Практична робота N 4

Тема: „Статистичні гіпотези. Перевірка параметричних статистичних гіпотез."

Основні теоретичні положення

Будь-які статистичні висновки, здобуті на підставі обробки вибірки, називають ***статистичними гіпотезами***.

Параметричні та непараметричні статистичні гіпотези

Статистичні гіпотези про значення параметрів ознак генеральної сукупності називають ***параметричними.***

Статистичні гіпотези, що висуваються на підставі обробки вибірки про закон розподілу ознаки генеральної сукупності, називаються ***непараметричними***.

Нульова й альтернативна гіпотези

Гіпотезу, що підлягає перевірці, називають ***основною***. Оскільки ця гіпотеза припускає відсутність систематичних розбіжностей (нульові розбіжності) між невідомим параметром генеральної сукупності і величиною, що одержана внаслідок обробки вибірки, то її називають ***нульовою гіпотезою*** і позначають H_0.

Зміст нульової гіпотези записується так:

$$H_0:\ \overline{x}_{\Gamma} = a;$$
$$H_0:\ \sigma_{\Gamma} = 2;$$
$$H_0:\ r_{xy} = 0{,}95.$$

Кожній нульовій гіпотезі можна протиставити кілька ***альтернативних (конкуруючих)*** гіпотез, які позначають символом H_α, що заперечують твердження нульової. Так, наприклад, нульова гіпотеза стверджує: $H_0 : \overline{x}_{\mathscr{T}} = a$, а альтернативна гіпотеза — $H_\alpha : \overline{x}_{\mathscr{T}} > a$, тобто заперечує твердження нульової.

Помилки першого та другого роду. Потужність критерію

Якою б не була малою величина α, потрапляння спостережуваного значення $K^*_{\text{сп}}$ у критичну область $\left(K^*_{\text{сп}} \in \overline{A}\right)$ ніколи не буде подією абсолютно неможливою. Тому не виключається той випадок, коли H_0 буде правильною, а $K^*_{\text{сп}} \in \overline{A}$, а тому нульову гіпотезу буде відхилено.

Отже, при перевірці правильності H_0 можуть бути допущені помилки. Розрізняють при цьому помилки першого і другого роду.

Якщо H_0 є правильною, але її відхиляють на основі її перевірки, то буде допущена помилка першого роду.

Якщо H_0 є неправильною, але її приймають, то в цьому разі буде допущена помилка другого роду.

Гіпотеза H_0	спростовується	приймається
вірна	помилка I типу	правильне рішення
невірна	правильне рішення	помилка II типу

Різницю $\pi = 1 - \beta$ (де β - ймовірність помилки другого роду) називають *імовірністю* обґрунтованого відхилення H_0, або *потужністю* критерію.

Статистичний критерій. Емпіричне значення критерію

Для перевірки правильності висунутої статистичної гіпотези вибирають так званий статистичний критерій, керуючись яким відхиляють або не відхиляють нульову гіпотезу. Статистичний критерій, котрий умовно позначають через K, є випадковою величиною, закон розподілу ймовірностей якої нам заздалегідь відомий. Так, наприклад, для перевірки правильності $H_0 : \overline{X}_{\mathcal{T}} = a$ як статистичний критерій K можна взяти випадкову величину, яку позначають через $K = Z$, що дорівнює

$$Z = \frac{\overline{x}_B - a}{\sigma(\overline{x}_B)},$$

і яка має нормований нормальний закон розподілу ймовірностей. При великих обсягах вибірки ($n > 30$) закони розподілу статистичних критеріїв наближатимуться до нормального.

Спостережуване значення критерію, який позначають через K^*, обчислюють за результатом вибірки.

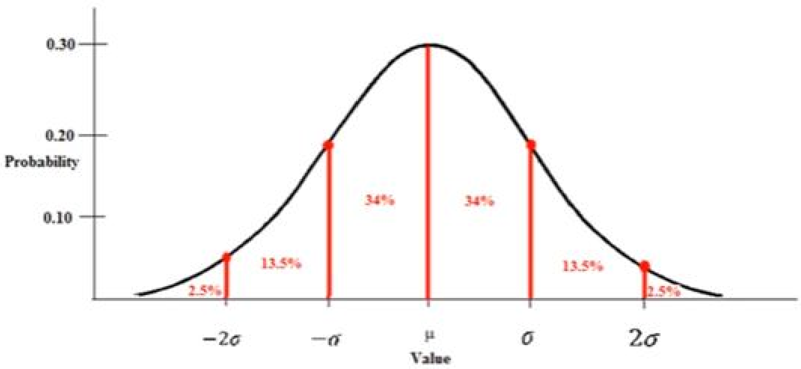
Область прийняття гіпотези. Критична область. Критична точка

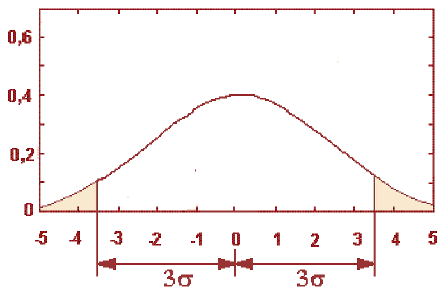
Множину Ω всіх можливих значень статистичного критерію K можна поділити на дві підмножини A і $\overline{A}$, які не перетинаються.

$$(A \bigcup \overline{A} = \Omega, \quad A \bigcap \overline{A} = \varnothing).$$

Сукупність значень статистичного критерію $K \in A$, за яких нульова гіпотеза не відхиляється, називають ***областю прийняття нульової гіпотези***.

Сукупність значень статистичного критерію $K \in \overline{A}$, за яких нульова гіпотеза не приймається, називають ***критичною областю***. Ймовірність попадання статистичного критерію в критичну область – мала величина, яку позначають α і називають ***рівнем значущості.*** Число α зазвичай беруть меншим або рівним за 0,05, адже за правилом трьох сигм ймовірність відхилення випадкової величини від свого середнього значення (математичного сподівання) на відстань, яка за модулем більша за 3σ майже дорівнює нулеві, а ймовірність відхилитись від середнього на відстань більшу за модулем від 2σ (але меншу за 3σ) в середньому складає не більше, ніж 0,05 (див. рис. нижче)





Доповняльну величину $\gamma = 1 - \alpha$ (ймовірність потрапляння у область прийняття нульової гіпотези) називають ***довірчою ймовірністю***.

Отже, A — область прийняття H_0,

$\overline{A}$ — критична область, де H_0 відхиляється.

Точку або кілька точок, що поділяють множину Ω на підмножини A і $\overline{A}$, називають ***критичними*** і позначають через $K_{кр}$.

Існують три види критичних областей:

Якщо при $K \le K_{кр}$ нульова гіпотеза відхиляється, то в цьому разі ми маємо ***лівобічну критичну область (ЛКО)***, яку умовно можна зобразити (рис. 1).

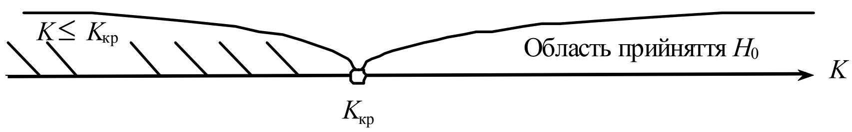


Рис. 1. *Лівобічна критична область (ЛКО)*

Якщо при $K \geq K_{\text{кр}}$ нульова гіпотеза відхиляється, то в цьому разі маємо ***правобічну критичну область (ПКО)*** (рис. 2).

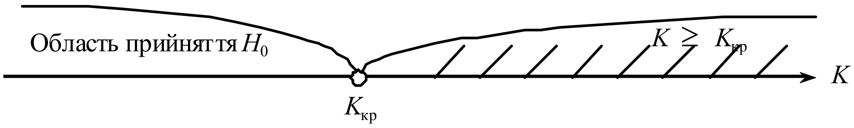


Рис. 2. *Правобічна критична область (ПКО)*

Якщо ж при $K \leq K'_{\text{кр}}$ і при $K \geq K''_{\text{кр}}$ нульова гіпотеза відхиляється, то маємо ***двобічну критичну область (ДКО)*** (рис. 3).

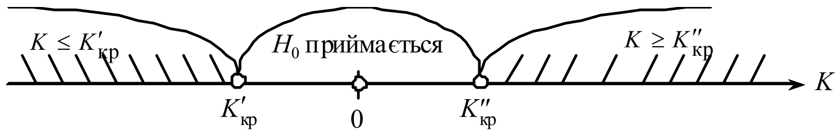


Рис. 3. *Двобічна критична область (ДКО)*

Лівобічна і правобічна області визначаються однією критичною точкою, двобічна критична область — двома критичними точками, симетричними відносно нуля.

Загальний алгоритм перевірки правильності нульової гіпотези

Для перевірки правильності H_0 задається так званий ***рівень значущості*** α, тобто мала ймовірність, якою наперед задаються. Вона може набувати значення $\alpha = 0{,}05;\ 0{,}01;\ 0{,}001$.

