

Міністерство освіти і науки України
Харківський національний університет імені В. Н. Каразіна

Економічний факультет
Кафедра статистики, обліку та аудиту

«ЗАТВЕРДЖУЮ»

Завідувач кафедри
д. е. н., проф. Оксана
НЕСТЕРЕНКО



Протокол № 1 від “26” серпня 2024 р.

НАВЧАЛЬНО-МЕТОДИЧНИЙ КОМПЛЕКС

дисципліни «Бізнес-статистика»

для студентів денної форми навчання
рівень вищої освіти другий (магістерський)
галузь знань 05 «Соціальні та поведінкові науки»
спеціальність 051 «Економіка»
освітня програма «Економічна аналітика та статистика»

Розроблено:

к.е.н.(PhD), доцент, доцент ЗВО кафедри статистики, обліку та аудиту
Олена КУЩЕНКО

2024 / 2025 навчальний рік

ЗМІСТ

1. Робоча програма навчальної дисципліни, Силабус та Сертифікат
2. Навчальний контент (розширений план лекцій, конспект лекцій)
3. Плани практичних (семінарських) занять, самостійної роботи (у т.ч. індивідуальні завдання)
4. Питання, задачі, завдання або кейси для поточного контролю знань і вмінь здобувачів вищої освіти, для контрольних робіт, передбачених навчальним планом, післятестаційного моніторингу набутих знань і вмінь з навчальної дисципліни
 - 4.1. Питання, задачі, завдання або кейси для поточного контролю знань і вмінь здобувачів вищої освіти
 - 4.2. Контрольні роботи, передбачені навчальним планом
5. Завдання підсумкового контролю (семестрових екзаменів)
6. Критерії оцінювання знань студентів та розподіл балів

1. РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ, СИЛАБУС ТА СЕРТИФІКАТ

Міністерство освіти і науки України

Харківський національний університет імені В.Н. Каразіна

Кафедра статистики, обліку та аудиту



Робоча програма навчальної дисципліни

Бізнес - статистика

(назва навчальної дисципліни)

рівень вищої освіти

другий (магістерський)

галузь знань

05 «Соціальні та поведінкові науки»

(шифр і назва)

спеціальність

051»Економіка»

(шифр і назва)

освітня програма

«Економічна аналітика та статистика»

(шифр і назва)

спеціалізація

(шифр і назва)

вид дисципліни

обов'язкова

[обов'язкова, вибіркова (міжфакультетська) чи факультативна]

факультет

економічний

2024 / 2025 навчальний рік

Програму рекомендовано до затвердження вченою радою факультету (навчально-наукового інституту, центру)

“27” серпня 2024 року, протокол № 11


РОЗРОБНИК(И) ПРОГРАМИ: (вказати авторів, їхні наукові ступені, вчені звання та посади)

Кущенко Олена Іванівна, к. е. н., доцент, доцент ЗВО кафедри статистики, обліку та аудиту

Програму схвалено на засіданні кафедри статистики, обліку та аудиту

Протокол від “26” серпня 2024 року № 1

Завідувач кафедри статистики, обліку та аудиту



(підпис)

Оксана НЕСТЕРЕНКО
(ім'я та прізвище)

Програму погоджено з гарантом освітньо-професійної програми

«Економічна-аналітика та статистика»
назва освітньої програми

Гарант освітньо-професійної(наукової) програми
(керівник проектної групи)



(підпис)

Олексій КОРЕПАНОВ
(ім'я та прізвище)

Програму погоджено науково- методичною комісією економічного факультету
(назва факультету, для здобувачів вищої освіти якого викладається навчальна дисципліна)

Протокол від «27» серпня 2024 року № 1

Голова науково-методичної комісії економічного факультету


(підпис)

Дар'я ЗАГОРСЬКА
(ім'я та прізвище)

ВСТУП

Програма навчальної дисципліни “**Бізнес-статистика**” укладено відповідно до освітньо-професійної (освітньо-наукової) програми підготовки фахівців другого (магістерського) рівня вищої освіти
(назва рівня вищої освіти, освітньо-кваліфікаційного рівня)

спеціальності (напряму) 051 «Економіка»

1. Опис навчальної дисципліни

1.1. **Мета** вивчення навчальної дисципліни надання студентам поглиблених знань з теорії та практики статистичного аналізу бізнес-процесів і формування навичок щодо застосування в бізнес-аналізі сучасних програмних продуктів MS-Excel, Statistika.

1.2. Основні **завдання** вивчення дисципліни засвоєння теоретичних та методологічних засад статистичного аналізу бізнес-процесів, набуття вмінь застосування ст атистичних методів щодо кількісного оцінювання стану розвитку бізнесу в країні.

Після завершення курсу студент має набути такі **загальні та фахові (предметні) компетентності**:

▪ **Компетентності загальні (ЗК)**

ЗК8. Здатність проводити дослідження на відповідному рівні.

▪ **Компетентності фахові (ФК)**

ФК2. Здатність до професійної комунікації в сфері економіки іноземною мовою.

ФК3. Здатність збирати, аналізувати та обробляти статистичні дані, науково-аналітичні матеріали, які необхідні для розв’язання комплексних економічних проблем, робити на їх основі обґрунтовані висновки.

▪ **Додаткові фахові (предметні) компетентності (ДФК).**

ДФК3. Здатність використовувати професійно профільовані знання й уміння щодо практичного використання комп’ютерних технологій в галузі економічної аналітики та статистики. Здатність формувати ресурсно-інформаційні бази щодо відповідної соціально-економічної ситуації для вирішення професійних аналітико-статистичних завдань.

1.3. Кількість кредитів 4

1.4. Загальна кількість годин 120

1.5. Характеристика навчальної дисципліни	
обов’язкова / вибіркова	
Денна форма навчання	Заочна форма навчання
Рік підготовки	
1-й	1-й
Семестр	
1-й	1-й
Лекції	
16 год.	6 год.
Практичні, семінарські заняття	
16 год.	2 год.
Лабораторні заняття	
год.	год.
Самостійна робота	
88 год.	112 год.
У тому числі індивідуальні завдання	
-	-

1.6. Заплановані **результати навчання**

Згідно з вимогами освітньо-професійної програми студенти повинні досягти таких **результатів навчання**:

Результати навчання (РН) та додаткові (ДРН)

РН1. Формулювати, аналізувати та синтезувати рішення науково-практичних проблем.

РН4. Розробляти соціально-економічні проекти та систему комплексних дій щодо їх реалізації з урахуванням їх цілей, очікуваних соціально-економічних наслідків, ризиків, законодавчих, ресурсних та інших обмежень.

РН7. Обирати ефективні методи управління економічною діяльністю, обґрунтовувати пропонувані рішення на основі релевантних даних та наукових і прикладних досліджень.

РН8. Збирати, обробляти та аналізувати статистичні дані, науково-аналітичні матеріали, необхідні для вирішення комплексних економічних завдань.

РН10. Застосовувати сучасні інформаційні технології та спеціалізоване програмне забезпечення у соціально-економічних дослідженнях та в управлінні соціально-економічними системами.

РН12. Обґрунтовувати управлінські рішення щодо ефективного розвитку суб'єктів господарювання, враховуючи цілі, ресурси, обмеження та ризики.

РН15. Організовувати розробку та реалізацію соціально-економічних проектів із врахуванням інформаційного, методичного, матеріального, фінансового та кадрового забезпечення.

ДРН17. Розглядати проблему наукового дослідження з використанням необхідної бібліографії та відповідних інформаційних джерел; демонструвати розуміння (детальне, системне) сучасних статистичних методів наукового дослідження.

2. Тематичний план навчальної дисципліни

Розділ 1. Статистична методологія етапів аналізу бізнес-процесів

Тема 1. Роль статистики в бізнесі

Чому саме статистика?; Чому необхідно вивчати статистику?; Чи складна статистика?; Чи впливає знання статистики на гнучкість прийняття рішень?; Що таке статистика?: Статистика розглядає загальну картину; Статистика не ігнорує окремі об'єкти; Подивимося на дані; Статистика в менеджменті. Чотири основні етапи статистичного аналізу: Оцінка невідомої величини; Перевірка гіпотез. Що таке ймовірність? Загальна порада. Додатковий матеріал: Резюме; Основні терміни; Контрольні питання; Задачі.

Тема 2. Структури даних. Класифікація, розподіл та перетворення даних

Класифікація різних типів наборів даних. Скільки змінних: Одномірні дані; Двовимірні дані; Багатовимірні дані. Кількісні дані: числа (Дискретні кількісні дані; Безперервні кількісні дані; Остерігайтеся чисел, які не мають змістовну інтерпретацію). Якісні дані: категорії (Порядкові якісні дані; Номінальні якісні дані). Тимчасові ряди і дані в одному часовому зрізі. Джерела даних, включаючи Internet. Додатковий матеріал: Резюме; Основні терміни; Контрольні питання; Задачі; Вправи з використанням даних.

Тема 3. Випадкові величини і закони розподілу як основа бізнес-аналітики

Робота з невизначеними значеннями. Дискретні випадкові величини: Обчислення середнього і стандартного відхилення. Біноміальний розподіл: Визначення біноміального розподілу і біноміального співвідношення; Обчислення середнього і стандартного відхилення: короткий шлях; Обчислення ймовірностей. Нормальний розподіл: Подання ймовірності як площі під кривою; Стандартний нормальний розподіл Z і відповідні ймовірності; Рішення задач на обчислення ймовірності при нормальному розподілі;

Чотири способи обчислення ймовірності; Увага! Не все розподілено нормально! Апроксимація біноміального розподілу нормально. Розподіл Пуассона і експоненціальний розподіл. Додатковий матеріал: Резюме; Основні терміни; Контрольні питання; Задачі; Вправи з використанням даних.

Тема 4. Статистичні гіпотези і статистичні висновки

Перевірка статистичних гіпотез: вибір між реальністю і збігом. Не всі гіпотези однакові: Нульова гіпотеза; Дослідницька гіпотеза; Про що свідчить результат; Приклади гіпотез. Перевірка гіпотези про рівність середнього генеральної сукупності деякого заданого значення: Використання довірчих інтервалів (простий спосіб, t-статистика: спосіб інший, результат той самий). Інтерпретація перевірки гіпотези: Помилки 1 і 11 роду; Умови застосовності; Гіпотези не можуть бути ймовірно істинними або ймовірно помилковими; Статистична значимість і рівні перевірки; Довірча ймовірність (Р-значення). Одностороння перевірка: Як виконати перевірку? Перевірка того, чи належить нове спостереження тієї ж генеральної сукупності? Порівняння двох вибірок: t-тест для залежних вибірок, t-тест для незалежних вибірок. Додатковий матеріал: Резюме; Основні терміни; Контрольні питання; завдання; Вправи з використанням даних.

Тема 5. Дисперсійний аналіз у дослідженні бізнес-процесів

Перевірка відмінностей для декількох вибірок і багато іншого. Використання блокових діаграм для одночасного подання декількох вибірок. F-тест визначає, чи значимо розрізняються середні: Дані та джерела варіації; допущення; гіпотези; F-статистика; F-таблиця; Результат F-тесту; Результат обчислень за допомогою комп'ютера: однофакторний ANOVA-таблиця. Тест найменшої значущої відмінності: які пари розрізняються? Більш складні плани дисперсійного аналізу: Різноманітність - ось, що надає смак життя; Двохфакторний дисперсійний аналіз; Три фактори і більш; Коваріаційний аналіз, ANCOVA; Багатовимірний дисперсійний аналіз (MANOVA); Як читати таблицю ANOVA. Додатковий матеріал: Резюме; Основні терміни; Контрольні питання; Задачі; Вправи з використанням бази даних.

Розділ 2. Статистична оцінка ефективності стану розвитку бізнесу в країні

Тема 6. Непараметричні методи в бізнес-аналізі

Перевірка гіпотез для порядкових даних або даних, що не підкорюються нормальному розподілу. Перевірка гіпотези про рівність медіани деякому заданому значенню: критерій знаків; гіпотези; Допущення. Тестування відмінностей у двох пов'язаних вибірках: Використання критерію знаків для різниць; гіпотези; Умова. Перевірка значущості відмінності двох незалежних вибірок: Процедур, заснована на ранжируванні всіх даних; гіпотези; Допущення. Додатковий матеріал: Резюме; Основні терміни; Контрольні питання; Задачі; Вправи з використанням бази даних.

Тема 7. Аналіз закономірностей для якісних даних (аналіз χ^2 -квадрат)

Узагальнення якісних даних за допомогою частот і відсотків. Перевірка того, що значення відсотків у генеральній сукупності дорівнюють деяким заданим значенням. Критерій « χ^2 -квадрат» щодо рівності відсотків. Перевірка взаємозв'язку між двома якісними змінними. Поняття незалежності змінних. Критерій « χ^2 -квадрат» незалежності. Додатковий матеріал: Резюме; Основні терміни; Контрольні питання; Задачі; Вправи з використанням бази даних.

Тема 8. Статистичні методи аналізу структури сукупності та її змін

Оцінка нерівномірності розподілу, яка вимірюється коефіцієнтами: локалізації та концентрації, а також за аналогією з коефіцієнтом концентрації розраховують коефіцієнт

подібності (схожості) структур. Для оцінка інтенсивності структурних зрушень у часі використовують абсолютні міри варіації – середнє лінійне або середнє квадратичне відхилення часток, які називають коефіцієнтами структурних зрушень. Додатковий матеріал: Резюме; Основні терміни; Контрольні запитання; Задачі.

Тема 9. Статистичне оцінювання ефективності бізнес-процесів

Зміст Визначення економічної ефективності використання ресурсів бізнес-процесів. Оцінка виконання планів, прогнозів, управлінських рішень, за ефективним використанням економічного потенціалу бізнес-процесів підприємства. Вивчення впливу об'єктивних та суб'єктивних, зовнішніх і внутрішніх чинників на результати бізнес-процесів підприємства чи фірми. Розробка й обґрунтування заходів, спрямованих на оптимізацію бізнес-процесів підприємства і прийняття ефективних управлінських рішень. Додатковий матеріал: Резюме; Основні терміни; Контрольні запитання; Задачі.

Назви розділів і тем	Кількість годин											
	денна форма						заочна форма					
	усь ого	у тому числі					усього	у тому числі				
		л	п	л	інд.	с. р.		л	п	лаб.	інд.	с. р.
1	2	3	4	5	6	7	8	9	10	11	12	13
Розділ 1. Статистична методологія етапів аналізу бізнес-процесів												
Тема 1.2.3. Роль статистики в бізнесі. Структури даних. Класифікація, розподіл та перетворення даних. Випадкові величини і закони розподілу як основа бізнес-аналітики.	30	4	4			22	29,3	1	0,3			28
Тема 4.5. Статистичні гіпотези і статистичні висновки. Дисперсійний аналіз у дослідженні бізнес-процесів.	30	4	4			22	30,7	2	0,7			28
Разом за розділом 1	60	8	8			44	60	3	1			56
Розділ 2. Статистична оцінка ефективності стану розвитку бізнесу в країні												
Тема 6.7.8. Непараметричні методи в бізнес-аналізі. Аналіз закономірностей для якісних даних (аналіз хі-квадрат). Статистичні методи аналізу структури сукупності та її змін.	30	4	4			22	29,5	1	0,5			28
Тема 9. Статистичне оцінювання ефективності бізнес-процесів.	30	4	4			22	30,5	2	0,5			28
Разом за розділом 2	60	8	8			44	60	3	1			56
Усього годин	120	16	16			88	120	6	2			112

4. Теми семінарських (практичних, лабораторних) занять

№ з/п	Назва теми	Кількість годин	
		Денне	Заочне
1	Кількісне оцінювання стану розвитку бізнесу в країні	1	0,1
2	Основні етапи статистичного аналізу бізнесу	1	0,1
3	Класифікації та структури даних	1	0,2
4	Способи перетворення даних	1	0,2
5	Статистичні гіпотези та критерії їх перевірки для різних даних	2	0,2
6	Дисперсійний аналіз	2	0,3
7	Непараметричні методи в бізнесі. Аналіз хі-квадрат	2	0,3
8	Статистичні методи аналізу структури сукупності	2	0,3
9	Статистична оцінка ефективності бізнесу	4	0,3
	Разом	16	2

5. Завдання для самостійної роботи

№ з/п	Види, зміст самостійної роботи	Кількість годин	
		Денне	Заочне
1	Теоретичні знання: класифікувати дані, що характеризують конкретний бізнес-процес	20	28
2	Виконання індивідуальних практичних завдань: формулювати нульову та альтернативну гіпотези відповідно до мети статистичного аналізу	20	28
3	Контрольні завдання: застосовувати методи одновимірного та багатовимірного дисперсійного аналізу	18	28
4	Написання контрольної роботи у вигляді тестів або реферату(есе) на тему «Статистичний аналіз бізнес ситуацій» з використанням: статистичної методології для оцінювання ефективності бізнесу та Internet ресурсів, який включає пакети прикладних програм для розв'язання задач кількісного оцінювання бізнес-процесів	30	28
	Разом	88	112

6. Індивідуальні завдання:

-

7. Методи навчання

При викладанні навчальної дисципліни «Бізнес-статистика» використовуються інформаційно-ілюстративний та проблемний методи навчання із застосуванням:

- лекцій з презентаціями ;
- відеофільми;
- розв'язування задач та кейсів MS- EXCEL, Statistika;
- складання графічних схем;
- виконання індивідуальних суцільних та вибірових статистичних спостережень;
- презентація контрольної роботи «Статистичний аналіз бізнес ситуацій» з використанням англійської мови;
- апробація наукових досліджень на Всеукраїнських та Міжнародних науково-практичних конференціях, а також публікації статей у фахових електронних виданнях.

8. Методи контролю

Основними завданнями контролю знань з дисципліни «Бізнес-статистика» є оцінювання засвоєння ними теоретичних знань і навичок, набутих під час навчання. При цьому, контрольні заходи мають:

- стимулювати систематичну самостійну індивідуальну роботу над навчальним матеріалом;
- забезпечити закріплення та реалізацію набутих теоретичних знань при проведенні семінарських занять та контрольної роботи у вигляді тестів або реферату(есе) на тему: «Статистичний аналіз бізнес ситуацій»;
- прищепити навички відповідального ставлення до своїх обов'язків, самостійного цілеспрямованого пошуку потрібної інформації;
- згідно з навчальним планом з дисципліни «Бізнес статистика» передбачена форма підсумкового контролю – залік. Оцінювання знань здійснюється за 100-бальною дворівневою шкалою.

Критерії оцінювання знань для дисципліни «Бізнес-статистика»

Порядок оцінювання знань студентів, розроблений: і впроваджений в навчальний процес у Харківському національному університеті імені В.Н.Каразіна унормовує основні принципи організації поточного, підсумкового та семестрового контролю знань студентів з усіх видів виконуваних робіт (теоретична підготовка, практична робота, самостійна робота, контрольна робота, наукова робота тощо), а також методику переведення показників академічної успішності студентів ХНУ імені В.Н.Каразіна в національну шкалу.

Контрольні заходи є необхідним елементом зворотного зв'язку у процесі навчання. Вони визначають відповідність рівня набутих студентами знань, умінь та навичок вимогам нормативних документів щодо вищої освіти.

Поточний контроль знань студентів

Система поточного контролю знань студентів передбачає оцінювання їх за трьома основними напрямками:

- а) перевірка теоретичних знань;
- б) перевірка практичних навичок (виконання контрольної роботи).

Оцінювання рівня засвоєння теоретичних знань може проводитись в одній з альтернативних форм:

- за підсумками тестування ;
- під час проведення усної співбесіди зі студентом, в процесі якої студент дає відповіді на питання з різних тем дисципліни.

Проміжний контроль рівня знань (тестування) передбачає виявлення опанування студентом лекційного матеріалу та вміння застосування його для вирішення практичної ситуації і проводиться у вигляді тестування.

Під час проведення усної співбесіди можуть міститися як запитання, що стосуються суто теоретичного матеріалу, так і запитання, спрямовані на вирішення невеличкого практичного завдання.

Оцінювання рівня набутих практичних навичок здійснюється за результатами виконання контрольної роботи(у вигляді тестів) або реферату(есе) на тему: «Статистичний аналіз бізнес ситуацій».

9. Схема нарахування балів та критерії оцінювання навчальних досягнень

Структура засобів контролю та розподіл балів із дисципліни «Бізнес-статистика» наведена у наступних таблицях:

Узагальнена схема нарахування балів (денна форма навчання)

Поточний контроль, самостійна робота, індивідуальні завдання													Залікова робота	Сума
Поточне оцінювання (проміжний контроль)									Контрольна робота, перед- бачена навчаль- ним планом	Індиві- дуальне зав- дання	Разом			
Розділ 1					Розділ 2									
T1	T2	T3	T4	T5	T6	T7	T8	T9						
4	4	4	4	4	5	5	5	5	20	-	60	40	100	

T1, T2 ... T9. – теми розділів.

Узагальнена схема нарахування балів (заочна форма навчання)

Поточний контроль, самостійна робота, індивідуальні завдання													Залікова робота	Сума
Поточне оцінювання (проміжний контроль)					Контрольна робота, перед- бачена навчаль- ним планом				Індиві- дуальне зав- дання	Разом				
Розділ 1		Розділ 2												
T1	T2	T3	T4	T5	T6	T7	T8	T9						
4	4	4	4	4	5	5	5	5	20	-	60	40	100	

T1, T2 ... T9. – теми розділів.

Для допуску до складання підсумкового заліку, здобувач вищої освіти повинен набрати не менше **15 балів** з навчальної дисципліни «Бізнес-статистика» під час поточного контролю, самостійної роботи та контрольної роботи.

Проміжний тестовий контроль проводиться два рази на семестр (в рамках вивчення Розділів). Загальна тривалість тестів за Розділами – по 1,5 години. Поточні тестування складаються з 10 тестів і 1 задачі. Одна правильна відповідь на кожен з тестів дорівнює 2 балу, за задачу – 10 балів. Тестове завдання містить запитання одиничного і множинного вибору різного рівня складності. Тести можуть бути застосовані як з метою контролю, так і для закріплення теоретичних знань і практичних навичок. Тести для проміжного контролю обираються із загального переліку тестів за відповідними темами.

Підсумковий контроль за курсом – у формі заліку

До складання заліку допускають студентів, що мають задовільну кількість балів зі складених тестів з основних навчальних двох розділів курсу, написання контрольної роботи та інших завдань, передбачених програмою дисципліни.

Залік здійснюють за заліковими завданнями, які містять теоретичну частину, в якій одне теоретичне питання та два тестових завдання і практичну частину, у складі якого передбачено розв'язання та аналітична оцінка конкретної бізнес ситуації. Саме вони дають можливість здійснити оцінювання знань студента за дисципліною «Бізнес-статистика».

Критерії оцінювання навчальних досягнень

Критерії оцінювання результативності роботи студентів при виконанні самостійної роботи

Якісними критеріями оцінювання виконання індивідуальних завдань студентами є:

1. Повнота виконання завдання: елементарна; фрагментарна; повна; неповна.

2. *Рівень самостійності студента:* під керівництвом викладача; консультація викладача;самостійно.

3. *Сформованість навчально-інформаційних умінь:* роботи з підручником, володіння різними способами читання, складання плану, рецензій, конспекту, вміння користуватися бібліотекою, спостереження, експеримент тощо.

4. *Сформованість навчально-інтелектуальних умінь:* визначення понять, аналіз, синтез, порівняння, класифікація, систематизація, узагальнення, абстрагування, вміння відповідати на запитання, виконувати творчі завдання тощо.

5. *Рівень сформованості фахових методичних вмінь (вміння застосовувати на практиці набуті знання):* низький – володіння умінням здійснювати первинну обробку навчальної інформації без подальшого її аналізу; середній – уміння вибирати відомі способи дій для виконання фахових завдань; достатній – застосовує набуті знання у стандартних практичних ситуаціях; високий – володіння умінням творчо-пошукової діяльності.

Критерії оцінювання здобувачів вищої освіти за відповіді на питання поточного контролю

Завданням поточного контролю є перевірка розуміння та засвоєння певного матеріалу, вироблених навичок проведення розрахункових робіт, умінь самостійно опрацьовувати тексти, здатності осмислити зміст теми чи розділу, умінь публічно чи письмово представити певний матеріал (презентація).

Якісними критеріями оцінювання виконання завдань поточного контролю є:

1. *Повнота відповіді або виконання завдання:* елементарна; фрагментарна; повна; неповна.

2. *Рівень сформованості логічних умінь:* елементарні дії; операція, правило, алгоритм; правила визначення понять; формулювання законів і закономірностей; структурування суджень, доводів, описів.

Критерії оцінювання здобувачів вищої освіти за виконання контрольної роботи

Оцінювання рівня набутих практичних навичок здійснюється за результатами виконання та захисту контрольної роботи в вигляді тестів.

Контрольна робота студента складається з 4-ох розділів завдання. Максимальна оцінка за виконання завдань – 20 балів.

Схема нарахування балів за виконання контрольної роботи

Кількість балів				Разом
3-1	3-2	3-3	3-4	20
5	5	5	5	

3-1, 3-2 ... 3-4 – складові контрольної роботи

Схема нарахування балів за контрольною роботою у вигляді реферату(есе)

№ п/п	Розділи завдання	max бал
1	Постановка проблеми та актуальність	5
2	Науково обґрунтований алгоритм збору, обробки та поширення інформації	5
3	Аналітична оцінка результатів дослідження	5
4	Бонусні бали за використання англійської мови	5
	Усього	20

Критерії оцінювання залікової роботи

Оцінювання навчальних досягнень студентів здійснюється за 100-бальною дворівневою («зараховано», «не зараховано») шкалою." За наслідками заліку студент може отримати максимум 40 балів:

1. Для отримання оцінки зараховано «відмінно» (35-40 балів) студент повинен: укластися у встановлений строк підготовки відповіді; викласти теоретичний матеріал чітко, коротко, зв'язано й обґрунтовано; навести вірне рішення задачі та тестів.

2. Для одержання оцінки зараховано «добре» (25-34 балів) студент повинен: укластися у встановлений строк підготовки відповіді; викласти теоретичний матеріал обґрунтовано й зв'язано; навести вірне рішення задачі; можливі помилки при відповіді на тести.

3. Для отримання оцінки зараховано «задовільно» (15-24 балів) студент повинен: викладати теоретичний матеріал у доступній для розуміння формі; можливі помилки при розв'язанні задачі та тестів.

4. Оцінку не зараховано «незадовільно» (1-14 балів) отримують студенти, відповіді яких можуть бути оцінені нижче вимог, сформульованих у попередньому пункті.

Шкала оцінювання залікової роботи:

- за правильну відповідь на теоретичне питання студент одержує - 20 балів;
- за правильно виконане 2-х тестових завдань – 20 балів;

Підсумкова оцінка з навчальної дисципліни визначається як загальна оцінка, яка враховує оцінки з кожного виду контролю (оцінки проміжного контролю за роботу протягом семестру та оцінка за результатами виконання залікового завдання).

Зведена шкала оцінювання роботи студентів з дисципліни «Бізнес-статистика»

Види робіт	Максимум балів
Поточне оцінювання	40
Контрольна робота	20
РАЗОМ	60
<i>Залікове завдання</i>	40
УСЬОГО	100

У відповідності до набраних студентом – бакалавром балів оцінка знання матеріалу проводиться за дворівневою шкалою оцінювання згідно з Методикою переведення показників успішності знань студентів.

