Technical Report

Rapid statistical calculations for determining the success of marketing campaigns

Received (in revised form): 7th January, 2002

Bruce Ratner

is President of DM STAT-1 CONSULTING, the leading firm for analysis and modelling in the DM industry, specialising in statistical methods and knowledge discovery and data-mining tools, in the areas of banking, insurance, finance, retail, telecommunications, business-to-business, and catalogue marketing. Bruce is the author of the *DM STAT-1 Newsletter* on the Web at www.dmstat1.com

Abstract People are always performing pencil-and-paper calculations, despite computers continuously evolving into smarter thinking machines for analysing data. Whether at home balancing the cheque book, or in the office working out the markdown to close the deal, people are, at one time or another, doing some kind of rapid pencil-and-paper calculations. In direct marketing, data analysts routinely perform rapid statistical calculations for analysing campaign results. This paper details rapid calculations for five popular statistical significance tests (decision rules) for determining the success of marketing campaigns. Additionally, a sixth decision rule, albeit novel and virtually unknown to statisticians, is presented to contrast today's thinking that information can only be extracted from big data.

This paper will be a welcomed addition to direct marketers who analyse campaign results. Statistical data analysts with a basic understanding of the concepts of significance level and critical values will use this paper as a handy reference guide, eliminating the need to look up the significance tests in their statistics books. Marketing analysts with limited statistical background can use the decision rules like well-tested recipes to yield the correct results every time.

DECISION RULE 1: FOR CAMPAIGN RESPONSE RATE GREATER THAN TARGET RESPONSE RATE

Situation: Campaign of sample size n achieves a response rate of $RR\%$. The campaign is successful if its response rate $RR\%$ is greater than the preselected target response rate $TR\%$.

*Method:*¹

Test-Statistic (TS) =

$$\frac{RR\% - TR\%}{\text{square root of } (TR\% \star (1 - TR\%)/n)}$$

Decision Rule (RR% is greater than TR%):

If TS is equal to or greater than the one-tailed critical value for significance level $\alpha\%$, $CV(1 - t, \alpha\%)$, then the campaign is successful at the $(100 - \alpha)\%$ confidence level; otherwise, if TS is less than $CV(1 - t, \alpha\%)$ then it cannot be concluded that the campaign is successful at the $(100 - \alpha)\%$ confidence level.² See Table 1.

Decision Rule (RR% is different than TR%): If absolute value³ of TS is equal to or greater than the two-tailed critical value for significance level $\alpha\%$, $CV(2 - t,$

Bruce Ratner
President, DM STAT-1
Consulting, 574 Flanders
Drive, North Woodmere,
NY 11581, USA.

Tel: +1 516 791 3544;
Fax: +1 516 791 5075;
e-mail: br@dmstat1.com

Table 1: Critical values

Significance level (a%)		CV(<i>i</i> - <i>t</i> , a%)
Two-tailed (2 - <i>t</i>) %	One-tailed (1 - <i>t</i>) %	
1.0	0.5	2.58
2.0	1.0	2.33
5.0	2.5	1.96
10.0	5.0	1.64
20.0	10.0	1.28

a%), then the campaign response rate $RR\%$ is different than the target response rate $TR\%$ at the $(100 - a)\%$ confidence level; otherwise, if TS is less than $CV(2 - t, a\%)$ then it cannot be concluded that the campaign response rate $RR\%$ is different than the target response rate $TR\%$ at the $(100 - a)\%$ confidence level.

Assumption: The test is appropriate for large samples (n greater than 30), and when $TR\% \star n$ greater than 5, and $(1 - TR\%) \star n$ greater than 5.

Example:

$$n = 1,000.$$

$$RR\% = 5.0\% \text{ and } TR\% = 4.0\%.$$

$$TS = 1.61$$

- TS is greater than the 10.0 per cent significance level/one-tailed critical value 1.28. Thus, the campaign is successful at the 90 per cent confidence level.
- TS is greater than the 20.0 per cent significance level/two-tailed critical value 1.28. Thus, the campaign response rate is different than the target response rate at the 80 per cent confidence level.

