

**Я.С. Бондаренко**

*Дніпровський національний університет імені Олеся Гончара*

## **ПОСЛІДОВНИЙ КРИТЕРІЙ ВІДНОШЕННЯ ЙМОВІРНОСТЕЙ ДЛЯ РАНДОМІЗОВАНИХ ОНЛАЙН ЕКСПЕРИМЕНТІВ**

В роботі запропоновано послідовний критерій відношення ймовірностей для перевірки простої гіпотези відносно множини альтернативних гіпотез при розв'язанні задачі оптимізації коефіцієнту конверсії. Застосовано байєсівський метод статистичного оцінювання невідомих параметрів.

**Ключові слова:** оптимізація коефіцієнту конверсії, послідовний аналіз, байєсівський метод статистичного оцінювання.

**Ya.S. Bondarenko**

*Oles Honchar Dnipro National University*

## **A SEQUENTIAL PROBABILITY RATIO TEST FOR RANDOMIZED ONLINE EXPERIMENTS**

In this paper a mixture sequential probability ratio test for simple hypothesis against a set of infinitely many alternatives for conversion rate optimization is proposed. Bayesian inference for estimation of unknown parameters is presented. Landing page optimization is the process of improving elements on a website to increase conversions. Landing page optimization is a subset of conversion rate optimization, and involves using methods such as A/B testing to improve the conversion goals of a given landing page. A/B testing is an experiment where two versions of landing page are shown to users at random, and statistical hypothesis testing is used to determine which version performs better for a given conversion goal.

In classical hypothesis testing the data collection is implemented without analysis and study of the data. Once all data are collected the statistical analysis is performed and conclusions are reached. Moving away from classical hypothesis testing to sequential analysis developed by A. Wald was presented in numerous papers. In sequential analysis every observation is analyzed immediately after being collected. The data are collected until that moment is compared with preassigned threshold values, which integrate the knowledge obtained from the newly collected observation. This technique allows to reach conclusions during the data collection. Final conclusion can be reached with much fewer observations than in the case of classical hypothesis testing.

In this paper sequential analysis as the technique of decision making during A/B testing by sequentially gathering and processing the data is presented. The advantages of sequential A/B testing are easy to see. As data collection can be terminated after fewer observations and decisions can be taken earlier, financial savings might be considerable. Sequential A/B testing allows to test different hypotheses for landing page optimization experiments and make changes based on actual data. Combining a priori information with empirical information of sequential actions of users on the website or application, allows the latter to dominate in the a posteriori distribution of parameters.

**Keywords:** conversion rate optimization, sequential analysis, Bayesian inference.

**Я.С. Бондаренко**

*Дніпровський національний університет імені Олеся Гончара*

## **ПОСЛЕДОВАТЕЛЬНЫЙ КРИТЕРИЙ ОТНОШЕНИЯ ВЕРОЯТНОСТЕЙ ДЛЯ РАНДОМИЗИРОВАННЫХ ОНЛАЙН ЭКСПЕРИМЕНТОВ**

**В работе предложен последовательный критерий отношения вероятностей для проверки простой гипотезы относительно множества альтернативных гипотез при решении задачи оптимизации коэффициента конверсии. Применен байесовский метод статистического оценивания неизвестных параметров.**

**Ключевые слова:** оптимизация коэффициента конверсии, последовательный анализ, байесовский метод статистического оценивания.

**Вступ.** Найважливішим ресурсом дата-орієнтованої компанії є дані. Вивчення даних та прийняття рішень на основі даних, а не на основі оцінок експертів або інтуїції, надає компанії безліч переваг. Компанії, які впровадили дата-орієнтовану культуру, значно випереджають своїх конкурентів за фінансовими показниками. Дата-орієнтовані компанії Airbnb, Amazon, Booking.com, eBay, Facebook, Google, LinkedIn, Lyft, Microsoft, Netflix, Twitter, Uber, Yahoo!/Oath, Yandex проводять сотні тисяч рандомізованих онлайн-експериментів щороку, залучаючи мільйони користувачів для тестування нових ідей та нових алгоритмів з метою оптимізації метрик успіху, зокрема, коефіцієнтів конверсії [1-5]. Оптимізація коефіцієнту конверсії – це процес збільшення проценту конверсії веб-сайту або мобільного додатку. Оптимізація коефіцієнту конверсії передбачає створення нових гіпотез щодо оновлення елементів веб-сайту або додатку, які можна покращити, а потім відхилення або не відхилення цих гіпотез за допомогою рандомізованих онлайн-експериментів з двома або багатьма групами користувачів.

Найпростіший рандомізований онлайн-експеримент з двома групами користувачів називають А/В тестуванням. Це метод оптимізації оригінальної цільової сторінки, на яку переходить відвідувач, залучений за допомогою реклами, систем пошуку, поштової розсилки. Оригінальна сторінка та її варіація (альтернативний варіант цільової сторінки) мають бути однакові за виключенням одного елементу, вплив якого на сприйняття відвідувачів перевіряється під час тестування. Оригінальна сторінка пропонується для перегляду базовій групі відвідувачів, варіація сторінки пропонується до перегляду експериментальній групі. Розглянемо потік відвідувачів: кожен відвідувач з імовірністю  $1/2$  може потрапити або до базової групи або до експериментальної групи. Після того, як відвідувач опинився в одній із двох груп, поведінка відвідувача однозначно визначається двома подіями: “успіх” – відвідувач виконав цільову дію (дія відвідувача, яка значуща для власника сайту: скачування безкоштовного контенту, реєстрація, підписка на розсилку, покупка товару, заказ послуги) або “неуспіх” – відвідувач не виконав цільову

дію. Якщо відвідувач належить базовій групі, то подія “успіх” відбувається з імовірністю  $p_1$ . Якщо відвідувач належить експериментальній групі, то подія “успіх” відбувається з імовірністю  $p_2$ . У зв’язку з цим виникає задача побудови критерію для перевірки статистичної гіпотези  $H_0 : p_1 = p_2$  про рівність імовірностей “успіхів” (коефіцієнтів конверсії).

