SIMULATION RESULTS OF A SEQUENTIAL FIXED WIDTH CONFIDENCE INTERVAL FOR A FUNCTION OF PARAMETERS

309

by

CHANG SOO PAIK

B.S., Oakland University, 1976

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1979

Approved by:

S. K. Perng

Major Professor.

TABLE OF CONTENTS

PART I

INTRODUCTION

1.1. Statement of the problem and the goal of this report

Let $X_1, X_2, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_n$ form two independent random samples for which the probability density functions or the probability functions are $f(x; \underline{\theta})$ and $h(y; \underline{\phi})$, respectively, where $\underline{\theta} = (\theta_1, \theta_2, \ldots, \theta_{d_1})$ and $\underline{\phi} = (\phi_1, \phi_2, \ldots, \phi_{d_2})$ are unknown parameter vectors.

Let $g(\underline{\gamma}) = g(\underline{\theta}, \underline{\phi})$ be a real valued function of the unknown parameters. Perng and Hasza ((3)) developed a sequential procedure to find a fixed width confidence interval which is of the form

$$I = (\hat{g}(\underline{\gamma}) - d,\ \hat{g}(\underline{\gamma}) + d) \quad \text{such that}$$

$$(1.1) \quad \lim_{d \to 0} P_r(g(\underline{\gamma}) \in I) = 1 - \alpha \quad \text{where } \alpha \text{ is preassigned.}$$

In order to understand the problem consider the following theorem and sequential procedure.

Theorem 1.1. (Perng and Hasza (3))

Let $\underline{\theta}_m$ be the maximum likelihood estimator (M.L.E.) of $\underline{\theta}$ based on $X_1, X_2, \ldots, X_m$ and $\underline{\phi}_n$ be the M.L.E. of $\underline{\phi}$ based on $Y_1, Y_2, \ldots, Y_n$. Assume that the regularity conditions (Zacks (4)) for the M.L.E. are satisfied and that

$$\underline{\theta}_m \longrightarrow \underline{\theta} \text{ with probability 1 as } m \longrightarrow \infty \text{ and}$$

$\underline{\hat{\phi}}_n \longrightarrow \underline{\phi}$ with probability 1 as $n \rightarrow \infty$ in such a manner that

$\frac{m}{m+n} \longrightarrow r$ where $0 < r < 1$.

Furthermore, assume that $g(\underline{\gamma}) = g(\underline{\theta}, \underline{\phi})$ has continuous derivatives of order zero, one and two at $\underline{\gamma} = (\underline{\theta}, \underline{\phi})$.

Then,

$$\sqrt{m+n}\,(g(\hat{\underline{\gamma}}_k) - g(\underline{\gamma}))\ \underrightarrow{\text{Asymp}}\ N(0, A) \text{ as } m, n \longrightarrow \infty$$

where $\hat{\underline{\gamma}}_k = (\hat{\underline{\theta}}_m, \hat{\underline{\phi}}_n)$ and $k = m + n$,

$$A = r^{-1} V_1(\underline{\theta}) + (1-r)^{-1} V_2(\underline{\phi}),$$

$$V_1(\underline{\theta}) = \underline{a}'\, I^{-1}(\underline{\theta})\, \underline{a},$$

$$V_2(\underline{\phi}) = \underline{b}'\, I^{-1}(\underline{\phi})\, \underline{b},$$

$$\underline{a} = (a_i) = \frac{\partial g(\underline{\gamma})}{\partial \theta_i} \quad \text{for } i = 1,2,\ldots,d_1,$$

$$\underline{b} = (b_i) = \frac{\partial g(\underline{\gamma})}{\partial \phi_i} \quad \text{for } i = 1,2,\ldots,d_2,$$

$I(\underline{\theta})$ is the Fisher Information Matrix of X and defined by

$$I(\underline{\theta}) = (I(\underline{\theta})_{ij}) = -E\left(\frac{\partial^2 \ln f(x; \underline{\theta})}{\partial \theta_i\, \partial \theta_j}\right) \quad \text{for } i = 1,2,\ldots,d_1;\ j = 1,2,\ldots,d_1.$$

Similarly,

$$I(\underline{\phi}) = -E\left(\frac{\partial^2 \ln h(y; \underline{\phi})}{\partial \phi_i\, \partial \phi_j}\right) \quad \text{for } i = 1,2,\ldots,d_2;\ j = 1,2,\ldots,d_2.$$

Given below are the five steps of the sequential procedure.

Step 1. Determine α and d.

Step 2. Take m_o observations from X and n_o observations from Y.

Step 3. (Sampling Rule) Suppose that at any stage there are p observations from X and q observations from Y. Then we take the next observation

(1.2) from X if $\frac{p}{p+q} \le \hat{r}$ and

from Y if $\frac{p}{p+q} > \hat{r}$,

where

(1.3) $$\hat{r} = \frac{V_1(\underline{\hat{\theta}}_p)^{\frac{1}{2}}}{V_1(\underline{\hat{\theta}}_p)^{\frac{1}{2}} + V_2(\underline{\hat{\phi}}_q)^{\frac{1}{2}}} ,$$

$$V_1(\underline{\hat{\theta}}_p) = \underline{\hat{a}}' \; I^{-1} \; (\underline{\hat{\theta}}_p) \; \underline{\hat{a}} ,$$

$$V_2(\underline{\hat{\phi}}_q) = \underline{\hat{b}}' \; I^{-1} \; (\underline{\hat{\phi}}_q) \; \underline{\hat{b}}' .$$

Step 4. (Stopping Rule) Stop sampling if either of the following conditions is satisfied:

(1.4) i) $k \ge \frac{C_k^2 \hat{A}_k}{d^2}$

or

$$(1.5) \qquad \text{ii)} \quad p \geq \frac{C_k^2 \, V_1(\hat{\underline{\theta}}_p)}{d^2} \quad \text{and} \quad q \geq \frac{C_k^2 \, V_2(\hat{\underline{\emptyset}}_q)}{d^2} \; ,$$

where $k = p+q$,

C_k is a sequence of positive real numbers such that $C_k \longrightarrow c$ and $\Phi\,(c) = 1 - \frac{\alpha}{2}$, where Φ is the cumulative density function of $N(0, 1)$,

$$(1.6) \qquad \hat{A}_k = \hat{r}^{-1} V_1(\hat{\underline{\theta}}_p) + (1 - \hat{r})^{-1} V_2(\hat{\underline{\emptyset}}_q) \; .$$

Notice that $k \longrightarrow \infty$ as $d \longrightarrow o$ under the assumption that $A < \infty$.

Step 5. Construct a confidence interval I_k as follows when sampling is stopped:

$$(1.7) \qquad I_k = (g(\hat{\underline{\gamma}}_k) - d, \; g(\hat{\underline{\gamma}}_k) + d) \; .$$

Perng and Hasza ((3)) have shown that the confidence interval I_k has the following asymptotic property as in (1.1):

$$\lim_{d \to o} \Pr(g(\underline{\gamma}) \in I_k) = 1 - \alpha \; .$$

However, it is not known how well the sequential procedure will behave with various values of d, in which experimenters or researchers may be interested. Also it is of interest to see whether the sample sizes are reasonable as a result of the sequential procedure. A Monte Carlo study was used to evaluate the performance of the proposed procedure and this paper reports the results of the Monte Carlo study.

PART II

SIMULATION RESULTS

The four cases of the real valued function g $(\underline{\gamma})$ to be studied are shown in the Table 2.1.

Table 2.1

Case	density of X	density of Y	$g(\underline{\gamma})$
1.	$N(\mu_1, \sigma_1^2)$	$N(\mu_2, \sigma_2^2)$	σ_1^2/σ_2^2
2.	Poisson (λ_1)	Poisson (λ_2)	$\lambda_1 - \lambda_2$
3.	Exponential (λ_1)	Exponential (λ_2)	λ_1/λ_2
4.	Exponential (λ_1)	Exponential (λ_2)	$\lambda_1 - \lambda_2$

In each case, 1,000 simulations were carried out for each set of parameters to determine the coverage probability, the average and the standard deviation of the total sample sizes, and the fixed sample sizes to achieve

$$(2.1) \qquad P_r(|\, g(\hat{\underline{\gamma}}_k) - g(\underline{\gamma})\, | \leq d) = 1 - \alpha .$$

To generate random samples, 'Super Duper' random number generator ((2)) was used for cases 1, 3 and 4, and Kemp's method ((1)) for case 2.

In the next four sections, each of the cases is studied and discussed. Tables and plots of the results are given for each case and an effective way of choosing parameter values is discussed.

Appendix A includes the computer programs used to generate random samples, get the coverage probability, and the average total sample sizes.

The computer programs used to compute the fixed sample size to achieve (2.1) for cases 2 and 4 are shown in Appendix B.

2.1. Normal distribution case

Let X_1, X_2, ..., and Y_1, Y_2, ... be two independent random samples from $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$, respectively.
Let $g(\underline{\gamma}) = g(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) = \sigma_1^2/\sigma_2^2$.

The steps of the sequential procedure are presented below.

Step 1. Let $\alpha = 0.05$.

Step 2. Take 7 observations from X and from Y.

(With various trials, it is felt that seven is appropriate for the initial observations).

Step 3. Compute the following:

$$\underline{a}' = \left(\frac{\partial g(\sigma_1^2/\sigma_2^2)}{\partial \mu_1} \quad \frac{\partial g(\sigma_1^2/\sigma_2^2)}{\partial \sigma_1^2}\right) = (0 \quad 1/\sigma_2^2) \ ,$$

$$\underline{b}' = \left(\frac{\partial g(\sigma_1^2/\sigma_2^2)}{\partial \mu_2} \quad \frac{\partial g(\sigma_1^2/\sigma_2^2)}{\partial \sigma_2^2}\right) = (0 \quad -\sigma_1^2/\sigma_2^4) \ ,$$

$$I(\mu_1, \sigma_1^2) = \begin{matrix} 1/\sigma_1^2 & 0 \\ 0 & 1/2\sigma_2^4 \end{matrix} \ ,$$

$$I(\mu_2, \sigma_2^2) = \begin{matrix} 1/\sigma_2^2 & 0 \\ 0 & 1/2\sigma_2^4 \end{matrix} \ .$$

Hence,

$$V_1(\underline{\theta}) = \underline{a}'I^{-1}(\mu_1, \sigma_1^2)\ \underline{a} = 2\sigma_1^4/\sigma_2^4 \quad \text{and}$$

$$V_2(\underline{\emptyset}) = \underline{b}'I^{-1}(\mu_2, \sigma_2^2)\ \underline{b} = 2\sigma_1^4/\sigma_2^4$$

Since $V_1(\underline{\theta}) = V_2(\underline{\emptyset})$, we get

(2.1.1) $\hat{r} = r = \frac{1}{2}$.

Thus, if there are p observations on X and q on Y, then we take the next observation from X if

(2.1.2) $\frac{p}{p+q} \leq \frac{1}{2}$ and from Y if

$$\frac{p}{p+q} > \frac{1}{2} .$$

Step 4. Compute A:

$$A = r^{-1}V_1(\underline{\theta}) + (1 - r)^{-1}V_2(\underline{\emptyset})$$

$$= (2)(2\sigma_1^4/\sigma_2^4) + (2)(2\sigma_1^4/\sigma_2^4)$$

$$= 8\sigma_1^4/\sigma_2^4 .$$

Hence,

$$\hat{A}_k = 8(S_1^2/S_2^2)^2$$

where $k = p+q$,

$$S_1^2 = \sum_{i=1}^{p} (X_i - \bar{X}_p)^2/p \ , \qquad \bar{X}_p = \sum_{i=1}^{p} X_i/p \ ,$$

$$S_2^2 = \sum_{j=1}^{q} (Y_j - \bar{Y}_q)^2/q \ , \qquad \bar{Y}_q = \sum_{j=1}^{q} Y_j/q \ .$$

Second, define $C_k^2 = (1.96)^2 \dfrac{p+q+13}{p+q-13}$.

It can now be seen that the minimum values of p and q that make C_k positive are $p = q = 7$, which were used in Step 2.

Thus, stop sampling if

$$(2.1.3) \qquad p+q \geq \frac{(1.96)^2 \dfrac{p+q+13}{p+q-13} \ 8(\dfrac{S_1^2}{S_2^2})}{d^2} \ .$$

Notice that the difference between p and q is either zero or one when the above is satisfied. This is clear since $r = ½$.

Step 5. Construct a confidence interval as follows when sampling is stopped:

$$(2.1.4) \qquad I_k = (\frac{S_1^2}{S_2^2} - d \ , \ \frac{S_1^2}{S_2^2} + d) \qquad \text{where } k = p+q \ .$$

For the purpose of simulation we now define the values of all the parameters μ_1, σ_1^2, μ_2, σ_2^2 and the confidence width d.

Because there is an infinite number of combinations of these values, it would be desirable to find an effective way of choosing them.

To do so we shall investigate how the random samples from $N(\mu_1, \sigma_1^2)$ are to be generated and how that relates to Step 4 and Step 5.

