ESTIMATION OF STAGE DEVELOPMENT TIME WITH MORTALITY

by

Yeong-Ling Hwang

B.S., National Chung Hsing University, 1980

A MASTER REPORT

Submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1988

Approved by:

John E Boyer Jr

TABLE OF CONTENTS

I: INTRODUCTION:

Consider an insect that displays observable stages throughout its lifetime. A common problem is to estimate the average or typical time required for the insect to reach a particular stage (e.g., the mature or adult stage) of development. One approach to this problem is to observe a cohort of such organisms periodically over time and record the number of insects in the various stages and average the times obtained.

It is not always possible to observe a population over time since, in some cases, the insects can not be recaptured to ascertain the difference between previous and current observing time. In other cases the insects may have to be killed or their habitat may have to be destroyed to determine which stage they are in.

We thus take samples from the population of insects at a sequence of times. We let t_1, $t_2, \ldots, t_F$ denote the times at which we select a sample. The samples will be taken randomly and without replacement, although we will assume that the population is large enough that the provision for sampling without replacement will not affect calculations and the samples then will be assumed to be independent.

Boyer and Deaton (1984) and Pontius (1987) derived and evaluated nonparametric estimators for the average time of development across various stages, under the assumption of no mortality in the course of the experiment. Here we extend their results to the situation with only one stage, i.e., from non-maturity to maturity, but where the sampled insects might be either dead or alive. In this case, death could occur either before or after maturity. We take independent samples of fixed size and record the numbers of both mature and non-mature organisms, noting mortality in each group, at various observing times and estimate the average time to maturity, and its corresponding standard error. We compare the mean with its theoretical value and make recommendations for appropriate estimation procedures.

II: THE RELATIONSHIP OF OBSERVATIONS TO THEORETICAL QUANTITIES:

Consider an experiment in which we are interested in progression of an organism from one stage to another stage, i.e., from the stage we label " non-mature" to the stage we label "mature". In the performance of this experiment, we collect independent samples of insects of size n_i at time t_i, $i=1,2,\ldots\ldots,F$. The insects collected may be either dead or alive, mature or non-mature. We then allocate the elements of the ith sample to one of four categories which are non-mature and dead (ND), non-mature and alive (NA), mature and dead (MD), mature and alive (MA). In each category at time i, the observations have the binomial (n_i , p_{ij}) distribution, j= ND, NA, MD, MA, and the overall data have the multinomial (n_{ij} , p_{iND} , p_{iNA} , p_{iMD} , p_{iMA}) distribution.

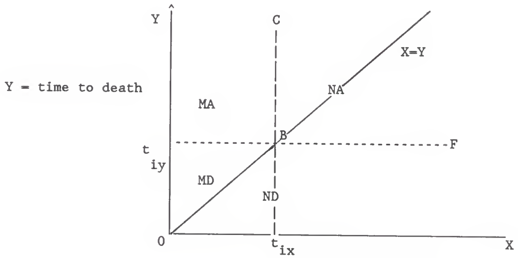
We let f(x,y) represent the bivariate density of X (time to maturity) and Y (time to death). We presume that f(x,y)>0 only if both x and y are positive. From the figure 2.1, we can see that at time t_i the probabilities for the categories are represented as outlined below.

ND= the area of FBOX , $\quad p_{iND}= \int_0^{t_{iy}} \int_y^{\infty} f(x,y)\,dx\,dy,$

NA= the area of CBF , $\quad p_{iNA}= \int_{t_{iy}}^{\infty} \int_{t_{ix}}^{\infty} f(x,y)\,dx\,dy,$

MD= the area of $Bt_{iy}0$, $\quad p_{iMD}= \int_0^{t_{iy}} \int_0^{y} f(x,y)\,dx\,dy,$

MA= the area of $Yt_{iy}BC$, $\quad p_{iMA} = \int_{t_{iy}}^{\infty} \int_{0}^{t_{ix}} f(x,y)\, dx\, dy.$



X = time to maturity

t_{ix}= the shorthand notation for

point $(t_i, 0)$

t_{iy}= the shorthand notation for

point $(0, t_i)$

t_i = time of ith sample

FIGURE 2.1

Note that n_i is the size of the sample taken at the ith sampling time. We let $p_i = p_{iMA} + p_{iMD} = \int_{0}^{t_{ix}} \int_{x}^{\infty} f(x,y)dy\, dx$

Consider an insect that displays observable stages throughout its lifetime. Then, again from figure 2.1, we let

$\hat{p}_i$ = the observed proportion of insects at time i which have reached the mature stage at time i

$$= \frac{\#MA_i + \#MD_i}{n_i} .$$

Here we let $\#MA_i$ be the number, from the sample of size n_i, found in the MA category, and $\#MD_i$ represents the corresponding number in the MD category at time t_i.

III: THE ESTIMATION PROCEDURE AND THE THEORETICAL CALCULATION:

a: Expected time to maturity

Let X represent time to maturity and let Y represent time to death. Since we are primarily interested in the length of time which it takes to reach maturity, we restrict our attention to those organisms that live long enough to do that. Hence the parameter that we shall estimate is the average time to maturity, among those that do reach maturity. Formally, the parameter of interest is $T = E[X|X<Y]$.

Then, one can show that

$$
\begin{aligned}
f(x|x<y) &= f(x|x-y<0) \\
&= \frac{\partial}{\partial X} P(X\le x|X-Y<0) \\
&= \frac{\partial}{\partial X} \frac{P(X\le x \text{ and } X-Y<0)}{P(X-Y<0)} \\
&= \frac{\partial}{\partial X} \frac{\int_{-\infty}^{x}\int_{s}^{\infty} f(s,t)\,dtds}{\int_{-\infty}^{\infty}\int_{s}^{\infty} f(s,t)\,dtds} \\
&= \frac{\int_{x}^{\infty} f(x,t)\,dt}{\int_{-\infty}^{\infty}\int_{s}^{\infty} f(s,t)\,dsdt} .
\end{aligned}
$$

Hence

$$
\begin{aligned}
T = E(X|X<Y) &= \int_{-\infty}^{\infty} xf(x|x<y)\,dx \\
&= \frac{\int_{0}^{\infty}\int_{x}^{\infty} xf(x,y)\,dydx}{\int_{0}^{\infty}\int_{x}^{\infty} f(x,y)\,dydx}
\end{aligned}
$$

equivalently, using change of order of integration,

$$= \frac{\int_0^\infty \int_0^y xf(x,y)dxdy}{\int_0^\infty \int_0^y f(x,y)dxdy} \qquad (3.a.1)$$

Note that we assume the density to be zero when either argument is negative, otherwise the lower limits in the expression above would all be $-\infty$.

b: An estimator of T

In the previous expression for T, let us denote the numerator by U and the denominator by V. Then note that

$$U = \int_0^\infty \int_0^y xf(x,y)dxdy = \int_0^\infty \int_x^\infty xf(x,y)dydx.$$

If the X-axis is broken into intervals $[0,t_1],[t_1,t_2],[t_2,t_3],\ldots..$ then clearly U is reasonably approximated (in a manner similar to Riemann Sums) by

$$\hat{U} = \hat{p}_1\left(\frac{t_1}{2}\right) + (\hat{p}_2-\hat{p}_1)\left(\frac{t_2+t_1}{2}\right) + (\hat{p}_3-\hat{p}_2)\left(\frac{t_3+t_2}{2}\right) +$$

$$\ldots\ldots. + (\hat{p}_F-\hat{p}_{F-1})\left(\frac{t_F+t_{F-1}}{2}\right).$$

(See Figure 3.1 below, and recognize that $\hat{p}_i$ is the fraction of the sample which lies above the line X = Y and to the left of t_{ix}).

