No. 220 (84-15)

Use of AICs in log linear model for contingency tables with Poisson and Multinomial designs

by

Sinji TSURUTA and Yoshiko NOGAMI

March 1984

Use of AICs in log linear model for contingency tables with Poisson and Multinomial designs.

by Sinji Tsuruta*) and Yoshiko Nogami

Sony Coop. Inc. and University of Tsukuba

1. Introduction.

Let $\underline{x} = (x_1, \ldots, x_n)'$ be a random sample of size n with each x having the probability density function $f(x;\theta)$. Let $L_{\underline{X}}(\theta) = \Pi_{i=1}^{n} f(x_i;\theta)$. Hereafter, we do not exhibit the subscript $\underline{X}$ in $L_{\underline{X}}(\theta)$ unless we need it. Akaike(1974) proposes the following information criterion:

(1) $\qquad \mathrm{AIC} = -2 \log L(\hat{\theta}_n) + 2c$

where $\hat{\theta}_n$ is the maximum likelihood estimate for θ. Let $\underline{y} = (y_1, \ldots, y_n)'$ be the future observations having the likelihood $\Pi_{i=1}^{n} f(y_i;\theta_0)$ where θ_0 is a true parameter. Noticing that 2c is an asymptotic bias removing from $-2 \log L(\hat{\theta}_n)$ we have that as $n \to \infty$, $-(2n)^{-1}(\mathrm{AIC}) \to I_n = E_{\underline{X}}[\int \log \Pi_{i=1}^{n} f(y_i;\hat{\theta}_n) \cdot \Pi_{i=1}^{n} f(y_i;\theta_0)\, d\underline{y}\,]$. Hence, Akaike's idea is to choose $\hat{\theta}_n$ which maximizes I_n's asymptotically unbiased estimates $-(2n)^{-1}(\mathrm{AIC})$ or equivalently minimizes AIC. This idea is equivalent to choosing $\hat{\theta}_n$ which maximizes I_n. These ideas coincide with minimizing the expectation of Kullback's discrimination information function;

$$E_{\underline{X}}[\int \log \Pi_{i=1}^{n}(f(y_i;\theta_0)/f(y_i;\hat{\theta}_n))\ \Pi_{i=1}^{n} f(y_i;\theta_0)\, d\underline{y}\,].$$

---

*) This paper is a refined version of part of Tsuruta's graduation thesis supervised by Dr. Y. Nogami for his Bachelor's degree at University of Tsukuba.

(These discussions are due to Sugiura(1978).)

Sugiura(1978) proposes a corrected (c-) AIC computing the exact bias directly. Namely, the corrected AIC is defined as follows:

(2) $$\text{c-AIC} = -2 \log L(\hat{\theta}_n) + 2d$$

where

$$d = E_{\underline{X}}[\log L(\hat{\theta}_n)] - I_n.$$

In this paper, we consider the log linear model for contingency tables under three sampling designs relating to Poisson and Multinomial distributions which are fully demonstrated by Bishop, Yvonne M. M., Fienberg, Stephen E. and Holland, Paul W. (1977), obtain the formulas for AIC and c-AIC , and demonstrate those by examples. Although we only treat 2-way contingency tables for complete data throughout the paper, we can proceed in the same way for 3 or more way contingency tables under either complete or incomplete data.

In Section 2 we exhibit the hypotheses under three sampling designs and in Section 3 we rephrase above hypotheses by unified hypothesis H_{12} under the log linear model. Section 4 states the likelihood ratio statistic z, quoted from E. B. Andersen(1980), testing independence (H_{12}). In Section 5 we consider AIC and in Section 6 we work for c-AIC, for testing H_{12} in both sections. In Section 7 we exhibit z-test from Andersen(1980) and find c-AIC for testing marginal effects. Section 8 gives three examples.

Main development is Sections 6 and 7.

We remark that S. Goto, S. Hatanaka and T. Tasaki(1980) exhibit AIC in log linear model for binary data with logistic transformation.

2. Three Sampling Designs.

From this section until the end of this paper we consider 2-way contingency table whose schematic form is shown in Table I below.

		Variable 2: 1 2 ... j ... J	Totals
	1		$x_{1.}$
	⋮	⋮	⋮
Variable 1	i	...... x_{ij}	$x_{i.}$
	⋮	⋮	⋮
	I		$x_{I.}$
Totals		$x_{.1}$ ... $x_{.j}$ ... $x_{.J}$	$x_{..}$

Table I

We consider tests of independence in a contingency table under three different sampling designs.