Let $e_1, e_2, \ldots,$ which will be generated, be a sequence of random samples from N(0, 1). Then X_i, the i th sample from $N(\mu_1, \sigma_1^2)$, would be generated by the following formula.

$$X_i = \sigma_1 e_i + \mu_1 .$$

Let $Se_1^2 = \sum_{i=1}^{p} (e_i - \bar{e}_1)^2/p$ and $Se_2^2 = \sum_{j=1}^{q} (e_j' - \bar{e}_2)^2/q$,

where $\bar{e}_1 = \sum_{i=1}^{p} e_i/p$, $e_2 = \sum_{j=1}^{q} e_j'/q$, and $\{e_i\}$ and $\{e_j'\}$ are two independent sequences of random samples from N(0, 1).

Now consider the stopping rule (Step 4),

$$p + q \geq \frac{8(1.96)^2 \frac{p+q+13}{p+q-13} \left(\frac{S_1^2}{S_2^2}\right)^2}{d^2} \text{ , in particular,}$$

$$\left(\frac{S_1^2}{S_2^2}\right)^2/d^2 = \left| \frac{\sum_{i=1}^{p} (X_i - \bar{X}_p)^2/p}{\sum_{j=1}^{q} (Y_j - \bar{Y}_q)^2/q} \right|^2 /d^2$$

$$= \left(\frac{\sigma_1^2 \; Se_1^2}{d\sigma_2^2 \; Se_2^2}\right)^2$$

Let $\lambda = \dfrac{\sigma_1^2}{d\sigma_2^2}$, then the above becomes

$$\left(\frac{\lambda \; Se_1^2}{Se_2^2}\right)^2 .$$

Next, in Step 5, the following inequality would be considered to determine, in each simulation, whether the confidence interval covers the ratio σ_1^2/σ_2^2 .

$$\frac{S_1^2}{S_2^2} - d \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{S_1^2}{S_2^2} + d \quad \text{which can be rewritten as}$$

$$\frac{\sigma_1^2 \; Se_1^2}{\sigma_2^2 \; Se_2^2} - d \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{\sigma_1^2 \; Se_1^2}{\sigma_2^2 \; Se_2^2} + d .$$

If we divide the above by d, it reduces to

$$(2.1.5) \quad \frac{\lambda Se_1^2}{Se_2^2} - 1 \leq \lambda \leq \frac{\lambda Se_1^2}{Se_2^2} + 1.$$

As has been seen the sampling rule (Step 3), the stopping rule (Step 4), and the coverage probability of the confidence interval (Step 5) do not depend on μ_1 and μ_2, and hence without loss of generality we may let $\mu_1 = \mu_2 = 0$. Furthermore, Steps 3, 4 and 5 would not be affected by specific values of σ_1^2, σ_2^2 and d as long as λ is the same.

In other words, if we use the same starting seed numbers to generate a sequence of random samples (via Super Duper) for any combination of σ_1^2, σ_2^2 and d with λ being equal, we would stop with the same values of p and q, respectively, and get the same inequality as in (2.1.5). Therefore, we need to consider only the values of λ , which reduces the difficulty of choosing a combination of μ_1, μ_2, σ_1^2, σ_2^2 and d. Now consider the fixed sample sizes needed to achieve

$$(2.1.6) \quad \Pr(|\, S_1^2/S_2^2 - \sigma_1^2/\sigma_2^2 | \leq d) = 0.95 .$$

The reason for this is to see if the sample sizes due to the sequential procedure are reasonable.

Suppose all the parameters are known. Then for any given d, p and q can be computed such that (2.1.6) is satisfied. Then compare these values to those in (2.1.3) for the goodness of the sequential procedure. Consider the following:

$$\Pr\left(\left| \frac{S_1^2}{S_2^2} - \frac{\sigma_1^2}{\sigma_2^2} \right| \le d\right)$$

$$= \Pr\left(\frac{p(q-1)\sigma_2^2}{q(p-1)\sigma_1^2} \left| \frac{S_1^2}{S_2^2} - \frac{\sigma_1^2}{\sigma_2^2} \right| \le \frac{p(q-1)\sigma_2^2\, d}{q(p-1)\sigma_1^2} \right)$$

$$= \Pr\left(\left| \frac{p\ Se_1^2/(p-1)}{q\ Se_2^2/(q-1)} - \frac{p(q-1)}{q(p-1)} \right| \le \frac{p(q-1)}{q(p-1)} \cdot \frac{1}{\lambda} \right)$$

$$(2.1.6) \qquad = \Pr\left(\left| F_{(p-1,\ q-1)} - \frac{p(q-1)}{q(p-1)} \right| \le \frac{p(q-1)}{q(p-1)} \cdot \frac{1}{\lambda}\right)$$

Where $F_{(p-1,\ q-1)}$ is a central F random variable with p-1 and q-1 degrees of freedom.

Notice the fact that the difference between p and q is either zero or one when sampling was stopped. Hence, p and q may be chosen such that (2.1.6) is near 0.95.

The Table 2.1.1 shows the coverage probability, the average and the standard deviation of the total sample sizes for the sequential procedure, based on 1,000 simulations, and the fixed sample size for each λ.

Table 2.1.1

λ	Coverage Probability	Average Total Sample Size	Standard Deviation	Fixed Sample Size*
.5	1.000	26.5	11.0	22
1.0	.998	50.4	25.7	48
1.5	.939	85.9	42.0	88
2.0	.911	137.4	60.3	140
2.5	.897	206.8	80.6	208
3.5	.932	391.7	110.1	390
4.5	.931	639.4	148.9	636
5.0	.936	777.7	163.8	782
5.5	.934	946.5	173.7	944

*To get the fixed sample size using (2.1.6), a Monroe Calculator (Model 1730) was used for $\lambda = 0.5$ through $\lambda = 3.5$. The function PROBF in SAS was used for $\lambda = 4.5$ through $\lambda = 5.5$.

Table 2.1.1 shows the coverage probability first decreases but then tends to increase a λ increases. The average total sample size matches very well with the fixed sample size regardless the value of λ.

The plots in Figure 2.1.1 through 2.1.4 show the frequency distributions of the sample sizes with different values of λ. The shaded area indicates that portion in which the confidence interval did not cover the ratio. This phenomenon occured when the sample size was either very small or very large.

This situation could be explained by studying the stopping rule in (2.1.2). If S_1^2/S_2^2 is small at an early stage, the stopping rule would be satisfied and the upper confidence limit may not be large enough to cover the true ratio σ_1^2/σ_2^2. On the other hand, if S_1^2/S_2^2 is not small, then the stopping rule would let us keep sampling until the stopping rule is met. However, if S_1^2/S_2^2 is still relatively large at a later stage, then the lower confidence limit may not be low enough to cover the ratio.

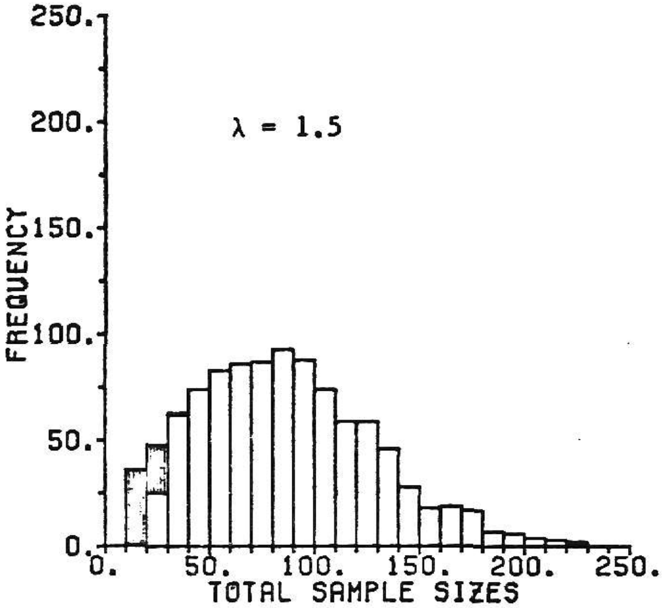

Figure 2.1.1

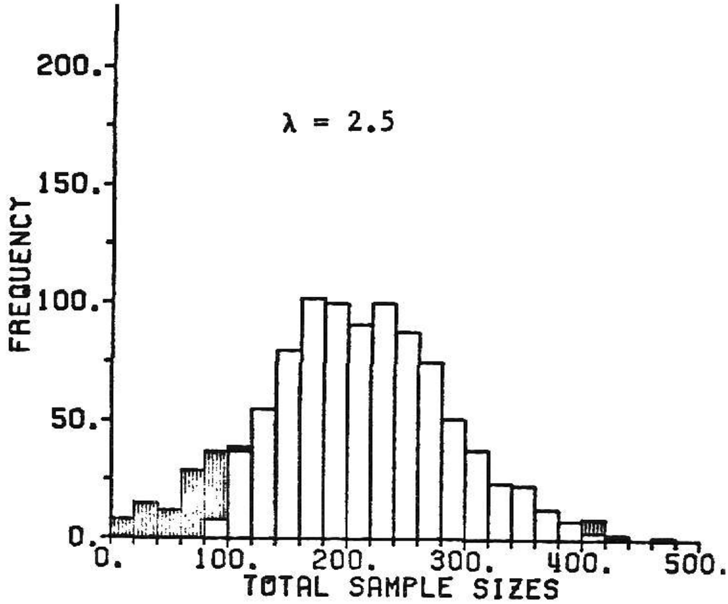

Figure 2.1.2

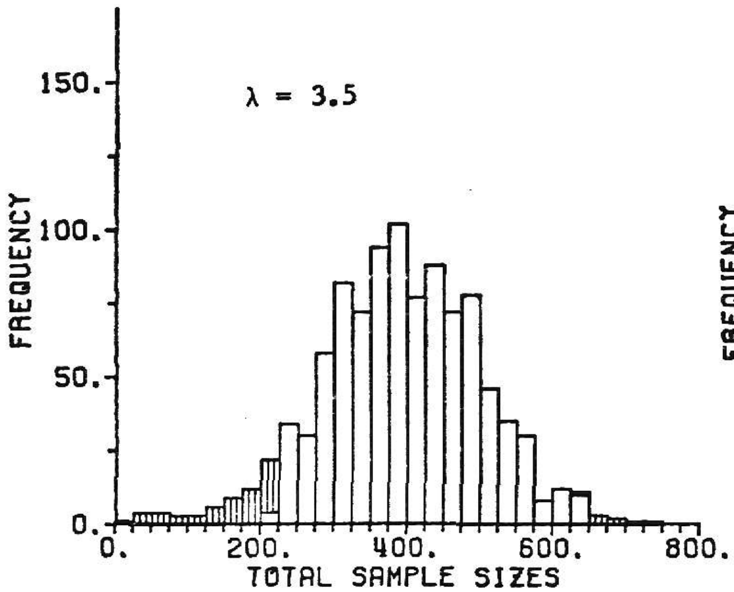

Figure 2.1.3

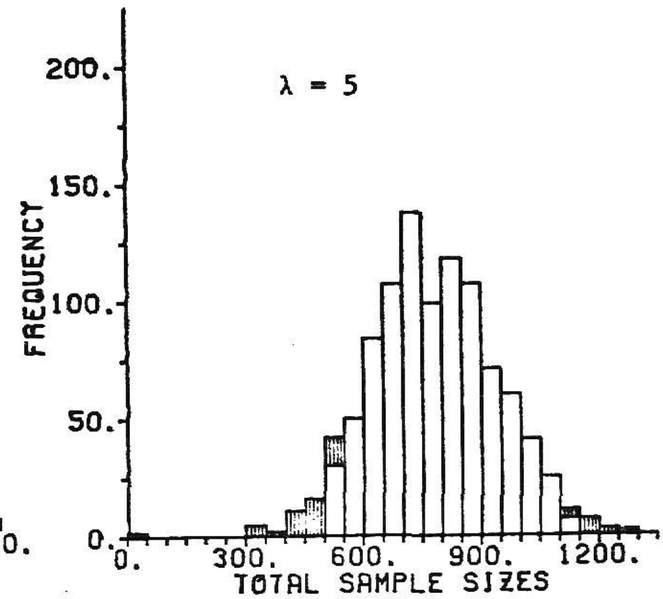

Figure 2.1.4

2.2. Poisson distribution case

Let $X_1, X_2, \ldots,$ and $Y_1, Y_2, \ldots$ be two independent random samples from Poisson (λ_1) and Poisson (λ_2), respectively, and $g(\underline{\gamma}) = g(\lambda_1, \lambda_2) = \lambda_1 - \lambda_2$.

Again, each step of the sequential procedure is presented.

Step 1. Let $\alpha = 0.05$

Step 2. Take 7 observations from X and from Y.

