Задачи для семинаров по математической статистике (ФФ, 6-й семестр)

§1. Распределения случайных векторов

1.1. Найти вероятность p и одномерные таблицы распределения случайных величин X и Y. Найти $\mathbf{P}\{X = 0,\ Y > 1\}$, $F_{(X,Y)}(2,\ 2)$:

$X\backslash Y$	1	2	3
0	0,1	p	0
1	0	0	0,02
2	0,03	0	0

1.2. Найти вероятности p и q, если известно, что $\mathbf{P}\{X = -1\} = 0,3$. Найти таблицу распределения случайной величины Y и вероятность $\mathbf{P}\{X > 1,\ Y < 1\}$:

$X\backslash Y$	0	0,5	2
-1	0,1	p	0,05
3	0,3	0,4	q

1.3. Найти одномерные таблицы распределения случайных величин X и Y. Найти условное распределение X при условии $Y = -1$; при условии $Y = 2$. Найти условное распределение Y при условии $X = -2$; при условии $X = 1$.

$X\backslash Y$	-1	0	2
-2	0	0,5	0,1
1	0,1	0	0
3	0,3	0	0

1.4. Найти одномерные таблицы распределения случайных величин X и Y. Найти условное распределение X при условии $Y = 1$; при условии $Y > 1$. Найти условное распределение Y при условии $X = 0$.

$X\backslash Y$	-2	1	2	3	5
0	0,1	0,1	0,1	0,1	0,1
1	0,1	0,1	0,1	0,1	0,1

1.5. Найти константу A такую, чтобы функция $f(t_1, t_2) = At_1t_2\exp(-t_1^2 - t_2^2)$ при $t_1,\ t_2 \geq 0$, и равная нулю для всех остальных значений вектора (t_1, t_2), являлась двумерной плотностью распределения. Найти двумерную функцию распределения. Найти $F(0,0), F(0,-1), F(3,-2)$.

1.6. Плотность распределения $f(t_1, t_2)$ двумерного случайного вектора $(X_1,\ X_2)$ равна $\frac{1}{2}t_1t_2$ в треугольнике $\{0 \le t_1 \le 1,$ $0 \le t_2 \le 4t_1\}$, и нулю вне этого треугольника. Найти двумерную функцию распределения. Найти одномерные функции и плотности распределения. Найти $\mathbf{P}\{X_2 > 2\}, \mathbf{P}\{X_1 + X_2 < 1\}$.

1.7. Точку бросают наудачу в круг радиуса R с центром в начале координат. Пусть $(X,\ Y)$ — декартовы координаты точки. Найти плотность двумерного распределения, функции и плотности одномерных распределений. Найти $\mathbf{P}\{X < 0, Y < X\}$.

1.8*. Точку бросают наудачу в шар радиуса R с центром в начале координат. Пусть $(X,\ Y,\ Z)$ — декартовы координаты точки. Найти плотность трехмерного распределения, функции и плотности одномерных распределений.

1.9*. Пусть $(X_1,\ X_2,\ X_3)$ — координаты точки, брошенной наудачу в тетраэдр $\{t_1 \ge 0,\ \ t_2 \ge 0,\ \ t_3 \ge 0,\ \ t_1 + t_2 + t_3 \le 2\}$. Найти плотность двумерного распределения $(X_1,\ X_2)$.

§2. Преобразования случайных векторов

2.1. Найти таблицы распределения случайных векторов $(X+Y,\ X-Y)$, $(\max(X,\ Y),\ \min(X,\ Y))$. Найти одномерные таблицы распределения случайных величин $X+Y$, $X-Y$, $\max(X,\ Y)$, $\min(X,\ Y)$, XY, Y/X.

$X\backslash Y$	-1	0	2
-2	0	0,5	0,1
1	0,1	0	0
3	0,3	0	0

2.2. Найти таблицы распределения случайных векторов $(X+Y,\ X-Y)$, $(\max(X,\ Y),\ \min(X,\ Y))$. Найти одномерные таблицы распределения случайных величин $X+Y$, $X-Y$, $\max(X,\ Y)$, $\min(X,\ Y)$, XY, Y/X.

$X\backslash Y$	-2	1	2	3	5
0	0,1	0,1	0,1	0,1	0,1
1	0,1	0,1	0,1	0,1	0,1

2.3. Случайные величины X_1, X_2, X_3 независимы, принимают значения 0 и 1 с равными вероятностями. Найти таблицы распределения случайных величин $X_1+X_2+X_3$, $X_1+X_2-X_3$, $2X_1-X_2-X_3$, $\max\{X_1,\ X_2, X_3\}$, $\min\{X_1,\ X_2, X_3\}$.

2.4. Случайные величины X_1, X_2 независимы, принимают значения -1, 0 и 1 с равными вероятностями. Найти таблицы распределения случайных величин X_1+X_2, X_1-X_2, $2X_1-X_2$, $\max\{X_1,\ X_2\}$, $X_1X_2^2$.

2.5. Случайная величина X имеет абсолютно непрерывное распределение с плотностью распределения f. Найти плотности распределения случайных величин $X+1$, $2-3X$, X^3, $\sqrt{|X|}$, $\sqrt{4+X^2}$.

2.6. Случайная величина X имеет абсолютно непрерывное распределение с плотностью распределения f. Найти плотности распределения случайных величин e^X, $1/X$, X^2, $\ln|X|$.

2.7. Найти плотность распределения суммы двух независимых случайных величин, имеющих показательное распределение с параметром α.