В основу перевірки H_0 покладено принцип $P(K \in \overline{A}) = \alpha$, тобто ймовірність того, що статистичний критерій потрапляє в критичну область $\overline{A}$, дорівнює малій імовірності α. Якщо ж виявиться, що $K \in \overline{A}$, а ця подія малоймовірна і все ж відбулася, то немає підстав приймати нульову гіпотезу.

Пропонується такий ***<u>алгоритм перевірки правильності H_0:</u>***

<u>Крок 1.</u> Сформулювати H_0 й одночасно альтернативну гіпотезу H_α.

<u>Крок 2.</u> Вибрати статистичний критерій, який відповідав би сформульованій нульовій гіпотезі.

<u>Крок 3.</u> Залежно від змісту нульової та альтернативної гіпотез будується правобічна, лівобічна або двобічна критична область, а саме:

нехай, наприклад, $H_0 : \overline{x}_\Gamma = a$, тоді, якщо

$H_\alpha : \overline{x}_\Gamma > a$, то вибирається правобічна критична область, якщо

$H_\alpha : \overline{x}_\Gamma < a$, то вибирається лівобічна критична область і коли

$H_\alpha : \overline{x}_\Gamma \neq a$, то вибирається двобічна критична область.

<u>Крок 4.</u> Для побудови критичної області (лівобічної, правобічної чи двобічної) необхідно знайти критичні точки. За вибраним статистичним критерієм та рівнем значущості α знаходяться критичні точки.

<u>Крок 5.</u> За результатами вибірки обчислюється спостережуване значення критерію $K^*_{\text{сп}}$.

<u>Крок 6.</u> Відхиляють чи приймають нульову гіпотезу на підставі таких міркувань:

у разі, коли $K^* \in \overline{A}$, а це є малоймовірною випадковою подією, $P(K^* \in \overline{A}) = \alpha$ і, незважаючи на це, вона відбулася, то в цьому разі H_0 відхиляється:

<u>для лівобічної критичної області</u>

$$P(K^*_{\text{сп}} \leq K_{\text{кр}}) = \alpha \quad (K_{\text{кр}} < 0 \;!!!) \qquad \textbf{(1)}$$

<u>для правобічної критичної області</u>

$$P(K^*_{\text{сп}} \geq K_{\text{кр}}) = \alpha \qquad \textbf{(2)}$$

для двобічної критичної області

$$P(K^*_{\text{сп}} \le K'_{\text{кр}}) + P(K^*_{\text{сп}} \ge K''_{\text{кр}}) = \alpha \quad \text{або} \quad P(K^*_{\text{сп}} \le K'_{\text{кр}}) = P(K^*_{\text{сп}} \ge K''_{\text{кр}}) = \frac{\alpha}{2} \qquad \textbf{(3)}$$

(враховуючи ту обставину, що критичні точки $K'_{\text{кр}} < 0$ і $K''_{\text{кр}} > 0$ симетрично розташовані відносно нуля, тобто $K''_{\text{кр}} = -K'_{\text{кр}}$).

В протилежному випадку H_0 приймається.

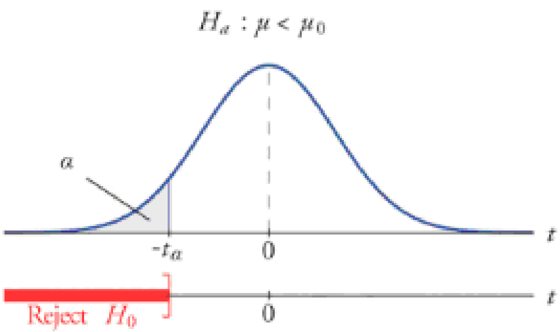
Зауваження: *Критична точка для лівобічної критичної області знаходиться з таблиць, але, оскільки в таблицях усі значення додатні, то у знайденому значенні змінюємо знак на протилежний, адже для лівобічної критичної області обов'язково* $K_{\text{кр}} < 0$.

Таблиця 1

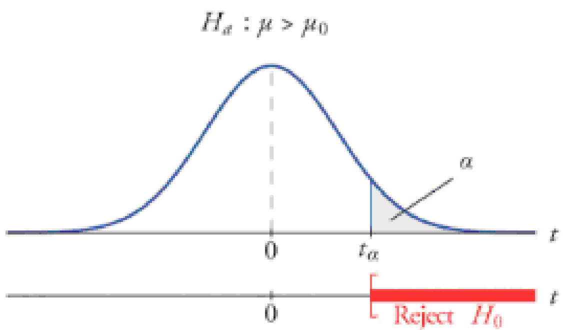
№	H_0	Додаткові умови	Статистика (критерій)	Розподіл статистики критерія, № додатку	H_α	Тип КО	Умови для знаходження критичних точок (меж КО)	Умова прийняття H_0
Перевірка гіпотези про значення генеральної середньої								
1.1	$H_0:\overline{X}_Г=a,$ $(M(X)=a)$	$\sigma_Г^2$ відома	$Z=\dfrac{\sqrt{n}\left(\overline{x}_B-a\right)}{\sigma_Г}$	нормальний (дод. 2)	$H_\alpha:\overline{X}_Г>a$	ПКО	$\Phi(z_{кр})=\dfrac{1-2\alpha}{2}$	$Z^*<Z_{кр}$
					$H_\alpha:\overline{X}_Г<a$	ЛКО		$Z^*>-Z_{кр}^+$
					$H_\alpha:\overline{X}_Г\neq a$	ДКО	$\Phi(z_{кр})=\dfrac{1-\alpha}{2}$	$\lvert Z^*\rvert<Z_{кр}$
1.2	$H_0:\overline{X}_Г=a,$ $(M(X)=a)$	$\sigma_Г^2$ невідома	$t=\dfrac{\sqrt{n}\left(\overline{x}_B-a\right)}{S}$	розподіл Стьюдента $St(t;k)$, $k=n-1$ (дод. 6)	$H_\alpha:\overline{X}_Г>a$	ПКО	$St(t_{кр})=\alpha$	$t^*<t_{кр}$
					$H_\alpha:\overline{X}_Г<a$	ЛКО		$t^*>-t_{кр}^+$
					$H_\alpha:\overline{X}_Г\neq a$	ДКО	$St(t_{кр})=\dfrac{\alpha}{2}$	$\lvert t^*\rvert<t_{кр}$
Перевірка гіпотези про рівність генеральних середніх двох вибірок обсягом n′ та n″								
2.1	$H_0:M(X)=M(Y)$	$n',n''>40$ генеральні $D_X=D_Y$, відомі	$Z=\dfrac{\overline{x}_B-\overline{y}_B}{\sqrt{\dfrac{D_x}{n'}+\dfrac{D_y}{n''}}}$	нормальний (дод. 2)	$H_\alpha:M(X)>M(Y)$	ПКО	$\Phi(z_{кр})=\dfrac{1-2\alpha}{2}$	$Z^*<Z_{кр}$
					$H_\alpha:M(X)<M(Y)$	ЛКО		$Z^*>-Z_{кр}^+$
					$H_\alpha:M(X)\neq M(Y)$	ДКО	$\Phi(z_{кр})=\dfrac{1-\alpha}{2}$	$\lvert Z^*\rvert<Z_{кр}$
2.2	$H_0:M(X)=M(Y)$	$n',n''>40$ генеральні $D_X=D_Y$ невідомі	$Z=\dfrac{\overline{x}_B-\overline{y}_B}{\sqrt{\dfrac{(n'-1)S_x^2+(n''-1)S_y^2}{n'+n''-2}}\sqrt{\dfrac{1}{n'}+\dfrac{1}{n''}}}$	нормальний (дод. 2)	$H_\alpha:M(X)>M(Y)$	ПКО	$\Phi(z_{кр})=\dfrac{1-2\alpha}{2}$	$Z^*<Z_{кр}$
					$H_\alpha:M(X)<M(Y)$	ЛКО		$Z^*>-Z_{кр}^+$
					$H_\alpha:M(X)\neq M(Y)$	ДКО	$\Phi(z_{кр})=\dfrac{1-\alpha}{2}$	$\lvert Z^*\rvert<Z_{кр}$
2.3	$H_0:M(X)=M(Y)$	$n',n''<40$ генеральні $D_X=D_Y$ невідомі	$t=\dfrac{\overline{x}_B-\overline{y}_B}{\sqrt{\dfrac{(n'-1)S_x^2+(n''-1)S_y^2}{n'+n''-2}}\sqrt{\dfrac{1}{n'}+\dfrac{1}{n''}}}$	розподіл Стьюдента $St(t;k)$, $k=n'+n''-2$ (дод. 6)	$H_\alpha:M(X)>M(Y)$	ПКО	$St(t_{кр})=\alpha$	$t^*<t_{кр}$
					$H_\alpha:M(X)<M(Y)$	ЛКО		$t^*>-t_{кр}^+$
					$H_\alpha:M(X)\neq M(Y)$	ДКО	$St(t_{кр})=\dfrac{\alpha}{2}$	$\lvert t^*\rvert<t_{кр}$

