

РЕГРЕССИОННЫЙ АНАЛИЗ РЕЗУЛЬТАТОВ КОНТРОЛЬНЫХ РАБОТ СТУДЕНТОВ ПО ТЕОРИИ ВЕРОЯТНОСТЕЙ

Е.С. Соболева
науч. рук. *Ю.Я. Кацман*
(г. Томск, Томский политехнический университет)
ess18@tpu.ru

REGRESSION ANALYSIS OF THE RESULTS OF EXAMINATIONS OF STUDENTS IN PROBABILITY THEORY

E.S. Soboleva
(Tomsk, Tomsk Polytechnik University)

Annotation. In this paper, is carried a study of the dependence of the results of tests with the help of linear and nonlinear regression analysis. As a result of an earlier statistical analysis with the same data, was found the relationship between them.

Keywords: regression analysis, research, dependent variable, model, coefficient of determination.

Введение. В настоящий момент методы сбора и обработки числовых данных, которые являются сутью статистического исследования, нужны для повседневной жизни в современном цивилизованном обществе. Аппарат математической статистики является наиболее мощным инструментом для отсеивания закономерностей от случайностей.

Целью данной работы является исследование влияния одной или нескольких независимых переменных на зависимую переменную (корреляционный анализ), построение и оценка качества стохастической модели, а также оценка параметров этой модели. Все исследования проведены с использованием пакета Statistica V10 [1].

Предпосылкой для применения регрессионного анализа стало выявление связи между переменными (оценками), полученными в результате выполнения контрольных работ [2]. Это доказывают полученные коэффициенты ранговой корреляции Спирмена и Кенделла, которые значимо отличаются от нуля на 5% -ном уровне [3].

Линейный регрессионный анализ. Одной из возможностей регрессионного анализа является предсказание зависимой переменной по одной или нескольким независимым переменным. Теоретическое линейное уравнение имеет вид [3]:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m + \varepsilon,$$

где Y – зависимая (объясняемая переменная), $X = (X_1, X_2, \dots, X_m)$ – вектор независимых (объясняющих) переменных, $\beta = (\beta_0, \beta_1, \dots, \beta_m)$ – вектор размерности $(m + 1)$ неизвестных параметров, ε – свободный коэффициент.

В исходных данных были представлены баллы (оценки) по 4 контрольным. Результаты анализа представлены в табл. 1.

Таблица 1. Линейная регрессионная модель ($m = 3$)

N = 160	R = 0,43287211, R ² = 0,18737826, Adjusted R ² = 0,17702639					
	Beta	Std. Err. of Beta	B	Std. Err. of B	t(157)	p-level
Intercept			7,338876	1,449179	5,064160	0,000001
Баллы_2	0,198831	0,075648	0,203742	0,077516	2,628387	0,009431
Баллы_3	0,327932	0,075648	0,340242	0,078487	4,334993	0,000026

Данные, приведенные в таблице, являются лучшим результатом проведенных исследований. Само уравнение в данном случае имеет вид:

$$\text{Баллы}_4 = 7,338876 + 0,203742 * \text{Баллы}_2 + 0,340242 * \text{Баллы}_3.$$

Значение R^2 численно выражает долю вариации зависимой переменной, объясненную с помощью регрессионного уравнения. Данное уравнение описывает всего 18,7% вариации зависимой переменной, что свидетельствует о том, что линейная модель не подходит для описания данной зависимости. Эти выводы подтверждаются результатами [4], где показана некорректность построения линейной модели для оценки качества обучения.

Нелинейный регрессионный анализ. Помимо линейной, связь между двумя признаками может быть и нелинейной. В общем виде нелинейная регрессионная модель может быть описана уравнением в полиномиальной форме:

$$Y = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \dots + \beta_m x_i^m.$$

Кроме полиномиальной формы широко используются логарифмическая, гиперболическая и тригонометрическая модели. Правильный выбор вида модели является опорной точкой для качественного ее анализа. На практике неизвестно, какая модель окажется корректной. Поэтому зачастую подбирают такую модель, которая наиболее точно соответствует реальным данным. В табл. 2 представлены характеристики модели.

Таблица 2. Характеристики нелинейной регрессионной модели

N = 277856	R = 0,66291196, R ² = 0,43945226, Adjusted R ² = 0,43944016					
	Beta	Std. Err. of Beta	B	Std. Err. of B	z	p-level
Intercept			-196,017	0,900687	-217,630	0,00
Баллы_1	-2,46086	0,016440	-2,782	0,018588	-149,690	0,00
Баллы_2	-1,70351	0,023141	-1,652	0,022444	-73,616	0,00
Баллы_3	-2,83961	0,022040	-2,998	0,023272	-128,837	0,00
LNБаллы_1	2,87395	0,016464	48,415	0,277356	174,559	0,00
LNБаллы_2	1,61755	0,023542	23,713	0,345131	68,708	0,00
LNБаллы_3	3,07862	0,022533	48,360	0,353956	136,626	0,00

Уравнение имеет вид:

$$\text{Баллы}_4 = -196,017 - 2,782 * \text{Баллы}_1 - 1,652 * \text{Баллы}_2 - 2,998 * \text{Баллы}_3 + 48,415 * \ln \text{Баллы}_1 - 23,713 * \ln \text{Баллы}_2 + 48,360 * \ln \text{Баллы}_3.$$

Данное уравнение описывает ~ 44% вариаций зависимой переменной.