2.8. Найти плотность распределения суммы двух независимых случайных величин, имеющих равномерное распределение на отрезке $[0,\ 1]$.

2.9. Найти плотность распределения суммы двух независимых случайных величин, имеющих стандартное нормальное распределение.

2.10. Случайные величины X_1, X_2 независимы и имеют абсолютно непрерывные распределения. Выразить плотность распределения их произведения через их плотности распределения.

2.11. Случайные величины X_1, X_2 независимы и имеют абсолютно непрерывные распределения. Выразить плотность распределения их разности через их плотности распределения.

2.12. Случайные величины X_1, X_2 независимы и имеют абсолютно непрерывные распределения. Выразить плотность распределения случайной величины $X_1^3 X_2^3$ через их плотности распределения.

2.13*. Случайные величины X_1, X_2 независимы и имеют равномерное распределение на отрезке $[0,\ 2]$. Найти плотность рапределения случайной величины $(X_1 - X_2)^{-1}$.

2.14*. Случайные величины X_1, X_2 независимы и имеют показательное распределение с параметром 1. Найти плотность рапределения случайной величины X_1/X_2.

2.15. Случайные величины $X_1, \ldots, X_n$ независимы и имеют одно и то же абсолютно непрерывное распределение с плотностью распределения f. Найти плотности распределения случайных величин $\max(X_1, \ldots, X_n)$ и $\min(X_1, \ldots, X_n)$.

§3. Моменты, ковариация, коэффициент корреляции

3.1. Найти математические ожидания и дисперсии случайных величин X и Y. Найти ковариацию и коэффициент корреляции случайных величин X и Y.

$X \backslash Y$	-1	0	2
-2	0	0,5	0,1
1	0,1	0	0
3	0,3	0	0

3.2. Найти математические ожидания и дисперсии случайных величин X и Y. Найти ковариацию и коэффициент корреляции случайных величин X и Y. Найти математические ожидания и дисперсии случайных величин $X + Y$ и $2X - 3Y - 2$.

$X \backslash Y$	-2	1	2	3
0	0,2	0	0	0,2
1	0,1	0,2	0,2	0,1

3.3. Наудачу выбирают цифру от 0 до 9. Найти коэффициент корреляции индикаторов событий «цифра делится без остатка на 3» и «цифра делится без остатка на 5».

3.4. Из 20 студентов 5 написали на «отлично» первую контрольную, 4 — вторую контрольную, и 3 — обе контрольные.

Для выбранного наудачу студента найти коэффициент корреляции индикаторов событий «первая контрольная написана на отличную оценку» и «вторая контрольная написана на отличную оценку».

3.5. Плотность распределения $f(t_1, t_2)$ двумерного случайного вектора $(X_1,\ X_2)$ равна $\frac{1}{2}t_1t_2$ в треугольнике $\{0 \le t_1 \le 1,$ $0 \le t_2 \le 4t_1\}$, и нулю вне этого треугольника. Найти математические ожидания и дисперсии компонент случайного вектора. Найти их ковариацию и коэффициент корреляции.

3.6. Точку бросают наудачу в круг радиуса R с центром в начале координат. Пусть $(X,\ Y)$ — декартовы координаты точки. Найти математические ожидания и дисперсии координат точки. Найти их ковариацию и коэффициент корреляции.

3.7. Случайные величины X и Y имеют дисперсии σ_X^2, σ_Y^2 и коэффициент корреляции ρ. Найти такую константу c, чтобы X и $Y - cX$ были некоррелированными.

3.8*. Найти коэффициент корреляции случайных величин X и X^2, если X имеет показательное распределение.

§4. Матрица ковариаций. Многомерное нормальное распределение

4.1. Найти вектор математического ожидания и матрицу ковариаций случайного вектора $(X,\ Y, X+Y)$:

$X \backslash Y$	-1	0	1
0	0,6	0	0
1	0	0	0,2
2	0	0,2	0

4.2. Найти вектор математического ожидания и матрицу ковариаций случайного вектора $(X,\ Y, XY)$. Найти математические ожидания и дисперсии случайных величин $X + Y$ и $XY - X - Y$.

$X \backslash Y$	0	2
0	0,1	0
1	0,7	0,2

4.3. Записать формулу плотности распределения 4-мерного стандартного нормального вектора. Найти плотность распределения суммы его компонент.

4.4. Найти вероятность того, что все компоненты стандартного нормального n-мерного вектора имеют один и тот же знак.

4.5. Пусть (X, Y, Z) — трехмерный стандартный нормальный вектор. С помощью таблицы нормального распределения найти приближенно $\mathbf{P}\{X > 1,96,\ Y < -2,33,\ Z < 0\}$.

4.6. Пусть (X, Y) — двумерный стандартный нормальный вектор. С помощью таблицы нормального распределения найти приближенно $\mathbf{P}\{X > -1,96,\ Y < 2,33\}$.

4.7. Пусть (X, Y) — двумерный стандартный нормальный вектор. Найти вероятность того, что $0 < X < Y$.

4.8. Пусть (X, Y) — двумерный стандартный нормальный вектор. Найти вероятность того, что $X < Y < X\sqrt{3}$.

4.9. Пусть

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix} + \begin{pmatrix} 3 & -4 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix},$$

вектор $(X_1,\ X_2)$ имеет стандартное нормальное распределение.

Записать плотность распределения вектора $(Y_1,\ Y_2)$.

