An Essay on Screening, or on Two-phase Sampling, applied to Surveys of a Community

W. Edwards Deming

Consultant in Statistical Studies, Washington 20016

International Statistical Review, 45 (1977) 29–37

An Essay on Screening, or on Two-phase Sampling, applied to Surveys of a Community

W. Edwards Deming

Consultant in Statistical Studies, Washington 20016

International Statistical Review, 45 (1977) 29–37

International Statistical Review, 45 (1977) 29–37

An Essay on Screening, or on Two-phase Sampling, applied to Surveys of a Community

W. Edwards Deming
Consultant in Statistical Studies, Washington 20016, USA

1 Purpose

A study will be conducted to estimate the overall proportion of people that are affected with some defined psychopathology. The final determination of the psychiatric and other medical characteristics of a person will be made by a psychiatrist. A plan to use the services of trained interviewers to screen and separate into two classes (with and without apparent psychopathology) a large preliminary sample in order to conserve the time of the psychiatrist, by letting him test mainly cases that are almost surely afflicted with psychopathology, is appealing wherever the cost per case is much lower for the screening than for the psychiatric examination. It is not generally appreciated, however, that the screening-test, to be economical, must be relatively cheap and must admit only a low proportion of false negatives. This principle is not new, but illustrative calculations that show how false negatives affect costs, and why false positives are not so important, are hard to find in the literature (Kish, 1965).

Guidance in any problem comes from calculations based on the appropriate theory. The purpose here is to present some theory and a simple illustration encountered in recent practice. The conclusions drawn here will be valid within a moderately wide band of conditions that border on those used here for illustration. Conditions far afield from those studied here might require fresh calculations by use of the appropriate costs and proportions in the equations that follow, or in modifications thereof.

The conclusions drawn here are worthy of consideration in the inspection of industrial product, in situations where the final test is relatively very expensive, and where a cheap screening test can be contrived.

It is presumed that a demographic screening has already taken place in which a roster is made of each family by age of person. People of age 60 or over can be serialized. These serial numbers constitute the frame.

The statistical procedure for screening (sometimes called two-phase sampling) may be described briefly in two steps.

Step 1 (1st phase). Screening. Draw from the frame a preliminary sample of N' people. Interview by a cheap test every person in the preliminary sample. Allot each person interviewed to one of two strata:

Stratum 1: negative on screening (no psychopathology indicated).

Stratum 2: positive on screening (psychopathology indicated).

Step 2 (2nd phase). Psychiatric interviews. A psychiatrist interviews samples from both strata. His decisions are final. Some people in Stratum 1, the psychiatrist will find, are pathologic. These are false negatives. Conversely, he will find that some people put into Stratum 2 are in his judgment not pathologic. These are false positives.

The final sample for the psychiatrist is drawn partly from Stratum 1 and partly from Stratum 2. The selections from each stratum are made by simple random sampling, one person at a time. Textbooks on statistical procedures describe two main ways to draw for the

psychiatrist a sample of people from the two strata: (1) proportionate allocation; (2) Neyman allocation.

We calculate also, for comparison, the amount of information to be expected from a plan that uses no screening at all. There is no preliminary sample in this plan: the psychiatrist interviews the entire sample.

We shall see that the distinction between proportionate allocation and Neyman allocation is important only if the screening is highly reliable, and that the best procedure may be no screening at all.

As we shall see, the proportion p_1 of false negatives in Stratum 1 is critical. The proportion q_2 of false positives in Stratum 2, on the other hand, is not critical, though it must not get out of hand.

In the ideal situation, there would be no psychopathology in Stratum 1 and nothing but psychopathology in Stratum 2. Interviews by the psychiatrist would yield identical results. This goal can of course not be achieved: there will in practice be false negatives in Stratum 1 and false positives in Stratum 2.

