

**Басалаева Анастасия Юрьевна**

студентка

**Гареева Гульнара Альбертовна**

канд. пед. наук, доцент

**Григорьева Диана Рамилевна**

канд. пед. наук, доцент

ФГАОУ ВО «Казанский (Приволжский)

федеральный университет»

г. Казань, Республика Татарстан

## **ПОСТРОЕНИЕ И ПРИМЕНЕНИЕ ЛОГИСТИЧЕСКОЙ РЕГРЕССИОННОЙ МОДЕЛИ В НЕФТЯНОЙ ПРОМЫШЛЕННОСТИ**

**Аннотация:** нефтедобывающие компании при добыче нефти встречаются с различными осложнениями. В работе уточняется, что логистическая регрессионная модель применяется для выявления парафиновых отложений на нефтепромысловых объектах и позволяет выявлять и предупреждать неисправности.

**Ключевые слова:** логистическая регрессионная модель, парафиновые отложения, нефтепромысловые объекты, логистическая функция.

В практике эксплуатации скважин нефтедобывающие компании встречаются с различными осложнениями (асфальтосмолопарафиновые отложения, вынос песка и образование песчаных пробок, отложение солей). Наиболее серьезные осложнения возникают в связи с отложениями асфальтенов, смол и парафинов (асфальтосмолопарафиновые отложения). В составе нефти, добываемой на месторождениях России, содержится различное количество высокомолекулярных соединений, таких как парафины, смолы, асфальтены. Особенno эта проблема актуальна для месторождений Татарстана и Башкортостана. Вследствие отложения асфальтосмолопарафинов уменьшается приток нефти, увеличивается нагрузка на станки-качалки (обрыв штанг, повышение расхода энергии, увеличение утечек через плунжер и клапан и, соответственно, уменьшение

коэффициента подачи) и ЭЦН (уменьшение производительности, увеличение затрат энергии, перегрев электродвигателя) [4].

Для выявления отложений парафинов в скважинах, предлагается использовать метод логистической регрессии. Это позволит своевременно выявлять отложения парафинов и смол на стенках стволов скважин, на штангах, выкидных линиях и нефтепромысловом оборудовании.

Логистическая регрессия – это статистическая модель, применяемая для предсказания вероятности возникновения некоторого события по значениям множества переменных, как количественных, так и категориальных [2, с. 321]. Для этого вводится так называемая зависимая переменная  $y$ , принимающая лишь одно из двух значений – как правило, это числа 0 (событие не произошло) и 1 (событие произошло), и множество независимых переменных (также называемых признаками, предикторами или регрессорами) – вещественных  $x_1, x_2, \dots, x_n$ , на основе значений которых требуется вычислить вероятность принятия того или иного значения зависимой переменной. Делается предположение о том, что вероятность наступления события  $y = 1$  равна:

$P\{y = 1 / x\} = f(z)$ , где  $z = \theta^T x = \theta_1 x_1 + \theta_n x_n$ ,  $x$  и  $\theta$  – векторы-столбцы значений независимых переменных  $x_1, \dots, x_n$  и параметров (коэффициентов регрессии) – вещественных чисел  $\theta_1, \dots, \theta_n$ , соответственно, а  $f(z)$  – так называемая логистическая функция:  $f(z) = \frac{1}{1+e^{-z}}$ . Логистическая функция представлена на рисунке 1 [3, с. 231].

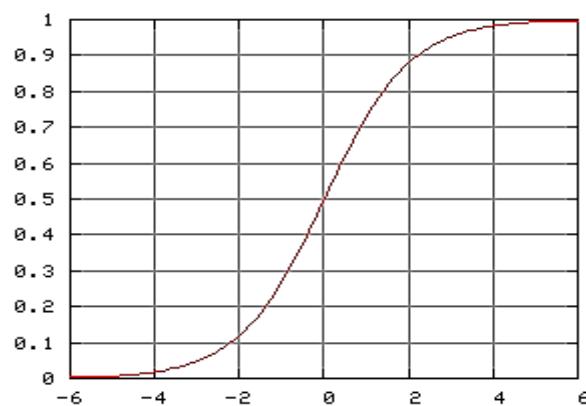


Рис. 1. Логистическая функция

В прикладном статистическом анализе логистическая регрессия используется для решения двух задач: моделирования взаимосвязи и классификации наблюдений. Логистическую регрессию применяют при проведении клинических исследований в медицине, для моделирования поведения покупателей и различных сферах. Рассчитать модель логистической регрессии можно практически во всех программах для профессионального статистического анализа данных таких, как SPSS, SAS, R, Statistica и других [1, с. 15].

В качестве признаков для построения модели были использованы признаки давление и пиковая нагрузка на насос, где  $p$  – наличие отложений,  $n$  – отсутствие отложений (табл. 1).