Step 3. Compute the following:

$$a = \left(\frac{\partial(\lambda_1 - \lambda_2)}{\partial\lambda_1} \right) = 1 ,$$

$$b = \left(\frac{\partial(\lambda_1 - \lambda_2)}{\partial\lambda_2} \right) = -1 ,$$

$$I(\lambda_1) = \lambda_1^{-1} , \; I(\lambda_2) = \lambda_2^{-1} .$$

Hence,

$$V_1(\lambda_1) = a\, I^{-1}(\lambda_1) a = \lambda_1 \quad \text{and}$$

$$V_2(\lambda_2) = b\, I^{-1}(\lambda_2) b = \lambda_2 .$$

Thus, $r = \dfrac{V_1(\lambda_1)^{\frac{1}{2}}}{V_1(\lambda_1)^{\frac{1}{2}} + V_2(\lambda_2)^{\frac{1}{2}}} = \dfrac{\lambda_1^{\frac{1}{2}}}{\lambda_1^{\frac{1}{2}} + \lambda_2^{\frac{1}{2}}}$

Because the M.L.E. of λ_1 is $\bar{X}_p$ and the M.L.E. of λ_2 is $\bar{Y}_q$,

$$(2.2.1)\quad r = \frac{\bar{X}_p^{\frac{1}{2}}}{\bar{X}_p^{\frac{1}{2}} + \bar{Y}_q^{\frac{1}{2}}}, \qquad \text{where } \bar{X}_p = \sum_{i=1}^{p} X_i/p \text{ and}$$

$$\bar{Y}_q = \sum_{i=1}^{q} Y_i/q .$$

Thus, if there are p observations from X and q from Y, then the next observation is taken from X if

$$\frac{q}{p} \geq \left(\frac{\bar{Y}_q}{\bar{X}_p}\right)^{\frac{1}{2}} \text{ and from Y if}$$

$$(2.2.2)\quad \frac{q}{p} < \left(\frac{\bar{Y}_q}{\bar{X}_p}\right)^{\frac{1}{2}} .$$

Step 4. Using ii) of (1.5) as the stopping rule (we could have used i) of (1.5) as the rule),

define $C_k = 1.96$, where $k = p + q$.

Then stop sampling if

$$p \geq (\bar{X}_p + \sqrt{\bar{X}_q + \bar{Y}_q})(1.96)^2/d^2 \text{ and}$$

$$(2.2.3)\quad q \geq (\bar{Y}_q + \sqrt{\bar{X}_p + \bar{Y}_q})(1.96)^2/d^2 .$$

Step 5. Construct a confidence interval as follows:

(2.2.4) $I_k = (\bar{X}_p - \bar{Y}_q - d,\ \bar{X}_p - \bar{Y}_q + d)$.

The Table 2.2.2 shows the coverage probability, the average and the standard deviation of the total sample sizes for the sequential procedure, and the fixed sample size for various values of λ_1, λ_2 and d, each based on 1000 simulations. Also given are the average sample sizes on X and on Y due to the sequential procedure, and the fixed sample sizes on X and on Y, in parenthesis. The frequency distributions of the sample sizes are given in Figure 2.2.1 thru 2.2.4.

To get the fixed sample size of X and Y, consider the following:

$$\Pr(|\bar{X}_p - \bar{Y}_q - (\lambda_1 - \lambda_2)| \le d)$$

$$= \Pr(|\sum_{i=1}^{p} X_i - \frac{p}{q}\sum_{j=1}^{q} Y_j - p(\lambda_1 - \lambda_2)| \le p\ d)$$

$$= \sum_{t=0}^{\infty} \Pr(|\sum_{i=1}^{p} X_i - \frac{p}{q}\sum_{j=1}^{q} Y_j - p(\lambda_1-\lambda_2)| \le p\ d|\sum_{j=1}^{q} Y_j = t)\Pr(\sum_{j=1}^{q} Y_j = t)$$

(2.2.5)
$$= \sum_{t=0}^{\infty} \Pr(|\sum_{i=1}^{p} X_i - \frac{pt}{q} - p(\lambda_1-\lambda_2)| \le p\ d)\ \Pr(\sum_{j=1}^{q} Y_j = t)\ .$$

It is known that $\sum_{i=1}^{p} X_i$ is a Poisson random variable with parameter $p\lambda_1$ and $\sum_{j=1}^{q} Y_j$ is a Poisson random variable with parameter $q\lambda_2$.

Note that sampling rule (2.2.2) and the stopping rule (2.2.3) suggest that the ratio q/p be approximately equal to $(\lambda_2/\lambda_1)^{\frac{1}{2}}$ when sampling is stopped. Combining these results, approximate values of p and q may be computed such that equation (2.2.5) is close to 0.95. A computer program was developed for this purpose and is given in Appendix B.

As Table 2.2.1 shows, the coverage probabilities in general are very close to 0.95, and the average total sample sizes, and the average sample sizes from X and Y agree very well with the fixed sample sizes. Again, the shaded area in the plots indicates that portion in which the confidence interval of (2.2.4) did not cover the difference of λ_1 and λ_2.

Table 2.2.1

λ_1	λ_2	d	Coverage Proba-bility	Average Total	Sample X	Size Y	Standard Devia-tion	Fixed Total	Sample X	Size Y
2	2	1	.942	31.2 (	15.6	15.5)	4.4	28 (	14	14)
		.5	.947	123.4 (	61.7	61.7	8.0	120 (	60	60)
		.25	.957	492.6 (	246.2	246.4)	15.5	484 (	242	242)
2	4	2.0	.977	14.4 (	7.0	7.4)	.7	11 (	5	6)
		1.0	.952	45.3 (	18.8	26.5)	4.2	43 (	18	25)
		.5	.941	179.7 (	74.5	105.2)	8.2	174 (	72	102)
		.33	.934	403.8 (	167.3	236.5)	11.5	396 (	165	231)
		.25	.945	717.6 (	297.3	420.3)	15.4	710 (	294	416)
		.2	.939	1121.1 (	464.4	656.7)	20.3	1113 (	461	652)
1	4	1.0	.938	34.8 (	11.7	23.1)	4.4	33 (	11	22)
		.5	.945	138.2 (	46.1	92.1)	9.1	132 (	44	88)
		.33	.947	311.7 (	104.1	207.6)	12.3	303 (	101	202)

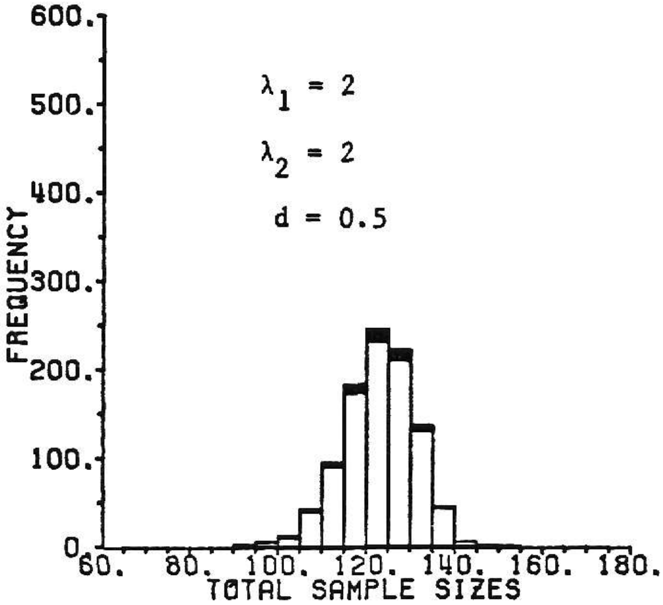

Figure 2.2.1

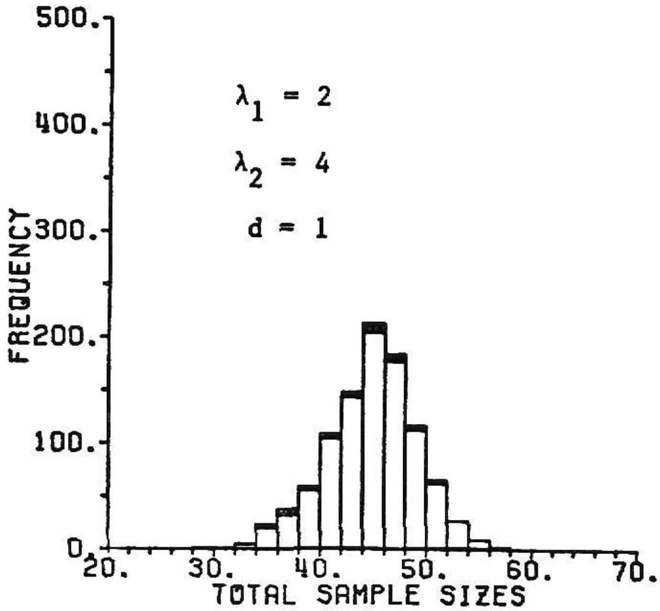

Figure 2.2.2

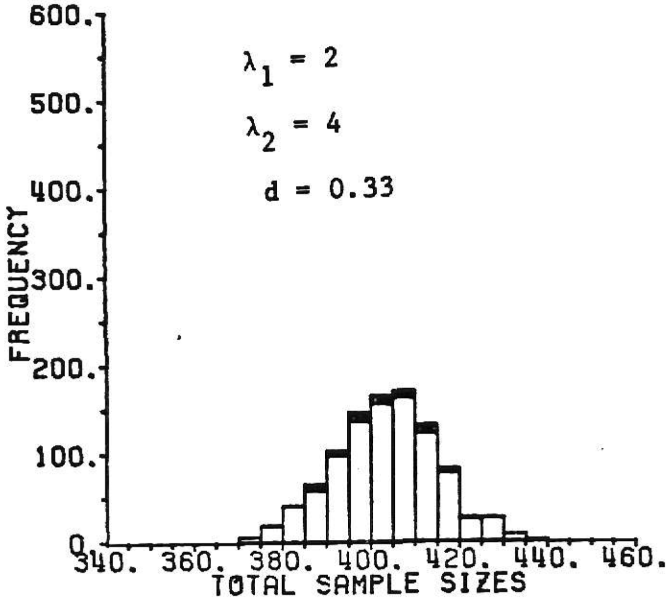

Figure 2.2.3

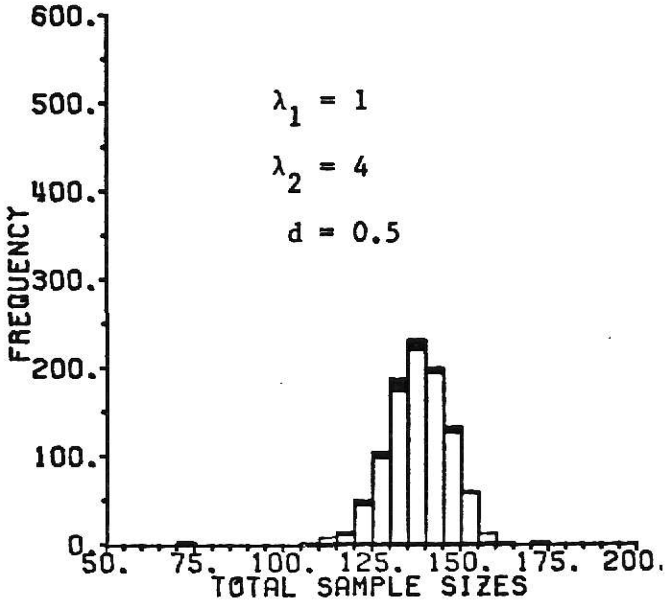

Figure 2.2.4

2.3. Exponential distribution case 1.

Let X_1, X_2, ... and Y_1, Y_2, ... be two independent random samples from Exp (λ_1) and Exp (λ_2), respectively, and $g(\underline{Y}) = g(\lambda_1, \lambda_2) = \lambda_1/\lambda_2$. Each step of the sequential procedure is given below.

Step 1. Let $\alpha = 0.05$

Step 2. Take 5 observations from X and from Y.

Step 3. First compute the following:

$$a = \frac{\partial(\lambda_1/\lambda_2)}{\partial\lambda_1} = \lambda_2^{-1} ,$$

$$b = \frac{\partial(\lambda_1/\lambda_2)}{\partial\lambda_2} = -\lambda_1/\lambda_2^2 ,$$

$$I^{-1}(\lambda_1) = \lambda_1^2 \quad \text{and} \quad I^{-1}(\lambda_2) = \lambda_2^2 .$$

Hence,

$$V_1(\lambda_1) = a\ I^{-1}(\lambda_1)a = (\lambda_1/\lambda_2)^2 \quad \text{and}$$

$$V_2(\lambda_2) = b\ I^{-1}(\lambda_2)b = (\lambda_1/\lambda_2)^2 .$$

Thus,

(2.3.1) $\hat{r} = r = \frac{1}{2}$.

