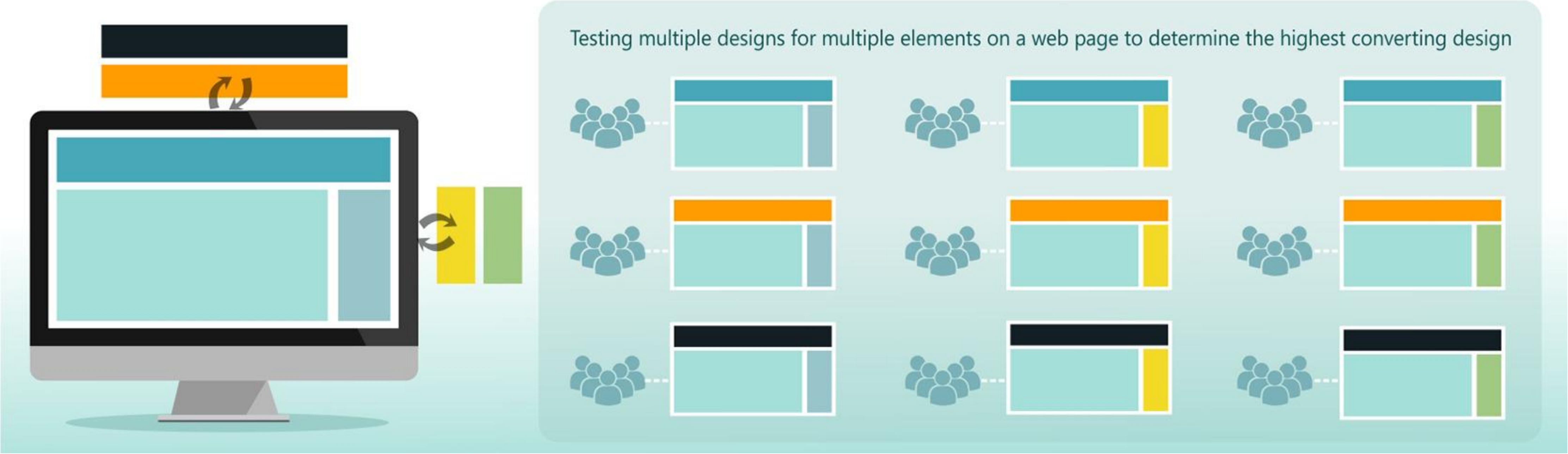
Bayesian A/B/C testing

Yana Bondarenko, Sergey Kravchenko

Department of Statistics and Probability Theory
Oles Honchar Dnipro National University
Dnipro, Ukraine

The Fifth Baltic-Nordic Conference on Survey Statistics
June 16-20, 2019, Örebro, Sweden

Multivariate testing



Multivariate testing (MVT) is a technique for testing a hypothesis in which multiple variables are modified. The goal of MVT is to determine which combination of variations performs the best out of all of the possible combinations.

1. Mathematical model of Frequentist testing
Problem formulation

Group 1

Conversion Rate p_1

Group 2

Conversion Rate p_2

…

Group m

Conversion Rate p_m

Bernoulli trials with two possible outcomes (success and failure) are conducted in m groups of visitors. Group 1 is baseline group. Probabilities of success $p_1, p_2, \ldots, p_m$ are **non-random** unknown variables.

Null hypotheses:

$$H_0^{ij}: p_i = p_j,\ \ i = 1, j = 2, \ldots, m.$$

Alternative hypotheses:

$$H_1^{ij}: p_i \neq p_j,\ \ i = 1, j = 2, \ldots, m.$$

Our goal is to accept or reject H_0^{ij} according to samples realization from Bernoulli distribution with parameters $p_1, p_2, \ldots, p_m$.

Two Proportion Z-Test

Group i

Group j

Frequencies of success:

$$\hat{p}_i = \mu_i / n_i \qquad\qquad \hat{p}_j = \mu_j / n_j$$

The weighted estimate of p_i and p_j is defined by:

$$\hat{p} = \frac{\hat{p}_i n_i + \hat{p}_j n_j}{n_i + n_j},$$

n_i, n_j - sample sizes for groups i, j

Two Proportion Z-Test

The null hypothesis $H_0^{ij}: p_i = p_j$ is rejected if

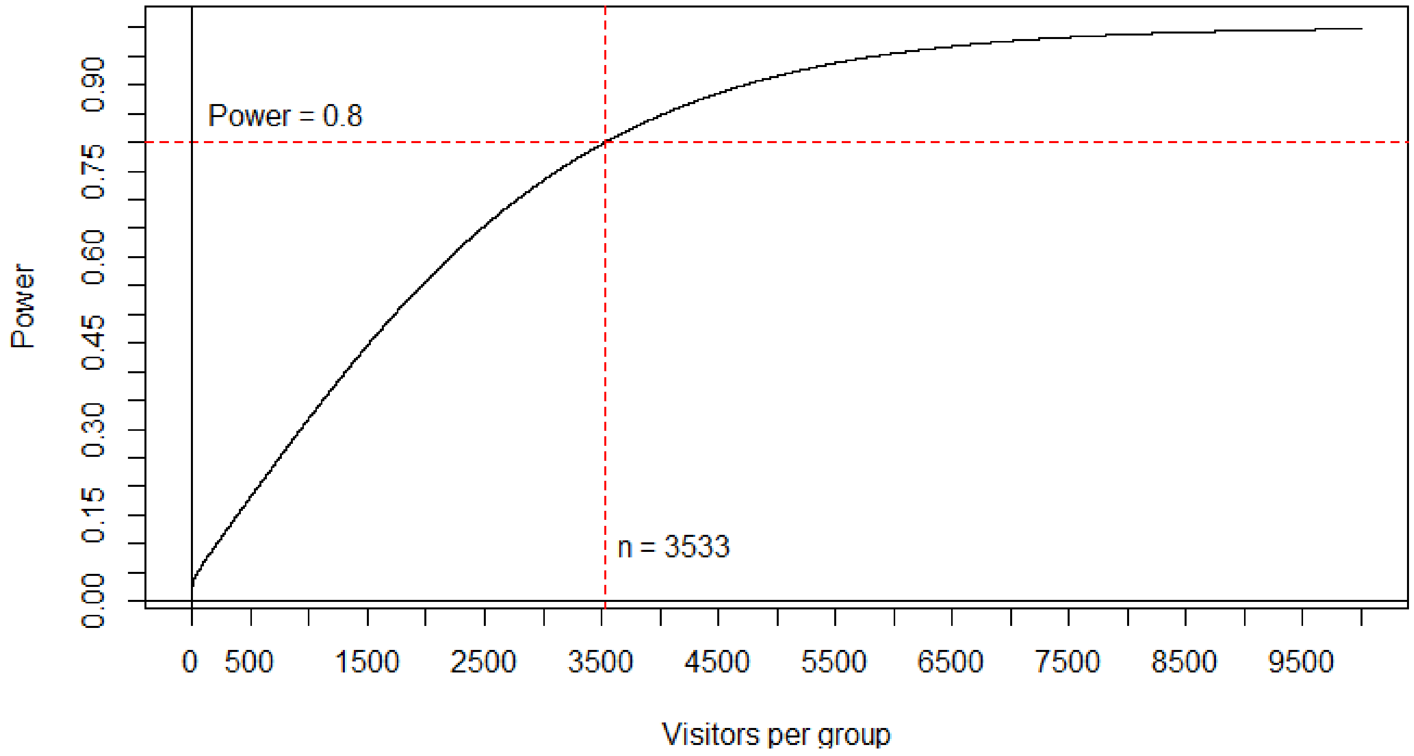
$$Z = \frac{|\hat{p}_i - \hat{p}_j|}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}} \geq z_{1-\alpha/2},$$

or, that is the same,

$$\hat{p}_i - \hat{p}_j \notin \left(-z_{1-\alpha/2}\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}, z_{1-\alpha/2}\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}\right),$$

$z_{1-\alpha/2}$ - critical value based on the standard normal distribution.

Number of unique visitors required for testing



$$Number\ of\ Unique\ Visitors = mn$$

Minimum sample size, required to prove that the probabilities of success in two groups of visitors are statistically different, is defined by:

$$n = \frac{2\left(z_{1-\alpha/2} + z_{1-\beta}\right)^2 \hat{p}(1-\hat{p})}{\theta^2}$$

where

α is probability of making a Type I error,

$1-\beta$ is Power of hypothesis testing,

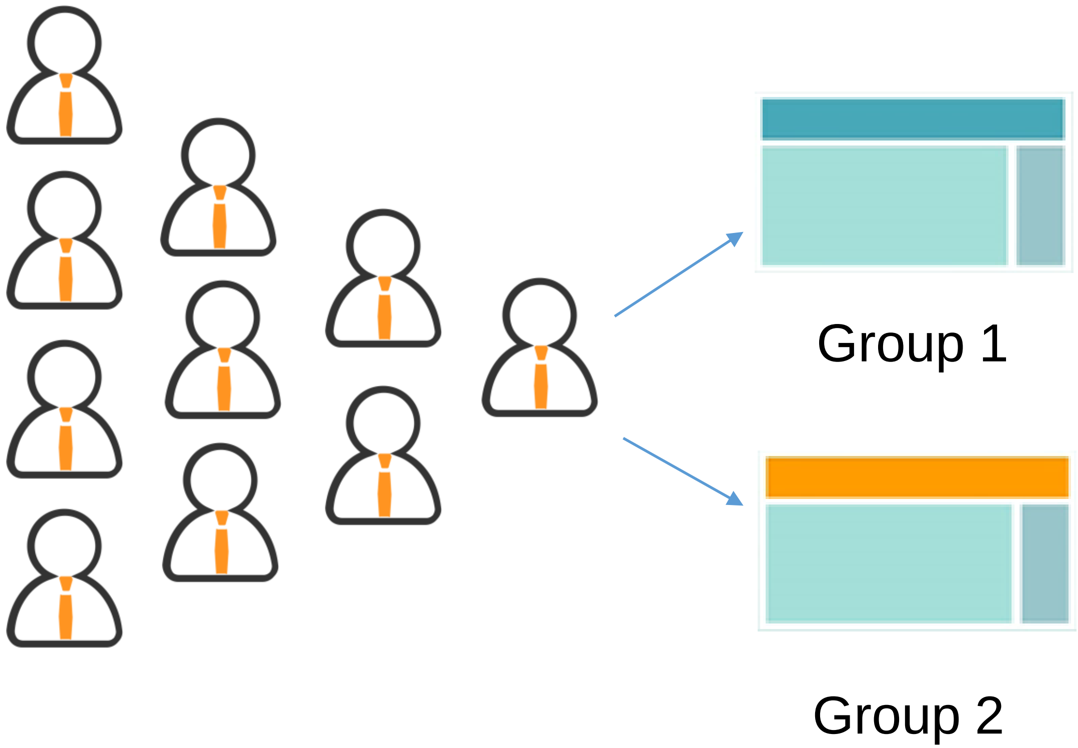
θ is expected improvement difference,

$\hat{p}$ is baseline probability of success.

