Estimating reliability in a strength-stress interference model: classical and bayesian perspectives

L. Lee Ho & A. Sá Peixoto Pinheiro
Depto de Engenharia de Produção - Escola Politécnica – USP
Caixa Postal 61548 - 05424-970 – SP. SP. Brazil

Abstract:
In a strength-stress interference model we are interested in estimating the reliability. Here we present two results: the first is a conservative lower bound of confidence interval of the reliability when stress and strength are independent normal random variables and the second is an empirical estimator of the reliability which inferential aspects were developed under a Bayesian perspective. Both results were illustrated by examples and simulations.

1 Introduction
In many situations, we are interested in calculating the probability of the event X < Y, where X and Y are independent random variables with means and variances $\mu_X$, $\sigma_X^2$ and $\mu_Y$, $\sigma_Y^2$ respectively. One of this situation is in a stress-strength interference model where X and Y are assumed respectively as strength and stress random variables. In engineering science, safety margin (SM) and safety factor (SF) are functions of stress and strength and they are respectively settled as SM = X - Y and SF = X/Y and the most interest concerns in obtaining the probability $p_r = P(SM > 0) = P(X-Y > 0)$ or $P(SF > 1) = P(X/Y > 1)$. (See Ebeling [1]). There is a large number of studies on this subject. And results under distribution functions for strength and stress such as normal, log-normal, exponential, gamma or Weibull are known. It is not difficult to verify that when X and Y are independent normally distributed, the probability $p_r = P(SM > 0)$ is given by

$$p_r = \Phi \left( \frac{\mu_X - \mu_Y}{\sigma_X^2 + \sigma_Y^2} \right)$$

where $\Phi (.)$ denotes the cumulative standard normal distribution. Similarly when X and Y are independent log-normally distributed, $p_r = P(SF > 1)$ is given by
\[ p_t = \Phi\left( \frac{\ln \xi_{0.5}(X) - \ln \xi_{0.5}(Y)}{\sigma_{\ln X}^2 + \sigma_{\ln Y}^2}^{0.5} \right) \]  \( \text{(2)} \)

where \( \xi_{0.5}(i) \) denotes the q-th quantile of i, i=X and Y; \( \sigma_i \) denotes the standard deviation of i, i=lnX, lnY. When X and Y are independent gamma random variables,

\[ p_t = \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} B(\gamma/(1 + \gamma), a, b) \]  \( \text{(3)} \)

where \( \gamma = \mu_X/\mu_Y, (a, \gamma_X) \) and \( (b, \gamma_Y) \) are respectively parameters of X and Y, and \( B(x, y, z) \) denotes \( P(U < x) \), where U is Beta distribution with parameters \( (y, z) \). When \( a = b = 1 \), X and Y are both exponential distributions, the eqn (3) reduces to

\[ p_t = \frac{\lambda_Y}{\lambda_Y + \lambda_X} = \frac{\mu_X}{\mu_X + \mu_Y} \]  \( \text{(4)} \)

where \( \lambda_i \) is the parameter of exponential distribution i, i=X and Y. And when X and Y are independent Weibull distributed, \( p_t \) does not have a closed expression and it is given by

\[ p_t = P(X > Y) = 1 - \int_0^\infty e^{-z} \exp\left\{ - \left[ \frac{\eta_X}{\eta_Y} z^{1/\beta_X} + \left( \frac{x}{\eta_Y} \right)^{\beta_Y} \right] \right\} dz \]  \( \text{(5)} \)

where \( (\beta_i, \eta_i) \) are parameters of i, i=X,Y; \( Z = \left( \frac{x}{\eta_X} \right)^{\beta_X} \), which implies

\[ dz = \frac{\beta_X}{\eta_X} \left( \frac{x}{\eta_X} \right)^{\beta_X - 1} dx \]. The major difficulty on this subject is not to calculate theoretical values of \( P( SM > 0) \) or \( P( SF > 1) \) but it is still to obtain estimates for (1)-(5), identify the distribution for the estimators of (1)-(5) as also their properties in order to make correct statistical inferences. Many authors have made efforts to propose estimates for (1)-(5). Under normal distributions for strength and stress, we may refer to Church & Harris [2]. They have proposed confidence limits for eqn(1) when \( \mu_X, \sigma_X^2 \) are unknown and \( \mu_Y = 0, \sigma_Y^2 = 1 \). Raiser & Guttman [3] have also presented confidence limits for eqn(1), considering the parameters \( \mu_X, \sigma_X^2 \) and \( \mu_Y, \sigma_Y^2 \) unknown, which require values of a non-central t-Student distribution when the sample sizes of X and Y are different. When the sample sizes are equal, the estimator of \( K = \frac{\mu_X - \mu_Y}{(\sigma_X^2 + \sigma_Y^2)^{0.5}} \)

