АВИАЦИОННАЯ И РАКЕТНО-КОСМИЧЕСКАЯ ТЕХНИКА

УДК 519.254
DOI: 10.26467/2079-0619-2018-21-2-132-142

ПРИМЕНЕНИЕ БАЙЕСОВСКОГО ПОДХОДА ПОСТРОЕНИЯ ЛОГИСТИЧЕСКОЙ РЕГРЕССИИ ПРИ ОБРАБОТКЕ РЕЗУЛЬТАТОВ ИСПЫТАНИЙ НА СТОЙКОСТЬ ЭЛЕМЕНТОВ АВИАЦИОННЫХ КОНСТРУКЦИЙ

Н.Ю. КОМРАКОВ[1], С.М. МУЖИЧЕК[2], А.А. СКРЫННИКОВ[2]
[1] *ЦНИИ ВВКО Минобороны России, г. Тверь, Россия*
[2] *Государственный научно-исследовательский институт авиационных систем, г. Москва, Россия*

Работа выполнена при поддержке РФФИ, проекты 16-08-00464а, 18-08-00060а

При решении задач прогнозирования стойкости элементов авиационных конструкций при столкновении с птицами, крупным градом, посторонними предметами необходимо знать зависимость вероятности пробития преграды от характеристик ударника и скорости удара. Результаты опытов имеют вид дихотомической переменной, имеющей лишь два возможных значения (преграда пробита или нет), поэтому искомая зависимость может быть представлена в виде логистической регрессии, задаваемой интегральной функцией логистического распределения. Для оценки параметров логистической регрессии при малом количестве опытов целесообразно использовать байесовский подход. Одной из составляющих байесовской оценки параметров является расчёт функции правдоподобия. Полученные выражения функции правдоподобия и её логарифма использовались для оценки параметров логистической регрессии методом максимального правдоподобия, который реализован с использованием метода Ньютона. Приведена расчётная итерационная схема метода максимального правдоподобия. По результатам опытных данных получены оценки, которые используются для сравнения с результатами, полученными с применением байесовского подхода. Рассмотрен частный случай – байесовская оценка параметра сдвига логистической кривой при заданном значении параметра масштаба, а также общий случай – оценка параметров сдвига и масштаба. В качестве априорного распределения для параметра сдвига использовано нормальное распределение, а для параметра масштаба – классическое (двухпараметрическое) гамма-распределение. Точечная оценка параметров логистической регрессии осуществлялась по модальному значению апостериорного распределения. Использование модального, а не среднего значения, позволило значительно сократить объём расчётов за счёт того, что в этом случае нет необходимости вычислять интегральную вероятность в знаменателе формулы Байеса.

Ключевые слова: бинарная регрессия, логистическая регрессия, функция правдоподобия, точечная оценка параметров, байесовский подход, стойкость элементов авиационных конструкций.

ВВЕДЕНИЕ

При проведении испытаний на стойкость элементов авиационных конструкций к внешним неблагоприятным воздействиям часто в качестве результатов опытов фиксируются альтернативные исходы – бинарные переменные: 1 – «успех», 0 – «неудача». Например, при действии ударника по дюралевой преграде фиксируется факт пробития; при действии ударника по топливному баку фиксируется факт возгорания и т. д. Задачей таких испытаний является определение вероятности «успешного» исхода в зависимости от условий опыта.

Например, при оценке стойкости элемента конструкции к действию ударника необходимо определить зависимость вероятности пробития преграды от скорости ударника при заданных значениях всех остальных параметров, характеризующих ударник и преграду. В этом случае независимая переменная x – скорость ударника; зависимая переменная y принимает значение 1, если ударник пробил преграду, и 0 в противном случае. По результатам опытов путём статистической обработки полученных данных необходимо определить значение скорости, при которой с вероятностью 0,5 ударник пробивает преграду, значения скоростей, при которых пробитие преграды становится практически невозможным и практически достоверным событиями, а также в целом – найти зависимость, по которой для произвольного значения скорости ударника можно найти вероятность пробития преграды. Эта зависимость должна изменяться в пределах от 0 до 1 и быть неубывающей функцией, поэтому в качестве такой зависимости может быть использована функция распределения $F(x)$ некоторой непрерывной случайной величины X, так как функция распределения обладает всеми этими свойствами. На практике чаще всего используют логит- и пробит-функции (соответственно функции распределения логистического и нормального распределения), а также кусочно-линейную зависимость (функцию распределения равномерно распределённой случайной величины) [7]. Преимущество логит-функции (логистической регрессии) заключается в том, что она вычисляется проще по сравнению с пробит-функцией, а по сравнению с кусочно-линейной – лучше отражает физику процесса [9, 10].