Анализ полученных результатов. Оценивая значения коэффициентов детерминации в случае линейного и нелинейного регрессионного анализа, можно сделать вывод о том, что полученные модели достаточно далеки от идеальных. Однако, нелинейное уравнение описывает модель на порядок лучше, т.е. модель является нелинейной.

Заключение. С помощью программы Statistica был произведен регрессионный анализ реальных данных. Исходные данные – оценки студентов кафедры ВТ за контрольные работы по теории вероятностей. Полученные результаты говорят о том, что можно узнать результат четвертой контрольной работы по первым трем с вероятностью 44%.

ЛИТЕРАТУРА

1. Боровиков В. Statistica. Искусство анализа данных на компьютере: Для профессионалов. 2-е изд. – СПб.: Питер, 2003. – 688 с.
2. Соболева Е.С. Статистический анализ результатов контрольных работ студентов по теории вероятностей./Сб. труд. XIII Межд.науч.-практ. конф. «Молодежь и современные ин-

формационные технологии». Томск, 9-13 ноября 2015 г. – 2016-Томск: Изд-во ТПУ. – Т.1, с. 69 – 70.

3. Кобзарь А.И. Прикладная математическая статистика.–М.: Физматлит, 2006.–816 с.

4. Кацман Ю.Я., Лепустин А.В., Илюхин Б.В. Влияние контекстных факторов на оценку результатов эффективности работы школ томской области [Электронный ресурс] // Современные проблемы науки и образования. – 2014. – №. 6. – С. 1-11. - Режим доступа: <http://www.science-education.ru/120-16117>

ТИПЫ DDOS-АТАК, МЕТОДЫ ПРОФИЛАКТИКИ И ЗАЩИТЫ ОТ НИХ

Фролов С.Г., Демин А.Ю.

(г. Томск, Томский политехнический университет)

TYPES OF DDOS-ATTACKS, METHODS OF PREVENTING AND DEFENDING AGAINST IT

Frolov S.G., Demin A.U.

(Tomsk, Tomsk Polytechnic University)

DDoS is the most popular type of attack on the company for disabling its service. This article presents classification of the most popular types of DDoS and methods of defending and preventing from it.

Keywords: DoS attack, DDoS attack, defending methods, methods of prevention of DDoS.

DoS-атака или атака типа «отказ в обслуживании» направлена на вычислительную систему с целью создать такие условия, при которых пользователи системы не могут получить данные к определенным ресурсам или сервисам. Одновременная атака с большого числа компьютеров свидетельствует о DDoS-атаке – распределенной атаке типа «отказ в обслуживании». Атаки выполняются с помощью зараженных специальными программами компьютеров, которые часто называют «компьютерами-зомби». [1]

Классификация DDoS-атак и защита от них. Все DDoS-атаки можно классифицировать на три больших типа:

- атаки, направленные на заполнение канала;
- атаки, использующие уязвимости протоколов;
- атаки, использующие уязвимости приложений.

Атаки, направленные на заполнение канала. Этот тип атак направлен на забивание полосы пропускания. Интенсивность данного вида атак измеряется в битах в секунду. В данный тип атак входят такие разновидности флудов, как UDP флуд, ICMP флуд и прочие направленные потоки фальшивых пакетов.

Существует способ защиты от таких атак — фильтрация паразитного трафика на уровне ЦОД или специализированных сервисов защиты. Для фильтрования остатков паразитного трафика рекомендуется также применять аппаратную защиту.

Атаки, использующие уязвимости протоколов. Эта категория направлена на существующие ограничения различного оборудования или уязвимости сетевых протоколов. Атаки данного типа засоряют ресурсы оборудования сфальсифицированными пакетами, в результате чего система оказывается неспособной обрабатывать полезный трафик. Сила атаки измеряется в пакетах в секунду. К этому типу атак относятся SYN флуд, *Ping of Death* и т.д.

Против таких атак самым эффективным средством является аппаратная защита. Специально разработанные устройства для фильтрации трафика помогут отсеять паразитный трафик от полезного.

Атаки на уровне приложений. Атаки данного типа направлены на различные уязвимости, которые присутствуют в программном обеспечении. Они приводят к выходу из строя какого-либо приложения или операционной системы в целом. Типичным представителем