Multiplicative Poisson Models with Unequal Cell Rates

Author(s): Erling B. Andersen

Source: *Scandinavian Journal of Statistics*, 1977, Vol. 4, No. 4 (1977), pp. 153-158

Published by: Wiley on behalf of Board of the Foundation of the Scandinavian Journal of Statistics

Stable URL: https://www.jstor.org/stable/4615671

Scand J Statist 4: 153–158, 1977

# Multiplicative Poisson Models with Unequal Cell Rates

ERLING B. ANDERSEN

University of Copenhagen

*Received April 1976, revised January 1977*

**ABSTRACT. A two way frequency table with independent Poisson distributed cell numbers is considered. The expected number in each cell is a product of a row effect, a column effect and a known constant. Methods are developed for estimation of the parameters by maximum likelihood. In addition asymptotic $\chi^2$-tests are considered for checking the model and testing equality of column effects and/or row effects. The proposed method is applied to Danish lung cancer data.**

*Key words:* multiplicative Poisson model, maximum likelihood, ancillarity, lung cancer data

## 1. Introduction

We consider a two dimensional contingency table

$$\mathbf{X} = \begin{pmatrix} X_{11} & \cdots & X_{1k} \\ \cdots & & \\ X_{m1} & \cdots & X_{mk} \end{pmatrix}$$

where the cell frequencies $X_{ij}$ are independent and Poisson distributed with expected values $\mu_{ij}$, $i=1, ..., m$, $j=1, ..., k$. If $X_{..}=n$ is given a priori we get a multinomial distribution of dimension $m \cdot k$ with parameters $(n, p_{11}, ..., p_{mk})$, where

$$p_{ij} = \mu_{ij}/\mu_{..}, \quad i = 1, ..., m, \quad j = 1, ..., k. \tag{1}$$

If we assume independence of the two entries of the contingency table, we get

$$p_{ij} = p_{i.}p_{.j}, \quad i = 1, ..., m, \quad j = 1, ..., k.$$

As is easily seen, this is equivalent to

$$\mu_{ij} = \mu_{i.}\mu_{.j}/\mu_{..},$$

or

$$\mu_{ij} = \varepsilon_i \delta_j, \tag{2}$$

where $\varepsilon_1, ..., \varepsilon_m$ are row effects and $\delta_1, ..., \delta_m$ are column effects. If we introduce the constraint

$$\delta_. = \sum_j \delta_j = 1, \tag{3}$$

we have $\varepsilon_i = \mu_{i.}$ and $\delta_. = \mu_{.j}/\mu_{..}$. The model (2) is generally known as the *multiplicative Poisson model.* The connection between independence in contingency tables and the multiplicative Poisson model is clearly demonstrated and investigated in Haberman (1975).

Slight complications arise in case the expected numbers in the cells of the contingency table are proportional to *known* constants $N_{ij}$, $i=1, ..., m$, $j=1, ..., k$. We then have the model

$$H_0: \mu_{ij} = N_{ij}\varepsilon_i\delta_j, \quad i = 1, ..., m, \quad j = 1, ..., k, \tag{4}$$

where we still retain the constraint (3).

This model is of great practical interest. As regards actuarial applications papers by Baily & Simon (1960) and Jung (1965) study the model in connection with primarily insurance risk models. Recently the applicability of the model to medical studies has been established by Osborn (1975).

In this paper we consider an epidemiological example concerning death rates caused by lung cancer.

Another Danish study under way is concerned with the registration of work accidents. We may assume that the observed number of accidents are Poisson distributed. The parameter of interest is not the expected number of accidents, but the expected number of accidents per hour. If we thus want to study the effect of, say, work type and age, we can divide the population by work type and age group. The total number of manhours $N_{ij}$ for a given work type $i$ and a given age group $j$ is known. If we also assume that age and work type influence the accidents independently, we get a model of the type

$$\mu_{ij} = N_{ij}\varepsilon_i\delta_j,$$

where $\mu_{ij}$ is the expected number of accidents in age group $j$ and for work type $i$, and where $\varepsilon_i$ is work type effect and $\delta_j$ is age group effect.

In addition to showing how the model can be applied to the lung cancer data mentioned above, the purpose of this paper is to discuss some of the theoretical properties of the model in relation to recent theory on exponential families and ancillarity.

