

USE OF STATISTICAL REFERENCE METHOD TO IMPROVE FLIGHT SAFETY

Irina Lazareva¹, Vladimir Shestakov²

Aviation Institute of Riga Technical University, Lomonosova St. 1, LV-1019 Riga, Latvia
E-mails: ¹lira112@inbox.lv (corresponding author); ²shestakov@inbox.lv

Received 12 March 2012; accepted 11 May 2012



Irina LAZAREVA

Date and place of birth: 1986 in Riga.

Education: Fire Safety and Civil Protection College, technician (2007), Fire Safety and Civil Defence of Riga Technical University, engineer (2009), master's programme student, Aviation Institute of Riga Technical University.

Research interests: flight safety and life protection sciences.



Vladimir SHESTAKOV, Prof. Dr Habil

Date and place of birth: 1938 in Voronezh oblast, Russia.

Education: Riga Institute of Civil Aviation Engineers in 1963, Technical Aircraft and Engine Exploitation. Professor, doctor of engineering sciences (Dr Habil Sci. Eng.)

Present job: head of the Aircraft Maintenance Division, Riga Technical University, Aviation Institute.

Education: Riga Civil Aviation Engineering Institute, mechanical engineer (1963); high doctor degree at Aviation Leningrad Civil Aviation Academy (1984); doctor habilitus degree in engineering (1992), full member of International Academy of Ecology and Life Protection Sciences (1994).

Affiliations: International Academy of Ecology and Life Protection Sciences full member (academic). Riga Technical University (RTU) Promotion Council 'Air Transport Operations' member. Latvian Association of University Professors.

Research interests: professional exploitation of air transport, flight safety, ecology and life protection sciences, ecologically clean non-traditional kinds of transport.

Publications: 242 published works, 25 teaching and methodical books, 9 certificates of inventions, including 3 textbooks (in co-authorship) and 2 monographs.

Abstract. Statistical comparison as a test of hypotheses about the coincidence of distribution parameters is considered. This allows defining whether the discrepancy in estimations of the distribution parameters of safety indicators is significant or can be justified by inessential changes in the conditions of the experiment-random drifts. Each adverse event is assigned a weight that reflects the probability that it can result in an adverse outcome in a particular flight. Risk (hazard) assessment is represented as a product of the mean value of weight n_0 of adverse event K_n and frequency of occurrence F_n of events during the period of time analysed.

Keywords: safety of flights, statistics, safety parameter, accident, incident, human factor, technical operation, risk factors.

1. Introduction

Over the past few years, methods of functional safety analysis (FSA) and functional risk assessment (FNA) have become widely used in different industries to identify and analyse risks. These methods allow identifying, investigating and monitoring any hazard that can help to determine the necessary measures to reduce risks to acceptable levels (Center... 1992; Kletz 1999).

General approaches to solving these problems are available in standard IEC 61508. In the aviation industry, these principles are standardised by rules of airworthiness. They are used in the design and certification of aircraft as well as the harmonisation of the airworthiness requirements of different countries (FAA...; SAE...) An important element in monitoring flight safety is the evaluation of the effectiveness of measures being taken at the airline, using the results of flight safety analysis over a given period.

It seems that one of the options for such an evaluation may be the method of statistical comparison of safety parameters obtained on the basis of statistical data before an event (T1) and after it (T2). Methods of comparison of statistical hypotheses are widely used in statistical investigations (www.wikipedia.org), and in our case its use in no way conflicts with the principles laid down in the standards of airworthiness. To solve the problem, let us represent suspended statistics in the form of samples from some general populations. By comparing them, we can answer whether the sample belongs to a general population or to different ones. In other words, we need to test the hypothesis that the two series of experiments in which samples were obtained were performed in the same conditions.

If the conditions of the experiment have not changed, the activities carried out can be considered as ineffective and vice versa. In this case, the safety record is represented by experimentally obtained parameters of distributions of random variables. Coming from that, the task of statistical comparisons can be viewed as the task of testing hypotheses about the concurrency of the distribution parameters. Determining whether the difference in estimates of the distribution parameters of safety indicators is significant or whether it is due to small variations in experimental conditions, i.e. random drift, is required. Hence to evaluate the effectiveness of interventions, it seems natural to adopt criteria that are used to statistically compare or test hypotheses in mathematical statistics.

2. Testing hypotheses in mathematical statistics

Two hypotheses are considered for the tasks of testing in mathematical statistics (Cohen 1994; Hubbard, Armstrong 2006):

- suggested;
- competing.

A specially selected random variable, the exact or approximate distribution of which is known, is used to test a suggested hypothesis.

