

From Andersen, E. (1977). Multiplicative Poisson Models with Unequal Cell Rates. Scandinavian Journal of Statistics, 4(4), 153-158.

Aa two dimensional contingency table: $X = \begin{bmatrix} X_{11} & \cdots & X_{1k} \\ \vdots & & \\ X_{m1} & \cdots & X_{mk} \end{bmatrix}$ where the cell frequencies X_{ij} are independent and Poisson distributed with expected values μ_{ij} , $i = 1, \dots, m$, $j = 1, \dots, k$.

$X_{..} = \sum_{j=1}^k \sum_{i=1}^m X_{ij} = n$ is given a priori we get a multinomial distribution of dimension $m \cdot k$ with parameters $(n, P_{11}, \dots, P_{mk})$, where

$$f_{p_{ij}}(x_{ij}) = \frac{n!}{x_{11}! \cdots x_{mk}!} p_{11}^{x_{11}} \cdots p_{mk}^{x_{mk}} \mathbf{1}_{\{\sum_{i,j} x_{ij} = n\}}$$

$$P_{ij} = \frac{E[X_{ij}]}{E[X_{..}]} = \frac{\mu_{ij}}{\mu_{..}}, \quad i = 1, \dots, m, \quad j = 1, \dots, k \quad (1)$$

Assume independence of the two entries of the contingency table.

$$p_{i.} = \sum_{j=1}^k P_{ij} = \sum_{j=1}^k \frac{\mu_{ij}}{\mu_{..}} = \frac{\mu_{i.}}{\mu_{..}}; \quad p_{.j} = \sum_{i=1}^m P_{ij} = \sum_{i=1}^m \frac{\mu_{ij}}{\mu_{..}} = \frac{\mu_{.j}}{\mu_{..}}$$

$$P_{ij} = p_{i.} p_{.j} = \frac{\mu_{i.} \mu_{.j}}{\mu_{..} \mu_{..}}, \quad i = 1, \dots, m, \quad j = 1, \dots, k$$

multiplicative Poisson model: $X_{ij} \stackrel{iid}{\sim} \text{Multi-Pois}(\mu_{ij}) \quad i = 1, \dots, m, \quad j = 1, \dots, k$, where

$$\mu_{ij} = P_{ij} \mu_{..} = \frac{\mu_{i.} \mu_{.j}}{\mu_{..}^2} \mu_{..} = \frac{\mu_{i.} \mu_{.j}}{\mu_{..}} \quad \text{or} \quad \mu_{ij} = \varepsilon_i \delta_j, \quad (2)$$

$\varepsilon_1, \dots, \varepsilon_m$ are row effects, $\varepsilon_i = \mu_{i.}$ and $\delta_1, \dots, \delta_k$ are column effects, $\delta_j = \mu_{.j} / \mu_{..}$. $\sum_{i=1}^m \sum_{j=1}^k P_{ij} = 1$ The constraint is

$$\delta_{.} = \sum_j \delta_j = \sum_j \frac{\mu_{.j}}{\mu_{..}} = \frac{\mu_{..}}{\mu_{..}} = 1, \quad (3)$$

Define the expected numbers in the cells of the contingency table are proportional to known constants N_{ij} , $i = 1, \dots, m$, $j = 1, \dots, k$

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

$$\begin{aligned} f(x_{11}, \dots, x_{mk}) &= \prod_{i=1}^m \prod_{j=1}^k \frac{1}{x_{ij}!} (N_{ij} \varepsilon_i \delta_j)^{x_{ij}} \exp[-N_{ij} \varepsilon_i \delta_j] \mathbf{1}_{x_{ij} \in \mathbf{N}_0} \\ &= \left(\prod_{i=1}^m \prod_{j=1}^k (x_{ij}!)^{-1} \mathbf{1}_{x_{ij} \in \mathbf{N}_0} \right) \exp \left[\sum_{i,j} x_{ij} \ln(N_{ij} \varepsilon_i \delta_j) - \sum_{i,j} N_{ij} \varepsilon_i \delta_j \right] \\ &= \left(\prod_{i=1}^m \prod_{j=1}^k (x_{ij}!)^{-1} \mathbf{1}_{x_{ij} \in \mathbf{N}_0} \right) \exp \left[\sum_i x_{i.} \ln(\varepsilon_i) + \sum_j x_{.j} \ln(\delta_j) + \sum_{i,j} x_{ij} \ln(N_{ij}) - \sum_{i,j} N_{ij} \varepsilon_i \delta_j \right] \end{aligned}$$

where $x_{i.} = \sum_{j=1}^k x_{ij}$ is the i^{th} row sum, $x_{.j} = \sum_{i=1}^m x_{ij}$ is the j^{th} column sum.

