

MIT Open Access Articles

Calculating Expectation Shift

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Frey, D. D. "Efficiently estimating mean shift due to variability." Proceedings of the 2011 IEEE International Conference on Quality and Reliability (ICQR), 2011. 95–100.

As Published: <http://dx.doi.org/10.1109/ICQR.2011.6031688>

Publisher: Institute of Electrical and Electronics Engineers (IEEE)

Persistent URL: <http://hdl.handle.net/1721.1/78662>

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Terms of use: Creative Commons Attribution-Noncommercial-Share Alike 3.0



Calculating Expectation Shift

Daniel D. Frey,

Associate Professor of Mechanical Engineering and Engineering Systems

Abstract

This paper concerns the problem of calculating expectation shift due to variability which tends to occur whenever the function of a random variable is non-linear and especially tends to occur in the neighborhood of a local maximum or minimum. The paper presents five theorems suggesting sampling points and formulae for estimating mean shift covering some of the most common cases of practical interest including multivariate normal distributions and uniform distributions as well as more general theorems covering all symmetric multivariate probability density functions.

1. Motivation

This paper concerns estimation of $E(y(\mathbf{x}))$ the mathematical expectation E of a function y of a vector of independent variables \mathbf{x} which are subject to variability or uncertainty. In engineering design, the function y can represent a performance criterion for an engineering system. The vector \mathbf{x} might represent a set of uncertain or variable parameters in the design with said uncertainty modeled by probability density functions $f_{\mathbf{x}}(\mathbf{x})$. The resulting uncertainty in $y(\mathbf{x})$ is modeled by probability density functions $f_y(y(\mathbf{x}))$.

Under many contemporary frameworks for design, accurate and efficient estimates of expected value under uncertainty are of central importance. In the case that $y(\mathbf{x})$ represents utility of the design, $E(y(\mathbf{x}))$ has special significance because a rational designer should seek to maximize expected value of utility [Hazelrigg, 1999]. In the case that y is a loss function, $E(y(\mathbf{x}))$ is important because the designer should minimize expected value of loss to society [Taguchi, 1989]. More generally, if $y(\mathbf{x})$ is a performance criterion of importance to the customer, the expected value $E(y(\mathbf{x}))$ will be of interest when it represents an economically significant indicator, for example, average fuel burned per mile given typical ranges of environmental and engine wear conditions.

It has long been understood that $E(y(\mathbf{x})) \neq y(E(\mathbf{x}))$. To simplify further discussion, let us define expectation shift due to variability, S , as

$$S = E(y(\mathbf{x})) - y(E(\mathbf{x})) \quad (1)$$

In the special case that y is an affine transformation, $y = \mathbf{A}\mathbf{x} + \mathbf{b}$, the expectation shift, S , is zero because the expected value of the function is the function of the expected value of the independent variables, $E(y) = \mathbf{A}E(\mathbf{x}) + \mathbf{b}$. However, for non-linear functions the expectation shift is generally not zero. In the special case that y is convex, Jensen's Inequality (1906) establishes that $E(y(\mathbf{x})) > y(E(\mathbf{x}))$ and so expectation shift S must be positive. In the context of engineering optimization, the expectation shift S is a cause for concern because it is almost always in an unfavorable direction. If the function $y(\mathbf{x})$ has been maximized and \mathbf{x} is set nominally at an interior optima of y , then the shift S generally reduces expectation below a local maximum (as shown in Figure 1). Similarly,

if the function $y(\mathbf{x})$ has been minimized and \mathbf{x} is set nominally at an interior optima of y , then the shift S brings a higher expected value of a property to be minimized like cost or weight (as shown in Figure 2).

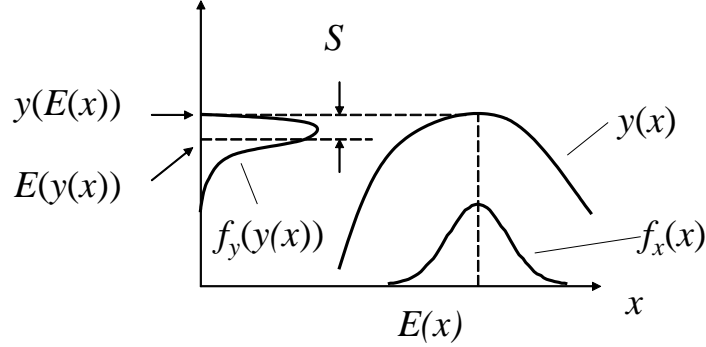


Figure 1. Expectation shift, S , reduces expected value of parameters to be maximized such as utility.

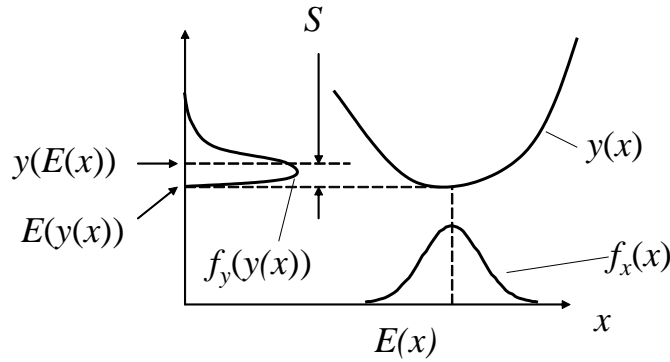


Figure 2. Expectation shift, S , raises expected value of parameters to be minimized such as loss functions.

