

Research Article

Data Transformation for Confidence Interval Improvement: An Application to the Estimation of Stress-Strength Model Reliability

Alessandro Barbiero

Department of Economics, Management and Quantitative Methods, Università degli Studi di Milano, Via Conservatorio, 7-20122 Milan, Italy

Correspondence should be addressed to Alessandro Barbiero; alessandro.barbiero@unimi.it

Received 31 July 2014; Revised 19 September 2014; Accepted 29 September 2014; Published 23 October 2014

Academic Editor: David Bulger

Copyright © 2014 Alessandro Barbiero. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In many statistical applications, it is often necessary to obtain an interval estimate for an unknown proportion or probability or, more generally, for a parameter whose natural space is the unit interval. The customary approximate two-sided confidence interval for such a parameter, based on some version of the central limit theorem, is known to be unsatisfactory when its true value is close to zero or one or when the sample size is small. A possible way to tackle this issue is the transformation of the data through a proper function that is able to make the approximation to the normal distribution less coarse. In this paper, we study the application of several of these transformations to the context of the estimation of the reliability parameter for stress-strength models, with a special focus on Poisson distribution. From this work, some practical hints emerge on which transformation may more efficiently improve standard confidence intervals in which scenarios.

1. Introduction

In many fields of applied statistics, it is often necessary to obtain an interval estimate for an unknown proportion or probability or, more generally, for a parameter whose natural space is the unit interval [1]. If p is the unknown parameter of a binomial distribution, the customary approximate two-sided confidence interval (CI) for p is known to be unsatisfactory when its true value is close to zero or one or when the sample size is small. In fact, estimation can cause difficulties because the variance of the corresponding point estimator is dependent on p itself and because its distribution can be skewed. A number of papers have been devoted to the development of more refined CIs for p (see, e.g., [1–4]). Here, we will consider the estimation of the probability $R = P(X < Y)$, where X and Y are two independent rv's. If Y represents the strength of a certain system and X the stress on it, R represents the probability that the strength overcomes the stress, and then the system works (R is then referred to as the “reliability” parameter). Such a statistical model is usually

called the “stress-strength model” and in the last decades has attracted much interest from various fields [5, 6], ranging from engineering to biostatistics. In these works, inferential issues have been dealt with, mainly in the parametric context. The problem of constructing interval estimators for R has been considered; when an exact analytical solution is not available, approximations based on the delta method and asymptotic normality of point estimators are carried out, some of them making use of some data transformation of the point estimate of R .

In this work, we concentrate on the case of Poisson-distributed stress and strength. Approximate large-sample CIs for the reliability R have already been built and assessed for different parameter and sample size configurations and have been proved to give satisfactory results unless R is close to 1 (or, symmetrically, 0) or the sample sizes are too small [7]. With the aim of improving the performance of such CIs, four transformation functions (logit, probit, arcsine root, and complementary log-log) are selected and applied to the maximum likelihood estimate of R , and the resulting CIs are

empirically compared in terms of coverage probability and expected length.

The paper is laid out as follows: in Section 2 a brief discussion on data transformations is presented, with a special focus on those ordinarily used in connection with the estimation of the reliability parameter of a stress-strength model. Section 3 introduces the stress-strength model with independent Poisson-distributed stress and strength, recalling the formulas for reliability, maximum likelihood estimator, and standard large-sample CI and introducing refinements for the latter based on data transformations. Section 4 is devoted to the Monte Carlo simulation study, which empirically assesses the statistical performance of CIs. An example of application on real data is provided in Section 5, and Section 6 concludes the paper with some final remarks.

2. Transformations and Application to the Estimation of a Stress-Strength Model

In almost all fields of research, one has to deal with data that are not normal. It is common practice to transform the nonnormal data at hand in order to exploit theoretical results that strictly hold only for the normal distribution, with the objective of building plausible or more efficient estimates. Citing [8], “transformations of statistical variables are used in applied statistics mainly for two purposes.” The first one is “variance stabilization.” “A transformation is applied in order to make it possible to use, at any rate approximately, the standard techniques associated with continuous normal variation, for example, the methods of analysis of variance. In particular, transformations are required which stabilize the variance, that is, which make the variance of the transformed variable approximately independent of, for example, the binomial probability or the mean value of the Poisson distribution. [...] The constant-variance condition has led to the introduction of the inverse sine, inverse, sinh and square-root transformations which are used nowadays in many fields of applications.” In linear regression models, the unequal variance of the error terms produces estimates that are unbiased but are no longer best in the sense of having the smallest variance [9]. The second purpose is “normalization.” “A transformation is used in order to facilitate the computation of tail sums of the distribution by the aid of the normal probability integral [...]. A review of the literature shows that a considerable number of transformations of binomial, negative binomial, Poisson, and χ^2 variables have been proposed.” In regression models, for example, nonnormality of the response variable invalidates the standard tests of significance with small samples since they are based on the normality assumption [9]. Reference [10] noted how “approximately symmetrizing transformation of a random variable may be a more effective method of normalizing it than stabilizing its variance.” Reference [11], although dated, and the more recent reference [12] provide an exhaustive review of transformation used in statistical data analysis.

If a random variable X , whose probability mass function or density function depends on a parameter θ , is transformed by a function $Y = f(X)$, which we suppose henceforth

to be strictly monotone, the standard deviation of Y , according to the delta method (see, e.g., [13]), is given approximately by

$$\sigma_{f(x)}(\theta) \approx \frac{\partial f}{\partial x}(\mu_x(\theta)) \cdot \sigma_x(\theta), \quad (1)$$

where $\mu_x(\theta)$ and $\sigma_x(\theta)$ are the expected value and standard deviation of X . Then the (approximate) standard deviation of the transformed random variable Y can be made equal to a constant if the function f is chosen so to satisfy the following relationship:

$$f(x) = \int^x \frac{k}{\sigma_x(\theta)} d\mu_x(\theta). \quad (2)$$

Thus, for example, if X is distributed as a Poisson random variable with parameter λ being $\mu_x = \sigma_x^2 = \lambda$, we obtain that the function $f(x) = \int^x k\lambda^{-1/2} d\lambda = k^* \sqrt{x}$ is a variance stabilizing function, with k and k^* proper positive constants. Formula (2) has been empirically shown to provide reasonable stabilization in various applications, as confirmed by its extensive use, but other criteria can be employed based on different “notions” of variance stabilization [8, 12]. Otherwise, modifications to the function derived by (2) can be proposed, for example, in order to reduce or remove the bias [14]. Reference [15] studied the root transformation $f(x) = \sqrt{x+c}$ for a Poisson-distributed random variable X and demonstrated that for $c = 1/4$ the root-transformed variable $\sqrt{X+c}$ has vanishing first-order bias and almost constant variance.

