J. Agri. & Fore. 2012. 61(1): 37-45

# Assessing the Application of the Jackknife and Bootstrap Techniques to the Estimation of the Variability of the Net Reproductive Rate and Gross Reproductive Rate: a Case Study in *Bactrocera cucurbitae* (Coquillett) (Diptera: Tephritidae)

Yu-Bing Huang [1)]  Hsin Chi [2)]

## Abstract

In the study of life tables, scientists usually begin with a single cohort and record survival, development, and fecundity until the death of all individuals. Because it is extremely time-consuming, the replication of life table studies is generally impractical. To estimate the variability of life table statistics, jackknife and bootstrap techniques are usually used. However, the biological meaning of these statistical procedures is not yet fully understood. In this paper, we assessed the use of the jackknife and the bootstrap in estimating the variability of the net reproductive rate and gross reproductive rate. Life table data for the melon fly, *Bactrocera cucurbitae* (Diptera: Tephritidae) reared on cucumber, sponge gourd, and carrot medium were used as examples. Our results show that the jackknife is inadequate for the estimation of the variability of both the net reproductive rate and gross reproductive rate and may overestimate the variability.

**Key words:** Life table, Variability, Jackknife, Bootstrap, Melon fly.

## Introduction

Life table studies are essential in population ecology in both theoretical research and practical applications. In most cases, researchers begin with a limited number of newborns and observe the development, survival, and fecundity at regular time intervals, e.g., days, from birth to

---

1) PhD Student, Department of Entomology, National Chung Hsing University, Taiwan, R.O.C.

2) Professor, Department of Entomology, National Chung Hsing University, Taiwan, R.O.C.
Corresponding author, E-mail: hsinchi@dragon.nchu.edu.tw.

death (4). To complete such a study usually requires a few months or a year (e.g., Tsai and Chi (26)). Because it is extremely time- and labor-consuming, replication is an impractical or even impossible method for detecting the variability of the life table. Therefore, only a single value can be obtained for each life table parameter (i.e., the intrinsic rate of increase, the net reproductive value, and the mean generation time) from the raw life table data, i.e., the daily survival rate, developmental rate, and fecundity for each individual of the cohort. Despite the thorough development of life table theories (1, 6, 17, 18, 20, 21, 22), no closed-form equations are available for the calculation of the variance of population parameters. Therefore, differences between population parameters of different cohorts cannot be statistically tested. Keyfitz (12) proposed the use of the jackknife method (27) for estimating the variance of the intrinsic rate of increase. Since then, this method has been discussed and used for many animal populations (3, 5, 8, 11, 16, 23, 24). Efron (9) reported that the bootstrap technique is more reliable than the jackknife technique for estimating variances. The variances obtained with the jackknife technique are usually much larger than those obtained with the bootstrap technique. Both techniques are based on the resampling procedure of deleting or repeatedly choosing all data for specific individuals. The statistical bases of these two techniques have been discussed in detail by Efron (9) and Efron and Tibshirani (10). To date, the biological meaning of the estimated means, variances, and standard errors has been discussed less frequently.

Kuczynski’s net reproduction rate (13) was one of the most important demographic parameters. Kuczynski’s contribution has been frequently reviewed since its original publication (25), although he identified the source of this concept as a work of Richard Boeckh published in 1890 (19). The concept of net reproduction rate linked population growth not only with economic pressure (13) but also population projection (14, 15). In view of its importance, the statistics of the net reproductive rate deserve special attention not only in ecology but also in pest management.

The sources of biological variability can be environmental, genetic, or both environmental and genetic (i.e., a combined effect). In life table studies, the limited number of individuals implies the limited genetic variability (small gene pool) of the experimental cohort. It may not be possible to incorporate the actual variability of a population, and the use of random sampling contributes somewhat to the uncertainty of the calculation of population parameters. In this paper, we explore the application of the jackknife and bootstrap techniques to the estimation of the net reproductive rate and the gross reproductive rate. Life table data on the melon fly (*Bactrocera cucurbitae*) are used as an example.