[Design I] X_{ij}'s are independently distributed according to Poisson distribution $P_o(\lambda_{ij})$ with parameter λ_{ij}.

[Design II] When the total count $x_{..}=n$ is fixed, $(X_{11}, \ldots, X_{IJ})$ are distributed according to Multinomial distribution $M(n;p_{11}, \ldots, p_{IJ})$ where p_{ij}'s are parameter such that $\Sigma_i \Sigma_j \, p_{ij} = 1$.

[Design III] When marginal counts $x_{i.} = n_i$ for $i=1, \ldots, I$ are fixed, $(X_{i1}, \ldots, X_{iJ})$ for $i= 1, \ldots, I$ are independent row observations and for each i, $(X_{i1}, \ldots, X_{iJ})$ is distributed according to Multinomial distribution $M(n_i;p^{\#}_{i1}, \ldots, p^{\#}_{iJ})$ where $\Sigma_{j=1}^{J} p^{\#}_{ij} =1$.

The null hypotheses for testing independence under above three designs respectively become as follows:

(3) $H_0^{I}: \lambda_{ij}=\lambda_{i.}\lambda_{.j}/\lambda_{..}$ where $\lambda_{i.}=\Sigma_{j=1}^{J}\lambda_{ij}$, $\lambda_{.j}=\Sigma_{i=1}^{I}\lambda_{ij}$ and $\lambda_{..}=\Sigma_i\Sigma_j\lambda_{ij}$,

(4) $H_0^{II}: p_{ij}=p_{i.}p_{.j}$

and

(5) $H_0^{III}: p_{ij}^{\#} = p_{.j}^{\#}/I$,

where the superscripts of H_0 mean Designs.

3. Log Linear Model.

Let $u_{12(ij)}$, $u_{1(i)}$, $u_{2(j)}$ and u be parameters satisfying the model

(6) $m_{ij}= E(X_{ij})= \exp\{u + u_{1(i)} + u_{2(j)} + u_{12(ij)}\}$

or equivalently,

(7) $m_{ij}^{*} = \log m_{ij} = u + u_{1(i)} + u_{2(j)} + u_{12(ij)}$

where

$$\Sigma_i u_{1(i)} = \Sigma_j u_{2(j)} = \Sigma_i u_{12(ij)} = \Sigma_j u_{12(ij)} = 0.$$

The basic decomposition of m_{ij}^{*} makes no assumptions about the m_{ij} and is called the saturated model. For above three sampling designs in Section 2, we have $m_{ij}=\lambda_{ij}=np_{ij} = n_i p_{ij}^{\#}$, respectively.

Letting

(8) $\overline{m^{*}_{i.}} = J^{-1}\Sigma_{j=1}^{J} m_{ij}^{*}$, $\overline{m^{*}_{.j}} = I^{-1}\Sigma_{i=1}^{I} m_{ij}^{*}$ and $\overline{m^{*}_{..}} = (IJ)^{-1}\Sigma_i\Sigma_j m_{ij}^{*}$

we know by e.g. Andersen(1980) that all of three hypotheses (3), (4) and (5) are equivalently expressed by

(9) $m_{ij}=m_{i.}m_{.j}/m_{..}$

or equivalently,

(10) H_{12}: $u_{12(ij)} = 0$ for all i and j

in the log linear model where

(11) $u_{12(ij)} = m^*_{ij} - \overline{m^*_{i.}} - \overline{m^*_{.j}} + \overline{m^*_{..}}$.

If the test has shown that all the interactions are 0, we may be interested in testing the hypotheses concerning the row and column effects.

By letting

(12) $u_{1(i)} = \overline{m^*_{i.}} - \overline{m^*_{..}}$, $u_{2(j)} = \overline{m^*_{.j}} - \overline{m^*_{..}}$ and $u = \overline{m^*_{..}}$,

the most common null hypotheses are

(13) H_1: $u_{1(i)} = 0$ for all i

and

(14) H_2: $u_{2(j)} = 0$ for all j.

In Sections 4, 5 and 6 we consider following three criterions for testing H_{12}: the likelihood ratio statistic z (χ^2 test), AIC and c-AIC, respectively.

4. z-test for testing H_{12} vs $\overline{H}_{12}$.

