Сделано в [Whisper AI](https://t.me/WhisperSummaryAI_bot)

Теория информации, Григорьев А.А., Лекция 03, 19.09.20

Чистая версия. Дата создания: ⁨03.10.2025⁩

с искажениями, но занимаемся пока фигней неравенствами Чебышева. Значит, я вам в прошлый раз доказал, допоминальник, я вам в прошлый раз доказал неравенство Чебышева общее, согласно которому вероятность того, что случайная величина х превышает некое А, она, видите ли, не превышает мат ожидания от φ от х поделить на φ от А. где φ любая функция такая, что значит, φ от x положительная, не отрицательная, больше или равна нулю, и из условия x больше а, то φ от x больше а. Вот приличная оценка. Мы ее в прошлый раз же использовали для того, чтобы доказать, границу Чебышева для дисперсии. Значит, мы поступили как? Мы взяли случайную величину х и ее матожидание m. Смотрели случайную величину х минус m. Ну, вот такую. Ну, и стало быть, в качестве функции f от х мы взяли х квадрат. Почему бы и нет? Подлеводворяет всем поставленным требованиям. Положительное и монотонное. Ну, в общем, для положительных величин монотонная. Ну и поэтому, значит, вероятность того, что x минус... Отличие x, случайной величины, от мат ожидания превышает а. То есть вероятность большого укронения получается так. Меньше или равно мат ожидания x. Мат ожидания x минус m. То есть это мат ожидания... Вот так, от модуль х-м, у меня вот такая случайная величина, в квадрате. Но это то же самое, что мат ожидания величины сигма квадрата. Так что в числителе у меня стоит сигма квадрата. А в знаменателе, ну, вот А квадрат. Вот это граница Чебышева для дисперсии. Кхм. Кхм. Итак, вероятность того, что случайная величина уклоняется от математического ожидания от среднего больше, чем на А, она не превышает вот этой величины. Где σ² – это дисперсия этой случайной величины, а А – это вот это число. Значит, А большое, то вероятность эта становится маленькой. Вот такая вот граница. И мы, значит, дальше так. Мы с вами в свое время рассмотрели... Вот такую на первой этой лекции было. Мы рассмотрели случайную величину эссента, которая есть сумма большого числа случайных величин х одинаково распределенных статистически независимых. Поделить на n. Ну, то есть случайная величина, которая получена как сумма большого числа случайных величин, ну, деленная на n. Значит, если каждая из случайных величин х-житая имеет математическое ожидание одинаковые они все, а дисперсия этой случайной величины, дисперсия от хжитая минус m в квадрате, математическое ожидание составляет сигма квадрат, то есть вот тут у меня случайные величины у всех мат ожидания m, а дисперсия сигма квадрата, они все одинаково распределены. То, что я могу сказать о дисперсии и мат ожидания вот этой случайной величины, такой вот Значит, мы показали, что мат ожидания этой суммы составляет m. Ну, банально, мат ожидания каждой составляет m, их n штук на n сокращается. То есть числитель получается m умножить на n в знаменателе n, поэтому n. А вот мат ожидания иосцентра минус m в квадрате Это я взял эту случайную величину, отнял от нее мат ожидания, получил случайную величину с нулевым уже мат ожиданием и вычисляю ее дисперсию. Вот это оно равно сигма квадрат делить на n. Ну, то есть вот эта случайная величина обладает, видите ли, маленькой дисперсией. Это у нас был закон больших чисел, слабый закон больших чисел, согласно которому, если суммируешь большое число независимых случайных величин с одинаковой дисперсией, то дисперсия вот этой нормированной суммы падает. Она при стремящемся к бесконечности стремится к нулю. Ну и получается вот так. Ну а теперь мы применим к этому всему делу границу Чебышева. Мы будем заниматься... Вот она, собственно, у нас уже есть. Значит, можно теперь... Прямо вот сюда написать, что вероятность того, что случайная величина S nt минус n по модулю превышает какое-то a, эта вероятность не превышает сигма квадрата, где сигма квадрат — это дисперсия вот этой случайной величины. А она вон у меня выписана. Это сигма квадрат делить на n. Вот оно получается так, сигма квадрат поделить на n на а. Квадрат. Вот это границы Чебышева для дисперсии суммы большого числа случайных величин. Вот так она выглядит. Если к бесконечности, вот это стремится к нулю, и какое бы а я ни взял, А это от хранения модуля вот этой суммы, от мотожидания этой суммы. Если какое бы ухранение я не взял, маленькое, при достаточно большом m, вероятность вот этого события, выхода за уклонение, стремится к нулю. То есть на самом деле случайная величина S n t сходится вот к этому мотожиданию m по вероятности. То есть вероятность уклонения равна нулю при достаточно больших n. Вот так вот. Это вот уже сильный закон больших чисел. Ну а теперь мы это все дело вынуждены применить для того, чтобы получить некие границы для вероятности, энтропийная граница для вероятности. Подключаем энтропию. Пока просто тупо неравенство Чебышева и все. Но вот тут, между прочим, так сказать, надо упомянуть, Что еще популярна граница чернова? Это вот когда φ от х берется в виде е в степени мю на х. Квадрат фи от х на е в степени мю получается чернов. Граница чернова. Выглядит она так. Вероятность того, что х превышает а, значение величина превышает а меньше или равно. Вот прямо по этой формуле. Мат ожидания E в степени μx поделить на E в степени μа. Фета. То есть равно E в степени минус μа на мат ожидания от E в степени μx. Интересно то, что она экспоненциальна. Видите, она по А падает, как Е в степени А. У нее параметр мил. Значит, смотрите, ее вычисляют так. Вычисляют вот это мат ожидания. Это на самом деле, это будет производящая функция для моментов. Если экспоненту Е в степени мил x разгожить в ряд, как сумму x, μ в n той степени на n факториал на x в n той степени. А n натурально от нуля до бесконечности. И взять от этого мат ожидания, то есть мат ожидания е в степени μx, это есть вот эта сумма по всем n, μ в n той степени на n факториал так остается, а здесь стоит мат ожидания от x в n той степени. Это момент n. Бывает первый момент, потом второй момент, потом третий. То есть вот тут стоят все возможные средние степеней момента. Ну и поэтому производящая функция моментов, ее надо вычислять, она будет тут какой-то константой. Эта константа есть просто численно, но она оказывается некой функцией от мюнг. И потом, забавность границы Чернова стоит в том, что она параметрическая. Ее можно по этому параметру минимизировать. То есть написать так, что на самом деле P от x больше a меньше или равно минимуму по всем u от этой величины e в степени mua нам от ожидания от e в степени mux. То есть вначале вы вычисляете мат ожидания E в степени μx, получается некая функция от μ. В результате вот здесь тоже возникает некая функция от μ. Вот это все от уха до уха. Ее нам надо минимизировать по всем μ. Получается довольно сильная граница. Но вот в плане домашнего задания, чтобы не сильно думать, я вас попрошу построить не границу черного, а гауссовское распределение. То есть... Возьмем гауссовское распределение P от X. Вот такое. Единица на корень из 