Discussion Paper

Central Bureau of Statistics, P.B. 8131 Dep, 0033 Oslo 1, Norway

No. 30 March 1988

ESTIMATING CHANGE IN A PROPORTION BY COMBINING MEASUREMENTS FROM A TRUE AND A FALLIBLE CLASSIFIER.

BY

ANDERS RYGH SWENSEN

Abstract

Consider a binary classification of a large population at two points in time. The classification is observed with error for the whole population using a fallible classifier and without error for a random sample using an accurate classifier. Following Tenenbein (1970), the population proportions are estimated by poststratification according to the fallible classifier for both the time points. Assuming a multinomial probability model, the joint asymptotic normality of the two estimators is demonstrated. Comparison is made with the estimator based on the survey data only. In particular the importance of including the same items in the samples at both time points is discussed.

The main part of the paper will appear in Scandinavian Journal of Statistcs. The editor, Søren Johansen, suggested several improvements that enhanced the quality of the paper.

1. Introduction.

The present paper is motivated by some practical considerations on the use of administrative registers in production of labor market statistics. From administrative records, mainly from the social security system, one can obtain data on whether an individual belongs to the labor force or not. For various reasons these figures are not accurate enough to be used as official statistics. Potential sources of degradation are belated updating and discrepancies in the definitions of labor market status.

In Norway labor force surveys are conducted quarterly by interviews. By matching the survey results with the classifications of the registers a fairly close correspondence is apparent. One may therefore expect to obtain estimators having smaller variance than the survey means by using the classification of the registers as a poststratification variable.

Formally, let the result of matching the survey and the register be $\{ n_{ij} \}_{i,j=0,1}$ where i denotes the classification of the survey, j the classification of the register. Let N_j, j=0,1 denote the number of persons belonging to category j in the register. The post-stratified estimator of the proportion of the population in state 1 is then

$$\frac{n_{10}}{n_{.0}} \frac{N_0}{N} + \frac{n_{11}}{n_{.1}} \frac{N_1}{N} \qquad (1)$$

where $n_{.0} = n_{00} + n_{10}$, $n_{.1} = n_{01} + n_{11}$ and $N = N_0 + N_1$.

Estimates of the change from the previous survey, and from the corresponding quarter last year are of considerable interest. To evaluate the variance of these estimates one needs some knowledge about the covariance between estimators of the form (1) taken at various epochs.

If we assume that the N units are drawn from a multinomial distribution and can be classified on two dimensions, the stochastic model is identical with the one treated by Tenenbein (1970), (1972). The observations consist of two parts: n units for which the complete classification is known and N - n units for which only the classification on the second dimension is known. Tenenbein (1970) showed that the estimator (1) is a maximum likelihood estimator in this sampling model. He also showed that it had an asymptotically normal distribution and gave a nice interpretation of the asymptotic variance.

The original motivation for introducing the model was a situation where two measuring devices were available; one fallible where measurements were cheap and easy to get, and another one which was more expensive to use but gave more accurate measurements. The double sampling procedure provides a method for estimating proportions as defined by the accurate measuring device.The analogy to the problem at hand is immediate. The survey is the accurate instrument, and the administrative records are the measurements of the fallible measuring device.

The framework can be generalized to include different variates along the completely and incompletely classified dimensions. A fairly substantial amount of research on this type of models has been directed towards testing and estimating structural models. One can mention Chen & Fienberg (1974 ,1976), Chen (1979), Espeland & Odoroff (1985) and Palmgren (1987).

A unifying theme of many population models with complicated schemes of observation is the assumption of a common probability model for each individual of the population, e. g. a Markov chain in case a dynamic situation is considered. Assuming statistical independence between the realizations, the distribution of the observations will be based on this probability model and the restrictions imposed by the scheme of observation. The model using incompletely classified data referred to in the previous paragraph is one example, combination of micro and macro data in economics is another. Rosenqvist (1986) provides a recent treatment of the latter.

In our case the probabilistic model is just a multinomial classification of each individual. The scheme of observation is more intricate. Some individuals are observed using the accurate device at both occasions and some using it only at one time point. For the rest of the population only the results obtained from the fallible classifier is available.

The paper is organized as follows. In section 2 we define the basic setup and give the main results. Section 3 is an illustration, discussing the importance of including the same individuals at both time points in the part classified by the accurate measuring device. Some technical details are collected in the appendix.

