GLOBAL JOURNAL OF MATHEMATICAL SCIENCES, VOL 2, NO. 1, 2003: 81-88
 ISSN 1596-6208

THE PROBABILITY DENSITY FUNCTION OF COMPLETED LENGTH OF SERVICE (CLS) DISTRIBUTION AND WASTAGE FUNCTION FOR SOME SECONDARY SCHOOLS IN ENUGU STATE

PETER I. UCHE and EVARESTUS O. OSSAI

(Received 18 March 2003; Revision accepted 18 May, 2003).

ABSTRACT

By investigating the existing relationships between the probability density function of CLS distribution and some other wastage functions this paper estimates the functions for some secondary schools in Enugu State. Wastage probabilities are calculated, survivor functions estimated. The accompanying standard errors are also obtained.

KEY WORDS: Manpower Planning, Length of Service, Modelling, Survivor Functions.

INTRODUCTION

In the study of certain aspects of manpower planning the length of service is considered above other rival factors such as age, sex, marital status, that determine the grading or classification of manpower systems, Bartholomew, et al (1991). This fact may intuitively be accepted if we consider that length of service phenomenon is compulsorily inherent in every dynamic manpower system.

Length of service distribution can be used in manpower planning in a number of ways. In the area of modelling, Ugwuowo and McClean (2000) have used models of length of service distribution in studying heterogeneity in manpower system. Some others, such as McClean, et al (1998), have modelled manpower system by the use of Markov processes with length of service dependent states. In these and other works the length of service distribution is used for prediction and for explanation of the underlying leaving process.

Perhaps, an aspect of manpower planning that has attracted a lot of interest, as seen in works like Bartholomew, et al (1979), is the configuration of wastage patterns or leaving experiences of employees. Strategic in this pursuit is the incorporation of length of service as a mainstream factor affecting propensity of workers to leave. Various researchers like Silcock (1954), Lane and Andrew (1955), Bartholomew (1973), have shown that propensity to leave decreases with increasing length of service. Also Forbes, (1971), McClean and Gribbin (1987) and Gribbin (1989), among others, have identified completed length of service (CLS) as the most important duration of interest in manpower planning. In the study of wastage in manpower planning two approaches have mostly been identified. One is through cohort analysis where a homogeneous group of employees entering at the same or about the same time is investigated while the second approach, the census method, is used where there exist several groups of different 'ages', and it is desired to reconstruct a composite picture of the wastage functions based only on the current leaving experiences of employees of the system, Bartholomew, et al. (1991). This study on teaching staff in some secondary schools utilizes the argument of the second approach. When we group teachers based on CLS, we then have the CLS distribution of the teachers.

PETER I. UCHE, Dept. of Statistics, University of Nigeria, Nsukka, Enugu State, Nigeria.
EVARESTUS O. OSSAI, Dept. of Statistics, University of Nigeria, Nsukka, Enugu State, Nigeria.

CLS Distribution and Other Wastage Functions

Consider a system with $[x_i, x_{i+1})$ as the ith length of service interval; L_i as the corresponding number of leavers in $[x_i, x_{i+1})$, and S_i as the average number in service in this interval. Also, according to Bartholomew, et al. (1979, 1991) we define the following probability functions and quantities and obtain the results that follow:

f(x) = Probability density function of CLS; F(x) = the distribution function of CLS; G(x) = the survivor function: the probability that an individual survives in the system for time x.

$\bar{D}(x)$ = the expected further duration; m = central wastage rate: number of leavers during the census period who were in the class when they left divided by the average number in this class during the census period.

From the above definitions, the following are immediate:

$$F(x) = 1 - G(x)$$

$$f(x) = \frac{d(1-G(x))}{dx} = \frac{-d\,G(x)}{dx}$$

or $$G(x) = \int_x^{\infty} f(y)dy \quad \ldots\ldots(1)$$

Also, since we can define f(x) as probability of leaving in a small interval per unit time.

$$f(x) = \lim_{\Delta x \to 0} \frac{P\{\text{an employee leaving in } [x, x+\Delta x)\}}{\Delta x}$$

and $$m(x) = \lim_{\Delta x \to 0} P\left\{\frac{\text{an employee with length of service x leaves in } [x, x+\Delta x)}{\Delta x}\right\}$$

then,

$$\Delta x f(x) = P\{\text{leaving in } [x, x+\Delta x)/\text{survival to } x\}.P\{\text{survival to } x\} = m(x).\Delta x.G(x)$$

Therefore, $$m(x) = \frac{f(x)}{G(x)} = \left(\left(\frac{d\,G(x)}{dx}\right)\right)\Big/ G(x) = \frac{-G'(x)}{G(x)} = -\frac{d\ln G(x)}{dx} \quad \ldots\ldots(2)$$

where $G'(x) = \frac{d\,G(x)}{dx}$

Equation (2) yields

$$G(x) = \exp[-H(x)] \quad \ldots\ldots\ldots\ldots\ldots (3)$$

where $H(x) = \int_0^x m(y)\,dy$ may be seen as the cumulative wastage rate function.