Записать в виде двукратного интеграла вероятность $\mathbf{P}\{Y_1 > -2,\ Y_1 + Y_2 < 2\}$.

4.10*. Пусть

$$Y_1 = X_1 + X_2, \quad Y_2 = X_1 - 2X_2,$$

вектор $(X_1,\ X_2)$ имеет нормальное распределение с нулевым вектором математического ожидания, единичными дисперсиями компонент и коэффициентом корреляции компонент, равным $1/2$. Записать плотность распределения вектора $(Y_1,\ Y_2)$. Записать в виде двукратного интеграла вероятность того, что обе компоненты вектора $(Y_1,\ Y_2)$ положительны.

§5. Характеристические и производящие функции

5.1. Найти характеристическую и производящую функции случайной величины, принимающей значения 0, 1 и 2 с равными вероятностями.

5.2. Найти характеристическую и производящую функции бернуллиевского распределения.

5.3. Найти характеристическую и производящую функции биномиального распределения.

5.4. Найти характеристическую и производящую функции пуассоновского распределения.

5.5. Пусть X — неотрицательная целочисленная случайная величина. Выразить $\mathbf{E}X$ и $\mathbf{D}X$ через производные производящей функции.

5.6. Найти характеристическую функцию показательного распределения.

5.7. Найти характеристическую функцию гамма-распределения.

5.8. Найти характеристическую функцию квадрата стандартной нормальной случайной величины.

5.9. Найти характеристическую функцию суммы квадратов n независимых стандартных нормальных случайных величин.

5.10. По характеристическим функциям восстановить распределения: $\cos t$, $(1-4it)^{-1}$, $\exp(2it-2t^2)$.

5.11*. Найти представление характеристической функции двумерного нормального вектора через вектор математического ожидания и ковариационную матрицу.

5.12*. Найти плотность двумерного распределения, соответствующего характеристической функции $\exp\left(-2t_1^2+t_1t_2-2t_2^2\right)$.

§6. Предельные теоремы

6.1 Игрок в каждой игре (независимо от результатов других игр) выигрывает 80 рублей с вероятностью 0,1, проигрывает 20 рублей с вероятностью 0,9. Найти, к какой величине сходится средний выигрыш за n игр при $n \to \infty$.

6.2 Пусть $X_1,\ X_2,\ \ldots$ — случайные числа, то есть независимые случайные величины, имеющие равномерное распределение на отрезке от 0 до 1. Найти пределы п. н. следующих выражений при $n \to \infty$:

$$\text{а) } \frac{{X_1}^2 + \ldots + {X_n}^2}{n}; \qquad \text{в) } \frac{1}{n}\left(\frac{1}{1+X_1} + \ldots + \frac{1}{1+X_n}\right);$$

$$\text{б) } \frac{\sqrt{{X_1}^2 + \ldots + {X_n}^2}}{n}; \qquad \text{г) } \operatorname{arctg}\left(\frac{2}{n}\left(X_1 + \ldots + X_n\right)\right).$$

6.3. Случайные величины $X_1, X_2, \ldots$ независимы и одинаково распределены по закону Пуассона с параметром λ. К чему сходится с вероятностью единица последовательность

$$\frac{X_1^2 + \ldots + X_n^2}{n} - \left(\frac{X_1 + \ldots + X_n}{n}\right)^2 \quad ?$$

6.4. Случайные величины $X_1,\ X_2,\ \ldots$ независимы и равномерно распределены на отрезке $[0,\ a]$. Доказать, что $Y_n \to a$ с вероятностью единица при $n \to \infty$ для последовательности случайных величин $Y_n = \max(X_1,\ \ldots,\ X_n)$ (указание: использовать тот факт, что для сходимости монотонной последовательности к константе a с вероятностью единица достаточно сходимости функций распределения во всех точках, отличных от точки a).

6.5 Какова вероятность того, что в 100 партиях одинаковых по силе противников один из них выиграет более 70 раз? Ничьих нет.

6.6. Вероятность выхода из строя за время T одного конденсатора равна 0,05. Определить вероятность того, что за время T из 100 конденсаторов выйдут из строя: а) не менее 5 конденсаторов; б) менее 13 конденсаторов.

6.7. Студент получает на экзамене 5 с вероятностью 0,2, 4 с вероятностью 0,4, 3 с вероятностью 0,3 и 2 с вероятностью 0,1.

За время обучения он сдает 100 экзаменов. Найти пределы, в которых с вероятностью 0,95 лежит средний балл.

6.8. Урожайность куста картофеля задается следующим распределением:

Урожай в кг	0	1	1,5	2	2,5
Вероятность	0,1	0,2	0,2	0,3	0,2

На участке высажено 900 кустов. В каких пределах с вероятностью 0,95 будет находиться урожай? Какое наименьшее число кустов нужно посадить, чтобы с вероятностью не менее 0,975 урожай был не менее тонны?

6.9*. Игральная кость подбрасывается до тех пор, пока общая сумма очков не превысит 700. Оценить вероятность того, что для этого потребуется более 210 бросаний.

6.10*. Пусть $X_1, X_2, \ldots$ — независимые одинаково распределенные случайные величины, $\mathbf{E}X_1 = 0, \quad \mathbf{D}X_1 < \infty$. Известно, что

$$\mathbf{P}\left(\frac{X_1 + \ldots + X_n}{\sqrt{n}} \geq 1\right) \to \frac{1}{3}$$

при $n \to \infty$. Найти $\mathbf{D}X_1$.

