SCB STATISTISKA CENTRALBYRÅN 1983-02-14

PROMEMORIOR FRÅN P/STM

NR 8

HOW LARGE MUST THE SAMPLE SIZE BE? NOMINAL CONFIDENCE LEVELS VERSUS ACTUAL COVERAGE PROBABILITIES IN SIMPLE RANDOM SAMPLING

AV JÖRGEN DALÉN

INLEDNING

TILL

Promemorior från P/STM / Statistiska centralbyrån. – Stockholm : Statistiska centralbyrån, 1978-1986. – Nr 1-24.

Efterföljare:

Promemorior från U/STM / Statistiska centralbyrån. – Stockholm : Statistiska centralbyrån, 1986. – Nr 25-28.

R & D report : research, methods, development, U/STM / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1987. – Nr 29-41.

R & D report : research, methods, development / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1988-2004. – Nr. 1988:1-2004:2.

Research and development : methodology reports from Statistics Sweden. – Stockholm : Statistiska centralbyrån. – 2006-. – Nr 2006:1-.

Promemorior från P/STM 1983:8. How large must the sample size be? Nominal confidence levels versus actual coverage probabilities in simple random sampling / Jörgen Dalén.
Digitaliserad av Statistiska centralbyrån (SCB) 2016.

urn:nbn:se:scb-PM-PSTM-1983-8

SCB STATISTISKA CENTRALBYRÅN

1983-02-14

PROMEMORIOR FRÅN P/STM

NR 8

HOW LARGE MUST THE SAMPLE SIZE BE? NOMINAL CONFIDENCE LEVELS VERSUS ACTUAL COVERAGE PROBABILITIES IN SIMPLE RANDOM SAMPLING

AV JÖRGEN DALÉN

C O N T E N T S

PREFACE

In survey sampling the prevailing estimation strategy for estimating a finite population parameter θ is based upon an (approximately) unbiased point estimator $\hat{\theta}$ and an (approximately) unbiased variance estimator $\hat{V}(\hat{\theta})$. Then the central limit theorem is referred to for an assumption that $\hat{\theta}$ is approximately normally distributed and it is stated that the interval

$$\hat{\theta} \pm 1.96 \times \sqrt{\hat{V}(\hat{\theta})}$$

covers the true value θ with a probability of approximately 95 %. Sometimes the value 1.96 is exchanged for the corresponding value taken from Student's t table with an appropriate number of degrees of freedom.

However, the central limit theorem does not say that this or that sample size is sufficient for the normal approximation to be reasonable in sampling from a particular finite population. For this reason much effort has been spent on the problem of convergence rates of sampling distributions to the normal distribution for simple random sampling from finite populations. But so far the attempts to reach exact theoretical results in this respect have more or less failed. This fact points at the necessity to seek other directions to approach this problem.

The non-applicability of the normal approximation is a problem that occurs frequently in sampling practice. One example is sampling from very skewed populations such as variables reported by enterprises (production, employment, investment, export, import). Another example is small area estimation where only a small number of observations form the basis of a certain figure in a table.

For this reason a work has been initiated in the Statistical Methods Unit of Statistics Sweden concerning sampling from very skewed populations. One line of research has been to try to establish rules for when the normal approximation is applicable by making empirical studies of the distributional behaviour of estimators. In this paper such rules based on the population skewness are presented for the simple random sampling case.

The work is a part of a larger project within the Methods Unit concerning sampling and estimation led by Bengt Swensson. He, Carl-Erik Särndal and Jan Hagberg have also contributed to this report by giving helpful comments and suggestions.

1 Introduction

Erdös and Rényi (1959) and Hájek (1960) have developed conditions in finite populations for the sampling distribution of the sample mean to converge to normality. Höglund (1978) has made the following remainder term estimate (the formula is slightly manipulated algebraically to serve our purpose):

$$\left| F(x) - \Phi\left(\frac{x-n\mu}{\sigma\sqrt{n(1-f)}}\right) \right| \leq \frac{C}{\sqrt{n(1-f)}} \cdot G_H, \text{ where}$$

F is the distribution function of the sum of a sample of n elements among the N population elements $(X_1 \ldots X_j \ldots X_N)$,

Φ the standard normal distribution function,

μ is the population mean,

σ is the population standard deviation,

$f = n/N$,

C is an absolute constant and

$$G_H = \frac{\frac{1}{N} \sum_{j=1}^{N} |X_j - \mu|^3}{\sigma^3}$$

We notice that the deviation from normality is bounded by a term containing the factor G_H, which may be considered to be a measure of the population skewness.

"For populations in which the principal deviation from normality consists of marked positive skewness", Cochran (1977) suggests the simple rule

$$n > 25\ G_C^2, \text{ where}$$

$$G_C = \frac{\frac{1}{N} \cdot \sum_{j=1}^{N} (X_j - \mu)^3}{\sigma^3}$$

"This rule is designed so that a 95 % confidence probability statement will be wrong not more than 6 % of the time." We notice that the rule includes G_c - a different measure of skewness (in fact the one most commonly used).

In this paper confidence intervals based on samples from the dichotomous population are systematically studied with respect to different population sizes, sample sizes and degrees of skewness. The reason for choosing the dichotomous population is that is is possible to obtain exact coverage probabilities for it. Also, the whole scale of different degrees of skewness is represented for a certain population size of the dichotomous population.

Exact probabilities are calculated for nominal 95 % confidence intervals based on the Student's t approximation to cover the true population mean. The t approximation is chosen instead of the normal one because it can be expected to give better approximations for small sample sizes and the 95 %-level because it is the most common one in practical survey sampling. Properties of these exact probabilities are demonstrated and simple rules of the Cochran type are proposed. The rules are formulated as $n > K_\alpha \cdot G^2$, where K_α is the constant required for obtaining at least α % probability for a supposed 95 % confidence interval to be true. The effects of using G_H as compared to G_C for G are studied. The rules are also found to work well for finite populations generated from some continuous parametric distributions.

2 The dichotomous population

For the dichotomous population studied the following notations are used:

Value	Number of elements in the population	Number of elements in the sample
0	N - M	n - k
1	M	k
Total	N	n

The population has the following characteristics:

Population mean = μ = P

Population variance = $\sigma^2 = P - P^2$

Skewness (Cochrans) = $G_C = (1 - 2P)/(P - P^2)^{0.5}$

Skewness (Höglunds) = $G_H = (1 - 2P + 2P^2)/(P - P^2)^{0.5}$

where P = M/N.

Notice that $G_H = G_C + 2P^{1.5}/(1 - P)^{0.5}$ so that $\lim_{P \to 0} (G_H - G_C) = 0$. Notice also that $G_c = 0$ and $G_H = 1$ when P = 0.5.

The sample has the following characteristics:

sample mean = $\overline{X} = k/n$

sample variance = $s^2 = (k-k^2/n)/(n-1)$

A supposed 95 % confidence interval for μ based on the sample outcome and the t-distribution would now be

$$\overline{X} - t_{0.975} \cdot s \cdot \sqrt{(1-f)/n} < \mu < \overline{X} + t_{0.975} \cdot s \cdot \sqrt{(1-f)/n}$$

(Since we are interested in how bad the t approximation could be at worst, the continuity correction is not used.)

Now, let I_k be the indicator for this confidence interval statement as a function of the sample outcome. That is:

$$I_k = \begin{cases} 1 \text{ for those k when the confidence interval statement is true} \\ 0 \text{ for those k when the confidence interval statement is false} \end{cases}$$

The actual coverage probability (ACP) is now defined as the probability for a sample of a certain size n from our population to produce a true confidence interval statement. In mathematical notation this becomes

$$ACP(N, M, n) = \sum_{k=0}^{n} p(k) \cdot I_k$$

$$\text{where} \quad p(k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

according to the hypergeometrical distribution.

Computer programs have been written, which compute these probabilities for various combinations of N, M and n. In table 1 one example of a result from such a program is shown.

There are many interesting features in this table. For example, there is not a steady increase in the ACP for increasing sample sizes. There are two reasons for this. One is due to the fact that the sampling distribution does not converge to the t-distribution when n approaches N. Actually, as pointed out by Plane and Gordon (1982), when $n > N/2$ the sampling distribution becomes more and more dissimilar to the normal distribution[1)].

The other reason for the jumps in the ACP as a function of sample size is understood when you look at the column "Outcomes where $I_k = 1$". There is, for example, a big downward jump from n=29 (ACP=95.6) to n=30 (ACP=82.6). This jump is due to the fact that the outcome k=1, which has a large probability to occur, for n=29 leads to a confidence interval, which just barely covers μ (the interval is from -0.0327 to 0.1016) but for n=30 the upper limit of the interval falls a little bit short of μ (it goes from -0.0313 to 0.0980).

