервалу та метод використання t-статистики. Розглянемо це за допомогою прикладу.

Відділ управління персоналом промислового підприємства збирається організувати виїзний тренінг співробітників з метою зміцнення командного духу та розробки інноваційних ідей. Виходячи з проведеного теоретичного аналізу, тренінг буде успішним і принесе плідні результати тільки в тому випадку, коли більше 30 % співробітників вирішать його відвідати. Ці 30,0 % є значенням µ0, яке було отримано без урахування вибірки. Для того, щоб вирішити питання щодо необхідності організації такої поїздки, HR-відділ провів опитування серед випадкової кількості співробітників підприємства (*п* = 43) і розрахував односторонній довірчий інтервал. Виходячи з проведеного опитування 45 % співробітників виказали бажання прийняти участь у тренінгу ( *х*  45 ). Керівництву відділу на

95 % впевнено, що щонайменше 37,8% працівників поїдуть на

тренінг. Оскільки задане значення µ0 = 30,0%, що характеризує ефективність тренінгу, знаходиться поза довірчого інтервалу (отже, неприйнятно вважати, що середнє значення дорівнює 30,0 %), існує переконливий доказ того, що середнє в генеральній сукупності є більшим за 30,0 %.

Ця ситуація коротко наведена в табл. 5.6.

## *Таблиця 5.6* Тест для відсотка працівників підприємства, що бажають прийняти участь у тренінгу (з використанням довірчого

**інтервалу)**

|  |  |  |
| --- | --- | --- |
| Показник | Загальний вираз | Конкретні значення |
| Нульова гіпотеза | Н0: µ  µ0 | Н0: µ  30 % |
| Альтернативна гіпотеза | Н1: µ > µ0 | Н1: µ > 30 % |
| Середня | *х* | 45% |
| Стандартна похибка (розрахована за  формулою з таблиці 4.4) | *Sx* | 3,324 % |
| Розмір вибірки | *n* | 43 |
| Задане значення | µ0 | 30 % |
| Довірчий інтервал | *х*  *tтабл Sx* до *х*  *tтабл Sx* | Керівництво відділу на  95 % впевнено, що середнє значення генеральної сукупності принаймні дорівнює  37,8 % |
| *Рішення* | *Прийняти альтернативну гіпотезу Н1* | Відділ сподівається, що більше ніж 30 % співробітників вирішать відвідати тренінг |

Таким чином, рішення задачі полягає в тому, щоб прийняти альтернативну гіпотезу Н1, оскільки задане значення знаходиться за межами довірчого інтервалу (тобто 30,0 % не означає щонайменше 37,8 %).

Наведемо загальний принцип односторонньої перевірки гіпотези для обох її типів (перевірок того, що *х* є значимо більшим чи значимо меншим за µ0) з використанням як методу довірчого інтервалу, так і *t*-статистики (табл. 5.7).

