C. R. Acad. Sci. Paris, Ser. I 351 (2013) 149–154

Contents lists available at SciVerse ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Statistics/Probability Theory

Goodness-of-fit test for homogeneous Markov processes

Test d'ajustement pour les processus homogènes de Markov

Vilijandas Bagdonavičius [a], Mikhail Nikulin [b]

[a] *University of Vilnius, 24, Naugarduko, Vilnius, Lithuania*
[b] *Université Victor-Segalen, Bordeaux-1, 351, cours de la Libération, 33405 Talence cedex, France*

ARTICLE INFO

Article history:
Received 19 July 2012
Accepted after revision 31 January 2013
Available online 13 February 2013

Presented by Paul Deheuvels

ABSTRACT

We give chi-squared goodness-of-fit tests for homogeneous Markov processes with unknown transition intensities or with transition intensities of known form depending on a finite-dimensional parameter.

RÉSUMÉ

On propose des tests d'ajustement du type chi deux de l'hypothèse selon laquelle un processus stochastique d'espace d'états fini est un processus de Markov homogène, dont les intensités de transition sont, ou inconnues, ou des fonctions spécifiées d'un paramètre de dimension finie.

© 2013 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

Version française abrégée

Nous considérons des tests du type chi deux de l'hypothèse H_0 selon laquelle un processus stochastique $X(t)$, $t \geqslant 0$, d'espace d'états fini $E = \{0, 1, 2, \ldots, s\}$, est un processus de Markov homogène, dont les intensités de transition sont, ou inconnues, ou des fonctions spécifiées d'un paramètre de dimension finie θ. On considère des données censurées à droite.

Les tests sont fondés sur le vecteur des différences entre les nombres de transitions observés et les nombres de transitions prédits par le processus de Markov homogène dans les intervalles de regroupement. Les statistiques du test sont des formes quadratiques et leur loi limite est de type chi deux. On considère le cas où le choix des intervalles de regroupement dépend des données.

1. Introduction

$X(t)$, $t \geqslant 0$, is a random process with finite state space $E = \{0, 1, 2, \ldots, s\}$, and such that $X(0) = 0$, and G is the set of pairs such that a direct transition from h to h' is possible at any time t. For example, if direct transition from state 2 to state 0 is not possible, then $(2, 0)$ is not included in the set G. Let us consider the hypothesis:

H_0: $X(t)$ is a homogeneous Markov process with the transition intensities:

$$\lambda_{hh'}(\theta), \quad \theta = (\theta_1, \ldots, \theta_m)^T \in \Theta \subset \mathbf{R}^m,$$

E-mail addresses: Vilijandas.Bagdonavicius@mif.vu.lt (V. Bagdonavičius), mikhail.nikouline@u-bordeaux2.fr (M. Nikulin).

1631-073X/$ – see front matter
http://dx.doi.org/10.1016/j.crma.2013.01.014

where θ is a vector of unknown parameters and $\lambda_{hh'}$ are specified functions of these parameters. In particular, when $\lambda_{hh'}(\theta) = \theta_{hh'}$, the parameter θ has the form $\theta = (\theta_{hh'}, (h, h') \in G)$ and the hypothesis means that the process X is simply a homogeneous Markov process with constant transition intensities. As the transition functions may often depend on the same parameters, we consider a more general case. Possible examples of such situations are as follows:

Example 1. Safety model of a duplex system [7].

Suppose that two processors have the same reliability with hazard rate $\lambda(t)$. A failure of each processor may lead to a system crash or a safe shutdown of the failed processor. The conditional probability of a safe shutdown of the failed processor given that a failure of one from two functioning processors occurs is p_2. The conditional probability of a safe shutdown of the failed processor given that a failure of the remaining processor occurs is p_1.

The states are: 0 (2 processors functioning), 1 (one processor functioning after safe shutdown of the other), 2 (both processors are safely shutdown), and 3 (the system is crashed). States 2 and 3 are absorbing.

Hypothesis H_0: $X(t)$ *is a homogeneous Markov process with non-zero transition intensities*:

$$\lambda_{01}(\theta) = 2p_2\lambda, \qquad \lambda_{03}(\theta) = 2(1 - p_2)\lambda, \qquad \lambda_{12}(\theta) = p_1\lambda, \qquad \lambda_{13}(\theta) = (1 - p_1)\lambda,$$

where $\theta = (p_1, p_2, \lambda)^T$. Four possible direct transition intensities depend on three parameters.

