Formulae for Bayesian A/B Testing: Walkthrough

Instructor: Jaeho Chang

cjhsfl@gmail.com

Client: 오댕

June 12, 2020

Original material

Table of Contents

A/B testing: binary outcomes

For a binary-outcome test (e.g. a test of conversion rates), the probability that B will beat A in the long run is given by:

$$\Pr(p_B > p_A) = \sum_{i=0}^{\alpha_B - 1} \frac{B(\alpha_A + i, \beta_B + \beta_A)}{(\beta_B + i) B(1 + i, \beta_B) B(\alpha_A, \beta_A)}$$

Where:

- α_A = 1 + S_A
- β_A = 1 + F_A
- α_B = 1 + S_B
- β_B = 1 + F_B
- B is the beta function. For arbitrary α >0 and β > 0, beta function is defines as:

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}.$$

Derivation

Suppose we have two independent experimental branches (A and B) and have a Bayesian belief for each one:

$$p_A \sim \mathrm{Beta}(\alpha_A, \beta_A)$$
$$p_B \sim \mathrm{Beta}(\alpha_B, \beta_B)$$

Using the pdf of the beta distribution, we can get the total probability that p_B is greater than p_A by integrating the joint distribution over all values for which $p_B > p_A$:

$$\begin{aligned} \Pr(p_B > p_A) &= \int_0^1 \Pr(p_B > p_A | p_A) \Pr(p_A) dp_A = \int_0^1 \int_{p_A}^1 \Pr(p_B | p_A) dp_B \Pr(p_A) dp_A \\ &= \int_0^1 \int_{p_A}^1 \frac{p_A^{\alpha_A - 1}(1 - p_A)^{\beta_A - 1}}{B(\alpha_A, \beta_A)} \frac{{p_B}^{\alpha_B - 1}(1 - p_B)^{\beta_B - 1}}{B(\alpha_B, \beta_B)} dp_B dp_A \end{aligned} \tag{1}$$

Define $\mathrm{B}(x;\, a, b) = \displaystyle\int_0^x t^{a-1}\,(1-t)^{b-1}\,dt$ and $I_x(a, b) = \frac{\mathrm{B}(x;\,a,b)}{\mathrm{B}(a,b)}$. $I_x(a, b)$ is the regularized incomplete beta function and note that $I_x(1, b) = \frac{\mathrm{B}(x;\,1,b)}{\mathrm{B}(1,b)} = b \cdot \int_0^x (1-t)^{b-1} dt = b \left\{ \frac{1}{b} - \frac{(1-x)^b}{b} \right\} = 1 - (1-x)^b$.

With these facts, evaluate the inner integral:

$$\int_{p_A}^1 \frac{p_B^{\alpha_B - 1}(1 - p_B)^{\beta_B - 1}}{B(\alpha_B, \beta_B)} dp_B = 1 - \int_0^{p_A} \frac{p_B^{\alpha_B - 1}(1 - p_B)^{\beta_B - 1}}{B(\alpha_B, \beta_B)} dp_B = 1 - \frac{B(p_A; \alpha_B, \beta_B)}{B(\alpha_B, \beta_B)} = 1 - I_{p_A}(\alpha_B, \beta_B). \tag{2}$$

So the equation becomes:

$$\Pr(p_B > p_A) = 1 - \int_0^1 \frac{p_A^{\alpha_A - 1}(1-p_A)^{\beta_A - 1}}{B(\alpha_A, \beta_A)} I_{p_A}(\alpha_B, \beta_B) dp_A. \tag{3}$$

Now, there is a recursive relationship

$$I_x(a,b) = I_x(a-1,b) - \frac{x^{a-1}(1-x)^b}{(a-1)B(a-1,b)} \tag{4}$$

because if we denote $f(a,b,x) = I_x(a,b) - I_x(a-1,b) + \dfrac{x^{a-1}(1-x)^b}{(a-1)\,B(a-1,b)}$, $f(a,b,0) = 0$ and

$$\begin{aligned} f'(a,b,x) &= I'_x(a,b) - I'_x(a-1,b) + \frac{x^{a-2}(1-x)^{b-1}\{(a-1)(1-x) - bx\}}{(a-1)B(a-1,b)} \\ &= \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)} + \frac{x^{a-2}(1-x)^{b-1}}{B(a-1,b)} + \frac{x^{a-2}(1-x)^{b-1}\{(a-1)(1-x) - bx\}}{(a-1)B(a-1,b)} \end{aligned} \tag{5}$$

so $f'(a,b,0) = 0$ and $f(a,b,x) = 0$.