follows asymptotically a normal distribution and approximated lower bound confidence limits of \( \Phi(K) \) can be obtained, since \( \Phi(.) \) is a strictly increasing monotonic function of K. And about inferential aspects when both are lognormally distributed, the procedures proposed by Raiser & Guttman [3] can be adopted after a logarithmic transformation on the observations. Under exponential distributions, we have the contributions from Tong [4]. Both have proposed unbiased minimum variance estimator for eqn(4). Kelley et all [5] have contributed also deriving the variance of Tong’s estimator. There are few studies concerning on estimates for eqn(3) and eqn(5). More details see Basu [6] and McCool [7], respectively.
What concerns us in this paper is to propose alternative methods to analyze a dataset from a strength-stress interference study, estimate the reliability and make inferences on this parameter. Here we present two results: a conservative confidence interval for \( p_r \) employing the lower bound \( \alpha_r \) proposed by Ichikawa [8]. This is appealing since \( \alpha_r \) is a lower bound of \( p_r \) for any distribution, estimates of \( \alpha_r \) can be handled and the lower confidence limit of \( \alpha_r \) can also be used as a conservative lower bound confidence limit for \( p_r \). In Section 2, confidence interval for \( \alpha_r \) is obtained under strength and stress normally distributed. When they are log-normally distributed, the results can easily be extended. To illustrate the coverage of the proposed conservative limits, simulations are performed under strength and stress independent normally distributed. When stress or strength are neither normal nor log-normal distributed, inference on reliability still remains not easily handled. Let us consider \( X_1, X_2, \ldots, X_n \) and \( Y_1, Y_2, \ldots, Y_n \) independent random samples of \( X \) and \( Y \), with \( n \) the sample size of \( X \) and \( Y \) and let \( W_i = \begin{cases} 1 & \text{if } X_i > Y_i \\ 0 & \text{otherwise} \end{cases} \) a Bernoulli random variable with parameter equal to \( p_r \). A natural estimator for \( p_r \) in this case is the empirical estimator \( \hat{p}_r = \frac{\sum W_i}{n} \). But for high reliable systems, \( \hat{p}_r \) is likely to be one and inference under statistical classical methods cannot be applied. Another alternative is to rewrite \( p_r \) as a function of \( h(\theta) \), where \( \theta \) is the parameter vector of the stress or strength distributions and make inference on \( p_r \) through \( \hat{p}_r = h(\hat{\theta}) \), where \( \hat{\theta} \) is, for example, maximum likelihood estimators of the unknown parameters. However, this approach can turn unmanageable since the distribution of \( \hat{p}_r \) may be difficult to study. For these reasons we shall propose inferences on \( p_r \) under Bayesian perspectives. Simulations under the most used distributions such as the earlier mentioned were obtained to evaluate the performance of the proposed estimator. In Section 3, the empirical estimator and its Bayesian confidence interval obtained such way are presented. The results of the simulations are described in Section 3 also. Some discussions and conclusions are drawn in the last section.

### 2 Conservative lower confidence limits for \( p_r \)

Ichikawa [8] has proposed an upper bound for \( p_r = P(X < Y) \) given by

\[
p_r \leq \frac{\sigma_X^2 + \sigma_Y^2}{(\mu_X - \mu_Y)^2 + \sigma_X^2 + \sigma_Y^2} = \alpha_r
\]

Reiser [9] has made a remark on eqn(6) adding a condition to hold the inequality. So eqn(6) only holds if \( \mu_X \geq \mu_Y \), otherwise \( p_r < 1 \). Manipulating eqn(6), we have the inequality

\[
p_r = 1 - p_r > 1 - \frac{\sigma_X^2 + \sigma_Y^2}{(\mu_X - \mu_Y)^2 + \sigma_X^2 + \sigma_Y^2} \Rightarrow p_r^{-1} < 1 + \frac{\sigma_X^2 + \sigma_Y^2}{(\mu_X - \mu_Y)^2} = 1 + CV^2_{(X-Y)} = \alpha_r^{-1}
\]
where CV(\cdot) denotes the coefficient of variation of a random variable, with X and Y independent random variables. Let us consider \(X_1, X_2, \ldots, X_n\) and \(Y_1, \ldots, Y_n\) independent random samples of X and Y and \(n\) the sample size of X and Y.

