BAYESIAN FIXED SAMPLE SIZE PROCEDURES FOR SELECTING THE BETTER OF TWO POISSON POPULATIONS

Saad A. Madhi **Samira Faisal Hathot**
Collage of Education, Babylon University

Abstract

In this paper an optimal (Bayesian) fixed sample size procedure for selecting the better of two Poisson populations is proposed and studied. Bayesian decision-theoretic approach with different loss functions and Gamma priors are used to construct this procedure. A suboptimal procedure that is based on posterior estimate of the parameters and a method of obtaining an approximation to the optimal procedure using Stirling's formula are also presented. Comparisons among these procedures are made using performance characteristics such as Bayes risk and the probability of correct selection of the better population.

1- Introduction

Suppose that Π_i (i=1,2) are two Poisson populations. The quality of the ith population is characterized by a positive real-valued parameter λ_i, the mean rate of occurrences in a unit of time, space, volume ...(i=1,2). The problem is to select the better of these Poisson populations on the basis of a fixed number of observations N which is partitioned into n_1 and n_2 (not necessarily are equal) the number of observations taken from populations Π_1 and Π_2 respectively.

The better population is defined to be the one with the largest mean rate of occurrences. The ranked mean rates are denoted by $\lambda_{[1]} \leq \lambda_{[2]}$, the values of $\lambda_{[i]}$ are assumed to be unknown to us. Moreover we don't know which population is associated with $\lambda_{[2]}$.

Our goal is to design fixed sample size selection procedures that enable us to select the population associated with $\lambda_{[2]}$, thus we have two-decision problem.

The statistical formulation as stated above is typical of many well-known practical problems in many situations in real life. For illustration, suppose we would like to compare the performance of two different machines that produce sheet of metal in continuous rolls, data may be collected on the number of defects x, that are spotted in several time segments of, say, 5 minutes duration. It is reasonable to assume that X has a Poisson distribution with parameter λ_i, (i=1,2) for the two machines. Our goal is to choose the machine that has high rate of defects. Further examples on the situations where the Poisson model applied and the problem of practical interest is to select the better of two Poisson populations may be found in Gibbons, Olkin and Sobel (1977).

The following experimental conditions should be met.

1. The observations produced by each population are independent each other.
2. λ_1 and λ_2, the mean rates of occurrences are constant during the experiment.

Most of the previous work on the Poisson selection problem seems to be limited to the classical approaches : the indifference zone approach and the subset selection approach.

Alam (1971) considers three procedures for selecting the process corresponding to the largest mean rate of occurrence using the indifference zone approach.

Goel (1972) considered inverse sampling rule to define a selection procedure for selecting a subset containing the population associated with largest value of mean rate.

Alam and Thompson (1973) proposed a procedure to select simultaneously the population associated with largest parameter and estimate this parameter.

Huang and Wen Tao (1973) suggested two selection procedures for choosing a random subset guaranteed to contain a 'best' population with a preassigned probability.

Gupta and Huang (1975) considered the selection from k Poisson populations of a variable size subset including that population with the largest parameter when fixed (equal) sample sizes are taken. In Gupta and Wong (1977), the problem of selecting a subset of k different Poisson processes including the best which is associated with the largest value of the mean rate is discussed.

Gupta, Leong and Wong (1978) proposed modified procedures to Gupta and Huang (1975) with tabulated approximate values (determined by numerical methods), for more details of this procedures, see Gibbons, Olkin and Sobel (1977).

Some contributions such as Nelson and Hong (2003) presented an indifference zone selection procedure which is sequential and has minimum number of switches. Nelson and Pichitlamken (2001) propose fully sequential indifference zone selection procedures.

Some contributions using Bayesian approach have been made by Madhi (1986) who presented Bayesian sequential procedures for Binomial and Multinomial selection problems. Bland and Bratcher (1968), Chick (1997). Chen (1995) and Chen et al. (1996) have formulated the R&S problem as multi-stage optimization problem.

In this paper, a Bayesian fixed sample size for Poisson selection procedure is proposed and studied using Bayesian decision theoretic formulation. The procedure is optimal in the sense that it minimizes the average risk with respect to certain prior distributions on the parameters, namely the family of Gamma distributions. Suboptimal procedures are also suggested. Results using Bayes risk and the probability of correct selection of the better population are also presented.

The paper is organized as follows : in section 2 we present an optimal (Bayesian) fixed sample size scheme for selecting the better of two Poisson populations using Gamma priors and various loss functions. The two-decision Poisson selection formulation is given in subsection 2.1. subsection 2.2 contains the Bayesian selection procedure (opm). In subsection 2.3 we drive the posterior expected losses of making decision d_1 and d_2 for constant, linear and quadratic loss functions. In section 3 we present the suboptimal selection procedure (subopm1) as an approximation to opm procedure. In section 4 we describe the suboptimal selection procedure (subopm2) based on posterior Bayes estimate of the parameter. Comparisons of the schemes using posterior expected loss-under linear loss function is given in section 5. Some Monte Carlo studies concerning the probability of correct selection of the better population are given in section 6. Conclusions and future research are given in section 7.

