БІЯЛОГІЯ

УДК 599.735.51:519.216.3:636(476)

Модельные математические функции в практической биологии

М.Н. Борисевич*, П.А. Красочко*, И.М. Прищепа**

**Учреждение образования*
«Витебская ордена “Знак Почета” государственная академия ветеринарной медицины»
***Учреждение образования «Витебский государственный*
университет имени П.М. Машерова»

В настоящей статье приведено решение задачи по построению регрессионных моделей общего поголовья новорожденных телят, полученных в Республике Беларусь в разные периоды времени, начиная с 1989 г. и по 1998 г. (всего за 10 лет).

Цель исследования – создание простых математических моделей для нужд практической биологии, описывающих реальные процессы и данные государственной статистической отчетности.

***Материал и методы.** Информационным материалом являются экспериментальные данные и числовые массивы государственной ветеринарной отчетности Главного управления ветеринарии Минсельхозпрода Республики Беларусь и Республиканской ветеринарной лаборатории. Использованы математические методы регрессионного анализа.*

***Результаты и их обсуждение.** Представлены графические зависимости совместно с данными, заимствованными из материалов официальной статистической отчетности. Кривые построены в соответствии с используемыми регрессионными моделями (полиномиальной функцией первой, второй, третьей и четвертой степеней; логарифмической функцией (по основанию 10); логарифмической функцией (по основанию Е) и экспоненциальной функцией).*

Графики позволяют визуально оценить имеющиеся различия между данными отчетности и расчетными данными, полученными в вычислительных экспериментах с группой указанных регрессионных моделей.

Регрессионные кривые, рассмотренные в настоящей работе, с различной степенью точности описывают данные официальной статистики.

***Заключение.** Наилучших результатов можно достичь при использовании полиномиальной кривой пятой степени. С ее помощью аппроксимация исходных данных может быть осуществлена почти с нулевой погрешностью. В ряде случаев, когда не требуется такая высокая точность регрессионного приближения, можно рекомендовать использование и более простых линий регрессии – полиномиальной функции второй степени (с погрешностью 9%) и логарифмических кривых (с ошибками приблизительно на том же уровне). В качестве первого приближения допускается применение линейного уравнения регрессии (полиномиальной функции первой степени), обеспечивающего уровень ошибки в пределах 24%.*

***Ключевые слова:** математическая модель, уравнение регрессии, функция, показательная функция, гиперболическая функция, степенная функция.*

Model Mathematic Functions in Practical Biology

M.N. Borisevich*, P.A. Krasochko*, I.M. Prishchepa**
**Educational Establishment «Vitebsk State Order of Honor Academy of Veterinary Medicine»*
***Educational Establishment «Vitebsk State P.M. Masherov University»*

The solution of the task of creating regression models of the common livestock of the newborn calves received in the Republic of Belarus during different periods of time between 1989 and 1998 is provided in the present article.

The purpose of the research is building up simple mathematic models for practical biology, which describe real processes and data of state statistic reports.

Material and methods. *The information material for the article is experimental data and numerical masses of the state veterinary reporting of the Head Veterinary Department of the Ministry of Agriculture of the Republic of Belarus and of the Republican Veterinary Laboratory. Mathematic methods of regression analysis are used.*

Findings and their discussion. *Graphic dependences together with the data borrowed from materials of the official statistical reporting are presented. The curves are constructed according to the used regression models (polynomial function of the first, second, third and fourth degrees; a logarithmic function (on the basis of 10); a logarithmic function (on the basis E) and exponential function).*

The schedules allow estimating visually available differences between the data of the reporting and the calculation data obtained in computing experiments with the group of the specified regression models.

The regression curves considered in the work with various degree of accuracy describe data of official statistics.