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

This is an exponential family with natural sufficient statistics.

$$\begin{aligned} T(\vec{x}) &= (X_{1.}, \dots, X_{m.}, X_{.1}, \dots, X_{.k}) \\ \eta &= (\ln \varepsilon_1, \dots, \ln \varepsilon_m, \ln \delta_1, \dots, \ln \delta_k) \\ A(\eta) &= \sum_i \sum_j N_{ij} \varepsilon_i \delta_j \end{aligned}$$

The likelihood equations for the estimation of $\varepsilon_1, \dots, \varepsilon_m$ and $\delta_1, \dots, \delta_k$ are

$$T(\vec{x}) = E[(x_{i.})_{i=1}^m, (x_{.j})_{j=1}^k] = \dot{A}(\eta) = \left(\sum_{j=1}^k N_{ij} \varepsilon_i \delta_j, \dots, \sum_{j=1}^k N_{ij} \varepsilon_m \delta_j, \sum_{i=1}^m N_{ij} \varepsilon_i \delta_1, \dots, \sum_{i=1}^m N_{ij} \varepsilon_i \delta_k \right)$$

where $i = 1, \dots, m, j = 1, \dots, k-1, \sum_j \delta_j = 1$. That are

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

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

$$\begin{aligned} \hat{\varepsilon}_i &= \frac{x_{i.}}{\sum_{j=1}^k N_{ij} \delta_j}, i = 1, \dots, m \\ \hat{\delta}_j &= \frac{x_{.j}}{\sum_{i=1}^m N_{ij} \varepsilon_i}, j = 1, \dots, k-1 \end{aligned}$$

(6) and (7) can be separated in ε_i and δ_i by substituting δ_j from (7) into (6) and by substituting ε_i from (6) into (7). In this way we get

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

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

Let $x_{i.} = \sum_{j=1}^k N_{ij} r_{ij}$, $x_{.j} = \sum_{i=1}^m N_{ij} r_{ij}$,

$$x_{.j} = \sum_{i=1}^m N_{ij} r_{ij} = \hat{\delta}_j \sum_{i=1}^m N_{ij} \hat{\epsilon}_i = \hat{\delta}_j \sum_{i=1}^m N_{ij} \frac{x_{i.}}{\sum_{j=1}^k N_{ij} \hat{\delta}_j} = \hat{\delta}_j \sum_{i=1}^m N_{ij} \frac{\sum_{j=1}^k N_{ij} r_{ij}}{\sum_{j=1}^k N_{ij} \hat{\delta}_j}$$

Look at the first row $i = 1$. When $k = 1$, $j = 1$, $\hat{\delta}_j = \sum_j \hat{\delta}_j = 1$ by the constraint

When $k = 2$, $j = 1, 2$,

$$\begin{aligned} x_{11} &= N_{11} r_{11} = \hat{\delta}_1 N_{11} \frac{N_{11} r_{11} + N_{12} r_{12}}{N_{11} \hat{\delta}_1 + N_{12} \hat{\delta}_2} \\ r_{11} &= \hat{\delta}_1 \frac{N_{11} r_{11} + N_{12} r_{12}}{N_{11} \hat{\delta}_1 + N_{12} (1 - \hat{\delta}_1)} \\ r_{11} &= \hat{\delta}_1 \frac{N_{11} r_{11} + N_{12} r_{12}}{(N_{11} - N_{12}) \hat{\delta}_1 + N_{12}} \\ [(N_{11} - N_{12}) \hat{\delta}_1 + N_{12}] r_{11} &= \hat{\delta}_1 (N_{11} r_{11} + N_{12} r_{12}) \\ -N_{12} r_{11} \hat{\delta}_1 + N_{12} r_{11} &= \hat{\delta}_1 N_{12} r_{12} \\ \hat{\delta}_1 &= \frac{r_{11}}{r_{12} + r_{11}} \end{aligned}$$

which is same as $\hat{\delta}_1 = \frac{N_{12} X_{11}}{N_{12} X_{11} + N_{11} X_{12}}$. By analogy or by constraint,

$$\hat{\delta}_2 = 1 - \hat{\delta}_1 = \frac{r_{12}}{r_{12} + r_{11}}$$

When $k = 3$, $j = 1, 2, 3$, we assume $\hat{\delta}_j = \frac{r_{1j}}{\sum_{j=1}^3 r_{1j}}$, check

