

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ  
ІМЕНІ ТАРАСА ШЕВЧЕНКА**

*механіко-математичний факультет  
кафедра загальної математики*

**Укладач:** доцент Кушніренко С.В.

**Методи багатомірної статистики**

**РОБОЧА НАВЧАЛЬНА ПРОГРАМА**

для студентів 4 курсу Військового інституту спеціальності "психологія"

**Затверджено**

на засіданні вченої ради  
механіко-математичного факультету  
протокол № 9 від 10.05.2011 р.

на засіданні кафедри  
загальної математики  
протокол № 9 від 27.04.2011 р.

*Зав. кафедри*

\_\_\_\_\_ Станжицький О.М.

Голова вченої ради  
механіко-математичного факультету

\_\_\_\_\_ Городній М.Ф.

**КИЇВ - 2011**

Робоча навчальна програма дисципліни  
«Методи багатомірної статистики»

**Укладач:** кандидат фізико-математичних наук,  
доцент Кушніренко С.В.

**Лектор:** кандидат фізико-математичних наук,  
доцент Кушніренко С.В.

**Викладач:** кандидат фізико-математичних наук,  
доцент Кушніренко С.В.

**Погоджено**  
з науково-методичною комісією  
протокол № 7 від 28.04.2011 р.

---

*Підпис голови НМК факультету/інституту*

## Методичні рекомендації по вивченню дисципліни

Навчальна дисципліна "Методи багатомірної статистики" є нормативною дисципліною для спеціальності "психологія" підготовки фахівців за освітньо-кваліфікаційним рівнем "бакалавр" Військового інституту.

Дисципліна викладається в VIII семестрі на 4 курсі в обсязі 1,5 кредитів (за Європейською Кредитно-Трансферною Системою ECTS) всього 58 годин, з них лекцій - 24 години, лабораторних - 24 години і 10 годин самостійної роботи.

Форма підсумкового контролю – іспит у письмовій формі у VIII семестрі.

### **Мета і завдання навчальної дисципліни.**

*Мета* викладання дисципліни "Методи багатомірної статистики" – ознайомити студентів з основними поняттями, теоретичними положеннями та сучасними математико-статистичними методами для розв'язування певних прикладних задач; сприяння розвитку логічного та аналітичного мислення студентів.

*Завдання* дисципліни "Методи багатомірної статистики" – теоретична та практична підготовка студентів, спрямована на оволодіння сучасними математико-статистичними методами, які повинні допомогти майбутнім військовим психологам розв'язувати певні прикладні задачі.

**Предмет навчальної дисципліни** "Методи багатомірної статистики" включає основні поняття теорії ймовірностей, теорії ігор та прийняття рішень у конфліктних ситуаціях, методи вибіркового обстеження, перевірки статистичних гіпотез, кластерного та факторного аналізу, моделі лінійної та нелінійної регресії, часові ряди. Поряд з вивченням математико-статистичних понять розглядаються їх практичні застосування в психології та інших соціальних науках.

### **Вимоги до знань та вмінь:**

*знати* основні поняття теорії ймовірностей, теорії ігор та прийняття рішень у конфліктних ситуаціях, методи вибіркового обстеження, методи прикладного статистичного аналізу;

*вміти* вибирати математико-статистичні методи та моделі для розв'язування прикладних задач; застосовувати сучасні методи прикладного статистичного аналізу, використовувати відповідне статистичне комп'ютерне забезпечення.

**Місце навчальної дисципліни в структурно-логічній схемі спеціальності.** Нормативна навчальна дисципліна "Методи багатомірної статистики" є складовою циклу професійної підготовки фахівців освітньо-кваліфікаційного рівня "бакалавр" спеціальності "психологія".

**Система контролю знань та умови складання іспиту чи заліку.** Навчальна дисципліна "Методи багатомірної статистики" оцінюється за модульно-рейтинговою системою. Вона складається з 2-х модулів: до першого входять 1-3 теми, до другого 4-7 теми.

Результати навчальної діяльності студентів оцінюються за 100-бальною шкалою.

**Форми поточного контролю:** написання та захист студентами домашніх самостійних завдань, виконання самостійних робіт під час лабораторних занять. Студент може отримати бали за усні відповіді та доповнення.

