BSCE 20 / WP 16
Helsinki, 21-25 May 1990

ON PREDICTING ACCIDENTS AND SERIOUS INCIDENTS TO CIVIL AIRCRAFT DUE TO BIRD STRIKES IN A FUTURE TIME PERIOD FROM KNOWN OBSERVATIONS

Presented by Nikolai A. Nechval

BSCE 20/WP
Helsinki, 21-25 May 1990

ON PREDICTING ACCIDENTS AND SERIOUS INCIDENTS TO CIVIL AIRCRAFT DUE TO BIRD STRIKES IN A FUTURE TIME PERIOD FROM KNOWN OBSERVATIONS

Nikolai A. Nechval

Civil Aviation Engineers Institute
Riga, USSR

SUMMARY

The problem of predicting the number of accidents and serious incidents to civil aircraft due to bird strikes during a future time period with the specified number of aircraft movements, knowing accidents and serious incidents during time intervals (with the known numbers of aircraft movements, respectively) in the past, is considered. It is known that in many familiar situations, the predictive estimators based on the principles of maximum likelihood and of minimum variance unbiased estimation are uniformly worst among all predictive estimators which one would consider using. In this paper, we suggest (as a particular possibility) the use of uniformly undominated predictive estimators and give the conditions that a predictive estimator must satisfy in order that it be uniformly undominated. It is assumed that accidents and serious incidents to civil aircraft due to bird strikes follow a binomial distribution. An illustrative example is presented.

1. INTRODUCTI

The binomial
on of the num
craft due to
fied number o
ned as: (a) l
on of aircraf
(e) uncontain
hole eg shatt
structural da
eg complete o
to helicopter
(1982) contai
due to bird s

Whatever may
decided to ac
will be found
binomial dist
sed on the ob
with the know

One of the pr
accidents and
strikes durin
aircraft move
ring time int
respectively)

2. PROBLEM ST

Frequently on
variable rath
this is to es
parameter) an
on; i.e., to
for the varia
dure if one a
that from thi
This however
uniform undom

Consider a pa
tion P_θ (with
family $P°$ of
to predict, s
value x of X
d(x), and the
that the loss
expectation o
infinite).

The risk asso
le) d(X) is d

$R_\theta(d) = E_\theta$

The choice of

1. INTRODUCTION

The binomial model can be widely used to describe the distribution of the number of accidents and serious incidents to civil aircraft due to bird strikes in a future time period with the specified number of aircraft movements. Here "serious" has been defined as: (a) loss of life, (b) injury to occupants, (c) destruction of aircraft, (d) damage/loss/shutdown of more than one engine, (e) uncontained engine failure, (f) fire, (g) significant sized hole eg shattered radome, holed windscreen, holed wing, (h) major structural damage, (i) particularly unusual or dangerous features eg complete obscuring of vision, multiple loss of system, damage to helicopter blades or transmissions. The paper of J. Thorpe (1982) contains brief details of accidents and serious incidents due to bird strikes world wide up to and including 1980.

Whatever may be the reasons for adopting a binomial model, having decided to accept such a model, results derived in this paper will be found appropriate. In practice, the true parameter of the binomial distribution is not known, and the inference must be based on the observed bird strike data during certain time periods with the known numbers of aircraft movements, respectively.

One of the problems considered here is to predict the number of accidents and serious incidents to civil aircraft due to bird strikes during a future time period with the specified number of aircraft movements, knowing accidents and serious incidents during time intervals (with the known numbers of aircraft movements, respectively) in the past.

2. PROBLEM STATEMENT

Frequently one is interested in estimating the value of a random variable rather than that of a parameter. A customary method for this is to estimate the expectation of the random variable (a parameter) and then to "identify" the variable and its expectation; i.e., to use the estimate of the expectation as a prediction for the variable. As we shall see below one is led to this procedure if one adopts the point of view of unbiased estimation, so that from this point of view prediction poses no new problem. This however is no longer true when one employs the principle of uniform undomination (see, in this connection, Nechval (1988)).

Consider a pair X,Y of random variables having a joint distribution P_θ (with $\theta \in \Theta^\circ$ (parameter space)) belonging to a parametric family P° of distributions. It is desired to use the observed X to predict, say, g(Y) where g is a some function of Y. If the value x of X is observed one makes an predictive estimate, say d(x), and thereby incurs a loss of $W[g(y),d(x)]$. We shall assume that the loss function is nonnegative. It then follows that the expectation of the loss will always exist (although it may be infinite).

The risk associated with the predictive estimator (decision rule) d(X) is defined to be the expected loss, as given by

$$R_\theta(d) = E_\theta(W[g(Y),d(X)]) \quad . \tag{1}$$

The choice of predictive estimator, d(X), should then be made according to the risk function (1). [illegible] a particular possibility [illegible] suggest the use of uniformly undominated predictive estimators.

