

**МІНІСТЕРСТВО ОХОРОНИ ЗДОРОВ'Я УКРАЇНИ
ВІННИЦЬКИЙ НАЦІОНАЛЬНИЙ МЕДИЧНИЙ УНІВЕРСИТЕТ**

Медична статистика

**МЕТОДИЧНІ ВКАЗІВКИ
навчальної дисципліни
з підготовки доктора філософії
на третьому (освітньо-науковому) рівні вищої освіти**

Галузі знань: 22 ОХОРОНА ЗДОРОВ'Я, 09 БІОЛОГІЯ

**Спеціальності:
229 Громадське здоров'я, 221 Стоматологія, 222 Медицина,
228 Педіатрія, 091 Біологія**

**Вінниця
2022 рік**

РОЗРОБЛЕНО ТА ВНЕСЕНО: Вінницький національний медичний університет ім. М.І. Пирогова

(повне найменування вищого навчального закладу)

РОЗРОБНИКИ:



д.мед.н, проф. О.М. Очередько

Обговорено на засіданні кафедри соціальної медицини та організації охорони здоров'я Вінницького національного медичного університету ім. М.І. Пирогова та рекомендовано до затвердження на центральній методичній раді / науковій комісії

“24” травня 2022 року, протокол №7

Затверджено на центральній методичній раді / науковій комісії

“24” травня 2022 року, протокол №7

Змістовий модуль

Тема 1. Планування статистичних досліджень. Валідність та достовірність, якісна репрезентативність та довірчі інтервали.

Генеральна сукупність. Вибірка, вибірковий результат. Гіпотеза. Основи статистичного висновку. Статистичні тести. Генерація статистичних даних проти генерації статистичних вимірів. Помилки першого та другого роду, потужність та p -рівень. Аналіз потужності. Авторська методологія аналізу потужності. Можливості інкорпорування елементів дизайнів, коваріат, різної природи зміщень. Графічні зображення. Буде приділена увага практичним аспектам використання аналізу потужності, традиційне програмне забезпечення та авторське в системі R.

Знати:

1. Відмінність вибіркового від істинного результату
2. Способи вираження істинного результату через вибірковий
3. Помилки першого та другого роду, потужність та p -рівень

Вміти:

Побудувати довірчий інтервал

Провести аналіз потужності за допомогою програмного забезпечення

Виразити результати аналізу потужності графічно

Література:

Тема 2. Підготовка зібраного матеріалу до статистичного аналізу. (Частина 1). Перевірка даних на припустимість. Обробка пропущених даних. Обробка віддалених (екстремальних) значень.

Вимоги до статистичних даних. Проблеми порушення властивостей статистичних даних. Проблеми аналізу лімітованих (constrained) даних. Проблеми аналізу даних з пропущеними значеннями. Проблеми аналізу даних з екстремальними значеннями. Способи перевірки даних на припустимість, трансформувannya і інверсне трансформувannya даних, метод якобіану. Моделі пропущених даних. Методи обробки пропущених даних. Способи виявлення та зменшення впливу екстремальних значень. Графічні зображення. Буде приділена увага практичним аспектам трансформації даних на основі традиційного програмного забезпечення в аналітичних системах R, EPIINFO.

Знати:

Перевірку на правильність заповнення таблиць даних.

Вплив типів змінних та їх групувань на зв'язок між змінними, поправки.

Використання кореляційної матриці за пропущеними даними.

Використання пропущених даних як даних.

Методи обробки пропущених даних.

Методи діагностики одномірних крайніх значень.

Методи діагностики багатомірних крайніх значень.

Методи та покази до трансформації даних.

Вміти:

Підготувати масиви даних за наявності крайніх значень.

Підготувати масиви з пропущеними даними.

Видаляти окремі групи одиниць спостережень, видаляти ознаки.

Заміщувати пропущені дані (логічне заміщення, заміщення середнім значенням, значенням регресії, значенням за оцінкою максимальної правдоподібності).

Описати крайні значення

Зменшувати вплив крайніх значень.

