Лекция 1.

Предмет теории вероятностей. Случайные события. Алгебра событий. Относительная частота и вероятность случайного события. Полная группа событий. Классическое определение вероятности. Основные свойства вероятности. Основные формулы комбинаторики.

В различных разделах науки и техники нередко возникают ситуации, когда результат каждого из многих проводимых опытов заранее предугадать невозможно, однако можно исследовать закономерности, возникающие при проведении серии опытов. Нельзя, например, точно сказать, какая сторона монеты окажется сверху при данном броске: герб или цифра – но при большом количестве бросков число выпадений герба приближается к половине количества бросков; нельзя заранее предсказать результат одного выстрела из данного орудия по данной цели, но при большом числе выстрелов частота попадания приближается к некоторому постоянному числу. Исследование вероятностных закономерностей массовых однородных явлений составляет предмет **теории вероятностей**.

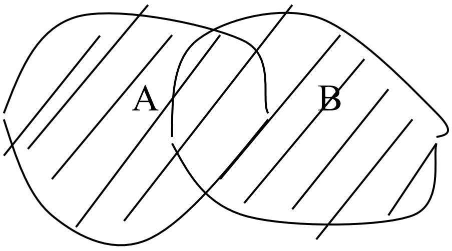
Основным интуитивным понятием классической теории вероятностей является **случайное событие**. События, которые могут произойти в результате опыта, можно подразделить на три вида:
а) **достоверное событие** – событие, которое всегда происходит при проведении опыта;
б) **невозможное событие** – событие, которое в результате опыта произойти не может;
в) **случайное событие** – событие, которое может либо произойти, либо не произойти. Например, при броске игральной кости достоверным событием является выпадение числа очков, не превышающего 6, невозможным – выпадение 10 очков, а случайным – выпадение 3 очков.

Алгебра событий.

Определение 1.1. **Суммой А+В** двух событий А и В называют событие, состоящее в том, что произошло хотя бы одно из событий А и В. **Суммой нескольких событий**, соответственно, называется событие, заключающееся в том, что произошло хотя бы одно из этих событий.

Пример 1. Два стрелка делают по одному выстрелу по мишени. Если событие А – попадание первого стрелка, а событие В – второго, то сумма А+В – это хотя бы одно попадание при двух выстрелах.
Пример 2. Если при броске игральной кости событием A_i назвать выпадение i очков, то выпадение нечетного числа очков является суммой событий $A_1+A_2+A_3$.

Назовем все возможные результаты данного опыта его *исходами* и предположим, что множество этих исходов, при которых происходит событие A (исходов, *благоприятных* событию A), можно представить в виде некоторой области на плоскости. Тогда множество исходов, при которых произойдет событие A+B, является объединением множеств исходов, благоприятных событиям A или B (рис. 1).



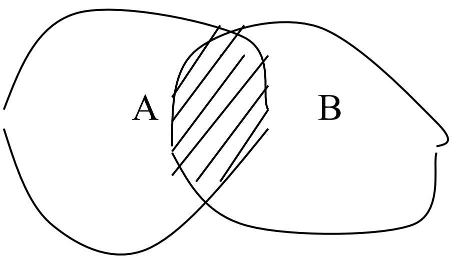
A + B

Рис.1.

Определение 1.2. **Произведением AB** событий A и B называется событие, состоящее в том, что произошло и событие A, и событие B. Аналогично **произведением нескольких событий** называется событие, заключающееся в том, что произошли все эти события.

Пример 3. В примере 1 (два выстрела по мишени) событием AB будет попадание обоих стрелков.
Пример 4. Если событие A состоит в том, что из колоды карт извлечена карта пиковой масти, а событие B – в том, что из колоды вынута дама, то событием AB будет извлечение из колоды дамы пик.

Геометрической иллюстрацией множества исходов опыта, благоприятных появлению произведения событий A и B, является пересечение областей, соответствующих исходам, благоприятным A и B.



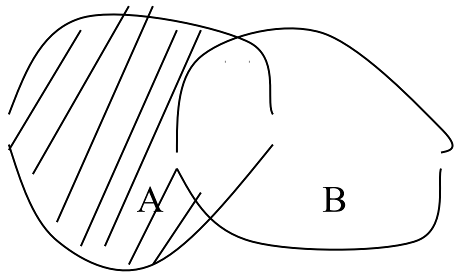
AB

Рис.2.

Определение 1.3. **Разностью A\B** событий A и B называется событие, состоящее в том, что A произошло, а B – нет.

Пример 5. Вернемся к примеру 1, где A\ B – попадание первого стрелка при промахе второго.

Пример 6. В примере 4 A\B – извлечение из колоды любой карты пиковой масти, кроме дамы. Наоборот, B \A – извлечение дамы любой масти, кроме пик.



A - B

Рис.3.

Введем еще несколько категорий событий.
Определение 1.4. События A и B называются **совместными**, если они могут произойти оба в результате одного опыта. В противном случае (то есть если они не могут произойти одновременно) события называются **несовместными.**
Примеры: совместными событиями являются попадания двух стрелков в примере 1 и появление карты пиковой масти и дамы в примере 4; несовместными – события A_1 – A_6 в примере 2.

Замечание 1. Если изобразить графически области исходов опыта, благоприятных несовместным событиям, то они не будут иметь общих точек.
Замечание 2. Из определения несовместных событий следует, что их произведение является невозможным событием.

Определение 1.5. Говорят, что события $A_1, A_2,\ldots,A_n$ образуют **полную группу**, если в результате опыта обязательно произойдет хотя бы одно из событий этой группы.

Замечание. В частности, если события, образующие полную группу, попарно несовместны, то в результате опыта произойдет *одно и только одно* из них. Такие события называют **элементарными событиями**.

Пример. В примере 2 события A_1 – A_6 (выпадение одного, двух,…, шести очков при одном броске игральной кости) образуют полную группу несовместных событий.

Определение 1.6. События называются **равновозможными**, если нет оснований считать, что одно из них является более возможным, чем другое.

Примеры: выпадение любого числа очков при броске игральной кости, появление любой карты при случайном извлечении из колоды, выпадение герба или цифры при броске монеты и т.п.

Классическое определение вероятности.

При изучении случайных событий возникает необходимость количественно сравнивать возможность их появления в результате опыта. Например, при последовательном извлечении из колоды пяти карт более возможна ситуация, когда появились карты разных мастей, чем появление пяти карт одной масти; при десяти бросках монеты более возможно чередование гербов и цифр, нежели выпадение подряд десяти гербов, и т.д. Поэтому с каждым таким событием связывают по определенному правилу некоторое число, которое тем больше, чем более возможно событие. Это число называется **вероятностью события** и является вторым основным понятием теории вероятностей.

Отметим, что само понятие вероятности, как и понятие случайного события, является аксиоматическим и поэтому не поддается строгому определению. То, что в дальнейшем будет называться различными определениями вероятности, представляет собой способы вычисления этой величины.

