МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«Курганский государственный университет»

Кафедра «Программное обеспечение автоматизированных систем»

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Методические рекомендации
к выполнению курсовых работ
для студентов направления (специальности)
09.03.04, 090303.65

Курган 2015

Кафедра: «Программное обеспечение автоматизированных систем»

Дисциплина: «Теория вероятностей и математическая статистика» (направление 09.03.04, специальность 090303.65).

Составил: канд. физ.-мат. наук, профессор В.А. Симахин, ассистент О.С. Черепанов.

Утверждены на заседании кафедры «11» июня 2015 г.

Рекомендованы методическим советом университета
«19» декабря 2014 г.

СОДЕРЖАНИЕ

1 ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Методические указания состоят из двух разделов: теории вероятностей и математической статистики. В каждом разделе даны задания на выполнение курсовой работы и подробные решения типовых задач.

Раздел 1 «Теория вероятностей» представлен задачами по темам: алгебра событий, классическое определение вероятности, формула полной вероятности, формула Бейеса, дискретная случайная величина и ее распределения, непрерывная случайная величина и ее распределения, предельные теоремы.

1.1 Варианты заданий

Вариант 1

1 Вероятность поражения мишени стрелком при одном выстреле равна 0,75. Найти вероятность того, что при 150 выстрелах мишень будет поражена не менее 70 и не более 80 раз.

2 В лотерее 30 билетов, из которых 4 выигрышные. Приобретено 3 билета. Найти вероятность всех исходов.

3 Дискретная величина X распределена по закону

X	1,2	2	2,5	3	4
P	0,1	0,2	0,3	0,2	0,2

Найти ее математическое ожидание, дисперсию и среднее квадратическое отклонение, построить многоугольник распределения.

4 Непрерывная случайная величина X задана интегральной функцией:

$$F(x) = \begin{cases} 0 \text{ при } x < 1, \\ \log_5 x \text{ при } 1 \le x \le 5, \\ 1 \text{ при } x > 5. \end{cases}$$

Найти дифференциальную функцию $f(x)$; математическое ожидание и дисперсию X; вероятность того, что X принимает значение, заключенное в интервале (2; 3) , построить графики функций $F(x)$ и $f(x)$.

5 В трех урнах находятся черные и белые шары: в первой - 3 черных и 2 белых, во второй - 3 белых и 1 черный, в третьей - 2 белых и 2 черных. Шар, извлеченный урны, оказался белым. Какова вероятность того, что этот шар извлечен из первой урны?

6 Исследованы две случайные величины X и Y. Доверительный интервал для неизвестного математического ожидания вычислен с доверительной вероятностью γ_X =0,9, для случайной величины Y получен такой же доверительный интервал с доверительной вероятностью γ_Y = 0,95. Во сколько раз отличаются средние квадратические отклонения случайных величин X и Y? Объемы выборок обеих величин одинаковы и больше 60.

Вариант 2

1 На сборку поступают однотипные изделия из 4 цехов. Вероятности

брака в каждом из цехов соответственно равны 0,03; 0,02; 0,05; 0,02. Первый цех дает 30% этого изделия, второй - 25%, третий - 20% и четвертый - 25%. Какова вероятность, что взятое наудачу изделие окажется бракованным.

2 В лотерее на каждые 100 билетов падает 5 выигрышных. Найти вероятность того, что на 6 билетов выпадет от двух до четырех выигрышей. Найти наивероятнейшее число выигрышей на 6 билетов.

3 Дискретная величина X распределена по закону

X	3	4	5	7	12
P	0,2	0,3	0,1	0,2	0,2

Найти ее математическое ожидание, дисперсию и среднее квадратическое отклонение, построить многоугольник распределения.

4 Вероятность соблюдения правил при прохождении пассажиров через автоматический контрольный пост метрополитена равна 0,9. Сколько пассажиров должно пройти через автоматический контрольный пост, чтобы с вероятностью 0,95 можно было ожидать отклонение относительной частоты соблюдения правил от вероятности не более, чем на 0,03.

5 Абонент забыл последнюю цифру номера телефона и потому набирает ее наудачу. Определить вероятность того, что ему придется звонить не более, чем три раза.

6 Исследуется две случайные величины X и Y, подчиняющиеся нормальному закону распределения. Известны средние квадратичные отклонения этих величин σ_X и σ_Y, причем $\sigma_X = 2\sigma_Y$. Для этих случайных величин получены две выборки объемом n. В каком соотношении находятся доверительные интервалы для неизвестных математических ожиданий этих случайных величин, если соответствующие доверительные вероятности равны γ_X =0,95; γ_Y =0,9 соответственно.

Вариант 3

1 В урне находится 25 шаров одинаковых по размеру, но отличающихся по цвету: 15 черных и 10 белых. Из урны извлекается одновременно 8 шаров. Какова вероятность того, что среди 8 шаров 5 черных и 3 белых?

2 Изделия изготовляются параллельно на двух станках. Вероятность брака на одном станке равна 0,04, на другом - 0,08. Определить вероятность того, что из 10 изделий, изготовленных по 5 на каждом станке, будет не менее 9 годных.

3 Непрерывная случайная величина X задана интегральной функцией:

$$F(x) = \begin{cases} 0 \text{ при } x \le 0, \\ \sin(x) \text{ при } 0 < x < \pi/2, \\ 1 \text{ при } x \ge \pi/2. \end{cases}$$

Найти дифференциальную функцию *f(x)*; математическое ожидание и дисперсию X; вероятность того, что X принимает значение, заключенное в интервале $\left(\frac{\pi}{6};\frac{\pi}{4}\right)$, построить графики функций F(x) и f(x).

4 Дискретная величина X распределена по закону:

X	1,3	1,5	1,7	2,2	2,8
P	0,1	0,3	0,2	0,1	0,3

Найти ее математическое ожидание, дисперсию и среднее квадратическое отклонение.

5 Вероятность изготовления стандартной детали на автомате равна 0,95. Изготовлена партия, состоящая из 200 деталей. Найти наиболее вероятное число нестандартных деталей в этой партии и его вероятность.

6 Задана выборка значений случайной величины X:

101; 102; 115; 113; 111; 119; 118; 117; 113; 112;
117; 119; 113; 116; 114; 117; 118; 111; 114; 118.

Одна из доверительных границ для неизвестного математического ожидания равна 111,8. Определить доверительную вероятность, с которой производились вычисления доверительного интервала.

Вариант 4

1 В цехе работает 4 станка, причем вероятность остановки для каждого из них в течение часа одна и та же и равна 0,8. Построить ряд распределения вероятности числа станков, останавливавшихся в течение данного часа.

2 Определить вероятность того, что среди 500 отобранных изделий число изделий первого сорта отклонится по абсолютной величине от наивероятнейшего их числа не больше, чем на 25 единиц, если доля изделия первого сорта равна P = 0,6.

3 Непрерывная случайная величина X задана интегральной функцией:

$$F(x)=\begin{cases}0 \text{ при } x \le 0,\\ 2\sin(x) \text{ при } 0 < x < \pi/6,\\ 1 \text{ при } x \ge \pi/6.\end{cases}$$

Найти дифференциальную функцию f(x); математическое ожидание и дисперсию X; вероятность того, что X принимает значение, заключенное в интервале $\left(\frac{\pi}{12};\frac{\pi}{8}\right)$, построить графики функций F(x) и $f(x)$.

4 В круг радиуса R вписан равносторонний треугольник. Какова вероятность того, что 4 точки, наудачу поставленные в данном круге, окажутся внутри треугольника.

5 Два стрелка поочередно стреляют в мишень. Вероятность попадания первыми выстрелами для них равна 0,4 и 0,5 соответственно, а вероятность попадания при последних выстрелах для каждого увеличивается на 0,05. Какова вероятность, что первым произвел выстрел первый стрелок, если при пятом выстреле произошло попадание в мишень.

6 Выборочная средняя и выборочная дисперсия, вычисленные по выборке объема $n = 8$, равны $\overline{X} = 12$, $D[x] = 49$ соответственно. Найти доверительный интервал для неизвестного математического ожидания с заданной надежностью $\gamma = 0{,}95$.

Вариант 5

1 Проводятся 4 независимых испытания, в каждом из которых событие A происходит с вероятностью 0,3. Событие B наступает с вероятностью, равной 1, если событие A произошло не менее 2 раз; не может наступить, если событие A не имело места, наступает с вероятностью 0,6, если событие A имело место один раз. Определить вероятность появления события B.

2 При установившемся технологическом процессе цех выпускает в среднем 80% продукции первого сорта. Какова вероятность того, что в партии из 125 изделий будет не менее 100 изделий первого сорта?

3 Непрерывная случайная величина X задана интегральной функцией:

$$F(x) = \begin{cases} 0 \text{ при } \mathrm{x} < 1, \\ \ln(\mathrm{x}) \text{ при } 1 \le \mathrm{x} \le e, \\ 1 \text{ при } \mathrm{x} \ge e. \end{cases}$$

Найти дифференциальную функцию f(x); математическое ожидание и дисперсию X; вероятность того, что X принимает значение, заключенное в интервале (1,8; 2), построить графики функций F(x) и f(x).

4 В коробке 15 яблок, одинаковых по внешнему виду, но отличающихся сортом, а именно: 10 яблок сорта А и 5 яблок сорта В. Из коробке берутся 6 яблок сразу. Какова вероятность того, что среди этих 6 яблок окажется 4 яблока сорта А и 2 яблока сорта В?

5 Мишень состоит из 2 концентрических кругов с радиусом r и R, $r<R$. Считая равновозможным попадание пули в любую часть круга R, определить вероятность того, что при 2 выстрелах будет одно попадание в круг радиуса r.

6 Имеется выборка значений случайной величины X: 12; 13; 14,5; 17; 12,7; 15,1; 15,5. Найти доверительный интервал для неизвестного математического ожидания с заданной надежностью $\gamma = 0{,}9$.

Вариант 6

1 Игральная кость бросается 20 раз. Какова вероятность того, что грань кости с цифрой 3 выпадает не более 5 раз.

2 На склад поступает продукция трех фабрик, причем изделия 1 фабрики составляют 35%, второй - 38% и третьей - 27%. В продукции первой фабрики 70% изделий высшего сорта, второй - 40% и третьей - 50%. Найти вероятность того, что среди 300 изделий, наудачу взятых со склада, число изделий высшего сорта заключено между 130 и 170.

3 Дискретная величина X распределена по закону:

X	2	2,4	3	3,5	4
P	0,2	0,1	0,3	0,2	0,2

Найти ее математическое ожидание, дисперсию и среднее квадратическое отклонение, построить многоугольник распределения.

4 Вероятность наступления события в каждом из независимых испытаний постоянна и равна 0,7. Сколько испытаний нужно произвести, чтобы вероятность отклонения частоты от 0,7 в ту или другую сторону менее, чем на 0,01, была равна 0,995.

5 Ящик содержит 90 годных и 10 бракованных деталей. Найти вероятность того, что среди 3 деталей, наугад вынутых из ящика, нет бракованных.

