Update: EXPLANATION OF THE STATISTICAL MODEL for STANDARD 4

There is a proposal to change the number of performance elements used in Standard 4, resulting in the need to update the statistical model. Previously, in large jurisdictions (jurisdictions with 10 or more inspectors), the evaluation was based on direct oversight of two inspections per inspector, with respect to 10 performance elements. However, the proposal contains 20 performance elements instead of 10.

Using the previous statistical model and assumptions, a team achieving 88 percent at each inspection would pass the evaluation 75 percent of the time. Therefore, this 88 percent level of performance was used as a simple representation of a team that is good enough that we want them to have a good chance of passing, but not so good that they would not find it advantageous to improve. But now with 20 items instead of 10, a jurisdiction with 88 percent level of performance would pass only 59 percent of the time. This would fail too many high performing jurisdictions.

In order to rectify this, for large jurisdictions (jurisdictions with 10 or more inspectors), the evaluation must now be based on direct oversight of three inspections per inspector, with respect to 20 performance elements. With the additional inspections evaluated, the 88 percent performing jurisdiction will pass 75% of the time.

Evaluation of performance of small jurisdictions

A statistical issue was to determine a reasonable standard for those jurisdictions with less than 10 inspectors. When the sample gets this small, the relative error in the estimated fractions gets so large that the "each of 20 items rule" will fail good programs too frequently. Therefore, the 88 percent level of performance at each inspection was the feature of the standard that was kept constant in designing the sample sizes for the smaller jurisdictions

In jurisdictions with less than 10 inspectors, the statistical solution is to group all of the individual ratings, disregarding the individual items. For 5 inspectors we would review 5 x 3 = 15 inspections, with respect to all 20 items combined. This gives 300 observations. It is not possible to make a total observation test mimic exactly a 20 item test, but the minimum passing rates will be about as stringent as the 75 percent for the 20 item test:

For 4 to 9 inspectors, conduct three joint inspections for each inspector. Chart 4-1 shows the lowest total passing score out of the complete set of combined items that would give at least a 75 percent chance of passing for a team with an 88 percent chance of getting any particular observation correct. For a team of three or less, it is recommended that extra oversight inspections be performed to produce a total of 12 inspections. This is an intuitive judgment call that any set smaller than 12 could randomly turn out to be odd enough to produce an unfair rating.

Standard 4: Uniform Inspection Program
Self-Assessment Worksheet

Chart 4-1: Method of Calculation for Jurisdictions with Less Than Ten Inspectors

# of inspectors	# inspections needed	# of items needed to be marked IN compliance in order to meet Standard 4 criteria
<4	12 minimum	200 (out of 240 possible Items)
4-9	3 per inspector	4 inspectors = 200 (out of 240 possible Items) 5 inspectors = 252 (out of 300 possible Items) 6 inspectors = 303 (out of 360 possible Items) 7 inspectors = 355 (out of 420 possible Items) 8 inspectors = 407 (out of 480 possible Items) 9 inspectors = 459 (out of 540 possible Items)

NOTE:

1. These minimum inspection program assessment criteria are comparable to the 75% IN Compliance rate for each of the ten inspection program areas for jurisdictions with 10 or more inspectors.

Example:

For 6 inspectors, there will be 3 field visits per inspector = 18 visits
18 visits X 20 Items per visit= 360 Total Possible Items