A predictive decision rule d_1 is said to be uniformly better than a predictive decision rule d_2 if $R_\theta(d_1) < R_\theta(d_2)$ for all $\theta \in \Theta^\circ$. A predictive decision rule d is said to be uniformly undominated if there exists no predictive decision rule uniformly better than d. Otherwise, it is uniformly dominated. The examples described in [illegible] (1988) may be of interest in that there the maximum likelihood and unbiased decision rules are uniformly worst among all decision rules which one would consider using.

The conditions that a predictive decision rule must satisfy in order that it be uniformly undominated are given by the following theorem.

<u>Theorem 2.1</u>. <u>Characterization of the uniformly undominated decision rules</u>. Let $(q_s; s=1,2, \dots)$ be a sequence of the prior distributions on the parameter space Θ°. Suppose that $(d_s; s=1,2, \dots)$ and $(Q(q_s,d_s): s=1,2, \dots)$ are the sequences of Bayes predictive decision rules and prior risks respectively. If there exists a predictive decision rule d^* such that its risk function $R_\theta(d^*)$, $\theta \in \Theta^\circ$, satisfies the relationship

$$\lim_{s \to \infty} [Q(q_s,d^*) - Q(q_s,d_s)] = 0, \tag{2}$$

where

$$Q(q_s,d) = \int_\Theta R_\theta(d)\, q_s(d\theta), \tag{3}$$

then d^* is an uniformly undominated predictive decision rule.

<u>Proof</u>. Suppose d^* is uniformly dominated. Then there exists a predictive decision rule d'' such that $R_\theta(d'') < R_\theta(d^*)$ for all $\theta \in \Theta^\circ$. Let

$$e = \inf_{\theta \in \Theta^\circ} [R_\theta(d^*) - R_\theta(d'')] > 0. \tag{4}$$

Then

$$Q(q_s,d^*) - Q(q_s,d'') \geq e. \tag{5}$$

Simultaneously,

$$Q(q_s,d'') - Q(q_s,d_s) \geq 0, \tag{6}$$

$s=1,2, \dots$, and

$$\lim_{s \to \infty} [Q(q_s,d'') - Q(q_s,d_s)] \geq 0. \tag{7}$$

On the other hand,

$$Q(q_s,d'') - Q(q_s,d_s) = [Q(q_s,d^*) - Q(q_s,d_s)] - [Q(q_s,d^*) - Q(q_s,d'')] \leq [Q(q_s,d^*) - Q(q_s,d_s)] - e \tag{8}$$

and

$\lim_{s \to \infty} [Q(q_s$

This contradi
predictive de

<u>Corollary 2.1</u>
ction is cons
on rule.

Suppose now t

$W[g(y),d(x$

Consider the
ness. A predi

$E_\theta(d(X)) =$

Subject to un

$E_\theta([g(Y) -$

But $Var_\theta(g(Y)$
minimizing (f
ces to that o
nimum varianc
triction to u
g(Y) is the s
expect that a
will coincide
the case sinc
in the two ca
lems are rath
distribution
account the d
rious prior d
g(Y) when θ i

The main purp
dominated pre
lems.

3. <u>PREDICTION</u>
<u>TO CIVIL A</u>

Consider an o
rious inciden
binomial dist
observation f
craft movemen
is recorded.
number of air
rious inciden
predict the v
variable $X(m_1$
ributions

and

$$\lim_{s \to \infty} [Q(q_s, d'') - Q(q_s, d_s)] < 0. \tag{9}$$

This contradiction proves that d* is an uniformly undominated predictive decision rule.

Corollary 2.1.1. A Bayes predictive decision rule, whose risk function is constant, is an uniformly undominated predictive decision rule.

Suppose now that X and Y are independent and that

$$W[g(y), d(x)] = [g(y) - d(x)]^2. \tag{10}$$

Consider the problem first from the point of view of unbiasedness. A prediction could reasonably be called unbiased if

$$E_\theta(d(X)) = E_\theta(g(Y)). \tag{11}$$

Subject to unbiasedness, the risk is given by

$$E_\theta([g(Y) - d(X)]^2) = \mathrm{Var}_\theta(g(Y)) + \mathrm{Var}_\theta(d(X)). \tag{12}$$

But $\mathrm{Var}_\theta(g(Y))$ is a known function of θ, and hence the problem of minimizing (for a particular θ) the expected squared error reduces to that of finding an unbiased estimate of $E_\theta(g(Y))$ with minimum variance at θ. In a similar way one sees, without any restriction to unbiased predictions, that the Bayes prediction for g(Y) is the same as the Bayes estimation for $E_\theta(g(Y))$. One might expect that as in the unbiased theory the predictive estimate will coincide with the unbiased estimate. This however is not the case since the prior distributions that give constant risk in the two cases will usually be distinct. In fact the two problems are rather different in that the "least favourable" prior distribution for the prediction problem must not only take into account the difficulty of finding the correct value of θ for various prior distributions but also the difficulty of predicting g(Y) when θ is known.

