Statistical Calculations in Target

This page documents the detailed statistical calculations used in manual A/Bn tests in Adobe Target. Definitions are provided for the Conversion Rate, Confidence Interval of Conversion Rate, Lift, Confidence Interval for Lift, and Confidence.

Experience	Visitors	Conversion Rate	Average Lift and Confidence Interval	Confidence
Experience A Control	45.42% 124	25.81% 32 ± 7.70%	Control	---
Experience B	54.58% 149	34.90% 52 ± 7.65%	↑ 35.23% -14.85% to 85.32%	89.81%
Activity	**100.00% 273**	**30.77% 84** --	---	---

Mean Performance

Conversion Rate and Revenue Per Visitor (RPV) Campaigns

The following illustration shows Conversion Rate, Confidence Interval of Conversion Rate and number of Conversions in Target report. For example, the first line shows that for Experience A: the conversion rate is 25.81% with a Confidence Interval of ±7.7% , and 32 conversions were recorded. Given that 124 Visitors saw the experience, this equates to 32/124 = 25.81%.

Conversion Rate
25.81% 32 ± 7.70%
34.90% 52 ± 7.65%
30.77% 84 --

The conversion rate or **mean**, μ_ν, for each experience ν in an Experiment is defined as a ratio of the sum of the metric to the number of units assigned to that metric, N_ν:

$$\hat{\mu}_\nu = \frac{1}{N_\nu} \sum_{i=1}^{N_\nu} Y_{i\nu}$$

Here,

- $Y_{i\nu}$ is the value of the metric for each unit i, that has been assigned to a given experience ν.
- The sum over units i depends on the choice of counting methodology.

- If *Visitors* is used as the counting methodology, each unit is a unique visitor defined as a unique participant in the activity for the life of the activity.
- If *Visits* is used as the counting methodology, each unit is a unique visit defined as a unique participant in an experience during an Adobe Target session (with a unique `sessionId`). When the `sessionId` changes, or the visitor reaches the conversion step, a new visit
- If *Activity Impressions* is used as the counting methodology, each unit is a unique impression defined as each time a visitor loads any page of the activity.

Confidence Interval of Mean/Conversion Rate

The confidence interval of the conversion rate is intuitively defined as range of possible conversion rates that is consistent with the underlying data.

Recall, when running experiments, the conversion rate we observe for a given Experience is an *estimate* of the "true" conversion rate. To quantify the uncertainty in this estimate, we can use a confidence interval. Adobe Target always reports a 95% confidence interval, which means that in the long run, 95% of confidence intervals calculated will include the true conversion rate of the Experience.

A 95% Confidence Interval of conversion rate μ_v is defined as the range of values:

$$[\hat{\mu}_v - 1.96\hat{\sigma}_{\hat{\mu}_v}, \hat{\mu}_v + 1.96\hat{\sigma}_{\hat{\mu}_v}]$$

where the standard error for the mean is defined as

$$\hat{\sigma}_{\hat{\mu}_v} = \frac{\hat{\sigma}_v}{N_v}$$

where an unbiased estimate of the sample standard deviation is used:

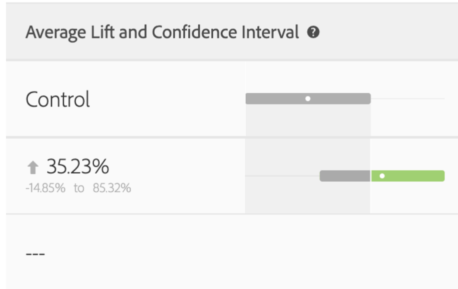
$$\hat{\sigma}_\nu = \sqrt{\frac{1}{N_\nu - 1}\sum_{i=1}^{N_\nu}(Y_{i\nu} - \hat{\mu}_\nu)^2}$$

Note that when the campaign is a conversion rate campaign (i.e., the conversion metric is binary), the standard error reduces to:

$$\hat{\sigma}_{\hat{\mu}_v} = \sqrt{\frac{\hat{\mu}_v(1 - \hat{\mu}_v)}{N_v}}$$

Lift

The following illustration shows Lift and Confidence Interval of Lift in Target Report. The number represents the average of the range of the lift bounds, and the arrow reflects if the lift is positive or negative. The arrow displays in grey until the confidence passes 95%. After confidence passes the threshold, the arrow would be green or red based on positive or negative lift.



