



Cairo University
Institute Of Statistical Studies & Research



THE ANNUAL CONFERENCE ON STATISTICS, COMPUTER SCIENCE AND OPERATIONS RESEARCH

APPLIED STATISTICS

**Institute of Statistical Studies and
Research
Cairo University
Egypt
22-25, Dec ,2008**



INTERNATIONAL ASSOCIATION OF AGRICULTURAL ECONOMISTS
P.O. Box 100, 1000 Amsterdam, The Netherlands
Tel: +31 (0)20 691 9111, Fax: +31 (0)20 691 9112
E-mail: iaae@iaae.nl

Index APPLIED STATISTICS

1	Measuring the Creditability of Borrowers in Egyptian Business Loans using a Generalized Partial Linear Model Magda Mohamed M. Haggag	1-21
2	Estimation of the Cardiac Time-Varying Parameters During the Ejection Phase of the Cardiac Cycle Using the Ito Calculus Ahmed S. Abutaleb, J. Melbin	22-41
3	On Shrinkage of the Simple Linear Regression .Parameters M. Ahmed AL_Robassi	42-49
4	Comparison Between Discriminant and Regression Analyses with an Application Hassan S.M.	50-61
5	Fuzzy Logistic Regression A Multi-objective Programming Approach Ramadan Hamed, Aly El-Hefnawy , Maha M. El-Ashram , Hesham A. Abdalla	62-78
6	An Application of Markov Random Walk to a Dialysis System M. A. El-Shehawey, M. F. Yassen, and N. G. Ragab	79-97
7	On Markov Chain Roulette Type Game EL-Shehawey M. A., EL-Shreef GH. A	98-107
8	E-Optmal Design For Detecting The Main Effects 2^k – Factorial Paired comparison Experiments With Full Profiles Essam A. Ahmed , Sayed R.El-Shair , Osama F. Abdelhalim	108-119

Measuring the Creditability of Borrowers in Egyptian Business Loans using a Generalized Partial Linear Model

BY

Magda Mohamed M. Haggag

*A Lecturer of Statistics,
Department of Statistics, Mathematics, and Insurance
Faculty of Commerce, Alexandria University (Damanhour Branch)*

Abstract

This paper considers improving the efficiency of regression estimators using a semiparametric generalized linear model, known as a generalized partial linear model (GPLM). The used method in estimation of GPLM is known as the profile-likelihood method, which is based on using kernel smoothing functions in the estimation of the nonparametric component of the model. I will derive the algorithms for the estimation process and develop these algorithms for the generalized partial linear model (GPLM) with a binary response. Data on business loans (credit scoring data) are used as a real application of GPLM with a binomially distributed response variable. The results give asymptotically efficient estimators of GPLM than that obtained using the parametric generalized linear model (GLM).

Key Words: asymptotically efficient estimators, Backfitting estimator, business loans, credit scoring, efficiency, generalized linear model, generalized partial linear model, kernel smoothing, misclassification rates, profile-likelihood, quasi-likelihood, semiparametric estimation.

1. Introduction

A parametric model is one where the set $\mathcal{P}$ is smoothly describable as the image of a set Θ in $\mathbb{R}^k$ by a map $\theta \rightarrow P_\theta$. Estimation of θ is determined to a large degree by smoothness properties of this map, where a smooth map means a continuously differentiable map on M , where M is a collection of probability measures on $(X, \mathcal{B})$, where X is the sample space of data and $\mathcal{B}$ is its Borel σ -field. Consider a model for i.i.d. random variables $X_1, \dots, X_n$ with common unknown probability distribution $P \in \mathcal{P}$ on a measurable space $(X, \mathcal{B})$. The probability

distribution P is dominated by μ ; i.e., there exists a set $M_\mu \equiv \{P \in \mathcal{M} : P \ll \mu\}$, where μ here is denoted to a fixed σ -finite measure on $(X, \mathcal{B})$. The model $\mathcal{P}$ is defined to be parametric if:

$$\mathcal{P} = \{P_\theta : \theta \in \Theta\}, \quad (1)$$

where $\Theta \subset \mathbb{R}^k$, and the parameterization $\theta \rightarrow P_\theta$ is smooth.

The model $\mathcal{P}$ is defined to be a regular parametric model if it has a regular parametrization, i.e.,

1. The map $\theta \rightarrow S(\theta)$, where $S(\theta) = \left[\frac{dP_\theta}{d\mu} \right]^{\frac{1}{2}} = (f_\theta)^{\frac{1}{2}}$ is continuously Frechet differentiable from Θ to $L_2(\mu)$, with derivative $\dot{S}(\theta)$ a k -vector of elements of $L_2(\mu)$, where $L_2(\mu)$ is a Hilbert space of μ -square integrable functions.
2. The $k \times k$ matrix $\int \dot{S}(\theta) \dot{S}^T(\theta) d\mu$ is nonsingular. (Ritov and Bickel, 1990).

Semiparametric models are naturally parameterized by both Euclidean and infinite dimensional parameters. Or, in another description, they are described as ones where every possible probability distribution is at least approximable by distributions in the model. In this sense, all regular parametric models are also semiparametric.

The model $\mathcal{P}$ is defined to be semiparametric if:

$$\mathcal{P} = \{P_{\theta,m} : \theta \in \Theta, m \in \mathcal{M}\} \quad (2)$$

where Θ is an open subset of $\mathbb{R}^p$ and $\mathcal{M}$ is an arbitrary set of nonparametric class of distributions or some other infinite-dimensional set.

The purpose of this paper is to measure the creditability of borrowers in Egyptian business loans using a generalized partial linear model (GPLM). Also, the algorithms used in the estimation of GPLM are derived using the profile-likelihood method and developed to the case of a binary response variable Y . A real application using data on business loans (credit scoring data) is shown with improved estimators, i.e., asymptotically efficient estimators. All results are obtained using XploRe (4.8). (Hardle et al. (1995), (2000a), and (2000b)).

This paper is organized as follows: a generalized partial linear model (GPLM) is identified as an extension of a parametric generalized linear model (GLM) in section (2). Estimation algorithms of a GLM and a GPLM are reviewed in section (3). Asymptotic statistical properties of the GPLM estimators are considered in section (4). Section (5) introduces an application of estimation algorithm methodology which is developed for GPLM in the case of a binary response variable Y , and measuring the creditability of borrowers of GPLM using data on Egyptian business loans (credit scoring data). Section (6) is for summary.

2. Generalized partial linear model (GPLM)

A semiparametric generalized linear model known as a generalized partial linear model (GPLM) is an extension of a generalized linear model (GLM),

$$E(Y|Z) = G(Z^T \beta), \quad (3)$$

where G is a known and monotone link function, and β is an unknown finite dimensional parameter. The generalized linear model has a wide propagation refers, in particular, to the ability of GLM to handle both discrete and continuous variables in the index $Z^T \beta$. In most cases, the nonlinear relation between the response variable and the regressors, known as the link function, in GLM is fixed and monotone. In the generalized partial linear model (GPLM), a more complex relationship between the response and the regressors is defined by a non monotone link function such that,

$$E(Y|Z) = G(U^T \beta + m(W)), \quad (4)$$

Where $Z = (U^T, W^T)^T$, G is a known and non monotone link function, β is an unknown finite dimensional parameter, and $m(\cdot)$ is an unknown infinite-dimensional function, i.e. a smooth function.

The first term of (4), $U^T \beta$, represents the parametric component of the model, and the second term, $m(W)$, represents the nonparametric component of the model. Therefore, the term semiparametric model refers to the model which contains these two components. (Sperlich et al. (2006)).

If the distribution of Y belongs to an exponential family, a likelihood method, known as a profile-likelihood introduced by Severini and Wang (1992), is used in the estimation of GPLM. On the other hand, if the exact distribution information of Y is not available but the first two moments are specified, then a quasi-likelihood method is used in the estimation of GPLM. A quasi-likelihood function has properties similar to those of a likelihood function since the exponential family distribution is specified by its first two moments. (Severini and Staniswalis (1994)).

3. An Overview of the Estimation Methods

3.1 Estimation of the GLM

Maximum likelihood estimate of the parameter β in a GLM is generally not available in closed form but can be obtained via an algorithm known as iteratively weighted least squares (IWLS). IWLS is one of the practical ways in which GLMs are in fact "general", providing MLEs for a wide class of commonly used models.

IWLS is based on regressing a new response variable γ on regressors, Z_j 's, which is a linearized form of the link function applied to the response Y and weights, D , which are functions of the fitted values $\hat{\mu}$. The process is iterative because both the adjusted response variable γ and the weights D depend on the fitted values, for which only a trial estimate of the parameters, $\hat{\beta}$, are available.

The iteratively weighted least squares (IWLS) algorithm is performed as follows:

1. Form the working responses with typical value

$$\gamma_i^{(0)} = \hat{\eta}_i^{(0)} + (y_i - \hat{\mu}_i^{(0)}) \dot{G}^{-1}(\hat{\mu}_i^{(0)}), \quad i=1, \dots, n,$$

where $\hat{\mu}_i^{(0)}$ and $\hat{\eta}_i^{(0)} = G^{-1}(\hat{\mu}_i^{(0)})$ are the initial estimators of the means and linear predictors respectively. $\dot{G}^{-1}(\hat{\mu}_i^{(0)}) = \frac{\partial \hat{\eta}_i^{(0)}}{\partial \hat{\mu}_i^{(0)}}$ is the derivative of the function $G^{-1}(\hat{\mu}_i^{(0)})$ evaluated at $\hat{\eta}_i^{(0)}$. The weights (quadratic) are defined by

$$\hat{D}_i^{(0)} = \frac{1}{[\dot{G}^{-1}(\hat{\mu}_i^{(0)})]^2 V_i^{(0)}}, \quad i=1, \dots, n,$$

where $V_i^{(0)}$ is the variance function of Y_i which is based on $\hat{\mu}_i^{(0)}$.

2. Regress $\gamma_i^{(0)}$ on X_j 's with weights $\hat{D}_i^{(0)}$ to give new estimators, $\hat{\beta}^{(1)}$, of the parameters. From these estimators form a new estimator $\hat{\eta}_i^{(1)}$, of the linear predictor. Given $\hat{\mu}_i^{(t)}$ and $\hat{\eta}_i^{(t)}$, form the adjusted response variable as follows

$$\gamma_i^{(t)} = \hat{\eta}_i^{(t)} + (y_i - \hat{\mu}_i^{(t)}) \dot{G}^{-1}(\hat{\mu}_i^{(t)}), \quad i=1, \dots, n,$$

and iterative weights:

$$\hat{D}_i^{(t)} = \frac{1}{[\dot{G}^{-1}(\hat{\mu}_i^{(t)})]^2 V_i^{(t)}}, \quad i=1, \dots, n.$$

3. Calculate $\hat{\beta}^{(t+1)}$ by weighted least squares regression of $\gamma^{(t)}$ on X as

$$\hat{\beta}^{(t+1)} = \left(X^T \hat{D}^{(t)} X \right)^{-1} X^T \hat{D}^{(t)} \hat{Y}^{(t)},$$

where $\hat{D}^{(t)} = \text{diag} \left(\hat{D}_i^{(t)} \right)$. Define $\hat{\eta}^{(t+1)}$ and $\hat{\mu}^{(t+1)}$ as:

$$\hat{\eta}^{(t+1)} = X \hat{\beta}^{(t+1)}$$

and

$$\hat{\mu}^{(t+1)} = G \left(\hat{\eta}^{(t+1)} \right)$$

4. Repeat steps 2 and 3 until changes in $\hat{\beta}$ are sufficiently small. The asymptotic variance of $\hat{\beta}$ is given by

$$V \left(\hat{\beta} \right) = \phi \left(X^T D X \right)^{-1},$$

where $D = \text{diag}(D_1, \dots, D_n)$ computed at the final iteration. The IWLS algorithm does not involve the dispersion parameter ϕ , which is usually either fixed a priori by the model (e.g., for binomial or Poisson models) or estimated after computing residuals with $\hat{\beta}$ (as in the Gaussian case). If ϕ is unknown, an estimator is required in order to compute the standard errors.

In Models with canonical link IWLS is simply the Newton-Raphson method, and more generally it coincides with the Fisher scoring method. (Nelder and Wedderburn (1972); McCullagh and Nelder (1989); Dobson (1990); and Muller (2004)).

3.2 Estimation of the GPLM

The estimation method that will be used for GPLM is the Profile likelihood method. The profile likelihood method introduced by Severini and Wang (1992) is based on the fact that the conditional distribution of Y given U and W is parametric. The idea of this method is as follows:

- a. Fix the parametric component of the model, i.e., the parameter vector β .
- b. Estimate the nonparametric component of the model which depends on this fixed β , i.e. $m_{\beta}(\cdot)$ by some type of smoothing method to obtain the estimate $\hat{m}_{\beta}(\cdot)$.
- c. Use the estimate $\hat{m}_{\beta}(\cdot)$ to construct a profile likelihood for the parametric component using either a true likelihood or quasi-likelihood function.
- d. The profile likelihood function is then used to obtain an estimator of the parametric component of the model using a maximum likelihood method.

Thus the profile likelihood method aims to separate the estimation process into two parts, the parametric part which is estimated by a parametric method and the nonparametric part which is estimated by a nonparametric method.

Murphy and Vaart (2000) showed that the full likelihood method fails in semiparametric models. In semiparametric models the observed information if it exists, would be at best an infinite-dimensional operator. They used a profile likelihood rather than a full likelihood to overcome this problem.

The algorithm for Profile likelihood method is derived as follows:

(I) Derivation of different likelihood functions:

*** For the parametric component of the model**

The objective function for the parametric component is the parametric profile likelihood function which is maximized to obtain an estimate for β is given by:

$$\ell(\beta) = \sum_{i=1}^n \ell(\mu_{i,\beta}, y_i) = \sum_{i=1}^n \ell_i, \quad (5)$$

where $\ell(\cdot)$ denotes the log-likelihood or quasi-likelihood function,

$$\mu_{i,\beta} = G[U_i^T \beta + m_\beta(W_i)], \quad (6)$$

and

$$\ell_i = \ell(\mu_{i,\beta}, y_i). \quad (7)$$

The function in (5) is maximized to obtain an estimate for β .

*** For the nonparametric component of the model**

The objective function for the nonparametric component is a smoothed or a local likelihood function which is given by

$$\ell_H[m_\beta(w)] = \sum_{i=1}^n \ell(\mu_{i,m_\beta(w)}, y_i) K_H(w - W_i), \quad (8)$$

where

$$\mu_{i,m_\beta(w)} = G[U_i^T + m_\beta(w)] = G(\eta), \quad (9)$$

and the local weight $K_H(w - W_i) = H^{-1} K(w - W_i)$ is the kernel weight with K denoting a multidimensional kernel function and H is a bandwidth matrix. The function in (8) is maximized to obtain an estimate for the smooth function $m_\beta(w)$ at a point w .

(II) Maximization of the likelihood functions:

The maximization of the local likelihood in (8) requires solving

$$0 = \sum_{i=1}^n \dot{\ell}(G(\eta), y_i) K_H(w - W_i), \quad (10)$$

with respect to $m_\beta(t)$, and $\dot{\ell}(\cdot)$ denotes the first derivative of $\ell(\cdot)$. The maximization of the profile likelihood in (3) requires solving

$$0 = \sum_{i=1}^n \dot{\ell}(G(\eta), y_i) (W_i + \nabla_{m_\beta}(W_i)), \quad (11)$$

with respect to the coefficient vector β . The vector $\nabla_{m_\beta}(W_i)$ denotes here the vector of all partial derivatives of m_β with respect to β . A further differentiation of (10) with respect to β leads to an explicit expression for ∇_{m_β} as follows:

$$\nabla_{m_\beta}(w) = - \frac{\sum_{i=1}^n \ddot{\ell}(G(\eta), y_i) K_H(w - W_i) U_i}{\sum_{i=1}^n \ddot{\ell}(G(\eta), y_i) K_H(w - W_i)}, \quad (12)$$

where $\ddot{\ell}$ denotes the second derivative of $\ell(G(\eta), y_i)$.

(III) Solving the Maximized likelihood functions:

Equations (10) and (11) can only be solved iteratively. Severini and Satniswalis (1994) presented a Newton-Raphson type algorithm for this problem as follows:
Let

$$\begin{aligned} m_j &= m_\beta(W_j), \\ \eta_i &= (U_i^T \beta + m_i), \\ \eta_{ij} &= (U_i^T \beta + m_j), \\ \ell_i &= \ell(G(\eta_i), y_i), \\ \ell_{ij} &= \ell(G(\eta_{ij}), y_i) \end{aligned}$$

Further let $\dot{\ell}_i, \ddot{\ell}_i, \dot{\ell}_{ij}$ and $\ddot{\ell}_{ij}$ be the first and second derivatives of ℓ_i and ℓ_{ij} with respect to their first argument. All values will be calculated at the observations W_i instead of the free parameter w . Then (10) and (11) are transformed to

$$0 = \sum_{i=1}^n \dot{\ell}_{ij} K_H(W_i - W_i), \quad (13)$$

$$0 = \sum_{i=1}^n \dot{\ell}_i (U_i + \nabla_{m_\beta}(W_i)). \quad (14)$$

The estimate of $\nabla_{m_\beta} = \nabla_{m_\beta}(W_j)$ based on (12) is necessary to estimate β :

$$\nabla_{m_j} = - \frac{\sum_{i=1}^n \ddot{\ell}_{ij} \mathcal{K}_H(W_i - W_j) U_i}{\sum_{i=1}^n \ddot{\ell}_{ij} \mathcal{K}_H(W_i - W_j)} \quad (15)$$

Equations (13) and (14) imply the following iterative Newton-Raphson Type algorithm:

First: Initial values:

Different strategies are found for obtaining start values of the estimators:

- Using $\hat{\beta}^{(0)}$ and $\hat{m}_j^{(0)}$ obtained from fitting a parametric generalized linear model (GLM).
- Using $\beta = 0$ and $m_j = G^{-1}[(Y_j + \bar{Y})/2]$ with the adjustment for Binomial responses as $m_j = G^{-1}[(Y_j + 0.5)/2]$.
- Using $\beta = 0$ and $m_j = G^{-1}[Y_j]$ with the adjustment for Binomial responses as $m_j = G^{-1}[(Y_j + 0.5)/2]$. (See Severini and Staniswalis (1994); and Hardle et al. (2000a)).

Second: The updating step for β :

$$\beta^{new} = \beta - B^{-1} \sum_{i=1}^n \ddot{\ell}_i \tilde{U}_i, \quad (16)$$

where B is a Hessian type matrix defined as

$$B = \sum_{i=1}^n \ddot{\ell}_i \tilde{U}_i \tilde{U}_i^T,$$

and

$$\tilde{U}_j = U_j + \nabla_{m_j} = U_j - \frac{\sum_{i=1}^n \ddot{\ell}_{ij} \mathcal{K}_H(W_i - W_j) U_i}{\sum_{i=1}^n \ddot{\ell}_{ij} \mathcal{K}_H(W_i - W_j)}.$$

The updating step for β can be summarized in a closed matrix form as follows:

$$\beta^{new} = (\tilde{U}^T D \tilde{U})^{-1} \tilde{U}^T D \tilde{Z}, \quad (17)$$

where

$$\tilde{U} = (I - S^p) U, \quad (18)$$

$$\tilde{Z} = \tilde{U} \beta - D^{-1} v, \quad (19)$$

where U is the design matrix with rows U_i^T , I is an (nxn) identity matrix,

$$v = (\ddot{\ell}_1, \dots, \ddot{\ell}_n), \quad (20)$$

$$D = \text{diag}(\ddot{\ell}_1, \dots, \ddot{\ell}_n), \quad (21)$$

and S^p is a smoother matrix with elements:

$$(S^p)_{ij} = \frac{\ddot{\ell}_{ij} K_H(W_i - W_j)}{\sum_{i=1}^n \ddot{\ell}_{ij} K_H(W_i - W_j)}. \quad (22)$$

Third: The updating-step for $m_j = m_\beta(w_j)$

The function $m_j = m_\beta(w_j)$ is updated by:

$$\hat{m}_j^{k+1} = \hat{m}_j^k - \frac{\sum_{i=1}^n \ddot{\ell}_{ij} K_H(W_i - W_j)}{\sum_{i=1}^n \ddot{\ell}_{ij} K_H(W_i - W_j)}. \quad (23)$$

It is noted that the function $\ddot{\ell}_{ij}(\cdot)$ can be replaced by its expectation with respect to Y to obtain a Fisher scoring type algorithm.

The previous procedure can be summarized as follows:

First: Updating step for β

$$\beta^{\text{new}} = (\tilde{U}^T D \tilde{U})^{-1} \tilde{U}^T D \tilde{Z}.$$

Second: Updating step for m_j

$$m_j^{\text{new}} = m_j - \frac{\sum_{i=1}^n \ddot{\ell}_{ij} K_H(W_i - W_j)}{\sum_{i=1}^n \ddot{\ell}_{ij} K_H(W_i - W_j)}.$$

3. Asymptotic Statistical Properties of GPLM Estimators

Severini and Staniswalis (1994) showed that under some regularity conditions the estimator $\hat{\beta}$ has the following properties:

1. $\sqrt{n}$ -consistent estimate for β .
2. Asymptotically normal.
3. Its limiting covariance has a consistent estimator.
4. Asymptotically efficient (has asymptotically minimum variance).

The nonparametric function m can be estimated (in the univariate case) with the usual univariate rate of convergence.

In summary the parametric and nonparametric estimators of GPLM have the following asymptotic properties:

$$\begin{aligned}\hat{\beta} &= \beta_0 + O_p\left(n^{-1/2}\right), \\ \|\hat{m}_{\hat{\beta}} - m_0\| &= o_p\left(n^{-1/4}\right), \quad \text{as } n \rightarrow \infty. \\ \hat{\sigma}^2 &= \sigma_0^2 + o_p(1),\end{aligned}$$

where β_0 , m_0 , and σ_0^2 are the true parameter values so that

$$\text{var}(Y | U, W) = \sigma_0^2 \text{var}[G(U^T \beta_0 + m_0(W))].$$

4. Applications

4.1 Developing the estimates of a univariate GPLM for the binomially distributed response

We develop the estimates of GPLM for the parametric component β and the nonparametric one $m(w)$ when the response variable Y belongs to a binomial distribution. The developing will be shown in the univariate case model using the profile likelihood method. The iterative procedure that will be used is the Newton-Raphson algorithm:

4.1.1 Updating step for m_j

Equation (23) will be

$$\hat{m}_j^{(k+1)} = \hat{m}_j^{(k)} - \frac{\sum_{i=1}^n \dot{\ell}_i(U_i \hat{\beta}^{(k)} + \hat{m}_j^{(k)}) K_h(W_i - W_j)}{\sum_{i=1}^n \ddot{\ell}_i(U_i \hat{\beta}^{(k)} + \hat{m}_j^{(k)}) K_h(W_i - W_j)}, \quad (24)$$

where

$$\dot{\ell}_i(U_i \hat{\beta}^{(k)} + \hat{m}_j^{(k)}) = \left[Y_i - \frac{\exp(U_i \hat{\beta}^{(k)} + \hat{m}_j^{(k)})}{1 + \exp(U_i \hat{\beta}^{(k)} + \hat{m}_j^{(k)})} \right],$$

and

$$\ddot{\ell}_i(U_i \hat{\beta}^{(k)} + \hat{m}_j^{(k)}) = \left[-\frac{\exp(U_i \hat{\beta}^{(k)} + \hat{m}_j^{(k)})}{[1 + \exp(U_i \hat{\beta}^{(k)} + \hat{m}_j^{(k)})]^2} \right].$$

4.1.2 Updating step for β

Equation (16) will be

$$\hat{\beta}^{(k+1)} = \hat{\beta}^{(k)} - \frac{1}{B} \sum_{i=1}^n \dot{\ell}_i(U, \hat{\beta}^{(k)} + \hat{m}_i^{(k+1)}) \tilde{U}_i^{(k)}, \quad (25)$$

Where

$$\dot{\ell}_i(U, \hat{\beta}^{(k)} + \hat{m}_i^{(k+1)}) = \left[Y_i - \frac{\exp(U, \hat{\beta} + \hat{m}_i^{(k+1)})}{1 + \exp(U, \hat{\beta} + \hat{m}_i^{(k+1)})} \right],$$

$$\tilde{U}_i^{(k)} = U_i - \frac{\sum_{j=1}^n \dot{\ell}_j(U, \hat{\beta}^{(k)} + \hat{m}_j^{(k+1)}) K_h(W_i - W_j) U_j}{\sum_{j=1}^n \dot{\ell}_j(U, \hat{\beta}^{(k)} + \hat{m}_j^{(k+1)}) K_h(W_i - W_j)},$$

$$\ddot{\ell}_i(U, \hat{\beta}^{(k)} + \hat{m}_i^{(k+1)}) = \left[- \frac{\exp(U, \hat{\beta} + \hat{m}_i^{(k+1)})}{[1 + \exp(U, \hat{\beta} + \hat{m}_i^{(k+1)})]^2} \right],$$

$$B = \sum_{i=1}^n \ddot{\ell}_i(U, \hat{\beta}^{(k)} + \hat{m}_i^{(k+1)}) \tilde{U}_i^{(k)2},$$

$j=1,2,\dots,n$ is the number of observations, and $k=0,1,2,\dots$ is the number of iterations. The procedure will be iterated until convergence. For $k=0$,

$$\hat{m}_j^{(1)} = \hat{m}_j^{(0)} - \frac{\sum_{i=1}^n \dot{\ell}_i(U, \hat{\beta}^{(0)} + \hat{m}_i^{(0)}) K_h(W_i - W_j)}{\sum_{i=1}^n \dot{\ell}_i(U, \hat{\beta}^{(0)} + \hat{m}_i^{(0)}) K_h(W_i - W_j)},$$

and

$$\hat{\beta}^{(1)} = \hat{\beta}^{(0)} - \frac{1}{B} \sum_{i=1}^n \dot{\ell}_i(U, \hat{\beta}^{(0)} + \hat{m}_i^{(1)}) \tilde{U}_i^{(0)},$$

where $\hat{m}_j^{(0)}$ and $\hat{\beta}^{(0)}$ are the initial values.

4.2 Measuring the creditability of borrowers in Egyptian business loans

Credit scoring is a method of evaluating the credit risk of loan applicants using historical data and statistical techniques. The method produces a "score" that banks and other financial institutions can use to rank its loan applicants or borrowers in terms of risk. Credit scoring methods are widely used to estimate and to minimize credit risk. The aim of credit scoring is to model or predict the probability that an

applicant with certain covariates is to be considered as a potential risk. (Muller (2001)).

Credit scoring have some obvious benefits that have led to its increasing use in loan evaluation:

- (1) Scoring greatly reduces the time needed in the loan approval process.
- (2) Scoring improves objectivity of the approval process. This objectivity helps lenders ensure they are applying the same underwriting criteria to all borrowers regardless of race, gender, or other factors prohibited by law from being used in credit decisions.

The aim of credit scoring is to model or predict the probability that an applicant with certain covariates is to be considered as a potential risk. For each applicant, the binary response variable "creditability" is available. This will be done using a generalized partial linear model (GPLM) compared with a generalized linear model (GLM). A well-designed model should give a higher percentage of high scores to borrowers whose loans will perform well and a higher percentage of low scores to borrowers whose loans won't perform well.

4.2.1 The Structure of the Data

The credit scoring data set determining the solidness of borrowers is collected from some Egyptian banks (in Alexandria, 2006). The data set consists of 500 observations on business loans and 17 covariates, which are assumed to influence creditability, classified as:

- (1) Non-metric covariates (categorical), which include 14 variables, the variables U1 to U5 are dichotomous, U6 to U10 have three categories, U11 to U13 have five categories, and U14 has seven categories. These variables are described as follows:

U1: the employment of credit for at least one year. (Yes/No)
 U2: existing of available savings. (Yes/No)
 U3: having a credit card. (Yes/No)
 U4: reputation. (Yes: good /No: not good)
 U5: existing judgments between the borrower and his clients. (Yes/No)
 U6: payment of previous credits. (Yes/No/There is no previous credits)
 U7: experience. (Good/Fair/There is no experience)
 U8: business partner relation. (Good/Bad/There are no partners)
 U9: the case of other enterprises.(Good/bad/There are no other enterprises)
 U10: Banker's guarantee type. (Real estate/Effects/others)
 U11: education. (Doctorate/Master/Bachelor/Medium/not educated)
 U12: marital status. (Married/Single/Widowed/Divorced/others)
 U13: project type. (Manufacturing/commercial/professional/agricultural/others)
 U14: purpose of credit. (Project construction/Project developing/Buying a car/Buying a building/Buying a land/Buying goods/others)

- (2) Metric covariates, which include three variables, W1 for age, W2 for the amount of loan in thousand of Egyptian bounds, and W3 for duration in months.

The response variable Y describes the creditability and takes two values, 1 for credit-worthy and 0 for not credit-worthy.

4.2.2 Comparing the Regression Estimates Using a Generalized Linear Model (GLM) and a Generalized partial Linear Model (GPLM):

For each applicant, the binary response variable "creditability" is obtained. This will be done using a generalized partial linear model (GPLM) compared with a generalized linear model (GLM).

The generalized linear model using a logistic cumulative distribution function (logit model) known as a parametric logistic regression model, PLRM, assumes that the probability of a good loan is given by

$$P[Y = 1 | U, W] = G(Z^T \beta) = \frac{\exp(Z^T \beta)}{1 + \exp(Z^T \beta)}, \quad (26)$$

where $Z = (U^T, W^T)^T$, and $G(\cdot)$ is the logistic cumulative distribution function and β is an unknown finite-dimensional parameter that will be estimated.

The semiparametric logistic regression model (SLRM), corresponds to a parametric logistic regression model (PLRM), is considered by generalizing the linear argument ($\beta^T Z$) to a partial linear argument ($\beta^T U + m(W)$), where $Z = (U^T, W^T)^T$, and $m(\cdot)$ is an unknown infinite-dimensional function, a smoothing function. SLRM assumes that the probability of a good loan is given by

$$p(Y = 1 | U, W) = G(U^T \beta + m(W)) = \frac{\exp(U^T \beta + m(W))}{1 + \exp(U^T \beta + m(W))}, \quad (27)$$

where $G(\cdot)$ denotes the logistic cumulative distribution function. The vector U contains all categorical variables and the vector W contains all the metric (continuous) variables which are linearly transformed into $[0, 1]$. The two components of (27), β and $m(\cdot)$, will be estimated.

4.2.3 The estimation results

The following table (Table (1)) shows The estimates of the parametric logistic regression model (PLRM) and the different semiparametric logistic regression models (seven SLRMs) which are obtained by including W_1 , W_2 , W_3 , W_1 and W_2 , W_1 and W_3 , W_2 and W_3 and all metric variables, nonparametrically in the model, respectively.

Table (1): The Parametric coefficient estimates of the parametric logistic regression model (PLRM) and the different semiparametric logistic regression models (SLRMs).

Parametric Variable	PLRM Estimates	SLRM Estimates By including different W_i 's nonparametrically						
		W1	W2	W3	W1, W2	W1, W3	W2, W3	W1, W2, W3
Constant	-4.1146	-	-	-	-	-	-	-
U1	*3.0085	*2.9317	*3.0343	*2.7863	*3.026	*2.6939	*2.8134	*2.7562
U2	0.11595	0.20215	0.22241	0.12575	0.30351	0.20217	0.239	0.33647
U3	*2.5159	*2.4709	*2.4475	*2.319	*2.4262	*2.2787	*2.2814	*2.2384
U4	*3.2842	*3.3563	*3.043	*3.3412	*3.105	*3.413	*3.0924	*3.0587
U5	*-3.5591	*-3.6297	*-3.5201	*-3.4538	*-3.5352	*-3.5419	*-3.4521	*-3.5224
U6	0.1825	0.048588	0.14683	0.20123	0.064899	0.068052	0.16982	0.052472
U7	0.55938	0.48648	0.70907	0.58304	0.67671	0.5253	0.75637	0.79087
U8	0.98888	0.99404	1.0197	0.97588	1.0305	0.9877	1.0013	1.0284
U9	-0.68568	-0.62661	-0.67504	-0.62874	-0.63078	-0.57547	-0.61879	-0.53241
U10	-0.62281	-0.74061	-0.43443	-0.75532	-0.46838	-0.86047	-0.55659	-0.55719
U11	-0.037022	-0.045497	0.17963	-0.072497	0.19833	0.077393	0.13815	0.20597
U12	0.72274	0.37397	0.67117	0.62432	0.45764	0.34042	0.58048	0.44347
U13	0.091592	0.065067	0.10846	0.094149	0.10071	0.065411	0.11457	0.098608
U14	*-0.44423	*-0.42469	*-0.45539	*-0.46177	*-0.45153	*-0.44909	*-0.4949	*-0.5086
W1	-0.032669	-	-0.02239	-0.033526	-	-	0.023525	-
W2	-9.7254e-06	-8.0622e-06	-	-8.1448e-06	-	-6.425e-06	-	-
W3	-0.01297	-0.013302	-0.014311	-	-0.014407	-	-	-
Deviance	132.35	132.98	129.07	129.71	128.28	130.21	126.5	125.37
Iteration No.	6	8	8	7	8	5	10	10

Note: Bold and starred values are significant at 5%.

From this table, it is noted:

- The only significant categorical variables are U1, U3, U4, U5, and U14, according to our data.
- All metric variables, W1, W2, and W3 are insignificant by including them parametrically in PLRM and SLRMs.
- It is noted that SLRMs outperform PLRM in terms of deviance.

4.2.4 Performance assessment of different scores

Different models or different scores can be compared using misclassification rates. (Muller and Ronz (1999)). For the parametric logistic regression model (PLRM), the score for a potential borrower is referred to

$$S = Z^T \hat{\beta} \quad (28)$$

and for the semiparametric logistic regression model (SLRM), the score is

$$S = U^T \hat{\beta} + \hat{m}(W). \quad (29)$$

Misclassification rates can be measured by performance curves such as Lorenz curves and Gini curves. The Lorenz curve is a plot of

$$P(S > t) \text{ versus } P(S > t | Y=0),$$

where t is a certain threshold value, $P(S > t)$ is the probability of observations classified as good, and $P(S > t | Y=0)$ is the probability of observations classified as good conditioned on bad. (Muller and Hardle (2002)).

Practically, the probability $P(S > t)$ is replaced by the relative frequency of good classifications, i.e. for $Y=1$, above a threshold value t . The probability $P(S > t | Y=0)$ is replaced by the relative frequency of faulty classifications, i.e. for $Y=0$, above a threshold value t .

In this work, a series of threshold values is taken and the scores are sorted from bad to good. The performance curves for both PLRM and SLRM are compared when using different combinations of metric variables in the nonparametric part of SLRM.

Figure (1) shows the performance curves of both PLRM and SLRM obtained by including all metric variables, W_1 , W_2 , and W_3 nonparametrically in SLRM. It is noted that the semiparametric model (SLRM) improves the performance with respect to the parametric model (PLRM) since the curve of SLRM is downward w.r.t. the curve of PLRM.

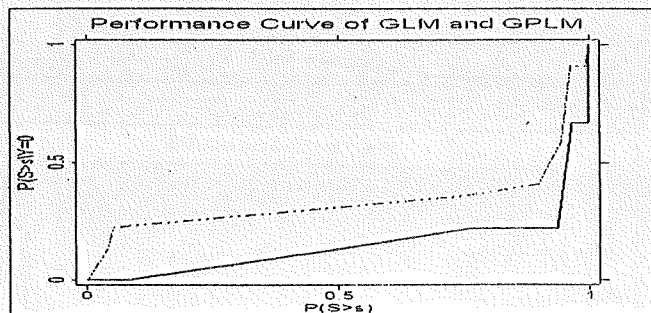


Figure (1): Performance curves for the generalized linear model, GLM (or parametric logistic regression model, PLRM): a thin dotted curve, and the generalized partial linear model, GPLM (or semiparametric logistic regression model, SLRM): a solid curve, with all metric variables included nonparametrically.

Figures (2) and (3) show the performance curves of both PLRM and SLRM obtained by including separately the variables, W_1 , W_2 , and W_3 nonparametrically in SLRM and by including two metric variables nonparametrically in SLRM, i.e. W_1 and W_2 , W_1 and W_3 , and W_2 and W_3 , respectively. These Figures show the influence of every metric variable and every two metric variables in improving the performance w.r.t. PLRM. It is shown that the effect of W_1 and W_2 in improving the performance is greater than the effect of W_3 . This can be interpreted economically that W_1 (age) means reputation, more experience, good previous business, and etc., for W_2 greater amount of a loan means more banker's guarantees, but for W_3 the credit risks for the duration of a loan is limited.

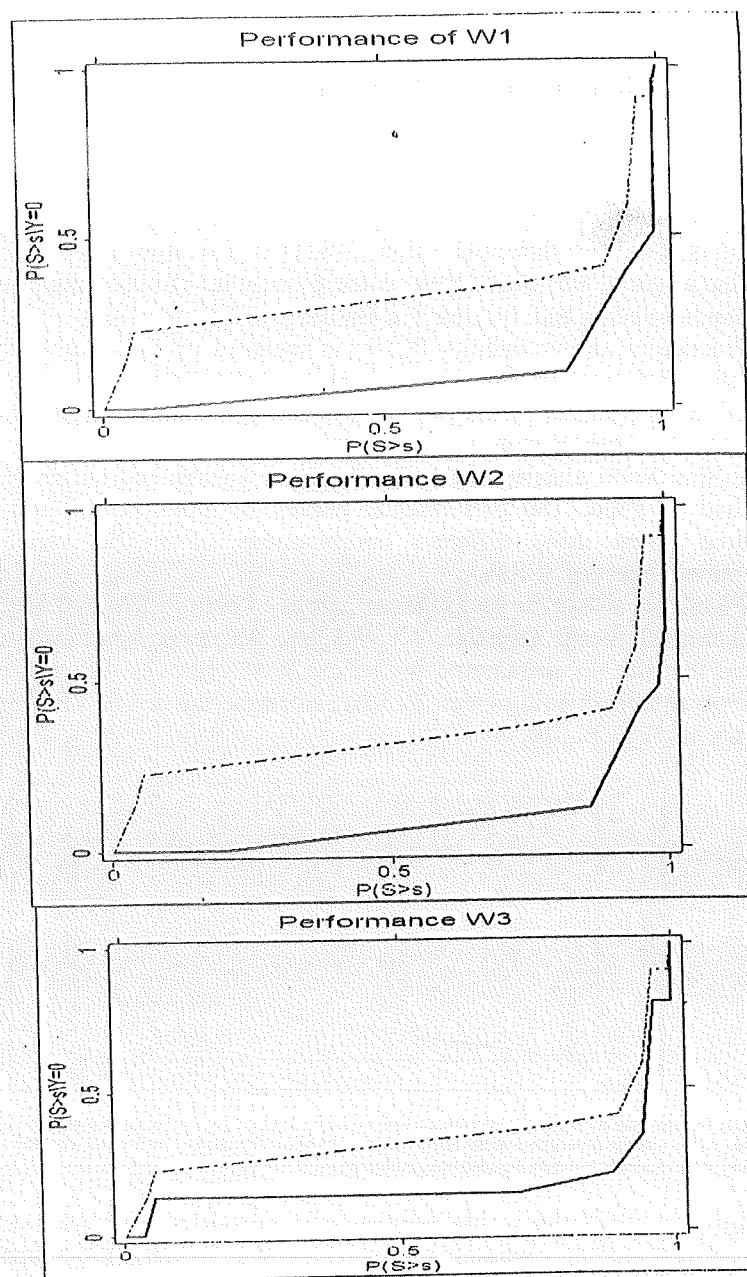


Figure (2): Performance curves for the generalized linear model, GLM (or parametric logistic regression model, PLRM): a thin dotted curve, and the generalized partial linear model, GPLM (or semiparametric logistic regression model, SLRM): a solid curve, by including, separately, the metric variables nonparametrically.