2. The model and main results.

To study the behavior of the estimator (1) at two different time points we shall assume a simple multinomial model for the total population consisting of N units. At each occasion there are four possible states corresponding to the combinations of the classifications of the accurate and fallible measuring device. We denote the states ij, i,j=0,1 where i refers to the first (accurate) dimension of the 2x2 classification and j to the second (inaccurate) dimension. Hence, when considering two points in time there are 16 combinations that must be taken into account. Let $p_{iji'j'}$ denote the probability of an individual being in state ij on the first occasion and in state i'j' on the second. The following table summarises the organisation of the parameters.

		Second occasion			
		00	01	10	11
First occasion	00	p_{0000}	p_{0001}	p_{0010}	p_{0011}
	01	p_{0100}	p_{0101}	p_{0110}	p_{0111}
	10	p_{1000}	p_{1001}	p_{1010}	p_{1011}
	11	p_{1100}	p_{1101}	p_{1110}	p_{1111}

Note that the model above implies a closed population, i.e. there is no immigration or emigration. We shall adopt the convention that summing over different states is denoted by a ., e.g. $p_{1.11}$ = $p_{1011} + p_{1111}$.

As explained in the introduction the complete classification of the units is not known. Only partial information is available. We shall indicate which observations are necessary to compute the poststratified estimators. These are the natural ones for estimating the fraction belonging to a particular state at each occasion. By keeping track of the units between the two points in time more information can be obtained. It is therefore possible to construct more efficient estimators. We are, however, mainly interested in the covariance structure of the poststratified estimator and shall not pursue the question of efficient estimation here.

The scheme of observation may, therefore, be described as follows:

For n_M units the complete 2x2 classification is known at both occasions. Thus the observations are $n_{Mij..}$ and $n_{M..i'j'}$, $i,j,i',j' = 0,1$.

For n_S units the complete 2x2 classification is known at the first occasion. At the second occasion only the classification on the second dimension is known. The observations are $n_{Sij..}$ and $n_{S...j'}$, $i,j,j'=0,1$.

For n_T units the complete 2x2 classification is known at the second occasion. At the first occasion only the classification on the second dimension is known. The observations are $n_{T..i'j'}$ and $n_{T.j..}$, $j,i',j'=0,1$.

For n_R units only the classification on the second dimension is known at both occasions. The observations are $n_{R.j..}$ and $n_{R...j'}$, $j,j'=0,1$.

Let $\underline{n}_M$, $\underline{n}_S$, $\underline{n}_T$ and $\underline{n}_R$ denote the 4x4 array having as elements the size of each of the four parts of the population that belongs to each of the 16 categories, e.g. $\underline{n}_M=\{ n_{Miji'j'} \}_{i,j,i',j'=0,1}$. We assume that $\underline{n}_M$, $\underline{n}_S$, $\underline{n}_T$ and $\underline{n}_R$ are independently multinomially distributed. Note that they are only partially observed although $\underline{n}_M$ may, as mentioned above, in principle be completely observed. Let $n = n_M + n_S + n_T$, hence $n_R = N - n$.

The parameters to be estimated are the relative numbers in state 1 on the first dimension at each occasion, i.e. $p_{1...} = \Sigma_{j,i',j'}\, p_{1ji'j'}$ and $p_{..1.} = \Sigma_{i,j,j'}\, p_{ij1j'}$. We shall consider the estimators

$$\hat{p}_{1...} = \frac{n_{M10..} + n_{S10..}}{n_{M.0..} + n_{S.0..}} \cdot \frac{n_{M.0..} + n_{S.0..} + n_{T.0..} + n_{R.0..}}{n_M + n_S + n_T + n_R}$$

$$+ \frac{n_{M11..} + n_{S11..}}{n_{M.1..} + n_{S.1..}} \cdot \frac{n_{M.1..} + n_{S.1..} + n_{T.1..} + n_{R.1..}}{n_M + n_S + n_T + n_R}$$

and

$$\hat{p}_{..1.} = \frac{n_{M..10} + n_{T..10}}{n_{M...0} + n_{T...0}} \cdot \frac{n_{M...0} + n_{S...0} + n_{T...0} + n_{R...0}}{n_M + n_S + n_T + n_R}$$

$$+ \frac{n_{M..11} + n_{T..11}}{n_{M...1} + n_{T...1}} \cdot \frac{n_{M...1} + n_{S...1} + n_{T...1} + n_{R...1}}{n_M + n_S + n_T + n_R}$$

which are exactly the poststratified estimators of the introduction. Suppose the following limits exist as n (and therefore also N) $\to \infty$:

$$\alpha_i = \lim n_i /n , \qquad i = M,S \text{ and } T \quad \text{and} \quad \alpha = \lim (N - n)/n.$$

We assume that $\alpha_M + \alpha_S > 0$ and $\alpha_M + \alpha_T > 0$. Using the δ-method, see e.g Rao (1973) p.385, one can show that the estimators $\hat{p}_{1...}$ and $\hat{p}_{..1.}$ are asymptotically normally distributed with mean $p_{1...}$ and $p_{..1.}$ and covariance matrix Σ_A/n as $n \to \infty$. The computation is straightforward but tedious, so we only give the main results. More details on the derivation can be found in the appendix.

