ISSN 1407-8015 TRANSPORT AND ENGINEERING
MAŠĪNZINĀTNE UN TRANSPORTS 2008-8015

TRANSPORT. AVIATION TRANSPORT
TRANSPORTS. AVIĀCIJAS TRANSPORTS

NOMINATION OF SPECIFIED LIFE USING BAYES-FIDUCIAL APPROACH

KONKRĒTA RESURSA NOTEIKŠANA, IZMANTOJOT BAIESA-FIDUCIALA METODI

Yu. Paramonov, Professor, Dr.habil.sc.ing
Aviation Institute, Riga Technical University
Address: Lomonosova 1, Riga, LV-1019, Latvia
Phone: +371 7255394; Fax: +371 7089990
E-mail rauprm@junik.lv

Keywords: aircraft reliability, statistical methods, dataset processing

Introduction

Any detail, element or assembly, which contributes significantly to carring flight, ground, pressure or control loads and whose failure could affect the structural integrity necessary for the safety of the airplane is classified as a structural significant item (SSI) (see [1]).

In this paper we consider only those SSI the operation reliability of which is ensured by discarding the SSI from service, if its service life exceeded the Retirement or Specified Life (SL), t_{SL}, measured, for example, in numbers of flights (the term "specified life" we'll use as more short equivalent of "maximum permitted life", defined by WORLD AIRLINES TECHNICAL OPERATION GLOSSARI (ATA/IATA/ICCAIA) as the time specified by an appropriate authority after which a particular item must be removed from service). Typical example of this type of SSI is landing gear (made of high strength steel by the use of welding) nearly for all types of aircrafts (by the way, accordingly to [2] a failure of the landing gear was the most common fatigue problem and accounted for 37% of the accidents involving fatigue fracture).

It should be mentioned that SL can be chosen as (1) some number from $[0,\infty)$ and as (2) some number from set $\{0,\ t^*_{SL}\}$. This corresponds to

(1) nomination of Specified Life, t_{SL},

(2) rejection or acceptance of predetermined (required) Specified Life, t^*_{SL}.

In this paper we consider only problem (1). For solution of this problems we use definition of so called p-bound for random variable, which was offered by author some years ago [Paramonov Yu.M., 1992, 1999].

1. The problem statement

Let $X = (X_1, X_2, ..., X_n)$ be a vector with cumulative distribution function (c.d.f.) $F_X(x, \theta)$, Z is random variable with c.d.f. $F_Z(z, \theta)$. Of the parameter θ, which is the same for both distributions, it is assumed known only that it lies in a certain set Ω, the parameter space. The function $\tau(x)$, where x denotes the observed value of vector X, is called a p-bound for random variable Z, if

$$\sup_\theta P_\theta \{Z < \tau(X)\} = p \, , \; \theta \in \Omega \, . \qquad (1)$$

There are two very important special types of p-bounds:
(1) parameter-free (p.f. p-bound), when

$$P_\theta \{Z < \tau(X) = p\} \text{ for all parameters } \theta \in \Omega \, , \qquad (2)$$

(2) right-hand binary (r.h.b. p-bound), when for each possible value x of X, function $\tau(x)$ assigns only one of two decisions:

$$\tau(x) = -\infty \text{ if } x \notin S \, ; \; \tau(x) = \tau^* \text{, if } x \in S^* \, , \qquad (3)$$

where τ^* is some number, S^* and S are two complementary regions of the sample space as S_0 and S_1 in the problem of hypotheses testing [3].

It is clear, that p.f. p-bound definition coincides with the definition of both lower prediction limit for future observation and a similar p-expectation tolerance limit. And we see also a close connection of the r.h.b. p-bound definition with the problem of statistical hypothesis testing. In this paper we use both names "p-bound" (because of its versatility) and "lower prediction limit" (in case of comparison with results of others authors).