In the example presented in table 1 you can also see that - due to the fluctuations of the ACP - you can never obtain a stable ACP-level of more than about 92 % by increasing the sample size (up to N/2).

---

1) Plane and Gordon prove that the sampling distributions for n and N-n are mirror images of each other except for a scale change. However, as could be seen in Table 1, this does not mean that the ACP is the same for n and N-n. The scale factor is obviously important here.

3 Guaranteed ACPS

A lot of tables like table 1 have been produced but for the sake of space they are not presented here. However, in table 2, we have calculated Cochran type constants leading to a certain guaranteed ACP. The procedure is as follows.

In deciding when a guaranteed ACP α is obtained we look for the smallest value of n (called n_α) such that $ACP(n) \geq \alpha$ for all $n \geq n_\alpha$. (For reasons mentioned above only $n \leq N/2$ are considered.) Then we calculate $K_{\alpha C} = n_\alpha/G_C^2$ and $K_{\alpha H} = n_\alpha/G_H^2$ whereupon we can obviously say that for all $n > K_\alpha \cdot G^2$ (indicated C or H) we have $ACP(n) \geq \alpha$.

For example look for a guaranteed ACP of 85 % in table 1. We see that $n_{0.85} = 32$, $K_{0.85C} = 32/2.67^2 = 4.5$ and $K_{0.85H} = 32/2.73^2 = 4.3$. These values could be found in the row in table 2 indicated N=300, M=30.

These calculations have been done for N=100 (100) 1 000, M=10, 20, 30, 40, 50, 75 and 100 (except for N=100 where M=5 (5) 50 is used), $2 \leq n \leq N/2$ and for α = 85 %, 90 %, 93 %, 94 % and 94.5 %. The results are presented in table 2.

As we calculate $K_\alpha = n_\alpha/G^2$ and consider only $n \leq N/2$, we realize that for one particular population you can never obtain larger values of K_α than $N/2G^2$. The last two columns of table 2 and 3 present the value of this limit. Values of K_α close to the limit (80 % of it has been chosen as a simple rule) are put within brackets, since the constants obtained in these cases may be merely accidental from the point of view of all possible populations.

Some striking features in table 2 are

- Cochrans measure of skewness (G_C) is not at all suited to establishing criteria for populations with almost symmetrical distributions. Since $G_C = 0$ for symmetrical distributions $K_{\alpha C} \to \infty$ as $G_C \to 0$. When $P > 0.25$ it seems advisable not to use G_C any longer. G_H, on the other hand, is useful for all degrees of population skewness. When the skewness is large G_C and G_H are almost identical.

- Guaranteed ACPs of 94 % or more seem to be very difficult to obtain for these populations. $K_{\alpha C} = 25$, as suggested by Cochran, is clearly not enough. (We recognize, of course, that Cochran's rule is not intended to apply to dichotomous populations. For the equivalent case of proportions he suggests other criteria.) For $\alpha = 85$ % $K_{\alpha H} = 5$ seems to be sufficient in most cases and $K_{\alpha H} = 6$ in all cases. For $\alpha = 90$ % the corresponding figures would be $K_{\alpha H} = 10$ or 11. For $\alpha = 93$ % it is already more difficult to pass a judgement but it seems as if $K_{\alpha H} = 40$ would be sufficient.

4 Average ACPS

Instead of dealing with guaranteed ACPs we may define another, more "liberal", concept, called the average ACP, in the following way.

β is an average ACP for a certain sample size n_0 in a certain population (M, N), if $ACP(n_0) \geq \beta$ and

$$\frac{\sum_{j=0}^{s} ACP(n_0+j)}{s+1} \geq \beta$$

for all values of s such that $0 \leq s \leq N/2 - n_0$.

In order to decide when an average ACP β is obtained we look for the smallest value of n_0 - called n_β - such that the above conditions are valid. In the same way as before we then calculate $K_{\beta C} = n_\beta/G_C^2$ and $K_{\beta H} = n_\beta/G_H^2$. We then state that for $n > K_\beta \cdot G^2$ (indicated C or H) there is an average ACP of at least β.

As an example we look for an 85 % average ACP in table 1. We see that ACP(18) > 85 %. There are only two ACPs smaller than 85 % for n > 18 (n = 30 and 31) and those are "averaged out" by the large ACPs surrounding them. We therefore certify that $n_{0.85} = 18$, $K_{0.85C} = 18/2.67^2 = 2.5$ and $K_{0.85H} = 18/2.73^2 = 2.4$. These values could be found in the row in table 3 indicated N=300, M=30.

What is the reason for bringing in the average ACP concept? It is obviously in order to even out the jumps in the ACP variation. In real life the exact population distribution is not known but you have an indication of the skewness from your sample. It is then intuitively reasonable to think of some kind of "expected ACP". Also, for larger and more diversified populations than the dichotomous one studied here, the jumps will be smaller and average ACP will be closer to the actual one.

The calculations based on the average ACP concept have been done for the same values of N, M, n and β (α) as was the case for the guaranteed ACPs. The results are presented in table 3.

What are the conclusions which could be drawn from this table?

- The relations between G_C and G_H is the same as in the preceding section.

- We generally get much lower constants than in the guaranteed ACP case. We also get useful results for the high β-levels. We see that a $K_{\beta H}$ of 3-4 (2-3 for very skew populations) seems to be sufficient for 85 % average ACP. For 90 % $K_{\beta H} = 5$ seems to be enough, for 93 % 10-12, for 94 % around 20, while for 94.5 % it is more difficult to make a definite statement but $K_{\beta H}$ = 35-40 seems to suffice.

5 Large populations

The populations studied above are comparatively small. N ranges from 100 up to 1 000. In certain instances there is a tendency of the constants K to rise with rising population size. It is therefore interesting to see what happens in much larger populations.

The rapid increase in computer time, however, preclude systematic investigations like those above for large populations. But it is possible to calculate ACPs for a limited selection of sample sizes. Tables 4-6 present some results of this kind.

In tables 4 and 5 some populations with different degrees of skewness are studied for some sample sizes. The population size is 10 000 in table 4 and 100 000 in table 5. For some sample sizes the corresponding value of the constant ($K_H = n/G_H^2$) is also given. Because of the scattered sample sizes no definite conclusions can be drawn from these tables but there is nothing in them contradicting the above conclusions which are based on smaller population sizes.

Table 6 presents a population with a size of eight millions with one percent ones and 99 percent zeroes. (This happens to be about the largest population of interest when sampling in Sweden.) ACPs are calculated for sample sizes from 50 to 6 000 with intervals of 50. $K_H \approx 3$ for n=300, 5 for n=500, 11 for n=1 050, 20 for n=1 950 and 40 for n=3 900. Also in this case, the conclusions based on the small populations seem to be approximately valid.

6 Recommended values of K

The investigations presented so far could be summarized in the following way. In order to ensure at least a certain actual coverage probability (ACP), when sampling from a dichotomous finite population and using a t approximation for a 95 % confidence interval, the sample size n must be larger than $K \cdot G_H^2$. The value of the constant K needed is a function of the desired ACP. It also depends upon the "average" or "guaranteed" interpretation of the ACP in about the following way:

	Value of K for ACP =				
	85	90	93	94	94.5
Guaranteed	5-6	10-11	(40)	?	?
Average	2-4	5	10-12	20	40

We immediately realize that it is very difficult to formulate any rules with a "guaranteed" interpretation attached to them. But with the "average" interpretation this becomes possible and the K-levels get practical significance. The 94 %-rules is also close to the rule suggested by Cochran.

In the next section we investigate ACPs for other types of distributions based on K-levels of 3, 5, 11, 20 and 40.

7 Populations based on some continuous parametric distributions

In order to find out if the constants K could be applied to populations with structures entirely different from the dichotomous one, finite populations based on fixed percentiles of some continuous theoretical distributions were generated. These are:

I) The beta distribution with probability density function

$$f(x) = \frac{x^{a-1}\,(1-x)^{b-1}}{B(a,b)} \;;\; 0 \leq X \leq 1,\; a > 0,\; b > 0$$

where B(a,b) is the beta function.

$$\gamma = \text{coefficient of skewness} = \frac{2(b-a)(a+b+1)^{1/2}}{(a+b+2)(ab)^{1/2}}$$

II) The lognormal distribution with probability density function

$$f(x) = (\sigma x)^{-1} \cdot (2\pi)^{-1/2} \cdot \exp\left[-(\log x - \theta)^2/2\sigma^2\right];\quad x > 0,\; \sigma > 0$$

and

$$\gamma = \left[\exp(\sigma^2) + 2\right] \cdot \left[\exp(\sigma^2) - 1\right]^{1/2} \;;$$

III) The power function with distribution function

$$F(x) = (x/\theta)^c;\; 0 < x < \theta,\; \theta > 0,\; c > 0$$

and

$$\gamma = \frac{2(1-c)(2+c)^{1/2}}{(3+c) \cdot c^{1/2}}$$

IV) The Weibull distribution with distribution function

$$F(x) = 1 - \exp\left[-\theta x^c\right]; \; x \geq 0, \; \theta > 0, \; c > 0$$

$$\gamma = \frac{\Gamma(1+3/c) - 3\Gamma(1+2/c)\,\Gamma(1+1/c) + 2\Gamma^3(1+1/c)}{\left[\Gamma(1+2/c) - \Gamma^2(1+1/c)\right]^{3/2}}$$

where $\Gamma(x)$ is the gamma function.