Whichever criterion is selected, very often the proper transformation to be adopted is tied to the particular statistical distribution underlying the data. However, as much as often the exact distribution of an estimator for a probability/proportion is not easily derivable, even if the distribution that the sample data came from is known. This sometimes happens, for example, for the maximum likelihood estimator (MLE) of the reliability parameter R of a stress-strength model. In this case, we do not know “a priori” which transformation may fit the data (i.e., the distribution of the MLE) best.

The focus of this paper is the estimation of a probability; thus we will confine ourselves to transformation of proportions. For this case, among the most used transformations, here we recall the logit, probit, arcsine, and complementary log-log. Table 1 reports the expression, the image, the first derivative, and the inverse of these four functions.

The logit and probit transformations are widely used in the homonym models [16] and more generally when dealing with skewed proportion distributions. They are similar since they are both antisymmetric functions around the point $x = 0.5$; that is, $f(x + 0.5) = -f(-x + 0.5)$. The difference stands in the fact that the logit function takes absolute values larger than the probit, as can be noted looking at Figure 1. With regard to the estimation of the reliability parameter for

TABLE 1: Transformations.

$f(x)$	Codomain	$f'(x)$	$f^{-1}(x)$
$\log[x/(1-x)]$	$(-\infty, +\infty)$	$1/[x(1-x)]$	$e^x/(1+e^x)$
$\Phi^{-1}(x)$	$(-\infty, +\infty)$	$1/\phi(\Phi^{-1}(x))$	$\Phi(x)$
$\arcsin(\sqrt{x})$	$(0, +\pi/2)$	$1/[2\sqrt{x(1-x)}]$	$\sin^2(x)$
$\log[-\log(1-x)]$	$(-\infty, +\infty)$	$1/[(1-x)\log(1-x)]$	$1 - e^{-e^x}$

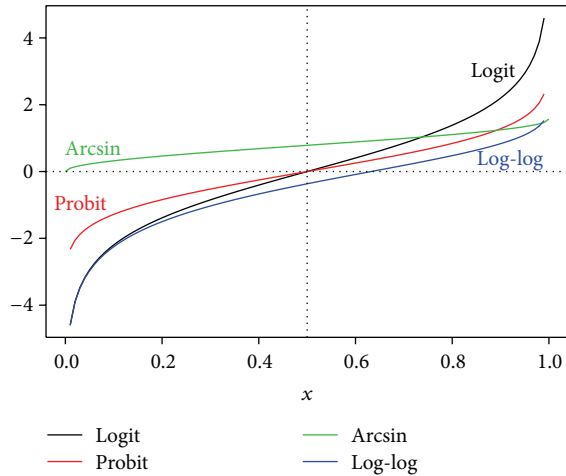


FIGURE 1: Plot of transformation functions and dotted horizontal lines of equation $y = 0$ and $x = 0.5$.

the stress-strength model, the logit transformation has been by far the most used transformation for improving the statistical performance of standard large-sample CIs; among others, it has been considered by [17–19], all these contributions concerning the Weibull distribution, by [20] for the Lindley distribution, by [21] when stress and strength follow a bivariate exponential distribution, and by [22, 23] in a nonparametric context.

Through the years, the arcsine root transformation has gained great favor and application among practitioners [24], perhaps more than its real merit. As noted previously, it should be chosen for stabilizing binomial data but is often used to stabilize sample proportions as well (i.e., relative binomial data). However, it presents a drawback highlighted in [25] and related to the fact that its codomain is the limited interval $(0, \pi/2)$; thus its normalizing effect may turn out to be meagre, as already pointed out in some studies where it has been used for estimating the reliability of stress-strength models [19, 25, 26].

The complementary log-log function, which is sometimes used in binomial regression models, is slightly different from logit and probit since it assumes negative values for $x < 1 - e^{-1}$ (and positive values for $x > 1 - e^{-1}$) and takes large positive values only for values of x very close to 1; for example, it takes values larger than 1 if and only if $x > 1 - e^{-e} = 0.934$.

In the next section, we will apply these transformations to the estimation of a stress-strength model with both stress and strength following a Poisson distribution.

3. Inference on the Reliability Parameter for a Stress-Strength Model with Independent Poisson Stress and Strength

Let X and Y be independent rv's modeling stress and strength, respectively, with $X \sim \text{Poisson}(\lambda_x)$ and $Y \sim \text{Poisson}(\lambda_y)$. Then, the reliability $R = P(X < Y)$ of the stress-strength model is given by (see [6, page 103])

$$\begin{aligned}
 R &= P(X < Y) = \sum_{i=0}^{+\infty} \frac{\lambda_x^i e^{-\lambda_x}}{i!} \left[1 - \sum_{j=0}^i \frac{\lambda_y^j e^{-\lambda_y}}{j!} \right] \\
 &= \lim_{k \rightarrow \infty} \sum_{i=0}^k \frac{\lambda_x^i e^{-\lambda_x}}{i!} \left[1 - \sum_{j=0}^i \frac{\lambda_y^j e^{-\lambda_y}}{j!} \right].
 \end{aligned} \tag{3}$$

If two simple random samples of size n_x and n_y from X and Y , respectively, are available, the reliability parameter can be estimated by the ML estimator, obtained substituting in (3) the MLEs of the unknown parameters λ_x and λ_y :

$$\hat{R} = \sum_{i=0}^{+\infty} \frac{\bar{x}^i e^{-\bar{x}}}{i!} \left[1 - \sum_{j=0}^i \frac{\bar{y}^j e^{-\bar{y}}}{j!} \right]. \tag{4}$$

Deriving the expression of the variance of the MLE is not straightforward; however, an approximate expression has been easily derived through the delta method in [7]. Based upon this approximation, a large-sample $1 - \alpha$ CI for R has been built. Such an interval estimator has the usual expression

$$(R_L, R_U) = \left(\hat{R} - z_{1-\alpha/2} \sqrt{v(\hat{R})}, \hat{R} + z_{1-\alpha/2} \sqrt{v(\hat{R})} \right), \tag{5}$$

where $v(\hat{R})$ is the sample estimate of the (asymptotic) variance of \hat{R} (see [7] for details).

Although such an estimator has been proved to have a satisfactory behavior in terms of coverage for several combinations of sample sizes and values of the reliability parameter, some decay of the performance is observed when R gets close to the extreme values 0 and 1 and when sample sizes are small. In these cases, the approximation to the normal distribution is in fact very rough, especially because of the skewness of the distribution of \hat{R} . Thus, transformations of the values of the estimates \hat{R} can be considered in order to “make the data more normal” and produce CIs based on transformed data that have a coverage closer to the nominal level.

TABLE 2: CIs of the reliability parameter R under variance-stabilizing/normalizing transformations.

