ПРАКТИКУМ

Целью практикума является выработка практических навыков расчета энтропии и количества измерительной информации, определения энтропийных коэффициентов, числа различимых градаций при применении средств измерений. В настоящем учебном пособии приведены четыре задания, выбранные из различных вариантов задач, прошедших апробацию в учебном процессе. Варианты задания даны из условия, что у каждого слушателя есть шифр, состоящий не менее чем из двух цифр.

Задание 1.

Получение любой информации, в том числе и измерительной, теория информации трактует как устранение некоторой части неопределенности. В случае измерения по шкалам порядка весь диапазон возможных значений измеряемой величины разбивается реперными точками на ряд интервалов. Неопределенность до измерения характеризуется тем, что состояние системы не определено, т.е. неизвестно в каком из интервалов лежит значение измеряемой величины. Результатом измерения является указание того, что измеряемая величина лежит в данном интервале, что означает сужение области неопределенности.

Таким образом, с точки зрения теории информации результат измерения заключается в выборе конкретного интервала из целого ряда возможных интервалов. В шкале Бофорта для определения силы ветра и минералогической шкале твердости каждому интервалу присвоен балл.

В предположении, что вероятности попадания измеряемой величины в любой из интервалов равны между собой, определить:

1. число интервалов шкалы n_o (число состояний системы);
2. число интервалов n (число состояний системы) после измерений;
3. Априорную (безусловную) энтропию $\boldsymbol{H_o}$;
4. Апостериорную (условную) энтропию $\boldsymbol{H}$;
5. Количество измерительной информации $\boldsymbol{I}$.

<u>Указание</u>. Слушатель выбирает шкалу по последней цифре шифра из таблицы 1. Нечетной цифре шифра соответствует минералогическая шкала твердости (таблица 2), четной цифре соответствует шкала Бофорта (таблица 3). Вариант интервала неопределенности после проведения измерений указан в баллах соответствующей шкалы, приведен в таблице 1 и определяется по предпоследней цифре шифра слушателя. При решении задачи используются только первый, второй и третий столбцы таблицы 3.

Таблица 1

Варианты, соответствующие шкале Бофорта и минералогической шкале твердости

Данные	Шкала Бофорта (четная последняя цифра шифра студента)									
	Варианты (предпоследняя цифра шифра студента)									
Результат измерения скорости ветра в баллах	0	1	2	3	4	5	6	7	8	9
	от 2 до 3	от 3 до 5	7	10	от 0 до 1	от 1 до 3	от 8 до 9	12	от 6 до 7	от 2 до 4
Данные	Минералогическая шкала твердости (нечетная последняя цифра шифра студента)									
	Варианты (предпоследняя цифра шифра студента)									
Результат измерения твердости в баллах	0	1	2	3	4	5	6	7	8	9
	от 1 до 2	от 4 до 5	10	7	от 7 до 8	от 0 до 3	от 3 до 6	9	от 2 до 3	5

Порядок расчета

Область неопределенности до измерения простирается на все интервалы шкалы, поэтому n_o соответствует числу интервалов, на которые разбита шкала. Число интервалов n определяется согласно таблице 3. При условии, что вероятности попадания измеряемой величины в любой из интервалов равны между собой, неопределенность ситуации до и после измерений характеризуется энтропией, равной логарифму числа интервалов (состояний системы).

$$\boldsymbol{H_o} = \log_2 n_o, \qquad \boldsymbol{H} = \log_2 n. \tag{1}$$

Таблица 2

Минералогическая шкала твердости

Баллы	Определение по твердости
1	2
0	Меньше твердости талька
1	Равна твердости талька или больше ее, но меньше твердости гипса
2	Равна твердости гипса или больше ее, но меньше твердости известкового шпата
3	Равна твердости известкового шпата или больше ее, но меньше твердости плавикового шпата
4	Равна твердости плавикового шпата или больше ее, но меньше твердости апатита
5	Равна твердости апатита или больше ее, но меньше твердости полевого шпата
6	Равна твердости полевого шпата или больше ее, но меньше твердости кварца
7	Равна твердости кварца или больше ее, но меньше твердости топаза
8	Равна твердости топаза или больше ее, но меньше твердости корунда
9	Равна твердости корунда или больше ее, но меньше твердости алмаза
10	Равна твердости алмаза или больше ее