Определение 1.7. Если все события, которые могут произойти в результате данного опыта,

а) попарно несовместны;

б) равновозможны;

в) образуют полную группу,

то говорят, что имеет место **схема случаев.**

Можно считать, что случаи представляют собой все множество исходов опыта. Пусть их число равно n (число возможных исходов), а при m из них происходит некоторое событие А (число благоприятных исходов).

Определение 1.8. **Вероятностью события А** называется отношение числа исходов опыта, благоприятных этому событию, к числу возможных исходов:

$$p(A) = \frac{m}{n} \quad - \qquad (1.1)$$

- *классическое определение вероятности*.

Свойства вероятности.

Из определения 1.8 вытекают следующие свойства вероятности:

Свойство 1. Вероятность достоверного события равна единице.

Доказательство. Так как достоверное событие всегда происходит в результате опыта, то все исходы этого опыта являются для него благоприятными, то есть $m = n$, следовательно, $P(A) = 1$.

Свойство 2. Вероятность невозможного события равна нулю.
Доказательство. Для невозможного события ни один исход опыта не является благоприятным, поэтому $m = 0$ и $p(A) = 0$.

Свойство 3. Вероятность случайного события есть положительное число, заключенное между нулем и единицей.
Доказательство. Случайное событие происходит при некоторых исходах опыта, но не при всех, следовательно, $0 < m < n$, и из (1.1) следует, что $0 < p(A) < 1$.

Пример. Из урны, содержащей 6 белых и 4 черных шара, наудачу вынут шар. Найти вероятность того, что он белый.
Решение. Будем считать элементарными событиями, или исходами опыта, извлечение из урны каждого из имеющихся в ней шаров. Очевидно, что эти события удовлетворяют всем условиям, позволяющим считать их схемой случаев. Следовательно, число возможных исходов равно 10, а число исходов, благоприятных событию А (появлению белого шара) – 6 (таково количество белых шаров в урне). Значит,

$$p(A) = \frac{m}{n} = \frac{6}{10} = 0{,}6.$$

Относительная частота. Статистическое определение вероятности.

Классическое определение вероятности применимо только для очень узкого класса задач, где все возможные исходы опыта можно свести к схеме случаев. В большинстве реальных задач эта схема неприменима. В таких ситуациях требуется определять вероятность события иным образом. Для этого введем вначале понятие **относительной частоты W(A)** события А как отношения числа опытов, в которых наблюдалось событие А, к общему количеству проведенных испытаний:

$$W(A) = \frac{M}{N}, \qquad (1.2)$$

где N – общее число опытов, M – число появлений события A.
Большое количество экспериментов показало, что если опыты проводятся в одинаковых условиях, то для большого количества испытаний относительная частота изменяется мало, колеблясь около некоторого постоянного числа. Это число можно считать вероятностью рассматриваемого события.

Определение 1.9. **Статистической вероятностью события** считают его относительную частоту или число, близкое к ней.

Замечание 1. Из формулы (1.2) следует, что свойства вероятности, доказанные для ее классического определения, справедливы и для статистического определения вероятности.
Замечание 2. Для существования статистической вероятности события A требуется:
1) возможность производить неограниченное число испытаний;
2) устойчивость относительных частот появления A в различных сериях достаточно большого числа опытов.

Замечание 3. Недостатком статистического определения является неоднозначность статистической вероятности.

Пример. Если в задаче задается вероятность попадания в мишень для данного стрелка (скажем, $p = 0,7$), то эта величина получена в результате изучения статистики большого количества серий выстрелов, в которых этот стрелок попадал в мишень около семидесяти раз из каждой сотни выстрелов.

Основные формулы комбинаторики.

При вычислении вероятностей часто приходится использовать некоторые формулы *комбинаторики* – науки, изучающей комбинации, которые можно составить по определенным правилам из элементов некоторого конечного множества. Определим основные такие комбинации.

Определение 1.10. **Перестановки** – это комбинации, составленные из всех n элементов данного множества и отличающиеся только порядком их расположения. Число всех возможных перестановок

$$P_n = n! \tag{1.3}$$

Пример. Сколько различных списков (отличающихся порядком фамилий) можно составить из 7 различных фамилий?
Решение. $P_7 = 7! = 2\cdot3\cdot4\cdot5\cdot6\cdot7 = 5040$.

Определение 1.11. **Размещения** – комбинации из m элементов множества, содержащего n различных элементов, отличающиеся либо составом элементов, либо их порядком. Число всех возможных размещений

$$A_n^m = n(n-1)(n-2)...(n-m+1). \tag{1.4}$$

Пример. Сколько возможно различных вариантов пьедестала почета (первое, второе, третье места), если в соревнованиях принимают участие 10 человек?
Решение. $A_{10}^3 = 10\cdot 9\cdot 8 = 720$.

Определение 1.12. **Сочетания** – неупорядоченные наборы из m элементов множества, содержащего n различных элементов (то есть наборы, отличающиеся только составом элементов). Число сочетаний

$$C_n^m = \frac{n!}{m!(n-m)!}. \tag{1.5}$$

Пример. В отборочных соревнованиях принимают участие 10 человек, из которых в финал выходят трое. Сколько может быть различных троек финалистов?
Решение. В отличие от предыдущего примера, здесь не важен порядок финалистов, следовательно, ищем число сочетаний из 10 по 3:

$$C_{10}^{3} = \frac{10!}{3!\cdot 7!} = \frac{8\cdot 9\cdot 10}{6} = 120.$$

Лекция 2.

Геометрические вероятности. Теорема сложения вероятностей. Противоположные события. Условные вероятности. Теорема умножения вероятностей. Независимые события. Вероятность появления хотя бы одного события.

Одним из недостатков классического определения вероятности является то, что оно неприменимо к испытаниям с бесконечным количеством исходов. В таких случаях можно воспользоваться понятием **геометрической вероятности**.

Пусть на отрезок L наудачу брошена точка. Это означает, что точка обязательно попадет на отрезок L и с равной возможностью может совпасть с любой точкой этого отрезка. При этом вероятность попадания точки на любую часть отрезка L не зависит от расположения этой части на отрезке и пропорциональна его длине. Тогда вероятность того, что брошенная точка попадет на отрезок l, являющийся частью отрезка L, вычисляется по формуле:

$$p = \frac{l}{L}, \qquad (2.1)$$

где l – длина отрезка l, а L – длина отрезка L.
Можно дать аналогичную постановку задачи для точки, брошенной на плоскую область S и вероятности того, что она попадет на часть этой области s:

$$p = \frac{s}{S}, \qquad (2.1`)$$

где s – площадь части области, а S – площадь всей области.
В трехмерном случае вероятность того, что точка, случайным образом расположенная в теле V, попадет в его часть v, задается формулой:

$$p = \frac{v}{V}, \qquad (2.1``)$$

где v – объем части тела, а V – объем всего тела.