2. The Bayesian (Optimal) selection procedure (Opm)

2.1 The Bayesian Decision-Theoretic Formulation

let Π_1 and Π_2 be two Poisson populations with unknown mean rates of occurrences λ_1 and λ_2 respectively, and consider the following two - decision problem with decisions :

$$d_1 : \lambda_1 < \lambda_2$$

and ....(1)

$$d_2 : \lambda_1 \geq \lambda_2$$

Corresponding to the two decision problem the parameter space $\Omega=\{(\lambda_1,\lambda_2): 0\leq\lambda_1<\infty,\ 0\leq\lambda_2<\infty\}$ is divided into two disjoint sets : $\Omega_1=\{(\lambda_1,\lambda_2): 0\leq\lambda_1<\lambda_2<\infty\}$ and $\Omega_2=\{(\lambda_1,\lambda_2): 0\leq\lambda_2\leq\lambda_1<\infty\}$.

To obtain, an explicit Bayes rule (Bayes selection procedure) for this two decision problem we must specify loss function and prior distributions. Suppose the loss functions proposed are as follows :

$$L_1(\lambda_1,\lambda_2;d_1)=\begin{cases}0 & if\ (\lambda_1,\lambda_2)\in\Omega_1\\ k_1|\lambda_1-\lambda_2|^r & if\ (\lambda_1,\lambda_2)\in\Omega_2\end{cases} \quad ....(2)$$

and

$$L_2(\lambda_1,\lambda_2;d_2)=\begin{cases}k_2|\lambda_1-\lambda_2|^r & if\ (\lambda_1,\lambda_2)\in\Omega_1\\ 0 & if\ (\lambda_1,\lambda_2)\in\Omega_2\end{cases} \quad ....(3)$$

Where $r=0,1,2$ gives the types of loss function, which are constant, linear and quadratic respectively. L_i $(i=1,2)$ is the loss function corresponding to decision d_i, and k_1 and k_2 are positive constants (the same for each pair of $\lambda's$).

The Bayesian approach requires that we specify a prior density function $\pi(\lambda_i)$, $i=1,2$, expressing our beliefs about λ_i before we obtain data. From a mathematical point of view, it would be very convenient if λ_i is assigned a prior distribution which is a member of the conjugate family, in this case is the family of Gamma distributions. Accordingly let λ_i, $(i=1,2)$ is assigned Gamma prior distribution with parameters $n_i', t_i', \lambda_i, Gamma\ (n_i', t_i')$. The normalized density function (Raiffa and Schlaifer (1968)) is given by :

$$\pi(\lambda_i)=\frac{(t_i')^{n_i'}}{\Gamma(n_i')}e^{-t_i'\lambda_i}\,\lambda_i^{n_i'-1} \quad ,n_i'>0, t_i'>0, \lambda_i>0 \quad ....(4)$$

Where

n_i' : the number of independent events occurred in a unit of time.

t_i' : time.

λ_i : the mean rate of occurrence

If $\underline{x}_i=(x_{i1},x_{i2},\ldots,x_{in})$, $(i=1,2)$ be a random sample from population Π_i, then the likelihood function is given by

$$f(\underline{x}_i / \lambda_i) = \frac{e^{-n\lambda_i} \lambda_i^{\sum_{j=1}^{n} x_{ij}}}{\prod_{j=1}^{n} x_{ij}!} \qquad \ldots\ldots\ldots\ldots(5)$$

The posterior density function is derived from the prior density function (4) and the assumed sampling model (5) by means of Bayes theorem as follows :

$$\pi(\lambda_i / \underline{x}_i) = \frac{f(\underline{x}_i / \lambda_i)\pi(\lambda_i)}{\int_{\lambda_i} f(\underline{x}_i / \lambda_i)\pi(\lambda_i)d\lambda_i}$$

where

$$\int_{\lambda_i} f(\underline{x}_i / \lambda_i)\pi(\lambda_i)d\lambda_i = \frac{(t_i')^{n_i'}\Gamma(n_i' + \sum_{j=1}^{n} x_{ij})}{\prod_{j=1}^{n} x_{ij}!\Gamma(n_i')(n + t_i')^{n_i' + \sum_{j=1}^{n} x_{ij}}}$$

let $t_i'' = t_i' + n$ and $n_i'' = n_i' + \sum_{j=1}^{n} x_{ij}$

then

$$\pi(\lambda_i / \underline{x}_i) = \frac{(t_i'')^{n_i''}}{\Gamma(n_i'')} e^{-t_i''\lambda_i} \lambda_i^{n_i''-1}$$

As the Gamma family is conjugate with the Poisson sampling, it is unnecessary to revise a Gamma prior on the basis of a sample from a Poisson population using Bayes' theorem. Given the prior distribution and sample results, we need simply note that

$$t_i'' = t_i' + n \quad \text{and} \quad n_i'' = n_i' + \sum_{j=1}^{n} x_{ij} \quad (i=1,2),$$

n is the number of observations taken from each population (N =2n is the total number of observations taken from both populations).