The elements of Σ_A are given by

$$\sigma_{11A} = p_{1...} (1-p_{1...}) ((1-K_1)/(\alpha_M+\alpha_S) + K_1/(1 + \alpha))$$

where $K_1 = (p_{11..} - p_{1...}p_{.1..})^2 / p_{0...}p_{1...}p_{.0..}p_{.1..}$,

$$\sigma_{22A} = p_{..1.} (1-p_{..1.}) ((1-K_2)/(\alpha_M+\alpha_T) + K_2/(1 + \alpha))$$

where $K_1 = (p_{..11} - p_{..1.}p_{...1})^2 / p_{..0.}p_{..1.}p_{...0}p_{...1}$,

and

$$\sigma_{12A} = (p_{1.1.} - p_{1...} p_{..1.}) ((1 - \varkappa_{12}) \alpha_M/(\alpha_M+\alpha_S)(\alpha_M+\alpha_T) + \varkappa_{12}/(1+\alpha))$$

where

$$\varkappa_{12} = 1 - K_{12}/ (p_{1.1.} - p_{1...} p_{..1.})$$

and

$$K_{12} = \Sigma_{i,j,i',j'} (-1)^{i-i'} p_{iji'j'} (1-p_{ij..} /p_{.j..})(1-p_{..i'j'} / p_{...j'}).$$

The variances σ_{11A}/n and σ_{22A}/n are those derived by Tenenbein (1970) for the univariate case. One can consult Tenenbein's paper for a more detailed discussion. We only mention that K_1 and K_2, the so-called reliability coefficients, measure the strength between the classifications along the two dimensions. More specifically, K_1 is the square of the correlation coefficient of classifying a particular unit in the same category on both dimensions at the first occasion. The variances, σ_{11A}/n say, may therefore be expressed as a convex combination of the variance between estimators based on $(1+\alpha)n$ and $(\alpha_M + \alpha_S)n$ completely classified observations.

We shall compare the estimators $\hat{p}_{1...}$ and $\hat{p}_{..1.}$ with the following simple ones which involve $\underline{n}_M$, $\underline{n}_S$ and $\underline{n}_T$ only:

$\tilde{p}_{1...} = (n_{M1...} + n_{S1...})/(n_M + n_S)$ and $\tilde{p}_{..1.} = (n_{M..1.} + n_{T..1.})/(n_M + n_T)$. These estimators are based on the accurate measurements, and make no use of the observations obtained by the fallible classifier. As $n \to \infty$, $\tilde{p}_{1...}$ and $\tilde{p}_{..1.}$ are asymptotically normally distributed with mean $p_{1...}$ and $p_{..1.}$ and covariance matrix Σ_B/n. The elements of Σ_B are given by

$$\sigma_{11B} = p_{1...}(1 - p_{1...})/(\alpha_M + \alpha_S),$$

$$\sigma_{22B} = p_{..1.}(1 - p_{..1.})/(\alpha_M + \alpha_T) \quad \text{and}$$

$$\sigma_{12B} = (p_{1.1.} - p_{1...}\,p_{..1.})\alpha_M/(\alpha_M + \alpha_S)(\alpha_M + \alpha_T).$$

As explained in the introduction, we are particularly interested in the variance of the estimators of change $\hat{p}_{1...} - \hat{p}_{..1.}$. The variance of the approximate distribution is given by $\sigma_{11A} + \sigma_{22A} - 2\sigma_{12A}$. For a discussion of the terms σ_{11A} and σ_{22A} we refer to the papers by Tenenbein (1970), (1972). Here we shall concentrate on σ_{12A} and compare it with the corresponding term, σ_{12B}, of the estimated change using the simple estimator.

From the results cited above it follows that

$$\sigma_{12B} - \sigma_{12A} = [\, p_{1.1.} - p_{1...}\, p_{..1.} - K_{12} \,]$$
$$[\alpha_M/(\alpha_M+\alpha_S)(\alpha_M+\alpha_T) - 1/(1+\alpha)\,].$$

The first factor of the product on the right hand side is $(1+\alpha)n$ times the approximate covariance of two poststratified estimators with no units among those classified by the accurate measuring device at both occasions (i.e $\alpha_M = 0$). Each term is dominated by 1 so the factor is less than 2 in absolute value. The second factor reflects the relative size of the various parts of the completely classified observations.

