Аналіз даних та статистичне виведення

Тиждень

5

тестування гіпотез

тестування гіпотез

Нехай дослідження вибірки з 57 студентів університету показало, що їх середній пульс становить 70 ударів за хвилину, середньоквадратичне відхилення 9.94.
Стандартне середнє значення становить 72 удари.
Дослідник хоче визначити чи відрізняються результати вибірки від загальних результатів

H_a: $\mu \neq 72$, $\mu < 72$, $\mu > 72$,

H_0: $\mu = 72$

тестування гіпотез для середнього значення

відключення електроенергії та народжуваність

9 листопада 1965 року в енергосистемі США сталася аварія. 30 мільйонів людей протягом 13 годин перебували без світла.
Через 9 місяців (10 серпня 1966) в NY Times опубліковане дослідження, яке стверджувало що значно зросла

кількість новонароджених
(Нью Йорк, серпень 1966)

Пн	Вт	Ср	Чт	Пт	Сб	Нд
452	470	431	448	467	377	344
449	440	457	471	463	405	377
453	499	461	442	444	415	356
470	519	443	449	418	394	399
451	468	432				

Перші два тижні

$\overline{X} = 432.21$
$S = 40.48$
$n = 14$

Ми хочемо дізнатись, чи відрізняється це значення від звичайного рівня народжуваності в Нью Йорку (середня кількість новонароджених 430 на добу)

відключення електроенергії та народжуваність

Нульова гіпотеза: Відключення електроенергії у листопаді 1965 року впливу на кількість новонароджених не має, тобто таке ж саме як і в інші місяці.
H_0: $\mu = 430$ (звична кількість новонароджених)

Альтернативна гіпотеза: Відключення електроенергії у листопаді 1965 має вплив на кількість новонароджених
H_a: $\mu \neq 430$

Це двостороння альтернатива, що означає що рівень народжуваності відрізняється. Можемо також розглядати односторонню альтернативу, наприклад H_a: $\mu > 430$

тестова статистика

Тестова статистика вимірює різницю між даними отриманої вибірки та нульовою гіпотезою.

Фактично тестова статистика відповідає на питання: ”Яка відстань у середньоквадратичних відхиленнях між середнім значенням отриманої вибірки та середнім значенням згідно нульової гіпотези”

Для рівня народжуваності в Нью Йорку середнє значення вибірки становить 432.21, а середнє значення згідно нульової гіпотези 430.

Щоб обрахувати тестову статистику нам потрібно знати середньоквадратичне відхилення для вибірки.

тестова статистика для середнього значення вибірки

Середньоквадратичне відхилення вибіркового розподілу $\frac{\sigma}{\sqrt{n}}$, наша тестова статистика Z буде обчислюватись за формулою:

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

Z – кількість середньоквадратичних відхилень між середнім значенням вибірки та середнього значення згідно нульової гіпотези

Для обрахування тестової статистики Z потрібно знати середньоквадратичне відхилення генеральної сукупності σ

Ми вже знаємо, що в цьому випадку можна використовувати t-розподіл.
Тобто потрібно обрахувати

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$

тестова статистика для рівня народжуваності

Для нашого прикладу середнє значення вибірки 432.21, середнє значення згідно нульової гіпотези середнє значення 430, середньоквадратичне відхилення вибірки 40.48. Розмір вибірки: 14.

Тестова статистика:

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{432.21 - 430}{40.48/\sqrt{14}} = 0.204$$

Тобто, середнє значення отриманої вибірки знаходиться на відстані 0.204 середньоквадратичних відхилень від середньоквадратичного значення нульової гіпотези.

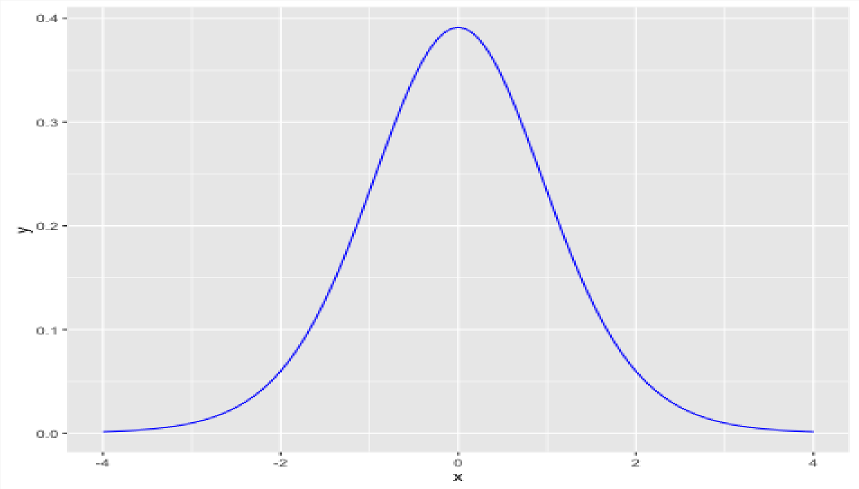
Чи ця різниця є статистично значимою? Чи можливо ми отримали це значення випадково?

