

NAMANGAN INSTITUTE OF ENGINEERING AND
TECHNOLOGY
M.V.LOMONOSOV MOSCOW STATE UNIVERSITY,
TASHKENT BRANCH
INSTITUTE OF MATHEMATICS named after V.I.
ROMANOVSKY OF THE ACADEMY OF SCIENCES OF THE
REPUBLIC OF UZBEKISTAN
INDIAN STATISTICAL INSTITUTE

PROCEEDINGS

of International Scientific Conference

STATISTICS

and its applications



19-20 October, 2022
Namangan Institute of Engineering and Technology
Namangan

УДК 519.2

Труды международной научной конференции "СТАТИСТИКА и ее применения-2022". Под редакцией профессора А.А.Абдушукурова - Наманган, Наманганский инженерно-технологический институт, 2022, 454 стр.

Организационный комитет:

Маматкаримов О.О. - Председатель, Ректор Наманганского инженерно-технологического института, профессор;

Аюпов Ш.А. - Сопредседатель, Директор Института математики АН РУз, академик;;

Часовских А.А. - Сопредседатель, Руководитель филиала МГУ, доцент;

Абдушукуров А.А. - Сопредседатель, профессор кафедры ПМИИ филиала МГУ (Организатор);

Sanghamitra Bandyopadhyay - Сопредседатель, Директор Индийского института статистики, профессор;

Маманазаров А.Б. - Заместитель председателя, заместитель руководителя филиала МГУ, доцент;

Дергач П.С. - Заместитель председателя, 1-й заместитель руководителя филиала МГУ;

Эргашев О.К. - Заместитель председателя, Проректор по научной работе и инновациям НамИТИ, профессор;

Сайдаматов Э.М. - Заместитель председателя, Исполнительный директор филиала МГУ, доцент;

Ботиров Г.И. - Заместитель председателя, Заместитель директора Института математики АН РУз, доцент;

Debasis Sengupta - Заместитель председателя, Декан учебных заведений Индийского статистического института, профессор.

Члены оргкомитета:

Мелибоев У.Х. - Проректор по учебной работе НамИТИ, доцент;

Каюмов М.У. - Проректор по работе с молодёжью НамИТИ, доцент;

Султонов С.Ю. - Проректор по финансам и экономике НамИТИ, доцент;

Маманазаров А.Б. - Заместитель руководителя филиала МГУ, доцент;

Каршиев Т.О. - Заместитель директора по учебной и научной работе, доцент;

Хусанбоев Я.М. - Заведующий научной лабораторией "Стохастический анализ" Института математики АН РУз, профессор;

Ayanendranath Basu -Заведующий кафедрой "Междисциплинарные статистические исследования" Индийского статистического института, профессор;

Обидов А.А. - Декан инженерно-технологического факультета НамИТИ, профессор;

Маткаримов К.Ж. - Декан экономического факультета НамИТИ, профессор;

Эргашев Ж.С. - Декан факультета технологии легкой промышленности НамИТИ, DSc.;

Мамаханов А.А. - Декан факультета автоматике и энергетике НамИТИ, DSc.;
 Уктамов Д. - Декан химико-технологического факультета НамИТИ, доцент;
 Пирназаров А. - Декан факультета технологии сельскохозяйственной продукции НамИТИ, DSc.;
 Мурадов Р.С. - Заведующий кафедрой "Высшая математика" НамИТИ, доцент (Организатор);
 Рахматуллаев М.М. - Заведующий Наманганским филиалом Института математики АН РУз, с.н.с.;
 Машраббоев А. - Заведующий кафедрой "Математический анализ" НамГУ, доцент;
 Рахманов А. - Доцент кафедры "Высшая математика" НамИТИ;
 Камолидинов М. - Доцент кафедры "Высшая математика" НамИТИ;
 Холмуродов М. - Доцент кафедры "Математический анализ" НамГУ;
 Шоюсупов Ш. - Доцент кафедры "Высшая математика" НамИТИ;
 Сайпиддинов Ш. - Стар. Пр. кафедры "Высшая математика" НамИТИ;
 Исмоилов Н. - Стар. Пр. кафедры "Высшая математика" НамИТИ;
 Абдурахмонов А. - Стар. Пр. кафедры "Высшая математика" НамИТИ;

