

### ЛАБОРАТОРНЕ ЗАНЯТТЯ 3

## ПЕРЕВІРКА СТАТИСТИЧНИХ ГІПОТЕЗ ІЗ ЗАСТОСУВАННЯМ ПАРАМЕТРИЧНИХ ТА НЕПАРАМЕТРИЧНИХ КРИТЕРІЇВ


**Мета:** ознайомитись із загальними принципами перевірки статистичних гіпотез; опанувати базові методи порівняння груп із застосуванням параметричного (Т-критерій) та непараметричного (U-критерій) критеріїв.

**Обладнання та матеріали:** ПЕОМ із встановленою ОС Microsoft Windows XP або Windows Seven, процесор таблиць Microsoft Excel 2007 або 2010, Інтернет браузер, методичні вказівки.

### План:

1. Формулювання нульової та альтернативної гіпотез.
2. Перевірка статистичних гіпотез із використанням параметричних критеріїв.
  - 2.1. Т-критерій для незалежних вибірок.
  - 2.2. Т-критерій для залежних вибірок.
3. Непараметричний критерій Мана-Уїтні.

### Теоретичні відомості.

 Загальні принципи порівняння відмінностей між групами ґрунтуються на аналізі так званої **нульової гіпотези**. Відповідно до цієї гіпотези, спочатку приймається припущення, що між даними показниками (або групами, на основі яких вони отримані) достовірної відмінності немає, тобто **обидві ці групи разом складають один і той же однорідний матеріал, одну сукупність**. Статистичний аналіз повинен привести або до відхилення нульової гіпотези, якщо доведена достовірність отриманих відмінностей, або до її збереження, якщо достовірність відмінностей не доведена, тобто виявлені відмінності визнаються **випадковими**. Але, так як всі статистичні показники та відмінності між ними характеризуються певними рівнями значущості, то відкидання нульової гіпотези має бути пов'язане з прийняттям певного рівня значущості ( $P$ ). Так, якщо необхідним визнаний рівень значущості 0,01 і якщо ймовірність достовірності (довірча імовірність) даного статистичного показника, або різниці між показниками, не задовольняє цій умові, тобто вона нижче 0,99 (наприклад, 0,97, 0,91 або 0,88), то немає підстав для відкидання нульової гіпотези. Її треба вважати правильною принаймні до тих пір, поки нові дані не дадуть можливості її спростувати, довівши, що існуючі відмінності не є чисто випадковими.

Звичайно, і у тому випадку, коли нульова гіпотеза вважається спростованою, якийсь шанс, що вона в дійсності вірна, залишається. При рівні значущості 0,01 цей шанс становить 1 на 100, тобто у 1% випадків відкидання нульової гіпотези буде помилкою. Якщо досягнуто рівня значимості не 0,01, а 0,001, то впевненість у тому, що нульова гіпотеза дійсно відкинута правильно, різко зростає (лише 1 шанс на 1000 випадків, що вона все ж вірна). При  $P = 0,05$

впевненість у правильності виведення становить лише 95 випадків з 100, а у 5 випадках можливий неправильний висновок.

Таким чином, якщо отримані дані характеризуються рівнем значущості  $P > 0,05$ , то немає підстав відхилити нульову гіпотезу. Якщо  $P < 0,01$ , то для відкидання нульової гіпотези підстави достатні. Але значно невизначений стан речей, якщо результати аналізу або порівняння задовольняють рівню значущості 0,05, але не задовольняють рівню значущості 0,01. Надійне судження виявляється неможливим. Очевидно, що у таких випадках повинні бути проведені додаткові дослідження, щоб вирішити, чи слід відкинути нульову гіпотезу. Взагалі треба мати на увазі, що збереження нульової гіпотези ще не означає її правильності. Може виявитися все ж, що вона неправильна. Збереження ж нульової гіпотези залишає питання відкритим.