6.11. Известно, что вероятность рождения мальчика приблизительно равна 0,515. Какова вероятность того, что среди 10 тыс. новорожденных окажется мальчиков не больше, чем девочек?

6.12. Для лица, дожившего до двадцатилетнего возраста, вероятность смерти на 21-м году жизни равна 0,006. Застрахована группа 10000 лиц 20-летнего возраста, причем каждый застрахованный внес 1200 рублей страховых взносов за год. В случае смерти застрахованного родственникам выплачивается 100000 рублей. Какова вероятность того, что:

а) к концу года страховое учреждение окажется в убытке;

б) его доход превысит 6000000 рублей?

Какой минимальный страховой взнос следует учредить, чтобы в тех же условиях с вероятностью 0,95 доход был не менее 4000000 рублей?

6.13. Суммируются 100 независимых одинаково распределенных векторов с нулевым вектором математического ожидания, равными единице дисперсиями компонент и коэффициентом корреляции компонент, равным -1/2. Записать в виде двойного интеграла приближенную вероятность того, что каждая из компонент суммы будет меньше 30.

6.14. На светофоре загорается красный, желтый или зеленый свет с равными вероятностями. Найти приближенно вероятность того, что при 90 наблюдениях светофора студент менее 20 раз заставал зеленый свет. Записать в виде двойного интеграла приближенную вероятность того, что при 90 наблюдениях светофора студент более 40 раз заставал красный свет и более 30 раз желтый.

§7. Выборка. Оценивание параметров

Выборка и вариационный ряд

7.1. По данной реализации выборки $\vec{x}$ =(0; 0; 1; 1; 0; 0; 0; 0; 0; 1):

а) построить график реализации эмпирической функции распределения;

б) вычислить реализации выборочного среднего и выборочной дисперсии.

7.2 По реализации выборки 1; 0; 1; 1; 0; 1; 0; 0; 0; 1 вычислить реализации выборочного среднего, выборочной дисперсии, выборочного среднеквадратического отклонения, несмещенной выборочной дисперсии, выборочных асимметрии и эксцесса.

7.3. Измерен рост (в см) студентов одной учебной группы. Результаты измерений дали выборку (171; 186; 164; 190; 158;

181; 176; 180; 174; 157; 176; 169; 164; 186).

а) Построить реализацию гистограммы.

б) Вычислить реализации выборочного среднего, выборочной дисперсии и выборочного стандартного отклонения S. На одном графике с гистограммой построить график плотности нормального закона с параметрами $\overline{X}$, S^2.

7.4. Пусть $\vec{X} \Subset \Phi_{a,\sigma^2}$. Вычислить $\mathbf{E}\overline{X}$, $\mathbf{D}\overline{X}$. Какое распределение имеет случайная величина $\overline{X}$?

7.5*. Пассажир маршрутного такси измерил 8 раз время ожидания такси и получил следующие результаты (в минутах): 8; 4; 5; 4; 2; 15; 1; 6. У него есть две гипотезы относительно графика движения такси: либо график движения соблюдается, и время ожидания имеет равномерное распределение на отрезке $[0;\ \theta]$, либо график движения не соблюдается, и время ожидания имеет показательное распределение с параметром λ.

а) Вычислить реализации оценок параметров θ и λ, использовав оценки $\widetilde{\theta}_2 = (n+1)X_{(n)}/n$ и $\widetilde{\lambda}_2 = \frac{n-1}{n\overline{X}}$.

б) Построить на одном графике реализацию эмпирической функции распределения и теоретические функции распределения равномерного и показательного законов, в которые вместо неизвестных параметров подставлены реализации их оценок.

в) Построить на одном графике реализацию гистограммы и теоретические плотности распределения равномерного и показательного законов, в которые вместо неизвестных параметров подставлены реализации их оценок.

г) На основании проведенного исследования сделать вывод о том, какая из гипотез выглядит более соответствующей экспериментальным данным.

7.6*. Дана выборка $\vec{X} \Subset \Pi_\lambda$, $\lambda > 0$ — неизвестный параметр. Проверить, что статистики

$$T_1 = \overline{X}, \quad T_2 = \frac{1}{n}\sum_{i=1}^{n} \mathbf{I}(X_i = k), \quad T_3 = \frac{X_1 + X_n}{2}$$

являются несмещенными оценками соответственно для

λ, $\frac{\lambda^k}{k!}e^{-\lambda}$ и λ. Являются ли эти оценки состоятельными?

7.7. По выборке $(X_1, \ldots, X_n)$ из бернуллиевского распределения B_p с неизвестным параметром $p \in (0;\ 1)$ построить оценки параметра p:

а) по первому моменту;

б) по второму моменту;

в) по произвольному k-му моменту.

Можно ли отдать предпочтение какой-либо из построенных оценок? Исследовать их состоятельность и несмещенность.

7.8. По выборке $(X_1, \ldots, X_n)$ из биномиального распределения $B_{m,p}$ построить оценки методом моментов:

а) параметра p по первому и по второму моменту при известном $m > 0$;

б) параметров p и m.

Исследовать состоятельность построенных оценок.

7.9. Используя метод моментов, построить бесконечную последовательность различных оценок параметра θ равномерного распределения на отрезке $[0;\ \theta]$. Будут ли полученные оценки состоятельными?