2. Review of Prior Work

The topic of this paper, assessment of a function in the presence of variability and uncertainty, is most often associated with robust parameter design (Taguchi, 1987; Wu and Hamada, 2001). In engineering application of robust design, the variables that affect a product's response are classified into control factors whose nominal settings can be specified by the designer and noise factors that cause unwanted deviations of the system response. In robust design, the space of control factors is searched to seek settings that are superior despite the effects of the noise factors. Robust design is often accomplished using a strategy employing two nested loops (Diwekar, 2003). An inner sampling loop is used to estimate robustness and an outer optimization loop serves to improve the design. A difficulty of robust design is that its resource demands typically scale as the *product* of the demands of the inner and outer loops.

Robust parameter design is increasingly performed using computer models of the product or process rather than laboratory experiments. Computer models can provide significant advantages in cost and speed of the robustness optimization process. Although random variations in noise factors can be simulated in a computer, the results of computer models generally lack pure experimental error. Therefore it is generally acknowledged that the techniques for robust design via computer models ought to be different than the techniques for robust design via laboratory experiments. In response to this fact, the field of Design and Analysis of Computer Experiments (DACE) has grown rapidly in recent decades providing a variety of useful techniques. Background on this subject is provided from various perspectives by Santner *et al.* (2003), by Giunta, *et al.*, 2003, and by Simpson *et al.* 2001. Three of the most promising techniques developed so far are reviewed in the next four paragraphs.

Latin Hypercube Sampling (LHS) is a stratified sampling technique which can be viewed as an extension of Latin square sampling and a close relative of highly fractionalized factorial designs (McKay *et al.*, 1979). LHS ensures that each input variable has all portions of its distributions uniformly represented by input values. This property (often referred to as a “projective” property) is especially advantageous when only a few of the many inputs turn out to contribute significantly to the output variance. McKay *et al.* (1979) proved that for large samples, LHS provides smaller variance in the estimators than simple random sampling as long as the response function is monotonic in all the noise factors. However, there is also evidence that LHS often provides little practical advantage over simple random sampling if the response function is highly non-linear (Giunta *et al.*, 2003).

An innovation called Hammersley Sequence Sampling (HSS) may provide significant advantages over LHS (Kalagnanam and Diwekar, 1997). HSS employs a quasi-Monte Carlo sequence with low discrepancy and good space filling properties. In applications to simple functions, it has been demonstrated that HSS converges to 1% accuracy faster than LHS. In an engineering application with six noise factors, HSS converged to 1% accuracy up to 40 times faster than LHS.

Quadrature and cubature are techniques for exactly integrating specified classes of functions by means of a small number of highly structured samples. A remarkable number of different methods have been proposed for different regions and weighting functions. See Cools and Rabinowitz (1993) and Cools (1999) for reviews. More recently, a new cubature method for Gaussian weighted integration was developed that scales better than other, previously published cubature schemes (Lu and Darmodai, 2000). If the number of uncertain factors in the inner loop is d , then the method requires d^2+3d+3 samples. The new rule provides exact solutions for the mean for polynomials of fifth degree including the effects of all multi-factor interactions. Used recursively to estimate transmitted variance, it provides exact solutions for polynomials of second degree including all two-factor interactions. On a challenging engineering application, the cubature method had less than 0.1% error in estimating mean and about 1% to 3% error in estimating variance. Quadratic scaling with dimensionality limits the method to applications with a relatively small number of uncertain factors.

Available sampling schemes still require too many function evaluations for many practical engineering applications. Computer models of adequate fidelity can require hours or days to execute. In some cases, changing input parameters to the computer

models is not an automatic process – it can often require careful attention to ensure that the assumptions and methods of the analysis remain reasonable. It is also frequently the case that the results of computer models must be reviewed by experts or that some human intervention is required to achieve convergence and accuracy. All of these considerations create a demand for techniques requiring a very small number of samples.

3. Expectation Shift Formulae

Mathematical expectation is defined by an integral which implies that computing expectation shift involves computing an integral

$$S = \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} y(\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}_1 d\mathbf{x}_2 \dots d\mathbf{x}_n \right] - y(E(\mathbf{x})) \quad (2)$$

where $f_{\mathbf{x}}(\mathbf{x})$ is the joint probability density function over the n dimensional random vector \mathbf{x} . Without loss of generality, one may perform a transformation of variables replacing \mathbf{x} by $\mathbf{z} = \mathbf{x} - E(\mathbf{x})$ and the expectation shift S as defined in Equation 1 becomes

$$S = \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} y(\mathbf{z} + E(\mathbf{x})) f_{\mathbf{z}}(\mathbf{z}) d\mathbf{z}_1 d\mathbf{z}_2 \dots d\mathbf{z}_n \right] - y(E(\mathbf{x})) \quad (3)$$

where $f_{\mathbf{z}}(\mathbf{z})$ is the joint probability density function over the random vector \mathbf{z} . This transformation of variables can, in some cases, greatly simplify evaluation of the integral in Equation 3 because many probability distributions have symmetries when transformed to be centered on their expected value. The following theorem is an example of such a simplification.

