VICTORIA UNIVERSITY
OF TECHNOLOGY

DEPARTMENT OF
MATHEMATICS, COMPUTING
AND OPERATIONS RESEARCH

THE VALIDITY OF SAMPLING THEORY ...... A RESPONSE

Neil S. Barnett

(10 E.Q.R.M. 6)

March 1991

TECHNICAL REPORT

DEPT OF MCOR
FOOTSCRAY INSTITUTE OF TECHNOLOGY
VICTORIA UNIVERSITY OF TECHNOLOGY
BALLARAT ROAD (P O BOX 64), FOOTSCRAY
VICTORIA, AUSTRALIA 3011
TELEPHONE (03) 688-4249/4225
FACSIMILE (03) 687-7632

THE VALIDITY OF SAMPLING THEORY ...... A RESPONSE

Neil S. Barnett

(10 E.Q.R.M. 6)

March 1991

Efficiency, Quality and Reliability
Management Centre

This report arose from an article by Allison and Ude that appeared in the January 1991 edition of 'Quality Australia'. My assessment of their article is that they have got it all wrong (I hasten to add, something that's happened to me on more than one occasion and perhaps even now!).

Although the material in the report is flavoured to refute the substance of Allison and Ude's paper, it is none-the-less self contained.

I'm indebted to Neil Diamond for his time, comment and reading of my script also to Peter Cerone and Len Armour for their interest and willingness to listen to my, at times, lengthy rhetoric on the material contained herein.

The Validity of Sampling Theory .... A Response.

The article, 'The Validity of Sampling Theory' that appeared in 'Quality Australia' January 1991 contains statements that are incorrect. Wrong inferences have been drawn from earlier work and comments made to point out supposed theoretical deficiencies. The material of this current article seeks to correct these errors.

The substance of Allison and Ude's article, 'in a nut shell', is described by the following diagram

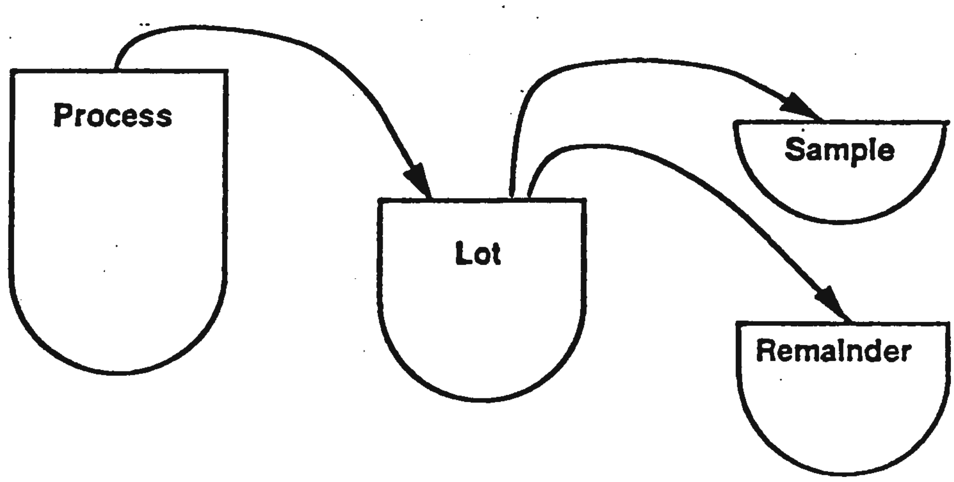

Figure 1: Selection of samples from lots drawn from a process.

It focuses on the reputed independence of the number of defectives found in a sample and the number of defectives found in the remainder. The proof of this independence has led to some bold assertions regarding the usefulness of sampling production 'lots'.

Although the bulk of the readership are not statisticians, since a statistical matter has been raised, it is necessary to deal with both the statistical and pragmatic issues in some detail. The typical practical situation is one of mass - production. A 'lot' is merely a sample produced by the process and invariably arises as a consequence of despatch quantities e.g. a 'lot' maybe a pallet of product or a standard truck load. The further sampling of 'lots' occurs in an attempt to re-assure the customer as to the quality of the product being consigned. This is a separate (but not independent) issue from the manufacturer wishing to sample the product produced by the process in order to monitor the proportion of defectives being made. He may know from past experience that a standard of say 3% defectives are produced and he wishes to have an early warning of a change in this proportion so that he may take appropriate action. These two situations are essentially examples of enumerative and analytic problems as defined by Deming in [1] (the terms themselves are not important but the types are).

