The Annals of Statistics
1978, Vol. 6, No. 1, 112–116

ESTIMATION OF A BIOMETRIC FUNCTION

BY GRACE L. YANG

University of Maryland

In the analysis of life tables one biometric function of interest is the life expectancy at age x, $e_x = E[X - x \mid X > x]$. Estimation of e_x is considered, the standard estimator used in life tables is shown to be asymptotically unbiased, uniformly strong consistent, and converges in distribution to a Gaussian process. The connections of the estimator studied in this article and that used in reliability theory are illustrated.

1. Introduction. Let X be a nonnegative random variable defined on a fixed probability space $(\Omega, \mathscr{A}, P)$. Assume that X has a mean η, a finite variance σ^2, and a density function $f(x) > 0$, $x \geqq 0$. Let $S(x) = 1 - F(x)$ be the survival function. Define

$$e_x = (\textstyle\int_x^\infty S(v)\,dv)/S(x) \qquad \text{for} \quad x \in [0, \infty)\,. \tag{1.1}$$

For life tables, e_x is called the life expectancy at age x, or more generally a biometric function (Chiang [3]). In biometry e_x is defined via the force of mortality $\mu(x) = f(x)/S(x)$,

$$e_x = \textstyle\int_0^\infty \exp\{-\int_0^v \mu(x + y)\,dy\}\,dv\,, \qquad \text{for} \quad x \in [0, \infty)\,.$$

Like $\mu(x)$, e_x also determines $S(x)$,

$$S(x) = e_0 e_x^{-1} \exp\{-\textstyle\int_0^x e_v^{-1}\,dv\}\,, \qquad x \in [0, \infty)\,.$$

Therefore $\mu(x)$, e_x and $f(x)$ are equivalent in determining $S(x)$.

While the theoretical investigation of $\mu(x)$ occupies a central position in reliability theory, little attention has been paid to e_x. The intent of this article is to study the estimation of e_x.

The estimator $\hat{e}_x$ of e_x, to be discussed (see (2.5)) corresponds to a cohort life table estimate. In Section 2, $\hat{e}_x$ is shown to have a multiplicative bias in a finite sample and to be uniformly strong consistent as sample size increases to infinity. In Section 3, $\hat{e}_x$, considered as a function of x, is shown to converge in distribution to a limiting Gaussian process. The major term in $\hat{e}_x$ is a statistic of the form

$$T_x = n \textstyle\int_x^\infty S_n(v)\,dv = \sum_{j=1}^n I(X_j - x)(X_j - x) \qquad \text{(cf. (2.1) to (2.4))} \tag{1.2}$$

where $S_n(v)$ denotes the empirical survival function. T_x is related to the total-time-on-test statistic $H_n(x)$ by

$$H_n(x) = n \textstyle\int_0^x S_n(v)\,dv = T_0 - T_x$$

in reliability theory (Barlow et al. [1]). But T_x and $H_n(x)$ differ in the range of

Received June 1975; revised August 1976.

AMS 1970 *subject classifications*. Primary 62P10, 62G05, 62E20.

Key words and phrases. Biometric function, life expectancy, reliability, total time on test, uniform strong consistency, bias, tightness, limiting normal process.

integration. Therefore, the techniques used in proving the limiting results differ in both cases. In our calculations it is advantageous to use the form on the extreme right of (1.2) which preserves the i.i.d. structure as opposed to the usual way of treating $H_n(x)$ as a function of the ordered X's. Another note is the analogy of T_x to $\sum_{j=1}^n I(X_j - x)$ used in Donsker's theorem, and T_x reduces to $\sum_{j=1}^n X_j$ at $x = 0$.