Якщо була отримана різниця між середніми арифметичними двох генеральних сукупностей, то, очевидно, не може стояти питання про статистичну помилку цієї різниці. Ця різниця завжди достовірна, навіть якщо вона і дуже мала. Інша річ, якщо порівнюються дві вибіркові сукупності, наприклад: дві групи морських свинок, яких піддавали впливу хімічних речовин або фізичних факторів, дві групи корів, порівнювані за масою і т.п. У цих випадках різниця між середніми має свою статистичну помилку, з якою її можна порівняти і встановити, чи достовірна ця різниця чи ні. Нульова гіпотеза в даному випадку буде зводитися до того, що дві досліджувані вибіркові сукупності походять з однієї і тієї ж генеральної сукупності і що різниця між їх середніми арифметичними випадкова, тобто лежить у межах помилки вибірковості. **Щоб мати право відкинути нульову гіпотезу, треба довести, що різниця між середніми арифметичними достовірна, тобто задовольняє необхідному рівню значущості.**

### Хід виконання роботи.

**Завдання 1.** Досліджували вплив двох видів комплексних добрив («Гуміноорм» та «Фітомакс») на врожайність картоплі (кг/м<sup>2</sup>) на 20 дослідних ділянках. Були отримані наступні результати:

Гуміноорм	27	20	16	18	22	19	23	21	17	19
Фітомакс	28	19	18	21	24	20	25	27	29	21



Картопля  
(*Solanum tuberosum*)

Визначити, чи якийсь з цих двох добрив збільшує врожайність картоплі у більшій мірі, ніж інше?



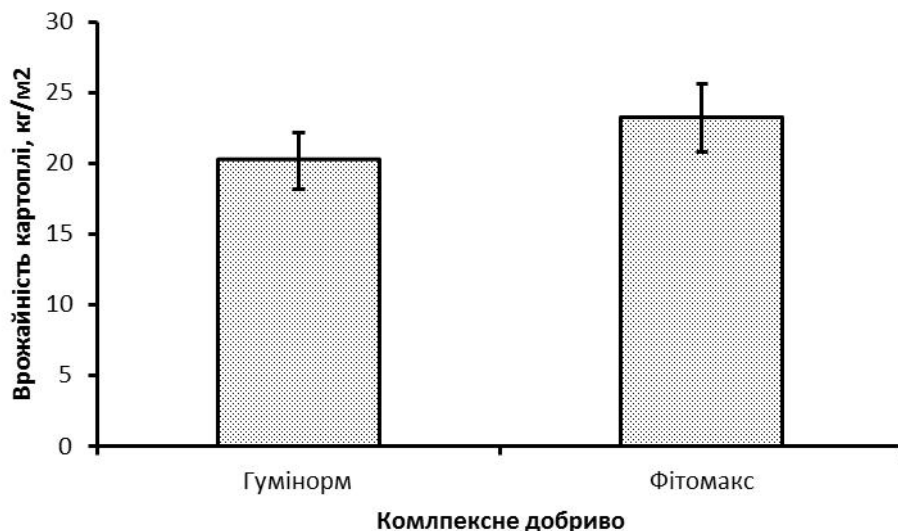
*Порівняння відмінностей між вибірковими сукупностями у MS Excel передбачає наступний загальний алгоритм дій:*

- Визначення типу вибірок (незалежні або залежні).
- Визначення типу даних.
- Формулювання нульової та альтернативної гіпотез.
- Вибір належного статистичного критерію.
- Перевірка нульової гіпотези на прийнятому рівні статистичної значимості за допомогою обраного статистичного критерію
- Інтерпретація результатів та формулювання висновку.

1.1. Створіть таблицю з даними. У комірку B1 уведіть текст «Врожайність картоплі, кг/м<sup>2</sup>» та поєднайте її з коміркою C1, у комірки B2 та C2 уведіть назви добрив: «Гуміноорм» та «Фітомакс» відповідно. У діапазон комірок уведіть дані про врожайність картоплі B3:C12.