У класичних статистичних критеріях обсяг вибірки відвідувачів заданий наперед, а це іноді призводить до великих зайвих витрат [6]. У послідовному аналізі Вальда обсяг вибірки відвідувачів не фіксується, а визначається у процесі аналізу статистичних даних, отриманих послідовно, в порядку їх надходження. Це призводить до помітного зменшення обсягу роботи у порівнянні з класичними критеріями [7-13, 19, 20].

**Постановка задачі.** Запропонувати послідовний критерій для перевірки простої гіпотези  $H_0 : p_1 = p_2$  відносно нескінченно великого числа альтернативних гіпотез  $H_\theta : p_1 - p_2 = \theta, \theta \neq 0$ , при цьому розподіл різниці коефіцієнтів конверсії  $\theta$  природньо задати щільністю нормального розподілу з випадковими параметрами  $(\mu, \sigma^2)$  (див. рис. 1).

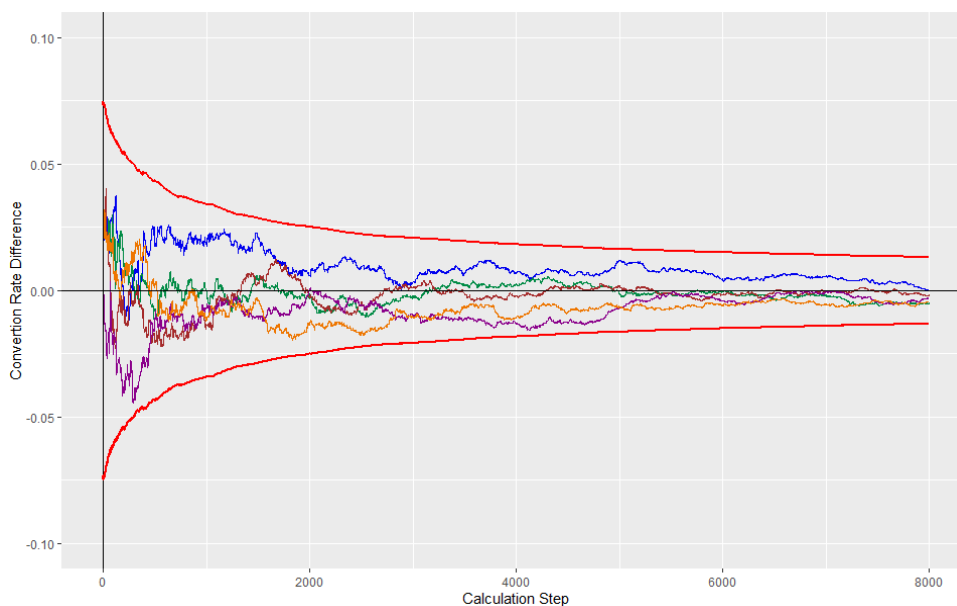


Рис. 1. Флуктуації різниці коефіцієнтів конверсії

**Відношення ймовірностей.** Нехай  $X_1, X_2, \dots, X_{n_1}$  – послідовність незалежних, однаково розподілених випадкових величин з невідомим параметром  $p_1$ :

$$P(x) = p_1^x (1 - p_1)^{1-x}, \quad x = \{0, 1\}, \quad p_1 \in (0, 1),$$

$Y_1, Y_2, \dots, Y_{n_2}$  – послідовність незалежних, однаково розподілених випадкових величин з невідомим параметром  $p_2$ :

$$P(y) = p_2^y (1 - p_2)^{1-y}, \quad y = \{0, 1\}, \quad p_2 \in (0, 1).$$

Спільний розподіл випадкових величин  $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$  набуває вигляду

$$L(p_1, p_2, x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) = p_1^{\sum_{i=1}^{n_1} x_i} (1-p_1)^{n_1 - \sum_{i=1}^{n_1} x_i} p_2^{\sum_{i=1}^{n_2} y_i} (1-p_2)^{n_2 - \sum_{i=1}^{n_2} y_i}.$$

Тоді оцінки максимальної правдоподібності для  $p_1, p_2$  дорівнюють

$$\hat{p}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} x_i, \quad \hat{p}_2 = \frac{1}{n_2} \sum_{i=1}^{n_2} y_i.$$

Здобуті оцінки підставимо в спільний розподіл:

$$L(\hat{p}_1, \hat{p}_2, x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) = \hat{p}_1^{n_1 \hat{p}_1} (1-\hat{p}_1)^{n_1 - n_1 \hat{p}_1} \hat{p}_2^{n_2 \hat{p}_2} (1-\hat{p}_2)^{n_2 - n_2 \hat{p}_2}.$$

При справедливій нульовій гіпотезі  $H_0: p_1 - p_2 = 0$  спільний розподіл випадкових величин  $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$  набуває вигляду:

$$L(p_0, x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) = p_0^{\sum_{i=1}^{n_1} x_i} (1-p_0)^{n_1 - \sum_{i=1}^{n_1} x_i} p_0^{\sum_{i=1}^{n_2} y_i} (1-p_0)^{n_2 - \sum_{i=1}^{n_2} y_i}.$$

Тоді оцінка максимальної правдоподібності для  $p_0$  дорівнює

$$\hat{p}_0 = \frac{1}{n_1 + n_2} \left( \sum_{i=1}^{n_1} x_i + \sum_{i=1}^{n_2} y_i \right) = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2}.$$

Здобуту оцінку підставимо в спільний розподіл:

$$L(\hat{p}_0, x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) = \hat{p}_0^{n_1 \hat{p}_1} (1-\hat{p}_0)^{n_1 - n_1 \hat{p}_1} \hat{p}_0^{n_2 \hat{p}_2} (1-\hat{p}_0)^{n_2 - n_2 \hat{p}_2}.$$

