THE UNIVERSITY OF MANITOBA

ON THE INADMISSIBILITY OF THE MLE OF A MATRIX OF POISSON MEANS UNDER A MULTIPLICATIVE MODEL FOR COMPLETE AND INCOMPLETE DATA PROBLEMS

by

© EMMANUEL AWUKU-DARKOH

A THESIS

SUBMITTED TO THE FACULTY OF

GRADUATE STUDIES IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Department of Statistics

WINNIPEG, MANITOBA

July, 1988

ISBN 0-315-48035-1

ON THE INADMISSIBILITY OF THE MLE OF A MATRIX OF POISSON MEANS UNDER A MULTIPLICATIVE MODEL FOR COMPLETE AND INCOMPLETE DATA PROBLEMS

BY

EMMANUEL AWUKU-DARKOH

A thesis submitted to the Faculty of Graduate Studies of
the University of Manitoba in partial fulfillment of the requirements
of the degree of

DOCTOR OF PHILOSOPHY

TABLE OF CONTENTS

ABSTRACT

Let X_{ij} $(i=1,2,\ldots,I;\ j=1,2,\ldots,J)$ be IxJ independent Poisson random variables arranged in a two-way table and let $\theta=\{\theta_{ij}\}$ be the corresponding matrix of means. The simultaneous estimation of the components of θ is examined under a multiplicative model in which $\theta_{ij}=\alpha_i\beta_j\tau$, where $0<\alpha_i<1$, $0<\beta_j<1$, $0<\tau<\infty$, $\Sigma\,\alpha_i=1$, and $\Sigma\,\beta_j=1$. Complete and incomplete tables are considered, the latter case arising when some of the X_{ij}'s are unobservable.

The admissibility of the maximum likelihood estimator (MLE) of θ is considered under a weighted squared-error loss function. For complete tables, it is shown that there exists a class P_D of estimators dominating the MLE. It is also shown that a particular generalized Bayes estimator of θ is a member of P_D, the (improper) prior being uniform on $\theta=(\alpha,\beta,\tau)$, where $\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_I)$, $\beta=(\beta_1,\beta_2,\ldots,\beta_J)$, and the components of α and β satisfy the constraints above. These results extend those in Clevenson and Zidek (1975).

For incomplete tables, the MLE and generalized Bayes estimators of θ cannot be obtained explicitly, and therefore iterative procedures must be used. An estimator of θ, based on an analogue of the expectation maximization (EM) algorithm (utilizing the form of the posterior mean in the complete data problem), is proposed as an alternative (approximation) to the generalized Bayes estimator with respect to the prior mentioned above. This estimator is called the EMGB estimator. Although it is based on an iterative procedure, the EMGB estimator is easier to evaluate than the generalized Bayes

estimator. Explicit expressions for the risks of the MLE and the EMGB estimator cannot be obtained. Therefore the risks of the estimators are compared in a simulation study in the special case of a 3x3 table in which cells are missing according to a balanced incomplete block design.

The simulation results show that the EMGB estimator dominates the MLE over the subset of the parameter space used in the simulations; the potential savings in risk which can be made by using the EMGB estimator instead of the MLE are appreciable over a wide range of the parameter space, particularly when τ is small.

ACKNOWLEDGEMENT

I wish to acknowledge with gratitude the sources of financial support which made this work possible: Ghana Government Scholarship, which fully funded my M.Sc degree program and the first three years of my Ph.D work; Department of Statistics at the University of Manitoba, for providing me a job, which was the main source of income for several years. Thank you Dr. Chan and Dr. Brewster.

Throughout this study, my wife has been a source of strength and emotional support. Lois, for all those sleepless nights, thanks for being there!

I would also like to express my sincere gratitude to my supervisor and chairman of my Advisory Committee, Dr. J. F. Brewster for suggesting the problem, and for his patience and guidance throughout the preparation of this dissertation. To the other members of my Advisory Committee, Drs. K. S. Mount, S. Cheng and H. Kesselman, a big thank you for your interest and helpful suggestions.

CHAPTER I

INTRODUCTION AND SUMMARY

Since its introduction by R. A. Fisher at the beginning of this century, the method of maximum likelihood (ML) has become one of the most widely used procedures in point estimation. ML estimators often possess desirable 'large sample' properties. In particular, if the density function satisfies certain regularity conditions, then the distribution of the ML estimator (MLE) may be approximated by a normal distribution for large n, and the sequence of ML estimators is best asymptotically normal (or BAN; see, for example, Theorem 5.5.3 in Zacks, 1971). Because of the asymptotic normality, an estimate of the precision associated with an MLE can be obtained even if its exact distribution cannot be determined.

In estimating the mean vector of a p-dimensional multivariate normal distribution with known covariance matrix of the form $\sigma^2 I_p$, the sample mean vector is the MLE. It is also the UMVUE and the method of moments estimator. Moreover, under squared-error loss, the sample mean vector is the best location-equivariant estimator and is minimax. For $p \leq 2$, the sample mean vector is also admissible under squared-error loss. For $p=1$, two proofs of the admissibility of the sample mean are given in Lehmann (1983). The proof of admissibility of the sample mean vector for $p=2$ is given in Stein (1956).

In the above problem, Stein (1956) proved the surprising result that the sample mean vector (denoted by $\bar{X}$) is inadmissible under squared-error loss when $p \geq 3$, although he gave no specific example of an estimator which is better than $\bar{X}$. Later, James and Stein (1961) presented an example of an estimator which dominates $\bar{X}$. While the James-Stein estimator dominates $\bar{X}$, it is also inadmissible under squared-error loss. However, it appears that no substantial improvements over the James-Stein estimator are possible (refer to Lehmann, 1983, page 304). Estimators which provide significant improvements over $\bar{X}$ have been presented in Bhattacharya (1966), Baranchik (1970), Strawderman (1971), Bock (1975) and in Berger (1975), among others.

The inadmissibility result in Stein (1956) is surprising because the coordinates of the random vector are independent. Thus, the problem of simultaneously estimating the p components of the mean vector can be considered as p independent problems. It is not intuitively clear how the combining of three or more independent problems can make unacceptable a procedure which is acceptable when applied to each coordinate. Brown (1975) has studied this apparent paradox and has observed that the p problems are not independent for the simple reason that the variances are measured on a common scale, as are the losses.

Stein's work on simultaneous estimation has since been extended along three main fronts. On the first front, researchers have been interested in a search for a wider class of estimators which are better than the MLE of the mean vector of the multivariate normal distribution. For example, Baranchik (1964) considered the case where the components are independently distributed with a common known variance σ^2 and obtained a class of dominating estimators for $p \geq 3$. In the same problem with σ^2 unknown, Baranchik (1970) obtained a class of dominating estimators for $p \geq 5$. In the known variance case, Strawderman (1971) and Berger (1975) extended the class of estimators proposed by Baranchik (1964) and, in addition, obtained proper Bayes (and hence admissible) dominating estimators. Strawderman (1972) showed that proper Bayes dominating estimators do not exist when $p=3$ or 4. More recently, Faith (1978) obtained a class of generalized Bayes estimators which dominate the MLE when $p=3$ or 4. Although the James-Stein estimator and other shrinkage estimators guarantee a decrease in total risk relative to the risk of the MLE, they perform poorly in estimating a component having an unusually large value compared to the other components. Efron and Morris (1971, 1972) considered this problem. They suggested that a 'limited translation' estimator, which is a compromise between the MLE and the James-Stein estimator, limits the risk associated with individual components, while only slightly reducing the overall savings in risk offered by the James-Stein rule.

On the second front, the interest is in finding a class of estimators that improve upon the best location-equivariant estimator in a general location parameter problem. Brown (1966) has done much in this regard and has shown that Stein's result (for the multivariate normal distribution under squared-error loss) is a special case of a general result which is true for a much wider class of underlying distributions and for many other loss functions. Later, Strawderman

(1973) found classes of dominating estimators for location vectors of certain symmetric density functions.

On the third front, the interest is in finding a class of estimators of non-location parameters that are better than the MLE's. In this regard, Clevenson and Zidek (1975) considered the simultaneous estimation of p Poisson means, and found a class of dominating estimators which shrink the MLE toward the origin. Since then, other classes of dominating estimators for a vector of Poisson means have been found by Tsui (1979, 1981), Hwang (1982) and Hudson (1985), among others. In the case of p independent multinomial distributions, Alam (1978) showed that the MLE is inadmissible for $p \geq 2$. More recently, Tsui and Press (1982) considered the general problem of admissibility for the discrete exponential family under several weighted squared-error loss functions and, in particular, obtained a class of dominating estimators for the negative binomial.

In all of these problems the real goal in looking for a class of estimators that dominate the MLE is to search for admissible estimators. In this regard, the search is for estimators that are either proper Bayes or are appropriate limits of Bayes procedures (Sacks, 1963). For example, as mentioned above, Strawderman (1971) obtained a class of proper Bayes (and hence admissible) dominating estimators when $p \geq 5$ in the multivariate normal problem. In the Poisson problem, Clevenson and Zidek (1975)) obtained a class of generalized Bayes dominating estimators. As a consequence, they were able to provide an estimator that is admissible under weighted squared-error loss relative to a wide class of estimators of a particular form.

Many of the admissibility problems that have been examined on the third front have involved the Poisson distribution. In this thesis, some problems concerning the admissibility of the MLE of a matrix of Poisson means will be examined.

Let X_{ij} ($i=1,2,\ldots,I$; $j=1,2,\ldots,J$) be p ($=I\text{x}J$) independent Poisson random variables arranged in a two-way table (of I rows and J columns) and let $\theta=\{\theta_{ij}\}$ denote the corresponding matrix of means. The simultaneous estimation of the components of θ will be considered under a multiplicative model in which $\theta_{ij}=\alpha_i\beta_j\tau$, where $0<\alpha_i<1$, $0<\beta_j<1$, $0<\tau<\infty$, $\sum \alpha_i=1$ and $\sum \beta_j=1$. This multiplicative structure on the means is analogous to the independence model in a two-way contingency table. Complete and incomplete tables will be discussed, the latter case arising when some of the X_{ij}'s are unobservable.

In Chapter III, we will find the MLE and obtain the form of generalized Bayes estimators of θ in the complete data problem.

In Chapter IV, the admissibility of the MLE under a weighted squared-error loss function will be considered for the complete data problem. It will be shown that there exists a class P_D of estimators having uniformly smaller risks than the MLE. It will also be shown that the generalized Bayes estimator of θ with respect to an (improper) uniform prior distribution on the parameters α, β and τ is a member of P_D. These results extend those in Clevenson and Zidek (1975).

In the case of incomplete tables, we will also be interested in searching for an estimator which improves upon the MLE. However, in the incomplete data problem, the MLE and the generalized Bayes estimators cannot be obtained explicitly and iterative methods must be used. Although the MLE can be obtained quite easily by using the expectation-maximization (EM) algorithm, the EM algorithm cannot be used to evaluate the generalized Bayes estimators. Generalized Bayes estimators can be obtained by applying numerical integration techniques, but it is difficult in this problem because of the multi-dimensional nature of the parameter space. Instead, an analogue of the EM algorithm, which utilizes the form of the generalized Bayes (GB) estimator in the complete data problem, will be proposed. In the special case of the uniform prior mentioned above, this estimator will be called the EMGB estimator. Computational details for finding the MLE and the EMGB estimator will be given in Chapter V.

It should be observed that the EMGB estimator is based on the same prior on $\theta=(\alpha,\beta,\tau)$ under which the generalized Bayes estimator in the complete data problem achieves dominance over the MLE. One might conjecture that the EMGB estimator will dominate the MLE in the incomplete data problem. We will use simulation to compare the risks of the MLE and EMGB estimator in the incomplete-data problem. The simulation results, to be discussed in Chapter VI, will confirm that, for the 3x3 incomplete table, the EMGB estimator dominates the MLE over the subset of the parameter space used in the simulations. The simulations will also indicate that the potential savings in risk which can be made by using the EMGB estimator instead of the MLE are appreciable over a wide range of the parameter space, particularly when τ is small. In the complete case, the risks of the MLE and the generalized Bayes estimator under consideration do not depend on α or

β, but only on τ; in the incomplete case, however, the risks of the MLE and the EMGB estimator depend on α, β and τ.

Dempster, Laird and Rubin (1977) (hereafter called DLR) introduced the EM algorithm for computing the maximum likelihood estimates from 'incomplete data'. The EM algorithm will be reviewed in Chapter II. DLR presented the EM algorithm as a numerical method for solving likelihood equations derived from incomplete data, although the procedure can also be used to find the mode of the posterior distribution in a less formal Bayesian context. Some of the theorems underlying the use of the algorithm will be discussed. Four examples will be provided to illustrate some applications of the EM algorithm. Apart from Example 3 (which involves the estimation of the mean of an exponential distribution from which only 'record-breaking' observations are available), all of the examples considered have previously been studied by other authors, although the methods used by these authors are different. It will be shown that computing the ML estimate by means of the EM algorithm may be more convenient than the existing method in the literature in view of the simple form of the estimators in the complete data problems. As far as we know, Example 3 presents a new application of the EM algorithm.

Chapter II

THE EM ALGORITHM

2.0 Introduction

The work done by DLR concerning the EM algorithm rightly deserves the wide recognition it is given in the literature, although the essential ideas underlying the use of the EM algorithm had been recognised previously by many authors in special circumstances. Among them we mention Hartley (1958), Blight (1970), Baum et al (1970), Orchard and Woodbury (1972), and Sundberg (1974). Hartley proposed an iterative procedure for obtaining ML estimates of the parameters in truncated Poisson and negative binomial distributions. The procedure is essentially an implementation of the EM algorithm. While Hartley's scheme was designed specifically for discrete densities belonging to the exponential family, Blight (1970) showed that the basic ideas in Hartley's work could be extended to continuous distributions of the exponential family. Another formulation of the EM algorithm was presented in Orchard and Woodbury (1972), where it was called 'the Missing Information Principle'--- presumably because they were considering the loss in precision associated with the ML estimator of the location parameter θ of a multivariate normal distribution, in the case where the samples have missing values. The striking representation of the incomplete data system of ML equations (2.2.16), which is also equation (2..13) in the DLR paper, had been recognized previously by Baum et al (1970) and Sundberg (1974), who attributed it to unpublished lecture notes of Martin Lof; see DLR for a detailed bibliography.

The main contributions of the DLR paper to numerical optimization theory in statistics are (i) the recognition that the EM algorithm can be formulated in a general way independent of the form of the underlying probability distribution, (ii) the presentation of a unified theory underlying the use of the algorithm dealing with the conditions necessary for the convergence of the iterative procedure and the rate of convergence near a stationary point, and (iii) the provision of a wide range of applications in statistics hitherto unrecognized.

In this chapter, we will present an overview of the EM technique, with particular reference to incomplete data from the exponential family of densities. This task is justifiable for two reasons. Firstly, this overview provides a summary of important applications of the EM algorithm. Secondly, the examples to be presented in this chapter should provide a better understanding of the simulation results presented in Chapter VI, which deal with the application of the EM technique for estimating a matrix of parameters of several Poisson populations using incomplete data from a two-way table.

In Section 2.2, the term 'incomplete data' will be explained in the light of its definition and usage in the literature. Some examples of incomplete data problems are presented. The general formulation of the EM algorithm given in DLR is presented in Section 2.2. Contrary to the general experience with many numerical optimization techniques, the implementation of the EM algorithm is relatively simple when the complete data likelihood equations can be solved exactly. In Section 2.4, we will give four numerical examples of estimation problems using incomplete data that can be solved by the application of the EM algorithm. In all these problems, the ML equations cannot be solved explicitly. A numerical method, such as Newton-Raphson's method or Fisher's method of scoring, must be used to obtain the estimates.

Compared to the Newton-Raphson method or the method of scoring, the EM algorithm is relatively simple to apply. When the ML estimator of θ in the complete data case can be obtained explicitly, the ML estimate in the incomplete data case (via the EM algorithm) is obtained by employing the same functional form of the estimator as in the complete data case. Thus, unlike the Newton-Raphson and the scoring methods, the calculation and inversion of the information matrix are avoided.

One drawback of the EM algorithm (in contrast to the Newton-Raphson method) that has been mentioned by many authors is that it does not automatically provide a means of estimating the information matrix associated with the estimates. This problem, however, has been examined by Hartley and Hocking (1971) and also by Louis (1983). The methods presented by these authors use the notion of observed information (see Efron and Hinkley, 1978; Orchard and Woodbury, 1971). In the procedure presented by Hartley and Hocking, the observed information matrix of the estimates is obtained from a Taylor's expansion, around the ML estimate, of the first derivative of the complete data likelihood function. In the procedure given in Louis (1983), the observed information matrix is obtained from the first and second derivatives of the complete-data likelihood function. It should be noted that the required matrix computations are carried out as a one-step calculation at the termination of the estimation step.

2.1 Incomplete Data

The term 'incomplete data' is often used in the application of the EM algorithm. Roughly speaking, it refers to data in which some observations are missing at random (Rubin, 1976), discarded, or are just not observable (Sundberg, 1976). A broader definition of incomplete data is given in the paper of DLR.

Formally, incomplete data may be defined as follows: let Y denote the sample space induced by the range of a mapping ϕ on a sample space X, which is assumed to be a σ-field of subsets of a k-dimensional Euclidean space R^k ; for every $y \in Y$, there exists $x \in X$ such that $y=\phi(x)$. Suppose that ϕ is many-to-one. Then ϕ^{-1} is a point-to-set map. We shall denote the set $\phi^{-1}(y)$ by $X(y)$. That is, $X(y)=\{x:x \in X, y=\phi(x)\}$. Incomplete data refers to $y \in Y$, which constitutes the observed data. The realization $x \in X$, which cannot be observed directly, but only through y and the map ϕ^{-1}, is called the 'complete' data corresponding to y.

For the point-to-set mapping, ϕ^{-1}, the complete data x for a particular y cannot be uniquely obtained. As a result, a set of sufficient statistics T may exist in X but not in Y. However, when the observed data is y, then T can only be approximated.

Many examples of estimation problems which involve incomplete data are encountered in the literature. Among these we mention problems involving finite mixtures of probability density functions, truncated distributions, and censored samples. In multivariate sampling, it may not be possible to observe all of the entries of the component variables for some of the sampled units. If the 'missing cells' in the data occur as if they were generated by a random process independent of the sampling procedure, then they are said to be missing at random. Rubin (1976) has presented missing at random multivariate data as another example of incomplete data. In this chapter, we will present four examples of estimation problems which involve incomplete data.

2.2 Formulation of the EM algorithm

The following formulation of the EM algorithm follows DLR. The EM algorithm may be viewed as a transformation of the parametric vector space. Formally, the EM algorithm may be defined as a point-to-set, injective mapping $M:\Omega \to \Omega$ on the parameter space Ω. With an assumed initial value of the parameter-vector, the EM algorithm defines a sequence of steps to locate a fixed point of the transformation that maximizes the conditional expected likelihood. In this section we examine this procedure in some detail. The underlying distribution is assumed to be a member of the regular exponential family.

Let the density function of X be $f(x|\theta)$, where $\theta \in \Omega$ is a vector of parameters, and let the density function of y be given by

$$g(y|\theta) = \int_{X} f(x|\theta)\, d\mu(x), \qquad x \in X(y). \tag{2.2.1}$$

The vector of parameters θ is to be estimated by the method of maximum likelihood. This requires maximizing $g(y|\theta)$ over $\theta \in \Omega$. In many statistical problems, maximizing the complete data likelihood function $f(x|\theta)$ is easier than maximizing the incomplete data likelihood function $g(y|\theta)$.

The essential feature of the EM algorithm is that the maximization of $g(y|\theta)$ is replaced by the maximization of $f(x|\theta)$ over $\theta \in \Omega$ in a two-step routine designated as the E-step and M-step. Since $x \in X$ is not observable, $\log f(x|\theta)$ is estimated (hence the E-step) by its conditional expectation given y and the current estimate $\theta^{(k)}$. Then the pseudo-complete data likelihood obtained in the E-step is maximized (hence the M-step) as if it were the observed data likelihood. The process is recursive. The E- and M-steps are performed in sequence until convergence is established.

Let $k(x|y,\theta)$ denote the conditional density of x given y and θ. That is, $k(x|y,\theta)=f(x|\theta)/g(y|\theta)$. Then the log of the likelihood of y may be written as

$$L(\theta')=\log g(y|\theta'),$$

$$=\log f(x|\theta')-\log k(x|y,\theta'), \tag{2.2.2}$$

$$=Q(\theta'|\theta)-H(\theta'|\theta), \tag{2.2.3}$$

where $Q(\theta'|\theta)=E(\log f(x|\theta')|y,\theta)$ and $H(\theta'|\theta)=E\{\log k(x|y,\theta')|y,\theta\}$ are assumed to exist for all pairs (θ',θ). Instead of maximizing $L(\theta')$ directly, the EM algorithm attempts to achieve the same objective by maximizing the function $Q(\theta'|\theta)$. Let $\Omega_M=\{\theta' \in \Omega \mid \theta'=M(\theta),\ \theta \in \Omega\}$.

For $k \geq 1$, and $\theta^{(k)} \in \Omega$, the EM algorithm can be thought of as a mapping $M: \theta^{(k)} \rightarrow \theta^{(k+1)} \in \Omega_M$, defined by the following two steps:

E-step: Determine $Q(\theta|\theta^{(k)})$.

M-step: Choose $\theta=\theta^{(k+1)}$ to be any value $\theta \in \Omega$ which maximizes

$$Q(\theta|\theta^{(k)}).$$

When $f(x|\theta)$ is a member of the exponential family, the E- and M-steps take on very simple forms. In accordance with Lehmann (1959), $f(x|\theta)$ is said to belong to the exponential family of distributions if

$$f(x|\theta)=b(x)\exp(\theta t'(x))/a(\theta), \tag{2.2.4}$$

with respect to a σ-finite measure μ over a Borel set X in the Euclidean space R^k, where θ is a vector of parameters and $t'(x)$ is the transpose of the vector t. The natural parameter space Ω is the set of points θ such that

$$a(\theta)=\int b(x)\exp(\theta t'(x))d\mu(x) < \infty. \tag{2.2.5}$$

Let $f(x|\theta)$ be given by (2.2.4). Starting with an initial value of θ (call it $\theta^{(0)}$, say), the sequence $\theta^{(0)}; t^{(1)}, \theta^{(1)}; t^{(2)}, \theta^{(2)}; \ldots;$ $t^{(k+1)}, \theta^{(k+1)}; \ldots,$ will be recursively produced, where $t^{(k+1)}$ and $\theta^{(k+1)}$ are obtained from the following E- and M- steps:

E-step: Estimate the complete data sufficient statistic t by

$$t^{(k+1)}=E(t|y,\theta^{(k)}). \qquad (2.2.6)$$

M-step: Determine the values of $\theta=\theta^{(k+1)} \in \Omega$ such that

$$t^{(k+1)}=E(t|\theta^{(k+1)}). \qquad (2.2.7)$$

In the E-step, the complete data sufficient statistic $t=t(x)$ is estimated by its conditional mean, given y and the current estimate $\theta^{(k)}$. In the M-step, a new estimate of θ is found by maximizing $\log f(x|\theta)$ as a function of θ given the newly estimated value of t in the E-step. For $f(x|\theta)$ in (2.2.4), the M-step is equivalent to equating the mean of t with the estimate of the complete data, $t^{(k)}$, and solving for θ. The sequence of E and M is repeated until a specified tolerance condition for convergence is satisfied.

In (2.2.4), t is a complete data sufficient statistic. Hence by the Neyman factorization theorem, $f(x|\theta)$ may be expressed as

$$f(x|\theta)/h_1(x) = h_2(t|\theta) = s(t)\exp(\theta t')/C(\theta), \qquad (2.2.8)$$

where $$C(\theta) = \int s(t)\exp(\theta t')d\tau(t) < \infty, \qquad (2.2.9)$$

$h_1(x)$ is independent of t and θ, $h_2(t|\theta)$ is the joint density function of t and s is a function of x only through t. By Lemma 3 of Ferguson (1967), the function $C(\theta)$ in (2.2.9) is an analytic function of θ in the interior of Ω. Hence differentiation w.r.t. θ may be performed under the integral sign, obtaining

$$DC(\theta)=\int ts(t)\exp(\theta t')d\tau(t). \tag{2.2.10}$$

$$[DC(\theta)]/C(\theta)=\int ts(t)\exp(\theta t')/C(\theta)d\tau(t) \tag{2.2.11}$$

$$=E(t|\theta)$$

That is, $D\log C(\theta)=E(t|\theta)$, (2.2.12)

where, in the preceding computations, D denotes the differential operator w.r.t. θ. From (2.2.8), after taking logs, we obtain

$$\log h(t|\theta) = \log s(t) + \theta' t - \log C(\theta). \tag{2.2.13}$$

Hence

$$D\log h(t|\theta)= t - D\log C(\theta). \tag{2.2.14}$$

Since the MLE of θ is obtained by equating the partial derivatives of the logarithm of the likelihood to zero and solving for θ, it follows from (2.2.14) that

$$t=D\log C(\theta), \tag{2.2.15}$$

which, by (2.2.12), becomes

$$t=E(t|\theta). \tag{2.2.16}$$

Thus, if the density function is of the exponential type, then the MLE of θ in the case of complete data can be obtained by equating a sufficient statistic $t(X)$ to the expected value $E(t(X)|\theta)$. This is a well known result which had been recognized earlier by Kale (1964). To obtain a similar result for the MLE of θ in the case of incomplete data, the log of the likelihood function, $g(y|\theta)$ may be rewritten in the form

$$\log g(y|\theta) = \log f(x|\theta) - \log k(x|y,\theta), \tag{2.2.17}$$

where

$$k(x|y,\theta)=[b(x)\exp(\theta t')/a(\theta)]/\{\int_{X(y)} b(x)\exp(\theta t')/a(\theta)\,d\mu(x)\},$$

$$=b(x)\exp(\theta t')a(\theta)/a(\theta|y). \tag{2.2.18}$$

and $$a(\theta|y)=\int_{X(y)} b(x)\exp(\theta t')\,d\mu(x). \tag{2.2.19}$$

Thus, similar to (2.2.8),

$$k(x|y,\theta)/h(x,y) = k_1(t|\theta,y) = l(t)\exp(\theta t')/C(\theta|y), \tag{2.2.20}$$

where $C(\theta|y)=\int_{X(y)} l(t)\exp(\theta t')\,d\mu(x)$,

$h(x,y)$ is independent of t and θ, and $k_1(t|\theta,y)$ is the joint distribution of t given y and $l(t)$ depends on x only through t. It follows from (2.2.8) and (2.2.20) that

$$g(y|\theta) = s(t)\exp(\theta t')/C(\theta)/\{l(t)\exp(\theta t')/C(\theta|y)\}$$

$$= s(t)C(\theta|y)/[l(t)C(\theta)], \tag{2.2.21}$$

$$\propto C(\theta|y)/C(\theta).$$

Hence $$\log g(y|\theta) = \text{constant} + \log C(\theta|y) - \log C(\theta), \tag{2.2.22}$$

and $$D\log g(y|\theta)= D\log C(\theta|y) - D\log C(\theta), \tag{2.2.23}$$

$$= E(t|y,\theta) - (E(t|\theta), \tag{2.2.24}$$

where, in a manner similar to (2.2.12), $D\log C(\theta|y)=E(t|y,\theta)$. Therefore $DL(\theta)=0 \Rightarrow E(t|y,\theta)=E(t|\theta)$, which is the M-step of the EM algorithm. Thus, if the EM algorithm sequence $\{\theta^{(k)}\}$ converges to $\theta^{*} \in \Omega$, then $t^{*}=E(t|y,\theta^{*})=E(t|\theta^{*})$, implying that θ^{*} is the ML estimate of θ.

In practice, it may not be computationally feasible to directly perform the M-step. DLR defined a generalized EM (GEM) algorithm to be a point-to-set map $\theta \rightarrow \theta' \in \Omega_M$, such that $Q(\theta'|\theta) \geq Q(\theta|\theta)$ for all θ'. Since EM is a special case of GEM, $Q(M(\theta)|\theta) \geq Q(\theta'|\theta)$, and $H(\theta|\theta) \geq H(\theta'|\theta)$, for every pair $(\theta',\theta) \in \Omega \times \Omega$. Hence for any sequence $\{\theta^{(k)}\}$ of GEM algorithm (in particular, an EM),

$$L(\theta^{(k+1)}) \geq L(\theta^{(k)}). \tag{2.2.25}$$

For a proof of (2.2.25) see Lemma 1 and Theorem 1 of DLR. The conclusion of equation (2.2.25) is that every sequence of EM iterations generates a monotonically increasing sequence $\{L(\theta^{(k)})\}$ of likelihood functions. If $\{L(\theta^{(k)})\}$ is assumed to be bounded on Ω, then (2.2.25) implies that $\{L(\theta^{(k)})\}$ converges monotonically to L^*, for instance. We want to know if L^* is the global maximum of $\{L(\theta)\}$ over Ω. If not, is it a local maximum or a stationary value? Let θ^* denote the point in Ω at which $L(\theta)$ attains its maximum. Then $DL(\theta^*)=0$, and the Hessian matrix $D^2L(\theta^*)$ is negative definite. Further, if θ^* is unique, it is also the global maximum. We note that the convergence of the sequence $\{L(\theta^{(k)})\}$ to L^* does not imply convergence of the sequence of EM estimates $\{\theta^{(k)}\}$. The problem has been studied by DLR who tried, in their Theorem 2, to provide a set of sufficient conditions for convergence. However Theorem 2 of DLR is incorrect, as has been pointed out by Wu (1983). Boyles (1983) has supplied an alternate set of conditions and provided a counter-example to Theorem 2 of DLR.

Although a global maximization of Q is involved in the M-step, this does not imply the maximization of L=Q-H, since the H term may cause difficulty. Let θ^* denote the limit of a convergent sequence $\{\theta^{(k)}\}$ in Ω_M, and assume that the Hessian matrices $D^{20}Q(\theta'|\theta)$ and $D^{20}H(\theta'|\theta)$ exist and are continuous. Then, by definition of the M-step of the EM algorithm, $-D^{20}Q(\theta^*|\theta^*)$ is non-negative definite. Also, by Lemma 2 of DLR, $-D^{20}H(\theta^*|\theta^*)$ is non-negative definite. However, $-D^{20}L(\theta^*|\theta^*) = -D^{20}Q(\theta^*|\theta^*) + D^{20}H(\theta^*|\theta^*)$, and θ^* may not be a local maximum, and therefore cannot be the global maximum. Theorem 1 of Wu (1983) states the required conditions. Important considerations are that the parameter space must be compact, and that the likelihood function must be continuous and differentiable in the interior of Ω (Wu, 1983). These are stronger conditions than the requirements in Theorem 2 in the DLR paper.

There are other theoretical considerations of the EM algorithm which touch on some practical problems in its implementation. For example, it is reported in the literature (Hasselblad, 1966; Habermann, 1974; Murray, 1977), that if the likelihood function has several local maxima or stationary points, the point of convergence of the EM sequence may depend on the initial starting value. This problem is not specific to the EM procedure, since no optimization technique is guaranteed to converge to a global maximum. However, unlike the

Newton-Raphson's method, a starting value close to a local maximum is not an essential requirement for the EM algorithm. For a complete discussion on this problem, refer to the discussion following the DLR paper.

Another major drawback of the EM algorithm is that the rate of convergence may be slow compared to other iterative techniques. In some problems, however, it is possible to speed up the process by incorporating other devices such as Aitken's acceleration process. These issues are dealt with by Smith, Tweedie and Little in the discussion following the DLR paper. While alternative iterative procedures can be found in many problems, numerous papers on the application of the EM algorithm have appeared in the literature since the work of DLR was published. This trend tends to support the increasing acceptance of the EM algorithm as a viable alternative to other methods.

2.3 Examples

In this section, we present four examples on the use of the EM algorithm. The first example deals with the estimation of the mean and variance of a univariate normal distribution from censored data. This problem is not new; it has been examined by Gupta (1952), and also by Cohen (1970). As an alternative to the methods suggested by these authors, we propose an iterative solution via the EM algorithm. It does not require the preparation of special tables and graphs, and it can be programmed easily on a computer.

The second example deals with a Genetics problem which is examined under a model of normal mixtures. Although the ML estimators of the parameters in mixtures of normal distributions cannot be obtained explicitly, we present a numerical solution via a procedure given in Day (1969), which is essentially the EM algorithm.

In the third example, we will consider the ML estimation of the mean of an exponential distribution from record-breaking observations. As far as we can determine, this example is a new application of the EM algorithm.

Our fourth example deals with the ML estimation of the parameter of a truncated Poisson distribution, a problem previously considered in Hartley (1958) and Cohen (1954), among others. This problem will also be considered from a Bayesian point of view. If one chooses the mode of the posterior distribution as an estimator, then one can use the EM algorithm to compute the estimate from the incomplete data. However, if the mean of the posterior distribution is chosen as the estimator, then the EM algorithm cannot be used to obtain the estimate, although the estimate could be obtained through numerical integration. We will demonstrate in the fourth example that an iterative solution based on an analogue of EM algorithm (utilizing the form of the posterior mean in the complete data problem) is an attractive alternative to the direct computation of the posterior mean through numerical integration. Although this iterative procedure does not yield the posterior mean, it does produce a good approximation.

2.3.1 Example 1: Estimation of the Parameters of a Normal Population from Censored Data

Consider a life-testing experiment in which n units are put on test, and assume that k units have failed up to the censoring time T. Suppose that the lifetimes y_i $(i=1,2,...,n)$ are independently distributed as normal with mean μ and variance σ^2. Without loss of generality, assume that the last (n-k) lifetimes y_s, $(s=k+1,k+2,...,n)$, are the running times of the unfailed units at time T, that is, $y_s>T$, $(s=k+1,...,n)$. The log of the likelihood function may be expressed in the form

$$\log L = Q - k\log\sigma - \sum_{i=1}^{k} [(y_i-\mu)/\sigma]^2/2 + (n-k)\log\Phi(z) \tag{2.3.1}$$

where Q is independent of μ and σ, $z=(T-\mu)/\sigma$, and

$$\Phi(z) = (1/\sqrt{(2\pi)})\int_z^\infty \exp(-x^2/2)\,dx. \tag{2.3.2}$$

The likelihood equations can then be written as

$$\partial/\partial\mu \log L = (n-k)R(z)/\sigma + \Sigma\, [(y_i-\mu)/\sigma^2] = 0, \tag{2.3.3}$$

$$\partial/\partial\sigma \log L = (n-k)zR(z)/\sigma + \Sigma\, (y_i-\mu)^2/\sigma^3 - k/\sigma = 0, \tag{2.3.4}$$

$$= (n^*/\sigma^3)\{\Sigma\, (y_i-\mu)^2/n^* - \sigma^2\}=0, \tag{2.3.5}$$

where

$$R(z) = \exp(-z^2/2)/\{\int_z^\infty \exp(-x^2/2)\,dx\}, \tag{2.3.6}$$

$$= \phi(z)/\Phi(z), \tag{2.3.7}$$

and $$n^* = (n-k)zR(z) - k. \tag{2.3.8}$$

Explicit solution of the likelihood equations is not possible. This problem has been considered by Gupta (1952), and Cohen (1970). Both authors used iterative procedures to obtain the MLE. However, both procedures utilized certain graphs and tables designed specifically for the normal density. Thus the procedures cannot be applied directly to censored data from non-normal distributions, whereas the EM algorithm can be used in these situations.

An iterative solution using the EM algorithm of DLR will now be described. Treating the k observed failure times as incomplete data, the 'missing' data consists of the unobserved failure times of the (n-k) units which were still operating at time T. Let Y_s denote the time-to-failure of unit s ($s \geq k+1$) that would have been observed. Then by the well known result (Johnson and Kotz, 1967), the expected value of Y_s given that $Y_s > T$, is

$$E(Y_s | Y_s > T) = \mu + \sigma R(z). \tag{2.3.9}$$

Starting with initial trial values of μ and σ (call these $\mu^{(0)}$ and $\sigma^{(0)}$), let $\mu^{(t)}$ and $\sigma^{(t)}$ denote the estimates of μ and σ at the t-th cycle of the EM algorithm. The estimates at the (t+1)-th cycle are given by the following two steps:

E-step: Obtain the 'complete' data $x=(x_1, x_2,...,x_n)$ by estimating the unobserved failure times by their conditional means given T, $\mu^{(t)}$ and $\sigma^{(t)}$, that is

$$
\begin{aligned}
x_i^{(t+1)} &= y_i, \qquad i=1,2,...,k \\
&= \mu^{(t)}+\sigma^{(t)}R(z^{(t)}), \qquad i=k+1,...,n.
\end{aligned} \tag{2.3.10}
$$

M-step: Obtain refined estimates of μ and σ, treating the $x^{(t)}$ as observed complete data:

$$\mu^{(t+1)}= \Sigma\, x_i^{(t)}/n, \tag{2.3.11}$$

$$(\sigma^2)^{(t+1)}= \sum_{i=1}^{k} [x_i^{(t)}- \mu^{(t)}]^2/n*^{(t)}, \tag{2.3.12}$$

where $$n*^{(t)} =(n-k)z^{(t)}R(z^{(t)}) - k, \tag{2.3.13}$$

and $$z^{(t)} = [T-\mu^{(t)}]/\sigma^{(t)}. \tag{2.3.14}$$

Repeated application of the E- and M- steps yields a sequence $\{\theta^{(i)}\}$, $\theta^{(i)}=(\mu^{(i)},\sigma^{(i)})$, which converges to the ML estimate of θ. The EM algorithm can be implemented without the use of the special tables or graphs, as presented by Gupta (1952). Hence the EM procedure can be used to estimate the parameters of other density functions using censored data.

<u>**Numerical Example**</u>: (see Gupta (1952) for details). 300 electric bulbs are put on test. The distribution of the failure times y_i of 119 bulbs that burned out up to time T hours are given in Table 2.1, which is a

slightly modified version of the distribution in Gupta (1952). The failure times in Gupta's paper were grouped into time intervals, whereas in Table 2,1 the failure times y_i are the computed midpoints of the class intervals used in Gupta's paper.

Table 2.1

Distribution of the Failure Times of 300 Electric Lamps in a Type II Censored Experiment. Failure times shown are for the Lamps which failed.

Failure Times y_i	**Number of Failures**
975	2
1025	2
1075	3
1125	6
1175	7
1225	12
1275	16
1325	20
1375	24
1425	27

Since the ungrouped data were not provided in Gupta's paper, the computed midpoint values, y_i, were taken to be the observed values. The termination time T was also not specified in Gupta's paper, although it was required for computing the estimates of the parameters; we assumed in our computation that the value of T is 1450 hours, which is the upper limit of the highest class interval '1400-1450' in Table 7 of Gupta's paper.

The source code of the Fortran program which can be used to perform the required computation of the ML estimates of μ and σ is given in Appendix A1. The program is designed for use on an IBM PC. In addition, the compiling should be done with the Microsoft Fortran Compiler Version 3.0, although any compiler which supports the full ANSI Fortran77 convention can also be used.

With initial trial values of $\mu^{(0)}=1500$ and $\sigma^{(0)}=1200$, convergence was achieved after 49 iterations. The results are summarized in Table 2.2. The stopping rule for convergence is that $|\mu^{(i)}-\mu^{(i-1)}| \leq 10^{-5}$ and $|\sigma^{(i)}-\sigma^{(i-1)}| \leq 10^{-5}$, where the superscripts denote the number of iterations. Thus the final estimates were $\hat{\mu}=1501.94$ and $\hat{\sigma}=202.02$, which are in close agreement with the values $\hat{\mu}=1502.05$, and $\hat{\sigma}=202.06$ obtained by Gupta (1952). We would not expect to get the same estimates as Gupta (1952) obtained because we did not have access to the original ungrouped data. We note parenthetically that, in the above computation, the EM algorithm could have been used to handle the problem caused by the absence of ungrouped data. The procedure would be similar to that described in Section 3.0

Table 2.2

Estimating the Mean and Standard Deviation of Normal Distribution from Censored Sample: Iterative Solution via the EM Algorithm

Iteration	$\hat{\mu}^{(i)}$	$\hat{\sigma}^{(i)}$
0	1500.000	1200.0000
1	1863.017	231.6831
2	1981.174	555.5859
3	1651.410	628.1521
4	1813.353	303.5665
5	1743.284	457.2563
6	1650.663	383.7957
7	1690.049	303.9314
8	1628.417	339.3252
9	1604.922	284.5963
10	1601.368	270.6659
11	1569.231	267.2572
12	1561.781	241.6117
13	1550.848	239.4721
14	1536.776	230.2853
15	1532.698	220.9252
16	1524.509	219.6593
17	1518.991	213.4017
18	1516.194	210.7825
19	1511.767	209.4138
20	1509.787	206.4547
21	1507.983	205.9447
22	1506.063	204.8382
23	1505.354	203.7977
24	1504.335	203.6972
25	1503.656	203.0250
26	1503.348	202.7821
27	1502.855	202.6879
28	1502.660	202.3750
29	1502.490	202.3610
30	1502.290	202.2627
32	1502.142	202.1703
34	1502.063	202.0818
36	1501.998	202.0449
38	1501.965	202.0417
40	1501.956	202.0325
42	1501.949	202.0238
44	1501.941	202.0207
46	1501.940	202.0209
48	1501.940	202.0208
49	1501.940	202.0208

2.3.2 Example 2: Estimating the Parameters of a Normal Mixture

Maximum likelihood estimation of parameters in a finite mixture of distributions does not admit closed-form solutions; for a given sample, the likelihood equation is non-linear in the parameters to be estimated. Therefore an iterative technique must be used, the Newton-Raphson method or Fisher's scoring technique being two of the most common in applications. In contrast, Day (1969) suggested an iterative procedure for estimating the parameters of mixtures of k-dimensional normal densities. The work of DLR shows that Day's method is a special case of the EM algorithm. In this example, the EM technique will be used to estimate the parameters μ_2 and σ^2 in a mixture of two univariate normal densities with means μ_1 and μ_2 and common variance σ^2. We will find the ML estimates of μ_2 and σ^2 in the special case when μ_1 and the mixing proportion p are known.

It may seem strange and artificial to assume a partial knowledge of the parameters. However, the motivation for this problem came from a study in an agricultural experiment which was conducted to investigate a problem in genetics.

Let $y=(y_1, \ldots, y_n)$ denote a random sample of n observations from a population with the density function

$$g(y|p,\mu,\sigma) = pf(y|\mu_1,\sigma) + (1-p)f(y|\mu_2,\sigma), \tag{2.3.15}$$

where

$$f(y|\mu_i,\sigma) = (\sigma\sqrt{(2\pi)})^{-1}\exp[-(y-\mu_i)^2/(2\sigma^2)], \quad i=1,2. \tag{2.3.16}$$

That is, the underlying sub-populations are $N(\mu_1,\sigma^2)$ and $N(\mu_2,\sigma^2)$ distributions, respectively. The density (2.3.15) is called a mixture of univariate normal densities, with mixing proportions p:1-p. Assume that the parameters p and μ_1 are known. We require the ML estimates of μ_2 and the common variance σ^2.

The likelihood function may be written as

$$L(\mu_2,\sigma|y) = \prod_j g(y_j|p,\mu,\sigma) \qquad (2.3.17)$$

$$=\prod_j \{pf(y_j|\mu_1,\sigma) + (1-p)f(y_j|\mu_2,\sigma)\} \qquad (2.3.18)$$

so that the log of the likelihood function is

$$l(\mu_2,\sigma|y) = \log L(\mu_2,\sigma|y) = \sum_j \log g(y_j|p,\mu,\sigma). \qquad (2.3.19)$$

The ML estimators of μ_2 and σ^2 are the solutions of the equations

$$\partial/\partial\mu_2 l(\mu,\sigma|y)= \sum_j \{p_2 f(y_j|\mu_2,\sigma)/g(y_j|p,\mu,\sigma)\}[(y_j-\mu_2)/\sigma^2]=0, \qquad (2.3.20)$$

$$\partial/\partial\sigma l(\mu,\sigma|y)=\sum_i \sum_j p_i f(y_j|\mu_i,\sigma)\{(y_j-\mu_i)^2 - \sigma^2\}/[g(y_j|p,\mu,\sigma)\sigma]=0, \qquad (2.3.21)$$

where p_1=p and p_2=1-p. Equations (2.3.20) and (2.3.21) may be rewritten in the form

$$\sum_j \mu_2\{p_2 f(y_j|\mu_2,\sigma)/g(y_j|p,\mu,\sigma)\} = \sum_j y_j\{p_2 f(y_j|\mu_2,\sigma)/g(y_j|p,\mu,\sigma)\}, \qquad (2.3.22)$$

and

$$\sum_i \sum_j p_i (y_j-\mu_i)^2 f(y_j|\mu_i,\sigma)/g(y_j|p,\mu,\sigma)=\sum_i \sum_j p_i\sigma^2 f(y_j|\mu_i,\sigma)/g(y_j|p,\mu,\sigma). \qquad (2.3.23)$$

It should be noted that $f(y|\mu_i,\sigma)$ and $g(y|p,\mu,\sigma)$ are non-linear functions of the unknown parameters μ_2 and σ; it is not possible to solve the above system of equations explicitly. The following procedure may be used to obtain an iterative solution.

Let $W(i,j)=\{p_i f(y_j|\mu_i,\sigma)/g(y_j|p,\mu,\sigma)\}$. (2.3.24)

Then, (2.3.22) and (2.3.23) may be rewritten respectively as

$$\mu_2=\sum_j y_j W(2,j)/[\sum_j W(2,j)], \qquad (2.3.25)$$

and $$\sigma^2=\sum_i\sum_j W(i,j)[y_j-\mu_i]^2/\sum_i\sum_j W(i,j). \qquad (2.3.26)$$

The EM algorithm is directed at solving (2.3.25) and (2.3.26) recursively. Choose trial values for μ_2 and σ; call these $\mu_2^{(0)}$ and $\sigma^{(0)}$. Then the sequence $W(i,j)^{(1)}$, $\mu_2^{(1)}$, $\sigma^{(1)}$; $W(i,j)^{(2)}$, $\mu_2^{(2)}$, $\sigma^{(2)}$; ...,; $W(i,j)^{(t+1)}$, $\mu_2^{(t+1)}$, $\sigma^{(t+1)}$,... will be obtained by the following two steps:

E-step: $$W^{(k+1)}(i,j) = p_i f^{(k)}(y_j|\mu_2,\sigma)/g^{(k)}(y|p,\mu,\sigma). \qquad (2.3.27)$$

M-step: Using $W^{(k+1)}(i,j)$, obtain

$$\mu_2^{(k+1)} = \sum_j y_j\, W^{(k+1)}(2,j)/\{\sum_j W^{(k+1)}(2,j)\}, \qquad (2.3.28)$$

and

$$(\sigma^2)^{(k+1)} = \sum_i\sum_j (y_j-\mu_i^{(k+1)})^2 W^{(k+1)}(i,j)/\{\sum_i\sum_j W^{(k+1)}(i,j)\}. \qquad (2.3.29)$$

In the E-step, estimates of the weights W(i,j) are computed using the 'incomplete' data from the mixture and the current (trial) values of the parameters. In the M-step, new estimates are obtained by adjusting the current estimates of the parameters. The entire process is recursively executed until a specified level of tolerance is achieved.

Numerical Example: A sample of 200 observations was generated from a mixture of two univariate normal populations with p=7/8. Population 1 has the $N(0,\sigma^2)$ distribution. Population 2 has the $N(2.0,\sigma^2)$ distribution, and $\sigma^2=1.0$. To illustrate the EM procedure, we treat μ_2 and σ^2 as unknown parameters so that only μ_1 and p are known. The source code of a WATFIV program to estimate μ_2 and σ^2 is given in Appendix A2. By a choice of $\mu_2^{(0)}=10.0$, $\sigma^{(0)}=50.0$, the iteration converges after 66 cycles, yielding $\hat{\mu}_2=1.53$, $\hat{\sigma}^2=1.10$. The stopping rule for convergence was that $|\mu^{(i)}-\mu^{(i-1)}| \leq 10^{-6}$ and $|\sigma^{(i)}-\sigma^{(i-1)}| \leq 10^{-6}$, where the superscripts denote the number of iterations. The results of the iterations are summarized in Table 2.3 below.

Table 2.3

Use of the EM Algorithm to obtain the ML Estimates of the Parameters of a Mixture of Normal Distributions

Iteration (i)	$\mu_2^{(i)}$	$\sigma^{(i)}$	$-\ l(\mu_2^{(i)},\sigma^{(i)}\mid y)$
1	0.148	3.664	11.147
2	0.157	1.223	99.854
3	0.280	1.221	99.941
4	0.388	1.217	100.397
5	0.487	1.213	100.970
6	0.576	1.207	101.705
7	0.659	1.202	102.496
8	0.736	1.196	103.352
9	0.809	1.190	104.259
10	0.877	1.184	105.205
15	1.161	1.153	110.135
20	1.351	1.128	114.317
30	1.501	1.107	118.093
40	1.527	1.103	118.786
50	1.530	1.102	118.885
60	1.531	1.102	118.899
66	1.531	1.102	118.900

2.3.3 Example 3. Estimating the Mean of an Exponential Distribution from Record-Breaking Observations

Let $X_1,\ldots,X_n$ denote independent, identically distributed random variables, each drawn from the exponential distribution with density function

$$f(x|\theta) = (1/\theta)\exp(-x/\theta), \quad x > 0. \tag{2.3.30}$$

For x>0, the cumulative distribution function of X is then given by

$$F(x|\theta)=1-\exp(-x/\theta). \tag{2.3.31}$$

Suppose that only selected X's can be observed. Indeed X_i is observable only if it is greater than every preceding X. The observed data then consists of $Y_1, Y_2,\ldots, Y_n$, where $Y_i=\max(X_1,X_2,\ldots,X_i)$, $(i=1,2,\ldots,n)$. To fix ideas, let R denote the number of distinct Y's, and $Z_1 < Z_2 < Z_3,\ldots,< Z_R$, the order statistics of the distinct values of Y's. In what follows, it will be convenient to denote the observed data by the sequence $\{Z_1,K_1,\ldots,Z_R,K_R\}$, where K_i $(i=1,2,\ldots,R)$ is the number of times Z_i occurs in the sample. By designating Z_i as the i-th record, K_i, $(I=1,2,\ldots,R-1)$, may be interpreted as the number of trials required to break the record Z_i. K_R is then the number of attempts made to break the current record Z_R.

The conditional density g of Z_i, given $Z_1,K_1,\ldots,Z_{i-1},K_{i-1}$, depends on Z's and K's preceding Z_i only through Z_{i-1}, and hence it is given by

$$g(z_i|Z_{i-1}=z_{i-1}) = f(z_i|\theta)/[1-F(z_{i-1})|\theta], \quad z_{i-1} < z_i < \infty \tag{2.3.32}$$

Also, the distribution of the number of trials K_i $(i=1,\ldots,R)$, given $Z_1,K_1,\ldots,Z_i$, depends only on the value of Z_i (Samaniego and Whitaker, 1984), and is given by the geometric distribution with the density function

$$P(K_i=k_i|Z_i=z_i) = [1-F(z_i|\theta)][F(z_i|\theta)]^{k_i-1} \qquad k_i=1,2,.. \tag{2.3.33}$$

Hence the likelihood function of the observations $\{Z_1,K_1,\ldots,Z_R,K_R\}$ is

$$L(\theta;z,k)$$

$$= \prod_{i=1}^{r} \{f(z_i|\theta)/[1-F(z_{i-1}|\theta)]\}[1-F(z_i|\theta)][F(z_i|\theta]^{k_i-1}, \tag{2.3.34}$$

where $z_0=0$. Using the expressions in (2.3.32), and (2.3.33), we obtain

$$L(\theta|z,k) = \theta^{-r}\prod_{i=1}^{r}\exp(-2z_i/\theta)\exp(-z_{i-1}/\theta)\prod_{i=1}^{r}[1-\exp(-z_i/\theta)]^{k_i-1},$$

$$= \theta^{-r}\exp(-2z_r/\theta)\exp(-\sum_{i=1}^{r-1}z_i/\theta)\prod_{i=1}^{r}[1-\exp(-z_i/\theta)]^{k_i-1}. \tag{2.3.35}$$

The maximum likelihood estimator, $\hat{\theta}$, is given by

$$\partial/\partial\theta[\log L(\theta|z,k)] = \sum_{i=1}^{r-1} z_i/\hat{\theta}^2 + \sum_{i=1}^{r} z_i(k_i-1)/\{\hat{\theta}^2[\exp(z_i/\hat{\theta})-1]\}$$

$$+ (2z_r/\hat{\theta} - r)/\hat{\theta} = 0. \tag{2.3.36}$$

It is not possible to solve for $\hat{\theta}$ in (2.3.36) explicitly. We now demonstrate the use of the EM algorithm to obtain an iterative solution. Treating $\{y_1,y_2,\ldots,y_n\}$ as incomplete data from (2.3.30), we define the complete data to be the set of n realizations $(x_1,\dots,x_n)$ of X. If all the X_i were observable, the ML estimator of θ would be given by the sample mean, $\hat{\theta}= \Sigma X_i/n$. The EM algorithm first estimates x_i by the conditional expectation of X_i given the current estimate of θ and z_i.

Indeed, let $\theta^{(0)}$ denote an initial trial value of θ and assume that after t iterations the estimate of θ is $\theta^{(t)}$. Then the (t+1)-th iteration is defined as follows:

E-step: $$x_i^{(t+1)} = E[X_i|X_i\le z_i,\theta^{(t)}], \tag{2.3.37}$$

$$=\theta^{(t)}[1-(1+z_i/\theta^{(t)})[1-F(y_i,\theta^{(t)})]/F(y_i,\theta^{(t)}), \tag{2.3.38}$$

M-step: $$\theta^{(t+1)} = [\sum_i^r z_i+(k_i-1)E(X_i|X_i\le z_i,\theta^{(t)})]/n, \tag{2.3.39}$$

$$=[\sum_i^r z_i+(k_i-1)x_i^{(t)}]/n, \tag{2.3.40}$$

where $F(y_i,\theta^{(t)})=1-\exp(-y_i/\theta^{(t)})$. The sequence of E- and M- steps is repeated until the sequence of estimates $\{\theta^{(t)}\}$ converges.

Numerical Example. A sample of size 50 was generated from the exponential distribution with mean θ=1.0. The relevant statistics for computing the ML estimate of θ are the following:

Z_i	1.70	2.11	2.39	3.00
K_i	4	5	5	36

The computations were carried out on an IBM PC; the required Fortran program, which includes a routine for generating random numbers from the exponential distribution, is listed in the Appendix **A3**. Although the estimation of θ does not require any data generation algorithm, we chose (in this example) to use computer generated data. It should be observed that the usual algorithms for generating uniform random numbers on the mainframe are unsuitable for the compilers available for use on micro-computers because of round-off errors. The uniform random number generator GGUBS in the IMS library, for example, utilizes the multiplicative congruential modulo $2^{31}-1$ integer arithmetic; the required degree of accuracy cannot be realized on a

Table 2.4

Using the EM algorithm to obtain the Estimate of the Parameter θ of the Exponential Distribution from Record Breaking Observations: EM Iterations

cycle (i)	$\theta^{(i)}$
1	1.406
2	1.061
3	0.953
4	0.909
5	0.888
6	0.879
7	0.873
8	0.871
9	0.870

micro-computer with the compilers available. For any number produced that is less than 5.9499 x 10^{-5}, the next number is simply 16807 times as much; similar problems occur for numbers close to one. As a solution, we used the random number generator proposed by Wichmann and Hill (1982), who have studied this problem. Their algorithm utilizes three simpler multiplicative congruential generators to produce numbers rectangularly distributed between 0 and 1.

With the choice of the trial values $\theta^{(0)}$=10.0, the EM iteration, given by (2.3.38) and (2.3.40), converges after 9 cycles. The condition for convergence is that $|\theta^{(t+1)}-\theta^{(t)}| \leq 10^{-6}$. The estimates (rounded off to 3 significant figures) at each cycle are given in Table 2.4. Hence, by the EM algorithm, $\hat{\theta}$=0.870.

2.3.4 Example 4: Estimation of the Parameter of a Poisson Population Truncated at Zero with a Single Observation.

In this section, we will consider how to obtain the ML estimate of the parameter θ of a Poisson population which is truncated at the point zero. We will also obtain the posterior mode and posterior mean estimators of θ.

In particular we will use this example to illustrate the fact that the EM algorithm's version of the posterior mean estimate (henceforth referred to as the EMGB estimate) of θ does not differ from the true posterior mean estimate by a significant amount. The difference also decreases as the parameter value gets farther away from zero.

Let X denote a Poisson random variable with mean θ. The probability density function of X is then given by

$$f(x)=\theta^{x}\exp(-\theta)/x!, \qquad x=0,1,2,.... \tag{2.3.41}$$

Hence $\quad f(0)=\exp(-\theta).$ (2.3.42)

The probability density function g(x) of X truncated at X=0, is obtained by dividing f(x) in (2.3.41) by [1-f(0)]. It follows that

$$g(y)=\theta^{y}\exp(-\theta)/\{y![1-\exp(-\theta)]\}, \qquad y=1,2.... \tag{2.3.43}$$

Consider a random sample y of size n observations from the density function in (2.3.43). Let n_y denote the frequency of the class X=y, and $N=n+n_o$, the total (complete) sample size, whereby n_o denotes the cell frequency associated with the unobservable outcome X=0. In order to estimate n_o, we may regard y as being a realization from a sequence of complete samples of fixed size N. Then n is a random variable whose distribution, given N, is binomial with parameters N and $p=[1-\exp(-\theta)]$. The log of the likelihood functions of n and y are respectively of the form

$$L_1^{*}(\theta,N) = \log L_1(\theta,N) = \log\binom{N}{n} + n\log p + (N-n)\log(1-p), \tag{2.3.44}$$

and
$$L_2^{*}(\theta,y) = \log L_2(\theta,y) = \sum_i^n y_i\log\theta - \sum_i^n \log(y_i!) - \sum_i^n \log[\exp(\theta)-1]. \tag{2.3.45}$$

Let $\hat{\theta}$ and $\hat{N}$ denote the ML estimates of θ and N respectively. Differentiating (2.3.44) w.r.t. θ and (2.3.45) w.r.t. p and equating the results to zero, we obtain

$$\partial/\partial p L_1^{*}(\theta,N) = (n/p)-[(N-n)/(1-p)]. \tag{2.3.46}$$

and $\partial/\partial\theta L_2^*(\theta,N) = \sum_i \{y_i/\theta-\exp(\theta)/[\exp(\theta)-1]\}.$ (2.3.47)

The solutions $\hat{\theta}$ and $\hat{N}$ must satisfy simultaneously

$$\bar{y} = \hat{\theta}/[1-\exp(\hat{\theta})], \qquad (2.3.48)$$

and

$$\hat{N} = n/[1-\exp(-\hat{\theta})], \qquad (2.3.49)$$

where $\bar{y}$ denotes the sample mean of the truncated data. The expression (2.3.48) was obtained by Cohen (1954) under the assumption that the truncated sample is of fixed size n. One could find the solution of (2.3.48) by means of graphical techniques, such as the one used in Cohen (1954) which utilizes a chart of the function $g(\theta)=\theta/[1-\exp(-\theta)]$.

Alternatively, the EM algorithm can be applied to this problem. We treat the realizations y as incomplete data from the density (2.3.43), with n_o (the number of zeros) being the missing data. For a 'complete' data sample of size N, the ML estimate of θ is simply the sample mean, that is

$$\hat{\theta} = \sum_i x_i n_i/(n+n_o). \qquad (2.3.50)$$

The EM algorithm estimates n_o (in the E-step). The maximum likelihood estimation of θ (M-step) then proceeds in the usual way through (2.3.50). Let $\theta^{(k)}$ denote the estimate of θ obtained after k cycles. The EM algorithm at the (k+1)-st iteration consists of the two steps:

E-step: Obtain estimates of n_o by its conditional mean given n and $\theta^{(k)}$, namely,

$$n_o^{(k)} = n\exp(-\theta^{(k)})/[1-\exp(-\theta^{(k)})]$$
$$= n/[\exp(\theta^{(k)})-1]. \qquad (2.3.51)$$

M-step: Treating $N^{(k)} = (n+n_o^{(k)})$ as the complete data sample size, find the ML estimate of θ as:

$$\theta^{(k+1)} = \sum_i x_i n_i / [n+n_o^{(k)}]. \qquad (2.3.52)$$

The E- and M- steps are repeated until the required condition of convergence is achieved. The process is initiated by assigning an arbitrary initial trial value to θ. This value is designated as $\theta^{(0)}$.

It should be observed that the EM algorithm was proposed by DLR as a means of obtaining the ML estimate of θ. In a Bayesian framework, the mode of a posterior distribution is sometimes used as a Bayes estimator although it does not correspond to a particular loss function in a formal decision theoretic framework. The EM algorithm can be used in incomplete data problems if one wishes to find the mode of the posterior distribution corresponding to a specified prior on θ. Sometimes a Bayesian would consider the posterior mode as an acceptable compromise in situations where the Bayes estimator (for example, the posterior mean) is difficult to evaluate. Laird (1981) and others have considered the use of the posterior mode in this context.

Under a squared-error loss, the Bayes estimator of θ is obtained by taking the ratio of integrals. This does not involve any maximization in the sense of the EM algorithm (i.e., the M-step of the

EM algorithm). In this regard, the EM procedure does not yield the Bayes estimate of θ.

Although the EM algorithm cannot be used to obtain the mean of the posterior distribution under a given prior, we will now demonstrate with the above example that the EM-derived estimate is likely to be close to the Bayes estimate. Suppose that (in the above problem) we wish to obtain the Bayes estimator (i.e., the mean of the posterior distribution) of θ under a squared-error loss. We assume that the prior distribution of θ is the Gamma distribution G(a,b) with probability density function

$$h(\theta|a,b)=\theta^{a-1}\exp(-\theta/b)/[\Gamma(a)b^{a}], \qquad \theta > 0. \qquad (2.3.53)$$

Then, from (2.3.43), the posterior distribution of θ, given the single observation y, is given by

$$g(\theta|y) \propto \theta^{(y+a-1)}\exp[-\theta(1+1/b)]/[1-\exp(-\theta)], \qquad \theta > 0. \qquad (2.3.54)$$

Hence the Bayes estimator of θ under a squared-error loss is given by

$$E(\theta|y)=\int \theta h(\theta|a,b)d\theta, \qquad (2.3.55)$$

$$=\int\theta^{(y+a)}\{\exp(-\theta(1+1/b))\}/[1-\exp(-\theta)]/h^*(a,b,y), \qquad (2.3.56)$$

where $$h^*(a,b,y)=\int\theta^{(y+a-1)}\{\exp(-\theta(1+1/b))\}/[1-\exp(-\theta)]d\theta, \qquad (2.3.57)$$

and the integration is from 0 to ∞. To evaluate the right hand side of (2.3.57), we will first obtain a Taylor's expansion of $1/(1-\exp(-\theta)$ about $\theta=0$, which is given by

$$[1-\exp(-\theta)]^{-1} = \sum_{k=0}^{\infty} \exp(-k\theta). \tag{2.3.58}$$

Hence $$E(\theta|y) = \sum_{k=0}^{\infty} \left[\int_0^{\infty} \theta^{(y+a)} \exp(-\theta(k+1+1/b))\,d\theta\right]/h^*(a,b,y), \tag{2.3.59}$$

$$= \sum_{k=0}^{\infty} \Gamma(y+a+1)(k+1+1/b)^{-(y+a+1)} \Big/ \sum_{k=0}^{\infty} \Gamma(y+a)(k+1+1/b)^{-(y+a)},$$

$$= (y+a)\ \Phi(y,a,b), \tag{2.3.60}$$

where $$\Phi(y,a,b) = \sum_{k=0}^{\infty} \Gamma(k+1+1/b)^{-(y+a+1)} \Big/ \sum_{k=0}^{\infty} \Gamma(k+1+1/b)^{-(y+a)}, \tag{2.3.61}$$

and $$h^*(a,b,y) = \sum_{k=0}^{\infty} \left[\int_0^{\infty} \theta^{(y+a-1)} \exp(-\theta(k+1+1/b))\,d\theta\right]. \tag{2.3.62}$$

Thus the Bayes estimate (posterior mean) can easily be obtained by evaluating numerically the quantity $\Phi(y,a,b)$ for specified values of a, b and y.

The EM version of the posterior mean estimate in the above problem consists of the following two steps:

E-step: Obtain the estimate of n_o by proportional fitting, that is

$$n_o^{(k)} = [\exp(-\theta^{(k)})-1]^{-1}. \tag{2.3.63}$$

M-step: $$\theta^{(k+1)} = (y+a)/[N^{(k)}+1/b], \tag{2.3.64}$$

where $$N^{(k)} = 1+n_o^{(k)}. \tag{2.3.65}$$

We compare the two estimates for some arbitrary values of the single observation y, and the prior densities defined by a and b. The EM and Bayes estimates will be denoted by $\hat{\theta}_E$ and $\hat{\theta}_B$ respectively. A summary of the computations are given in Table 2.5. A close examination of the last entries in the table shows that $\hat{\theta}_B$ and $\hat{\theta}_E$ get closer and closer as y increases, but are noticeably different for small values of y. This trend persists over all the selected values of a and b. Thus, although the estimates are different, good approximation can be obtained when the observation y is away from zero.

Table 2.5

A Comparison of Bayes Estimates and the Estimate obtained by using the EM Algorithm

a	b	y	$\hat{\theta}_B$	$\hat{\theta}_E$
1.0	1.0	3	1.7943	1.8246
1.0	1.0	5	3.4184	3.4432
1.0	1.0	10	5.4776	5.4886
1.0	1.0	17	8.9980	8.9994
1.0	1.0	30	15.5000	15.5000
1.0	2.0	3	2.4656	2.5196
1.0	2.0	5	3.9104	3.9482
1.0	2.0	10	8.6162	8.7173
1.0	2.0	17	11.9995	12.0000
1.0	2.0	30	20.6667	20.6667
2.0	1.0	3	2.3482	2.3785
2.0	1.0	5	3.4184	3.4482
2.0	1.0	10	5.9839	5.9925
2.0	1.0	17	9.4986	9.4996
2.0	1.0	30	16.0000	16.0000
4.0	2.0	3	4.6072	4.6364
4.0	2.0	5	5.9742	5.9900
4.0	2.0	10	9.3304	9.3328
4.0	2.0	17	13.9998	14.0000
4.0	2.0	30	22.6667	22.6667

Chapter III

BAYESIAN AND MAXIMUM LIKELIHOOD ESTIMATION OF POISSON MEANS UNDER A MULTIPLICATIVE MODEL: THE CASE OF COMPLETE DATA

3.0 Introduction.

In this chapter, the estimation of the means of Poisson populations under a multiplicative model will be considered. The MLE will be derived and Bayes estimators will be obtained with respect to a class of prior distributions, under two weighted squared-error loss functions. We will consider the complete-data case in which the sample space X is of dimension p=IxJ, so that every entry x_{ij} in the table is observed. The case of incomplete data (in which some of the x_{ij} are not observable) will be considered in Chapter V.

3.1 The Model and the Likelihood Function

Let $x=\{x_{ij}\}$ $(i=1,2,\ldots,I;\ j=1,2,\ldots,J)$ denote a two-dimensional array of p=IxJ observations. The entries, x_{ij}, are realizations of the random variables X_{ij} which are independently distributed as Poisson with means θ_{ij}. The probability density function of X_{ij} is then

$$f(x_{ij}|\theta_{ij})=(\theta_{ij})^{x_{ij}}\exp(-\theta_{ij})/x_{ij}! \qquad x_{ij}=0,1\ldots, \tag{3.1.1}$$

Suppose that there exists a set of parameters $\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_I)$, $\beta=(\beta_1,\beta_2,\ldots,\beta_J)$, and τ such that

$$\theta_{ij}=\alpha_i\beta_j\tau, \tag{3.1.2}$$

where $\tau > 0$, $0 < \alpha_i < 1$, $0 < \beta_j < 1$, $\sum_i \alpha_i = 1$ and $\sum_j \beta_j = 1$.

Model (3.1.2), being multiplicative in the parameter space, corresponds to the log-linear model of 'independence' of Rasch (1960), where $\log \tau$ denotes the overall mean effect, $\log \alpha_i$ denotes the row i effect and $\log \beta_j$ denotes the column j effect. Under the above model, the likelihood function may be expressed as

$$L(\alpha,\beta,\tau|x) = K(x)\ \tau^z \exp(-\tau) \prod_i \alpha_i^{r_i} \prod_j \beta_j^{s_j}, \tag{3.1.3}$$

where $r_i = \sum_j x_{ij}$ and $s_j = \sum_i x_{ij}$ denote the i-th row and j-th column totals of the observation matrix $x = \{x_{ij}\}$, $z = \sum\sum x_{ij}$ and $K(x)$ is a statistic. It is easy to show (also see Corollary 3.3) that the ML estimators of τ and the components of α and β are given by

$$\hat{\alpha}_i = r_i/z, \tag{3.1.4}$$

$$\hat{\beta}_j = s_j/z, \tag{3.1.5}$$

and

$$\hat{\tau} = z, \tag{3.1.6}$$

provided that $z>0$. If $z=0$, then $\hat{\tau}=0$ but the $\hat{\alpha}_i$'s and $\hat{\beta}_j$'s are indeterminate.

It can be shown that the ML estimator of $\theta = \{\theta_{ij}\}$ is inadmissible under various loss functions, and that there exists a generalized Bayes estimator that dominates the ML estimator (see Chapter IV). Therefore, in the following sections we will be looking at the form of generalized Bayes estimators.

3.2 Specifying the Prior on $\theta=(\alpha,\beta,\tau)$ and the Loss Function

In the Bayesian framework, an essential element is the assumption that a prior distribution can be specified for the unknown parameter $\theta=(\alpha, \beta, \tau)$. In this problem, one possibility is for α, β and τ to be independently distributed as follows:

$$\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_I) \sim D_I(a_1,\ldots,a_I),$$

$$\beta=(\beta_1,\beta_2,\ldots,\beta_J) \sim D_J(b_1,\ldots,b_J),$$

$$\tau \sim M(d\tau),$$

where M is a Lebesgue measurable function on $(0,\infty)$. Here, $D_K(a_1,\ldots,a_K)$ denotes the K-variate Dirichlet distribution with parameters a_1+1, $a_2+1,\ldots,a_K+1$, for which the probability density function $f(u|a,K)$ is given by

$$f(u|a,K) = Q(a)\ u_1^{a_1}\ u_2^{a_2}\ldots u_I^{a_K}, \qquad 0 < u_i < 1,\ \sum u_i=1,$$

where $Q(a)= \Gamma(K+a_1+a_2+\ldots+a_K)/[\Gamma(a_1+1)\ldots\Gamma(a_K+1)]$ and $a_i> -1$. The joint posterior density of α, β and τ, given $x=\{x_{ij}\}$, is then

$$g(\alpha,\beta,\tau|x) \propto M(d\tau)\ \tau^z\exp(-\tau)\ \prod_i \alpha_i^{r_i+a_i} \prod_j \beta_j^{s_j+b_j} \tag{3.2.1}$$

where $\quad r_i=\sum_j x_{ij},\quad s_j=\sum_i x_{ij},\quad$ and $z=\sum_i \sum_j x_{ij}$.

If the problem is to be examined from a decision-theoretic point of view, then the loss incurred in estimating $\theta=\{\theta_{ij}\}$ by the estimate $d=\{d_{ij}\}$ must be specified. Two possible loss functions which we will consider are the weighted squared-error loss functions l_1 given by

$$l_1(d;\alpha,\beta,\tau) = \sum_i \sum_j \theta_{ij}^{-1}(d_{ij} - \theta_{ij})^2, \tag{3.2.2}$$

where $\theta_{ij}=\alpha_i\beta_j\tau$,

and $l_2(t,v;\alpha,\beta,\tau) =$

$$\sum_i (\theta_{i\cdot})^{-1}(d_{i\cdot} - \theta_{i\cdot})^2 + \sum_j (\theta_{\cdot j})^{-1}(d_{\cdot j} - \theta_{\cdot j})^2, \tag{3.2.3}$$

where $d_{i\cdot} = \sum_j d_{ij}$, $d_{\cdot j} = \sum_i d_{ij}$, $\theta_{i\cdot} = \sum_j \theta_{ij} = \alpha_i\tau$ and $\theta_{\cdot j} = \sum_i \theta_{ij} = \beta_j\tau$.

In using the loss function l_1, we are concerned with fitting each of the cell means θ_{ij}, whereas in using l_2 our goal is to estimate the means $\theta_{i\cdot} = \sum_j \theta_{ij}$ and $\theta_{\cdot j} = \sum_i \theta_{ij}$ along the margins. Under these circumstances, we will consider the Bayes estimators of τ and the vectors $\alpha=(\alpha_1,\ldots,\alpha_I)$ and $\beta=(\beta_1,\ldots,\beta_J)$.

The work presented in this chapter parallels that of Clevenson and Zidek (1975), who studied a similar problem in a one-way setting. However, the problem they studied does not impose the above multiplicative structure on the parameters. When I=1, our model (3.1.2) reduces to the case considered by Clevenson and Zidek.

3.3 Generalized Bayes Estimators under the loss functions l_1 and l_2

The generalized Bayes estimate of $\theta=\{\theta_{ij}\}$ w.r.t. to the loss function l_1 is denoted by $d=\{d_{ij}\}$, where $d_{ij}=u_iv_jw$. The values of $u=(u_1,\ldots,u_I)$, $v=(v_1,\ldots,v_J)$ and w are chosen to minimize the posterior risk for a fixed $x=\{x_{ij}\}$ and a specified joint (possibly improper) prior on $\theta=(\alpha,\beta,\tau)$, subject to the constraints $u_i>0$, $v_j>0$,

$w>0$, $\sum u_i=1$ and $\sum v_j=1$. The posterior risk is given by

$$R(u,v,w|x) = \int l_1(u,v,w;\alpha,\beta,\tau)\ g(\alpha,\beta,\tau|x)\ d\alpha\ d\beta\ d\tau. \tag{3.3.1}$$

If the prior on $\theta=(\alpha,\beta,\tau)$ is proper, then the resulting Bayes estimator will also minimize the (overall) risk

$$R = \int R(u,v,w|x)h(x)dx, \tag{3.3.2}$$

where $h(x)$ is the marginal density function of X.

Theorem 3.1: Assume that the prior distribution on $\theta=(\alpha,\beta,\tau)$ is as given in Section 3.2. Then, the (generalized) Bayes estimators of α, β and τ (and hence θ), w.r.t. the loss function l_1 are given by

$$\tilde{\alpha}_i = (R_i+a_i)/(Z+A), \tag{3.3.3}$$

$$\tilde{\beta}_j = (S_j+b_j)/(Z+B), \tag{3.3.4}$$

and $$\tilde{\tau} = [(z+A)/(z+A+I-1)][(z+B)/(z+B+J-1)]H(z) \tag{3.3.5}$$

where $A=\sum a_i$, $B=\sum b_j$,

and $$H(Z) = \int \tau^{Z}\exp(-\tau)M(d\tau)/[\int \tau^{Z-1}\exp(-\tau)M(d\tau)], \tag{3.3.6}$$

provided that $Z+A>0$, $Z+B>0$ and $H(z)$ is well defined. If $Z=0$, $Z+A\leq 0$ and $Z+B\leq 0$, then $\tilde{\tau}=0$ but the α_i's and β_j's are indeterminate.

