4. Построение выборки
4.1. Генеральная совокупность и выборка
В статистике вы обычно хотите изучить определенный набор объектов или предметов. Этот набор объектов известен в статистике как **генеральная совокупность**. Например, ранее вы изучили данные, полученные из Интернета о хлопьях для завтрака. Было бы полезно получить информацию обо всех хлопьях для завтрака, произведенных из любой точки мира, но это нереально, поскольку такое исследование заняло бы слишком много времени, а проведение такого крупного исследования стоило бы слишком дорого. Более реалистичной целью было бы изучение какой-нибудь выборки из генеральной совокупности.
Генеральная совокупность — это набор объектов, о которых необходимо получить информацию.
Выборка — это небольшой набор объектов, извлеченных из генеральной совокупности.
В идеале выборка должна иметь все характеристики генеральной совокупности. Другими словами, это должна быть миниатюрная версия генеральной совокупности с одинаковыми с ней свойствами. К сожалению, практически невозможно для любой выборки любой генеральной совокупности обладать <u>всеми</u> характеристиками генеральной совокупности. Например, при исследовании по хлопьям для завтрака информация была собрана в США, поэтому многие хлопья для завтрака, произведенные в других странах, не были представлены. Вполне вероятно, что другие хлопья для завтрака имеют свойства, отличные от тех, которые производятся и продаются в США, но по ним не было «никакой информации». Иногда у вас может быть время, деньги и ресурсы, чтобы <u>полностью</u> исследовать генеральную совокупность. Такое исследование называется **переписью**. Правительства многих стран проводят переписи населения. Данные переписи могут использоваться при планировании, например, предоставления жилья и образования.
В целом выбор выборки предпочтительнее переписи, поскольку проведение переписи занимает слишком много времени и требует слишком много денег. Есть несколько других причин, по которым вы можете предпочесть взять выборку, а не проводить перепись.
- Вероятно, при взятии выборки будет меньше ошибок при записи, поскольку в нем, вероятно, будет задействовано меньше людей. Эти люди, скорее всего, будут иметь больше времени для проведения измерений и записи для обследования, поскольку размер и сложность задачи будут не такими большими, как при проведении переписи.
- Анализ результатов может быть проведен быстрее.
- Иногда измерение на отдельном элементе выборки может привести к его уничтожению.

Например, если вы записывали время работы батарейки определенной марки, то каждая выбранная батарейка будет использоваться до тех пор, пока она не перестанет работать. Если бы была проведена перепись, то каждая батарейка, когда-либо созданная этим производителем, была бы уничтожена. Вряд ли какой-либо производитель захочет согласиться на такую перепись!
Когда вы делаете выборку из генеральной совокупности, вы хотите убедиться, что ваш метод отбора сделан таким образом, чтобы ни одна особенность генеральной совокупности не была пере представлена или не до представлена в выборке.
Например, предположим, что вы отбирали учеников из школы с целью определения среднего роста всех учеников в школе. Вы не построили бы выборку, состоящую исключительно из первоклассников, потому что вы ожидаете, что такая выборка будет недооценивать истинный средний рост. Когда метод построения выборки пере представляет или не до представляет особенность генеральной совокупности, он считается предвзятым методом построения выборки. Любой хороший метод отбора должен стараться максимально снизить вероятность смещения.
Наиболее распространенным подходом к решению задачи предотвращения смещения является построение случайной **выборки**.

Случайную выборку строят таким образом, что:
1) каждый объект генеральной совокупности имеет одинаковую вероятность быть отобранным;
2) объекты отбирают независимо друг от друга.
Через N будем обозначать объем генеральной совокупности, а через n — объем выборки.

4.2.Построение случайной выборки
Для того, чтобы построить случайную выборку, вам нужен список всех объектов генеральной совокупности. Этот список называется **рамкой выборки**.

Предположим, например, что генеральная совокупность состоит из 450 различных объектов и что вы хотите построить случайную выборку размером 20. Вы должны назначить каждому объекту генеральной совокупности уникальный идентификационный номер от 1 до 450 включительно. Затем вы должны написать каждое из чисел от 1 до 450 на отдельных билетах и положить их в большую шляпу (или другой контейнер). Чтобы получить случайную выборку размером 20, вы вытягиваете из шляпы 20 билетов, один за другим, так же, как и в конкурсе лотереи. Вы должны убедиться, что билеты тщательно перетасовываются между каждым выбором, и вы должны выбирать билеты «без возвращения».
Описанный выше метод довольно трудоемкий, если объем генеральной совокупности велик, и поэтому более распространено использование альтернативных методов, которые используют таблицы случайных чисел или генератор случайных чисел на калькуляторе или компьютере.
Функция случайных чисел в калькуляторе предоставляет десятичное число от 0 до 1, заданное фиксированному числу десятичных знаков. Количество знаков после запятой варьируется в зависимости от марки калькулятора. Некоторые калькуляторы дают десятичную дробь с 10 цифрами после запятой. Чтобы калькулятор мог генерировать это случайное десятичное число, вам, вероятно, потребуется поискать команду [rand] или (rnd] на своем калькуляторе. Каждый раз, когда вы даете калькулятору эту команду, он будет производить десятичное число в диапазоне от 0,00000 00000 до 0,99999. 99999 включительно. Каждый раз, когда вы используете [rand], появляется новая десятизначная десятичная дробь. Строго говоря, десятичные дроби, создаваемые таким образом, не случайны, потому что калькулятор использует систему правил для получения каждой десятичной дроби, и если бы вы знали правила, вы могли бы предсказать каждое десятичное число. Это означает, что полученные цифры на самом деле не зависят друг от друга. Для действительно случайной последовательности цифр процесс, который производит цифры, должен гарантировать, что каждая цифра в последовательности с равной вероятностью будет любой из цифр 0, 1, 2, ..., 9, независимо от любых других цифр. Тем не менее последовательности, полученные таким образом, имеют некоторые свойства случайных цифр, и их обычно считают приемлемыми в качестве альтернативы относительно действительно случайной последовательности цифр.
Мы будем использовать случайные числа, которые записаны блоками по пять цифр.
Таблица случайных чисел представляет собой список целых чисел 0, 1, 2, … , 9, который создается таким образом, что каждая цифра одинаково может появиться в любой позиции в списке. Этот список цифр был взят из набора таблиц случайных чисел.

2 4 3 5 9	7 4 0 2 5	9 0 8 3 1	8 8 6 1 0
1 4 6 6 8	7 8 2 9 2	5 1 4 7 0	1 7 5 0 5
4 0 5 8 0	9 6 4 1 8	7 3 3 8 1	2 3 1 1 2

Хотя таблицы организованы в блоки из пяти цифр, вы можете считать, что они образуют один длинный список. Процедура построения случайной выборки приведена ниже. Объем генеральной совокупности составляет 450, и каждый объект генеральной совокупности имеет уникальный идентификационный номер от 1 до 450 включительно.
Выберите произвольную начальную точку в таблицах (используя кубики или подобным способом). Цифра, выделенная жирным шрифтом, показывает случайно выбранную начальную точку.