In a similar manner V is just the fraction of the population which lies above the line X = Y, and is approximated by $\hat{p}_F$, the fraction of mature organisms found at the last sampling time. Of course, this presumes that sampling time F is sufficiently large that nearly all organisms that will mature have done so.

Thus we arrive at an estimator $\hat{T}$ by taking $\hat{T} = \hat{U} / \hat{V}$, this gives us

$$\hat{T} = \Big[\hat{p}_1\Big(\frac{t_1}{2}\Big) + (\hat{p}_2 - \hat{p}_1)\Big(\frac{t_2 + t_1}{2}\Big) + (\hat{p}_3 - \hat{p}_2)\Big(\frac{t_3 + t_2}{2}\Big) + \ldots + (\hat{p}_{F-} - \hat{p}_{F-2})\Big(\frac{t_{F-1} + t_{F-2}}{2}\Big) + (p_F - p_{F-1})\hat{p}_F \frac{t_F + t_{F-1}}{2}\Big)\Big] / \hat{p}_F \qquad (3.b.1)$$

and with a little simplification we get an equivalent form which is slightly better for some computations.

$$= \frac{-t_2}{2}\Big(\frac{\hat{p}_1}{\hat{p}_F}\Big) + \sum_{i=2}^{F-1}\Big(\frac{t_{i-1} - t_{i+1}}{2}\Big)\Big(\frac{\hat{p}_i}{\hat{p}_F}\Big) + \Big(\frac{t_F + t_{F-1}}{2}\Big).$$

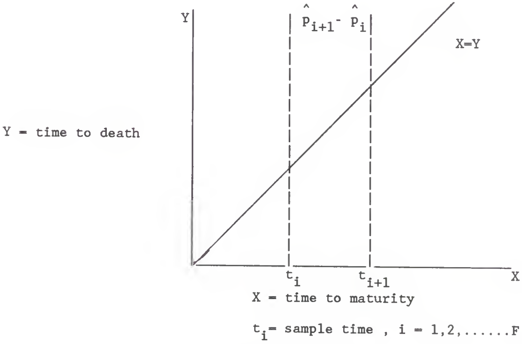


FIGURE 3.1

c: Approximate mean of $\hat{T}$

Using the approximate formula (Mood, Graybill and Boes, 1974), we find

$$E\left(\frac{A}{B}\right) \approx \frac{\mu_a}{\mu_b} - \frac{1}{\mu_b^2}\,\mathrm{Cov}(A,B) + \frac{\mu_a}{\mu_b^3}\,\mathrm{Var}(B).$$

Hence we note that after a little simplification (because $\mathrm{cov}(\hat{p}_i,\hat{p}_F) = 0$)

$$E\left(\frac{\hat{p}_i}{\hat{p}_F}\right) \approx \frac{p_i}{p_F}\left(\frac{n_F p_F + 1 - p_F}{n_F p_F}\right) \text{--->} \frac{p_i}{p_F} \qquad \text{as } n_F \text{--->} \infty .$$

Therefore $\frac{\hat{p}_i}{\hat{p}_F}$ is an approximately unbiased estimator of the corresponding ratio of population parameters as n_F---> ∞, and, since $\hat{T}$ is nearly a weighted average of such quantities we are led to believe that $\hat{T}$ will be unbiased or nearly unbiased as an estimator of T.

d: Approximate variance of $\hat{T}$

Again we use an approximation (Mood, Graybill, and Boes, 1974)

$$\mathrm{Var}(\hat{T}) = \mathrm{Var}\left(\frac{A}{B}\right)$$

$$\approx \frac{\mu_A^2}{\mu_B^2}\left[\frac{\mathrm{Var}(A)}{\mu_A^2} + \frac{\mathrm{Var}(B)}{\mu_B^2} - \frac{2\mathrm{Cov}(A,B)}{\mu_A \mu_B}\right] \quad (3.d.1).$$

In this case

$$A = \hat{U} = \frac{-t_2}{2}\hat{p}_1 + \sum_{i=2}^{F-2}\left(\frac{t_{i-1} - t_{i+1}}{2}\right)\hat{p}_i + \left(\frac{t_F + t_{F-1}}{2}\right)\hat{p}_F,$$

$$B = \hat{V} = \hat{p}_F.$$

Thus $$\mu_A^2 = \left[\left(\frac{-t_2}{2}\right)p_1 + \sum_{i=2}^{F-1} \left(\frac{t_{i-1} - t_{i+1}}{2}\right)p_i + \left(\frac{t_F + t_{F-1}}{2}\right)p_F \right]^2,$$

$$\mu_B^2 = p_F^2 ,$$

$$Var(A) = \left(\frac{-t_2}{2}\right)^2 \frac{p_1(1-p_1)}{n_1} + \sum_{i=2}^{F-1} \left(\frac{t_{i-1} - t_{i+1}}{2}\right)^2 \frac{p_i(1-p_i)}{n_i}$$

$$+ \left(\frac{t_F + t_{F-1}}{2}\right)^2 \frac{p_F(1-p_F)}{n_F},$$

$$Var(B) = \frac{p_F(1-p_F)}{n_F},$$

$$Cov(A,B) = \left(\frac{t_F + t_{F-1}}{2}\right) Var(\hat{p}_F)$$

$$= \left(\frac{t_F + t_{F-1}}{2}\right) \frac{p_F(1-p_F)}{n_F}.$$

In estimating the variance of $\hat{T}$, we will insert the corresponding sample quantities ($\hat{p}_i$'s) into each term and replace in (3.d.1).

e: The comparison to the estimator of Pontius

The $\hat{p}_i$ we have defined represents the proportion of organisms sampled at time i that have reached the mature stage, so that $1-\hat{p}_i$ (i.e., $\hat{q}_i$) represents the proportion that have not yet reached the mature stage.

Since $$\hat{T} = \frac{\left[\frac{t_1}{2}\hat{p}_1 + \sum_{i=2}^{F}(\hat{p}_i - \hat{p}_{i-1})\left(\frac{t_i + t_{i-1}}{2}\right)\right]}{\hat{p}_F}$$

and $\hat{p}_i = 1-\hat{q}_i$.

Then

$$\hat{T} = \frac{\left[\frac{t_1}{2} + \sum_{i=2}^{F-1}\hat{q}_i\left(\frac{t_{i+1} - t_{i-1}}{2}\right) - \hat{q}_F\left(\frac{t_F + t_{F-1}}{2}\right)\right]}{1-\hat{q}_F}.$$

Under the assumptions made by Pontius, t_0 must be zero and $\hat{p}_{F,S}$ (our $\hat{q}_F$) must be zero. Using this condition the expression for $\hat{T}$ reduces to

$$\hat{T} = \frac{t_1}{2} + \sum_{i=2}^{F-1}\hat{q}_i\left(\frac{t_{i+1} - t_{i-1}}{2}\right).$$

This is exactly the same as the estimator derived by Pontius (1987) for the case in which it is assumed that no insects have died. Hence, although we have extended the method for computations to populations which may contain dead insects, the final result is still the same as the result derived by Pontius if no mortality is observed during the experiment.