Conclusion. *The best results can be achieved when using the fifth degree polynomial curve. With its help approximation of input data can be carried out with almost zero error. In some cases, when such high precision of the regression approximation is not required, it is possible to recommend using simpler lines of regression – polynomial function of the second degree (with a margin error 9%) as well as logarithmic curves (with mistakes approximately at the same level). As the first approximation the use of the simple equation of regression (polynomial function of the first degree) providing mistake level within 24% is allowed.*

Key words: *mathematical model, regression equation, function, exponential function, hyperbolic function, power function.*

В настоящей статье приведено решение задачи по построению регрессионных моделей общего поголовья новорожденных телят, полученных в Республике Беларусь в разные периоды времени, начиная с 1989 г. и по 1998 г. (всего за 10 лет) [1; 2].

В результате численных экспериментов установлены зависимости, которые представлены на рис. 1 и рис. 2. Здесь отложены: по оси абсцисс – годы, цифре 1 соответствует 1989 г., 2 – 1990 г., 3 – 1991 г. и т.д.; по оси ординат – поголовье телят, полученных в Республике Беларусь (тыс. голов). Точками показаны значения, заимствованные из материалов официальной статистической отчетности, кривые построены в соответствии с используемыми регрессионными моделями (перечень моделей представлен в табл.).

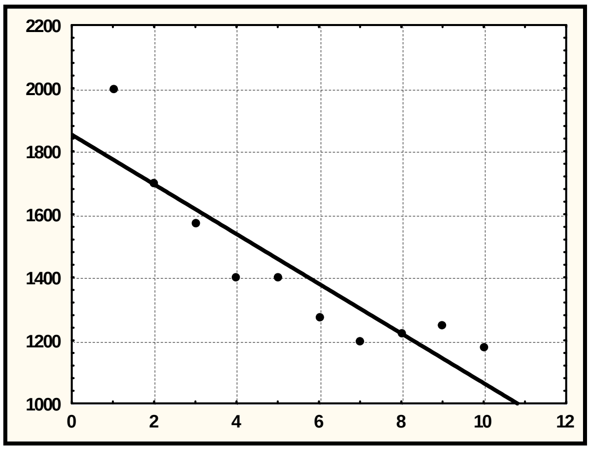
Графики позволяют визуально оценить имеющиеся различия между данными ветеринарной отчетности и расчетными данными, полученными в вычислительных экспериментах с группой регрессионных моделей: на рис. 1 это серия полиномиальных моделей, различающихся значением показателя степени (m = 1, 2, 3, 4, 5), на рис. 2 – группа из трех моделей – двух логарифмических (десятичной и натуральной) и экспоненциальной.

Цель исследования – создание простых математических моделей для нужд практической биологии, описывающих реальные процессы и данные государственной статистической отчетности.

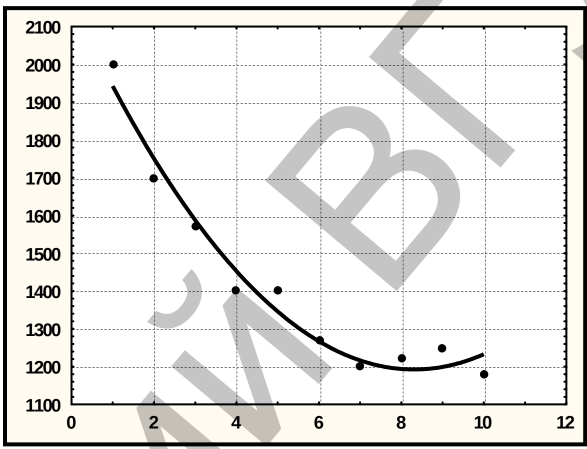
Материал и методы. Информационным материалом являются экспериментальные данные и числовые массивы государственной ветеринарной отчетности Главного управления ветеринарии Минсельхозпрода Республики Беларусь и Республиканской ветеринарной лаборатории. Использованы математические методы регрессионного анализа.

