

Fefelov_Final Project-5

Розрахунок довірчого інтервалу для різниці конверсії Cart→Purchase

Executive Summary

Analysis Type: 95% Confidence Interval for Cart→Purchase Conversion Rate Difference

Key Results

Metric	Value	Interpretation
Control Conversion	40.21%	Baseline performance
Test Conversion	59.13%	Improved performance
Point Estimate	+18.91 pp	Best estimate of difference
95% CI	[18.14%, 19.69%]	Range of true difference
Statistical Significance	✓ YES	CI does not contain zero
Practical Significance	✓ HIGH	+47% relative improvement

What This Means

We are 95% confident that the test campaign improves cart-to-purchase conversion by 18-20 percentage points.

Even in the **worst case** (lower bound), test campaign is 45% better at converting cart additions to purchases.

Business Impact (Actual Test Results)

- Additional Purchases: 476 (+3.14%)
- Additional Spend: \$10,074
- Marginal ROAS: 2.36x (at \$50 AOV)
- Verdict: Profitable and statistically robust

1. Вступ

1.1. Мета аналізу

Розрахувати 95% довірчий інтервал (Confidence Interval, CI) для різниці у конверсії Add to Cart → Purchase між тестовою та контрольною групами.

1.2. Чому ця метрика?

Cart-to-Purchase Rate — це ключова метрика, що показує:

- Якість залученого трафіку
- Intent to purchase (намір купити)
- Ефективність процесу checkout
- Реальну цінність користувачів, які додали товар в кошик

Це один з найважливіших індикаторів того, чи новий підхід до реклами приводить більш "готових до покупки" користувачів.

2. Метод розрахунку

2.1. Формула для різниці двох пропорцій

Для порівняння двох незалежних пропорцій (конверсій) використовуємо:

Різниця пропорцій:

$$\Delta p = p_{\text{test}} - p_{\text{control}}$$

Стандартна похибка (Standard Error) різниці:

$$SE(\Delta p) = \sqrt{[p_{\text{control}}(1-p_{\text{control}})/n_{\text{control}} + p_{\text{test}}(1-p_{\text{test}})/n_{\text{test}}]}$$

95% Довірчий інтервал:

$$CI_{95\%} = \Delta p \pm Z_{\{\alpha/2\}} \times SE(\Delta p)$$


Де:

- $Z_{\{\alpha/2\}} = 1.96$ для 95% рівня довіри (two-tailed)


- **p_control** = конверсія в контрольній групі
- **p_test** = конверсія в тестовій групі
- **n_control** = кількість Add to Cart в контрольній групі
- **n_test** = кількість Add to Cart в тестовій групі

2.2. Інтерпретація CI

Якщо CI не містить нуль (0):

-  Різниця статистично значуща на рівні 95%
- Можемо з 95% впевненістю сказати, що є реальна різниця між групами

Якщо CI містить нуль (0):

-  Різниця НЕ статистично значуща
- Не можемо виключити, що різниця виникла випадково

3. Вхідні дані

3.1. Контрольна група

Метрика	Значення
Кількість Add to Cart	37,700
Кількість Purchases	15,161
Конверсія (p_control)	$15,161 / 37,700 = 0.4021$ (40.21%)

3.2. Тестова група

Метрика	Значення
Кількість Add to Cart	26,446
Кількість Purchases	15,637
Конверсія (p_test)	$15,637 / 26,446 = 0.5913$ (59.13%)

3.3. Перевірка даних

Умови для застосування методу:

- ☒ Незалежні групи (Control та Test — окремі кампанії)
- ☒ Великі вибірки ($n > 30$ для обох груп)
- ☒ Випадкова вибірка (рандомізація через ad platform)
- ☒ Біноміальна природа даних (покупка/не покупка)

Перевірка правила п'яти (Rule of Five):

- Control: $\min(n \times p, n \times (1-p)) = \min(15,161, 22,539) = 15,161$ ☒ > 5
- Test: $\min(n \times p, n \times (1-p)) = \min(15,637, 10,809) = 10,809$ ☒ > 5

Висновок: Всі умови виконані, можна застосовувати нормальне наближення.

4. Покрокові розрахунки

Крок 1: Розрахунок конверсій

Control:

$$p_{\text{control}} = 15,161 / 37,700 = 0.402148 \text{ (40.21\%)}$$

Test:

$$p_{\text{test}} = 15,637 / 26,446 = 0.591280 \text{ (59.13\%)}$$

Крок 2: Різниця пропорцій (Point Estimate)

$$\Delta p = p_{\text{test}} - p_{\text{control}}$$

$$\Delta p = 0.591280 - 0.402148$$

$$\Delta p = 0.189132$$

Інтерпретація: Тестова група має на **18.91 percentage points (pp)** вищу конверсію.