№	H_0	Додаткові умови	Статистика (критерій)	Розподіл статистики критерію, № додатку	H_α	Тип КО	Умови для знаходження критичних точок (меж КО)	Умова прийняття H_0
Перевірка гіпотези про рівність генеральних дисперсій двох вибірок								
3.	$H_0:\sigma^2(X)=\sigma^2(Y)$, (де $\sigma^2(X)=D_X$, $\sigma^2(Y)=D_Y$ - дисперсії генеральних сукупностей)	$S_1^2>S_2^2$	$F=\dfrac{S_1^2}{S_2^2}=\dfrac{S_б^2}{S_м^2}$	розподіл Фішера-Снедекора $F(\alpha;k_1,k_2)$, $k_1=n'-1$, $k_2=n''-1$ (дод. 7)	$\sigma_1^2>\sigma_2^2$	ПКО	$P(F>F_{кр}(\alpha;k_1,k_2))=\alpha$	$F^*<F_{кр}$
Перевірка гіпотези про рівність двох часток ознаки								
4.	$H_0:p_1=p_2$ (де p_1, p_2 - частки певної ознаки у першій та другій генеральній сукупностях відповідно)	$n_1,n_2>30$	$Z=\dfrac{p_{S_1}-p_{S_2}}{\sqrt{\overline{p}(1-\overline{p})\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)}}$, де n_1, n_2 - обсяги першої та другої вибірок; m_1, m_2 - кількість «успіхів» у першій та другій вибірці відповідно, $p_{S_i}=\dfrac{m_i}{n_i}$ $(i=1,2)$, $\overline{p}=\dfrac{m_1+m_2}{n_1+n_2}$	нормальний (дод. 2)	$H_\alpha:p_1>p_2$	ПКО	$\Phi(z_{кр})=\dfrac{1-2\alpha}{2}$	$Z^*<Z_{кр}$
					$H_\alpha:p_1<p_2$	ЛКО		$Z^*>-Z^+_{кр}$
					$H_\alpha:p_1\neq p_2$	ДКО	$\Phi(z_{кр})=\dfrac{1-\alpha}{2}$	$\lvert Z^*\rvert<Z_{кр}$

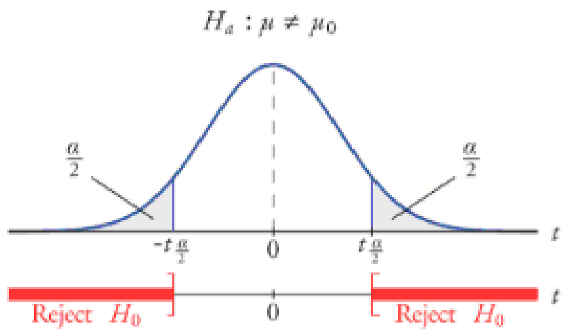
Різновиди критичних областей **(**$H_0:\mu=\mu_0$**)**



Лівобічна критична область (ЛКО)



Правобічна критична область (ПКО)



Двобічна критична область (ДКО)

Задачі для розв'язання в аудиторії

Задача 1. За результатами аналізу темпів зростання продуктивності 10 підприємств галузі було встановлено, що середній темп зростання становить $\overline{x}_B = 2,5\%$. Припускаючи, що темп зростання є випадковою величиною, розподіленою за нормальним законом з $\sigma_Г = 0,3$ на рівні значущості $\alpha = 0,01$ перевірити гіпотезу $H_0 : \overline{x}_Г = 2,8$ проти альтернативної гіпотези $H_\alpha : \overline{x}_Г \neq 2,8$.

Задача 2. В умовах задачі 1 та припущенні, що $\sigma_Г$ невідоме, а за вибіркою обчислено значення $S = 0,4$ виправленого СКВ, на рівні значущості $\alpha = 0,05$ перевірити гіпотезу $H_0 : \overline{x}_Г = 2,6$ проти альтернативної $H_\alpha : \overline{x}_Г = 2,2$.

Задача 3. Вимірювання пульсу 10 хворих, проведені після деякої процедури, і 12 хворих контрольної групи дали наступні результати: $\overline{x}_B = 70$ *(уд/хв)*, $S_x^2 = 9$ (*(уд/хв)*^{2}) для першої групи, $\overline{y}_B = 68$ *(уд/хв)*, $S_y^2 = 4$ (*(уд/хв)*^{2}). На рівні значущості $\alpha = 0,05$ перевірити гіпотезу про рівність середніх значень пульсу хворих, які взяли процедуру, і хворих, які її не брали.

Задача 4. Продуктивність праці деякої бригади з 50 осіб, що працюють у денну зміну, дорівнює 14 одиниць продукції з генеральною дисперсією $\sigma_X^2 = 4$. Продуктивність праці цієї ж бригади в нічну зміну дорівнює 13 одиниць продукції з генеральною дисперсією $\sigma_Y^2 = 3$. Чи можна на рівні значущості ***a*** **=** 0,05 вважати, що нічна робота менш ефективна?

Задача 5. Для дослідження середньої ваги зерна двох сортів були зроблені вибірки по 10 випадково відібраних зерен кожного сорту. Середня вага зерен одного сорту склала 68,2 *мг* при середньоквадратичному відхиленні S_1=0,7 *мг*, а для зерен іншого сорту – 67,0 *мг* зі стандартним відхиленням S_2=5,9 *мг*. У припущенні про нормальні закони розподілу генеральних сукупностей, чи можна з довірчою ймовірністю $\gamma = 0,99$ вважати, що середня вага зерна двох сортів різниться?

Задача 6. Визначалися річні середні витрати електроенергії на комунально-побутові вимоги для одного мешканця у двох містах**.** Результати розрахунків подано двома статистичними розподілами для першого і другого міст***:***

y_i, Вт/м.	700	708	716	724	732	740
n'_i	5	6	9	6	3	1

x_j, Вт/м.	706	710	714	718	722	726	730
n''_j	8	10	12	5	2	2	1

Ознаки ***X*** і **Y** (річні витрати в *кВт/особу*) є незалежними між собою і мають нормальний закон розподілу. При рівні значущості α = 0,001 перевірити правильність нульової гіпотези.

$$H_0 : D_y = D_x,$$

якщо альтернативна гіпотеза

$$H_\alpha : D_y > D_x \ .$$

Задача 7. Аналіз роботи кадрової служби показав, що число тих, хто прийнятий на роботу і тих, кому в цьому було відмовлено, розподілено наступним чином:

Стать	*Прийнято*	*Відмовлено*
Чоловіки	247	113
Жінки	198	110

На рівні значущості ***a***=0,05 перевірити, чи не обмежуються права жінок при прийомі на роботу.

Задача 1. *Задані дві генеральні сукупності, ознаки яких X і Y мають нормальний закон розподілу і при цьому незалежні одна від одної. З кожної генеральної сукупності здійснена вибірка відповідно з обсягами n' і n'': вимірювалась температура в морозильній камері за допомогою двох електронних термометрів. Результати вимірювання подано статистичними розподілами:*

x_i,°C	-7,6	-7,5	-7,2	-6,9	-6,4	-6,1
n_i'	4	8	15	10	7	6

y_i,°C	-8,2	-7,8	-7,3	-7,0	-6,5	-6,2
n_i''	4	6	16	11	8	5

Завдання:

За заданими статистичними розподілами двох вибірок на рівні значущості $\alpha = 0,01$ перевірити правдивість таких нульових гіпотез:

1) $H_0 : D(X) = D(Y)$, *якщо альтернативна гіпотеза*
$H_\alpha : D(X) > D(Y)$ *(або $D(Y) > D(X)$).*

2) $H_0 : M(X) = M(Y)$, *якщо альтернативна гіпотеза*
$H_\alpha : M(X) \neq M(Y)$.

3) $H_0 : M(X) = M(Y)$, *якщо альтернативна гіпотеза*
$H_\alpha : M(X) < M(Y)$.

