УДК 519.226 doi: 10.15421/321901

Ya.S. Bondarenko, S.V. Kravchenko
Department of Statistics and Probability Theory
Oles Honchar Dnipro National University

BAYESIAN APPROACH TO LANDING PAGE TESTING

В роботі байєсівський метод статистичного оцінювання застосовується до оптимізації коефіцієнта конверсії. Запропоновано методологію байєсівського A/B/C тестування цільової сторінки сайту, при цьому математичне сподівання функції втрат знайдено аналітично. Результати байєсівського A/B/C тестування представлено графічно та описово.
Ключові слова: A/B/C тестування, байєсівський метод статистичного оцінювання, оптимізація коефіцієнта конверсії, математичне сподівання функції втрат, оптимізація цільової сторінки, функція втрат, щільність розподілу.

В работе байесовский метод статистического оценивания применяется для оптимизации коэффициента конверсии. Предложена методология байесовского A/B/C тестирования целевой страницы сайта, при этом математическое ожидание функции потерь найдено аналитически. Результаты байесовского A/B/C тестирования представлены графически и описательно.
Ключевые слова: A/B/C тестирование, байесовский метод статистического оценивания, оптимизация коэффициента конверсии, математическое ожидание функции потерь, оптимизация целевой страницы, функция потерь, плотность распределения.

In this paper a Bayesian inference to conversion rate optimization is considered. Bayesian A/B/C testing methodology with the expected value of the loss function computed analytically is proposed. Bayesian A/B/C testing results are presented graphically and descriptively.
Keywords: A/B/C testing, Bayesian inference, conversion rate optimization, expected loss, landing page optimization, loss function, probability density function.

1. Introduction

Each company directs its activity to generate profit. It is important for the company to have an effective website at the time when a significant part of the target audience receives information about the products and services via the Internet. Company website must be constantly optimized to increase profit. The object of optimization is conversion rate (percent of website visitors who engage in desired target action for the website owner). A/B, A/B/C, and multivariate landing page testing are the ways to conversion rate optimization. A/B testing is an experiment of showing two variants of the landing page to website visitors at the same time and comparing which variants give more conversion actions. A/B/C testing is an experiment of showing three variants of the landing page to website visitors, and statistical analysis is used to determine which one performs better for a given conversion aim. In multivariate landing page testing, several elements are selected on the landing page and their modifications are created. All kinds of combinations for these modified elements are further created, therefore landing page variations are generated. Website visitors are equally divided between landing page variations. A variation is suggested viewing for each visitor and visitor's behavior is monitored. All possible combinations of page elements simultaneously are tested. It makes possible to assess the effect of each element and their interaction to conversion rate. Landing page variation (the optimal combination of elements) is selected depending on the outcome of the testing. For more details, see [1], [2], [3] and [4].

Moving away from classical hypothesis testing, like the t-test, to sequential testing developed by A. Wald [5] and by M. Girshick [6] was presented in papers [7], [8], [9], [10], [11]. The number of visitors required for landing page testing isn't fixed in advanced, conversions are evaluated as they are collected, and testing is stopped in accordance with a predefined stopping rule as soon as the significant result is observed. But experimental design without prior information about the statistical properties of the observed random variable occurs rarely. Classical methods of unbiased estimation and maximum likelihood method don't give insight adjusted for prior knowledge. Theory of Bayesian estimation allows combining prior information with observed random variables. The main difference between Bayesian and Frequentist approaches is that parameters of the distribution are considered as random variables in the first case and are defined as constant variables in the second one. Mathematical models of Bayesian testing were developed in [12]. The expected value of the loss function was proposed to calculate numerically and arising problems were described in this case.

In this paper Bayesian A/B/C testing implementation with the expected value of the loss function computed analytically is proposed. The paper consists of 5 sections. Section 2 is devoted to the main definitions. Posterior distributions for probabilities of success are considered. Section 3 deals with the analytical solution for expected value of the loss function. Bayesian testing implementation is described in section 4 and Bayesian testing results are discussed in section 5.

2. The posterior distribution for the probability of success

Bernoulli trials with probabilities $\theta_A,\theta_B,\theta_C$ of success are conducted in three groups of visitors (A, B, C). Probabilities $\theta_A,\theta_B,\theta_C$ of success are unknown random variables. Let $p(\theta_A), p(\theta_B), p(\theta_C)$ be the prior probability density functions for $\theta_A,\theta_B,\theta_C$, respectively. Let $p(\theta_A|x_1,x_2,\ldots,x_n)$, $p(\theta_B|y_1,y_2,\ldots,y_n)$, $p(\theta_C|z_1,z_2,\ldots,z_n)$ be the posterior probability density functions for $\theta_A,\theta_B,\theta_C$ given the sample vectors $x=(x_1,x_2,\ldots,x_n)$, $y=(y_1,y_2,\ldots,y_n)$, $z=(z_1,z_2,\ldots,z_n)$ are observed. Our goal is to find Bayesian estimators for probabilities $\theta_A,\theta_B,\theta_C$ of success concerning the loss function $L(\theta_A,\theta_B,\theta_C,\cdot)$.

Bernoulli trials with two possible outcomes (success $\{x=1\}$ and failure $\{x=0\}$) are conducted in group A of visitors. The number of successes in one trial has Bernoulli distribution with parameter θ_A:

$$P(x,\theta_A)=\theta_A^x(1-\theta_A)^{1-x},\ 0<\theta_A<1.$$

The function θ_A defined by:

$$p(x_1,x_2,\ldots,x_n|\theta_A)=\theta_A^{\sum_{i=1}^{n}x_i}(1-\theta_A)^{n-\sum_{i=1}^{n}x_i},\ i=1,\ldots,n,$$

is the likelihood function. The prior information about probability θ_A of success determined by:

$$p(\theta_A)=\frac{\theta_A^{a-1}(1-\theta_A)^{b-1}}{B(a,b)},\ 0\le\theta_A\le 1,$$

is Beta distribution with parameters (a,b).