In all cases γ stands for the coefficient of skewness corresponding to G_c above, that is

$$\gamma = E(x-\mu)^3/\left[E(x-\mu)^2\right]^{3/2}$$

The reference used for these distributions is Patel et al (1976).

A common feature of all these distributions is that for some value(-s) of the involved parameter(-s) γ can take on at least any value > 0.

For each of these four distributions six different finite populations have been generated with different degrees of skewness. The finite populations are all of size 500 and have been generated by taking the percentiles from 0.001 to 0.999 with intervals of 0.002.

For each population five different sample sizes have been chosen so that they correspond as closely as possible to the average K-levels above. This means that n has been chosen so that

$$n \geq K_{\beta H} \cdot G_H^2 > n - 1$$

for values of K of 3, 5, 11, 20 and 40 respectively.

For every sample size 1 000 Monte-Carlo-simulations of a simple random sample without replacement have been made. For each sample the population mean has been estimated and a confidence interval based on the sample standard deviation and the t-distribution has been calculated. The number of cases when this interval covers the true population mean has been counted. This figure divided by 1 000 becomes our estimated actual coverage probability (EACP). EACP is of course stochastic in this case with a standard error of 0.7 % to 1.1 % when EACP ranges from 95 % to 85 %.

In table 7 - 10 the outcome of these simulations is presented in terms of the EACP for a certain combination of population and sample size. We see that in almost all cases the β-levels expected from the studies of the dichotomous population are surpassed, often by large margins. Only in five cases are the presupposed ACP-levels not obtained (those cases are indicated with an asterisk). The EACPs are in these cases 0.1 - 0.4 % below the expected level. One case is for K=11 (0.1 below), three cases are for K=20 (0.1-0.3 below) and one case is for K=40 (0.4 below). These small deviations in our results are not enough to cast any serious doubts on the usefulness of our proposed rules as a whole. The deviations may very well be entirely due to stochastic effects.

However, we must also admit that our rules for the higher ACP-levels (β= 94 and 94.5 %) are not equally well founded compared with those for the lower levels (β= 85 and 90 %). A more solid foundation for these levels would require a study of larger populations and sample sizes and also more replications for each case, since we are so close to 95 % that the stochastic effect of the simulations becomes an important disturbance.

8 Conclusions

Our empirical investigations have supported the following rule for the use of the Student's t approximation (and thereby also the normal approximation for reasonable sample sizes) in simple random sampling.

If you apply the t approximation for a 95 % confidence interval and the sample size

$$n > K_\beta \cdot G_H^2 ,$$

you could be reasonably certain that the coverage probability is at least β. Compared with the rule suggested by Cochran a different measure of skewness is used. For populations with large positive skewness the two measures approximately coincide but for close-to-symmetrical populations it is necessary to use G_H instead of G_c.

Five values of K have been derived, namely

$K_{0.85} = 3$,
$K_{0.90} = 5$,
$K_{0.93} = 11$,
$K_{0.94} = 20$ and
$K_{0.945} = 40$.

$K_{0.94}$ and $K_{0.945}$ remain not equally well founded compared with the others.

As a principal rule for allowing the t approximation in simple random sampling for a 95 % confidence interval, we propose

$$n > 20 \cdot G_H^2$$

If this rule is applied you should be able to count on being correct about 94 % of the time.

The values of K_β are not guaranteed in the sense that no instances could be found, where they are not sufficient. But in the average sense defined above they generally hold good and they are therefore sound to use in practical work.

However, one warning is necessary to issue. The rules above are based upon knowledge of the population skewness. It is not sufficiently known to what extent estimates of this skewness from the sample could replace it. For example, the sample skewness is a consistent but biased estimate of the population skewness. The bias could be considerable precisely in those cases where you would want to apply the sample skewness in the rules above. Work in this area is going on at the present time and results will be published in a forthcoming report.

What, then, are the theoretical reasons for these rules? In this area more research is undoubtedly needed in the future. Here, let us only point out the fact that the relevant convergence theorems and remainder term estimates are stated in terms of the standardized third moment of the population distribution. Therefore there is a reason to believe that two populations with the same degree of skewness should not differ radically with respect to the degree of correspondence between their sampling distributions and the t-distribution.

9 References

Cochran, William G., (1977), Sampling Techniques, third edition, Wiley

Erdös, P & Rényi, A, (1959), On the central limit theorem for samples from a finite population. Publ. Math. Inst. Hung. Acad. Sci. 4, 49-57.

Hájek, J, (1960), Limiting distributions in simple random sampling from a finite population. Publ. Math. Inst. Hung. Acad. Sci. 5, 361-374.

Höglund, (1978), Sampling from a finite population. A remainder term estimate. Scand J Statist 5:69-71.

Patel, J.K., Kapadia, C. H. & Owen, D.B., (1976), Handbook of statistical distributions, Marcel Dekker, Inc.

Plane, D.R. & Gordon, K.R., (1982), A simple proof of the non-applicability of the central limit theorem to finite populations. The American statistician, Vol. 36, No. 3, Part 1.

Table 1: Actual coverage probability (ACP) for a dichotomous population. N=300. M=30. n=2 - 299. μ=0.1. G_C=2.67. G_H=2.73

Sample size	ACP	Outcomes (k) where I_k=1		Sample size	ACP	Outcomes (k) where I_k=1	
2	18.060	1 -	1	68	93.625	4 -	12
3	27.090	1 -	2	69	94.000	4 -	12
4	34.529	1 -	3	70	94.328	4 -	12
5	41.133	1 -	3	71	94.612	4 -	12
6	47.047	1 -	3	72	94.851	4 -	12
7	52.533	1 -	4	73	95.902	4 -	13
8	57.373	1 -	4	74	90.214	5 -	13
9	61.713	1 -	4	75	90.826	5 -	13
10	65.594	1 -	4	76	91.387	5 -	13
11	69.051	1 -	4	77	91.899	5 -	13
12	72.421	1 -	5	78	92.361	5 -	13
13	75.263	1 -	5	79	92.776	5 -	13
14	77.802	1 -	5	80	93.143	5 -	13
15	80.063	1 -	5	81	93.464	5 -	13
16	82.067	1 -	5	82	94.637	5 -	14
17	83.832	1 -	5	83	94.970	5 -	14
18	85.779	1 -	6	84	95.267	5 -	14
19	87.250	1 -	6	85	95.529	5 -	14
20	88.554	1 -	6	86	95.756	5 -	14
21	89.703	1 -	6	87	90.805	6 -	14
22	90.710	1 -	6	88	91.296	6 -	14
23	91.583	1 -	6	89	91.739	6 -	14
24	92.331	1 -	6	90	92.135	6 -	14
25	93.491	1 -	7	91	93.401	6 -	15
26	94.142	1 -	7	92	93.804	6 -	15
27	94.709	1 -	7	93	94.168	6 -	15
28	95.196	1 -	7	94	94.496	6 -	15
29	95.608	1 -	7	95	94.787	6 -	15
30	82.636	2 -	7	96	95.043	6 -	15
31	83.943	2 -	7	97	95.264	6 -	15
32	85.120	2 -	7	98	95.450	6 -	15
33	86.880	2 -	8	99	90.883	7 -	15
34	87.952	2 -	8	100	91.300	7 -	15
35	88.923	2 -	8				
36	89.798	2 -	8				
37	90.580	2 -	8				
38	91.273	2 -	8				
39	91.880	2 -	8	101	92.683	7 -	16
40	93.101	2 -	9	102	93.111	7 -	16
				103	93.500	7 -	16
				104	93.852	7 -	16
				105	94.167	7 -	16
41	93.667	2 -	9	106	94.446	7 -	16
42	94.170	2 -	9	107	94.689	7 -	16
43	94.611	2 -	9	108	94.897	7 -	16
44	94.994	2 -	9	109	95.070	7 -	16
45	95.321	2 -	9	110	96.197	7 -	17
46	86.543	3 -	9	111	92.108	8 -	17
47	87.401	3 -	9	112	92.555	8 -	17
48	88.936	3 -	10	113	92.965	8 -	17
49	89.742	3 -	10	114	93.336	8 -	17
50	90.479	3 -	10	115	93.671	8 -	17
51	91.147	3 -	10	116	93.969	8 -	17
52	91.749	3 -	10	117	94.231	8 -	17
53	92.287	3 -	10	118	94.457	8 -	17
54	92.763	3 -	10	119	94.648	8 -	17
55	93.177	3 -	10	120	95.845	8 -	18
56	94.320	3 -	11	121	96.067	8 -	18
57	94.729	3 -	11	122	92.133	9 -	18
58	95.091	3 -	11	123	92.559	9 -	18
59	95.409	3 -	11	124	92.947	9 -	18
60	95.683	3 -	11	125	93.298	9 -	18
61	89.082	4 -	11	126	93.613	9 -	18
62	89.737	4 -	11	127	93.891	9 -	18
63	90.328	4 -	11	128	94.133	9 -	18
64	90.856	4 -	11	129	94.340	9 -	18
65	92.213	4 -	12	130	95.585	9 -	19
66	92.732	4 -	12	131	95.822	9 -	19
67	93.203	4 -	12	132	96.032	9 -	19