A natural estimator for \(CV(X,Y)\) is \(CV(X,Y) = \frac{(S^2_X + S^2_Y)^{0.5}}{\bar{X} - \bar{Y}}\) where \(S_i\) denotes the sample standard deviation of \(i\), \(i = X\) and \(Y\); \(\bar{X}\) and \(\bar{Y}\) are the sample means. Here we propose a conservative lower confidence limit for \(p_r\). When X and Y are independent normal distributions, \((X-Y)\) is still a normal distribution and an approximated confidence interval for \(CV\) under a normal distribution is given by

\[
CV\left[1 \pm z_{\alpha/2} \sqrt{\frac{(0.5 + CV^2)}{n}}\right]^{-1},
\]

where \(CV\) is an estimate of \(CV\) and \(z_{\alpha/2}\) is such that \(P\left(Z < z_{\alpha/2}\right) = \alpha/2\) (Johnson, Kotz & Balakrishnan [10]) and \(Z\), a standard normal distribution. Manipulating the above expression, approximated confidence intervals for \(p_r\) can be obtained which is given by

\[
\left(1 + \left\{CV \left[1 \pm z_{\alpha/2} \sqrt{\frac{(0.5 + CV^2)}{n}}\right]^{-1}\right\}^2\right)^{-1}
\]  

(8)

Since \(\alpha_r < p_r\), the lower bound of eqn(8) can be used as a lower confidence limit for \(p_r\) \((SM > 0)\). Extending these results when X and Y are lognormally distributed, we have \((lnX - lnY)\) is still also a normal distribution and lower bound of eqn(8) can be used as a conservative lower confidence limit for \(p_r\). \((lnX - lnY > 0) = P(X/Y > 1) = P(SF > 1)\). If different sample sizes are taken for X and Y, an alternative is to consider \(n = \min\{m, n\}\) to keep the conservative property of eqn(8). In practical applications similar sample sizes for X and Y are strongly recommendable. In order to illustrate the coverage of the lower bound proposed in eqn(8), simulations were conducted for this purpose. Independent normal distributions for strength and stress were simulated considering the following parameters: \(\mu_X = 80, \sigma_X = 22.4\) and \(\mu_Y = 20, \sigma_Y = 10\) which corresponds to \(p_r = 0.99492\) and \(\alpha_r = 0.86903\). Sample sizes \(n= 10, 50, 100, 200\) and \(1000\) were taken and for each sample, estimate of \(CV(X,Y)\) was performed as also confidence interval of \(\alpha_r\) This procedure was replicated one thousand times. Table 1 shows the absolute frequency of the uncoverage lower bound of the intervals for \(p_r\) and \(\alpha_r\). According to Table 1, even when the sample size is small, the lower limit proposed in eqn(8) can be used since it was observed zero % of uncoverage.

<table>
<thead>
<tr>
<th>Sample size</th>
<th># &lt; 0.99492</th>
<th># &lt; 0.86903</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>50</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1000</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

The main advantage of the present result is the easiness to get the results and the possibility to be applied either to different sample sizes or equal sample sizes of X and Y, although the conservative property.
3 Bayesian estimator for \( p_r \)

Let us consider \( X_1, X_2, \ldots, X_n \) and \( Y_1, Y_2, \ldots, Y_n \) independent random samples of \( X \) and \( Y \) and \( n \), the sample size of \( X \) and \( Y \) and \( W_i = \begin{cases} 1 & \text{if } X_i > Y_i \\ 0, & \text{otherwise} \end{cases} \)

Bernoulli random variable with parameter equal to \( p_r \). In Bayesian analysis, the uncertainty concerning \( p_r \) may be described by a Beta prior distribution with parameters \( a \) and \( b \). And it is not difficulty to show that the posterior distribution of \( (p_r|s) \) follows also a Beta distribution with parameters \( (a+s) \) and \( (n-s+b) \), \( s = \sum_{i=1}^{n} W_i \) is the number of successes in \( n \) trials and a Bayesian estimation for \( p_r \) under a squared error loss function is given by

\[
\hat{p}_r = \frac{a + s}{n + b + a}
\]

(More details see Martz & Waller [11]). Notice that under this perspective, inferential procedure follows even when \( s=n \) and Bayesian confidence intervals can also be easily handled. Let \( [L, U] \) a Bayesian confidence interval for \( p_r \) with confidence coefficient \( (1-\gamma) \) such that \( P(L < p_r < U \mid W_1=w_1, \ldots, W_n=w_n) = 1 - \gamma \). The values of \( L \) and \( U \) can be obtained using the well-known relation between Beta and Fisher-Snedecor distributions and they are respectively

\[
L = \frac{f(\gamma)(a+s)}{(n-s+b)+(a+s)f(\gamma)} \quad \text{and} \quad U = \frac{f(1-\gamma)(a+s)}{(n-s+b)+(a+s)f(1-\gamma)}
\]