Результаты и их обсуждение. Выполним визуальный анализ рис. 1. Видно, что полиномиальная функция первой степени (рис. 1, а) неудовлетворительно описывает статистические данные. Отчетливо выражен разброс фактических точек в позициях модельной кривой, почти все они находятся за ее пределами. Есть, таким образом, основания предполагать, что значения коэффициента детерминации D, описывающего величину разброса, будут невысокими. Анализируя соответствующие данные табл., приходим к заключению: выражения $\sum_{i=1}^{n}(y_i^{+} - \bar{y})^2$ и $\sum_{i=1}^{n}(y_i - \bar{y})^2$, задающие D, действительно различаются между собой незначительно, как следствие этого, D<1 и равно 0,807. Это значит, что полиномиальная функция первой степени описывает данные официальной ветеринарной статистики с точностью $\eta = 76\%$, погрешность представления $(1-\eta) = 24\%$. Если ветеринарного специалиста такая точность устраивает, то он может воспользоваться предложенной моделью для аналитического описания поголовья новорожденных телят. В противном случае ему следует обратиться к другим регрессионным моделям, представленным в табл.

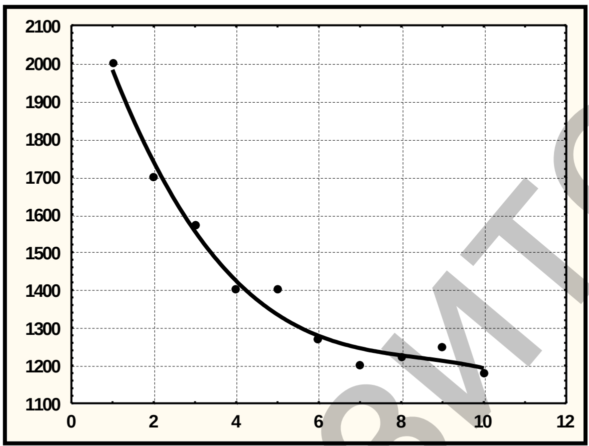
Из табл. видно также, что математическая модель регрессии для обсуждаемого случая имеет самый простой вид. Она является линейной моделью и определяется только двумя коэффициентами: $\theta_0 = 1852,67$ и $\theta_1 = 78,85$.



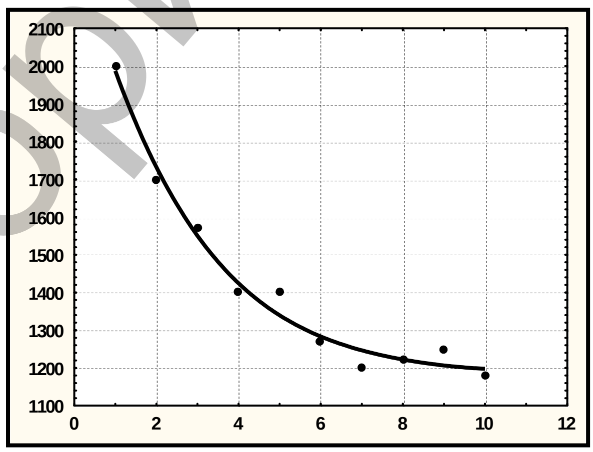
а)



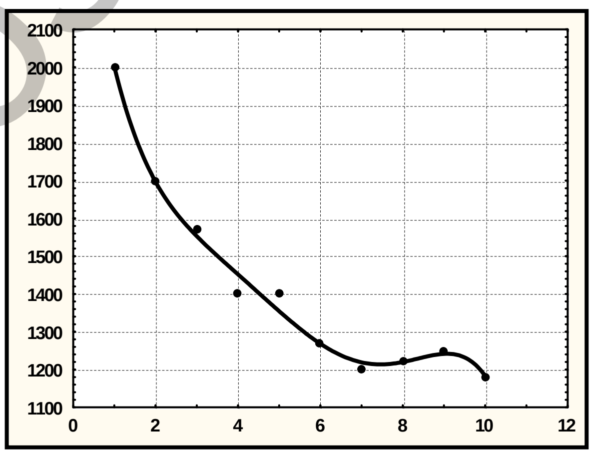
б)



в)



г)



д)