Example 2. Redundant system with one main and $n - 1$ stand-by elements.

Let us consider a redundant system with one main and $n - 1$ stand-by units. Suppose that stand-by units function under lower stress than the main one. So reserving is "warm". Let $X(t)$ be the number of failed elements up to time t.

Hypothesis H_0: $X(t)$ *is a homogeneous Markov process with non-zero transition intensities*:

$$\lambda_{i-1,i} = \lambda + (n - i)\mu, \quad i = 1, \ldots, n.$$

So $\theta = (\lambda, \mu)^T$. n possible direct transition intensities depend on two parameters.

In the case of censored survival data, the idea of comparing observed and expected numbers of failures in time intervals is due to Akritas [2] and was developed by Hjort [5], Bagdonavičius and Nikulin [4].

Kalbfleisch and Lawless [6], Aguirre-Hernandez and Farewell [1] give Pearson-type goodness-of-fit test for stationary and time-continuous Markov regression models when stochastic processes are observed at fixed time moments.

Titman and Sharples [8] give a general goodness-of-fit test for Markov and hidden Markov models.

We give chi-squared type goodness-of-fit tests for hypothesis H_0. Choice of random grouping intervals as data functions is considered.

2. The data and the MLE estimators

Suppose that n independent copies $X^{(1)}(t), \ldots, X^{(n)}(t)$ of the random processes $X(t)$ are observed from time 0 to time τ and the ith process may be right censored at time C_i independent of X_i. Let $N_{hh'}^{(i)}(t)$ be the stochastic process counting the number of direct transitions from a state h to a state h', $(h, h') \in G$, and $N_h^{(i)}(t)$ be the number of sojourn times in state h of $X^{(i)}$ inside $[0, t]$.

The number of direct transitions of all n processes from a state h to a state h' in the interval $[0, t]$ is $N_{hh'}(t) = \sum_{i=1}^n N_{hh'}^{(i)}(t)$. Denote by $U_{hh'} = N_{hh'}(\tau)$ the number of direct transitions of all n processes from a state h to a state h' in the interval $[0, \tau]$, i.e. during the experiment, and by $Y_h(t) = \sum_{i=1}^n Y_h^{(i)}(t)$, $Y_h^{(i)}(t) = \mathbf{1}_{\{X^{(i)}(t-)=h\}}$ the number of processes in the state h just before the moment t.

Let $T_{hh's}$, $s = 1, \ldots, U_{hh'}$ be transition moments from state h to state h' in the interval $[0, \tau]$ of all Markov processes.

Denote by X_{h1s} the times of transition to the state h in the interval $[0, \tau]$ and X_{h2s} the time of the first transition from the state h after the $t = X_{h1s}$ of the same process if it occurs before censoring. X_{h2s} is the censoring time otherwise, $s = 1, \ldots, n_h$. Then $Y_h(t) = \sum_{s=1}^{n_h} \mathbf{1}_{\{t \in (X_{h1s}, X_{h2s}]\}}$.

The loglikelihood function up to time τ is (see Andersen et al. [3]):

$$\ell(\theta, \tau) = \sum_{hh' \in G} \left\{ \ln \lambda_{hh'}(\theta) N_{hh'}(\tau) - \lambda_{hh'}(\theta) \int_0^\tau Y_h(s)\,ds \right\}.$$

The score function and the matrix of the second derivatives are:

$$\dot{\ell}(\theta) = \sum_{hh' \in G} \frac{\partial}{\partial \theta} \ln \lambda_{hh'}(\theta) \left\{ N_{hh'}(\tau) - \int_0^\tau Y_h(s)\,ds \right\}, \qquad \ddot{\ell}(\theta) = \sum_{hh' \in G} \frac{\partial^2}{\partial \theta^2} \ln \lambda_{hh'}(\theta) \left\{ N_{hh'}(\tau) - \int_0^\tau Y_h(s)\,ds \right\}.$$