Finally, we have that $f(x) = m(x)\exp[-H(x)] \ldots\ldots\ldots\ldots (4)$

Equation (4) expresses the probability density function of CLS in terms of the wastage function, m (x), alone. Actually, m(x) tells about the propensity of employees to leave. The above result or similar results were also obtained by Bartholomew et al. (1979,1991) and Lee, E. T. (1992).

Estimation of the Functions

To apply the above relationships represented in continuous time to grouped data we redefine the estimates of the functions as follows, for i = 1, 2, … , k

f_i = P {entrant leaves during a unit interval within $[x_i, x_{i+1})$}; G_i = P {an entrant survives to x_i}; D_i = expected remaining service for an employee with length of service x;

m_i = step function approximation of m (x) for x within $[x_i, x_{i+1})$.

$\hat{m}_i = \frac{L_i}{S_i}$; an estimator of m_i

Following the stability curve method of Lane and Andrew (1955)

a. H(x) can be seen as a cumulative estimate of m (x) for various groups from o up to x_i and then from x_i to x in the last interval $[x_i, x_{i+1})$ to obtain:

$$\hat{H}(x) = \sum_{j=0}^{i-1} c_j\hat{m}_j + (x - x_i)\,\hat{m}_i \quad \ldots\ldots\ldots\ldots\ldots (5)$$

b. From (3) and (5), an estimator of the survivor function can be obtained as:

$$\hat{G}(x) = \exp[-\sum_{j=0}^{i-1} c_j\,\hat{m}_j - (x - x_i)\,\hat{m}_i\,] \quad \ldots\ldots\ldots\ldots\ldots (6)$$

where $c_j = x_{j+1} - x_j$

c. According to Bartholomew, et al (1979, 1991), we define

$$f_i = \int_{x_i}^{x_{i+1}} \frac{f(x)}{x_{i+1} - x_i}$$

The estimator of f_i is then obtained as

$$\hat{f}_i = \frac{1}{c_i} \int_{x_i}^{x_{i+1}} \hat{m}_i \exp\left(-\sum_{j=0}^{i-1} c_j \hat{m}_j - (x - x_i)\hat{m}_i\right) dx.$$

$$= \frac{\exp\left(-\sum_{j=0}^{i-1} c_j \hat{m}_j\right) - \exp\left[-\sum_{j=0}^{i-1} c_j \hat{m}_j (x_{i+1} - x_i)\hat{m}_i\right]}{c_i} = \frac{\hat{G}(x_i) - \hat{G}(x_{i+1})}{c_i}$$

$$= \frac{\hat{G}_i - \hat{G}_{i+1}}{c_i} \qquad \text{.........................(7)}$$

d. It has also been shown by Bartholomew, et al. (1979, 1991), that by defining the expected remaining service for an employee with length of service x as

$D(x) = \int_x^{\infty} \frac{uf(u)\,du}{G(x)}$, an estimator of D_i is

$$\hat{D}i = \hat{G}_i^{-1}\left(\sum_{j=1}^{k-1} (\hat{G}_j - \hat{G}_{j+1})/\hat{m}_j + {}^1/_2\, C_k \hat{G}_k\right] \qquad \text{.....................(8)}$$

for i= 0. 1, 2, …, k - 1. Where D_i estimates $D(x)$ at the points x_i.

Estimates of the Standard Errors of $\hat{m}_i$ and $\hat{G}_i$

a. Following the result from Forbes (1971), the number of leavers in $[x_i, x_{i+1})$, L_i, is approximately poisson with $L_i \sim$ poisson $(m_i S_i)$.

This means that the Minimum Variance Unbiased Estimator for m_i is $\hat{m}_i = \frac{L_i}{S_i}$

Assuming S_i to be constant over $[x_i, x_{i+1})$ throughout the census period, then

$$Var(\hat{m}_i) = \frac{Var(L_i)}{S_i^2} = \frac{m_i S_i}{S_i^2} = \frac{m_i}{S_i^2}$$

Hence, an estimate of the standard error can be given by $Se(\hat{m}i) = \sqrt{\frac{\hat{m}_i}{S_i}}$

b. An approximate result for the variance of a function of random variables, also given in Forbes (1971), is as follows.