The foregoing is all that need in order to obtain a Bayes rule (Bayes selecting procedure) for the component two-decision problem.

.2 The Procedure (opm)

For the two-decision problem considered above, the Bayesian selection procedure is given as follows :

make decision d_1: $\lambda_1 < \lambda_2$ that is selecting Π_2 as the better population if $R_1(\lambda_1,\lambda_2;d_1) < R_2(\lambda_1,\lambda_2;d_2)$

and

make decision d_2 : $\lambda_1 \geq \lambda_2$ that is selecting Π_1 as the better population if $R_1(\lambda_1,\lambda_2;d_1) \geq R_2(\lambda_1,\lambda_2;d_2)$

Where $R_i(\lambda_1,\lambda_2;d_i)$, $(i=1,2)$ is the posterior expected loss for the decision d_i calculated as follows :

$$R_i(\lambda_1,\lambda_2;d_i)=E_{\pi(\lambda_1,\lambda_2/n''_1,t''_1,n''_2,t''_2)}[L_i(\lambda_1,\lambda_2;d_i)] \ \ , \ i=1,2$$

where $\pi(\lambda_1,\lambda_2 / n''_1,t''_1,n''_2,t''_2)$ on the expectation sign is the joint posterior of λ_1 and λ_2 with respect to which the expectation is being performed.

The optimal posterior expected loss using opm is $R(\lambda_1,\lambda_2) = min(R_1,R_2)$

2.3 The Posterior Expected Loss for Some Loss Functions

In this subsection, we derive the posterior expected losses of making decision d_1 and d_2 for constant, linear and quadratic loss functions , suppose the losses in making decision d_1 and d_2 are constant, then the posterior expected losses are defined as follows :

$$R_1(\lambda_1,\lambda_2;d_1) = k_1\left[1-\frac{(t''_1)^{n''_1}}{\Gamma(n''_1)}\sum_{j=0}^{n''_2-1}\frac{(t''_2)^j\,\Gamma(n''_1+j)}{j!(t''_1+t''_2)^{n''_1+j}}\right]$$

Where $\Gamma()$ is a gamma function with a similar form for d_2, $R_2(\lambda_1,\lambda_2;d_2)$.

If the losses are linear, then the posterior expected losses will be :

$$R_1(\lambda_1,\lambda_2;d_1)=k_1\left[\frac{n''_1}{t''_1}-\frac{n''_2}{t''_2}+\frac{(t''_1)^{n''_1}(t''_2)^{n''_2-1}(n''_1+n''_2-1)!}{(n''_1-1)!(n''_2-1)!(t''_1+t''_2)^{n''_1+n''_2}}\right.$$

$$\left.+(t''_1)^{n''_1}\sum_{j=0}^{n''_2-1}\frac{(n''_1+j-1)!(t''_2)^{j-1}(n''_2t''_1+n''_2t''_2-t''_2n''_1-t''_2j)}{j!(t''_1+t''_2)^{n''_1+j+1}(n''_1-1)!}\right]$$

With a similar form for the posterior expected loss for d_2, $R_2(\lambda_1,\lambda_2;d_2)$

Similarly for the Quadratic losses, the posterior expected losses are given by :

$$R_1(\lambda_1,\lambda_2;d_1)=k_1\left[\frac{(n''_1+1)n''_1}{(t''_1)^2}-\sum_{j=0}^{n''_2-1}\frac{(t''_1)^{n''_1}(t''_2)^j\,\Gamma(n''_1+j+2)}{j!\,\Gamma(n''_1)(t''_1+t''_2)^{n''_1+j+2}}\right.$$

$$-\frac{2n''_2n''_1}{t''_2t''_1}+\frac{2n''_2}{t''_2}\sum_{j=0}^{n''_2}\frac{(t''_1)^{n''_1}(t''_2)^j\,\Gamma(n''_1+j+1)}{j!\,\Gamma(n''_1)(t''_1+t''_2)^{n''_1+j+1}}$$

$$\left.+\frac{(n''_2+1)n''_2}{(t''_2)^2}-\frac{(n''_2+1)n''_2}{(t''_2)^2}\sum_{j=0}^{n''_2+1}\frac{(t''_1)^{n''_1}(t''_2)^j\,\Gamma(n''_1+j)}{\Gamma(n''_1)\,j!\,(t''_1+t''_2)^{n''_1+j}}\right]$$

With a similar form for the posterior expected loss for d_2, $R_2(\lambda_1,\lambda_2;d_2)$.

3. The Suboptimal Selection Procedure (Subopm1)

Using stirling's formula for approximating factorials, we can obtain the approximations posterior expected loss $R^{*}_1(\lambda_1,\lambda_2;d_1)$ and $R^{*}_2(\lambda_1,\lambda_2;d_2)$ for constant, linear and quadratic losses given in section 2.3.