Шкала оцінювання

Сума балів за усі види навчальної діяльності протягом семестру	Оцінка
	для дворівневої шкали оцінювання
90–100	зараховано
70–89	
50–69	
1–49	не зараховано

10. Рекомендована література (основна, допоміжна)

Основна література

1. **Ануфрієва О. Л., Пальчевська Т. Г., Лагоцька Г. М.** Підприємницька діяльність. Навчальний посібник. Івано-Франківськ: «Лілея-НВ», 2014. 304 с.
2. **Бізнес-статистика.** Навч. посібник / Матковський С. О., Гринькевич О. С. [та ін.]. К.: Алерта, 2016. 280 с.
3. **Бізнес-статистика.** Підручник / К. Д. Семенова, К. І. Тарасова. К.: ФОП Гуляєва В. М. 2018. 210 с.
4. **Кущенко О.І.** Бізнес-статистика: навчально-методичний посібник українською та англійською мовами. Харків: ХНУ імені В.Н. Каразіна, 2019. 116 с.
5. **Кущенко О.І.** Бізнес-статистика: дистанційний курс: навчально-методична праця. Сертифікат № 374/2023. Рішення вченої ради університету(науково-методичної ради)_Протокол № 8 від 16.05.2023 року. URL: <http://www.dist.karazin.ua>
6. **Міжнародний бізнес.** Підручник./ за ред. проф. В. А. Вергуна. К.: ВАДЕКС, 2014. 810 с.
7. **Міжнародний бізнес.** Навчальний посібник / Файчук О. М., Сидоренко С. В. К.: НУБіП України, 2016. 236 с.
8. **Основи організації бізнесу.** Навчальний посібник / [Мельников А.](#) М. ЦУЛ, 2017. 200 с.
9. **Положення про Державну службу статистики України,** Київ, 06.03.2011, № 396/2011
10. **Економічна статистика:** навчальний посібник, 2-ге вид., перероб. та доп. / О.О. Нестеренко, О.С. Корепанов, Т. Г. Чала та ін. ; за ред. О. О. Нестеренко, В. М. Соболева, Т. Г. Чалої. Х.: ХНУ імені В. Н. Каразіна, 2023. 448 с. **Кущенко О.І. Розділ 16.** Митна статистика (С. 405–425), **Розділ 17.** Статистика туризму (С.426-438).
11. **Business Statistics.** Dr. S.M. Shukla and Dr. S.P. Sahai URL: <https://sahityabhawanpublications.com/product/business-statistics>
12. **Devid Spiegelhalter The Art of Statistics: Learning from Data.** (2020). Penguin, P. 448.
13. **Kushchenko O.I.** Sources of development business of education Ukraine(Витоки розвитку бізнес освіти в Україні). *Східна Європа: економіка, бізнес та управління: Електронне наукове фахове видання.* Випуск 4(15) 2018. URL: <http://www.easterneurope-ebm.in.ua/15-2018-ukr>.
14. **Kushchenko O.I.** Possibilities of the eKhNUIR use of statistical methodology are in educational process. ISSN 2310-8665. [eKhNUIR](#). Kharkiv. 2019. URI: <http://dspace.univer.kharkov.ua/handle/123456789/14605>
15. **Kushchenko E.I. Methodological bases of statistical study of seasonality of market of tourist product.** *The world of science and innovation.* Abstracts of the 1 st International scientific and practical conference. Cognum Publishing House. London, United Kingdom. 2020. Pp. 44-48. URL: <https://sci-conf.com.ua/i-mezhdunarodnaya-nauchno-prakticheskaya-konferentsiya-the-world-of-science-and-innovation-19-21-avgusta-2020-goda-london-velikobritaniya-arhiv>;
16. **Olena Kushchenko, Larisa Izotova.** Analytical methodology of international comparisons of the European Union Countries (Аналітична методологія міжнародних співставлень країн Європейського союзу). *Комерціалізація інновацій: захист інтелектуального капіталу, маркетинг та комунікації монографія* / за ред. к.е.н., доц. Сагер Л.Ю., к.е.н., доц. Сигиди Л.О. Суми: Сумський державний університет, 2022. С. 133-140. URL: <http://dspace.univer.kharkov.ua/handle/123456789/17621>
17. **Olena Kushchenko.** Theoretical and methodological bases of custom market research. Development of Ukraine's customs policy in the context of European integration: Scientific

- monograph. Riga, Latvia: Baltija Publishing, 2023. 276 p.C.130-148.
URL: <http://www.baltijapublishing.lv/omp/index.php/bp>
18. **Practical Business Statistics** Andrew F. Siegel, Michael R. Wagner. (2016). P. 642
19. **Practical Problems in Business Statistics** Dr. S.M. Shukla, Dr. Shiv Pujan Sahai URL: <https://www.flipkart.com/practical-problems-business-statistics/p/itm4mekkf2xehbk>

Допоміжна література

1. **Кущенко О.І., Воловод А.А.** Методологічні засади бізнес аналітичного прогнозування ринку послуг. Електронне наукове фахове видання «Економіка та суспільство». 2021 №29
URL: <https://economyandsociety.in.ua/index.php/journal/article/view/559>
DOI: <https://doi.org/10.32782/2524-0072/2021-29-24>
2. **Кущенко О. І., Гаврюшенко А.А.** Статистичні бази даних як джерело інформації щодо проведення аналітичних досліджень. Technologies and strategies for the implementation of scientific achievements: collection of scientific papers «SCIENTIA» with Proceedings of the I International Scientific and Theoretical Conference (Vol.1), May 27, 2022. Stockholm, Kingdom of Sweden: European Scientific Platform.C.60-64. URL: <https://ojs.ukrlogos.in.ua/index.php/scientia/issue/view/27.05.2022/746>
3. **Кущенко О. І., Оніщенко К.** Аналітика сталого розвитку підприємств. *Transformacje cyfrowe i technologie innowacyjne w ekonomii*. Część 2: międzynarodowej naukowo-praktycznej konferencji internetowej (14–15.03.2024, Łomża-Charków): zbiór materiałów = Цифрові трансформації та інноваційні технології в економіці. Ч. 2: міжнародної науково-практичної інтернет-конференції (14–15.03.2024 Ломжа-Харків): збірник матеріалів. Ломжа: MANS, 2024. С. 118–122. URI: <https://ekhnuir.karazin.ua/handle/123456789/18410> Зібрання [Наукові роботи студентів та аспірантів](#). Економічний факультет.
4. **Кущенко О. І., Оніщенко К.** Сталий розвиток підприємств: стратегічне управління та інноваційні технології в економіці України: Презентація. Харків: Харківський національний університет імені В.Н. Каразіна, 2024. 17 слайдів. URI: <https://ekhnuir.karazin.ua/handle/123456789/18417>. Зібрання [Презентації](#). [Економічний факультет](#)
5. **Сидорова А. В., Біленко Д. В., Буркіна Н. В.** Бізнес-аналітика: навчально-методичний посібник. Вінниця: ДонНУ імені Василя Стуса. 2019. 104 с.
6. **Управління міжнародним бізнесом.** Конспект лекцій для студентів економічного спрямування, аспірантів, викладачів. / Н. Г. Ушакова, К. Ю. Величко, О. І. Печенка. Харків: Видавництво «Форт», 2016. 126 с.

11. Посилання на інформаційні ресурси в Інтернеті, відео-лекції, інше методичне забезпечення

1. Головне управління статистики у Харківській області
URL: <http://uprstat.kharkov.ukrtel.net>
2. Державна служба статистики України URL: <http://www.ukrstat.gov.ua>
3. Огляд діяльності Євростату. URL: <https://ec.europa.eu/eurostat/about/overview>
4. Ukrstat.org – публікація документів Державної служби статистики України. URL: <http://ukrstat.org/about.html>
5. International Association for Official Statistics – IAOS. URL: <https://www.iaos-isi.org/>
6. International statistical institute. URL: <http://xn--h1ahbi.com.ua/info/150-isi-international-statistical-institute-mezhdunarodnyy-statisticheskiy-institut-msi.html>
7. Statistical office of the European Union. URL: <https://ec.europa.eu/eurostat/home>
8. United Nations Statistical Commission. URL: <https://unstats.un.org/unsd/statcom/>

12. Особливості навчання за денною (заочною) формами в умовах дії обставин непоборної сили

В умовах дії воєнного стану, освітній процес в університеті здійснюється:

– дистанційно (за затвердженим електронним розкладом занять у особистому кабінеті) на платформі **ZOOM** проводяться усі лекційні заняття:

- дистанційно на платформі **Moodle**. URL: <https://dist.karazin.ua/moodle/course/view.php?id=2729> проводяться практичні (семінарські), індивідуальні заняття та консультації, контроль самостійної роботи.

Складання підсумкового семестрового контролю: надається можливість скласти **екзамен дистанційно на платформі Moodle** в дистанційному курсі «Бізнес-статистика», URL: <https://dist.karazin.ua/moodle/course/view.php?id=2729>

СИЛАБУС

ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМ. В.Н. КАРАЗІНА Економічний факультет



Кафедра статистики, обліку та аудиту СИЛАБУС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ БІЗНЕС-СТАТИСТИКА

Рівень вищої освіти	–	другий (магістерський)
Галузь знань	–	05 «Соціальні та поведінкові науки»
Спеціальність	–	051 «Економіка»
Освітньо-професійна програма	–	Економічна аналітика та статистика
Рік підготовки, семестр	–	1 рік, 2 семестр
Вид дисципліни	–	Обов'язкова
Сторінка дисципліни на платформі LMS Moodle	–	https://moodle.karazin.ua/course/view.php?id=2729

ВИКЛАДАЧІ



канд.екон.наук(PhD), доцент ЗВО кафедри статистики, обліку та аудиту

Кушченко Олена Іванівна o.kushchenko@karazin.ua,
+380503239415

Досвід роботи – 47 років. Автор понад 135 наукових та навчально - методичних праць.

Провідний лектор з дисциплін: «Вступ до фаху», «Економічна та бізнес-статистика», «Статистика ринку товарів та послуг», «Статистика туризму», «Бізнес-статистика», «Статистичне забезпечення управління економікою», «Аналіз банківської діяльності» та ін.

ОБСЯГ ДИСЦИПЛІНИ

(денна та заочна форми навчання)

Всього годин / кредитів ЄКТС	Аудиторні заняття, год				Самостійна робота		Вид підсумкового контролю
	Лекції		Практичні заняття				
120/4	Денна	Заочна	Денна	Заочна	Денна	Заочна	Залік
	16	6	16	2	88	112	

ЗАГАЛЬНА ІНФОРМАЦІЯ ПРО ДИСЦИПЛІНУ

Мета дисципліни	Надання студентам поглиблених знань з теорії та практики статистичного аналізу бізнес-процесів і формування навичок щодо застосування у бізнес-аналізі сучасних програмних продуктів MS-Excel, Statistika, SPSS
Завдання дисципліни	Засвоєння теоретичних та методологічних засад статистичного аналізу бізнес-процесів, набуття вмінь застосування статистичних методів для кількісного оцінювання стану розвитку бізнесу в країні
Пререквізити	Загальні та фахові знання, отримані на першому (бакалаврському) рівні вищої освіти зі «Вступу до фаху», «Макроекономіки», «Статистики», «Основи підприємництва та міжнародного бізнесу», «Маркетингу», «Інформатика», «Візуалізація даних» «Управління ризиками»
Компетенції відповідно до освітньо-професійної програми	
Загальні компетентності (ЗК)	Фахові компетентності (ФК)
ЗК8. Здатність проводити дослідження на відповідному рівні.	<p>ФК2. Здатність до професійної комунікації в сфері економіки іноземною мовою.</p> <p>ФК3. Здатність збирати, аналізувати та обробляти статистичні дані, науково-аналітичні матеріали, які необхідні для розв'язання комплексних економічних проблем, робити на їх основі обґрунтовані висновки.</p> <p>Додаткові фахові (предметні) компетентності (ДФК).</p> <p>ДФК3. Здатність використовувати професійно профільовані знання й уміння щодо практичного використання комп'ютерних технологій в галузі економічної аналітики та статистики. Здатність формувати ресурсно-інформаційні бази щодо відповідної соціально-економічної ситуації для вирішення професійних аналітико-статистичних завдань.</p>
Результати навчання (РН) та додаткові (ДРН)	
<p>РН1. Формулювати, аналізувати та синтезувати рішення науково-практичних проблем.</p> <p>РН4. Розробляти соціально-економічні проекти та систему комплексних дій щодо їх реалізації з урахуванням їх цілей, очікуваних соціально-економічних наслідків, ризиків, законодавчих, ресурсних та інших обмежень.</p> <p>РН7. Обирати ефективні методи управління економічною діяльністю, обґрунтовувати пропонувані рішення на основі релевантних даних та наукових і прикладних досліджень.</p> <p>РН8. Збирати, обробляти та аналізувати статистичні дані, науково-аналітичні матеріали, необхідні для вирішення комплексних економічних завдань.</p> <p>РН10. Застосовувати сучасні інформаційні технології та спеціалізоване програмне забезпечення у соціально-економічних дослідженнях та в управлінні соціально-економічними системами.</p> <p>РН12. Обґрунтовувати управлінські рішення щодо ефективного розвитку суб'єктів господарювання, враховуючи цілі, ресурси, обмеження та ризики.</p> <p>РН15. Організовувати розробку та реалізацію соціально-економічних проектів із врахуванням інформаційного, методичного, матеріального, фінансового та кадрового забезпечення.</p> <p>ДРН17. Розглядати проблему наукового дослідження з використанням необхідної бібліографії та відповідних інформаційних джерел; демонструвати розуміння (детальне,</p>	

системне) сучасних статистичних методів наукового дослідження.

СТРУКТУРА ДИСЦИПЛІНИ

Назви розділів і тем	Кількість годин											
	денна форма						заочна форма					
	усього	у тому числі					усього	у тому числі				
		л	п	лаб.	інд.	с.р.		л	п	лаб.	дист.	с.р.
1	2	3	4	5	6	7	8	9	10	11	12	13
РОЗДІЛ 1. Статистична методологія етапів аналізу бізнес-процесів												
Тема 1.2.3. Роль статистики в - бізнесі. Структури даних. Класифікація, розподіл та перетворення даних. Випадкові величини і закони розподілу як основа бізнес-аналітики	30	4	4	-	-	22	29,3	1	0,3			28
Тема 4.5. Статистичні гіпотези і статистичні висновки. Дисперсійний аналіз у дослідженні бізнес-процесів	30	4	4	-	-	22	30,7	2	0,7			28
Разом за розділом 1	60	8	8	-	-	44	60	3	1			56
РОЗДІЛ 2. Статистична оцінка ефективності стану розвитку бізнесу в країні												
Тема 6.7.8. Непараметричні методи в бізнес-аналізі. Аналіз закономірностей для якісних даних (аналіз хі-квадрат). Статистичні методи аналізу структури сукупності та її змін.	30	4	4	-	-	22	29,5	1	0,5			28
Тема 9.10. Статистичні методи контролю якості. Статистичне оцінювання ефективності бізнес-процесів.	30	4	4	-	-	22	30,5	2	0,5			28
Разом за розділом 2	60	8	8	-	-	44	60	3	1			56
Усього годин	120	16	16	-	-	88	120	6	2			112

РЕЗУЛЬТАТИ НАВЧАННЯ

Здобувачі другого рівня вищої освіти магістри повинні набути таких **результатів навчання**:

знати: роль статистики в бізнесі; основні етапи статистичного аналізу бізнес-процесів; класифікації та структури даних; способи перетворення даних; основні закони розподілу випадкових величин; сутність статистичних гіпотез та необхідність їх перевірки; критерії перевірки статистичних гіпотез для різних типів даних; основи дисперсійного аналізу; статистичні методи аналізу структури сукупності та її змін; методи статистичного оцінювання ефективності бізнес-процесів.

вміти: класифікувати дані, що характеризують конкретний бізнес-процес; перетворювати первинні дані щодо забезпечення їх зі ставності та відповідності вимогам статистичного аналізу; формулювати нульову та альтернативну гіпотези відповідно до мети статистичного аналізу; обґрунтовувати вибір критеріїв для перевірки статистичних гіпотез; застосовувати методи одновимірної та багатовимірної дисперсійного аналізу; здійснювати порівняльний аналіз структур; здійснювати статистичне оцінювання змін у структурі сукупності; використовувати статистичні методи оцінювання ефективності бізнес-процесів; застосовувати пакети прикладних програм для розв'язання задач кількісного оцінювання бізнес-процесів.

СИСТЕМА ОЦІНЮВАННЯ ТА КОНТРОЛЮ РЕЗУЛЬТАТІВ НАВЧАННЯ

Оцінка знань та умінь з дисципліни здійснюється за даними поточного та підсумкового контролю. Форми оцінювання поточної навчальної діяльності включає контроль теоретичної та практичної підготовки під час проведення навчальних занять. Підсумковий контроль проводиться у формі заліку.

Рейтинг студента визначається за 100 бальною шкалою і складається з оцінки поточного контролю (max 60 балів) і підсумкового контролю (max 40 балів) відповідно до «Положення про організацію освітнього процесу у Харківському національному університеті імені В.Н. Каразіна».

Зведена шкала оцінювання роботи студентів з дисципліни «Бізнес-статистика»

Види робіт	Максимум балів
Поточне оцінювання	40
Контрольна робота	20
РАЗОМ	60
<i>Залік</i>	40
УСЬОГО	100

ПОЛІТИКА ДИСЦИПЛІНИ

Інструменти та обладнання. Спеціалізовані інформаційні системи і програмні продукти. Система дистанційного навчання і контролю LMS Moodle за посиланням: <https://moodle.karazin.ua/course/view.php?id=2729>.

Дедлайни. Усі завдання, передбачені програмою, мають бути виконані у встановлені терміни. Роботи, які здаються із порушенням термінів без поважних причин, оцінюються на нижчу оцінку (75% від можливої максимальної кількості балів за вид діяльності балів).

Академічна доброчесність. Учасники освітнього процесу керуються принципами академічної доброчесності згідно положень про «Порядок проведення перевірки кваліфікаційних робіт, наукових праць та навчальних видань щодо наявності запозичень з інших документів»

https://karazin.ua/storage/documents/552_N7Fu8UFiFvAjGBQhM08p5U5ww.pdf та «Положення про систему внутрішнього забезпечення якості освіти Харківського національного університету імені В. Н. Каразіна» <http://rada.karazin.ua/public/uploads/2020/06.pdf>).

Відвідування. Відвідування занять є обов'язковим компонентом оцінювання, за яке нараховуються бали. В умовах дії форс-мажорних обмежень освітній процес в університеті здійснюється відповідно до наказів/ розпоряджень ректора/ проректора або за змішаною формою навчання або повністю дистанційно у синхронному режимі.

РЕКОМЕНДОВАНА ЛІТЕРАТУРА

1. Кущенко О.І. Бізнес-статистика: навчально-методичний посібник українською та англійською мовами. Харків: ХНУ імені В.Н. Каразіна, 2019. 116 с.

2. Kushchenko Olena Statistical methodology of analytic geometry of Green economy. «Зелена економіка» – необхідна умова сталого розвитку національної економіки України: колективна монографія / За заг. ред. О.Л. Гальцової. Херсон: Видавничий дім «Гальветика», 2020. С.34-51. URI: <http://dspace.univer.kharkov.ua/handle/123456789/17620>

3. Кущенко Олена, Душина Діана, Інформаційні можливості платформи Alteryx щодо оптимізації ефективного управління ринком. *Obszary rozwoju współczesnego biznesu*. Monograph. Redakcja Dorota Burzyńska Piotr Idrian. Skierniewice, 2022. PP.162-169. URL: <http://dspace.univer.kharkov.ua/handle/123456789/17619>

4. Olena Kushchenko, Larisa Izotova. Analytical methodology of international comparisons of the European Union Countries (Аналітична методологія міжнародних співставлень країн Європейського союзу). *Комерціалізація інновацій: захист інтелектуального капіталу, маркетинг та комунікації монографія* / за ред. к.е.н., доц. Сагер Л.Ю., к.е.н., доц. Сигиди Л.О. Суми: Сумський державний університет, 2022. С. 133-140. URL: <http://dspace.univer.kharkov.ua/handle/123456789/17621>

5. Olena Kushchenko. Theoretical and methodological bases of custom market research. Development of Ukraine's customs policy in the context of European integration: Scientific monograph. Riga, Latvia: Baltija Publishing, 2023. 276 p. С.130-148. URL: <http://www.baltijapublishing.lv/omp/index.php/bp>

СЕРТИФІКАТ



2. НАВЧАЛЬНИЙ КОНТЕНТ

РОЗДІЛ 1

СТАТИСТИЧНА МЕТОДОЛГІЯ ЕТАПІВ АНАЛІЗУ БІЗНЕС- ПРОЦЕСІВ

1.1. Структури даних та випадкові величини і закони розподілу як основа бізнес-аналітики

*Структури даних. Класифікація, розподіл та перетворення даних.
Основні переваги статистичних обстежень бізнес активності підприємств.
Випадкові величини і закони розподілу як основа бізнес-статистики.*

Структури даних.

Статистичні дані можуть бути представлені в різних формах. Набори даних мають в собі одне чи декілька значень для кожного окремого об'єкту. В якості таких об'єктів можуть виступати люди, міста, комп'ютери, книги і все, що має інтерес для вивчення. Такі об'єкти називають *елементарними одиницями*. Для кожного об'єкту реєструють одні і ті самі ознаки. Наприклад, реєструють ріст і масу людей; чисельність населення, рівень народжуваності і смертності для окремих міст; об'єм пам'яті, швидкість роботи комп'ютеру; рік видання, кількість сторінок, видавництво книг і т.п. Ознака, що реєструється для кожного з об'єктів називають *змінною*.

Класифікація, розподіл та перетворення даних

Набори даних класифікують за наступними ознаками:

- За кількістю змінних(одномірні, двомірні, багатомірні);
- За типом даних(кількісні, якісні);
- За тим, чи потрібна впорядкованість даних у часі чи ні.

Одномірні набори даних містять тільки одну ознаку для кожного об'єкта. Ці дані дозволяють визначити типове значення ознаки, наскільки значення відрізняються один від одного, вимагають чи окремі дані на особливу увагу. Прикладом одновимірних даних є інформація про середню зарплату в регіоні по галузям. Вона дозволяє назвати галузь з найвищим рівнем зарплати, зрозуміти, наскільки відрізняються рівні середніх зарплат в

різних галузях друг від друга, звернути увагу на галузі, де рівень зарплати найнижчий.

Набори *двовимірних* даних містять інформацію про двох ознаках для кожного з об'єктів. Крім того, що вони дають можливість отримати два набори *одновимірних* даних, двовимірні дані дозволяють встановити, чи існує зв'язок між двома змінними, наскільки сильно пов'язані змінні, чи можна передбачити значення однієї змінної за значенням іншого і якщо так, то з якою надійністю. Наприклад, дані опитування студентів про те, чи задоволені вони рівнем теоретичної і практичної підготовки, одержуваної у вузі (значення обох змінних записуються у вигляді так / ні, або 1/0), дозволяють встановити, чи є зв'язок між рівнями теоретичної і практичної підготовки в вузі.

Багатовимірні дані містять інформацію про три або більше ознаках для кожного об'єкта. На додаток до тієї інформації, яку можна отримати з *одновимірних* і *двовимірних* наборів, багатовимірні дані можна використовувати для отримання інформації про те, чи існує проста залежність між цими ознаками, наскільки вони взаємопов'язані (мова йде не тільки про попарний взаємозв'язок ознак, але і про залежність в сукупності), чи можна передбачити значення однієї змінної на підставі значень інших. Прикладом багатовимірних даних може бути число іноземних громадян, які навчаються в ХНУ на певну дату, по країнам.

Значення змінних, які реєструються за допомогою чисел, що мають змістовний сенс, називають кількісними даними. З кількісними даними можна виконувати всі звичайні операції над числами, такі, як обчислення середнього і оцінку мінливості.

Залежно від того, які значення може потенційно приймати змінна, виділяють два типи кількісних даних: *дискретні* і *безперервні*.

Дискретна - це така змінна, яка може набувати значень тільки з деякого списку певних чисел. Прикладами дискретної змінної є число дітей в сім'ї; число викликів "швидкої допомоги", що надходять до лікарні; число

відмов виробу; число клієнтів, які звернулися у фірму за певний проміжок часу, і т.д.

Безперервною будемо вважати будь-яку змінну, яка не є дискретною. Вона приймає значення з деякого проміжку. Прикладами безперервної змінної є зріст дорослої людини (наприклад від 140 до 230см), фактична маса буханки хліба (наприклад від 750 до 830г), дальність польоту снаряда, врожайність культури, вирощеної в господарстві і т.п.

Є дані, які реєструють певну якість, яким володіє об'єкт. Такі дані називають *якісними*. Навіть якщо значенням цієї якості можна приписати числа (наприклад, статтю людини приписати відповідно числа 0 і 1), то обробляти ці числа як кількісні дані не можна. Прикладами якісних даних є тип школи, де навчається дитина (ліцей, гімназія, спеціалізована фізико-математична школа); посаду, яку займає співробітник на підприємстві; назви газет, які читають в певному місті, і т. п.

Якісні дані бувають двох типів: *порядкові*, для яких існує має змістовний сенс порядок, і *номінальні*, для яких немає змістовно інтерпретується порядку.

Порядкові дані можна ранжувати і використовувати це ранжування при проведенні статистичного аналізу. Прикладом порядкових даних є відповіді на питання анкети, що містить наступні варіанти відповідей: так; більше так, аніж ні; більше ні, ніж так; немає. Хоч і можна виразити ці відповіді числами (наприклад, 4, 3, 2, 1), але запропонована шкала оцінок носить суб'єктивний характер. Не можна вважати, що різниця між відповідями 4 і 3 така ж, як і між відповідями 2 і 1. Також не можна вважати, що відповідь 3 в три рази краще відповіді 1.

Для номінальних даних немає числових значень і немає основи для ранжирування. Прикладами номінальних даних є регіони України, з яких приїхали студенти ХНУ; назви фірм, що виготовляють м'які засоби; стать працівників фірми.

Якщо порядок запису значень даних у часі має змістовний сенс, то кажуть, що ці дані являють собою *часовий ряд*. Ці дані представляють інформацію про об'єкт в різні моменти часу. Якщо порядок запису даних в часі не суттєвий, то говорять про один часовий зріз. Ці дані представляють інформацію про об'єкти в певний момент часу. Прикладами тимчасового ряду є дані про злочинність в регіоні за кілька років; про рівень безробіття за кілька років; динаміка успішності учнів з математики і т. п. Прикладами одного тимчасового зрізу є дані про кількість злочинів в районах області в певний рік; число безробітних серед різних вікових груп населення за даними на певний день; дані про виконання тексту учнями класу і т. п.

Іноді дані класифікують також по тому, чи збиралися вони спеціально для запланованого аналізу або збиралися раніше для інших потреб. У першому випадку їх називають *первинними даними*, а в другому - *вторинними даними*. Отримання первинних даних пов'язане найчастіше з великими витратами коштів та часу, але зате дослідник отримує те, що йому потрібно. Вторинні дані, як правило, обходяться дешевше, їх швидше можна отримати, але при цьому не завжди в них можна знайти те, що потрібно.

Багато статистичні дані отримують в процесі вимірів. Метою вимірювань є отримання інформації про ознаки об'єктів, організмів, подій. Вимірюється не сам об'єкт, а тільки властивості або відмінні ознаки об'єкта. Наприклад, вимірюється не учень, а його зростання, маса, його швидкість читання, досягнення з математики, спортивні досягнення і т. і. Вимірювання здійснюються шляхом встановлення відповідності між числами та об'єктами, які є носіями підлягають вимірюванню властивостей. Вимірювання можуть проводитися на різних рівнях. Різних рівнів вимірювань відповідають різні шкали:

- 1) Номінальна шкала;
- 2) Порядкова(рангова);
- 3) Шкала інтервалів;
- 4) Шкала відношень(шкала пропорцій).

Номинальна шкала використовується для реєстрації самого нижчого рівня вимірювань, який передбачає наявність мінімальних передумов для вимірювання. При вимірах на даному рівні практично не використовуються числа. Тут важливо встановити подібність або відмінність об'єктів за певною ознакою, тобто, при цьому мають справу з якісними даними. Розглянемо приклади:

1. Розподіл учнів по класах, за статевою ознакою, за місцем проживання, за видами спорту, якими вони займаються, по числу дітей в сім'ї є прикладами величин номінальної шкали. При цьому можливий розподіл учнів по двом або більше ознаками (двовимірні або багатовимірні дані).
2. Перелік фірм, що займаються виробництвом вантажних і легкових автомобілів, автомобілів спеціального призначення, автобусів; відмінні ознаки автомобілів є прикладами величин номінальної шкали.

За допомогою підрахунку можна встановити частоту тієї чи іншої категорії (число хлопчиків і дівчат в школі; число учнів, які проживають в кожному мікрорайоні; число учнів у кожному класі; число учнів, що займаються тим чи іншим видом спорту; кількість фірм, що займаються виробництвом автобусів і т. д.). При цьому можна визначити найбільш часто зустрічається величину (клас, в якому навчається найбільша кількість учнів; вид спорту, який користується найбільшою популярністю в учнів; тип автомобіля, виробництвом якого займається найбільше число фірм). Категорії даних номінальної шкали позначаються, як правило, словесно (вербально).

Порядкова(рангова) шкала вказує лише послідовність носіїв ознаки або напрямок ступеня вираженості ознаки.

Наприклад, учнів можна ранжувати за кількістю правильно виконаних тестових завдань. Нехай учні А, Б, В, Г, Д правильно виконали відповідно 21, 16, 12, 9 і 3 завдання. Графічно це можна зобразити так:

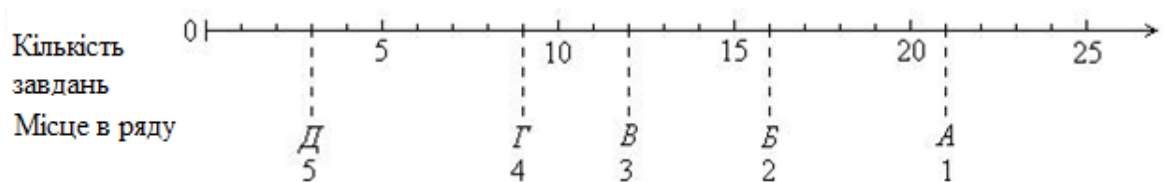


Рис. 1. Кількість вірно виконаних завдань учнями

Ця порядкова шкала має величини від 1 до 5, і учні на ній розміщені в залежності від кількості правильно виконаних завдань: А - перший, Д - п'ятий. З малюнка видно, що інтервали, що розділяють місця в ряду, різні за величиною. З цієї причини недоцільно складати, віднімати, множити і ділити порядкові місця.