## 2. Maximum likelihood estimation

Assume now that $X_{11}, ..., X_{mk}$ are independent Poisson distributed with mean values $\mu_{11}, ..., \mu_{mk}$, where $\mu_{ij}$ is given by (4). Let further $\boldsymbol{\epsilon} = (\varepsilon_1, ..., \varepsilon_m)$ and $\boldsymbol{\delta} = (\delta_1, ..., \delta_k)$. The loglikelihood function of the observed data then becomes

$$\ln L(\boldsymbol{\epsilon}, \boldsymbol{\delta}) = \sum_i \sum_j x_{ij} \ln\{N_{ij} \varepsilon_i \delta_j\}. \\ - \sum_i \sum_j \ln (x_{ij}!) - \sum_i \sum_j N_{ij} \varepsilon_i \delta_j.$$

We notice that the model forms an exponential family, since $\ln L$ can be written on the alternative form

$$\ln L(\boldsymbol{\epsilon}, \boldsymbol{\delta}) = \sum_i x_{i.} \ln \varepsilon_i + \sum_j x_{.j} \ln \delta_j \\ + \sum_i \sum_j (x_{ij} \ln N_{ij} - \ln(x_{ij}!)) - \sum_i \sum_j N_{ij} \varepsilon_i \delta_j. \quad (5)$$

From standard results for exponential families follow then, that the likelihood equations for the estimation of $\varepsilon_1, ..., \varepsilon_m$ and $\delta_1, ..., \delta_k$ are

$$x_{i.} = E[X_{i.}] = \varepsilon_i \sum_{j=1}^{k} N_{ij} \delta_j, \quad i = 1, ..., m \quad (6)$$

and

$$x_{.j} = E[X_{.j}] = \delta_j \sum_{i=1}^{m} N_{ij} \varepsilon_i, \quad j = 1, 2, ... k-1 \quad (7)$$

Only $m + k - 1$ equations need to be considered since (6) and (7) sum to the same number over their indicies. Because of (3) we solve only $k-1$ of the eqs. (7). The solutions of our likelihood equations will be denoted by $\hat{\boldsymbol{\epsilon}}$ and $\hat{\boldsymbol{\delta}}$. Eqs. (6) and (7) can be separated in $\varepsilon_i$ and $\delta_i$ by substituting $\delta_j$ from (7) into (6) and by substituting $\varepsilon_i$ from (6) into (7). In this way we get

$$x_{i.} = \varepsilon_i \sum_{j=1}^{k} \left\{ N_{ij} x_{.j} / \sum_{i=1}^{m} N_{ij} \varepsilon_i \right\} \quad (8)$$

and

$$x_{.j} = \delta_j \sum_{i=1}^{m} \left\{ N_{ij} x_{i.} / \sum_{j=1}^{k} N_{ij} \delta_j \right\}. \quad (9)$$

Eqs. (6) and (7) and alternatively (8) and (9) were first derived by Jung (1965) and discussed in detail by Ajne (1975). In both cases the quantities on the left hand sides were expressed in terms of $r_{ij} = x_{ij}/N_{ij}$, in which case the equations involve the two sums $\sum_j N_{ij} r_{ij}$ and $\sum_i N_{ij} r_{ij}$.

Statistical properties of the solutions to (8) and (9) are obtained from the general results of Barndorff-Nielsen (1973). Eqs. (8) and (9) thus have a set of unique solutions if $x_{.1}, ..., x_{.k}$, and $x_{1.}, ..., x_{m.}$ all stay away from 0. This result follows from Barndorff-Nielsen (1973), Theorem 7.1.

In order to find $\hat{\varepsilon}_i$ and $\hat{\delta}_j$, we solve (9) for $j = 1, ..., k$ observing that $\delta_k = 1 - \delta_1 - ... - \delta_{k-1}$. From the soluions $\hat{\delta}_1, ..., \hat{\delta}_k$, we can then determine $\hat{\varepsilon}_1$ directly from (6) for $i = 1, ..., m$.