Proof: Let $r_i^+=r_i+a_i$ and $s_j^+=s_j+b_j$ $(i=1,2,\ldots,I;\ j=1,2,\ldots,J)$. From (3.2.1) and (3.3.1), it follows that the posterior risk $R(d|x)$ is proportional to

$$\int \sum_i \sum_j \theta_{ij}^{-1}(d_{ij}-\theta_{ij})^2 \prod_i \alpha_i^{r_i^+} \prod_j \beta_j^{s_j^+} \tau^z \exp(-\tau)M(d\tau)\prod_i d\alpha_i \prod_j d\beta_j, \tag{3.3.7}$$

$$=\int \sum_i \sum_j (\theta_{ij}^{-1}d_{ij}^2-2d_{ij}+\theta_{ij})\prod_i \alpha_i^{r_i^+} \prod_j \beta_j^{s_j^+} \tau^z \exp(-\tau)M(d\tau)\prod_i d\alpha_i \prod_j d\beta_j. \tag{3.3.8}$$

In what follows, it will be convenient to define the quantity $R'(d|x)$ by

$$R'(d|x) \quad =\int \sum_i \sum_j (\theta_{ij}^{-1}d_{ij}^2-2d_{ij})\prod_i \alpha_i^{r_i^+} \prod_j \beta_j^{s_j^+} \tau^z \exp(-\tau)M(d\tau) \prod_i d\alpha_i \prod_j d\beta_j$$

$$+ k_1(1-\sum_i u_i) \quad +k_2(1-\sum_j v_j), \tag{3.3.9}$$

where the integration is over the (I+J-1) dimensional space of α, β and τ, and k_1 and k_2 are Lagrangian multipliers. Thus

$$R'(u,v,w|x) =\int M(d\tau)\tau^Z\exp(-\tau)\sum_i \sum_j [\theta_{ij}^{-1}u_i^2v_j^2w^2-2u_iv_jw] \prod_i \alpha_i^{r_i^+}\prod_j \beta_j^{s_j^+}d\alpha_i d\beta_j$$

$$+ k_1(1-\sum_i u_i) + k_2(1-\sum_j v_j), \tag{3.3.10}$$

$$=\int\tau^z\exp(-\tau)M(d\tau)\int \sum_i \sum_j [(\alpha_i\beta_j\tau)^{-1}u_i^2v_j^2w^2- 2u_iv_jw]\prod_i \alpha_i^{r_i^+}\prod_j \beta_j^{s_j^+} d\alpha_i d\beta_j$$

$$+ k_1(1-\sum_i u_i) + k_2(1-\sum_j v_j), \tag{3.3.11}$$

$$=\int \tau^z\exp(-\tau)M(d\tau) \sum_i \sum_j [(u_iv_jw)^2\tau^{-1}(z+A+I-1)(z+B+J-1)/r_i^+s_j^+-2u_iv_jw]$$

$$+ k_1(1-\sum_i u_i) + k_2(1-\sum_j v_j). \tag{3.3.12}$$

The evaluation of (3.3.11) involves (I+J-1) integrations w.r.t. the components of α, β and τ. The expression (3.3.12) is easily obtained by observing that the marginal posterior distribution of α_i, given z, is a beta distribution with parameters $(r_i^+ +1)$, and $\sum_{j\neq i}^{I}(r_j^+ +1)$. Hence $E[(1/\alpha_i)|z]=(z+A+I-1)/(r_i^+)$, $(r_i^+ > 0\ ;\ i=1,\ldots,I)$. Similarly, $E(1/\beta_j)|z)=(z+B+J-1)/s_j^+$, $(s_j^+ > 0\ ;\ j=1,\ldots,J)$. Further, it should be observed that minimizing $R(u,v,w|x)$ in (3.3.1) w.r.t. $u=(u_1,\ldots,u_I)$, $v=(v_1,\ldots,v_J)$ and w is equivalent to minimizing $R'(u,v,w|x)$ w.r.t. u, v, w, k_1 and k_2 unconditionally. On differentiating (3.3.12) separately w.r.t. k_1, k_2, w and the components of u and v we obtain

$$\partial/\partial u_i\ R'(u,v,w|x) =$$

$$2\int\tau^z\exp(-\tau)M(d\tau)\ \sum_j [u_i(v_jw)^2(z+A+I-1)(z+B+J-1)/(\tau r_i^+ s_j^+)-v_jw]-k_1, \tag{3.3.13}$$

$$\partial/\partial v_j\ R'(u,v,w|x)=$$

$$2\int\tau^z\exp(-\tau)M(d\tau)\ \sum_i [v_j(u_iw)^2(z+A+I-1)(z+B+J-1)/(\tau r_i^+ s_j^+)-u_iw]-k_2, \tag{3.3.14}$$

$$\partial/\partial w\ R'(u,v,w|x)=$$

$$=2\int\tau^z\exp(-\tau)M(d\tau)\sum_i\sum_j [w(u_iv_j)^2(z+A+I-1)(z+B+J-1)/(\tau r_i^+ s_j^+)-u_iv_j], \tag{3.3.15}$$

$$\partial/\partial k_1 \; R'(u,v,w|x) = 1-\sum_i u_i, \tag{3.3.16}$$

and

$$\partial/\partial k_2 \; R'(u,v,w|x) = 1-\sum_j v_j. \tag{3.3.17}$$

Equating each partial derivative to zero and solving the resulting equations simultaneously, we obtain

$$u_i = r_i^+\{k_1/2+Q(z)\sum_j v_j w\}/\{(z+A+I-1)(z+B+J-1)Q(z-1)\sum_j v_j^2 w^2/s_j^+\} \tag{3.3.18}$$

$$v_j = s_j^+\{k_2/2+Q(z)\sum_i u_i w\}/\{(z+B+J-1)(z+A+I-1)Q(z-1)\sum_i u_i^2 w^2/r_i^+\}, \tag{3.3.19}$$

and $$w = H(z) \sum_i\sum_j u_i v_j]/[(z+A+I-1)(z+B+J-1)\sum_i\sum_j (u_i v_j)^2/r_i^+ s_j^+)], \tag{3.3.20}$$

where $Q(a) = \int \tau^a \exp(-\tau) M(d\tau)$.

From (3.3.18) and (3.3.19), it follows immediately that

$$u_i \propto r_i^+, \qquad (i=1,\ldots,I)$$

and $$v_j \propto s_j^+. \qquad (j=1,\ldots,J)$$

Letting $u_i=k_1 r_i^+$, and noting that $\sum u_i=1$, we get $k_1=1/(z+A)$. Hence

$$u_i = r_i^+/(z+A). \tag{3.3.21}$$

Similarly,

$$v_j = s_j^+/(z+B). \tag{3.3.22}$$

Substituting the expressions in (3.3.21) and (3.3.22) into (3.3.20) and noting that $\sum_i\sum_j u_i v_j=1$, we obtain

$$w = H(z)(z+A)^2(z+B)^2/[(z+A+I-1)(z+B+J-1)\sum_i\sum_j r_i^+ s_j^+], \tag{3.3.23}$$

$$= [(z+A)/(z+A+I-1)][(z+B)/(z+B+J-1)]H(z), \tag{3.3.24}$$

which, together with (3.3.21) and (3.3.22) is the statement of the theorem.

≡

Corollary 3.1 considers the special case in which the prior distribution is an improper uniform distribution on $\theta=(\alpha,\beta,\tau)$, that is, $a_i=0$, $\forall$ i, $b_j=0$ $\forall$ j and $M(d\tau)=d\tau$ is a Lebesgue measure on $(0,\infty)$.

Corollary 3.1: If the prior distribution on $\theta=(\alpha,\beta,\tau)$ is uniform on $\theta=(\alpha,\beta,\tau)$, then the generalized Bayes estimators of α, β and τ are

$$\tilde{\alpha}_i = R_i/Z,$$

$$\tilde{\beta}_j = S_j/Z,$$

and $$\bar{\tau} = z[1-\kappa(I)][1-\kappa(J)], \tag{3.3.25}$$

where $$\kappa(S) = (S-1)/(z+S-1), \tag{3.3.26}$$

Proof: Substituting $M(d\tau)=d\tau$ into (3.3.6), we obtain

$$w = [z/(z+I-1)][z/(z+J-1)]\int\tau^{z}\exp(-\tau)d\tau/\{\int\tau^{z-1}\exp(-\tau)d\tau\}, \tag{3.3.27}$$

$$= z[1-\kappa(I)][1-\kappa(J)]. \tag{3.3.28}$$

≡

We will determine next the form of the generalized Bayes estimators of $\alpha=(\alpha_1,\ldots,\alpha_I)$, $\beta=(\beta_1,\ldots,\beta_J)$ and τ w.r.t. to the loss function l_2. We state the main result in Theorem 3.2 which follows.

Theorem 3.2: For $x=\{x_{ij}\}$, assume that the posterior on $\theta=(\alpha,\beta,\tau)$ is $g(\alpha,\beta,\tau|x)$ given by equation (3.2.1). Then the generalized Bayes estimators of α, β, and τ w.r.t. l_2 are given by

$$\bar{\bar{\alpha}}_i = R_i^+/(Z+A) \qquad (i=1,\ldots,I),$$

$$\bar{\bar{\beta}}_j = S_j^+/(Z+B) \qquad (j=1,\ldots,J),$$

$$\bar{\bar{\tau}} \equiv w = H(Z)/[(Z+A+I-1)/(z+A) + (Z+B+J-1)/(Z+B)] \tag{3.3.29}$$

Proof: For a given $x=\{x_{ij}\}$, the generalized Bayes estimates $u=(u_1,\ldots,u_I)$, $v=(v_1,\ldots,v_J)$ and w are such that the posterior Bayes risk

$$R_2(u,v,w|x) = \int l_2(u,v,w;\alpha,\beta,\tau)g(\alpha,\beta,\tau|u,v,w) \prod_i d\alpha_i \prod_j d\beta_j \, d\tau. \tag{3.3.30}$$

$$= \int[\sum_i W_i^+(u_i w-\theta_{i\cdot})^2 g(\alpha,\beta,\tau|u,v,w) \prod_i d\alpha_i \prod_j d\beta_j \, M(d\tau)$$

$$+ \int[\sum_j W_j^{++}(v_j w-\theta_{\cdot j})^2] g(\alpha,\beta,\tau|u,v,w) \prod_i d\alpha_i \prod_j d\beta_j \, M(d\tau)$$

where $W_i^+ = \theta_{i\cdot}^{-1} = (\alpha_i\tau)^{-1}$ and $W_j^{++} = \theta_{\cdot j}^{-1} = (\beta_j\tau)^{-1}$,

is a minimum. Then equivalent to the above is the minimization of

$$R_2(u,v,w|x) = \int\{\sum_i W_i^+(u_i^2 w^2 - 2u_i w\alpha_i\tau)\} \; g(\alpha,\beta,\tau|u,v,w) \, d\tau \prod_i d\alpha_i \prod_j d\beta_j$$

$$+ \int\{\sum_j W_j^{++}(v_j^2 w^2 - 2v_j w\beta_j\tau)\} \; g(\alpha,\beta,\tau|u,v,w) \, d\tau \prod_i d\alpha_i \prod_j d\beta_j$$

$$+ \delta_1(1-\sum u_i) + \delta_2(1-\sum v_j) \tag{3.3.31}$$

$$=\sum_i \int \{u_i^2 w^2 (\alpha_i \tau)^{-1} - 2u_i w\} \prod \alpha_i^{r_i^+} \prod \beta_j^{s_j^+} \tau^z \exp(-\tau) M(d\tau) \prod d\alpha_i \prod d\beta_j$$

$$+ \sum_j \int \{v_j^2 w^2 (\beta_j \tau)^{-1} - 2v_j w)\} \prod \alpha_i^{r_i^+} \prod \beta_j^{s_j^+} \tau^z \exp(-\tau) M(d\tau) \prod d\alpha_i \prod d\beta_j$$

$$+ \delta_1 (1-\sum u_i) + \delta_2 (1-\sum v_j), \qquad (3.3.32)$$

where δ_1 and δ_2 are Lagrangian multipliers. Thus,

$$R_2(u,v,w|x) = \int \tau^z \exp(-\tau) [\sum u_i^2 w^2 (z+A+I-1)/r_i^+ \tau - 2u_i w] M(d\tau)$$

$$+ \int \tau^z \exp(-\tau) [\sum v_j^2 w^2 (z+B+J-1)/s_j^+ \tau - 2v_j w)] M(d\tau)$$

$$+ \delta_1 (1-\sum_i u_i) + \delta_2 (1-\sum_j v_j). \qquad (3.3.33)$$

Hence

$$\partial R_2(u,v,w|x)/\partial u_k = 2\int \tau^z \exp(-\tau) [u_k w^2 (z+A+I-1)/r_k^+ \tau - w] M(d\tau) - \delta_1 = 0, \qquad (3.3.34)$$

$$\partial R_2(u,v,w|x)/\partial v_t = 2\int \tau^z \exp(-\tau) [v_t w^2 (z+B+J-1)/s_j^+ \tau - w] M(d\tau) - \delta_2 = 0 \qquad (3.3.35)$$

$$\partial R_2(u,v,w|x)/\partial w = 2\int \tau^z \exp(-\tau) [w(z+A+I-1) \sum_i u_i^2/(r_i^+ \tau) - \sum_i u_i] M(d\tau)$$

$$+ 2\int \tau^z \exp(-\tau) w(z+B+J-1) \sum_j v_j^2/(s_j^+ \tau) - \sum_j v_j] M(d\tau) = 0, \qquad (3.3.36)$$

$$\partial R_2(u,v,w|x)/\partial \delta_1 = 1-\sum u_i = 0, \qquad (3.3.37)$$

and $\partial R_2(u,v,w|x)/\partial \delta_2 = 1- \sum v_j = 0.$ (3.3.38)

The Bayes estimators of α, β and τ are obtained by solving equations (3.3.34) through (3.3.38) simultaneously. From (3.3.34) and (3.3.35), it follows that

$$u_k = r_k^+ [wQ(z) + \delta_1/2)]/[w^2 Q(z-1)(z+A+I-1)], \quad (k=1,\ldots,I) \qquad (3.3.39)$$

and $v_t = s_t^+ [wQ(z) + \delta_2/2)]/[w^2 Q(z-1)(z+B+J-1)], \quad (t=1,\ldots,J).$ (3.3.40)

Hence $\bar{\alpha}_k \equiv u_k \propto r_k^+$, $(k=1,\ldots,I)$

and $\bar{\beta}_t \equiv v_t \propto s_t^+$, $(t=1,2,\ldots,J)$

$$\Rightarrow \quad \bar{\alpha}_k \equiv u_k = r_k^+/(z+A), \quad (k=1,\ldots,I) \tag{3.3.41}$$

$$\text{and} \quad \bar{\beta}_t \equiv v_t = s_t^+/(z+B). \quad (t=1,\ldots,J) \tag{3.3.42}$$

After substituting u_i and v_j from (3.3.41) and (3.3.42) into (3.3.36) and solving for w, it follows that

$$\bar{\tau} \equiv w = H(Z)/[(Z+A+I-1)/(z+A) + (Z+B+J-1)/(Z+B)] \tag{3.3.43}$$

≡≡

Corollary 3.2: If the prior on $\theta=(\alpha,\beta,\tau)$ is such that $M(d\tau)=d\tau$, $\tau \in (0,\infty)$, $a_i=0$, $\forall$ i, $b_j=0$ $\forall$ j, then the generalized Bayes estimator of τ, w.r.t. the loss function l_2, is

$$\bar{\bar{\alpha}}_i = R_i/Z \quad (i=1,\ldots,I),$$

$$\bar{\bar{\beta}}_j = S_j/Z \quad (j=1,\ldots,J),$$

$$\text{and} \quad \bar{\bar{\tau}} = Z/[(Z+I-1)/Z+(Z+J-1)/Z]. \tag{3.3.44}$$

Proof: Similar to Corollary 3.1.

≡≡

3.4 Posterior Mode Estimators:

Sometimes a Bayesian would use a posterior mode estimator as a natural extension of the maximum likelihood technique, although the posterior mode estimator is not formally related to a specific loss function. From our point of view, the posterior mode estimator may also be easier to evaluate than the posterior mean estimator under a given loss function and a specified prior on θ. For example, we will show in Chapter V that, under a specified improper prior on θ, the generalized Bayes estimator in the incomplete multiplicative Poisson problem cannot be obtained in closed form; even finding the generalized Bayes estimate by numerical methods is computationally difficult, whereas it is relatively easy to compute the mode of the posterior distribution via the EM algorithm.

In Chapter V, we shall examine fully how to obtain the posterior mode estimator and the generalized Bayes estimator, with respect to a uniform (improper) prior on $\theta=(\alpha,\beta,\tau)$, for the incomplete multiplicative Poisson using the EM algorithm. Therefore we would like to know the form of the posterior mode estimator of α, β and τ for the above complete data problem.

In the following derivation of the posterior mode estimator of θ, we will assume that the prior distribution of θ is the same as that given in Section 3.2. In addition, we will assume that the measure $M(d\tau)$ on τ corresponds to an exponential distribution. The posterior mode estimator of θ, w.r.t. the above prior, is defined to be a value of θ at which the posterior distribution attains its maximum value (assuming that it exists).

Theorem 3.3: Let $x=\{x_{ij}\}$, where the x_{ij} are realizations of the random variables X_{ij} $(i=1,..,I;\ j=1,...,J)$ which are independently distributed with probability distribution functions given by (3.1.1). Assume that the prior distribution on $\theta=(\alpha,\ \beta,\tau)$ is as given in Section 3.2, with $M(d\tau)=k^{-1}\exp(-\tau/k)d\tau$, the exponential distribution with mean k. Then the posterior mode estimators of α, β, and τ are given by

$$\bar{\alpha}_i = R_i^+/\sum_j R_j^+, \qquad (i=1,...,I),$$

$$\bar{\beta}_j = S_j^+/\sum_i S_i^+, \qquad (j=1,...,J),$$

and $\quad \bar{\tau}=Z/(1+k^{-1})$,

where $R_i^+ = R_i+a_i$ and $S_j^+ = S_j+b_j$.

Proof: From (3.2.1), the joint posterior distribution of α, β an τ is given by

$$h(\alpha,\beta,\tau)\,|\,r,s,k) \propto \prod_i \alpha_i^{r_i^+} \prod_j \beta_j^{s_j^+}\ \tau^z \exp(-\tau(1+1/k)). \qquad (3.4.1)$$

The result follows after taking the log of (3.4.1), equating to zero the partial derivatives w.r.t. the components of α, β and τ and solving the resulting simultaneous equations.

==

Corollary 3.3: If the prior distribution on $\theta=(\alpha,\beta,\tau)$ is as given in Section 3.2, then the ML estimators of α, β and τ can be obtained from Theorem 3.3 by setting $a_i=0$, $b_j=0$ and $k=\infty$.

Proof: The ML estimators are formally equivalent to the mode of the posterior distribution w.r.t. a uniform prior on the parameters. This follows from the fact that, w.r.t. a uniform prior density, the joint posterior density of α, β and τ is proportional to the likelihood function (3.1.3). Since the (improper) joint uniform prior on α, β and τ can be obtained from Theorem 3.3 by setting $a_i=0$, $(i=1,\ldots,I)$, $b_j=0$, $(j=1,\ldots,J)$ and $k=\infty$, it follows from Theorem 3.3 that the ML estimators of α, β and τ are respectively

$$\hat{\alpha}_i = R_i/Z, \qquad \text{if} \quad Z > 0 \qquad (i=1,\ldots,I),$$

$$\hat{\beta}_j = S_j/Z, \qquad \text{if} \quad Z > 0 \qquad (j=1,\ldots,J),$$

and $\hat{\tau} = Z.$

==

Chapter IV

SIMULTANEOUS ESTIMATION OF POISSON PARAMETERS UNDER A MULTIPLICATIVE MODEL: SOME DECISION-THEORETIC RESULTS FOR COMPLETE DATA.

4.0 Introduction

In this chapter, the admissibility of the MLE under a weighted squared-error loss function will be considered for the complete data problem. It will be shown that there exists a class P_D of estimators having uniformly smaller risks than the MLE. It will also be shown that the generalized Bayes estimator of $\theta = (\alpha, \beta, \tau)$ with respect to an (improper) uniform prior distribution on θ is a member of P_D. These results extend those of Clevenson and Zidek (1975) who considered a similar problem. The difference in our approach lies in the treatment of the parameters. Unlike their procedure, ours assumes that there is a multiplicative structure on the p=IxJ population means θ_{ij}, that is, $\theta_{ij}=\alpha_i\beta_j\tau$, $0<\alpha_i<1$, $0<\beta_j<1$, $\tau>0$, $\sum \alpha_i=1$ and $\sum \beta_j=1$ $(i=1,\ldots,I;$ $j=1,\ldots,J)$, and we consider the simultaneous estimation of $\alpha=(\alpha_1,\ldots,\alpha_I)$, $\beta=(\beta_1,\ldots,\beta_J)$ and τ.

4.1 Dominating Estimators

Let $X=\{X_{ij}\}$ denote a matrix of independent Poisson random variables with means $\theta=\{\theta_{ij}\}$. In Chapter III, it was shown that the MLE of the matrix $\theta=\{\theta_{ij}\}$ is given by $\hat{\Theta}=\{\hat{\Theta}_{ij}\}$, where $\hat{\Theta}_{ij}=\hat{\alpha}_i\hat{\beta}_j\hat{\tau}$ and the $\hat{\alpha}_i$'s, $\hat{\beta}_j$'s and $\hat{\tau}$ are given in equations (3.1.4),(3.1.5) and (3.1.6), respectively. In this chapter, we will use $\hat{\Theta}$, rather than $\hat{\theta}$, to represent the maximum likelihood estimator in order to emphasize that $\hat{\Theta}$ is a random variable (matrix) and that we are examining sampling properties of $\hat{\Theta}$.

An estimator $\tilde{\Theta} = \{\tilde{\Theta}_{ij}\}$ is said to **dominate** $\hat{\Theta} = \{\hat{\Theta}_{ij}\}$, w.r.t a specified loss function, if the risk function of $\tilde{\Theta}$ is less than or equal to the risk function of $\hat{\Theta}$, for all θ, and with strict inequality for at least one θ. Consider a family P_D of estimators $\tilde{\Theta} = \{\tilde{\Theta}_{ij}\}$, where $\tilde{\Theta}_{ij}$ is defined by

$$\tilde{\Theta}_{ij} = \hat{\Theta}_{ij}[1 - G(Z)] = \hat{\Theta}_{ij}W(Z), \tag{4.2.1}$$

where $Z = \sum_i \sum_j X_{ij}$, $0 \le G(Z) \le 1$, for all Z, and $G(0)=0$. We remark that $\hat{\Theta}_{ij}=0$ for Z=0 and hence G(0) is arbitrary; for convenience G(0) is taken to be zero. Theorem 4.1 asserts that, if G(Z) satisfies certain conditions, then equation (4.2.1) defines an estimator which dominates the MLE. The proof of Theorem 4.1 will be established by comparing the risk function of $\tilde{\Theta}$ with that of $\hat{\Theta}$.

In the following proofs, we will make use of the fact that, conditional on Z, $R_i = \sum_j X_{ij}$ and $S_j = \sum_i X_{ij}$ are independent, the conditional distribution of R_i, given Z, is binomial with parameters Z and α_i, and the conditional distribution of S_j, given Z, is binomial with parameters Z and β_j. Hence, for example,

$$E(R_i | Z) = Z\alpha_i,$$

and $$E(R_i^2 | Z) = Z\alpha_i(1-\alpha_i) + (Z\alpha_i)^2, \quad (i=1,2,\ldots,I).$$

The following lemmas will be required in the derivation of the risk functions of $\hat{\Theta}$ and $\tilde{\Theta}$.

Lemma 4.1. $E(\tilde{\Theta}_{ij}) = \alpha_i\beta_j[\tau - E(ZG(Z))]$. (4.2.2)

Proof:

$$E(\tilde{\Theta}_{ij}) = E[\hat{\Theta}_{ij}W(Z)], \tag{4.2.3}$$

$$= E\{W(Z)E(\hat{\Theta}_{ij}|Z)\}, \tag{4.2.4}$$

$$= \alpha_i\beta_j E\{ZW(Z)\},$$

$$= \alpha_i\beta_j[E(Z) - E(ZG(Z))], \tag{4.2.5}$$

$$= \alpha_i\beta_j[\tau - E(ZG(Z))]. \tag{4.2.6}$$

==

Lemma 4.2. For Z>0,

$$E(\hat{\Theta}^2_{ij}|Z) = Z^2\alpha_i^2\beta_j^2 + Z[\alpha_i^2\beta_j(1-\beta_j) + \beta_j^2\alpha_i(1-\alpha_i)] + \alpha_i\beta_j(1-\alpha_i)(1-\beta_j). \tag{4.2.7}$$

Proof: For Z>0,

$$E(\hat{\Theta}^2_{ij}|Z) = E(R_i^2S_j^2/Z^2|Z), \tag{4.2.8}$$

$$=(1/Z)^2E(R_i^2|Z)E(S_j^2|Z), \tag{4.2.9}$$

$$=(1/Z)^2\{[Z\alpha_i(1-\alpha_i) + (Z\alpha_i)^2][Z\beta_j(1-\beta_j) + (Z\beta_j)^2], \tag{4.2.10}$$

$$= [\alpha_i(1-\alpha_i) + Z\alpha_i^2][\beta_j(1-\beta_j) + Z\beta_j^2], \tag{4.2.11}$$

$$= (Z\alpha_i\beta_j)^2 + Z\alpha_i\beta_j[\alpha_i(1-\beta_j) + \beta_j(1-\alpha_i)] \tag{4.2.12}$$

$$+ \alpha_i\beta_j(1-\alpha_i)(1-\beta_j). \tag{4.2.13}$$

≡≡

Lemma 4.3. $E(\tilde{\Theta}_{ij}^2) = \alpha_i^2\beta_j^2 E\{Z^2W^2(Z)\}$

$$+ \alpha_i\beta_j[\alpha_i(1-\beta_j) + \beta_j(1-\alpha_i)]E\{ZW^2(Z)\}$$

$$+ \alpha_i\beta_j(1-\alpha_i)(1-\beta_j)[E(W^2(Z)) - \exp(-\tau)]. \tag{4.2.14}$$

Proof:

$$E(\tilde{\Theta}_{ij}^2) = E\{E(\hat{\Theta}_{ij}^2 W^2(Z)|Z)\}, \tag{4.2.15}$$

$$= E\{W^2(Z)E(\hat{\Theta}_{ij}^2|Z)\}, \tag{4.2.16}$$

$$= \sum_{z>0} \{W^2(z)E(\hat{\Theta}_{ij}^2|Z=z)\}P(Z=z). \tag{4.2.17}$$

Hence, using Lemma 4.2, equation (4.2.17) becomes

$$E(\tilde{\Theta}_{ij}^2) =$$

$$\sum_{z>0} W^2(z)\{z^2\alpha_i^2\beta_j^2 + z[\alpha_i^2\beta_j(1-\beta_j) + \beta_j^2\alpha_i(1-\alpha_i)]\}P(Z=z)$$

$$+ \alpha_i\beta_j(1-\alpha_i)(1-\beta_j)\sum_{z>0} W^2(z)P(Z=z), \tag{4.2.18}$$

$$= E\{W^2(Z)(Z^2\alpha_i^2\beta_j^2 + Z\alpha_i^2\beta_j(1-\beta_j) + Z\alpha_i\beta_j^2(1-\alpha_i)\}$$

$$+ \alpha_i\beta_j(1-\alpha_i)(1-\beta_j)E\{W^2(Z)\} - \alpha_i\beta_j(1-\alpha_i)(1-\beta_j)\exp(-\tau), \quad (4.2.19)$$

$$= \alpha_i^2\beta_j^2E\{Z^2W^2(Z)\} + \alpha_i\beta_j[\alpha_i(1-\beta_j) + \beta_j(1-\alpha_i)]E\{ZW^2(Z)\}$$

$$+ \alpha_i\beta_j(1-\alpha_i)(1-\beta_j)[E\{W^2(Z)\} - \exp(-\tau)]. \quad (4.2.20)$$

==

Theorem 4.1: Let $L(Z)=(Z+J-1)(Z+I-1)$ and P_D denote the class of estimators $\tilde{\Theta}=\{\tilde{\Theta}_{ij}\}$, where the $\tilde{\Theta}_{ij}$'s are given by (4.2.1), with the function G satisfying the following conditions:

(1) $G(Z) \geq 0, \quad \forall\ Z>0,$

(2) $[1-G(Z)/2]L(Z)/Z \geq Z, \quad \forall\ Z>0,$ and

(3) $ZG(Z)$ is a monotonically increasing function of Z.

Then, w.r.t. the loss function l_1, $\tilde{\Theta} \in P_D$ dominates $\hat{\Theta}$; in fact, the risk of $\tilde{\Theta}$ is uniformly smaller (over the parameter space Ω) than the risk of $\hat{\Theta}$.

Proof: With respect to the loss function l_1, which is given by

$$l_1(\tilde{\Theta};\theta) = \sum_i \sum_j \theta_{ij}^{-1} (\tilde{\Theta}_{ij} - \theta_{ij})^2, \quad (4.2.21)$$

the risk of $\tilde{\Theta}$ is defined by

$$E(l_1(\tilde{\Theta};\theta) = E[\sum_i \sum_j \theta_{ij}^{-1}(\tilde{\Theta}_{ij} - \theta_{ij})^2] \quad (4.2.22)$$

$$= \sum_i \sum_j \theta_{ij}^{-1}\{E[\tilde{\Theta}_{ij} - \theta_{ij}]^2\}. \quad (4.2.23)$$

$$= \sum_i \sum_j \theta_{ij}^{-1} E(\tilde{\Theta}_{ij}^2) - 2\sum_i \sum_j E(\tilde{\Theta}_{ij}) + \tau. \tag{4.2.24}$$

From Lemma 4.1, it follows that

$$\sum_i \sum_j E(\tilde{\Theta}_{ij}) = \sum_i \sum_j \alpha_i \beta_j [\tau - E(ZG(Z))],$$

$$= \tau - E(ZG(Z)). \tag{4.2.25}$$

Applying Lemma 4.3 to the first term in (4.2.24), we obtain

$$\sum_i \sum_j \theta_{ij}^{-1} E(\tilde{\Theta}_{ij}^2) = \tau^{-1} \sum_i \sum_j \alpha_i \beta_j E\{Z^2 W^2(Z)\} + \tau^{-1} \sum_i \sum_j (1-\alpha_i)(1-\beta_j) E\{W^2(Z)\}$$

$$+ \tau^{-1} \sum_i \sum_j [\alpha_i (1-\beta_j) + \beta_j (1-\alpha_i)] E\{ZW^2(Z)\}$$

$$- \tau^{-1} \sum_i \sum_j (1-\alpha_i)(1-\beta_j) \exp(-\tau), \tag{4.2.26}$$

$$= \tau^{-1} E\{W^2(Z)(Z^2 + Z(I+J-2)\} \tag{4.2.27}$$

$$+ \tau^{-1} (I-1)(J-1)[E\{W^2(Z)\} - \exp(-\tau)].$$

Using the results in equations (4.2.27) and (4.2.25), equation (4.2.24) can be expressed in the form

$$r_G(\tau) = E(1_1(\tilde{\Theta};\theta)) = \tau^{-1} E\{W^2(Z)(Z^2 + Z(I+J-2)\} + 2E\{ZG(Z)\}$$

$$+ \tau^{-1}(I-1)(J-1)[E\{W^2(Z)\} - \exp(-\tau)] - \tau. \tag{4.2.28}$$

Note that the risk function of $\tilde{\Theta}$ depends on θ only through τ, and hence will be denoted by $r_G(\tau)$. Since $\hat{\Theta}$ can be obtained from equation (4.2.1) by setting $G(Z)=0$ for all Z, the risk of $\hat{\Theta}$, denoted by $r_M(\tau)$, may be obtained from (4.2.28) by setting $W(Z)=1$. Thus

$$r_M(\tau) = E(l_1(\hat{\Theta};\theta)) = \tau^{-1}E\{Z^2+Z(I+J-2)\} + \tau^{-1}(I-1)(J-1)(1-\exp(-\tau))-\tau,$$

$$= (I+J-1) - (I-1)(J-1)\tau^{-1}(1-\exp(-\tau)). \qquad (4.2.29)$$

Rewriting (4.2.28) in terms of G(Z), we obtain

$$r_G(\tau) = \tau^{-1}E\{G^2(Z)(Z^2+Z(I+J-2))\} - \tau^{-1}2E\{G(Z)(Z^2+Z(I+J-2))\}$$

$$+ \tau^{-1}(I-1)(J-1)E\{G^2(Z)-2G(Z)\} + 2E\{ZG(Z)\}$$

$$+ \tau^{-1}[E\{Z^2+Z(I+J-2)\} + (I-1)(J-1)(1-\exp(-\tau)) - \tau^2]. \qquad (4.2.30)$$

From equation (4.2.29), it is clear that the last term (in square brackets) of (4.2.30) is equal to $r_M(\tau)$. Hence, upon simplification, (4.2.30) may be written in the form

$$r_G(\tau)= r_M(\tau) + \tau^{-1}E\{(G^2(Z)-2G(Z))(Z+I-1)(Z+J-1)\} + 2E\{ZG(Z)\}. \quad (4.2.31)$$

Now let $L(Z)=(Z+I-1)(Z+J-1)$, and $G^*(Z)=(1-G(Z)/2)$. Then (4.2.31) becomes

$$r_G(\tau) = r_M(\tau) - 2\tau^{-1}E\{ZG(Z)[L(Z)G^*(Z)/Z - \tau]\}. \qquad (4.2.32)$$

Hence $\tau[r_M(\tau) - r_G(\tau)] = 2E\{ZG(Z)[L(Z)G^*(Z)/Z - \tau]\},$ (4.2.33)

$$\geq 2E\{ZG(Z)[Z-\tau]\}, \qquad (4.2.34)$$

$$> 2E\{ZG(Z)\}E[Z-\tau], \qquad (4.2.35)$$

$$=0.$$

The inequality (4.2.34) follows from (4.2.33) by replacing $L(Z)G^*(Z)/Z$ by Z, and applying condition (2) of the theorem. Also, by hypothesis, $ZG(Z)$ is a strictly increasing function of Z, as is $Z-\tau$. Therefore the inequality (4.2.35) is easily obtained by noting that $\mathrm{cov}(ZG(Z), Z-\tau) > 0$. The proof is completed by noting that $E(Z-\tau)=0$.

≡≡

Theorem 4.2 (below) will establish the existence of a generalized Bayes estimator (of θ) which is a member of P_D, and which hence dominates the MLE.

Theorem 4.2: Let $\tilde{\Theta}_B$ denote the generalized Bayes estimator of θ w.r.t. the loss function l_1 and a prior on $\theta=(\alpha,\beta,\tau)$ under which α, β and τ are independent and:

$$\alpha=(\alpha_1,\ldots,\alpha_I) \sim D_I(0,\ldots,0),$$
$$\beta=(\beta_1,\ldots,\beta_J) \sim D_J(0,\ldots,0),$$
$$\tau \sim \text{improper } U(0,\infty).$$

Then $\tilde{\Theta}_B$ dominates the MLE, $\hat{\Theta}$.

Proof: Let $\nu_1=(I-1)$ and $\nu_2=(J-1)$. From equations (3.3.3), (3.3.4) and (3.3.5), and setting $a_i=0$, $\forall i$, $b_j=0$, $\forall j$, it follows that the generalized Bayes estimator of $\theta=\{\theta_{ij}\}$, w.r.t. l_1 and the given priors on α, β and τ, may be written in the form

$$\tilde{\Theta}_{Bij} = R_i S_j [1-G(Z)]/Z,$$

where $G(Z)=H(Z)/L(Z)$,

$H(Z)=[Z(\nu_1+\nu_2)+\nu_1\nu_2]$,

and $L(Z)=[(Z+\nu_1)(Z+\nu_2)]$.

The function G satisfies the following conditions:

(1). $G(Z)=[Z(\nu_1+\nu_2)+\nu_1\nu_2]/[(Z+\nu_1)(Z+\nu_2)] \geq 0, \quad Z > 0.$

(2). $(1-G(Z)/2)L(Z)/Z \geq Z$, implies that

$$H(Z) \leq 2[Z(\nu_1+\nu_2)+\nu_2\nu_1], \tag{4.2.36}$$

$$= 2H(Z),$$

(3). $$ZG(Z) = ZH(Z)/L(Z)=Z[Z(\nu_1+\nu_2)+\nu_1\nu_2]/[(Z+\nu_1)(Z+\nu_2)], \tag{4.2.37}$$

$$=(\nu_1+\nu_2)\{Z/(Z+\nu_1)\}\{[Z+\nu_1\nu_2/(\nu_1+\nu_2)]/(Z+\nu_2)\}. \tag{4.2.38}$$

Note that in equation (4.2.38), $ZG(Z)$ is a monotonically increasing function of Z since both $Z/(Z+\nu_1)$ and $[Z+\nu_1\nu_2/(\nu_1+\nu_2)]/[Z+\nu_2]$ are increasing functions of Z. Thus, $\tilde{\Theta}_B=\{\tilde{\Theta}_{Bij}\}$ is a member of the class of estimators P_D. Hence, by Theorem 4.1, the risk function of $\tilde{\Theta}_B$ is uniformly smaller than that of the MLE.

==

Theorem 4.3: Let $r_B(\tau)$ denote the risk function of $\tilde{\Theta}_B$ w.r.t. l_1. Then

$$r_M(\tau) - r_B(\tau)$$

$$= 2(\nu_1+\nu_2+\nu_1\nu_2/\tau) - \tau^{-1}E\{[Z(\nu_1+\nu_2)+\nu_1\nu_2]^2/[(Z+\nu_1)(Z+\nu_2)]\}$$

$$- 2E\{[Z^2(\nu_1+\nu_2)+Z\nu_1\nu_2]/[(Z+\nu_1)(Z+\nu_2)]\}. \tag{4.2.39}$$

Proof: The generalized Bayes estimator of θ is a member of P_D, with $G(Z)= [Z(\nu_1+\nu_2) + \nu_1\nu_2]/[(Z+\nu_1)(Z+\nu_2)]$. Hence, after by substituting $(I-1)$ for ν_1, $(J-1)$ for ν_2 and $r_G(\tau)$ for $r_B(\tau)$, equation (4.3.31) may be written as

$$r_M(\tau) = r_B(\tau) - 2E\{[Z^2(\nu_1+\nu_2)+Z\nu_1\nu_2]/[(Z+\nu_1)(Z+\nu_2)]\}$$

$$- \tau^{-1}E\{[Z(\nu_1+\nu_2)+\nu_1\nu_2]^2/[(Z+\nu_1)(Z+\nu_2)]\}$$

$$+ 2\tau^{-1}E\{Z(\nu_1+\nu_2)+\nu_1\nu_2\}, \qquad (4.2.40)$$

Hence $r_M(\tau) - r_B(\tau)$

$$= 2(\nu_1+\nu_2+\nu_1\nu_2/\tau) - \tau^{-1}E\{[Z(\nu_1+\nu_2) + \nu_1\nu_2]^2/[(Z+\nu_1)(Z+\nu_2)]\}$$

$$- 2E\{[Z^2(\nu_1+\nu_2) + Z\nu_1\nu_2]/[(Z+\nu_1)(Z+\nu_2)]\}.$$

==

We now show that the dominating generalized Bayes estimator $\tilde{\Theta}_B$ is not unique; there are many other estimators of the form (4.2.1) which are also members of P_D, and hence dominate $\hat{\Theta}$. Consider the class of estimators $\bar{\bar{\Theta}} = \{\bar{\bar{\Theta}}_{ij}\}$, where

$$\bar{\bar{\Theta}}_{ij} = \hat{\Theta}_{ij}[1-Q(Z)], \qquad \text{if } Z>0,$$

$$= 0 \qquad \text{if } Z=0, \qquad (4.2.41)$$

and $Q(Z)=[(Z+\nu_1)c_2+(Z+\nu_2)c_1-c_1c_2]/[(Z+\nu_1)(Z+\nu_2)]$, for some constants c_1 and c_2. We will show that by appropriate choices of c_1 and c_2 the estimators (4.2.41) also dominate the ML estimator of $\theta= (\alpha,\beta,\tau)$. Indeed, since (4.2.41) is of the same form as (4.2.1), we only need to choose c_1 and c_2 such that

$$(Z+\nu_1)c_2 + (Z+\nu_2)c_1 - c_1c_2 = Z(c_1+c_2) + \nu_1c_2 + \nu_2c_1 - c_1c_2$$
$$\leq 2[Z(\nu_1+\nu_2) + \nu_1\nu_2]. \tag{4.2.42}$$

The inequality (4.2.42) holds provided that

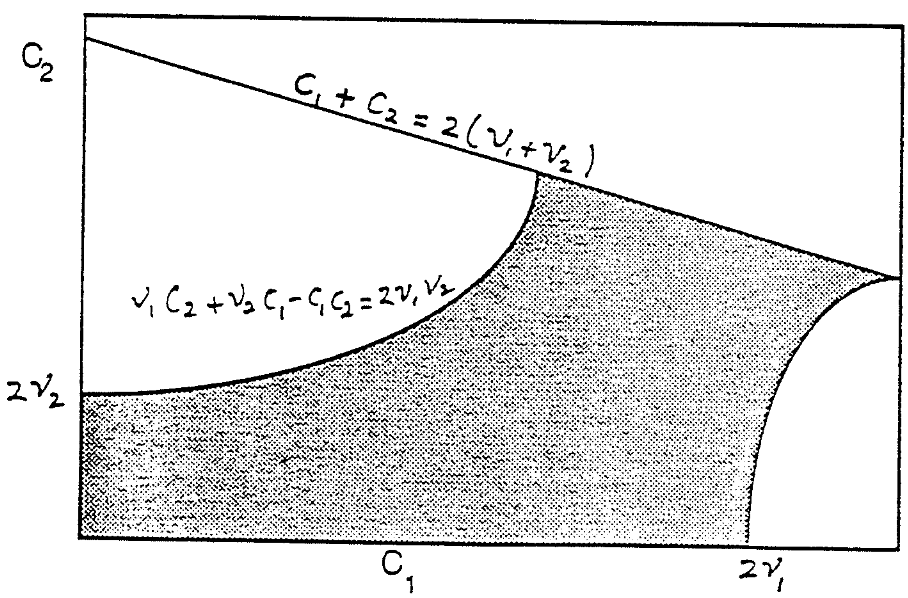
$$c_1 + c_2 \leq 2(\nu_1+\nu_2),$$

and
$$\nu_1c_2 + \nu_2c_1 - c_1c_2 \leq 2\nu_1\nu_2.$$

The set of points (c_1,c_2) which simultaneously satisfies the above inequalities defines a class of estimators which dominate $\hat{\Theta}$. The dominating region defined by c_1 and c_2 are shown in Figure 4.1.

Figure 4.1

A Region Defining a Class of Estimators Dominating the MLE of a Matrix of Poisson Means under a Multiplicative Model for Complete Data Problems



Theorem 4.4: Assume that the prior on $\theta=(\alpha,\beta,\tau)$ is such that α, β and τ are independent and

$$\alpha=(\alpha_1,\ldots,\alpha_I) \sim D_I(0,\ldots,0),$$

$$\beta=(\beta_1,\ldots,\beta_J) \sim D_J(0,\ldots,0),$$

and $\tau \sim \mathrm{Exp}(k)$, exponential distribution with mean k.

Then, w.r.t the loss function l_1 and the given prior on θ, the posterior mode estimator of θ (refer to Theorem 3.3) dominates $\hat{\Theta}$, if k is such that $2k+1 > Z^2/[Z(n_1+n_2) + \nu_1 n_2]$.

Proof: From Theorem 3.3, with $a_i=0 \quad \forall\, i,$, the posterior mode estimator of θ w.r.t. the above priors on α, β and τ is given by

$$\bar{\Theta}_{ij} = \hat{\Theta}_{ij}/Z[1+(k+1)^{-1}], \qquad (4.2.43)$$

and thus is of the same form as the estimator (4.2.1), with $G(Z)=(k+1)^{-1}$, $k>0$. Hence it suffices to show that the function G satisfies the conditions of Theorem 4.1, namely:

1. $G(Z)=(1+k)^{-1}$ implies that $G(Z) \geq 0$ for all $k > 0$.

2. $(1-G(Z)/2)L(Z) \geq Z^2$ implies that

$$(Z+n_1)(Z+n_2)/Z[1-1/(2k+2)]$$

$$=[(2k+1)/(2k+2)][Z^2 + Z(n_1+n_2) + n_1 n_2] \geq Z^3. \qquad (4.2.44)$$

Therefore the prior parameter k is such that $2k+1 > Z^2/[Z(n_1+n_2)+n_1 n_2]$.

As an alternative to considering the component-wise loss l_1, one might consider the loss incurred in estimating the means along the row and column margins of the table. Then the loss function l_2 might be preferred to l_1.

Theorem 4.5: Let $\dot{\Theta}^{\star}_{B} = \{\dot{\Theta}^{\star}_{Bi.}\}$ denote the generalized Bayes estimator of $\theta_{i.} = \sum_j \theta_{ij}$ and let $\Theta^{\star}_{B} = \{\Theta^{\star}_{B.j}\}$ denote the generalized Bayes estimator of $\theta.j = \sum_i \theta_{ij}\}$, w.r.t. the loss function l_2 and the (improper) uniform prior on $\theta = (\alpha, \beta, \tau)$ given in Theorem 4.2. Then

(1) $\dot{\Theta}^{\star}_{Bi.} = R_i[1-n_1/(Z+n_1)]$,

(2) $\Theta^{\star}_{B.j} = S_j[1-n_2/(Z+n_2)]$,

(3) $\dot{\Theta}^{\star}_{B}$ and $\Theta^{\star}_{B}$ dominate the corresponding MLE.

Proof: (1) & (2). Given $Z = \sum_i \sum_j X_{ij}$, the joint distribution of $R_i = \sum_j X_{ij}$ $(i=1,2,\ldots,I)$ and $S_j = \sum_i X_{ij}$ $(j=1,2,\ldots,J)$, w.r.t. the given prior on θ, can be factored as a product of the joint marginal of the R_i's and the joint marginal of the S_j's. The proof is completed by noting that Theorem 2.2 of Clevenson and Zidek (1975) can be applied separately to the R_i's and S_j's.

(3). This follows from Theorem 2.1 of Clevenson and Zidek (1975).

==

Theorem 4.6: With respect to the loss function l_2, the MLE of $\theta = (\alpha, \beta, \tau)$ has a constant risk.

Proof: The MLE of $\theta_{i\cdot} = \sum_j \theta_{ij} = \alpha_i \tau$ is $R_i = \hat{\theta}_{i\cdot} = \sum_j X_{ij}$, and that of $\theta_{\cdot j} = \sum_i \theta_{ij} = \beta_j \tau$ is $\hat{\theta}_{\cdot j} = S_j = \sum_i X_{ij}$. Thus, under the loss function

$$l_2(\hat{\theta};\theta) = \Sigma(\alpha_i\tau)^{-1}[R_i - \alpha_i\tau]^2 + \Sigma(\beta_j\tau)^{-1}[S_j - \beta_j\tau]^2, \tag{4.2.45}$$

the risk of $\hat{\theta}^* = (\hat{\theta}_{i\cdot}, \hat{\theta}_{\cdot j})$ is given by

$$E[l_2(\hat{\theta};\theta)] = E[\sum_i (\alpha_i\tau)^{-1}(R_i - \alpha_i\tau)^2 + \sum_j (\beta_j\tau)^{-1}(S_j - \beta_j\tau)^2], \tag{4.2.46}$$

$$= \Sigma\ (\alpha_i\tau)^{-1} E(R_i - \alpha_i\tau)^2 + \Sigma\ (\beta_j\tau)^{-1}(S_j - \beta_j\tau)^2$$

$$= \Sigma\ (\alpha_i\tau)^{-1}\alpha_i\tau + \Sigma\ (\beta_j\tau)^{-1}\beta_j\tau$$

$$= I + J. \tag{4.2.47}$$

≡≡

Thus, the risk of the estimator $\hat{\theta}$ w.r.t l_2 is independent of θ.

Chapter V

BAYESIAN AND MAXIMUM LIKELIHOOD ESTIMATION OF POISSON MEANS UNDER A MULTIPLICATIVE MODEL: THE CASE OF INCOMPLETE DATA..

5.0 Introduction

In Chapter III, we considered the simultaneous estimation of the IxJ coordinates of θ, the means of independent Poisson random variables, where the observations are arranged in a two-way table of I rows and J columns and the means have a multiplicative structure. We derived the MLE and a class of (generalized) Bayes estimators of θ for a particular family of joint prior distributions of θ under two loss functions l_1 and l_2. In Theorem 4.1, we proved that there exists a class of estimators of θ every member of which dominates the MLE under the loss l_1. Moreover, in Corollary 4.1 we showed that this class contains a particular generalized Bayes estimator corresponding to an (improper) uniform prior on $\theta=(\alpha,\beta,\tau)$. In proving these results, we assumed that each cell had exactly one observation, and that the observations were drawn independently from the Poisson populations. We call this the complete data case, in contrast to the situation in which some cell entries are unavailable.

Bishop, Fienberg and Holland (1975) considered a related problem, namely, the ML estimation of Poisson means in a two-way array under a log linear model for incomplete data. In their model, the cells which have no data in them were assumed to be either 'structural zero' cells or 'random zero' cells. The latter case is attributed to sampling variability and the fact that the cell probabilities are small in these cells, whereas, in the former case, the cell probabilities are known a priori to be zero.

Another possibility of incomplete data is one in which a cell is empty because an observation was not taken in that cell. This situation is analogous to a missing data problem in a two-factor analysis of variance in which certain factor level combinations are not observed. An example of this type of incomplete data is given in Johnston and Brewster (1982). One of the interests in their study was the determination of the number of boats launched at each of three launching sites on each of twenty one selected days. However, on each day, it was possible to place observers at a given site for only two of the seven possible time periods; on each day, there were no observations taken from the remaining five time periods, although it would have been possible to observe the number of launchings in these periods if sufficient manpower were available. In Table 5.1, the data from one of the observed sites has been reproduced. This is the type of incomplete data that we will be considering.

The system of equations defining the MLE for the above type of incomplete-data problem cannot be solved explicitly, but we will show that the EM algorithm can be used to solve the ML equations iteratively. The equations for finding the generalized Bayes estimators also cannot be solved explicitly, and numerical procedures must be used. In this case, however, the EM algorithm cannot be used to find the generalized Bayes estimates. (It should be noted that the EM algorithm can be used to find the generalized Bayes procedure, if one were to define the Bayes procedure to be the mode of the posterior distribution. However, this is not the case under the loss functions l_1 and l_2.)

Although numerical integration techniques could be used to find the form of (generalized) Bayes estimators, these techniques would not be easy to implement because of the multi-dimensional nature of the parameter space.

As an alternative, we will consider an iterative procedure for obtaining an estimator which should, in many cases, be close to the generalized Bayes estimator. The procedure is analogous to the EM algorithm. The M-step (although it is not maximization in this situation) utilizes the form of the generalized Bayes estimator in the complete data problem; the E-step estimates the missing data using the generalized Bayes estimates obtained in the M-step. For convenience, we will designate the estimator so obtained, under the uniform prior on $\theta=(\alpha,\beta,\tau)$, given in Section 3.2, by the EMGB estimator. This is the prior under which the generalized Bayes estimator dominates the MLE in the complete data problem.

One might conjecture that the EMGB estimator would dominate the MLE in the incomplete data problem. This conjecture will be investigated in Chapter VI.

In this chapter we will find the likelihood equations and show how the EM algorithm can be used to find the MLE. We will describe the equations one would need to solve in order to find the generalized Bayes estimator and we will then describe the EMGB estimator.

5.1 The Model and the Likelihood Equations

In this section, we are interested in the estimation of the means $\theta_{ij}=\alpha_i\beta_j\tau$ of the independent Poisson random variable X_{ij} whose distribution is given by (3.1.1). However, not all the cell entries have data in them. For convenience, let S denote the set of all cells (i,j) with no missing data and let

$$\delta_{ij} =1, \text{ if cell } (i,j) \in S$$
$$=0, \quad \text{otherwise.} \tag{5.1.1}$$

That is, $\delta_{ij}=1$ or 0 according as x_{ij} is observed or unobserved. The likelihood is then given by

$$l(\alpha,\beta,\tau\,|\,y,\delta)=\prod^{*} [(\alpha_i\beta_j\tau)^{y_{ij}} \exp(-\alpha_i\beta_j\tau)/y_{ij}!], \tag{5.1.2}$$

$$\propto \prod_i \prod_j \alpha_i^{r_i^*}\beta_j^{s_j^*} \tau^{z^*} \exp(-\delta_{ij}\alpha_i\beta_j\tau),$$

where

$$r_i^*=\sum_j \delta_{ij}y_{ij}, \tag{5.1.3}$$

$$s_j^*=\sum_i \delta_{ij}y_{ij}, \tag{5.1.4}$$

$$z^*=\sum_i \sum_j \delta_{ij}y_{ij}, \tag{5.1.5}$$

and $\prod^*$ denotes the product over cells $(i,j) \in S$. This likelihood equation assumes that, if δ is random, then the distribution of δ is independent of α, β and τ. This would be appropriate, for example, if

Table 5.1

Boat Launching Data arranged in a BIBD
(entries are the number of launchings)

Period \ Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1				9	9				3	7					0			4			
2		9				9				2			0			5				6	
3				14			11	7				24				8					31
4	5				37			10						34			35			24	
5		11	48						51			50					38		7		
6			17			36					47			40				25			33
7	4						4				37		6		13				8		

the missing cells had been selected at random according to one of the following plans:

1. Decide a priori that k cells, will be missing and that all of the ${}^{IJ}C_k$ incomplete samples with k missing cells have the same probability of being selected.

2. There is a fixe. known probability with which any particular cell will be missing, and the missing cells are the outcomes of IJ independent Bernoulli trials.

3. There is some structure to the missing cells. For example, one might decide that non-missing cells would correspond to a balanced incomplete block design (BIBD) with some specified design parameters and that the design would be randomly chosen from those designs satisfying the appropriate conditions. This is the case in the boating problem in Johnston and Brewster (1982).

The third possibility stated above, is the case we will consider in Chapter VI. Rubin (1976) and Little (1979) have discussed other examples of experimental situations in which data can be considered to be missing 'at random'.

From (5.1.2), the log of the likelihood function reduces to

$$l(\alpha,\beta,\tau \mid y,\delta) = \sum_i r_i^* \log\alpha_i + \sum_j s_j^* \log\beta_j$$

$$+ z^* \log\tau - \sum_i \sum_j \delta_{ij}\alpha_i\beta_j\tau + Q(y,\delta), \qquad (5.1.6)$$

$Q(y,\delta)$ is a function of only y and δ. Hence the MLE of α, β and τ, subject to the usual constraints, may be obtained by maximizing the function

$$l^* \equiv l^*(\alpha,\beta,\tau\,|\,y,\delta)$$

$$= l(\alpha,\beta,\tau\,|\,y,\delta) + k_1(1-\Sigma\,\alpha_i) + k_2(1-\Sigma\,\beta_j),$$

$$=\Sigma\, r_i^*\log\alpha_i + \Sigma\, s_j^*\log\beta_j + z^*\log\tau - \Sigma\,\Sigma\,\delta_{ij}\alpha_i\beta_j\tau + Q(y,\delta)$$

$$+ k_1(1-\Sigma\,\alpha_i) + k_2(1-\Sigma\,\beta_j), \tag{5.1.7}$$

where k_1 and k_2 are Lagrangian multipliers. Differentiating (5.1.7) partially w.r.t. α, β, τ, k_1 and k_2 and setting the result to zero, we obtain the equations

$$\partial l^*/\partial\alpha_i = r_i^*/\alpha_i - \tau\sum_j \delta_{ij}\beta_j - k_1 = 0, \tag{5.1.8}$$

$$\partial l^*/\partial\beta_j = s_j^*/\beta_j - \tau\sum_i \delta_{ij}\alpha_i - k_2 = 0, \tag{5.1.9}$$

$$\partial l^*/\partial\tau = z^*/\tau - \sum_i\sum_j \delta_{ij}\alpha_i\beta_j = 0, \tag{5.1.10}$$

$$\partial l^*/\partial k_1 = (1-\sum_i \alpha_i) = 0, \tag{5.1.11}$$

$$\partial l^*/\partial k_2 = (1-\sum_j \beta_j) = 0. \tag{5.1.12}$$

Equations (5.1.8) through (5.1.10) may be rewritten respectively as

$$\hat{\alpha}_i = r_i^*/(k_1 + \tau\sum_j \delta_{ij}\hat{\beta}_j), \tag{5.1.13}$$

$$\hat{\beta}_j = s_j^*/(k_2 + \tau\sum_i \delta_{ij}\hat{\alpha}_i), \tag{5.1.14}$$

and

$$\hat{\tau} = z^*/\sum_i\sum_j \delta_{ij}\hat{\alpha}_i\hat{\beta}_j. \tag{5.1.15}$$

We note that each denominator term in (5.1.13) depends on the β_j corresponding to observations in the non-empty cells in row i. That is, α_i is not proportional to the row i total, R_i, as was the case for the complete data problem. The same remark can be made of the denominators in (5.1.14). Explicit solutions cannot be obtained for the above system of equations. In Section 5.3, we will describe an iterative solution based on the EM algorithm.

5.2 Generalized Bayes Estimators under Loss Functions l_1 and l_2:

We will assume in this section that the joint prior density function of α, β and τ, is the same as stated in Section 3.2 of Chapter III, whether or not the two-dimensional table is complete. Hence the joint posterior density of α, β and τ, given the incomplete data $\{y_{ij}\}$ is

$$g(\alpha,\beta,\tau|y) \propto \prod_i \prod_j \alpha_i^{r_i^{*+}} \beta_j^{s_j^{*+}} \tau^z \exp(-\delta_{ij}\alpha_i\beta_j\tau) M(d\tau),$$

where $r_i^{*+} = r_i^* + a_i$, $s_j^{*+} = s_j^* + b_j$ and $z = \sum r_i^*$. The generalized Bayes estimate of $\theta=\{\theta_{ij}\}$ w.r.t. to the loss function l_1 is denoted by $d=\{d_{ij}\}$, where $d_{ij}=u_i v_j w$. The values of $u=(u_1,\ldots,u_I)$, $v=(v_1,\ldots,v_J)$ and w are chosen to minimize the posterior risk for a fixed $x=\{x_{ij}\}$ and a specified joint (possibly improper) prior on $\theta=(\alpha,\beta,\tau)$, subject to the constraints $u_i>0$, $v_j>0$, $w>0$, $\sum u_i=1$ and $\sum v_j=1$. The posterior risk is given by

$$R(u,v,w|y) = \int l_1(u,v,w;\alpha,\beta,\tau)\; g(\alpha,\beta,\tau|y)\; d\alpha\; d\beta\; d\tau, \qquad (5.2.1)$$

If the prior on $\theta=(\alpha,\beta,\tau)$ is proper, then the resulting Bayes estimator will also minimize the (overall) risk

$$R = \int R(u,v,w|y)h(y)dy, \tag{5.2.2}$$

where $h(y)$ is the marginal density function of Y. The generalized Bayes estimator of θ may be obtained by differentiating $R(u,v,w|y)$ w.r.t. w and the components of u and v, equating the results to zero and solving the resulting equations simultaneously.

However, it is clear from equation (5.2.1), that the contributions from the exponential factors $\exp(-\delta_{ij}\alpha_i\beta_j\tau)$ do not collapse to one term, as was the case in the complete data problem. The α, β and τ terms in the exponential factor are inextricably tied together and therefore appropriate factoring of the posterior distribution is not possible. There is no feasible way to integrate out the terms individually.

5.3 Iterative Solutions via the EM Algorithm:

We have already shown in Section 3.1 that the complete data likelihood equations can be solved explicitly. Clearly, the complexity of the ML equations in the incomplete data case is due to a lack of symmetry introduced by the empty cells. The EM algorithm was introduced in Chapter II as an iterative scheme for finding the MLE in incomplete data problems. For the present problem, the first step of the EM algorithm is the estimation of observations in the missing cells. This requires the arbitrary specification of values to be used as initial estimates of α, β and τ. Denote these by $\alpha^{(0)}$, $\beta^{(0)}$ and

$\tau^{(0)}$ respectively, and suppose that cell (s,t) is missing. Then the missing entry in cell (s,t) is estimated by $x_{st}^{(0)}=\alpha_s^{(0)}\beta_t^{(0)}\tau^{(0)}$. A new set of estimates $\alpha^{(1)}$, $\beta^{(1)}$ and $\tau^{(1)}$ is computed using the observed data and the estimated values of the missing cells. Suppose that $\theta_{ij}^{(k)}$ is the estimate of θ_{ij} at the k-th iteration. The conditional expectation of X_{ij}, given that $\theta_{ij} = \theta_{ij}^{(k)}$, is then

$$E[X_{ij}|\theta_{ij}^{(k)}]=\theta_{ij}^{(k)}= \alpha_i^{(k)}\beta_j^{(k)}\ \tau^{(k)} \qquad (5.3.1)$$

so that current estimate of the missing entry in cell (s,t) is $\theta_{st}^{(k)}$. In general the EM algorithm is defined by the following steps:

E-step: Estimate the complete data $x=\{x_{ij}\}$:

$$x_{ij}^{(k+1)}= \delta_{ij}y_{ij}+(1-\delta_{ij})\theta_{ij}^{(k)}. \quad (i=1,...,I;\ j=1,...,J) \qquad (5.3.2)$$

M-step: $\theta_{ij}^{(k+1)} = r_i^{(k+1)}\ s_j^{(k+1)}/z^{(k+1)}$, (5.3.3)

where $r_i^{(k+1)}= \sum_j x_{ij}^{(k+1)}$, $s_j^{(k+1)}= \sum_i x_{ij}^{(k+1)}$, $z^{(k+1)}= \sum_i \sum_j x_{ij}^{(k+1)}$,

and $\theta_{ij}^{(k+1)}$ is given by (5.3.1). The sequence of E- and M- steps is repeated until a specified level of accuracy is obtained. For the types of data we have encountered so far, this iterative process appears to be largely unaffected by the choice of initial values, provided that reasonable values are used. For example, the choice of $\alpha^{(0)}$ must be such that $\sum \alpha_i^{(0)}=1$. It should be observed that it not necessary to obtain estimates of α, β and τ separately at each step. In fact it is computationally efficient to obtain the θ_{ij} first. Since

the estimates of the complete data sufficient statistics r_i, s_j and z are now available (from the M step), the MLE of α, β and τ can be obtained
quite simply as $\hat{\alpha}_i=r_i/z$, $\hat{\beta}_j=s_j/z$ and $\hat{\tau}=z$.

We now consider the EM-version of the generalized Bayes procedure for incomplete data. For the loss function l_1, we propose the following iterative scheme:

E-step: $$x_{ij}^{(k)} = y_{ij}\delta_{ij} + (1-\delta_{ij})\theta_{ij}^{(k)}. \tag{5.3.4}$$

M-step: $$\theta_{ij}^{(k+1)}=r_i^{(k)}s_j^{(k)}/z^{(k)}\{[\ z^{(k)}/(z^{(k)}+A)][z^{(k\ )}/(z^{(k)}+B)]\}, \tag{5.3.5}$$

where $$r_i^{(k)}=\sum_j x_{ij}^{(k)}+a_i, \tag{5.3.6}$$

$$s_j^{(k)}=\sum_i x_{ij}^{(k)}+b_j, \tag{5.3.7}$$

$$z^{(k)}=\sum_i\sum_j x_{ij}^{(k)}, \tag{5.3.8}$$

$A = \sum a_i$ and $B = \sum b_j$. The EMGB estimator is defined as the limit of the sequence $\{\theta_{ij}^{(k)}\}$ given by (5.3.4) and (5.3.5). In the E-step, the complete table observations x_{ij} are estimated. This is achieved by filling in each missing entry (i,j) by the conditional expectation given $\theta_{ij}^{(k)}$ of the associated random variable X_{ij}. That is,

$$E[X_{ij} | \theta_{ij}^{(k)}] = \theta_{ij}^{(k)} = \alpha_i^{(k)} \beta_j^{(k)} \tau^{(k)}. \qquad (5.3.9)$$

In the M step, the $x_{ij}^{(t)}$ are treated as observed data of a complete two-way table with IJ entries, a situation for which equations (3.3.3) and (3.3.4) yield the generalized Bayes estimates. Thus, each statistic in (3.3.4) replaced by the estimate computed with the $x_{ij}^{(k)}$'s.

In particular, if the prior distribution on $\theta = (\alpha, \beta, \tau)$ is an improper uniform distribution, that is, $a_i=0$, $\forall$ i, $b_j=0$, $\forall$ j and $M(d\tau)=d\tau$ on $(0,\infty)$, then the generalized Bayes estimator of θ may be obtained from (5.3.4) and (5.3.5) by setting the a_i's and b_j's equal to zero.

We remark that the EMGB estimator is not the generalized Bayes estimator in the above incomplete data problem. However, the EMGB procedure deserves some consideration because it is computationally appealing. As shown in the simulation study in Example 4 of Chapter II, (in the complete-data case), the EMGB estimates are likely to be close to the (generalized) Bayes estimates.

CHAPTER VI

A COMPARISON OF THE RISKS OF THE MLE AND THE EMGB ESTIMATOR OF A MATRIX OF POISSON MEANS: A SIMULATION STUDY FOR INCOMPLETE DATA.

6.0 Introduction

In Theorem 4.1 of Chapter IV, we presented a class of estimators which dominate the ML estimator of θ, a matrix of Poisson means under a multiplicative model. Included in this class is the generalized Bayes estimator corresponding to an improper uniform prior on $\theta=(\alpha,\beta,\tau)$. Being generalized Bayes, there is some hope that this estimator might be admissible, although this has not been proven. It should be observed that Theorem 4.1 pertains to complete-data problems.

In the incomplete-data problem, it is quite natural to conjecture that the corresponding generalized Bayes estimator (with respect to the same prior as above) dominates the MLE of θ. The proof of this assertion would involve a comparison of the risks of the estimators, as was done in the proof of Theorem 4.1 (for the complete-data problem). However, in this case, the risks cannot be obtained explicitly and simulation will thus be used. Moreover, rather than considering the actual generalized Bayes estimator, which is difficult to evaluate, we will consider the EMGB estimator as given in Chapter V.

In this chapter, we will use simulation to compare the risks of the EMGB estimator and the MLE.

6.1 Creating a Balanced Incomplete Table with Independent Poisson Observations

Because of cost considerations, the simulation study will be restricted to 3x3 incomplete tables. Moreover, we will consider the situation in which the observed cells correspond to a balanced incomplete block design, with two observations in each row and column. Thus, there are six possible designs, one of which has the missing cells on the main diagonal.

For a 3x3 table, there are 9 unknown means θ_{ij} $(i=1,2,3;\ j=1,2,3)$ to be estimated. However, because of the multiplicative structure on the means, that is, $\theta_{ij}=\alpha_i\beta_j\tau$, $(\alpha_i>0,\ \beta_j>0,\ \Sigma\,\alpha_i=1,\ \Sigma\,\beta_j=1)$, only the 5 unknown parameters α_1, α_2, β_1, β_2 and τ need to be considered. We showed in Chapter IV that the risk functions of the MLE and the generalized Bayes estimator in the complete-data problem are independent of the parameters α and β, and depend on θ only through τ. We will show that this is not true in the incomplete-data problem.

The parameter space is five-dimensional, with continuous ranges of values for τ and the components of α and β; this makes it difficult and expensive to examine extensively the set of possible parameter values. In this study, the parameter space will be restricted by the assumption that two of the coordinates of $\alpha=(\alpha_1,\alpha_2,\alpha_3)$ and $\beta=(\beta_1,\beta_2,\beta_3)$ are equal. Without loss of generality, it will be assumed that $\alpha_1=\alpha_2$ and $\beta_1=\beta_2$.

Define $A=B=\{0.1,\ 0.2,\ 0.3,\ 1/3,\ 0.4\}$ and $C=\{0,\ 25,\ 50,\ldots,\ 550\}$. Let $\alpha^*=\alpha_1=\alpha_2$ and $\beta^*=\beta_1=\beta_2$ so that $\alpha_3=1-2\alpha^*$ and $\beta_3=1-2\beta^*$. To allow for a reasonable coverage of the range of values for the parameters, and also because of the practical difficulties mentioned in the last paragraph, the following scheme will be used to select the values of the parameters:

Choose $\theta_{ij}=\alpha^*\beta^*\tau$, where $\alpha^* \in A$, $\beta^* \in B$ and $\tau \in C$.

That is, we will consider only Poisson densities for which α^* is restricted to values in the set A, β^* is restricted to the values in the set B for each value of α^*, and τ is restricted to the values in the set C for every combination of α^* and β^*. For example, if $\alpha^*=0.2$, $\beta^*=0.1$ and $\tau=150$, then the means θ_{ij} for the cells in the table will be as shown in Table 6.1 below.

The observations in the cells are independently generated from the respective Poisson density functions according to the values of the cell means and also depending on whether the cell is designated to be non-empty or empty. The observations are generated only for the non-empty cells.

Table 6.1

The Means of Poisson Populations in a 3x3 Table for $\alpha^*=0.2$, $\beta^*=0.1$ and $\tau=150$.

	$\beta_1=0.1$	$\beta_2=0.1$	$\beta_3=0.8$
$\alpha_1=0.2$	3.0	3.0	24.0
$\alpha_2=0.2$	3.0	3.0	24.0
$\alpha_3=0.6$	9.0	9.0	72.0

In order to keep track of empty and non-empty cells, we define two 'incidence matrices' IM1 and IM2 as follows:

$$IM1=\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

$$IM2=\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

An entry is 0 or 1 according as the associated cell is empty or not empty. For any combination of α^*, β^* and τ, there are 6 different configurations of balanced incomplete tables. However, each of these configurations can be transformed into either IM1 or IM2 by a suitable permutation of rows and columns. For example, suppose $\alpha^*=0.2$, $\beta^*=0.1$ and $\tau=1.0$. Then the following balanced incomplete tables are possible:

I:

	$\beta_1=0.1$	$\beta_2=0.1$	$\beta_3=0.8$
$\alpha_1=0.2$	0	1	1
$\alpha_2=0.2$	1	0	1
$\alpha_3=0.6$	1	1	0

II:

	$\beta_1=0.1$	$\beta_2=0.1$	$\beta_3=0.8$
$\alpha_1=0.2$	0	1	1
$\alpha_2=0.2$	1	1	0
$\alpha_3=0.6$	1	0	1

III:

	$\beta_1=0.1$	$\beta_2=0.1$	$\beta_3=0.8$
$\alpha_1=0.2$	1	0	1
$\alpha_2=0.2$	0	1	1
$\alpha_3=0.6$	1	1	0

IV:

	$\beta_1=0.1$	$\beta_2=0.1$	$\beta_3=0.8$
$\alpha_1=0.2$	1	0	1
$\alpha_2=0.2$	1	1	0
$\alpha_3=0.6$	0	1	1

V:

	$\beta_1=0.1$	$\beta_2=0.1$	$\beta_3=0.8$
$\alpha_1=0.2$	1	1	0
$\alpha_2=0.2$	1	0	1
$\alpha_3=0.6$	0	1	1

VI:

	$\beta_1=0.1$	$\beta_2=0.1$	$\beta_3=0.8$
$\alpha_1=0.2$	1	1	0
$\alpha_2=0.2$	0	1	1
$\alpha_3=0.6$	1	0	1

It should be noted that in each of the above tables, there is exactly one empty cell in every row and in every column. By rearranging the rows and columns, it can be shown that the incidence matrices I and III are the same as IM1. Similarly, it can be shown that the incidence matrices II, IV, V and VI are the same as IM2. Thus, it is sufficient to examine only two types of balanced incomplete tables for every set of permissible values of α^*, β^* and τ ---those using IM1 and IM2.

For each of IM1 and IM2, a balanced incomplete table of independent Poisson observations was generated using the Poisson random variate generator GGPON from the IMSL subroutine. For each table, the ML and the EMGB estimates were computed. The loss incurred by each estimate under the loss l_1 was also computed. This procedure was repeated 2000 times, for each combination of α^*, β^* and τ.

6.2 Computing the Estimated Risks of the MLE and the EMGB Estimator

We are interested in a comparison of the risk functions of the MLE and the EMGB estimator. Thus, in estimating the two risk functions through simulation, it makes sense to evaluate both estimators on the same data set. If a particular data set produces poor estimates using one estimator, then it will likely produce poor estimates using the other estimator. That is, we are using what corresponds to a paired experiment and this has bearings on our assessment of the reliability of the comparisons (see Section 6.5 below).

Two different risks can be defined for an estimator in the above incomplete-data problem. In the first case, the risk is the usual expected loss with respect to the particular table that has been selected. In the second case, which corresponds to a random table, the risk is the expected loss over all the possible tables that could have been selected (that is, over the six possible tables in a 3x3 situation). We will consider both of these risks in comparing the estimators. In order to distinguish between them, the risk in the latter case will be called the 'ensemble risk' of the estimator. The risks will be computed separately for each of IM1 and IM2. As shown above, tables corresponding to IM2 occur twice as often as those corresponding to IM1, if the tables are selected at random. Thus, the ensemble risk of each estimator will be computed as the weighted average of the risks from IM1 and IM2, the weights being in a 1:2 ratio. A measure of performance of the EMGB estimator relative to the MLE is defined to be the 'gain' in risk using the EMGB estimator instead of the MLE. This gain is obtained by subtracting the risk of the EMGB estimator from the risk of the MLE, and dividing by the risk of the MLE.

The values of α^*, β^* and τ that were used in the simulations are shown in Appendix C2. A Fortran program, which performed all of these operations on the Amdahl 360/370 computer at the University of Manitoba, is listed in Appendix C3.

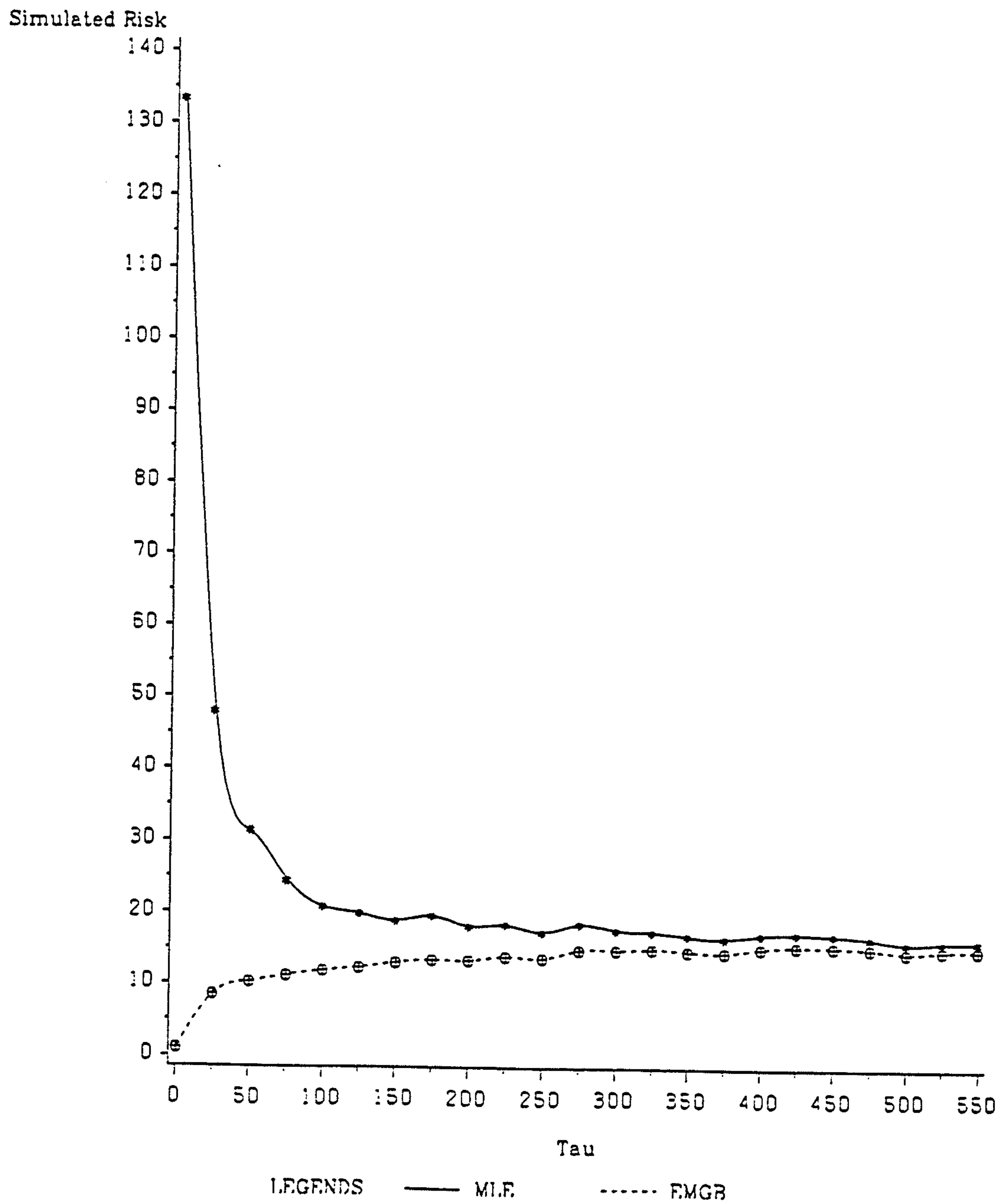
6.3 Displaying the Results of the Simulations

The results of the simulations are presented in a series of two-dimensional plots of the simulated risks against θ using the GPLOT procedure in the SAS/GRAPH statistical package. For $\alpha^*=0.1$ and $\beta^*=.4$, the graph of the simulated risk functions of MLE and the EMGB estimator are shown in Figure 6.2. The graphs for other parameter settings are shown in Appendix E. The two-dimensional graphs were obtained by plotting the risks against τ, for each combination of $\alpha^* \in$ A and $\beta^* \in$ B. Thus, each graph may be interpreted as a plot of the relationship between the risk of the estimator and the parameter τ for

Figure 6.2

A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Matrix of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.
Incidence Matrix is IM2.