Transformation	$\hat{\theta}$	CI for $\theta: (\theta_L, \theta_U)$	CI for $R: (R_L, R_U)$
logit	$\log[\hat{R}/(1 - \hat{R})]$	$\left(\hat{\theta} \mp z_{1-\alpha/2} \sqrt{v(\hat{R})/[\hat{R}(1 - \hat{R})]} \right)$	$(\exp(\theta_L) / (1 + \exp(\theta_L)), \exp(\theta_U) / (1 + \exp(\theta_U)))$
Probit	$\Phi^{-1}(\hat{R})$	$\left(\hat{\theta} \mp z_{1-\alpha/2} \sqrt{v(\hat{R})/\phi(\Phi^{-1}(\hat{R}))} \right)$	$(\Phi(\theta_L), \Phi(\theta_U))$
Arcsin	$\arcsin \sqrt{\hat{R}}$	$\left(\hat{\theta} \mp z_{1-\alpha/2} \sqrt{v(\hat{R})/[2\sqrt{\hat{R}(1 - \hat{R})}]} \right)$	$(\sin^2(\theta_L), \sin^2(\theta_U))$
Loglog	$\log[-\log(1 - \hat{R})]$	$\left(\hat{\theta} \mp z_{1-\alpha/2} \sqrt{v(\hat{R})/[(1 - \hat{R})\log(1 - \hat{R})]} \right)$	$(1 - \exp(-\exp(\theta_L)), 1 - \exp(-\exp(\theta_U)))$

Recalling Table 1 and considering the logit transformation of the reliability parameter R , $\theta = \log[R/(1 - R)]$, the MLE of θ is $\hat{\theta} = \log[\hat{R}/(1 - \hat{R})]$ and an approximate $(1 - \alpha)$ CI for θ is

$$(\theta_L, \theta_U) = \left(\hat{\theta} - z_{1-\alpha/2} \cdot \frac{\sqrt{v(\hat{R})}}{[\hat{R}(1 - \hat{R})]}, \right. \\ \left. \hat{\theta} + z_{1-\alpha/2} \cdot \frac{\sqrt{v(\hat{R})}}{[\hat{R}(1 - \hat{R})]} \right). \quad (6)$$

Then, an approximate $(1 - \alpha)$ CI for R is

$$(R_L, R_U) = \left(\frac{\exp(\theta_L)}{[1 + \exp(\theta_L)]}, \frac{\exp(\theta_U)}{[1 + \exp(\theta_U)]} \right). \quad (7)$$

For the other transformations, following the same steps just outlined, approximate “transformed” CIs for R can be obtained, which are alternative to the standard naïve CI of (5); they are synthesized in Table 2.

An alternative way to make inference on the reliability parameter R for the stress-strength model with independent Poisson random variables can be summarized as follows: (1) to transform (normalize) samples from X and Y according to the a proper transformation for the Poisson distribution; (2) to compute point and interval estimates of R using the methods for a stress-strength model with independent normal variables with known variances [6, page 112]. Letting $\Xi = \sqrt{X + 1/4}$ and $\Upsilon = \sqrt{Y + 1/4}$, then Ξ and Υ are independent and approximately distributed as normal random variables with variance $\sigma_\xi^2 = \sigma_\nu^2 = 1/4$ and expected value $\mu_\xi = \sqrt{\lambda_x}$ and $\mu_\nu = \sqrt{\lambda_y}$, respectively. Then, instead of estimating $R = P(X < Y)$, one can estimate $R' = P(\Xi < \Upsilon)$; a $1 - \alpha$ confidence interval for R' is given by

$$\left(\Phi \left[\hat{\eta} - \frac{z_{1-\alpha/2}}{\sqrt{M}} \right], \Phi \left[\hat{\eta} + \frac{z_{1-\alpha/2}}{\sqrt{M}} \right] \right), \quad (8)$$

with $M = (\sigma_\xi^2 + \sigma_\nu^2)/(\sigma_\xi^2/n_x + \sigma_\nu^2/n_y)$ and $\hat{\eta} = (\bar{v} - \bar{\xi})/\sigma$ being an estimator of $\eta = (\mu_\nu - \mu_\xi)/\sigma$, where $\bar{\xi}$ and \bar{v} are the sample means of Ξ and Υ , respectively, and $\sigma = \sqrt{\sigma_\xi^2 + \sigma_\nu^2} = \sqrt{1/2}$.

An alternative procedure to make inference about the reliability parameter R can be provided by parametric bootstrap

[27, pages 53–56]. In the basic version, it works as follows for the estimation problem at hand.

- (1) Estimate the unknown parameters of the Poisson rv's X and Y through their sample means \bar{x} and \bar{y} .
- (2) Draw independently a bootstrap sample x^* of size n_x from a Poisson rv X^* with parameter \bar{x} and a bootstrap sample y^* of size n_y from a Poisson rv Y^* with parameter \bar{y} .
- (3) Estimate the reliability parameter, say \hat{R}^* , on samples x^* and y^* , using the very same expression in (4).
- (4) Repeat steps 2 and 3 B times (B sufficiently large, e.g., 2,000), thus obtaining the bootstrap distribution of \hat{R}^* .
- (5) Estimate a $(1 - \alpha)$ bootstrap percentile CI for R from \hat{R}^* distribution, taking the $\alpha/2$ and $1 - \alpha/2$ quantiles:

$$(\hat{R}_{\alpha/2}^*, \hat{R}_{1-\alpha/2}^*). \quad (9)$$

4. Simulation Study

4.1. Scope and Design. The simulation study aims at empirically comparing the performance of the interval estimators presented in the previous section, namely, the standard CI of (5), labeled “AN,” those of Table 2 (labeled “logit,” “probit,” “arcsine,” “loglog”), and the CIs of (8) and (9), in terms of coverage rate (and also lower and upper uncovrage rates) and expected length. In this Monte Carlo (MC) study, the value of parameter λ_x of the Poisson distribution for stress X is set equal to a “reference” value 2, and the parameter λ_y of the Poisson distribution modeling strength is varied in order to obtain—according to (3)—several different levels of reliability R , namely, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 0.95.

For each couple (λ_x, λ_y) , a large number ($nS = 5,000$) of samples of size n_x and of size n_y are drawn from $X \sim \text{Poisson}(\lambda_x)$ and $Y \sim \text{Poisson}(\lambda_y)$ independently. Different and unequal sample sizes are here considered (all the 16 possible combinations between the values $n_x = 5, 10, 20, 50$, and $n_y = 5, 10, 20, 50$). On each pair of samples, the 95% CIs for R listed above are built, and the MC coverage rate and the length of the CIs are computed over the nS MC samples. Moreover, the lower and upper uncovrage or error rates are computed, that is, the proportion of CIs for which $R < R_L$ and $R > R_U$, respectively.