The procedure subopm1 is given as follows :

make decision d_1 (select Π_2 as better population) if

$R^{*}_1(\lambda_1,\lambda_2;d_1) < R^{*}_2(\lambda_1,\lambda_2;d_2)$

and

make decision d_2 (select Π_1 as better population) if

$R^{*}_1(\lambda_1,\lambda_2;d_1) \geq R^{*}_2(\lambda_1,\lambda_2;d_2)$

Under constant losses, R_1^* and R_2^* are given by :

$$R_1^*(\lambda_1,\lambda_2;d_1)=k_1\left[1-\frac{(t_1'')^{n_1''}}{(t_1''+t_2'')^{n_1''}}-\frac{(t_1'')^{n_1''}}{(2\pi)^{\frac{1}{2}}(n_1''-1)^{n_1''-\frac{1}{2}}}\sum_{j=1}^{n_2''-1}\frac{(t_2'')^{j}(n_1''+j-1)^{n_1''+j-\frac{1}{2}}}{j^{j+\frac{1}{2}}(t_1''+t_2'')^{n_1''+j}}\right]$$, with a similar form for $R_2^*(\lambda_1,\lambda_2;d_2)$

Under linear losses, R_1^* and R_2^* will be :

$$R_1^*(\lambda_1,\lambda_2;d_1)=k_1\left[\frac{n_1''}{t_1''}-\frac{n_2''}{t_2''}+\frac{(t_1'')^{n_1''}(t_2'')^{n_2''-1}(n_1''+n_2''-1)^{n_1''+n_2''-\frac{1}{2}}}{(n_1''-1)^{n_1''-\frac{1}{2}}(2\pi)^{\frac{1}{2}}(n_2''-1)^{n_2''-\frac{1}{2}}e\,(t_1''+t_2'')^{n_1''+n_2''}}\right.$$

$$\left.-\frac{(t_1'')^{n_1''}(n_1''t_1''+n_2''t_2''-t_2''n_1'')}{(t_2'')(t_1''+t_2'')^{n_1''+1}}+\sum_{j=1}^{n_2''-1}\frac{(n_1''+j-1)^{n_1''+j-\frac{1}{2}}(t_2'')^{j-1}(n_1''t_1''+n_2''t_2''-t_2''n_1''-t_2''j)}{j^{j+\frac{1}{2}}(t_1''+t_2'')^{n_1''+j+1}(2\pi)^{\frac{1}{2}}(n_1''-1)^{n_1''-\frac{1}{2}}}\right]$$

Similar form for $R_2^*(\lambda_1,\lambda_2;d_2)$.

Similarly under quadratic losses, R_1^* and R_2^* have the forms :

$$R_1^*(\lambda_1,\lambda_2;d_1)=k_1\left[\frac{(n_1''+1)n_1''}{(t_1'')^2}-\frac{(t_1'')^{n_1''}(n_1''+1)n_1''}{(t_1''+t_2'')^{n_1''+2}}-\sum_{j=1}^{n_2''-1}\frac{(t_1'')^{n_1''}(t_2'')^{j}(n_1''+j+1)^{n_1''+j+\frac{3}{2}}}{(n_1''-1)^{n_1''-\frac{1}{2}}(2\pi)^{\frac{1}{2}}j^{j+\frac{1}{2}}e^{2}(t_1''+t_2'')^{n_1''+j+2}}\right.$$

$$-\frac{2n_2''n_1''}{t_2''t_1''}+\frac{2n_2''(t_1'')^{n_1''}n_1''}{t_2''(t_1''+t_2'')^{n_1''+1}}+\frac{2n_1''}{t_2''}\sum_{j=1}^{n_2''}\frac{(t_1'')^{n_1''}(t_2'')^{j}(2\pi)^{\frac{1}{2}}(n_1''+j)^{n_1''+j+\frac{1}{2}}e^{-n_1''-j}}{j^{j+\frac{1}{2}}e^{-j}(2\pi)^{\frac{1}{2}}(n_1''-1)^{n_1''-\frac{1}{2}}e\,(t_1''+t_2'')^{n_1''+j+1}}$$

$$\left.+\frac{(n_2''+1)n_2''}{(t_2'')^2}-\frac{(t_1'')^{n_1''}(n_2''+1)n_2''}{(t_2'')^2(t_1''+t_2'')n_1''}-\sum_{j=1}^{n_2''+1}\frac{(t_1'')^{n_1''}(t_2'')^{j}(n_2''+1)n_2''(n_1''+j-1)^{n_1''+j-\frac{1}{2}}}{j^{j+\frac{1}{2}}e^{-j}(2\pi)^{\frac{1}{2}}(n_1''-1)^{n_1''-\frac{1}{2}}(t_2'')^2(t_1''+t_2'')^{n_1''+j}}\right]$$

With a similar form for $R_2^*(\lambda_1,\lambda_2;d_2)$.