Программный комитет:

Алимов Ш.А. - Академик АН РУз. (Узбекистан);
 Садуллаев А.С. - Академик АН РУз. (Узбекистан);
 Орлов А.И. - Профессор МГТУ им. Н.Э.Баумана (Россия);
 Subhamoy Maitra - Профессор Indian Statistical Institute (Индия);
 Чилин В.И. - Профессор НУУз (Узбекистан);
 Зупаров Т.М. - Профессор НУУз (Узбекистан);
 Лемешко Б.Ю. - Профессор НГТУ (Россия);
 V.L.S. Prakasa Rao - Профессор AIMSCS (Индия);
 Арипов М.М. - Профессор НУУз (Узбекистан);
 Шарахметов Ш. - Профессор Институт математики АН РУз (Узбекистан);
 Чупрунов А. - Профессор КФУ (Россия);
 Алексей Мин - Профессор Technical University of Munich (Германия);
 Базаров Т.Ю. - Профессор филиала МГУ (Узбекистан);

Азамов А.А. - Академик АН РУз. (Узбекистан);
 Фармонов Ш.К. - Академик АН РУз. (Узбекистан);
 Somnath Datta - Профессор University of Florida (США);
 Allan Gut - Профессор Uppsala University (Швеция);
 Ejaz Ahmed - Профессор Brock University (Канада);
 Peihua Qiu - Профессор University of Florida (США);
 Гафуров М.У. - Профессор ТГТУ (Узбекистан);
 Qing Lu - Профессор University of Florida (США);
 Назаров А.А. - Профессор НИТГУ (Россия);
 Ибрагимов Р. - Профессор Imperial College London (Великобритания);
 Мирахмедов Ш. - Профессор Институт математики АН РУз (Узбекистан);
 Arup Bose - Профессор Indian Statistical Institute (Индия);
 Воинов В. - Профессор KIMEP University (Казахстан);

Алоев Р.Д. - Профессор НУУз (Узбекистан);

Семенова Д.В. - Профессор СФУ (Россия);

Ходжибаев В. - Профессор НамИСИ (Узбекистан);

Сагидуллаев К.С. - Доцент Технический институт ЁДЖУ (Узбекистан);

Имомов А. - Доцент КарГУ (Узбекистан);

Нурмухамедова Н.С. - Доцент НУУз (Узбекистан);

Хамдамов И. - Доцент НУУз (Узбекистан);

Уринов А. - Доцент филиала МГУ (Узбекистан);

Розиков У. - Профессор Институт математики АН РУз (Узбекистан);

Шарипов О. - Профессор НУУз (Узбекистан);

Biswabrata Pradhan - Доцент Indian Statistical Institute (Индия);

Чичагов В.В. - Доцент ПГНИУ (Россия);

Abhik Ghosh - Доцент Indian Statistical Institute (Индия);

Турсунов Г.Т. - Доцент НУУз (Узбекистан);

Альхамов Р.Р. - Доцент Институт математики АН РУз (Узбекистан);

Жалилов Б. - И.о. доцент НамИТИ (Узбекистан);

Секции конференции

1. Математическая статистика.

2. Вероятность и смежные вопросы.

3. Дискретная и непрерывная математика;

4. Математические - статистические методы в экономике, медицине, естествознании и методике преподавания.

5. Фундаментальная информатика и информационные технологии.

ISBN 978-9943-305-86-1

Спонсор:

Конференция проходит при поддержке Международного узбекско-индийского совместного гранта UZB-Ind-2021-97.

UDK 519.2

Proceedings of the international scientific conference "STATISTICS and its applications-2022". Editor of Professor A.A.Abdushukurov - Namangan, Namangan Institute of Engineering and Technology, 2022, 454 pp.