According to Bayes' theorem, the posterior distribution for probability θ_A of success is given by

$$p(\theta_A|x_1,x_2,\ldots,x_n)=\frac{p(\theta_A)\,p(x_1,x_2,\ldots,x_n|\theta_A)}{\int_0^1 p(\theta_A)\,p(x_1,x_2,\ldots,x_n|\theta_A)\,d\theta_A},$$

$$p(\theta_A|x_1,x_2,\ldots,x_n)=\frac{\theta_A^{a+\sum_{i=1}^{n}x_i-1}(1-\theta_A)^{b+n-\sum_{i=1}^{n}x_i-1}}{B\left(a+\sum_{i=1}^{n}x_i,b+n-\sum_{i=1}^{n}x_i\right)},$$

consequently, the posterior distribution for probability θ_A of success in group A is Beta distribution with parameters $(\tilde{a},\tilde{b})$:

$$p(\theta_A|x_1,x_2,\ldots,x_n)=\frac{\theta_A^{\tilde{a}-1}(1-\theta_A)^{\tilde{b}-1}}{B(\tilde{a},\tilde{b})},$$

where

$$\tilde{a}=a+\sum_{i=1}^{n}x_i,\ \tilde{b}=b+n-\sum_{i=1}^{n}x_i.$$

The posterior distribution for probability θ_B of success in group B is Beta distribution with parameters $(\tilde{c},\tilde{d})$:

$$p(\theta_B|y_1,y_2,\ldots,y_n)=\frac{\theta_B^{\tilde{c}-1}(1-\theta_B)^{\tilde{d}-1}}{B(\tilde{c},\tilde{d})},$$

where

$$\tilde{c}=c+\sum_{i=1}^{n}y_i,\ \tilde{d}=d+n-\sum_{i=1}^{n}y_i.$$

The posterior distribution for probability θ_C of success in group C is Beta distribution with parameters $\left(\tilde{e}, \tilde{f}\right)$:

$$p\left(\theta_C | z_1, z_2, \ldots, z_n\right) = \frac{\theta_C^{\tilde{e}-1}\left(1-\theta_C\right)^{\tilde{f}-1}}{B\left(\tilde{e}, \tilde{f}\right)},$$

where

$$\tilde{e} = e + \sum_{i=1}^{n} z_i, \ \tilde{f} = f + n - \sum_{i=1}^{n} z_i.$$

3. Expected loss $EL\left(\theta_A, \theta_B, \theta_C, \cdot\right)$

The loss function $L\left(\theta_A, \theta_B, \theta_C, \cdot\right)$ describes the loss under decision making about choosing landing page variant which can be published on website. The loss functions are

$$L\left(\theta_A, \theta_B, \theta_C, A\right) = max\left\{\theta_B - \theta_A, \theta_C - \theta_A, 0\right\},$$
$$L\left(\theta_A, \theta_B, \theta_C, B\right) = max\left\{\theta_A - \theta_B, \theta_C - \theta_B, 0\right\},$$
$$L\left(\theta_A, \theta_B, \theta_C, C\right) = max\left\{\theta_A - \theta_C, \theta_B - \theta_C, 0\right\}$$

for group A, B and C, respectively [12].

The expected value of the loss function is computed analytically in this section. The expected loss for group A is determined by:

$$EL\left(\theta_A, \theta_B, \theta_C, A\right) = \int_0^1\int_0^1\int_0^1 L\left(x, y, z, A\right) p\left(x, y, z\right) dxdydz,$$

where $p\left(x, y, z\right)$ is the joint density of $\theta_A, \theta_B, \theta_C$.

By definition of the loss function for group A, we have

$$EL\left(\theta_A, \theta_B, \theta_C, A\right) =$$
$$= \int_0^1 dx \int_x^1 dy \int_0^y \left(y - x\right) \frac{x^{a-1}\left(1-x\right)^{b-1}}{B\left(a,b\right)} \frac{y^{c-1}\left(1-y\right)^{d-1}}{B\left(c,d\right)} \frac{z^{e-1}\left(1-z\right)^{f-1}}{B\left(e,f\right)} dz +$$
$$+ \int_0^1 dx \int_0^x dy \int_x^1 \left(z - x\right) \frac{x^{a-1}\left(1-x\right)^{b-1}}{B\left(a,b\right)} \frac{y^{c-1}\left(1-y\right)^{d-1}}{B\left(c,d\right)} \frac{z^{e-1}\left(1-z\right)^{f-1}}{B\left(e,f\right)} dz +$$
$$+ \int_0^1 dx \int_x^1 dy \int_y^1 \left(z - x\right) \frac{x^{a-1}\left(1-x\right)^{b-1}}{B\left(a,b\right)} \frac{y^{c-1}\left(1-y\right)^{d-1}}{B\left(c,d\right)} \frac{z^{e-1}\left(1-z\right)^{f-1}}{B\left(e,f\right)} dz.$$

The expected value of the loss function for group A has the following form:

$$EL\left(\theta_A, \theta_B, \theta_C, A\right) =$$
$$= \frac{B\left(c+1,d\right)}{B\left(c,d\right)}\left(1 - h\left(a,b,c+1,d\right)\right) - \frac{B\left(a+1,b\right)}{B\left(a,b\right)}\left(1 - h\left(a+1,b,c,d\right)\right) -$$