Запишемо відношення ймовірностей:

$$\Lambda(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) = \frac{\hat{p}_0^{n_1 \hat{p}_1} (1-\hat{p}_0)^{n_1 - n_1 \hat{p}_1} \hat{p}_0^{n_2 \hat{p}_2} (1-\hat{p}_0)^{n_2 - n_2 \hat{p}_2}}{\hat{p}_1^{n_1 \hat{p}_1} (1-\hat{p}_1)^{n_1 - n_1 \hat{p}_1} \hat{p}_2^{n_2 \hat{p}_2} (1-\hat{p}_2)^{n_2 - n_2 \hat{p}_2}}.$$

Логарифм відношення ймовірностей дорівнює:

$$\begin{aligned} & -\ln \Lambda(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) = \\ & = n_1 \hat{p}_1 (\ln \hat{p}_1 - \ln \hat{p}_0) + n_1 (1-\hat{p}_1) (\ln(1-\hat{p}_1) - \ln(1-\hat{p}_0)) + \\ & + n_2 \hat{p}_2 (\ln \hat{p}_2 - \ln \hat{p}_0) + n_2 (1-\hat{p}_2) (\ln(1-\hat{p}_2) - \ln(1-\hat{p}_0)). \end{aligned}$$

Скористаємося лінійною апроксимацією для різниці логарифмів:

$$\ln(1-\hat{p}_i) - \ln(1-\hat{p}_0) = \ln \frac{1-\hat{p}_i}{1-\hat{p}_0} = \ln \left( 1 - \frac{\hat{p}_i - \hat{p}_0}{1-\hat{p}_0} \right) \approx -\frac{\hat{p}_i - \hat{p}_0}{1-\hat{p}_0}, \quad i=1,2,$$

$$\ln \hat{p}_i - \ln \hat{p}_0 = \ln \frac{\hat{p}_i}{\hat{p}_0} = \ln \left( 1 + \frac{\hat{p}_i - \hat{p}_0}{\hat{p}_0} \right) \approx \frac{\hat{p}_i - \hat{p}_0}{\hat{p}_0}, \quad i=1,2.$$

Підставимо перетворені різниці в логарифм відношення ймовірностей:

$$\begin{aligned} & -\ln \Lambda(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) \approx \\ & \approx n_1 \hat{p}_1 \left( \frac{\hat{p}_1 - \hat{p}_0}{\hat{p}_0} \right) + n_1 (1-\hat{p}_1) \left( -\frac{\hat{p}_1 - \hat{p}_0}{1-\hat{p}_0} \right) + \end{aligned}$$

$$\begin{aligned}
& +n_2\hat{p}_2\left(\frac{\hat{p}_2-\hat{p}_0}{\hat{p}_0}\right)+n_2(1-\hat{p}_2)\left(-\frac{\hat{p}_2-\hat{p}_0}{1-\hat{p}_0}\right)= \\
& =\frac{n_1(\hat{p}_1-\hat{p}_0)^2+n_2(\hat{p}_2-\hat{p}_0)^2}{\hat{p}_0(1-\hat{p}_0)}.
\end{aligned}$$

Перетворимо вирази у чисельнику:

$$\begin{aligned}
n_1(\hat{p}_1-\hat{p}_0)^2 & =n_1\left(\hat{p}_1-\frac{n_1\hat{p}_1+n_2\hat{p}_2}{n_1+n_2}\right)^2=n_1n_2^2\left(\frac{\hat{p}_1-\hat{p}_2}{n_1+n_2}\right)^2, \\
n_2(\hat{p}_2-\hat{p}_0)^2 & =n_2\left(\hat{p}_2-\frac{n_1\hat{p}_1+n_2\hat{p}_2}{n_1+n_2}\right)^2=n_2n_1^2\left(\frac{\hat{p}_2-\hat{p}_1}{n_1+n_2}\right)^2.
\end{aligned}$$

Логарифм відношення ймовірностей запишеться у вигляді:

$$\begin{aligned}
& -\ln\Lambda(x_1,\dots,x_{n_1},y_1,\dots,y_{n_2})\approx \\
& =\frac{n_1n_2^2\left(\frac{\hat{p}_1-\hat{p}_2}{n_1+n_2}\right)^2+n_2n_1^2\left(\frac{\hat{p}_2-\hat{p}_1}{n_1+n_2}\right)^2}{\hat{p}_0(1-\hat{p}_0)}=\frac{(\hat{p}_1-\hat{p}_2)^2}{\hat{p}_0(1-\hat{p}_0)\frac{n_1+n_2}{n_1n_2}}.
\end{aligned}$$

Отже, логарифм відношення ймовірностей набуває вигляду:

$$-\ln\Lambda(x_1,\dots,x_{n_1},y_1,\dots,y_{n_2})\approx\left|\frac{\hat{p}_1-\hat{p}_2}{\sqrt{\hat{p}_0(1-\hat{p}_0)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}\right|^2.$$

**Послідовність відношень ймовірностей.** Нехай  $\{\Omega,\sigma,P\}$  – імовірнісний простір,  $T$  – множина дійсних чисел. Для кожного  $t\in T$  розглянемо підсигма-алгебру  $\mathfrak{F}_t\subset\sigma$ , де  $t$  будемо інтерпретувати як моменти часу, в кожен з яких відбувається рандомізований експеримент. Сигма-алгебра  $\mathfrak{F}_t$  – це сигма-алгебра подій, які спостерігаються до моменту часу  $t$ . Природно, що  $\mathfrak{F}_{t_1}\subset\mathfrak{F}_{t_2}$ ,  $t_1\leq t_2$ .

Розглянемо сім'ю випадкових величин  $\xi(t)$ ,  $t\in T$  з властивістю: їхні значення визначаються експериментами, проведеними до моменту часу  $t$ . Для цього природньо вимагати, щоб при кожному  $t$  випадкова величина  $\xi(t)$  була вимірною відносно  $\mathfrak{F}_t$ .