It would be easy to construct a system of screening that would hardly ever put positive psychopathology into Stratum 1. It is only necessary to specify that a person that exhibits in the screening a shred of evidence of psychopathology shall be placed in Stratum 2. Such a procedure could easily get out of hand: the proportion q_2 of false positives in Stratum 2 would reach an alarming proportion and would defeat the purpose of screening. We now proceed with the calculations.

2 Notation

P_1 the expected proportion of cases placed in Stratum 1 in Step 1

P_2 the expected proportion of cases placed in Stratum 2 in Step 1

p_1 the expected proportion of false negatives in Stratum 1

q_2 the expected proportion of false positives in Stratum 2

p the proportion pathologic in the entire frame

$\hat{p}$ an estimate of p.

We define q_1 and p_2 so that

$$p_1+q_1=1 \quad p_2+q_2=1$$

n_1 the number of persons selected from Stratum 1 for the final sample (for examination by the psychiatrist)

n_2 likewise for Stratum 2

$n_1+n_2=n$ the size of the final sample (for the psychiatrist); n will depend on the plan adopted

$\sigma_1^2=p_1q_1$ the expected variance between people in Stratum 1

$\sigma_2^2=p_2q_2$ the expected variance between people in Stratum 2

$\sigma_w^2=P_1\sigma_1^2+P_2\sigma_2^2$ the average variance between sampling units within strata

$\bar{\sigma}_w=P_1\sigma_1+P_2\sigma_2$ the average standard deviation between sampling units within strata

$\sigma_b^2=P_1P_2(p_2-p_1)^2$ the variance between the means of the two strata

N' the number of people screened (the preliminary sample)

c_1 the cost to screen one person

c_2 the cost for the psychiatrist to interview one person.

The relative proportions of the two strata in any one study give the estimates $\hat{P}_1$ and $\hat{P}_2$ of P_1 and P_2. Step 2 gives the estimates $\hat{p}_1$ and $\hat{q}_2$ of p_1 and q_2. Then

$$\hat{p} = \hat{P}_1\hat{p}_1 + \hat{P}_2\hat{p}_2 \tag{1}$$

will be an unbiased estimate of the overall proportion p that are pathologic in the entire frame.

$\hat{P}_1$ and $\hat{P}_2$ come from the preliminary sample, the screening: they are the proportions in the two strata. $\hat{p}_1$ and $\hat{p}_2$ come from the psychiatric interviews.

A 2×2 diagram may be helpful. The psychiatric interviews separate Stratum 1 into two groups with proportions p_1 and q_1, and separate Stratum 2 into two groups whose expected proportions will be p_2 and q_2.

Psychiatric interview	Screening	
	Stratum 1	Stratum 2
No psychopathology	P_1q_1	P_2q_2
Psychopathology	P_1p_1	P_2p_2
Total	P_1	P_2

We now examine the variance of $\hat{p}$ under the three possible methods for selection of the final sample from the preliminary sample.

3 Proportionate allocation

In this procedure, we draw, for the final sample, the same proportion of people from Stratum 1 as from Stratum 1 as from Stratum 2.

$$\left.\begin{aligned} n_1 &= n\hat{P}_1 \\ n_2 &= n\hat{P}_2 \end{aligned}\right\}. \tag{2}$$

When the results are in, we form by eqn. (1) the estimate $\hat{p}$ of p. For this plan

$$\text{var } \hat{p} = \frac{\sigma_b^2}{N'} + \frac{\sigma_w^2}{n}. \tag{3}$$

The optimum relation between n and N' for proportionate allocation is

$$\text{opt } n/N' = \frac{\sigma_w}{\sigma_b}\sqrt{\frac{c_1}{c_2}}. \tag{4}$$

4 Neyman allocation

Here, we aim at the allocation

$$\left.\begin{aligned} n_1 &= n\sigma_1 P_1/\bar{\sigma}_w \\ n_2 &= n\sigma_2 P_2/\bar{\sigma}_w \end{aligned}\right\}. \tag{5}$$