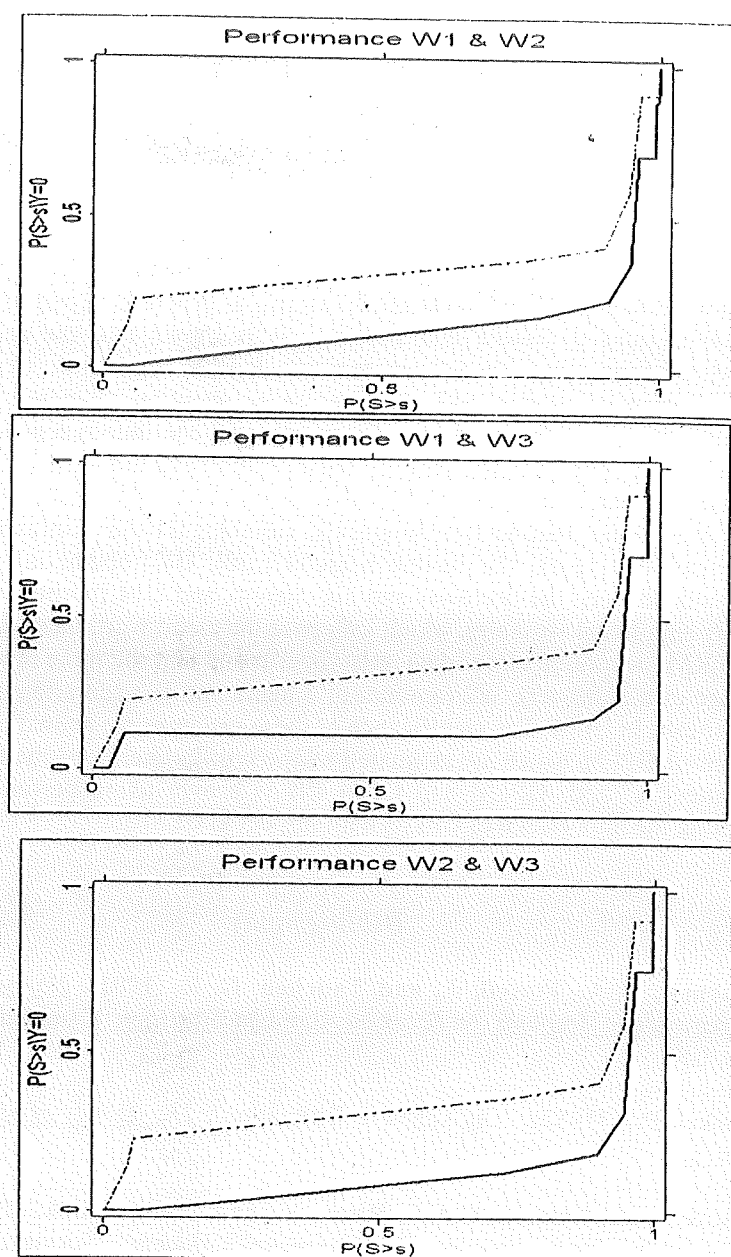


Figure (3): Performance curves for the generalized linear model, GLM (or parametric logistic regression model, PLRM): a thin dotted curve, and the generalized partial linear model, GPLM (or semiparametric logistic regression model, SLRM): a solid curve, by including two metric variables nonparametrically.

5. Summary

In this article, we measured the creditability of borrowers in Egyptian business loans using the generalized partial linear model (GPLM). We developed the iterative procedure for estimating the semiparametric logistic regression model in the case of existing one variable in each component of the model. The profile likelihood method gives estimates which are asymptotically efficient. The

efficiency of these estimates is achieved if the nonparametric part is a "least favorable curve".

Credit scoring represents a set of common techniques to decide whether a bank should grant a loan to an applicant (borrower) or not. The generalized linear model (GLM) and semiparametric generalized linear model or a generalized partial linear model (GPLM) are used and compared. The power of the semiparametric methods is investigated using real data on "business loans". Therefore, semiparametric generalized linear model known as generalized partial linear model (GPLM) is used to handle this situation. It is shown that the semiparametric methods perform better than the parametric logistic regression. Semiparametric logistic regression model (SLRM) gives more attention to metric variables by estimating them nonparametrically. Practically, the continuous regressor variables chosen for the nonparametric part are those which have the most influence on the response variable.

It is found that the five categorical variables U1 (the employment of credit for at least one year), U3 (having a credit card), U4 (reputation), U5 (existing judgments between the borrower and his clients), and U14 (purpose of credit) are the only significant variables, which influence the credit worthiness, in both PLRM and SLRMs according to our data. Also, it is found that the three metric variables, W1 (age of a borrower), W2 (amount of a loan), and W3 (duration of a loan), are insignificant in GLM. The significance of these variables are shown when using all versions of GPLM which are estimated by the profile-likelihood method. Also, the assessment of both GPLM and GLM fits is shown using performance curves indicating the importance of GPLM with respect to GLM.

Bibliography

1. Barnett, W.A., Powell, J., and Tauchen, G. E. (1991). *Nonparametric and Semiparametric Methods in Econometrics and Statistics*. Cambridge University Press, UK.
2. Bickel, P.J. (1982). "On Adaptive Estimation", *The Annals of Statistics*, 10, 3, 647-671.
3. Bickel, P.J., Klassen, C.A.J., Ritov, Y., and Wellner, J.A. (1998). *Efficient and Adaptive Estimation for Semiparametric Models*, Springer-Verlag, New York, Inc.
4. Bickel, P.J., Ritov, Y. (1990). "Achieving Information Bounds in Non and Semiparametric Models", *The Annals of Statistics*, 18, 2, 925-938.
5. Blundell, R., Duncan, A., and Pendakur, K. (1998). "Semiparametric Estimation and consumer demand", *Journal of Applied Econometrics*, Vol. 13, No. 5, 435-461.
6. Burda, M., Hardle, W., Muller, M., and Werwatz, A. (1998). "Semiparametric analysis of German East-West migration: Facts and Theory", *Journal of Applied Econometrics*, Vol. 13, No. 5, 525-542.

7. Chen, H. (1995). "Asymptotically Efficient Estimation in Semiparametric Generalized Linear Models", *The Annals of Statistics*, 23,4, 1102-1129.
8. Cuzick, J. (1992). "Semiparametric Additive Regression", *Journal of the Royal Statistical Society*, 54,3,831-843.
9. Dobson, A.J. (1990). *An Introduction to Generalized linear Models*, Chapman & Hall, New York.
10. Engle, R.F., Granger, C.W.J., Rice, J., and Weiss A. (1986). "Semiparametric Estimates of the Relation Between Weather and Electricity Sales", *Journal of the American Statistical Association*, 81,394,310-320.
11. Gill, R.D. (1989). "Non-and Semi-parametric Maximum Likelihood Estimators and the von Mises Method (Part I)", *Scandinavian Journal of Statistics*, 16, 97-128.
12. Gill, R.D., and Van Der Vaart, A.W. (1993). "Non-and Semi-parametric Maximum Likelihood Estimators and the von Mises Method (Part II)", *Scandinavian Journal of Statistics*, 20, 271-288.
13. Green, P.J. (1984). "Iteratively Reweighted Least Squares for Maximum Likelihood Regression and Some Robust Resistant Alternatives (with Discussion)", *Journal of the Royal Statistical Society, B*, 46, 149-192.
14. Green, P.J., and Silverman, B.W. (1994). *Nonparametric Regression and Generalized Linear Models*, Chapman & Hall, New York.
15. Hand, D.J., and Henley, W.E., (1997). "Statistical Classification methods in Consumer Credit Scoring: a Review", *Journal of the Royal Statistical Society, A*, 160, 523-541.
16. Hardle, W., Klinke, S., and Muller, M. (1995). *XploRe: An interactive Statistical computing Environment*, Springer-Verlag, New York.
17. Hardle, W., Hlavka, Z., and Klinke, S. (2000a). *XploRe Application Guide*, Springer-Verlag, Berlin, Germany.
18. Hardle, W., Klinke, S., and Muler, M. (2000b). *XploRe Learning Guide*, Springer-Verlag, Berlin, Germany.
19. Hardle, W., Muller M., Sperlich, S., and Werwatz, A. (2004). *Nonparametric and Semi-parametric Models (An Introduction)*, Springer, New York.
20. McCullagh, P. (1983). "Quasi-Likelihood Functions", *The Annals of Statistics*, 11, 1, 59-67.
21. McCullagh, P., and Nelder, J.A. (1989). *Generalized Linear Models*, Chapman & Hall, New York.

22. Monahan, J.F. (2001). *Numerical Methods of Statistics*, Cambridge University Press, UK.
23. Muller, H.-G., and Stadtmuller, U. (1987). "Variable Bandwidth Kernel Estimators of Regression Curves", *The Annals of Statistics*, 15, 1, 182-201.
24. Muller, M. and Ronz, B. (1999). "Credit Scoring using semiparametric methods", in *Measuring Risk in complex stochastic systems*, by Frank, J., Hardle, W., and Stahl, G. (2000), Springer Verlage.
25. Muller, M., (2001). "Estimation and Testing in Generalized Partial Linear Models—A comparative Study", *Statistics and Computing*, 11, 299-309.
26. Muller, M., (2004). "Generalized Linear Models", *Handbook of Computational Statistics*, Vol. 1.
27. Murphy, S.A., and Van Der Vaart, A.W. (2000). "On Profile Likelihood", *Journal of the American Statistical Association*, 95, 449-485.
28. Nadaraya, E.A. (1964). "On Estimating Regression", *Theory of Probability and its Applications*, Vol. 10, 186-190.
29. Nelder, J.A., and Pregibon, D. (1987). "An Extended quasi-likelihood Functions", *Biometrika*, 74, 221-232.
30. Nelder, J.A., and Wedderburn, R.W.M. (1972). "Generalized Linear Models", *Journal of the Royal Statistical Society, A*, 135, 370-384.
31. Pagan, A., and Ullah, A. (1999). *Nonparametric Econometrics*, Cambridge University Press, UK.
32. Peng, H. (2003). "Efficient Estimation in Semiparametric Generalized Linear Models", to appear in proceedings of JSM2003.
33. Peng, H. (2004). "Efficient Inference in Semiparametric Generalized Linear Models", <http://WWW.home.olemiss.edu/~mmpeng/preprints>.
34. Peng, H., and Wang, X. (2004). "Moment Estimation in Semiparametric Generalized Linear Models", *Nonparametric Statistics*, 1-22.
35. Powell, J., L. (1994). "Estimation of Semiparametric Models", *Handbook of Econometric Vol. 4, Ch. 41*.
36. Robinson, P.M. (1988). "Root-N-Consistent Semiparametric Regression", *Econometric*, 56, 4, 931-954.
37. Ruppert, D., Wand, M.P. (1994). "Multivariate Locally Weighted Least Squares Regression", *The Annals of Statistics*, Vol.12, No.2, 1346-1370.

38. Ruppert, D., Wand, M.P., and Carroll, R.J. (2003). Semiparametric Regression, Cambridge University Press.
39. Severini, T.A., and Staniswalis, J.G. (1994). "Quasi-likelihood Estimation in Semiparametric Models", Journal of the American Statistical Association, 89, 426, 501-511.
40. Severini, T.A., and Wong, W.H., (1992). "Profile Likelihood and Conditionally Parametric Models", The Annals of Statistics, 20, 4, 1768-1802.
41. Signorini, D.F., and Jones, M.C. (2004). "Kernel Estimators for Univariate Binary Regression", Journal of the American Statistical Association, 99, 465, 199-126.
42. Speckman, P. (1988). "Kernel Smoothing in Partial Linear Models", Journal of the Royal Statistical Society, B, Vol. 50, No.3, 413-436.
43. Sperlich, S., Hardle, W., and Aydinli, G. (2006). The Art of Semiparametrics, Physica-Verlag, Germany.
44. Staniswalis, J.G. (1989). "The Kernel Estimate of a Regression Function in Likelihood-Based Models", Journal of the American Statistical Association, 84, 405, 276-283.
45. Van Der Vaart, A.W. (1988). "Estimating a Real Parameter in a Class of Semiparametric Models", The Annals of Statistics, 16, 4, 1450-1474.
46. Van Der Vaart, A.W. (2000). Asymptotic Statistics, Cambridge University Press, UK.
47. Wand, M.P., and Jones, M.C. (1995). Kernel Smoothing, Chapman & Hall, London.
48. Watson, G.S. (1964). "Smooth Regression Analysis", Sankhya, Series A, Vol.26, 359-372.
49. Wedderburn, R.W.M. (1974). "Quasi-Likelihood Functions, Generalized Linear Models, and the Gauss-Newton method", Biometrika, 61, 439-447.
50. Wedderburn, R.W.M. (1976). "On the Existence and uniqueness of the maximum likelihood Estimates for Certain Generalized Linear Models", Biometrika, 63, 27-32.
51. Wong, W.H., and Severini, T.A. (1991). "On Maximum Likelihood Estimation in Infinite Dimensional Parameter Spaces", The Annals of Statistics, 19, 2, 603-632.
52. Yatchew, A. (2003). Semiparametric Regression for the Applied Econometrician, Cambridge University Press, UK.

Estimation of the Cardiac Time-Varying Parameters During the Ejection Phase of the Cardiac Cycle Using the Ito Calculus

By

AHMED S. ABUTALEB* and J. MELBIN**

* Cairo University, School of Engineering, Systems and Bioengineering Depts., Giza, Egypt. E-Mail: ataleb@mcit.gov.eg

** Professor Emeritus, University of Pennsylvania, Philadelphia, Pennsylvania, PA 19083, USA.

Abstract: With the assumption that noisy measurements are possible for the Pulmonary and Aortic flows and pressures, we seek to estimate the time-varying parameters (Compliance, Resistance, and Inertance) as descriptors of the right and left ventricles. A polynomial in time model is assumed for each parameter. The coefficients of the polynomials are estimated from the observed data and the maximum likelihood method and the stochastic calculus are utilized. Some but not all the results agree with results previously reported in the literature.

Key words: Aortic Pressure, Pulmonary Pressure, Left Ventricle, right Ventricle, Cardiac Dynamics, Time-Varying Parameters, Ito Calculus

I. Introduction:

The ejection of blood from the heart is influenced by its filling and the receiving arterial system (load), aspects recognized since [Roy; 1870]. Furthermore, the contractile properties of the ventricle are sensitive to neural and endocrine control and heart rate adds other fluid-mechanical considerations to the pumping process. Consequently, hemodynamic variables, such as ventricular pressure and flow, reflect complex interactions of these contributors [Abutaleb et. al; 1984, Abutaleb et. al; 1986].

Over the years, several models for the left ventricle were introduced. Two compliances were utilized by [Warner; 1959] to model a simple ventricle: one compliance for diastole and a lesser value for systole. Expressions for left ventricle compliance also included the instantaneous ratio between ventricular pressure and volume, the ratio between alterations of pressure and corresponding volume within small time intervals, and the slope of the tangent to the ventricular pressure-volume loop. These latter methods have yielded positive, negative, or infinite compliance values. Some popular models that derive the aortic pressure are reviewed.

(1) The elastance-resistance model [Shroff and Motz; 1989]:

In this model the relationship between the aortic pressure, $P(t)$, and left ventricle volume, $Q(t)$, is described by the nonlinear differential equation:

$$P(t) = a(t)Q(t) + b(t)\frac{dQ}{dt} + c(t)Q(t)\frac{dQ}{dt} + g(t)$$

Where $a(t)$ represents elastance, $b(t)$ represents resistance, $c(t)$ and $g(t)$ represent known quantities derived from the pressure and the flow.

(2) This model introduces nonlinearity explicitly by the term " $Q(t)\frac{dQ}{dt}$ ". An inertial component is not represented. A nonlinear elastance model [Suga and Sagawa; 1988]:

$$P(t) = E(t)[Q(t) - Q(0)]f\left[Q(t), \frac{dQ}{dt}\right]$$

Where $E(t)$ is the elastance, $Q(0)$ is the ventricular volume at zero pressure, and " $f\left[Q(t), \frac{dQ}{dt}\right]$ " is a corrective function.

(3) A different nonlinear elastance model [Zhong et. al; 2005] establishes elastance as a combination of time-varying passive and time-varying and volume-dependent active elastance. .

Time-varying heart (myocardial) properties were also estimated using mechanical models [Karam and Abche; 2006]. These studies demonstrated that both a time-variant compliance and a time-variant resistance are manifested during the cardiac cycle.

Approximate, closed form relations between left ventricular time-varying resistance; compliance, and pressure and volume were derived using optimality principle [Abutaleb et. al; 1981]. That work yielded a negative, "seemingly" physically unrealizable value of inertance during a portion of ejection, suggesting inadequacy of the physical representation of the ventricle, the optimization procedure, or the constraints. A different physical meaning for the inertance, however, was able to explain these negative values [Abutaleb; 1984].

Inertial properties as well as losses of both the blood and ventricle are often ignored on the assumption that at particular sites or conditions either may represent only small

fractions of total energy. Although kinetic energy may constitute a small part of myocardium mechanical energy, flow and its derivative may achieve appreciable values particularly during systole where blood acceleration at the onset of ejection is appreciable. Consequently, a comprehensive study of myocardial behavior during ejection suggests inclusion of all time-varying parameters, Elastance, Resistance, and Inertance [Abutaleb et. al; 1981, Abutaleb et. al; 1982]. For a noise-free environment, these may be related as [Noordergraaf and Melbin; 1982]:

$$L(t) \frac{d^2 Q(t)}{dt^2} + \frac{Q(t)}{C(t)} + R(t) \frac{dQ(t)}{dt} = P(t) \quad (I. 1)$$

Or more adequately, with the inclusion of the derivative of the time-varying elastance " $\frac{dL(t)}{dt}$ ", as:

$$L(t) \frac{d^2 Q(t)}{dt^2} + \frac{Q(t)}{C(t)} + \left[\frac{dL(t)}{dt} + R(t) \right] \frac{dQ(t)}{dt} = P(t) \quad (I. 2)$$

Where the unknown $L(t)$, $C(t)$, and $R(t)$ represent the inertia of blood, the compliance of left/right ventricular wall, and the losses during ejection respectively. $Q(t)$ is the blood volume in the left/right ventricle and $P(t)$ is the pressure at the root of the Aorta or the Pulmonary pressure for the right ventricle.

In [Abutaleb; 1986], the pump equation (I.1) was modified by relating inertance to an equivalent area in a force balance relation. By modeling the unknown time-varying cardiac parameters as polynomials in time, uniqueness of their values could be established. We utilize methods of the Ito calculus to estimate the coefficients of the polynomials that represent the time-varying cardiac parameters $L(t)$, $C(t)$, and $R(t)$ during the ejection phase of the cardiac cycle. The maximum likelihood method is utilized to accommodate the system whether or not noise exists

in the measurements of flow and pressure. In the estimation of the time-varying parameters, constraints are added to ensure nonnegative values.

Estimation Methods of Time-Varying Parameters:

The problem of the estimation of time-varying parameters has, in general, four methods for solution:

- (1) Assuming that the system coefficients are varying sufficiently slowly, they can be tracked using the localized (weighted or windowed) versions of the least squares or maximum likelihood estimators [Niedzwiecki; 1984, 1990, 2000].
- (2) Approximation of the time-varying coefficients by a weighted combination of a certain number of known functions (basis functions). If the unknown weights are assumed to be constants, well known identification techniques could be used [Grenier; 1983, Van Trees; 1968].
- (3) Assumption that the time-varying coefficients evolve as Markovian. In this case, the Kalman filter technique and its modifications could be used for the estimation of the time-varying parameters [Chow; 1987, Abutaleb and Papaioannou; 2000].
- (4) The time-varying coefficients could be treated as unknown controllers to be estimated to track the observed data. The method of Pontryagin maximum principle could be used to find the desired values [Abutaleb; 1986, Chen et al; 1998].

And most recently the Malliavin calculus was used to estimate time-varying parameters in other applications [Abutaleb and Papaioannou; 2007, 2008].

In this report a modification of the second method is introduced. A time dependent polynomial is assumed for each unknown time-varying parameter. The coefficients of the polynomials are estimated using the maximum likelihood method of the Ito calculus. This is because the system dynamics and the measurements are noisy.

In section II, the system dynamics or the pump equation is presented as a linear model of the left/right ventricle during ejection of the blood. In section III, we introduce the proposed method that is based on the stochastic calculus (Ito calculus). In section IV, the simulated and the real data is presented. Section V presents results, summary and conclusions.

II. The Left/Right Ventricle as a Linear Time-Varying Model:

It is assumed that a linear time-varying model describes the left/right ventricle during the ejection phase of the cardiac cycle. The model is given as:

$$L(t) \frac{d^2 Q(t)}{dt^2} + \frac{Q(t)}{C(t)} + \left[\frac{dL(t)}{dt} + R(t) \right] \frac{dQ(t)}{dt} = P(t) \quad (\text{II. 1})$$

Where the unknown $L(t)$, $C(t)$, and $R(t)$ represent the inertia of blood, the compliance of the ventricular wall, and the losses during ejection respectively. $Q(t)$ is the blood volume in the ventricle and $P(t)$ is the pressure at the root of the Aorta or the pulmonary pressure for the right ventricle. It is desired to estimate these unknown time-varying parameters.

III. Estimation of the Time-Varying Parameters:

In this section the proposed Stochastic-Calculus-based-method is explained. As before, the differential equation that governs the blood ejection from the left/right ventricle (the system dynamics) is taken as:

$$L(t) \frac{d^2 Q(t)}{dt^2} + \frac{Q(t)}{C(t)} + \left[\frac{dL(t)}{dt} + R(t) \right] \frac{dQ(t)}{dt} = P(t)$$

A stochastic differential equation (Ito Calculus) is first established for the evolution of $Q(t)$ over time to accommodate noise in $Q(t)$ and $P(t)$. Since these variables are treated as stochastic processes, deterministic calculus would be inadequate.

$$\text{Let } \underline{X}(t) = \begin{bmatrix} X_1(t) \\ X_2(t) \end{bmatrix} = \begin{bmatrix} Q(t) \\ \frac{dQ(t)}{dt} \end{bmatrix} \quad (\text{III. 1})$$

$$\begin{aligned} \text{Since } \frac{d^2 Q(t)}{dt^2} &= \frac{P(t) - \left\{ \frac{Q(t)}{C(t)} + \left[\frac{dL(t)}{dt} + R(t) \right] \frac{dQ(t)}{dt} \right\}}{L(t)} \\ &= - \frac{\left[\frac{dL(t)}{dt} + R(t) \right]}{L(t)} \frac{dQ(t)}{dt} - \frac{1}{C(t)L(t)} Q(t) + \frac{1}{L(t)} P(t) \end{aligned} \quad (\text{III. 2})$$

We have the deterministic system of equations:

$$\frac{dX_1(t)}{dt} = X_2(t) \quad (\text{III. 3a})$$

$$\frac{dX_2(t)}{dt} = - \frac{\left[\frac{dL(t)}{dt} + R(t) \right]}{L(t)} X_2(t) - \frac{1}{L(t)C(t)} X_1(t) + \frac{1}{L(t)} P(t) \quad (\text{III. 3b})$$

The stochastic differential equation that describes blood ejection is derived as:

$$\begin{bmatrix} dX_1(t) \\ dX_2(t) \end{bmatrix} = \left\{ \begin{bmatrix} 0 & 1 \\ -\frac{1}{L(t)C(t)} & -\frac{\left(\frac{dL(t)}{dt} + R(t) \right)}{L(t)} \end{bmatrix} \begin{bmatrix} X_1(t) \\ X_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{P(t)}{L(t)} \end{bmatrix} \right\} dt + \begin{bmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{bmatrix} \begin{bmatrix} dW_1(t) \\ dW_2(t) \end{bmatrix} \quad (\text{III. 4})$$

where $W_1(t), W_2(t)$ are Wiener processes, σ_{11} is a constant that represents the noise

level in the measurements of $\frac{dQ(t)}{dt}$ and σ_{22} is a constant that represents the noise

level in the measurements of $\frac{d^2 Q(t)}{dt^2}$.

III. 1, Maximum Likelihood Estimates of the Cardiac Parameters:

Each parameter module is assumed to have a polynomial expansion in time.

Specifically;

$$\left(\frac{dL(t)}{dt} + R(t) \right) = \sum_{i=0}^{I-1} R_i t^i \quad (\text{III. 5a})$$

$$L(t) = \sum_{j=0}^{J-1} L_j t^j \quad (\text{III. 5b})$$

$$\text{And } E(t) = \sum_{k=0}^{K-1} E_k t^k \quad (\text{III. 5c})$$

We observe $\underline{X}(t) = \begin{bmatrix} X_1(t) \\ X_2(t) \end{bmatrix} = \begin{bmatrix} Q(t) \\ \frac{dQ(t)}{dt} \end{bmatrix}$, and the pressure $P(t)$.

With Equation (III. 1) the stochastic differential equation (SDE) has the form:

$$d\underline{X}(t) = \underline{b}(\underline{X}(t); \underline{\theta}) dt + \Sigma(t, \underline{X}(t)) d\underline{W}(t) \quad (\text{III. 6})$$

$$\text{Where } \underline{b}(\underline{X}(t); \underline{\theta}) = \left\{ \begin{bmatrix} 0 \\ -\frac{E(t)}{L(t)} \end{bmatrix} - \frac{\begin{bmatrix} 1 \\ \left(\frac{dL(t)}{dt} + R(t) \right) \end{bmatrix}}{L(t)} \begin{bmatrix} X_1(t) \\ X_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{P(t)}{L(t)} \end{bmatrix} \right\} \quad (\text{III. 7})$$

$$\text{of the format: } \underline{b}(\underline{X}(t); \underline{\theta}) = A(t; \underline{\theta}) \underline{X}(t) + \underline{d}(t; \underline{\theta}) \quad (\text{III. 8})$$

$$\text{and } A(t; \underline{\theta}) = \begin{bmatrix} 0 & 1 \\ -\frac{\sum_{k=0}^{K-1} E_k t^k}{\sum_{j=0}^{J-1} L_j t^j} & -\frac{\sum_{i=0}^{I-1} R_i t^i}{\sum_{j=0}^{J-1} L_j t^j} \end{bmatrix}, \quad \underline{d}(t; \underline{\theta}) = P(t) \begin{bmatrix} 0 \\ 1 \\ \frac{1}{\sum_{j=0}^{J-1} L_j t^j} \end{bmatrix}$$

$$\underline{\theta} = [R_0 \quad \dots \quad R_{I-1} \quad L_0 \quad \dots \quad L_{J-1} \quad E_0 \quad \dots \quad E_{K-1}]^T,$$

$$\text{and } \Sigma(t, \underline{X}(t)) = \begin{bmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{bmatrix}. \quad (\text{III. 9})$$

Notice that both A and $\underline{d}$ are nonlinear in the vector of unknown parameters $\underline{\theta}$.

According to Girsanov theory, if $\Sigma(t, \underline{X}(t)) = \begin{bmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{bmatrix}$ is completely known, the

Radon-Nikodym derivative which is the likelihood function, $\mathcal{L}(t, \underline{X}(t); \underline{\theta})$, is given by

[Abutaleb and Papaioannou; 2008, and Oksendal; 1998]:

$$\begin{aligned} \mathcal{L}(t, \underline{X}(t); \underline{\theta}) = \exp \left\{ -\frac{1}{2} \int_0^t \underline{b}^T(s, \underline{X}(s); \underline{\theta}) [\Sigma^{-1}(s, \underline{X}(s))]^T [\Sigma^{-1}(s, \underline{X}(s))] \underline{b}(s, \underline{X}(s); \underline{\theta}) ds \right. \\ \left. + \int_0^t \underline{b}^T(s, \underline{X}(s); \underline{\theta}) [\Sigma^{-1}(s, \underline{X}(s))]^T [\Sigma^{-1}(s, \underline{X}(s))] d\underline{X}(s) \right\} \end{aligned} \quad (III. 10)$$

The log likelihood function $\lambda = \ln \mathcal{L}$ becomes:

$$\begin{aligned} \lambda(t, \underline{X}(t); \underline{\theta}) = -\frac{1}{2} \int_0^t \underline{b}^T(s, \underline{X}(s); \underline{\theta}) [\Sigma^{-1}(s, \underline{X}(s))]^T [\Sigma^{-1}(s, \underline{X}(s))] \underline{b}(s, \underline{X}(s); \underline{\theta}) ds \\ + \int_0^t \underline{b}^T(s, \underline{X}(s); \underline{\theta}) [\Sigma^{-1}(s, \underline{X}(s))]^T [\Sigma^{-1}(s, \underline{X}(s))] d\underline{X}(s) \end{aligned} \quad (III. 11)$$

The maximum likelihood estimate for the unknown parameters $\underline{\theta}$ is obtained by

maximizing equation (III. 11). This is true as long as $\Sigma(t, \underline{X}(t)) = \begin{bmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{bmatrix}$ is

completely known. In reality, we estimate $\Sigma(t, \underline{X}(t)) = \begin{bmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{bmatrix}$ first and use its

value in the log likelihood function to find the estimates of $\underline{\theta}$.

Since $\lambda(t, \underline{X}(t); \underline{\theta})$ is nonlinear in $\underline{\theta}$, the maximization of the log likelihood with respect to

$\underline{\theta}$ is difficult. The stochastic annealing method is used for this task. Without constraints

placed on the shape of $R(t)$, $L(t)$, and $E(t)$, certain of these estimated functions could attain

negative values at some portions of the ejection phase of the cardiac cycle. Constraints ensure

that this does not occur with the addition of penalty terms that preclude parameters from

becoming negative.

IV. Simulation and Real Data:

To test the approach, simulated pressure and flow for the Canine heart was first used

followed by real data [Abutaleb; 1984]. All data, simulated and real, were used to

determine the unknown parameters elastance, inertance, and resistance.

IV. 1, Simulation:

Ventricular volume (in cubic centimeter) and Aortic/Pulmonary pressure (in millimeter Hg) was each modeled as a polynomial in time. They are as follows;

Left ventricle flow: define S as the maximum value of the flow= 320 cc/sec, $t_{\max}$ as the instant when the maximum flow occurs= 0.2/3 sec, and T=0.2 sec as the total duration of the ejection, then:

$$\text{For the first third of the ejection phase: } \frac{dQ(t)}{dt} = \frac{S}{t_{\max}} t \quad (\text{IV. 1a})$$

$$\text{For the rest of the ejection phase: } \frac{dQ(t)}{dt} = S - \frac{S}{T - t_{\max}} (t - t_{\max}) \quad (\text{IV. 1b})$$

Q(t) is the integration of the flow with initial conditions, end diastolic volume, set to 50 cc.

Aortic Pressure:

$$P(t) = \beta_0 + \beta_1 t + \beta_2 t^2, \beta_0 = 98, \beta_1 = 200, \beta_2 = -1000 \quad (\text{IV. 2})$$

The right ventricle volume: S=120 cc/sec and the same equations as the left ventricle.

Pulmonary pressure:

$$\beta_0 = 10, \beta_1 = 240, \beta_2 = -1200 \quad (\text{IV. 3})$$

The observation period is 0.2 seconds, and the sampling interval is 0.004 seconds yielding 50 samples. The simulated data are shown in Figure 2 for the left ventricle and in Figure 5 for the right ventricle. For all the parameters, and for simulated and real data, the polynomial order was gradually increased till we obtained the minimum error between the estimated pressure and the observed/simulated volume. It was found that I=J=K=4 yield good estimates. Changing the order between 4 and 5 had a minor effect on the results. A possible improved criterion for order selection is currently under investigation.

IV. 2. Real Data:

Data were obtained from mongrel dogs, weighing between 14.5 and 25.0 kg. Anesthesia was induced with Na thiopental (0.5 ml kg. And, after tracheal intubation, maintained with an oxygen-halothane (0.5-2.5 percent) mixture [Abutaleb; 1984].

Ten animals were instrumented, chronically, on the systemic side. For this group, an electromagnetic flow probe was placed around the root of the aorta and a radiopaque no. 5 French catheter was introduced via the internal thoracic artery such that the tip was located at the center of the flow prob. A statham (SP-37) transducer provided manometric data, flat to 25 Hz. And electromagnetic flow cuff (In-Vivo-Metric) was used, with Biotronex Lab flow meters set at 50 Hz cutoff frequency for flow measurements. In addition, a bipolar electrode was sutured to the right atrial appendage.

Six dogs of this group were also bolus infused intravenously with 0.5 cc 20 μ g/ml norepinephrine (Levophed bitrate. Wintrop Labs). Data were recorded for 2-3 min after the new stable heart rate (persisting for about 3-4 min) was reached.

For the awake group (four of the ten dogs), all leads and the catheter were routed subcutaneously and exteriorized through skin connectors caudal to the scapulae. Data were recorded with the dogs under light anesthesia and, over a period of days, with the animals standing and lightly restrained by a support sling. Depending on the animal condition, about 2-4 days elapsed after surgery before data were acquired from the awake animal.

Another group of six dogs was instrumented on the pulmonic side. Here, the flow robe was placed around the main pulmonary trunk and a Millar catheter tip transducer was routed via the right ventricular outflow tract to locate the tip at the center of the

flow probe. A pacing electrode was also secured to the right atrial appendage. For this group, data were obtained only during the anesthetized state.

Except for the awake group, heart rates were controlled by external pacing. Steady-state data, obtained with pacing frequencies of 100, 140, and 180/min, were digitized at 333 Hz.

Flow from the left/right ventricle and pressure at the root of the Aorta and Pulmonary trunk were obtained [Abutaleb; 1984]. These data were used with the proposed algorithm to define the properties of the Cardiac muscle as explained above (Section III).

V. RESULTS AND DISCUSSION

Cardiac parameters were estimated for both ventricles under the assumption that they are nonnegative.

V. 1. The Left Ventricle:

Figure 2 presents the raw data, $Q(t)$ and $P(t)$, for the left ventricle and the estimated parameters are shown in Figure 3. The values of $R(t)$ and $E(t) = 1/C(t)$ are similar to that reported in [Abutaleb; 1984, Abutaleb et. al; 1986]. The contribution of each parameter to the observed Aortic pressure is shown in Figure 4. The shape of the contribution and percentage agree with the results reported in [Abutaleb; 1986]. The following conclusions are drawn:

- (1) At the onset of ejection, the major manifestation of energy utilization resides in the cardiac compliance that ejects blood from the ventricle, overcoming the inertia and resistance of the blood and large proximal vessels.

As momentum develops inertial forces also move blood and the role of the compliance is reduced. We expect the Compliance to be high at the onset of systole, reducing as ejection progresses. Conversely $L(t) \frac{d^2 Q(t)}{dt}$ opposes outflow at the onset of ejection and aids outflow towards end. of ejection. Inertial forces contribute about 20% of the Aortic pressure in the last two thirds of the ejection phase. Thus, this quantity changes sign from negative (opposing) at the onset of ejection to positive (aiding) towards the end of ejection. Resistance consistently opposes ejection and manifests as losses (heat).

(2) The Left Ventricle elastance, $E(t)$, as estimated for the ejection period agrees, in order of magnitude and in its general time course, with elastance derived by other methods, specifically, with instantaneous $P(t)/Q(t)$ [Sagawa et. al; 1988]. The elastance, however, is approximately 10% greater than that of $P(t)/Q(t)$ due to the influence of inertance, $L(t)$, on the calculations.

(3) The resistive parameter, " $\frac{dL(t)}{dt} + R(t)$ ", during ejection agrees with the lower values reported elsewhere [Hunter; 1982], [Suga et. al; 1980, Schroff; 1981, Ringo et. al; 1982] but is not in agreement with the higher values that reported in [Vaartjes; 1983].

(4) The inertial parameter, $L(t)$, cannot be ignored in any analysis. The surge of blood from the left ventricle develops a large inertial impact that influences the Aortic pressure.

V. 2, The Right Ventricle:

Figure 5 presents the raw data for the right ventricular volume, $Q(t)$, and the pulmonary pressure, $P(t)$. The estimated parameters (all constrained to be

nonnegative) are shown in Figure 6. The values of $R(t)$ and $E(t) = 1/C(t)$ are similar to that reported in [Abutaleb; 1984, Abutaleb et. al; 1986]. The contribution of each parameter to the observed Pulmonary pressure is shown in Figure 7. The shape of the contribution and percentage agree with the results reported in [Abutaleb; 1986]. The following conclusions are drawn:

(1) As with the left ventricle at the onset of ejection, the major manifestation of energy utilization resides in the cardiac compliance that ejects blood from the ventricle, overcoming the inertia and resistance of the blood and large proximal vessels.

As momentum develops, inertial forces dominate movement of blood and the role of the compliance becomes virtually nonexistent. This differs from what is observed in the left ventricle where inertia is important but not dominating. $L(t) \frac{d^2 Q(t)}{dt^2}$ opposes outflow of blood at commencement of ejection and aids the outflow of blood towards the end of ejection. Inertial forces contribute about 80% of the pulmonary pressure during the last two thirds of the ejection phase. Thus, this quantity changes sign from negative (opposing) at the onset of ejection to positive (aiding) towards the end of ejection. The resistance (losses) opposes ejection.

(2) The right ventricular elastance, $E(t)$, as estimated for the ejection period agrees, in order of magnitude and its general time course, with elastance derived by other methods, specifically, with instantaneous $P(t)/Q(t)$ for the first one third of the ejection phase. Elastance, during the last two thirds, is about 50% lower than that of $P(t)/Q(t)$ due to the influence of inertance, $L(t)$, on the calculations.

(3) The resistive parameter, " $\frac{dL(t)}{dt} + R(t)$ ", during ejection agrees with the lower values reported elsewhere [Abutaleb; 1986].

(4) The inertial parameter, $L(t)$, cannot be ignored in any analysis. The surge of blood from the right ventricle develops a large inertial impact that influences pulmonary pressure particularly during the last two thirds of the ejection phase.