Simple hypothesis: $H_0^{ij}: p_i = p_j$

simple alternative: $H_1^{ij}: p_i - p_j = \theta$

A/B testing implementation



The flow of visitors has been simulated.

Each visitor can belong to each group with probability $1/2$.

Visitor's behavior is simulated after identification of visitor belonging to group.

Visitor's behavior is determined with two outcomes: success – conversion action is done, failure – conversion action isn't done.

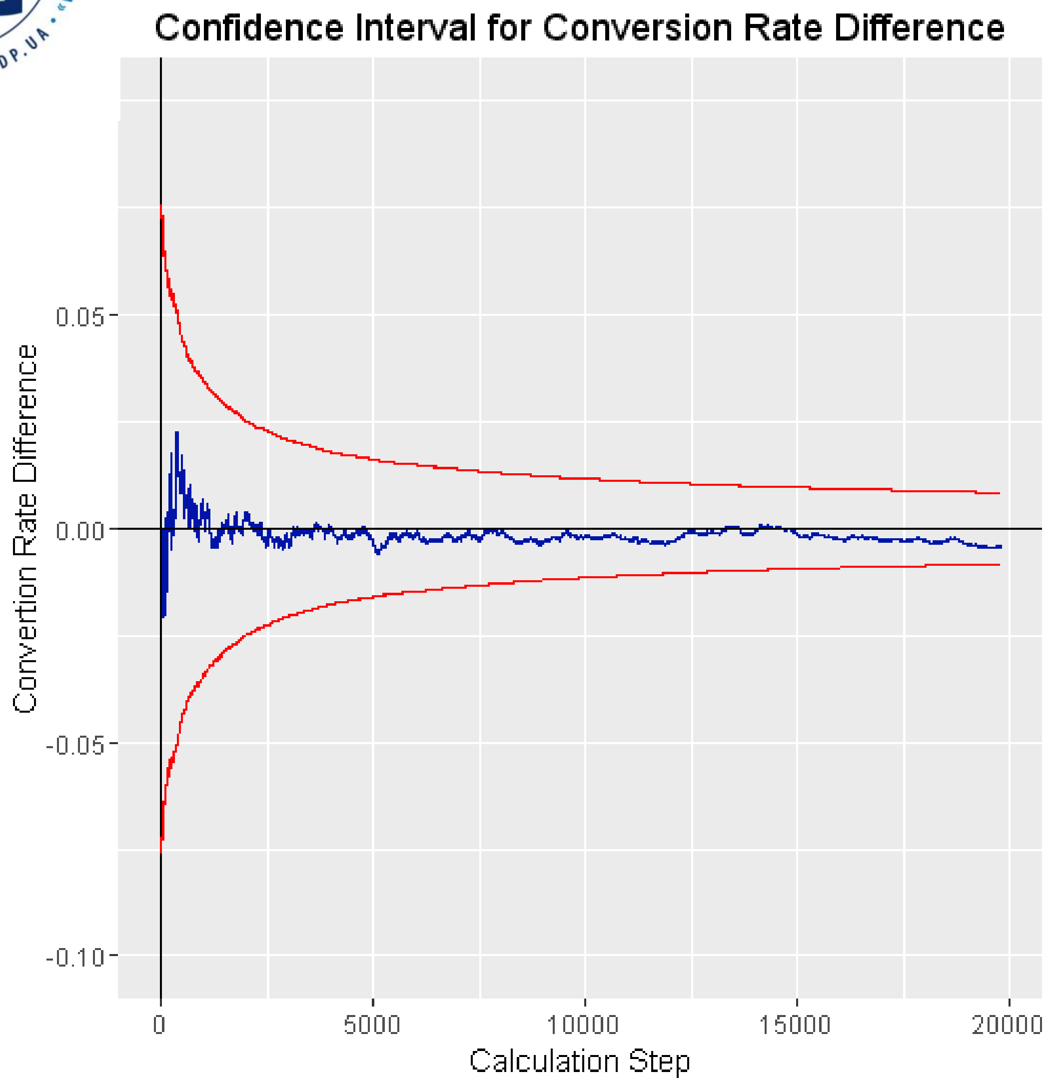
$H_0^{12}: p_1 = p_2$

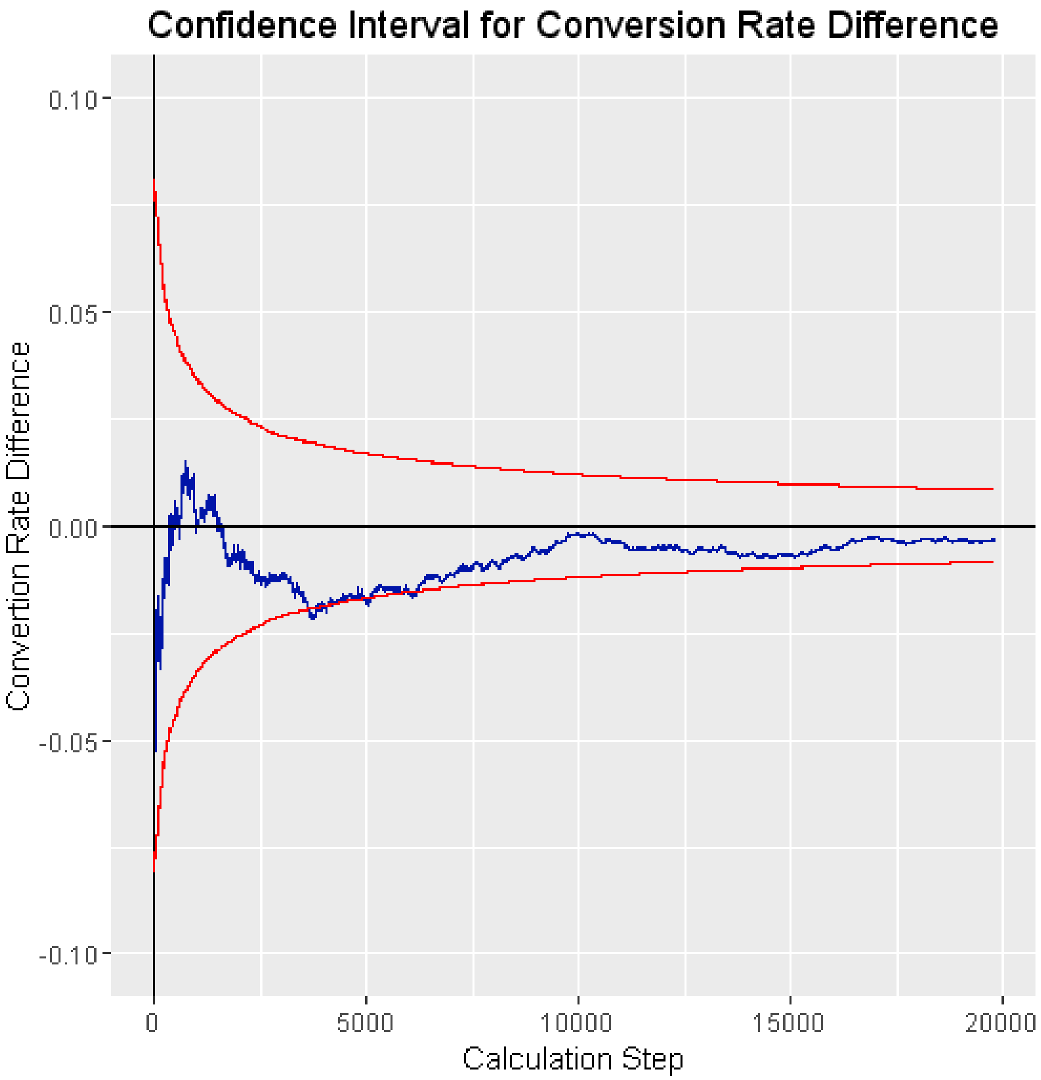
p_1 – conversion rate for group 1

p_2 – conversion rate for group 2

