

“Empirical Estimation of Risk Measures and Related Quantities,” Bruce L. Jones and Ricārdas Zitikis, October 2003

VYTARAS BRAZAUSKAS* AND
THOMAS KAISER†

We congratulate Dr. Jones and Dr. Zitikis for making a very valuable contribution to an important area of actuarial research, which involves the price determination of an insurance risk. In this article the authors introduced a unifying representation of risk measures originally defined in terms of expectations with respect to distorted probabilities, and they proposed a general (and intuitively appealing) approach for estimation of these quantities. In particular, using asymptotic theory for L -statistics, they have developed the empirical nonparametric estimators for these risk measures.

In this discussion we examine finite-sample performance of the proposed (asymptotic) confidence intervals. Using Monte Carlo simulations, we investigate the following questions:

- How fast do these intervals attain the intended confidence level?
- How much does one gain/lose if, instead of empirical intervals, parametric confidence intervals based on maximum likelihood estimators are used?
- How bad are the consequences if one of the conditions necessary for the asymptotic normality of the empirical estimator to hold is ignored?

1. INTRODUCTION

We consider a risk measure based on the proportional hazard transform (PHT) and investigate its interval estimation problem when data are gener-

ated by three similar shape parametric families that in addition have equal PHT measures. Confidence intervals are constructed by applying (1) the empirical approach proposed by Jones and Zitikis (2003), and (2) standard asymptotic theory for the maximum likelihood estimators (MLEs). From a practical standpoint it is of interest to see how accurately we can estimate the PHT measure when the underlying scenarios of data generation are very similar (but not identical).

The PHT measure is given by formula (2) of Jones and Zitikis (2003),

$$\text{PHT}(F, r) = \int_0^\infty [1 - F(x)]^r dx.$$

In this formula we choose $r = 0.55, 0.70, 0.85, 0.95$ and the following three families for F :

- *Exponential* with the cdf given by $F_1(x) = 1 - e^{-(x-x_0)/\theta}$, for $x > x_0$ and $\theta > 0$
- *Pareto* with the cdf given by $F_2(x) = 1 - (x_0/x)^\gamma$, for $x > x_0$ and $\gamma > 0$
- *Lognormal* with the cdf given by $F_3(x) = \Phi(\log(x - x_0) - \mu)$, for $x > x_0$ and $-\infty < \mu < \infty$, where $\Phi(\cdot)$ denotes the cdf of the standard normal distribution.

Here the parameter x_0 can be interpreted as a deductible or a retention level and, thus, assumed to be known. (Note that, due to x_0 , distributions F_1, F_2 , and F_3 have the same support.) The remaining parameters γ, θ , and μ are unknown, and we must estimate them from the data. For the purposes of simulation, they are chosen so that, except for the choice $r = 0.55$ and F_2 , for all other combinations of r and F the assumptions of Theorem 3.2 (see Jones and Zitikis 2003) are satisfied, and thus the empirical estimator of the PHT measure is asymptotically normal. The requirement that all three distributions possess the same PHT measure, that is, $\text{PHT}(F_1, r) = \text{PHT}(F_2, r) = \text{PHT}(F_3, r)$, is equivalent to

$$x_0 + \frac{\theta}{r} = x_0 + \frac{x_0}{\gamma r - 1} = x_0 + C_r e^\mu, \quad (1)$$

where, for fixed r , the integral $C_r = \int_{-\infty}^\infty [1 - \Phi(z)]^r e^z dz$ is found numerically. (For example,

* Vytautas Brazauskas is Assistant Professor, Department of Mathematical Sciences, University of Wisconsin–Milwaukee, P.O. Box 413, Milwaukee, Wisconsin 53201, e-mail: vytautas@uwm.edu.

