

Міністерство освіти і науки України

Харківський національний університет імені В.Н. Каразіна

Кафедра генетики і цитології

“ЗАТВЕРДЖУЮ”

Проректор з науково-педагогічної роботи

_____ Пантелеймонов А. В.

“ _____ ” _____ 2018 р.

Робоча програма навчальної дисципліни

Статистичні методи в біології

(назва навчальної дисципліни)

рівень вищої освіти _____ перший (бакалаврський) _____

галузь знань _____ 09 біологія _____
(шифр і назва)

спеціальність _____ 091 біологія _____
(шифр і назва)

освітня програма _____ Біологія _____
(шифр і назва)

спеціалізація _____ _____
(шифр і назва)

вид дисципліни _____ обов'язкова _____
обов'язкова / за вибором

факультет _____ Біологічний _____

2018 / 2019 навчальний рік

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

“29” серпня 2018 року, протокол № 8

РОЗРОБНИКИ ПРОГРАМИ:

Утевська Ольга Михайлівна, д.б.н., доцент кафедри генетики і цитології біологічного факультету ХНУ імені В.Н.Каразіна.

Програму схвалено на засіданні кафедри генетики і цитології

Протокол від “28” серпня 2018 року № 1

Завідувач кафедри генетики і цитології

_____ Атраментова Л. О.
(підпис) (прізвище та ініціали)

Програму погоджено методичною комісією біологічного факультету

Протокол від “28” серпня 2018 року № 1

Голова методичної комісії біологічного факультету

_____ Мартиненко В. В.
(підпис) (прізвище та ініціали)

ВСТУП

Програма навчальної дисципліни “Статистичні методи в біології” складена відповідно до освітньо-професійної (освітньо-наукової) програми підготовки *бакалаврів* спеціальності (напряму) *біологія*

1. Опис навчальної дисципліни

1.1. Мета викладання навчальної дисципліни

Навчання базовим методам статистичного аналізу даних, отриманих під час біологічних експериментів або спостережень, а також засобам презентації результатів дослідницької роботи. Ознайомлення з базовими принципами організації наукових досліджень.

1.2. Основні завдання вивчення дисципліни

- Надати уявлення про спектр статистичних методів, що застосовуються в біології.
- Навчити коректному вибору статистичних методів відповідно до мети дослідження, дизайну експерименту та типу експериментальних даних.
- Сформувати навички користування методами аналізу даних, зокрема із застосуванням статистичних програмних пакетів.
- Навчити засобам коректної наукової інтерпретації отриманих результатів.

1.3. Кількість кредитів – 3.

1.4. Загальна кількість годин – 90.

1.5. Характеристика навчальної дисципліни	
Нормативна	
Денна форма навчання	Заочна (дистанційна) форма навчання
Рік підготовки	
2-й	2-й
Семестр	
3-й	3-й
Лекції	
32 год.	4 год.
Практичні, семінарські заняття	
16 год.	4 год.
Лабораторні заняття	
-	-
Самостійна робота	
42 год.	82 год.
Індивідуальні завдання	
30 год.*	30 год.*

Примітка. * – 3 годин, виділених на самостійну роботу.

1.6. Заплановані результати навчання

Згідно з вимогами освітньо-професійної програми студенти повинні знати основні засоби організації наукового дослідження; основні методи систематизації даних; основні статистичні методи порівняння груп; основні статистичні методи аналізу зв'язку між ознаками; основні статистичні методи аналізу дії фактору на ознаку; принципи застосування багатовимірних статистичних методів; головні статистичні критерії. Студенти повинні визначати необхідний тип статистичного аналізу у залежності від особливостей вибірки, типу даних, завдання дослідження; виконувати необхідний тип статистичного аналізу, у тому числі з використанням програмного забезпечення; користуватися таблицями статистичних критеріїв; біологічно інтерпретувати отримані статистичні результати; застосовувати отримані теоретичні і практичні знання при аналізі даних під час виконання курсових і дипломних робіт.

2. Тематичний план навчальної дисципліни

Розділ 1. Описова статистика

Тема 1. Організація досліджень та методологія експерименту.

Експеримент та спостереження. Планування експерименту. Формування вибірок. Принцип випадковості. Контрольна група. Об'єм досліджень. Часовий тренд. Помилки досліджень. Програмне забезпечення для статистичного аналізу.