DECISION RULE 2: FOR TEST CAMPAIGN RESPONSE RATE GREATER THAN CONTROL CAMPAIGN RESPONSE RATE

Situation: Test and control campaigns of sample sizes n_t and n_c achieve response

rates, $RR_t\%$ and $RR_c\%$, respectively. The test campaign is successful if its response rate $RR_t\%$ is greater than the control response rate $RR_c\%$.

*Method:*⁴

Test-Statistic (TS) =

$$\frac{RR_t\% - RR_c\%}{\text{square root of } (RR\% \star (1 - RR\%) \star (1/n_t + 1/n_c))}$$

$$\text{where } RR\% = \frac{RR_t\% \star n_t + RR_c\% \star n_c}{n_t + n_c}$$

Decision Rule ($RR_t\%$ is greater than $RR_c\%$):

If TS is equal or greater than the one-tailed critical value for significance level $a\%$, $CV(1 - t, a\%)$, then the test campaign is successful at the $(100 - a)\%$ confidence level; otherwise, if TS is less than $CV(1 - t, a\%)$ then it cannot be concluded that the test campaign is successful at the $(100 - a)\%$ confidence level.

Decision Rule ($RR_t\%$ is different than $RR_c\%$): If absolute value of TS is equal or greater than the two-tailed critical value for significance level $a\%$, $CV(2 - t, a\%)$, then the test response rate $RR_t\%$ is different than the control response rate $RR_c\%$ at the $(100 - a)\%$ confidence level; otherwise, if TS is less than $CV(2 - t, a\%)$ then it cannot be concluded that the test campaign response rate $RR_t\%$ is different than the control response rate $RR_c\%$ at the $(100 - a)\%$ confidence level.

Assumption: The test is appropriate for large samples (n greater than 30), and when $RR\% \star n_t$ greater than 5, $RR\% \star n_c$ greater than 5, $(1 - RR\%) \star n_t$ greater than 5, and $(1 - RR\%) \star n_c$ greater than 5.

Example:

$$n_t = 1,000 \text{ and } n_c = 1,000.$$

$$RR_t\% = 5.0\% \text{ and } RR_c\% = 4.0\%.$$

$$RR\% = 4.5\%$$

$$TS = 1.08.$$

- TS is less than all critical values in Table 1. Thus, it cannot be concluded that the test campaign is successful at any of the corresponding confidence levels.
- Additionally, it cannot be concluded that test campaign response rate is different than the control response rate at any of the corresponding confidence levels.

DECISION RULE 3: FOR CAMPAIGN MEAN PROFIT GREATER THAN TARGET MEAN PROFIT

Situation: Campaign of sample size n achieves a mean profit of MP. The campaign is successful if its mean profit MP is greater than the preselected target mean profit TP.

*Method:*⁵

$$\text{Test-Statistic (TS)} = \frac{\text{MP} - \text{TP}}{\text{stdDev}/\text{square root of } (n)}$$

stdDev = standard deviation of n profit values, P_i

$$\text{stdDev} = \text{square root} (\text{sum } ((P_i - \text{MP}) * (P_i - \text{MP})) / (n - 1))$$

Decision Rule (MP is greater than TP): If TS is equal or greater than the one-tailed critical value for significance level $\alpha\%$, $CV(1 - t, \alpha\%)$, then the test campaign is successful at the $(100 - \alpha)\%$ confidence level; otherwise, if TS is less than $CV(1 - t, \alpha\%)$ then it cannot be concluded that the test campaign is successful at the $(100 - \alpha)\%$ confidence level.

Decision Rule (MP is different than TP): If TS is equal or greater than the two-tailed critical value for significance level $\alpha\%$, $CV(2 - t, \alpha\%)$, then the campaign mean profit is different than the target mean profit at the $(100 - \alpha)\%$

confidence level; otherwise, if TS is less than $CV(2 - t, \alpha\%)$ then it cannot be concluded that the campaign mean profit is different than the target mean profit at the $(100 - \alpha)\%$ confidence level.

Assumption: The test is appropriate for large samples (n greater than 30) and profit values that are bell-shaped.⁶ This test is an approximate guide for profit values that are not bell-shaped.

Example:

$$n = 1,000.$$

$$\text{MP} = \$50.00 \text{ and } \text{TP} = \$48.00.$$

$$\text{stdDev} = 35.$$

$$\text{TS} = 1.81$$

- TS is greater than the 5.0 per cent significance level/one-tailed critical value 1.64. Thus, the campaign is successful at the 95 per cent confidence level.
- TS is greater than the 10.0 per cent significance level/two-tailed critical value 1.64. Thus, the campaign mean profit is different than the target mean profit at the 90 per cent confidence level.

