

«ЗАТВЕРДЖУЮ»
Завідувач кафедру
інформатики та кінезіології

_____ І.П. Заневский

(підпис, ініціали, прізвище)

_____ 200__р

ЛЕКЦІЯ №5
з навчальної дисципліни
«КОМП'ЮТЕРНА ТЕХНІКА І МАТЕМАТИЧНА СТАТИСТИКА»

Тема: Статистичні гіпотези

Навчальний потік

для студентів першого року навчання факультету спорту,
фізичного виховання, здоров'я людини і туризму

Навчальна мета: Ознайомити студентів з основними поняттями і значеннями, етапами перевірки статистичної гіпотези, зокрема з гіпотезою про рівність вибіркового середнього і заданого числа, гіпотезою про рівність середніх двох незалежних сукупностей, гіпотезою про рівність середніх двох зв'язаних сукупностей.

Виховна мета: Розвивати логічне мислення, увагу, пам'ять, спостережливість, зацікавити математичною статистикою як наукою, розширювати кругозір з математичної статистики в галузі ФКіС, а також формувати систематизовані знання з математичної статистики.

Навчальні питання і розподілення часу:

Вступ _____ -10...хв.

1. Основні поняття і значення. _____ 10хв.

2. Загальний порядок перевірки статистичної гіпотези. __ 15хв.

3. Гіпотеза про рівність вибіркового середнього і заданого числа. __ 15хв.

4. Гіпотези про рівність середніх двох незалежних сукупностей. ____ 15хв.

5. Гіпотези про рівність середніх двох зв'язаних сукупностей. ____ 15 хв.

Заключення та відповіді на запитання _____ - 10хв.

Навчальна література

1. Ільків О.С. Матвіїв В.І. Інформатика та комп'ютерна техніка (з елементами математичної статистики): Навч. посіб. – Львів: ЛДУФК, 2010.
 2. Дибкова Л.М. Інформатика та комп'ютерна техніка: Навч. Посіб. – Київ: Академвидав, 2005.
 3. Литвин І.І. Інформатика: теоретичні основи і практикум. – Львів: Новий світ, 2004.
 4. Глинський Я.М. Практикум з інформатики. – Львів: Деол, 2005.
 5. Русіло П.О., Заневський І.П. Математична статистика. Обробка і аналіз результатів спортивних вимірів. - Львів, 1995.
 6. Статистика. Підручник за ред. С.С. Герасименка. Київ: КНЕУ, 2000.
 7. Основы математической статистики. Под ред. В.С.Иванова, -М.: ФиС, 1990.
 8. Глинський Я.М. Інформатика: інформаційні технології. - Львів: Деол, 2003.
-

Додаткова

1. Петров П.К. Информационные технологии в физической культуре и спорте. – М.: ФКиС, 2008.
2. Височанський В.С., Кардаш А.І., Костів О.В., Черняхівський В.В. “Елементи інформатики”, - Львів: “Світ”, 1990.
3. Андреев А.Г. и др. Microsoft Windows XP: Home Edition и Professional. Русские версии / Под общ. ред. А.Н. Чекмарева. – СПб.: БХВ - Петербург, 2003.
4. Руденко В.Д., Макаруч О.М., Патланжоглу М.О. Практичний курс інформатики / За ред. Мадзігога В.М. - К.: Фенікс, 1997.
5. Фигурнов В.Е. IBM PC для пользователя. - М.: Финансы и статистика, 1992.
6. Зайцева Т.И., Смирнова О.Ю. В сб.: Информационные технологии в образовании. – М., 2000.
7. Kosmol A., Kosmol I. Komputery - nowoczesne technologie w sporcie. - Warszawa: AWF, 1999
8. Волков В.Ю. Компьютерные технологии в образовательном процессе по физической культуре в вузе: Монография. – СПб.: СПбГТУ, 1997.
9. Макарова М.В. Електронна комерція. Посібник. Київ. Видавничий центр "Академія". 2002.

Тема 5. Статистичні гіпотези

1. Основні поняття і значення.
2. Загальний порядок перевірки статистичної гіпотези.
3. Гіпотеза про рівність вибіркового середнього і заданого числа.
4. Гіпотези про рівність середніх двох незалежних сукупностей.
5. Гіпотези про рівність середніх двох зв'язаних сукупностей.