Once the study is completed, we again use eqn. (1) to form $\hat{p}$, for which

$$\text{var } \hat{p} = \frac{\sigma_b^2}{N'} + \frac{(\bar{\sigma}_w)^2}{n}. \tag{6}$$

The optimum relation between n and N' in Neyman allocation is

$$\text{opt } n/N' = \frac{\bar{\sigma}_w}{\sigma_b}\sqrt{\frac{c_1}{c_2}}. \tag{7}$$

Any non-zero sizes of sample n_1 and n_2 for the psychiatrist, when used in eqn. (1), will give an unbiased estimate of p. One of our aims here, however, is to find the optimum relationship between n_1, n_2 and N'. This we do by use of eqns. (2), (4), (5) and (7). Use of samples other than the optimum indicated in the tables would yield less information per unit cost than the optimum sizes will yield.

We introduce now specific numbers for our calculations. We use $q_2 = 0.1$, and choose a few values of p and p_1. For costs, we set $c_1 = \$5$ and $c_2 = \$45$; then $c_1 : c_2 = 1/9$ and by eqns. (3) and (6),

$$\text{Proportionate sampling} \quad \text{opt}\,\frac{n}{N'} = \frac{\sigma_w}{\sigma_b}\sqrt{\tfrac{1}{9}} = \tfrac{1}{3}\frac{\sigma_w}{\sigma_b} \tag{8}$$

$$\text{Neyman allocation} \quad \text{opt}\,\frac{n}{N'} = \tfrac{1}{3}\frac{\bar{\sigma}_w}{\sigma_b}. \tag{9}$$

The total cost of Steps 1 and 2 will be

$$K = 5N' + 45n = n\,(5N'/n + 45) \tag{10}$$

wherein n and N' are specific to the plan adopted.

The amount of information in an estimate $\hat{p}$ was defined by Sir Ronald Fisher as

$$I = 1/\text{var}\ \hat{p}. \tag{11}$$

The efficiency of the procedure that delivers the estimate $\hat{p}$ was defined by Morris Hansen as

$$I/K = \frac{1}{K\ \text{var}\ \hat{p}}, \tag{12}$$

which is the amount of information per unit cost. For no screening at all,

$$\text{var}\ \hat{p} = \frac{\sigma^2}{n} = \frac{p\,(1-p)}{n} \tag{13}$$

$$K = \$45n$$

$$I/K = \frac{1}{45p\,(1-p)}. \tag{14}$$

All the above formulas are in any book in statistical theory.

We should emphasize that it is the ratio $c_1 : c_2$ and not the absolute costs c_1 and c_2 that are important for the relationship between n and N'. Moreover, as the ratio $c_1 : c_2$ appears only under the square-root sign, the relationship between n and N' is not very sensitive to the costs within a moderate range of $c_1 : c_2$.

Costs in absolute numbers are necessary in order to compare the efficiency of proportionate allocation or Neyman allocation with a plan that uses no screening at all. Costs in absolute numbers are also necessary for prediction of the total cost of a study, once the plan is decided.

The calculations are shown in Tables 1 and 2. The important lines in the tables are lines 15, 21 and 24, which compare the amount of information per unit cost for the three plans under consideration. The tables show also the optimum relationships between the sizes of the samples N', n_1, n_2 under proportionate allocation and Neyman allocation.

5 Conclusion from the calculations

Comparisons of I/K, the amount of information per unit cost for the three plans – screening with proportionate allocation for the final sample; screening with Neyman allocation for the

final sample; no screening at all – leads to the following conclusions, which are valid over moderate intervals above and below the proportions and costs that the calculations were based on.

1. The proportion p_1 of false negatives in Stratum 1 is critical in consideration of choice of plan.

2. Screening is most effective in the reduction of costs when the overall proportion p of the disease under investigation is low, and when the screening is highly successful in the separation of cases, leaving the proportion p_1 of psychiatric cases very low in Stratum 1, while holding the proportion q_2 of non-psychiatric cases to a moderately low level in Stratum 2.