Organizing Committee:

Mamatkarimov O.O. - Chairman, Rector of the Namangan Institute of Engineering and Technology, Professor;

Ayupov Sh.A. - Co-Chairman, Director of the Institute of Mathematics of the Academy of Sciences of the Republic of Uzbekistan, academician;

Chasovskikh A.A. - Co-Chairman, Head of the branch of Moscow State University, Associate Professor;

Abdushukurov A.A. - Co-Chairman, Professor of the Department of PMI of the Moscow State University Branch (Organizer);

Sanghamitra Bandyopadhyay - Co-Chairman, Director of the Indian Institute of Statistics, Professor;

Mamanazarov A.B. - Vice-Chairman, Deputy Head of the branch of Moscow State University, Associate Professor;

Dergach P.S. - Vice-Chairman, 1st Deputy Head of the branch of Moscow State University;

Ergashev O.K. - Vice-Chairman, Vice-Rector for Research and Innovation NamMITI, Professor;

Saydamatov E.M. - Vice-Chairman, Executive Director of the branch of Moscow State University, Associate Professor;

Botirov G.I. - Vice-Chairman, Deputy Director of the Institute of Mathematics of the Academy of Sciences of the Republic of Uzbekistan, Associate Professor;

Debasis Sengupta - Vice-Chairman, Dean of Educational Institutions of the Indian Statistical Institute, Professor.

Organizing committee members:

Meliboev U.Kh. - Vice-Rector for Academic Affairs NamMITI, Associate Professor;

Kayumov M.U. - Vice-rector for work with youth at NamMITI, associate professor;

Sultonov S.Yu. - Vice-Rector for Finance and Economics NamMITI, Associate Professor;

Mamanazarov A.B. - Deputy head of the Moscow State University branch, associate professor;

Karshiev T.O. - Deputy director for academic and scientific work, associate professor;

Khusanboev Ya.M. - Head of the scientific laboratory "Stochastic analysis" of the Institute of Mathematics of the Academy of Sciences of the Republic of Uzbekistan, professor;

Ayanendranath Basu - Head of the Department of Interdisciplinary Statistical Research, Indian Statistical Institute, Professor;

Obidov A.A. - Dean of the Faculty of Engineering and Technology NamMITI, Professor;

Matkarimov K.Zh. - Dean of the Faculty of Economics NamMITI, professor;

Ergashev Zh.S. - Dean of the Faculty of Light Industry Technology NamMITI, DSc.;
 Mamakhanov A.A. - Dean of the Faculty of Automation and Energy NamMITI, DSc.;
 Uktamov D. - Dean of the Faculty of Chemical Technology NamMITI, Associate Professor;
 Pirnazarov A. - Dean of the Faculty of Agricultural Technology NamMITI, DSc.;
 Muradov R.S. - Head of the Department "Higher Mathematics" NamMITI, Associate Professor (Organizer);
 Rakhmatullaev M.M. - Head of the Namangan branch of the Institute of Mathematics of the Academy of Sciences of the Republic of Uzbekistan, Senior Researcher;
 Mashrabboev A. - Head of the Department of Mathematical Analysis, NamSU, Associate Professor;
 Rakhmanov A. - Associate Professor of the Department "Higher Mathematics" NamITI;
 Kamolidinov M. - Associate Professor of the Department "Higher Mathematics" NamITI;
 Kholmurodov M. - Associate Professor of the Department of Mathematical Analysis, NamSU;
 Shoyusupov Sh. - Associate Professor of the Department "Higher Mathematics" NamITI;
 Saipiddinov Sh. - Star. Etc. Department "Higher Mathematics" NamMITI;
 Ismoilov N. - Star. Etc. Department "Higher Mathematics" NamMITI;
 Abdurakhmonov A. - Star. Etc. Department "Higher Mathematics" NamMITI;

Program Committee:

Alimov Sh.A. - Academician of the Academy of Sciences of the Republic of Uzbekistan. (Uzbekistan);
 Sadullaev A.S. - Academician of the Academy of Sciences of the Republic of Uzbekistan. (Uzbekistan);
 Orlov A.I. - Professor of MSTU im. N.E. Bauman (Russia);
 Subhamoy Maitra - Professor of the Indian Statistical Institute (India);
 Chilin V.I. - Professor NUUZ (Uzbekistan);
 Zuparov T.M. - Professor NUUZ (Uzbekistan);
 Lemeshko B.Yu. - Professor of NSTU (Russia);
 B.L.S. Prakasa Rao - Professor AIMSCS (India);
 Aripov M.M. - Professor NUUZ (Uzbekistan);
 Sharakhmetov Sh. - Professor Institute of Mathematics of the Academy of Sciences of the Republic of Uzbekistan (Uzbekistan);