Theorem 1

If the joint probability density function $f_{\mathbf{z}}(\mathbf{z})$ is symmetric about every plane $\mathbf{z}_i=0$ and the response surface $y(\mathbf{x})$ is approximated by a fourth order polynomial $\eta(\mathbf{z}) \approx y(\mathbf{z} + E(\mathbf{x}))$ where

$$\eta(\mathbf{z}) = \beta_0 + \sum_{i=1}^n \beta_i \mathbf{z}_i + \sum_{j=1}^n \sum_{\substack{i=1 \\ i \leq j}}^n \beta_{ij} \mathbf{z}_i \mathbf{z}_j + \sum_{j=1}^n \sum_{\substack{i=1 \\ i \leq j}}^n \sum_{\substack{k=1 \\ k \leq j}}^n \beta_{ijk} \mathbf{z}_i \mathbf{z}_j \mathbf{z}_k + \sum_{j=1}^n \sum_{\substack{i=1 \\ i \leq j}}^n \sum_{\substack{k=1 \\ k \leq j}}^n \sum_{\substack{l=1 \\ l \leq k}}^n \beta_{ijkl} \mathbf{z}_i \mathbf{z}_j \mathbf{z}_k \mathbf{z}_l \quad (4)$$

then the expectation shift S which is defined as $E(y(\mathbf{x})) - y(E(\mathbf{x}))$ is

$$S = \sum_{i=1}^n \int_{-\infty}^{\infty} \beta_{ii} \mathbf{z}_i^2 f_{\mathbf{z}}(\mathbf{z}) d\mathbf{z}_i + \sum_{i=1}^n \sum_{\substack{j=1 \\ j \leq i}}^n \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \beta_{ijij} \mathbf{z}_i^2 \mathbf{z}_j^2 f_{\mathbf{z}}(\mathbf{z}) d\mathbf{z}_i d\mathbf{z}_j + E_T. \quad (5)$$

where E_T is the truncation error due to the polynomial approximation.

Proof --

Substitution of the polynomial approximation (Equation 4) into Equation 3 leads to

$$S = \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \eta(\mathbf{z}) f_{\mathbf{z}}(\mathbf{z}) d\mathbf{z}_1 d\mathbf{z}_2 \dots d\mathbf{z}_n \right] - \beta_0 \quad (6)$$

Any integral of a function of a variable \mathbf{z}_i which is anti-symmetric the plane $\mathbf{z}_i=0$ evaluates to zero over the limits shown in Equation 6. The product of a symmetric function and an anti-symmetric function is an anti-symmetric function. The probability density function $f_{\mathbf{z}}(\mathbf{z})$ was assumed to be symmetric about every plane $\mathbf{z}_i=0$. Therefore every term of the polynomial that is anti-symmetric about $\mathbf{z}_i=0$ can be dropped. This includes any term that contains any odd power of \mathbf{z}_i . Dropping those terms and canceling β_0 with the second term in Equation 6 leads to equation 5.●

Theorem 1 represents a substantial reduction in the total number of polynomial coefficients which must be estimated. The sixth order polynomial of Equation 4 contains $1 + 4n + \binom{n}{2} + \binom{n}{3} + \binom{n}{4}$ terms. The formula for expectation shift (Equation 5) contains only $2n + \binom{n}{2}$ terms. As will be discussed in Section 5, the dropping of polynomial terms has a significant impact in determining efficient computational methods for estimating expectation shift. Theorem 1 has an impact in any design scenario in which the probability density functions can be approximated as symmetric functions such as the multivariate normal distribution and the uniform distribution (as shown in the next in corollaries 1 and 2 below).

Corollary 1 – Independent normal distribution

If the response surface $y(\mathbf{x})$ is approximated by a polynomial $\eta(\mathbf{z}) \approx y(\mathbf{z} + E(\mathbf{x}))$ (as in equation 4) and the joint probability density function $f_{\mathbf{z}}(\mathbf{z})$ is normal with covariance matrix \mathbf{K} and the elements of \mathbf{z} are probabilistically independent then the expectation shift S is

$$S = \sum_{i=1}^n \beta_{ii} \mathbf{K}_{ii} + 3 \sum_{i=1}^n \beta_{iii} \mathbf{K}_{ii}^2 + \sum_{i=1}^n \sum_{\substack{j=1 \\ i < j}}^n \beta_{ijj} \mathbf{K}_{ii} \mathbf{K}_{jj} + E_{\text{TN}} \quad (7)$$

where E_{TN} is the truncation error due to the polynomial approximation.

Proof -- The probability density function $f_{\mathbf{z}}(\mathbf{z})$ has a mean of zero by definition since $\mathbf{z} = \mathbf{x} - E(\mathbf{x})$. If \mathbf{z} is normally distributed then

$$f_{\mathbf{z}}(\mathbf{z}) = \frac{1}{\sqrt{2\pi}^n \sqrt{|\mathbf{K}|}} e^{-\frac{1}{2}\mathbf{z}^T \mathbf{K}^{-1} \mathbf{z}} \quad (8).$$

If the elements of \mathbf{z} are probabilistically independent then \mathbf{K} is diagonal and therefore Equation 8 is symmetric about planes $\mathbf{z}_i=0$, therefore Theorem 1 applies and the expectation shift is

$$\begin{aligned} S = & \sum_{i=1}^n \int_{-\infty}^{\infty} \beta_{ii} \mathbf{z}_i^2 \frac{1}{\sqrt{2\pi} \sqrt{\mathbf{K}_{ii}}} e^{-\frac{1}{2\mathbf{K}_{ii}} \mathbf{z}_i^2} d\mathbf{z}_i \\ & + \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \beta_{ij} \mathbf{z}_i^2 \mathbf{z}_j^2 \frac{1}{2\pi \sqrt{\mathbf{K}_{ii} \mathbf{K}_{jj}}} e^{-\frac{1}{2\mathbf{K}_{ii}} \mathbf{z}_i^2} e^{-\frac{1}{2\mathbf{K}_{jj}} \mathbf{z}_j^2} d\mathbf{z}_i d\mathbf{z}_j \cdot (9) \\ & + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} E_T \frac{1}{\sqrt{2\pi}^n \sqrt{|\mathbf{K}|}} e^{-\frac{1}{2}\mathbf{z}^T \mathbf{K}^{-1} \mathbf{z}} d\mathbf{z}_1 d\mathbf{z}_2 \dots d\mathbf{z}_n \end{aligned}$$

The first two terms can be evaluated symbolically. The last term for now is simply renamed E_{TN} .•

Note that Equation 7 can be applied to correlated normally distributed random variables by using a change of variables. If the probability density function $f_{\mathbf{z}}(\mathbf{z})$ is normal and \mathbf{K} is *not* diagonal, then a linear transformation $\mathbf{z}' = \mathbf{K}^{-1} \mathbf{z}$ can be performed so that the probability density function $f_{\mathbf{z}'}(\mathbf{z}')$ is normal and \mathbf{z}'_i are independent with all $\mathbf{K}_{ii}=1$. The coefficients of the polynomial must then be evaluated in the transformed space in order to use Equation 7.