1.2. У комірки A13 та A14 уведіть текст «Середнє значення» та «95% ДІ», та обрахуйте відповідні показники у комірках B13:C13 та B14:C14. Виділіть вміст комірок A13:C14 червоним кольором.

1.3. Побудуйте графік, на якому відображено середнє значення врожайності картоплі за умов внесення зазначених добрив та 95% довірчі інтервали за наступним зразком:



Середня врожайність картоплі за умов внесення комплексного добрива «Фітомакс» відповідно до графіку є майже на 15% більшою, ніж за умов використання добрива «Гуміноорм». Але чи дійсно «Фітомакс» збільшує врожайність картоплі?

1.4. Визначте, чи існують **статистично значимі** відмінності у врожайності картоплі за умов внесення комплексних добрив на прийнятому рівні

статистичної значущості. Сформулюємо нульову гіпотезу ( $H_0$ ): «**Вплив кожного з добрив на врожайність картоплі однаковий**». Альтернативна гіпотеза ( $H_1$ ) у такому випадку буде сформульована наступним чином: «**За умов внесення одного з добрив врожайність картоплі більша, ніж за умов внесення іншого**». Рівень статистичної значущості приймемо на рівні 5% ( $p = 0,05$ ), що є достатнім для досліджень подібного роду. Отже, дуже спрощуючи: якщо обраховане значення статистичного критерію буде меншим за 0,05, то нульова гіпотеза буде відхилена і підтверджена альтернативна, якщо більшим за 0,05 – то вірною залишиться нульова гіпотеза.

Для перевірки висунутих гіпотез застосуємо параметричний Т-критерій (відомий як Т-критерій Ст'юдента). Для розрахунку даного критерію у Excel є формула ТТЕСТ, яка має наступні аргументи ТТЕСТ(масив1; масив2; хвости; тип), де масив1 та масив2 – дані вибірок сукупностей для порівняння; хвости – може приймати значення 1 (для одностороннього розподілу) та 2 (для двостороннього розподілу); тип – може приймати значення 1 (для залежних вибірок), 2 (для незалежних вибірок) та 3 (для незалежних вибірок із нерівними дисперсіями).

Отже, у комірку A15 уведіть текст « $p =$ », а у комірку B15 – формулу для обрахування Т-критерію (=ТТЕСТ(B3:B12;C3:C12;2;3)):

11		17	29
12		19	21
13	<b>M</b>	20,2	23,2
14	<b>95% ДІ</b>	2,00	2,44
15	$p =$	=ТТЕСТ(B3:B12;C3:C12;2;3)	
16			

Отриманий результат ( $p = 0,0794$ ) є більшим за прийнятий рівень статистичної значущості (0,05 або 5%). Інакше кажучи, вірогідність, що дані у вибірках не є дійсно різними становить майже  $(0,0794 * 100) = 8\%$ , що більше за обраний нами рівень статистичної значущості.

Отже, ми не можемо підтвердити альтернативну гіпотезу та відхилити нульову. Таким чином висновок, підтверджений результатом статистичного аналізу, буде наступним: «**Вплив кожного з добрив на врожайність картоплі однаковий**», або, інакше кажучи, ми не можемо стверджувати що будь-яке з добрив суттєво збільшує врожайність картоплі, ніж інше; виявлені варіації врожайності – випадкові.