2 4 3 5 9	7 4 0 2 5	9 0 8 3 1	8 8 6 1 0
1 4 6 6 8	7 8 2 9 2	5 **1** 4 7 0	1 7 5 0 5
4 0 5 8 0	9 6 4 1 8	7 3 3 8 1	2 3 1 1 2

Объем генеральной совокупности представляет собой трехзначное число, поэтому отсчитайте первые три цифры, начиная с цифры, выделенной жирным шрифтом, и получите число 147. Это означает, что первым объектом в вашей выборке является объект, идентификационный номер которого равен 147. Поскольку объем выборки должна быть равен 20, вам нужно выбрать еще 19 объектов из генеральной совокупности. Возьмите следующий блок из трех цифр, 017, и выберите объект с номером 17. Третий блок - 505; это не соответствует идентификационному номеру какого-либо объекта генеральной совокупности, поэтому проигнорируйте его и переходите к следующему блоку 405. Продолжайте таким образом, выбирая блоки из трех цифр, пока у вас не будет выборки из 20 объектов. Если трехзначный блок встречается более одного раза, игнорируйте любое последующее вхождение и переходите к следующему блоку.
Вы можете подумать, что расточительно отбрасывать все блоки цифр, которые дают числа, превышающие 450. Вы можете избежать этой потери, выделив второй блок из трех цифр каждому объекту генеральной совокупности. Тогда 900 из 1000 возможных трехзначных комбинаций будут

соответствовать объектам в вашей генеральной совокупности. Вы не можете распределить оставшиеся 100 блоков среди объектов генеральной совокупности, потому что это даст некоторым объектам генеральной совокупности больше шансов быть избранным, чем другим. Для случайной выборки каждый объект генеральной совокупности должен иметь равные шансы быть выбранным.
Есть много способов, которыми вы можете выделить два блока для каждого объекта генеральной совокупности. В идеале вам нужен способ, позволяющий легко определить, какой объект генеральной совокупности будет выбран. Следующая таблица показывает один из возможных способов.

Идентификационный номер объекта генеральной совокупности	Случайные блоки цифр
1	001 002
2	003 004
. . .	. . .
450	899 900
Выделение случайных чисел объектам генеральной совокупности	

Легко определить, какой объект генеральной совокупности будет выбран, используя систему этой таблицы. Для блока с четным номером вы делите его значение пополам, чтобы найти подходящего объекта генеральной совокупности, а для блока с нечетным номером вы добавляете один, а затем делите число на два.
Например, блок 420 соответствует номеру объекта $\frac{1}{2} \cdot 420 = 210$ и блок 555 соответствует номеру объекта $\frac{1}{2} \cdot (555 + 1) = 278$
Если объем генеральной совокупности больше 1000, но меньше 10 000, вы можете адаптировать описанный метод построения выборки, взяв блоки, содержащие четыре случайные цифры вместо трех.
Метод легко настроить, если вы используете функцию случайных чисел на калькуляторе.
Предположим, например, что вы получили случайное десятичное число 0.33936 62525. Вы можете игнорировать десятичную точку и рассматривать 10 цифр как 10 случайных цифр 33936 62525. После этого вы можете использовать метод, который был задан для использования с таблицами случайных чисел.
В некоторых учебниках вы можете увидеть альтернативный метод, предложенный для использования функции случайных чисел на калькуляторе. Приведенные ниже инструкции суммируют этот метод для объема генеральной совокупности размера N.

- Выберите случайное десятичное число, r, на вашем калькуляторе
- Умножьте r на N
- Возьмите целую часть этого ответа и добавьте 1.

Этот процесс выражается в следующей формуле:
Идентификационный номер выбранного объекта генеральной совокупности =
$int(N \cdot r) + 1$
где int (x) - это функция, результатом которого является наибольшее целое число, не превышающее x (целая часть x).
Например, если N=450 и r=0.04399 19875
int(N x r)+1= int(450 x 0.04399 19875)+1
= int(19.79...) + 1= 19 +1 = 20
Этот метод однозначно позволяет каждому объекту генеральной совокупности быть выбранным, но одинаковые ли шансы каждого объекта быть выбранным? Строго говоря, ответ на этот вопрос - «Нет» по следующей причине.
Существует 10^{10} различных комбинаций из 10-ти случайных цифр, и каждый из указанных выше методов генерирует целое число от 1 до 450 включительно. Тем не менее, 10^{10} не делится на 450, поэтому некоторые значения от 1 до 450 должны произойти с большей вероятностью, чем другие. На

практике различия между частотой появления каждой цифры невелики, и этот метод часто используется в качестве быстрого практического метода построения случайной выборки.
Вы должны понимать, что даже если вы взяли случайную выборку, вам не гарантируется, что ваша выборка будет репрезентативной. Например, после выбора случайной выборки из 10 учащихся из школы с целью оценки их среднего роста вы можете обнаружить, что каждый член выборки является первоклассником. Это не очень вероятно, но это, безусловно, возможно. Метод случайной выборки не гарантирует, что выбранная выборка является репрезентативной. Что гарантировано, так это то, что метод отбора свободен от предвзятости.
Вы также должны понимать, что не всегда можно построить случайную выборку. Вы помните, что вам нужна **рамка выборки**, прежде чем вы сможете выбрать случайную выборку из генеральной совокупности, иногда это невозможно. Например, предположим, что ваша генеральная совокупность - это все дальтоники. Для этой группы населения список недоступен, поэтому случайная выборка не может быть взята. В таких случаях были разработаны другие неслучайные методы построения выборки.
Пример 4.2.1.
Страховая компания получает большое количество претензий за ущерб от шторма. После периода особенно штормовой погоды были получены 42 претензии за один день. У страховой компании имеется достаточный персонал для расследования только 6-ти претензий из этих 42-х претензий. Претензии пронумерованы от 01 до 42, и есть несколько предложений о том, как построить выборку из 6-ти претензий. Прокомментируйте каждый из следующих методов построения выборки, с объяснением того, даст ли метод случайную выборку или нет. В каждом случае требуется построить выборку из 6-ти претензий.
(а) Выберите шесть крупнейших претензий.
(б) Выберите двузначные случайные числа, игнорируя 00 и любое число больше 42. Когда будут получены шесть различных чисел, выберите соответствующие претензии.
(в) Выберите двузначные случайные числа. Разделите на 42, возьмите остаток, добавьте 1 и выберите соответствующие претензии. (Например, если выбрано значение 44, будет выбран номер претензии 03).
(г) Как же как в части (в), но при выборе двухзначных случайных чисел игнорировать числа равные 84 и более.
(д) Произвольно выберите случайную цифру, игнорируя 0, 8 и 9. Выберите претензию, соответствующую этому номеру, и каждую седьмую претензию после этого. (Например, если выбрано 3, выберите претензии с номерами 03, 10, 17, 24, 31, 38.)
Решение:
(а) Выбор шести крупнейших претензий был бы хорошей идеей, если у страховой компании есть какие-либо сомнения относительно претензий, поскольку компания не хотела бы выплачивать большие суммы, если претензии были необоснованными. Тем не менее, этот метод не дает каждому из 42 претензий равную вероятность быть выбранным, поэтому он не является методом построения случайной выборки.
(б) Этот метод предоставит случайную выборку. Однако многие пары случайных цифр будут игнорироваться, так как вполне вероятно, что некоторые пары будут больше 42, поэтому метод неэффективен и расточителен.
(в) Этот метод даст результаты от 1 до 42 включительно, но не все числа от 1 до 42 имеют одинаковый шанс быть выбранными. Например, претензия 1 имеет три возможные пары случайных цифр, которые ему соответствуют. Это 00, 42 и 84. Претензия 30 имеет только две пары случайных цифр, которым она соответствует, 29 и 71. Поэтому предлагаемый метод не даст случайную выборку.
(г) Возражения о том, что некоторые пары встречаются чаще, чем другие, более не будут обоснованными. Поэтому описанный метод предоставит случайную выборку.
(д) Этот метод обычно известен как систематическая выборка. Каждая отдельная претензия имеет равные шансы быть выбранным, но метод, тем не менее, не дает случайной выборки. Для случайной выборки каждая возможная <u>группа</u> из 6-ти претензий должна иметь одинаковый шанс выбора. Однако, используя этот метод, было бы невозможно, например, выбрать первые 6 претензий (претензии 01, 02, 03, 04, 05 и 06). При условии, что список претензий не организован в каком-либо конкретном порядке, это вполне может быть очень разумным методом выбора выборки, но он не обеспечивает действительно случайную выборку в соответствии со строгим определением случайной выборки.

