Extending the Piecewise Exponential Estimator of the Survival Function

Ganesh B Malla
Department of Mathematics & Statistics, Fairmount College of Liberal Arts and Sciences

Abstract – We propose a new way of extending the piecewise exponential estimator (PEXE) beyond the last observation, where the original PEXE and the Kaplan-Meier estimator (KME) of the true survival function (SF) are undefined. We propose additional exponential tail to the original PEXE so that the estimate of the true mean by our estimator, which we will call extended-PEXE (EPEXE), and the KME are equal. By simulations, with various survival functions, we have been able to show that EPEXE improves on both the original PEXE and the KME.

1. Introduction - The survival function of a life time, T, is defined by S (t) = P (T > t) for t ≥ 0. The common estimator of S is the empirical survival function which is defined as the proportion of n subjects in the study surviving time t. This is a step function. The KME, S_n (t) that utilizes randomly selected right-censored data is by far the most popular estimator of S. This too is a step function that jumps only at the points where the observations are uncensored. In case of no censoring, it reduces to the empirical survival function.
Proschan and others have developed a method of estimation that is piecewise exponential between successive jump points of the estimator is of the form S (t) = C exp (-rt) for some constants C and r. This estimator, called PEXE, is continuous, as it should be for a life distribution. However, from the method of construction, the PEXE is undefined beyond the last score, X_n. We propose a method of estimating the SF beyond X_n.

2. Estimation- Since the PEXE is piecewise exponential and is undefined on $[X_{n,} \infty)$, it suggests adding an exponential tail to PEXE, S^*_n so that the PEXE becomes EPEXE, S^{**}_n which is well–defined on $[0, \infty)$. The problem is what hazard rate to use in the tail. We propose a hazard rate in the exponential tail by requiring that the mean of the resulting SF be equal to the mean of the KME of the SF. This is feasible as we have proven

$$\Delta \equiv \int_0^{X_n} \left(S_n(t) - S_n^*(t) \right) dt > 0 \quad . \qquad (2.1)$$

Thus, S^{**}_n (t) = S^*_n (t) when $0 \le t \le X_n$ and for $t > X_n$ the estimate S^{**}_n is given by

$$S_n^{**}(t) = S_n^*(X_n) \ \exp[-\frac{S_n^*(X_n)(t - X_n)}{\Delta}] \ . \qquad (2.2)$$

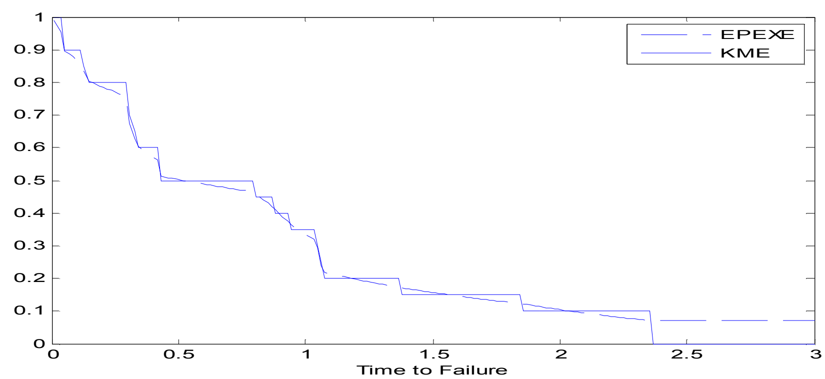
3. Simulation - In our simulation study we use the Weibull distribution with scale parameter α and shape parameter β as our life distribution. Throughout the study, we chose various values of β so that our comparison could be valid for both IFR and DFR Survival models. The value of α is chosen to be 1, but can be chosen arbitrarily since the estimators under comparison are scale invariant. The censoring distribution is the exponential distribution with parameter λ. By choosing the suitable value of λ, we can get different proportions of censoring in each simulation. For comparison purposes, we calculate and present the mean sums of squares (MSE) ratio R≡ MSE (KME)/ MSE (EPEXE) and biases of the estimators in the Table 3.1 and plot the two estimators in a single graph in figure 3.