**Завдання 2.** У вісьмох осіб двічі була виміряна частота пульсу: до та після прийому великої кількості їжі:

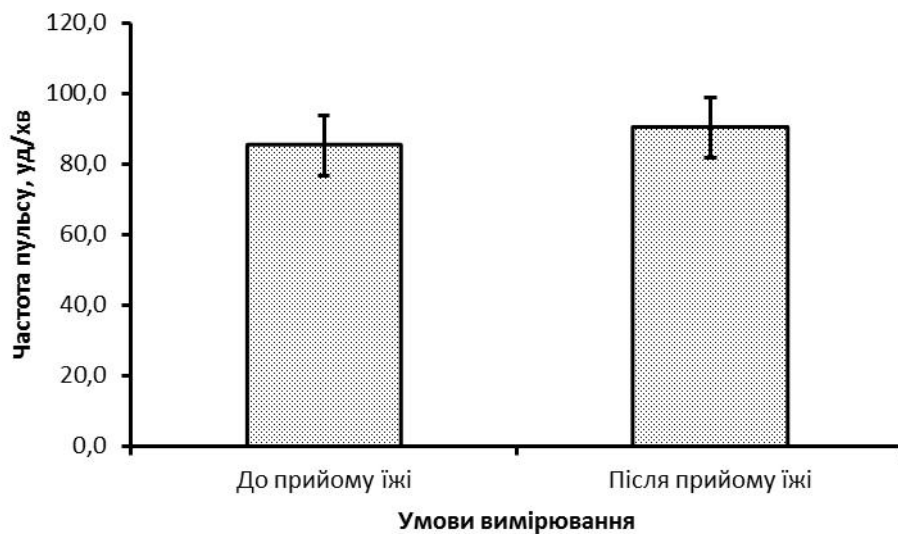
Умови вимірювання	Частота пульсу, уд/хв.							
	До прийому їжі	105	79	79	103	87	74	73
Після прийому їжі	109	83	86	109	90	78	78	89

Встановити, чи впливає прийом великої кількості їжі на частоту пульсу?

2.1. Створіть таблицю з даними. У комірку B1 уведіть текст «Частота пульсу, уд/хв» та поєднайте її з коміркою C1, у комірки B2 та C2 уведіть умови визначення частоти пульсу: «До прийому їжі» та «Після прийому їжі» відповідно. У діапазон комірок B3:C10 уведіть дані про частоту пульсу.

2.2. У комірки A11 та A12 уведіть текст «Середнє значення» та «95% ДІ», та обчисліть відповідні показники у комітках B11:C11 та B12:C12. Виділіть вміст комірок A13:C14 червоним кольором.

2.3. Побудуйте графік, на якому відображено середнє значення частоти пульсу у осіб до і після прийому їжі:



Середнє значення частоти пульсу у осіб після прийому їжі збільшено лише на 5,9%. То ж чи дійсно прийом їжі впливає на частоту пульсу? Чи навпаки, не впливає?

Сформулюємо нульову гіпотезу: «**Прийом їжі не змінює частоту пульсу**», та альтернативну – «**Прийом їжі змінює частоту пульсу (збільшує)**». Рівень статистичної значущості ( $P$ ) приймемо на рівні 0,05 ( $P = 0,05$ ).

Для перевірки висунутих гіпотез знову застосуємо параметричний Т- критерій. Для цього у комірку A13 уведіть текст «р =», а у комірку B13 – формулу для обчислення Т-критерію =ТТЕСТ(B3:B10;C3:C10;2;1). **Зверніть увагу**, що тип тесту (останній аргумент функції) змінився на «1», тому що у даному випадку ми маємо справу із парною (синоніми – пов'язаною, залежною) вибіркою.

Оцініть результат тесту.  $P = 0,00003$ , отже вірогідність, що дані у вибірках не є дійсно різними, складає лише  $0,00003 * 100 = 0,003\%$ , що значно нижче, за обраний нами рівень статистичної значимості ( $p = 0,05$ , або 5%). Таким чином ми можемо відхилити нульову гіпотезу і прийняти альтернативну «**Прийом їжі змінює частоту пульсу (збільшує її)**».

**Завдання 3.** Визначали врожайність двох сортів пшениці (*Triticum turgidum* та *Triticum polonicum*) при культивуванні у засушливих умовах.