Hence, if there are p observations from X and q observations from Y, the next observation is taken from X if

$$\frac{p}{p+q} \leq \frac{1}{2} \text{, and from Y if}$$

(2.3.2) $$\frac{p}{p+q} > \frac{1}{2} .$$

Step 4. First, define $C_k^2 = (1.96)^2 \dfrac{p+q+9}{p+q-9}$.

$$A = r^{-1}V_1(\lambda_1) + (1 - r)^{-1}V_2(\lambda_2)$$

$$= 4(\lambda_1/\lambda_2)^2 .$$

Because the M.L.E. of λ_1 is $\bar{X}_p^{-1}$ and the M.L.E. of λ_2 is $\bar{Y}_q^{-1}$ where $\bar{X}_p = \sum_{i=1}^{p} X_i/p$ and $\bar{Y}_q = \sum_{j=1}^{q} Y_j/q$,

$$\hat{A}_k = 4 \left(\frac{\bar{Y}_q}{\bar{X}_p}\right)^2 .$$

Thus, stop sampling if

(2.3.3) $$p + q \geq \frac{4(1.96)^2(\bar{Y}_q/\bar{X}_p)^2}{d^2} \; \frac{p+q+9}{p+q-9} .$$

Step 5. Construct a confidence interval as follows:

$$(2.3.4) \qquad I_k = \left(\frac{\bar{\bar{Y}}_q}{\bar{\bar{X}}_p} - d,\ \frac{\bar{\bar{Y}}_q}{\bar{\bar{X}}_p} + d\right)$$

For the purpose of simulation, we shall investigate how a sequence of random samples from a exponential distribution with parameter λ_1 could be generated and how that relates to (2.2.3) and (2.3.4). Let $e_1, e_2, \ldots,$ to be generated, be a sequence of random samples from the standard exponential distribution (Exp (1)). Then X_i, the i th sample from Exp (λ_1), could be generated by:

$$X_i = \lambda_1^{-1} e_i .$$

Now let $\Delta = \dfrac{\lambda_1}{d\lambda_2}$, then stopping rule (2.3.3) can be written as:

Stop sampling if

$$p + q \geq 4(1.96)^2 \, \Delta^2 \, (\bar{e}_2/\bar{e}_1)^2 \, \frac{p+q+9}{p+q-9} ,$$

where $\bar{e}_1 = \sum_{i=1}^{p} e_i/p$, $\bar{e}_2 = \sum_{j=1}^{q} e_j'/q$, and

$\{e_i\}$ and $\{e_j'\}$ are two independent sequences of random samples from Exp (1).

Again consider the following inequality to check whether the confidence interval I_k covers the ratio λ_1/λ_2.

$$\frac{\bar{Y}_q}{\bar{X}_p} - d \leq \frac{\lambda_1}{\lambda_2} \leq \frac{\bar{Y}_q}{\bar{X}_p} + d \text{ which can be written as}$$

$$\Delta(\bar{e}_2/\bar{e}_1) - 1 \leq \Delta \leq \Delta\ (\bar{e}_2/\bar{e}_1) + 1\ .$$

As in the normal case the sampling rule (2.3.2), the stopping rule (2.3.3), and the coverage probability of the confidence interval in (2.3.4) do not depend on specific values of λ_1, λ_2, and d as Δ is constant.

Thus, we need to consider only values of Δ instead of choosing combinations of λ_1, λ_2, and d.

To get the fixed sample size, consider the following:

$$\Pr(|\bar{Y}_q/\bar{X}_p - \lambda_1/\lambda_2| \leq d)$$

$$= \Pr(|\lambda_1\bar{e}_2/(\lambda_2\bar{e}_1) - \lambda_1/\lambda_2| \leq d)$$

(2.3.5)

$$= \Pr(|\bar{e}_2/\bar{e}_1 - 1| \leq \Delta^{-1}).$$

Since $\bar{e}_2$ is a gamma random variable with parameters q and q, and $\bar{e}_1$ is a gamma random variable with parameters p and p, it can easily be shown that $\bar{e}_2/\bar{e}_1$ is a central F random variable with 2q and 2p degrees of freedom. Therefore, (2.3.5) may be rewritten as:

(2/3/6) $\Pr(|F_{(2q,\ 2p)} - 1| \leq \Delta^{-1})$.

Considering $r = \frac{1}{2}$, p and q can be choosen such that (2.3.6) is as near 0.95 as possible.

The Table 2.3.1 shows the simulation results with various values of Δ. The frequency distribution of the total sample sizes are given in Figure 2.3.1 through 2.3.4. Table 2.3.1 shows that the coverage probabilities are close to 0.95 and the average total sample sizes are reasonable compare to those from the fixed sample size procedure.

Table 2.3.1

λ	Coverage Probability	Average Total Sample Size	Standard Deviation	Fixed Sample Size*
.5	1.000	16.7	5.5	10
1.0	1.000	29.9	12.2	24
1.5	.961	48.7	20.3	42
2.0	.942	74.8	29.1	68
2.5	.943	109.9	37.3	102
3.0	.930	151.3	46.7	144
3.5	.924	198.9	56.6	194
4.0	.931	258.6	64.1	252
4.5	.934	327.1	72.5	318
5.0	.927	397.9	84.2	390
7.0	.961	768.1	107.3	758
9.0	.951	1255.3	140.2	1250

*To compute the fixed sample size, a Monro Calculator (Model 1730) was used for λ = 0.5 through λ = 5.0 and PROBF in SAS was used for λ = 7.0 and λ = 9.0.

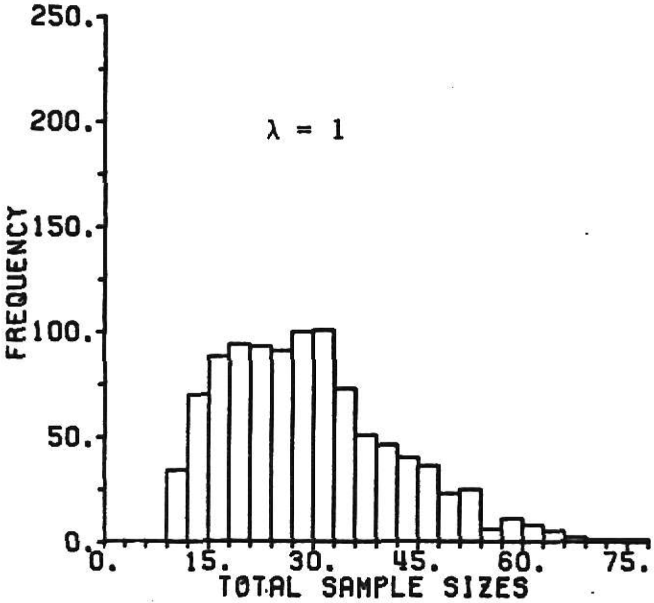

Figure 2.3.1

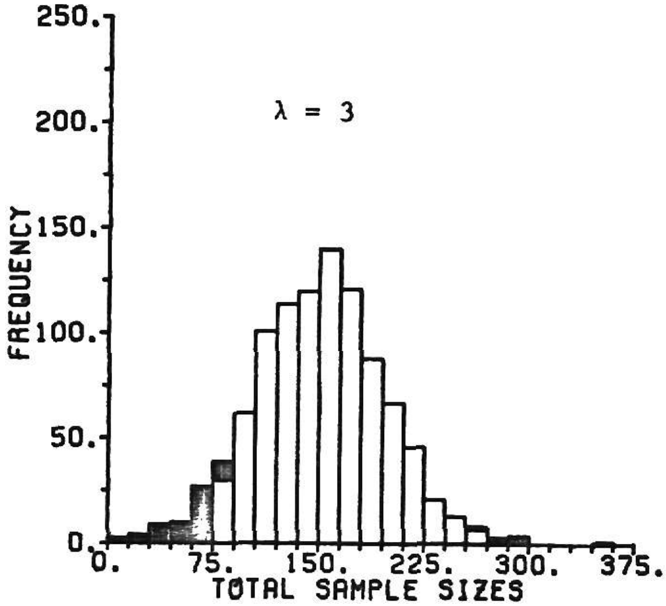

Figure 2.3.2

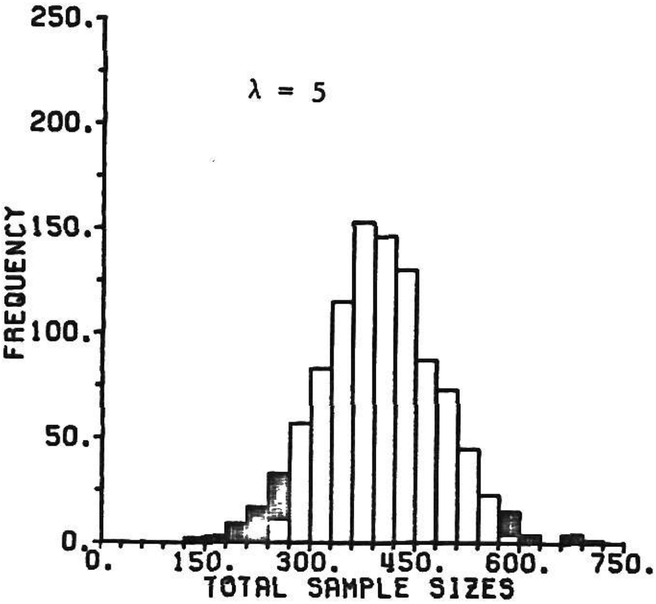

Figure 2.3.3

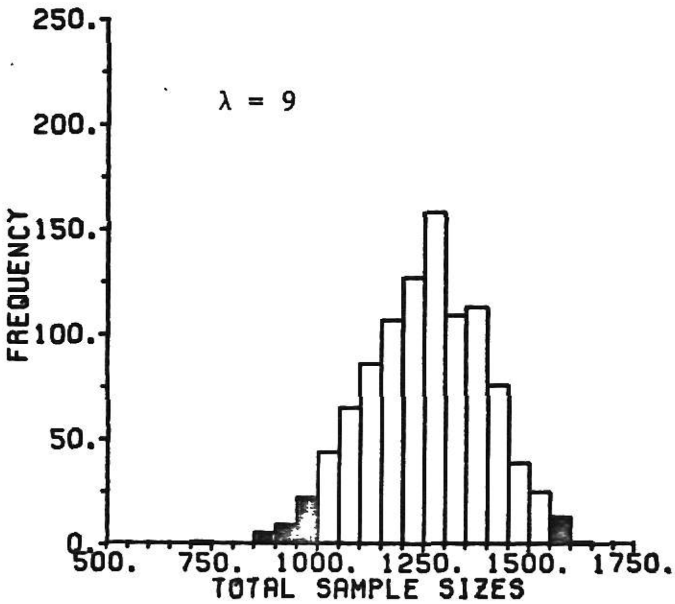

Figure 2.3.4

2.4. Exponential distribution case 2.

Let $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ be two independent random samples from Exp (λ_1) and Exp (λ_2), respectively, and $g(\underline{\gamma}) = g(\lambda_1, \lambda_2) = \lambda_1 - \lambda_2$.

The sequential procedure is given below.

Step 1. Let $\alpha = 0.05$

Step 2. Take 7 observations from X and from Y.

Step 3. Compute the following:

$$a = \frac{\partial(\lambda_1 - \lambda_2)}{\partial\lambda_1} = 1 ,$$

$$b = \frac{\partial(\lambda_1 - \lambda_2)}{\partial\lambda_2} = -1 ,$$

$$I^{-1}(\lambda_1) = \lambda_1^2 \text{ and } I^{-1}(\lambda_2) = \lambda_2^2 .$$

Hence,

$$r = \frac{\lambda_1}{\lambda_1 + \lambda_2} \text{ and}$$

(2.4.1) $$\hat{r} = \frac{\bar{X}_p^{-1}}{\bar{X}_p^{-1} + \bar{Y}_q^{-1}} = \frac{\bar{Y}_q}{\bar{X}_p + \bar{Y}_q}$$

where $\bar{X}_p = \sum_{i=1}^{p} X_i/p$ and $\bar{Y}_q = \sum_{j=1}^{q} Y_j/q$.

Thus, the sampling rule is: if there are p observations from X and q from Y, take the next observation from X if

(2.4.2) $$\frac{q}{p} \geq \frac{\bar{X}_p}{\bar{Y}_q}$$ and from Y if

$$\frac{q}{p} < \frac{\bar{X}_p}{\bar{Y}_q} .$$

Step 4. Let $C_k = 1.96$ with $k = p + q$.

Next, compute A:

$$A = r^{-1}V_1(\lambda_1) + (1 - r)^{-1}V_2(\lambda_2)$$

$$= (\lambda_1 + \lambda_2)^2 .$$

Hence,

$$\hat{A}_k = (\bar{X}_p^{\,-1} + \bar{Y}_q^{\,-1})^2 .$$

Thus, stop sampling if

(2.4.3) $$p + q \geq \frac{(1.96)^2(\bar{X}_p^{\,-1} + \bar{Y}_q^{\,-1})^2}{d^2}$$

Step 5. Construct the confidence interval as follows:

$$I = (\bar{X}_p^{\,-1} - \bar{Y}_q^{\,-1} - d,\ \bar{X}_p^{\,-1} - \bar{Y}_q^{\,-1} + d).$$

For the purpose of simulation, let $\psi = \lambda_2/\lambda_1$ and $\zeta = \lambda_1/d$.

Then (2.4.2) can be rewritten as: Take the next observation from X or from Y according to whether $\frac{q}{p} \geq \psi \frac{\bar{e}_1}{\bar{e}_2}$ or $\frac{q}{p} < \psi \frac{\bar{e}_1}{\bar{e}_2}$, respectively, where $\bar{e}_1$ and $\bar{e}_2$ are defined as in the previous section. Stop sampling if $p + q \geq (1.96)^2 \zeta^2(\bar{e}_1^{-1} + \psi \bar{e}_2^{-1})^2$ according to stopping rule (2.4.3).