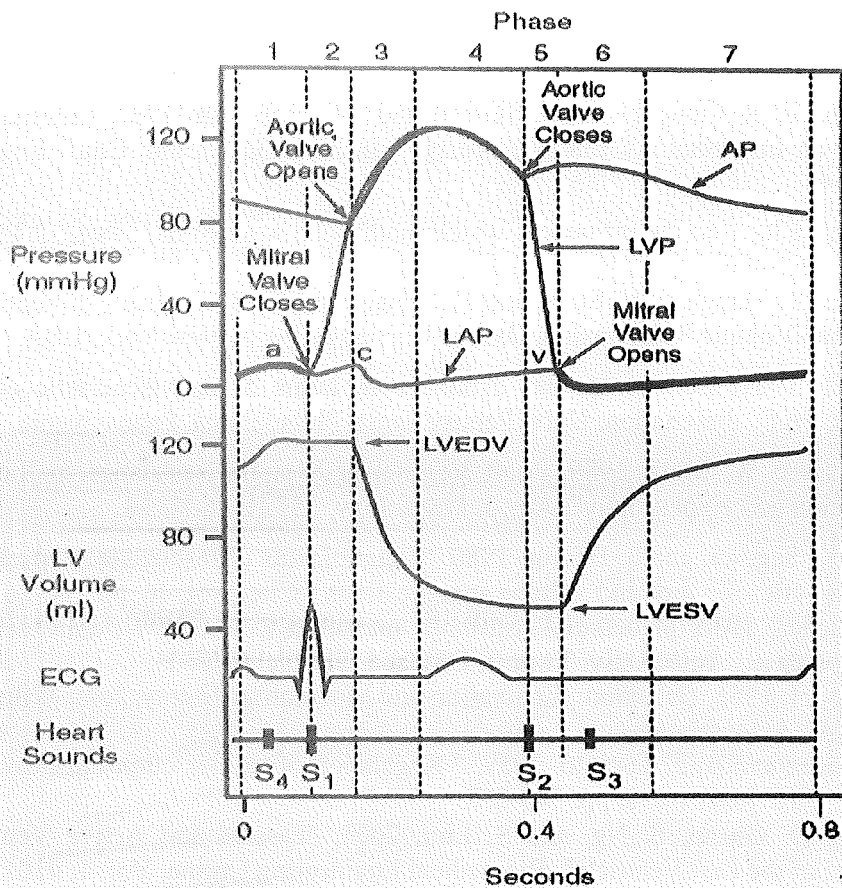
References:

- A. Abutaleb, J. Melbin, and A. Noordergraaf, 1981, "Continuous energy conversion during ventricular ejection, " in Proc. 34th Annu. Conf. Eng. Med. Biol., Houston, Texas.
- A. Abutaleb, J. Melbin, and A. Noordergraaf, 1982, "Sensitivity of unconstrained cardiac time-varying parameters to errors in measured afterload and preload, "in Proc. 10th Annu. NE Bioeng. Conf., pp. 283-288.
- A. Abutaleb, 1984, Identification of the ventricular pump parameters during the ejection phase of the cardiac cycle, Ph.D. dissertation, University of Pennsylvania, Philadelphia, Pennsylvania, USA.
- A. Abutaleb, J. Melbin, And A. Noordergraaf, 1986, "Identification Of The Cardiac Parameters During The Ejection Phase", IEEE Trans. Biomedical Engineering, March.
- A. Abutaleb, 1986, "Target Image Tracking Using a Nonlinear Filter Based on Pontryagin Minimum Principle", IEEE Trans. Automatic Control, Vol. AC-31, #12.
- A. Abutaleb, and M. Papaioannou, 2000, "Maximum Likelihood Estimation of Time-Varying Parameters: An Application to the Athens Exchange Index", Applied Economics, Vol. 32, #10, pp. 1323-1329.
- A. Abutaleb, and M. Papaioannou, 2007, "Malliavin Calculus for the Estimation of Time-Varying Regression Models Used in Financial Applications", Int. J. Theoretical and Applied Finance.
- A. Abutaleb, and M. Papaioannou, 2008, Introduction to the Stochastic and the the Malliavin Calculus with Applications to Engineering and Finance, IDEA press, in press.
- Y. Chen, C. Wen, J. Xu, and M. Sun, 1998, "High Order Iterative Learning Identification of Projectile's Aerodynamic Drag Coefficient Curve from Radar Measured Velocity Data", IEEE Trans. Control Systems Technology, Vol. 6, #4, pp. 563-570.
- G. Chow, 1987, Econometrics, John Wiley, New York.
- Y. Grenier, 1983, " Time-Dependent ARMA Modeling of Nonstationary Signals", IEEE Trans. ASSP, Vol. 31, p. 899-911.
- W. C. Hunter, 1982, A new approach to ventricular dynamics: The flowpulse response, Ph.D. dissertation, Univ. Pennsylvania, Philadelphia, Pennsylvania, USA.
- E. Karam and A. Abche, 2006, "Left Ventricular Model to Study the Combined Viscoelastic, Heart Rate, and Size Effects", Int. J. Biomedical Sciences, Vol. 1, #2, pp. 75-84.
- P. Kloeden, and E. Platen, 1992, Numerical Solution of Stochastic Differential Equations, Springer-Verlag, New York.
- Y. Kutoyants, 2003, Statistical Inference for Ergodic Diffusion Processes, Springer, USA.
- M. Niedzwiecki, 2000, Identification of Time-Varying Processes, John Wiley, New York, USA.
- M. Niedzwiecki, 1990, "Identification of Time-Varying Systems Using Combined Parameter Estimation and Filtering", IEEE Trans. ASSP, Vol. 38, #4, pp. 679-686.
- M. Niedzwiecki, 1984, "On the Localized Estimators and Generalized Akaike's

- Criterion", IEEE Trans. Automatic Control, Vol. AC-29, pp. 970-983.
- A. Noordergraaf and J. Melbin, 1982, "Introducing the pump equation," in Cardiovascular System Dynamics, T. Kenner, R. Busse, and H. Hinghofer-Szalkay, Eds. New York: Plenum., pp. 19-35.
- B. Oksendal, 1998, Stochastic Differential Equations: An Introduction with Applications, Universitext, NY, USA.
- J. A. Ringo, K. R. Campbell, B. K. Slinker, and J. D. Robinette, 1982, "Internal series resistance-an important LV pump property. " in Proc. 35th Annu. Conf. Eng. Med. Biol., p. 64.
- C. S. Roy, 1870, "on the influences which modify the work of the heart." J. Physiol., vol. 1. Pp. 452-496.
- K. Sagawa, L. Maughan, H. Suga, and K. Sunagawa., 1988, Cardiac Contraction and the Pressure-Volume Relationship, Oxford University Press, New York, USA.
- S. Schroff, 1981, Left ventricular dynamics in terms of its viscoelastic chamber properties, Ph.D. dissertation, Univ. Pennsylvania, Philadelphia, Pennsylvania, USA.
- S. Shroff, and W. Motz, 1989, "Left Ventricular Systolic Resistance in Rats with Hypertension and Hypertrophy", Am. J. Physiology, 257:H386-H394.
- H. Suga, K. Sagawa, and L. Demur, 1980, "Determinants of instantaneous pressure in canine left ventricle: time and volume specifications," Circ., vol. 46, pp.256-263.
- S. R. Vaartjes, 1983, Myocardial resistance measured with a flow-clamp technique, Ph.D. dissertation, Tech. Univ. Twente, Ensched, The Netherlands.
- Van Trees, H, 1968, Detection , Estimation and Modulation Theory, John Wiley, New York, USA.
- H. R. Warner and A. F. Toronto, 1959, " Regulation of cardiac output through stroke volume, " Circ. Res., vol. 8, pp. 549-552.
- L. Zhong, D. Ghista, E. Ng, and S. Lim, 2005, "Passive and Active Ventricular Elastances of the Left Ventricle", Biomedical Engineering Online, Vol. 4, #10.

List of Figures:

- Fig. 1, The Cardiac Cycle of the Human
- Fig. 2, Left ventricular volume and Aortic Pressure during the Ejection Phase of the Cardiac Cycle (simulated).
- Fig. 3, Estimated Elastance, Resistance, and Inertance for the simulated volume and pressure.
- Fig. 4, Contribution of Elastance, Inertance, and Resistance to the Aortic Pressure.
- Fig. 5, Right ventricular volume and Pulmonary Pressure during the Ejection Phase of the Cardiac Cycle (simulated).
- Fig. 6, Estimated Elastance, Resistance, and Inertance for the simulated volume and pressure (Right Ventricle).
- Fig. 7, Contribution of Elastance, Inertance, and Resistance to the Pulmonary Pressure.



Typical Cardiac Cycle for the Human. LV=left ventricle, ECG=Electro CardioGram, AP=Aortic Pressure, LVP=Left Ventricular Pressure, LAP=Left Atrial Pressure, LVEDV=Left Ventricular End Diastolic Volume, LVESV=Left Ventricular End Systolic Volume,.

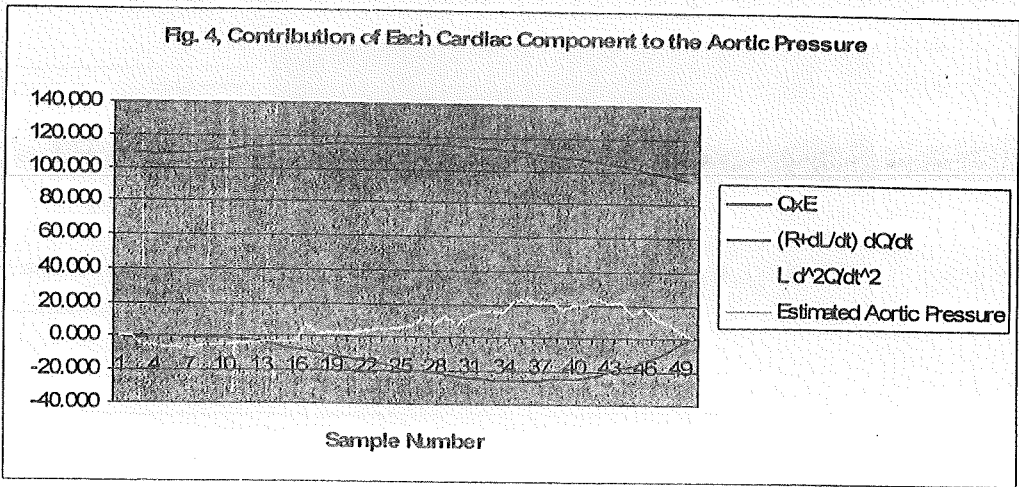
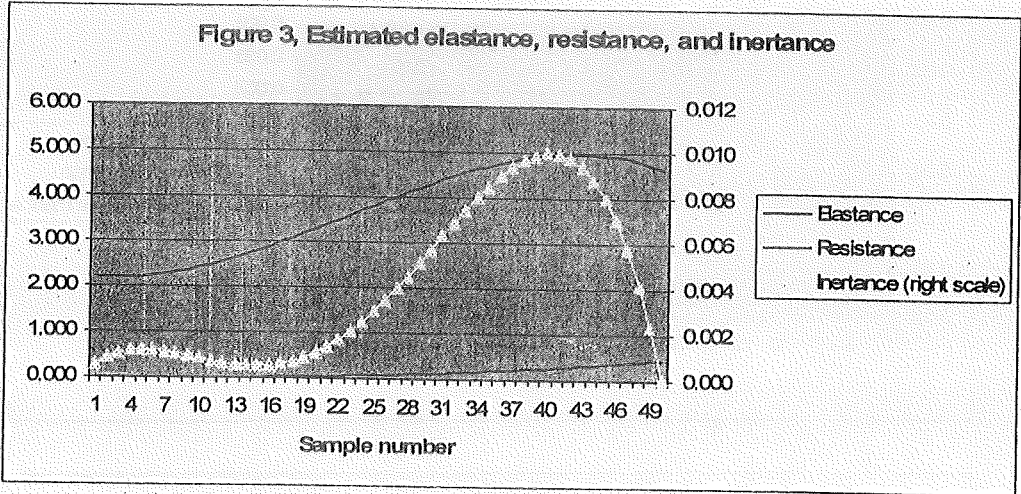
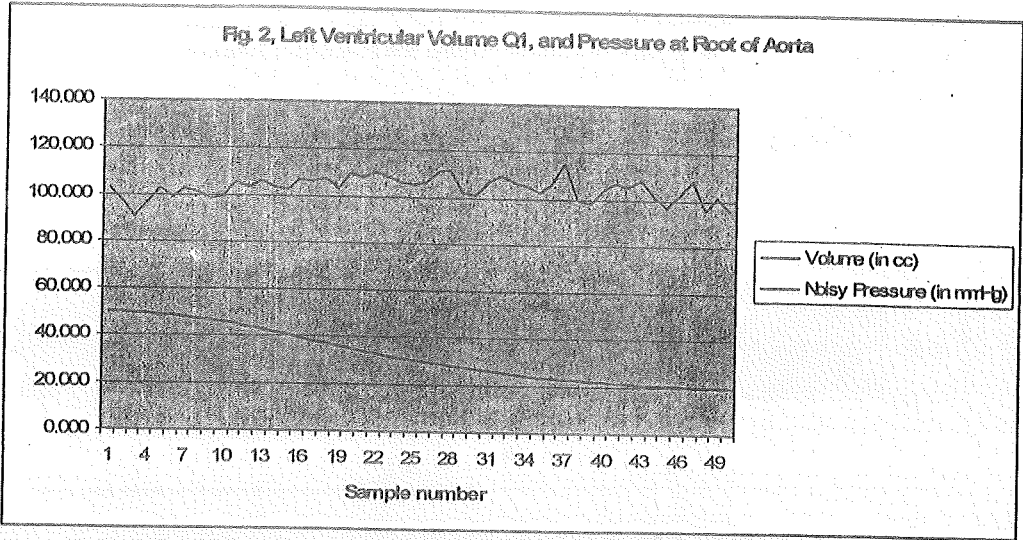


Fig. 5, Right Ventricular Volume and Pulmonary Pressure

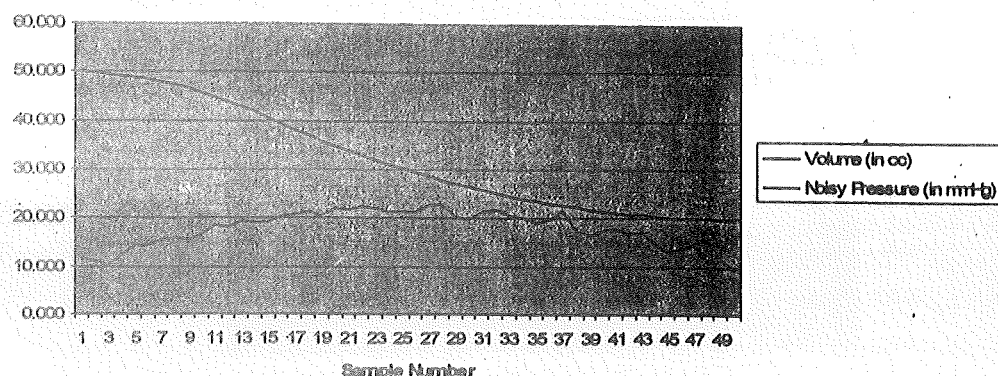


Fig. 6, Estimated Elastance, Resistance, and Inertance (Right Ventricle)

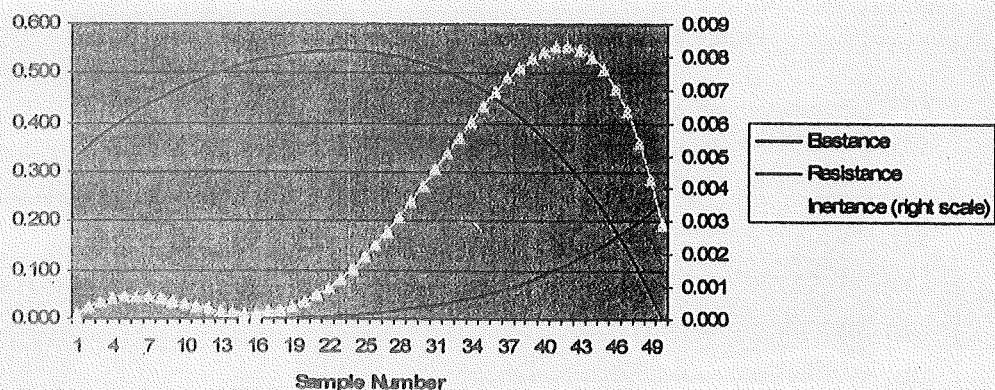
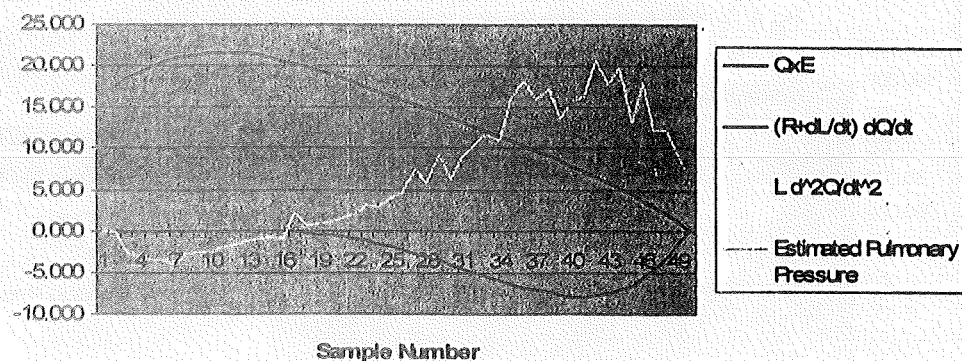


Fig. 7, Contribution to the Pulmonary Pressure



On Shrinkage of the Simple Linear Regression Parameters.

M. Ahmed AL_Robassi

Faculty of Administration science, Taiz University, Taiz, Yemen.

Abstract :

In this paper the estimation of simple linear regression parameters by shrunken method is suggested, when an initial estimate θ_0 of θ is available from the past. In this estimator a variable shrinkage weight function is taken. The expressions for the bias and mean squared error of the proposed estimators are given.

The new estimator is shown to be more efficient than usual estimators when θ is Close to θ_0 .

Introduction

Let $y_i = \theta + \beta x_{i1} + e_i$ $i = 1, 2, \dots, n$... (1)

Represent the simple linear regression equation, where θ, β are unknown parameters and e_i is random error.

Al-Bayyati H.A and Arnold, J.C., (1972), Proposed the Double Stage shrunken Estimator of Simple Linear Regression parameters Using Prior Knowledge.

Their estimator is:

$$\tilde{\theta} = \begin{cases} k(\hat{\theta}_1 - \theta_0) + \theta_0 & \text{if } \hat{\theta}_1 \in R \\ \frac{n_1 \hat{\theta}_1 + n_2 \hat{\theta}_2}{n_1 + n_2} & \text{if } \hat{\theta}_1 \notin R \end{cases} \quad \dots(2)$$

Where $0 \leq k \leq 1$, is the shrinkage factor, R is the parameter, and $\hat{\theta}_1$ is MLE (maximum likelihood estimator) of θ based on n_1 .

Rahim(1976) proposed a shrinkage procedure of estimating the simple linear regression parameters, his estimators i

$$\tilde{\theta} = \begin{cases} k\theta_0 + (1-k)\hat{\theta}, \hat{\beta} & \text{if } |t_{\theta}| \leq t_{\alpha} \\ \hat{\theta}, \hat{\beta} & \text{if } |t_{\theta}| > t_{\alpha} \end{cases} \quad \dots(3)$$

On Shrinkage of the Simple Linear Regression Parameters.

M. Ahmed AL_Robassi

Faculty of Administration science, Taiz University, Taiz, Yemen.

Email:alrobasi5@yahoo.com

Abstract :

In this paper the estimation of simple linear regression parameters by shrunk method is suggested ,when an initial estimate θ_0 of θ is available from the past .In this estimator a variable shrinkage weight function is taken. The expressions for the bias and mean squared error of the proposed estimators are given.

The new estimator is shown to be more efficient than usual estimators when θ is Close to θ_0 .

Introduction

Let $y_i = \theta + \beta x_{i1} + e_i$ $i = 1, 2, \dots, n$... (1)

Represent the simple linear regression equation, where θ, β are unknown parameters and e_i is random error.

Al-Bayyati H.A and Arnold ,J.C. , (1972) ,Proposed the Double Stage shrunk Estimator of Simple Linear Regression parameters Using Prior Knowledge.

Their estimator is:

$$\tilde{\theta} = \begin{cases} k(\hat{\theta}_1 - \theta_0) + \theta_0 & \text{if } \hat{\theta}_1 \in R \\ \frac{n_1 \hat{\theta}_1 + n_2 \hat{\theta}_2}{n_1 + n_2} & \text{if } \hat{\theta}_1 \notin R \end{cases} \quad \dots (2)$$

Where $0 \leq k \leq 1$, is the shrinkage factor, R is the parameter, and $\hat{\theta}_1$ is MLE (maximum likelihood estimator) of θ based on n_1 .

Rahim(1976) proposed a shrinkage procedure of estimating the simple linear regression parameters ,his estimators i

$$\tilde{\theta} = \begin{cases} k\theta_0 + (1-k)\hat{\theta} & \text{if } |\hat{\theta}| \leq t_\alpha \\ \hat{\theta} & \text{if } |\hat{\theta}| > t_\alpha \end{cases} \quad \dots (3)$$

$$\tilde{\beta} = \begin{cases} k\beta_0 + (1-k)\hat{\beta}, \hat{\theta} & \text{if } |t_{\beta}| \leq t_{\alpha} \\ \hat{\beta}, \hat{\theta} & \text{if } |t_{\beta}| \geq t_{\alpha} \end{cases} \quad \dots (4)$$

Where $0 \leq k \leq 1$ is the shrinkage factor and $\hat{\theta}, \hat{\beta}$ are O.L.S. estimators of θ and β respectively.

Al-knane I.H., (1997) Considered the Double Stage shrunk estimator of Simple Linear Regression parameters Using shrinkage weighted function of the form:

$$\Psi(\hat{\theta}) = \begin{cases} 1 - ae^{\frac{b(\hat{\theta} - \theta_0)}{\text{var}(\hat{\theta})}} & \text{if } \hat{\theta} \in R \\ 1 & \text{if } \hat{\theta} \notin R \end{cases} \quad \dots (5)$$

Where a, b are constant and R is a suitable region in the parameter space including θ_0 .

In this paper we suggested the variable shrinkage weighed function of the form:

$$\Psi(\hat{\theta}) = \begin{cases} ae^{\frac{b(\hat{\theta} - \theta_0)}{\text{var}(\hat{\theta})}} & \text{if } \hat{\theta} \in R \\ 1 & \text{if } \hat{\theta} \notin R \end{cases} \quad \dots (6)$$

So we choice R as the region of acceptance of the hypothesis H_0 :

$\theta = \theta_0$ against $H_1: \theta \neq \theta_0$. Then for given level of significance α , a choice for R the pretest region is given by:

$$R = \left[\theta_0 - t_{n-2, \frac{\alpha}{2}} \sqrt{\frac{\sigma^2 \sum x_i^2}{nSS_x}}, \theta_0 + t_{n-2, \frac{\alpha}{2}} \sqrt{\frac{\sigma^2 \sum x_i^2}{nSS_x}} \right] \quad \dots (7)$$

Where $t_{n-2, \frac{\alpha}{2}}$ is the value table for t -distribution with $n-2$ degree freedom.

Therefore a pretest estimators of the parameter θ and β are given by:

$$\tilde{\theta} = \begin{cases} (ae^{\frac{-b(\hat{\theta}-\theta_0)^2}{\text{var}(\hat{\theta})}})(\hat{\theta} - \theta_0) + \theta_0 & \text{if } \hat{\theta} \in R \\ \hat{\theta} & \text{if } \hat{\theta} \notin R \end{cases} \quad \dots(8)$$

$$\tilde{\beta} = \begin{cases} (ae^{\frac{-b(\hat{\beta}-\beta_0)^2}{\text{var}(\hat{\beta})}})(\hat{\beta} - \beta_0) + \beta_0 & \text{if } \hat{\beta} \in R \\ \hat{\beta} & \text{if } \hat{\beta} \notin R \end{cases} \quad \dots(9)$$

Bias, Mean Squared Error, and efficiency

(1) The expressions of the bias ,mean squared error and efficiency for $\tilde{\theta}$ estimator are follows :

$$\text{Bias}(\tilde{\theta} / \theta, R) = E(\tilde{\theta} - \theta)$$

$$= \frac{a\sigma\sqrt{\sum x_i^2}}{\sqrt{nss_x}(2b+1)^3} e^{\frac{-b\lambda^2}{2b+1}} \left[\sqrt{(2b+1)} N_1(L, U) + \lambda N_0(L, U) \right] \quad \dots(10)$$

$$\text{Where } \lambda = (\theta - \theta_0)\sqrt{\text{var}(\hat{\theta})}, \quad N_i = \int_L^U \frac{1}{\sqrt{2\pi}} z^i e^{\frac{z^2}{2}} dz, \quad i = 0, 1$$

$$L = \frac{1}{\sqrt{2b+1}} \left[-\lambda - t_{n-2, \frac{\alpha}{2}}(2b+1) \right] \quad U = \frac{1}{\sqrt{2b+1}} \left[-\lambda + t_{n-2, \frac{\alpha}{2}}(2b+1) \right]$$

$$\begin{aligned}
 MSE(\tilde{\theta} / \theta, R) &= E(\tilde{\theta} - \theta)^2 \\
 &= \frac{\sigma^2 \sum x_i^2}{nss_x} - \frac{2a\sigma^2 \sum x_i^2}{nss_x(2b+1)^{5/2}} e^{-\frac{b\lambda^2}{2b+1}} [(2b+1)N_2(L, U) + \lambda(1-2b)\sqrt{(2b+1)} N_1(L, U) \\
 &\quad - 2b\lambda^2 N_0(L, U)] + \frac{a^2\sigma^2 \sum x_i^2}{nss_x(4b+1)^{5/2}} e^{-\frac{2b\lambda^2}{4b+1}} [(4b+1)N_2(L_1, U_1) + 2\lambda\sqrt{4b+1} N_1(L_1, U_1) \\
 &\quad + \lambda^2 N_0(L_1, U_1)]
 \end{aligned}$$

$$\begin{aligned}
 R.EF(\tilde{\theta} / \theta, R) &= \frac{MSE(\hat{\theta} / \theta)}{MSE(\tilde{\theta} / \theta, R)} \\
 &= 1 / \{1 - \frac{2a}{(2b+1)^{5/2}} e^{-\frac{b\lambda^2}{2b+1}} [(2b+1)N_2(L, U) + \lambda(1-2b)\sqrt{(2b+1)} N_1(L, U) \\
 &\quad - 2b\lambda^2 N_0(L, U)] + \frac{a^2}{(4b+1)^{5/2}} e^{-\frac{2b\lambda^2}{4b+1}} [(4b+1)N_2(L_1, U_1) + 2\lambda\sqrt{4b+1} N_1(L_1, U_1) \\
 &\quad + \lambda^2 N_0(L_1, U_1)] \}
 \end{aligned}$$

(2) The expressions of the bias ,mean squared error and efficiency for $\tilde{\beta}$ estimator are follows :

$$Bias(\tilde{\beta} / \beta, R) = \frac{-a\sigma e^{-\frac{b\lambda_2^2}{2b+1}}}{\sqrt{ss_x}(2b+1)^{3/2}} [\sqrt{2b+1} N_1(L, U) + \lambda_2 N_0(L, U)] \quad \dots(11)$$

Where $\lambda_2 = (\beta - \beta_0) / \text{var}(\hat{\beta})$

$$MSE(\tilde{\beta} / \beta, R) = \frac{\sigma^2}{ss_x} - \frac{2a\sigma^2 e^{\frac{-b\lambda^2}{2b+1}}}{ss_x(2b+1)^{5/2}} [(2b+1)N_2(L, U) + \lambda_2(1-2b)N_1(L, U) - 2b\lambda_2 N_0(L, U)]$$

$$+ \frac{a^2\sigma^2 e^{\frac{-2b\lambda^2}{4b+1}}}{ss_x(4b+1)^{5/2}} [(4b+1)N_2(L, U) + 2\lambda_2 N_1(L, U) + \lambda_2^2 N_0(L, U)] \quad \dots(12)$$

$$R.EF(\tilde{\theta} / \theta, R) = \frac{MSE(\hat{\theta} / \theta)}{MSE(\tilde{\theta} / \theta, R)}$$

$$= 1 / \{ 1 - \frac{2a}{(2b+1)^{5/2}} e^{\frac{-b\lambda^2}{2b+1}} [(2b+1)N_2(L, U) + \lambda(1-2b)N_1(L, U) - 2b\lambda^2 N_0(L, U)] + \frac{a^2}{(4b+1)^{5/2}} e^{\frac{-2b\lambda^2}{4b+1}} [(4b+1)N_2(L_1, U_1) + 2\lambda N_1(L_1, U_1) + \lambda^2 N_0(L_1, U_1)] \} \quad \dots(13)$$

Numerical results

The computation of mean squared error and efficiency of the estimator $\tilde{\theta}$ with respect to O.L.S. estimator $\hat{\theta}$ considered in this section by taking $n = 5(5)30$, $b = 0.5, 0.05, 0.1, 0.01$, $\alpha = 0.1, 0.01, 0.05$.

The following conclusions are based on these computations :

- (1) Mean squared error of $\tilde{\theta}_b$ is increasing function of the shrinkage factor, and decreasing function with the first sample n_1 (see table (2)).
- (2) As expected, the efficiency of the shrunken estimator is better than the efficiency of the O.L.S. estimator, when θ is close to θ_0 .
- (3) Table (3) indicates that the efficiency of the suggested estimator is decreasing function of λ , and increasing function of n .

$$MSE(\tilde{\theta}_B / \theta, R) \text{ for } \theta = 10, \sigma = 15, \lambda = (\theta - \theta_0) \sqrt{\text{var}(\hat{\theta})}$$

$n = 5$				$n = 10$			$n = 15$			$n = 20$		
$\lambda \backslash b$	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.01	0.01	0.50	0.1
0.0	0.938	0.171	0.043	0.103	0.097	0.117	0.046	0.89	0.108	0.043	0.087	0.103
0.2	0.917	0.186	0.051	0.123	0.118	0.128	0.057	0.115	0.126	0.051	0.113	0.123
0.4	0.895	0.203	0.078	0.137	0.133	0.143	0.071	0.128	0.139	0.078	0.126	0.137
0.6	0.872	0.221	0.095	0.155	0.155	0.164	0.103	0.147	0.157	0.095	0.146	0.155
0.8	0.859	0.237	0.117	0.175	0.178	0.181	0.121	0.173	0.176	0.117	0.170	0.175
1.0	0.847	0.251	0.141	0.187	0.195	0.204	0.149	0.192	0.198	0.141	0.189	0.187
1.2	0.819	0.304	0.163	0.221	0.237	0.233	0.169	0.219	0.223	0.163	0.214	0.221
1.4	0.789	0.327	0.184	0.243	0.261	0.274	0.187	0.247	0.257	0.184	0.239	0.243
1.6	0.764	0.339	0.215	0.280	0.315	0.301	0.219	0.293	0.294	0.215	0.270	0.280
1.8	0.735	0.348	0.247	0.327	0.341	0.356	0.249	0.330	0.330	0.247	0.312	0.327
2.0	0.697	0.427	0.278	0.368	0.407	0.419	0.287	0.387	0.379	0.278	0.358	0.368

$$Eff(\tilde{\theta} / \theta, R), \theta = 10, \sigma = 15, \lambda = (\theta - \theta_0) \sqrt{\text{var}(\hat{\theta})}$$

$\lambda \backslash b$	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.01	0.01	0.50	0.1
0.0	7.526	5.011	2.953	7.821	4.189	3.413	8.201	5.203	3.932	8.900	5.813	4.310
0.2	4.703	3.715	2.341	6.395	3.507	2.971	7.053	4.047	3.045	8.015	4.965	3.591
0.4	3.250	2.512	1.932	4.395	2.915	2.576	5.917	3.341	2.610	7.217	4.072	2.843
0.6	2.401	1.903	1.057	3.317	2.014	1.789	4.731	2.865	1.969	6.044	3.001	2.169
0.8	2.53	1.142	0.980	2.650	1.581	1.316	3.021	2.054	1.497	4.351	2.379	1.629
1.0	1.830	0.873	0.789	1.813	0.899	0.983	2.450	1.631	1.000	3.190	1.832	1.110
1.2	1.012	0.705	0.703	1.259	0.758	0.837	1.534	0.971	0.872	2.418	1.006	0.901
1.4	0.831	0.571	0.639	0.985	0.696	0.721	1.069	0.823	0.763	1.505	0.879	0.829
1.6	0.679	0.485	0.561	0.739	0.587	0.649	0.869	0.765	0.703	0.892	0.781	0.740
1.8	0.485	0.445	0.483	0.590	0.569	0.573	0.753	0.682	0.649	0.782	0.718	0.683
2.0	0.401	0.406	0.415	0.543	0.550	0.525	0.614	0.603	0.583	0.631	0.662	0.616

References

- [1] AL—Bayyati H.A and Arnold ,J.C. , (1972) ,On Double Stage Estimation in Simple Linear Reggression Using Prior Knowledge .Technometrics, Vol(14),P.405-414.
- [2] AL-Kanane I.H.(1997), Single and Double Stage Shrunkn Estimators for the Linear Reggression Models.” Ph.d.Thesis, Al-Mustansiryah University Iraq.
- [3] Katte ,S.K., (1962) , Use Of Some Prior Knowledge In Estimation Means From Double Samples , Biometrics , Vol.,(18) ,P 139.
- [4] Rahim, M. A. (1976) ”Pretest estimation of simple linear regression model using prior knowledge of one of the parameters” “Technometrics, Vol(18), p. 439-445.
- [5] Waikar, V.R., Schuurman, F.J., and Raghunathann ,(1984), On a Two Stage Shrinkage Testimator of The Mean of a Normal Distribution. Communication in Statistics.-Theory and Methods., 13(15),p, 1901-1913.

Comparison Between Discriminant and Regression Analyses with an Application

By

Hassan S. M.

Thebes Higher Institute for Management and
Information Technology, Al Haram, Sakkra Rd.,
Giza, Egypt.

E. Mail: drhassn_242@yahoo.com

Abstract

The main objectives of this study is to compare between the discriminant analysis and the regression analysis as two statistical predictive techniques. A discriminant function and a regression equation for a sample of graduates in some higher institutes in Egypt, as an application, are constructed. The parameters of both discriminant function and regression equation are estimated. The significance tests for both models and their predictors are carried out. The estimated discriminant function is used to determine the group membership for any cases that are or not included in the analysis. The estimated regression equation is used to estimate the GPA of any graduates are or not included in the analysis. The high school grade (Sgrade) and the high school type (Stype) are the most important predictors for classifying the graduates in these institutes and determining the GPA of these graduates.

Keywords: Discriminant function, Data mining, Regression analysis, Significance tests.

1. Introduction

Data mining is the process of selecting, exploring, and modeling large amounts of data to uncover new trends and patterns in massive data base. These analyses lead to proactive decision making and knowledge discovery in large databases by stressing data exploration to thoroughly study the structure of data and to check the validity of statistical models that fit [2].

Discriminant Analysis (DA), a multivariate statistical technique is commonly used to build a predictive / descriptive model of group discrimination based on observed predictor variables and to classify each observation into one of the groups. In DA multiple quantitative attributes are used to discriminate single classification variable. DA is different from both the cluster analysis and the multivariate analysis of variance (MANOVA). It is different from the cluster analysis because prior knowledge of the classes, usually in the form of a sample from each class is required [2]. On the other hand, DA is a reversed multivariate analysis of variance, MANOVA. In MANOVA, the independent variables are the groups and the dependent variables are the predictors, while in DA, the

independent variables are the predictors and the dependent variables are the groups [3].

A discriminant function, also called a canonical root, is a latent variable which is created as a linear combination of discriminating (independent) variables, such that

$$L = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_n X_n \quad (1)$$

Where the b 's are discriminant coefficients, the X 's are discriminating variables, and b_0 is a constant. This is analogous to multiple regression, but the b 's are discriminant coefficients which maximize the distance between the means of the criterion (dependent) variable. There is one discriminant function for 2-group discriminant analysis, but for k -group discriminant analysis, the number of discriminant functions will be $k-1$, and each discriminant function is orthogonal to the others [3] and [5].

On the other hand, the multiple-regression analysis is another statistical method for establishing an equation that allows a dependent variable to be estimated from two or more explanatory variables (predictors). The linear regression model takes the form,

$$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + \dots + b_k X_k + e \quad (2)$$

where Y is the dependent variable

X_i are the explanatory variables, $i = 1, 2, \dots, k$

b_0 is the true intercept

b_i are the true slopes, $i = 1, 2, \dots, k$

k is the number of explanatory variables

e is the random error

The common objectives of DA are i) to investigate differences between groups, ii) to discriminate groups effectively, iii) to identify important discriminating variables, iv) to perform hypothesis testing on both the significance of the model and the significance of the independent variables, and v) to classify new observations into pre-existing groups.

In contrast, the common objectives of regression analysis are i) to estimate the regression equation that describes the relation between dependent and independent variables, ii) to perform hypothesis testing of the overall regression model, iii) to perform hypothesis testing of the independent variables and iv) to predict the value of dependent variable for any cases that are not included in the analysis

2. Discriminant and Regression Models

The study population consists of all graduates data from the year of 2005 until the year of 2008 in three higher institutes; Thebes Higher Institute of Management, Thebes Higher Institute of Computer and Thebes Higher Institute of Engineering. We select a random samples of size 311 from this population. We divide these individuals into two categories; group 1 represents pass graduates whose GPA belongs to the interval [1, 2], and group 2 represents good graduates whose GPA belongs to the interval (2, 4]. These two categories are the dependent variables of the discriminant function. The independent variables (predictors) are the score of high school, Sgrade, Gender (male = 1, female = 0), Age, type of high school (public = 1, technical = 0) Stype. Since we have 2-group discriminant analysis, and one discriminant function. The discriminant function will take the form

$$L = b_0 + b_1 \times Sgrade + b_2 \times Gender + b_3 \times Age + b_4 \times Stype \quad (3)$$

That is, we have two groups, four independent variables and one discriminant function.

In contrast, the multiple-regression equation will take the form

$$Y = b_0 + b_1 \times Sgrade + b_2 \times Gender + b_3 \times Age + b_4 \times Stype \quad (4)$$

Where Y is the GPA of the graduate.

3. Estimating the Coefficients of Predictors

The data is run by SPSS-win package using entering all independents together method, where we obtain the numerical results (outputs) as shown in the tables through this study. One of these tables, (Table1), gives un standardized discriminant coefficients which are used in the formula for making the classifications in DA, much as b coefficients are used in regression in making predictions. The constant plus the sum of products of the un standardized coefficients with the observations yields the discriminant scores. These coefficients are partial coefficients, reflecting the unique contribution of each variable to the classification of the criterion variable (GPA) [3] and [4].

Table 1
Estimated Coefficients of Predictors

Predictor	Discriminant Function	Regression Equation	ρ - value	
			Disc. Fn.	Reg. Eq.
Gender	-0.063	- 0.0191	0.965	0.613
Age	0.005	- 0.00088	0.012	0.949
Sgrade	0.052	0.0199	0.000	0.000
Stype	0.547	0.224	0.000	0.000
Constant	- 12.880	- 2.832		

That is, the estimated discriminant function is

$$L = -12.880 + 0.052 \times Sgrade - 0.063 \times Gender + 0.005 \times Age + 0.547 \times Stype . \quad (5)$$

and the estimated regression equation is

$$Y = -2.832 + 0.0199 \times Sgrade - 0.0191 \times Gender - 0.00088 \times Age + 0.224 \times Stype . \quad (6)$$

4. Test the Significance of the Model

To test the significance of the model as a whole, we have the following two subsections:

4.1 Test the Significance of the Discriminant Function

To test the significance of the Discriminant Function, we have the following two hypotheses:

Null hypothesis: the two groups have the same mean discriminant function scores;

$$H_0: \mu_1 = \mu_2$$

Alternate hypothesis: the two groups do not have the same mean discriminant function scores;

$$H_1: \mu_1 \neq \mu_2$$

Table 2 shows that the value of wilks' lambda statistic for the test of discriminant function (chi-square = 249.297) has probability of 0.000 which is less than or equal to the level of significance of 0.05. That is the null hypothesis is rejected and the model differentiates scores among the groups significantly.

Table 2

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.444	249.297	4	.000

4.2 Test the Significance of the Regression Equation

To test the significance of the overall regression, we have the following two hypotheses:

Null hypothesis: All the b_i s equal 0.

Alternate hypothesis: At least one of b_i s is not equal 0.

Since F statistic (273.853) is greater than the critical value of the test (2.83) and the p -value is less than the level of significance, 0.05, the null hypothesis is rejected. That is the overall regression is highly significant (See: Table 3).

Table 3

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	104.348	4	26.087	273.853	.000 ^a
	Residual	29.149	306	.095		
	Total	133.497	310			

a. Predictors: (Constant), secondary type, GENDER, secondary grade, AGE

b. Dependent Variable: GPA

Moreover, the coefficient of multiple correlation between the dependent variable and the independent variables is 0.884, the standard error of the estimate is very small (0.3086401) and the coefficient of determination and the adjusted coefficient of determination are 0.782 and 0.779 respectively. That is the estimated model is more predictive model as shown in Table 4 below.

Table 4

Regression Model Summary

Model	R	R ²	Adj.R ²	Std. Error of Estimate	Change Statistics				
					R Square Change	F Change	df 1	df 2	Sig. F Change
1	.884	.782	.779	.3086401	0.782	273.853	4	306	0.000

5. Test the Significance of the Independent Variables

5. 1 Test the Significance of Predictors of the Discriminant Function

As we mentioned before, the smaller the variable wilks' lambda for an independent variable, the more that variable contributes to the discriminant. Since the value of wilks' lambda statistic for the Gender (0.980), Sgrade (0.456) and Stype (0.895) are the smallest, these predictors contribute to the discriminant function more than the other independent variables. Moreover, since the p -value of the test of each of these predictors is smaller than the level of significance (0.05), these predictors are more significant. That is these predictors discriminate among groups and make $\mu_1 \neq \mu_2$ (See: Table 5).

Table 5

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
GENDER	1.000	.002	1	309	.965
AGE	.980	6.365	1	309	.012
secondry grade	.456	368.928	1	309	.000
secondry type	.895	36.297	1	309	.000

5. 2 Test the Significance of Predictors of the Regression Equation

It is obvious from Table 6 below that there is an inverse relation between GPA and each of gender and age. In contrast, there is a direct relationship between GPA and each of secondary grade and secondary type. On the other hand, Since the p -value of each of secondary grade and secondary type is less than the level of significance (0.05), the two independent variables are highly significant.

Table 6

Estimated Coefficients of Regression Equation

Model	Un standardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
Constant	-2.832	0.355		-7.975	.000
Gender	-1.91E-02	0.038	-0.014	0.506	.613
Age	-8.80E-04	0.014	-0.002	-0.064	.949
Sgrade	1.99E-02	0.001	0.831	29.527	.000
Stype	0.224	0.045	0.142	4.926	.000

6. Important discriminating variables

The standardized discriminant function coefficients should be used to assess each variable's unique contribution to discriminant function [2]. The standardized canonical discriminant function coefficients which are presented in Table 7 below reflect the contribution of one independent variable in the context of the other variables in the model. A low standardized coefficient might mean that the groups do not differ much on that variable or it might just mean that that variable's correlation with the grouping variable is redundant with that of another variable in the model. In other words, the larger the standardized coefficient, the greater is the contribution of the respective variable to the discrimination between groups [6] and [8].

Table 7

Standardized Canonical Discriminant Function Coefficients

	Function
	1
GENDER	-.030
AGE	.007
secondry grade	.958
secondry type	.216

It is obvious from Table (7) above that the independent variables Sgrade and Stype are more contribution in the model than the other variables. Another way to determine which independent variables define a particular discriminant function is to look at the factor structure. Structure matrix below gives coefficients of correlation between independent variables and discriminant function. Generally, any independent with a loading (a score) of 0.30 or more is considered to be important in defining the discriminant dimension [7]. It is obvious from Table 8 below that the independent variables Sgrade and Stype are more correlated with the discriminant function.