The Fisher information matrix is $I(\theta) = -E\ddot{\ell}(\theta)$. Set $i(\theta) = I(\theta)/n$. This matrix is estimated by

$$\hat{i} = \frac{1}{n}\sum_{hh'\in G}\int_0^\tau \frac{\partial}{\partial\theta}\ln\lambda_{hh'}(\hat{\theta})\left(\frac{\partial}{\partial\theta}\ln\lambda_{hh'}(\hat{\theta})\right)^T \mathrm{d}N_{hh'}(u) = \frac{1}{n}\sum_{hh'\in G}\frac{\partial}{\partial\theta}\ln\lambda_{hh'}(\hat{\theta})\left(\frac{\partial}{\partial\theta}\ln\lambda_{hh'}(\hat{\theta})\right)^T n_{hh'}.$$

Under regularity conditions (Andersen et al. [3]):

$$\sqrt{n}(\hat{\theta}-\theta_0) = i^{-1}(\theta_0)\frac{1}{\sqrt{n}}\dot{\ell}(\theta_0) + o_P(1), \qquad \hat{i} = i(\theta_0) + o_P(1). \tag{1}$$

3. Useful asymptotic results

If $X^{(1)}(t), \dots, X^{(n)}(t)$ are homogeneous Markov processes with the transition intensities $\lambda_{hh'}(\theta)$, then the stochastic processes (Andersen et al. [3]) $M_{hh'}(t) = N_{hh'}(t) - \lambda_{hh'}(\theta)\int_0^t Y_h(u)\,\mathrm{d}u$ are zero mean **F**-martingales in the interval $[0,\tau]$, where **F** is the filtration generated by the processes $N_{hh'}^{(i)}(t)$ and $Y_h^{(i)}(t)$, $hh' \in G$, $i = 1, \dots, n$. The equality:

$$EN_{hh'}(t) = \lambda_{hh'}(\theta)E\int_0^t Y_h(u)\,\mathrm{d}u$$

implies that the mean $EN_{hh'}(t)$ can be estimated by a nonparametric estimator $N_{hh'}(t)$ or a parametric estimator $\lambda_{hh'}(\hat{\theta})\int_0^t Y_h(u)\,\mathrm{d}u$, where $\hat{\theta}$ is the ML estimator of θ under hypothesis H_0. If the hypothesis H_0 holds then the difference between these two estimators should be small.

For this reason let us consider at first the r-dimensional stochastic process:

$$H(t) = \big(H_{hh'}(t), hh' \in G\big), \quad H_{hh'}(t) = \frac{1}{\sqrt{n}}\left(N_{hh'}(t) - \lambda_{hh'}(\hat{\theta})\int_0^t Y_h(u)\,\mathrm{d}u\right),$$

and find its limit distribution under the hypothesis H_0.

Denote by θ_0 the true value of θ. Suppose that $\frac{1}{n}\int_0^t Y_h(u)\,\mathrm{d}u \xrightarrow{P} y_h(t)$ uniformly for $t \in [0,\tau]$ and set:

$$A_{hh'}(t) = \lambda_{hh'}(\theta_0)y_h(t), \qquad C_{hh'}(t) = \frac{\partial}{\partial\theta}\lambda_{hh'}(\theta_0)y_h(t).$$

By the formulas (1):

$$\sqrt{n}(\hat{\theta}-\theta_0) = i^{-1}(\theta_0)\frac{1}{\sqrt{n}}\sum_{hh'\in G}\frac{\partial}{\partial\theta}\ln\lambda_{hh'}(\theta_0)M_{hh'}(t) + o_P(1).$$

So using the Taylor formula we have:

$$\begin{aligned}
H_{hh'}(t) &= \frac{1}{\sqrt{n}}M_{hh'}(t) + \frac{1}{\sqrt{n}}\big[\lambda_{hh'}(\theta_0) - \lambda_{hh'}(\hat{\theta})\big]\int_0^t Y_h(u)\,\mathrm{d}u \\
&= \frac{1}{\sqrt{n}}M_{hh'}(t) - C_{hh'}(t)i^{-1}(\theta_0)\frac{1}{\sqrt{n}}\sum_{ll'\in G}\frac{\partial}{\partial\theta}\ln\lambda_{ll'}(\theta_0)M_{ll'}(\tau) + o_P(1) \\
&=: M^*_{1hh'}(t) - C_{hh'}(t)i^{-1}(\theta_0)M^*_2(\tau) + o_P(1).
\end{aligned}$$