where \( f(k) \) is a constant such that \( P[ F < f(k) ] = k \), \( 0 < k < 1 \) and \( F \) is Fisher-Snedecor distribution with parameters \( 2(a+s); 2(n-s+b) \). When intervals \( [L, U] \) are constructed satisfying \( P( p_r < L ) = P( p_r > U ) \), they are called as symmetric Bayesian confidence intervals for \( p_r \). In reliability studies, it is reasonable taken a large number of the samples. Let suppose that the first sample of size \( n_1 \) provides us \( s_1 \) successes, which results a Bayesian estimate \( \hat{p}_r^{(1)} = (s_1 + a)/(n_1 + a + b) \) in accordance with eqn(9). In Bayesian analysis, Beta posteriori distribution \( (p_r \mid s_1) \) can be used as a prior distribution to obtain a new recalibrated estimate of \( p_r \) when a second sample of size \( n_2 \) is taken and \( s_2 \) successes are observed. In this case, the posterior distribution of \( (p_r|s_2) \) is also Beta distribution with parameters \( [ s_2+a_1; n_2-s_2+b_1] \) where \( a_1=a+s_1 \) and \( b_1=n_1-s_1+b \) are parameters of the prior distribution of \( (p_r \mid s_1) \). The recalibrated estimate of \( p_r \) including the new information is given by \( \hat{p}_r^{(2)} = (s_1 + a_1)/(n_1 + a_1 + b_1) \). And this adjustment makes Bayesian estimate be upgraded at each new sample and after \( j \)-th recalibrations, Bayesian estimate for \( p_r \) storing the earlier informations is given as

\[
\hat{p}_r^{(j)} = \frac{\sum_{i=1}^{j} s_i + a}{\sum_{i=1}^{j} n_i + b}
\]

To illustrate the recalibration procedure, simulations are conducted. Samples of sizes equal to 5 and 10 are taken two hundred times from normal, exponential,
gamma and Weibull distributions. The parameters of these distributions are displayed in Table 2. Note that the value of $p_r$ under Weibull distribution was obtained by numerical integration and its value is not exact. The estimate of $p_r$ at the first sample was obtained under a non-informative Beta prior distribution, that is $a=1$ and $b=1$. For each new sample, the reliability $p_r$ is recalibrated and its Bayesian confidence interval at 95% are also obtained. Figures 1-4 show the results from these simulations. According to Figures 1-4, data drawn from normal distribution provide us stable Bayesian estimates before the first hundred sample; for others distributions the stability is reached before the second hundred samples. As expected, the width of confidence limits became narrower as more samples are extracted.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Strength(X)</th>
<th>E(X)</th>
<th>Stress(Y)</th>
<th>E(Y)</th>
<th>$p_r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>$\mu=80, \sigma=22.4$</td>
<td>80</td>
<td>$\mu=20; \sigma=10$</td>
<td>20</td>
<td>0.99492</td>
</tr>
<tr>
<td>Exponential</td>
<td>$\lambda=0.0125$</td>
<td>80</td>
<td>$\lambda_r=0.05$</td>
<td>20</td>
<td>0.8</td>
</tr>
<tr>
<td>Gamma</td>
<td>$a=80; \gamma_r=2$</td>
<td>160</td>
<td>$b=20; \gamma_r=2$</td>
<td>40</td>
<td>0.896</td>
</tr>
<tr>
<td>Weibull</td>
<td>$\eta_s=80; \beta_r=2$</td>
<td>70.9</td>
<td>$\eta_r=20; \beta_r=2$</td>
<td>17.7</td>
<td>&gt; 0.90</td>
</tr>
</tbody>
</table>

4 Conclusions

The reliability estimation proposed in this study relies on two aspects. The first is a conservative result which can be applied only when strength and stress are normally or lognormally distributed. Although this restriction, the lower bounds of the reliability are easily handled by this procedure. For other distributions the sampling distribution of eqn(7) is still a question to be answered. The second is about the choice of the prior distribution for $p_r$. Here we adopted Beta distribution as a prior distribution. The main reasons for this choice are: Beta densities can take different forms when values of the parameters $a$ and $b$ are appropriately chosen; it is conjugate with the parameter $p_r$, that is, the posterior distributions of $p_r$ has also Beta distribution and this property turns easy to obtain Bayesian estimates. Choices other than Beta distribution as prior distributions certainly are alternatives but the conjugate property mentioned before can not be hold and to obtain Bayesian point estimate and its confidence intervals may have to appeal to numerical methods such as Monte Carlo Markov Chain methods. And to close this section, it is important to remind that all these results are valid thinking on stable processes.

References


Figure 1: Normal distribution - Bayesian estimates and confidence limits of $p_f$
Figure 2: Exp. distribution-Bayesian estimates and confidence limits of $p_t$
Figure 3: Gamma distribution- Bayesian estimate and confidence limits of $p_r$. 
Figure 4: Weibull distribution - Bayesian estimate and confidence limits of \( p_t \)