2π на e в степени минус x квадрат по полам. Это гауссовское распределение с единичной дисперсией. Гауссовская кривая. Вот такая. Вот P от X. Здесь у меня X. Вот она вот так вот выглядит. Ну и, стало быть, построить, вот взять число А и построить границу Чернова вероятность того, что X... Больше а, ну вот, меньше или равна. Довольно легкая задача. Значит, смотрите, когда станете считать это матожидание, е в степени мюх, вот сюда вот допишите е в степени мюх, интеграга тут берется. Ну, то есть, производящие функции моментов считаются в явном виде. Ну, тут полный квадрат получается, короче говоря, нам добавить константа, так, чтобы в показателе полный квадрат возник. И потом, когда вычислите этот квадрат, производящую функцию момента в основание это вопрос перенормировки этого алгоритма то есть алгоритм когда множишь на константу, у него меняется основание то есть эта вещь не принципиальная она зависит от того в чем вы мерите энтропию в битах или в натах тогда натуральный алгоритм можно в q-битах тогда куличный алгоритм по основанию q но, так как быть, если у нас есть две случайные величины x и y и они статистически независимы есть p от x на пат и на совместное распределение п от x , то есть есть произведение вот такие две случайные величины называются статистически независимыми если их совместное распределение есть произведение распределения x а распределение y ну то есть у нас будет p от xy для краткости есть px умножить на py. Вот если я стану считать энтропию этой исключенной величины, то есть h от xy, то окажется, что так. Это есть минус сумма по всем x и по всем y. pxy на логарифм pxy. Но это pxy я произведению по x на по игрок а варихам произведения от сумма удалив поэтому получается минус сумма по всем x и крик но так напишу по x на по игрок на логарифм x плюс логарифм при но и в этом каждый узнает вот смотрите вот эту сумму когда по игроку просуммирую я единицу получу а по иксу я получаю энтропию То есть это есть h от x плюс h от y. То есть энтропия векторной случайной величины, состоящей из двух статистически независимых координат, есть сумма энтропии. Вот это самое главное свойство энтропии, которое обусловлено тем, что тут стоит его гарифм. Вот почему гаегорифм. Гаегорифм произведения, видите ли, равен сумме гаегорифмов. И поэтому энтропия совместного распределения равна энтропии маргинальных распределений. Ну вот это дальше обобщается на троизвольное число этих самых координат. И если у меня есть векторная случайная величина х, которая есть вектор, состоящий из х1, хн, и у каждой случайной величины есть синтропия, h1, а у этой hн, то h от х большое есть сумма h от х житая по всем g от 1 до n. Ну, при условии, что эти случайные величины статистически независимы. То есть, совместное распределение P от X один эксцентра равно произведению P от X житых по всем G от единицы дарима. Вот если это так, то энтропии суммируются. Ну вот, теперь мы займемся вот чем. Меня будет интересовать... Значит, это я вам... рассказал о том, что энтропия это аддитивная мера. То есть энтропия в совокупности независимых случайных величин равна их сумме энтропии. Теперь мы займемся блоками. То есть у нас будут вот такие блоки. Давайте считать так, что у меня есть координата какая-то х. Это случайная величина. У нее есть энтропия h от х. Вполне определенное сто раз писание. Я образую N блок из этих координат. Буду его так обозначать. Это блок из X1, Xn. И меня будет интересовать вот такая оценка. Меня будет интересовать оценка логарифма 1 на P от X1, Xn. У этого блока есть совместная вероятность. Вот она. p от x1, xn, d. Я буду интересовать минус логарифм этой вероятности. Меня будет интересовать оценка для минус логарифм этой вероятности. Смотрите, что я тут имею. Это самая вероятность. Она ведь на самом деле... Поскольку вот эта вероятность представляется произведением, у меня случайные величины предполагаются независимыми, поэтому это логарифм единицы на произведение. p от x житых. p от 1 до n. А этот логарифм я могу представить суммой логарифмов. Сумма по всем j от 1 до n. Логарифм единицы на p j. Вот она такая, вот случайная величина. То есть, видите, она уже имеет вид вот этой самой суммы, с которой я начинал, с n. Она была вот такая у меня сумма, x1 плюс и так далее, плюс xn, деленная на n. Вот она, эта случайная величина, уже представлена суммой случайных величин, а ставить только на n. Это все дело поделить. Ну и поделим. Вот так. То есть, вот этот нормированный логарифм вероятности представляет собой сумму большого числа случайных величин. Вот таких. Каждая случайная величина, вот она такая. Значит, что я про нее могу сказать? Каково мат ожидания этой случайной величины? А мат ожидания этой самой случайной величины, вот мат ожидания логарифма единица на px. Это есть не более не менее по определению, как сумма P житая на логарифм единицы деленные на P житая. Матожидание этой случайной величины аккуратно равно энтропии. Энтропия на одну координату. То есть у меня были случайные величины, каждая была как-то распределена, у каждой была вот такая энтропия H от X. И вот от ожидания этой величины это и есть эта самая энтропия. Вот так. То есть у меня в числителе стоят случайные величины, у которых среднее значение равно энтропии. Какова действия? Ну вот осталось как-то разобраться с дисперсией. Значит, для этого надо... Рассмотреть случайную величину логарифма единицы на p, g, t. Отнять от нее энтропию. Это ее мат ожидания. Получится вот эта случайная величина с нулевым мат ожиданием. Ее возвести в квадрат и взять от этого делом от ожидания. И вот это обозначим через сигму квадрата. Дисперсия этой самой случайной величины. Итак, я вернулся к той самой диспозиции, что и была. У меня есть сумма большого числа случайных величин. У каждой случайной величины есть мат ожидания h. И у каждой случайной величины есть дисперсия, сигма квадрат. h – это энтропия, а сигма квадрат – это дисперсия. Ну и всё. Теперь я могу просто вот эту оценку применить к оцениванию вот этой величины. Делается так. Я просто тупо переписываю, что там у меня написано. Значит, вероятность того, что единица на n на логарифм единица на p от вектора x, p от x1, xn, минус h. Это мат ожидания. Вот оно, модуль Вот эта величина по модулю больше бета. Я букву А заменю на букву Б. Это меньше или равно по нашей теореме вот этой. Дисперсии этой самой случайной величины. А дисперсии этой самой случайной величины есть сигма квадрат. Сигма квадрат на n на бета квадрат. Вот такая вот оценка. Это и есть энтропийная граница для уклонения логарифма вероятности от энтропии. Вот смотрите, тут написан логарифм вероятности на одну координату, делённую на n, приведённую в логарифм вероятности. Оказывается, что он мало отличается от энтропии, совсем мало. Потому что вероятность того, что вот это отличается от энтропии, более чем на бета, не превышает вот этой величины. А эта величина стремится к нулю при больших Н. Стремится к нулю по Н. По Н, в больших бесконечности. Вот это неравенство, оно у нас будет пользоваться для определения