►**Розв'язання.** Обчислимо точкові статистичні оцінки параметрів генеральних сукупностей (числові характеристики вибірок). Перша вибірка задається дискретним статистичним розподілом

x_i,°C	-7,6	-7,5	-7,2	-6,9	-6,4	-6,1
n_i'	4	8	15	10	7	6

☑ Обсяг вибірки

$$n' = 4 + 8 + 15 + 10 + 7 + 6 = 50 .$$

☑ Обчислимо вибіркове середнє (незміщену ТСО для генеральної середньої $\overline{x}_Г = M(X)$):

$$\overline{x}_B = \frac{\sum_{i=1}^{K} x_i n_i}{n'} = \frac{x_1 n_1 + x_2 n_2 + \ldots + x_K n_K}{n'} =$$

$$= \frac{-7,6 \cdot 4 + (-7,5) \cdot 8 + (-7,2) \cdot 15 + (-6,9) \cdot 10 + (-6,4) \cdot 7 + (-6,1) \cdot 6}{50} = -6,976$$

☑ Обчислимо вибіркову дисперсію, яка є зміщеною ТСО генеральної дисперсії $D(X) = \sigma^2(X)$:

$$D_B(X) = \frac{\sum_{i=1}^{K} x_i^2 n_i}{n} - \left(\overline{x}_B\right)^2 = \frac{x_1^2 n_1 + x_2^2 n_2 + \ldots + x_K^2 n_K}{n'} - \left(\overline{x}_B\right)^2 =$$

$$= \frac{(-7,6)^2 \cdot 4 + (-7,5)^2 \cdot 8 + (-7,2)^2 \cdot 15 + (-6,9)^2 \cdot 10 + (-6,4)^2 \cdot 7 + (-6,1)^2 \cdot 6}{50} - (-6,976)^2 \approx 0,23$$

☑ Знайдемо виправлену дисперсію та виправлене середньоквадратичне відхилення

$$S_x^2 = \frac{n}{n-1} D_B(X) = \frac{50}{50-1} \cdot 0,23 \approx 0,235$$ - виправлена дисперсія;

$$S_x = \sqrt{\frac{n}{n-1} D_B(X)} = \sqrt{S_x^2} = \sqrt{0,235} \approx 0,485$$ - виправлене СКВ.

Друга вибірка задається дискретним статистичним розподілом

y_i,°C	-8,2	-7,8	-7,3	-7,0	-6,5	-6,2
n_i''	4	6	16	11	8	5

☑ Обсяг другої вибірки

$$n'' = 4 + 6 + 16 + 11 + 8 + 5 = 50.$$

☑ Обчислимо вибіркове середнє (незміщену ТСО для генеральної середньої $\overline{y}_Г = M(Y)$) :

$$\overline{y}_B = \frac{\sum_{i=1}^{K} y_i n_i}{n''} = \frac{y_1 n_1 + y_2 n_2 + \ldots + y_K n_K}{n''} =$$

$$= \frac{-8,2 \cdot 4 + (-7,8) \cdot 6 + (-7,3) \cdot 16 + (-7,0) \cdot 11 + (-6,5) \cdot 8 + (-6,2) \cdot 5}{50} = -7,128$$

☑ Обчислимо вибіркову дисперсію, яка є зміщеною ТСО генеральної дисперсії $D(Y) = \sigma^2(Y)$:

$$D_B(Y) = \frac{\sum_{i=1}^{K} y_i^2 n_i}{n''} - \left(\overline{y}_B\right)^2 = \frac{y_1^2 n_1 + y_2^2 n_2 + \ldots + y_K^2 n_K}{n''} - \left(\overline{y}_B\right)^2 =$$

$$= \frac{(-8,2)^2 \cdot 4 + (-7,8)^2 \cdot 6 + (-7,3)^2 \cdot 16 + (-7,0)^2 \cdot 11 + (-6,5)^2 \cdot 8 + (-6,2)^2 \cdot 5}{50} - (-7,128)^2 \approx 0,31$$

☑ Знайдемо виправлену дисперсію та виправлене середньоквадратичне відхилення

$$S_y^2 = \frac{n}{n-1} D_B(Y) = \frac{50}{50-1} \cdot 0,31 \approx 0,316$$ - виправлена дисперсія;

$$S_y = \sqrt{\frac{n}{n-1} D_B(Y)} = \sqrt{S_y^2} = \sqrt{0,316} \approx 0,562$$ - виправлене СКВ.

<u>Приступаємо до перевірки гіпотез.</u>

1) **Перевіримо гіпотезу про рівність генеральних дисперсій**. Дотримуємось алгоритму перевірки гіпотез, описаному вище.

<u>**Крок 1.**</u>*Сформулювати H_0 й одночасно альтернативну гіпотезу H_α.*

$H_0 : D(Y) = D(X)$ - нульова гіпотеза

$H_\alpha : D(Y) > D(X)$ - альтернативна гіпотеза (зупиняємось на такому варіанті, адже $S_y^2 = 0,316 > S_x^2 = 0,235$).

<u>**Крок 2.**</u> *Вибрати статистичний критерій, який відповідав би сформульованій нульовій гіпотезі.*

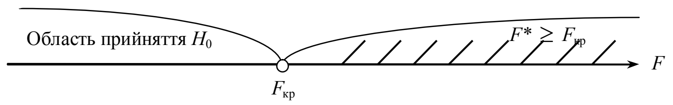
Оскільки перевіряється нульова гіпотеза про рівність двох дисперсій, то користуємось F - критерієм (рядок під номером 3 у таблиці 1). Статистика

$$F = \frac{S_1^2}{S_2^2} = \frac{S_б^2}{S_м^2}$$

має розподіл Фішера-Снедекора $F(\alpha; k_1, k_2)$, $k_1 = n' - 1$, $k_2 = n'' - 1$.

<u>**Крок 3.**</u> *Залежно від змісту нульової та альтернативної гіпотез будується правобічна, лівобічна або двобічна критична область.*

Оскільки альтернативна гіпотеза $H_\alpha : D(Y) > D(X)$, то будуємо <u>*правобічну*</u> критичну область (ПКО):



***<u>Умова прийняття/відхилення нульової гіпотези</u>*:** Якщо $F^* = F_{сп} < F_{кр}$, то приймається H_0 з ймовірністю помилки $\alpha = 0{,}01$, інакше H_0 відхиляється на користь H_α .

<u>Крок 4.</u> *Для побудови критичної області (лівобічної, правобічної чи двобічної) необхідно знайти критичні точки. За вибраним статистичним критерієм та рівнем значущості α з відповідних таблиць знаходяться критичні точки.*

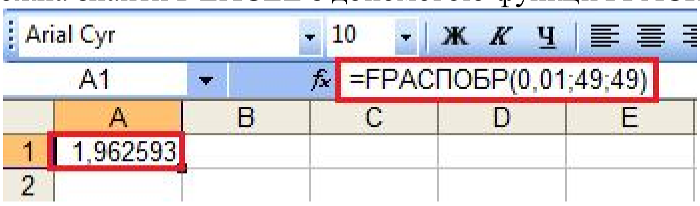
Критичну точку розподілу Фішера-Снедекора

$$F_{кр} = F(\alpha; k_1, k_2) = F(0{,}01; 50-1, 50-1) = F(0{,}01; 49, 49),$$

$k_1 = n'' - 1$, $k_2 = n' - 1$ знаходимо з дод. 7 (Жлуктенко, с. 328)

Рівень значущості 0,01										
k_2 \ k_1	1	2	3	4	5	6	8	12	24	–
1	4052	4999	5403	5625	5764	5859	5981	6106	6234	6366
2	98,5	99,0	99,2	99,3	99,3	99,4	99,3	99,4	99,5	99,5
3	34,1	30,8	29,5	28,7	28,2	27,9	27,5	27,1	26,6	26,1
30	7,6	5,4	4,5	4,0	3,7	3,5	3,2	2,8	2,5	2,0
40	7,3	5,2	4,3	3,8	3,5	3,3	3,0	2,7	2,3	1,8
60	7,1	5,0	4,1	3,7	3,3	3,1	2,8	2,5	2,1	1,6

Але, оскільки в цій таблиці немає саме таких значень $k_1 = k_2 = 49$, то шукане критичне значення буде близьким до середнього значення чотирьох чисел, виділених жовтим на малюнку вище. Точне значення можна знайти в EXCEL з допомогою функції FРАСПОБР:



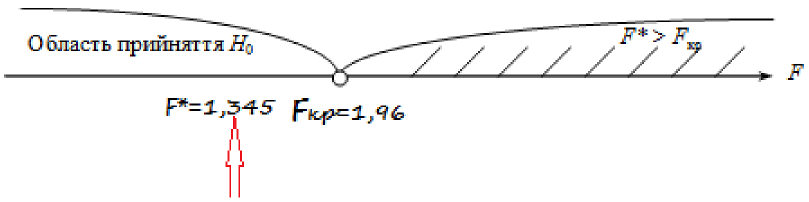
Отже, $F_{кр} = 1{,}96$.

<u>Крок 5.</u> *За результатами вибірки обчислюється спостережуване значення критерію* F^*.

Обчислимо розрахункове значення статистики. Оскільки $S_y^2 = 0{,}316 > S_x^2 = 0{,}235$, то $S_б^2 = S_y^2 = 0{,}316$, $S_м^2 = S_x^2 = 0{,}235$. Тому

$$F^* = F_{сп} = \frac{S_б^2}{S_м^2} = \frac{0{,}316}{0{,}235} = 1{,}345.$$

<u>Крок 6.</u> *Відхиляють чи приймають нульову гіпотезу.*



Оскільки $1{,}345 = F^* < F_{кр} = 1{,}96$, то приймається H_0 з ймовірністю помилки $\alpha = 0{,}01$.