The lift between a experience *v*, and the control experience v_0 is the relative "delta" in conversion rates, defined as

$$\delta_{\nu,\nu_0} = \frac{\hat{\mu}_{\nu} - \hat{\mu}_{\nu_0}}{\hat{\mu}_{\nu_0}}$$

where the individual conversion rates are as defined above. More simply,

```
Lift(Experience N) = (Performance_Experience_N - Performance_Control)/ Performance_Control
```

If conversion rate of control experience v_0 is 0, then there is no lift.

Confidence Interval of Lift

The boxplot graph in the `Average Lift and Confidence Interval` column represents the average value and 95% Confidence Interval of Lift. The boxplot is grey when there is any overlap in the confidence interval of a given non-control experience with the confidence interval of control experience, and is in green or red when the range of given experience's confidence interval is above or below the confidence interval of control experience.

The standard error of the lift between a experience *v*, and the control experience v_0 is defined as:

$$\sigma_{lift} = \frac{\hat{\mu}_v}{\hat{\mu}_{v_0}} * \sqrt{\left(\frac{\hat{\sigma}_{\hat{\mu}_v}}{\hat{\mu}_v}\right)^2 + \left(\frac{\hat{\sigma}_{\hat{\mu}_{v_0}}}{\hat{\mu}_{v_0}}\right)^2}$$

Then the 95% Confidence Interval of the lift is:

$$[\delta_{v,v_0} - 1.96 * \sigma_{lift}, \delta_{v,v_0} + 1.96 * \sigma_{lift}]$$

This calculation uses the "Delta" method, and is described in more detail in this document

Confidence

The final column shows the confidence in Target report.The confidence of an experience is a probability (denoted as a percentage) of obtaining a result less extreme than the one that is actually observed given the null hypothesis is true. In terms of p-values, the confidence displayed is *1 - p-value*. Intuitively, higher confidence means that it is less likely that the control and non-control experience have equal conversion rates.

In Adobe Target, a two-tailed **Welch's t-test** is performed between the test experience and the control experience to test if the means of test and control experiences are the same. Since we usually do not know if sample sizes and variances of two groups are the same before running the experiment, and Adobe Target also allows you to have unequal percentages of traffic sent to each experience, we do not assume that the variance for each experience is equal. Thus,Welch's t-test is chosen instead of Student's t-test.

To perform Welch's t-test, we first start calculating the t-statistic and the degrees of freedom, then run a two-tailed t-test to generate the p-value. Finally, we calculate the confidence based on p-value.

The *t*-statistic is defined to be the difference of the means of any two independent random variables, v and v_0, divided by the standard error of the difference:

$$t = \frac{\hat{\mu}_\nu - \hat{\mu}_{\nu_0}}{\hat{\sigma}_{\Delta\mu}}$$

where μ_v and μ_{v0} are the means of v and v_0 respectively, and the standard error of the difference between μ_v and μ_{v0} are given by:

$$\hat{\sigma}_{\Delta\mu} = \sqrt{\frac{\hat{\sigma}_\nu^2}{N_\nu} + \frac{\hat{\sigma}_{\nu_0}^2}{N_{\nu_0}}}$$

where $\sigma^2{}_v$ and $\sigma^2{}_{v_0}$ are the variances of two experiences v and v_0 respectively, and N_v and N_{v_0} are sample sizes for v and v_0 respectively.

For Welch's t-test, the degree of freedom is calculated as following:

$$df = \frac{(\frac{\sigma_v^2}{N_v} + \frac{\sigma_{v_0}^2}{N_{v_0}})^2}{(\frac{\sigma_v^4}{N_v^2 df_v} + \frac{\sigma_{v_0}^4}{N_{v_0}^2 df_{v_0}})^2}$$

and degree of freedom for v and v_0 are defined as:

$$df_v = N_v - 1$$

$$df_{v_0} = N_{v_0} - 1$$

Then the p-value can be computed from the area in the tails of the *t*-distribution:

$$p = t.dist.2t(|t|, df)$$

Finally, the Confidence reported in Target is defined as:

$$Confidence = 1 - p$$