Azamov A.A. - Academician of the Academy of Sciences of the Republic of Uzbekistan. (Uzbekistan);
 Farmonov Sh.K. - Academician of the Academy of Sciences of the Republic of Uzbekistan. (Uzbekistan);
 Somnath Datta - Professor of University of Florida (USA);
 Allan Gut - Professor of Uppsala University (Sweden);
 Ejaz Ahmed - Professor of Brock University (Canada);
 Peihua Qiu - Professor of University of Florida (USA);
 Gafurov M.U. - Professor of TSTU (Uzbekistan);
 Qing Lu - Professor of the University of Florida (USA);
 Nazarov A.A. - Professor NITSU (Russia);
 Ibragimov R. - Professor Imperial College London (Great Britain);
 Arup Bose - Professor Indian Statistical Institute (India);

Chuprunov A. - Professor KFU (Russia);
 Aleksey Min - Professor Technical University of Munich (Germany);
 Bazarov T.Yu. - Professor of branches of Moscow State University (Uzbekistan);
 Alov R.D. - Professor NUUz (Uzbekistan);
 Semenova D.V. - Professor SibFU (Russia);
 Khodzhibaev V. - Professor NamISI (Uzbekistan);
 Sagidullaev K.S. - Associate Professor of the Yeoju Technical Institute (Uzbekistan);
 Imomov A. - Associate Professor of KarSU (Uzbekistan);
 Nurmukhamedova N.S. - Associate Professor of NUUz (Uzbekistan);
 Khamdamov I. - Associate Professor of NUUz (Uzbekistan);
 Urinov A. - Associate Professor of branches of Moscow State University (Uzbekistan);
 Zhalilov B. - Acting Associate Professor NamITI (Uzbekistan);

Mirakhmedov Sh. - Professor Institute of Mathematics of the Academy of Sciences of the Republic of Uzbekistan (Uzbekistan);
 Voinov V. - Professor of KIMEP University (Kazakhstan);
 Rozikov U. - Professor of the Institute of Mathematics of the Academy of Sciences of the Republic of Uzbekistan (Uzbekistan);
 Sharipov O. - Professor of the National University of Uzbekistan (Uzbekistan);
 Biswabrata Pradhan - Associate Professor of the Indian Statistical Institute (India);
 Chichagov V.V. - Associate Professor of PSNIU (Russia);
 Abhik Ghosh - Associate Professor of Indian Statistical Institute (India);
 Tursunov G.T. - Associate Professor of NUUz (Uzbekistan);
 Alkhamov R.R. - Associate Professor Institute of Mathematics of the Academy of Sciences of the Republic of Uzbekistan (Uzbekistan);

Section of conference

1. Mathematical Statistics.
2. Probability and related problems.
3. Discrete and continuous mathematics;
4. Mathematical-Statistical methods in economics, medicine, natural sciences and teaching methodology.
5. Fundamental informatics and information technology;.

ISBN 978-9943-305-86-1

Sponsor:

The conference is supported by the International Uzbek-Indian joint grant UZB-Ind-2021-97.

ПРЕДИСЛОВИЕ

Конференция "СТАТИСТИКА и ее применения" была приурочена к Всемирному Дню Статистики, который отмечается каждый год 20 октября, начиная с 2010 года во многих странах в соответствии с Резолюцией Генеральной Ассамблеи ООН. Конференция призвана объединить ученых не только Республики Узбекистан, которые занимаются актуальными вопросами статистики и ее практического применения, а также и других стран. I научно-практическая конференция "СТАТИСТИКА и ее применения" проводилась 17-18 октября 2012 года, в материалы которой были включены 86 статей. 17-18 октября 2013 года была организована II научно-практическая конференция "СТАТИСТИКА и ее применения". Конференция была приурочена к 22-годовщине Независимости Республики Узбекистан, а также тому, что 2013 год был объявлен Всемирным Годом Статистики. Материалы конференции включили в себя 87 статей. III научно-практическая конференция "СТАТИСТИКА и ее применения" проводилась 16-17 октября 2015 года, и в материалах конференции было опубликовано 110 статей.