## Materials and Methods

### Life tables of the melon fly

Melon flies were collected in a field used to grow vegetables. The flies were subsequently reared on cucumber (*Cucumis sativus* L.) in the laboratory. Before the beginning of the life table studies, the colony was maintained for two generations on cucumber, carrot medium (mashed *Daucus carota* L. mixed with sucrose and yeast hydrolysate), and sponge gourd (*Luffa cylindrica* Roem). The detailed methods used for insect rearing and life table analysis are reported in Huang and Chi (11).

### Demographic analysis

The raw life history data were analyzed according to the age-stage, two-sex life table theory (6), and the method described by Chi (4). The net reproductive rate ($R_0$) is calculated as

$$R_0 = \sum_{x=0}^{\infty} l_x m_x \,, \tag{1}$$

where $l_x$ is the age-specific survival rate and $m_x$ is the age-specific fecundity at age $x$. The gross reproductive rate (*GRR*) is calculated as

$$GRR = \sum_{x=0}^{\infty} m_x \,. \tag{2}$$

### Jackknife technique

For the jackknife and bootstrap techniques, we adopted the procedures of Meyer et al. (24) and Efron and Tibshirani (10) in this and the next section. To apply the jackknife technique, we first calculate the net reproductive rate for all $n$ individuals of the cohort ($R_{0,all}$) with Equation 1 as

$$R_{0,all} = \sum_{x=0}^{\infty} l_x m_x \,, \tag{3}$$

where $l_x$ and $m_x$ are calculated including all individuals in the cohort. We then omit individual $i$ and use the other $n - 1$ individuals to calculate the jackknife value of $R_{0,i}$ as

$$R_{0,i} = \sum_{x=0}^{\infty} l_x m_x \,, \tag{4}$$

where $l_x$ and $m_x$ are calculated by omitting the data for individual $i$. Next, we calculate $R_{0,i\text{-}pseudo}$ as

$$R_{0,i-pseudo} = n \cdot R_{0,all} - (n-1) \cdot R_{0,i} \,. \tag{5}$$

The mean $R_{0,J}$, variance $s_J^2$, and standard error se($R_{0,J}$) according to the jackknife technique are calculated as

$$R_{0,J} = \frac{\sum_{i=1}^{n} R_{0,i-pseudo}}{n} \tag{6}$$

$$s_J^2 = \frac{\sum_{i=1}^{n}\left(R_{0,i-pseudo} - R_{0,J}\right)^2}{n-1} \tag{7}$$

$$\mathrm{se}\left(R_{0,J}\right) = \sqrt{\frac{s_J^2}{n}} \,. \tag{8}$$

The same method is used for the corresponding estimates of the gross reproductive rate, i.e., $GRR_{i\text{-}pseudo}$, $GRR_J$, and the variance and standard error of $GRR_J$.

### Bootstrap technique

In the bootstrap procedure, we randomly take a sample of $n$ individuals from the cohort with replacement and calculate the $R_{0,i\text{-}boot}$ for this bootstrap sample as

$$R_{0,i-boot} = \sum_{x=0}^{\infty} l_x m_x \,, \tag{9}$$

where the subscript *i-boot* represents the *i*th bootstrap and $l_x$ and $m_x$ are calculated from the $n$ individuals selected randomly with replacement. Generally, the data on the same individual are repeatedly selected. We repeat this procedure $m$ times ($m > 200$, according to Meyer et al. (24)) and compute the mean of these $m$ bootstraps as

$$R_{0,B} = \frac{\sum_{i=1}^{m} R_{0,i-boot}}{m}. \tag{10}$$

In this paper, we repeat the bootstrap 10,000 times ($m = 10{,}000$). The variance $s_B^2$ and standard error $\mathrm{se}(R_{0,B})$ of these $m$ bootstraps can be calculated as

$$s_B^2 = \frac{\sum_{i=1}^{m}\left(R_{0,i-boot} - R_{0,B}\right)^2}{m-1} \tag{11}$$

$$\mathrm{se}(R_{0,B}) = \sqrt{s_B^2}\,. \tag{12}$$

The same method is used for the corresponding estimates of the gross reproductive rate, i.e., $GRR_{i\text{-}boot}$, $GRR_B$, and the variance and standard error of $GRR_B$.