$$-\frac{B(c+1,d)}{B(c,d)}\sum_{i=0}^{e-1}\frac{B(c+1+i,d+f)}{(f+i)B(1+i,f)B(c+1,d)}(1-h(a,b,c+i+1,d+f)+$$
$$+\frac{B(a+1,b)}{B(a,b)}\sum_{i=0}^{e-1}\frac{B(c+i,d+f)}{(f+i)B(1+i,f)B(c,d)}(1-h(a+1,b,c+i,d+f)+$$
$$+\frac{B(e+1,f)}{B(e,f)}\left(1-h(a,b,e+1,f)\right)-\frac{B(a+1,b)}{B(a,b)}(1-h(a+1,b,e,f)-$$
$$-\frac{B(e+1,f)}{B(e,f)}\sum_{i=0}^{e}\frac{B(a+i,b+f)}{(i+f)B(1+i,f)B(a,b)}\left(1-h(a+i,b+f,c,d)\right)+$$
$$+\frac{B(a+1,b)}{B(a,b)}\sum_{i=0}^{e-1}\frac{B(a+1+i,b+f)}{(i+f)B(1+i,f)B(a+1,b)}\left(1-h(a+1+i,b+f,c,d)\right)+$$
$$+\frac{B(e+1,f)}{B(e,f)}\sum_{i=0}^{e}\frac{B(c+i,d+f)}{(i+f)B(1+i,f)B(c,d)}\left(1-h(a,b,c+i,d+f)\right)-$$
$$-\frac{B(a+1,b)}{B(a,b)}\sum_{i=0}^{e-1}\frac{B(c+i,d+f)}{(i+f)B(1+i,f)B(c,d)}\left(1-h(a+1,b,c+i,d+f)\right).$$

where the function h is equal to:

$$h(a,b,c,d)=1-\sum_{i=0}^{c-1}\frac{B(a+i,b+d)}{(d+i)B(1+i,d)B(a,b)}.$$

Expected loss for group B is determined by:

$$EL(\theta_A,\theta_B,\theta_C,B)=\int_0^1\int_0^1\int_0^1 L(\theta_A,\theta_B,\theta_C,B)\,p(x,y,z)\,dxdydz,$$

where $p(x,y,z)$ is the joint density of $\theta_A,\theta_B,\theta_C$.

By definition of the loss function for group B, we have

$$EL(\theta_A,\theta_B,\theta_C,B)=$$
$$=\int_0^1 dx\int_0^x dy\int_0^x (x-y)\frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}\frac{y^{c-1}(1-y)^{d-1}}{B(c,d)}\frac{z^{e-1}(1-z)^{f-1}}{B(e,f)}dz+$$
$$+\int_0^1 dx\int_0^x dy\int_x^1 (z-y)\frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}\frac{y^{c-1}(1-y)^{d-1}}{B(c,d)}\frac{z^{e-1}(1-z)^{f-1}}{B(e,f)}dz+$$
$$+\int_0^1 dx\int_x^1 dy\int_y^1 (z-y)\frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}\frac{y^{c-1}(1-y)^{d-1}}{B(c,d)}\frac{z^{e-1}(1-z)^{f-1}}{B(e,f)}dz.$$

Expected value of the loss function for group B has the following form:

$$EL(\theta_A,\theta_B,\theta_C,B)=$$

$$=\frac{B(a+1,b)}{B(a,b)}h(a+1,b,c,d)-\frac{B(a+1,b)}{B(a,b)}\left(1-h(a+1,b,e,f)\right)+$$

$$+\frac{B(a+1,b)}{B(a,b)}\sum_{i=0}^{e-1}\frac{B(a+i+1,b+f)}{(i+f)B(i+1,f)B(a+1,b)}\left(1-h(a+i+1,b+f,c,d)\right)-$$

$$-\frac{B(c+1,d)}{B(c,d)}h(a,b,c+1,d)+\frac{B(c+1,d)}{B(c,d)}(1-h(a,b,e,f)-$$

$$-\frac{B(c+1,d)}{B(c,d)}\sum_{i=0}^{e-1}\frac{B(a+i,b+f)}{(i+f)B(i+1,f)B(a,b)}\left(1-h(a+i,b+f,c+1,d)\right)+$$

$$+\frac{B(e+1,f)}{B(e,f)}(1-h(a,b,e+1,f)-\frac{B(c+1,d)}{B(c,d)}\left(1-h(a,b,e,f)\right)-$$

$$-\frac{B(e+1,f)}{B(e,f)}\sum_{i=0}^{e}\frac{B(a+i,b+f)}{(i+f)B(i+1,f)B(a,b)}\left(1-h(a+i,b+f,c,d)\right)+$$

$$+\frac{B(c+1,d)}{B(c,d)}\sum_{i=0}^{e-1}\frac{B(a+i,b+f)}{(i+f)B(i+1,f)B(a,b)}\left(1-h(a+i,b+f,c+1,d)\right)+$$

$$+\frac{B(e+1,f)}{B(e,f)}\sum_{i=0}^{e}\frac{B(c+i,d+f)}{B(c,d)(i+f)B(i+1,f)}\left(1-h(a,b,c+i,d+f)\right)-$$

$$-\frac{B(c+1,d)}{B(c,d)}\sum_{i=0}^{e-1}\frac{B(c+i+1,d+f)}{B(c+1,d)(i+f)B(i+1,f)}\left(1-h(a,b,c+i+1,d+f)\right),$$

where the function h is equal to:

$$h(a,b,c,d)=1-\sum_{i=0}^{c-1}\frac{B(a+i,b+d)}{(d+i)B(1+i,d)B(a,b)}.$$

The expected loss for group C is determined by:

$$EL(\theta_A,\theta_B,\theta_C,C)=\int_0^1\int_0^1\int_0^1 L(\theta_A,\theta_B,\theta_C,C)\,p(x,y,z)\,dxdydz,$$

where $p(x,y,z)$ is the joint density of $\theta_A,\theta_B,\theta_C$.