ALPHA=(0.1, 0.1, 0.8); BETA=(0.4, 0.4, 0.2)



known values α^* and β^*. To allow the risks of the estimators to be compared visually, the plots have been overlaid on the same page.

We remark that the curves shown in Figure 6.2 and Appendix E pass through the actual values obtained in the simulations. In obtaining the plots, we simply joined adjacent points using a spline algorithm in SAS/GRAPH, which forces the graph to go through each of the data points. Thus, the algorithm does not 'fit' the data, in the sense of regression analysis. Although curve-fitting procedures, such as non-linear regression, can be explored to obtain smoother estimates of the risk functions, it will be difficult to implement them in this problem. The least squares fitting-algorithm, for example, would require the specification of an equation describing the functional relationship between the risk of the estimator and the parameters α, β and τ. Because there are many parameters involved, the selection of an appropriate equation is a formidable practical problem.

6.4 Discussion of the Results of the Simulations

As mentioned above, for $\alpha^*=0.1$ and $\beta^*=.4$, the graphs of the simulated risk functions of the MLE and the EMGB estimator are shown in Figure 6.2. The risk functions for the parameter settings we have considered are shown in Appendix E.

A close study of the graphs of the risk functions shows that, over the subset of the parameter space used in the simulations, the risk of the MLE is consistently greater than the risk of the EMGB estimator. The difference in the risk functions is more pronounced for values of τ in the neighbourhood of zero. However, this difference tends to decrease with increasing values of τ, with the result that the risk functions are almost identical for $\tau \geq 550$.

6.5 Reliability of the Simulated Risks

The graphs shown in Figure 6.2 and Appendix E are estimates of the risk functions of the MLE and the EMGB estimator, where each point is based on 2000 observations (incomplete 3x3 tables). It is of interest to study the precision associated with these estimates in order to determine the adequacy of the sample size used in the simulations. In order to assess the reliability of the results obtained above, we will study the distribution of the losses at each of four selected points of $\theta=(\alpha,\beta,\tau)$.

Univariate plots (histograms) of the losses of the MLE and EMGB estimator corresponding to the four settings of $\alpha^*=0.4$, $\beta^*=0.4$ and $\tau=2.0$, 20.0, 200.0, 500.0 were obtained using the GCHART procedure in SAS/GRAPH (see pages 157-164). Histograms of DIFF = Loss(MLE) - Loss(EMGB) were also obtained. The plots are shown in Appendix E. It should be noted that the class intervals are not uniformly scaled; values around zero were assigned shorter interval widths than those further away from zero, with the result that it was possible to compare the losses associated with the estimates over a wider range of values.

From the histograms of the losses, it is observed that most of the losses associated with the EMGB estimator fall in a narrow interval around zero, while the losses associated with the MLE are widely scattered over a wider interval. Also, a significant number of big losses are observed for the MLE. These observations are also evident from the distributions of DIFF. The risk (average loss) of the MLE is higher for all of the cases considered.

The mean and the standard deviation of the 2000 losses associated with the MLE and the EMGB estimator were obtained for each of the four selected values of θ: $\alpha^*=0.4$, $\beta^*=0.4$ and $\tau=2.0$, 20.0, 200.0, 500. The mean and the standard deviation of the 2000 values of DIFF were also computed. These are shown in Table A1, A2, and A3 (refer to Appendix A).

Because of the paired nature of the experiment, the mean and the standard deviation of DIFF are the appropriate statistics to use in comparing the risks of the estimators.

Table 6.4

Comparing the Distribution of Losses of the MLE and the EMGB Estimates of a Vector of Poisson Means under a Multiplicative Model in a 3x3 Table. Summary Statistics for 2000 Tables.

$\alpha^*=0.4$; $\beta^*=0.4$; $\tau=20.0$

VARIABLE	MLE	EMGB	DIFF
N	2000	2000	2000
MEAN	88.419	6.768	81.651
STANDARD DEV	354.143	4.869	352.720
MINIMUM VALUE	0.200	0.200	-7.400
MAXIMUM VALUE	5659.600	65.300	5609.200
STD ERROR OF MEAN	7.919	0.109	7.887
SUM	176837.400	13536.000	163301.600
VARIANCE	125417.388	23.702	124411.111
C.V.	400.530	71.934	431.985
LOWER BOUND	72.898	6.554	66.193
UPPER BOUND	103.940	6.982	97.109

From these results, 95% confidence intervals for the risk of the MLE, the risk of the EMGB estimator and the difference between the two risks were obtained. For example, with $\alpha^*=0.4$, $\beta^*=0.4$ and $\tau=20.0$, the average loss associated with the MLE was 88.419 and that of the EMGB estimator was 6.768 (refer to Table 6.4 above); the standard errors of these estimated risks were 7.919 and 0.109 respectively. Hence a 95% confidence interval for the risk of the MLE is 88.419 ± 15.521 or (72.898, 103.940); a 95% confidence interval of the risk of the EMGB estimator is 6.768 ± 0.214 or (6.554, 6.982). It would appear that the risk functions of the estimators have different values at the specified θ value, although a formal test of hypothesis is based on the analysis of DIFF. In fact, from Table 6.4, the mean of DIFF is 81.651 and the standard error is 7.881. Thus a 95% confidence interval for the difference between the risks of the MLE and the EMGB estimator (for $\alpha^*=0.4$, $\beta^*=0.4$ and $\tau=20.0$) is 81.651 ± 15.447; a significant difference clearly exists between the risks of the estimators at this value of θ. The confidence intervals corresponding to other parameter settings were obtained in a similar manner. For the values of $\tau=2.0$, 20.0 and 200, a significant difference between the risks of the estimators was observed, whereas no significant difference was observed for $\tau=550$.

Although only four values of θ were considered in this analysis, similar results were obtained for other values. It does not seem unreasonable to conclude that the risk functions of the MLE and the EMGB estimator are different for small and moderate values of τ.

6.6 Limitations and Scope for Further Work

In this study, we have investigated the problem of improving on the MLE of $\underset{\sim}{\theta}$, a matrix of Poisson means under a multiplicative model for complete and incomplete-data problems. In the complete-data case, a class of estimators dominating the MLE has been found and, in particular, it has been shown that the generalized Bayes estimator under a uniform improper prior on θ is included in this class. In the incomplete-data case, it has been demonstrated by simulation that the EMGB estimator dominates the MLE in the special case of a balanced 3x3 table.

However, there are other related problems which require further study. For example, in the complete-data problem, it is of interest to find the form of other generalized Bayes estimators dominating the MLE. Also, in this problem, the admissibility of the generalized Bayes estimator (with respect to the above prior on θ) has not been established.

In the incomplete-data case, our conclusions have been based on simulation results using a balanced 3x3 table. It would thus be desirable to have a formal proof of the dominance of the EMGB estimator, although such a proof will be very difficult in this case. It would also be desirable to consider other incomplete tables (obtained, for example, by removing the restriction of a balanced design) and to examine tables of higher dimensions. Again, in the incomplete-data problems, the simulation results have demonstrated that the risk functions of the estimators depend on α, β and τ, although the exact functional relationship between the risk functions and the parameters α, β and τ has not been obtained. There is thus a need for further research to explore the use of non-linear regression techniques to obtain smoother estimates of the risk functions in the incomplete-data problems.

BIBLIOGRAPHY

AITKIN, M. (1981). A Note on the Regression Analysis of Censored Data. Technometrics **23** 161-163.

AITKIN, M., & WILSON, G. T. (1980). Mixture Models, Outliers, and the EM Algorithm. Technometrics **22** 325-331.

ALAM K. (1975) Minimax and Admissible Estimators of the Mean of a Multivariate Normal Distribution for unknown Covariance Matrix. J. Multiv. Anal. **5** 83-95.

BAUM, L. E., PETRIE, T., SOULES, G., & WEISS, N. (1970). A Maximization Technique occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains. Ann. Math. Statist. **41** 164-171.

BARANCHIK, A. J. (1964). Multiple Regression and Estimation of the Mean of a Multivariate Normal Distribution. Technical Report No. **51**, Stanford University.

BARANCHIK, A. J. (1970). A Family of Minimax Estimators of the Mean of a Multivariate Normal Distribution. Ann. Math. Statist. **41** 642-645.

BHATTACHARYA, P. K. (1966). Estimating the Mean of the Multivariate Normal Population with General Quadratic Loss Function. Ann. Math. Statist **37** 1819-1824.

BERGER, J. (1974). Minimax Estimation of Location Vectors For a Wide Class of Densities. Ann. Statist. **3** 1318-1328.

BERGER, J. (1976). Admissible Minimax Estimation of a Multivariate

Normal Mean with Arbitrary Quadratic Loss. Ann. Statist. **4** 223-226.

BIRCH, M. W. (1963). Maximum Likelihood in Three-Way Contingency Tables. J. R. Statist.Soc. B **25** 220-233.

BISHOP, Y. M. M., & FIENBERG, S. E. (1969). Incomplete Two-Dimensional Contingency Tables. Biometrics **25** 119-128.

BISHOP, Y. M. M., FIENBERG, S. E., & HOLLAND, P. W. (1975). Discrete Multivariate Analysis: Theory and Practice. The MIT Press, Cambridge, Massachusetts.

BLIGHT, B. J. N. (1970). Estimation from a Censored Sample for the Exponential Distribution. Biometrika **57** 389-395.

BOCK, M. E. (1975). Minimax Estimators of the Mean of the Multivariate Normal Distribution. Annals of Statistics **3** 209-218.

BOYLES, R. A. (1983). On the Convergence of the EM Algorithm. J. R. Statist. Soc. B **44** 47-50.

BROWN, L. D. (1966). On the Inadmissibility of Invariant Estimators of One or More Location Parameters. Ann. Math. Statist. **37** 1087-1136.

BROWN, L. D. (1975). Estimation with Incompletely Specified Loss Functions (The Case of Several Location Parameters). J. Amer. Statist. Assoc. **70** 417-427.

BROWN, M. L. (1974). Identification of the Sources of Significance in Two-Way Tables. Appl. Statist. **23** 405-413.

CHAN-FU-CHEN (1979). Bayesian Inference for a Normal Dispersion Matrix and its Application to Stochastic Multiple Regression. J. R. Statist. Soc. B **41** 235-248.

CHENG, S. W. & FU, J. C. (1982). Estimation of Mixed Weibull Parameters in Life Testing. IEEE Trans. **R-31** 377-381.

CLEVENSON, M. L., & ZIDEK, J.V. (1975). Simultaneous Estimation of the Means of Independent Poisson Laws. J. Amer. Statist. Assoc. **70** 698-705.

COHEN, A. C. (1954). Estimation of the Poisson Parameter from Truncated Samples and from Censored Samples. J.Amer. Statist. Assoc. **49** 158-168.

COHEN, A. C. (1960). Estimation in Truncated Poisson Distributions when Zeros and some Ones are Missing. J. Amer. Statist. Assoc. **55** 342-348.

COHEN, A. C. (1970). On the Solution of Estimating Equations for Truncated and Censored Samples from Normal Populations. Biometrika **44** 225-236.

DAY, N. E. (1969). Estimating the Components of a Mixture of Normal Distributions. Biometrika **56** 463-474.

DEMPSTER, A. P., LAIRD, N. M., & RUBIN, D. B. (1977). Maximum Likelihood from Incomplete Data via the EM algorithm (with discussion). J. R. Statist.Soc. B **39** 1-38.

DEMPSTER, A. P., RUBIN, D. B., & TSUTAKAWA, R. K. (1981). Estimation in Covariance Models. J. Amer. Statist. Assoc. **76** 341-353.

EFRON, B. & MORRIS, C. (1971). Limiting the Risk of Bayes and Empirical Bayes Estimators, Part I: The Bayes Case. J. Amer. Statist Assoc. **70,** 311-319.

EFRON, B. & MORRIS, C. (1972). Limiting the Risk of Bayes and Empirical Bayes Estimators, Part II: The Empirical Bayes Case.

J. Amer. Statist Assoc. **70** 311-319.

EL-SAYYAD, G. M. (1967). Estimation of the Parameters of an Exponential Distribution. J. R. Statist. Soc. B, **29** 525-532.

FAITH, R. E. (1978). Minimax Bayes Estimators of a Multivariate Normal Mean J. Mult. Anal. **8** 372-379.

FERGUSON, T. S. (1967). Mathematical Statistics: A Theoretical Decision Approach. Academic Press, New York.

FIENBERG, S. E. (1970). Quasi-Independence and Maximum Likelihood Estimation in Incomplete Contingency Tables. J.Amer. Statist. Ass. **65** 1610-1616.

FIENBERG, S. E. & HOLLAND, P. W. (1975). Methods of Estimating Zero Counts in Contingency Tables. In Random Counts in Models and Structures (Editor: G. P. Patil), Pennslyvania State University Press, Pa.

FUCHS, C. (1982). Maximum Likelihood Estimation and Model Selection in Contingency Tables with Missing Data. J.Amer. Statist. Ass. **77** 270-278.

GHOSH, M., HWANG, J. T., & TSUI, K. W. (1983). Construction of Improved Estimators in Multiparameter Estimation for Discrete Exponential Families. Ann. Statist. **11** 351-367.

GHOSH, M., & PARSIAN, A. (1981). Bayes Minimax Estimation of Multiple Parameters. J. Multivariate Anal. **11** 280-288.

GOOD, I. J. (1956). On the Estimation of Small Frequencies in Contingency Tables. J. R. Statist. Soc. B **18** 113-124.

GOODMAN, L. A. (1968). The Analysis of Cross-Classified Data: Independence, Quasi-Independence, and Interactions in Contingency

Tables with or without Missing Entries. J. Amer. Statist. Ass. **63** 1091-1131.

GRIFFIN, B. S. & KRUTCHKOFF, R. G. (1971). Optimal Linear Estimators: An Empirical Bayes Version with Application to the Binomial Distribution. Biometrika **58** 195-201.

GUPTA, A. K. (1952). Estimation of the Mean and Standard Deviation of a Normal Population from a Censored Sample. Biometrika, **39** 260-273.

HABERMANN, S. J. (1974). Loglinear Models for Frequency Tables Derived by Indirect Observations: Maximum Likelihood Equations. Ann. Statist. **2** 911-924.

HAIGHT, F. A. & FISHER G. R. (1968). Maximum Likelihood Estimation for the Truncated Poisson. Biometrics **22** 620-624.

HARTLEY, H. O. (1958).Maximum Likelihood Estimation from Incomplete Data. Biometrics **14** 174-194.

HARTLEY, H. O., & HOCKING, R. R. (1971). The Analysis of Incomplete Data (with discussions). Biometrics **27** 783-823.

HASSELBLAD, V. (1966). Estimation of Parameters for a Mixture of Normal Distributions. Technometrics **8** 431-444.

HODGES, J. L. & LEHMANN, E. L. (1951). Some Applications of the Cramer-Rao Inequality. Proc. Second Berkeley Symp. Math. Statist. Prob. 13-20: University of California Press, Berkeley.

HOSMER, D. W. (1973). On the MLE of the Parameters of a Mixture of Two Normal Distributions when the Sample Size is Small. Comm. Statist. **1** 217-227.

HOSMER, D. W. (1973). A Comparison of Iterative Maximum Likelihood

Estimates of the Parameters of a Mixture of Two Normal Distributions under Three Different Types of Samples. Biometrics **29** 761-770.

HUDSON, H. M. (1985). Adaptive Estimators for Simultaneous Estimation of Poisson Means. Ann. Statist. **13** 246-261.

HWANG, J. T. (1982). Improving upon Standard Estimators in Discrete Exponential Families with Application to Poisson and Negative Binomial Cases. Ann. Statist. **10** 857-867.

IRWIN, J. O. (1959). On the Estimation of the Mean of a Poisson Distribution from a Sample with the Zero Class Missing. Biometrics **15** 324-326.

JAMES, W & STEIN, C. (1961). Estimation with Quadratic Loss. Proc. Fourth Berkeley Symp. Math. Statist. Prob. **1** 361-379, Univ. of California Press.

JOHNSTON, B. & BREWSTER, J. F. (1982). The use and Analysis of a BIBD in a Sample Survey. Appl. Statist. **31** 101-107.

JOHNSON, N. L., & KOTZ, S. (1967). Distributions in Statistics - Continuous Distributions, Houghton Miffin Company, Boston.

JOHNSON, B. M. (1971).)n the Admissible Estimators for Certain Fixed Sample Binomial Problems. Ann. Math. Statist. **42** 1579-1582.

KALE, B. K. (1964). Some Remarks on Truncation and Admissibility in Exponential Families. Sankhya **26** 55-62.

LAIRD, N. M. (1978). Nonparametric Maximum Likelihood Estimation of a Missing Distribution. J. Amer. Statist. Ass. **73** 805-811.

LAIRD, N. M. (1978). Empirical Bayes Methods for Two-Way Contingency Tables. Biometrika **65** 581-590.

LAIRD, N. M. & LOUIS, T. A. (1982). Approximate Posterior Distributions for Incomplete Data Problems. J. R. Statist. Soc. B **44** 190-200.

LEONARD, T. (1975). Bayesian Estimation Methods for Two-Way Contingency Tables. J. R. Statist. Soc. **B** 23-37.

LEONARD, T. (1986). Bayesian Full Rank Marginalization for Two-Way Contingency Tables. J. Educ. Statist. **11** 33-56.

LEHMANN, E. L. (1959). Testing Statistical Hypothesis. John Wiley, New York.

LEHMANN, E. L. (1983). Theory of Point Estimation, John Wiley, New York.

LITTLE, R. J. A. (1979). Maximum Likelihood Inference for Multiple Regression with Missing Values: A Simulation Study. J. R. Statist. Soc. B **41** 76-87.

LOUIS, T.A. (1983). Finding the Observed Information Matrix when using the EM Algorithm. J. Amer. Statist. Assoc. B **44** 226-233.

MENDENHALL, W., & HADER, R. J. (1958). Estimation of parameters of Mixed Exponentially Distributed Failure-Distributions from Censored Life Test Data. Biometrika **45** 504-526.

MOOD, A. M., GRAYBILL, F. A., BOES, D. C. (1974). Introduction To the Theory of Statistics. McGraw-Hill, London.

MORGAN, B. J. T., & TITTERINGTON, D. M. (1977). A Comparison of Iterative Methods for Obtaining Maximum Likelihood Estimates in Contingency Tables with a Missing Diagonal. Biometrika **64** 265-269.

MURRAY, G. D. (1977). Contribution to Discussion of Paper by A. P.

Dempster, N. M. Laird and D. B. Rubin. J. Roy. Statist. Soc. B **39** 27-28.

OLSON, D. M. (1974). A Sequential Simplex Program for Solving Minimization Problems. J. Qual. Tech. **6** 53-57.

OLSON, D. M. (1979). Estimation for Mixtures of Distributions by Direct Maximization of the Likelihood Function. Comp. Prog. **11** 153-159.

ORCHARD, T. & WOODBURY, M. A. (1972). A Missing Information Principle: Theory and Applications. Proc. 6th Berkeley Symp. Math. Statist. **1** 697-715, University of California Press.

PAUL, S. R. (1985). Estimation of the Parameters of a Clumped Binomial Model via the EM algorithm. Amer. Statist. **39** 136-139.

RASCH, G. (1960). Probabilistic Models for Some Intelligence and Attainment Tests. Nielson and Lydiche, Copenhagen.

RAYNER, J. W. C. (1985). Maximum Likelihood of the Mean and the Variance from a Multivariate Normal Distribution. Amer. Statist. **39** 12-125.

RIDER, P. R. (1953). Truncated Poisson Distributions. J. Amer. Statist. Ass. **32** 143-147.

RUBIN, D. B. (1976). Inference and Missing Data. Biometrika **63** 581-592.

RUBIN, D. B. (1980). Using Empirical Bayes Techniques in the Law School Validity Studies (with comments). J. Amer. Statist. Ass. **75** 801-827.

SCHMEE, J. & HAHN, G. J. (1979). A Simplified Method for Regression Analysis with Censored Data. Technometrics. **21,** 417-432.

SAMANIEGO, F. J., & WHITAKER, L. R. (1984). On Estimating Population Characteristics from Record-Breaking Observations. **Technical Report 84-01**, University of California, Davies.

STEIN, C. (1956). Inadmissibility of the Usual Estimators for the Mean of a Multivariate Normal Distribution. Proc. Third Berkeley Symp. Math. Statist. Prob. **I** 197-206, University of California Press, Berkeley.

STRAWDERMAN W. E. & COHEN, A. (1971). Admissibility of Estimators of the Mean Vector of a Multivariate Normal Distribution with Quadratic Loss. Ann. Math. Statist. **42** 270-296.

STRAWDERMAN, W. E. (1970). On the Existence of a Proper Bayes Minimax Estimators of the Mean of a Multivariate Normal Distribution. Proc. Sixth Berkeley Symp. Math. Statist. Prob.

STRAWDERMAN, W. E. (1971). Proper Bayes Minimax Estimators of the Multivariate Normal Mean. Ann. Math. Stat. **42** 385-388.

STRAWDERMAN, W. E. (1973). Proper Bayes Minimax Estimators of the Multivariate Normal Mean Vector for the Case of Common Unknown Variances. Annals of Statistics **44** 1189-1194.

SUNDBERG, R. (1974). Maximum Likelihood Theory for Incomplete Data from an Exponential Family. Scan. J. Statist. **1** 49-58.

SUNDBERG, R. (1976). An Iterative Method for Solution of the Likelihood Equations for Incomplete Data from Exponential Families. Comm. Statist. Simula. Computa. B **5** 55-64.

TAN, W. Y. & CHANG, W. C. (1970). Comparison of the Method of Moments and Maximum Likelihood in Estimating Parameters of a Mixture of Two Normal Densities. J. Amer. Statist. Ass. **67** 702-709.

TSUI, K. W. (1979). Estimation of Poisson Means under Weighted Squared Error Loss. Canadian Journal of Statistics **7** 201-204.

TSUI, K. W. (1981). Simultaneous Estimation of Several Poisson Parameters. Ann. Inst. Statist. Math. **33** 215-224.

TSUI, K. W. & PRESS, S. J. (1982). Simultaneous Estimation of Several Poisson Parameters under k-Normalized Squared Error Loss. Ann. Statist., **10** 93-100.

WICHMANN, B. A. & HILL, I. D., (1982). An Efficient and Portable Pseudo-Random Number Generator. Appl. Statist. **31** 188-190.

WU, C. F. J. (1983). On the Convergence Properties of the EM Algorithm. Ann. Statist. **11** 95-103.

ZACKS, S. (1971). The Theory of Statistical Inference, Wiley, New York.

ZACKS, S. & ZIMMER W. J. (1972). Estimators of Severity Factors in a Multiplicative Poisson Model. J. Amer. Statist. Assoc. **67** 192-195.

Appendix A1

ESTIMATION OF THE MEAN AND THE STANDARD DEVIATION OF A NORMAL POPULATION USING CENSORED SAMPLES: AN ITERATIVE SOLUTION VIA THE EM ALGORITHM. DATA FROM GUPTA(1952). THE PROGRAM IS DESIGNED FOR USE ON AN IBM PC USING THE MICROSOFT COMPILER VERSION 4.1

```
program censor

real x(300),y(119),f(10),yd(10),yp(10),xd(300)

real limit,a1,a2,a3,a,exp,sigma1,sigma,mu1,mu ,xvar,sqrt

real bgfeta,smfeta,xmean,nstar,atan,trunc,t,eta

integer k,i,n ,count

open (1,file='censor.dat' ,status='old' )

open (2,file='censor.res', status='new' )

read (1,*) (yp(i),i=1,10)

write (*,*) 'now reading input data from data files: '

read (1,*) (f(i),i=1,10)

k=119

n=300

mu=1400

mu1=1500

sigma=1000

sigma1=1200

a1=0.4361836

a2=-.1201676

a3=.937298

trunc=1450

limit=0.001
```

```
do 1  i=1,k
 if (i.le. 2) then
 y(i)=yp(1)
 end if
 if ((i .ge. 3) .AND. (i .le. 4))  then
 y(i)=yp(2)
 end if
 if ((i .ge. 5) .AND. (i .le. 7))  then
 y(i)=yp(3)
 end if
 if ((i .ge. 8) .AND. (i .le. 13))  then
 y(i)=yp(4)
 end if
 if ((i .ge. 14) .AND. (i .le. 20)) then
 y(i)=yp(5)
 end if
 if ((i .ge. 21) .AND. (i .le. 32))  then
 y(i)=yp(6)
 end if
 if ((i .ge. 33) .AND. (i .le. 48)) then
 y(i)=yp(7)
 end if
 if ((i .ge. 49) .AND. (i .le. 68))  then
 y(i)=yp(8)
 end if
 if ((i .ge. 69) .AND. (i .le. 92)) then
```

```
      y(i)=yp(9)
      end if
      if ((i .ge. 93) .AND. (i .le. 119))  then
      y(i)=yp(10)
      end  if
      write (*,*) '    '
1     continue
      write (*,*) 'data entry completed'
      write (*,*) '  '
      write (*,*) 'computation of parameters begins here'
      write (*,*) 'please wait a moment'
      do 2  i=1,k
        x(i)=y(i)
2     continue
      count=0
10    eta=(trunc-mu)/sigma
      smfeta=sqrt(8.0*atan(1.0))
      smfeta=smfeta*exp((eta**2)/2)
      smfeta=1.0/smfeta
      t=(1+.33267*eta)
      t=1.0/t
      bgfeta=(a1*t+a2*t**2+a3*t**3)*smfeta
      a=smfeta/bgfeta
      nstar=k-(n-k)*a*eta
      if (nstar .le. 0) then
      nstar=10.0
```

```
      end if
      do 3 i=k+1,n
       x(i)=mu+sigma*a
3     continue
      xmean=0.0
      xvar=0.0
      mu=mu1
      sigma=sigma1
      do 4  i=1,n
       xd(i)=x(i)-x(1)
       xmean=xmean+xd(i)
4     continue
      xmean=xmean/n + x(1)
      mu1=xmean
      do 5  i=1,k
       xvar=xvar+((x(i)-mu)**2)
5     continue
      xvar=xvar/nstar
      sigma1=sqrt(xvar)
      write (*,*) mu,sigma
      write (2,*) mu,sigma
      if (abs(sigma-sigma1) .gt. limit) then
      count=count+1
      goto 10
      end if
      write (2,*) '         '
      write (2,*) '    '
      write (2,*) mu,sigma
      write (2,*) 'number of cycles: ',count
      end
```

Appendix A2

This Program estimates the Mean and Variance of a Mixture of two Univariate Normal Densities. One of the Means and the mixing proportion are assumed known.

```
//NORMAL JOB  ',,,T=1M',DARKO
/*ROUTE PRINT NEMP
// EXEC WATFIV
$JOB WATFIV    AWUKU,NOEXT
C
C   ML estimates of the parameters of mix of 2 normal; EM
algorithm.
C
C  Model:  p and  Mu1  are known; Mu2 =Mu1+delta
C          to estimate  delta and common variance
C
       REAL R1(200),P,LIKE,ALOG,CRITC,LIKL,DSIG ,UNIF(1)
       REAL FUNC1,FUNC2,SIGMA,MU(2),EXP,SQRT,NORM1(1),NORM2(1)
       REAL LIKE1,LIKE2 ,ABS,W(2,200),X ,EMU,ESIGMA,SIGMA1
       INTEGER  I ,J,NR
       DOUBLE PRECISION DSEED1,DSEED2
       EXTERNAL FUNC1,FUNC2
       DSEED1=123467.D0
       DSEED2=234259.D0
       D=1.0
       READ ,(MU(I),I=1,2),SIGMA
        P=7.0/8
```

```
      NR=1
      I=1
      SIGMA1=SQRT(SIGMA)
      WHILE ( I .LE. 200) DO
      DSEED1=DSEED1+23*I
      DSEED2=DSEED2+12*I
      CALL GGUBS(DSEED1,NR,UNIF)
      IF ( UNIF(1) .GT. P) THEN DO
       CALL GGNML (DSEED2,NR,NORM2)
       R1(I)=NORM2(1)*SIGMA1+MU(2)
       ELSE
       DSEED2=DSEED2+17.D0
       CALL GGNML (DSEED2,NR,NORM1)
       R1(I)=NORM1(1)*SIGMA1
        END IF
        I=I+1
        END WHILE
      PRINT 1
1     FORMAT('1'//20X,'Values of parameters are as follows')
      PRINT ,'      '
      PRINT , 'Value of MU(1) is: ', MU(1)
      PRINT ,'Value of MU(2) is :  ',MU(2)
      PRINT, 'Value of std dev is : ',SIGMA1
      PRINT , '  '
      PRINT, '      '
      PRINT , '     '
```

```
      PRINT,'    '
      READ ,EMU,ESIGMA
      PRINT ,'Initial estimate of DELTA  :',EMU
      PRINT,'Initial estimate  of std dev : ',ESIGMA
      PRINT 8
       I=1
       LIKE1=0.1
       LIKE2=0.2
       MU(2)=10.0
       SIGMA=ESIGMA
       CRITC=ABS((LIKE2-LIKE1)/LIKE1)
       WHILE (CRITC .GT.1.0E-6  .AND. I .LE. 100) DO
         CALL  ESTIMS (R1,MU,SIGMA,P,W)
         EXECUTE LIKEHD
         LIKE1=LIKE2
         LIKE2=LIKE
         CRITC=ABS((LIKE2-LIKE1)/LIKE1)
         CALL CHECK (R1,MU,SIGMA,P,W,DMU,DSIG)
         PRINT, 'I=',I,  'Mu2=' ,MU(2),'SIGMA=',SIGMA
         PRINT,'DLDMU2=',DMU,'DLDSIG',DSIG
         PRINT,'LOG LIKEHD =',LIKE
         PRINT 8
 8       FORMAT('-')
         I=I+1
      END WHILE
7     FORMAT('1')
```

```
      PRINT  6
      PRINT 7
      STOP
 6    FORMAT('-')
C
C
C
C
C
C
C
C
C REMOTE BLOCK LIKEHD
      REMOTE BLOCK LIKEHD
      J=1
      LIKE=0.0
      WHILE(J .LE.  200) DO
        X=R1(J)
        LIKL=P*FUNC1(X,MU,SIGMA)+(1-P)*FUNC2(X,MU,SIGMA)
       LIKE=LIKE+ALOG(LIKL)
        J=J+1
      END  WHILE
      END BLOCK
      END
C
C
C
      REAL  FUNCTION FUNC1(X,MU,SIGMA)
      REAL  X,SIGMA,MU(2),SQRT,NUM
```

```
      NUM=(X-MU(1))*(X-MU(1))/(2.0*SIGMA**2)
      FUNC1=EXP(-NUM)
      RETURN
      END
C
C
C
C
      REAL FUNCTION   FUNC2(X,MU,SIGMA)
      REAL  SIGMA,MU(2),EXP,X,NUM
      NUM=(X-MU(2))*(X-MU(2))/(2.0*SIGMA**2)
      FUNC2=EXP(-NUM)
      RETURN
      END
C
C
C
C
      SUBROUTINE ESTIMS (R1,MU,SIGMA,P,W)
      REAL   W(2,200),D2,D,P,DENOM ,SIGMA,R1(200),X
      REAL  FUNC1,FUNC2,MU(2),SUMSQ1,P11,P22,SUMSQ2
      INTEGER L, M
      L=1
      SUMSQ1=SUMSQ2=P11=P22=D1=D2=0.0
      WHILE (L .LE. 2) DO
        M=1
```

```
          WHILE (M .LE. 200) DO
          X=R1(M)
          DENOM=P*FUNC1(X,MU,SIGMA) +(1-P)*FUNC2(X,MU,SIGMA)
          IF(L .EQ. 1) THEN DO
          W(L,M)=P*FUNC1(X,MU,SIGMA)/DENOM
           SUMSQ1=SUMSQ1+W(L,M)*(R1(M)-MU(L))*(R1(M)-MU(L))
           P11=P11+W(L,M)
          END IF
          IF(L .EQ. 2) THEN DO
          W(L,M)=(1-P)*FUNC2(X,MU,SIGMA)/DENOM
          P22=P22+W(L,M)
          SUMSQ2=SUMSQ2+W(L,M)*(R1(M)-MU(L))*(R1(M)-MU(L))
          D2=D2+W(L,M)*X
           END IF
          M =M +1
          END WHILE
          L=L+1
      END WHILE
      SIGMA=(SUMSQ1+SUMSQ2)/(P11+P22)
      SIGMA=SQRT(SIGMA)
      MU(2)=D2/P22
      RETURN
      END
C
C
      SUBROUTINE CHECK  (R1,MU,SIGMA,P,W,DMU,DSIG)
```

```
      REAL  DMU,R1(200),MU(2),SIGMA,W(2,200),P,DSIG,TERM
      INTEGER  I,J
      DMU=DSIG=0.0
      I=1
      WHILE (I .LE.2) DO
       J=1
       WHILE (J .LE. 200) DO
         IF (I.EQ. 2) THEN DO
         TERM=(R1(J)-MU(I))/(SIGMA**2)
         DMU=DMU+TERM*W(I,J)
         END IF
         TERM=(R1(J)-MU(I))**2
         TERM=TERM-(SIGMA**2)
         TERM=TERM/(SIGMA**3)
         TERM=TERM*W(I,J)
         DSIG=DSIG+TERM
         J=J+1
         END WHILE
         I=I+1
      END WHILE
      RETURN
      END
$ENTRY
 0 3.0 2.0
10.0 50.0  3.0
20.0
```

Appendix B1

This program computes the Risks of the MLE and Generalized Bayes Estimator of a Vector of Poisson Means under a Multiplicative Model for a 3x3 Table.

```
//EXPMLBA JOB ',,L=10,T=20,I=15',AWUKU
/*ROUTE PRINT NEMP
// EXEC WATFIV,REGION=650K
//FT08F001    DD  DSN='AWUKU.RIMLBA.DATA',DISP=OLD
//SYSIN       DD   *
$JOB WATFIV    AWUKU,NOEXT
C
C  1.   Compare the risks of the ML estimator and the Bayes
C       estimator of a vector of Poisson Means under a
C       multiplicative model inan IxJ (complete) Table.
C
C
C  2.   For this run ROWS=I=3, COLUMNS=J=3 and TAU is from 0
C       to 70.
C
C
        IMPLICIT INTEGER*4 (I)
        REAL TEMP1,TEMP2,PARA,P,MLE,BAYES,RATIO,FNG
        REAL TEMP,MAX,EXP,MEAN ,LIMIT
        INTEGER N1,N2,FN1,FN2,IROW,ICOL
        MAX=70.0
```

```
      READ ,IROW,ICOL,PARA
      LIMIT=0.000001
      N1=IROW-1
      N2=ICOL-1
11    P=1.0/EXP(PARA)
      MEAN=0.0
      TEMP1=10.0
      TEMP2=1000.0
      IL=-1
10    IL=IL+1
      FNH=IL*(N1+N2)+N1*N2
      FNL=(IL+N1)*(IL+N2)
      FNG=FNH/FNL
      IF (IL .EQ. 0 ) THEN
        P=1.0/EXP(PARA)
      ELSE
        P=P*PARA/IL
      ENDIF
      TEMP=FNH*(2-FNG)
      TEMP=TEMP-2*IL*FNG*PARA
      TEMP=TEMP*P
      MEAN=MEAN+TEMP
      TEMP2=TEMP1
      TEMP1=MEAN
      IF (ABS((TEMP1-TEMP2)/TEMP1) .GT. LIMIT ) THEN
        GOTO 10
```

```
      ENDIF
      IF ( IL .LT. 30 ) THEN
       GOTO 10
      ENDIF
      MLE=N1+N2+1+N1*N2/PARA
      BAYES=MLE-TEMP1/PARA
      RATIO=(MLE-BAYES)/MLE*100.0
      PRINT 3,PARA,MLE,BAYES,RATIO,IL
      WRITE (8,2) PARA,MLE,BAYES,RATIO
3     FORMAT  (F10.2,2F9.5,F10.5,I5)
      IF (PARA .LE. MAX ) THEN
        PARA=PARA+2.5
        GOTO  11
      END IF
2     FORMAT  (F10.2,3X,2F9.3,3X,F10.3)
      STOP
      END
$ENTRY
   3,3 1.0
/*
```

Appendix B2

THIS SAS PROGRAM PLOTS THE RISKS OF THE ESTIMATORS OF THE MULTIPLICATIVE POISSON MEANS IN THE COMPLETE-DATA PROBLEM

```
//S1 EXEC SAS,REGION=3260K,OPTIONS=NONOTES
//DATA1   DD   DSN='AWUKU.RIMLBA.DATA',DISP=OLD
//SYSIN DD  *
*                                                              *;
*                                                              *;
*  The following SAS program plots the risks using data from  *;
*  the FORTRAN program in Appendix A2.                         *;
*                                                              *;
*                                                              *;
GOPTIONS DEVICE=XEROX colors=(bl);
DATA RISKS;
INFILE  DATA1 ;
INPUT  LAMDA MLRISK RISBAY  RELGAIN ;
proc format ;value typfmt   1='MLE'    2= 'BAYES';
data one ;set risks;
risk=mlrisk; legends=1;output;risk=risbay;legends=2;OUTPUT ;
keep legends lamda risk;
PROC GPLOT DATA=ONE;
legend1 label=(f=complex h=2) value=(f=complex h=1)
 shape=line(4) across=2;
PLOT  RISK*LAMDA=LEGENDS  / haxis=axis2 vaxis=axis1
legend=legend1;format legends typfmt.;
```

```
23. axis1 label=(h=1 f=complex ' Risk') ;

24. axis2 label=(h=1 f=complex 'Tau') order=(0 to 70 by 10) ;

25. SYMBOL1 I=SM20 L=1 V=:;

26. SYMBOL2 I=SM20 L=2 V=+;

27. title1 h=1 f=complex 'A Comparison of the Risks of the MLE and
the';

28. title2 h=1  f=complex 'EMGB Estimator of a Vector of Poisson
Means';

29. title3 h=1 f=complex ' under a Multiplicative Model for a
Balanced';

30. title4 h=1 f=complex '(Complete) 3x3 Table.';

31. PROC GPLOT DATA=RISKS;

32. PLOT RELGAIN*LAMDA /HAXIS=axis2 vaxis=axis1;

33. axis1 label=(h=1 f=complex ' Relative Risk') ;

34. axis2 label=(h=1 f=complex 'Tau') order=(0 to 70 by 10) ;

35. SYMBOL  I=SM20  L=1 V=:;

36. title1 h=1 f=complex 'The Gain in Risk using the EMGB
Estimator';

37. title2 H=1 F=complex 'rather than the MLE of a Vector of
Poisson';

38. title3 h=1 f=complex 'Means under a Multiplicative Model for a
';

39. title4 h=1 f=complex 'Balanced (Complete) 3x3 Table.
';

40. //
```

Appendix C1

THIS PROGRAM COMPUTES THE MLE OF AN EXPONENTIAL MEAN USING RECORD-BREAKING OBSERVATIONS. IN THIS RUN, 50 OBSERVATIONS WERE GENERATED FROM THE EXPONENTIAL DISTRIBUTION WITH MEAN=1.0

```
program censor

open (1,file= 'expmean.dat',status=old)

real  expo(500,50),r(50) ,mean,z(50),x(50),abs

real the ,sum,exp,limit,ht,zs,estim(500),max

integer dseed

logical go

integer i,nr ,m,k(50),s,j,count

i=1

nr=50

dseed=23457

mean=1.0

while (i .le. 500 )  do

   call  expos (dseed,mean,nr)

   j=1

   while (j .le.  nr) do

   expo(i,j)=r(j)

   j=j+1

   end while

   dseed=dseed+3*i

  i=i+1

  end while

  i=1

  while (i .le. 50) do
```

```
 s=m=1
 count=1
j=2
max=expo(i,1)
z(1)=max
while (j  .le. 50) do
 if ( j .eq. 50) then
  cuunt=count+1
  z(count)=max
  k(count)=s+1
 end if
 if (expo(i,j)  .gt. max ) then
   m=m+1
   z(m)=expo(i,j)
   k(m-1)=s
   max=expo(i,j)
    s=1
    count=count+1
else
s=s+1
 end if
 j=j+1
 end while
 go=.true.
 the=10.0
 while (go)  do
```

```
        sum=0.0
        m=1
        while ( m .le. count) do
         ht=exp(z(m)/the)
         ht=1.0/ht
         zs=1.0+z(m)/the
         zs=zs*ht
         zs=(1.0-zs)/(1.0-ht)
        x(m)=zs*the
         sum=sum+z(m)+(k(m)-1.0)*x(m)
         m=m+1
         end  while
         the1=the
         the=sum/nr
         if  (abs(the1-the)  .le.  1.0e-3)  then
             go=.false.
         end if
         end while
         estim(i)=the
         i=i+1
      end while
      write  (1,10)  (estim(i),i=1,50)
      write  (*,10)  (estim(i),i=1,50)
10    format (4x,  5f11.5)
      stop
c
```

```
c       The following subroutine generates exponential variates
c       using the algorithm of Wichmann and Hill (1982).
c

        subroutine expos (iran, pz, ia)

        implicit integer*2  (i)

        intrinsic amod,mod,float,exp

        real ran,ui,p,t,q,pz,b

        dimension ilits(1000)

        ix=101+iran

        iy=1001+iran

        iz=3004+iran

        do   5   il=1,ia

        il=0

        q=pz

        p=1.0/exp(q)

        t=1.0

1       ix=171*mod (ix,177)-2*(ix/177)

        iy=172*mod (iy,176)-35*(iy/176)

        iz=170*mod(iz,178) -63*(iz/178)

        if (ix .lt. 0 ) then

          ix=ix+30269

        end if

        if (iy .lt. 0) then

          iy=iy+30307

        end if

        if (iz .lt. 0) then

         iz=iz+30323
```

```
      end if
      ran=float(ix)/30269.0+float(iy)/30307.0+float(iz)/30323.0
      ui= amod(ran,1.0)
      t=t*ui
      if (t .le. p)  then
        goto 2
      end if
      il=il+1
      goto 1
2     ilits(i1)=il
5     continue
      return
      end

         end
```

COMBINATIONS OF PARAMETERS WHICH WERE USED IN SIMULATION EXPERIMENTS DESCRIBED IN CHAPTER VI

α_1	α_2	α_3	β_1	β_2	β_3
0.1	0.1	0.8	0.1	0.1	0.8
0.1	0.1	0.8	0.2	0.2	0.6
0.1	0.1	0.8	0.3	0.3	0.4
0.1	0.1	0.8	0.4	0.4	0.2
0.1	0.1	0.8	1/3	1/3	1/3
0.2	0.2	0.6	0.2	0.2	0.6
0.2	0.2	0.6	0.3	0.3	0.4
0.2	0.2	0.6	0.4	0.4	0.2
0.2	0.2	0.6	1/3	1/3	1/3
0.3	0.3	0.4	0.3	0.3	0.4
0.3	0.3	0.4	0.4	0.4	0.2
0.3	0.3	0.4	1/3	1/3	1/3
0.4	0.4	0.2	0.4	0.4	0.2
0.4	0.4	0.2	1/3	1/3	1/3
1/3	1/3	1/3	1/3	1/3	1/3

Appendix C3

SIMULATION PROGRAM TO COMPARE THE RISKS OF MLE AND EMGB ESTIMATOR OF A VECTOR OF POISSON MEANS UNDER A MULTIPLICATIVE MODEL WHEN OBSERVATIONS ARE ARRANGED IN AN INCOMPLETE 3X3 TABLE ACCORDING TO A BALANCED INCOMPLETE BLOCK DESIGN.

```
RISKS JOB ',,L=20,T=58M',AWUKU
/*ROUTE PRINT NEMP
//STEP1 EXEC WATFIV,REGION=650K
//FT01F001       DD    DSN='AWUKU.DUMMY1.DATA',DISP=MOD
//FT08F001       DD    DSN='AWUKU.DUMMY2.DATA',DISP=MOD
//SYSIN       DD    *
$JOB WATFIV    AWUKU,NOEXT
C
C
C  1.    Generate a 3x3 table of independent Poisson observations.
C        The observation in cell (1,j) comes from a Poisson population
C        with mean m(i,j)=a(i)*b(j)*lamda.
C
C  2.    Make an incomplete table by deleting some cell according
C        to the incidence matrix IM1 or IM2
C
C  3.    For each incomplete table, compute
C
```

```
C         (a)  EMGB estimates
C
C         (b)  ML estimates
C
C   4.   Repeat steps 1-3 2000 times
C
C   5.   Compute the average losses of the MLE and EMGB estimates
C
C   6.   Calculate the 'gain' in risk when the EMGB estimator is used
C        instead of the MLE:
C
C         gain =(Risk of MLE-Risk of EMGB)/Risk of MLE
C
C
C
C
      REAL RLAM,P(3),L(3),YROW(3),YCOL(3),LAMDA,BETBAY(3)
      REAL  LEAST(3),PEAST(3),TAUMLE,TAUBAY,RATMLE,RATBAY
      REAL LOBAY,LOMLE ,SUPMLE,SUPBAY,DIFTOT,MLETOT,BAYTOT
      REAL  ALFMLE(3),ALFBAY(3),BETMLE(3)
      INTEGER NR,K(1),IER,I,J,POIS(3,3),IZ
      INTEGER DUMMY(3,3),G,Q,COUNT
      COMMON /ITER/ P,L,IZ  /POIST/POIS
      COMMON /ESTIM/ DUMMY,YROW,YCOL
```

```
      COMMON  /INIT/PEAST,LEAST
      COMMON /FINAL/ TAUMLE,TAUBAY
      DOUBLE PRECISION  DSEED
C
C      ***** MAIN PROGRAM  *****
C
      EXECUTE PARAS
      COUNT=1
      WHILE (COUNT .LE. 2) DO
      EXECUTE  MATRIX
      LAMDA=1.0
      WHILE (LAMDA  .LE. 551.0) DO
       IZ=2000
       DIFTOT=MLETOT=BAYTOT=0.0
       Q= 1
       WHILE  (Q .LE.  IZ)  DO
5      EXECUTE  POISS
       EXECUTE  MARGIN
       CALL BAYES (ALFBAY,BETBAY)
       CALL MLES (ALFMLE,BETMLE)
       EXECUTE LOSSES
       MLETOT=MLETOT+SUPMLE
       BAYTOT=BAYTOT+SUPBAY
       DIFTOT=DIFTOT+(SUPMLE-SUPBAY)
       Q=Q+1
       END WHILE
```

```
      MLETOT=MLETOT/(Q-1)
      BAYTOT=BAYTOT/(Q-1)
      DIFTOT=DIFTOT/(Q-1)
      IF (COUNT  .EQ. 1) THEN
       WRITE (1,2) LAMDA,MLETOT,BAYTOT,DIFTOT
      END IF
      IF (COUNT .EQ. 2) THEN
       WRITE (8,2) LAMDA,MLETOT,BAYTOT,DIFTOT
      END IF
      LAMDA=LAMDA+25.0
     END WHILE
     COUNT=COUNT+1
     END WHILE
2    FORMAT (F9.2,2F9.3,F9.3)
     STOP
C
C    ********************************
C
C
C
C    Create a 3X3 complete table  of Poisson;
C    cell (i,j) is from a Poisson with mean p(i)*l(j)*lamda.
C
C
C
C   Read  parameters of the Poisson vars.
```

```
C
      REMOTE BLOCK PARAS
      READ, (P(I),I=1,3)
      READ, (L(I),I=1,3)
      READ, (PEAST(I),I=1,3), (LEAST(I),I=1,3)
      ISS=1
      WHILE (ISS .LE. 3) DO
      L(ISS)=1.0/3.0
      P(ISS)=1.0/3.0
      ISS=ISS+1
      END WHILE
      END BLOCK
C
C
C     Read  'Incidence matrix' of Incomplete table
C
      REMOTE BLOCK  MATRIX
      READ,(( DUMMY(I,J),J=1,3),I=1,3)
      END  BLOCK
C
C
C
C
C
C  Obtain a table of Poisson observations
C
```

```
      REMOTE BLOCK  POISS
C
      DSEED=291735.D0+13.D0*Q
      NR=1
      I=1
      WHILE (I .LE. 3) DO
       J=1
       WHILE ( J .LE. 3 ) DO
         RLAM= P(I)*L(J)*LAMDA
         CALL GGPON  (RLAM,DSEED,NR,K,IER)
         POIS(I,J)=K(1)
         J=J+1
        END WHILE
        I=I+1
      END WHILE
      END BLOCK
      END
C
C      Calculate marginal totals of incomplete table
C
      REMOTE BLOCK MARGIN
      I=1
      WHILE (I .LE. 3 ) DO
       YCOL(I)=YROW(I)=0.0
       I=I+1
      END WHILE
```

```
      I=1
      WHILE (I .LE. 3) DO
       J=1
       WHILE (J .LE. 3 ) DO
         IF (DUMMY(I,J) .GT. 0 ) THEN DO
           YROW(I)=YROW(I)+POIS(I,J)
           YCOL(J)=YCOL(J)+POIS(I,J)
         END IF
         J=J+1
       END WHILE
       I=I+1
      END WHILE
      END BLOCK
C
C
C     Calculate the loss associated with each estimate
C
       REMOTE  BLOCK LOSSES
        IT=1
        SUPMLE=SUPBAY=0.0
        WHILE (IT  .LE.  3) DO
         IST=1
         WHILE (IST .LE. 3) DO
          AMLE=ALFMLE(IT)*BETMLE(IST)*TAUMLE
          ABAY=ALFBAY(IT)*BETBAY(IST)*TAUBAY
          PARA =P(IT)*L(IST)*LAMDA
```

```
                LOMLE=(1/PARA)*((AMLE-PARA)**2)
                LOBAY=(1.0/PARA)*((ABAY-PARA)**2)
                SUPMLE=SUPMLE +LOMLE
                SUPBAY=SUPBAY +LOBAY
                IST=IST+1
                END WHILE
                IT=IT+1
              END WHILE
              END BLOCK
C
C
C  Calculate the EMGB estimate using data from incomplete
table
C
        SUBROUTINE BAYES (ALFBA,BETBA)
        REAL YROW(3),YCOL(3),ABS,P(3),L(3),LAMEST,X(3,3)
        REAL LO,XROW(3),XCOL(3),SUM,PEST(3),LEST(3)
        REAL LIMIT,LLIMIT,PLIMIT,PI,ALFBA(3),BETBA(3)
        REAL PEAST(3),LEAST(3)
        INTEGER I,J,DUMMY(3,3),F,POIS(3,3)
        LOGICAL AZERO,BZERO
        COMMON /ITER/ P,L,IZ    /POIST/ POIS
        COMMON /ESTIM/ DUMMY,YROW,YCOL
        COMMON /INIT/ PEAST,LEAST
        COMMON /FINAL/ TAUMLE,TAUBAY
        I=1
```

```
WHILE (I .LE. 3)  DO
 LEST(I)=LEAST(I)
 PEST(I)=PEAST(I)
 I=I+1
END WHILE
LO=PI=1.0
LAMEST=10.0
I=1
LIMIT=1.0E-4
LLIMIT=ABS(LEST(1)-LO)
PLIMIT=ABS(PEST(1)-PI)
WHILE (LLIMIT .GT. LIMIT .AND. PLIMIT .GT. LIMIT) DO
I=1
WHILE (I .LE. 3 ) DO
 XROW(I)=YROW(I)
 XCOL(I)=YCOL(I)
 I=I+1
END WHILE
I=1
SUM=0.0
WHILE (I .LE. 3 ) DO
 J=1
 WHILE (J .LE. 3) DO
   AZERO=PEST(I) .GT. 1.0E-7
   BZERO=LEST(J) .GT. 1.0E-7
  IF (  DUMMY(I,J) .LE. 0 ) THEN DO
```

```
    IF (AZERO .AND. BZERO  ) THEN DO
     X(I,J)=PEST(I)*LEST(J)*LAMEST
    ELSE DO
     X(I,J)=0.0
    END IF
  ELSE DO
   X(I,J)=0.0
  END IF
  XROW(I)=X(I,J)+XROW(I)
  XCOL(J)=X(I,J)+XCOL(J)
  J=J+1
 END WHILE
 SUM=SUM+XROW(I)
 I=I+1
 END WHILE
  I=1
  LO=LEST(1)
  PI=PEST(1)
  WHILE (I .LE. 3 ) DO
   IF (SUM .LE. 1.0E-7) THEN DO
    PEST(I)=1.0/3
    LEST(I)=1.0/3
   ELSE DO
    PEST(I)=XROW(I)/SUM
    LEST(I)=XCOL(I)/SUM
   END IF
```

```
            I=I+1
          END WHILE
            IF (SUM .LE. 1.0E-7) THEN DO
              LAMEST=0.0
            ELSE DO
          LAMEST=1.0-2./(SUM+2.0)
          LAMEST=SUM*LAMEST*LAMEST
          END IF
          LLIMIT=ABS( LEST(1)-LO)
          PLIMIT=ABS(PEST(1)-PI)
        END WHILE
        I=1
        WHILE (I .LE. 3)  DO
         ALFBA(I)=PEST(I)
         BETBA(I)=LEST(I)
         I=I+1
        END WHILE
        TAUBAY=LAMEST
        RETURN
        END
C
C
C  Calculate the MLE using data from incomplete table
C
      SUBROUTINE MLES (ALFML,BETML)
      REAL YROW(3),YCOL(3),ABS,P(3),L(3),LAMEST,X(3,3)
```

```
      REAL LO,XROW(3),XCOL(3),SUM,PEST(3),LEST(3)
      REAL LIMIT,LLIMIT,PLIMIT,PI,ALFML(3),BETML(3)
      REAL PEAST(3),LEAST(3)
      INTEGER I,J,DUMMY(3,3),F,Q,POIS(3,3)
      LOGICAL AZERO, BZERO
      COMMON /ITER/ P,L,IZ   /POIST/ POIS
      COMMON /ESTIM/ DUMMY,YROW,YCOL
      COMMON /INIT/ PEAST,LEAST
      COMMON /FINAL/ TAUMLE,TAUBAY
      LO=PI=1.0
      LAMEST=10.0
      I=1
      WHILE (I .LE. 3) DO
       LEST(I)=LEAST(I)
       PEST(I)=PEAST(I)
       I=I+1
      END WHILE
      I=1
      LIMIT=1.0E-4
      LLIMIT=ABS(LEST(1)-LO)
      PLIMIT=ABS(PEST(1)-PI)
      WHILE (LLIMIT .GT. LIMIT .AND. PLIMIT .GT. LIMIT) DO
      I=1
      WHILE (I .LE. 3 ) DO
       XROW(I)=YROW(I)
       XCOL(I)=YCOL(I)
```

```
 I=I+1
END WHILE
I=1
SUM=0.0
WHILE (I .LE. 3 ) DO
 J=1
 WHILE (J .LE. 3) DO
   AZERO=PEST(I) .GT. 1.0E-6
   BZERO=LEST(J) .GT. 1.0E-6
  IF (  DUMMY(I,J) .LE. 0 ) THEN DO
    IF (AZERO .AND. BZERO  ) THEN DO
     X(I,J)=PEST(I)*LEST(J)*LAMEST
    ELSE DO
     X(I,J)=0.0
    END IF
  ELSE DO
    X(I,J)=0.0
  END IF
  XROW(I)=X(I,J)+XROW(I)
  XCOL(J)=X(I,J)+XCOL(J)
  J=J+1
 END WHILE
 SUM=SUM+XROW(I)
 I=I+1
END WHILE
I=1
```

```
        LO=LEST(1)
        PI=PEST(1)
        WHILE (I .LE. 3 ) DO
         IF (SUM .LE. 1.0E-7) THEN DO
          PEST(I)=1.0/3
          LEST(I)=1.0/3
         ELSE DO
          PEST(I)=XROW(I)/SUM
          LEST(I)=XCOL(I)/SUM
         END IF
          I=I+1
        END WHILE
        IF (SUM .LE.  1.0E-7) THEN  DO
         LAMEST =0.0
        ELSE DO
         LAMEST=SUM
        END IF
        LLIMIT=ABS(LEST(1)-LO)
        PLIMIT=ABS(PEST(1)-PI)
    END WHILE
     TAUMLE=LAMEST
    I=1
    WHILE (I .LE. 3)  DO
     ALFML(I)=PEST(I)
     BETML(I)=LEST(I)
     I=I+1
```

```
            END WHILE
            RETURN
            END
$ENTRY
0.4  0.4   0.2
0.4  0.4   0.2
0.1  0.3   0.6   0.2   0.3   0.5
   0    1    1
   1    0    1
   1    1    0
   1    0    1
   0    1    1
   1    1    0
/*
```

Appendix C4

A SAS Program to Plot the Frequency Distribution of 2000 Losses of MLE and EMGB estimates of a Vector of Poisson Means in the case of Incomplete Table.

```
//COMPRISK JOB ',,L=10,T=80',AWUKU
/*ROUTE PRINT NEMP
//S1 EXEC SAS,REGION=3260K,OPTIONS=NONOTES
//DATA1    DD    DSN='AWUKU.RISK.DATA',DISP=OLD
//SYSIN DD  *
OPTIONS NODATE;goptions device=xerox;
pattern1  c=bl v=e ; pattern2  v=x1  c=bl ;
proc format ;value typfmt   1='MLE'    2= 'EMGB';
DATA LOSSES;
INFILE  DATA1 ;
INPUT MLELOSS BAYLOSS  DIFF;
data one ;set losses;
loss=mleloss;types=1;output;loss=bayloss;types=2;output;
keep loss types;
proc sort data=one ;by types;
PROC GCHART DATA=one;
VBAR LOSS/TYPE=PCT MIDPOINTS=0 TO 10 BY 5 20 TO 300 BY 50
SPACE=0 group=types patternid=group;
title1 f=complex h=1
 'Frequency Distribution of the Losses of ML Estimates of ';
title2 f=complex h=1
```

```
'Poisson Means under a Multiplicative Model in an Incomplete';
 title3 f=complex h=1  '3x3 Table.
Loss Function is the Sum of the Losses for the    ';
 title4 f=complex h=1
 'Estimates of the Cell Means. 2000 Tables used.  ';
 title5 f=complex h=1
 'ALPHA=(0.4, 0.4, 0.2);   BETA=(0.4, 0.4, 0.2)';
PROC GCHART DATA=LOSSES; VBAR DIFF /TYPE=PCT MIDPOINTS=-10 TO
10 BY 10 50 TO 200 BY 50 SPACE=0 ;
title1 f=complex h=1
 'Frequency Distribution of the Difference between the Risks';
title2 f=complex h=1
 'of the ML estimates and the EMGB Estimators of Poisson    ';
title3 f=complex h=1
 'Means under a Multiplicative Model in a 3x3 Incomplete    ';
title4 f=complex h=1
 'Table. Data are based on 2000 Random Tables.              ';
title5 '                     ';
title6 ' DIFF=MLE-BAYES                ';
PROC MEANS data=losses MAXDEC=3;
title1 'Comparing the Distributions of the Losses of ML     ';
title2 'and EMGB Estimates of a Vector of Poisson Means under';
title3 'a Multiplicative Model in a 3x3 Incomplete Table.   ';
title4 'ALPHA=(0.4, 0.4, 0.2); BETA=(0.4, 0.4, 0.2);
LAMDA=50.';
Title5 '                                                  ';
```

```
title6 ' Summary Statistics for 2000 Random Tables.          ';
data _null_; put _page_;
data _NULL_; set losses;
if _N_ =1 then put 'Losses of ML estimates which exceed 400 ';
if Mleloss >400 then
put _N_  mleloss ;
data _NULL_ ;set losses;
if _N_ =1 then put
'Losses of the EMGB estimates which exceed 400 ';put '   ' ;
if bayloss > 400 then
put _N_  Bayloss  ;
data _NULL_ ;set losses;
if _N_ =1 then
put  'Observations for which DIFF exceed 400.'; put '   ' ;
if diff > 400 then
put _N_  mleloss bayloss diff ;
```

Appendix D

Table D1

Comparing the Distribution of Losses of the MLE and the EMGB Estimates of a Matrix of Poisson Means under a Multiplicative Model in a 3x3 Table. Summary Statistics for 2000 Tables.

ALPHA=(0.4, 0.4, 0.2); BETA=(0.4,0.4,0.2); TAU=2.0

VARIABLE	MLE	EMGB	DIFF
N	2000	2000	2000
MEAN	125.807	1.279	124.528
STANDARD DEV	490.008	1.084	489.778
MINIMUM VALUE	1.800	0.400	1.000
MAXIMUM VALUE	7379.500	20.500	7364.500
STD ERROR OF MEAN	10.957	0.024	10.952
SUM	251612.700	2557.600	249055.600
VARIANCE	240108.700	1.175	239882.354
C.V.	389.494	84.752	393.308
LOWER 95% BOUND	104.330	1.232	103.062
UPPER 95% BOUND	147.282	1.326	145.994

Table D2

Comparing the Distribution of Losses of the MLE and the EMGB Estimates of a Matrix of Poisson Means under a Multiplicative Model in a 3x3 Table. Summary Statistics for 2000 Tables.

ALPHA=(0.4, 0.4, 0.2); BETA=(0.4,0.4,0.2); TAU=20.0

VARIABLE	MLE	EMGB	DIFF
N	2000	2000	2000
MEAN	88.419	6.768	81.651
STANDARD DEV	354.143	4.869	352.720
MINIMUM VALUE	0.200	0.200	-7.400
MAXIMUM VALUE	5659.600	65.300	5609.200
STD ERROR OF MEAN	7.919	0.109	7.887
SUM	176837.400	13536.000	163301.600
VARIANCE	125417.388	23.702	124411.111
C.V.	400.530	71.934	431.985
LOWER 95% BOUND	72.898	6.554	66.193
UPPER 95% BOUND	103.940	6.982	97.109

Table D3

Comparing the Distribution of Losses of the MLE and the EMGB Estimates of a Matrix of Poisson Means under a Multiplicative Model in a 3x3 Table. Summary Statistics for 2000 Tables.

ALPHA=(0.4, 0.4, 0.2); BETA=(0.4,0.4,0.2); TAU=200.0

VARIABLE	MLE	EMGB	DIFF
N	2000	2000	2000
MEAN	11.409	10.136	1.272
STANDARD DEV	11.946	9.262	4.100
MINIMUM VALUE	0.200	0.300	-73.800
MAXIMUM VALUE	196.600	138.300	58.300
STD ERROR OF MEAN	0.267	0.207	0.092
SUM	22817.400	20271.900	2544.900
VARIANCE	142.706	85.792	16.807
C.V.	104.709	91.381	322.184
LOWER 95% BOUND	10.887	9.730	1.092
UPPER 95% BOUND	11.931	10.542	1.452

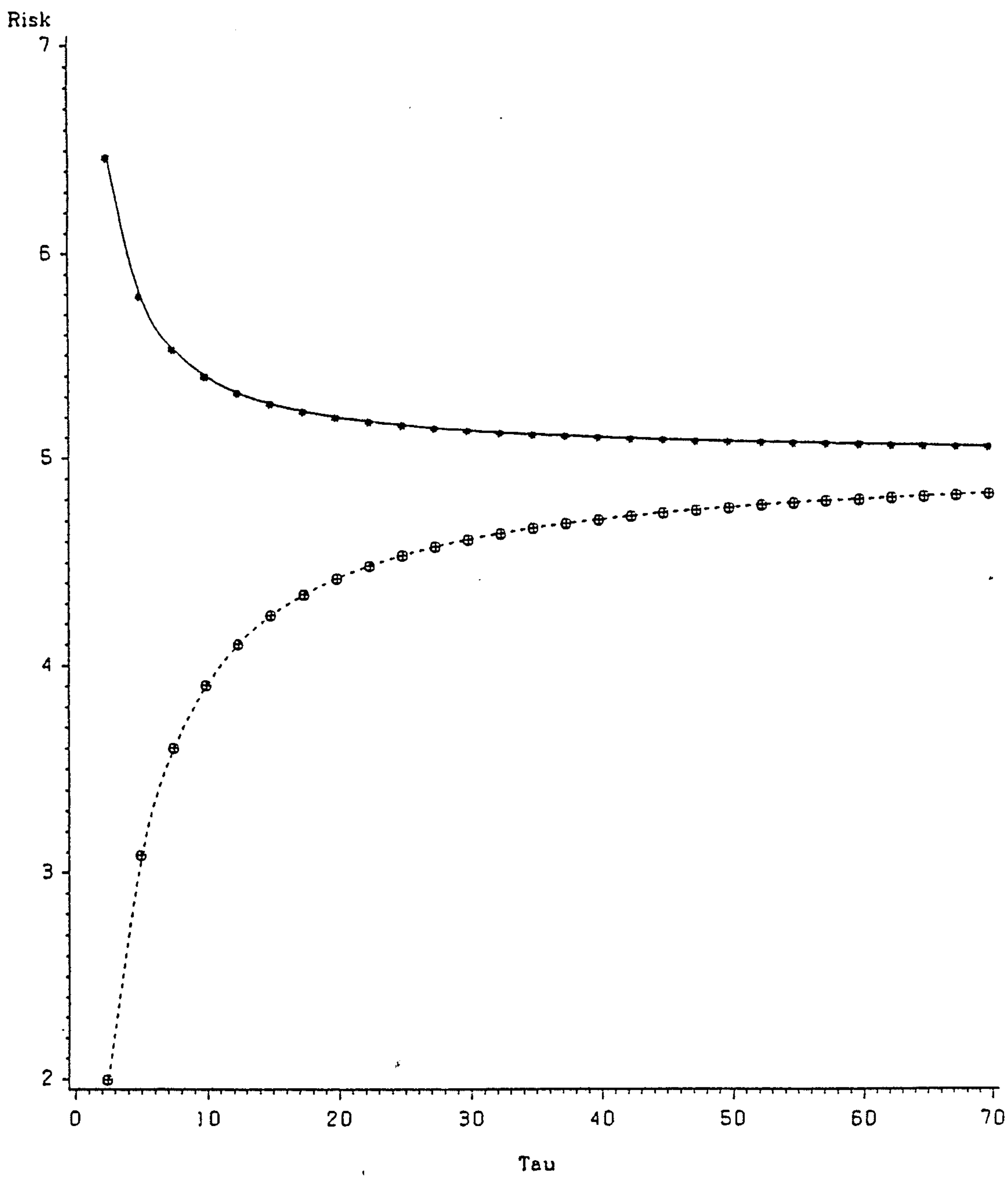
Table D3

Comparing the Distribution of Losses of the MLE and the EMGB Estimates of a Matrix of Poisson Means under a Multiplicative Model in a 3x3 Table. Summary Statistics for 2000 Tables.

ALPHA=(0.4, 0.4, 0.2); BETA=(0.4,0.4,0.2); TAU=500.0

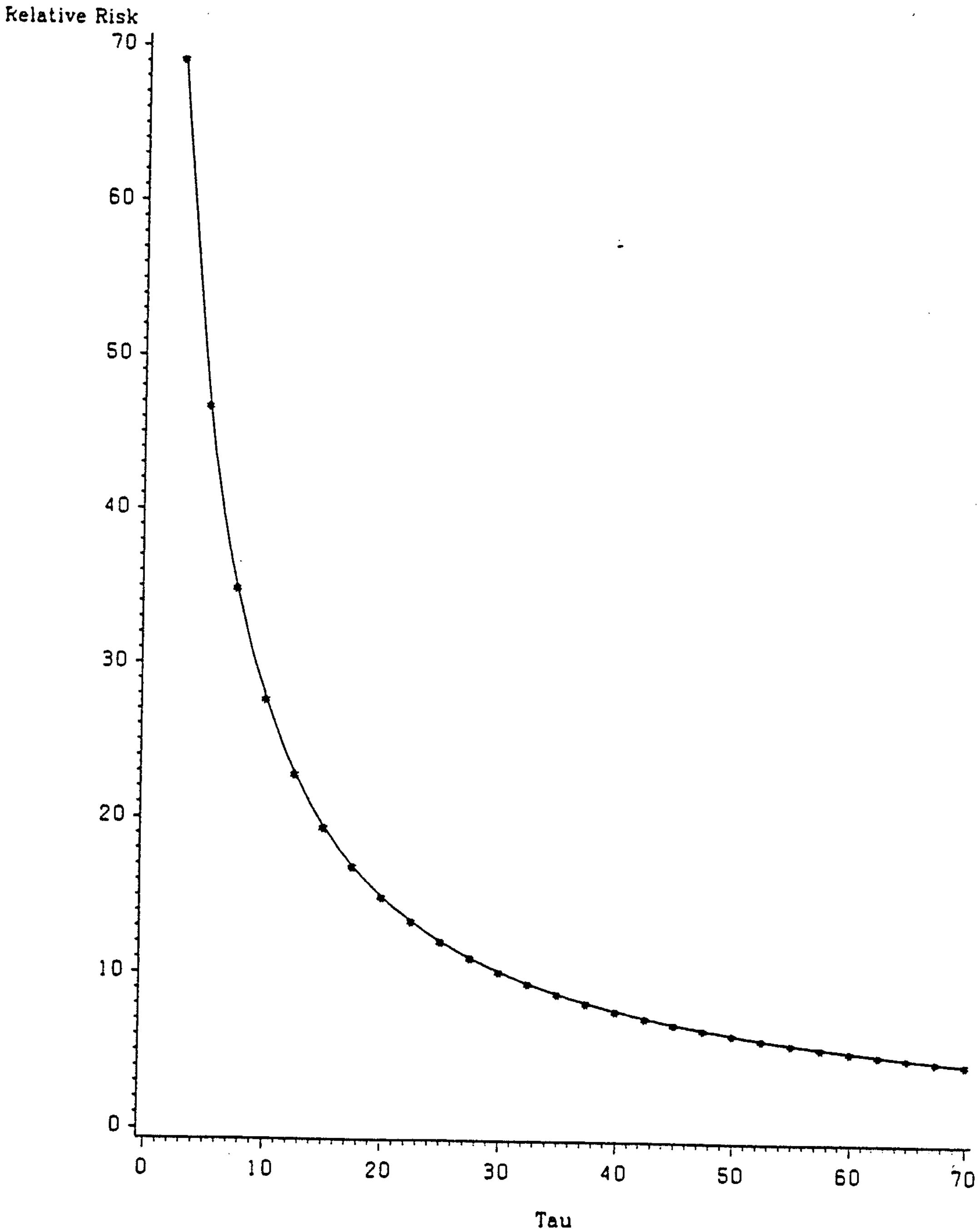
VARIABLE	MLE	EMGB	DIFF
N	2000	2000	2000
MEAN	10.892	10.518	0.374
STANDARD DEV	9.769	9.317	1.318
MINIMUM VALUE	0.100	0.200	-12.500
MAXIMUM VALUE	96.100	99.500	10.900
STD ERROR OF MEAN	0.218	0.208	0.029
SUM	21783.500	21035.400	748.400
VARIANCE	95.442	86.809	1.736
C.V.	89.696	88.585	352.151
LOWER 95% BOUND	11.319	10.925	0.431
UPPER 95% BOUND	10.465	10.518	0.317

A COMPARISON OF THE RISKS OF THE MLE AND THE GENERALIZED BAYES ESTIMATOR OF A VECTOR OF POISSON MEANS UNDER A A MULTIPLICATIVE MODEL FOR A COMPLETE 3X3 TABLE.



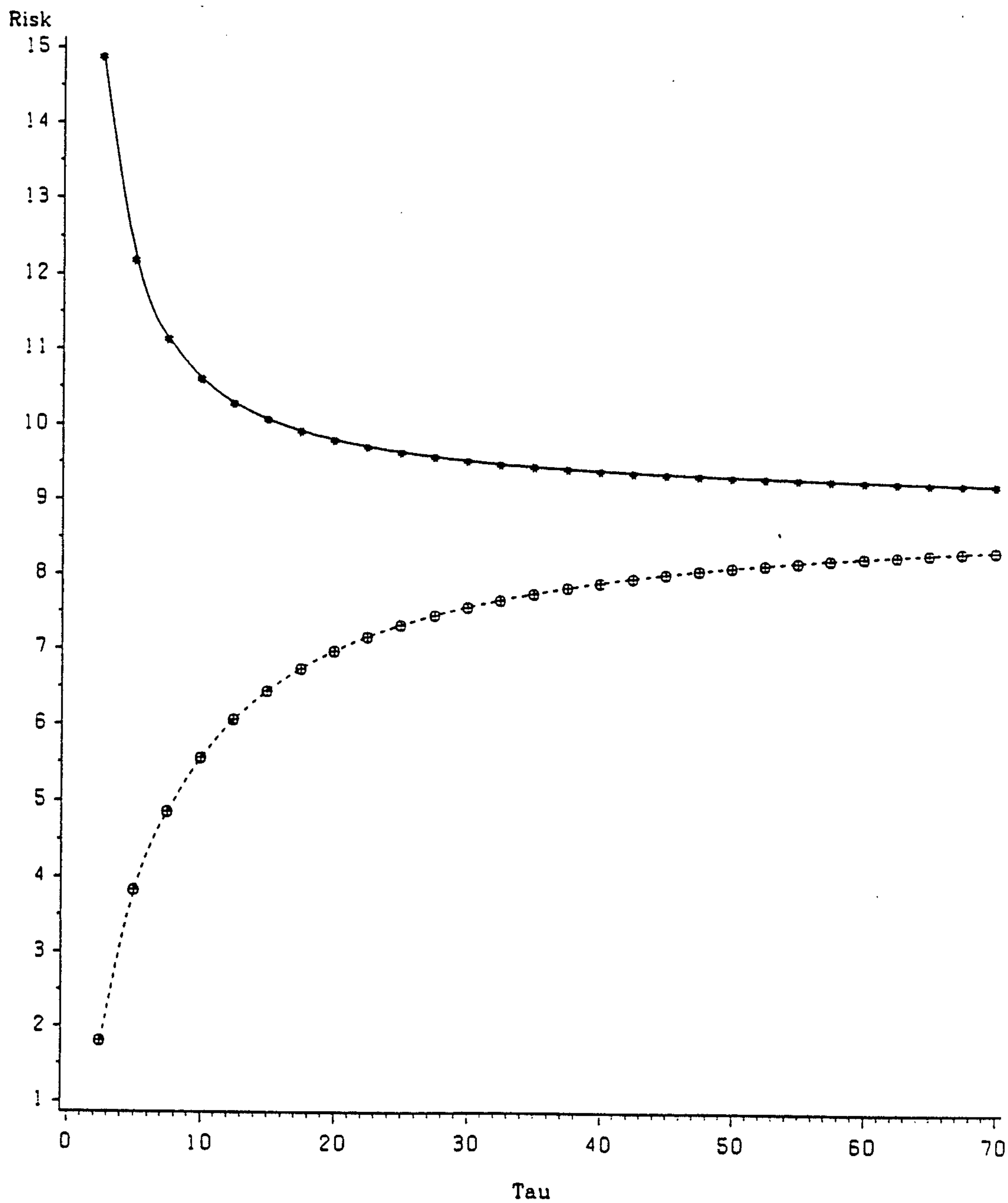
LEGENDS ---------- MLE .... GEN BAYES

THE GAIN IN RISK USING THE GENERALIZED BAYES ESTIMATOR RATHER THAN THE MLE OF A MATRIX OF POISSON MEANS UNDER A MULTIPLICATIVE MODEL FOR A COMPLETE 3X3 TABLE.