Consider now the hypothesis

$$H_1: \delta_j = \delta_{j0}, \quad j = 1, ..., k. \quad (10)$$

Under $H_1$, the likelihood equations are obtained by (6) with $\delta_j = \delta_{j_0}$. We thus easily get

$$\tilde{\varepsilon}_i = x_{i.} / \sum_j N_{ij} \delta_{j0}, \quad i = 1, ..., m. \quad (11)$$

For the important case

$$H_1^*: \delta_j = 1/k, \quad j = 1, ..., k \quad (12)$$

we get

$$\tilde{\varepsilon}_i = k x_{i.} / N_{i.}, \quad i = 1, ..., m. \quad (13)$$

We thus have explicit solutions to the likelihood equations under the hypothesis (10).

Asymptotic properties of the estimatiors are easily obtained from the exponential form (5). By Barndorff-Nielsen's theorem (Barndorff-Nielsen (1973), Theorem 7.1), the transformation from the sufficient statistic $x_{1.}, ..., x_{m.}, x_{.1}, ..., x_{.k}$ to the ML-estimators constitute continuous functions, which are one–one when $x_{i.} > 0$ and $x_{.j} > 0$ for all $i$ and $j$. Hence convergence of $x_{i.}/x_{..}$ to its mean value and of $x_{.j}/x_{..}$ to its mean value ensures that the ML-estimators converge to the corresponding transformations of the mean values, which by definition are equal to the true values of the parameters. The ML-estimators are thus consistent.

Asymptotic normality of the estimators follow by a Taylor-expansion of the log-likelihood function centered at the true values.

Since the second derivatives of the log-likelihood function are not stochastic and continuous functions of the parameters, the usual regularity conditions (cf. for example Rao (1973), chap. 5f. 2) are easily seen to hold for this case. Hence it follows, that when $C(\boldsymbol{\epsilon}, \boldsymbol{\delta})$ is the matrix of second derivatives of $\ln L(\boldsymbol{\epsilon}, \boldsymbol{\delta})$ with respect to the parameters, then $(\hat{\boldsymbol{\epsilon}}, \hat{\boldsymbol{\delta}})$ is asymptotically normally distributed with mean values $(\varepsilon_0, \delta_0)$ and covariance matrix $[-C(\boldsymbol{\epsilon}_0, \boldsymbol{\delta}_0)]^{-1}$. Here $\boldsymbol{\epsilon}_0$ and $\boldsymbol{\delta}_0$ are the true parameter values.

## 3. Hypothesis testing

We shall now consider various tests. Firstly, a goodness of fit test for the model (4). This test may also under the assumption of Poisson distributed $X_{ij}$'s be viewed as a test for the hypothesis $H_0$ of multiplicative expected numbers $\mu_{ij}$.

We may test how well the data fits the model either by a log-likelihood ratio test or by a Pearson $\chi^2$-test. Log-likelihood ratio test variables are in this paper denoted by $z$'s and Pearsons $\chi^2$-test variables by $q$'s. The log-likelihood function is given by (5) under the multiplicative model (4). Without $H_0$ we have

$$\ln L(\mu_{11}, \ldots, \mu_{mk}) = \sum_i \sum_j (x_{ij} \ln \mu_{ij} - \mu_{ij}) - \sum_i \sum_j \ln(x_{ij}!)$$

and hence the ML-estimators

$$\hat{\mu}_{ij} = x_{ij}. \tag{14}$$

If we under the model (4) replace $\varepsilon_i$ by $\hat{\varepsilon}_i$ and $\delta_j$ by $\hat{\delta}_j$ in (5) and note that $\sum\sum N_{ij}\hat{\varepsilon}_i\hat{\delta}_j = x_{..}$ (cf. (6) and (7)), we get the test statistic

$$\begin{aligned} z &= -2\{\ln L(\hat{\boldsymbol{\varepsilon}}, \hat{\boldsymbol{\delta}}) - \ln L(\hat{\mu}_{11}, \ldots, \hat{\mu}_{mk})\} \\ &= 2\sum_i \sum_j x_{ij}\{\ln x_{ij} - \ln (N_{ij}\hat{\varepsilon}_i\hat{\delta}_j)\} \end{aligned} \tag{15}$$

This test quantity is by standard results asymptotically $\chi^2$-distributed with $(m-1)(k-1)$ degrees of freedom. The result may e.g. be obtained from the corresponding result for a multinomial distribution since $X_{11}, \ldots, X_{mk}$ given $X_{..} = x_{..}$ is multinomially distributed with cell probabilities $\mu_{ij}/\mu_{..}$. As an alternative to $z$, we may compute the corresponding Pearson $\chi^2$-test quantity

$$q = \sum_i \sum_j (x_{ij} - N_{ij}\hat{\varepsilon}_i\hat{\delta}_j)^2/(N_{ij}\hat{\varepsilon}_i\hat{\delta}_j), \tag{16}$$

which is equivalent to $z$ in large samples and thus also asymptotically $\chi^2$-distributed with $(m-1)(k-1)$ degrees of freedom.