Шкала шкільних оцінок з одного предмета є порядковою шкалою, так як інтервали між окремими балами не відображають розриву між реальними результатами. Ми знаємо тільки, що учень, який отримав оцінку "5" за якогось предмету, знає цей предмет краще того, хто отримав "4". Але не можна стверджувати, що різниця в знаннях цих учнів таке ж, як і в знаннях тих, хто отримав "4" і "3". Так як шкала оцінок є порядковою шкалою, то некоректно виставляти підсумкову оцінку як середню арифметичну поточних оцінок.

Іншим прикладом порядкової шкали є шкала посад, які займають працівники на підприємстві: директор, заступник директора, начальник відділу, завідувач лабораторією, старший інженер, інженер, молодший інженер, технік, старший лаборант, лаборант.

На шкалі *інтервалів* рівні інтервали відображають однакову міру величини вимірюваного ознаки. Наприклад, 1см між 3-м і 4-м сантиметрами на шкалі вимірювань довжин має такий же зміст, як і 1см між 82-м і 83-м сантиметрами. Іншими словами, на шкалі інтервалів відстані між сусідніми поділами рівні. На інтервальною шкалою цілком осмисленим є питання "на скільки?". Але не завжди, користуючись інтервальною шкалою, можна формулювати питання "у скільки разів?". Справа в тому, що на шкалі інтервалів встановлюються довільно початок відліку (нуль шкали), одиниця виміру і напрямок відліку. Прикладом інтервальною шкали є температурна

шкала за Цельсієм. Різниця між температурами повітря $+30$ і $+20$ ° С настільки ж велика, як і між -10 і -20 ° С. Однак не можна стверджувати, що при температурі повітря $+30$ ° С в півтора рази тепліше, ніж при температурі $+20$ ° С. Навіть якщо температура повітря дорівнює 0 ° С, не можна стверджувати, що тепла немає зовсім: адже початок відліку вибрано довільно.

Шкали на більшості фізичних приладів (амперметр, вольтметр і ін.) Є інтервальними. Шкала коефіцієнта інтелекту IQ є шкалою інтервалів.

Шкала інтервалів є метричною, з її допомогою можна виконувати додавання і віднімання. Вона має значні переваги в порівнянні з номінальною і порядковою шкалами.

Шкала відношень, або шкала пропорцій, дає можливість встановлювати відносини значень вимірюваного ознаки завдяки тому, що значенням шкали "0" відповідає величина, для якої вимірюється ознака відсутня. Іншими словами, початок відліку на цих шкалах вибирають мимоволі. Прикладами шкали відносин є міри довжини (м, см і т. Д.) І маси (кг, г і т. Д.). Предмет довжиною 100 см удвічі довше предмета довжиною 50см.

Іноді дані потребують перетворення. Зокрема, потреба в цьому виникає, коли в ряду даних одне або кілька даних істотно перевищують інші. Якщо дані явно несиметричні, то замінюють кожне значення наведеного набору даних логарифмом цього значення з метою спростити статистичний аналіз.

Основні переваги статистичних обстежень бізнес активності підприємств

Функціонування будь-якого суб'єкта господарювання на ринку не може відбуватися успішно без постійного використання інформації про власний стан, стан навколишнього середовища, партнерів, конкурентів, що й сприяло формуванню нового напрямку у практиці національної статистики — статистичного моніторингу. Моніторинг (від лат. monitor — що нагадує, що наглядає) — у широкому розумінні є функцією управління поряд з функціями планування, стратегічного та операційного менеджменту.

Удосконалення управління зумовлює використання менеджерами результатів статистичного моніторингу ділової активності підприємств для розробки подальших можливих корпоративних стратегій розвитку: реструктуризації, диверсифікації, вертикальної інтеграції та портфельного аналізу.

Ці корпоративні стратегії розвитку досить вичерпно охарактеризовані у відповідних літературних джерелах . Зазначимо тут лише, що портфельним називають аналіз, для проведення якого потрібне статистичне забезпечення у вигляді статистичного моніторингу ділової активності підприємств. Статистичний моніторинг ділової активності підприємств має цілком визначену мету, алгоритм функціонування і використовує систему якісних і кількісних статистичних показників.

Традиційно портфельний аналіз використовується з метою оптимізації портфеля цінних паперів інвестора. Фінансист під терміном «портфоліо» розуміє оптимальний з погляду комбінації ризику та прибутковості набір інвестицій. Отже, в основу портфельного аналізу покладено два оцінні критерії: теперішня вартість очікуваних доходів від володіння цінними паперами (проценти, дивіденди) та рівень ризикованості вкладень.

За аналогією з цінними паперами можна проводити аналіз портфеля продуктів (товарів і послуг) підприємства . Це означає, що портфельний аналіз доцільний на тих підприємствах, які виробляють багато видів продуктів. У процесі аналізу окремі групи продуктів розглядаються як відповідні стратегічні «бізнес-одиниці», причому кожна з них оцінюється щодо прибутковості та ризикованості виробництва. На підставі результатів портфельного аналізу приймаються рішення про додаткові інвестиції в окремі виробничі програми, реінвестиції чи дезінвестиції, а також вивчається стратегія управління ризиками за кожною зі стратегічних бізнес-одиниць.

Під час аналізу використовуються чотири широковідомі в теорії маркетингу та на практиці портфельні матриці:

- матриця «зростання ринку — частка ринку» (матриця Бостонської консультативної групи);

- матриця «привабливість ринку — конкурентні переваги» (матриця Мак-Кінсі);
- матриця General Electric «конкурентоспроможність стратегічних бізнес-одиниць» «привабливість галузі»;
- матриця напрямленої політики «перспективи розвитку ділового сектору (погані, середні або добрі) — конкурентоспроможність суб'єктів господарювання (низька, середня або висока).

Саме остання, четверта матриця — матриця напрямленої політики — має за систему координат перспективи розвитку ділового сектору. Тому результати статистичного моніторингу ділової активності підприємств мають бути спрямовані саме на агреговану оцінку перспектив розвитку ділового сектору.

Згідно з планом статистичних робіт Державна служба статистики України проводить регулярні щоквартальні обстеження ділової активності підприємств України, за результатами яких видається Бюлетень обстежень ділової активності підприємств України. Обстеження ділової активності охоплює промислові, будівельні, торговельні, сільськогосподарські та транспортні підприємства.

В Україні щомісячно опитується 1313 промислових підприємств, 244 будівельних, 726 торговельних, 590 сільськогосподарських і 251 транспортне підприємство. Щодо причин, які спонукають респондентів до участі в опитуваннях, то 37 % від загальної їхньої кількості вказало на те, що обстеження дають привід обмірковувати стан підприємства, майже кожний п'ятий із опитуваних отримує корисну інформацію в обмін на участь в обстеженнях .

Які ж основні переваги обстежень ділової активності? По-перше, — це швидкість отримання необхідної інформації. Обстеження, за наявності комп'ютерних технологій обробки, роблять доступними ключові дані протягом 15—20 днів, тоді як офіційні цифри отримуються значно пізніше.

По-друге, інформація з обстежень доповнює дані офіційної статистики, що дає змогу краще і швидше показувати поширення економічних тенденцій.

По-третє, інформація, яку отримують за допомогою обстежень, і суб'єктивною, ці результати вимірюють розуміння ділової ситуації респондентами — менеджерами підприємств. Наприклад, за результатами обстеження легко отримати інформацію про очікування підприємців і виявити на її підставі не тільки їхнє сприйняття економічної ситуації, а й спрогнозувати їхню економічну поведінку. І остання, найбільш очевидна перевага обстежень: вони дозволяють отримувати інформацію зі сфер економічної діяльності, що не охоплені кількісною офіційною статистикою, наприклад про наявність і використання виробничих потужностей, оцінку запасів готової продукції та ін.

Обстеження проводяться за методикою Комісії Європейського союзу, рекомендованою для застосування у країнах з перехідною економікою.

Випадкові величини і закони розподілу як основа бізнес-статистики

Змінна величина називається *випадковою*, якщо в результаті досвіду вона може приймати дійсні значення з певними ймовірностями. *Закон розподілу* - функція (таблиця, графік, формула), що дозволяє визначати ймовірність того, що випадкова величина X приймає певне значення x_i або потрапляє в певний інтервал.

Випадкова величина X називається дискретною, якщо існує така невід'ємна функція

$$P\{X = x_i\} = p_i, \quad i = 1, 2, \dots, \sum_{i=1}^{\infty} p_i = 1, \quad (3.1)$$

яка ставить у відповідність значенням x_i змінної X ймовірність p_i , з якої вона приймає це значення.

Випадкова величина X називається *безперервною*, якщо для будь-яких $a < b$ існує така невід'ємна функція $f(x)$

$$P\{a \leq X \leq b\} = \int_a^b f(x)dx. \quad (3.2)$$

Функція $f(x)$ називається *щільністю розподілу* неперервної випадкової величини. Імовірність того, що випадкова величина X (дискретна або безперервна) приймає значення, менше x , називається *функцією розподілу* випадкової величини X і позначається $F(x)$:

$$F(x) = P\{X < x\}. \quad (3.3)$$

Кожен закон розподілу - це деяка функція, що повністю описує випадкову величину з імовірнісної точки зору. На практиці про розподіл ймовірностей випадкової величини X часто доводиться судити лише за результатами випробувань. Повторюючи випробування, будемо кожен раз реєструвати, чи відбулося випадкова подія A , чи ні. Відносною частотою (або просто частотою) випадкової події A називається відношення числа n_A появ цієї події до загального числа n проведених випробувань.

Проводячи серії випробувань, які дають емпіричні значення величини X , треба фіксувати числа n_x влучень результатів в кожен інтервал. При великому числі випробувань n ставлення n_x / n (частоти потрапляння в інтервали) повинні бути близькі до можливостям попадання в ці інтервали. Залежність частот n_x / n інтервалів визначає *емпіричний розподіл* ймовірностей випадкової величини X , графічне представлення якої називається гістограмою (рис. 2).

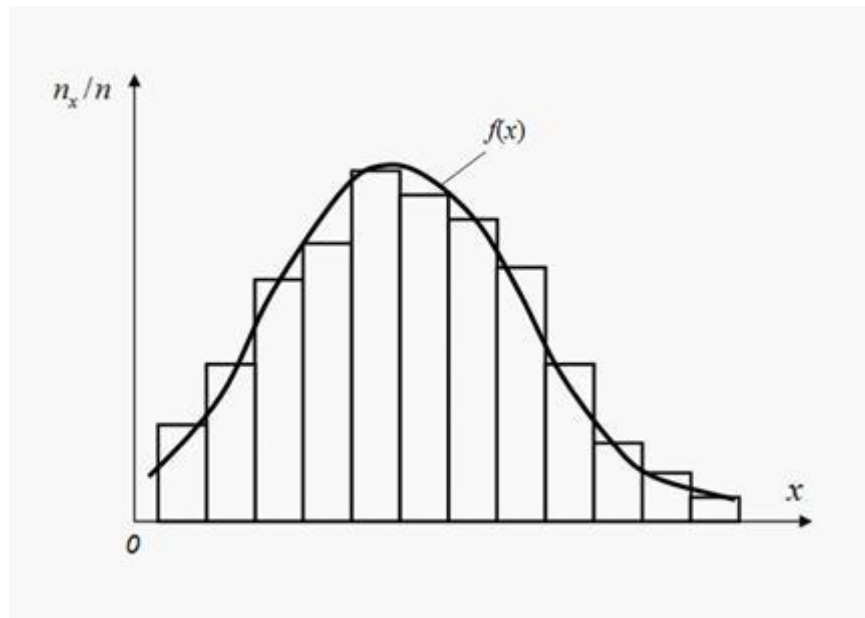


Рис.2. Гістограма і щільність розподілу

Для побудови гістограми по осі абсцис відкладають інтервали рівної довжини, на які розбивається весь діапазон можливих значень випадкової величини X , а по осі ординат відкладають частоти n_x / n . Тоді висота кожного стовпчика гістограми дорівнює відповідній частоті. Таким чином, виходить наближене уявлення закону розподілу ймовірностей для випадкової величини X у вигляді ступінчастої функції, апроксимація (вирівнювання) якої деякої кривої $f(x)$ дасть щільність розподілу.

Розподіл дискретних випадкових величин

Біноміальний розподіл. Дискретна випадкова величина X має біноміальний розподіл, якщо її можливі значення $0, 1, 2, \dots, m, \dots, n$, а відповідні їм ймовірності рівні:

$$P_m = P\{X = m\} = C_n^m p^m q^{n-m}, \quad (3.4)$$

Як видно з (4), ймовірності P_m обчислюються, як члени розкладання бінома Ньютона $(p+q)^n$, звідки і назва «біноміальний розподіл».

Прикладом є вибіркового контроль якості виробничих виробів, при якому відбір виробів для проби проводиться за схемою *випадкової повторної* вибірки, тобто коли перевірені вироби повертаються у вихідну партію. Тоді

кількість нестандартних виробів серед відібраних є випадкова величина з біноміальним законом розподілу ймовірностей.

Біноміальний розподіл визначається двома параметрами: n і p . Випадкова величина, розподілена за біноміальним законом, має наступні основні числові характеристики:

$$m = np, D = npq, \sigma = \sqrt{npq}. \quad (3.5)$$

Геометричний розподіл. Дискретна випадкова величина X має геометричний розподіл, якщо її можливі значення $0, 1, 2, \dots, m, \dots$, а ймовірності цих значень:

$$m = np, D = npq, \sigma = \sqrt{npq}. \quad (3.6)$$

де $0 < p < 1, q = 1 - p; m = 0, 1, 2, \dots$

Розподіл неперервних випадкових величин

Рівномірний розподіл. Величина X розподілена рівномірно на інтервалі (a, b) , якщо всі її можливі значення знаходяться на цьому інтервалі і щільність розподілу ймовірностей постійна:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{при } x \in (a, b), \\ 0 & \text{при } x \notin (a, b). \end{cases} \quad (3.7)$$

Для випадкової величини X , рівномірно розподіленою в інтервалі (a, b) (рис. 3), ймовірність попадання в будь-який інтервал (x_1, x_2) , що лежить всередині інтервалу (a, b) , дорівнює:

$$P\{x_1 < X < x_2\} = \frac{x_2 - x_1}{b - a}. \quad (3.8)$$

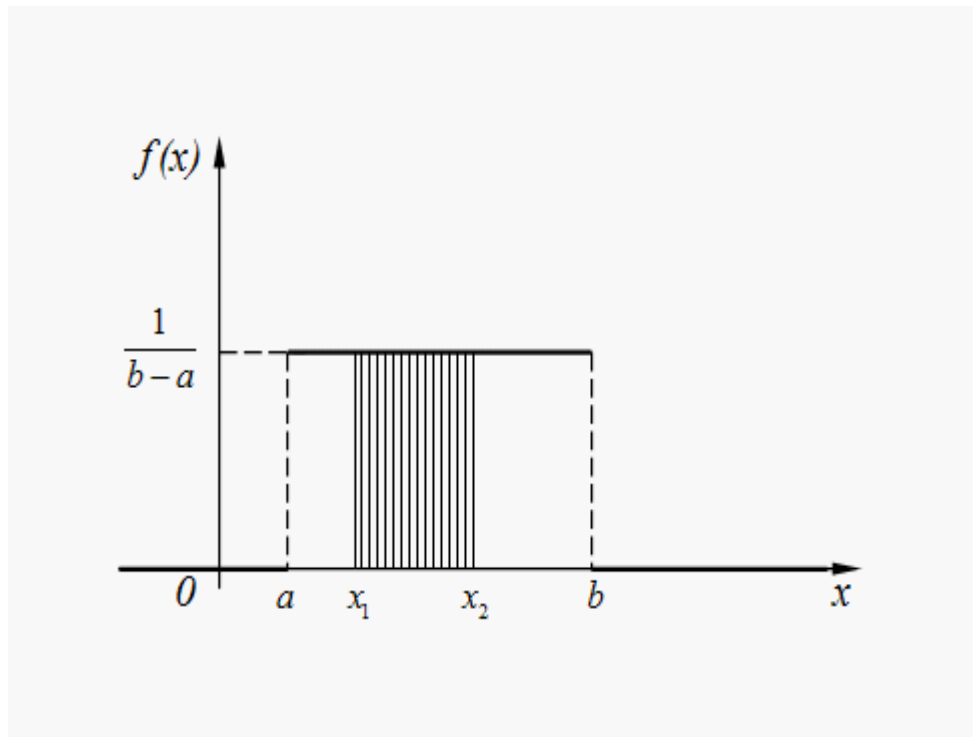


Рис. 3. Графік щільності рівномірного розподілу

Прикладами рівномірно розподілених величин є помилки округлення. Так, якщо всі табличні значення деякої функції округлені до одного і того ж розряду 10^{-m} , то вибираючи навмання табличне значення, ми вважаємо, що помилка округлення обраного числа є випадкова величина.

Показовий розподіл. Безперервна випадкова величина X має показовий розподіл, якщо щільність розподілу її ймовірностей виражається формулою:

$$f(t) = \begin{cases} \lambda * e^{-\lambda t} & \text{при } t > 0, \\ 0 & \text{при } t \leq 0. \end{cases} \quad (3.9)$$

Графік щільності розподілу ймовірностей (3.9) представлений на рис. 4.

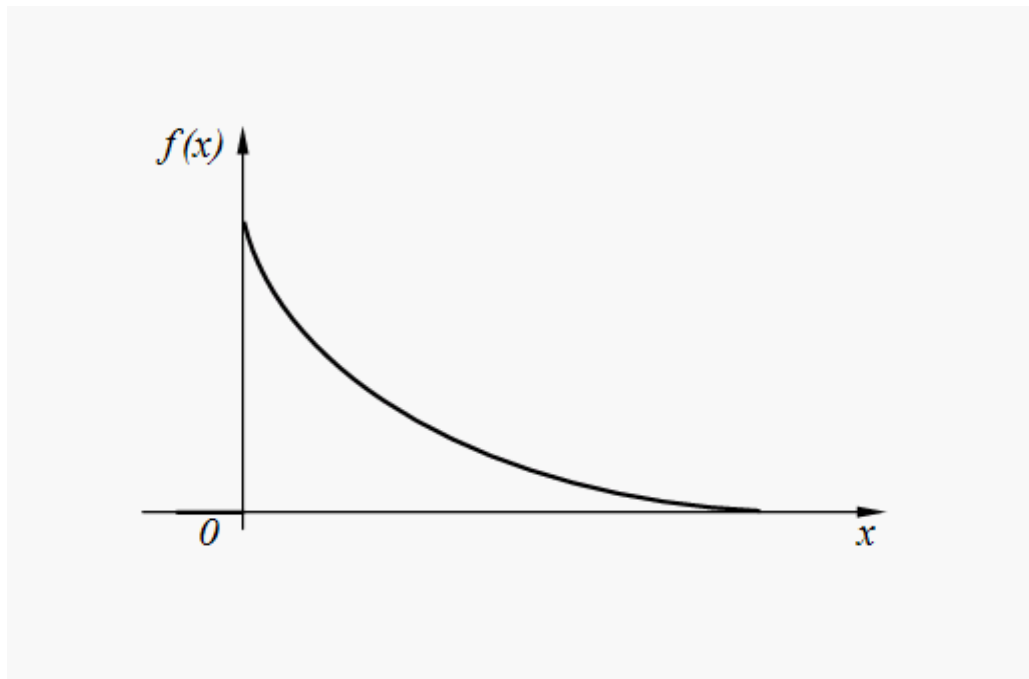


Рис. 4. Графік щільності показового розподілу

Час T безвідмовної роботи комп'ютерної системи є випадкова величина, що має показовий розподіл з параметром λ , фізичний зміст якого - середнє число відмов в одиницю часу, беручи до уваги простоїв системи для ремонту.

Нормальне (гауссовий) розподіл. Випадкова величина X має нормальний (гауссовий) розподіл, якщо щільність розподілу її ймовірностей визначається залежністю:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad (3.10)$$

де $m = M(X)$, $\sigma = \sqrt{D(X)}$.

При $m=0$, $\sigma=1$ нормальний розподіл називається *стандартним*.

Графік щільності нормального розподілу (3.10) представлений на рис.

5.

Нормальний розподіл є найбільш часто зустрічається в різних випадкових явищах природи. Так, помилки виконання команд автоматизованим пристроєм, помилки виведення космічного корабля в задану точку простору, помилки параметрів комп'ютерних систем і т.п. в більшості випадків мають нормальне або близьке до нормального розподіл.

Більш того, випадкові величини, утворені підсумовуванням великої кількості випадкових доданків, розподілені практично по нормальному закону.

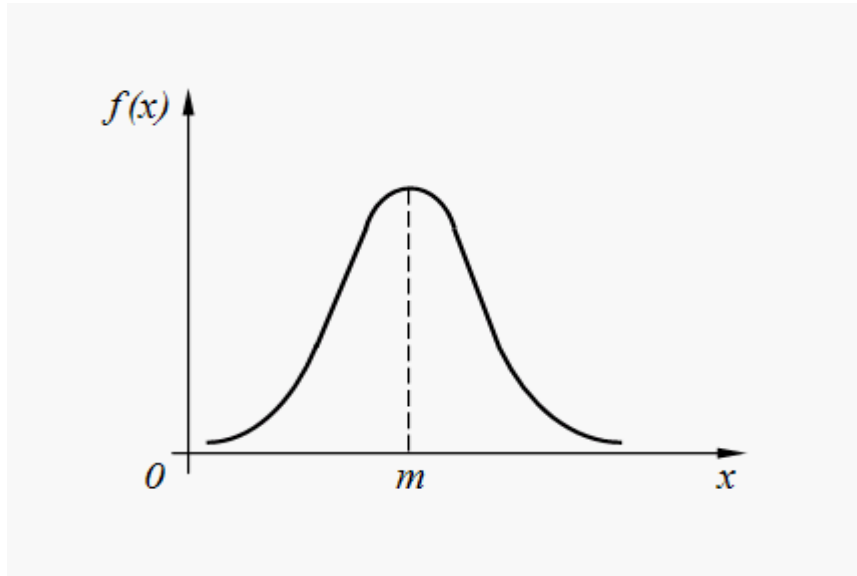


Рис. 5. Графік щільності нормального розподілу

Системи випадкових величин

Істотний інтерес в математичній статистиці представляє розгляд системи двох і більше випадкових величин і їх статистична взаємозв'язок один з одним. За аналогією з рядом розподілу однієї дискретної величини X для двох дискретних випадкових величин X і Y будується матриця розподілу - прямокутна таблиця, в якій записані всі ймовірності $p_{ij} = P\{X = x_i, Y = y_j\}$, $i = 1, \dots, n$; $j = 1, \dots, m$. Вони кажуть події (або досліди) називаються незалежними, якщо ймовірність появи (результату) кожного з них не залежить від того, які події (результати) мали місце в інших випадках (дослідах). Дві випадкові величини X і Y називаються незалежними, якщо незалежні все пов'язані з ними події: наприклад, $\{X < a\}$ і $\{Y < b\}$ або $\{X = x_i\}$ і $\{Y = y_j\}$ і т.д. Спільною функцією розподілу системи двох випадкових величин (X, Y) називається ймовірність спільного виконання нерівностей $X < x$ і $Y < y$:

$$F(x, y) = P\{X < x, Y < y\}. \quad (3.11)$$

Подія означає твір (спільне виконання) подій $\{X < x\}$ і $\{Y < y\}$.

Геометричній інтерпретацією спільної функції розподілу $F(x, y)$ є ймовірність попадання випадкової точки (X, Y) на площині всередину квадранта з вершиною в точці (x, y) (заштрихована область на рис. 6).

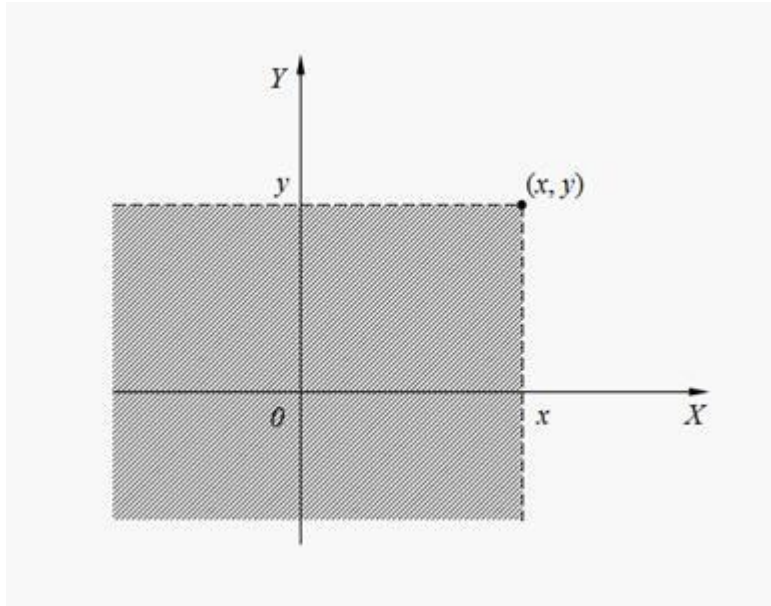


Рис. 6. Геометрична інтерпретація спільної функції розподілу $F(x, y)$

В якості числових характеристик системи двох випадкових величин X і Y зазвичай розглядаються початкові і центральні моменти різних порядків. Порядком моменту називається сума його індексів $k + s$.

Початковим моментом порядку $k + s$ системи двох випадкових величин X і Y називається математичне очікування добутку X^k на Y^s :

$$\alpha_{k,s} = M(X^k Y^s). \quad (3.12)$$

Центральним моментом порядку $k + s$ системи двох випадкових величин (X, Y) називається математичне сподівання добутку $(X - m_x)^k$ на $(Y - m_y)^s$:

$$\mu_{k,s} = M[(X - m_x)^k (Y - m_y)^s], \quad (3.13)$$

де $m_x = M(X)$, $m_y = M(Y)$.

В інженерному застосуванні математичної статистики найчастіше використовуються моменти першого і другого порядків.

Початкові моменти першого порядку

$$\alpha_{1,0} = M(X^1 Y^0) = M(X) = m_x; \quad \alpha_{0,1} = M(X^0 Y^1) = M(Y) = m_y \quad (3.14)$$

є математичними очікуваннями випадкових величин X і Y .

Центральні моменти першого порядку завжди дорівнюють нулю:

$$\begin{aligned}\mu_{1,0} &= M[(X - m_x)^1(Y - m_y)^0] = M(X - m_x) = M(X) - M(m_x) = m_x - m_x \\ &= 0, \\ \mu_{0,1} &= M[(X - m_x)^0(Y - m_y)^1] = M(Y - m_y) = M(Y) - M(m_y) = m_y - m_y \\ &= 0 \quad (3.15)\end{aligned}$$

Початкові моменти другого порядку:

$$\begin{aligned}\alpha_{2,0} &= M(X^2Y^0) = M(X^2) = \alpha_2(X), \\ \alpha_{0,2} &= M(X^0Y^2) = M(Y^2) = \alpha_2(Y), \\ \alpha_{1,1} &= M(X^1Y^1) = M(XY). \quad (3.16)\end{aligned}$$

Центральні моменти другого порядку:

$$\begin{aligned}\mu_{2,0} &= M[(X - m_x)^2(Y - m_y)^0] = M[(X - m_x)^2] = D_x, \\ \mu_{0,2} &= M[(X - m_x)^0(Y - m_y)^2] = M[(Y - m_y)^2] = D_y, \\ \mu_{1,1} &= M[(X - m_x)^1(Y - m_y)^1] = M[(X - m_x)(Y - m_y)]. \quad (3.17)\end{aligned}$$

Тут D_x , D_y - дисперсії випадкових величин X і Y .

Центральний момент другого порядку називається коваріацією випадкових величин X і Y . Позначимо його:

$$K_{x,y} = M[(X - m_x)(Y - m_y)]. \quad (3.18)$$

З визначення коваріації (18) слідує:

$$K_{xy} = K_{yx} \quad (3.19)$$

Дисперсія випадкової величини є по суті окремим випадком коваріації:

$$\begin{aligned}D_x &= M[(X - m_x)(X - m_x)] = K_{xx}, \\ D_y &= M[(Y - m_y)(Y - m_y)] = K_{yy} \quad (3.20)\end{aligned}$$

Коваріація двох випадкових величин X і Y характеризує ступінь їх залежності і міру розсіювання навколо точки. Часто буває зручно виразити коваріацію у вигляді:

$$K_{xy} = M(X * Y) - M(X) * M(Y) \quad (3.21)$$

Вираз (52) впливає з визначення коваріації (22).

