The Annals of Statistics
1981, Vol. 9, No. 4, 697–704

ON A GOODNESS-OF-FIT TEST FOR MULTIPLICATIVE POISSON MODELS

By Åke Svensson

University of Stockholm

A model for random $n \times k$ matrices X is considered. The elements X_{ij} are assumed to be independent and Poisson-distributed random variables with means $\alpha_i \exp(\gamma t'_{ij})$, where t_{ij} are known m-dimensional vectors and (α, γ) is an unknown parameter. A goodness-of-fit test is proposed and an approximation is derived when the number n of rows is large.

1. Introduction. Multiplicative Poisson models can be applied in situations where an $n \times k$ matrix x of nonnegative integers x_{ij} is observed. In the simplest case the elements of x are outcomes of experiments that are influenced by two factors, one row factor which is described by the n-dimensional parameter $\alpha = (\alpha_1, \cdots, \alpha_n)$, and one column-factor which is described by the k-dimensional parameter $\beta = (\beta_1, \cdots, \beta_k)$. It is assumed that the elements of x are independent observations on Poisson distributed random variables, X_{ij}, with expectations $\alpha_i\beta_j$ respectively. In the sequel we shall denote random variables with capital letters and observations with the corresponding small letters.

The simple multiplicative Poisson model can be generalized so that factors other than row and column can be taken into account. Let $t_{ij} = (t_{ij1}, \cdots, t_{ijm})$, $i = 1, \cdots, n$, $j = 1, \cdots, k$, be m-dimensional real vectors. Assume that the elements of x are independent observations on Poisson distributed random variables with expectations

$$\alpha_i \exp(\gamma t'_{ij}) \tag{1.1}$$

respectively. Here $\gamma = (\gamma_1, \cdots, \gamma_m)$ is a m-dimensional parameter.

The simple multiplicative Poisson model is clearly included in the class of models defined by (1.1). The models are treated in the theory of contingency tables and log-linear models (cf. Haberman, 1974, 1977).

The main purpose of this paper is to construct goodness-of-fit tests for these models, i.e., we shall test the hypothesis that a given set of observations is generated by a multiplicative Poisson model with a certain structure. Interest is focused on results which are valid when the number of rows is large.

In order to make things more clear we shall discuss an example. An experiment with speed restrictions was carried through in Sweden during the summers of 1961 and 1962. Periods with free speed alternated with periods with a general speed limit (90 km/h or 100 km/h). The number of traffic accidents with personal injuries that occurred (and were reported to the police) in day i of year j, x_{ij}, is given in the table below. By using a multiplicative Poisson model to analyse these data we can relate the nonrandom variations in the x_{ij}'s to the time of the year (day), the year and the speed restriction. To derive day effects it is important that days with corresponding characteristics are compared. To ensure this the days have been numbered so that Monday in week 22 of 1961 has the same number as Monday in week 22 of 1962, and so on. The observations consist of 92 days (n = 92) and two years (k = 2). The elements of x are assumed to be observations on independent Poisson distributed random variables, X_{ij}. Define the vector t_{ij} so that

Received February 15, 1978; revised July 28, 1980.

AMS 1970 *subject classifications.* Primary 62H15; secondary 62P99.

Key words and phrases. Multiplicative Poisson model, exact test, asymptotic distribution, conditional limit theorem.

$$
(1.2)\qquad
\begin{aligned}
E\,X_{i1} &= \begin{cases} \alpha_i \exp(\gamma_1) & \text{if there was free speed day } i \text{ year 1} \\ \alpha_i \exp(\gamma_1 + \gamma_2) & \text{otherwise} \end{cases} \\
E\,X_{i2} &= \begin{cases} \alpha_i & \text{if there was free speed day } i \text{ year 2} \\ \alpha_i \exp(\gamma_2) & \text{otherwise.} \end{cases}
\end{aligned}
$$

The α_i's described the day-effects, γ_1 the differences between the two years, and γ_2 the