3. An illustration: Poststratification and design of repeated surveys.

In designing repeated surveys it is common practice to include overlapping parts in order to reduce the variance of certain estimates of change, e. g. between successive survey periods. When estimating fairly stable population characteristics, this can result in substantial gains compared to estimates based on surveys with no common elements. The problem we want to throw some light on is how poststratification affects this fact.

The covariance of the poststratified estimator is a complicated function of the parameters $p_{0000},\ldots,p_{1111}$. It may therefore be difficult to find a simple interpretation of the variance of the estimated change in terms of the parameters. To get an idea on this dependence we shall therefore consider three constructed numerical examples.

Let the parameters be given by:

Case 1	0.30	0.05	0.005	0.005
	0.05	0.03	0.005	0.005
	0.005	0.005	0.03	0.05
	0.005	0.005	0.05	0.40

Case 2

0.35	0.02	0.005	0.005
0.02	0.04	0.005	0.005
0.005	0.005	0.04	0.02
0.005	0.005	0.02	0.45

Case 3

0.40	0.01	0.005	0.005
0.01	0.01	0.005	0.005
0.005	0.005	0.01	0.01
0.005	0.005	0.01	0.50 .

Note that the correspondence between the classifications increases. Also, the 2x2 subtables corresponding to the classification of the accurate measuring device are the same in all three cases. Hence the variances and covariances of the sample means, i. e. method B, will be identical.

To simplify the situation we shall assume that α is so large that we need only consider the term involving $\alpha_M/(\alpha_M+\alpha_S)(\alpha_M+\alpha_T)$ in the covariance formula of the poststratified estimator. This corresponds to a situation where the part of the population for which only the results of the fallible measuring device are available, is large compared to the part for which accurate measurements are taken. We recall that the accurate classifier is the sample survey. In comparing different rotation plans according to how precisely they allow a change between two points in time to be measured, it is most natural to keep the sample size fixed and introduce a parameter for the fraction of the sample that is retained. Thus, let the sample size at each occasion be m, and assume that rm units are observed at both points in time. The number of distinct units is then (2-r)m, corresponding to what we denoted by n in the previous section. In terms of r, $\alpha_M/(\alpha_M+\alpha_S)(\alpha_M+\alpha_T) = (2-r)r$ and $1/(\alpha_M+\alpha_S) = 2-r$. Hence, the variance of the approximate distribution of the estimator of change based on the sample mean is $[\ 2x0.55x0.45 - 2r(0.53-0.55)^2\ ]/m$. The variance, when the estimator of change is based on the poststratified estimator, is [$(2 - K_1 - K_2)$x0.55x0.45 - $2rK_{12}$]/m. In Table 1 we have collected some quantities related to the variance of the estimators for the three numerical examples. K_1, K_2 and K_{12} are given in the first and second row. The third row shows the correlation between the poststratified estimators using completely overlapping samples (i.e. r=1).

Table 1. Some quantities related to the variance of the estimators of change.

	Case 1	Case 2	Case 3
$K_1=K_2$	0.40	0.51	0.77
K_{12}	0.083	0.075	0.013
ϱ_{PS}	0.55	0.61	0.22

Table 2 indicates the effect of varying the proportion of the common part of repeated surveys when estimating change. All quantities are computed as the percentage of the variance of the estimator based on the survey mean with no common units at the two occasions (i.e. r=0). The figures of the first and second row are based on half of the sample being common at both time points (i.e. r=1/2). The first row shows the variance of the estimator using poststratification, the second the variance of the survey based estimator. The variance of the poststratified estimator when the sample contains no common units is displayed in the third row.

Table 2. The (relative) variance of estimators of change based on the sample mean (M) and on the poststratified estimator (PS). The fraction of the sample being common at both occasions is denoted by r.

Estimator	r	Case 1	Case 2	Case 3
PS	0.5	44	34	21
M	0.5	55	55	55
PS	0.0	60	49	23
M	0.0	100	100	100

One conclusion to be drawn from Table 2 is that the correlation structure of the poststratified estimator can differ substantially from that of the mean of random samples. Furthermore, it is worth noting that the gain using overlapping samples when poststratification is the method of estimation, is smaller in case 3 than in case 2 and 1. This may indicate that the effect of classifying the same elements by the accurate measuring device at both occasions decreases as the correspondence between the two measuring devices becomes closer. Although this conclusion is rather tentative, it has a certain intuitive appeal: Closer correspondence between the two measuring devices implies that the importance of the accurate measurements diminishes, hence also the gain which may be obtained by a skillful sampling design.

Appendix.