Розмірність коваріації дорівнює добутку розмірностей випадкових величин X і Y .

Безрозмірна величина, що характеризує тільки залежність випадкових величин X і Y :

$$r_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y} \quad (3.23)$$

називається *коефіцієнтом кореляції* випадкових величин X і Y . Цей параметр характеризує ступінь лінійної залежності випадкових величин X і Y . Для будь-яких двох випадкових величин X і Y коефіцієнт кореляції $-1 \leq r \leq 1$. Якщо $r > 0$, то лінійна залежність між X і Y зростаюча, якщо, $r < 0$ то лінійна залежність між X і Y зворотня, при $r = 0$ лінійної залежності між X і Y немає. Відсутність лінійної кореляції не означає відсутність будь-якої іншої залежності між X і Y . Якщо має місце жорстка лінійна залежність $Y = aX + b$, то $r = +1$ при $a > 0$ і $r = -1$ при $a < 0$.

Контрольні питання

1. За якими критеріями класифікують дані? Наведіть приклади.
2. Назовіть шкали вимірювань на різних рівнях.
3. Основні переваги статистичних обстежень бізнес активності підприємств.
4. Поняття про «портфоліо» та матриці портфельного аналізу.
5. Що таке випадкова величина і чим вона відрізняється від числа?
6. Вкажіть відмінності між дискретною та неперервною випадковою величиною.
7. Системи випадкових величин та їх графічна інтерпретація.
8. Поняття про «коваріацію» та її характеристика.

1.2. Статистичні гіпотези та дисперсійний аналіз у дослідженні бізнес-процесів

Поняття статистичної гіпотези та методика її перевірки. Критерії перевірки статистичних гіпотез та типи та їх властивості. Статистичний висновок як основа узагальнення даних. Дисперсійний аналіз і дослідження бізнес-процесів

Поняття статистичної гіпотези та методика її перевірки

Статистична гіпотеза - певне припущення про властивості розподілу ймовірностей, що лежить в основі спостережуваних випадкових явищ. Результати спостережень представляються зазвичай у вигляді реалізації деякої сукупності випадкових величин, кінцевою або нескінченною. При цьому спільний розподіл цих випадкових величин відомо в повному обсязі, і передбачає належність його до деякого певного класу розподілів. У такій ситуації ставиться завдання статистичних гіпотез перевірки.

Перевірка статистичної гіпотези - це процес прийняття рішення про те, чи суперечить розглянута статистична гіпотеза спостерігається вибірці даних.

Статистичний тест або статистичний критерій - строгий математичний правило, за яким приймається або відкидається статистична гіпотеза.

Статистичні гіпотези перевіряються на підставі вибірових дослідів або спостережень. Тому ймовірність тієї чи іншої гіпотези ніколи не може бути прийнята однозначно рівною 0 або 1. Замість цього використовуються допустимі рівні ймовірності: α і $1-\alpha$ (або β). У практиці статистичного аналізу найчастіше мають справу з двома конкуруючими гіпотезами: нульовою гіпотезою, що позначається H_0 , і альтернативної гіпотезою, що позначається H_a або H_1 .

Нульова гіпотеза - це твердження, що підлягає перевірці. Альтернативна гіпотеза - це протилежне твердження, яке ми намагаємося спростувати в ході аналізу. Залежно від величини ймовірності альтернативна гіпотеза може бути прийнята або відкинута. Нульова гіпотеза може бути

відкинута, але не прийнята! У разі високого рівня ймовірності правдивості нульової гіпотези, кажуть, що гіпотеза не може бути відкинута.

Нехай нульова гіпотеза полягає в тому, що випадкова величина належить розподілу 1 на рис. 1.1 Це записується в такий спосіб: $H_0: X \sim 1$ (1). Конкуруючої гіпотезою в цьому випадку може бути $H_1: X \sim 2$ (випадкова величина X належить розподілу 2 на рис. 1.1).

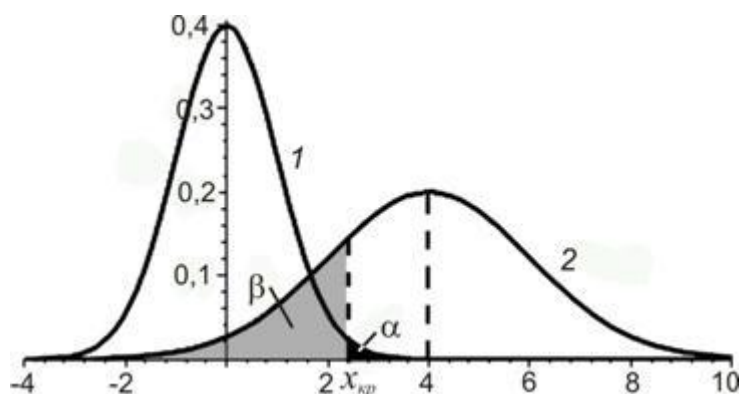


Рис.1.1. Вид гіпотез: 1-нульова; 2- конкуруюча

Гіпотези бувають прості і складні. Простий називають гіпотезу, яка містить тільки одне припущення. У наведеному вище прикладі обидві гіпотези прості. Складною називають гіпотезу, яка складається з кінцевого або нескінченного числа нульових гіпотез (припущень).

Вибір гіпотези. Щоб з'ясувати, чи відповідає розподіл в генеральній сукупності заданої теоретичного розподілу, необхідно сформулювати досліджувану гіпотезу. В цьому випадку можуть випробовувати дві гіпотези: 1) $H: \chi^2 = 0$; 2) $H: \chi^2 > 0$.

Для вибору досліджуваної гіпотези необхідно враховувати сліdstва можливих помилок. Обчислюючи теоретичне розподіл, дослідник намагається встановити закономірність даного явища, довести, що досліджуване розподіл підпорядковується запропонованого теоретичного закону. Тому важливо не зробити помилкового відкриття: чи не приписувати розподілу даний теоретичний закон, якщо насправді воно відповідає іншим законом. Такий висновок відповідає помилкового відхилення гіпотези $\chi^2 > 0$, якщо вона вірна, тобто коли теоретичне розподіл відрізняється від розподілу

генеральної сукупності. Альтернативною повинна бути гіпотеза $\chi^2 = 0$. Гіпотеза $\chi^2 > 0$ є складовою, і для її перевірки необхідно було б виділити просту гіпотезу, яка відповідає практично незначну різницю між розподілами. Однак існує багато теоретичних розподілів, близьких до вибраного. Вказати мінімальне значення χ^2 , відповідне різниці між обраним теоретичним і іншим, найближчим до нього, неможливо. Тому в цьому випадку доводиться відступати від загальних правил вибору гіпотез. Використовувана гіпотеза буде нульовою через відсутність різниці між теоретичним і генеральним розподілом, тобто: $H_0: \chi^2 = 0$. У зв'язку з цим слід вибрати досить великий рівень значущості, щоб зменшити ймовірність, що має більш серйозні наслідки помилок II-го роду, тобто помилкового визнання відповідності розподілу ознаки в генеральній сукупності запропонованого теоретичного розподілу. З огляду на те, що альтернативна гіпотеза $\chi^2 > 0$, проводиться правобічна перевірка.

Якщо вірна альтернативна гіпотеза, то статистична характеристика підпорядковується *нецентрального розподілу* χ^2 . Користуючись цим розподілом, можна обчислити потужність критерію χ^2 . Однак з огляду на особливості вибору гіпотези, такі розрахунки потужності в цьому випадку мають невелике практичне значення і не будуть розглядатися.

Методика перевірки статистичних гіпотез. Нехай задана випадкова вибірка $x^m = (x_1, \dots, x_m)$ - послідовність m об'єктів з множини X . Передбачається, що в множині X існує деяка невідома імовірнісна міра P .

Методика полягає в наступному.

1. Формулюється нульова гіпотеза H_0 про розподіл ймовірностей на множині. Гіпотеза формулюється виходячи з вимог прикладної задачі. Найчастіше розглядаються дві гіпотези - основна або нульова (H_0) і альтернативна (H_1). Іноді альтернатива формулюється в явному вигляді; тоді передбачається, що H_1 означає «не H_0 ». Іноді розглядається відразу кілька альтернатив. У математичній статистиці добре вивчено кілька десятків

«найбільш часто зустрічаються» типів гіпотез, і відомі ще сотні спеціальних варіантів і різновидів. Приклади наводяться нижче.

2. Задається деяка статистика (функція вибірки) $T: X^m \rightarrow R$, для якої в умовах справедливості гіпотези H_0 виводиться функція розподілу $F(T)$ і / або щільність розподілу $p(T)$. Питання про те, яку статистику треба взяти для перевірки тієї чи іншої гіпотези, часто не має однозначної відповіді. Є цілий ряд вимог, яким повинна задовольняти «хороша» статистика T . Висновок функції розподілу $F(T)$ при заданих H_0 і T є суворим математичним завданням, яке вирішується методами теорії ймовірностей; в довідниках наводяться готові формули для $F(T)$; в статистичних пакетах є готові обчислювальні процедури.

3. Фіксується *рівень значимості* - допустима для даного завдання ймовірність *помилки першого роду*, тобто того, що гіпотеза насправді вірна, але буде відкинута процедурою перевірки. Це повинно бути досить мале число $\alpha \in [0,1]$. На практиці часто вважають $\alpha=0,05$.

4. На безлічі допустимих значень статистики T виділяється критична множина Ω_α найменш імовірних значень статистики T , таке, що $P\{T \in \Omega_\alpha | H_0\} = \alpha$ Обчислення меж критичного безлічі як функції від рівня значущості α є суворим математичним завданням, яка в більшості практичних випадків має готове просте рішення.

5. Власне *статистичний тест (статистичний критерій)* полягає в перевірці умови:

□ якщо $T(X^m) \in \Omega_\alpha$, то робиться висновок «дані суперечать нульовій гіпотезі при рівні значущості α ». Гіпотеза відкидається.

□ якщо $T(X^m) \notin \Omega_\alpha$, то робиться висновок «дані не суперечать нульовій гіпотезі при рівні значущості α ». Гіпотеза приймається.

Отже, *статистичний критерій* визначається статистикою T і критичною множиною Ω_α , яке залежить від рівня значущості.

Зауваження. Якщо дані не суперечать нульовій гіпотезі, це ще не означає, що гіпотеза вірна. Для цього є дві причини.

□ У міру збільшення довжини вибірки нульова гіпотеза може спочатку прийматися, але потім виявляться більш тонкі невідповідності даних гіпотезі, і вона буде відкинута. Тобто багато що залежить від обсягу даних; якщо даних не вистачає, можна прийняти навіть саму неправдоподібну гіпотезу.

□ Обрана статистика може відображати не всю інформацію, що міститься в гіпотезі H_0 . В такому випадку збільшується ймовірність помилки другого роду - нульова гіпотеза може бути прийнята, хоча насправді вона не вірна. Припустимо, наприклад, що $H_0 =$ «розподіл нормально»; $T(X^m) =$ «Коефіцієнт асиметрії»; тоді вибірка з будь-яким симетричним розподілом буде визнана нормальною. Щоб уникати таких помилок, слід користуватися більш потужними критеріями.

Критерії перевірки статистичних гіпотез, типи та їх властивості

Ряди розподілу є найповнішою характеристикою статистичних сукупностей. Кожному ряду розподілу притаманна певна закономірність, яка виражається залежністю між варіантами і частотами. Закономірності розподілу можна зобразити графічно, їх можна також висловити аналітично у вигляді закону розподілу, який встановлює відношення варіант ймовірностей. Використовуючи закон розподілу, можна глибше проаналізувати досліджуване явище, спрогнозувати розподіл і т.п. Крім того, необхідною умовою застосування багатьох критеріїв є нормальність розподілу. У зв'язку з цим нерідко потрібно перевіряти відповідність розподілу цієї величини нормальному закону.

Надалі ряд розподілу буде називатися *емпіричним*, а відповідний йому закон - *теоретичним розподілом*. Існує багато законів розподілу: нормальний, логарифмічно-нормальний, Пуассона та ін. Вибір теоретичного закону розподілу повинен базуватися на розумінні механізму формування досліджуваного явища. Його можуть також полегшити графіки емпіричних розподілів.

Частки емпіричного розподілу відрізняються від ймовірностей теоретичного розподілу. Це може бути викликано, наприклад, помилковим

вибором теоретичного розподілу (в генеральній сукупності відповідає іншому закону) або випадковими похибками, які обумовлені відхиленнями вибіркової часткою від відповідних їй ймовірностей. Тому виникає завдання за допомогою *критеріїв узгодження* перевірити за даними вибіркового емпіричного розподілу гіпотезу про закон розподілу у генеральній сукупності.

У логарифмічно-нормальному розподілі розбіжності між теоретичними і емпіричними частотами значно менше, ніж при нормальному розподілі.

Для перевірки гіпотези про відповідність обраних законів розподілу і розподілу в генеральній сукупності в більшості критеріїв використовуються відхилення емпіричних частот від теоретичних. Очевидно, що чим більше ці відхилення, тим гірше теоретичне розподіл описує емпіричне. Статистичні характеристики більшості критеріїв узгодження є деякими функціями цих відхилень.

Критерій узгодження χ^2 , статистична характеристика. З усіх критеріїв узгодження критерій χ^2 застосовується частіше інших. Його статистичною характеристикою є сума часткою від ділення квадрата різниці між емпіричними і теоретичними частотами на теоретичні частки:

$$\chi^2 = \sum_{i=1}^m \frac{(f_i - f'_i)^2}{f'_i}, \quad (4.2)$$

де f – частки емпіричного розподілу; f' – частки теоретичного розподілу.

При розрахунку χ^2 частки можна замінити частотами:

$$\chi^2 = n \sum_{i=1}^m \frac{(\omega_i - P_i)^2}{P_i}. \quad (4.3)$$

Якщо розподіл в генеральній сукупності відповідає обраному теоретичному закону розподілу, то в генеральній сукупності $\chi^2 = 0$. Чим більше відхилення між цими розподілами, тим більше величина χ^2 . Отже, χ^2 є ступенем відхилення запропонованого теоретичного розподілу від розподілу в генеральній сукупності, а значення χ^2 , обчислене за даними вибірки, – оцінкою цього відхилення. За умови правильно обраного теоретичного

розподілу (в генеральній сукупності $\chi^2 = 0$), розподіл статистичної характеристики гіпотези при $n \rightarrow \infty$ зводиться до розподілу χ^2 з $m-l$ ступенями свободи. Критерій можна застосовувати тільки за умови, що всі $f' > 5$. В іншому випадку розподіл статистичної характеристики буде помітно відхилятися від розподілу.

Під час визначення кількості ступенів свободи m - кількість груп в емпіричному розподілі, а l - кількість параметрів, які розраховуються за допомогою емпіричного розподілу, і число незалежних лінійних зв'язків, яким підкоряються емпіричні частки.

Властивості статистичних критеріїв.

Потужність критерію $1 - \beta(H) = P\{T \in \Omega_\alpha | H\}$: - ймовірність відхилити гіпотезу H_0 , якщо насправді вірна альтернативна гіпотеза H_1 . Потужність критерію є числовою функцією від альтернативної гіпотези H .

Незміщений критерій $1 - \beta(H) \geq \alpha$: для всіх альтернатив H або, що те ж саме $P\{T \in \Omega_\alpha | H\} \geq P\{T \in \Omega_\alpha | H_0\}$, для всіх альтернатив

Рівномірно більш потужний критерій. Кажуть, що критерій з потужністю $1 - \beta(H)$ є рівномірно більш потужним, ніж критерій з потужністю $1 - \beta'(H)$, якщо виконуються дві умови:

1. $\beta(H_0) = \beta'(H_0)$;
2. $\beta(H_1) \leq \beta'(H_1)$ для всіх розглянутих альтернатив $H_1 \neq H_0$, причому хоча б для однієї альтернативи нерівність суворе.

Типи статистичних критеріїв. Залежно від конкретних *нульової гіпотези* статистичні критерії діляться на групи, перераховані нижче по розділах.

Поряд з нульовою гіпотезою, яка приймається або відкидається по результату аналізу вибірки, статистичні критерії можуть спиратися на додаткові припущення, які апріорі припускаються виконаними.

□ *Параметричні критерії* передбачають, що вибірка породжена розподілом із заданого параметричного сімейства. Зокрема, існує багато критеріїв, призначених для аналізу вибірок з нормального розподілу.

Перевага цих критеріїв в тому, що вони більш потужні. Якщо вибірка дійсно задовольняє додатковим припущеннями, то параметричні критерії дають більш точні результати. Однак якщо вибірка їм не задовольняє, то ймовірність помилок (як I, так і II роду) може різко зрости. Перш ніж застосовувати такі критерії, необхідно переконатися, що вибірка задовольняє додатковим припущеннями. Гіпотези про вид розподілу перевіряються за допомогою критеріїв згоди.

□ *Непараметричні критерії* не спираються на додаткові припущення про розподіл. Зокрема, до цього типу критеріїв відноситься більшість рангових критеріїв.

Критерії згоди. Критерії згоди перевіряють, чи відповідає задана вибірка з заданим фіксованим розподілом, із заданим параметричним сімейством розподілів, або з іншого вибіркою. (Критерій χ^2 -квадрат (Пірсона)).

Критерії зсуву. Спеціальний випадок двохвибіркових критеріїв згоди. Перевіряється гіпотеза зсуву, згідно з якою розподілу двох вибірок мають однакову форму і відрізняються тільки зрушенням на константу.

Критерії нормальності. Критерії нормальності - це виділений окремий випадок критеріїв згоди. Нормально розподілені величини часто зустрічаються в прикладних задачах, що обумовлено дією закону великих чисел. Якщо про вибірки заздалегідь відомо, що вони підкоряються нормальному розподілу, то до них стає можливо застосовувати більш потужні параметричні критерії. Перевірка нормальності часто виконується на першому кроці аналізу вибірки, щоб вирішити, використовувати далі параметричні методи або непараметричні.

Критерії однорідності. Критерії однорідності призначені для перевірки нульової гіпотези про те, що дві вибірки (або кілька) взяті з одного розподілу, або їх розподілу мають однакові значення математичного очікування, дисперсії, або інших параметрів.

Критерії симетричності. Критерії симетричності дозволяють перевірити симетричність розподілу.

Критерії тренду, стаціонарності і випадковості. Критерії тренда і випадковості призначені для перевірки нульової гіпотези про відсутність залежності між вибірковими даними і номером спостереження у вибірці. Вони часто застосовуються в аналізі часових рядів, зокрема, при аналізі регресійних залишків.

Статистичний висновок як основа узагальнення даних

Статистичні висновки використовуються для узагальнення даних з вибірки щодо всієї генеральної сукупності. Випадкові помилки, характерні для вибіркового дослідження, можуть привести до того, що вибірка не буде [досить точної] моделлю генеральної сукупності. Насправді вибірка ніколи не є моделлю генеральної сукупності на всі 100%, а лише її більш-менш спотвореним варіантом. Для того, щоб оцінити такі спотворення і, отже, зробити більш точні висновки про генеральну сукупність і використовуються статистичні висновки. Перш за все, вони дозволяють оцінити ймовірність того, що виявлені у вибірці взаємозв'язку, відмінності, величини і т.д. характерні виключно для вибірки, але не для генеральної сукупності. Логіка тут наступна: якщо така ймовірність висока, то приймається рішення, згідно з яким параметри вибірки не характерні для генеральної сукупності і навпаки - якщо така ймовірність низька, то прийнято вважати, що відповідні параметри вибірки говорять про параметри генеральної сукупності.

Важливо пам'ятати - досягнення 100% гарантії того, що результати, отримані в дослідженні, характерні для генеральної сукупності, можливо лише в тому випадку, коли проведено суцільне дослідження, тобто опитування всіх представників генеральної сукупності. Але це вже не вибіркове дослідження і воно не передбачає використання статистичних висновків. У найзагальнішому вигляді статистичні висновки можна розділити на дві групи: 1) інтервальний оцінювання (побудова інтервалу, в який із заданою вірогідністю має потрапити середнє значення або пропорція

генеральної сукупності); 2) перевірка статистичних гіпотез (імовірнісний висновок про те, що певні параметри вибіркової сукупності відображають (або ж ні) параметри генеральної сукупності).

Інтервальне оцінювання. У багатьох випадках буває необхідно на основі одиничного параметра вибірки (середнє значення або пропорція) оцінити відповідний параметр генеральної сукупності. Якщо вибірка достатня велика (100 і більше спостережень), використовуючи властивості кривої нормального розподілу, а також центральну бокові теорему, можна побудувати інтервал, в який із заданою вірогідністю потрапить справжнє значення. Розподіл вибірових середніх має нормальний вигляд. Відповідно ймовірність отримання вибірки із середнім близьким до середнього значення генеральної сукупності досить висока. Але навіть в тих випадках, коли середнє для вибірки буде досить сильно відрізнятися від середнього генеральної сукупності, довірчий інтервал [в переважній більшості випадків] буде включати справжнє значення. І лише в дуже рідкісних випадках будуть отримані вибірки з настільки сильно відрізняються параметрами вибірки і генеральної сукупності, що справжнє значення не потрапить в довірчий інтервал.

Довірчі інтервали для середніх значень:

$$c.i. = \bar{X} \pm Z \frac{s}{\sqrt{N-1}} ; \quad (4.6)$$

де Z - стандартизоване значення, яке визначається рівнем альфа або p -значенням (ймовірністю того, що істинне значення не потрапить в довірчий інтервал); s - стандартне відхилення за вибіркою; n - розмір вибірки. Очевидно що s і n відомі нам з самого дослідження. У свою чергу, Z визначається за допомогою таблиці стандартизованих значень, які рекомендовано застосовувати в аналізі бізнес-процесів:

Таблиця 4.2.

Довірчий рівень	Альфа (p)	Z-значення
90%	0,10	±1,65
95%	0,05	±1,96

99%	0,01	$\pm 2,58$
99,9%	0,001	$\pm 3,29$

Довірчий рівень говорить про те, з якою ймовірністю істинне значення потрапить в побудований інтервал. Розглянемо приклад. Припустимо з вибірки ($n = 300$) відомо, що середнє значення для місячного заробітку киян становить 4000 грн., А стандартне відхилення - 75 грн. Якщо нас задовольнить ймовірність помилки дорівнює 5% (альфа - 0,05), то $Z = \pm 1,96$.

Звідси:

$$c.i. = 4000 \pm 1,96 \frac{75}{\sqrt{300-1}} = 4000 \pm 8,5; \quad (4.7)$$

Таким чином, справжнє значення для киян з імовірністю 95% має потрапити в інтервал від 3991,5грн. до 4008,5грн.

$$c.i. = P_s \pm Z \sqrt{\frac{P_u(1-P_u)}{N}}; \quad (4.8)$$

У порівнянні з попередньою, у цій формулі залишився розмір вибірки і Z -значення. Останнє визначається воно також за допомогою наведеної вище таблиці. Інші складові такі: P_s - значення пропорції для вибірки; P_u - значення пропорції для генеральної сукупності. Фактично дослідник повинен піти на явне збільшення інтервалу, адже більший інтервал з більшою ймовірністю буде включати шукану справжню величину для генеральної сукупності. Таким значенням є 0,5: $0,5 (1-0,5) = 0,5 * 0,5 = 0,25$.

Тепер розглянемо приклад. Нехай, згідно з результатами передвиборчого опитування, за опозиційну партію готові проголосувати 17% населення ($P_s = 0,17$), розмір вибірки становить 1200 осіб, а рівень альфа - 0,01. тоді:

$$c.i. = 0,17 \pm 2,58 \sqrt{\frac{0,5(1-0,5)}{1200}} = 0,17 \pm 0,02; \quad (4.9.)$$

Отже, з імовірністю в 99% за опозиційну партію проголосує від 15% до 19% населення.

Дисперсійний аналіз у дослідженні бізнес-процесів

Дисперсійний аналіз (Analysis of variance, або скорочено ANOVA) призначений для перевірок гіпотез в складних випадках. Що являє собою групу методів порівняння розмірів варіації, обумовлених різними факторами.

Дисперсію використовують не тільки для оцінки варіації, а й при вимірюванні взаємозв'язків, для перевірки статистичних гіпотез і т.п. Для ознак метричної шкали розрахунок дисперсії ведеться за формулою:

$$\sigma^2 = \frac{\sum_1^m (x_j - \bar{x})^2 f_j}{\sum_1^m f_j} = \overline{x^2} - \bar{x}^2. \quad (5.1)$$

Як і будь-яка середня, дисперсія має певні математичні властивості:

- а) якщо всі значення ознаки x_j зменшити (збільшити) на певну величину, дисперсія не зміниться;
- б) якщо всі значення ознаки змінити в k раз, то дисперсія зміниться в k^2 разів;
- в) у разі заміни частот частотами дисперсія не зміниться.

Для альтернативної ознаки, варіація якого має два взаємовиключних значення - "1" і "0", а розподіл характеризується відповідно двома частотами - d_1 і d_0 , дисперсія розраховується як твір частот $\sigma^2 = d_1 d_0 = d_1 (1 - d_1)$.

У табл. 5.1 показаний розрахунок абсолютних показників варіації на прикладі терміну обігу облігацій.

Таблиця 5.1

Термін обігу облігацій, міс.	f	x_j	$x_j - \bar{x}$	$ x_j - \bar{x} f$	$(x_j - \bar{x})^2 f$
До 2	15	1	-4,6	69,0	317,40
2—4	13	3	-2,6	33,8	87,88
4—6	29	5	-0,6	17,4	10,44
6—8	22	7	1,4	30,8	43,12
8—10	12	9	3,4	40,8	138,72
10 і більше	9	11	5,4	48,6	262,44
Всього	100	X	X	240,4	860,00

Середній термін обігу облігацій - 5,6 міс .; середнє лінійне відхилення становить $= 240,4: 100 = 2,4$ міс .; дисперсія - $s^2 = 860: 100 = 8,6$; середньоквадратичне відхилення - $s = 2,9$ міс.

Частка облігацій з терміном обігу менше 2 міс. становить $d_1 = 0,15$. Дисперсія частки $s^2 = 0,15 \cdot (1 - 0,15) = 0,1275$.

При порівнянні варіацій різних ознак або однієї ознаки в різних сукупностях, використовують відносні характеристики варіації. Коефіцієнти варіації розраховуються як відношення абсолютних, іменованих характеристик варіації (s , R) до центру розподілу і часто виражаються в процентах, отже:

лінійний коефіцієнт варіації:

$$V_t = \frac{\bar{l}}{\bar{x}} 100; \quad (5.2)$$

$$V_\sigma = \frac{\sqrt{256}}{80} * 100 = 20\%; \quad (5.3)$$

Дисперсія, на відміну від інших характеристик варіації, є адитивною величиною. Тобто в структурованій сукупності, розділеної на групи за ознакою x , дисперсія ознаки y може бути розкладена на: дисперсію в кожній групі (внутригрупповую) і дисперсію між групами (міжгруппова). Загальна дисперсія характеризує варіацію ознаки y в цілому по сукупності, міжгруппова - варіацію групових середніх, а внутрішньо групові - варіацію індивідуальних значень щодо групових середніх.

Міжгруппова дисперсія обчислюється за формулою:

$$\delta^2 = \frac{\sum_1^m (\bar{y}_j - \bar{y})^2 f_j}{\sum_1^m f_j}; \quad (5.4)$$

де \bar{y} - відповідно середня j -ї групи і загальна середня варіюючого ознаки y ; f_j - частота j -ї групи.

Внутрішньогруппова дисперсія розраховується окремо для кожної j -ї групи:

$$\delta^2 = \frac{\sum_1^{f_j} (y_j - \bar{y}_j)^2}{f_j}; \quad (5.5)$$

де y - значення ознаки в окремих елементів сукупності.

Для всіх груп в цілому обчислюється середня з внутрішньогрупових дисперсій, зважених на частоти відповідних груп:

$$\overline{\delta^2} = \frac{\sum_1^m \sigma_j^2 f_j}{\sum_1^m f_j}; \quad (5.6)$$

Взаємозв'язок між трьома дисперсіями отримав назву правила складання дисперсій, відповідно до якого

$$\sigma^2 = \delta^2 + \overline{\sigma^2}; \quad (5.7)$$

Загальну дисперсію можна визначити і безпосередньо по формулі

$$\sigma^2 = \overline{y^2} - \bar{y}^2; \quad (5.8)$$

Ставлення між груповий дисперсії до загальної називається *кореляційним відношенням*, η^2 .

Розрахунок зазначених дисперсій показаний в табл. 1 на прикладі варіації врожайності винограду.