† Thomas Kaiser is a graduate teaching assistant, Department of Mathematical Sciences, University of Wisconsin–Milwaukee, P.O. Box 413, Milwaukee, Wisconsin 53201, e-mail: kaisert@uwm.edu.

for the values of r considered here, one finds $C_{0.55} = 3.896$, $C_{0.70} = 2.665$, $C_{0.85} = 2.030$, and $C_{0.95} = 1.758$.) Thus, taking into consideration equation (1) and the above discussion, we select

$$x_0 = 1, \quad \gamma = 5.5, \quad \theta = \frac{r}{5.5r - 1},$$

$$\mu = -\log[C_r(5.5r - 1)]. \quad (2)$$

Further, as follows from formulas (10), (18), (19), and (28) of Jones and Zitikis (2003), the $100(1 - \alpha)$ percent confidence interval, based on the empirical estimator of the PHT measure, is given by

$$L_n[X] \pm z_{\alpha/2} \sqrt{\frac{Q_n(\psi, \psi)}{n}},$$

where

$$Q_n(\psi, \psi) = \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} c_n(i, j) \psi(i/n) \\ \times \psi(j/n) (X_{i+1:n} - X_{i:n})(X_{j+1:n} - X_{j:n})$$

with $c_n(i, j) = \min\{i/n, j/n\} - (i/n)(j/n)$ and $\psi(s) = r(1 - s)^{r-1}$,

$$L_n[X] = \sum_{i=1}^n c_{in} X_{i:n}$$

with $c_{in} = (1 - (i - 1)/n)^r - (1 - (i/n))^r$, and $z_{\alpha/2}$ is the $\alpha/2$ -critical value of the standard normal distribution, and $X_{1:n} \leq \dots \leq X_{n:n}$ denote the ordered values of the sample X_1, \dots, X_n .

Furthermore, as follows from asymptotic theory for the maximum likelihood procedures, the $100(1 - \alpha)$ percent confidence intervals for the parameters θ , γ , and μ are given by

$$\hat{\theta}(1 \pm z_{\alpha/2} \sqrt{1/n}), \quad \hat{\gamma}(1 \pm z_{\alpha/2} \sqrt{1/n}), \\ \hat{\mu} \pm z_{\alpha/2} \sqrt{1/n}, \quad (3)$$

where

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n (X_i - x_0), \quad \hat{\gamma} = \left[\frac{1}{n} \sum_{i=1}^n \log(X_i/x_0) \right]^{-1}, \\ \text{and } \hat{\mu} = \frac{1}{n} \sum_{i=1}^n \log(X_i - x_0)$$

are the MLEs of θ , γ , and μ , respectively. Finally, to get the parametric confidence intervals for the

PHT measure, one just has to transform the end points of each interval in equation (3) according to the corresponding formula of equation (1) and keep in mind that x_0 is known.

2. SIMULATIONS

We used the following design for the Monte Carlo simulation study. Ten thousand samples of size n were generated from a distribution F . For each sample, a $(1 - \alpha)$ level confidence interval for the PHT measure was constructed using the empirical and parametric approaches. Then, based on these 10,000 intervals for each approach, the average length of the interval and the proportion of times the interval covers the true value of the PHT measure were evaluated. This procedure was repeated 10 times, and the means and standard errors of the average length and the proportion of coverage were recorded. The study was performed for the following (specific) choices of parameters:

- Sample sizes: $n = 100, 250, 500, 1,000$
- Confidence levels: $1 - \alpha = 0.90, 0.95, 0.99$
- $r = 0.55, 0.70, 0.85, 0.95$
- Distribution functions: exponential (F_1), Pareto (F_2), and lognormal (F_3), where parameters x_0 , γ , θ , and μ are chosen according to equation (2)
- The true values of the PHT measure, $\text{PHT}(F, r) = x_0 + x_0/(\gamma r - 1)$: $\text{PHT}(F, 0.55) = 1.494$, $\text{PHT}(F, 0.70) = 1.351$, $\text{PHT}(F, 0.85) = 1.272$, $\text{PHT}(F, 0.95) = 1.237$.

In Table 1 we report the average *length* (denoted L) and the average proportion of *coverage* (denoted C) for the 90% empirical and parametric confidence intervals. For other choices of confidence level $1 - \alpha$ patterns are similar. We graphically summarize patterns for C in Figure 1; those for L are easily predictable, and we do not present them here.