The main purpose of the present paper is to obtain uniformly undominated predictive estimators for a number of specific problems.

3. PREDICTION OF THE NUMBER OF ACCIDENTS AND SERIOUS INCIDENTS TO CIVIL AIRCRAFT DUE TO BIRD STRIKES

Consider an ornithological situation in which accidents and serious incidents to civil aircraft due to bird strikes follow a binomial distribution with parameter p. The situation is under observation for time interval with the known number m_1 of aircraft movements, where $X(m_1)$ of accidents and serious incidents is recorded. In some future time interval with the specified number of aircraft movements m_2, the number of accidents and serious incidents is denoted by $Y(m_2)$. The specific problem is to predict the value of a random variable $Y(m_2)$ observing a random variable $X(m_1)$, where $X(m_1)$ and $Y(m_2)$ have the probability distributions

$$f(X(m_1)=x;p) = \binom{m_1}{x} p^x(1-p)^{m_1-x}, \quad x=0,1,2, \dots ,m_1, \tag{13}$$

and

$$f(Y(m_2)=y;p) = \binom{m_2}{y} p^y(1-p)^{m_2-y}, \quad y=0,1,2, \dots ,m_2, \tag{14}$$

respectively, which are dependent on the same (unknown) parameter p, $0 \leq p \leq 1$. Here we consider the situation when a statistician predicts systematically the value of Y observing $X_1(m_1)$, $X_2(m_2)$, ... ,$X_n(m_n)$ at the stages 1,2, ... ,n, respectively, and when the loss function is the sum of losses at the particular stages. There are many problems of this type which can be stated and solved (and some of them have been actually solved). We restrict ourselves to one of them.

Let $X_1(m_1)$, ... ,$X_{n+1}(m_{n+1})$ be independent random variables with the distributions

$$f(X_i(m_i)=x_i) = \binom{m_i}{x_i} p^{x_i}(1-p)^{m_i-x_i}, \quad x_i=0,1,2, \dots ,m_i; \quad i=1(1)n+1. \tag{15}$$

Let

$$\underline{X}_k = \sum_{i=1}^{k} X_i(m_i). \tag{16}$$

We want to predict the random variable $\underline{Y}_n$ on the basis of observations of $\underline{X}_1, \underline{X}_2, \dots ,\underline{X}_n$. Since at the kth stage we know the values of random variables $X_1(m_1)$, ... ,$X_k(m_k)$ and $\underline{X}_k$ is sufficient for p, it is sufficient to predict the values of

$$\underline{Y}_k = \sum_{i=k+1}^{n+1} X_i(m_i), \ k=1(1)n, \tag{17}$$

on the basis of $\underline{X}_k$. Let $d_k=d_k(\underline{X}_k)$ be a kth predictive estimator (decision rule) for $\underline{Y}_k$ and let the loss function be

$$W[\underline{Y}^n,d^n] = \sum_{k=1}^{n} c_k(\underline{Y}_k-d_k)^2, \tag{18}$$

where

$$\underline{Y}^n = (\underline{Y}_1, \dots ,\underline{Y}_n), \tag{19}$$

$$d^n = (d_1, \dots ,d_n), \tag{20}$$

and $c_k \geq 0$ for k=1, ... ,n, $c_k > 0$ for at least one k.

Let

$$d_k = \underline{M}_k(a_k \frac{\underline{X}_k}{\underline{m}_k} + b_k), \ k=1(1)n, \tag{21}$$

where

$$\underline{m}_k = \sum_{i=1}^{k} m_i \tag{22}$$

and

$$\underline{M}_k = \sum_{i=k+1}^{n+1} m_i . \tag{23}$$

Then the risk function takes the form

$$R_p(d^n) = E_p(W[\underline{Y}^n, d^n])$$

$$= \sum_{k=1}^{n} c_k E_p([\underline{Y}_k - \underline{M}_k(a_k \frac{X_k}{\underline{m}_k} + b_k)]^2) = \check{a}p^2 + \check{b}p + \check{c}, \tag{24}$$

where

$$\check{a} = \sum_{k=1}^{n} c_k \underline{M}_k \left[\underline{M}_k \frac{\underline{m}_k - 1}{\underline{m}_k} a_k^2 - 2\underline{M}_k a_k + \underline{M}_k - 1 \right], \tag{25}$$

$$\check{b} = \sum_{k=1}^{n} c_k \underline{M}_k \left[\frac{\underline{M}_k}{\underline{m}_k} a_k^2 + 2\underline{M}_k b_k (a_k - 1) + 1 \right], \tag{26}$$

