

План A/B тесту: Вплив знижки 50% на конверсію підписки

1. Гіпотеза

Нульова гіпотеза (H0): Додавання повідомлення про знижку 50% до екрану підписки вартістю \$4.99 не впливає на показник конверсії у підписку.

Альтернативна гіпотеза (H1): Додавання повідомлення про знижку 50% до екрану підписки вартістю \$4.99 збільшує показник конверсії у підписку порівняно з поточним дизайном екрану.

Очікуваний результат: Ми припускаємо, що відображення знижки 50% (що імплікує початкову ціну \$9.98) створить сприйняття вищої цінності продукту та підвищить конверсію з поточних 17% до 20-22%.

2. Аудиторія для тесту

Критерії включення:

- Всі нові користувачі, які встановили застосунок під час періоду тестування
- Користувачі, які успішно завершили онбординг
- Користувачі, які дійшли до екрану з пропозицією підписки (34% від встановлень)

Критерії виключення:

- Користувачі, які вже мають активну підписку
- Користувачі, які раніше бачили будь-який з варіантів екрану підписки (щоб уникнути ефекту навчання)
- Користувачі з технічних тестових акаунтів

Розподіл:

- **Контрольна група (А):** 50% користувачів бачать поточний дизайн екрану з підпискою \$4.99
- **Тестова група (В):** 50% користувачів бачать новий дизайн з підпискою \$4.99 та повідомленням про знижку 50%

Метод рандомізації: Випадковий розподіл на рівні користувача за user_id для забезпечення консистентності досвіду.

3. Метрики

Цільова метрика

Conversion Rate (CR) до тижневої підписки

- **Формула:** (Кількість користувачів, що купили підписку / Кількість користувачів, що побачили екран підписки) \times 100%
- **Поточне значення:** 17%
- **Мінімальна детектована різниця (MDE):** 2.5 в.п. (відносне підвищення ~15%)

Обґрунтування: Це основна бізнес-метрика, яка напряду відображає ефективність зміни дизайну. Підвищення конверсії безпосередньо впливає на виручку та окупність витрат на залучення користувачів.

Допоміжні метрики

1. Revenue per User (RPU) на екрані підписки

- **Формула:** Загальний дохід від підписок / Кількість користувачів, що побачили екран
- **Поточне значення:** $\$4.99 \times 0.17 = \0.85
- **Обґрунтування:** Допомогає оцінити фінансовий вплив зміни. Навіть якщо CR зростає, важливо переконатися, що це не призведе до збільшення відмов/рефандів.

2. Click-through Rate (CTR) на кнопку підписки

- **Формула:** (Кількість кліків на кнопку "Підписатися" / Кількість показів екрану) \times 100%
- **Обґрунтування:** Показує, чи зацікавлює нова пропозиція користувачів на рівні інтенції. Якщо CTR зростає, але CR залишиться незмінним, це може вказувати на проблеми на етапі оплати.

3. Time on Screen (середній час на екрані підписки)

- **Обґрунтування:** Допомогає зрозуміти, чи читають користувачі інформацію про знижку, чи це їх збиває з пантелику.

4. Refund Rate (протягом перших 3 днів)

- **Формула:** (Кількість рефандів / Кількість покупок) \times 100%

- **Обґрунтування:** Критична метрика для перевірки, чи не сприймають користувачі знижку як маніпуляцію. Високий рівень повернень може вказувати на негативний досвід.

5. Retention Day 7

- **Формула:** Користувачі, активні на 7-й день / Користувачі, що купили підписку
- **Обґрунтування:** Перевіряє, чи залучає нова пропозиція "правильних" користувачів, які продовжують користуватися застосунком.

4. Розрахунок розміру вибірки та тривалості тесту

Вихідні дані:

- Базова конверсія (p_1): 17% = 0.17
- Очікувана конверсія в тестовій групі (p_2): 20% = 0.20
- Рівень значущості (α): 0.05 (ймовірність помилки I роду)
- Потужність тесту ($1-\beta$): 0.80 (ймовірність виявити ефект, якщо він є)
- Двосторонній тест

Розрахунок (калькулятор Evan Miller)

Question: How many subjects are needed for an A/B test?

Baseline conversion rate:	<input type="text" value="17"/>	%	<div><div></div></div> 17%	[link]
Minimum Detectable Effect:	<input type="text" value="3"/>	%	<div><div></div><div></div></div> 14% – 20%	
<small>The Minimum Detectable Effect is the smallest effect that will be detected (1-β)% of the time.</small>			<input checked="" type="radio"/> Absolute <input type="radio"/> Relative	
			<small>Conversion rates in the gray area will not be distinguishable from the baseline.</small>	

Sample size:

2,510

per variation

Statistical power $1-\beta$: 80% Percent of the time the minimum effect size will be detected, assuming it exists

Significance level α : 5% Percent of the time a difference will be detected, assuming one does NOT exist

n ≈ 2,510 користувачів на групу (підтверджено через Evan's A/B Tools Calculator)

Загальний розмір вибірки: 5,020 користувачів (обидві групи разом)

Розрахунок тривалості тесту:

Щоденний трафік:

- Встановлень на день: 2,000
- До екрану підписки доходять: $2,000 \times 34\% = 680$ користувачів/день
- На одну групу: $680 / 2 = 340$ користувачів/день

Кількість днів:

- Мінімальна кількість: $2,510 / 340 \approx 7.4$ днів
- **Рекомендована тривалість: 14 днів (2 тижні)**

Обґрунтування тривалості:

- Покриття двох повних тижнів дозволяє врахувати сезонність (різниця між робочими днями та вихідними)
 - Додатковий час забезпечує запас для досягнення статистичної значущості
 - Це дає нам ~9,520 користувачів (4,760 на групу), що майже вдвічі перевищує мінімальний розмір вибірки та забезпечує надійність результатів
-

Критерії успішності тесту

Тест вважається успішним, якщо:

1. **Статистична значущість:** $p\text{-value} < 0.05$ для різниці в конверсії між групами А і В
2. **Практична значущість:** Conversion Rate в групі В вищий мінімум на 2 в.п. (з 17% до 19%+)
3. **Допоміжні метрики в межах норми:**
 - Refund Rate не перевищує поточний показник більше ніж на 20%
 - Retention Day 7 не погіршується більше ніж на 5%
 - Revenue per User зростає пропорційно до CR

Сценарії дій залежно від результатів

Сценарій 1: Тест успішний (CR зріс, $p < 0.05$, метрики в нормі)

Дії:

- Розгорнути варіант В на 100% користувачів
- Моніторити метрики протягом наступних 2 тижнів для підтвердження результатів

- Провести якісне дослідження (опитування, інтерв'ю) для розуміння, чому знижка спрацювала
- Розглянути тестування знижок в інших точках користувацького шляху

Сценарій 2: Тест неуспішний (немає статистично значущого покращення)

Дії:

- Залишити поточний варіант А
- Проаналізувати дані сегментовано (за платформою, країною, джерелом трафіку) — можливо, ефект є в певних сегментах
- Перейти до тестування альтернативних варіантів (див. нижче)

Сценарій 3: Негативний результат (CR впав або погіршилися допоміжні метрики)

Дії:

- негайно зупинити тест
- Залишити поточний варіант А
- Провести якісне дослідження для розуміння, чому знижка відштовхнула користувачів
- Переглянути стратегію ціноутворення

Альтернативні варіанти для тестування

Якщо основний тест не буде успішним, пропоную наступні гіпотези:

Варіант 1: Соціальний доказ + обмежена пропозиція

Зміна: Додати текст "Понад 50,000 користувачів вже підписалися" + "Пропозиція дійсна 24 години" **Гіпотеза:** Соціальний доказ та FOMO підвищать довіру та терміновість покупки

Очікуваний ефект: CR 17% → 20%

Варіант 2: Безкоштовний триал на 3 дні + \$4.99/тиждень

Зміна: Спочатку надати 3 дні безкоштовного доступу, потім \$4.99/тиждень **Гіпотеза:** Зниження бар'єру входу через безкоштовний період збільшить кількість підписок

Очікуваний ефект: CR 17% → 25-30% (початкова конверсія), але потрібно моніторити trial-to-paid конверсію

Варіант 3: Тестування ціни + value proposition

Зміна: Залишити ціну \$4.99, але додати детальний список преміум-функцій з візуалізацією

Гіпотеза: Користувачі не розуміють цінність підписки, тому не купують **Очікуваний ефект:** CR 17% → 21%

Варіант 4: Багаторівнева пропозиція

Зміна: Показати три опції: \$4.99/тиждень, \$14.99/місяць (знижка 25%), \$49.99/рік (знижка 50%) **Гіпотеза:** Наявність варіантів та ефект якоря підвищать загальну конверсію та ARPU **Очікуваний ефект:** Загальна CR до будь-якої підписки 17% → 22%

Рекомендація щодо пріоритетності

Якщо основний тест не вдасться:

1. **Спочатку тестувати Варіант 2** (триал) — найбільший потенціал підвищення конверсії
 2. **Потім Варіант 3** — покращує value proposition без зміни ціноутворення
 3. **Нарешті Варіант 4** — найскладніший для реалізації, але з високим потенціалом ARPU
-

Додаткові рекомендації

Технічна реалізація

- Використати feature flag систему для швидкого перемикавання між варіантами
- Логувати всі події (show_screen, click_button, purchase, refund)
- Забезпечити консистентність групи для користувача протягом усього життєвого циклу

Моніторинг під час тесту

- Щоденний моніторинг метрик для раннього виявлення проблем
- Weekly review з командою для оцінки проміжних результатів
- Sample Ratio Mismatch (SRM) перевірка — співвідношення груп має бути 50/50

Ризики та мітигація

- **Ризик:** Знижка може сприйматися як маніпуляція → **Мітигація:** Моніторити refund rate та відгуки
- **Ризик:** Сезонність (свята, вихідні) → **Мітигація:** Тривалість тесту 14 днів
- **Ризик:** Технічні збої → **Мітигація:** Killswitch для швидкої зупинки тесту