# Results and Discussion

The estimated means and standard errors of the net reproductive rates and gross reproductive rates of *B. cucurbitae* on three rearing media were listed in Table 1. The means and standard errors estimated with both techniques are very close. The values of these estimates are close because both methods are resampling procedures and estimates are obtained from the raw life history data of the same cohorts. If all $R_{0,i\text{-}pseudo}$, $GRR_{i\text{-}pseudo}$, $R_{0,i\text{-}boot}$, and $GRR_{i\text{-}boot}$ are plotted as a frequency distribution and fitted to a normal distribution, the high coefficients of determination of normality obtained with the bootstrap technique (Fig. 1 to 3) showed the normal distribution of the estimates and indicated that the bootstrap technique is a better choice than the jackknife technique.

If the bootstrap technique is used, every bootstrap sample can be generated differently by sampling with replacement. This property implies that variability in $R_{0,B}$ and $GRR_B$ is

Table 1. Means, standard errors, variance of net reproductive rate ($R_0$), and gross reproductive rate ($GRR$) estimated with the jackknife and bootstrap techniques

| Rearing medium | | $R_0$ | | $GRR$ | |
|---|---|---|---|---|---|
| | | Jackknife | Bootstrap | Jackknife | Bootstrap |
| Cucumber | Mean ± SE<br>Variance | 137.8 ± 30.9<br>59116 | 138.8 ± 30.5<br>932 | 227.0 ± 54.7<br>185152 | 226.4 ± 52.7<br>2782 |
| Sponge gourd | Mean ± SE<br>Variance | 172.3 ± 26.0<br>67484 | 173.0 ± 26.2<br>686 | 309.5 ± 39.6<br>157169 | 308.8 ± 38.9<br>1517 |
| Carrot medium | Mean ± SE<br>Variance | 46.8 ± 8.8<br>7662 | 47.0 ± 8.5<br>72.69 | 114.9 ± 20.0<br>40060 | 114.8 ± 19.2<br>367.7 |

always possible. Even with the same number of replicates *m*, different $R_{0,B}$ and $GRR_B$ will be obtained. The variability is significant with smaller *m*. With increasing *m*, the variability will diminish, but it can always be observed in the variance and standard error as well as in the frequency distribution of $R_{0,B}$ and $GRR_B$. In contrast, there is always exactly one jackknife result, and all the frequency distributions from the jackknife technique in Fig. 1 to 3 will be the same.

Life table methodology is based on solid mathematical theory. Lotka (21) gave a mathematical proof of the relationship between the net reproductive rate and the intrinsic rate of increase. Lewis (20) derived the relationship between the net reproductive rate and the finite rate. Although the jackknife technique has been used in demographic statistics for several decades, Chi and Yang (7) noticed the problem and explicitly stated, "It results in some degree of discrepancy between the estimated means and their definition". This observation reminds us that the application of resampling methods, such as the jackknife and bootstrap, should be performed with particular caution.