The *critical area* is a set of values of the criteria under which the suggested hypothesis is rejected. The *area of the hypothesis* is a set of values of the criteria under which the hypothesis is accepted. The *critical points* are the boundaries of the critical area.

A sufficiently small probability, significance level α , is given to find the critical area. Critical points are found based on the requirement that the veracity of the basic hypothesis of the sum of the probabilities that the criterion K will accept a value less than \tilde{K}_{1tca} or greater than \tilde{K}_{2tca} is equal to the accepted level of significance.

$$P(K < \tilde{K}_{1tca}) + P(K > \tilde{K}_{2tca}) = \alpha.$$

Let us consider the relative indicators, the number of dangerous events in the conditional unit of uptime (in hours or miles) as a safety indicator, which can be compared:

$$K_1 = \frac{n}{t},$$

where K_1 is marker; n – number of dangerous events; t – uptime.

In same way, the ratio of the number of dangerous events, n , to the number of flights, N , is the relative frequency of the occurrence of dangerous events. In compliance with this, we formulate comparative tasks for each indicator of flight safety.

3. Testing the hypothesis of equality of mathematical expectation $K_1^{(1)}$ and $K_1^{(2)}$ and the two random variables n_1 and n_2

Since hazardous events are infrequent and independent of each other, it is acceptable to assume that their distribution is subject to Poisson distribution (or the Poisson law of small numbers). In this case, the index of K_1 acquires the meaning of mathematical expectation in the law of Poisson. In general, to test the hypothesis of equality of mathematical expectation \bar{x} and \bar{y} and two random variables X and Y , let us assume a random variable:

$$Z' = \frac{\bar{x} - \bar{y}}{\sqrt{D(\bar{x}) + D(\bar{y})}},$$

where \bar{x} and \bar{y} are evaluations of the mathematical expectation of values of X and Y ; $D(\bar{x})$, $D(\bar{y})$ – evaluations of the mathematical expectation of values \bar{x} and \bar{y} .

The criterion Z' is distributed approximately normally with parameters $M(Z') = 0$ and $\sigma(Z') = 1$, and the law on the distribution of general sets can be arbitrary provided the samples are independent and the series of samples is not smaller than 30.

With regard to the conditions of our problem $\bar{x} = K_1^{(1)}$, $\bar{y} = K_1^{(2)}$ and to find $D(K_1^{(1)})$ and $D(K_1^{(2)})$, we proceed as follows: divide period T into r parts ΔT_j ($j = 1, \dots, r$). On each count of the ΔT_j , calculate Δn_j and Δt_j ($j = 1, \dots, r$), that is, obtain r values of the random variable n . At each interval ΔT_j , we calculate the estimate.

$$K = \frac{\Delta n_j}{\Delta t_j} \quad (j = 1, \dots, r).$$

Further, it is easy to calculate the evaluation of dispersion K_1 :

$$D(K_1) = \frac{\sum_{j=1}^r (k_{1j} - k_1)^2}{(r-1)}.$$

In light of the abovementioned calculations, criterion Z' will be equal to:

$$Z'_{ob} = \frac{\hat{E}_1^{(1)} - K_1^{(2)}}{\sqrt{\frac{\sum_{j=1}^{r_1} \left(\frac{\Delta n_j^{(1)}}{\Delta t_{1j}} - K_1^{(1)}\right)^2}{r_1 - 1} + \frac{\sum_{j=1}^{r_2} \left(\frac{\Delta n_j^{(2)}}{\Delta t_{2j}} - K_1^{(2)}\right)^2}{r_2 - 1}}},$$

Z'_{ob} – values of the criterion observed as a result of the experiment,

$$\sum_{j=1}^r \Delta n_j = n,$$

$$\sum_{j=1}^r \Delta t_j = t,$$

r_1 and r_2 – some arbitrary numbers whose values should be rather large (12+15).

The critical area is defined depending on the type of competing hypotheses. Let us formulate the decision rule for testing the basic hypothesis $K_1^{(1)} = K_1^{(2)}$ by the criterion of $K = Z'_{ob}$ at the given significance level α :

Case 1. Competing hypothesis $K_1^{(1)} = K_1^{(2)}$; the observed value of criterion Z'_{ob} is calculated.

On the table of the Laplace function, the critical point of the equation is:

$$F(Z_{tca}) = (1 - \alpha) / 2,$$

F – Laplace function.

If $Z'_{ob} / < Z_{tca}$ – there is no reason to reject the suggested hypothesis.

If $Z'_{ob} / > Z_{tca}$ – the suggested hypothesis is rejected.