Поняття гіпотези, з яким має справу математична статистика, більш вузьке, ніж загальне поняття наукової гіпотези. Наприклад, висуваються наукові гіпотези про походження Всесвіту, про відвідування інопланетянами Землі кілька тисячоліть тому назад чи про існування нафти на глибині 15 км під поверхнею Землі. Хоча подібні гіпотези заслуговують особливої уваги, їх неможливо піддати перевірці з допомогою методів математичної статистики.

Статистичні гіпотези торкаються перевірки спостережуваних випадкових величин або випадкових подій. Середній зріст юнаків за останні 20 років збільшився на 10 см; або: в місті спостерігається епідемія грипу; або: недоліки при виготовленні болтів на двох станках, які спостерігалися, однакові – це приклади статистичних гіпотез. Їх перевірка здійснюється шляхом співставлення з результатами спостережень.

Але результати спостережень залежать від випадку. Тому статистичні гіпотези носять не категоричний, однозначний характер, а характер правдоподібного твердження, яке теж має цілком визначену ймовірність.

І так завжди: всі висновки про правильність чи помилковість статистичних гіпотез, які можна висловити на основі кінцевого числа спостережень носять ймовірнісний, а не детермінований, однозначний характер. Тому, якщо не знайдено протиріч, на основі яких гіпотеза повинна бути відхилена, то статистична гіпотеза відкрита для подальшої перевірки.

Статистична гіпотеза – це твердження про характер розподілу параметрів генеральної сукупності, яке перевіряється математичними методами.

Гіпотези є: *нульова і альтернативна*.

Нульову гіпотезу прийнято позначати H_0 , а альтернативну гіпотезу – H_A .

Прийнято нульову гіпотезу вважати початковою.

Нульова гіпотеза передбачає, що між параметрами генеральних сукупностей немає очікуваних розбіжностей. Запис нульової гіпотези має, наприклад, вигляд $H_0: (\bar{x} = \bar{y})$ що означає: між сукупностями X і Y статистично немає відмінностей варіацій.

Альтернативна гіпотеза передбачає протилежне твердження до нульової гіпотези, наприклад, $H_A: (\bar{x} \neq \bar{y})$, тобто варіації генеральних сукупностей статистично неоднакові.

Точність перевірки гіпотези вибирається з практичних міркувань. Відповідна похибка визначається рівнем значущості α .

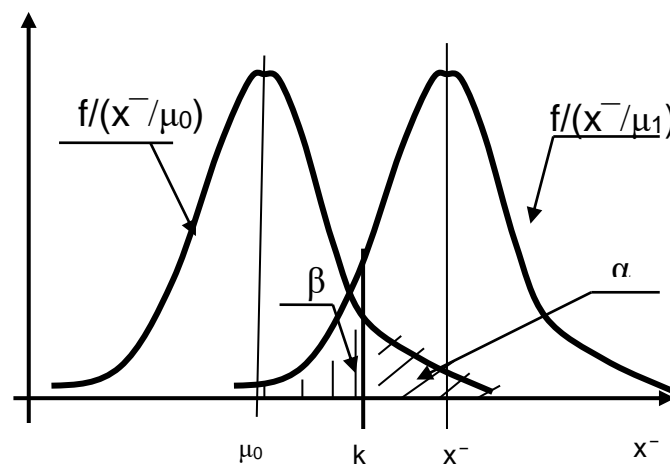
Імовірність прийняття правильної гіпотези називають **вірогідністю** β і визначають за формулою $\beta = (1 - \alpha) \cdot 100 \%$.

При перевірці гіпотез похибки є двох родів.

Похибка першого роду – відхилення істинного припущення.

Похибка другого роду – прийняття помилкового твердження.

Процедура перевірки гіпотез зводиться до обчислення за вибірковими даними величини, яку називають статистичним критерієм.



Критерії

Рис. 7. Похибка першого і другого роду.

яє перевірити

статистичну гіпотезу і забезпечує прийняття істинної і відхилення невірної гіпотези з наперед заданою імовірністю.

Статистичні критерії значущості бувають трьох типів:

1. Параметричні критерії, які оперують фізичними величинами (м, кг, с).
2. Непараметричні критерії, в яких використовують величини, що не мають розмірностей фізичних величин (наприклад, місця, ранги, спортивні розряди).