6 Среднее квадратическое отклонение нормально распределенной случайной величины X равно 3,5. По выборке объемом n=40 найден доверительный интервал для неизвестного математического ожидания с доверительной вероятностью γ = 0,85. Какой объем выборки потребуется для того, чтобы получить такой же доверительный интервал с доверительной вероятностью γ = 0,95?

Вариант 7

1 Вероятность появления события в каждом из 10000 независимых испытаний постоянна и равна 0,75. Найти вероятность того, что относительная частота появления события отклонится от вероятности по абсолютной величине не более, чем на 0,001.

2 Два автомата производят детали, которые поступают на общий конвейер. Вероятность получения нестандартной детали на 1 автомате равна 0,06, на 2-м - 0,09. Производительность второго автомата в 2 раза больше, чем первого. Найти вероятность того, что изделие, наудачу взятое с конвейера, нестандартно.

3 Непрерывная случайная величина X задана интегральной функцией:

$$F(x) = \begin{cases} 0 \text{ при } x < 1, \\ \ln(x) \text{ при } 1 \le x \le 10, \\ 1 \text{ при } x \ge 10. \end{cases}$$

Найти дифференциальную функцию f(X), математическое ожидание и дисперсию X, вероятность того, что X принимает значение, заключенное в интервале (5,6) , построить графики функций F(x) и $f(x)$.

4 В семье 10 детей. Считая вероятности рождения мальчика и девочки равными 0,5, определить вероятность того, что в данной семье: а) 5 мальчиков, б) мальчиков не менее 3, но не более 8.

5 В студии телевидения имеются три телевизионные камеры. Для каждой камеры вероятность того, что она включена в данный момент, равна 0,6. Найти вероятность того, что в данный момент включена хотя бы одна камера.

6 Нормально распределенная случайная величина имеет среднее квадратическое отклонение σ=2,5. По выборке объема n с данной доверительной вероятностью γ=0,95 найдены границы доверительного интервала [11,58; 13,22] . Определить чему равен объем выборки n?

Вариант 8

1 С помощью карточек, на которых написано по одной букве, составлено слово «карета». Карточки перемешиваются, затем наугад извлекаются по одной. Какова вероятность того, что в порядке поступления букв образуется слово «ракета».

2 Три стрелка независимо друг от друга стреляют по мишени. Вероятность попадания в мишень первого стрелка - 0,8, для второго стрелка - 0,7, для третьего - 0,6. После стрельбы в мишени обнаружена одна пробоина. Найти вероятность того, что эта пробоина принадлежит первому стрелку.

3 Дискретная величина X распределена по закону:

X	1,2	1,5	2,3	2,7	3,3
P	0,1	0,4	0,1	0,2	0,2

Найти ее математическое ожидание, дисперсию и среднее квадратическое отклонение, построить многоугольник распределения.

4 Вероятность того, что изделие не выдержит испытания, равна 0,001. Найти вероятность того, что из 5000 изделий не менее двух не выдержат испытаний.

5 В трех урнах находятся черные и белые шары; в первой - 3 черных и 2 белых, во второй - 3 белых и один черный, в третьей - 2 белых и 2 черных. Некто выбирает наугад одну из урн и извлекает из нее шар. Найти вероятность того, что этот шар белый.

6 Исследуется случайная величина X, подчиняющаяся нормальному закону. Получены две выборки значений случайной величины объема n_1 и n_2 значений, $n_2 = 40$. В каком соотношении находятся доверительные интервалы для неизвестного математического ожидания этой случайной величины, если соответствующие доверительные вероятности равны $\gamma_1 = 0,90$ и $\gamma_2 = 0,95$ соответственно.

Вариант 9

1 Из партии деталей, среди которых 10 штук годных и 5 бракованных, для контроля наудачу взято 10 штук. При контроле оказалось, что первые 4 детали - годные. Определить вероятность того, что следующая деталь будет годная.

2 Вероятность того, что изготовленная на первом станке деталь будет первого сорта, равна 0,7. При изготовлении такой же детали на втором станке вероятность равна 0,8. На первом станке изготовлено 2 детали, на втором - 3 детали. Найти вероятность того, что все детали первого сорта.

3 Дискретная величина X распределена по закону:

X	2	4	5	5,3	6,1
P	0,15	0,2	0,25	0,3	0,1

Найти ее математическое ожидание, дисперсию и среднее квадратическое отклонение, построить многоугольник распределения.

4 Численность выборки - 800 единиц. Доля признака 0,4. Найти с вероятностью 0,8 в каких пределах находится отклонение частоты от доли.

5 В ящике 1000 билетов, из них на один билет падает выигрыш 500 рублей, на 10 билетов - выигрыш по 100 рублей, на 50 билетов - выигрыши по 20 рублей, на 100 билетов - выигрыши по 5 рублей. Некто покупает один билет. Найти вероятность выигрыша на сумму не менее 20 рублей.

6 При исследовании случайной величины X получены две выборки объема n_1 и n_2: $n_1 = 10$; $n_2 = 15$. В каком соотношении находятся величины доверительных интервалов для неизвестного математического ожидания этой случайной величины, если соответствующие доверительные вероятности равны $\gamma_1 = 0{,}9$ и $\gamma_2 = 0{,}95$ соответственно? При решении задачи можно считать, что исправленная выборочная дисперсия при увеличении объема выборки осталась прежней.

Вариант 10

1 Имеются две партии изделий по 12 и 10 штук, причем в каждой партии одно изделие бракованное. Изделие, взятое наудачу из первой партии, переложено во вторую. Затем выбирается изделие, наудачу взятое из второй партии. Определить вероятность извлечения бракованного изделия.

2 Вероятность появления некоторого события в каждом из 18 независимых опытов равна 0,2. Определить вероятность появления этого события по крайней мере три раза.

3 Непрерывная случайная величина X задана интегральной функцией:

$$F(x) = \begin{cases} 0 \text{ при } x < 0, \\ x^2 \text{ при } 0 \le x \le 1, \\ 1 \text{ при } x > 1. \end{cases}$$

Найти дифференциальную функцию $f(x)$; математическое ожидание и дисперсию X; вероятность того, что X принимает значение, заключенное в интервале (0,2; 0,3), построить графики функций $F(x)$ и $f(x)$.

4 В группе 20 спортсменов: 13 лыжников, 5 велосипедистов и 2 бегуна. Вероятность выполнения квалификационной нормы равна для лыжников 0,9; для велосипедистов - 0,8; для бегуна - 0,75. Найти вероятность того, что спортсмен, вызванный наудачу, выполнит норму.

5 На станке-автомате изготовили 90 изделий. Чему равна вероятность изготовления на этом станке изделия первого сорта, если наивероятнейшее число таких изделий в данной партии равно 82.

6 Выборочная средняя и исправленная выборочная дисперсия, вычисленные по малой выборке (n<20), равны $\overline{X} = 11$; $S^2 = 20{,}25$ соответственно. Одна из доверительных границ для неизвестного математического ожидания, вычисленная с доверительной вероятностью γ = 0,99, равна 14,85. Найти объем выборки.

Вариант 11

1 В тире имеется 4 ружья, вероятности попадания из которых равны 0,5; 0,6; 0,7; 0,8 соответственно. Определить вероятность попадания при одном выстреле, если стреляющий берет одно из ружей наудачу.

2 Вероятность появления хотя бы одного события при четырех независимых опытах равна 0,59. Какова вероятность появления события *A* при одном опыте, если при каждом из них эта вероятность одинакова.

3 Дискретная величина *X* распределена по закону:

X	1,2	1,5	1,8	2,1	2,4
P	0,1	0,2	0,2	0,2	0,3

Найти ее математическое ожидание, дисперсию и среднее квадратическое отклонение, построить многоугольник распределения.

4 В первом кармане имеются 3 монеты по 20 копеек и 4 монеты по 3 копейки. Из первого кармана во второй перекладывается пять монет, взятых наудачу. Определить вероятность извлечения из первого кармана после перекладывания монеты в 20 копеек, если монета берется наудачу.

5 Для уничтожения танка требуется не менее двух попаданий. Найти вероятность уничтожения танка десятью выстрелами, если вероятность попадания в танк при каждом выстреле равна 0,4.

6 Выборочная средняя и исправленная выборочная дисперсия, вычисленные по двум выборкам объема n_1 и n_2, равны $\overline{X} = 12$; $S^2 = 16$. Известно, что n_1=8. По обеим выборкам вычислены доверительные интервалы для неизвестного математического ожидания с доверительной вероятностью γ=0,95. Определить, чему равно n_2, если величина второго доверительного интервала меньше в 1,5 раза.

Вариант 12

1 Известно, что на один лотерейный билет выпал выигрыш, вероятность того, что выигрышем будет велосипед или стиральная малина, равны 0,03 и 0,02 соответственно. Найти вероятность выигрыша хотя бы одного из этих предметов из 10 выигрышных билетов, выбранных из разных серий.

2 Вероятность поражения мишени равна 0,75. Найти вероятность того, что при 100 выстрелах мишень будет поражена: а) не более 80 и не менее 70 раз, б) не более 70 раз.

3 Непрерывная случайная величина *X* задана интегральной функцией:

$$F(x) = \begin{cases} 0 \text{ при } x < 0, \\ x^3 \text{ при } 0 \le x \le 1, \\ 1 \text{ при } x > 1. \end{cases}$$

Найти дифференциальную функцию *f(x)*; математическое ожидание и дисперсию *X*; вероятность того, что *X* принимает значение, заключенное в интервале (0,1; 0,2), построить графики функций F(x) и f(x).

4 На семи карточках написаны буквы: на четырех «т», «б», «л», «с» соответственно и на трех оставшихся буква «и». Какова вероятность того, что при произвольном извлечении букв в порядке поступления образуется слово «тбилиси».

5 Определить вероятность того, что номер первой встретившейся автомашины не содержит: а) цифры 5; б) двух цифр 5, если известно, что все номера четырехзначные и неповторяющиеся.

6 Исправленная выборочная дисперсия, вычисленная по выборке объема n=11, равна S=9. Одна из доверительных границ (меньшая по абсолютной величине) для неизвестного математического ожидания равна 43,95. Вычислен доверительный интервал с доверительной вероятностью 0,95. Найти величину выборочной средней, использованной в расчетах.

Вариант 13

1 Какова вероятность того, что герб появится не менее 15 раз и не более 25 раз при условии, что монета брошена 40 раз.

2 Среди продукции, изготовленной на данном станке, брак составляет 2%. Сколько изделий необходимо взять, чтобы с вероятностью 0,987 можно было ожидать, что частота бракованных деталей среди них отличается от 0,02 по абсолютной величине не более, чем на 0,03.