Варіація урожайності окремо для кожного сорту винограду визначається трьома внутрішньогруповими дисперсіями. Для сорту винограду "Аліготе" внутригруповая дисперсія становить

$$\sigma_1^2 = \frac{(2,3-2,4)^2 + (2,2-2,4)^2 + (2,7-2,4)^2 + (2,6-2,4)^2 + (2,2-2,4)^2}{5} = 0,044; \quad (5.9)$$

$$\text{де } \bar{y}_1 = \frac{\sum y}{f_1} = \frac{12}{5} = 2,4$$

Таблиця 5.2

Сорт винограду	Кількість кущів, f_i	Урожайність з одного куща, кг y	$\sum y$	\bar{y}_i
Аліготе	5	2,3; 2,2; 2,7; 2,6; 2,2	12	2,4
Фетяска	8	2,9; 3,3; 2,7; 2,8; 3,2; 3,4; 2,8; 2,9	24	3,0
Рислінг	7	3,6; 4,0; 4,1; 4,4; 3,8; 3,9; 4,2	28	4,0
Всього	20	X	64	3,2

Сорт винограду Кількість кущів Урожайність з одного куща, кг $\sum y$

Аналогічно обчислені внутрішньо групові дисперсії для другої і третьої груп:
= 0,06. Середня з групових дисперсій

$$\overline{\sigma^2} = \frac{0,044*5 + 0,06*8 + 0,06*7}{20} = 0,056;$$

Міжгрупова дисперсія становить:

$$\delta^2 = \frac{(2,4-3,2)^2*5+(3,0-3,2)^2*8+(4,0-3,2)^2*7}{20} = 0,4;$$

$$\bar{y} = \frac{\sum \bar{y}_j f_j}{\sum f_j} = \frac{64}{20} = 3,2.$$

Загальну дисперсію врожайності визначимо за правилом розкладання дисперсій як суму груповий і середньої з групових дисперсій:

$$\sigma^2 = 0,4 + 0,056 = 0,456;$$

Безпосереднє обчислення загальної дисперсій за спрощеною формулою

$$\sigma^2 = \frac{213,92}{20} - 3,2^2 = 10,696 - 10,24 = 0,456;$$

дає той же результат.

У практичній діяльності однофакторний аналіз дисперсій здійснюється за допомогою *MS Excel*. Результат програма виводить у вигляді двох таблиць SUMMARY та ANOVA. При цьому вихідні дані формуються за допомогою блок команд *Пакету аналізу(Data Analysis)*[1. С.106-108]

Якщо ваші дані представляють більш складну структуру, дисперсійний аналіз може бути перетворений для відповіді на більш складні питання. [2.С.828-834].

Контрольные вопросы

1. Сформулюйте визначення поняття «статистична гіпотеза». Наведіть приклади.
2. Охарактеризуйте поняття статистичного критерію.
3. Які типові значення рівня статистичної значущості рекомендовано застосовувати в аналізі бізнес-процесів?
4. Значення статистичного висновку у бізнес-аналізі.
5. Які завдання вирішують за допомогою дисперсійного аналізу?
6. Математичні властивості загальної дисперсії.
7. Охарактеризуйте внутригрупову та між групову дисперсії.
8. Застосування пакету прикладних програм для проведення однофакторного та двофакторного аналізу.

РОЗДІЛ 2

СТАТИСТИЧНА ОЦІНКА ЕФЕКТИВНОСТІ СТАНУ РОЗВИТКУ БІЗНЕСУ В КРАЇНІ

2.1. Непараметричні методи та аналіз закономірностей для якісних даних у бізнес-аналізі

Сутність і передумови використання непараметричних методів у бізнес-статистиці. Рангова кореляція. Оцінка зв'язків між альтернативними ознаками у бізнес-статистиці. Якісні дані у бізнес-статистиці та особливості їх аналізу. Статистичний критерій хі-квадрат: характеристика та методи розрахунку. Виявлення прихованих закономірностей для вирішення бізнес-задач за допомогою Data-Mining

Сутність і передумови використання непараметричних методів у бізнес-статистиці. Рангова кореляція.

Непараметрична статистика – це самостійна галузь математичної статистики, яка набула популярності у другій половині XIX століття. Саме тоді були розроблені такі класичні методи аналізу даних, що можуть бути вираженілише у номінальній та ранговій шкалах вимірювання, про що йшла мова у темі 1.1. Методологічною основою непараметричної статистики є непараметричні методи дослідження. До них відноситься статистична перевірка гіпотез на основі даних, які не мають характеру нормального розподілу, оскільки передбачають використання частот або рангів.

Загалом непараметричні методи є більш гнучкими і не вимагають точного визначення ситуації. Прикладами можуть бути значення експорту чи імпорту продукції до і після введення митного тарифу чи квот, результати опитувань інвесторів щодо рівня інвестиційної привабливості підприємств до та після проведення реформ, ставлення чоловіків та їх дружин до купівлі-продажу нерухомості за участю маклерів тощо.

Виявлення причинно-наслідкових зв'язків у характеристиках бізнес-процесів та прийняття ефективних управлінських рішень часто потребують використання математичного апарату, який доволі простий та якому

притаманна стійкість висновків. Найбільш популярним є коефіцієнт кореляції рангів Спірмена, який було названо на честь Чарльза Едварда Спірмена – англійського психолога, відомого своїми працями у галузі статистики.

Розглянемо розрахунок цього коефіцієнта на прикладі.

Приклад. За даними таблиці. 6.4 оцінимо тісноту зв'язку між рівнем ефективності економіки і надійністю ділового партнерства для семи країн Східної Європи. Оскільки інформація представлена у формі інтегральних показників (оцінки в балах), необхідно провести ранжування країн. Найменшим значенням інтегрального показника надається ранг 1, найбільшому - ранг 7. Сума квадратів відхилень рангів становить

$$\sum_1^n d_j^2 = 10,$$

а коефіцієнт рангової кореляції

$$\rho = 1 - \frac{6 \cdot 10}{7(7^2 - 1)} = 1 - \frac{60}{336} = 0,8;$$

Значення коефіцієнта рангової кореляції свідчить про наявність прямого і досить заповітної зв'язку між зазначеними параметрами ризику іноземного інвестування економіки. За дод. 6 критичне значення коефіцієнта рангової корекції для $\alpha = 0,05$ і $n = 7$ становить $r_{0,95}(7) = 0,71$, що значно менше фактичного. Отже, істотність зв'язку доведена з імовірністю 0,95.

Таблиця 6.4

Країна	Інтегральні показники		Ранги показників		Відхилення рангів d_j	d_j^2
	Ефективність економіки (max = 10)	Надійність ділового партнерства (max = 100)	R_x	R_y		
A	5,9	54,9	6	7	-1	1
B	7,1	54,8	7	6	1	1
C	4,2	45,3	4	5	-1	1
D	3,4	36,9	3	4	-1	1
K	4,9	35,8	5	3	2	4
M	2,7	26,4	1	2	-1	1
P	2,9	24,8	2	1	1	1
Всього	X	X	X	X	0	10

Оцінка зв'язків між альтернативними ознаками у бізнес-статистиці

Аналіз взаємозв'язків між атрибутивними ознаками проводиться на основі таблиць взаємної спряженості (взаємозалежності), що описують комбінаційні розподіли сукупностей за двома ознаками - факторному x і результативному y . При наявності стохастичною зв'язку умовні розподіли змінюються від групи до групи. Оцінка тісноти стохастичною зв'язку ґрунтується на відхиленнях частот (частостей) умовних розподілів від безумовного, тобто на відхиленнях фактичних частот f_{ij} від теоретичних F_{ij} , пропорційних підсумковим частотам безумовного розподілу:

$$F_{ij} = \frac{f_{i0}f_{0j}}{n},$$

де f_{i0} - підсумкові частоти за ознакою x ; f_{0j} - підсумкові частоти за ознакою y , n - обсяг сукупності.

Очевидно, що

$$n = \sum_i^{m_x} f_{i0} = \sum_j^{m_y} f_{0j};$$

Абсолютну величину відхилень ($f_{ij} - F_{ij}$) характеризує квадратична спряженість Пірсона χ^2 :

$$\chi^2 = \sum_i \sum_j \frac{(f_{ij} - F_{ij})^2}{F_{ij}} = n \left[\sum_i \sum_j \frac{f_{ij}^2}{f_{i0}f_{0j}} - 1 \right];$$

Якщо стохастична зв'язок відсутній, то $\chi^2 = 0$. Для того, щоб зробити висновок про суттєвості зв'язку, необхідно порівняти фактичне значення χ^2 з критичним для заданої ймовірності $1 - \alpha$ і числа ступенів свободи $k = (m_x - 1)(m_y - 1)$, де m_x і m_y - відповідно кількість груп за ознаками x та y .

Відносною мірою тісноти стохастичною зв'язку служать коефіцієнти взаємної спряженості C , які за змістом ідентичні коефіцієнтам кореляції. Якщо $m_x = m_y$, використовують коефіцієнт спряженості Чупрова:

$$C = \sqrt{\frac{\chi^2}{n\sqrt{(m_x-1)(m_y-1)}}},$$

якщо $m_x \neq m_y$, перевагу віддають коефіцієнту спряженості Крамера:

$$C = \sqrt{\frac{\chi^2}{n(m_{\min}-1)}};$$

де m_{\min} - мінімальна кількість груп за ознакою x або y .

Значення коефіцієнта C коливаються в межах від 0 до 1.

Приклад. У табл. 6.5 наведено комбінаційний розподіл респондентів як потенційних інвесторів за віком і схильності до ризику. До групи ризикують віднесені респонденти, які мають намір придбати цінні папери, незважаючи на ризик, обережні не ризикують без гарантій, неріскующі уникають ризику взагалі. Концентрація частот близько діагоналі з верхнього лівого кута в правий нижній свідчить про наявність стохастичного зв'язку. Фактичне значення χ^2 становить:

$$\chi^2 = 200 \left[\frac{24^2}{40 \cdot 50} + \frac{12^2}{40 \cdot 80} + \frac{4^2}{40 \cdot 70} + \frac{20^2}{100 \cdot 50} + \frac{50^2}{100 \cdot 80} + \frac{30^2}{100 \cdot 70} + \frac{6^2}{60 \cdot 50} + \frac{18^2}{60 \cdot 80} + \frac{36^2}{60 \cdot 70} \right] = 49,5$$
що значно вище критичного $\chi^2_{0,95}(4) = 9,49$. Отже, істотність зв'язку між віком респондентів і схильністю їх до ризику доведена з імовірністю 0,95.

Таблиця 6.5

Вік, років	Тип інвестора			Всього
	Ризиковий	Обережний	Не ризиковий	
До 30	24	12	4	40
30—50	20	50	30	100
50 і більше	6	18	36	60
Всього	50	80	70	200

Оскільки $m_x = m_y = 3$ для оцінки тісноти зв'язку використовуємо коефіцієнт взаємної спряженості Чупрова:

$$C = \sqrt{\frac{49,5}{200 \sqrt{2 \cdot 2}}} = 0,352;$$

тобто зв'язок між ознаками помірна.

У разі, коли $m_x = m_y = 2$, розрахунок коефіцієнта взаємної спряженості спрощується:

$$C = \sqrt{\frac{\chi^2}{n}} = \frac{f_{11}f_{22} - f_{12}f_{21}}{\sqrt{f_{01}f_{02}f_{10}f_{20}}}$$

У статистичній літературі коефіцієнт С для 4-клітинної таблиці називають *коефіцієнтом контингенції (асоціації)*. Очевидно, що $\chi^2 = nC^2$.

Для аналізу таблиць такого типу використовують також ставлення перехресних творів, або *відношення шансів*:

$$W = \frac{f_{11}f_{22}}{f_{12}f_{21}}$$

Ставлення шансів характеризує міру відносного ризику (виграшу), пов'язаного з фактором х.

Приклад. За даними табл. 6.6 оцінимо тісноту зв'язку між сприйняттям реклами та придбанням рекламованого товару, а також результативність реклами.

Таблиця 6.6

Сприйняття реклами	Кількість респондентів		Всього
	Купили товар	Не купили товар	
Запам'ятали рекламу	9	31	40
Не запам'ятали рекламу	2	38	40
Всього	11	69	80

Коефіцієнт контингенції свідчить про наявність стохастичного зв'язку

$$C = \frac{9 \cdot 38 - 2 \cdot 31}{\sqrt{11 \cdot 69 \cdot 40 \cdot 40}} = 0,254.$$

фактичне значення $\chi^2 = nC^2 = 80 \cdot 0,2542 = 5,16$, що вище критичної позначки $\chi^2_{0,95}(1) = 3,84$. Отже, істотність зв'язку доведена.

Ставлення шансів становить

$$W = \frac{9 \cdot 38}{2 \cdot 31} = 5,5.$$

тобто шанси реалізувати рекламований товар в 5,5 рази вище в порівнянні з нерекламованим.

Методи аналізу таблиць взаємної спряженості можна використовувати і для кількісних ознак, проте слід зазначити, що заходи тісноти кореляційної зв'язку - коефіцієнт детермінації і кореляційне відношення більш чутливі.

Якісні дані у бізнес-статистиці та особливості їх аналізу.

Якісні дані, як і кількісні, описують властивості елементів, процесів чи явищ. Для обробки масивів якісних даних часто використовують статистичні, комбінаторні та евристичні показники. Перелік показників, які використовують для аналітичного оцінювання якісних даних подано у таблиці

Таблиця - **Статистичні показники аналізу якісних даних у бізнесі** [BS с. 142]

Параметри для визначення	Показники	Характеристика	Приклади використання
Середнього значення сукупності	Мода та середній нормований бал	Типовий рівень варіативної ознаки	Модальне значення освіти працівників
Варіації ознак	Дисперсія альтернативної ознаки	Коливання, мінливість значень ознаки	Дисперсія якісної продукції (контроль якості продукції)
Концентрації, диференціації	Коефіцієнт локалізації, коефіцієнт концентрації	Ступінь концентрації ознак у сукупності загалом, в окремих групах	Коефіцієнт концентрації власного капіталу
Аналітичні оцінки взаємозв'язків між ознаками	Коефіцієнти взаємного спряження, коефіцієнт рангової кореляції	Щільність зв'язку між ознаками	Вплив віку менеджерів на прийняття рішень на основі інтуїції

Очевидно, що перелік показників, які обчислюють у бізнес-аналізі даних, є дещо обмежений. Це зумовлено статистичною природою атрибутивних ознак. Статистичні висновки щодо якісних даних, у випадку, якщо результати спостереження подано не числами, а категоріями, роблять на основі перевірки гіпотез.

Статистичний критерій хі-квадрат: характеристика та методи розрахунку

Критерій хі-квадрат використовують для перевірки гіпотез про якісні дані, представлених не числовими, а категоріальними змінними. Тут прийнято оперувати підрахунком частоти (оскільки ранжування або арифметичні дії виконувати неможливо). Критерій (тест) «хі-квадрат»

заснований на частотах, які представляють собою кількість об'єктів вибірки, що потрапляють в ту чи іншу категорію. Суть показника χ^2 -квадрат (χ^2): він вимірює різницю між що спостерігаються (експериментальними) частотами f_E і очікуваними (теоретичними) частотами f_T . Саме він розраховується як сума квадратів різниці цих частот, виражена в частках частоти теоретичної.

Використання такого статистичного підходу можливо в різних обставинах. Розглянемо найбільш поширені.

Комбінація: нинішні і минулі події (критерій « χ^2 -квадрат» відповідності)

Даний спосіб широко застосовується в тих випадках, коли потрібно визначити, чи є наш нинішній досвід (виражений в частотах або відсотках) типовим по відношенню до минулого досвіду (набір так званих опорних величин). Таку ситуацію можна умовно позначити фразою «Це було недавно, а то було давно. Між ними є відповідність? »Тест« χ^2 -квадрат »щодо відповідності відсотків використовується для перевірки гіпотези про те, що комбінація спостережуваних частот або відсотків (що характеризують одну якісну змінну) побудована на даних з деякою генеральної сукупності з уже відомими значеннями відсотків (опорними величинами). Можна сформулювати висловлені міркування і по-іншому: ті результати, які ми спостерігаємо зараз (фактичні дані, тобто наш нинішній досвід), за характером аналогічні минулим даними (опорним величинам). Це пояснюється тим, що і ті, і інші ставляться до однієї і тієї ж генеральної сукупності, але витягувалися в різний час (зараз і колись давно). Очікуване значення частоти для кожної категорії розраховується як твір заданого опорного значення відсотка в генеральній сукупності на розмір вибірки n . На підставі наявних знань про що спостерігається очікуваної частотами аналізованого події визначається власне показник χ^2 -квадрат. Розрахункове значення χ^2 -квадрат потім порівнюють з критичним (табличним) показником для відповідного числа ступенів свободи, який визначається як кількість категорій мінус одиниця. Якщо виявляється справедливим нерівність $\chi^2_{\text{расч}} >$

$\chi^2_{кріт}$, то із заданою вірогідністю (або рівнем значущості) можна стверджувати, що спостерігаються частоти (наш досвід) значимо відрізняються від тих, які очікуються виходячи з відомих нам опорних значень відсотків (частот). Отже, обґрунтовано можна робити висновок про те, що спостерігаються вибіркові відсотки значимо відрізняються від заданих опорних значень. Якщо маємо співвідношення $\chi^2_{расч} < \chi^2_{кріт}$, то спостережувані значення незначно відрізняються від опорних показників і, отже, наші фактичні результати не мають значущих відмінностей від заданих опорних значень. При виконанні такого аналізу прийнято дотримуватися наступного емпіричного правила: очікувані частоти в кожній категорії повинні бути, по крайній мірі, не менше п'яти (оскільки критерій хі-квадрат залишається приблизною, а не точною оцінкою). Аналіз критерію відповідності відсотків (частот) зручно виконувати, дотримуючись такої схеми.

1. Є табличні дані частот для кожної категорії однієї якісної змінної. Обговорюються такі гіпотези:

а) частоти (відсотки) нинішнього досвіду рівні набору відомих, фіксованих опорних величин (з минулого досвіду);

б) частоти (відсотки) нинішнього досвіду не рівні набору опорних величин (даних минулого досвіду).

2. Очікувані частоти обчислюються так: потрібно для кожної категорії помножити відоме значення її частки в загальній кількості (генеральної сукупності) на розмір вибірки n . При цьому передбачається, що а) набір даних представляє собою випадкову вибірку з даної генеральної сукупності і б) очікується наявність, принаймні, п'яти об'єктів в кожній з категорій.

3. Проводиться аналіз «хі-квадрат».

4. Інтерпретація результату тесту "хі-квадрат": наявність значущої зв'язку відзначається тоді, коли розрахункове значення "хі-квадрат" більше табличного або критичного (тобто $\chi^2_{расч} > \chi^2_{кріт}$), в іншому випадку значущою зв'язку немає. Тепер приступимо до конкретного аналізу критерію

відповідності частот. Розглянемо наступний приклад. Серед студентів металургійного факультету, які здавали на першому курсі в літню сесію іспит з математики, було проведено опитування з метою з'ясування того, які чинники впливають на отримання незадовільної оцінки. Число опитаних студентів становило 50 осіб. Найбільш часто згаданими причинами були наступні:

1. Сам винен, треба було краще займатися.
2. Строгий викладач.
3. На жаль, не вдалося списати.
4. Позначилося вплив фатальних прийме (дістався квиток № 13, по-зустрічав чорного кота, забув надіти «щасливий» светр і ін.).

Ці відповіді можна умовно розділити на наступні категорії:

1. Сам винен.
2. Дуже строгий викладач.
3. Шпаргалки.
4. Прикмети.

У табл.1.1 наведені дані про причини отримання «незадовільно» з математики за минулу сесію, а також вказані значення опорних величин, взяті з екзаменаційних відомостей з цього предмету за минулі роки (по таким же категоріям). Як видно з даних таблиці, за кількісними показниками всі аналізовані причини формально відрізняються від опорних значень. Однак ця відмінність виявляється далеко нерівноцінних. Так, можна визнати, що в категорії самооцінки («Сам винен») фактичні дані відрізняються від відповідних опорних величин відносно слабо (наприклад, 57% в порівнянні з 59% для минулих сесій). У той же час по іншим категоріям відносна відмінність виглядає більш помітним. Особливо впадає в очі невідповідність по позиції «Шпаргалки».

Таблиця 7.1.

Підсумкові дані про причини отримання незадовільної оцінки з математики за дану сесію і сесії минулих років

Причина	Спостережувані дані (за минулу сесію)		Опорні значення, % (Очікувані дані)
	Частота	Відсоток від загального числа	
Сам винуватий	28	57,0	59,0
Дуже строгий викладач	10	19,0	14,0
Шпаргалки	7	14,0	20,0
Прикмети	5	10,0	7,0
Всього:	50	100	100

Питання полягає в тому, значима ця різниця? Інакше кажучи, чи можуть отримані за підсумками минулої сесії «незадовільно» розглядатися як результат вилучення випадкової вибірки з генеральної сукупності, в якій відсотки «незадовільно» відповідають опорним величинам? Або ще по-іншому: чи достатньо велика спостерігається різниця, щоб її не можна було пояснити тільки випадковістю? Тест хі-квадрат відповідності відсотків дозволить дати відповідь на це питання. Стверджувальне висновок отримаємо за умови, коли виявиться справедливим співвідношення $\chi^2_{расч} > \chi^2_{кріт}$. Його потрібно буде тлумачити так: результати нинішньої сесії і результати минулих сесій відрізняються між собою принципово, оскільки відмінність між ними не носить випадкового характеру. Якщо виявиться справедливим нерівність $\chi^2_{расч} < \chi^2_{кріт}$, то із заданою вірогідністю можна буде говорити про не значимі відмінності між аналізованими результатами. У табл.7.2 вкажемо частотні величини для обох інформаційних позицій - поточні дані («Спостереження») і відомості за минулі роки («Очікування»). Розрахунок частот для графі «Очікування» (тобто очікувані частоти) проведемо шляхом множення значень опорних величин відсотків (59%, 14%, 20% і 7%) на розмір вибірки ($n = 50$). В результаті отримаємо такі значення

частот: $0,59 \times 50 = 29,5$; $0,14 \times 50 = 7,0$ і т.д. Зауважимо, що в підсумковому рядку для обох колонок загальна сума частот однакова - дорівнює 50.

Таблиця 7.2

**Спостережувані і очікувані дані (частоти) про причини
незадовільних оцінок**

Причина	Спостереження	Очікування
Сам винуватий	28	29,5
Дуже строгий викладач	10	7,0
Шпаргалки	7	10,0
Прикмети	5	3,5
Всього	50	50,0

Ці дані і будемо використовувати для вирішення питання про значне відповідно (або невідповідність) фактичних і очікуваних результатів.

Про коефіцієнти взаємної спряженості

На основі хі-квадрата прийнято також оцінювати показники ступеня тісноти зв'язку - коефіцієнти взаємної спряженості К.Пирсона і А.Чупрова.

Коефіцієнт Пірсона розраховується за формулою:

$$K_{\pi} = \sqrt{\frac{\chi^2}{n + \chi^2}};$$

де χ^2 - розрахункове значення хі-квадрата, n - загальне число спостережень (обсяг вибірки).

Коефіцієнт Чупрова дозволяє врахувати число груп по кожному ознакою і визначається наступним чином:

$$K_{\text{ч}} = \sqrt{\frac{\chi^2}{n \sqrt{(k_1 - 1)(k_2 - 1)}}};$$

де k_1 і k_2 - відповідно число значень (груп) для першого і другого ознак або, по-іншому, число рядків і стовпців в таблиці, а n - загальне число спостережень (обсяг вибірки). Спробуємо виконати такі розрахунки для нашого прикладу.

$$K_{\Pi} = \sqrt{\frac{\chi^2}{n+\chi^2}} = \sqrt{\frac{2,905}{50+2,905}} = 0,234;$$

$$K_{\chi} = \sqrt{\frac{\chi^2}{n\sqrt{(k_1-1)(k_2-1)}}} = \sqrt{\frac{2,905}{50*\sqrt{(4-1)(2-1)}}} = 0,184;$$

Розрахунок обох коефіцієнтів дає вельми малі величини, що свідчить про відсутність зв'язку між досліджуваними характеристиками. Це ж підтверджують і оцінки по таблиці Чеддока: розраховані коефіцієнти, по модулю менші 0,3, говорять про відсутність кореляційної зв'язку. Інакше кажучи, використання і цих коефіцієнтів підтверджує раніше висловлену міркування: досліджуєма ситуація за своїми параметрами відповідає опорним (очікуваним) показниками і тому не вимагає введення будь-яких коригувань.

Статистичний критерій хі-квадрат: характеристика та методи розрахунку.

**Перевірка наявності взаємозв'язку між двома якісними змінними
(критерій «хі-квадрат» незалежності)**

Можливі ситуації, коли є дві якісні змінні, тобто набір експериментальних даних являє собою двовимірні якісні дані. Після вивчення кожної з них окремо за допомогою аналізу частот (або відсотків) може виникнути питання про наявність зв'язку між ними. Вважається, що дві якісні змінні є незалежними, якщо знання значення однієї змінної не допомагає передбачити значення іншої. Уявімо собі, що ваша фірма розробила технологію гальванічного покриття нікелем сталевих деталей автомобільного кузова. В середньому відсоток браку, пов'язаного з відшаруванням покриття, становить 3,1%. Однак коли працює технолог Іванов, розмір браку досягає 11,2%. У цьому випадку знання значення однієї змінної (ім'я конкретного технолога) допомагає спрогнозувати значення іншої змінної (обсяг шлюбу певного типу), оскільки 3,1% і 11,2% розрізняються між собою. Отже, ці дві змінні не є незалежними. Використання критерію «хі-квадрат» дозволяє вирішити питання про те, чи є що розглядаються якісні сукупності залежними або ж незалежними один від одного. У цьому випадку застосовується так званий критерій «хі-квадрат»

незалежності, який встановлює наявність (або відсутність) зв'язку між двома якісними змінними. Для такого аналізу використовується таблиця частот, які можна було б очікувати в тому випадку, якщо змінні виявилися б незалежними. У загальному випадку критерій «хі-квадрат» незалежності прийнято представляти наступним чином:

1. Є вихідні дані у формі табличного списку частот всіх комбінацій категорій двох якісних змінних. Обговорюються такі гіпотези:

- а) дві змінні не залежать одна від одної;
- б) дві змінні пов'язані, вони не є незалежними один від одного.

2. Складається таблиця очікуваних (теоретичних) частот. Для їх розрахунку частоту однієї категорії (результат експерименту) слід помножити на частоту іншої категорії (також експериментальний показник) і отриманий добуток поділити на загальний обсяг вибірки n :

$$\text{Очікувана частота } f_{\text{оч(т)}} = \frac{\text{Частота категорії } f_{e1} \times \text{Частота категорії } f_{e2}}{\text{Загальний об'єм вибірки } n};$$

$$\text{або більш компактно, в символічній формі: } f_{\text{оч(т)}} = \frac{f_{e1} \times f_{e2}}{n};$$

При цьому вважається, що а) набір даних представляє собою випадкову вибірку з даної генеральної сукупності і б) для кожної комбінації категорій очікувана частота, по крайній мере, не менше п'яти. 3. Далі проводиться аналіз «хі-квадрат», розрахунок виконується з використанням знайомого виразу:

$$\chi^2 = \sum \frac{(f_e - f_t)^2}{f_t};$$

Ступінь свободи обчислюється таким чином:

$$f = (k_1 - 1) \times (k_2 - 1);$$

де k_1 і k_2 - число категорій відповідно для першої та другої змінної.

4. Результат тесту «хі-квадрат» трактується так: наявність значущої зв'язку проявляється тоді, коли розрахункове значення «хі-квадрат» більше критичного (тобто $\chi^2_{\text{расч}} > \chi^2_{\text{крит}}$), в іншому випадку значущою зв'язку немає. Давайте познайомимося з цим видом статистичного аналізу, для чого розглянемо наступний приклад. Кіт Матроскін, зайнявшись молочним

бізнесом, вирішив провести маркетингове дослідження, щоб з'ясувати, який вид молочної продукції вважають за краще ті чи інші покупці. Для кожної покупки фіксувалися дві якісні змінні - вид продукції і тип покупця. Як продається молочної продукції фігурували молоко, сметана і сир. Покупців Матроскін умовно розділив на дві категорії - практичні і імпульсивні. До перших він відніс тих покупців, які йдуть на ринок уже з чітко сформульованим наміром щодо того, що купити і скільки саме. Другу ж категорію склали покупці, які рішення приймають на місці, безпосередньо перед покупкою. Отримані дані статистичного опитування Матроскін представив в табличній формі (табл.1.3), в якій для кожного виду молочної продукції вказав кількість скоєних покупок тих чи інших покупцем, тобто привів фактичну частоту. Суд має надати висновок за підсумками статистичної перевірки за критерієм «хі-квадрат», тобто сформулювати висновок і пояснити результат з практичної точки зору - визначити яку ринкову стратегію має обрати кіт Матроскін і, отже, на якого покупця і на який вид молочної продукції йому належить орієнтуватися.

