center, the angular flux tends to be isotropic, and it may not be necessary to choose μ_m as in case 2 in order to integrate quadratic polynomials exactly.

Thus the choice of $\mu_m = \mu_{mean}$ leads to a better representation of the neutron flow across the physical faces than the choice in case 2 near the center. Of course, the net flow across the faces of the phase-space cell, which is the sum of the first two terms on the left side of Eq. (1), may not be affected by the choice of μ_m in case 2. This can be easily seen in the light of the following relation:

$$\alpha_{m+1/2} - \alpha_{m-1/2} = -\omega_m \mu_m \quad . \tag{7}$$

Nevertheless, the choice of μ_m, as in case 2, may lead to some incorrect redistribution of neutrons in angular and space variables. When regions away from the center are considered, the choice of μ_m is not so crucial. In that case, ΔE is obtained by replacing A_1 by $(A_{i+1} - A_i)$ in Eqs. (4) and (6). As $(A_{i+1} - A_i)/V_{i+1/2}$ is of order $(1/r)$ and $\partial N/\partial\mu$ need not tend to zero, it is seen for both cases that ΔE is of order $(\Delta\mu^2/r)$, which is small for large r.

R. S. Modak

Bhabha Atomic Research Centre
Reactor Analysis and Systems Division
Bombay - 400 085, India

October 31, 1984

REFERENCES

1. W. H. REED and K. D. LATHROP, *Nucl. Sci. Eng.*, **41**, 237 (1970).

2. G. I. BELL and S. GLASSTONE, *Nuclear Reactor Theory*, Van Nostrand Reinhold Company, New York (1970).

On Efficient Estimation of Variances

In Ref. 1 Dubi addresses the problem of the optimum estimation of variance from a given number of realizations of a random variable. The question is posed in the following way. Let $x_1, x_2, \ldots, x_n$ be independent realizations of the random variable x with an expectation μ and variance σ^2. Let the realizations be divided into groups ("batches") containing k realizations each and let

$$\bar{S}_i = \frac{1}{k}\sum_{j=1}^{k} x_{(i-1)k+j} \quad . \tag{1}$$

If the variance of x is estimated in terms of the batchwise averages $\bar{S}_i$, what is the value of k that minimizes the variance of the estimated variance? In other words, what is the optimum batch size that gives the most reliable estimate of the variance? Dubi proves that *if the expectation* μ *of* x *is known, the optimum value of* k *is 1*, i.e. the "one-particle" estimation of the variance is the most reliable. He also shows that in the optimum case, the variance of the estimate is

$$D^2 = \frac{1}{n}\left[\langle (x-\mu)^4\rangle - \langle (x-\mu)^2\rangle^2\right] \quad , \tag{2}$$

where angle brackets stand for expectation.

In most practical cases, however, the expectation of the random variable is not known and therefore cannot be used in the estimation of the variance. In this letter we show that $k = 1$ is optimal also if the expectation is to be estimated, and this estimate is used in the estimate of the variance. It will also be shown that the variance of the entirely empirical estimate of the variance is higher than that in Eq. (2).

Let

$$\hat{S} = \frac{1}{p}\sum_{i=1}^{p}\bar{S}_i = \frac{1}{n}\sum_{i=1}^{n} x_i \quad , \tag{3}$$

where $\bar{S}_i$ is the batchwise average in Eq. (1) and $p = n/k$, the number of batches formed from the n realizations. Obviously,

$$\langle \hat{S}\rangle = \mu \quad ;$$

i.e., $\hat{S}$ is an unbiased estimate of the mean. Let

$$\bar{v}_i = \frac{n}{(p-1)}(\bar{S}_i - \hat{S})^2 \quad ; \tag{4}$$

then it is easily seen that

$$\langle \bar{v}_i\rangle = \langle (x-\mu)^2\rangle = \sigma^2 \quad , \quad (j = 1,2,\ldots,p) \quad ;$$

i.e., $\bar{v}_i$ represents a realization of a random variable that has the expectation σ^2. The sample average formed from these realizations is

$$\hat{v} = \frac{1}{p}\sum_{i=1}^{p}\bar{v}_i \quad , \tag{5}$$

which is obviously also unbiased with respect to σ^2. The question again is how to choose the value p (or equivalently the value of $k = n/p$) in order to minimize the variance of the estimate $\hat{v}$ in Eq. (5). The question is answered by calculating the variance in question. The variance of the estimate is

$$\hat{D}^2 = \langle \hat{v}^2\rangle - \langle \hat{v}\rangle^2 = \left[\frac{n}{(p-1)}\right]^2 \left\langle \left[\frac{1}{p}\sum_{i=1}^{p}(\bar{S}_i - \hat{S})^2\right]^2 \right\rangle - \sigma^4 \quad . \tag{6}$$