ОЦЕНКА ПАРАМЕТРОВ ЛОГИСТИЧЕСКОЙ РЕГРЕССИИ МЕТОДОМ МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

Рассмотрим в качестве функции $F(x_i) = P\{y_i = 1\}$ логистическую функцию. Логит-преобразование представляет собой логарифм отношения вероятности того, что случайное событие произойдёт, к вероятности того, что это событие не произойдёт (логарифм отношения шансов: $\ln(P\{y_i = 1\} \,/\, P\{y_i = 0\})$). Для линейного логит-преобразования

$$\ln\frac{F(x_i)}{1 - F(x_i)} = \beta_0 + \beta_1 x_i ,$$

откуда

$$F(x_i) = \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}}; \quad 1 - F(x_i) = \frac{1}{1 + e^{\beta_0 + \beta_1 x_i}}. \tag{1}$$

Выражение $F(x_i)$ можно записать в виде

$$F(x_i) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_i)}}. \tag{2}$$

Использование логистического распределения (2) позволяет выразить его параметры – математическое ожидание и среднее квадратическое отклонение – через коэффициенты β_0 и β_1. Сопоставляя выражение (2) и стандартную форму записи логистического распределения [1]

$$F(x_i) = \frac{1}{1 + e^{-\frac{\pi(x-\xi)}{\sigma\sqrt{3}}}}, \tag{3}$$

где ξ и σ – математическое ожидание и среднее квадратическое отклонение соответственно, получим

$$\sigma = \frac{\pi}{\sqrt{3}\ \beta_1}. \tag{4}$$

Так как плотность логистического распределения симметрична относительно математического ожидания, то, решая (1) при $F(x) = 0,5,$ найдём

$$\xi = -\frac{\beta_0}{\beta_1}. \tag{5}$$

И наоборот:

$$\beta_0 = \frac{-\pi\xi}{\sqrt{3}\ \sigma}; \tag{6}$$

$$\beta_1 = \frac{\pi}{\sqrt{3}\ \sigma}. \tag{7}$$

Стандартная форма записи логистической кривой удобна тем, что значение ξ соответствует скорости, при которой в 50 % случаев наблюдается положительный исход и в 50 % случаев – отрицательный исход; при скоростях $\xi - 3\sigma$ и менее практически всегда наблюдается отрицательный исход, а при скоростях $\xi + 3\sigma$ и более практически всегда наблюдается положительный исход.

Оценка неизвестных параметров β_0 и β_1 по данным выборки может быть проведена с использованием метода максимального правдоподобия, предполагая, что наблюдения независимы [2].

С учётом (1) функция правдоподобия будет иметь вид

$$\begin{aligned} L(x_1,...,x_n,y_1,...,y_n;\beta_0,\beta_1) &= \prod_{i=1}^{n} F(x_i)^{y_i}[1-F(x_i)]^{1-y_i} = \\ &= \prod_{i=1}^{n}\left[\frac{e^{\beta_0+\beta_1 x_i}}{1+e^{\beta_0+\beta_1 x_i}}\right]^{y_i}\left[\frac{1}{1+e^{\beta_0+\beta_1 x_i}}\right]^{1-y_i}, \end{aligned} \tag{8}$$

а логарифм функции правдоподобия –

$$\ln L(x_1,...,x_n,y_1,...,y_n;\beta_0,\beta_1) = \sum_{i=1}^{n} y_i(\beta_0 + \beta_1 x_i) - \ln\left[1+e^{\beta_0+\beta_1 x_i}\right].$$

Продифференцируем логарифм функции правдоподобия по переменным β_0 и β_1:

$$\frac{\partial \ln L}{\partial \beta_0} = \sum_{i=1}^{n} \frac{y_i(1+e^{\beta_0+\beta_1 x_i}) - e^{\beta_0+\beta_1 x_i}}{1+e^{\beta_0+\beta_1 x_i}};$$

$$\frac{\partial \ln L}{\partial \beta_1} = \sum_{i=1}^{n} \frac{x_i(y_i(1+e^{\beta_0+\beta_1 x_i}) - e^{\beta_0+\beta_1 x_i})}{1+e^{\beta_0+\beta_1 x_i}}$$