3. Критерії узгодженості, які використовують для перевірки узгодженості розподілу генеральної сукупності з прийнятою раніше теоретичною моделлю (наприклад, з нормальним розподілом).

Етапи перевірки статистичної гіпотези:

1. Формулювання нульової та альтернативної гіпотез.
2. Розрахунок необхідних статистичних характеристик.
3. Вибір рівня значущості α .
4. Вибір та обчислення розрахункового значення критерію для перевірки гіпотез.

Критичне значення t-критерію Стьюдента визначається за таблицею.

5. Порівняння розрахункового і табличного значень критерію, на основі чого робиться висновок про прийняття чи відхилення гіпотези.

Статистичні гіпотези в спорті використовуються:

- при порівнянні вибіркового середнього арифметичного з середнім значенням генеральної сукупності;
- при перевірці однаковості результатів в різних групах (незв'язані вибірки). При порівнянні припускають, що і перша і друга вибірки належать до однієї генеральної сукупності, і їх вибіркові значення не відрізняються;
- при перевірці результатів одних і тих же спортсменів на різних етапах підготовки (перевіряють, чи змінився стан спортсменів). В цьому випадку вибірки рівнозначні, а результати можуть бути з'єднані в пари (кожна пара – це результат вимірів на одній людині на початку та в кінці експерименту).

Порівняння середнього арифметичного із заданим числом.

Завдання: Визначити, чи відповідає рівень фізичного розвитку школярів певного віку загальноприйнятій нормі.

Вихідні дані:

1. Результати вимірювань рівня фізичного розвитку школярів (вибірка обсягом n).

2. (μ – норма фізичного розвитку школярів.

Порядок виконання обчислень:

1. Сформулювати нульову гіпотезу про рівень фізичного розвитку школярів.

2. Визначити вибіркові числові характеристики, а саме:

а) середнє арифметичне значення:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

б) оцінку значення дисперсії генеральної сукупності:

$$D = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

в) стандартну похибку середнього арифметичного:

$$S_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

3. Обчислити розрахункове значення статистичного t -критерію Стьюдента за формулою:

$$t_p = \frac{\bar{x} - \mu}{S_{\bar{x}}}$$

4. Обрати рівень значущості, наприклад, $\alpha = 0,05$, що відповідає вірогідності $\beta = (1 - \alpha) \cdot 100 \% = 95 \%$.

5. Визначити значення t -критерію Стьюдента $t(\alpha, \nu)$ згідно таблиці при $\nu = n - 1$, де ν – число ступенів свободи, n – обсяг вибірки.

6. Порівняти табличні значення t – критерію з розрахунковими і зробити висновок про прийняття чи відхилення гіпотези:

а) якщо $t_p < t_\alpha$, то з вірогідністю $\beta = (1 - \alpha) \cdot 100\% = 95\%$. можна стверджувати, що нульова гіпотеза приймається і рівень фізичного розвитку школярів відповідає нормі;

б) якщо $t_p > t_\alpha$, то з вірогідністю $\beta = (1 - \alpha) \cdot 100\% = 95\%$. можна стверджувати, що нульова гіпотеза відхиляється і рівень фізичного розвитку школярів не відповідає нормі.

Порівняння **середніх арифметичних** 2-х незв'язаних сукупностей.

Завдання: порівняти рівень фізичного розвитку студентів 2-х спеціалізацій.

Вихідні дані:

1. Вибіркові параметри x , які характеризують рівень фізичного стану студентів 1-ї спеціалізації.
2. Вибіркові параметри y , які характеризують рівень фізичного стану студентів 2-ї спеціалізації.