Let $\underline{u} = (u; u_{1(i)}, i=1, \ldots, I; u_{2(j)}, j=1, \ldots, J; u_{12(ij)}, i=1, \ldots, I, j=1, \ldots, J)'$. By using new parameters u, $u_{1(i)}$, $u_{2(j)}$ and $u_{12(ij)}$ the log-likelihood functions in three sampling designs are written in the form

(15)
$$\log L(\underline{u}) = x_{..}u + \Sigma_i x_{i.} u_{1(i)} + \Sigma_j x_{.j} u_{2(j)} + \Sigma_i \Sigma_j x_{ij} u_{12(ij)} - c(\underline{u}) - h(\underline{x})$$

where $c(\underline{u})$ and $h(\underline{x})$ are some functions of $\underline{u}$ and $\underline{x}$, respectively. By Theorem 5.7(Andersen(1980)) MLE's $\hat{\underline{u}}^0$ for $\underline{u}$ under H_{12} are given by

(16) $$\hat{u}^0_{1(i)} = \log x_{i.} - I^{-1}\Sigma_{i=1}^{I}\log x_{i.},\quad \hat{u}^0_{2(j)} = \log x_{.j} - J^{-1}\Sigma_{j=1}^{J}\log x_{.j},$$
$$\hat{u}^0 = I^{-1}\Sigma_{i=1}^{I}\log x_{i.} + J^{-1}\Sigma_{j=1}^{J}\log x_{.j} - \log x_{..}.$$

Also, from Theorem 5.5 (Andersen(1980)), by denoting $\overline{H}_{12}$ as a compliment of H_{12} MLE's $\underline{\hat{u}}$ for $\underline{u}$ under $\overline{H}_{12}$ are given as follows:

(17) $$\hat{u}_{12(ij)} = x^*_{ij} - \overline{x^*_{i.}} - \overline{x^*_{.j}} + \overline{x^*_{..}},\quad \hat{u}_{1(i)} = \overline{x^*_{i.}} - \overline{x^*_{..}}$$
$$\hat{u}_{2(j)} = \overline{x^*_{.j}} - \overline{x^*_{..}},\quad \text{and}\quad \hat{u} = \overline{x^*_{..}}\ .$$

Inserting (16) and (17) into (15) leads to the likelihood ratio statistic

$$z = -2(\log L(\underline{\hat{u}}^0) - \log L(\underline{\hat{u}}))$$
(18) $$= 2\,\Sigma_i\Sigma_j\, x_{ij}\,\{\log x_{ij} - \log(x_{i.}x_{.j}/x_{..})\}.$$

Hence, at a level of significance α, the test is to reject H_{12} if $z > \chi^2((I-1)\cdot(J-1))$. Note that for Design II, $x_{..}$ may be replaced by n and for Design III $x_{i.}$ and $x_{..}$ may replaced by n_i and n, respectively.

5. AIC for testing H_{12} vs $\overline{H}_{12}$.

Letting AIC_{12} and AIC_0 be AIC's under H_{12} and $\overline{H}_{12}$, respectively. Testing the hypotheses using AIC is to reject H_{12} if $AIC_{12} - AIC_0 > 0$. In the definition (1) of AIC, let c_{12} and c_0 be c's under H_{12} and $\overline{H}_{12}$, respectively. c here represents the number of parameters whose values can be changed freely. Then, in any design,

(19) $$AIC_{12} - AIC_0 = z + 2(C_{12} - c_0) = z-2(I-1)(J-1)\ (= z \doteqdot 2E(z))$$

where the second equality follows from Table II below.

	Design I	II	III
c_{12}	(I-1)+(J-1)+1	(I-1)+(J-1)	J-1
c_0	IJ	IJ-1	I(J-1)

Table II.

Hence, testing hypotheses using AIC is to reject H_{12} if $z > 2E(z)$.

6. c-AIC for testing H_{12} vs $\bar{H}_{12}$.