3 Непрерывная случайная величина X задана интегральной функцией:

$$F(x)=\begin{cases}0 \text{ при } x<0,\\ \sqrt{x} \text{ при } 0\le x\le 1,\\ 1 \text{ при } x>1.\end{cases}$$

Найти дифференциальную функцию $f(x)$; математическое ожидание и дисперсию X; вероятность того, что X принимает значение, заключенное в интервале (0,15; 0,2), построить графики функций F(x) и $f(x)$.

4 Производится три выстрела по одной мишени. Вероятность попадания при первом, втором и третьем выстрелах равна P_1 = 0,6, P_2 = 0,7, P_3 = 0,85 соответственно. Найти вероятность того, что в результате этих трех выстрелов в мишени будет одна пробоина.

5 В урне 4 красных, 5 зеленых и 5 синих шаров. Найти вероятность того, что два шара, одновременно вынутых из урны, будут одного цвета.

6 Задана выборка значений случайной величины X: 56; 55; 58; 60; 52; 51; 53; 52. С доверительной вероятностью 0,95 вычислить доверительный интервал для неизвестного математического ожидания. Во сколько раз необходимо увеличить объем выборки, чтобы при неизменных выборочной средней $\overline{X}$, исправленной выборочной дисперсии S и неизменном доверительном интервале доверительная вероятность оценки возросла до 0,99?

Вариант 14

1 В урне 10 белых и 8 черных шаров. Найти наивероятнейшее число

выхода белых и черных шаров при 30 и 60 извлечениях, если шары извлекаются с возвращением. Найти соответствующие вероятности.

2 Вероятность изготовления стандартной детали на станке-автомате равна 0,9. Найти вероятность того, что из 100 изделий, взятых наудачу, окажется не менее 84 стандартных.

3 Дискретная величина X распределена по закону:

X	2	4	5,3	6,1	7,2
P	0,1	0,35	0,15	0,2	0,2

Найти ее математическое ожидание, дисперсию и среднее квадратическое отклонение, построить многоугольник распределения.

4 В партии, состоящей из n изделий, k бракованные. Определить вероятность того, что среди m изделий, выбранных наудачу для проверки, y окажутся бракованными.

5 Партия, состоящая из 10 стандартных и 5 нестандартных деталей, разбивается наудачу на 5 частей по три детали в каждой. Найти вероятность того, что в каждой части будет по 2 стандартных и 1 нестандартной детали.

6 Среднее квадратическое отклонение и выборочная средняя нормально распределенной случайной величины X равны σ =3,5; $\overline{X} = 10{,}6$ соответственно. Известна одна из доверительных границ доверительного интервала с доверительной вероятностью γ = 0,85. Эта доверительная граница равна 9,8. Требуется определить объем выборки n, по которой вычислялся доверительный интервал.

Вариант 15

1 Вероятность выигрыша по одному лотерейному билету составляет 0,01. Что вероятнее, выиграть по двум билетам из 4 или по 3 из 5?

2 В цехе имеется 125 станков, работающих независимо друг от друга. Каждый станок оказывается включенным 0,85 рабочего времени. Какова вероятность того, что в некоторый момент времени окажутся включенными не менее 100 станков.

3 Непрерывная случайная величина X задана интегральной функцией:

$$F(x) = \begin{cases} 0 \text{ при } \mathrm{x} < 0, \\ \sqrt[3]{\mathrm{x}} \text{ при } 0 \le \mathrm{x} \le 1, \\ 1 \text{ при } \mathrm{x} > 1. \end{cases}$$

Найти дифференциальную функцию f(x); математическое ожидание и дисперсию X; вероятность того, что X принимает значение, заключенное в интервале (0,1; 0,4), построить графики функций F(x) и f(x).

4 Нормально распределенная случайная величина X имеет математическое ожидание, равное 2,8, и среднее квадратическое отклонение, равное 4,5. Найти вероятность того, что X принимает значение из интервала (1;3).

5 15 экзаменационных билетов содержат по два вопроса, которые не повторяются. Экзаменующийся может ответить на 25 вопросов. Определить вероятность того, что экзамен будет сдан, если для этого достаточно ответить на два вопроса из одного билета или на один вопрос из первого билета и на указанный дополнительный вопрос из другого билета.

6 Среднее квадратическое отклонение, выборочная средняя и объем выборки нормально распределенной случайной величины X равны $\sigma = 2$; $\overline{X} = 12{,}2$; $n = 75$ соответственно. Найти доверительный интервал дня неизвестного математического ожидания с доверительной вероятностью γ=0,95.

Вариант 16

1 Прибор выходит из строя, если перегорит не менее 5 ламп первого типа или не менее 2 ламп второго типа. Определить вероятность выхода из строя прибора, если известно, что перегорело 5 ламп, а вероятность перегорания ламп 1 и 2 типа равна 0,7 и 0,3 соответственно.

2 Вероятность того, что деталь нестандартна, равна 0,1. Отобрано 400 деталей. Найти ту относительную частоту появления нестандартных деталей, которая с вероятностью, равной 0,954, отклоняется от постоянной вероятности P по абсолютной величине не более, чем на 0,3.

3 Дискретная величина X распределена по закону:

X	1	3	4	5,2	7,3
P	0,3	0,2	0,1	0,15	0,25

Найти ее математическое ожидание, дисперсию и среднее квадратическое отклонение, построить многоугольник распределения.

4 Нормально распределенная случайная величина X имеет математическое ожидание, равное 6, и среднее квадратическое отклонение, равное 5. Найти вероятность того, что X принимает значение из интервала (3;4).

5 В ящике находятся 15 теннисных мячей, из которых 9 новых. Для первой игры наугад берутся 3 мяча, которые после игры возвращаются в ящик. Для второй игры также наугад берутся три мяча. Найти вероятность того, что все мячи, взятые для второй игры, новые.

6 Исследована случайная величина X со средним квадратическим отклонением σ_X. По выборке объема n = 40 получена выборочная средняя $\overline{X} = 10{,}2$. С доверительной вероятностью γ = 0,9 вычислены доверительные границы для математического ожидания [9,55; 10,85]. Вычислить по этим параметрам величину σ_X.

Вариант 17

1 На сборку поступают однотипные изделия из 4 цехов, вероятность брака в каждом из них равна 0,04; 0,03; 0,06; 0,02 соответственно. Первый цех доставляет 30, второй - 20, третий - 50 и четвертый - 25 изделий. Какова вероятность того, что изделие, взятое наудачу, окажется бракованным.

2 Вероятность появления хотя бы одного события при 4 независимых опытах равна 0,59. Какова вероятность появления события *A* при одном опыте, если при каждом опыте эта вероятность одинакова.

3 Непрерывная случайная величина *X* задана интегральной функцией:

$$F(x)=\begin{cases}\frac{1}{4}\left(x+|x|\right) \text{ при x} < 2,\\ 1 \text{ при x} \geq 2.\end{cases}$$

Найти дифференциальную функцию *f(x)*, математическое ожидание и дисперсию *X*; вероятность того, что *X* принимает значение, заключенное в интервале (1,2), построить графики функций *F*(x) и *f*(x).

4 Определить вероятность получения не менее 28 очков при 3 выстрелах из спортивного пистолета по мишени с максимальным числом очков, равным 10, если вероятность получения 30 очков равна 0,008. Известно, что при одном выстреле вероятность получения восьми очков равна 0,15, а менее восьми очков - 0,4.

5 Нормально распределенная случайная величина *X* имеет математическое ожидание, равное 4,3, и среднее квадратическое отклонение, равное 9,2. Найти вероятность того, что *X* принимает значение из интервала (5,6).

6 Задана выборка значений случайной величины *X*: 15; 16; 18; 15; 5; 14; 11; 18; 17; 16; 5; 14; 13,5; 19; 17,5; 12; 13,5; 17,5; 18,5; 19; 11; 13. Вычислить доверительные границы для неизвестного математического ожидания с доверительной вероятностью $\gamma = 0{,}99$.

Вариант 18

1 Вероятность того, что лампа останется исправной после 1000 часов работы, равна 0,2. Какова вероятность того, что хотя бы одна из трех ламп останется исправной после 1000 часов работы.

2 Вероятность того, что любой из 300 абонентов позвонит на коммутатор в течение часа, равна 0,01. Какова вероятность того, что в течение часа позвонит 4 абонента.

3 Дискретная величина *X* распределена по закону:

X	0,2	1,2	2,2	3,3	5,8
P	0,1	0,1	0,4	0,2	0,2

Найти ее математическое ожидание, дисперсию и среднее квадратическое отклонение, построить многоугольник распределения.

4 Нормально распределенная случайная величина *X* имеет математическое ожидание, равное 10, и среднее квадратическое отклонение, равное 4. Найти вероятность того, что *X* принимает значение из интервала (5;7).

5 В ящике находится 60 годных и 10 бракованных деталей. Найти вероятность того, что среди трех деталей, наудачу вынутых из ящика, 1) нет

бракованных; 2) хотя бы одна годная.

6 Задана выборка значений случайной величины X: 11,2; 10,8; 10; 12; 12,5; 13; 12,4; 14; 13,1. С доверительной вероятностью 0,95 вычислить доверительный интервал для неизвестного математического ожидания. Как надо увеличить объем выборки, чтобы при неизменных выборочной, средней и исправленной выборочной дисперсии уменьшить доверительный интервал вдвое? Доверительная вероятность так же считается неизменной.

Вариант 19

1 Работница текстильной фабрики обслуживает 4 станка. Вероятность того, что в течение часа станок не потребует внимания работницы, равна для первого станка 0,8, для второго - 0,9, для третьего - 0,85, четвертого - 0,95. Найти вероятность того, что в течение часа ни один станок не потребует внимания работницы и хотя бы один станок потребует внимания работницы?

2 Вероятность появления события в каждом из независимых испытаний равна 0,64. Произведено 144 испытания. Найти вероятность того, что относительная частота появления события отклонится от вероятности 0,64 по абсолютной величине не более, чем на 0,04.

3 Дискретная величина X распределена по закону:

X	1,2	2,3	2,6	2,8	2,9
P	0,1	0,1	0,3	0,4	0,1

Найти ее математическое ожидание, дисперсию и среднее квадратическое отклонение, построить многоугольник распределения.

4 Технический контроль проверяет из партии готовой продукции не более 5 изделий последовательно одно за другим. При обнаружении бракованного изделия бракуется вся партия. Определить вероятность того, что партия будет забракована, если процент брака в этой партии составляет 4%.

5 Нормально распределенная случайная величина X имеет математическое ожидание, равное 6, и среднее квадратическое отклонение, равное 9. Найти вероятность того, что X принимает значение из интервала (6;9).

6 Имеется выборка значений случайной величины X: 101; 110; 112; 105; 103; 115; 120; 111. Одна из доверительных границ для неизвестного математического ожидания равна 121,6. Определить доверительную вероятность, с которой вычислялся доверительный интервал.

Вариант 20

1 У сборщика имеется 20 деталей, 6 изготовлено заводом №1 г. Москвы, 10 - заводом №2 г. Москвы, 4 - заводом г. Ленинграда. Наудачу взяты две детали. Найти вероятность того, что хотя бы одна из них окажется изготовленной в г. Москве.