Тема 2. Систематизація даних.

Статистичні сукупності (генеральні та вибіркові). Ознаки (кількісні, якісні, рангові). Частотний розподіл (абсолютний, відносний). Графічне відображення частотних розподілів (гістограма, частотний полігон, стовбчаста діаграма, кругова діаграма).

Тема 3. Характеристики розподілів.

Середні величини (середня арифметична, мода, медіана). Показники варіації (стандартне відхилення, дисперсія, коефіцієнт варіації). Біноміальний розподіл. Нормальний розподіл. Асиметрія. Експес. Розподіл Пуасона.

Тема 4. Оцінки генеральних параметрів.

Оцінки генеральних параметрів. Статистичні похибки. Формули статистичних помилок. Довірчий інтервал.

Розділ 2. Статистичні прогнози

Тема 5. Перевірка розподілу даних на нормальність.

Методи перевірки розподілу на нормальність для великих та малих вибірок.

Тема 6. Статистичні гіпотези.

Статистичні гіпотези (нульова та альтернативна). Рівень значущості. Статистичні критерії (параметричні та непараметричні). Таблиці статистичних критеріїв.

Тема 7. Аналіз зв'язку.

Кореляційний аналіз (параметричний – критерій Пірсона r , непараметричний – критерій Спірмена, для якісних ознак – тетрахоричний та поліхоричний показники зв'язку. Регресійний аналіз (загальні поняття).

Тема 8. Порівняння двох груп.

Обрання методу. Залежні та незалежні групи. Порівняння двох груп за кількісною ознакою з нормальним типом розподілу (порівняння дисперсій – F-критерій, порівняння середніх арифметичних – t-критерій). Порівняння груп за ранговими та кількісними ознаками з будь-яким типом розподілу (критерії знаків, Уїлкоксона, Манна-Уїтні, тощо). Порівняння вибірових часток (F-критерій, кутове перетворення). Порівняння рядів розподілів (критерій χ^2).

Тема 9. Порівняння більш ніж двох груп.

Поправка Бонферроні. Дисперсійний аналіз (принципи параметричного дисперсійного аналізу, однофакторний дисперсійний аналіз).

Тема 10. Основи багатовимірної статистики.

Принципи и застосування методів багатовимірної статистики (множинна регресія, багатофакторний дисперсійний аналіз, факторний, кластерний, дискримінантний аналіз).

3. Структура навчальної дисципліни

Назви розділів і тем	Кількість годин											
	денна форма						заочна форма					
	усього	у тому числі					усьог о	у тому числі				
		л	п	лаб.	інд.	с. р.		л	п	ла б.	інд.	с. р.
1	2	3	4	5	6	7	8	9	10	11	12	13
Розділ 1. Описова статистика												
Тема 1. Організація досліджень та методологія експерименту	4	2	-			2		0,25	-			2
Тема 2. Систематизація даних	8	2	2			4		0,25	0,5			8
Тема 3. Характеристика розподілів	6	2	-			4		0,25	-			8
Тема 4. Оцінки генеральних параметрів	8	2	2			4		0,25	0,5			8
Разом за розділом 1	26	8	4			14		1	1			26

Розділ 2. Статистичні прогнози											
Тема 5. Перевірка розподілу даних на нормальність	8	2	2			4		0,5	0,5		8
Тема 6. Статистичні гіпотези	6	2	-			4		0,5	0,5		8
Тема 7. Аналіз зв'язку	16	6	4			6		0,5	0,5		12
Тема 8. Порівняння двох груп	18	8	4			6		0,5	0,5		12
Тема 9. Порівняння більш ніж двох груп	8	2	2			4		0,5	0,5		8
Тема 10. Основи багатовимірної статистики	8	4	-			4		0,5	0,5		8
Разом за розділом 2	64	24	12			28		3	3		56
Усього годин	90	32	16			42		4	4		82

4. Теми семінарських (практичних, лабораторних) занять

№ з/п	Назва теми	Кількість годин, д.в./з.в.
1	Систематизація даних. Частотний розподіл для кількісних і якісних ознак	2 / 0,5
2	Оцінки генеральних параметрів – статистичні похибки і довірчий інтервал	2 / 0,5
3	Перевірка розподілу на відповідність до нормального закону. Критерій Стюдента	2 / 0,5
4	Кореляційний зв'язок між кількісними ознаками	2 / 0,5
5	Кореляційний зв'язок між якісними ознаками	2 / 0,5
6	Порівняння двох груп за кількісною ознакою	2 / 0,5
7	Порівняння двох груп за якісною ознакою	2 / 0,5
8	Порівняння декількох груп і дія фактору на ознаку	2 / 0,5
	Разом	16 / 4