Table 8

Structure Matrix

	Function
	1
secondry grade	.976
secondry type	.306
AGE	-.128
GENDER	.002

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions
Variables ordered by absolute size of correlation within function.

On the other hand, it is obvious that the independent variables Sgrade and Stype are more correlated with the estimated regression equation than other variables. It is obvious also that the independent variable Sgrade is more correlated with discriminant function than the regression equation, whereas the independent variable Stype is more correlated with the regression equation than the discriminant function (See: Table 9)

Table 9

Correlation Coefficients Between Dependent Variable (GPA) and Each of Independent Variables

	GPA	Gender	Age	Secondary grade	Secondary type
Pearson Correlation GPA	1.000	0.006	-0.175	0.873	0.392

7. Relation Between Dependent and Independent Variables

The table 10 below shows the eigenvalues, λ , which verify the relation between dependent and independent variables. The larger the eigenvalue, the more of the variance in the dependent variable is explained by that function. Since the dependent in this application has two groups, we have one discriminant function. The second column lists the percent of variance explained by the discriminant function. The third column is the cumulative percent of variance explained. The last column is canonical correlation, where the squared canonical correlation is the percent of variation in the dependent discriminated by the independents in DA. That is the estimated discriminant function is an important function because: 1) 100 % of variance between the the two groups can be explained by this function and 2) the degree of relationship between the predictors and groups (canonical correlation) due to function is 0.746. The canonical correlation which is denoted by η can be obtained for the discriminant function as [1]

$$\eta = \sqrt{\frac{\lambda}{1+\lambda}} = \sqrt{\frac{1.253}{1+1.253}} = 0.746 \quad (7)$$

Table 10
Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	1.253 ^a	100.0	100.0	.746

a. First 1 canonical discriminant functions were used in the analysis

On the other hand, there is a strong positive relation between GPA and both high school (Sgrade), 0.873 and high school type (Stype), 0.392 (See: Table 9).

8. Classification the Cases

Table 11 is used to assess how well the discriminant function work, and if it works equally well for each group of the dependent variable. Here it correctly classifies about 89.1 %. This is rather satisfactory. As results of the estimated discriminant functions, the cases are classified as given in Table 11 below. We notice that 89.9 % of cases were classified into the first group, GPA of interval [1, 2] and 88.2 % of cases were classified into the second group, GPA of (2, 4]. That is estimated discriminant functions gets almost all cases of G_1 and G_2 correctly classified.

Table 11
Classification Results^{b,c}

			Predicted Group Membership		Total
			1	2	
Original	Count	1	142	16	158
		2	18	135	153
	%	1	89.9	10.1	100.0
		2	11.8	88.2	100.0
Cross-validated ^a	Count	1	142	16	158
		2	19	134	153
	%	1	89.9	10.1	100.0
		2	12.4	87.6	100.0

a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

b. 89.1% of original grouped cases correctly classified.

c. 88.7% of cross-validated grouped cases correctly classified.

Fisher's classification function coefficients are used to classify the cases between the groups. The coefficients of the independent variables which are shown in Table 12 below are used to construct a discriminant function for each group, say G_1 and G_2 where,

$$G_1 = -237.348 - 3.864 \times \text{Gender} + 14.224 \times \text{Age} + 0.703 \times \text{Sgrade} + 12.581 \times \text{Stype} \quad (8)$$

$$G_2 = -266.160 - 4.005 \times \text{Gender} + 14.236 \times \text{Age} + 0.818 \times \text{Sgrade} + 13.803 \times \text{Stype} \quad (9)$$

For each case, a G is computed for each group and the case is classified into the group for which G is highest. That is for example if $G_1 > G_2$ then the case is classified into group 1. Moreover, the estimated relations (8) and (9) can be used to classify new observations into pre-existing groups [7].

Table 12

Classification Function Coefficients

	GROUP	
	1	2
GENDER	-3.864	-4.005
AGE	14.224	14.236
secondary grade	.703	.818
secondary type	12.581	13.803
(Constant)	-237.348	-266.160

Fisher's linear discriminant functions

Another technique to determine the group membership of any case is to determine the posterior probabilities of this case. That is each case's discriminant score is used to determine the posterior probabilities of being in each of the two groups. The case is then classified (predicted) to be in the group with higher posterior probability. SPSS-WIN11 output gives group membership of the cases that are included in the analysis.

9. Example

As for example, the estimated discriminant function is applied to determine the group membership of any graduate who included or not in the analysis. The estimated regression equation is also applied to determine the GPA of any graduate who included or not in the analysis. Using estimated equations 6, 8 and 9, we can predict the GPA or the the group of any cases are or not included in the analysis. For example, Table 13 below presents the predicted GPA and group of some cases. That is there is a coincide between the two estimated models for prediction.

Table 13
Case wise Diagnostics

Case Number	GPA	Predicted GPA	Predicted Group
109	1.5111	2.877313	Group 2
110	2.7683	3.816916	Group 2
176	2.4024	4.020673	Group 2

Discussion the Results

1. Both discriminant and regression analyses are two important statistical predictive techniques. The discriminant analysis is more convenient predictive technique in the situations in which the independent variable is a categorical variable. On the other hand, the regression analysis is more convenient predictive technique in the situations in which the independent variable is a single quantitative variable.
2. Both analyses are applied to estimate the parameters. They give good results where the estimated discriminant function and the regression equation are highly significant.
3. The estimated discriminant function has a high predictive ability where the eigenvalue of the function is equal to 1.253. Also the regression equation has a rather high predictive ability where the coefficient of determination of the estimated model is equal to 0.782.
4. The independent variables Sgrade and Stype are more correlated with the estimated discriminant function where a) the correlation between the predictor, Sgrade, and the function is 0.976 and b) the correlation between

the predictor, Stype, and the function is 0.306. Whereas there is a converse relation between the independent variables Age and the estimated discriminant function, that is the youngest graduates are more classified into the group of higher GPA than the oldest ones (see: Table 8). On the other hand, it is obvious that the independent variables Sgrade and Stype are more correlated with the estimated regression equation than other variables (See: Table 9).

5. The estimated discriminant function correctly classifies about 89.1 % of the cases that are included in the analysis. 89.9 % of the classified cases are classified into the first group, GPA of interval [1, 2]. 88.2 % of the classified cases are classified into the second group, GPA of interval (2, 4]. (See: Table 11).

6. The predictors Sgrade and Stype have the most contribution in the classification dependent variable GPA, where a) the canonical discriminant function coefficient between the estimated discriminant function and the Sgrade is 0.958, so there is a direct relation between the predictor, Sgrade and the groups of GPA. b) the canonical discriminant function coefficient between the estimated discriminant function and the Stype is 0.216, so there is a direct relation between the predictor, Stype, and the groups of GPA. That is the graduates whose high schools certificate are more classified into the group of high GPA than others. Also the graduates of high scores in high schools certificate are more classified into the group of high GPA than the graduates who have educated in technical secondary schools. . So the author recommend with modifying the curriculum of these technical schools to satisfy the fulfillment of requirements for studying in these institutes such as addition of some courses or development of the current courses (See: Table 7).

References

[1] Ashour, S. K. and Salem, S. A. (2005). Statistical Presentation and Analysis using SPSSWIN, Part two: Advanced Applied Statistics. Cairo University: ISSR.

[2] George C. J. F (2000). Discriminant Analysis, A Powerful Classification Technique in Data Mining. Statistics and Data Analysis, Paper 247-27.

[3] <http://www2.chass.ncsu.edu>

[4] <http://www.colorado.edu>

[5] <http://www.epa.gov>

[6] Huberty, Carl J. (1994). Applied Discriminant Analysis. NY: Wiley-Interscience.

[7] Karl Wuensch L. (2001). Two Group Discriminant Function Analysis.
<http://core.ecu.edu>

[8] Klecka, William R. (1980). Discriminant Analysis. Quantitative Applications in the Social Sciences Series, No 19. Thousand Oaks, CA: Sage Publications.

[9] McLachlan, Geoffrey J. (2004). Discriminant Analysis and Statistical Pattern Recognition. NY: Wiley- Interscience.

[10] Tabachnick, Barbara G. and Linda Fidell S. (2001). Using Multivariate Statistics, 4th ed., Boston: Allyn and Bacon.

Acknowledgement

The idea of this paper is taken from [1], so I would like to thank the authors;

Fuzzy Logistic Regression: A Multi-objective Programming Approach*

Ramadan Hamed¹, Aly El-Hefnawy², Maha M. El-Ashram³, Hesham A. Abdalla⁴

Abstract

This study extends logistic regression model to fuzzy logistic regression model to be an alternate to Maximum Likelihood approach in situations of small sample size and /or sparse datasets. Maximum likelihood (ML) is the most commonly used approach for estimating logistic regression parameters, and generally perform well for large sample sizes, but it may fail to converge or at least produce poor results especially for small and/or sparse datasets. We propose a new approach based on fuzzy concepts to estimate parameters of the logistic regression model and to be an effective alternative to ML in cases of small and/or sparse datasets. A nonlinear Bi-Objective model is suggested and its performance is evaluated versus ML approach by a simulation study. Results show that the new proposed model outperforms ML in cases of small sample size for models of normally and uniformly distributed covariates. Furthermore, this model is applicable to datasets with fuzzy entities.

Keywords: Fuzzy Numbers; Extension Principle; Logistic Regression; Maximum Likelihood; Small Data Sets; Sparse Data; Multi-objective Programming; Monte Carlo Approach; Similarity Measures.

1 Introduction

Logistic regression is part of a class of statistical models called generalized linear models [1]. This broad class of models encompasses linear regression, analysis of variance, and analysis of covariance, as well as regression models for categorical data. Logistic regression is widely used in biomedical, behavioral and physical researches.

Parameters estimation for logistic regression is usually based on maximizing the unconditional likelihood function. The maximum likelihood estimates (MLEs) of the parameters in logistic regression are asymptotically unbiased with minimum variance when the model is correctly specified. For large well-balanced datasets ML estimation is a satisfactory approach. Unfortunately, ML can fail completely or at least produce poor

* This Paper is a part of PH.D. thesis submitted to Faculty of Economics and Political Sciences, Cairo University.

¹ Professor, Department of Statistics, Faculty of Economics and Political Sciences, Cairo University.

² Associate Professor, Department of Statistics, Faculty of Economics and Political Sciences, Cairo University.

³ Assistant Professor, Department of Statistics, Faculty of Economics and Political Sciences, Cairo University.

⁴ PH.D. Student, Department of Statistics, Faculty of Economics and Political Sciences, Cairo University.

results in terms of estimated probabilities and confidence interval of parameters, specially for small datasets or unbalanced ones where the percentage of observations with Y values coded 1, is close to zero or one [4],[9][12], and [14].

To overcome the special problems that logistic regression analysis of small and sparse data sets faces, special methods have been developed, including estimation and exact inference based on conditional likelihood [9] and penalized maximum likelihood estimation and inference [4],[6]. These methods are updated versions of the classical ML but they provide less biased estimates and more accurate inference.

In this study, a new logistic regression approach based on fuzzy set theory is developed. This approach is based on the idea of Fuzzy Linear Regression proposed by Tanaka et al [20] where, we assume that the unfitted errors in logistic regression can be viewed as the fuzziness of the model components. This means that the residuals between predictions and actual observations are not produced by measurements errors, but rather by the parameters uncertainty in the model, respectively a possibility distribution can be used instead of probability distribution. The new developed approach produces an alternate to the Maximum Likelihood method to estimate the model's parameters, especially in situations where the ML method fails to converge as situations of small or unbalanced (sparse) data sets. In addition, this new approach will be able to deal with fuzzy entities.

This paper is organized as follows: concepts of logistic regression model and the limitations of the Maximum Likelihood approach are reviewed in section 2. In section 3, essential definitions of fuzzy numbers and extension principle are presented. In section 4, the fuzzy logistic regression is formulated and proposed. A Monte Carlo simulation study is designed and presented in section 5 to evaluate the proposed model as an alternate to ML approach. Section 6 presents the measures used to evaluate the proposed model against the ML. Finally, results and conclusions are discussed in sections 7 and 8.

2 Maximum Likelihood Estimates for Logistic Regression Model.

Consider the logistic regression model:

$$\pi(x_i) = \frac{1}{1 + \exp[-\sum_{j=0}^{k-1} \beta_j x_{ij}]} \quad i = 1, 2, \dots, n \quad (1)$$

which represents the relationship between $\pi(x_i) = P(y_i = 1 | x_i, \beta)$ (probability of success) and covariates x 's. Where β_j 's are unknown model parameters. The maximum likelihood method is the most commonly used method of parameter estimation in logistic regression. ML estimates $\hat{\beta}_j$ of the regression parameters β_j 's $j = 0, \dots, k-1$ which can be interpreted as log odds ratio estimates, are obtained as solutions to the score equations:

$$\frac{\partial l(\beta)}{\partial \beta} = X'(Y - \pi) = 0 \quad (2)$$

where $l(\beta) = \log(L(\beta_0, \beta_1, \beta_2, \dots, \beta_{k-1}))$ is the log-likelihood function, X is a $(n \times k)$ matrix of the covariates' values, and Y is a $(n \times 1)$ vector of response variable values. The ML is an iterative approach which requires initial estimates for the parameters β_j 's, and will generally perform well for large sample sizes. But for small data sets or data sets in which the average value of Y is close to zero or one, the method may produce poor results, or even fail to converge [4],[9],[12],[14], and [17].

Some methods have been developed to overcome problems that logistic regression analysis of small and sparse data sets faces, including estimation and exact inference based on conditional likelihood [9] and penalized maximum likelihood estimation and inference [4],[6]. They generally are updated versions of the classical ML. Although, exact logistic regression can provide finite and accurate estimates in some situations, it cannot generally be used as a tool to deal with small and sparse data [8]. While, the availability of the penalized maximum likelihood estimation in statistical package software is very limited and rare.

3 Fuzzy numbers and Extension principle

This section recalls the basic definitions of fuzzy number and fuzzy operations. The concept of fuzzy numbers was first introduced by Zadeh [23]-[25] with the purpose of analyzing and manipulating approximate numeric values. The concept has been refined by Dubois & Prade [5], and several definitions exist in [13],[18],[27].

Definition: A fuzzy number is a fuzzy subset defined on the real line $\mathbb{R}$. A fuzzy number $\tilde{M}$ is generally assumed to have a membership function which satisfies that [7]:

Normality: it exists exactly one $m \in \mathbb{R}$ with $\mu_{\tilde{M}}(m) = 1$ (m is called the mean value of $\tilde{M}$).

Convexity: $\forall x_1 \in \mathbb{R}, \forall x_2 \in \mathbb{R}$, and $\forall \lambda \in [0, 1]$,

$$\mu_{\tilde{M}}(\lambda x_1 + (1 - \lambda)x_2) \geq \min(\mu_{\tilde{M}}(x_1), \mu_{\tilde{M}}(x_2)),$$

$\mu_{\tilde{M}}(x)$ is upper semicontinuity,

Where $\mu_{\tilde{M}}(\cdot)$ represents the membership function of $\tilde{M}$ and λ is any weighting factor in $[0, 1]$. $\mu_{\tilde{M}}(\cdot)$ has different forms, consequently, there are many kinds of fuzzy numbers. L - R representation of Fuzzy Numbers [5] is a general form that covers many types of membership function. L and R are functions defining the left and right sides of a fuzzy number and are usually denoted $L(\cdot)$ and $R(\cdot)$. $L(\cdot)$ and $R(\cdot)$ are called reference functions of fuzzy numbers and satisfy the following conditions:

1. $L(x) = L(-x)$, $R(x) = R(-x)$,

2. $L(0) = 1, R(0) = 1,$
3. $L(x)$ and $R(x)$ are strictly decreasing for $x \geq 0$

Triangular Fuzzy Numbers (T.F.N.) is a special case of L-R fuzzy numbers.

Definition: A Triangular Fuzzy Number (T.F.N.) $\tilde{M}$ is defined by its lower limit m_l , upper limit m_u , and central value m , where $m_l < m < m_u$. So that, the membership function of $\tilde{M}$ has a triangular shape and it is defined as:

$$\mu_{\tilde{M}}(x) = \begin{cases} 0 & \text{if } x < m_l \\ \frac{x - m_l}{m - m_l} & \text{if } m_l \leq x \leq m \\ \frac{m_u - x}{m_u - m} & \text{if } m < x \leq m_u \\ 0 & \text{if } x > m_u \end{cases} \quad (3)$$

If the triangular shape is symmetric, i.e., $(m - m_l) = (m_u - m) = s$, the membership function can be written as:

$$\mu_{\tilde{M}}(x) = \max(0, 1 - |x - m|/s) \quad c > 0 \quad (4)$$

Basic operations and functions on these fuzzy numbers can be defined by the help of the extension principle stated by Zadeh [22].

Definition: If X and Y are two sets, and f is a mapping $f : X \rightarrow Y$ such that for all $x \in X, f(x) = y \in Y$, then f can be extended to operate on fuzzy subsets of X in the following way:

Let $\tilde{A}$ be a fuzzy subset of X with membership function $\mu_{\tilde{A}}$, then the image of $\tilde{A}$ in Y is the fuzzy subset $\tilde{B}$ with the membership function $\mu_{\tilde{B}}$:

$$\mu_{\tilde{B}}(y) = \begin{cases} \sup \{ \mu_{\tilde{A}}(x); x : y = f(x), x \in X \}, \\ 0 & \text{if there is no } x \in X \text{ such that } f(x) = y. \end{cases} \quad (5)$$

Applying this principle to functions like $x + y$ or $x \cdot y$, we can define the basic operations on fuzzy numbers.

4 A Bi-Objective Model for Fuzzy Logistic Regression.

Consider the logistic regression model (1). Assuming the model parameters are fuzzy numbers, the model can be expressed as the following fuzzy logistic regression model:

$$\pi(x_i) = \frac{1}{1 + \exp[-\sum_{j=0}^{k-1} \tilde{\beta}_j x_{ij}]} \quad i = 1, 2, \dots, n \quad (6)$$

where $\tilde{\beta}_j$ is assumed to be a symmetrical triangular fuzzy number (4) denoted as $(c_j, s_j)_L$. Accordingly, the estimated values of π will be symmetric triangular fuzzy numbers. To estimate these fuzzy parameters, the proposed approach is based on the proposition 3 in Bárdossy [2],[3].

Bárdossy proved that if there is a non-decreasing function for any x and j such that:

$$g(u) = f(x, \beta_0, \dots, \beta_{j-1}, u, \beta_{j+1}, \dots, \beta_{k-1})$$

And using Zadeh' extension principle (5), the conditions of fuzzy regression can be written as

$$\begin{aligned} y_{iL} &\geq f(x_i, \beta_L) \quad i = 1, \dots, n \\ \text{and} \\ y_{iU} &\leq f(x_i, \beta_U) \quad i = 1, \dots, n \end{aligned} \quad (7)$$

where

$$\begin{aligned} y_{iL} &= \inf_z \{z : \mu_{y_i}(z) \geq h\} \\ y_{iR} &= \sup_z \{z : \mu_{y_i}(z) \geq h\} \end{aligned}$$

Assuming the parameters β_j 's are L-R fuzzy numbers $\beta_j = (c_j, s_{jL}, s_{jR})$, the β_L and β_R are expressed as

$$\begin{aligned} \beta_L &= \{(c_0 - s_{0L} L_0^{-1}(h)), \dots, (c_{k-1} - s_{k-1,L} L_{k-1}^{-1}(h))\}, \\ \beta_U &= \{(c_0 - s_{0R} L_0^{-1}(h)), \dots, (c_{k-1} - s_{k-1,R} L_{k-1}^{-1}(h))\}. \end{aligned}$$

This proposition expresses the lower (upper) limit of the fuzzy number y_{ih} as functions of x and the lower (upper) limit of β respectively.

According to this proposition, assuming we have crisp data from the logistic model, i.e., x_i 's and $\pi(x_i)$ (denoted p_i) are crisp and β is composed of symmetric triangular fuzzy numbers with membership function $L(z) = \max(0, 1 - |z|)$, the conditions of Fuzzy Logistic Regression can be expressed as:

$$\begin{aligned} \frac{1}{1 + \exp[-(c^t x_i - (1-h)s^t |x_i|)]} &\leq p_i, \quad i = 1, \dots, n \\ \frac{1}{1 + \exp[-(c^t x_i + (1-h)s^t |x_i|)]} &\geq p_i, \quad i = 1, \dots, n \end{aligned} \quad (8)$$

In case we have fuzzy data from the logistic model, such that x_i 's are crisp and each $\pi(x_i)$ is a symmetric triangular fuzzy number denoted by $(p_i, e_i)_L$, the conditions of Fuzzy Logistic Regression will be:

$$\begin{aligned} \frac{1}{1 + \exp(-(c^t x_i - (1-h)s^t |x_i|))} &\leq p_i - (1-h)e_{iL}, & i = 1, \dots, n \\ \frac{1}{1 + \exp(-(c^t x_i + (1-h)s^t |x_i|))} &\geq p_i + (1-h)e_{iR}, & i = 1, \dots, n \end{aligned} \quad (9)$$

In the current study, we consider the case of crisp x_i ($i = 1, \dots, n$) to ensure obtaining symmetric fuzzy estimates of π_i .

Based on Özelkan and Duckstein idea [16], we suggest a Bi-objective fuzzy logistic model based on the tradeoff between unfitted errors of boundaries of the fuzzy logistic model and unfitted errors of the center of the model. The proposed fuzzy logistic model is formulated as follow:

$$\begin{aligned} \min - \text{dominate } \{D^p, \varepsilon^p\} \\ \text{s.t.} \\ E_{iL}(h) - O_{iL}(h) &\leq \varepsilon_{iL} & i = 1, 2, \dots, n \\ O_{iU}(h) - E_{iU}(h) &\leq \varepsilon_{iU} & i = 1, 2, \dots, n \\ \varepsilon_{iL}, \varepsilon_{iU} &\geq 0 & i = 1, 2, \dots, n \end{aligned} \quad (10)$$

where

O_{iL} and O_{iU} are the lower and upper limits of the fuzzy observations represented as intervals at chosen h -level.

E_{iL} and E_{iU} are the lower and upper limits of the fuzzy predictions represented as intervals at chosen h -level.

$\varepsilon^p = \sum_i (\varepsilon_{iL}^p + \varepsilon_{iU}^p)$ is L_p norm of relaxation variables where $1 \leq p \leq \infty$.

$D^p = \sum_i \{O_{iC} - E_{iC}\}^p$ is L_p norm of distance between the center of the fuzzy observations and the center of the fuzzy predictions.

Using the Euclidean norm ($p = 2$) and conditions (8) and (9), we can reformulate model (10) in two different cases as follow:

Case I. Crisp x_i and Crisp π_i .

$$\begin{aligned} \min \sum_{i=1}^n (\varepsilon_{iu}^2 + \varepsilon_{il}^2) + \sum_{i=1}^n \left(p_i - \frac{1}{1 + \exp(-\mathbf{c}^t \mathbf{x}_i)} \right)^2 \\ \text{s.t.} \\ \frac{w_i}{1 + \exp(-(\mathbf{c}^t \mathbf{x}_i - (1-h)s^t | \mathbf{x}_i |))} \leq p_i + \varepsilon_{il}, \quad i = 1, \dots, n \\ \frac{w_i}{1 + \exp(-(\mathbf{c}^t \mathbf{x}_i + (1-h)s^t | \mathbf{x}_i |))} \geq p_i - \varepsilon_{iu}, \quad i = 1, \dots, n \\ \varepsilon_{iu}, \varepsilon_{il} \geq 0, \quad i = 1, \dots, n \\ s \geq 0 \end{aligned} \quad (11)$$

Case II. Crisp x_i and Fuzzy π_i .

$$\begin{aligned} \min \sum_{i=1}^n (\varepsilon_{iu}^2 + \varepsilon_{il}^2) + \sum_{i=1}^n \left(p_i - \frac{1}{1 + \exp(-\mathbf{c}^t \mathbf{x}_i)} \right)^2 \\ \text{s.t.} \\ \frac{w_i}{1 + \exp(-(\mathbf{c}^t \mathbf{x}_i - (1-h)s^t | \mathbf{x}_i |))} \leq p_i - (1-h)e_i + \varepsilon_{il}, \quad i = 1, \dots, n \\ \frac{w_i}{1 + \exp(-(\mathbf{c}^t \mathbf{x}_i + (1-h)s^t | \mathbf{x}_i |))} \geq p_i + (1-h)e_i - \varepsilon_{iu}, \quad i = 1, \dots, n \\ \varepsilon_{iu}, \varepsilon_{il} \geq 0, \quad i = 1, \dots, n \\ e_i \geq 0, \quad i = 1, \dots, n \\ s \geq 0 \end{aligned} \quad (12)$$

where

- ε_{iu} and ε_{il} are relaxation variables.
- c_j center of the j^{th} parameter of the logistic model.
- s_j spread of the j^{th} parameter of the logistic model.
- p_i center of the dependent variables $P(y_i = 1)$.
- e_i spread of the dependent variables $P(y_i = 1)$.
- w_i weight corresponding to data after being grouped.

In this context, n represents the number of grouped data that are set up while computing the probability $p_i = P(y_i = 1)$ from the original data set, and w_i 's are weights corresponding to that grouped data.

To implement this model, we consider two main aspects. First, how can the probabilities p_i be obtained or computed if the data used to estimate the logistic model parameters consist of the covariates $\mathbf{x}$ and the response variable represented as 0,1 variable rather than probabilities. Second, which value of h - level should be chosen.

Real data sets of logistic models often consist of the values of the covariates x and the values of the response variable represented in binary form rather than probabilities. Therefore, to implement the proposed model, the values 0,1 of the response variable should be expressed as probabilities. To obtain these probabilities, we suggest the following steps:

- Break down each covariate into a limited number of levels (intervals).
- Compute the mean value of each interval.
- Identify all possible combinations of covariates levels or intervals
- Identify the number of combinations n and the number of observations into each combination m_i , $i = 1, \dots, n$ where the total number of observations

$$M = \sum_{i=1}^n m_i .$$

- Compute the probability p_i for each combination which is equivalent to the average of y values corresponding to combination i .

In this context, n new observations (combinations) are then obtained; each of them has the modified values of x_i and the corresponding p_i . Consequently, we can implement the proposed models with these new observations.

Reducing the covariates into a limit number of levels leads to reduction of the number of observations from M to n observations ($M > n$). Consequently, weighting can be used to compensate for sampling schemes that stratify on the covariates, giving results that more accurately reflect the population. The new modified values of the covariates are weighted by their actual weights in the original data set. Therefore, the proposed models are weighted by the weight w_i where $w_i = \exp\left(\frac{m_i}{M}\right)$.

Selection of h - level is the second issue to discuss for implementing these models. Conditions (8) and (9) in those models take into account that the membership degree of each observation is greater than an imposed threshold possibility h where $h \in [0,1]$. This criterion simply expresses the fact that the fuzzy output of the model should cover all data points to a certain h - level. The selection of h - level value will influence the spread of the parameters, consequently the estimation errors. Therefore, we adopt the algorithm proposed by Nasrabadi and Nasrabadi [15] to select h - level that minimizes the estimation error.

5 A Monte Carlo Simulation Study.

A simulation study is designed to empirically assess the performance of the proposed Fuzzy Logistic Regression model in a number of situations and compares its performance to that of the Maximum Likelihood approach (when appropriate). The simulation study is structured in two situations: the first situation, when all the covariates

and the response variable are crisp (non fuzzy). The second situation, where the response variable is generated as a fuzzy variable while the covariates are crisp. Three factors are considered in each situation:

1. Sample size: small sample size ($n = 30$), moderately large sample size ($n=200$) and reasonably large sample size ($n=600$).
2. Model type (based on number of parameters): simple model (two parameters and one covariate) and multiple model (four parameters and three covariates).
3. Type and Distribution of covariates: continuous with Normal distribution, continuous with Uniform distribution, Binary (Bernoulli) and mixed (Normal, Uniform, Binary).

Table 1 shows the details of 21 groups arising from the different combinations of the categories of the three factors.

Table 1. Combination of Simulated Data
According to Simulation Factors

Combination	Sample Size	Number of Covariates (k-1)	Distribution of Covariates
Group 1	30	1	N(0,2)
Group 2	30	3	N(0,2)
Group 3	200	1	N(0,2)
Group 4	200	3	N(0,2)
Group 5	600	1	N(0,2)
Group 6	600	3	N(0,2)
Group 7	30	1	U(-6,6)
Group 8	30	3	U(-6,6)
Group 9	200	1	U(-6,6)
Group 10	200	3	U(-6,6)
Group 11	600	1	U(-6,6)
Group 12	600	3	U(-6,6)
Group 13	30	1	Binary
Group 14	30	3	Binary
Group 15	200	1	Binary
Group 16	200	3	Binary
Group 17	600	1	Binary
Group 18	600	3	Binary
Group 19	30	3	Mixed
Group 20	200	3	Mixed
Group 21	600	3	Mixed

Therefore, in this simulation, we have 42 groups of factors combination; 21 for the crisp simulated data and 21 for the fuzzy simulated data. The factor that is held fixed in the simulation study is the number of replications for each combination of design factors. Since the objective of the simulation study is to assess the performance of the proposed model compared to ML, rather than studying the asymptotic properties of estimated parameters, we select only 50 replications for each combination of design factors. We

think the selected number of replications is enough to investigate the comparison between ML and the proposed model.

Situation I: Crisp Data generation

In this situation, the set of covariates values x are randomly generated from a specified underlying distribution $N(0,2)$, $U(-6,6)$, Bernoulli or mixed for each independent sampling unit in the sample, rather than fixing the values of covariates at pre-selected values. On the other hand, the logistic model parameters are generated randomly to be in the range $[-3,3]$. Thus, the simulation sitting can be considered more reflective of observational studies rather than traditional experimental designs in which the investigator takes control over parameters and covariates sitting.

After generating the covariates and parameters, the true probability of "success" is then computed for each independent sampling unit to be $\pi_i(x_i, \beta)$ ($i = 1, \dots, n$). The outcome variable Y is generated accordingly by comparing an independently generated $U(0,1)$ variate, u , to true underlying computed logistic probability using the rule

$$y = \begin{cases} 1 & u_i \leq \pi_i(x_i, \beta) \\ 0 & \text{otherwise} \end{cases}$$

Situation II: Fuzzy Data generation

Through this situation, we generate fuzzy data sets from logistic regression models which assumed to have fuzzy parameters, a fuzzy response variable and crisp covariates. We assume the underlying logistic model has fuzzy parameters with symmetric triangular fuzzy numbers or $\tilde{\beta}_j = (c_j, s_j)$ $j = 1, \dots, k$ where c_j is the center value and s_j is the spread of parameter j . We adopt the technique suggested by Sun and Wu [19] to use Monte Carlo approach to generate a fuzzy parameter j . The steps of the adopted technique are the following::

1. Generate randomly the center c_j in the range $[-3,3]$.
2. Generate randomly the spread s_j as a percent of c_j .
3. Compute β_{jL}, β_{jU} the lower and upper limit of $\tilde{\beta}_j$ where $\beta_{jL} = c_j - s_j$ and $\beta_{jU} = c_j + s_j$.
4. Generate crisp numbers randomly on $[\beta_{jL}, \beta_{jU}]$ to represent values of $\tilde{\beta}_j$ at different levels of α -cut.

As in situation I, the set of covariates values is randomly generated from a specified underlying distribution $N(0,2)$, $U(-6,6)$, Bernoulli or mixed for each independent sampling unit in the sample. The number of generated parameters sets for simple models is 2 and 4 for multiple models. Sample size is chosen to be 30, 200, and 600.

Due to generating fuzzy parameters with lower and upper limits, the true probability of "success" will be fuzzy with lower and upper limits. These limits are computed for each independent sampling unit to be $\pi_{iL}(x_i, \beta_L)$ and $\pi_{iU}(x_i, \beta_U)$ $i = 1, \dots, n$. Using the same rule as in situation I, the outcome variable y can be generated.

6 Performance Measures.

Since, our aim is to compare the performance of the proposed fuzzy model against the ML approach, the selection of the goodness of fit measure that will be used here, should take into account two main aspects:

- The proposed model is built on fuzzy concepts consequently the selected measures should be applicable to both possibilistic and probabilistic bases.
- The selected measures should not be built on the concept of ML approach, in order not to tendentious to that approach.

Accordingly, a fuzzy similarity measure is used to evaluate the performance of the proposed as well as the classical predictive measure of correct classification. We use Yang et al [21] similarity measure $S(\tilde{A}, \tilde{B})$ that measures the distance between two L-R type fuzzy numbers $\tilde{A}_i = (a_i, a_{iL}, a_{iR})_{LR}$ and $\tilde{B}_i = (b_i, b_{iL}, b_{iR})_{LR}$ as:

$$S(\tilde{A}_i, \tilde{B}_i) = \begin{cases} 1 & \text{if } \tilde{A}_i = \tilde{B}_i \\ \exp(-d_{LR}^2(\tilde{A}_i, \tilde{B}_i)/\sigma_i) & \text{if } \tilde{A}_i \neq \tilde{B}_i \end{cases}$$

where

$$\begin{aligned} d_{LR}^2(\tilde{A}_i, \tilde{B}_i) &= (a_i - b_i)^2 + ((a_i - \ell_i a_{iL}) - (b_i - \ell_i b_{iL}))^2 + ((a_i - \tau_i a_{iR}) - (b_i - \tau_i b_{iR}))^2 \\ \sigma_i &= \{ |(a_i - \ell_i a_{iL}) - (b_i - \ell_i b_{iL})| + |(a_i - \tau_i a_{iR}) - (b_i - \tau_i b_{iR})| \} / 2 \\ &\quad + \{ |(a_i - a_{iL}) - (b_i - b_{iL})| + |(a_i - a_{iR}) - (b_i - b_{iR})| \} / 2^3 \end{aligned} \quad (13)$$

and

$$\ell_i = \int_0^1 L^{-1}(v_i) dv \quad \tau_i = \int_0^1 R^{-1}(v_i) dv$$

when $v = 1$ both $\tilde{A}_i, \tilde{B}_i$ will be the common triangular fuzzy numbers. This measure is defined for i^{th} observation, respectively the index for the whole data set can be computed as:

$$SIM = \frac{\sum_i S(\tilde{A}_i, \tilde{B}_i)}{n} \quad (14)$$

This index is used to measure the similarity between the original probability π_i (true probability of simulated data) and the estimated probability of the proposed

model $\hat{\pi}_i$. In case of crisp data, the original probabilities π_i are crisp numbers, thus the left and right spreads are zeros. To make an equitable comparison between the performance of the proposed model and ML, we use the statistical confidence interval of $\hat{\pi}_i$ of ML as well as using the fuzzy interval of $\hat{\pi}_i$ of the proposed model to compute the similarity index for each of them. Basically, the computed interval of the proposed model is different from the interval of ML and hence must be adjusted to ensure equitable comparison. To solve this issue, the idea of Kim et al [11] is used to find a value of h -level (h^*), by which the area between the upper and lower limit of the fuzzy interval of $\hat{\pi}_i$ estimated from the proposed models is equal to that of 95% prediction interval of ML.

Another intuitively appealing way is to summarize the results of a fitted logistic regression model by a classification table. This table is the result of cross-classifying the outcome variable Y with $\hat{Y}$ the variable whose values are derived from the estimated logistic probabilities. To obtain $\hat{y}$, a cut point must be defined, and then each estimated probability is compared to this cut point. We select here this cut point at .5 which is the most commonly used value. Hence the estimated values of y will be:

$$\hat{y}_i = \begin{cases} 1 & \text{if } \hat{\pi}_i \geq .5 \\ 0 & \text{if } \hat{\pi}_i < .5 \end{cases}$$

7 Results

Results of Situation I.

The relative performance of the proposed fuzzy logistic model versus the ML approach is analyzed by obtaining the differences between each of the proposed models and the ML with respect to both similarity measure (14) and overall rate of correct classification. Paired T-test is employed to analyze the results at 5% significance level. In addition, we count for each group, how many times the proposed model results are superior the ML results (the superiority of proposed model on ML)

Similarity Measure.

For small size samples, Table 2. shows that the proposed model significantly outperforms ML in three groups (group 1, 8 and 19) out of seven groups in this category. Two of them are multiple models with a constant term and three covariates, and the third is a simple model with a constant term and one covariate. In four groups (2, 7, 13 and 14), performance of the Bi-objective Model is non-significantly different from ML where group 7 and 13 stand for simple models with uniform and binary covariates respectively. The Percentage of superiority of the Bi-objective Model on ML within each group ranged from 44% to 84%.

For large size samples, Table 2. clearly shows that ML significantly outperforms the suggested model in all groups with Moderately Large and Reasonable Large samples,

except group 5 which represents simple model with Normal distribution covariate and sample size 600. This group records 66% of the superiority of suggested model on ML.

Table 2. Bi-objective Model versus ML Performances
Similarity Measure.

	ML	Bi-objective Model	T-Test	Sig.	Bi-objective Model do better
Group 01	.3662	.4665	-3.415	.001	58%
Group 02	.3235	.3503	-.892	.377	52%
Group 03	.8703	.5781	12.384	.000	0%
Group 04	.8605	.4012	19.572	.000	0%
Group 05	.4662	.5957	-4.290	.000	66%
Group 06	.9277	.4204	22.970	.000	0%
Group 07	.3535	.3799	-1.077	.287	48%
Group 08	.2407	.3786	-5.210	.000	84%
Group 09	.8973	.4492	17.240	.000	2%
Group 10	.8960	.2733	50.023	.000	0%
Group 11	.9432	.4219	18.614	.000	0%
Group 12	.9393	.2966	57.885	.000	0%
Group 13	.5085	.5290	-.755	.454	44%
Group 14	.3896	.4282	-.929	.357	44%
Group 15	.8446	.5705	10.024	.000	6%
Group 16	.7785	.4845	9.148	.000	2%
Group 17	.9080	.6098	9.865	.000	0%
Group 18	.8874	.5698	12.811	.000	0%
Group 19	.3235	.3955	-2.509	.015	58%
Group 20	.8847	.3929	20.122	.000	0%
Group 21	.9249	.4094	21.943	.000	0%

Overall Rate of Correct Classification.

The overall rate of correct classification index measures the degree of classification accuracy rather than the model fit. Therefore, we use this index as supplementary tool to investigate the classification capability of the proposed models. Results show the followings:

For small size samples, Table 3. shows that Bi-objective Model significantly outperforms ML in two groups of the small sample groups (1 and 7). However, its performance is non-significantly different from the ML in two groups (13 and 14). The majority of these groups (1,7 and 13) stand for simple models with different distributions of covariates. Percentage of superiority of the Bi-objective model on ML for these groups ranged form 78% (group 14) to 98% (group 1).

For large size samples, the suggested model does not outperform ML for all large samples. However, its performance is approximately equivalent to the ML especially for groups with large sample and one covariate. As a general result, we can conclude that the overall rate of correct classification of the Bi-objective Model is the same as over all rate of ML for most of the groups specially for those representing simple models. In addition, we can notice that the average overall rate of correct classification for this model is always above 70%.

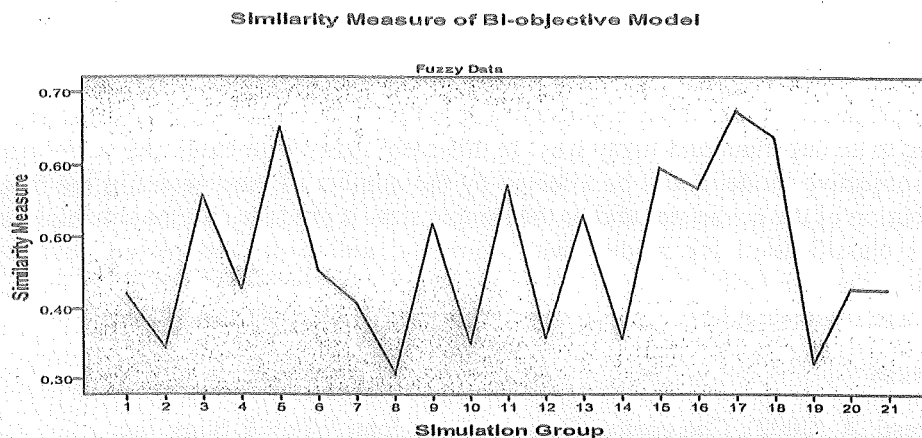
Table 3. Bi-objective Model versus ML Performances
Overall Rate of Correct Classification

	ML	Bi-objective Model 1	T-Test	Sig.	Bi-objective Model do equal or better
Group 01	.5013	.8380	-14.417	.000	98%
Group 02	.9493	.7047	6.078	.000	14%
Group 03	.8448	.8466	-1.498	.141	76%
Group 04	.9051	.8987	3.944	.000	30%
Group 05	.5183	.8459	-22.892	.000	100%
Group 06	.9146	.9135	1.339	.187	44%
Group 07	.4947	.8320	-8.211	.000	94%
Group 08	.9787	.7073	6.517	.000	14%
Group 09	.9150	.9079	1.010	.317	74%
Group 10	.9534	.9466	4.503	.000	32%
Group 11	.9086	.9007	1.021	.312	58%
Group 12	.9494	.9448	2.492	.016	30%
Group 13	.7727	.7767	-.092	.927	90%
Group 14	.7727	.8100	-.862	.393	78%
Group 15	.7929	.7789	1.000	.322	98%
Group 16	.8420	.8422	-1.429	.159	100%
Group 17	.8185	.8185	NA	NA	100%
Group 18	.8180	.8183	-1.429	.159	100%
Group 19	.9480	.7953	5.036	.000	28%
Group 20	.9356	.9239	4.087	.000	46%
Group 21	.9196	.9097	1.531	.132	34%

Results of Situation II.