Table 1 (cont.)

Sample size	ACP	Outcomes (k) where I_k=1	Sample size	ACP	Outcomes (k) where I_k=1
133	92.279	10 - 19	198	95.002	16 - 25
134	92.681	10 - 19	199	95.291	16 - 25
135	93.047	10 - 19	200	95.546	16 - 25
136	93.376	10 - 19	201	95.769	16 - 25
137	93.669	10 - 19	202	95.959	16 - 25
138	93.925	10 - 19	203	92.618	17 - 25
139	94.146	10 - 19	204	92.995	17 - 25
140	94.331	10 - 19	205	93.327	17 - 25
141	95.668	10 - 20	206	93.617	17 - 25
142	95.891	10 - 20	207	93.864	17 - 25
143	96.086	10 - 20	208	94.069	17 - 25
144	92.536	11 - 20	209	95.822	17 - 26
145	92.914	11 - 20	210	96.086	17 - 26
146	93.257	11 - 20	211	96.317	17 - 26
147	93.562	11 - 20	212	96.518	17 - 26
148	93.832	11 - 20	213	93.390	18 - 26
149	94.066	11 - 20	214	93.780	18 - 26
150	94.265	11 - 20	215	94.127	18 - 26
151*	94.428	11 - 20	216	94.432	18 - 26
152	95.838	11 - 21	217	94.696	18 - 26
153	96.046	11 - 21	218	94.917	18 - 26
154	92.508	12 - 21	219	95.098	18 - 26
155	92.898	12 - 21	220	95.236	18 - 26
156	93.252	12 - 21	221	93.340	19 - 27
157	93.570	12 - 21	222	93.824	19 - 27
158	93.852	12 - 21	223	94.267	19 - 27
159	94.099	12 - 21	224	94.667	19 - 27
160	94.310	12 - 21	225	95.026	19 - 27
161	94.486	12 - 21	226	95.345	19 - 27
162	95.873	12 - 22	227	95.623	19 - 27
163	96.090	12 - 22	228	95.861	19 - 27
164	96.281	12 - 22	229	96.060	19 - 27
165	92.995	13 - 22	230	92.652	20 - 27
166	93.360	13 - 22	231	93.064	20 - 27
167	93.690	13 - 22	232	93.423	20 - 27
168	93.983	13 - 22	233	93.729	20 - 27
169	94.242	13 - 22	234	93.982	20 - 27
170	94.465	13 - 22	235	95.992	20 - 28
171	94.654	13 - 22	236	96.318	20 - 28
172	94.808	13 - 22	237	96.607	20 - 28
173	96.217	13 - 23	238	96.858	20 - 28
174	96.417	13 - 23	239	93.579	21 - 28
175	93.202	14 - 23	240	94.066	21 - 28
176	93.578	14 - 23	241	94.501	21 - 28
177	93.917	14 - 23	242	94.885	21 - 28
178	94.222	14 - 23	243	95.218	21 - 28
179	94.491	14 - 23	244	95.501	21 - 28
180	94.726	14 - 23	245	95.734	21 - 28
181	94.927	14 - 23	246	95.918	21 - 28
182	95.093	14 - 23	247	92.367	22 - 28
183	95.226	14 - 23	248	92.794	22 - 28
184	91.791	15 - 23	249	93.153	22 - 28
185	93.516	15 - 24	250	95.534	22 - 29
186	93.900	15 - 24	251	95.980	22 - 29
187	94.249	15 - 24	252	96.379	22 - 29
188	94.564	15 - 24	253	96.731	22 - 29
189	94.844	15 - 24	254	97.038	22 - 29
190	95.089	15 - 24	255	93.470	23 - 29
191	95.301	15 - 24	256	94.074	23 - 29
192	95.480	15 - 24	257	94.614	23 - 29
193	95.625	15 - 24	258	95.090	23 - 29
194	92.328	16 - 24	259	95.502	23 - 29
195	92.646	16 - 24	260	95.851	23 - 29
196	94.322	16 - 25	261	96.139	23 - 29
197	94.679	16 - 25	262	96.364	23 - 29

* For n > 150 the normal approximation is used.

Table 1 (cont.)

Sample size	ACP	Outcomes (k) where I_k=1		
263	92.634	24	-	29
264	93.172	24	-	29
265	93.622	24	-	29
266	96.201	24	-	30
267	96.751	24	-	30
268	97.244	24	-	30
269	97.681	24	-	30
270	93.725	25	-	30
271	94.594	25	-	30
272	95.382	25	-	30
273	96.093	25	-	30
274	96.727	25	-	30
275	97.289	25	-	30
276	97.781	25	-	30
277	93.478	26	-	30
278	94.506	26	-	30
279	95.431	26	-	30
280	96.253	26	-	30
281	96.975	26	-	30
282	97.601	26	-	30
283	92.321	27	-	30
284	93.689	27	-	30
285	94.913	27	-	30
286	95.990	27	-	30
287	96.920	27	-	30
288	97.708	27	-	30
289	91.409	28	-	30
290	93.304	28	-	30
291	94.975	28	-	30
292	96.408	28	-	30
293	97.592	28	-	30
294	88.741	29	-	30
295	92.002	29	-	30
296	94.883	29	-	30
297	97.267	29	-	30
298	98.995	29	-	30
299	90.000	30	-	30

Table 2: Cochran type constants (K_α) leading to a certain guaranteed ACP for all $n > K_\alpha \cdot G^2$ for nominal 95 % confidence intervals

	p	α (%) 85		90		93		94		94.5		$\frac{N}{2} \cdot \frac{1}{G^2}$	
		G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H
N=100, M=5	0.050											2.9	2.9
, M=10	0.100	4.2	4.0	(6.3)	(6.0)	(6.8)	(6.4)	(6.9)	(6.6)	(7.0)	(6.7)	7.0	6.7
, M=15	0.150	3.1	2.8	10.2	9.0	(12.8)	(11.3)	(13.0)	(11.5)	-	-	13.0	11.5
, M=20	0.200	7.1	5.5	10.7	8.3	(21.3)	(16.6)	(21.8)	(17.0)	(22.2)	(17.3)	22.2	17.3
, M=25	0.250	5.3	3.4	15.0	9.6	(36.8)	(23.5)	(36.8)	(23.5)	(36.8)	(23.5)	37.5	24.0
, M=30	0.300	7.9	3.8	15.8	7.5	44.6	21.2	(59.1)	(28.1)	(64.3)	(30.6)	65.6	31.2
, M=35	0.350	12.6	3.8	15.2	4.6	48.0	14.6	-	-	-	-	126	38.3
, M=40	0.400	24.0	3.6	30.0	4.4	174	25.7	(280)	(42.6)	(280)	(42.6)	300	44.4
, M=45	0.450	99.0	3.9	124	4.9	371	14.6	(1 064)	(41.7)	(1 064)	(41.7)	1 238	48.5
, M=50	0.500	∞	4	∞	8	∞	-	∞	-	∞	-	∞	50
N=200, M=10	0.050	3.5	3.5	(5.3)	(5.2)	(5.7)	(5.6)	(5.8)	(5.7)	(5.8)	(5.7)	5.9	5.8
, M=20	0.100	4.5	4.3	8.4	8.0	(13.9)	(13.3)	(14.1)	(13.4)	(14.1)	(13.4)	14.1	13.4
, M=30	0.150	5.5	4.8	8.6	7.6	21.3	18.8	(25.5)	(22.5)	(25.8)	(22.7)	26.0	23.0
, M=40	0.200	4.0	3.1	11.1	8.7	33.3	26.0	(43.1)	(33.6)	(43.6)	(33.9)	44.4	34.6
, M=50	0.250	5.3	3.4	15.0	9.6	42.0	26.9	56.3	36.0	(75.0)	(48.0)	75	48
, M=75	0.375	18.8	4.2	37.5	8.3	120	26.6	255	56.5	(371)	(82.2)	375	83
, M=100	0.500	∞	4	∞	11	∞	31	∞	-	∞	-	∞	100