понятия типическое множество. Множество типичных групп. Все, больше мне, собственно, ничего не нужно, кроме вот этой границы. Мы теперь с ней будем разбираться. Значит, мыслим так. У нас есть множество n броков. Множество х, которое состоит из броков. Х1, так далее, хн. Каждая координата этого брока – это случайная величина с энтропией h. Ну и, стало быть, есть какое-то распределение вероятностей на этих блоках. p от x. И вот у нас есть такая граница, что если взять легорифм этого p икса и поделить на n, то получится что-то типа h. Ну вот в этой связи можно попытаться ввести такое определение. Вот в этом множестве давайте выделим множество тех блоков, вероятности которых отличаются от мага то есть назовем и бета это множество так эпичных блоков тб это есть множество блоков x прямо вот из этого пространства вот таких вот блоков таких что вероятность этого самого блока вот так таких что вероятность этого блока то есть Единица на n на логарифм единица на p от х брока минус h по модулю множество и меньше бета. Вот такое вот определение. То есть это множество броков х, приведенные вероятности которых отличаются от энтропии меньше, чем на бета. Это множество типичных блоков. Сейчас окажется, что мы с вами отсюда видим, что вероятность не принадлежать этому блоку, то есть вот у меня множество х большое разбилось на вот это множество tβ, оно состоит из тех блоков, вероятности которых близки к nh, к n умножить на h. И все остальные блоки, tβ с крышкой, которые... имеют нетипичные вероятности. Так вот, вот эта граница говорит, что вероятность от t-beta с чертой, ну то есть данный блок, он либо удовлетворяет вот этой границы, то есть он отличается от аж больше на бета, тогда он лежит вне t-beta, то есть он лежит t-beta с чертой. Так вот, эта граница говорит, что полная мера t-beta с чертой полная вероятность принадлежать не принадлежать множество тб ты лежать вот здесь она ограничена вот этой величиной меньше или равно сигма квадрат на n бета квадрат и стремится к нулю ну то есть оказывается что если я здесь случайно выбираю какие-то блоки они скорее всего с вероятностью близкой к единице попадают вот сюда и А попасть вот сюда, вот из границы Чебышева, вероятность крайне мала, стремится к нулю при больших. Поэтому вот эти броки называются типичными. Брок называется типичными. Можно считать так, что у меня распределение вероятностей на этих броков построено таким образом, что вот здесь лежат приличные вероятности, а здесь где-то бета-чертой, они мавы экспоненциарно. убывают с ростом n, стремятся к нулю. То есть вероятность попасть вот сюда, она никакая. Полная мера вот этого дополнения порядка нуля. Ну а полная мера t-бета, она, естественно, стремится к единице. Значит, полная мера множества t-бета это дополнение до этого множества. Она просто больше или равна единице минус сигма квадрат на n бета квадрат. Ну а что касается оценок самих вероятностей вот этих p от x для элементов этого множества, то они вот легко получаются отсюда. Значит, смотрите. Вот логарифм этой вероятности поделить на n минус h, это по модулю меньше β. Если это расписать, это получается так. Значит, этот самый единица на n, логарифм единицы на p от x, Это меньше или равно и больше или равно. Значит, было меньше или равно бета и больше или равно минус бета. Значит, модуль х, модуль там чего-нибудь, а он меньше а и больше минуса. Поэтому это меньше бета. Вот тут бета напишу, а тут напишу минус бета. еще там аж было вот это я его направо и налево перетащу лишь тут получается эндж плюс бета а тут получается и минус это пока ничего не сделал я просто вот это неравенство расписал на левую и правую граница граница для модуля надо левую границу и правую границу вот мы стали быть а дальше занимаюсь так умножаю все на n вот здесь и вот здесь n исчезает Дальше гагарифм P от X. Я возвожу это в степень. Если это гагарифм по основанию 2, то все возвожу в степень по основанию 2. Получается отсюда вот такая граница. Вероятность P от X меньше или равно 2 в степени минус n на h плюс бета и больше или равно 2 в степени минус n edge минус b ну то есть вот смотрите да это просто выражение этого факта что вы говорили но этих вероятностей мало отличаются от и наш вот они могут отличаются от и наш значит вероятности это они заключены в интервале 2 в степени и на что ты 20 педе но что-то вот это может так написать здесь два степени н-пич на 2 в степени плюс бета только тогда тут плюс, а здесь минус. Вот так. И меньше или равно 2 в степени NH с минусом на 2 в степени минус Nβ. А, плюс Nβ вот здесь, а тут минус Nβ. Ну, то есть, вот это одинаковые вещи, 2 в степени NH, а здесь малая величина 2 в степени минус Nβ, а здесь большая величина 2 в степени плюс Nβ. Ну вот, то есть они вот, ну и стало быть, КБТ маленькая, то фактически вероятность оказывается зажата между вот этими двумя границами. То есть можно сказать, что во множестве типичных броков все вероятности P от X почти одинаковы. То есть каждая P от X внутри этого множества, оно имеет вероятность порядка 2 в степени минус n на h. И отличается от него множителем 2 в степени n-бета в положительную сторону возрастания, либо 2 в степени минус n-бета в сторону убывания. Вот так. Ну то есть получается, что все типичные элементы этого множества почти имеют одинаковую вероятность. Ну то есть вот картина вырисовывается такая. Если у нас есть какое-то множество блоков с антропией h на n, один элемент этого блока, то в нем вырисовывается множество типичных блоков, которые устроены так. Все вероятности во множестве этих типичных блоков примерно одинаковы. То есть они почти все равновероятны. А те, которые не принадлежат множеству типичных блоков, являются нетипичными. Их полная вероятность стремится к нулю, а следовательно и каждая вероятность стремится к нулю. То есть вот эти все имеют нулевую вероятность. Оказывается, что в этом множестве блоков далеко не все блоки встречаются. Мы, собственно, об этом уже много говорили. Вот когда мы станем делать из черных и белых шаров с вероятностью p длинные блоки, то мы увидим, что среди этих блоков в основном встречаются те, у которых отношение числа черных блоков к n совпадает с вероятностью выпадения черного шара, как n равно p. Иначе вероятность стремится к нулю. Ну вот здесь ровно тот же самый эффект. Какое бы у вас ни было распределение на х, оказывается, что появляется множество типичных блоков. Ну и стало быть, элементы этого множества, их вероятность уничтожают верхние и нижние границам. Ну и они, собственно говоря, при малых бета, они почти два в степени минусы наше. То есть почти все одинаковые. Ну а прочие, которые нетипичны, их вероятность стремится к нулю. Вот так. Для нас будет важно следующее, что вот само это множество типичных блоков, оно существует. Нас сейчас будет интересовать оценка мощности этого самого множества. То есть мы заодно сумеем оценить мощность этого множества Тбета, количество элементов в нем. Значит, вот смотрите, я вот это обозначу через Pmax, а вот это обозначу через Pmin. Значит, у меня вероятность P