Пример 1. Найти вероятность того, что точка, наудачу брошенная в круг, не попадет в правильный шестиугольник, вписанный в него.

Решение. Пусть радиус круга равен R , тогда сторона шестиугольника тоже равна R. При этом площадь круга $S = \pi R^2$, а площадь шестиугольника $s = \frac{3\sqrt{3}}{2}R^2$. Следовательно,

$$p = \frac{S - s}{S} = \frac{\pi R^2 - \frac{3\sqrt{3}}{2}R^2}{\pi R^2} = \frac{\pi - 3\sqrt{3}}{2\pi} \approx 0{,}174.$$

Пример 2. На отрезок *AB* случайным образом брошены три точки: *C, D* и *M*. Найти вероятность того, что из отрезков *AC, AD* и *AM* можно построить треугольник.

Решение. Обозначим длины отрезков *AC, AD* и *AM* через x, y и z и рассмотрим в качестве возможных исходов множество точек трехмерного пространства с координатами (x, y, z). Если принять длину отрезка равной 1, то эти множество возможных исходов представляет собой куб с ребром, равным 1. Тогда множество благоприятных исходов состоит из точек, для координат которых выполнены неравенства треугольника: $x + y > z$, $x + z > y$, $y + z > x$. Это часть куба, отрезанная от него плоскостями $x + y = z$, $x + z = y$, $y + z = x$

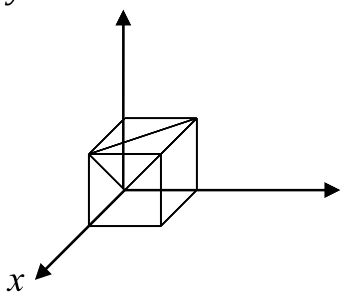


Рис.1.

(одна из них, плоскость $x + y = z$, проведена на рис.1). Каждая такая плоскость отделяет от куба пирамиду, объем которой равен $\frac{1}{3}\cdot\frac{1}{2}\cdot 1 = \frac{1}{6}$.

Следовательно, объем оставшейся части $v = 1 - 3\cdot\frac{1}{6} = \frac{1}{2}$. Тогда $p = \frac{v}{V} = \frac{1}{2} : 1 = \frac{1}{2}$.

Теорема сложения вероятностей.

Теорема 2.1 (теорема сложения).

Вероятность $p\,(A + B)$ суммы событий A и B равна

$$P\,(A + B) = p\,(A) + p\,(B) - p\,(AB). \qquad (2.2)$$

Доказательство.

Докажем теорему сложения для схемы случаев. Пусть n – число возможных исходов опыта, m_A – число исходов, благоприятных событию A, m_B – число исходов, благоприятных событию B, а m_{AB} – число исходов опыта, при которых происходят оба события (то есть исходов, благоприятных произведению AB). Тогда число исходов, при которых имеет место событие

$A + B$, равно $m_A + m_B - m_{AB}$ (так как в сумме $(m_A + m_B)$ m_{AB} учтено дважды: как исходы, благоприятные A, и исходы, благоприятные B). Следовательно, вероятность суммы можно определить по формуле (1.1):

$$p(A+B) = \frac{m_A + m_B - m_{AB}}{n} = \frac{m_A}{n} + \frac{m_B}{n} - \frac{m_{AB}}{n} = p(A) + p(B) - p(AB),$$

что и требовалось доказать.

Следствие 1. Теорему 2.1 можно распространить на случай суммы любого числа событий. Например, для суммы трех событий A, B и C

$$P(A + B + C) = p(A) + p(B) + p(C) - p(AB) - p(AC) - p(BC) + p(ABC) \qquad (2.3)$$

и т.д.

Следствие 2. Если события A и B несовместны, то $m_{AB} = 0$, и, следовательно, вероятность суммы несовместных событий равна сумме их вероятностей:

$$P(A + B) = p(A) + p(B). \qquad (2.4)$$

Определение 2.1. **Противоположными событиями** называют два несовместных события, образующих полную группу. Если одно из них назвать A, то второе принято обозначать $\bar{A}$.

Замечание. Таким образом, $\bar{A}$ заключается в том, что событие A не произошло.

Теорема 2.2. Сумма вероятностей противоположных событий равна 1:

$$p(A) + p(\bar{A}) = 1. \qquad (2.5)$$

Доказательство.

Так как A и $\bar{A}$ образуют полную группу, то одно из них обязательно произойдет в результате опыта, то есть событие $A + \bar{A}$ является достоверным. Следовательно,

$P(A + \bar{A}) = 1$. Но, так как A и $\bar{A}$ несовместны, из (2.4) следует, что $P(A + \bar{A}) = p(A) + p(\bar{A})$. Значит, $p(A) + p(\bar{A}) = 1$, что и требовалось доказать.

Замечание. В ряде задач проще искать не вероятность заданного события, а вероятность события, противоположного ему, а затем найти требуемую вероятность по формуле (2.5).

Пример. Из урны, содержащей 2 белых и 6 черных шаров, случайным образом извлекаются 5 шаров. Найти вероятность того, что вынуты шары разных цветов.

Решение. Событие $\bar{A}$, противоположное заданному, заключается в том, что из урны вынуто 5 шаров одного цвета, а так как белых шаров в ней всего два, то этот цвет может быть только черным. Множество возможных исходов опыта найдем по формуле (1.5):

$$n = C_8^5 = \frac{8!}{5!\cdot 3!} = \frac{6\cdot 7\cdot 8}{6} = 56,$$

а множество исходов, благоприятных событию $\overline{A}$ - это число возможных наборов по 5 шаров только из шести черных:

$$m_{\overline{A}} = C_6^5 = 6.$$

Тогда $p(\overline{A}) = \frac{6}{56} = \frac{3}{28}$, а $p(A) = 1 - \frac{3}{28} = \frac{25}{28}$.

Теорема умножения вероятностей.

Определение 2.2. Назовем **условной вероятностью** $\boldsymbol{p(B/A)}$ **события** $\boldsymbol{B}$ вероятность события B при условии, что событие A произошло.
Замечание. Понятие условной вероятности используется в основном в случаях, когда осуществление события A изменяет вероятность события B.
Примеры:

1) пусть событие A – извлечение из колоды в 32 карты туза, а событие B – то, что и вторая вынутая из колоды карта окажется тузом. Тогда, если после первого раза карта была возвращена в колоду, то вероятность вынуть вторично туз не меняется: $p(B) = p(A) = \frac{4}{32} = \frac{1}{8} = 0{,}125.$ Если же первая карта в колоду не возвращается, то осуществление события A приводит к тому, что в колоде осталась 31 карта, из которых только 3 туза. Поэтому $p(B/A) = \frac{3}{31} \approx 0{,}097.$
2) если событие A – попадание в самолет противника при первом выстреле из орудия, а B – при втором, то первое попадание уменьшает маневренность самолета, поэтому $p(B/A)$ увеличится по сравнению с $p(A)$.