In the similar fashion to above Section 5, in case of c-AIC , by letting d_{12} and d_0 be d's under H_{12} and $\bar{H}_{12}$, respectively, the test is to reject H_{12} if

(20) $$\text{c-AIC}_{12} - \text{c-AIC}_0 = z + 2(d_{12}-d_0) > 0$$

where

(21) $$d_{12} = E_{\underline{X}}(\log L(\hat{\underline{u}}^0)) - E_{\underline{X}}E_{\underline{Y}}(\log L_{\underline{Y}}(\hat{\underline{u}}^0))$$

and

$$d_0 = E_{\underline{X}}(\log L(\hat{\underline{u}})) - E_{\underline{X}}E_{\underline{Y}}(\log L_{\underline{Y}}(\hat{\underline{u}})).$$

Applying (15) with (16), (6) and Table III below for estimates m_{ij}'s under H_{12} we can easily obtain

(22) $$d_{12} = \Sigma_i\Sigma_j E[(x_{ij} - m_{ij})\log \hat{m}^0_{ij}]$$

where $\hat{m}^0_{ij}$ is the MLE for m_{ij} under H_{12}. Similarly,

(23) $$d_0 = \Sigma_i\Sigma_j E[(x_{ij} - m_{ij})\log \hat{m}_{ij}]$$

where $\hat{m}_{ij}$ is the MLE for m_{ij} under $\bar{H}_{12}$.

To obtain d_{12} and d_0 we need to consider the sampling distributions of x_{ij}. Now, from (16) and (17) and by using the inverse transformation (6), we can easily get the following estimators for m_{ij}, $m_{i.}$ and $m_{.j}$ in Table III:

	m_{ij}	$m_{i.}$	$m_{.j}$
$\overline{H}_{12}$	x_{ij}		
		$x_{i.}$	$x_{.j}$
H_{12}	$x_{i.}\, x_{.j}/x_{..}$		

Table III. Estimators.

For Design I, we have (22) with m_{ij} replaced by λ_{ij}. Substituting $\hat{\lambda}^{0}_{ij} = x_{i.}x_{.j}/x_{..}$ from Table III leads to

(24) $$d_{12} = \Sigma_i E_{\underline{X}}((x_{i.}-\lambda_{i.})\log x_{i.}) + \Sigma_j E_{\underline{X}}((x_{.j}-\lambda_{.j})\log x_{.j}) - E_{\underline{X}}((x_{..}-\lambda_{..})\log x_{..}).$$

Expanding log w about the neighborhood of $w=\eta$, we get

$$\log w \simeq \log \eta + (w-\eta)\eta^{-1} - 2^{-1}(w-\eta)^2\eta^{-2} + 3^{-1}(w-\eta)^3\eta^{-3}$$

and hence

(25) $$E((w-\eta)\log w) \simeq \log \eta . E(w-\eta) + E\{(w-\eta)/\sqrt{\eta}\}^2 - (2\eta^{\frac{1}{2}})^{-1}E\{(w-\eta)/\sqrt{\eta}\}^3 + (3\eta)^{-1}E\{(w-\eta)/\sqrt{\eta}\}^4.$$

Substituting moments for Poisson in Table IV below with $\eta=\lambda$ into (25) gives

(26) $$E((w-\lambda)\log w) \simeq 1 + \frac{1}{2\lambda} + \frac{1}{3\lambda\sqrt{\lambda}} .$$

	$E(w)$	$E((w-\eta)/\sqrt{\eta})^2$	$E((w-\eta)/\sqrt{\eta})^3$	$E((w-\eta)/\sqrt{\eta})^4$
Poisson ($P_o(\eta)$)	η	1	$\eta^{-\frac{1}{2}}$	$3+\eta^{-\frac{1}{2}}$
Binomial ($B_i(n,p)$)	$np(=\eta)$	$1-p$	$\frac{1-2p}{\sqrt{np(1-p)}}$	$3-6n^{-1}+\frac{1}{np(1-p)}$

Table IV. Moments.

Thus, applying (26) three times to the rhs of (24) leads to

(27) $$d_{12}-c_{12} \simeq 2^{-1}\{\Sigma_i\lambda_{i.}^{-1}+\Sigma_j\lambda_{.j}^{-1}-\lambda_{..}^{-1}\}+3^{-1}\{\Sigma_i\lambda_{i.}^{-3/2}+\Sigma_j\lambda_{.j}^{-3/2}-\lambda_{..}^{-3/2}\}.$$

Similarly, we obtain

(28) $$d_0-c_0 \simeq 2^{-1}\Sigma_i\Sigma_j\lambda_{ij}^{-1}+3^{-1}\Sigma_i\Sigma_j\lambda_{ij}^{-3/2}.$$

In the similar fashion to above, for Design II, since by (25) and Table IV $E((w-np)\log(w/n)) \simeq 1-p+\frac{1-p}{2np}+O(n^{-2})$, using this relation several times gives