2. Bias and uniform consistency of $\hat{e}_x$. Let $X_1, \cdots, X_n$ be i.i.d. random variables from F. Let

(2.1) $$I(a) = 1 \quad \text{or} \quad 0 \qquad \text{according to} \quad a > 0 \quad \text{or} \quad a \leqq 0\,,$$

(2.2) $$l_x = \sum_{j=1}^n I(X_j - x) \qquad \text{for every} \quad x \in [0, \infty)\,,$$

(2.3) $$S_n(x) = l_x/n\,,$$

and

(2.4) $$T_x = n \int_x^\infty S_n(v)\,dv = \sum_{j=1}^n I(X_j - x)(X_j - x)\,.$$

When F is not specified in a parametric form, the proposed estimator $\hat{e}_x$ is

(2.5) $$\hat{e}_x = (S_n(x))^{-1} \int_x^\infty S_n(v)\,dv\, I(X_{(n)} - x)$$

where $X_{(n)} = \max_{1\leqq j\leqq n} X_j$ [4]. Interpreting x as age, $\hat{e}_x$ gives the average time to be lived by those l_x individuals having age larger than x.

In the sequel, the limits are taken as $n \to \infty$ unless otherwise stated. The following lemmas discuss the bias and uniform strong consistency of $\hat{e}_x$.

LEMMA 1. *For every $x \in [0, \infty)$, $E\hat{e}_x = e_x P[l_x > 0]$. Also, $\hat{e}_x$ is an unbiased estimator for e_0 and asymptotically unbiased for $x > 0$.*

PROOF. It suffices to notice that given l_x, there is a sample of size l_x from $(f(y)/S(x))I(y - x)$ and l_x is a binomial random variable. Since $l_x = 0$ with positive probability, on occasion the data contains no observation. Moreover, note that no unbiased estimator for e_x exists for $x > 0$.

LEMMA 2. *Let b be any fixed positive number; $\hat{e}_x$ in* (2.5) *is uniformly strong consistent for e_x in the sence that*

(2.6) $$P[\sup_{0\leqq x\leqq b} |\hat{e}_x - e_x| \to 0, \ \text{as} \ n \to \infty] = 1\,.$$

PROOF. Since $EX < \infty$, applying a lemma due to Marshall and Proschan [6] in the form given in [1] (page 237),

(2.7) $$\int_x^\infty S_n(v)\,dv \to \int_x^\infty S(v)\,dv \qquad \text{uniformly in} \quad x \quad \text{with probability one.}$$

Combining (2.7) with the fact that $S_n^{-1}(x) \to S^{-1}(x)$ uniformly in $x \in [0, b]$ with probability one gives (2.6).

3. Weak convergence of $\hat{e}_x$. For notational convenience, denote

(3.1) $$\theta(t, u) = EI(t < F(X) \leqq u)X \qquad \text{and}$$
$$\sigma^2(t, u) = \operatorname{Var}\,(I(t < F(X) \leqq u)X)\,,$$

where $I(t < F(X) \leqq u)$ is the indicator of $[t < F(x) \leqq u]$ for $0 \leqq t < u \leqq 1$.

THEOREM 1. *Let $t = F(x)$. Let e_x and $\hat{e}_x$ be as given in* (1.1) *and* (2.5). *The process*

$$n^{\frac{1}{2}}(\hat{e}_{F^{-1}(t)} - e_{F^{-1}(t)}) \qquad \text{for} \quad t \in [0, b], \quad b < 1,$$

converges in distribution to a Gaussian process $U(t)$ with mean zero and covariance function

$$(3.2) \qquad \Gamma(s, t) = (1-s)^{-2}(1-t)^{-2}\{(1-s)(1-t)\sigma^2(t, 1) - t(1-s)\theta^2(t, 1)\} \qquad 0 \leqq s \leqq t \leqq b.$$

The proof of Theorem 1 relies on Lemmas 3, 4 and Theorem 2 below.