2 Сколько надо произвести независимых испытаний с вероятностью появления события A в каждом испытании, равной 0,2, чтобы наивероятнейшее

число появления события А в этих испытаниях было равно 20.

3 Случайная величина X задана дифференциальной функцией:

$$f(x)=\begin{cases}\frac{1}{2}\cos(x) \text{ при } -\frac{\pi}{2}\le x\le\frac{\pi}{2},\\ 0 \text{ при } x<-\frac{\pi}{2} \text{ и } x>\frac{\pi}{2}.\end{cases}$$

Найти интегральную функцию F(x): математическое ожидание и дисперсию X; вероятность того, что X принимает значение, заключенное в интервале $\left(0;\frac{\pi}{4}\right)$, построить графики функций F(x) и f(x).

4 Нормально распределенная случайная величина X имеет математическое ожидание, равное 3, и среднее квадратическое отклонение, равное 4. Найти вероятность того, что X принимает значение из интервала (1;3).

5 На участке АВ для мотоциклиста имеются 12 препятствий, вероятность остановки на каждом из которых равна 0,1. Вероятность того, что от пункта В до конечного пункта С мотоциклист проедет без остановки, равна 0,7. Определить вероятность того, что на участке АС не будет ни одной остановки.

6. Выборочная средняя, вычисленная по выборке объемом n = 12, равна $\overline{X}=10{,}8$. Одна из доверительных границ для неизвестного математического ожидания, вычисленная с доверительной вероятностью γ = 0,999, равна 8,24. Найти исправленную выборочную дисперсию, использованную в расчетах.

Вариант 21

1 Сборщик получил 3 ящика деталей; в первом ящике 40 деталей, из них 20 окрашенных; во втором - 50, из них 10 окрашенных, в третьем - 30 деталей, из них 15 окрашенных. Найти вероятность того, что деталь, наудачу извлеченная из ящика, взятого наудачу, окажется окрашенной.

2 Вероятность осуществления события равна 0,36. Произведено 1000 испытаний. Какова вероятность того, что частота появления события отклонится от его вероятности не более, чем на 0,02.

3 Дискретная величина X распределена по закону:

X	2	3,2	5,1	4,2	4
P	0,1	0.1	0,3	0,4	0,1

Найти ее математическое ожидание, дисперсию и среднее квадратическое отклонение, построить многоугольник распределения.

4 Имеется $к_1$ урн, в каждой из которых $м_1$ белых и $н_1$ черных шаров, и $к_2$ урн, содержащих по $м_2$ белых и $н_2$ черных шаров. Извлеченный из наудачу взятой урны один шар оказался белым. Какова вероятность того, что данный шар извлечен из первой урны?

5 Два баскетболиста делают по два броска мячом в корзину. Вероятность попадания мяча при каждом броске равна 0,6 и 0,7 соответственно. Найти вероятность того, что у обоих баскетболистов будет равное количество

попаданий.

6 Имеется выборка значений случайной величины X: 152; 160; 151; 153; 152; 158; 155; 156. С доверительной вероятностью γ = 0,95 вычислить доверительные границы для неизвестного математического ожидания. Как изменятся доверительные границы, если будет известно среднее квадратическое отклонение σ_X, равное 10?

Вариант 22

1 Вероятность того, что деталь нестандартна, равна 0,2. Найти вероятность того, что среди случайно отобранных 500 деталей относительная частота появления нестандартных деталей отклонится от вероятности 0,2 по абсолютной величине не более, чем на 0,03.

2 Имеются пять билетов стоимостью по 100 рублей, три билета - по 300 рублей и два билета - по 200 рублей. Наудачу берутся 3 билета. Определить вероятность того, что а) хотя бы 2 из этих билетов имеют одинаковую стоимость; б) все три билета стоят 700 рублей.

3 Непрерывная случайная величина X задана интегральной функцией:

$$F(x)=\begin{cases} 0 \quad \text{при x} < -\dfrac{\pi}{2}, \\ \dfrac{1}{2}(\sin(x)+1) \text{ при } \quad -\dfrac{\pi}{2} \le x \le \dfrac{\pi}{2}, \\ 1 \text{ при x} > \dfrac{\pi}{2}. \end{cases}$$

Найти дифференциальную функцию f(x), математическое ожидание и дисперсию X; вероятность того, что X принимает значение, заключенное в интервале $\left(0; \dfrac{\pi}{4}\right)$, построить графики функций F(x) и $f(x)$.

4 Для некоторой местности число дождливых дней в августе равно 11. Чему равна вероятность того, что первые два дня августа будут дождливыми?

5 Нормально распределенная случайная величина X имеет математическое ожидание, равное 2,5, и среднее квадратическое отклонение, равное 0,5. Найти вероятность того, что X принимает значение из интервала (0;1).

6 Среднее квадратическое отклонение и объем выборки случайной величины X равны σ =3; n=35 соответственно. Доверительные границы для неизвестного математического ожидания известны и равны 9,4; 11,4. Найти доверительную вероятность γ.

Вариант 23

1 Рабочий обслуживает 4 станка. Вероятность того, что в течение часа первый станок не требует внимания рабочего, равна 0,7. Для второго эта вероятность равна 0,8, для третьего - 0,9 и для четвертого - 0,85. Найти вероятность того, что в течение некоторого часа по крайней мере один станок

потребует к себе внимания рабочего.

2 Вероятность выхода из строя изделия за время испытаний на надежность равна 0,05. Какова вероятность, что за время испытаний 100 изделий выйдут из строя: а) не менее 5 изделий; б) менее 5 изделий; в) от 10 до 20 изделий.

3 Непрерывная случайная величина X задана дифференциальной функцией:

$$f(x) = \frac{1}{\pi(1+x^2)}.$$

Найти интегральную функцию F(x): математическое ожидание и дисперсию X; вероятность того, что X принимает значение, заключенное в интервале (0; 1), построить графики функций F(x) и $f(x)$.

4 Нормально распределенная случайная величина X имеет математическое ожидание, равное 5, и среднее квадратическое отклонение, равное 2. Найти вероятность того, что X принимает значение из интервала (2;5).

5 Вероятность изготовления изделия первого сорта равна 0,8. Сколько должно быть изготовлено изделий, чтобы с вероятностью, не меньшей 0,95, можно было ожидать, что среди них есть хотя бы одно изделие не первого сорта?

6 Выборочная средняя и исправленная дисперсия, вычисленные по выборке объема n = 10, равны $\overline{X} = 11{,}2$; S^2=30,25 соответственно. Вычислить доверительный интервал для неизвестного математического ожидания с заданной доверительной вероятностью γ = 0,95. Как изменятся доверительные границы, если при неизменных параметрах задачи объем выборки возрастет до 48?

Вариант 24

1 Вероятности перегорания первой, второй и третьей ламп равны 0,1; 0,2; 0,3 соответственно. Вероятности выхода из строя прибора при перегорании одной, двух и трех ламп равны 0,25; 0,6; 0,9 соответственно. Определить вероятность выхода из строя прибора.

2 Найти вероятность того, что при 400 испытаниях событие наступит ровно 104 раза, если вероятность его появления при каждом испытании равна 0,2.

3 Непрерывная случайная величина X задана интегральной функцией:

$$F(x) = \frac{1}{\pi}\arg tg(x) + \frac{1}{2}.$$

Найти дифференциальную функцию F(x); математическое ожидание и дисперсию X; вероятность того, что X принимает значение, заключенное в интервале (-1; 1), построить графики функций F(x) и f(x).

4 Монета бросается до тех пор, пока 2 раза подряд она не выпадет одной и той же стороной. Найти вероятность того, что опыт завершится до шестого бросания.

5 Из 18 стрелков 5 попадают в мишень с вероятностью 0,8; 7 стрелков - с

вероятностью 0,7; 4 стрелка - с вероятностью 0,6 и 2 стрелка - с вероятностью 0,5. Наудачу выбранный стрелок произвел выстрел, но в мишень не попал. К какой из групп вероятнее всего принадлежал этот стрелок?

6 Среднее квадратическое отклонение и выборочная средняя нормально распределенной случайной величины *X* равны σ_X; $\overline{X} = 11{,}6$ соответственно. По выборке объема n = 42 вычислены доверительные границы для неизвестного математического ожидания с доверительной вероятностью γ = 0,85. Одна из доверительных границ равна 12,49. По приведенным данным требуется восстановить значение σ_X.

1.2 Примеры решения и оформления задач

Задача 1. В урне 10 белых и 5 черных шаров. Из урны извлекают два шара. Найти вероятность того, что шары а) белые; б) одного цвета; в) разного цвета.

Решение. Пусть событие A означает извлечение белого шара, B - извлечение черного шара и пусть индекс есть номер извлечения.

Тогда в случае а) искомое событие имеет вид A_1A_2 (первый шар – белый и второй шар – белый). Поскольку A_1 и A_2 зависимы, используем формулу вероятности произведения для произвольных событий:

$$P(A_1A_2) = P(A_1)\cdot P_{A_1}(A_2).$$

$P(A_1)$ находим согласно классическому определению вероятности:

$$P(A_1) = \frac{m}{n},$$

где n - общее число случаев,

m - число случаев, благоприятных событию A_1.

Так как среди 15 шаров 10 белых, получаем $P(A_1) = \frac{10}{15}$.

$P_{A_1}(A_2)$ есть условная вероятность события A_2 (второй шар – белый) при условии, что A_1 (первый шар – белый) произошло. Но если первым взят белый шар, то среди 14 оставшихся шаров белых – 9, поэтому $P_{A_1}(A_2) = \frac{9}{14}$.

Получаем: $P(A_1A_2) = P(A_1)P_{A_1}(A_2) = \frac{10}{15}\cdot\frac{9}{14} = \frac{3}{7}$.

В случае б) искомое событие есть $A_1A_2 + B_1B_2$ (оба шара белые или оба шара черные), причем слагаемые несовместны, а сомножители зависимы. Тогда вероятность:

$$P(A_1A_2 + B_1B_2) = P(A_1A_2) + P(B_1B_2) = P(A_1)P_{A_1}(A_2) + P(B_1)P_{B_1}(B_2) =$$

$$= \frac{10}{15}\cdot\frac{9}{14} + \frac{5}{15}\cdot\frac{4}{14} = \frac{11}{21}.$$

В случае в) будем находить вероятность события $A_1B_2 + B_1A_2$ (первый шар белый, второй - черный или наоборот, первый – черный, а второй – белый).

Получим

$$P(A_1B_2 + B_1A_2) = P(A_1B_2) + P(B_1A_2) =$$

$$= P(A_1)P_{A_1}(B_2) + P(B_1)P_{B_1}(A_2) = \frac{10}{15}\cdot\frac{5}{14} + \frac{5}{15}\cdot\frac{10}{14} = \frac{10}{21}.$$

Задача 2. Один завод производит приборов в 3 раза меньше, чем второй. Вероятность безотказной работы прибора первого завода – 0,9, второго – 0,7. Случайным образом выбранный прибор отказал. Какова вероятность, что он сделан на втором заводе?

