МИНОБРНАУКИ РОССИИ
Федеральное государственное автономное образовательное учреждение
высшего профессионального образования
«ЮЖНЫЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

Шубарин М. А.

МЕТОДИЧЕСКОЕ ПОСОБИЕ
Теория вероятностей и математическая статистика, часть 2

для студентов института Социологии и Регионоведения

Ростов-на-Дону
2016

Учебное пособие разработано доцентом кафедры математического анализа М. А. Шубариным.

Ответственный редактор канд. физ.-мат. наук А. И. Луценко (ЮФУ)

Печатается в соответствии с решением кафедры математического анализа факультета математики, механики и компьютерных наук ЮФУ, протокол № 1 от 5 сентября 2016.

Рекомендовано для передачи и хранения в банк изданий ЮФУ на заседании учебно–методического совета института Социологии и Регионоведения ЮФУ (протокол №3 от 19 октября 2016).

1 Теория

1 Предельные теоремы Муавра–Лапласа и Пуассона

Пусть λ — дискретная случайная величина, имеющая биномиальное рапределение и определяемая пераметрами n, p, $q = 1 - p$. Ряд распределения этой случайной величины имеет вид

ζ	0	1	$\dots$	k	$\dots$	$n-1$	n
p	$P_n(0)$	$P_n(1)$	$\dots$	$P_n(k)$	$\dots$	$P_n(n-1)$	$P_n(n)$

где

$$P_n(k) = C_n^k p^k q^{n-k}. \tag{1}$$

Известно, что $\mathbf{M}\lambda = np$, $\mathbf{D}\lambda = npq$.

Через $P(k_0 \leqslant k \leqslant k_1)$ обозначим вероятность того, что значение случайной величины λ не меньше k_0 и не больше k_1. События $\lambda = k$ (состоящие в том, что значение случайной величины λ равно k) попарно несовместны. Поэтому

$$P(k_0 \leqslant k \leqslant k_1) = \sum_{j=k_0}^{k_1} P(\lambda = k) = \sum_{j=k_0}^{k_1} C_n^k p^k q^{n-k} \tag{2}$$

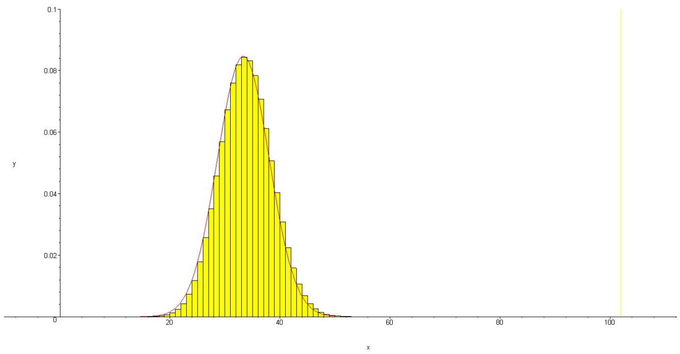
Применение формул (1)–(2) для вычисления вероятностей оправдано только для малых значений числа n.

Рассмотрим вспомагательную случайную величину μ, имеющую нормальное распределение и такую, что $\mathbf{M}\mu = np$, $\sigma_\mu = \sqrt{npq}$. Эту непрерывную случайную величину можно рассматривать как хорошую аппоксимацию дискретной случайной величины.

Пример 1. *Рассмотрим случайные величины λ и μ:*

- *λ — биномиально распределённая случайная величина, определяемая параметрами $n = 100$, $p = 1/3$, $q = 2/3$;*
- *μ — нормально рампределённая случайная величина, определяемая параметрами $m_0 = np \approx 33,333$, $\sigma = \sqrt{npq} \approx 4.714$.*

На следующем графике изображены графики плотности распределения случайной величины μ и многоугольник ряда распределения случайной величины λ:



Теорема 1 (Локальная теорема Муавра – Лапласа). *Если* $k \approx pn$, *то* $p_n(k) \approx \frac{1}{\sqrt{npq}}\varphi(X)$, *где*

a) $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$,

b) $X = \frac{k - np}{\sqrt{npq}}$

Чем меньше $\frac{|k - np|}{(npq)^{3/2}}$, *тем точнее приближённое это равенство.*

Теорема 2 (Интегральная теорема Муавра – Лапласа). *Если* $0 \leqslant k_0 \leqslant k_1 \leqslant n$, *то* $P(k_0 \leqslant k \leqslant k_1) \approx \Phi(X_1) - \Phi(X_0)$, *где*

a) $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int\limits_0^x e^{-t^2/2}\,dt$,

b) $X_0 = \frac{k_0 - np}{\sqrt{npq}}$, $X_1 = \frac{k_1 - np}{\sqrt{npq}}$.

Пример 2. *Пусть* λ *— биномиально распределённая случайная величина, определяемая пераметрами* $n = 1000$, $p = 1/4$, $q = 3/4$. *Вычислите (приближённо) вероятности следующих событий:*

i) значение случайной величины λ *равно* 251,

ii) значение случайной величины λ *не меньше* 240 *и не больше* 262.

По условию $np = 250$, $\sqrt{npq} = \sqrt{\frac{375}{2}} \approx 13.693$. *Применение точных формул дают следующие значения искомых веротностей:*

$$p_{1000}(251) \approx 0,02900807359(\text{ если применить формулу (1)}),$$
$$P(240 \leqslant \lambda \leqslant 262) \approx 0.5972702962(\text{ если применить формулу (2)})$$

С другой строны, если применить теоремы Муавра–Лапласа, то *i)* $X = \dfrac{251 - np}{\sqrt{npq}} \approx 0,073029674$, $p_{1000}(251) \approx \dfrac{1}{\sqrt{npq}}\varphi(X) \approx 0,029057035$;

ii) $X_0 = \dfrac{240 - np}{\sqrt{npq}} \approx -0,730296743$, $X_1 = \dfrac{262 - np}{\sqrt{npq}} \approx 0,876356092$, $P(240 \leqslant \lambda \leqslant 262) \approx 0,5769773505$.

Предположим, что в серии из n экспериментов количество благоприятных исходов для события A равно k. Числа m и $\dfrac{m}{n}$ называют соответственно частотой и относительной частой появления события A в рассматриваемой серии.

Вероятность того, что относительная частота появления события A отличается от p не более, чем на $\varepsilon > 0$, обозначим через $P\left(\left|\dfrac{m}{n} - p\right| \leqslant \varepsilon\right)$. Для этой вероятности можно написать точное равенство (которое не целесообразно применять для больших значений n и k) и приближённое равенство:

$$P\left(\left|\frac{m}{n} - p\right| \leqslant \varepsilon\right) = \sum_{m: p - n\varepsilon \leqslant m \leqslant p + n\varepsilon} C_n^m p^m q^{n-m}; \qquad (3)$$

$$P\left(\left|\frac{m}{n} - p\right| \leqslant \varepsilon\right) \approx 2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right). \qquad (4)$$

Из равенства (4) получается приближнная формула для вычисления вероятности $P\left(|m - pn| \leqslant n\varepsilon\right)$ того, что частота появления события A в серии из n экспериметов от личается от теретического среднего (равного в рассматриваемом случае np) не более чем на $n\varepsilon$:

$$P\left(|m - pn| \leqslant n\varepsilon\right) \approx 2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right). \qquad (5)$$

Пример 3. *Используя приближённую формула (4), решим следующую задачу:*

> *Вероятность благоприятного исхода события в каждом из* 1000 *экспериментов равна* $p = \dfrac{9}{10}$. *Найдите вероятность того, что относительная частота появления этого событий в серии отличается от* p *не более чем на* $\dfrac{1}{100}$.