Сім'ю сигма-алгебр  $\mathfrak{F}_t$ ,  $t\in T$ , яка монотонно не спадає ( $\mathfrak{F}_{t_1}\subset\mathfrak{F}_{t_2}$ ,  $t_1\leq t_2$ ) називають потоком сигма-алгебр. Сім'ю випадкових величин  $\xi(t)$ ,  $t\in T$  називають узгодженою з потоком сигма-алгебр  $\mathfrak{F}_t$ , якщо для кожного  $t$  випадкова величина  $\xi(t)$  вимірна відносно  $\mathfrak{F}_t$ . Сім'ю випадкових величин  $\{\xi(t), t\in T\}$ ,

узгоджену з потоком сигма-алгебр  $\mathfrak{F}_t$ , називають мартингалом, якщо для кожного  $t: M|\xi(t)| < \infty$  и  $M(\xi(t)|\mathfrak{F}_s) = \xi(s), s < t$ .

Послідовність відношень ймовірностей  $\{\Lambda_n\}$  є мартингалом при справедливій нульовій гіпотезі  $H_0: p_1 - p_2 = 0$ .

Дійсно, повинні виконуватися дві умови:  $M|\Lambda_n| < \infty, n = 0, 1, 2, \dots$ ;  $M(\Lambda_{n+1} | \Lambda_1, \dots, \Lambda_n) = \Lambda_n, n = 0, 1, 2, \dots$ . Розглянемо послідовності відношень ймовірностей. При  $k = 0$  відношення ймовірностей  $\Lambda_0$  дорівнює 1.

При  $k = 1$  відношення ймовірностей  $\Lambda_1$  набуває вигляду

$$\Lambda_1 = \frac{L(\hat{p}_1, \hat{p}_2, x_1, y_1)}{L(\hat{p}_0, x_1, y_1)} = \Lambda_0 \frac{\hat{p}_1^{x_1} (1 - \hat{p}_1)^{1-x_1} \hat{p}_2^{y_1} (1 - \hat{p}_2)^{1-y_1}}{\hat{p}_0^{x_1} (1 - \hat{p}_0)^{1-x_1} \hat{p}_0^{y_1} (1 - \hat{p}_0)^{1-y_1}},$$

при цьому математичне сподівання дорівнює

$$\begin{aligned} M\Lambda_1 &= M\Lambda_0 M \left( \frac{\hat{p}_1^{x_1} (1 - \hat{p}_1)^{1-x_1} \hat{p}_2^{y_1} (1 - \hat{p}_2)^{1-y_1}}{\hat{p}_0^{x_1} (1 - \hat{p}_0)^{1-x_1} \hat{p}_0^{y_1} (1 - \hat{p}_0)^{1-y_1}} \right) = \\ &= 1 \cdot \sum_{i=0}^1 \sum_{j=0}^1 \frac{\hat{p}_1^i (1 - \hat{p}_1)^{1-i} \hat{p}_2^j (1 - \hat{p}_2)^{1-j}}{\hat{p}_0^i (1 - \hat{p}_0)^{1-i} \hat{p}_0^j (1 - \hat{p}_0)^{1-j}} \hat{p}_0^i (1 - \hat{p}_0)^{1-i} \hat{p}_0^j (1 - \hat{p}_0)^{1-j} \\ &= \sum_{i=0}^1 \sum_{j=0}^1 \hat{p}_1^i (1 - \hat{p}_1)^{1-i} \hat{p}_2^j (1 - \hat{p}_2)^{1-j} = 1. \end{aligned}$$

При  $k = 2$  відношення ймовірностей  $\Lambda_2$  набуває вигляду

$$\Lambda_2 = \frac{L(\hat{p}_1, \hat{p}_2, x_1, x_2, y_1, y_2)}{L(\hat{p}_0, x_1, x_2, y_1, y_2)} = \Lambda_1 \frac{\hat{p}_1^{x_2} (1 - \hat{p}_1)^{1-x_2} \hat{p}_2^{y_2} (1 - \hat{p}_2)^{1-y_2}}{\hat{p}_0^{x_2} (1 - \hat{p}_0)^{1-x_2} \hat{p}_0^{y_2} (1 - \hat{p}_0)^{1-y_2}},$$

при цьому математичне сподівання дорівнює

$$M\Lambda_2 = M\Lambda_1 M \left( \frac{\hat{p}_1^{x_2} (1 - \hat{p}_1)^{1-x_2} \hat{p}_2^{y_2} (1 - \hat{p}_2)^{1-y_2}}{\hat{p}_0^{x_2} (1 - \hat{p}_0)^{1-x_2} \hat{p}_0^{y_2} (1 - \hat{p}_0)^{1-y_2}} \right) = 1.$$

При  $k = n$  відношення ймовірностей  $\Lambda_n$  набуває вигляду

$$\Lambda_n = \frac{L(\hat{p}_1, \hat{p}_2, x_1, \dots, x_n, y_1, \dots, y_n)}{L(\hat{p}_0, x_1, \dots, x_n, y_1, \dots, y_n)} = \Lambda_{n-1} \frac{\hat{p}_1^{x_n} (1 - \hat{p}_1)^{1-x_n} \hat{p}_2^{y_n} (1 - \hat{p}_2)^{1-y_n}}{\hat{p}_0^{x_n} (1 - \hat{p}_0)^{1-x_n} \hat{p}_0^{y_n} (1 - \hat{p}_0)^{1-y_n}},$$

при цьому математичне сподівання дорівнює

$$M\Lambda_n = M\Lambda_{n-1} M \left( \frac{\hat{p}_1^{x_n} (1 - \hat{p}_1)^{1-x_n} \hat{p}_2^{y_n} (1 - \hat{p}_2)^{1-y_n}}{\hat{p}_0^{x_n} (1 - \hat{p}_0)^{1-x_n} \hat{p}_0^{y_n} (1 - \hat{p}_0)^{1-y_n}} \right) = 1.$$

При  $k = n + 1$  відношення ймовірностей набуває вигляду

$$\Lambda_{n+1} = \frac{L(\hat{p}_1, \hat{p}_2, x_1, \dots, x_{n+1}, y_1, \dots, y_{n+1})}{L(\hat{p}_0, x_1, \dots, x_{n+1}, y_1, \dots, y_{n+1})} = \Lambda_n \frac{\hat{p}_1^{x_{n+1}} (1 - \hat{p}_1)^{1-x_{n+1}} \hat{p}_2^{y_{n+1}} (1 - \hat{p}_2)^{1-y_{n+1}}}{\hat{p}_0^{x_{n+1}} (1 - \hat{p}_0)^{1-x_{n+1}} \hat{p}_0^{y_{n+1}} (1 - \hat{p}_0)^{1-y_{n+1}}}.$$