Сорт пшениці	Врожайність, ц/г						
<i>Triticum turgidum</i>	15	19	20	18	21	25	31
<i>Triticum polonicum</i>	16	17	22	21	24	23	28



Пшениця  
(*Triticum spp*)

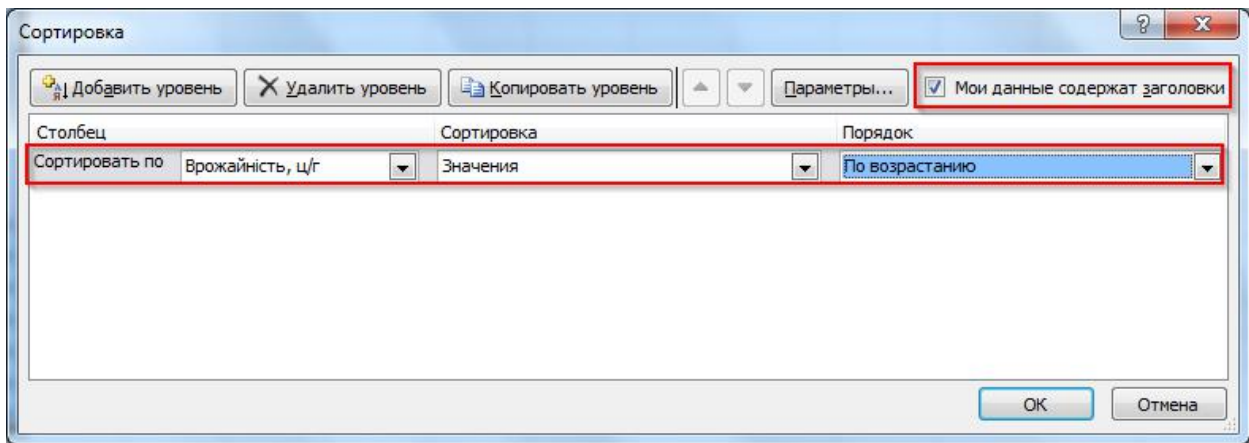
Визначите за допомогою U-критерію Мана-Уїтні, чи відрізняється врожайність вищезазначених сортів пшениці.

3.1 Створіть робочу таблицю із даними. У комірку A1 уведіть текст «Врожайність пшениці», у комірку B2 уведіть текст «Група». Отже, комірки у стовпчику будуть містити значення окремих варіант, а комірки стовпчику B – так звану **групувальну змінну**.

У комірки A2:A8 уведіть дані про врожайність сорту *Triticum turgidum*, а у комірки B2:B8 – назву сорту «*Triticum turgidum*». Відповідні дії виконайте для іншого сорту пшениці («*Triticum polonicum*»), увівши дані у діапазон комірок A9:A15 та B9:B15 відповідно.

	A	B
1	Врожайність, ц/г	Група
2	15	<i>Triticum turgidum</i>
3	19	<i>Triticum turgidum</i>
4	20	<i>Triticum turgidum</i>
5	18	<i>Triticum turgidum</i>
6	21	<i>Triticum turgidum</i>
7	25	<i>Triticum turgidum</i>
8	31	<i>Triticum turgidum</i>
9	16	<i>Triticum polonicum</i>
10	17	<i>Triticum polonicum</i>
11	22	<i>Triticum polonicum</i>