If $x = (x_1, ..., x_n)$ is result of full-scale fatigue test of airframe $Z = \min(Y_1, ..., Y_m)$ smallest fatigue life of airplane airframe in operation then p.f. p-bound $\tau(x)$ defines the SL for this park of airplanes such that probability of at least one airplane failure in operation is equal to p. Of course the value p should be very small. So the problem of SL nomination is the problem of a choice of p.f. p-bound $\tau(x)$, which is some statistical decision function.

A lot of papers discussed the problem of calculation lower prediction interval. Only for the case of Weibull distribution of lifetime the 15 references can be found in [4]. As a rule the offered methods are based on the use Monte Carlo method. In recently published paper [9], also devoted to the Weibull distribution, analytical method is studied but (as author of [9] thinks) for calculation of conditional lower prediction interval (we show here that really this is unconditional lower prediction interval). In this paper we consider the use of Bias-Fiducial method for distributions with location and scale parameters. This method was offered in [5,6]. Its application specifically for the p-bound calculation for distribution with location and scale parameters was discussed in [7, 8]. But in [7,8] there are not numerical examples of using this method. In this paper we consider specific examples for (log) normal and (Weibull) smallest extreme value (sev) distribution. Numerical comparison of these results with already published results is given also.

2. P–bound for distributions with location and scale parameters

Let

$$F_{X_i}(x, \theta) = F_{\overset{o}{X}}\left(\frac{x - \theta_0}{\theta_1}\right), \; i = 1, ..., n \, , \qquad (4)$$

$$F_Z(x, \theta) = F_{\overset{o}{Z}}\left(\frac{x - \theta_0}{\theta_1}\right),$$

where $F_{\overset{o}{X}}(x)$, $F_{\overset{o}{Z}}(x)$ are known c.d.f. of $\overset{o}{X}$, $\overset{o}{Z}$, θ_0, θ_1 – are unknown location and scale parameters.

This means that X_i can be described by the structural formula

$$X_i = \theta_0 + \theta_1 \overset{o}{X}_i, \; i = 1, ..., n. \tag{5}$$

Vector $X = (X_1, ..., X_n)$ is used for estimation of parameters θ_0, θ_1. Let these random variables, estimates of θ_0, θ_1, can be described by the similar structural formulas

$$\hat{\theta}_0 = \theta_0 + \theta_1 \overset{o}{\theta}_0, \; \hat{\theta}_1 = \theta_1 \overset{o}{\theta}_1, \tag{6}$$

where $\overset{o}{\theta}_0$, $\overset{o}{\theta}_1$ – are random variables –estimates of θ_0, θ_1, corresponding to a sample of the same size *n*, but for the case when $\theta_0 = 0$, $\theta_1 = 1$. It is well known that estimates of maximum likelihood (ML), moments and other methods has such structure.

P-bound τ can be considered as p-quantile estimate of cdf $F_Z((z-\theta_0)/\theta_1)$. If parameters θ_0, θ_1 are known then p-quantile τ_p is defined by equation

$$F_Z((\tau_p - \theta_0)/\theta_1) = p. \tag{7}$$

If parameters θ_0, θ_1 are not known but there is sample $x = (x_1, ..., x_n)$ then instead of parameters θ_0, θ_1 we can use the estimates of these parameters. The problem is a choice of theses estimates taking into account that they are random variable.

In [8] it is offered to use fiducial "estimates", because always they are function of sufficient statistics and so for small *p* they guarantee the maximum of expectation value of SL. Let us make this idea clear. Instead of vector $x = (x_1, ..., x_n)$ without loss of information we can consider vector $\varpi = (\hat{\theta}_0, \hat{\theta}_1, w_1, ..., w_{n-2})$, where $w_i = (x_i - \hat{\theta}_0)/\hat{\theta}_1$, $i = 1, ..., n-2$. Then fiducial distribution of random variables $\tilde{\theta}_0$, $\tilde{\theta}_1$ is defined by formula [Paramonov, 1992].

$$f_{\tilde{\theta}_0, \tilde{\theta}_1 | w_1, ..., w_n}(s_0, s_1) = h \frac{\hat{\theta}_1^{n-1}}{s_1^{n+1}} \prod_{i=1}^{n} f\left(\frac{\hat{\theta}_0 + \hat{\theta}_1 w_i - s_0}{s_1} \right) ds_0 ds_1, \tag{8}$$

where h is just normalization factor, $w_i = (x_i - \hat{\theta}_0)/\hat{\theta}_1$, $i = 1, ..., n$. (Note: w_{n-1}, , w_n are functions of vector ϖ).