Порахуємо умовне математичне сподівання відношення ймовірностей:

$$M(\Lambda_{n+1} | \Lambda_1, \dots, \Lambda_n) = M \left( \Lambda_n \frac{\hat{p}_1^{x_{n+1}} (1 - \hat{p}_1)^{1-x_{n+1}} \hat{p}_2^{y_{n+1}} (1 - \hat{p}_2)^{1-y_{n+1}}}{\hat{p}_0^{x_{n+1}} (1 - \hat{p}_0)^{1-x_{n+1}} \hat{p}_0^{y_{n+1}} (1 - \hat{p}_0)^{1-y_{n+1}}} \middle| \Lambda_1, \dots, \Lambda_n \right) =$$

$$= \Lambda_n M \left( \frac{\hat{p}_1^{x_{n+1}} (1 - \hat{p}_1)^{1-x_{n+1}} \hat{p}_2^{y_{n+1}} (1 - \hat{p}_2)^{1-y_{n+1}}}{\hat{p}_0^{x_{n+1}} (1 - \hat{p}_0)^{1-x_{n+1}} \hat{p}_0^{y_{n+1}} (1 - \hat{p}_0)^{1-y_{n+1}}} \middle| \Lambda_1, \dots, \Lambda_n \right).$$

Оскільки  $X_1, \dots, X_n, Y_1, \dots, Y_n$  – послідовності незалежних випадкових величин, то умовне математичне сподівання збігається з безумовним:

$$\begin{aligned} M(\Lambda_{n+1} | \Lambda_1, \dots, \Lambda_n) &= \Lambda_n M \left( \frac{\hat{p}_1^{x_{n+1}} (1 - \hat{p}_1)^{1-x_{n+1}} \hat{p}_2^{y_{n+1}} (1 - \hat{p}_2)^{1-y_{n+1}}}{\hat{p}_0^{x_{n+1}} (1 - \hat{p}_0)^{1-x_{n+1}} \hat{p}_0^{y_{n+1}} (1 - \hat{p}_0)^{1-y_{n+1}}} \right) = \\ &= \Lambda_n \sum_{i=0}^1 \sum_{j=0}^1 \frac{\hat{p}_1^i (1 - \hat{p}_1)^{1-i} \hat{p}_2^j (1 - \hat{p}_2)^{1-j}}{\hat{p}_0^i (1 - \hat{p}_0)^{1-i} \hat{p}_0^j (1 - \hat{p}_0)^{1-j}} \hat{p}_0^i (1 - \hat{p}_0)^{1-i} \hat{p}_0^j (1 - \hat{p}_0)^{1-j} = \\ &= \Lambda_n \sum_{i=0}^1 \sum_{j=0}^1 \hat{p}_1^i (1 - \hat{p}_1)^{1-i} \hat{p}_2^j (1 - \hat{p}_2)^{1-j} = \Lambda_n \cdot 1 = \Lambda_n, \end{aligned}$$

$$M(\Lambda_{n+1} | \Lambda_1, \dots, \Lambda_n) = \Lambda_n, \quad n = 0, 1, 2, \dots$$

Отже, послідовність відношень ймовірностей  $\{\Lambda_n\}$  є мартингалом при справедливій нульовій гіпотезі  $H_0 : p_1 - p_2 = 0$ .

Послідовність відношень ймовірностей  $\{\Lambda_n\}$  задовольняє нерівності Дуба для ймовірності перевищення випадковим процесом заданого рівня. Для будь-якого  $\lambda > 0$  та  $0 \leq n < \infty$  справедлива нерівність:

$$P \left\{ \max_{0 \leq k \leq n} \Lambda_k \geq \lambda \right\} \leq \frac{M \Lambda_n}{\lambda},$$

зокрема, якщо  $M \Lambda_n = M \Lambda_0 = 1$ , то

$$0 < P \left\{ \max_{0 \leq k \leq n} \Lambda_k \geq \lambda \right\} \leq \frac{1}{\lambda}.$$

**Послідовна перевірка простої гіпотези відносно нескінченно великого числа альтернативних гіпотез.** Побудуємо послідовний критерій відношення ймовірностей для перевірки нульової гіпотези  $H_0 : p_1 - p_2 = 0$  відносно сім'ї альтернативних гіпотез  $H_\theta : p_1 - p_2 = \theta, \theta \neq 0$ .

Нехай розподіл  $p_{0n}$  – спільний розподіл випадкових величин  $(X_1, Y_1), \dots, (X_n, Y_n)$  при справедливій нульовій гіпотезі  $H_0$ :

$$p_{0n} = L(0; x_1, \dots, x_n, y_1, \dots, y_n),$$

розподіл  $p_{1n}$  визначається як зважене середнє спільних розподілів випадкових величин  $(X_1, Y_1), \dots, (X_n, Y_n)$ , що відповідають різним значенням параметра  $\theta$  в області  $\omega_0$  відхилення нульової гіпотези  $H_0$ :

$$p_{1n} = \int_{\omega_0} L(\theta; x_1, \dots, x_n, y_1, \dots, y_n) \omega(\theta) d\theta,$$