и приравняем производные нулю:

$$\begin{cases} \sum_{i=1}^{n} \dfrac{y_i(1+e^{\beta_0+\beta_1 x_i}) - e^{\beta_0+\beta_1 x_i}}{1+e^{\beta_0+\beta_1 x_i}} = 0; \\ \sum_{i=1}^{n} \dfrac{x_i(y_i(1+e^{\beta_0+\beta_1 x_i}) - e^{\beta_0+\beta_1 x_i})}{1+e^{\beta_0+\beta_1 x_i}} = 0. \end{cases} \tag{9}$$

Полученная система нелинейных уравнений может быть решена относительно переменных β_0 и β_1 методом Ньютона. Обозначим

$$f_1(\beta_0,\beta_1) = \frac{\partial \ln L}{\partial \beta_0}; \qquad f_2(\beta_0,\beta_1) = \frac{\partial \ln L}{\partial \beta_1}.$$

В соответствии с методом Ньютона [3]

$$\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}^{(k+1)} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}^{(k)} - J^{-1} \begin{pmatrix} f_1(\beta_0,\, \beta_1) \\ f_2(\beta_0,\, \beta_1) \end{pmatrix}^{(k)}, \quad k = 0, 1, \dots,$$

где J – матрица Якоби:

$$J = \begin{pmatrix} \dfrac{\partial f_1}{\partial \beta_0} & \dfrac{\partial f_1}{\partial \beta_1} \\ \dfrac{\partial f_2}{\partial \beta_0} & \dfrac{\partial f_2}{\partial \beta_1} \end{pmatrix};$$

$$J^{-1} = \frac{1}{|J|} \begin{pmatrix} \dfrac{\partial f_1}{\partial \beta_0} & -\dfrac{\partial f_1}{\partial \beta_1} \\ -\dfrac{\partial f_2}{\partial \beta_0} & \dfrac{\partial f_2}{\partial \beta_1} \end{pmatrix}; \qquad |J| = \frac{\partial f_1}{\partial \beta_0}\frac{\partial f_2}{\partial \beta_1} - \frac{\partial f_1}{\partial \beta_1}\frac{\partial f_2}{\partial \beta_0}.$$

Элементы матрицы Якоби рассчитываются следующим образом:

$$J = \begin{pmatrix} \dfrac{\partial^2 \ln L}{\partial \beta_0^2} & \dfrac{\partial^2 \ln L}{\partial \beta_0 \partial \beta_1} \\ \dfrac{\partial^2 \ln L}{\partial \beta_0 \partial \beta_1} & \dfrac{\partial^2 \ln L}{\partial \beta_0^2} \end{pmatrix} = \begin{pmatrix} -\sum_{i=1}^{n} \dfrac{e^{\beta_0+\beta_1 x_i}}{(1+e^{\beta_0+\beta_1 x_i})^2} & -\sum_{i=1}^{n} \dfrac{x_i e^{\beta_0+\beta_1 x_i}}{(1+e^{\beta_0+\beta_1 x_i})^2} \\ -\sum_{i=1}^{n} \dfrac{x_i e^{\beta_0+\beta_1 x_i}}{(1+e^{\beta_0+\beta_1 x_i})^2} & -\sum_{i=1}^{n} \dfrac{x_i^2 e^{\beta_0+\beta_1 x_i}}{(1+e^{\beta_0+\beta_1 x_i})^2} \end{pmatrix}.$$

Решение в поэлементной форме имеет вид:

$$\beta_0^{(k+1)} = \beta_0^{(k)} - \left[\frac{f_1}{|J|} \frac{\partial f_2}{\partial \beta_1} - \frac{f_2}{|J|} \frac{\partial f_1}{\partial \beta_1} \right]^{(k)};$$

$$\beta_1^{(k+1)} = \beta_1^{(k)} - \left[\frac{f_2}{|J|} \frac{\partial f_1}{\partial \beta_0} - \frac{f_1}{|J|} \frac{\partial f_2}{\partial \beta_0} \right]^{(k)}, \quad k = 0, 1, \ldots$$

В качестве примера рассмотрим результаты десяти опытов по пробитию преграды ударником (см. табл. 1).