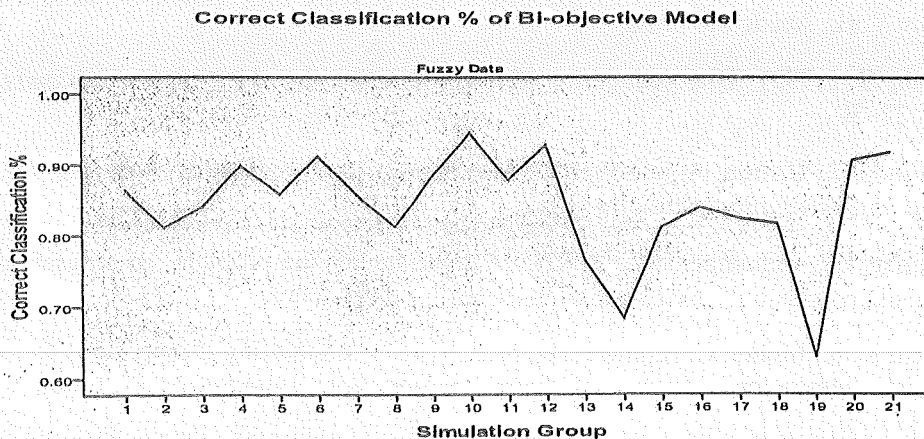
The similarity measure of the proposed Bi-objective Model (Graph 1.) indicates that the performance of the model is affected by type and distribution of covariates, the sample size and number of parameters. It records the minimum value (.31) in data sets with uniformly distributed covariates and small sample size. The highest value (.68) is recorded for simple model with binary covariates and large sample (group 17). Graph 1. indicates that small samples record a low level of similarity measure compared to moderate and reasonable large samples with the same distribution type of covariates. In

general, multiple models (three covariates) have a low level of similarity compared to simple models.



Graph 1. Bi-Objective Model Performance (Similarity Measure).

Concerning the correct classification rate (Graph 2.), the proposed Bi-objective Model rates are always more than 80% for all groups except some groups with binary covariates (group 13 and 14) and group 19 with mixed distributions and small sample. Generally speaking, the proposed model yields reasonable high rates of correct classification for models with normally, uniformly and mixed covariates versus moderate high rates for models with binary covariates.



Graph 2. Bi-Objective Model Performance (Correct Classification Rate).

8 Conclusions

A Bi-objective mathematical model is formulated to extend logistic regression model to fuzzy logistic regression model to overcome the failure of the Maximum Likelihood approach in cases of small sample size and /or spare datasets. The new

proposed model is evaluated through a simulation study. We conclude that the proposed model performs successfully for small datasets. Moreover, it doesn't fail if the average value of Y is exactly 0 or 1. Generally, It outperforms ML in cases of small sample size for models of normally and uniformly distributed covariates.

The proposed model is an applicable approach when variables in the dataset are interacting in an uncertain and fuzzy way. In this case(fuzzy data case), the performance of the Bi-objective model could be affected by the number of parameters in the model, the distribution of the covariates and/or the sample size. It produces reasonable high rates of correct classification for models with normally, uniformly and mixed covariates. Considering the similarity measure, the proposed model performs better in cases of moderate and reasonable large sample and simple models with one covariates.

References

- [1] Agresti, A. (2002), *Categorical Data Analysis*, John Wiley & Sons, Inc.
- [2] Bárdossy, A., Bogárdi, I. and Duckstein, L. (1993), "Fuzzy nonlinear regression analysis of dose-response relationships", *European J. of Oper. Res.*, 66, 36-51.
- [3] Bárdossy, A. (1990), "Note on fuzzy regression", *Fuzzy Sets and Systems*, 37, 65-75.
- [4] Bull, S. B., Mak, C., and Greenwood, C. M. T. (2002), "A modified score function estimator for multinomial logistic regression in small samples", *Computational Statistics & Data Analysis*, 39, 57-74.
- [5] Dubois, D. and Prade, H. (1980), *Fuzzy sets and systems: Theory and applications*. New York: Academic Press.
- [6] Firth, D. (1993) "Bias reduction of maximum likelihood estimates." *Biometrika* 80, 27-38.
- [7] Galindo, J., Urrutia, A. and Piattini, M. (2006), *Fuzzy Databases Modeling, Design And Implementation*, Idea Group Inc.
- [8] Heinze, G., Schemper, M., (2002), "A solution for the problem of separation in logistic regression", *Statistics in Medicine* 21, 2409-2419.
- [9] Hirji, K. F., Mehta, C. R. and Patel, N. R. (1987) "computing distributions for exact logistic regression", *Journal of American Statistical Association*, 82 (400), 1110-1117.
- [10] Hosmer, D. W. and Lemeshow, S. (2000), *Applied Logistic Regression*, John Wiley & sons, Inc., USA.
- [11] Kim, K. J., Moskowitz, H. and Koksalan, M. (1996), "Fuzzy versus statistical linear regression" *European Journal of Operational Research* 92, 417-434.

- [12] King, E. N. and Ryan, T. P. (2002), "A preliminary investigation of maximum likelihood logistic regression versus exact logistic regression", *The American Statistician*, 56 (3), 99-170.
- [13] Klir, G.J. and Yuan B. (1995), *Fuzzy Sets And Fuzzy Logic*, Prentice-Hall Inc.
- [14] Mehta, C. R., Patel, N. R. and Senchaudhuri, P. (2000), "Efficient Monte Carlo methods for conditional logistic regression" *Journal of American Statistical Association*, 95 (449) 99-108.
- [15] Nasrabadi, M. M. and Nasrabadi E. (2004), "A mathematical-programming approach to fuzzy linear regression analysis" *Applied Mathematics and Computation* 155, 873-881.
- [16] Özelkan, E.C. and Duckstein, L. (2000), "Multiobjective fuzzy regression: a general framework", *Comput. Oper. Res.* 27, 635-640.
- [17] Ryan, T. P. (1997), *Modern Regression Methods*, John Wiley & sons, Inc., New York.
- [18] Sakawa, M. (1993), *Fuzzy Sets and Interactive Multiobjective Optimization*, Plenum Press, New York.
- [19] Sun H. and Wu, J. (2006), "A new approach for ranking fuzzy numbers based on fuzzy simulation analysis method", *Applied Mathematics and Computation* 174, 755-767.
- [20] Tanaka, H., Uejima, S. and Asai, K. (1982), "Linear regression analysis with fuzzy model", *IEEE Trans. Systems, Man, Cybernet.* 12, 903-907.
- [21] Yang, M.-S., Hung, W.-L. and Chang-Chien, S.-J. (2005), "On a similarity measure between LR-type fuzzy numbers and its application to database acquisition" *International Journal of Intelligent Systems* 20, 1001- 1016.
- [22] Zadeh, L. A. (1965), "Fuzzy sets", *Information and Control* 8, 338-353.
- [23] Zadeh, L. A. (1975)a, "The concept of linguistic variable and its application to approximate reasoning-I". *Information Sciences*, 8, 199-249.
- [24] Zadeh, L. A. (1975)b, "The concept of linguistic variable and its application to approximate reasoning-II". *Information Sciences*, 8, 301-357.
- [25] Zadeh, L. A. (1975)c, "The concept of linguistic variable and its application to approximate reasoning-III". *Information Sciences*, 9, 43-60.
- [26] Zadeh, L. A. (1978), "Fuzzy Sets as a basis for a theory of possibility", *Fuzzy Sets and Systems*, 1, 3-28.
- [27] Zimmermann, H.-J. (1991), *Fuzzy Set Theory And Its Applications*, 2nd ed., Kluwer Academic.

An Application of Markov Random Walk to a Dialysis System

M. A. El-Shehawey, M. F. Yassen, and N. G. Ragab
Math. Dept., Damietta Faculty of Science, Mansoura University
"Damietta Branch",
P.O. Box 6, New Damietta, Egypt.
E-mail address: el_shehawey@mans.edu.eg

Abstract

The application of the Markov random walk with a partially reflecting and an absorbing barrier has been demonstrated to a biological example arising from a system of countercurrent dialysis. Explicit expressions for the mean of the occupation total as well as the expected time to absorption and the absorption probabilities have been investigated.

Keywords: Markov Random walk; Tridiagonal matrix inversion;
Dialysis system; Hemodialysis; Peritoneal dialysis; Absorption probability
AMS subject classifications. 60G40, 60J80

1. Introduction

Markov random walk (MRW) between an absorbing and a partially reflecting barrier has been discussed in earlier works (cf. Weesakul (1961), Watterson (1961), Parzen (1962), Neuts (1963, 1964), Cox and Miller (1965), Feller (1968), Hardin and Sweet (1970), Kemeny and Snell (1976), Srinivasan and Mehata (1978), Iosifescu (1980), Percus (1985), Rudolph (1999, 2001), El_Shehawey (1994, 2000, 2008_a, 2008_b). A biological example of such a MRW, arising from a system of countercurrent dialysis has been discussed in several papers (cf. Craig and King (1955, 1956), Craig et al. (1957), Kastenbaum (1960, 1966), Davies (1978, 1985), Baigent et al. (2001)).

Kastenbaum (1960), suggested a technique for obtaining the probabilities of being in a specific state after a given number of transitions, if the probability is unity of being in the reflecting state at time zero, using the eigenvalues of the stochastic matrix. These probabilities could be achieved, using the concept of path lengths, by Kemperman (1961). Kastenbaum (1966) demonstrated how his technique may be extended to yield full Toeplitz matrix of higher transition probabilities.

In this paper, we consider a (MRW) $\{X_n, n = 0, 1, 2, \dots\}$ with state space $\{1, 2, \dots, m+1\}$, absorbing barrier at $m+1$, and variable transition probabilities. Following the notation of Kemeny and Snell (1976) and Iosifescu (1980), we let T_{ij} be the number of times the chain is in state j , given $x_i = i$, before absorption at $m+1$, and b_j be the probability of absorption at $j = m+1$, given $x_i = i$, and e is the column vector of units. We write the one-step transition matrix of the (MRW) in the form

EL-Shehawey et al.
Stochastic Modeling Dialysis

$$P = (p_{ij}) = \begin{pmatrix} Q & I & R \\ 0 & I & I \end{pmatrix}$$

Then $N = (n_{ij}) = (ET_{ij})_{ij} = (I - Q)^{-1}$, $N_2 = E(T_{ij}^2) = N(2N_{00} - I)$, $T_i = \sum_j T_{ij}$,

$$m_1 = (E(T_i))_i = Ne, \quad m_2 = (2N - I)Ne, \quad \text{and} \quad B = (b_{ij})_{ij} = NR.$$

Countercurrent dialysis is a diffusion system in which a linear series of cells are used. Each dialysis cell is termed a stage, and each dialysis period is termed a cycle. After each cycle, the dialout at each stage is concentrated and used as the dialin in the succeeding stage, whereas the diallin at each stage is concentrated and returned to the dialysis sac at the previous stage. The dialin at the first stage is retained at the first stage; and the dialout at the final stage is taken out of the system. In the nomenclature of stochastic processes, the first stage represents a partially reflecting barrier and the last stage an absorbing barrier. The objective of the present work is to find the expected of the occupation total the probability of being absorbed at $m+1$ after a given number of cycles, and the probability of absorption.

The paper is organized as follows. In Section 2, we introduce the basics of dialysis. In Section 3, we formulate the model of dialysis system, which divided into the diffusion matrix and the transfer matrix. In Section 4, we discuss some auxiliary concepts of matrices. The last section is devoted to the main results, in which the closed formulas for our objective in this work have been obtained; some very simple forms have been immediately provided lead to closed formulas for some special cases.

2. Basics of dialysis

Dialysis is a method of removing toxic substances (impurities or wastes) from the blood when the kidneys are unable to do so. Dialysis works on the principles of the diffusion of solutes and convection of fluid across a semi-permeable membrane. (CF. Glicksman (2000), Goldman, Ausiello (2007)).

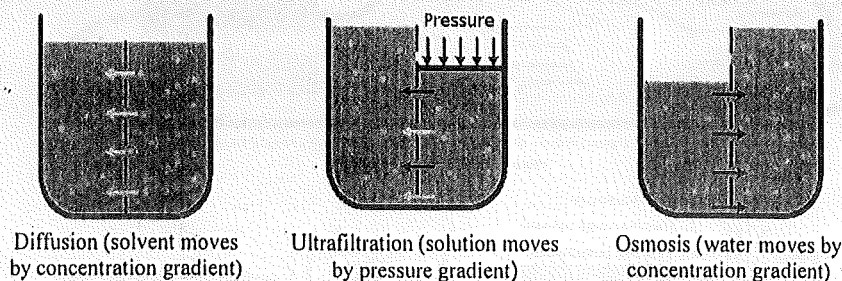
The fluid used to clean the blood flows in the opposite direction to the blood on the opposite side of the membrane, while waste and extra fluid are removed from the blood and end up in the dialysate by controlling three processes: (see Fig.1)

(1) Diffusion is the transfer of dissolved solutes in fluid across the membrane due to concentration gradient. By controlling the chemicals in the dialysate, the dialysis machine controls this transfer of solutes according to the doctor's prescription. Dialysis machines control the chemicals in the dialysate by mixing dialysis fluid concentrates, which are strong versions of the chemicals, (acetate or sodium bicarbonate plus acetic acid based solutions) in purified water. The mixing is generally controlled using either conductivity or volumetric measured of the water and concentrates. Sodium profiling,

changing the amount of sodium in the dialysis fluid at different times during the treatment, is used in many modern dialysis machines to vary the conductivity of the dialysate over the course of the treatment to improve the total amount of solute removed by diffusion.

(2) **Ultrafiltration**, or convection, is a forced fluid flow through the membrane, due to a difference in pressure on the two sides of the dialyzer (pressure gradient). This controls the patient's weight loss over the course of the treatment. The earlier dialysis machines controlled dialysate pressure or the pressure difference across the membrane in order to achieve ultrafiltration, while modern dialysis machines are generally volumetric. Volumetric control is generally achieved either by controlling the flow of dialysate in and out of the dialyzer at different rates with two flow controllers. Volumetric control allows the doctor to take advantage of more effective "high flux" dialyzers, which allow a great deal of fluid movement with a very little pressure differences.

(3) **Osmosis** is the net movement of water across a selectively permeable membrane driven by a difference in the amounts of solute on the two sides of the membrane. In dialysis, this refers not to water movement across the hemodialyzer membrane, but across cell membranes within the body, within the red cells to the blood plasma, or from cells of the various tissues in the body (like muscles) to interstitial fluid. Sodium profiling can be used to increase the rate of osmosis early in the treatment by increasing the sodium level of the plasma.



"Fig.1 Processes of the controlling of the dialysate" [www.toltecint.com]

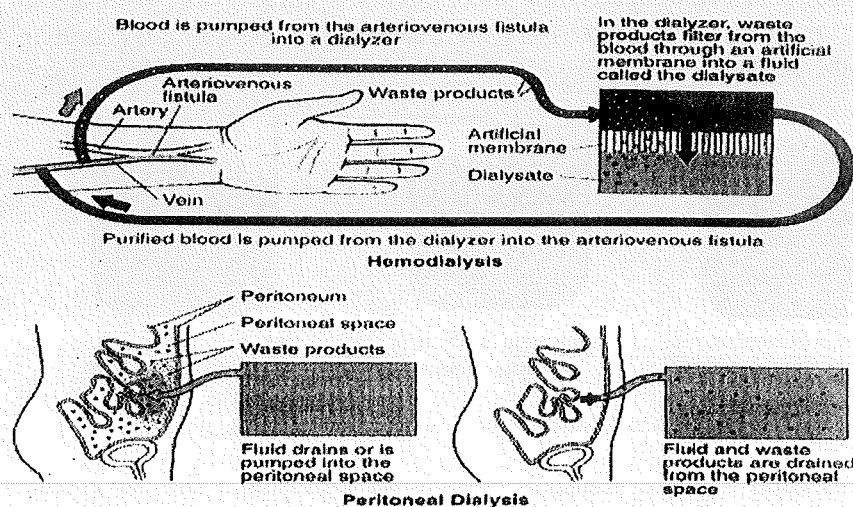
Diffusion is part of transport phenomena. Diffusion is the movement of molecules from a region of higher concentration to one of lower concentration by random molecular motion. In a phase with uniform temperature, diffusion processes tend to lead towards even distributions of molecules. Diffusion is a statistical phenomenon i.e. the chance of a molecule "jumping" from one volume to another depends on the number of molecules in the first volume, so molecules in volumes, which have a relatively high initial concentration, tend to disperse to less concentrated areas until a balance of exchange (equilibrium) is reached. 'Pure' diffusion is generally superimposed on other transport phenomena such as

convection.

Blood is diverted from the access point in the body to a dialysis machine. Here, the blood flows counter-current to a special solution called the dialysate. The chemical imbalances and impurities of the blood are corrected and the blood is then returned to the body.

There are three types of dialysis, hemodialysis, peritoneal dialysis and renal transplant: (CF. Glicksman (2000), Goldman and Ausiello (2007)):

- 1- Hemodialysis works by circulating the blood through special filters outside the body. The blood flows across a filter along with solutions that help to remove toxins. This form of dialysis is normally done 3 sessions every week, where each session takes about 4 hours at hospital, or a dialysis clinic.
- 2- Peritoneal dialysis filters toxic substances using the peritoneal membrane inside the abdomen. The abdomen is filled with special solutions that help remove toxins. The solutions remain in the abdomen for a time and then drained out. This form of dialysis can be performed at home, but must be done every day. The comparing between hemodialysis and peritoneal dialysis shown in Fig.2.
- 3- Renal Transplant is the organ transplant of a kidney in a patient with end stage renal disease.



"Fig.2 Comparing hemodialysis With peritoneal dialysis" [www.merck.com]

3. Formulation of the model of dialysis system

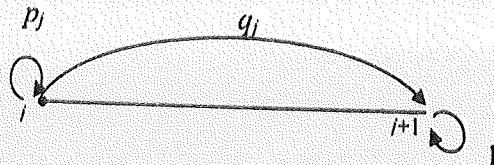
3.1 The diffusion matrix

Let m be the number of stages in the system, p_i be the probability that, in any stage j , a particle inside the dialysis sac will not penetrate to the outside volume, and $q_i = 1 - p_i$, the probability that a particle will penetrate to the

outside. Also specify that, once outside, a particle cannot return inside, except by manual transfer. Then the matrix of diffusion probabilities for m stages is a square matrix of order $2m$ given by

$$D = \begin{pmatrix} D_1 & 0 & & 0 \\ 0 & D_1 & & \\ & & \ddots & \\ & & & D_{m-1} & 0 \\ 0 & & & 0 & D_m \end{pmatrix}, \text{ where } D_j = \begin{pmatrix} p_j & q_j \\ 0 & 1 \end{pmatrix} \text{ for all } j = 1, 2, \dots, m. \quad (3.1)$$

D_j is the matrix of diffusion probabilities for the j th stage (Fig.3).



"Fig.3 Transitions for the j th stage of the diffusion matrix"

The odd rows and columns of D correspond to events taking place inside the dialysis sac for successive stages and the even rows and columns correspond to events taking place on the outside for successive stages.

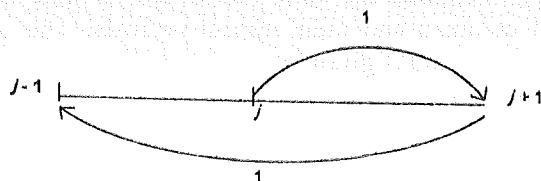
3.2 The transfer matrix

A cycle is defined as a constant interval of time commencing with the transfer of material into the dialysis sac (the beginning of the diffusion process) to the end of the diffusion process.

The transfer matrix, T , is a square matrix of order $2m$ whose elements are probabilities equal to zero or unity and which may be constructed according to the following scheme:

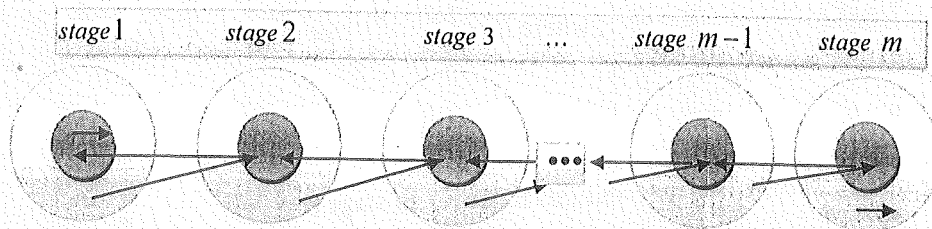
At the end of each cycle:

- (i) a particle in the inside at stage 1 will remain in the inside of stage 1 with probability 1;
- (ii) a particle in the inside at stage j will be transferred to the inside of stage $j-1$ with probability 1, for $j = 2, 3, \dots, m$ see (Fig.4);
- (iii) a particle in the outside of stage j will be transferred to the inside of stage $j+1$ with probability 1, for $j = 1, 2, \dots, m-1$ see (Fig.4);
- (iv) a particle in the outside of stage m will remain in the outside of stage m with probability 1. i.e. $p_{11} = 1$, $p_{mm} = 1$.



"Fig.4 Transitions for the manual transfer matrix"

This transfer mechanism is described pictorially in (Fig.5). In the nomenclature of MRW, the inside of the first stage is a partially reflecting barrier, and the outside of the m th stage is an absorbing barrier.



"Fig.5 Manual transfer after each dialysis cycle"

The matrix T has the form:

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & & 0 \\ 0 & 0 & 1 & 0 & 0 & & 0 \\ 1 & 0 & 0 & 0 & 0 & & 0 \\ 0 & 0 & 0 & 0 & 1 & & 0 \\ & & & & & 0 & 0 & 1 & 0 \\ & & & & & 1 & 0 & 0 & 0 \\ 0 & & & & & 0 & 0 & 0 & 1 \end{pmatrix} \quad (3.2)$$

The product DT is a square matrix of order $2m$ with non-negative elements and unit row sums. The elements of DT are the conditional probabilities that a particle will make a transition from one of the $2m$ states to another after both dialysis and transfer have taken place. Such a matrix is stochastic matrix

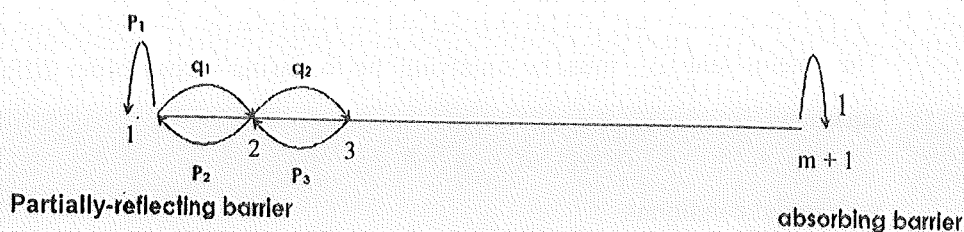
$$DT = \begin{pmatrix} p_1 & 0 & q_1 & 0 & 0 & & 0 \\ 0 & 0 & 1 & 0 & 0 & & 0 \\ p_2 & 0 & 0 & 0 & q_2 & & 0 \\ 0 & 0 & 0 & 0 & 1 & & 0 \\ & & & & & 0 & 0 & 1 & 0 \\ & & & & & p_m & 0 & 0 & q_m \\ 0 & & & & & 0 & 0 & 0 & 1 \end{pmatrix} \quad (3.3)$$

By the rules for its construction, the transfer matrix T will have $m-1$ null columns, namely, all those columns, except the $2m$ th, involving transfers to the outside volume. Hence the matrix DT will have $m-1$ null columns whose positions are known. This property suggests that an initial simplification would be to reduce the matrix DT of order $2m$ to a matrix DT of order $m+1$, by eliminating the $m-1$ null columns and their respective rows. The resulting $m+1 \times m+1$ matrix is of full rank, and is given by

EL-Shehawey et al.
Stochastic Modeling Dialysis

$$DT = \begin{pmatrix} p_1 & q_1 & 0 & 0 & \dots & \dots & \dots & 0 \\ p_2 & 0 & q_2 & 0 & & & & \vdots \\ 0 & p_3 & 0 & q_3 & & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & & p_{m-1} & 0 & q_{m-1} & 0 \\ \vdots & & & & 0 & p_m & 0 & q_m \\ 0 & \dots & \dots & \dots & 0 & 0 & 0 & 1 \end{pmatrix} \quad (3.4)$$

This matrix can be described by the following diagram:



"Fig.6 one absorbing one semi-reflecting barriers at $m+1$ and 1 respectively"

We can see from this diagram that we have (MRW) with one absorbing and one partially-reflecting barriers at $m+1$ and 1 , respectively.

4. Some auxiliary concepts of matrices

Consider the following Jacobi tridiagonal matrix $J \equiv J(a_1, c_1; b_1, a_1, c_1; b_n, a_n)$:

$$J = \begin{pmatrix} a_1 & c_1 & 0 & 0 & \dots & 0 \\ b_2 & a_2 & c_2 & 0 & & \vdots \\ 0 & b_3 & a_3 & c_3 & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & c_{n-1} \\ 0 & \dots & \dots & 0 & b_n & a_n \end{pmatrix} \quad (4.1)$$

The matrix (4.1) is called Toeplitz tridiagonal matrix, $J_1 \equiv J(a, c; b, a, c; b, a)$, if the matrix elements $a_i = a, b_i = b, c_i = c, 1 \leq i \leq n$, called perturbed Toeplitz tridiagonal matrix, $J_2 \equiv J(a_1, c_1; b, a, c; b_n, a_n)$, if the "boundary" elements a_1 and a_n are different from a ; c_1 different from c ; and b_n different from b .

Lemma 4.1 In the case $a_i + b_i + c_i = 1, 2 \leq i \leq n-1$, the continuant

$\Delta = \det(I - J), 1 \leq i, j \leq n$, J defined in (4.1), is given by:

$$\Delta = \sum_{i \rightarrow j} \left(\prod_{k=i+1}^j (c_k) \right) \left(\prod_{k=i}^j (b_k) \right), \quad (4.2)$$

where $i, j = 2, \dots, n$ in the case $a_1 + c_1 = 1$, $a_n + b_n < 1$; $i, j = 1, \dots, n-1$ in the case $a_1 + c_1 < 1$, $a_n + b_n = 1$ and $i, j = 1, \dots, n$ in the case $a_1 + c_1 < 1$, $a_n + b_n < 1$.

Note that the value of the continuant $\Delta_{i \rightarrow j} = \det(I - J), 1 \leq i, j \leq n$ does not exist in the case $a_1 + c_1 = a_n + b_n = 1$.

Proof

In the case $a_1 + c_1 = 1$, $a_n + b_n < 1$, we extend the assumption $a_i + b_i + c_i = 1$, $2 \leq i \leq n-1$ to include the case $i = n$, by assuming that there is a constant $c_n : c_n > 0$ satisfies $c_n = 1 - (a_n + b_n)$, then

$$a_i + b_i + c_i = 1, \quad 2 \leq i \leq n. \quad (4.3)$$

The determinant $\Delta_{i \rightarrow j}$ must be satisfying the following second order difference equation:

$$\Delta_{i \rightarrow j} = (1 - a_j) \Delta_{i \rightarrow j-1} - b_j c_{j-1} \Delta_{i \rightarrow j-2}, \quad (4.4)$$

with initial conditions

$$\Delta_{i \rightarrow i} = 1 - a_i, \quad \Delta_{i \rightarrow i-1} = 1, \quad \Delta_{i \rightarrow i-2} = 0. \quad (4.5)$$

From (4.3), the difference equation (4.4) with (4.5) can be reduced to the following linear difference equation of first order:

$$\Delta_{i \rightarrow j} - c_j \Delta_{i \rightarrow j-1} = b_j \left[\Delta_{i \rightarrow j-1} - c_{j-1} \Delta_{i \rightarrow j-2} \right]. \quad (4.6)$$

Putting $\Delta_{i \rightarrow j} - c_j \Delta_{i \rightarrow j-1} = \delta_j$, (then $\Delta_{i \rightarrow j-1} - c_{j-1} \Delta_{i \rightarrow j-2} = \delta_{j-1}$), formula (4.6) becomes

$$\delta_j = b_j \delta_{j-1}, \quad (\text{and } \delta_{j-1} = b_{j-1} \delta_{j-2}). \quad \text{Then } \delta_j = b_j b_{j-1} \delta_{j-2}.$$

The difference equation (4.6) can be solved by simple iteration as

$$\delta_j = \prod_{s=i}^j b_s,$$

then

$$\Delta_{i \rightarrow j} - c_j \Delta_{i \rightarrow j-1} = \prod_{s=i}^j (b_s). \quad (4.7)$$

The general solution of the homogeneous equation for (4.7) is

$$\Delta_{i \rightarrow j} = \prod_{k=i}^j (c_k). \quad (4.8)$$

By dividing both sides of equation (4.7) by $(\prod_{k=i}^j (c_k))$, then we have

$$\frac{\Delta_{i \rightarrow j}}{\prod_{k=i}^j (c_k)} - \frac{\Delta_{i \rightarrow j-1}}{\prod_{k=i}^{j-1} (c_k)} = \frac{\prod_{s=i}^j (b_s)}{\prod_{k=i}^j (c_k)}. \quad (4.9)$$

By replacing $j-1$ instead of j in equation (4.9), we have

$$\frac{\Delta_{i \rightarrow j-1}}{\prod_{k=i}^{j-1} (c_k)} - \frac{\Delta_{i \rightarrow j-2}}{\prod_{k=i}^{j-2} (c_k)} = \frac{\prod_{s=j}^{j-1} (b_s)}{\prod_{k=i}^{j-1} (c_k)}. \quad (4.10)$$

Repeating this process, $j-i$ times, we get the following particular solution of the equation (4.7)

$$\Delta_{i \rightarrow j} = \prod_{k=i}^j (c_k) \sum_{s=i}^j \left(\prod_{k=i}^s \left(\frac{b_k}{c_k} \right) \right). \quad (4.11)$$

From (4.8) and (4.11), then the general solution of linear difference equation of second order (4.7) is given by:

$$\Delta_{i \rightarrow j} = F \prod_{k=i}^j (c_k) + \prod_{k=i}^j (c_k) \sum_{s=i}^j \left(\prod_{k=i}^s \left(\frac{b_k}{c_k} \right) \right), \quad (4.12)$$

where F is constant.

Using the initial condition $\Delta_{i \rightarrow i-1} = 1$, we get $F = 1$.

Therefore

$$\Delta_{i \rightarrow j} = \prod_{k=i}^j (c_k) + \prod_{k=i}^j (c_k) \sum_{s=i}^j \left(\prod_{k=i}^s \left(\frac{b_k}{c_k} \right) \right),$$

or

$$\Delta_{i \rightarrow j} = \prod_{k=i}^j (c_k) + \sum_{s=i}^j \left(\prod_{k=s+1}^j (c_k) \right) \left(\prod_{k=i}^s (b_k) \right),$$

which can be easily simplified to get formula (4.2).

In the case $a_1 + c_1 < 1$, $a_n + b_n = 1$, we extend the assumption $a_i + b_i + c_i = 1$, $2 \leq i \leq n-1$ to include the case $i=1$, by assuming that there is a constant $b_1 : b_1 > 0$ satisfies $b_1 = 1 - (a_1 + c_1)$, then

$$a_i + b_i + c_i = 1, \quad 1 \leq i \leq n-1. \quad (4.13)$$

By the same argument of the proof of the first case, we obtain the corresponding result.

In the case $a_1 + c_1 < 1$, $a_n + b_n < 1$, we extend the assumption $a_i + b_i + c_i = 1$, $2 \leq i \leq n-1$ to include two cases $i=1, n$, by assuming that there exist two constants b_1 and $c_n : b_1 > 0, c_n > 0$ satisfying $b_1 = 1 - (a_1 + c_1)$, and $c_n = 1 - (a_n + b_n)$, then

$$a_i + b_i + c_i = 1, \quad 1 \leq i \leq n. \quad (4.14)$$

By the same argument of the proof of the first case, we obtain the corresponding result. These complete the proof of Lemma 4.1.

To confirm the results in Lemma 4.1, we have the following proof by induction as follows

- At $j = i$, then formula (4.2) will be

$$\Delta_{i \rightarrow i} = c_i + b_i,$$

since $\Delta_{i \rightarrow i} = (1 - a_i) = b_i + c_i$,

then the formula (4.2) holds for $j = i$.

- At $j = i + 1$, then formula (4.2) will be

$$\Delta_{i \rightarrow i+1} = c_i c_{i+1} + c_{i+1} b_i + b_i b_{i+1},$$

since

$$\Delta_{i \rightarrow i+1} = \det \begin{pmatrix} 1 - a_i & -c_i \\ -b_{i+1} & 1 - a_{i+1} \end{pmatrix} = b_i b_{i+1} + b_i c_{i+1} + c_i c_{i+1},$$

then the formula (4.2) holds for $j = i + 1$.

- At $j = i + 2$, then formula (4.2) will be

$$\Delta_{i \rightarrow i+2} = c_i c_{i+1} c_{i+2} + c_{i+1} c_{i+2} b_i + c_{i+2} b_i b_{i+1} + b_i b_{i+1} b_{i+2},$$

since

$$\begin{aligned} \Delta_{i \rightarrow i+2} &= \det \begin{pmatrix} 1 - a_i & -c_i & 0 \\ -b_{i+1} & 1 - a_{i+1} & -c_{i+1} \\ 0 & -b_{i+2} & 1 - a_{i+2} \end{pmatrix} \\ &= b_i b_{i+1} b_{i+2} + b_i b_{i+1} c_{i+2} + b_i c_{i+1} c_{i+2} + c_i c_{i+1} c_{i+2}, \end{aligned}$$

then formula (4.2) holds for $j = i + 2$.

Let formula (4.2) holds for $j - 1$ and $j - 2$, then

$$\Delta_{i \rightarrow j} = b_j \left(\Delta_{i \rightarrow j-1} - c_{j-1} \Delta_{i \rightarrow j-2} \right) + c_j \Delta_{i \rightarrow j-1}. \quad (4.15)$$

Since the formula (4.2) holds for $j - 1$, then

$$\Delta_{i \rightarrow j-1} = \prod_{k=i}^{j-1} (c_k) + \sum_{s=i}^{j-1} \left(\prod_{k=s+1}^{j-1} (c_k) \right) \left(\prod_{k=i}^s (b_k) \right), \quad (4.16)$$

then

$$c_j \Delta_{i \rightarrow j-1} = \prod_{k=i}^j (c_k) + \sum_{s=i}^{j-1} \left(\prod_{k=s+1}^j (c_k) \right) \left(\prod_{k=i}^s (b_k) \right). \quad (4.17)$$

Similarly,

$$c_{j-1} \Delta_{i \rightarrow j-2} = \prod_{k=i}^{j-1} (c_k) + \sum_{s=i}^{j-2} \left(\prod_{k=s+1}^{j-1} (c_k) \right) \left(\prod_{k=i}^s (b_k) \right). \quad (4.18)$$

Substituting from (4.16), (4.17) and (4.18) into (4.15), then we have

$$\Delta_{i \rightarrow j} = \prod_{k=i}^j (c_k) + \sum_{s=i}^j \left(\prod_{k=s+1}^j (c_k) \right) \left(\prod_{k=i}^s (b_k) \right),$$

This is confirmed the formula (4.2).

Lemma 4.2 The elements of $(I - J)^{-1} = (N_{ij})_{1 \leq i, j \leq n}$, where J defined in (4.1), can be expressed as

$$(N_{ij})_{1 \leq i, j \leq n} = \begin{cases} \frac{\Delta_{1 \rightarrow i-1, j+1 \rightarrow n}}{\Delta_{1 \rightarrow n}} \prod(c_k) & \text{for } (i \leq j), \\ \frac{\Delta_{1 \rightarrow j-1, i+1 \rightarrow n}}{\Delta_{1 \rightarrow n}} \prod(b_k) & \text{for } (i > j), \end{cases} \quad (4.19)$$

where $\Delta_{1 \rightarrow j}$ is given as in Lemma 4.1.

Proof

The proof can be easily obtained by following the proof as cf. Mallik (2001) and El-Shehawey (2008)_a.

Assume that J_1 be a Toeplitz tridiagonal matrix of (4.1),

$J_1 \equiv J(a, c; b, a, c; b, a)$, i.e the elements $a_i = a, b_i = b, c_i = c, 1 \leq i \leq n$, we get the following result

Corollary 4.1 The elements of the inverse matrix $(I - J_1)^{-1} = (N_{ij})_{1 \leq i, j \leq n}$, can be expressed as

$$(N_{ij})_{1 \leq i, j \leq n} = \begin{cases} \frac{\Delta_{1 \rightarrow i-1, j+1 \rightarrow n}}{\Delta_{1 \rightarrow n}} (c)^{j-i} & \text{for } (i \leq j), \\ \frac{\Delta_{1 \rightarrow j-1, i+1 \rightarrow n}}{\Delta_{1 \rightarrow n}} (b)^{i-j} & \text{for } (i > j), \end{cases} \quad (4.20)$$

where the continuant $\Delta_{1 \rightarrow j} = \det(I - J_1), 1 \leq i, j \leq n$, is given by:

$$\Delta_{1 \rightarrow j} = \frac{1}{b-c} [b^{j-i+2} - c^{j-i+2}], \quad (4.21)$$

where $i, j = 2, \dots, n$ in the case $a+c=1, a+b<1$; $i, j = 1, \dots, n-1$ in the case $a+c<1, a+b=1$ and $i, j = 1, \dots, n$ in the case $a+c<1, a+b<1$.

Note that the value of the continuant $\Delta_{1 \rightarrow j} = \det(I - J_1), 1 \leq i, j \leq n$ does not exist in the case $a+c=a+b=1$.

Assume that J_2 be a perturbed Toeplitz tridiagonal matrix of (4.1),

$J_2 \equiv J(a_1, c_1; b, a, c; b_n, a_n)$, i.e the "boundary" elements a_1 and a_n are different from a ; c_1 different from c ; and b_n different from b , we get the following result

Corollary 4.2 The elements of the inverse matrix $(I - J_2)^{-1} = (N_{ij})_{1 \leq i, j \leq n}$, can be expressed as

$$(N_{ij})_{1 \leq i, j \leq n} = \begin{cases} \frac{\Delta_{1 \rightarrow j-1} \Delta_{j+1 \rightarrow n}}{\Delta_{1 \rightarrow n}} (c)^{j-i} & \text{for } (i \leq j), \\ \frac{\Delta_{1 \rightarrow j-1} \Delta_{i+1 \rightarrow n}}{\Delta_{1 \rightarrow n}} (b)^{j-i} & \text{for } (i > j), \end{cases} \quad (4.22)$$

where for $i = 1, \dots, n$

$$\Delta_{1 \rightarrow i} = (1 - a_1) \Delta_{2 \rightarrow i} - b c_1 \Delta_{3 \rightarrow i}, \quad (4.23)$$

$$\Delta_{i \rightarrow n} = (1 - a_n) \Delta_{i \rightarrow n-1} - b_n c \Delta_{i \rightarrow n-2}, \quad (4.24)$$

$$\Delta_{1 \rightarrow n} = (1 - a_n) \Delta_{1 \rightarrow n-1} - b_n c \Delta_{1 \rightarrow n-2}, \quad (4.25)$$

and $\Delta_{i \rightarrow j} = \det(I - J_2), 2 \leq i, j \leq n-1$ is given as in Corollary 4.1.

5. Main results

In the dialysis model, we have an absorbing (MRW) represented by the transition probability matrix given in (3.4), following the notation of Kemeny and Snell (1976) and Iosifescu (1980), the sub-matrix Q is the $m \times m$ matrix, defined by

$$Q = \begin{pmatrix} p_1 & q_1 & 0 & & 0 \\ p_2 & 0 & q_2 & & \\ 0 & \ddots & \ddots & \ddots & \\ & \ddots & \ddots & \ddots & 0 \\ & & p_{m-1} & 0 & q_{m-1} \\ 0 & & 0 & p_m & 0 \end{pmatrix}_{m \times m}. \quad (5.1)$$

The fundamental matrix may be written as:

$$N = (I - Q)^{-1}, \quad (5.2)$$

Our main purpose of this contribution is to find closed form formulas for the mean of the occupation total, the expected time to absorption, and the absorption probability. So we first compute the elements of $(I - Q)^{-1}$, where Q is defined in (5.1), it is formulated in the next theorem.