Table 2 (cont.)

	p	α (%)										$\frac{N}{2} \cdot \frac{1}{G^2}$	
		85		90		93		94		94.5			
		G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H
N=300, M=10	0.033	3.3	3.3	(5.0)	(5.0)	(5.4)	(5.4)	(5.4)	(5.4)	(5.5)	(5.5)	5.5	5.5
, M=20	0.067	4.0	3.9	7.5	7.4	(12.3)	(12.0)	(12.3)	(12.1)	-	-	12.4	12.2
, M=30	0.100	4.5	4.3	8.9	8.4	(20.5)	(19.5)	(21.0)	(19.9)	-	-	21.1	20.1
, M=40	0.133	5.2	4.7	10.3	9.4	(31.2)	(28.3)	(31.4)	(28.5)	(31.4)	(28.5)	32.2	29.3
, M=50	0.167	3.4	2.9	9.4	8.0	28.8	24.5	(46.9)	(39.9)	(46.9)	(39.9)	46.9	39.9
, M=75	0.250	5.3	3.4	11.3	7.2	57.0	36.5	(112)	(71.5)	(112)	(71.5)	113	72
, M=100	0.333	10.0	3.6	22.0	7.9	80.0	28.8	(258)	(92.9)	(258)	(92.9)	300	108
N=400, M=10	0.025	3.2	3.3	(4.9)	(4.9)	(5.2)	(5.2)	(5.3)	(5.3)	(5.4)	(5.4)	5.4	5.4
, M=20	0.050	3.7	3.7	8.5	8.4	(11.5)	(11.4)	(11.7)	(11.6)	-	-	11.7	11.6
, M=30	0.075	4.1	4.0	8.1	7.9	(18.8)	(18.3)	(19.2)	(18.7)	(19.2)	(18.7)	19.2	18.7
, M=40	0.100	4.5	4.3	8.9	8.4	(27.1)	(25.8)	(27.3)	(26.0)	(27.6)	(26.2)	28.1	26.8
, M=50	0.125	5.1	4.7	9.9	9.1	(32.3)	(29.8)	(38.9)	(35.8)	-	-	38.9	35.8
, M=75	0.188	3.9	3.2	10.5	8.5	34.7	28.0	(77.6)	(62.7)	-	-	78	63.0
, M=100	0.250	5.3	3.4	11.3	7.2	39.8	25.4	(127)	(81.1)	(127)	(81.1)	150	96

Table 2 (cont.)

	p	α (%) 85		90		93		94		94.5		$\frac{N}{2} \cdot \frac{1}{G^2}$	
		G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H
N=500, M=10	0.020	3.2	3.2	(4.8)	(4.8)	(5.1)	(5.1)	(5.2)	(5.2)	(5.3)	(5.3)	5.3	5.3
, M=20	0.040	3.6	3.6	8.3	8.2	(11.1)	(11.0)	(11.3)	(11.3)	-	-	11.3	11.3
, M=30	0.060	3.9	3.9	9.1	9.0	(17.8)	(17.6)	(18.1)	(17.8)	(18.1)	(17.8)	18.2	17.9
, M=40	0.080	4.3	4.2	10.0	9.7	(22.3)	(21.7)	(25.5)	(24.7)	(25.7)	(24.9)	26.1	25.3
, M=50	0.100	4.5	4.3	10.8	10.3	26.2	24.9	-	-	-	-	35.2	33.5
, M=75	0.150	5.7	5.1	11.2	9.9	33.1	29.2	(65.1)	(57.4)	(65.1)	(57.4)	65.1	57.4
, M=100	0.200	4.0	3.1	11.1	8.7	29.3	22.8	(105)	(82.0)	(111)	(86.2)	111	86.5
N=600, M=10	0.017	3.2	3.2	(4.8)	(4.8)	(5.1)	(5.1)	(5.2)	(5.2)	(5.3)	(5.3)	5.3	5.3
, M=20	0.033	3.5	3.5	(9.3)	(9.3)	(10.9)	(10.8)	(11.1)	(11.0)	-	-	11.1	11.0
, M=30	0.050	3.8	3.8	8.8	8.7	(17.3)	(17.1)	(17.4)	(17.2)	(17.5)	(17.3)	17.6	17.4
, M=40	0.067	4.1	4.0	9.5	9.3	(22.6)	(22.2)	(24.3)	(23.8)	(24.6)	(24.1)	24.9	24.4
, M=50	0.083	4.3	4.2	10.2	9.9	(23.2)	(22.5)	-	-	-	-	33.0	31.9
, M=75	0.125	5.1	4.7	10.1	9.3	36.8	33.9	(56.6)	(52.2)	(58.3)	(53.8)	58.3	53.8
, M=100	0.167	6.3	5.3	9.7	8.3	29.4	25.0	(90.9)	(77.5)	(90.9)	(77.5)	93.7	79.9

Table 2 (cont.)

	p	α (%)										$\frac{N}{2} \cdot \frac{1}{G^2}$	
		85		90		93		94		94,5			
		G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H
N=700, M=10	0.014	3.1	3.1	(4.7)	(4.7)	(5.0)	(5.0)	(5.2)	(5.1)	(5.2)	(5.2)	5.2	5.2
, M=20	0.029	3.5	3.5	(9.2)	(9.2)	(10.7)	(10.7)	-	-	-	-	10.9	10.9
, M=30	0.043	3.7	3.7	10.0	9.9	(16.9)	(16.8)	(16.9)	(16.8)	(17.1)	(17.0)	17.2	17.0
, M=40	0.057	3.9	3.9	9.3	9.1	(21.9)	(21.6)	(23.6)	(23.2)	(23.9)	(23.6)	24.0	23.7
, M=50	0.071	4.2	4.1	8.1	7.9	22.1	21.6	-	-	-	-	31.6	30.9
, M=75	0.107	4.8	4.5	9.5	8.9	32.7	30.9	(46.6)	(44.0)	-	-	54.2	51.2
, M=100	0.143	5.5	4.9	8.6	7.7	28.6	25.6	67.7	60.6	(83.5)	(74.7)	84.0	75.2
N=800, M=10	0.013	3.1	3.1	(4.7)	(4.7)	(5.0)	(5.0)	(5.1)	(5.1)	(5.2)	(5.2)	5.2	5.2
, M=20	0.025	3.4	3.4	9.1	9.1	(10.6)	(10.6)	-	-	-	-	10.8	10.8
, M=30	0.038	3.6	3.6	9.8	9.8	(16.6)	(16.5)	(16.7)	(16.6)	(16.8)	(16.7)	16.9	16.8
, M=40	0.050	3.8	3.8	9.0	8.9	(21.4)	(21.2)	(23.1)	(22.8)	(23.3)	(23.1)	23.5	23.2
, M=50	0.063	4.0	3.9	9.4	9.3	24.1	23.7	-	-	-	-	30.6	30.1
, M=75	0.094	4.5	4.3	9.0	8.6	30.9	29.6	(51.5)	(49.3)	-	-	51.5	49.3
, M=100	0.125	5.1	4.7	10.1	9.3	30.1	27.8	59.5	54.8	-	-	77.8	71.7

Table 2 (cont.)