3. There is little choice between proportionate allocation and Neyman allocation in drawing the final sample from the preliminary sample, unless the screening finds 85 per cent or more of the psychiatric cases and places them in Stratum 1. Neyman allocation creeps ahead of proportionate allocation as the screening improves beyond this point.

4. Lines 11, 12, 17 and 18 in the tables show that for best efficiency (optimum balance) most of the cases for the final interviews will come from Stratum 1 unless the screening is extremely effective. This seems reasonable on reflection, because Stratum 1, intended to be pure, no psychopathology, will otherwise contain far more psychopathology than Stratum 2.

6 Remarks

Calculations of the kind shown here indicate that screening (or two-phase sampling) will not pay off under any usual circumstances unless the costs c_1 and c_2 are widely different. Specifically, 1 : 6 for the ratio $c_1 : c_2$ is usually a rough break-even point. The ratio $c_1 : c_2$ is likely to be especially low when the screening and classification are to be done on the basis of records, and where the final investigation may require costly field-work or costly interviews. In some of my own experience with screening and stratification carried out by perusing records on hand, the ratio $c_1 : c_2$ has run in the neighborhood of 1 : 40, or even 1 : 100.

Use of random numbers one after another to place people into Stratum 1 or into Stratum 2 would be equivalent to no screening at all, as the expected proportion of psychopathology would then be p in each stratum. It is not enough that the screening merely be better than random numbers. The tables indicate that there is no economy to realize from screening unless it be sufficiently effective to render $p_1 \leqq \frac{1}{4}p$. The proportion q_2 of false positives in Stratum 2 deserves reasonable care, but under conditions in any way similar to those studied here, q_2 is nowhere near as critical as the proportion p_1 of false negatives in Stratum 1.

Neyman allocation should not be attempted without calculating in advance the possible loss of efficiency from use of a predicted value of p_1 that turns out to be wide of the mark. There is a lot of leeway, but only within limits. A sample designed for Neyman allocation by use of a value of p_1 that turns out to be wide of the mark may end up with greater variance than proportionate allocation. Proportionate allocation is foolproof and simple to apply (Deming 1960, Hasel 1954). Unfortunately, the more reliable be the screening, the more difficult it is to measure how good it is.

It is easy to fall into a trap in the planning-stages by putting unwarranted credence into an advance estimate of p_1. A large experiment, or a long history of usage of the exact plan of screening, is required. An experiment, for example, wherein 30 persons were screened and all interviewed by a psychiatrist may not furnish enough information for a rational decision on whether to use screening, unless the correspondence is nigh perfect. This is so because an estimate of a small proportion like p_1 is subject to a wide standard error. Unfortunately, an experiment that requires 30 interviews is a heavy work-load for the psychiatrist.

The results of screening, as with any interviews, depend heavily on the questions and on the techniques of asking the questions. A few interviews conducted with one method of screening,

Table 1

Results of calculation, where $p = 0.1$, $\sigma^2 = 0.09$, $q_2 = 0.1$, $\sigma_2^2 = p_2q_2 = 0.09$ and $\sigma_2 = 0.3$