Таблиця 7.3

Результати опитування про перспективи молочного бізнесу

Вид молочної продукції	Частота переваг	
	Практичний покупець	Імпульсивний покупець
Молоко	38	15
Сметана	24	31
Творог	18	27

Для цього доповнимо таблицю з вихідними даними підсумкової рядком і додатковим «сумарним» стовпцем, заповнимо їх, виконавши нескладні розрахунки (табл.7.4):

Таблиця 7.4

**Доповнені дані за результатами опитування
про перспективи молочного бізнесу**

Вид молочної продукції	Частота переваг		Всього
	Практичний покупець	Імпульсивний покупець	
Молоко	38	15	53
Сметана	24	31	55
Творог	18	27	45
Всього	80	73	153

Чисто візуально важко відповісти, чи є взаємозв'язок між цими ознаками: різними категоріями покупців і видами молочної продукції. Тому необхідно дати аналіз розподілу частот в таблиці по рядках і графах. Будемо виходити з такого становища. Якщо ознака, покладений в основу угруповання по рядках (вид молочної продукції), не залежить від ознаки, покладеної в основу угруповання за стовпцями (тип покупця), то в кожному рядку (стовпці) розподіл частот має бути пропорційно розподілу їх в підсумковому рядку (стовпці). Такий розподіл можна розглядати як теоретичне (очікуване), частоти якого розраховані в припущенні відсутності зв'язку між досліджуваними сукупностями.

Розрахуємо очікувані частоти всередині таблиці пропорційно розподілу частот в підсумковому рядку. Так, молоко як один з видів молочної продукції в залежності від поведінки відвідувачів ринку по частоті потрапляння в категорії «Практичний покупець» і «Імпульсивний покупець» має наступні показники:

$$f_{11} = \frac{53 \times 80}{153} = 27,7;$$

$$f_{12} = \frac{53 \times 73}{153} = 25,3;$$

Для другого рядка, тобто для категорії сметана, ці показники мають вже такі значення:

$$f_{21} = \frac{55 \times 80}{153} = 28,8;$$

$$f_{22} = \frac{55 \times 73}{153} = 26,2;$$

Для третього рядка (категорія творог):

$$f_{31} = \frac{45 \times 80}{153} = 23,5;$$

$$f_{32} = \frac{45 \times 73}{153} = 21,5.$$

Отримані результати (обчислені значення частот) помістимо в табл 7.5

Таблиця 7.5

Дані про перспективи молочного бізнесу з урахуванням очікуваних частот

Вид молочної продукції	Очікувана частота переваг		Всього
	Практичний покупець	Імпульсивний покупець	
Молоко	27,7	25,3	53
Сметана	28,8	26,2	55
Творог	23,5	21,5	45
Всього	80	73	153

Розрахункове значення критерію хі-квадрат визначимо за формулою:

$$\chi^2 = \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} \frac{(f_{ij} - f_{ij}^*)^2}{f_{ij}^*};$$

де f_{ij} і f_{ij}^* - відповідно фактичні та теоретичні (очікувані) частоти в i -му рядку і j -го стовпця; k_1 і k_2 - відповідно число категорій в рядках і стовпцях таблиці.

Виконаємо відповідні розрахунки:

$$\chi_{\text{розрах}}^2 = \frac{(38-27,7)^2}{27,7} + \frac{(15-25,3)^2}{25,3} + \frac{(24-28,8)^2}{28,8} + \frac{(31-26,2)^2}{26,2} + \frac{(18-23,5)^2}{23,5} + \frac{(27-21,5)^2}{21,5} = 12,4$$

Далі потрібно було порівняти розрахункове значення $\chi^2_{\text{расч}}$ з табличним показником (зазвичай для рівня значущості 0,05 або 0,01). У розглянутому прикладі число ступенів свободи дорівнює двом, тобто $f = (3 - 1)(2 - 1) = 2$. При $\alpha = 0,05$ табличне значення $\chi^2_{\text{табл}}$ при $f = 2$ складає 5,991, а для $\alpha = 0,01$ відповідно 9,210. Оскільки $\chi^2_{\text{расч}} > \chi^2_{\text{табл}}$, то з упевненістю на 95% (навіть на 99%) можна стверджувати, що вплив

психологічного типу покупця очевидним чином позначається на результатах комерційної діяльності кота Матроскіна.

Виявлення прихованих закономірностей для вирішення бізнес-задач за допомогою Data-Mining

Проведення автоматичного аналізу даних передбачає використання Data-Mining (добування знань) – технологію інтелектуального аналізу даних щодо виявлення прихованих закономірностей у вигляді значущих особливостей кореляцій, тенденцій, шаблонів тощо.

Data-Mining – міждисциплінарна галузь, що виникла та розвинулась на основі статистики, теорії баз даних та штучного інтелекту.

Технології Data-Mining гармонійно поєднує формалізовані та непараметричні методи аналізу, які впливають на її можливість розв'язати такі задачі: класифікацію (classification), кластеризацію (clustering), прогнозування (forecasting), асоціацію (associations), візуалізацію (visualization), аналіз та виявлення відхилень (analysis and identification of deviations), аналіз взаємозв'язків (correlation analysis), підведення підсумків (summarization) тощо.

Data-Mining складається з трьох стадій:

1. Виявлення закономірностей (вільний пошук).
2. Використання виявлених закономірностей для прогнозу невідомих значень (прогностичне моделювання).
3. Аналіз виключень (виявлення та пояснення аномалій, знайдених у закономірностях).

Data-Mining широко використовують щодо вирішення бізнес-задач у: банківській справі, сфері електронної комерції, страховій та фондовій діяльності, а також в маркетингових дослідженнях.

Контрольні питання

1. Що вивчає непараметрична статистика?
2. Назвіть переваги та недоліки непараметричних методів аналізу бізнес-процесів.

3. Як можна визначити наявність та напрям зв'язку за допомогою коефіцієнта кореляції рангів Спірмена?
4. Які коефіцієнти у бізнес-статистиці розраховують за допомогою «чотириклітинної» таблиці?
5. Що характеризує відношення шансів?
6. Статистичні показники аналізу якісних даних у бізнесі.
7. Що характеризує статистичний критерій χ^2 ?
8. Використання Data-Mining в аналізі бізнес-процесів.

2.2. . Статистичне оцінювання ефективності бізнес-процесів

Поняття ефективності та реінжинірингу бізнес-процесів. Інформаційне забезпечення аналізу бізнес-процесів. Методика оцінки малого бізнесу.

Статистичне оцінювання ефективності бізнес-процесів

Ефективність в економіці часто асоціюють з оптимальністю цілеспрямованої дії, способу, механізму реалізації або стану суб'єкту за наявності альтернатив. Вона проявляється через співвідношення між метою та отриманими результатами, використаними ресурсами, цілями, обставинами зовнішнього середовища та часовими рамками. Що ж стосується ефективності бізнес-процесів, то вона показує, наскільки добре виконуються процеси, які тільки через удосконалення можуть бути більш ефективними.

Покращення бізнес-процесів і відповідно їх ефективності, здійснюється шляхом реінжинірингу. Родоначальником терміну вважають Майкла Хаммера [БС, С. 245]. За його визначенням, реінжиніринг – це фундаментальне переосмислення і радикальне перепроєктування бізнес-процесів з метою досягнення істотного поліпшення ефективності їх функціонування. Досягти оцінки функціонування процесу можливо лише на основі досконалого бізнес-аналізу.

Аналіз повинен бути комплексним дослідженням внутрішніх і зовнішніх чинників, що впливають на процес функціонування бізнес-процесів підприємства, оскільки основна мета проведення аналізу –

підвищення ефективності функціонування бізнес-процесів та прийняття оптимальних управлінських рішень по забезпеченню їх стійкого розвитку.

Для успішної реалізації мети аналізу бізнес-процесів необхідно вирішити наступні завдання:

- визначення економічної ефективності використання ресурсів бізнес-процесів;
- оцінка виконання планів, прогнозів, управлінських рішень, за ефективним використанням економічного потенціалу бізнес-процесів підприємства;
- вивчення впливу об'єктивних і суб'єктивних, зовнішніх і внутрішніх чинників на результати бізнес-процесів підприємства;
- розробка й обґрунтування заходів спрямованих на оптимізацію бізнес-процесів підприємства і прийняття ефективних управлінських рішень.

Необхідною умовою якісного проведення аналізу бізнес-процесів є його інформаційне забезпечення.

Інформаційне забезпечення аналізу – це сукупність інформаційних ресурсів, необхідних для досягнення цілей аналізу.

Як інформаційні ресурси виступає інформація, що циркулює в системі управління і є основою управлінських рішень.

Інформаційне забезпечення бізнес-аналізу можна розділити на наступні основні блоки:

1. Нормативно-довідкова – законодавчі акти, нормативи урядових і старанних структур, норми, нормативи, ліміти і інші регламенти бізнес-процесів.
2. Планова інформація – міститься в бізнес-плані підприємства або в плані економічного (соціального) розвитку підприємства.
3. Фактографічна інформація – є сукупність даних економічного характеру про факти, які дійсно здійснювалися у бізнес-процесах підприємства. Ця інформація, як правило, формується обліково-

фінансовою службою. Фактографічну інформацію можна підрозділити на три блоки:

- Облікова інформація, що міститься в облікових документах підприємства - первинних документах, регістрах бухгалтерського обліку, відомості і ін.
- Звітна інформація, що міститься в статистичній і фінансовій звітності, наприклад, форма 1 "Баланс", форма 2 "Звіт про фінансові результати", "Звіт про працю" і тому подібне. Саме цей вид інформації має основне значення при проведенні зовнішнього аналізу.
- Позаоблікова інформація, що міститься в актах перевірок, аудиторських висновках, а також в документах, регулюючих бізнес-процеси підприємства (офіційні документи, договори, угоди, рекламації, технічна і технологічна інформація).
- Нормативно-довідкова інформація, яка регулює усі бізнес-процеси підприємства це Закон України "Про бухгалтерський облік та фінансову звітність", Положення (стандарти) бухгалтерського обліку в Україні, господарсько-правові документи, технічна й технологічна документація, інформація, отримана в ході маркетингових досліджень, тощо.

Джерелами інформації для аналізу процесу постачання матеріальних ресурсів та процесу розрахунків з покупцями виступають:

- довідкова інформація: плановані обсяги запасів та обсяги надходження (портфель замовлень), план матеріально-технічного постачання, нормативи товарних запасів, нормативи товарних витрат.
- облікова інформація: накладні на відвантаження продукції; товаротранспортні накладні; рахунки; журнал реєстрації виставлених рахунків – фактур; оперативні дані відділу матеріально-технічного постачання; відомості, звіти матеріально-відповідальних осіб, регістри синтетичного обліку по рахункам "Каса", "Товари", "Розрахунки з постачальниками".

- звітна інформація: форми №1 "Баланс", форма №2 "Звіт про фінансові результати", форми статистичної звітності № 1-СН, 3-СН, 4-СН, 11-СН, 12-СН про наявність і використання матеріальних ресурсів, ф.№ 5-з про витрати на виробництво.
- позаоблікова інформація: акти експертів, протоколи зборів, договори, угоди, контракти на постачання сировини та матеріалів.

Джерелами інформації для аналізу процесу реалізації готової продукції є:

- довідкова інформація: індекси цін на споживчі товари, нормативи товарних запасів, нормативи товарних витрат, нормативи обороту на 1 і 2 торговій площі.
- планова інформація: данні бізнес-плану, план реалізації у розрізі товарних груп, план постачання товарів.
- облікова інформація: накладні на відвантаження продукції; товаротранспортні накладні; рахунки – фактури; журнал реєстрації виставлених рахунків – фактур; книга продажів; відомості, звіти матеріально-відповідальних осіб, регістри синтетичного обліку по рахункам "Каса", "Товари", "Розрахунки з постачальниками".
- звітна інформація: форми №1 "Баланс", форма №2 "Звіт про фінансові результати", Форма №3 - торг - (коротка) місячна, Форма №7 - торг (роздріб) - "Звіт про наявність торгової мережі" - річна, Форма №1-торг "Звіт про товарообіг".
- позаоблікова інформація: акти експертів, протоколи зборів, договір на поставку товарів.

У зв'язку з особливостями етапу становлення малого бізнесу країни і недостатнім рівнем розвитку статистики малого бізнесу у деяких випадках використовуються показники, які можуть дати опосередковану інформацію.

1. Показники, що відображають стан сектора МСБ в цілому і за окремими галузями:

- розподіл числа малих і середніх підприємств за галузями економіки;
- розподіл показника "частка малих/середніх підприємств в загальній

кількості підприємств в регіоні" за галузями економіки;

- розподіл показника "кількість малих/середніх підприємств на 1000 чоловік населення регіону" за галузями економіки;
- розподіл середньої чисельності працівників малих і середніх підприємств за галузями економіки;
- розподіл показника "середня чисельність працівників на одному малому/середньому підприємстві" за галузями економіки;
- розподіл обсягу виробництва продукції (робіт, послуг) малих підприємств за галузями економіки.

2. Показники, що відображають ефективність функціонування сектора МСБ в цілому і за окремими галузями:

- розподіл виручки від реалізації продукції (робіт, послуг) малих і середніх підприємств, зокрема у розрахунку на одне підприємство, за галузями економіки;
- розподіл прибутку малих і середніх підприємств, зокрема у розрахунку на одне підприємство, за галузями економіки;
- розподіл показника "частка прибуткових малих/середніх підприємств у галузі" за галузями економіки.

3. Показники ефективності використання трудових ресурсів сектора МСБ в цілому і за окремими галузями:

- розподіл продуктивності праці малих і середніх підприємств (обсяг виробництва у розрахунку на одного працівника підприємства) за галузями економіки;
- розподіл прибутку малих і середніх підприємств у розрахунку на одного працівника підприємства за галузями економіки.

4. Показники динаміки малих підприємств в цілому і за окремими галузями:

- розподіл темпів зростання числа малих підприємств, чисельності працівників і обсягу виробництва продукції (робіт, послуг) малих підприємств за галузями економіки.

Статистичні методи передбачають широке застосування абсолютних і

відносних показників. Для дослідження необхідно використовувати декілька типів відносних показників з метою вивчення особливостей функціонування і розвитку малого бізнесу:

- показники інтенсивності розвитку – характеризують поширення суб'єктів малого бізнесу у просторі (забезпеченість малих підприємств у розрахунку на 10 тис. осіб населення);
- показники співвідношень – засвідчують про співвідношення одних показників до інших.

Методика оцінки малого бізнесу містить такі складові:

1. Визначення змін у малому бізнесі. Для цього використовуються показники:

- Абсолютний приріст (спад);
- Темпи зростання (спаду);
- Коефіцієнт дивергенції:

$$K_d = \frac{\Delta B_{p-n}}{\Delta B_{\text{країна}}}; \quad (10.1)$$

де ΔB_{p-n} – приріст виробництва продукції (робіт, послуг) малих підприємств на одного жителя регіону за певний період;

$\Delta B_{\text{країна}}$ – приріст виробництва продукції (робіт, послуг) малих підприємств на одного жителя країни за цей період.

Інтегральна оцінка розвитку малих підприємств базується на розрахунку індексу рівня розвитку малих підприємств, який враховує добуток часткових індексів:

$$I_{pМП} = I_{кМП} \times I_{зМП} \times I_{пМП}; \quad (10.2)$$

де $I_{pМП}$ – інтегральний індекс розвитку малих підприємств;

$I_{кМП}$ – індекс кількості малих підприємств (частка малих підприємств у загальній кількості підприємств);

$I_{зМП}$ – індекс зайнятості на малих підприємствах (частка зайнятих на МП до загальної кількості зайнятих);

$I_{nМП}$ – індекс обсягів виробленої продукції малими підприємствами (частка малих підприємств у загальному обсязі виробленої продукції (робіт, послуг) в країні).

Визначення реальних масштабів діяльності малих підприємств країни. Враховується тіньовий компонент діяльності за умов зіставлення фактичних основних показників діяльності малих підприємств ("реальних" або "потенційних"). "Реальний" дохід від реалізації продукції (робіт, послуг) D_p , обчислюють за формулою:

$$D_{pi} = K_{pi} \times D_i; \quad (10.3)$$

де K_{pi} – коефіцієнт "реальної" діяльності i -го року;

D_i – фактичний дохід від реалізації продукції (робіт, послуг) i -го року.

Коефіцієнт "реальної діяльності" обчислюють як співвідношення кількості зареєстрованих суб'єктів, зокрема малі підприємства, до кількості діючих:

$$K_{pi} = \frac{З_{МПi}}{Д_{МПi}}; \quad (10.4)$$

де $З_{МПi}$ – кількість зареєстрованих малих підприємств i -го року;

$Д_{МПi}$ – кількість діючих малих підприємств i -го року (які звітували).

Для оцінка впливу чинників становлення і розвитку малого бізнесу використовують метод побудови множинної регресії. Загальне призначення моделі множинної регресії полягає в аналізі зв'язку між декількома незалежними змінними (пояснювальними факторами) та залежною змінною (результативним фактором). Для типізація одиниць адміністративно-територіального поділу країни за рівнем та характером розвитку малого бізнесу використовуються методи факторного та таксономічного аналізу.

Потенційні кредитори та інвестори для гарантування прибутку на вкладений капітал детально аналізують фінансово-господарську діяльність позичальника. Одним із важливих етапів є дослідження фінансової стійкості та стабільності підприємства-позичальника, що включає такі показники:

- коефіцієнт фінансової незалежності
($K_{ф.н.}$);
- коефіцієнт фінансової стабільності
($K_{ф.с.}$);
- коефіцієнт фінансової залежності
($K_{ф.з.}$);
- коефіцієнт фінансового ризику($K_{ф.р.}$);
- коефіцієнт забезпечення власними коштами($K_{з.р.}$).

Ринкову стійкість підприємства характеризує коефіцієнт фінансової незалежності ($K_{ф.н.}$):

$$K_{ф.н.} = \frac{\text{Власний капітал}}{\text{Вартість майна підприємства}}; \quad (10.5)$$

Коефіцієнт незалежності бажано підтримувати на достатньо високому рівні, що дасть можливість забезпечити стабільну структуру капіталу. Саме такій структурі віддають перевагу інвестори та кредитори.

На практиці питання оптимального співвідношення власного і позикового капіталу не вирішується однозначно. Кожне підприємство визначає його у взаємозв'язку зі структурою основного й оборотного капіталу, швидкістю оборотності обігових коштів та інших показників.

Коефіцієнт фінансової стабільності ($K_{ф.с.}$):

$$K_{ф.с.} = \frac{\text{Власний капітал}}{\text{Позиковий капітал}}; \quad (10.6)$$

Значення цього коефіцієнта не повинно перевищувати 1.

Коефіцієнт фінансової залежності ($K_{ф.з.}$):

$$K_{ф.з.} = \frac{\text{Позиковий капітал}}{\text{Загальний капітал}}; \quad (10.7)$$

Він характеризує залежність діяльності підприємства від позикового капіталу. Оптимальним значенням цього коефіцієнта вважається 0,4.

Коефіцієнт фінансового ризику ($K_{ф.р.}$):

$$K_{ф.р.} = \frac{\text{Позиковий капітал}}{\text{Власний капітал}}; \quad (10.8)$$

Це – найвагоміший показник, який свідчить про фінансову

незалежність (автономність) підприємства від залучення позикових коштів. Він показує, скільки позикових коштів залучає підприємство на 1 гривню власного капіталу.

Оптимальне значення $K_{ф.р.}$ – 0,3 – 0,5. Критичним значенням цього коефіцієнта вважається одиниця. Однак, економісти вважають, що за високих показників оборотності обігових коштів критичне значення коефіцієнта може перевищувати одиницю без істотних наслідків для ринкової стійкості.

Малі і середні підприємства майже всіх галузей є досить залежними від позикових коштів.

Для вирішення питання щодо забезпечення підприємства власними коштами розраховується коефіцієнт забезпечення власними коштами ($K_{з.к.}$):

$$K_{з.к.} = \frac{\text{Власний капітал} - \text{Необоротні активи}}{\text{Оборотний капітал}}; \quad (10.9)$$

Значення цього коефіцієнта повинно бути більше ніж 0,1.

На останок хотілося зробити висновок. Формування ринкової системи господарювання в Україні пов'язане із зростанням підприємницької активності в усіх сферах економіки. Один із перспективних напрямів створення конкурентно-ринкового середовища – розвиток малого та середнього бізнесу. Світовий досвід і практика господарювання показують, що найважливішим компонентом ринкової економіки є існування, взаємодія та оптимальне співвідношення великих, середніх та малих підприємств, а статистична оцінка ефективності бізнес-процесів у середніх та малих підприємствах є дуже важливою для з'ясування того, чи правильно функціонують ці підприємства.

Контрольні питання

1. Поясніть суть понять «ефективність» та реінжиніринг бізнес-процесів».
2. Завдання статистичного оцінювання ефективності бізнес-процесів.
3. Алгоритм інформаційного забезпечення бізнес-аналізу, що циркулює у системі управління.
4. Методика оцінки ефективності малого бізнесу.

ЛІТЕРАТУРА

1. Бізнес-статистика: навч. посібник. /С.О. Матковський, О.С. Гринькевич, М.Л. Вдовин, О.М. Вільчинська, О.Р. Марець, О.З. Сорочак. К.: Алерта, 2016. 280 с.
2. Статистична інформація. Офіційний сайт державної служби статистики України. URL: <http://www.ukrstat.gov.ua>
3. Additive and multiplicative models. URL: <http://www-ist.massey.ac.nz/dstirlin/CAST/CAST/Hmultiplicative/multiplicative1.html>

**3. ПЛАНИ ПРАКТИЧНИХ
(СЕМІНАРСЬКИХ) ЗАНЯТЬ,
САМОСТІЙНОЇ РОБОТИ
(у т.ч. ІНДИВІДУАЛЬНІ ЗАВДАННЯ)**

ПЛАН ПРАКТИЧНИХ ЗАНЯТЬ

Назва теми практичного заняття
<p>СЗ 1. Кількісне оцінювання стану розвитку бізнесу в країні</p> <p>1.1. Проведення семінарського заняття з теми 1 з відповідями на 10 контрольних запитань</p> <p>1.2. Обговорення проблемних запитань з теми</p> <p>1.3. Тестування 1-го та 2-го рівня складності</p>
<p>СЗ 2. Основні етапи статистичного аналізу бізнесу</p> <p>2.1. Перевірка рівня знань студентів з наступних питань:</p> <p>2.1.1. Планування та організація збору даних</p> <p>2.1.2. Попереднє вивчення даних</p> <p>2.1.3. Оцінка невідомих параметрів</p> <p>2.1.4. Перевірка статистичної гіпотези</p>
<p>ПЗ 3. Класифікації та структури даних</p> <p>3.1. Виконання індивідуальних завдань з теми 2 та обговорення їх результатів:*</p> <p>Назвіть два різних набори двовимірних та багатовимірних даних, які безпосередньо чи опосередковано пов'язані з Вашими обов'язками. У кожному випадку визначте характер залежності між двома факторами і встановіть, чи можна і чи мати можливість передбачати один фактор на підставі іншого.</p>
<p>ПЗ 4. Способи перетворення даних</p> <p>4.1. Практичне завдання з цензурування та перетворення даних: Виберіть деяку фірму та опишіть (загалом) базу даних цієї фірми. Всередині цієї бази даних визначте три набори різних видів даних. Для кожного з цих трьох наборів вкажіть, що є елементарною одиницею, та визначте, що можна дізнатися у результаті відповідного аналізу.</p> <p>4.2. Тестування з теми 1-го та 2-го рівня складності</p>
<p>ПЗ 5. Статистичні гіпотези та критерії їх перевірки для різних даних:</p> <p>5.1. Виконання практичних завдань (5 задач різного рівня складності)</p> <p>5.2. Перевірка статистичних гіпотез на конкретному прикладі: Виберіть два рекламних ролика, один значимий, другий нет. Какова Ваша інтерпретація ефективності реклами? Что бы Вы порекомендовали в этой ситуации, используя методологию проверки статистических гипотез?</p>
ПЗ 6. Дисперсійний аналіз
ПЗ 7. Непараметричні методи в бізнесі. Аналіз хі-квадрат
ПЗ 8. Статистичні методи аналізу структури сукупності
ПЗ 9. Статистична оцінка ефективності бізнесу

*Приклади виконання практичних завдань

**Практичні заняття з тем 6-10 проводяться в аналогічному ключі

ПРИКЛАДИ ВИКОНАННЯ ПРАКТИЧНИХ ЗАВДАНЬ

ТЕМА 2

1. Назвіть два різних набори двовимірних та багатовимірних даних, які безпосередньо чи опосередковано пов'язані з Вашими обов'язками. У кожному випадку визначте характер залежності між двома факторами і встановіть, чи можна і чи мати можливість передбачати один фактор на підставі іншого.

Приклад двовимірних даних: Будучи працівником банку або іншої фінансової установи, на Українській фондовій біржі визначається ціна однієї акції фірми (перша змінна) і дата (друга змінна), зареєстровані щодня протягом останніх місяців. Зв'язок між ціною та часом дозволяє побачити тенденції у зміні вартості інвестицій. Проте, важко сказати, чи можна на підставі таких даних передбачити майбутню вартість інвестицій (це, зокрема, залежить від того, чи є зміна вартості непередбачуваною «випадковою помилкою», чи існує певна закономірність).

Приклад багатовимірних даних: Будучи спеціалістом з нерухомості, працівник стикається з необхідністю визначити ринкову вартість будинку (квартири). Так, для кожного з будинків у районі ціна цього будинку, квартири (виділена змінна) та ряд змінних, від яких залежить вартість нерухомості (кількість будинків, квартир подібного типу, площа будинку, кількість кімнат, рік будівництва будинку, наявність або відсутність ландшафтного дизайну та тощо). Аналіз показав би, що як оцінюється нерухомість у певному районі. Такий результат можна було б використовувати для визначення реальної ринкової вартості будинку в цьому районі або при будівництві, щоб визначити яка комбінація характеристик нового будинку підвищує його ціну.

Висновок: наведені приклади довели, що завжди можна виявити залежність між двома чи більше чинниками. Якщо залежність між чинниками можна визначити, то передбачити, спрогнозувати має сенс, але про точності, прогнозу, його достовірності не ведеться, оскільки можуть «втрутитися» особливості ринкової кон'єктури, особливості політичних, економічних подій.

2. Виберіть деяку фірму та опишіть (загалом) базу даних цієї фірми. В середині цієї бази даних визначте три набори різних видів даних. Для кожного з цих трьох наборів вкажіть, що є елементарною одиницею, та визначте, що можна дізнатися у результаті відповідного аналізу.

Виберемо «Печерськ Плаза» - фірма-забудовник, який пропонує квартири на певних умовах (див. табл. 1):

Таблиця 1

Продаж квартир у «Печерськ Плаза» в Києві в 2022 році

Кімнатність	Площа, м ²	Ціна за м ²	Вартість квартири, млн. грн.
1 к.	52	38500	2 002 000
2 к.	75	38500	2 887 500
3 к.	92	38500	3 542 000

Слід врахувати, що виділеною змінною буде вартість квартири, а відповідно ряд змінних – кількість кімнат, площа, ціна за метр квадратний, а також розміщення будинку. Так, як будинок розташований у центрі Києва, то ціна істотно відрізнятиметься від будинків, які розташовані в інших частинах міста.

Інші фірми-забудовники, які можна переглянути на сайті «ЛУН.ЮА», включають інші фактори: клас (бізнес/економ) черга будівництва, поверховість, планування, вид з вікна, наявність ландшафтного парку, технологія будівництва, наявність паркінгу, близькість до метро.

Необхідно зазначити, що одиницею проведеного статистичного спостереження є фірма-збудувник "Печерськ Плаза".

Висновок: в результаті наведеного аналізу, можна дізнатися ринкову вартість квартири, враховуючи сукупність різних факторів. Так, при придбанні квартири покупцеві буде легше орієнтуватися в ціні на той чи інший житловий комплекс, а саме: чи не завищує ціну якийсь забудувник?

ТЕМА 3

1. Знайдіть в Internet і напишіть результати спостереження щонайменше п'яти різних випадкових величин, таких, як біржові котирування, відсоткові ставки, обсяги продажів корпорацій, або будь-яких інших величин, пов'язаних з комерційною діяльністю та цікавих для вас.

ПАТ «Укрнафта» – найбільша нафтовидобувна компанія України. Частка суспільства у загальному видобутку нафти з газовим конденсатом в Україні у 2019 році становить 67,9%, частка у загальному видобутку газу – 7,6%.

Станом на 31.12.2022 р. ПАТ «Укрнафта» володіє 82 спеціальними дозволами на видобуток (промислову розробку родовищ) вуглеводнів.

На балансі бурових управлінь ВАТ "Укрнафта" знаходиться 58 бурових установок.