IV: NUMERICAL INTEGRATION AND THE SCOPE OF THE SIMULATION STUDY:

a: The numerical integration

Since the quantity $T = E(X|X<Y)$ (which is the parameter of interest in this problem) is not readily derived for the bivariate normal distribution, we used the IMSL subroutine DBLIN to numerically perform the two double integrations required (expression 3.a.1). A copy of the numerical integration program is in the appendix.

b: The scope of the simulation study

We developed a computer program to simulate the experiment described above. As an underlying distribution we used the bivariate normal distribution, which has a large number of shapes depending on one's choices of parameters. Since we wanted to investigate the lifetimes of insects which mature before death, and, since in most populations, lifetime is much longer than maturation time, we chose mean of "time to death" (i.e. μ_y) to be larger than mean of "time to maturity" (i.e. μ_x). For normal distributions, a distance of three standard deviations from the mean in both directions will cover most of the observations. Hence we arbitrarily chose μ_x and μ_y, and then picked standard deviations of "time to death" and "time to maturity" (i.e. σ_y, σ_x, resp.) to be values such that $\mu_x \pm 3\sigma_x$ and/or $\mu_y - 3\sigma_y$ were in between the first and the last observation time. (We were willing to allow some of the insects to die after our experiment was over; that's why we are not concerned with $\mu_y+3\sigma_y$ being inside our observation time.) We let μ_x be fixed at 6 and σ_x be fixed at 2. We then let $\mu_y= 9$ with $\sigma_y= 2$ and $\sigma_y= 3$, $\mu_y= 12$ with $\sigma_y= 3$ and $\sigma_y= 4$ and

$\mu_y = 15$ with $\sigma_y = 4$ and $\sigma_y = 5$. We also allowed the correlation coefficient ρ to be 0, 0.25 and 0.5. We considered samples of size 15 and 30 at each time period and we had sampling regimens which had 10 sampling times equally spaced (every two time units from time t = 2 to time t = 20) and unequally spaced (also from t = 2 to t = 20). For each of the 72 parameter configurations, 100 repetitions of the experiment were run, and computations were made.

For each repetition, $\hat{T}$ and its estimated standard error, which will be denoted by $std(\hat{T})$, were calculated. Additionally the intervals $\hat{T} \pm 2std(\hat{T})$ and $\hat{T} \pm 3std(\hat{T})$ were computed.

For each parameter configuration, several summary statistics over the 100 repetitions were retained and examined. They include:

(1) Average value of $\hat{T}$.

(2) Number of intervals of the form $\hat{T} \pm 2std(\hat{T})$ that contained T.

(3) Number of intervals of the form $\hat{T} \pm 3std(\hat{T})$ that contained T.

(4) Sample standard deviation of $\hat{T}$.

(5) The interval for the mean of $\hat{T}$, given by average value of $\hat{T} \pm 2$ * sample standard deviation of $\hat{T}$.

(6) Average value of estimated standard error of $\hat{T}$.

A copy of the simulation program is found in the appendix.

c: Two possible computations:

The population p_i in formula (3.b.1) should be monotonically increasing as a function of i, but when we dealt with the sample proportions, we occasionally found $\hat{p}_i$ to be smaller than $\hat{p}_{i-1}$. We used both the negative value and zero respectively for the quantity $\hat{p}_i - \hat{p}_{i-1}$ in our simulation study when $\hat{p}_i$ was smaller than $\hat{p}_{i-1}$.

V: CONCLUSIONS:

We discovered early that the quantity $\hat{p}_i - \hat{p}_{i-1}$ would be negative sometimes, even though the corresponding population quantity is not. We considered replacing these negative values of $\hat{p}_i - \hat{p}_{i-1}$ by zero in the computations. In fact in early stages of the simulation, we did computations both ways. After comparing the results, we recommend the use of the negative value instead of using zero when this occurs. Our recommendation is based on the comparison of the simulation results to the numerical integration results, i.e., the average $\hat{T}$ from 100 simulations with $\hat{p}_i - \hat{p}_{i-1}$ allowed to be negative is considerably closer to the numerical intergration than $\hat{p}_i - \hat{p}_{i-1}$ set to zero. That is, the average $\hat{T}$ from 100 simulations tends to be biased with $\hat{p}_i - \hat{p}_{i-1}$ set to zero.

Using the computer output of the simulation program, we have obtained an estimated value of $\hat{T}$ for each different configuration of μ_x, μ_y, σ_x, σ_y, ρ. We have calculated $\bar{\hat{T}}$ as the average of the estimated $\hat{T}$'s and formed an interval of the form $\bar{\hat{T}} \pm 2\hat{\sigma}_{\bar{\hat{T}}}$. Since we have 100 independent $\hat{T}$'s from 100 independent repetitions, by the Central Limit Theorem $\bar{\hat{T}}$ will be apporximately normally distributed and therefore, the quantity $\bar{\hat{T}} \pm 2\hat{\sigma}_{\bar{\hat{T}}}$ is approximately a 95% confidence interval for $E(\hat{T})$. We found that for (1) 29 out of 36 parameter configurations with sample size 15, (2) 26 out of 36 with sample size 30, (3) 27 out of 36 with equally spaced sampling time, (4) 28 out of 36 with unequally spaced sampling times,the confidence intervals contained the parameter of interest (T or $E[X|X<Y]$).

Note that the standard error of $\hat{T}$ is estimated in two ways in this study. The first is from looking at 100 independent repetitions of the experiment and estimating the standard deviation of the population of $\hat{T}$'s. This quantity is used in the confidence interval calculations described above. The second method is from the series expansion for an individual repetition, and uses formula (3.d.1). We have averaged these values and compared them with the lengths of the confidence intervals described above. (Note that the length of the confidence interval should be $4 \times \hat{\sigma}_{\hat{T}} / \sqrt{100}$ or approximately 0.4 times the standard error of $\hat{T}$). We find good agreement with this criterion in our tables, indicating that the short series expansion gives a good estimate of the actual standard error.

Note also that the standard errors for samples of size 15 are approximately 1.4 (or $\sqrt{2}$) times the corresponding standard errors for samples of size 30. This is an indication to us that even though there is no such term directly found in the Taylor series expansion which gave us (3.d.1), (at least for the situation where all the samples are of the same fixed sample size n), the standard error of the estimator may be proportional to $1/\sqrt{n}$, as are standard errors of means.

Another study that we performed is summarized in table 2. There we used the standard errors for each repetition, using the formula (3.d.1) and denoted here by "std($\hat{T}$)" , and formed the intervals $\hat{T}\pm2$std($\hat{T}$) and $\hat{T}\pm3$std($\hat{T}$). Then we determined how many of the 100 repetitions generated confidence intervals which contained the parameter values. Although the results do not show perfect agreement with the nominal coverage proportions, they always show around 80 to 95 percent of confidence intervals which contained the parameter values.

As a nonparametric estimator, it was not possible to show that $\hat{T}$ is a Uniform Minimum Variance Unbiased Estimator (UMVUE), but we did demonstrate that it is nearly unbiased for the parameters and distribution used in this study. We also showed that confidence interval coverage is near the nominal levels.