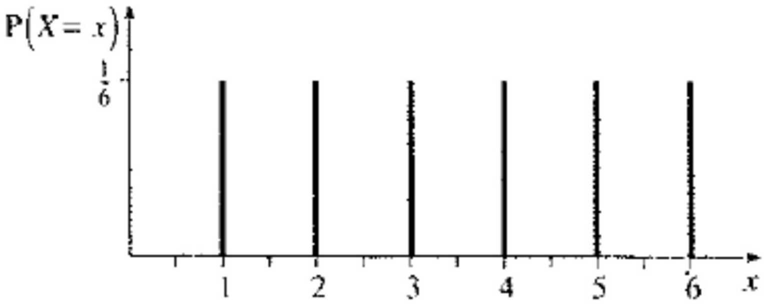
4.3. Распределение выборочного среднего выборки из генеральной совокупности

Чтобы исследовать связь между генеральной совокупностью и выборками, извлеченными из нее, полезно начать с простой практической ситуации.
Предположим, что вы делаете выборочные броски из кубиков. Лучший способ сделать это случайно - бросить кости. Если вам нужен выборка размера один, вы бросаете одну кость: если вы хотите выборку размера два, вы бросаете две кости, что равносильно бросанию одной кости два раза подряд. Для выборки размера n вы должны бросить n кубиков или один кубик n раз подряд. Чтобы прояснить ситуацию, представьте следующую ситуацию, начиная с самого простого случая.

Выборка размера один

Предположим, вы играете в игру, в которой вы получаете приз в $, равный количеству очков, полученному за один бросок кости. Перед тем, как начать играть в игру, вы не знаете точно, какой будет сумма вашего приза, поэтому ваш приз является случайной величиной.
Для одного броска кости случайная величина количества очков, X, имеет закон распределение вероятностей, приведенное в таблице и на рисунке.

Значение, *x*	1	2	3	4	5	6
P(*X*=*x*)	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
Закон распределение вероятностей для X, количества очков при одном подбрасывании кости						



Среднее значение (математическое ожидание), μ , случайной величины X, равна $3\frac{1}{2}$. Это можно увидеть из симметричности распределения.
Для нахождения дисперсии случайной величины X, используем формулу нахождения дисперсии дискретной случайной величины
$\sigma^2 = Д(X) = \sum x_i^2 p_i - \mu^2$. В данном случае:

$$\sigma^2 = Д(X) = 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + 4^2 \cdot \frac{1}{6} + 5^2 \cdot \frac{1}{6} + 6^2 \cdot \frac{1}{6} - \left(3\frac{1}{2}\right)^2 =$$

$$= (1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) \cdot \frac{1}{6} - \left(3\frac{1}{2}\right)^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$$

Выборка размера два

Предположим теперь, что правилом игры является то, что вы получаете приз в $, равный среднему значению очков, полученных за два броска костей. Например, если выяснилось, что вы бросили 4, а затем 1, ваш приз составил бы $2.50, тогда как если вы набрали 6 и 4, то ваш приз составил бы $5.
Перед тем, как начать играть в игру, вы не знаете точно, какой будет сумма вашего приза, поэтому стоимость вашего приза является случайной величиной.
Пусть X_1 – количество очков при первом броске кости, а X_2–количество очков при втором броске кости, то **выборочное среднее** – среднее значение, которое выражается через X_1 и X_2 как $\bar{X} = \frac{X_1+X_2}{2}$. Обратите внимание, что каждый из случайных величин X_1 и X_2, имеет такое же распределение, как и распределение случайной величины X, рассмотренной выше.
Таким образом, в этом примере $\frac{X_1+X_2}{2}$ является средним значением, но также является случайной величиной, поскольку она зависит от двух отдельных показателей, которые сами являются случайными величинами. Естественный символ для обозначения этой случайной величины - $\bar{X}$. В дальнейшем, мы будем обсуждать размер выборки, поэтому обозначим выборочное среднее как $\bar{X}(2)$, поскольку она

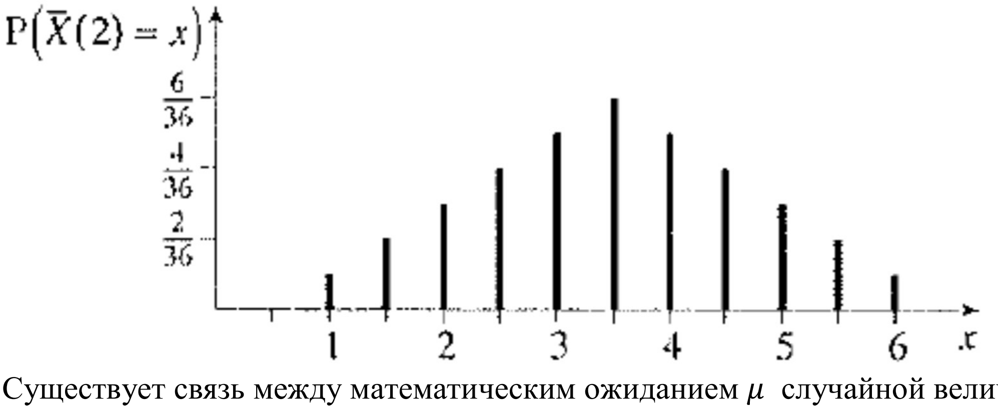
представляет собой среднее из **двух** оценок количества очков кости или, что эквивалентно, среднее значение выборки размера **два**, взятой из распределения случайной величины X.
Поскольку $\bar{X}(2)$ - является случайной величиной, у нее есть свой закон распределения вероятностей. В следующей таблице показаны все 36 одинаково вероятных возможных результатов двух бросков кости. Записи представляют средние оценки 36 возможных пар.

		x_2, значения случайной величины X_2					
		1	2	3	4	5	6
x_1, значения случайной величины X_1	1	1	$1\frac{1}{2}$	2	$2\frac{1}{2}$	3	$3\frac{1}{2}$
	2	$1\frac{1}{2}$	2	$2\frac{1}{2}$	3	$3\frac{1}{2}$	4
	3	2	$2\frac{1}{2}$	3	$3\frac{1}{2}$	4	$4\frac{1}{2}$
	4	$2\frac{1}{2}$	3	$3\frac{1}{2}$	4	$4\frac{1}{2}$	5
	5	3	$3\frac{1}{2}$	4	$4\frac{1}{2}$	5	$5\frac{1}{2}$
	6	$3\frac{1}{2}$	4	$4\frac{1}{2}$	5	$5\frac{1}{2}$	6
*Возможные **выборочные средние** при броске двух костей.*							

Из 36 возможных пар (x_1, x_2) в четырех случаях средний балл составляет $2\frac{1}{2}$. Четырьмя случаями являются пары: (1,4), (2,3), (3,2) и (4,1).
Поэтому $P\left(\bar{X}(2) = 2\frac{1}{2}\right) = \frac{4}{36} = \frac{1}{9}$. Подсчитав все пары (x_1, x_2) пары, которые дают 11 возможных средних баллов, вы сможете убедиться, что в следующей таблице представлен закон распределения вероятностей случайной величины $\bar{X}(2)$. Это распределение называется распределением **выборочного среднего** значения, в данном случае для среднего значения двух бросков кости.