ПАТ «Укрнафта» володіє однією з найбільших в Україні мереж заправних станцій, є у більшості регіонів країни. Станом на 31.12.2022 р. ПАТ «Укрнафта» експлуатувала 537 АЗС.

Центральний офіс "Укрнафти" розташований у Києві. 50% + 1 акція компанії належить НАК "Нафтогаз України".

Акції "Укрнафти" котируються на "Українській біржі" (Київ). Вони включені до індексу УБ, що є головним показником активності ринку акцій України.

Є такі показники ПАТ «Укрнафта» станом на 20.09.2022:

1. Ціна покупки за акцію складає 104,10.
2. Ціна продажу за акцію складає 102,00.
3. Ціна сьогодні (20.09.2022) за акцію має значення 104,10.
4. Обсяг торгів набуває значення 1119930.00 грн.
5. Процентна ставка протягом дня становить +3,89%.

ТЕМА 4

Виберіть два рекламні ролики, один значимий, інший немає. Якою є Ваша інтерпретація ефективності реклами? Що б Ви порекомендували у цій ситуації, використовую методологію перевірки статистичних гіпотез?

Для перевірки статистичних гіпотез про ефективність реклами слід використовувати двосторонню перевірку, так як альтернативна гіпотеза розглядає можливість того, що значення середньої генеральної сукупності може бути розташоване по обидва боки (як і більше, так і менше - мають на увазі, що вихід нових рекламних роликів може спричинити до збільшення обсягу продажу, у разі реклама є ефективною, або ж до зменшення) від заданого значення.

Слід зазначити, що може не цікавити перевірка того, чи відрізняється середнє генеральної сукупності від заданого значення. Так, може цікавити більш конкретне питання: чи середнє генеральної сукупності є більше (в одних випадках) або менше (в інших випадках) заданого значення.

Щоб використовувати односторонній тест, слід бути впевненим, що незалежно від того, як поведуться дані, односторонній тест продовжуватиме застосовуватися на тій же стороні («більше, ніж» або «менше ніж»). Якщо внаслідок зміни характеру даних виникає бажання використовувати тест з іншого боку замість раніше планованої, необхідно перейти до використання двостороннього тесту. Якщо є сумніви, слід також використовувати двосторонній тест.

Зокрема використання одностороннього тесту залишає місце для критики. Оскільки рішення, що вас цікавить, може бути суб'єктивним, то бажання зосередитися на тому, що цікавить безпосередньо вас, може не відповідати думці тих, кого ви хочете переконати. Якщо необхідно переконати людей, яких можуть мати іншу точку зору, слід використовувати двосторонній тест і робити на його основі односторонній висновок. З іншого боку, якщо необхідно переконати «дружньо» налаштованих людей з інтересами, аналогічними до ваших, то у разі виконання зазначеної раніше умови можна використовувати переваги одностороннього тесту.

Висновок: З вищевказаної інформації випливає, для перевірки статистичних гіпотез про ефективність реклами необхідно використовувати двосторонню перевірку, оскільки альтернативна гіпотеза розглядає можливість того, що значення середньої генеральної сукупності може бути розташоване по обидва боки; тобто враховуються всі варіанти результату. Проте слід зазначити, якби рекламний ролик було реалізовано, «протестовано», поширено серед цільової аудиторії, то одностороння перевірка була б доречнішою. (Як зазначалося раніше, якщо необхідно переконати «дружньо» налаштованих людей з інтересами, аналогічними вашим, то разі виконання зазначеного раніше умови можна використовувати переваги одностороннього тесту). Наприкінці, можна назвати, що вибір варіантів перевірок статистичних гіпотез (односторонньої чи двосторонньої) найчастіше залежить від особливостей вихідної ситуації, які потрібно знати.

ТЕМА 5

«Дисперсійний аналіз у дослідженні бізнес-процесів»

Дисперсійний аналіз (скорочено ANOVA) задає загальну схему перевірки статистичних гіпотез, що ґрунтується на ретельному вивченні різних джерел варіації у складній ситуації. Для перевірки кожної з гіпотез у дисперсійному аналізі використовують F-тест, заснований на F-статистиці, яка є відношенням двох дисперсій. Чисельник, у цьому відношенні є варіацію, обумовлену конкретним цікавим для нас ефектом, що ми і перевіряємо, а знаменник – варіацію, обумовлену випадковістю. Якщо це відношення більше табличного F-значення, ефект значущий. Однофакторний дисперсійний аналіз, зокрема, використовують для перевірки значущості відмінностей між собою середніх значень, що характеризують різні ситуації.

Не забувайте про первинний розвідувальний аналіз даних. Блокова діаграма допоможе порівняти відразу кілька розподілів, що дозволить вам побачити структуру даних, визначити проблеми (якщо вони є) та перевірити початкові умови, необхідні для дисперсійного аналізу, а саме нормальність розподілів та рівність варіації.

Дані для однофакторного дисперсійного аналізу є k незалежних одновимірних вибірок, елементи яких виміряні в однакових одиницях. Однофакторний дисперсійний аналіз порівнює два наступні джерела варіації:

- міжгрупова варіація (між вибірками);
- Внутрігрупова варіація (всередині кожної вибірки).

Для коректного застосування однофакторного дисперсійного аналізу потрібне виконання двох умов.

1. Набір даних складається з генеральних сукупностей.

2. Кожна генеральна сукупність підпорядковується нормальному розподілу, і стандартні відхилення переважають у всіх генеральних сукупностях однакові, тобто:

$$\sigma_1 = \sigma_2 = \dots = \sigma_i.$$

Нульова гіпотеза стверджує, що між генеральними сукупностями немає відмінностей, а альтернативна гіпотеза стверджує, що деякі відмінності мають місце:

- $H_0 : \mu_1 = \mu_2 = \dots = \mu_i$ (усі середні рівні між собою);

- $H_1 : \mu_1 \neq \mu_i$ принаймні для однієї пари генеральних сукупностей (не всі середні рівні між собою).

Загальне (головне) середнє є середнє всіх значень з усіх вибірок:

$$\bar{X} = \frac{n_1 \bar{X}_1 + n_2 \bar{X}_2 + \dots + n_i \bar{X}_i}{n} = \frac{\sum_{i=1}^i n_i \bar{X}_i}{n}.$$

Міжгрупова (міжвибіркова) варіація вимірює відмінність вибірових середніх, а внутрішньогрупова (внутрішньовибіркова) варіація вимірює варіабельність кожної з вибірок:

$$\text{міжгрупова варіація} = \frac{n_1(\bar{X}_1 - \bar{X})^2 + n_2(\bar{X}_2 - \bar{X})^2 + \dots + n_k(\bar{X}_k - \bar{X})^2}{k-1} = \frac{\sum_{i=1}^k n_i(\bar{X}_i - \bar{X})^2}{k-1}.$$

Число ступенів свободи: $= k - 1$.

$$\text{Внутрішньогрупова варіація} = \frac{(n_1-1)(S_1)^2 + (n_2-1)(S_2)^2 + \dots + (n_k-1)(S_k)^2}{n-k} = \frac{\sum_{i=1}^k (n_i-1)(S_i)^2}{n-k}.$$

Число ступенів свободи: $= n - k$.

F-статистика є відношенням цих двох значень варіації і показує міру відмінності вибірових середніх (числитель) по відношенню до загального рівня варіації вибірок (знаменник).

$$F = (\text{Міжгрупова варіація}) / (\text{Внутрішньогрупова варіація}).$$

Число ступенів свободи $= k - 1$ (для чисельника)

та $n - k$ (для знаменника).

F-таблиця містить критичні значення для розподілу F-статистики. Таким чином, якщо справедлива нульова гіпотеза, значення F-статистики перевищує значення з F-таблиці в контрольованому відсотку випадків (наприклад, 5%). F-тест виконують, порівнюючи значення F-статистики (розраховане з даних) з критичним значенням F-таблиці.

F-тест визначає лише наявність чи відсутність відмінностей. Якщо F-тест фіксує значущість відмінностей, то тест найменшої значущої відмінності використовується для порівняння кожної пари вибірок, щоб визначити, чи вони різняться між собою. Цей тест заснований на середній різниці між двома порівнюваними групами, стандартній помилці цієї різниці та кількості ступенів свободи ($n-k$):

$$\text{Стандартна помилка} = \sqrt{(\text{внутрішньогрупова варіація}) \left(\frac{1}{n_i} + \frac{1}{n_j} \right)},$$

де n_i і n_j - розміри двох порівнюваних вибірок.

Існує багато складніших методів дисперсійного аналізу, включаючи двофакторний план та плани вищого порядку. Для того, щоб будь-який з таких методів дисперсійного аналізу можна було застосувати, дані повинні складатися з набору вибірок, що представляють результати вимірювання в тих самих одиницях, як і при звичайному дисперсійному однофакторному аналізі. Пам'ятайте, що ANOVA-таблиця не повністю розкриває всю картину; завжди необхідно дивитися на значення середніх, щоб зрозуміти, що справді відбувається. де n_i і n_j - розміри двох порівнюваних вибірок.

Завдання самостійної роботи для студентів з Бізнес-статистики

Види та зміст самостійної роботи:

1. Перевірка теоретичних знань: класифікувати дані, що характеризують конкретний напрямок економіки.
2. Перевірка виконання індивідуальних практичних завдань: формулювати нульову та альтернативну гіпотези відповідно до мети статистичного аналізу.
3. Перевірка контрольних завдань: застосовувати методи одновимірного та багатовимірного дисперсійного аналізу.
4. Написання есе: використовувати статистичні методи для оцінювання ефективності бізнесу.
5. Використання Internetресурсів: застосовувати пакети прикладних програм для розв'язання задач кількісного оцінювання економічних процесів.

Контрольна робота з «Бізнес-статистики» «Статистичний аналіз бізнес ситуацій»

Управлінська діяльність в будь-якій предметній області (виробництво, торгівля, фінанси і т.п.) постійно пов'язана з прийняттям рішень по виникаючих проблемах. Для розробки і вибору єдино правильного рішення необхідна повна і достовірна статистична інформація, що характеризує проблемну ситуацію. Інакше кажучи, повинна бути повна визначеність щодо всіх факторів, які прямо або побічно впливають на рішення. Однак на практиці дію багатьох чинників можна передбачити з повною визначеністю: рівень попиту на товар, погодні умови, курс акцій, ціни на нафту, політична обстановка, стихійні лиха і т.п. Тому і вибір рішень передбачає:

1. Опис проблемної ситуації;
 2. Виявлення можливих варіантів дій (альтернативних рішень або альтернатив);
 3. Опис випадкових факторів, тобто передбачуваних майбутніх подій, здатних вплинути на рішення, що приймаються;
 4. Оцінка результатів (наслідків) прийнятих рішень.
- Виконати в формі Есе.

**4. ПИТАННЯ, ЗАДАЧІ, ЗАВДАННЯ АБО КЕЙСИ
ДЛЯ ПОТОЧНОГО КОНТРОЛЮ ЗНАНЬ І ВМІНЬ
ЗДОБУВАЧІВ ВИЩОЇ ОСВІТИ, ДЛЯ
КОНТРОЛЬНИХ РОБІТ, ПЕРЕДБАЧЕНИХ
НАВЧАЛЬНИМ ПЛАНОМ,
ПІСЛЯТЕСТАЦІЙНОГО МОНІТОРИНГУ
НАБУТИХ ЗНАНЬ І ВМІНЬ З НАВЧАЛЬНОЇ
ДИСЦИПЛІНИ**

Практичний тренінг розділу 1

Тестові завдання щодо самоконтролю

Структури даних. Класифікація, розподіл та перетворення даних.

1. Багатовимірні дані містять інформацію:
А) про дві ознаки для кожного з об'єктів;
В) про три або більше ознак для кожного об'єкта;
С) тільки про одну ознаку для кожного об'єкта;
Д) усі відповіді вірні.
2. Яких типів бувають якісні дані?:
А) порядкові і комбіновані;
В) номінальні і комбіновані;
С) номінальні і порядкові;
Д) порядкові і рангові.
3. Яких типів бувають кількісні дані?:
А) дискретні та безперервні;
В) дискретні та порядкові;
С) рангові і безперервні;
Д) безперервні і комбіновані.
4. Який вид шкали використовується для реєстрації самого нижчого рівня вимірювань?:
А) шкала інтервалів ;
В) номінальна шкала ;
С) логарифмічна шкала;
Д) Шкала пропорцій.
5. Прикладом якої шкали є міри довжини (м, см. і тд.):
А) шкала відносин;
В) рангова шкала ;
С) порядкова шкала;
Д) шкала інтервалів.
6. Удосконалення управління зумовлює використання менеджерами результатів статистичного моніторингу ділової активності підприємств для розробки подальших можливих корпоративних стратегій розвитку:
А) реструктуризації;
В) диверсифікації;
С) вертикальної інтеграції;
Д) портфельного аналізу.
7. В основу портфельного аналізу покладено такі оцінні критерії:
А) теперішня вартість очікуваних доходів від володіння цінними паперами (проценти, дивіденди);
В) рівень ризикованості вкладень;
С) аналіз портфеля продуктів (товарів і послуг) підприємства;
Д) усі відповіді вірні.

Випадкові величини і закони розподілу як основа бізнес-статистики

8. Імовірність того, що випадкова величина X (дискретна або безперервна) приймає значення, менше x , називається:
- А) функцією розподілу;
 - В) величиною розподілу;
 - С) законом розподілу;
 - Д) гіпотезою розподілу.
9. Відносною частотою випадкової події A називається:
- А) відношення числа n проведених випробувань до загальної кількості nA появ цієї події;
 - В) відношення числа nA появ цієї події до загального числа n проведених випробувань;
 - С) відношення числа nA появ цієї події до середнього числа n проведених випробувань;
 - Д) усі відповіді вірні.
10. Який вид розподілу визначається двома параметрами, n і p :
- А) геометричний розподіл;
 - В) рівномірний розподіл;
 - С) біноміальний розподіл;
 - Д) показовий розподіл.
11. Дві випадкові величини X і Y називаються незалежними, якщо:
- А) якщо незалежні всі пов'язані з ними події;
 - В) якщо незалежним є хоча б одна пов'язана з ним подія;
 - С) якщо жодна пов'язана з ним подія не є незалежною.
 - Д) усі відповіді невірні.
12. Коваріація двох випадкових величин X і Y характеризує:
- А) ступінь їх залежності навколо точки (m_x, m_y) ;
 - В) міру їх розсіювання навколо точки (m_x, m_y) ;
 - С) ступінь їх залежності і міру розсіювання навколо точки (m_x, m_y) ;
 - Д) не одна з відповідей не є вірною
13. Безрозмірна величина, що характеризує тільки залежність випадкових величин X і Y , а не розкид називається:
- А) коефіцієнт детермінації;
 - В) коефіцієнт Спірмена;
 - С) коефіцієнт Чупрова;
 - Д) коефіцієнт кореляції.
14. Коефіцієнт кореляції характеризує:
- А) ступінь лінійної залежності випадкових величин X і Y ;
 - В) ступінь логарифмічної залежності випадкових величин X і Y ;

- С) ступінь експоненційної залежності випадкових величин X і Y ;
- Д) ступінь інтегральної залежності випадкових величин X і Y .

Статистичні гіпотези та методика їх перевірки

15. Статистична гіпотеза – це:

- А) певне твердження, яке потрібно підтвердити або відкинути на основі наявних статистичних даних;
- В) інформація, яку отримують на підставі даних генеральної сукупності;
- С) ступінь ризику, що відповідає тому чи іншому висновку;
- Д) ймовірність того, що прийняті рішення будуть помилковими.

16. За прикладним змістом можна виділити такі групи статистичних гіпотез:

- А) щодо типу розподілу досліджуваної випадкової величини;
- В) щодо числових характеристик досліджуваної генеральної сукупності;
- С) щодо однорідності двох і більше вибірок або характеристик досліджуваних сукупностей;
- Д) щодо ступеню та типу залежності між досліджуваними ознаками.

17. Гіпотезу, яка заперечує нульову, називають:

- А) параметричною;
- В) нульовою;
- С) конкуруючою;
- Д) альтернативною.

18. Прийняти неправильне рішення під час перевірки статистичної гіпотези можна у:

- А) одному випадку;
- В) двох випадках;
- С) трьох випадках;
- Д) трьох і більше випадках.

19. Точки, що відділяють критичну область від області прийняття гіпотез називають:

- А) критичними;
- В) односторонніми;
- С) двохсторонніми;
- Д) правосторонніми.

20. До завдань перевірки нульової гіпотези не входить:

- А) вибір рівня істотності;
- В) визначення критичної області;
- С) обчислення статистичного критерію;
- Д) оцінка похибки вибірки.

21. Якщо перевіряють гіпотези щодо вибірових середніх за невідомого значення дисперсії генеральної сукупності, застосовують критерій, що відповідає:

- А) розподілу Фішера;
- В) розподілу Стюдента;

- С) нормальному розподілу;
- Д) варіаційному розпод

Дисперсійний аналіз у дослідженні бізнес-процесів

22. З англійської мови ANOVA перекладається як:

- А) кореляційний аналіз;
- В) регресійний аналіз;
- С) математичний аналіз;
- Д) дисперсійний аналіз.

23. Виберіть вірну математичну властивість, характерну для дисперсії:

- А) якщо всі значення ознаки змінити в k раз, то дисперсія зміниться в k 2 разів;
- В) якщо всі значення ознаки збільшити в k раз, то дисперсія зменшиться в k раз;
- С) якщо всі значення ознаки зменшити в k раз, то дисперсія збільшиться в k раз;
- Д) жодне з вище згаданих властивостей не характерна для дисперсії.

24. Якщо усі значення ознаки x_j зменшити або збільшити на певну величину- то: А) при збільшенні ознаки дисперсія збільшиться, при зменшенні - зменшиться;

- В) дисперсія не зміниться;
- С) при зменшенні ознаки - дисперсія збільшиться;
- Д) при збільшенні ознаки - дисперсія зменшиться.

25. У структурованій сукупності, розділеної на групи за ознакою x , дисперсія ознаки u може бути розкладена на:

- А) міжгрупові та зовнішньогрупові;
- В) зовнішньогрупову та внутрішньогрупову;
- С) міжгрупову та внутрішню групу;
- Д) міжгрупову та середню групову.

26. Щодо усіх груп у цілому обчислюється:

- А) середня з між групових;
- В) середня з внутрішньогрупових;
- С) середня з зовнішньогрупових;
- Д) середня арифметична.

27. Взаємозв'язок між трьома дисперсіями має такий вигляд:

- А) загальногруппова дисперсія = міжгруппова – внутрішньогруппова;
- В) загальногруппова дисперсія = міжгруппова + внутрішньогруппова;
- С) внутрішньогруппова дисперсія = міжгруппова * загальну дисперсію;
- Д) міжгруппова дисперсія = загальна дисперсія : середню з внутрішньогруппових.

28. Ставлення групової дисперсії до загальної називається:

- А) варіаційним відношенням η^2 ;
- В) дисперсійним відношенням η^2 ;
- Д) кореляційним відношенням η^2 .
- С) емпіричним відношенням η^2 ;

Практичний тренінг розделу 2

Тести щодо самоконтролю

Непараметричні методи у бізнес-статистиці

- Непараметрична статистика - це:
А) самостійна галузь математичної статистики;
В) класичний метод аналізу даних;
С) статистична перевірка гіпотез;
Д) усі відповіді вірні.
- Статистичний метод, який доволі простий та якому притаманна стійкість висновків:
А) коефіцієнт кореляції рангів Спірмена;
В) квадратична спряженість Пірсона χ^2 ;
С) коефіцієнт спряженості Чупрова;
Д) усі відповіді невірні.
- Ранги оцінки восьми підприємств іноземними інвесторами та зовнішніми незалежними експертами у сфері інвестування розподілені так:

Наданий ранг	
Іноземними інвесторами	Зовнішніми незалежними експертами
8	8
7	4
6	1
4	7
1	3
2	6
3	2
5	5

Коефіцієнт рангової кореляції Спірмена дорівнює:

А) 0,24; В) 0,87 С) 0,35; Д) 0,51

- Коефіцієнти асоціації та контингенції використовують для оцінки зв'язку у бізнес-статистиці між:
А) атрибутивними ознаками;
В) кількісними ознаками;
С) альтернативними ознаками;
Д) ознаками порядкової шкали;

5. Розподіл молодих працівників за ступенем задоволеності умовами праці та професійною мобільністю характеризується такими даними:

Ступінь задоволеності умовами праці	Мають намір змінити професію	Не мають наміру змінювати професію	Разом
Задоволені	10	30	40
Незадоволені	20	5	25
Разом	30	35	65

Коефіцієнт контингенції становить:

А) -0,54; В) -0,84; В) 0,25; Д) 0,47.

6. Коефіцієнт спряженості Чупрова розраховується за формулою:

$$А) K = \frac{\sum_1^n xy - n\bar{x}\bar{y}}{n\sqrt{\sigma_x^2\sigma_y^2}}; В) K = \sqrt{\frac{\chi^2}{n\sqrt{(k_1-1)*(k_2-1)}}}; С) K = \frac{\sum C - \sum H}{\sum C + \sum H}; Д) K = \frac{2S}{n(n-1)}.$$

7. Коефіцієнт спряженості Крамера розраховується за формулою:

$$А) K = \frac{2S}{n(n-1)}; В) K = \frac{\sum C - \sum H}{\sum C + \sum H}; С) K = \frac{\sum_1^n xy - n\bar{x}\bar{y}}{n\sqrt{\sigma_x^2\sigma_y^2}}; Д) K = \sqrt{\frac{\chi^2}{n(m_{\min}-1)}}.$$

Аналіз закономірностей для якісних даних

8. Для обробки масивів якісних даних часто використовують показники:

- А) статистичні;
- В) комбіновані;
- С) евристичні;
- Д) усі відповіді вірні.

9. Щодо аналітичного оцінювання якісних даних, використовують наступні показники:

- А) моду та середній нормований бал;
- В) дисперсію альтернативної ознаки;
- С) коефіцієнти локалізації та концентрації;
- Д) коефіцієнти взаємного спряження та рангової кореляції.

10. Показник, який вимірює різницю між частотами що спостерігаються (експериментальними) f_E і очікуваними (теоретичними) частотами f_T , називається:

- А) «хі-квадрат»;
- В) коефіцієнт Пірсона;
- С) коефіцієнтом Чупрова;
- Д) коефіцієнтом детермінації.

11. Значення коефіцієнтів Чупрова, Крамера та Пірсона у бізнес-статистиці змінюється у межах від:

- А) $[0, 1]$; В) $[-1, 1]$; С) $[-\infty, +\infty]$; Д) $[0, +\infty]$.

12. Якщо виявляється справедливим нерівність $\chi^2_{\text{розр}} > \chi^2_{\text{крит}}$, то із заданою вірогідністю (або рівнем значущості) можна стверджувати, що частоти котрі спостерігаються:

А) незначно відрізняються від опорних показників, отже, наші фактичні результати не мають значущих відмінностей від заданих опорних значень;

В) значимо відрізняються від тих, які очікуються виходячи з відомих нам опорних значень відсотків (частот);

С) зовсім не відрізняються;

Д) усі відповіді вірні.

13. При обчисленні статистики «хі-квадрат» для перевірки гіпотези про незалежність, число ступенів свободи розраховують за такою формулою:

А) число ступенів свободи = (Число категорій для першої змінної -2) * (Число категорій для другої змінної-2);

В) число ступенів свободи = (Число категорій для першої змінної -1) + (Число категорій для другої змінної-1);

С) число ступенів свободи = (Число категорій для першої змінної -1) / (Кількість категорій для другої змінної-1);

Д) число ступенів свободи = (Число категорій для першої змінної -1) * (Число категорій для другої змінної-1).

14. Data-Mining - технологія інтелектуального аналізу даних щодо виявлення прихованих закономірностей у вигляді значущих особливостей кореляцій, тенденцій, шаблонів тощо, виникла та розвинулась на основі:

А) статистики;

В) теорії баз даних;

С) штучного інтелекту;

Д) усі відповіді вірні.

Статистичне оцінювання ефективності бізнес-процесів

15. Покращення бізнес-процесів, тобто підвищення їх ефективності, здійснюється шляхом їх:

А) модернізації;

В) автоматизації;

С) реінжинірингу;

Д) реструктуризації.

16. Родоначальником терміну «реінжиніринг» вважається:

А) К. Чупров;

В) Л. Козерод;

С) Д. Антіпов;

Д) М. Хаммер.

17. Щодо успішної реалізації мети аналізу бізнес-процесів необхідно:

А) визначення економічної ефективності використання ресурсів бізнес-процесів;

В) оцінка виконання планів, прогнозів, управлінських рішень тощо;

С) вивчення впливу об'єктивних і суб'єктивних, зовнішніх і внутрішніх чинників на результати бізнес-процесів підприємства;

Д) розробка й обґрунтування заходів спрямованих на оптимізацію бізнес-процесів підприємства і прийняття ефективних управлінських рішень.

18. Інформаційне забезпечення бізнес - аналізу можна розділити на:

А) нормативно-довідкове;

В) планово - інформаційне;

С) фактографічно- інформаційне;

Д) звітно-облікове тощо.

19. У якості основних складових системи показників, які характеризують бізнес-процеси приладобудівних підприємств, виділяють наступні їх групи:

А) показники задоволеності споживачів результатами діяльності підприємства;

В) показники результативності бізнес-процесів;

С) показники вартості бізнес-процесів;

Д) показники фрагментації.

20. Методика оцінки малого бізнесу містить такі складові:

А) визначення змін у малому бізнесі;

В) інтегральна оцінка розвитку малих підприємств;

С) визначення реальних масштабів діяльності малих підприємств країни;

Д) усі відповіді вірні.

21. В сучасній фаховій літературі, залежно від масштабів проведення реінжинірингу, вчені виділяють такі його базові види:

А) «разові поліпшення»;

В) «кризовий»;

С) «адаптивний»;

Д) «тотальне моделювання».

**5. ЗАВДАННЯ ПІДСУМКОВОГО КОНТРОЛЮ –
СЕМЕСТРОВИЙ ЗАЛІК
(дворівнева шкала оцінювання)**

Харківський національний університет імені В.Н. Каразіна

Факультет ЕкономічнийСпеціальність (напрямок) 051 «Економіка»

Освітня програма «Економічна аналітика та статистика»

Семестр ХФорма навчання деннаРівень вищої освіти (освітньо-кваліфікаційний рівень): другий «магістерський»Навчальна дисципліна: «**Бізнес-статистика**»

ЗАЛКОВЕ ЗАВДАННЯ № 1

1. Теоретична частина:

1.1. Дати повну відповідь на питання: *Поняття ефективності та реінжинірингу бізнес-процесів.* (max 15)

2. Визначити єдину правильну відповідь на кожне з двох тестових завдань: 2
тести (max 10)

2-а. Якщо перевіряють гіпотези щодо вибірових середніх за невідомого значення дисперсії генеральної сукупності, застосовують критерій, що відповідає:

- А) розподілу Фішера;
- В) розподілу Стюдента;
- С) нормальному розподілу;
- Д) варіаційному розподілу.

2-б. Статистичний метод, який доволі простий та якому притаманна стійкість висновків:

- А) коефіцієнт кореляції рангів Спірмена;
- В) квадратична спряженість Пірсона χ^2 ;
- С) коефіцієнт спряженості Чупрова;
- Д) усі відповіді невірні.

3. Практична частина:

3.1. Розв'язати задачу: (max 15)

Вартість міжнародних турів у розрізі окремих категорій туристських підприємств характеризується наступними даними:

Таблиця 1.1

Категорія туристського підприємства	Кількість туристських підприємств	Загальна вартість придбаних турів, тис. ум. од.	Вартість одного міжнародного туру, тис. ум. од.
Турфірми	48	480	3,20
Турагенства	25	280	2,80
Итого	73	760	X

Визначте середню вартість міжнародного туру. Зробіть висновки.

Затверджено на засіданні кафедри статистики, обліку та аудиту
протокол № 1 від “26 ” серпня 2024 р.

Викладач

Олена КУЩЕНКО

Харківський національний університет імені В.Н. Каразіна

Факультет ЕкономічнийСпеціальність (напрямок) 051 «Економіка»

Освітня програма «Економічна аналітика та статистика»

Семестр ХФорма навчання деннаРівень вищої освіти (освітньо-кваліфікаційний рівень): другий «магістерський»Навчальна дисципліна: «**Бізнес-статистика**»

ЗАЛКОВЕ ЗАВДАННЯ № 2

1. Теоретична частина:

1.1. Дати повну відповідь на питання *Загальна характеристика статистичних підходів до управління якістю та їх наочний аналіз.* (max 15)2. Визначити єдину правильну відповідь на кожне з двох тестових завдань: 2
тести (max 10)

2-а. Прикладом якої шкали є міри довжини (м, см. і т.п.):

- А) шкала відносин;
- В) рангова шкала ;
- С) порядкова шкала;
- Д) шкала інтервалів.