Corollary 2 – Uniform Distributions

If the joint probability density function $f_{\mathbf{z}}(\mathbf{z})$ is uniformly distributed over a region where $-\Delta_i < \mathbf{z}_i < \Delta_i$ then the expectation shift S is

$$S = \frac{1}{\prod_{i=1}^n \Delta_i} \left[\frac{1}{3} \sum_{i=1}^n \beta_{ii} \Delta_i^3 + \frac{1}{5} \sum_{i=1}^n \beta_{iii} \Delta_i^5 + \frac{1}{3} \sum_{i=1}^n \sum_{\substack{j=1 \\ i < j}}^n \beta_{ij} \Delta_i^3 \Delta_j^2 + \frac{1}{3} \sum_{i=1}^n \sum_{\substack{j=1 \\ i < j}}^n \beta_{ij} \Delta_i^2 \Delta_j^3 \right] + E_{TU} \cdot \quad (10)$$

where E_{TU} is the truncation error due to the polynomial approximation.

Proof -- If the joint probability density function $f_{\mathbf{z}}(\mathbf{z})$ is uniformly distributed over a region where $-\Delta_i < \mathbf{z}_i < \Delta_i$ then the probability density is $\frac{1}{\prod_{i=1}^n 2\Delta_i}$ everywhere in the support

set. The probability density function $f_{\mathbf{z}}(\mathbf{z})$ is symmetric about planes $\mathbf{z}_i=0$, therefore Theorem 1 applies and the expectation shift is

$$S = \int \sum_{i=1}^n \beta_{ii} \mathbf{z}_i^2 \frac{1}{\prod_{i=1}^n 2\Delta_i} d\mathbf{z} + \int \sum_{i=1}^n \sum_{\substack{j=1 \\ j \leq i}}^n \beta_{ijij} \mathbf{z}_i^2 \mathbf{z}_j^2 \frac{1}{\prod_{i=1}^n 2\Delta_i} + \int E_T \frac{1}{\prod_{i=1}^n 2\Delta_i}. \quad (11)$$

The first two terms can be evaluated symbolically. The last term for now is simply renamed E_{TU} .•

4. Computation Methods and Modern DOE

In order to use the formulae in section 3, one must form a polynomial response surface approximation to the function of interest $f_{\mathbf{x}}(\mathbf{x})$. There are many ways to form such an approximation most of which require sampling the function $f_{\mathbf{x}}(\mathbf{x})$ at a set of points and fitting the polynomial to those points either exactly or approximately. The arrangement of the points is of great practical interest because it can have a significant on accuracy and computational expense. In this section, four different methods are proposed.

In Design of Experiments, one often invokes the hierarchy principle. This principle leads one to focus on main effects (first order polynomial terms) before interactions and higher order polynomial terms. Theorem 1 shows that linear sensitivities and simple two-factor interactions have no impact on expectation shift if the probability distribution is symmetric. Therefore different experimental designs may be needed. Approaches appropriate to the formulae in section 3 must be designed to provide estimates of simple quadratic terms, simple 4th order terms, and sometimes two-factor interactions between quadratic terms.

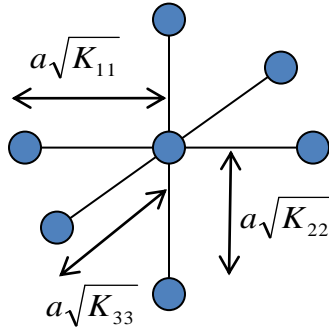


Figure 3. A “star pattern” for sampling $2n+1$ points in an n -dimensional space (in this case $n=3$).

One efficient design for estimating pure quadratic terms is a center point and $2n$ axial points (a.k.a. star points) as shown in Figure 3. This design requires $1+n$ function evaluations. Let us assume that the distance from the center point to the axial points is proportional to the standard deviation of the variable \mathbf{z}_i so that the distance is $a\sqrt{K_{ii}}$. In this case, the polynomial coefficients are given by

$$\beta_{ii} = \frac{1}{a^2 \mathbf{K}_{ii}} \left[y(E(\mathbf{x}) - a\sqrt{\mathbf{K}_{ii}}) - 2y(E(\mathbf{x})) + y(E(\mathbf{x}) + a\sqrt{\mathbf{K}_{ii}}) \right] \quad (12)$$

It is possible to select the parameter a so that the accuracy of the method is optimized. One reasonable criterion for the choice of a is to select a value that broadens the class of functions for which the method provides exact solutions. This leads to the following theorem.