Значение, *x*	1	$1\frac{1}{2}$	2	$2\frac{1}{2}$	3	$3\frac{1}{2}$	4	$4\frac{1}{2}$	5	$5\frac{1}{2}$	6
$P(\bar{X}(2) = x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$
*Закон распределения вероятностей **выборочного среднего** двух бросков кости*											

Распределение иллюстрируется на этом рисунке:



Существует связь между математическим ожиданием μ случайной величины X и математическим ожиданием **выборочного среднего** $\bar{X}(2)$. Из симметричности распределения $\bar{X}(2)$: $M(\bar{X}(2) = 3\frac{1}{2}$, поэтому $M(\bar{X}(2) = \mu$.
Дисперсии случайных величин X и $\bar{X}(2)$также связаны. Дисперсия $\bar{X}(2)$ определяется как:

$$Д\left(\bar{X}(2)\right) = 1^2 \cdot \frac{1}{36} + \left(1\frac{1}{2}\right)^2 \cdot \frac{2}{36} + 2^2 \cdot \frac{3}{36} + \cdots + 6^2 \cdot \frac{1}{36} - \left(3\frac{1}{2}\right)^2 = \frac{35}{24}$$

Как вы можете заметить: $\sigma^2 = \frac{35}{12}$, поэтому $Д\left(\bar{X}(2)\right) = \frac{1}{2}\sigma^2$

Подводя итог и возвращаясь к языку выборок из генеральной совокупности, для выборки размера два: $\mathrm{М}(\bar{X}(2)) = \mu$и $\mathrm{Д}(\bar{X}(2)) = \frac{1}{2}\sigma^2$, где $\mathrm{М}(\mathrm{X}) = \mu$ и $\mathrm{Д}(\mathrm{X}) = \sigma^2$

Эти результаты верны для *всех* распределений, а не только для этого конкретного. Результаты следуют из правил, установленных для линейных комбинаций случайных величин:

$$\mathrm{М}(\bar{\mathrm{X}}(2)) = \mathrm{М}\left(\frac{1}{2}(\mathrm{X}_1 + \mathrm{X}_2)\right) = \mathrm{М}\left(\frac{1}{2}\mathrm{X}_1 + \frac{1}{2}\mathrm{X}_2\right) = \frac{1}{2}\mathrm{М}(\mathrm{X}_1) + \frac{1}{2}\mathrm{М}(\mathrm{X}_2) =$$

$$= \frac{1}{2}\mu + \frac{1}{2}\mu = \mu = \mathrm{М}(\mathrm{X})$$, и

$$\mathrm{Д}(\bar{\mathrm{X}}(2)) = \mathrm{Д}\left(\frac{1}{2}(\mathrm{X}_1 + \mathrm{X}_2)\right) = \mathrm{Д}\left(\frac{1}{2}\mathrm{X}_1 + \frac{1}{2}\mathrm{X}_2\right) = \left(\frac{1}{2}\right)^2 \mathrm{Д}(\mathrm{X}_1) + \left(\frac{1}{2}\right)^2 \mathrm{Д}(\mathrm{X}_2) =$$

$$= \frac{1}{4}\sigma^2 + \frac{1}{4}\sigma^2 = \frac{1}{2}\sigma^2 = \frac{1}{2}\mathrm{Д}(\mathrm{X})$$

Выборка размера три

Возможно, теперь вы можете выразить значения $M(\bar{X}(3)$и $\mathrm{Д}(\bar{X}(3)$ через μ и σ^2для выборки размера три или, что эквивалентно, трех бросков кости. Пусть $\bar{X}(3)$ обозначает среднее количество выпавших очков при трех бросках кости.

Пусть X_1, X_2и X_3 – количество выпавших очков соответственно при первом, втором и третьем бросании кости. Тогда выборочное среднее этих трех бросков: $\bar{X}(3) = \frac{1}{3}(X_1 + X_2 + X_3)$.

Также $\bar{X}(3)$ – является случайной величиной , так как X_1, X_2и X_3 – являются случайными величинами.

Одним из возможных значений случайной величины $\bar{X}(3)$ является $2\frac{1}{3}$.

Существует четыре различных комбинаций выпавших очков, которые дают средние значение $2\frac{1}{3}$: это (1,1,5), (1,2,4), (1,3,3) и (2,2,3).

Комбинация (1,1,5) может встречаться в любом из 3-х различных вариантах: (1,1,5), (1,5,1), (5,1,1);
аналогично (1,2,4) могут встречаться в 6-ти различных вариантах: (1,2,4), (1,4,2), (2,1,4), (2,4,1), (4,1,2), (4,2,1);
(1,3,3) может встречаться в 3-х различных вариантах: (1,3,3), (3,1,3), (3,3.1);
и (2.2,3) могут встречаться в 3-х различных вариантах: (2,2,3), (2.3,2), (3,2.2).

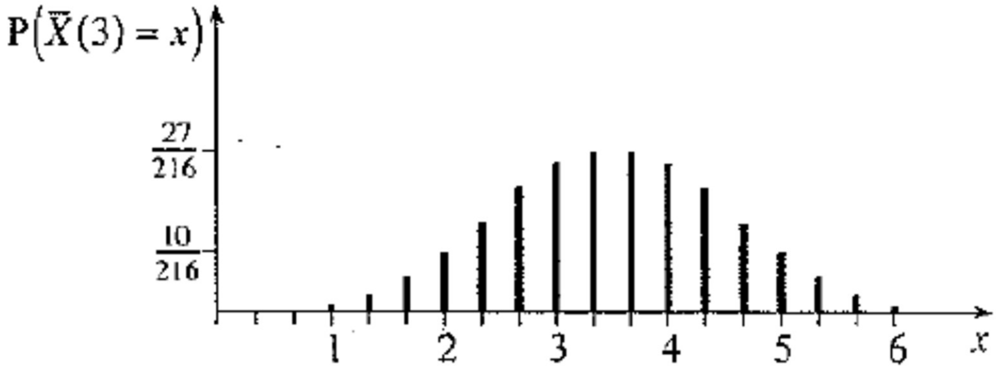
Поэтому: $P\left(\bar{X}(3) = 2\frac{1}{3}\right) = \frac{3}{216} + \frac{6}{216} + \frac{3}{216} + \frac{3}{216} = \frac{15}{216}$

Рассматривая все (x_1, x_2, x_3) тройки, которые дают каждый из 16 возможных средних значений очков, вы можете получить закон распределения вероятностей случайной величины $\bar{X}(3)$. Этот закон распределения приведен в следующей таблице. Это распределение **выборочного среднего** значения для выборок из 3 бросков кости.