**Модульний контроль.** Дві модульні контрольні роботи.

### Змістовий модуль 1.

<i>Вид роботи</i>	<i>Бали</i>
Домашнє самостійне завдання 1. <i>Випадкові події та їх ймовірності. Числові характеристики випадкових величин.</i>	20
Домашнє самостійне завдання 2. <i>Вибірковий метод.</i>	20
Домашнє самостійне завдання 3. <i>Перевірка статистичних гіпотез.</i>	20
Усні відповіді та доповнення на заняттях.	10
Модульна контрольна робота 1.	30
<b>Сума</b>	<b>100</b>

### Змістовий модуль 2.

<i>Вид роботи</i>	<i>Бали</i>
Домашнє самостійне завдання 4. <i>Задачі регресійного аналізу.</i>	20
Домашнє самостійне завдання 5. <i>Прогнозування – часові ряди.</i>	20
Домашнє самостійне завдання 6. <i>Кластерний та факторний аналіз.</i>	20
Усні відповіді та доповнення на заняттях.	10
Модульна контрольна робота 2.	30
<b>Сума</b>	<b>100</b>

За результатами навчальної діяльності студент отримує підсумкову оцінку за 100-бальною шкалою, яка розраховується як середньозважене оцінок за кожен з двох модулів та оцінки за іспит за наступною формулою:

	<i>Змістовий модуль 1 (ЗМ<sub>1</sub>)</i>	<i>Змістовий модуль 2 (ЗМ<sub>2</sub>)</i>	<i>Комплексний підсумковий модуль (КПМ) (іспит)</i>	<i>Разом (підсумкова оцінка)</i>
Вагові коефіцієнти (%)	30% $k_1=0,3$	30% $k_2=0,3$	40% $k_3=0,4$	100%
Максимальна оцінка в балах	100	100	100	100

Розрахунок підсумкової оцінки:

$$ПО = ЗМ_1 \times k_1 + ЗМ_2 \times k_2 + КПМ \times k_3.$$

### Шкала відповідності

За 100-бальною шкалою	Оцінка за національною шкалою	
90 – 100	5	відмінно
85 – 89	4	добре
75 – 84		
65 – 74	3	задовільно
60 – 64		
35 – 59	2	незадовільно
1 – 34		

При цьому кількість балів відповідає оцінці:

- **1 - 34** - «незадовільно» з обов'язковим повторним вивченням дисципліни (F);
- **35 - 59** - «незадовільно» з можливістю повторного складання (FX) ;
- **60 - 64** - «задовільно» («достатньо») (E) ;
- **65 - 74** - «задовільно» (D);
- **75 - 84** - «добре»(C);
- **85 - 89** - «добре» («дуже добре») (B);
- **90 - 100** - «відмінно» (A).

Якщо за результатами модульно-рейтингового контролю студент отримав за два змістовні модулі менше 30 балів за роботу в семестрі, то він не допускається до іспиту і вважається таким, що не виконав усі види робіт, які передбачаються навчальним планом з дисципліни "Методи багатомірної статистики".

## НАВЧАЛЬНО-ТЕМАТИЧНИЙ ПЛАН ЛЕКЦІЙ І ЛАБОРАТОРНИХ ЗАНЯТЬ

№ теми	НАЗВА ТЕМИ	Кількість годин		
		Лекції	Лабораторні	Самостійна робота
<b>ЗМІСТОВИЙ МОДУЛЬ І. „Методи теорії ймовірностей та вибіркові методи”</b>				
1	Основи теорії ймовірностей	4	4	2
2	Вибірковий метод	4	4	2
3	Перевірка статистичних гіпотез	4	4	1
<b>ЗМІСТОВИЙ МОДУЛЬ ІІ. « Методи регресійного, кластерного, факторного аналізу, аналіз часових рядів»</b>				
4	Регресійні моделі	2	2	2
5	Ієрархічна кластерна техніка	4	4	1
6	Факторний аналіз	2	2	1
7	Часові ряди	4	4	1
<b>Всього годин за семестр 58, з них:</b>		<b>24</b>	<b>24</b>	<b>10</b>

# **Теми лекцій, лабораторних занять та завдання для самостійної роботи**

## **ЗМІСТОВИЙ МОДУЛЬ I**

### **Методи теорії ймовірностей та вибіркові методи ТЕМА №1. Основи теорії ймовірностей**

**Лекції 1-2.** Основи теорії ймовірностей. - 4 год.

Предмет та задачі теорії ймовірностей. Класичне та статистичне визначення ймовірності. Властивості ймовірності. Випадкові події та їх ймовірності.