значення ймовірності (p-value)

За припущення, що нульова гіпотеза правдива, p-value відповідає на питання “яка ймовірність отримати значення більш екстремальне ніж наще спостережуване середнє значення”

Чим менше p-value тим більш нереалістичною є нульова гіпотеза
Для нашого прикладу t =0.204.

За припущення що середнє значення нашої генеральної сукупності 430, яка ймовірність отримати вибірку, з t статистикою 0.204 або більш екстремальне?



p.val <- 2*pt(0.204, df=13, lower.tail = FALSE)
p.val = 0.841

статистична значущість

Ми порівнюємо отримане p-value з фіксованим значенням, яке є вирішальним наскільки ми маємо доказів, щоб відкинути нульову гіпотезу.

Це вирішальне значення має назву рівень значущості та позначаєтся α.

Загально прийнятим α рівнем є $\alpha = 0.05$. Це означає, що докази які ми отримали проти нульової гіпотези настільки сильні, що можуть бути отримані в результаті випадкового збігу не більше ніж в 5% (якщо нульова гіпотеза справедлива).

Якщо p-value менше ніж α,
говорять, що різниця **статистично значима для рівня α**

обчислення p-value

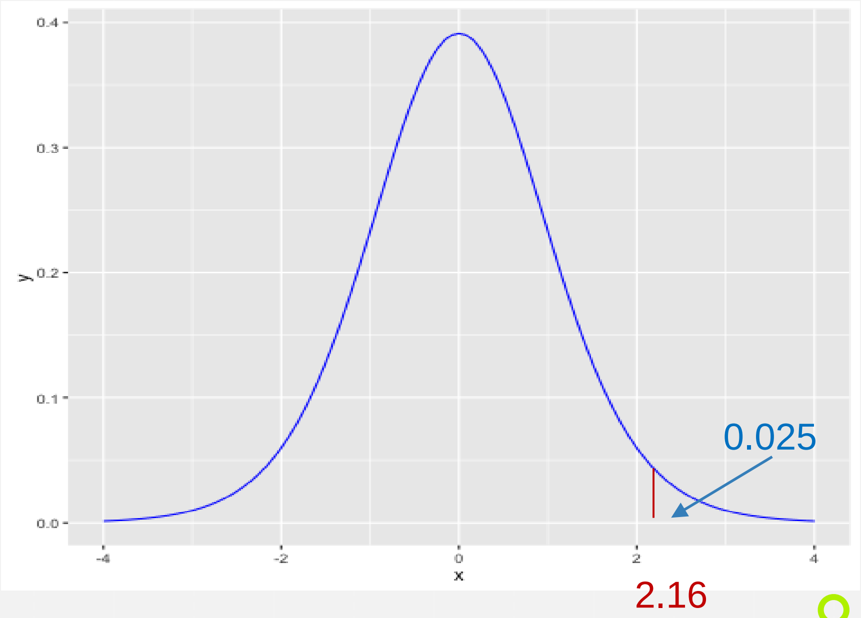
$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{432.21 - 430}{40.48/\sqrt{14}} = 0.204$$

$H_0\colon \mu = 430, H_a\colon \mu \neq 430$

$\alpha = 0.05$

df = n - 1 = 13

p = 0.814, p > α



трактування результатів

Для кількості новонароджених у Нью Йорку p-значення = 0.814, що не дозволяє відкинути нульову гіпотезу для рівня значущості $\alpha = 0.05$.

Іншими словами, можна сказати що різниця між нульовою гіпотезою та даними вибірки не є статистично значущою.

Тобто наші дані не підтверджують гіпотезу, що рівень народжуваності у перші два тижні серпня 1966 відрізняється від звичного. Тобто, відсутність електроенергії не мало впливу на рівень народжуваності.

тестування гіпотез для пропорції

тестування гіпотез для пропорцій

Чи відрізняється відсоток новонароджених хлопчиків від 50%? У вибірці 200 новонароджених, з них 96 хлопчики.