Perusing once again the diagram of figure 1, consider a single item produced by the process. Assuming that the process is in a stable mode of operation, then the assumed fixed proportion of defectives produced is merely the probability that this single item will be defective. Should we take, for example, four items directly from the process, each individual item has this same probability of being defective, which means that the probability of having a certain number defective in the four is given by the binomial distribution.

A 'lot', as defined in figure 1, is such a number of selected process individuals and therefore the number of defectives in it follows likewise a binomial distribution. It is as if the process is a generator of items from an infinite source containing a fixed proportion of defectives.

Taking a sample from a 'lot' however, is a different process - the 'lot' is of a clearly determined size. In considering samples from it we cannot assume that the 'lot' is a generator of items from an infinite source - it is clearly finite. The 'lot' contains a fixed number of defectives (whether we know this number or not). If we consider the items constituting the sample selected from the 'lot' one at a time, the probability of getting a defective at any selection is changing depending on the outcome of previous selections. The appropriate method for ascertaining the probability of the number of defectives in the sample is to use the so-called hypergeometric distribution which is pointedly dependent on the number of defectives in the 'lot'. The calculation consists of taking the ratio of the total number of different ways samples can be taken from the 'lot' of specified size having the specified number of defectives, to the total number of ways the samples of specified size can be taken from the 'lot'. The essential difference in a 'lot' and a 'sample', as defined in figure 1, is that a 'lot' is a sample from an infinite population and a 'sample' is a sample from a finite population.

Returning to the two issues of process monitoring, as opposed to assessing the quality of 'lots', if process monitoring was the sole aim then 'lots' would not be sampled but examined in total as direct measures of process performance, and a conventional 'p' chart plotted (lot sizes would be chosen to satisfy the usual criteria). This is shown later.

The situation shown in figure 1 is, therefore, assumed to be a regime for assessing the quality of 'lots' (the issue of being able to use the information obtained to also monitor the process is dealt with later). Let's now examine the mathematics discussed by Allison and Ude in their appendix 1 using the same notation. As the authors point out, the fundamental definition of independence is that the joint probability of two events is the product of their unconditional probabilities. A definition of independence which is equivalent to this but which has much more intuitive appeal, is to say that two events are independent if the probability of one given that the other has occurred is the same as the probability of the one occurring irrespective of the others' occurrence.
i.e. for events A and B

if $P(A/B) = P(A)$ and $P(B/A) = P(B)$ then A and B are independent.

Taking N and n to be 'lot' and 'sample' sizes respectively and X and r to be the corresponding number of defectives actually in each, it follows that $r \leq X$. This fact must 'flavour' the calculation of p(r), the probability that the sample (taken from a finite source) contains exactly r defectives.

Elementary probability theory shows that:-

$$p(r) = \sum_{X=r}^{N} p(r/X)\, p(X)$$

Because the N, constituting the 'lot', are selected from an infinite source, p(X) follows a binomial distribution. Because the r defectives in the sample are drawn from the finite number of defectives in the 'lot', X, p(r/X) follows the hypergeometric distribution and is equal to:

$$\frac{\binom{X}{r}\binom{N-X}{n-r}}{\binom{N}{n}}$$

and by necessity,

$r \leq X$ and $n - r \leq N - X$.