<u>Висновок:</u> *гіпотеза про рівність генеральних дисперсій приймається з ймовірністю помилки* $\alpha = 0{,}01$.

г) Далі, оскільки генеральні дисперсії рівні, то можна перейти до перевірки гіпотези про рівність генеральних середніх (проти альтернативної гіпотези $H_\alpha : M(X) \neq M(Y)$).

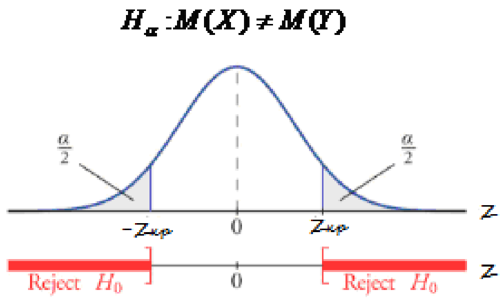
Крок 1. *Сформулювати H_0 й одночасно альтернативну гіпотезу H_α.*
$H_0 : M(X) = M(Y)$ - нульова гіпотеза;
$H_\alpha : M(X) \neq M(Y)$ - альтернативна гіпотеза.

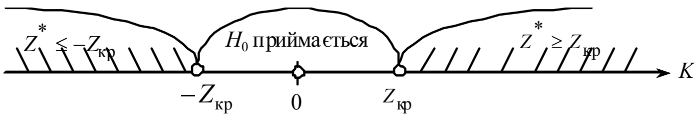
Крок 2. *Вибрати статистичний критерій, який відповідав би сформульованій нульовій гіпотезі.*

Оскільки $n' = n'' = 50 > 40$, генеральні дисперсії рівні $D_X = D_Y$, але невідомі, то перевірка здійснюватиметься з допомогою Z -тесту (рядок з номером 2.2 таблиці 1).

Крок 3. *Залежно від змісту нульової та альтернативної гіпотез будується правобічна, лівобічна або двобічна критична область.*

Оскільки альтернативна гіпотеза $H_\alpha : M(X) \neq M(Y)$, то критична область буде *двобічною* (ДКО)





Умова прийняття/відхилення нульової гіпотези: Якщо $|Z|^* < Z_{кр}$, то приймається H_0 з ймовірністю помилки $\alpha = 0,01$, інакше H_0 відхиляється на користь H_α .

Крок 4. *Для побудови критичної області (лівобічної, правобічної чи двобічної) необхідно знайти критичні точки. За вибраним статистичним критерієм та рівнем значущості α з відповідних таблиць знаходяться критичні точки.*

Критичні точки знаходимо з рівності $\Phi(z_{кр}) = \dfrac{1-\alpha}{2}$ з використанням додатку 2 (Жлуктенко, с. 319). Маємо:

$$\Phi(z_{кр}) = \frac{1-\alpha}{2} = \frac{1-0,01}{2} = 0,495\,.$$

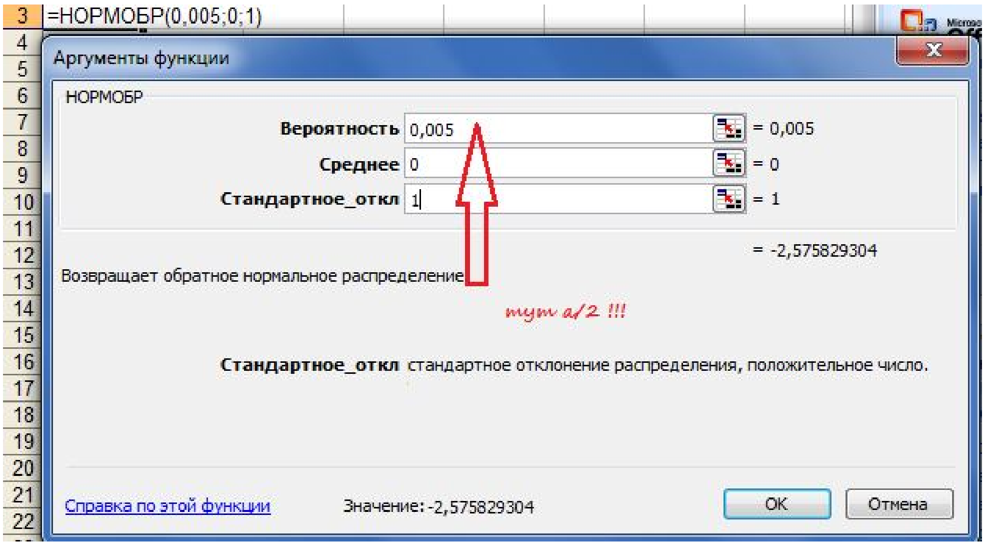
З додатку 2 знаходимо

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2,40	0,4918	2,60	0,4953	2,80	0,4974	3,20	0,49931
2,42	0,4922	2,62	0,4956	2,82	0,4976	3,40	0,49966
2,44	0,4927	2,64	0,4959	2,84	0,4977	3,60	0,49984
2,46	0,4931	2,66	0,4961	2,86	0,4979	3,80	0,499928
2,48	0,4934	2,68	0,4963	2,90	0,4981	4,00	0,499968
2,50	0,4938	2,70	0,4965	2,92	0,4982	5,00	0,499997
2,52	0,4941	2,72	0,4967	2,94	0,4984		
2,54	0,4945	2,74	0,4969	2,96	0,49846		
2,56	0,4948	2,76	0,4971	2,98	0,49856		
2,58	0,4951	2,78	0,4973	3,00	0,49865	$x > 5$	0,5

З таблиць знаходимо $Z^+_{кр} = 2,58$, друга критична точка симетрична першій відносно початку координат:

$$Z^-_{кр} = -Z^+_{кр} = -2,58\,.$$

Від'ємне значення $Z_{кр}^{-}$ критичної точки можна знайти в EXCEL з допомогою функції НОРМОБР:



Крок 5. *За результатами вибірки обчислюється спостережуване значення критерію* Z^*.

Обчислимо спостережуване значення статистики за формулою

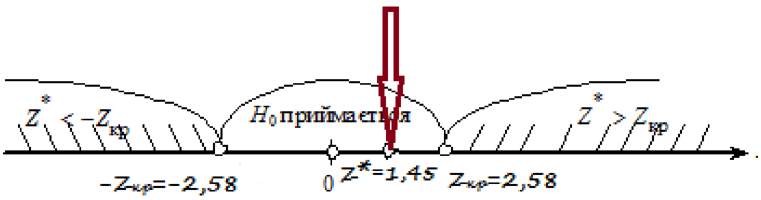
$$Z = \frac{\overline{x}_B - \overline{y}_B}{\sqrt{\frac{(n'-1)S_x^2 + (n''-1)S_y^2}{n'+n''-2}}\sqrt{\frac{1}{n'}+\frac{1}{n''}}}$$

Маємо:

$$Z^* = \frac{-6,976-(-7,128)}{\sqrt{\frac{(50-1)\cdot 0,235+(50-1)\cdot 0,316}{50+50-2}}\sqrt{\frac{1}{50}+\frac{1}{50}}} \approx 1,45$$

Крок 6. *Відхиляють чи приймають нульову гіпотезу*

Оскільки $1,45 = |Z^*| \le Z_{кр} = 2,58$,



то немає підстав відхиляти нульову гіпотезу.

Висновок: *гіпотеза про рівність генеральних середніх приймається з ймовірністю помилки* $\alpha = 0,01$.

3) **Перевіримо гіпотезу про рівність генеральних середніх (проти альтернативної** $H_\alpha : M(X) < M(Y)$**).** *(це можна робити, адже, як ми переконались у першому завданні, генеральні дисперсії рівні).*

Крок 1. *Сформулювати H_0 й одночасно альтернативну гіпотезу H_α.*

$H_0 : M(X) = M(Y)$ - нульова гіпотеза,

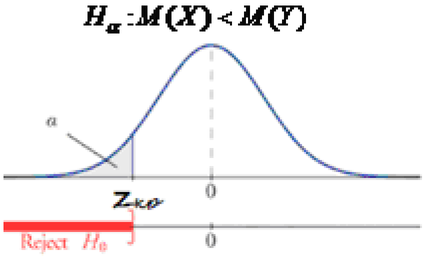
$H_\alpha : M(X) < M(Y)$ - альтернативна гіпотеза.

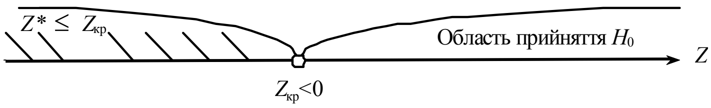
Крок 2. *Вибрати статистичний критерій, який відповідав би сформульованій нульовій гіпотезі.*

Оскільки $n' = n'' = 50 > 40$, генеральні дисперсії рівні $D_X = D_Y$, але невідомі, то перевірка здійснюватиметься з допомогою Z-тесту (рядок з номером 2.2 таблиці 1).