A/A testing

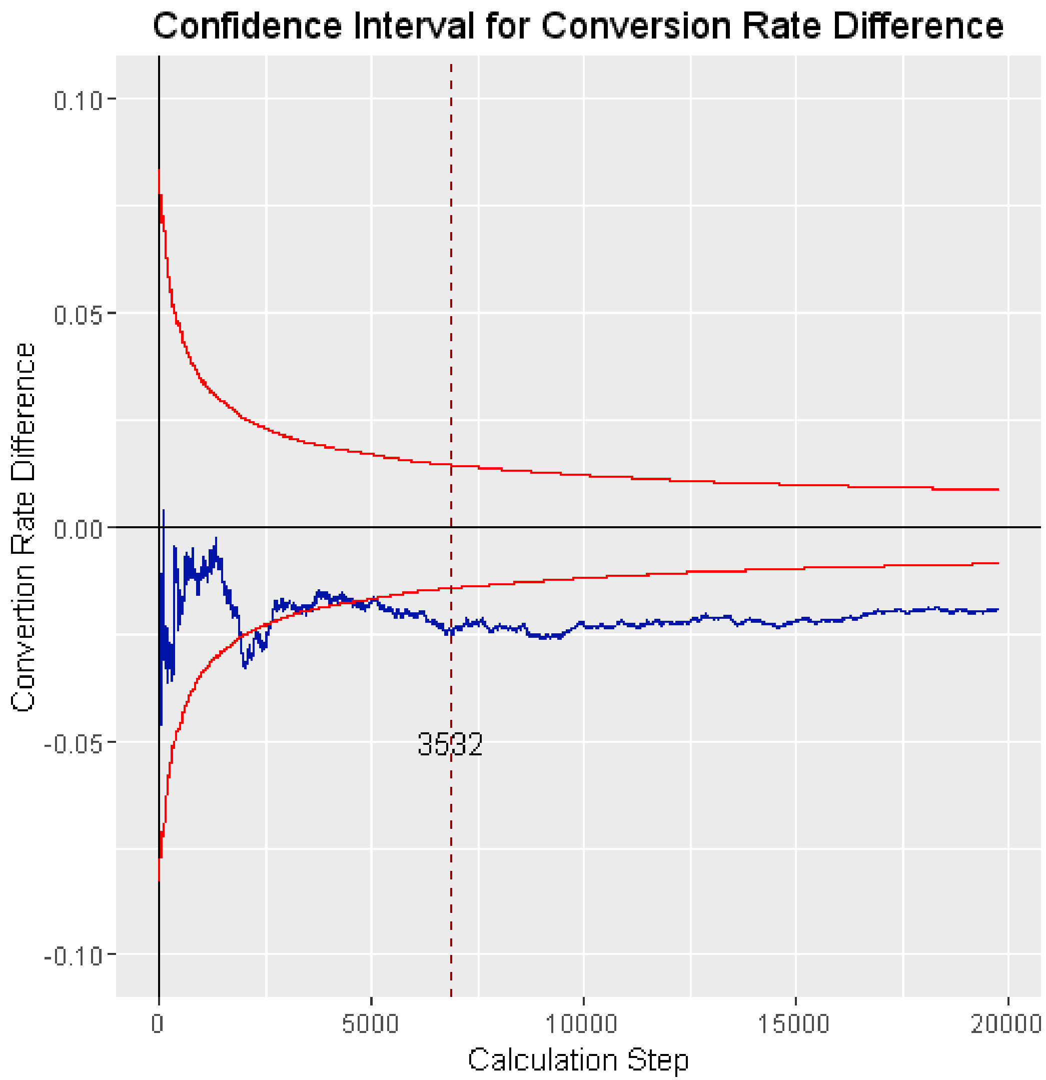
$\alpha = 0{,}05, \beta = 0{,}2, \hat{p} = 0{,}1$



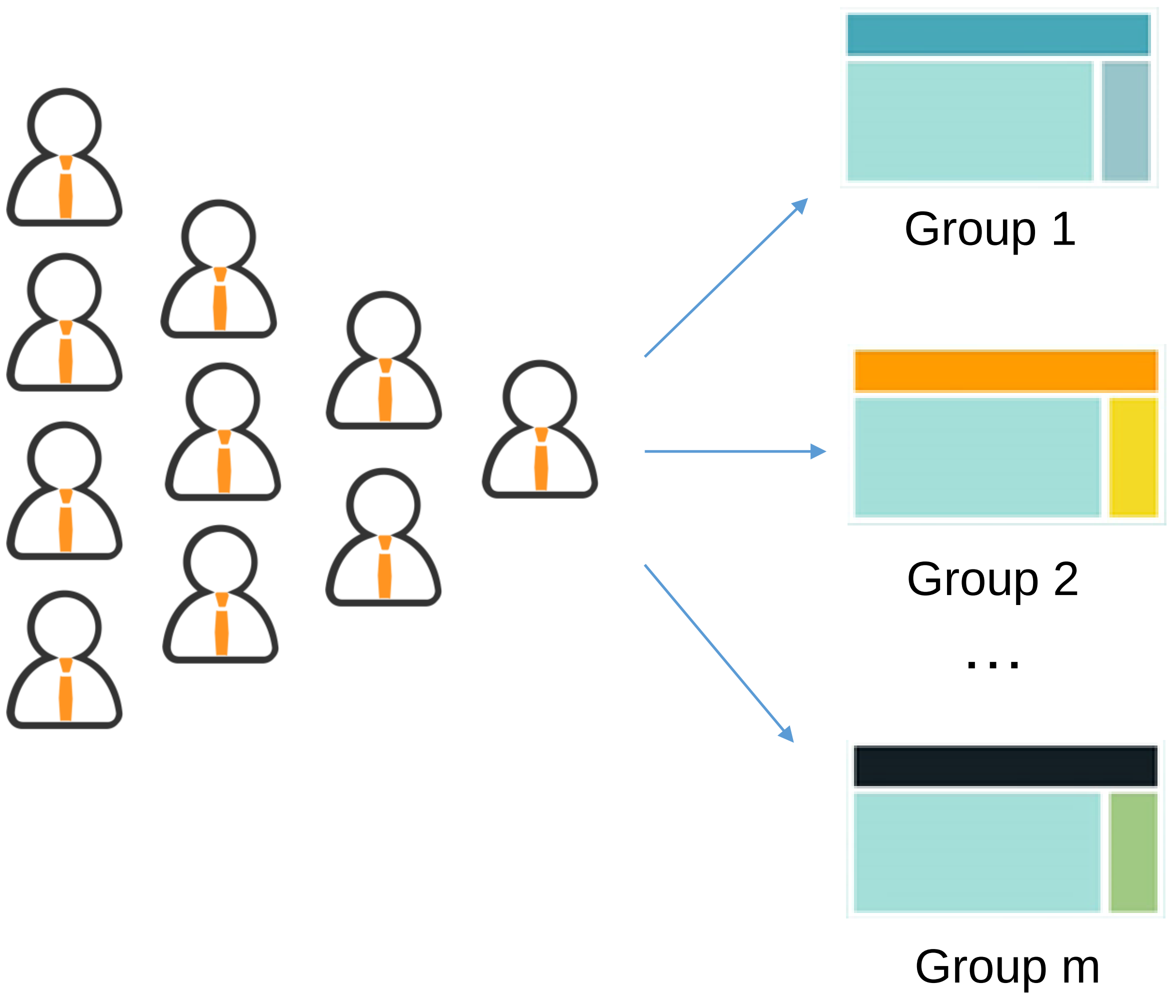


A/B testing

$$\alpha = 0{,}05, \beta = 0{,}2, \hat{p} = 0{,}1, \theta = 0{,}02$$



Multivariate testing implementation



The flow of visitors has been simulated.

Each visitor can belong to each group with probability $1/m$.

Visitor’s behavior is simulated after identification of visitor belonging to group.

Visitor’s behavior is determined with two outcomes: success – conversion action is done, failure – conversion action isn’t done.