Case 2. Competing hypothesis $K_1^{(1)} > K_1^{(2)}$; the observed value of criterion Z'_{ob} is calculated.

$$F(Z_{kp}) = (1 - 2\alpha) / 2,$$

if $Z'_{ob} / < Z_{tca}$ – there is no reason to reject the suggested hypothesis.

If $Z'_{ob} / > Z_{tca}$ – the suggested hypothesis is rejected.

4. A comparison of the relative frequencies of the occurrence of hazardous events

It is necessary to establish whether there is a significant discrepancy between the values of $K_2^{(1)}$ and $K_2^{(2)}$ that represents the relative frequencies of the occurrence of hazardous events.

$$K_2^{(1)} = \frac{n_1}{N_1}, \quad K_2^{(2)} = \frac{n_2}{N_2}.$$

In this case, as a criterion for testing the main hypothesis, $K_2^{(1)} = K_2^{(2)} = K_2$, we will assume a random variable:

$$U = \frac{\frac{n_1}{N_1} - \frac{n_2}{N_2}}{\sqrt{K_2(1 - K_2) \left(\frac{1}{N_1} + \frac{1}{N_2}\right)}}.$$

It is known that U is distributed approximately normally with parameters $M(U) = 0$ and $\sigma(U) = 1$; since the probability of K_2 is unknown, replace it with the maximum likelihood estimate:

$$K_2 = \frac{(n_1 + n_2)}{(N_1 + N_2)}.$$

As a result, we find a working formula for calculating the observed value of the criterion.

$$U_{ob} = \frac{\frac{n_1}{N_1} - \frac{n_2}{N_2}}{\sqrt{\frac{n_1 + n_2}{N_1 + N_2} \cdot \left(1 - \frac{n_1 + n_2}{N_1 + N_2}\right) \cdot \left(\frac{1}{N_1} + \frac{1}{N_2}\right)}}.$$

Let us consider a rule for testing the hypothesis of the equality of probabilities (relative frequency) of one event in the two samples according to the Poisson distribution (or Poisson law of small numbers) for a given significance level α .

Case 1. Competing hypothesis $K_2^{(1)} \neq K_2^{(2)}$.

Calculation of the observed value of criterion U_{ob} is carried out.

According to the table of the Laplace function, we find critical point U_{kp} according to equality $F(U_{tca}) = (1 - \alpha) / 2$.

If $U_{ob} / < U_{tca}$ – there is no reason to reject the suggested hypothesis.

If $U_{ob} / > U_{tca}$ – the suggested hypothesis is rejected.

Case 2. Competing hypothesis $K_2^{(1)} > K_2^{(2)}$.

Finding the critical point of right-handed critical area accordingly to equation

$$F(U_{tca}) = (1 - \alpha) / 2$$

if $U_{ob} / < U_{tca}$ – there is no reason to reject the suggested hypothesis.

If $U_{ob} / > U_{tca}$ – the suggested hypothesis is rejected.

5. Assessment of hazard level of adverse factors

The most common method of assessing risk is expert estimates (Ivanek 1996; Monks 1992; Krokhin *et al.* 1987). However, using assessment of risks based on expert analysis as a rule makes it difficult to obtain reliable conclusions regarding the significance of the events studied and their causes. Additionally, the constantly evolving changes in the transportation system create the necessity for a systematic correction of the results in experts' evaluations, i.e., the need to conduct follow-up examinations.

It seems that good reliability and efficiency of risk assessment may be obtained based on

- assessing the severity of each specific adverse event,
- establishing a list of all possible causes (factors) that led to the particular adverse event.

This will allow the parameter of weight to be attributed to each adverse event. Weight reflects the probability that there could end up being an adverse outcome on this particular flight. The output information can be presented in the following form.

Table 1 shows that during this flight two independent events, C₂ and C₃, took place, and one could result in an accident with the probability of 0.4, while the other could end with an air crash.

Table 1. Presentation of output information on events

C	C ₁	C ₂	C ₃	C ₄		C _n
D	0	1	0.4	0		0
C	F						
	F ₁	F ₂	F ₃	F ₄	...	F _m	
	α ₁	α ₂	α ₃	α ₄	...	α _m	
C ₁	C ₁₁	α ₁₁₁	α ₂₁₁				α _{m11}
	C ₁₂	α ₁₁₂	α ₂₁₂				α _{m12}
	C _{1c}	α _{11c}	α _{21c}				α _{m1c}
C ₂	C ₂₁	α ₁₂₁	α ₂₂₁				α _{m21}
	C ₂₂	α ₁₂₂	α ₂₂₂				α _{m22}
	C _{2s}	α _{12s}	α _{22s}				α _{m2s}
C _n	C _{n1}	α _{1n1}	α _{2n1}				α _{mn1}
	C _{n2}	α _{1n2}	α _{2n2}				α _{mn2}
	C _{nt}	α _{1nt}	α _{2nt}				α _{mnt}

Providing output information in this case can be recommended in the form of Table 2.