The predictable variations and covariations of $M^*_{1hh'}$ and M^*_2 are:

$$\langle M^*_{1hh'}\rangle(t) = \frac{1}{n}\lambda_{hh'}(\theta_0)\int_0^t Y_h(u)\,\mathrm{d}u \xrightarrow{P} A(t), \qquad \langle M^*_{1hh'}, M^*_{1ll'}\rangle(t) = 0, \quad hh' \neq ll',$$

$$\langle M^*_{1hh'}, M^*_2\rangle(t) = \frac{1}{n}\frac{\partial}{\partial\theta}\lambda_{hh'}(\theta_0)\int_0^t Y_h(u)\,\mathrm{d}u \xrightarrow{P} C_{hh'}(t),$$

$$\langle M^*_2\rangle(\tau) = \frac{1}{n}\sum_{hh'\in G}\frac{\partial}{\partial\theta}\ln\lambda_{hh'}(\theta_0)\left(\frac{\partial}{\partial\theta}\ln\lambda_{hh'}(\theta_0)\right)^T\lambda_{hh'}(\theta_0)\int_0^\tau Y_h(u)\,\mathrm{d}u \xrightarrow{P} i(\theta_0).$$

Suppose that the $(r+1)$-dimensional martingale $(M^*_{1hh'}, hh' \in G, M^*_2)$ satisfies the Lindeberg condition (Andersen et al. [3]). The CLT for martingales implies that:

$$\left(M^*_{1hh'}, hh' \in G, M^*_2\right) \xrightarrow{d} V^* = \left(V^*_{1hh'}, hh' \in G, V^*_2\right) \quad \text{on } D^{r+1}[0,\tau];$$

here V^* is zero mean Gaussian martingale such that for all $0 \leqslant s \leqslant t$

$$\mathbf{cov}\left(V^*_{1hh'}(s), V^*_{1hh'}(t)\right) = A_{hh'}(s), \qquad \mathbf{cov}\left(V^*_{1hh'}(s), V^*_{1ll'}(t)\right) = 0, \quad hh' \neq ll',$$
$$\mathbf{cov}\left(V^*_{1hh'}(s), V^*_2\right) = C_{hh'}(s), \qquad \mathbf{var}\left(V^*_2\right) = i(\theta_0).$$

So the following convergence holds:

$$H(t) = \left(H_{hh'}(t), hh' \in G\right) \xrightarrow{d} V(t) = \left(V_{hh'}(t), hh' \in G\right) \quad \text{on } D^r[0,\tau] \text{ as } n \to \infty; \tag{2}$$

here V is zero mean r-dimensional Gaussian process such that for all $0 \leqslant s \leqslant t$:

$$\mathbf{cov}\left(V_{hh'}(s), V_{hh'}(t)\right) = A_{hh'}(s) - C^T_{hh'}(s) i^{-1}(\theta_0) C_{hh'}(t),$$
$$\mathbf{cov}\left(V_{hh'}(s), V_{ll'}(t)\right) = -C^T_{hh'}(s) i^{-1}(\theta_0) C_{ll'}(t), \quad hh' \neq ll'.$$

4. Chi-squared test construction

The idea of the test statistic construction is as follows. For any $hh' \in G$, consider a partition of the interval $[0,\tau]$ into $k_{hh'}$ smaller intervals:

$$I_{hh'j} = (a_{hh',j-1}, a_{hh'j}], \qquad a_{hh'0} = 0, \qquad a_{hh'k_{hh'}} = \tau, \quad j = 1, \ldots, k_{hh'}.$$

We allow different partitions for different hh', because in some cases the most of transitions for one specified hh' may occur at the beginning of the observation, whereas the most of observed transitions for some other ll' may occur at the end of the observation. So, taking the same partition for both cases, it can occur that intervals with zero transitions of one or other type may be observed.