(29) $$d_{12}-c_{12} \simeq 2^{-1}\{\Sigma_i(1-p_{i.})/(np_{i.})+\Sigma_j(1-p_{.j})/(np_{.j})\}+O(n^{-2})$$

and

(30) $$d_0-c_0 \simeq 2^{-1}\Sigma_i\Sigma_j(1-p_{ij})/(np_{ij})+O(n^{-2}).$$

For Design III, we can proceed in the same way as above by using the fact that $X_{i.} \sim Bi(n,J^{-1}p^{\#}_{i.})$ and $X_{.j} \sim Bi(n,I^{-1}p^{\#}_{.j})$. Then, we get

(31) $$d_{12}-c_{12} \simeq 2^{-1}\Sigma_j(I-p^{\#}_{.j})/(np^{\#}_{.j})+O(n^{-2})$$

and

(32) $$d_0-c_0 \simeq 2^{-1}\Sigma_i\Sigma_j(1-p^{\#}_{ij})/(np^{\#}_{ij})+O(n^{-2}).$$

From (20) and (27) through (32), the test using c-AIC is to reject H_{12} if

(33) $$z-2E(z) > -\Sigma_{i=1}^{I}\lambda_{i.}^{-1}-\Sigma_{j=1}^{J}\lambda_{.j}^{-1}+\lambda_{..}^{-1}+\Sigma_i\Sigma_j\lambda_{ij}^{-1} \qquad \text{for Design I}$$

or

(34) $$> -\Sigma_{i=1}^{I}\frac{1-p_{i.}}{np_{i.}}-\Sigma_{j=1}^{J}\frac{1-p_{.j}}{np_{.j}}+\Sigma_i\Sigma_j\frac{1-p_{ij}}{np_{ij}} \qquad \text{for Design II}$$

or

(35) $$> -\Sigma_{j=1}^{J}\frac{I-p^{\#}_{.j}}{np^{\#}_{.j}}+\Sigma_i\Sigma_j\frac{1-p^{\#}_{ij}}{np^{\#}_{ij}} \qquad \text{for Design III}$$

where for the parameters on the rhs's we substitute the values of the estimators from Table III.

7. Tests for H_1 and for H_2.

In this section, we consider z-test and c-AIC for testing the hypotheses H_1: $u_{1(i)}=0$ for all i against H_{12} and for testing the hypotheses H_2: $u_{2(j)}=0$ for all j against H_{12}. In sampling designs I, II and III corresponding H_1 and H_2 are shown as in Table V, below.

		H_1	H_2
	I	$\lambda_{i.}=\lambda_{..}/I, \forall i$	$\lambda_{.j}=\lambda_{..}/J, \forall j$
Designs	II	$p_{i.}=I^{-1}, \forall i$	$p_{.j}=J^{-1}, \forall j$
	III	———	$p^{\#}_{.j}=J^{-1}, \forall j$

Table V. Hypotheses.

In the similar fashion to (18) in Section 4, we can see from Theorem 5.9 (Andersen(1980)) that the LR-test for H_1 vs H_{12} is

(36) $$z_1 = 2\Sigma_{i=1}^{I} x_{i.}\{\log x_{i.} - \log(x_{..}/I)\}$$

which is asymptotically $\chi^2(I-1)$ under H_1 and the LR-test for H_2 vs H_{12} is

(37) $$z_2 = 2\Sigma_{j=1}^{J} x_{.j}\{\log x_{.j} - \log(x_{..}/J)\}$$

which is asymptotically $\chi^2(J-1)$ under H_2.

In the definition of c-AIC, let d_1 and d_2 be d's under H_1 and under H_2, respectively.

For Design I, in view of (22) and (23), we can easily check

(38) $$d_1 = \Sigma_j E_{\underline{X}}\{(x_{.j} - \lambda_{.j}) \log x_{.j}\} = J + 2^{-1}\Sigma_j\lambda_{.j}^{-1} + 3^{-1}\Sigma_j\lambda_{.j}^{-3/2}$$

and

(39) $$d_2 = \Sigma_i E_{\underline{X}}\{(x_{i.} - \lambda_{i.})\log x_{i.}\} = I + \tfrac{1}{2}\Sigma_i \lambda_{i.}^{-1} + 3^{-1}\Sigma_i \lambda_{i.}^{-3/2}.$$

In the definition (1) of AIC , let c_1 and c_2 be c's under H_1 and H_2, respectively. c_1 and c_2 become as follows:

	Designs		
	I	II	III
c_1	J	J-1	___
c_2	I	I-1	0

Table VI. D. F.'s

(40) $$d_1 - c_1 \simeq 2^{-1}\Sigma_j\lambda_{.j}^{-1} + 3^{-1}\Sigma_j\lambda_{.j}^{-3/2}$$

and

(41) $$d_2 - c_2 \simeq 2^{-1}\Sigma_i\lambda_{i.}^{-1} + 3^{-1}\Sigma_i\lambda_{i.}^{-3/2}.$$

Similarly, for Design II,

(42) $$d_1 - c_1 \simeq 2^{-1}\Sigma_j(1-p_{.j})/(np_{.j})$$

and

(43) $$d_2 - c_2 \simeq 2^{-1}\Sigma_i(1-p_{i.})/(np_{i.})$$

and for Design III, since $d_2=0$ and $c_2=0$ by Table VII,

(44) $$d_2 - c_2 = 0.$$

Therefore, in view of (20), (27), (29) and (31) and by noticing $c_{12}-c_1=E(z)$ the tests using c-AIC to reject H_1 if $z+2(d_1-d_{12}) > 0$ or equivalently,

(45) $$z - E(z) > -\Sigma_{i=1}^{I}\lambda_{i.}^{-1} + \lambda_{..}^{-1} \quad \text{for Design I,}$$

$$> \Sigma_i(1-p_{i.})/(np_{i.}) \quad \text{for Design II,}$$

and the tests using c-AIC to reject H_2 if $z + 2(d_2-d_{12}) > 0$ or equivalently

(46) $$z - 2E(z) > -\Sigma_j\lambda_{.j}^{-1} + \lambda_{..}^{-1} \quad \text{for Design I,}$$

$$> \Sigma_j(1-p_{.j})/(np_{.j}) \quad \text{for Design II,}$$

$$> \Sigma_j(1-p_{.j}^{\#})/(np_{.j}^{\#}) + O(n^{-2}) \quad \text{for Design III.}$$

where the last inequality follows from (44) and (31)

We remark that if the sample size is small, then we may want to get further expanded terms in d_1 and d_2 for c-AIC. In this case we may use the estimators for m_{ij}, $m_{i.}$ and $m_{.j}$ under H_1 and H_2 in Table VII below.

	m_{ij}	$m_{i.}$	$m_{.j}$
H_1	$x_{.j}/I$	$x_{..}/I$	$x_{.j}$
H_2	$x_{i.}/J$	$x_{i.}$	$x_{..}/J$

Table VII. Estimators for m_{ij}, $m_{i.}$ and $m_{.j}$.

8. Examples.

In this section we introduce three examples as applications of Designs I, II and III. In all three examples we can see that the tests using AIC and c-AIC give the same results as those using z-test for large n. Let ΔAIC =(AIC under the null hypothesis H_0) −(AIC under the alternative hypothesis H') and Δc-AIC =(c-AIC under H_0) − (c-AIC under H'). (Examples are taken from Andersen(1980).)

Example 1.) As an application of Design I, we consider the data in Table VIII from Rasch(1973). As part of a major investigation carried out in Sweden in 1962 to evaluate the influence of speed roads the number of accidents on Swedish roads were counted in 15 consecutive weeks. The roads were counted into groups: State highways, Country roads and Other roads.

Week	State highways	Country roads	Other roads	Total
1	2	7	4	13
2	8	8	4	20
3	7	9	9	25
4	7	4	8	19
5	3	5	7	15
6	5	4	4	13
7	4	5	7	16
8	4	4	12	20
9	7	3	8	18
10	3	8	12	23
11	4	12	15	31
12	4	5	14	23
13	9	12	10	31
14	10	9	17	36
15	10	9	14	33
Total	87	104	145	336

Table VIII. Accidents on Swedish roads classified according to type of road for 15 weeks in 1962.

The tests according to z-test, AIC and c-AIC can be summarized in the following variation table:

Variation	Null Hypotheses	Z	D.F.	$\chi^2_{.05}$	ΔAIC	Δc-AIC
Between cells	H_{12}: $u_{12(ij)}=0$	27.3	28	41.3	-28.7	-35.7
Between rows	H_1: $u_{1(i)}=0$	33.7	14	23.7	5.7	4.96
Between columns	H_2: $u_{2(j)}=0$	15.5	2	5.99	11.5	11.475
Total	all = 0	75.5	44			

Table IX. Table of variation for traffic data.