Using this relationship and the fact that α and β are integers, we can express I_x as:

$$\begin{aligned} I_x(a,b) &= I_x(a-1,b) - \frac{x^{a-1}(1-x)^b}{(a-1)B(a-1,b)} \\ &= I_x(a-2,b) - \frac{x^{a-2}(1-x)^b}{(a-2)B(a-2,b)} - \frac{x^{a-1}(1-x)^b}{(a-1)B(a-1,b)} = \cdots = 1 - (1-x)^b - \sum_{j=1}^{a-1} \frac{x^{a-j}(1-x)^b}{(a-j)B(a-j,b)} \end{aligned} \tag{6}$$

Or equivalently:

$$I_x(a,b) = 1 - \sum_{i=0}^{a-1} \frac{x^i(1-x)^b}{(b+i)B(1+i,b)} \tag{7}$$

The probability integral (3) can therefore be written:

$$\begin{aligned} \Pr(p_B > p_A) &= 1 - \int_0^1 \tfrac{p_A^{\alpha_A-1}(1-p_A)^{\beta_A-1}}{B(\alpha_A,\beta_A)} \left(1 - \sum_{i=0}^{\alpha_B-1} \tfrac{p_A^i(1-p_A)^{\beta_B}}{(\beta_B+i)B(1+i,\beta_B)}\right) dp_A \\ &= 1 - 1 + \int_0^1 \tfrac{p_A^{\alpha_A-1}(1-p_A)^{\beta_A-1}}{B(\alpha_A,\beta_A)} \sum_{i=0}^{\alpha_B-1} \tfrac{p_A^i(1-p_A)^{\beta_B}}{(\beta_B+i)B(1+i,\beta_B)} dp_A \\ &= \int_0^1 \sum_{i=0}^{\alpha_B-1} \tfrac{p_A^{\alpha_A-1+i}(1-p_A)^{\beta_A+\beta_B-1}}{(\beta_B+i)B(\alpha_A,\beta_A)B(1+i,\beta_B)} dp_A \\ &= \sum_{i=0}^{\alpha_B-1} \int_0^1 \tfrac{p_A^{\alpha_A-1+i}(1-p_A)^{\beta_A+\beta_B-1}}{(\beta_B+i)B(\alpha_A,\beta_A)B(1+i,\beta_B)} dp_A \\ &= \sum_{i=0}^{\alpha_B-1} \tfrac{B(\alpha_A+i,\beta_A+\beta_B)}{(\beta_B+i)B(\alpha_A,\beta_A)B(1+i,\beta_B)} \int_0^1 \tfrac{p_A^{\alpha_A-1+i}(1-p_A)^{\beta_A+\beta_B-1}}{B(\alpha_A+i,\beta_A+\beta_B)} dp_A \end{aligned}$$

Finally:

$$\Pr(p_B > p_A) = \sum_{i=0}^{\alpha_B-1} \frac{B(\alpha_A+i, \beta_A+\beta_B)}{(\beta_B+i)B(1+i,\beta_B)B(\alpha_A,\beta_A)} \tag{8}$$

Equivalent Formulas

It's possible to derive similar formulas that sum over the other three parameters:

$$\Pr(p_B > p_A) = 1 - \sum_{i=0}^{\alpha_A-1} \frac{B(\alpha_B+i, \beta_B+\beta_A)}{(\beta_A+i)B(1+i,\beta_A)B(\alpha_B,\beta_B)} \tag{9}$$

$$\Pr(p_B > p_A) = \sum_{i=0}^{\beta_A-1} \frac{B(\beta_B+i, \alpha_A+\alpha_B)}{(\alpha_A+i)B(1+i,\alpha_A)B(\alpha_B,\beta_B)} \tag{10}$$

$$\Pr(p_B > p_A) = 1 - \sum_{i=0}^{\beta_B-1} \frac{B(\beta_A+i, \alpha_A+\alpha_B)}{(\alpha_B+i)B(1+i,\alpha_B)B(\alpha_A,\beta_A)} \tag{11}$$

The above formulas can be found with symmetry arguments.

A/B/C testing: binary outcomes

It is possible to extend the binary-outcome formula to three test groups, call them A, B, and C. The probability that C will beat both A and B in the long run is:

$$\Pr(p_C > \max\{p_A, p_B\}) =$$
$$1 - \Pr(p_A > p_C) - \Pr(p_B > p_C) + \sum_{i=0}^{\alpha_A-1}\sum_{j=0}^{\alpha_B-1} \frac{B(i+j+\alpha_C, \beta_A+\beta_B+\beta_C)}{(\beta_A+i)B(1+i,\beta_A)(\beta_B+j)B(1+j,\beta_B)B(\alpha_C,\beta_C)}$$

Where:

- α_X is one plus the number of successes for $X \in \{A, B, C\}$
- β_X is one plus the number of failures for $X \in \{A, B, C\}$
- $\Pr(p_X > p_C)$ is the formula for the two-group case, given by (8)

Note that this formula can be computed in $O(\alpha_A \alpha_B)$ time (see the implementation section below).