From the simulations we have, we did not find any significant difference in results between 10 equally spaced and unequally spaced observation times. It seems to be a convenience for an experimenter not to have to follow a strict schdule to observe and collect data.

The preceding work is not a complete study of the proposed estimator. Future work could fill several gaps. Foremost among these might be a study using alternative underlying bivariate distributions. (The bivariate normal was used in this case because of ready access to it, the ease of computations, and the versatility in terms of the number of shapes availables.) Other aspects which should be investigated include a more thorough study of the effect of sample size, a study of the effects of unequal sample sizes (at the various observation times), and a more thorough study of the lengths of the intervals between sampling times (how big should they be, and do they need to be equal ?).

REFERENCES

Boyer, J. E., Jr. and M. L. Deaton. 1984. Estimation of duration from stage frequency data. Tech. Rep., Dept. of Statistics, Kansas State University.

Mood, A. M., F. A. Graybill and D. C. Boes. 1974. Introduction to the theory of statistics, 3rd ed. McGraw-Hill, Inc., New York.

Pontius, J. S. 1987. Nonparametric estimation of stage transition time from stage frequency data. M. S. report. Kansas State University, Manhattan, Kansas.

The IMSL LIBRARY FORTRAN Subroutines in the area of Mathematics and Statistics, edition 7, International Mathematical & Statistical Libraries, Inc., Houston, Texas.