To find the coverage probability of the confidence interval I_k, consider the following:

$$Pr(\bar{X}_p^{-1} - \bar{Y}_q^{-1} - d \leq \lambda_1 - \lambda_2 \leq \bar{X}_p^{-1} - \bar{Y}_q^{-1} + d)$$

$$= Pr(\lambda_1\bar{e}_1^{-1} - \lambda_2\bar{e}_2^{-1} - d \leq \lambda_1 - \lambda_2 \leq \lambda_1\bar{e}_1^{-1} - \lambda_2\bar{e}_2^{-1} + d)$$

$$(2.4.4) \quad = Pr(\bar{e}_1^{-1} - \psi \bar{e}_2^{-1} - \zeta^{-1} \leq 1 - \psi \leq \bar{e}_1^{-1} - \psi \bar{e}_2^{-1} + \zeta^{-1})$$

As has been seen, the sampling rule, the stopping rule, and the coverage probability depend on λ_1, λ_2 and, d only through ψ and ζ. Thus, only various values of ψ and ζ are to be considered without specifying the values of λ_1, λ_2 and d.

To find the fixed sample size, consider the following.

$$\Pr(|\bar{X}_p^{-1} - \bar{Y}_q^{-1} - (\lambda_1 - \lambda_2)| \leq d)$$

$$= \Pr(|\lambda_1\bar{e}_1^{-1} - \lambda_2\bar{e}_2^{-1} - (\lambda_1 - \lambda_2)| \leq d)$$

$$= \Pr(|\bar{e}_1^{-1} - \psi\bar{e}_2^{-1} - (1 - \psi)| \leq \zeta^{-1})$$

$$= \int_0^\infty \Pr(|\bar{e}_1^{-1} - \psi\bar{e}_2^{-1} - (1 - \psi)| \leq \zeta^{-1} | \bar{e}_2 = y) h(y;\, q,\, q)\, dy$$

$$= \int_0^\infty \Pr(|\bar{e}_1^{-1} - \psi y^{-1} - (1 - \psi)| \leq \zeta^{-1}) h(y;\, q,\, q)\, dy$$

$$= \int_0^\infty \Pr(\psi y^{-1} + (1 - \psi) - \zeta^{-1} \leq \bar{e}_1^{-1} \leq y^{-1} + (1-\psi) + \zeta^{-1}) h(y;\, q,q)\, dy$$

(2.4.5)
$$= \int_0^\infty \Pr\left|(y^{-1} + (1-\psi) - \zeta^{-1})^{-1} \leq \bar{e}_1 \leq (y^{-1} + (1-\psi) + \zeta^{-1})^{-1}\right| h(y;\, q,\, q)\, dy,$$

where $h(y;\, q,\, q)$ is the probability density function of a gamma distribution with parameters q and q.

We know that $\bar{e}_1$ is a gamma random variable with parameters p and p. Also note that the sampling rule and the stopping rule suggest the ratio q/p be approximately equal to λ_2/λ_1 when sampling is stopped. Hence, it is possible to choose approximate values of p and q such that (2.4.5) is near 0.95.

Because equation (2.4.5) is mathematically complex, a computer program was developed to solve for p and q. In Appendix B, a modified form of (2.4.5) which was used for computer solution is given.

Table 2.4.1 shows the simulation results with various values of ψ and ζ. Figures 2.4.1 through 2.4.8 show the frequency distribution of the total sample sizes.

Table 2.4.1

$\psi(=\frac{\lambda_2}{\lambda_1})$	$\zeta(=\frac{\lambda_1}{d})$	Coverage Probability	Sample Size Total	X	Y	Fixed Sample Size Total	X	Y
1.0	.5	.994	14.1	(7.0	7.1)	10	(5	5)
	1.0	.976	19.2	(9.6	9.6)	22	(11	11)
	1.5	.964	36.2	(18.1	18.1)	44	(22	22)
	2.0	.957	62.7	(31.3	31.4)	70	(35	35)
	2.5	.962	97.2	(48.7	48.5)	106	(53	53)
	3.0	.941	140.2	(70.0	70.2)	148	(74	74)
	3.5	.960	189.6	(94.9	94.7)	198	(99	99)
	4.0	.959	246.9	(123.4	123.5)	258	(129	129)
	4.5	.954	311.6	(156.1	155.5)	322	(161	161)
.5	.5	.992	14.0	(7.0	7.0)	9	(6	3)
	1.0	.990	15.2	(8.1	7.1)	15	(10	5)
	1.5	.982	22.6	(14.0	8.6)	27	(18	9)
	2.0	.953	36.3	(23.7	12.6)	45	(30	15)
	2.5	.960	55.1	(36.4	18.7)	63	(42	21)
	3.0	.947	79.0	(52.1	26.8)	87	(58	29)
	3.5	.946	107.2	(71.1	36.1)	114	(76	38)
	4.0	.959	139.9	(92.8	47.1)	147	(98	49)
	4.5	.949	176.6	(117.1	59.4)	186	(124	62)
	5.0	.960	217.0	(144.4	72.7)	225	(150	75)
	5.5	.956	262.3	(174.3	88.0)	270	(180	90)

Contd..

$\psi(=\frac{\lambda_2}{\lambda_1})$	$\zeta(=\frac{\lambda_1}{d})$	Coverage Proba-bility	Sample Size Total	X	Y	Fixed Sample Size Total	X	Y
.25	1.5	.984	18.0	(10.9	7.1)	25	(20	5)
	2.0	.957	26.4	(18.8	7.6)	30	(24	6)
	2.5	.958	39.4	(30.2	9.2)	45	(36	9)
	3.0	.939	55.3	(43.5	11.8)	65	(52	13)
	3.5	.949	75.1	(59.5	15.6)	85	(68	17)
	4.0	.954	97.2	(77.0	20.2)	105	(84	21)
	4.5	.945	122.8	(97.7	25.1)	130	(104	26)
	5.0	.957	151.7	(120.7	31.0)	160	(128	32)