Количество измерительной информации ***I*** рассчитывается по формуле

$$I = H_o - H \tag{2}$$

Для удобства расчетов в приложении 1 приведены логарифмы по основанию 2 для чисел от 1 до 100. В формулах (1,2) энтропия и количество измерительной информации измеряется в битах. Представить полученные результаты в натах и дитах, используя следующие приближенные соотношения между единицами энтропии и количества информации:

1дит = 2,3нат = 3,3бит
1нат = 1,45бит = 0,43дит
1бит = 0,69нат = 0,3бит

Таблица 3

Шкала Бофорта для определения силы ветра

Баллы	Название	Признаки (действие)	Скорость (от – до), м/с	Протяженность балла, м/с
1	2	3	4	5
0	Штиль	Дым идет вертикально	0-0,9	0,9
1	Тихий	Дым идет слегка наклонно	0,9-2,4	1,5
2	Легкий	Ощущается лицом, шелестят листья	2,4-4,4	2,0
3	Слабый	Развеваются флаги	4,4-6,7	2,3
4	Умеренный	Поднимает пыль	6,7-9,3	2,6
5	Свежий	Вызывает волны на воде	9,3-12,3	3,0
6	Сильный	Свистит в вантах, гудят провода	12,3-15,6	3,3
7	Крепкий	На волнах образуется пена	15,6-18,9	3,3
8	Очень крепкий	Трудно идти против ветра	18,9-22,6	3,7
9	Шторм	Срывает черепицу	22,6-26,4	3,8
10	Сильный шторм	Вырывает деревья с корнем	26,4-30,5	4,1
11	Жестокий шторм	Большие разрушения	30,5-34,8	4,3
12	Ураган	Опустошительное действие	34,8-39,2	4,4

Задание 2.

Измерительная информация по шкале порядка получается путем сравнения друг с другом двух размеров Q_i и Q_j одной физической величины (например, массы) с помощью компаратора (например, равноплечих весов). Результатом каждого отдельного сравнения (отсчета) является одно из трех выражений:

$$Q_i > Q_j\ ,\ \ Q_i = Q_j\ ,\ Q_i < Q_j$$

т.е. решение о том, больше (меньше) или равен один размер другому.
Для повышения точности результатов (уточнения силы неравенства) измерения проводят n раз (многократно). В каждой серии из n отсчетов каждый из трех возможных результатов измерения ($Q_i > Q_j$, $Q_i = Q_j$, $Q_i < Q_j$) встречается m_k раз, где k = 1, 2, 3 (возможный результат $Q_i > Q_j$ встречается m_1 раз, возможный результат $Q_i = Q_j$ встречается m_2 раз, возможный результат $Q_i < Q_j$ встречается m_3 раз).

Рассчитать:
a) частость (вероятность) каждого из возможных результатов;
b) априорную (безусловную) энтропию $\boldsymbol{H_o}$;
c) апостериорную (условную) энтропию $\boldsymbol{H_n}$;
d) количество измерительной информации $\boldsymbol{I_n}$ в каждой серии.

Построить по расчетным данным графики:
a) зависимости апостериорной энтропии $\boldsymbol{H_n}$ от числа измерений n $\boldsymbol{H_n}= f(n)$;
b) зависимости количества информации $\boldsymbol{I_n}$ от числа измерений n $\boldsymbol{I_n}= f(n)$.

<u>Указание</u>. Вычисления ведутся для четырех серий измерений. При четной предпоследней цифре шифра студента в первой серии измерений n = 10; во второй серии n = 20, в третьей серии n = 40 и в четвертой серии n = 80. При нечетной предпоследней цифре шифра студента в первой серии измерений n = 15; во второй серии n = 25, в третьей серии n = 50 и в четвертой серии n = 100. Студент выбирает значение m_k из таблицы 4 по последней цифре шифра.