Дискретні випадкові величини та їх розподіли. Геометричний розподіл. Розподіл Пуассона. Числові характеристики дискретних випадкових величин. Математичне сподівання випадкової величини, його властивості. Дисперсія випадкової величини, її властивості. Коефіцієнт кореляції випадкових величин, його властивості.

Неперервні випадкові величини. Щільність розподілу та функція розподілу, їх властивості. Рівномірний розподіл. Показниковий розподіл. Нормальний розподіл. Числові характеристики неперервних випадкових величин.

**Лабораторні заняття.** Зміст, кількість, тема та назва лабораторних занять повністю відповідають змісту лекційного матеріалу.

**Завдання для самостійної роботи за темою № 1.** – 2 год.

Прикладні задачі та соціально-економічні явища, що описуються даними моделями.

### **ТЕМА № 2. Вибірковий метод**

**Лекції 3-4.** Вибірки, їх числові характеристики, методи оцінювання параметрів розподілу. - 4 год.

Задачі математичної статистики. Математична статистика як базовий кількісний метод вивчення соціального явища. Вибірковий метод. Варіаційні ряди. Полігон. Гістограма. Оцінка щільності та функції розподілу. Емпірична функція розподілу .

Оцінювання невідомих параметрів розподілу. Вибіркове середнє. Вибіркова дисперсія. Вибірковий коефіцієнт кореляції. Їх властивості. Інтервальні оцінки: інтервал довіри, рівень надійності оцінки. Побудова інтервалу довіри для математичного сподівання нормального розподілу при відомій дисперсії.

**Завдання для самостійної роботи за темою № 2.** – 2 год.

Графічні засоби оцінювання невідомих параметрів розподілу в пакеті “Statistica”.

*Література* [1-6, 8, 9, 14].

### **ТЕМА № 3. Перевірка статистичних гіпотез**

**Лекції 5-6.** Перевірка статистичних гіпотез - 4 год.

Поняття про перевірку статистичних гіпотез. Задача вибору гіпотези: ймовірнісний підхід та методологія. Загальна схема перевірки статистичних гіпотез. Помилки 1-го та 2-го роду. Перевірка гіпотез про параметри розподілів. Критерії згоди про вигляд розподілу. Критерії згоди К.Пірсона: дискретний та неперервний випадки. Непараметричні тести. Перевірка гіпотез про однорідність вибірових даних.

**Завдання для самостійної роботи за темою № 3.** – 1 год.

Можливості перевірки статистичних гіпотез в пакеті “Statistica”.

*Література* [1-6, 8, 11, 13, 14].

**Модульна контрольна робота № 1.** – 1 год.

1. Впорядкована за величиною послідовність вибірових значень називається ...
2. Сума частот всіх варіант у дискретному розподілі вибірки = ...
3. Графічно інтервальні варіаційні ряди зображують у вигляді ...
4. Функція розподілу вибірки називається ... функцією розподілу.
5. Випадкова величина, яка використовується для перевірки статистичних гіпотез, називається ...
6. Площа гістограми відносних частот = ...
7. При збільшенні об'єму вибірки емпірична функція розподілу наближається до ...
8. Для знаходження похибки точкової оцінки параметра розподілу використовують ...
9. Множину значень статистичного критерію розбивають на дві області, які називаються ...
10. З яких міркувань має вибиратися нульова гіпотеза?



## **ЗМІСТОВИЙ МОДУЛЬ II**

### **Методи регресійного, кластерного, факторного аналізу, аналіз часових рядів**

#### **ТЕМА № 4. Регресійні моделі**

**Лекція 7.** Регресійний аналіз. - 2 год.

Задача регресійного аналізу. Лінійна регресія. Метод найменших квадратів побудови прямої регресії. Багатофакторна регресія. Перевірка значимості рівняння регресії. Нелінійне оцінювання: експоненційна та шматково-лінійна моделі. Загальна регресійна модель, прогнозування засобами регресійного аналізу.

**Завдання для самостійної роботи за темою № 4.** – 2 год.

Графічні ілюстрації регресійних моделей в пакеті “Statistica”.

*Література* [1-6, 8, 11, 13, 14].