If the jackknife technique is used and the omitted individual is a male or one that died in the preadult stage, the $R_{0,J}$ will be zero. This effect can be observed in all the figures showing the jackknife results (Fig. 1 to 3). A net reproductive rate of zero means that the population does not produce any offspring. According to life table theory, we cannot

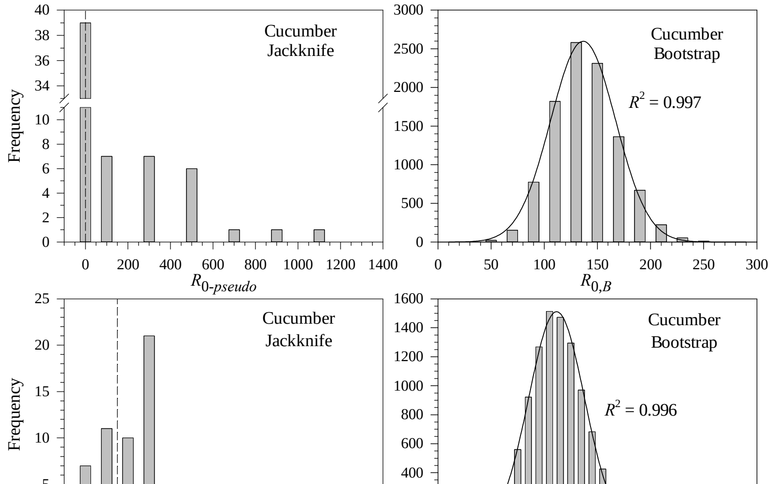

Fig. 1. Frequency distribution of the $R_{0\text{-}pseudo}$, $R_{0,B}$, $GRR_{pseudo}$, and $GRR_B$ of *B. cucurbitae* reared on cucumber.

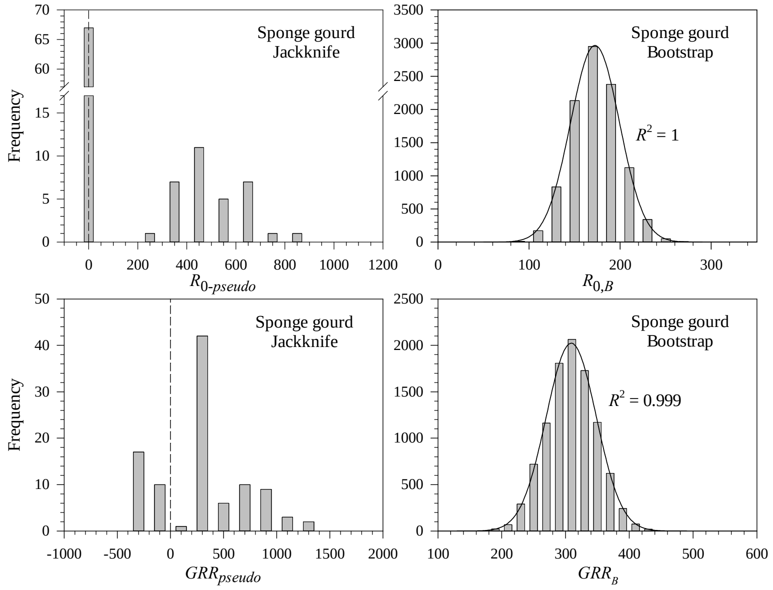

Fig. 2. Frequency distribution of the $R_{0\text{-}pseudo}$, $R_{0,B}$, $GRR_{pseudo}$, and $GRR_B$ of *B. cucurbitae* reared on sponge gourd.