7.10. С помощью метода моментов построить оценку параметра $\theta > 0$, если распределение выборки имеет плотность:
а) $\theta t^{\theta-1}$ при $t \in [0;\ 1]$; б) $2t/\theta^2$ при $t \in [0;\ \theta]$.

Исследовать полученные оценки на состоятельность.

7.11. Дана выборка из распределения с плотностью

$$f_\theta(t) = \begin{cases} 3t^2\theta^{-3}, & t \in [0;\ \theta]; \\ 0, & t \notin [0;\ \theta]. \end{cases}$$

Найти оценку параметра $\theta > 0$ методом моментов, исследовать ее на несмещенность и состоятельность.

7.12. Методом моментов найти оценку параметра $\alpha > 0$ по выборке из показательного распределения с плотностью $f_\alpha(t) = \alpha e^{-\alpha t}$, $t > 0$. Будет ли оценка несмещенной и состоятельной?

7.13*. По выборке $(X_1, \ldots, X_n)$ методом моментов найти две различные оценки параметра $p \in (0, 1)$, если известно, что:

$P\{X_1 = 1\} = p/2;\ P\{X_1 = 2\} = p/2;\ P\{X_1 = 3\} = 1 - p.$

Будут ли полученные оценки несмещенными и состоятельными?

7.14*. При каких значениях параметра $\theta > 0$ распределения Парето с плотностью

$$f_\theta(t) = \begin{cases} \frac{\theta}{t^{\theta+1}}, & t \geq 1; \\ 0, & t < 1 \end{cases}$$

существует оценка параметра по первому моменту? Можно ли построить состоятельную оценку методом моментов в случае, когда оценки по первому моменту не существует?

7.15*. По выборке $(X_1, \ldots, X_n)$ из распределения Лапласа с плотностью $f_\lambda(t) = \frac{\lambda}{2} e^{-\lambda|t|}, \quad t \in \mathbf{R}$, построить оценку параметра $\lambda > 0$ методом моментов.

7.16. Пусть дана выборка из нормального распределения с параметрами α и σ^2. Используя метод моментов, построить оценки:

а) неизвестного математического ожидания α;

б) неизвестной дисперсии σ^2, если α известно;

в) неизвестной дисперсии σ^2, если α неизвестно.

Исследовать полученные оценки на несмещенность и состоятельность.

7.17*. Используя метод моментов, оценить параметр θ равномерного распределения на отрезке:

а) $[-\theta;\ \theta]$, $\theta > 0$; б) $[\theta;\ \theta + 1]$.

Исследовать полученные оценки на несмещенность и состоятельность.

§8. Оценки максимального правдоподобия

8.1. По выборке $(X_1, \ldots, X_n)$ из бернуллиевского распределения B_p с неизвестным параметром $p \in (0;\ 1)$ построить оценку параметра p методом максимального правдоподобия. (Указание: показать, что вероятность попадания в точку t для элементов выборки равна $f(t,\ p) = p^t(1-p)^{1-t}$, где t может принимать только два значения — 0 и 1). Исследовать состоятельность и несмещенность полученной оценки.

8.2. По выборке $(X_1, \ \ldots, \ X_n)$ из биномиального распределения $B_{m,p}$ построить оценку максимального правдоподобия параметра p при известном $m > 0$. Исследовать состоятельность и несмещенность оценки.

8.3. По выборке из показательного распределения E_α построить оценку максимального правдоподобия параметра $\alpha > 0$. Исследовать состоятельность оценки.

8.4. Построить оценку максимального правдоподобия по выборке из распределения Парето с плотностью

$$f_\theta(t) = \begin{cases} \frac{\theta}{t^{\theta+1}}, & t \geq 1; \\ 0, & t < 1. \end{cases}$$

Доказать состоятельность полученной оценки.

8.5*. С помощью метода максимального правдоподобия построить оценку параметра $\theta > 0$, если элементы выборки имеют плотность распределения:

а) $\theta t^{\theta-1}$ при $t \in [0; \ 1]$; б) $2t/\theta^2$ при $t \in [0; \ \theta]$.

Исследовать полученные оценки на состоятельность.

8.6*. Дана выборка из распределения с плотностью

$$f_\theta(t) = \begin{cases} 3t^2\theta^{-3}, & t \in [0; \ \theta]; \\ 0, & t \notin [0; \ \theta]. \end{cases}$$

Найти оценку параметра $\theta > 0$ методом максимального правдоподобия, исследовать ее на несмещенность и состоятельность.

8.7*. По выборке $(X_1, \ldots, X_n)$ методом максимального правдоподобия найти оценку параметра $p \in (0, 1)$, если известно, что $P\{X_1 = 1\} = p/2, \ P\{X_1 = 2\} = p/2, \ P\{X_1 = 3\} = 1 - p$.

Будет ли полученная оценка несмещенной и состоятельной?

8.8. Дана выборка из распределения с плотностью

$$f_\theta(t) = \begin{cases} e^{\theta-t}, & t \geq \theta; \\ 0, & t < \theta. \end{cases}$$

Найти оценку для θ:

а) методом моментов;

б) методом максимального правдоподобия.

Будут ли полученные оценки состоятельными? Вычислить смещения оценок и получить исправленные несмещенные оценки.

8.9*. По выборке $(X_1, \ldots, X_n)$ из распределения Лапласа с плотностью $f_\lambda(t) = \frac{\lambda}{2}e^{-\lambda|t|}, \quad t \in \mathbf{R}$, построить оценку параметра $\lambda > 0$ методом максимального правдоподобия.