Таблица 1
Table 1

i	x_i	y_i
1	988,9	1
2	1026,9	1
3	1024,5	1
4	1034,0	1
5	961,5	1
6	1022,9	1
7	1024,8	1
8	967,7	0
9	1067,2	1
10	1006,6	0

Начальное приближение можно выбрать достаточно грубо: ξ = 1000; σ = 100; по формулам (6), (7) найдём значения параметров: $\beta_0^{(0)}$ = –18; $\beta_1^{(0)}$ = 0,018. Решая систему (9) методом Ньютона, уже через 4 шага находим значения оценок неизвестных параметров: β_0^* = –37,0324; β_1^* = 0,0383. Теперь по формулам (4), (5) можно найти значения оценок параметров стандартной формы записи логистической кривой: ξ^* = 966,707; σ^* = 47,348. График полученной по результатам опытов зависимости (1) приведён на рис. 1.

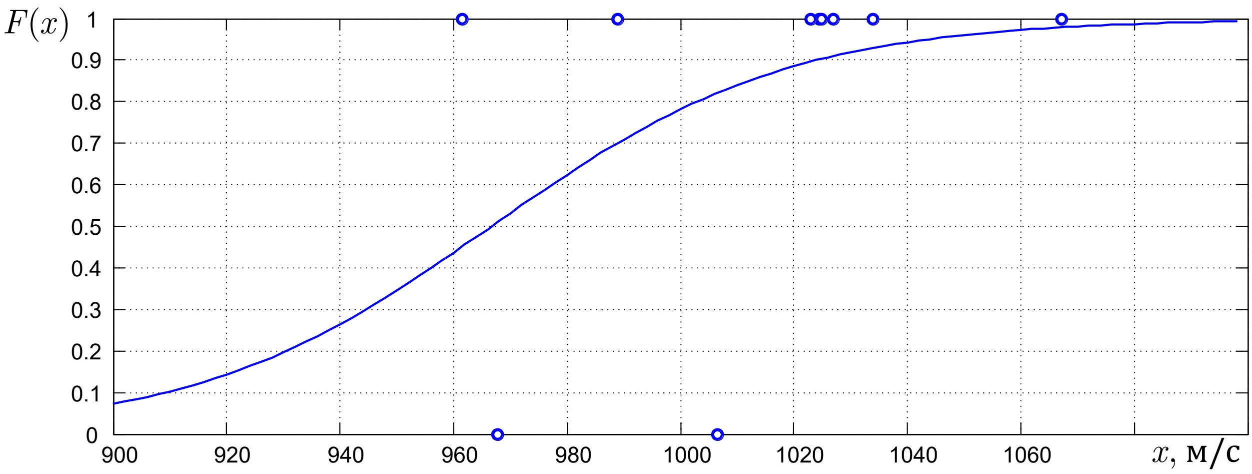

Рис. 1
Fig. 1

ОЦЕНКА ПАРАМЕТРОВ ЛОГИСТИЧЕСКОЙ РЕГРЕССИИ С ИСПОЛЬЗОВАНИЕМ БАЙЕСОВСКОГО ПОДХОДА

В условиях относительно малых выборок заметным преимуществом по сравнению с классическими методами (метод максимального правдоподобия, метод моментов и др.) обладает байесовский подход [4, 5]. Байесовская оценка базируется на априорной информации об оцениваемых параметрах и на данных экспериментов. Априорная информация задаётся в виде распределения вероятностей неизвестного параметра как степень уверенности статистика в том, что этот параметр примет то или иное значение, ещё до начала сбора данных. Новая информация, поступающая с данными результатов экспериментов, позволяет пересмотреть это распределение, переходя от априорного к апостериорному распределению [6].

При байесовском подходе к оценке параметров логистической регрессии принимается, что параметры β_0 и β_1 являются случайными величинами и для них задаётся априорное распределение – распределение, полученное на основе данных предыстории или на основании суждений [5].

Рассмотрим сначала частный случай, когда параметр σ логистического распределения известен; тогда по формуле (7) можно найти значение параметра β_1, т. е. в рассматриваемом случае β_1 – величина известная.

Множество возможных значений параметра ξ – множество действительных чисел. Примем априорное распределение параметра ξ нормальным с числовыми характеристиками m_ξ и σ_ξ. Тогда из (5) получим

$$\beta_0 = -\beta_1 \xi, \tag{10}$$

следовательно,

$$M[\beta_0] = -\beta_1 m_\xi; \qquad \sigma[\beta_0] = \beta_1 \sigma_\xi.$$

Так как априорное распределение случайной величины ξ – нормальное, то и β_0 как линейная функция случайного аргумента также будет иметь нормальное распределение:

$$f(\beta_0) = \frac{1}{\sqrt{2\pi}\ \beta_1 \sigma_\xi} \exp\left\{-\ \frac{(\beta_0 + \beta_1 m_\xi)^2}{2\beta_1^2 \sigma_\xi^2}\right\}. \tag{11}$$