В решаемой задаче $n = 1000$, $p = \dfrac{9}{10}$, $q = \dfrac{1}{10}$, $\varepsilon = \dfrac{1}{100}$. *Поэтому*

- $X = \sqrt{\dfrac{n}{pq}} \approx 105,4092553,$

- $\varepsilon X \approx 1,054092553,$

- $\Phi(\varepsilon X) \approx 0,3540797273.$

Следовательно, искомая вероятность равна

$$P\left(\left|\frac{m}{n} - p\right| \leqslant \varepsilon\right) \approx 2\Phi(\varepsilon X) \approx 0,7081594546.$$

Пример 4. *Пусть вероятность благоприятного исхода события в каждом из $n = 1000$ экспериментов равна $p = \dfrac{9}{10}$. Найдите вероятность того, что частота появления этого события отличается от среднего значения не более чем на 5.*

Решение задачи разбивается на несколько шагов:

- *число* $\varepsilon = \dfrac{1}{200}$ *находим из уравнения* $1000 \cdot \varepsilon = 5.$

- $X = \varepsilon\sqrt{\dfrac{n}{pq}} \approx 0,5270462765$

- $\Phi(X) \approx 0,2009192736$

Поэтому $P\left(|m - pn| \leqslant n\varepsilon\right) \approx 0,4018385472.$

2 Неравенства Чебышева

Первое неравенство Чебышева. Пусть λ — случайная величина, принимающя только неотрицательные значения, и математическое ожидание её конечно. Тогда $P(\lambda \geqslant \varepsilon) \leqslant \dfrac{\mathbf{M}\lambda}{\varepsilon}$ для произвольного $\varepsilon > 0$.

Пример 5. *Пусть случайная величина β равна скорости ветра, измеренной на высоте 100 над уровнем моря. Известно, что среднее значения этой случайной величины равна 5 м/сек. Оцените вероятность того, что скорость ветра на этой высоте, измеренная в случайный момент времени, более 30 м/сек.*

По условию $\mathbf{M}\beta = 5$, $\varepsilon = 30$. *Искомая оценка следует из первого неравенств Чебышева:* $P(\beta \geqslant 30) \leqslant \dfrac{5}{30} = \dfrac{1}{6}.$

Второе неравенство Чебышева Пусть λ — случайная величина с конечной дисперсией. Неравенство $P(|\lambda - \mathbf{M}\lambda| \geqslant \varepsilon) \leqslant \dfrac{\mathbf{D}\lambda}{\varepsilon^2}$ верно для всех $\varepsilon > 0$.

Пример 6. *Предположим, что дисперсия случайной величины γ равна 100. Найдите вероятность того, что значение этой случайной величины отклоняется от своего среднего значения больше, чем на 20.*

По условию $\mathbf{D}\gamma = 100$, $\varepsilon = 20$. Из второго неравенства Чебышева следует, что $P(|\gamma - \mathbf{M}\gamma| \geqslant 20) \leqslant \dfrac{100}{20^2} = \dfrac{1}{4}$.

Математическая статистика

2 Числовые характеристики экспериментальных данных

1 Выборки

Пусть λ — случайная величина. Предположим, что проводится эксперимент, результатом которого должно быть значение этой случайной величины. Совокупность всевозможных результатов проведения этого эксперимента называют **генеральной выборкой** или **генеральной совокупностью**.

Если эксперимент проводится конечное число раз, то полученная совокупность значений случайной величины называют **выборкой**, а количество проведённых экспериментов — **объёмом выборки** (который обозначим через n).

Значения случайной величины, которые получаются в результате проведённого эксперимента, называются **вариантами**.

Число повторений отдельной варианты называют **частотой**, если x_j — j–ая по порядку варианта, то её частота равна m_j. **Относительной частотой** варианты x_j называют число $p_j^* = \dfrac{m_j}{n}$. Очевидно, что $\sum\limits_j n_j = n$, $\sum\limits_j \dfrac{n_j}{n} = 1$. Предполагается, что в предыдущих выражениях суммирование ведётся по множеству всех вариант.

Дискретным вариационным рядом (или статистическим рядом или статистическим распределением) будем называть таблицу, включающую в себя перечень всех вариант, попадающих в выборку (выписанных в порядке возрастания), и соответствующих им частот и относительных частот (обычно в эту таблицу помещаю только одну из этих характеристик):

X	x_1	x_2	$\dots$	x_k
	n_1	n_2	$\dots$	n_k
	p_1^*	p_2^*	$\dots$	p_k^*

(6)

Некоторыми авторами (например, [4], §6) вариационным рядом называют последовательность вариант, записанную в возрастающем порядке.

Пример 7. *Генератор случайных (точнее, псевдослучайных) чисел - алгоритм, порождающий последовательность чисел, элементы которой почти независимы друг от друга и подчиняются заданному распределению (обычно равномерному).*

Предположим, что в результате работы генератора случайных чисел была получен набор чисел:

$$1, 8, 2, 3, 10, 4, 6, 8, 1, 6, 7, 2, 3, 5, 4, 2, 8, 7, 6, 4, 5, 10, 7, 3, 5, 9, 5$$
$$6, 3, 8, 1, 1, 2, 2, 6, 4, 9, 6, 8, 7, 6, 8, 2, 10, 4, 10, 3, 1, 8, 9, 1, 9, 6,$$
$$4, 9, 3, 3, 4, 8, 6, 6, 1, 3, 8, 9, 8, 8, 4, 8, 3, 1, 1, 6, 2, 10, 5, 6, 6, 3, 2, 4,$$
$$4, 8, 2, 2, 3, 4, 1, 7, 4, 6, 5, 5, 10, 6, 10, 10, 5, 9, 2.$$

Объём полученной выборки равен $n = 100$. *Дискретный вариационный ряд этой выборки имеет вид*

X	1	2	3	4	5	6	7	8	9	10
	10	11	11	12	8	15	5	13	7	8
	$\frac{10}{100}$	$\frac{11}{100}$	$\frac{11}{100}$	$\frac{12}{100}$	$\frac{8}{100}$	$\frac{15}{100}$	$\frac{5}{100}$	$\frac{13}{100}$	$\frac{7}{100}$	$\frac{8}{100}$

(7)

Интервальным вариационным рядом называют таблицу

X	$[t_1, t_2]$	$[t_2, t_3]$	$\dots$	$[t_k, t_{k+1}]$
	n_1	n_2	$\dots$	n_k

(8)

в которой используются следующие обозначения:

1. $[t_1, t_2], \dots, [t_k, t_{k+1}]$ — отрезки, на которые разбит отрезок $[X_{min}, X_{max}]$,

2. X_{min}, X_{max} — соответственно наименьшая и наибольшая варианта, входящая в рассматриваемую выборку,

3. n_j — число элементов выборки, содержащихся в отрезке $[t_j, t_{j+1}]$,

Предполагается, что числа t_j не являются вариантами для рассматриваемой выборки.

Некоторые рекомендации относительно выбора числа k, содержатся в [6], п. 3.1.1.1. Сформулируем некоторые из них:

(1) $k \approx 3,7(n-1)^{2/5}$, если $n > 200$;

(2) $k \approx 1 + 3,32\lg(n)$;

(3) $k \leqslant n/5$;

(4) если p_j — теоретическая вероятность попадания рассматриваемой случайной величины на j-ый отрезок разбиения, $np_j \geqslant 5$. Для приближенного анализа статистических данных можно рассматривать разбиения, для которых неравенство $1 \leqslant np_j \leqslant 5$ выполняется не более чем $1/5$ всех отрезков.