Теорема 2.3 (теорема умножения). Вероятность произведения двух событий равна произведению вероятности одного из них на условную вероятность другого при условии, что первое событие произошло:

$$p\,(AB) = p\,(A) \cdot p\,(B/A). \tag{2.6}$$

Доказательство.
Воспользуемся обозначениями теоремы 2.1. Тогда для вычисления $p(B/A)$ множеством возможных исходов нужно считать m_A (так как A произошло), а множеством благоприятных исходов – те, при которых произошли и A, и B (m_{AB}). Следовательно,

$p(B/A) = \frac{m_{AB}}{m_A} = \frac{m_{AB}}{n} \cdot \frac{n}{m_A} = p(AB) : p(A),$ откуда следует утверждение теоремы.

Пример. Для поражения цели необходимо попасть в нее дважды. Вероятность первого попадания равна 0,2, затем она не меняется при промахах, но после первого попадания увеличивается вдвое. Найти вероятность того, что цель будет поражена первыми двумя выстрелами.

Решение. Пусть событие A – попадание при первом выстреле, а событие B – попадание при втором. Тогда $p(A) = 0{,}2$, $p(B/A) = 0{,}4$, $p(AB) = 0{,}2 \cdot 0{,}4 = 0{,}08$.

Следствие. Если подобным образом вычислить вероятность события BA, совпадающего с событием AB, то получим, что $p(BA) = p(B) \cdot p(A/B)$. Следовательно,

$$p(A) \cdot p(B/A) = p(B) \cdot p(A/B). \qquad (2.7)$$

Определение 2.3. Событие B называется **независимым** от события A, если появление события A не изменяет вероятности B, то есть $p(B/A) = p(B)$.
Замечание. Если событие B не зависит от A, то и A не зависит от B. Действительно, из (2.7) следует при этом, что $p(A) \cdot p(B) = p(B) \cdot p(A/B)$, откуда $p(A/B) = p(A)$. Значит, **свойство независимости событий взаимно**.

Теорема умножения для независимых событий имеет вид:

$$p(AB) = p(A) \cdot p(B), \qquad (2.8)$$

то есть вероятность произведения независимых событий равна произведению их вероятностей.
При решении задач теоремы сложения и умножения обычно применяются вместе.

Пример. Два стрелка делают по одному выстрелу по мишени. Вероятности их попадания при одном выстреле равны соответственно 0,6 и 0,7. Найти вероятности следующих событий:
A – хотя бы одно попадание при двух выстрелах;
B – ровно одно попадание при двух выстрелах;
C – два попадания;
D – ни одного попадания.
Решение. Пусть событие H_1 – попадание первого стрелка, H_2 – попадание второго. Тогда
$A = H_1 + H_2$, $B = H_1 \cdot \overline{H}_2 + \overline{H}_1 \cdot H_2$, $C = H_1 \cdot H_2$, $D = \overline{H}_1 \cdot \overline{H}_2$. События H_1 и H_2 совместны и независимы, поэтому теорема сложения применяется в общем виде, а теорема умножения – в виде (2.8). Следовательно,

$$p(C) = 0{,}6 \cdot 0{,}7 = 0{,}42, \qquad p(A) = 0{,}6 + 0{,}7 - 0{,}42 = 0{,}88,$$

$p(B) = 0{,}6 \cdot 0{,}3 + 0{,}7 \cdot 0{,}4 = 0{,}46$ (так как события $H_1 \cdot \overline{H}_2$ и $\overline{H}_1 \cdot H_2$ несовместны), $p(D) = 0{,}4 \cdot 0{,}3 = 0{,}12$. Заметим, что события A и D являются противоположными, поэтому $p(A) = 1 - p(D)$.

Вероятность появления хотя бы одного события.

Теорема 2.4. Вероятность появления хотя бы одного из попарно независимых событий $A_1, A_2, \ldots, A_n$ равна

$$p(A) = 1 - q_1 q_2 \ldots q_n, \qquad (2.9)$$

где q_i – вероятность события $\overline{A_i}$, противоположного событию A_i.

Доказательство.
Если событие A заключается в появлении хотя бы одного события из $A_1, A_2, \ldots, A_n$, то события A и $\overline{A}_1\overline{A}_2...\overline{A}_n$ противоположны, поэтому по теореме 2.2 сумма их вероятностей равна 1. Кроме того, поскольку $A_1, A_2, \ldots, A_n$ независимы, то независимы и $\overline{A}_1, \overline{A}_2, ..., \overline{A}_n$, следовательно,

$$p(\overline{A}_1\overline{A}_2...\overline{A}_n) = p(\overline{A}_1)p(\overline{A}_2)...p(\overline{A}_n) = q_1q_2...q_n.$$

Отсюда следует справедливость формулы (2.9).

Пример. Сколько нужно произвести бросков монеты, чтобы с вероятностью не менее 0,9 выпал хотя бы один герб?
Решение. Вероятность выпадения герба при одном броске равна вероятности противоположного события (выпадения цифры) и равна 0,5. Тогда вероятность выпадения хотя бы одного герба при n выстрелах равна $1- (0,5)^n$. Тогда из решения неравенства $1- (0,5)^n > 0,9$
следует, что $n > \log_2 10 \geq 4$.

Лекция 3.

Формула полной вероятности и формула Байеса. Схема и формула Бернулли. Приближение Пуассона для схемы Бернулли.

Определение 3.1. Пусть событие A может произойти только совместно с одним из событий $H_1, H_2, \ldots, H_n$, образующих полную группу несовместных событий. Тогда события $H_1, H_2, \ldots, H_n$ называются **гипотезами.**

Теорема 3.1. Вероятность события A, наступающего совместно с гипотезами $H_1, H_2, \ldots, H_n$, равна:

$$p(A) = \sum_{i=1}^{n} p(H_i)p(A/H_i), \tag{3.1}$$

где $p(H_i)$ – вероятность i- й гипотезы, а $p(A/H_i)$ – вероятность события A при условии реализации этой гипотезы. Формула (3.1) носит название **формулы полной вероятности**.
Доказательство.
Можно считать событие A суммой попарно несовместных событий $AH_1, AH_2, \ldots, AH_n$. Тогда из теорем сложения и умножения следует, что

$$p(A) = p(AH_1 + AH_2 + ... + AH_n) = p(AH_1) + p(AH_2) + ... + p(AH_n) = \sum_{i=1}^{n} p(H_i)p(A/H_i),$$

что и требовалось доказать.
Пример. Имеются три одинаковые урны с шарами. В первой из них 3 белых и 4 черных шара, во второй – 2 белых и 5 черных, в третьей – 10 черных шаров. Из случайно выбранной урны наудачу вынут шар. Найти вероятность того, что он белый.
Решение. Будем считать гипотезами H_1, H_2 и H_3 выбор урны с соответствующим номером. Так как по условию задачи все гипотезы

равновозможные, то $p(H_1)=p(H_2)=p(H_3)=\frac{1}{3}$. Найдем условную вероятность A при реализации каждой гипотезы:
$p(A/H_1)=\frac{3}{7}$, $p(A/H_2)=\frac{2}{7}$, $p(A/H_3)=0$. Тогда $p(A)=\frac{1}{3}\cdot\frac{3}{7}+\frac{1}{3}\cdot\frac{2}{7}+\frac{1}{3}\cdot 0=\frac{5}{21}\approx 0{,}238$.