TABLE 1

The observation matrix x for the speed limit experiments 1961 *and* 1962. (+ *indicates days with speed limit*).

day no	1961	1962	day no	1961	1962
1	9	9	47	37	29+
2	11	20	48	32	17+
3	9	15	49	25	17+
4	20	14	50	20	15+
5	31	30	51	40	25+
6	26	23	52	21	9+
7	18	15	53	18	16+
8	19	14	54	35	25+
9	18	16	55	21	25+
10	13	20	56	25	16+
11	29	17+	57	34	22+
12	40	23+	58	42	21+
13	28	16+	59	27	17+
14	17	20+	60	34	26+
15	15	13+	61	47	41+
16	21	13+	62	36	25+
17	24	9+	63	15	12+
18	15	10+	64	26	17+
19	32	17+	65	27	21
20	22	12+	66	18	19
21	24	7+	67	16	24
22	11	11+	68	32	44
23	27	15+	69	28	31
24	12+	19+	70	17	21
25	41+	32+	71	16	20
26	15+	22+	72	19	19
27	18+	24+	73	18	20
28	11+	9+	74	22	29
29	19+	10	75	37	48
30	19+	14	76	29	36
31	9+	18	77	18	15
32	21+	26	78	14	16
33	22+	38	79	14	29
34	23+	31	80	18	12
35	14+	12	81	21	24+
36	19+	8	82	39	26+
37	15+	22	83	39	16+
38	13+	17	84	21	15+
39	22+	31	85	15	12+
40	42+	49	86	17	22+
41	29+	23	87	20	24+
42	21+	14	88	24	16+
43	12+	25	89	30	25+
44	16+	24	90	25	14+
45	17	18	91	8	15+
46	27	19	92	21	9+

effect of the speed restriction. In fact $1 - \exp(\gamma_2)$ is a measure of the relative effect of the speed restriction on the expected number of accidents.

2. Test of the model. The probability of obtaining the observation x is

$$P_{\alpha\gamma}(x) = \{\prod_{i,j}(x_{ij}!)^{-1}\}(\prod_i \alpha r_i)\exp(\gamma u')\exp\{-\sum_{i,j}\alpha_i\exp(\gamma t'_{ij})\}, \tag{2.1}$$

where

$$r_i = \sum_j x_{ij}, \qquad i = 1, \cdots, n \tag{2.2}$$

and

$$u = \sum_{i,j} x_{ij}t_{ij}. \tag{2.3}$$

We shall use the notation T_i for the $k \times m$ matrix which has the vectors t_{ij} as rows. There is some arbitrariness in the choice of the t_{ij}'s. The same probabilities can be obtained (with other parameters) for different t_{ij}'s. It is in fact always possible to choose the t_{ij}'s so that the last row of the T_i's contain only zero. This will often be a convenient choice.

The probabilities (2.1) form an exponential family of distributions with the random variables $R_i = \sum_j X_{ij}$, $i = 1, \cdots, n$, and $U = \sum_{i,j} X_{ij}t_{ij}$ as sufficient statistics. We shall assume that $R = (R_1, \cdots, R_n)$ and U are linearly independent. It follows from the sufficiency that the conditional probabilities of X given $R = r$ and $U = u$

$$P(x \mid r, u) = P_{\alpha\gamma}(x)/P_{\alpha\gamma}(r, u) \tag{2.4}$$

do not depend on the values of the parameters.

In this paper we shall study a conditional goodness-of-fit test that rejects the model when the conditional probability of obtaining the observation x given the sufficient statistics is too small, i.e.,

$$P(x \mid r, u) \leq \varepsilon \tag{2.5}$$

where ε is chosen so that the desired level of the test is obtained, i.e.,

$$P(\{X;\, P(X \mid r, u) \leq \varepsilon\} \mid r, u) = \delta \tag{2.6}$$

This is an absolute test in the sense of Cox and Hinkley (1974, page 83) applied to the conditional distribution. Tests of this kind are also called exact tests (cf. Fisher, 1934, and Martin-Löf, 1973).