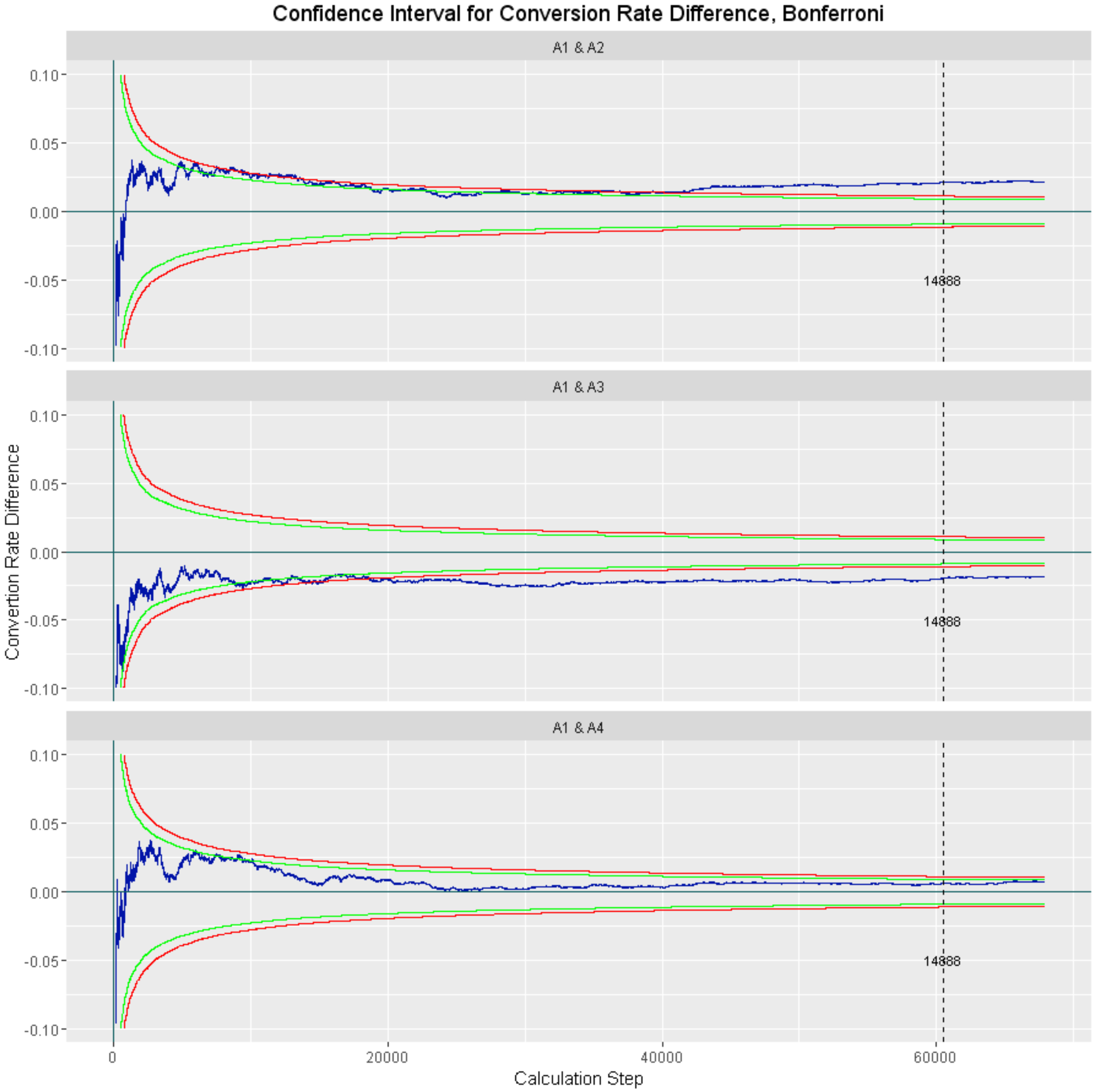
$H_0^{ij}: p_i = p_j$

p_i – conversion rate for group i

p_j – conversion rate for group j

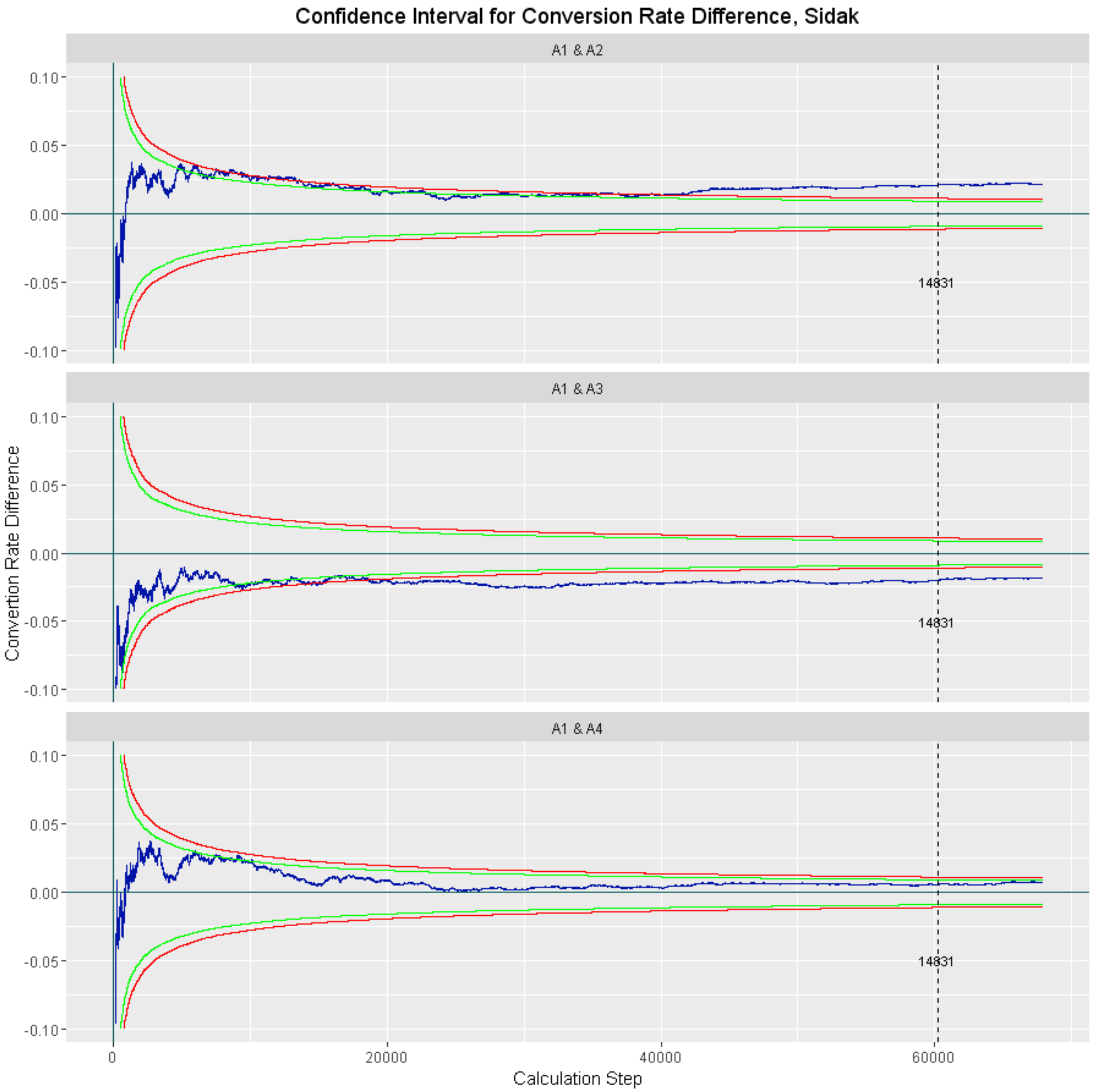
$A_1/A_2/A_3/A_4$ testing

$\hat{p} = 0.2;\ \theta = 0{,}015; \alpha = 0{,}05; \beta = 0{,}2$



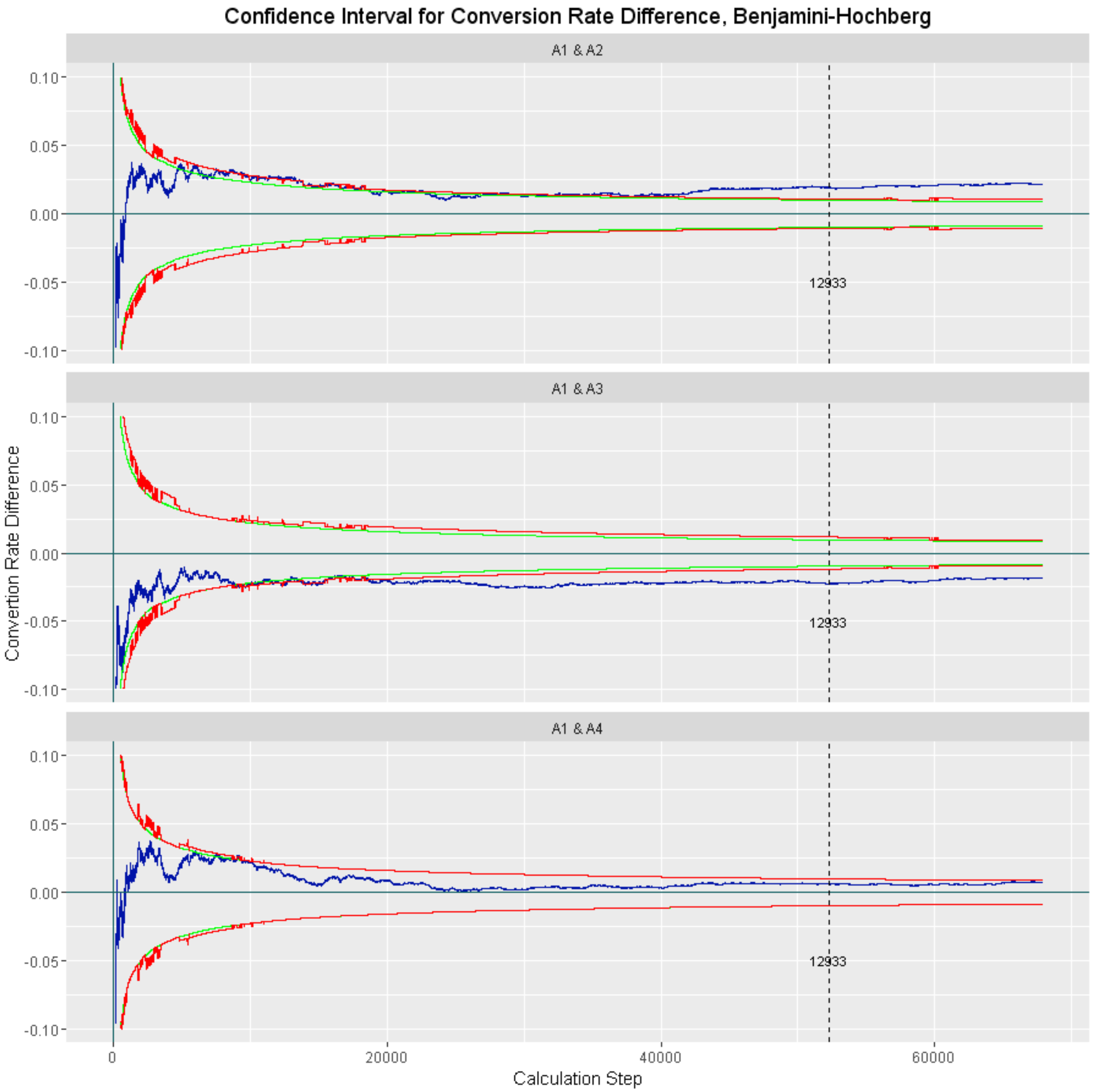
$A_1/A_2/A_3/A_4$ testing

$\hat{p} = 0.2;\ \theta = 0{,}015;\ \alpha = 0{,}05;\ \beta = 0{,}2$



$A_1/A_2/A_3/A_4$ testing

$\hat{p} = 0.2;\ \theta = 0{,}015;\ \alpha = 0{,}05;\ \beta = 0{,}2$



Hypothesis	Conv.Rate	Improvement	Multiple Comparison Adjustment Method		
			Bonferroni	Sidak	Benjamini Hochberg
			Statistical Significance		
H_0^{12}	$\hat{p}_1 = 0{,}2068$ $\hat{p}_2 = 0{,}2249$	8,75 %	0,9972	-	-
H_0^{13}	$\hat{p}_1 = 0{,}2068$ $\hat{p}_3 = 0{,}1852$	-10,44 %	0,9998	-	-
H_0^{14}	$\hat{p}_1 = 0{,}2068$ $\hat{p}_4 = 0{,}2132$	3,09 %	0,3059	-	-
H_0^{12}	$\hat{p}_1 = 0{,}2068$ $\hat{p}_2 = 0{,}2249$	8,75 %	-	0,9998	-
H_0^{13}	$\hat{p}_1 = 0{,}2068$ $\hat{p}_3 = 0{,}1852$	-10,44 %	-	0,9973	-
H_0^{14}	$\hat{p}_1 = 0{,}2068$ $\hat{p}_4 = 0{,}2132$	3,09 %	-	0,4641	-
H_0^{12}	$\hat{p}_1 = 0{,}2068$ $\hat{p}_2 = 0{,}2249$	8,75 %	-	-	0,9991
H_0^{13}	$\hat{p}_1 = 0{,}2068$ $\hat{p}_3 = 0{,}1852$	-10,44 %	-	-	0,9999
H_0^{14}	$\hat{p}_1 = 0{,}2068$ $\hat{p}_4 = 0{,}2132$	3,09 %	-	-	0,7686

2. Mathematical model of Bayesian testing
Problem formulation

Group A

Conversion Rate θ_A

Group B

Conversion Rate θ_B

Group C

Conversion Rate θ_C

Bernoulli trials with probabilities $\theta_A, \theta_B, \theta_C$ of success are conducted in three groups of visitors.