Таблица 4

Варианты для задания 2

Число измерений n	Значения m_k	Варианты (последняя цифра шифра слушателя)										
		0	1	2	3	4	5	6	7	8	9	
10	m_1 m_2 m_3	8 1 1	1 8 1	1 1 8	7 1 2	1 7 2	1 2 7	6 2 2	2 6 2	2 2 6	5 3 2	четная предпоследняя цифра шифра слушателя
20	m_1 m_2 m_3	18 1 1	1 18 1	1 1 18	17 1 2	1 17 2	2 1 17	16 2 2	2 16 2	2 2 16	15 2 3	
40	m_1 m_2 m_3	38 1 1	1 38 1	1 1 38	37 2 2	1 37 2	1 2 37	36 1 3	1 36 3	2 2 36	34 3 3	
80	m_1 m_2 m_3	78 1 1	1 78 1	1 1 78	77 2 1	1 77 2	2 1 77	76 2 2	3 76 1	2 2 76	75 2 3	
15	m_1 m_2 m_3	9 3 3	3 9 3	3 3 9	10 2 3	2 10 3	2 3 10	11 2 2	2 11 2	2 2 11	12 1 2	нечетная предпоследняя цифра шифра слушателя
25	m_1 m_2 m_3	20 3 2	3 20 2	3 2 20	21 2 2	2 21 2	2 2 21	22 1 2	1 22 2	1 2 22	23 1 1	
50	m_1 m_2 m_3	46 2 2	2 46 2	2 2 46	47 1 2	1 47 2	1 2 47	48 1 1	1 48 1	1 1 48	45 2 3	
100	m_1 m_2 m_3	94 3 3	3 94 3	2 4 94	95 2 3	2 95 3	2 3 95	96 2 2	2 96 2	1 96 3	97 2 1	

Порядок расчета

Для расчета априорной энтропии используется формула

$$H_o = -\sum_{k=1}^{3} p_k \log_2 p_k \;;$$

считая, что все три результата измерения равновероятны, т.е. $P_1 = P_2 = P_3 = 1/3$.

Расчет апостериорной энтропии выполняют по формуле

$$H_n = -\sum_{k=1}^{3} p_{nk} \log_2 p_{nk} \;;$$

где $P_{n\,k} = m_k/n$ = частость (вероятность) каждого результата в серии из n отсчетов.

Количество измерительной информации $\boldsymbol{I_n}$, полученной в каждой серии, рассчитывается по формуле

$$I_n = H_o - H_n\,.$$

При расчете апостериорной энтропии для первой серии измерений (n = 10 или n = 15) воспользоваться таблицей логарифмов по основанию 2, приведенной в приложении 1. При расчете апостериорной энтропии для остальных серий измерений можно пользоваться таблицей приложения 1 или таблицей приложения 2, в которой приведены значения $p_i\log_2 p_i$ для значений p_i от 0,01 до 1.

Задание 3.

Измерительная информация по шкале отношений получается с помощью аналогового прибора. Согласно одному из постулатов метрологии при любом измерении обязательно использование априорной информации. Априорной информацией является:

1. Интервал [Q_1, Q_2] в пределах которого находится значение измеряемой величины;
2. Закон распределения вероятности (ЗРВ) значений измеряемой величины в этом интервале;
3. числовые характеристики неизвестного ЗРВ (например, среднее квадратическое отклонение σ_Q).

Если ЗРВ неизвестен, то его заменяют ситуационной моделью, в качестве которой чаще всего используют равномерный ЗРВ.

После измерения происходит уточнение значения измеряемой величины:

1.Меньше становится апостериорный интервал неопределенности $[Q_1, Q_2]$ и уточняются параметры ЗРВ;
2.Становится известным ЗРВ (при многократном измерении).