Consider now the hypothesis $H_1$ given by (10). Under $H_1$ we get from (5) the log-likelihood function

$$\begin{aligned} \ln L(\tilde{\boldsymbol{\varepsilon}}, \boldsymbol{\delta}_0) &= \sum_i x_{i.} \ln \tilde{\varepsilon}_i + \sum_j x_{.j} \ln \delta_{j0} \\ &\quad + \sum_i \sum_j (x_{ij} \ln N_{ij} - \ln(x_{ij}!)) - \sum\sum N_{ij}\tilde{\varepsilon}_i\delta_{j0}, \end{aligned}$$

where $\tilde{\varepsilon}_i$ is given by (11). Since by (11) and (6) we have

$$\sum_i \sum_j N_{ij}\tilde{\varepsilon}_i\delta_{j0} = \sum_i \sum_j N_{ij}\hat{\varepsilon}_i\hat{\delta}_j = x_{..},$$

we get

$$z_1 = 2\textstyle\sum_i x_{i.} (\ln \hat{\varepsilon}_i - \ln \tilde{\varepsilon}_i) + 2\sum_j x_{.j} (\ln \hat{\delta}_j - \ln \delta_{j_0}) \tag{17}$$

as the log-likelihood ratio test statistic for the hypothesis (10) against the multiplicative model (4). This test statistic is asymptotically $\chi^2$-distributed with $k-1$ degrees of freedom.

We may finally test the hypothesis

$$H_2\colon \varepsilon_i = \varepsilon_{i0},\ \delta_j = \delta_{j0},$$

that all the parameters of the model have prescribed values. Against $H_1$ the test quantity for $H_2$ becomes

$$z_2 = \sum_i x_{i.}(\ln \tilde{\varepsilon}_i - \ln \varepsilon_{i0}), \tag{18}$$

which is asymptotically $\chi^2$-distributed with $m-1$ degrees of freedom. As we shall see in the example of Section 4, it depends on the problem at hand which hypotheses are relevant to test. It should also be noted that the sequence $z$, $z_1$, $z_2$ represents successive testing such that each hypothesis is tested under the assumption that all previous tested hypotheses are found to be true.

## 4. An example

The present paper was initiated by a Danish set of data, which were at the center of public interest in Denmark in 1974. In this section we shall give an analysis of these data in order to illustrate the use of the multiplicative model (4). Clemmensen et al. (1974) reports the number of recognized lung cancer cases in 4 Danish cities from 1968 to 1971. The purpose of the study was to find out whether there were grounds for a local suspicion that one of the cities (Fredericia) had a substantially higher rate of lung cancer cases than the other cities. It was recognized beforehand that age was a significant factor. Hence the cancer cases were reported in age groups. The data are reproduced from Clemmensen et al. (1974) in Table 1.

In Table 2 is shown the number of inhabitants for the four cities.

The numbers of Table 2 are relevant figures since we are interested in the risk of lung cancer for each individual. We shall, therefore, assume

1) The number of lung cancer cases in each cell of Table 1 is Poisson distributed.
2) The individual risk of getting lung cancer is composed as a product of an age factor and a city factor.

Assumptions 1) and 2) mean that with $x_{ij}$ = number of lung cancer cases in age group $i$ and for city $j$, we have

$$E[x_{ij}] = N_{ij}\varepsilon_i\delta_j, \tag{19}$$

where $N_{ij}$ are the numbers of Table 2. The model (19) is identical with (4), which means that we can apply the methods of sections 2 and 3.