DECISION RULE 4: FOR TEST CAMPAIGN MEAN PROFIT GREATER THAN CONTROL CAMPAIGN MEAN PROFIT

Situation: Test and control campaigns of sample sizes n_t and n_c achieve mean profits, MP_t and MP_c , respectively. The test campaign is successful if its mean profit MP_t is greater than the control mean profit MP_c .

*Method:*⁷

$$\text{Test-Statistic (TS)} = \frac{\text{MP}_t - \text{MP}_c}{\text{square root of } ((\text{stdDev}_t^2/n_t) + (\text{stdDev}_c^2/n_c))}$$

stdDev_t = standard deviation of n_t Test profit values, T_i

stdDev_c = standard deviation of n_c
 Control profit values, C_i
 stdDev_t = square root (sum
 $((T_i - MP_t) \star (T_i - MP_t)) / (n_t - 1))$
 stdDev_c = square root (sum
 $((C_i - MP_c) \star (C_i - MP_c)) / (n_c - 1))$

Decision Rule (MP_t is greater than MP_c): If TS is equal or greater than the one-tailed critical value for significance level $\alpha\%$, $CV(1 - t, \alpha\%)$, then the test campaign is successful at the $(100 - \alpha)\%$ confidence level; otherwise, if TS is less than $CV(1 - t, \alpha\%)$ then it cannot be concluded that the test campaign is successful at the $(100 - \alpha)\%$ confidence level.

Decision Rule (MP_t is different than MP_c): If TS is equal or greater than the two-tailed critical value for significance level $\alpha\%$, $CV(2 - t, \alpha\%)$, then the test mean profit is different than the control mean profit at the $(100 - \alpha)\%$ confidence level; otherwise, if TS is less than $CV(1 - t, \alpha\%)$ then it cannot be concluded that the test mean profit is different than the control mean profit at the $(100 - \alpha)\%$ confidence level.

Assumptions: The test is appropriate if both the test and control profit values are bell-shaped, or if both sample sizes are large (both n_t and n_c are greater than 30). This test is an approximate guide for profit values that are not bell-shaped.

Example:

$n_t = 1,000$ and $n_c = 1,000$.
 $MP_t = \$50$ and $MP_c = \$48$
 $\text{stdDev}_t = 35$ and $\text{stdDev}_c = 25$.
 $TS = 1.47$

- TS is greater than the 10.0 per cent significance level/one-tailed critical value 1.28. Thus, the campaign is successful at the 90 per cent confidence level.
- TS is greater than the 20.0 per cent significance level/two-tailed critical value 1.28. Thus, the campaign mean

profit is different than the target mean profit at the 80 per cent confidence level.

DECISION RULE 5: FOR EQUALITY BETWEEN TWO CORRELATED RESPONSE RATES

For the situation below, the previous decision rule 2 is often erroneously used, as the correct decision rule 5 is not well known.

Situation: Campaign of sample size n offers two products, A and B. Individuals respond 'yes' or 'no' to their preference for/likelihood-to-buy each product. A two-by-two table can be constructed indicating the cross-counts of individual responses: a individuals prefer both products A and B, b individuals prefer product B but not product A; c individuals prefer product A but not product B; and d individuals prefer neither product A nor B. Product A and B response rates are now easily calculated: $RR_A\%$ ($= (a + c)/n$) and $RR_B\%$ ($= (a + b)/n$), respectively. See Table 2.

Is product A response rate $RR_A\%$ greater than product B response rate $RR_B\%$? Is product A response rate $RR_A\%$ different than product B $RR_B\%$?

*Method:*⁸

$$\text{Test-Statistic (TS)} = \frac{b - c}{\text{square root of } (b + c)}$$

Decision Rule ($RR_A\%$ is greater than $RR_B\%$): If TS is equal or greater than the one-tailed critical value for significance level $\alpha\%$, $CV(1 - t, \alpha\%)$, then

Table 2: Cross-counts of responses

		Product A	
Product B	Yes	a	b
	No	c	d
		n	

Table 3

Significance level (a%)					
c	20%	15%	10%	5%	1%
	2.42	3.23	4.84	9.68	48.39

the product A response rate $RR_A\%$ is greater than product B response rate $RR_B\%$ at the $(100 - a)\%$ confidence level; otherwise, if TS is less than $CV(1 - t, a\%)$ then it cannot be concluded that the product A response rate $RR_A\%$ is greater than product B response rate $RR_B\%$ at the $(100 - a)\%$ confidence level.