Рассмотренная информационная модель измерения графически представлена на рис.1 для четырех случаев:

1.Уменьшился интервал неопределенности значений измеряемой величины при известном равномерном ЗРВ;
2.Уменьшился интервал неопределенности значений измеряемой величины при известном треугольном ЗРВ;
3.Ситуационная модель (равномерный ЗРВ) после измерения заменена найденным из опыта нормальным ЗРВ;
4.Ситуационная модель (равномерный ЗРВ) после измерения заменена найденным из опыта экспоненциальным ЗРВ.

Рассчитать: априорную (безусловную) энтропию $\boldsymbol{H_o}$, апостериорную (условную) энтропию $\boldsymbol{H}$ и количество измерительной информации $\boldsymbol{I}$ для следующих комбинаций априорного и апостериорного законов распределения информации:

a) равномерный – равномерный;
b) треугольный – треугольный;
c) равномерный (ситуационная модель) – нормальный;
d) равномерный (ситуационная модель) – экспоненциальный.

Построить в выбранном масштабе априорные $p_o(Q)$ и апостериорные $p(Q)$ плотности распределения вероятности для всех четырех случаев аналогично тому, как это показано на рис.1.

Рассчитать энтропийный коэффициент K для всех четырех ЗРВ, используемых в задании.

Указание. Исходные данные для расчета слушатель выбирает из табл. 5 в соответствии со своим шифром. Значение среднего квадратического отклонения σ дано для нормального и экспоненциального апостериорных законов распределения.

Порядок расчета

В табл. 6 приведены все необходимые справочные данные для расчетов. Для расчета априорной $\boldsymbol{H_o}$ и апостериорной $\boldsymbol{H}$ энтропии воспользоваться формулами, приведенными в третьем столбце табл.6.
Количество измерительной информации рассчитывается по формуле

$$\boldsymbol{I = H_o - H}$$

Таблица 5

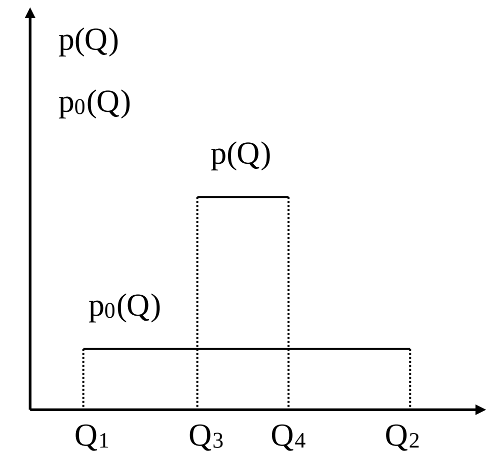
Варианты для задания 3

Параметры распределения	Варианты										Цифра шифра слушателя
	0	1	2	3	4	5	6	7	8	9	
Q_1	2	3	4	5	6	7	8	9	10	11	последняя
Q_2	14	15	16	17	18	19	20	21	22	23	последняя
Q_3	7	8	9	10	11	12	13	14	15	16	последняя
Q_4	9	10	11	12	13	14	15	16	17	18	последняя
σ	1	0,5	1,2	1	0,5	1,2	1	0,5	1,2	1,5	предпоследняя