Item		p_1					
		0.15	0.10	0.05	0.025	0.01	0.005
1	$P_1 = (p_2-p)/(p_2-p_1)$	Not applicable		0.9412	0.9143	0.8989	0.8938
2	$P_2 = 1-P_1$			0.0588	0.0857	0.1011	0.1062
3	$\sigma_1^2 = p_1q_1$			0.0475	0.0244	0.0099	0.0050
4	σ_1			0.2179	0.1561	0.0995	0.0701
5	$\sigma_w^2 = P_1\sigma_1^2+P_2\sigma_2^2$			0.0500	0.0300	0.0180	0.0140
6	σ_w			0.2236	0.1732	0.1342	0.1183
7	$\bar{\sigma}_w = P_1\sigma_1+P_2\sigma_2$			0.2228	0.1685	0.1198	0.0949
8	σ_b^2			0.0400	0.0600	0.0720	0.0760
9	σ_b			0.2000	0.2449	0.2683	0.2757
Proportionate allocation							
10	$N' = n(\sigma_w/\sigma_b)\sqrt{c_1 : c_2}$			2.683n	4.242n	6.000n	6.990n
11	$n_1 = nP_1$			0.941n	0.914n	0.899n	0.894n
12	$n_2 = nP_2$			0.059n	0.086n	0.101n	0.106n
13	var $\hat{p}$			0.0649/n	0.0441/n	0.0300/n	0.0249/n
14	K, total cost			58.42n	66.21n	75.00n	79.95n
15	$I/K = 1/K$ var $\hat{p}$			0.264	0.342	0.444	0.503
Neyman allocation							
16	$N' = n(\bar{\sigma}_w/\sigma_b)\sqrt{c_1 : c_2}$			2.693n	4.362n	6.721n	8.716n
17	$n_1 = n\sigma_1P_1/\bar{\sigma}_w$			0.921n	0.847n	0.747n	0.660n
18	$n_2 = n\sigma_2P_2/\bar{\sigma}_w$			0.079n	0.153n	0.253n	0.336n
19	var $\hat{p}$			0.0645/n	0.0421/n	0.0251/n	0.0177/n
20	K, total cost			58.46n	66.81n	78.61n	88.58n
21	$I/K = 1/K$ var $\hat{p}$			0.265	0.355	0.508	0.637
No screening							
22	var $\hat{p} = \sigma^2/n = p(1-p)/n$	$= 0.09/n$					
23	K	$= 45n$					
24	$I/K = 1/K$ var $\hat{p}$	$= 0.247$					

Table 2

Results of calculations, where $p = 0.2$, $\sigma^2 = 0.16$, $q_2 = 0.1$, $\sigma_2^2 = p_2q_2 = 0.09$ *and* $\sigma_2 = 0.3$

Item		p_1					
		0.15	0.10	0.05	0.025	0.01	0.005
1	$P_1 = (p_2-p)/(p_2-p_1)$	0.9333	0.8750	0.8235	0.8000	0.7865	0.7821
2	$P_2 = 1-P_1$	0.0667	0.1250	0.1765	0.2000	0.2135	0.2179
3	$\sigma_1^2 = p_1q_1$	0.1275	0.0900	0.0475	0.0244	0.0099	0.0050
4	σ_1	0.3571	0.3000	0.2179	0.1561	0.0995	0.0705
5	$\sigma_w^2 = P_1\sigma_1^2+P_2\sigma_2^2$	0.1250	0.0900	0.0550	0.0375	0.0270	0.0235
6	σ_w	0.3536	0.3000	0.2345	0.1936	0.1643	0.1533
7	$\bar{\sigma}_w = P_1\sigma_1+P_2\sigma_2$	0.3533	0.3000	0.2324	0.1849	0.1423	0.1205
8	σ_b^2	0.0350	0.0700	0.1050	0.1225	0.1330	0.1365
9	σ_b	0.1871	0.2646	0.3241	0.3500	0.3647	0.3695
Proportionate allocation							
10	$N' = n(\sigma_w/\sigma_b)\sqrt{c_1 : c_2}$	1.587n	2.646n	4.146n	5.422n	6.659n	7.231n
11	$n_1 = nP_1$	0.933n	0.875n	0.823n	0.800n	0.786n	0.782n
12	$n_2 = nP_2$	0.067n	0.125n	0.177n	0.200n	0.214n	0.218n
13	var $\hat{p}$	0.147/n	0.116/n	0.080/n	0.060/n	0.047/n	0.042/n
14	K, total cost	52.94n	58.23n	65.73n	72.11n	78.30n	81.15n
15	$I/K = 1/K$ var $\hat{p}$	0.128	0.147	0.189	0.231	0.272	0.291
Neyman allocation							
16	$N' = n(\bar{\sigma}_w/\sigma_b)\sqrt{c_1 : c_2}$	1.589n	2.649n	4.184n	5.679n	7.689n	9.199n
17	$n_1 = n\sigma_1P_1/\bar{\sigma}_w$	0.943n	0.875n	0.772n	0.675n	0.550n	0.458n
18	$n_2 = n\sigma_2P_2/\bar{\sigma}_w$	0.057n	0.125n	0.228n	0.324n	0.450n	0.542n
19	var $\hat{p}$	0.147/n	0.116/n	0.079/n	0.056/n	0.038/n	0.029/n
20	K, total cost	52.94n	58.23n	65.92n	73.39n	83.44n	91.00n
21	$I/K = 1/K$ var $\hat{p}$	0.129	0.147	0.192	0.244	0.319	0.374
No screening							
22	var $\hat{p} = \sigma^2/n = p(1-p)/n$	$= 0.16/n$					
23	K	$= 45n$					
24	$I/K = 1/K$ var $\hat{p}$	$= 0.139$					