Table 1. *Observed number of lung cancer cases between 1968 and 1971 for four Danish cities*

| Age | City Fredericia | Horsens | Kolding | Vejle | Total |
|---|---|---|---|---|---|
| 40–54 | 11 | 13 | 4 | 5 | 33 |
| 55–59 | 11 | 6 | 8 | 7 | 32 |
| 60–64 | 11 | 15 | 7 | 10 | 43 |
| 65–69 | 10 | 10 | 11 | 14 | 45 |
| 70–74 | 11 | 12 | 9 | 8 | 40 |
| >75 | 10 | 2 | 12 | 7 | 31 |
| Total | 64 | 58 | 51 | 51 | 224 |

Before we proceed it should be remarked that we are assuming independence between the $x_{ij}$'s. This will be the case if appearance of a lung cancer case does not increase the risk of lung cancer for other individuals in the neighbourhood. It is thus assumed that lung cancer is not contagious. This assumption cannot be checked based on the given data.

For the present data, eqs. (9) were solved by the Newton–Raphson iterative method. The method proved to be fast and efficient for the present data. As initial values we used

$$\delta_{j0} = x_{.j}/x_{..},$$

which are the solutions to (9) in case $N_{ij} = N_{i.}/k$. With these initial values the following solutions to (9) were found to

$$\hat{\boldsymbol{\delta}} = (0.315, 0.227, 0.218, 0.240). \tag{20}$$

From (6) we then get

$$\hat{\boldsymbol{\epsilon}} = (0.01138, 0.03414, 0.05184, 0.06650, 0.07469, 0.04695). \tag{21}$$

The previous papers (Jung, 1965; Ajne, 1975; Osborn, 1975) which has dealt with obtaining ML-

Table 2. *Number of inhabitants for the four cities distributed over age groups*

| Age | City Fredericia | Horsens | Kolding | Vejle | Total |
|---|---|---|---|---|---|
| 40–54 | 3 059 | 2 879 | 3 142 | 2 520 | 11 600 |
| 55–59 | 800 | 1 083 | 1 050 | 878 | 3 811 |
| 60–64 | 710 | 923 | 895 | 839 | 3 367 |
| 65–69 | 581 | 834 | 702 | 631 | 2 748 |
| 70–74 | 509 | 634 | 535 | 539 | 2 217 |
| >75 | 605 | 782 | 659 | 619 | 2 665 |

Table 3. *Expected number of lung cancer cases*

| Age | City Fredericia | Horsens | Kolding | Vejle |
|---|---|---|---|---|
| 40–54 | 10.9 | 7.4 | 7.8 | 6.9 |
| 55–59 | 8.6 | 8.4 | 7.8 | 7.2 |
| 60–64 | 11.6 | 10.9 | 10.1 | 10.5 |
| 65–69 | 12.2 | 12.6 | 10.2 | 10.1 |
| 70–74 | 11.7 | 10.5 | 8.5 | 9.4 |
| >75 | 9.0 | 8.3 | 6.7 | 7.0 |

estimates only considers eqs. (6) and (7), where the $\varepsilon_i$'s and $\delta_i$'s are not separated. In addition it is advocated to solve (6) and (7) by successive adjustment of marginals. This method, also known as the Deming–Stephan method (Deming & Stephan, 1940) is widely used for the analysis of contingency tables. The method is described in detail by Osborn (1975). In the present author's experience, the Newton–Raphson solution (9) combined with (6) is the most efficient way of obtaining the ML-estimators. In addition the standard errors of the estimates are easily obtained from the matrix of second derivatives used in the Newton–Raphson procedure. Expressions for the standard errors of the ML-estimators are also given by Osborn (1975).

The expected numbers $N_{ij}\hat{\varepsilon}_i\hat{\delta}_j$ are shown in Table 3.

In order to test the hypothesis $H_0$, that a Poisson model with multiplicative mean values of the form (4) describes the data, we apply the test quantity (15), which for the data in Tables 1, 2 and 3 becomes

$$z = 23.4, \quad df = 15.$$

At a 5%-level this value is not significant and we can not reject the null hypothesis $H_0$. Hence we must conclude that a Poisson model with multiplicative cancer risks seems to fit the data. The Pearson $\chi^2$-test statistic attains the value

$$q = 22.6, \quad df = 15,$$

which is below the 95%-fractile of a $\chi^2$-distribution with 15 degrees of freedom, and confirms the conclusion we drew from the $z$-value above.