"If x and y are random variables with x~(mean = μ, variance = σ^2), y~(μ, r^2), and correlation coefficient ρ, and if Z is a function of x and y such that Z = θ(x, y), then for sufficiently small δ and r and well - behaved θ.

$$\text{Variance } (Z) \approx \delta^2 \left(\frac{\partial\theta}{\partial x}\right)^2 + 2\rho\delta\left(\frac{\partial\theta}{\partial x}\frac{\partial\theta}{\partial y}\right) + r^2\left(\frac{\partial\theta}{\partial y}\right)^2$$

where all the partial differentials are evaluated at x = μ, y = μ"

By extending the above results to a function of more than two variables and recalling that

(i) $\hat{G}i = \exp\left[-\sum_{j=0}^{i-1} \hat{m}_j c_j\right] = e^{-\hat{m}_0 c_0} \cdot e^{-\hat{m}_1 c_1} \cdots e^{-\hat{m}_{i-1} c_{i-1}}$

(ii) $\text{Var}(\hat{m}_j) = \dfrac{\hat{m}_j}{S_j}$

Then for m_j, m_k independent,

$$\text{Var}(\hat{G}_i) \approx \sum_{j=0}^{i-1} \text{var}(\hat{m}_j) \cdot \left(\frac{\partial\hat{G}_i}{\partial\hat{m}_j}\right)^2 = \sum_{j=0}^{i-1} \text{var}(\hat{m}_j) \cdot c_j^2 \cdot \hat{G}_i^2 = \hat{G}_i^2 \sum_{j=0}^{i-1} c_j^2 \frac{\hat{m}_j}{S_j}$$

Hence, an estimate of the standard error of $\hat{G}_i$ can be approximated as

$$Se(\hat{G}_i) \approx \hat{G}_i \sqrt{\sum_{j=0}^{i-1} \frac{c_j^2 \hat{m}_j}{S_j}}$$

APPLICATION

We applied the preceding work on the data of a sample of five secondary schools in Enugu State. Information was collected, from the schools' records, about years of first appointment and date of leaving as well as reason for leaving, on every teaching staff that has been a member of the five schools from the date of inception till 1994. Census method involves a record of the stocks and flows (leavers); in this case of teachers, in an interval of time. This interval may be a year, but in this study we have used two years – 1993 to 1994, in conformity with Bartholomew, et al (1991), due to the observed low wastage rates. In transition probability studies two consecutive years may be used The use of the above interval means that we consider leavers as teachers who even though have started work in some previous years or in the period 1993 to 1994 left in the accounting period 1993 to 1994. The data are as shown in Tables 1 and 2. The choice of our length of service interval follows the prescription for satisfactory length of service groups by Lane and Andrew (1955), which is to suit the structure of the CLS curve which has its critical part as the beginning portion. This needs fine groupings, while the end portion of the curve with slower changes needs more spaced intervals.

For Table 2, we define the following

$$L_i = L_{i1} + L_{i2};\ S_{i2} = S_{i1} + S_{i2}^{*} - L_{i1} - L_{i2};\ S_i = (S_{i1} + S_{i2})/2$$

where

S_{i1} = number in the class $[x_i, x_{i+1})$ at the beginning of the census period;

S_{i2}^{*} = number of entrants during the course of the census period; S_{i2} = number in the class at the end of the census period; S_i = average number in the class during the census period;

TABLE 1: Number of Teachers Employed and Number that left in each Class Interval (During the Census Period 1993 – 1994)

Length of service interval (in years)	Data considered in 1993 (beginning of census period)			Data considered in 1994 (During the course and end of census period)		
	Year of employment	Number employed	Number who left in 1993	Year of employment	Number employed including those that remained after 1993	Number who left in 1994
0 – 1	93	8	2	94	0	0
1 – 2	92	25	1	93	6	0
2 – 3	91	2	0	92	24	4
3 – 4	90	14	0	91	3	1
4 – 5	89	1	0	90	14	1
5 – 8	88, 87, 86	21	1	89, 88, 87	18	3
8 – 10	85, 84	5	1	86, 85	8	0
10 – 15	83, 82, 81, 80, 79	28	1	84, 83, 82, 81, 80	20	3
15 – 35	78, 77, 76, …, 64	19	2	79, 78, 77, …, 65	24	0

(Note: Under year of employment, 93, for instance, means 1993)

TABLE 2: Stock and Flow Data and the Estimates of m_i for the Secondary School Academic Manpower System