де  $\omega(\theta)$  – вагова функція з властивостями:  $\omega(\theta) \geq 0$  для всіх  $\theta$  в області  $\omega_0$  відхилення нульової гіпотези, і

$$\int_{\omega_0} \omega(\theta) d\theta = 1.$$

Статистика послідовного критерію відношення ймовірностей дорівнює:

$$\begin{aligned} \frac{p_{1n}}{p_{0n}} &= \frac{\int L(\theta; x_1, \dots, x_n, y_1, \dots, y_n) \omega(\theta) d\theta}{L(0; x_1, \dots, x_n, y_1, \dots, y_n)} \\ &= \frac{\int_{\omega_0} \frac{L(\theta; x_1, \dots, x_n, y_1, \dots, y_n)}{L(p_1, p_2; x_1, \dots, x_n, y_1, \dots, y_n)} \omega(\theta) d\theta}{\frac{L(0; x_1, \dots, x_n, y_1, \dots, y_n)}{L(p_1, p_2; x_1, \dots, x_n, y_1, \dots, y_n)}} = \frac{\int \Lambda(\theta; x_1, \dots, x_n, y_1, \dots, y_n) \omega(\theta) d\theta}{\Lambda(0; x_1, \dots, x_n, y_1, \dots, y_n)}. \end{aligned}$$

Гіпотеза  $H_0$  відхиляється, якщо

$$\frac{p_{1n}}{p_{0n}} \geq \frac{1 - \beta}{\alpha},$$

гіпотеза  $H_0$  не відхиляється, якщо

$$\frac{p_{1n}}{p_{0n}} \leq \frac{\beta}{1 - \alpha},$$

і проводиться додаткове спостереження, якщо

$$\frac{\beta}{1 - \alpha} < \frac{p_{1n}}{p_{0n}} < \frac{1 - \beta}{\alpha}.$$

**Байєсівське оцінювання невідомого параметра  $\sigma^2$ .** Нехай вагова функція  $\omega(\theta)$  задається щільністю нормального розподілу з параметрами  $(0, \sigma^2)$ . Розглянемо  $\theta_1, \theta_2, \dots, \theta_n$  – реалізації незалежних, нормально розподілених випадкових величин з відомим середнім  $\mu_0 = 0$  та невідомою дисперсією  $\sigma^2$ . Функція максимальної правдоподібності параметру  $\sigma^2$  дорівнює

$$\begin{aligned} L(\theta_1, \dots, \theta_n; \sigma^2) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(\theta_i - \mu_0)^2}{2\sigma^2}\right\} \\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (\theta_i - \bar{\theta} + \bar{\theta} - \mu_0)^2\right\} \\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n ((\theta_i - \bar{\theta})^2 + (\bar{\theta} - \mu_0)^2)\right\} \\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left\{-\frac{1}{2\sigma^2} (ns^2 + n(\bar{\theta} - \mu_0)^2)\right\} \\ &\propto (\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2} (ns^2 + n(\bar{\theta} - \mu_0)^2)\right\}, \end{aligned}$$

де  $\bar{\theta}$  – вибіркове середнє,  $s^2$  – вибіркова дисперсія.



Нехай апіорна щільність розподілу параметру  $\sigma^2$  задається щільністю оберненого гамма розподілу з параметрами  $(\alpha_0, \beta_0)$ :

$$\pi(\sigma^2) = \frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)} (\sigma^2)^{-(\alpha_0+1)} \exp\left\{-\frac{\beta_0}{\sigma^2}\right\}.$$

Тоді апостеріорна щільність розподілу параметру  $\sigma^2$  дорівнює

$$\begin{aligned} \pi(\sigma^2 | x_1, \dots, x_n) &\propto L(\theta_1, \dots, \theta_n; \sigma^2) \pi(\sigma^2) \\ &\propto (\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2} (ns^2 + n(\bar{\theta} - \mu_0)^2)\right\} (\sigma^2)^{-(\alpha_0+1)} \exp\left\{-\frac{\beta_0}{\sigma^2}\right\} \\ &\propto (\sigma^2)^{-(\alpha_0+n/2+1)} \exp\left\{-\frac{1}{\sigma^2} \left(\beta_0 + \frac{ns^2 + n(\bar{\theta} - \mu_0)^2}{2}\right)\right\}. \end{aligned}$$

Здійснено перехід від апіорної щільності оберненого гамма розподілу з параметрами  $(\alpha_0, \beta_0)$  до апостеріорної щільності оберненого гамма розподілу з параметрами

$$\left(\alpha_0 + \frac{n}{2}, \beta_0 + \frac{ns^2 + n(\bar{\theta} - \mu_0)^2}{2}\right).$$

Отже, байєсівська точкова оцінка параметру  $\sigma^2$

$$\hat{\sigma}^2 = \beta_0 (\alpha_0 + 1)^{-1}$$

після спостереження  $n$  пар послідовних дій відвідувачів дорівнює:

$$\hat{\sigma}^2 = \left(\beta_0 + \frac{ns^2 + n(\bar{\theta} - \mu_0)^2}{2}\right) \left(\alpha_0 + \frac{n}{2} + 1\right)^{-1}.$$

Тоді статистика послідовного критерію відношення ймовірностей

$$\frac{p_{1n}}{p_{0n}} = \frac{\int_{-\infty}^{+\infty} \exp\left\{-\frac{(\hat{p}_1 - \hat{p}_2 - \theta)^2}{2\hat{p}_0(1-\hat{p}_0)/n}\right\} \cdot \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} \exp\left\{-\frac{\theta^2}{2\hat{\sigma}^2}\right\} d\theta}{\exp\left\{-\frac{(\hat{p}_1 - \hat{p}_2)^2}{2\hat{p}_0(1-\hat{p}_0)/n}\right\}}$$

набуває вигляду:

$$\begin{aligned} \frac{p_{1n}}{p_{0n}} &= \sqrt{\frac{V}{V + \hat{\sigma}^2}} \exp\left\{\frac{1}{2} \frac{(\hat{p}_1 - \hat{p}_2)^2 \hat{\sigma}^2}{(V + \hat{\sigma}^2)V}\right\}, \\ V &= \frac{\hat{p}_0(1-\hat{p}_0)}{n}, \quad \hat{p}_0 = \frac{\hat{p}_1 + \hat{p}_2}{2}. \end{aligned}$$

Згідно з нерівністю Дуба помилка I роду контролюється протягом послідовного тестування:

$$P\left\{\max_{0 \leq k \leq n} \frac{p_{1k}}{p_{0k}} \geq \frac{1}{\alpha}\right\} \leq \alpha, \quad \alpha > 0.$$

Тоді момент зупинки спостережень набуває вигляду

$$\inf \left\{ n > 0 : \frac{p_{1n}}{p_{0n}} < \frac{1}{\alpha} \right\},$$

і нульова гіпотеза  $H_0 : p_1 - p_2 = 0$  відхиляється.