Таблица 1

## Обучающая выборка

<b>id</b> <i>(идентификатор нефтепромыслового объекта)</i>	<b>conclusion</b> <i>(результат)</i>	<b>pressure</b> <i>(давление)</i>	<b>peak load</b> <i>(пиковая нагрузка)</i>
2908	<i>p</i>	217	4013
1231	<i>n</i>	123	3000
4543	<i>n</i>	109	3202
4545	<i>n</i>	132	3589
6564	<i>n</i>	93	2630
6554	<i>p</i>	229	4700
6561	<i>p</i>	197	4329
6754	<i>n</i>	102	3122
7887	<i>n</i>	89	2938
7658	<i>n</i>	90	3124
7878	<i>n</i>	118	3319
9099	<i>p</i>	203	4567
9098	<i>n</i>	121	3410
2053	<i>p</i>	230	4783
6890	<i>p</i>	243	4957
9087	<i>n</i>	101	3109
8765	<i>p</i>	200	4567
9876	<i>n</i>	101	3105
0909	<i>n</i>	113	3240
3217	<i>n</i>	97	3089
5464	<i>p</i>	213	4856
7908	<i>n</i>	105	3248
6436	<i>n</i>	111	3301
6742	<i>n</i>	92	3001
5632	<i>n</i>	94	3121
1088	<i>p</i>	257	4976
7889	<i>p</i>	254	4889
9877	<i>n</i>	76	2785
9000	<i>n</i>	79	2856
9021	<i>n</i>	107	3287

Получены коэффициенты регрессии  $\theta$  для построения логистической кривой. Теперь модель готова для выявления парафиновых отложений на интересующих нефтяных промысловых объектах. Скрипт реализован на языке *R*. Реализация регрессионной логистической модели изображено на рисунке 2.

```
model<-glm(formula = clients$V2[i] ~ clients$V3[i] + clients$V4[i],
family = binomial, data = clients) // построение модели
teta=coef(model)
```

Рис. 2. Модель на языке R

Нанесем облако точек на график, изображенный на рисунке 3. Скважины с наличием парафиновых отложений будут окрашены в синий цвет, а без отложений в красный.

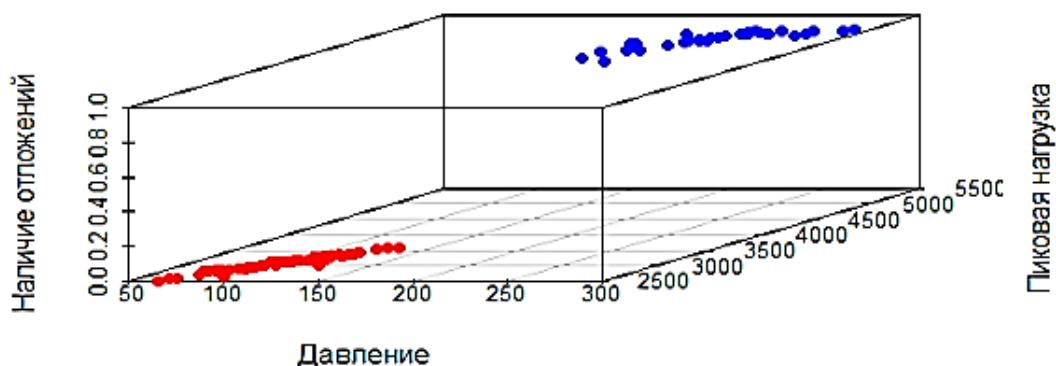


Рис. 3. Зависимость наличия отложений в скважине  
от давления и пиковой нагрузки

Определим наличие парафиновых отложений с помощью полученной модели. Полученные результаты изображены в таблице 2.

Таблица 2

Результат работы модели

давление	пиковая нагрузка	вероятность
217	4013	0,994221
123	3000	0,022532
229	4700	0,997389
92	3001	0,019326
293	5021	0,998372

Модель показала результат без ошибок и достаточно точно.

Логистическая регрессионная модель с высокой степенью вероятности подходит для выявления парафиновых отложений на нефтепромысловых объектах и позволяет обнаруживать и предупреждать неисправности. Вовремя обнаруженное критическое состояние объектов позволяет сэкономить миллионы рублей нефтедобывающим компаниям.

### ***Список литературы***

1. Hosmer D.W., Lemeshow S. Applied logistic regression. NY: John Wiley & Sons, 2013.
2. Harrell Frank. Regression modeling strategies. NY: Springer, 2012.
3. Елисеева И.И. Статистика: Учебник для вузов / И.И. Елисеева; отв. ред. И.И. Елисеева. – М.: Юрайт, 2012. – 565 с.
4. Причины отложения асфальтенов, смол и парафинов в скважинах коммуникациях. Методы удаления АСПО [Электронный ресурс]. – Режим доступа: <https://studopedia.org/11-35810.html>