We shall show how the asymptotic distributions of the estimators of methods A and B are derived. Consider first method B.

We write the estimators

$$\tilde{p}_{1...} = \frac{(n_M/n)(n_{M1...}\ /n_M) + (n_S/n)(n_{S1...}\ /n_S)}{(n_M/n) + (n_S/n)}$$

$$\tilde{p}_{..1.} = \frac{(n_M/n)(n_{M..1.}\ /n_M) + (n_T/n)(n_{T..1.}\ /n_T)}{(n_M/n) + (n_T/n)}$$

For simplicity of notation we define $\tilde{\underline{n}}_M = (n_{M0.0.}, n_{M0.1.}, n_{M1.0.}, n_{M1.1.})'$ and define $\tilde{\underline{n}}_S$, $\tilde{\underline{n}}_T$, $\tilde{\underline{n}}_R$ and $\tilde{\underline{p}}$ similarly. From the central limit theorem

$$n^{1/2} \begin{vmatrix} n_M^{-1}(\tilde{\underline{n}}_M - n_M\tilde{\underline{p}}) \\ n_S^{-1}(\tilde{\underline{n}}_S - n_S\tilde{\underline{p}}) \\ n_T^{-1}(\tilde{\underline{n}}_T - n_T\tilde{\underline{p}}) \end{vmatrix} \rightarrow N(0, \Omega_B)$$

where Ω_B is block diagonal with elements $\alpha_M^{-1}\Omega$, $\alpha_S^{-1}\Omega$ and $\alpha_T^{-1}\Omega$ where $\Omega = D_{\tilde{p}} - \tilde{\underline{p}}\,\tilde{\underline{p}}'$. $D_{\tilde{p}}$ is the 4x4 diagonal matrix with non-zero elements equal to the elements of the vector $\tilde{\underline{p}}$.

The estimators $\tilde{p}_{1...}$ and $\tilde{p}_{..1.}$ are linear functions of the stochastic variables $(\underline{n}_M, \underline{n}_S, \underline{n}_T)'$. We define a 12x1 vector $\underline{w}_1 = (w_{1,1}, \ldots, w_{1,12})'$ so that

$$w_{1,3} = w_{1,4} = \alpha_M/(\alpha_M + \alpha_S)$$

$$w_{1,7} = w_{1,8} = \alpha_S/(\alpha_M + \alpha_S)$$

and $w_{1,i} = 0$ otherwise. Similarly we define $\underline{w}_2$ by setting the elements equal to 0 except

$$w_{2,2} = w_{2,4} = \alpha_M/(\alpha_M + \alpha_T)$$

$$w_{2,10} = w_{2,12} = \alpha_T/(\alpha_M + \alpha_T)$$

Then $n^{1/2}(\tilde{p}_{1...} - p_{1...}, \tilde{p}_{..1.} - p_{..1.})'$ converges towards a bivariate Gaussian random variable with mean 0 and covariance matrix

$$\{\sigma_{ijB}\}_{i,j=0,1} = \begin{pmatrix} \underline{w}_1' \\ \underline{w}_2' \end{pmatrix} \Omega_B \, (\underline{w}_1 \vdots \underline{w}_2).$$

Remark that Ω_B/n is the exact covariance matrix of $\tilde{p}_{1...}$ and $\tilde{p}_{..1.}$.

Let us now consider method A. The estimators $\hat{p}_{1...}$ and $\hat{p}_{..1.}$ are not linear functions of the stochastic variables $\underline{n}_M$, $\underline{n}_S$, $\underline{n}_T$ and $\underline{n}_R$. The δ-method is based on using a linear approximation. This program can be carried out in the present context. The details are rather lengthy and tedious so we present only the main steps. Introduce the 1x64 vector $\hat{\underline{\theta}}$ with elements given by $\underline{n}_M/n_M$, $\underline{n}_S/n_S$, $\underline{n}_T/n_T$ and $\underline{n}_R/n_R$. Let $\underline{\theta} = E\hat{\underline{\theta}}$. We can then write

$$f_1(\underline{\theta}) = p_{1...} = (a_{11}/b_{11})c_{11} + (a_{12}/b_{12})c_{12} = n\, h_1(\underline{\theta})/N$$

$$f_2(\underline{\theta}) = p_{..1.} = (a_{21}/b_{21})c_{21} + (a_{22}/b_{22})c_{22} = n\, h_2(\underline{\theta})/N$$

where

$$a_{11} = (n_M/n)(\theta_9 + \ldots + \theta_{12}) + (n_S/n)(\theta_{25} \ldots + \theta_{28}),$$