	p	α (%)										$\frac{N}{2} \cdot \frac{1}{G^2}$	
		85		90		93		94		94.5			
		G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H
N=900, M=10	0.011	3.1	3.1	(4.7)	(4.7)	(5.0)	(5.0)	(5.1)	(5.1)	(5.2)	(5.2)	5.2	5.2
, M=20	0.022	3.4	3.4	(9.0)	(9.0)	(10.5)	(10.5)	-	-	-	-	10.7	10.7
, M=30	0.033	3.6	3.6	9.7	9.7	(16.4)	(16.3)	(16.5)	(16.4)	(16.6)	(16.5)	16.6	16.6
, M=40	0.044	3.8	3.8	8.8	8.7	(22.2)	(22.0)	(22.7)	(22.5)	(23.0)	(22.8)	23.0	22.8
, M=50	0.056	3.9	3.9	9.2	9.0	(24.8)	(24.5)	-	-	-	-	29.9	29.5
, M=75	0.083	4.3	4.2	8.7	8.4	28.3	27.4	-	-	-	-	49.5	47.9
, M=100	0.111	4.9	4.6	9.6	9.1	35.3	33.1	(73.0)	(68.6)	(73.5)	(69.0)	73.5	69.0
N=1 000, M=10	0.010	3.1	3.1	(4.7)	(4.7)	(5.0)	(5.0)	(5.1)	(5.1)	(5.2)	(5.2)	5.2	5.2
, M=20	0.020	3.4	3.4	(9.0)	(9.0)	(10.4)	(10.4)	-	-	-	-	10.6	10.6
, M=30	0.030	3.6	3.5	9.6	9.6	(16.2)	(16.1)	(16.3)	(16.2)	(16.5)	(16.4)	16.5	16.4
, M=40	0.040	3.7	3.7	8.6	8.6	(21.9)	(21.8)	(22.3)	(22.2)	(22.7)	(22.5)	22.7	22.5
, M=50	0.050	3.9	3.8	9.0	8.9	(26.9)	(26.6)	-	-	-	-	29.3	29.0
, M=75	0.075	4.2	4.1	10.0	9.7	35.9	35.0	-	-	-	-	48.0	46.8
, M=100	0.100	4.6	4.4	9.3	8.8	32.1	30.5	(70.2)	(66.8)	(70.2)	(66.8)	70.3	66.9

Table 3: Cochran type constants (K_β) leading to a certain average ACP for all $n > K_\beta\, G^2$ for nominal 95 % confidence intervals

	p	β (%)											
		85		90		93		94		94.5		$\frac{N}{2} \cdot \frac{1}{G^2}$	
		G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H
N=100, M=5	0.050	1.8	1.8	2.2	2.2	-	-	-	-	-	-	2.9	2.9
, M=10	0.100	2.4	2.3	2.8	2.7	(6.8)	(6.4)	(6.9)	(6.6)	(7.0)	(6.7)	7.0	6.7
, M=15	0.150	3.1	2.8	3.6	3.2	6.5	5.8	(13.0)	(11.5)	-	-	13.0	11.5
, M=20	0.200	4.0	3.1	4.4	3.5	8.9	6.9	14.7	11.4	(22.2)	(17.3)	22.2	17.3
, M=25	0.250	5.3	3.4	6.0	3.8	7.5	4.8	19.5	12.5	(33.8)	(21.6)	37.5	24.0
, M=30	0.300	7.9	3.8	9.2	4.4	10.5	5.0	18.4	8.7	24.9	11.9	65.6	31.2
, M=35	0.350	12.6	3.8	15.2	4.6	17.7	5.4	17.7	5.4	50.6	15.3	126	38.3
, M=40	0.400	24.0	3.6	30.0	4.4	42.0	6.2	42.0	6.2	84.0	12.4	300	44.4
, M=45	0.450	99.0	3.9	124	4.9	124	4.9	149	5.8	446	17.5	1 238	48.5
, M=50	0.500	∞	4	∞	5	∞	5	∞	11	∞	11	∞	50
N=200, M=10	0.050	2.0	2.0	3.9	3.9	(5.7)	(5.6)	(5.8)	(5.7)	(5.8)	(5.7)	5.9	5.8
, M=20	0.100	2.5	2.4	5.1	4.8	(9.1)	(8.7)	(12.5)	(11.9)	(14.1)	(13.4)	14.1	13.4
, M=30	0.150	3.1	2.8	3.6	3.2	9.4	8.3	18.0	15.9	(25.8)	(22.7)	26.0	23.0
, M=40	0.200	4.0	3.1	4.9	3.8	8.9	6.9	20.9	16.3	31.1	24.2	44.4	34.6
, M=50	0.250	5.3	3.4	6.0	3.8	16.5	10.6	21.0	13.4	36.8	23.5	75	48
, M=75	0.375	18.8	4.2	18.8	4.2	26.3	5.8	56.3	12.5	93.8	20.8	375	83.0
, M=100	0.500	∞	4	∞	5	∞	5	∞	11	∞	14	∞	100

Table 3 (cont.)

	p	β (%)										$\frac{N}{2} \cdot \frac{1}{G^2}$	
		85		90		93		94		94.5			
		G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H
N=300, M=10	0.033	1.9	1.9	3.7	3.7	(5.4)	(5.4)	(5.4)	(5.4)	(5.5)	(5.5)	5.5	5.5
, M=20	0.067	2.2	2.2	4.5	4.4	8.2	8.0	(12.3)	(12.1)	-	-	12.4	12.2
, M=30	0.100	2.5	2.4	3.1	2.9	9.4	9.0	(18.0)	(17.1)	-	-	21.1	20.1
, M=40	0.133	3.0	2.7	3.4	3.1	9.0	8.2	19.1	17.4	24.9	22.7	32.2	29.3
, M=50	0.167	3.4	2.9	4.1	3.5	10.3	8.8	18.1	15.4	33.8	28.8	46.9	39.9
, M=75	0.250	5.3	3.4	6.8	4.3	12.0	7.7	25.5	16.3	25.5	16.3	113	72
, M=100	0.333	10.0	3.6	14.0	5.0	14.0	5.0	42.0	15.1	44.0	15.8	300	108
N=400, M=10	0.025	1.9	1.9	3.7	3.6	(5.2)	(5.2)	(5.3)	(5.3)	(5.4)	(5.4)	5.4	5.4
, M=20	0.050	2.1	2.1	4.2	4.2	9.0	8.9	(11.7)	(11.6)	-	-	11.7	11.6
, M=30	0.075	2.3	2.2	4.7	4.6	10.4	10.1	(16.5)	(16.1)	(19.2)	(18.7)	19.2	18.7
, M=40	0.100	2.5	2.4	5.2	5.0	9.7	9.2	15.1	14.3	(26.2)	(24.9)	28.1	26.8
, M=50	0.125	2.9	2.7	3.5	3.2	10.9	10.0	19.1	17.6	27.8	25.6	38.9	35.8
, M=75	0.188	3.9	3.2	4.7	3.8	11.7	9.5	17.6	14.2	35.5	28.7	78	63.0
, M=100	0.250	5.3	3.4	6.8	4.3	12.0	7.7	25.5	16.3	40.5	25.9	150	96

Table 3 (cont.)

	p	β (%)										$\frac{N}{2} \cdot \frac{1}{G^2}$	
		85		90		93		94		94.5			
		G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H
N=500, M=10	0.020	1.8	1.8	3.6	3.6	(5.1)	(5.1)	(5.2)	(5.2)	(5.3)	(5.3)	5.3	5.3
, M=20	0.040	2.0	2.0	4.1	4.1	8.7	8.6	(11.3)	(11.3)	-	-	11.3	11.3
, M=30	0.060	2.2	2.2	4.5	4.4	9.7	9.5	15.5	15.3	(18.1)	(17.8)	18.2	17.9
, M=40	0.080	2.4	2.3	4.9	4.8	9.1	8.8	17.3	16.8	(25.7)	(24.9)	26.1	25.3
, M=50	0.100	2.5	2.4	5.3	5.1	11.5	11.0	20.4	19.4	(30.0)	(28.5)	35.2	33.5
, M=75	0.150	3.1	2.8	3.9	3.5	12.2	10.8	21.1	18.6	35.9	31.7	65.1	57.4
, M=100	0.200	4.0	3.1	4.9	3.8	12.4	9.7	21.8	17.0	31.1	24.2	111	86.5
N=600, M=10	0.017	1.8	1.8	3.6	3.6	(5.1)	(5.1)	(5.2)	(5.2)	(5.3)	(5.3)	5.3	5.3
, M=20	0.033	2.0	2.0	4.0	4.0	8.5	8.5	(11.1)	(11.0)	-	-	11.1	11.0
, M=30	0.050	2.2	2.2	4.4	4.4	9.3	9.2	(15.0)	(14.8)	(17.5)	(17.3)	17.6	17.4
, M=40	0.067	2.2	2.2	4.6	4.6	10.1	9.9	16.3	16.0	(24.6)	(24.1)	24.9	24.4
, M=50	0.083	2.4	2.3	5.0	4.8	9.1	8.8	20.7	20.0	(31.0)	(30.0)	33.0	31.9
, M=75	0.125	2.9	2.7	3.5	3.2	10.9	10.0	17.3	16.0	33.8	31.2	58.3	53.8
, M=100	0.167	3.4	2.9	4.1	3.5	13.4	11.5	20.9	17.8	37.2	31.7	93.7	79.9

