Alternative Approaches to Significance Testing with Weighted Means

Dr. Albert Madansky
Vice President, The Analytical Group, Inc.
and
H.G.B. Alexander Professor Emeritus of Business Administration
Graduate School of Business
University of Chicago

The creators of statistical processing software for the marketing research community have confronted them with a variety of approaches in dealing with significance testing relating to weighted sample means. Each of these approaches produces a different standard error of the weighted sample mean, and thus a different test statistic. The purpose of this note is to sort through these approaches, explain their bases in as nontechnical way as I can, and make some recommendations.

We start with a random sample of n observations, say $x_1, x_2, \ldots, x_n$, drawn from a population. For some reason we wish to weight each of these observations, with weights $w_1, w_2, \ldots, w_n$, and calculate a weighted mean $\overline{x}_w$ given by the expression

$$\overline{x}_w = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}$$

Finally, we wish to perform some statistical test of significance using this weighted mean (e.g., compare it to some hypothetical value μ_0, or compare it to another weighted mean $\overline{y}_w$).

As we learned in our Statistics courses, it is necessary to determine the standard error of $\overline{x}_w$ to perform such a test. In general the variance of the weighted mean is

$$Var(\overline{x}_w) = \frac{\sum_{i=1}^{n} w_i^2 Var(x_i)}{(\sum_{i=1}^{n} w_i)^2}$$

If each of the x=s has the same variance, σ^2, then this reduces to

$$Var(\overline{x}_w) = \frac{\sum_{i=1}^{n} w_i^2 \sigma^2}{(\sum_{i=1}^{n} w_i)^2} = \frac{\sigma^2}{e}$$

where the "effective sample size" e is given by

$$e = \frac{(\sum_{i=1}^{n} w_i)^2}{\sum_{i=1}^{n} w_i^2}$$

Thus the standard error of the weighted mean $\overline{x}_w$ is given by $\sigma / \sqrt{e}$. (This is the analogue to the formula for the standard error of the unweighted mean, namely $\sigma / \sqrt{n}$.)

The problem for the market researcher arises because three different software systems, WinCross, SPSS, and CfMC=s Mentor, take different tacks in estimating σ^2. SPSS uses the estimator

$$s_w^2 = \frac{\sum_{i=1}^{n} w_i (x_i - \overline{x}_w)^2}{\sum_{i=1}^{n} w_i - 1}$$

The rationale behind this estimator is that one should treat each weight w_i as the number of times the observations x_i is to be "replicated" in the sample, and calculate the estimate of the variance the way one would if one had "replicated" data. (The w_i are not necessarily integers; SPSS uses the noninteger values of w_i in this part of the computation.)

Unfortunately, s_w^2 is a biased estimate of σ^2. SPSS has been aware of this for sometime, but unfortunately has not made the appropriate correction. Moreover, instead of using s_w^2/e as its estimate of the variance of $\overline{x}_w$, SPSS calculates a "weighted sample size" c, as the sum of the weights

$$c = \sum_{i=1}^{n} w_i$$

and uses s_w^2/c as its estimate of the variance of $\overline{x}_w$, further contributing to the bias of its estimate of the variance of $\overline{x}_w$.

The Mentor software corrects for the bias in s_w^2 and uses as its estimate of σ^2 the estimator

$$s_c^2 = \frac{(\sum_{i=1}^{n} w_i)\sum_{i=1}^{n} w_i (x_i - \overline{x}_w)^2}{(\sum_{i=1}^{n} w_i)^2 - \sum_{i=1}^{n} w_i^2}$$

It also correctly divides s_w^2 by e to estimate the variance of $\overline{x}_w$ as s_c^2/e.

WinCross uses the usual unweighted estimate of the variance, σ^2, namely

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}$$

in its software. It, too, is an unbiased estimate of σ^2, and so WinCross estimates the variance of $\overline{x}_w$ as s^2/e.