Theorem 5.1 If Q is a non-singular matrix defined as (5.1), then the mean of the occupation total is given by

$$(N_{ij})_{1 \leq i, j \leq m} = (I - Q)^{-1} = \begin{cases} \frac{(q_1 \Delta_{2 \rightarrow j-1} \Delta_{j+1 \rightarrow m} - p_2 q_1 \Delta_{3 \rightarrow j-1} \Delta_{j+1 \rightarrow m})}{(q_1 \Delta_{2 \rightarrow m} - p_2 q_1 \Delta_{3 \rightarrow m})} \prod_{k=i}^{j-1} (q_k) & \text{for } (i \leq j), \\ \frac{(q_1 \Delta_{2 \rightarrow j-1} \Delta_{i+1 \rightarrow m} - p_2 q_1 \Delta_{3 \rightarrow j-1} \Delta_{i+1 \rightarrow m})}{(q_1 \Delta_{2 \rightarrow m} - p_2 q_1 \Delta_{3 \rightarrow m})} \prod_{k=j+1}^i (p_k), & \text{for } (i > j), \end{cases} \quad (5.3)$$

where

EL-Shehawey et al.
Stochastic Modeling Dialysis

$$\Delta_{i \rightarrow j} = \sum_{s=i-1}^i \left(\prod_{k=s+1}^j (q_k) \right) \left(\prod_{k=i}^s (p_k) \right), i, j = 2, \dots, m. \quad (5.4)$$

Proof

For the transition probability matrix (5.1), we have

$$I - Q = \begin{pmatrix} q_1 & -q_1 & 0 & & 0 \\ -p_2 & 1 & -q_2 & & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ & \ddots & \ddots & \ddots & 0 \\ 0 & & -p_{m-1} & 1 & -q_{m-1} \\ 0 & & 0 & -p_m & 1 \end{pmatrix}_{m \times m}, \quad (5.5)$$

Applying Lemma 4.2, we have

$$(N_{ij})_{1 \leq i, j \leq m} = (I - Q)^{-1} = \begin{cases} \frac{\Delta_{1 \rightarrow i-1} \Delta_{j+1 \rightarrow m}}{\Delta_{1 \rightarrow m}} \prod_{k=i}^{j-1} (q_k) & \text{for } (i \leq j), \\ \frac{\Delta_{1 \rightarrow j-1} \Delta_{i+1 \rightarrow m}}{\Delta_{1 \rightarrow m}} \prod_{k=j+1}^i (p_k) & \text{for } (i > j), \end{cases} \quad (5.6)$$

where

$$\Delta_{1 \rightarrow i-1} = q_1 \Delta_{2 \rightarrow i-1} - p_2 q_1 \Delta_{3 \rightarrow i-1}, \Delta_{1 \rightarrow m} = q_1 \Delta_{2 \rightarrow m} - p_2 q_1 \Delta_{3 \rightarrow m},$$

such that $\Delta_{i \rightarrow j}$ is given as in Lemma 4.1.

These complete the proof.

Corollary 5.1 The expected time to absorption in this model after a given number of cycles is given by:

$$m_i = \frac{1}{(q_1 \Delta_{2 \rightarrow m} - p_2 q_1 \Delta_{3 \rightarrow m})} \left[\sum_{j=1}^i (q_1 \Delta_{2 \rightarrow j-1} \Delta_{j+1 \rightarrow m} - p_2 q_1 \Delta_{3 \rightarrow j-1} \Delta_{j+1 \rightarrow m}) \prod_{k=j+1}^i (p_k) + \sum_{j=i+1}^m (q_1 \Delta_{2 \rightarrow j-1} \Delta_{j+1 \rightarrow m} - p_2 q_1 \Delta_{3 \rightarrow j-1} \Delta_{j+1 \rightarrow m}) \prod_{k=i}^{j-1} (q_k) \right]_{1 \leq i \leq m} \quad (5.7)$$

where $\Delta_{i \rightarrow j}$ is defined in (5.4).

Corollary 5.2 The probability of being absorbed at state $m+1$ after a given number of cycles is given by:

$$B_{m+1} = \left[\frac{(q_1 \Delta_{2 \rightarrow i-1} - p_2 q_1 \Delta_{3 \rightarrow i-1})}{(q_1 \Delta_{2 \rightarrow m} - p_2 q_1 \Delta_{3 \rightarrow m})} \prod_{k=i}^m (q_k) \right]_{1 \leq i \leq m}. \quad (5.8)$$

where $\Delta_{i \rightarrow j}$ is defined in (5.4).

Proof

$$B_{m+1} = [b_{m+1}]_{1 \leq i \leq m} = \sum_{k=1}^m [N_{ik} R_{km+1}]_{1 \leq i \leq m}.$$

Since $R_{km+1} = 0 \forall k \neq m$, then

EL-Shehawey et al.
Stochastic Modeling Dialysis

$$B_{m+1} = [N_{im} R_{mm+1}]_{1 \leq i \leq m} \quad (5.9)$$

Therefore, B_{m+1} becomes

$$B_{m+1} = q_m \left[\frac{(q_1 \Delta_{2 \rightarrow i-1, m+1 \rightarrow m} \Delta_{3 \rightarrow j-1, m+1 \rightarrow m} - p_2 q_1 \Delta_{3 \rightarrow j-1, m+1 \rightarrow m} \Delta_{2 \rightarrow i-1, m+1 \rightarrow m})}{(q_1 \Delta_{2 \rightarrow m} \Delta_{3 \rightarrow m} - p_2 q_1 \Delta_{3 \rightarrow m} \Delta_{2 \rightarrow m})} \prod_{k=i}^{m-1} (q_k) \right]_{1 \leq i \leq m} \quad (5.10)$$

where $\Delta_{m+1 \rightarrow m} = 1$, formula (5.8) immediately follows.

5.1 Explicit Expressions for the Mean Time to Absorption

Many results as explicit expressions can be emanating from our results as particular cases for example.

Case 1. Semi-homogeneity Case

If a semi-homogeneity is presented, i.e. $p_j = p, q_j = q, j = 2, 3, \dots, m-1$ and p_j, q_j for $j = 1, m$ are different from p, q , respectively. The corresponding result is a perturbed Toeplitz tridiagonal matrices Q such that:

$$Q = \begin{pmatrix} p_1 & q_1 & 0 & \dots & \dots & 0 \\ p & 0 & q & & & \vdots \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & p & 0 & q \\ 0 & \dots & \dots & 0 & p_m & 0 \end{pmatrix}_{m \times m} \quad (5.11)$$

we have the following corollaries.

Corollary 5.3 If Q is the $m \times m$ non-singular matrix defined as (5.11), then the mean of the occupation total becomes

$$(N_{ij})_{1 \leq i, j \leq m} = \frac{1}{(q_1 \Delta_{2 \rightarrow m} - p q_1 \Delta_{3 \rightarrow m})} \begin{cases} \Delta_{2 \rightarrow m} & \text{for } (i = j = 1), \\ \Delta_{j+1 \rightarrow m} q^{j-2} q_1 & \text{for } (i = 1, 2 \leq j \leq m-1), \\ q_1 q^{m-2} & \text{for } (i = 1, j = m), \\ (q_1 \Delta_{2 \rightarrow i-1, j+1 \rightarrow m} \Delta_{3 \rightarrow i-1, j+1 \rightarrow m} - p q_1 \Delta_{3 \rightarrow i-1, j+1 \rightarrow m} \Delta_{2 \rightarrow i-1, j+1 \rightarrow m}) q^{i-j} & \text{for } (2 \leq i \leq m-1, i \leq j \leq m), \\ (q_1 \Delta_{2 \rightarrow j-1, i+1 \rightarrow m} \Delta_{3 \rightarrow j-1, i+1 \rightarrow m} - p q_1 \Delta_{3 \rightarrow j-1, i+1 \rightarrow m} \Delta_{2 \rightarrow j-1, i+1 \rightarrow m}) p^{i-j} & \text{for } (2 \leq i \leq m-1, i > j \geq 1), \\ p_m p^{m-2} & \text{for } (i = m, j = 1), \\ (q_1 \Delta_{2 \rightarrow j-1} - p q_1 \Delta_{3 \rightarrow j-1}) p^{m-j-1} p_m & \text{for } (i = m, 2 \leq j \leq m-1), \\ (q_1 \Delta_{2 \rightarrow m-1} - p q_1 \Delta_{3 \rightarrow m-1}) & \text{for } (i = j = m), \end{cases} \quad (5.12)$$

where

$$\begin{aligned}
 \Delta_{i \rightarrow j} &= \frac{1}{p-q} [p^{i-i+2} - q^{i-i+2}], 2 \leq i \leq j \leq m-1, \\
 \Delta_{2 \rightarrow m} &= \frac{1}{p-q} [(p^{m-1} - q^{m-1}) - p_m q (p^{m-2} - q^{m-2})], \\
 \Delta_{2 \rightarrow i-1} &= \frac{1}{p-q} [(p^{i-1} - q^{i-1})], i = 3, \dots, m-1, \\
 \Delta_{j+1 \rightarrow m} &= \frac{1}{p-q} [(p^{m-j} - q^{m-j}) - p_m q (p^{m-j-1} - q^{m-j-1})], j = 1, \dots, m-1
 \end{aligned} \tag{5.13}$$

Corollary 5.4 The mean time to absorption in this semi-homogeneity model after a given number of cycles is

$$m_i = \frac{1}{(q_1 \Delta_{2 \rightarrow m} - p q_1 \Delta_{3 \rightarrow m})} \begin{cases} \Delta_{2 \rightarrow m} + q_1 q^{m-2} + q_1 \sum_{j=2}^{m-1} \Delta_{j+1 \rightarrow m} q^{j-2} & \text{for } (i=1), \\ \left[\sum_{j=2}^i (\Delta_{2 \rightarrow j-1} \Delta_{j+1 \rightarrow m} - p \Delta_{3 \rightarrow j-1} \Delta_{j+1 \rightarrow m}) p^{i-j} \right] \\ + \sum_{j=i+1}^m (\Delta_{2 \rightarrow j-1} \Delta_{j+1 \rightarrow m} - p \Delta_{3 \rightarrow j-1} \Delta_{j+1 \rightarrow m}) q^{i-j} & \text{for } (2 \leq i \leq m-1), \\ q_1 (\Delta_{2 \rightarrow m-1} - p \Delta_{3 \rightarrow m-1}) + p_m p^{m-2} + q_1 p_m \sum_{j=2}^{m-1} (\Delta_{2 \rightarrow j-1} - p \Delta_{3 \rightarrow j-1}) p^{m-i-1} & \text{for } (i=j=m), \end{cases} \tag{5.14}$$

where $\Delta_{i \rightarrow j}, \Delta_{2 \rightarrow m}, \Delta_{2 \rightarrow i-1}, \Delta_{j+1 \rightarrow m}$ defined in (5.13).

Corollary 5.5 The absorption probabilities become

$$B_{m+i} = \frac{q_m}{(q_1 \Delta_{2 \rightarrow m} - p q_1 \Delta_{3 \rightarrow m})} \begin{cases} q_1 q^{m-2} & \text{for } (i=1), \\ q_1 q^{m-i} (\Delta_{2 \rightarrow i-1} - p \Delta_{3 \rightarrow i-1}) & \text{for } (2 \leq i \leq m-1), \\ q_1 (\Delta_{2 \rightarrow m-1} - p \Delta_{3 \rightarrow m-1}) & \text{for } (i=m), \end{cases} \tag{5.15}$$

where $\Delta_{i \rightarrow j}, \Delta_{2 \rightarrow m}, \Delta_{2 \rightarrow i-1}, \Delta_{j+1 \rightarrow m}$ defined in (5.13).

Case 2. Homogeneous Case

If homogeneity is presented, i.e. $p_j = p, q_j = q, j = 1, 2, 3, \dots, m$. The transition probability Q corresponding to the following Toeplitz tridiagonal matrix:

EL-Shehawey et al.

Stochastic Modeling Dialysis

$$Q = \begin{pmatrix} p & q & 0 & \dots & \dots & 0 \\ p & 0 & q & & & \vdots \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & q \\ 0 & \dots & \dots & 0 & p & 0 \end{pmatrix}_{m \times m}, \quad (5.16)$$

the following corollaries follow.

Corollary 5.6 If Q is the matrix defined as (5.16), then the mean of the occupation total becomes

if $p \neq q, u = \frac{p}{q}$:

$$(N_{ij})_{1 \leq i, j \leq m} = \frac{1}{q(u-1)} \begin{cases} u^{m-j+1} - 1 & \text{for } i \leq j, \\ u^{i-j} (u^{m-i+1} - 1) & \text{for } j < i, \end{cases} \quad (5.17)$$

if $p = q = \frac{1}{2}$:

$$(N_{ij})_{1 \leq i, j \leq m} = \begin{cases} 2(m-j+1) & \text{for } i \leq j, \\ 2(m-i+1) & \text{for } j < i. \end{cases} \quad (5.17a)$$

Corollary 5.7 The mean time to absorption in the homogeneity model is

$$m_1 = \begin{cases} \frac{1}{q(u-1)^2} [u^{m+1} - (u^i + u + 1)]_{1 \leq i \leq m} & \text{if } p \neq q, \\ [(m-i+1)(m+i)]_{1 \leq i \leq m} & \text{if } p = q = \frac{1}{2}. \end{cases} \quad (5.18)$$

These results agree with that of Kastenbaum (1966) "see also El-Shehawey (2000), (2008)_a, (2008)_b".

Note that. The absorption probability in this homogeneous case is one.

References

- [1] Baigent S., Unwin R., and Yeng C. C., (2001) Mathematical modeling of profiled hemodialysis: a simplified approach, *Journal of Theoretical Medicine*; 3: 143-160.
- [2] Cox, D. R. and Miller, H. D. (1965) *The Theory of Stochastic Processes*. Chapman and Hall, London.
- [3] Cottle R. W., (1974) Manifestations of Schur complement, *Linear Algebra Appl.* 8: 189-211.
- [4] Craig L. C. and King, T. P. (1955) Some dialysis experiments with polypeptides, *J. Amer. Chem. Soc.* 77, 6620-6624.
- [5] Craig L. C. and King, T. P. (1956) Fractional dialysis with cellophane membranes, *J. Amer. Chem. Soc.* 78, 4171- 4172.

El-Shehawey et al.
Stochastic Modeling Dialysis

- [6] Craig L. C, King, T. P. and Stracher A. (1957) Dialysis studies. II. Some experiments dealing with the problem of selectivity. J. Amer. Chem. Soc. 79, 3729- 3737.
- [7] Davies R. (1978) A Study of a renal unit, J. Opl Res. Soc., 30, 873- 884.
- [8] Davies R. (1985) An assessment of models of a health system, J. Opl Res. Soc. 36, 679- 687.
- [9] El_Shehawey, M. A. (1994) On The Frequency Count For a Random Walk With Absorbing Boundaries: A Carcinogenesis Example I. J. Phys. A: Math. Gen. Vol. 27, 7035-7046.
- [10] El_Shehawey, M. A. (2000) Absorption Probabilities for a Random Walk Between Two Partially Absorbing Boundaries: I. J. Phys. A: Math. Gen. Vol. 33, 9005-9013.
- [11] El_Shehawey, M. A. (2008), On Inverses of Tridiagonal Matrix Arising From Markov Chain Random Walk I, SIAM J. Matrix Anal. Appl. V. 30 N. 2, P. 497-508.
- [12] El_Shehawey, M. A. (2008), Trinomial Random Walk With One Or Two Imperfect Absorbing Barriers, Math. Slovaca Vol. 58 N. 3, P. 353-376.
- [13] El-Shehawey M. A., El-Shreef Gh. A., and Al-Henawy A. Sh., (2008) Analytical inversion of general periodic tridiagonal matrices. J. Math Anal; 345: 123-134.
- [14] Feller, W. (1968) An Introduction to Probability Theory and Its Application. Vol. I. (third edition) John Wiley & Sons, New York.
- [15] Glicksman M.E., (2000) Diffusion in Solids, Wiley, New York, p. 391.
- [16] Grimmett, G. R. and Stirzaker, D. R. (1987) Probability and Random Processes. Oxford Science Publications, Oxford.
- [17] Goldman L., and Ausiello D., (2007) Cecil Textbook of Medicine, 23rd Philadelphia, Pa: Saunders Elsevier.
- [18] Hardin J. C. and Sweet A. L. (1970) Moments of the Time to Absorption in the Random Walk Between a Reflecting and an Absorbing Barrier, Siam Review, V. 12 N.1, 140- 142.
- [19] Horn R. A., and Johnson C. R., (1994) Topic in Matrix Analysis, Cambridge University Press, Cambridge.
- [20] Iosifescu, M. (1980) Finite Markov Processes and Their Applications. John Wiley and Sons, Chichester, New York.
- [21] Kastenbaum, M. A. (1960) The separation of molecular compounds by countercurrent dialysis: a stochastic process, Biometrika; 47: 69-77.
- [22] Kastenbaum, M. A. (1966) A dialysis system with one absorbing and one semi-reflecting state, J. Appl. Prob. 3, 265- 374.
- [23] Kemeny, J. G. and Snell, J. L. (1976) Finite Markov Chains. Princeton, NJ: Van Nostrand-Reinhold.
- [24] Kemperman, J. H. B. (1961) On a dialysis problem, Technique Summary Report No. 235, Mathematics Research Center, U. S. Army, Univ. of Wisconsin, Madison, Wisconsin.
- [25] Mallik, R. K. (2001) The Inverse of a Tridiagonal Matrix. Linear Algebra Appl. Vol. 325, 109-139.
- [26] Neuts M. F. (1963) Absorption Probabilities for a Random Walk Between a Reflecting and An Absorbing Barrier. Bull. Soc. Math. Belg. 15, 253- 258.
- [27] Neuts M. F. (1964) General Transition Probabilities For Finite Markov Chains. Proc. Camb. Phil. Soc., 60, 83- 91.

EL-Shehawey et al.

Stochastic Modeling Dialysis

- [28] Parzen, E. (1962) Stochastic Process, Holden-Day, Inc., London.
- [29] Percus, O. E. (1985) Phase Transition in One- Dimensional Random Walk With Partially Reflecting Boundaries, Adv. Appl. Prob. 17, 594- 606.
- [30] Rudolph, G. (1999) The fundamental matrix of the general random walk with absorbing boundaries. Technical report of the collaborative research center "Computational Intelligence", university of Dortmund. 1-5.
- [31] Rudolph, G. (2001) Takeover times of noisy Non-Generational selection rules that undo extinction. Springer, Vienna. 268-271.
- [32] Srinivasan, S. R. and Mehata, K. M. (1978) Stochastic Processes. Tata, McGraw-Hill, New Delhi.
- [33] Watterson G. A. (1961) Markov Chain With Absorbing States: A Genetic Example. Ann. Math. Stats. 32, 716- 729.
- [34] Weesakul B. (1961) The Random Walk Between a Reflecting and an Absorbing Barrier. Ann. Math. Stat. 32, 765- 769.

المخلص العربي الخاص بالبحث

تطبيق للمشى العشوائي لماركوف علي نظام الغسيل الكلوي

An Application of Markov Random Walk to a Dialysis System

د. محمد أحمد أنور الشهلاوي د. منصور فتحي ياسين علي نجوي جمال الدين رجب

في هذا البحث تم تناول تطبيق بيولوجي للمشى العشوائي لماركوف $\{X_n, n=0,1,2, \dots\}$ بفضاء عينة $\{1,2,\dots,m+1\}$ ذات طرفين إحداهما شبه عاكس عند المرحلة الأولى والأخر ماص عند المرحلة الأخيرة $m+1$. هذا التطبيق البيولوجي ينشأ من نظام التيار المضاد للغسيل الكلوي، حيث تم عرض صور واضحة وصريحة وجديدة (لم تغطي بالبحث من قبل) وذلك للوسط الحسابي لزمن انشغال الحالة قبل الامتصاص والقيمة المتوقعة لاحتمالات الامتصاص بمعلومية حالة البداية، وكذلك احتمالات الامتصاص التي تتناظر هذا النظام.

يلاحظ أن نموذج المشى العشوائي تحت الدراسة تم مناقشته في كثير من المراجع والابحاث في حالات خاصة قديما وحديثا، وهو يناظر سلسلة خطية من الخلايا، كل خلية غسيل كلوي محددة بمرحلة وكل فترة غسيل كلوي محددة بدورة. بعد كل دورة الحلقة الخارجية في كل مرحلة تستعمل كحلقة داخلية في المرحلة القادمة و الحلقة الداخلية في كل مرحلة تعاد إلي كيس الغسيل الكلوي في المرحلة السابقة، الحلقة الداخلية في المرحلة الأولى محتفظة بالمرحلة الأولى والحلقة الخارجية في المرحلة النهائية تكون خارج النظام. حيث تمثل المرحلة الأولى حد شبه عاكس والمرحلة الأخيرة حد ماص. كما تناولنا بعض الحالات الخاصة التي تمدنا بصيغ واضحة وصريحة وقابلة للتطبيق.

ON MARKOV CHAIN ROULETTE TYPE GAME

El-Shehawey, M. A. and Al-Shreef, Gh. A.

Math. Dept., Damietta Faculty of Science, Mansoura University

"Damietta Branch", P.O. Box 6, New Damietta, Egypt.

E-mail address: el_shehawey@mans.edu.eg

Abstract: A Markov chain on non-negative integers which arises in the roulette - type game is discussed. The transition probabilities are $p_{01} = \rho$, $p_{Nj} = \delta_{Nj}$, $p_{i,i+W} = q$, $p_{i,i-1} = p = 1 - q$, $1 \leq W < N$, $0 \leq \rho \leq 1$, $N - W < j \leq N$ and $i = 1, 2, \dots, N - W$. Using formulae for determinant of a partitioned matrix, closed form expression for the solution of the Markov chain roulette type game is deduced.

Keywords: Gambler's ruin; Game of roulette; Markov chain; Partition matrix.

AMS Classif. 05.40.Fb, 02.50.Cw, 02.50.Ga, 02.50.Le.

1. Introduction

The classical two- player gambler's ruin problem, in which one unit changes hands until one player goes bankrupt, is well known (cf. Feller [9], chapter XIV.8, Srinivasan and Mehata [17] chapter three, Percus [15], El-Shehawey [8], and Lefebvre [14]), has a very simple, closed form solution, in the cited literature and also referred to in Corollary 2 of the present paper in the case $\rho = 0$. There are numerous applications of this type of problem, and has attracted some attention in the seventies in the context of cancer growth modeling (cf. Bell [1] and Beyer and Waterman [2]). In particular, there was significant interest in estimating the expected time for a cancerous clone of cells to reach tumor size from a single way-ward cell, rather than dying off before reaching tumor size. The analysis can help to determine the appearance of the first wayward cancerous cell, and perhaps to identify reasons for the cancerous growth (cf. El-Shehawey [4] and references therein). Here, we generalize the classical problem by introducing the following assumptions:

- (1) The number of units the gambler may win in each single game is W , $1 \leq W < N$,
- (2) The initial capital of the gambler is i units, $0 \leq i \leq N - W$,
- (3) The ultimate fortune which the gambler wants to accumulate is N units,
- (4) In each single game the gambler has a probability q of winning W units from his adversary, and probability $p = 1 - q$ of losing one unit to his adversary,
- (5) If the gambler fortune drops down to 0 units he is allowed with probability ρ , $0 \leq \rho \leq 1$, to start the game again with one unit,
- (6) After a series of single games, the game ends
 - i-with probability $1 - \rho$ when the gambler goes broke (loses all his fortune), or
 - ii-when he reaches his objective (wins all his adversary fortune).

In the language of Markov chain random walk- that is, $\{X_n, n = 0, 1, 2, \dots\}$ whose state space is $\{0, 1, 2, \dots, N\}$ with transition probabilities

$$p_{i,i+W} = q, p_{i,i-1} = p = 1 - q \quad \text{for } i \in \{1, 2, \dots, N - W\}.$$

When the particle hits the origin absorption (ruin) partially occurs with probability $1-\rho$ or reflects to the state one with probability ρ , $0 \leq \rho \leq 1$. At the boundary state j , $N-W < j \leq N$ there is an absorbing barrier.

We look for the probability the gambler loses all his fortune, that is the probability of ruin partially

$$P(i) := \Pr\{\tau_0 < \tau_{N-W} | X_0 = i\} \quad \text{for } i \in \{1, 2, \dots, N-W\}$$

where

$$\tau_0 := \inf\{n \geq 0 : X_n = 0\} \quad \text{and} \quad \tau_{N-W} := \inf\{n \geq 0 : X_{N-W} = 0\}.$$

In the present paper, using formulae for determinant of a partitioned matrix, we seek a closed form solution to this generalized roulette-type game, i.e. We compute the probability of ruin partially. This problem is solved in several special cases (cf. Feller [9] p. 352, Hill and Gulati [10], Percus [15] and El-Shehawey [5, 8]).

Using generating functions and Lagrange's theorem for the expansion of a function as a power series, El-Shehawey [6] deduced an explicit expression for the probability of the player's capital at the n -th step with the assumption that the adversary is infinitely rich ($N \rightarrow \infty$) (see also Hill and Gulati [10]).

Kozek [13] has derived similar formulae for the complementary probability of ruin,

$$q(i) := \Pr\{\tau_{N-W} < \tau_0 | X_0 = i\}, \quad \text{in the particular case when } \rho = 0.$$

2. Closed form solution for Markov chain roulette-type game

Let $P(i)$ be the probability of ruin partially for a gambler starting with i units, $0 \leq i \leq N-W$, with the probability p of losing a unit and $q = 1-p$ of winning W units and playing against an adversary with initial capital $N-i$ units. The probability $P(i)$ satisfies the following difference equation

$$P(i) = pP(i-1) + qP(i+W), \quad 1 \leq i \leq N-W \quad (2.1)$$

with the boundary conditions

$$\begin{cases} P(0) = \rho P(1) + (1-\rho), \\ P(i) = 0, \end{cases} \quad N-W < i \leq A \quad (2.2)$$

The equation (2.1) and (2.2) can be expressed as the following system of $N-W$ equations:

$$\begin{aligned} pP(1) - qP(W+1) &= p(\rho P(1) + (1-\rho)) & \text{at } i=1 \\ -pP(1) - P(2) - qP(W+2) &= 0 & \text{at } i=2 \\ -pP(2) + P(3) - qP(W+3) &= 0 & \text{at } i=3 \\ &\vdots & \\ -pP(N-1-W) + P(N-W) - qP(N-W) &= 0 & \text{at } i=N-W \end{aligned} \quad (2.3)$$

$$\begin{aligned}
 -pP(N-2W)+P(N-2W+1) &= 0 & \text{at } i = N-2W+1 \\
 -pP(N-2W+1)+P(N-2W+2) &= 0 & \text{at } i = N-2W+2 \\
 &\vdots \\
 -pP(N-W-1)+P(N-W) &= 0 & \text{at } i = N-W
 \end{aligned}$$

This system can be written in matrix form as

$$P_{N-W}Y_{N-W} = R_{N-W}, \quad (2.4)$$

where P_{N-W} is a band matrix of order $N-W$, defined by

$$P_{N-W} = \begin{matrix} & \begin{matrix} 1 & 2 & \dots & W+1 & \dots & N-W \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ \vdots \\ N-2W \\ \vdots \\ N-W \end{matrix} & \begin{pmatrix} 1 & 0 & \dots & 0 & -q & 0 & \dots & 0 \\ -p & 1 & 0 & & -q & & & \vdots \\ 0 & -p & \ddots & & & & & 0 \\ & \ddots & \ddots & \ddots & & & & -q \\ \vdots & & \ddots & \ddots & \ddots & & & 0 \\ & \vdots & & \ddots & \ddots & \ddots & & \vdots \\ & & & & -p & 1 & 0 & 0 \\ 0 & & & & 0 & -p & 1 & 1 \end{pmatrix} \end{matrix}, \quad (2.5)$$

With the elements of the matrix $P_{N-W} = (P_{N-W}(i, j))$ is

$$P_{N-W}(i, j) = \begin{cases} 1, & i = j, j = 1, \dots, N-W \\ -p, & i = j+1, j = 1, \dots, N-W-1 \\ -q, & i = j-W, j = W+1, \dots, N-W \\ 0 & \text{otherwise.} \end{cases}$$

Y_{N-W} is a column vector of probabilities $P(i)$, $i = 1, 2, \dots, N-W$, and R_{N-W} is a column vector with the first element equal to $p(\rho P(1) + (1-\rho))$ and last $N-W-1$ elements equal to zero.

Denote

$$\Delta_{N-W} = |P_{N-W}| \quad (2.6)$$

Lemma 1. Let A be an invertible matrix of order n , partitioned as

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

in which A_{22} is non-singular. Then the determinant

$$\det A = \det A_{22} \cdot \det (A_{11} - A_{12}A_{22}^{-1}A_{21}). \quad (2.7)$$

The matrix with the form of $A_{11} - A_{12}A_{22}^{-1}A_{21}$ is called a Schur complement of A_{22} in A .

An excellent review of Schur complement and their applications is given by Horn and Johnson [11], Section 07.3, p. 18, (see also El-Shehawey et al. (2008)).

Lemma 2. The determinant Δ_{N-W} of the matrix P_{N-W} obeys the following recursive relation

$$\Delta_{N-W} = \Delta_{N-W-1} - qp^W \Delta_{N-2W-1} \quad (2.8)$$

with initial conditions

$$\Delta_1 = \Delta_2 = \dots = \Delta_W = 1. \quad (2.9)$$

Proof. See appendix.

Lemma 3. The determinant $\Delta_{N-W} = P_{N-W}$ can be expressed in the following explicit form

$$\Delta_{N-W} = \sum_{i=0}^{[(N-W-1)/(W+1)]} (-1)^i \binom{N-(i+1)W-1}{i} (qp^W)^i \quad (2.10)$$

where $[x]$ denotes the greatest integer not exceeding x .

Proof. See appendix.

Lemmas 2 and 3, with appropriate change of notation, agree with that of Kozek [13].

2.1. Main result. The main purpose of this contribution is to find closed form solution for the Markov chain roulette- type game, i.e., solving the system of equations (2.4). It is formulated in the next theorem:

Theorem (1)

The probability $P(i)$ takes the form

$$P(i) = \begin{cases} \rho P(1) + (1-\rho) & i = 0 \\ p^i (\rho P(1) + (1-\rho)) \frac{\Delta_{N-W-1}}{\Delta_{N-W}} & 1 < i \leq N-W \\ 0 & N-W < i \leq N \end{cases}, \quad (2.11)$$

where

$$P(1) = \frac{p(1-\rho)\Delta_{N-W-1}}{\Delta_{N-W} - \rho p \Delta_{N-W-1}}, \quad (2.12)$$

and Δ_{N-W} given by (2.10)

Proof. (by induction)

Let $P_{N-W}^{(i)}$ denote the matrix obtained from the matrix P_{N-W} by replacing its i^{th} column by the vector R_{N-W} . We denote by $\Delta_{N-W}^{(i)}$ the determinant of the matrix $P_{N-W}^{(i)}$.

Basis steps. When $i = 1$, partition the matrix $P_{N-W}^{(1)}$ by the first row and the first column in the form

$$P_{N-W}^{(1)} = \begin{pmatrix} 1 & p(\rho P(1) + (1-\rho)) & 0 & \dots & 0 & -q & 0 & \dots & 0 \\ 2 & 0 & 1 & 0 & & -q & & & \\ \vdots & \vdots & -p & \ddots & & & & & 0 \\ \vdots & \vdots & & \ddots & \ddots & & & & -q \\ \vdots & \vdots & & & \ddots & \ddots & & & 0 \\ \vdots & \vdots & & & & \ddots & \ddots & & \vdots \\ N-W & 0 & \dots & \dots & & -p & 1 & 0 \\ & & & & & 0 & -p & 1 \end{pmatrix} \quad (2.13)$$

Applying Lemma 1, we get

$$\Delta_{N-W}^{(1)} = p(\rho P(1) + (1-\rho)) \Delta_{N-W-1}. \quad (2.14)$$

Dividing both sides by Δ_{N-W} , we have

$$\frac{\Delta_{N-W}^{(1)}}{\Delta_{N-W}} = p(\rho P(1) + (1-\rho)) \frac{\Delta_{N-W-1}}{\Delta_{N-W}}. \quad (2.15)$$

Using Cramer's rule, we obtain

$$P(1) = \frac{\Delta_{N-W}^{(1)}}{\Delta_{N-W}}. \quad (2.16)$$

Hence the theorem is true for $i = 1$.

Substituting (2.16) into (2.15), we get

$$\left(1 - \rho p \frac{\Delta_{N-W-1}}{\Delta_{N-W}}\right) P(1) = p(1-\rho) \frac{\Delta_{N-W-1}}{\Delta_{N-W}}. \quad (2.17)$$

Thus formula (2.12) follows.

Now let $i = 2$. Partition the matrix $P_{N-W}^{(2)}$ by the first two rows and the first two columns in the form

$$P_{N-W}^{(2)} = \begin{pmatrix} 1 & 1 & p(\rho P(1) + (1-\rho)) & 0 & \dots & 0 & -q & 0 & \dots & 0 \\ 2 & -p & 0 & 1 & 0 & & -q & & & \\ \vdots & 0 & \vdots & -p & \ddots & & & & & 0 \\ \vdots & \vdots & \vdots & & \ddots & \ddots & & & & -q \\ \vdots & \vdots & \vdots & & & \ddots & \ddots & & & 0 \\ \vdots & \vdots & \vdots & & & & \ddots & \ddots & & \vdots \\ N-W & 0 & 0 & \dots & \dots & & -p & 1 & 0 \\ & & & & & & 0 & -p & 1 \end{pmatrix} \quad (2.18)$$

Applying Lemma 1, we obtain

$$\Delta_{N-W}^{(2)} = p^2(\rho P(1) + (1-\rho)) \Delta_{N-W-2}. \quad (2.19)$$

Dividing both sides by Δ_{N-W} , we have

$$\frac{\Delta_{N-W}^{(2)}}{\Delta_{N-W}} = p^2 (\rho P(1) + (1-\rho)) \frac{\Delta_{N-W-2}}{\Delta_{N-W}}. \quad (2.20)$$

Using Cramer's rule, we obtain

$$P_2 = \frac{\Delta_{N-W}^{(2)}}{\Delta_{N-W}}. \quad (2.21)$$

Then the theorem is true for $i = 2$ as well.

Induction step. Suppose that the formula (2.11) is true, in the case $1 < i \leq N - W$, for $i = j - 1$ and $i = j$, and we try to prove it's holds at $j + W$, i.e. Let

$$P(j-1) = p^{j-1} (\rho P(1) + (1-\rho)) \frac{\Delta_{N-W-j+1}}{\Delta_{N-W}}, \quad (2.22)$$

$$P(j) = p^j (\rho P(1) + (1-\rho)) \frac{\Delta_{N-W-j}}{\Delta_{N-W}}. \quad (2.23)$$

From (2.22), (2.23), and the difference equation (2.1) with $i = j$, it is easy to verify that

$$qP(j+W) = \frac{p^j (\rho P(1) + (1-\rho))}{\Delta_{N-W}} \{ \Delta_{N-W-j} - \Delta_{N-W-j+1} \}. \quad (2.24)$$

Using lemma 2, formula (2.24) finally becomes

$$P(j+W) = p^{j+W} (\rho P(1) + (1-\rho)) \frac{\Delta_{N-2W-j}}{\Delta_{N-W}}. \quad (2.25)$$

Hence the theorem holds for $i = j + W$. This completes the proof.

The following two corollaries are simple consequence of Theorem 1. In Corollary 1. The game terminates when the player loses all his fortune or when his capital reaches, or exceeds level $N - W$.

Corollary 1. Let $P(i)$ be the probability of ruin ($\rho = 0$) for a gambler with an initial capital of i units, $i = 0, 1, 2, \dots, N$, with the probability p of losing a unit and $q = 1 - p$ of winning W units and playing against an adversary with initial capital $N - i$ units. Then we have

$$P(i) = \begin{cases} 1 & i = 0 \\ p^i \frac{\Delta_{N-W-i}}{\Delta_{N-W}} & 1 < i \leq N - W \\ 0 & N - W < i \leq N \end{cases}, \quad (2.26)$$

And Δ_{N-W} given by (2.10).

The complementary probability of ruin is

$$q(i) = 1 - P(i) \quad \text{for } i = 0, 1, 2, \dots, N \quad (2.27)$$

Where $P(i)$ is given by (2.26).

Formula (2.27), with the appropriate change of notation, agrees with that of Kozek [13].

Corollary 2. Let $P(i)$ be the probability of ruin partially ($W=1$) or the classical gambler's ruin in which the gambler loses or wins one unit with the probabilities p and $q=1-p$, respectively. If the gambler's initial is i units, $i=0, 1, 2, \dots, N$, and he plays against an adversary with initial capital $N-i$ units. Then we get

$$P(i) = \begin{cases} \frac{(1-\rho) \left[\left(\frac{p}{q} \right)^i - \left(\frac{p}{q} \right)^N \right]}{1 - \rho \frac{p}{q} - (1-\rho) \left(\frac{p}{q} \right)^N} & p \neq q \\ \frac{N-i}{N + \frac{\rho}{1-\rho}} & p = q \end{cases} \quad (2.28)$$

Proof. Using Lemma 3, with $W=1$, it is easy to verify that

$$\Delta_{N-1} = q^{N-1} \sum_{s=0}^{N-1} \left(\frac{p}{q} \right)^s = q^{N-1} \frac{1 - \left(\frac{p}{q} \right)^N}{1 - \frac{p}{q}}, \quad (2.29)$$

$$\Delta_{N-1-1} = q^{N-1-1} \frac{1 - \left(\frac{p}{q} \right)^{N-1}}{1 - \frac{p}{q}} \quad \text{and} \quad \Delta_{N-2} = q^{N-2} \frac{1 - \left(\frac{p}{q} \right)^{N-1}}{1 - \frac{p}{q}} \quad (2.30)$$

Substituting from (2.29) and (2.30) into Theorem 1, with $W=1$, after simplification, formula (2.28) follows in the case $p \neq q$. The case $p=q$ immediately follows from the case $p \neq q$ after taking the limit as $\frac{p}{q} \rightarrow 1$.

It is straightforward to show that the present results are generalized some well known results, see for example, formula (2.28), with appropriate change of notation, agrees with that of El-Shehawey [5] (see also Percus [15]), agrees in the case $\rho=0, N \rightarrow \infty$ with that of El-Shehawey [6] and agrees in the case $\rho=0$ with that of Lefebvre [14] (see also Feller [9], p. 345, and El-Shehawey [4]).

Conclusion. In this paper, we have considered the so-called Markov chain roulette-type game that generalizes the classical gambler's ruin problem. We derived exact and explicit expression for the probability ruin partially. It has been derived by means of elementary partitioned matrix. The present formula enables us to find many interesting particular cases through as appropriate choice of the boundary probability. The current results are especially useful in various applications in countercurrent dialysis, tumor growth, multigene evolution, queuing, war with bargaining, ... (cf. Kastenbaum [12], Bell [1], Beyer and Waterman [2], El-shehawey [4], Bohm [3], and Smith and Stam [16]).

3. Appendix

Proof 2. Expanding the determinant Δ_{N-W} of the matrix given in (2.5) by the elements of the first row, we obtain

$$\Delta_{N-W} = \Delta_{N-W-1} + (-1)^W (-q) \begin{vmatrix} -p & 1 & 0 & \dots & -q & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & -p & \dots & -q & 0 & \dots & 0 \\ 0 & \dots & \dots & 0 & 1 & 0 & \dots & 0 & -q & 0 & \vdots \\ \vdots & \vdots & \vdots & \vdots & -p & 1 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & 0 & -p & \ddots & \ddots & \ddots & \ddots & -q \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 0 & \dots & 0 & -p & 1 \end{vmatrix} \quad (3.1)$$

Applying Lemma 1, we get

$$\Delta_{N-W} = \Delta_{N-W-1} + (-1)^W (-q) \begin{vmatrix} 1 & 0 & \dots & 0 & -q & 0 & \dots & 0 \\ -p & 1 & 0 & \dots & -q & \vdots & \vdots & \vdots \\ 0 & -p & 1 & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & -q \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & -p & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \\ 0 & \dots & \dots & 0 & 0 & \dots & 0 & -p & 1 \end{vmatrix} \quad (3.2)$$

and formula (2.28) immediately follows.