$$b_{11} = (n_M/n)(\theta_1 + \ldots + \theta_4 + \theta_9 + \ldots + \theta_{12}) + (n_S/n)(\theta_{17} + \ldots + \theta_{20} + \theta_{25} + \ldots + \theta_{28}),$$

$$c_{11} = b_{11} + (n_T/n)(\theta_{33} + \ldots + \theta_{36} + \theta_{41} + \ldots + \theta_{44}) + (n_R/n)(\theta_{49} + \ldots + \theta_{52} + \theta_{57} + \ldots + \theta_{60}),$$

and

$$a_{12} = (n_M/n)(\theta_{13} + \ldots + \theta_{16}) + (n_S/n)(\theta_{29} + \ldots + \theta_{32}),$$

$$b_{12} = (n_M/n)(\theta_5 + \ldots + \theta_8 + \theta_{13} + \ldots + \theta_{16}) + (n_S/n)(\theta_{21} + \ldots + \theta_{24} + \theta_{29} + \ldots + \theta_{32}),$$

$$c_{12} = b_{12} + (n_T/n)(\theta_{37} + \ldots + \theta_{40} + \theta_{45} + \ldots + \theta_{48}) + (n_R/n)(\theta_{53} + \ldots + \theta_{56} + \theta_{61} + \ldots + \theta_{64})$$

with similar expressions for a_{21}, b_{21}, c_{21}, a_{22} , b_{22} , c_{22}.

Carrying out the differentiations and inserting we get

$$dh_1/d\theta_i = -[\alpha_M(\alpha_T + \alpha)/(\alpha_M + \alpha_S)]\, p_{10..}/p_{.0..}, \quad i=1,\ldots,4,$$

$$= -[\alpha_M(\alpha_T + \alpha)/(\alpha_M + \alpha_S)]\, p_{11..}/p_{.1..}, \quad i=5,\ldots,8,$$

$$= [\alpha_M/(\alpha_M + \alpha_S)]\,[1 - (\alpha_T + \alpha)p_{10..}/p_{.0..}], \quad i=9,\ldots,12,$$

$$= [\alpha_M/(\alpha_M + \alpha_S)]\,[1 - (\alpha_T + \alpha)p_{11..}/p_{.1..}], \quad i=13,\ldots,16,$$

$$= -[\alpha_S(\alpha_T + \alpha)/(\alpha_M + \alpha_S)]\, p_{10..}/p_{.0..}, \quad i=17,\ldots,20,$$

$$= -[\alpha_S(\alpha_T + \alpha)/(\alpha_M + \alpha_S)]\, p_{11..}/p_{.1..}, \quad i=21,\ldots,24,$$

$$= [\alpha_S/(\alpha_M + \alpha_S)]\,[1 - (\alpha_T + \alpha)p_{10..}/p_{.0..}], \quad i=25,\ldots,28,$$

$$= [\alpha_S/(\alpha_M + \alpha_S)]\,[1 - (\alpha_T + \alpha)p_{11..}/p_{.1..}], \quad i=29,\ldots,32,$$

$$= \alpha_T\, p_{10..}/p_{.0..}, \quad i=33,\ldots,36,41,\ldots,44,$$

$$= \alpha_T\, p_{11..}/p_{.1..}, \quad i=37,\ldots,40,45,\ldots,48,$$

$= \alpha\ p_{10..}/p_{.0..}$, $i=49,\ldots,52,57,\ldots,60$,

$= \alpha\ p_{11..}/p_{.1..}$, $i=53,\ldots,56,60,\ldots,64$

and

$dh_2/d\theta_i = -[\alpha_M(\alpha_S + \alpha)/(\alpha_M + \alpha_T)]\ p_{..10}/p_{...0}$, $i=1,5,9,13$,

$= -[\alpha_M(\alpha_S + \alpha)/(\alpha_M + \alpha_T)]\ p_{..11}/p_{...1}$, $i=2,6,10,14$,

$= [\alpha_M/(\alpha_M + \alpha_T)]\ [1-(\alpha_S+\alpha)p_{..10}/p_{...0}]$, $i=3,7,11,15$,

$= [\alpha_M/(\alpha_M + \alpha_T)]\ [1-(\alpha_S+\alpha)p_{..11}/p_{...1}]$, $i=4,8,12,16$,

$= \alpha_S\ p_{..10}/p_{...0}$, $i=17,19,21,23,25,27,29,31$,

$= \alpha_S\ p_{..11}/p_{...1}$, $i=18,20,22,24,26,28,30,32$,

$= -[\alpha_T(\alpha_S + \alpha)/(\alpha_M + \alpha_T)]\ p_{...10}/p_{..0}$, $i=33,37,41,45$,