8.10. Пусть дана выборка из нормального распределения с параметрами α и σ^2. Используя метод максимального правдоподобия, построить оценки:

а) неизвестного математического ожидания α;

б) неизвестной дисперсии σ^2, если α известно;

в) неизвестной дисперсии σ^2, если α неизвестно.

Исследовать полученные оценки на несмещенность и состоятельность.

8.11*. Используя метод максимального правдоподобия, оценить параметр θ равномерного распределения на отрезке:

а) $[-\theta;\ \theta]$, $\theta > 0$;

б) $[\theta;\ \theta + 1]$.

Исследовать полученные оценки на несмещенность и состоятельность.

§9. Сравнение оценок: среднеквадратический подход

9.1. Имеется выборка четного объема n из распределения с конечной ненулевой дисперсией. По этой выборке построены 2 оценки математического ожидания: среднее по всей выборке и среднее по первой половине выборки. Сравнить их в среднеквадратическом смысле.

9.2. Пусть $\vec{X}$ — выборка из распределения с математическим ожиданием θ и конечной ненулевой дисперсией σ_θ^2. Выяснить, каковы должны быть константы $C_1, \dots, C_n$, чтобы оценки вида $\widetilde{\theta} = C_1X_1 + C_2X_2 + \dots + C_nX_n$ были несмещенными. Показать, что оценка $\theta_1^* = \overline{X}$ является наилучшей в среднеквадратическом в этом классе оценок.

9.3. Для выборок из следующих распределений найти оценку максимального правдоподобия $\widehat{\theta}$, проверить ее несмещенность и вычислить $\mathbf{E}(\widehat{\theta} - \theta)^2$:

1) распределение Бернулли с параметром p;

2) биномиальное распределение с параметрами 2, p;

3) геометрическое распределение с параметром $1/\theta$, $\theta > 1$ (напомним, что $\mathbf{P}_\theta\{X = k\} = \frac{1}{\theta}\left(1 - \frac{1}{\theta}\right)^{k-1}$, $k \geq 1$, $\mathbf{E}X_1 = \theta$, $\mathbf{D}X_1 = \theta(\theta - 1)$);

4) показательное распределение с параметром $1/\theta$, $\theta > 0$;

5) нормальное распределение с параметрами a, 1;

6) нормальное распределение с параметрами 0, σ^2.

9.4. Дана выборка $\vec{X} \Subset \mathrm{U}_{[0,\theta]}$; $\theta > 0$ — неизвестный параметр. Сравнить, какая из оценок для параметра θ лучше в среднеквадратическом смысле: $\theta_1^* = 2\overline{X}$, $\widetilde{\theta} = \frac{n+1}{n}X_{(n)}$.

9.5. Для распределений из задачи 9.3 проверить условие регулярности, вычислить информацию Фишера и исследовать эффективность полученных в задаче 9.3 оценок максимального правдоподобия.

9.6*. Дана выборка из распределения с плотностью

$$f_\theta(t) = \begin{cases} e^{\theta - t}, & t \geq \theta; \\ 0, & t < \theta. \end{cases}$$

Найти оценки для θ методом моментов и методом максимального правдоподобия. Сравить найденные оценки в среднеквадратическом.

9.7*. По выборке $(X_1, \dots, X_n)$ из распределения Лапласа с плотностью $f_\lambda(t) = \frac{\lambda}{2}e^{-\lambda|t|}$, $t \in \mathbf{R}$, построить оценки параметра $\lambda > 0$ на основании второго момента и методом максимального правдоподобия. Сравнить эти оценки в среднеквадратическом смысле.

§10. Оценивание параметров в задачах линейной регрессии

10.1. Пусть $Y_i = x_i + \theta + \varepsilon_i$, $i = 1, \ldots, n$. Здесь $x_i,\ \theta \in \mathbf{R}$. Найти оценку для θ по методу наименьших квадратов. Найти оценку дисперсии регрессионных ошибок σ^2.

10.2. Пусть $Y_i = \theta x_i + \varepsilon_i$, $i = 1, \ldots, n$. Здесь $x_i,\ \theta \in \mathbf{R}$. Выяснить, для каких значений x_i выполнены предположения теоремы Гаусса—Маркова. Найти оценку для θ по методу наименьших квадратов. Найти оценку дисперсии регрессионных ошибок σ^2.

10.3. Концентрация лекарства $Y > 0$ в крови пациента обратно пропорциональна массе тела $x > 0$. Найти оценку коэффициента пропорциональности для следующих моделей:

1) $Y_i = \theta / x_i + \varepsilon_i$;

2) $\ln Y_i = \ln(\theta / x_i) + \varepsilon_i$;

$i = 1, \ldots, n$. Найти оценку параметра θ в каждой модели. Найти дисперсию оценки в первой модели и дисперсию логарифма оценки во второй модели.

10.4. Для регрессионной модели $Y_i = a + b x_i + \varepsilon_i$, $i = 1, \ldots, n$, найти оценки параметров $a,\ \ b$ по методу наименьших квадратов. Найти ковариационную матрицу оценок. Найти оценку дисперсии регрессионных ошибок σ^2.

10.5. По реализации двумерной выборки $x_1 = 1$, $Y_1 = 0$, $x_2 = 2$, $Y_2 = 2,5$, $x_3 = 3$, $Y_3 = 0,5$, найти реализации оценок параметров модели из задачи 11.4. Вычислить реализацию коэффициента детерминации.