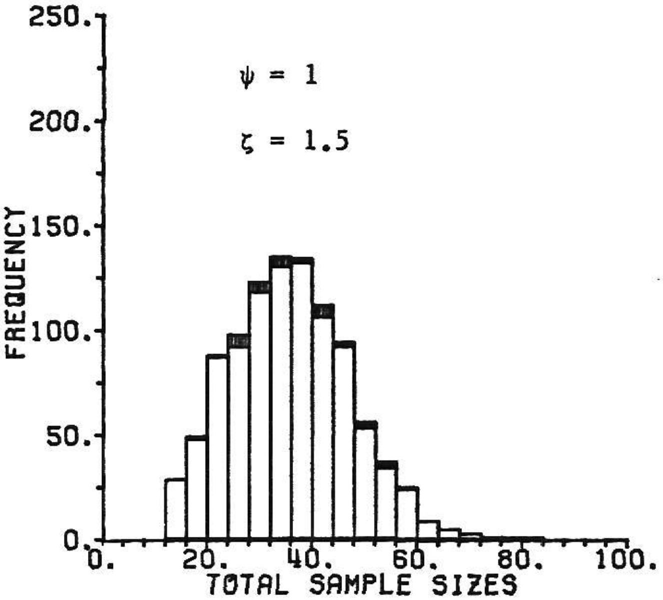

Figure 2.4.1

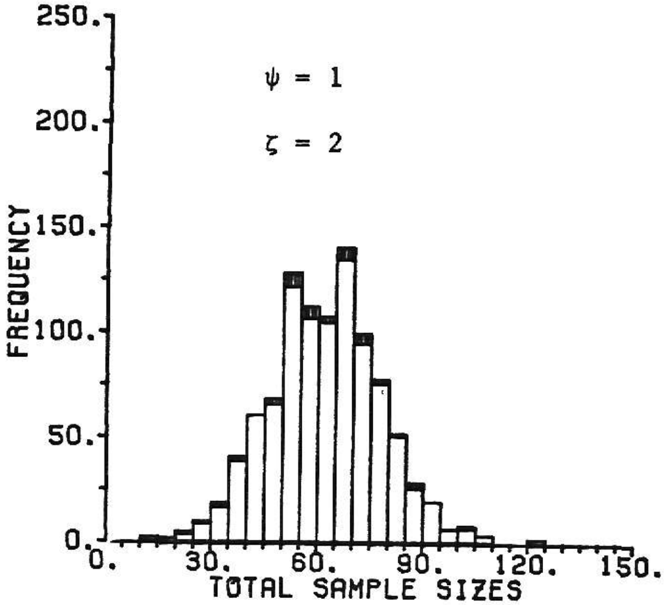

Figure 2.4.2

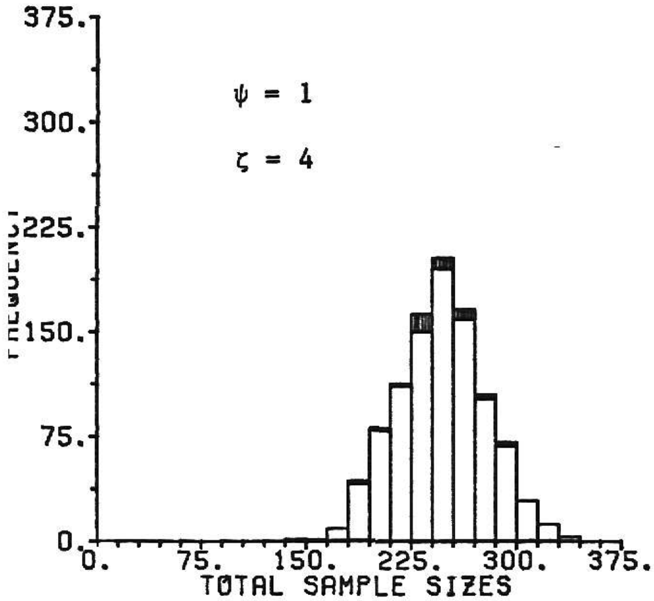

Figure 2.4.3

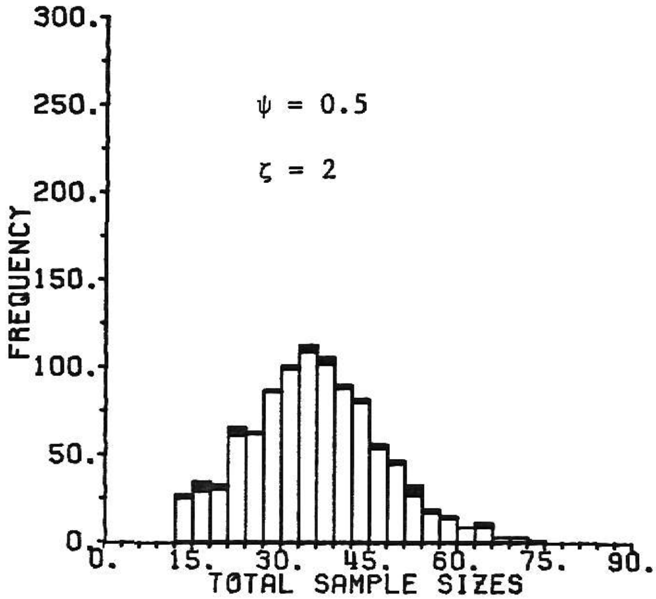

Figure 2.4.4

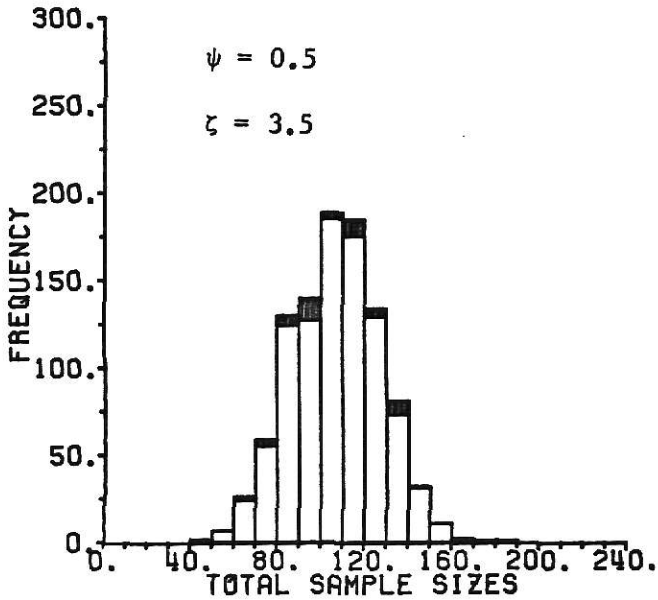

Figure 2.4.5

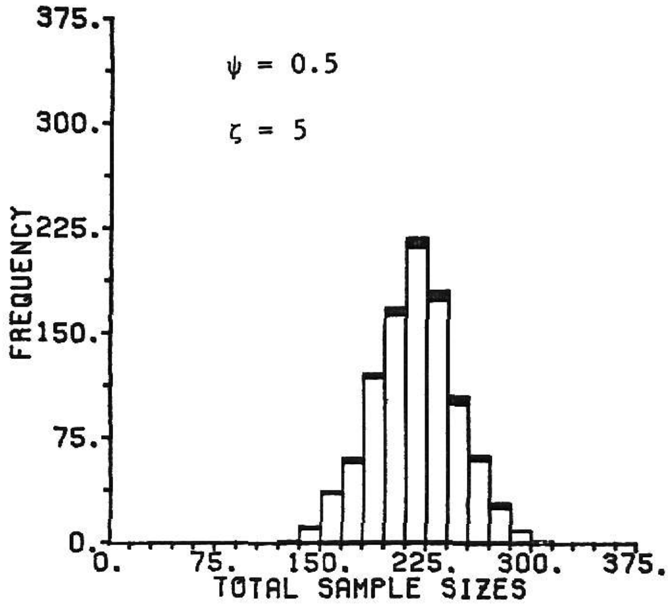

Figure 2.4.6

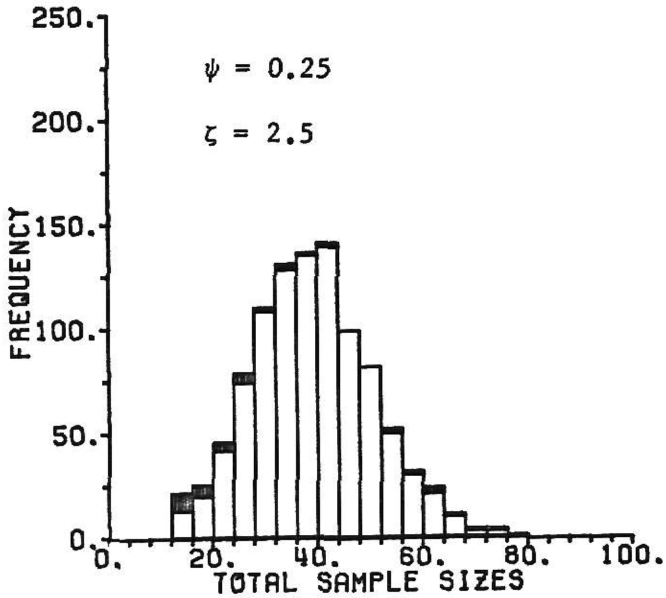

Figure 2.4.7

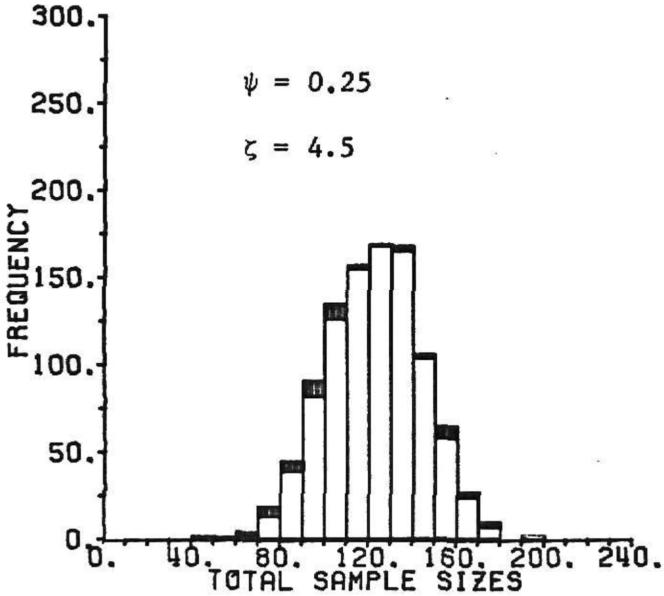

Figure 2.4.8

2.5. Conclusion

A Monte Carlo study was used to evaluate finite sample sizes performance of the proposed sequential procedure.

Four cases of real valued function of parameters were considered. In all four cases and all parameters the simulation results indicated that the coverage probabilities were very close to 0.95, the expected coverage probability. The average total sample sizes were very reasonable compare to those from the fixed sample size procedure.

REFERENCES

(1) Kemp, K. (1977). Class note on 'Digital Statistical Analysis', Department of Statistics, Kansas State University.

(2) Marsaglia, G., Ananthanarayanan, K., Paul, N. (1977). The McGill Random Number Package "Super-Duper", School of Computer Science, McGill University

(3) Perng, S., Hasza, D. (1978), Paik, C. Sequential Fixed-Width Confidence Interval for a Function of Parameters, Unpublished Paper, Department of Statistics, Kansas State University.