Крок 3. *Залежно від змісту нульової та альтернативної гіпотез будується правобічна, лівобічна або двобічна критична область.*

Оскільки альтернативна гіпотеза $H_\alpha : M(X) < M(Y)$, то критична область буде *лівобічною* (ЛКО).





Умова прийняття/відхилення нульової гіпотези: Якщо $Z^* > -Z^+_{кр}$, де $Z^+_{кр} > 0$ - критичне значення, знайдене з таблиць, то приймається H_0 з ймовірністю помилки $\alpha = 0,01$, інакше H_0 відхиляється на користь H_α .

Крок 4. *Для побудови критичної області (лівобічної, правобічної чи двобічної) необхідно знайти критичні точки. За вибраним статистичним критерієм та рівнем значущості α з відповідних таблиць знаходяться критичні точки.*

Критичні точки знаходимо з рівності $\Phi(z_{кр}) = \dfrac{1-2\alpha}{2}$ з використанням додатку 2 (Жлуктенко, с. 319). Маємо:

$$\Phi(z_{кр}) = \frac{1-2\alpha}{2} = \frac{1-2\cdot 0,01}{2} = 0,49$$

З додатку 2 знаходимо у стовпці «$\Phi(x)$» значення, найближче до 0,49. Таких значень 2 – це 0,4898 та 0,4904; відповідні їм значення – це $x = 2,32$ та $x = 2,34$.

1,28	0,3997	1,57	0,4418	1,86	0,4686	2,30	0,4893
1,29	0,4015	1,58	0,4429	1,87	0,4693	2,32	0,4898
1,30	0,4032	1,59	0,4441	1,88	0,4699	2,34	0,4904
1,31	0,4049	1,60	0,4452	1,89	0,4706	2,36	0,4909

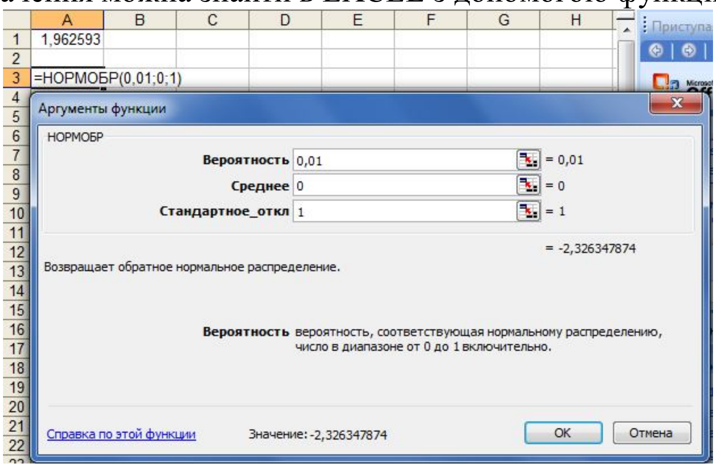
В такому разі шукане критичне значення буде середнім арифметичним знайдених чисел 2,32 та 2,34:

$$Z^+_{кр} = \frac{2,32+2,34}{2} = 2,33 .$$

Знайдена точка була би критичною у разі правобічної критичної області, але, у нас критична область лівобічна, тому критична точка має бути від'ємною але такою ж за модулем, отже, міняємо знак і знаходимо шукане значення

$$Z_{кр} = -Z^+_{кр} = -2,33 .$$

Шукане критичне значення можна знайти в EXCEL з допомогою функції НОРМОБР:



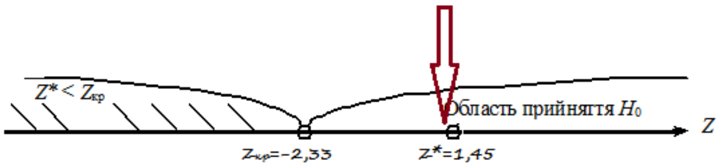
Крок 5. *За результатами вибірки обчислюється спостережуване значення критерію Z^*.*

Спостережуване значення статистики було обчислено у частині 2) цієї задачі:

$$Z^* = \frac{-6,976-(-7,128)}{\sqrt{\frac{(50-1)\cdot 0,235+(50-1)\cdot 0,316}{50+50-2}}\sqrt{\frac{1}{50}+\frac{1}{50}}} \approx 1,45.$$

Крок 6. *Відхиляють чи приймають нульову гіпотезу*

Оскільки $1,45 = Z^* > Z_{кр} = -Z_{кр}^{+} = -2,33$,



то немає підстав відхиляти нульову гіпотезу.

Висновок: *гіпотеза про рівність генеральних середніх приймається з ймовірністю помилки* $\alpha = 0,01$. ■

Задача 2. *При вступі абітурієнти розв'язали 130 із 200 запропонованих задач з алгебри, а з 300 запропонованих задач з геометрії розв'язали 120. Чи можна з довірчою ймовірністю* $\gamma = 0,95$ *стверджувати, що шкільну алгебру було засвоєно краще за геометрію?*

►***Розв'язання.*** Маємо задачу на порівняння часток у двох вибірках. За умовою нам дано 2 вибірки: одна обсягом $n_1 = 200$ (задачі з алгебри), а друга – обсягом $n_2 = 300$ (задачі з геометрії). Число «успіхів» у першій вибірці $m_1 = 130$, а у другій вибірці $m_2 = 120$. Тоді $p_{S_1} = \frac{130}{200} = \frac{13}{20}$, $p_{S_2} = \frac{120}{300} = \frac{2}{5}$. Нехай p_1 та p_2 - реальні частки успіхів у першій та другій генеральних сукупностях відповідно.

Крок 1. *Сформулювати H_0 й одночасно альтернативну гіпотезу H_α.*

$H_0 : p_1 = p_2$ -нульова гіпотеза (рівень засвоєння розділів однаковий);

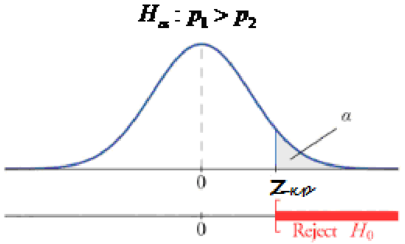
$H_\alpha : p_1 > p_2$ - альтернативна гіпотеза (перший розділ засвоєний краще, тобто частка успіхів там більша).

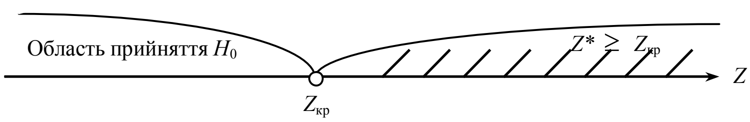
Крок 2. *Вибрати статистичний критерій, який відповідав би сформульованій нульовій гіпотезі.*

Оскільки $n_1 = 200$, $n_2 = 300$>40, то перевірка здійснюватиметься з допомогою Z -тесту (рядок з номером 4 таблиці 1).

Крок 3. *Залежно від змісту нульової та альтернативної гіпотез будується правобічна, лівобічна або двобічна критична область.*

Оскільки альтернативна гіпотеза $H_\alpha : p_1 > p_2$, то критична область буде <u>*правобічною*</u> (ПКО)





Умова прийняття/відхилення нульової гіпотези: Якщо $Z^* < Z_{кр}$, то приймається H_0 з ймовірністю помилки $\alpha = 0,01$, інакше H_0 відхиляється на користь H_α .

Крок 4. *Для побудови критичної області (лівобічної, правобічної чи двобічної) необхідно знайти критичні точки. За вибраним статистичним критерієм та рівнем значущості α з відповідних таблиць знаходяться критичні точки.*

Критична точка знаходиться з умови $\Phi(z_{\text{кр}}) = \dfrac{1-2\alpha}{2}$. Знайдемо спочатку рівень значущості α, оскільки замість α в умові нам дано довірчу ймовірність γ. Рівень значущості знаходимо з умови

$$\alpha + \gamma = 1.$$

За умовою $\gamma = 0,95$, отже,

$$\alpha = 1 - \gamma = 1 - 0,95 = 0,05.$$

Підставимо знайдене значення α в рівність для знаходження критичної точки:

$$\Phi(z_{\text{кр}}) = \frac{1-2\alpha}{2} = \frac{1-2\cdot 0,05}{2} = 0,45.$$

Знайдемо $Z_{кр}$ з таблиці функції Лапласа (дод. 2, Жлуктенко, с. 319)

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1,04	0,3508	1,33	0,4082	1,62	0,4474	1,91	0,4719
1,05	0,3531	1,34	0,4099	1,63	0,4484	1,92	0,4726
1,06	0,3554	1,35	0,4115	1,64	0,4495	1,93	0,4732
1,07	0,3577	1,36	0,4131	1,65	0,4505	1,94	0,4738

У стовпці «$\Phi(x)$» додатку 2 є два значення, найближчих до 0,45: 0,4495 та 0,4505; відповідні їм значення – це $x = 1,64$ та $x = 1,65$.