Denote by $U_{hh'j} = N_{hh'}(a_{hh'j}) - N_{hh'}(a_{hh',j-1})$ the number of observed transitions from state h to state h' of all processes in the interval $I_{hh'j}$, and

$$e_{hh'j} = \lambda_{hh'}(\hat{\theta}) \int_{a_{hh',j-1}}^{a_{hh'j}} Y_h(u)\,\mathrm{d}u,$$

the number of observed transitions predicted by the Markov model,

$$Z_{hh'j} = \frac{1}{\sqrt{n}}(U_{hh'j} - e_{hh'j}), \quad j = 1, \ldots, k_{hh'},$$

the normalized differences of observed and predicted numbers of transitions and by

$$Z_{hh'} = (Z_{hh'1}, \ldots, Z_{hh'k_{hh'}})^T,$$

the $k_{hh'}$-dimensional vector of the differences. Suppose that $G = \{(h_1 h'_1), \ldots, (h_r h'_r)\}$. Set

$$Z = \left(Z^T_{h_1 h'_1}, \ldots, Z^T_{h_r h'_r}\right)^T = (Z_{h_1 h'_1 1}, \ldots, Z_{h_1 h'_1 k_{h_1 h'_1}}, \ldots, Z_{h_r h'_r 1}, \ldots, Z_{h_r h'_r k_{h_r h'_r}})^T.$$

This vector characterizes the differences between observed and predicted numbers of all possible transitions in all intervals of partitions of the interval $[0,\tau]$.

The test statistic is based on the $K = k_{h_1 h'_1} \cdot \cdots \cdot k_{h_r h'_r}$-dimensional random vector Z. The convergence (2) implies that:

$$Z \xrightarrow{d} Y \sim N_K(0, V), \qquad Z^T V^- Z \xrightarrow{d} \chi^2\left(rang(V)\right),$$

where V is the limit covariance matrix which is consistently estimated by

$$\hat{V} = \hat{\mathrm{A}} - \hat{C}^T \hat{i}^{-1} \hat{C}, \quad \hat{C} = (\hat{C}_{h_1 h'_1} | \hat{C}_{h_2 h'_2} | \ldots | \hat{C}_{h_r h'_r}),$$

where $\hat{C}$ is $m \times K$ matrix,

$$\hat{C}_{hh'} = \begin{pmatrix} \hat{C}_{hh'11} & \cdots & \hat{C}_{hh'1k_{hh'}} \\ \cdots & \cdots & \cdots \\ \hat{C}_{hh'm1} & \cdots & \hat{C}_{hh'mk_{hh'}} \end{pmatrix}, \qquad \hat{C}_{hh'ij} = \frac{1}{n} \sum_{s:T_{hh's} \in I_{hh'j}} \frac{\partial}{\partial \theta_i} \ln \lambda_{hh'}(\hat{\theta}),$$

$\hat{A}$ – $K \times K$ diagonal matrix – with the diagonal elements $\hat{A}_{h_1 h'_1 1}, \ldots, \hat{A}_{h_1 h'_1 k_{h_1 h'_1}}, \ldots, \hat{A}_{h_r h'_r 1}, \ldots, \hat{A}_{h_r h'_r k_{h_r h'_r}}$, $\hat{A}_{hh'j} = U_{hh'j}/n$.

A test for hypothesis H_0 can be based on the statistic:

$$Y^2 = Z^T \hat{V}^- Z.$$

The random variable Y^2 can be interpreted as a distance between observed and predicted numbers of transitions. Note that:

$$\hat{C}\hat{\mathring{A}}^{-1}\hat{C}^T = \sum_{hh' \in G} \hat{C}_{hh'} \hat{\mathring{A}}_{hh'}^{-1} \hat{C}_{hh'}^T, \quad \hat{C}_{hh'} = [C_{hh'ij}]_{m \times k},$$

$\hat{\mathring{A}}_{hh'}^{-1}$ being $k_{hh'} \times k_{hh'}$ diagonal matrix with elements $\hat{A}_{hh'j}$ on the diagonal, $j = 1, \ldots, k_{hh'}$. So,

$$\begin{aligned} \hat{C}\hat{\mathring{A}}^{-1}\hat{C}^T &= \sum_{hh' \in G} \frac{\partial}{\partial \theta} \ln \lambda_{hh'}(\hat{\theta}) (\hat{A}_{hh'1}, \ldots, \hat{A}_{hh'k_{hh'}}) \mathring{A}_{hh'}^{-1} (\hat{A}_{hh'1}, \ldots, \hat{A}_{hh'k_{hh'}})^T \left(\frac{\partial}{\partial \theta} \ln \lambda_{hh'}(\hat{\theta}) \right)^T \\ &= \frac{1}{n} \sum_{hh' \in G} \frac{\partial}{\partial \theta} \ln \lambda_{hh'}(\hat{\theta}) \left(\frac{\partial}{\partial \theta} \ln \lambda_{hh'}(\hat{\theta}) \right)^T n_{hh'} = \hat{i}, \end{aligned}$$

and consequently,

$$\hat{V} = \hat{A} - \hat{C}^T \left(\hat{C}\hat{\mathring{A}}^{-1}\hat{C}^T \right)^{-1} \hat{C}.$$