Derivation

Start with a Bayesian belief for each of three experimental branches (A, B, and C):

$$p_A \sim \text{Beta}(\alpha_A, \beta_A)$$
$$p_B \sim \text{Beta}(\alpha_B, \beta_B)$$
$$p_C \sim \text{Beta}(\alpha_C, \beta_C)$$

Calling the pdf of the beta distribution $f(p|\alpha,\beta) = f(p)$, we can get the total probability that p_C is greater than both p_A and p_B by integrating the joint distribution over all values for which $p_C > p_A$ and $p_C > p_B$:

$$\Pr(\max\{p_A,p_B\} < p_C) = \int_0^1 \Pr(p_C)\Pr(p_C|\max\{p_A,p_B\} < p_C)dp_C = \int_0^1 \Pr(p_C)\Pr(p_A < p_C)\Pr(p_B < p_C)dp_C$$
$$= \int_0^1 \int_0^{p_C} \int_0^{p_C} f(p_A)f(p_B)f(p_C)dp_A dp_B dp_C$$

Evaluating the inner two integrals, the equation becomes:

$$\Pr(p_C > \max\{p_A, p_B\}) = \int_0^1 I_{p_C}(\alpha_A,\beta_A) I_{p_C}(\alpha_B,\beta_B) f(p_C)dp_C \tag{12}$$

Using the identity for I_X (7), we have:

$$\Pr(p_C > \max\{p_A,p_B\}) = \int_0^1 \left(1 - \sum_{i=0}^{\alpha_A-1} \frac{p_C^i(1-p_C)^{\beta_A}}{(\beta_A+i)B(1+i,\beta_A)}\right)\left(1 - \sum_{i=0}^{\alpha_B-1} \frac{p_C^i(1-p_C)^{\beta_B}}{(\beta_B+i)B(1+i,\beta_B)}\right) f(p_C)dp_C$$

Multiplying out the parenthetical terms and integrating them separately:

$$\Pr(p_C > \max\{p_A,p_B\}) = 1 - \int_0^1 \sum_{i=0}^{\alpha_A-1} \frac{p_C^i(1-p_C)^{\beta_A}}{(\beta_A+i)B(1+i,\beta_A)} f(p_C)dp_C - \int_0^1 \sum_{i=0}^{\alpha_B-1} \frac{p_C^i(1-p_C)^{\beta_B}}{(\beta_B+i)B(1+i,\beta_B)} f(p_C)dp_C$$
$$+ \int_0^1 \sum_{i=0}^{\alpha_A-1} \frac{p_C^i(1-p_C)^{\beta_A}}{(\beta_A+i)B(1+i,\beta_A)} \sum_{i=0}^{\alpha_B-1} \frac{p_C^i(1-p_C)^{\beta_B}}{(\beta_B+i)B(1+i,\beta_B)} f(p_C)dp_C$$

From the previous derivation, we can rewrite the first two integrals as $\Pr(p_A > p_C)$ and $\Pr(p_B > p_C)$, and consolidate the terms inside the third integral:

$$\begin{aligned}
\Pr(p_C > \max\{p_A,p_B\}) &= 1 - \Pr(p_A > p_C) - \Pr(p_B > p_C) \\
&\quad + \textstyle\int_0^1 \sum_{i=0}^{\alpha_A-1} \sum_{j=0}^{\alpha_B-1} \frac{p_C^{i+j}(1-p_C)^{\beta_A+\beta_B}}{(\beta_A+i)(\beta_B+j)B(1+i,\beta_A)B(1+j,\beta_B)} \frac{p_C^{\alpha_C-1}(1-p_C)^{\beta_C-1}}{B(\alpha_C,\beta_C)} dp_C \\
&= 1 - \Pr(p_A > p_C) - \Pr(p_B > p_C) \\
&\quad + \textstyle\int_0^1 \sum_{i=0}^{\alpha_A-1} \sum_{j=0}^{\alpha_B-1} \frac{p_C^{i+j+\alpha_C-1}(1-p_C)^{\beta_A+\beta_B+\beta_C-1}}{(\beta_A+i)(\beta_B+j)B(1+i,\beta_A)B(1+j,\beta_B)B(\alpha_C,\beta_C)} dp_C
\end{aligned}$$

Finally, evaluating the integral we have:

$$
\begin{aligned}
\Pr(p_C > \max\{p_A, p_B\}) &= 1 - \Pr(p_A > p_C) - \Pr(p_B > p_C) \\
&\quad + \sum_{i=0}^{\alpha_A - 1} \sum_{j=0}^{\alpha_B - 1} \frac{B(i+j+\alpha_C, \beta_A+\beta_B+\beta_C)}{(\beta_A+i)(\beta_B+j)B(1+i,\beta_A)B(1+j,\beta_B)B(\alpha_C,\beta_C)}
\end{aligned}
$$