Obviously it is difficult to calculate the exact value of ε in every situation since it will depend on the t_{ij}'s, r, and u in a very complicated manner. It is, however, possible to find good approximations in certain asymptotic situations. A familiar case is when k, n, and m are kept fixed and the expected values, $\alpha_i\exp\{\gamma t'_{ij}\}$, of the cells grow large. The usual goodness-of-fit test is the one that rejects the model when

$$\sum_{i,j}\frac{\{x_{ij} - \hat{\alpha}_i\exp(\hat{\gamma}t'_{ij})\}^2}{\hat{\alpha}_i\exp(\hat{\gamma}t'_{ij})} > \chi^2_{1-\delta}(kn - m - n)$$

where $\chi^2_{1-\delta}(kn - m - n)$ is the $(1 - \delta)$ percentile of a χ^2 distribution with $kn - m - n$ degrees of freedom. The $\hat{\alpha}$'s and $\hat{\gamma}$'s are the ML estimates of the parameters. This test can be derived as an approximation to the exact test (cf. Martin-Löf, 1973).

In this paper we shall consider the case when k and m are fixed but the number of rows, n, is large. It is also possible in this situation to find an approximation to the exact test. From (2.1) and (2.4) it follows that

$$P(x \mid r, u) = \prod_{i,j}(x_{ij}!)^{-1}H(r, u) \tag{2.7}$$

where H is a function of (r, u) only. Since the test defined by (2.5) is conditional on (R, U) it is equivalent to reject the model when the observed value of the test statistic

$$Q = \sum_{i,j}\ln X_{ij}! \tag{2.8}$$

is larger than some number $\varepsilon(r, u)$.

3. The asymptotic distribution of the test statistic. The vectors $X_i = (X_{i1}, \cdots, X_{ik})$ conditional on $R_i = r_i$ are independent and k-nomially distributed random vectors with the probabilities

$$p_\gamma^{(i)}(x_i \mid r_i) = r_i!(\prod_j x_{ij}!)^{-1} \prod_j p_{ij}^{r_i} \tag{3.1}$$

respectively, where

$$p_{ij} = \exp(\gamma t'_{ij})/\sum_{\nu} \exp(\gamma t'_{i\nu}), \qquad j = 1, \cdots, k. \tag{3.2}$$

Conditionally on the row-sums the test statistic Q is thus a sum of a large number, n, of independent random variables $Q_i = \sum_j \ln X_{ij}!$. The same is true for the statistic U (cf. (2.3)). With this in mind it should be possible to impose conditions on the behaviour of the r_i's and the t_{ij}'s so that the $(m + 1)$-dimensional random vector (Q, U) is asymptotically normally distributed conditional on r.

If we heuristically condition on $U = u$ in this limiting distribution we obtain a normal distribution with mean

$$E_\gamma(Q \mid r) + \mathrm{Cov}_\gamma (Q, U \mid r)\{\mathrm{Var}_\gamma (U \mid r)\}^{-1}\{U - E_\gamma(U \mid r)\}'$$

and variance

$$\mathrm{Var}_\gamma (Q \mid r) - \mathrm{Cov}_\gamma (Q, U \mid r)\{\mathrm{Var}_\gamma (U \mid r)\}^{-1}\mathrm{Cov}'_\gamma (Q, U \mid r).$$

We shall prove that under certain conditions this will indeed be the true limiting distribution of Q given $R = r$ and $U = u$ provided γ is chosen as the solution of the equation

$$u = E_\gamma(U \mid r). \tag{3.3}$$

The solution of this equation is the ML estimate of γ. With the help of some lemmata of a technical nature, proved in Section 6, we can prove the following theorem:

THEOREM 3.1. *Let X be a random matrix with probabilities given by* (2.1) *such that*
(a) $0 < A \le \alpha_i \le B < \infty \qquad i = 1, 2, \cdots;$
(b) *there exist a positive definite matrix J and a finite number n_b such that*

$$\sum_{i=1}^n \alpha_i T'_i T_i / \sum_{i=1}^n \alpha_i \ge J \qquad \textit{for all } n \ge n_b;$$

(c) *the matrices T_i take only one of a finite number of values $V_i, \cdots, V_c$ and there exists a positive number η such that*

$$\#\{i \le n;\ T_i = V_d\}/n^\eta$$

are bounded away from zero as $n \to \infty$. Then $\Pr(Q \le m + vz \mid r, u) \to \Phi(z)$ *as $n \to \infty$ where*

$$m = E_{\hat\gamma}(Q \mid r),\ v^2 = \mathrm{Var}_{\hat\gamma} (Q \mid r) - \mathrm{Cov}_{\hat\gamma} (Q, U \mid r)\{\mathrm{Var}_{\hat\gamma} (U \mid r)\}^{-1} \mathrm{Cov}'_{\hat\gamma} (Q, U \mid r),$$

and $\hat\gamma$ is a solution of (3.3).

PROOF: See Section 6.

4. Some special models. We shall apply Theorem 1 to two special cases which are closely related to the example discussed in the introduction. In both cases $k = 2$.

4.1 *The simple multiplicative Poisson model.* A simple multiplicative Poisson model with $k = 2$ occurs when $m = 1$, $t_{i1} = 1$ and $t_{i2} = 0$. According to (1.1) the Poisson distributed random variables X_{ij} have the expectations $\alpha_i \exp(\gamma_j)$ respectively ($\gamma_2 = 0$). The conditions (b) and (c) are trivially satisfied. If we assume that (a) holds we can apply the theorem. Now

$$U = \sum_{i=1}^n X_{i1} \qquad \text{and} \qquad \hat\gamma_1 = \ln\{u/(t - u)\}, \qquad \text{where } t = \sum_{i=1}^n (x_{i1} + x_{i2}).$$

The test statistic

$$Q = \sum_{i=1}^{n} (\ln X_{i1}! + \ln X_{i2}!)$$

is, according to the theorem, asymptotically normally distributed with mean and variance respectively

$$\sum_{i=1}^{n} E(\hat{p}, r_i), \qquad \sum_{i=1}^{n} D(\hat{p}, r_i) - \{\sum_{i=1}^{n} C(\hat{p}, r_i)\}^2/\{u(t-u)/t\}.$$

Here

$$\hat{p} = \exp(\hat{\gamma}_1)/\{1 + \exp(\hat{\gamma}_1)\} = u/t.$$

The functions E, C, and D are defined by

$$E(p, h) = \sum_{x=0}^{h} \{\ln x! + \ln(h-x)!\} \binom{h}{x} p^x(1-p)^{h-x},$$

$$C(p, h) = \sum_{x=0}^{h} x\{\ln x! + \ln(h-x)!\} \binom{h}{x} p^x(1-p)^{h-x} - hpE(p, h),$$

and

$$D(p, h) = \sum_{x=0}^{h} \{\ln x! + \ln(h-x)!\}^2 \binom{h}{x} p^x(1-p)^{h-x} - E^2(p, h).$$

4.2 *Model with one additional factor.* Assume that in addition to the row and column effects there exists a two-level factor (levels A and B) which has a multiplicative effect on the expected numbers in the cells. In the model discussed in the introduction the additional factor was the speed restriction. Let h_{ij} be the level of the additional factor in cell (i, j) and define the 2-dimensional t_{ij}-vectors by

$$t_{i1} = \begin{cases} (1, 0) & \text{if } h_{i1} = A \\ (1, 1) & \text{if } h_{i1} = B \end{cases}$$

$$t_{i2} = \begin{cases} (0, 0) & \text{if } h_{i2} = A \\ (0, 1) & \text{if } h_{i2} = B. \end{cases}$$