Формула Байеса (теорема гипотез).

Пусть известен результат опыта, а именно то, что произошло событие A. Этот факт может изменить априорные (то есть известные до опыта) вероятности гипотез. Например, в предыдущем примере извлечение из урны белого шара говорит о том, что этой урной не могла быть третья, в которой нет белых шаров, то есть $p\,(H_3/A)=0$. Для переоценки вероятностей гипотез при известном результате опыта используется **формула Байеса**:

$$p(H_i/A)=\frac{p(H_i)p(A/H_i)}{p(A)}. \tag{3.2}$$

Действительно, из (2.7) получим, что $p(A)p(H_i/A)=p(H_i)p(A/H_i)$, откуда следует справедливость формулы (3.2).

Пример. После двух выстрелов двух стрелков, вероятности попаданий которых равны 0,6 и 0,7, в мишени оказалась одна пробоина. Найти вероятность того, что попал первый стрелок.
Решение. Пусть событие A – одно попадание при двух выстрелах, а гипотезы:

H_1 – первый попал, а второй промахнулся,
H_2 – первый промахнулся, а второй попал,
H_3 – оба попали,
H_4 – оба промахнулись.

Вероятности гипотез: $p(H_1)=0{,}6\cdot 0{,}3=0{,}18$,
$p(H_2)=0{,}4\cdot 0{,}7=0{,}28$,
$p(H_3)=0{,}6\cdot 0{,}7=0{,}42$,
$p(H_4)=0{,}4\cdot 0{,}3=0{,}12$.

Тогда $p(A/H_1)=p(A/H_2)=1$, $p(A/H_3)=p(A/H_4)=0$.
Следовательно, полная вероятность
$p(A)=0{,}18\cdot 1+0{,}28\cdot 1+0{,}42\cdot 0+0{,}12\cdot 0=0{,}46$. Применяя формулу Байеса, получим:

$$p(H_1/A)=\frac{0{,}18\cdot 1}{0{,}46}=\frac{9}{23}\approx 0{,}391.$$

Схема повторения испытаний. Формула Бернулли.

Рассмотрим серию из n испытаний, в каждом из которых событие A появляется с одной и той же вероятностью p, причем результат каждого испытания не зависит от результатов остальных. Подобная постановка задачи называется **схемой повторения испытаний**. Найдем вероятность того, что в такой серии событие A произойдет ровно $к$ раз (неважно, в какой последовательности). Интересующее нас событие представляет собой сумму равно-вероятных несовместных событий, заключающихся в том, что *A* произошло в некоторых *к* испытаниях и не произошло в остальных $n-\kappa$ испытаниях. Число таких событий равно числу сочетаний из *n* по *к*, то есть C_n^κ, а вероятность каждого из них: $p^k q^{n-k}$, где $q = 1 - p$ – вероятность того, что в данном опыте *A* не произошло. Применяя теорему сложения для несовместных событий, получим **формулу Бернулли**:

$$p_n(k) = C_n^k \cdot p^k \cdot q^{n-k}. \qquad (3.3)$$

Пример. Для получения приза нужно собрать 5 изделий с особым знаком на этикетке. Найти вероятность того, что придется купить 10 изделий, если этикетки с этим знаком имеют 5% изделий.

Решение. Из постановки задачи следует, что последнее купленное изделие имеет особый знак. Следовательно, из предыдущих девяти эти знаки имели 4 изделия. Найдем вероятность этого по формуле Бернулли:

$p_9(4) = C_9^4 \cdot (0{,}05)^4 \cdot (0{,}95)^5 = 0{,}0006092$ Тогда

$p = 0{,}0006092 \cdot 0{,}05 = 0{,}0000304$.

Приближение Пуассона для схемы Бернулли.

Формула Бернулли требует громоздких расчетов при большом количестве испытаний. Можно получить более удобную для расчетов приближенную формулу, если при большом числе испытаний вероятность появления *A* в одном опыте мала, а произведение $np = \lambda$ сохраняет постоянное значение для разных серий опытов (то есть среднее число появлений события *A* в разных сериях испытаний остается неизменным). Применим формулу Бернулли:

$$p_n(k) = \frac{n(n-1)(n-2)...(n-k+1)}{k!} p^k (1-p)^{n-k} = \frac{n(n-1)...(n-k+1)}{k!} \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}.$$

Найдем предел полученного выражения при $n \to \infty$:

$$p_n(k) \approx \frac{\lambda^k}{k!} \lim_{n\to\infty} \left(1 \cdot \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)...\left(1-\frac{k-1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{n-k}\right) = \frac{\lambda^k}{k!} \lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-k} = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \cdot 1.$$

Таким образом, **формула Пуассона**

$$p_n(k) = \frac{\lambda^k e^{-\lambda}}{k!} \qquad (3.4)$$

позволяет найти вероятность *к* появлений события *A* для массовых (*n* велико) и редких (*p* мало) событий.

Лекция 4.

Случайные величины. Закон распределения и функция распределения дискретной случайной величины. Биномиальное распределение и распределение Пуассона.

Наряду с понятием случайного события в теории вероятности используется и более удобное понятие *случайной величины*.

Определение 4.1. **Случайной величиной** называется величина, принимающая в результате опыта одно из своих возможных значений, причем заранее неизвестно, какое именно.

Будем обозначать случайные величины заглавными буквами латинского алфавита (X, Y,Z,…), а их возможные значения – соответствующими малыми буквами (x_i, y_i, …).

Примеры: число очков, выпавших при броске игральной кости; число появлений герба при 10 бросках монеты; число выстрелов до первого попадания в цель; расстояние от центра мишени до пробоины при попадании.

Можно заметить, что множество возможных значений для перечисленных случайных величин имеет разный вид: для первых двух величин оно конечно (соответственно 6 и 11 значений), для третьей величины множество значений бесконечно и представляет собой множество натуральных чисел, а для четвертой – все точки отрезка, длина которого равна радиусу мишени. Таким образом, для первых трех величин множество значений из отдельных (дискретных), изолированных друг от друга значений, а для четвертой оно представляет собой непрерывную область. По этому показателю случайные величины подразделяются на две группы: дискретные и непрерывные.

Определение 4.2. Случайная величина называется **дискретной**, если она принимает отдельные, изолированные возможные значения с определенными вероятностями.

Определение 4.3. Случайная величина называется **непрерывной**, если множество ее возможных значений целиком заполняет некоторый конечный или бесконечный промежуток.

Дискретные случайные величины.