By definition of the loss function for group C, we have

$$EL(\theta_A,\theta_B,\theta_C,C)=$$

$$=\int_0^1 dx\int_0^x dy\int_0^x (x-z)\frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}\frac{y^{c-1}(1-y)^{d-1}}{B(c,d)}\frac{z^{e-1}(1-z)^{f-1}}{B(e,f)}dz+$$

$$+\int_0^1 dx\int_x^1 dy\int_0^y (y-z)\frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}\frac{y^{c-1}(1-y)^{d-1}}{B(c,d)}\frac{z^{e-1}(1-z)^{f-1}}{B(e,f)}dz.$$

Expected value of the loss function for group C has the following form:

$$EL(\theta_A,\theta_B,\theta_C,C)=$$

$$=\frac{B(a+1,b)}{B(a,b)}h(a+1,b,c,d)-\frac{B(a+1,b)}{B(a,b)}\big(1-h(a+1,b,e,f)\big)+$$

$$+\frac{B(a+1,b)}{B(a,b)}\sum_{i=0}^{e-1}\frac{B(a+i+1,b+f)}{(i+f)B(i+1,f)B(a+1,b)}\big(1-h(a+i+1,b+f,c,d)\big)-$$

$$-\frac{B(e+1,f)}{B(e,f)}+\frac{B(c+1,d)}{B(c,d)}\big(1-h(a,b,c+1,d)\big)+\frac{B(e+1,f)}{B(e,f)}\big(1-h(a,b,e+1,f)\big)-$$

$$-\frac{B(e+1,f)}{B(e,f)}\sum_{i=0}^{e}\frac{B(a+i,b+f)}{(i+f)B(i+1,f)B(a,b)}\big(1-h(a+i,b+f,c,d)\big)-$$

$$-\frac{B(c+1,d)}{B(c,d)}\sum_{i=0}^{e-1}\frac{B(c+i+1,d+f)}{(f+i)B(1+i,f)B(c,d)}\big(1-h(a,b,c+i+1,d+f)\big)-$$

$$+\frac{B(e+1,f)}{B(e,f)}\sum_{i=0}^{e}\sum_{j=0}^{c+i-1}\frac{B(c+i,d+f)B(a+j,d+f+b)}{B(a,b)(i+f)B(i+1,f)B(c,d)(j+d+f)B(j+1,d+f)},$$

where the function h is equal to:

$$h(a,b,c,d)=1-\sum_{i=0}^{c-1}\frac{B(a+i,b+d)}{(d+i)B(1+i,d)B(a,b)}.$$

The foregoing results for three group of visitors A, B, C can be easily converted into results for two groups. Expected value of the loss function for group A has the following form:

$$EL(\theta_A,\theta_B,A)=\frac{B(c+1,d)}{B(c,d)}\big(1-h(a,b,c+1,d)\big)-\frac{B(a+1,b)}{B(a,b)}\big(1-h(a+1,b,c,d)\big)$$

and expected value of the loss function for group B has the following form:

$$EL(\theta_A,\theta_B,B)=\frac{B(a+1,b)}{B(a,b)}h(a+1,b,c,d)-\frac{B(c+1,d)}{B(c,d)}h(a,b,c+1,d).$$

4. Bayesian testing implementation

Landing page variants A, B, C are suggested viewing for the first, second and third group of visitors. We need to identify users during testing for clear experiment and suggest them the same landing page variant that they have viewed earlier in case of repeated visits. The flow of visitors has been simulated. Each visitor can belong to the first, second and third group with probability $1/3$. Visitor behavior is simulated after identification of visitor belonging to group. Visitor behavior is determined with two outcomes: success – conversion action is done, failure – conversion action isn't done. If visitor belongs to the first group, success will happen with probability θ_A, and if visitor belongs to the second group, success will happen with probability θ_B, and if visitor belongs to the third group, success will happen with probability θ_C. Prior information about unknown distribution for probabilities θ_A, θ_B and θ_C of success is defined with Beta distribution with parameters $a=1, b=1$ and $c=1, d=1$, and $e=1, f=1$, respectively.

According to Bayes' theorem, posterior distribution for probability θ_A of success is Beta-distribution with $\left(\tilde{a},\tilde{b}\right)$ parameters:

$$\tilde{a}=a+x_i,\ \tilde{b}=b+\left(1-x_i\right),$$

where x_i is number of successes in one trial $\left(x_i=0,1\right)$, the posterior distribution for probability θ_B of success is Beta-distribution with $\left(\tilde{c},\tilde{d}\right)$ parameters:

$$\tilde{c}=c+y_i,\ \tilde{d}=d+\left(1-y_i\right),$$

where y_i is number of successes in one trial $\left(y_i=0,1\right)$,
the posterior distribution for probability θ_C of success is Beta-distribution with $\left(\tilde{e},\tilde{f}\right)$ parameters:

$$\tilde{e}=e+z_i,\ \tilde{f}=f+\left(1-z_i\right),$$

where z_i is number of successes in one trial $\left(z_i=0,1\right)$.

Expected loss $L\left(\theta_A,\theta_B,\theta_C,\cdot\right)$ compares with threshold of loss ε after each visit of landing page variant A, B, C. If

$$EL\left(\hat{\theta}_A,\hat{\theta}_B,\hat{\theta}_C,A\right)\le\varepsilon, EL\left(\hat{\theta}_A,\hat{\theta}_B,\hat{\theta}_C,B\right)>\varepsilon, EL\left(\hat{\theta}_A,\hat{\theta}_B,\hat{\theta}_C,C\right)>\varepsilon,$$

then testing will be stopped, landing page variant A will be chosen for publishing on website, if

$$EL\left(\hat{\theta}_A,\hat{\theta}_B,\hat{\theta}_C,A\right)>\varepsilon, EL\left(\hat{\theta}_A,\hat{\theta}_B,\hat{\theta}_C,B\right)\le\varepsilon, EL\left(\hat{\theta}_A,\hat{\theta}_B,\hat{\theta}_C,C\right)>\varepsilon,$$

then testing will be stopped, landing page variant B will be chosen for publishing on website, if

$$EL\left(\hat{\theta}_A,\hat{\theta}_B,\hat{\theta}_C,A\right)>\varepsilon, EL\left(\hat{\theta}_A,\hat{\theta}_B,\hat{\theta}_C,B\right)>\varepsilon, EL\left(\hat{\theta}_A,\hat{\theta}_B,\hat{\theta}_C,C\right)\le\varepsilon,$$

then testing will be stopped, landing page variant C will be chosen for publishing on website [12].