Пример 8. *В ходе социологического опроса, в котором исследовалось различные аспекты потребления кофе, была получена таблица, содержащее количество кофе (в килограммах), которое опрашиваемые употребили в течении года*

$$\begin{gathered}
1,30; 0,67; 9,10; 5,80; 4,20; 0,73; 1,90; 0,87; 3,10; 2,40; 8,40; 0,91; 5,80;\\
6,90; 9,40; 5,40; 7,01; 5,20; 9,80; 0,13; 2,40; 9,60; 5,10; 9,40; 6,40; 4,80;\\
5,20; 1,50; 7,50; 3,40; 4,30; 8,90; 1,90; 4,80; 0,05; 6,10; 8,40; 2,90; 2,10;\\
8,90; 8,40; 4,10; 8,20; 8,60; 0,53; 2,60; 7,50; 6,50; 4,40; 7,20; 1,60; 9,40;\\
4,50; 8,80; 4,90; 3,40; 9,0; 6,20; 0,57; 7,40; 5,90; 5,60; 0,00; 0,57; 9,20;\\
1,60; 9,00; 2,00; 3,20; 5,50; 8,10; 3,01; 2,10; 1,80; 1,10; 9,80; 9,60; 7,30;\\
6,50; 9,30; 1,80; 2,50; 8,90; 6,01; 0,71; 3,70; 9,00; 7,20; 6,70; 6,20; 9,40;\\
4,40; 2,30; 1,60; 8,70; 3,30; 8,60; 9,40; 8,30; 10,00.
\end{gathered}$$

Если применить перечисленные выше рекомендации для выбора числа разбиений, получим следующие оценки числа k:

(1) $k \approx 3,7(100-1)^{2/5} \approx 24$, *но у нас* $n < 200$;

(2) $k \approx 1 + 3,32\lg(n) \approx 7.64$;

(3) $k \leqslant 20$.

Построим интервальный ряд распределения для построенной выборки. В рассматриваемом случае объём выборки $n = 100$, $X_{min} = 0$, $X_{max} = 10$, *отрезок* $[0, 10]$ *разбит на 8 отрезков одинаковой длины. Непосредственный подсчёт позволяет построить следующую таблицу*

X	[0;1,25]	[1,25;2,5]	[2,5;3,75]	[3,75;5]	[5;6,25]	[6,25;7,5]	[7,5;8,75]	[8,57;10]
	11	16	10	9	15	9	4	26

(9)

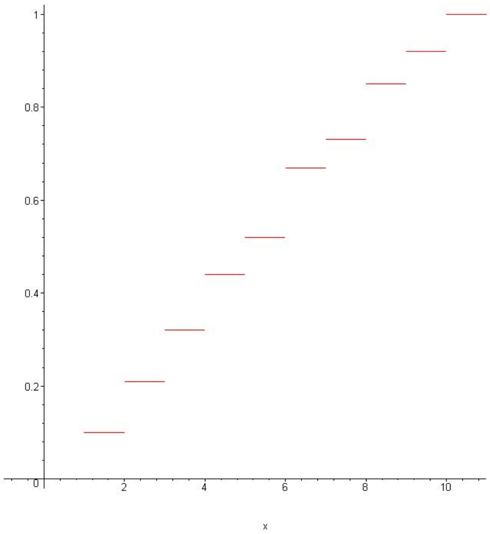
Заметим, что некоторые числа из выборки попадают в разные отрезки разбиения. Например, $7,5 \in [6,25; 7,5]$ *и* $7,5 \in [7,5; 8,75]$. *Для определённости буду учитываться попадания таких значений только в левый отрезок из двух, в которые это число попадает.*

2 Эмпирическая функция распределения

Эмпирической функцией распределения выборки X (которая задана вариационным рядом (6)) называют функцию F_X^*, определяемую равенством: $F_X^*(t) = \sum\limits_{j: X_j < t} p_j^*$, где суммирование ведётся по множеству индексов j таких, что $X_j < t$. Если множество таких индексов пусто, то $F_X(t) = 0$ по определению.

Пример 9. *Необходимо найти эмпирическую функцию распределения и построить её график для выборки, определённой вариационным рядом (7) из примера 7. Из определения эмпирической функции распределения следует, что*

$$F_X(t) = \begin{cases} 0 & , t \leqslant 1 \\ 10/100 & , 1 < t \leqslant 2 \\ 21/100 & , 2 < t \leqslant 3 \\ 32/100 & , 3 < t \leqslant 4 \\ 44/100 & , 4 < t \leqslant 5 \\ 52/100 & , 5 < t \leqslant 6 \\ 67/100 & , 6 < t \leqslant 7 \\ 72/100 & , 7 < t \leqslant 8 \\ 85/100 & , 8 < t \leqslant 9 \\ 92/100 & , 9 < t \leqslant 10 \\ 1 & , 10 < t \end{cases}$$



Эмпирическая функция распределения F^*_X, построенная по выборке X, удовлетворяет следующим условиям

1. значение этой функции принадлежат отрезку $[0, 1]$;

2. она неубывает;

3. она принимает постоянное значения на интервалах между последовательными вариантами;

4. если $X_{\min}$ и $X_{\max}$ соответственно наименьшая и наибольшая варианта, то $F^*_X(t) = 0$ для любого $x < X_{\min}$ и $F^*_X(t) = 1$ для любого $x > X_{\max}$.

3 Полигон

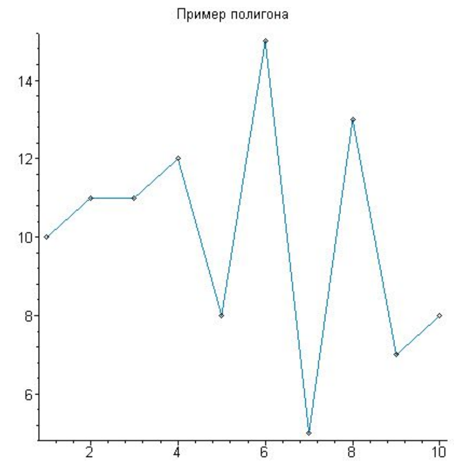
Предположим, что в плоскости фиксирована прямоугольная система координат, в которой на оси абсцисс будем откладывать варианты, а на оси - частоты (или относительные частоты).

Пусть дана выборка, определяемая своим дискретным вариационным рядом

X	x_1	x_2	$\ldots$	x_k
	n_1	n_2	$\ldots$	n_k
	p^*_1	p^*_2	$\ldots$	p^*_k

Полигоном этой выборки называют ломанную, соединяющую последовательно точки (x_1, n_1), (x_2, n_2),. . . , (x_k, n_k) (или (x_1, p^*_1), (x_2, p^*_2),. . . , (x_k, p^*_k)).

Пример 10. *Пусть дан дискретный вариационный ряд (7). Постоим её полигон:*



4 Гистограмма

Пусть для выборки дан её интервальный вариационный ряд:

X	$[t_1, t_2]$	$[t_2, t_3]$	$\ldots$	$[t_k, t_{k+1}]$
	n_1	n_2	$\ldots$	n_k

Предположим, что в плоскости фиксирована прямоугольная система координат, в которой на оси абцисс будем откладывать варианты, а на оси - частоты.

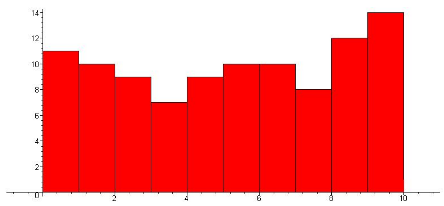
Гистограмма выборки X (которая задана вариационным рядом (8)) состоит из последовательности прямоугольников, которые строятся следующим образом: основание j–ого по порядку прямоугольника является отрезок $[t_j, t_{j+1}]$, высота этого прямоугольника равна частоте j-ой варианты или её относительной частоте.