$$\begin{aligned} x_{11} &= N_{11} r_{11} = \hat{\delta}_1 N_{11} \frac{N_{11} r_{11} + N_{12} r_{12} + N_{13} r_{13}}{N_{11} \hat{\delta}_1 + N_{12} \hat{\delta}_2 + N_{13} \hat{\delta}_3} \\ r_{11} &= \frac{r_{11}}{r_{11} + r_{12} + r_{13}} \cdot \frac{N_{11} r_{11} + N_{12} r_{12} + N_{13} r_{13}}{N_{11} \frac{r_{11}}{r_{11} + r_{12} + r_{13}} + N_{12} \frac{r_{12}}{r_{11} + r_{12} + r_{13}} + N_{13} \hat{\delta}_3} \\ 1 &= \frac{N_{11} r_{11} + N_{12} r_{12} + N_{13} r_{13}}{N_{11} r_{11} + N_{12} r_{12} + N_{13} \hat{\delta}_3 (r_{11} + r_{12} + r_{13})} \\ N_{11} r_{11} + N_{12} r_{12} + N_{13} \hat{\delta}_3 (r_{11} + r_{12} + r_{13}) &= N_{11} r_{11} + N_{12} r_{12} + N_{13} r_{13} \\ \hat{\delta}_3 &= \frac{r_{13}}{r_{11} + r_{12} + r_{13}} \end{aligned}$$

Thus, it is true for $j = 1, 2, \dots, k$

$$\hat{\delta}_j = \frac{r_{1j}}{\sum_{j=1}^k r_{1j}} = \frac{r_{1j}}{r_{1.}}$$

We can get

$$\hat{\epsilon}_1 = \frac{x_{1.}}{\sum_{j=1}^k N_{1j} \hat{\delta}_j} = \frac{x_{1.}}{\sum_{j=1}^k N_{1j} \frac{r_{1j}}{r_{1.}}} = \frac{x_{1.} r_{1.}}{x_{1.}} = r_{1.}$$

which is not involved specific j . In other words, Each row effect is a marginalized value across columns and will not be affected by column effects. Thus, $\hat{\delta}_j = \frac{r_{1j}}{\sum_{j=1}^k r_{1j}}$ can be applied on any row.

Therefore, we can inductive that

$$\begin{aligned}\hat{\delta}_j &= \frac{r_{ij}}{\sum_{j=1}^k r_{ij}} = \frac{r_{ij}}{r_{i.}} \\ \hat{\varepsilon}_i &= \frac{x_{i.}}{\sum_{j=1}^k N_{ij} \hat{\delta}_j} = \frac{x_{i.}}{\sum_{j=1}^k N_{ij} \frac{r_{ij}}{r_{i.}}} = \frac{x_{i.} r_{i.}}{x_{i.}} = r_{i.}\end{aligned}$$

Or

$$\begin{aligned}\hat{\delta}_j &= \frac{x_{ij} N_{i.}}{x_{i.} N_{ij}} \\ \hat{\varepsilon}_i &= \frac{x_{i.}}{N_{i.}}\end{aligned}$$

Consider now the hypothesis

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

Under H_1 , the likelihood equations are obtained by $\hat{\varepsilon}_i = \frac{x_{i.}}{\sum_{j=1}^k N_{ij} \delta_j}$, $i = 1, \dots, m$ with $\delta_j = \delta_{j0}$.

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

For the important case

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

we get

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

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

```

#data
y<-c(3062,587,284,103,33,4,2)
S<-length(y)
w <- sum(y)
w_0 <- y[1]
w_1to6 <- sum(y[2:7])
c<- sum(y[2:7]*c(1:6))
lambda0<-2; eta0=0.1
PHI<-matrix(nrow=S,ncol=2)
PHI[1,]<-phi<-c( eta0, lambda0)
### iteration
for(s in 1:S) {
  # generate a new eta value
  phi[1] <- (w-w_0*exp(phi[2]))/(w*(1-exp(phi[2])))
  # generate a new lambda value
  phi[2]<- (c*phi[1]*(lambda0-1)*exp(lambda0)-w_1to6*phi[1]*(lambda0)^2*exp(lambda0)-c*(1-phi[1]))/
    ((c-w_1to6)*phi[1]*exp(lambda0)-w_1to6*phi[1]*(lambda0)*exp(lambda0)-w*(1-phi[1]))
  PHI[s,]<-phi
}
PHI

```

```

##           [,1]      [,2]
## [1,] 0.7125025 1.538327
## [2,] 0.6834310 1.515827
## [3,] 0.6814488 1.514257
## [4,] 0.6813078 1.514145
## [5,] 0.6812978 1.514137
## [6,] 0.6812971 1.514137
## [7,] 0.6812970 1.514137