Likelihood that probability of success in one group is greater than the probability of success in another two groups is defined by (see [13])

$$P\left\{\theta_A>\max\left\{\theta_B,\theta_C\right\}\right\}=1-P\left\{\theta_B>\theta_A\right\}-P\left\{\theta_C>\theta_A\right\}+$$
$$+\sum_{i=0}^{c-1}\sum_{j=0}^{e-1}\frac{B\left(i+j+a,b+d+f\right)}{\left(d+i\right)\left(f+j\right)B\left(1+i,d\right)B\left(1+j,f\right)B\left(a,b\right)},$$

$$P\left\{\theta_B > \max\left\{\theta_A, \theta_C\right\}\right\} = 1 - P\left\{\theta_A > \theta_B\right\} - P\left\{\theta_C > \theta_B\right\} +$$
$$+\sum_{i=0}^{a-1}\sum_{j=0}^{e-1}\frac{B\left(i+j+c, b+d+f\right)}{\left(b+i\right)\left(f+j\right)B\left(1+i,b\right)B\left(1+j,f\right)B\left(c,d\right)},$$
$$P\left\{\theta_C > \max\left\{\theta_A, \theta_B\right\}\right\} = 1 - P\left\{\theta_A > \theta_C\right\} - P\left\{\theta_B > \theta_C\right\} +$$
$$+\sum_{i=0}^{a-1}\sum_{j=0}^{c-1}\frac{B\left(i+j+e, b+d+f\right)}{\left(b+i\right)\left(d+j\right)B\left(1+i,b\right)B\left(1+j,d\right)B\left(e,f\right)}.$$

5. Bayesian testing results

Minimally-informative prior densities are shown in Figure 1. Prior densities for probability $\theta_A, \theta_B, \theta_C$ of success are represented with blue, red and green color, respectively. Likelihoods that probability of success in one group is greater than the probabilities of success in two other groups are equal to

$$P\left\{\hat{\theta}_A > \max\left\{\hat{\theta}_B, \hat{\theta}_C\right\}\right\} = P\left\{\hat{\theta}_B > \max\left\{\hat{\theta}_A, \hat{\theta}_C\right\}\right\} = P\left\{\hat{\theta}_C > \max\left\{\hat{\theta}_A, \hat{\theta}_B\right\}\right\} = 0,33.$$

Bayesian testing results on several iterations are considered. On 21st iteration posterior density for probability θ_A of success (Beta-distribution with $\tilde{a} = 7, \tilde{b} = 3$ parameters) is represented with blue color, posterior density for probability θ_B of success (Beta-distribution with $\tilde{c} = 6, \tilde{d} = 2$ parameters) is represented with red color, posterior density for probability θ_C of success (Beta-distribution with $\tilde{e} = 3, \tilde{f} = 5$ parameters) is represented with green color (see Fig. 2). Bayesian estimators for probabilities $\theta_A, \theta_B, \theta_C$ of success are equal to

$$\hat{\theta}_A = 0,75; \hat{\theta}_B = 0,83; \hat{\theta}_C = 0,33.$$

Likelihoods that probability of success in one group is greater than the probabilities of success in two other groups are equal to

$$P\left\{\hat{\theta}_A > \max\left\{\hat{\theta}_B, \hat{\theta}_C\right\}\right\} = 0,39,$$
$$P\left\{\hat{\theta}_B > \max\left\{\hat{\theta}_A, \hat{\theta}_C\right\}\right\} = 0,60,$$
$$P\left\{\hat{\theta}_C > \max\left\{\hat{\theta}_A, \hat{\theta}_B\right\}\right\} = 0,01.$$

On 39th iteration posterior density for probability θ_A of success (Beta-distribution with $\tilde{a} = 13, \tilde{b} = 6$ parameters) is represented with blue color, posterior density for probability θ_B of success (Beta-distribution with $\tilde{c} = 10, \tilde{d} = 2$ parameters) is represented with red color, posterior density for probability θ_C of success (Beta-distribution with $\tilde{e} = 3, \tilde{f} = 10$ parameters) is represented with green color (see Fig. 4). Bayesian estimators for probabilities $\theta_A, \theta_B, \theta_C$ of success are equal to

$$\hat{\theta}_A = 0,70; \hat{\theta}_B = 0,90; \hat{\theta}_C = 0,18.$$

Likelihoods that probability of success in one group is greater than the probabilities of success in two other groups are equal to

$$P\left\{\hat{\theta}_A > \max\left\{\hat{\theta}_B, \hat{\theta}_C\right\}\right\} = 0,151,$$

$$P\left\{\hat{\theta}_B > \max\left\{\hat{\theta}_A, \hat{\theta}_C\right\}\right\} = 0,848,$$

$$P\left\{\hat{\theta}_C > \max\left\{\hat{\theta}_A, \hat{\theta}_B\right\}\right\} = 0,001.$$