APPENDIX

```
      REAL*8 DSEED,UX,UY,VX,VY,PXY,A,B,C,D,DEXP,DSQRT,FD,PP,X,YCD,CN
     *,AREA,FN,AY,F,BY
      DIMENSION SIGMA(3),XYMA(200,2),WKVEC(2),KKVEC(2),T(30),MD(30)
     *          ,MA(30),ND(30),NA(30),P(30),F(30),G(30),G0(30),ET(100)
     *          ,ET0(100)
       COMMON   UX, UY, VX, VY,PXY
       INTEGER   IER
       EXTERNAL FN, FD, AY, BY
C
C      ***************************************
C      *                                     *
C      *    T(i) = SAMPLE TIME AT TIME i     *
C      *       i = 1,2,........10            *
C      *    AX, BX ARE LOWER AND UPPER LIMIT *
C      *      ON X                           *
C      *                                     *
C      ***************************************
       AX=0.D0
       BX=30.D0
       AERR=0.000001
C
C     *******************************************
C     *                                         *
C     *      UNEQUALLY SPACED OBSERVATION TIME  *
C     *         ALTERNATIVELY                   *
C     *       SET T(i) = 2 TO 20 BY 2 IS        *
C     *         EQUALLY SPACED OBSERVATION TIME *
C     *                                         *
C     *******************************************
C
      T(1)=2.D0
      T(2)=3.D0
      T(3)=6.D0
      T(4)=9.D0
      T(5)=10.D0
      T(6)=11.D0
      T(7)=12.D0
      T(8)=15.D0
      T(9)=16.D0
      T(10)=20.D0
       DSEED     = 466364003.D0
       NR        = 30
       K         = 2
       IR        = 200
```

```
      KKVEC (1) =1.D0
      WKVEC (1) = 0.D0
C
C     ************************************
C     *                                  *
C     *   READ 18 PARAMETER CONFIGURATION  *
C     *                                  *
C     ************************************
C
  501 READ(5,105,END=500) UX,UY,VX,VY,PXY
  105 FORMAT(5F5.2)
C
C     ***********************************************
C     *                                             *
C     *     SET THE FOLLOWING VARIABLES TO BE ZERO  *
C     *     NSAM = # OF SAMPLE SIZE.                *
C     *     SUM  = SUM OF T.                        *
C     *     SUMV = SUM OF STD(T).                   *
C     *     ETSQ = SUM OF SQUARE OF T.              *
C     *     V    = CROSS PRODUCT OF T IN ORDER TO   *
C     *            COMPUTE THE VARIANCE OF T.       *
C     *     ETMIN = MIN OF T.                       *
C     *     ETMAX = MAX OF T.                       *
C     *     COUNT2 = COUNT THE # OF 2STD(T) INTERVALS *
C     *              THAT CONTAIN THE PARAMETER OF  *
C     *              INTEREST.                      *
C     *     COUNT3 = COUNT THE # OF 3STD(T) INTERVALS *
C     *              THAT CONTAIN THE PARAMETER OF  *
C     *              INTEREST.                      *
C     *     VV     = CROSS PRODUCT OF STD(T) IN ORDER *
C     *              TO COMPUTE THE VARIANCE OF STE(T)*
C     *   THE ABOVE VARIABLES ARE COMPUTED WITH THE *
C     *     QUANTITY P_i - P_i-1 SET TO NEGATIVE , IN- *
C     *     STEAD OF THAT SUMO, SUMOV, ETSQO, ETOMIN *
C     *     ETOMAX, COUNTO2, COUNTO3, VVO, VO ARE    *
C     *     COMPUTED WITH THE QUANTITY P_i - P_i-1 SET *
C     *     TO ZERO.                                *
C     *                                             *
C     ***********************************************
C
```

```
      NSAM=0
      SUM=0
      SUM0=0
      SUMV=0
      SUM0V=0
      ETSQ=0
      ETSQ0=0
      V=0
      V0=0
      ETMIN=100.D0
      ETMAX=0.D0
      ET0MIN=100.D0
      ET0MAX=0.D0
      COUNT2=0.D0
      COUNT3=0.D0
      COUN02=0.D0
      COUN03=0.D0
      SUMV=0.D0
      VV=0.D0
      VV0=0.D0
      CALL ERRSET(208, 256, -1, 1)
C
C     *****************************************
C     *                                       *
C     *    DOUBLE INTEGRATION BY USING DBLIN  *
C     *      SUBROUTINE.                      *
C     *    CN = NUMERATOR.                    *
C     *    CD = DENOMINATOR.                  *
C     *    AREA = E(X|X<Y).                   *
C     *                                       *
C     *****************************************
C
      CN = DBLIN(FN, AX, BX, AY, BY, AERR, ERROR, IER)
      CD = DBLIN(FD, AX, BX, AY, BY, AERR, ERROR, IER)
      AREA = CN/CD
      WRITE(6,121)
  121 FORMAT('-',14X,'AY',9X,'BY',10X,'AERR',8X,'IER',11X,'CD',11X,'CN'
     * ,10X,'AREA')
      WRITE(6,122) AX, BX, AERR, IER, CD, CN, AREA
  122 FORMAT('-',10X,7(F10.6,2X))
C
C     *****************************************
C     *                                       *
C     *      100 REPETITIONS.                 *
C     *      SET TT AND VARA = 0.             *
C     *                                       *
C     *****************************************
C
```

```
      DO 6 KK=1,100
      TT=0
      TT0=0
      VARA=0.D0
C     *********************************************
C     *                                           *
C     *   GERERATE BIVARIATE NORMAL RANDOM NUMBER *
C     *   NR = # OF OBSERVATIONS                  *
C     *   K  = # OF RADAM VARIABLE                *
C     *   SIGMA = VARIANCE-COVARIANCE MATRIX      *
C     *             (UPPER TRIANGAL)              *
C     *   IR    = # OF ROW DIMENSION              *
C     *   XYMA = OUTPUT NR BY K MATRIX,           *
C     *   IER   = ERROR PARAMETER (OUTPUT)        *
C     *                                           *
C     *********************************************

      SIGMA(1)=VX
      SIGMA(2)=PXY
      SIGMA(3)=VY
      IF (KK.GT. 1) GO TO 103
      CALL GGNSM (DSEED, NR, K, SIGMA, IR, XYMA, WKVEC, IER)
C
C     *****************************************
C     *                                       *
C     *     DETERMINE THE GERERATED DATA THAT *
C     *      BELONG TO WHICH OF THE FOUR      *
C     *      CATEGORIES.                      *
C     *     NA = NON-MATURE AND ALIVE.        *
C     *     ND = NON-MATURE AND DEAD.         *
C     *     MA = MATURE AND ALIVE.            *
C     *     MD = MATURE AND DEAD.             *
C     *                                       *
C     *****************************************
C
      NA(1)=0.D0
      ND(1)=0.D0
      MA(1)=0.D0
      MD(1)=0.D0
      DO 1 I=1,15
      X = UX+ XYMA (I,1)
      Y = UY+ XYMA (I,2)
      IF (X.LE.T(1) .AND. Y.LE.T(1)) GO TO 5
      IF (X.LE.T(1) .AND. Y.GT.T(1)) GO TO 2
      IF (X.GT.T(1) .AND. Y.LE.T(1)) GO TO 3
      IF (X.GT.T(1) .AND. Y.GT.T(1)) GO TO 4
 5    IF (X.GT.Y) GO TO 3
      MD(1)=MD(1)+1
```

```
      GO TO 1
  2   MA(1)=MA(1)+1
      GO TO 1
  3   ND(1)=ND(1)+1
      GO TO 1
  4   NA(1)=NA(1)+1
   1  CONTINUE
      K=1
C
C     ****************************************
C     *                                      *
C     *    CALCULATE P , F , G  AND          *
C     *               i   i   i              *
C     *              F                       *
C     *      TT = Σ  F * G                   *
C     *          i=1  i   i                  *
C     *      ^      ^                        *
C     *     (T = TT/P  LABELED BY ET(KK))    *
C     *              F                       *
C     *                                      *
C     ****************************************
C
      P(K)=(MA(K)+MD(K))/15.D0
      IF ((K-1).EQ.0) T(K-1)=0.D0
      F(K)=(T(K)+T(K-1))/2.D0
      IF((K-1).EQ.0) P(K-1)=0.D0
      G(K)=P(K)-P(K-1)
      IF(G(K).LT.0) THEN
      G0(K)=0.D0
      ELSE
      G0(K)=G(K)
      END IF
      TT=TT+F(K)*G(K)
      TT0=TT0+F(K)*G0(K)
      GO TO 109
 103  DO 7 LL=1,10
      L=LL
      GO TO 108
109   DO 11 L=2,10
108   CALL GGNSM (DSEED, NR, K, SIGMA, IR, XYMA, KKVEC, IER)
      NA(L)=0.D0
      ND(L)=0.D0
      MA(L)=0.D0
      MD(L)=0.D0
      DO 12 I=1,15
      X = UX + XYMA (I,1)
      Y = UY + XYMA (I,2)
      IF (X.LE.T(L) .AND. Y.LE.T(L)) GO TO 15
      IF (X.LE.T(L) .AND. Y.GT.T(L)) GO TO 25
```

```
      IF (X.GT.T(L) .AND. Y.LE.T(L)) GO TO 35
      IF (X.GT.T(L) .AND. Y.GT.T(L)) GO TO 45
 15   IF (X.GT.Y) GO TO 35
      MD(L)=MD(L)+1
      GO TO 12
 25   MA(L)=MA(L)+1
      GO TO 12
 35   ND(L)=ND(L)+1
      GO TO 12
 45   NA(L)=NA(L)+1
 12   CONTINUE
      P(L)=(MA(L)+MD(L))/15.D0
      IF ((L-1).EQ.0) T(L-1)=0.D0
      F(L)=(T(L)+T(L-1))/2.D0
      IF ( L.EQ.10 ) GO TO 16
      VARA=VARA+(((T(L-1)-T(L+1))/2)**2)*P(L)*(1-P(L))/15.D0
      GO TO 17
C
C     ****************************************
C     *                                      *
C     *    CALCULATE VARA, VARB, COVAB, UA, UB *
C     *     BY INSERTING SAMPLE QUANTITIES IN  *
