



БІАЛОГІЯ

УДК 636.2:51-7:001.891.5(476)

Полиномиальные, логарифмические и экспоненциальные математические функции при построении моделей в биологических исследованиях

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 поголовья телят в Республике Беларусь, заболевших колибактериозом. Данные взяты за 10 лет, начиная с 1989 г. и по 1998 г.

Цель работы – построение несложных математических моделей, описывающих реальные процессы в биологических исследованиях.

Материал и методы. Материал: данные официальной статистической отчетности Главного управления ветеринарии Минсельхозпрода Республики Беларусь и Республиканской ветеринарной лаборатории. Использованы методы регрессионного математического анализа.

Результаты и их обсуждение. В результате серии вычислительных экспериментов построены двумерные графические зависимости: по оси абсцисс – годы, по оси ординат – поголовье больных телят (тыс. голов); отмечены значения, заимствованные из материалов официальной статистической отчетности (на примере Главного управления ветеринарии Минсельхозпрода Республики Беларусь и Республиканской ветеринарной лаборатории); 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 E) и экспоненциальной функцией).

Графические зависимости позволяют визуально оценить имеющиеся различия между данными ветеринарной отчетности и данными вычислений, полученными в экспериментах с группой указанных регрессионных моделей.

Заключение. Регрессионные кривые, представленные в статье, описывают большое поголовье новорожденных телят с различной степенью точности. Наилучшего результата можно достичь при использовании полиномиальной функции пятой степени. Аппроксимация статистических данных с ее помощью может быть осуществлена почти с нулевой ошибкой. Для остальных регрессионных моделей картина складывается неудовлетворительная (точность/ошибка): полиномиальные функции первой степени – (18%/82%), второй – (25%/75%); третьей – (28%/72%); четвертой – (21%/79%); логарифмические функции – (10%/90%); экспоненциальная функция – (23%/77). Сказанное позволяет окончательно заключить, что для практических приложений можно с большой долей уверенности рекомендовать математическую модель, описываемую полиномиальной функцией пятой степени. Остальные модели, рассмотренные в данной статье, мало пригодны или почти непригодны для этих целей.

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

Polynomial, Logarithmic and Exponential Mathematical Functions in Building Models in Biological Research

M.N. Borisevich*, P.A. Krasochko*, I.M. Prishchepa**

*Educational Establishment «Vitebsk State Academy of Veterinary Medicine»

**Educational Establishment «Vitebsk State P.M. Masherov University»

In the present article the solution of the problem of creation of regression models of a livestock of calves in the Republic of Belarus, ill with colibacteriosis, is provided. The data are taken over ten years, from 1989 to 1998.

The objective of the research is building simple mathematical models which describe real processes in biological research.

Material and methods. *The research material is data of the official statistical reporting of Head Department of Veterinary Medicine of Ministry of Agriculture and Food Production of the Republic of Belarus and the Republican Veterinary Laboratory. Methods of regressive mathematical analysis are used.*

Findings and their discussion. *As a result of a series of computing experiments two-dimensional graphic dependences are constructed: on an abscissa axis – years, on an axis of ordinates – a livestock of sick calves (thousands of heads); the values are borrowed from materials of the official statistical reporting (on the example of Head Department of Veterinary Medicine of Ministry of Agriculture and Food Production of the Republic of Belarus and Republican Veterinary Laboratory); curves are constructed according to the used regression models (polynomial function of the first, second, third and fourth degrees; the logarithmic function (on the basis of 10); the logarithmic function (on the basis of E) and the exponential function.*

Graphic dependences allow to estimate visually available differences between the data of the veterinary reporting and data of calculations received in experiments with the group of the specified regression models.

Conclusion. *The regression curves presented in article describe sick livestock of newborn calves with various degree of accuracy. The best result can be achieved when using polynomial function of the fifth degree. Approximation of statistical data with its help can be carried out with almost a zero mistake. For the other regression models the picture develops unsatisfactory (accuracy/error: polynomial functions – the first degree (18%/82%), the second – (25%/75%); the third – (28%/72%); the fourth – (21%/79%); logarithmic functions – (10%/90%); exponential function – (23%/77). This allows to conclude finally that for practical applications it is possible to recommend the mathematical model described by the polynomial function of the fifth degree with a big share of confidence. Other models considered in this article are little suitable or almost unsuitable for these purposes.*

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

В настоящей статье приведено решение задачи по построению регрессионных моделей поголовья телят в Республике Беларусь, заболевших колибактериозом. Данные взяты за 10 лет, начиная с 1989 г. и по 1998 г. [1; 2].