от X заключена между максимально возможной вероятностью, вот она, и минимально возможной вероятностью, вот она. Они отличаются знаком вот тут в показателе. Тут плюс NBT, тут минус NBT. Вот. Значит, воспользовавшись этим представлением, я подумаю, как я могу оценить мощность этого множества. Значит, смотрите, картинка такая. Это количество элементов в нем. Если я это количество умножу на p максимум, то есть я сделаю вид, что все они имеют максимальную вероятность. Максимально возможную вероятность для элемента типичного множества. то эта величина будет меньше единицы. Ну потому что полная вероятность меньше единицы. Это как бы у меня будет верхняя оценка для полной вероятности вот этого множества. Вот эта вероятность не может превышать единицу. Отсюда я получаю оценку на вероятность этого множества. А вот если я это самое t-бета помножу на по минимуму, то по минимуму. и n мими на минимальных вероятность то я получу оценку снизу на полную вероятность этого множества то есть это меньше или равно на на п от тб то смотрите тб то входит какие-то элементы у каждого своя вероятность если просуммирует все эти вероятности то я получу то что называется птб то ну а я взял в тб то может напоминать на по минимум то есть каждый из этих из вероятностей каждого элемент это уже сумма вероятностей настоящих значит в птб то вошли вероятности всех этих было элементов всех этих блоков истинные настоящие а здесь у меня та же самая сумма но все вероятности заменены минимум на минимально возможную поэтому вот так и Мощность Тбэта на P максимум меньше единицы. И мощность Тбэта на P минимум меньше P Тбэта. Это дает мне возможность построить оценки для мощности Тбэта. Значит, P максимум это у меня. Вот оно. 2 в степени минус N на H плюс бэта. H минус бэта. 2 в степени минус n на h минус бета. Отсюда следует, что мощность t бета не превышает. Давайте равно поставим. 2 в степени n на h минус бета. Так, мощность t бета на p максимума. скажу мощность этих самых на табы значит это некая вероятность и она не превышает единицы от обетки на по минус по минимум сейчас подумаю значит как ты я тут немного зашелся сейчас я посмотрю так как там этому делается то вот они оценки для табыта оценки множество так ну ты макс помин ясно что ты бы это на поймать больше или равно птб тогда а птб ты на помин меньше единицы так делается то бета по минимум меньше единицы а вот это больше или равно п а вот я вот в этом месте всегда Начну это поскольку вероятность, то она заведомо не превышает единицу, а вот это больше или равно p от t-бета. Вот так, вот отсюда я буду границы выводить на самом деле. Ну то есть я взял множество t-бета, в нем сколько-то элементов, вот столько, помножим на максимальную вероятность. Заведомо я при этом получил вероятность большую, чем вероятность множества t-бета. А вероятность множества t-бета Я вот уже написал, вот она, 2, то есть P, T, B, то мы с вами оценили. Ну а здесь я как поступил? Я T, B, то умножен на минимальную вероятность, и то, что я получил, заведомо эта вероятность, это оцененная снизу вероятность вот этого множества, она заведомо не больше 1. При минимум это у меня 2 в степени минус N на H плюс бета. Поэтому получается оценка такая. Мощность β, я здесь пишу, она с этой стороны меньше или равна на 2 в степени n на h плюс β. 2 в степени n на h плюс β. А с этой стороны, это я отсюда оценку получу. Это мощность p от t β, это вероятность p от t β на 2 в степени n На h минус бета. Вероятность множества тбета мы оценивали. Она составляет почти единицы. И составляет она единица минус сигма квадрат на n бета квадрат. Вот такие границы у меня получаются для мощности. Окончательно. Они нам сейчас будут сильно нужны, поэтому я их выпишу. Мощность множества типичных последовательностей меньше или равна 2 в степени nH плюс бета и больше или равна 1 минус сигма квадрат на n бета квадрат, вот такая тут поправка, на 2 в степени n на H минус бета. Вот это граница для мощности множества типичных броков. Только их всего. Ну все, значит, после перерыва сейчас мы немножко погуляем. Я займусь доказательством теоремы копирования, опираясь вот на это. Ну, в общем, уже понятно, откуда уши-то растут. Растут вот откуда. Оказывается, что вот это множество типичных броков, оно имеет большую вероятность. Вот такую вероятность, почти единичную. Но элементов в нем относительно мало. два в этой степени а два в степени и наш где аж энтропия и поэтому для кодирования я могу использовать относительно мало битов вот Вот, ладно. А надо продолжать. Наконец, у нас с вами есть все, вот эти границы, для того, чтобы заняться кодированием источника с потерями. Итак, модель все та же самая. Вот я ее просто почищу. Значит, есть сет источник, чего угодно, изображений, звуков. Выражает некие блоки данных. Сроки данных из множества х, имеющие вид х1 и так далее, и хн. Опять все координаторы, вероятно, независимы и имеют энтропию h. Как-то распределены и имеют энтропию h на отдельную координацию. Вот у меня множество этих самых всех блоков. Х. Вот это оно. Идея кодирования встает в следующем. Значит, смотрите, в этом множестве ужасно много блоков. Если я стану передавать все их, вот, выпад такой блок, я его вам кодирую. Ну, я при этом на кодирование буду... Вот если мощность этого множества составляет m, значит, смотрите, если каждая координата принимает q каких-то значений, то всего имеется m, равное q в n-той степени этих самых броков. вот самого m битов. То есть n на логарифм по основанию 2 q битов на элемент. Ну или если это поделить на n, то скорость передачи r, которая определена как логарифм по основанию m делить на n, ну то есть количество битов, затрачиваемых на координатах, оно составит как раз логарифм по основанию 2 q. Вот. Ну вот это совершенно естественно. Я хочу сэкономить, то есть передавать с меньшей скоростью. Для этого есть простое предложение. Вот как уменьшить количество битов, необходимых для кодирования каждого блока? Вот самый простой вариант – это пойти на некие искажения. То есть отказаться от передачи всех блоков, а передавать только те, которые более-менее вероятные. говорим такие слова, что давайте мы в этом множестве всем выберем множество S дельта. Такое, чтобы его вероятность полная, P от S дельта, ну то есть сумма вероятности всех броков, входящих сюда, была порядка единицы. Больше, чем или равно, чем единица минус дельта. Дельта это будет малый параметр в моей теории. Ну то есть вот здесь полная вероятность S, Единица минус дельта, почти единица, а все остальное, оно укладывается в вероятность дельта. Ну, то есть, P от S дельта дополнение — это малая вероятность дельта. И вот мы говорим такие слова потребителю, что если поочечайнее у нас выпало так, что я вам должен сообщить X вот этот, то я буду вместо него передавать что-нибудь отсюда, при этом я буду совершать искажение какое-то передаваемых данных, Но искажение это будет крайне маловероятно. То есть дельта мы будем стремить к нулю. Вот такая вот картина. Ну и, стало быть, хотелось бы выбрать множество из дельта так, чтобы в нем