У відносних термінах:

$$\text{Relative improvement} = (p_{\text{test}} / p_{\text{control}} - 1) \times 100\%$$

$$= (0.5913 / 0.4021 - 1) \times 100\%$$

$$= 47.03\%$$

Це означає +47% покращення конверсії у тестовій групі!

Крок 3: Стандартна похибка (Standard Error)

Формула:

$$SE(\Delta p) = \sqrt{[p_c(1-p_c)/n_c + p_t(1-p_t)/n_t]}$$

Підставляємо значення:

Для Control:

$$\begin{aligned} \text{variance_control} &= p_c \times (1 - p_c) / n_c \\ &= 0.402148 \times (1 - 0.402148) / 37,700 \\ &= 0.402148 \times 0.597852 / 37,700 \\ &= 0.240365 / 37,700 \\ &= 0.00000637636 \end{aligned}$$

Для Test:

$$\begin{aligned} \text{variance_test} &= p_t \times (1 - p_t) / n_t \\ &= 0.591280 \times (1 - 0.591280) / 26,446 \\ &= 0.591280 \times 0.408720 / 26,446 \\ &= 0.241671 / 26,446 \\ &= 0.00000913936 \end{aligned}$$

Сумарна варіація:

$$\begin{aligned} \text{variance_diff} &= \text{variance_control} + \text{variance_test} \\ &= 0.00000637636 + 0.00000913936 \\ &= 0.00001551572 \end{aligned}$$

Стандартна похибка:

$$SE(\Delta p) = \sqrt{0.00001551572}$$

$$= 0.00393902$$

$$= 0.3939 \text{ percentage points}$$

Крок 4: Z-значення для 95% CI

Для 95% довірчого інтервалу (two-tailed):

$$Z_{\{\alpha/2\}} = Z_{\{0.025\}} = 1.96$$

Крок 5: Margin of Error

$$\text{Margin of Error} = Z_{\{\alpha/2\}} \times SE(\Delta p)$$

$$= 1.96 \times 0.00393902$$

$$= 0.00772048$$

$$= 0.7720 \text{ percentage points}$$

Крок 6: Довірчий інтервал

Нижня межа (Lower Bound):

$$CI_{\text{lower}} = \Delta p - \text{Margin of Error}$$

$$= 0.189132 - 0.00772048$$

$$= 0.181412$$

$$= 18.14 \text{ percentage points}$$

Верхня межа (Upper Bound):

$$CI_{\text{upper}} = \Delta p + \text{Margin of Error}$$

$$= 0.189132 + 0.00772048$$

$$= 0.196852$$

= 19.69 percentage points

5. Результат

5.1. Фінальний довірчий інтервал

95% Confidence Interval для різниці конверсії Cart→Purchase:

CI_95% = [18.14%, 19.69%]

Або у форматі:

$\Delta p = 18.91 \text{ pp} \pm 0.77 \text{ pp}$ (95% CI: [18.14, 19.69])

5.2. Інтерпретація результату

Ключові висновки:

1. Point estimate (найкраща оцінка):

- Тестова група має на 18.91 pp вищу конверсію Cart→Purchase

2. 95% Confidence Interval:

- З 95% імовірністю справжня різниця між групами знаходиться в межах 18.14% до 19.69%

3. Статистична значущість:

- ☒ CI НЕ містить нуль (0)
- ☒ Різниця СТАТИСТИЧНО ЗНАЧУЩА на рівні 95%
- ☒ З 95% впевненістю можемо стверджувати, що Test група має кращу конверсію

4. Практична значущість:

- Навіть у найгіршому сценарії (lower bound) покращення становить +18.14 pp
- Це відповідає мінімум +45% відносного покращення $((1+0.1814/0.4021)-1)$
- Таке покращення має величезне бізнес-значення