C     *                   ^                    *
C     *     ORDER TO CALCULATE VAR(T),         *
C     *     LABELED BY VARET.                  *
C     *                                      *
C     ****************************************
C
 16   VARA=VARA+(((T(L)+T(L-1))/2)**2)*P(L)*(1-P(L))/15.D0
      VARB=P(10)*(1-P(10))/15.D0
      COVAB=((T(L)+T(L-1))/2)*VARB
 17   IF((L-1).EQ.0) P(L-1)=0.D0
      G(L)=P(L)-P(L-1)
      IF ( G(L).LT.0) THEN
      G0(L)=0.D0
      ELSE
      G0(L)=G(L)
      END IF
      TT=TT+F(L)*G(L)
      TT0=TT0+F(L)*G0(L)
      IF (LL .GE. 1) GO TO 7
 11   CONTINUE
      UA=(TT)**2
      U0A=(TT0)**2
      UB=(P(10))**2
      GO TO 110
  7   CONTINUE
      UA=(TT)**2
```

```
      U0A=(TT0)**2
      UB=(P(10))**2
      LL=0
  110 NSAM=NSAM+1
      ET(KK) =TT/P(10)
      ET0(KK) =TT0/P(10)
      VARET=(UA/UB)*((VARA/UA)+(VARB/UB)-(COVAB/(UA*UB)))
      VAR0ET=(U0A/UB)*((VARA/U0A)+(VARB/UB)-(COVAB/(U0A*UB)))
      STDET=SQRT(VARET)
      STD0ET=SQRT(VAR0ET)
      SUMV=SUMV+STDET
      SUM0V=SUM0V+STD0ET
C
C     ********************************************
C     *                                          *
C     *                   ^                      *
C     *     CALCULATE 2 & 3STD(T) INTERVALS      *
C     *      WITH LABEL DL AND DU AS LOWER       *
C     *       AND UPPER LIMIT.                   *
C     *                                          *
C     ********************************************
C
      DL2=ET(KK)-2*STDET
      DU2=ET(KK)+2*STDET
      DL3=ET(KK)-3*STDET
      DU3=ET(KK)+3*STDET
      D0L2=ET0(KK)-2*STD0ET
      D0U2=ET0(KK)+2*STD0ET
      D0L3=ET0(KK)-3*STD0ET
      D0U3=ET0(KK)+3*STD0ET
      IF ( AREA.LE.DU2.AND.AREA.GE.DL2) THEN
      COUNT2=COUNT2+1
      ELSE
      COUNT2=COUNT2
      END IF
      IF ( AREA.LE.DU3.AND.AREA.GE.DL3) THEN
      COUNT3=COUNT3+1
      ELSE
      COUNT3=COUNT3
      END IF
      IF ( AREA.LE.D0U2.AND.AREA.GE.D0L2) THEN
      COUN02=COUN02+1
      ELSE
      COUN02=COUN02
      END IF
      IF ( AREA.LE.D0U3.AND.AREA.GE.D0L3) THEN
      COUN03=COUN03+1
      ELSE
```

```
      COUN03=COUN03
      END IF
      SUM=SUM+ET(KK)
      SUMO=SUMO+ETO(KK)
      ETSQ=ETSQ+ET(KK)*ET(KK)
      ETSQO=ETSQO+ETO(KK)*ETO(KK)
      IF (ET(KK).LT.ETMIN) THEN
      ETMIN=ET(KK)
      ELSE
      ETMIN=ETMIN
      END IF
      IF (ET(KK).GT.ETMAX) THEN
      ETMAX=ET(KK)
      ELSE
      ETMAX=ETMAX
      END IF
      IF (ETO(KK).LT.ETOMIN) THEN
      ETOMIN=ETO(KK)
      ELSE
      ETOMIN=ETOMIN
      END IF
      IF (ETO(KK).GT.ETOMAX) THEN
      ETOMAX=ETO(KK)
      ELSE
      ETOMAX=ETOMAX
      END IF
      IF (NSAM .EQ. 1) GO TO 119
      V=V+(NSAM*ET(KK)-SUM)**2/(NSAM*NSAM-NSAM)
      VO=VO+(NSAM*ETO(KK)-SUMO)**2/(NSAM*NSAM-NSAM)
      VV=VV+((NSAM*STDET-SUMV)**2)/(NSAM*(NSAM-1))
      VVO=VVO+((NSAM*STDOET-SUMOV)**2)/(NSAM*(NSAM-1))
      GO TO 120
119   V=0.D0
      VO=0.D0
      VV=0.D0
      VVO=0.D0
120   IF (KK.LT.100) GO TO 6
      PER2=COUNT2/100
      PER3=COUNT3/100
      PERO2=COUN02/100
      PERO3=COUN03/100
      XBAR=SUM/100
      XOBAR=SUMO/100
      STBAR=SUMV/100
      STOBAR=SUMOV/100
      VAR=V/99
      VARO=VO/99
      VARST=VV/99
```

```
      VAR0ST=VV0/99
      STD=SQRT(VAR/100)
      STD0=SQRT(VAR0/100)
      STDST=SQRT(VARST)
      STD0ST=SQRT(VAR0ST)
C
C     ******************************************
C     *                                        *
C     *     CALCULATE THE UPPER AND LOWER LIMIT *
C     *                   ^  ^                  *
C     *      OF THE INTERVAL FORMED T±2σ^ .     *
C     *                              T          *
C     *                                        *
C     ******************************************
C
      UPPER=XBAR+2*STD
      DLOW=XBAR-2*STD
      UPPER0=X0BAR+2*STD0
      DLOW0=X0BAR-2*STD0
      WRITE(6,107)
  107 FORMAT('1',10X,'UX',10X,'UY',10X,'VX',10X,'VY',10X,'PXY')
C
C     ****************************************
C     *                                      *
C     *      WRITE THE PARAMETER CONFIGURATIONS*
C     *       AND THE SUMMARY STATISTICS.    *
C     *                                      *
C     ****************************************
C
      WRITE(6,106) UX,UY,VX,VY,PXY
  106 FORMAT('-',9X,5(F5.2,7X))
      WRITE(6,100)
  100 FORMAT('-',10X,'NR', 5X,'K',2X,'SIGMA',18X,'IR')
      WRITE(6,101) NR, K,(SIGMA(J),J=1,3), IR
  101 FORMAT('-', 8X, I4, 2X, I4, 2X, 3(F5.3,2X), I4/)
      WRITE(6,115)
  115 FORMAT('-',5X,'THE FOLLOWING VALUE IS WHEN THE PROBABILITY IS NEG
     *ATIVE:')
      WRITE(6,104) TT, ET(KK), SUM,XBAR,ETSQ,VAR,STD,ETMIN,ETMAX,DLOW,
     * UPPER,COUNT2,PER2,COUNT3,PER3,STBAR,VARST,STDST
  104 FORMAT('-',10X,'TT=',F10.5,10X,'EXPECTATION OF TIME TO THE MATURI
     *TY =', F10.5/10X,'SUM=',F10.5,10X,'XBAR=',F10.5 /10X,'SUM OF SQUAR
     *E =',F10.5,/10X,'VARIANCE =',F10.5,10X,'STD =',F10.5/10X,
     *'MIN OF ET =',F10.5,10X,'MAX OF ET =',F10.5/10X,'LOWER LIMIT =',
     *F10.5,10X,'UPPER LIMIT =',F10.5/10X,'# OF 2STD CONTAIN TRUE VALUE
     *=',F10.5,10X,'PERCENT OF 2STD CONTAIN TRUE VALUE =',F10.5/10X,
     *'# OF 3STD CONTAIN TRUE VALUE =',F10.5,10X,'PERCENT OF 3STD CONTAI
     *N TRUE VALUE =',F10.5/10X,'MEAN OF STD OF EXPECTED TIME TO MATURIT
```

```
     *Y =',F10.5/10X,'VARIANCE OF STD OF EXPECTED TIME TO MATURITY =',
     *F10.5,10X,'STD OF STD OF EXPECTED TIME TO MATURITY =',F10.5)
      WRITE(6,116)
  116 FORMAT('-',5X,'THE FOLLOWING VALUE IS WHEN THE PROBABILITY IS ZER
     *O: ')
      WRITE(6,117) TTO, ETO(KK), SUMO,XOBAR,ETSQO,VARO,STDO,ETOMIN,
     * ETOMAX,DLOWO,UPPERO,COUNO2,PERO2,COUNO3,PERO3,STOBAR,VAROST
     *,STDOST
  117 FORMAT('-',10X,'TT=',F10.5,10X,'EXPECTATION OF TIME TO THE MATURI
     *TY =', F10.5/10X,'SUM=',F10.5,10X,'XBAR=',F10.5 /10X,'SUM OF SQUAR
     *E =',F10.5,/10X,'VARIANCE =',F10.5,10X,'STD =',F10.5/10X,
     *'MIN OF ET =',F10.5,10X,'MAX OF ET =',F10.5/10X,'LOWER LIMIT =',
     *F10.5,10X,'UPPER LIMIT =',F10.5/10X,'# OF 2STD CONTAIN TRUE VALUE
     *= ',F10.5,10X,'PERCENT OF 2STD CONTAIN TRUE VALUE =',F10.5/10X
     *,'# OF 3STD CONTAIN TRUE VALUE =',F10.5,10X,'PERCENT OF 3STD CONTA
     *IN TRUE VALUE =',F10.5/10X,'MEAN OF STD OF EXPCTED TIME TO MATURI
     *TY =',F10.5/10X,'VARIANCE OF STD OF EXPECTED TIME TO MATURITY ='
     *,F10.5,10X,'STD OF STD OF EXPECTED TIME TO MATURITY =',F10.5)
    6 CONTINUE
      WKVEC(1)=1.D0
      GO TO 501
  500 STOP
      END
C
C          ***********************************
C          *                                 *
C          *        DEFINE FN FUNCTION       *
C          *                                 *
C          ***********************************
C
          FUNCTION FN(X,Y)
          IMPLICIT REAL*8 (A-H,O-Z)
          COMMON   UX, UY, VX, VY, PXY
          PP=PXY*PXY
          A=DSQRT(1-PP)
          B=(X-UX)/VX
          C=(Y-UY)/VY
          D = (B*B) - (2*PXY*B*C) + (C*C)
          FN = X * (1/(2*3.14159*VX*VY*A))*DEXP((-1/(2*(1-PP)))*D)
          RETURN
          END
C
C            ***************************
C            *                         *
C            *     DEFINE FD FUNCTION  *
C            *                         *
C            ***************************
C
```

```
      FUNCTION FD(X,Y)
      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON  UX, UY, VX, VY, PXY
      PP=PXY*PXY
      A=DSQRT(1-PP)
      B=(X-UX)/VX
      C=(Y-UY)/VY
      D = (B*B) - (2*PXY*B*C) + (C*C)
      FD = (1/(2*3.14159*VX*VY*A))*DEXP((-1/(2*(1-PP)))*D)
      RETURN
      END
C
C     ************************************
C     *                                  *
C     *       DEFINE LOWER LIMIT         *
C     *            ON Y                  *
C     *                                  *
C     ************************************
C
      FUNCTION AY(X)
      IMPLICIT REAL*8 (A-H,O-Z)
      AY=X
      RETURN
      END
C
C     ********************************
C     *                              *
C     *     DEFINE UPPER LIMITS      *
C     *         ON Y                 *
C     *                              *
C     ********************************
C
      FUNCTION BY(X)
      IMPLICIT REAL*8  (A-H,O-Z)
      BY=30.D0
      RETURN
      END
/*
//LKED.SYSLIB    DD
//               DD
//               DD
//               DD DSN=SYS1.IMSL.SPFLIB,DISP=SHR
//GO.SYSIN   DD *
  6    9    2    2    0
  6    9    2    2     25
  6    9    2    2     50
  6    9    2    3    0
  6    9    2    3     25
```