Note that for the smallest sample size, that is, when $n_x = 5$ (or $n_y = 5$), a practical problem arises as the sample values of X (or Y) may all be 0; in this case, the variance estimate $v(\hat{R})$ cannot be computed and a standard approximate CI cannot be built. We decided to discard these samples from the 5,000 MC samples planned for the simulation study and to compute the quantities of interest only on the “feasible” samples. Indeed, the rate of such “nonfeasible” samples is in any case very low under each scenario. For the worst ones, characterized by $R = 0.1$ and $n_y = 5$, the theoretical rate of “nonfeasible” samples is about 5%.

Since results for the coverage rates of some CIs are often close to the nominal level 0.95, we performed a test in order to check if the actual coverage rate is significantly different from 0.95, that is, to state if such CIs are conservative or liberal. The null hypothesis is that the true rate η of CIs that do not cover the real value R is equal to $\alpha = 0.05$: $H_0 : \eta = \alpha = 0.05$, whereas the alternative hypothesis is that the rate η is different from α : $H_1 : \eta \neq \alpha$. We employ the test suggested in [28, pages 518-519], which is based on the statistic

$$\mathcal{F} = \frac{nR + 1}{nS - nR} \cdot \frac{1 - \alpha}{\alpha}, \tag{10}$$

where nR is the number of rejections of H_0 in nS (5,000 in our case unless there are some nonfeasible samples) iterations of a MC simulation plan. Under H_0 , \mathcal{F} follows an F distribution with $\nu_1 = 2(nS - nR)$ and $\nu_2 = 2(nR + 1)$ degrees of freedom. The test rejects H_0 at level γ if either $\mathcal{F} \leq \mathcal{F}_{\gamma/2, \nu_1, \nu_2}$ or $\mathcal{F} \geq \mathcal{F}_{1-\gamma/2, \nu_1, \nu_2}$, where $\mathcal{F}_{\gamma, \nu_1, \nu_2}$ denotes the γ percentile point of an F distribution with ν_1 and ν_2 degrees of freedom. The test was performed on each scenario for each kind of CI at a level $\gamma = 1\%$.

4.2. Results and Discussion. The simulation results are (partially) reported in Table 3 ($\lambda_y = 2$, varying R and the common size $n_x = n_y$). The results about the CIs (8) and (9) are not reported here as their performances are overall unsatisfactory. Even if theoretically appealing, the procedure leading to the CI in (8) practically fails; some of the scenarios of the simulation plan were considered, and the CI built following this alternative procedure never provides satisfactory results. For the smallest sample size ($n = 5$) the coverage proves to be larger than that provided by the AN CI but however smaller than the nominal one; for the larger sample size ($n = 10, 20, 50$) the coverage rate dramatically decreases (even to values as low as 60%). Paradoxically, the discrete nature of the Poisson variable and thus the quality of the normal approximation of step (1) affect results to a more relevant extent as the sample size increases. As to the bootstrap procedure yielding the CI in (9), this solution is computationally cumbersome; it becomes even more time consuming if bias-corrected accelerated CIs [27, pages 184–188] have to be calculated and it provides barely satisfactory results, as proven by a preliminary simulation study not reported here that confirms the findings reported in [25] for parametric bootstrap inference on the reliability parameter in the bivariate normal case. The rejection of H_0 based on the test statistic (10) described in the previous subsection

is indicated by a “*” near the actual coverage rate of each CI. Figures 2 and 3 graphically display the values of lower and upper undercoverage rates for the five interval estimators, varying R in $\{0.5, 0.6, 0.7, 0.8, 0.9, 0.95\}$, for $n_x = n_y = 5$ and $n_x = n_y = 50$, respectively. These results comprise only the equal sample size scenarios: those for unequal sample size scenarios do not add much value to the general discussion we are going to outline.

- (i) First, the improvement provided by the four transformations to the standard naïve CI of R in terms of coverage is evident. This improvement is considerable for a small sample size and extreme values of R (say, 0.95) when the coverage rate of the standard naïve CI tends to decrease dramatically, even below 90%. Note that for the sample sizes 5, 10, and 20, the actual coverage rate of the standard naïve CI is always significantly different (smaller) than the nominal level. Under some scenarios, namely, for $R = 0.3-0.7$, the increase in coverage rate is accompanied also by a reduction in the CI average length. On the contrary, for the other values of R , there is an increase in the average width of the modified CIs, which is much more apparent for the logit transformation.
- (ii) Examining closely Table 3, one can note that the CI based on the arcsine root transformation is, except for one scenario, uniformly worse in terms of actual coverage than the other three CIs based on logit, probit, and complementary log-log transformations. For small sample sizes ($n_x = 5, 10$), the coverage rate actually provided is significantly different (smaller) than the nominal one. It can also be claimed that this unsatisfactory performance is due to its incapability of symmetrizing the distribution, as can be seen by glancing at the lower and upper undercoverage rates for values of R getting close to 1; there is an apparent undercoverage effect on the left side (just a bit smaller than that of the “original” standard approximate CI).
- (iii) Logit, probit, and complementary log-log transformed CIs have overall a satisfactory performance in terms of closeness to the nominal confidence level. Logit and probit CIs exhibit a similar behavior as both the lower and upper undercoverage rates they provide are close to the nominal value (2.5%). On the contrary, complementary log-log function (as well as arcsine root) often produces a larger undercoverage on one side (here, left), which is partially balanced by an overcoverage on the other side (here, right). This is clear evidence of the higher symmetrizing capability of the logit and probit transformations. However, taking into account the results of the hypothesis test based on the statistic (10), the probit and complementary log-log functions are those that overall perform best; the statistical hypothesis that their actual coverage rate is equal to the nominal one is always accepted except for one case ($R = 0.1$ and $n = 5$), whereas the same hypothesis is sometimes rejected (10 times out of 40) for the logit function (which tends to produce significantly larger coverage for small sample sizes).

TABLE 3. Simulation results ($\lambda_x = 2$).