Proof 3. (by induction)

Using Lemma 2, it is clear that for $N-W=1$, $\Delta_1=1$.

Assume that Lemma 3 is valid for $N-W-1$, $N-W-2$, ..., $N-2W-1$. For every $N-W \geq 1$. We will show that this implies its validity for $N-W$. From formula (2.8) we have

$$\begin{aligned} \Delta_{N-W} &= \sum_{i=0}^{[(N-W-2)/(W+1)]} (-1)^i \binom{N-(i+1)W-2}{i} (qp^W)^i - qp^W \sum_{i=0}^{[(N-2W-2)/(W+1)]} (-1)^i \binom{N-2W-iW-2}{i} (qp^W)^i \\ &= \sum_{i=0}^{[(N-W-2)/(W+1)]} (-1)^i \binom{N-(i+1)W-2}{i} (qp^W)^i + \sum_{i=0}^{[(N-W-1)/(W+1)-1]} (-1)^{i+1} \binom{N-2W-iW-2}{i} (qp^W)^{i+1} \quad (3.3) \\ &= 1 + \sum_{i=1}^{[(N-W-2)/(W+1)]} (-1)^i \binom{N-(i+1)W-2}{i} (qp^W)^i + \sum_{i=1}^{[(N-W-1)/(W+1)]} (-1)^i \binom{N-W-iW-2}{i-1} (qp^W)^i. \end{aligned}$$

and formula (2.10) immediately follows if the integer parts of $(N-W-1)/(W+1)$ and $(N-W-2)/(W+1)$ are equal, using the fact

$$\binom{N-(i+1)W-2}{i} + \binom{N-(i+1)W-2}{i-1} = \binom{N-(i+1)W-1}{i}$$

Note that there is one, last term in the last sum in (c), not matched, if the former integer is strictly smaller than the latter one, i.e.

$$\text{i.e. } \left\lfloor \frac{A-W-2}{W+1} \right\rfloor < \left\lfloor \frac{A-W-1}{W+1} \right\rfloor \quad \text{or} \quad \left\lfloor \frac{A-W-1}{W+1} - \frac{1}{W+1} \right\rfloor < \left\lfloor \frac{A-W-1}{W+1} \right\rfloor.$$

This happens only when $N-W-1$ is a multiple of $W+1$, say $N-W = k(W+1)+1$. Then the last value of index i in the last sum in (c) equal k . Hence, the coefficient of $(qp^W)^k$ is $(-1)^k \binom{k(W+1)-1-kW}{k-1} = (-1)^k \binom{k-1}{k-1} = (-1)^k$, and the not-matched term of the last sum in (c) coincide with the last term of (2.10). Consequently formula (2.10) holds true.

References

- [1] Bell G. I., 1976. *Models of carcinogenesis as an escape from mitotic inhibitors*, Science 192 (4239) 569- 572.
- [2] Beyer, W. A., and Waterman, M. S., 1977. *Symmetries for conditional ruin problem*. Math. Mag. 50 (1), 42-45.
- [3] Bohm, W., 1994. *On discrete time Markovian N- policy queues involving bathes*, Sankya A, 56 N1, 144- 163.
- [4] El-Shehawey, M. A., 1994. *The frequency count for a random walk with absorbing boundaries: a carcinogenesis example. I*, J. Phys. A: Math. Gen. Vol. 27, 7035-7046.
- [5] El-Shehawey, M. A., 2000. *Absorption probabilities for a random walk between two partially absorbing boundaries: I*, J. Phys. A: Math. Gen. 33, 9005- 9013.
- [6] El-Shehawey, M. A., 2002. *A semi-infinite random walk associated with the game of roulette*, J. Phys. A: Math. Gen. Vol. 35, 1813-1820.
- [7] El-Shehawey, M. A., El- Shreef Gh. A., and Al- Henawy, A. Sh. 2008. *Analytical inversion of general periodic tridiagonal matrices*, J. Math. Anal. Appl. 345, 123-134.
- [8] El-Shehawey, M. A., 2008. *Tridiagonal random walk with one or two imperfect absorbing barriers*. Math. Slovaca. 58, 353- 376.
- [9] Feller, W., 1968. *An Introduction to Probability Theory and Its Applications*, 3rd.vol. 1. Wiley, New York.
- [10] Hill, J. M. and Gulati, C. M. 1981. *The random walk associated with the game of roulette*, J. Appl. Prob. Vol.18, 931-936.
- [11] Horn, R. A., and Johnson, C. R., 1994. *Topics in Matrix Analysis*, Cambridge University Press, Cambridge.
- [12] Kastenbaum, M. A., 1966. *A Dialysis system with one absorbing and one semi- reflecting state*, J. Appl. Prob. 3, 365- 374.
- [13] Kozek, A. S., 2002. *Solution of a roulette type ruin problem a correction to A rule of thumb(not only) for gambler's*, Stochastic Process. Appl. Vol. 100, 301-311.
- [14] Lefebvre, M., 2008. *The gambler's ruin problem for a Markov chain related to the Bessel process*. Statistics and Probability Letters. 78, 2314- 2320.
- [15] Percus, O. E., 1985. *Phase transition in one- dimensional random walk with partially reflecting boundaries*. Adv. Appl. Prob. 17, 594- 606.
- [16] Smith, A. and Stam, A. C., 2004. *Bargaining and the nature of war*, J. Conflict Resolution, V. 48 N-6, 783- 813
- [17] Srinivasan, S. K., and Mehata, K., M., 1978. *Stochastic Processes*, Mc Graw Hill, New Delhi.

دراسة للعبة الروليت كسلسلة ماركوف

"ON MARKOV CHAIN ROULETTE TYPE GAME"

في هذا البحث تناولنا دراسة:

نموذج المشي عشوائي في بعد واحد كتطبيق للعبة الروليت في وجود:

(i) حاجزين أحدهما عند الموضع 0 وهو حاجز ماص جزئياً بينما الحاجز الآخر عند الموضع j حيث $A - W < j \leq A$ تام الامتصاص.

(ii) الخطوات غير متساوية الطول حيث عند الحالة i يتحرك إلى الحالة $i + w$ أو إلى الحالة $i - 1$ باحتمالات p, q على الترتيب

حيث تم:

استنباط صور عامة تحليلية دقيقة لاحصاءات هذا النموذج (لم تغطي بالبحث من قبل إلا في حالات خاصة) حيث تم استخدام عدة طرق لحل نظام المعادلات المناظر لهذه النماذج وهي طريقة كرامر (Cramer) وتقنية كوزك (Kozek).

لقد تم التأكد من صحة النتائج بمقارنة الحالات الخاصة من نتائجنا مع نتائج

.Kozek(2002), El-Shehawey(2000).

E-OPTMAL DESIGN FOR DETECTING THE MAIN EFFECTS IN 2^k -FACTORIAL PAIRED COMPARISON EXPERIMENTS WITH FULL PROFILES

Essam A. Ahmed, Sayed R. El-Shair and Osama F. Abdelhalim

Department of Mathematics,
Insurance & Applied Statistics,
Helwan University

ABSTRACT

In this paper we construct the form of the E-optimal paired comparison design when there are k factors, each with two levels in a factorial experiment for detecting the main effects in case that all pairs with the same number of factors different appear equally often. So, paired comparisons are used in which full profiles of the treatments are presented. In this setting the E-optimal design has pairs in which all factors are different, this pairs called the fold-over pairs.

Key Words: paired comparisons; optimal designs; factorial treatments; full profile.

1. Introduction

El-Helbawy & Bradley (1978), determined optimal 2^3 paired comparison designs for testing whether effect of any two factor interaction is null, and whether all effects of the interactions involving specified single factor are null. In El-Helbawy & Ahmed (1984), optimal 2^k paired comparison designs for four situations are given. These situations are:

1. for determining if the effect of one m factor interaction is null.
2. for determining if the effects of m interactions involving specified single factor are null.

3. for determining if the effects of m odd order interactions are null.
4. for determining if the effects of m higher order interactions are null.

In each of these four situations the possible pairs are divided into two classes, the classes depending on the situation. For instance, for the first situation, one class of pairs consists of those pairs in which only one treatment in the pair has an even number of the factor listed in the m factor interaction at the high level. The other class of pairs consists of the remaining pairs (so those where both treatments have an even number of the factors listed in the m factor interaction at the high level and those where neither treatment has an even number of the factors listed in the m factor interaction at the high level). El-helbawy et al. (1994), extended the cases of El-Helbawy & Ahmed (1984) to asymmetrical factorial paired comparison experiments. Van Berkum (1987) constructs D-optimal paired comparison designs for quadratic response surfaces and for the estimation of main effects and two-factor interactions. In Street et al. (2001), the forms of the D- and A-optimal designs for 2^k paired comparison experiments are established, for testing main effects, main effects and two factor interactions and for testing main effects and two and three factor interactions.

In a paired comparison experiments 2^k treatments $T_1, \dots, T_{2^k}$ considered with n_{ij} of T_i and T_j , $i, j = 1, \dots, 2^k$. Bradley & Terry (1952) postulated that a preference parameter π_i , $\pi_i > 0$, is associated with T_i such that the probability of selecting (preferring) T_i when compared with T_j is $Pr(T_i \text{ preferred to } T_j) = \pi_i / (\pi_i + \pi_j)$, $i \neq j$, $i, j = 1, \dots, 2^k$. A convenient scale-determining constraint for the preference parameters is $\sum_{i=1}^{2^k} \gamma_i = 0$, where $\gamma_i = \ln \pi_i$, $i = 1, \dots, 2^k$. We let $\Lambda^{(k)}$ be the matrix of second derivatives of the likelihood function (assuming the Bradley-Terry model is

used), corresponding to $\pi = (\pi_1, \dots, \pi_{2^k})$. We let $\lambda_{ij} = \lim_{N \rightarrow \infty} n_{ij}/N$,

$i, j = 1, \dots, 2^k$. Then we have that

$$\Lambda_{ii}^{(k)} = \pi_i \sum_{j \neq i} \frac{\lambda_{ij} \pi_j}{(\pi_i + \pi_j)^2}, \quad i = 1, \dots, 2^k \quad (1.1)$$

and

$$\Lambda_{ij}^{(k)} = -\frac{\lambda_{ij} \pi_i \pi_j}{(\pi_i + \pi_j)^2}, \quad i \neq j, \quad i, j = 1, \dots, 2^k \quad (1.2)$$

We let B_m and B_n be a $m \times 2^k$ and $n \times 2^k$ matrices respectively with rows that sum to zero and are orthonormal. We also let B_m^* be a $(2^k - m - 1) \times 2^k$ matrix that completes the rows of B_m and $\underline{1}/\sqrt{2^k}$ to form a $2^k \times 2^k$ orthonormal matrix, where $\underline{1}$ is a row Vector with all entries equal to 1 and with a length appropriate to the context. Assuming that $B_m \gamma(\pi) = \underline{0}$, El-helbawy & Bradley (1978) show that the unique solution of the likelihood equations, π_L say, is a consistent estimator of π_0 and that $\gamma(\pi_L)$ is a consistent estimator of $\gamma(\pi_0)$. Finally they show that $\sqrt{N} B_n [\gamma(\pi_L) - \gamma(\pi_0)]$ has a limiting distribution n-variate normal with mean Vector zero and with covariance matrix $B_n B_m'^* [C^k(\pi)]^{-1} B_m^* B_n'$, where $C^k(\pi) = B_m^* \Lambda^{(k)} B_m'^*$. When B_m^* has B_n for its first n rows then this covariance matrix is the first principal minor of order n of $[C^k(\pi)]^{-1}$.

In some fields of applications like psychology, when the treatments are described by a considerable number of different factors, it is usually too difficult for a judge to express his preference probably. Therefore, partial profiles are shown to the judged, i.e. within each single paired comparison only a few factors

are specified while the others remain undetermined. The number of factors to be presented is bounded by the profile strength. For instance, when there are five factors in an experiment and a profile strength of two is employed, only pairs of treatments have to be evaluated these pairs in which (at most) two factors differ. The factors which are allowed to vary may however change from pair to pair.

Optimal 2^k paired comparison design for partial profiles is presented without proof by Schwabe et al. (2002). In Grasshoff et al. (2003), optimal paired comparison designs in case that the same number of levels is available for each factor independently are determined such that paired comparisons are used in which partial profiles of the treatments are presented.

In our paper, we investigate the form of the E-optimal paired comparison design with full profiles for detecting main effects.

We treat the case when there are k factors, each with two levels, and the set of possible choice pairs is restricted so that each pair of treatment combinations in which there are r factors with different levels appears equally often. For example, if $k = 3$ then the pairs with $r = 3$ factors different, specifically $\{000 \text{ Vs. } 111\}$, $\{001 \text{ Vs. } 110\}$, $\{010 \text{ Vs. } 101\}$ and $\{100 \text{ Vs. } 011\}$, would each appear the same number of times in the design (although that number might be 0), and the pairs with $r = 2$ factors different, specifically $\{000 \text{ Vs. } 110\}$, $\{000 \text{ Vs. } 101\}$, $\{000 \text{ Vs. } 011\}$, $\{001 \text{ Vs. } 111\}$, $\{001 \text{ Vs. } 100\}$, $\{001 \text{ Vs. } 010\}$, $\{010 \text{ Vs. } 111\}$, $\{010 \text{ Vs. } 100\}$, $\{011 \text{ Vs. } 110\}$, $\{011 \text{ Vs. } 101\}$, $\{100 \text{ Vs. } 111\}$ and $\{101 \text{ Vs. } 110\}$, would each appear the same number of times in the design (although that number might be 0), and the pairs with $r = 1$ factor different, specifically $\{000 \text{ Vs. } 100\}$, $\{000 \text{ Vs. } 010\}$, $\{000 \text{ Vs. } 001\}$, $\{001 \text{ Vs. } 101\}$, $\{001 \text{ Vs. } 011\}$, $\{010 \text{ Vs. } 110\}$, $\{010 \text{ Vs. } 011\}$, $\{011 \text{ Vs. } 111\}$, $\{100 \text{ Vs. } 110\}$, $\{100 \text{ Vs. } 101\}$, $\{101 \text{ Vs. } 111\}$ and $\{110 \text{ Vs. } 111\}$, would each appear the same number of times, again possibly 0 times, in the design.

Anderson et al. (1992), Chrzan & Orme (2000) and Bunch et al. (1996) assumed that participants must choose one of the options in each pair. Thus there is no "neither of these" option in the experiments we investigate.

This paper aims to construct the E-optimal design for estimating the main effects, that is, for estimating B_n which indicate to the treatment contrasts of the main effects, under the assumption that $\pi_0 = \underline{1}$.

In section 2, we show the $C^k(\pi_0)$ matrix for the class of designs that we are considering in our paper. In section 3, we use this $C^k(\pi_0)$ matrix to demonstrate that the E-optimal design for testing hypotheses about the main effects is the design in which the only pairs are the fold-over pairs.

2. The derivation of $[C^k(\pi)]^{-1}$ variance-covariance matrix

Street et al. (2001) used the method of El-helbawy, Ahmed and Alharbey (1994) to drive the $C^k(\pi_0)$ matrix for the model in which no contrasts are assumed to be null (that is, B_m is empty). Let $a^{(k,r)}$ denotes the proportion of frequency of pairs in which there are r factors with different levels. The treatment combinations placed in Yates standard order (that is, when represented as a binary k -tuple the first factor appears in the last position, the second in the second last position and so on).

Definition 2.1:

Let $D^{(k,r)}$ be a $(0,1)$ matrix of order 2^k such that $d_{ij}^{(k,r)}$ is the element in position (i, j) where,

$$d_{ij}^{(k,r)} = \begin{cases} 1 & \text{If } T_i, T_j \text{ have } r \text{ factor with different levels} \\ 0 & \text{Otherwise} \end{cases}$$

So,

$$D^{(k+1,r)} = \begin{pmatrix} D^{(k,r)} & D^{(k,r-1)} \\ D^{(k,r-1)} & D^{(k,r)} \end{pmatrix}, \quad k \geq r \geq 1,$$

$$D^{(k+1,0)} = I_{2^k} \quad (2.1)$$

where I_{2^t} is the identity matrix of order 2^{k+1} .

Assuming $\pi_0 = \underline{1}$, and in view of (1.1) and (1.2), the $\Lambda^{(k)}$ matrix can be written as follows:

$$4\Lambda^{(k)} = \left(k a^{(k,1)} + \binom{k}{2} a^{(k,2)} + \dots + \binom{k}{r} a^{(k,r)} + \dots + a^{(k,k)} \right) I_t - \sum_{r=1}^k a^{(k,r)} D^{(k,r)} \quad (2.2)$$

Definition 2.2:

Let $B_{t-1}^{(k)}$ be the $(2^k - 1) \times 2^k$ orthonormal matrix of the usual contrasts associated with a 2^k factorial experiments, where t represents number of treatments (i.e. $t = 2^k$), and let the column labels be the treatment combination in standard order.

Then,

$$B_{t-1}^{(k)} = \frac{1}{\sqrt{t}} \begin{bmatrix} B_{\frac{t}{2}-1}^{(k-1)} & B_{\frac{t}{2}-1}^{(k-1)} \\ -\underline{1}' & \underline{1}' \\ -B_{\frac{t}{2}-1}^{(k-1)} & B_{\frac{t}{2}-1}^{(k-1)} \end{bmatrix} \quad (2.3)$$

Since B_m is empty, we have that $B_m^* = B_{t-1}^{(k)}$. We begin by evaluating

$D^{(k,r)} B_{t-1}'^{(k)}$. First we observe that $D^{(2,0)} B_3'^{(2)} = B_3'^{(2)}$ and that

$$D^{(2,1)} B_3'^{(2)} = \frac{1}{\sqrt{4}} \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & 2 \\ 0 & 0 & 2 \\ 0 & 0 & -2 \end{pmatrix}, \quad D^{(2,2)} B_3'^{(2)} = \frac{1}{\sqrt{4}} \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & -1 \\ -1 & -1 & 1 \end{pmatrix} \quad (2.4)$$

Hence the column of $B_3^{(2)}$ together with $\underline{1}$, are eigen vectors of $D^{(2,r)}$. Given the definitions of $B_{t-1}^{(k)}$ and $D^{(k,r)}$, we have by mathematical induction that the treatment contrasts for each k are the eigen vectors of $D^{(k,r)}$. Thus, we can express the $D^{(k,r)}$, and hence $\Lambda^{(k)}$, in terms of treatment contrasts. Then $C^k(\pi_0) = B_m^* \Lambda^{(k)} B_m^{*'} is a diagonal, and the diagonal entries for effects of the same order are equal. Now we can evaluate the first $k \times k$ principal minor of $C^k(\pi_0)$ matrix without difficulty.$

3. Detecting the main effects

From the previous section, we know the general form of $C^k(\pi_0)$. We now evaluate the first $k \times k$ principal minor of $C^k(\pi_0)$ associated with the main effects so that we can determine the E-optimal design (i.e. the design which minimizing the largest eigenvalue of a defined covariance matrix) for detecting the main effects, then,

$$H_0: B_{k,t}^{(k)} \gamma(\pi) = \underline{0} \quad (3.1)$$

$$H_1: B_{k,t}^{(k)} \gamma(\pi) \neq \underline{0} \quad (3.2)$$

Where $B_{k,t}^{(k)}$ denotes the rows of $B_{t-1}^{(k)}$ that correspond to main effects, then:

$$B_{k,t}^{(k)} D^{(k,r)} = \left[\binom{k-1}{r} - \binom{k-1}{r-1} \right] B_{k,t}^{(k)}, \quad 1 \leq r \leq k \quad (3.3)$$

When $k = 3$, we find that

$$B_{3,8}^{(3)} D^{(3,1)} = (1) \times B_{3,8}^{(3)} \quad (3.4)$$

$$B_{3,8}^{(3)} D^{(3,2)} = (-1) \times B_{3,8}^{(3)} \quad (3.5)$$

$$B_{3,8}^{(3)} D^{(3,3)} = (-1) \times B_{3,8}^{(3)} \quad (3.6)$$

In view of (3.3) when $B_{t-1}^{(k)}$ has $B_{k,t}^{(k)}$ for its first k rows, then the first

$k \times k$ principal minor of $C^k(\pi_0)$ associated with the main effects is:

$$C_k = 2^{-1} \sum_{r=1}^k a^{(k,r)} \binom{k-1}{r-1} I_k \quad (3.7)$$

We can show the first $k \times k$ principal minor of the $C^k(\underline{1})$ matrix as the following:

$$C_k = B_{k,t}^{(k)} \Lambda^{(k)}(\underline{1}) B_{k,t}'^{(k)} \quad (3.8)$$

In view of (2.2), we can write (3.8) as

$$C_k = \frac{1}{4} B_{k,t}^{(k)} \left[\sum_{r=1}^k \binom{k}{r} a^{(k,r)} I_t - \sum_{r=1}^k a^{(k,r)} D^{(k,r)} \right] B_{k,t}'^{(k)} \quad (3.9)$$

In view of (3.3), we can show that

$$C_k = \frac{1}{4} \left\{ \sum_{r=1}^k \binom{k}{r} a^{(k,r)} I_k - \sum_{r=1}^k a^{(k,r)} \left[\binom{k-1}{r} - \binom{k-1}{r-1} \right] B_{k,t}^{(k)} B_{k,t}'^{(k)} \right\} \quad (3.10)$$

and we can write (3.10) in a simple form to be

$$\begin{aligned} C_k &= \frac{1}{4} \sum_{r=1}^k a^{(k,r)} \left[\binom{k}{r} - \binom{k-1}{r} + \binom{k-1}{r-1} \right] I_k \\ &= \frac{1}{4} \sum_{r=1}^k a^{(k,r)} \left[\binom{k-1}{r-1} + \binom{k-1}{r-1} \right] I_k \end{aligned} \quad (3.11)$$

$$C_k = 2^{-1} \sum_{r=1}^k a^{(k,r)} \binom{k-1}{r-1} I_k$$

Theorem 1:

Consider 2^k paired comparison experiments with full profiles and consider the hypothesis testing situations by (3.1), (3.2) under the assumption of

each pair of treatment combinations in which there are r factors with different levels appears equally often, $r = 1, 2, \dots, k$. The E-optimal design for detecting the main effects consist of the pairs in which all factors are different, this pairs called the fold-over pairs.

Proof:

Recall that the constraint of the proportion of the pairs is

$$\sum_{j \neq i}^t \lambda_{ij} = \sum_{r=1}^k \binom{k}{r} a^{(k,r)}, \quad i, j = 1, \dots, t, \quad i \neq j \quad (3.12)$$

Then,

$$\sum_{i=1}^t \sum_{j \neq i}^t \lambda_{ij} = 2^k \sum_{r=1}^k \binom{k}{r} a^{(k,r)} \quad (3.13)$$

Since, $\sum_{i < j}^t \lambda_{ij} = 1$, and in view of (3.13), we can write that:

$$\sum_{i < j}^t \lambda_{ij} = \frac{1}{2} \cdot 2^k \sum_{r=1}^k \binom{k}{r} a^{(k,r)} = 1 \quad (3.14)$$

Thus, (3.14) can be written as follows:

$$2^{k-1} \sum_{r=1}^k \binom{k}{r} a^{(k,r)} = 1 \quad (3.15)$$

In view of (3.7), to find the E-optimal design, we must maximize

$$\Psi = 2^{-1} \sum_{r=1}^k a^{(k,r)} \binom{k-1}{r-1} \quad (3.16)$$

Where Ψ is the equally eigenvalue of C_k matrix in (3.7) subject to the constraint (3.15). Considering expression in (3.16), it is clear that the maximum value of Ψ is attained when $a^{(k,k)} = 1/2^{k-1}$ and all other $a^{(k,r)}$ are zero. That is, the E-optimal paired comparison design for detecting main effects consists of the fold-over pairs only.

Numerical example:

Consider a paired comparisons experiment includes three factors each of them has two levels 0,1, and we want to construct the E-optimal design for detecting the main effects such that no contrasts are assumed to be null under

the assumption of each pair of treatment combinations in which there are r factors with different levels appears equally often, $r = 1, 2, 3$.

We construct three designs as follows:

Design 1:

Table (3.1) shows number of judgments for each comparison in design 1.

Table (3.1) number of judgments for each comparison								
	T_{000}	T_{001}	T_{010}	T_{011}	T_{100}	T_{101}	T_{110}	T_{111}
T_{000}	0	10	10	15	10	15	15	20
T_{001}		0	15	10	15	10	20	15
T_{010}			0	10	15	20	10	15
T_{011}				0	20	15	15	10
T_{100}					0	10	10	15
T_{101}						0	15	10
T_{110}							0	10
T_{111}								0

Design 2:

Table (3.2) shows number of judgments for each comparison in this design.

Table (3.2) number of judgments for each comparison								
	T_{000}	T_{001}	T_{010}	T_{011}	T_{100}	T_{101}	T_{110}	T_{111}
T_{000}	0	0	0	5	0	5	5	20
T_{001}		0	5	0	5	0	20	5
T_{010}			0	0	5	20	0	5
T_{011}				0	20	5	5	0
T_{100}					0	0	0	5
T_{101}						0	5	0
T_{110}							0	0
T_{111}								0

Design 3:

Table (3.3) shows number of judgments for each comparison in this design.

Table (3.3) number of judgments for each comparison

	T_{000}	T_{001}	T_{010}	T_{011}	T_{100}	T_{101}	T_{110}	T_{111}
T_{000}	0	0	0	0	0	0	0	20
T_{001}		0	0	0	0	0	20	0
T_{010}			0	0	0	20	0	0
T_{011}				0	20	0	0	0
T_{100}					0	0	0	0
T_{101}						0	0	0
T_{110}							0	0
T_{111}								0

The values of $a^{(3,r)}$, $r = 1, 2, 3$. (proportion of frequency of pairs in which there are r factors with different levels) and eigenvalues of C matrix (ψ) for all above designs are illustrated in the following table:

Table (3.4) the results of designs

The Design	$a^{(3,1)}$	$a^{(3,2)}$	$a^{(3,3)}$	Eigenvalues of C matrix (ψ)
Design 1	1/38	3/76	1/19	3/38
Design 2	0	1/28	1/7	3/28
Design 3	0	0	1/4	1/8

From table (3.4), we notice that the value of the eigenvalue of Design 3 is the largest eigenvalue and it is the E-optimal design for detecting the main effects.

References

- [1] Anderson, D.; Borgers, A.; Ettema, D. and Timmermans, H. (1992). Estimating availability effects in travel choice modeling: A stated choice approach, *Transportation research record*, 1357, 51-65.
- [2] Bunch, D. S.; Louviere, J. J. and Anderson, D. A. (1996). A comparison of experimental design strategies for choice-based conjoint analysis with generic-attribute multinomial logit models, working paper, Graduate School of Management, UC Davis.
- [3] Bradley, R. A. and Terry, M. E. (1952). Rank analysis of incomplete block designs. I. the method of paired comparisons. *Biometrika*, 39, 324-345.
- [4] Chrzan, K and Orme, B. (2000). An overview and comparison of design strategies for choice-based conjoint analysis, From <http://www.sawtoothsoftware.com/Techpap.html/gencon>.
- [5] El-Helbawy, A. T. and Ahmed, E. A. (1984). Optimal design results for 2^n factorial paired comparisons experiments. *Communication in statistics, Theory and methods*, 13, 2827-2845.
- [6] El-Helbawy, A. T.; Ahmed, E. A. and Alharbey, A. H. (1994). Optimal designs for asymmetrical factorial paired comparison experiments. *Communications in statistics, simulation and computation*, 23, 663-681.
- [7] El-Helbawy, A. T. and Bradley, R. A. (1978). Treatment contrasts in paired comparisons: Large sample results applications and some optimal designs. *Jour. Amer. Stat. Assoc.*, 73, 831-839.
- [8] Grasshoff, U., Grossmann, H., Holling, H. and Schwabe, R. (2003). Optimal paired comparison designs for first-order interactions. *Statistics*, 37, 373-386.
- [9] Schwabe, R.; Grasshoff, U.; Grossmann, H. and Holling, H. (2002). Optimal paired comparison designs for partial profiles. *Tatra Mountains Math.* (submitted).
- [10] Street, D. J. Bunch, D. S. and Moore, B. J. (2001). Optimal designs for 2^k paired comparison experiments. *Communication Stat, Theory and methods*, 30, 2149-2171.
- [11] Van Berkum, E. E. (1987). Optimal paired comparison designs for factorial and quadratic models. *Journal of statistical planning and inference*, 15, 265-278.

المصادر

(References)

- 1- Aguilar , O. and M.West (2000) , " Bayesian dynamic factor models and Portfolio allocation " , Journal of Business and Economic Statistics 18 : 338 – 357.
- 2- Gerlach , R . , C . Carter and R.Kohn (2000) , " Efficient Bayesian inference for dynamic mixture models " , Journal of the American Statistical Association 95 : 819 – 828 .
- 3- Harrison , P . J . and C.F.Stevens (1976) , " Bayesian forecasting " , Journal of the Royal Statistical Society . series B (Methodological) , Vol . 50 , No . 3 : 205 – 247 .
- 4- Petridis , V . , A . Kehagias , L . P etrou , A . Bakirtzis A . , S . Kiartzis , H . Panagiotou and N . Maslaris (2001) , " A Bayesian multiple models combination method for time series prediction " , Journal of Intelligent and Robotic Systems 31 : 69 – 89 .
- 5- Smith , A . F . (1982) , " Monitoring Kidney transplants Patients " , the Statistic , vol. 32 : 46 – 54 .
- 6- West , M . and J . Harrison (1997) , " Bayesian Forecasting and Dynamic Models " , (Second Edition , Springer , New York) .

$$c_{1,t} = r_{1,t} - \frac{(r_{1,t})^2}{1+r_{1,t}}$$

$$c_{2,t} = r_{2,t} - \frac{(r_{1,t})(r_{2,t})}{1+r_{1,t}}$$

$$c_{3,t} = r_{3,t} - \frac{(r_{2,t})^2}{1+r_{1,t}}$$

$$\hat{x}_{t+1} = u_t + Z b_t \quad \text{وكذلك :}$$

$$\hat{x}_{t+1} = r_{1,t+1} + 1$$

حيث ان $r_{1,t+1}$ نحصل عليها من العلاقة (21) وتكون بالصيغة التالية :

$$r_{1,t+1} = c_{1,t} + 2 Z c_{2,t} + Z^2 c_{3,t} + n r_\beta w_1^2 + n r_\mu$$

ونشير الى اننا نحتاج جميع العلاقات المتكرره المذكوره اعلاه لايجاد قيمة u_t ، ومن الممكن برمجة هذه العلاقات باستخدام احدى لغات الحاسوب ، وذلك في حال التطبيقات العمليه .

بهذا نكون قد اوضحنا كيفية ايجاد العلاقات المتكرره للنموذج الحركي الخطي المركب لفترات زمنية غير متساويه ، وذلك لدراسة سياسة الفحص بدلالة الزمن في حالة ان كل من $(r_\mu , r_\beta , \sigma^2)$ معلوم ، ونوصي بايجاد التوزيعات الاحصائية سالفة الذكر في حالة ان كل من $(r_\mu , r_\beta , \sigma^2)$ غير معلوم .

$$\begin{aligned}
 &= E \left(\begin{matrix} \mu_t \\ \beta_t \end{matrix} \setminus D_t \right) \\
 &= \begin{pmatrix} u_t \\ b_t \end{pmatrix} \\
 R_t &= \begin{bmatrix} r_{1,t} & r_{2,t} \\ r_{2,t} & r_{3,t} \end{bmatrix} \quad (24)
 \end{aligned}$$

وبموجب قوانين المتواليه الهندسية فان :

$$H_t = \begin{bmatrix} 1 & \frac{h^{n+1}-hz^n}{hz^n-z^{n+1}} \\ 0 & \left(\frac{h}{z}\right)^n \end{bmatrix}$$

وعلى ضوء ذلك فانه من العلاقات (15) ، (16) ، (17) ، (24) ، يمكننا التوصل الى العلاقات التعاقبيه (المتجدده) ، (Recursive) التالية :

$$a_{1,t} = \frac{r_{1,t}}{1+r_{1,t}}$$

$$a_{2,t} = \frac{r_{2,t}}{1+r_{2,t}}$$

وللسهوله نضع :

$$Z = \frac{h^{n+1}-hz^n}{hz^n-z^{n+1}}$$

وبالتالي يكون :

$$f_t = u_{t-1} Z b_{t-1}$$

$$u_t = u_{t-1} + Z b_{t-1} + a_{1,t} (x_t - f_t)$$

$$b_t = \left(\frac{h}{z}\right)^n b_{t-1} + a_{2,t} (x_t - f_t)$$

$$r_{1,t} = c_{1,t-1} + 2Zc_{2,t-1} + Z^2c_{3,t-1} + nr_{\beta}w_1^2 + nr_{\mu}$$

$$r_{2,t} = \left(\frac{h}{z}\right)^n c_{2,t-1} + \left(\frac{h}{z}\right)^n Zc_{3,t-1} + r_{\beta}w_1w_2$$

$$r_{3,t} = \left(\frac{h}{z}\right)^{2n} c_{3,t-1} + r_{\beta}w_3$$

التوزيع التنبؤي : Prediction Distribution

ان ايجاد التوزيع التنبؤي ذو اهمية ، وذلك لانه يغني بالمعلومات المستقبلية ، اذ يمكن على الاقل معرفة ماذا سيكون عليه معدل النوعية للماده التي سوف تنتج ، واذا كان تباين هذا التوزيع كبيرا ، فهذا يزيد من عدم التاكديه (*Uncertainty*) حول العملية الانتاجيه ، ومن هنا نستعد لاتخاذ القرار المناسب . وللحصول على التوزيع التنبؤي خطوه واحده الى امام (*One - Step Ahead Prediction Distribution*) لابد من كتابة صيغة النموذج الحركي الخطي المركب لفترات زمنية غير متساويه ، المعرف بالعلاقين (13) ، (14) بالصيغه المستقبلية كالتالي :

$$x_{t+1} = F\theta_{t+1} + v_{t+1} \quad , \quad v_{t+1} \sim N(0, \sigma^2) \quad (18)$$

$$\theta_{t+1} = H_t\theta_t + \gamma_{t+1} \quad \gamma_{t+1} \sim N(0, \sigma^2 W) \quad (19)$$

حيث أن v_{t+1} ، γ_{t+1} احداث متنافيه ومستقله التوزيع . وعلى ضوء ذلك يمكننا التوصل إلى النتائج التالية : من العلاقه (19) نحصل على ان :

$$(\theta_{t+1} \setminus D_t) \sim N(H_t m_t, \sigma^2 R_{t+1}) \quad (20)$$

حيث ان :

$$R_{t+1} = H_t C_t H' + W \quad (21)$$

كذلك من العلاقات (18) ، (19) ، (20) ، يمكننا التوصل إلى أن :

$$(x_{t+1} \setminus D_t) \sim N(\hat{x}_{t+1}, \sigma^2 \hat{\lambda}_{t+1})$$

حيث ان :

$$\hat{x}_{t+1} = FH_t m_t$$

$$\hat{\lambda}_{t+1} = FR_{t+1}F' + 1$$

الان في نموذجنا هذا اذا وضعنا :

$$C_t = \begin{bmatrix} c_{1,t} & c_{2,t} \\ c_{2,t} & c_{3,t} \end{bmatrix} \quad (22)$$

$$A_t = \begin{bmatrix} a_{1,t} \\ a_{2,t} \end{bmatrix} \quad (23)$$

$$m_t = E(\theta_t \setminus D_t)$$

$$E(x_t \setminus \theta_t) = E[(F\theta_t + V_t) \setminus D_t]$$

$$= F\theta_t$$

$$var(x_t \setminus \theta_t) = var[(F\theta_t + V_t) \setminus D_t]$$

$$= \sigma^2$$

ونصيف هذا التوزيع الاحتمالي رياضيا بالشكل التالي :

$$(x_t \setminus \theta_t) \sim N(F\theta_t, \sigma^2)$$

وبالتالي التوزيع الاحتمالي البعدي نحصل عليه كالتالي :

$$P(\theta_t \setminus D_t) \propto \text{Exp} \left[\frac{-1}{2\sigma^2} (\theta_t - H_t m_{t-1})' R_t^{-1} (\theta_t - H_t m_{t-1}) \right]$$

$$\cdot \text{Exp} \left[\frac{-1}{2\sigma^2} (x_t - F\theta_t)' (x_t - F\theta_t) \right]$$

وباجراء عملية الضرب ودمج الكميات المتشابهة مع بعض ، والكميات الثابتة مع ثابت التناسب ، يمكننا التوصل الى ان :

$$P(\theta_t \setminus D_t) \propto \text{Exp} \left[\frac{-1}{2\sigma^2} (\theta_t - m_t)' C_t^{-1} (\theta_t - m_t) \right]$$

وذلك بوضع :

$$m_t = H_t m_{t-1} + A_t (x_t - f_t) \quad (15)$$

حيث m_t يمثل توقع التوزيع الاحتمالي البعدي عند الزمن t ، يسمى بترشيح كالمن (Kalmen Filter) ، وان $(x_t - f_t)$ يدعى بخطأ التنبؤ ، حيث $f_t = FH_t m_{t-1}$. وان :

$$C_t' = R_t - A_t F R_t \quad (16)$$

حيث C_t' تمثل مصفوفة التباين والتباين المشترك للتوزيع الاحتمالي البعدي عند الزمن t ، وان :

$$A_t = R_t F' (1 + F R_t F')^{-1} \quad (17)$$

حيث A_t يدعى بمعامل كالمن (Kalmen Factor) ، وعليه يمكن صياغة التوزيع الاحتمالي البعدي رياضيا بالشكل التالي :

$$(\theta_t \setminus D_t) \sim N(m_t, \sigma^2 C_t)$$

وبذلك يتحقق المطلوب .

البرهان : باستخدام الاستقراء الرياضي ، نستنتج انه عند الزمن ($t-1$) يكون :

$$(\theta_{t-1} \setminus D_{t-1}) \sim N(m_{t-1}, \sigma^2 C_{t-1})$$

حيث ان المعلومات المتوافره D_{t-1} هي :

$$D_{t-1} = (x_{t-1}, x_{t-2}, \dots, x_1, r_\mu, r_\beta, \sigma^2)$$

ومن هذا يمكن الحصول على التوزيعات التاليه :

1 (التوزيع الاحتمالي القبلي للمعلمه θ_t عند الزمن t هو :

من العلاقه (14) نجد ان :

$$E(\theta_t \setminus D_{t-1}) = E[(H_t \theta_{t-1} + \gamma_t) \setminus D_{t-1}]$$

$$= H_t m_{t-1}$$

$$var(\theta_t \setminus D_{t-1}) = var[(H_t \theta_{t-1} + \gamma_t) \setminus D_{t-1}]$$

$$= H_t var(\theta_{t-1} \setminus D_{t-1}) H_t' + var(\gamma_t \setminus D_{t-1})$$

$$= H_t \sigma^2 C_{t-1} H_t' + \sigma^2 W$$

$$= \sigma^2 (H_t C_{t-1} H_t' + W)$$

$$= \sigma^2 R_t$$

$$R_t = H_t C_{t-1} H_t' + W$$

وذلك بوضع :

حيث ان R_t مصفوفه موجبه مؤكده (*Posetive Definite*) ومتماثله (*Symmetric*) ، ونصيح هذا التوزيع الاحتمالي رياضيا بالشكل التالي :

$$(\theta_t \setminus D_{t-1}) \sim N(H_t m_{t-1}, \sigma^2 R_t)$$

2 (التوزيع الاحتمالي البعدي ، نحصل عليه باستخدام نظريه بيز التي لها الصيغه التاليه :

$$P(\theta_t \setminus D_t) \propto P(\theta_t \setminus D_{t-1}) . P(x_t \setminus D_t)$$

حيث ان : $P(x_t \setminus D_t)$ تمثل دالة الترجيح ، ونحصل على التوقع و التباين لهذه الداله من العلاقه

(13) وكما يلي :

ومن أجل التبسيط نضع :

$$w_1 = 1 + \sum_{j=1}^n \left(\frac{h}{z}\right)^d = \frac{h^{n+1} - z^{n+1}}{hz^n - z^{n+1}}$$

$$w_2 = \sum_{j=1}^n \left(\frac{h}{z}\right)^{n-j} = \frac{h^n - z^n}{hz^{n-1} - z^n}$$

$$w_3 = \sum_{j=1}^n \left(\frac{h}{z}\right)^{2(n-j)} = \frac{h^{2n} - z^{2n}}{h^2 z^{2n-2} - z^{2n}}$$

وعليه فإن المصفوفة W تصبح بالشكل التالي :

$$W = \begin{bmatrix} nr_\beta w_1^2 + nr_\mu & r_\beta w_1 w_2 \\ r_\beta w_1 w_2 & r_\beta w_3 \end{bmatrix} \quad (12)$$

وعلى ضوء ذلك يمكن إعادة صياغة النموذج الحركي الخطي المركب لفترات زمنية غير متساوية ، عند زمن الفحص t كالآتي :

$$x_t = F\theta_t + v_t, \quad v_t \sim N(0, \sigma^2) \quad (13)$$

$$\theta_t = H_t \theta_{t-1} + \gamma_t, \quad \gamma_t \sim N(0, \sigma^2 W) \quad (14)$$

وبالتالي بعد الحصول على العلاقتين (13) ، (14) وباستخدام نظريه بيز والاستفادة من عملية ماركوف (Markov Process) وباستخدام أسلوب التمثيل المتسلسل (sequenteil Analysis) ، والاستفادة من كتاب (West and Harrison (1997) ، يمكن الحصول على التوزيعات الاحتمالية البعديه لمعطيات العمليه من خلال الاجراءات المتكرره ، وباعتبار كل من (r_β, r_μ, σ^2) معلومه ، فاننا نصيغ النظرية التالية :

نظريه (2) : اذا كان التوزيع القبلي (Prior Distrbution) لـ θ_0 قبل أي مشاهدته ، يتبع التوزيع الطبيعي ($N(m_0, \sigma^2 C_0)$) ، فإن التوزيع البعدي (Posterior Distrbution) لـ θ_t عند الزمن t يتبع ايضا التوزيع الطبيعي أي ان :

$$(\theta_t \setminus D_t) \sim N(m_t, \sigma^2 C_t)$$

حيث ان المعلومات المتوافره D_t هي :

$$D_t = (x_t, x_{t-1}, \dots, x_1, r_\mu, r_\beta, \sigma^2)$$

وهي عبارته عن المعلومات السابقة (المشاهدات) ، والثوابت المعروفة الاخرى .

$$= G_t^n \theta_{t-1} + G_t^{n-1} \omega_{1,t-1} + G_t^{n-2} \omega_{2,t-1} + \dots + G_t \omega_{n-1,t-1} + \omega_{n,t-1}$$

$$H_t = G_t^n \quad \text{وبوضع كل من :}$$

$$\gamma_t = G_t^{n-1} \omega_{1,t-1} + G_t^{n-2} \omega_{2,t-1} + \dots + G_t \omega_{n-1,t-1} + \omega_{n,t-1}$$

$$\theta_t = H_t \theta_{t-1} + \gamma_t \quad \text{نحصل على :}$$

وبذلك يتحقق المطلوب ..