В такому разі шукане критичне значення буде середнім арифметичним знайдених чисел 1,64 та 1,65:

$$Z_{кр} = \frac{1,64 + 1,65}{2} = 1,645.$$

Крок 5. *За результатами вибірки обчислюється спостережуване значення критерію Z^*.*

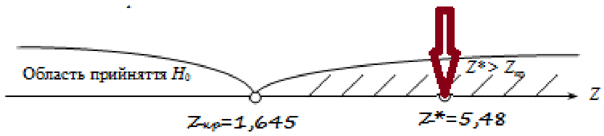
$$Z^* = \frac{p_{S_1} - p_{S_2}}{\sqrt{\overline{p}(1-\overline{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}},$$

де n_1, n_2 - обсяги першої та другої вибірок; m_1, m_2 - кількість «успіхів» у першій та другій вибірці відповідно, $p_{S_i} = \dfrac{m_i}{n_i}$ $(i = 1, 2)$, $\overline{p} = \dfrac{m_1 + m_2}{n_1 + n_2}$. Маємо:

$$Z^* = \frac{p_{S_1} - p_{S_2}}{\sqrt{\dfrac{m_1+m_2}{n_1+n_2}\left(1 - \dfrac{m_1+m_2}{n_1+n_2}\right)\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} = \frac{\dfrac{13}{20} - \dfrac{2}{5}}{\sqrt{\dfrac{130+120}{200+300}\left(1 - \dfrac{130+120}{200+300}\right)\left(\dfrac{1}{200} + \dfrac{1}{300}\right)}} = \frac{\dfrac{1}{4}}{\sqrt{\dfrac{1}{2}\cdot\dfrac{1}{2}\cdot\dfrac{1}{120}}} = \frac{1}{2}\cdot 2\sqrt{30} \approx 5,48$$

Крок 6. *Відхиляють чи приймають нульову гіпотезу*

Оскільки $5,48 = Z^* > Z_{кр} = 1,645$,



то нульова гіпотеза про рівність часток вірно розв'язаних задач в обох генеральних сукупностях відхиляється на користь альтернативної гіпотези про те, що у першій генеральній сукупності частка розв'язаних задач більша.

Висновок: *з імовірністю $\gamma = 0,95$ можна прийняти гіпотезу про те, що перший розділ (алгебра) абітурієнти засвоїли краще.* ■

Задача 1. *Задані дві генеральні сукупності, ознаки яких X і Y мають нормальний закон розподілу і при цьому незалежні одна від одної. З кожної генеральної сукупності здійснена вибірка відповідно з обсягами n' і n''.*

Завдання:

За заданими статистичними розподілами двох вибірок на рівні значущості $\alpha = 0,05$ для непарних варіантів та $\alpha = 0,01$ - для парних перевірити правдивість таких нульових гіпотез:

1) $H_0 : D(X) = D(Y)$, *якщо альтернативна гіпотеза*
   $H_\alpha : D(X) > D(Y)$ *(або $D(Y) > D(X)$).*
2) $H_0 : M(X) = M(Y)$, *якщо альтернативна гіпотеза*
   $H_\alpha : M(X) \neq M(Y)$.
3) $H_0 : M(X) = M(Y)$, *якщо альтернативна гіпотеза*
   $H_\alpha : M(X) < M(Y)$.

Варіант 1 (Байрамов Алі)

Задача 1. Електролампочки на 220 *В* виготовляються двома електроламповими заводами. З партії, виготовленої заводом «Прометей», здійснили вибірку обсягом $n' = 25$, а з партії заводу «Ілліч» – обсягом $n'' = 36$. Першу і другу партії електролампочок перевірили на тривалість роботи. Результати перевірки наведено у вигляді статистичних розподілів такого вигляду:

x_i, діб	48	50	52	54	56
n_i'	2	3	14	5	1

y_i, діб	53	56	59	62	65
n_i''	4	6	10	12	4

Задача 2. Перед експертами поставлено задачу оцінити порівняльну активність електорату Києва та Чернівців на виборах у Верховну Раду. Для цього було здійснено випадкову репрезентативну вибірку у цих містах з тих, хто має право голосу. У Києві з 1500 опитаних взяли участь у виборах 480 людей, а в Чернівцях з 1630 потенційних виборців дільниці відвідали 490 людей. На рівні значущості $\alpha = 0,01$ перевірити гіпотезу про те, що виборці у Києві були активнішими за виборців Чернівців.

Варіант 2 (Беленчук Олексій)

Задача 1. Визначався обсяг валової продукції на підприємствах однієї і тієї самої галузі в двох регіонах України. Результати розрахунків подано двома статистичними розподілами:

x_i, млн. грн.	380	400	420	440	460
n_i'	5	15	30	40	10

y_i, млн.грн.	360	400	440	480	500	540
n_i''	10	20	30	20	15	5

Задача 2. Головний бухгалтер великої корпорації провів обстеження за даними минулого року з метою з'ясування частки некоректних рахунків. З 2000 обраних рахунків у 25 ви-

явилися некоректні проводки. Для зменшення частки помилок він впровадив нову систему. Рік по тому він вирішив перевірити, як працює нова система, і відібрав для перевірки в порядку випадкового відбору 3000 рахунків компанії. Серед них виявилося 30 некоректних. Чи можна стверджувати, що нова система дозволила зменшити частку некоректних проводок в рахунках? Рівень значущості прийняти рівним $\alpha = 0,05$.

Варіант 3 (Березний Ігор)

Задача 1. Вимірювалось споживання масла за одну добу одним мешканцем у двох регіонах країни. Результати вимірювання подано двома статистичними розподілами:

x_i, г	16	19	22	25	28
n_i'	4	6	20	10	5

y_i, г	14,5	20,5	26,5	30,5	36,5
n_i''	6	14	16	6	4

Задача 2. В ході медичного обстеження стояло завдання перевірити алергенність нового препарату. З 250 пацієнтів з одним і тим же захворюванням частина приймала старий загальновідомий препарат X, а частина - новий препарат Y. З тих, хто приймав старий препарат: у 67 осіб була нормальна реакція, а у 33 осіб виявлено алергію. Серед тих, хто приймав новий препарат: у 100 зафіксована нормальна реакція, а у 50 осіб алергія. Перевірити гіпотезу про рівність ймовірностей виникнення алергії при застосуванні препаратів X і Y, коли рівень значущості дорівнює $\alpha = 0,01$.

Варіант 4 (Бужак Андірй)

Задача 1. Досліджувався місячній прибуток робітників двох заводів однієї і тієї самої галузі виробництва. Результати досліджень подано двома статистичними розподілами:

x_i, грн.	2500	2700	2900	3100	3300
n_i'	12	28	40	18	2

y_i, грн.	2400	2700	3000	3300	3600
n_i''	2	6	32	8	2

Задача 2. З 200 проектів, виконаних на IT-компанії «Альфа», цілком задовольняють заданим технічним вимогам 80. При контролі 100 проектів, виконаних на IT-компанії «Бета», заданим технічним вимогам задовольняє 92. Перевірити гіпотезу про рівність можливостей виконання проекту, що повністю відповідає технічним вимогам на цих IT-компаніях з довірчою ймовірністю $\gamma = 0,95$.

Варіант 5 (Бурле Павло)

Задача 1. Пружність вимірювалась на зразках, виготовлених з однієї й тієї ж самої марки сталі. Зразки було вибрані з двох партій. Результати вимірювання подано двома статистичними розподілами:

x_i, ум.од.	36,8	38,8	40,8	42,8	44,8
n_i'	2	4	6	5	3

y_i ,ум.од.	34,2	38,2	42,2	46,2	50,2
n_i''	2	5	10	4	4

Задача 2. Вступний іспит проводився на двох факультетах інституту. На фінансово-кредитному факультеті з 900 абітурієнтів витримали іспит 500 чоловік; а на обліково-статистичному факультеті з 800 абітурієнтів - 408. На рівні значущості α = 0,01 перевірити гіпотезу про відсутність істотних відмінностей у рівні підготовки абітурієнтів двох факультетів.

Варіант 6 (Василевич Павло)

Задача 1. Протягом року вимірювалась продуктивність праці (в тис. грн./працівн.) в двох будівельних фірмах. Результати вимірювання подано статистичними розподілами:

x_i	120	150	180	210	240	270
n_i'	10	20	30	20	15	5

y_i	90	130	170	210	250	290
n_i''	10	20	40	20	5	5

Задача 2. В ході соціологічних досліджень, що стосуються ставлення до релігії, проведених у місті *А* та селі *Б* були отримані наступні результати:

Населений пункт	*Вірю в Бога*	*Переконаний атеїст*
Село Б	63	27
Місто А	46	54

На рівні значущості $\alpha = 0,05$ перевірити гіпотезу про рівність часток тих, хто вірить у Бога у місті *А* та у селі *Б*.