Обозначим: $p(y_1,\ldots,y_n \mid x_1,\ldots,x_n;\beta_0)$ – вероятность наблюдения значений $y_1,\ldots,y_n$ при конкретном значении β_0 и при условии, что опыты проводились в точках $x_1,\ldots,x_n$; это не что иное, как функция правдоподобия (8). Тогда теорему Байеса для рассматриваемого случая можно сформулировать следующим образом: если B_0 – множество возможных значений параметра β_0 с априорной плотностью распределения (11) на множестве B_0 и если $p(y_1,\ldots,y_n \mid x_1,\ldots,x_n;\beta_0)$ обозначает правдоподобие, то апостериорная плотность вероятностей $f(\beta_0 \mid x_1,\ldots,x_n,y_1,\ldots,y_n)$ задаётся выражением

$$f(\beta_0 \mid x_1,\ldots,x_n,y_1,\ldots,y_n) = \frac{p(y_1,\ldots,y_n \mid x_1,\ldots,x_n;\beta_0) f(\beta_0)}{p(y_1,\ldots,y_n \mid x_1,\ldots,x_n)}, \tag{12}$$

где $p(y_1,...,y_n \mid x_1,...,x_n) = \int\limits_{B_0} p(y_1,...,y_n \mid x_1,...,x_n;\beta_0) f(\beta_0) d\beta_0$.

Подставив (8), (11) в (12), получим

$$f(\beta_0 \mid x_1,...,x_n,\ y_1,...,y_n) = \\ = \frac{\prod\limits_{i=1}^{n}\left[\frac{e^{\beta_0+\beta_1 x_i}}{1+e^{\beta_0+\beta_1 x_i}}\right]^{y_i}\left[\frac{1}{1+e^{\beta_0+\beta_1 x_i}}\right]^{1-y_i}\frac{1}{\sqrt{2\pi}\ \beta_1\sigma_\xi}\exp\left\{-\ \frac{(\beta_0+\beta_1 m_\xi)^2}{2\beta_1^2\sigma_\xi^2}\right\}}{\int\limits_{-\infty}^{\infty}\prod\limits_{i=1}^{n}\left[\frac{e^{\beta_0+\beta_1 x_i}}{1+e^{\beta_0+\beta_1 x_i}}\right]^{y_i}\left[\frac{1}{1+e^{\beta_0+\beta_1 x_i}}\right]^{1-y_i}\frac{1}{\sqrt{2\pi}\ \beta_1\sigma_\xi}\exp\left\{-\ \frac{(\beta_0+\beta_1 m_\xi)^2}{2\beta_1^2\sigma_\xi^2}\right\}d\beta_0}. \tag{13}$$

В качестве точечной оценки β_0^* неизвестного параметра β_0 примем модальное значение случайной величины β_0; для этого необходимо найти такое значение аргумента β_0, при котором функция $f(\beta_0 \mid x_1,...,x_n,\ y_1,...,y_n)$ (13) будет иметь максимальное значение. Тогда достаточно построить зависимость числителя формулы (13) от β_0, а знаменатель можно не рассчитывать, так как он играет роль нормирующего множителя.

Оценка β_0^* неизвестного параметра β_0 по данным табл. 1 при значениях параметров априорного распределения m_ξ = 1000; σ_ξ = 100 даёт значение β_0^* = –35,448, что соответствует значению оценки ξ^* = 977,18.

Перейдём теперь к общему случаю – байесовской оценке двух неизвестных параметров β_0 и β_1. Так как для получения точечных оценок нет необходимости вычислять знаменатель формулы Байеса, задача значительно упрощается.

В этом случае нужно задавать совместную плотность распределения $f(\beta_0,\beta_1)$. Если параметры ξ и σ логистического распределения (3) независимы, то параметры β_0 и β_1 – уже зависимые (см. (10)), и, задаваясь априорным распределением параметров ξ и σ, нужно решать самостоятельную задачу определения совместного закона распределения $f(\beta_0,\beta_1)$. Поэтому задачу оценки параметров логистического распределения целесообразно решать в независимых переменных ξ и σ.