LEMMA 3. *Let $X_1, \cdots, X_n$ be* i.i.d. *satisfying assumptions specified in Section* 1. *Let*

$$(3.3) \qquad V_n(t) = n^{-\frac{1}{2}} \sum_{j=1}^n [I(F(X_j) - t)X_j - E(I(F(X_j) - t)X_j)], \qquad \text{for} \quad t \in [0, 1].$$

Then, for any positive integer k and $0 \leqq t_1 < t_2 < \cdots < t_{k+1} = 1$ the finite dimensional distribution of $V_n(t)$ converges to a multivariate normal, $[V_n(t_1), \cdots, V_n(t_k)]' \to_{\mathscr{L}} \mathscr{N}_k(0, A\Sigma A')$, where $\mathscr{N}_k(0, A\Sigma A')$ denotes a k-dimensional normal distribution, A a $k \times k$ matrix, A' its transpose, with components

$$(3.4) \qquad a_{ij} = 1 \quad \text{if} \quad i \leqq j \quad \text{and} \quad 0 \quad \text{otherwise}$$

and Σ a $k \times k$ positive definite covariance matrix with components

$$(3.5) \qquad \sigma_{ii} = \sigma^2(t_i, t_{i+1}), \qquad \sigma_{ij} = -\theta(t_i, t_{i+1})\theta(t_j, t_{j+1}), \qquad i \neq j = 1, \cdots, k.$$

PROOF. Let $\zeta_n(t) = \sum_{j=1}^n I(F(X_j) - t)X_j$, for brevity,

$$D_n' = [\zeta_n(t_1) - \zeta_n(t_2), \cdots, \zeta_n(t_k) - \zeta_n(t_{k+1})],$$

$$\Theta' = [\theta(t_1, t_2), \theta(t_2, t_3), \cdots, \theta(t_k, 1)] \qquad \text{(cf. (3.1))},$$

and

$$\lambda' = (\lambda_1, \cdots, \lambda_k) \in R^k, \quad \text{a } k\text{-dimensional scalar.}$$

Then the inner product $\lambda' D_n$ may be expressed as a sum of i.i.d. random variables,

$$\lambda' D_n = \sum_{j=1}^n Y_j X_j,$$

where $Y_j = \sum_{\alpha=1}^k \lambda_\alpha I(t_\alpha < F(X_j) \leqq t_{\alpha+1})$.

Direct computation gives

$$EY_jX_j = \lambda'\Theta \qquad \text{and} \qquad \mathrm{Var}\,(Y_jX_j) = \lambda'\Sigma\lambda.$$

It then follows from the central limit theorem that

$$(n\lambda'\Sigma\lambda)^{-\frac{1}{2}}\lambda'(D_n - ED_n) \to_{\mathscr{L}} \mathscr{N}(0, 1), \qquad \text{for every} \quad \lambda \in R^k.$$

Application of the Cramér–Wold theorem (Billingsley [2]) gives

$$n^{-\frac{1}{2}}(D_n - ED_n) \to_{\mathscr{L}} \mathscr{N}_k(0, \Sigma),$$

and $F(x)$ being strictly increasing insures Σ being positive definite. Since

$$n^{\frac{1}{2}}V_n(t_\alpha) = \sum_{j=\alpha}^k [\zeta_n(t_j) - \zeta_n(t_{j+1}) - E(\zeta_n(t_j) - \zeta_n(t_{j+1}))],$$

applying transformation A in (3.4) to D_n completes the proof of Lemma 3.

Next we show that the sequence $V_n(t)$ given in (3.3) is tight (Le Cam [5]). The sample functions of V_n are right continuous and have left-hand limits, so we shall establish the tightness condition via verifying the hypothesis of Theorem 15.6 in [2].

LEMMA 4. *For arbitrarily chosen values t_1, t_2, t_3 subject to $0 \leqq t_1 < t_2 < t_3 \leqq 1$, $V_n(t)$ in* (3.3) *satisfies*

$$E|V_n(t_2) - V_n(t_1)|^2 |V_n(t_3) - V_n(t_2)|^2 \leqq (g(t_3) - g(t_1))^2 , \tag{3.6}$$

where $g(t) = \int_{F^{-1}(t)}^{\infty} v^2 f(v)\,dv$ for t in [0, 1].