Варіант 7 (Волощук Назарій)

Задача 1. У двох таксопарках виміряли витрати палива за зміну автомобілем. Результати вимірювання показано двома статистичними розподілами:

x_i,л/год	35	35,2	35,4	35,6	35,8	36
n_i'	2	8	10	6	4	3

y_i,л/год	35,4	35,8	36,2	36,6	37	37,7
n_i''	4	5	6	13	6	2

Задача 2. Проводилися випробування нових ліків. В експерименті брали участь 2000 чоловіків та 2500 жінок. У 40 чоловіків і 70 жінок спостерігалися побічні ефекти. Чи можна стверджувати, що побічні ефекти від нових ліків у жінок виникають частіше, ніж у чоловіків? Довірча ймовірність $\gamma = 0,99$.

Варіант 8 (Георгіян Євген)

Задача 1. Визначалися річні середні витрати електроенергії на комунально-побутові потреби для одного мешканця у двох містах. Результати розрахунків подано двома статистичними розподілами для першого і другого міст :

x_i, Вт/р.	700	750	800	900	1000	1200
n_i'	5	6	9	6	3	1

y_i, Вт/р.	710	740	800	840	920	980	1150
n_i''	8	10	12	5	2	2	1

Задача 2. В ході соціологічних досліджень, що стосуються ставлення до використання кредитних продуктів представлених в регіоні, проведених у Васильківському та Бузковому районах були отримані наступні результати:

Район	*Користуюсь*	*Не користуюсь*
Васильківський	874	451
Бузковий	654	678

На рівні значущості $\alpha = 0{,}1$ перевірити гіпотезу про те, що у Васильківському районі частка тих, хто користується кредитними продуктами більша, ніж у Бузковому.

Варіант 9 (Гончаров Олександр)

Задача 1. Визначалась урожайність зеленої маси вівса, зібраного в двох районах області. Результати розрахунків подано двома статистичними розподілами:

x_i, грн./га	58	60	62	64	66
n_i'	2	3	10	4	1

y_i, грн./га	56	60	64	70	74
n_i''	4	6	3	2	1

Задача 2. За деякий період часу в місті Березневе у нічний час було скоєно 125 злочинів, з яких виявилося 40 квартирних крадіжок. За той же проміжок часу в Травневому в нічний час було скоєно 102 злочини, серед яких виявилося 35 квартирних крадіжок. Перевірити гіпотезу про рівність можливостей здійснення квартирних крадіжок вночі в Березневому та Травневому при рівні значущості $\alpha = 0{,}01$.

Варіант 10 (Григорчук В'ячеслав)

Задача 1. Дві кондитерські фабрики виробляють печиво «Марія» у пачках, на яких надруковано: «маса нетто 200 *г*». Вимірювалась маса печива у пачках, що вироблені першою та другою фабрикою відповідно. Результати вимірювання подано двома статистичними розподілами:

x_i, г	194	197	199	200	202	203	205
n_i'	7	11	15	20	12	6	4

y_i, г	195	198	199	200	202	204	206
n_i''	4	12	18	16	15	7	3

Задача 2. З проконтрольованих 100 телевізорів, випущених на першому заводі, цілком задовольняють заданим технічним вимогам 85. При контролі 105 телевізорів, випущених на другому заводі, заданим технічним вимогам задовольняє 98 телевізорів. Перевірити гіпотезу про рівність можливостей випуску придатного телевізора на цих заводах при рівні значущості $\alpha = 0{,}05$.

Варіант 11 (Денис Денис)

Задача 1. За допомогою двох радіодальномірів було здійснено 50 вимірювань однієї і тієї самої відстані. Результати вимірювань подано двома статистичними розподілами:

x_i, м	190	195	200	203	206	210
n_i'	3	7	15	13	8	4

y_i, м	191	195	198	201	204	207	211
n_i''	2	4	10	14	10	7	3

Задача 2. Контрольну роботу з вищої математики виконували студенти двох груп першого курсу. В першій групі студентам було запропоновано 105 задач, з яких студенти вірно розв'язали 60, у другій групі з 140 запропонованих задач студенти вірно розв'язали 69. На рівні значущості $\alpha = 0{,}01$ перевірити гіпотезу про відсутність істотних відмінностей у засвоєнні матеріалу студентами обох груп (добре засвоїли матеріал рівні частки студентів обох груп).

Варіант 12 (Дісар Іван)

Задача 1. У двох партіях містяться однотипні шарикопідшипники, виготовлені двома заводами. Вимірювання їх діаметрів дали результати, які наведено у вигляді двох статистичних розподілів:

x_i, мм	6,64	6,7	6,74	6,78	6,82
n_i'	2	4	8	6	4

y_i, мм	6,58	6,6	6,8	7,0	7,2
n_i''	6	8	10	4	2

Задача 2. Порівнювали ефективність двох різних методик лікування. При застосуванні першої не вилікувались 8 хворих з 61, а із застосуванням другої – 10 з 67. На рівні значущості $\alpha = 0{,}05$ перевірити гіпотезу про відсутність істотних відмінностей у методах лікування (про рівність часток (не)одужавших пацієнтів).

Варіант 13 (Дручук Роман)

Задача 1. Вимірювання барометром атмосферного тиску протягом 100 діб у двох містах України подано двома статистичними розподілами:

x_i, мм.рт.ст	744	746	748	750	752	754
n_i'	10	20	30	20	15	5
y_i, мм.рт.ст	740	744	748	752	756	760
n_i''	5	15	30	35	10	5

Задача 2. Проводилися випробування нових ліків. В експерименті брали участь 2000 чоловіків і 2500 жінок. У 40 чоловіків і 70 жінок спостерігалися побічні ефекти. Чи можна стверджувати, що побічні ефекти від нових ліків у жінок виникають частіше, ніж у чоловіків? Довірча ймовірність $\gamma = 0{,}99$.

Варіант 14 (Дубець Василь)

Задача 1. З двох партій монет вартістю 5 коп. було вибрано 50 і 60 штук, які зважували на терезах. Результати цих зважувань подано у вигляді двох статистичних розподілів:

x_i, мг	9,4	9,6	9,6	9,8	10,2
n_i'	5	15	20	8	2

y_i, мг	9,33	9,63	9,93	10,23	10,53
n_i''	8	12	26	10	4

Задача 2. В ході соціологічних досліджень серед студентів ЧНУ було виявлено поділ студентів на дві групи - «автомобілісти» і «велосипедисти». На основі частоти появи цих ознак в обстежуваних групах була складена наступна таблиця:

Факультет/інститут	*авто*	*велосипед*
Юридичний	100	12
ІФТКН	50	55

З довірчою ймовірністю $\gamma = 0{,}99$ перевірити гіпотезу про те, що «автомобілістів» серед студентів юридичного факультету більше, ніж серед студентів ІФТКН.

Варіант 15 (Дуплава Олександр)

Задача 1. Досліджується ефективність двох видів добрів. Для цього вимірювалась урожайність культури на однорідних ділянках, які були оброблені добривами першого та другого видів відповідно. Результати вимірювання подано статистичними розподілами:

x_i, ц/га	5,6	6,2	6,4	7,7	8,0	8,4	8,6
n_i'	1	2	4	9	5	3	1

y_i, ц/га	5,0	5,5	6,0	6,4	7,0	7,4	8,0
n_i''	1	1	2	3	9	5	4

Задача 2. Контрольну роботу з математики за індивідуальними варіантами виконували студенти двох груп першого курсу. У першій групі було запропоновано 105 задач, з яких правильно розв'язано 60, в другій групі зі 140 запропонованих завдань вірно вирішено 69. Перевірити гіпотезу про те, що студенти першої групи засвоїли матеріал краще, ніж студенти другої. Довірча ймовірність $\gamma = 0{,}98$.

Варіант 16 (Жупник Евеліна)

Задача 1. Кров'яний тиск був виміряний (в умовних одиницях) у 20 осіб віком 40 років із одного району міста та у 18 осіб того самого віку з іншого району міста. Результати вимірювання надано двома статистичними розподілами:

x_i	114	116	118	120	122	124
n_i'	2	4	6	5	2	1

y_i	115	118	121	124	127	130
n_i''	1	3	6	4	3	1

Задача 2. При опитуванні викладачів та студентів ІФТКН отримали дані про те, носять чи не носять вони джинси. Дані опитування в таблиці:

Користування джинсами	*Носять*	*Не носять*
Викладачі	120	80
Студенти	160	40

З довірчою ймовірністю $\gamma = 0,95$ перевірити гіпотезу про рівність часток тих, хто носить джинси, серед викладачів та студентів ІФТКН.