Працювати з головним меню програми EPIINFO, основні розділи. Робота з розділом

Make\Edit View:

Меню File
 Меню Edit
 Меню View
 Меню Insert
 Меню Format
 Меню Tools

Створити файлу-бази та таблиць даних за допомогою розділу Make\Edit EPIINFO, базової бібліотеки обміну даних R.

Література:

1. Chave, S.P.W. (2001). In: Oxford Textbook of Public Health. Vol 1, Oxford University Press
2. Concern for Europe's Tomorrow: health and the environment in the WHO European region / WHO European Center for Environment and Health. [Publ. on behalf of the World Health Organization Regional Office for Europe].-Stuttgart: Wiss.Verl.-Ges., 2005.- 537 p.
3. Hogarth, J. (1998). Glossary for Health Care Terminology, Public Health in Europe-4
4. Last, J.M. (2001). A Dictionary of Epidemiology, Oxford University Press
5. Mausner, J.S. and Kramer, S. (2005). Mausner and Bahn: Epidemiology - An Introductory Text, Saunders
6. Nalini Sathiakumar, Olga Henaо (2010). Методи исследования (пер. с англ. осущ. Центром спаркмана по Международному Здравоохранению, Бирмингем, шт. Алабама, США), США, - 189с.
7. Nalini Sathiakumar, Olga Henaо (2002). EPI INFO 2000. Лабораторное руководство (пер. с англ. осущ. Центром спаркмана по Международному Здравоохранению, Бирмингем, шт. Алабама, США), США, - 267с.
8. Park, K. (2007). Textbook of Preventive and social medicine. 15th Eddition.-Jabalpur
9. Roger Detels and Lester Breslow (2004). In: Oxford Textbook of Public Health. Vol 1. W.W. Holland, et al (eds) Oxford University Press

Тема 3. Підготовка зібраного матеріалу до статистичного аналізу. (Частина 2). Нормальність, лінійність і гомоскедастичність. Перетворення даних. Мультиколінеарність і сингулярність

Перевірка даних на нормальність розподілу. Підготовка масивів з ненормальним характером розподілу ознак. Перевірка змінних на лінійність попарного зв'язку. Трансформація змінних. Перевірка змінних на гомоскедастичність, гомогенність матриці коваріацій. Підготовка масивів з порушенням гомоскедастичності. Мультиколінеарність та сингулярність. Наслідки. Підготовка даних з мультиколінеарними змінними. Повний приклад проведення розвідувального аналізу даних та підготовки до подальшої статистичної обробки. Графічні зображення. Практичні аспекти трансформації даних на основі прикладу продемонстровані на основі традиційного програмного забезпечення в аналітичних системах R, EPIINFO.

Знати:

Нормальний розподіл

Коефіцієнт кореляції

Методи трансформування змінних

Матриці кореляцій та коваріацій

Поняття гомоскедастичності, гомогенності матриці коваріацій

Поняття мультиколінеарності та сингулярності

Вміти:

Провести розвідувальний аналіз даних за допомогою аналітичних систем R, EPIINFO.

Підготувати первинні дані до наступної обробки за допомогою аналітичних систем R, EPIINFO

Література:

1. Chave, S.P.W. (2001). In: Oxford Textbook of Public Health. Vol 1, Oxford University Press
2. Concern for Europe's Tomorrow: health and the environment in the WHO European region / WHO European Center for Environment and Health. [Publ. on behalf of the World Health Organization Regional Office for Europe].-Stuttgart: Wiss.Verl.-Ges., 2005.- 537 p.
3. Hogarth, J. (1998). Glossary for Health Care Terminology, Public Health in Europe-4
4. Last, J.M. (2001). A Dictionary of Epidemiology, Oxford University Press
5. Mausner, J.S. and Kramer, S. (2005). Mausner and Bahn: Epidemiology - An Introductory Text, Saunders
6. Nalini Sathiakumar, Olga Hena (2010). Методи исследования (пер. с англ. осущ. Центром спаркмана по Международному Здравоохранению, Бирмингем, шт. Алабама, США), США, - 189с.
7. Nalini Sathiakumar, Olga Hena (2002). EPI INFO 2000. Лабораторное руководство (пер. с англ. осущ. Центром спаркмана по Международному Здравоохранению, Бирмингем, шт. Алабама, США), США, - 267с.
8. Park, K. (2007). Textbook of Preventive and social medicine. 15th Eddition.-Jabalpur
9. Roger Detels and Lester Breslow (2004). In: Oxford Textbook of Public Health. Vol 1. W.W. Holland, et al (eds) Oxford University Press

Тема 4. Основи статистичного висновку. Тестування гіпотез.