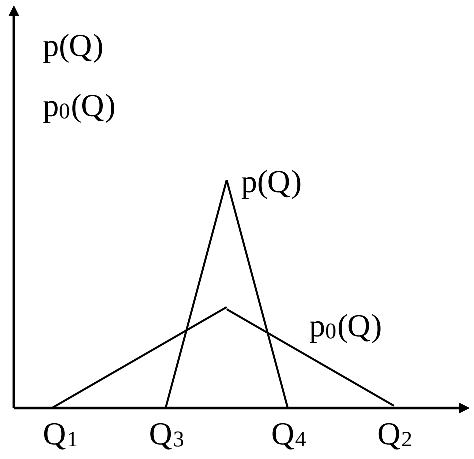
Для построения графиков плотности распределения вероятности $p_o(Q)$ и $p(Q)$ использовать формулы, приведенные во втором столбце табл.6. Для нормального и экспоненциального законов распределения графики строятся по точкам $\left| Q-\overline{Q} \right| = 0$; σ; 2σ; 3σ. Числовые значения σ необходимо выбрать из табл.5 (при использовании формул из табл.6 принять $Q = x; \overline{Q} = \overline{x}$). Величину, стоящую под знаком логарифма в формуле энтропии (табл.6, столбец 3) называют энтропийным интервалом неопределенности $\boldsymbol{d}$. Отношение половины энтропийного интервала неопределенности к среднеквадратическому отклонению (СКО) называют энтропийным коэффициентом ЗРВ.

$$K = \frac{d}{2\sigma}$$

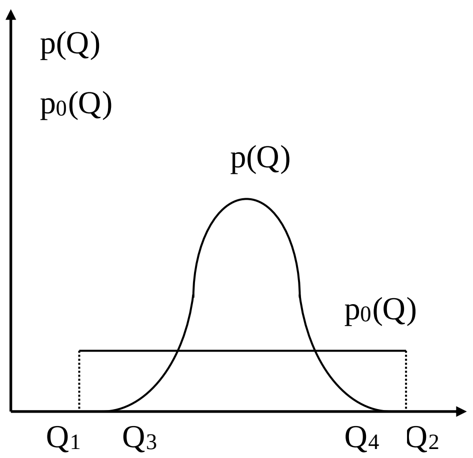
Для расчета энтропийного коэффициента $\boldsymbol{K}$ надо приравнять выражения, стоящие под знаком логарифма в формуле для энтропии H(x) (табл.6, столбец 3), значению $\boldsymbol{d}$ и подставить в формулу для энтропийного коэффициента.



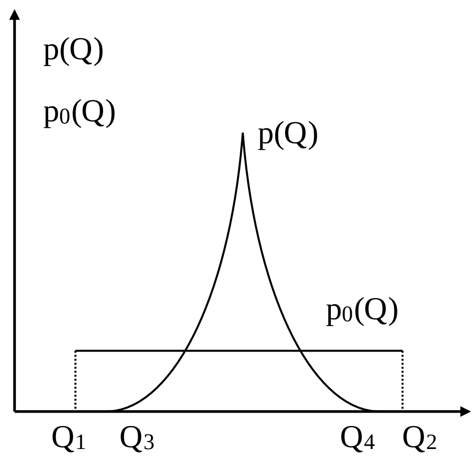
а) априорный ЗРВ – равномерный
апостериорный ЗРВ - равномерный



b) априорный ЗРВ – треугольный
апостериорный ЗРВ - треугольный



а) априорный ЗРВ – равномерный
апостериорный ЗРВ - нормальный

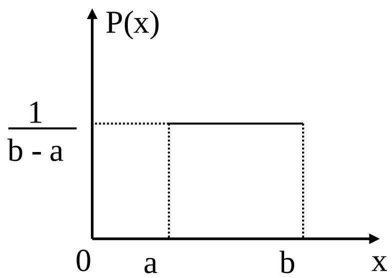
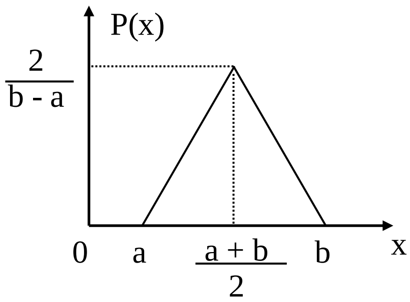