Now instead of (7) let us consider equation

$$E\left(F\left((\tau - \tilde{\theta}_0)/\tilde{\theta}_1\right)\right) = p, \tag{9}$$

where pmf of random variables $\tilde{\theta}_0$, $\tilde{\theta}_1$ is defined by (8).

If in corresponding integral we change the variables of integration:

$$u_0 = (\hat{\theta}_0 - s_0)/\hat{\theta}_1, \; u_1 = \hat{\theta}_1 / s_1,$$

then instead of (9) we get equation

$$E\left(F\left((\tau - \hat{\theta}_0)/\hat{\theta}_1\right) - U_0)/U_1\right) = p, \tag{10}$$

where $\varpi = \left(\hat{\theta}_0, \hat{\theta}_1, w_1, ..., w_{n-2}\right)$ is some constant, but random variables U_0, U_1 has pmf

$$f_{U_0, U_1 | w_1, ..., w_n}\left(u_0, u_1\right) = h_w u_0^{n-2} \prod_{i=1}^{n} f\left(u_0 + w_i u_1\right), \tag{11}$$

where h_w is just normalization factor which depends only on vector $w = \left(w_1, ..., w_{n-2}\right)$.

Let us denote $\overset{0}{\tau} = \left(\tau - \hat{\theta}_0\right) / \hat{\theta}_1$ then instead of equation (10) we get equation

$$E\left(F\left(\left(\overset{0}{\tau} - U_0\right) \Big/ U_1\right)\right) = p. \tag{12}$$

If $\overset{0}{\tau}$ is solution of this equation then

$$\tau = \hat{\theta}_0 + \overset{0}{\tau} \hat{\theta}_1 \tag{13}$$

is p-bound for random variable Z, because equation (12) takes place for every vector $w = \left(w_1, ..., w_{n-2}\right)$, cdf of which does not depend on $\theta = \left(\theta_0, \theta_1\right)$. So if (12) is true then (14) is true also:

$$E_{W_1, ..., W_n} E_{U_0 U_1}\left(F\left(\overset{0}{\tau} - U_0\right) \Big/ U_1\right) = p. \tag{14}$$

Here $E_X\left(f(X)\right)$ is expected value of $f(X)$ in accordance with cdf of X.

It is very important that

1) $\overset{0}{\tau}$, as solution of equation (14), does not depend on true value of $\theta = \left(\theta_0, \theta_1\right)$ and we can set $\theta_0 = 0$, $\theta_1 = 1$.
2) Result does not depend on the choice of the statistics $\hat{\theta}_0$, $\hat{\theta}_1$ also because all vectors of the type $\varpi = \left(\hat{\theta}_0, \hat{\theta}_1, w_1, ..., w_{n-2}\right)$ and vector $x = \left(x_1, ..., x_n\right)$ have one-to-one correspondence. It is important only that $\hat{\theta}_0$, $\hat{\theta}_1$ have the structures defined by (6) but $w_i = \left(x_i - \hat{\theta}_0\right) / \hat{\theta}_1$, $i = 1, ..., n-2$.