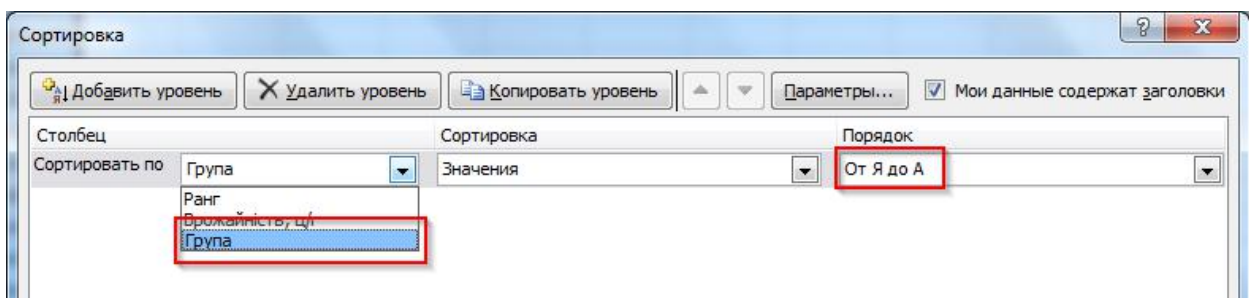
3.2. Наступним кроком є ранжування даних. Для цього спочатку відсортуйте дані про врожайність обох сортів пшениці за зростанням. Виділіть діапазон комірок A1:B15 та на стрічці табличного процесору Excel оберіть вкладку «Данные», панель інструментів «Сортировка и фильтр» кнопка «Сортировка»:



Після того, як ваші дані будуть відсортовані, додайте новий стовпчик А до таблиці; у комірку А1 уведіть текст «Ранг», та за допомогою автозаповнення надайте відсортованим значенням ранг починаючи з 1 та з кроком, який також дорівнює 1. Отже, дані у таблиці повинні виглядати наступним чином:

	A	B	C
1	Ранг	Врожайність, ц/г	Група
2	1	15	Triticum turgidum
3	2	16	Triticum polonicum
4	3	17	Triticum polonicum
5	4	18	Triticum turgidum
6	5	19	Triticum turgidum
7	6	20	Triticum turgidum
8	7	21	Triticum turgidum
9	8	21	Triticum polonicum
10	9	22	Triticum polonicum
11	10	23	Triticum polonicum
12	11	24	Triticum polonicum
13	12	25	Triticum turgidum
14	13	28	Triticum polonicum
15	14	31	Triticum turgidum

3.3. Далі, розподіліть ранжовані дані у дві групи. Для цього знову відсортуйте дані, але за значенням групувальної змінної «Група» (стовпчик С):



Отже, таблиця із даними прийме наступний вид:

	A	B	C
1	Ранг	Врожайність, ц/г	Група
2	1	15	Triticum turgidum
3	4	18	Triticum turgidum
4	5	19	Triticum turgidum
5	6	20	Triticum turgidum
6	7	21	Triticum turgidum
7	12	25	Triticum turgidum
8	14	31	Triticum turgidum
9	2	16	Triticum polonicum
10	3	17	Triticum polonicum
11	8	21	Triticum polonicum
12	9	22	Triticum polonicum
13	10	23	Triticum polonicum
14	11	24	Triticum polonicum
15	13	28	Triticum polonicum

3.4. Наступний крок – підрахунок суми рангів у кожній групі. Для цього, додайте новий рядок до таблиці із даними, клацнувши ПКМ по номеру рядка 9, та оберіть у контекстному меню пункт «Вставити». У комірці A9 обчисліть суму рангів (R1) = СУММ(A2:A8), а у комірці B9 – кількість варіант у групі (N1) = СЧЕТ(A2:A8). Відповідно у комірках A17 та B17 обчисліть суму рангів для іншого сорту пшениці.

3.5. Далі, у комірку A18 уведіть позначення критерію – літеру «U» а у комірці B18 обчисліть значення критерію Мана-Уїтні за відомою формулою:

$$U = N1 * N2 + ( ( N1 * (N1 + 1) ) / 2 ) - R1$$

15	11	24	Triticum polonicum		
16	13	28	Triticum polonicum		
17	56	7			
18	U=	=B9*B17+((B9*(B9+1)/2))-A9			
19					