Порядок виконання обчислень:

1. Сформулювати нульову гіпотезу про рівність фізичного стану студентів 2-х спеціалізацій.
2. Визначити вибіркові числові характеристики:
 - а) середні арифметичні значення кожної вибірки:

$$\bar{x} = \frac{\sum_{i=1}^{n_x} x_i}{n_x} \qquad \bar{y} = \frac{\sum_{i=1}^{n_y} y_i}{n_y}$$

- б) оцінки дисперсії для кожної вибірки:

$$D_x = \frac{1}{n_x} \sum_{i=1}^{n_x} (x_i - \bar{x})^2 \qquad D_y = \frac{1}{n_y} \sum_{i=1}^{n_y} (y_i - \bar{y})^2$$

3. Визначити стандартну похибку різниці середніх арифметичних незв'язаних сукупностей:

$$S_{\bar{x}-\bar{y}} = \sqrt{\frac{D_x}{n_x} + \frac{D_y}{n_y}}$$

4. Обчислити розрахункове значення статистичного t-критерію Стьюдента за формулою:

$$t_p = \frac{|\bar{x} - \bar{y}|}{S_{\bar{x}-\bar{y}}}$$

5. Обрати рівень значущості α , що відповідає $\beta = (1 - \alpha) \cdot 100 \% = 95 \%$, і визначити табличне значення t-критерію Стьюдента. Число ступенів свободи ν розраховується за формулою:

$$\nu = (n_x - 1) + (n_y - 1)$$

6. Порівняти розрахункове значення t-критерію Стьюдента з табличним і зробити висновок про прийняття чи відхилення нульової гіпотези:

а) якщо розрахункове значення t-критерію менше за табличне, то фізичний стан цих 2-х груп є однаковий і нульова гіпотеза приймається;

б) якщо розрахункове значення t-критерію більше або рівне табличному, то фізичний стан цих двох груп не є однаковий і нульова гіпотеза відхиляється.

Порівняння середніх арифметичних 2-х зв'язаних сукупностей.

Завдання: визначити, чи відбуваються зміни спортивних показників студентів за час тренувань, що проходять за даним методом.

Вихідні дані:

1. Результати спортивних показників студентів:

x_1 – значення показника першого виміру (початок експерименту);

x_2 – значення показника другого виміру (кінець експерименту).

Порядок виконання обчислень:

1. На основі припущення про нормальний розподіл різниці $d_i = x_{1i} - x_{2i}$ сформулювати нульову гіпотезу про рівність двох середніх арифметичних спортивних показників.

2. Визначити вибіркові характеристики:

а) різницю між результатами першого і другого вимірів

$$d_i = x_{1i} - x_{2i}$$

б) середню арифметичну різницю

$$\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i$$

в) дисперсію різниці

$$D_d = \frac{1}{n-1} \sum_{i=1}^n (d_i - \bar{d})^2$$

г) стандартну похибку середньої різниці

$$S_{\bar{d}} = \sqrt{\frac{D_d}{n}}$$

3. Обчислити розрахункове значення статистичного t-критерію Стьюдента за формулою:

$$t_p = \frac{|\bar{d}|}{S_{\bar{d}}}$$

4. Обрати рівень значущості, наприклад, $\alpha = 0,05$, що відповідає вірогідності $\beta = 95\%$ і визначити табличне значення t-критерію Стьюдента $t_{\alpha\nu}$ де $\nu = n-1$ – число ступенів свободи.

5. Порівняти розрахункове значення t-критерію з табличним і зробити висновок про суттєві зміни спортивних показників студентів:

а) якщо $t_p < t_{\alpha}$ суттєвих змін за час тренувань не сталося;

б) якщо $t_p \geq t_{\alpha}$ спортивний рівень або покращився, або погіршився.

Питання для самоконтролю:

1. Дайте визначення поняттю «**Статистична гіпотеза**».
2. Похибки якого роду виникають при перевірці статистичних гіпотез.
3. Яка процедура перевірки гіпотез.
4. Який порядок виконання обчислень у випадку порівняння середнього арифметичного із заданим числом?
5. Який порядок виконання обчислень у випадку порівняння середніх арифметичних 2-х зв'язаних сукупностей?

Навчально-матеріальне забезпечення

Мультимедійний проектор

Самостійна робота:

1. Перевірка гіпотези про рівність середніх двох зв'язаних сукупностей.
2. Оцінка відповідності середнього рівня фізичного розвитку групи молоді загальноприйнятій нормі.
3. Порівняння між собою двох різних методик підготовки. Оцінка достовірності змін в часі.

Лекцію розробили: к.пед. н., доц. О.С. Ільків.

(посада, вчений ступінь, вчене звання, підпис, ініціали, прізвище)

Обговорено на засіданні кафедри: інформатики та кінезіології

(назва кафедри)

Протокол від _____ 20__ р. № _____