Решение. Обозначим A - событие, состоящее в том, что выбранный прибор отказал. Возможны две гипотезы: H_1- прибор сделан на первом заводе, H_2 - на втором.

Задача решается по формуле Бейеса, так как событие «прибор отказал» произошло. Запишем формулу Бейеса для случая двух гипотез:

$$P_A(H_2) = \frac{P(H_2)P_{H_2}(A)}{P(H_1)P_{H_1}(A) + P(H_2)P_{H_2}(A)}.$$

Найдем вероятности гипотез H_1 и H_2 до опыта. На один прибор первого завода приходится 3 прибора второго завода, значит, доля первого завода $P(H_1) = \frac{1}{4}$, второго $P(H_2) = \frac{3}{4}$.

Найдем условные вероятности. Вероятность отказа прибора при условии, что он изготовлен на первом заводе, равна $P_{H_1}(A) = 1 - 0,9 = 0,1$. Если же прибор сделан на втором заводе, то вероятность отказа составляет $P_{H_2}(A) = 1 - 0,7 = 0,3$.

Осталось найти вероятность гипотезы H_2 после опыта (то есть при условии, что A произошло):

$$P_A(H_2) = \frac{\frac{3}{4}\cdot 0,3}{\frac{1}{4}\cdot 0,1 + \frac{3}{4}\cdot 0,3} = 0,9.$$

Задача 3. Дискретная случайная величина задана законом распределения:

X	-1	4
P	p_1	p_2

Найти p_1, p_2, если $M(X) = 0,5$.

Решение. Используя формулу для математического ожидания дискретной случайной величины, согласно которой $M(X) = \sum_i x_i p_i$, имеем

$(-p_1 + 4p_2) = 0{,}5$. Кроме того, $p_1 + p_2 = 1$. Решая получившуюся систему уравнений:

$$\begin{cases} -p_1 + 4p_2 = 0{,}5; \\ p_1 + p_2 = 1, \end{cases}$$

находим $p_1 = 0{,}7$, $p_2 = 0{,}3$.

Задача 4. Устройство содержит 2000 одинаковых элементов с вероятностью отказа, равной 0,001, для каждого элемента за время T. Найти вероятность того, что за время T откажут: а) меньше трех элементов; б) не меньше одного элемента.

Решение. Это биномиальное распределение, но поскольку число элементов велико, а вероятность отказа каждого мала, можно применить формулу Пуассона:

$$P_n(m) \approx \frac{a^m}{m!} e^{-a}, \; a = np.$$

Так как $n = 2000$, $p = 0{,}001$, то $np = 2000 \cdot 0{,}001 = 2$.

а) искомая величина:

$$P_{2000}(X < 3) = P_{2000}(0) + P_{2000}(1) + P_{2000}(2) =$$
$$= \frac{2^0}{0!} e^{-2} + \frac{2^1}{1!} e^{-2} + \frac{2^2}{2!} e^{-2} = e^{-2}(1 + 2 + 2) = \frac{5}{e^2};$$

б) используя вероятность противоположного события, получим:

$$P_{2000}(X \ge 1) = 1 - P_{2000}(X < 1) = 1 - P_{2000}(0) = 1 - \frac{2^0}{0!} e^{-2} = 1 - e^{-2} = \frac{e^2 - 1}{e^2}.$$

Задача 5. Непрерывная случайная величина задана плотностью распределения:

$$f(x) = \begin{cases} Ax^3, & x \in [0;1]; \\ 0, & x \notin [0;1]. \end{cases}$$

Найти $P(0{,}5 \le X \le 3)$.

Решение. Так как $\int\limits_{-\infty}^{\infty} f(x)dx = 1$.

Найдем константу А:

$$\int\limits_0^1 Ax^3 dx = 1, \; \left.\frac{Ax^4}{4}\right|_0^1 = 1, \; \frac{A}{4} = 1 \Rightarrow A = 4.$$

Тогда

$$P(a \le X \le b) = \int\limits_a^b f(x)dx,$$

получаем:

$$P(0{,}5 \le X \le 3) = \int\limits_{0{,}5}^{3} f(x)dx = \int\limits_{0{,}5}^{1} 4x^3 dx + \int\limits_{1}^{3} 0dx = x^4 \Big|_{0{,}5}^{1} = 1 - \frac{1}{16} = \frac{15}{16}.$$

Задача 6. Детали, выпускаемые цехом, по размеру диаметра распределяются по нормальному закону с параметрами $M(X) = 3{,}3\,см$ и $D(x) = 0{,}49\,см^2$. Деталь считается годной, если ее диаметр не менее 1,9 см и не более 4 см. Определить процент брака.

Решение. Если случайная величина распределена по нормальному закону, то вероятность попадания её на интервал (α, β) определяется формулой:

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta - m}{\sigma}\right) - \Phi\left(\frac{\alpha - m}{\sigma}\right),$$

где $m = M(X)$ - математическое ожидание,

$\sigma = \sqrt{D(x)}$ - среднее квадратическое отклонение,

$\Phi(x)$ - функция Лапласа, значения которой находят по таблице.

Подставляя в эту формулу заданные значения, получаем (с учетом нечетности функции $\Phi(x)$):

$$P(1{,}9 < X < 4) = \Phi\left(\frac{4 - 3{,}3}{\sqrt{0{,}49}}\right) - \Phi\left(\frac{1{,}9 - 3{,}3}{\sqrt{0{,}49}}\right) = \Phi(1) - \Phi(-2) =$$
$$= \Phi(1) + \Phi(2) = 0{,}3413 + 0{,}4772 = 0{,}8185.$$

Тогда процент годных деталей равен 81,85; соответственно брак составит 18,15%.

Задача 7. Вероятность поражения мишени при одном выстреле 0,8. Найти вероятность того, что при 100 выстрелах мишень будет поражена 88 раз.

Решение. Для решения этой задачи можно воспользоваться локальной теоремой Лапласа:

$$P_n(k) = \frac{1}{\sqrt{npq}}\varphi\left(\frac{k - np}{\sqrt{npq}}\right),$$

где $\varphi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

По условию $n = 100$, $p = 0{,}8$, $q = 1 - p = 0{,}2$, $k = 88$.

Тогда

$$P_{100}(88) = \frac{1}{\sqrt{100 \cdot 0{,}8 \cdot 0{,}2}}\varphi\left(\frac{88 - 100 \cdot 0{,}8}{\sqrt{100 \cdot 0{,}8 \cdot 0{,}2}}\right) = \frac{1}{4}\varphi(2) = \frac{1}{4} \cdot 0{,}054 = 0{,}0135.$$

2 МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Методические указания состоят из двух разделов: теории вероятностей и математической статистики. В каждом разделе даны задания на выполнение курсовой работы и подробные решения типовых задач.

В разделе 2 «Математическая статистика» рассматриваются задания для студентов по курсовой работе и методические указания по исследованию свойств оценок методом статистических испытаний (бутстреп – процедуры).

2.1 Варианты заданий раздела

На основе бутстреп-метода провести исследование и сравнение оценок параметров $\hat{\theta}_1$ и $\hat{\theta}_2$ (из таблицы 1 и номера варианта) параметра θ распределения $F(x)$:

1 Реализовать датчики псевдослучайных величин из распределения $F(x)$ и сформировать выборки заданного объема.

2 Построить оценки плотности вероятностей $f(x)$ случайной величины X по исходным выборкам.

3 Построить оценки параметров $\hat{\theta}_1$ и $\hat{\theta}_2$.

Исследовать оценки параметров $\hat{\theta}_1$ и $\hat{\theta}_2$:

1 На объемах выборки 50 и 500.

2 Без выбросов, при наличии симметричных и асимметричных выбросов модели Тьюки.

3 Используя бутстреп-метод, оценить среднее, дисперсию, среднеквадратическое отклонение (СКО) и доверительные интервалы оценок параметров $\hat{\theta}_1$ и $\hat{\theta}_2$.

4 Провести сравнение оценок $\hat{\theta}_1$ и $\hat{\theta}_2$.

5 Сделать выводы.

Таблица 1 – Варианты заданий

№	Оценка $\hat{\theta}_1$	Оценка $\hat{\theta}_2$	Параметр распределения	Распределение P
1	выборочное среднее	выборочная медиана	параметр сдвига	нормальное распределение
2	оценка Пирсона-Тьюки	выборочная медиана	параметр сдвига	нормальное распределение
3	быстре оценки Кенуя. Средний квартальный размах	быстре оценки Кенуя по пяти квантилям	параметр сдвига	нормальное распределение
4	оценка Ходжеса-Лемана	выборочное среднее	параметр сдвига	нормальное распределение

Продолжение таблицы 1

5	выборочное среднеквадратическое отклонение	оценка на основе среднего абсолютного отклонения	параметра масштаба	нормальное распределение
6	выборочное среднеквадратическое отклонение	линейная оценка Даутона	параметра масштаба	нормальное распределение
7	линейная оценка Даутона	оценка Джини	параметра масштаба	нормальное распределение
8	оценка Джини	оптимальная комплексная оценка на порядковых статистиках	параметра масштаба	нормальное распределение
9	оценка максимального правдоподобия	оценка на порядковых статистиках	параметра масштаба	экспоненциальное распределение
10	выборочное среднее	выборочная медиана	параметр сдвига	Распределение Лапласа
11	выборочное среднее	выборочная медиана	параметр сдвига	Равномерное распределение
12	выборочное среднее	выборочная медиана	параметр сдвига	распределения Коши
13	выборочное среднее	оценка Ходжеса-Лемана	параметр сдвига	распределения Лапласа
14	выборочное среднее	оценка Ходжеса-Лемана	параметр сдвига	равномерное распределения
15	выборочное среднее	оценка Ходжеса-Лемана	параметр сдвига	распределения Коши
16	оценка Ходжеса-Лемана	оценка на порядковых статистиках	параметр сдвига	распределения Лапласа
17	оценка Ходжеса-Лемана	оценка на порядковых статистиках	параметр сдвига	равномерное распределения
18	оценка Ходжеса-Лемана	оценка на порядковых статистиках	параметр сдвига	распределения Коши

Продолжение таблицы 1

19	оценка Ходжеса-Лемана	оценка на порядковых статистиках	параметр сдвига	нормальное распределение
20	оценки усеченного среднего		параметр сдвига	равномерное распределения
21	оценки усеченного среднего		параметр сдвига	нормальное распределение
22	оценки усеченного среднего		параметр сдвига	распределение Лапласа
23	оценки усеченного среднего		параметр сдвига	распределение Коши
24	оценки обобщенно-нормального распределения		параметр сдвига	нормальное распределение
25	оценки обобщенно-нормального распределения		параметр сдвига	равномерное распределение
26	оценки обобщенно-нормального распределения		параметр сдвига	распределение Лапласа
27	оценки обобщенно-нормального распределения		параметр сдвига	распределение Коши
28	выборочная медиана	оценка Коши	параметр сдвига	нормальное распределение

2.2 Методические материалы

2.2.1 Оценка параметров распределения

2.2.1.1 Оценки. Основные определения

Основным объектом исследования в математической статистике является случайная величина, которая полностью характеризуется законом распределения. К настоящему времени предложен достаточно большой набор математических моделей законов распределения: нормальный закон распределения, равномерный закон распределения, показательный закон распределения, χ^2 распределение, распределение Стьюдента и т.д. Каждый закон распределения зависит от набора параметров, например, нормальный закон распределения имеет два параметра: математическое ожидание μ и среднеквадратическое отклонение σ.