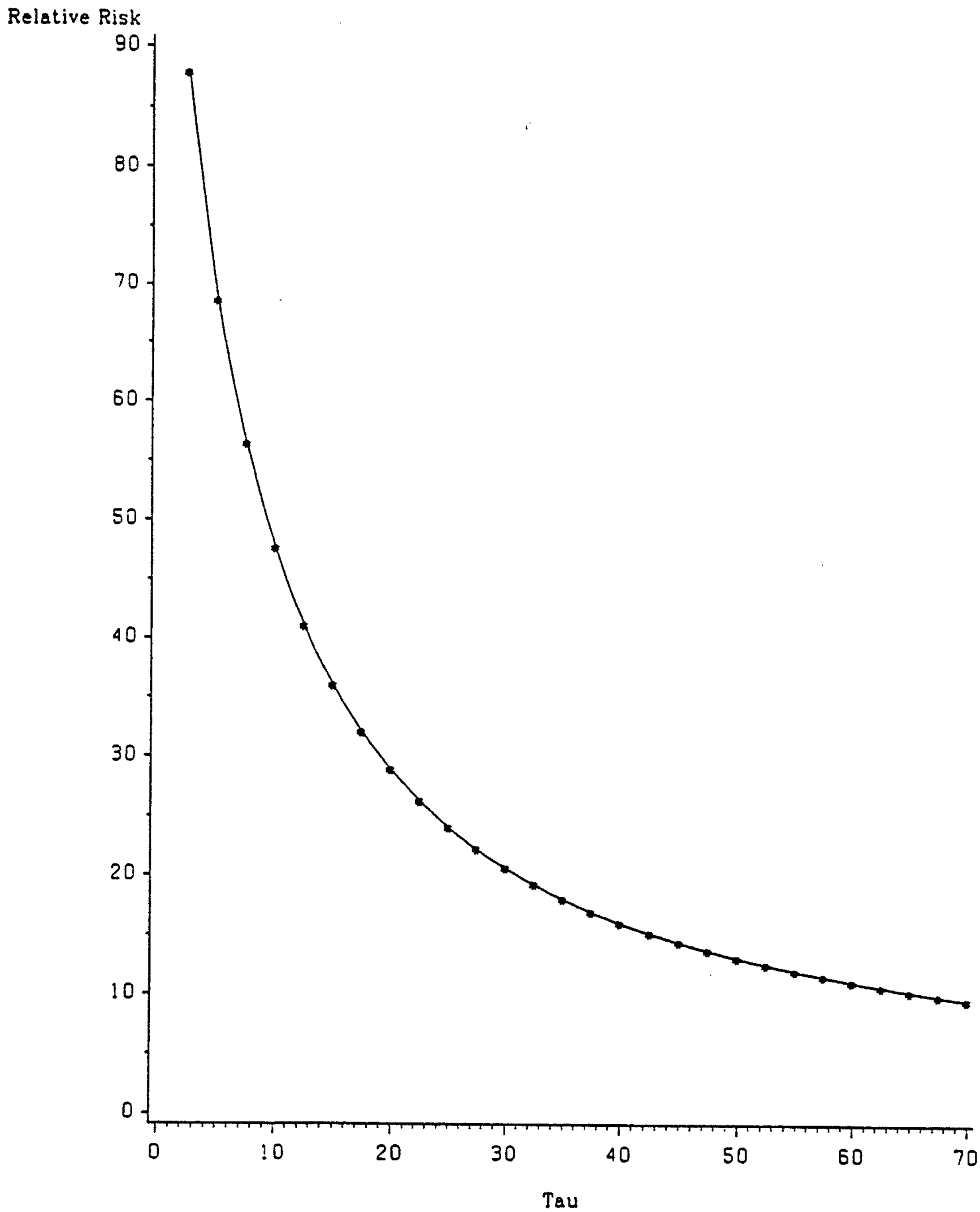
LEGENDS ---------- MLE .... GEN BAYES

A COMPARISON OF THE RISKS OF THE MLE AND THE GENERALIZED BAYES ESTIMATOR OF A VECTOR OF POISSON MEANS UNDER A A MULTIPLICATIVE MODEL FOR A COMPLETE 5X5 TABLE.



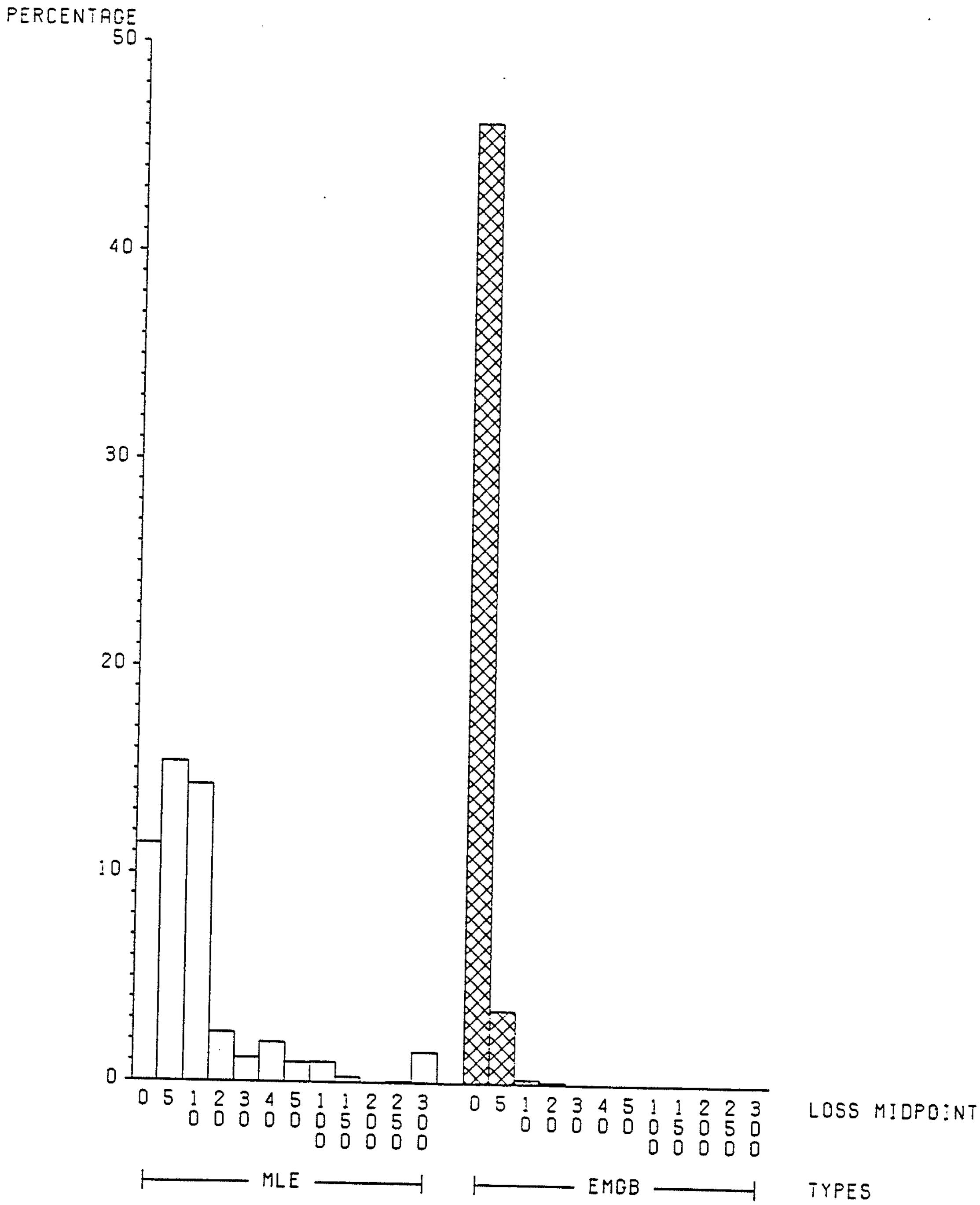
LEGENDS ---------- MLE .... GEN BAYES

THE GAIN IN RISK USING THE GENERALIZED BAYES ESTIMATOR RATHER THAN THE MLE OF A MATRIX OF POISSON MEANS UNDER A MULTIPLICATIVE MODEL FOR A COMPLETE 5X5 TABLE.

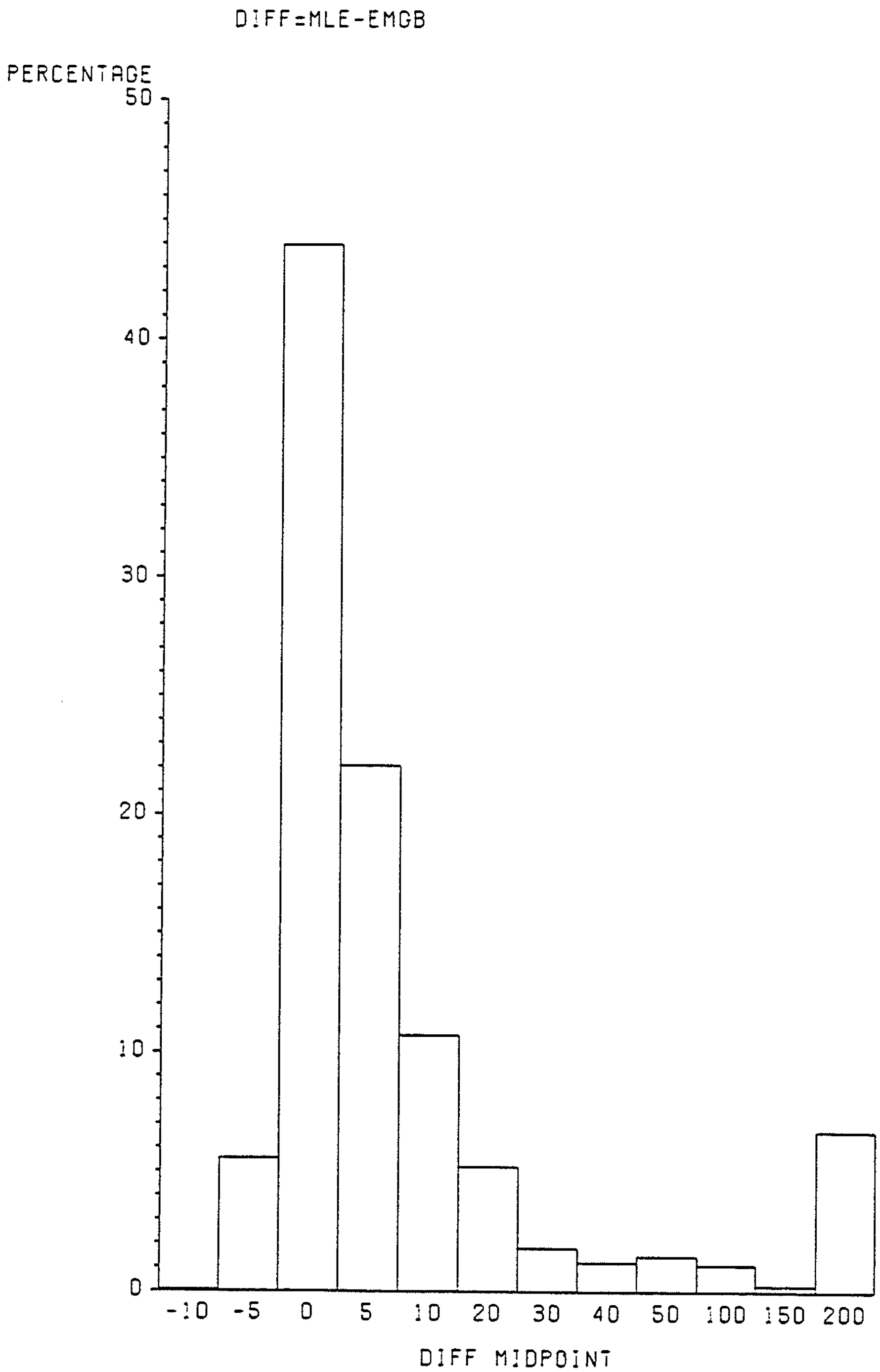


LEGENDS ---------- MLE .... GEN BAYES

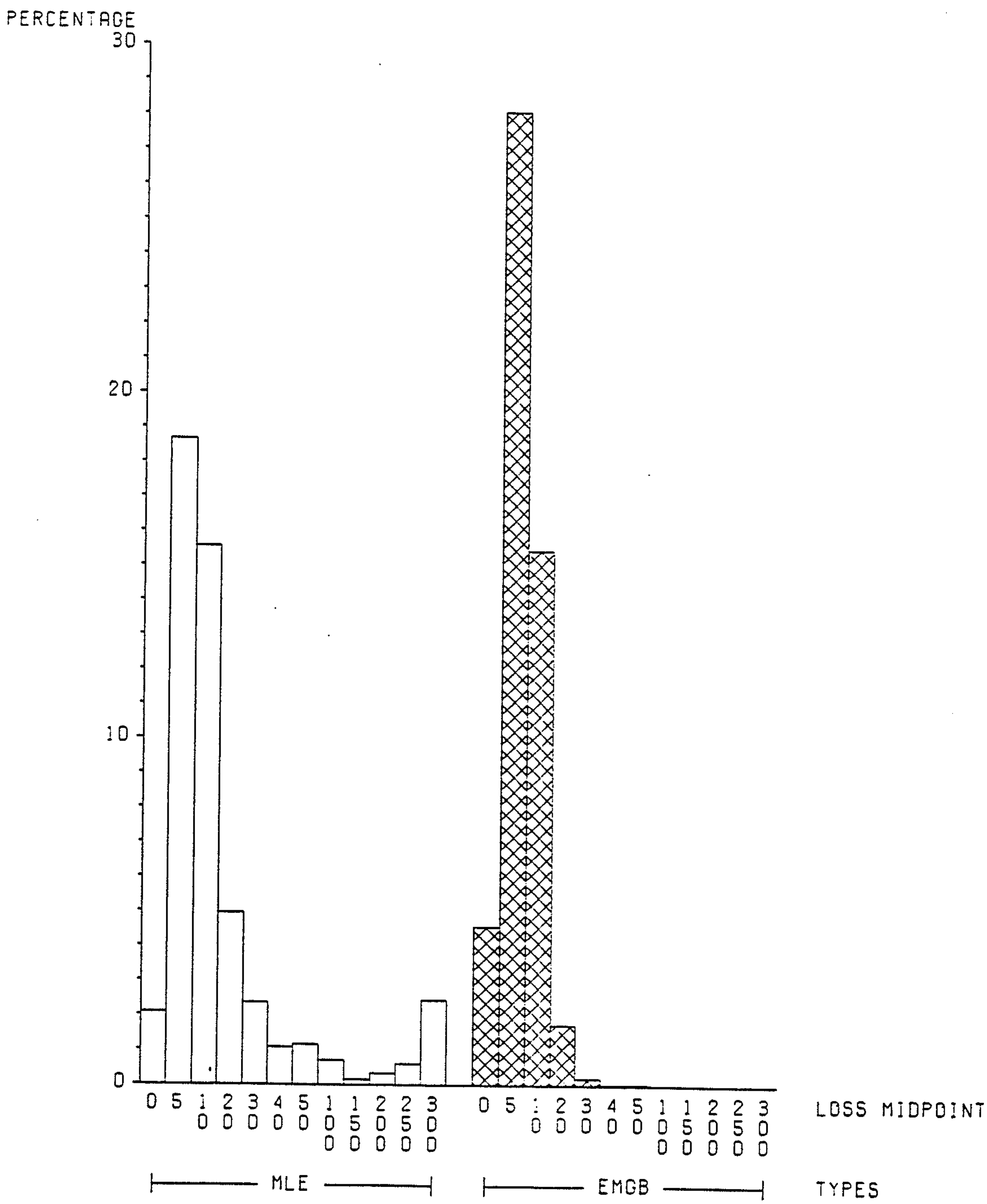
Frequency Distribution of the Losses of ML Estimates of Poisson Means under a Multiplicative Model in an Incomplete 3x3 Table. Loss Function is the Sum of the Losses for the Estimates of the Cell Means. 2000 Tables used.
ALPHA=(0.1, 0.1, 0.8); BETA=(0.4, 0.4, 0.2) ; TAU=2.0



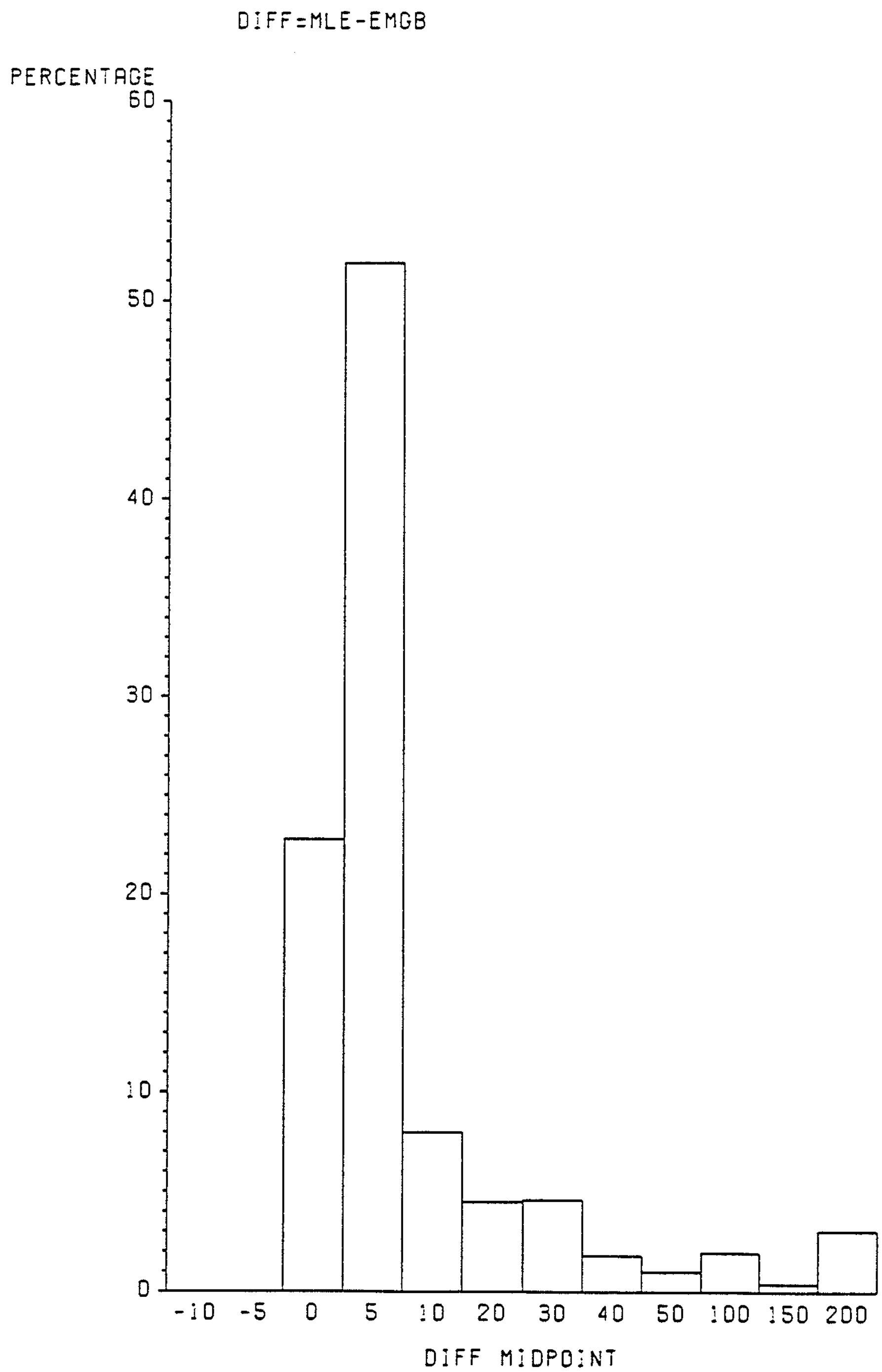
Frequency Distribution of the Difference between the Risks of the ML Estimates and the EMGB Estimators of Poisson Means under a Multiplicative Model in a 3x3 Incomplete Table. Summary Statistics for 2000 Tables.



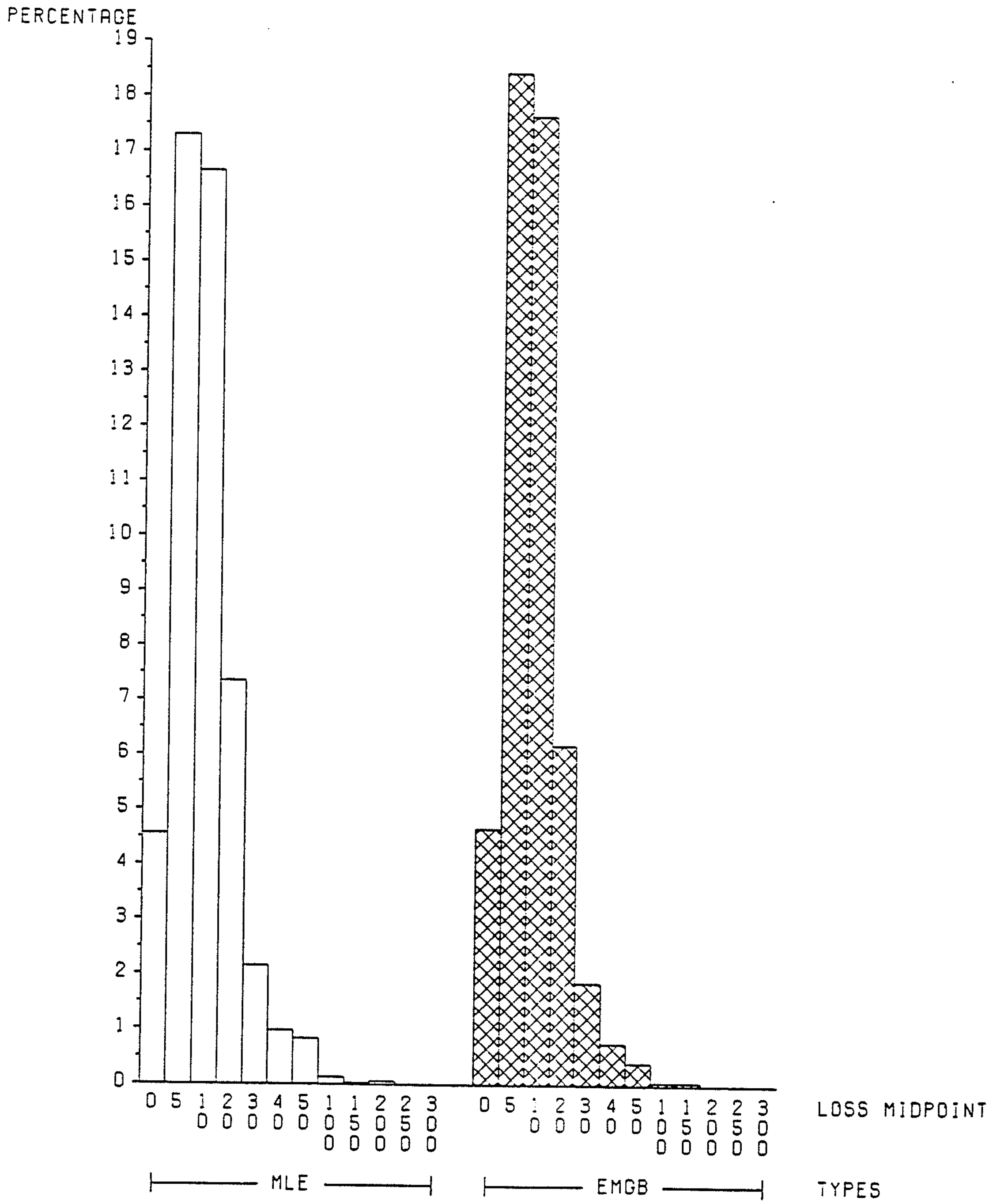
Frequency Distribution of the Losses of ML Estimates of Poisson Means under a Multiplicative Model in an Incomplete 3x3 Table. Loss Function is the Sum of the Losses for the Estimates of the Cell Means. 2000 Tables used.
ALPHA=(0.1, 0.1, 0.8); BETA=(0.4, 0.4, 0.2) ; TAU=20.0



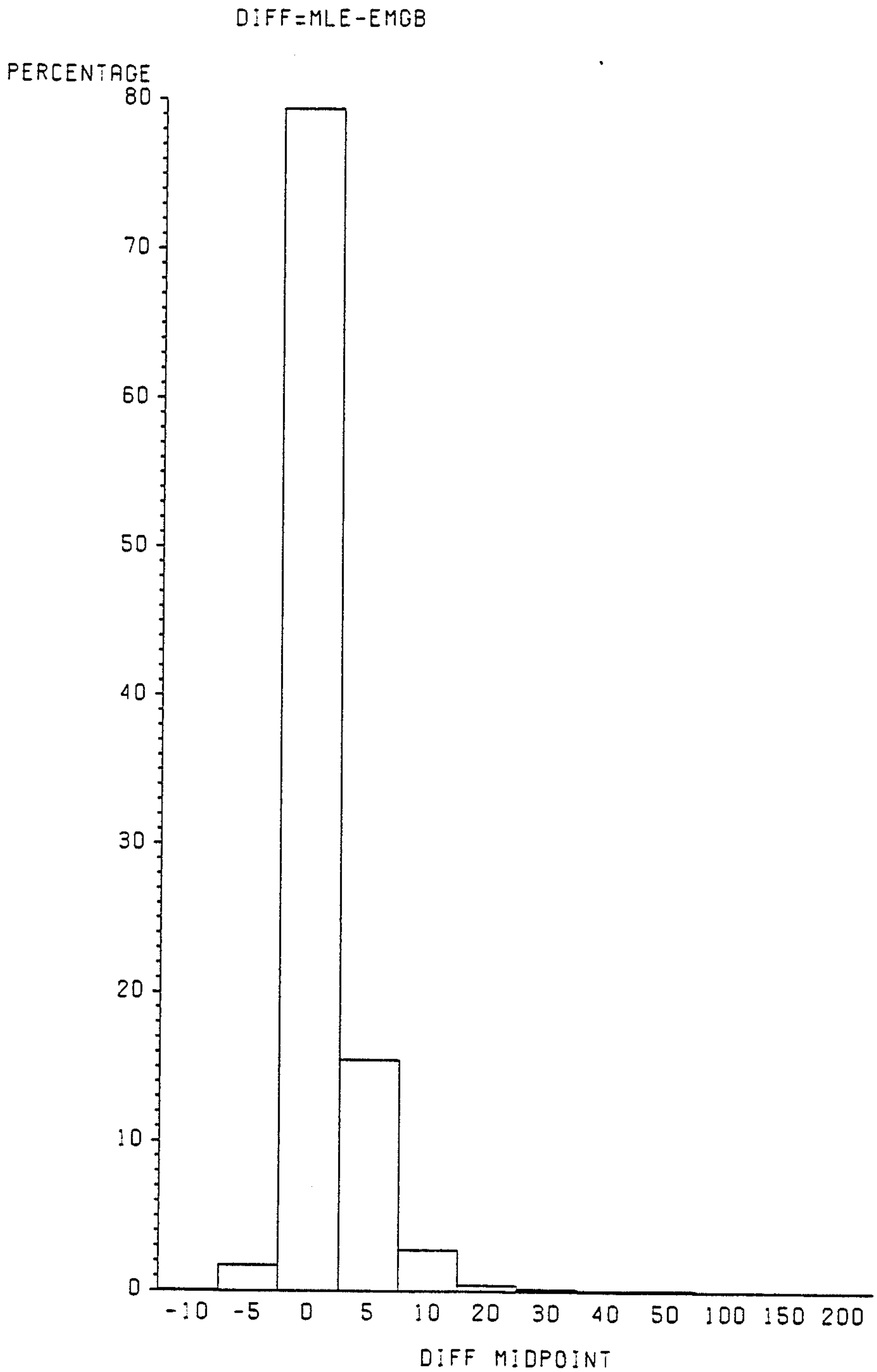
Frequency Distribution of the Difference between the Risks of the ML Estimates and the EMGB Estimators of Poisson Means under a Multiplicative Model in a 3x3 Incomplete Table. Summary Statistics for 2000 Tables.



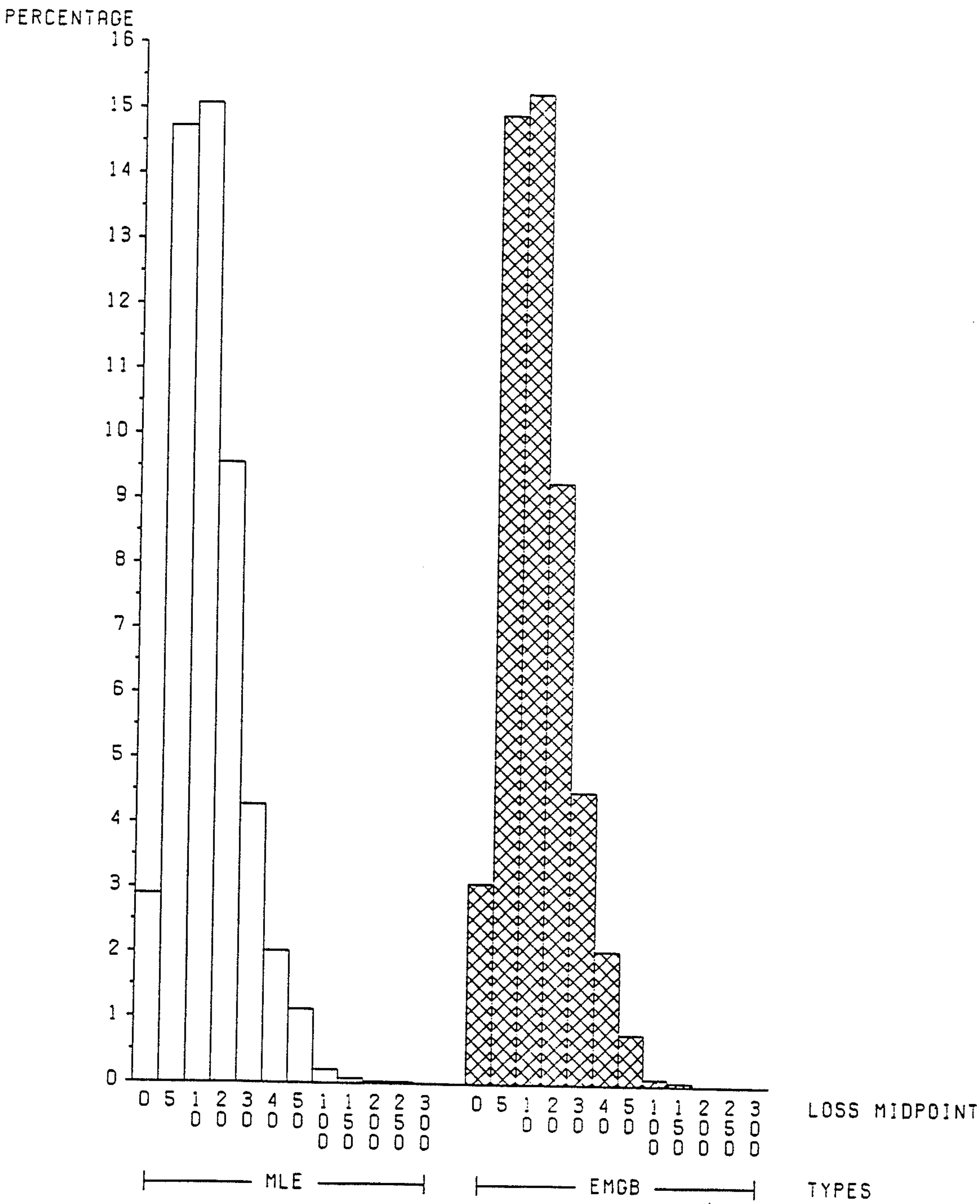
Frequency Distribution of the Losses of ML Estimates of Poisson Means under a Multiplicative Model in an Incomplete 3x3 Table. Loss Function is the Sum of the Losses for the Estimates of the Cell Means. 2000 Tables used.
ALPHA=(0.4, 0.4, 0.2); BETA=(0.4, 0.4, 0.2) ; TAU=200.0



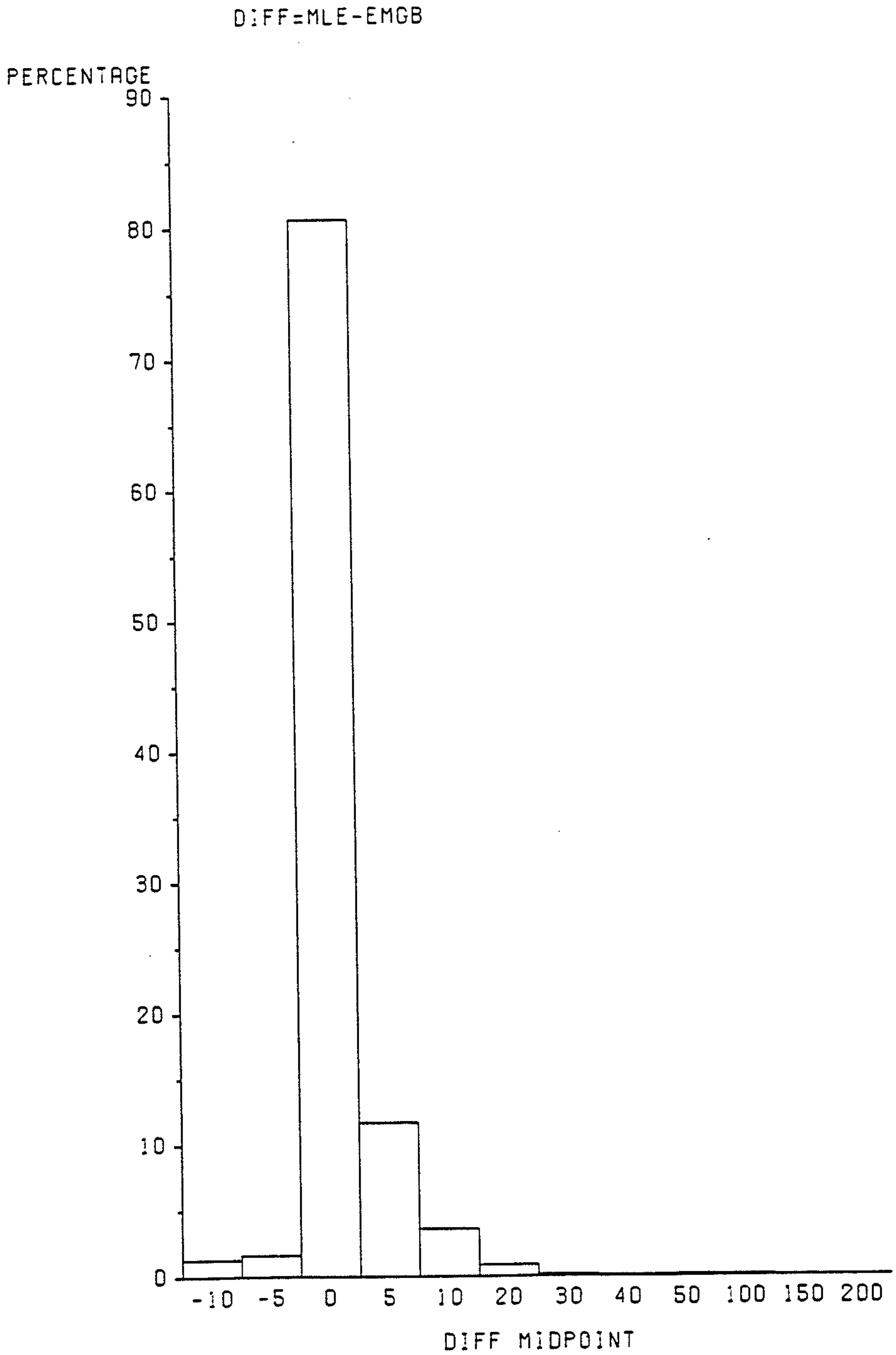
Frequency Distribution of the Difference between the Risks of the ML Estimates and the EMGB Estimators of Poisson Means under a Multiplicative Model in a 3x3 Incomplete Table. Summary Statistics for 2000 Tables.



Frequency Distribution of the Losses of ML Estimates of Poisson Means under a Multiplicative Model in an Incomplete 3x3 Table. Loss Function is the Sum of the Losses for the Estimates of the Cell Means. 2000 Tables used.
ALPHA=(0.1, 0.1, 0.8); BETA=(0.4, 0.4, 0.2) ; TAU=500.0

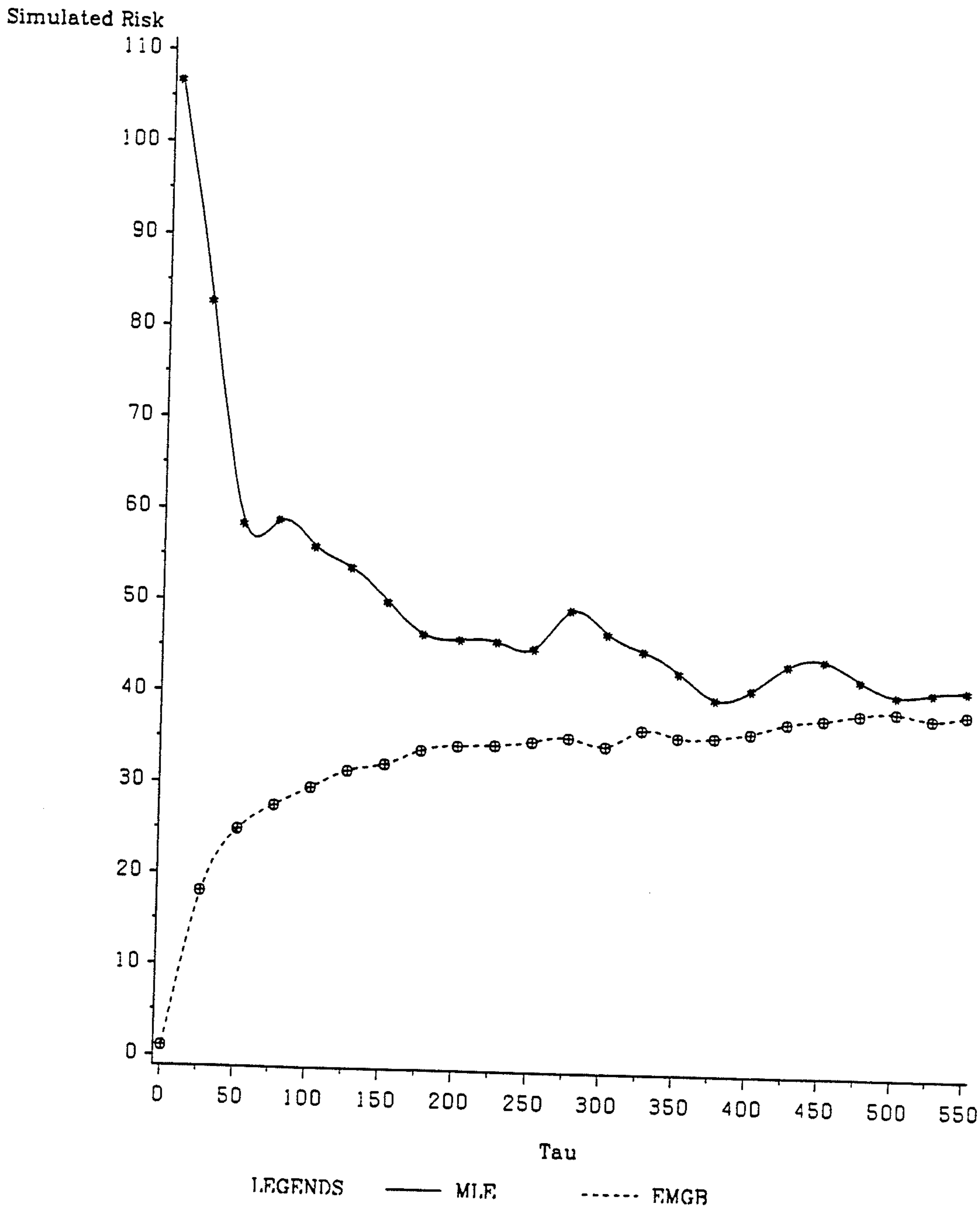


Frequency Distribution of the Difference between the Risks of the ML Estimates and the EMGB Estimators of Poisson Means under a Multiplicative Model in a 3x3 Incomplete Table. Summary Statistics for 2000 Tables.



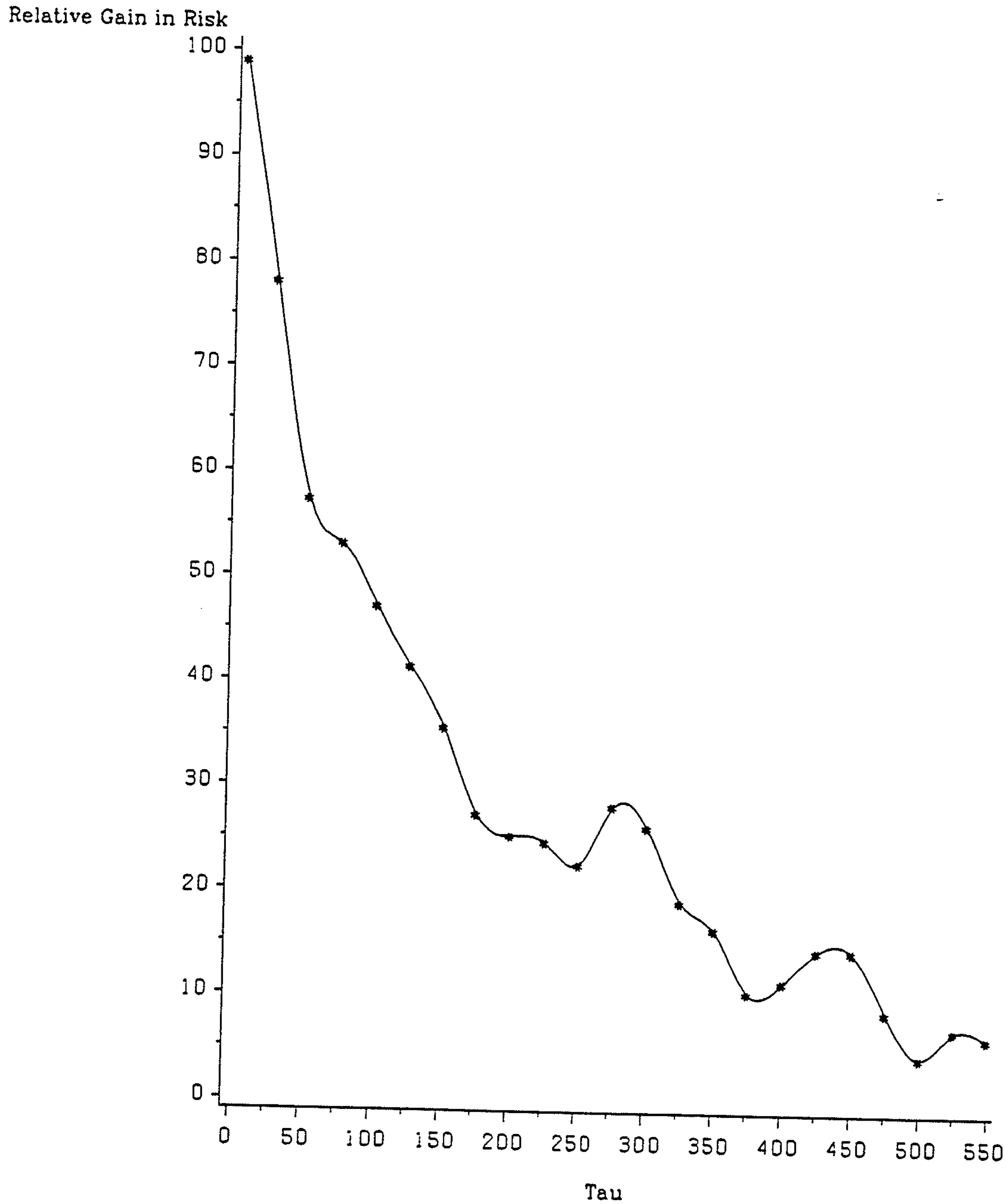
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Vector of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1.
ALPHA=(0.1, 0.1, 0.8); BETA=(0.1, 0.1, 0.8)



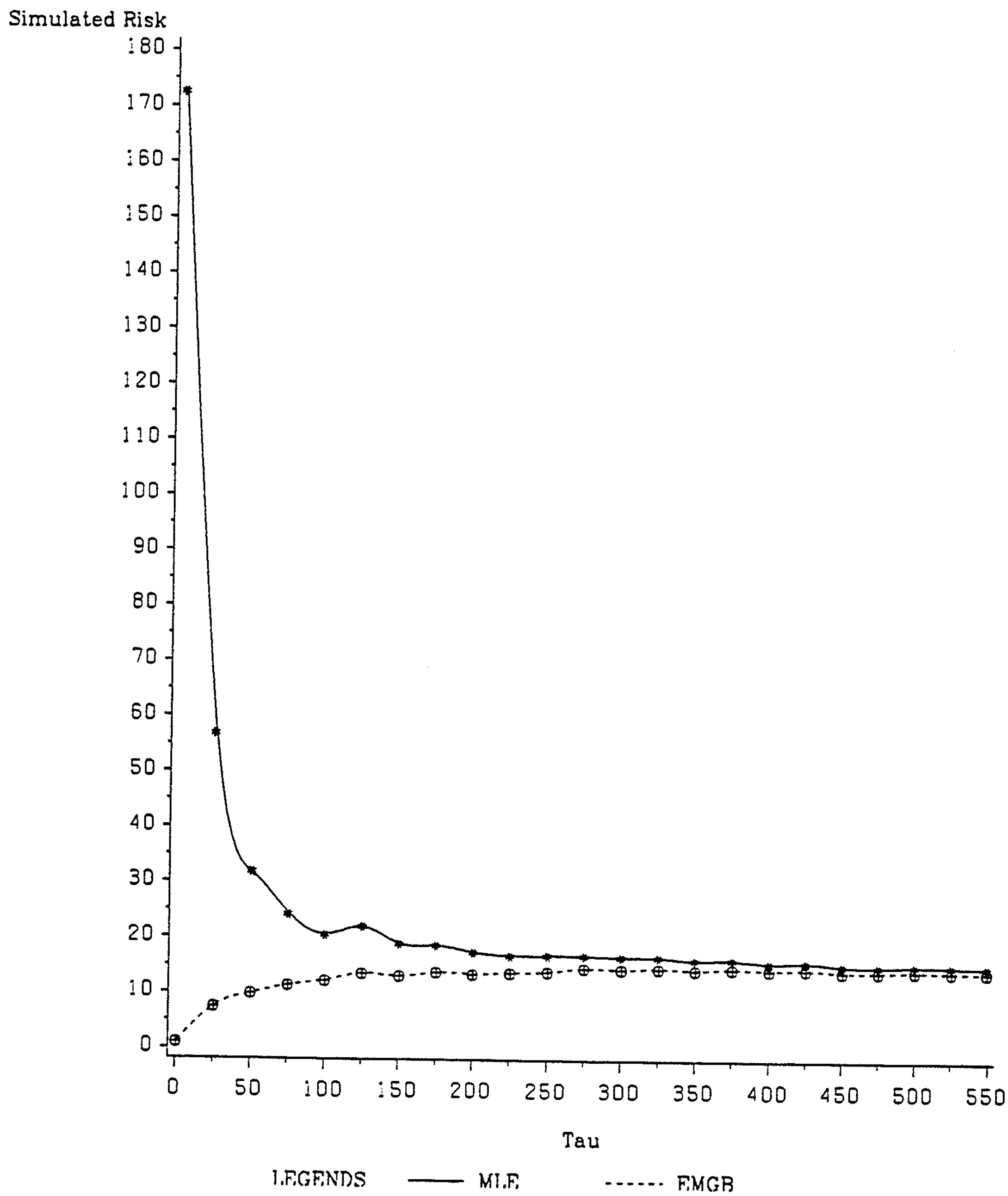
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1
ALPHA=(0.1, 0.1, 0.8);BETA=(0.1, 0.1, 0.8)



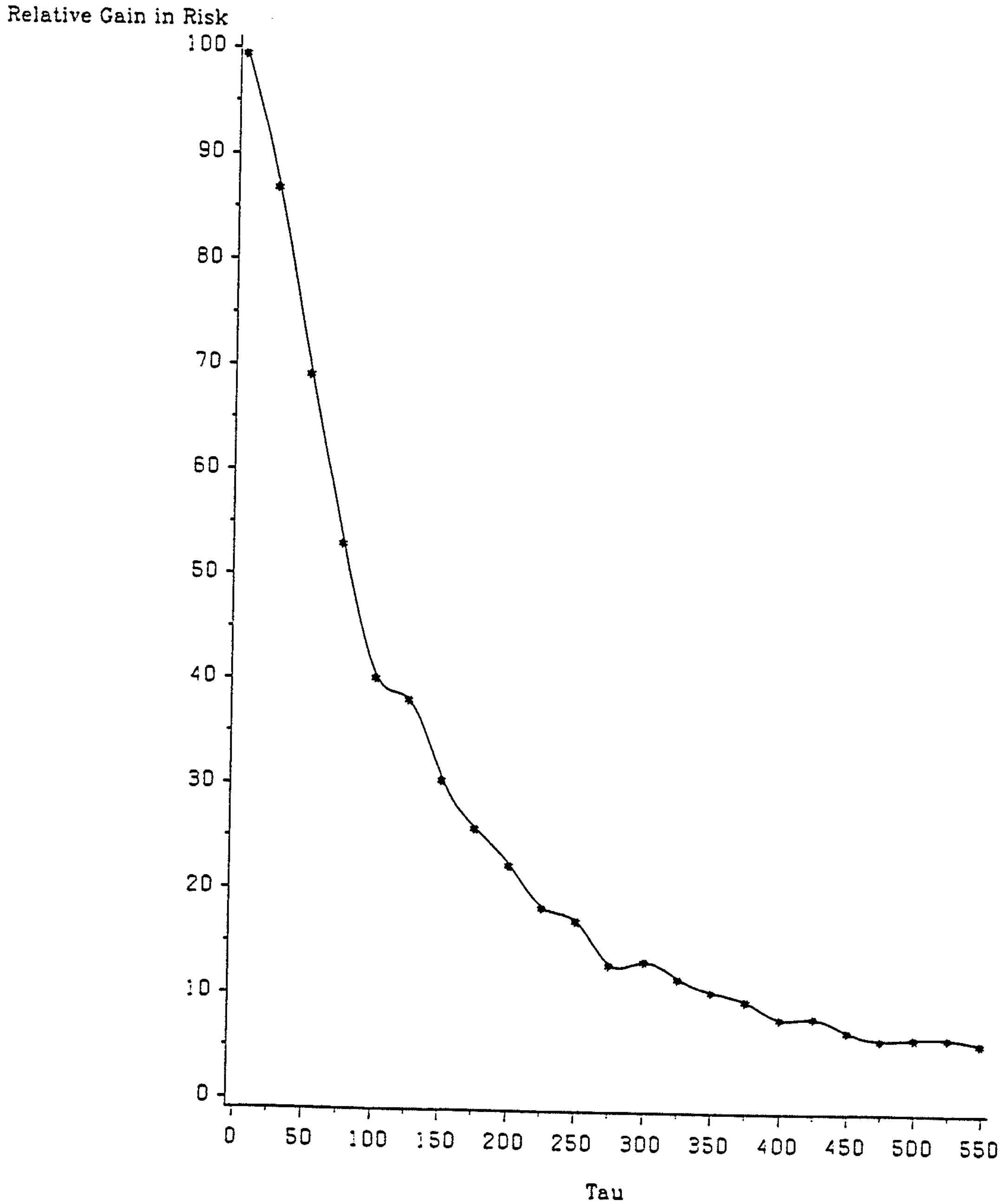
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Vector of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.
Incidence Matrix is IM2.

ALPHA=(0.1, 0.1, 0.8); BETA=(0.1, 0.1, 0.8)



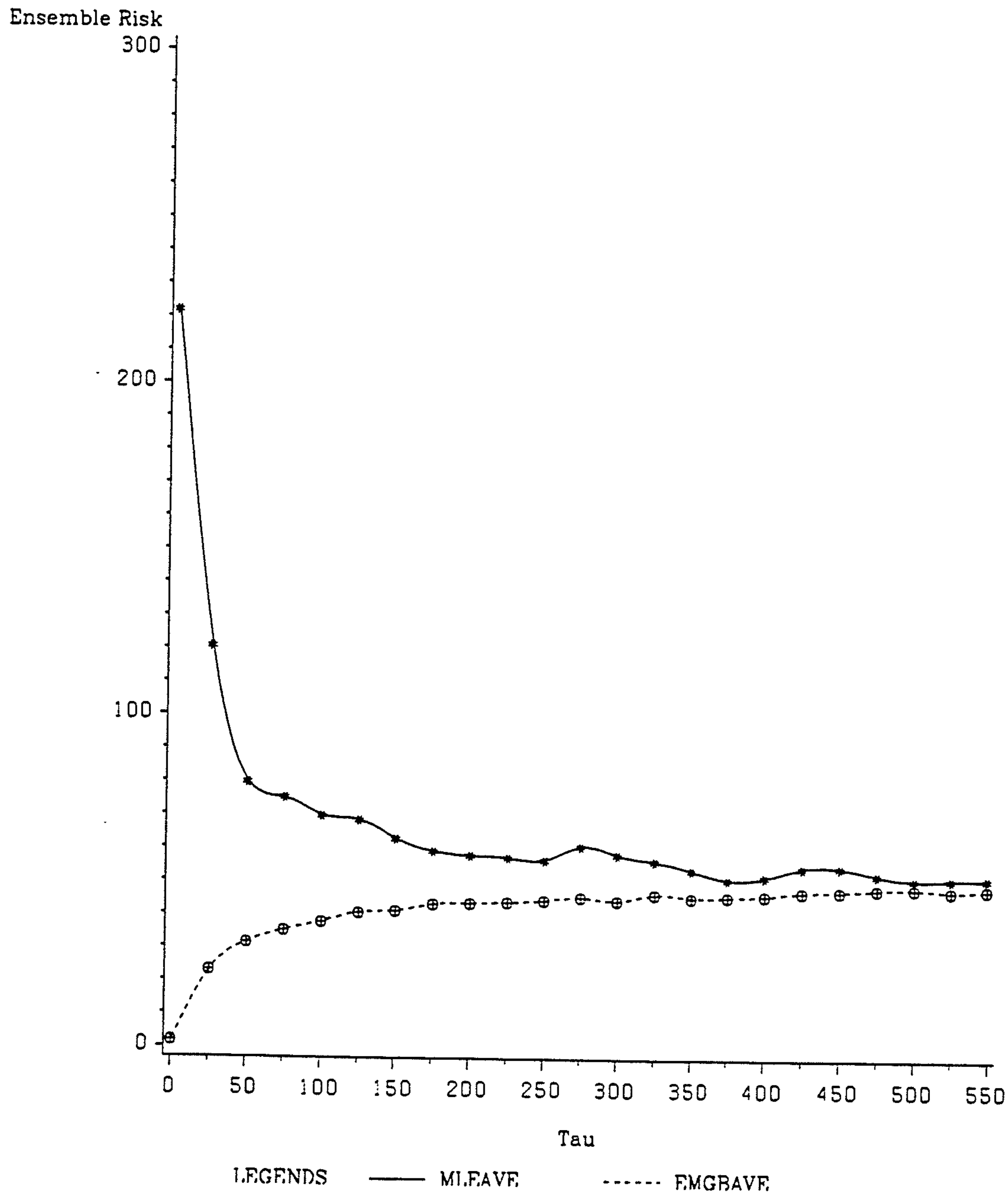
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM2
ALPHA=(0.1, 0.1, 0.8);BETA=(0.1, 0.1, 0.8)



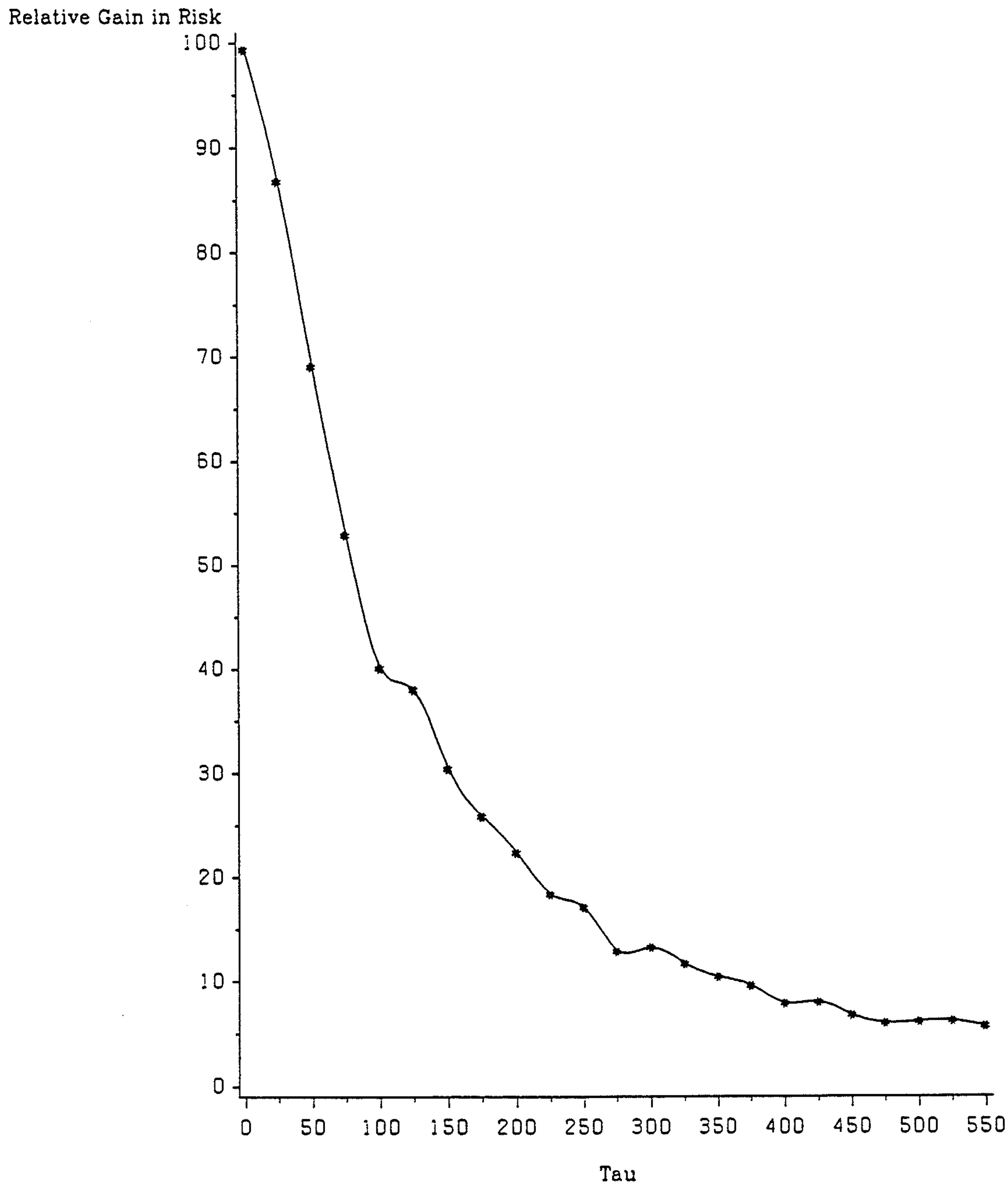
A Comparison of the Ensemble Risks of the MLE and the EMGB Estimator of a Vector of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

MLEAVE=1/3 MLRIS1 + 2/3 MLRIS2
EMGBAVE=1/3 BARIS1+2/3 BARIS2
ALPHA=(0.1, 0.1, 0.8); BETA=(0.1, 0.1, 0.8)



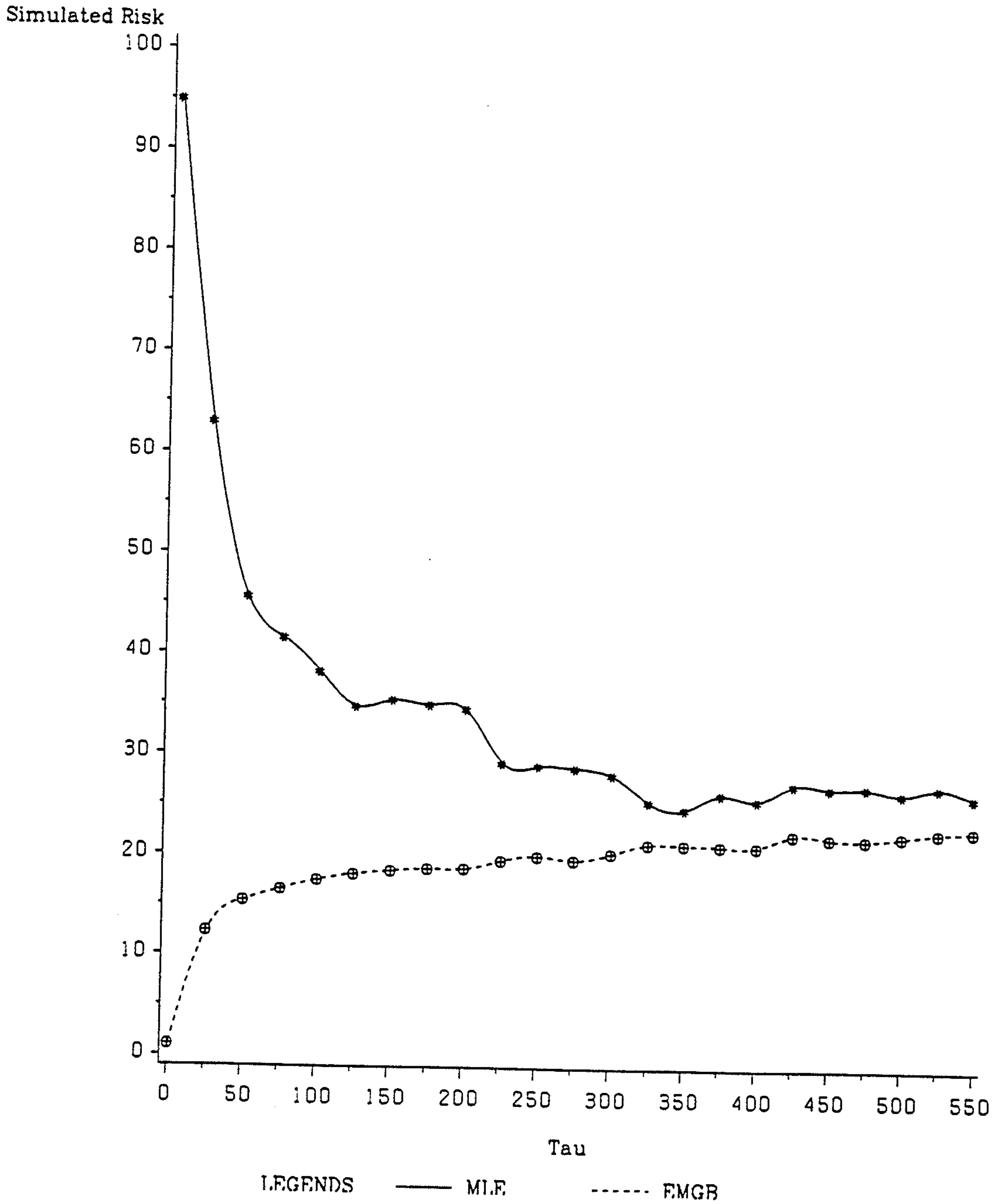
The Simulated Relative Gain in Ensemble Risk using the EMGB Estimator rather than the MLE of a Vector Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

ALPHA=(0.1, 0.1, 0.8); BETA=(0.1, 0.1, 0.8)



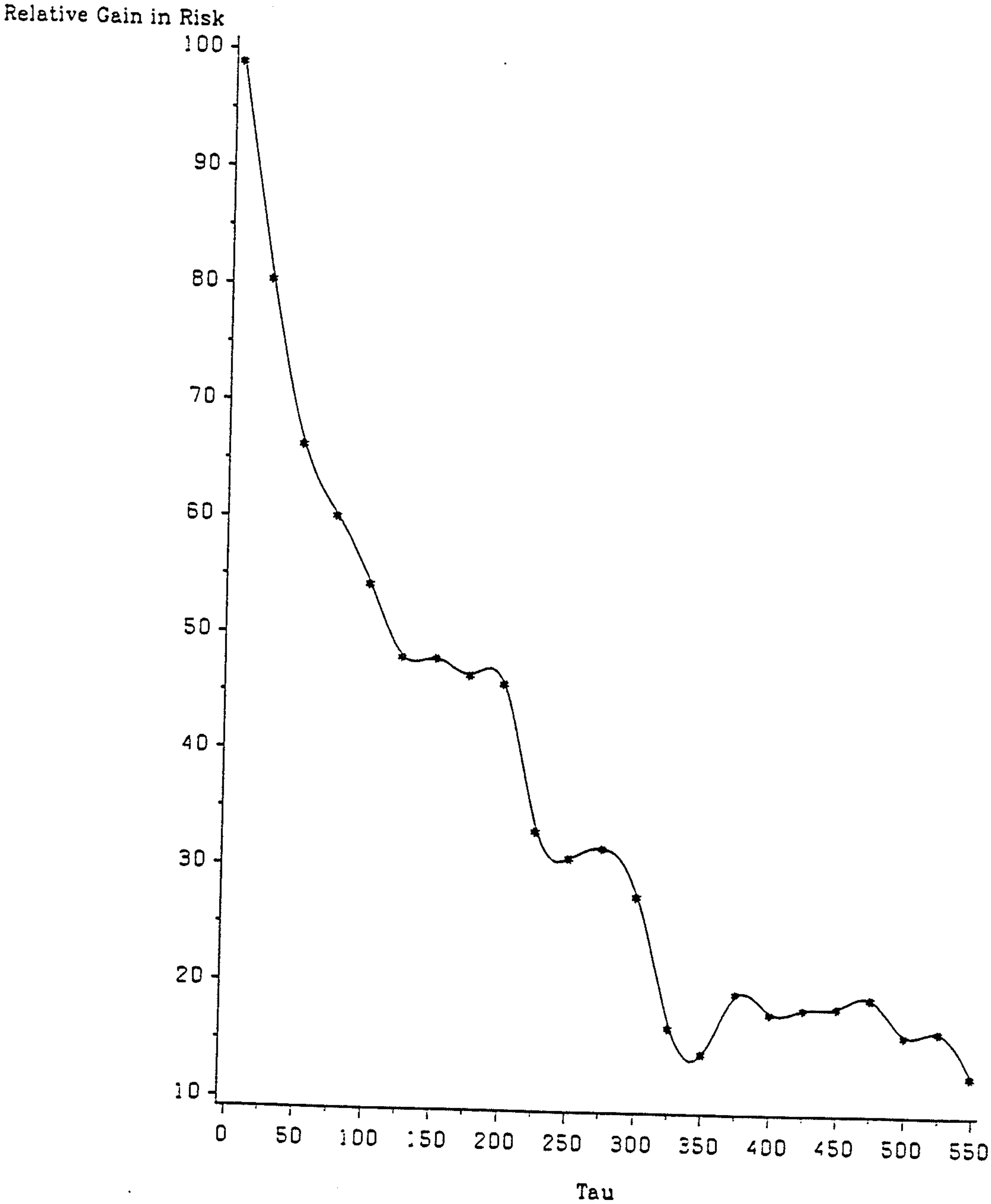
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Vector of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1.
ALPHA=(0.1, 0.1, 0.8); BETA=(0.2, 0.2, 0.6)



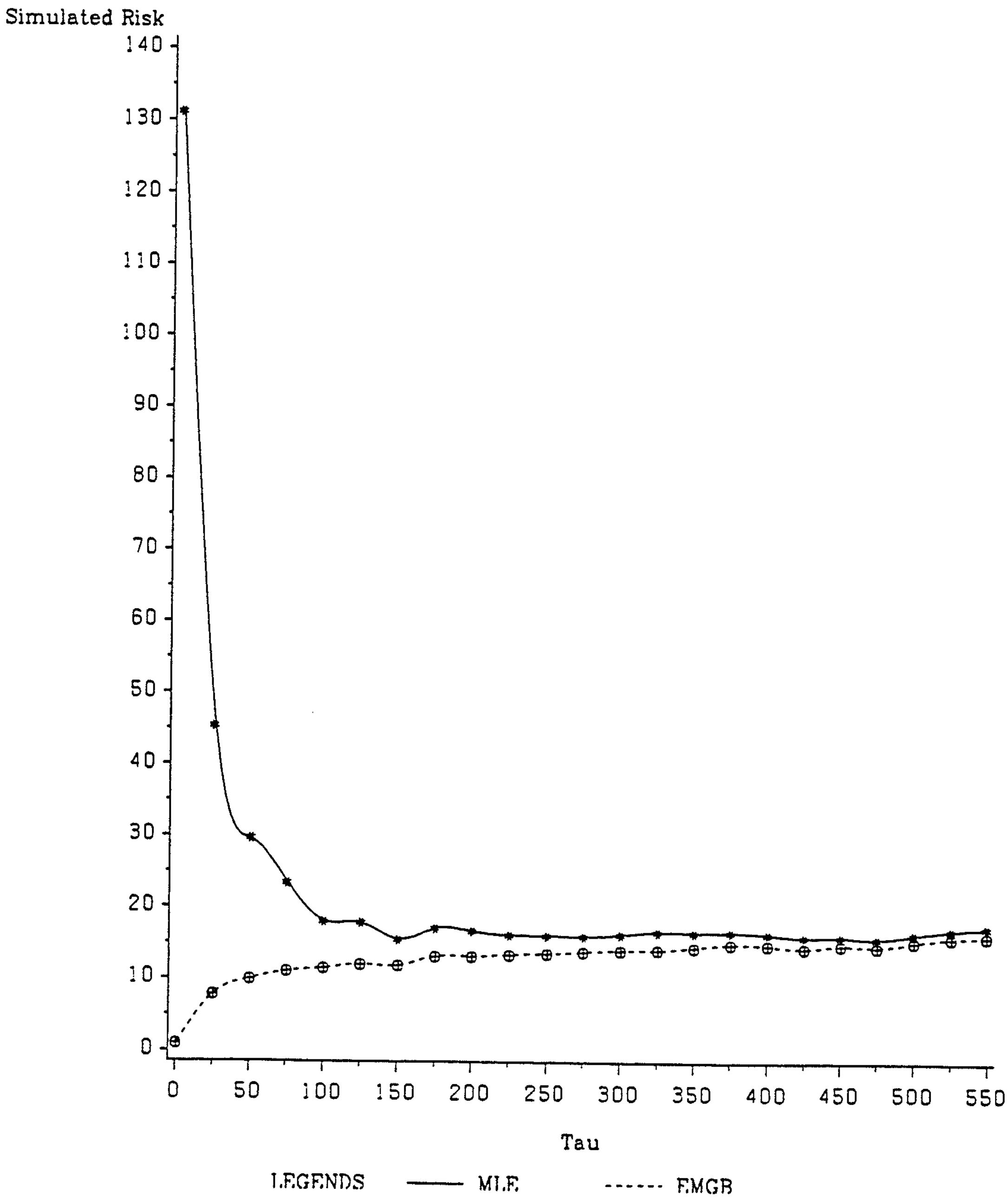
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1
ALPHA=(0.1, 0.1, 0.8);BETA=(0.2, 0.2, 0.6)



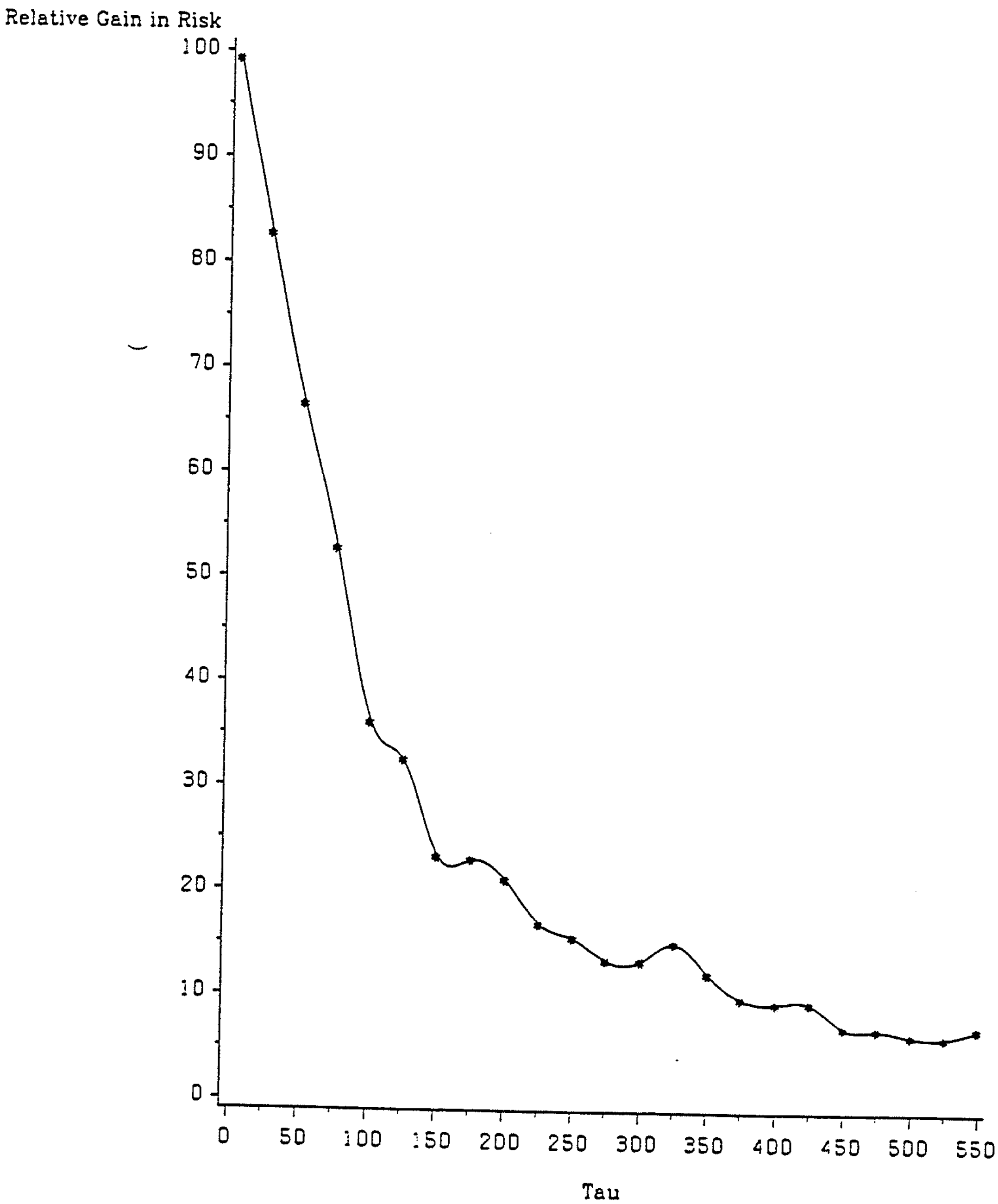
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Vector of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table. Incidence Matrix is IM2.