$x_i - x_{i+1}$	L_{i1}	L_{i2}	L_i	S_{i1}	S_{i2}	S_{i2}	$\hat{S}_i$	$\hat{m}_i$	Se(m_i)
0 – 1	2	0	2	8*	0	6	7.0	0.286	0.202
1 – 2	1	0	1	25	6	30	27.5	0.036	0.036
2 – 3	0	4	4	2	24	22	12.0	0.333	0.167
3 – 4	0	1	1	14	3	16	15.0	0.067	0.067
4 – 5	0	1	1	1	14	14	7.5	0.133	0.133
5 – 8	1	3	4	21	18	35	28.0	0.143	0.071
8 – 10	1	0	1	5	8	12	8.5	0.118	0.118
10 – 15	1	3	4	28	20	44	36.0	0.111	0.056
15 – 35	2	0	2	19	24	41	30.0	0.054	0.042

Table 3: Other Estimated Wastage Functions for the Secondary School Academic Manpower Data.

x_i	$\hat{G}_i$	$Se(\hat{G}_i)$	f_i	$\hat{D}_i$
0	1.000	0.000	0.249	7.1
1	0.751	0.152	0.026	8.2
2	0.725	0.149	0.206	7.5
3	0.519	0.137	0.033	9.3
4	0.486	0.133	0.061	9.0
5	0.425	0.129	0.049	9.2
8	0.277	0.103	0.029	10.3
10	0.219	0.096	0.019	10.8
15	0.126	0.066	0.004	12.2
35	0.043	0.043	-	0

L_{i1} = number of leavers from S_{i1}; L_{i2} = number of leavers from* S_{i2} ; L_i = number of leavers during th census period who were in the class when they left.

Tables 1 and 2 illustrate the above account. Table 2 gives the stock and flow data and the estimate of the central wastage rate m_i. From the table, m_i can be considered to be generally low even though n standard rates for comparison exists. Its values can be seen to alternate in the first five intervals but afte decrease as length of service increases. It is expected that the longer people stay in employment th propensity to leave decreases.

Table 3 gives the estimates of the other wastage probabilities and functions, all calculated the point $x = x_i$. It could be observed from the estimates of D_i, especially in the longer length of servic intervals, that the expected further duration of teachers increases as they remain longer in the service b decreases to zero as retirement is approached. For instance, a teacher who has spent four years is expected stay a further period of 9 years while some one with five years is expected to stay 9.2 years more. T estimates of G_i and f_i all conform to the earlier observation about m_i, decreasing as length of servi increases.

CONCLUSION

We have been able to show the relationships existing between CLS and some other wastage probabilities and functions. It is easily seen that with very little mathematical manipulation one can go from one function to another.

The application has shown that the teachers' urge to leave decreases as length of service increases, especially, in the lower length of service intervals. However the wastage rate is observably low.

ACKNOWLEDGEMENT

This work was sponsored by University of Nigeria Senate Research Grants Committee.

REFERENCES

Bartholomew, D. J., 1973. Stochastic Models For Social Processes (2nd ed.), Wiley, New York.

Bartholomew, D. J. and A. F. Forbes 1979. Statistical Techniques for Manpower Planning, John Wiley, New York.

Bartholomew, D. J., Forbes, A.F. and McClean. S. I., 1991. Statistical Techniques for Manpower Planning. Wiley, New York.

Forbes, A. F. 1971. Non-Parametric Methods of Estimating the Survivor Function. Statistician,. 20(1): 27- 52.

Gribbin, J. O., 1989. Modelling the Nurse Manpower System. Ph.D. Thesis, University of Ulster, Coleraine, BT52 ISA, Northern Ireland, U.K.

Lane, K. F. and Andrew , J. E., 1955. A Method of Labour Turnover Analysis". Royal Statistical Society, pp. 296 – 322.

Lee, E. T., 1992. Statistical Methods for Survival Data Analysis. John Wily, New York.

McClean, S. I. and Gribbin, J. O.,1987. Estimation for Incomplete Manpower Data. Applied Stochastic Models and Data Analysis, 3: 13 – 25.

McClean, S. I., Montgomery, E. and Ugwuowo, F. I., 1998. Non-Homogeneous Continuous –Time Markov and Semi-Markov Manpower Models. Appl. Stochastic Models and Data Analysis. 13, 191 – 198.

Silcock, H., 1954. The Phenomenon of Labour Turnover. J. Roy. Statist. Soc. A 117: 429 – 440.

Ugwuowo, F. I. and McClean, S. I., 2000. Modelling Heterogeneity in a Manpower System: A Review. Appl. Stochastic Models in Bus. and Ind., 16: 99 – 110.