Складові статистичної гіпотези. Види гіпотез. Параметричні і непараметричні тести. Довірчі інтервали. Поправки на множинні тести. Апостеріорні розподіли параметрів гіпотез. Основні статистичні тести. Графічні зображення тестів. Бокс-віскер графіки. Тестування гіпотез в аналітичній системі R.

Знати:

Складові гіпотез

Параметричні і непараметричні тести.

Довірчі інтервали.

Вміти:

Будувати одномірний та багатомірний довірчий інтервал

Користуватися поправками при множинних тестах

Тестувати гіпотези в аналітичній системі R

Створювати графічні зображення результатів тестів

Література:

Тема 5. Види прикладних статистичних аналізів даних. Класичні оптимізатори та МСМС інтегратори.

Класифікація статистичних аналізів даних. Параметричні і непараметричні аналізи. Загальні принципи оцінювання. Види оцінників. Переваги МСМС інтеграторів. Апостеріорні розподіли. Включення апріорної інформації. Накопичення інформації різних досліджень. Основи мета-аналізу.

Знати:

Класифікація статистичних аналізів даних.

Параметричні і непараметричні аналізи.
 Загальні принципи оцінювання. Види оцінщиків.
 Вміти:
 Обрати відповідний вид аналізу
 Конкретизувати оцінщик
 Література:

Тема 6. Аналіз табульованих даних. Лог-лінійний аналіз.

Формування і види табульованих даних. Види аналізів і оцінщиків табульованих даних. Класичний пропорційний ітеративний алгоритм. Лог-лінійний аналіз: дані, трансформації даних, формулювання лог-лінійної моделі, оцінювані регресійні контрасти і ефекти, аналітичні форми, асимптотики, тестування комбінованих ефектів, дисперсійний аналіз. Бібліотеки аналізу табульованих даних R. Авторська бібліотека *ltable* (розробник і утримувач Очередько О.М.) в складі аналітичної системи R.

Знати:
 Поняття основ табуляції та табульованих даних
 Види аналізів і оцінщиків табульованих даних
 Основні тести для табульованих даних
 Вміти:
 Здійснити табуляцію різних видів даних.
 Застосовувати програмне забезпечення для аналізу табульованих даних
 Розуміти результати аналізу табульованих даних
 Література:

Тема 7. Аналіз даних, отриманих за планом «випадок-контроль»

Звіти по випадкам/серіям випадків. Умови застосування, методика проведення, аналіз результатів. Кореляційне (екологічне) дослідження. Умови застосування, методика проведення, аналіз результатів. Перехресне дослідження. Умови застосування, методика проведення, аналіз результатів. Стратифікаційне дослідження. Умови застосування, методика проведення, аналіз результатів. Аналітичні дослідження. Типи, завдання, що вирішується. Характерні риси дослідження за планом випадок-контроль. Основні етапи дослідження за планом випадок-контроль: відбір випадків і контролю, підбір пар, вимір експозиції, аналіз і інтерпретація результатів. Приклади.

Знати:
 Основні епідеміологічні дизайни
 Дизайн «випадок-контроль»
 Властиві зміщення
 Вміти:
 Провести відбір випадків і контролю
 Провести підбір пар випадок-контроль
 Застосовувати програмне забезпечення для аналізу даних, отриманих за планом «випадок-контроль»
 Інтерпретувати результати аналізу
 Література:

Тема 8. Аналіз даних, отриманих за когортним планом

Характерні риси і принципи когортних досліджень. Типи когортних досліджень. Основні етапи когортних досліджень. Визначення об'єкту дослідження. Вимір експозиції.