The optimum value of $k = n/p$ with n given is the one that minimizes the quantity

$$Q(k) = \left[\frac{n}{(p-1)}\right]^2 \left\langle \left[\frac{1}{p}\sum_{i=1}^{p}(\bar{S}_i - \hat{S})^2\right]^2 \right\rangle \quad . \tag{7}$$

In order to make the derivations simpler, we introduce random variables with zero expectations. Thus, let

$$y = x - \mu \quad , \quad y_i = x_i - \mu \quad , \quad \bar{Z}_i = \bar{S}_i - \mu \quad .$$

Then

$$\langle \bar{Z}_i\rangle = \langle \bar{Z}_i\bar{Z}_j\bar{Z}_r\bar{Z}_s\rangle = 0 \quad , \quad (i \neq j,r,s) \quad , \tag{8}$$

since the realizations x_i are assumed to be independent and therefore so are the sample averages $\bar{Z}_i$. Now, in view of Eqs. (1) and (3), the quantity in Eq. (7) reads

$$Q(k) = \left[\frac{n}{p(p-1)}\right]^2 \left\langle \left[\sum_{i=1}^{p}\bar{Z}_i^2 - \frac{1}{p}\left(\sum_{i=1}^{p}\bar{Z}_i\right)^2\right]^2 \right\rangle$$
$$= \left[\frac{n}{p(p-1)}\right]^2 \left\langle \left(\frac{p-1}{p}\sum_{i=1}^{p}\bar{Z}_i^2 - \frac{1}{p}\sum_{i=1}^{p}\sum_{\substack{j=1\\ j\neq 1}}^{p}\bar{Z}_i\bar{Z}_j\right)^2 \right\rangle \quad .$$

Explicit calculation of the square in brackets yields

$$Q(k) = \left[\frac{n}{p^2(p-1)}\right]^2 \left\langle (p-1)^2\sum_{i=1}^{p}\bar{Z}_i^2 + [(p-1)^2 + 1] \times \sum_{i=1}^{p}\sum_{\substack{j=1\\ j\neq 1}}^{p}\bar{Z}_i^2\bar{Z}_j^2 \right\rangle \quad ,$$

as the odd powers of $\bar{Z}_i$ have zero expectation according to Eq. (8). Finally, since from a statistical point of view, all the $\bar{Z}_i$'s are equivalent, we have

$$\langle \bar{Z}_i^r \rangle = \langle \bar{Z}_1^r \rangle = \left\langle \left(\frac{1}{k} \sum_{i=1}^{k} y_i \right)^r \right\rangle \quad (9)$$

and

$$Q(k) = \frac{n^2}{p^3} \langle \bar{Z}_1^4 \rangle + \frac{n^2[(p-1)^2+1]}{p^3(p-1)} \langle \bar{Z}_1^2 \rangle^2 \ . \quad (10)$$

It remains to determine the expectations in Eq. (9) for $r = 2$ and $r = 4$. For $r = 2$ the result is commonplace:

$$\langle \bar{Z}_1^2 \rangle = \left\langle \left(\frac{1}{k} \sum_{i=1}^{k} y_i \right)^2 \right\rangle = \left\langle \frac{1}{k^2} \sum_{i=1}^{k} y_i^2 \right\rangle = \frac{1}{k} \langle y^2 \rangle \ . \quad (11)$$

Similarly for the fourth moment,

$$\langle \bar{Z}_1^4 \rangle = \left\langle \left(\frac{1}{k} \sum_{i=1}^{k} y_i \right)^4 \right\rangle = \frac{1}{k^4} \left\langle \left(\sum_{i=1}^{k} y_i^2 + \sum_{i=1}^{k} \sum_{\substack{j=1 \\ j \neq i}}^{k} y_i y_j \right)^2 \right\rangle$$

$$= \frac{1}{k^4} \left\langle \sum_{i=1}^{k} y_i^4 + 2 \sum_{i=1}^{k} \sum_{\substack{j=1 \\ j \neq i}}^{k} y_i^2 y_j^2 \right\rangle$$

$$= \frac{1}{k^3} \langle y^4 \rangle + \frac{2(k-1)}{k^3} \langle y^2 \rangle^2 \ . \quad (12)$$