6	9	2	3	50
6	12	2	3	0
6	12	2	3	25
6	12	2	3	50
6	12	2	4	0
6	12	2	4	25
6	12	2	4	50
6	15	2	4	0
6	15	2	4	25
6	15	2	4	50
6	15	2	5	0
6	15	2	5	25
6	15	2	5	50

TABLE 1

simulate 100 times

$\mu_x = 6, \quad \sigma_x = 2$

$\rho = 0.0$

μ_y	σ_y	n =	equally spaced 15	equally spaced 30	unequally spaced 15	unequally spaced 30
9	2	true value		5.63		
		estimated $\hat{T}$	5.74	5.70	5.77	5.73
		confidence interval of $\hat{T}$	(5.50, 5.98)	(5.51, 5.90)	(5.53, 6.01)	(5.54, 5.92)
		mean of std($\hat{T}$)	1.59	1.16	1.53	1.11
	3	true value		5.62		
		estimated $\hat{T}$	5.37	5.62	5.38	5.63
		confidence interval of $\hat{T}$	(5.06, 5.68)	(5.42, 5.81)	(5.09, 5.68)	(5.44 5.82)
		mean of std($\hat{T}$)	1.68	1.19	1.62	1.15

TABLE 1 (continued)

simulated 100 times

$\mu_x = 6$, $\sigma_x = 2$

$\rho = 0.0$

μ_y	σ_y		equally spaced, n = 15	equally spaced, n = 30	unequally spaced, n = 15	unequally spaced, n = 30
12	3	true value	5.89			
		estimated $\hat{T}$	5.94	5.95	5.94	5.94
		confidence interval of $\hat{T}$	(5.83, 6.05)	(5.89, 6.02)	(5.83, 6.05)	(5.88, 6.01)
		mean of std($\hat{T}$)	0.49	0.34	0.53	0.37
	4	true value	5.85			
		estimated $\hat{T}$	6.06	6.05	6.06	6.05
		confidence interval of $\hat{T}$	(5.97, 6.15)	(5.98, 6.11)	(5.97, 6.16)	(5.98, 6.13)
		mean of std($\hat{T}$)	0.45	0.34	0.48	0.36

TABLE 1 (continued)

simulated 100 times

$\mu_x = 6, \quad \sigma_x = 2$

$\rho = 0.0$

μ_y	σ_y	n =	equally spaced 15	equally spaced 30	unequally spaced 15	unequally spaced 30
15	4	true value	5.96			
		estimated $\hat{T}$	6.02	6.03	6.04	6.05
		confidence interval of $\hat{T}$	(5.94, 6.09)	(5.97, 6.08)	(5.96, 6.12)	(5.99, 6.11)
		mean of std($\hat{T}$)	0.38	0.27	0.43	0.31
	5	true value	5.93			
		estimated $\hat{T}$	6.02	6.03	6.01	6.01
		confidence interval of $\hat{T}$	(5.95, 6.10)	(5.97, 6.08)	(5.92, 6.09)	(5.95, 6.06)
		mean of std($\hat{T}$)	0.37	0.27	0.42	0.30

TABLE 1 (continued)

simulated 100 times

$\mu_x = 6$, $\sigma_x = 2$

$\rho = 0.25$

μ_y	σ_y		equally spaced n = 15	equally spaced 30	unequally spaced 15	unequally spaced 30
9	2	true value	5.75			
		estimated $\hat{T}$	5.54	5.65	5.56	5.69
		confidence interval of $\hat{T}$	(5.29, 5.79)	(5.48, 5.82)	(5.32, 5.79)	(5.53, 5.85)
		mean of std($\hat{T}$)	1.65	1.20	1.59	1.16
	3	true value	5.77			
		estimated $\hat{T}$	5.55	5.61	5.55	5.64
		confidence interval of $\hat{T}$	(5.30, 5.80)	(5.41, 5.81)	(5.30, 5.80)	(5.45, 5.82)
		mean of std($\hat{T}$)	1.58	1.17	1.52	1.12

TABLE 1 (continued)

simulated 100 times

$\mu_x = 6, \quad \sigma_x = 2$

$\rho = 0.25$

μ_y	σ_y	n =	equally spaced 15	equally spaced 30	unequally spaced 15	unequally spaced 30
		true value		5.96		
	3	estimated $\hat{T}$	5.98	5.97	5.99	5.97
		confidence interval of $\hat{T}$	(5.87, 6.08)	(5.90, 6.04)	(5.89, 6.09)	(5.89, 6.04)
		mean of std($\hat{T}$)	0.50	0.35	0.53	0.37
12		true value		5.94		
	4	estimated $\hat{T}$	5.90	5.93	5.91	5.94
		confidence interval of $\hat{T}$	(5.80, 6.00)	(5.87, 6.00)	(5.80, 6.01)	(5.87, 6.01)
		mean of std($\hat{T}$)	0.53	0.36	0.55	0.38

TABLE 1 (continued)

simulated 100 times

$\mu_x = 6, \quad \sigma_x = 2$

$\rho = 0.25$

μ_y	σ_y	n =	equally spaced 15	equally spaced 30	unequally spaced 15	unequally spaced 30
		true value		5.99		
		estimated $\hat{T}$	5.94	5.94	5.92	5.95
	4	confidence interval of $\hat{T}$	(5.86, 6.02)	(5.88, 5.99)	(5.84, 6.01)	(5.89, 6.00)
		mean of std($\hat{T}$)	0.37	0.26	0.42	0.30
15		true value		5.98		
		estimated $\hat{T}$	6.01	6.01	6.00	6.03
	5	confidence interval of $\hat{T}$	(5.93, 6.09)	(5.96, 6.06)	(5.91, 6.08)	(5.97, 6.09)
		mean of std($\hat{T}$)	0.38	0.27	0.43	0.30