and a few interviews conducted with another method of screening, do not combine statistically to form a base for advance knowledge of what value of p_1 to plan on.

Interpretation of the results of an experiment on screening carried out by a small group of workers is sometimes difficult because the inevitable scientific cross-talk between the people that do the screening, and the psychiatrists, will exert day by day strong influence on everybody involved, with constant changes in the procedures. In the end, wisdom accumulates, but at the expense of a valid estimate of p_1.

General impressions about screening can be a hazard on a par with scanty figures.

In the end, when the results are in, it is then possible to compute by hindsight estimates of all the proportions involved, including $\hat{p}$ and its standard error.

In the absence of elaborate measures of quality control, the contribution to var $\hat{p}$ from the variances between investigators (interviewers and psychiatrists) are sure to be shocking. These variances can be evaluated by allotting a random portion of the whole area or of a subportion thereof to each interviewer and to each psychiatrist (Deming, 1972; Tepping and Bailar, 1968).

One might summarize the conclusions from the equations and the tables by saying that, in the absence of sound information about the screening and a clear indication that proportionate or Neyman allocation would pay off, it is perhaps best to use no screening at all. If screening be adopted, it is best to use proportionate allocation unless there be a firm basis for Neyman allocation.

In addition to the guidance supplied by the equations, there are some arguments to bear in mind about screening that are not expressible mathematically. Some of these arguments are negative on screening; some are positive. I may remind the reader first on the negative side that use of screening (unless the screening be carried out on the basis of records collected in a previous study) requires a second interview (the one by the psychiatrist) of the people that are selected into the final sample. There is always the possibility that this second appointment may encounter resistance and loss of the psychiatric interview. This means a total loss of the case, except for information of secondary importance that was already elicited in the screening interview. The loss from refusals at the second interview undoubtedly varies widely between communities, and with public interest in respect to the disease under investigation. Resistance may be serious in one place and not in another.

A further negative point to bear in mind is that screening necessitates some extra administrative attention in the field-work. Besides, with screening, there is the selection of the final sample to accomplish. These costs are in the equations, supposedly incorporated in the symbol c_1, but the equations do not take care of the circumstance in which the organization is small and overworked, with no one to take on with diligence the extra duties involved.

On the positive side of the ledger, also not in the equations, is the insight that a preliminary sample yields about the material in the frame, and about the problems that one will encounter in the investigation. A fairly large preliminary sample, even though the equations do not indicate any economical advantage of screening, puts before the investigator a miniature display of the frame. One will often find in this display problems that no one could otherwise foresee. It may bring out, for example, cases that do not belong in the investigation at all. It may bring out the existence of difficult cases. It may indicate errors in the delineation of sampling units, and need of more care in preparation. In a preliminary sample of hospital records intended for an investigation of adults 21 to 60, the preliminary sample contained admissions of age 20 and under. There were also cases beyond the intended age-limit, and emergency cases of various kinds not intended for investigation. Some cases were transfers from other hospitals, and would require requests for additional notes. Without the preliminary sample, the investigators would have had no warning that 15 per cent of the frame was made up of a spectrum of blanks to be discarded, nor that 10 per cent of the frame came from transfers. A fairly large preliminary sample, screened by use of the case-notes, permits one to

throw out the blanks and to stratify the valid cases by type of ailment indicated. The final sample for further study may then be balanced in the main categories of hospital diagnosis (Kuriansky, Deming and Gurland, 1974).