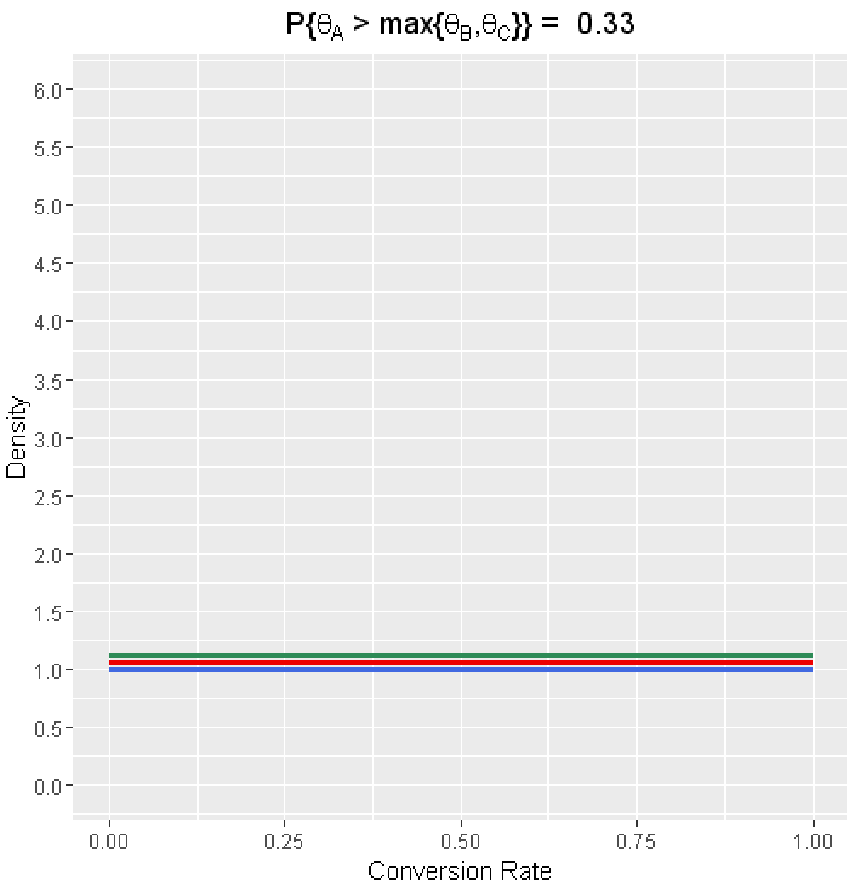

Fig. 1. Prior densities for probabilities $\theta_A, \theta_B, \theta_C$ of success

On 60th iteration posterior density for probability θ_A of success (Beta-distribution with $\tilde{a} = 16, \tilde{b} = 9$ parameters) is represented with blue color, posterior density for probability θ_B of success (Beta-distribution with $\tilde{c} = 14, \tilde{d} = 2$ parameters) is represented with red color, posterior density for probability θ_C of success (Beta-distribution with $\tilde{e} = 7, \tilde{f} = 17$ parameters) is represented with green color (see Fig. 5). Bayesian estimators for probabilities $\theta_A, \theta_B, \theta_C$ of success are equal to $\hat{\theta}_A = 0,65; \hat{\theta}_B = 0,93; \hat{\theta}_C = 0,27$.

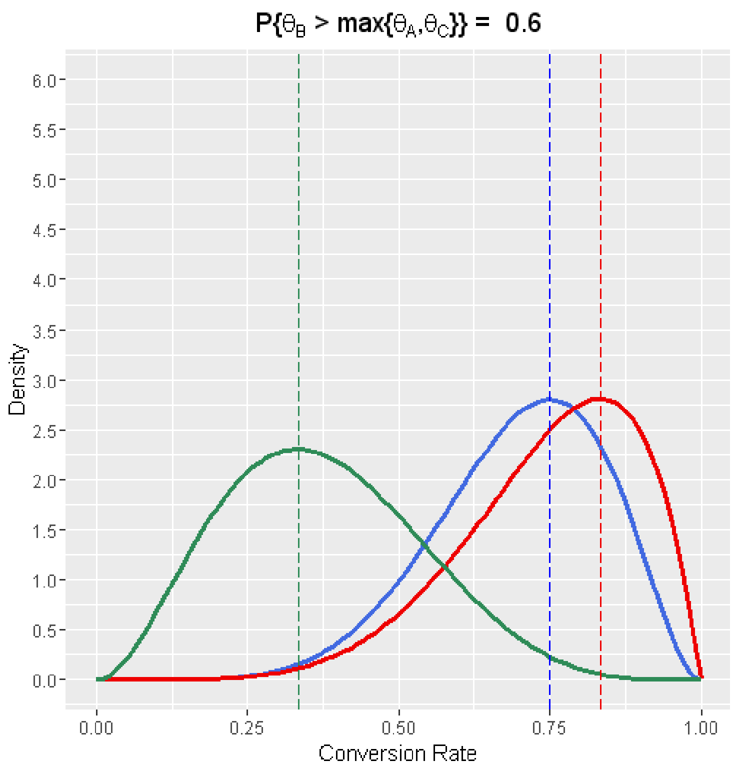

Fig. 2. Posterior densities for probabilities $\theta_A, \theta_B, \theta_C$ of success on 21st iteration

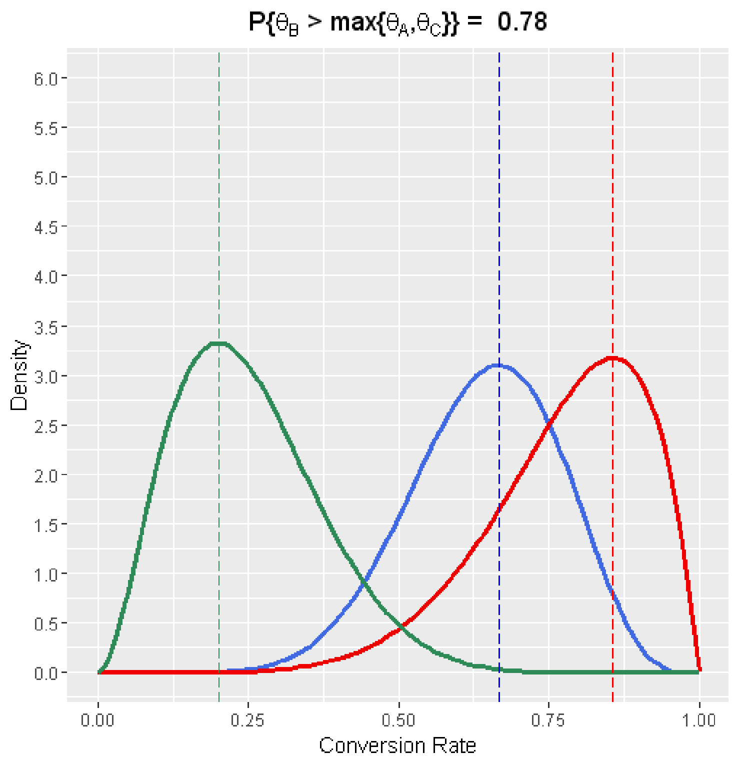

Fig. 3. Posterior densities for probabilities $\theta_A, \theta_B, \theta_C$ of success on 30th iteration