5. Завдання для самостійної роботи

№ з/п	Види, зміст самостійної роботи	Кількість годин, д.в./з.в.
1	Методи формування статистичних вибірок	2 / 2
2	Побудова частотного розподілу кількісної ознаки, розрахунки характеристик розподілу.	8 / 8
3	Оцінка генеральних параметрів розподілу	4 / 8
4	Перевірка розподілу даних на нормальність	4 / 8
5	Користування таблицями статистичних критеріїв і перевірка статистичних гіпотез	4 / 8
6	Розрахунки коефіцієнтів кореляції	6 / 8
7	Методи порівняння двох груп за кількісними і якісними ознаками	6 / 12
8	Теорія і практика дисперсійного аналізу	4 / 8
9	Методи багатовимірної статистики	4 / 8
	Разом	42 / 82

6. Індивідуальні завдання

Розрахунково-графічна робота

Приклад варіанту індивідуальних завдань

1. Описова статистика. Систематизація даних

Дано частотний розподіл ознаки:

Причина непрацездатності	Серцево-судинні захворювання	Онкологічні захворювання	Цукровий діабет	Травми
Відносна частота	44	21	8	27

- Дана ознака кількісна або якісна?
- Скільки класів виділено в мінливості ознаки?
- В яких одиницях виражено частоту?
- Визначте модальний клас.
- Зобразіть даний ряд графічно.

2. Пошук зв'язку між ознаками. Параметричні і непараметричні методи кореляційного аналізу

Для встановлення зв'язку між вмістом фосфору в ґрунті x і вмістом фосфору в злакових рослинах y проведено 9 аналізів. Визначте коефіцієнт кореляції r_{xy} , оцініть його статистичну значущість, зробіть висновок про зв'язок між ознаками.

x	1	4	5	9	13	11	23	23	28
y	64	71	54	81	93	76	77	95	109

3. Порівняння груп за кількісною ознакою

З двох ділянок ґрунту був проведений висів бактерій на 10 чашок з живильним середовищем, де кожна бактерія дає початок колонії. Число колоній, які утворилися на кожній чашці, склало:

Ділянка 1	7	4	8	10	10	7	16	11	7	12
Ділянка 2	5	10	9	4	7	5	1	11	12	15

Встановіть, чи достовірні відмінності в числі бактерій на порівнюваних ділянках.

4. Порівняння груп за якісною ознакою

При схрещуванні чорних курей, гетерозиготних по білому забарвленню, отримано наступне розщеплення: 18 чорних гомозиготних, 42 чорних гетерозиготних і 26 білих.

Чи відповідає воно очікуваному 1: 2: 1?

5. Дисперсійний аналіз

Випадковим чином були обрані 5 зразків крові кожної групи системи АВ0. Зміст лейкоцитів (тис.) наведено в таблиці. Чи впливає група крові на вміст лейкоцитів?

Групи крові			
А	В	АВ	О
5,0	7,0	7,0	5,3
5,5	8,0	7,1	8,0
8,0	5,0	9,0	6,7
7,7	9,9	9,2	9,3
10,0	6,3	7,7	6,7

7. Методи контролю

Методами контролю є самоконтроль, усне опитування, виконання індивідуальних завдань (рішення практичних задач) та поточні контрольні роботи (перевірка теоретичних знань).

8. Схема нарахування балів

Підсумковий семестровий контроль в формі заліку без виконання залікової роботи

Поточний контроль, самостійна робота, індивідуальні завдання		Сума
Поточний контроль	Індивідуальне завдання	
50	50	100

Шкала оцінювання

Сума балів за всі види навчальної діяльності протягом семестру	Оцінка для дворівневої шкали оцінювання
50 – 100	зараховано
1 – 49	не зараховано

9. Рекомендована література

Основна література

1. Атраментова Л. О., Утєвська О. М. Статистичні методи в біології. – Харків: ХНУ, 2007.
2. Атраментова Л. О., Утєвська О. М. Біометрія: Підручник. В 2-х ч. – Харків: Вид-во «Ранок», 2007.