3. Examples

3.1. Example 1

Let r.v. T, lifetime of some SSI (in cycles), has a lognormal distribution and $t = (t_1, t_2, t_3)$ = (45 952, 54 143, 65 440) is the sample from the same distribution. Then r.v. $X = \log(T)$ has a normal distribution $N(\theta_0, \theta_1^2)$ and $x = (x_1, x_2, x_3)$ = (10.735 10.899 11.089) is the sample from this distribution. The problem is to calculate the p.f. p-bound for independent r.v. $Z = \min(Y_1, ..., Y_m)$, where r.v. Y_i, $i = 1, ..., m$, has the normal distribution $N(\theta_0, \theta_1^2)$ also. We consider here only the case, when m = 1, because for this case there is general analytical solution (see, for example, [7, p. 172])

$$\tau(x) = \hat{\theta}_0 + \hat{\theta}_1 t_{n-1,p} (1 + 1/n)^{1/2}, \tag{15}$$

where

$$\hat{\theta}_0 = \bar{x}\,,\ \hat{\theta}_1 = (\sum (x_i - \bar{x})^2 /(n-1))^{1/2} \tag{16}$$

are estimates of expected value and standard deviation, $t_{k,q}$ is q-quantile of Student's distribution with k degree of freedom. So we can make comparison of this solution with the solution which we get using new approach.

For considered data, using equation (15) for $p = 0.01$ we calculate $t_{St} = \exp(\tau(x)) = 13\,162$, which is the value of p-bound for r.v. T on the base of the considered observations (t_1, t_2, t_3).

Now let us consider the offered new approach. For normal distribution the equation (11) has following form

$$f_{U_0, U_1 | w_1, \dots, w_n}(u_0, u_1) = h_w u_0^{n-2} \prod_{i=1}^{n} \varphi(u_0 + w_i u_1), \tag{17}$$

where $\varphi(x) = \exp(-x^2/2)/(2\pi)^{1/2}$.

After transformation the equation (12) has the following form

$$1 - a\left(\overset{0}{\tau}, \bar{z}, D_z\right)\Big/\Gamma((n-1)/2) = p\,, \tag{18}$$

where:

$a\left(\overset{0}{\tau}, \bar{z}, D_z\right) = \int_0^\infty u^{(n-3)/2} \exp(-u) \Phi\left((2u/D_z(n+1))^{1/2} \left(\bar{z} - \overset{0}{\tau}\right)\right) du$, $\bar{z} = \sum_1^n z_i / n$,

$D_z = \sum_{i=1}^{n} (z_i - \bar{z})^2 / n$, $\Gamma(\cdot)$ is gamma function, $\Phi(\cdot)$ is c.d.f. of standard normal distribution.

Now we consider two types of statistics $\hat{\theta}_0$, $\hat{\theta}_1$ which for considered data has following values:

$$\text{a) } \hat{\theta}_0 = \bar{x} = 10.908\,,\ \hat{\theta}_1 = \left(\sum (x_i - \bar{x})^2 /(n-1)\right)^{1/2} = 0.177 = 0.177, \tag{19 a}$$

$$\text{b) } \hat{\theta}_0 = x_{1,n} = 10.735,\ \hat{\theta}_1 = x_{n,n} - x_{1,n} = 0.354, \tag{19 b}$$

where $x_{i,n}$ is *i*-th order statistic of vector $x = (x_1, \dots, x_n)$.

In case a) we have $\overset{0}{\tau} = -7.889$, in case b) we have $\overset{0}{\tau} = -3.560$.

Corresponding values of p-bound for r.v. T on the base of observations (t_1, t_2, t_3) are:

$$t_a = \exp(\tau(x)) = 13\,523\,,\ t_b = \exp(\tau(x)) = 13\,050\,.$$

It seems that the difference between t_a, t_b and $t_{St} = 13\,162$ is produced only by the problem to get required calculation accuracy.

3.2. Example 2

Let we have the same sample $t = (t_1, t_2, t_3) = (45\,952,\quad 54\,143,\quad 65\,440)$ or $x = (x_1, x_2, x_3) = \; = (10.735,\quad 10.899,\quad 11.089)$ but r.v. T has a Weibull distribution and, correspondingly $X = \log(T)$ has distribution of smallest extreme value with cdf $F_X(x) = 1 - \exp(-\exp((x - \theta_0)/\theta_1))$. In this case the equation (12) has following form

$$1 - a\left(\overset{0}{\tau}, \bar{z}, D_z\right) / b(\bar{z}, D_z) = p\,, \tag{20}$$