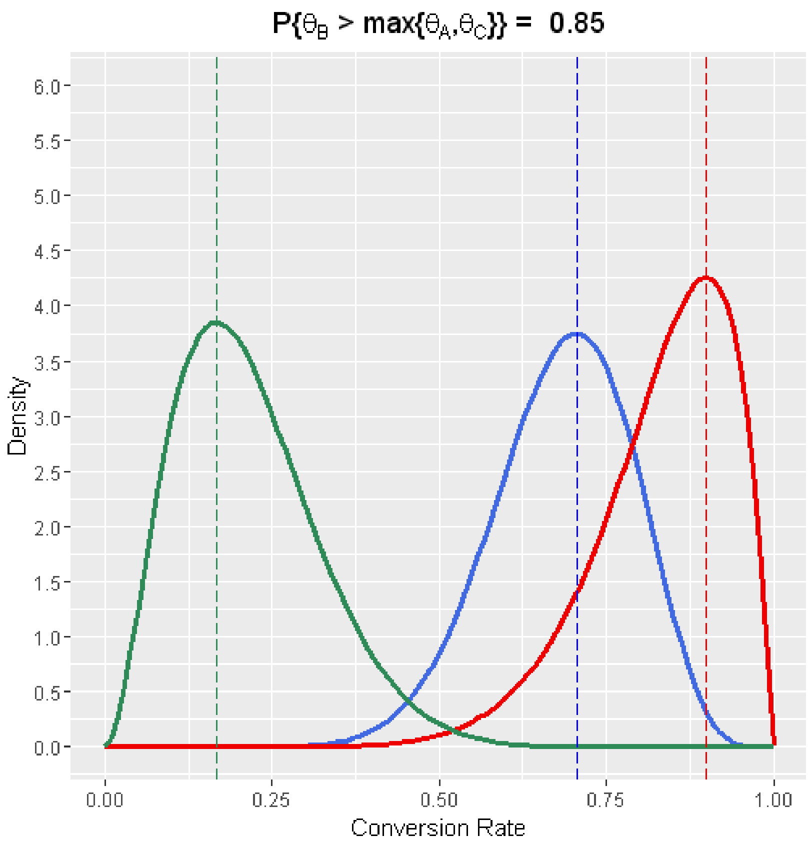

Fig. 4. Posterior densities for probabilities $\theta_A, \theta_B, \theta_C$ of success on 39th iteration

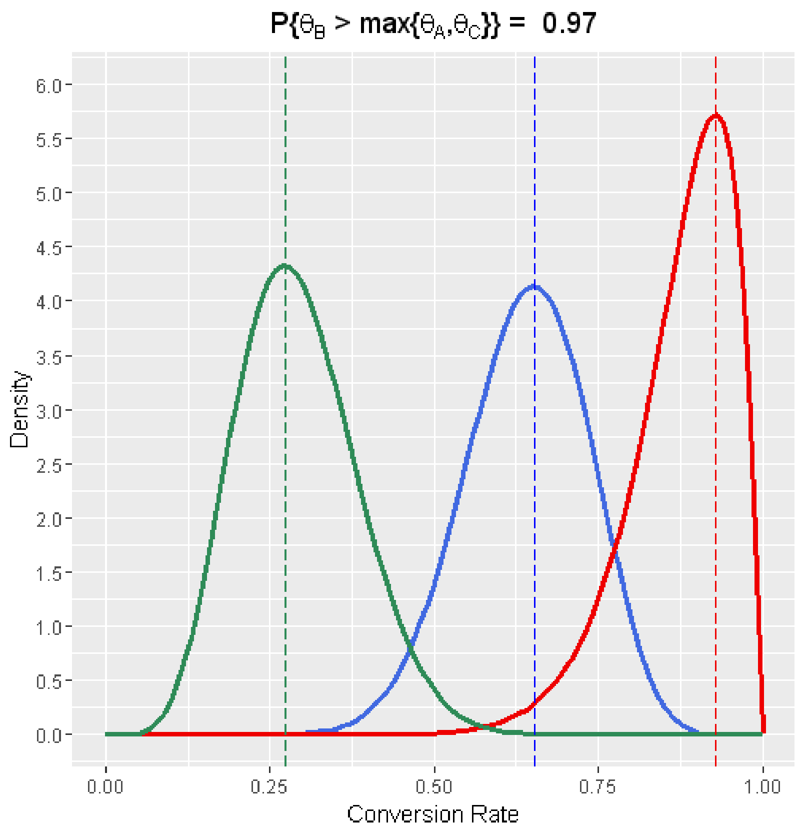

Fig. 5. Posterior densities for probabilities $\theta_A, \theta_B, \theta_C$ of success on 60th iteration