From Table IX, since 27.3 < 41.3, ΔAIC<0 and Δc-AIC<0, we cannot reject H_{12}. Hence, there is no relation between roads and the number of accidents. Similarly, we reject both H_1 and H_2 according to any test . Rejecting H_1 tells significant increase in the number of accidents, but rejecting H_2 does not tell us much.

Example 2.) In Table X is shown the distribution of 1517 married women in Denmark according to area of residence and type of occupation. The data are from Noordhoek(1969). Here, only the total sample is given. So we have Design II. We want to test if the distribution over occupations is the same for all three living areas.

Living area	Occupation				Total
	House wife	Parttime work	Fulltime work	In husbands business	
Copenhagen	188	66	139	16	409
Suburban Copenhagen	145	32	98	14	289
Cities	401	114	239	65	819
Total	734	212	476	95	1517

Table X. Married women in Denmark in 1964 distributed according to living area and occupation.

As above, the tests according to z-test, AIC and c-AIC can be summarized in the following variation table:

Variation	Null Hypotheses	Z	D.F.	$\chi^2_{.05}$	ΔAIC	Δc-AIC
Between cells	H_{12}: $u_{12(ij)}=0$	15.21	6	12.59	3.21	2.99

Table XI. Table of variation for woman's occupation data.

The result is clearly significant. Hence, occupation pattern is different in the city areas. Since we reject H_{12}, we do not analyze furthermore.

Example 3.) As an example of the test for homogeneity under Design III, we consider the data of Table XII in which is shown the results of the polls in September 1974 as published by four survey companies: Gallup, Observa, AIM and K. Vilstrup. It is known that the samples are not random. However, we shall for the purpose of illustrating the theory by an example consider each sample as randomly drawn from the Danish population.

It is a reasonable hypothesis that the four multinomial distribution corresponding to the columns of Table XII have the same cell probabilities, since each sample is taken from the same population at about the same time.

Party	Gallup	Observa	AIM	K. Vilstrup	Total
Labour	444	403	282	182	1311
Liberals	129	104	93	63	389
Conservatives	105	104	68	56	333
Socialists	68	65	60	42	235
Communists	75	65	52	56	248
Christian democrats	80	52	47	35	214
Farmers	225	286	196	140	847
Progress party	280	143	132	84	639
Others	94	78	70	42	284
Sample size	1500	1300	1000	700	4500

Table XII. A comparison of 4 Danish political polls.

The tests according to z-test, AIC and c-AIC can be summarized in the following variation table:

Variation	Null Hypothesis	Z	D.F.	$\chi^2_{.05}$	ΔAIC	Δc-AIC
Between cells	H_{12}: $u_{12(ij)}=0$	76.6	24	36.4	28.6	28.53
Between columns	H_2: $u_{2(j)}=0$	1880.75	8	15.5	1864.75	1864.73

Table XIII. Table of variation for poll data.

Since the z-test fails at the level of significance .05, there is no homogeneity of the columns of Table XII. Unfortunately, since H_{12} is not accepted, it is no use of testing for H_2. However, for the sake of an example for the theory, we compute the values of test statistics. As we can see from the Table XIII, uniformity assumption H_2 is obviously rejected.

9. Acknowledgement.

Authors wish to thank for valuable suggestions Professor Yoshiyuki Sakamoto at the Institute of Statistical Mathematics and Dr. Noriyuki Matsuda at the Institute of Socio-Economic Planning, University of Tsukuba.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on automatic control, Vol.AC-19, No. 6

Andersen, E. B. (1980). Discrete Statistical Methods with Social Science Applications. North Holland.

Bishop, Yvonne M. M., Fienberg, Stephen E. and Holland, Paul W. (1977). Discrete Multivariate Analysis. The MIT Press.

Goto, S., Hatanaka, S. and Tasaki, S.(1980). The Analysis of Binary Data (by D. R. Cox), Chapman and Hall (Translation; Asakura Shoten)

Noordhoek, J. A. (1969). Married Women, Family and Work. Vol. I, Employment. Copenhagen: Teknisk Forlag.

Rasch, G. (1973). Two applications of the multiplicative Poisson model in road accident statistics. Proc. 39th Sess. Int. Stat. Inst., 4, 31-43.

Sugiura, N. (1978). Further analysis of the data by AKAIKE's information criterion and the finite corrections. Commun, Statist. Theor. Meth., A7(1), 13-26.