Table 3 (cont.)

	p	β (%) 85		90		93		94		94.5		$\frac{N}{2} \cdot \frac{1}{G^2}$	
		G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H	G_C	G_H
N=700, M=10	0.014	1.8	1.8	3.6	3.6	(5.0)	(5.0)	(5.2)	(5.1)	(5.2)	(5.1)	5.2	5.2
, M=20	0.029	2.0	2.0	4.0	4.0	8.4	8.3	-	-	-	-	10.9	10.9
, M=30	0.043	2.1	2.1	4.2	4.2	9.1	9.1	(14.5)	(14.4)	(17.1)	(17.0)	17.2	17.0
, M=40	0.057	2.2	2.2	4.5	4.4	9.8	9.7	17.0	16.8	(23.9)	(23.6)	24.0	23.7
, M=50	0.071	2.4	2.3	4.7	4.6	10.5	10.2	18.5	18.1	(29.6)	(28.9)	31.6	30.9
, M=75	0.107	2.6	2.5	3.3	3.1	12.2	11.6	19.5	18.4	33.2	31.3	54.2	51.2
, M=100	0.143	3.1	2.8	3.6	3.2	9.4	8.4	18.7	16.8	33.1	29.6	84.0	75.2
N=800, M=10	0.013	1.8	1.8	3.5	3.5	(5.0)	(5.0)	(5.1)	(5.1)	(5.2)	(5.2)	5.2	5.2
, M=20	0.025	1.9	1.9	3.9	3.9	8.3	8.3	-	-	-	-	10.8	10.8
, M=30	0.038	2.1	2.1	4.2	4.2	9.0	8.9	(15.5)	(15.4)	(16.8)	(16.7)	16.9	16.8
, M=40	0.050	2.2	2.2	4.4	4.4	11.0	10.9	15.3	15.1	(23.3)	(23.1)	23.5	23.2
, M=50	0.063	2.2	2.2	4.6	4.5	10.3	10.1	19.1	18.7	(28.6)	(28.1)	30.6	30.1
, M=75	0.094	2.6	2.5	5.2	4.9	9.7	9.3	18.5	17.8	34.8	33.3	51.5	49.3
, M=100	0.125	2.9	2.7	3.5	3.2	13.2	12.2	21.2	19.5	34.6	31.9	77.8	71.7

Table 3 (cont.)

	p	β (%) 85 G_C	85 G_H	90 G_C	90 G_H	93 G_C	93 G_H	94 G_C	94 G_H	94.5 G_C	94.5 G_H	$\frac{N}{2} \cdot \frac{1}{G^2}$ G_C	$\frac{N}{2} \cdot \frac{1}{G^2}$ G_H
N=900, M=10	0.011	1.8	1.8	3.5	3.5	(5.0)	(5.0)	(5.1)	(5.1)	(5.2)	(5.2)	5.2	5.2
, M=20	0.022	1.9	1.9	3.9	3.9	8.2	8.2	-	-	-	-	10.7	10.7
, M=30	0.033	2.0	2.0	4.1	4.1	8.9	8.8	(15.3)	(15.2)	(16.6)	(16.5)	16.6	16.6
, M=40	0.044	2.1	2.1	4.3	4.3	9.5	9.4	16.4	16.2	(23.0)	(22.8)	23.0	22.8
, M=50	0.056	2.2	2.2	4.5	4.5	10.0	9.9	17.4	17.2	(27.9)	(27.5)	29.9	29.5
, M=75	0.083	2.4	2.3	5.0	4.8	11.0	10.6	19.7	19.1	34.5	33.4	49.5	47.9
, M=100	0.111	2.8	2.6	3.3	3.1	12.4	11.7	22.2	20.9	40.7	38.2	73.5	69.0
N=1 000, M=10	0.010	1.8	1.8	3.5	3.5	(5.0)	(5.0)	(5.1)	(5.1)	(5.2)	(5.2)	5.2	5.2
, M=20	0.020	1.9	1.9	3.9	3.9	8.2	8.2	-	-	-	-	10.6	10.6
, M=30	0.030	2.0	2.0	4.1	4.0	8.8	8.8	(14.0)	(13.9)	(16.5)	(16.4)	16.5	16.4
, M=40	0.040	2.1	2.1	4.3	4.2	9.4	9.3	16.2	16.0	(22.7)	(22.5)	22.7	22.5
, M=50	0.050	2.2	2.2	4.4	4.4	9.8	9.7	17.1	16.9	(27.3)	(27.0)	29.3	29.0
, M=75	0.075	2.4	2.3	4.8	4.7	10.7	10.4	20.5	19.9	32.2	31.3	48.0	46.8
, M=100	0.100	2.5	2.4	3.1	2.9	10.0	9.5	21.1	20.1	36.1	34.4	70.3	66.9

Table 4: ACPs for some populations and sample sizes when N=10 000. Normal approximation is used.

	M =																	
n =	5	(K_H)	10	(K_H)	50	(K_H)	100	(K_H)	200	(K_H)	500	(K_H)	1000	(K_H)	2000	(K_H)	5000	(K_H)
10	0.5		1.0		4.9		9.6		18.3		40.0		65.0		88.6	(3.5)	89.1	(10)
50	2.5		4.9		22.2		39.6		63.6		92.0	(2.9)	88.0	(6.7)	93.8	(17.3)	93.6	(50)
100	4.9		9.6		39.6		63.5		86.8	(2.1)	87.9	(5.8)	93.4	(13.4)	93.4	(34.6)	94.4	
300	14.1		26.3		78.2		80.4	(3.1)	93.5	(6.4)	92.9		94.1	(40.2)	95.0		94.7	
500	22.6		40.1		92.0	(2.6)	87.9	(5.2)	92.8	(10.6)	93.8	(29)	94.9		94.6		95.2	
1 000	41.0		65.1		88.6	(5.1)	93.4	(10.3)	93.6	(21.2)	94.9	(58)	95.2		94.9			
2 000	67.2		89.2	(2)	94.6	(10.2)	94.3	(20.6)	94.4		94.6		94.9		95.1			
3 000	83.2		84.9		94.8		94.1		94.6		94.8		94.9		95.0			
4 000	92.2	(2)	95.2		93.9	(20.4)	94.5	(41.2)	94.9		94.4		94.7		95.0			
5 000	81.3		94.4	(5)	95.2		93.9		94.4		95.1		95.0		94.9			
Skewness (G_C)	44.7		31.6		14.0		9.8		6.9		4.1		2.7		1.5		0	
Skewness (G_H)	44.7		31.6		14.0		9.9		6.9		4.2		2.7		1.7		1	

Table 5: ACPs for some populations and sample sizes when N=100 000. Normal approximation is used.

	M =																	
n =	10	(K_H)	50	(K_H)	100	(K_H)	500	(K_H)	1000	(K_H)	2000	(K_H)	5000	(K_H)	10000	(K_H)	20000	(K_H)
10	0.1		0.5		1.0		4.9		9.6		18.3		40.0		65.0		88.6	(3.5)
50	0.5		2.5		4.9		22.2		39.5		63.5		92.0	(2.9)	87.9	(6.7)	93.8	(17.3)
100	1.0		4.9		9.5		39.4		63.4		86.7	(2.1)	87.8	(5.8)	93.3	(13.4)	93.3	(34.6)
300	3.0		14.0		26.0		77.7		79.9	(3.1)	93.2	(6.4)	92.6		93.8	(40.2)	94.7	
500	4.9		22.2		39.4		91.5	(2.6)	87.2	(5.2)	92.9	(10.6)	93.2	(29)	94.3		94.9	
1 000	9.6		39.5		63.4		87.2	(5.1)	92.8	(10.3)	92.4	(21.2)	94.3	(58)	95.4		94.8	
2 000	18.3		63.5		86.7	(2)	92.9	(10.2)	92.4	(20.6)	94.7	(42.4)	94.9		94.9		94.8	
3 000	26.3		78.1		80.2		92.4		93.6		95.0		94.8		95.1		95.1	
4 000	33.5		86.9	(2)	91.1		92.6	(20.4)	94.9	(41.2)	94.7		94.8		94.7		94.9	
5 000	40.1		92.0		87.8	(5)	93.2		94.3		94.9		94.9		95.0		95.0	
10 000	65.1		88.5	(5)	93.3	(10)	94.3	(51)	94.6									
20 000	89.2	(2)	94.6	(10)	94.2	(20)	95.4		95.2									
50 000	94.4	(5)	95.1	(25)	93.8	(50)	94.5		94.7									
Skewness (G_C)	100.0		44.7		31.6		14.0		9.8		6.9		4.1		2.7		1.5	
Skewness (G_H)	100.0		44.7		31.6		14.0		9.9		6.9		4.2		2.7		1.7	