Поиск эффективных оценок неизвестных параметров распределения является одним из основных разделов математической статистики. Задача точечного оценивания параметров распределения заключается в нахождении приближенных значений неизвестных параметров на основе обработки реализаций случайной величины X (выборки). Точечной оценкой $\hat{\theta}$ параметра θ называют приближенное значение параметра, полученное на основании выборки $\vec{X}_N = (x_1, x_2, ..., x_N)^T$ объёма N. Важно отметить, что точечная оценка является статистикой (функцией от выборки) и случайной величиной, которая определяется некоторым законом распределения. Рассмотрим ряд требований, которым должны удовлетворять оценки:

1 Состоятельность. Оценка $\hat{\theta}$ называется состоятельной, если она стремится по вероятности к истинному значению параметра θ. Состоятельность означает, что отличие оценки от истинного значения параметра будет уменьшаться с ростом объема выборки

$$\Pr\{|\theta_N - \theta| > \varepsilon\} \to 0, \quad N \to \infty, \varepsilon > 0 .$$

2 Несмещенность.

Под несмещенной оценкой понимается такая оценка $\hat{\theta}$ параметра θ, математическое ожидание которой равно θ, т.е.

$$M(\hat{\theta}) = \theta .$$

Для несмещенных оценок отсутствует систематическая оценка.

3 Эффективность.

Эффективной оценкой называют оценку с минимальной вариацией. Для несмещенных оценок вариация совпадает с дисперсией. В ряде случаев рассматривают асимптотически эффективные оценки, то есть такие, которые обладают минимальной дисперсией при $N \to \infty$.

2.2.1.2 Оценка параметров положения

2.2.1.2.1 Оценка параметра положения нормального распределения

Напомним, что плотность вероятностей нормально распределенной случайной величины определяется формулой вида:

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} ,$$

где μ и σ – параметры сдвига и масштаба соответственно.

2.2.1.2.1.1 Оценка максимального правдоподобия

Оценка максимального правдоподобия определяется оценочным уравнением вида:

$$\sum_{i=1}^{N} \frac{\partial}{\partial \mu} \ln f(x_i, \mu)\Big|_{\mu=\hat{\mu}} = 0 .$$

Найдем частную производную от логарифма плотности по μ:

$$\frac{\partial}{\partial \mu} \ln f(x, \mu) = \frac{\partial}{\partial \mu} f(x, \mu) \frac{1}{f(x, \mu)} = (x - \mu) .$$

Подставляя производную по параметру от логарифма плотности в оценочное уравнение, получим:

$$\sum_{i=1}^{N} (x_i - \hat{\mu}) = 0 \Rightarrow \hat{\mu} = \frac{1}{N} \sum_{i=1}^{N} x_i .$$

2.2.1.2.1.2 Оценка с помощью медианы

В качестве оценки параметра положения μ может быть использована выборочная медиана:

$$\hat{\mu} = \begin{cases} \frac{1}{2}\left(x_{[N/2]} + x_{[N/2+1]}\right), \text{если N/2 - целое,} \\ x_{\left[\frac{N+1}{2}\right]}, \text{если } \frac{\text{N}+1}{2} \text{ - целое,} \end{cases}$$

где $x_{[k]}$ – k-я порядковая статистика, равная k-му по величине значению выборочного ряда $x_1 \le x_2 \le ... \le x_N$, ранжированного по возрастанию.

2.2.1.2.1.3 Оценка Пирсона-Тьюки

Оценки Пирсона-Тьюки – упрощенные оценки, основанные на расстоянии между процентными точками частотных кривых распределений. Для параметра μ нормального распределения предложена оценка вида:

$$\hat{\mu} = x_{[N/2]} + 0{,}185\Delta,$$

$$\Delta = x_{[0{,}95N]} + x_{[0{,}05N]} - 2x_{[0{,}5N]},$$

где $x_{[k]}$ - κ-я порядковая статистика.

2.2.1.2.1.4 Быстрые оценки Кенуя

Средний квартильный размах:

$$\hat{\mu} = \frac{1}{2}\left(x_{[0{,}25N]} + x_{[0{,}75N]}\right),$$

где $x_{[k]}$ - κ-я порядковая статистика.

Быстрая оценка по пяти квантилям:

$$\hat{\mu} = 0{,}2x_{[N/16]} + 0{,}6x_{[N/2]} + 0{,}2x_{[15N/16]},$$

где $x_{[k]}$ - κ-я порядковая статистика.

2.2.1.2.2 Оценки параметра положения обобщенно-нормального распределения

Плотность обобщенно-нормального распределения имеет вид:

$$f(x) = \frac{\beta}{s\Gamma(1/\beta)} e^{-\left(\frac{|x-\mu|}{s}\right)^{\beta}},$$

где $\Gamma(x)$ – гамма-функция; μ – параметр положения; s – параметр масштаба; β – параметр формы.

Оценка максимального правдоподобия определяется оценочным уравнением вида:

$$\sum_{i=1}^{N} \frac{\partial}{\partial \mu} \ln f(x_i, \mu) \Big|_{\mu=\hat{\mu}} = 0 .$$

Найдем частную производную по μ от логарифма плотности:

$$\frac{\partial}{\partial \mu} \ln f(x) = \frac{\partial}{\partial \mu} f(x) \frac{1}{f(x)} = \frac{1}{s} Sign(x-\mu) \left(\frac{|x-\mu|}{s} \right)^{\beta-1} .$$

Подставляя производную в оценочное уравнение, получим:

$$\sum_{i=1}^{N} | x_i - \hat{\mu} |^{\beta-1} = 0 .$$

2.2.1.2.3 Оценка параметра положения распределения Коши

Плотность распределения Коши имеет вид:

$$f(x) = \frac{1}{\pi s \left[1 + \left(\frac{x-\mu}{s} \right)^2 \right]} ,$$

где μ – параметр положения,

s – параметр масштаба.

Оценка максимального правдоподобия определяется оценочным уравнением вида:

$$\sum_{i=1}^{N} \frac{\partial}{\partial \mu} \ln f(x_i) \Big|_{\mu=\hat{\mu}} = 0 .$$

Найдем частную производную от логарифма плотности по μ:

$$\frac{\partial}{\partial \mu} \ln f(x) = \frac{\partial}{\partial \mu} f(x) \frac{1}{f(x)} = 2\pi \left(\frac{x-\mu}{s} \right) f(x, \mu) .$$

Подставляя производную в оценочное уравнение, получим:

$$\sum_{i=1}^{N} (x_i - \hat{\mu}) f(x, \hat{\mu}) = 0 .$$

2.2.1.2.4 Другие оценки параметра положения

Пусть X – случайная величина с функцией распределения с *F(x-μ)* и симметричной унимодальной плотностью *f(x-μ)*. Требуется по выборке независимых и одинаково распределенных случайных величин (н.о.р.) $\vec{X}_N = (x_1, ..., x_N)$ оценить неизвестный параметр положения μ.

2.2.1.2.4.1 Выборочное среднее

Оценка определяется выражением вида:

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^{N} x_i .$$

Выборочное среднее является оптимальной оценкой параметра положения для нормального распределения. Оценка очень чувствительна к наличию выбросов.

2.2.1.2.4.2 Усеченное среднее

Усеченное среднее - выборочное среднее по выборке, из которой исключены α% наибольших и наименьших значений. Оценка определяется выражением вида:

$$\hat{\mu} = \frac{1}{N(1-2\alpha)} \sum_{i=N\alpha}^{N-N\alpha} x_{[i]},$$

где α – параметр усечения,

$x_{[k]}$ – k-я порядковая статистика.

Усеченное среднее менее чувствительно к наличию выбросов, чем выборочное среднее.

2.2.1.2.4.3 Выборочная медиана

Выборочная медиана – 0,5 порядковая статистика, определяемая выражением:

$$\hat{\mu} = \begin{cases} \frac{1}{2}\left(x_{[N/2]} + x_{[N/2+1]}\right), \text{если N/2 - целое,} \\ x_{\left[\frac{N+1}{2}\right]}, \text{если } \frac{\mathrm{N}+1}{2} \text{ - целое.} \end{cases}$$

Выборочная медиана является оптимальной оценкой для распределения Лапласа. Оценка обладает робастными свойствами и поэтому часто рекомендуется для оценки параметра положения симметричных распределения при условии наличия выбросов.

2.2.1.2.4.4 Оценка Ходжеса-Лемана

Определим по выборке $\vec{X}_N = (x_1,...,x_N)$ набор из $\frac{N(N+1)}{2}$ средних вида $z_{i,j} = \frac{x_i + x_j}{2}$ $(i \le j)$, называемых средними Уолша. Оценка Ходжеса-Лемана определяется как медиана средних Уолша, то есть медиана ряда $z_1 \le z_2 \le ... \le z_{N(N+1)/2}$.

Следует отметить высокую устойчивость оценки к наличию аномальных наблюдений в выборке.

2.2.1.2.4.5 Полусумма порядковых статистик

Оценка параметра положения на порядковых статистиках определяется следующим выражением:

$$\hat{\mu} = \frac{x_{[r]} + x_{[1-r]}}{2},$$

где $x_{[r]}$ – r-я порядковая статистика.

2.2.1.3 Оценка параметра масштаба

2.2.1.3.1 Оценка параметра масштаба нормального распределения

2.2.1.3.1.1 Оценка максимального правдоподобия нормального распределения

Оценка максимального правдоподобия параметра масштаба σ определяется формулой:

$$\hat{\sigma} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{\mu})^2} \, ,$$

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} x_i \, .$$

Оценка является состоятельной и эффективной, но асимптотически несмещенной. Поэтому рекомендуется использовать несмещенную оценку вида:

$$\hat{\sigma} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \hat{\mu})^2} \, .$$

2.2.1.3.1.2 Среднее абсолютное отклонение

Оценка вычисляется по формулам:

$$\hat{\sigma} = \frac{1}{N \cdot c}\sum_{i=1}^{N}| x_i - \hat{\mu} |,$$

$$c = \left(\frac{2}{\pi}\left(1 - \frac{1}{N}\right)\right)^{1/2},$$

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} x_i \, .$$

2.2.1.3.1.3 Линейная оценка Даутона

Даутон предложил достаточно простую, но весьма эффективную оценку вида:

$$\hat{\sigma} = \frac{1{,}77245}{N(N-1)}\sum_{i=1}^{N}(2i - N - 1)x_i \, .$$

2.2.1.3.1.4 Оценка Джини

Используется как быстрая оценка параметра σ и подсчитывается по формуле:

$$\hat{\sigma} = \frac{1}{N(N-1)}\sum_{i,j=1}^{N}| x_i - x_j | = \sum_{i=1}^{[N/2]}\frac{N - 2i + 1}{N(N-1)} w_i \, ,$$

где w_i – i-й подразмах.