and

$$\check{c} = \sum_{k=1}^{n} c_k \underline{M}_k^2 b_k^2 . \tag{27}$$

(24) is constantly equal to $\check{c}$ (i.e., (24) is independent on p) whenever

$$\check{a} = \check{b} = 0. \tag{28}$$

It can be shown that (21) is the Bayes solution corresponding to the prior distribution of p,

$$g(p;a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} p^{a-1}(1-p)^{b-1}, \quad 0 \leq p \leq 1 \quad (a,b > 0), \tag{29}$$

with

$$a_k = \frac{\underline{m}_k}{a+b+\underline{m}_k}, \quad b_k = \frac{a}{a+b+\underline{m}_k}, \quad k=1(1)n. \tag{30}$$

Suppose that (a^*,b^*) is a solution of equation (28). It follows from Corollary 2.1.1 that (21) with

$$a_k = \frac{\underline{m}_k}{a^*+b^*+\underline{m}_k}, \quad b_k = \frac{a^*}{a^*+b^*+\underline{m}_k}, \quad k=1(1)n, \tag{31}$$

is the uniformly undominated decision rule for $\underline{Y}^n$.

For example, in the case n=1, the uniformly undominated predictive estimator of $Y=X_2(m_2)$ based on $X=X_1(m_1)$ with respect to the loss function

$$W[Y,d_1] = c_1[Y - d_1]^2 \tag{32}$$

is

$$d_1 = m_2\left(a_1 \frac{X}{m_1} + b_1\right), \tag{33}$$

where

$$a_1 = \frac{m_1}{m_1-1}\left[1 - (1/m_1 + 1/m_2 - 1/(m_1m_2))^{1/2}\right] \tag{34}$$

and

$$b_1 = (1 - a_1)/2. \tag{35}$$

Note that (33) is the Bayes solution corresponding to the prior distribution of p (29) with

$$a^* = b^* = (m_1/2)((1-a_1)/a_1) \tag{36}$$

and hence uniformly undominated.

It is interesting to compare the risk of the above uniformly undominated predictive estimator (33) with that of the standard unbiased estimator

$$d_0 = m_2 \frac{X}{m_1}. \tag{37}$$

We have

$$R_p(d_1) = E_p(W[Y,d_1]) = E_p(c_1[Y - d_1]^2)$$
$$= c_1 \frac{m_2^2}{4}\left[1 - \frac{m_1}{m_1-1}(1 - (1/m_1 + 1/m_2 - 1/(m_1m_2))^{1/2})\right]^2 \tag{38}$$

and

$$R_p(d_0) = E_p(W[Y,d_0]) = E_p(c_1[Y - d_0]^2)$$
$$= c_1 \frac{m_2}{m_1}(m_1 + m_2)p(1 - p). \tag{39}$$

As is easily seen,

$$R_p(d_0) \leqslant R_p(d_1) \tag{40}$$

if and only if

$$\left|p - \frac{1}{2}\right| \geqslant \left[1 - \frac{m_1m_2}{m_1+m_2}\left(1 - \frac{m_1}{m_1-1}\left(1 - \left(\frac{1}{m_1} + \frac{1}{m_2} - \frac{1}{m_1m_2}\right)^{1/2}\right)\right)^2\right]^{1/2}/2. \tag{41}$$

If, say, m_1=

$$\left[1 - \frac{m_1m_2}{m_1+m\ldots}\right.$$

Thus the sta...
uniformly und...

$$\left|p - \frac{1}{2}\right| \geqslant$$

REFERENCES

Nechval, N.A
the minim
from the
8th IFAC
timation,
Thorpe, J. (
craft due
Europe. M

If, say, m_1=70,000 and m_2=156,463 of aircraft movements, then

$$\left[1 - \frac{m_1 m_2}{m_1+m_2}\left(1 - \frac{m_1}{m_1-1}\left(1 - \left(\frac{1}{m_1} + \frac{1}{m_2} - \frac{1}{m_1 m_2}\right)^{1/2}\right)\right)^2\right]^{1/2} /2 = 0.0396188. \tag{42}$$

Thus the standard unbiased estimator d_0 (37) is better than the uniformly undominated predictive estimator d_1 (33) if and only if

$$\left| p - \frac{1}{2} \right| \geqslant 0.0396188. \tag{43}$$

REFERENCES

Nechval, N.A. (1988). A new efficient approach to constructing the minimum risk estimators of state of stochastic systems from the statistical data of small samples. Preprint of the 8th IFAC Symposium on Identification and System Parameter Estimation, 27-31 August 1988, Beijing, P.R. CHINA, pp. 1-6.

Thorpe, J. (1982). Accidents and serious incidents to civil aircraft due to bird strikes. 16th Meeting Bird Strike Committee Europe. Moscow, Working Paper 16.