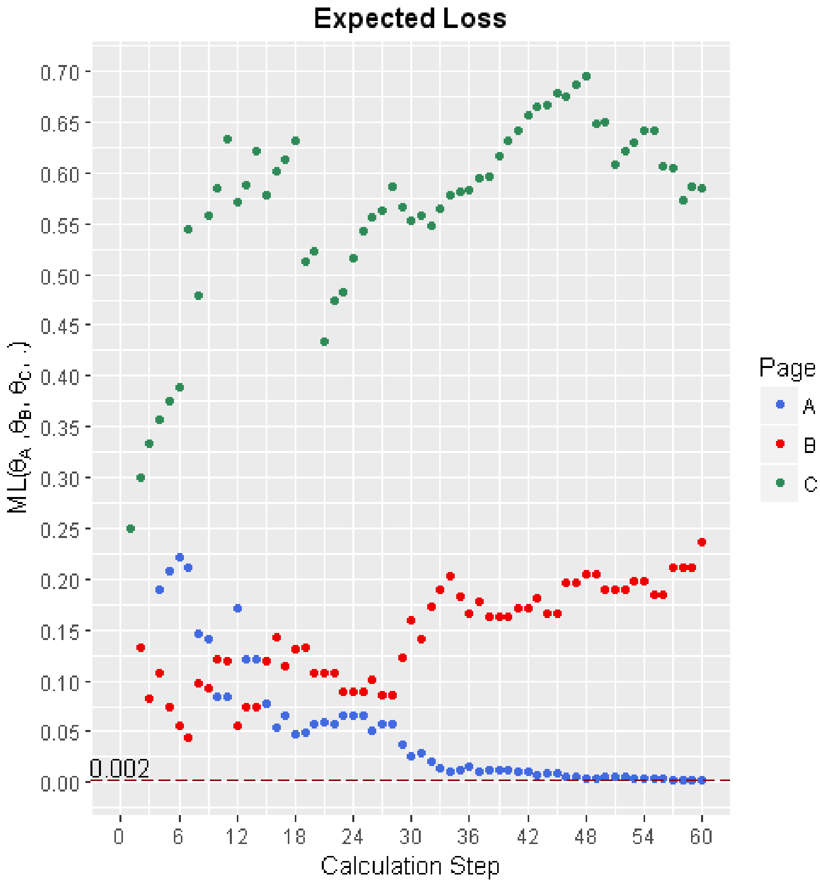

Fig. 6. Expected loss $L(\theta_A, \theta_B, \theta_C, \cdot)$ on each iteration

Likelihoods that probability of success in one group is greater than the probabilities of success in two other groups are equal to

$$P\left\{\hat{\theta}_A > \max\left\{\hat{\theta}_B, \hat{\theta}_C\right\}\right\} = 0{,}03,$$

$$P\left\{\hat{\theta}_B > \max\left\{\hat{\theta}_A, \hat{\theta}_C\right\}\right\} = 0{,}97,$$

$$P\left\{\hat{\theta}_C > \max\left\{\hat{\theta}_A, \hat{\theta}_B\right\}\right\} = 9{,}65 \cdot 10^{-6}.$$

Testing has been finished with expected loss:

$$EL\left(\hat{\theta}_A, \hat{\theta}_B, \hat{\theta}_C, A\right) = 0{,}251,\ EL\left(\hat{\theta}_A, \hat{\theta}_B, \hat{\theta}_C, B\right) = 0{,}002,\ EL\left(\hat{\theta}_A, \hat{\theta}_B, \hat{\theta}_C, C\right) = 0{,}585.$$

6. Conclusion

We proposed a Bayesian A/B/C testing implementation with the expected value of the loss function computed analytically. A/B/C testing procedure has several advantages over the multivariate testing. A/B/C testing can be used when we have enough traffic for A/B/C test, but not for a multivariate test. A/B/C testing should be used when we wanted to improve element on a website that may have multiple versions: A, B, C. A/B/C must be used to answer the question ”what is probability that landing page variant A is better than landing page variants B and C?”

Bayesian A/B/C testing can easily be converted into Bayesian A/B testing with a closed-form solution.

References

1. **R. Kohavi, R. M. Henne, and D. Sommerfield.** Practical guide to controlled experiments on the web: listen to your customers not to the hippo [Text] // Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 959–967, ACM, 2007.
2. **R. Kohavi, R. Longbotham, D. Sommerfield, and R. M. Henne**. Controlled experiments on the web: survey and practical guide [Text] // Data mining and knowledge discovery, vol. 18, no. 1, pp. 140–181, 2009.
3. **R. Kohavi, A. Deng, B. Frasca, T. Walker, Y. Xu, and N. Pohlmann.** Online controlled experiments at large scale [Text] // Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 1168–1176. ACM, 2013.
4. **R. Kohavi, A. Deng, R. Longbotham, and Y. Xu.** Seven rules of thumb for web site experimenters [Text] // Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, 2014.
5. **A. Wald.** Sequential tests of statistical hypotheses [Text] // Ann. Math. Statist., vol. 16, no. 2, pp. 117–186, 1945.
6. **M. Girshick.** Contributions to the theory of sequential analysis [Text] // I, Ann. Math. Statist., vol. 17, no. 2, pp. 123–143, 1946.
7. **L. Pekelis, D. Walsh, R. Johari.** The new stats engine [Text] // Technical report, Optimizely, 2015.
8. **R. Johari, P. Koomen, L. Pekelis, and D. Walsh.** Peeking at a/b tests: Why it matters, and what to do about it [Text] // Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1517–1525, 2017.
9. **R. Johari, L. Pekelis, and D. Walsh.** Always valid inference: Bringing sequential analysis to A/B testing, arXiv preprint arXiv:1512.04922v3, Optimizely, 2019.
10. **E. Miller.** Simple sequential a/b testing, URL http://www.evanmiller.org/sequential-ab-testing.html, blog post, 2015.
11. **Ya. Bondarenko, and O. Yashchuk.** Sequential analysis in the A/B testing [Text] // Problems of Applied Mathematics and Mathematical modelling, pp. 192–203, 2016.
12. **C. Stucchio.** Bayesian A/B Testing at VWO [Text] // Technical report, Visual Website Optimizer, 2015.
13. **E. Miller.** Formulas for Bayesian A/B testing, URL http://www.evanmiller.org/bayesian-ab-testing.html, blog post, 2015.
14. **Ya. Bondarenko, and S. Kravchenko.** On the Frequentist Approach to Multivariate Landing Page Testing [Text] // Visnyk DNU. Series: Modelling, vol. 9, no. 8, pp. 142–151, 2017.
15. **Ya. Bondarenko, and S. Kravchenko.** Bayesian A/B/C testing [Text] // The Fifth Baltic-Nordic Conference on Survey Statistics, 16-20 June 2019 in Örebro, Sweden.

Received 17.09.2019.