How does one choose between the two unbiased estimates of σ^2, WinCross's s^2 and Mentor's s_c^2? By selecting the one that most closely estimates σ^2, that is, the one whose sampling distribution has the smallest variance. One can derive mathematically that the variance of s^2 is $2\sigma^4/(n-1)$ and that the variance of s_c^2 is

$$Var(s_c^2) = 2\sigma^4 \frac{\sum_{i=1}^{n} w_i^2 (\sum_{i=1}^{n} w_i)^2 - 2\sum_{i=1}^{n} w_i^3 \sum_{i=1}^{n} w_i + (\sum_{i=1}^{n} w_i^2)^2}{(\sum_{i=1}^{n} w_i)^4 - 2\sum_{i=1}^{n} w_i^2 (\sum_{i=1}^{n} w_i)^2 + (\sum_{i=1}^{n} w_i^2)^2}$$

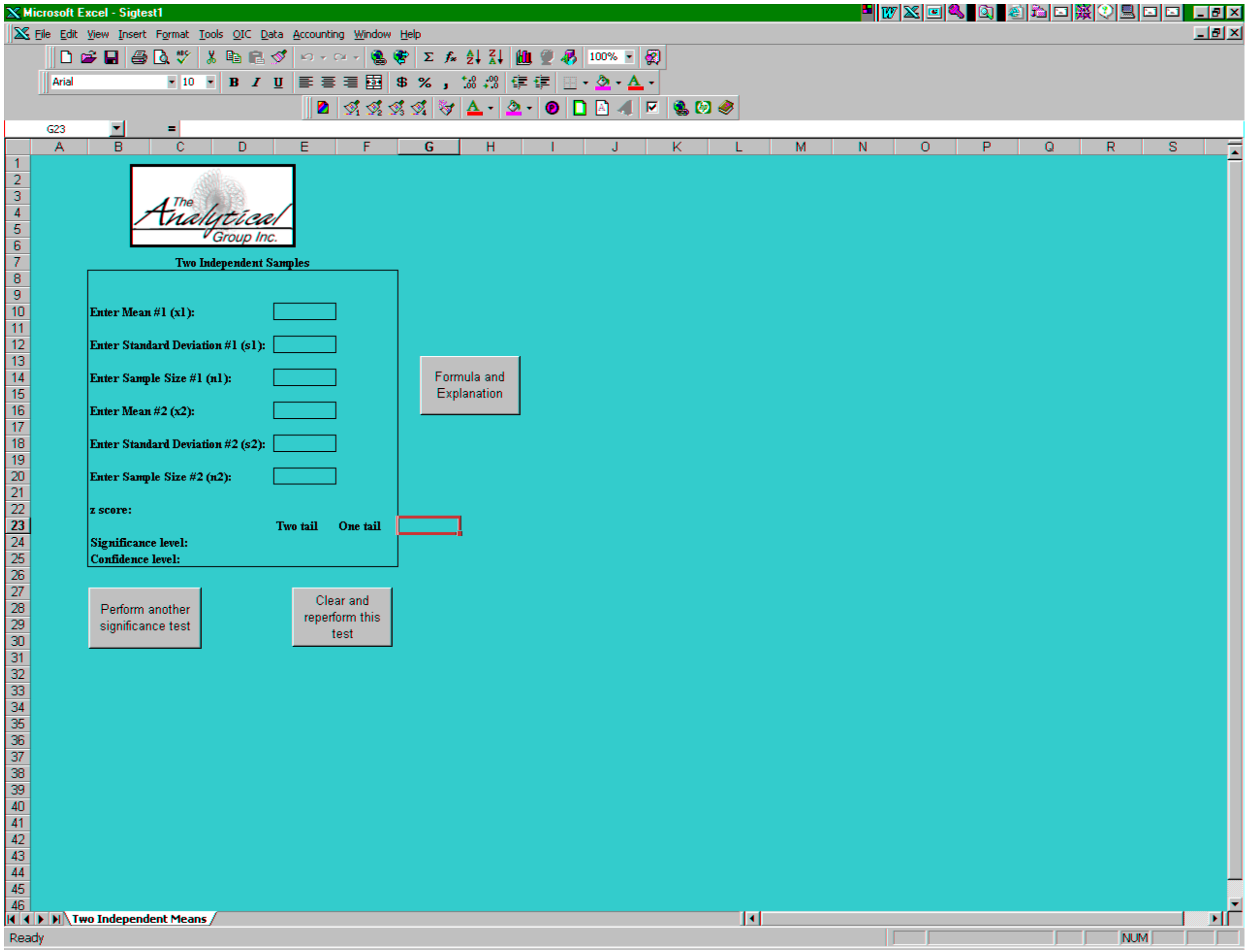
The derivation of these results, and the proof that $Var(s^2) \leq Var(s_c^2)$, so that WinCross's estimator is the preferred estimator of the variance of $\overline{x}_w$, can be found on The Analytical Group, Inc. website:

http://www.analyticalgroup.com/download/WEIGHTED MEAN.pdf

The following example illustrates how discrepant these two estimators will be. I selected as weights 100 random numbers from a uniform distribution between 0 and 1. These weights, along with their squares and cubes, are given on The Analytical Group, Inc. website. The variance of s^2, excluding the factor $2\sigma^4$, is 1/99= 0.010101. The variance of s_c^2, again excluding the factor $2\sigma^4$, is 0.014756. Thus use of s_c^2 will produce an estimate of the variance of $\overline{x}_w$ which is 1.46 times the variance produced using s^2 A simulation experiment using 1,000 replicates of 100 observations drawn from a single population and these weights, illustrating the three approaches and their resulting estimates of the variance of $\overline{x}_w$, can be found on The Analytical Group, Inc. website:

http://www.analyticalgroup.com/download/simulation.pdf

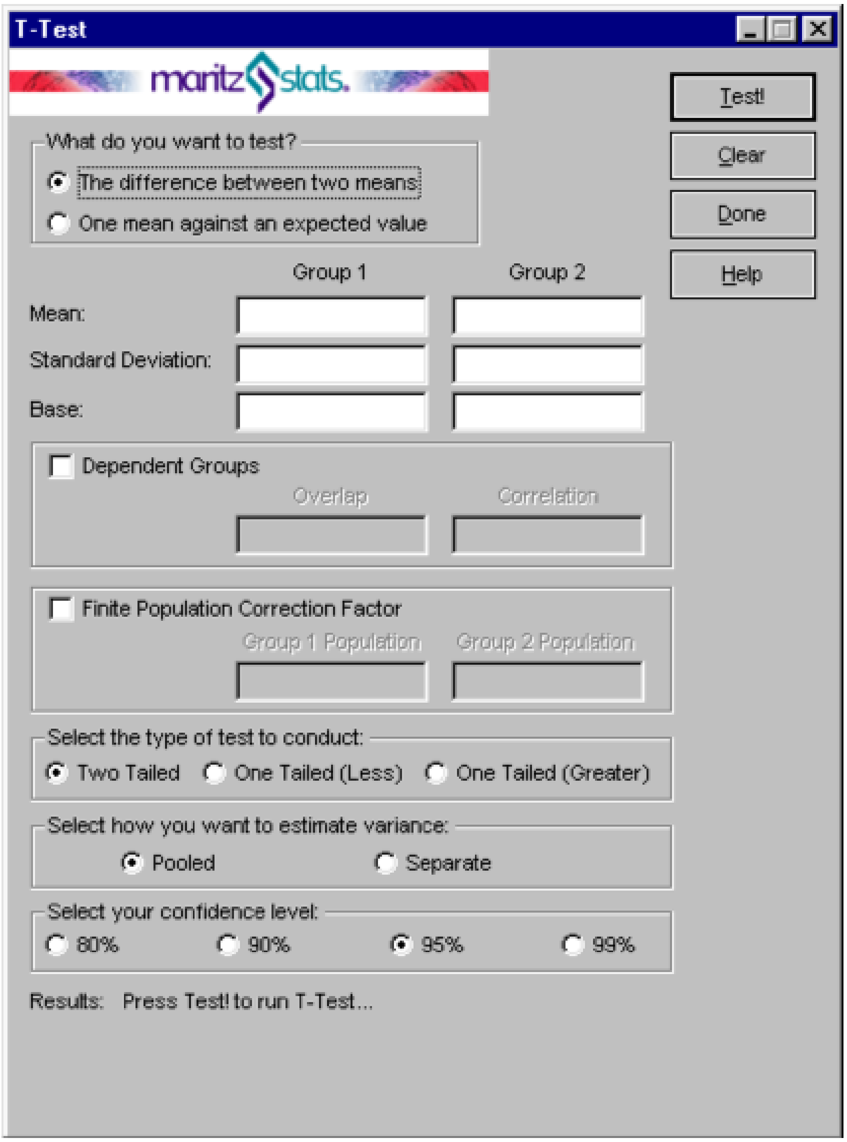
How does one use these statistics in standard statistical testing templates? Here is the T-Test template provided by The Analytical Group. To obtain a copy of this program visit: http://www.analyticalgroup.com/download/sigtest.xls