6. Верифікація через Python

6.1. Код для перевірки

```
import numpy as np

from scipy import stats

# Data

n_control = 37700

n_test = 26446

purchases_control = 15161

purchases_test = 15637

# Conversion rates

p_control = purchases_control / n_control

p_test = purchases_test / n_test

# Difference

diff = p_test - p_control

# Standard errors

se_control = np.sqrt(p_control * (1 - p_control) / n_control)

se_test = np.sqrt(p_test * (1 - p_test) / n_test)

se_diff = np.sqrt(se_control**2 + se_test**2)

# 95% CI

z_score = stats.norm.ppf(0.975) # 1.96

ci_lower = diff - z_score * se_diff
```



```
ci_upper = diff + z_score * se_diff
```

```
# Output
```

```
print(f"Control conversion: {p_control*100:.2f}%")
```

```
print(f"Test conversion: {p_test*100:.2f}%")
```

```
print(f"Difference: {diff*100:.2f} pp")
```

```
print(f"Standard Error: {se_diff*100:.2f} pp")
```

```
print(f"95% CI: [{ci_lower*100:.2f}%, {ci_upper*100:.2f}%]")
```

```
print(f"Contains zero: {ci_lower <= 0 <= ci_upper}")
```

6.2. Результат виконання

Control conversion: 40.21%

Test conversion: 59.13%

Difference: 18.91 pp

Standard Error: 0.39 pp

95% CI: [18.14%, 19.69%]

Contains zero: False

✅ Результати співпадають з ручними розрахунками!

7. Альтернативні методи (для порівняння)

7.1. Z-Test для пропорцій

Можна також використати z-test для перевірки гіпотези:

```
from statsmodels.stats.proportion import proportions_ztest
```

```
count = np.array([purchases_test, purchases_control])
```

```
nobs = np.array([n_test, n_control])

z_stat, p_value = proportions_ztest(count, nobs)

print(f"Z-statistic: {z_stat:.4f}")

print(f"P-value: {p_value:.6f}")
```

Результат:

Z-statistic: 48.0089

P-value: 0.000000

Інтерпретація: p-value практично нуль (< 0.000001), що підтверджує дуже високу статистичну значущість різниці.

7.2. Bootstrap Confidence Interval (альтернативний метод)

Для додаткової перевірки можна використати bootstrap:

```
from scipy.stats import bootstrap

def conversion_diff(control_data, test_data):

    return test_data.mean() - control_data.mean()

# Simulate data (bernoulli trials)

rng = np.random.default_rng(42)

control_data = rng.binomial(1, p_control, n_control)

test_data = rng.binomial(1, p_test, n_test)

# Bootstrap CI

res = bootstrap(

    (control_data, test_data),

    conversion_diff,
```

```
n_resamples=10000,  
  
method='percentile'  
  
)  
  
print(f"Bootstrap 95% CI: [{res.confidence_interval.low*100:.2f}%,  
{res.confidence_interval.high*100:.2f}%]")
```

Очікуваний результат: Схожий до параметричного CI ([18.1%, 19.7%])

8. Чутливість до помилок вимірювання

8.1. Що якщо є похибка в даних?

Припустимо, є $\pm 2\%$ похибка у вимірюванні покупок (measurement error).

Песимістичний сценарій (worst case):

- Test purchases: $15,637 \times 0.98 = 15,324$
- Control purchases: $15,161 \times 1.02 = 15,464$

$p_{\text{test_worst}} = 15,324 / 26,446 = 0.5794$ (57.94%)

$p_{\text{control_worst}} = 15,464 / 37,700 = 0.4101$ (41.01%)

$\text{diff_worst} = 0.5794 - 0.4101 = 0.1693$ (16.93 pp)

Навіть у цьому сценарії різниця **дуже значна** (+41% відносно).

Висновок: Результати **robust** до можливих похибок вимірювання.

9. Порівняння з іншими метриками

9.1. Контекст відносно інших конверсій

Метрика	Control	Test	Різниця (pp)	Relative Change	CI Contains Zero?
CTR	4.86%	8.09%	+3.23 pp	+66.5%	❌ NO (significant)
Cart→Purchase	40.21%	59.13%	+18.91 pp	+47.0%	❌ NO (significant)
Overall Conv	9.83%	8.64%	-1.19 pp	-12.1%	✅ YES (not significant)

Спостереження:

- Cart→Purchase має найбільше абсолютне покращення (+18.91 pp)
- Це друга за значущістю метрика після CTR
- Підтверджує, що Test кампанія приводить більш якісний трафік

10. Бізнес-імплікації

10.1. Що це означає для бізнесу?

⚠️ **ВАЖЛИВО:** Наступні сценарії є гіпотетичними прогнозами для ілюстрації потенційного ефекту при масштабуванні. Фактичні результати тесту описані в секції 10.2.