#### **ТЕМА № 5. Ієрархічна кластерна техніка**

**Лекції 8-9.** Кластерний аналіз - 4 год.

Методи вимірювання схожості та відстані між окремими об'єктами. Подібність змінних з неперервними значеннями. Методи найближчих та найвіддаленіших сусідів. Метод середніх групових відстаней. Кластеризація за центрами. Метод Уорда. Дендрограма. Вибір кількості кластерів. Відстані між кластерами. Агломеративні та подрібнюючі методи кластеризації в пакеті “Statistica”.

**Завдання для самостійної роботи за темою № 5.** – 1 год.

Можливості кластеризації в пакеті “Statistica”.

*Література* [1-6, 8, 11, 13, 14].

#### **ТЕМА № 6. Факторний аналіз**

**Лекція 10.** Методи факторного аналізу - 2 год.

Загальний огляд. Методи зменшення кількості змінних. Визначення найбільш важливих факторів. Графічні зображення внеску факторів. Методи класифікації. Факторний аналіз в пакеті “Statistica”.

**Завдання для самостійної роботи за темою № 6.** – 1 год.

Можливості факторного аналізу в пакеті “Statistica”.

*Література* [1-6, 8, 11, 13, 14].

## ТЕМА № 7. Часові ряди

**Лекції 11-12.** Аналіз часових рядів - 4 год.

Приклади та структура часового ряду. Тренд, сезонна та циклічна компоненти. Кореляційний аналіз та аналітичне вирівнювання часового ряду. Білий шум, ARMA, ARIMA процеси. Коваріація та часткова кореляція. Прогнозування поведінки часового ряду. Адекватність моделі, аналіз залишків. Часові ряди з інтервенцією.

Дослідження часових рядів у пакеті “Statistica”.

**Завдання для самостійної роботи за темою № 7.** – 1 год.

Можливості аналізу часових рядів засобами пакету “Statistica”.

*Література* [1-6, 8, 11, 13, 14].

**Модульна контрольна робота № 2.** – 1 год.

1. Побудувати емпіричну функцію розподілу за вибіркою

$x_i$	-2	-1	1	3	5
$n_i$	10	20	13	17	40

Оцінити за побудованою функцією розподілу  $P\{-1 \leq \xi < 3\}$ .

2. За вибіркою об'єму  $n = 16$  з нормально розподіленої генеральної сукупності  $\xi$  обчислено вибіркове середнє  $\bar{x} = 10,2$ . Знайти надійний інтервал для оцінки з надійністю  $\gamma = 0,99$  невідомого математичного сподівання  $a$ , якщо: а)  $\sigma = 4$ ;

б)  $\sigma$  - невідоме, а за вибіркою обчислено  $D_B = \overline{x^2} - \bar{x}^2 = 4,6$ .

3. Перевірити, чи узгоджується вибірка при рівні значущості  $\alpha = 0,05$  з певним дискретним розподілом.

## Зразок завдання для самостійної роботи

За результатами 30 вимірів певного параметра отримали варіаційний ряд:

$x_i$	$x_1$	$x_2$	...	$x_k$
$n_i$	$n_1$	$n_2$	...	$n_k$

$x_i$  - значення параметра,  $n_i$  - відповідна частота.

- 1) Знайти точкові незміщені оцінки математичного сподівання та дисперсії досліджуваного параметра.
- 2) За припущення нормальності вибірових даних знайти надійний інтервал для оцінки математичного сподівання параметра з надійністю 0,95.
- 3) Побудувати відповідний даному інтервальному варіаційний ряд.
- 4) Побудувати полігон частот та емпіричну функцію розподілу параметра для дискретного варіаційного ряду та гістограму відносних частот за інтервальним варіаційним рядом.
- 5) Користуючись критерієм  $\chi^2$  при рівні значущості  $\alpha = 0,05$  перевірити гіпотезу про нормальний розподіл досліджуваного параметра.

### Практичні завдання в пакеті "Statistica":

1. Знаходження числових характеристик вибірки у модулі Basic Statistics.
2. Знаходження 0,95-надійного інтервалу для вибіркового середнього.
3. Обчислення кореляційної матриці та візуалізація лінійної залежності даних.
4. Обчислення ймовірності  $P\{\xi \in [a, b]\}$  за допомогою ймовірнісного калькулятора.
5. Перевірка гіпотези про рівність середніх двох незалежних вибірок.
6. Перевірка гіпотези про рівність середніх двох залежних вибірок.
7. Перевірка гіпотези про нормальний розподіл вибірових даних.
8. Застосування критерію  $\chi^2$  для перевірки гіпотези про незалежність двох показників.
9. Побудова регресійної моделі для даних з викидами.
10. Класифікація даних у модулі Cluster Analysis.
11. Прогнозування в багатофакторній регресійній моделі.
12. Побудова моделі часового ряду та перевірка її адекватності.