Значение, x	$P(\bar{X}(3) = x)$	Значение, x	$P(\bar{X}(3) = x)$
1	$\frac{1}{216}$	$3\frac{2}{3}$	$\frac{27}{216}$
$1\frac{1}{3}$	$\frac{3}{216}$	4	$\frac{25}{216}$
$1\frac{2}{3}$	$\frac{6}{216}$	$4\frac{1}{3}$	$\frac{21}{216}$
2	$\frac{10}{216}$	$4\frac{2}{3}$	$\frac{15}{216}$
$2\frac{1}{3}$	$\frac{15}{216}$	5	$\frac{10}{216}$
$2\frac{2}{3}$	$\frac{21}{216}$	$5\frac{1}{3}$	$\frac{6}{216}$
3	$\frac{25}{216}$	$5\frac{2}{3}$	$\frac{3}{216}$
$3\frac{1}{3}$	$\frac{27}{216}$	6	$\frac{1}{216}$

Закон распределения вероятностей ***выборочного среднего*** *трех бросков кости*

Распределение иллюстрируется на этом рисунке:



Из симметричности распределения, математическое ожидание случайной величины $\bar{X}(3)$ также равна $3\frac{1}{2}$, то есть $M(\bar{X}(3)) = \mu$.
Дисперсия $\bar{X}(3)$ определяется как:

$$Д(\bar{X}(3)) = 1^2 \cdot \frac{1}{216} + \left(1\frac{1}{3}\right)^2 \cdot \frac{3}{216} + \left(1\frac{2}{3}\right)^2 \cdot \frac{6}{216} + \dots + 6^2 \cdot \frac{1}{216} - \left(3\frac{1}{2}\right)^2 =$$

$$= \left(1 + \frac{16}{9} \cdot 3 + \frac{25}{9} \cdot 6 + \dots + 36\right) \cdot \frac{1}{216} - \left(3\frac{1}{2}\right)^2 = \frac{35}{36}$$

Таким образом, $Д(\bar{X}(3)) = \frac{1}{3}\sigma^2$

Попробуйте доказать эти формулы используя свойства линейных комбинаций случайных величин.

Формулы математического ожидания и дисперсии **выборочного среднего** для выборки размера два и выборки размера три могут быть обобщены для выборки размера n: $M(\bar{X}(n)) = \mu = M(X)$и$Д(\bar{X}(n)) = \frac{\sigma^2}{n} = \frac{Д(X)}{n}$.

Эти формулы можно доказать следующим образом:

$$M(\bar{X}(n)) = M\left(\frac{1}{n}(X_1 + X_2 + \dots + X_n)\right) = \frac{1}{n}M(X_1) + \frac{1}{n}M(X_2) + \dots + \frac{1}{n}M(X_n) == \frac{1}{n}\mu + \frac{1}{n}\mu + \dots + \frac{1}{n}\mu$$

$$= n \cdot \frac{1}{n}\mu = \mu = M(X)$$

$$Д(\bar{X}(n)) = Д\left(\frac{1}{n}(X_1 + X_2 + \dots + X_n)\right) = \frac{1}{n^2}Д(X_1) + \frac{1}{n^2}Д(X_2) + \dots + \frac{1}{n^2}Д(X_n) =$$

$$= \frac{1}{n^2}\sigma^2 + \frac{1}{n^2}\sigma^2 + \dots + \frac{1}{n^2}\sigma^2 = n \cdot \frac{1}{n^2}\sigma^2 = \frac{\sigma^2}{n} = \frac{Д(X)}{n}$$

Если случайная выборка состоит из n наблюдений над случайной величиной X, и выборочное среднее этой выборки – случайная величина $\bar{X}$, то:$M(\bar{X}) = \mu$ и $Д(\bar{X}) = \frac{\sigma^2}{n}$, где $\mu = M(X)$и $\sigma^2 = Д(X)$.

Пример 4.3.1.

Неправильная монета, с вероятностью выпадения ГЕРБа при одном подбрасывании равной $\frac{2}{3}$, подброшена 20 раз. Обозначим $\bar{X}$ среднее число выпавших ГЕРБов при этих 20-ти подбрасываний. Вычислите: $M(\bar{X})$и $Д(\bar{X})$.

Решение:

Случайная величина X в этом случае представляет количество ГЕРБов за одно подбрасывание этой неправильной монеты. Закон распределение случайной величины X приведен в таблице:

Значение	0	1
P(X=x)	$\frac{1}{3}$	$\frac{2}{3}$

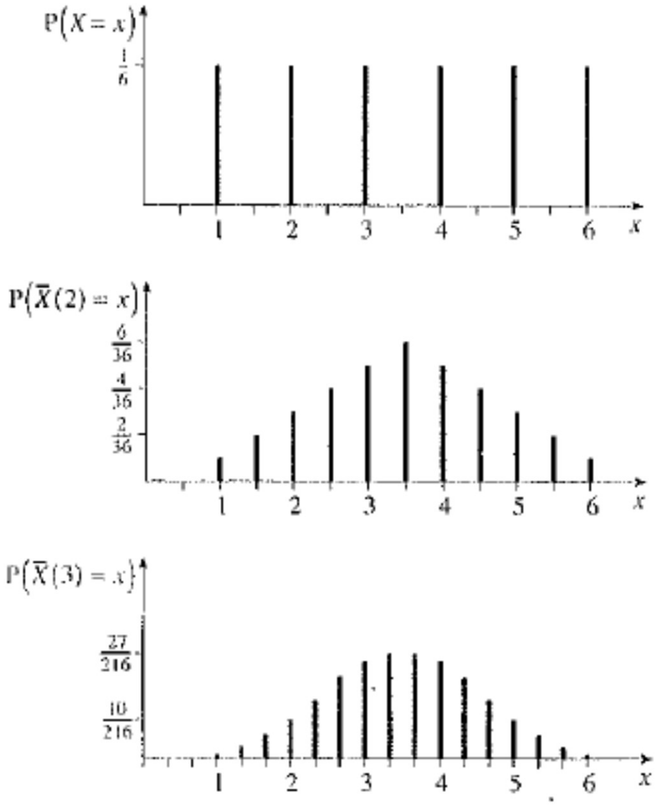
Математическое ожидание: $\mu = M(X) = 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3}$, и

Дисперсия: $\sigma^2 = Д(X) = 0^2 \cdot \frac{1}{3} + 1^2 \cdot \frac{2}{3} - \left(\frac{2}{3}\right)^2 = \frac{2}{9}$.

Используя формулы математического ожидания и дисперсии выборочного среднего, получим:
$M(\bar{X}(20)) = \mu = \frac{2}{3}$и $Д(\bar{X}(20)) = \frac{\sigma^2}{20} = \frac{1}{20} \cdot \frac{2}{9} = \frac{1}{90}$.