If one wanted to use this to compare weighted means, one need only fill in the weighted mean $\bar{x}_w$ in the box denoted **Enter Mean**, the unweighted standard deviation s in the box denoted **Enter Standard Deviation**, and the effective sample size e in the box denoted **Enter Sample Size** and you will have emulated the WinCross procedure for hypothesis testing using weighted means. (To emulate the Mentor procedure, use the bias-corrected weighted standard deviation s_c in the box denoted **Enter Standard Deviation** , and the effective sample size e in the box denoted **Enter Sample Size**. To emulate the SPSS procedure, use the biased weighted standard deviation s_w in the box denoted **Enter Standard Deviation**, and the weighted sample size c in the box denoted **Enter Sample Size**.)

And here is the T-Test template for using Maritz Stats. To obtain a copy of this program visit:

http://www.maritzresearch.com/contactformresearchstats.asp

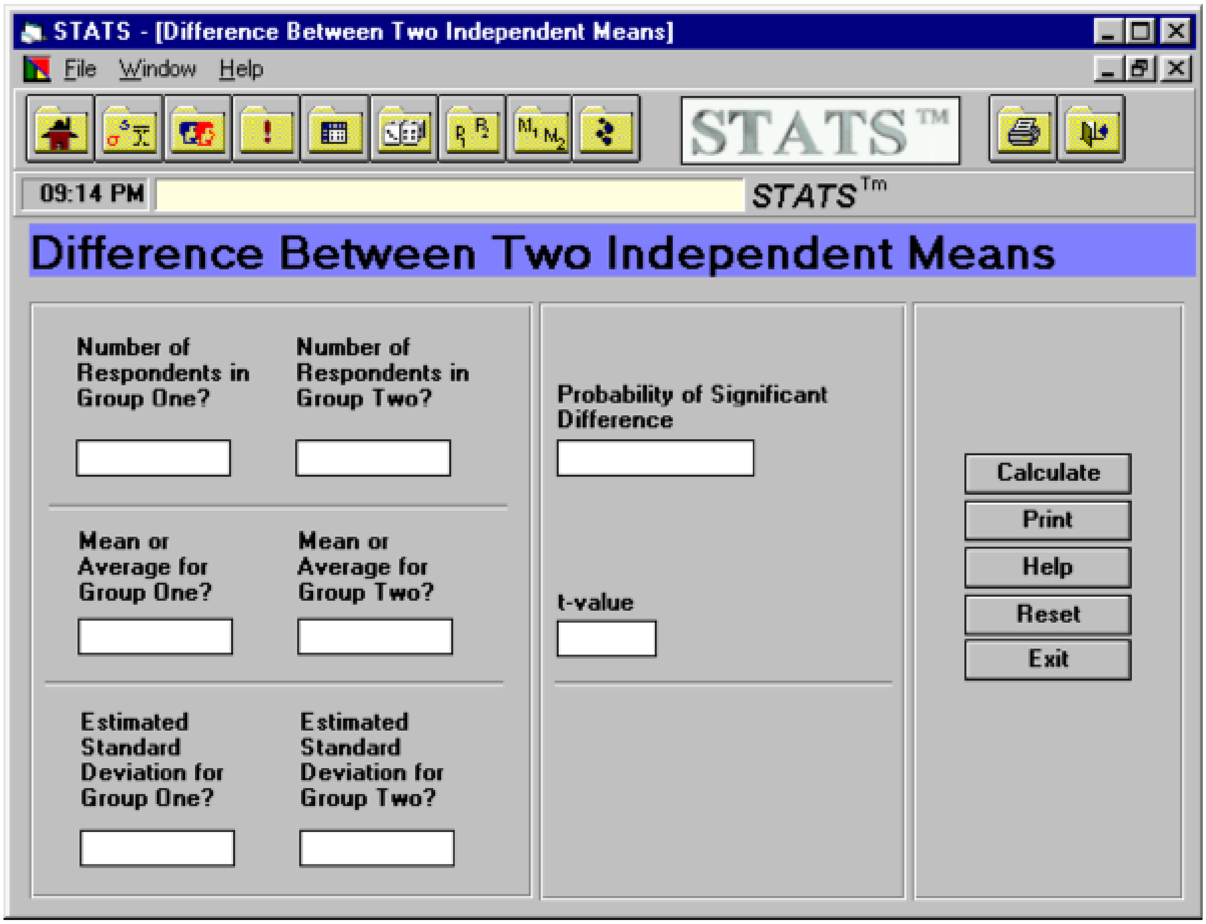


If one wanted to use this to compare weighted means, one need only fill in the weighted mean $\bar{x}_w$ in the box denoted **Mean:**, the unweighted standard deviation s in the box denoted **Standard Deviation:**, and the effective sample size e in the box denoted **Base:** and you will have emulated the WinCross procedure for hypothesis testing using weighted means. (To emulate the Mentor procedure, use the bias-corrected weighted standard deviation s_c in the box denoted **Standard Deviation:** , and the effective sample size e in the box denoted **Base:**. To emulate the SPSS procedure, use the biased weighted standard deviation s_w in the box denoted **Standard**