Theorem 2 – Location of $2n+1$ axial points for independent normal distribution

If the response surface $y(\mathbf{x})$ is exactly a polynomial $\eta(\mathbf{z})=y(\mathbf{z}+E(\mathbf{x}))$ as in Equation 4 and $\beta_{ijj} = 0$ where $i \neq j$ and the joint probability density function $f_{\mathbf{z}}(\mathbf{z})$ is normal with covariance matrix \mathbf{K} and the elements of \mathbf{z} are probabilistically independent then there exists a unique value a for sampling along each axis in both directions around the expected value such that expectation shift S is exactly

$$S = \sum_{i=1}^n [\beta_{ii} + 3\beta_{iii} \mathbf{K}_{ii}] \mathbf{K}_{ii} \quad (13)$$

where

$$\beta_{ii} + 3\beta_{iii} \mathbf{K}_{ii} = \frac{1}{2a^2 \mathbf{K}_{ii}} \left[y \left(E(\mathbf{x}) - \begin{Bmatrix} 0 \\ \vdots \\ a\sqrt{\mathbf{K}_{ii}} \\ \vdots \\ 0 \end{Bmatrix} \right) - 2y(E(\mathbf{x})) + y \left(E(\mathbf{x}) + \begin{Bmatrix} 0 \\ \vdots \\ a\sqrt{\mathbf{K}_{ii}} \\ \vdots \\ 0 \end{Bmatrix} \right) \right] \quad (14)$$

and the unique value of a is

$$a = \sqrt{3} \quad (15)$$

Proof – Under the conditions described in Theorem 2, the result of Corollary 1 hold and the truncation error E_{TN} is zero since the function $y(\mathbf{x})$ is exactly a fourth order polynomial therefore

$$S = \sum_{i=1}^n \beta_{ii} \mathbf{K}_{ii} + 3 \sum_{i=1}^n \beta_{iii} \mathbf{K}_{ii}^2 + \sum_{i=1}^n \sum_{\substack{j=1 \\ i < j}}^n \beta_{ijj} \mathbf{K}_{ii} \mathbf{K}_{jj} . \quad (16)$$

In addition it was stipulated that $\beta_{ijj} = 0$ where $i \neq j$ so the exact solution is

$$S = \sum_{i=1}^n \beta_{ii} \mathbf{K}_{ii} + 3 \sum_{i=1}^n \beta_{iii} \mathbf{K}_{ii}^2 . \quad (17)$$

Simplifying Equation 17 gives Equation 13. If the function $y(\mathbf{x})$ is sampled at star points as stipulated in Theorem 2, then the samples of the function will be

$$y(E(\mathbf{x}) + a\sqrt{\mathbf{K}_{ii}}) = \beta_0 + \beta_i a\sqrt{\mathbf{K}_{ii}} + \beta_{ii} a^2 \mathbf{K}_{ii} + \beta_{iii} a^3 \mathbf{K}_{ii}^{3/2} + \beta_{iiii} a^4 \mathbf{K}_{ii}^2 \quad (18)$$

$$y(E(\mathbf{x}) - a\sqrt{\mathbf{K}_{ii}}) = \beta_0 - \beta_i a\sqrt{\mathbf{K}_{ii}} + \beta_{ii} a^2 \mathbf{K}_{ii} - \beta_{iii} a^3 \mathbf{K}_{ii}^{3/2} + \beta_{iiii} a^4 \mathbf{K}_{ii}^2 \quad (19)$$

$$y(E(\mathbf{x})) = \beta_0 \quad (20)$$

Adding Equation 18 and 19, substituting for β_0 using Equation 20 and rearranging we find

$$\beta_{ii} + a^2 \beta_{iiii} \mathbf{K}_{ii} = \frac{1}{2a^2 \mathbf{K}_{ii}} [y(E(\mathbf{x}) - a\sqrt{\mathbf{K}_{ii}}) - 2y(E(\mathbf{x})) + y(E(\mathbf{x}) + a\sqrt{\mathbf{K}_{ii}})] \quad (21)$$

The result in Equation 21 provides the needed value to compute the expectation shift S exactly via Equation 13 when $a = \sqrt{3}$. •

The result of Theorem 2 is closely related to Gaussian weighted quadrature formulae in single variables. The location of the sample points and the weightings are identical to the formula tabulated in Stroud and Seacrest (1966) for $\int_{-\infty}^{\infty} e^{-x^2} f(\mathbf{x}) d\mathbf{x}$ if the locations and weightings are adjusted to transform the weight function to a normal distribution. The key difference is that Theorem 2 applies to multidimensional integration with the modest constraint on the class of functions – that $\beta_{ijj} = 0$ where $i \neq j$. Therefore, Theorem 2 shows how to apply single dimensional integration rule to multivariate cases in order to make benefit of the more favorable scaling law. Theorem 2 requires $2n+1$ function evaluations whereas the most efficient multidimensional quadrature rule requires n^2+3n+3 function evaluations (Lu and Darmofal, 2005).

An extension of the previous result can be made to relax the restriction to polynomials wherein that $\beta_{ijj} = 0$ where $i \neq j$. Because the terms in question are all functions of \mathbf{z}_i^2 , the effect that needs to be estimated can come from a sample in any half space in R_i for all i . This leads to the following theorem.