было поменьше элементов. То есть, чтобы мощность из дельта была бы как можно меньше, минимум. Иногда бы я для кодирования того, что я должен передавать без искажений, использовал малое количество битов. Я бы использовал для этого гагарифм по основанию 2 мощности s δ битов. Ну и вот скорость бы моя составляла, если в единицах на n вот такую величину. Ну и стало быть... Гогарифм по основанию s δ делить на n – это моя скорость. Ну вот, ну, стало быть, вот смотрите, предельная скорость, если я передаю всё, она вот такая, а скорость, которую я передаю, если я согласился на потери с вероятностью δ, она вот такая, гогарифм по основанию s δ. Значит, выигрыш в скорости, значит, это отношение этой скорости к этой скорости, Это ровно отношение гаорифма мощности S δ к гаорифму мощности всего множества. Ну, то есть, вот смотрите, если по этой... Вот теперь давайте думать. Значит, у нас получается такая картина, что вот по этой оси у меня будет стоять выигрыш в скорости, то есть отношение скорости к гаорифму по основанию 2Q. Галифм по основанию 2Q, это у меня скорость, которая, когда имеет место, столько битов нужно, когда я передаю без потери. А когда я передаю с потерями, мне вот столько нужно битов. Ну, стало быть, вот это относительно, так сказать, по этой оси отложена скорость с потерями. И максимальное ее значение равно единице. А по этой оси мы будем откладывать вот эту самую меру, Меру вероятности искажений. Параметр дельта. Он тоже лежит от нуля до единицы. Ну вот уже более-менее эту кривую можно набросать. Она выглядит примерно так. Совершенно ясно, что если дельта равно нулю, то в качестве S-дельта я должен выбрать вообще все множество. И при этом я нахожусь вот в этой точке. Единица, ноль. То есть, когда я передаю без искажений, скорость у меня вот такая, единичная. Когда я передаю с безумными искажениями дельта порядка единицы, то множество S-дельта состоит вообще из одного элемента. То есть, я в S-дельта включил... Только одного или вообще никого, наверное, одного. А все остальные лежат вне него, я их не передаю. То есть у меня искажения очень большие, вот они единичные. А скорость равна нулю, потому что мощность S δ равна единице, и гогарифм S δ равен нулю. Ну, то есть, вот ожидается, что между параметром δ и скоростью будет какая-то кривая вот такая. Так вот, мы сейчас покажем то, что я еще анонсировал на первой лекции, что тут возникает удивительный эффект. Эффект состоит в следующем, что если у моего этого источника есть энтропия, то вот здесь стоит уровень скорости, равный энтропии. И вот тут мы поставим границу с величиной h плюс ε, ε бесконечно малый, а тут границу с величиной h минус ε. Так вот, оказывается, что с ростом n, при достаточно больших n, чем мы сейчас и станем заниматься, доказательством чего, кривая эта, она отнюдь не так себя ведет. Она ведет себя так. Она быстренько уходит в эту зону, тут как-то болтается, и потом уходит вот в эту зону. Ну, то есть, получается так, что стоит вам допустить возможность хоть каких-то искажений, как вы получаете возможность передавать со скоростью равной энтропии. Вот с такой чуть больше энтропии. Вот это превышение зависит от того, какую меру искажений вы допустите. Но с ростом n она, это мера искажений, стремится к нулю. То есть эта кривая с ростом n неограниченно приближается к энтропии. Потом она где-то пересекает энтропию. Если вы задумаете передавать со скоростью меньше энтропии, вот с этой h-eps, то значит... Вам это не удастся при сколь угодно больших N, при сколь угодно больших дельта. Вот вплоть до сюда. То есть дельта надо будет сделать порядка единицы, чтобы выйти на этот спад. Ну то есть появляется возможность уменьшить скорость, только тогда, когда искажение дельта совершенно невозможно. Вот такая картинка. Значит, вот мы с вами собираемся оценивать мощности. Значит, на самом деле, вот... Если конструктивно подходить к этой проблеме, то она решается так. Как нам выбрать это множество из дельта? Значит, у него должна быть большая вероятность полная, больше, чем единица минус дельта, и в нем должно быть как можно меньше элементов. То есть оптимальный вариант его выбора такой. Вы просматриваете это все множество и берете там элемент с наибольшей вероятностью, его включаете сюда. В том, что осталось, выбираете следующий элемент с наибольшей вероятностью и включаете его сюда. И так вы набираете элементы, пока сумма вероятностей набранных элементов не совпадет, не превысит вот этот порог. После этого вы заканчиваете построение множества S-дельта, оно будет содержать минимально возможное количество блоков элементов. Ну и стало быть, нам надо сейчас с вами доказать следующее, что... Существует, вот мы сначала будем доказывать верхнюю границу, то есть существует множество S δ, такое, что гагарифм по основанию 2 его мощности, мощности S δ поделить на n, больше или равен, чем вот это вот H, меньше или равен, чем H плюс эпсимума. Вот такая граница меня для начала будет волновать. Ну, то есть, я утверждаю, что при достаточно большой длине я смогу выбрать множество из дельта, такое, что для него вот эта скорость передачи будет отличаться от энтропии не более чем на эпсилон. Ну, то есть, гагарифм по основанию 2 из дельта будет меньше или равен 2 в степени n на h плюс эпсилон. Вот такая штука. Ну, стало быть, прекрасно, как это доказать? Дело в том, что для множества S-дельта я предложил алгоритм построения вполне конкретный, путем выбора элементов с наибольшими вероятностями. Но, к сожалению, я вряд ли смогу оценить его мощность в какой-нибудь, сколь-нибудь общей ситуации. А мне нужна именно оценка мощности. Так вот, На роль этого самого множества S δ мы возьмем t β, множество типичных блоков. Видите, для него у нас имеется оценка сверху для мощности. И мы скажем такие слова, что давайте не будем выбирать множество S δ оптимально, а выберем его как-нибудь. Например, включим в него все типичные блоки. Вероятность этого множества будет порядка единицы. То есть, если я в качестве множества S δ возьму множество типичных блоков tβ, то что я про него знаю? Что его полная вероятность P от tβ будет больше или равна, чем единица минус σ2 на nβ2. то есть будет стремиться к единице. Вот это я возьму равно единице минус дельта, а вот это возьму дельтой. Вот это условие у меня будет выполнено, то есть оно годится, но оно может быть не очень хорошее в том понедельник, что мощность у него не минимально возможна. Она может быть и поменьше сделана. Но во всяком случае эта мощность не превышает вот этой границы. То есть мощность Тбета... Ну, то есть у меня мощность S δ вполне возможно может быть и поменьше, чем мощность T β, а мощность T β меньше или равно 2 в степени n на h плюс β. Он по той границе. Ну и все. Значит, осталось поступить так. Зафиксирую некое δ, получу отсюда оценку для β, и стало быть... Вот смотрите тут, как получится. Зафиксирую некая дельта, безумно маленькая, и мне это даст оценку для бета, что это самое бета, это есть сигма квадрат на сигма просто на корень из n, делённое на сигма, делённое на корень из n и делённое на дельта. Вот так. Ну вот, то есть бета будет с ростом n стремиться к нулю. Ну и, так же быть, если я сюда это подставлю, то тут у меня получится, вот это бета я сюда подставлю, будет оценка типа 2 в степени n на h плюс, ну вот это, сигма на корень из n на дельта. И мне надо, чтобы это было меньше любого ε. Вон у меня какая оценка-то мне нужна. 