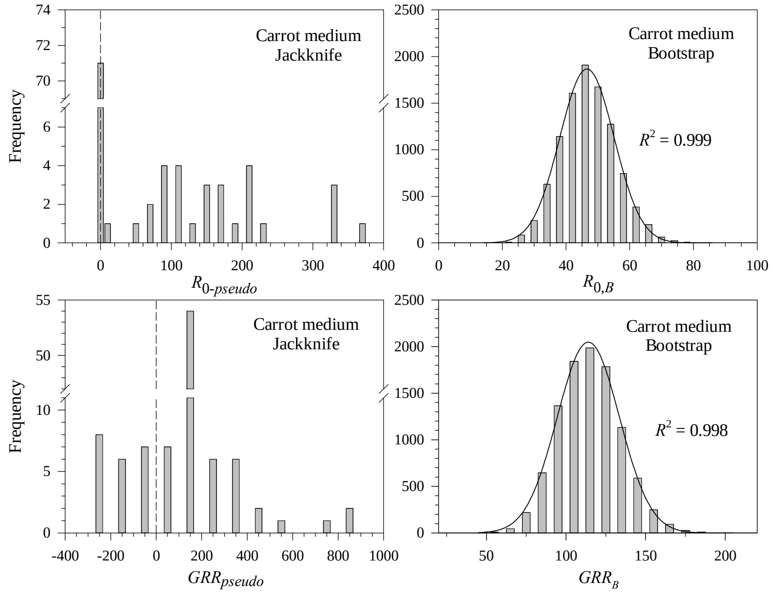

Fig. 3. Frequency distribution of the $R_{0\text{-}pseudo}$, $R_{0,B}$, $GRR_{pseudo}$, and $GRR_B$ of *B. cucurbitae* reared on carrot medium.

calculate the intrinsic rate if the net reproductive rate is zero. However, omitting one male individual or one that died before the adult stage still allows us to calculate the intrinsic rate of increase. This outcome contradicts life table theory. We therefore conclude that the jackknife technique should not be used for the estimation of the standard error of the net reproductive rate.

In life table statistics, we usually calculate the mean generation time ($T$) as

$$T = \frac{\ln R_0}{r}. \qquad (13)$$

The mean generation time is defined as the length of time that a stable population needs to increase to $R_0$-fold of its original size. It is a parameter usually reported in life table research (11). If the jackknife technique should not be used for the estimation of standard errors of the net reproductive rate, we cannot use Eq. 13 to estimate the mean generation time. Consequently, the jackknife should not be used for the estimation of population parameters in demography.

## Conclusions

Most insect populations are two-sex populations and are stage structured. However, the traditional female age-specific life tables (1, 2, 17, 18, 20) ignore the male population and the variation of developmental rates among individuals. Their application to insect populations usually results in errors in population parameters and erroneous relationships among the net reproductive rate, the intrinsic rate, the mean fecundity, and other derived quantities. In this study, we used the age-stage, two-sex life table theory developed by Chi and Liu (6) and Chi (4). Because both sexes and stage differentiation are incorporated, our analysis includes the variability due to changes in sex ratio and developmental rate. These important factors are, however, ignored if the traditional female age-specific life table is used. As Yu et al. (28) observed, *GRR* ignores the age-specific survival rate; it is not a biologically meaningful statistic. Accordingly, the gross reproductive rate should be used with caution. Our analyses demonstrate that the jackknife technique should not be used in the estimation of the mean and standard errors of the net reproductive rate and other population parameters.

## Acknowledgments

We thank Shyng Jung Wu for his help on the life table experiment. This research was supported partly by grants from the Bureau of Animal and Plant Health Inspection and Quarantine, Taiwan (90AS-6.2.3-BQ-B1[9], 91AS-7.2.3-BQ-B1[2], 92AS-1.8.1-BQ-B4[2], 96AS-14.2.1-BQ-B4[6], 97AS-14.2.1-BQ-B3[2]) and the National Science Council (NSC98-2313-B-005-020-MY3, NSC95-2621-B-005-009, NSC94-2621-B-005-003, NSC93-2621-B-005-008).