10.6*. Для регрессионной модели $Y_i = a_1 \cos x_i + b_1 \sin x_i + \varepsilon_i$, $i = 1, \ldots, n$, найти оценки параметров $a_1,\ \ b_1$ по методу наименьших квадратов. Рассмотреть случай $x_i = \pi i/2$, $n = 4$. Найти ковариационную матрицу оценок.

§11. Интервальное оценивание

11.1. Пусть элементы выборки $\vec{X}$ имеют плотность распределения

$$f(t) = \frac{1}{\pi(1+(t-\theta)^2)}, \quad t \in \mathbf{R}.$$

Здесь θ — неизвестный параметр, $\theta \in \mathbf{R}$. Построить точный доверительный интервал для параметра θ по одному наблюдению ($n = 1$).

11.2. $\vec{X} \Subset B_p$, $0 < p < 1$. Построить асимптотический доверительный интервал для параметра p.

11.3. Дана выборка из геометрического распределения с параметром p, $0 < p < 1$. Построить асимптотический доверительный интервал для параметра p.

11.4. По выборке из распределения Пуассона с параметром $\lambda > 0$ построить асимптотический доверительный интервал для параметра λ.

11.5*. Дана выборка из распределения с плотностью $e^{-|t-a|}/2$, $a \in \mathbf{R}$. Построить асимптотический доверительный интервал для параметра a.

11.6*. Пусть $\vec{X} \Subset U_{[0;\,\theta]}$, где $\theta > 0$. С помощью статистик $\overline{X}$ и $\overline{X^2}$ построить асимптотические доверительные интервалы (соответственно (θ_1^-, θ_1^+) и (θ_2^-, θ_2^+)) уровня $1-\varepsilon$ и показать, что случайный интервал (θ_2^-, θ_2^+) асимптотически короче соответствующего (θ_1^-, θ_1^+).

11.7. Известно, что измерения величины a независимы, имеют нормальное распределение с математическим ожиданием a (то есть отсутствует систематическая погрешность) и стандартным отклонением 10 мм. Результаты 4 измерений дали среднее значение 512 мм. Найти доверительный интервал для параметра a уровня 0,95; уровня 0,998.

11.8. Известно, что измерения величины a независимы, имеют нормальное распределение с математическим ожиданием a

(то есть отсутствует систематическая погрешность) и стандартным отклонением σ. Результаты 100 измерений эталонной длины 1 м дали выборочное среднее 1,01 м и выборочный второй момент 1,04 м^2. Найти доверительный интервал для стандартного отклонения уровня 0,9; уровня 0,99.

11.9. По выборке объема 25 из нормального распределения подсчитаны выборочное среднее 2,1 и выборочный второй момент 4,42. Построить точные доверительные интервалы уровня 0,95 для параметров нормального распределения.

§12. Статистические гипотезы и критерии

12.1. Крупная партия товаров может содержать долю дефектных изделий. Поставщик полагает, что эта доля составляет 3%, а покупатель — 10%. Условия поставки: если при проверке 20 случайным образом отобранных товаров обнаружено не более одного дефектного, то партия принимается на условиях поставщика, в противном случае — на условиях покупателя. Требуется определить:
1) каковы статистические гипотезы, статистика критерия, область ее значений, критическая область;
2) какое распределение имеет статистика критерия, в чем состоят ошибки первого и второго рода и каковы их вероятности.

12.2. Имеется выборка объема 1 из нормального распределения $\Phi_{a,1}$. Проверяются простые гипотезы $H_0 : \ a = 0, \ H_1 : \ a = 1$. Используется следующий критерий (при заданной постоянной c):

$$H_0 \Leftrightarrow X_1 \leq c.$$

Вычислить, в зависимости от c, вероятности ошибок первого и второго рода.

12.3. Используя конструкции доверительного интервала, построить критерий точного уровня ε для проверки гипотезы $\mathrm{H} : \theta = 1$, если:

а) $\vec{X} \Subset \Phi_{\theta,1}$;
б) $\vec{X} \Subset \Phi_{1,\theta}$.

12.4. Построить критерий, обладающий нулевыми вероятностями ошибок, для проверки гипотез $\mathrm{H}_0 : \vec{X} \Subset \Phi_{0,1}$ против $\mathrm{H}_1 : \vec{X} \Subset \Pi_{\lambda}$.

12.5. Пусть $\vec{X} \Subset \Phi_{a,1}$. Для проверки гипотез $\mathrm{H}_0 : a = 0$ против $\mathrm{H}_1 : a = 1$ используется следующий критерий: H_0 принимается, если $X_{(n)} < 3$, и отвергается в противном случае. Найти вероятности ошибок.

12.6*. Используя конструкции доверительного интервала, построить критерий асимптотического уровня ε для проверки гипотезы $\mathrm{H} : \theta = 1$, если а) $\vec{X} \Subset \mathrm{E}_{\theta}$; б) $\vec{X} \Subset \mathrm{B}_{\theta/2}$; в) $\vec{X} \Subset \Pi_{\theta}$.

12.7. Вычислить значение статистики Колмогорова по реализации выборки (1,1; 0,4; 0,2; 3,2), если основная гипотеза состоит в том, что распределение элементов выборки — равномерное на $[0, \ 4]$.

12.8. Вычислить достигнутый уровнень значимости критерия Колмогорова, если объем выборки равен 100, а $\sup_{-\infty<t<\infty} |F_n^*(t) - F_0(t)| = 0,2$.