Рис. 1. **Полиномиальные функции различных степеней: первой (а), второй (б), третьей (в), четвертой (г), пятой (д)**

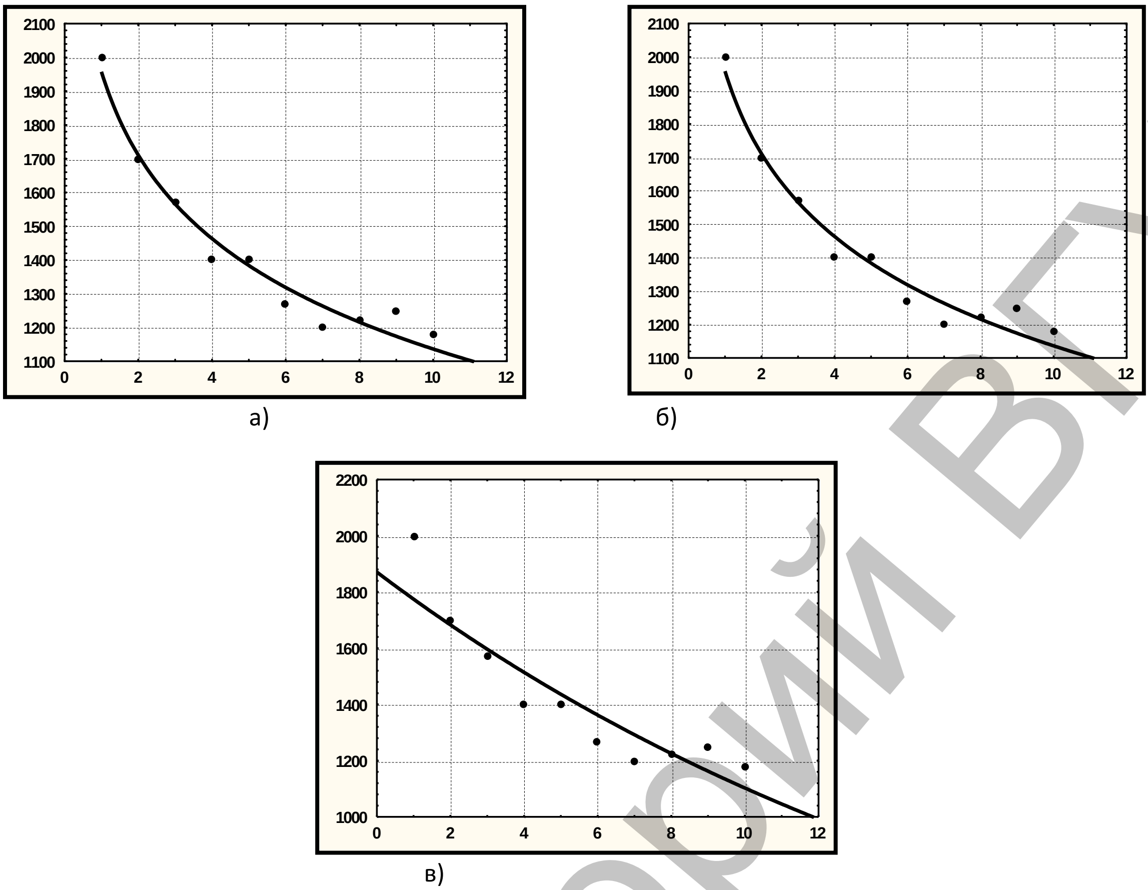


Рис. 2. **Логарифмические функции:**
по десятичному основанию (а), по натуральному основанию (б); экспоненциальная функция (в)

Чтобы воспользоваться моделью на практике, достаточно вместо переменной *x* подставить в формулу порядковый номер года. После несложных вычислений можно прийти к нужному поголовью.