25-27 апреля 2016 года был проведен совместный научно-практический семинар "СТАТИСТИКА и ее применения" кафедры "Теория Вероятностей и Математическая Статистика" НУУз и отдела "Биостатистика" Университета Флориды США с участием профессоров Сомнат Датта, Сусмита Датта и их учеников. IV конференция "СТАТИСТИКА и ее применения" проводилась 19-20 октября 2017 года. В материалах этой конференции были представлены 90 статей. V конференция "СТАТИСТИКА и ее применения" проводилась 17-18 октября 2019 года в Филиале МГУ в г. Ташкенте и она была приурочена также и 60 летию проф. А.А.Абдушукурова, профессора филиала и основного организатора этих конференций.

Настоящая, VI международная конференция "СТАТИСТИКА и ее применения" проводится 19-20 октября 2022 года в Наманганском инженерно-технологическом институте совместно с Филиалом МГУ имени М.В. Ломоносова в городе Ташкенте, Институтом математики им. В.И.Романовского АН РУз. и Индийским институтом статистики. Труды данной конференции включили в себя 88 работ не только ученых из Узбекистана но также и США, Индии, Великобритании, Франции, Кувейта и России. Оргкомитет выражает глубокую благодарность всем участникам конференции любезно представившим свои научные работы в данный сборник.

Оргкомитет

PREFACE

The conference "STATISTICS and its applications" was dedicated to the World Day of Statistics which celebrated on October 20 every year, since 2010 in many countries according to the UN General Assembly Resolution. Conference is aimed to collaboration of scientists not only of the Republic of Uzbekistan, which are engaged in current issues of statistics and its practical application, but also from other countries. I scientific-applied conference "STATISTICS and its applications" was held on 17-18 October, 2012 and materials of the conference include 86 articles. On 17-18 October, 2013 was organized II scientific-applied conference "STATISTICS and its applications". This conference was dedicated to the 22nd anniversary of Independence of the Republic of Uzbekistan, , also that 2013 year declared the International Year of Statistics. Materials of this conference consists of 87 articles. III scientific-applied conference "STATISTICS and its applications" was held on 16-17 October, 2015 and materials of the conference include 110 articles.

25-27 April 2016 was organizes jointly scientific-applied seminar "STATISTICS and its applications" of the Department of "Probability Theory and Mathematical Statistics" of National University of Uzbekistan and Department of "Biostatistics" of University of Florida of USA with participating of professors Somnath Datta and Susmita Datta and their graduate students. IV scientific-applied conference "STATISTICS and its applications" organized on 19-20 October, 2017. Materials of this conference include 90 papers. V conference "STATISTICS and its applications" was organized on October 17-18, 2019 in Tashkent Branch of MSU and it was dedicated also to 60th Birthday of Prof. A.A.Abdushukurov, the Principle Organizer of these conferences.

This, VI International conference "STATISTICS and its applications" organized on October 19-20, 2022 in Namangan Institute of Engineering and Technology together with the Branch of Moscow State University named after M.V. Lomonosov in the city of Tashkent, the Institute of Mathematics named after V.I. Romanovsky of the Academy of Sciences of the Republic of Uzbekistan and the Indian Statistical Institute. Proceedings of this conference include 88 papers of scientists not only from Uzbekistan, also from USA, India, UK, France, KUWAIT and Russia. The organizing committee expresses its deep appreciation to all the participants of the conference who kingly submit their articles to this book.

Organizing Committee

СПОСОБ СГЕНЕРИРОВАТЬ ТЕКСТ С ПОМОЩЬЮ ЦЕПИ МАРКОВА

Ёдгоров У., Эркинов Ф., Юлдашов А.

ТАШКЕНТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ УЗБЕКСКОГО
ЯЗЫКА И ЛИТЕРАТУРЫ ИМЕНИ АЛИШЕРА НАВОИ

yodgorov@navoiy-uni.uz

Существует много генераторов, которые генерируют какой-то случайный текст. Проблема с этими генераторами в том, что нужно вручную прописывать все варианты, последовательности и комбинации из слов и предложений. Главный минус такого подхода - вы делаете всю работу за компьютер. Вы сами жестко задаете структуру каждого предложения. Такой текст вас ничем не удивит, потому что вы четко знаете, сколько предложений у вас в шаблоне и из каких частей состоит каждое. Чтобы генерировать много похожего текста, но без таких заморочек, используют нейросети или алгоритмы на цепях Маркова.