ALPHA=(0.1, 0.1, 0.8); BETA=(0.2, 0.2, 0.6)



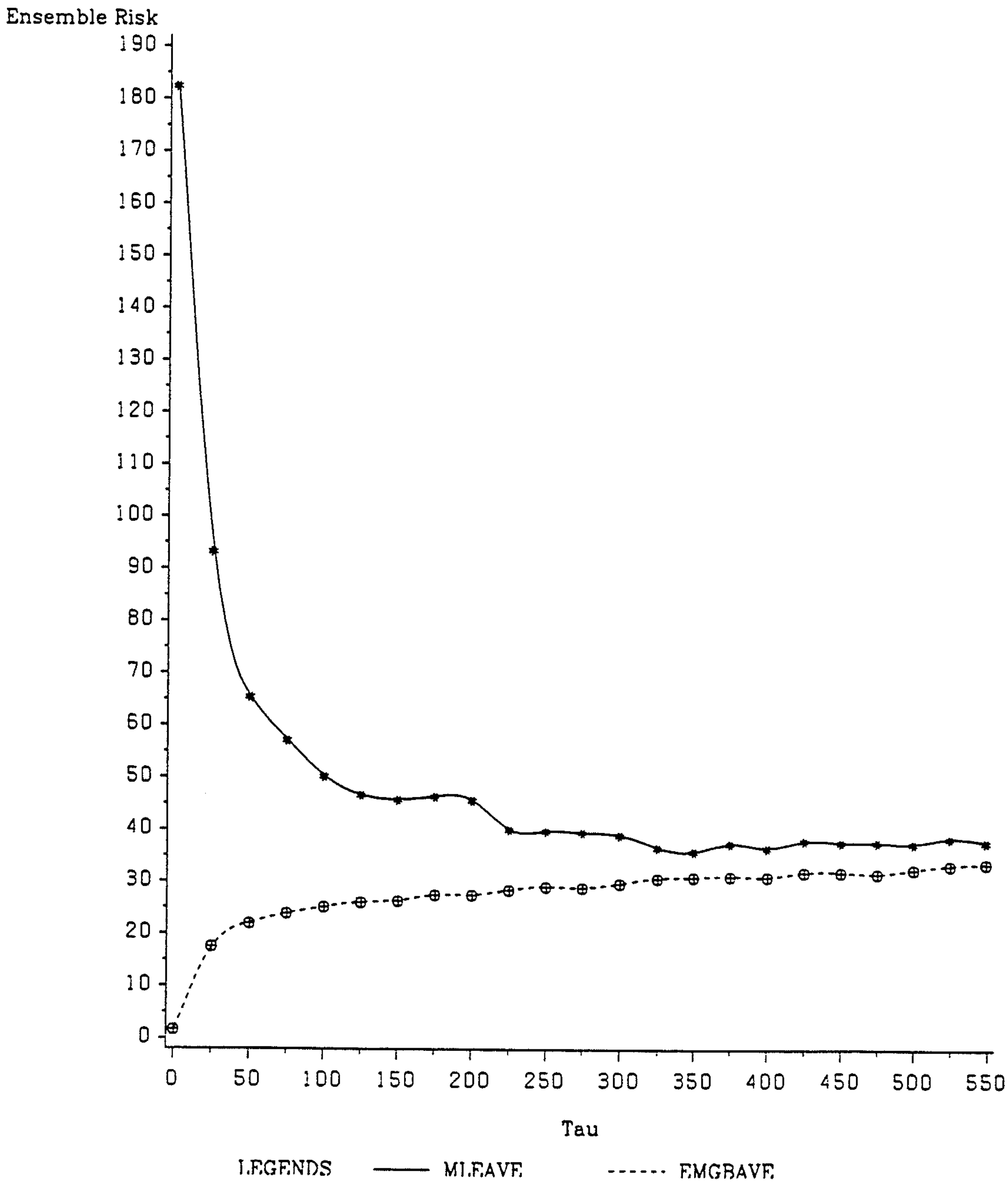
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM2
ALPHA=(0.1, 0.1, 0.8);BETA=(0.2, 0.2, 0.6)



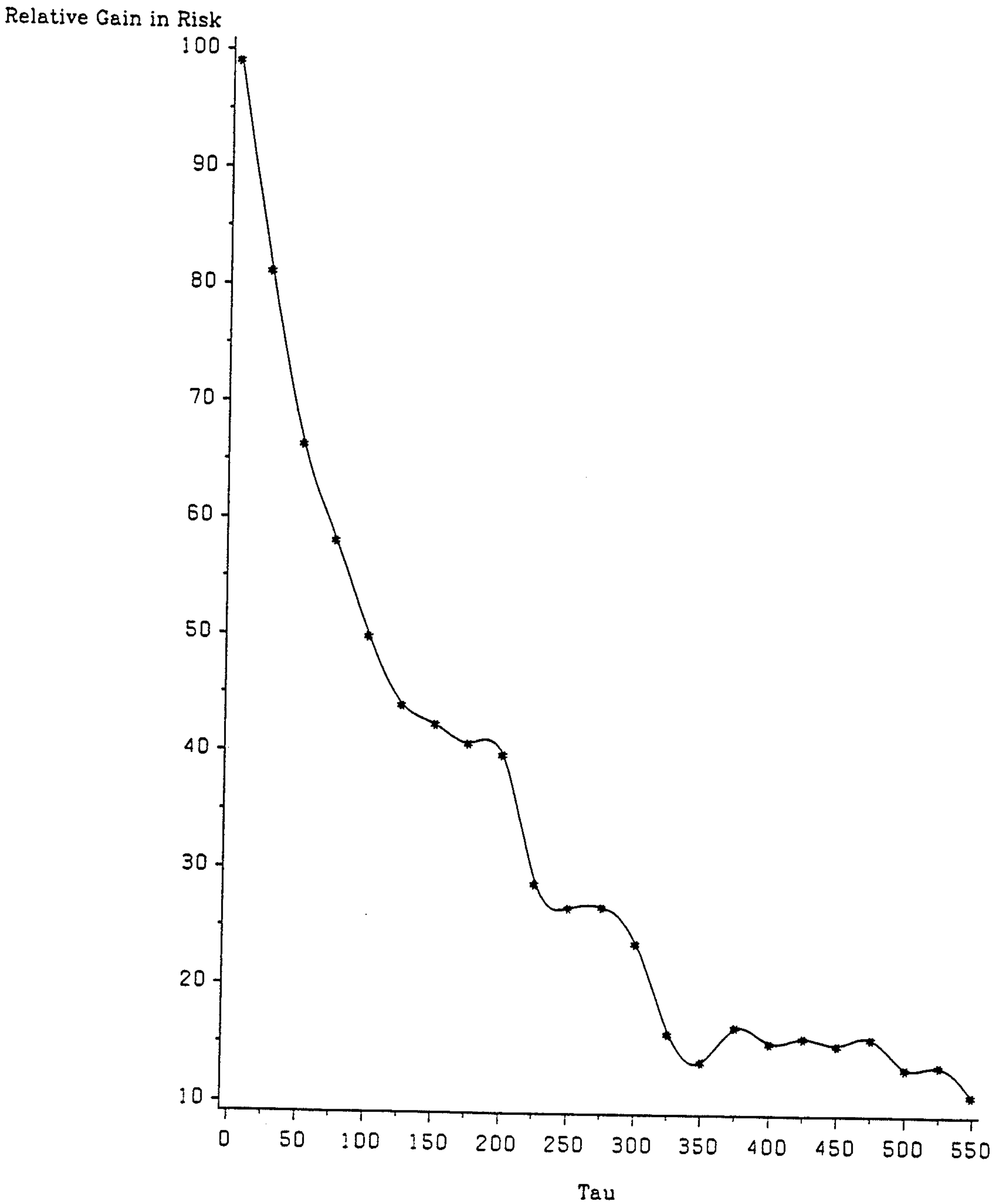
A Comparison of the Ensemble Risks of the MLE and
the EMGB Estimator of a Vector of Poisson Means
under a Multiplicative Model for a Balanced
Incomplete 3x3 Table.

MLEAVE=1/3 MLRIS1 + 2/3 MLRIS2
EMGBAVE=1/3 BARIS1+2/3 BARIS2
ALPHA=(0.1, 0.1, 0.8); BETA=(0.2, 0.2, 0.6)



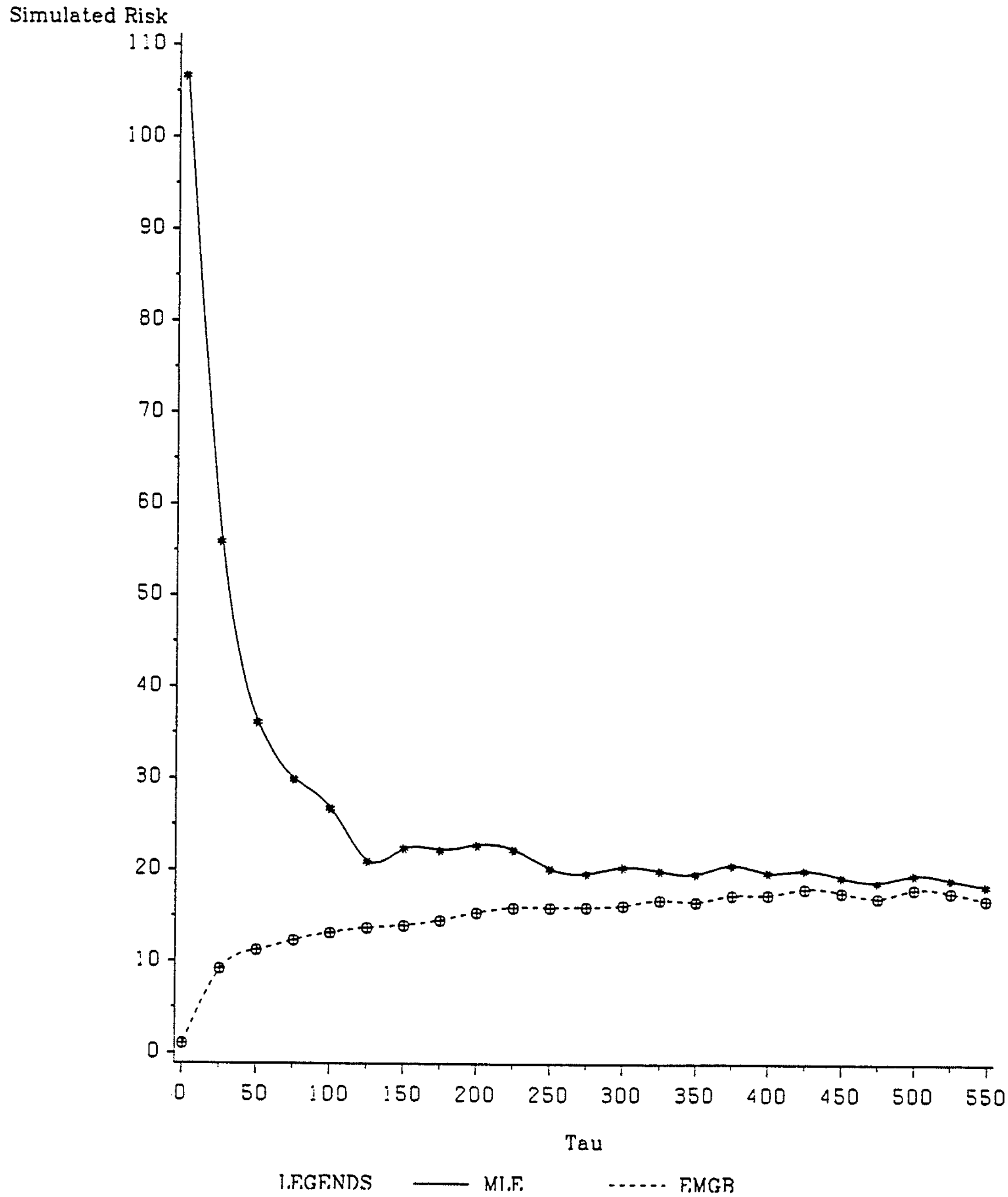
The Simulated Relative Gain in Ensemble Risk using the EMGB Estimator rather than the MLE of a Vector Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

ALPHA=(0.1, 0.1, 0.8); BETA=(0.2, 0.2, 0.6)



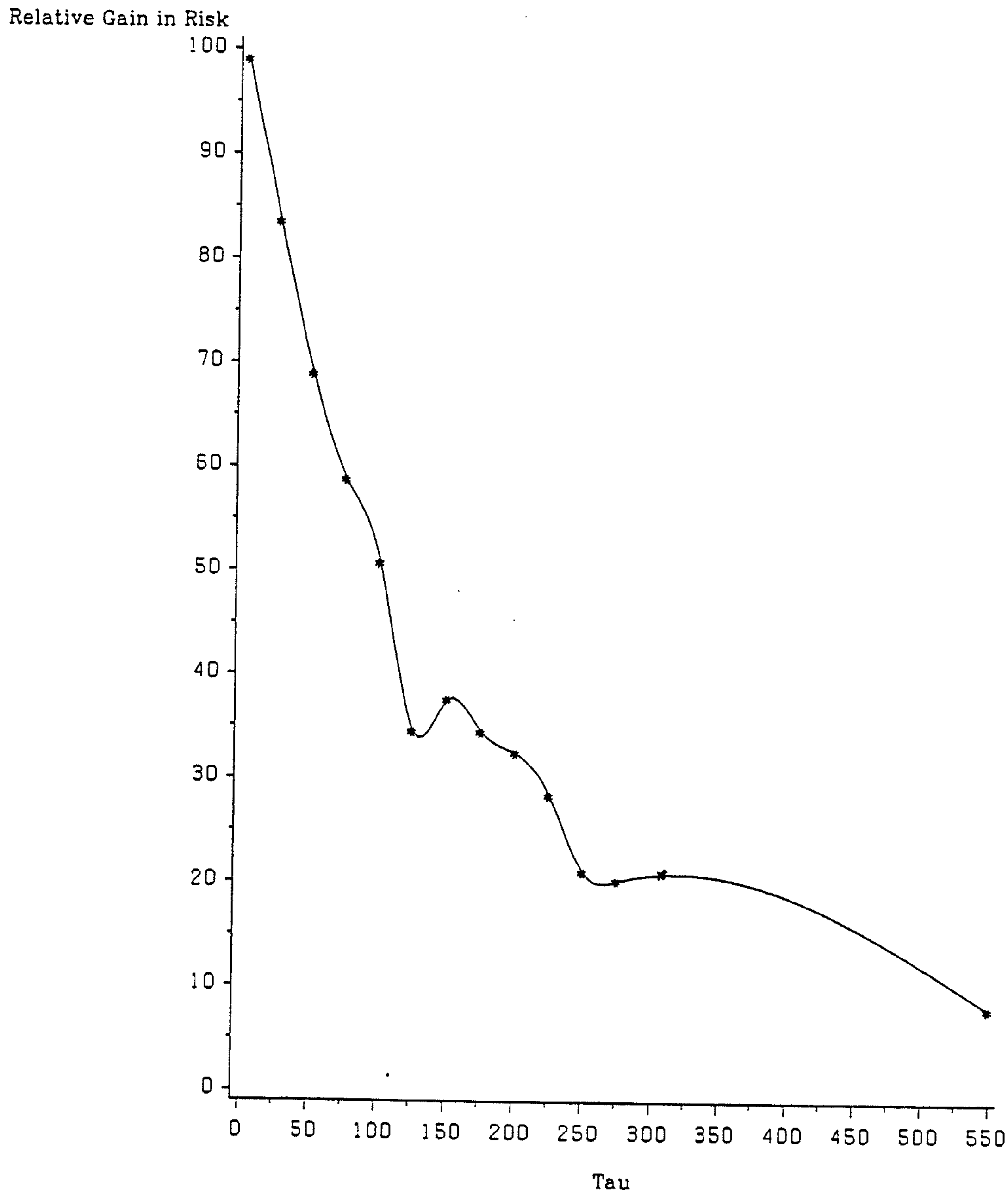
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Vector of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1.
ALPHA=(0.1, 0.1, 0.8); BETA=(0.3, 0.3, 0.4)



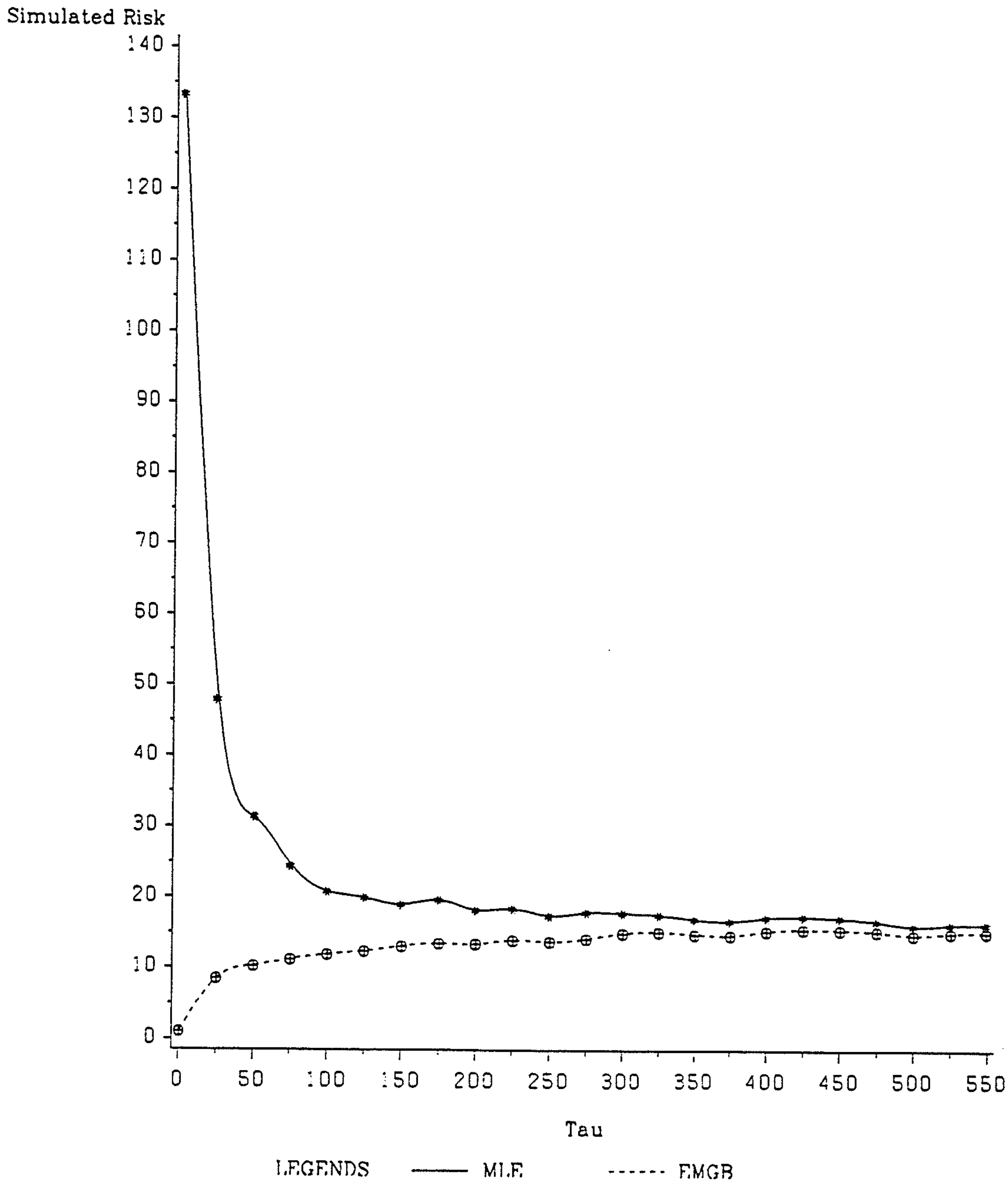
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1
ALPHA=(0.1, 0.1, 0.8);BETA=(0.3, 0.3, 0.4)



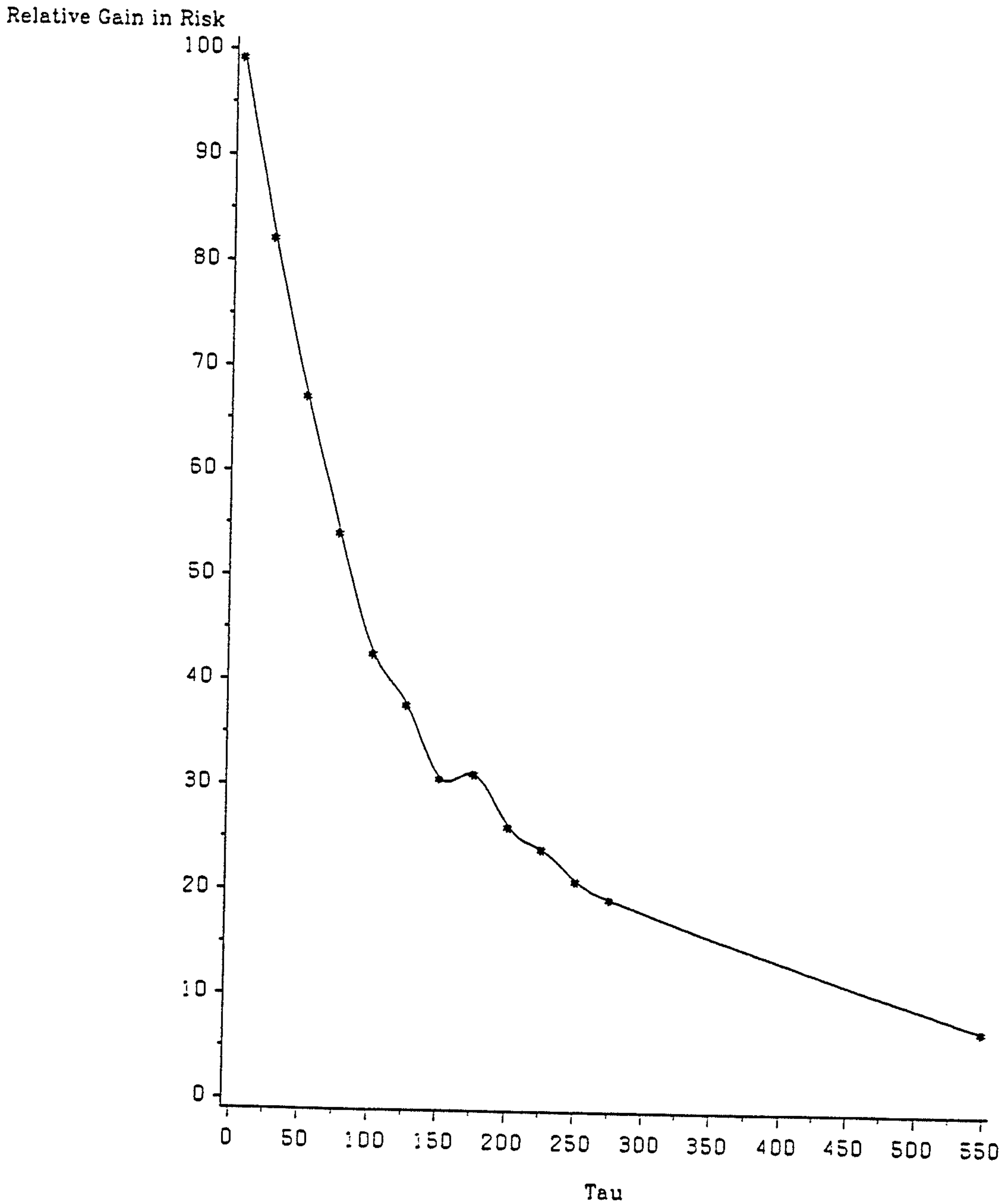
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Vector of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table. Incidence Matrix is IM2.

ALPHA=(0.1, 0.1, 0.8); BETA=(0.3, 0.3, 0.4)



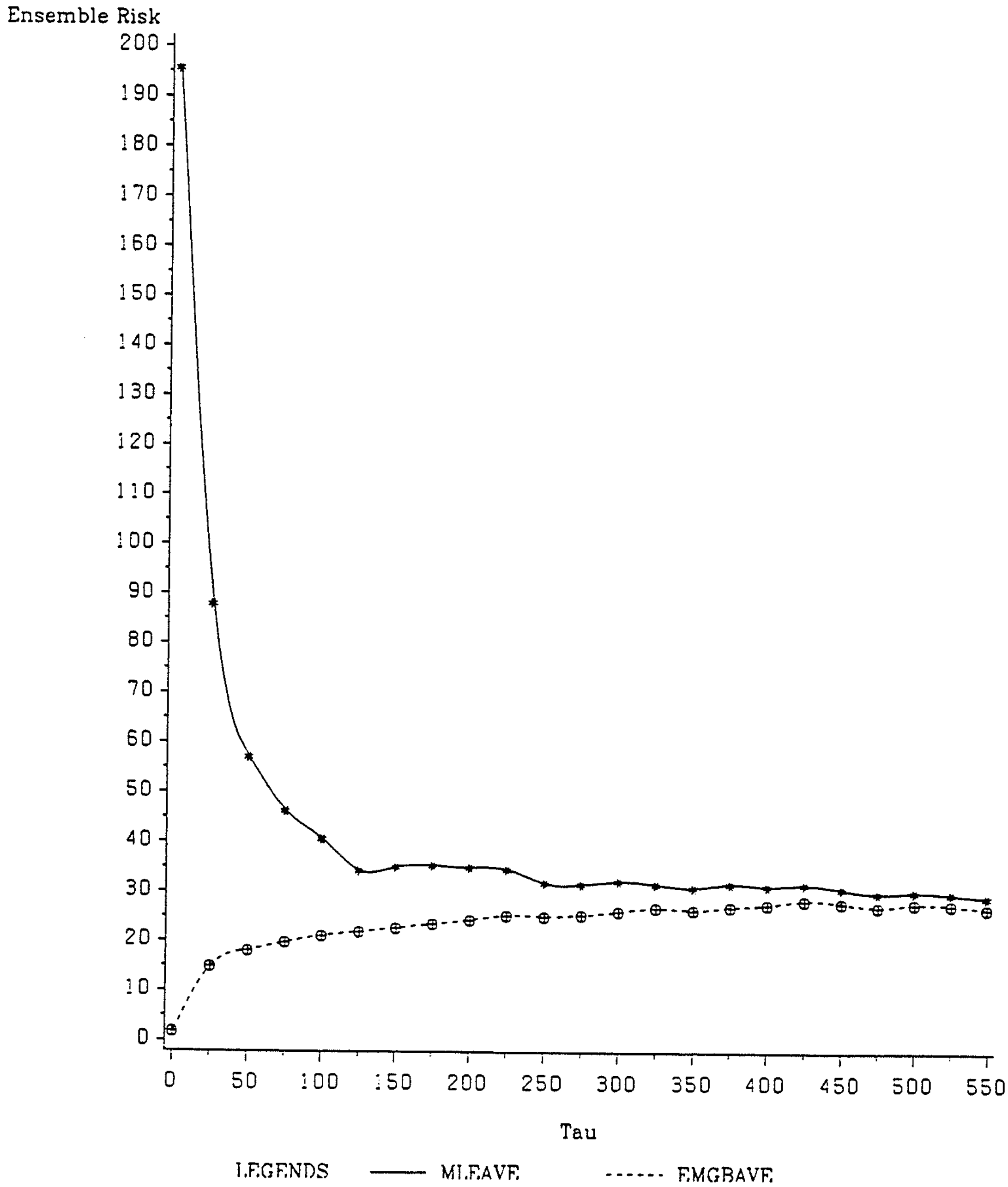
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM2
ALPHA=(0.1, 0.1, 0.8);BETA=(0.3, 0.3, 0.4)



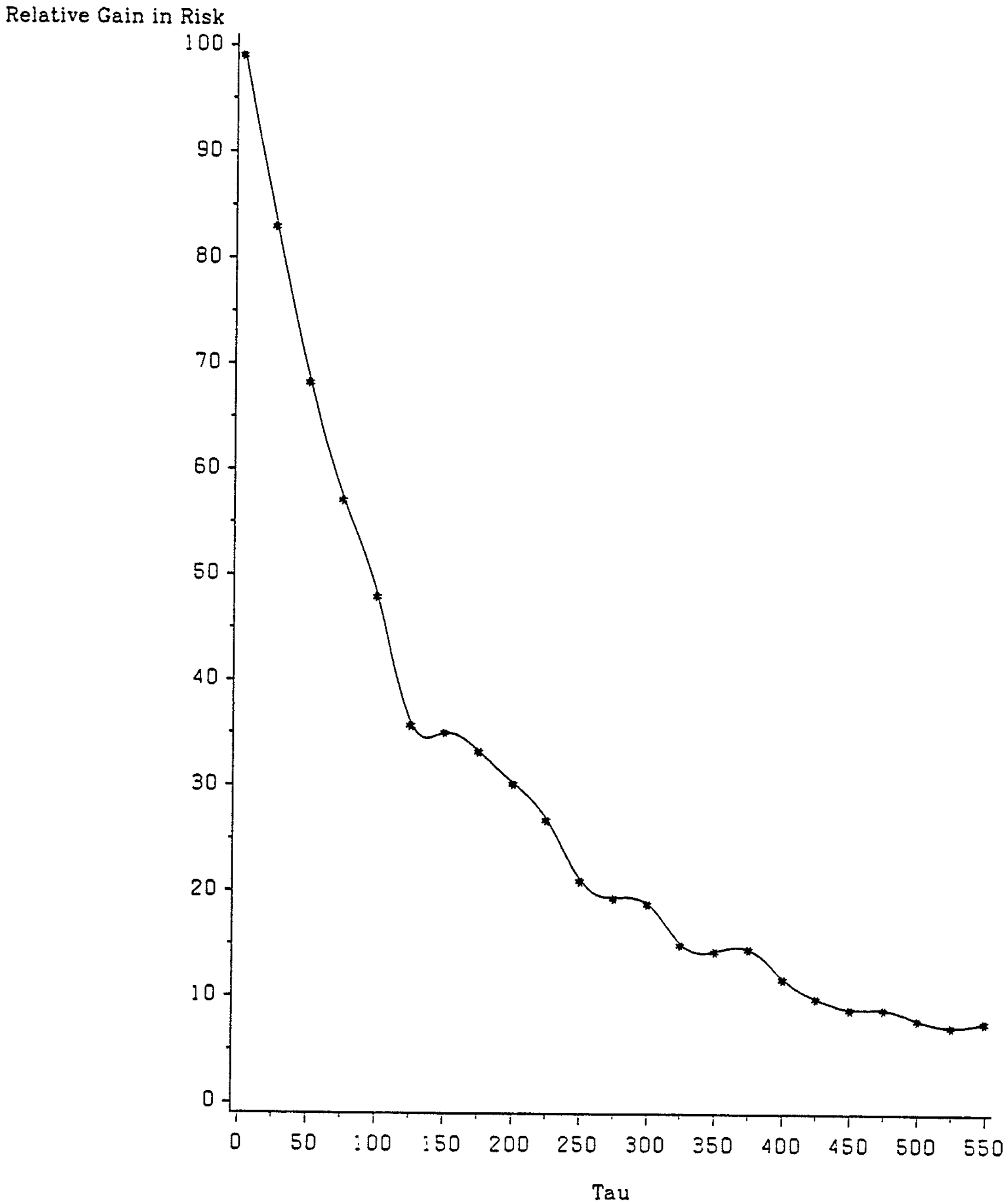
A Comparison of the Ensemble Risks of the MLE and the EMGB Estimator of a Vector of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

MLEAVE=1/3 MLRIS1 + 2/3 MLRIS2
EMGBAVE=1/3 BARIS1+2/3 BARIS2
ALPHA=(0.1, 0.1, 0.8); BETA=(0.3, 0.3, 0.4)



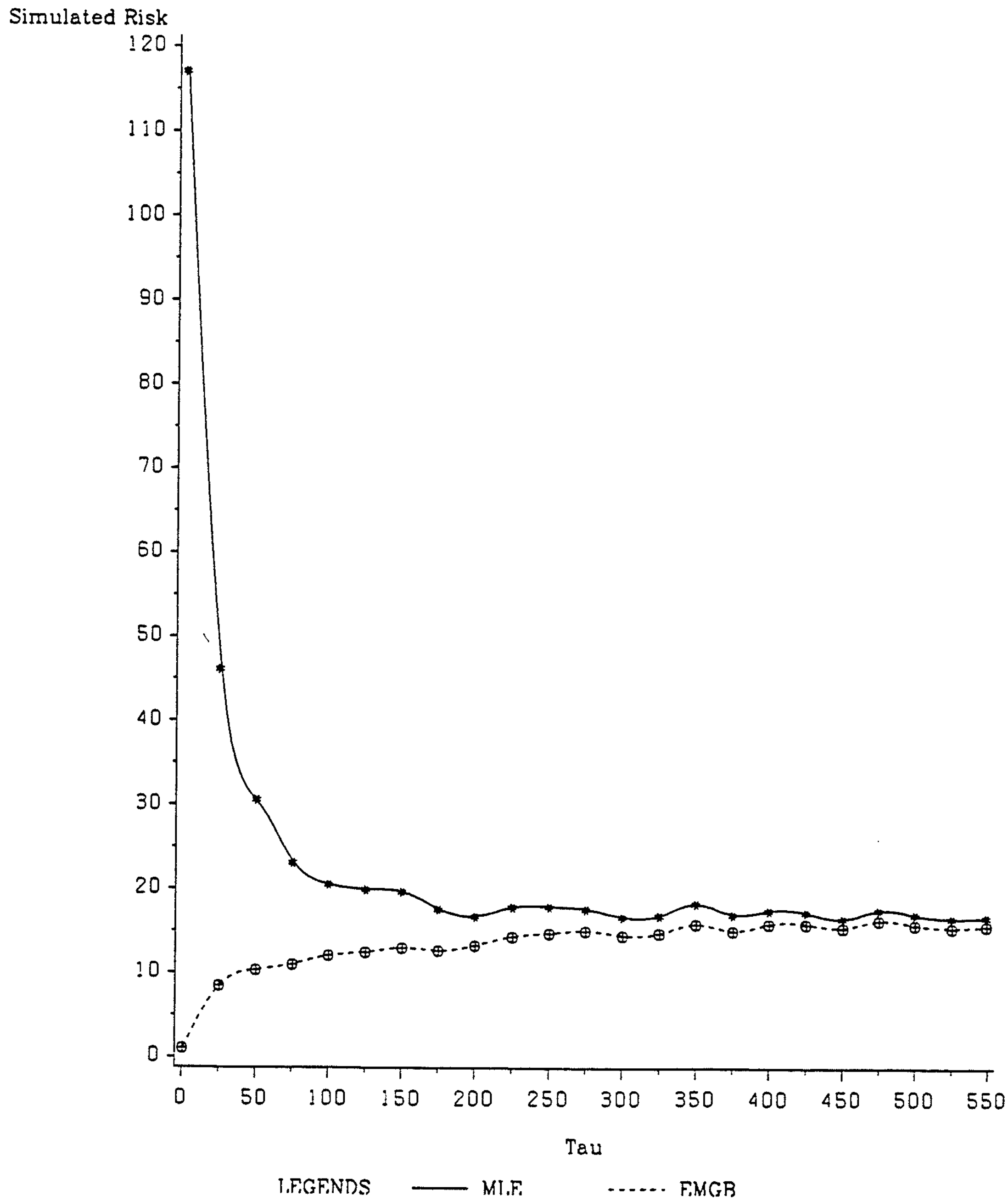
The Simulated Relative Gain in Ensemble Risk using the EMGB Estimator rather than the MLE of a Vector Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

ALPHA=(0.1, 0.1, 0.8); BETA=(0.3, 0.3, 0.4)



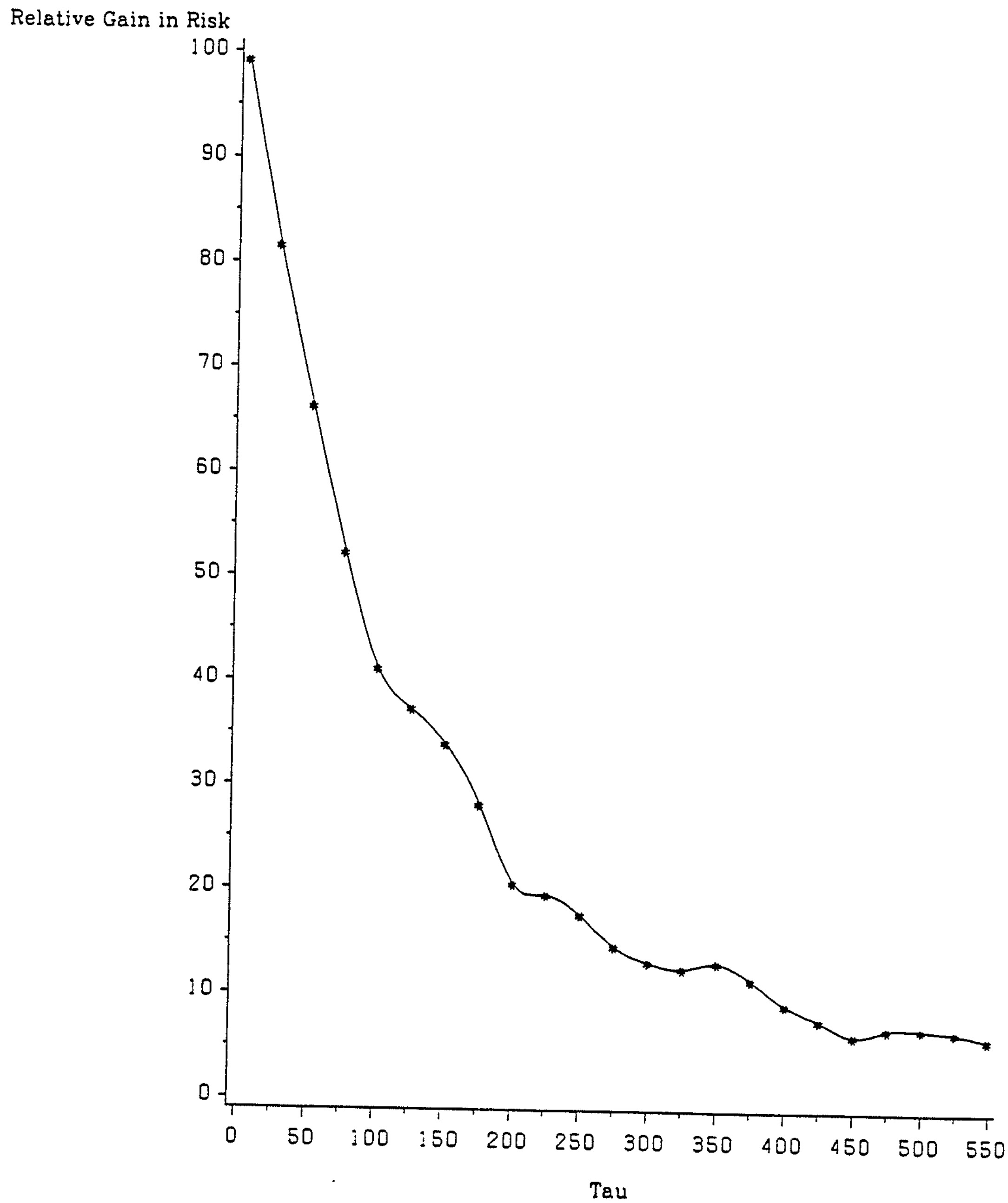
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Vector of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1.
ALPHA=(0.1, 0.1, 0.8); BETA=(1/3, 1/3, 1/3)



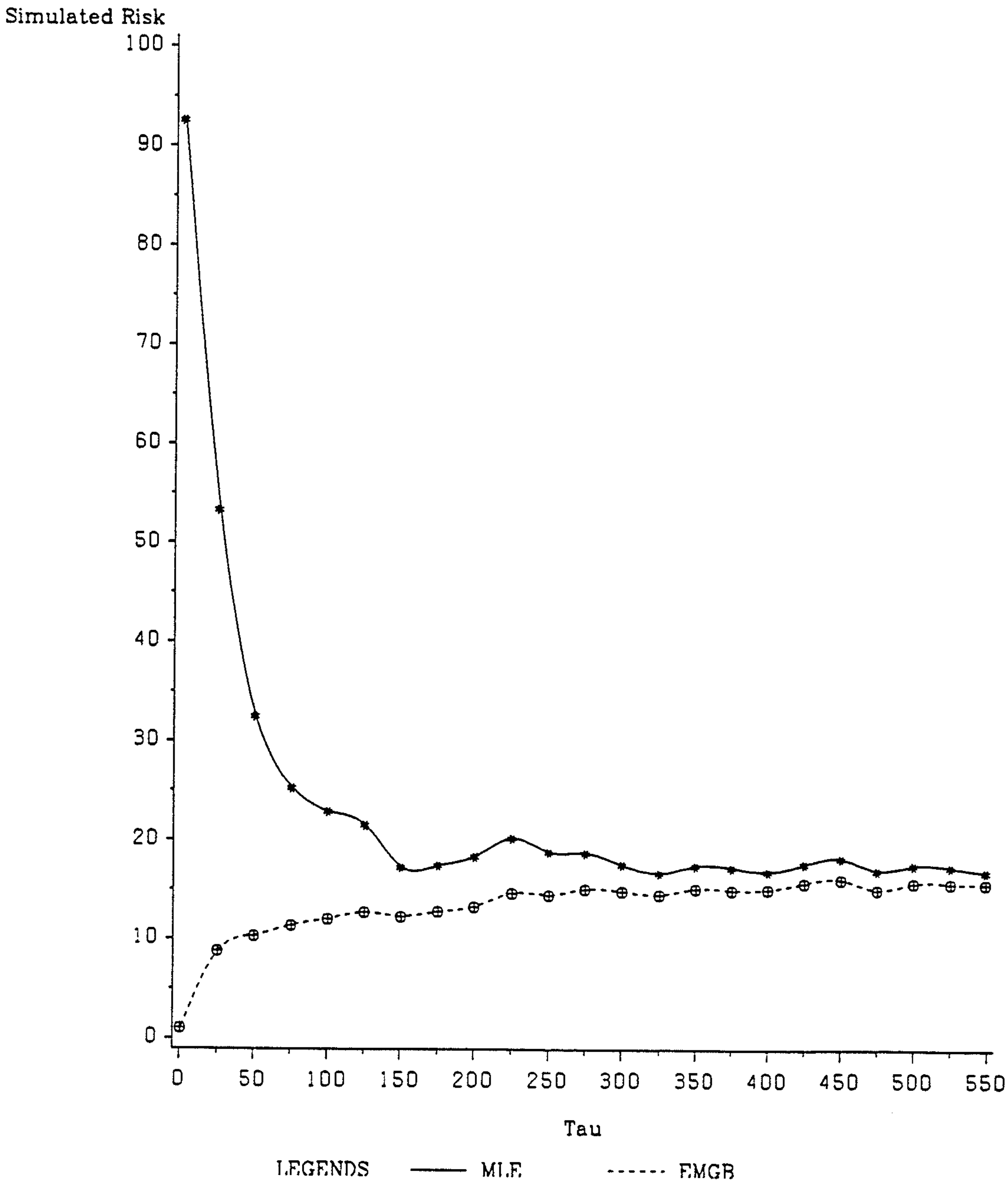
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1
ALPHA=(0.1, 0.1, 0.8);BETA=(1/3, 1/3, 1/3)



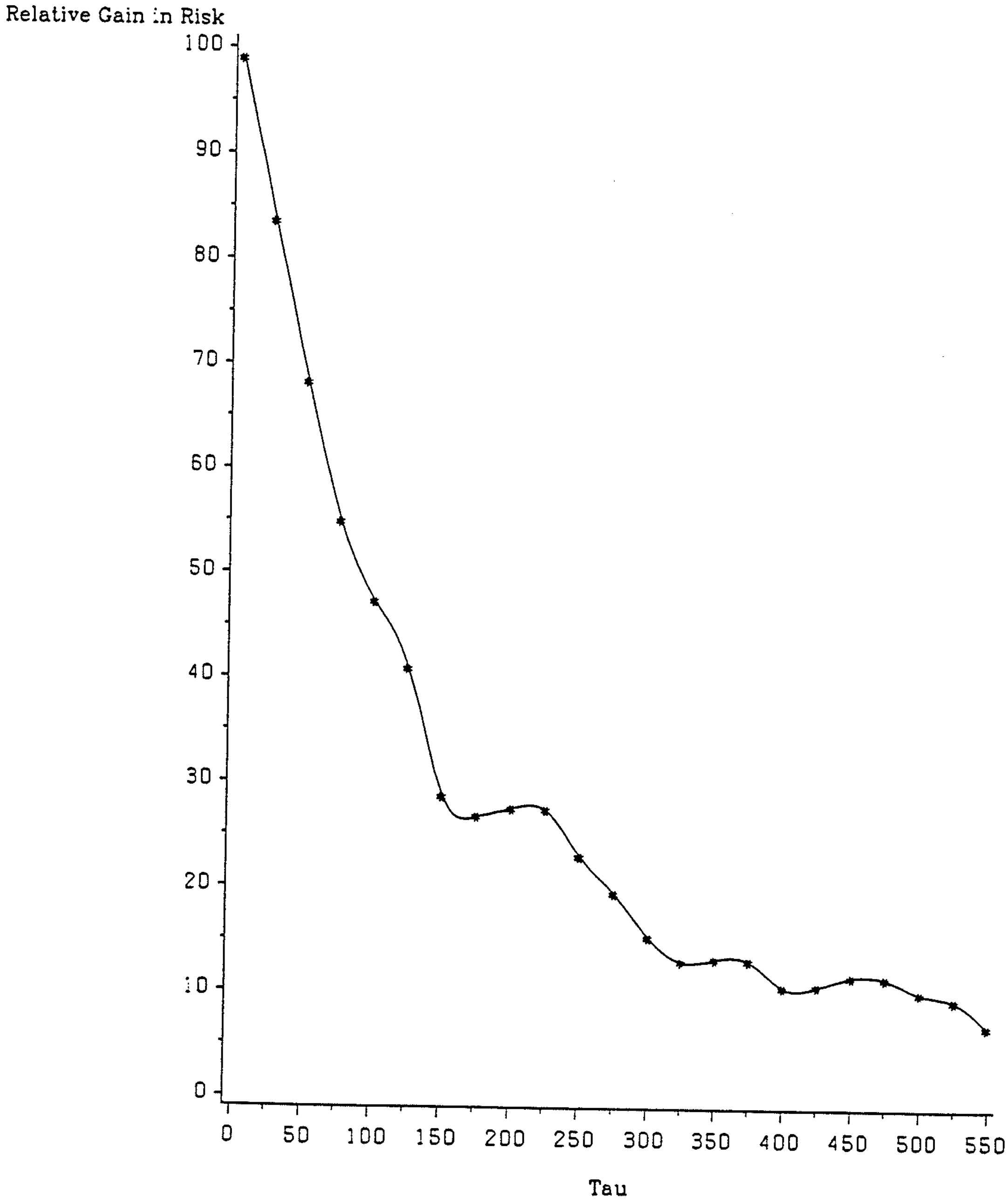
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Vector of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table. Incidence Matrix is IM2.

ALPHA=(0.1, 0.1, 0.8); BETA=(1/3, 1/3, 1/3)



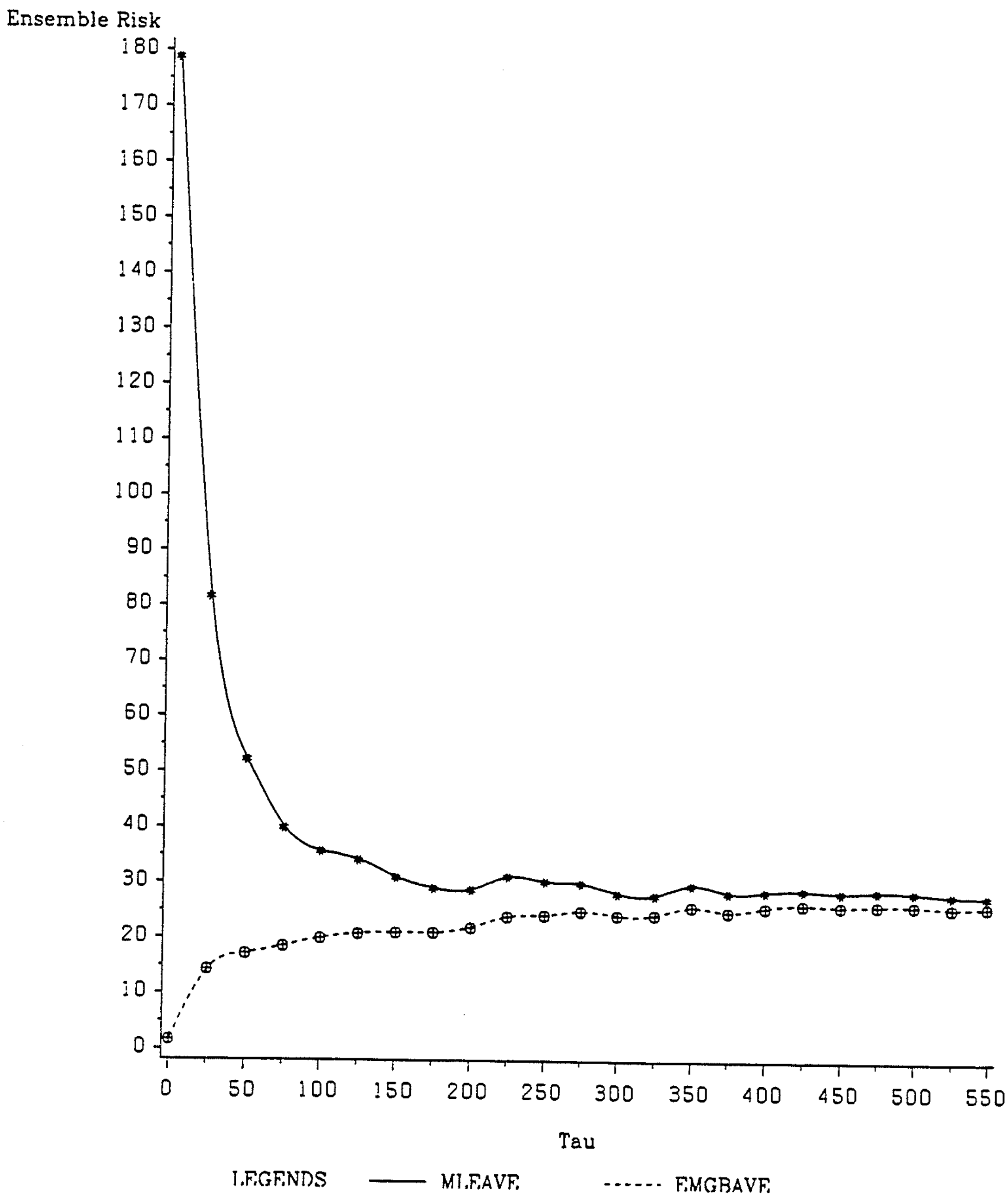
The Simulated Relative Gain in Risk using the
EMGB Estimator rather than the MLE of a Matrix
of Poisson Means under a Multiplicative Model
in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM2
ALPHA=(0.1, 0.1, 0.8);BETA=(1/3, 1/3, 1/3)



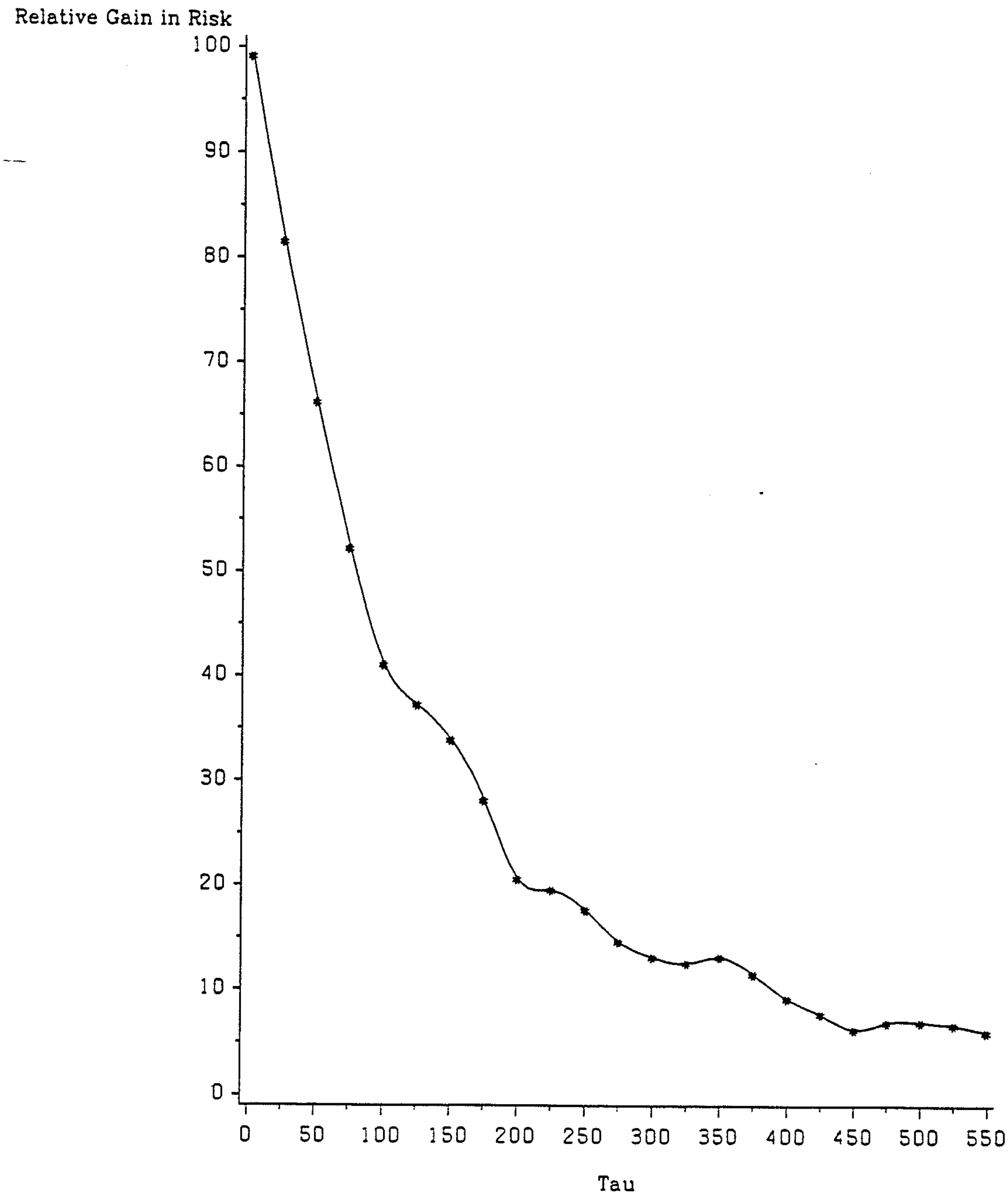
A Comparison of the Ensemble Risks of the MLE and the EMGB Estimator of a Vector of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

MLEAVE=1/3 MLRIS1 + 2/3 MLRIS2
EMGBAVE=1/3 BARIS1+2/3 BARIS2
ALPHA=(0.1, 0.1, 0.8); BETA=(1/3, 1/3, 1/3)



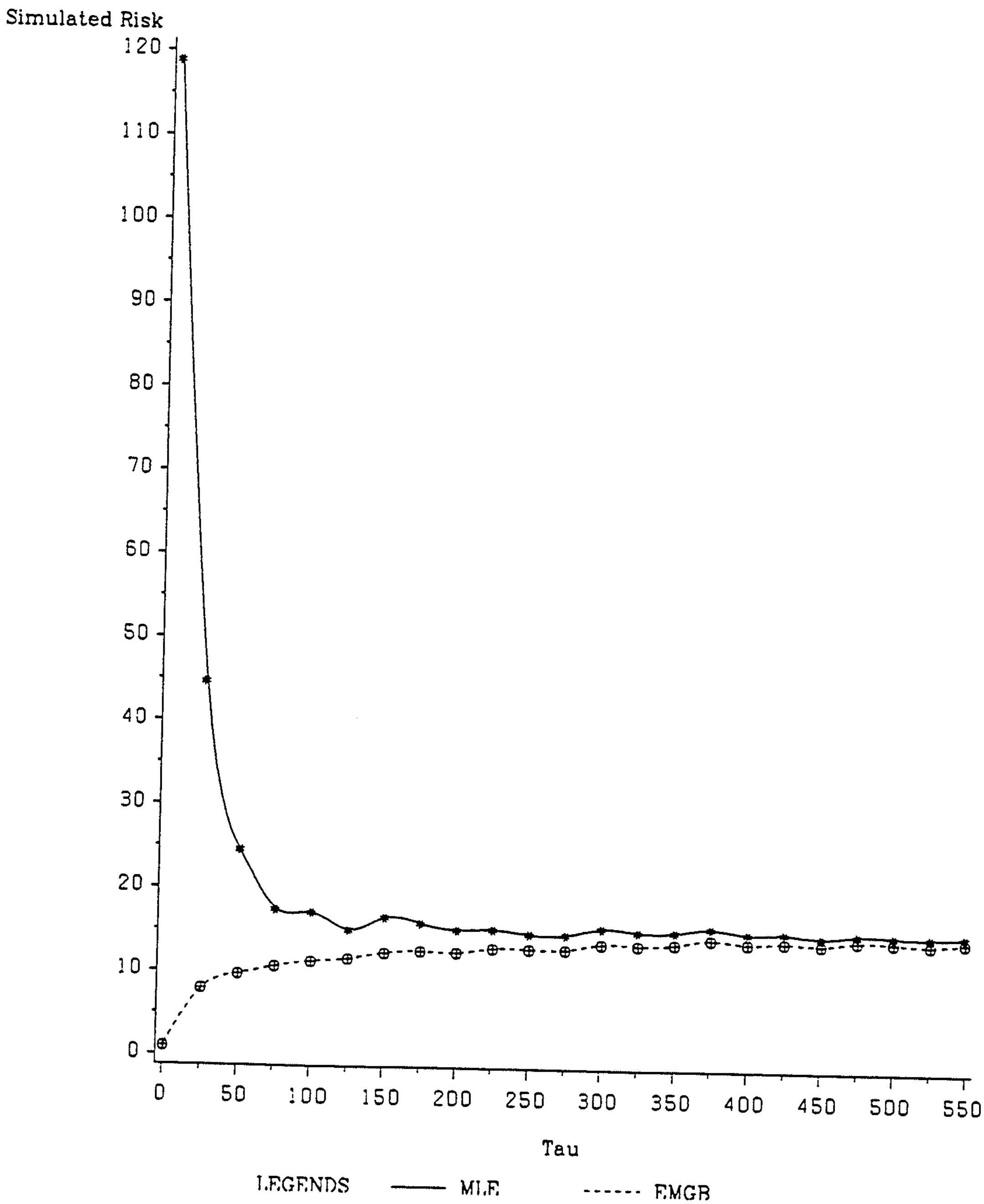
The Simulated Relative Gain in Ensemble Risk using the EMGB Estimator rather than the MLE of a Vector Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

ALPHA=(0.1, 0.1, 0.8); BETA=(1/3, 1/3, 1/3)



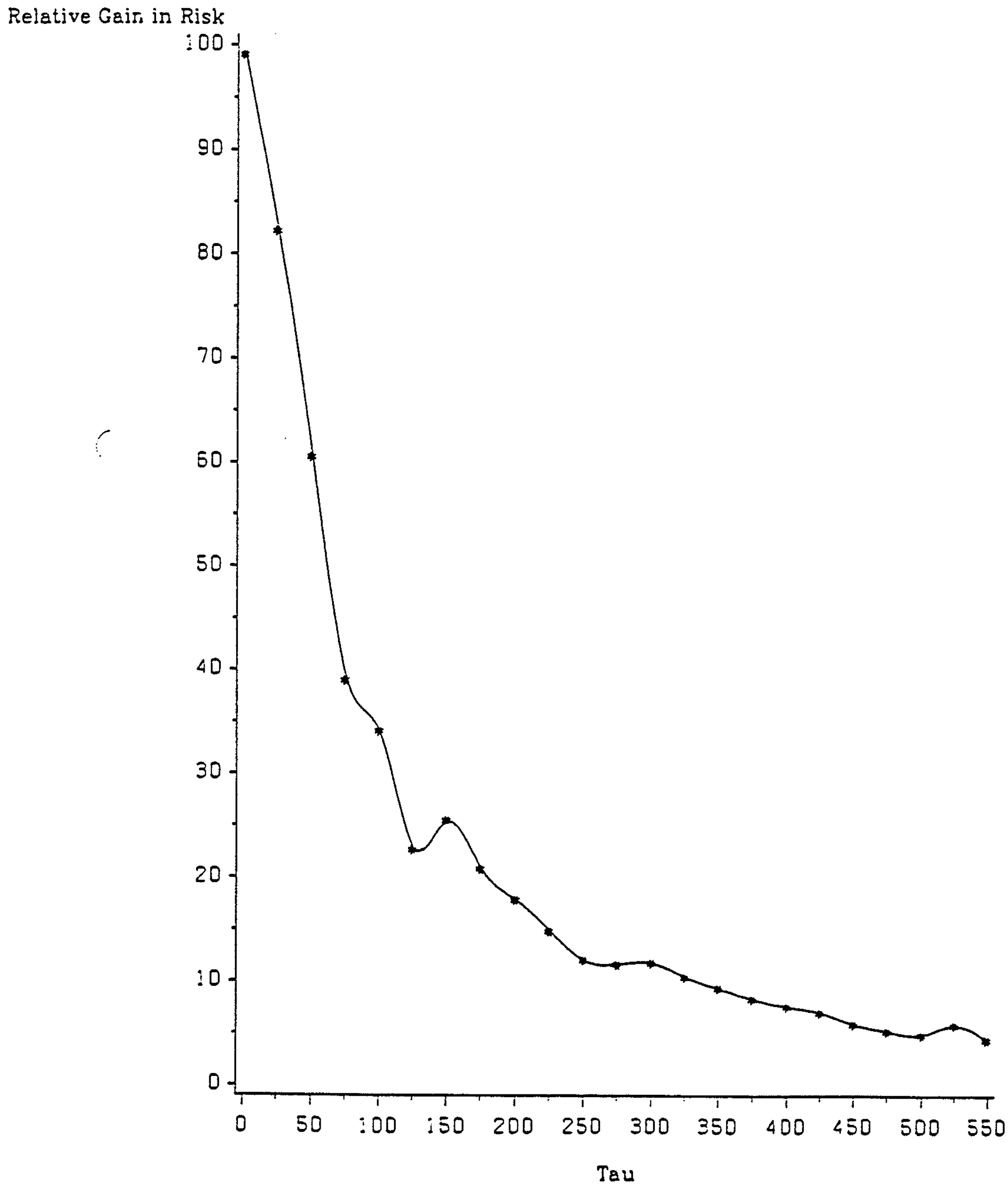
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Vector of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1.
ALPHA=(0.1, 0.1, 0.8); BETA=(0.4, 0.4, 0.2)



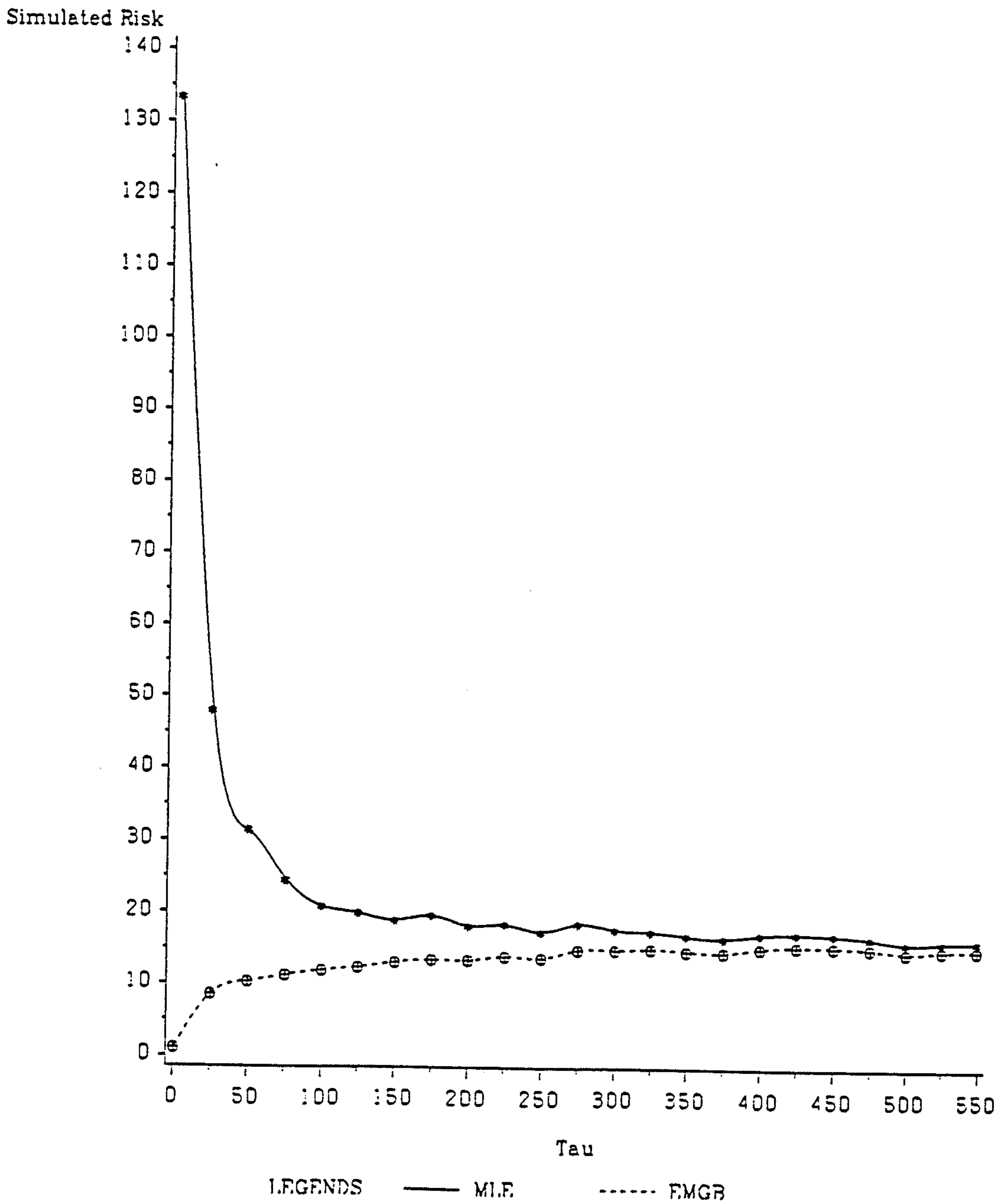
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1
ALPHA=(0.1, 0.1, 0.8);BETA=(0.4, 0.4, 0.2)



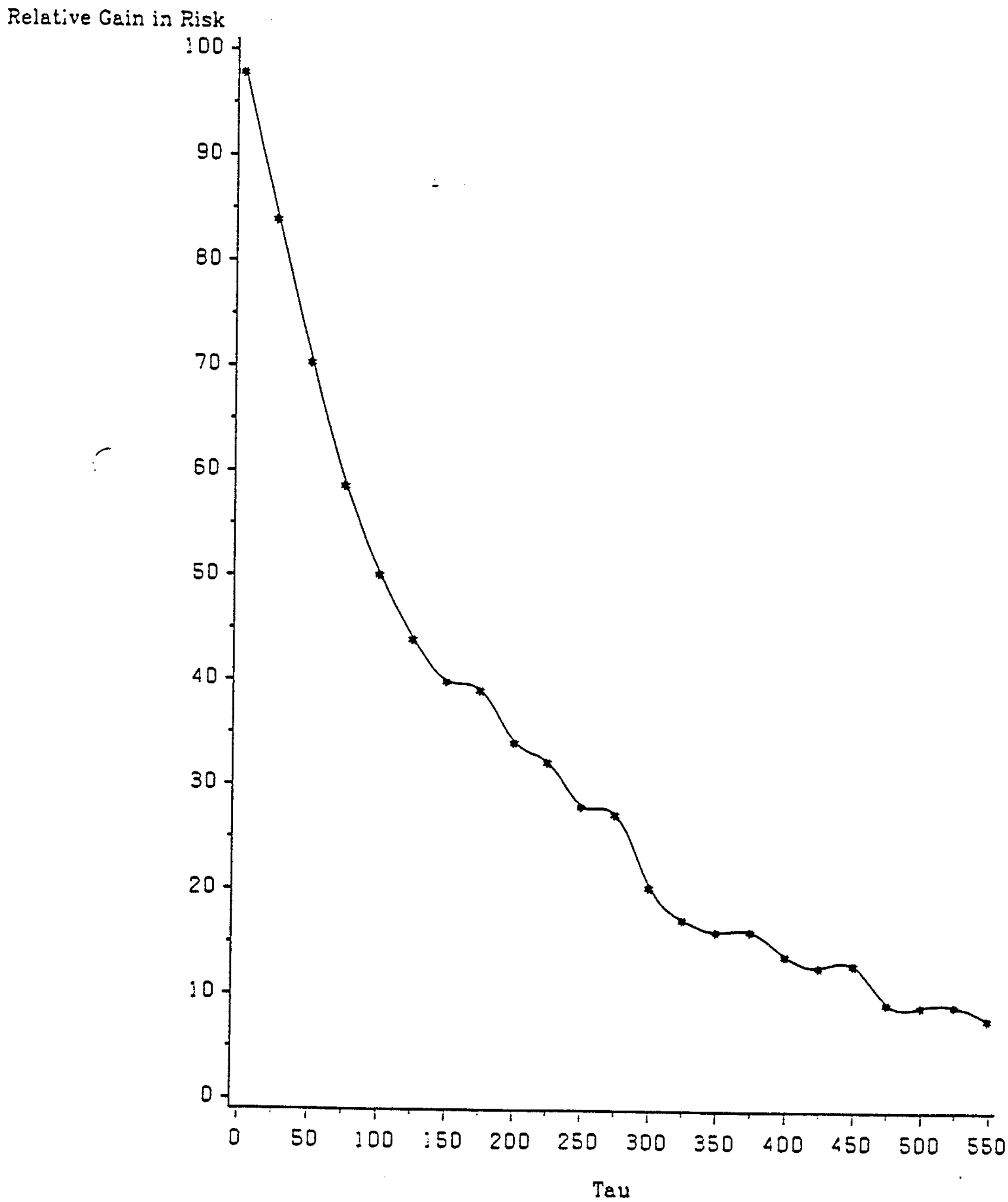
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Matrix of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.
Incidence Matrix is IM2.