2-б. Гіпотезу, яка заперечує нульову, називають:

- А) параметричною;
- В) нульовою;
- С) конкуруючою;
- Д) альтернативною.

3. Практична частина:

3.1. Розв'язати задачу: (max 15)

Динаміка виробництва прокату чорних металів характеризується даними:

Таблиця 2.1

Прокат	Обсяг виробництва, тис.т		Собівартість виробництва, грн.	
	Базовий період	Поточний період	Базовий період	Поточний період
Середньокатаний	40	44	550	500
Тонкокатаний	25	35	380	360

Визначити: абсолютну зміну витрат виробництва в цілому і за рахунок зміни в собівартості. **Зробіть** висновки.Затверджено на засіданні кафедри статистики, обліку та аудиту
протокол № 1 від “26 ” серпня 2024 р.

Викладач

Олена КУЩЕНКО

Харківський національний університет імені В.Н. Каразіна

Факультет ЕкономічнийСпеціальність (напрямок) 051 «Економіка»

Освітня програма «Економічна аналітика та статистика»

Семестр XФорма навчання деннаРівень вищої освіти (освітньо-кваліфікаційний рівень): другий «магістерський»Навчальна дисципліна: «**Бізнес-статистика**»

ЗАЛІКОВЕ ЗАВДАННЯ № 3

1. Теоретична частина:

1.1. Дати повну відповідь на питання: *Виявлення прихованих закономірностей для вирішення бізнес-задач за допомогою Data-Mining.* (max 15)

2. Визначити єдину правильну відповідь на кожне з двох тестових завдань: 2 тести (max 10)

2-а. Удосконалення управління зумовлює використання менеджерами результатів статистичного моніторингу ділової активності підприємств для розробки подальших можливих корпоративних стратегій розвитку:

- А) реструктуризації;
- В) диверсифікації;
- С) вертикальної інтеграції;
- Д) портфельного аналізу.

2-б. Виберіть вірну математичну властивість, характерну для дисперсії:

- А) якщо всі значення ознаки змінити у k раз, то дисперсія зміниться у k^2 разів;
- В) якщо всі значення ознаки збільшити у k раз, то дисперсія зменшиться у k раз;
- С) якщо всі значення ознаки зменшити у k раз, то дисперсія збільшиться у k раз;
- Д) жодне з вище згаданих властивостей не характерно щодо дисперсії.

3. Практична частина:

3.1. Розв'язати задачу: (max 15)

Результати контрольної перевірки удою молока від 10 корів породистого стада приведені нижче:

Таблиця 3.1

Стадо корів	Денний удой молока, кг					
Чистопорідне	10	14	9	11	-	-
Племінне	12	9	10	7	13	9

Визначте: міжгрупову дисперсію денного удою молока. **Зробіть** аналітичні висновки.

Затверджено на засіданні кафедри статистики, обліку та аудиту
протокол № 1 від “26” серпня 2024 р.

Викладач

 Олена КУЩЕНКО

Харківський національний університет імені В.Н. Каразіна

Факультет ЕкономічнийСпеціальність (напрямок) 051 «Економіка»

Освітня програма «Економічна аналітика та статистика»

Семестр ХФорма навчання деннаРівень вищої освіти (освітньо-кваліфікаційний рівень): другий «магістерський»Навчальна дисципліна: «**Бізнес-статистика**»

ЗАЛКОВЕ ЗАВДАННЯ № 4

1. Теоретична частина:

1.1. Дати повну відповідь на питання: *Оцінка зв'язків між альтернативними ознаками у бізнес-статистиці.* (max 15)

2. Визначити єдину правильну відповідь на кожне з двох тестових завдань: 2 тести (max 10)

2-а. Ранги оцінки восьми підприємств іноземними інвесторами та зовнішніми незалежними експертами у сфері інвестування розподілені так:

Таблиця 4.1

Наданий ранг	
Іноземними інвесторами	Зовнішніми незалежними експертами
8	8
7	4
6	1
4	7
1	3
2	6
3	2
5	5

Коефіцієнт рангової кореляції Спірмена дорівнює: А) 0,24; В) 0,87 С) 0,35; Д) 0,51.

2-б. Показник, який вимірює різницю між частотами що спостерігаються (експериментальними) f_e та очікуваними (теоретичними) частотами f_m , називається:

А) «хі-квадрат»; В) коефіцієнт Пірсона; С) коефіцієнтом Чупрова; Д) коефіцієнтом детермінації.

3. Практична частина:

3.1. Розв'язати задачу: (max 15)

Динаміка курсу цін на акції двох емітентів характеризується наступними даними:

Таблиця 4.2

Емітенти	Курс цін на акції		Кількість проданих акцій, тис. шт.	
	Базисний період	Поточний період	Базисний період	Поточний період
А	0,9	1,4	3,0	2,0
В	1,8	2,0	1,5	2,5

Визначте індекс структурних зрушень курсу цін, %.

Зробіть обґрунтовані висновки.

Затверджено на засіданні кафедри статистики, обліку та аудиту

протокол № 1 від "26" серпня 2024 р.

Викладач  Олена КУЩЕНКО

Харківський національний університет імені В.Н. Каразіна
 Факультет Економічний
 Спеціальність (напряом) 051 «Економіка»
 Освітня програма «Економічна аналітика та статистика»
 Семестр X
 Форма навчання денна
 Рівень вищої освіти (освітньо-кваліфікаційний рівень): другий «магістерський»
 Навчальна дисципліна: «**Бізнес-статистика**»

ЗАЛІКОВЕ ЗАВДАННЯ № 5

1. Теоретична частина:

1.1. Дати повну відповідь на питання: *Сутність і передумови використання непараметричних методів у бізнес-статистиці.* (max 15)

2. Визначити єдину правильну відповідь на кожне з двох тестових завдань: 2 тести (max 10)

2-а. Покращення бізнес-процесів, тобто підвищення їх ефективності, здійснюється шляхом їх:

- А) модернізації;
- В) автоматизації;
- С) реінжинірингу;
- Д) реструктуризації.

2-б. За прикладним змістом можна виділити такі групи статистичних гіпотез:

- А) щодо типу розподілу досліджуваної випадкової величини;
- В) щодо числових характеристик досліджуваної генеральної сукупності;
- С) щодо однорідності двох і більше вибірок або характеристик досліджуваних сукупностей;
- Д) щодо ступеню та типу залежності між досліджуваними ознаками.

3. Практична частина:

3.1. Розв'язати задачу: (max 15)

За даними про випуск продукції малими підприємствами області:

Таблиця 5.1

Обсяг виробництва продукції, тис. грн.	Кількість підприємств
До 12	5
12 - 16	7
16 - 20	5
20 - 24	4
24 - 28	2
28 і більше	3

Визначте: 1) середній випуск продукції у вартісному виразі та **обґрунтувати** вибір середньої величини, яку при цьому використано; 2) дисперсію та квадратичний коефіцієнт варіації. **Проаналізувати** одержані показники.

Затверджено на засіданні кафедри статистики, обліку та аудиту
 протокол № 1 від “26 ” серпня 2024 р.

Викладач



Олена КУЩЕНКО

Харківський національний університет імені В.Н. Каразіна

Факультет ЕкономічнийСпеціальність (напрямок) 051 «Економіка»

Освітня програма «Економічна аналітика та статистика»

Семестр ХФорма навчання деннаРівень вищої освіти (освітньо-кваліфікаційний рівень): другий «магістерський»Навчальна дисципліна: «**Бізнес-статистика**»

ЗАЛІКОВЕ ЗАВДАННЯ № 6

1. Теоретична частина:

1.1. Дати повну відповідь на питання: *Застосування пакету прикладних програм для проведення однофакторного та двофакторного аналізу.* (max 15)2. Визначити єдину правильну відповідь на кожне з двох тестових завдань: 2
тести (max 10)

2-а. Родоначалником терміну «реінжиніринг» вважається:

- А) К. Чупров;
- В) Л. Козерод;
- С) Д. Антіпов;
- Д) М. Хаммер.

2-б. Методика оцінки малого бізнесу містить такі складові:

- А) визначення змін у малому бізнесі;
- В) інтегральна оцінка розвитку малих підприємств;
- С) визначення реальних масштабів діяльності малих підприємств країни;
- Д) усі відповіді вірні.

3. Практична частина:

3.1. Розв'язати задачу: (max 15)

На основі наведених даних про групування регіонів України за чисельністю населення й обсягом прямих іноземних інвестицій у малий бізнес:

Таблиця 6.1

Групи регіонів за чисельністю населення, млн. осіб	Частка регіонів у загальній їх кількості, %	Частка прямих іноземних інвестицій в загальному їх обсязі, %
До 1	7,4	0,9
1 – 2	59,3	31,0
2 – 3	25,9	55,2
3 – 4	3,7	5,8
Понад 4	3,7	7,1
Разом	100,0	100,0

Розрахувати коефіцієнти локалізації та концентрації розподілів прямих іноземних інвестицій і регіонів за чисельністю населення.

Затверджено на засіданні кафедри статистики, обліку та аудиту
протокол № 1 від “26 ” серпня 2024 р.

Викладач

Олена КУЩЕНКО

Харківський національний університет імені В.Н. Каразіна

Факультет ЕкономічнийСпеціальність (напрямок) 051 «Економіка»

Освітня програма «Економічна аналітика та статистика»

Семестр XФорма навчання деннаРівень вищої освіти (освітньо-кваліфікаційний рівень): другий «магістерський»Навчальна дисципліна: «**Бізнес- статистика**»

ЗАЛІКОВЕ ЗАВДАННЯ № 7

1. Теоретична частина:**1.1.** Дати повну відповідь на питання: *Значення статистичного висновку у бізнес-аналізі.* (max 15)**2.** Визначити єдину правильну відповідь на кожне з двох тестових завдань: 2 тести (max 10)**2-а.** Яких типів бувають якісні дані?:

- А) порядкові і комбіновані;
- В) номінальні і комбіновані;
- С) номінальні і порядкові;
- Д) порядкові і рангові.

2-б. В сучасній фаховій літературі, залежно від масштабів проведення реінжинірингу, вчені виділяють такі його базові види:

- А) «разові поліпшення»;
- В) «кризовий»;
- С) «адаптивний»;
- Д) «тотальне моделювання».

3. Практична частина:**3.1.** Розв'язати задачу: (max 15)

Дані, зареєстровані районною службою зайнятості, свідчать про наявність зв'язку між рівнем утворення безробітних і терміном перерви в роботі:

Таблиця 7.1

Рівень освіти	Кількість безробітних, осіб	Середній термін перерви в роботі, міс.
Середнє	50	3
Середнє спеціальне	65	6
Вище	85	8
В цілому	200	6,1

Загальна дисперсія терміну перерви в роботі складає 5,8.

Визначте кореляційне відношення та зробіть висновки.Затверджено на засіданні кафедри статистики, обліку та аудиту
протокол № 1 від "26" серпня 2024 р.

Викладач



Олена КУЩЕНКО

Харківський національний університет імені В.Н. Каразіна

Факультет Економічний

Спеціальність (напрямок) 051 «Економіка»

Освітня програма «Економічна аналітика та статистика»

Семестр X

Форма навчання денна

Рівень вищої освіти (освітньо-кваліфікаційний рівень): другий «магістерський»

Навчальна дисципліна: «Бізнес-статистика»

ЗАЛІКОВЕ ЗАВДАННЯ № 8

1. Теоретична частина:

1.1. Дати повну відповідь на питання: *Дисперсійний аналіз у дослідженні бізнес-процесів (max 15)*2. Визначити єдину правильну відповідь на кожне з двох тестових завдань: 2
тести (max 10)

2-а. Прийняти неправильне рішення під час перевірки статистичної гіпотези можна у:

- А) одному випадку;
- В) двох випадках;
- С) трьох випадках;
- Д) трьох і більше випадках.

2-б. Коефіцієнт спряженості Чупрова розраховується за формулою:

$$A) K = \frac{\sum_1^n xy - n\bar{x}\bar{y}}{n \sqrt{\sigma_x^2 \sigma_y^2}}; \quad B) K = \sqrt{\frac{\chi^2}{n\sqrt{(k_1-1)(k_2-1)}}}; \quad C) K = \frac{\sum C - \sum H}{\sum C + \sum H}; \quad D) K = \frac{2S}{n(n-1)}.$$

3. Практична частина:

3.1. Розв'язати задачу: (max 15)

Маємо дані про аналітичне групування малих підприємств, що характеризує залежність обсягу продукції від вартості основних виробничих засобів:

Таблиця 8.1

Групи підприємств за обсягом середньорічної вартості основних виробничих засобів, млн. грн	Кількість підприємств	Обсяг виробленої продукції в середньому на одне підприємство, млн. грн
1,6 – 4,3	19	2,62
4,3 – 7,0	5	4,20
7,0 – 9,6	7	5,70
Разом	31	3,57

Виявити наявність і напрям кореляційного зв'язку між результативною та факторною ознаками. **Оцінити** щільність зв'язку; **Перевірити** його істотність з рівнем імовірності $P = 0,05$. Обчислена за первинними даними загальна дисперсія результативної ознаки $\sigma^2 = 2,38$. **Зробіть** обґрунтовані висновки.

Затверджено на засіданні кафедри статистики, обліку та аудиту
протокол № 1 від “26” серпня 2024 р.

Викладач

Олена КУЩЕНКО

Харківський національний університет імені В.Н. Каразіна

Факультет ЕкономічнийСпеціальність (напрямок) 051 «Економіка»

Освітня програма «Економічна аналітика та статистика»

Семестр ХФорма навчання деннаРівень вищої освіти (освітньо-кваліфікаційний рівень): другий «магістерський»Навчальна дисципліна: «**Бізнес-статистика**»

ЗАЛІКОВЕ ЗАВДАННЯ № 9

1. Теоретична частина:

1.1. Дати повну відповідь на питання: *Поняття про «портфоліо» та матриці портфельного аналізу.* (max 15)2. Визначити єдину правильну відповідь на кожне з двох тестових завдань: 2
тести (max 10)

2-а. З англійської мови ANOVA перекладається як:

А) кореляційний аналіз;

В) регресійний аналіз;

С) математичний аналіз;

Д) дисперсійний аналіз.

2-б. Коефіцієнт спряженості Крамера розраховується за формулою:

$$\text{А) } K = \frac{2S}{n(n-1)}; \text{ В) } K = \frac{\sum C - \sum H}{\sum C + \sum H}; \text{ С) } K = \frac{\sum_1^n xy - n\bar{x}\bar{y}}{n\sqrt{\sigma_x^2\sigma_y^2}}; \text{ Д) } K = \sqrt{\frac{\chi^2}{n(m_{\min}-1)}}.$$

3. Практична частина:

3.1. Розв'язати задачу: (max 15)

Для визначення якості продукції, виготовленої малим підприємством «Фортуна», за допомогою 5%-ної вибірки відібрано 200 виробів. Серед них виявили 40 бракованих.

Визначте граничну помилку вибірки для частки першосортних виробів за умови, що $P = 0,997$. **Врахуйте**, що відбір проводився механічною вибіркою. **Зробіть** обґрунтовані висновки.

Затверджено на засіданні кафедри статистики, обліку та аудиту
протокол № 1 від "26" серпня 2024 р.

Викладач



Олена КУЩЕНКО

Харківський національний університет імені В.Н. Каразіна

Факультет ЕкономічнийСпеціальність (напрямок) 051 «Економіка»

Освітня програма «Економічна аналітика та статистика»

Семестр ХФорма навчання деннаРівень вищої освіти (освітньо-кваліфікаційний рівень): другий «магістерський»Навчальна дисципліна: «**Бізнес-статистика**»

ЗАЛІКОВЕ ЗАВДАННЯ № 10

1. Теоретична частина:

1.1. Дати повну відповідь на питання *Випадкові величини і закони розподілу як основа бізнес-статистики.* (max 15)

2. Визначити єдину правильну відповідь на кожне з двох тестових завдань: 2 тести (max 10)

2-а. Концепція TQM (Total Quality Management) базується на необхідності постійного та одночасного удосконалення трьох складових:

- А) якості продукції;
- В) якості організації процесів;
- С) рівня кваліфікації персоналу;
- Д) якості ресурсів.

2-б. Розподіл молодих працівників за ступенем задоволеності умовами праці та професійною мобільністю характеризується такими даними:

Таблиця 10.1

Ступінь задоволеності умовами праці	Мають намір змінити професію	Не мають наміру змінювати професію	Разом
Задоволені	10	30	40
Незадоволені	20	5	25
Разом	30	35	65

Коефіцієнт контингенції становить: А) -0,54; В) -0,84; В) 0,25; Д) 0,47.

3. Практична частина:

3.1. Розв'язати задачу: (max 15)

Випущено продукції підприємствами міста (тис. грн.):

Таблиця 10.2

Номер підприємства	I квартал 2023 року	II квартал 2023 року
1	5	7
2	20	21
Разом	25	28

Визначити: відносні величини динаміки та структури.

Зробити обґрунтовані висновки.

Затверджено на засіданні кафедри статистики, обліку та аудиту
протокол № 1 від "26" серпня 2024 р.

Викладач



Олена КУЩЕНКО

Харківський національний університет імені В.Н. Каразіна

Факультет ЕкономічнийСпеціальність (напрямок) 051 «Економіка»

Освітня програма «Економічна аналітика та статистика»

Семестр ХФорма навчання деннаРівень вищої освіти (освітньо-кваліфікаційний рівень): другий «магістерський»Навчальна дисципліна: «**Бізнес-статистика**»

ЗАЛІКОВЕ ЗАВДАННЯ № 11

1. Теоретична частина:**1.1.** Дати повну відповідь на питання: *Основні переваги статистичних обстежень бізнес активності підприємств.* (**max 15**)**2.** Визначити єдину правильну відповідь на кожне з двох тестових завдань: 2 тести (**max 10**)**2-а.** Відповідність створеного продукту встановленим стандартам чи вимогам споживача дає змогу:

А) діаграма Парето;

В) контрольна карта;

С) стратифікація (розшарування);

Д) діаграма Ісікави (причинно-наслідкова діаграма).

2-б. Значення коефіцієнтів Чупрова, Крамера та Пірсона у бізнес-статистиці змінюється у межах від:А) $[0, 1]$; В) $[-1, 1]$; С) $[-\infty, +\infty]$; Д) $[0, +\infty]$.**3. Практична частина:****3.1.** Розв'язати задачу: (**max 15**)За оцінками експертів, ранги країн за їх політичною стабільністю R_x і станом платіжного балансу R_y розподілились таким чином:

Таблиця 11.1

Країна	A	B	C	D	O	M	N
R_x	1	3	2	4	7	6	5
R_y	2	3	1	5	6	7	4

Визначити: коефіцієнт рангової кореляції Спірмена. Доведіть істотність зв'язку з вірогідністю 0,95 критичне його значення $p_{0,95}(7) = 0,714$.**Зробіть** обґрунтовані аналітичні висновки.Затверджено на засіданні кафедри статистики, обліку та аудиту
протокол № 1 від "26" серпня 2024 р.

Викладач



Олена КУЩЕНКО

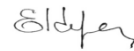
6. КРИТЕРІЇ ОЦІНЮВАННЯ ЗНАНЬ СТУДЕНТІВ ТА РОЗПОДІЛ БАЛІВ

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені В.Н.КАРАЗІНА

Економічний факультет
Кафедра статистики, обліку та аудиту

**КРИТЕРІЇ ОЦІНКИ
УСПІШНОСТІ ТА РЕЗУЛЬТАТІВ НАВЧАННЯ
З НАВЧАЛЬНОЇ ДИСЦИПЛІНИ
«Бізнес – статистика»**

Укладач –к.е.н.(PhD), доцент ЗВО **Олена КУЩЕНКО**



Затверджено
на засіданні кафедри статистики,
обліку та аудиту, протокол № 1
від 26 серпня 2024 року

Напрямок підготовки 051 «Економіка»
Освітня програма «Економічна аналітика та статистика»

З метою удосконалення системи контролю якості знань студентів, а також для стимулювання їх систематичної, ритмічної самостійної роботи и підвищення об'єктивності оцінювання знань та адаптації до вимог, визначених розробленим: «Порядком оцінювання знань студентів» і впровадженням в навчальний процес у Харківському національному університеті імені В.Н.Каразіна.

Порядок оцінювання унормовує основні принципи організації поточного, підсумкового та семестрового контролю знань студентів з усіх видів виконуваних робіт (теоретична підготовка, практична робота, самостійна робота, контрольна робота, наукова робота, тощо), а також методику переведення показників академічної успішності студентів ХНУ імені В.Н.Каразіна в національну шкалу

Таблиця 1

Шкала оцінювання

Сума балів за усі види навчальної діяльності протягом семестру	Оцінка
	для дворівневої шкали оцінювання
90–100	зараховано
70–89	
50–69	
1–49	не зараховано

Контрольні заходи є необхідним елементом зворотного зв'язку у процесі навчання. Вони визначають відповідність рівня набутих студентами знань, умінь та навичок вимогам нормативних документів щодо вищої освіти.

Таблиця 2

Критерії оцінки рівня знань при складанні заліку

Оцінка «зараховано» 35-40 балів	Студент має систематичні та глибокі знання учбового матеріалу, вміє без помилок виконувати практичні завдання та дослідження, які передбачені програмою курсу, засвоїв основну й ознайомився з додатковою літературою, викладає матеріал у логічній послідовності, робить узагальнення і висновки, наводить практичні приклади у контексті тематичного теоретичного матеріалу.
Оцінка «зараховано» 25-34 балів	Студент повністю засвоїв учбовий матеріал, знає основну літературу, вміє виконувати практичні завдання, викладає матеріал у логічній послідовності, робить певні узагальнення і висновки, але не наводить практичних прикладів у контексті тематичного теоретичного матеріалу або допускає незначні помилки у формуванні термінів, категорій, невеликі арифметичні помилки у розрахунках при вирішенні практичних завдань
Оцінка «зараховано» 15-24 балів	Студент засвоїв матеріал не у повному обсязі, дає неповну відповідь на поставлені теоретичні питання, допускається грубих помилок у розрахунках при розв'язанні практичного завдання
Оцінка «незараховано» 1-14 балів	Студент не засвоїв учбовий матеріал, дає неправильні відповіді на поставлені теоретичні питання, не вміє або неправильно виконує розрахунки при розв'язанні практичних завдань

**Форми організації контролю знань з навчальної дисципліни
«Бізнес - статистика»**

Поточний контроль знань студентів

Система поточного контролю студентів передбачає оцінювання їх за трьома основними напрямками:

- а) перевірка теоретичних знань;
- б) перевірка виконання індивідуальних практичних завдань;
- в) перевірка контрольних завдань.

Оцінювання рівня засвоєння теоретичних знань може проводитись в одній з альтернативних форм:

- за результатами відповідей про вивчення відповідних тем дисципліни або написання рефератів для іноземців згідно з тематичним планом;
- за підсумками тестування ;
- під час проведення усної співбесіди зі студентом, в процесі якої студент дає відповіді на питання з різних тем дисципліни.

Оцінювання рівня набутих практичних навичок здійснюється за результатами виконання та захисту контрольної роботи «Статистичний аналіз бізнес ситуацій», а також розрахункових завдань з кожної теми курсу.

Таблиця 3

Схема нарахування балів за індивідуальне завдання

№ п/п	Розділи завдання	max бал
1	Постановка проблеми та актуальність	5
2	Науково обґрунтований алгоритм збору, обробки та поширення інформації	5
3	Аналітична оцінка результатів дослідження	5
4	Бонусні бали за використання англійської мови	5
	Усього	20

Підсумковий контроль знань студентів

Формою підсумкового контролю знань студентів з дисципліни «Бізнес - статистика» є залік у письмовій формі з урахуванням поточної успішності студентів.

Таблиця 4

Розподіл балів, які отримують студенти

Поточний контроль, самостійна робота, індивідуальні завдання							Залікова робота	Сума
Розділ 1		Розділ 2		Контрольна робота, передбачена навчальним планом	Індивідуальне завдання (розрахунково-графічні роботи)	Разом		
T1. 2.3	T4.5	T6.7. 8	T9 .					
10	10	10	10	20	-	60	40	100

T1, T2 ... T10 – теми розділів.

T1. Роль статистики в бізнесі.

T2. Структури даних. Класифікація, розподіл та перетворення даних.

- T3. Випадкові величини і закони розподілу як основа бізнес-аналітики.
- T4. Статистичні гіпотези і статистичні висновки.
- T5. Дисперсійний аналіз у дослідженні бізнес-процесів.
- T6. Непараметричні методи в бізнес-аналізі.
- T7. Аналіз закономірностей для якісних даних (аналіз хі-квадрат).
- T8. Статистичні методи аналізу структури сукупності та її змін.
- T9. Статистичне оцінювання ефективності бізнес-процесів.

Для допуску до складання підсумкового заліку, здобувач вищої освіти повинен набрати не менше **15 балів** з навчальної дисципліни **«Бізнес-статистика»** під час поточного контролю, самостійної роботи та контрольної роботи.

Робоча навчальна програма дисципліни **«Бізнес - статистика»** передбачає як лекційні, так і практичні заняття, а також контрольну роботу, тому сума балів, які студент може набрати за поточним контролем, дорівнює 60, а сума балів підсумкового семестрового контролю – 40. Загальна сума балів поточного і підсумкового семестрового контролю складає 100.

ПОРЯДОК ПРОВЕДЕННЯ ЗАЛІКУ З КУРСУ «Бізнес - статистика» ТА КРИТЕРІЇ ОЦІНЮВАННЯ ЗНАНЬ СТУДЕНТІВ

Залік проводиться за єдиними білетами(див. приклад білету)

Додаток 20

Харківський національний університет імені В.Н. Каразіна

Факультет економічний

Спеціальність 051 «Економіка»

Освітня програма «Економічна аналітика та статистика»

Семестр 2

Форма навчання денна

Рівень вищої освіти (освітньо-кваліфікаційний рівень): другий (магістерський)

Навчальна дисципліна: «Бізнес-статистика»

ЗАЛІКОВЕ ЗАВДАННЯ № 4

I. Теоретична частина:

1.1. Дати повну відповідь на питання: Особливості становлення та розвитку бізнесу в Україні.

1.2. Визначити єдину правильну відповідь на кожне з тестових завдань:

1.2.1. Державна політика підтримки малого бізнесу –це сукупність пріоритетних рішень, які визначають основні напрями і форми сприяння розвитку малого підприємництва:

- А) правового характеру;
- В) економічного;
- С) організаційного;
- Д) усі відповіді вірні.

1.2.2. Основними причинами, які сприяють посиленню і зростанню тіньового сектору економіки на сучасному етапі, можна вважати:

- А) брак стабільного й збалансованого законодавства;
- В) недосконалість системи обліку та звітності;
- С) потужний податковий прес;
- Д) відсутність інвестиційної альтернативи тіньовим капіталам;
- Е) безконтрольне зростання кількості кримінальних структур;

Г) усі відповіді вірні.

11. Практична частина:

11.1. Розв'язати задачу:

Моємо дані про аналітичне групування малих підприємств, що характеризує залежність обсягу продукції від вартості основних виробничих засобів:

Таблиця 4.1

Групи підприємств за обсягом середньорічної вартості основних виробничих засобів, млн. грн	Кількість підприємств	Обсяг виробленої продукції в середньому на одне підприємство, млн. грн
1,6 – 4,3	19	2,62
4,3 – 7,0	5	4,20
7,0 - 9,6	7	5,70
Разом	31	3,57

Виявити наявність і напрям кореляційного зв'язку між результативною та факторною ознаками. **Оцінити** щільність зв'язку;

Перевірити його істотність з рівнем імовірності $\alpha = 0,05$.

Обчислена за первинними даними загальна дисперсія результативної ознаки $Q = 2,38$.

Зробіть обґрунтовані висновки.

Порядок наступний:

1. Завдання білета виконуються студентами виключно самостійно за 2 академічні години.
2. Після завершення заліку, роботи, які виконуються на бланках білету, приймаються викладачем від кожного студента.
3. Результати заліку, як правило, оголошуються у наступний день після його проведення, але не пізніше дня, що передує наступному заліку за розкладом.
4. У разі позитивної оцінки результати залікового завдання фіксуються у заліковій книжці за національною шкалою.
5. Кожен білет складається з трьох завдань (див. приклад).
6. Диференціація відповідей на завдання білетів у максимальних балах проводиться за шкалою, наведеною у таблиці 5:

Таблиця 5

№ питання	Максимальна кількість балів	Зміст відповіді
1	15	Глибокі знання учбового матеріалу
2(1,2)	10	У логічній послідовності знайти відповідь на тестові завдання серед запропонованих варіантів
3	15	Розв'язання практичного завдання

Максимальна кількість балів для залікового завдання дорівнює 40.

Переведення суми балів до національної шкали оцінювання здійснюється відповідно таблиць 1-5.