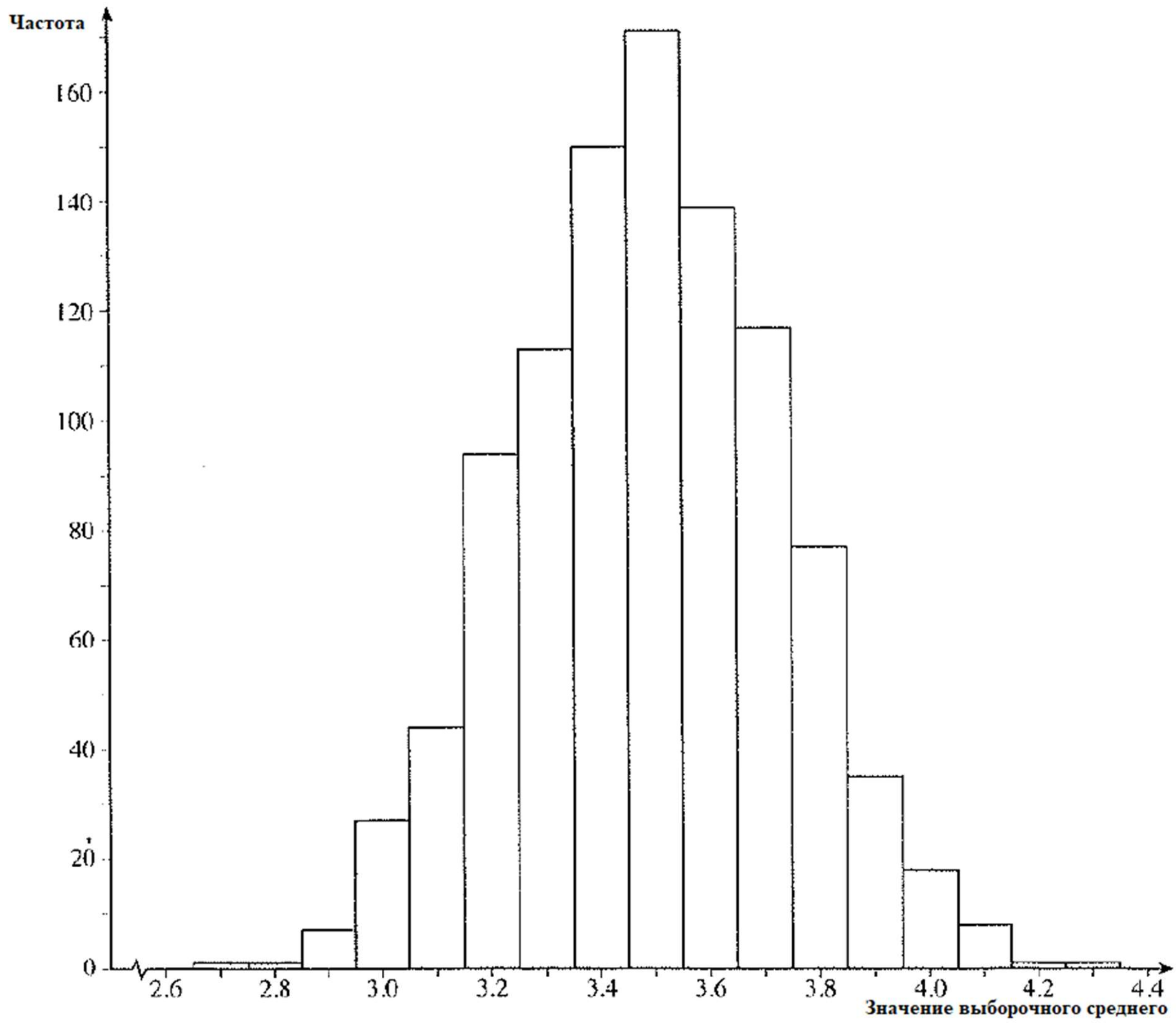
4.4.Центральная предельная теорема

Следующий рисунок сравнивает гистограммы, полученные для распределения очков одного броска кости и среднего значения очков при 2-х и 3-х бросках кости.



Распределение для одного броска является равномерным, в то время как для $\bar{X}(2)$и для $\bar{X}(3)$ не является равномерным. «Пик» распределение $\bar{X}(2)$ находится в середине, и значения в центре 3, $3\frac{1}{2}$ и 4 более вероятны, чем другие значения. Когда бросают кость три раза, центральные значения встречаются все чаще. Фактически, гистограмма для $\bar{X}(3)$ очень похожа на график нормального распределения.

По мере увеличения n, распределение $\bar{X}(n)$ становится все более похожим на нормальное распределение. На следующем рисунке показан результат моделирования с использованием компьютерных электронных таблиц (Excel) для выборки размером 50.



Диаграмма, показывающая распределение частот среднего значения 50 бросков справедливой кости для 1000 симуляций в электронной таблице

Итак, из рассмотренных нами примеров было бы разумно сделать вывод, что распределение **выборочного среднего** $\bar{X}(n)$ приблизительно нормальное распределение для больших значений n. Из доказанных ранее формул:

$\mathrm{M}(\overline{\mathrm{X}}) = \mu$ и $Д(\overline{\mathrm{X}}) = \frac{\sigma^2}{n}$

Эти результаты являются частными случаями более общей теоремы, называемой **центральной предельной теоремой**.

<u>Центральная предельная теорема.</u> Для любой последовательности независимых одинаково распределенных случайных величин $X_1, X_2, \ldots, X_n$ с конечным математическим ожиданием μи ненулевой дисперсией σ^2, при условии, что n достаточно велико, $\overline{X}$ имеет приблизительно нормальное распределение со средним μ и дисперсией $\frac{\sigma^2}{n}$, где $\overline{X} = \frac{X_1+X_2+\cdots+X_n}{n}$, то есть $\overline{X} \sim N\left(\mu; \frac{\sigma^2}{n}\right)$.

Обратите внимание, что размер выборки n был опущен из $\bar{X}$, поскольку из контекста обычно ясно, насколько большой размер выборки.

Возможно, вы задавались вопросом, насколько большим должно быть n, чтобы нормальное распределение было достаточно близко к распределению $\bar{X}$. К сожалению, простого ответа нет. Это зависит от распределения самого X. Если распределение X имеет некоторые особенности нормального распределения, то, вероятно, n вообще не обязательно будет очень большим. С другой стороны, если распределение X очень «смещено», тогда n, возможно, должно быть достаточно большим. При решении задач вы обычно можете судить о целесообразности использования центральной предельной теоремы. Правильное практическое правило заключается в применении теоремы, когда n больше 30.

Эта теорема является фундаментальной теоремой теории статистики и объясняет, почему нормальное распределение так широко изучено. Существенным моментом является то, что не имеет значения, какое распределение $X_1, X_2, \ldots, X_n$ имеет по отдельности: пока все они имеют одинаковое распределение и не зависят друг от друга, распределение выборочного среднего - $\bar{X}$ будет приблизительно нормальным, пока n достаточно велико. Центральная предельная теорема слишком сложна, чтобы привести строгое доказательство; однако предыдущее обсуждение средних значений очков по одной, двум и трех бросанием кости показывает, что теорема является разумной. Следующие примеры показывают, как теорема применима к другим случайным величинам.

Пример 4.4.1.

Непрерывная случайная величина X имеет функцию плотности вероятности f(x), определяемую как:

$$f(x) = \begin{cases} 2x, & \text{для } 0 \le x \le 1 \\ 0, & \text{в противном случае} \end{cases}$$

Найдите: (а) Математическое ожидание $\mu = M(X)$

(б) Дисперсию $\sigma^2 = Д(X)$

Была построена случайная выборка из 100 наблюдений над этой случайной величиной и найдено выборочное среднее значение $\bar{X}$. Рассчитать вероятность: $P(\bar{X} < 0.68)$

Решение:

Используя определения математического ожидания и дисперсии, получим:

(а) $\mu = \int_0^1 x \cdot 2x dx = \int_0^1 2x^2 dx = \left[\frac{2}{3}x^3\right]_0^1 = \frac{2}{3}$

(б) $\sigma^2 = \int_0^1 x^2 \cdot 2x dx - \left(\frac{2}{3}\right)^2 = \int_0^1 2x^3 dx - \frac{4}{9} = \left[\frac{2}{4}x^4\right]_0^1 - \frac{4}{9} = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$

По центральной предельной теореме распределение $\bar{X}$ примерно $N\left(\frac{2}{3}; \frac{1}{1800}\right)$

Стандартизируя, используя $Z = \frac{\bar{X} - \frac{2}{3}}{\sqrt{\frac{1}{1800}}}$, получим:

$P(\bar{X} < 0.68) \approx P\left(Z < \frac{0.68 - \frac{2}{3}}{\sqrt{\frac{1}{1800}}}\right) = P(Z < 0.56568 \dots) = 07143 = 0.714$, с точностью до трех значащих цифр.