Пример 11. *Постройте гистограмму выборки, для которой дан интервальный вариационный ряд*

X	$[0,1]$	$[1,2]$	$[2,3]$	$[3,4]$	$[4,5]$	$[5,6]$	$[6,7]$	$[7,8]$	$[8,9]$	$[9,10]$
	11	10	9	7	9	10	10	8	12	14

(10)

Из определения интервального вариационного ряда следует, что



3 Точечные оценки

Пусть для выборки X задан объём n и дискретный вариационный ряд

X	x_1	x_2	$\dots$	x_k
	n_1	n_2	$\dots$	n_k
	p_1^*	p_2^*	$\dots$	p_k^*

Выборочная средняя $\overline{X}$ этой выборки определяется равенством

$$\overline{X} = \frac{1}{n}\sum_{j=1}^{k} x_j n_j = \sum_{j=1}^{k} x_j p_j^*.$$

Выборочной дисперсией $d_{\text{в}} = d_{\text{в}}X$ выборки $\overline{X}$ называют число

$$d_{\text{в}} = d_{\text{в}}X = \frac{1}{n}\sum_{j=1}^{k}(x_j - \overline{X})^2 n_j = \sum_{j=1}^{k}(x_j - \overline{X})^2 p_j^* = \overline{X^2} - (\overline{X})^2.$$

Для вычисления числа $\overline{X^2}$ можно воспользоваться таблицей

X	x_1^2	x_2^2	$\dots$	x_k^2
	n_1	n_2	$\dots$	n_k
	p_1^*	p_2^*	$\dots$	p_k^*

которая в общем случае не будет дискретым вариационным рядом. Например среди чисел x_j^2 могут встречаться одинаковые.

Наконец, **выборочным средне квадратичным отклонением** $\sigma_{\text{в}}$ и **исправленным выборочным средне квадратичным отклонением** s называют соответственно выражения $\sigma_{\text{в}} = \sqrt{d_{\text{в}}}$ и $s = \sqrt{\dfrac{n}{n-1}}\sigma_{\text{в}}$.

Пример 12. *Вычислите точечные оценки выборки, построенной в примере 1. В этом примере построен дискретный вариационный ряд этой выборки:*

X	1	2	3	4	5	6	7	8	9	10
	10	11	11	12	8	15	5	13	7	8
	$\frac{10}{100}$	$\frac{11}{100}$	$\frac{11}{100}$	$\frac{12}{100}$	$\frac{8}{100}$	$\frac{15}{100}$	$\frac{5}{100}$	$\frac{13}{100}$	$\frac{7}{100}$	$\frac{8}{100}$

Из определения выборочной средней следует, что

$$\overline{X} = \frac{1}{100}(1 \cdot 10 + 2 \cdot 11 + 3 \cdot 11 + 4 \cdot 12 + 5 \cdot 8 + \\ + 6 \cdot 15 + 7 \cdot 5 + 8 \cdot 13 + 9 \cdot 7 + 10 \cdot 8) = \frac{21}{4};$$

$$\overline{X^2} = \frac{1}{100}(1^2 \cdot 10 + 2^2 \cdot 11 + 3^2 \cdot 11 + 4^2 \cdot 12 + 5^2 \cdot 8 + \\ + 6^2 \cdot 15 + 7^2 \cdot 5 + 8^2 \cdot 13 + 9^2 \cdot 7 + 10^2 \cdot 8) = \frac{3529}{100};$$

$$d_в = \frac{3529}{100} - \left(\frac{21}{4}\right)^2 = \frac{26027}{80} \approx 325.3375,$$

$$\sigma_в \approx 18.0371; s = \sqrt{\frac{100}{99}} \cdot \frac{26027}{80} \approx 18.1270.$$

4 Интервальные оценки

1. Пусть выборка X построена по случайной величине λ. Предположим, что для числовой характеристики τ этой случайной величины (которую будем называть теоретической) существует аналогичная характеристика выборки τ^* (называемую выборочной характеристикой). Например, теоретической средней $\overline{\lambda}$ случайной величины соответствует выборочная средняя $\overline{X}$.

Пусть дано число $\gamma \in (0, 1)$ — уровень доверия. Доверительным интервалом для теоретической характеристики называют интервал (α, β) называют любой интервал, для которого выполняются условия:

1. числа α, β — определяются по случайной величине и выборке (следовательно, сами являются случайными величинами, определёнными на подходящем вероятностном пространстве);

2. $\tau^* \in (\alpha, \beta)$,

3. $P(\tau \in (\alpha, \beta)) = \gamma$, т.е. вероятность принадлежности числа τ интервалу (α, β) равна γ.

Ниже будут построены доверительные интервалы для теоретической средней и теоретического средне квадратичного отклонения нормально распределённой случайной величины.

2. Если известно среднеквадратичное отклонение σ_λ рассматриваемой случайной величины, то доверительный интервал для её математического ожидания находится из условия $P(|\overline{X_{\text{в}}} - m| < \delta) = 2\Phi\,(z)$, где $z = \dfrac{\delta}{\sigma_\lambda}\sqrt{n}$. Поэтому, если γ — уровень доверия, то соответствующий ему доверительный интервал $\left(\overline{X_{\text{в}}} - z\dfrac{\sigma_\lambda}{\sqrt{n}}, \overline{X_{\text{в}}} + z\dfrac{\sigma_\lambda}{\sqrt{n}}\right)$ определяется из уравнения $\Phi(z) = \gamma/2$.

Пример 13. *Пусть X — выборка, построенная по нормально распределённой случайной величине. Известно, что объём выборки равен $n = 100$, выборочным средним $\overline{X} = 5$ и средне квадратично отклонение равно $\sigma = 5$. Найдите доверительные интервал для математического ожидания этой случайной величины с уровнем доверия $\gamma = 0,99$.*

При сделанных предположения доверительный интервал для математического отклонения (с заданным уровнем значимости) имеет вид

$$\left(\overline{X_{\textit{в}}} - z\frac{\sigma_\lambda}{\sqrt{n}}, \overline{X_{\textit{в}}} + z\frac{\sigma_\lambda}{\sqrt{n}}\right),$$

где число $z = 2,585$ найдено из уравнения $\Phi(z) = \dfrac{0,99}{2} = 0,495$ при помощи таблицы из приложения 1. Поэтому искомый доверительный интервал равен $(3,706; 6,292)$.

3. Предположим, что для данной случайной величины дана только выборка, для вычислены объём n, выборочное среднее $\overline{x}$ и несмещённое выборочное средне квадратичное s. Доверительный интервал для теоретической средней $\overline{\mu}$ рассматриваемой случайной величины имеет вид ([6], п. 2.1.1.2.2):

$$\left(\overline{x} - t_\gamma\frac{s}{\sqrt{n}}, \overline{x} + t_\gamma\frac{s}{\sqrt{n}}\right),$$

где t_γ — γ–квантиль случайной величины, имеющей распределения Стьюдента с $l = n - 1$ степенью свободы. Другими словами, число t_γ находится из условия $P(T < t_\gamma) = \gamma$. Квантили распределения Стьюдента с l степенями свободы можно вычислять, используя таблицы или следующие приближённые формулы:

$$t_\gamma \approx u_\gamma\left(1 - \frac{u_\gamma + 1}{4l}\right)^{-1}, \;\; u_\gamma \approx 4,91[\gamma^{0,14} - (1-\gamma)^{0,14}]; \tag{11}$$

$$t_{0,975} \approx 1,96 + \frac{2,5}{l - 0,8}. \tag{12}$$

Пример 14. *Пусть X — выборка, построенная по нормально распределённой случайной величине λ. Известно, что $\overline{\lambda} = 5$, объём выборки равен $n = 100$ и несмещённое средне квадратично отклонение равно $s = 1$. Найдите доверительные интервал для среднего значения этой случайной величины с уровонями доверия $\gamma = 0,99$ и $\gamma = 0,975$.*