Decision Rule ($RR_A\%$ is different than $RR_B\%$): If TS is equal or greater than the two-tailed critical value for significance level $a\%$, $CV(2 - t, a\%)$, then product A response rate $RR_A\%$ is different than product B response rate $RR_B\%$ at the $(100 - a)\%$ confidence level; otherwise, if TS is less than $CV(2 - t, a\%)$ then it cannot be concluded that product A response rate $RR_A\%$ is different than product B response rate $RR_B\%$ at the $(100 - a)\%$ confidence level.

Assumption: The test is appropriate when $b + c$ is greater than 10.

Example:

$a = 20$, $b = 40$, $c = 25$ and $d = 35$.
 $RR_A\% = 37.5\%$ and $RR_B\% = 50\%$.
 $TS = 1.86$

- TS is greater than the 5.0 per cent significance level/one-tailed critical value 1.64. Thus, product A response rate is greater than product B response rate at the 95 per cent confidence level.
- TS is greater than the 10.0 per cent significance level/two-tailed critical value 1.64. Thus, product A response rate is different than product B response rate at the 90 per cent confidence level.

DECISION RULE 6: FOR CONFIDENCE INTERVAL FROM A SAMPLE OF SIZE 1⁹

Decision rule 6 is virtually unknown to statisticians. Although discovered in 1966 by an electrical engineer,¹⁰ this decision rule remains unknown in database marketing perhaps because it has no practical application. The author feels, however, that it does have 'significance' when discussing statistical tests. This rule is presented to contrast today's thinking that 'golden nuggets' of information can only be extracted from big data via current data mining-computer learning procedures. When a single observation is the only available datum, the numerical value itself could be the golden nugget.

The following situation is counterintuitive in that it contradicts the standard statistical intuition that at least two observations are necessary to calculate a variance. It is possible, however, to construct a confidence interval for a population mean from a sample of size 1. The implication is that the single observation itself tells something about the variability of the data as well as the mean.

Situation: Consider a sample of size one, the single observation X . What is the $(100 - a)\%$ confidence level for the population mean?

Method: Confidence interval for a population mean from a sample of size 1.

$X \pm c|X|$ is the $(100 - a)\%$ confidence interval for population mean.¹¹

Assumption: The single observation is from a normal population with variance

unknown. This test is an approximate guide for a unimodal and symmetric population.

Example:

$X = 20$.

90% confidence interval for the population mean:

$(20 \pm 4.84 \cdot 20) = (-76.8, 116.80)$.

References

- 1 Z-test for proportion.
- 2 When statistical significance is not reached, as indicated when TS is less than $CV(1 - t, \alpha\%)$, the conclusion is: 'fail-to-reject' the null hypothesis H_0 ($H_0: RR\% = TR\%$ vs. $H_1: RR\% > TR\%$). This is not equivalent to accepting H_0 (eg $RR\% = TR\%$).
- 3 Absolute value of TS, $abs(TS)$, is the nonnegative numeric value of TS, eg if $TS = -4$ then $abs(TS) = +4$; if $TS = +4$ then $abs(TS) = +4$.
- 4 Z-test for equality of two proportions.
- 5 t -test for a population mean with variance unknown.
- 6 Formally, profit values from a normal population.
- 7 t -test for two population means with variances unknown and unequal.
- 8 Z-test for two correlated proportions (Kanji, G. K. (1993) '100 statistical tests', Sage Publications).
- 9 Machol, R. E. (1996) 'Confidence interval based on single observation', *Proceedings of the IEEE*, Vol. 55, p. 122.
- 10 Wall, M. M., Boen, J. and Tweedie, R. (2001) 'An effective confidence interval for the mean with samples of size one and two', *The American Statistician*, Vol. 55, No. 2, May.
- 11 $\Pr(X - c|X| < M < X + c|X|) > (100 - \alpha)\%$, where M is the population mean.