Пример 4.4.2.

Сорок учеников бросали кубики по 12 раз. Затем каждый ученик записал, сколько раз в его 12-ти бросках выпало «шесть» очков. Затем учитель рассчитал среднее количество «шести» очков, полученных на одного студента. Найти вероятность того, что это среднее было больше 2.2.

Решение:

Пусть X_i – количество выпавших «шесть» очков i-го ученика, для i=1,2,…40.

Каждая случайная величина X_i – распределена по биномиальному закону распределения с параметрами n=12 и $p = \frac{1}{6}$.

Напомним, что для случайной величины X, которая распределена по биномиальному закону распределения, M(X)=np и Д(X)=npq. В данном случае: $M(X) = 12 \cdot \frac{1}{6} = 2$ и $Д(X) = 12 \cdot \frac{1}{6} \cdot \frac{5}{6} = \frac{5}{3}$

Используя центральную предельную теорему, получим: $\bar{X} \sim N\left(2; \frac{\frac{5}{3}}{40}\right)$, то есть приблизительно:

$\bar{X} \sim N\left(2; \frac{1}{24}\right)$.

$\bar{X}$ представляет собой выборочное среднее из 40-ка случайных величин, распределенных по биномиальному закону распределения, поэтому его можно записать в виде:

$\bar{X} = \frac{1}{40}(X_1 + X_2 + \dots + X_{40})$.

Нам необходим найти $P(\bar{X} > 2.2)$. Это выражение можно выразить через случайную величину T, где T – сумма 40-ка случайных величин, то есть:

$T = X_1 + X_2 + \dots + X_n$.

$P(\bar{X} > 2.2) = P(T > 2.2 \cdot 40) = P(T > 88)$

Заметим, что T - это общее количество выпавших «шесть» очков, полученных за 480 бросков кости, поэтому $T \sim B\left(480; \frac{1}{6}\right)$.

Напомним, что, когда вы используете нормальное распределение для приближенного вычисления биномиального распределения, вам требуется коррекция непрерывности, поэтому P (T> 88) приблизительно равно

P (V> 88 +0,5), где V - соответствующее нормальное приближение биномиальной случайной величины T.
То есть, если применить коррекцию непрерывности, то нам необходимо найти:
$P\left(\bar{X} > 2.2 + \frac{1}{80}\right)$.
Таким образом, когда применяется коррекция непрерывности к выборочному среднему значению $\bar{X}$ линейной комбинации n дискретных случайных величин, то коррекция непрерывности составляет $\frac{1}{2n}$ а не $\frac{1}{2}$.
Стандартизируя $\bar{X}$, получим: $Z = \frac{\bar{X}-2}{\sqrt{\frac{1}{24}}}$, где Z~N(0;1).
Тогда:
$P\left(\bar{X} > 2.2 + \frac{1}{80}\right) = P\left(Z > 0.2125\sqrt{24}\right) = P(Z > 1.041 \ldots) =$
$= 1 - \Phi(1.041 \ldots) = 0.1490 \ldots = 0.149$, с точностью до трех значащих цифр.

4.5. Распределение выборочного среднего $\bar{X}$, если $X_1, X_2, \ldots, X_n$ случайные величины, распределенные по нормальному закону распределения с параметрами $N(\mu; \sigma^2)$
Центральная предельная теорема утверждает, что распределение выборочного среднего $\bar{X}$ приблизительно равно $N\left(\mu; \frac{\sigma^2}{n}\right)$, когда $X_1, X_2, \ldots, X_n$ - независимые одинаково распределенные случайные величины с конечным математическим ожиданием μ и ненулевой дисперсией σ^2.
Необходимо, чтобы значение n было достаточно большим, чтобы это приближение было разумным.
Однако, если случайные величины $X_1, X_2, \ldots, X_n$ независимые и одинаково распределенные случайные величины, распределенные по нормальному закону распределения с параметрами $N(\mu; \sigma^2)$, то распределение выборочного среднего $\bar{X}$ - точно будет нормально распределенной случайной величиной с параметрами $N\left(\mu; \frac{\sigma^2}{n}\right)$, независимо от размера выборки n. Это следует из того факта, что линейные комбинации нормально распределенных случайных величин тоже будет нормально распределенной случайной величиной.
Пример 4.6.1.
(а) Масса случайно выбранного 15-летнего мальчика большой средней школы может быть смоделирована с помощью нормального распределения со средним значением 55 кг и стандартным отклонением 2,2 кг. Четыре 15-летних мальчика этой школы выбираются случайным образом. Рассчитайте вероятность того, что средняя масса этих четырех мальчиков будет:
(i) меньше 58 кг, (ii) от 52 кг до 57,5 кг.
(б) Вторая выборка размера n выбрана из 15-летних мальчиков этой школы. Насколько велик должен быть размер выборки n, чтобы вероятность того, что средняя масса данной выборки отличается от средней массы генеральной совокупности более чем на 0,6 кг, составляла не более 2%?
Решение:
(а) Пусть M_1, M_2, M_3 и M_4 – массы этих случайно выбранных четырех 15-летних мальчиков. Тогда $M_i \sim N(55; 2.2^2)$, для i=1,2,3,4.
Следовательно, $\bar{M} \sim N\left(55; \frac{2.2^2}{4}\right) = N(55; 1.21)$
Обратите внимание, что, хотя n очень мало, поскольку распределение для каждого M_i является нормальным, распределение выборочного среднего $\bar{M}$ тоже является нормальным.
Стандартизируем: $Z = \frac{\bar{M}-55}{1.1}$, где Z~N(0;1).
(i) $P(\bar{M} < 58) = P\left(Z < \frac{58-55}{1.1}\right) = P(Z < 2.727 \ldots) =$
$= \Phi(2.727 \ldots) = 0.9968 = 0.997$, с точностью до трех значащих цифр
(ii) $P(52 < \bar{M} < 57.5) = P\left(\frac{52-55}{1.1} < Z < \frac{57.5-55}{1.1}\right) =$
$= P(-2.727 \ldots < Z < 2.272 \ldots) = \Phi(2.272 \ldots) - \Phi(-2.727 \ldots) =$
$= \Phi(2.272 \ldots) - \left(1 - \Phi(2.727 \ldots)\right) = 0.9884 - (1 - 0.9968) =$
$= 0.985$, с точностью до трех значащих цифр.
(б) Пусть $M_1, M_2, \ldots, M_n$ – массы случайно выбранных n15-летних мальчиков этой школы. Тогда $\bar{M} \sim N\left(55; \frac{2.2^2}{n}\right)$. Следовательно нам необходимо найти минимальное значение n, для которого:

$$P(54.4 \leq \bar{M} \leq 55.6) \geq 0.98$$

Стандартизируем, $P\left(\frac{-0.6}{2.2/\sqrt{n}} < Z < \frac{0.6}{2.2/\sqrt{n}}\right) \geq 0.98.$

$$\Phi\left(\frac{0.6}{2.2/\sqrt{n}}\right) - \Phi\left(\frac{-0.6}{2.2/\sqrt{n}}\right) \geq 0.98$$

$$\Phi\left(\frac{0.6}{2.2/\sqrt{n}}\right) - \left(1 - \Phi\left(\frac{0.6}{2.2/\sqrt{n}}\right)\right) \geq 0.98$$

$$\Phi\left(\frac{0.6}{2.2/\sqrt{n}}\right) \geq 0.99$$

Используя обратную функцию, получим:

$$\frac{0.6}{2.2/\sqrt{n}} \geq \Phi^{-1}(0.99)$$

Из таблицы нормального распределения:

$$\frac{0.6}{2.2/\sqrt{n}} \geq 2.236$$

$$\sqrt{n} \geq 8.528 \ldots$$

$$n \geq 72.7 \ldots$$

Поскольку n должно быть целым числом, оно должно быть больше или равно 73. Таким образом, выборка должна содержать не менее 73 мальчика для ее средней массы, чтобы приблизить среднюю массу генеральной совокупности с требуемой точностью и достоверностью.