Что такое цепь Маркова. Представьте, что у нас есть набор каких-то событий, связанных друг с другом. Например, первое наступает только после второго, второе - после третьего или четвертого, а третье - после четвертого с вероятностью 30

Текст и цепи Маркова. Возьмем такую скороговорку: "Osmonda ikkita kalxat. Birining oq dumi kalta kalxat, birining qora dumi kalta kalxat. Oq dumi kalta kalxat qora dumi kalta kalxatga xalaqit beradi, qora dumi kalta kalxat oq dumi kalta kalxatga xalaqit beradi." В теории цепей Маркова это называется корпус - исходный материал, по которому будут составляться взаимосвязи. В нашем случае по этому предложению мы поймем, как одни слова будут связаны с другими. Для простоты мы опустим знаки препинания и большие буквы. В нормальных алгоритмах это, конечно, учитывается, но пока мы учимся, можно и так. Наша задача - понять на этом примере, как устроены и как работают цепи Маркова. Для этого мы шаг за шагом смоделируем работу алгоритма и посмотрим, что получится в итоге.

Посчитаем, сколько раз встречается каждое слово в нашем тексте:

Osmonda-1

ikkita-1

kalxat-5

kalxatga-2

birining-2

oq-3

dumi-6

kalta-6

qora-3

xalaqit-2

beradi-2

Еще нам нужно пометить, что после "beradi" ничего нет, просто ставим точку.

Такие отдельные элементы в цепи Маркова называют звеньями. А цепь как раз и состоит из звеньев.

Теперь составим пары связанных событий. Связные - это когда одно событие идет только после конкретного другого события. В нашем тексте первая пара будет такая:

Osmonda → ikkita

Так как в исходном тексте слово "Osmonda" только одно и после него идет "ikkita", то алгоритм запомнит, что после "Osmonda" нужно обязательно поставить только "ikkita", без вариантов.

А вот со следующей парой уже поинтереснее: kalxat → (birining, oq, qora)

Получается, что наш алгоритм, когда дойдет до "kalxat", с какой-то долей вероятности выберет одно из трех следующих слов.

Сразу разберемся, как считать доли вероятности - это пригодится нам на следующем этапе. У нас есть пара kalxat → (birining, oq, qora), при этом "birining" у нас встречается 2 раз, "oq" - 3 раза, и "qora" - 3 раза, а всего эти слова вместе в тексте встречаются 7 раз. Получается, что после "kalxat" вероятность продолжения текста:

-словом "birining" - 2/7,

-словом "oq" - 3/7,

-словом "qora" - 3/7.

Именно потому, что существуют вероятности, текст в цепях Маркова может каждый раз получаться разным. Алгоритм с разной вероятностью случайным образом идет по разным путям и дает нам каждый раз разный текст. Значения этих вероятностей пригодятся нам при составлении схемы связей. Составим точно так же остальные пары:

ikkita → kalxat

birining → (oq, qora)

oq → dumi

dumi → kalta

kalta → kalxat

kalxat → (birining, oq, qora)

qora → dumi

kalta → kalxatga

kalxatga → xalaqit

xalaqit → beradi.

Алгоритм с той или иной долей вероятности выбирает продолжение для каждого слова, и в результате у нас получается разный текст.

Как сделать текст более осмысленным. У нас в примере алгоритм всегда выдает чушь, потому что исходный текст (корпус) слишком маленький для составления связей и в нем мало слов. Если мы возьмем текст побольше, то алгоритм сможет составлять текст, похожий на язык оригинала, как будто его написал человек. Еще можно сделать так: составлять пары не из одних слов, а из словосочетаний. Мы перебираем все сочетания из двух слов, которые есть в тексте, и составляем пары к каждому из них. Это более сложно и ресурсоемко, зато текст получается похож на человеческий. Если комбинировать, например, цепочки из пар и одного слова, получится более читаемый текст (или нет, тут дело случая).

TEXT GENERATOR BASED ON MARKOV CHAINS

Yodgorov U.

TASHKENT STATE UNIVERSITY OF UZBEK LANGUAGE AND LITERATURE

yodgorov@navoiy-uni.uz

This article gives a general idea of how to generate texts using Markov process modeling. We will master the basic concept of a Markov chain, and also define what a corpus, links, probability distribution and histograms are in our context.