Inserting Eqs. (11) and (12) into Eq. (10), the quantity to be minimized becomes

$$Q(k) = \frac{1}{n} \left[\langle y^4 \rangle + \left(n - 2 + k + \frac{k^2}{n-k} \right) \langle y^2 \rangle^2 \right] \ , \quad (13)$$

where we have put $p = n/k$. Obviously $Q(k)$ is minimal with $k = 1$; i.e., *the variance of the estimated variance is minimal if every batch consists of a single realization.*

Therefore, when estimating the theoretical variance of a random variable from the realizations $x_1, x_2, \ldots, x_n$, the most efficient estimate follows from Eqs. (4) and (5) as

$$\hat{V} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \hat{S})^2 \ , \quad (14)$$

where $\hat{S}$ is the empirical mean of the realizations as given in Eq. (3).

The minimum value of $Q(k)$ is

$$Q(1) = \frac{1}{n} \left[\langle y^4 \rangle + \left(n - 1 + \frac{1}{n-1} \right) \langle y^2 \rangle^2 \right]$$

$$= \frac{1}{n} [\langle (x-\mu)^4 \rangle - \langle (x-\mu)^2 \rangle^2] + \left[\frac{1}{n(n-1)} + 1 \right] \langle (x-\mu)^2 \rangle^2 \ ,$$

and according to Eq. (6), the variance of the optimal variance estimate is

$$\hat{D}^2 = Q(1) - \langle (x-\mu)^2 \rangle^2$$

$$= \frac{1}{n} [\langle (x-\mu)^4 \rangle - \langle (x-\mu)^2 \rangle^2] + \frac{1}{n(n-1)} \langle (x-\mu)^2 \rangle^2 \ .$$

Comparing it with Eq. (2), it is apparent that

$$\hat{D}^2 = D^2 + \sigma^4/[n(n-1)] \ ;$$

i.e., *the variance is increased by* $\sigma^4/[n(n-1)]$ *if the mean of the realizations is not known but is also to be estimated.*

Iván Lux

Central Research Institute for Physics
P.O. Box 49
H-1525 Budapest, Hungary

January 2, 1985

REFERENCE

1. A. DUBI, *Nucl. Sci. Eng.*, **72**, 108 (1979).

Limitations on the Use of the THOR Critical Assembly for Validation of n + ^{239}Pu Cross Sections

A major revision of the ENDF/B-V evaluation of neutron-induced nuclear data for ^{239}Pu was recently provided by Arthur et al.[1] The revised data were validated by calculating measured quantities for the five fast critical assemblies JEZEBEL, JEZEBEL-PU, FLATTOP-PU, THOR, and ZPR-6/7. The integral parameters calculated are k_{eff} and certain fission ratios. The revised data set improved the agreement between calculated and measured integral parameter values ("the agreement") for all the assemblies except the JEZEBEL-PU assembly. Table I gives the extent of improvement obtained[1] using the revised set for the five assemblies. One can see that the improvement of the agreement is maximum for the THOR assembly and the agreement has worsened in the case of JEZEBEL-PU. The authors[1] have emphasized that while their new inelastic, elastic, and total cross-section results are based on a thorough analysis, the $\bar{\nu}_p(E_n)$ and fission spectrum modifications in their paper are of an interim nature, because in both cases entire data bases were not considered. They have also suggested a new analysis of the resolved and unresolved resonance regions that extends the resolved resonance region to as high an energy as feasible and also an analysis of smooth (n,f) and (n,γ) cross sections that accounts for energy correlations in the data.

The purpose of our letter is to point out that the good improvement in the agreement obtained in the case of the THOR assembly may be fortuitous. The comment is only on the weakness of using the THOR assembly for testing ^{239}Pu cross sections and not on the quality of the evaluation of ^{239}Pu cross sections themselves.

TABLE I

Deviations of k_{eff} from Unity for the Critical Assemblies When ENDF/B-V and Revision 2 Data Sets for ^{239}Pu Are Used

Critical Assembly	ENDF/B-V (%)	Revision 2 (%)
JEZEBEL	0.68	−0.18
JEZEBEL-PU	−0.20	−0.83
FLATTOP-PU	0.93	0.50
THOR	2.28	0.70
ZPR-6/7	−0.44	−0.42