The resulting expectations are given by (1.2). Conditional on $R_i = r_i$ the random variables X_{i1} are binomially distributed with probability parameters (cf. (3.2))

$$(4.1) \qquad p_{i1} = \begin{cases} \pi_1 = \exp(\gamma_1)/\{1 + \exp(\gamma_1)\} & \text{if } h_{i1} = h_{i2} \\ \pi_2 = \exp(\gamma_1 + \gamma_2)/\{1 + \exp(\gamma_1 + \gamma_2)\} & \text{if } h_{i1} = B \text{ and } h_{i2} = A \\ \pi_3 = \exp(\gamma_1)/\{\exp(\gamma_1) + \exp(\gamma_2)\} & \text{if } h_{i1} = A \text{ and } h_{i2} = B. \end{cases}$$

All conditions of the theorem are satisfied if (a) holds and if $\#\{i \le n;\ p_{i1} = \pi_v\}/n$, $v = 1, 2, 3$, are bounded away from zero as $n \to \infty$. If $\delta_h^c = 1$ when $h = c$ and $\delta_h^c = 0$ otherwise then

$$U = (\sum_{i=1}^{n} X_{i1}, \sum_{i=1}^{n} \sum_{j=1}^{k} \delta_{h_{ij}}^{B} X_{ij}).$$

The ML estimates of γ solve the equations

$$u_1 = r_1^{(n)}\pi_1 + r_2^{(n)}\pi_2 + r_3^{(n)}\pi_3$$

$$u_2 = r_2^{(n)}\pi_2 + r_3^{(n)}(1 - \pi_3),$$

where $r_v^{(n)} = \sum_{i=1}^{n} \delta_{p_{i1}}^{\pi_v} r_i$, $v = 1, 2, 3$. The test statistic Q is asymptotically normally distributed with mean and variance respectively

$$\sum_{i=1}^{n} E(\hat{p}_{i1}, r_i), \qquad \sum_{i=1}^{n} D(\hat{p}_{i1}, r_i) - C[\mathrm{Var}_{\hat{\gamma}}\,(U \mid r)]^{-1}C',$$

where $\hat{p}_{i1}$ is defined by (4.1) with the estimated $\hat{\gamma}$'s and

$$C = (\sum_{i=1}^{n} C(\hat{p}_{i1}, r_i), \sum_{i=1}^{n} \{\delta_{h_{i1}}^{B} C(\hat{p}_{i1}, r_i) + \delta_{h_{i2}}^{B} C(1 - \hat{p}_{i1}, r_i)\}).$$

5. An application. We shall now return to the empirical data discussed in the introduction. They will be analysed with the models treated in the previous section. First we disregard the possible effects of the speed restriction and try to apply a simple multiplicative Poisson model to the data. This results in

Q_{obs}	$\hat{\gamma}_1$	m	v^2	$(Q_{obs} - m)/v$
8969.85	0.11	8940.75	44.65	4.36

The hypothesis that the model fits the data can thus be rejected.

The bad fit of the model might be the result of the fact that the speed limit has had a real effect on the number of accidents or of some other deviation from the model. If we try to apply a simple model to days with equal speed restrictions both years, we will get three disjoint sets of days. Using a goodness-of-fit test on each of these sets we find:

(a) days with speed limit both years or free speed both years

Q_{obs}	$\hat{\gamma}_1$	m	v^2	$(Q_{obs} - m)/v$
2975.73	0.02	2979.46	15.42	−0.95

(b) days with speed limit 1961 and free speed 1962

Q_{obs}	$\hat{\gamma}_1$	m	v^2	$(Q_{obs} - m)/v$
1529.29	−0.14	1524.23	7.73	1.82

(c) days with free speed 1961 and speed limit 1962

Q_{obs}	$\hat{\gamma}_1$	m	v^2	$(Q_{obs} - m)/v$
4464.83	0.36	4463.82	21.08	0.22.