Для задания дискретной случайной величины нужно знать ее возможные значения и вероятности, с которыми принимаются эти значения. Соответствие между ними называется **законом распределения** случайной величины. Он может иметь вид таблицы, формулы или графика.

Таблица, в которой перечислены возможные значения дискретной случайной величины и соответствующие им вероятности, называется **рядом распределения**:

x_i	x_1	x_2	…	x_n	…
p_i	p_1	p_2	…	p_n	…

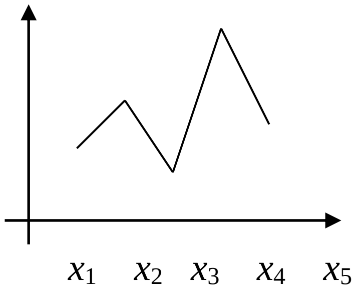
Заметим, что событие, заключающееся в том, что случайная величина примет одно из своих возможных значений, является достоверным, поэтому $\sum_{i=1}^{n(\infty)} p_i = 1$.

Пример. Два стрелка делают по одному выстрелу по мишени. Вероятности их попадания при одном выстреле равны соответственно 0,6 и 0,7. Составить ряд распределения случайной величины X – числа попаданий после двух выстрелов.

Решение. Очевидно, что X может принимать три значения: 0, 1 и 2. Их вероятности найдены в примере, рассмотренном в лекции 3. Следовательно, ряд распределения имеет вид:

x_i	0	1	2
p_i	0,12	0,46	0,42

Графически закон распределения дискретной случайной величины можно представить в виде **многоугольника распределения** – ломаной, соединяющей точки плоскости с координатами (x_i, p_i).



Функция распределения.

Определение 4.4. **Функцией распределения *F*(*x*)** случайной величины X называется вероятность того, что случайная величина примет значение, меньшее x:

$$F\,(x) = p\,(X < x). \qquad (4.1)$$

Свойства функции распределения.

1) $0 \le F(x) \le 1$.

Действительно, так как функция распределения представляет собой вероятность, она может принимать только те значения, которые принимает вероятность.

2) Функция распределения является неубывающей функцией, то есть

$$F(x_2) \ge F(x_1) \text{ при } x_2 > x_1.$$

Это следует из того, что

$$F(x_2) = p(X < x_2) = p(X < x_1) + p(x_1 \le X < x_2) \ge F(x_1).$$

3) $\lim\limits_{x\to-\infty} F(x)=0,\ \lim\limits_{x\to+\infty} F(x)=1.$ В частности, если все возможные значения X лежат на интервале $[a, b]$, то $F(x) = 0$ при $x \le a$ и $F(x) = 1$ при $x \ge b$. Действительно, $X < a$ – событие невозможное, а $X < b$ – достоверное.
4) Вероятность того, что случайная величина примет значение из интервала $[a, b]$, равна разности значений функции распределения на концах интервала:

$$p\ (\ a < X < b\ ) = F(b) - F(a).$$

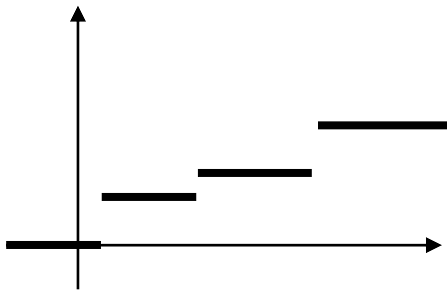
Справедливость этого утверждения следует из определения функции распределения (см. свойство 2).

Для дискретной случайной величины значение $F(x)$ в каждой точке представляет собой сумму вероятностей тех ее возможных значений, которые меньше аргумента функции.

Пример. Найдем $F(x)$ для предыдущего примера:

$$F(x)=\begin{cases} 0,\ \ x\le 0 \\ 0{,}12,\ \ 0<x\le 1 \\ 0{,}12+0{,}46=0{,}58,\ \ 1<x\le 2 \\ 0{,}58+0{,}42=1,\ \ x>2 \end{cases}$$

Соответственно график функции распределения имеет ступенчатый вид:



Биномиальное распределение.

Вернемся к схеме независимых испытаний и найдем закон распределения случайной величины X – числа появлений события A в серии из n испытаний. Возможные значения A: 0, 1, …, n. Соответствующие им вероятности можно вычислить по формуле Бернулли:

$$p(X=k)=C_n^k p^k q^{n-k} \qquad (4.2)$$

(p – вероятность появления A в каждом испытании).

Такой закон распределения называют **биномиальным**, поскольку правую часть равенства (4.2) можно рассматривать как общий член разложения бинома Ньютона:

$$(p+q)^n = C_n^n p^n + C_n^{n-1} p^{n-1} q + \ldots + C_n^k p^k q^{n-k} + \ldots + C_n^0 q^n.$$

Пример. Составим ряд распределения случайной величины X – числа попаданий при 5 выстрелах, если вероятность попадания при одном выстреле равна 0,8.

$p(X{=}0) = 1\cdot(0{,}2)^5 = 0{,}00032;$

$p(X=1) = 5\cdot0{,}8\cdot(0{,}2)^4 = 0{,}0064;$
$p(X=2) = 10\cdot(0{,}8)^2\cdot(0{,}2)^3 = 0{,}0512;$
$p(X=3) = 10\cdot(0{,}8)^3\cdot(0{,}2)^2 = 0{,}2048;$
$p(X=4) = 5\cdot(0{,}8)^4\cdot0{,}2 = 0{,}4096;$
$p(X=5) = 1\cdot(0{,}8)^5 = 0{,}32768.$

Таким образом, ряд распределения имеет вид:

x	0	1	2	3	4	5
p	0.00032	0.0064	0.0512	0.2048	0.4096	0.32728

Распределение Пуассона.

Рассмотрим дискретную случайную величину X, принимающую только целые неотрицательные значения (0, 1, 2,…, m,…), последовательность которых не ограничена. Такая случайная величина называется распределенной **по закону Пуассона**, если вероятность того, что она примет значение m, выражается формулой:

$$p(X=m) = \frac{a^m}{m!}e^{-a}, \qquad (4.3)$$

где a – некоторая положительная величина, называемая *параметром* закона Пуассона.

Покажем, что сумма всех вероятностей равна 1:

$$\sum_{m=0}^{\infty} p(X=m) = e^{-a}\sum_{m=0}^{\infty}\frac{a^m}{m!} = e^{-a}\cdot e^{a} = 1$$

(использовано разложение в ряд Тейлора функции e^x).

Рассмотрим типичную задачу, приводящую к распределению Пуассона. Пусть на оси абсцисс случайным образом распределяются точки, причем их распределение удовлетворяет следующим условиям:

1) вероятность попадания некоторого количества точек на отрезок длины l зависит только от длины отрезка и не зависит от его расположения на оси (то есть точки распределены с одинаковой средней плотностью);
2) точки распределяются независимо друг от друга (вероятность попадания какого-либо числа точек на данный отрезок не зависит от количества точек, попавший на любой другой отрезок);
3) практическая невозможность совпадения двух или более точек.