Функция правдоподобия $p(y_1,...,y_n \mid x_1,...,x_n;\xi,\sigma)$ будет иметь вид

$$p(y_1,...,y_n \mid x_1,...,x_n;\xi,\sigma) = \prod_{i=1}^{n}\left[\frac{1}{1+e^{-\frac{\pi(x_i-\xi)}{\sigma\sqrt{3}}}}\right]^{y_i}\left[\frac{e^{-\frac{\pi(x_i-\xi)}{\sigma\sqrt{3}}}}{1+e^{-\frac{\pi(x_i-\xi)}{\sigma\sqrt{3}}}}\right]^{1-y_i}, \tag{14}$$

а совместная плотность $f(\xi,\sigma) = f(\xi)f(\sigma)$ в силу их независимости.

Примем распределение величины ξ нормальным с параметрами m_ξ, σ_ξ. Распределение величины σ должно быть одним из распределений с возможными значениями на положительной полуоси; примем классическое (двухпараметрическое) гамма-распределение [7, 8] с плотностью

$$f(\sigma) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)} \sigma^{\alpha-1} \exp\{-\lambda\sigma\},$$

где λ – параметр масштаба ($\lambda > 0$);
α – параметр формы ($\alpha > 0$).

Математическое ожидание и дисперсия случайной величины σ, имеющей гамма-распределение, вычисляются по формулам $m_{\sigma} = \alpha / \lambda$; $D_{\sigma} = \alpha / \lambda^2$.

Тогда

$$f(\xi,\sigma) = \frac{\lambda^{\alpha}\sigma^{\alpha-1}}{\sqrt{2\pi}\sigma_{\xi}\Gamma(\alpha)} \exp\left\{-\frac{(\xi - m_{\xi})^2}{2\sigma_{\xi}^2} - \lambda\sigma\right\}. \tag{15}$$

В этом случае числитель формулы Байеса будет равен произведению функций $p(y_1,...,y_n \mid x_1,...,x_n;\xi,\sigma)$ и $f(\xi,\sigma)$, вычисляемых соответственно по формулам (14) и (15). Необходимо найти такие значения ξ^* и σ^*, при которых функция $p(y_1,...,y_n \mid x_1,...,x_n;\xi,\sigma)f(\xi,\sigma)$ принимает максимальное значение.

Ввиду того, что численные значения рассматриваемой функции на всей области практически возможных значений параметров ξ и σ могут быть очень малыми, то целесообразно строить график логарифма функции.

Оценки ξ^* и σ^* неизвестных параметров ξ и σ по данным табл. 1 при значениях параметров априорного распределения m_{ξ} = 1000; σ_{ξ} = 100, α = 2,50; λ = 0,05 даёт значения ξ^* = 980,6 и σ^* = 31,44. Линии равного уровня в окрестностях точки максимума приведены на рис. 2.

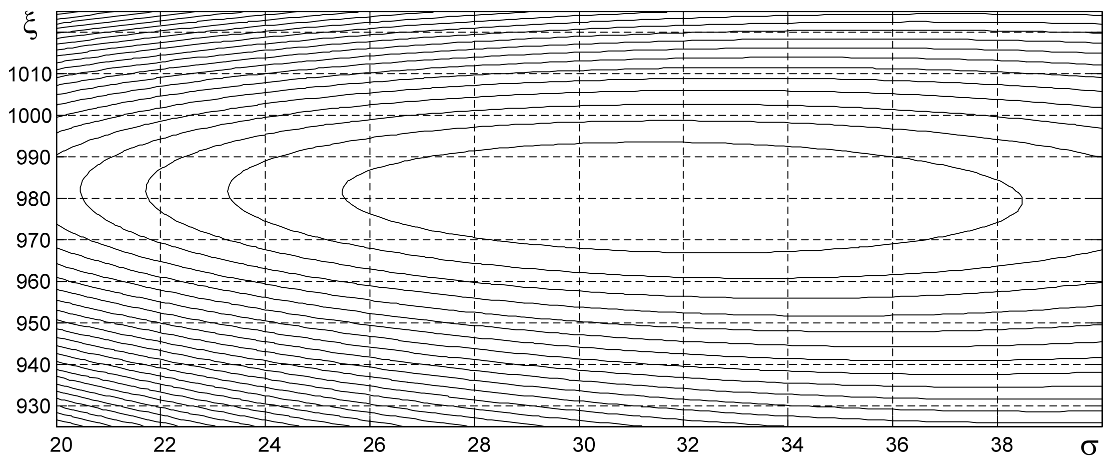

Рис. 2
Fig. 2

Приведённые на рис. 2 линии равного уровня подтверждают независимость переменных ξ и σ и в апостериорной функции распределения.