This latter expression can be written alternatively (see Allison and Ude) as:

$$\frac{\binom{n}{r}\binom{N-n}{X-r}}{\binom{N}{X}}$$

$$\therefore\ p(r) = \sum_{X=r}^{N+r-n} \frac{\binom{n}{r}\binom{N-n}{X-r}\binom{N}{X} p^X q^{N-X}}{\binom{N}{X}}$$

$$= \binom{n}{r} (q)^N \sum_{X=r}^{N+r-n} \binom{N-n}{X-r} \left(\frac{p}{q}\right)^X$$

$$= \binom{n}{r} (q)^N \left[\binom{N-n}{o}\left(\frac{p}{q}\right)^r + \binom{N-n}{1}\left(\frac{p}{q}\right)^{r+1} + \ldots\ldots\ldots \binom{N-n}{N-n}\left(\frac{p}{q}\right)^{r+N-n} \right]$$

$$= \binom{n}{r} p^r (q)^{N-r} \left[\binom{N-n}{0} + \binom{N-n}{1}\frac{p}{q} + \ldots\ldots\ldots + \binom{N-n}{N-n}\left(\frac{p}{q}\right)^{N-n} \right]$$

$$= \binom{n}{r} p^r q^{N-r} \left(1 + \frac{p}{q}\right)^{N-n}$$

$$= \binom{n}{r} p^r q^{n-r}$$

which is binomial in form and is identical to the probability of the same, had the sample been taken directly from the process. This shows that there is no advantage, if our sole aim is to monitor p, to take lots and sample these; we may as well sample directly from the process. As the derivation shows however, this is indeed pertinent to the 'sample' coming from a 'lot', contrary to the comment made by Allison and Ude in the appendix to their article that this expression relates just to a sample taken from a process not from a 'lot'. The two probabilities just happen to be the same. Their further comment on the validity of p(X-r) is therefore, also invalid.

Now $p\left[X \cap r\right] = p\left[X/r\right] . p\left[r\right] = p\left[r/X\right] . p\left[X\right]$

$$\therefore\ p\left[X/r\right] = \frac{p\left[r/X\right] . p\left[X\right]}{p\left[r\right]}$$

$$\therefore\ p\left[X/r\right] = \frac{\binom{n}{r}\binom{N-n}{X-r}}{\binom{N}{X}} \times \frac{\binom{N}{X} p^{X} q^{N-X}}{\binom{n}{r} p^{r} q^{n-r}}$$

$$= \binom{N-n}{X-r} p^{X-r} q^{N-X-n+r} \qquad \text{\_\_\_\_\_\_\_\_} \quad (*)$$

$$X = r, r+1, \ldots\ldots N-n+r$$

which is also of binomial form.

Since $p(X = \binom{N}{X} p^{X} q^{N-X}$ and $p\left[X/r\right] \neq p\left[X\right]$

this implies a dependency of X on r i.e. the number of defectives in the 'lot' is DEPENDENT on the number of defectives in the sample. The theory presented in Appendix 1 of Allison and Ude is in fact correct, contrary to their remarks; Mood's paper [2] is quite clear in this regard. It is thus fact, as the previous result shows, that the number of defectives in the sample and in the remainder are uncorrelated, yet there is a positive correlation between the number of defectives in the 'sample' and the number of defectives in the 'lot'! This at first seems strange since the 'lot' consists of a sum total of the sample and the remainder. Strange it may seem but fact it is and it is this subtlety that the authors' have missed. The fact that the sum total includes the known number of defectives in the sample induces this correlation. The conclusion that Allison and Ude make, that if the number of defectives in the sample and in the remainder are independent then this implies that the sample contains no information about the lot, is fallacious. Clearly the dependency of X on r implies that p(X/r) is a means of making inferences about X from knowledge of r.

It can easily be shown that

$$E\left[X/r\right] = (N-n)\,p + r\ ,$$

$$V\left[X/r\right] = (N-n)\,pq\ ,$$

and that the correlation between r and X is $\sqrt{\frac{n}{N}}$. Without sampling the lot, the number of defectives contained in it can be estimated using the distribution $p\,(X) = \binom{N}{X} p\ q^{N-X}$, the variance being Npq.

The proportional decrease in estimation variance, induced by adopting a routine of sampling the lots, can be seen to be $\frac{Npq - (N - n)\,pq}{Npq} = \frac{n}{N}$ which is the square of the correlation or the ratio of the sample size to the batch size.