Probabilities $\theta_A, \theta_B, \theta_C$ of success are unknown **random** variables.

$p(\theta_A), p(\theta_B), p(\theta_C)$ are prior densities for $\theta_A, \theta_B, \theta_C$

$p(\theta_A|x_1, \dots, x_n), p(\theta_B|y_1, \dots, y_n), p(\theta_C|z_1, \dots, z_n)$ are posterior densities for $\theta_A, \theta_B, \theta_C$ given the sample vectors $x = (x_1, x_2, \dots, x_n)$, $y = (y_1, y_2, \dots, y_n)$, $z = (z_1, z_2, \dots, z_n)$ are observed

Our goal is to find Bayesian estimators $\hat{\theta}_A, \hat{\theta}_B, \hat{\theta}_C$ and Bayesian credible intervals for $\theta_A, \theta_B, \theta_C$.

Prior distribution for probability of success

The likelihood function for group A: $p(x_1, x_2, \dots, x_n|\theta_A) = \theta_A^{\sum_{i=1}^n x_i}(1-\theta_A)^{n-\sum_{i=1}^n x_i}$

The likelihood function for group B: $p(y_1, y_2, \dots, y_n|\theta_B) = \theta_B^{\sum_{i=1}^n y_i}(1-\theta_B)^{n-\sum_{i=1}^n y_i}$

The likelihood function for group C: $p(z_1, z_2, \dots, z_n|\theta_C) = \theta_C^{\sum_{i=1}^n z_i}(1-\theta_C)^{n-\sum_{i=1}^n z_i}$

The prior information about probability θ_A of success: $p(\theta_A) = \frac{\theta_A^{a-1}(1-\theta_A)^{b-1}}{B(a,b)}, a > 0, b > 0$

The prior information about probability θ_B of success: $p(\theta_B) = \frac{\theta_B^{c-1}(1-\theta_B)^{d-1}}{B(c,d)}, c > 0, d > 0$

The prior information about probability θ_C of success: $p(\theta_C) = \frac{\theta_C^{e-1}(1-\theta_C)^{f-1}}{B(e,f)}, e > 0, f > 0$

Posterior distribution for probability of success

Posterior distribution for probability θ_A of success:

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

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

x_i is number of successes in one trial ($x_i = 0$ or 1).

Posterior distribution for probability θ_B of success:

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

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

y_i is number of successes in one trial ($y_i = 0$ or 1).

Posterior distribution for probability θ_C of success:

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

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

z_i is number of successes in one trial ($z_i = 0$ or 1).

Loss function $\boldsymbol{L(\theta_A, \theta_B, \theta_C, \cdot)}$

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

Loss function for group A (landing page variant A is suggested viewing for the first group of visitors)

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

Loss function for group B (landing page variant B is suggested viewing for the second group of visitors)

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

Loss function for group C (landing page variant C is suggested viewing for the third group of visitors)

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

Expected loss $L(\theta_A, \theta_B, \cdot)$ is proposed to compute with numerical approach for A/B testing [Chris Stucchio, Bayesian A/B Testing at VWO, Whitepaper, Visual Website Optimizer, 2015].

Expected loss $\boldsymbol{EL(\theta_A, \theta_B, \theta_C, \cdot)}$

Expected loss for group A : $EL(\theta_A, \theta_B, \theta_C, A) = \int_0^1 \int_0^1 \int_0^1 \max\{y - x, z - x, 0\} p(x, y, z) dx dy dz$

Expected loss for group B: $EL(\theta_A, \theta_B, \theta_C, B) = \int_0^1 \int_0^1 \int_0^1 \max\{x - y, z - y, 0\} p(x, y, z) dx dy dz$

Expected loss for group C: $EL(\theta_A, \theta_B, \theta_C, C) = \int_0^1 \int_0^1 \int_0^1 \max\{x - z, y - z, 0\} p(x, y, z) dx dy dz$

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

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

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

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

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

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

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

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

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

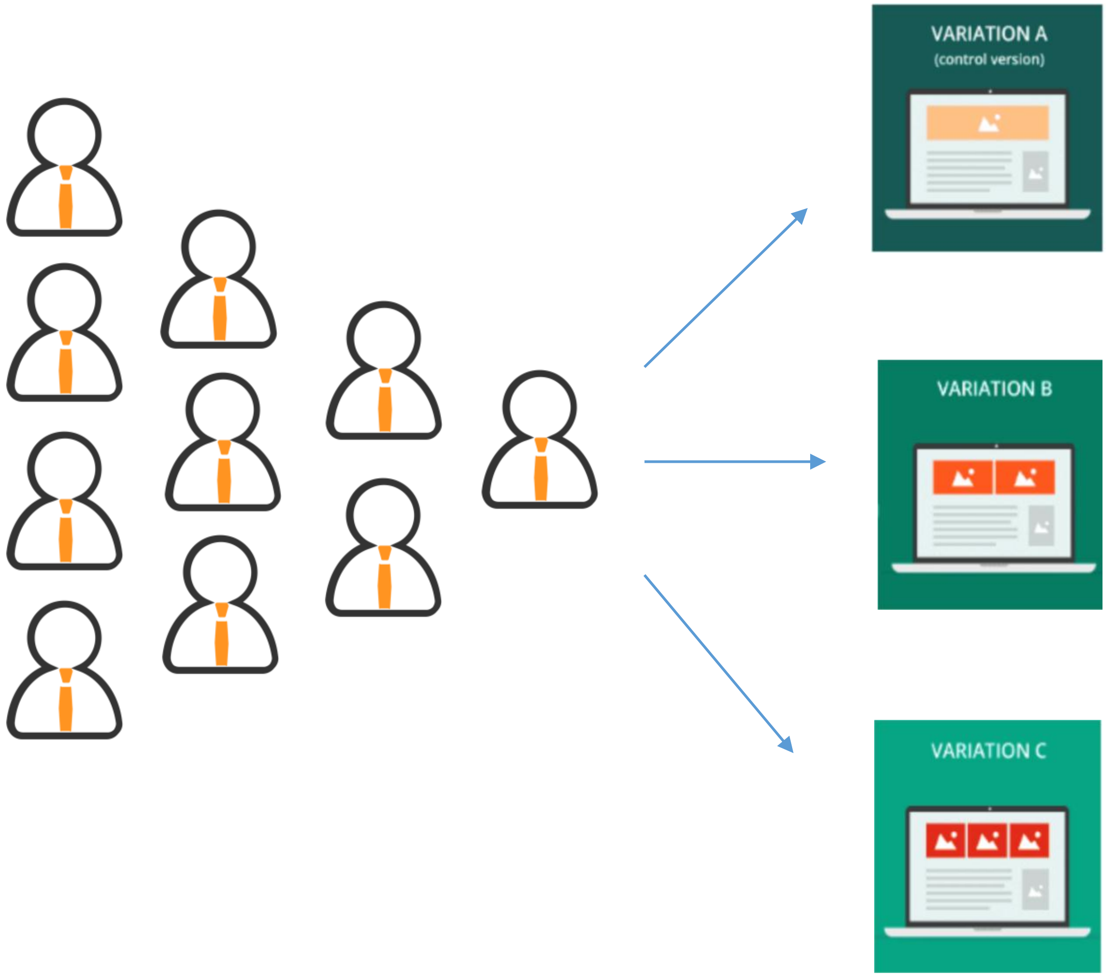
$$EL(\boldsymbol{\theta_A},\boldsymbol{\theta_B},\boldsymbol{\theta_C},\mathrm{C}) = \frac{B(a+1,b)}{B(a,b)}h(a+1,b,c,d) - \frac{B(a+1,b)}{B(a,b)}\left(1-h(a+1,b,e,f)\right) + \frac{B(a+1,b)}{B(a,b)}\sum_{i=0}^{e-1}\frac{B(a+i+1,b+f)}{(i+f)B(i+1,f)B(a+1,b)}(1-h(a+i+1,b+f,c,d)) -$$

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

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

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

Bayesian testing implementation



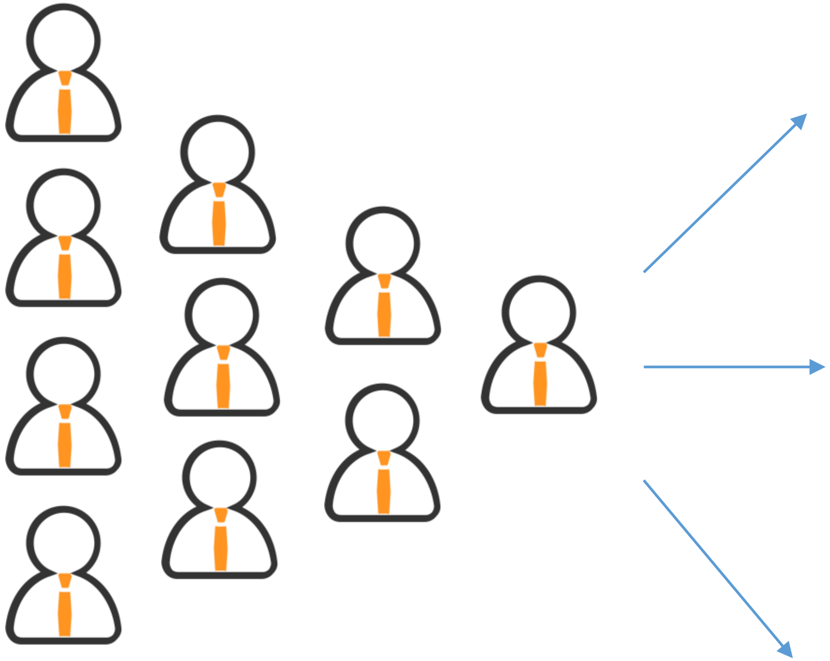
The flow of visitors has been simulated.

Each visitor can belong
to the first group with probability $1/3$,
to the second group with probability $1/3$,
to the third group with probability $1/3$.

Visitor’s behavior is simulated after identification of visitor belonging to group.

Visitor’s behavior is determined with two outcomes:
success – conversion action is done,
failure – conversion action isn’t done.