2.2.1.3.1.5 Оптимальные комплексные оценки на порядковых статистиках

$$\hat{\sigma} = 0{,}2581(x_{[0{,}9332N]} - x_{[0{,}0668N]}) + 0{,}2051(x_{[0{,}7088N]} - x_{[0{,}2912N]}),$$

где $x_{[k]}$ – k-я порядковая статистика.

2.2.1.3.2 Оценки параметра масштаба экспоненциального распределения

Напомним, что плотность вероятностей экспоненциального распределения случайной величины описывается формулой вида:

$$f(x,\lambda) = \frac{1}{\lambda} e^{-\frac{x}{\lambda}},$$

где λ – параметр распределения.

2.2.1.3.2.1 Оценка максимального правдоподобия

Оценка максимального правдоподобия определяется оценочным уравнением вида:

$$\sum_{i=1}^{N} \frac{\partial}{\partial \lambda} \ln f(x_i, \lambda)\Big|_{\lambda=\hat{\lambda}} = 0.$$

Найдем частную производную от логарифма плотности по λ:

$$\frac{\partial}{\partial \lambda} \ln f(x,\lambda) = \frac{\partial}{\partial \lambda} f(x,\lambda) \frac{1}{f(x,\mu)} = \left(\frac{x}{\lambda} - 1\right).$$

Подставляя производную от логарифма плотности по параметру в оценочное уравнение, получим:

$$\sum_{i=1}^{N} \left(\frac{x_i}{\hat{\lambda}} - 1\right) = 0 \Rightarrow \hat{\lambda} = \frac{1}{N} \sum_{i=1}^{N} x_i.$$

2.2.1.3.2.2 Оценка по одной порядковой статистике

Хартер и Эпштейн предложили оценку параметра λ, основанную на одной порядковой статистке x_r. Оценка находится по формуле:

$$\hat{\lambda} = \frac{x_r}{\sum\limits_{i=1}^{N} a_i},$$

где $a_i = \dfrac{1}{N - i + 1}$, x_r – r-я порядковая статистика.

Оценка имеет высокую эффективность при $\dfrac{r}{N} \le \dfrac{2}{3}$ или $\dfrac{r}{N} \le \dfrac{1}{2}$.

2.2.2 Исследование свойств оценок

2.2.2.1 Асимптотический подход

Пусть имеются две состоятельные несмещенные оценки параметра θ. Какую из этих оценок предпочесть? Так как любая оценка как функция от результатов наблюдения является случайной величиной, то ее свойства определяются функцией распределения. Вследствие того, что оценка строится по выборке конечного объема N, то ее функция распределения зависит от N. Однако теоретическое определение законов распределения оценок при конечных объемах выборки в большинстве случаев весьма затруднительно, гораздо проще вычислить асимптотическое распределение оценок. В силу центральной предельной теоремы большинство оценок параметров распределения являются асимптотически нормальными.

Оценка $\hat{\theta}$ называется асимптотически нормальной оценкой параметра θ с коэффициентом (асимптотической дисперсией) $V(\hat{\theta})$, если

$$\sqrt{N}(\theta - \hat{\theta}) \Rightarrow N(0, V).$$

Асимптотический подход для сравнения двух асимптотически нормальных оценок предлагает следующий критерий. Оценка $\hat{\theta}_1$ эффективней оценки $\hat{\theta}_2$, если $V(\hat{\theta}_1) < V(\hat{\theta}_2)$. Для количественного сравнения эффективности оценок используется относительный коэффициент эффективности вида:

$$\xi = \frac{V_1}{V_2}.$$

Необходимо отметить, что поиск асимптотической дисперсии оценки теоретически сложен и зависит от знания и вида закона распределения случайной величины X.

Для примера рассмотрим исследование двух оценок параметра положения μ: $\hat{\theta}_1$ - выборочное среднее и $\hat{\theta}_2$ - выборочная медиана. Обозначим через $\mu = MX$ среднее значение X, $\sigma^2 = DX$ - дисперсию X, $\lambda = Med\ X$ - медиану X. Можно показать, что

$$V(\hat{\theta}_1) = \sigma^2,$$

$$V(\hat{\theta}_2) = [2f(\lambda)]^{-2}.$$

Пусть X - нормальная случайная величина:

$$f(x, \mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu)^2}.$$

В этом случае $\sigma^2 = DX = 1$ и $f(\lambda) = (\sqrt{2\pi})^{-1}$. В результате получаем:

$$\xi = \frac{V(\hat{\theta}_1)}{V(\hat{\theta}_2)} = \frac{1}{(2f(0))^{-2}} = \frac{2}{\pi} \approx 0{,}64.$$

Следовательно, выборочное среднее на 36% эффективнее медианы.

Пусть X - случайная величина с распределением Лапласа

$$f(x,\mu)=\frac{1}{2}e^{-|x-\mu|}.$$

В этом случае $\sigma^2 = DX = 2$ и $f(\lambda) = 2$. В результате получаем:

$$\xi=\frac{V(\hat{\theta}_1)}{V(\hat{\theta}_2)}=\frac{8}{(2f(0))^{-2}}=2.$$

Следовательно, медиана в два раза эффективнее выборочного среднего.

Пусть X - случайная величина с распределением Коши:

$$f(x)=\frac{1}{\pi\left[1+(x-\mu)^2\right]}.$$

В этом случае $\sigma^2 = DX = \infty$ и $f(\lambda)=(\pi)^{-1}$. В результате получаем, что выборочное среднее имеет нулевую эффективность по сравнению с выборочной медианой. Это связано с тем, что распределение Коши имеет сильно затянутые хвосты.

2.2.2.2 Метод статистических испытаний

Одним из перспективных численных методов моделирования и исследования сложных стохастических систем, получивший большое распространение, является метод статистических испытаний (метод Монте-Карло). Оценка как функция от выборки является случайной величиной. При подстановке в оценку конкретной выборки объёма N получаем одно значение оценки $\hat{\theta}_1$ искомого параметра. Пусть имеется следующая серия выборочных значений объёма N. При подстановке в оценку этой серии конкретной выборки объёма N получаем следующее значение оценки $\hat{\theta}_2$ искомого параметра. Предположим, что имеется М серий таких выборок объёма N. Вычислим для каждой серии выборочных значений объёма N значение оценки искомого параметра. В результате получаем выборку оценок $(\hat{\theta}_1,\hat{\theta}_2,...,\hat{\theta}_M)$ из случайной величины θ.

На основе выборки $(\hat{\theta}_1,\hat{\theta}_2,...,\hat{\theta}_M)$ можно найти оценку среднего, дисперсии и СКО оценки θ по формуле:

$$V=\frac{1}{M}\sum_{i=1}^{M}\left(\hat{\theta}_i-\overline{\theta}_M\right)^2,$$

$$СКО = V + b^2,$$

$$b=\overline{\theta}_M-\theta,$$

$$\overline{\theta}_M=\frac{1}{M}\sum_{i=1}^{M}\hat{\theta}_i,$$

где $\overline{\theta}_M$ – выборочное среднее, θ - истинное значение параметра.

Таким образом, удается численно оценить неизвестную дисперсию оценки θ. Но такой подход требует наличия *M* серий повторных выборок объёма *N* каждая, который называют «ресемплингом» (Re-sample). В реальности имеется только одна исходная выборка. Возникает вопрос: «Как получить *M* серий выборок из распределения *F(x)*?»

2.2.3 Построение датчиков псевдослучайных величин

Для исследования свойств оценок методом статического моделирования или методами на основе ресемплинга необходимо получать на ЭВМ последовательности значений случайной величины *X* с заданным распределением *F(x)*. Случайные величины обычно моделируют с помощью преобразования одного или нескольких значений случайной величины U, равномерно распределенной в [0,1]. Поэтому вопросы построения датчиков равномерной случайной величины *U* играют важнейшее значение в задачах статистического моделирования. На всех современных ЭВМ имеются генераторы (датчики) псевдослучайных величин из равномерного распределения. Они входят в набор стандартного программного обеспечения, поставляемого с ЭВМ. В дальнейшем под U_1 ,U_2, U_3 ,...,U_i будем обозначать случайные числа из равномерного в [0,1] распределения случайной величины *U*.

В настоящее время существует большое разнообразие методов построения датчиков случайных чисел: метод обратного преобразования функции распределения; метод Неймана; метод кусочной аппроксимации законов распределения; метод предельных теорем; метод суперпозиции и др.

2.2.3.1 Метод обратного преобразования функции распределения

Рассмотрим наиболее распространенный метод обратного преобразования функции распределения, который является классическим для построения датчиков псевдослучайных величин.

Пусть *U* – равномерная в [0,1] случайная величина и *X* – случайная величина с непрерывной функцией распределения *F(x)*, значения которой нам требуется сгенерировать. Из теории вероятностей известен следующий факт: при применении преобразования *F(x)* имеет место следующее соотношение:

$$U=F(x).$$

Следовательно, если обозначить через $F^{-1}(.)$ функцию, обратную к *F(x)*, то получим:

$$X = F^{-1}(U),$$

где *X* имеет функцию распределения *F(x)*.

Таким образом, имея датчик случайной величины *U* и зная функцию $F^{-1}(.)$, получим следующий алгоритм генерирования случайных чисел $x_1, x_2 ....$ из распределения *F(x):*

$$X_i = F^{-1}(U_i), i = 1,2,...,$$

где $U_1, U_2 ...$ - случайные числа из *U*.

Сложность применения данного метода связана в основном с трудностью нахождения обратной функции $F^{-1}(.)$.

Пример. Требуется сгенерировать случайные числа из экспоненциального распределения с плотностью:

$$f(x,\lambda)=\begin{cases}0 & x<0,\\ \lambda\exp(-\lambda x) & x\ge 0, \lambda>0,\end{cases}$$

и функцией распределения:

$$F(x)=\begin{cases}0 & x<0,\\ 1-\exp(-\lambda x) & x\ge 0.\end{cases}$$

Тогда $X=F^{-1}(U)$ будет иметь вид $X=-\frac{1}{\lambda}\ln(1-U)$ и алгоритм датчика псевдослучайных величин с экспоненциальным распределением примет вид:

$$X_i=-\frac{1}{\lambda}\ln(1-U_i),\ i=1,2\ldots$$

2.2.3.2 Алгоритмы моделирования некоторых типовых распределений

2.2.3.2.1 Равномерное распределение

Параметры распределения: a – нижняя граница области значений, b – верхняя граница области значений.

