**План А/В тестування**

**Вхідні дані:**

В нашому мобільному застосунку користувачу після онбордингу пропонують купити тижневу підписку вартістю $4.99, що надає доступ до преміум-функцій.

Зараз підписку на цьому екрані купують 17% з користувачів, що отримали пропозицію. Ми вирішили протестувати альтернативний дизайн екрана з підпискою, на якому ми також пропонуємо підписку вартістю $4.99, але кажемо користувачеві, що це зі знижкою 50%.

Кожного дня застосунок встановлюють близько 2 тис користувачів, до екрану з підпискою при цьому доходять 34% користувачів.

**Параметри тестування:**

1. **Рівень значущості** (significance level) **α** = 0,05;
2. **Потужність тесту** (Ймовірність виявити реальний ефект, якщо він справді існує) power = 1 – β = 80%.
3. Щоденні інсталяції = 2000, до екрану доходять 34% (**680 користувачів в день)** доступні для тесту;
4. базова конверсія на екрані = **17%**.

**1. Гіпотеза**

**H0:** Зміна дизайну (повідомлення «50% off» при тій же ціні $4.99) не змінює конверсію покупки.  
**H1:** Нова версія підвищує конверсію покупки.

**2. Популяція учасників**

* Включаємо **усіх нових користувачів**, які дійшли до екрану пропозиції підписки (тобто 34% від 2 тис.).
* Виключаємо: команди QA/dev, бот-трафік, внутрішні корпоративні акаунти.
* Розподіл трафіку: **рандомно 50/50** (контроль - поточний екран; варіант - зі знижкою).

**3. Цільові та допоміжні метрики**

**Цільова (primary):**

Conversion rate на екрані (доля тих, хто купив підписку) - тестована метрика.

**Допоміжні (secondary):**

* ARPU або Avg Revenue per reached user (щоб бачити вплив на доходи), кількість покупок в перші 7 та 30 днів (retention of paying users);
* CTR частка користувачів, які клікнули на кнопку знижки;
* CTA (ключова кнопка або заклик до дії на екрані);
* Негативні сигнали (відміни, скарги) - безпековий чек.

**Обґрунтування**: Цільова метрика **CR** - безпосередньо бізнес-мета; допоміжні метрики - щоб не прийняти фальшивий результат, що знижує LTV.

**4. Розмір вибірки та тривалість**

Параметр MDE (мінімально значущий ефект) потрібно вибрати практично. Розраховуємо для трьох варіантів, приблизно, нормальна апроксимація:

**Поточна Базова конверсія** у контрольній групі,   
тобто "від чого ми стартуємо": **p₀** = 0.17.

**Розрахунки (α=0.05, power=0.8):**

1. **MDE = +2** процентні пункти **(рр) = 19%**  
   n (розмір вибірки) ≈ **5783** на групу; **11 566** всього; при 680/день ≈ **17–18 днів**.
2. **MDE = +3 рр (20%)** - рекомендований компроміс  
   n ≈ **2618** на групу; **5236** всього; при 680/день ≈ **8 днів**.
3. **MDE = +5 рр (22%)**  
   n ≈ **984** на групу; **1968** всього; при 680/день ≈ **3 дні**.

**Рекомендація:** вибрати MDE = **+3 pp** (реалістичний і бізнес-значущий).

Мінімальний розмір вибірки: **≈2618 користувачів у кожній групі** – період тестування **8 днів** за ідеальних умов.

**Обов’язкова практична поправка:**

* Запускати тест **мінімум 14 днів**, навіть якщо за розрахунком вистачає 8 днів, щоб покрити тижневі цикли, сезонні хвилі та стабілізувати метрики.
* Переконатися, що щоденний трафік не коливається різко; якщо є флуктуації, продовжити тест до періоду стабільності.

**5. Критерії успіху й дії**

**Успішність тесту:**

* **Статистична значущість**: p-value < 0.05 (двосторонній тест) **та** різниця ≥ обраного MDE (наприклад, ≥ +3 pp).
* **Бізнес-умова**: додатково перевірити, що ARPU або ROMI не знизились (або загальний дохід зріс або не знизився).

**Дії:**

* Якщо успіх (p<0.05 і бізнес-метрики не погіршилися): впроваджуємо новий дизайн для всіх користувачів.
* Якщо статистично значуще, але **негативний вплив** на доходи/LTV – не впроваджуємо, аналізуємо причини (можливо, купують більше одноразово, але потім відписуються).
* Якщо незначущий (p≥0.05): закриваємо тест як неуспішний; аналіз – сегментація, перевірка платформи, країни, джерела трафіку. Можливі подальші дії з іншим креативом або більшою вибіркою (зниження MDE).

**6. Альтернативний варіант, якщо тест не спрацює**

Якщо повідомлення «50% off» не підвищить продажі:

* **Тестувати іншу пропозицію**: наприклад, **пробний період 7 днів безплатно**, а потім $4.99 — протестувати вплив trial на придбання і утримання.
* **Тестувати зміну CTA та дизайн**: інша кнопка (скільки часу пропозиція діє), соціальний доказ (reviews), спліт за порядком блоків (ціна без знижки/зі знижкою).
* **Сегментовані експерименти**: показати знижку лише певним сегментам (нові користувачі з iOS, або з певних країн) — можливо, позитивний ефект в окремих групах.

**Короткий план дій перед запуском тестування (checklist)**

1. Реалізувати рандомізацію 50/50 та логувати group\_id.
2. Поставити events: impression (дійшов до показу екрану знижки), view, click, purchase, revenue, user\_id, date, device, country.
3. Протестувати метрику baseline – визначити daily\_reached стабільність.
4. Запустити тест мінімум 14 днів або поки не буде досягнуто обчисленого розміру вибірки на групу.
5. Моніторинг за «стоп-критеріями»: критичні баги, різка зміна відтоку, технічні проблеми.
6. Після завершення провести статистичний тест (z-test для двох пропорцій), зробити аналіз secondary метрик і сегментів.