Тогда случайная величина X – число точек, попадающих на отрезок длины l – распределена по закону Пуассона, где a – среднее число точек, приходящееся на отрезок длины l.

Замечание. В лекции 3 говорилось о том, что формула Пуассона выражает биномиальное распределение при большом числе опытов и малой вероятности события. Поэтому закон Пуассона часто называют *законом редких явлений*.

Лекция 5.
Функция распределения и плотность распределения непрерывной случайной величины, их взаимосвязь и свойства. Равномерное распределение вероятностей.

Определение и свойства функции распределения сохраняются и для непрерывной случайной величины, для которой функцию распределения можно считать одним из видов задания закона распределения. Но для непрерывной случайной величины вероятность каждого отдельного ее значения равна 0. Это следует из свойства 4 функции распределения: $p(X = a) = F(a) - F(a) = 0$. Поэтому для такой случайной величины имеет смысл говорить только о вероятности ее попадания в некоторый интервал.

Вторым способом задания закона распределения непрерывной случайной величины является так называемая плотность распределения (плотность вероятности, дифференциальная функция).

Определение 5.1. Функция $f(x)$, называемая **плотностью распределения** непрерывной случайной величины, определяется по формуле:

$$f(x) = F'(x), \tag{5.1}$$

то есть является производной функции распределения.

Свойства плотности распределения.

1) $f(x) \geq 0$, так как функция распределения является неубывающей.
2) $F(x) = \int\limits_{-\infty}^{x} f(t)dt$, что следует из определения плотности распределения.
3) Вероятность попадания случайной величины в интервал (a, b) определяется формулой $p(a < X < b) = \int\limits_{a}^{b} f(x)dx.$

   Действительно, $p(a < X < b) = F(b) - F(a) = \int\limits_{-\infty}^{b} f(x)dx - \int\limits_{-\infty}^{a} f(x)dx = \int\limits_{a}^{b} f(x)dx.$
4) $\int\limits_{-\infty}^{+\infty} f(x)dx = 1$ (условие нормировки). Его справедливость следует из того, что $\int\limits_{-\infty}^{+\infty} f(x)dx = F(+\infty)$, а $\lim\limits_{x\to+\infty} F(x) = 1.$
5) $\lim\limits_{x\to\pm\infty} f(x) = 0,$ так как $F(x) \to const$ при $x \to \pm\infty.$

Таким образом, график плотности распределения представляет собой кривую, расположенную выше оси Оx, причем эта ось является ее горизонтальной асимптотой при $x \to \pm\infty$ (последнее справедливо только для случайных величин, множеством возможных значений которых является все множество действительных чисел). Площадь криволинейной трапеции, ограниченной графиком этой функции, равна единице.

Замечание. Если все возможные значения непрерывной случайной величины сосредоточены на интервале $[a, b]$, то все интегралы вычисляются в этих пределах, а вне интервала $[a, b]$ $f(x) \equiv 0$.

Пример 1. Плотность распределения непрерывной случайной величины задана формулой

$$f(x) = \frac{C}{1+x^2}, \ -\infty < x < +\infty.$$

Найти:

а) значение константы C;

б) вид функции распределения;

в) p (-1 < x < 1).

Решение. а) значение константы C найдем из свойства 4:

$$\int_{-\infty}^{+\infty} \frac{C}{1+x^2} dx = Carctgx \Big|_{-\infty}^{+\infty} = C\left(\frac{\pi}{2} + \frac{\pi}{2}\right) = C\pi = 1, \text{ откуда } C = \frac{1}{\pi}.$$

б) $F(x) = \frac{1}{\pi}\int_{-\infty}^{x} \frac{1}{1+t^2} dt = \frac{1}{\pi} arctg\ t \Big|_{-\infty}^{x} = \frac{1}{\pi}\left(arctgx + \frac{\pi}{2}\right) = \frac{1}{\pi} arctgx + \frac{1}{2}.$

в) $p(-1 < x < 1) = \frac{1}{\pi}\int_{-1}^{1} \frac{1}{1+x^2} dx = \frac{1}{\pi} arctgx \Big|_{-1}^{1} = \frac{1}{\pi}\left(\frac{\pi}{4} + \frac{\pi}{4}\right) = 0{,}5.$

Пример 2. Функция распределения непрерывной случайной величины имеет вид:

$$F(x) = \begin{cases} 0, & x \le 2 \\ \dfrac{x-2}{2}, & 2 < x \le 4 \\ 1, & x > 4. \end{cases}$$

Найти плотность распределения.

Решение.

$$f(x) = \begin{cases} 0', & x \le 2 \\ \left(\dfrac{x-2}{2}\right)', & 2 < x \le 4 \\ 1', & x > 4 \end{cases} = \begin{cases} 0, & x \le 2 \\ 0{,}5, & 2 < x \le 4 \\ 0, & x > 4. \end{cases}$$

Равномерный закон распределения.

Часто на практике мы имеем дело со случайными величинами, распределенными определенным типовым образом, то есть такими, закон распределения которых имеет некоторую стандартную форму. В прошлой лекции были рассмотрены примеры таких законов распределения для дискретных случайных величин (биномиальный и Пуассона). Для непрерывных случайных величин тоже существуют часто встречающиеся виды закона распределения, и в качестве первого из них рассмотрим равномерный закон.

Определение 5.2. Закон распределения непрерывной случайной величины называется **равномерным**, если на интервале, которому принадлежат все возможные значения случайной величины, плотность распределения сохраняет постоянное значение

$$f(x) = \text{const при } a \le x \le b,$$
$$f(x) = 0 \text{ при } x < a,\ x > b.$$

Найдем значение, которое принимает $f(x)$ при $x \in [a, b]$. Из условия нормировки следует, что $\int_a^b f(x)dx = \int_a^b c dx = c(b-a) = 1$, откуда $f(x) = c = \frac{1}{b-a}$.

Вероятность попадания равномерно распределенной случайной величины на интервал $[\alpha, \beta]$ $(a \le \alpha < \beta \le b)$ равна при этом $\int_\alpha^\beta \frac{1}{b-a} dx = \frac{\beta - \alpha}{b-a}$.

Вид функции распределения для нормального закона: $F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \le x \le b \\ 1, & x > b. \end{cases}$

Пример. Автобусы некоторого маршрута идут с интервалом 5 минут. Найти вероятность того, что пришедшему на остановку пассажиру придется ожидать автобуса не более 2 минут.

Решение. Время ожидания является случайной величиной, равномерно распределенной в интервале [0, 5]. Тогда $f(x) = \frac{1}{5}$, $p(0 \le x \le 2) = \frac{2}{5} = 0{,}4$.

Лекция 6.

Нормальный закон распределения вероятностей. Нормальная кривая. Функция Лапласа. Вычисление вероятности попадания в заданный интервал нормальной случайной величины. Правило трех сигм. Показательное распределение. Функция надежности. Показательный закон надежности.