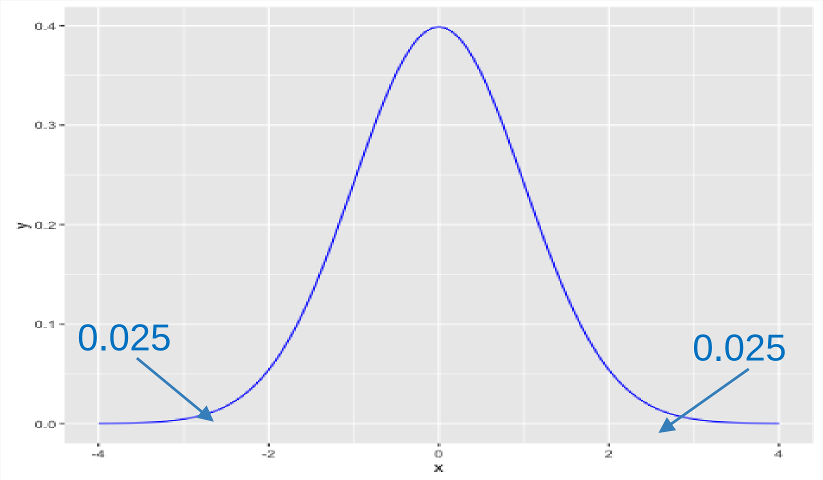
H_0: p = 0.5, H_a: p ≠ 0.5

Рівень довіри: $\alpha = 0.05$

Тестова статистика: $z = \frac{\bar{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$

$\bar{p}$ = 96/200 = 0.48, p= 0.5, n=200

$$z = \frac{\bar{p} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.48 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{200}}} = -0.566$$



p.val = 2*pnorm(-0.566),
p.val = 0.571

трактування результатів

Для кількості новонароджених p-значення = 0.571, що не дозволяє відкинути нульову гіпотезу для рівня значущості $\alpha = 0.05$

Іншими словами, можна сказати що різниця між нульовою гіпотезою та даними вибірки не є статистично значущою.

Тобто наші дані не підтверджують гіпотезу, що відсоток хлопчиків серед новонароджених відрізняється від 50%.

тестування гіпотез покроковий план

покроковий план тестування гіпотез

Середнє чи пропорція?

Сформулювати гіпотези

Визначити рівень значущості

Обрахувати тест статистику

Зобразити вибірковий розподіл

Знайти p-value та розташування статистики

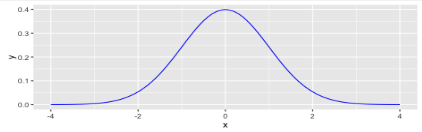
Зробити висновок

середнє

H_0: $\mu = 430, H_a$: $\mu > 430$

$\alpha = 0.05$

$\mathrm{t} = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}}$ $\quad \mathrm{t} = 0.204$



p value = 0.425

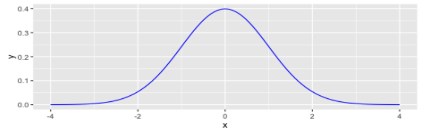
не можемо відкинути H_0

пропорція

H_0: p = 0.5, H_a: p ≠ 0.5

$\alpha = 0.05$

$\mathrm{z} = \dfrac{\bar{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$ $\quad \mathrm{z} = -0.566$



p value = 0.571

не можемо відкинути H_0

тестування гіпотез та довірчі інтервали

зв'язок між тестом на значущість та довірчим інтервалом

P-значення для двостороннього тесту ≤ 0.05

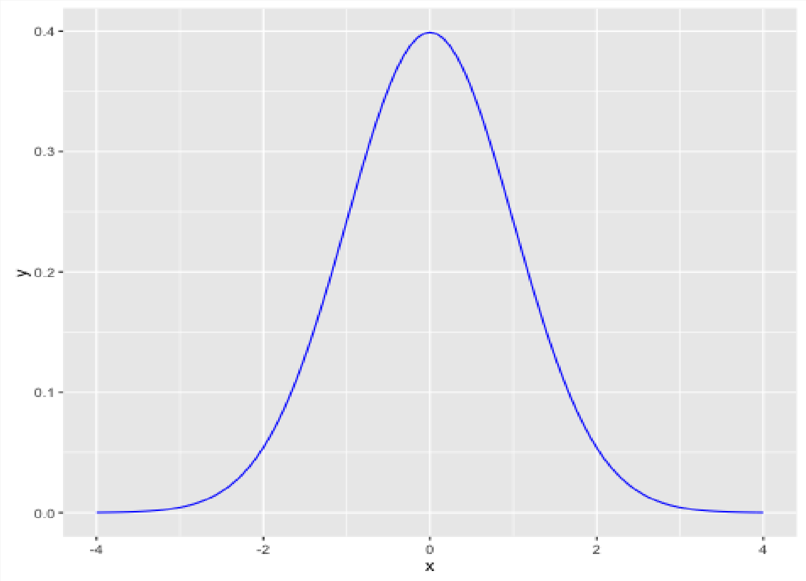
⇕

95% довірчий інтервал не містить H_0 значення

P-значення для двостороннього тесту > 0.05

⇕

95% довірчий інтервал буде містити H_0 значення



зв’язок між тестом на значущість та довірчим інтервалом

Чи відрізняється відсоток новонароджених хлопчиків від 50%? У вибірці 200 новонароджених, з них 96 хлопчики.