The optimal posterior expected loss under subopm1 is $R^* = min(R^*_1, R^*_2)$.

. The Suboptimal Procedure (subopm2)

A Bayesian suboptimal scheme is proposed with decision criteria based on the sterior probabilities of λ_1 and λ_2, the posterior Bayes estimator of λ_i, $(i=1,2)$ with ect to the gamma posterior distribution is given by $E(\lambda_i / \underline{x}_i) = \frac{n_i''}{t_i''}$. This is mpted by the need for a quick, easy procedure, to select the better of two Poisson ulations, which allow for the incorporation of information about the parameters sampling information, ignoring the decision-theoretic structure and indifference- formulation. Suppose n observations are taken from each population and are assumed to be independent, this procedure is given as follows : Select $d_1 : \lambda_1 < \lambda_2$ if $\hat{\lambda}_1 < \hat{\lambda}_2$ and Select $d_2 : \lambda_1 \geq \lambda_2$ if $\hat{\lambda}_1 \geq \hat{\lambda}_2$

For the sake of risk comparison, we use the following procedure

Let $R' = R_1(\lambda_1,\lambda_2;d_1)$ if $\hat{\lambda}_1 < \hat{\lambda}_2$, $R' = R_2(\lambda_1,\lambda_2;d_2)$ if $\hat{\lambda}_1 \geq \hat{\lambda}_2$

So R' will be the optimal risk using subopm2

5. Comparisons and Discussion Under Posterior Expected Losses

This section contains some numerical results about the efficiency of these schemes relative to opm for the linear loss function, various N and various priors.

From table I we note that, the posterior expected loss for the optimal procedure is less than or equal to the posterior expected loss for the suboptimal 1 and suboptimal 2 procedures. Also it is clear from the table that as N increases, the Bayes risk decreases in all schemes.

2.1 The Algorithms for the opm and subopm1 Procedures

1. Specify prior parameters n_i', t_i' , $i = 1,2$, sample size n, parameters for populations $\lambda_i > 0$ and constant losses k_i.
2. Generate a random sample of size n from populations Π_i , $(x_{i1}, x_{i2}, \ldots, x_{in})$ and find $s = \sum_{j=1}^{n} x_{ij}$ as follows :
   i. For i=1 to n
   ii. Generate a random number $U_i \in [0,1]$
   iii. If $\prod_i U_i \geq e^{-\lambda_i}$ then j=j+1 and goto ii else s=s+j
3. Calculate the posterior parameter for populations Π_i, $i=1,2$

   $n_i'' = n_i' + s$, $t_i'' = t_i' + n$
4. Find the posterior expected loss for decision d_1 and d_2 ($R_1(\lambda_1,\lambda_2;d_1)$, $R_2(\lambda_1,\lambda_2;d_2)$ for opm procedure and $R^*_1(\lambda_1,\lambda_2;d_1)$, $R^*_2(\lambda_1,\lambda_2;d_2)$ for subopm1 procedure).

Table I

Comparisons of the schemes using Bayes risk under linear loss function for various prior and various N

$k_1 = k_2 = 5$, $\underline{\lambda} = (9,6)$				
Prior prob.	N	opm	Subopm1	Subopm2
(3,5),(3,5)	10	2.444837E-02	7.027235E-02	2.444837E-02
	14	3.549442E-03	4.755408E-02	3.549442E-03
	20	1.281756E-03	3.324223E-02	1.281756E-03
(6,5),(6,5)	10	3.163718E-02	7.330076E-02	3.163718E-02
	14	4.736941E-03	4.559913E-02	4.736941E-03
	20	1.624834E-03	3.207349E-02	1.624834E-03

(7,5),(7,5)	10	3.423650E-02	7.468765E-02	3.423650E-02
	14	5.182229E-03	4.510168E-02	5.182229E-03
	20	1.751910E-03	3.173133E-02	1.751910E-03
(9,5),(9,5)	10	3.972983E-02	7.797140E-02	3.972983E-02
	14	6.149522E-03	4.431819E-02	6.149522E-03
	20	2.026059E-03	3.111388E-02	2.026059E-03
(10,5),(10,5)	10	4.262032E-02	7.985240E-02	4.262032E-02
	14	6.672257E-03	4.402663E-02	6.672257E-03
	20	2.173426E-03	3.083736E-02	2.173426E-03