**Байєсівське оцінювання невідомих параметрів  $\mu$  та  $\sigma^2$ .** Нехай вагова функція  $\omega(\theta)$  задається щільністю нормального розподілу з параметрами  $(\mu, \sigma^2)$ . Розглянемо  $\theta_1, \theta_2, \dots, \theta_n$  – реалізації незалежних, нормально розподілених випадкових величин з невідомим середнім  $\mu$  та невідомою дисперсією  $\sigma^2$ . Функція максимальної правдоподібності параметрів  $(\mu, \sigma^2)$  дорівнює

$$\begin{aligned} L(\theta_1, \dots, \theta_n; \mu, \sigma^2) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(\theta_i - \mu)^2}{2\sigma^2}\right\} \\ &\propto (\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}(ns^2 + n(\bar{\theta} - \mu)^2)\right\} \\ &\propto \tau^{n/2} \exp\left\{-\frac{\tau}{2}(ns^2 + n(\bar{\theta} - \mu)^2)\right\}, \end{aligned}$$

де  $\bar{\theta}$  – вибіркове середнє,  $s^2$  – вибіркова дисперсія,  $\tau = \sigma^{-2}$ .

Нехай апіорна щільність розподілу параметрів  $(\mu, \tau)$  задається щільністю оберненого гамма розподілу з параметрами  $(\mu_0, \lambda_0, \alpha_0, \beta_0)$ :

$$\begin{aligned} \pi(\mu, \tau) &= \frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)} \tau^{\alpha_0-1} \exp\{-\beta_0\tau\} \frac{\sqrt{\lambda_0\tau}}{\sqrt{2\pi}} \exp\left\{-\frac{\lambda_0\tau(\mu - \mu_0)^2}{2}\right\} \\ &\propto \tau^{\alpha_0-1/2} \exp\{-\beta_0\tau\} \exp\left\{-\frac{\lambda_0\tau(\mu - \mu_0)^2}{2}\right\}. \end{aligned}$$

Тоді апостеріорна щільність розподілу параметрів  $(\mu, \tau)$  обчислюється згідно з теоремою Байєса:

$$\begin{aligned} \pi(\mu, \tau | \theta_1, \dots, \theta_n) &\propto L(\theta_1, \dots, \theta_n; \mu, \tau) \pi(\mu, \tau) \\ &\propto \tau^{n/2} \exp\left\{-\frac{\tau}{2}(ns^2 + n(\bar{\theta} - \mu)^2)\right\} \tau^{\alpha_0-1/2} \exp\{-\beta_0\tau\} \exp\left\{-\frac{\lambda_0\tau(\mu - \mu_0)^2}{2}\right\} \\ &\propto \tau^{n/2+\alpha_0-1/2} \exp\left\{-\tau\left(\frac{ns^2}{2} + \beta_0\right)\right\} \exp\left\{-\frac{\tau}{2}(\lambda_0(\mu - \mu_0)^2 + n(\bar{\theta} - \mu)^2)\right\} \\ &\propto \tau^{n/2+\alpha_0-1/2} \exp\left\{-\tau\left(\beta_0 + \frac{ns^2}{2} + \frac{\lambda_0 n(\bar{\theta} - \mu_0)^2}{2(\lambda_0 + n)}\right)\right\} \exp\left\{-\frac{\tau}{2}(\lambda_0 + n)\left(\mu - \frac{\lambda_0\mu_0 + n\bar{\theta}}{\lambda_0 + n}\right)^2\right\}. \end{aligned}$$

Здійснено перехід від апіорної щільності гамма нормального розподілу з параметрами  $(\mu_0, \lambda_0, \alpha_0, \beta_0)$  до апостеріорної щільності гамма нормального розподілу з параметрами

$$\left( \frac{\lambda_0 \mu_0 + n \bar{\theta}}{\lambda_0 + n}, \lambda_0 + n, \alpha_0 + \frac{n}{2}, \beta_0 + \frac{ns^2}{2} + \frac{\lambda_0 n (\bar{\theta} - \mu_0)^2}{2(\lambda_0 + n)} \right).$$

Отже, байєсівські точкові оцінки параметрів  $(\mu, \tau)$

$$\hat{\mu} = \mu_0, \hat{\tau} = \alpha_0 \beta_0^{-1}$$

після спостереження  $n$  пар послідовних дій відвідувачів дорівнюють:

$$\hat{\mu} = \frac{\lambda_0 \mu_0 + n \bar{\theta}}{\lambda_0 + n},$$

$$\hat{\tau} = \left( \alpha_0 + \frac{n}{2} \right) \left( \beta_0 + \frac{ns^2}{2} + \frac{\lambda_0 n (\bar{\theta} - \mu_0)^2}{2(\lambda_0 + n)} \right)^{-1}.$$

Байєсівська оцінка  $\hat{\mu}$  є зваженим середнім апріорного середнього  $\mu_0$  різниці коефіцієнтів конверсії та вибіркового середнього  $\bar{\theta}$  різниці коефіцієнтів конверсії. Параметр  $\lambda_0$  можна розглядати як обсяг вибірки, необхідний для обчислення  $\mu_0$ , а  $n$  – обсяг вибірки, необхідний для обчислення  $\bar{\theta}$ . Байєсівська оцінка  $\hat{\tau}$  є відношенням параметрів апостеріорного гамма розподілу.