3.7 Наступний етап – обчислення у комірці A19 середнього арифметичного ( $\mu U$ ) та у комірці A20 – стандартного відхилення ( $\sigma U$ ), за наступними формулами:

$$\mu U = N1 * N2 / 2$$

$$\sigma U = \text{КОРЕНЬ} ( ( N1 * N2 * ( N1 + N2 + 1) ) / 12 )$$

3.8. У комірці обчисліть Z оцінку за наступною формулою:

$$Z \text{ оцінка} = ( U - \mu U ) / \sigma U$$



15	11	24	Triticum polonicum		
16	13	28	Triticum polonicum		
17	56	7			
18	U=	28			
19	$\mu U$ =	24,5			
20	$\sigma U$ =	7,826			
21	Z=	0,447			
22					

3.9. Останній крок при порівнянні груп за допомогою критерію Мана-Уїтні полягає у порівнянні обрахованої Z оцінки з Z критичним для заданого рівня статистичної значимості  $\alpha$ .

Для тесту із однобічним розподілом Z критичне =НОРМСТОБР ( $\alpha$ ), а для тесту із двобічним розподілом Z критичне =НОРМСТОБР ( $\alpha/2$ ).

Якщо потрібен 95% рівень довірчої імовірності, тоді рівень статистичної значимості = 0,05; тоді для тесту з однобічним розподілом, Z критичне буде обчислюватись за формулою: =НОРМСТОБР ( $\alpha$ ) = НОРМСТОБР (0,05), а для тесту з двобічним розподілом, Z критичне обчислюється за формулою: =НОРМСТОБР ( $\alpha/2$ ) = НОРМСТОБР (0,025).

18	U=	28		
19	$\mu U$ =	24,5		
20	$\sigma U$ =	7,826		
21	Z=	0,447		
22	Zкр.1бічн.=	-1,644853627		
23	Zкр.2бічн.=	-1,959963985		
24				



*Результат статистичного тесту є значущим у випадку якщо*

$$| Z \text{ оцінка} | > | Z \text{ критичного} |$$

У нашому випадку  $| Z \text{ оцінка} | < | Z \text{ критичного} |$  (1,96 для двобічного розподілу та 1,64 – для однобічного). Отже, результат не є статистично значимим, і ми не можемо стверджувати із 95% ступенем довірчої вірогідності, що є відмінності між цими двома групами. Таким чином остаточний висновок такий: «**При культивуванні у засушливих умовах врожайність сортів пшениці *Triticum turgidum* та *Triticum polonicum* не відрізняється**».

### Завдання для самостійного виконання.

1. Були отримані наступні дані про масу тіла тушканчиків (*Dipus aegyptius*):

Самці	186	190	165	182	182	182	180	173	157	179	164	146	173
	144	156	156	165	160	160	161	144	153	152	151	173	
Самиці	162	163	190	188	147	146	145	157	162	186	175	147	164
	145	145	155	174	180	148	175	145	144	153	165	141	

Чи відрізняються за вагою самці від самок? Сформулюйте нульову гіпотезу та зробіть статистично обґрунтований висновок.

2. Були отримані наступні дані про температуру тіла тушканчиків (*Dipus aegyptius*):

Самці	37,4	37,5	37,9	37,8	36,8	37,8	37,5
Самиці	37,0	37,8	38,1	37,5	37,7	37,8	37,6

Чи відрізняються за температурою тіла самці від самок? Сформулюйте нульову гіпотезу та зробіть статистично обґрунтований висновок.

### Контрольні питання.

1. Поясніть сутність нульової гіпотези та наведіть приклади.
2. Яким чином оцінюється достовірність різниці між середніми арифметичними? Чи однакові способи оцінки при малих і великих  $n$ ?
3. Наведіть приклад формулювання нульової гіпотези при порівнянні двох середніх арифметичних.
4. Поясніть, у чому перевага попарного порівняння даних? Наведіть приклади з біології.
5. Вкажіть відмінності між графічним та математичним методами порівняння середніх арифметичних.
6. У чому полягає нульова гіпотеза при порівнянні дисперсій?
7. Чи можна вважати достовірним відмінність між середніми арифметичними, якщо фактичне значення  $T$  більше критичного? Якщо воно менше критичного?