Having accepted the model we can proceed to test the hypothesis (12) of identical column effects. $H_1^*$, given by (12) for the present problem means that the individual risk of getting lung cancer is the same for all 4 cities. The log-likelihood ratio test statistic $z_1$ for $H_1^*$ is given by formula (17). With $\hat{\varepsilon}_i$ and $\hat{\delta}_j$ given by (20) and (21) and $\tilde{\varepsilon}_i$ computed as (13) we get

$$z_1 = 4.9, \quad df = 3.$$

From this value of $z_1$ we may conclude that the hypothesis of identical column effects cannot be rejected.

In the original study, Clemmensen et al. (1974), it was assumed known that Fredericia had a higher intensity of lung cancer cases, if there was any difference at all. This suggests a one-sided test for (12) against the alternative

$$\delta_1 > \delta_2, \delta_3, \delta_4.$$

Such a test cannot be carried out by a $\chi^2$-test since $z_1$ can not detect which of the significant values are in accordance with $\delta_1 > \delta_2, \delta_3, \delta_4$.

An alternative would be to reformulate (12) as

$$H_1^{**}: \delta_1 = (\delta_2 + \delta_3 + \delta_4)/3$$

while leaving $\delta_2$, $\delta_3$ and $\delta_4$ free. This would require (6) and (7) to be solved under the constraint

$$\delta_1 = (\delta_2 + \delta_3 + \delta_4)/3.$$

A simpler approach would be to reiterate the complete analysis with Kolding, Horsens and Vejle as one city. Since in the authors' opinion there is no reason to believe a priori that Fredericia is the more dangerous city, no detailed results are reported here.

## 5. Sufficiency and ancillarity

The fact that the likelihood eqs. (6) and (7) can be written as (8) and (9) in which the $\varepsilon$'s and $\delta$'s are separated was not mentioned in the previous papers by Jung (1965), Ajne (1975) and Osborn (1975). Eqs. (8) and (9) can be derived directly if the $\varepsilon$'s and the $\delta$'s are obtained as *conditional* ML-estimators as suggested by Andersen (1970). Since $X_{1.}, ..., X_{m.}$ are sufficient for $\varepsilon_1, ..., \varepsilon_m$ we can base inference concerning $\delta_1, ..., \delta_k$ on the conditional distribution of $X_{.1}, ..., X_{.k}$ given $X_{1.} = x_{1.}, ..., X_{m.} = x_{m.}$ This distribution is independent of the $\varepsilon$'s by sufficiency and leads to (9) by maximization. A further consequence is that statistical tests concerning the $\delta$'s can be based on a conditional distribution that only depends on the $\delta$'s. Usually there is a loss in efficiency inflicted by considering conditional inference. Only when the conditioning statistics are in addition ancillary can we be sure that no such loss is inflicted. We shall now prove that $X_{1.}, ..., X_{m.}$ is ancillary for $\delta_1, ..., \delta_k$ in the sense that the marginal distribution of $X_{1.}, ..., X_{m.}$ contains no available information about $\delta_1, ..., \delta_k$.

We shall call a statistic $T$ weakly ancillary for $\theta$ in the presence of $\tau$, if for each $\theta_0$, $\tau_0$ and $\theta$, there exists a $\tau = \tau(\theta)$ such that the distribution of $T$ given $(\theta_0, \tau_0)$ is identical with the distribution of $T$ given $(\theta, \tau(\theta))$. This formulation was suggested by Andersen (1973), chapter 3.6. In Barndorff-Nielsen (1973), chapter 3, this form of ancillarity was termed $S$-ancillarity. The original definition was suggested by Sverdrup (1966) and Sandved (1967). Since the marginal distribution of $T = (X_{1.}, ..., X_{m.})$ is given by the density

$$f(x_{1.}, \ldots, x_{m.}) = \prod_{i=1}^{m} \left\{ \left( \sum_{j=1}^{k} \delta_j N_{ij} \right)^{x_{i.}} \varepsilon_i^{x_{i.}} \right\} \times \exp\{ -\sum_i \sum_j \varepsilon_i \delta_j N_{ij} \} / \prod_i x_{i.}!,$$