Table 2. Presentation of output information for reasons (factors)

C ₂	F ₁	F ₂	F ₃	F ₄	...	F _m
C ₂	0.3	0	0.5	0	0	0.7
C ₃	F ₁	F ₂	F ₃	F ₄		F _m
C ₃	0	1	0	0		0

From Tables 1 and 2, it is evident that with different probability the cause of event C₂ could be factors F₁, F₂, F₃, and C₃, and event C₃ was caused by factor F₂.

Assuming that the output information of the investigation of incidents is input information for the analysis of the safety of flights, the assessment of risk of adverse events can be presented as a product of the average severity of nth adverse event K_n and the frequency of occurrence of F_n events during the period of time analysed.

$$D_n = K_n \cdot F_n .$$

Respectively:

$$K_n = \frac{K_{n1} + K_{n2} + \dots + K_{nr}}{r} ,$$

where K_{n1}, K_{n2} are values of the coefficients of the weight of the nth adverse event in the first, second and nth flight; r – number of nth adverse events during the period of time analysed.

Presentation of investigation data for the period of time analysed

$$F_n = \frac{r}{V + p + \dots + r} = \frac{r}{N} ,$$

where N is a total number of adverse events during the period of time analysed. Finally the formula for determining the risk of the nth event can be presented as:

$$D_n = K_n \cdot F_n = \frac{(K_{n1} + K_{n2} + \dots + K_{nr}) \cdot r}{r \cdot N}$$

or

$$D_n = \frac{\sum_{w=1}^r K_n \cdot W}{N} .$$

Assessment of the hazardous factors of adverse events is based on a similar scheme and can be represented by two groups of values of the hazards:

- the hazard of the nth factor in each adverse event /b_{mn}/;
- the absolute hazard of the nth factor of /b_m/.

The hazard of the mth factor in the nth adverse event is a ratio of the sum of hazards (risks) of the mth factor

in each case, in every n^{th} adverse event / α_{mnt} / to the number of factors.

$$L_{mn} = \frac{(\alpha_{mn1} + \alpha_{mn2} + \dots + \alpha_{mnt})}{mt},$$

where t is a number of adverse m -type events that occur during the period of time analysed;

m – the number of different factors.

The total hazard of the m^{th} factor can be expressed as the ratio of the sum of individual hazards of the m^{th} factor to the total number of adverse events.

$$L_m = \frac{(L_{m1} + L_{m2} + \dots + L_{mn})}{n}$$

6. Conclusion

This approach to risk assessment allows forming an array of risk assessment data at the stage of the investigation of individual events by the possibility of their occurrence in the form of accidents, as well as by estimating causal factors for the possible formation of one event or another and developing appropriate measures to evaluate the significance of each undesirable event or factor.

References

- Center for Chemical Process Safety Guidelines for Hazard Evaluation Procedures with Worked Examples. 1992. 2nd ed. Wiley-American Institute Of Chemical Engineers. ISBN 0-8169-0491-X.
- Cohen, J. 1994. The Earth is round, *American Psychologist* 49(12): 997–1003.
<http://dx.doi.org/10.1037/0003-066X.49.12.997>
- FAA AC25.1309 [online]. Available from Internet: www.faa.gov.
- Hubbard, R.; Armstrong, J. S. 2006. Why we don't really know what statistical significance means: implications for educators, *Journal of Marketing Education* 28(2): 114.
<http://dx.doi.org/10.1177/0273475306288399>
- Ivanek, J. 1996. *An Expert System Recommending Suitable Mathematical Decision Method. Computers and Artificial Intelligence*.
- Kletz, T. 1999. *Hazop and Hazan*. 4th ed. Taylor & Francis. ISBN 0-85295-421-2.
- Krokhin, Z. T.; Skripnik, F. I.; Shestakov, B. 1987. *Inzhenerno-organizatsionnye osnovy obespecheniia bezopasnosti poliotov*. Moskva: Transport (in Russian).
- Monks, J. G. 1992. *Operations Decision Making. Operations Management*. New York.
- SAE ARP 4761, EASA CS-25, 1309 [online]. Available from Internet: www.aaiib.gov.uk/