Используя формулы (11)–(12) находим квантили $t_{0,99}$ и $t_{0,975}$:

- $u_{0,99} \approx 4,91[0,99^{0,14} - (1-0,99)^{0,14}] \approx 2,326291628$;
- $t_{0,99} \approx 2,326291628 \left(1 - \dfrac{2,326291628+1}{4 \cdot 100}\right)^{-1} \approx 2.345798654$;
- $t_{0,975} \approx 1,96 + \dfrac{2,5}{100-0,8} \approx 1.985201613$

Поэтому доверительные интервалы для рассматриваемой случайной величины имеют следующий вид:

- $(4.765420135; 5.234579865)$ *для уровня значимости* $\gamma = 0,99$;
- $(4.801479839; 5.198520161)$ *для уровня значимости* $\gamma = 0,975$;

4. Доверительный интервал для средне квадратичного отклонения σ_λ случайной величины λ с заданным уровнем доверия γ имеет вид

$$(\frac{s}{1+q}, \frac{s}{1-q}),$$

где

- s — несмещённое выборочное среднеквадратичное отклонение,
- q — находится по таблице (см. Приложение 2) по заданному объёму выборки n и уровню значимости γ.

Пример 15. *Пусть X — выборка, построенная по нормально распределённой случайной величине. Известно, что объём выборки равен $n = 100$ и несмещённое средне квадратично отклонение равно $s = 5$. Найдите доверительные интервал для средне квадратичного отклонения этой случайной величины с уровнем доверия $\gamma = 0,99$.*

Из таблицы (приложение 2) по заданному объёму выборки и уровню значимости находим значение параметра q: $q = 0,198$. Поэтому доверительный интервал для средне квадратичного отклонения данной случайной величины имеет вид $(\frac{5}{1+0,198}; \frac{5}{1-0,198}) = (4,17; 6,23)$.

Приложение 1: Таблица значений функции $\varphi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-x^2/2}$

В таблице указаны десятичные знаки значения функции ϕ. Например, если $x = 0,12$, то в таблице находим число 04776. Поэтому $\varphi(x) = 0,04776$.

Если $x > 0$, то $\varphi(x) \approx 0$, функция φ — чётная

	0	1	2	3	4	5	6	7	8	9
0,0	3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3725	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3208	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1022	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0668
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0378	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0204	0198	0194	0189	0184	0180
2,5	0176	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0089	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0059	0058	0056	0055	0053	0051	0045	0048	0047	0046
3,0	0044	0043	0042	0040	0039	0038	0037	0036	0035	0034

Приложение 2: Нормированная функция Лапласса

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2}\, dt$$

В таблице указаны десятичные знаки значения функции Лапласса. Например, если $x = 0,12$, то в таблице находим число 04776. Поэтому $\Phi(x) = 0,04776$.

x	0	1	2	3	4	5	6	7	8	9
0,0	00000	00399	00789	01197	01595	01994	02392	02790	03188	03586
0,1	03983	04380	04776	05172	05567	05962	06356	06749	07142	07535
0,2	07926	08317	08706	09095	09483	09871	10257	10642	11026	11409
0,3	11791	12172	12552	12930	13307	13683	14058	14431	14803	15173
0,4	15542	15910	16276	16640	17003	17364	17724	18082	18439	18793
0,5	19146	19497	19847	20194	20540	20884	21226	21566	21904	22240
0,6	22575	22907	23237	23565	23891	24215	24537	24857	25175	25490
0,7	25804	26115	26424	26730	27035	27337	27637	27935	28230	28524
0,8	28814	29103	29389	29673	29955	30234	30511	30785	31057	31327
0,9	31594	31859	32121	32381	32639	32894	33147	33398	33646	33891
1,0	34134	34375	34614	34850	35083	35314	35543	35769	35993	36214
1,1	36433	36650	36864	37076	37286	37493	37698	38000	38100	38298
1,2	38493	38686	38877	39065	39251	39435	39617	39796	39973	40147
1,3	40320	40490	40658	40824	40988	41149	41308	41466	41621	41774
1,4	41924	42073	42220	42364	42507	42647	42786	42922	43056	43189
1,5	43319	43448	43574	43699	43822	43943	44062	44179	44295	44408
1,6	44520	44630	44738	44845	44950	45053	45154	45254	45352	45449
1,7	45543	45637	45728	45818	45907	45994	46080	46164	46246	46327
1,8	46407	46485	46562	46638	46712	46784	46856	46926	46997	47062
1,9	47128	47193	47257	47320	47381	47441	47500	47558	47615	47670
2,0	47725	47778	47831	47822	47932	47982	48030	48077	48124	48169
2,1	48214	48257	48300	48341	48382	48422	48461	48500	48537	48574
2,2	48610	48645	48679	48713	48745	48778	48809	48840	48870	48899
2,3	48928	48956	48983	49010	49036	49061	49086	49111	49134	49158
2,4	49180	49202	49224	49245	49266	49286	49305	49324	49343	49361
2,5	49379	49396	49413	49430	49446	49461	49477	49492	49506	49520
2,6	49534	49547	49560	49573	49585	49598	49609	49621	49632	49643
2,7	49653	49664	49674	49683	49693	49702	49711	49720	49728	49736
2,8	49744	49752	49760	49767	49774	49781	49788	49795	49801	49807

x	0	1	2	3	4	5	6	7	8	9
2,9	49813	49819	49825	49831	49836	49841	49846	49851	49856	49861
3,0	49865	49869	49874	49878	49882	49886	49889	49893	49896	49900
3,1	49903	49906	49910	49913	49916	49918	49921	49924	49926	49929
3,2	49931	49934	49936	49938	49940	49942	49944	49946	49948	49950
3,3	49952	49953	49955	49957	49958	49960	49961	49962	49964	49965
3,4	49966	49968	49969	49970	49971	49972	49973	49974	49975	49976
3,5	49977	49978	49978	49979	49980	49981	49981	49982	49983	49983
3,6	49984	49985	49985	49986	49986	49987	49987	49988	49988	49989
3,7	49989	49990	49990	49990	49991	49991	49992	49992	49992	49992
3,8	49993	49993	49993	49994	49994	49994	49994	49995	49995	49995
3,9	49995	49995	49996	49996	49996	49996	49996	49996	49997	49997
4,0	49997									
5,0	49999									

Приложение 3: Значение чисел q в зависимости от объёма выборки n и уровня доверия γ для определения доверительного интервала средне квадратичного отклонения σ_λ.

	γ				γ		
n	0,95	0,99	0,999	n	0,95	0,99	0,999
7	0,92	–	–	25	0,32	0,49	0,73
8	0,80	–	–	30	0,28	0,43	0,63
9	0,71	–	–	35	0,26	0,38	0,56
10	0,65	–	–	40	0,24	0,35	0,50
11	0,59	0,98	–	45	0,22	0,32	0,46
12	0,55	0,90	–	50	0,21	0,30	0,43
13	0,52	0,83	–	60	0,188	0,269	0,38
14	0,48	0,78	–	70	0,174	0,245	0,34
15	0,46	0,73	–	80	0,161	0,226	0,31
16	0,44	0,70	–	90	0,151	0,211	0,29
17	0,42	0,66	–	100	0,143	0,198	0,27
18	0,40	0,63	0,96	150	0,115	0,160	0,211
19	0,39	0,60	0,92	200	0,099	0,136	0,185
20	0,37	0,58	0,88	250	0,089	0,120	0,162

5 Задачи

1 Непрерывные случайные величины

1. *Пусть функция ρ определяется равенством*

$$\rho(x) = \begin{cases} 0, & x \leqslant 1 \\ \alpha x, & x \in (1, 2] \\ 0, & x > 2 \end{cases}.$$

При каких условия она будет плотностью распределения непрерывной случайной величины ζ.