## Питання до екзамену

1. Випадкові події та операції над ними.
2. Класичне означення ймовірності.
3. Статистичне означення ймовірності.
4. Геометричні ймовірності.
5. Властивості ймовірності.
6. Дискретні випадкові величини та їх розподіли.
7. Числові характеристики дискретних випадкових величин.
8. Неперервні випадкові величини.
9. Нормальний розподіл та його числові характеристики.
10. Функція розподілу випадкової величини, її властивості.
11. Випадкові вектори, сумісні розподіли дискретних випадкових величин.
12. Коефіцієнт кореляції та його властивості.
13. Лінійна регресійна модель.
14. Основи вибіркового методу. Поняття генеральної та вибіркової сукупностей. Приклади суцільних та вибірових обстежень.
15. Вибірка та її числові характеристики.
16. Дискретні варіаційні ряди, полігон частот та відносних частот.
17. Інтервальні варіаційні ряди, гістограма частот та відносних частот.
18. Емпірична функція розподілу, її властивості та графік.
19. Точкові оцінки невідомих параметрів розподілу.
20. Інтервальні оцінки для математичного сподівання нормального розподілу.
21. Вибірковий коефіцієнт кореляції та його властивості.
22. Метод найменших квадратів для лінійної регресії.
23. Перевірка статистичних гіпотез. Поняття статистичного критерію, його потужність та рівень значущості.
24. Загальна схема перевірки статистичних гіпотез.
25. Перевірка гіпотези про узгодженість вибірки із дискретним розподілом.
26. Перевірка гіпотези про узгодженість вибірки із нормальним розподілом.
27. Критерій  $\chi^2$  як критерій незалежності.
28. Кластерний аналіз: методи вимірювання відстані між окремими об'єктами та кластерами.
29. Агломеративні методи та методи подрібнення в кластерному аналізі. Дендрограма.
30. Прикладні соціально-психологічні задачі, що описуються вивченими моделями.

## ДЖЕРЕЛА ІНФОРМАЦІЇ

1. Вища математика: Підручник: У 2 кн. – 2-ге вид., - За ред. Кулініча Г.Л. К.: Либідь, 2003.
2. Вища математика: Основні означення, приклади і задачі. За ред. Кулініча Г.Л. К.: Либідь, 1992.
3. Айвазян С.А., Енюков И.С., Мешкалин Л.Д. Прикладная статистика: Основы моделирования и первичная обработка данных. Справочное издание под ред. Айвазяна С.А. – М.: Финансы и статистика, 1983. – 471с.
4. Айвазян С.А., Енюков И.С., Мешкалин Л.Д. Прикладная статистика: Исследование зависимостей. Справочное издание под ред. Айвазяна С.А. – М.: Финансы и статистика, 1985. – 427с.
5. Боровиков В.П. Популярное введение в программу “Statistica”. М., 2001.
6. Вентцель Е.С. Теория вероятностей. М.: «Высшая школа», 1999, 575 с.
7. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике. М., Высшая школа, 1999.
8. Гмурман В.Е. Теория вероятностей и математическая статистика. М., Высшая школа, 1999.
9. Зінченко Н.М., Оленко А.Я. Аналітичні моделі та методи у соціології.- К.: ВПЦ “Київський університет”, 2000. - 106с.
10. Кремер Н.Ш. Теория вероятностей и математическая статистика. Изд. Юнити, М. 2000. – 543с.
11. Оленко А.Я. Комп’ютерна статистика.- К.: ВПЦ “Київський університет”, 2007. - 174с.
12. Черняк О.І. , Обушна О.М., Ставицька А.В. Теорія ймовірностей та математична статистика: Збірник задач: Навч. посібн.-К.:Т-во “Знання”, 2001.
13. Електронний підручник зі статистики: [www.statsoft.ru/home/textbook](http://www.statsoft.ru/home/textbook).