Статистика послідовного критерію відношення ймовірностей

$$\frac{p_{1n}}{p_{0n}} = \frac{\int_{-\infty}^{+\infty} \exp \left\{ -\frac{(\hat{p}_1 - \hat{p}_2 - \theta)^2}{2\hat{p}_0(1 - \hat{p}_0)/n} \right\} \cdot \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} \exp \left\{ -\frac{(\theta - \hat{\mu})^2}{2\hat{\sigma}^2} \right\} d\theta}{\exp \left\{ -\frac{(\hat{p}_1 - \hat{p}_2 - \hat{\mu})^2}{2\hat{p}_0(1 - \hat{p}_0)/n} \right\}}$$

набуває вигляду:

$$\frac{p_{1n}}{p_{0n}} = \sqrt{\frac{V}{V + \hat{\sigma}^2}} \exp \left\{ \frac{1}{2} \frac{(\hat{p}_1 - \hat{p}_2 - \hat{\mu})^2 \hat{\sigma}^2}{(V + \hat{\sigma}^2)V} \right\},$$

$$V = \frac{\hat{p}_0(1 - \hat{p}_0)}{n}, \hat{p}_0 = \frac{\hat{p}_1 + \hat{p}_2}{2}.$$

Згідно з нерівністю Дуба помилка I роду контролюється протягом послідовного тестування:

$$P \left\{ \max_{0 \leq k \leq n} \frac{p_{1k}}{p_{0k}} \geq \frac{1}{\alpha} \right\} \leq \alpha, \quad \alpha > 0.$$

Тоді момент зупинки спостережень набуває вигляду

$$\inf \left\{ n > 0 : \frac{p_{1n}}{p_{0n}} < \frac{1}{\alpha} \right\},$$

і нульова гіпотеза  $H_0 : p_1 - p_2 = 0$  відхиляється.

**Висновки.** В роботі запропоновано узагальнення послідовного критерію відношення ймовірностей для перевірки простої гіпотези відносно множини альтернативних гіпотез при розв'язанні задачі оптимізації коефіцієнту конверсії, при цьому застосовано два статистичних методи, які широко використовуються для аналізу даних рандомізованих онлайн експериментів: послідовний аналіз Вальда та байєсівський метод статистичного оцінювання невідомих параметрів.

Послідовний аналіз Вальда є методом перевірки статистичних гіпотез, характерною рисою якого є те, що кількість спостережень, необхідних в процесі перевірки, не задана наперед, а залежить від результату самих спостережень, тобто є випадковою величиною, на відміну від традиційних критеріїв. Перевагою даного методу є те, що він потребує, меншої кількості спостережень, ніж рівносильна йому перевірка з фіксованою кількістю спостережень.

Байєсівський метод оцінювання дає суттєвий вигравш в точності при обмежених обсягах вибірок в порівнянні з традиційним частотним підходом. Об'єднання апріорної інформації з емпіричною інформацією послідовних дій користувачів сайту або додатку, дозволяє останній домінувати в апостеріорному розподілі параметрів.

#### Бібліографічні посилання

1. Kohavi R., Henne R.M., Sommerfield D. Practical guide to controlled experiments on the web: listen to your customers not to the hippo. *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*. Pp. 959–967. ACM, 2007.
2. Kohavi R., Longbotham R., Sommerfield D., Henne R. M. Controlled experiments on the web: survey and practical guide. *Data mining and knowledge discovery*. Vol. 18. №. 1. Pp. 140–181. 2009.
3. Kohavi R., Deng A., Frasca B., Walker T. Online controlled experiments at large scale. *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1168–1176. ACM. 2013.
4. Kohavi R., Deng A., Longbotham R., Xu Y. Seven rules of thumb for web site experimenters. *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. 2014.
5. Kohavi R. Tang D., Xu Y. Trustworthy Online Controlled Experiments: A Practical Guide to A/B Testing. *Cambribge University Press*. 2020.
6. Bondarenko Ya.S., Kravchenko S.V. On the frequentist approach to multivariate landing page testing (English). *Visnyk DNU. Series: Modelling*. Dnipro: DNU. Issue 9. No. 8. Pp. 142–151. 2017.
7. Wald A. Sequential tests of statistical hypotheses. *Ann. Math. Statist.* Vol. 16. № 2. Pp. 117–186. 1945.
8. Robbins H. Statistical methods related to the law of the iterated logarithm. *The Annals of Mathematical Statistics*. 41(5): 1397-1409.
9. Pekelis L., Walsh D., Johari R. The new stats engine. *Technical report, Optimizely*. 2015.
10. Johari R., Koomen P., Pekelis L., Walsh D. Peeking at a/b tests: Why it matters, and what to do about it. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Pp. 1517–1525. 2017.
11. Johari R., Pekelis L., Walsh D. Always valid inference: Bringing sequential analysis to A/B testing. arXiv:1512.04922v3. 2019.

12. Miller E. Simple sequential a/b testing. URL <http://www.evanmiller.org/sequential-ab-testing.html>, blog post. 2015.
13. Abhishek V., Mannor S., Abhishek V. A nonparametric sequential test for online randomized experiments. arXiv: 1610.02490v4. 2017.
14. Stucchio C. Bayesian A/B Testing at VWO. *Technical report, Visual Website Optimizer*, 2015.
15. Miller E. Formulas for Bayesian A/B testing. URL <http://www.evanmiller.org/bayesian-ab-testing.html>, blog post. 2015.
16. Bernardo J.M., Smith A.F.M. Bayesian Theory. Wiley, 1994.
17. Bondarenko Ya.S., Kravchenko S.V. Bayesian A/B/C testing. *The Fifth Baltic-Nordic Conference on Survey Statistics*, 16-20 June 2019 in Örebro, Sweden.
18. Bondarenko Ya.S., Kravchenko S.V. Bayesian approach to landing page testing. *Питання прикладної математики та математичного моделювання*. Д.: Вид-во ДНУ, 2019. С. 3-16.
19. Bondarenko Ya. Sequential A/B testing. *Proceedings of the 2019 IEEE International Conference on Advanced Trends in Information Theory*, 18-20.12.2019. Kyiv. Ukraine.
20. Bondarenko Ya. A sequential probability ratio test for online experiments. *Proceedings of the Summer School of Survey Statistics 2021*. Statistics Lithuania. Vilnius, 2021. Pp. 35-39.

Надійшла до редколегії 22.04. 2021.