Вибір контролю. Динамічне спостереження. Аналіз і інтерпретація результатів. Проблеми зміщення і методи їх усунення. Приклади.

Знати:

Характерні риси і принципи когортних досліджень.

Типи когортних досліджень.

Основні етапи когортних досліджень.

Вміти:

Здійснювати аналіз і інтерпретацію результатів когортних досліджень.

Усувати зміщення результатів когортних досліджень.

Застосовувати програмне забезпечення для аналізу даних, отриманих за когортним планом

Інтерпретувати результати аналізу

Література:

Тема 9. Аналіз панельних даних.

Типи панельних даних. Формати панельних даних. Вимоги до панельних даних. Ієрархічна композиція панельних даних. Методи аналізу різних типів панельних даних. Можливі зміщення результатів аналізу. Класичні оцінщики. МСМС оцінщики. Методи корекції зміщень. Практичні аспекти аналізу панельних даних на основі прикладів продемонстровані на основі традиційного програмного забезпечення в аналітичній системі R.

Знати:

Формати панельних даних.

Вимоги до панельних даних.

Ієрархічна композиція панельних даних.

Методи аналізу різних типів панельних даних.

Можливі зміщення результатів аналізу.

Вміти:

Здійснювати аналіз і інтерпретацію результатів аналізу панельних даних.

Усувати зміщення результатів аналізу панельних даних.

Застосовувати програмне забезпечення для аналізу панельних даних

Інтерпретувати результати аналізу панельних даних

Література:

Тема 10. Аналіз даних, отриманих за дизайнами КРД

Експериментальні (клінічні) дослідження. Завдання, що вирішуються. Характерні риси і принципи клінічних досліджень. Аспекти експериментів на тваринах. Умови експерименту на людях. Контрольовані рандомізовані дослідження (КРД). Етапи проведення КРД: розробка протоколу, визначення референтної і експериментальної популяцій, рандомізація, створення експериментальної експозиції, динамічне спостереження, оцінка наслідку, аналіз і інтерпретація результатів. Основні плани КРД. Типи КРД з прикладами. Види не випадкових експериментальних досліджень. Практичні аспекти аналізу панельних даних на основі прикладів продемонстровані на основі традиційного програмного забезпечення в аналітичній системі R і EPITABLE.

Знати:

Умови експерименту на людях

Етапи проведення КРД

Основні плани КРД

Вміти:

Працювати з підменю Study програми EPITABLE.

Аналіз результатів експериментального дослідження при вимірі наслідків кумулятивним рівнем. Пункт меню (Cohort-> Cumulative Incidence).

Аналіз результатів експериментального дослідження при вимірі наслідків первинним рівнем. Пункт меню (Cohort-> Incidence Density).

Аналіз результатів стратифікованого експериментального дослідження при вимірі наслідків кумулятивним рівнем. Пункт меню (Cohort-> Stratified, Cumulative).

Аналіз результатів стратифікованого експериментального дослідження при вимірі наслідків первинним рівнем. Пункт меню (Cohort-> Stratified, Density).

Література:

Тема 11. Аналіз тестів скринінгу

Концепція тестів скринінгу і діагностичних. Типи скринінгу. Критерії скринінгу. Джерела можливої варіації. Валідність тесту. Чутливість і специфічність. Прогностичне значення і діагностична потужність тесту. Хибно-позитивні та хибно-негативні результати. Продуктивність тесту. Комбінація тестів скринінгу. Проблема граничної лінії. ROC аналіз. Дизайни скринінгових досліджень. Приклади ROC аналізу та дизайнів скринінгових досліджень. Практичні аспекти аналізу скринінгових даних на основі прикладів продемонстровані на основі традиційного програмного забезпечення в аналітичній системі R

Знати:

Типи скринінгу.

Критерії скринінгу.

Джерела можливої варіації.

Валідність тесту.

Чутливість і специфічність.

Прогностичне значення і діагностична потужність тесту.

Продуктивність тесту

Вміти:

Визначити дизайн і провести оцінку діагностичного чи скринінгового тесту

Використовувати програмне забезпечення для оцінки тестів та побудови ROC кривої.