Theorem 3 – Location of $2n+2$ axial points for independent normal distribution

If the response surface $y(\mathbf{x})$ is exactly a polynomial $\eta(\mathbf{z})=y(\mathbf{z}+E(\mathbf{x}))$ as in Equation 4 and the joint probability density function $f_{\mathbf{z}}(\mathbf{z})$ is normal with covariance matrix \mathbf{K} and the elements of \mathbf{z} are probabilistically independent then there exists a unique value a for sampling along each axis in both directions around the expected value such that expectation shift S is exactly

$$S = \sum_{i=1}^n [\beta_{ii} + 3\beta_{iiii} \mathbf{K}_{ii}] \mathbf{K}_{ii} + \sum_{i=1}^n \sum_{\substack{j=1 \\ i < j}}^n \beta_{ijj} \mathbf{K}_{ii} \mathbf{K}_{jj} \quad (22)$$

where $\beta_{ii} + 3\beta_{iii}\mathbf{K}_{ii}$ is given by Equation 14 with $a = \sqrt{3}$ and

$$\sum_{i=1}^n \sum_{\substack{j=1 \\ i < j}}^n \beta_{ijj} \mathbf{K}_{ii} \mathbf{K}_{jj} = \frac{1}{2a^2} \left[y \left(E(\mathbf{x}) + \begin{Bmatrix} \pm a\sqrt{\mathbf{K}_{11}} \\ \vdots \\ \pm a\sqrt{\mathbf{K}_{ii}} \\ \vdots \\ \pm a\sqrt{\mathbf{K}_{nn}} \end{Bmatrix} \right) \right] - \sum_{i=1}^n (\beta_{ii} + 3\beta_{iii}\mathbf{K}_{ii}) \quad (23)$$

Proof – From Theorem 1 Corollary 1, relaxing the stipulation that $\beta_{ijj} = 0$ where $i \neq j$, so the exact solution is

$$S = \sum_{i=1}^n \beta_{ii} \mathbf{K}_{ii} + 3 \sum_{i=1}^n \beta_{iii} \mathbf{K}_{ii}^2 + \sum_{i=1}^n \sum_{\substack{j=1 \\ i < j}}^n \beta_{ijj} \mathbf{K}_{ii} \mathbf{K}_{jj} + E_{\text{TN}}. \quad (24)$$

If the function $y(\mathbf{x})$ is sampled at star points as stipulated in Theorem 2 and also in one additional point as given in Theorem 3, then the samples of the function will be as given in Equations 18, 19, and 20, and also

$$y \left(E(\mathbf{x}) + \begin{Bmatrix} \pm a\sqrt{\mathbf{K}_{11}} \\ \vdots \\ \pm a\sqrt{\mathbf{K}_{ii}} \\ \vdots \\ \pm a\sqrt{\mathbf{K}_{nn}} \end{Bmatrix} \right) = \beta_0 + \sum_{i=1}^n \left[\beta_i a \sqrt{\mathbf{K}_{ii}} + \beta_{ii} a^2 \mathbf{K}_{ii} + \beta_{iii} a^3 \mathbf{K}_{iii}^{3/2} + \beta_{iiii} a^4 \mathbf{K}_{iiii}^2 \right] + \sum_{i=1}^n \sum_{\substack{j=1 \\ i < j}}^n \beta_{ijj} \mathbf{K}_{ii} \mathbf{K}_{jj} \quad (25)$$

$$y(E(\mathbf{x}) - a\sqrt{\mathbf{K}_{ii}}) = \beta_0 - \beta_i a \sqrt{\mathbf{K}_{ii}} + \beta_{ii} a^2 \mathbf{K}_{ii} - \beta_{iii} a^3 \mathbf{K}_{ii}^{3/2} + \beta_{iiii} a^4 \mathbf{K}_{ii}^2 \quad (19)$$

$$y(E(\mathbf{x})) = \beta_0 \quad (20)$$

Another option is to select a design that can resolve pure quadratic coefficients and also pure 4th order coefficients. One design is a center point and $4n$ equally spaced axial points (a.k.a. star points). This design requires $1+4n$ function evaluations. Let us assume that the distance from the center point to the axial points is proportional to the standard deviation of the variable \mathbf{z}_i so that the distance from the center point is $a\sqrt{K_{ii}}$ and $2a\sqrt{K_{ii}}$. . In this case, the polynomial coefficients are given by

$$\beta_{ii} = \frac{1}{a^2 K_{ii}} [y(E(\mathbf{x}) - a\sqrt{K_{ii}}) - 2y(E(\mathbf{x})) + y(E(\mathbf{x}) + a\sqrt{K_{ii}})] \quad (19)$$

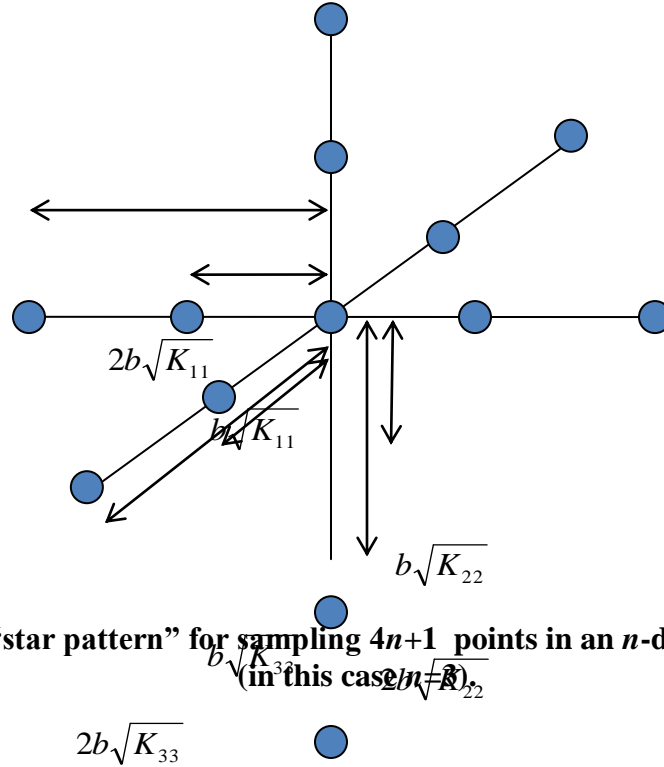


Figure 4. A “star pattern” for sampling $4n+1$ points in an n -dimensional space (in this case $n=4$).