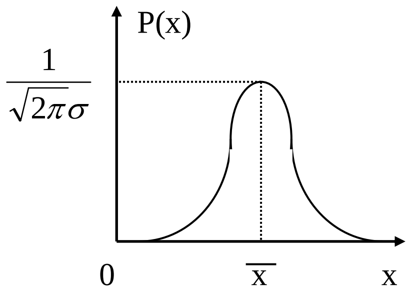
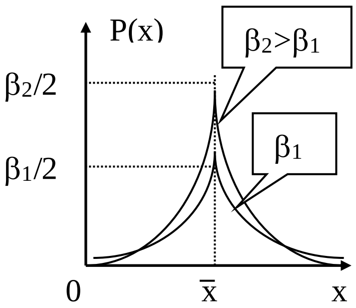
а) априорный ЗРВ – равномерный
апостериорный ЗРВ –
экспоненциальный

Рис.1 Априорная p_0(Q) и апостериорная p(Q) плотности распределения вероятности значений измеряемой

Таблица 6

Справочные данные для расчетов

№	Вид закона распределения	Числовые характеристики	
		Энтропия	Дисперсия
1	2	3	4
1	РАВНОМЕРНЫЙ  Плотность вероятности $p\ (x) = \dfrac{1}{b-a}$ a < x < b	$H\,(x) = \ln(b-a) =$ $= \ln(\sigma\ 2\sqrt{3})$	$\sigma^2 = \dfrac{(b-a)^2}{12}$ Стандартизованная длина половины доверительного интервала – 1,73σ
2	ТРЕУГОЛЬНЫЙ (Симпсона)  Плотность вероятности $p(x) = \begin{cases} \dfrac{4(x-a)}{(b-a)^2}, & a < x < \dfrac{a+b}{2} \\ \dfrac{4(b-x)}{(b-a)^2}, & \dfrac{a+b}{2} < x < b \end{cases}$	$H\,(x) = \ln\dfrac{(b-a)\sqrt{e}}{2} =$ $= \ln(\sigma\sqrt{6\,e})$	$\sigma^2 = \dfrac{(b-a)^2}{24}$ Стандартизованная длина половины доверительного интервала – 2,45σ

1	2	3	4
3	НОРМАЛЬНЫЙ (Гаусса)  Плотность вероятности $p(x)=\frac{1}{\sigma\sqrt{2\pi}}\,e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$	$H(x)=\ln(\sigma\sqrt{2\pi\, e})$	Стандартизированная длина половины доверительного интервала - 3σ
	ЭКСПОНЕНЦИАЛЬНЫЙ ДВУХСТОРОННИЙ (Лапласса)  Плотность вероятности $p(x)=\frac{\beta}{2}e^{-\beta\lvert x-\bar{x}\rvert},\ \ \beta=\frac{\sqrt{2}}{\sigma}$	$H(x)=\ln\frac{2e}{\beta}=$ $=\ln(\sigma\, e\sqrt{2})$	$\sigma^2=\frac{2}{\beta^2}$ При доверительной вероятности 0,925 длина половины доверительного интервала - 4σ

Приложение 1

Таблица двоичных логарифмов целых чисел от 1 до 100

x	$\log_2 x$	x	$\log_2 x$	x	$\log_2 x$
1	0,00000	35	5,12928	68	6,08746
2	1,00000	36	5,16993	69	6,10852
3	1,58496	37	5,20945	70	6,12928
4	2,00000	38	5,24793	71	6,14975
5	2,32193	39	5,28540	72	6,16992
6	2,58496	40	5,32193	73	6,18982
7	2,80735	41	5,35755	74	6,20945
8	3,00000	42	5,39232	75	6,22882
9	3,16993	43	5,42626	76	6,24793
10	3,32193	44	5,45943	77	6,26679
11	3,45943	45	5,49185	78	6,28540
12	3,58496	46	5,52356	79	6,30378
13	3,70044	47	5,55459	80	6,32193
14	3,80735	48	5,58496	81	6,33985
15	3,90689	49	5,61471	82	6,35755
16	4,00000	50	5,64386	83	6,37504
17	4,08746	51	5,67242	84	6,39232
18	4,16993	52	5,70044	85	6,40939
19	4,24793	53	5,72792	86	6,42626
20	4,32193	54	5,75489	87	6,44294
21	4,39232	55	5,78136	88	6,45943
22	4,45943	56	5,80735	89	6.47573
23	4,52356	57	5,83289	90	6,49185
24	4,58496	58	5,85798	91	6,50779
25	4,64386	59	5,88264	92	6,52356
26	4,70044	60	5,90689	93	6,53916
27	4,75489	61	5,93074	94	6,55459
28	4,80735	62	5,95420	95	6,56986
29	4,85798	63	5,97728	96	6,58496
30	4,90689	64	6,00000	97	6,59991
31	4,95420	65	6,02237	98	6,61471
32	5,00000	66	6,04439	99	6,62936
33	5,04439	67	6,06609	100	6,64386
34	5,08746				