Mood [2] concludes correctly that the independence of r and X-r makes nonsense of the practice of sampling and inspecting the remainder on the basis of the number of defectives found in the sample. Nowhere does he conclude that sample information tells nothing about the 'lot'. Deming [3] similarly addresses the issue of justification for further inspection of the remainder on the strength of sample findings.

The distribution of X given r shown in (*) is appropriate for use if we wish to have a basis for discussion of the quality of the whole 'lot'. It should be noted that it is a function of (X - r) which is the number of defectives in the remainder.

Allison and Ude point out that the customer purchases 'lots' not remainders. It is because of this that the above distribution is pertinent. Calculations of course need to allow for the replacement of any defectives found in the sample. It is assumed that this would be normal practice. In considering replacement when sampling preceding despatch detects defectives, there are three obvious procedures:-

(i) The defectives are not replaced but are included when the 'sample' is returned to the 'lot'.

(ii) The defectives are replaced by others (untested) from the process before returning the 'sample' to the 'lot'.

(iii) The defectives are replaced by non-defective (tested) items from the process before returning the 'sample' to the 'lot'.

Following procedure (i), although an unlikely practical occurrence, the situation after sampling is the same as before. Following (iii) implies additional sampling and testing, the end result being that the number of defectives in the 'lot' is reduced by the number found in the 'sample' once sampling and replacement are complete. Under such circumstances the number of defectives in the 'lot' following sampling and replacement is merely the number of defectives in the remainder (that part of the 'lot' not tested). As has been seen this number is independent of the number of defectives orginally found in the 'sample' and is thus no basis for inference. Hence if procedure (iii) is adopted, although the same sample information is available as for (i) and (ii) the replacement strategy adopted renders the original sample information ineffective as a basis for inference of the number of defectives in the 'lot'. As previously pointed out, the procedure itself implies additional sampling which destroys the original sampling strategy as the sole basis for inference. The most likely strategy to adopt would be (ii). At the conclusion of sampling and replacement the inferential probabilities would be

P[Lot contains X defectives/original sample contains r defectives]

$$= \sum_{s=0}^{r} p\,(X+r-s/r) \binom{r}{s} p^s q^{r-s}$$

$$= p^X q^{N-n-X+r} \sum_{s=0}^{r} \binom{N-n}{X-s} \binom{r}{s}$$

Should it be required to monitor the stability of the process with respect to p then since $p\,(r) = \binom{n}{r} p^r q^{n-r}$ is binomial, a standard p or np chart could be plotted based on the number of defectives found in samples which, it has been shown, could be taken directly from the process.

The illustration given on page 12 of Allison and Ude's article when X is known shows numerically the dependency of the number of defectives in the sample on the number of defectives in the remainder. This, however, does not relate to the issue raised in their paper, since in practice, as they point out, X is unknown.

When X is known the situation is different because

$$p\left[r/X\right] = p\left[X-r/X\right] = \frac{\binom{n}{r}\binom{N-n}{X-r}}{\binom{N}{X}}$$

Clearly, under such circumstances

$$p\left[(r \cap X-r)/X\right] = \frac{\binom{n}{r}\binom{N-n}{X-r}}{\binom{N}{X}}$$

and the dependency that the authors illustrate numerically is seen to be theoretically correct

$$\left(p\left[(r \cap X-r)/X\right] \neq p\left[r/X\right] . \, p\left[X-r/X\right]\right).$$

The issue, however, is not the dependency of the number of defectives in the sample and the number of defectives in the remainder when X is known, it is their independence when X is unknown.

Mood's result is correct and it in no-way confronts or seeks to discredit standard sampling theory. What it does do is say that for an in control process an acceptance sampling procedure that tests lot remainders on the basis of the number of defectives found in the sample, is unsound. The sample findings already contain information that can be used directly to assess the quality of the lot.

References

[1] W.Edwards Deming, 'Some Theory of Sampling', Wiley 1950.

[2] Alexander Mood, 'On the Dependence of Sampling Inspection Plans upon Population Distributions' Annals of Mathematical Statistics, 14, 1943.

[3] W. Edwards Deming, 'Out of the Crisis', M.I.T. Centre for Advanced Engineering Study, 1986.