```

Introduction

We consider a two dimensional contingency table: $X = \begin{bmatrix} X_{11} & \cdots & X_{1k} \\ \vdots & & \vdots \\ X_{m1} & \cdots & X_{mk} \end{bmatrix}$ where the cell frequencies X_{ij} are independent and Poisson distributed with expected values μ_{ij} , $i = 1, \dots, m$, $j = 1, \dots, k$. If $X_{..} = n$ is given a priori we get a multinomial distribution of dimension $m \cdot k$ with parameters $(n, P_{11}, \dots, P_{mk})$, where

$$P_{ij} = \frac{\mu_{ij}}{\mu_{..}}, \quad i = 1, \dots, m, \quad j = 1, \dots, k \quad (14)$$

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

$P_{ij} = p_{i.}p_{.j}$, $i = 1, \dots, m, j = 1, \dots, k$ As is easily seen, this is equivalent to

$$\mu_{ij} = \mu_{i.}\mu_{.j}/\mu_{..} \quad \text{or} \quad \mu_{ij} = \varepsilon_i\delta_j, \quad (15)$$

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

$$\delta_{.} = \sum_j \delta_j = 1, \quad (16)$$

we have $\varepsilon_i = \mu_{i.}$ and $\delta_j = \mu_{.j}/\mu_{..}$. 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} , $i = 1, \dots, m, j = 1, \dots, k$. We then have the model

$$H_0 : \mu_{ij} = N_{ij}\varepsilon_i\delta_j, \quad i = 1, \dots, m, j = 1, \dots, k, \quad (17)$$

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 μ_{ij} is the expected number of accidents in age group j and for work type i , and where ε_i is work type effect and δ_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.

Maximum likelihood estimation

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

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

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

$$\ln L(\vec{\varepsilon}, \vec{\delta}) = \sum_i x_{i.} \ln \varepsilon_i + \sum_j x_{.j} \ln \delta_j + \sum_i \sum_j (x_{ij} \ln N_{ij} - \ln(x_{ij}!)) - \sum_i \sum_j N_{ij}\varepsilon_i\delta_j \quad (18)$$

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

$$x_{i.} = E[X_{i.}] = \varepsilon_i \sum_{j=1}^k N_{ij}\delta_j, i = 1, \dots, m \quad (19)$$

$$x_{.j} = E[X_{.j}] = \delta_j \sum_{i=1}^m N_{ij}\varepsilon_i, j = 1, \dots, k-1 \quad (20)$$

Only $m+k-1$ equations need to be considered since (6) and (7) sum to the same number over their indices. Because of (3) we solve only $k-1$ of the eqs. (7). The solutions of our likelihood equations will be denoted by $\hat{\varepsilon}$ and $\hat{\delta}$ Eqs.

(6) and (7) can be separated in ε_i and δ_i by substituting δ_j , from (7) into (6) and by substituting ε_i from (6) into (7). In this way we get

$$x_{i.} = \varepsilon_i \sum_{j=1}^k \left\{ N_{ij} x_{.j} / \sum_{i=1}^m N_{ij} \varepsilon_i \right\} \quad (21)$$

$$x_{.j} = \delta_j \sum_{i=1}^m \left\{ N_{ij} x_{i.} / \sum_{j=1}^k N_{ij} \delta_j \right\} \quad (22)$$

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 $x_{.1}, \dots, x_{.k}$, and $x_{1.}, \dots, x_{m.}$ 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 $j = 1, \dots, k$ observing that $\delta_k = 1 - \delta_1 - \dots - \delta_{k-1}$. From the solutions $\hat{\delta}_1, \dots, \hat{\delta}_k$, we can then determine $\hat{\varepsilon}_1$ directly from (6) for $i = 1, \dots, m$.

Consider now the hypothesis

$$H_1 : \delta_j = \delta_{j0}, j = 1, \dots, k. \quad (23)$$

Under H_1 , the likelihood equations are obtained by (6) with $\delta_j = \delta_{j0}$. We thus easily get

$$\tilde{\varepsilon}_i = x_{i.} / \sum_j N_{ij} \delta_{j0}, i = 1, \dots, m. \quad (24)$$

For the important case

$$H_1^* : \delta_j = 1/k, j = 1, \dots, k. \quad (25)$$

we get

$$\tilde{\varepsilon}_i = kx_{i.}/N_{i.}, i = 1, \dots, m. \quad (26)$$

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