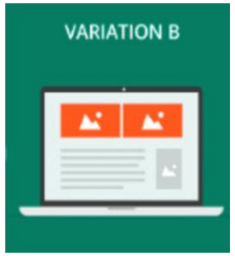
Bayesian testing implementation



$\theta_A \sim Beta(a, b)$

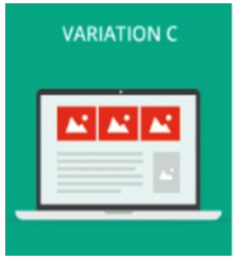
Prior information about probability θ_A of success is defined with Beta-distribution with (a, b) parameters.

$\theta_B \sim Beta(c, d)$



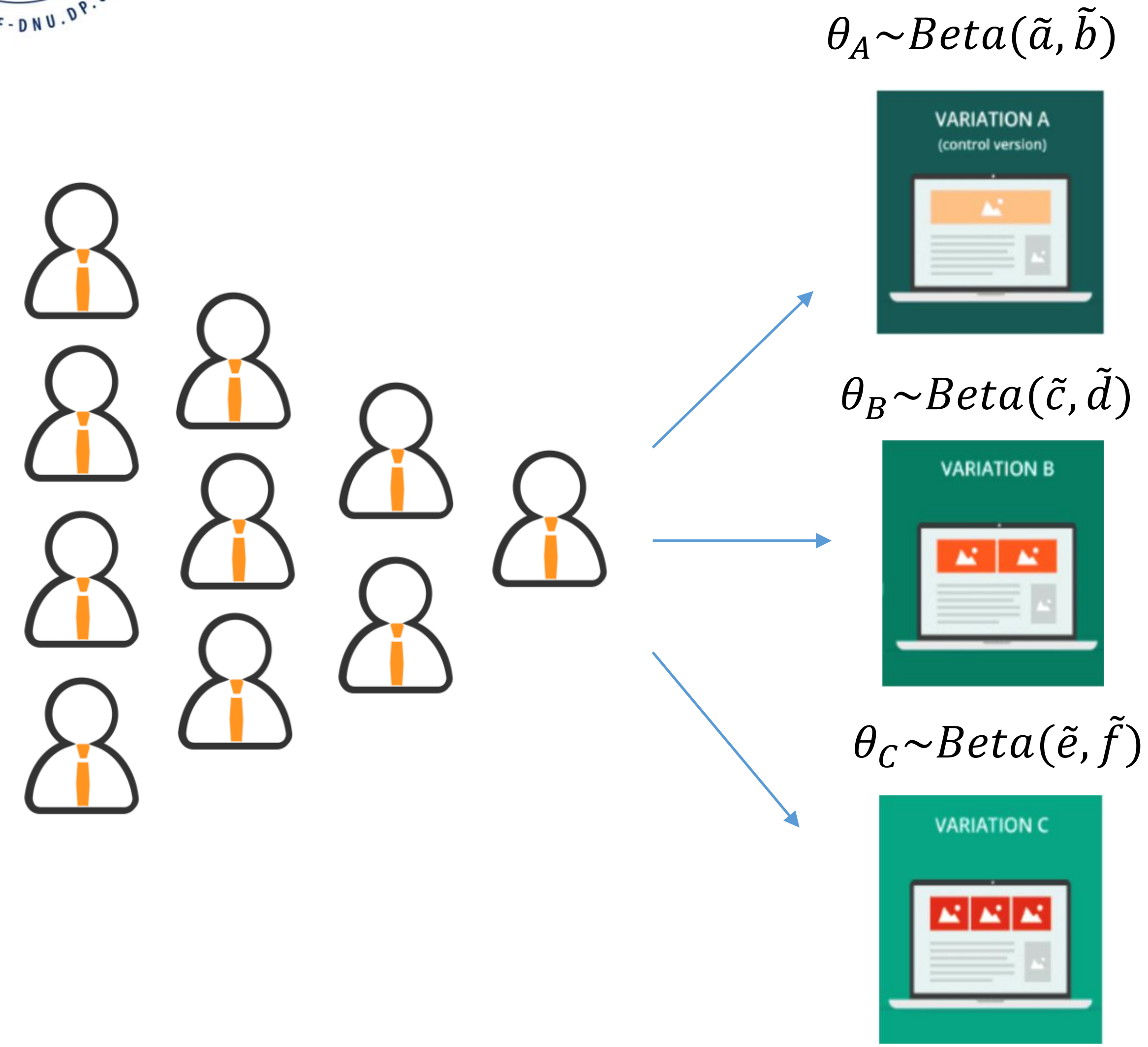
Prior information about probability θ_B of success is defined with Beta-distribution with (c, d) parameters.

$\theta_C \sim Beta(e, f)$



Prior information about probability θ_C of success is defined with Beta-distribution with (e, f) parameters.

Bayesian testing implementation



Posterior distributions for $\theta_A, \theta_B, \theta_C$ are

Beta-distribution with $(\tilde{a}, \tilde{b})$ parameters:

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

where x_i is number of successes in one trial $(x_i = 0 \; or \; 1)$,

Beta-distribution with $(\tilde{c}, \tilde{d})$ parameters:

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

where y_i is number of successes in one trial $(y_i = 0 \; or \; 1)$,

Beta-distribution with $(\tilde{e}, \tilde{f})$ parameters:

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

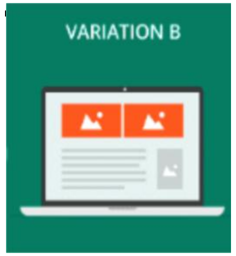
where z_i is number of successes in one trial $(z_i = 0 \; or \; 1)$.

Bayesian testing implementation

Expected loss is computed and compared with threshold of loss ε after each visit of landing page variant A or B or C.

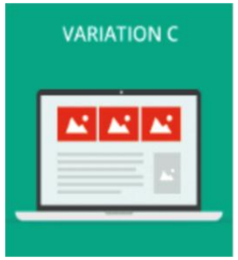
If $EL(\hat{\theta}_A, \hat{\theta}_B, \hat{\theta}_C, A) \leq \varepsilon$, $EL(\hat{\theta}_A, \hat{\theta}_B, \hat{\theta}_C, B) > \varepsilon$, $EL(\hat{\theta}_A, \hat{\theta}_B, \hat{\theta}_C, C) > \varepsilon$

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



If $EL(\hat{\theta}_A, \hat{\theta}_B, \hat{\theta}_C, A) > \varepsilon$, $EL(\hat{\theta}_A, \hat{\theta}_B, \hat{\theta}_C, B) \leq \varepsilon$, $EL(\hat{\theta}_A, \hat{\theta}_B, \hat{\theta}_C, C) > \varepsilon$

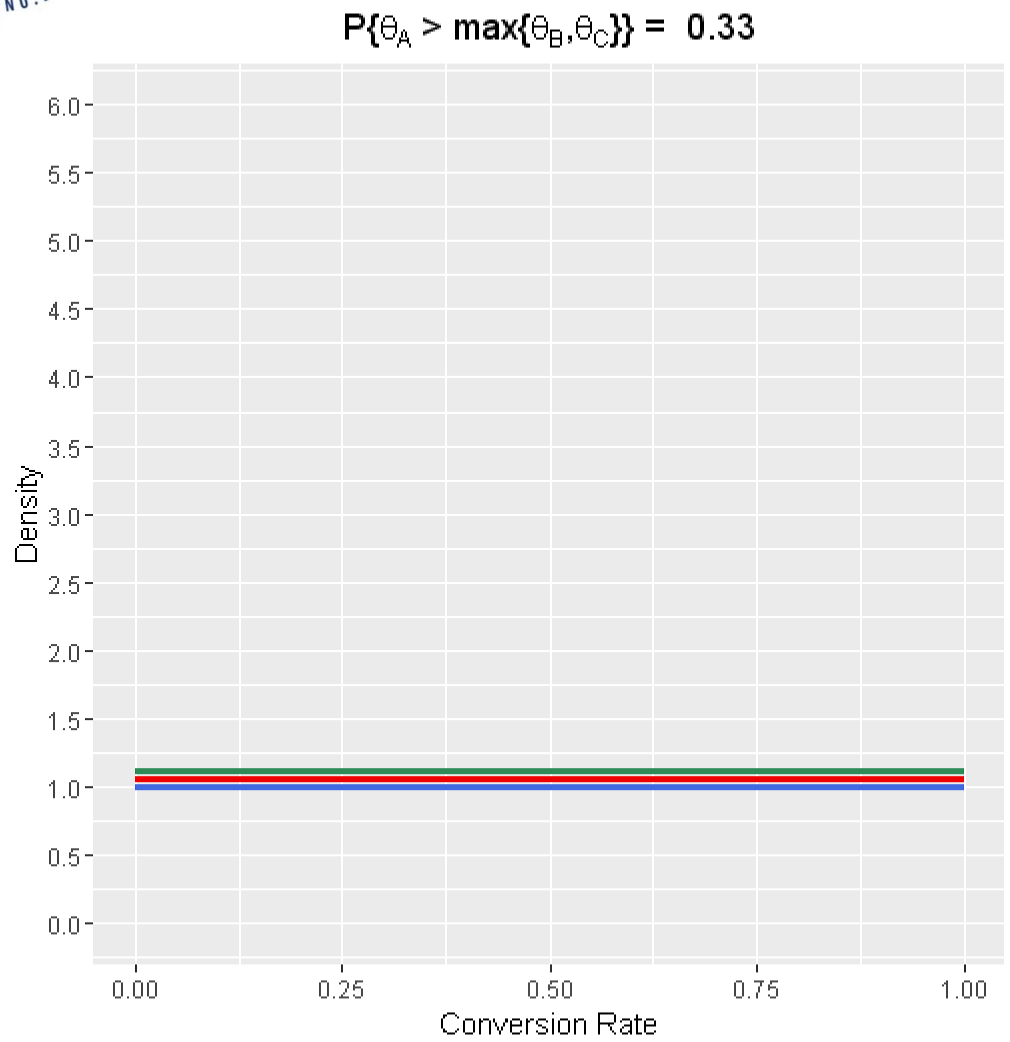
testing will be stopped, landing page variant B will be chosen for publishing on website.



