Market Basket Comparison Evaluation Statistics

NOTE: The purpose of this section is to provide a reader with an understanding of the inputs NYS used for the statistical test. This document is not intended to inform the reader of the general principles of statistical analyses.

NYS is using a *One-Sample Test for Proportion*[1] to determine the competiveness of one bidder's Published Price List over another in a sub-lot. Specifically, NYS is using the *Z Statistic*[2] which has the following criteria:

1. Population distribution approximates the *Normal Distribution*[3] (required). In a proportion test, each trial has only two possible outcomes – Bidder wins or loses. Hence, the distribution follows that of a *Binomial Distribution*[4]. Based on the *Central Limit Theorem*[5], the *Normal Distribution* can approximate the *Binomial Distribution* with reasonably large sample sizes. Therefore, criteria one for using the *Z Statistic* is met.

2. Sample size needs to be larger than 30 (preferred)[6]. NYS has elected to use a sample size of 40 items. Therefore, criteria two for using the Z *Statistic* is met.

With any statistical test, confidence level needs to be pre-set based on the level of error the tester is willing to allow. NYS has chosen an industry standard of 99% confidence[7] for its confidence level. This means that the State is willing to allow for a 1% chance that the conclusion is inaccurate.

In a *One-Sample Test for Proportion*, a sample of the larger population is taken and a *Sample Proportion* is calculated. In this case, the *Sample Proportion* equates to the number of items one bidder wins on price (x) divided by the total sample size (n) or *Sample Proportion* = x/n.

Based on a 99% confidence level and a sample size of 40, *Z Statistics* were run to determine the *Sample Proportion* required to conclude that a given bidder will have a lower price more than 50% of the time. The result of the *Z Statistics* showed that the minimum *Sample Proportion* for this test is 28/40. This means that the bidder must have lower pricing on at least 28 of the 40 items in the sample.

The following elaborates on the process used with *Z Statistics* to determine the minimum *Sample Proportion* stated earlier:

1. Calculated *True Proportion Intervals*[8] for various *Sample Proportions*.

2. Identified the minimum *Sample Proportion* that would result in a *True Proportion Interval* that had a lower limit greater than .5 (if the lower limit is more than .5, this means that this Sample Proportion will yield a winner more than 50% of the time).

   For example, a 25/40 sample proportion at a 95% confidence level results in a *True Proportion Interval* of 0.475 < True Proportion < 0.775. This means that for a 100,000 item Published Price List, there is a 95% chance that the tested bidder has anywhere from 47,500 – 77,500 items with lower pricing. Since this range contains 50,000 (break-even point), we cannot conclude that the tested bidder will have lower pricing more often on any given item.

   By increasing the sample proportion to 28/40 and the confidence level to 99%, the *True Proportion Interval* increases to .513 < True Proportion < .887. Since the lower end of this interval does is greater than.5, which equates to above the 50,000 item break-even point in the 100,000 item Published Price List example, we can conclude that the tested bidder will have lower pricing more often on any given item.

   Any sample proportion greater than 28/40 will continue to increase both ends of the *True Proportion Interval*. This translates to an even greater chance that the tested bidder will have lower pricing on any given item.

It is important to note that this statistical evaluation method only tests for the frequency by which one bidder beats another but does not consider by how much. The approach assumes that regardless of how frequent one bidder wins over the other, the probability that a given win is a $1 or $100 price difference is equal for all bidders being compared.

[1] *One Sample Test for Proportion* is a statistical test that can determine the likelihood of a certain proportion occurring in a greater population.
[2] *Z Statistic* is a statistical construct used to perform statistical testing, such as the *One Sample Test for Proportion*. Confidence Intervals, such as the *True Proportion Interval*, are common outputs of statistical testing with the *Z Statistic.*
[3] *Normal Distribution* is commonly referred to as the “bell curve”. The center contains the greatest number of occurrences of a possible outcome.
[4] *Binomial Distribution* characterizes a probability curve that results from multiple trials of an experiment that has only 2 possible outcomes (win or loss, heads or tails).
[5] *Central Limit Theorem* allows the *Normal Distribution* to approximate the *Binomial Distribution*. Typically for normal approximated *Binomial Distributions*, the following conditions also need to be met: $n*p>5$ and $n(1-p)>5$, where n is the sample size and p is the *Sample Proportion*. This is because the *Normal Approximation* becomes less accurate towards the proportion extremes of 1 and 0. Since NYS is only concerned with which bidder wins more than 50% of the time, these additional conditions do not apply. Exact binomial *True Proportion Intervals* for *Sample Proportions* close to 1 are still greater than .5.
[6] Statistics in Research and Development Second Edition by R. Caulcutt cites “…one-sample [proportion] test must only be used with a large sample (i.e. greater than 30)...”
[7] Hardwick Market Research Services cites 95% confidence level as industry standard for most research studies. 99% confidence level increases the accuracy and reduces the margin of error.
[8] *True Proportion Interval* is a statistical output of the Z *Statistic* and provides a range of values where the true proportion of the population lies. The true proportion is the actual proportion of items one bidder has lower price than another if every item on each bidder’s Published Price List were compared/sampled.