R	n_x	n_y	AN			Logit			Probit			Arcsin			Loglog			AN	Logit	Probit	Arcsin	Loglog
			LE	UE	cov	LE	UE	cov	LE	UE	cov	LE	UE	cov	LE	UE	cov					
0.1	5	5	89.94*	0.34	9.72	97.17*	2.66	0.17	96.73*	2.07	1.20	94.35	1.12	4.53	96.20*	3.63	0.17	0.2836	0.3835	0.3491	0.3152	0.4204
	10	10	88.49*	0.40	11.11	96.23*	2.51	1.26	95.57	1.85	2.59	92.52*	1.06	6.42	95.65	3.23	1.12	0.2153	0.2501	0.2354	0.2214	0.2622
	20	20	92.26*	0.36	7.38	95.20	3.06	1.74	94.94	2.32	2.74	93.62*	1.44	4.94	95.12	3.26	1.62	0.1620	0.1709	0.1653	0.1600	0.1750
	50	50	93.34*	0.94	5.72	95.44	2.32	2.24	95.28	1.98	2.74	94.50	1.56	3.94	95.48	2.54	1.98	0.1020	0.1044	0.1029	0.1015	0.1054
0.2	5	5	87.51*	1.66	10.83	96.01*	3.11	0.88	95.11	2.53	2.37	91.66*	1.66	6.68	95.19	4.29	0.52	0.4360	0.4717	0.4553	0.4425	0.5123
	10	10	90.68*	1.12	8.20	95.30	2.68	2.02	94.72	2.22	3.06	93.14*	1.74	5.12	94.86	3.64	1.50	0.3395	0.3429	0.3358	0.3312	0.3597
	20	20	93.18*	1.82	5.00	95.34	2.56	2.10	95.04	2.36	2.60	94.32	1.80	3.88	94.90	3.22	1.88	0.2469	0.2473	0.2444	0.2428	0.2537
	50	50	93.96*	1.38	4.66	95.42	2.68	2.10	95.16	2.20	2.64	94.78	1.96	3.26	95.20	2.76	2.04	0.1569	0.1572	0.1563	0.1559	0.1588
0.3	5	5	89.11*	2.50	8.39	96.42*	2.24	1.34	94.72	2.24	3.04	92.22*	2.24	5.54	94.72	4.48	0.80	0.5440	0.5284	0.5245	0.5264	0.5664
	10	10	91.10*	1.92	6.98	95.40	2.48	2.12	94.90	2.06	3.04	93.70*	1.92	4.38	94.28	4.12	1.60	0.4149	0.3991	0.3981	0.4002	0.4166
	20	20	93.18*	1.82	5.00	95.34	2.56	2.10	95.04	2.36	2.60	93.98*	2.32	3.70	94.96	3.48	1.56	0.3012	0.2947	0.2945	0.2956	0.3018
	50	50	94.38	1.82	3.80	95.46	2.42	2.12	95.32	2.20	2.48	95.10	1.96	2.94	95.40	2.80	1.80	0.1927	0.1909	0.1909	0.1912	0.1929
0.4	5	5	88.98*	3.86	7.16	96.52*	2.10	1.38	94.48	2.32	3.20	92.40*	2.90	4.70	95.04	4.14	0.82	0.6068	0.5592	0.5627	0.5733	0.5891
	10	10	92.14*	2.70	5.16	95.86*	2.08	2.06	94.78	2.28	2.94	93.72*	2.46	3.82	94.98	3.82	1.62	0.4572	0.4305	0.4335	0.4397	0.4455
	20	20	93.64*	2.60	3.76	95.58	2.44	1.98	95.22	2.44	2.34	94.70	2.48	2.82	95.14	3.32	1.54	0.3319	0.3210	0.3226	0.3254	0.3274
	50	50	94.72	2.18	3.10	95.48	2.20	2.32	95.28	2.20	2.52	95.06	2.20	2.74	95.64	2.52	1.84	0.2128	0.2098	0.2103	0.2111	0.2116
0.5	5	5	89.32*	5.04	5.64	96.72*	1.56	1.72	95.04	2.38	2.58	92.78*	3.48	3.74	94.90	4.34	0.76	0.6285	0.5691	0.5752	0.5889	0.5864
	10	10	91.98*	3.82	4.20	95.54	2.04	2.42	94.94	2.36	2.70	93.84*	2.78	3.38	94.88	3.62	1.50	0.4709	0.4406	0.4450	0.4526	0.4501
	20	20	93.50*	3.32	3.18	95.36	2.30	2.34	94.80	2.54	2.66	94.64	2.62	2.74	94.76	3.50	1.74	0.3413	0.3290	0.3312	0.3345	0.3332
	50	50	94.62	2.46	2.92	95.38	2.18	2.44	95.16	2.32	2.52	94.96	2.34	2.70	95.42	2.72	1.86	0.2191	0.2157	0.2163	0.2173	0.2169
0.6	5	5	89.44*	6.36	4.20	96.52*	1.44	2.04	94.70	2.50	2.80	92.66*	4.04	3.30	94.78	4.28	0.94	0.6095	0.5587	0.5628	0.5741	0.5601
	10	10	91.78*	4.84	3.38	95.52	1.78	2.70	94.88	2.32	2.80	93.92*	3.06	3.02	94.92	3.48	1.60	0.4576	0.4305	0.4337	0.4401	0.4324
	20	20	93.40*	4.02	2.58	95.44	2.02	2.54	95.12	2.34	2.54	94.46	2.98	2.56	95.00	3.34	1.66	0.3304	0.3197	0.3212	0.3240	0.3206
	50	50	95.06	2.68	2.26	95.54	1.98	2.48	95.38	2.14	2.48	95.24	2.28	2.48	95.58	2.58	1.84	0.2120	0.2090	0.2095	0.2103	0.2093
0.7	5	5	89.52*	7.54	2.94	96.06*	1.48	2.46	95.04	2.50	2.46	92.56*	4.76	2.68	94.54	4.54	0.92	0.5490	0.5266	0.5242	0.5279	0.5094
	10	10	91.70*	5.82	2.48	95.32	1.64	3.04	95.02	2.22	2.76	93.78*	3.66	2.56	94.92	3.44	1.64	0.4162	0.3991	0.3987	0.4014	0.3919
	20	20	93.28*	4.68	2.04	95.62	1.90	2.48	95.18	2.34	2.48	94.44	3.26	2.30	95.30	3.26	1.44	0.2984	0.2920	0.2918	0.2930	0.2889
	50	50	94.74	3.24	2.02	95.48	1.88	2.64	95.40	2.16	2.44	95.24	2.52	2.24	95.38	2.64	1.98	0.1910	0.1893	0.1893	0.1896	0.1885
0.8	5	5	88.94*	9.60	1.46	96.00*	1.24	2.76	94.92	2.32	2.76	92.54*	5.78	1.68	95.00	4.18	0.82	0.4419	0.4657	0.4526	0.4436	0.4289
	10	10	91.66*	7.02	1.32	95.18	1.56	3.26	94.96	2.34	2.70	93.70*	4.18	2.12	95.22	3.36	1.42	0.3411	0.3408	0.3349	0.3315	0.3242
	20	20	93.22*	5.64	1.14	95.22	1.88	2.90	94.94	2.52	2.54	94.32	3.64	2.04	95.00	3.34	1.66	0.2424	0.2425	0.2399	0.2386	0.2356
	50	50	94.60	3.94	1.46	95.34	1.94	2.72	95.20	2.32	2.48	95.04	2.86	2.10	95.40	2.70	1.90	0.1543	0.1544	0.1537	0.1533	0.1525
0.9	5	5	88.32*	11.14	0.54	95.28	1.14	3.58	95.04	2.58	2.38	91.96*	6.76	1.28	94.70	4.38	0.92	0.2799	0.3557	0.3295	0.3043	0.3029
	10	10	90.80*	8.50	0.70	95.12	1.44	3.44	94.76	2.52	2.72	93.40*	4.98	1.62	95.08	3.30	1.62	0.2175	0.2410	0.2294	0.2187	0.2181
	20	20	92.48*	6.92	0.60	95.24	1.74	3.02	95.12	2.36	2.52	94.22*	4.28	1.50	95.20	3.20	1.60	0.1550	0.1621	0.1574	0.1532	0.1531
	50	50	94.52	4.62	0.86	95.10	1.98	2.92	95.30	2.30	2.40	95.06	3.22	1.72	95.40	2.62	1.98	0.0976	0.0996	0.0983	0.0971	0.0971
0.95	5	5	86.86*	12.96	0.18	95.16	1.12	3.72	95.10	2.62	2.28	91.52*	7.72	0.76	95.02	4.12	0.86	0.1691	0.2593	0.2289	0.1973	0.2058
	10	10	89.90*	9.84	0.26	94.80	1.48	3.72	94.98	2.50	2.52	92.62*	6.28	1.10	94.84	3.46	1.70	0.1277	0.1611	0.1489	0.1359	0.1398
	20	20	92.30*	7.50	0.20	95.22	1.74	3.04	95.12	2.48	2.40	94.02*	4.76	1.22	94.92	3.32	1.76	0.0931	0.1021	0.0975	0.0925	0.0941
	50	50	94.20*	5.32	0.48	95.30	1.88	2.82	95.24	2.40	2.36	95.08	3.46	1.46	95.32	2.78	1.90	0.0579	0.0602	0.0590	0.0577	0.0582