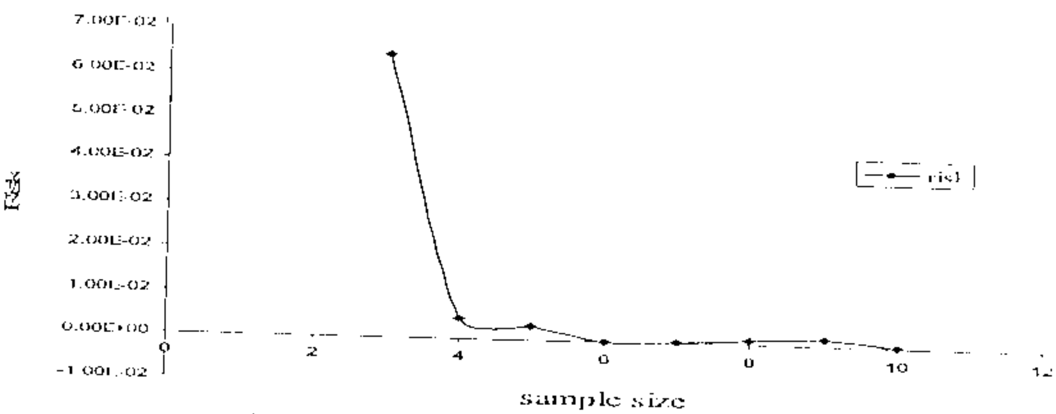

Figure (1) : The influence of the sample size on the posterior expected loss in opm procedure

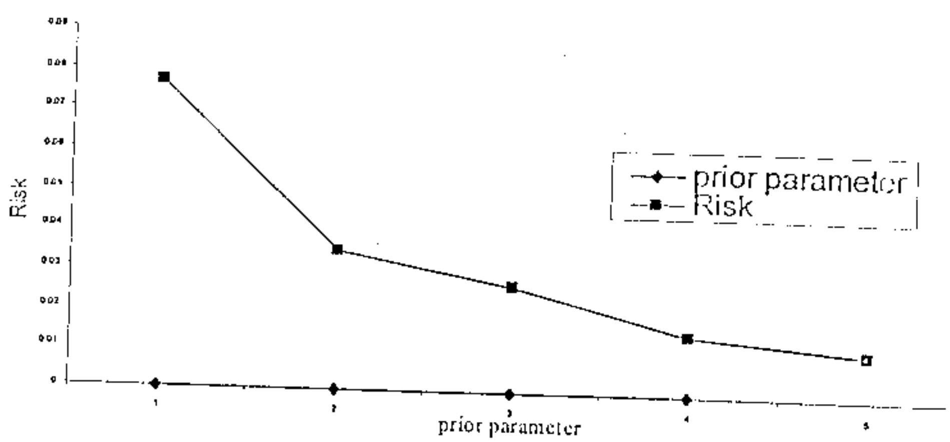

Figure (2) : The influence of the prior information on the posterior expected loss in opm procedure

6. The MC Studies

For some applications it is of interest to compare the procedures mentioned in the previous sections under criteria other than risk such as the probability of correct selection of the better population. Using this performance measure, some Monte Carlo (MC) simulations were carried out to evaluate the effectiveness of our procedures.

Subsection 6.1 contains description of the MC studies. In subsection 6.2 we present some MC results of estimates and discussion.

6.1 Description of the MC Studies

In this subsection we briefly describe the method of MC simulation as it is applied to our procedures. MC studies have been carried out to investigate the probability of correctly selecting the better of two Poisson populations.

The simulation program performs large number of runs *(t=5000)*, which are assumed to be independent, in order to obtain MC estimates with high precision. At each run mutually independent Poisson observations are generated by using the assumed probability model with λ_1 and λ_2 specified in advance and then the selection procedure is applied. The observed values of probability of correct selection measure are accumulated. At the end of all runs, these accumulated values are divided by t to obtain the MC estimates of probability of correct selection. On run of 5000 trials the same λ_1 and λ_2 are used.

The values of X can be considered as the observed values of random variable, possessing the Poisson distribution that should be simulated according to the sampling scheme, the following quantities are required for input k_1,k_2, priors, loss function, N. As measure of performance of the proposed procedures we shall use the probability of correct selection $P(CS)$. In a MC experimentation the population that has greater mean rate is known to us, so we can check if the procedure gives a correct selection. After t repetitions we estimate $P(CS)$ by the fraction of correct selections in the t repetitions. It can computed as follows :

$P(d_i/d_i)$ is the proportions of number of times when the procedure takes decision d_i given decision d_i is true in t repetitions.

$$P(CS)=\sum_{i=1}^{2}P(d_i/d_i) \quad , where\ d_1:\lambda_1<\lambda_2 \ \ ,\ d_2:\lambda_1\geq\lambda_2$$

6.2 The Algorithms to Find the Probability of Correct Selection

1. specify prior parameter $n'_i,t'_i\ ,i=1,2$, sample size n from each population, and the parameter for populations $\lambda_i>0$.
2. take t=1,…,5000 replication
2. take an i.i.d. sample of size n from populations Π_i, $(x_{i1},x_{i2},\dots,x_{in})$ and find $s=\sum_{j=1}^{n}x_{ij}$.
4. calculate the posterior Bayes estimators $\hat{\lambda}_i$ where $\hat{\lambda}_i=\dfrac{n'_i+\sum_{j=1}^{n}x_{ij}}{t'_i+n}$, $i=1,2$
5. compare the posterior Bayes estimators $\hat{\lambda}_i$, $i=1,2$ and the parameters of populations λ_i to find the number of correct selection and the number of not correct selection for the subopm2 procedure. for the opm and subopm1 procedures

compare $\hat{\lambda}_i$, $i=1,2$ and $R_1(\lambda_1,\lambda_2;d_1)$, $R_2(\lambda_1,\lambda_2;d_2)$ for opm procedure and $R^*_1(\lambda_1,\lambda_2;d_1)$, $R^*_2(\lambda_1,\lambda_2;d_2)$ for subopm1 procedure.