Література:

John M. Chambers and Trevor J. Hastie eds. (1992) *Statistical Models in S*, Wadsworth and Brooks/Cole, Pacific Grove, CA

K. B. Westlund \& L. T. Kurland (1953), Studies on multiple sclerosis in Winnipeg, Manitoba and New Orleans, Louisiana, *American Journal of Hygiene*, 57, 380–396

Тема 12. Аналіз дизайнів складної структури з ієрархічними експериментальними одиницями і рандомізованими ефектами

Рандомізовані і фіксовані ефекти. Основні складові дизайну. Експериментальні одиниці дизайну. Кроки перетворення дизайну в модель. Кроки тестування моделі з ієрархічними експериментальними одиницями і рандомізованими ефектами. Приклади. Приклади авторських досліджень. Авторські інновації. Практичні аспекти аналізу дизайнів складної структури з ієрархічними експериментальними одиницями і рандомізованими ефектами на основі прикладів продемонстровані на основі традиційного програмного забезпечення в аналітичній системі R.

Знати:

Рандомізовані і фіксовані ефекти.

Основні складові дизайну.

Експериментальні одиниці дизайну.

Вміти:

Перетворити дизайн в модель

Тестувати параметри моделі з ієрархічними експериментальними одиницями за допомогою традиційного програмного забезпечення в аналітичній системі R.

Інтерпретувати результати.

Література:

Тема 13. Аналіз даних з селекційним зміщенням вибірки.

Джерела селекційного зміщення. Загальний напрямок зміщення внаслідок селекції. Цензурування даних: кутове рішення, верхнє/нижнє кодування. Транкування даних. Селекція вибірки (sample selection). Зміщення Берксона. Виснаження вибірки. Методи корекції зміщень внаслідок селекції. Тобіт I, тобіт II, hurdle моделі. Оцінка відмінності різниць (DD-estimator), фрейлті моделі виживання, моделювання фракції виживання, моделі конкуруючих ризиків. Приклади. Практичні аспекти аналізу даних з селекційним зміщенням вибірки на основі прикладів продемонстрованих на основі традиційного програмного забезпечення в аналітичній системі R.

Знати:

Джерела селекційного зміщення.

Загальний напрямок зміщення внаслідок селекції.

Вміти:

Виявляти наявність і види селекційного зміщення вибірки

Тестувати параметри моделі з селекційним зміщенням вибірки за допомогою традиційного програмного забезпечення в аналітичній системі R.

Інтерпретувати результати.

Література:

Тема 14. Основні нелінійні регресії: тобіт, пробіт, логіт моделі.

Області практичного застосування. Формуляція індексної моделі. Функції правдоподібності та ітеративна процедура нелінійної оптимізації Ньютона-Рафсона. Інтерпретація ефектів та їх деривативів. Тестування ефектів та контрастів ефектів. Дисперсійний аналіз. Приклади. Приклади авторських досліджень. Реалізація на основі традиційного програмного забезпечення в аналітичній системі R.

Знати:

Суть тобіт, пробіт, логіт моделей

Області практичного застосування

Вміти:

Тестувати параметри тобіт, пробіт, логіт моделей за допомогою традиційного програмного забезпечення в аналітичній системі R.

Інтерпретувати результати.

Література:

Тема 14. Моделювання процесів виживаності. Модель Кокса, семипараметричні та фрейлті моделі.

Процеси виживаності. Практичні ситуації використання. Види розподілів часу до настання події та функцій ризику. Правостороннє, лівостороннє, інтервальне усічення та цензурування. Параметричні, семи-параметричні та непараметричні моделі виживання. Моделі акселерації. Фрейлті моделі виживання. Моделі виживання з фразією нечутливих. Приклади. Приклади авторських досліджень. Реалізація на основі традиційного програмного забезпечення в аналітичній системі R.

Знати:

Практичні ситуації використання процесів виживаності.

Види розподілів часу до настання події та функцій ризику.

Правостороннє, лівостороннє, інтервальне усічення та цензурування.