the marginal distribution of $(X_{1.}, ..., X_{m.})$ only depends on the parameters through $(\varepsilon_1 \sum_j \delta_j N_{1j}, ..., \varepsilon_m \sum_j \delta_j N_{mj})$. Hence for given values $(\delta_{10}, ..., \delta_{k0})$ and $(\varepsilon_{10}, ..., \varepsilon_{m0})$ of $(\delta_1, ..., \delta_k)$ and $(\varepsilon_1, ..., \varepsilon_m)$, we can put

$$\varepsilon_i' = \varepsilon_i(\delta_1, \ldots, \delta_k) = \varepsilon_{i0} \sum_j \delta_{j0} N_{ij} / \sum_j \delta_j N_{ij}.$$

With this definition, the distribution of $(X_{1.}, ..., X_{m.})$ is the same at the parameter point $(\delta_1, .., \delta_k, \varepsilon_1', ..., \varepsilon_m')$ as at the parameter point $(\delta_{10}, ..., \delta_{k0}, \varepsilon_{10}, ..., \varepsilon_{m_0})$. This means, however, according to the given definition that $(X_{1.}, ..., X_{m.})$ is weakly ancillary for $(\delta_1, ..., \delta_k)$ and thus contains no obtainable information about the value of $\delta_1, ..., \delta_k$.

As shown by Andersen (1973), chapter 3.6, a consequence of weak ancillarity is that conditional ML-estimators coincide with direct ML-estimators.

Thus it follows from known results that the solutions to (8) and (9) also maximize the conditional distributions given the sufficient statistics $X_{.1}, ..., X_{.k}$ and $X_{1.}, ..., X_{m.}$. It also follows that all the tests of section 3 are conditional tests and accordingly uniformly most powerful unbiased by the results in Lehmann (1959).

It is well known that the multiplicative Poisson model with all $N_{ij}$'s equal has ancillary as well as sufficient marginals. In this section we have thus established that the presence of unequal $N_{ij}$'s does not destroy this property.

## Acknowledgement

The author wishes to thank the referee for pointing out a number of relevant references and for many helpful comments on the paper.

## References

Almer, B. (1962). Individual risk theory and risk statistics as applied to fire insurance. *Astin Bull.* **2**.
Ajne, B. (1975). A note on the multiplicative ratemaking model. *Astin Bull.* **8**, 144–153.
Andersen, E. B. (1970). Asymptotic properties of conditional

maximum likelihood estimators. *J. Royal Stat. Soc. B.* **32**, 283–301.
Andersen, E. B. (1973). *Conditional inference and models for measuring*. Mentalhygiejnisk Forlag, Copenhagen.
Barndorff-Nielsen, O. (1973). *Exponential families and conditioning*. S. Møller Christensen, Århus.
Baily, R. A. & Simon, I. J. (1960). Two studies in automobile insurance ratemaking. *Astin Bull.* **1**.
Clemmensen, J., Hansen, G., Nielsen, A., Røjel, J., Steensberg, J., Sørensen, S. & Toustrup, J. (1974). Lung cancer and air pollution in Fredericia. *Ugeskrift for Læger* **136**, 2260–2268.
Deming, W. E. & Stephan, F. F. (1940). On a least square adjustment of a sampled frequency when the expected marginal totals are known. *Ann. Math. Stat.* **11**, 427–444.
Haberman, S. (1975). *The analysis of frequency data*. The University of Chicago Press, Chicago.
Jung, J. (1965). On automobile insurance ratemaking. *Astin Bull.* **5**, 41–48.
Lehmann, E. L. (1959). *Testing statistical hypothesis*. Wiley, New York.
Osborn, J. (1975). A multiplicative model for the analysis of vital statistics rates. *Appl. Statist.* **24**, 75–84.
Rao, C. R. (1973). *Linear statistical inference and its applications*. Wiley, New York.
Sandved, E. (1967). A principle for conditioning on an ancillary statistics. *Skand. Aktuarietidsskrift* **50**, 39–47.
Sverdrup, E. (1966). The present state of the decision theory and the Neyman–Pearson theory. *Rev. Int. Stat. Inst.* **34**, 309–333.

Erling B. Andersen
Department of Statistics
University of Copenhagen
Studiestræde 6
1455 Copenhagen
Denmark