1. *Постройте графики плотности распределения и функции распределения случайной величины ζ.*
2. *Найдите математическое ожидание и дисперсию этой случайной величины.*
3. *Найдите вероятность того, значение построенной случайной величины принадлежит отрезку $[3/2, 5/2]$.*

2. *Пусть функция ρ определяется равенством*

$$\rho(x) = \begin{cases} 0, & x \leqslant 1 \\ \alpha, & x \in (1, 2] \\ 0, & 2 < x \leqslant 4 \\ \beta, & x \in (4, 7] \\ 0, & x > 7 \end{cases}$$

При каких условия эта функция будет плотностью распределения непрерывной случайной величины ξ так, что $\mathbf{M}\xi = 4$.

1) *Постройте графики плотности распределения и функции распределения случайной величины ξ.*
2) *Найдите дисперсию этой случайной величины.*
3) *Найдите вероятность того, значение построенной случайной величины принадлежит отрезку $[3/2, 9/2]$.*

3. *Пусть функция ρ определяется равенством*

$$\rho(x) = \begin{cases} 0, & x \leqslant -1 \\ \alpha(1 - x^2), & x \in (-1, 1] \\ 0, & 1 < x \end{cases}.$$

При каких условия она будет плотностью распределения непрерывной случайной величины ψ. Постройте графики плотности распределения и функции распределения случайной величины ψ. Найдите математическое ожидание и дисперсию этой случайной величины.

4. *Пусть $\lambda > 0$. Покажите, что функция ρ_τ, определяемая равенством*

$$\rho_\tau(x) = \begin{cases} 0, & x \leqslant 0 \\ \lambda e^{-\lambda x}, & x > 0 \end{cases} \tag{13}$$

является плотностью распределения непрерывной случайной величины τ.

1) *Постройте графики плотности распределения и функции распределения случайной величины τ.*

2) *Найдите математическое ожидание и дисперсию этой случайной величины.*

Замечание 1. *Если плотность распределения случайной величины определяется равенством (13), то говорят, что эта случайная величина имеет экспоненциальное (или показательное) распределение.*

5. *Пусть $n = 1, 2, 3$ и $\beta > 0$ — фиксированные числа и функция ρ определяется равенством*

$$\rho(x) = \begin{cases} 0, & x \leqslant 0 \\ \lambda x^n e^{-x/\beta}, & x > 0 \end{cases} \tag{14}$$

При каких условия эта функция будет плотностью распределения непрерывной случайной величины ξ.

1) *Постройте графики плотности распределения и функции распределения случайной величины ξ.*

2) *Найдите математическое ожидание и дисперсию этой случайной величины.*

Замечание 2. *1. Если плотность распределения случайной величины определяется равенством (14), то говорят, что эта случайная величина имеет γ – распределение.*

2. Пусть $n = r-1$. Если время работы прибора (до первой поломки) имеет экспоненциальное распределение, то время, прошедшее до r-ого отказа этого прибора есть случайная величина, имеющая γ-распределение.

3. Для решении первой части задачи необходимо неоднократно применить формулу интегрирования по частям для определённого интеграла

6. *Найдите значение параметра α, для которого функция $F(x) = \frac{1}{2} + \alpha \operatorname{arctg}(X)$ будет функцией распределения некоторой случайной величины. Найдите вероятность того, что значение этой случайной величины принадлежит интервалу $(0, 1)$.*

7. *Пусть $\alpha > 0$ и функция F определяется равенством*

$$F(x) = \begin{cases} 0, & x \leqslant \alpha \\ \beta(x-\alpha), & \alpha \leqslant x < \alpha + 1 \\ 1, & 1+\alpha \leqslant x \end{cases}.$$

При каких значениях параметра β эта функция будет плотностью распределения непрерывной случайной величины. найдите вероятность того, что значение этой случайной величины удовлетворяет следующим значениям:

- *оно меньше* 0,
- *оно меньше* $1/2$,
- *оно больше* $1/3$ *и меньше* $2/3$,
- *оно больше* $1/5$,
- *оно больше* 2.

8. *Пусть функция случайной величины имеет вид*

$$F(x) = \begin{cases} 0, & x \leqslant 0 \\ 2x, & 0 \leqslant x < 1 \\ 1, & 1 \leqslant x \end{cases}.$$

Найдите вероятность того, что в серии из 4 независимых испытаний значение этой случайной величины 2 раза попадет на отрезок $[1/4, 3/4]$.

9. *Пусть функция случайной величины имеет вид*

$$F(x) = \begin{cases} 0, & x \leqslant 0 \\ 1 - e^{-x}, & 0 \leqslant x \end{cases}.$$

Пусть, кроме того, задано число $p_0 \in (0, 1)$. Найдите число X так, что рассматриваемая случайная величина принимала значение, большее чем X, с вероятностью p_0.

10. *Пусть случайная величина имеет равномерное распределение на интервале* $(-2, 5)$.

- *Найдите вероятность того, что значение этой случайной величины попадает на интервал* $(0, 3)$.
- *Какое из двух событий имеет большую вероятность: попадание этой случайной величины на отрезок* $[-1, 3]$ *или на отрезок* $[4, 6]$.

11. *Пусть случайная величина имеет равномерное распределение на интервале* $(2, 5)$. *Найдите вероятность того, что в серии из* 6 *независимых испытаний значение этой случайной величины не менее трёх раз раза попадет на отрезок* $[3, 4]$.

12. *Пусть случайная величина имеет равномерное распределение на интервале* (a, b). *Найдите это интервал из условия, что математическое ожидание и среднеквадратичное отклонение этой случайной величины равны соответственно заданным числам* m_0 *и* s_0.

2 Нормальное распределение

1. *Пусть случайная величина* λ *имеет нормальное распределение и* $\mathbf{M}\lambda = 2$, $\sigma_\lambda = 3$. *Найдите плотность распределения этой случайной величины.*

2. *Пусть плотность распределения нормально распределённой случайной величины* λ *имеет вид* $\rho_\lambda(x) = \dfrac{1}{3\sqrt{2\pi}} e^{-(x-4)^2/18}$.

3. *Пусть случайная величин* α *имеет нормальное распределение и* $\mathbf{M}\alpha = 1$, $\mathbf{D}\alpha = 4$.

1) Найдите вероятность того, что значение этой случайной величины принадлежит отрезку $[0, 2]$.

2) Найдите положительное число t *такое, что вероятность события* $\lambda > t$ *равна* $1/4$.

3) Найдите отрезок, симметричный относительно математического значения этой случайной величины (т.е. отрезок вида $[\mathbf{M}\alpha - t, \mathbf{M}a + t]$*) такой, что значение данной случайной величины принадлежит этому отрезку с вероятностью* $0,1234$.

4. *Пусть случайная величин* α *имеет нормальное распределение и* $\mathbf{M}\alpha = -1$, $\mathbf{D}\alpha = 16$.

1) Найдите вероятность того, что значение этой случайной величины принадлежит отрезку $[1, 2]$.

2) Найдите вероятность того, что $\alpha > 0$.

3) Найдите отрезок, симметричный относительно математического значения этой случайной величины (т.е. отрезок вида $[\mathbf{M}\alpha - t, \mathbf{M}\alpha + t]$*) такой, что значение данной случайной величины принадлежит этому отрезку с вероятностью* $0,1234$.

