

Заключение по содержанию диссертации

Савватеев Алексей Владимирович

(Ф.И.О. члена диссертационного совета)

Волков Никита Алексеевич

(Ф.И.О. соискателя ученой степени)

О некоторых свойствах вероятностных распределений

и их применении в задачах машинного обучения

(Название диссертации, ученая степень, на которую представлена диссертация, специальность)

представлена на соискание степени кандидата физико-математических наук по специальности 05.13.17 (теоретические основы информатики)

Дата защиты 15.12.20

Оценка соответствия диссертации требованиям Положения о присуждении ученых степеней кандидата наук, доктора наук в МФТИ (далее - Положение):

1. Актуальность тематики диссертации:

Изучению свойств биномиального и пуассоновского распределений ранее были посвящены различные работы, в частности, исследовалось поведение возле медианы. В случае биномиального распределения эта задача мотивируется широко известной гипотезой Самуэльса об экстремуме вероятности малого отклонения (на константу) суммы независимых случайных величин от ее математического ожидания. Первая часть диссертации посвящена именно исследованию свойств биномиального распределения вблизи медианы. Вторая часть диссертации посвящена оцениванию параметров смеси многомерных распределений Стьюдента. Распределение Стьюдента играет большую роль в современном анализе данных, в частности при наличии выбросов в данных. Теоретические результаты диссертации были применены в задаче оценки качества проб пластовых флюидов, которая важна для работы инженеров-геологов.

2. Научная новизна выносимых на защиту результатов:

Все результаты, полученные в диссертации, являются новыми. Получено существенное обобщение ранее известного свойства биномиального распределения на случай параметров n , $b/(n+c)$, где c – число из отрезка $[0,1]$. Ранее аналогичный результат был известен лишь для $c=0$. Разработана процедура оценки параметров смеси многомерных распределений Стьюдента по выборкам с пропусками. Реализован метод машинного обучения для решения различных задач, показано его применение для оценки качества пластовых флюидов.

3. Теоретическая и практическая значимость диссертационной работы:

Диссертационная работа носит как теоретический, так и практический характер. Ее теоретические результаты могут быть использованы в теории вероятностей и математической статистике. Практические результаты могут быть использованы для решения различных задач анализа данных.

4. Полнота опубликования основных результатов диссертации в рецензируемых научных изданиях в соответствии с требованиями Положения:

Все результаты диссертации опубликованы: имеется 5 статей в рецензируемых журналах (все входят в перечень ВАК), три из которых индексируются Scopus. Работа апробирована на различных представительских семинарах.

5. Вопросы и замечания (в соответствии с п. 4.13 Положения соискатель отвечает на сформулированные здесь вопросы и замечания на заседании по защите диссертации):

Почему при исследовании биномиального распределения рассматривается только случай c из $(0,1]$? Что происходит при $c > 1$?

Почему в модели смеси распределений Стьюдента число степеней свободы считается известным? Можно ли подбирать число степеней свободы по данным? Могут ли разные компоненты иметь разное число степеней свободы?

Можно ли сделать оценку параметров смесью нормальных распределений, а затем — прорядить кластеры?

6. Общая характеристика диссертации (не включает резолютивную часть):

В диссертации Н.А. Волкову удалось существенно обобщить ранее известные результаты о поведении функции биномиального распределения возле медианы. Кроме того, Н.А. Волков обобщил известные ранее методы оценки параметров, что позволяет их использовать для данных, в которых есть пропуски и выбросы. Наконец, Н.А. Волков разработал метод машинного обучения, который успешно может быть применен геологами. Достоверность теоретических выводов и заключений, сформулированных в диссертации, подтверждается строгими математическими доказательствами. Адекватность практических выводов, сформулированных в диссертации, подтверждается численными метриками качества в сравнении с другими моделями. Диссертация является научно-квалификационной работой, результаты которой вносят весомый вклад в развитие теории вероятностей, математической статистики, анализа данных.

Дата

01.12.2020г.

Подпись

А.С.А.У.

/ расшифровка (полностью)