12.9. При 4040 бросаниях монеты Бюффон получил $\nu_1 = 2048$ выпадений герба и $\nu_2 = n - \nu_1 = 1992$ выпадений решетки. Согласуется ли это с гипотезой о том, что монета правильная, при уровне значимости 0,1? С каким предельным уровнем значимости может быть принята эта гипотеза?

12.10. При $n = 4000$ независимых испытаний события $A1$, $A2$, $A3$, составляющие полную группу, осуществились соответственно 1905, 1015 и 1080 раз. Проверить, согласуются ли эти данные при уровне значимости 0,05 с гипотезой H_0: $p_1 = 1/2$, $p_2 = p_3 = 1/4$, где $p_j = \mathbf{P}(A_j)$. Найти достигнутый уровень значимости.

12.11. В экспериментах с селекцией гороха Мендель наблюдал частоты различных видов семян, полученных при скрещивании растений с круглыми желтыми семенами и растений с морщинистыми зелеными семенами. Эти данные и значения теоретических вероятностей по теории наследственности приведены в следующей таблице:

Семена	Частота	Вероятность
Круглые и желтые	315	9/16
Морщинистые и желтые	101	3/16
Круглые и зеленые	108	3/16
Морщинистые и зеленые	32	1/16
Σ	n=556	1

Следует проверить гипотезу H_0 о согласовании частотных данных с теоретическими вероятностями (на уровне значимости 0,1) и найти достигнутый уровень значимости.

12.12. В таблице приведены числа m_i участков равной площади 0,25 км^2 южной части Лондона, на каждый из которых приходилось по i попаданий самолетов-снарядов во время второй мировой войны. Проверить согласие опытных данных с законом распределения Пуассона, приняв за уровень значимости $\alpha = 0,05$:

i	0	1	2	3	4	5 и более	Итого
m_i	229	211	93	35	7	1	$\Sigma m_i = 576$

§13. Статистические критерии для нескольких выборок

13.1. Пусть $\vec{X}, \vec{Y}$ — независимые выборки объема 2 из непрерывного распределения. Составить таблицу распределения случайной величины $d_{2,2}(\vec{X}, \vec{Y})$.

13.2. По следующим реализациям выборок вычислить реализацию статистики $d_{n,m}$ Колмогорова—Смирнова:

$\vec{X}$ = (1,2, 0,4, -0,2, 0,9), $\vec{Y}$ = (0,2, -0,5, 1, -0,9, 0,3, 0,5).

13.3. По реализациям независимых выборок $\vec{X}, \vec{Y}$ объемов 40 и 50 вычислено значение $\sup_{t\in\mathbf{R}} |F^*_{1,n}(t) - F^*_{2,m}(t)| = 0,1$. Найти достигнутый уровень значимости гипотезы об однородности. Сделать вывод о том, принимается ли эта гипотеза на уровне 0,05.

13.4. По реализациям независимых выборок $\vec{X}, \vec{Y}$ объемов 20 и 30 из нормального распределения вычислены значения статистик $S^2_x = 15$ и $S^2_y = 10$. Найти реально достигнутый уровень значимости гипотезы о равенстве дисперсий против двусторонней альтернативы, а также против каждой из односторонних альтернатив.

13.5. Пусть в условиях предыдущей задачи предполагается равенство дисперсий, и известны значения $\overline{X} = 2, \overline{Y} = 12$. Найти реально достигнутый уровень значимости гипотезы о равенстве математических ожиданий против двусторонней альтернативы, а также против каждой из односторонних альтернатив.

13.6. Пусть $(X_1, Y_1), \ldots, (X_4, Y_4)$ — выборка объема 4 из двумерного нормального распределения. Найти, при каких значениях выборочного коэффициента корреляции гипотеза о независимости компонент отвергается на уровне 0,1 при двусторонней альтернативе. Вычислить значение выборочного коэффициента корреляции по реализации двумерной выборки (1, 2), (2, 3), (-1, 0), (0, 0).

13.7*. Пусть $(X_1, Y_1), \ldots, (X_{100}, Y_{100})$ — реализация выборки объема 100 из двумерного нормального распределения. Значение выборочного коэффициента корреляции $\widehat{r}_n$ равно –0,25. Найти реально достигнутый уровень значимости гипотезы о независимости компонент против двусторонней альтернативы, а также против каждой из односторонних альтернатив.

Таблица нормального распределения

Значения функции $\Phi(t) = \frac{1}{\sqrt{2\pi}} \cdot \int\limits_{-\infty}^{t} e^{-\frac{u^2}{2}} du$ и функции

$\overline{\Phi}(t) = \Phi(-t) = 1 - \Phi(t)$.

t	$\Phi(-t)$	$\Phi(t)$
4,75	0,000001	0,999999
4,26	0,00001	0,99999
3,72	0,0001	0,9999
3,09	0,001	0,999
2,58	0,005	0,995
2,33	0,01	0,99
2,05	0,02	0,98
1,96	0,025	0,975
1,88	0,03	0,97
1,75	0,04	0,96
1,64	0,05	0,95
1,28	0,1	0,9
0,84	0,2	0,8
0,52	0,3	0,7
0,25	0,4	0,6
0,00	0,5	0,5

Для $|t| > 4,75$ можно использовать аппроксимацию $\overline{\Phi}(t) \sim \frac{e^{-t^2/2}}{t\sqrt{2\pi}}$.