$= -[\alpha_T(\alpha_S + \alpha)/(\alpha_M + \alpha_T)]\ p_{..11}/p_{..1}$, $i=34,38,42,46$,

$= [\alpha_T/(\alpha_M + \alpha_T)]\ [1-(\alpha_S+\alpha)p_{..10}/p_{...0}]$, $i=35,39,43,47$,

$= [\alpha_T/(\alpha_M + \alpha_T)]\ [1-(\alpha_S+\alpha)p_{..11}/p_{...1}]$, $i=36,40,44,48$,

$= \alpha\ p_{..10}/p_{...0}$, $i=49,51,53,55,57,59,61,63$,

$= \alpha\ p_{..11}/p_{...1}$, $i=50,52,54,56,58,60,62,64$.

By the central limit theorem

$$n^{-1/2}\left|\begin{array}{c} n_M^{-1}(\underline{n}_M - n_M\underline{p}) \\ n_S^{-1}(\underline{n}_S - n_S\underline{p}) \\ n_T^{-1}(\underline{n}_T - n_T\underline{p}) \\ n_R^{-1}(\underline{n}_R - n_R\underline{p}) \end{array}\right| \rightarrow N(0,\Omega_A)$$

where Ω_A is a block diagonal matrix whith matrices $\alpha_M^{-1}\Omega$, $\alpha_S^{-1}\Omega$, $\alpha_T^{-1}\Omega$ and $\alpha_R^{-1}\Omega$ along the diagonal. Here $\Omega = D_p - \underline{p}\ \underline{p}'$ where $\underline{p}$ is the 16x1 vector having as elements $p_{iji'j'}$ lexicographically ordered, and D_p is the diagonal matrix with the elements of $\underline{p}$ on the diagonal. Since $\hat{p}_{1...} = f_1(\hat{\underline{\theta}}) = f_1(n_M^{-1}\underline{n}_M, n_S^{-1}\underline{n}_S, n_T^{-1}\underline{n}_T, n_R^{-1}\underline{n}_R)$ and $\hat{p}_{..1.} = f_2(\hat{\underline{\theta}}) = f_2(n_M^{-1}\underline{n}_M, n_S^{-1}\underline{n}_S, n_T^{-1}\underline{n}_T, n_R^{-1}\underline{n}_R)$ it follows by the δ-method that

$$n^{-1/2}\left|\begin{array}{c} (\hat{p}_{1...} - p_{1...}) \\ (\hat{p}_{..1.} - p_{..1.}) \end{array}\right|$$

converges towards a bivariate Gaussian random variable with mean 0 and covariance matrix { σ_{ijA} }. Denote by $\underline{u}_1$, $\underline{u}_2$, $\underline{u}_3$ and $\underline{u}_4$ the 16x1 vectors having elements $dh_1/d\theta_i$, i=1,...,16; $dh_1/d\theta_i$, i=17,...,32; $dh_1/d\theta_i$, i=33,...,48; $dh_1/d\theta_i$, i=49,...,64. Define vectors $\underline{v}_1$, $\underline{v}_2$, $\underline{v}_3$ and $\underline{v}_4$ similarly with respect to the derivatives of h_2.

Then
$$\sigma_{12A} = \{ \alpha_M^{-1} \underline{u}_1' D_p \underline{v}_1 + \alpha_S^{-1} \underline{u}_2' D_p \underline{v}_2 + \alpha_T^{-1} \underline{u}_3' D_p \underline{v}_3 + \alpha^{-1} \underline{u}_4' D_p \underline{v}_4 - \alpha_M^{-1}(\underline{u}_1' \underline{p})(\underline{v}_1' \underline{p}) - \alpha_S^{-1}(\underline{u}_2' \underline{p})(\underline{v}_2' \underline{p}) - \alpha_T^{-1}(\underline{u}_3' \underline{p})(\underline{v}_3' \underline{p}) - \alpha^{-1}(\underline{u}_4' \underline{p})(\underline{v}_4' \underline{p}) \}/(1 + \alpha)^2$$

which after some straightforward calculation gives the result claimed in section 2.

References.

Chen, T. T. (1979). Log-linear models for categorical data with misclassifications and double sampling. J. Amer. Statist. Assoc. 74 481-488.

Chen, T. T. and Fienberg, S. E. (1974). Two-dimensional contingencey tables with both completely and partially cross-classified data. Biometrics 30 629-642.

Chen, T. T. and Fienberg, S. E. (1976). The analysis of contingency tables with incompletely classified data. Biometrics 32 133-144.