Функция распределения: $F(x)=\frac{x-a}{b-a}$.

Плотность вероятности: $f(x)=\frac{1}{b-a}$, $x\in[a,b]$.

Моделирование случайной величины X на отрезке $[a,b]$ осуществляется по формуле:

$$x=(b-a)\cdot U+a,$$

где U – равномерно распределенная случайная величина на [0,1].

2.2.3.2.2 Экспоненциальное распределение

Область значений случайной величины $x > 0$. λ – величина, обратная параметру $M(X)$.

Функция распределения: $F(x)=1-e^{-\lambda x}$.

Плотность вероятности: $f(x)=\lambda e^{-\lambda x}$.

Моделирование случайной величины X происходит по формуле:

$$x=-\frac{1}{\lambda}\cdot\ln u.$$

2.2.3.2.3 Распределение Коши

Область значений случайной величины X – множество действительных чисел.

Параметры: μ – параметр расположения; s – параметр масштаба.

Функция распределения: $F(x) = \frac{1}{2} + \frac{1}{\pi} arctg\left(\frac{x-A}{B}\right)$.

Плотность вероятности: $f(x) = \frac{1}{\pi s \cdot \left(1 + \left(\frac{x-\mu}{s}\right)^2\right)}$.

Моделирование случайной величины X происходит по формуле:

$$x = s \cdot tg(\pi \cdot (U - 0.5)) + \mu .$$

2.2.3.2.4 Нормальное распределение

Область значений случайной величины X – любая.

Параметры: μ – параметр расположения, математическое ожидание; s – параметр масштаба.

Плотность вероятности:

$$f(x) = \frac{1}{s\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{s}\right)^2}$$

Моделирование случайной величины X происходит по формуле:

$$x = [(-2 \cdot \ln U_1)^{\frac{1}{2}} \cdot \cos(2\pi U_2)]s + \mu$$

или

$$x = [(-2 \cdot \ln U_1)^{\frac{1}{2}} \cdot \sin(2\pi U_2)]s + \mu ,$$

где U_1, U_2 – независимые случайные величины, равномерно распределенные в [0,1].

2.2.3.3 Генераторы псевдослучайных величин модели выбросов Тьюки

2.2.3.3.1 Описание модели выбросов Тьюки

Модель выбросов Тьюки может быть представлена следующим образом:

$$F(x) = (1-\varepsilon)G(x) + \varepsilon H(x) ,$$

где *G(x)* - основное распределение,

H(x) - распределение выбросов,

ε - доля выбросов.

Рассмотрим частные модели выбросов Тьюки для симметричных распределений $G(x)$.

1 Симметричные выбросы:

$$F(x) = (1-\varepsilon)G(x,\mu,s) + \varepsilon H(x,\mu,s_1) .$$

В данном случае основное распределение и распределение выбросов имеют одинаковые значения параметра положения, но разное значение параметров масштаба.

2 Асимметричные выбросы:

$$F(x) = (1-\varepsilon)G(x, \mu_1) + \varepsilon H(x, \mu_2).$$

В данном случае основное распределение и распределение выбросов имеют разные значения параметров положения.

2.2.3.3.2 Датчик псевдослучайных величин модели выбросов Тьюки

Пусть имеются два псевдослучайных датчика случайных величин g и h из распределения *G(x)* и *H(x)* соответственно. Требуется построить датчик псевдослучайных величин из распределения:

$$F(x) = (1-\varepsilon)G(x) + \varepsilon H(x).$$

Генератор псевдослучайных величин из смеси распределений должен на каждом такте своей работы выбирать из какого распределения в данный момент необходимо сгенерировать число. Для этого воспользуемся датчиком случайных величин u из равномерного распределения U(0,1). Если он сгенерировал число из диапазона $[0,\varepsilon]$, то формируем случайное число датчиком h, иначе датчиком g.

2.2.3.4 Бутстреп – метод

Для исследования свойств оценок параметра θ методом статистических испытаний требуется механизм создания M серий повторных выборок объёма N каждая - «ресемплинг».

Метод исследования свойств оценок методом статистических испытаний с помощью ресемплинга получил название бутстреп – метода. Исходя из априорной информации о знании вида исходного распределения *F(x),* механизм ресемплинга бутстреп – метода делят на два типа: параметрический бутстреп и непараметрический бутстреп. При параметрическом бутстрепе вид исходного распределения *F(x)* известен точно или с точностью до конечного числа неизвестных параметров. При непараметрическом бутстрепе вид исходного распределения *F(x)* неизвестен, известна лишь общая информация о классе функций, к которой принадлежит *F(x)*, например, класс непрерывных симметричных функций распределения.

2.2.3.4.1 Параметрический бутстреп

Пусть априорно вид исходного распределения *F(x)* известен точно и имеется датчик случайных чисел из распределения *F(x).* В этом случае генерируем M повторных выборок объема N каждая $\vec{X}_j^*, \vec{X}_j^* = (x_{1j}^*, \ldots, x_{1j}^*)^T \;\; j = \overline{1, M}$, M ≈500-1000.

Пусть априорно вид исходного распределения *F(x)* известен с точностью до конечного числа неизвестных параметров. По исходной выборке $\vec{X}_N = (x_1, x_2, \ldots, x_N)^T$ объёма N находим состоятельные оценки неизвестных параметров распределения и подставляем их в распределение. Далее ресемплинг осуществляется по вышеописанной схеме.

2.2.3.4.2 Непараметрический бутстреп

При непараметрическом бутстрепе вид исходного распределения *F(x)* неизвестен. Именно в такой постановке возникает центральная проблема построения непараметрического датчика случайных чисел из неизвестного распределения *F(x).* Изначально бутстреп был сформулирован и развивался именно в непараметрической постановке задачи и этим принципиально отличался от классических задач статистического моделирования. В реальности имеем только одну исходную выборку $\vec{X}_N = (x_1, x_2, ..., x_N)^T$ объёма *N*. Возникает вопрос: как получить М серий выборок (ресемплинг) из распределения *F(x)*? Рассмотренные выше методы исследования свойства оценок требуют знания распределения случайной величины *F(x)*. В непараметрической постановке задачи такая информация отсутствует, поэтому требуется универсальная процедура исследования свойств оценок, независящая от распределения случайной величины *X*. К таким простейшим процедурам относятся метод складного ножа (jackknife) и «наивный» бутстреп-метод.

2.2.3.4.2.1 Метод складного ножа

Основная идея метода складного ножа заключается в последовательном исключении из имеющейся выборки каждого элемента с возвратом. Пусть имеется выборка $x=(x_1, x_2, \ldots, x_N)$. Согласно методу складного ножа получим *N* ре-выборок объема *N* вида:

$$z_1=(x_2, x_3, x_4,..,x_N),$$
$$z_2= (x_1, x_3, x_4,..,x_N),$$
$$\ldots.$$
$$z_N= (x_1, x_2, x_3,..,x_{N-1}).$$

На основе полученных ре-выборок находятся оценки параметра $\hat{\theta}_i$, оценки среднего ($\overline{\theta}$), дисперсии (*V)* и СКО оценки следующим образом:

$$\overline{\theta} = \frac{1}{N}\sum_{i=1}^{N}\hat{\theta}_i ,$$

$$\text{V} = \frac{1}{N}\sum_{i=1}^{N}(\hat{\theta}_i - \overline{\theta})^2 ,$$

$$СКО = V + b^2 ,$$
$$b = \overline{\theta} - \theta .$$

Следовательно, в данном методе получаем *M=N* серий ре-выборок объема *N* каждая.

Главным недостатком метода складного ножа является применение его только к гладким дифференцируемым статистикам. Не рекомендуется применять данный метода для оценивания свойств распределения медианы и порядковых статистик.

2.2.3.4.2.2 «Наивный» бутстреп-метод

Бутстреп-метод был предложен Эфроном как некоторое обобщение алгоритма складного ножа. Основная идея бутстрепа состоит в том, чтобы методом статистических испытаний многократно извлекать повторные выборки из исходных наблюдений. Пусть имеется исходная выборка $x=(x_1, x_2, \ldots, x_N)$. Требуется сгенерировать M ре-выборок. На основе генератора псевдослучайных величин, равномерно распределенных на $[1,N]$, генерируются случайные числа, соответствующие индексам элементов исходной выборки, которые помещаются в ре-выборку. Таким образом, можно сформировать любое, сколь угодно большое, число ре-выборок. При этом в одной ре-выборке какие-то элементы исходной выборки могут повторяться, тогда, как другие элементы могут отсутствовать. Например, используя данный метод генерации выборок, можно получить следующие ре-выборки:

$$z_1=(x_{37}, x_{76}, x_{59},..,x_{13}),$$
$$z_2= (x_3, x_{89}, x_{15},..,x_3),$$
$$z_3= (x_{11}, x_{26}, x_{41},..,x_{14}),$$
$$\ldots.$$
$$z_M= (x_{31}, x_{98}, x_{39},..,x_1).$$

Как и для метода складного ножа, на основе полученных ре-выборок строятся оценки среднего $\overline{\theta}$ и дисперсии V вида:

$$\overline{\theta}=\frac{1}{M}\sum_{i=1}^{M}\hat{\theta}_i\ ,$$

$$\mathrm{V}=\frac{1}{M}\sum_{i=1}^{M}(\hat{\theta}_i-\overline{\theta})^2\ ,$$

$$СКО=V+b^2,$$
$$b=\overline{\theta}-\theta,$$

где $\hat{\theta}_i$ - оценка параметра θ вычисленная на основе i-й ре-выборки.

Данный метод ресемплинга основан на построении непараметрического датчика из эмпирического распределения $F_N(x)$, где $F_N(x)$ - эмпирическая функция распределения. Можно использовать другие непараметрические оценки неизвестной $F(x)$ и на их основе строить непараметрические датчики.

ПРИЛОЖЕНИЕ А

Титульный лист

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ «КУРГАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

Кафедра «Программное обеспечение автоматизированных систем»

Курсовая работа
по дисциплине
«Теория вероятностей и математическая статистика»

РФ. КГУ. КР.09.03.04 123456

Разработал студент гр. Т - 20012 __________________/ И.И. Иванов/
Руководитель,
Профессор, к. ф.- м. н. __________________/ В.А. Симахин /

Члены комиссии __________________/В.А. Симахин/

__________________/О.С. Черепанов/

Работа защищена с оценкой _________ «___» _________ 2015 г.

Курган 2015 г.

Симахин Валерий Ананьевич
Черепанов Олег Сергеевич

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Методические рекомендации
к выполнению курсовых работ
для студентов направления (специальности)
09.03.04, 090303.65

Редактор Н.Л. Борисова

Подписано в печать	Формат 60х84 1/16	Бумага 65 г/м^2
Печать цифровая	Усл. печ. л. 2,75	Уч.–изд. л.2,75
Заказ	Тираж 25	Не для продажи

РИЦ Курганского государственного университета.
640000, г. Курган, ул. Советская, 63/4
Курганский государственный университет.