Legend: "cov" = actual coverage rate; "LE" = lower error rate; "UE" = upper error rate.
 *The rejection of H0.

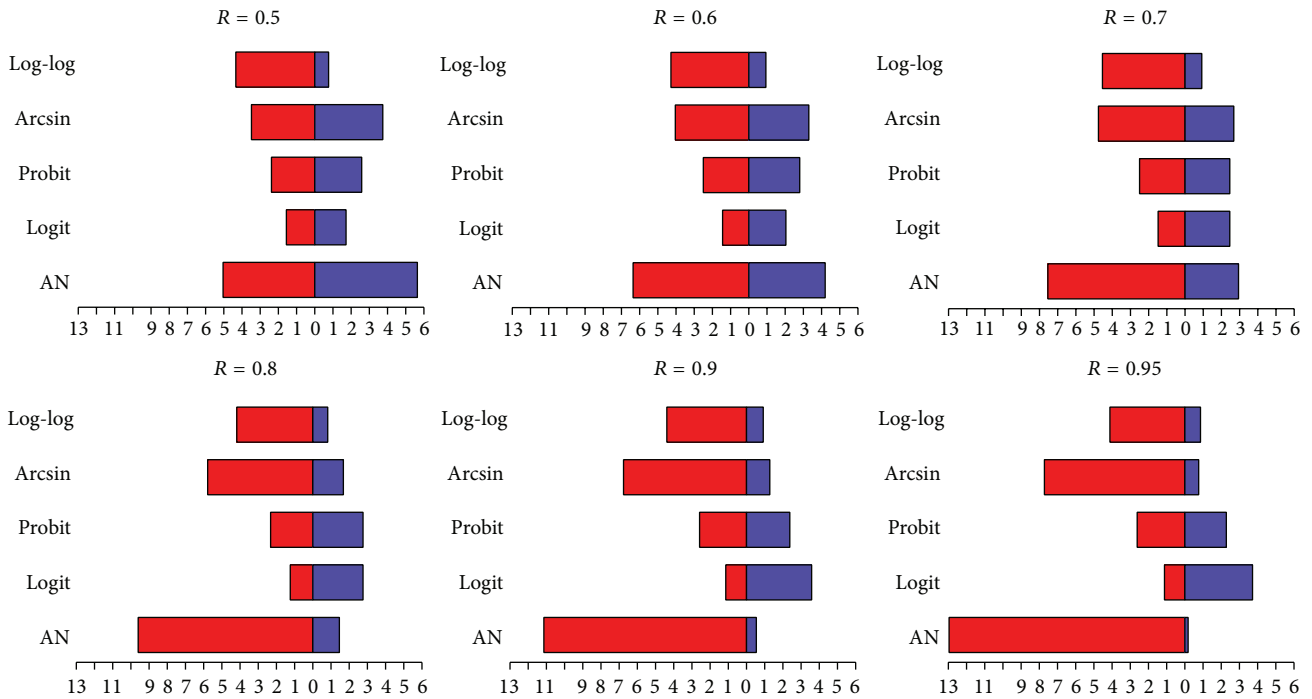


FIGURE 2: Lower and upper uncoverage rates: $\lambda_x = 2, R \geq 0.5, n_x = n_y = 5$.

- (iv) Differences in behavior among the four “transformed” CIs tend to become more apparent for small sample sizes and for extreme values of R (close to 0 or 1) as can be seen looking at the bottom right graph of Figure 2. Clearly, the same differences tend to vanish as the sample sizes increase and R tends to 0.5 (see the top left graph in Figure 3).
- (v) Contrary to what one would predict, the performances in terms of coverage rate of logit, probit, and complementary log-log-based CIs do not seem to be negatively affected by sample size; in fact, as underlined before, an (sometimes significant) increase of the coverage rate of the logit CIs is noticed when the sample size is small ($n_x = 5$), whereas the other two CIs show an almost constant trend. The CI exploiting the arcsine root transformation is the one taking most advantage from the increase in sample size. Obviously, there is an increase in the average length of the CIs moving to larger sample sizes.
- (vi) As one could expect, there is a symmetry in the behavior of each CI when considering two complementary values of R (i.e., values of R summing to 1) keeping the sample size fixed; the values of the coverage rate and average width are similar. As to the uncoverage rates, they are nearly exchangeable for AN, logit, probit, and arcsin CIs; that is, a lower uncoverage error for a fixed value of R is similar to the upper uncoverage error for its complementary value (or at least their relative magnitude is exchangeable). This feature does

not hold for log-log CIs, which always present a left-side uncoverage error larger than that of the right-side. These features are weakened for R equal to 0.9 (and $1 - R$ then equal to 0.1) probably because of the presence of “nonfeasible” samples as discussed in Section 4.1, which distorts this symmetry condition.