One fish two fish red fish blue fish

This sentence is the corpus, that is, the base on the basis of which the text will be generated in the future. It consists of eight words, but there are only five unique words - these are links (we are talking about a Markov chain, after all). For clarity, let's color each link in its own color:

One fish two fish red fish blue fish

Write down the number of occurrences of each of the links in the text:

One: 1
 fish: 4
 two: 1
 red: 1
 blue: 1

It can be seen that the word "fish" appears in the text 4 times more often than each of the other words ("One" "two" "red" "blue"). That is, the probability of meeting the word fish in our corpus is 4 times higher than the probability of meeting every other word from the figure. Speaking in the language of mathematics, we can determine the law of distribution of a random variable and calculate with what probability one of the words will appear in the text after the current one. The probability is calculated as follows: we need to divide the number of occurrences of the word we need in the corpus by the total number of all words in it. For the word "fish" this probability is 50

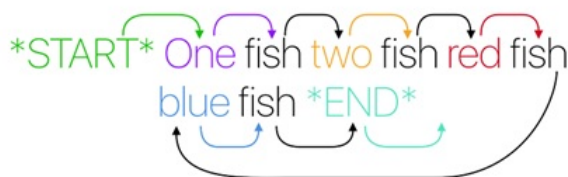
Now let's add to our text elements that are always implied, but not called in everyday speech - the beginning and the end of a sentence:

START One fish two fish red fish blue fish *END*

Any sentence contains these invisible "beginning" and "end" let's add them as links to our distribution:

START: 1
 One: 1
 fish: 4
 two: 1
 red: 1
 blue: 1
 END: 1

We model a process in which the state of the system at the next moment of time depends only on its state at the current moment, and does not depend in any way on all previous states. Imagine that you have a window in front of you that displays only the current state of the system (in our case, this is one word), and you need to determine what the next word will be, based only on the data presented in this window. In our corpus, the words follow one after the other according to the following pattern:



Thus, pairs of words are formed (even the end of the sentence has its own pair - an empty value):

- (*START*, One)
- (One, fish)
- (fish, two)
- (two, fish)
- (fish, red)
- (red, fish)
- (fish, blue)
- (blue, fish)
- (fish, *END*)
- (*END*, none)

Let's group these pairs by the first word. We will see that each word has its own set of links that, in the context of our sentence, can follow it:

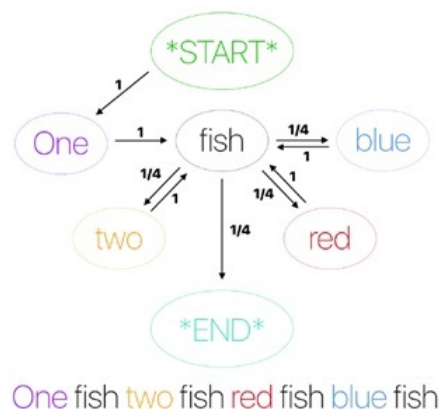
- (*START*, One)
- (One, fish)
- (fish, two) (fish, red) (fish, blue) (fish, *END*)
- (two, fish)

(red, fish)
 (blue, fish)
 (*END*, none)

Let's present this information in a different way - each link will be assigned an array of all words that may appear in the text after this link:

START : [One]
 One : [fish]
 Fish : [two, red, blue, *END*]
 two : [fish]
 red : [fish]
 blue : [fish]
 END : [none]

We see that each link has words that can come after it in a sentence. If we were to show the diagram above to someone else, that person could, with some probability, reconstruct our initial sentence, i.e. the corpus. We have just modeled a Markov process - we have determined each next word only on the basis of knowledge about the current one. For complete assimilation of the material, let's build diagrams that show the dependencies between the elements within our corpus. Ovals represent links. The arrows lead to potential links that can follow the word in the oval. Near each arrow - the probability with which the next link will appear after the current one:



So far, we have only considered one-word windows. You can increase the size of the window so that the text generator produces more "verified" sentences. This means that the larger the window, the less deviation from the corpus will be during generation. Increasing the window size corresponds to the transition of the Markov chain to a higher order.