5. after 5000 replication compute the probability of correct selection P(CS) and the probability of not correct selection P(NCS).

P(CS)=total number of correct selection /number of replications

P(NCS)= total number of not correct selection / number of replications

5.3 The MC Results of P(CS): Comparison and Discussion

In this subsection we compare and discuss the MC estimates of P(CS) of the schemes opm,subopm1 and subopm2. The numerical results are presented in table II. From this table we note that the P(CS) increases as N increase for all procedures for all sets of simulations. The table also shows that the P(CS) for the opm procedure is greater than or equal the P(CS) under the procedures subopm1 and subopm2 for all set of simulations. The P(CS) for the subopm1 is greater than P(CS) for the subopm2 for all sets of simulations.

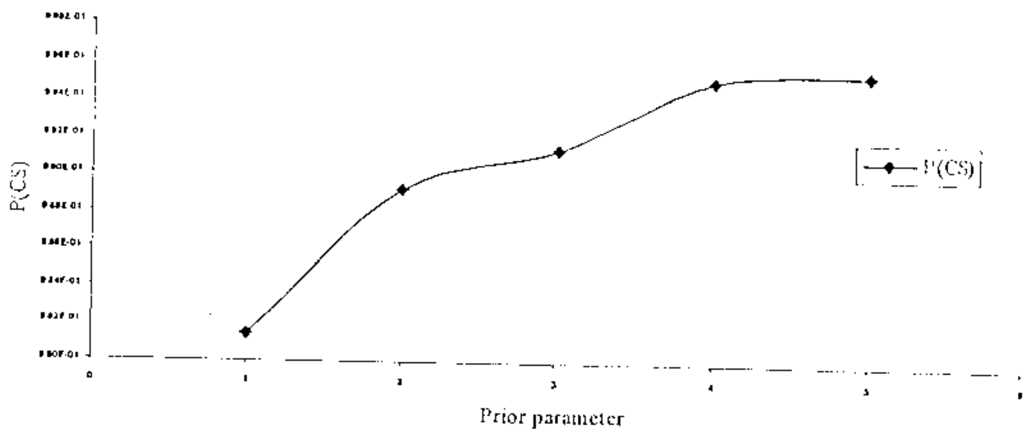

Figure (3) : The influence of the prior information on the probability of correct selection in opm procedure

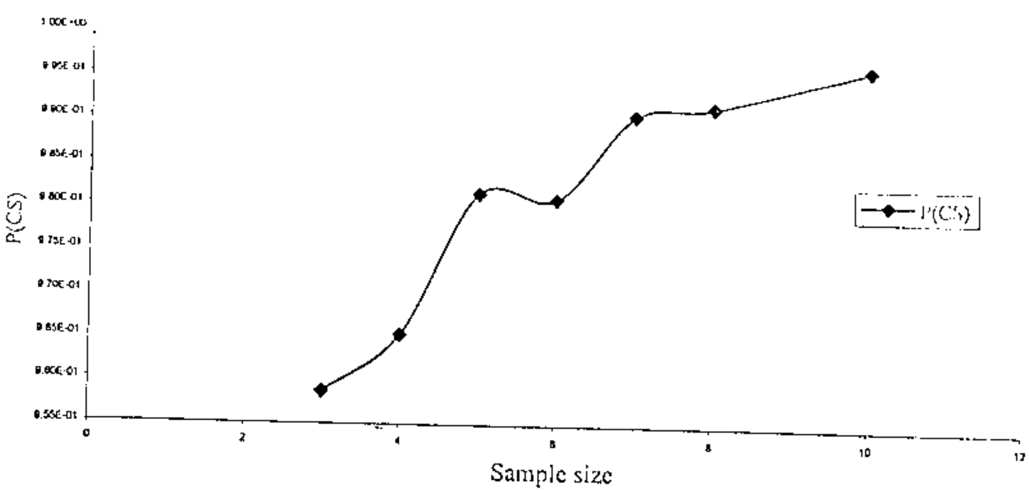

Figure (4) : The influence of the sample size on the probability of correct selection in opm procedure

Table II
Comparisons of the schemes using P(cs) under linear loss function for various prior probabilities, various N and various λ_1 and λ_2.