PROOF. For simplicity, denote

$$\theta_i = \theta(t_i, t_{i+1}) \quad \text{and} \quad \sigma_i^2 = \sigma^2(t_i, t_{i+1}) \quad \text{for} \quad i = 1, 2 \qquad (\text{cf. } (3.1)) .$$

Proceed from the left-hand side of (3.6)

$$\begin{aligned} &E|V_n(t_2) - V_n(t_1)|^2 |V_n(t_3) - V_n(t_2)|^2 \\ &\quad = n^{-2} E[|\textstyle\sum_{j=1}^n (I(t_1 < F(X_j) \leqq t_2)X_j - \theta_1)|^2 |\sum_{j=1}^n (I(t_2 < F(X_j) \leqq t_3)X_j - \theta_2)|^2] \\ &\quad = \sigma_1^2\sigma_2^2 + 2\theta_1^2\theta_2^2 + n^{-1}[\theta_1^2\sigma_2^2 + \theta_2^2\sigma_1^2 - \sigma_1^2\sigma_2^2 - 3\theta_1^2\theta_2^2] \\ &\quad \leqq 2(\sigma_1^2 + \theta_1^2)(\sigma_2^2 + \theta_2^2) = 2[g(t_1) - g(t_2)][g(t_2) - g(t_3)] \\ &\quad \leqq [g(t_3) - g(t_1)]^2 . \end{aligned}$$

□

THEOREM 2. *Let $V_n(t)$ be given in* (3.3). *Then*

$$V_n \to_{\mathscr{L}} V$$

where V is a Gaussian process with

$$EV(t) = 0$$

$$EV(s)V(t) = \sigma^2(t, 1) - \theta(s, t)\theta(t, 1) , \qquad \text{for} \quad 0 \leqq s \leqq t \leqq 1 ,$$

and σ^2, $\theta(s, t)$ are defined in (3.1).

PROOF. Theorem 2 follows from Theorem 15.6 in [2] which is applicable because of Lemmas 3 and 4. □

PROOF OF THEOREM 1. Using (2.4) and (2.5), write, for $x < X_{(n)}$,

$$\begin{aligned} n^{\frac{1}{2}}(\hat{e}_x - e_x) = n^{\frac{1}{2}}(S_n(x)S(x))^{-1}[S(x) \textstyle\int_x^\infty (S_n(v) - S(v))\,dv \\ - (S_n(x) - S(x)) \textstyle\int_x^\infty S(v)\,dv] . \end{aligned} \tag{3.7}$$

Let $t = F(x)$, $V_n(t)$ be given in (3.3), and

$$\begin{aligned} W_n(t) &= n^{-\frac{1}{2}} \textstyle\sum_{j=1}^n [I(F(X_j - t)) - EI(F(X_j - t))] \\ &= n^{\frac{1}{2}}(S_n(F^{-1}(t)) - (1 - t)) . \end{aligned} \tag{3.8}$$

Then,

$$\begin{aligned} n^{\frac{1}{2}}(\hat{e}_x - e_x) &= n^{\frac{1}{2}}(\hat{e}_{F^{-1}(t)} - e_{F^{-1}(t)}) \\ &= (S_n(F^{-1}(t))(1 - t))^{-1}\{(1 - t)V_n(t) - [F^{-1}(t)(1 - t) \\ &\qquad + \textstyle\int_t^1 (1 - u)\,dF^{-1}(u)]W_n(t)\} \qquad \text{for} \quad t \in [0, 1) . \end{aligned} \tag{3.9}$$