Допоміжна література

1. Лакин Г. Ф. Биометрия. – Москва: ВШ, 1990.
2. Бейли Н. Статистические методы в биологии. – Москва: ВШ, 1962.
3. Плохинский Н. А. Алгоритмы биометрии. – Москва: ВШ, 1980.
4. Рокицкий П. Ф. Биологическая статистика. – Минск, 1973.
5. Терентьев П. В., Ростова Н. С. Практикум по биометрии. – Ленинград, 1977.
6. Урбах В.Ю. Статистический анализ в биологических и медицинских исследованиях. Москва: Медицина, 1975.
7. Бейли Н. Математика в биологии и медицине. Москва: МИР, 1970.
8. Дерябин В.Е. Многомерная статистика. – Москва: ВШ,., 1991.
9. Боровиков В.П. Популярное введение в программу Статистика. – Москва: Финансы и статистика. - 1998.

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

1. Електронний учебник StatSoft <http://www.statsoft.ru/home/textbook/default.htm>
2. <http://statosphere.ru/books-arch/statistica-books.html>
3. <http://www.statsoft.ru/home/textbook/modules/stbasic.html>
4. http://www.libedu.ru/l_b/rokickii_p_f/biologicheskaja_statistika_3e_izd_.html
5. <https://www.twirpx.com/file/18966/>
6. <http://www.lib.uniyar.ac.ru/edocs/iuni/20110318.pdf>

Інше методичне забезпечення: пакети статистичних програм Excel, SPSS, Statistica.

Додаток

Приклад контрольних питань (Тема 7. Аналіз зв'язку)

1. Що таке кореляція?
2. Що таке кореляційний аналіз?
3. Що таке часткова (парціальна) кореляція?
4. Що таке множинна кореляція?
5. За яких умов використовується коефіцієнт кореляції Пірсона?
6. За яких умов використовується коефіцієнт кореляції Спірмена?
7. За яких умов використовується бісеріальний коефіцієнт кореляції?
8. Для чого використовується тетрагорічеській показник зв'язку?
9. Для чого використовується полігорічеській показник зв'язку?
10. Якими методами кореляційного аналізу встановлюють наявність зв'язку між кількісними ознаками?
11. Якими методами кореляційного аналізу встановлюють наявність зв'язку між якісними ознаками?
12. Якими методами кореляційного аналізу встановлюють наявність зв'язку між кількісним і якісним ознаками?
13. Які значення може приймати коефіцієнт кореляції Пірсона?
14. Як знак коефіцієнта кореляції характеризує співвідношення ознак?
15. Яку інформацію можна отримати з числового значення коефіцієнта кореляції Пірсона?
16. Які значення коефіцієнта кореляції Пірсона вказують на сильний зв'язок між ознаками?
17. Які значення коефіцієнта кореляції Пірсона вказують на слабкий зв'язок між ознаками?
18. Які значення коефіцієнта кореляції Пірсона вказують на зв'язок середньої сили між ознаками?
19. Що таке пряма (позитивна) залежність між ознаками?
20. Що таке зворотна залежність між ознаками?
21. Які значення коефіцієнта кореляції вказують на зворотну залежність між ознаками?
22. Які значення коефіцієнта кореляції вказують на пряму залежність між ознаками?
23. Які значення приймає коефіцієнт кореляції при відсутності залежності?
24. У чому полягає разниця між лінійної і нелінійної регресією?
25. Що таке регресія?
26. Що таке регресійний аналіз?
27. Для чого використовується лінія регресії?
28. Що таке коефіцієнт регресії?
29. Що описує рівняння регресії?
30. Що таке скатер-діаграма?
31. Які особливості лінії регресії при прямій залежності між ознаками?
32. Які особливості лінії регресії при зворотній залежності між ознаками?
33. Зобразіть схематично лінію регресії для лінійної залежності.
34. Зобразіть схематично лінію регресії для прямої (позитивної) залежності.
35. Зобразіть схематично лінію регресії для зворотній залежності.
36. Зобразіть схематично скатер-діаграму.
37. Що говорить про зв'язок між ознаками коефіцієнт кореляції $r = -0,4$?
38. Який з коефіцієнтів кореляції характеризує найслабший зв'язок?
-0,71 0,08 0,62 -0,12