ЗАКЛЮЧЕНИЕ

Использование байесовского подхода даёт возможность учёта знаний, полученных в подобных испытаниях на стойкость элементов авиационных конструкций. Степень нашей уверенности в значениях скорости, при которой в 50 % случаев происходит пробитие преграды, скорости, ниже которой пробитие никогда не происходит, и скорости, выше которой пробитие достоверно, численно выражается в виде априорной вероятности.

Рассчитанные апостериорные функции распределения позволяют провести не только точечную, но и интервальную оценку неизвестных параметров логистической регрессии, дать рекомендации об условиях проведения опытов.

СПИСОК ЛИТЕРАТУРЫ

1. Джонсон Н.Л., Коц С., Балакришнан Н. Одномерные непрерывные распределения: в 2-х ч. Ч. 2. М.: БИНОМ. Лаборатория знаний, 2010. 600 с.

2. Буре В.М., Парилина Е.М. Теория вероятностей и математическая статистика. СПб.: Лань, 2013. 416 с.

3. Квасов Б.И. Численные методы анализа и линейной алгебры. Использование Matlab и Simulink. СПб.: Лань, 2016. 328 с.

4. Айвазян С.А. Байесовский подход в эконометрическом анализе // Прикладная эконометрика. 2008. №1 (9). С. 93–130.

5. Справочник по прикладной статистике. В 2-х т. Т. 2 / под ред. Э. Ллойда, У. Лидермана. М.: Финансы и статистика, 1990. 526 с.

6. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики. М.: ЮНИТИ, 1998. 1005 с.

7. Вадзинский Р.Н. Справочник по вероятностным распределениям. СПб.: Наука, 2001. 295 с.

8. Хастингс Н., Пикок Дж. Справочник по статистическим распределениям. М.: Статистика. 95 с.

9. О методике автоматизированного проведения наземных испытаний малокалиберных боеприпасов / С.М. Мужичек, А.А. Скрынников, С.А. Абрамов, В.В. Ефанов // Вопросы оборонной техники. Серия 16. 2017. Вып. 111–112. С. 48–53.

10. Мужичек С.М., Скрынников А.А., Абрамов С.А. Автоматизированная технология наземных испытаний боеприпасов // Юбилейная всероссийская научно-техническая конференция «Авиационные системы в XXI веке». Москва, 26–27 мая 2017 г.: сборник докладов. М.: ГосНИИАС, 2017. С. 100–109.

СВЕДЕНИЯ ОБ АВТОРАХ

Комраков Николай Юрьевич, кандидат технических наук, старший научный сотрудник, старший научный сотрудник ЦНИИ ВВКО Минобороны России, komr.valentina@yandex.ru.

Мужичек Сергей Михайлович, доктор технических наук, профессор, учёный секретарь ФГУП ГосНИИАС, msm19@yandex.ru.

Скрынников Андрей Александрович, кандидат технических наук, старший научный сотрудник, начальник сектора ФГУП ГосНИИАС, a1260@mail.ru.

APPLICATION OF BAYESIAN APPROACH OF LOGISTIC REGRESSION FORMATION WHILE PROCESSING THE TESTS RESULTS ON THE DURABILITY OF AIRCRAFTCONSTRUCTIONS ELEMENTS

Nikolay Yu. Komrakov[1], **Sergey M. Muzhichek**[2], **Andrey A. Skrynnikov**[2]
[1] *The Central Scientific Research Institute of Air Defense Forces, Tver, Russia*
[2] *State Research Institute Of Aviation Systems, Moscow, Russia*

The study was conducted with support of Russian Foundation for Basic Research (RFBR) projects 16-08-00464a, 18-08-00060a

ABSTRACT

When solving the problems of predicting the durability of aircraft structures elements in case of a bird strike, in a collision with large hail, foreign objects, it is necessary to know the dependence of the probability of the barrier penetration on the characteristics of the impactor and the impact speed. The results of the experiments have the form of a dichotomous variable having only two possible values (the barrier is punched or not), so the required dependence can be represented in the form of a logistic regression given by the integral function of the logistic distribution. To evaluate the parameters of logistic regression with a small number of experiments, it is advisable to use the Bayesian approach. One of the components of the Bayesian parameter estimation is the calculation of the likelihood function. The resulting expressions for the likelihood function and its logarithm were used to estimate the logistic regression parameters using the maximum likelihood method, which was realized by Newton method. The calculated iterative scheme of the maximum likelihood method is given. Based on the results of experimental data, estimates are obtained that are used to compare with the results obtained by the Bayesian approach. A special case is considered – the Bayesian estimation of the shift parameter of the logistic curve for a given value of the scale parameter, and also a general case – the estimation of the shift and scale parameters. A normal distribution is used as a priori distribution for the shift parameter, and a classical (two-parameter) gamma distribution is used for the scale parameter. Point estimation of logistic regression parameters was carried out by the modal value of a posteriori distribution. Using the modal rather than the mean value, made it possible to significantly reduce the amount of calculations due to the fact that in this case there is no need to calculate the integral probability in the denominator of the Bayesian formula.