It is possible to select the parameter a so that the accuracy of the method is optimized. One reasonable criterion for the choice of a is to select a value that broadens the class of functions for which the method provides exact solutions. This leads to the following theorems.

Theorem 3 – Location of $4n$ axial points for independent normal distribution

If the response surface $y(\mathbf{x})$ is exactly a polynomial $\eta(\mathbf{z})=y(\mathbf{z}+E(\mathbf{x}))$

$$\eta(\mathbf{z}) = \beta_0 + \sum_{i=1}^n \beta_i \mathbf{z}_i + \sum_{j=1}^n \sum_{\substack{i=1 \\ i \leq j}}^n \beta_{ij} \mathbf{z}_i \mathbf{z}_j + \sum_{j=1}^n \sum_{\substack{i=1 \\ i \leq j}}^n \sum_{\substack{k=1 \\ k \leq j}}^n \beta_{ijk} \mathbf{z}_i \mathbf{z}_j \mathbf{z}_k + \sum_{i=1}^n \beta_{iiii} \mathbf{z}_i^4 + \sum_{i=1}^n \beta_{iiiii} \mathbf{z}_i^5$$

and the joint probability density function $f_{\mathbf{z}}(\mathbf{z})$ is normal with covariance matrix \mathbf{K} and the elements of \mathbf{z} are probabilistically independent, then there exists a unique value a such that expectation shift S is exactly

$$S = \sum_{i=1}^n \beta_{ii} \mathbf{K}_{ii} + 3 \sum_{i=1}^n \beta_{iiii} \mathbf{K}_{ii}^2 \quad (20)$$

where

$$\beta_{ii} = \frac{1}{24a^2 \mathbf{K}_{ii}} \left[-y(E(\mathbf{x}) - 2a\sqrt{\mathbf{K}_{ii}}) + 16y(E(\mathbf{x}) - a\sqrt{\mathbf{K}_{ii}}) - 2y(E(\mathbf{x})) + 30y(E(\mathbf{x})) \right] \quad (21)$$

$$\beta_{iiii} = \frac{1}{24a^4 \mathbf{K}_{ii}^2} \left[y(E(\mathbf{x}) - 2a\sqrt{\mathbf{K}_{ii}}) - 4y(E(\mathbf{x}) - a\sqrt{\mathbf{K}_{ii}}) - 2y(E(\mathbf{x})) + 6y(E(\mathbf{x})) \right] \quad (22)$$

and the unique value of a is

$$a = \sqrt{\frac{15}{11}} \quad (23)$$

Proof – Under the conditions described in Theorem 3, the result of Corollary 1 hold and the truncation error E_{TN} is zero since the function $y(\mathbf{x})$ is exactly a fourth order polynomial therefore

$$\eta(\mathbf{z}) = \beta_0 + \sum_{i=1}^n \beta_i \mathbf{z}_i + \sum_{j=1}^n \sum_{\substack{i=1 \\ i \leq j}}^n \beta_{ij} \mathbf{z}_i \mathbf{z}_j + \sum_{j=1}^n \sum_{\substack{i=1 \\ i \leq j}}^n \sum_{\substack{k=1 \\ k \leq j}}^n \beta_{ijk} \mathbf{z}_i \mathbf{z}_j \mathbf{z}_k + \sum_{j=1}^n \sum_{\substack{i=1 \\ i \leq j}}^n \sum_{\substack{k=1 \\ k \leq j}}^n \sum_{\substack{l=1 \\ l \leq k}}^n \beta_{ijkl} \mathbf{z}_i \mathbf{z}_j \mathbf{z}_k \mathbf{z}_l$$

Given the structure of the samples, the polynomial coefficients in each individual variable can be computed exactly and are given by

$$\beta_{ii} = \frac{1}{a\sqrt{\mathbf{K}_{ii}}} \left[y(E(\mathbf{x}) - a\sqrt{\mathbf{K}_{ii}}) - 2y(E(\mathbf{x})) + y(E(\mathbf{x}) + a\sqrt{\mathbf{K}_{ii}}) \right]$$

$$\beta_{iii} = \frac{1}{a^2\mathbf{K}_{ii}} \left[y(E(\mathbf{x}) - a\sqrt{\mathbf{K}_{ii}}) - 2y(E(\mathbf{x})) + y(E(\mathbf{x}) + a\sqrt{\mathbf{K}_{ii}}) \right]$$

Based on the formula of Equation 12 the computed value of expectation shift will be

$$S' = (\beta_{ii} + 2a^2\mathbf{K}_{ii}\beta_{iii})\mathbf{K}_{ii} \quad (24)$$

In addition, it was stipulated that $\beta_{ijj} = 0$ where $i \neq j$ so the exact solution is given by Equation 16. Equating the exact solution (Eqn. 16) and the computed value (Eqn. 24)

and solving for a gives $a = \sqrt{\frac{15}{11}}$. •

5. Summary

The theorems developed here deal with expectation shift due to variability. This sort of shift occurs whenever the function of a random variable is non-linear and especially tends to occur in the neighborhood of a local maximum or minimum. The sampling locations and formulae given here should be able to help practitioners more quickly assess quality losses due to variation when the main source of loss is due to average performance rather than variability. The methods are exact for simple cases and make mild assumptions about continuity and smoothness of the response surface.

References

Cools, R., 1999, "Monomial Cubature Rules Since "Stroud": A Compilation – Part 2,"

Journal of Computational and Applied Mathematics **112**: 21-27.