2m в степени n плюс h плюс ε. Но я выберу он достаточно большим, так чтобы вот это отношение было меньше ε. Оно уже падает, как корень из n, поэтому его можно сделать сколь угодно маленьким. Поэтому вот эта граница выполняется. То есть оказывается, что если выбирать... для передачи только типичные блоки, то это, возможно, будет неоптимальный выбор в смысле мощности, но он даст достаточно малую мощность. То есть мощность Тбета удовлетворяет той границе, которая нам нужна. То есть мы доказали с вами теорему кодирования вот в эту сторону, что для передачи с искажениями дельта элементов вот этого множества мы можем иметь скорость меньше вот этой границы h плюс ε, но при сколь угодно малом δ. То есть при сколь угодно малом δ у нас эта кривая окажется ниже вот этой границы h плюс ε. То есть фактически мы передаем энтропию. Энтропия – это и есть мера скорости передачи. Ну а в обратную сторону это доказывается ничуть не сложнее. Значит, нам надо доказать следующее, что Если мы собрались передавать со скоростью меньше h плюс и минус, со скоростью меньше вот этой, то вероятность такого множества окажется не стремящейся к единице. То есть мы возьмем множество s δ, такое, что его количество элементов в нем, мощность s δ, ну вот меньше или равно 2 в степени n, на H минус ε, вот этой нижней границы. То есть вот у меня тут есть 2 в степени H минус ε, вот сюда попаду. И займусь я исследованием вероятности этого множества. Я покажу, что P от этой мощности, полная вероятность этого множества, стремится к нулю. Вот такая теорема. То есть выбрать... Для кодирования множества с малой мощностью, с мощностью меньше, чем h-ε, 2 в степени NH-ε, невозможно, потому что у вас мера неискажённых броков будет сильно отличаться от единицы. будете стремиться к нулю. То есть вы с искажениями будете передавать почти все. А вот это делается так. Значит, мы рассматриваем какое-нибудь множество t-бета и занимаемся вот такой вещью. Значит, мы пишем, что это самое множество s-дельта есть s-дельта пересечения с t-бета и прибавить объединение s-дельта пересечение с дополнением Т-бета. Ну, то есть, картинка вот такая. Вот у меня есть где-то тут множество Т-бета, а вот есть множество С-дельта. Ну, я, стало быть, рассмотрел пересечение С-дельта вот с этим. Это вот это множество. Вот оно. И пересечение С-дельта с не Т-бета. Не Т-бета – это вот это множество дополнения. Ну, то есть, это просто ничего. Я просто представил... Это множество раздражил его на то, вот в S-дельта входят какие-то элементы. Какие-то из них входят и в t-бета, а какие-то не входят в t-бета. Каждый элемент может входить в t-бета, либо входить в его дополнение. То есть это ничего не знаю. Это просто я разбил S-дельта на два непересекающиеся множества. Вот это и вот это. Только и всего. Ну а теперь я пишу, что вероятность множества S δ есть сумма и от этой вероятности, и от этой вероятности. P от S δ пересечения с T β плюс P от S δ пересечения с не Т β. И каждую из этих вероятностей мне предстоит оценивать. этой совсем просто дело в том что ее можно мне надо оценивать эту сверху уж тем чтобы показывать что п от с дельта мало маленькая но у меня мощность дополнения тб достаточно мала поэтому вот эта вероятность значит п от с дельта пересечении с нет и бета она меньше или равна заведомо Вероятности от не Тбета. Ну, я вот в этом множестве его расширю. Оно состоит из тех элементов, которые принадлежат не Тбета, и к тому же из дельта. А я туда добавил вообще все элементы, которые принадлежат не Тбета. А у меня есть оценка сверху на мощность Тбета. Мне Тбета, она, как известно, меньше или равна сигма квадрат на n на бета квадрата. Тбэта – это у нее типичное множество, у него вероятность порядка единицы, а мощность не Тбэта – она вот маленькая. Поэтому с этим у меня проблем нет, и при больших N вот эта срыгаемая стремится к нулю. Ну а здесь смотрите как. Я предположил, у меня есть оценка, у меня есть мощность, у меня есть... оценка на мощность S δ, я его оценю сверху. Мощностью S δ на мощность T β. Ну, то есть, напишу так, это S δ, это T β. Для S δ я предпоказал, что оно меньше, чем 2 в степени n h минус ε. А мощность T β, она не больше, чем 2 в степени n на... А вероятность, то есть немножко не так. Всё правильно, но я оценивать это буду так. В это множество. У меня входят элементы, которые принадлежат S-дельта и T-бета одновременно. Вероятности их не больше, чем вероятности T-бета. А сами они, ну вот. Ну, то есть эту вероятность надо оценивать сверху, как мощность S-дельта на вероятность... элемента из множества t-бета, конкретного элемента из множества t-бета. Вот сюда входят все элементы, которые сюда входят, их вот столько, мощность s-дельта. И у каждого вероятность не больше, чем вероятность элемента s-бета. А это мы знаем, какая вероятность. Она вот такая, 2 в степени минус n на h плюс бета. А s-дельта, я договорился, иметь... Меньше или равно 2 в степени n на h минус эпсидум. Ну вот, тут получается 2 в степени минус n на h плюс бета. Значит, получается так. 2 в степени nh и 2 в степени минус nh уходят. Остается 2 в степени n на бета минус эпсидум. На эпсидум. Вот тут у меня минус ε, а тут 2 в степени минус n на β минус ε. Вот так. Ну и стало быть, если у меня ε достаточно мало, то есть оно больше β, то там показатель положительный, и штука эта стремится тоже к нулю. Поэтому вот так получается доказательство обратной теоремы. Довольно... Такое оно трудное для понимания. Тут запомнить трудно. Значит, смотрите, идея состоит в следующем, что мы предполагаем, что мощность множества s дельта мара, то есть скорость передачи будет ниже энтропии существенно, на епсилон, и доказываем, что для любого такого множества его вероятность стремится к нулю, полная вероятность. Для этого мы его разбиваем на пересечении с типичным множеством и дополнением к типичному множеству. Дополнение к типичному множеству оцениваем его полной вероятностью. Она вот у нас маленькая и вот такая. А вот эту вероятность оцениваем хитро. Мы ее оцениваем так. Мы берем тут мощность множества S δ. умноженную на вероятность каждого из элементов из этого множества. Для вероятности каждого элемента мы имеем вот такую оценку, 2 минус n на h плюс β. Ну и, значит, если мощность s δ маленькая, то у нас справедлива вот эта оценка, и, значит, вероятность такого множества стремится к нулю. Так что вот получается такой удивительный факт. Значит, сухой остаток из этого такой, что какой бы у вас небогосточник, у него есть энтропия. Если вы намереваетесь кодировать длинные бруси... Вот, энтропия на символ. Для независимого источника это понятно, что такое. Для независимого, немножко посложнее, скоро мы на это посмотрим. Значит, у