## References

1. Birch, L. C. 1948. The intrinsic rate of natural increase of an insect population. J. Anim. Ecol. 17: 15-26.
2. Carey, J. R. 1993. Applied demography for biologists with special emphasis on insects. Oxford University Press, New York.
3. Caswell, H. 2001. Matrix population models: construction, analysis, and interpretation, 2nd ed. Sinauer Associates, Sunderland, Massachusetts.
4. Chi, H. 1988. Life table analysis incorporating both sexes and variable development rates among individuals. Environ. Entomol. 17: 26-34.
5. Chi, H. 1989. Two-sex life table of the silkworm, *Bombyx mori* L. Chinese J. Entomol. 9: 141-150.
6. Chi, H. and H. Liu. 1985. Two new methods for the study of insect population ecology. Acad. Sin., Bull. Inst. Zool. 24: 225-240.
7. Chi, H. and T. C. Yang. 2003. Two-sex life table and predation rate of *Propylaea japonica* Thunberg (Coleoptera: Coccinellidae) fed on *Myzus persicae* (Sulzer) (Homoptera: Aphidiae). Environ. Entomol. 32: 327-333.
8. Chi, H. and W. M. Getz. 1988. Mass rearing and harvesting based on an age-stage, two-sex life table: A potato tuberworm (Lepidoptera: Gelechiidae) case study. Environ. Entomol. 17: 18-25.
9. Efron, B. 1982. The Jackknife, the bootstrap and other resampling plans. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, USA.
10. Efron, B. and R. J. Tibshirani. 1993. An introduction to the bootstrap. Chapman and Hall, New York.
11. Huang, Y. B. and H. Chi. 2012. Age-stage, two-sex life tables of *Bactrocera cucurbitae* (Coquillett) (Diptera: Tephritidae) with a discussion on the problem of applying female age-specific life tables to insect populations. Insect Sci. 19: 263-273.
12. Keyfitz, N. 1977. Introduction to the mathematics of populations. Addison-Wesley, Reading, Massachusetts.
13. Kuczynski, R. R. 1930. Population growth and economic pressure. Annals of the American Academy of Political and Social Science, Vol. 150, Economics of World Peace, pp. 1-6.
14. Kuczynski, R. R. 1937. Population trends. Math. Gaz. 21(243): 99-105.
15. Kuczynski, R. R. 1944. World population problems. Intl. Aff. 20: 449-457.
16. Lenski, R. E. and P. M. Service. 1982. The statistical analysis of population growth rates calculated from schedules of survivorship and fecundity. Ecology 63: 655-662.

17. Leslie, P. H. 1945. On the use of matrices in certain population mathematics. Biometrika 33: 183-212.
18. Leslie, P. H. 1948. Some further notes on the use of matrices in population mathematics. Biometrika 35: 213-245.
19. Lewes, F. M. M. 1984. A note on the origin of the net reproduction ratio. Popul. Stud. 38: 321-324.
20. Lewis, E. G. 1942. On the generation and growth of a population. Sankhya 6: 93-96.
21. Lotka, A. J. 1907a. Studies on the mode of growth of material aggregates. Am. J. Sci. 24: 199-216.
22. Lotka, A. J. 1907b. Relation between birth rates and death rates. Science 26: 21-22.
23. Maia, A. H. N., A. J. B. Luiz and C. Campanhola. 2000. Statistical inference on associated fertility life table parameters using jackknife technique: computational aspects. J. Econ. Entomol. 93: 511-518.
24. Meyer, J. S., C. G. Ingersoll, L. L. McDonald and M. S. Boyce. 1986. Estimating uncertainty in population growth rate: Jackknife vs. bootstrap techniques. Ecology 67: 1156-1166.
25. Population Council. 2008. Robert Kuczynski on population growth and economic pressure. Popul. Dev. Rev. 34: 539-546.
26. Tsai, T. J. and H. Chi. 2007. Temperature-dependent demography of *Supella longipalpa* (Blattodea: Blattellidae). J. Med. Entomol. 44: 772-778.
27. Tukey, J. 1958. Bias and confidence in not quite large samples. Ann. Math. Statist. 29: 614.
28. Yu, J. Z., H. Chi and B. H. Chen. 2005. Life table and predation of *Lemnia biplagiata* (Coleoptera: Coccinellidae) fed on *Aphis gossypii* (Homoptera: Aphididae) with a proof on relationship among gross reproduction rate, net reproduction rate, and preadult survivorship. Ann. Entomol. Soc. Am. 98: 475-482.

Received: November 15, 2011.
Accepted: February 18, 2012.