If $EL(\hat{\theta}_A, \hat{\theta}_B, \hat{\theta}_C, A) > \varepsilon$, $EL(\hat{\theta}_A, \hat{\theta}_B, \hat{\theta}_C, B) > \varepsilon$, $EL(\hat{\theta}_A, \hat{\theta}_B, \hat{\theta}_C, C) \leq \varepsilon$

testing will be stopped, landing page variant C will be chosen for publishing on website.

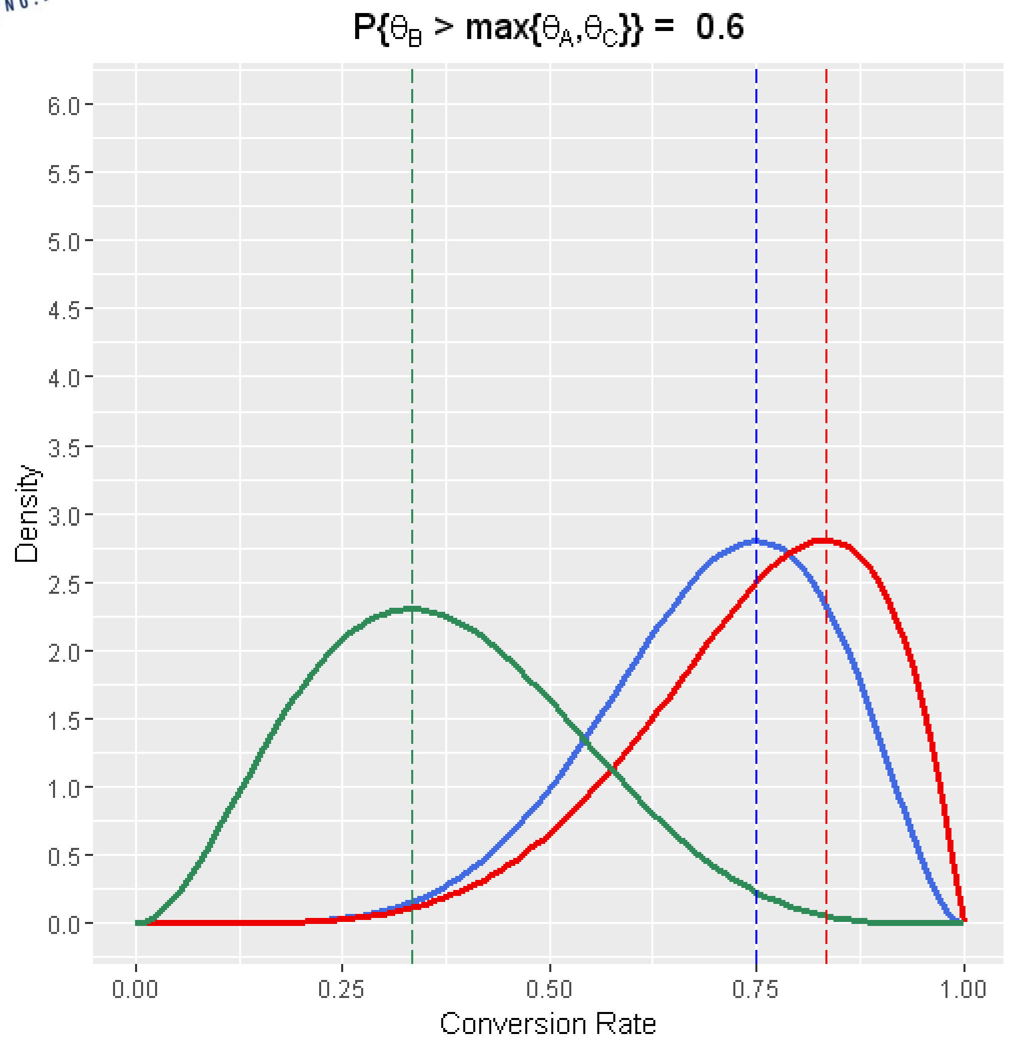
Results (1st iteration)



Prior density for Conversion Rate θ_A (Beta-distribution with $(1,1)$ parameters) is represented with blue color,
prior density for Conversion Rate θ_B (Beta-distribution with $(1,1)$ parameters) is represented with red color,
prior density for Conversion Rate θ_C (Beta-distribution with $(1,1)$ parameters) is represented with green color.

$$P\{\theta_A > \max\{\theta_B, \theta_C\}\} = 0{,}33,$$
$$P\{\theta_B > \max\{\theta_A, \theta_C\}\} = 0{,}33,$$
$$P\{\theta_C > \max\{\theta_A, \theta_B\}\} = 0{,}33.$$

Results (21st iteration)



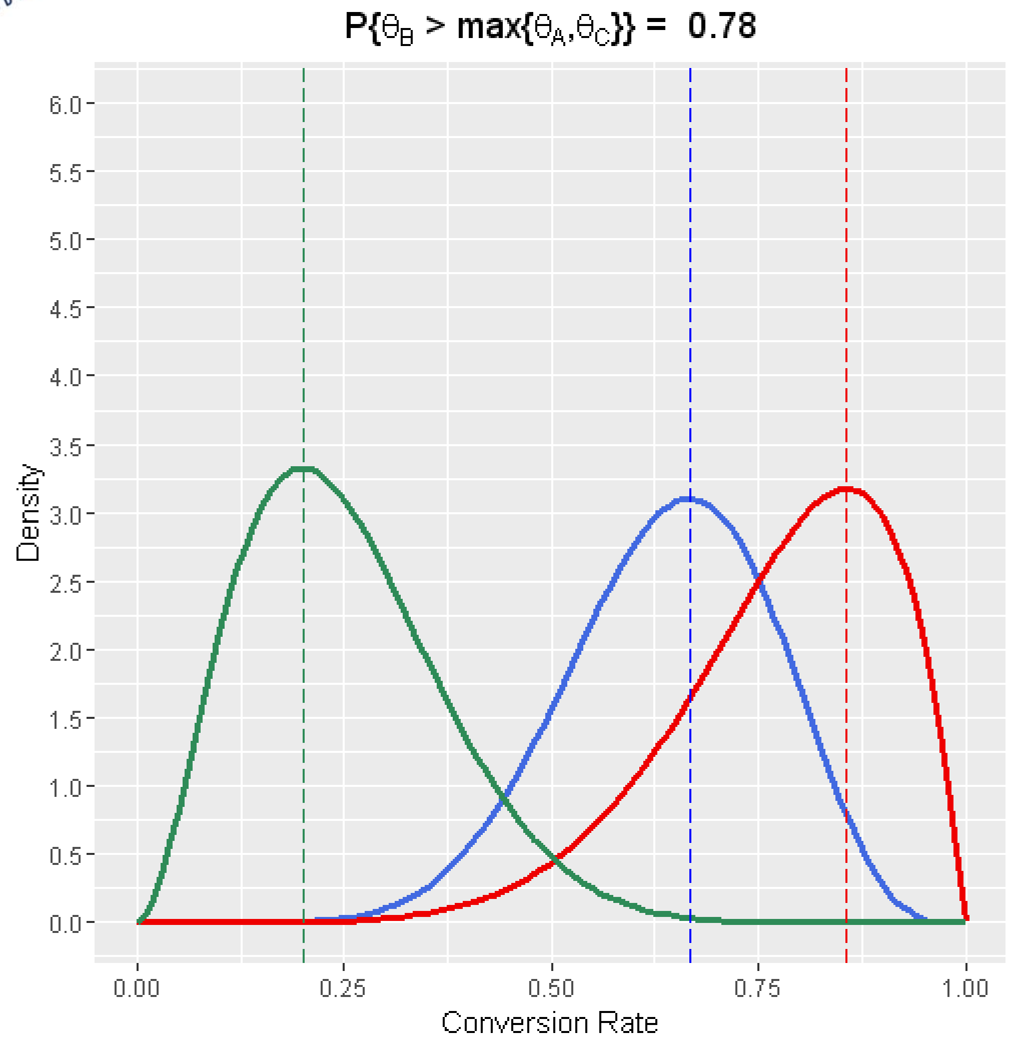
Posterior density for Conversion Rate θ_A (Beta-distribution with $(7{,}3)$ parameters) is represented with blue color,
posterior density for Conversion Rate θ_B (Beta-distribution with $(6{,}2)$ parameters) is represented with red color,
posterior density for Conversion Rate θ_C (Beta-distribution with $(3{,}5)$ parameters) is represented with green color.

Bayesian estimators:

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

$$P\{\theta_A > \max\{\theta_B, \theta_C\}\} = 0{,}39,$$
$$P\{\theta_B > \max\{\theta_A, \theta_C\}\} = 0{,}60,$$
$$P\{\theta_C > \max\{\theta_A, \theta_B\}\} = 0{,}01.$$

Results (30th iteration)