TABLE 1 (continued)

simulated 100 times

$\mu_x = 6, \quad \sigma_x = 2$

$\rho = 0.5$

μ_y	σ_y	n =	equally spaced 15	equally spaced 30	unequally spaced 15	unequally spaced 30
		true value		5.87		
		estimated $\hat{T}$	5.81	5.67	5.79	5.66
	2	confidence interval of $\hat{T}$	(5.57, 6.06)	(5.49, 5.86)	(5.55, 6.03)	(5.48, 5.83)
		mean of std($\hat{T}$)	1.46	1.15	1.42	1.10
9		true value		5.92		
		estimated $\hat{T}$	5.58	5.61	5.56	5.60
	3	confidence interval of $\hat{T}$	(5.32, 5.84)	(5.43, 5.80)	(5.30, 5.82)	(5.42, 5.78)
		mean of std($\hat{T}$)	1.55	1.17	1.50	1.12

TABLE 1 (continued)

simulated 100 times

$\mu_x = 6$, $\sigma_x = 2$

$\rho = 0.5$

μ_y	σ_y		equally spaced n = 15	equally spaced n = 30	unequally spaced n = 15	unequally spaced n = 30
		true value		6.00		
		estimated $\hat{T}$	5.82	5.94	5.79	5.94
	3	confidence interval of $\hat{T}$	(5.71, 5.93)	(5.88, 6.01)	(5.67, 5.91)	(5.87, 6.01)
		mean of std($\hat{T}$)	0.54	0.36	0.57	0.38
12		true value		6.01		
		estimated $\hat{T}$	5.97	5.98	5.96	5.95
	4	confidence interval of $\hat{T}$	(5.88, 6.07)	(5.92, 6.05)	(5.87, 6.06)	(5.87, 6.02)
		mean of std($\hat{T}$)	0.46	0.33	0.50	0.35

TABLE 1 (continued)

simulated 100 times

$\mu_x = 6, \quad \sigma_x = 2$

$\rho = 0.5$

μ_y	σ_y		equally spaced n = 15	equally spaced n = 30	unequally spaced n = 15	unequally spaced n = 30
	4	true value	6.01			
		estimated $\hat{T}$	5.99	6.00	5.96	5.98
		confidence interval of $\hat{T}$	(5.92, 6.07)	(5.95, 6.05)	(5.88, 6.05)	(5.93, 6.03)
		mean of std($\hat{T}$)	0.37	0.27	0.42	0.30
15	5	true value	6.02			
		estimated $\hat{T}$	5.96	6.01	5.95	5.98
		confidence interval of $\hat{T}$	(5.89, 6.03)	(5.95, 6.07)	(5.87, 6.04)	(5.91, 6.05)
		mean of std($\hat{T}$)	0.37	0.27	0.42	0.30

TABLE 2

simulated 100 times

$\mu_x = 6, \quad \sigma_x = 2$

μ_y	σ_y			ρ = 0.0	0.25	0.5
		true value		5.63	5.75	5.87
			n = 15	(0.84)	(0.87)	(0.83)
		equally		[0.91]	[0.95]	[0.92]
		spaced	30	(0.89)	(0.95)	(0.93)
				[0.91]	[0.98]	[0.96]
	2					
			15	(0.84)	(0.91)	(0.85)
		unequally		[0.94]	[0.96]	[0.93]
		spaced	30	(0.89)	(0.96)	(0.94)
				[0.92]	[0.97]	[0.96]
9						
		true value		5.62	5.77	5.92
			15	(0.88)	(0.86)	(0.93)
		equally		[0.94]	[0.97]	[0.99]
		spaced	30	(0.99)	(0.95)	(0.96)
				[0.99]	[0.96]	[0.97]
	3					
			15	(0.91)	(0.90)	(0.97)
		unequally		[0.96]	[0.98]	[0.99]
		spaced	30	(0.98)	(0.95)	(0.96)
				[0.99]	[0.96]	[0.99]

TABLE 2 (continued)

simulated 100 times

$\mu_x = 6, \quad \sigma_x = 2$

μ_y	σ_y			ρ = 0.0	ρ = 0.25	ρ = 0.5
		true value		5.89	5.96	6.00
			n = 15	(0.89)	(0.91)	(0.97)
		equally		[0.99]	[0.99]	[1.00]
		spaced	30	(0.94)	(0.94)	(0.98)
				[1.00]	[0.99]	[1.00]
	3					
			15	(0.92)	(0.93)	(0.96)
		unequally		[1.00]	[0.99]	[1.00]
		spaced	30	(0.95)	(0.93)	(0.97)
				[1.00]	[0.97]	[1.00]
12						
		true value		5.85	5.94	6.01
			15	(0.90)	(0.91)	(0.96)
		equally		[0.98]	[0.99]	[1.00]
		spaced	30	(0.86)	(0.96)	(0.95)
				[0.97]	[1.00]	[1.00]
	4					
			15	(0.89)	(0.94)	(0.93)
		unequally		[0.97]	[1.00]	[0.98]
		spaced	30	(0.85)	(0.97)	(0.95)
				[0.97]	[1.00]	[0.99]

TABLE 2 (continued)

simulated 100 times

$\mu_x = 6, \quad \sigma_x = 2$

μ_y	σ_y		n	ρ = 0.0	0.25	0.5
		true value		5.96	5.99	6.01
		equally spaced	n = 15	(0.95)	(0.93)	(0.92)
				[1.00]	[1.00]	[1.00]
			30	(0.93)	(0.94)	(0.97)
				[0.98]	[0.99]	[1.00]
	4	unequally spaced	15	(0.93)	(0.93)	(0.94)
				[1.00]	[0.98]	[0.99]
			30	(0.93)	(0.96)	(0.97)
				[0.99]	[0.99]	[1.00]
15		true value		5.93	5.98	6.02
		equally spaced	15	(0.91)	(0.95)	(0.97)
				[1.00]	[1.00]	[1.00]
			30	(0.94)	(0.97)	(0.92)
				[1.00]	[1.00]	[1.00]
	5	unequally spaced	15	(0.91)	(0.94)	(0.92)
				[0.99]	[0.99]	[1.00]
			30	(0.97)	(0.96)	(0.90)
				[0.99]	[1.00]	[1.00]

The value in the parenthesis [bracket] is the percent of 2 std C.I.'s [3 std C.I.'s] which contain the true value.

ESTIMATION OF STAGE DEVELOPMENT

TIME WITH MORTALITY

by

Yeong-Ling Hwang

B.S., National Chung Hsing University, 1980

AN ABSTRACT OF 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

1988

ABSTRACT

Consider an insect that progresses through developmental stages throughout its lifetime. We here consider only one transition, i.e., from non-maturity to maturity. We take independent samples of fixed size and record the numbers of both mature and non-mature organisms, noting mortality in each group, at various observing times. We estimate the average time to maturity, and its corresponding standard error. We also present a simulation study comparing the estimated time to maturity with its theoretical value.