Гіпотетичний Scenario 1: Поточні обсяги (місячні)

Якщо у місяць є ~30,000 Add to Cart (близько до тестових даних):

- Control: $30,000 \times 0.4021 = 12,063$ purchases
- Test: $30,000 \times 0.5913 = 17,739$ purchases
- Додаткові purchases: +5,676 per month (+47%)

Гіпотетичний Scenario 2: Масштабування (при 100k Add to Cart/month)

- Control: $100,000 \times 0.4021 = 40,210$ purchases
- Test: $100,000 \times 0.5913 = 59,130$ purchases
- Додаткові purchases: +18,920 per month

Якщо Average Order Value (AOV) = \$50:

- Додатковий revenue = $18,920 \times \$50 = \$946,000/\text{month}$
- Річний додатковий revenue = \$11.35 million

Застереження: Ці розрахунки припускають, що покращення конверсії зберігається при масштабуванні, що потребує валідації.

10.2. ROI оцінка для фактичних результатів тесту

Додатковий revenue на investment dollar:

Marginal ROAS = Additional Revenue / Additional Cost

Фактичні результати тесту (серпень 2019):

- Control Spend: \$66,818
- Test Spend: \$76,892
- Additional Spend: \$10,074
- Control Purchases: 15,161
- Test Purchases: 15,637
- Additional Purchases: 476

При AOV = \$50:

- Additional revenue: $476 \times \$50 = \$23,800$
- Additional ad cost: \$10,074
- Marginal ROAS: 2.36x (кожен додатковий долар витрат приносить \$2.36 revenue)

Висновок: Інвестиція у Test підхід **рентабельна** з позитивним ROI.

Примітка: Scenario 2 вище (100k Add to Cart/month) є гіпотетичним прогнозом при масштабуванні. Фактичний тест показав набагато менші абсолютні обсяги, але підтвердив покращення ефективності конверсії.

11. Обмеження та застереження

11.1. Припущення методу

✓ Виконані:

- Незалежність між групами
- Достатній розмір вибірки
- Біноміальна природа даних

⚠ Потенційні обмеження:

1. **Сезонність:** Дані тільки за серпень 2019
 - Рішення: Повторити тест в інші сезони
2. **Novelty effect:** Можливий ефект новизни
 - Рішення: Подовжити період спостереження до 60-90 днів
3. **External validity:** Чи узагальнюються результати?
 - Рішення: Тестувати на інших ринках/сегментах

11.2. Рекомендації щодо інтерпретації

Правильно:

- ✓ "З 95% впевненістю, Test група має на 18-20 pp кращу конверсію"
- ✓ "Різниця статистично та практично значуща"
- ✓ "Очікується +47% покращення Cart→Purchase rate"

Неправильно:

- ✗ "Є 95% ймовірність, що різниця між 18.14% і 19.69%"
 - Правильніше: "Якщо повторити експеримент 100 разів, в 95 випадках справжня різниця буде в цих межах"
- ✗ "Test група завжди буде на 18.91% краща"
 - Правильніше: "Наша найкраща оцінка покращення - 18.91 pp"

12. Висновки

12.1. Підсумок аналізу

Основний результат:

95% Довірчий інтервал для різниці конверсії Cart→Purchase (Test - Control):

CI_95% = [18.14%, 19.69%]

Point Estimate = 18.91 pp

Standard Error = 0.39 pp

Статистичні висновки:

- ☒ Різниця **СТАТИСТИЧНО ЗНАЧУЩА** (CI не містить нуль)
- ☒ Z-statistic = 48.01, p-value < 0.000001
- ☒ Ефект дуже великий (Cohen's h \approx 0.38, що вважається mid-large)
- ☒ Результати **robust** та **надійні**

Практичні висновки:

- ☒ Test підхід покращує конверсію на **+47%** (відносно)
- ☒ Навіть у найгіршому випадку покращення **+45%**
- ☒ Очікуваний додатковий revenue: **мільйони доларів** при масштабуванні
- ☒ **НАСТІЙНО РЕКОМЕНДУЄТЬСЯ** запуск Test підходу на 100%

12.2. Відповідь на завдання

Завдання 5: Розрахувати довірчий інтервал в різниці конверсії кроку покупки після додавання продукту до кошика.

Відповідь:

При 95% рівні довіри, різниця у конверсії Cart→Purchase між Test та Control групами становить:

- **Point Estimate:** +18.91 percentage points
- **95% Confidence Interval:** [+18.14%, +19.69%]
- **Інтерпретація:** З 95% впевненістю, Test кампанія має на 18.14% до 19.69% вищу конверсію з кошика в покупку порівняно з Control кампанією

- Статистична значущість: Різниця статистично значуща ($p < 0.001$)
- Практична значущість: Покращення на 47% має величезне бізнес-значення