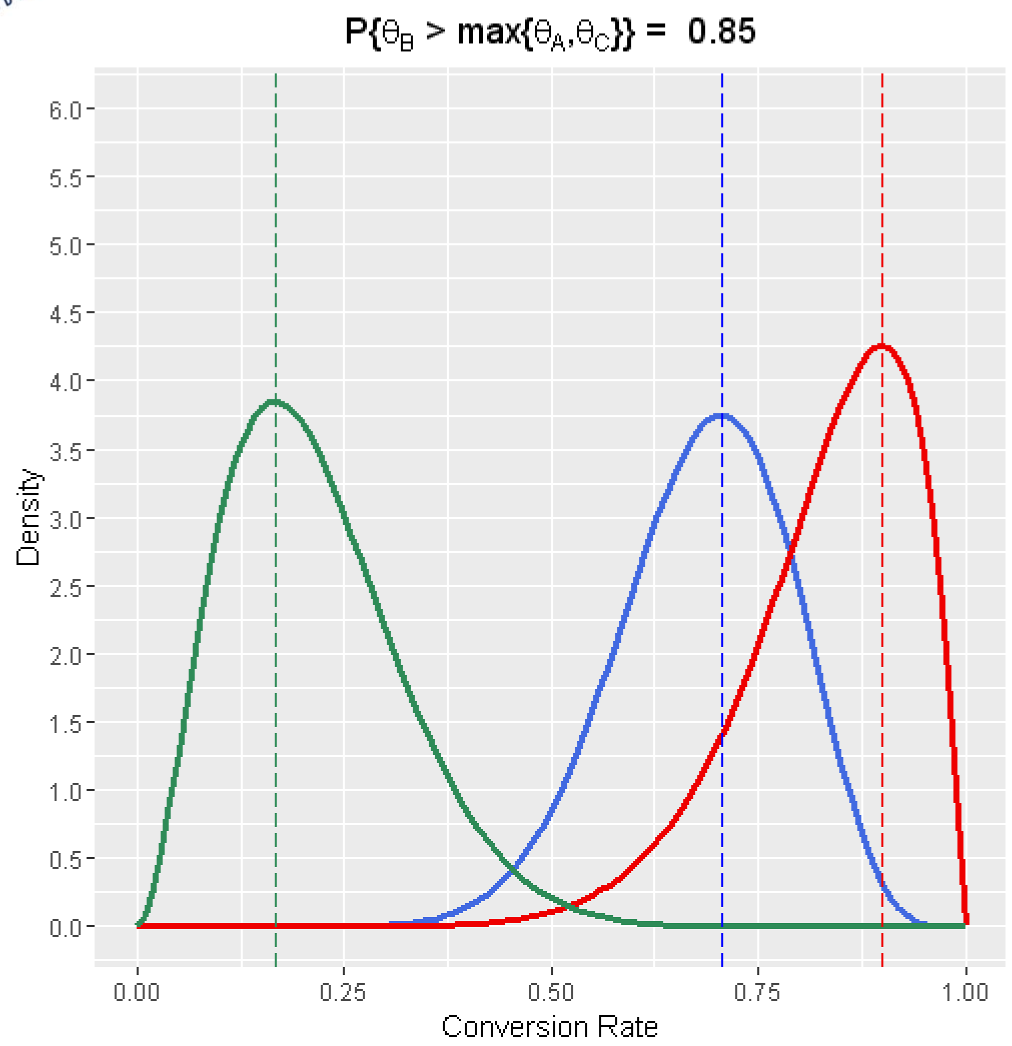
Posterior density for Conversion Rate θ_A (Beta-distribution with $(9,5)$ parameters) is represented with blue color,
posterior density for Conversion Rate θ_B (Beta-distribution with $(7,2)$ parameters) is represented with red color,
posterior density for Conversion Rate θ_C (Beta-distribution with $(3,9)$ parameters) is represented with green color.

Bayesian estimators:

$$\hat{\theta}_A = 0{,}67, \hat{\theta}_B = 0{,}85, \hat{\theta}_C = 0{,}20$$

$$P\{\theta_A > \max\{\theta_B, \theta_C\}\} = 0{,}220,$$
$$P\{\theta_B > \max\{\theta_A, \theta_C\}\} = 0{,}778,$$
$$P\{\theta_C > \max\{\theta_A, \theta_B\}\} = 0{,}002.$$

Results (39th iteration)



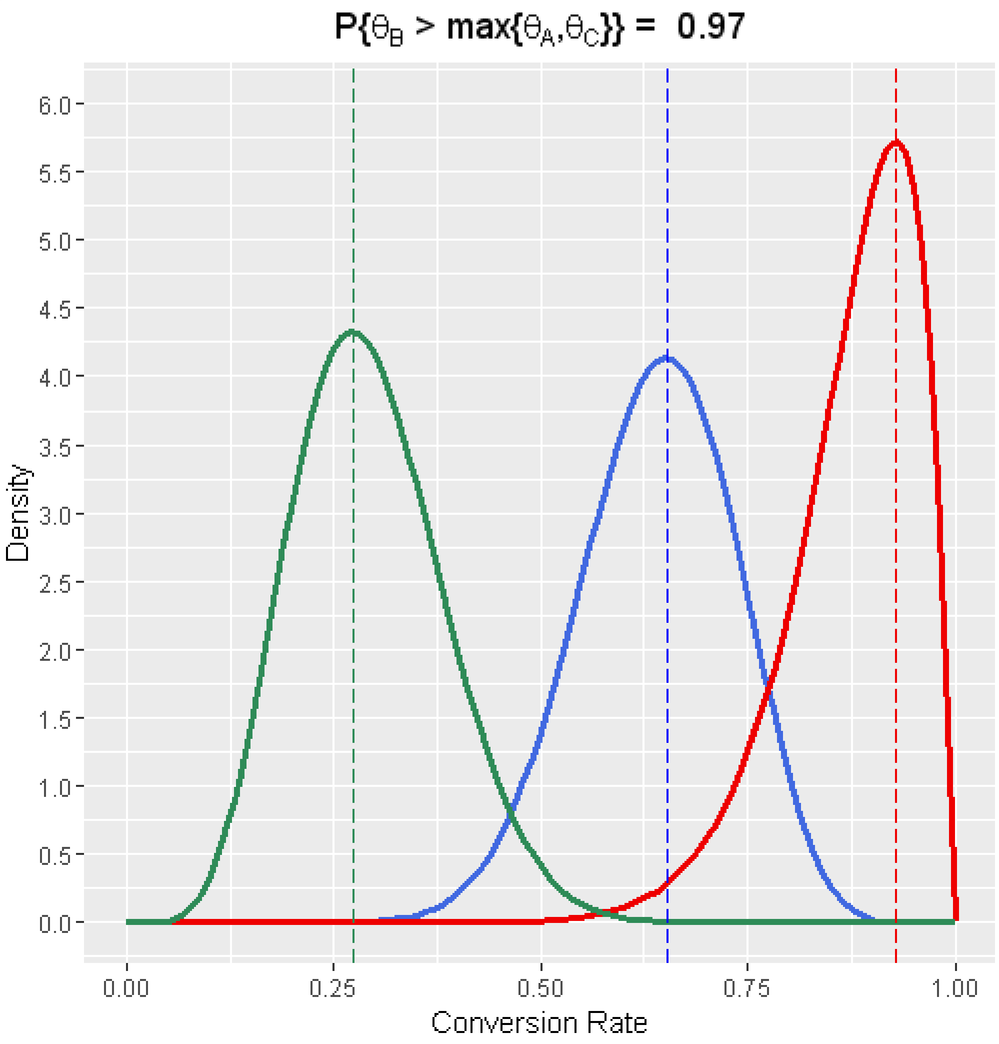
Posterior density for Conversion Rate θ_A (Beta-distribution with $(13,6)$ parameters) is represented with blue color,
posterior density for Conversion Rate θ_B (Beta-distribution with $(10,2)$ parameters) is represented with red color,
posterior density for Conversion Rate θ_C (Beta-distribution with $(3,10)$ parameters) is represented with green color.

Bayesian estimators:

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

$$P\{\theta_A > \max\{\theta_B, \theta_C\}\} = 0{,}151,$$
$$P\{\theta_B > \max\{\theta_A, \theta_C\}\} = 0{,}848,$$
$$P\{\theta_C > \max\{\theta_A, \theta_B\}\} = 0{,}001.$$

Results (60th iteration)



Posterior density for Conversion Rate θ_A (Beta-distribution with $(16,9)$ parameters) is represented with blue color,
posterior density for Conversion Rate θ_B (Beta-distribution with $(14,2)$ parameters) is represented with red color,
posterior density for Conversion Rate θ_C (Beta-distribution with $(7,17)$ parameters) is represented with green color.

Bayesian estimators:

$$\hat{\theta}_A = 0{,}65, \hat{\theta}_B = 0{,}93, \hat{\theta}_C = 0{,}27$$

$$P\{\theta_A > \max\{\theta_B, \theta_C\}\} = 0{,}03,$$
$$P\{\theta_B > \max\{\theta_A, \theta_C\}\} = 0{,}97,$$
$$P\{\theta_C > \max\{\theta_A, \theta_B\}\} = 9{,}65 \cdot 10^{-6}.$$

Results

23 visitors of group A, 14 visitors of group B and 22 visitors of group C have taken part in Bayesin testing. Test has been finished with expected loss:

$$EL\left(\hat{\theta}_A, \hat{\theta}_B, \hat{\theta}_C, A\right) = 0{,}251 > \varepsilon = 0{,}002,$$

$$EL\left(\hat{\theta}_A, \hat{\theta}_B, \hat{\theta}_C, B\right) = 0{,}002 \le \varepsilon = 0{,}002,$$

$$EL\left(\hat{\theta}_A, \hat{\theta}_B, \hat{\theta}_C, C\right) = 0{,}585 > \varepsilon = 0{,}002,$$

landing page variant B is chosen for publishing on website.

Probability that Conversion Rate in one group is greater than Conversion Rate in other groups is

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

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

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

Bayesian estimator for Conversion Rate θ_B is $\hat{\theta}_B = 0{,}93$.

Bayesian credible interval for Conversion Rate θ_B is $P\{0{,}68 \le \theta_B \le 0{,}98\} = 0{,}95$.

References

1. L. Pekelis, D. Walsh, R. Johari, The New Stats Engine, Whitepaper, Optimizely, 2015.
2. R. Johari, L. Pekelis, D. Walsh, Always valid inference: Bringing sequential analysis to A/B testing. *arXiv preprint arXiv:1512.04922,* Optimizely, 2015.
3. C. Stucchio, Bayesian A/B Testing at VWO, Whitepaper, Visual Website Optimizer, 2015.
4. G. Georgiev, Issues with Current Bayesian Approaches to A/B Testing in Conversion Rate Optimization, Whitepaper, Analytics-Toolkit, 2017.
5. G. Georgiev, Efficient A/B Testing in Conversion Rate Optimization: The AGILE Statistical Method, Whitepaper, Analytics-Toolkit, 2017.
6. G. Georgiev, Non-Inferiority Designs in the A/B Testing, Whitepaper, Analytics-Toolkit, 2017.
7. J. Hickey, Beyond A/B testing: Automation within testing in Adobe Target, Whitepaper, Adobe Target, 2017.
8. Ya. Bondarenko, O. Yashchuk, Sequential analysis in the A/B testing // Problems of Applied Mathematics and Mathematical modelling. - Dnipro: DNU. - 2016. – pp. 192-203.
9. Ya. Bondarenko, S. Kravchenko, On the Frequentist Approach to Multivariate Landing Page Testing // Visnyk DNU. Series: Modelling. - Dnipro: DNU. - 2017. - Issue 9. - No. 8. - pp. 142-151.

Thank you for attention!