A window is that data in the current state of the system that is used to make decisions. If we combine a large window and a small data set, then we are likely to get the same sentence every time. Let's take the dictionary base from our first example and expand the window to size 2:

```
(*Start*, one) : [fish : 1]
(one, fish)     : [two : 1]
(fish, two)     : [fish : 1]
(two, fish)     : [red : 1]
(fish, red)     : [fish : 1]
(red, fish)     : [blue : 1]
(fish, blue)    : [fish : 1]
(blue, fish)    : [*END* : 1]
```

The extension has resulted in each window now having only one option for the next system state - no matter what we do, we will always get the same sentence, identical to our corpus. Therefore, in order to experiment with windows, and for the text generator to return unique content, stock up on a vocabulary base of 500,000 words or more.

Оглавление

Предисловие	8
<i>Abdushukurov A.A., Muradov R.S.</i>	
About the works of V.I. Romanovsky in the field of mathematical statistics	15
<i>B.L.S. Prakasa Rao</i>	
Cramer-Rao Inequality Revisited for Randomly Censored Data	25
<i>Ayanendranath Basu</i>	
Minimum Distance and Robust Inference: A Review	29
<i>Abhik Ghosh, Ayanendranath Basu</i>	
Robust inference under the Parametric proportional hazard models in Survival Analyses	33
<i>Subhasish Basak, Partha Sarathi Mukherjee</i>	
Image denoising by jump regression and local clustering	39
<i>Jayant Jha, Meysam Hashemi, Anirudh Nihalani Vattikonda, Huifang Wang, Viktor Jirsa</i>	
Fully Bayesian estimation of virtual brain parameters with self-tuning Hamiltonian Monte Carlo	41
<i>Nilotpall Sanyal</i>	
Iterative variable selection for high-dimensional data with binary outcomes	48
<i>Sancharee Basak</i>	
On the "Optimal" Density Power Divergence Tuning Parameter	56
<i>Soumya Chakraborty</i>	
A Comparative Review of some Robust Tests of Hypothesis based on Density Power Divergence	63
<i>Семенова Д.В., Куликов В.Р.</i>	
Моделирование систем событий с заданной структурой независимости	71

<i>Уринов А.А., Таштемирова Д.Х.</i>	
Прогнозирование ожидаемых экономических изменений с помощью корреляционно-регрессионного анализа	394
<i>Уринов А.А., Дلاءерова В.А.</i>	
Корреляционно - регрессионный анализ как способ прогнозирования экономического развития предприятия	399
<i>Azizov Sh.M., Mirzakarimov M.M.</i>	
Yangi konstruksiyadagi jinda tolaning shikastlanishi bo'yicha excelda statistik taxlil	403
<i>Azizov Sh.M., Uzoqov F.G'.</i>	
Kolosnik modelini loyihalash	406
<i>Алимов О.</i>	
Искусственный интеллект в экономике	413
<i>Bekberganova M.R.</i>	
Hayot sug'urtasiga bo'lgan talabni modellashtirish yo'nalishlari	420
<i>Fayzullaev D.M.</i>	
O'zbekistonda elektron tijoratni rivojlantirishning ustuvor yo'nalishlari	423
<i>Hamroyeva D.Y.</i>	
Hayot davomiyligiga ta'sir qiluvchi omillar. hayot sug'urtasini rivojlantirishda jahon tajribasi	430
<i>Ядгаров Т.Г., Абдуганиева О.И.</i>	
Метод применения двоичной системы исчисления в теории вероятностей и математической статистики .	434
<i>Ёдгоров У., Эркинов Ф., Юлдашов А.</i>	
Способ сгенерировать текст с помощью цепи Маркова	437
<i>Yodgorov U.</i>	
Text generator based on Markov chains	441
<i>Джаббаров Р.А.</i>	
Ўзбекистонда рақамли иқтисодиётни ривожланиши-ни статистик таҳлили	446
<i>Холмуродов М.К., Бахромов Р.</i>	
Эҳтимоллардан тузилган қаторларнинг яқкинлашиш тезликлари	448

Холмуродов М.К., Асқаров А.

Индекслари кўп ўлчовчи тасодифий миқдорлар йиғиндиси учун катта сонлар қонунида асимптотик масалалар	450
Приложения	451