$S_n(F^{-1}(t))$ converges to $(1 - t)$ a.s. It suffices to show that the random function in the bracket of (3.9) converges weakly to a Gaussian process. First, observe that the tightness of the sequence $n^{\frac{1}{2}}(\hat{e}_{F^{-1}(t)} - e_{F^{-1}(t)})$ is a result of the tightness of $W_n(t)$ and $V_n(t)$; the former is well known and the latter is given in Lemma 4. Secondly, using the Cramér–Wold technique we can show that the finite-dimensional distributions of the vector process $[V_n(t), W_n(t)]'$ converge to multivariate normal distributions. To complete the proof of Theorem 1 what remains is the computation of the covariance function of the limiting process of $n^{\frac{1}{2}}(\hat{e}_{F^{-1}(t)} - e_{F^{-1}(t)})$. The procedure is straightforward, we shall only present the final result in (3.2). □

REMARK. Applying (3.8), rewrite (3.7) as

$$\begin{aligned} n^{\frac{1}{2}}(\hat{e}_x - e_x) &= n^{\frac{1}{2}}(\hat{e}_{F^{-1}(t)} - e_{F^{-1}(t)}) \\ &= [S_n(F^{-1}(t))(1-t)]^{-1}[(1-t)\textstyle\int_t^1 W_n(u)\,dF^{-1}(u) \\ &\quad - W_n(t)\textstyle\int_t^1 (1-u)\,dF^{-1}(u)]\,. \end{aligned} \tag{3.10}$$

$\int_t^1 (1-u)\,dF^{-1}(u)$ exists because $EX < \infty$. Since $W_n(t)$ converges weakly to the Brownian bridge $W(t)$, (3.10) suggests that

$$\begin{aligned} n^{\frac{1}{2}}(\hat{e}_{F^{-1}(t)} - e_{F^{-1}(t)}) \to_{\mathscr{L}} (1-t)^{-2}[(1-t)\textstyle\int_t^1 W(u)\,dF^{-1}(u) \\ - W(t)\textstyle\int_t^1 (1-u)\,dF^{-1}(u)]\,. \end{aligned} \tag{3.11}$$

However, this is not immediate since in general $\int_t^1 dF^{-1}(u)$ is infinite unless X is bounded. Under the assumption that variance of X is finite using integration by parts we can show that

$$\int_t^1 \int_t^1 \gamma(u, v)\,dF^{-1}(u)\,dF^{-1}(v) < \infty \tag{3.12}$$

where $\gamma(u, v) = u(1-v)$, for $0 \leqq u \leqq v \leqq 1$, is the covariance function of $W(t)$. Condition (3.12) implies that the integral $\int_t^1 W(u)\,dF^{-1}(u)$ exists in quadratic mean and the $\operatorname{Var}[\int_t^1 W(u)\,dF^{-1}(u)] = (3.12)$, explicitly

$$\begin{aligned} \operatorname{Var}(\textstyle\int_t^1 W(u)\,dF^{-1}(u)) &= \textstyle\int_t^1 (F^{-1}(u))^2\,du - (\int_t^1 F^{-1}(u)\,du)^2 \\ &\quad + t(1-t)(F^{-1}(t))^2 - 2tF^{-1}(t)\textstyle\int_t^1 F^{-1}(u)\,du\,. \end{aligned}$$

We can then verify directly that the normal process on the right-hand side of (3.11) has the same covariance function as $\Gamma(s, t)$ given in (3.2) of Theorem 1.

REFERENCES

[1] BARLOW, R. E., BARTHOLOMEW, D. J., BREMNER, J. M. and BRUNK, H. D. (1972). *Statistical Inference Under Order Restrictions.* Wiley, New York.

[2] BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[3] CHIANG, C. L. (1960). A stochastic study of the life table and its applications: I. Probability distributions of the biometric functions. *Biometrics* **16** 618–635.

[4] CHIANG, C. L. (1968). *Introduction of Stochastic Process in Biostatistics.* Wiley, New York.

[5] LE CAM, L. (1957). *Convergence in Distribution of Stochastic Processes.* Univ. of Calif. Publ. in Statist. **2**, No. 11, 207–236. Univ. of California Press.

[6] MARSHALL, A. W. and PROSCHAN, E. (1965). Maximum likelihood estimation for distribution with monotone failure rate. *Ann. Math. Statist.* **36** 69–77.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND 20742