The fit of the model seems to be good in cases (a) and (c) but somewhat dubious in case (b). It now seems reasonable to apply the model with one additional factor (cf. Section 4.2) to the complete set of observations. This results in:

Q_{obs}	$\hat{\gamma}_1$	$\hat{\gamma}_2$	m	v^2	$(Q_{obs} - m)/v$
8969.85	0.02	−0.29	8962.83	44.44	1.05

The model can not be rejected by this test with the evidence from these data. The estimate $\hat{\gamma}_2 = -0.29$ can be interpreted as saying that the speed limit has decreased the expected number of accidents by a factor $1 - \exp(-0.29) = 0.25$.

6. Theoretical results. The proof of Theorem 3.1 is based on the following theorem. (Z^k is the k-dimensional lattice of integers.)

THEOREM 6.1. (Conditional limit theorem). *Let (Z_{in}, Y_{in}), $i = 1, \cdots, n$ be a sequence of independent random variables with $Z_{in} \in R$ and $Y_{in} \in Z^k + a_{in}$, where a_{in} are k-dimensional real vectors. Assume that $E\, Z_{in} = 0$, $E\, Y_{in} = \mathrm{Cov}(\sum_i Z_{in}, \sum_i Y_{in}) = 0, s_n^2 = \sum_i E\, Z_{in}^2$, $L_n = \sum_i E\, Y'_{in} Y_{in}$, and that there exist sequences c_n, d_n, and e_n which all tend to 0 as $n \to \infty$ and*

(a) $\sum_i E\,|\,Z_{in}\,|^3 \le c_n s_n^3$

(b) $\sum_i E\,|\,vY'_{in}\,|^3 \le d_n(vL_nv')^{3/2}$ *for all* $v \in R^k$

(c) *there exist a positive definite matrix J such that* $e_n L_n^{1/2} \ge J$,

(d) *there exist a $\xi > 0$ such that if*

$m_n = \#\{i \le n;\ \inf_v \sup_y \Pr(Y_{in} = y)\Pr(Y_{in} = y + e_v) \ge \xi\}$

(e_v is the vth unit vector) then for all $|\,\rho\,| < 1$

$\rho m_n L_n \to 0$ *as* $n \to \infty$,

then for any sequence y_n contained in a compact set such that $y_n - \sum_i a_{in} \in Z^k$

$$\Pr(\textstyle\sum_i Z_{in} \le x s_n \mid \sum_i Y_{in} = y_n) \to \Phi(x)$$

as $n \to \infty$.

A proof of this conditional limit theorem is given in von Bahr and Svensson (1979). We shall use it here to derive the asymptotic distribution of Q given $R = r$ and $U = u$. If we consider the conditions of the theorem it is easy to see that (a) and (b) are Liapounov conditions which will guarantee the asymptotic normality of the sums $\sum_i Y_{in}$ and $\sum_i Z_{in}$. Condition (c) implies that $\sum_i Y_{in}$ is asymptotically nonsingular. The fourth condition, (d), has obviously to do with the arithmetic structure of the distributions.

The conditional limit theorem will be used to justify the heuristic argument in Section 3. The technical parts of the proof will be left out where they only use standard statistical arguments or straight-forward derivations of inequalities.

The proof is divided into steps. We remind the reader that conditional on $R = r$ the k-dimensional vectors X_i are independent random vectors with probabilities $P_\gamma^{(i)}(x_i \mid r_i)$ (cf. (3.1)). Conditional also on $U = u$ the distribution of X does not depend on the value of γ. The same is true for Q since it is a function of X only. This fact allow us to work with $\hat\gamma$ (the solution of $u = E_\gamma(U \mid r)$) instead of γ. The justification for this is the following argument. If for any array $\tilde{X}_i^{(n)}$ of k-dimensional random vectors, $i = 1, \cdots, n$, with probabilities $P_{\hat\gamma}^{(i)}(\tilde{x}_i^{(n)} \mid r_i)$ we can prove that

$$\Pr(\textstyle\sum_{i,j} \ln \tilde{X}_{ij}^{(n)}! \le m + vz \mid \sum_j \tilde{X}_{ij}^{(n)} = r_i,\ i = 1, \cdots, n,\ \sum_{i,j} \tilde{X}_{ij}^{(n)} t_{ij} = u) \rightarrow \Phi(z)$$

as $n \rightarrow \infty$ we have also proved the conditional convergence of Q.