- (vii) Finally—the relative results are not reported here for the sake of brevity—moving to larger values (i.e., larger than 2) of λ_x , keeping R constant, seems to bring benefit to the performance of all the CIs as can be quantitatively noted by the reduced number of rejections of the null hypothesis of equivalence between actual and theoretical coverage. This may be explained by the fact that for a large parameter λ , the Poisson distribution tends to a normal distribution; therefore, larger values of λ_x imply a better normal approximation to Poisson and, presumptively, to the MLE \hat{R} .

In Figure 4, for illustrative purposes, the MC distribution of \hat{R} and the transformed data according to the four transformations are displayed ($R = 0.8$ and $n_x = n_y = 5$). We can note at a glance that logit and probit produce distributions closer to normality than arcsine root and log-log complementary functions, which do not seem completely able to “correct” the skewness of the original distribution. However, implementing the Shapiro-Wilk test of normality on the four transformed distributions leads to very low P -values, practically equal to 0, except for the probit function whose P -value is 0.03132, thus proving that in this case all the transformed data distributions are still far from normality.

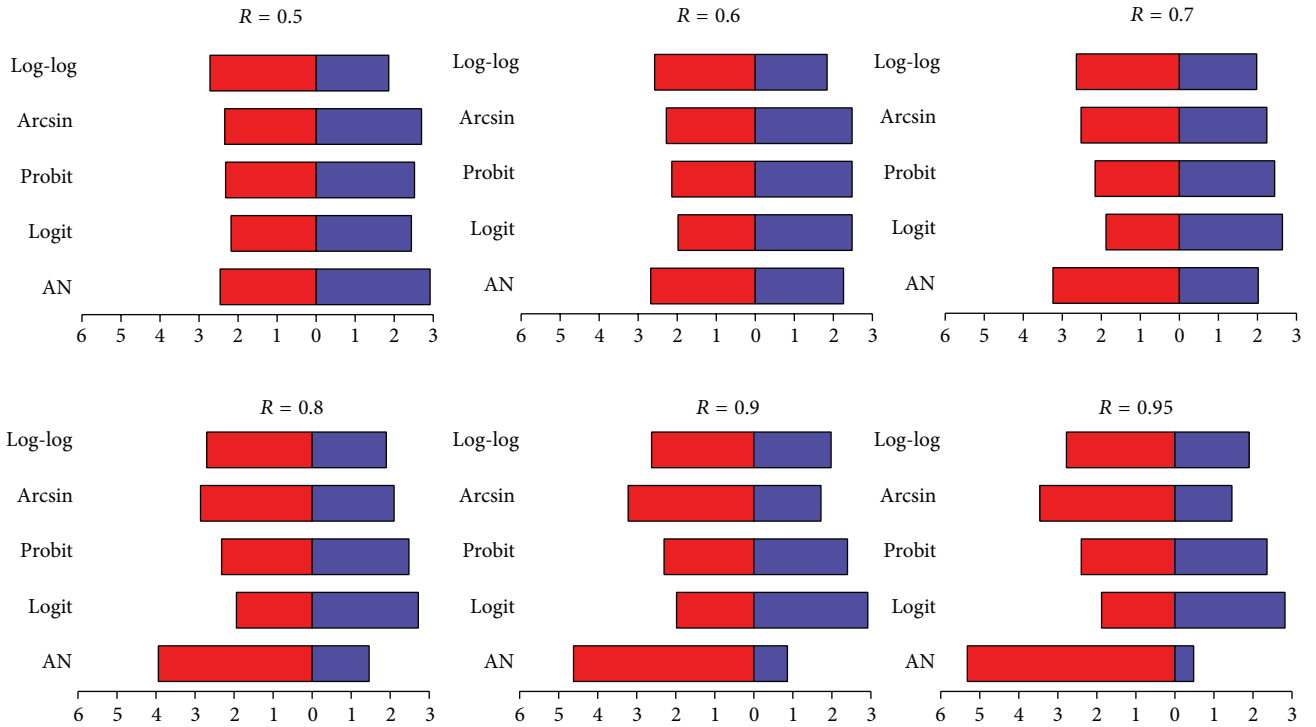


FIGURE 3: Lower and upper uncoverage rates: $\lambda_x = 2, R \geq 0.5, n_x = n_y = 50$.

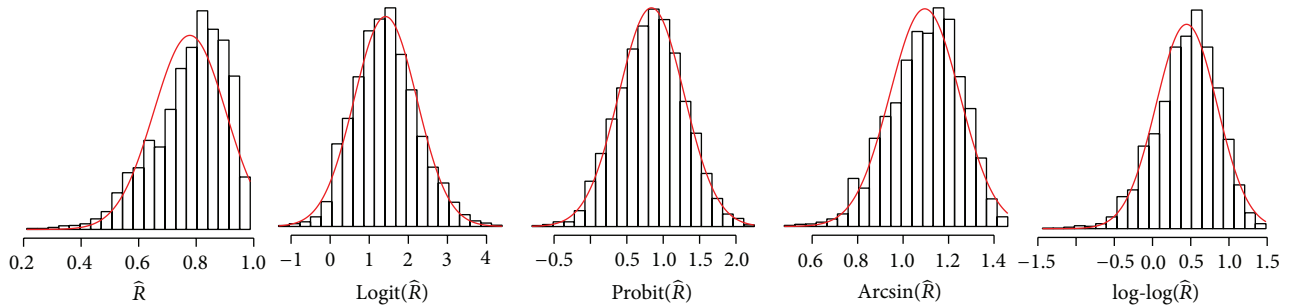


FIGURE 4: Histogram of the MC empirical distribution of \hat{R} and $\hat{\theta}$ according to the transformations of Table 2 ($\lambda_x = 2, R = 0.8, n_x = n_y = 5$). Superimposed, the density function of the normal distribution with mean and standard deviation equal to the mean and standard deviation of the data. Note that neither the y -axis nor x -axis shares the same scale in the different plots.

5. An Application

The application we illustrate here is based on the data described in [29] and already used in [7], to which we redirect for full details. On these data, we build the four 95% CIs based on the transformations of Table 2, along with the standard one. The results (lower and upper bounds and length of the CIs) are reported in Table 4. Although the five CIs are not much different from each other, nevertheless we can note that all four transformation-based intervals have both lower and upper bounds smaller than the corresponding bound for the standard naïve interval (i.e., they are left shifted with respect to it); the logit transformation yields an interval a bit wider than the standard one, whereas the other three transformations produce a slight decrease in its length.

TABLE 4: Results for the application data.