*Таблиця 5.7*

## Схема вирішення поставленого завдання

|  |  |
| --- | --- |
| Одностороння перевірка того, що µ є більшим за µ0 | Одностороння перевірка того, що µ є меншим за µ0 |
| Перевіряємо нульову гіпотезу Н0:  µ  µ0 проти альтернативної Н1: µ >µ0 | Перевіряємо нульову гіпотезу Н0: µ   µ0 проти альтернативної Н1: µ < µ0 |
| Твердження про довірчий інтервал: керівництво відділу на 95% впевнено, що середнє генеральної сукупності  принаймні не менше за *х*  *tтабл Sр* | Твердження про довірчий інтервал: керівництво відділу на 95% впевнено, що середнє генеральної сукупності не  більше за *х*  *tтабл Sx* |
| t-статистика визначається за формулою 5.3. Варто пам’ятати, що в односторонній перевірці абсолютне (модальне) значення не  використовується | t-статистика визначається за формулою  5.3. Варто пам’ятати, що в односторонній перевірці абсолютне (модальне) значення не  використовується |
| Чи дійсно *х*  *tтабл Sр*  µ0? У рамках методу довірчого інтервалу формулюється питання: чи  знаходиться значення µ0 всередині  нього. У рамках методу t-статистики формулюємо питання, чи є розраховане значення t меншим, чи  дорівнює t-табличному | Чи дійсно *х*  *tтабл Sx*  µ0? У рамках методу довірчого інтервалу формулюється питання: чи знахо- диться значення µ0 всередині нього. У рамках методу t-статистики форму- люємо питання, чи є розраховане  значення t більшим, чи дорівнює t- табличному, взятому з мінусом |
| ***Якщо відповідь на це питання***  ***позитивна, то:*** | ***Якщо відповідь на це питання***  ***позитивна, то:*** |
| Нульова гіпотеза *Н0* приймається як допустима можливість | Нульова гіпотеза *Н0* приймається як допустима можливість |
| Альтернативна гіпотеза *Н1*  спростовується | Альтернативна гіпотеза *Н1*  спростовується |
| Середнє значеннями *х* не є значимо більшим за µ0 | Середнє значеннями *х* не набагато менше за µ0 |
| Якщо *х* є більшим за µ0, то це можна пояснити випадковістю | Якщо *х* є меншим за µ0, то це можна пояснити випадковістю |
| Результат перевірки не є статистично значущим | Результат перевірки не є статистично значущим |
| Чи дійсно *х*  *tтабл Sx* > µ0? У рамках методу довірчого інтервалу формулюється питання: чи знаходиться значення µ0 поза нього. У рамках методу t-статистики формулюємо питання, чи є розрахо-  ване значення t більше t-табличного. | Чи дійсно *х*  *tтабл Sx* < µ0? У рамках методу довірчого інтервалу формулю- ється питання: чи знаходиться зна- чення µ0 поза нього. У рамках методу t- статистики формулюємо питання, чи є розраховане значення t меншим  t- табличного, взятого з мінусом. |

*Продовження таблиці 4.7*

|  |  |
| --- | --- |
| Одностороння перевірка того, що µ є більшим за µ0 | Одностороння перевірка того, що µ є меншим за µ0 |
| ***Якщо відповідь на це питання позитивна, то:*** | ***Якщо відповідь на це питання позитивна, то:*** |
| Альтернативна гіпотеза *Н1*  приймається як допустима можливість | Альтернативна гіпотеза *Н1*  приймається як допустима можливість |
| Нульова гіпотеза *Н0* спростовується | Нульова гіпотеза *Н0* спростовується |
| Середнє значення *х* є значно більшим за µ0 | Середнє значення *х* є значно меншим за µ0 |
| Відмінність між значеннями  неможливо пояснити простою випадковістю | Відмінність між значеннями  неможливо пояснити простою випадковістю |
| Результат перевірки є статистично значущим | Результат перевірки є статистично значущим |

Отже, під час використання методу довірчого інтервалу варто пам’ятати, що існують два різних односторонніх довірчих інтервали. Необхідно обрати той, який відноситься до сторони, стосовно якої сформульована альтернативна гіпотеза. Наприклад, якщо досліджувана альтернативна гіпотеза має вигляд Н1: µ > µ0, то необхідний односторонній довірчий інтервал буде складатися з усіх значень µ, які принаймні не є меншими за відповідне число,

розраховане за формулою взятого із таблиці.

*х*  *tтабл Sx*

з використанням значення t,

На рис. 4.4 показано, що для прийняття рішення про те, що

*х* значно більше за µ0, відстань між ними повинна бути достатньо великою, для того, щоб цю відмінність неможливо було пояснити лише випадковістю [16, c.472]. Односторонній довірчий інтервал використовує той самий напрям, що й альтернативна гіпотеза, тобто ті значення µ, які не є меншими за граничне значення довірчого інтервалу. Прийняти рішення про те, що середнє вибірки є значимо більшим можна тільки в тому випадку, коли задане значення µ0 знаходиться набагато нижче від середнього вибірки.

На рис. 5.5. наведено схематичне зображення односторонньої перевірки у протилежному напрямку. Односторонній довірчий інтервал використовує той самий напрямок, що й альтернативна гіпотеза, тобто ті значення µ, які є меншими або дорівнюють



### Рис. 5.4. Схема використання односторонньої перевірки для прийняття рішення про те, чи дійсно значення µ є більшим за задане значення µ0.

граничному значенню довірчого інтервалу. Прийняти рішення про те, що середнє значення вибірки є значимо меншим, можна тільки в тому випадку, коли задане значення µ0 знаходиться набагато вище від середнього з вибірки.



### Рис. 5.5. Схема використання односторонньої перевірки для прийняття рішення про те, чи дійсно значення µ є меншим за задане значення µ0.