Table 6: ACPs when N=8 000 000 and M=80 000 for some sample sizes $G_C \approx G_H = 9.85$. (Normal approximation is used.)

n	ACP	Outcomes where I_k=1	n	ACP	Outcomes where I_k=1	n	ACP	Outcomes where I_k=1
50	39.5	1-4	1 750	92.0	12-27	3 450	94.6	25-47
100	63.3	1-5	1 800	93.6	12-28	3 500	93.8	26-48
150	77.8	1-6	1 850	94.3	12-28	3 550	94.8	26-49
200	86.5	1-7	1 900	92.9	13-29	3 600	93.4	27-49
250	91.5	1-7	1 950	94.3	13-30	3 650	94.4	27-50
300	79.9	2-8	2 000	94.9	13-30	3 700	94.8	27-50
350	86.2	2-9	2 050	93.6	14-31	3 750	94.1	28-51
400	90.7	2-10	2 100	94.8	14-32	3 800	95.0	28-52
450	93.3	2-10	2 150	93.0	15-32	3 850	93.8	29-52
500	87.1	3-11	2 200	94.3	15-33	3 900	94.7	29-53
550	90.8	3-12	2 250	94.8	15-33	3 950	95.0	29-53
600	93.0	3-12	2 300	93.7	16-34	4 000	94.4	30-54
650	88.3	4-13	2 350	94.9	16-35	4 050	94.7	30-54
700	91.4	4-14	2 400	93.2	17-35	4 100	94.1	31-55
750	93.7	4-15	2 450	94.4	17-36	4 150	95.0	31-56
800	89.4	5-15	2 500	94.9	17-36	4 200	93.8	32-56
850	92.0	5-16	2 550	93.9	18-37	4 250	94.7	32-57
900	94.1	5-17	2 600	94.4	18-37	4 300	95.0	32-57
950	90.4	6-17	2 650	93.4	19-38	4 350	94.4	33-58
1 000	92.7	6-18	2 700	94.5	19-39	4 400	94.7	33-58
1 050	93.9	6-18	2 750	95.0	19-39	4 450	94.1	34-59
1 100	91.3	7-19	2 800	94.1	20-40	4 500	95.0	34-60
1 150	93.3	7-20	2 850	94.6	20-40	4 550	93.8	35-60
1 200	94.4	7-20	2 900	93.7	21-41	4 600	94.7	35-61
1 250	92.2	8-21	2 950	94.7	21-42	4 650	95.0	35-61
1 300	94.0	8-22	3 000	93.2	22-42	4 700	94.4	36-62
1 350	94.8	8-22	3 050	94.3	22-43	4 750	94.7	36-62
1 400	93.0	9-23	3 100	94.7	22-43	4 800	94.2	37-63
1 450	94.5	9-24	3 150	93.9	23-44	4 850	95.0	37-64
1 500	92.0	10-24	3 200	94.9	23-45	4 900	93.9	38-64
1 550	93.7	10-25	3 250	93.5	24-45	4 950	94.7	38-65
1 600	94.5	10-25	3 300	94.5	24-46	5 000	95.0	38-65
1 650	92.9	11-26	3 350	94.9	24-46	5 050	94.5	39-66
1 700	94.3	11-27	3 400	94.2	25-47	5 100	94.8	39-66

Table 6 (cont.)

n	ACP	Outcomes where I_k=1
5 150	94.3	40-67
5 200	95.0	40-68
5 250	94.0	41-68
5 300	94.8	41-69
5 350	93.8	42-69
5 400	94.6	42-70
5 450	94.8	42-70
5 500	94.4	43-71
5 550	95.1	43-72
5 600	94.2	44-72
5 650	94.9	44-73
5 700	93.9	45-73
5 750	94.7	45-74
5 800	94.9	45-74
5 950	94.3	47-76
6 000	95.0	47-77

Table 7: Simulations from populations based on the beta distribution.

K	β(%)	b=1/a=1		b=1/a=1.756		b=1/a=2.390		b=1/a=4.083		b=1/a=6.16		b=1/a=8.509	
		G_H=1.299		G_H=1.592		G_H=1.978		G_H=2.946		G_H=3.882		G_H=4.714	
			n		n		n		n		n		n
		G_C = 0		G_C = 1.000		G_C = 1.598		G_C = 2.767		G_C = 3.780		G_C = 4.647	
3	85	94.2	6	90.2	8	90.9	12	89.8	27	88.5	46	89.3	67
5	90	94.6	9	92.6	13	90.9	20	90.9	44	91.7	76	91.6	112
11	93	94.8	19	93.7	28	93.3	44	94.1	96	94.2	166	94.6	245
20	94	95.1	34	95.5	51	93.9*	79	94.5	174				
40	94.5	94.7	68	95.0	102	94.6	157						

Table 8: Simulations from populations based on the lognormal distribution.

K	β(%)	$\sigma=0.1$		$\sigma=0.4$		$\sigma=0.72$		$\sigma=0.92$		$\sigma=1.06$		$\sigma=1.3$	
		$G_H=1.597$		$G_H=1.982$		$G_H=2.943$		$G_H=3.883$		$G_H=4.695$		$G_H=6.331$	
			n		n		n		n		n		n
		$G_C=0.220$		$G_C=1.25$		$G_C=2.577$		$G_C=3.654$		$G_C=4.533$		$G_C=6.243$	
3	85	95.9	8	92.7	12	92.1	26	89.6	46	90.6	67	89.5	121
5	90	94.7	13	93.6	20	91.4	44	92.8	76	91.8	111	91.7	201
11	93	95.9	28	95.6	44	94.4	96	95.2	166	94.5	243		
20	94	94.8	51	95.5	79	93.9*	174						
40	94.5	95.4	102	96.2	158								

Table 9: Simulations from populations based on the power function distribution

K	β(%)	C=0.316		C=0.1933		C=0.0905		C=0.0540		C=0.0376		C=0.0213	
		G_H=1.590		G_H=1.981		G_H=2.940		G_H=3.877		G_H=4.698		G_H=6.319	
			n		n		n		n		n		n
		G_C=1.117		G_C=1.702		G_C=2.829		G_C=3.820		G_C=4.663		G_C=6.304	
3	85	89.6	8	88.5	12	87.7	26	88.8	46	87.0	67	88.2	120
5	90	92.3	13	91.0	20	91.1	44	92.0	76	92.0	111	90.9	200
11	93	93.7	28	93.3	44	93.1	96	92.9*	166	94.9	243		
20	94	94.0	51	95.0	79	94.5	173						
40	94.5	94.7	102	95.1	157								

Table 10: Simulations from populations based on the Weibull distribution.

K	β_B(%)	C=2.15		C=1.24		C=0.795		C=0.615		C=0.53		C=0.425	
		G_H=1.585		G_H=1.984		G_H=2.923		G_H=3.967		G_H=4.831		G_H=6.554	
			n		n		n		n		n		n
		G_C=0.528		G_C=1.406		G_C=2.645		G_C=3.809		G_C=4.724		G_C=6.499	
3	85	94.4	8	92.8	12	90.7	26	90.4	48	88.9	71	91.0	129
5	90	94.1	13	93.8	20	92.0	43	92.0	79	91.6	117	91.0	215
11	93	95.0	28	94.3	44	93.8	94	93.2	174				
20	94	94.5	51	93.7*	79	95.3	171						
40	94.5	95.2	101	94.1*	158								

G-90693-30001-001-K-H

CATARINA ELFFORS

A/BF-S

Tidigare nummer av Promemorior från P/STM:

Nr	
1	Bayesianska idéer vid planeringen av sample surveys. Lars Lyberg (1978-11-01)
2	Litteraturförteckning över artiklar om kontingenstabeller. Anders Andersson (1978-11-07)
3	En presentation av Box-Jenkins metod för analys och prognoser av tidsserier. Åke Holmén (1979-12-20)
4	Handledning i AID-analys. Anders Norberg (1980-10-22)
5	Utredning angående statistisk analysverksamhet vid SCB: Slutrapport.P/STM, Analysprojektet (1980-10-31)
6	Metoder för evalvering av noggrannheten i SCBs statistik. En översikt. Jörgen Dalén (1981-03-02)
7	Effektiva strategier för estimation av förändringar och nivåer vid föränderlig population. Gösta Forsman och Tomas Garås (1982-11-01)

Kvarvarande exemplar av ovanstående nummer kan rekvireras från SCB, P/STM, Elseliv Lindfors, 115 81 Stockholm, eller per telefon 08 14 05 60 ankn 4178.

Garnisonstryckeriet ★ Stockholm ★ 1983