Key words: binary regression, logistic regression, likelihood function, point estimation of parameters, Bayesian approach, durability of aircraft structures elements.

REFERENCES

1. Johnson N.L., Kotz S., Balakrishnan N. *Odnomerniye nepreryvniye raspredeleniya* [Continuous univariate distributions]. In 2 vol. Vol. 2, M.: BINOM, *Laboratoria znaniy* [Lab of knowledge], 2010, 600 p. (in Russian)

2. Bure V.M., Parilina E.M. *Teorija verojatnostej i matematicheskaja statistika* [Theory of probabilities and mathematical statistics]. SPb.: Lan', 2013, 416 p. (in Russian)

3. Kvasov B.I. *Chislennye metody analiza i linejnoj algebry. Ispol'zovanie Matlab i Simulink* [Numeral methods of an analysis and linear algebra. Use of Matlab and Simulink]. SPb.: Lan', 2016. 328 p. (in Russian)

4. Ajvazjan S.A. *Bajesovskij podhod v ekonometricheskom analize* [Bayesian approach in the econometrical analysis]. *Prikladnaja ekonometrika* [Applied econometrics], 2008, No 1 (9), pp. 93–130. (in Russian)

5. *Spravochnik po prikladnoy statistike* [Handbook of Applied Statistics]. In 2 vol., Volume II. Statistics, Ed. by E.Lloid, Y.Liderman. M.: Finance and Statistics. 1990, 526 p. (in Russian)

6. Ajvazjan S.A., Mhitarjan V.S. *Prikladnaja statistika i osnovy jekonometriki* [Applied statistics and fundamentals of econometrics]. M.: JuNITI, 1998, 1005 p. (in Russian)

7. Vadzinskij R.N. *Spravochnik po verojatnostnym raspredelenijam* [Reference book on probable distributions]. SPb.: Nauka, 2001, 295 p. (in Russian)

8. Hastings N.A.J., Peacock J.B. *Spravochnik po statesticheskim raspredileniyam* [Handbook on Statistical Distributions]. M.: Statistics, 95p. (in Russian)

9. Muzhichek S.M., Skrynnikov A.A., Abramov S.A., Efanov V.V. *O metodike avtomatizirovannogo provedenija nazemnyh ispytanij malokalibernyh boepripasov* [On the method of automated ground testing of small-caliber ammunition]. *Voprosy oboronnoj tehniki Seria 16* [Milittary Enginery Issue 16: Counter-terrorism technical devices], 2017, pp. 48–53. (in Russian)

10. Muzhichek S.M., Skrynnikov A.A., Abramov S.A. *Avtomatizirovannaja tehnologija nazemnyh ispytanij boepripasov* [Automated technology of ground weapons trials]. *Jubilejnaja vserossijskaja nauchno-tehnicheskaja konferencija «Aviacionnye sistemy v XXI veke»* [Anniversary all-Russian scientific and technical Conference "Aviation systems in XXI century"]. M.: May 26-27th 2017 *Sbornik dokladov* [Collection of reports]. M.: FGUP "GosNIIAS" State Research Institute of Aviation Systems, pp.100–109. (in Russian)

INFOTMATION ABOUT THE AUTORS

Nikolay Yu. Komrakov, Candidate of Technical Sciences, Senior Researcher, Senior Researcher of Central Scientific Research Institute of Air Defense Forces, komr.valentina@yandex.ru.

Sergey M. Muzhichek, Doctor of Technical Sciences, Professor, Scientific Secretary, State Research Institute of Aviation Systems, msm19@yandex.ru.

Andrey A. Skrynnikov, Candidate of Technical Sciences, Senior Researcher, Head of Sector, State Research Institute of Aviation Systems, a1260@mail.ru.

Поступила в редакцию	25.10. 2017	Received	25.10. 2017
Принята в печать	14.03.2018	Accepted for publication	14.03.2018