كما يمكننا الحصول على مصفوفة التباين والتباين المشترك لـ $\omega_{j,t-1}$ كالآتي :

$$\omega_{j,t-1} = \begin{pmatrix} \Delta_{j,t-1} + \delta_{j,t-1} \\ \Delta_{j,t-1} \end{pmatrix} \quad \text{بما أن :}$$

$$\text{var}(\omega_{j,t-1}) = E \begin{pmatrix} \Delta_{j,t-1} + \delta_{j,t-1} \\ \Delta_{j,t-1} \end{pmatrix} (\Delta_{j,t-1} + \delta_{j,t-1} \quad \Delta_{j,t-1}) \quad \text{فان :}$$

$$= E \begin{bmatrix} (\Delta_{j,t-1} + \delta_{j,t-1})^2 & (\Delta_{j,t-1} + \delta_{j,t-1}) \Delta_{j,t-1} \\ \Delta_{j,t-1} (\Delta_{j,t-1} + \delta_{j,t-1}) & (\Delta_{j,t-1})^2 \end{bmatrix}$$

$$= \sigma^2 \begin{bmatrix} r_\beta + r_\mu & r_\beta \\ r_\beta & r_\beta \end{bmatrix}$$

$$= \sigma^2 r$$

(11)

$$r = \begin{bmatrix} r_\beta + r_\mu & r_\beta \\ r_\beta & r_\beta \end{bmatrix}$$

وذلك بوضع :

$$\text{var}(\gamma_t) = [G_t^{n-1} r (G_t^{n-1})' + G_t^{n-2} r (G_t^{n-2})' + \dots + G_t r G_t' + r] \sigma^2 \quad \text{وبالتالي :}$$

باستخدام خواص ضرب وجمع المصفوفات نستطيع التوصل إلى أن :

$$\text{var}(\gamma_t) = \sigma^2 W$$

حيث أن :

$$W = \begin{bmatrix} nr_\beta \left[1 + \sum_{j=1}^n \left(\frac{h}{z} \right)^j \right]^2 + nr_\mu & r_\beta \left[1 + \sum_{j=1}^n \left(\frac{h}{z} \right)^j \right] \sum_{j=1}^n \left(\frac{h}{z} \right)^{n-j} \\ r_\beta \left[1 + \sum_{j=1}^n \left(\frac{h}{z} \right)^j \right] \sum_{j=1}^n \left(\frac{h}{z} \right)^{n-j} & r_\beta \sum_{j=1}^n \left(\frac{h}{z} \right)^{2(n-j)} \end{bmatrix}$$

وبالتالي فان :

$$\mu_{j,t-1} = \mu_{0,t-1} + \left[\frac{h^{n+1}-hz^n}{hz^n-z^{n+1}} \right] \beta_{0,t-1} + \left[\frac{h^n-z^n}{hz^{n-1}-z^n} \right] \Delta_{1,t-1} +$$

$$+ \left[\frac{h^{n-1}-z^{n-1}}{hz^{n-2}-z^{n-1}} \right] \Delta_{2,t-1} + \dots + \Delta_{j,t-1} + \delta_{1,t-1} + \delta_{2,t-1} + \dots + \delta_{j,t-1}$$

⋮

$$\mu_{n,t-1} = \mu_{0,t-1} + \left[\frac{h^{n+1}-hz^n}{hz^n-z^{n+1}} \right] \beta_{0,t-1} + \sum_{j=1}^n \frac{h^{n-j+1}-z^{n-j+1}}{hz^{n-j}-z^{n-j+1}} \Delta_{j,t-1} + \sum_{j=1}^n \delta_{j,t-1}$$

وباستخدام العرف في العلاقة (8) نحصل على :

$$\mu_t = \mu_{t-1} + \left[\frac{h^{n+1}-hz^n}{hz^n-z^{n+1}} \right] \beta_{t-1} + \sum_{j=1}^n \frac{h^{n-j+1}-z^{n-j+1}}{hz^{n-j}-z^{n-j+1}} \cdot \Delta_{j,t-1} + \sum_{j=1}^n \delta_{j,t-1} \quad (10)$$

وبالتالي من العلاقتين (9) ، (10) يمكن التوصل الى الصيغه النهائيه المطلوبه كالاتي :

$$\theta_t = \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix}$$

$$= \begin{bmatrix} \mu_{t-1} + \left[\frac{h^{n+1}-hz^n}{hz^n-z^{n+1}} \right] \beta_{t-1} + \sum_{j=1}^n \frac{h^{n-j+1}-z^{n-j+1}}{hz^{n-j}-z^{n-j+1}} \Delta_{j,t-1} + \sum_{j=1}^n \delta_{j,t-1} \\ \left(\frac{h}{z} \right)^n \beta_{t-1} + \sum_{j=1}^n \left(\frac{h}{z} \right)^{n-j} \Delta_{j,t-1} \end{bmatrix}$$

$$= \begin{bmatrix} \mu_{t-1} + \left[\frac{h^{n+1}-hz^n}{hz^n-z^{n+1}} \right] \beta_{t-1} \\ \left(\frac{h}{z} \right)^n \beta_{t-1} \end{bmatrix} + \begin{bmatrix} \sum_{j=1}^n \frac{h^{n-j+1}-z^{n-j+1}}{hz^{n-j}-z^{n-j+1}} \Delta_{j,t-1} + \sum_{j=1}^n \delta_{j,t-1} \\ \sum_{j=1}^n \left(\frac{h}{z} \right)^{n-j} \Delta_{j,t-1} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \frac{h^{n+1}-hz^n}{hz^n-z^{n+1}} \\ 0 & \left(\frac{h}{z} \right)^n \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \beta_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & \frac{h^n-hz^{n-1}}{hz^{n-1}-z^n} \\ 0 & \left(\frac{h}{z} \right)^{n-1} \end{bmatrix} \begin{bmatrix} \Delta_{1,t-1} + \delta_{1,t-1} \\ \Delta_{1,t-1} \end{bmatrix} +$$

$$+ \begin{bmatrix} 1 & \frac{h^{n-1}-hz^{n-2}}{hz^{n-2}-z^{n-1}} \\ 0 & \left(\frac{h}{z} \right)^{n-2} \end{bmatrix} \begin{bmatrix} \Delta_{2,t-1} + \delta_{2,t-1} \\ \Delta_{2,t-1} \end{bmatrix} + \dots +$$

$$+ \begin{bmatrix} 1 & \frac{h}{z} \\ 0 & \frac{h}{z} \end{bmatrix} \begin{bmatrix} \Delta_{n-1,t-1} + \delta_{n-1,t-1} \\ \Delta_{n-1,t-1} \end{bmatrix} + \begin{bmatrix} \Delta_{n,t-1} + \delta_{n,t-1} \\ \Delta_{n,t-1} \end{bmatrix}$$

$$= \left(\frac{h}{z}\right)^n \beta_{0,t-1} + \sum_{j=1}^n \left(\frac{h}{z}\right)^{n-j} \Delta_{j,t-1}$$

وباستخدام العرف في العلاقة (8) نحصل على :

$$\beta_t = \left(\frac{h}{z}\right)^n \beta_{t-1} + \sum_{j=1}^n \left(\frac{h}{z}\right)^{n-j} \Delta_{j,t-1} \quad (9)$$

الآن باستخدام العلاقة (6) و النتائج اعلاه الموصلة للعلاقة (9) نحصل على :

$$\mu_{1,t-1} = \mu_{0,t-1} + \beta_{1,t-1} + \delta_{1,t-1}$$

$$= \mu_{0,t-1} + \frac{h}{z} \beta_{0,t-1} + \Delta_{1,t-1} + \delta_{1,t-1}$$

$$\mu_{2,t-1} = \mu_{1,t-1} + \beta_{2,t-1} + \delta_{2,t-1}$$

$$= \mu_{0,t-1} + \left[\frac{h}{z} + \left(\frac{h}{z}\right)^2 \right] \beta_{0,t-1} + \left[1 + \frac{h}{z} \right] \Delta_{1,t-1} + \Delta_{2,t-1} + \delta_{1,t-1} + \delta_{2,t-1}$$

$$\mu_{3,t-1} = \mu_{2,t-1} + \beta_{3,t-1} + \delta_{3,t-1}$$

$$= \mu_{0,t-1} + \left[\frac{h}{z} + \left(\frac{h}{z}\right)^2 + \left(\frac{h}{z}\right)^3 \right] \beta_{0,t-1} + \left[1 + \frac{h}{z} + \left(\frac{h}{z}\right)^2 \right] \Delta_{1,t-1} + \left(1 + \frac{h}{z} \right) \Delta_{2,t-1} + \Delta_{3,t-1} + \delta_{1,t-1} + \delta_{2,t-1} + \delta_{3,t-1}$$

⋮

$$\begin{aligned} \mu_{j,t-1} &= \mu_{0,t-1} + \left[\sum_{d=1}^n \left(\frac{h}{z}\right)^d \right] \beta_{0,t-1} + \left[\sum_{d=1}^n \left(\frac{h}{z}\right)^{d-1} \right] \Delta_{1,t-1} + \\ &+ \left[\sum_{d=2}^n \left(\frac{h}{z}\right)^{d-2} \right] \Delta_{2,t-1} + \dots + \Delta_{j,t-1} + \delta_{1,t-1} + \delta_{2,t-2} + \dots + \delta_{j,t-1} \end{aligned}$$

وباستخدام قوانين المتواليه الهندسيه نحصل على :

$$\sum_{d=1}^n \left(\frac{h}{z}\right)^d = \frac{h^{n+1} - z^{n+1}}{hz^n - z^{n+1}}$$

$$\sum_{d=1}^n \left(\frac{h}{z}\right)^{d-1} = \frac{h^n - z^n}{hz^{n-1} - z^n}$$

$$\sum_{d=2}^n \left(\frac{h}{z}\right)^{d-2} = \frac{h^{n-1} - z^{n-1}}{hz^{n-2} - z^{n-1}}$$

نظريه (1) : لتكن $\theta_t = \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix}$ وباستخدام العلاقتين (6) ، (7) نحصل على :

$$\theta_t = H_t \theta_{t-1} + \gamma_t$$

$$H_t = G_t^n = \begin{pmatrix} 1 & \sum_{d=1}^n \left(\frac{h}{z}\right)^d \\ 0 & \left(\frac{h}{z}\right)^n \end{pmatrix} \quad \text{حيث ان :}$$

وذلك بوضع : $h = N_t$, $z = N_{t-1}$ ، وان :

$$\gamma_t = G_t^{n-1} \omega_{1,t-1} + G_t^{n-2} \omega_{2,t-1} + \dots + G_t \omega_{n-1,t-1} + \omega_{n,t-1}$$

$$\omega_{j,t-1} = \begin{pmatrix} \Delta_{j,t-1} + \delta_{j,t-1} \\ \Delta_{j,t-1} \end{pmatrix} , \quad j = 1, 2, \dots, n \quad \text{حيث ان :}$$

البرهان : لـ n من الوحدات المنتجة خلال الفترات الزمنية غير المتساوية وقصيرة الامد ، فانه

باستخدام العلاقه (7) نحصل على :

$$\beta_{1,t-1} = \frac{h}{z} \beta_{0,t-1} + \Delta_{1,t-1}$$

$$\beta_{2,t-1} = \frac{h}{z} \beta_{1,t-1} + \Delta_{2,t-1}$$

$$= \left(\frac{h}{z}\right)^2 \beta_{0,t-1} + \frac{h}{z} \Delta_{1,t-1} + \Delta_{2,t-1}$$

$$\beta_{3,t-1} = \frac{h}{z} \beta_{2,t-1} + \Delta_{3,t-1}$$

$$= \left(\frac{h}{z}\right)^3 \beta_{0,t-1} + \left(\frac{h}{z}\right)^2 \Delta_{1,t-1} + \frac{h}{z} \Delta_{2,t-1} + \Delta_{3,t-1}$$

:

$$\beta_{j,t-1} = \left(\frac{h}{z}\right)^j \beta_{0,t-1} + \left(\frac{h}{z}\right)^{j-1} \Delta_{1,t-1} + \left(\frac{h}{z}\right)^{j-2} \Delta_{2,t-1} + \dots + \frac{h}{z} \Delta_{j-1,t-1} + \Delta_{j,t-1}$$

:

$$\beta_{n,t-1} = \left(\frac{h}{z}\right)^n \beta_{0,t-1} + \left(\frac{h}{z}\right)^{n-1} \Delta_{1,t-1} + \left(\frac{h}{z}\right)^{n-2} \Delta_{2,t-1} + \dots + \frac{h}{z} \Delta_{n-1,t-1} + \Delta_{n,t-1}$$

في بحثنا هذا سوف نستعرض دراسة رياضية شاملة لحالة الفحص بدلالة الزمن في حالة الفترات الزمنية غير المتساوية ، ففي هذه الحالة تفحص النوعية خلال فترات زمنية ، مثلا تفحص وحده واحده في ازمته قصيره الامد وغير متساويه ، وتتمثل هذه السياسه من خلال العلاقتين (1) ، (2) ، فعند ماده المنتجه j حيث $(j=1,2,...,n)$ ، وبعد الزمن t ، تصبح هاتين العلاقتين كما يلي :

$$x_{j,t} = \mu_{j,t} + v_{j,t} \quad , \quad v_{j,t} \sim N(0, \sigma^2) \quad (5)$$

$$\mu_{j,t} = \mu_{j-1,t} + \beta_{j,t} + \delta_{j,t} \quad , \quad \delta_{j,t} \sim N(0, \sigma^2 r_\mu) \quad (6)$$

$$\beta_{j,t} = \frac{N_t}{N_{t-1}} \beta_{j-1,t} + \Delta_{j,t} \quad , \quad \Delta_{j,t} \sim N(0, \sigma^2 r_\beta) \quad (7)$$

حيث أن : $v_{j,t}$ ، $\delta_{j,t}$ ، $\Delta_{j,t}$ أخطاء عشوائيه غير مترابطه .

علما بأن :

t : يمثل الزمن للوحده المفحوصه .

n : يمثل عدد الوحدات المنتجه .

j : يمثل دليل الوحده المنتجه .

$x_{j,t}$: يمثل قيمه لنوعيه ماده المنتجه j بعد الزمن t .

$\mu_{j,t}$: يمثل الميل الحقيقي لنوعيه ماده المنتجه j بعد الزمن t .

$\beta_{j,t}$: يمثل عامل الميل (Slope Tactor) لنوعيه ماده المنتجه j بعد الزمن t .

σ^2 : يمثل تباين المشاهدات .

وليكن y رمزا شاملا (يمثل جميع المتغيرات العشوائيه المستخدمه في هذا البحث) فاننا سنستخدم العرف (Convention) التالي :

$$y_{n,t-1} = y_{0,t} = y_t \quad (8)$$

وللمزيد من المعلومات حول النماذج الحركيه واحصاء بيز ومجال تطبيقهما انظر Aguilar and petridis and et.al (2001) و Gerlach and et.al (2000) ، et.al (2000) .

مما سبق يمكننا التعبير عن النموذج الحركي الخطي المركب لفترات زمنية غير متساويه عند ازمته الفحوصات (او المشاهدات) الحقيقيه بشكل رياضي اوضح ، من خلال صياغة النظرية التاليه ، التي سنقوم باثبات صحتها في هذا البحث .

النوعيه ، الا اذا لم يكن هناك أي خطأ في المشاهده (*Free - Error - Observation*) . أما العلاقه (2) فانها تعبر عن ترابط او عن علاقات خطيه لمعاملات العمليه μ_t ، β_t وتطورها مع الزمن .

الان يمكن كتابة العلاقتين (1) ، (2) على صيغة النموذج الحركي الخطي المركب لفترات زمنية غير متساويه كالآتي :

$$x_t = F\theta_t + v_t \quad , \quad v_t \sim N(0, \sigma^2) \quad (3)$$

$$\theta_t = G_t\theta_{t-1} + \omega_t \quad , \quad \omega_t \sim N(0, \sigma^2 r) \quad (4)$$

حيث ان : v_t ، ω_t : متجهان طبيعيان عشوائيان مستقلان .

$$r = E(\omega_t \cdot \omega_t') \quad \text{وأن :}$$

$$= \begin{bmatrix} r_\beta + r_\mu & r_\beta \\ r_\beta & r_\beta \end{bmatrix}$$

$$F = \begin{pmatrix} 1 & 0 \end{pmatrix} \quad \text{علما بان :}$$

$$G_t = \begin{bmatrix} 1 & \frac{N_t}{N_{t-1}} \\ 0 & \frac{N_t}{N_{t-1}} \end{bmatrix}$$

$$\theta_t = \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix}$$

$$\omega_t = \begin{pmatrix} \Delta_t + \delta_t \\ \Delta_t \end{pmatrix}$$

سياسات الفحص : *Inspection Policies*

عادة يكون هناك العديد من سياسات الفحص التي تجري على نوعية الماده المنتجه في الحياه العمليه منها : الفحص بدلالة الوحدات المنتجه و الفحص بدلالة الزمن .

سنتناول سياسة الفحص بدلالة الزمن ، وفي هذه الحاله تفحص الوحدات المنتجه خلال فترات زمنية متساويه او غير متساويه فمثلا :

1) فحص وحده واحده (او عينه حجمها n من الوحدات) في فترات زمنية ثابتة ومتساويه (*Fixed Interval Inspection*) .

2) فحص وحده واحده (او n من الوحدات) في فترات زمنية غير متساويه وقصيرة الامد (*Unequally Spaced Inspection*) .

وأن ميل المستوى عند الزمن t ، يعبر عنه كالآتي :

$$\beta_t = \frac{N_t}{N_{t-1}} \beta_{t-1}$$

من أجل بقاء الصيغة الخطية والحصول على نتائج مقبولة ودقيقة ، يفضل ان تكون الفروقات بين (N_t, N_{t-1}) ليست كبيرة ، ولان تغير الميل ممكن في أي لحظة ، يضاف حد الخطأ العشوائي δ_t لمعادلة الميل اعلاه ، وبذلك نحصل على :

$$\beta_t = \frac{N_t}{N_{t-1}} \beta_{t-1} + \delta_t, \delta_t \sim N(0, \sigma^2 r_\beta)$$

بناء على ذلك فان النموذج الحركي الخطي المركب لفترات زمنية غير متساوية ، يكتب رياضيا بالشكل التالي :

$$x_t = \mu_t + v_t, v_t \sim N(0, \sigma^2) \quad (1)$$

$$\begin{aligned} \mu_t &= \mu_{t-1} + \beta_t + \delta_t, \delta_t \sim N(0, \sigma^2 r_\mu) \\ \beta_t &= \frac{N_t}{N_{t-1}} \beta_{t-1} + \Delta_t, \Delta_t \sim N(0, \sigma^2 r_\beta) \end{aligned} \quad (2)$$

حيث أن :

t : دليل الزمن $(t=1,2,\dots)$.

x_t : تمثل قيمة المشاهدة لنوعية المادة المنتجة عند الزمن t .

μ_t : يمثل مستوى او المعدل الحقيقي لنوعية المادة المنتجة المفحوصه عند الزمن t ، ويكون عادة غير معروف بالضبط .

β_t : يمثل الميل الحقيقي لنوعية المادة المنتجة المفحوصه عند الزمن t .

المعادله (1) تسمى معادلة المشاهده ، وتشير الى ان مشاهدات العملية تعتمد عشوائيا على معلمات تلك العملية ، والمعادله (2) تسمى معادلة النظام ، توضح تطور معلمات العملية بمرور الزمن .

كما ان جميع الاخطاء العشوائية v_t, δ_t, Δ_t يفترض انها حوادث مستقلة ، حيث ان :

v_t : يكون ناتج عن الخطأ العشوائي في المشاهدات او المقاييس المستخدمه او كلاهما .

δ_t, Δ_t : يكونان نتيجة الاخطاء العشوائية (Disturbances) في العملية الانتاجية نفسها .

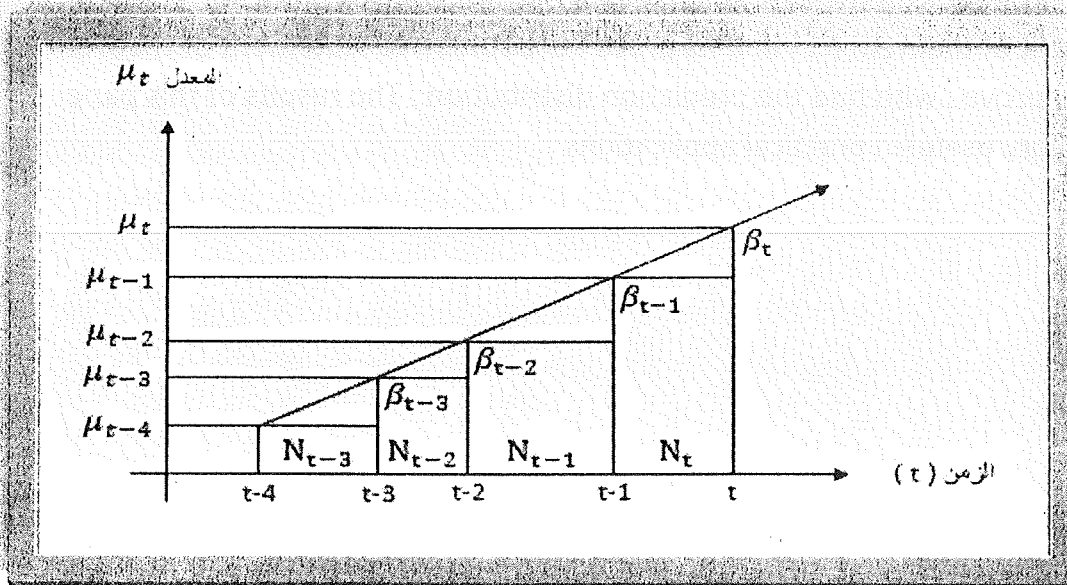
يلاحظ في العلاقة (1) أن قيمة المشاهدة لنوعية المادة المنتجة تساوي معدل النوعية للمادة المنتجة ، بالاضافه الى الخطأ العشوائي في المشاهده نفسها ، ولايمكن ان تتساوى قيمة المشاهده الحقيقيه ومعدل

نمذجة العملية الانتاجية : Modeling of Production Proses

ان نوعية المادة المنتجة في أي عملية انتاجية تكون معرضه للعديد من التغيرات ، نتيجة اسباب عشوائية او غير عشوائية وفي الغالب يحصل نوعان من التغيرات في نوعية المادة المنتجة في اية عملية انتاجية ، الاول تغيرات مفاجئة (Sudden Changes) والذي يسبب تغير مستوى النوعية الى مستوى اخر اسوء بشكل مفاجى ويسمى في السيطره النوعية بـ (Jump) ، تكون هذه التغيرات عادة معنوية (Significant) ، ناتجة عن اسباب فعلية (Assignable causes) في العملية الانتاجية ، مثل (كسر جزء من اجزاء الماكينه ، عامل غير ماهر ، نوعية المواد الاوليه رديئه جدا ... الخ) ، والنوع الثاني تغيرات تدريجية (Gradual Changes) ، أي تغير في الميل (Slope Change) ، عادة تكون هذه التغيرات طفيفة في البدايه ، لكن مع مرور الزمن تتحول الى تغيرات معنوية ، مثل (تآكل جزء او اجزاء الماكينه ، عامل العمر الزمني ، التغير التدريجي في درجة الحرارة والرطوبة ، ... الخ) وتدعى في الاقتصاد بـ (Trend) ، وفي السيطره النوعية بـ (Drift) . في هذا البحث سنأخذ هذه التغيرات بنظر الاعتبار وسنستفيد من النماذج الحركية الخطية لـ Harrison and Stevens (1976) وبالاخص النموذج الحركي الخطي المركب (Combined Dynamic Linear Model) لفترات زمنية غير متساوية ، حيث اخذ المشاهدات في فترات زمنية غير متساوية مع النماذج الحركية الخطية اقترحت من قبل (Smith (1982) .

ان إسقاط القيد (تساوي الفترات الزمنية) في النموذج تزيد من مرونة استخدامه في مجالات الحياة العملية ، فاذا فرض ان الميل في المستوى يسير على نحو ثابت ، فان النمو في المستوى يكون متناسبا مع حجم الفترات الزمنية بين المشاهدات ، انظر الشكل ادناه :

(مع ملاحظة ان $t_i - t_{i-1} \neq t_{i+1} - t_i$, $i=1,2,\dots$)



((النمذجة وسياسة الفحص بدلالة الزمن))

((Modeling And Inspection Policy By The Time))

د . سلطان علي محمد سالم

Dr . sultan ali mohammed salim

قسم الرياضيات – كلية التربية

جامعة الحديدة – اليمن

((الخلاصة))

في هذا البحث تم استخدام النموذج الحركي الخطي المركب لفترات زمنية غير متساوية لإيجاد التوزيعات الاحصائية لسياسة الفحص بدلالة الزمن في حالة الفترات الزمنية غير المتساوية ، وتركز عملنا بصياغة النظرية (1) وإثبات صحتها ، بالإضافة الى صياغة النظرية (2) وإثبات صحتها ، مع ايجاد التوزيع التنبؤي ، والنتائج الرياضية المتوصل اليها في متن هذا البحث تفيدنا في حالة التطبيقات العملية .

((Abstract))

In this paper ,we use the combined dynamic linear model of unequally spaced intervals in order to find the statistical distribution for the inspection policy by the time in the case of unequally spaced intervals . our work was statement of theorem (1) and its prove . In addition to statement of theorem (2) , its prove , with find the prediction distribution , The results of this paper will be very useful in practical applications .

فهرس الإحصاء التطبيقي

٢٥-١٠	النمذجة وسياسة الفحص بدلالة الزمن سلطان على محمد سالم	١
-------	--	---

اللجنة الفنية للمؤتمر

- د. أحمد أمين الشيخ
- أ. نادية حلمي عواد
- أ. زينب مصطفى قرني
- سكرتارية المؤتمر
- سكرتارية المؤتمر

مساعدو اللجنة الفنية للمؤتمر

- أ. جليلة كامل
- أ. سوسن عبد الحليم
- أ. سلمى محمود أحمد
- أ. صفاء محمود
- أ. هالة اسماعيل
- أ. سمر صبري
- أ. ريهام أحمد
- أ. شيماء إسماعيل
- أ. سمير يونان
- أ. مصطفى عباس

العلاقات العامة (التدريب)

- أ. مجدي حسن
- أ. رضا أحمد
- أ. هناء عبد الحليم
- أ. دعاء عبد الفتاح
- أ. أحمد خلف

وحدة التصوير والطباعة والتجليد

- أ. غريب محمود سالم

الإستعلامات

- أ. سلامة عبد الواحد

- د. مرفت حسن غيث
د. محمد بيومي
أ.د. محمد جمال الدين درويش
أ.د. محمد حسن رسمى
د. محمد مصطفى حسن
أ.د. محمد نجيب عبد الفتاح
أ.د. محمود رياض محمود
أ.د. معتز محمد حسنى خورشيد
أ.د. منى عبد الفتاح خليفة
د. مها مصطفى كمال
د. ليلي ناصف عطيه
د. ناجى درويش رمضان
د. نجلاء عبد المنعم مراد
أ.د. نيفين درويش
أ.د. هدى محمد أنسى
د. هشام نبيه المهدى
د. هشام احمد حفننى
أ.د. هشام حسن مخلوف
أ.د. وائل عبد الواحد
د. وليد محمد عرفة
أ.د. يوسف مصطفى محبوب
- معهد الإحصاء - جامعة القاهرة
كلية الحاسبات - جامعة القاهرة
عميد كلية الحاسبات - جامعة الاهرام الكندية
كلية الحاسبات - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
عميد معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
نائب رئيس اتحاد الجامعات العربية
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
كلية الهندسة - جامعة القاهرة
وكيل كلية الحاسبات - جامعة القاهرة
كلية الحاسبات - جامعة القاهرة
قائم بأعمال رئيس قسم علوم الحاسب - معهد الإحصاء -
جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
وكيل كلية الحاسبات - جامعة المنوفية
معهد الإحصاء - جامعة القاهرة
وكيل معهد الإحصاء لشئون خدمة المجتمع وتنمية البيئة -
معهد الإحصاء - جامعة القاهرة

- أ.د. أمانى موسى محمد موسى
 أ.د. الهام شكري محمد
 د. توفيق محمد بونس
 أ.د. حجازي محمد زاهر
 أ.د. ريم بهجت رضا
 د. رمضان عبد الحميد زين الدين
 د. سحر أحمد نيازي
 أ.د. سعد الدين محمد الشيبال
 أ.د. سمير كامل عاشور
 د. سمية العطار
 د. سمية محمود السعدني
 أ.د. سناء العلا حنفى احمد
 أ.د. سيد مشعال
 د. صلاح مهدى محمد
 أ.د. طه مرسي علي
 أ.د. عاطف محمد خليفة
 أ.د. عاطف محمد عبد المنعم
 أ.د. عبد الغنى محمد عبدالغنى
 أ.د. عثمان حجازي محمد
 د. علاء أحمد عبد العزيز
 أ.د. علي علي فهمي
 أ.د. غزال عبد العزيز عامر
 أ.د. فاروق تمام على شعيب
 د. فاطمة عبده الليثي
 أ.د. ليلي عبد العزيز
 أ.د. مجدى سيد مصطفى
 معهد الإحصاء - جامعة القاهرة
 معهد الإحصاء - جامعة القاهرة
 معهد الإحصاء - جامعة القاهرة
 معهد الإحصاء - جامعة القاهرة
 وكيل كلية الحاسبات - جامعة القاهرة
 معهد الإحصاء - جامعة القاهرة
 معهد الإحصاء - جامعة القاهرة
 معهد الإحصاء - جامعة القاهرة
 معهد الإحصاء - جامعة القاهرة
 قائم بأعمال رئيس قسم الإحصاء السكاني - معهد الإحصاء
 - جامعة القاهرة
 وكيل كلية الحاسبات - جامعة القاهرة
 رئيس قسم الإحصاء التطبيقي - معهد الإحصاء - جامعة
 القاهرة
 معهد الإحصاء - جامعة القاهرة
 كلية العلوم - جامعة أسيوط
 معهد الإحصاء - جامعة القاهرة
 معهد الإحصاء - جامعة القاهرة
 مدير المركز الديموجرافي - القاهرة
 كلية الحاسبات - جامعة القاهرة
 معهد الإحصاء - جامعة القاهرة
 عميد كلية الحاسبات - جامعة القاهرة
 معهد الإحصاء - جامعة القاهرة
 معهد الإحصاء - جامعة القاهرة
 معهد الإحصاء - جامعة القاهرة
 معهد الإحصاء - جامعة القاهرة
 معهد الإحصاء - جامعة القاهرة

أمانة المؤتمر والهيئة العلمية

رعاية المؤتمر

تحت رعاية الاستاذ الدكتور رئيس الجامعة أ.د. حسام كامل

رئاسة المؤتمر

رئيس المؤتمر عميد معهد الدراسات والبحوث الإحصائية أ.د. محمد نجيب عبد الفتاح

أمانة المؤتمر

امين عام المؤتمر وكيل معهد الدراسات والبحوث الإحصائية
لشئون خدمة المجتمع و تنمية البيئة أ.د. أماني موسى محمد
أ.د. يوسف مصطفى محجوب

الهيئة العلمية مرتبة أبجدياً

الاسم	الجهة
د.أبو العلا عطيفي	كلية الحاسبات - جامعة القاهرة
د.أحمد أمين الشيخ	معهد الإحصاء - جامعة القاهرة
أ.د. أحمد فؤاد محمد عطية	وكيل معهد الإحصاء لشئون الدراسات العليا والبحوث - رئيس قسم الإحصاء الرياضي - معهد الإحصاء - جامعة القاهرة
أ.د. أحمد حسن يوسف	معهد الإحصاء - جامعة القاهرة
أ.د. الحسينى عبد البر راضى	معهد الإحصاء - جامعة القاهرة

الدكتور/ حسام محمد كامل رئيس الجامعة علي كريم رعايته لمؤتمرنا هذا ودعاه الدائم للعون المادي والمعنوي اللذين كان لها أبعاد الأثر في تنظيم مؤتمرنا هذا. كما أتقدم بالشكر لكل من:

- أكاديمية البحث العلمي
- صندوق الأمم المتحدة للسكان
- المشروع العربي لصحة الأسرة بالجامعة العربية
- جمعية النور الخيرية (مغربى)

أدعو الله العزيز القدير أن يوفقنا جميعا لما فيه الخير لمصرنا العزيزة

والسلام عليكم ورحمة الله وبركاته

عميد المعهد

أ.د. محمد نجيب عبد الفتاح

ويتفاعل المؤتمر بكثير من القضايا والهموم التى يعانى منها مجتمعنا المصرى من خلال ٧ ندوات علمية تتركز حول :

- التعديلات الإدارية لمحافظات مصر ونتائجها السكانية وتتم من خلال شراكة فاعلة من جمعية الديموجرافيين المصريين والجمعية الجغرافية المصرية.
- العشوائيات فى محافظات مصر: الوضع القائم والاسلوب الامثل للتعامل. وتنظمها مركز المعلومات واتخاذ القرار بمجلس الوزراء.
- تحديد الألف قرية الأكثر فقرا - وينظمها مع الجهاز المركزى للتعبئة العامة والإحصاء.
- امراض العيون من المنظور الديموجرافى - وتنظم بالتعاون مع مؤسسة النور الخيرية (مغربى).
- تداعيات الأزمة المالية العالمية على التجارة الخارجية المصرية.

ويتواصل المؤتمر بمحاضرتين عامتين من خيرة علماء مصر بالخارج بل سفرائها من خلال مشروع نقل المعرفة والخبرة اكااديمية البحث العلمى والتى نتشرف بهما وهما:

- الاستاذ الدكتور سمير فريد للحديث عن:
"النمو السكانى وقدرات وأمن الأمم"
- الاستاذ الدكتور سعيد المهدى للحديث عن:

"Active Learning Use of Technology in Teaching Intro Statistics"

السادة الحضور الكريم

بدأت كلمتى بالشكر الواجب .. وفى النهاية تملى على الامانة العلمية لتقديم خالص الشكر والتقدير للجهات والهيئات الداعمة لمؤتمرنا السنوى لهذا العام وعلى رأسها سعادة الأستاذ

الشكر والتقدير الى الأستاذة الدكتورة أمانى موسى والتي تحملت عبء امانة المؤتمر السنوى لهذا العام بكل تفانى واقتدار. والتي حالت ظروفها الصحية حضورها وأنتهز هذه المناسبة بالدعاء لله بأن يمن عليها نعمة الشفاء. واشكر فريق العمل لامانة المؤتمر الذى عمل فى صمت واستمر فى الاداء حتى تحقق حلمنا بل عيدنا السنوى فى أظهار اعمال المؤتمر فى أبهى صورة.

الساده الحضور الكريم

لعل من حسن الطالع ان يواكب مؤتمرنا هذا فى ان يدرج معهدنا فى مشروع الجودة للتعليم ويخطو خطواته الاولى والذى نرى فيه املا كبيرا فى ان يحقق معهدنا رؤيته ورسالته فى أقصر وقت ممكن. فقد انبثقت رؤية المعهد من فكر اساتذة فى :

" التميز فى الأداء الاكاديمي والمهنى لتحقيق الريادة فى المجالات الاحصائية وعلوم الحاسب ونظم المعلومات وبحوث العمليات ليكون احدى المنارات العلمية والبحثية للارتقاء بمستوى أداء الخريج طبقا للمعايير القومية والعالمية".

وبالتالى تملى علينا تلك الريادة من خلال افراد معهد الإحصاء بمنح درجة الدبلوم والماجستير والدكتوراة فى الإحصاء وعلوم الحاسب وبحوث العمليات أهمية مواكبة العصر من تكنولوجيا المعلومات وطرق التعليم والتعلم والتي تضيف الى علوم الإحصاء وجه جديد.

ولعل من المناسب ذكره فى ذلك المقام استضافة معهد الإحصاء ورشة العمل عن التعليم الإلكتروني بالتعاون مع وحدة ضمان الجودة بجامعة القاهرة فى أطار مؤتمرنا السنوى.

أساتذتى وزملائى الأعزاء:

تتعد أنشطة مؤتمرنا السنوى لهذا العام ليشمل جلسات علمية لتشمل ١٢ جلسة علمية تحتوى على (مايقرب من خمسون بحثا) فى الإحصاء السكانى والتطبيقات والرياضى وعلوم الحاسب والمعلومات وبحوث العمليات.

بسم الله الرحمن الرحيم
كلمة المؤتمر السنوى الثالث والأربعين
"الإحصاء وعلوم الحاسب وبحوث العمليات"
٢٢-٢٥ ديسمبر ٢٠٠٨

أساتذتى وزملائى الاعزاء
السيدات والسادة الحضور الكريم

تتعقد فعاليات أنشطة مؤتمرنا السنوي الثالث والأربعين
فى الإحصاء وعلوم الحاسب وبحوث العمليات

تحت شعار
لا تنمية بدون إحصاء

خلصت خبرات السنوات الماضية عن حقيقة دامغة مفادها ان أى خطه أو برنامج أو مشروع لا يمكن انجازه فى غياب قواعد البيانات وبالتالي الإحصاء كأداة لاتخاذ القرارات. وفى هذا السياق فان اعتزازنا بهذا الشعار ينبع من أهمية الدور الذى يقوم به معهدنا العريق فى المشاركة فى تحقيق اهداف التنمية الاقتصادية والاجتماعية والتى تتصدر اولويات اهتمام القيادة السياسية إن لم يكن شاغلها الاول.

وفى البداية اتوجه بالشكر والتقدير لحضوركم الكريم كل باسمه ويمتد هذا الشكر والتقدير إلى اساتذتنا الاجلاء. وأخص بالشكر والتقدير الى الأستاذة الدكتورة الهام شكرى عميد معهدنا السابق والتى وضعت كثير من المعايير ارست بها قواعد العمل التنظيمى للمعهد. كذلك كل



جامعة القاهرة
معهد الدراسات والبحوث الإحصائية



المؤتمر السنوى الثالث والأربعون
للإحصاء وعلوم الحاسب
وبحوث العمليات

إحصاء تطبيقي

ينظمه معهد الدراسات والبحوث الإحصائية
جامعة القاهرة

٢٢-٢٥ ديسمبر ٢٠٠٨