Визуальный анализ рис. 1, б показывает, что полиномиальная функция второй степени описывает статистические данные лучше, чем линейная модель. Можно, например, отметить меньший разброс точек в области модельной кривой и их близкое взаимное расположение. По этой причине значение D здесь выше, чем в случае предыдущем (для линейной модели D = 0,807, для квадратичной – 0,971). Аналитическое соотношение сложнее, модель относится к классу нелинейных регрессионных моделей и описывается тремя параметрами (в предыдущем случае их было два): $\theta_0 = 2161{,}83$, $\theta_1 = -233{,}43$ и $\theta_2 = 14{,}05$.

Точность, с которой квадратичная модель воспроизводит исходные данные, составляет $\eta = 91\%$, что почти в 1,2 раза больше, чем для предыдущего случая. На практике модель можно использовать лишь с одной оговоркой – получаемые с ее помощью данные отличаются от исходных в пределах 9%. Полиномиальная функция третьей степени (рис. 1в) также неплохо приближает статистические данные. Для нее коэффициент детерминации D = 0,967 (табл.), что лишь немногим меньше предыдущего случая. Зато математическое соотношение сложнее – требуется уже четыре параметра для описания исходной статистической зависимости: $\theta_0 = 2297{,}33$, $\theta_1 = -353{,}61$, $\theta_2 = 40{,}11$ и $\theta_3 = -1{,}57$. Точность приближения фактических данных и погрешности – на уровне предыдущей модели.

Полиномиальная функция четвертой степени (рис. 1г) лучшим образом определяет статистические точки по сравнению с предыдущей моделью. Для нее коэффициент детерминации близок к единице (D = 0,996), отличаясь лишь сотыми долями 0,004. Это составляет 0,4% регрессионных погрешностей. Хотя для корректного построения модели требуется пять параметров ($\theta_0 = 2335{,}83$, $\theta_1 = -402{,}97$, $\theta_2 = 58{,}17$, $\theta_3 = -4{,}04$ и $\theta_4 = 0{,}11$), ее можно рекомендовать к практическому применению. При этом достигается почти 94% – точность совпадения с фактическими данными, уровень погрешностей – не выше 6%.

Таблица

Вид регрессии, ее математическая модель, значения вспомогательных выражений и ошибка регрессии D

№ п/п	Вид регрессии	Математическая модель регрессии	$\sum_{i=1}^{n}(y_i^+ - \bar{y})^2$	$\sum_{i=1}^{n}(y_i - \bar{y})^2$	D	η
1.	Полиномиальная функция первой степени	$y = 1852{,}67 - 78{,}85 * x$	512929,1	635490	0,807	0,76
2.	Полиномиальная функция второй степени	$y = 2161{,}83 - 233{,}43 * x + 14{,}05 * x^2$	617548,1	635490	0,971	0,91
3.	Полиномиальная функция третьей степени	$y = 2297{,}33 - 353{,}61 * x + 40{,}11 * x^2 - 1{,}57 * x^3$	614593,1	635490	0,967	0,91
4.	Полиномиальная функция четвертой степени	$y = 2335{,}83 - 402{,}97 * x + 58{,}17 * x^2 - 4{,}04 * x^3 + 0{,}11 * x^4$	632963,7	635490	0,996	0,94
5.	Полиномиальная функция пятой степени	$y = 2693{,}33 - 992{,}30 * x + 367{,}54 * x^2 - 73{,}21 * x^3 + 6{,}98 * x^4 - 0{,}25 * x^5$	676348,5	635490	1,064	1,00
6.	Логарифмическая функция (по основанию 10)	$y = 1957{,}73 - 821{,}27 * \log(x)$	615204,2	635490	0,968	0,91
7.	Логарифмическая функция (по основанию ***e***)	$y = 1957{,}73 - 356{,}67 * \ln(x)$	615193,7	635490	0,968	0,91
8.	Экспоненциальная функция	$y = 1870{,}87 * \exp(-0{,}0529 * x)$	460012,7	635490	0,723	0,68

Наилучшее приближение к исходным данным достигается с помощью полиномиальной функции пятой степени (рис. 1д). Для ее построения требуется уже шесть параметров ($\theta_0 = 2693{,}33$, $\theta_1 = -992{,}30$, $\theta_2 = 367{,}54$, $\theta_3 = -73{,}21$, $\theta_4 = 6{,}98$ и $\theta_5 = -0{,}25$). За счет этого обеспечивается почти 100% точность воспроизведения фактических данных.

