Basic Engineering Data Collection and Analysis

Stephen B. Vardeman
Iowa State University

J. Marcus Jobe
Miami University

DUXBURY
THOMSON LEARNING

Australia • Canada • Mexico • Singapore • Spain • United Kingdom • United States

Example 23 *(continued)*

consider what means and standard deviations are associated with the probability distributions of the sample average, $\overline{S}$, of first the next 4 and then the next 100 excess service times.

$S_1, S_2, \ldots, S_{100}$ are, to the extent that the service process is physically stable, reasonably modeled as independent, identically distributed, exponential random variables with mean $\alpha = 16.5$. The exponential distribution with mean $\alpha = 16.5$ has variance equal to $\alpha^2 = (16.5)^2$. So, using formulas (5.55) and (5.56), for the first 4 additional service times,

$$E\overline{S} = \alpha = 16.5 \text{ sec}$$

$$\sqrt{\operatorname{Var}\overline{S}} = \sqrt{\frac{\alpha^2}{4}} = 8.25 \text{ sec}$$

Then, for the first 100 additional service times,

$$E\overline{S} = \alpha = 16.5 \text{ sec}$$

$$\sqrt{\operatorname{Var}\overline{S}} = \sqrt{\frac{\alpha^2}{100}} = 1.65 \text{ sec}$$

Notice that going from a sample size of 4 to a sample size of 100 decreases the standard deviation of $\overline{S}$ by a factor of 5 $(= \sqrt{\frac{100}{4}})$.

Relationships (5.55) and (5.56), which perfectly describe the random behavior of $\overline{X}$ under random sampling with replacement, are also approximate descriptions of the behavior of $\overline{X}$ under simple random sampling in enumerative contexts. (Recall Example 18 and the discussion about the approximate independence of observations resulting from simple random sampling of large populations.)

5.5.4 The Propagation of Error Formulas

Proposition 1 gives exact values for the mean and variance of $U = g(X, Y, \ldots, Z)$ only when g is linear. It doesn't seem to say anything about situations involving nonlinear functions like the one specified by the right-hand side of expression (5.52) in the solar collector example. But it is often possible to obtain useful approximations to the mean and variance of U by applying Proposition 1 to a first-order multivariate Taylor expansion of a not-too-nonlinear g. That is, if g is reasonably well-behaved, then for $x, y, \ldots, z$ (respectively) close to $EX, EY, \ldots, EZ$,

$$\left.\begin{aligned} g(x, y, \ldots, z) \approx g(EX, EY, \ldots, EZ) + \frac{\partial g}{\partial x} \cdot (x - EX) + \frac{\partial g}{\partial y} \cdot (y - EY) \\ + \cdots + \frac{\partial g}{\partial z} cdot (z - EZ) \end{aligned}\right\} \quad (5.57)$$

where the partial derivatives are evaluated at $(x, y, \ldots, z) = (EX, EY, \ldots, EZ)$. Now the right side of approximation (5.57) is linear in $x, y, \ldots, z$. Thus, if the variances of $X, Y, \ldots, Z$ are small enough so that with high probability, $X, Y, \ldots, Z$ are such that approximation (5.57) is effective, one might think of plugging $X, Y, \ldots, Z$ into expression (5.57) and applying Proposition 1, thus winding up with approximations for the mean and variance of $U = g(X, Y, \ldots, Z)$.

Proposition 2 *(The Propagation of Error Formulas)*

If $X, Y, \ldots, Z$ are independent random variables and g is well-behaved, for small enough variances Var X, Var $Y, \ldots,$ Var Z, the random variable $U = g(X, Y, \ldots, Z)$ has approximate mean

$$EU \approx g(EX, EY, \ldots, EZ) \tag{5.58}$$

and approximate variance

$$\text{Var}\, U \approx \left(\frac{\partial g}{\partial x}\right)^2 \text{Var}\, X + \left(\frac{\partial g}{\partial y}\right)^2 \text{Var}\, Y + \cdots + \left(\frac{\partial g}{\partial z}\right)^2 \text{Var}\, Z \tag{5.59}$$

Formulas (5.58) and (5.59) are often called the **propagation of error** or **transmission of variance** formulas. They describe how variability or error is propagated or transmitted through an exact mathematical function.

Comparison of Propositions 1 and 2 shows that when g is exactly linear, expressions (5.58) and (5.59) reduce to expressions (5.53) and (5.54), respectively. (a_1 through a_n are the partial derivatives of g in the case where $g(x, y, \ldots, z) = a_0 + a_1 x + a_2 y + \cdots + a_n z$.) Proposition 2 is purposely vague about when the approximations (5.58) and (5.59) will be adequate for engineering purposes. Mathematically inclined readers will not have much trouble constructing examples where the approximations are quite poor. But often in engineering applications, expressions (5.58) and (5.59) are at least of the right order of magnitude and certainly better than not having any usable approximations.

Example 24

A Simple Electrical Circuit and the Propagation of Error

Figure 5.35 is a schematic of an assembly of three resistors. If R_1, R_2, and R_3 are the respective resistances of the three resistors making up the assembly, standard theory says that

$$R = \text{the assembly resistance}$$