Sampling to measure the prevalence of a rare characteristic is a subject all by itself, beyond the scope of this paper, and must be dismissed here with the statement that for a characteristic that has a high probability of being treated in an institution, samples might be taken from clinics and hospitals, accompanied by a sample from the general population. Statistical procedures to determine with a prescribed probability that the prevalence of a certain rare disease does not exceed some small proportion such as $p \leqq 1/50$ call forth still further theory, also not to be covered here.

This paper should also mention circumstances that often face investigators in small research organizations where there is a shortage of psychiatrists or of men with other specialized knowledge. It is then imperative to carry out screening. In fact, the optimum plan in such circumstances may be to use a preliminary sample that is double or treble the size that the equations indicate; then to adjust the sizes n_1 and n_2 of the final sample by proportionate allocation or by Neyman allocation, holding n_1+n_2 to the maximum number that the psychiatrists can handle. The information per unit cost will be less than that indicated by the optimum ratio of N' to n, but it will be valid statistical information bought at the lowest price consistent with the restraints.

7 Note in respect to the tables

The symbol n in Tables 1 and 2 has a different meaning from one panel to another. Thus, the number n required to reach a given precision with proportionate allocation would not be the same number required to reach the same precision with Nayman allocation, or with no screening at all. Comparison between plans is possible only in lines 15, 21 and 24 which show the efficiency I/K of the plans, wherein n does not appear.

Acknowledgements

I am indebted to my friend and colleague Dr Morris H. Hansen for many important suggestions on this paper, and for the good fortune to work with him on several studies that have required use and extension of the theory presented here. It has been my good fortune to work with Dr Barry Gurland of the New York State Psychiatric Institute, under whose direction a number of studies have required consideration of screening and allocation of sample.

References

Deming, W.E., 1960. *Sample design in Business Research.* New York: Wiley, p. 295.

Deming, W.E., 1972. Some theory on the influence of the inspector and environmental conditions. *Statistica Neerlandica*, **26/3**, 101–112.

Hasel, A.A. (1954). Problems in inventory, Chapter 19 in O. Kempthorne (editor), *Statistics and Mathematics in Biology*. Iowa State College Press, p. 267.

Kish, Leslie (1965). *Sampling Techniques.* John Wiley, p. 407.

Kuriansky, J., Deming, W.E. and Gurland, B. (1974). On trends in the diagnosis of schizophrenia. *American Journal of Psychiatry*, **4**, 402–408.

Tepping, B.J. and Bailar, B.A. (1968). Effects of interviewers and crew leaders. Series ER No. 7, Bureau of the Census, Washington.

Résumé

L'application de cette thésis est à l'édude de psychopathologie des habitants d'une ville, age (e.g.) de 65 ans ou meilleur. Un dessein pour utilisation les services des interviewers versés spécialement pour de mettre à l'épreuver et sortir en avant entre deux classes (avec ou sans psychopathologie, à luis avis) chaque personne d'un grand

échantion, afin que d'conserver le temps de psychiatriste, par presentation à lui rien que les personnes qui sont presque sans contredit psychopathologiques, est attrayant. Il est nécessaire, néanmois, fin qu'obtenir un estimateur sans biais, que le psychiatriste emanine un échantillon de chaque classe. De plus, il est nécessaire que l'épreuve, fait en avant, (*a*) soit à bon marché, et (*b*) admette seul en trés petite proportion des faux négatifs. La proportion des faux positifs n'est pas si importante. L'auteur illustre ces principes par calculations avec variés proportions des cas psychopathologiques verifiés.