5. *Пусть случайная величин* α *имеет нормальное распределение и* $\mathbf{M} = 5$, $\mathbf{D} = 4$. *Какое из событий*

A: значение случайной величины α *принадлежит отрезку* $[1, 3]$*;*

B: значение случайной величины α *принадлежит отрезку* $[7, 8]$*;*

имеет большую вероятность.

6. *Пусть значение случайной величины* α *равно весу сахара, потребляемого случайным образом выбранного человека в течении суток, имеет нормальное распределение и* $\mathbf{M}\alpha = 50$ *грамм. Известно, что с вероятностью* $9/10$ *суточное потребление сахара колеблется от* 44 *до* 56 *грамм. Найдите средне квадратичное отклонение этой случайной величины.*

7. *Пусть случайная величина* β *равна длине ступни, измеренная у случайным образом выбранного человека. Известно, что эта случайная величина имеет нормальное распределение и* $\mathbf{M}\alpha = 300$ *миллиметров,* $\sigma_\alpha = 20$ *миллиметров. Какое из двух событий более вероятно:*

- *длина ступни не менее* 250 *и не более* 260 *миллиметров;*
- *длина ступни не менее* 320 *и не более* 360 *миллиметров;*

8. *Случайная величина измерения имеет нормальное распределения имеет математическое ожидание и средне квадратичное отклонение, равные соответственно* $m = 0$ *мм и* $s = 11$. *Найдите вероятность того, что при трёх независимых измерениях ошибка хотя бы одного по абсолютной величине не более* 4 *мм.*

9. *Случайная величина* λ *равна погрешности измерения длины шнурка для турель. Известно, что эта случайная величина имеет нормальное распределение и* $\mathbf{M}\alpha = 0$ *миллиметров,* $\sigma_\alpha = 5$ *миллиметров. Найдите вероятности того, что дли двух случайных шнурков погрешность измерения их длин по абсолютной величине не более* 2 *миллиметров.*

10. *Известно, что длина детали имеет нормально распределение (со средним значением* $x_0 = 100$ *миллиметров) и вероятность того, что длина детали отличается от среднего не более, чем на* 20 *миллиметров, равна* $0,49795$. *Найдите*

1. *средне квадратичное отклонение этой случайной величины,*

2. *вероятность того, что значение этой случайной величины попадает на отрезок* $[50, 60]$.

11. *Предположим, что результаты измерения длины и ширины прямоугольного земляного участка есть независимые случайные величины* ξ *и* μ, *имеющие нормальное распределение. Известны, средние значения этих случайных величин равны нулю, а средне квадратичные отклонения — соответственно* 3 *и* 5. *Найдите вероятность того, что вычисленная площадь земельного участка отличается от реального не более, чем на* 10 *квадратных метров.*

3 Предельные теоремы Муавра–Лапласа и Пуассона

1. *Вероятность попадания в мишень при одном выстреле равна* $19/20$. *найдите вероятность того, что после* 200 *выстрелов будет* 50 *попаданий.*

2. *Найдите вероятность того, что после* 1000 *бросаний монеты "герб" выпадет* 500 *раз.*

3. *Найдите вероятность того, что после* 1000 *бросаний монеты количество выпадений "герба" больше числа выпадений "решки больше на* 100.

4. *Вероятность благоприятного исхода в каждом из* 225 *испытаний равна* $7/10$. *Найдите вероятность того, что вероятность отклонения относительной частоты благоприятного исхода в этой серии отклоняется от среднего значения по абсолютной величине не более* $1/100$.

5. *Вероятность благоприятного исхода в каждом из* 300 *испытаний равна* $6/10$. *Найдите вероятность того, что вероятность отклонения частоты благоприятного исхода в этой серии отклоняется от среднего значения по абсолютной величине не более* 10.

6. *Пусть вероятность благоприятного исхода для эксперимента* A *в каждом из* 1000 *независимых экспериментов равна* $1/4$. *Найдите вероятность того, что*

1. *число благоприятных исходов равно* 256,

2. *число благоприятных исходов равно* 260,

3. *число благоприятных исходов не менее* 240 *и не более* 260.

7. *Пусть вероятность благоприятного исхода для эксперимента* A *в каждом из* 2000 *независимых экспериментов равна* 1/5. *Найдите вероятность того, что*

1. *число благоприятных исходов равно* 100,

2. *число благоприятных исходов больше* 260,

3. *число благоприятных исходов не менее* 200 *и не более* 300.

8. *Найдите вероятность того, что после* 1000 *бросаний монеты "герб" выпадет не менее* 977 *и не более раз* 1023 *раз.*

9. *Вероятность неблагоприятного исхода в каждом независимом испытании равна* 1/10. *Найдите наименьшее число проведённых экспериментов, достаточное для того, что бы с вероятностью* 99/100 *относительная частота благоприятного исхода эксперимента отличается от теоретического среднего не более чем* 1/100.

10. *В урне лежат шары:* 4 *белых и* 6 *чёрных. Из урны берутся* 3 *шара и возвращаются назад. Благоприятный исход проведённого эксперимента состоит в том, что среди выбранных шаров будет* 2 *белых шара. Найдите вероятность того, что в серии из* 1000 *проведённых экспериментов относительная частота благоприятных исходов будет отличаться от теоретического среднего больше, чем на* 1/2.

4 Выборки

1. *В результате проведения эксперимента была получена следующая выборка:*

$$\begin{gathered}3,4,2,8,0,7,5,2,9,1,5,6,6,2,7,6,0,0,8,4,\\0,1,7,0,8,1,8,7,9,0,2,6,5,0,3,3,0,5,3,6,\\7,9,8,9,7,0,8,0,7,6,3,4,3,8,2,6,8,2,9,5,\\1,8,0,8,6,4,3,2,8,8,3,4,9,5,2,4,1,5,6,9,\\7,0,8,0,2,2,3,6,6,1,5,3,9,4,3,0,9,3,0,3.\end{gathered}$$

Постройте дискретный вариационный ряд этой выборки.

2. *В результате проведения эксперимента была получена следующая выборка:*

$$\begin{gathered}1, 10, 17, 3, 16, 18, 5, 8, 1, 9, 15, 3, 17, 0, 14, 5, 16, 1, 7, 9,\\ 6, 18, 6, 9, 13, 1, 14, 6, 13, 2, 11, 13, 0, 11, 18, 13, 0, 4, 5, 11,\\ 15, 14, 9, 0, 2, 12, 13, 17, 9, 15, 1, 8, 3, 3, 15, 13, 9, 17, 17, 19,\\ 14, 8, 15, 15, 7, 11, 2, 17, 8, 2, 17, 18, 7, 7, 14, 5, 7, 2, 9, 12,\\ 4, 5, 5, 18, 18, 16, 15, 14, 3, 0, 19, 1, 9, 7, 13, 4, 11, 16, 19, 4.\end{gathered}$$

Постройте дискретный вариационный ряд этой выборки.

3. *В результате проведения эксперимента была получена следующая выборка:*

$$\begin{gathered}-9, 6, 5, -7, -6, 3, 9, -4, 0, -4, 0, 5, -5, 6, 6, -4, -8, -3, 1, -6,\\ 5, 6, -3, -5, -4, -10, -8, -10, 8, 5, -6, 9, 0, 9, -6, -4, -3, 6, 0, -2,\\ -10, -7, 6, 1, 9, 9, -3, 0, 0, -4, -9, -10, -10, 8, 7, -9, 7, 0, -2, -6,\\ 1, 8, -3, 5, 7, -10, 9, 3, 1, 7, -5, 7, 5, 7, -7, -1, 3, 3, -10, -10,\\ 6, 9, 6, -6, -3, -8, 4, 3, -6, 3, 7, -1, 4, -4, -2, 5, 9, -9, 2, -10.\end{gathered}$$

Постройте дискретный вариационный ряд этой выборки.