where

$$a\left(\overset{0}{\tau}, \bar{z}, D_z\right) = \int_0^{\infty} u^{(n-2)} \left(\exp\left(-u\sum_{i=1}^{n} z_i\right) \Big/ \left(\sum_{i=1}^{n} \exp(uz_i) + m\exp\left(u\overset{0}{\tau}\right)\right)^{n} \right) du\,, \tag{21}$$

$$b(\bar{z}, D_z) = \int_0^{\infty} u^{(n-2)} \left(\exp\left(-u\sum_{i=1}^{n} z_i\right) \Big/ \left(\sum_{i=1}^{n} \exp(uz_i)\right)^{n} \right) du\,, \tag{22}$$

$$\bar{z} = \sum_{1}^{n} z_i / n\,,\; D_z = \sum_{i=1}^{n} (z_i - \bar{z})^2 / n\,.$$

For $m = 1$, $p = 0.01$, using statistics (19a) we get $\overset{0}{\tau} = -11.929$, using statistics (19b) we get $\overset{0}{\tau} = -5.424$. Corresponding values of p-bound for r.v. T on the base of observations (t_1, t_2, t_3) are:

$$t_a = \exp(\tau(x)) = 6\,616\,,\; t_b = \exp(\tau(x)) = 6\,752\,.$$

For $m = 500$, $p = 0.2$ using statistics (19a) we get $\overset{0}{\tau} = -12.889$, using statistics (19b) we get $\overset{0}{\tau} = -5.970$. Corresponding values of p-bound for r.v. T on the base of observations (t_1, t_2, t_3) are:

$$t_a = \exp(\tau(x)) = 5\,584\,,\; t_b = \exp(\tau(x)) = 5\,568\,.$$

Again, it seems that the difference between t_a and t_b is produced only by the problem to get required calculation accuracy.

Considered data really was considered in several papers and for $m = 500$, $p = 0.2$ Lowless (1973) obtained prediction limit of 5 623, Mee and Kushary (1994) – 5 225. The Mann and Saunders (1969) result was only 766. Our result really coincides with the results Lowless and Mee & Kushary.

References

1. Hall, J. and Goranson U. Structural damage tolerance of commercial jet transports // Airliner, July-September 1984 – p. 16 – 20.
2. Compbell, G.S. and Lahey, R. A. Survey of serious aircraft accidents involving fatigue fracture." Int. J. Fatigue. vol.5., 1984 – p. 25 – 30.
3. Lehmann, E.L. Testing statistical hypotheses. New York: John Wiley & Sons Inc.
4. Mee, R. W. Prediction limits for the Weibull distribution utilizing simulation // Computational statistics&Data Analysis. vol.17, 1994 – p. 327 – 336.

5. Nechval, K.N., Nechval, N.A., Berzins, G. and Purgailis, M. Planning Inspections in Service of Fatigue-Sensitive Aircraft Structure Components for Initial Crack Detection // Maintenance and Reliability, vol. 35, No. 3, 2007 – p. 76 – 80.
6. Paramonov, Yu.M. Using of Mathematical statistics method for aircraft reliability estimation and ensuring. (in Russian). Riga: RKIIGA. 1976 – 212 p.
7. Paramonov, Yu.M. (1992), Mathematical statistics application to the problem of aircraft fatigue life estimation and ensuring (in Russian). Riga: RAU, 1992 – 248 p.
8. Paramonov, Yu., Bayes-fiducial approach for quantile estimation and specified life nomination // Proceedings of the international conference STATISTICAL METHODS FOR BIOMEDICAL AND TECHNICAL SYSTEMS, Limassol, Cyprus, 2006 – p. 421 – 425.
9. Paramonov, Yu., Bayes-fiducial approach for quantile estimation and specified life nomination// ABSTRACTS of communications. 9th international Vilnius conference on PROBABILITY THEORY AND MATHEMATCAL STATISTICS, Vilnius, 25-30 June, 2006 – p. 253 – 254.