Тест на значущість:

H_0: p = 0.5, H_a: p ≠ 0.5

Рівень довіри: $\alpha = 0.05$

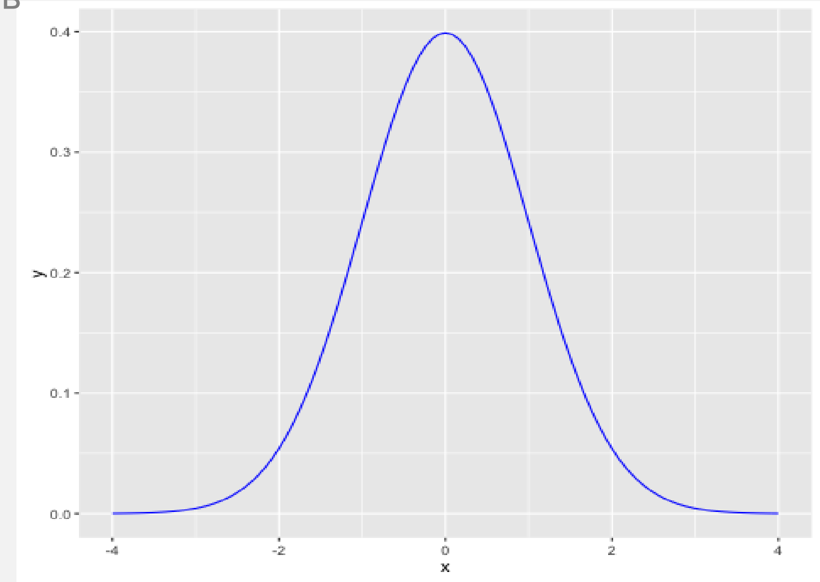
P value 0.571

Довірчий інтервал:

$p \pm Z_{95\%} * se(\bar{p})$, де $se(\bar{p}) = \sqrt{\frac{p(1-p)}{n}}$

$0.48 \pm 1.96 * \sqrt{\frac{0.48(1-0.48)}{200}}$

[0.41, 0.55]



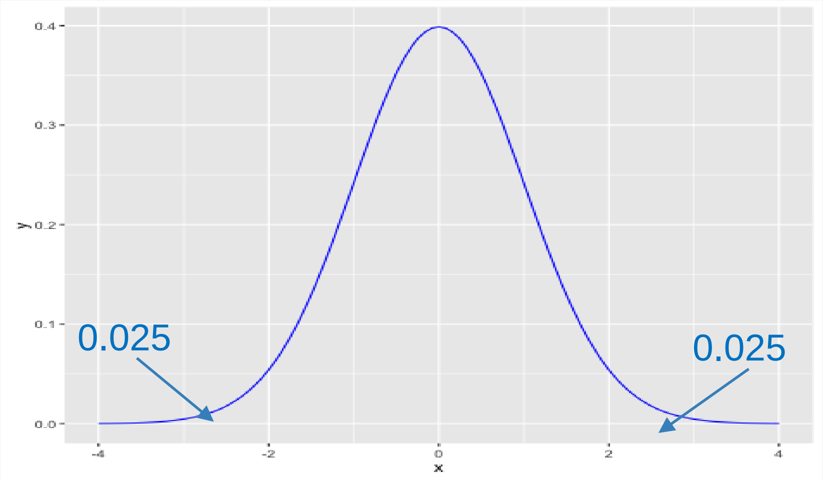
зв’язок між тестом на значущість та довірчим інтервалом

Чи відрізняється відсоток новонароджених хлопчиків від 60%? У вибірці 200 новонароджених, з них 96 хлопчики.

H_0: p = 0.6, H_a: p ≠ 0.6
Рівень довіри: $\alpha = 0.05$
Тестова статистика: $z = \frac{\bar{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$

$\bar{p}$ = 96/200 = 0.48, p= 0.6, n=200

$$z = \frac{\bar{p} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.48 - 0.6}{\sqrt{\frac{0.6(1-0.6)}{200}}} = -3.464$$



p.val = 2*pnorm(-0.566),
p.val = 0.0005

ПОМИЛКИ

помилки І та ІІ типу

	не відхиляти	відхилити
H_0 правдива	v	x (помилка І типу, α)
H_0 хибна	x(помилка ІІ типу, β)	v (потужність, 1-β)