4. *В результате проведения эксперимента была получена следующая выборка:*

$$\begin{gathered}9, 60, 88, 74, 54, 28, 23, 72, 70, 91, 97, 79, 18, 44, 43, 82, 13, 79, 91, 49,\\ 7, 66, 10, 6, 60, 79, 64, 75, 63, 69, 71, 11, 5, 89, 76, 93, 50, 33, 6, 97,\\ 33, 83, 16, 98, 8, 38, 16, 21, 54, 53, 22, 92, 57, 95, 30, 16, 53, 67, 31, 8,\\ 1, 29, 79, 33, 85, 32, 50, 29, 17, 38, 54, 18, 52, 86, 9, 54, 25, 20, 53, 88,\\ 12, 65, 1, 2, 46, 83, 80, 48, 72, 92, 99, 23, 24, 47, 90, 74, 86, 55, 87, 1.\end{gathered}$$

постройте интервальный вариационный ряд этой выборки, разбив интервал между минимальной и максимальной вариантами на 7 интервалов одинаковой длины

5. *В результате проведения эксперимента была получена следующая вы-*

борка:

$$93, 144, 42, 8, 70, 137, 75, 52, 29, 151, 55, 46, 36, 52, 187, 66, 60,$$
$$170, 128, 144, 40, 31, 17, 40, 108, 151, 88, 197, 39, 70, 172, 36, 95, 160,$$
$$53, 73, 190, 55, 53, 56, 67, 29, 188, 49, 87, 60, 98, 150, 27, 196, 33,$$
$$4, 183, 88, 72, 6, 158, 72, 149, 5, 71, 68, 20, 8, 56, 184, 143, 172,$$
$$18, 68, 193, 54, 169, 155, 12, 54, 31, 65, 46, 169, 117, 150, 198, 60, 42,$$
$$2, 23, 196, 126, 91, 175, 153, 29, 114, 43, 170, 109, 153, 60, 193.$$

постройте интервальный вариационный ряд этой выборки, разбив интервал между минимальной и максимальной вариантами на 11 *интервалов одинаковой длины*

6. *В результате проведения эксперимента была получена следующая выборка:*

$$0,77; 0,60; -0,89; -0,48; 1,49; -0,23; 0,78; -1,19; -0,52;$$
$$-1,06; -0,22; 0,97; -0,71; 0,55; 0,31; -0,89; 0,85; 0,53;$$
$$-0,48; -0,98; 0,82; -0,69; -1,31; -0,12; -0,64; 0,18; -0,31;$$
$$-0,96; -1,92; 1,77; 1,18; -0,51; 0,48; 0,21; 1,50; -0,12;$$
$$-1,59; -0,78; -0,43; 1,65; 0,58; 0,63; -0,80; -0,71; -0,54;$$
$$1,09; -0,65; 1,24; -0,10; -0,54; -0,77; 0,26; -1,01; -0,25;$$
$$-0,20; 1,11; -0,36; -0,25; 1,63; 0,82; 0,62; 0,53; 1,99;$$
$$0,56; 2,21; 0,90; -0,42; 2,59; -0,76; 0,91; -0,53; -2,06;$$
$$0,42; 1,47; 0,18; 1,57; 0,66; -0,88; 1,46; 0,55; -2,36;$$
$$0,52; -0,69; 1,25; -0,58; 0,19; -0,35; 0,37; 0,97; -0,96;$$
$$1,60; 0,86; 0,92; -2,14; -0,40; -0,40; 2,41; 0,79; -0,47; 1,01.$$

постройте интервальный вариационный ряд этой выборки, разбив интервал между минимальной и максимальной вариантами на 11 *интервалов одинаковой длины*

5 Эмпирическая функция распределения. Полигон. Гистограмма

1. *Найдите эмпирическую функцию распределения и постройте её график для выборок, описанных в задачах 1–3 из предыдущего пункта.*

2. *Найдите гистограммы распределения для выборок, описанных в задачах 4–6 из предыдущего пункта.*

3. *Пусть выборка X задана своим дискретным вариационным рядом:*

X	-2	-1	0	2	1	2	3
	6	8	11	12	17	9	8

Найдите эмпирическую функцию распределения этой выборки и постройте её график.

4. *Пусть выборка X задана своим дискретным вариационным рядом:*

X	1	2	3	4	5	6
	5	45	65	66	45	5

Найдите эмпирическую функцию распределения этой выборки и постройте её график.

5. *Пусть выборка X задана своим интервальным вариационным рядом:*

X	$(-2,-1)$	$(-1,0)$	$(0,1)$	$(1,2)$	$(2,3)$	$(3,4)$
	5	45	65	66	45	5

Постройте гистограмму этой выборки.

6 Точечные оценки

1. *Найдите выборочную среднюю, выборочную дисперсию, выборочное средне квадратичное и несмещённое выборочное среднеквадратичное выборок, описанных в задачах 1–3 в разделе "Выборки".*

7 Интервальные оценки

1. *Найдите доверительные интервалы для выборок с уровнями значимости $\gamma = 0,975$, $\gamma = 0,99$, $\gamma = 0,999$, описанных в задачах 1–3 в разделе "Выборки".*

8 Итоговые контрольные задания

Дана выборка значений случайной величины. Следует выполнить следующие задания:

1. постройте дискретный вариационный ряд этой выборки;

2. постройте интервальный вариационный ряд этой выборки, разбив интервал между минимальной и максимальной вариантами на 7 интервалов одинаковой длины;

3. постройте по дискретному вариационному ряду полигон данной выборки;

4. постройте по интервальному вариационному ряду гистограмму данной выборки;

5. найдите точечные оценки данной выборки (выборочную среднюю, выборочную дисперсию, выборочное средне квадратичное отклонение, несмещённое выборочное средне квадратичное отклонение);

6. предполагая, что рассматриваемая случайная величина имеет нормальное распределение, найдите доверительные интервалы для её средне квадратичного отклонения с заданным уровнем значимости γ.

7. предполагая, что рассматриваемая случайная величина имеет нормальное распределение, найдите доверительные интервалы для её математическое ожидание с заданным уровнем значимости γ (предполагая, что её средне квадратичное отклонение совпадает с серединой доверительного интервала, найденного на предыдущем шаге).

8. предположим, что рассматриваемая случайная величина имеет нормальной распределение и среднеквадратичное отклонение совпадает с серединой доверительного интервала, найденного в п.6. Можно ли с уровнем значимости $\alpha =$ принять гипотезу о равенстве $\mathbf{M} = m_0$ (где m_0 - середина доверительного интервала, найденного в п. 7). Проверить все типы альтернативных гипотез.

9. предположим, что рассматриваемая случайная величина имеет нормальной распределение. Можно ли с уровнем значимости $\alpha =$ принять гипотезу о равенстве $\mathbf{M} = m_0$ (где m_0 — середина доверительного интервала, найденного в п. 7). Проверить все типы альтернативных гипотез.

Список литературы

[1] Луценко А. И. Теория вероятностей //Ростов-на-Дону, изд. Феникс. 2009 г.

[2] Общий курс высшей математики для экономистов, под ред. В. И. Ермакова, любое издание

[3] Сборник задач по высшей математике для экономистов // под ред. В. И. Ермакова, любое издание

[4] Гмурман В. Е. Теория вероятностей и математическая статистика // М., Высшая школа, 1975.

[5] Гмурман В. Е. Руководство к решению задач по теории вероятностей и математической статистике // М., Высшая школа, 1979

[6] Кобзарь А. И. Прикладная математическая статистика. Для инженеров и научных работников // М.: ФИЗМАТЛИТ.- 2006.- 816 с.

[7] Орлов А. И Прикладная статистика // М., изд. "Экзамен" .- 2004 .- 482 с.