DISCUSSION OF TABLE 1

For all n and r , the proportions of coverage of the parametric intervals attain the nominal level of 0.90. Except for a few cases at the exponential model, the proportions of coverage of the empirical intervals do not attain the nominal level even for sample sizes as large as 1,000. For $n = 100$ at the exponential model and for $n \leq 250$ at the Pareto model, parametric intervals have the same

Table 1
Performance of 90% Confidence Intervals for PHT(F, r), for Selected F, r , and n

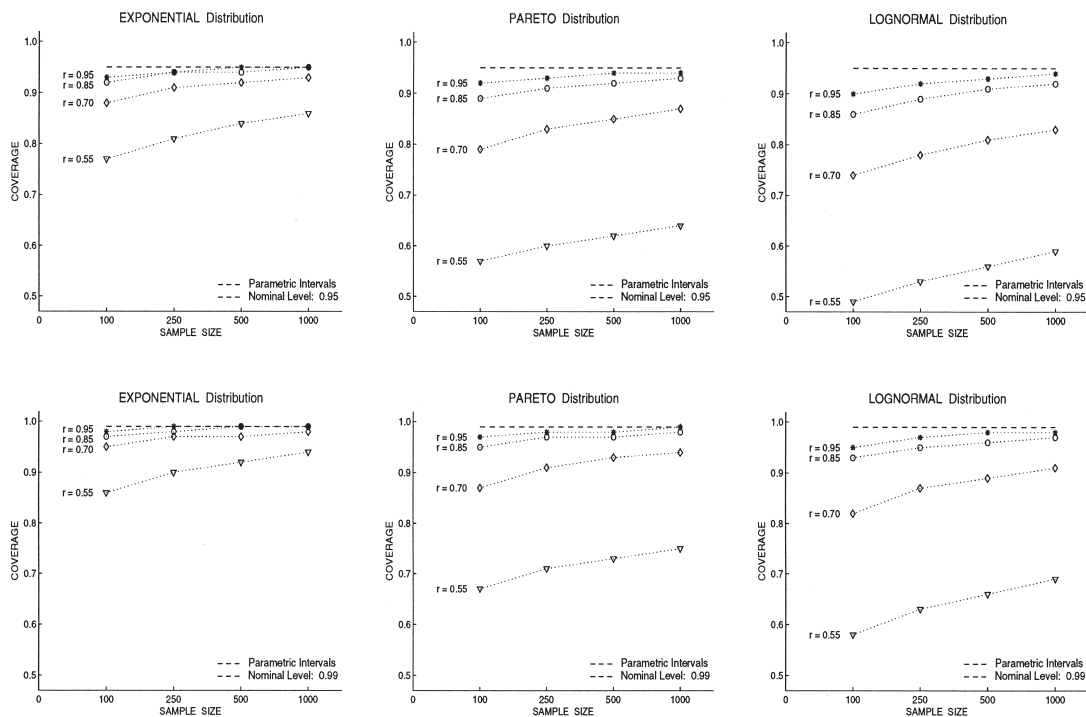
n	r	Exponential (F_1)				Pareto (F_2)				Lognormal (F_3)			
		Empirical		Parametric		Empirical		Parametric		Empirical		Parametric	
		L	C	L	C	L	C	L	C	L	C	L	C
100	0.55	.14	.70	.16	.90	.18	.50	.26	.90	.19	.44	.16	.90
	0.70	.11	.82	.12	.90	.14	.73	.16	.90	.15	.68	.12	.90
	0.85	.09	.87	.09	.90	.11	.83	.12	.90	.12	.81	.09	.90
250	0.95	.08	.88	.08	.90	.09	.87	.10	.90	.10	.85	.08	.90
	0.55	.10	.75	.10	.90	.14	.54	.16	.90	.15	.47	.10	.90
	0.70	.07	.85	.07	.90	.10	.77	.10	.90	.11	.72	.07	.90
500	0.85	.06	.89	.06	.90	.07	.86	.07	.90	.08	.84	.06	.90
	0.95	.05	.90	.05	.90	.06	.88	.06	.90	.06	.87	.05	.90
	0.55	.07	.77	.07	.90	.11	.56	.11	.90	.13	.50	.07	.90
1,000	0.70	.05	.87	.05	.90	.07	.79	.07	.90	.09	.75	.05	.90
	0.85	.04	.89	.04	.90	.05	.87	.05	.90	.06	.86	.04	.90
	0.95	.03	.90	.03	.90	.04	.89	.04	.90	.05	.88	.03	.90
	0.55	.06	.79	.05	.90	.09	.57	.08	.90	.10	.52	.05	.90
	0.70	.04	.88	.04	.90	.06	.81	.05	.90	.07	.77	.04	.90
	0.85	.03	.90	.03	.90	.04	.88	.04	.90	.04	.87	.03	.90
	0.95	.02	.90	.02	.90	.03	.89	.03	.90	.03	.89	.02	.90

Note: Standard errors for all entries are $\leq .001$ (for L) and between $.002$ and $.007$ (for C).

length as (or are longer than) empirical ones, and thus their better coverage can be viewed as a trade-off between L and C . As n increases, how-

ever, they become as good as (or even better than) empirical intervals with respect to the L criterion and uniformly better with respect to C .