Цель работы – построение несложных математических моделей, описывающих реальные процессы в биологических исследованиях.

Материал и методы. Материал: данные официальной статистической отчетности Главного управления ветеринарии Минсельхозпрода Республики Беларусь и Республиканской ветеринарной лаборатории. Использованы методы регрессионного математического анализа.

Результаты и их обсуждение. В результате серии вычислительных экспериментов получены зависимости, представленные 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 больных телят (тыс. голов); точками отмечены значения, заимствованные из материалов официальной статистической отчетности (Главного управления ветеринарии Минсельхозпрода Республики Беларусь и Республиканской ветеринарной лаборатории); сплошные кривые построены в соответствии с используемыми регрессионными моделями (перечень моделей представлен в табл.).

Графические зависимости позволяют визуально оценить имеющиеся различия между данными ветеринарной отчетности и данными вычислений, полученными в 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 – трехмодельная группа из двух логарифмических (десятичной и натуральной) и одной экспоненциальной моделей).

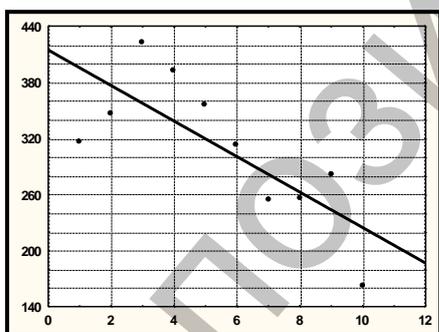
Полиномиальная функция первой степени (рис. 1, а) (по-другому ее называют линейной функцией или просто линейной регрессионной моделью) неудовлетворительно описывает фактические данные – точки разбросаны почти по всей области рисунка, ни одна из них не принадлежит уравнению регрессии. Вследствие этого имеет место и невысокое значение коэффициента детерминации (для анализируемой кривой он равен 0,577). Фактические значения, соответствующие статистической отчетности, описываются с крайне низкой

точностью – 18%, при этом погрешность аппроксимации недопустимо большая ($1 - \eta = 82\%$). Вместе с тем модель регрессии имеет простой вид, конструкция уравнения обеспечивается только двумя коэффициентами – $\theta_0 = 414,43$ и $\theta_1 = -18,99$. Рекомендовать кривую к практическому применению не имеет смысла – слишком велика погрешность и низка точность модельного представления.

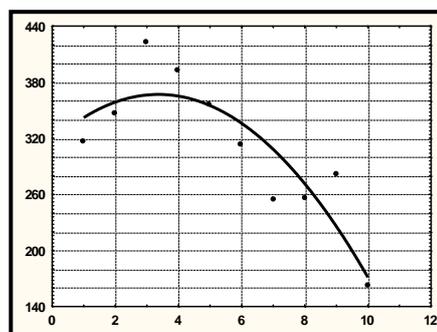
Полиномиальная функция второй степени (изображена на рис. 1, б) относится к классу нелинейных регрессионных моделей. Для ее конструирования требуются три модельных коэффициента – $\theta_0 = 316,70$, $\theta_1 = 29,87$ и $\theta_2 = -4,44$ (последние, как известно, находятся из фактических данных по формулам регрессионного анализа). Коэффициенты определяют структуру регрессионного уравнения, применяемого в анализе с целью аппроксимации исходных данных, выполняя в нем роль параметров. Анализ рис. 1, б показывает, в частности, что для данного вида регрессии степень приближения теоретических точек к точкам фактическим может быть оценена как неудовлетворительная (точность – 25%, ошибка – 75%; для первого случая эти же показатели были равны соответственно 18% и 82%). Коэффициент детерминации 0,777. Это выше, чем в предыдущем случае, но недостаточно для корректного описания исходных данных. По обозначенной причине применять полученные здесь формулы для практических вычислений не следует.