Espeland, M. A. and Odoroff, C. L. (1985). Log-linear models for doubly sampled categorical data fitted by the EM algoritm. J. Amer. Statist. Assoc. 80 663-670.

Palmgren, J. (1987). Precision of double sampling estimators for comparing two probabilities. Biometrika 74 687-694.

Rao, C. R. (1973). Linear Statistical Inference and Its Applications. 2'nd ed. Wiley, New York.

Rosenqvist, G. (1986). Micro and macro data in statistical inference on Markov chains. Thesis, Swedish School of Economics and Business Administration, Helsinki.

Tenenbein, A. (1970). A double sampling scheme for estimating from binomial data with misclassifications. J. Amer. Statist. Assoc. 65 1350-1361.

Tenenbein, A. (1972). A double sampling scheme for estimating from misclassified multinomial data with application to sampling inspection. Technometrics 14 187-202.

ISSUED IN THE SERIES DISCUSSION PAPER

No. 1 I. Aslaksen and O. Bjerkholt: Certainty Equivalence Procedures in the Macroeconomic Planning of an Oil Economy.

No. 3 E. Biørn: On the Prediction of Population Totals from Sample surveys Based on Rotating Panels.

No. 4 P. Frenger: A Short Run Dynamic Equilibrium Model of the Norwegian Prduction Sectors.

No. 5 I. Aslaksen and O. Bjerkholt: Certainty Equivalence Procedures in Decision-Making under Uncertainty: an Empirical Application.

No. 6 E. Biørn: Depreciation Profiles and the User Cost of Capital.

No. 7 P. Frenger: A Directional Shadow Elasticity of Substitution.

No. 8 S. Longva, L. Lorentsen, and Ø. Olsen: The Multi-Sectoral Model MSG-4, Formal Structure and Empirical Characteristics.

No. 9 J. Fagerberg and G. Sollie: The Method of Constant Market Shares Revisited.

No.10 E. Biørn: Specification of Consumer Demand Models with Stocahstic Elements in the Utility Function and the first Order Conditions.

No.11 E. Biørn, E. Holmøy, and Ø. Olsen: Gross and Net Capital, Productivity and the form of the Survival Function . Some Norwegian Evidence.

No.12 J. K. Dagsvik: Markov Chains Generated by Maximizing Components of Multidimensional Extremal Processes.

No.13 E. Biørn, M. Jensen, and M. Reymert: KVARTS - A Quarterly Model of the Norwegian Economy.

No.14 R. Aaberge: On the Problem of Measuring Inequality.

No.15 A-M. Jensen and T. Schweder: The Engine of Fertility - Influenced by Interbirth Employment.

No.16 E. Biørn: Energy Price Changes, and Induced Scrapping and Revaluation of Capital - A Putty-Clay Approach.

No.17 E. Biørn and P. Frenger: Expectations, Substitution, and Scrapping in a Putty-Clay Model.

No.18 R. Bergan, Å. Cappelen, S. Longva, and N. M. Stølen: MODAG A - A Medium Term Annual Macroeconomic Model of the Norwegian Economy.

No.19 E. Biørn and H. Olsen: A Generalized Single Equation Error Correction Model and its Application to Quarterly Data.

No.20 K. H. Alfsen, D. A. Hanson, and S. Glomsrød: Direct and Indirect Effects of reducing SO_2 Emissions: Experimental Calculations of the MSG-4E Model.

No.21 J. K. Dagsvik: Econometric Analysis of Labor Supply in a Life Cycle Context with Uncertainty.

No.22 K. A. Brekke, E. Gjelsvik, B. H. Vatne: A Dynamic Supply Side Game Applied to the European Gas Market.

No.23 S. Bartlett, J. K. Dagsvik, Ø. Olsen and S. Strøm: Fuel Choice and the Demand for Natural Gas in Western European Households.

No.24 J. K. Dagsvik and R. Aaberge: Stochastic Properties and Functional Forms in Life Cycle Models for Transitions into and out of Employment.

No.25 T. J. Klette: Taxing or Subsidising an Exporting Industry.

No.26 K. J. Berger, O. Bjerkholt and Ø. Olsen: What are the Options for non-OPEC Producing Countries.

No.27 A. Aaheim: Depletion of Large Gas Fields with Thin Oil Layers and Uncertain Stocks.

No.28 J. K. Dagsvik: A Modification of Heckman's Two Stage Estimation Procedure that is Applicable when the Budget Set is Convex.

No.29 K. Berger, Å. Cappelen and I. Svendsen: Investment Booms in an Oil Economy - The Norwegian Case.

No.30 A. Rygh Swensen: Estimating Change in a Proportion by Combining Measurements from a True and a Fallible Classifier.