Параметричні, семи-параметричні та непараметричні моделі виживання.

Моделі акселерації.

Вміти:

Працювати з дизайнами, побудованих на процесах виживаності

Обробляти дані виживаності на основі моделі Кокса, тестувати параметри моделі Кокса за допомогою традиційного програмного забезпечення в аналітичній системі R.

Контрольні заходи та питання до диференційованого заліку/заліку/іспиту

Відмінність вибіркового від істинного результату

Способи вираження істинного результату через вибіркового

Помилки першого та другого роду, потужність та р-рівень

Побудувати довірчий інтервал

Провести аналіз потужності за допомогою програмного забезпечення

Виразити результати аналізу потужності графічно

Перевірку на правильність заповнення таблиць даних.

Вплив типів змінних та їх групувань на зв'язок між змінними, поправки.

Використання кореляційної матриці за пропущеними даними.

Використання пропущених даних як даних.

Методи обробки пропущених даних.

Методи діагностики одновимірних крайніх значень.

Методи діагностики багатовимірних крайніх значень.

Методи та покази до трансформації даних.

Підготовка масивів даних за наявності крайніх значень.

Обробка масивів з пропущеними даними.

Заміщення пропущених даних (логічне заміщення, заміщення середнім значенням, значенням регресії, значенням за оцінкою максимальної правдоподібності).

Способи діагностики та корекції крайніх значень.

Властивості нормального розподілу

Коефіцієнт кореляції

Методи трансформування змінних

Матриці кореляцій та коваріацій

Поняття гомоскедастичності, гомогенності матриці коваріацій

Поняття мультиколінеарності та сингулярності

Складові гіпотез

Параметричні і непараметричні тести.

Довірчі інтервали.

Побудова одновимірного та багатовимірного довірчих інтервалів

Поправки при множинних тестах

Класифікація статистичних аналізів даних.

Параметричні і непараметричні аналізи.

Загальні принципи оцінювання параметрів статистичної моделі. Види оцінщиків.

Поняття основ табуляції та табульованих даних

Види аналізів і оцінщиків табульованих даних

Основні тести для табульованих даних

Основні епідеміологічні дизайни

Дизайн «випадок-контроль». Властиві зміщення
 Характерні риси і принципи когортних досліджень.
 Типи когортних досліджень.
 Основні етапи когортних досліджень
 Формати панельних даних.
 Вимоги до панельних даних.
 Ієрархічна композиція панельних даних.
 Методи аналізу різних типів панельних даних.
 Можливі зміщення результатів аналізу панельних даних.
 Умови експерименту на людях
 Етапи проведення КРД
 Основні плани КРД
 Типи скринінгу, критерії скринінгу, джерела можливої варіації.
 Валідність тесту діагностичного чи скринінгу. Чутливість і специфічність.
 Прогностичне значення і діагностична потужність тесту скринінгу. Продуктивність тесту
 Рандомізовані і фіксовані ефекти.
 Основні складові дизайну.
 Експериментальні одиниці дизайну.
 Рандомізовані і фіксовані ефекти.
 Перетворення дизайну в модель
 Особливості тестування параметрів моделі з ієрархічними експериментальними одиницями.
 Суть тобіт, пробіт, логіт моделей
 Області практичного застосування тобіт, пробіт, логіт моделей
 Особливості тестування параметрів тобіт, пробіт, логіт моделей.
 Інтерпретація результатів тобіт, пробіт, логіт моделей.
 Суть лог-лінійної моделі
 Області практичного застосування лог-лінійної моделі
 Особливості тестування параметрів лог-лінійної моделі.
 Інтерпретація результатів лог-лінійної моделі
 Практичні ситуації використання процесів виживаності.
 Види розподілів часу до настання події та функцій ризику.
 Правостороннє, лівостороннє, інтервальне усічення та цензурування.
 Параметричні, семи-параметричні та непараметричні моделі виживання.
 Моделі акселерації.
 Особливості дизайнів, побудованих на процесах виживаності
 Особливості тестування параметрів моделі виживаності Кокса.
 Інтерпретація результатів моделі виживаності Кокса