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

Appendix A

```
                C                                                              *
                C  COMPUTER PROGRAM FOR NORMAL DISTRIBUTION CASE               *
                C  INPUT - SEED NOS,LAMBDA,C(=1.96),INITIAL OBS.               *
                C  OUTPUT :                                                    *
                C     PRINTER - SAMPLE SIZE ON TOTAL,ON X AND ON Y,VARIA*
                C       NCE OF X AND OF Y,COVERAGE INDICATOR OF THE C.I.,*
                C       AVERAGE SAMPLE SIZE ON TOTAL,ON X AND ON Y,COVERA*
                C       GE PROBABILITY                                         *
                C     TAPE - ALL INPUT,SIMULATION NO,SAMPLE SIZE ON X          *
                C       AND ON Y,MEAN AND VARIANCE OF X AND OF Y               *
                C                                                              *
0001                 IMPLICIT REAL*8(B-H,O-Z)
0002                 INTEGER CRDR/5/,PRTR/6/,TAPE/10/
0003                 INTEGER STAR/'*'/
0004                 DATA TPO,TP,COUNT/3*0.0/
0005             100 FORMAT(2I15,3F5.0,T31,3F5.0)
0006             200 FORMAT('1 SEED NOS. =',2I18/'- LAMBDA=',F5.1,'  C'
                    *,'=',F5.3,'  INITIAL OBS =',F3.0)
0007             251 FORMAT(' ',4I9,F10.4,F9.4)
0008             351 FORMAT(' ',5X,'SIM.NO.    TOT.     POP.1     POP.2',
                    *  '     VAR.1     VAR.2  UB<T.R.  LB>T.R.'/)
0009             400 FORMAT('-',6X,'TOT.AVG.SAMP.SIZE =',F6.1/22X,
                    *'POP.1=',F8.1/22X,'POP.2=',F8.1/6X,'COVERAGE',
                    *' PROB. =',F6.3)
0010             450 FORMAT('+',64X,'-    ',A1)
0011             451 FORMAT('+',60X,A1,'    -')
0012                 READ(5,100) II,JJ,RAMDA,C,DIN,ARAMDA,AC,ADIN
0013                 WRITE(6,200)II,JJ,RAMDA,C,DIN
0014                 WRITE(6,351)
0015                 WRITE(10)ARAMDA,AC,ADIN
0016                 CCDD=C*C*8.D0*RAMDA*RAMDA
0017                 ODR=1.D0/RAMDA
0018                 N=DIN
0019                 SIMNO=1000.D0
0020                 CALL RSTART(II,JJ)
0021                 DO 1000 IJK=1,1000
0022                 SUMX=0.D0
0023                 SUMY=0.D0
0024                 SSQX=0.D0
0025                 SSQY=0.D0
0026                 VARX=0.D0
0027                 VARY=0.D0
0028                 PQ=1.D0
0029                 P=1.D0
0030                 SUMX=RNOR(0)
0031                 Q=1.D0
0032                 SUMY=RNOR(0)
0033                 PQ=PQ+1.D0
0034                 DO 20 J=2,N
0035                 DJ=J
0036                 DJJ=DJ*DJ-DJ
0037                 P=P+1.D0
0038                 PQ=PQ+1.D0
0039                 OBS=RNOR(0)
0040                 SUMX=SUMX+OBS
0041                 SQ=DJ*OBS-SUMX
0042                 SSQX=SSQX+SQ*SQ/DJJ
0043                 VARX=SSQX/P
```

```
              Q=Q+1.D0
              PQ=PQ+1.D0
              OBS=RNOR(0)
              SUMY=SUMY+OBS
              SQ=DJ*OBS-SUMY
              SSQY=SSQY+SQ*SQ/DJJ
              VARY=SSQY/Q
           20 CONTINUE
              SXDSY=VARX/VARY
              COMP=CCDD*SXDSY*SXDSY*(PQ+13.D0)/(PQ-13.D0)
              IF(PQ.GE.COMP)GO TO 999
              U=UNI(0)
              IF(U.GE. .5)GOTO 2
            1 P=P+1.D0
              PQ=PQ+1.D0
              OBS=RNOR(0)
              SUMX=SUMX+OBS
              SQ=P*OBS-SUMX
              SSQX=SSQX+SQ*SQ/(P*P-P)
              VARX=SSQX/P
              SXDSY=VARX/VARY
              COMP=CCDD*SXDSY*SXDSY*(PQ+13.D0)/(PQ-13.D0)
              IF(PQ.GE.COMP)GO TO 999
            2 Q=Q+1.D0
              PQ=PQ+1.D0
              OBS=RNOR(0)
              SUMY=SUMY+OBS
              SQ=Q*OBS-SUMY
              SSQY=SSQY+SQ*SQ/(Q*Q-Q)
              VARY=SSQY/Q
              SXDSY=VARX/VARY
              COMP=CCDD*SXDSY*SXDSY*(PQ+13.D0)/(PQ-13.D0)
              IF(PQ.GE.COMP)GO TO 999
              GO TO 1
          999 CONTINUE
              NPQ=PQ
              NP=P
              NQ=Q
              WRITE(PRTR,251)IJK,NPQ,NP,NQ,VARX,VARY
              TPQ=TPQ+PQ
              TP=TP+P
              IF(DABS(SXDSY-ODR)-1.D0)30,30,25
           25 IF(SXDSY-ODR)40,40,35
           35 WRITE(6,450)STAR
              COUNT=COUNT+1.D0
              GO TO 30
           40 WRITE(6,451)STAR
              COUNT=COUNT+1.D0
           30 AIJK=IJK
              APQ=PQ
              AP=P
              ASUMX=SUMX
              ASUMY=SUMY
              AVARX=VARX
              AVARY=VARY
              WRITE(TAPE)AIJK,APQ,AP,ASUMX,AVARX,ASUMY,AVARY
         1000 CONTINUE
              TQ=TPQ-TP
              TPQ=TPQ/SIMNO
              TP=TP/SIMNO
              TQ=TQ/SIMNO
              TC=1.-COUNT/SIMNO
              WRITE(PRTR,400)TPQ,TP,TQ,TC
              STOP
              END
```

```
FORTRAN IV G LEVEL  21                    MAIN                    DATE = 78176
C                                                                        *
C COMPUTER PROGRAM FOR POISSON DISTRIBUTION CASE                         *
C INPUT - SEED NOS,LAMBDA1,LAMBDA2,WIDTH,C(=1.96),INIT*
C     IAL OBS.                                                           *
C OUTPUT :                                                               *
C   PRINTER - SAMPLE SIZE ON TOTAL,ON X AND ON Y,MEAN *
C     OF X AND OF Y,CONFIDENCE LIMITS,COVERAGE PROB.                     *
C   TAPE - ALL INPUT,SIMULATION NO,SAMPLE SIZE ON X                      *
C     AND ON Y SUM OF X AND OF Y                                         *
C                                                                        *
      IMPLICIT REAL*8(B-H,O-Z)
      DIMENSION CDFX(60),CDFY(60)
      INTEGER STAR/'*'/
      DATA TP,TQ,COUNT/3*0.D0/
  100 FORMAT(2I15,5F5.0,T31,5F5.0)
  200 FORMAT('1 SEED NOS. =',2I18/'- LAMBDA 1 =',F5.1,
     *'  LAMBDA 2 =',F5.1,'  D =',F6.3,'  C =',F5.3,
     *'  INITIAL OBS. =',F3.0)
  250 FORMAT(' ',5X,'SIM.NO.    TOT.     POP.1     POP.2',
     *'    MEAN.1    MEAN.2            L             U'/)
  251 FORMAT(' ',4I9,2F9.3,F13.3,F8.3)
  300 FORMAT('+',77X,A1)
  400 FORMAT('-',6X,'TOT.AVG.SAMP.SIZE =',F6.1/22X,
     *'POP.1=',F8.1/22X,'POP.2=',F8.1/6X,'COVERAGE',
     *' PROB. =',F6.3)
      READ(5,100)II,JJ,DLMDAX,DLMDAY,D,C,DIN,ALMDAX,
     *     ALMDAY,AD,AC,ADIN
      WRITE(6,200)II,JJ,DLMDAX,DLMDAY,D,C,DIN
      WRITE(10)ALMDAX,ALMDAY,AD,AC,AIN
      CALL POISSN(CDFX,DLMDAX)
      IF(DLMDAY .NE. DLMDAX)GO TO 25
      DO 23 I=1,60
   23 CDFY(I)=CDFX(I)
      GO TO 26
   25 CALL POISSN(CDFY,DLMDAY)
   26 CCDD=C*C/(D*D)
      WRITE(6,250)
      DIFF=DLMDAX-DLMDAY
      SIMNO=1000.D0
      CALL RSTART(II,JJ)
      INOBS=DIN
      DO 1000 IJK=1,1000
      SUMX=0.D0
      SUMY=0.D0
      DO 500 ILL=1,INOBS
      UNFX=UNI(0)
      DO 35 I=1,60
      IF(CDFX(I)-UNFX)35,40,40
   35 CONTINUE
   40 X=DFLOAT(I-1)
      SUMX=SUMX+X
      UNFY=UNI(0)
      DO 75 I=1,60
      IF(CDFY(I)-UNFY)75,80,80
   75 CONTINUE
   80 Y=DFLOAT(I-1)
      SUMY=SUMY+Y
  500 CONTINUE
```

```
      P=DIN
      Q=DIN
      XBAR=SUMX/P
      YBAR=SUMY/Q
    9 PQ=P+Q
      XY=DSQRT(XBAR*YBAR)
      XCOND=(XBAR+XY)*CCDD
      YCOND=(YBAR+XY)*CCDD
      IF(P.GT.XCOND .AND. Q.GT.YCOND)GOTO 900
      IF(XBAR.EQ.0.D0)GO TO 2
      YDX=YBAR/XBAR
      QDP=Q*Q/(P*P)
      IF(YDX-QDP)1,1,2
    1 UNFX=UNI(0)
      DO 13 I=1,60
      IF(CDFX(I)-UNFX)13,14,14
   13 CONTINUE
   14 X=DFLOAT(I-1)
      SUMX=SUMX+X
      P=P+1.D0
      XBAR=SUMX/P
      GO TO 9
    2 UNFY=UNI(0)
      DO 18 I=1,60
      IF(CDFY(I)-UNFY)18,19,19
   18 CONTINUE
   19 Y=DFLOAT(I-1)
      SUMY=SUMY+Y
      Q=Q+1.D0
      YBAR=SUMY/Q
      GO TO 9
  900 CONTINUE
      NPQ=PQ
      NP=P
      NQ=Q
      DL=XBAR-YBAR-D
      DU=DL+D+D
      WRITE(6,251)IJK,NPQ,NP,NQ,XBAR,YBAR,DL,DU
      TP=TP+P
      TQ=TQ+Q
      IF(DL.LE.DIFF .AND. DU.GE.DIFF) GO TO 111
      WRITE(6,300)STAR
      COUNT=COUNT+1.D0
  111 AIJK=IJK
      AP=P
      AQ=Q
      ASUMX=SUMX
      ASUMY=SUMY
      WRITE(10)AIJK,AP,AQ,ASUMX,ASUMY
 1000 CONTINUE
      TPQ=TP+TQ
      TPQ=TPQ/SIMNO
      TP=TP/SIMNO
      TQ=TQ/SIMNO
      TC=1.-COUNT/SIMNO
      WRITE(6,400)TPQ,TP,TQ,TC
      STOP
      END
```

```
FORTRAN IV G LEVEL  21                  POISSN              DATE = 78176

                   SUBROUTINE POISSN(CDF,LAMBDA)
      C                                                                *
      C THIS SUBROUTINE COMPUTES THE PROBABILITY FUNCTION              *
      C AND THE CUMMULATIVE DENSITY FUNCTION OF THE POISSON            *
      C DISTRIBUTION WITH PARAMETER LAMBDA.                            *
      C                                                                *
                   IMPLICIT REAL*8(A-H,O-Z)
                   REAL*8 LAMBDA,CDF(60),PDF(60)
                   IF(LAMBDA.LE.0.00001D0)GOTO3
                   P=DEXP(-LAMBDA)
                   CDF(1)=P
                   PDF(1)=P
                   DO 1 I=1,58
                   P=P*LAMBDA/DFLOAT(I)
                   CDF(I+1)=CDF(I)+P
                   PDF(I+1)=P
                   PR=CDF(I+1)
                   IF(PR.GE. 0.999999D0)GOTO2
                 1 CONTINUE
                   WRITE(6,100)CDF(59)
               100 FORMAT(' 59 INTERVALS ARE NOT ENOUGH.',F11.8)
                   STOP
                 2 CDF(I+2)=1.D0
                   PDF(I+2)=0.D0
                   RETURN
                 3 WRITE(6,400)
               400 FORMAT('*** LAMDA TOO SMALL FOR THE ALGORITHM')
                   STOP
                   END
```

```
FORTRAN IV G LEVEL  21                    MAIN                    DATE = 78176

      C COMPUTER PROGRAM FOR EXPONENTIAL DISTRIBUTION CASE 1*
      C INPUT - SEED NOS.LAMBDA.C(=1.96).INITIAL OBS.          *
      C OUTPUT :                                               *
      C    PRINTER - SAMPLE SIZE ON TOTAL.ON X AND ON Y.MEAN   *
      C      OF X AND OF Y.COVERAGE INDICATOR OF THE C.I..AVER*
      C      AGE TOTAL SAMPLE SIZE, SAMPLE SIZE ON X AND ON Y  *
      C    TAPE - ALL INPUT,SIMULATION NO,SAMPLE SIZE ON X     *
      C      AND ON Y,SUM OF X AND OF Y                        *
            IMPLICIT REAL*8(B-H,O-Z)
            INTEGER STAR/'*'/
            DATA TP,TQ,COUNT/3*0.D0/
        100 FORMAT(2I15,3F5.0,T31,3F5.0)
        200 FORMAT('1 SEED NOS. =',2I18/'- LAMBDA=',F5.1,'   C'
           $,'=',F5.3,'   INITIAL OBS =',F3.0)
        251 FORMAT(' ',4I9,F10.4,F9.4)
        351 FORMAT(' ',5X,'SIM.NO.    TOT.     POP.1     POP.2'
           $,'    MEAN.1    MEAN.2.  UB<T.R. LB>T.R.'/)
        400 FORMAT('-',6X,'TOT.AVG.SAMP.SIZE =',F6.1/22X,
           $'POP.1=',F8.1/22X,'POP.2=',F8.1/6X,'COVERAGE',
           $' PROB. =',F6.3)
        450 FORMAT('+',64X,'-     ',A1)
        451 FORMAT('+',60X,A1,'    -')
            READ(5,100)II,JJ,RAMDA,C,DIN,ARAMDA,AC,ADIN
            WRITE(6,200)II,JJ,RAMDA,C,DIN
            WRITE(10)ARAMDA,AC,ADIN
            WRITE(6,351)
            CC=C*C*RAMDA*RAMDA*4.D0
            E=1.D0/RAMDA
            SIMNO=1000.D0
            ISIM=SIMNO
            CALL RSTART(II,JJ)
            INOBS=DIN
            DO 1000 IJK=1,ISIM
            SUMX=REXP(O)
            SUMY=REXP(O)
            DO 500 ILL=2,INOBS
            SUMX=SUMX+REXP(O)
            SUMY=SUMY+REXP(O)
        500 CONTINUE
            P=DIN
            Q=DIN
            PQ=P+Q
            UX=SUMX/P
            UY=SUMY/Q
            U=UNI(O)
            IF(U.GE.0.5D0)GOTO 2
          1 RATIO=UY/UX
            RR=RATIO*RATIO*CC*(PQ+ 9.D0)/(PQ- 9.D0)
            IF(PQ.GE.RR)GOTO 900
            P=P+1.D0
            PQ=PQ+1.D0
            SUMX=SUMX+REXP(O)
            UX=SUMX/P
          2 RATIO=UY/UX
            RR=RATIO*RATIO*CC*(PQ+ 9.D0)/(PQ- 9.D0)
            IF(PQ.GE.RR)GOTO 900
            Q=Q+1.D0
            PQ=PQ+1.D0
```

```
          SUMY=SUMY+REXP(Q)
          UY=SUMY/Q
          GO TO 1
      900 CONTINUE
          TP=TP+P
          TQ=TQ+Q
          NPQ=PQ
          NP=P
          NQ=Q
          WRITE(6,251)IJK,NPQ,NP,NQ,UX,UY
          COMP=RATIO-1.D0
          IF(DABS(COMP)-E)30,30,20
       20 IF(COMP)40,40,35
       35 WRITE(6,450)STAR
          COUNT=COUNT+1.D0
          GO TO 30
       40 WRITE(6,451)STAR
          COUNT=COUNT+1.D0
       30 AIJK=IJK
          AP=P
          AQ=Q
          ASUMX=SUMX
          ASUMY=SUMY
          WRITE(10)AIJK,AP,AQ,ASUMX,ASUMY
     1000 CONTINUE
          TPQ=TP+TQ
          TPQ=TPQ/SIMNO
          TP=TP/SIMNO
          TQ=TQ/SIMNO
          TC=1.D0 - COUNT/SIMNO
          WRITE(6,400)TPQ,TP,TQ,TC
          STOP
          END
```

```
FORTRAN IV G LEVEL  21                  MAIN                DATE = 78176
C                                                                     *
C COMPUTER PROGRAM FOR EXPONENTIAL DISTRIBUTION CASE 2*
C   INPUT - SEED NOS.,LAMBDA,PHI,C(=1.96),INITIAL OBS.*
C OUTPUT :                                                            *
C   PRINTER - SAMPLE SIZE ON TOTAL,ON X AND ON Y, MEAN*
C     OF X AND OF Y,COVERAGE INDICATOR OF THE C.I.,COVE*
C     RAGE PROBABILITY                                                *
C   TAPE - ALL INPUT,SIMULATION NO,SAMPLE SIZE ON X    *
C     AND ON Y,SUM OF X AND OF Y                                      *
C                                                                     *
      IMPLICIT REAL*8(B-H,O-Z)
      INTEGER STAR/'*'/
      DATA TP,TQ,COUNT/3*0.D0/
  100 FORMAT(2I15,4F5.0,T31,4F5.0)
  200 FORMAT('1 SEED NOS. =',2I18/'- LAMBDA =',F5.1,
     *'   PHI=',F5.3,'  C=',F5.3,'  INITIAL OBS.=',F3.0)
  251 FORMAT(' ',4I9,F10.4,F9.4)
  351 FORMAT(' ',5X,'SIM.NO.    TOT.    POP.1    POP.2'
     *,'   MEAN.1   MEAN.2  UB<T.R.  LB>T.R.'/)
  400 FORMAT('-',6X,'TOT.AVG.SAMP.SIZE =',F6.1/22X,
     *'POP.1=',F8.1/22X,'POP.2=',F8.1/6X,'COVERAGE',
     *' PROB. =',F6.3)
  450 FORMAT('+',64X,'-    ',A1)
  451 FORMAT('+',60X,A1,'    -')
      READ(5,100)II,JJ,RAMDA,PHI,C,DIN,ARAMDA,APHI,
     *      AC,ADIN
      WRITE(6,200)II,JJ,RAMDA,PHI,C,DIN
      WRITE(10)ARAMDA,APHI,AC,ADIN
      WRITE(6,351)
      CCRR=C*C*RAMDA*RAMDA
      CALL RSTART(II,JJ)
      INOBS=DIN
      SIMNO=1000.D0
      F=1.D0/RAMDA
      PHIS=1.D0-PHI
      DO 1000 IJK=1,1000
      SUMX=REXP(0)
      SUMY=REXP(0)
      DO 500 ILL=2,INOBS
      SUMX=SUMX+REXP(0)
      SUMY=SUMY+REXP(0)
  500 CONTINUE
      P=DIN
      Q=DIN
      PQ=P+Q
      UX=SUMX/P
      UY=SUMY/Q
   10 DX=1.D0/UX
   11 PY=PHI/UY
      SPCOND=CCRR*(DX+PY)*(DX+PY)
      IF(PQ.GE.SPCOND)GO TO 900
       IF(SUMY-PHI*SUMX)2,1,1
    1 SUMX=SUMX+REXP(0)
      P=P+1.D0
      PQ=PQ+1.D0
      UX=SUMX/P
      GOTO10
    2 SUMY=SUMY+REXP(0)
```

```
      Q=Q+1.D0
      PQ=PQ+1.D0
      UY=SUMY/Q
      GO TO 11
  900 CONTINUE
      TP=TP+P
      TQ=TQ+Q
      NPQ=PQ
      NP=P
      NQ=Q
      WRITE(6,251)IJK,NPQ,NP,NQ,UX,UY
      COMP=DX-PY-PHIS
      IF(DABS(COMP)-E)30,30,20
   20 IF(COMP)40,40,35
   35 WRITE(6,450)STAR
      COUNT=COUNT+1.D0
      GO TO 30
   40 WRITE(6,451)STAR
      COUNT=COUNT+1.D0
   30 AIJK=IJK
      AP=P
      AQ=Q
      ASUMX=SUMX
      ASUMY=SUMY
      WRITE(10)AIJK,AP,AQ,ASUMX,ASUMY
 1000 CONTINUE
      TPQ=TP+TQ
      TPQ=TPQ/SIMNO
      TP=TP/SIMNO
      TQ=TQ/SIMNO
      TC=1.-COUNT/SIMNO
      WRITE(6,400)TPQ,TP,TQ,TC
      STOP
      END
```

APPENDIX B

In this part, we whall illustrate how we can compute (2.4.5) given p and q. We write the equation again.

