

The Population Stability Index (PSI) was a great idea when it was developed around 1975. However, it is well past its use-by date. The best measure of population stability is the Kolmogorov-Smirnov statistic and associated p-level. This is the standard statistical test for the hypothesis that two samples (standard and current) come from the same population.

There is no reason to continue to use the Population Stability Index. The problems with PSI are:

1) There is no measure of its statistical properties (bias and variance notably). This matters a lot where your current sample is small.

2) PSI (like Information Value) is upwardly biased because it picks up "noise" (random fluctuation) as well as signal ("real" differences). Based on some simulations with normally distributed data and no "real" difference between current and standard, for a sample of 300 split into 10 bands, the median PSI was found to be .027. It is anticipated that the bias is larger with "real" score distributions.

3) The PSI ignores rank ordering. Random shifts between adjacent bands will give just as big a PSI as clear upward or downward shifts in the population. This means that it is less powerful in detecting the shifts which are most important i.e. either an upward or downward shift in the overall score distribution.

4) PSI is sensitive to the number of bands. Too many bands greatly increases noise (and hence over-estimate PSI), too few and shifts within bands are missed (and hence under-estimate PSI). There is no known formula to determine the "right" number of bands.

Note (1) on the KS: if you band the data before calculating the KS, you will slightly underestimate KS. This also applies if you use the KS at the characteristic level and have discrete attributes.

Note (2) on the KS: in the original form the KS is an absolute value. By eliminating this constraint, the value indicates whether the population shift is positive (higher average scores in the current sample) or negative (worse profiles than expected).

In summary Kolmogorov-Smirnov statistic and associated p-level provides a standard statistical test for the hypothesis that two samples (standard and current) come from the same population. It is less affected by changes in the number of scorebands analysed than the PSI, it gives the direction of change, it detects changes in rank ordering and the p value measures significance of any change detected.