Details of the Survey Simulator web application

https://iecodesign.shinyapps.io/survey_simulator

Kevin Little, Ph.D. 17 October 2015

Purpose

I created this app to help participants in an Institute for Healthcare Improvement (www.ihi.org) program develop their intuition about variation in patient experience survey results.

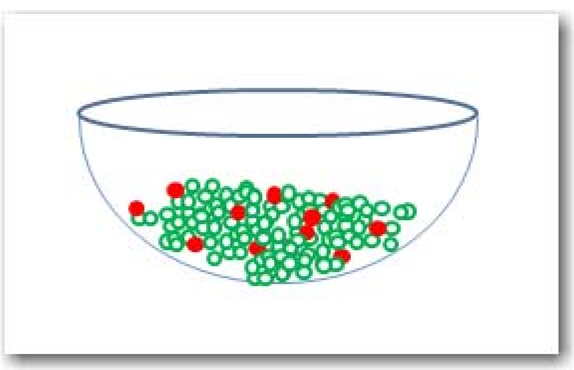
Background

In the U.S., hospitals are required to participate in a formal patient survey, the Hospital Consumer Assessment of Healthcare Providers and Systems survey, HCAHPS.   There are other CAHPS surveys developed and required for primary care, nursing facilities, and hospice.

Healthcare organizations study reports of patient experience survey results. Analysts and managers need basic knowledge of sampling variation to minimize two errors: (1) seeing signals of change where there is no strong evidence for change and (2) missing signals of change when in fact there is evidence for change.   The primary tool to minimize the costs of these two errors is an appropriate control chart.

To help motivate people to learn about and use control charts, we can start to build intuition about variation in measures that they typically review.

Computing Details-1



The web app shows the sampling variation in proportion of successes from a population of items.

Imagine a bowl with some red beads (shaded) and green beads (unshaded).   If we could count all the beads and correctly categorize each bead as red or green, we have no sampling variation.

In many situations, you can't count all the beads; you take a sample and seek to estimate the proportion of green beads in the bowl based on the sample.   If you can sample according to a random procedure you can characterize the sampling variation in observed proportion of successes.

To survey patients about their experience, we requires a sampling frame—a list that uniquely identifies each patient and enables a survey vendor or the organization to apply a random selection procedure to obtain a sample. The frame provides the operational definition of the population to be sampled.

The app uses two sampling models. It simulates values to show sampling variation with the assumption of simple random sampling.

(1) A binomial sampling model, named the **simple** model. The simulation requires only two values:
    a. the reference per cent of success in the population (called p) and
    b. the number of items sampled (called n).
(2) A hypergeometric sampling model, named the **adjusted** model. The simulation requires three values:
    a. the proportion of success in the population,
    b. the number of items sampled, and
    c. the sampling fraction—the size of the sample as a fraction of the total population.

As the sampling fraction gets smaller and smaller, samples from the adjusted model look more and more like samples from the simple model.

If you want to see formulas for the two models, here's an article that is not overly technical: J. Wroughton and T. Cole (2013), "Distinguishing Between Binomial, Hypergeometric and Negative Binomial Distributions", *Journal of Statistics Education*, **21**, 1, www.amstat.org/publications/jse/v21n1/wroughton.pdf accessed 15 October 2015.

Computing notes
I used R version 3.2.2 (R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org/)

Code is available on GitHub, https://github.com/klittle314/survey_simulator.git

Statistical Details-2

Sampling method
HCAPHS sampling allows three methods of sampling: simple random sampling, proportionate stratified random sampling, and disproportionate stratified random sampling. The simulation applies directly to either simple random sampling or to sample results within a stratum.

(See *HCAHPS Quality Assurance Guidelines*, Version 10.0, March 2015, Chapter IV "Sampling Protocol". http://www.hcahpsonline.org/Files/QAG_V10_0_2015.pdf).