In the following we will use the random vectors $\tilde{X}_i^{(n)}$. To simplify notation we will write X_i instead of $\tilde{X}_i^{(n)}$. We have to bear in mind that the distribution of X_i depends on n (through $\hat\gamma$). Define

$$Y_{in} = \{X_i - E_{\hat\gamma}(X_i \mid r_i)\} T_i$$

and

$$Z_{in} = \textstyle\sum_j \ln X_{ij}! - E_{\hat\gamma}(\sum_j \ln X_{ij}! \mid r_i) - Y_{in} B_n'$$

where

$$B_n = \mathrm{Cov}_{\hat\gamma}(Q, U \mid r)\{\mathrm{Var}_{\hat\gamma}(U \mid r)\}^{-1}.$$

The idea is to verify that for any well-behaved (in a sense to be defined later) sequence $r_1, r_2, \cdots$ we can apply the conditional limit theorem to the variables (Z_{in}, Y_{in}) conditional on $R_i = r_i$. In this way we can derive the asymptotic distributions of $\sum_i Z_{in}$ conditional on $R = r$ and $\sum_i Y_{in} = 0$ or equivalently $U = u$.

To prove the main theorem we have to verify:

(i) The sequence $r_1, r_2, \cdots$ is well-behaved with probability 1,

(ii) the solution $\hat\gamma$ on which the construction is based does really exist, and

(iii) the conditions (a)–(d) are satisfied if the sequence is well-behaved.

Using Kolmogorov's strong law of large numbers it is possible to prove the following lemma about the well-behaviour of the sequence $r_1, r_2, \cdots$

LEMMA 1. *If* (a)–(c) *of Theorem* 3.1 *hold then with probability* 1 *it is true that there exist a finite number* n_1 *such that if* $n \ge n_1$

(A) $\sum_{i=1}^n r_i/n$, $\sum_{i=1}^n r_i^2/n$, *and* $\sum_{i=1}^n r_i^3/n$ *are bounded away from* 0 *and* ∞;

(B) $\#\{i \le n;\ r_i = h,\ T_i = V_d\}/n^\eta$ *and* $\#\{i \le n,\ r_i = h\}/n$ *are bounded away from* 0 *for all integers* h *and* $d = 1, \cdots, c$;

(C) *there exist a positive definite matrix* J_1 *such that* $\sum_{i=1}^n r_i T_i' T_i / \sum_{i=1}^n r_i \ge J_1$.

With arguments which are standard to prove the consistency of ML estimates we can also verify (ii) with the following lemma.

LEMMA 2. *If* (a)–(c) *of Theorem* 3.1 *hold then it is true with probability* 1 *that*

(D) *for any neighborhood* Γ *of* γ *there exist a finite number* $n(\Gamma)$ *such that the equation* (3.3) *has a solution contained in* Γ *for all* $n \ge n(\Gamma)$.

The remaining parts of the proof is given in the following lemma.

LEMMA 3. *If* (A)–(D) *hold then*

$$\Pr(\textstyle\sum_i Z_{in} \le zs_n \mid \sum_i Y_{in} = 0) \to \Phi(z)$$

as $n \to \infty$, *where* $s_n^2 = \sum_i \operatorname{Var} Z_{in}$.