Логарифмы, основные формулы

$$a^{\log_2 N} = N$$

$$\log_a a = 1, \quad \log_a 1 = 0$$

$$\log_a (MN) = \log_a M + \log_a N$$

$$\log_a \frac{M}{N} = \log_a M - \log_a N$$

$$\log_a N^m = m \log_a N$$

$$\log_a \sqrt[m]{N} = \frac{1}{m} \log_a N$$

обозначения: $\log_{10} = \lg N, \quad \log_e N = \ln N$

соотношения:

$$\log_b a = \frac{1}{\log_a b}$$

$$\log_a N = \frac{\log_b N}{\log_b a}$$

(число $\log_b a$ в последней формуле называется модулем перехода от системы логарифмов с основанием b к системе с основанием a)

Приложение 2

Таблица значений функции $-p\ log_2 p$

p	$-p\ log_2 p$	p	$-p\ log_2 p$	p	$-p\ log_2 p$	p	$-p\ log_2 p$
0	0						
0,01	0,0664	0,26	0,5053	0,51	0,4954	0,76	0,3009
0,02	0,1128	0,27	0,5100	0,52	0,4906	0,77	0,2903
0,03	0,1518	0,28	0,5142	0,53	0,4854	0,78	0,2796
0,04	0,1858	0,29	0,5179	0,54	0,4800	0,79	0,2687
0,05	0,2161	0,30	0,5211	0,55	0,4744	0,80	0,2575
0,06	0,2435	0,31	0,5238	0,56	0,4685	0,81	0,2462
0,07	0,2686	0,32	0,5260	0,57	0,4623	0,82	0,2348
0,08	0,2915	0,33	0,5278	0,58	0,4558	0,83	0,2231
0,09	0,3126	0,34	0,5292	0,59	0,4491	0,84	0,2112
0,10	0,3322	0,35	0,5301	0,60	0,4422	0,85	0,1992
0,11	0,3503	0,36	0,5306	0,61	0,4350	0,86	0,1871
0,12	0,3671	0,37	0,5307	0,62	0,4276	0,87	0,1748
0,13	0,3826	0,38	0,5305	0,63	0,4199	0,88	0,1623
0,14	0,3971	0,39	0,5298	0,64	0,4121	0,89	0,1496
0,15	0,4105	0,40	0,5288	0,65	0,4040	0,90	0,1398
0,16	0,4230	0,41	0,5274	0,66	0,3957	0,91	0,1238
0,17	0,4346	0,42	0,5256	0,67	0,3871	0,92	0,1107
0,18	0,4453	0,43	0,5236	0,68	0,3784	0,93	0,0974
0,19	0,4552	0,44	0,5210	0,69	0,3694	0,94	0,0839
0,20	0,4244	0,45	0,5184	0,70	0,3602	0,95	0,0703
0,21	0,4728	0,46	0,5153	0,71	0,3508	0,96	0,0565
0,22	0,4806	0,47	0,5120	0,72	0,3412	0,97	0,0426
0,23	0,4877	0,48	0,5083	0,73	0,3314	0,98	0,0286
0,24	0,4941	0,49	0,5043	0,74	0,3215	0,99	0,0144
0,25	0,5000	0,50	0,5000	0,75	0,3113	1,00	0