Note that

$$\hat{V}^- = \hat{A}^{-1} + \hat{A}^{-1}\hat{C}^T\hat{C}$$

is a general inverse of $\hat{V}$ because $\hat{V}\hat{V}^-\hat{V} = \hat{V}$. The limit distribution of the test statistic:

$$Y^2 = Z^T \hat{V}^- Z = Z^T \hat{A}^{-1} Z + Q, \quad Q = Z^T \hat{A}^{-1} \hat{C}^T \hat{C} Z$$

is chi-square with $Tr(V^- V)$ degrees of freedom. Note that:

$$V^- V = E_K - \mathring{A}^{-1} C^T \left(C\mathring{A}^{-1} C^T \right)^{-1} C + \mathring{A}^{-1} C^T C \mathring{A} - \mathring{A}^{-1} C^T C C^T \left(C\mathring{A}^{-1} C^T \right)^{-1} C,$$

$Tr(PQ) = Tr(QP)$ for any $m \times n$ matrix P and $n \times m$ matrix Q; so,

$$\begin{aligned} &Tr(E_K) = K, \quad Tr\left(\mathring{A}^{-1} C^T C \mathring{A}\right) = Tr\left(C\mathring{A}\mathring{A}^{-1} C^T\right) = Tr\left(CC^T\right), \\ &Tr\left(\mathring{A}^{-1} C^T \left(C\mathring{A}^{-1} C^T\right)^{-1} C\right) - Tr\left(C\mathring{A}^{-1} C^T \left(C\mathring{A}^{-1} C^T\right)^{-1}\right) = Tr(E_m) = m, \\ &Tr\left(\mathring{A}^{-1} C^T C C^T \left(C\mathring{A}^{-1} C^T\right)^{-1} C\right) = Tr\left(\left(C\mathring{A}^{-1} C^T\right)^{-1} C\mathring{A}^{-1} C^T C C^T\right) = Tr\left(CC^T\right). \end{aligned}$$

We have $Tr(V^- V) = K - m$.

The statistic Y^2 can be written in the form:

$$Y^2 = \sum_{hh' \in G} \sum_{j=1}^{k_{hh'}} \frac{(U_{hh'j} - e_{hh'j})^2}{U_{hh'j}} + Q,$$

where

$$Q = Z^T \hat{A}^{-1} \hat{C}^T \hat{C} Z = \sum_{hh' \in G} Z_{hh'}^T \hat{A}_{hh'}^{-1} \hat{C}_{hh'}^T \sum_{ll' \in G} \hat{C}_{ll'} Z_{ll'}, \quad Z_{hh'} = (Z_{hh'1}, \ldots, Z_{hh'k})^T.$$

Note that:

$$\begin{aligned} \sum_{hh' \in G} Z_{hh'}^T \hat{A}_{hh'}^{-1} \hat{C}_{hh'}^T &= \sum_{hh' \in G} Z_{hh'}^T \hat{A}_{hh'}^{-1} (\hat{A}_{hh'1}, \ldots, \hat{A}_{hh'k})^T \left(\frac{\partial}{\partial \hat{\theta}} \ln \lambda_{hh'}(\hat{\theta}) \right)^T \\ &= \sum_{hh' \in G} Z_{hh'}^T (1, \ldots, 1)^T \left(\frac{\partial}{\partial \hat{\theta}} \ln \lambda_{hh'}(\hat{\theta}) \right)^T \\ &= \frac{1}{\sqrt{n}} \sum_{hh' \in G} (n_{hh'} - e_{hh'}) \left(\frac{\partial}{\partial \hat{\theta}} \ln \lambda_{hh'}(\hat{\theta}) \right)^T = \frac{1}{\sqrt{n}} \dot{\ell}(\hat{\theta}) = 0. \end{aligned}$$

So, $Q = 0$ and the test statistic has the form:

$$Y^2 = Z^T \hat{V}^- Z = \sum_{hh' \in G} \sum_{j=1}^{k} \frac{(U_{hh'j} - e_{hh'j})^2}{U_{hh'j}} \xrightarrow{d} \chi^2(K - m).$$