Composite Survey Measures

The public reporting of HCAHPS results uses composite measures. For example the overall nurse communication"top box" result is the arithmetic average of the fraction of "always" answers to the three HCAHPS questions about nursing communication (questions 1-3 in the 2015 survey). The pain management composite "top box" result is the arithmetic average of the fraction of always answers to the two HCAHPS questions about pain (questions 13 and 14 in the 2015 survey).

The simulations do not apply directly to composite results. However, the simulations will approximate the sampling variation of the composite more and more closely as the agreement between or among the composite components increases.

For example, consider the the pain management composite.

If X is a random variable that describes the per cent of "top box" answers to HCAHPS question 13 and Y is a random variable that describes the per cent of "top box" answers to HCAHPS question 14, we need to characterize $Z = (X + Y)/2$, assuming the same number of responses for each question. (If the number of answers differ, then the variables are weighted by the number of responses divided by the total number of responses for both variables).

To illustrate, consider the means, variances and correlation for the two variables.

Let μ_X stand for the expected value of the random variable X, σ_X^2 stand for the variance of X and ρ_{XY} stand for the covariance of X and Y.

We expect X and Y to be highly positively correlated. Look at the expected value and variance of Z:

(1) $\mu_Z = (\mu_X + \mu_Y)/2$

(2) $\sigma_Z^2 = \sigma_X^2/4 + \sigma_Y^2/4 + 2\,\rho_{XY}\,(\sigma_X/2)(\sigma_Y/2)$ using the formula for variance of two random variables.

If X and Y are very similar, then $\mu_X \approx \mu_Y \approx \mu$; also $\sigma_X \approx \sigma_Y \approx \sigma$ and $\rho_{XY} \approx 1$ and formulas (1) and (2) match what we should expect to see: Z has approximately the same mean and the same variance as either X or Y so we can use the sampling distribution of X or Y to approximate the sampling distribution of Z.

Adjusted Per cents

For some surveys, the reported per cents are adjusted for patient-mix. This is true of the publicly reported HCAHPS composite results, which also adjusts results for survey administration; for example, phone surveys yield higher results than mail surveys. See e.g. http://www.hcahpsonline.org/modeadjustment.aspx accessed 17 October 2015.

As the adjustments are typically derived from properties of the patients sampled and characteristics of the organization and so the adjustments should be considered a random variable, say A.

The variance of an adjusted score Z = X + A is:

(3) $\sigma_Z^2 = \sigma_X^2 + \sigma_A^2 + 2\,\rho_{XA}\,\sigma_X\sigma_A$

The adjustments are typically small, 10% or less of the unadjusted score, which suggests σ_A^2 will be a small fraction of σ_X^2. ρ_{XA} may be positive or negative, depending on the nature of the adjustments.

Simple Model may be good enough

Interpreting CMS documentation (*Mode and Patient-mix Adjustment of the CAHPS® Hospital Survey (HCAHPS) April 30, 2008*, http://www.hcahpsonline.org/files/Final%20Draft%20Description%20of%20HCAHPS%20Mode%20and%20PMA%20with%20bottom%20box%20modedoc%20April%2030,%202008.pdf) pp. 12-13), suggests that the simple model (binomial sampling) is used to support the power and significance calculations for sample sizes of at least 300—evidently the survey developers are prepared to ignore relatively small effects that might arise from computing composite scores and adjusting for patient-mix and survey mode.

> "At least 300 completed surveys over four quarters are necessary to ensure adequate statistical power to compare hospitals to one another and to national benchmarks. At least 300 completed surveys are required to ensure an 80% chance that two hospitals that truly differ by 12% are reported as statistically different, or that a hospital that is truly 8% above the national benchmark is reported as statistically significantly above average (both using 5% thresholds of significance and two-sided tests). Observed differences of 6% between two hospitals and observed differences of 9% from a benchmark will be significant with at least 300 completed surveys."

The maximum sample size to compare two proportions will occur when we try to detect a difference of 12% symmetric to 50%, the value for which binomial variation is a maximum.

Using the R function `power.prop.test` with $p_1 = 0.44$, $p_2=0.56$, n=300, a significance level of 0.05, and a two-sided test structure gives a power calculation of 0.838.