$k_1 = k_2 = 5$				
Prior prob.	N	opm	Subopm1	Subopm2
(3,5),(2,4) $\underline{\lambda}=(6,9)$	10	1.000000	1.000000	9.794000E-01
	14	1.000000	1.000000	9.910000E-01
	20	1.000000	1.000000	9.970000E-01
(6,3),(4,2) $\underline{\lambda}=(5,8)$	10	1.000000	1.000000	9.868000E-01
	14	1.000000	1.000000	9.926000E-01
	20	1.000000	1.000000	9.988000E-01
(7,9),(3,3) $\underline{\lambda}=(9,3)$	10	1.000000	1.000000	9.830000E-01
	14	1.000000	1.000000	9.984000E-01
	20	1.000000	1.000000	9.998000E-01
(6,4),(4,5) $\underline{\lambda}=(3,5)$	10	9.980000E-01	9.946000E-01	8.00200E-01
	14	9.980000E-01	9.980000E-01	8.914000E-01
	20	1.000000	9.992000E-01	9.560000E-01

7. Conclusions and Future Research

In this paper we proposed Bayesian fixed sample size procedures for selecting the better of two Poisson populations using Bayesian decision-theoretic formulation under different loss functions. A method of obtaining approximation (subopm1) to the optimal procedure (opm) is suggested which is simpler in computation with little increasing in risk. The suboptimal scheme (subopm1) that based on posterior estimate is suitable when the number of alternatives are large as the computations procedures are simple and fast.

Future research directions include

1. Only Bayesian fixed sample procedures have been considered, an extension to Bayesian sequential procedures would be useful.
2. The case of two populations has been taken up, an extension to more than two populations would be straight forward, although not without computational difficulties.
3. Applications of our approach can be explored in other task.
4. The exact formula for the probability of a correct selection (P(CS)) can be explored.
5. The work can be extend using exponential family and general loss function.

References

(1) Alam, K. (1971). Selection from Poisson processes. Ann. Inst. Statist. Math. 23, 411-418.

(2) Alam, K. and Thompson, J.R. (1973). A problem of ranking and estimation with Poisson processes Technometrics 15, 801-808.

(3) Chen, CH. (1995). An effective approach to smartly allocate computing budget for discrete-event simulation. Proceedings of the 34th IEEE conference on decision and control, 2598-2605.

(4) Chen, C.H and et al. (1996). A gradient approach for smartly allocating computing budget for discrete event simulation. Proceedings of the 1996 Winter simulation conference, 398-405.

(5) Chick, S.E. (1997). Selecting the best system : A decision-theoretic approach. Proceedings of the 1997 Winter simulation conference, 326-333.

(6) Dudewicz, D.J. and Koo, J-k (1982). The complete categorized guide to statistical selection and ranking procedures. Americal Sciences press Inc, Columbus, Ohio, USA.

(7) Gibbons, J.D.. Olkin, I. and Sobel, M. (1977). Selecting and ordering populations : A new statistical methodology. John Wiley and Sons, New York.

(8) Gupta, S.S. and Panchapakesan, S. (1979). Multiple decision procedures : Theory and methodology of selecting and ranking populations. John Wiley and Sons, New York.

(9) Horrace, W.C. (2002). Selection procedures for order statistics in Empirical Economic studies, 1-26.

(10) Huang, Wen Tao (1973). Some selection procedures of Poisson populations conditioned on the sum of observations. Tamkang J. Math. 4, No. 2, 195-202.

(11) Liese, F., (2001), selection of the best population based on rare data, Purdue University.

(12) Madhi, S.A. (1986), Bayesian sequential procedures for Binomial and Multinomial selection problems. PHD thesis Manchester University.

(13) Nelson, B.L. and Hong, L.J. (2003), an indifference zone selection procedure with minimum switching and sequential sampling, USA 474-480.

(14) Nelson, L. and Pichitlamken J., (2001) selection of the best procedures for optimization-via simulation. USA, 401-407.

(15) Prasanta S.B., (2005), Bayesain and computational statistics, Montana State University, USA.

(16) Raiffa, H. and Schlaifer, R. (1968). Applied statistical decision theory. Massachusetts Institute of Technology Press. USA.

(17) Ross, M.S. (1987). Introduction to probability and statistics for engineers and scientists. John Wiely & Sons, Inc. New York.

اجراءات بيزينية لاختيار افضل مجتمع من بين مجتمعين بوسونين

الخلاصة

في هذا البحث تم اقتراح ودراسة طريقة مثلى باستخدام حجم عينة ثابت لاختيار أفضل مجتمع من بين مجتمعين بوسونيين . لقد استخدمنا منهج نظرية القرار البيزينية مع دوال خسارة مختلفة مع معلومات سابقة للتجربة بصورة التوزيع الاحتمالي كاما لبناء هذا الأجراء. واحتوى البحث أيضاً على طرق مثلى جزئياً أحدها يعتمد على تقدير بعدي للمعالم والأخرى طريقة تقريبية للطريقة المثلى باستخدام قاعدة سترلنك للتقريب. وتضمن البحث أيضاً مقارنات بين هذه الإجراءات باستخدام خصائص أنجاز مثل الخطورة البيزينية واحتمال الاختيار الصحيح للمجتمع الأفضل.