Перевірка односторонньої гіпотези за допомогою t-статистики, в свою чергу, буде проводитись наступним чином: дослідник повинен порівняти розраховану *t*-статистику із табличним значенням *t* або з табличним значенням *t* взятим із мінусом (залежно від того, з якого боку робиться перевірка). Більше того, табличне значення t- статистики є однаковим, як для односторонньої, так і для двосторонньої перевірки; відмінність лише в тому, яким чином її використовують.

Нами був розглянутий приклад відвідування співробітниками підприємства виїзного тренінгу, де стверджувалось, що ця подія буде ефективною тільки в тому випадку, коли її відвідають не менше 30 % працівників. Відповідні дані наведені в таблиці 4.6. Виконаємо таку саму перевірку прийняття гіпотези, тільки за допомогою методу t- статистики (табл. 5.8).

*Таблиця 5.8*

### Тест для відсотка працівників підприємства, що бажають прийняти участь у тренінгу (з використанням методу

***t-статистики)***

|  |  |  |
| --- | --- | --- |
| Показник | Загальний вираз | Конкретні значення |
| Нульова гіпотеза | *Н0: µ*  *µ0* | *Н0: µ*  *30 %* |
| Альтернативна гіпотеза | *Н1: µ > µ0* | *Н1: µ > 30 %* |
| Середнє | *х* | 45 % |
| Стандартна похибка (розрахована за формулою з таблиці 4.4) | *Sx* | 3,324 % |
| Задане значення | µ0 | 30 % |
| t-статистика | *t*  *х*  0  *статистика S*  *x* | *t*  45,0  30  4,513  *статистика* 3,324 |
| Критичне значення | *tтабл* | 1,625 |
| *Рішення* | *Прийняти альтернативну гіпотезу Н1* | Відділ сподівається, що більше ніж 30 % співробітників вирішать  відвідати тренінг |

Таким чином, виходячи із даних табл. 5.8, ми приймаємо альтернативну гіпотезу Н1: µ > µ0 та спростовуємо нульову гіпотезу

Н0: µ  µ0, оскільки *tстатистика* є більшою за *tтабл* (тобто 4,513 є більшим

за 1,625). Отже, як і в прикладі із двосторонньою перевіркою, обидва методи дали один і той самий результат.

Підводячи висновки вищевикладеного, відмітимо, що перевірку статистичних гіпотез використовують для обґрунтування рішення при наявності альтернативного варіанту розвитку подій. Досить часто процедуру перевірки використовують для виявлення закономірностей і спростування випадковостей.

## Питання для самоконтролю

1. Надайте визначення статистичній гіпотезі.
2. Яку статистичну гіпотезу називають простою, а яку – складною? Наведіть приклади.
3. Чим відрізняються одна від одної параметричні та непараметричні гіпотези?
4. Дайте визначення нульовій гіпотезі.
5. Що таке альтернативна гіпотеза?
6. Наведіть ситуації з прикладами нульової та альтернативної гіпотези.
7. У чому сутність статистичного критерію? Які статистичні критерії ви знаєте?
8. Перелічіть етапи перевірки статистичної гіпотези за допомогою статистичного критерію.
9. У чому сутність помилки І роду під час перевірки гіпотези?
10. Охарактеризуйте помилки ІІ роду під час перевірки гіпотези.
11. Надайте поняття критичної області та області допустимих значень.
12. Які існують види критичної області залежно від формулювання альтернативної гіпотези?
13. Які основні методи використовують для перевірки гіпотези?
14. Що таке довірчий інтервал?
15. Коротко опишіть етапи перевірки двосторонньої гіпотези за допомогою довірчого інтервалу.
16. Коротко опишіть етапи перевірки двосторонньої гіпотези за допомогою методу t-статистики.
17. Що таке задане значення? Звідки воно береться? Чи можна взяти його із таблиці?
18. Дайте характеристику односторонній перевірці.
19. В яких ситуаціях краще використовувати односторонню перевірку?
20. Чи можна використовувати двосторонню перевірку гіпотези замість односторонньої?
21. Сформулюйте гіпотези для односторонньої перевірки.
22. Як виконують односторонню перевірку гіпотези на основі довірчого інтервалу?
23. Як виконують односторонню перевірку гіпотези на основі t-статистики?