неё есть энтропия на символ. То есть, вот, каждый источник порождает символы, и эти символы имеют какую-то энтропию на символ. Значит... Для того, чтобы передавать его, вот, сжимать его, передавать в сжатом виде, значит, смысл этой энтропии такой. Вы можете это сжать до такой степени, чтобы каждый символ у вас переносился вот количеством битов, равным h, делить на логарифм по основанию 2 от q, если это q-ичный источник. h – это его энтропия на символ. И H делить на ГО. Если хотите, то энтропию можно определять по основанию КУ, тогда это просто КУИЧная энтропия. То есть, каждой буквой источник, каждый символ источника переносит вот такое количество битов, если это КУИЧный источник. И, в общем, получается так, что допустив хоть какие-нибудь искажения источника, вы получаете возможность передавать... со скоростью чуть больше этой. Кали превышение зависит от дельта. То есть эта кривая, которую я тут рисую, она все же ведет себя плавненько. И где-то она при каких-то искажениях выше h на ε, а где-то она ниже h на ε. Она пересекает эту линию. То есть если вы допускаете хоть малейшие искажения, то вы получаете возможность передавать данные со скоростью меньше, чем... Вот это. Вот так напишу. H плюс ε. Многогарифм по основанию 2Q. То есть фактически вы передаете энтропию, порождаемую этим источником. А если вы захотите передавать со скоростью меньше, чем эта величина, то у вас появятся искажения очень большие. Такие, что при любой конечном, то вероятность вот этого множества, P от S δ, но δ будет стремиться к... вероятность вот этого множества будет стремиться к нулю, то есть дельта будет стремиться к единице. То есть вот с другой стороны R зажато между вот этой границей. H минус ε, гарифм по основанию 2Q. Ну так, где ε, любое? Ну то есть получается так, что при достаточно больших мощностях, объемах кодируемых броков, вы, в общем, передаете... Такое количество битов, как будто бы каждый символ переносит энтропию битов. Вот энтропию делить на гарифм по основанию 2Q, это количество битов, которые нужно для кодирования одного символа. Ну вот в этом как бы смысл энтропии в смысле сжатия данных. То есть, если вы хотите этот источник передавать компактно, малым битовым потоком, то знайте, что у вас потребуется... Вот у него на символ есть какая-то энтропия С, на символ примерно столько битов вам потребуется для передачи одного бита. При этом вы можете передавать со сколь угодно малыми искажениями. Вот это теорема кодирования источника с искажениями. Практически её смысл, я бы сказал, никакой. Дело в том, что она просто обозначает какую-то предельную границу для степени сжатия данных. Дальше идут всякие, очень многочисленные алгоритмы сжатия данных с потерями. Это всякие JPEG, MPEG и так далее. Кодирование аудиофайлов, изображений и прочее. 100% они не достигают этой границы. То есть там эвристические схемы кодирования, сжатия данных, которые далеки от совершенства. То есть это как бы... путеводная звезда, что вот к этому надо стремиться. Дело в том, что когда вы работаете с какими-нибудь JPEG-ами, то очень трудно применить эту теорию по той причине, что JPEG- это, как правило, мегабит данных какой-нибудь, да, и как-то распределение вероятности этих блоков очень трудно что-то о нем внятное сказать. Ну, то есть тут вероятностные оценки, вероятностные категории становятся не очень адекватными. Более-менее ясно, что эти биты случайны, но вот как написать распределение вероятностей мегабитовых блоков, внятно, ну и оценить его энтропию, ну вот этим очень трудно заниматься, и смотреть потом, количество битов в выходном файле у вас совпадает с энтропией на символ или нет, или не совпадает. Ну то есть, чтобы получить какие-то оценки близости вот этих, ну, методов предельным оценкам надо потратить ужасно много сил в этом проблема поэтому скорее все это все вот эти соображения они просто иллюстрируют важность суть энтропию энтропия это В общем, когда работает какой-либо источник, он порождает энтропию. И вы передаете в сжатом виде именно вот эту энтропию. То есть, вам надо стремиться к тому, чтобы в сжатом виде энтропия на символ была максимальная. Тем самым вы достигнете сжатия данных. Ну вот, а дальше идут... Уже более понятные вещи – это символьное кодирование, так называемое, сжатие данных по методу преобразования символьных источников. Ну вот это второй подход, о котором я тоже уже когда-то говорил. Здесь не будет никаких потерь, мы будем пытаться восстанавливать сжатые данные однозначно, без потерь. То есть постановка та же самая. Есть блок «Подобный». Вот множество, в котором имеется m элементов, m точек, которые я намереваюсь вам передавать битовыми, ну вот, вам передавать какими-то битовыми строками, битовыми блоками, потоком битами, комбинациями нулей и единиц. Значит, если никаких мер не предпринимать, то у меня алфавит двоичный, то мне потребуется грифм по основанию 2 m битов для кодирования каждой точки. каждый из точек. Или если эти точки имеют длину, вот как у нас оно было, x1, xn, то есть n-броки, то вот столько битов на брок. Вот столько битов на брок. Ну и стало быть, хотим по-прежнему передавать побыстрее, поменьше битов на блок. Как это сделать? Ну вот первый подход мы только что использовали. Мы отказались от передачи всех блоков и занялись передачей только типичного множества блоков фактически. Ну то есть мы выбрали множество S-дельта, такое, что его вероятность почти единица, P от S-дельта порядка единицы, Ну и вот оказывается, что если это p от p от s дельта, все же поменьше единицы, то передавать можно со скоростью равной энтропии. Энтропии на символ, на бит, на символ. Вот так. А теперь мы будем эксплуатировать всем другую идею. Значит, смотрите, вот в этом множестве уже никаких t-бета нет, там есть просто блоки x, и у них у каждого своя вероятность p от x. Так вот идея состоит в том, чтобы... Теретина – кодирование с разными длинными битовых блоков. Тем блокам, которые высоко вероятны, у которых вероятности большие, надо присвоить короткие битовые комбинации. А тем символам, у которых вероятности маленькие, можно присвоить вероятности и подлиннее. И вот у нас с вами получается такая картина. У нас есть элементы этого множества, буду обозначать их х. Их много, им штук. У каждого есть какая-то вероятность Px. И я попытаюсь каждому присвоить какой-то битовый блок длины Lx. Это количество битов в этом блоке. То есть элементы x буду кодировать двоичным кодом, двоичным блоком какой-то там 0,1, 0,1, 0,1 длины Lx. L двоичных символов в этом блоке. У меня появится понятие средней длины L. Это сумма по всем x, Px на Lx. И вот я буду пытаться вот эту сумму минимизировать. То есть попытаться достичь минимума этой суммы. Ну, то есть научиться кодировать этот вот источник. Источник порождает вот такие символы. Символов М штук много. Этими символами могут быть, в частности, вот такие блоки. И, значит, каждому символу я хочу присвоить... его закодировать для сжатия, для передачи какой-то строку битов в количестве elix. И хочу это сделать так, чтобы средняя длина кодового