5. Choice of random grouping intervals

Similarly as in [4] it can be shown that the limit law of the test statistics does not change if the limits of grouping intervals are chosen as random data functions in the following way: $a_{hh'j}$ verifies the equation:

$$\int_0^{a_{hh'j}} Y_h(s)\,\mathrm{d}s = \frac{j}{k} \int_0^{\tau} Y_h(s)\,\mathrm{d}s.$$

Under such a choice

$$e_{hh'1} = \cdots = e_{hh'k_{hh'}} = \frac{1}{k_{hh'}} \lambda_{hh'}(\hat{\theta}) \int_0^{\tau} Y_h(s)\,\mathrm{d}s = \frac{1}{k} \lambda_{hh'}(\hat{\theta}) S_h, \quad S_h = \sum_{s=1}^{n_h} [X_{h2s} - X_{h1s}].$$

Numerically, the point $a_{hh'j}$ is found very simply by the bisection method.

6. Chi-squared tests for homogeneous Markov processes

Chi-squared test: the null hypothesis is rejected with approximate significance level α if $Y^2 > \chi^2_\alpha(K - m)$.

6.1. Examples

1. Transition intensities of unknown form. Suppose that $\theta = (\lambda_{hh'}, hh' \in G)$, the intensities $\lambda_{hh'}$ being completely unknown. Then:

$$\ell_{hh'}(\lambda_{hh'}) = U_{hh'} \ln \lambda_{hh'} - \lambda_{hh'} S_h, \quad S_h = \sum_{s=1}^{n_h} [X_{h2s} - X_{h1s}],$$

and

$$\hat{\lambda}_{hh'} = \frac{U_{hh'}}{S_h}.$$

The null hypothesis is rejected with approximate significance level α if $Y^2 > \chi^2_\alpha(K - r)$.

2. Safety model of a duplex system.

Under H_0

$$\ell(\theta) = U \ln \lambda - \lambda(2S_0 + S_1) + U_{01} \ln(2p_2) + U_{12} \ln p_1 + U_{13} \ln(1 - p_1) + U_{03} \ln\bigl(2(1 - p_2)\bigr),$$

where $U = \sum_{hh' \in G} U_{hh'}$.

The ML estimators are explicit:

$$\hat{p}_1 = \frac{U_{12}}{U_{12} + U_{13}}, \qquad \hat{p}_2 = \frac{U_{01}}{U_{01} + U_{03}}, \qquad \hat{\lambda} = U/S,$$

where $S = 2S_0 + S_1$.

The null hypothesis is rejected with approximate significance level α if $Y^2 > \chi^2_\alpha(K - 3)$.

References

[1] R. Aguirre-Hernandez, V.T. Farewell, A Pearson-type goodness-of-fit test for stationary and time-continuous Markov regression models, Stat. Med. 21 (2002) 1899–1911.
[2] M.G. Akritas, Pearson-type goodness-of-fit tests: the univariate case, J. Am. Stat. Assoc. 83 (1988) 222–230.
[3] P.K. Andersen, O. Borgan, R.D. Gill, N. Keiding, Statistical Models Based on Counting Processes, Springer-Verlag, New York, 1993.
[4] V. Bagdonavičius, M. Nikulin, Chi-squared goodness-of-fit test for right censored data, Int. J. Appl. Math. Stat. 24 (2011) 30–50.
[5] N.L. Hjort, Goodness of fit tests in models for life history data based on cumulative hazard rates, Ann. Stat. 18 (1990) 1221–1258.
[6] J.D. Kalbfleisch, J.F. Lawless, The analysis of panel data under a Markov assumption, J. Am. Stat. Assoc. 80 (1985) 863–871.
[7] A. Rindos, S. Woolet, I. Viniotis, K. Trivedi, Exact methods for the transient analysis of nonhomogeneous continuous time Markov chains, in: W.J. Stewart (Ed.), Computations with Markov Chains, Kluwer Academic Publishers, Boston, 1995, pp. 121–133 (Chapter 8).
[8] A.C. Titman, L.D. Sharples, A general goodness-of-fit test for Markov and hidden Markov models, Stat. Med. 27 (2008) 2177–2195.