Paramonovs Ju. Konkrēta resursa noteikšana, izmantojot Baiesa-fiduciala metodi
Raksts ir veltīts lidmašīnas konstrukcijas spēku elementu noguruma sagrūšanas novēršanas problēmas risināšanai, izmantojot lidmašīnas ekspluatācijas laika ierobežošanu. Lidmašīnu nepieciešami ņemt no ekspluatācijas neatkarīgi no konstrukcijas stāvokļa, ja lidmašīnas ekspluatācijas laiks ir vienāds ar noteiktu resursu. Lai iegūt sakuma informāciju aplūkotai problēmai risināšanai tiek veikti lidmašīnas konstrukcijas noguruma izmēģinājumi. Izmēģinājumu gaitā lidmašīnas konstrukcijas spēku elementu noguruma ilgizturības ir fiksētas. Izmēģinājumu rezultāti ir apstrādāti, izmantojot modernas matemātiskās statistikas metodes. Šajā rakstā ir pieņemts ka ekspluatācijas laika ilgums jāizvēlē no nosacījuma noguruma sagrūšanas varbūtības ierobežošanas. Šajā gadījumā noteikt resurss ir lidmašīnas ekspluatācijā minimālai ilgizturībai p-robeža. P-robežas aprēķinam ir piedāvāta Baiesa-fiduciala metode. Izmantojot tādu metodi mēs vienmēr varam dabūt atrisinājumu kā pietiekamas statistikas funkciju. Aplūkotai problēmai šis atrisinājums dod maksimālu noteikta resursa sagaidāmu vērtību, ja p ir pietiekami maz lielums. Šajā rakstā skaitliskie piemēri ir aplūkoti, kad ir (log-) normālais un (Veibula-) minimāla gaidījuma lieluma sadalījumi, piemēru salīdzināšanas ar jau publicētiem rezultātiem ir izdarīti.

Paramonov Yu. Nomination of specified life using bayes-fiducial approach
The paper is devoted to solution of the problem of aircraft structural significant item fatigue failure prevention in the frame of fatigue safe-life approach. In this case the aircraft should be discarded from operation if the specified life is reach. The flights must be ceased independently on real airframe state. Full-scale fatigue tests must be carried out in order to get initial information for solution of considered problem. As the result of these tests we get the observations of fatigue life of aircraft structural significant items. For processing of these data the modern mathematical statistics method are used. In this paper it is assumed, that we should choose the specified life under the condition of fatigue failure probability limitation. In this case the specified life is the p-bound for the smallest fatigue life of aircraft in operation. For its calculation the Bayes-fiducial approach is offered, which always provides the solution as function of sufficient statistics and, correspondingly, the maximum of expected value of specified life if p is small enough. In this paper numerical examples for (log) normal and (Weibull) smallest extreme value (sev) distribution are studied. Numerical example comparison of this paper results with already published results is given also.

Парамонов Ю. Определение назначенного ресурса, используя Байес-фидуциальный метод
Статья посвящена решению проблемы предотвращения усталостного разрушения силовых элементов самолета за счёт ограничения длительности его эксплуатации величиной назначенного ресурса. По достижении назначенного ресурса эксплуатация самолёта прекращается независимо от его фактического состояния. Для получения исходной информации для решения рассматриваемой задачи проводятся натурные усталостные испытания, в ходе которых фиксируется долговечность силовых элементов самолета. Данные об усталостной долговечности обрабатываются, используя методы современной математической статистики. В настоящей статье принято, что допускаемая длительность эксплуатации должна быть выбрана из условия ограничения вероятности отказа. В этом случае назначенный ресурс является p-границей для минимальной долговечности самолёта в эксплуатации. Для её расчета предложен Байес-фидуциальный метод решения задачи. Используя этот метод, мы всегда можем получить решение как функцию достаточной статистики. Для рассматриваемой задачи это решение обеспечивает максимум математического ожидания назначенного ресурса, если p достаточно мало. Численные примеры приведены для (лог-) нормального распределения и (Вейбулла) распределения левого экстремума. Приведено сравнение с ранее опубликованными результатами.