ALPHA=(0.1, 0.1, 0.8); BETA=(0.4, 0.4, 0.2)



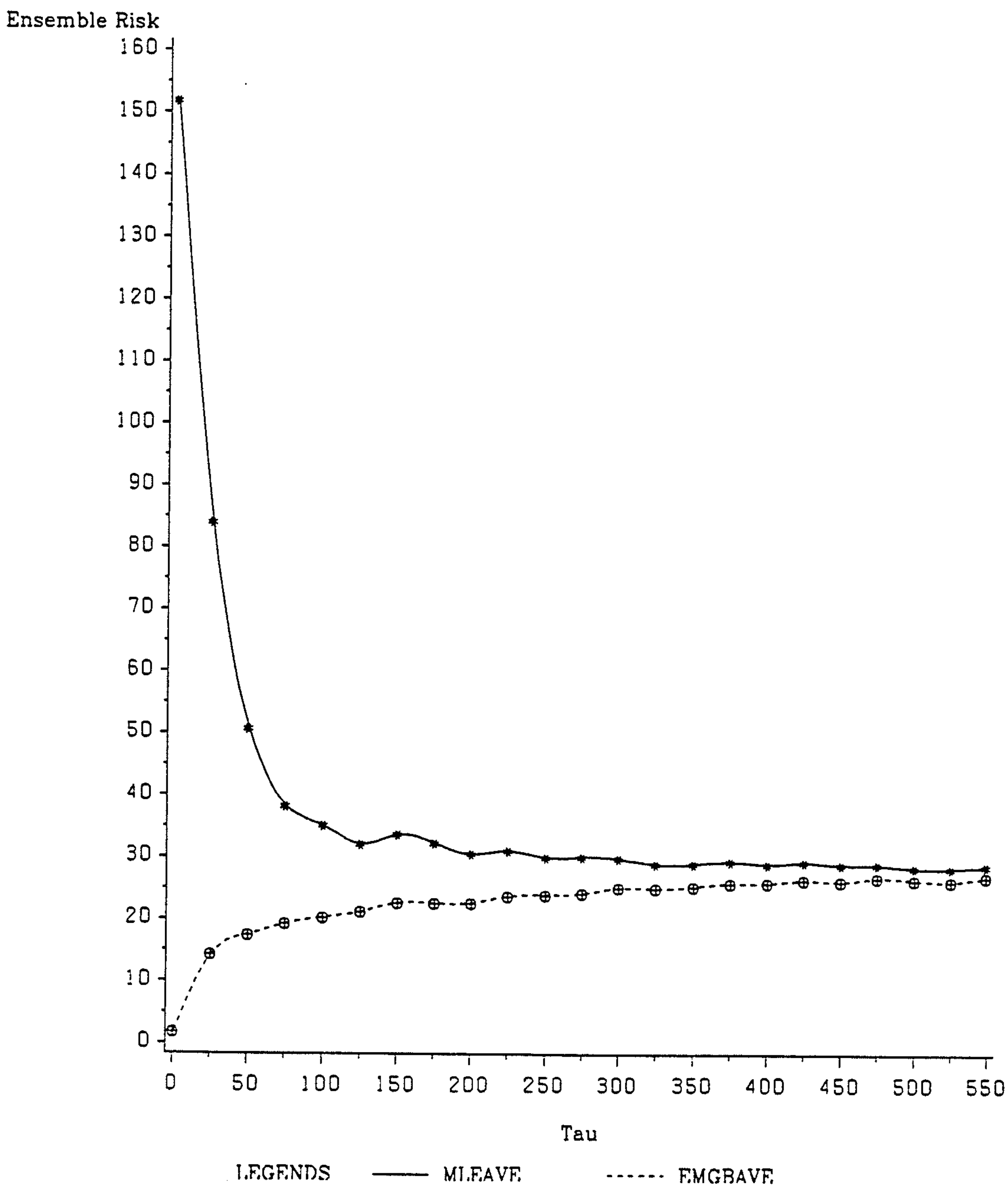
The Simulated Relative Gain in Risk using the
EMGB Estimator rather than the MLE of a Matrix
of Poisson Means under a Multiplicative Model
in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM2
ALPHA=(0.1, 0.1, 0.8);BETA=(0.4, 0.4, 0.2)



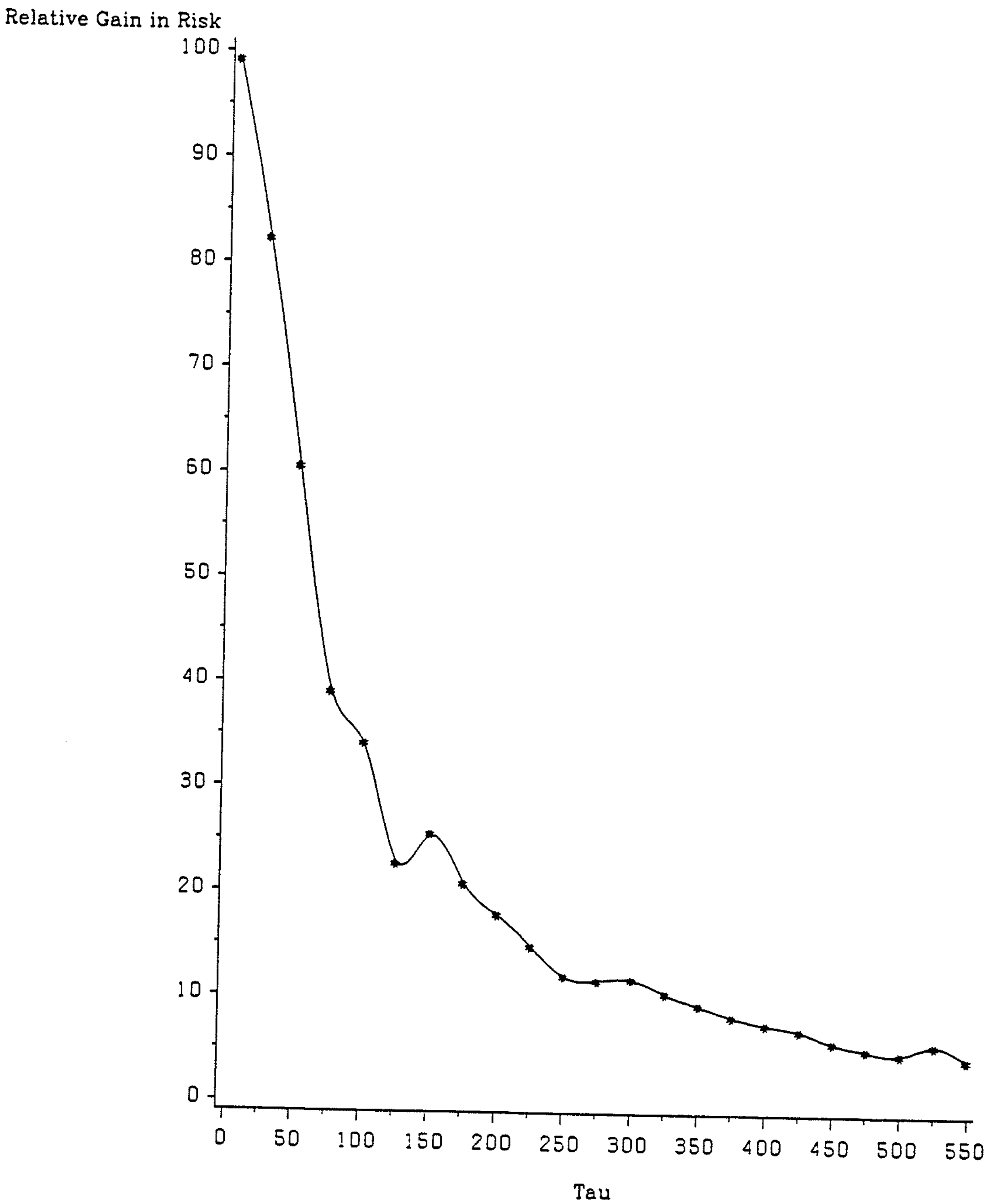
A Comparison of the Ensemble Risks of the MLE and the EMGB Estimator of a Vector of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

MLEAVE=1/3 MLRIS1 + 2/3 MLRIS2
EMGBAVE=1/3 BARIS1+2/3 BARIS2
ALPHA=(0.1, 0.1, 0.8); BETA=(0.4, 0.4, 0.2)



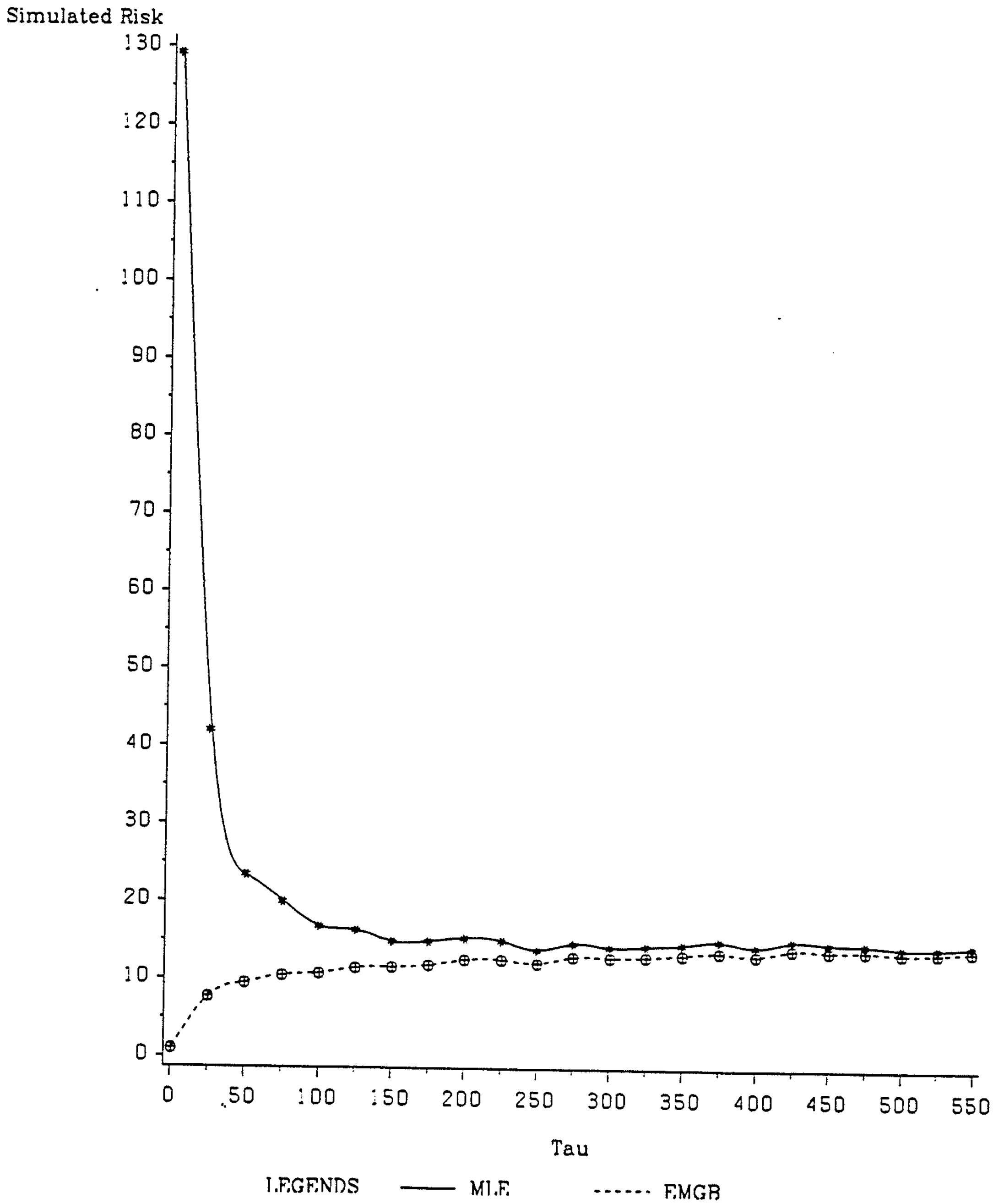
The Simulated Relative Gain in Ensemble Risk using the EMGB Estimator rather than the MLE of a Vector Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

ALPHA=(0.1, 0.1, 0.8); BETA=(0.4, 0.4, 0.2)



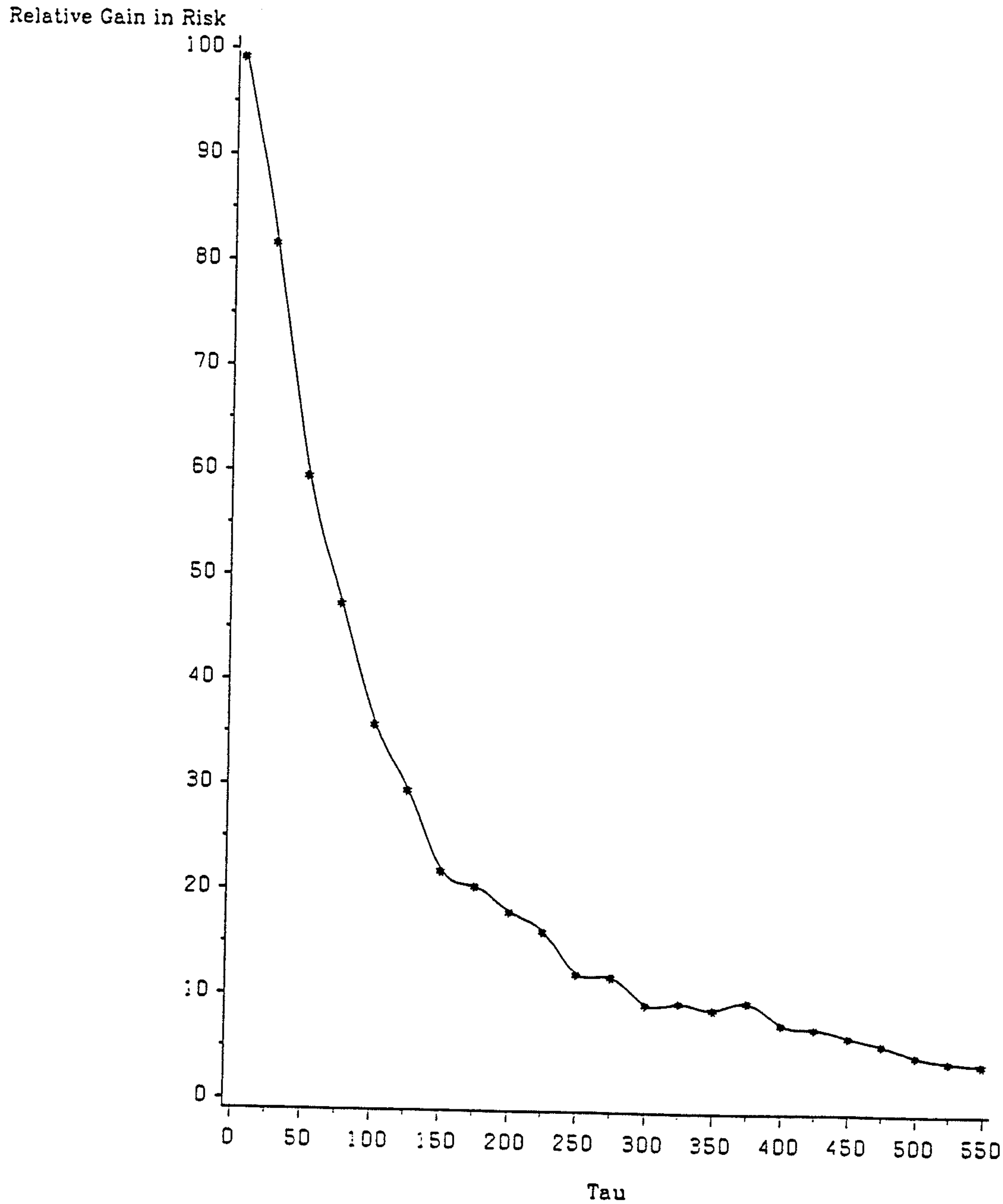
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Matrix of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1.
ALPHA=(0.2, 0.2, 0.6); BETA=(0.2, 0.2, 0.6)



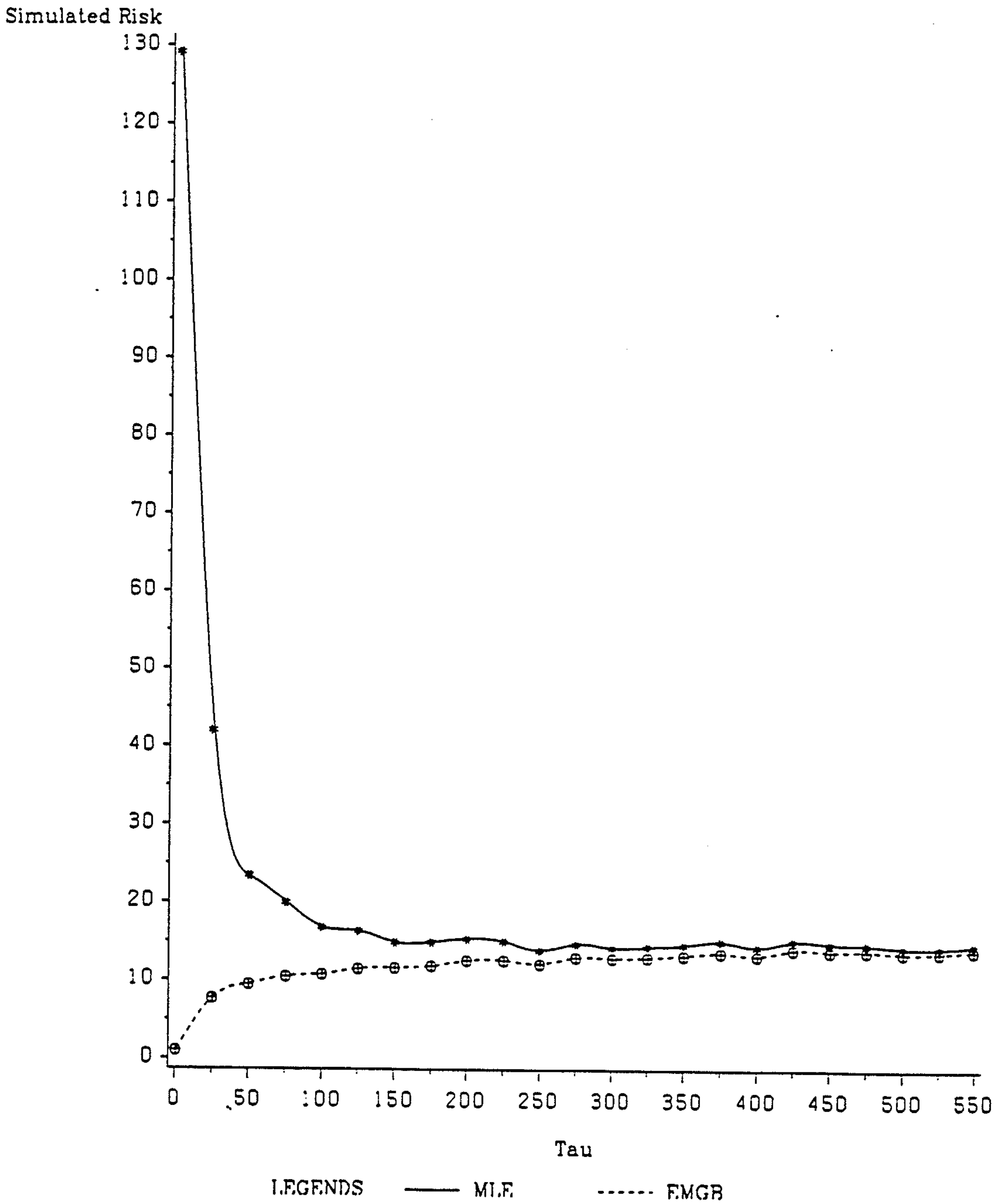
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1
ALPHA=(0.2, 0.2, 0.6);BETA=(0.2, 0.2, 0.6)



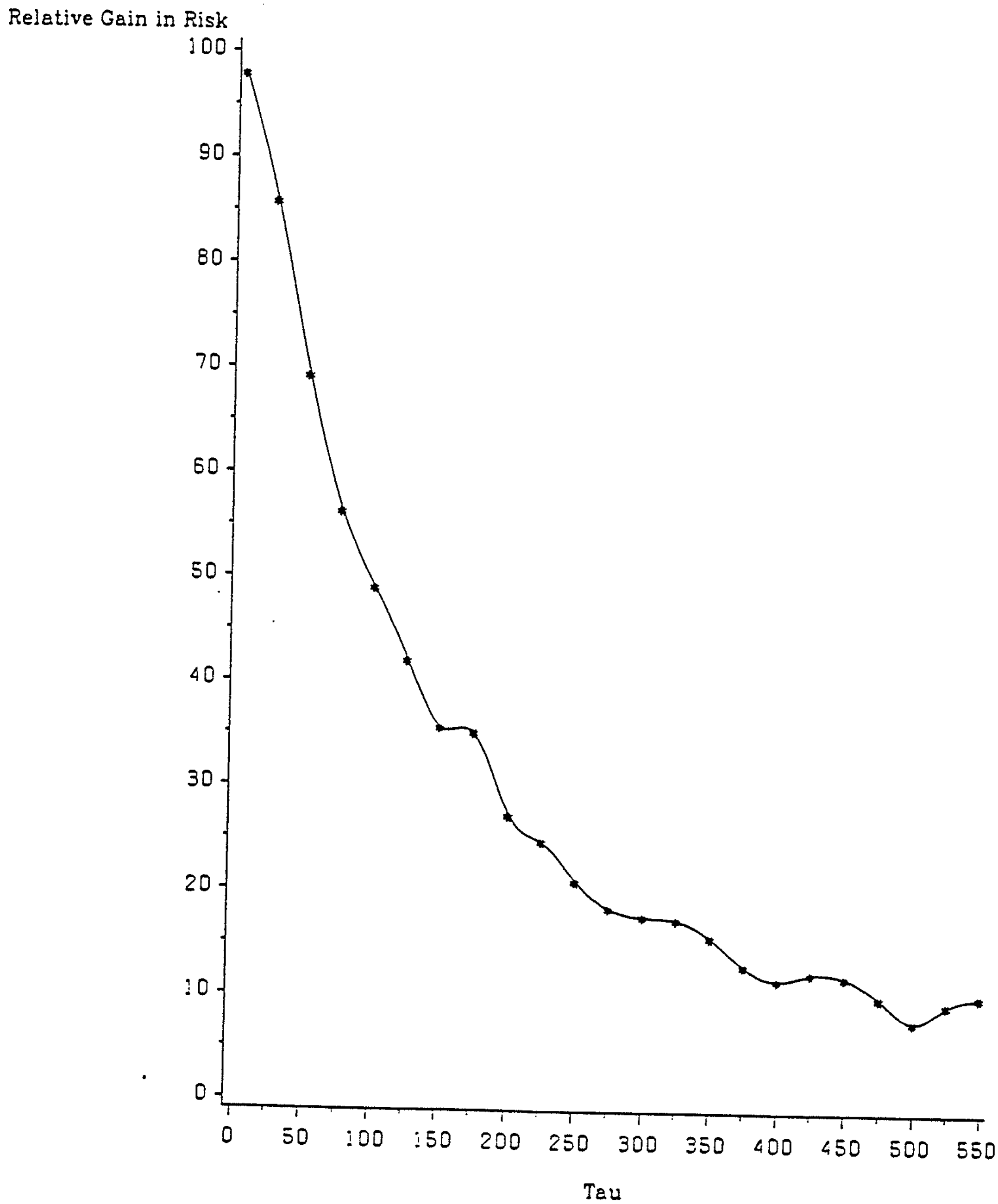
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Matrix of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1.
ALPHA=(0.2, 0.2, 0.6); BETA=(0.2, 0.2, 0.6)



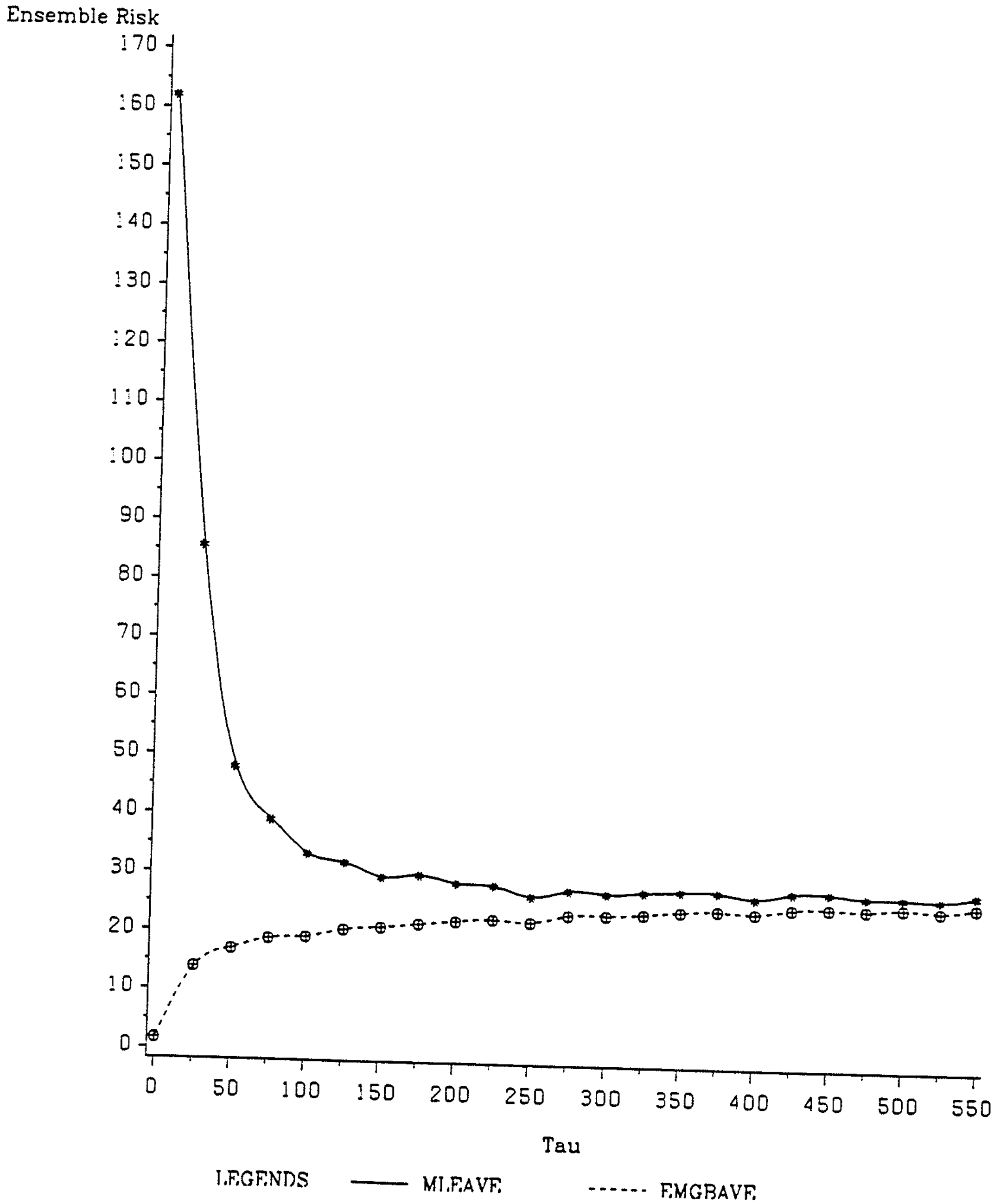
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM2
ALPHA=(0.2, 0.2, 0.6);BETA=(0.2, 0.2, 0.6)



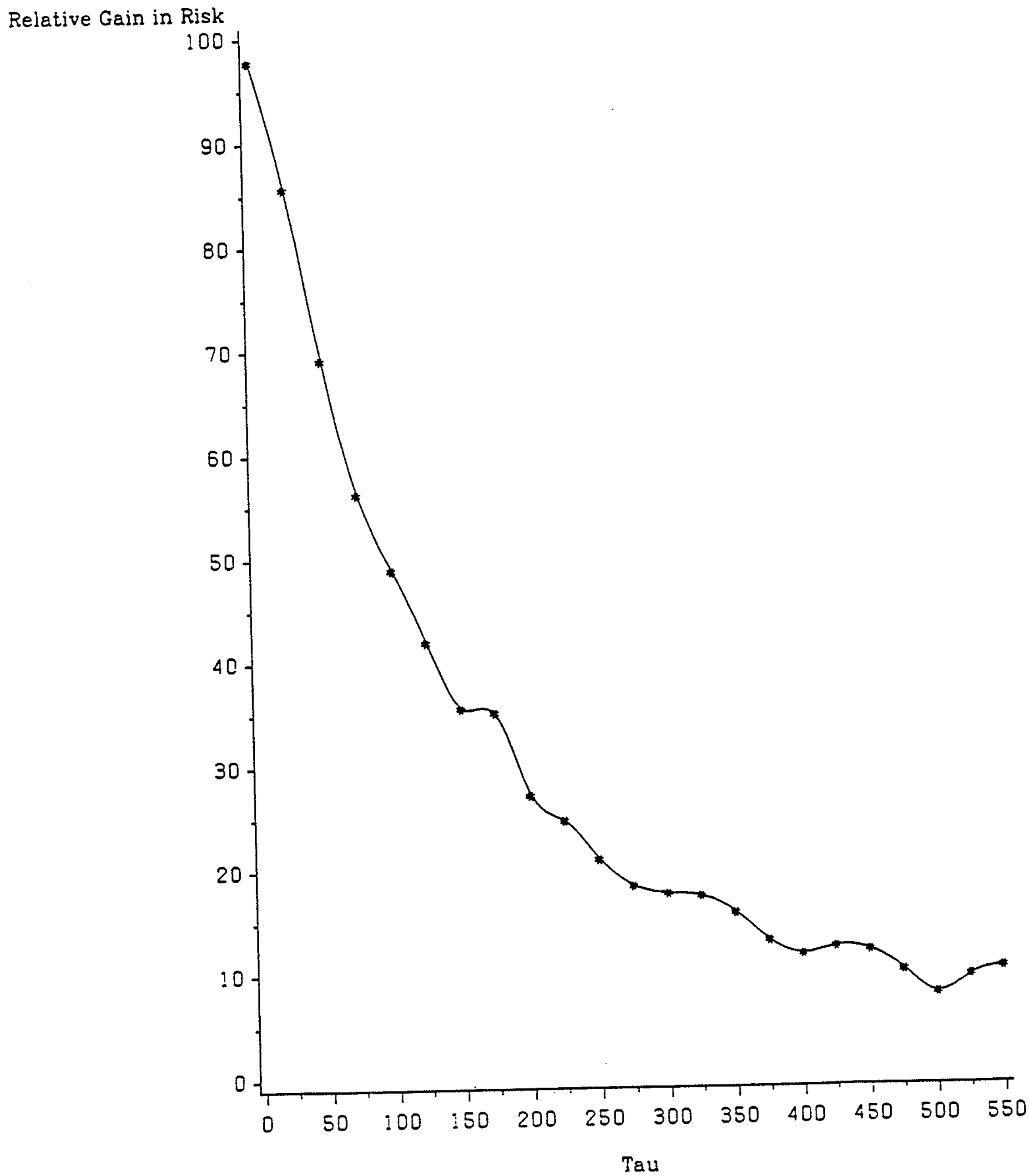
A Comparison of the Ensemble Risks of the MLE and the EMGB Estimator of a Matrix of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

MLEAVE=1/3 MLRIS1 + 2/3 MLRIS2
EMGBAVE=1/3 BARIS1+2/3 BARIS2
ALPHA=(0.2, 0.2, 0.6);BETA=(0.2, 0.2, 0.6)



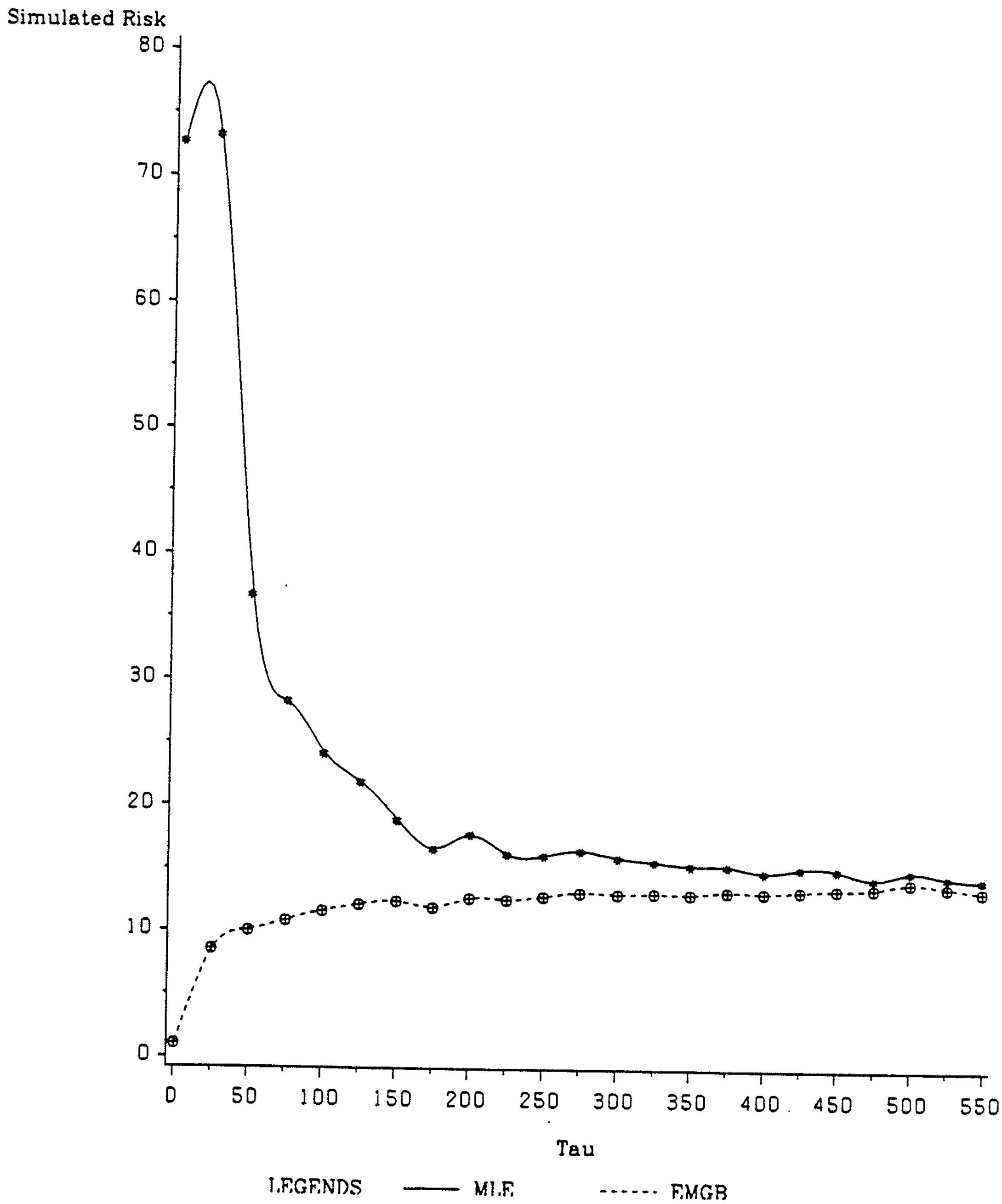
The Simulated Relative Gain in Ensemble Risk using the EMGB Estimator rather than the MLE of a Vector Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

ALPHA=(0.2, 0.2, 0.6); BETA=(0.2, 0.2, 0.6)



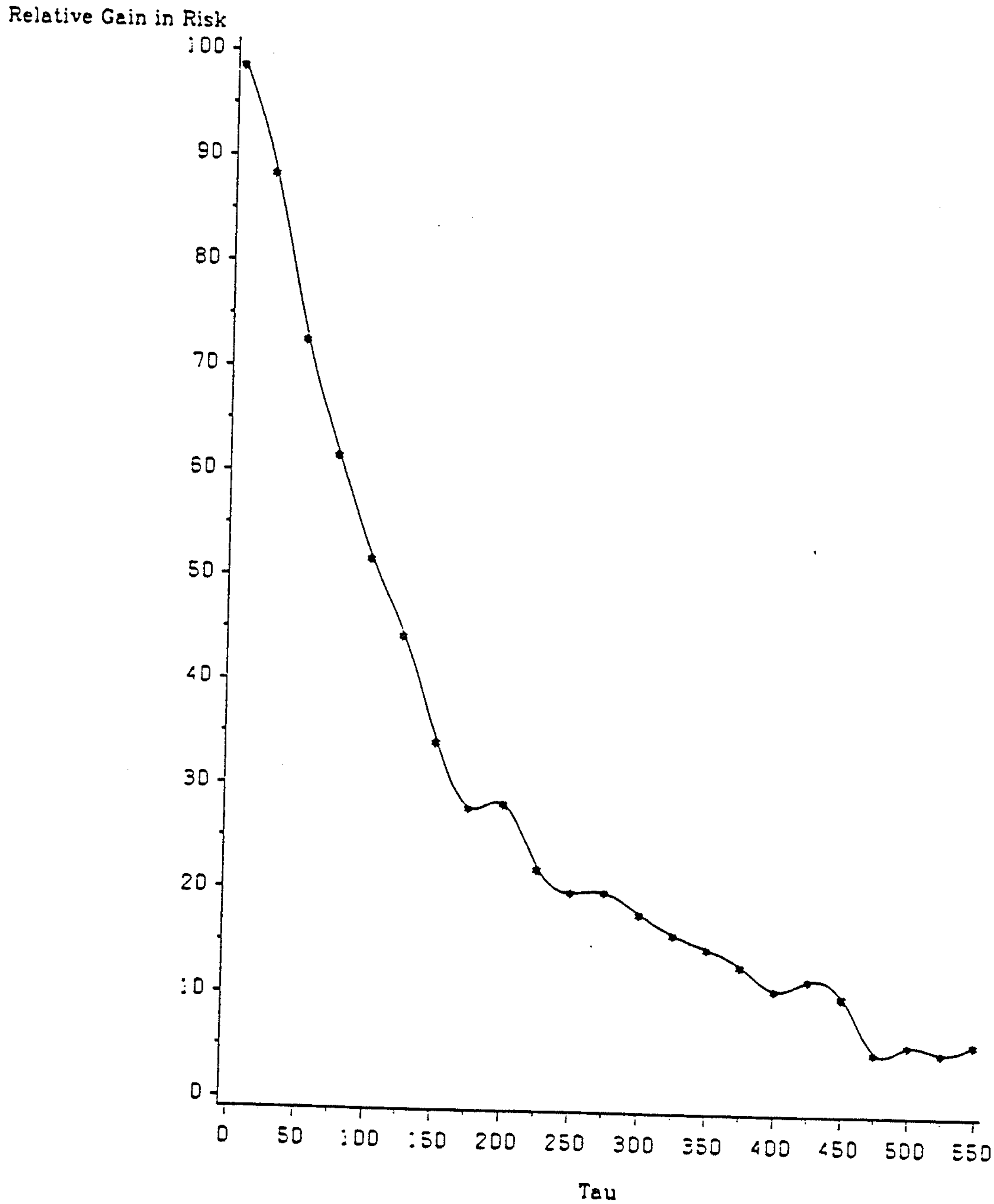
A Comparison of the Simulated Risks of the
MLE and the EMGB Estimator of a Vector of
Poisson Means under a Multiplicative Model
for a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1.
ALPHA=(0.2, 0.2, 0.6); BETA=(0.3, 0.3, 0.4)



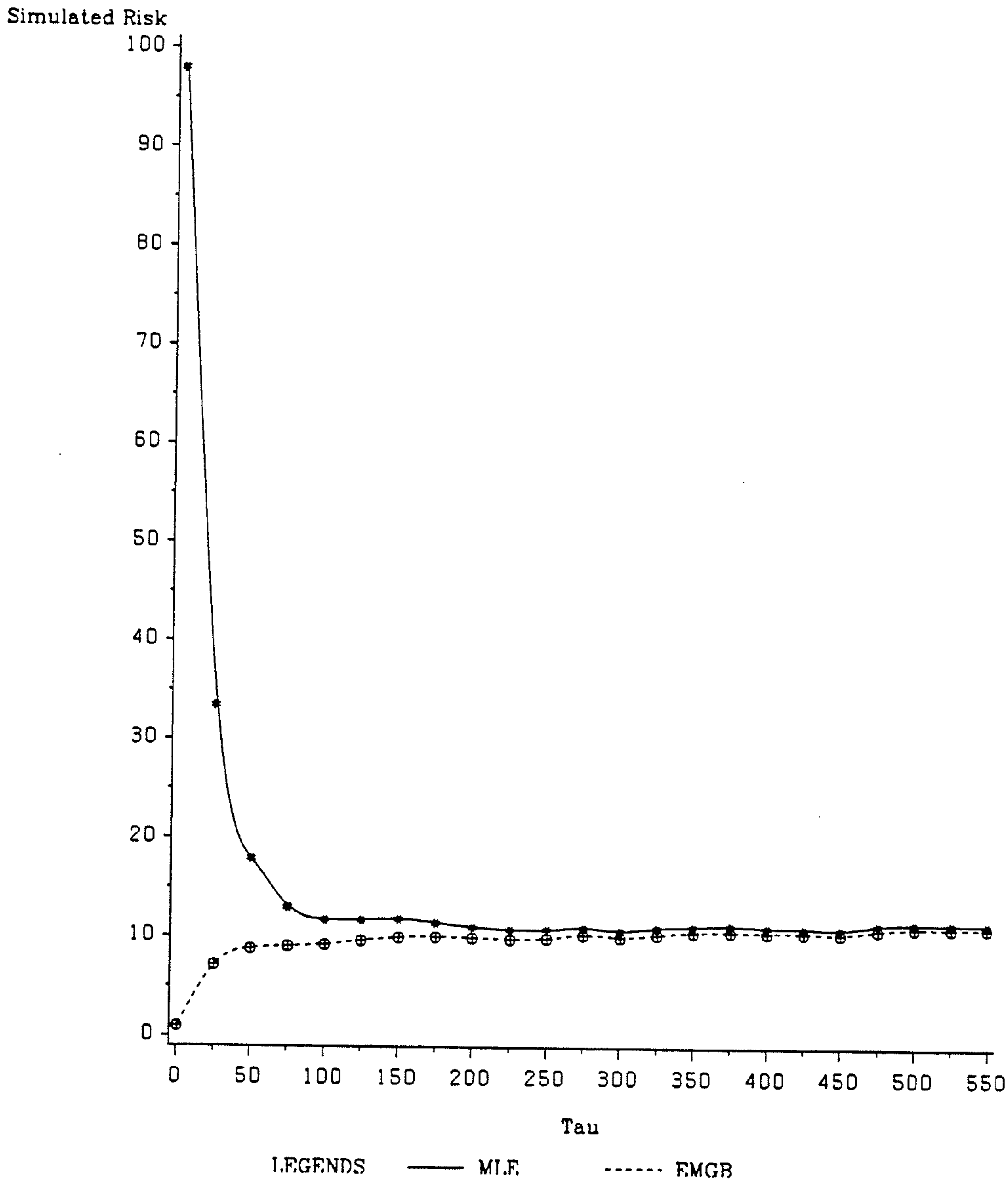
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1
ALPHA=(0.2, 0.2, 0.6);BETA=(0.3, 0.3, 0.4)



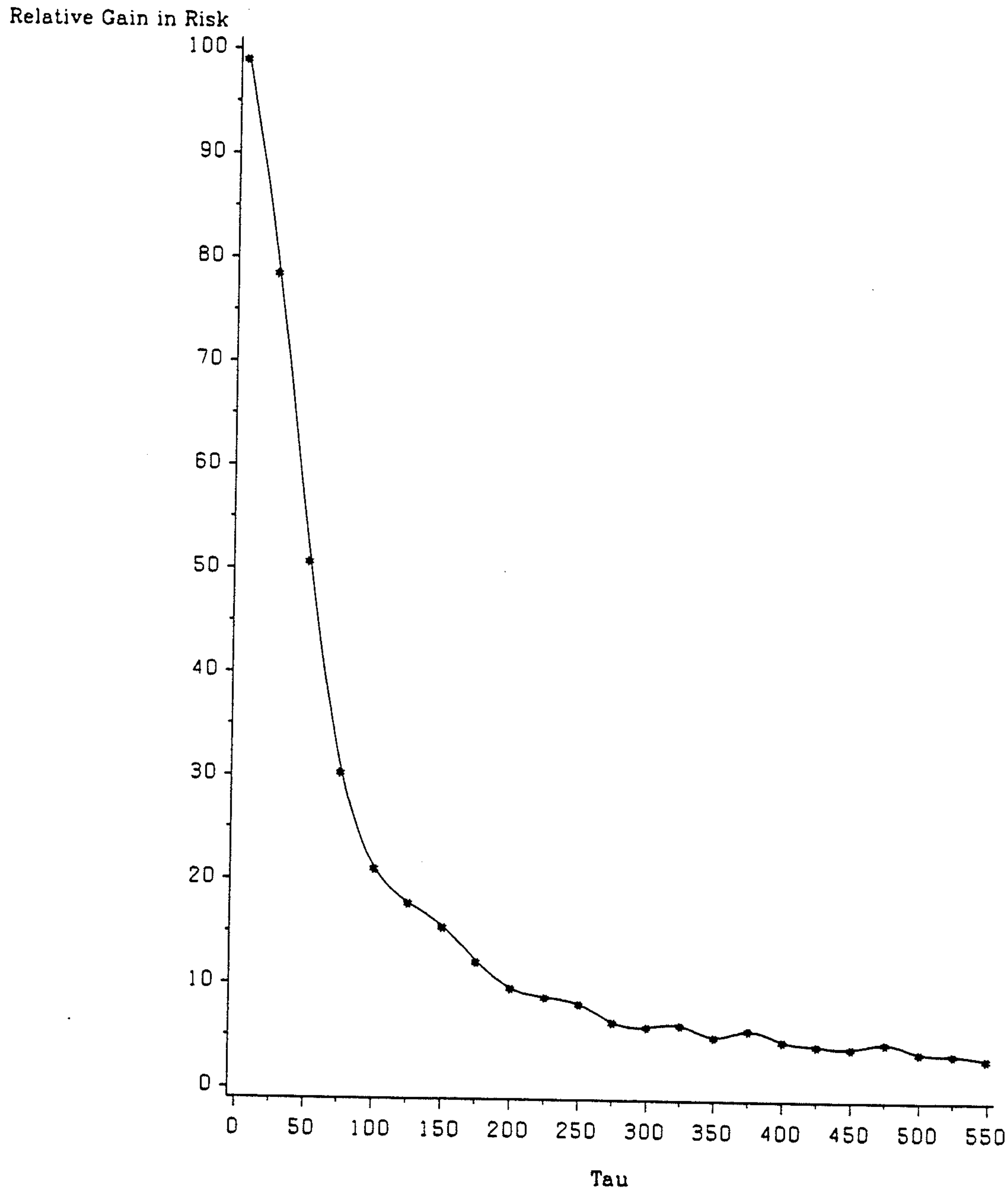
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Vector of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table. Incidence Matrix is IM2.

ALPHA=(0.2, 0.2, 0.6); BETA=(0.3, 0.3, 0.4)



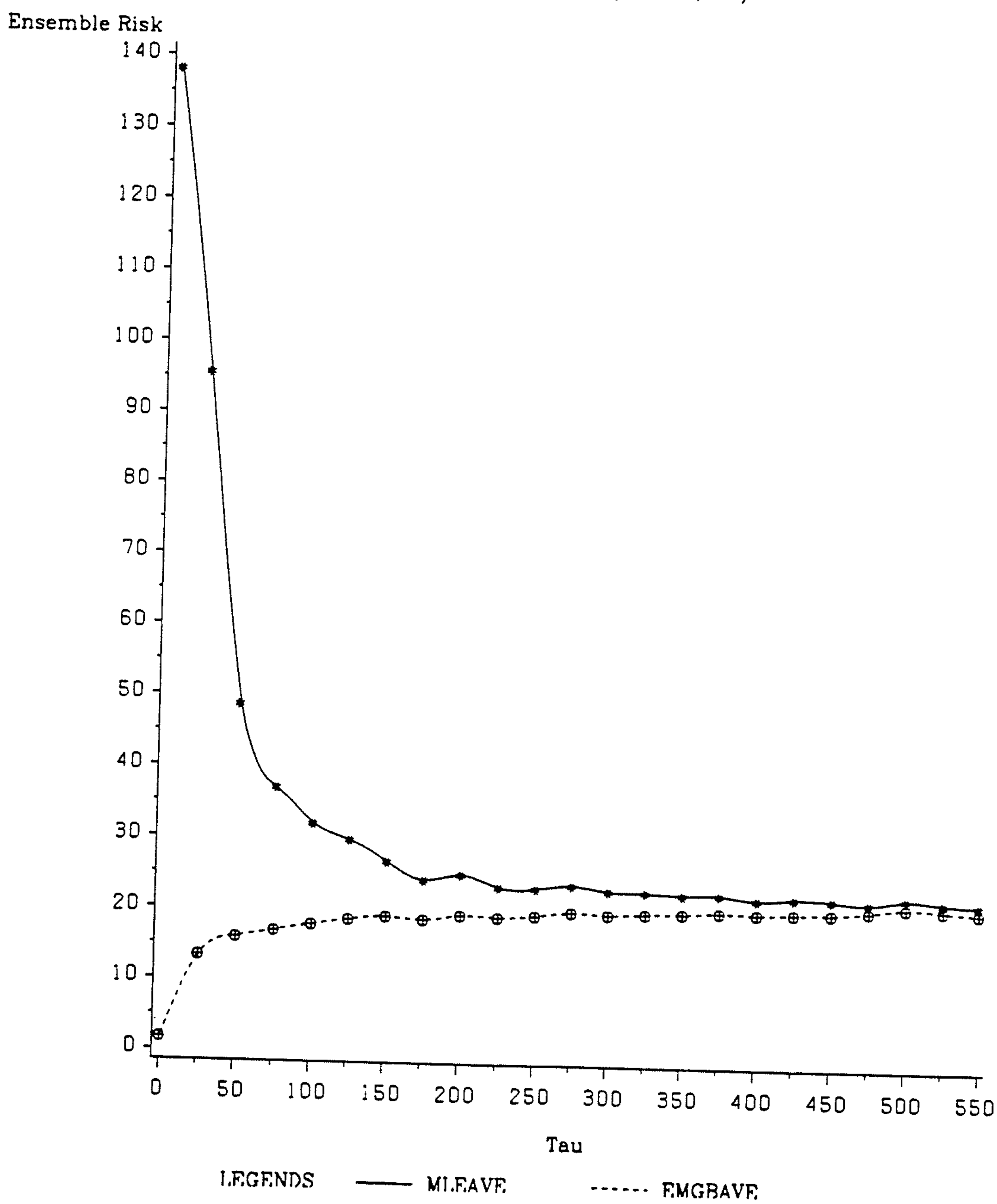
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM2
ALPHA=(0.2, 0.2, 0.6);BETA=(0.3, 0.3, 0.4)



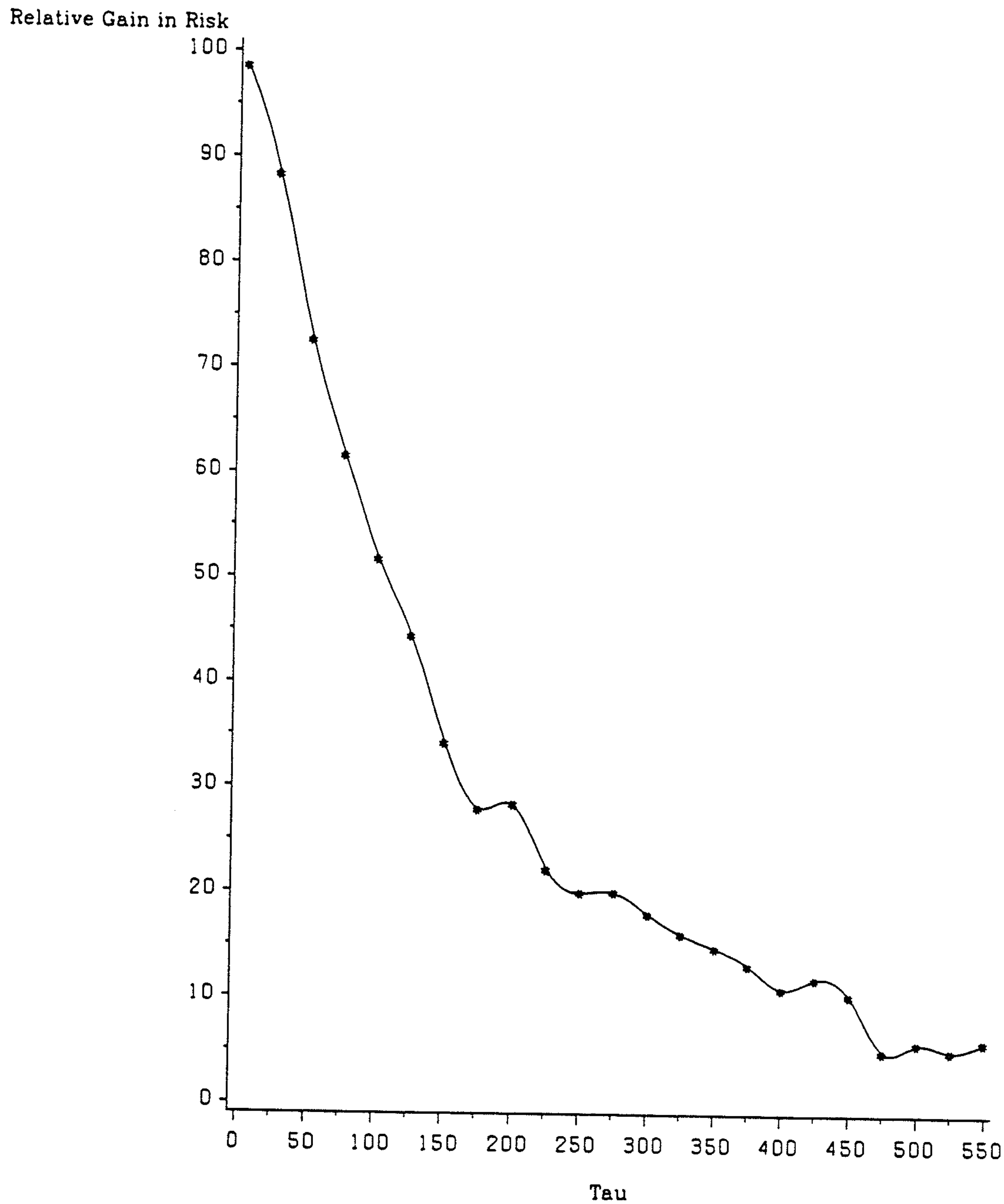
A Comparison of the Ensemble Risks of the MLE and the EMGB Estimator of a Matrix of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

MLEAVE=1/3 MLRIS1 + 2/3 MLRIS2
EMGBAVE=1/3 BARIS1+2/3 BARIS2
ALPHA=(0.2, 0.2, 0.6);BETA=(0.3, 0.3, 0.4)



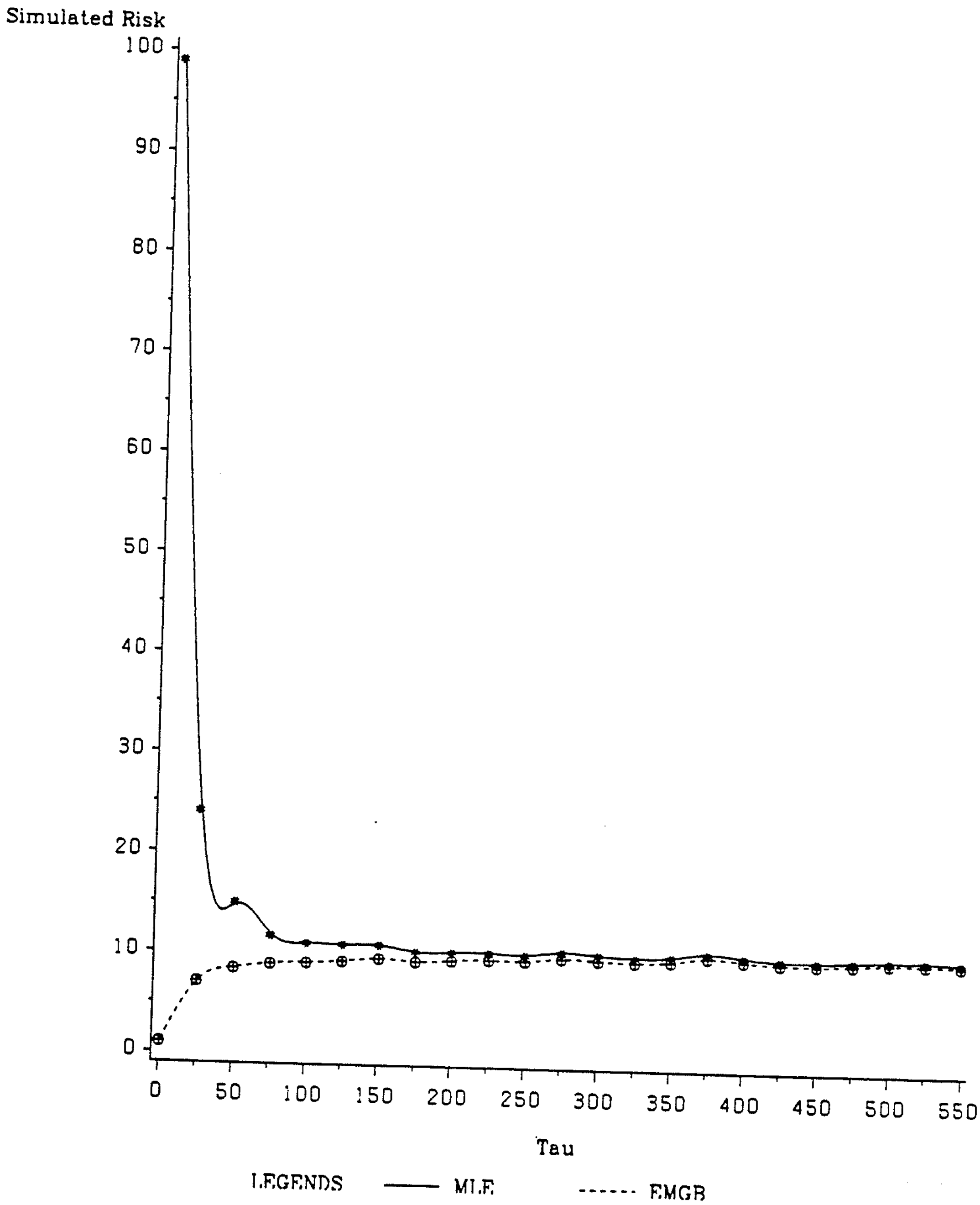
The Simulated Relative Gain in Ensemble Risk using the EMGB Estimator rather than the MLE of a Vector Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

ALPHA=(0.2, 0.2, 0.6); BETA=(0.3, 0.3, 0.4)



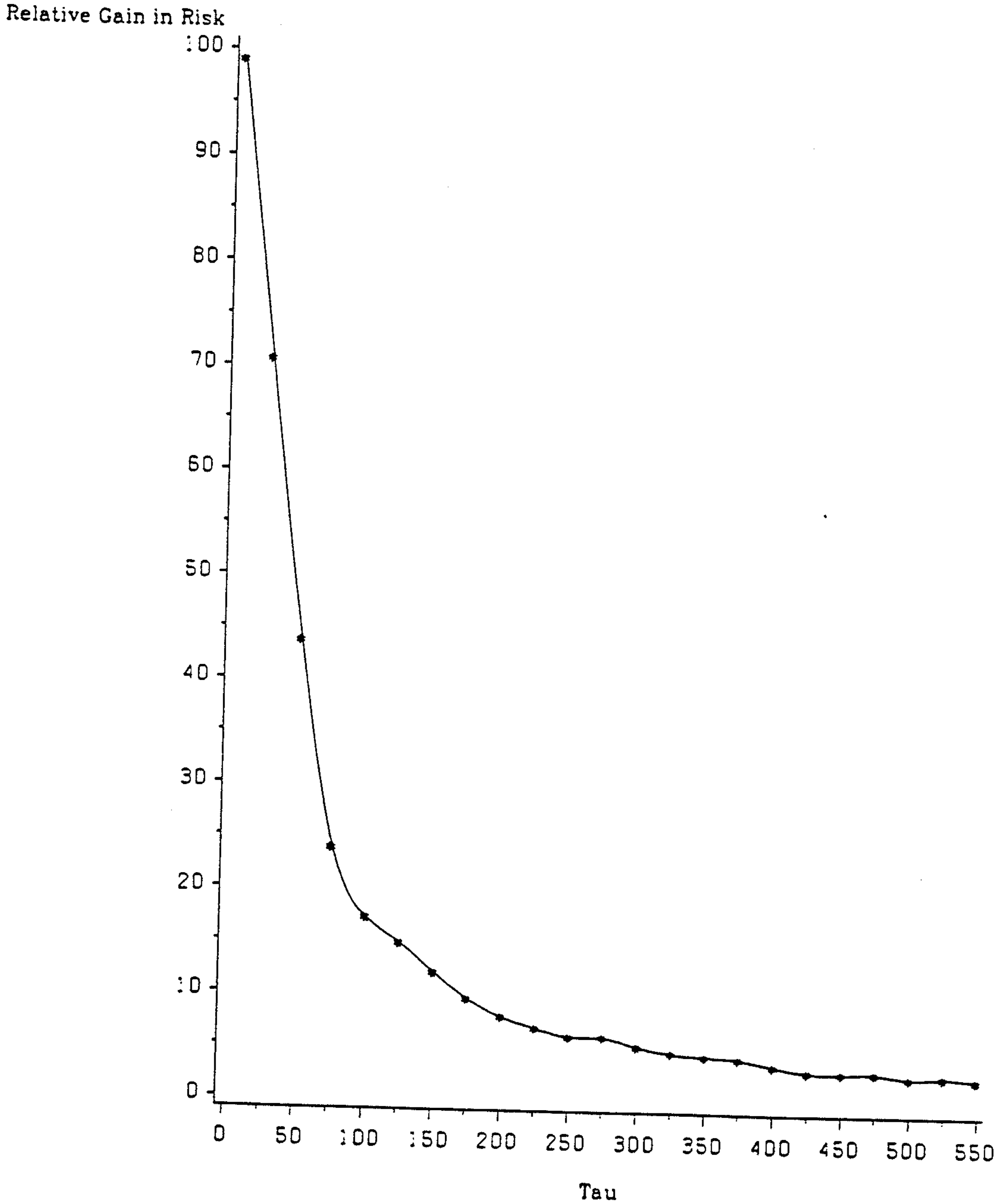
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Vector of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1.
ALPHA=(0.2, 0.2, 0.6); BETA=(1/3, 1/3, 1/3)



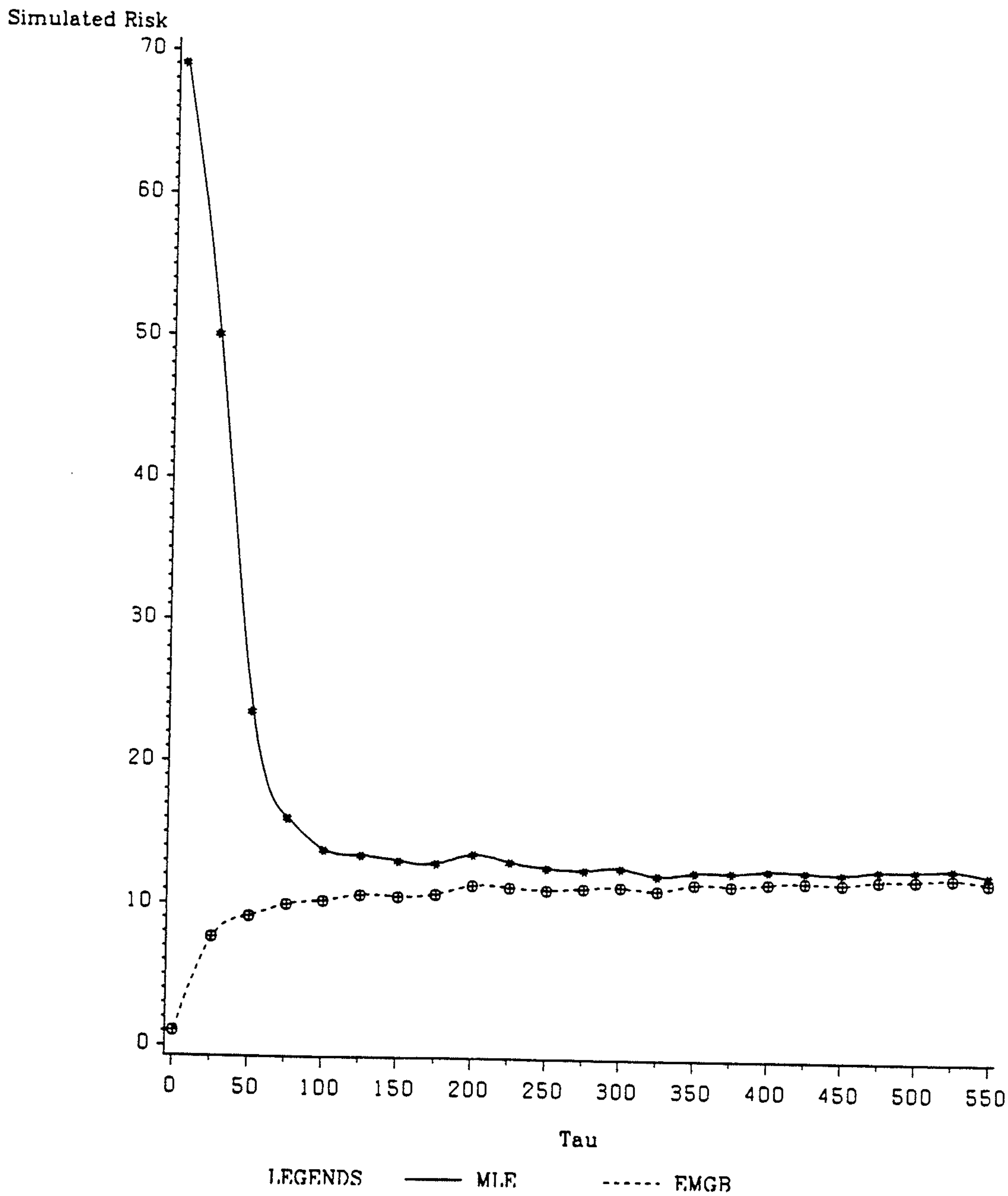
The Simulated Relative Gain in Risk using the
EMGB Estimator rather than the MLE of a Matrix
of Poisson Means under a Multiplicative Model
in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1
ALPHA=(0.2, 0.2, 0.6);BETA=(1/3, 1/3, 1/3)



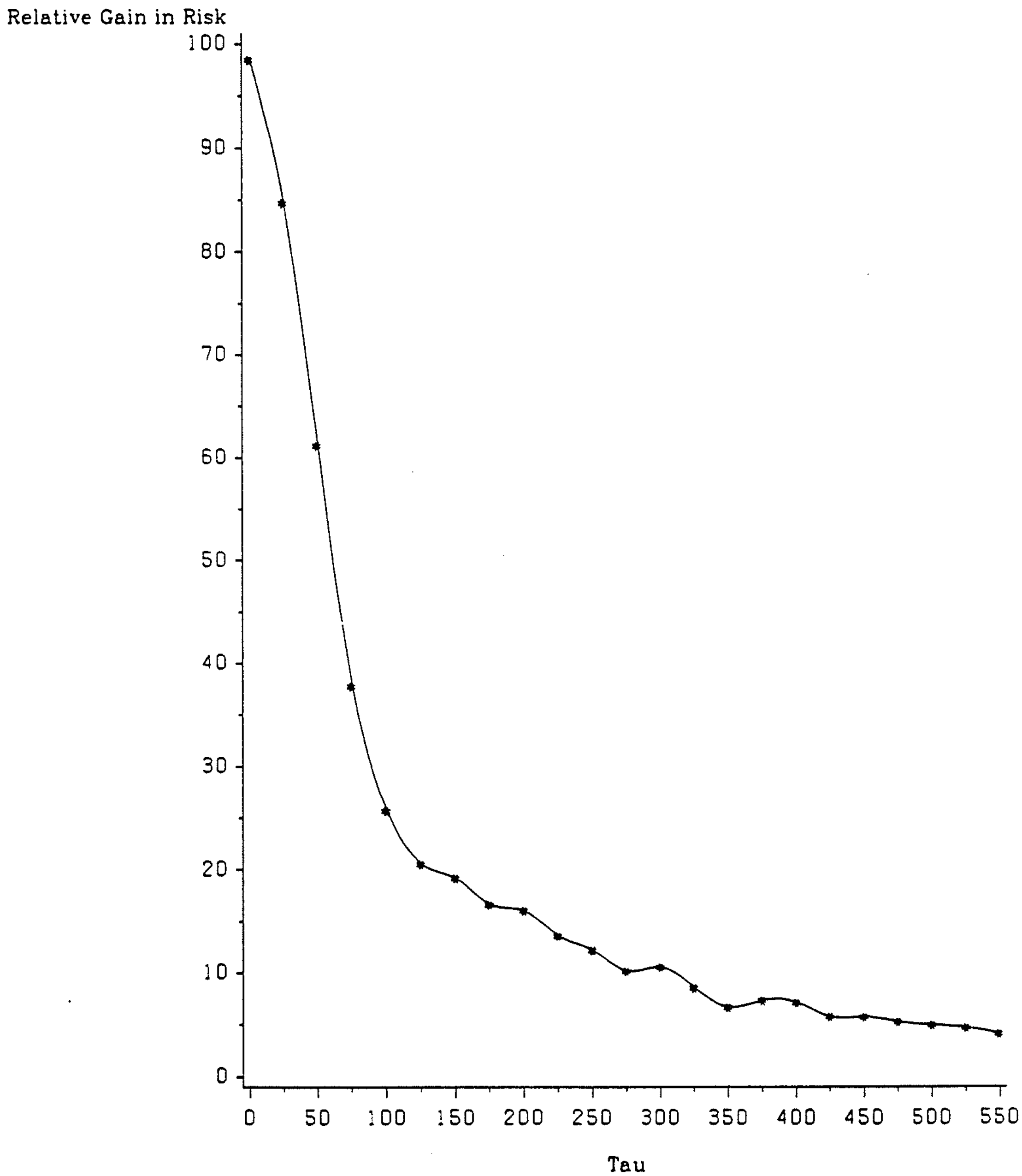
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Vector of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.
Incidence Matrix is IM2.

ALPHA=(0.2, 0.2, 0.6); BETA=(1/3, 1/3, 1/3)



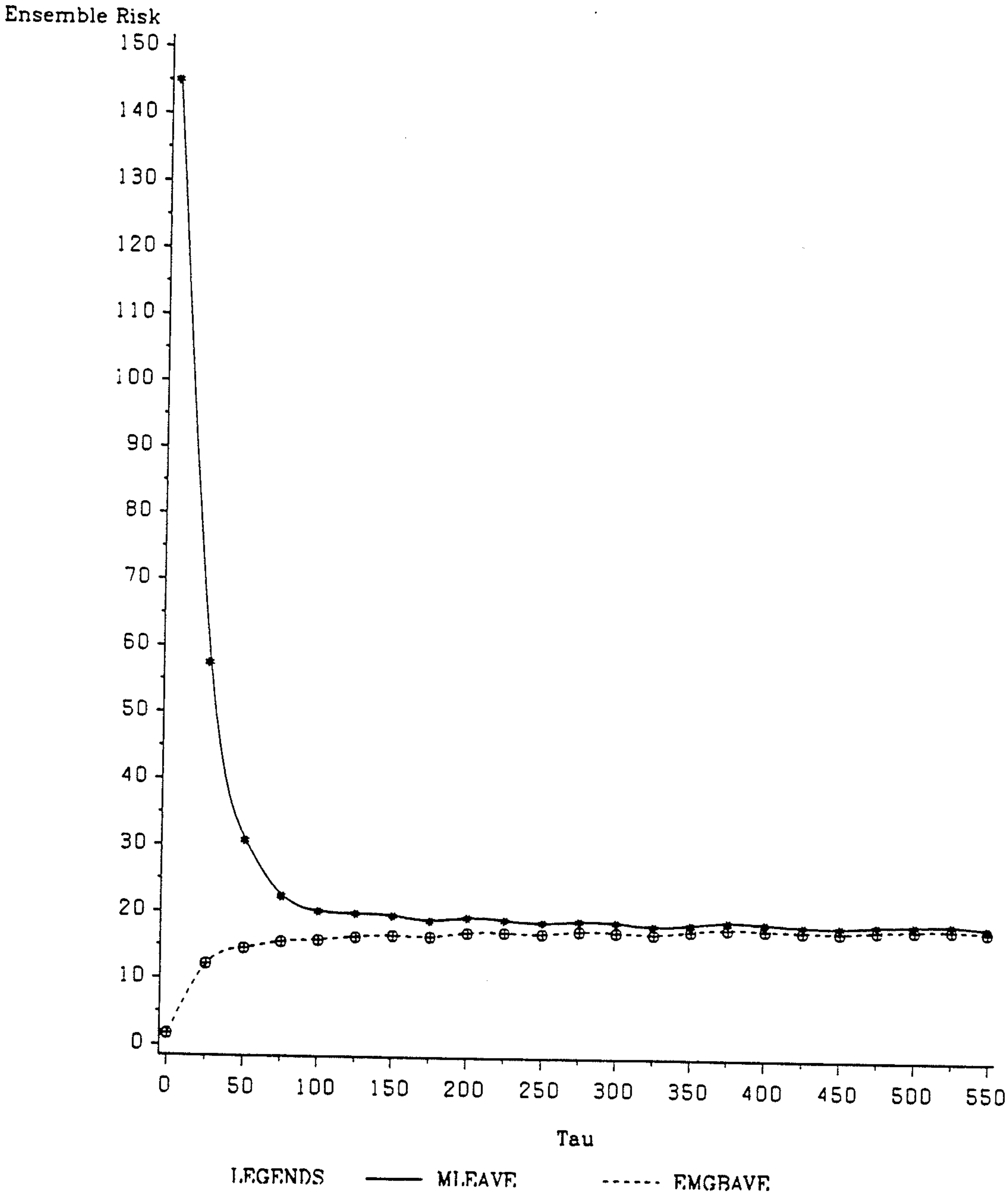
The Simulated Relative Gain in Risk using the
EMGB Estimator rather than the MLE of a Matrix
of Poisson Means under a Multiplicative Model
in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM2
ALPHA=(0.2, 0.2, 0.6);BETA=(1/3, 1/3, 1/3)



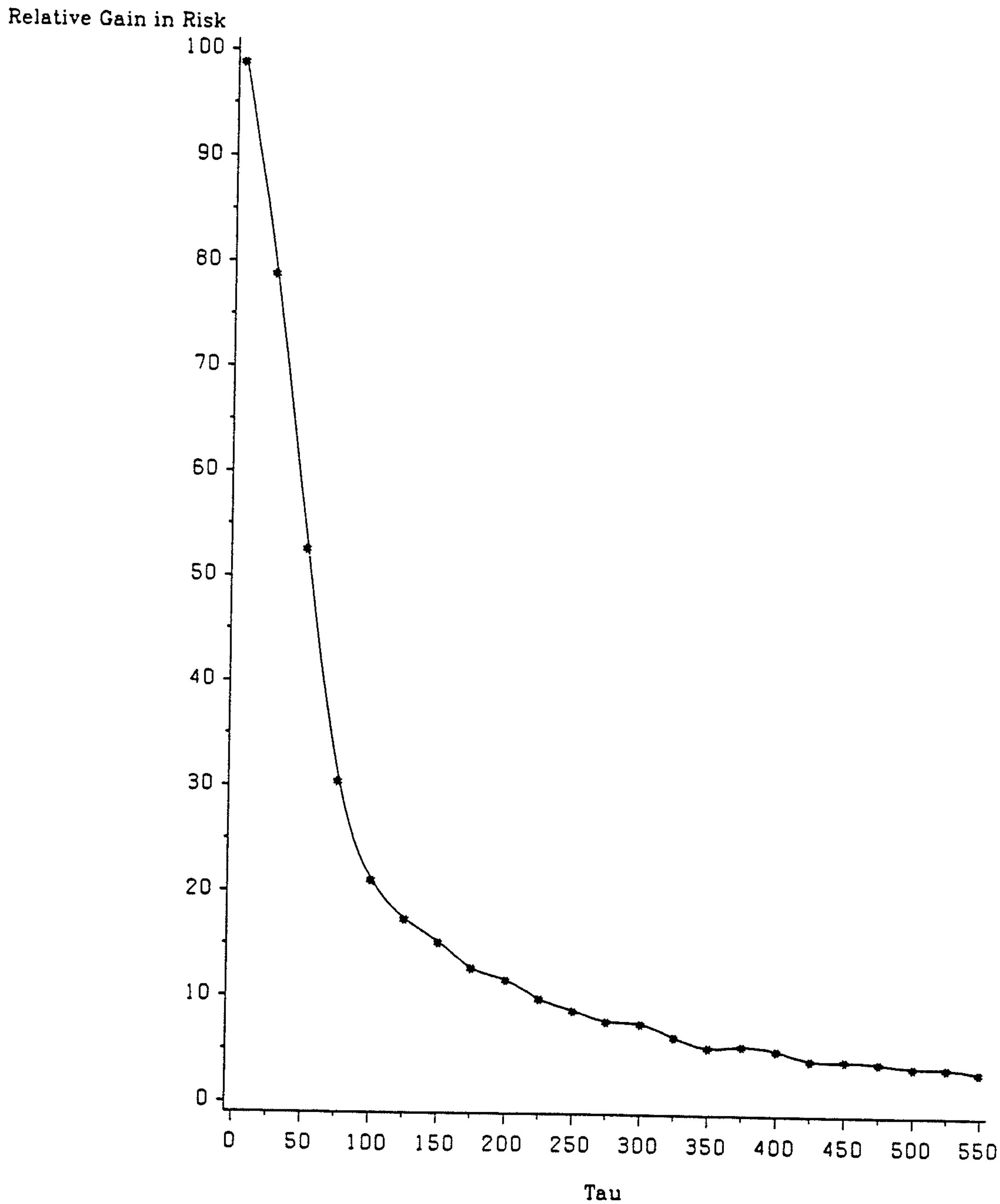
A Comparison of the Ensemble Risks of the MLE and the EMGB Estimator of a Matrix of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

MLEAVE=1/3 MLRIS1 + 2/3 MLRIS2
EMGBAVE=1/3 BARIS1+2/3 BARIS2
ALPHA=(0.2, 0.2, 0.6);BETA=(1/3, 1/3, 1/3)



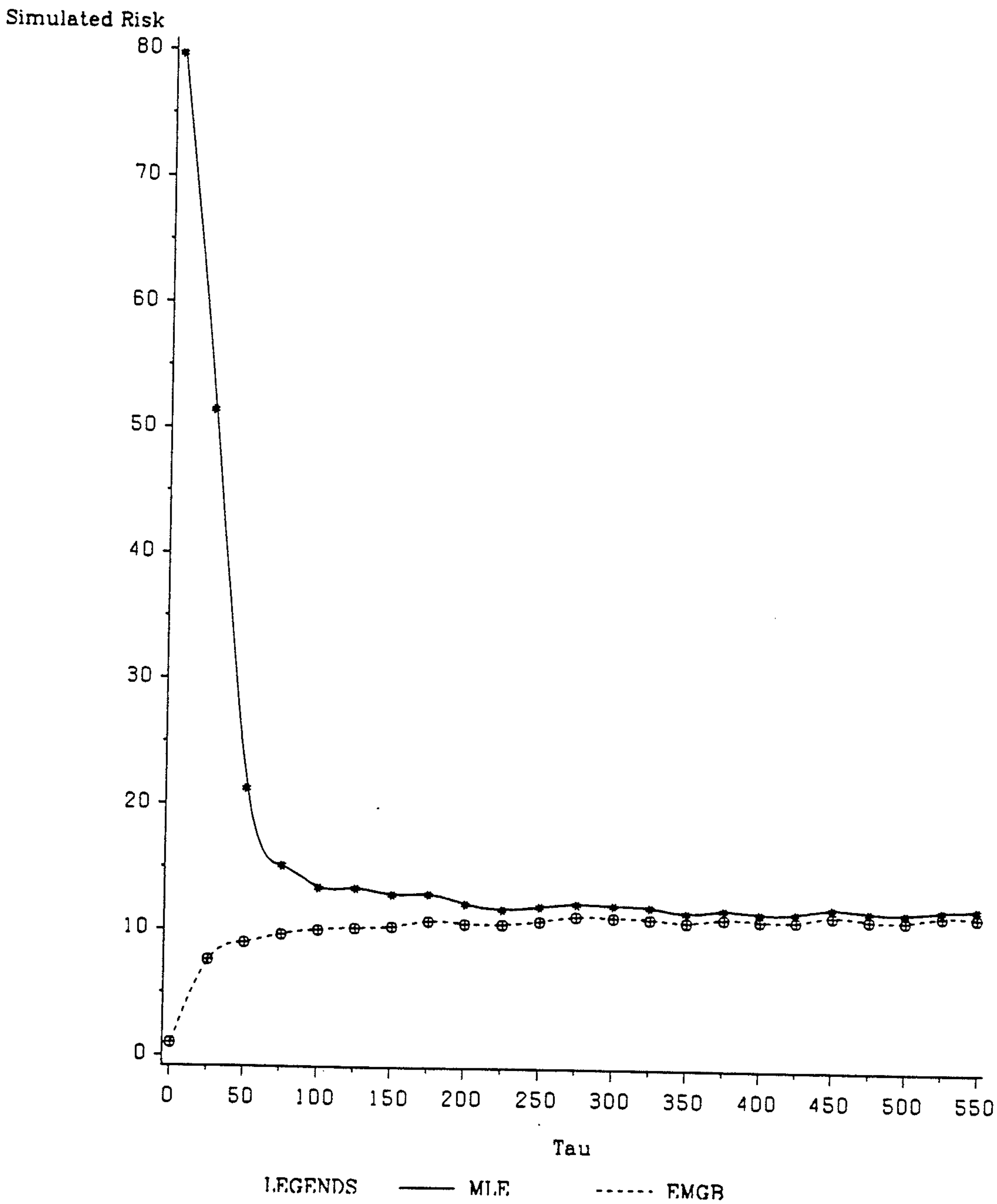
The Simulated Relative Gain in Ensemble Risk using the EMGB Estimator rather than the MLE of a Vector Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

ALPHA=(0.2, 0.2, 0.6); BETA=(1/3, 1/3, 1/3)



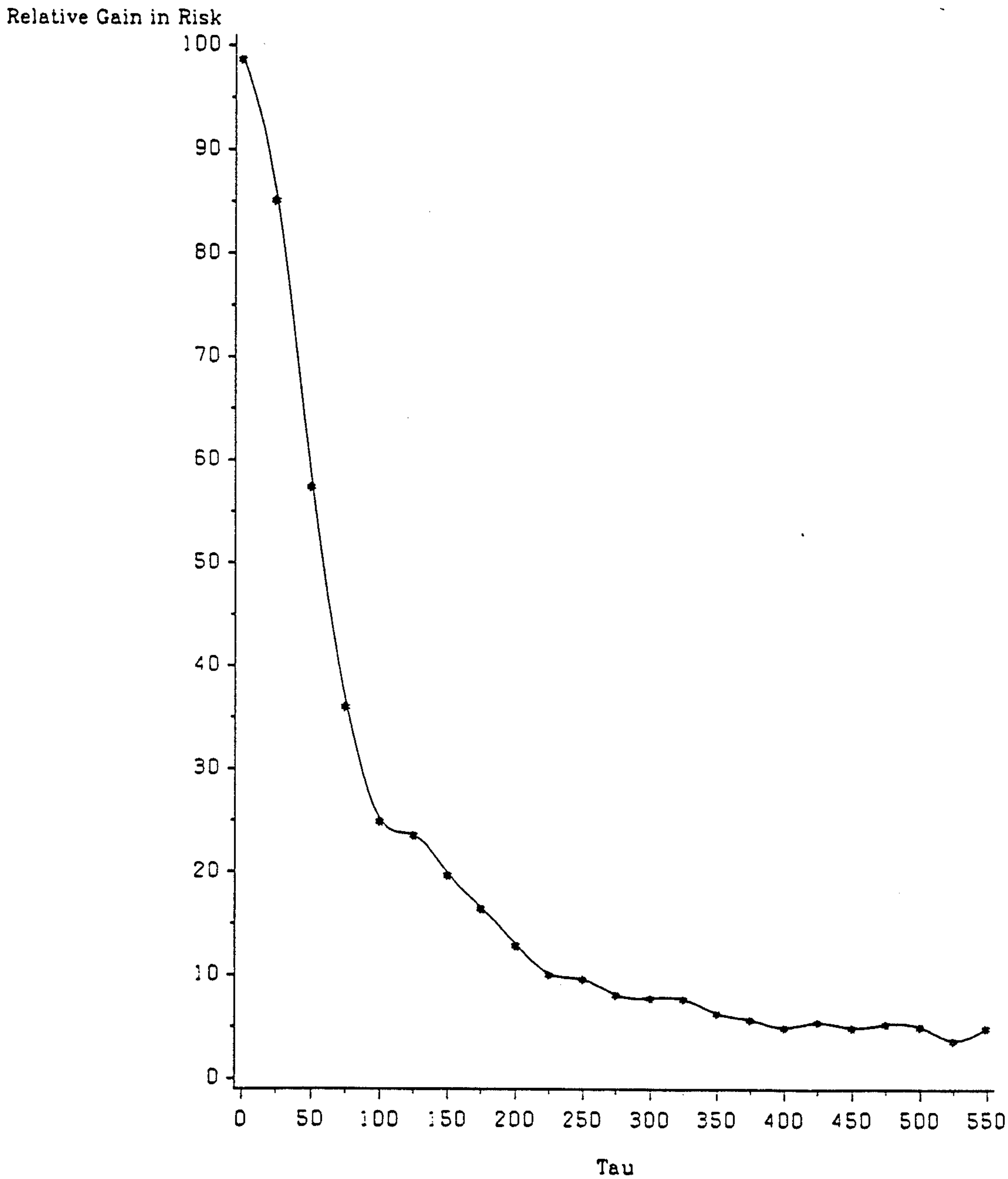
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Vector of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1:
ALPHA=(0.2, 0.2, 0.6); BETA=(0.4, 0.4, 0.2)



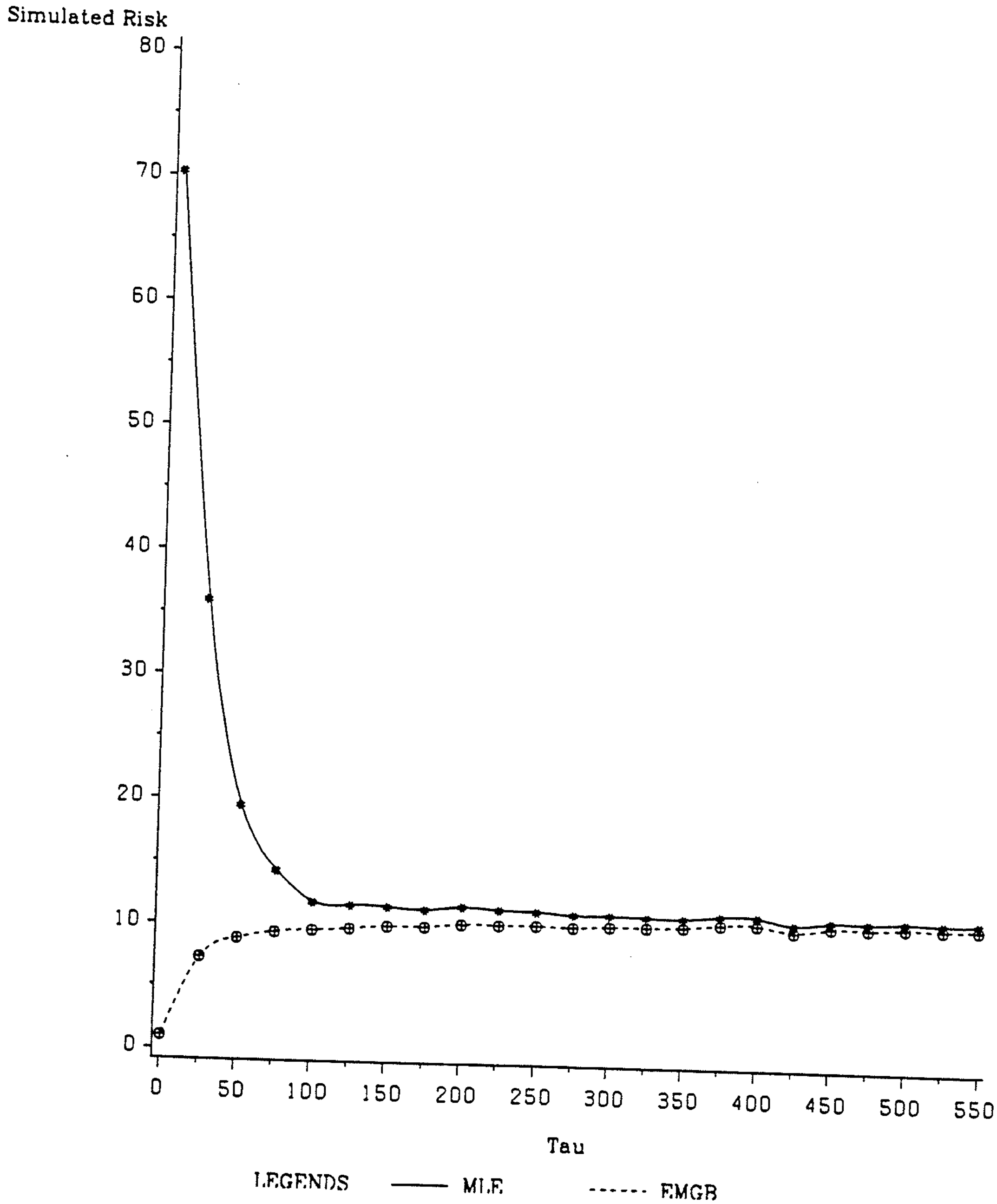
The Simulated Relative Gain in Risk using the
EMGB Estimator rather than the MLE of a Matrix
of Poisson Means under a Multiplicative Model
in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1
ALPHA=(0.2, 0.2, 0.6);BETA=(0.4, 0.4, 0.2)



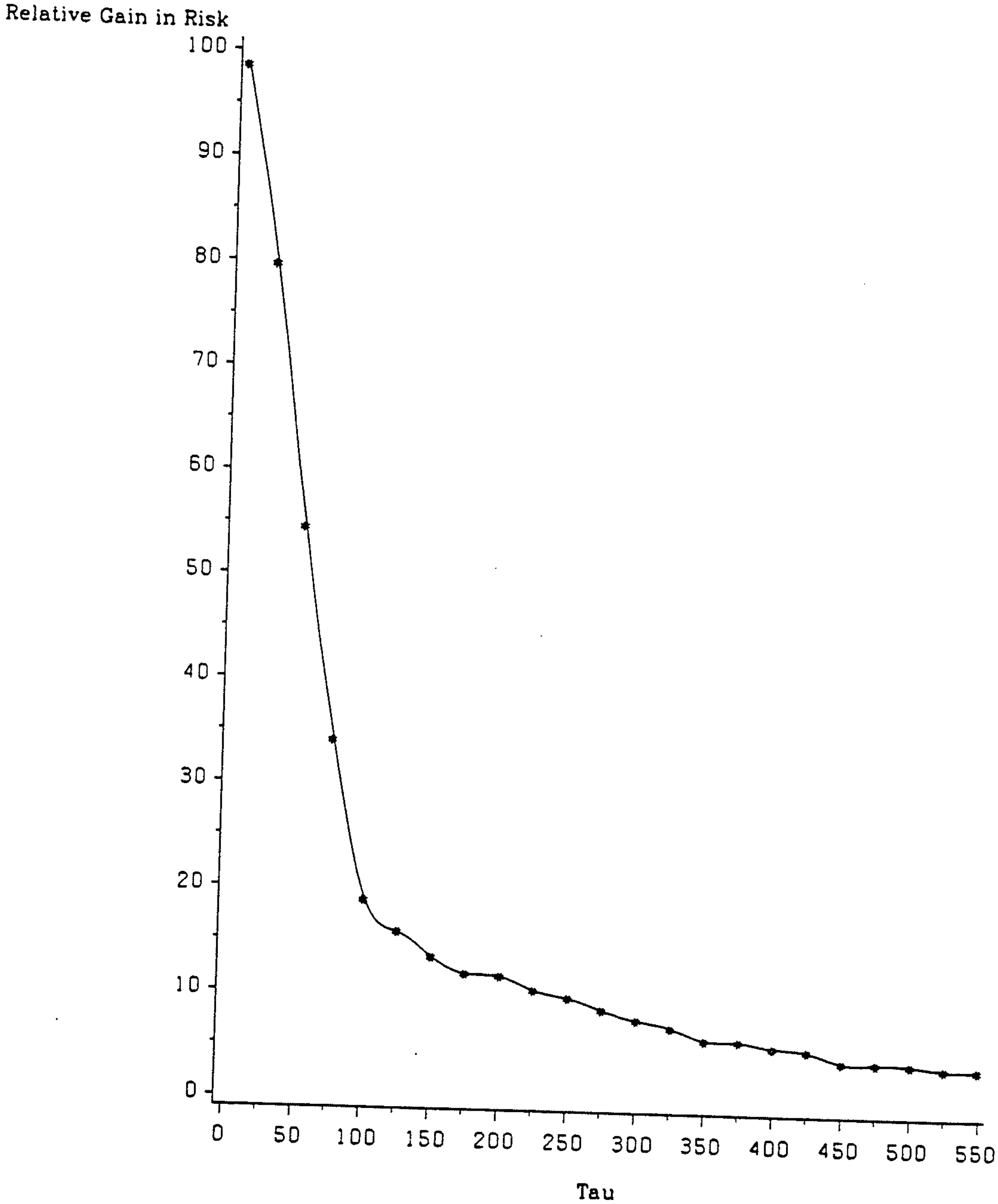
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Vector of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.
Incidence Matrix is IM2.