Задачи.
№1. Яйца, продаваемые в супермаркете, упакованы в коробки по 12 штук. Для каждого яйца вероятность того, что оно треснет, равна 0,05, независимо от всех остальных яиц. Выбирается случайная выборка из n коробок, и выборочная дисперсия числа треснувших яиц в коробке составляет 0,019. Найдите значение n.
Ответ №1: $n = 30$

№2. Среднее значение случайной выборки из n наблюдений, взятых из распределения $N(\mu, \sigma^2)$, обозначается $\bar{X}$. Дано, что $P(|\bar{X} - \mu| > 0{,}5\sigma) < 0{,}05$.
а) Найдите наименьшее значение n.
б) С этим значением n, найдите $P(\bar{X} < \mu + 0{,}1\sigma)$.
Ответ №2: а) 16 б) 0,6554

№3. Ботаник хочет оценить среднее значение μ и стандартное отклонение σ глубины почвы на большом прямоугольном поле. Прокомментируйте следующие методы получения выборочных точек.
а) Ботаник стоит в точке около центра поля, лицом к определенному направлению, и бросает камень через плечо. Новая точка отсчета – это место, где приземляется камень. Это повторяется, лицом к другому направлению.
б) Ботаник разбивает поле на квадраты со стороной 1 метр и использует таблицу случайных чисел для определения точки выборки в центре таких квадратов.
Предположим, что ботаник использует подходящую процедуру получения случайной выборки. Ей требуется, чтобы средняя глубина выборки отличалась от μ менее чем на 10% от σ с вероятностью не менее 97,5%. Найдите наименьший размер выборки, который она должна будет получить.
Ответы №3:
а) Выборка ограничена дальностью броска. Удовлетворительно, если исследователь перемещается в разные части поля перед броском.
б) Выборочные точки ограничены центрами квадратов сетки, а также точками, удаленными более чем на 0,5 м от границы. Однако получение такой выборки удовлетворительна.
Минимальный размер такой выборки равен 502

№4. Функция плотности распределения вероятности времени T часов, затраченных на ремонт оборудования может быть смоделирована как:

$$f(t) = \begin{cases} \dfrac{24}{7t^4}, & \text{для } 1 \le t \le 2, \\ 0, & \text{в противном случае.} \end{cases}$$

а) Найдите $M(T)$ и $Д(T)$
б) $\bar{T}$ обозначает среднее значение 30-ти случайно выбранных ремонтов. Предполагая, что центральная теорема справедлива, оцените $P(\bar{T} < 1{,}2)$.
Объясните, есть ли в вашем ответе небольшая или значительная ошибка.
Ответы №4: а) $\frac{9}{7}$ и $\frac{3}{49}$ б) 0,0299; распределение случайной величины T очень смещено(искажено), поэтому ответ не очень точен.

№5. Станок предназначен для производства шарикоподшипников со средним диаметром 1,2 см. Каждый день выбирается случайная выборка из 50 шарикоподшипников и точно измеряются их диаметры. Если выборочное среднее диаметров образцов лежит за пределами диапазона от 1,18 см до 1,22 см, то это будет считаться доказательством того, что средний диаметр изготовленных шарикоподшипников не составляет 1,2 см. Затем станок будет остановлена и в нее будут внесены изменения. Предполагая, что диаметры имеют стандартное отклонение 0,075 см, найдите вероятность того, что:
а) станок останавливается без необходимости.
б) машина не останавливается, когда средний диаметр шарикоподшипников составляет 1,15 см.
Ответ №5: а) 0,0594 б) 0,0023

№6. Количество ночных вызовов на пожарную станцию, обслуживающую небольшой город, можно смоделировать с помощью распределения Пуассона со средним значением 2,7 вызова за ночь.

а) Укажите математическое ожидание и дисперсию среднего числа ночных звонков за период n ночей.
б) Оцените вероятность того, что в течение данного года из 365 дней общее количество ночных звонков превысит 1050.

Ответ №6: а) 2,7 $\frac{2,7}{n}$ б) 0,0192

№7. Можно предположить, что прочность тротуарной плитки в общественных местах распределена по нормальному закону распределения со средним значением 50 единиц (условных) и стандартным отклонением 8 единиц (условных). Берется случайная выборка из n тротуарных плит и средняя прочность обозначается $\bar{X}$.
а) Укажите распределение случайной величины $\bar{X}$, приведя его среднее значение и дисперсию.
б) Найдите вероятность того, что $\bar{X}$ превышает 54 в случае $n = 25$.
в) Найдите наименьший размер выборки, если вероятность того, что $\bar{X}$ превышает 54 единицы, меньше 0,01.

Ответ №7: а) $N\left(50, \frac{64}{n}\right)$ б) 0,0062 в) 22

Приблизительно распределено по нормальному закону распределения, если n достаточно большая величина и приблизительно не распределено по нормальному закону распределения если n маленькая величина.

№8. Взята случайная выборка из $2n$ наблюдений случайной величины $X \sim B(n, p)$, где $p > 0{,}5$. Выборочное среднее значение обозначается $\bar{X}$. Дано, что $M(\bar{X}) = 64$ и $Д(\bar{X}) = 0{,}08$. Найдите:
а) значения p и n б) $P(\bar{X} > 64{,}5)$
Ответ №8: а) $p = 0{,}8$ и $n = 80$ б) 0,037

№9. Срок службы батареек Powerlong, продаваемых в упаковках по 6 штук, можно считать что распределена по нормальному закону распределения со средним значением 32 часа и стандартным отклонением σ часов. Найдите значение σ такое, что для одной коробки из 100 батареек (упаковок?) (в среднем) средний срок службы батареек составлял менее 30 часов.
Ответ №9: 2,11 часов

№10. Выборочное среднее 64-х наблюдений над случайной величиной X ($M(X) = 9$ и $Д(X) = 4$) обозначается как $\bar{X}$. Найдите границы интервала, такое что вероятность того, что $\bar{X}$ лежит в этом интервале равна 0,96.
Неравенство Чебышева гласит, что для любой случайной величины Y с математическим ожиданием μ и дисперсией σ^2 справедливо: $P(|Y - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}$, для $k \geq 1$.
Используя неравенство Чебышева, найдите другие границы интервала для $\bar{X}$. Прокомментируйте эти границы интервала по отношению к уже найденным границам интервала.
Ответ №10: $(8{,}49; 9{,}51)$, $(7{,}75; 10{,}25)$. (Подсказка: $k = 5$). Интервал шире, так как вероятность больше 0,96.