Полиномиальная функция третьей степени (рис. 1, в) лишь немногим лучше предыдущей функции (коэффициент детерминации – 0,868, точность приближения – 28%, ошибка аппроксимации – 72%). Функция сложнее квадратичной (полиномиальной функции второй степени). Для ее задания требуется уже четыре параметра – $\theta_0 = 233,06$, $\theta_1 = 104,06$, $\theta_2 = -20,52$ и $\theta_3 = 0,97$. Перечисленные факторы не позволяют рекомендовать эту модель к практическому применению.

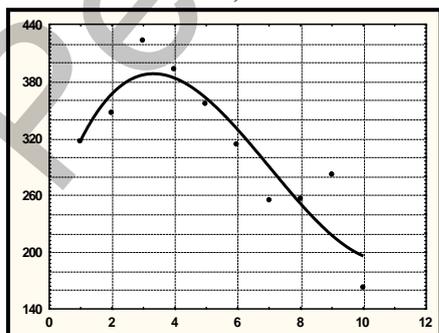
Полиномиальная функция четвертой степени (рис. 1, г) еще хуже определяет исходные точки по сравнению с предыдущей моделью. Для нее коэффициент детерминации $D = 0,650$. Хотя для корректного построения этой модели требуется пять параметров (а не восемь или десять) ($\theta_0 = 94,67$, $\theta_1 = 281,48$, $\theta_2 = -85,44$, $\theta_3 = 9,85$ и $\theta_4 = -0,40$), ее трудно рекомендовать к практическому использованию. С ее помощью достигается лишь 21%-ная точность совпадения с фактически данными, уровень погрешностей достаточно высок – 79%.



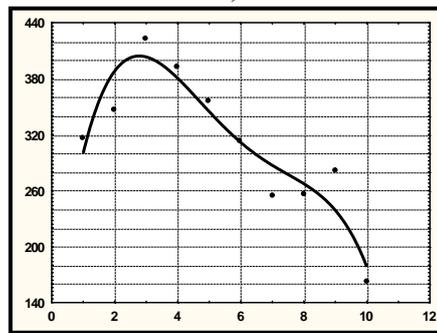
а)



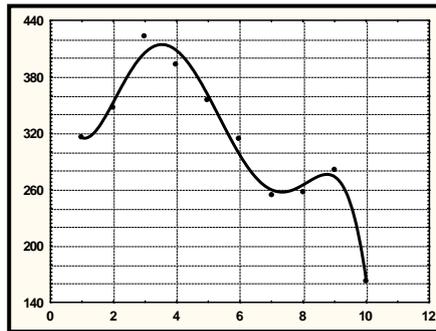
б)



в)

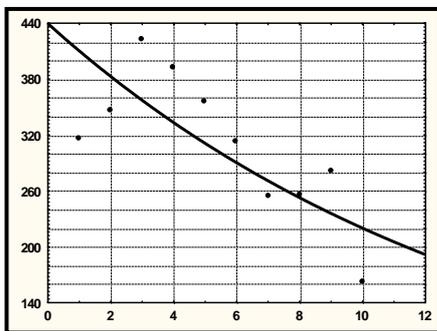


г)

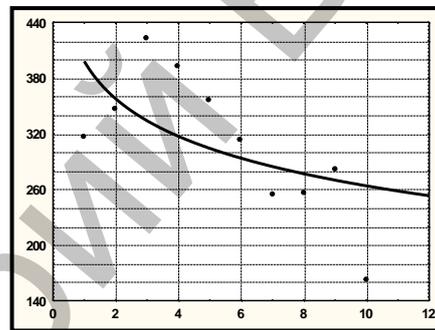


д)

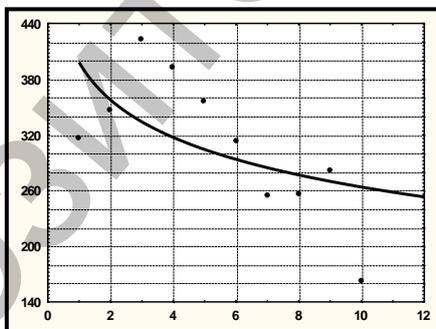
Рис. 1. Полиномиальные функции различных степеней: первой (а), второй (б), третьей (в), четвертой (г), пятой (д).



а)



б)



в)

Рис. 2. Логарифмические функции: по основанию 10 (а), по основанию e (б); экспоненциальная функция (в).