Type	R_L	R_U	Width
AN	0.7452	0.9232	0.1780
Logit	0.7256	0.9054	0.1799
Probit	0.7302	0.9080	0.1777
Arcsine	0.7365	0.9128	0.1763
Log-log	0.7363	0.9113	0.1750

6. Conclusions

This work provided an empirical analysis of the convenience of data transformations for estimating the reliability

parameter in stress-strength models. We focused on the model involving two independent Poisson random variables since this is a case where the distribution of the MLE is not easily derivable and exact CIs cannot be built; thus, transformation of the estimates is a viable tool to refine the standard naïve interval estimator derived by the central limit theorem. Four transformations were considered (logit, probit, arcsine root, and complementary log-log) and empirically assessed under a number of scenarios in terms of the coverage and average length of the CI they produced. Results are in favor of logit, probit, and complementary log-log functions—although to varying degrees—which ensure a coverage rate close to the nominal coverage even with small sample sizes. On the contrary, the arcsine root function, although improving the performance of the standard CI, often keeps its coverage rates under the nominal confidence level. These findings were to some extent predictable since logit and probit transformations are very popular link functions in generalized regression models but are a bit surprising with regard to the complementary log-log function whose use is limited.

Future research will ascertain if such results hold true for other distributions for the stress-strength model and will eventually inspect alternative data transformations.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The author would like to thank the editor and two anonymous referees whose suggestions and comments improved the quality of the original version of the paper.

References

- [1] J. L. Fleiss, B. Levin, and M. C. Paik, *Statistical Methods for Rates and Proportions*, John Wiley & Sons, Hoboken, NJ, USA, 3rd edition, 2003.
- [2] A. Agresti and B. A. Coull, “Approximate is better than “exact” for interval estimation of binomial proportions,” *The American Statistician*, vol. 52, no. 2, pp. 119–126, 1998.
- [3] L. D. Brown, T. T. Cai, and A. Dasgupta, “Confidence intervals for a binomial proportion and asymptotic expansions,” *The Annals of Statistics*, vol. 30, no. 1, pp. 160–201, 2002.
- [4] A. M. Pires and C. Amado, “Interval estimators for a binomial proportion: comparison of twenty methods,” *REVSTAT Statistical Journal*, vol. 6, no. 2, pp. 165–197, 2008.
- [5] B. V. Gnedenko and I. A. Ushakov, *Probabilistic Reliability Engineering*, John Wiley & Sons, New York, NY, USA, 1995.
- [6] S. Kotz, Y. Lumelskii, and M. Pensky, *The Stress-strength Model and Its Generalizations: Theory and Applications*, World Scientific, New York, NY, USA, 2003.
- [7] A. Barbiero, “Inference for reliability of Poisson data,” *Journal of Quality and Reliability Engineering*, vol. 2013, Article ID 530530, 8 pages, 2013.
- [8] G. Blom, “Transformations of the binomial, negative binomial, poisson and X^2 distributions,” *Biometrika*, vol. 41, pp. 302–316, 1954.
- [9] D. C. Montgomery, *Design and Analysis of Experiment*, John Wiley & Sons, New York, NY, USA, 6th edition, 2005.
- [10] Y. P. Chaubey and G. S. Mudholkar, “On the symmetrizing transformations of random variables,” Concordia University, unpublished preprint, Mathematics and Statistics, 1983.
- [11] M. H. Hoyle, “Transformations: an introduction and a bibliography,” *International Statistical Review*, vol. 41, no. 2, pp. 203–223, 1973.
- [12] A. Foi, “Optimization of variance-stabilizing transformations,” submitted.
- [13] W. Greene, *Econometric Analysis*, Prentice Hall, New York, NY, USA, 7th edition, 2012.
- [14] A. Dasgupta, *Asymptotic Theory of Statistics and Probability*, Springer, New York, NY, USA, 2008.
- [15] L. Brown, T. Cai, R. Zhang, L. Zhao, and H. Zhou, “The root-unroot algorithm for density estimation as implemented via wavelet block thresholding,” *Probability Theory and Related Fields*, vol. 146, no. 3-4, pp. 401–433, 2010.
- [16] J. H. Aldrich and F. D. Nelson, *Linear Probability, Logit, and Probit Models*, Sage, Beverly Hills, Calif, USA, 1984.
- [17] A. Baklizi, “Inference on $Pr(X < Y)$ in the two-parameter Weibull model based on records,” *ISRN Probability and Statistics*, vol. 2012, Article ID 263612, 11 pages, 2012.
- [18] K. Krishnamoorthy and Y. Lin, “Confidence limits for stress-strength reliability involving Weibull models,” *Journal of Statistical Planning and Inference*, vol. 140, no. 7, pp. 1754–1764, 2010.
- [19] S. P. Mukherjee and S. S. Maiti, “Stress-strength reliability in the Weibull case,” in *Frontiers in Reliability*, vol. 4, pp. 231–248, World Scientific, Singapore, 1998.
- [20] D. K. Al-Mutairi, M. E. Ghitany, and D. Kundu, “Inferences on stress-strength reliability from Lindley distributions,” *Communications in Statistics: Theory and Methods*, vol. 42, no. 8, pp. 1443–1463, 2013.
- [21] M. M. Shoukri, M. A. Chaudhary, and A. Al-Halees, “Estimating $P(Y < X)$ when X and Y are paired exponential variables,” *Journal of Statistical Computation and Simulation*, vol. 75, no. 1, pp. 25–38, 2005.
- [22] A. Baklizi and O. Eidous, “Nonparametric estimation of $(X < Y)$ using kernel methods,” *Metron*, vol. 64, no. 1, pp. 47–60, 2006.
- [23] G. Qin and L. Hotilovac, “Comparison of non-parametric confidence intervals for the area under the ROC curve of a continuous-scale diagnostic test,” *Statistical Methods in Medical Research*, vol. 17, no. 2, pp. 207–221, 2008.
- [24] W. H. Ahrens, “Use of the arcsine and square root transformations for subjectively determined percentage data,” *Weed Science*, vol. 38, no. 4-5, pp. 452–458, 1990.
- [25] A. Barbiero, “Interval estimators for reliability: the bivariate normal case,” *Journal of Applied Statistics*, vol. 39, no. 3, pp. 501–512, 2012.
- [26] A. Barbiero, “Confidence intervals for reliability of stress-strength models in the normal case,” *Communications in Statistics—Simulation and Computation*, vol. 40, no. 6, pp. 907–925, 2011.

- [27] B. Efron and R. Tibshirani, *An Introduction to the Bootstrap*, Chapman and Hall, London, UK, 1993.
- [28] I. R. Cardoso de Oliveira and D. F. Ferreira, "Multivariate extension of chi-squared univariate normality test," *Journal of Statistical Computation and Simulation*, vol. 80, no. 5-6, pp. 513–526, 2010.
- [29] P. C. Wang, "Comparisons on the analysis of Poisson data," *Quality and Reliability Engineering International*, vol. 15, pp. 379–383, 1999.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