PROOF. We have to verify that conditions (a)–(d) of Theorem 6.1 are satisfied. According to Lemma 2 the variables Y_{in} and Z_{in} are well-defined for large n-values. By construction the moments satisfy the required restrictions. Now $\operatorname{Var}_{\hat{\gamma}}(Z_{in} \mid R_i = 2)$ is a positive continuous function of $\hat{\gamma}$. It follows from Lemma 2 and (B) that s_n^2/n is bounded away from 0 for large n. Due to (C) the matrix L_n/n is larger than a positive definite matrix for large n. The two inequalities $|X_{ij}| \le r_i$ and $|\sum_j \ln X_{ij}! - \ln r_i!| \le \text{const } r_i$ implies together with (A) that the individual Y_{in} and Z_{in} are small compared to their sums and that the Liapounov conditions (a) and (b) are satisfied. Condition (c) is a consequence of (A) and (B). It remains to verify (d).

Let us for the moment assume that the lattices $Z^k V_d$, $d = 1, \cdots, c$, span Z^m. If this is the case there will exist at least one $\zeta \in Z^k$ and a V_d such that $\zeta V_d = e_v$. Let ζ^* be any member of Z^k such that all elements of both ζ^* and $\zeta^* + \zeta$ are nonnegative. If the T_i's are chosen so that all elements of the last column equal 0, we can always obtain that $h_d^* = \sum_i \zeta_i^* = \sum_i \zeta_i + \zeta_i^*$ by varying the last element of ζ. If $T_i = V_d$ and $r_i = h_d^*$ then Y_{in} can take any of the values $y = (\zeta^* - E_{\hat{\gamma}}(X_i \mid h_d^*))V_d$ and $y + e_v = (\zeta + \zeta^* - E_{\hat{\gamma}}(X_i \mid h_d^*))V_d$ with positive probability. If ξ is chosen small enough we can obtain that $m_n = \inf_d \#\{i \le n;\ r_i = h_d^*,\ T_i = V_d\}$. According to (b) $m_n \sim n^\eta$. Since $L_n \sim \sqrt{n}$ (d) is satisfied.

If $Z^k V_d$, $d = 1, \cdots, c$, do not span Z^m we can find a nonsingular matrix C so that $Z^k V_d C$, $d = 1, \cdots, c$, span Z^m. (For a proof of this fact the reader is referred to van der Waerden (1959, page 149).) We can show that $\Pr(\sum_i Z_{in} \le zs_n \mid \sum_i Y_{in} C = 0) \to \Phi(z)$ as $n \to \infty$. This is equivalent to the statement of the lemma. Lemma 3 is thus proved.

By definition the event $\sum_i Y_{in} = 0$ is the same as $U = u$. If this is the case it also follows that $\sum_i Z_{in} = Q - E_{\hat{\gamma}}(Q \mid r)$, and

$$s_n^2 = \operatorname{Var}_{\hat{\gamma}}(\textstyle\sum_i Z_{in}) = \operatorname{Var}_{\hat{\gamma}}(Q \mid r) - \operatorname{Cov}_{\hat{\gamma}}(Q, U \mid r)[\operatorname{Var}_{\hat{\gamma}}(U \mid r)]^{-1} \operatorname{Cov}_{\hat{\gamma}}(Q, U \mid r).$$

The theorem in Section 3 is thus a consequence of these three lemmata.

REFERENCES

VON BAHR, B. and SVENSSON Å. (1978). A conditional limit theorem. Univ. Stockholm, Inst. Math. Stat. Res. Report No. 107.

COX, D. R. and HINKLEY, D. V. (1974). *Theoretical Statistics.* Chapman and Hall, London.

FISHER, R. A. (1934) *Statistical Methods for Research Workers.* Oliver and Boyd, Edinburgh.

HABERMAN, S. (1974) *The Analysis of Frequency Data.* The Univ. Chicago Press.

HABERMAN, S. (1977) Log-Linear models and frequency tables with small expected cell counts. *Ann. Statist.* **5** 1148–1169.

MARTIN-LÖF, P. (1973). Statistiska modekker. Univ. Stockholm.

VAN DER WAERDEN, B. L. (1959) *Algebra* 2. 4th ed. Springer, Berlin.

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF STOCKHOLM
BOX 6701
S-113 85 STOCKHOLM, SWEDEN