Наилучшее приближение к статистическим данным достигается при помощи полиномиальной функции пятой степени (рис. 1, д). Ее конструкция требует максимального числа параметров – шести ($\theta_0 = 450,25$, $\theta_1 = -304,68$, $\theta_2 = 222,27$, $\theta_3 = -58,94$, $\theta_4 = 6,43$ и $\theta_5 = -0,25$). За счет этого обеспечивается почти 100%-ная точность воспроизведения отчетных данных, ошибка воспроизведения равна нулю. Модель является идеальной для практических применений.

Резюмируя анализ рис. 1, отметим, что в целом полиномиальная функция не пригодна для аппроксимации данных ветеринарной отчетности. Для нее характерен низкий уровень точности регрессионных кривых (18% – для линейной модели, 25% – для квадратичной, 28% – для кубической, 21% – для четверичной). Исключение составляет лишь функция пятой степени – она описывает данные с 100%-ной точностью и может быть рекомендована для использования на практике без ограничений.

Таблиця

Вид регресии, 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	$y = 414,43 - 18,99 * x$	29751,16	51602	0,577	0,18
2.	Полиномиальная функция второй степени	$y = 316,70 + 29,87 * x - 4,44 * x^2$	40097,3	51602	0,777	0,25
3.	Полиномиальная функция третьей степени	$y = 233,06 + 104,06 * x - 20,52 * x^2 + 0,97 * x^3$	44804,33	51602	0,868	0,28
4.	Полиномиальная функция четвертой степени	$y = 94,67 + 281,48 * x - 85,44 * x^2 + 9,85 * x^3 - 0,40 * x^4$	33551,47	51602	0,650	0,21
5.	Полиномиальная функция пятой степени	$y = 450,25 - 304,68 * x + 222,27 * x^2 - 58,94 * x^3 + 6,43 * x^4 - 0,25 * x^5$	161994,8	51602	3,139	1,00
6.	Логарифмическая функция (по основанию 10)	$y = 398,10 - 134,32 * \log(x)$	16456,17	51602	0,318	0,10
7.	Логарифмическая функция (по основанию e)	$y = 398,10 - 58,34 * \ln(x)$	16459,31	51602	0,318	0,10
8.	Экспоненциальная функция	$y = 439,67 * \exp(-0,07 * x)$	37571,40	51602	0,728	0,23

Перейдем теперь к анализу зависимостей рис. 2 (здесь рассматриваются три регрессионные модели: логарифмическая по основанию 10 (рис. 2, а), логарифмическая по основанию e (рис. 2, б) и экспоненциальная (рис. 2, в).

Логарифмические функции (рис. 2, а и б) непригодны для аналитического приближения данных (значения коэффициента детерминации одинаковы и равны 0,318). Хотя аналитический вид функций несложен (для построения моделей требуются всего два параметра θ_0 и θ_1), погрешность, с которой они описывают большое поголовье животных, чрезмерно велика (составляет 90%). Одновременно точность описания данных крайне низкая – 10%.

Экспоненциальная кривая (рис. 2, в) дает лучший результат (коэффициент детерминации – 0,728, точность приближения – 23%, ошибка – 77%). По сравнению с логарифмическими функциями она решает задачу аппроксимации почти в 2,3 раза надежнее. Несмотря на это, руководствоваться ею на практике не следует.

Таким образом, следует отметить, что регрессионные кривые, рассмотренные в статье, описывают большое поголовье новорожденных телят с различной степенью точности. Наилучшего результата можно достичь при использовании полиномиальной функции пятой степени:

$$y = 450,25 - 304,68 * x + 222,27 * x^2 - 58,94 * x^3 + 6,43 * x^4 - 0,25 * x^5, \quad (1)$$

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

Заключение. Аппроксимация статистических данных с помощью (1) может быть осуществлена почти с нулевой ошибкой. Для остальных регрессионных моделей картина складывается неудовлетворительная (точность/ошибка): полиномиальные функции первой степени – (18%/82%), второй – (25%/75%); третьей – (28%/72%); четвертой – (21%/79%); логарифмические функции – (10%/90%); экспоненциальная функция – (23%/77%).

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. Русиневич, А.А. Математическая модель-прогноз эпизоотического процесса инфекции ВЛКРС в спонтанных условиях / А.А. Русиневич // Весті 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. *Vesti 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.