Cools, R., P. Rabinowitz, 1993, "Monomial Cubature Rules Since "Stroud": A

Compilation," *Journal of Computational and Applied Mathematics* **48**(3): 309-326.

Diwekar U. M., 2003, "A novel sampling approach to combinatorial optimization under uncertainty" *Computational Optimization and Applications* **24** (2-3): 335-371.

- Giunta, A.A.; Wojtkiewicz, S.F., Jr.; and Eldred, M.S., 2003, "Overview of Modern Design of Experiments Methods for Computational Simulations," *AIAA 2003-0649*, 41st AIAA Aerospace Sciences Meeting and Exhibit, Reno, NV.
- Kalagnanam, J. R. and U. M. Diwekar, 1997, "An Efficient Sampling Technique for Off-Line Quality Control", *Technometrics* **39**(3)308-319.
- Lu, J. L., and D. Darmofal., 2005, "Numerical Integration in Higher Dimensions with Gaussian Weight Function for Application to Probabilistic Design," *SIAM Journal of Scientific Computing*, **26** (2) 613-624.
- McKay, M.D., R. J. Beckman, and W. J. Conover, 1979, "Comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics* **21**(2)239-245.
- Santner, T. J., B. J. Williams, and W. Notz, 2003, *The Design and Analysis of Computer Experiments*, Springer-Verlag, New York.
- Simpson, T.W.; Peplinski, J.; Koch, P. N.; and Allen, J.K., 2001, "Meta-models for Computer-based Engineering Design: Survey and Recommendations," *Engineering with Computers* **17**(2)129-150.
- Taguchi, G., 1987, *System of Experimental Design: Engineering Methods to Optimize Quality and Minimize Costs*, American Supplier Institute, Dearborn, MI.
- Wu, C. F. J., and M. Hamada, 2000, *Experiments: Planning, Design, and Parameter Optimization*, John Wiley & Sons, New York.

-
- Chipman, H. M., M. Hamada, and C. F. J. Wu, 1997, "Bayesian variable-selection approach for analyzing designed experiments with complex aliasing," *Technometrics* **39**(4) 372-381.
- Christoffel, E. B., 1877, "Sur une classe particuliere de fonctions entieres et de fractions continues," *Annali di Matematica Pura ed Applicata*, series 2, vol. 8, pp. 1-10.
- F.Mistree, U. Lautenschläger, S. O. Erikstad, and J. K. Allen. *Simulation reduction using the Taguchi method*. Contractor Report 4542, NASA, October 1993.
- Fieller, E. C. (1932). The distribution of the index in a normal bivariate population. *Biometrika* **24**, pp. 428-40.
- Gauss, 1814, "Methodus nova integralium valores per approximationem inveniendi," *Comment. Soc. Regiae Sci. Gottiningensis Recentiores*, vol. 3, pp. 202-206.
- Gautchi, W., 1983, "How and How Not to Check Gaussian Quadrature Formulae," *BIT* **23** 209-216.
- Gautchi, W., 1968, "Construction of Gauss-Christoffel Quadrature Formulas," *Mathematics of Computation* **22** 102 251-270.
- Hatton, L., 1997, "The T-experiments in Quality of Numerical Software, Assessment and Enhancement," *IEEE Computational Science & Engineering*, **4**(2), 27-38.
- Hazelrigg, G., 2003, "Thoughts on Model Validation for Engineering Design," in *Proceedings of the ASME Design Theory and Methodology Conference*, Chicago, IL, DETC2003/DTM-48632.

- Huang, B. and Du, X. 2005, "Uncertainty analysis by dimension reduction integration and saddlepoint approximations," *ASME Design Engineering Technical Conferences*, Sept. 24-28, Long Beach, CA.
- Jensen, J. L. W. V. (1906). "Sur les fonctions convexes et les inégalités entre les valeurs moyennes". *Acta Mathematica* 30 (1): 175–193.
- Joseph, V. R., 2006, "A Bayesian approach to the design and analysis of fractionated experiments," *Technometrics*, v 48, n 2, p 219-229.
- Li, X. and D. Frey, 2005, "Study of Effect Structures in Engineering Systems," *ASME Design Engineering Technical Conferences*, Sept. 24-28, Long Beach, CA.
- Magnus, J., 1986, "The Exact Moments of a Ratio of Quadratic Forms of Normal Variables", *Annales d'Economie et de Statistique*, 4, pp. 95-109.
- Manteufel, R. D., 2000, "Evaluating the Convergence of Latin Hypercube Sampling", *AIAA-2000-1636, 41st AIAA Structures, Structural Dynamics, and Materials Conference*, Atlanta, GA, 3-8 April.
- Phadke, Madhav S., 1989, *Quality Engineering Using Robust Design*, Prentice Hall, Englewood Cliffs, NJ.
- Rahman, S. and Wei D. 2005, "A univariate approximation at most probable point for higher-order reliability analysis" Submitted to *International Journal of Solids and Structures*.
- Stroud, A. H., 1971, *Approximate Calculation of Multiple Integrals*, Prentice Hall Series in Automatic Computation, Englewood Cliffs, NJ.
- Stroud, A. H., and D. Secrest, 1966, *Gaussian Quadrature Formulas*, Prentice Hall, Englewood Cliffs, NJ.

- Taguchi, G., 1987, *System of Experimental Design: Engineering Methods to Optimize Quality and Minimize Costs*, American Supplier Institute, Dearborn, MI.
- Yu, J.; Ishii. K., 1998, "Design optimization for robustness using quadrature factorial models" *Engineering Optimization*, v 30, n 3/4, p 203-225.