Figure 1
Proportions of Coverage of 95% and 99% Confidence Intervals for PHT(F, r) for Selected Distribution Function F , Sample Size n , and r



DISCUSSION OF FIGURE 1

Similarly to the 90% case (see Table 1), the proportions of coverage of the parametric intervals attain the nominal levels of 0.95 and 0.99, for all r and n . Convergence of the proportions of coverage of the empirical intervals is slow, but it improves as intervals become wider (i.e., as confidence level $1 - \alpha$ increases, for fixed n and r). For $\text{PHT}(F, r)$ with $r = 0.95$ and 0.85 , performances of the empirical intervals are very good at the exponential distribution and good at Pareto and lognormal families. For $\text{PHT}(F, r)$ with $r = 0.70$, empirical intervals perform reasonably well at the exponential distribution, but their performance significantly deteriorates at Pareto and lognormal families. For instance, for $n = 1,000$, the proportion of coverage of the 99% empirical interval at the lognormal model barely reaches 0.90. The case of $r = 0.55$ is extreme and very interesting. While at the Pareto model the empirical estimator of $\text{PHT}(F, r)$ is *not* asymptotically normal, and, thus, one can anticipate poor performance of the corresponding intervals, that performance is even worse at the lognormal model and still unsatisfactory at the exponential model.

3. CONCLUSIONS

Simulation results in Section 2 suggest the following answers to the questions raised at the beginning of this discussion:

- a. Convergence of the proportion of coverage of the empirical intervals is slow and depends on the value of r . For “large” r (e.g., $0.85 \leq r < 1$), the coverage levels of these intervals get reasonably close to the nominal level for $n \geq 500$ and for all distributions F that we considered. For $r \leq 0.70$, however, their performances are unacceptable even for $n = 1,000$. (Moreover, a smaller simulation study suggests that there is not much improvement even for $n = 1,500$.)
- b. Parametric intervals attain the intended confidence levels for all sample sizes under consideration, that is, for $n \geq 100$. (In additional studies we found that this is true even for $n = 50$.) For heavier-tailed distributions (Pareto, lognormal) or for $n > 250$, these intervals dominate empirical counterparts with respect to the length criterion too.
- c. Investigations of this question yielded somewhat puzzling results. On the one hand, the

choice of $r = 0.55$ and F_2 (Pareto) was introduced to check what happens if violations of certain theoretical assumptions occur. As expected, performance of the empirical intervals in this case was poor. On the other hand, although no violations of the assumptions for $r = 0.55$ and other choices of F were found, the empirical intervals performed similarly to the Pareto case. Even for such a “nice” distribution as an exponential one, their performance was still unsatisfactory.

Hence, to achieve the intended coverage levels in applications, we recommend that the proposed empirical intervals for $\text{PHT}(F, r)$ should be used for $r \geq 0.85$ and $n \geq 500$. In situations when one of these conditions is not satisfied, parametric intervals should be relied on, though these procedures are (typically) sensitive to model misspecifications. Finally, similar conclusions were also reached for the right-tail risk measure defined by formula (4) of Jones and Zitikis (2003).

REFERENCE

- JONES, BRUCE L., AND RICÁRDAS ZITIKIS. 2003. “Empirical Estimation of Risk Measures and Related Quantities,” *North American Actuarial Journal* 7(4): 44–54.

AUTHORS’ REPLY

We are truly delighted to see the thought-provoking analysis by Brazauskas and Kaiser (2004) of our recent results on empirical estimation of risk measures and related quantities.

In Jones and Zitikis (2003) we concentrate on developing a nonparametric approach for estimating risk measures and in this way contribute to filling the gap in the actuarial literature on the topic. Thus, in the literature we now find two competing and complementary approaches for analyzing risk measures: parametric and nonparametric. Given the choice of approaches, the researcher and, especially, the practitioner now face a dilemma: which approach should be preferred?

Sufficiently many examples in the statistical, actuarial, and econometric literature clearly prove that neither parametric nor nonparametric approaches work well in every situation and at