ALPHA=(0.2, 0.2, 0.6); BETA=(0.4, 0.4, 0.2)



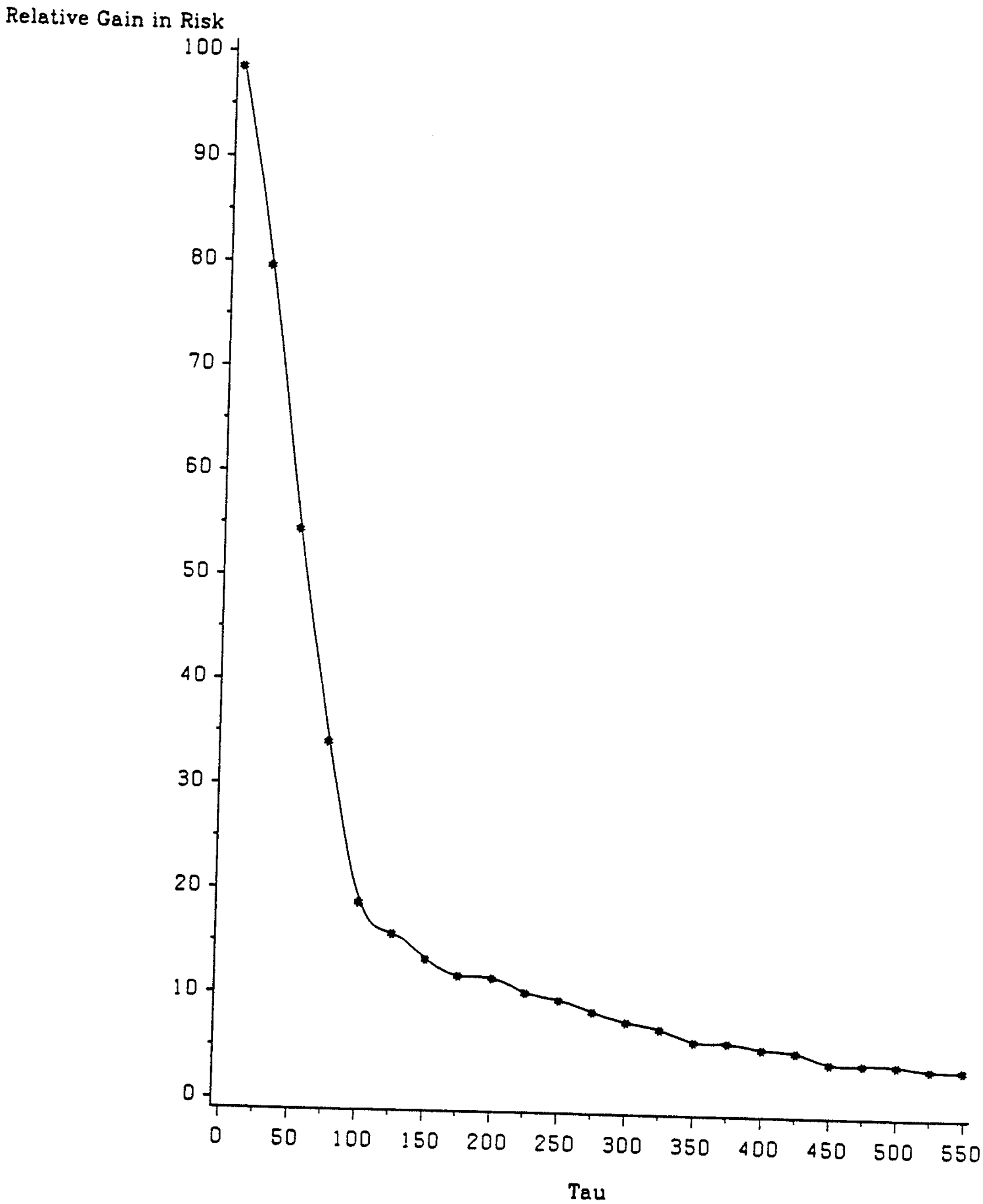
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM2
ALPHA=(0.2, 0.2, 0.6);BETA=(0.4, 0.4, 0.2)



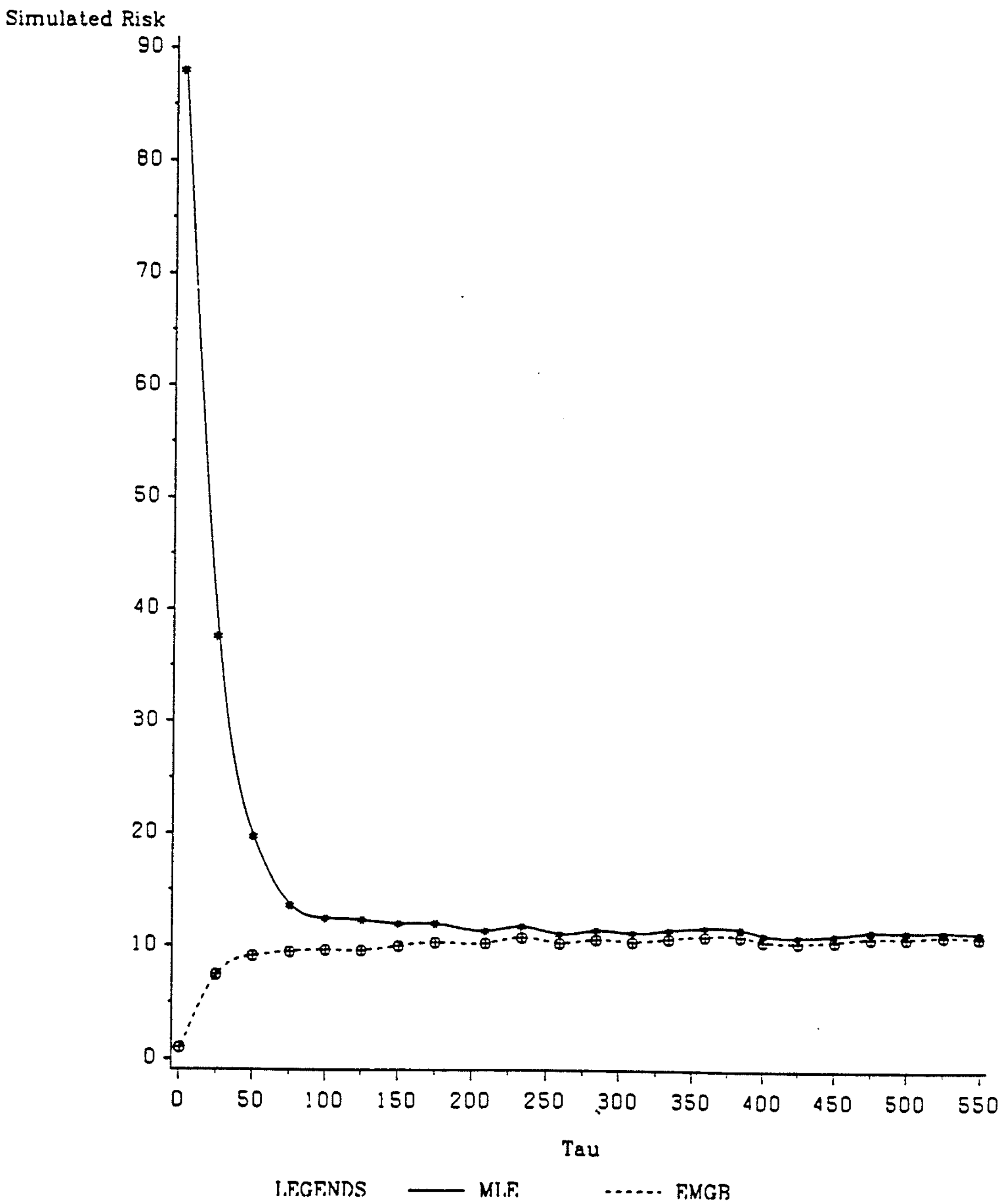
The Simulated Relative Gain in Ensemble Risk using the EMGB Estimator rather than the MLE of a Vector Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

ALPHA=(0.2, 0.2, 0.6); BETA=(0.4, 0.4, 0.2)



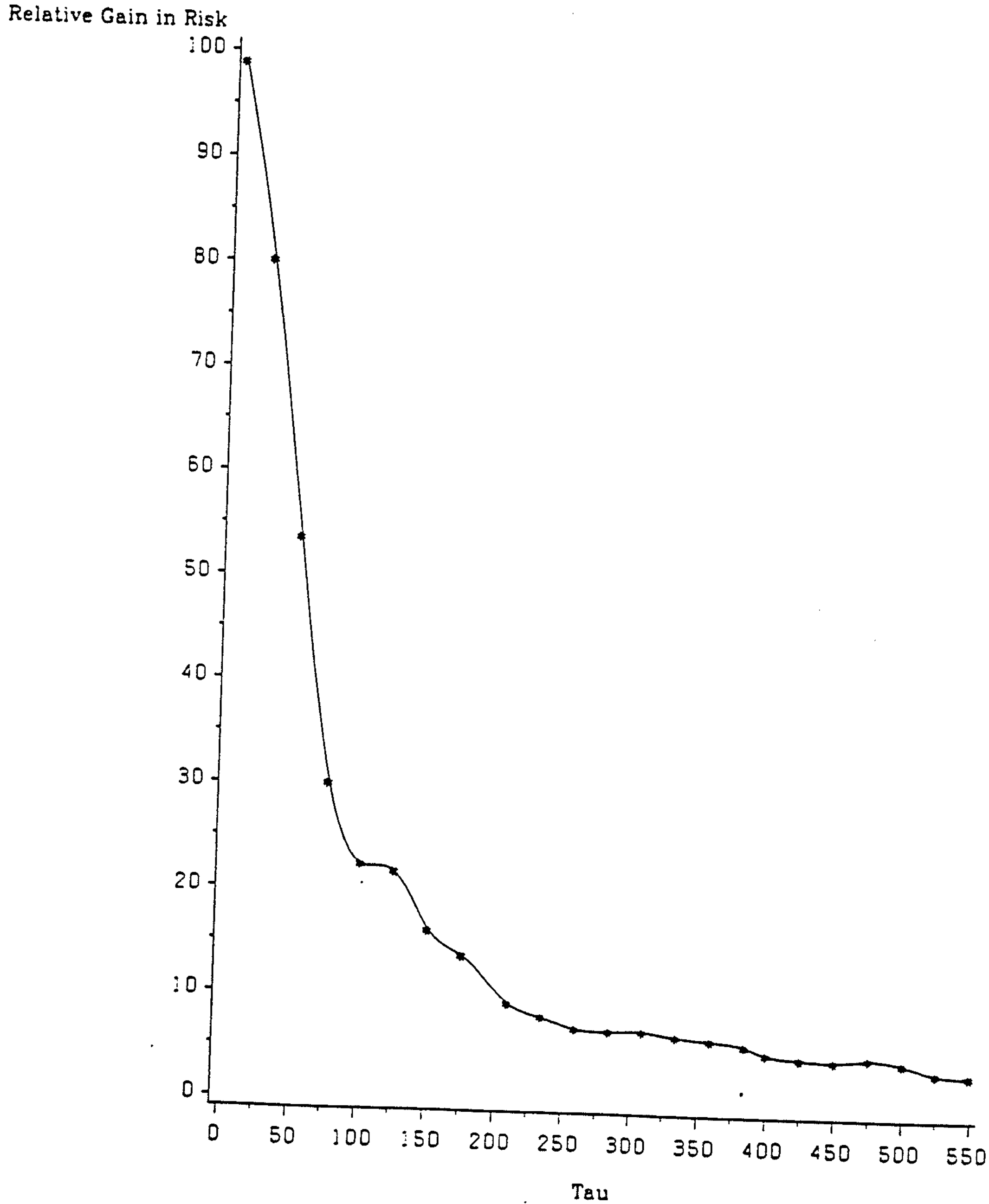
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Matrix of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1.
ALPHA=(0.3, 0.3, 0.4); BETA=(0.3, 0.3, 0.4)



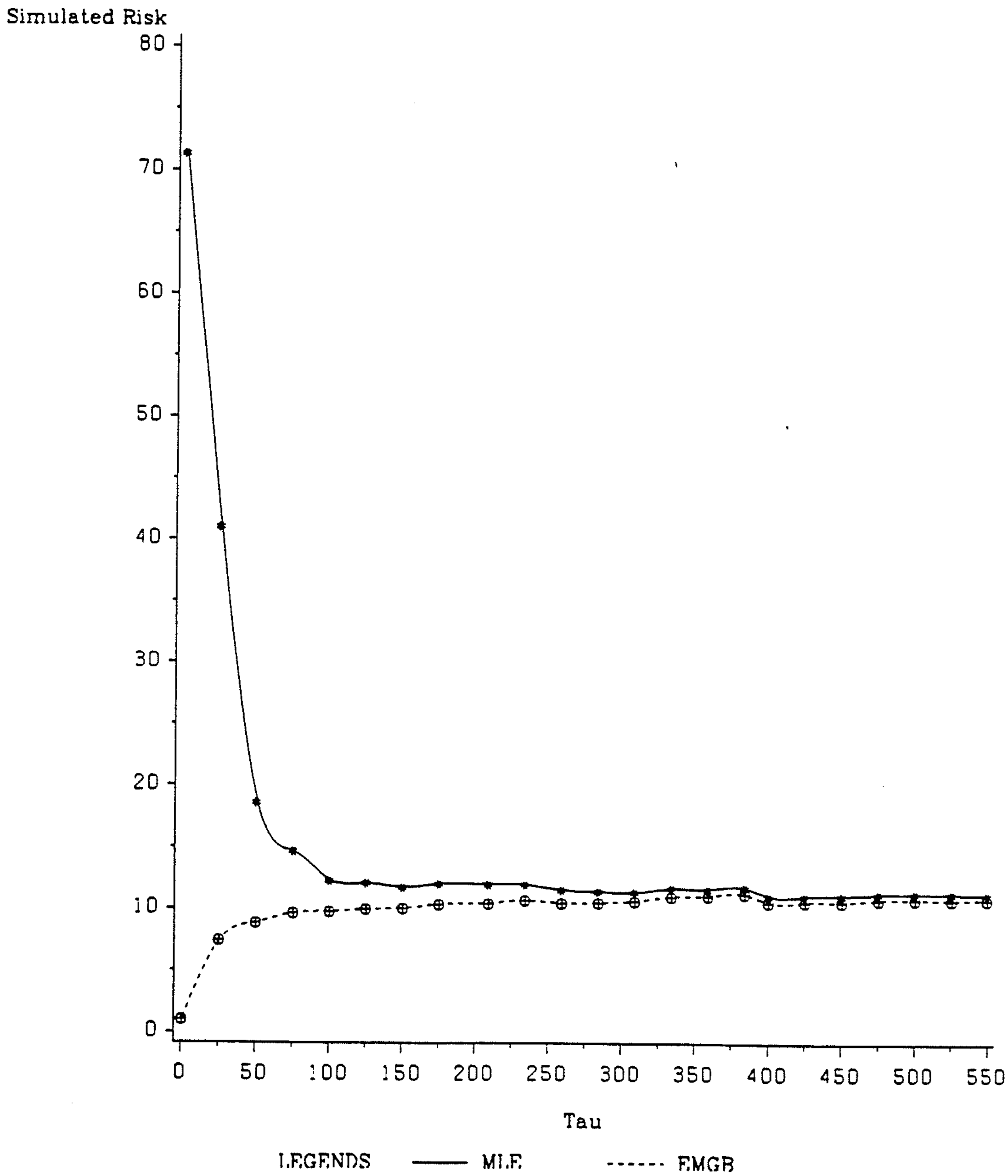
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1
ALPHA=(0.3, 0.3, 0.4);BETA=(0.3, 0.3, 0.4)



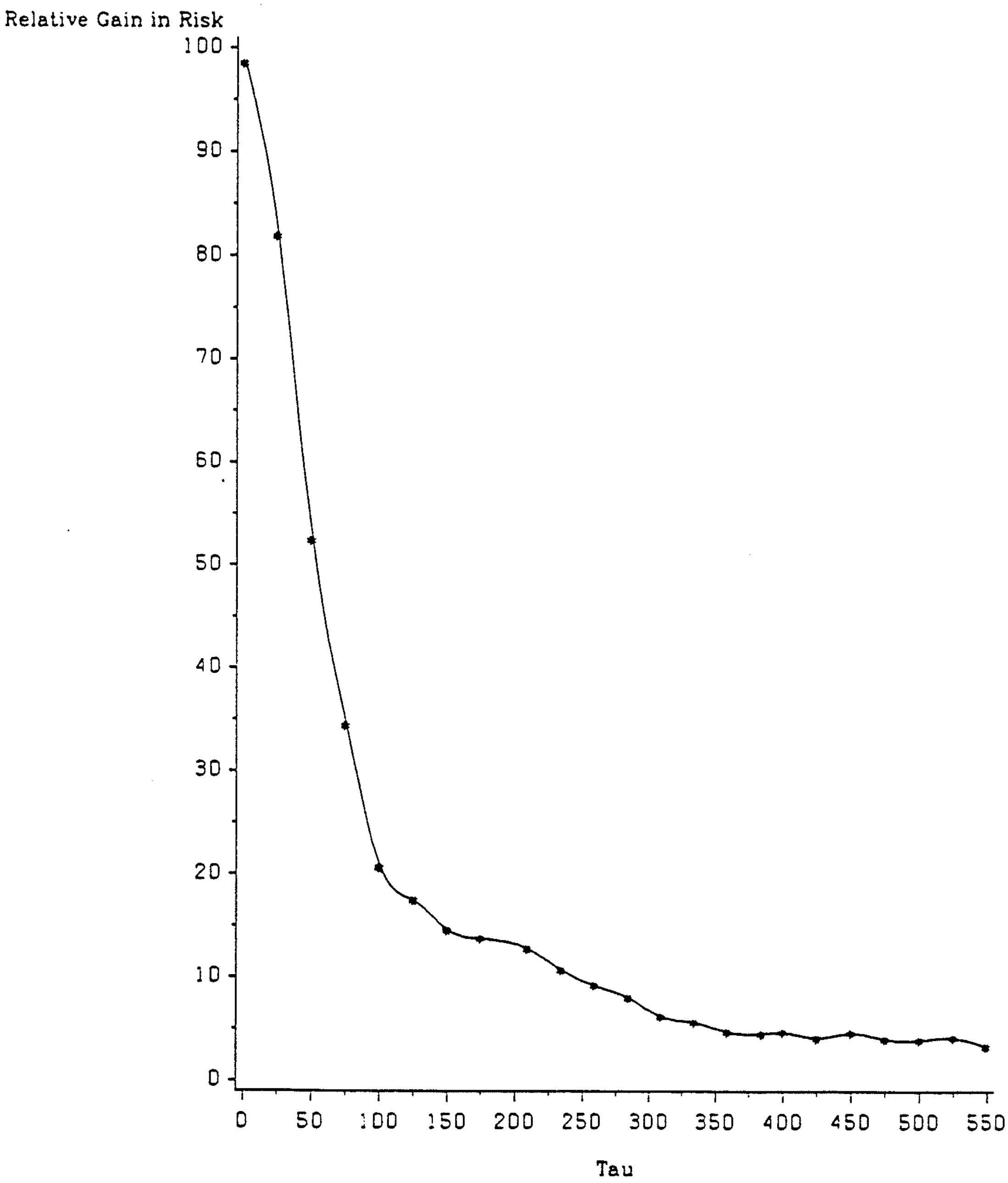
A Comparison of the Simulated Risks of the
MLE and the EMGB Estimator of a Matrix of
Poisson Means under a Multiplicative Model
for a Balanced Incomplete 3x3 Table.
Incidence Matrix is IM2.

ALPHA=(0.3, 0.3, 0.4); BETA=(0.3, 0.3, 0.4)



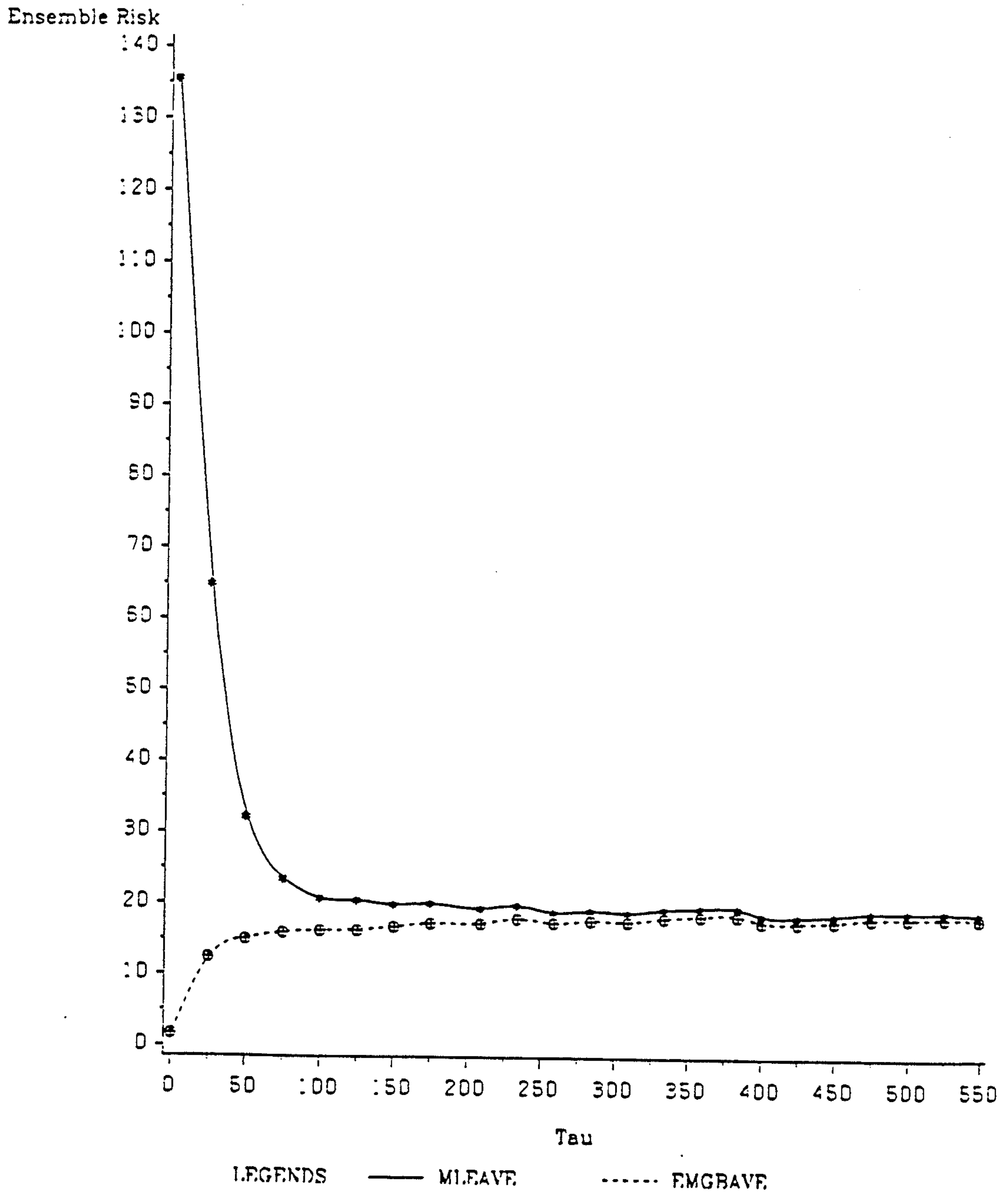
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM2
ALPHA=(0.3, 0.3, 0.4);BETA=(0.3, 0.3, 0.4)



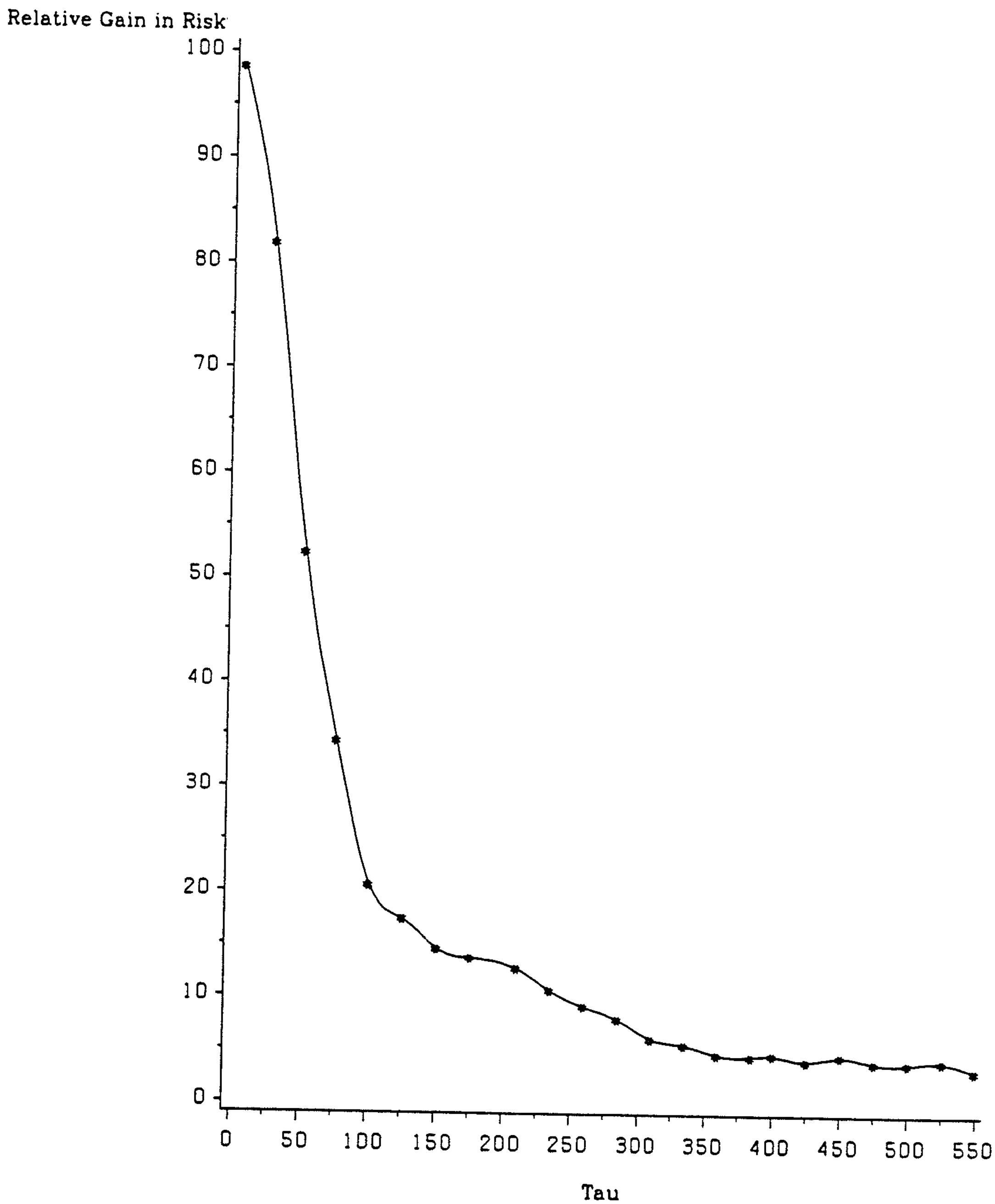
A Comparison of the Ensemble Risks of the MLE and the EMGB Estimator of a Matrix of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

MLEAVE=1/3 MLRIS1 + 2/3 MLRIS2
EMGBAVE=1/3 BARIS1+2/3 BARIS2
ALPHA=(0.3, 0.3, 0.4); BETA=(0.3, 0.3, 0.4)



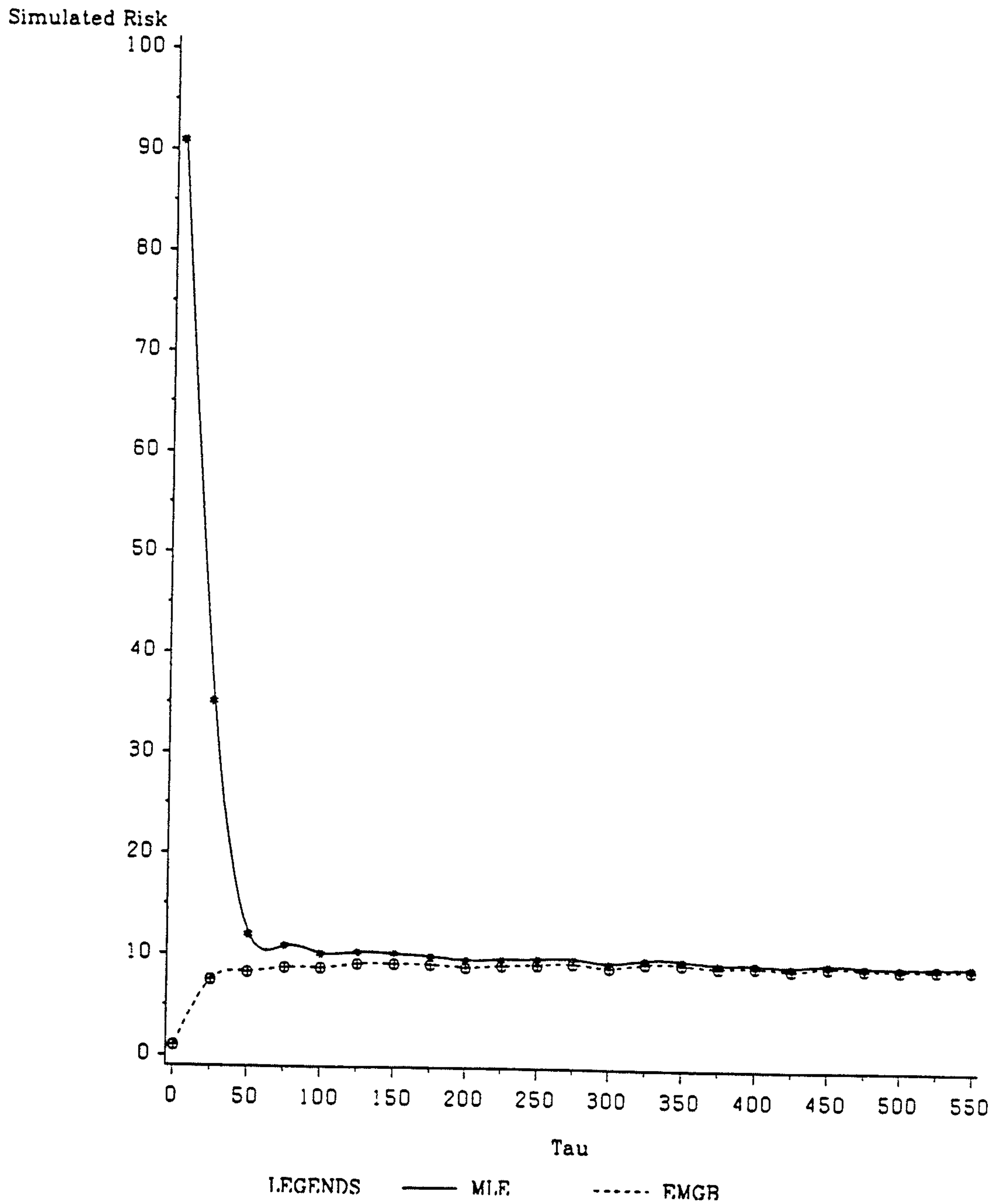
The Simulated Relative Gain in Ensemble Risk using the EMGB Estimator rather than the MLE of a Vector Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

ALPHA=(0.3, 0.3, 0.4); BETA=(0.3, 0.3, 0.4)



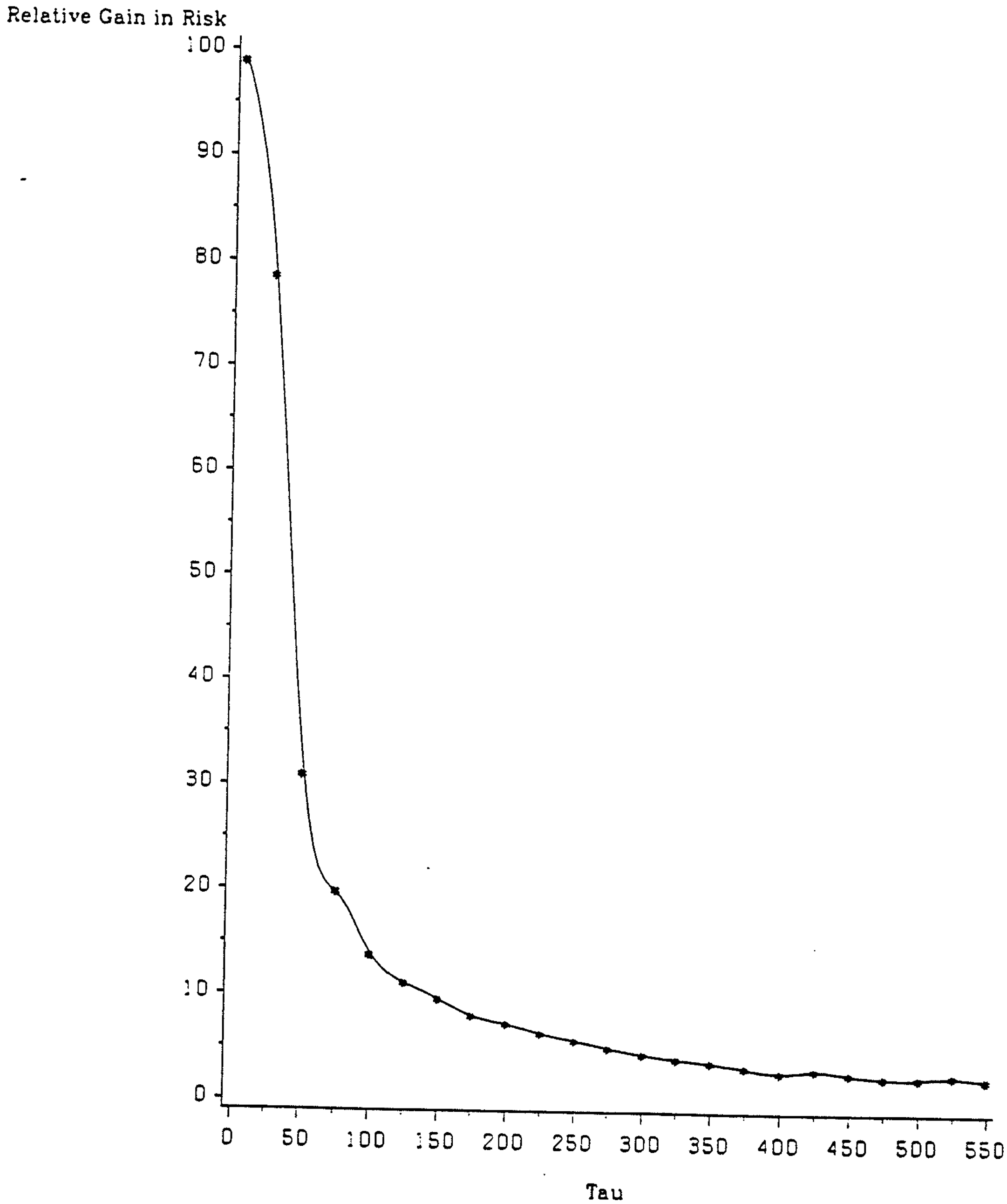
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Matrix of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1.
ALPHA=(0.3, 0.3, 0.4); BETA=(1/3, 1/3, 1/3)



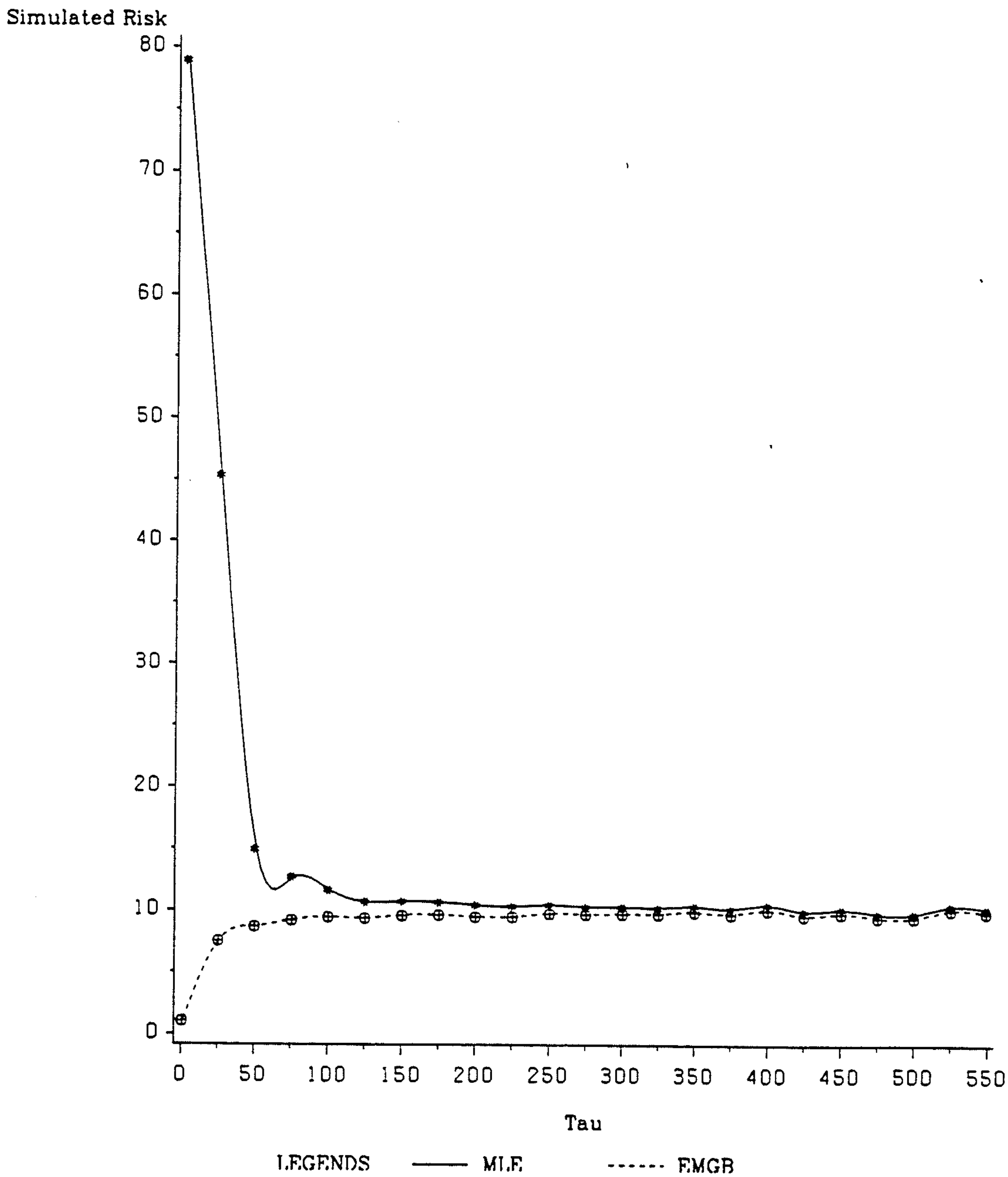
The Simulated Relative Gain in Risk using the
EMGB Estimator rather than the MLE of a Matrix
of Poisson Means under a Multiplicative Model
in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1
ALPHA=(0.3, 0.3, 0.4);BETA=(1/3, 1/3, 1/3)



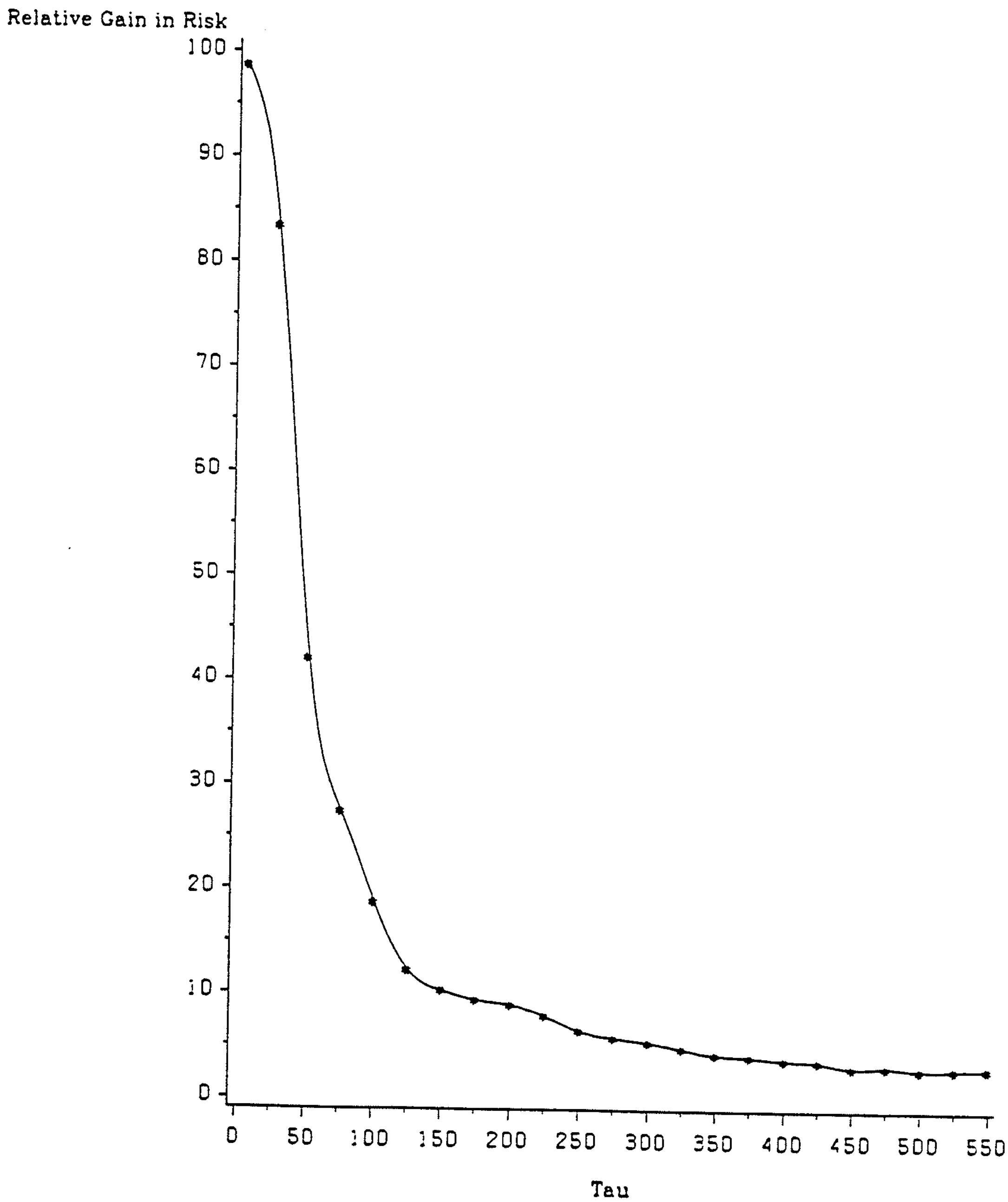
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Matrix of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.
Incidence Matrix is IM2.

ALPHA=(0.3, 0.3, 0.4); BETA=(1/3, 1/3, 1/3)



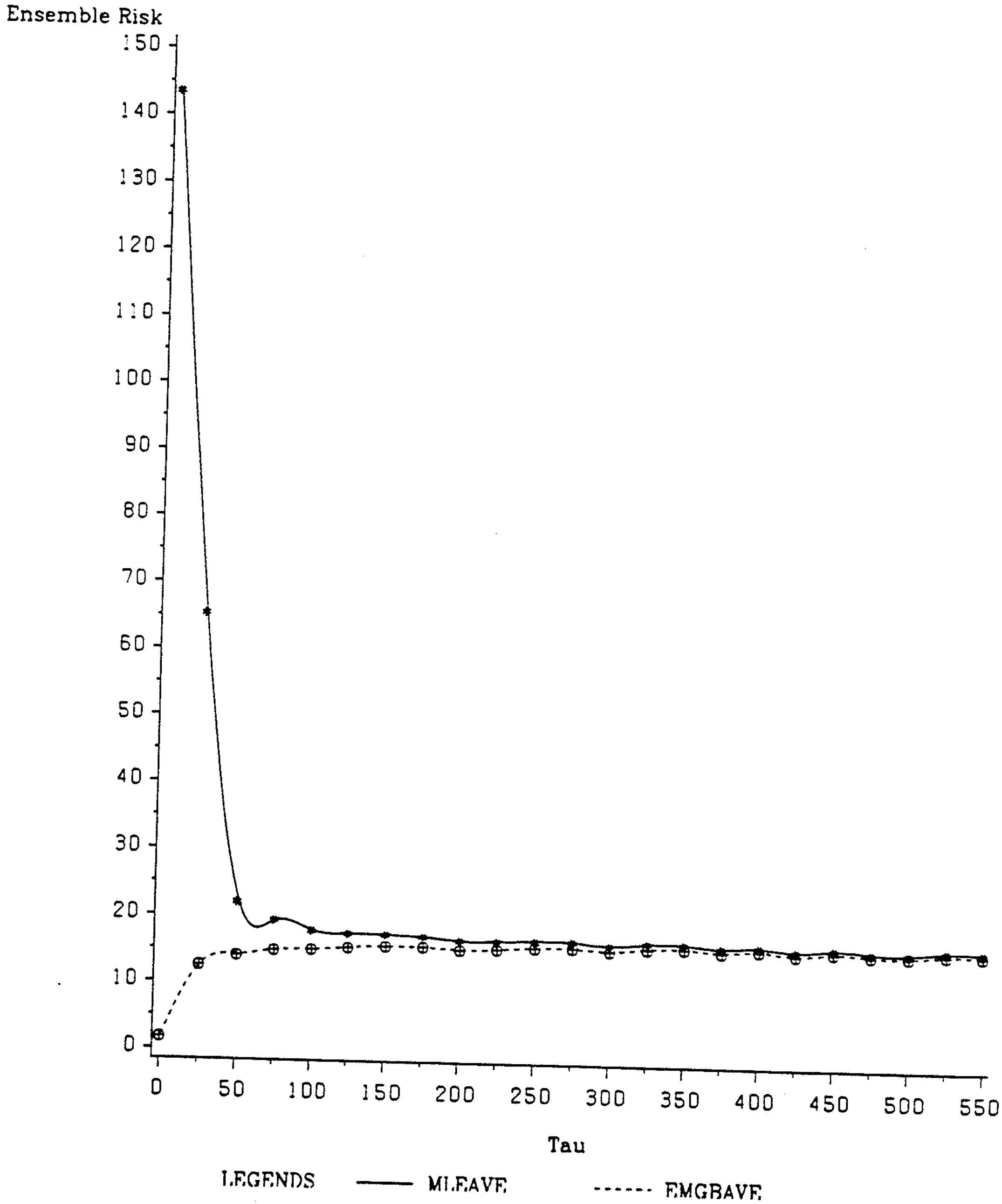
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM2
ALPHA=(0.3, 0.3, 0.4);BETA=(1/3, 1/3, 1/3)



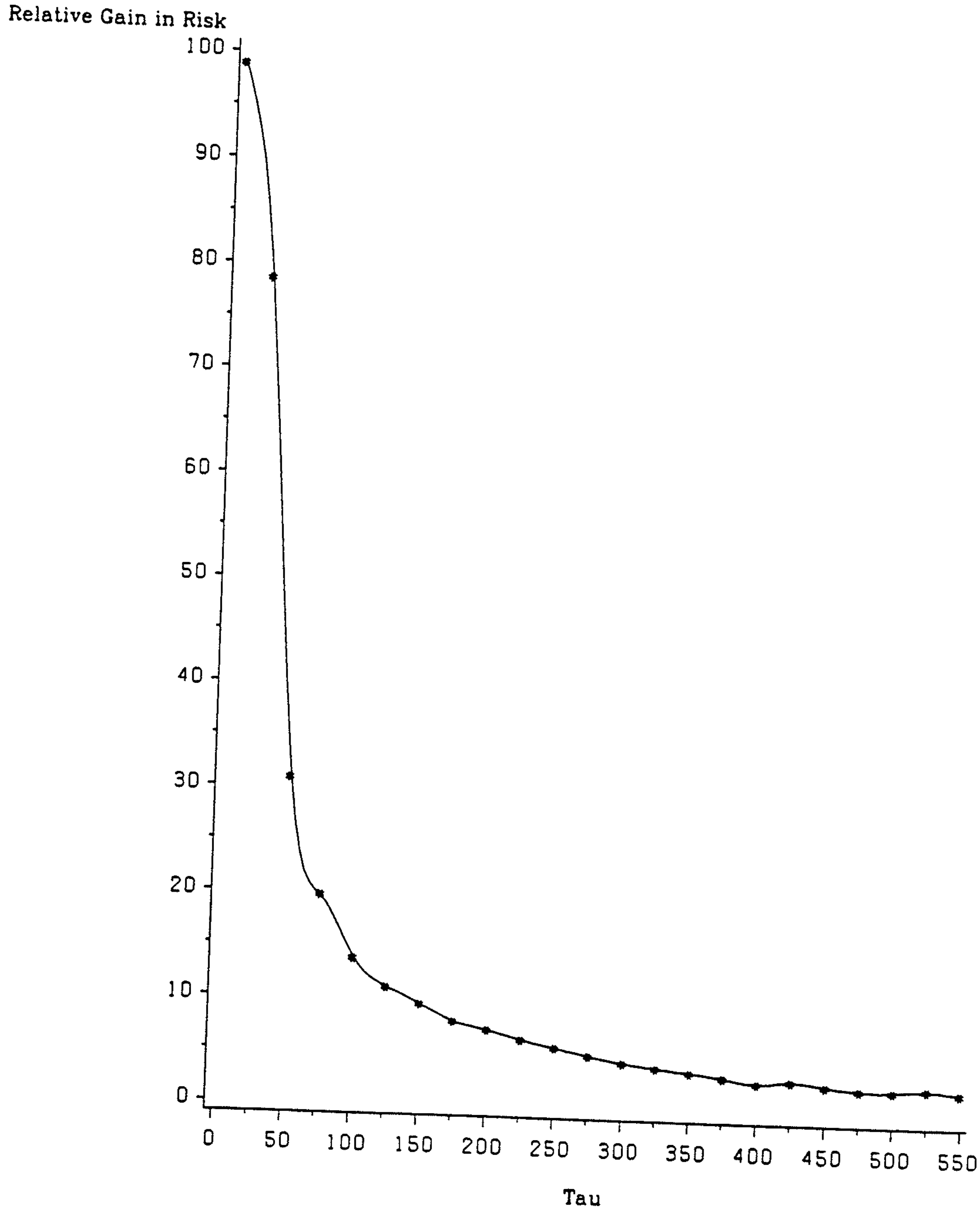
A Comparison of the Ensemble Risks of the MLE and the EMGB Estimator of a Matrix of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

MLEAVE=1/3 MLRIS1 + 2/3 MLRIS2
EMGBAVE=1/3 BARIS1+2/3 BARIS2
ALPHA=(0.3, 0.3, 0.4); BETA=(1/3, 1/3, 1/3)



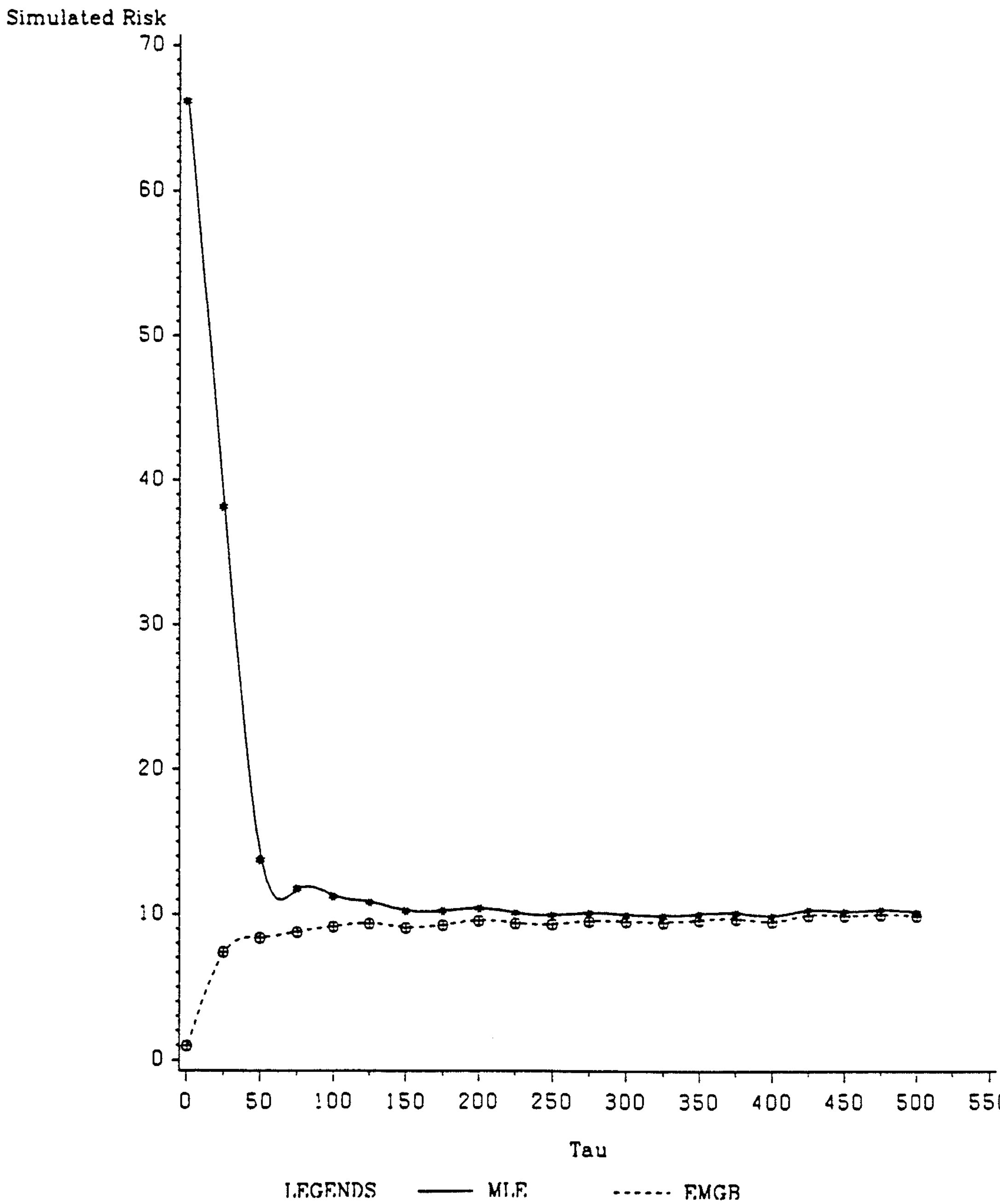
The Simulated Relative Gain in Ensemble Risk using the EMGB Estimator rather than the MLE of a Vector Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

ALPHA=(0.3, 0.3, 0.4); BETA=(1/3, 1/3, 1/3)



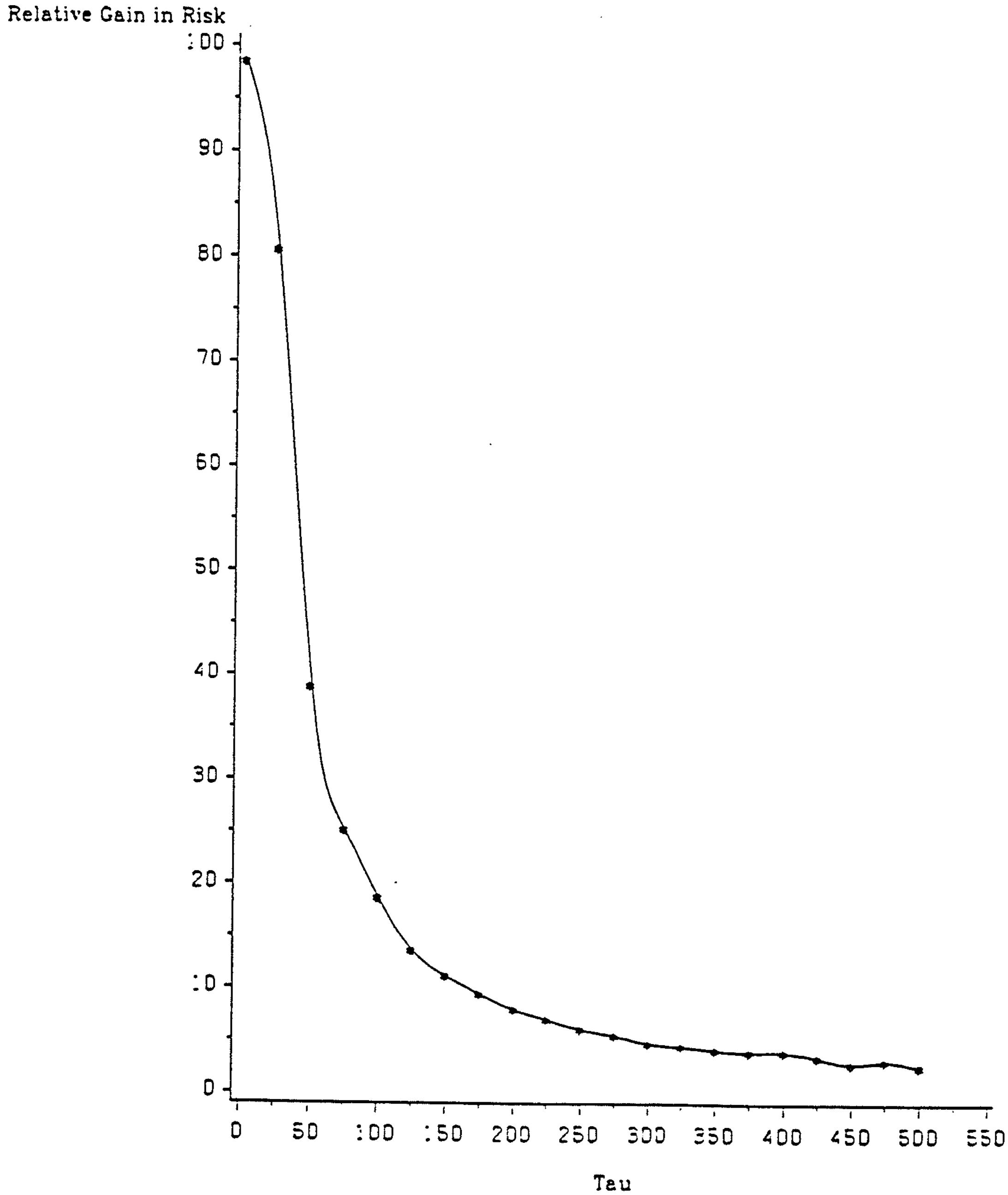
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Matrix of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1.
ALPHA=(0.3, 0.3, 0.4); BETA=(0.4, 0.4, 0.2)



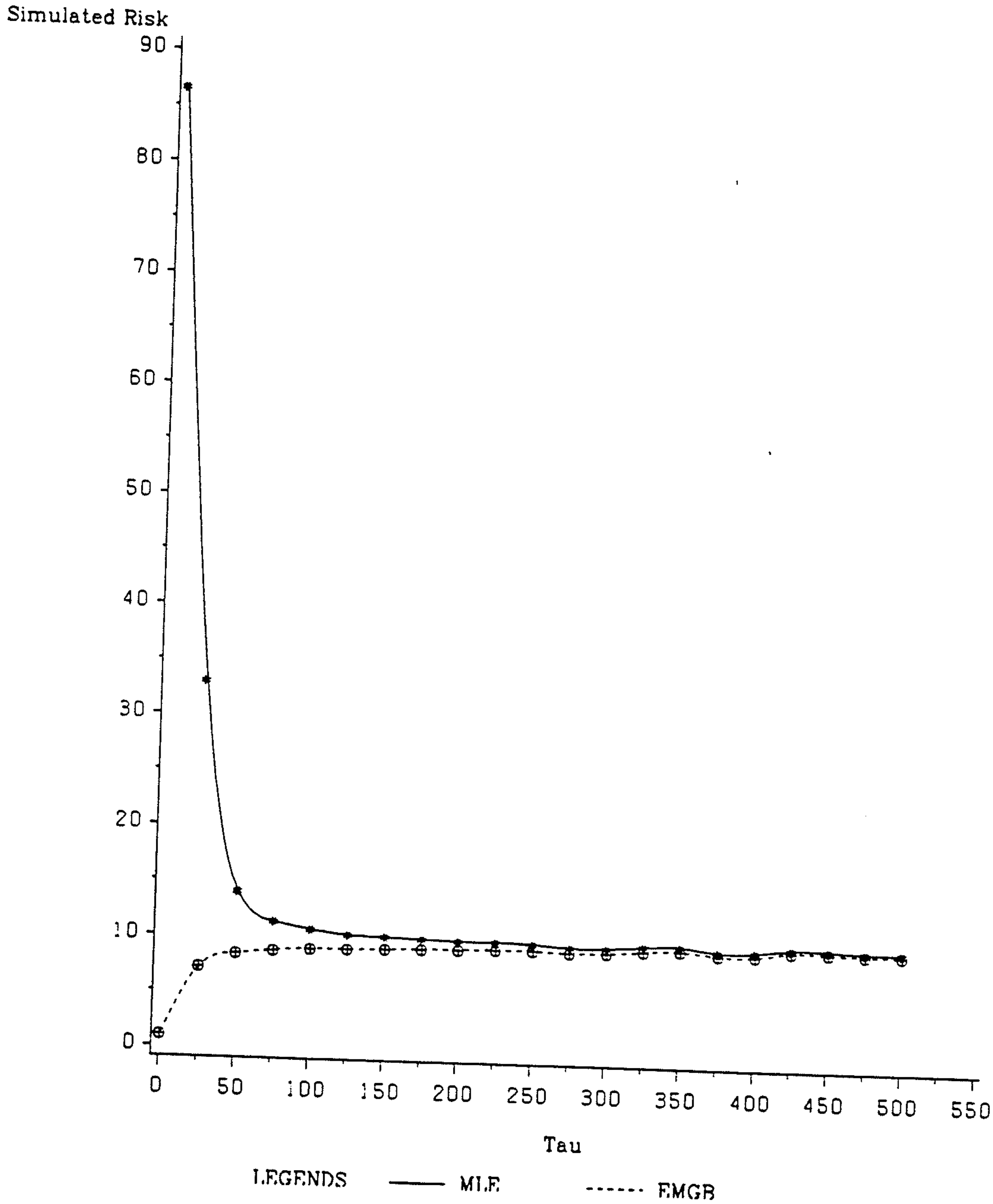
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1
ALPHA=(0.3, 0.3, 0.4);BETA=(0.4, 0.4, 0.2)



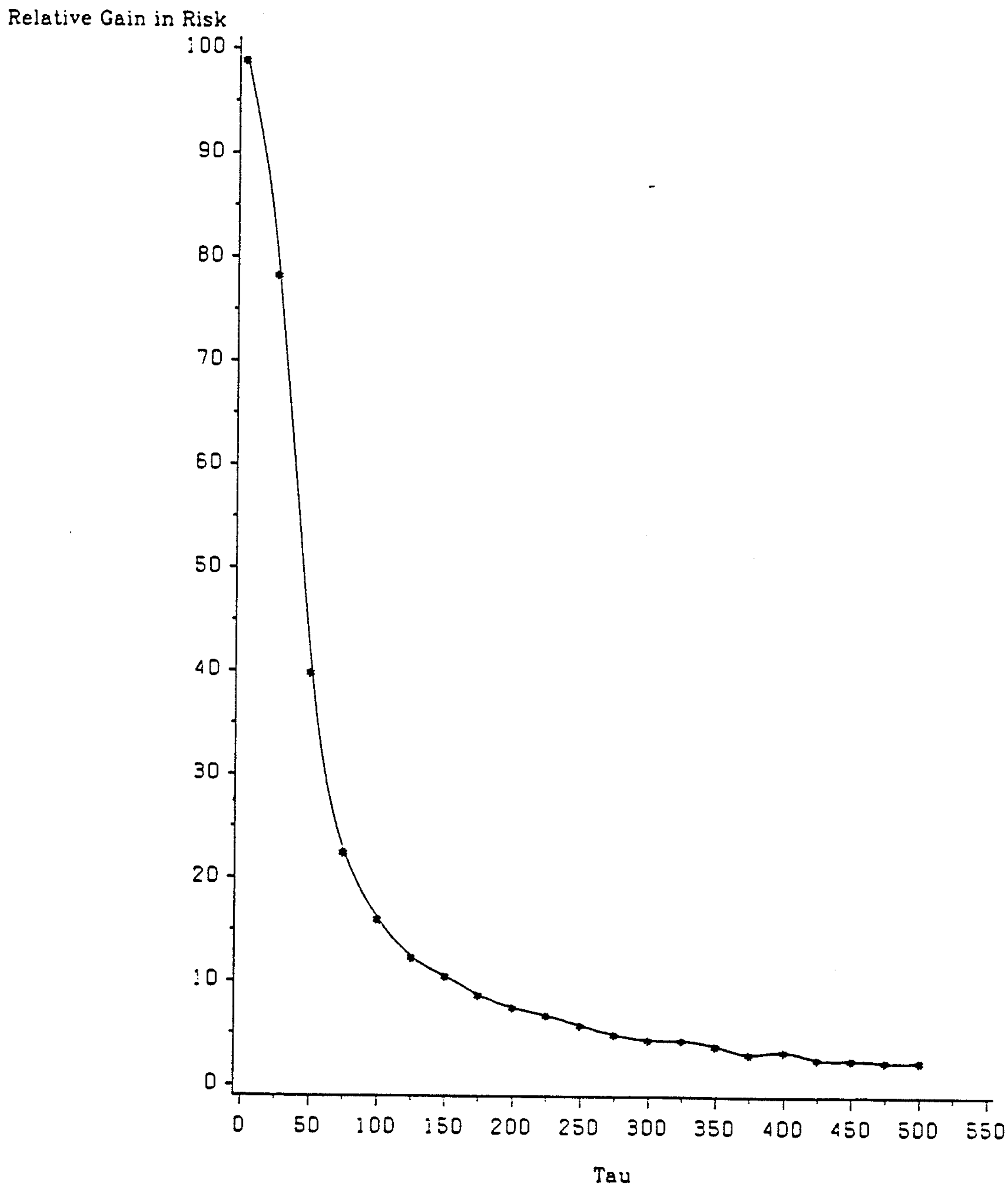
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Matrix of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.
Incidence Matrix is IM2.