Table 3.1 The MSE Ratio and Biases of the EPEXE and the KME , a Sample Size of 50, # of iterations = 10, 000

Life distribution	Quintiles	R			Bias(EPEXE)			Bias(KME)		
		cp = 0	cp = 50	cp = 80	cp = 0	cp = 50	cp=80	cp = 0	cp = 50	cp=80
Weibull (0.5, 1)	.10	1. 000	1.000	1.028	-0.0040	-0.0026	-0.0013	-0.0041	-0.0039	-0.0018
	.25	1.000	1.022	1.036	-0.0001	-0.0019	-0.0015	-0.0011	-0.0023	-0.0071
	.50	1.020	1.025	1.039	0.0001	-0.0177	-0.0410	0.0001	-0.0202	-0.0876
	.75	1.025	1.024	1.038	0.0007	-0.0032	-0.0943	0.0009	-0.0048	-0.0196
	.90	1.029	1.026	1.042	0.0007	-0.0103	-0.0497	0.0008	-0.0107	-0.0300
	.95	1.028	1.028	1.044	0.0002	-0.0014	-0.0300	0.0010	-0.0079	-0.0100
	.99	1.033	1.023	1.042	0.0001	-0.0015	-0.0100	-0.0006	-0.0080	-0.0102
Weibull (1, 1)	.10	1.000	1.001	1.000	-0.0001	-0.0022	-0.0012	0.0003	-0.0033	-0.0015
	.25	1.011	1.003	1.001	-0.0008	-0.0036	-0.0016	0.0009	-0.0043	-0.0077
	.50	1.015	1.090	1.004	-0.0007	-0.0110	-0.0012	0.0066	-0.0125	-0.0128
	.75	1.016	1.011	1.010	0.0033	-0.0039	-0.0103	0.0289	-0.0046	-0.0210
	.90	1.019	1.017	1.010	0.0033	-0.0102	-0.0012	0.0590	-0.0114	-0.0222
	.95	1.020	1.012	1.016	0.0034	-0.0049	-0.0110	0.0654	-0.0067	-0.0320
	.99	1.024	1.018	1.022	0.0192	-0.0066	-0.0111	0.0343	-0.0082	-0.0132
Weibull (1.5, 1)	.10	1.000	1.011	1.001	-0.0021	-0.0027	-0.0011	-0.0039	-0.0041	-0.0014
	.25	1.001	1.009	1.005	-0.0002	0.0034	-0.0023	-0.0015	-0.0039	-0.0072
	.50	1.015	1.009	1.009	0.0003	-.0056	-0.0431	0.0009	-0.0222	-0.0865
	.75	1.006	1.010	1.007	0.0010	-0.0033	-0.0152	0.0011	-0.0050	-0.0469
	.90	1.011	1.012	1.010	0.0010	-0.0223	-0.0341	0.0022	-0.0317	-0.0450
	.95	1.013	1.013	1.014	0.0002	-0.0077	-0.0230	0.0012	-0.0080	-0.0410
	.99	1.014	1.015	1.017	0.0009	-0.0065	-0.0200	-0.0015	-0.0081	-0.0122

Notes: cp = Censoring Proportion, R ≡ MSE (KME) / MSE (EPEXE)

Figure3.1 The EPEXE and the KME of the SF of the Weibull (β=2, α=1,) distribution with α = 0.749 (cp =50%).



4. Conclusions- We note that the PEXE and the EPEXE are the same as long as t $\in$ [0, X_n]. The PEXE has several advantages over the KME for estimating the SF, a complete description of it can be found in [1]. The simulation results show that the MSE ratio is uniformly ≥ 1 for the estimation of the various SF with different sample sizes and the censoring proportions. The absolute bias of the EPEXE is consistently smaller than that of the KME. The EPEXE provides a well-defined, smooth and continuous estimator of the SF even beyond the last observation and retains all the advantages of the PEXE over the KME, which makes it more appealing and realistic to the users.

References

[1] Frank Proschan, Jee Kim, "Piecewise Exponential estimator of the Survival Function", IEEE Trans. on Reliability, vol.40, 1989, pp 134-139.
[2] David M Reineke, Edward A Pohl, W. Paul Murdock Jr., "Maintenance-Policy Cost-Analysis for a Series System with Highly-Censored Data", IEEE Transactions on Reliability, Vol. 48, pp 413 - 420.
[3] Ulf Westberg, Bengt Klefsjö, "TTT-Plotting for Censored Data Based on the PEXE", International Journal of Reliability, Quality and Safety Engineering, Vol.1, No.1, (1994) pp 1-13.