Резюмируя анализ рис. 1, отметим, что в целом полиномиальная функция аппроксимирует исходные данные с удовлетворительной степенью точности. Исключение составляет линейная модель – ее представление не отличается высоким значением точности (около 76%), на заметной отметке остается и уровень погрешностей – 24%. Вместе с тем для практических вычислений можно использовать любую модель полиномиальной функции. При этом важно помнить, с какой ошибкой она описывает исходные данные. Для почти безошибочных расчетов предпочтительнее использовать функцию пятой степени.

Перейдем теперь к визуальному анализу зависимостей рис. 2. Напомним, что здесь рассматриваются три регрессионные модели: рис. 2а – логарифмическая десятичная (по основанию 10), рис. 2б – логарифмическая натуральная (по основанию *e*) и рис. 2в – экспоненциальная.

Логарифмические функции (рис. 2а, б) неплохо аппроксимируют статистические данные (значения D для обеих функций одинаковы и равны 0,968). Вид функций несложен, для построения моделей требуются всего два параметра – θ_0 и θ_1. Ошибка, с которой функции описывают поголовье телят, не выше 9%.

Экспоненциальная функция (рис. 2, в) не приводит к желаемому результату. Она крайне плохо приближает расчетные данные к данным официальной статистики (для нее характерен самый низкий коэффициент детерминации среди всех обсуждаемых моделей, равный 0,723), а следовательно, уровень ошибок аппроксимации здесь достаточно высок, чтобы не применять эту модель в практических целях (32%).

Заключение. Таким образом, регрессионные кривые, рассмотренные в настоящей работе, с различной степенью точности описывают данные официальной ветеринарной статистики. Наилучших результатов можно достичь при использовании полиномиальной кривой пятой степени:

$$y = 2693{,}33 - 992{,}30 * x + 367{,}54 * x^2 - 73{,}21 * x^3 + 6{,}98 * x^4 - 0{,}25 * x^5 , \qquad (1)$$

где *x* – порядковый номер года, начиная с 1989 г.

С ее помощью аппроксимация исходных данных может быть осуществлена почти с нулевой погрешностью. В ряде случаев, когда не требуется такая высокая точность регрессионного приближения, можно рекомендовать использование и более простых линий регрессии – полиномиальной функции второй степени (с погрешностью 9%) и логарифмических кривых (с ошибками приблизительно на том же уровне). В качестве первого приближения допускается применение линейного уравнения регрессии (полиномиальной функции первой степени), обеспечивающего уровень ошибки в пределах 24%.

ЛИТЕРАТУРА

1. Ломако, Ю.В. Эпизоотический мониторинг колибактериоза новорожденных телят в Республике Беларусь / Ю.В. Ломако, Н.Н. Андросик. – Ветеринарная медицина Беларуси. – 2002. – № 2. – С. 15–17.
2. Русинович, А.А. Математическая модель-прогноз эпизоотического процесса инфекции ВЛКРС в спонтанных условиях / А.А. Русинович // Весці Акадэміі аграрных навук Рэспублікі Беларусь. – 2001. – № 3. – С. 60–63.

REFERENCES

1. Lomako Yu.V., Androsik N.N. *Veterinarnaya meditsina Belarusi* [Veterinary Medicine of Belarus], 2002, 2, pp. 15–17.
2. Rusinovich A.A. *Vestsi natsiyanalnai akademii navuk Respubliki Belarus* [Journal of the National Academy of Sciences of the Republic of Belarus], 2001, 3, pp. 60–63.

Поступила в редакцию 04.04.2018
Адрес для корреспонденции: e-mail: bomini@mail.ru – Борисевич М.Н.