ALPHA=(0.3, 0.3, 0.4); BETA=(0.4, 0.4, 0.2)



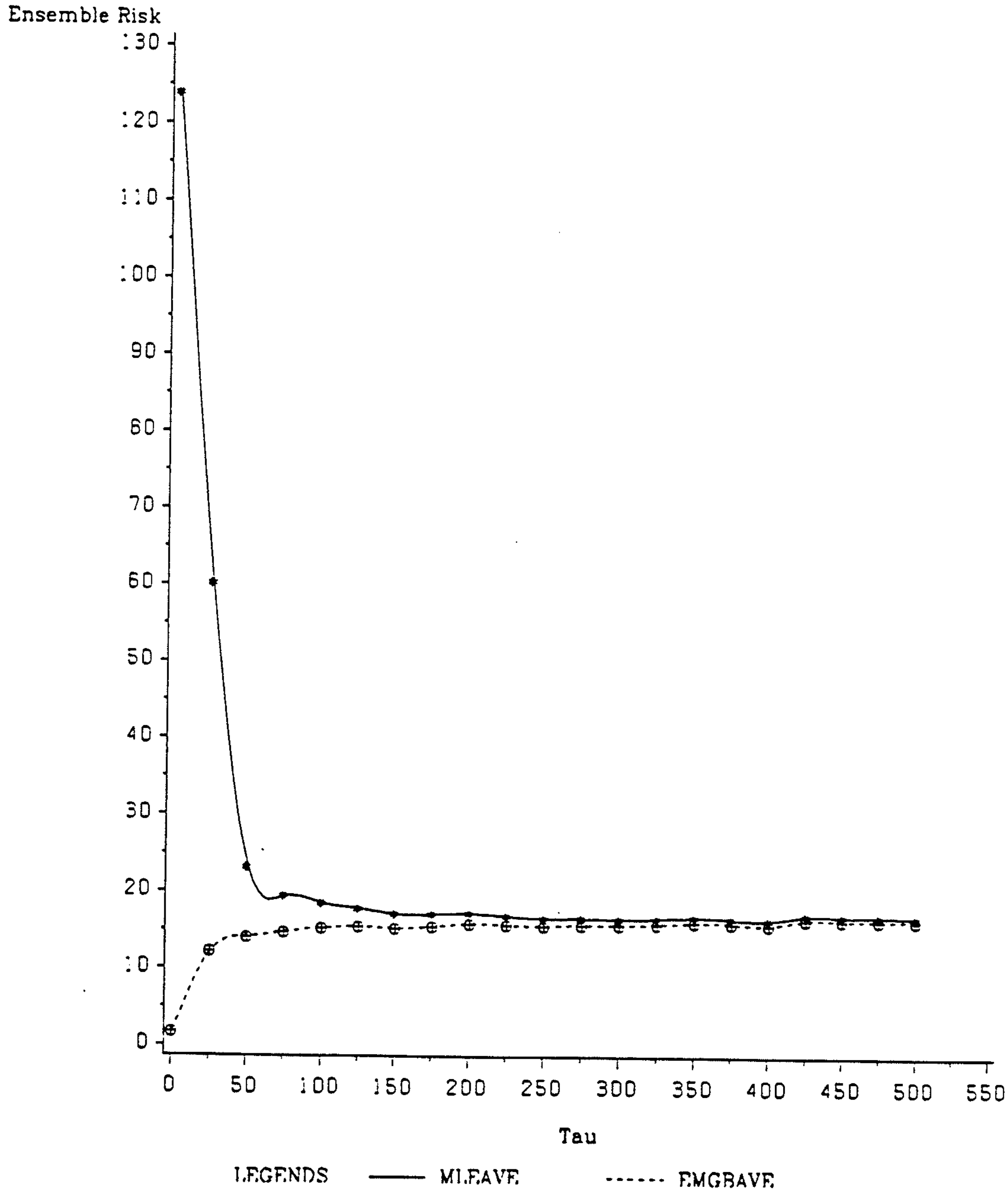
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM2
ALPHA=(0.3, 0.3, 0.3);BETA=(0.4, 0.4, 0.2)



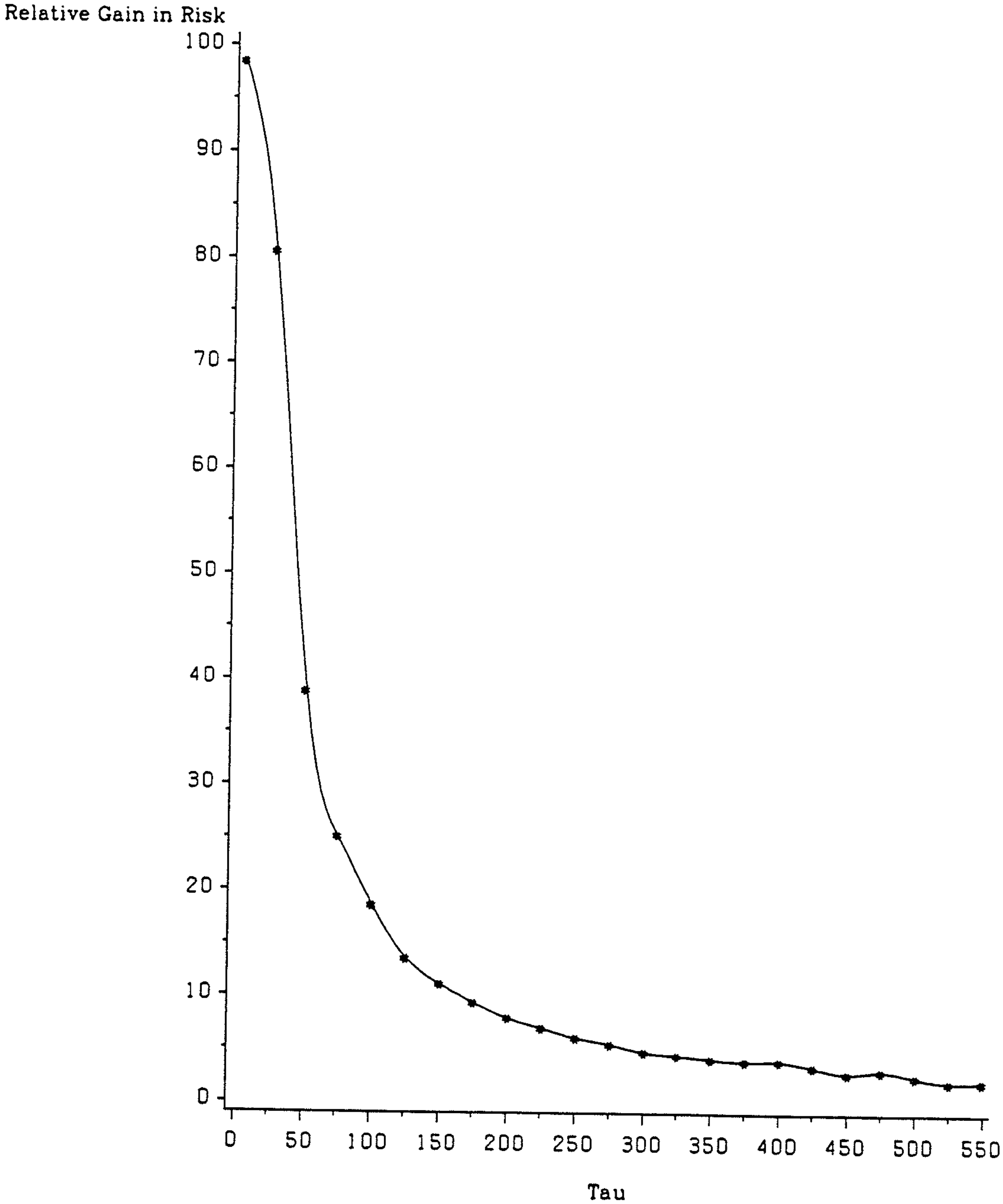
A Comparison of the Ensemble Risks of the MLE and the EMGB Estimator of a Matrix of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

MLEAVE=1/3 MLRIS1 + 2/3 MLRIS2
EMGBAVE=1/3 BARIS1+2/3 BARIS2
ALPHA=(0.3, 0.3, 0.4); BETA=(0.4, 0.4, 0.2)



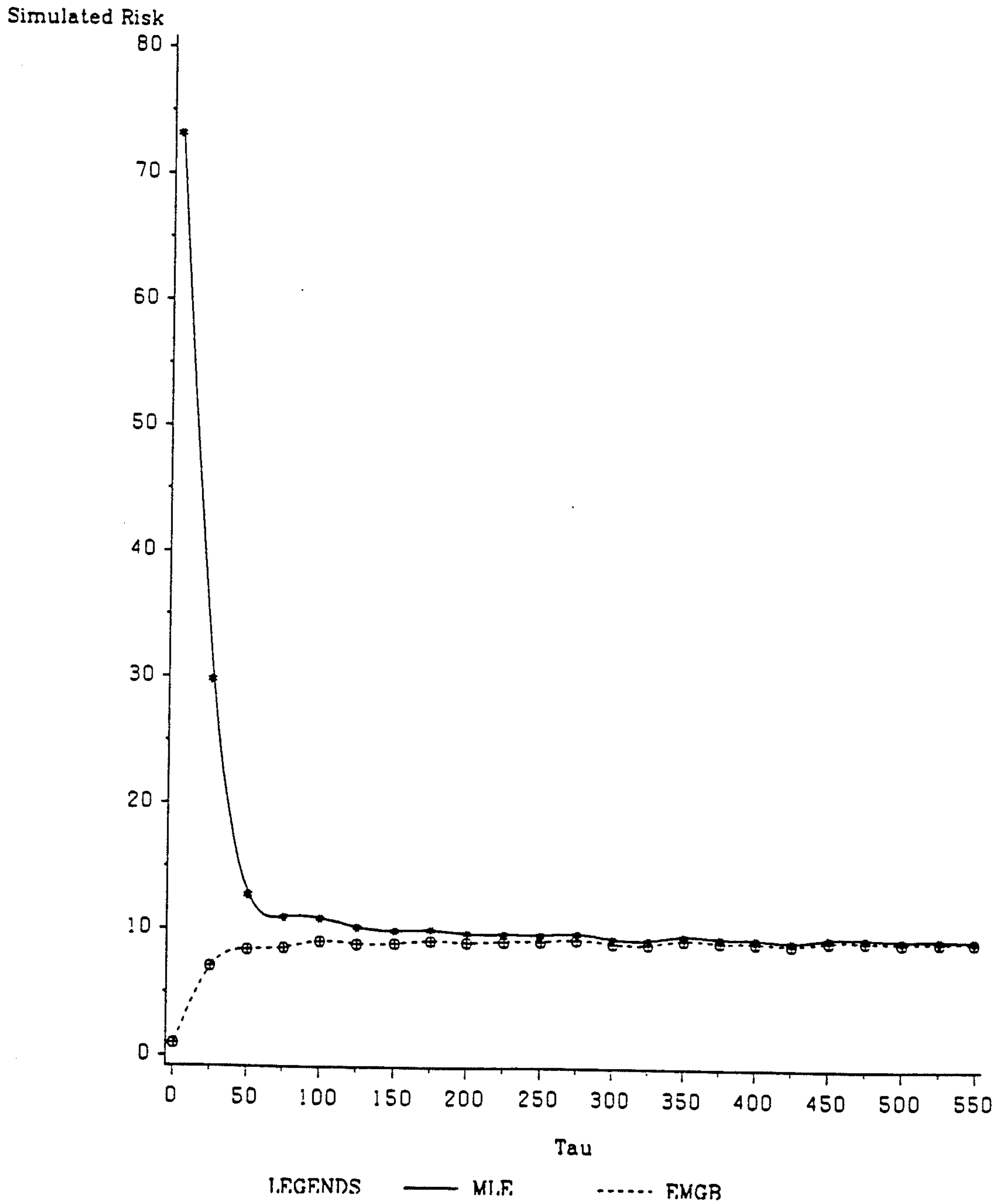
The Simulated Relative Gain in Ensemble Risk using the EMGB Estimator rather than the MLE of a Vector Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

ALPHA=(0.3, 0.3, 0.4); BETA=(0.4, 0.4, 0.2)



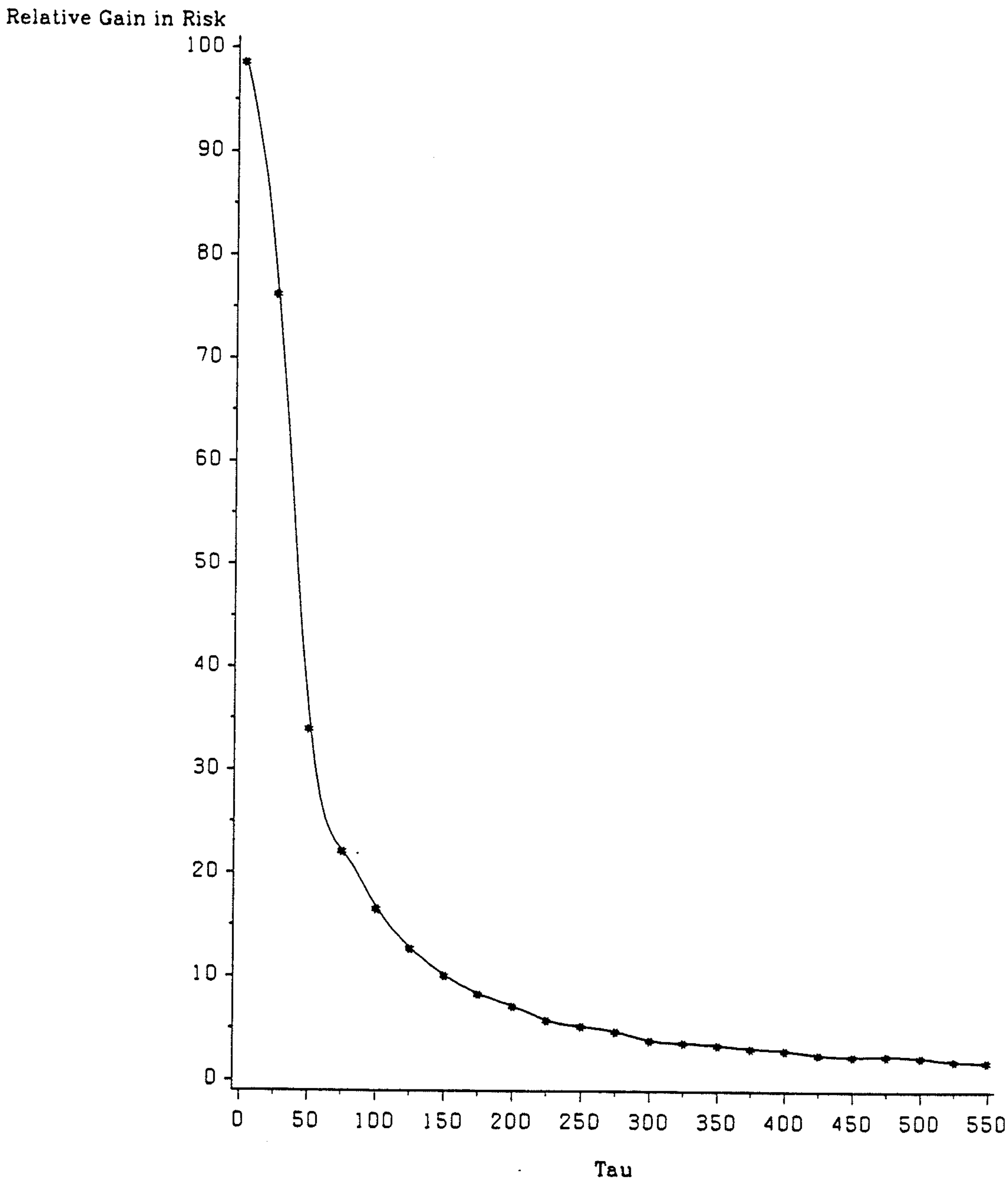
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Matrix of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1.
ALPHA=(1/3, 1/3, 1/3);BETA=(1/3, 1/3, 1/3)



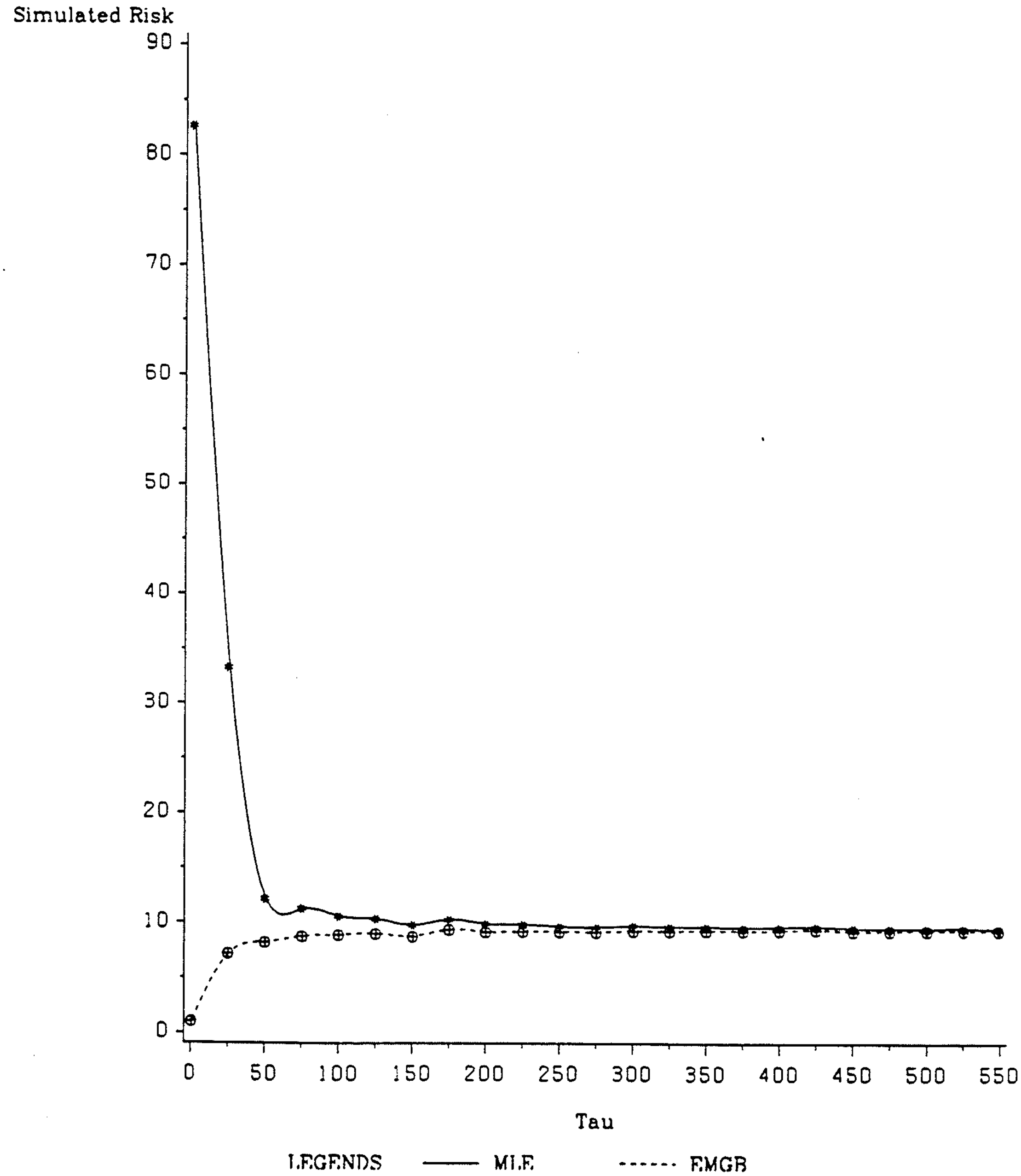
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1
ALPHA=(1/3, 1/3, 1/3); BETA=(1/3, 1/3, 1/3)



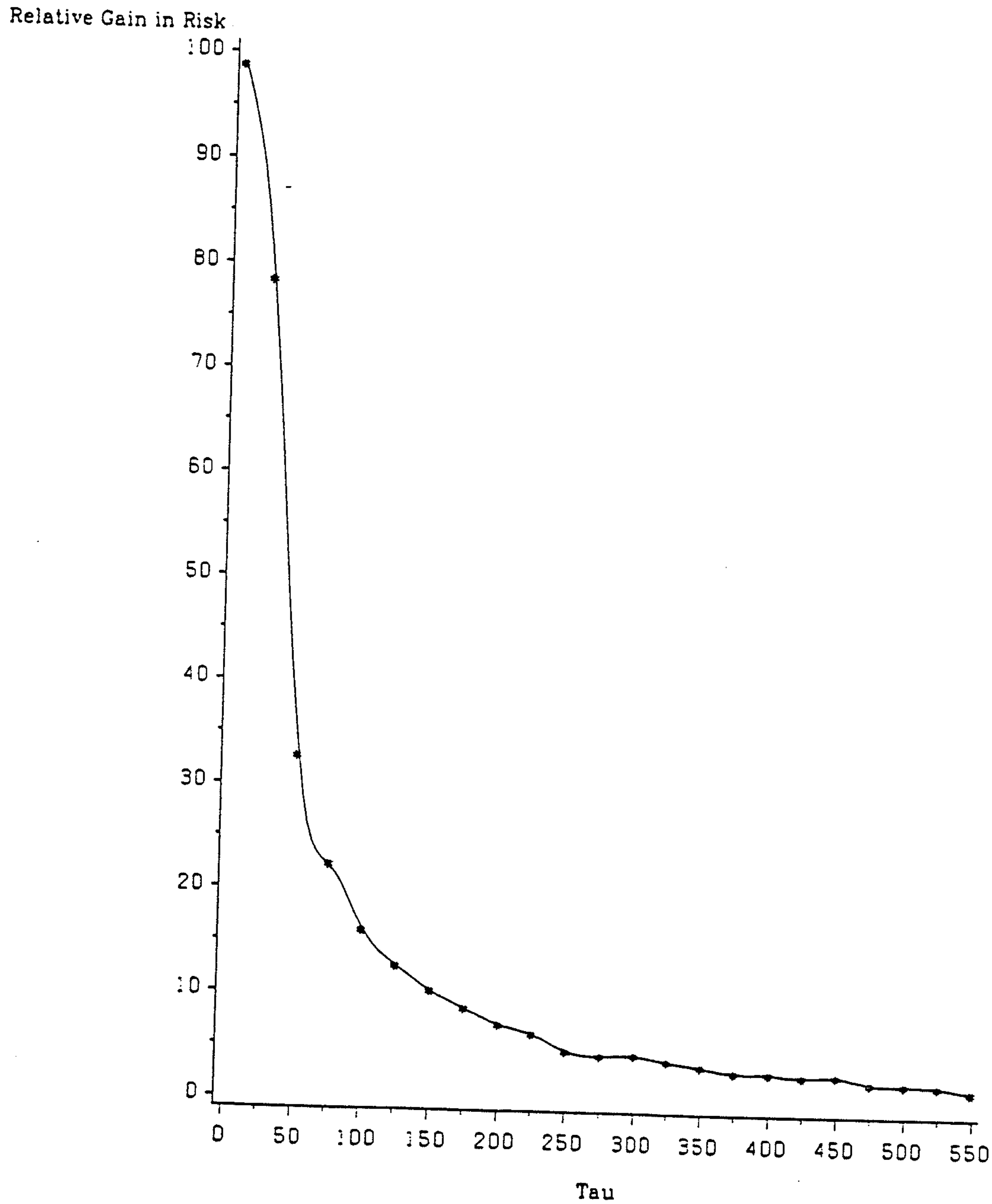
A Comparison of the Simulated Risks of the
MLE and the EMGB Estimator of a Matrix of
Poisson Means under a Multiplicative Model
for a Balanced Incomplete 3x3 Table.
Incidence Matrix is IM2.

ALPHA=(1/3, 1/3, 1/3);BETA=(1/3, 1/3, 1/3)



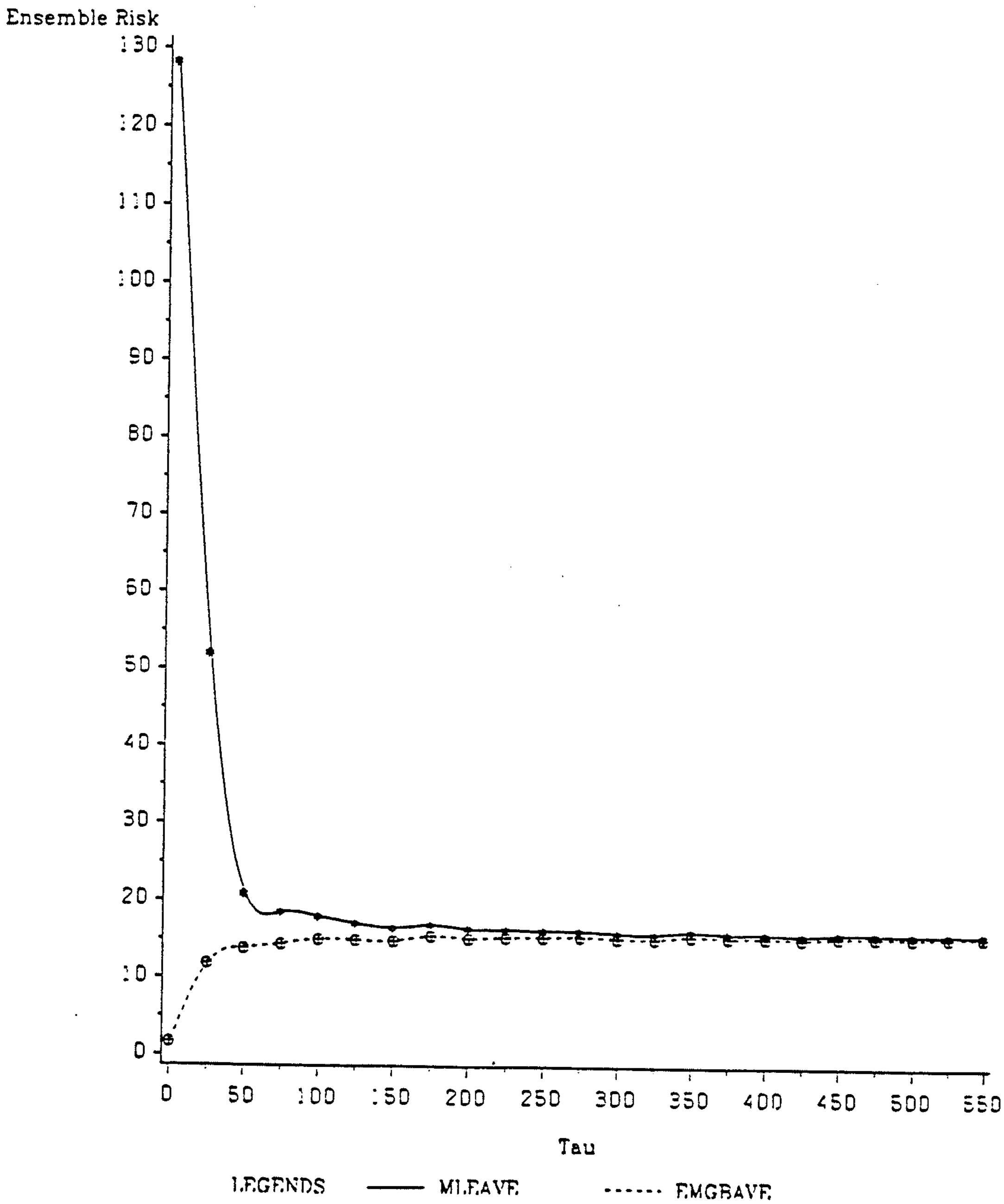
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM2
ALPHA=(1/3, 1/3, 1/3); BETA=(1/3, 1/3, 1/3)



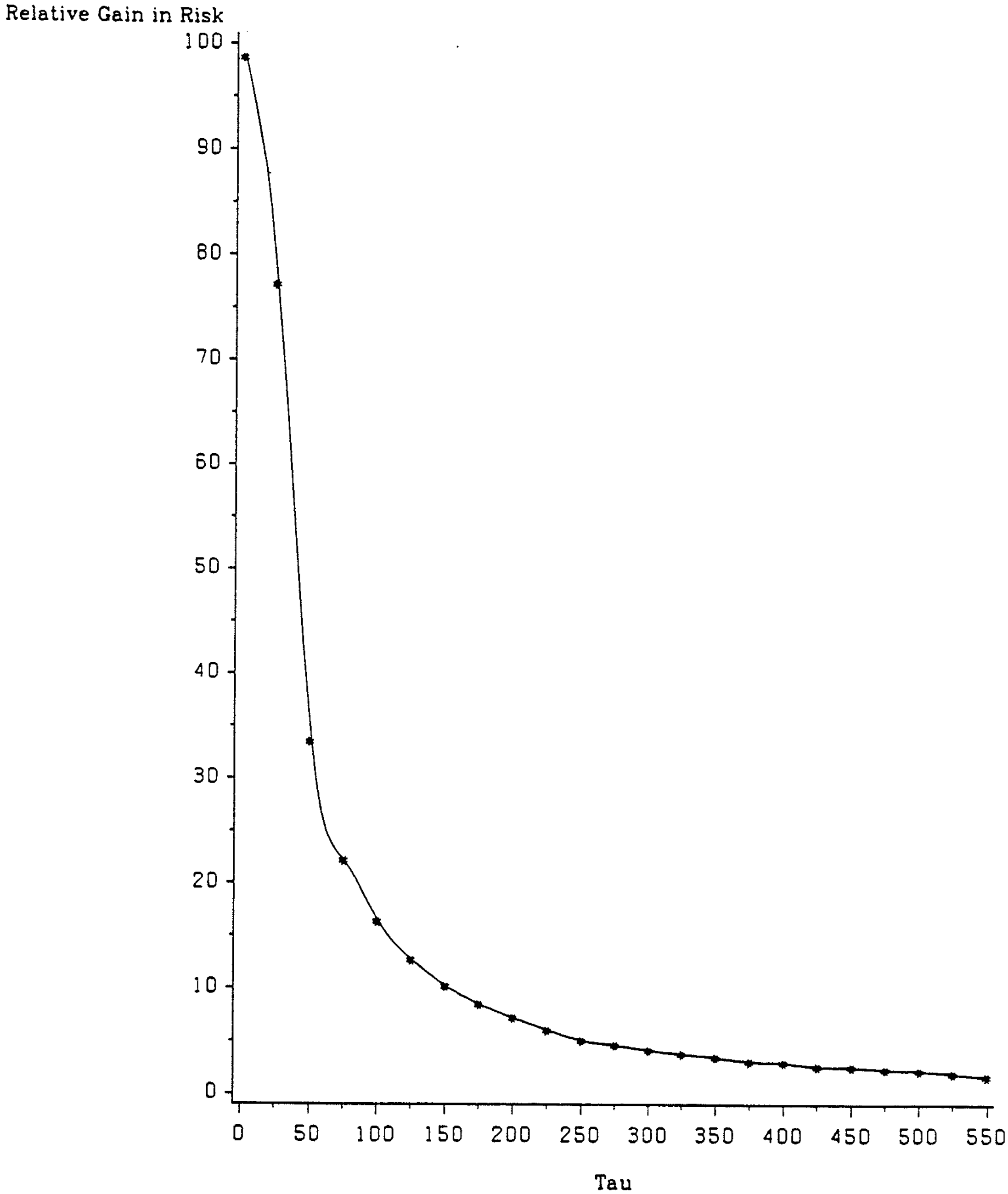
A Comparison of the Ensemble Risks of the MLE and the EMGB Estimator of a Matrix of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

MLEAVE=1/3 MLRIS1 + 2/3 MLRIS2
EMGBAVE=1/3 BARIS1+2/3 BARIS2
ALPHA=(1/3, 1/3, 1/3); BETA=(1/3, 1/3, 1/3)



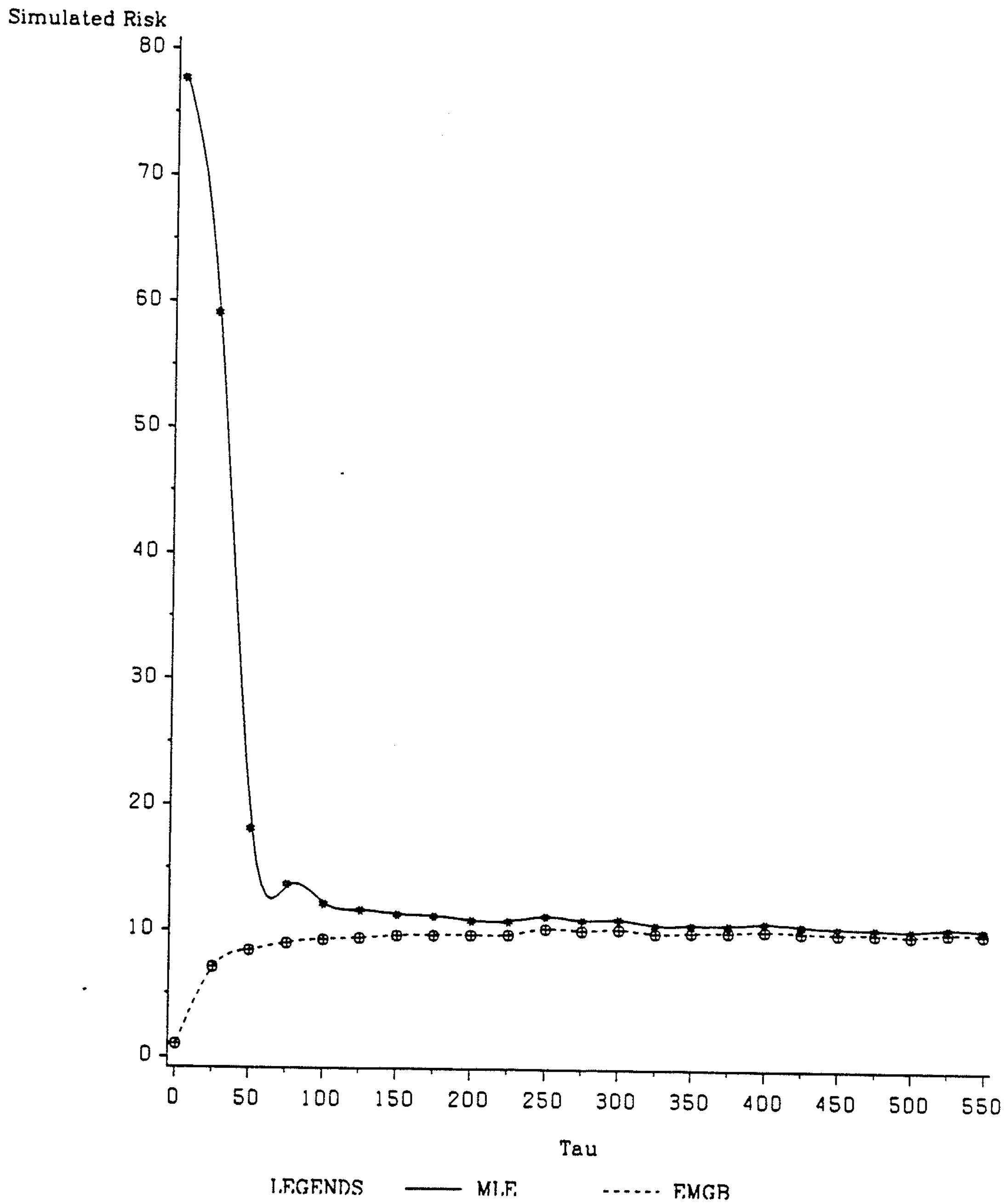
The Simulated Relative Gain in Ensemble Risk using the EMGB Estimator rather than the MLE of a Matrix Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

ALPHA=(1/3, 1/3, 1/3); BETA=(1/3, 1/3, 1/3)



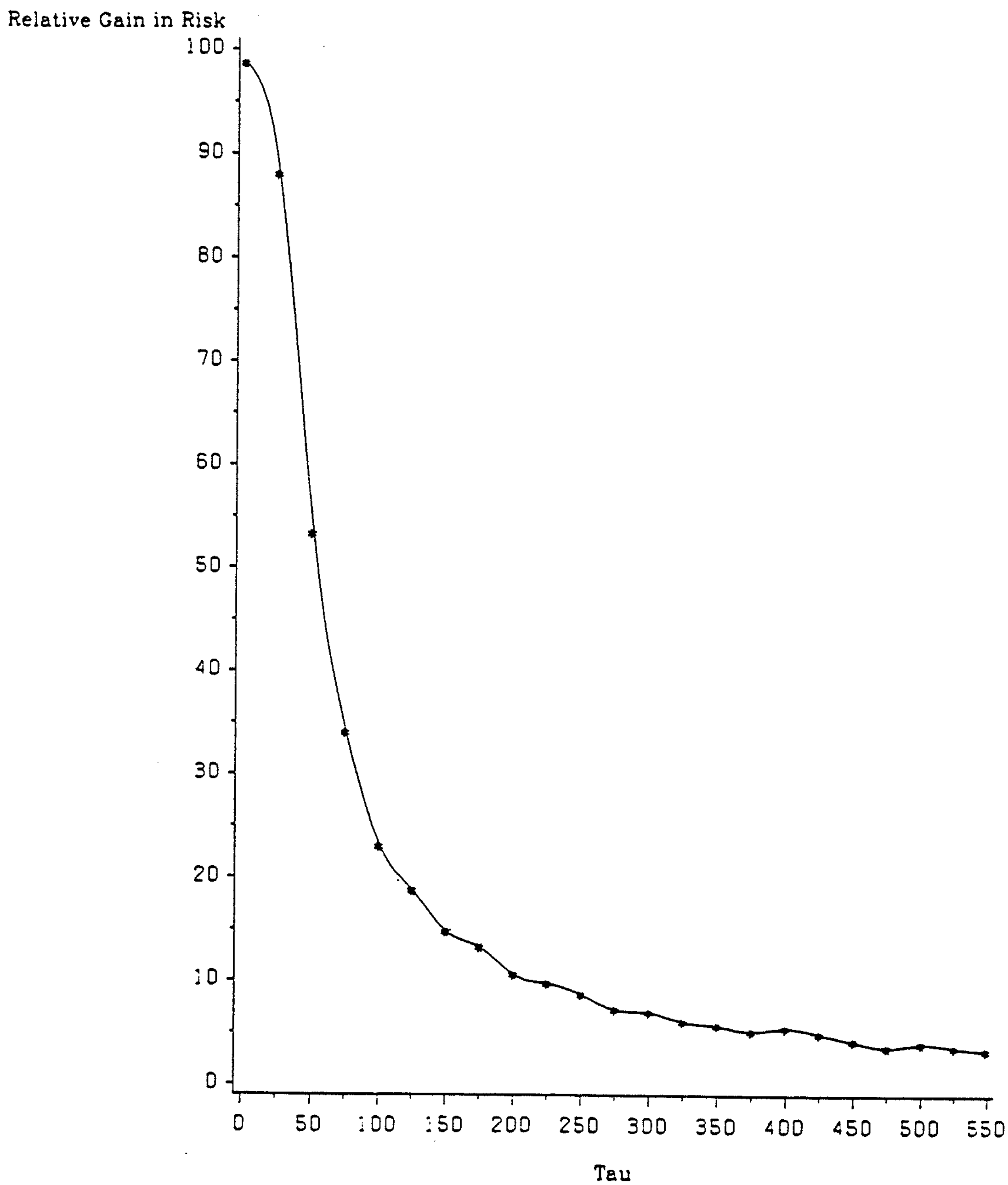
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Vector of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1.
ALPHA=(0.4, 0.4, 0.2); BETA=(1/3, 1/3, 1/3)



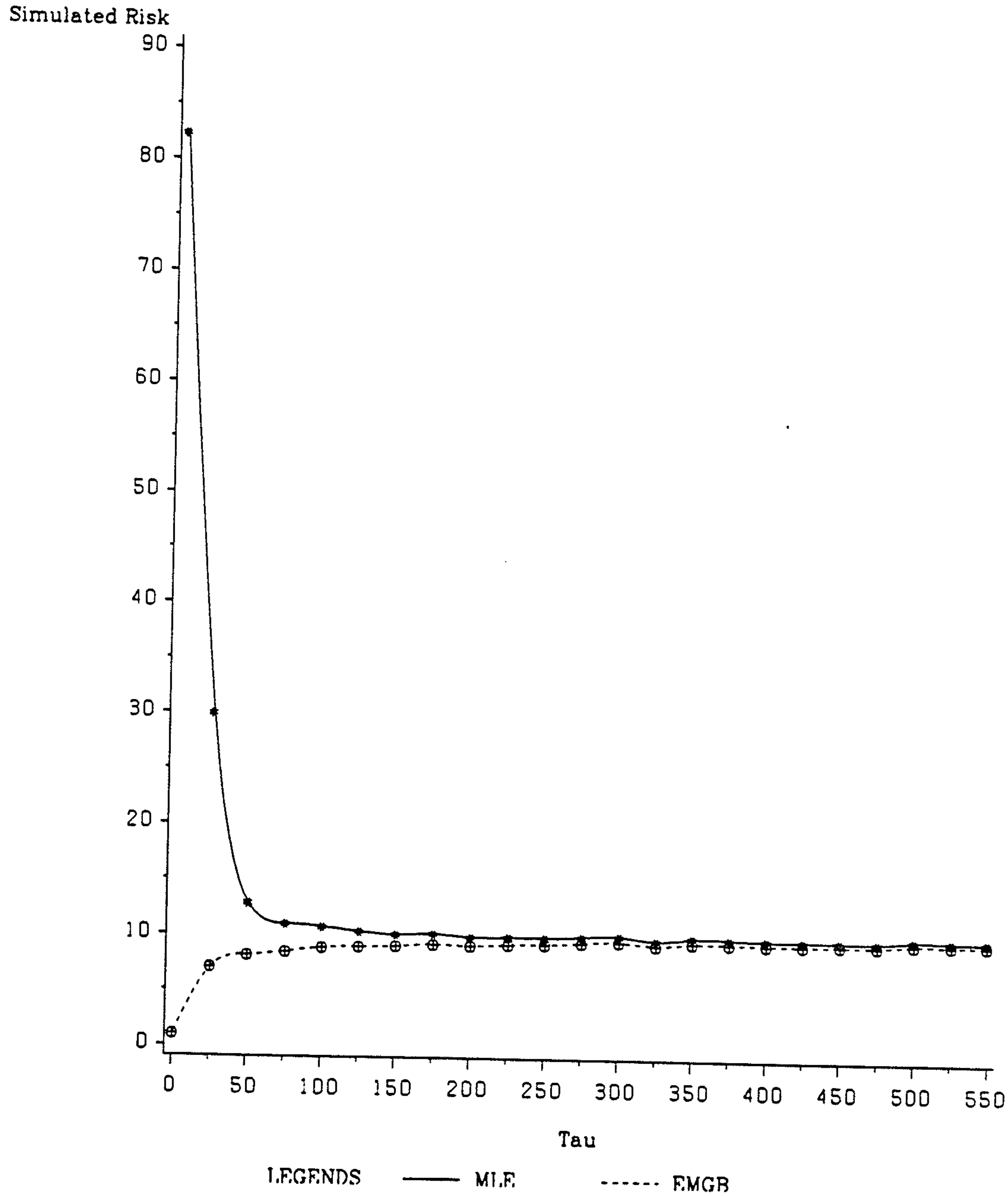
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1
ALPHA=(0.4, 0.4, 0.2); BETA=(1/3, 1/3, 1/3)



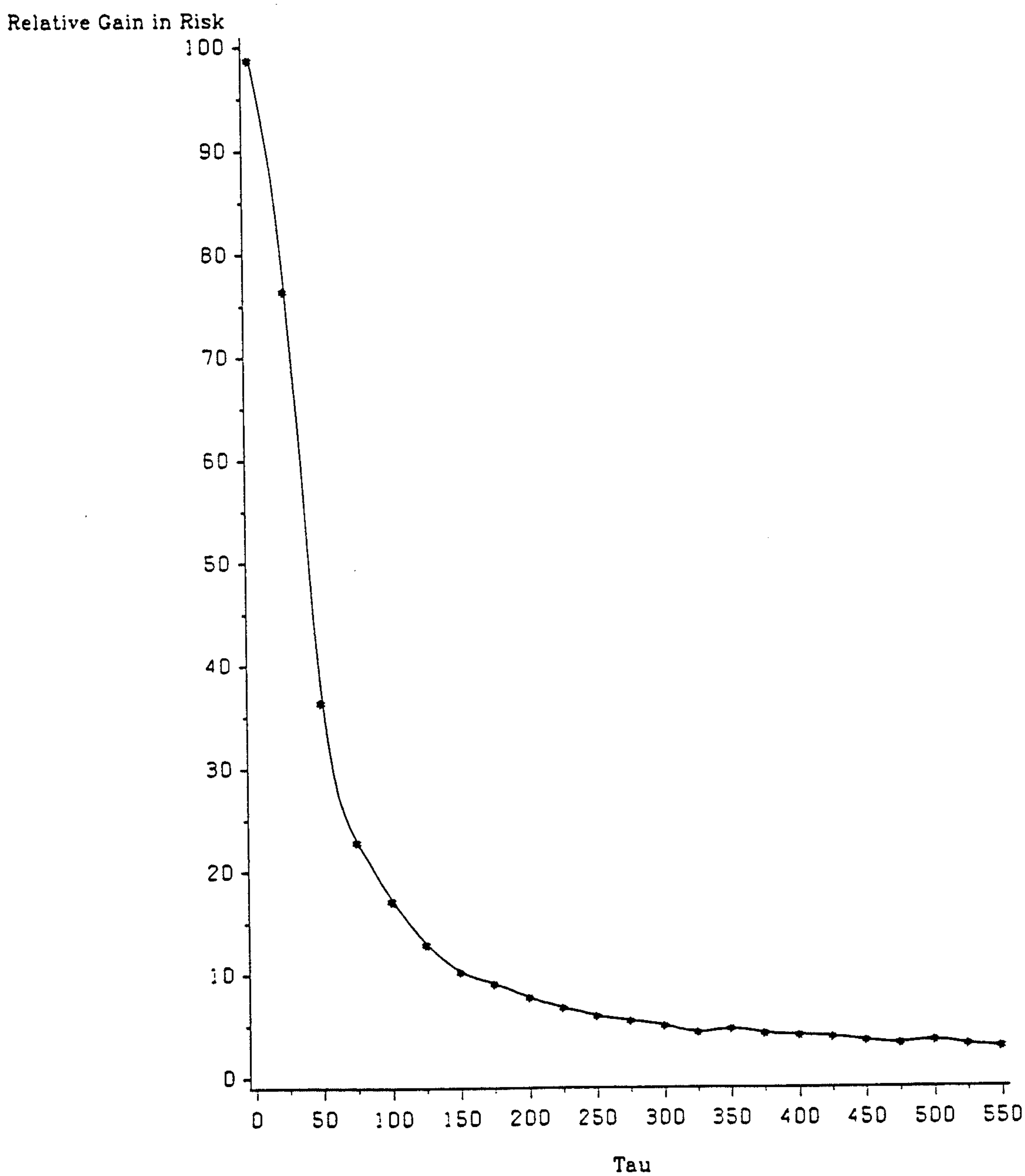
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Vector of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.
Incidence Matrix is IM2.

ALPHA=(0.4, 0.4, 0.2); BETA=(1/3, 1/3, 1/3)



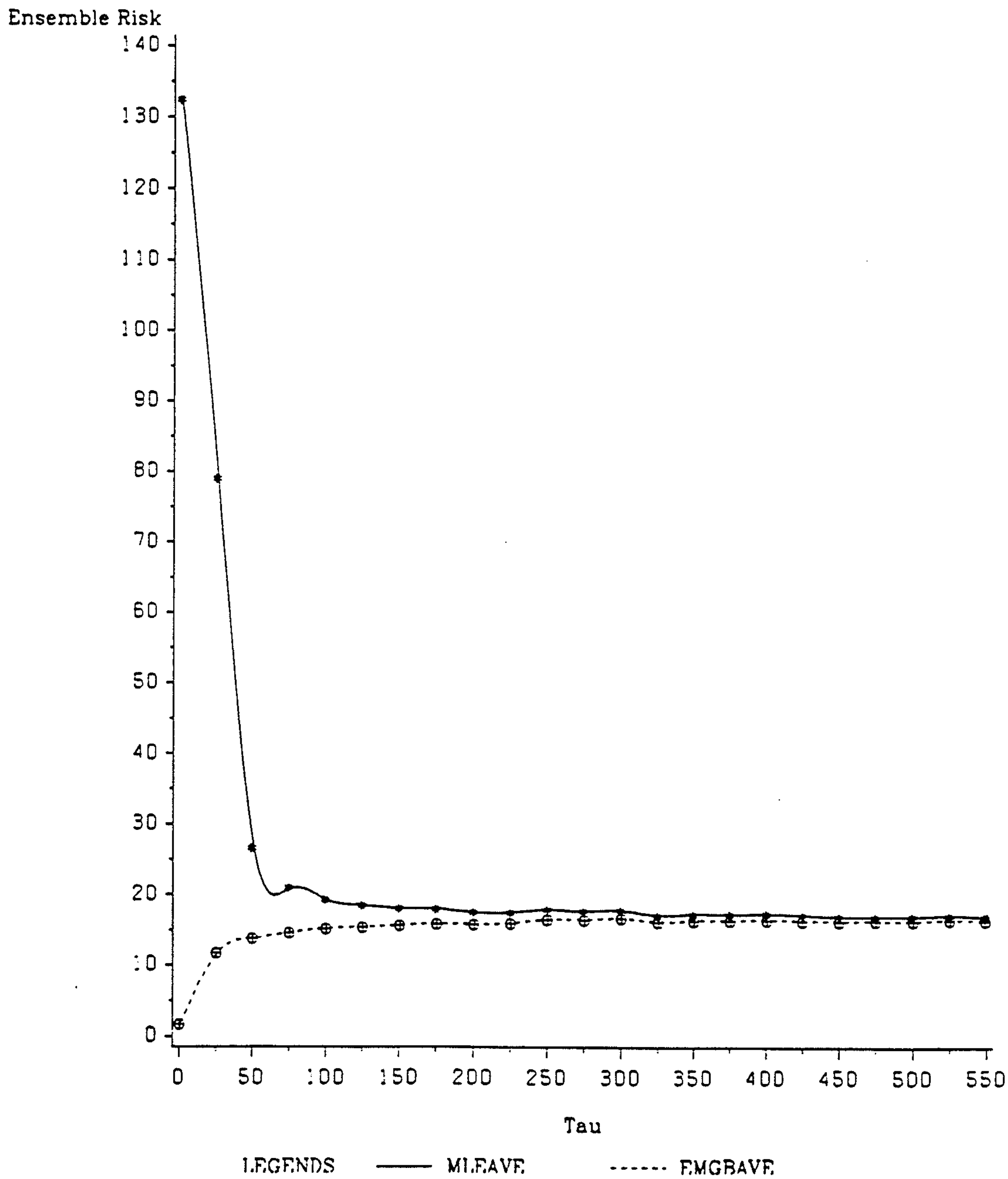
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM2
ALPHA=(0.4, 0.4, 0.2); BETA=(1/3, 1/3, 1/3)



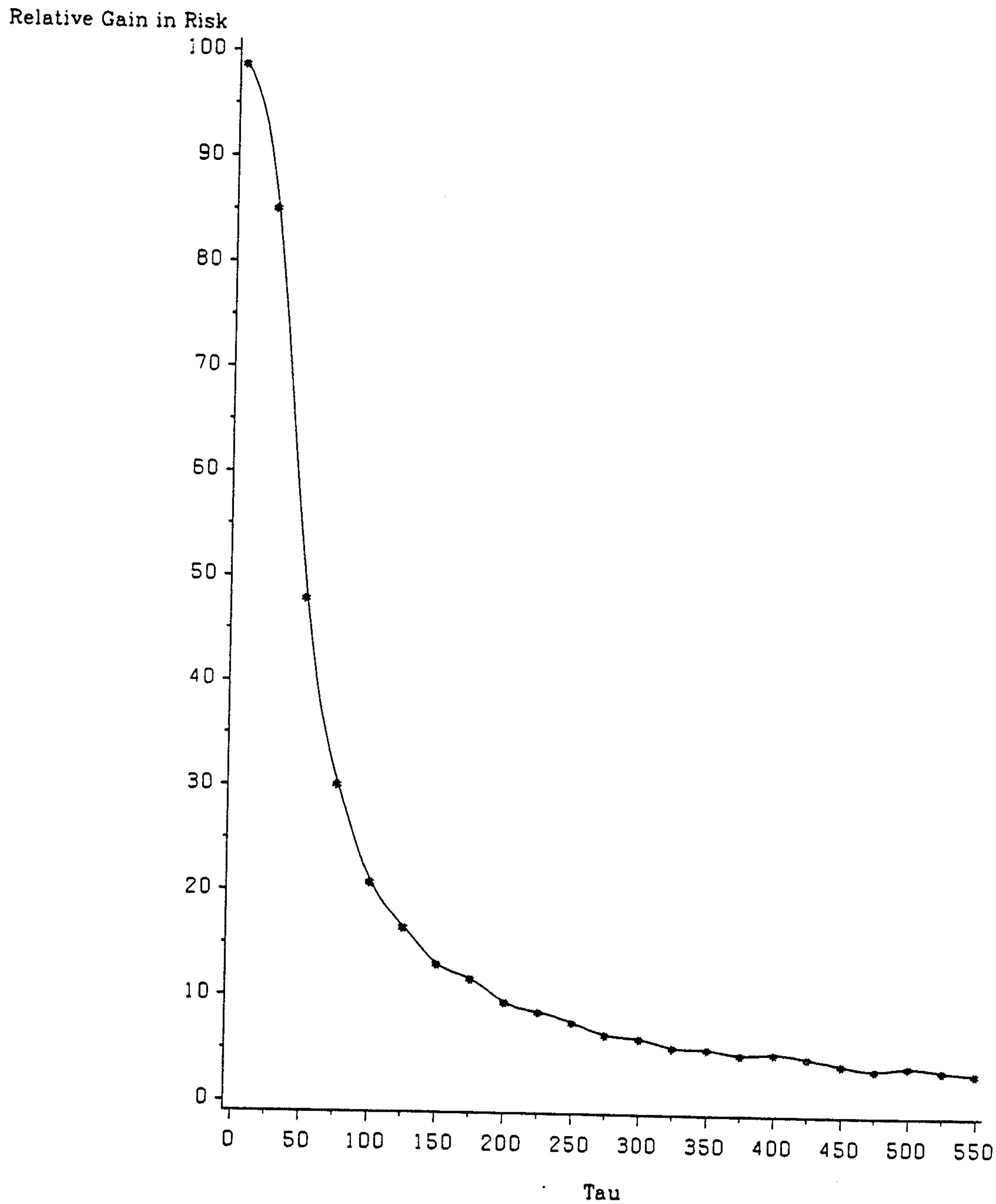
A Comparison of the Ensemble Risks of the MLE and the EMGB Estimator of a Matrix of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

MLEAVE=1/3 MLRIS1 + 2/3 MLRIS2
EMGBAVE=1/3 BARIS1+2/3 BARIS2
ALPHA=(0.4, 0.4, 0.2); BETA=(1/3, 1/3, 1/3)



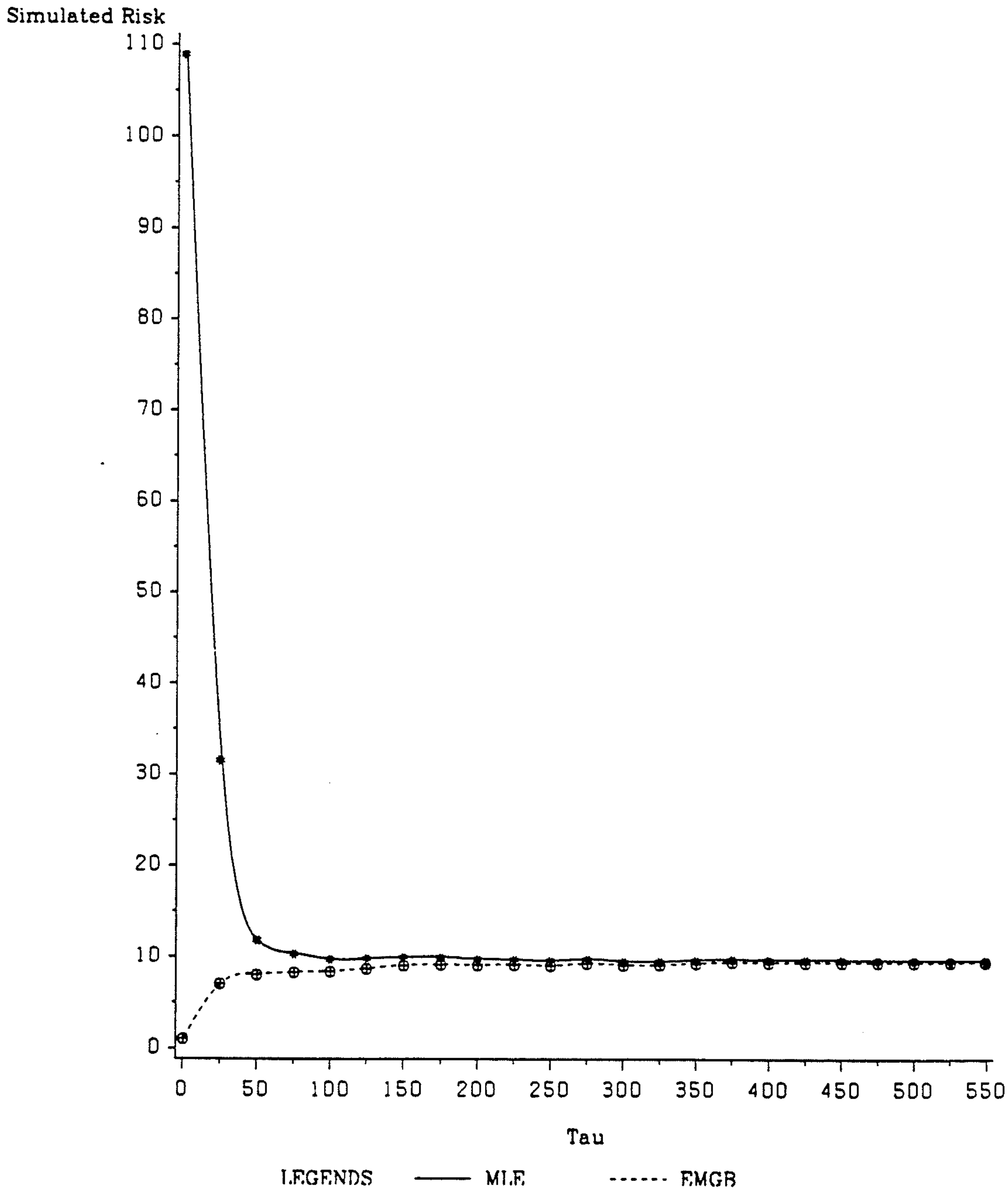
The Simulated Relative Gain in Ensemble Risk using the EMGB Estimator rather than the MLE of a Matrix Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

ALPHA=(0.4, 0.4, 0.2); BETA=(1/3, 1/3, 1/3)



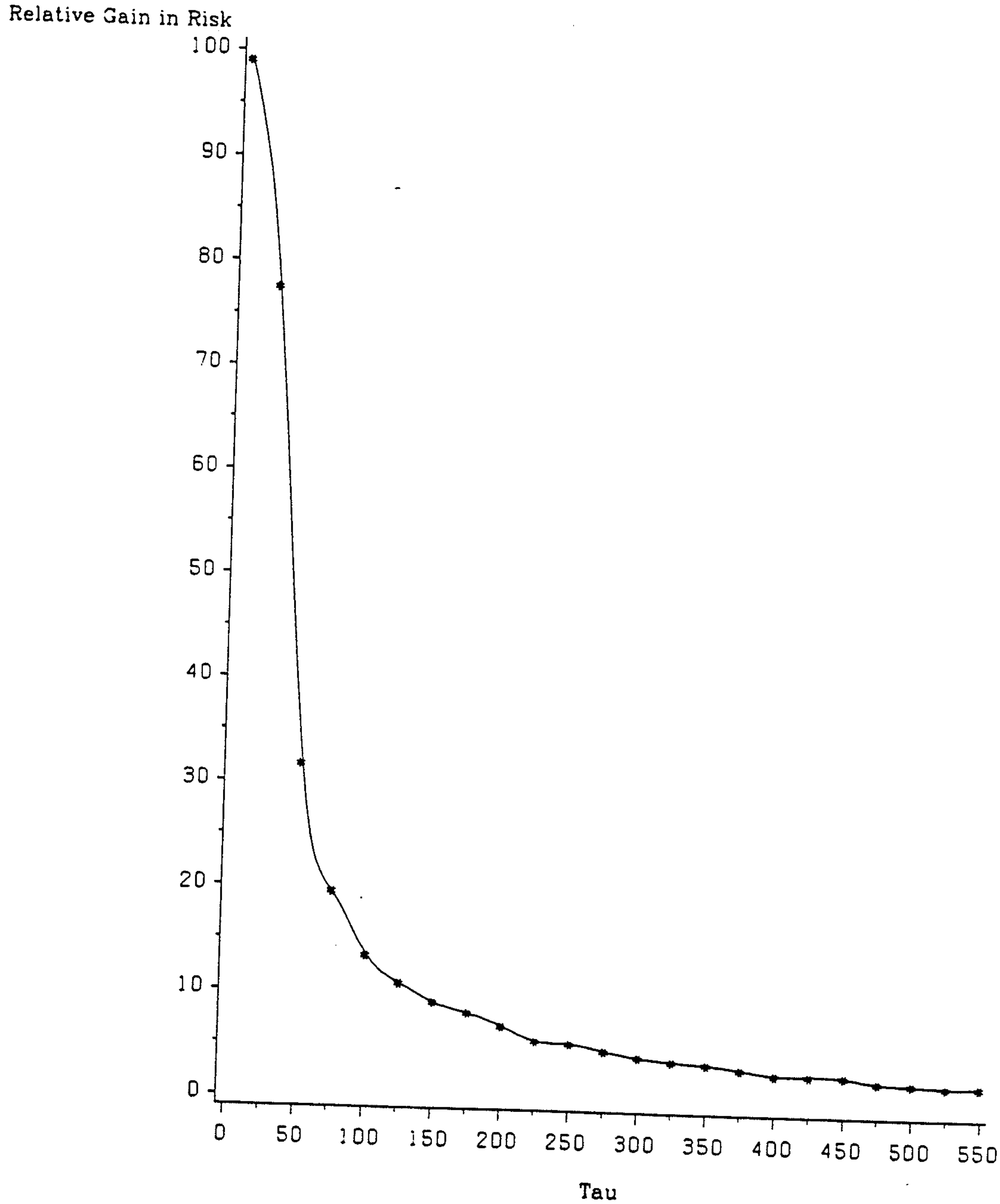
A Comparison of the Simulated Risks of the MLE and the EMGB Estimator of a Vector of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1.
ALPHA=(0.4, 0.4, 0.2); BETA=(0.4, 0.4, 0.2)



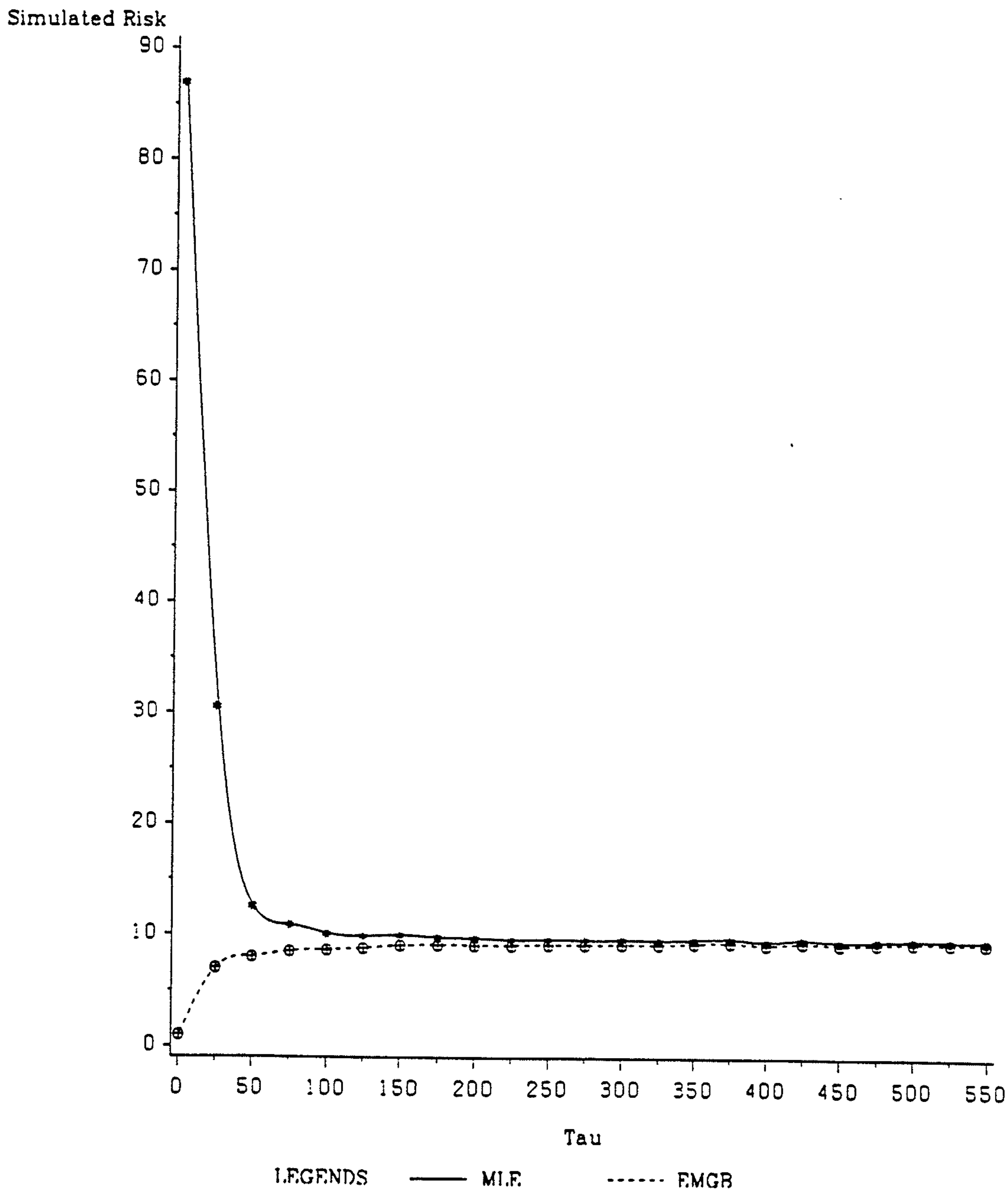
The Simulated Relative Gain in Risk using the EMGB Estimator rather than the MLE of a Matrix of Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

Incidence Matrix is IM1
ALPHA=(0.4, 0.4, 0.2);BETA=(0.4, 0.4, 0.2)



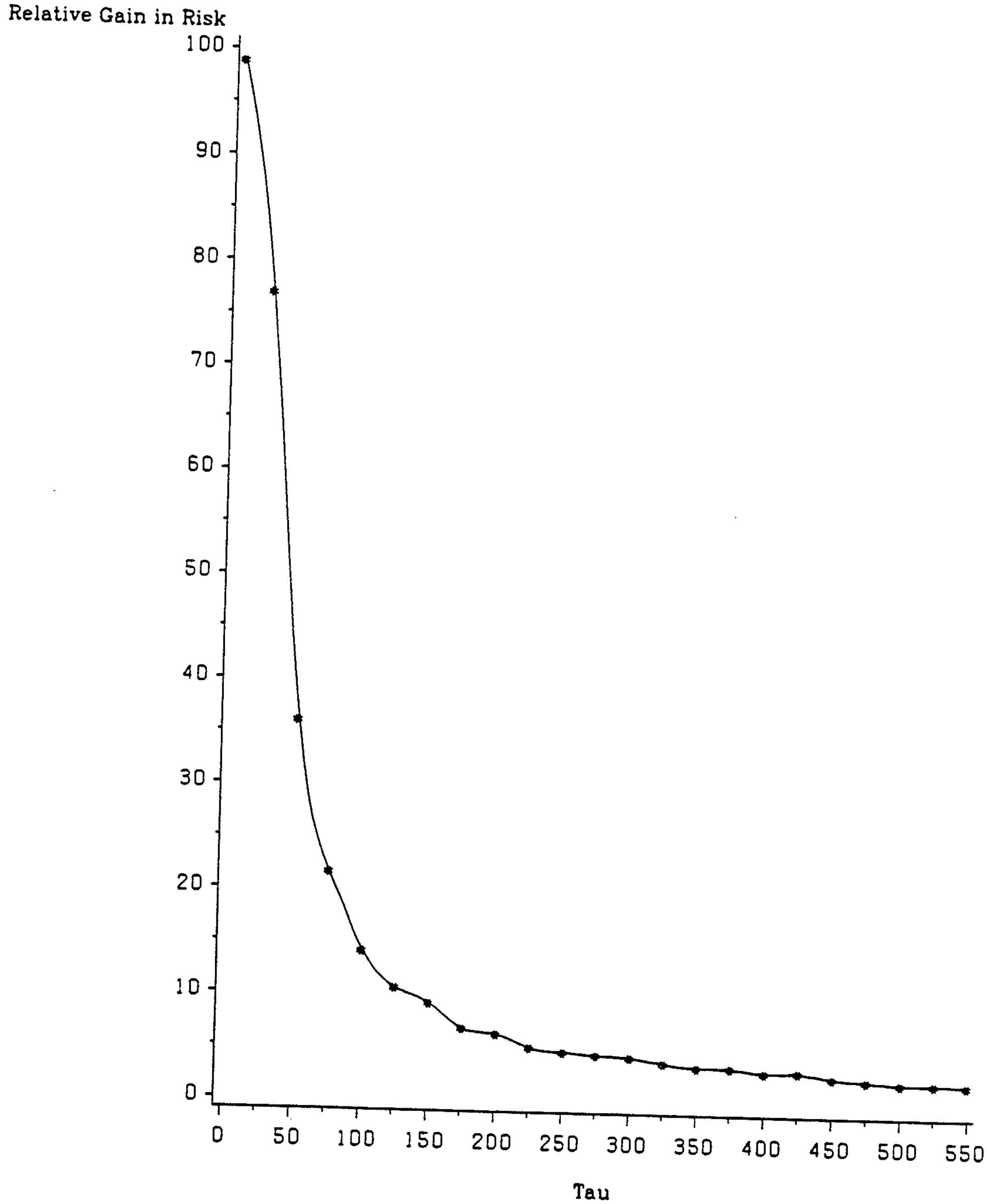
A Comparison of the Simulated Risks of the
MLE and the EMGB Estimator of a Matrix of
Poisson Means under a Multiplicative Model
for a Balanced Incomplete 3x3 Table.
Incidence Matrix is IM2.

ALPHA=(0.4, 0.4, 0.2);BETA=(0.4, 0.4, 0.2)



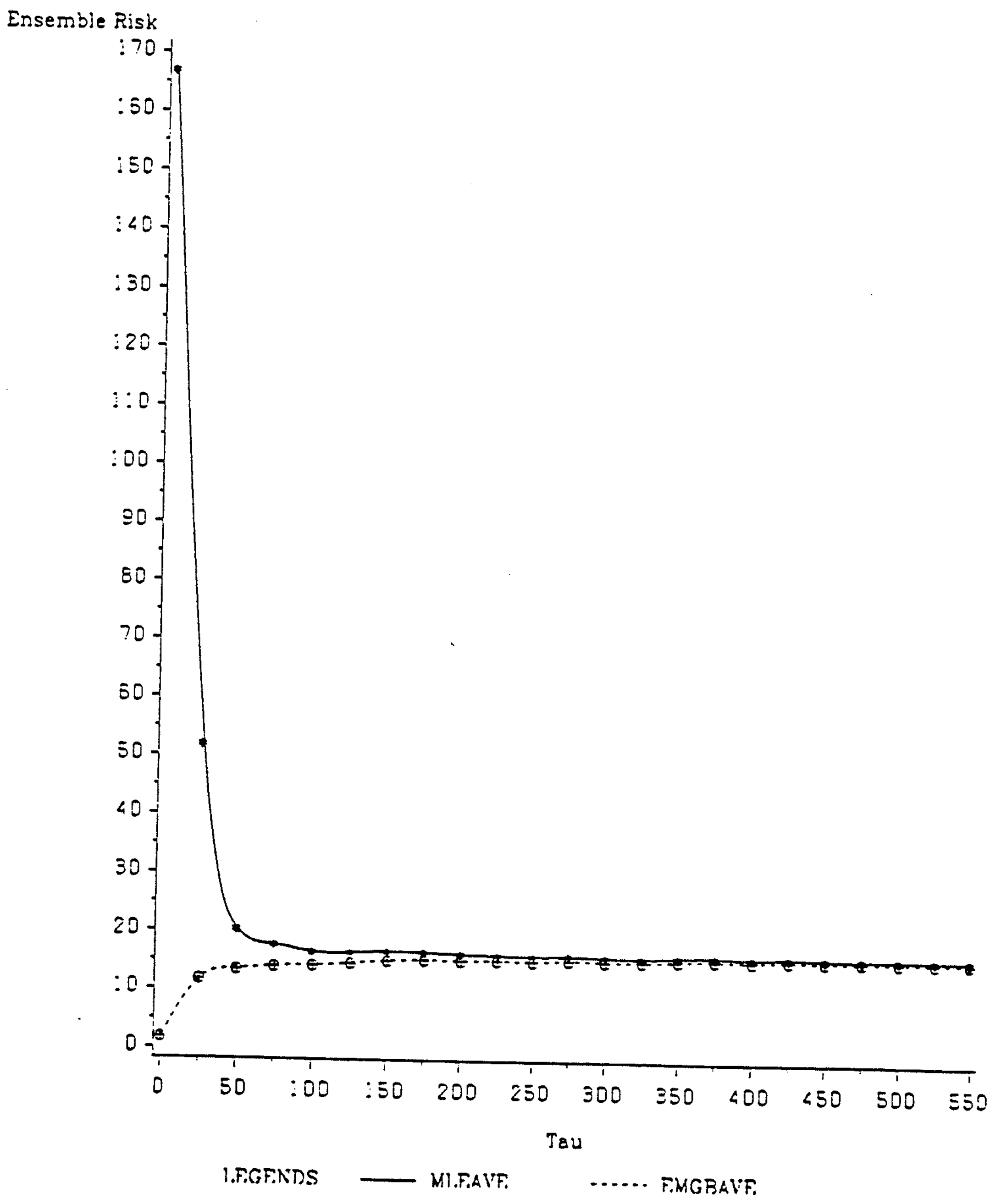
The Simulated Relative Gain in Ensemble Risk using the EMGB Estimator rather than the MLE of a Vector Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

ALPHA=(0.4, 0.4, 0.2); BETA=(0.4, 0.4, 0.2)



A Comparison of the Ensemble Risks of the MLE and the EMGB Estimator of a Matrix of Poisson Means under a Multiplicative Model for a Balanced Incomplete 3x3 Table.

MLEAVE=1/3 MLRIS1 + 2/3 MLRIS2
EMGBAVE=1/3 BARIS1+2/3 BARIS2
ALPHA=(0.4, 0.4, 0.2); BETA=(0.4, 0.4, 0.2)



The Simulated Relative Gain in Ensemble Risk using the EMGB Estimator rather than the MLE of a Matrix Poisson Means under a Multiplicative Model in a Balanced Incomplete 3x3 Table.

ALPHA=(0.4, 0.4, 0.2);BETA=(0.4, 0.4, 0.2)