Deviation:, and the weighted sample size c in the box denoted **Base:**.)

Finally here is the T-Test template for using Decision Analyst=s STATS. To obtain a copy of this program visit

http://www.decisionanalyst.com/download.asp



If one wanted to use this to compare weighted means, one need only fill in the weighted mean $\bar{x}_w$ in the box denoted **Mean or Average**, the unweighted standard deviation s in the box denoted **Estimated Standard Deviation**, and the effective sample size e in the box denoted **Number of Respondents** and you will have emulated the WinCross procedure for hypothesis testing using weighted means. (To emulate the Mentor procedure, use the bias-corrected weighted standard deviation s_c in the box denoted **Estimated Standard Deviation**, and the effective sample size e in the box denoted **Number of Respondents**. To emulate the SPSS procedure, use the biased weighted standard deviation s_w in the box denoted **Estimated Standard Deviation**, and the weighted sample size c in the box denoted **Number of Respondents**.)

Given both the bias in the SPSS estimate of σ^2 and its incorrect denominator in determining the standard error of $\overline{x}_w$, the probabilities calculated based on the t-statistic will be incorrect. The probabilities based on both the WinCross and Mentor statistics will be correct, but, because Mentor uses an estimate of the variance of $\overline{x}_w$ with a larger variance than that of the estimate used by WinCross, it is more likely that one will find fewer significant differences using the Mentor procedure than using the WinCross procedure.