слова, вот эта, была по возможности минимальна. Ну вот, забегая вперед, мы с вами скоро покажем, вот прямо немедленно почти, что имеют место границы для этой самой средней длины. Вот она, это самая средняя длина, которая есть сумма по всем хам, по х и по х. Она, видите ли, больше или равна энтропии, и меньше или равна энтропии плюс единица. Где энтропия? Это энтропия вот этого источника. Вот так. Вот еще один смысл этой самой энтропии. Оказывается, сжимать данные можно... до границ, связанных с энтропией. То есть нельзя передавать меньше, чем энтропию, нельзя передавать броки никаким кодированием, вот таким, когда каждому элементу алфавита ставится в соответствие кодовый блок какой-то из нулей единиц, нельзя сделать среднюю длину меньше энтропии вот этой случайной величины. Но можно сделать среднюю длину меньше, чем эта энтропия плюс единица. Ну вот довольно хорошие границы. И самое главное, что тут работают уже конструктивные алгоритмы. Ну вот, начинается вся эта наука с некого неравенства, которое называется неравенство... Так, кода Хаффмана. А неравенство, неравенство... Вылетели из гагалы, как же оно называется-то? Я не сщебышевая. Неравенство крафта. Крафта на кволите. Итак, крафт. Значит, смотрите, картинка такая. Эта граница, она вообще очень общая. Она доказывается в рамках теории языков. Значит, смотрите, вот обыгрывается это все так. Есть у меня некий алфавит, а состоящий из букв букв. там ABC и так далее какие-то буквы ну допустим английский алфавит вот значит из этих букв я составлю слова словарь ну вот как слова мюллера вот ABC это будет слово там ABB это тоже слово ну то есть всякая конечная последовательность букв есть слово у слова есть длина L Ну и есть у меня объем словаря, это Q. Значит так, алфавит словарь. Высказывание. Высказывание это когда несколько слов вот так вот приставлено одно к другому. Слово 1, слово 2, слово 3. Ну и стало быть в результате мы получаем то, что называется язык, language. Мы из слов можем строить высказывание, предложение. Язык называется однозначно декодируемым, если всякое высказывание допускает только одно разбиение на слова. Вот смотрите, когда вы получили высказывание, это длинный поток символов алфавита. Вот его предстоит интерпретировать, это высказывание. Для этого вы должны глянуть словарь и разобраться, сколько существует вариантов составить именно вот это высказывание из элементов этого словаря. Таких вариантов может быть несколько. Например, у вас вот это высказывание составлено вот из таких слов, а еще оно составлено вот из таких слов, а еще оно составлено вот из таких двух слов. Тогда это не язык с однозначным декодированием. Язык с однозначным декодированием таков, что он допускает разбиение любого высказывания на слова лишь единственным образом. Ну, то есть, он не допускает двузначной интерпретации. Ну, то есть, в таком языке пробегы между словами не обязательно. Вы всё равно прочтёте этот текст, разобьёте его на слова, даже если слова не разделены пробегами. Вот это понятие языка с однозначным декодированием. И есть неравенство Крафта, которое утверждает, что если язык с однозначным декодированием, то длиннокодовых сров, LGT, которые в нем содержатся, удликают неравенство Крафта. Вот оно такое. Сумма в степени минус LGT, процент G, от 0,5 до N, меньше. или равно 1. Крафт и накволити. Для любого языка с однозначным декодированием длиннокодовых сров не могут быть любыми, а должны отзывать вот этой границы. Это в теории языков фундаментальная граница. Крафт. Граница крафта. Крафт и накволити. Ну вот доказательство, оно довольно противное, но, значит, тем не менее. Значит, делается так. берется эта сумма и возводится в ужасно большую степень. Значит, сумма по всем словам Q на Q в степени минус Lg. А вот поступим так. Рассмотрим эту сумму по всем j Q в степени минус Lg и возведем ее в какую-нибудь ужасную степень m. Ну, то есть перемножим m таких сумм. В результате у меня получится m-кратная сумма. Вот. Первая сумма будет по индексу j1, вторая сумма будет по индексу j2 и так далее. Много сумм. Последняя будет по индексу jm. И сюда будет входить q в степени минус l1, q в степени минус l2. Ну вот, неправда, j1, j2, И так далее. Q в степени минус LJMT. Вот так. Дальше я все это объединю и получу, что это есть Q в степени минус LJ1 плюс LJ2 плюс LJMT. Ну, то есть, вот так. Ну, и, так же быть, надо посмотреть на эту сумму, что она собой представляет. Значит, смотрите, здесь у меня каждое срагаемое в этой сумме, оно представляет некую конкатенацию слов из кодового словаря. Вот в данном случае здесь вот такое вот последовательное соединение, я называю конкатенацией. Значит, это слово длины LJ1, приставлено к слову J2, приставлено к слову LJMT. Вот такая конкретинация. Всего получилось слово длиной равной сумме вот этих длин. Давайте я вот эту сумму разобью на сумму, на суммирование не по множественному индексу, а суммирование по полной длине вот этого слова. Это будет так выглядеть. Это будет сумма по всем L, от L минимума, что L-максимум. Что такое L-миниум, что L-максимум? Q в степени… вот A от L, некое число, на Q в степени минус L. Значит, смотрите, я взял конкретную вот эту длину, чтобы сумма L-жита равнялась L. Значит, это L заключено у меня от какого-то минимального значения, это когда тут самые короткие слова все стоят. Вот. Вреди, говоря, есть короткие слова и есть длинные слова. Вот сумма самых коротких слов и даёт вот это самое l минимум. То есть сумма этих l житых, она не может быть слишком маленькой, она больше или равна l минимум. И меньше или равна l максимум. Это сумма самых длинных слов. А от l – это количество слов, у которых количество вариантов, которые, если вот так приставить, то длина срова будет этого равна L, а вот вариантов содержится A в L-той степени. Ну вот количество таких сров, оно не превышает Q в степени минус L. Значит, вот мы оценим вот это количество. Значит, сколько у меня сров можно... Вот есть у меня... высказывание длины L конкретной, сколькими способами его можно разбить на отдельные слова. Количество этих способов, ну, вот оно не превышает Q в степени, это длина L, значит, оно не превышает Q в степени L. Меньше или равно, ну, потому что всего у меня вот на длине L имеется Q в L-той степени слов. И каждая такая Слово может быть разбито на подслова единственным образом. В силу однозначности декодирования. Поэтому это меньше или равно сумма по L равная L минимум до L максимум Q в степени L на Q в степени минус L. А это есть единица. Ну, то есть, это сумма по L максимум минус L мин. Вот такая. Это не важно, какая она, а важно, что она конечна. Смотрите, если бы вот эта сумма была больше единицы, вот она была бы, вот сумма Q в степени минус N была какой-нибудь гамма и больше единицы, то гамма в м-той степени стремилась бы к бесконечности. А это не так, оно в бесконечной стремится, оно стремится где-то заключено между L минимумом и L максимумом. Поэтому делается отсюда вывод, что сумма Q в степени минус L житая, это гамма, и она меньше или равна единице. Вот это и есть.