(1) $\int_0^\infty \Pr(\{y^{-1}+(1-\psi)-\Lambda^{-1}\}^{-1} \leq \bar{e}_1 \leq \{y^{-1}+(1-\psi)+\Lambda^{-1}\}^{-1})h(y;\, q,\, p)dy$

For convenience, let

$$\{ y^{-1} + (1 - \psi) - \Lambda^{-1} \}^{-1} = c_1 ,$$

$$\{ y^{-1} + (1 - \psi) + \Lambda^{-1} \}^{-1} = c_2 , \text{ and}$$

$$\bar{e}_1 = Z$$

Then (1) becomes

(2) $\Pr(c_1 \leq Z \leq c_2)\ h\ (y;\, q,\, q)dy$

where $Z \sim$ Gamma (p, p) and

$Y \sim$ Gamma (q, q).

Now, $\int_0^\infty \Pr(c_1 \leq Z \leq c_2) = \int_{c_1}^{c_2} \frac{p^p}{\Gamma(p)} Z^{p-1} e^{-pZ} dZ$

Integrating by terms, the above can be rewritten as

$$-p^p\, e^{-pZ} \sum_{k=1}^{p} \frac{Z^{p-k}}{(p-k)\,!\, p^k} \Bigg|_{c_1}^{c_2} .$$

Thus, (2) can be expressed by

$$\int_0^\infty \left| - p^p \, e^{-pZ} \sum_{k=1}^{p} \frac{Z^{p-k}}{(p-k)\,!\,p^k} \right|_{c_1}^{c_2} \left| \frac{q^q}{\Gamma(q)} y^{q-1} e^{-qy} \, dy \right. .$$

And the final form can be rewritten as

$$(3) \qquad \int_0^\infty \left| \frac{p^p \, q^q}{\Gamma(q)} \, e^{-pc_1-qy} \sum_{k=1}^{p} \frac{c_1^{\,p-k}}{(p-k)\,!\,p^k} - \right.$$

$$\left. \frac{p^p \, q^q}{\Gamma(q)} \, e^{-pc_2-qy} \sum_{k=1}^{p} \frac{c_2^{\,p-k}}{(p-k)\,!\,p^k} \right| dy$$

Since c_1 and c_2 are functions of Y only, we can use numerical methods for integration by using a computer.

The curvature of the function (3) turned out to be smooth enough to use the trapezoidal method. However, each term of (3) can be very large or very small.

Thus, we only can get an approximate value of (3).

By applying various values of p and q, we now can achieve (3) to be close to 0.95 and hence we get the approximate p and q.

```
FORTRAN IV G LEVEL  21                MAIN                DATE = 78176

           C                                                            *
           C COMPUTER PROGRAM TO COMPUTE (2.2.5) GIVEN P AND Q           *
           C                                                            *
                 IMPLICIT REAL*8(A-H,O-Z)
                 DIMENSION PF1(800),CF(800),PF2(800)
                 REAL*8 LABDA1,LABDA2,LAMDA1,LAMDA2
             100 FORMAT(5F5.0,I5)
             200 FORMAT('1 LAMBDA 1=',F5.1,'  LAMBDA 2=',F5.1,'  D'
                *,'=',F6.3//)
                 READ(5,100)LAMDA1,LAMDA2,D,P,Q,K
                 WRITE(6,200)LAMDA1,LAMDA2,D
                 DIFF=LAMDA1-LAMDA2
                 DO 1000 IJK=1,K
                 LABDA1=LAMDA1*P
                 LABDA2=LAMDA2*Q
                 TP=0.D0
                 PDQ=P/Q
                 PDIFF=P*DIFF
                 PD=P*D
                 CALL POISSN(LABDA1,PF1,CF,LL1,LU1)
                 IF(LABDA1.NE.LABDA2)GOTO 6
                 LL2=LL1
                 LU2=LU1
                 N=LU2-LL2+1
                 DO 25 I=1,N
              25 PF2(I)=PF1(I)
                 GOTO 4
               6 CONTINUE
                 CALL POISSN(LABDA2,PF2,PF1,LL2,LU2)
               4 CONTINUE
                 N=LU2-LL2+1
                 DWN=(LL2-1)*PDQ+PDIFF-PD
                 UP=DWN+PD+PD
                 DO 55 I=1,N
                 DWN=DWN+PDQ
                 UP=UP+PDQ
                 NU=UP+1.D-8
                 NL=DWN+1.D-8
                 KU=NU-LL1
                 KL=NL-LL1
                 IF(KU)55,55,2
               2 IF(KL)3,3,5
               3 IF(NU-LU1)7,9,9
               7 PR=CF(KU+1)
                 GO TO 20
               9 PR=1.D0
                 GO TO 20
               5 IF(NU-LU1)11,13,13
              11 PR=CF(KU+1)-CF(KL)
                 GO TO 20
              13 PR=1.D0-CF(KL)
              20 PP=PR*PF2(I)
                 TP=TP+PP
              55 CONTINUE
                 WRITE(6,300)P,Q,TP
             300 FORMAT('0     P=',F5.0,'  Q=',F5.0,'  PROB.=',F9.5)
                 P=P+1.D0
                 Q=P*DSQRT(LAMDA2/LAMDA1)
                 JJ=Q+.5D0
                 Q=JJ
            1000 CONTINUE
                 STOP
                 END
```

```
FORTRAN IV G LEVEL  21                  POISSN              DATE = 78176

            SUBROUTINE POISSN(LAMBDA,PF,CF,LL,LU)
      C                                                                    *
      C THIS SUBROUTINE COMPUTES THE PROBABILITY FUNCTION                   *
      C AND THE CUMMULATIVE DENSITY FUNCTION OF THE POISSON                 *
      C DISTRIBUTION WITH PARAMETER LAMBDA.                                *
      C LL AND LU ARE THE BOUNDARIES WHOSE P.F. VALUES ARE                  *
      C LESS THAN EXP(-25).                                                *
      C                                                                    *
            IMPLICIT REAL*8(A-H,O-Z)
            DIMENSION  PF(800),CF(800)
            REAL*8 LAMBDA
            MM=1
            LL=0
            J=3
            SUMLOG=0.D0
            DLG=DLOG(LAMBDA)
            IF(LAMBDA.LT.25)GOTO 10
            PF(1)=0.D0
            CF(1)=0.D0
            PF(2)=0.D0
            CF(2)=0.D0
            GO TO 20
         10 P=DEXP(-LAMBDA)
            DLG=DLOG(LAMBDA)
            PF(1)=P
            CF(1)=P
            P=P*LAMBDA
            PF(2)=P
            CF(2)=CF(1)+P
            MM=0
         20 I=2
         21 CONTINUE
            DI=DFLOAT(I)
            SUMLOG=SUMLOG+DLOG(DI)
            P=-LAMBDA+DI*DLG-SUMLOG
            IF(P.LT.-25.D0)GOTO 25
            P=DEXP(P)
            IF(MM.EQ.0)GOTO 26
            MM=0
            LL=I
            PF(1)=P
            CF(1)=P
            J=2
            GOTO 22
         25 CONTINUE
            P=0.D0
            IF(MM.EQ.1)GOTO 22
         26 CONTINUE
            PF(J)=P
            CF(J)=CF(J-1)+P
            IF( CF(J).GE. 0.9999999D0)GOTO 2
            J=J+1
            IF(J.GT.2000)GOTO33
         22 CONTINUE
            I=I+1
            GO TO 21
         33 CONTINUE
            WRITE(6,100) CF(800)
        100 FORMAT('- 800 INTERVALS ARE NOT ENOUGH.',F15.8)
            STOP
          2 CF(J)=1.D0
            LU=I
            RETURN
            END
```

```
C                                                                      *
C COMPUTER PROGRAM TO COMPUTE (2.4.5) GIVEN P AND Q                    *
C                                                                      *
      IMPLICIT REAL*8(A-H,O-Z)
      REAL*8 LAMBDA
      DIMENSION Y(1000)
  150 FORMAT(4F5.0,I5)
  200 FORMAT('1   PHI=',F5.2,'   LAMBDA=',F5.1//)
      COMP=40.D0
      SMALL=1.D-4
      Y(1)=0.D0
      READ(5,150)PHI,LAMBDA,P,Q,K
      WRITE(6,200)PHI,LAMBDA
      A=1.D0/LAMBDA
      C1=1.D0-PHI-A
      C2=C1+A+A
      B=PHI
      FINC=.01D0
      DO 1000 IIJJ=1,K
      SLNQ=0.D0
      SLNP=0.D0
      N=Q-.9D0
      DO 1 I=2,N
    1 SLNQ=SLNQ+DLOG(DFLOAT(I))
      N=P+.1D0
      DO 2 I=1,N
    2 SLNP=SLNP+DLOG(DFLOAT(I))
      M=N+1
      PLN=DLOG(P)
      QLN=DLOG(Q)
      PLNP=P*PLN
      QLNQ=Q*QLN
      E=PLNP+QLNQ-SLNQ
      FX=0.D0
      DO 100  IJ=2,1000
      FX=FX+FINC
      FXLN=DLOG(FX)
      C2B=C2+B/FX
      PC2B=P/C2B
      C2BLN=DLOG(C2B)
      QLNFX=E+(Q-1.D0)*FXLN-Q*FX
      D2=QLNFX-PC2B
      D=-P*DLOG(C2B)-SLNP
      SUM2=0.D0
      DO 50 IK=1,N
      G=D+DLOG(DFLOAT(M-IK))-PLN
      D=G+C2BLN
      D2D=D2+D
      IF(D2D+COMP)50,50,4
    4 SUM2=SUM2+DEXP(D2D)
   50 CONTINUE
      IF(SUM2 - Y(IJ-1))5,7,7
    5 IF(SUM2 - SMALL)15,15,17
   15 SUM2=0.D0
      Y(IJ)=SUM2
      INDX=IJ-1
      GO TO 6
    7 IF(SUM2.LE.SMALL) SUM2=0.D0
```

```
FORTRAN IV G LEVEL  21                    MAIN                    DATE = 78176

              17 Y(IJ)=SUM2
                 INDX=IJ-1
                 IF(SUM2)100,100,6
               6 C1B=C1+B/FX
                 IF(C1B)100,100,8
               8 PC1B=P/C1B
                 C1BLN=DLOG(C1B)
                 D1=QLNFX-PC1B
                 S=-P*DLOG(C1B)-SLNP
                 SUM1=0.D0
                 DO 75 IJK=1,N
                 T=S+DLOG(DFLOAT(M-IJK))-PLN
                 S=T+C1BLN
                 S2S=D1+S
                 IF(S2S+COMP)75,75,14
              14 SUM1=SUM1+DEXP(S2S)
              75 CONTINUE
                 IF(SUM1.LE.SMALL) SUM1=0.D0
                 Y(IJ) =SUM2-SUM1
             100 CONTINUE
                 IF(SUM2)25,25,20
              20 WRITE(6,111)SUM2,SUM1
             111 FORMAT('-*** 1000 INTERVALS ARE NOT ENOUGH. '
                $,'THE LAST VALUES ARE ',F15.10,' AND',F15.10)
                 GO TO 88
              25 CONTINUE
                 AREA=0.D0
                 DO 66 I=2,INDX
              66 AREA=AREA+Y(I)+Y(I)
                 AREA=AREA+Y(INDX+1)
                 AREA=AREA*FINC/2.D0
              88 CONTINUE
                 WRITE(6,300)P,Q,AREA
             300 FORMAT('0     P=',F5.0,'  Q=',F5.0,'   PROB.=',F9.5)
                 Q=Q+1.D0
                 IP=Q/PHI + .1D0
                 P=DFLOAT(IP)
            1000 CONTINUE
              99 CONTINUE
                 STOP
                 END
```

ACKNOWLEDGEMENT

I am sincerely grateful to Dr. Peter S. K. Perng, major professor for his suggestion of the topic and for his guidance. My sincere appreciation is due to Dr. Arthur D. Dayton, Professor and Head, Department of Statistics, for providing financial aid during my graduate study. Thanks are extended to Dr. Kenneth E. Kemp, associate professor, who read the manuscripts and gave me handful of advices.

I am indebted most to my wife, Inok, for her endless encouragement, patience and sacrifice which she had to endure.

SIMULATION RESULTS OF A SEQUENTIAL FIXED WIDTH CONFIDENCE INTERVAL FOR A FUNCTION OF PARAMETERS

by

CHANG SOO PAIK

B.S., Oakland University, 1976

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requitements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1979

ABSTRACT

Let $X_1, X_2, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_n$ be two independent random samples from the density functions $f(x; \underline{\theta})$ and $h(y; \underline{\phi})$, respectively, where $\underline{\theta}$ and $\underline{\phi}$ are unknown parameter vectors. Let $g(\underline{\gamma}) = g(\underline{\theta}, \underline{\phi})$ be a real valued function of the parameters.

A sequential procedure which was developed by Perng and Hasza to find a fixed width confidence interval of $g(\underline{\gamma})$ was investigated by Monte Carlo study.

Cases of $g(\underline{\gamma})$ were the variance ratio in normal distributions, the parameter difference in poisson distributions, the parameter ratio in exponential distributions and the parameter difference in exponential distributions. It was of interest to determine the coverage probabilities and the average total sample sizes due to the suggested procedure.