Определение 6.1. Непрерывная случайная величина называется распределенной по **нормальному закону**, если ее плотность распределения имеет вид:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}. \qquad (6.1)$$

Замечание. Таким образом, нормальное распределение определяется двумя параметрами: a и σ.

График плотности нормального распределения называют **нормальной кривой (кривой Гаусса)**. Выясним, какой вид имеет эта кривая, для чего исследуем функцию (6.1).

1) Область определения этой функции: (-∞, +∞).
2) $f(x) > 0$ при любом x (следовательно, весь график расположен выше оси Оx).
3) $\lim\limits_{|x|\to\infty} f(x) = 0,$ то есть ось Оx служит горизонтальной асимптотой графика при $x \to \pm\infty$.
4) $f'(x) = -\dfrac{x-a}{\sigma^3\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} = 0$ при $x = a$;

   $f'(x) > 0$ при $x > a$, $f'(x) < 0$ при $x < a$. Следовательно, $\left(a, \dfrac{1}{\sigma\sqrt{2\pi}}\right)$ - точка максимума.
5) $f(x - a) = f(a - x)$, то есть график симметричен относительно прямой $x = a$.
6) $f''(x) = -\dfrac{1}{\sigma^3\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} \left(1 - \dfrac{(x-a)^2}{\sigma^2}\right) = 0$ при $x = a \pm \sigma$, то есть точки $\left(a \pm \sigma, \dfrac{1}{\sigma\sqrt{2\pi e}}\right)$ являются точками перегиба.

Примерный вид кривой Гаусса изображен на рис.1.

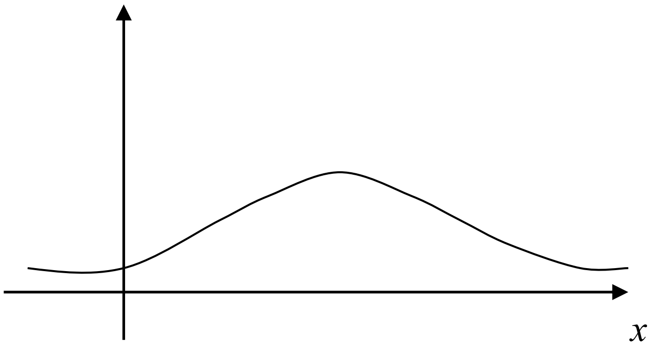


Рис.1.

Найдем вид функции распределения для нормального закона:

$$F(x) = \int_{-\infty}^{x} f(t)dt = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-a)^2}{2\sigma^2}} dt. \qquad (6.2)$$

Перед нами так называемый «неберущийся» интеграл, который невозможно выразить через элементарные функции. Поэтому для вычисления значений $F(x)$ приходится пользоваться таблицами. Они составлены для случая, когда $a = 0$, а $\sigma = 1$.

Определение 6.2. Нормальное распределение с параметрами $a = 0$, $\sigma = 1$ называется **нормированным**, а его функция распределения

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt \quad - \qquad (6.3)$$

- **функцией Лапласа**.

Замечание. Функцию распределения для произвольных параметров можно выразить через функцию Лапласа, если сделать замену: $t = \dfrac{x-a}{\sigma}$, тогда

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-a}{\sigma}} e^{-\frac{t^2}{2}} dt.$$

Найдем вероятность попадания нормально распределенной случайной величины на заданный интервал:

$$p(\alpha < x < \beta) = F(\beta) - F(\alpha) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right). \qquad (6.4)$$

Пример. Случайная величина X имеет нормальное распределение с параметрами a = 3, σ = 2. Найти вероятность того, что она примет значение из интервала (4, 8).

Решение.

$$p(4 < x < 8) = F(8) - F(4) = \Phi\left(\frac{8-3}{2}\right) - \Phi\left(\frac{4-3}{2}\right). = \Phi(2{,}5) - \Phi(0{,}5) = 0{,}9938 - 0{,}6915 = 0{,}3023.$$

Правило «трех сигм».

Найдем вероятность того, что нормально распределенная случайная величина примет значение из интервала (a - 3σ, a + 3σ):

$$p(a - 3\sigma < x < a + 3\sigma) = \Phi(3) - \Phi(-3) = 0{,}9986 - 0{,}0014 = 0.9973.$$

Следовательно, вероятность того, что значение случайной величины окажется *вне* этого интервала, равна 0,0027, то есть составляет 0,27% и может считаться пренебрежимо малой. Таким образом, на практике можно считать, что *все* возможные значения нормально распределенной случайной величины лежат в интервале (a - 3σ, a + 3σ).

Полученный результат позволяет сформулировать **правило «трех сигм»**: *если случайная величина распределена нормально, то модуль ее отклонения от x = a не превосходит 3σ.*

Показательное распределение.

Определение 6.3. **Показательным (экспоненциальным)** называют распределение вероятностей непрерывной случайной величины X, которое описывается плотностью

$$f(x) = \begin{cases} 0, & x < 0 \\ \lambda e^{-\lambda x}, & x \ge 0. \end{cases} \qquad (6.5)$$

В отличие от нормального распределения, показательный закон определяется только одним параметром λ. В этом его преимущество, так как обычно параметры распределения заранее не известны и их приходится оценивать приближенно. Понятно, что оценить один параметр проще, чем несколько.

Найдем функцию распределения показательного закона:

$F(x)=\int\limits_{-\infty}^{x} f(t)dt=\int\limits_{-\infty}^{0} 0\cdot dt+\lambda\int\limits_{0}^{x} e^{-\lambda t}dt=1-e^{-\lambda x}$. Следовательно,

$$F(x)=\begin{cases} 0, & x<0 \\ 1-e^{-\lambda x}, & x\ge 0. \end{cases} \qquad (6.6)$$

Теперь можно найти вероятность попадания показательно распределенной случайной величины в интервал (a, b):

$$p(a<x<b)=e^{-\lambda a}-e^{-\lambda b}. \qquad (6.7)$$

Значения функции e^{-x} можно найти из таблиц.

Функция надежности.

Пусть *элемент* (то есть некоторое устройство) начинает работать в момент времени $t_0 = 0$ и должен проработать в течение периода времени t. Обозначим за T непрерывную случайную величину – время безотказной работы элемента, тогда функция $F(t) = p(T > t)$ определяет вероятность отказа за время t. Следовательно, вероятность безотказной работы за это же время равна

$$R(t) = p(T > t) = 1 – F(t). \qquad (6.8)$$

Эта функция называется **функцией надежности**.

Показательный закон надежности.

Часто длительность безотказной работы элемента имеет показательное распределение, то есть $F(t) = 1 – e^{-\lambda t}$.

Следовательно, функция надежности в этом случае имеет вид:

$$R(t) = 1 – F(t) = 1 – (1 – e^{-\lambda t}) = e^{-\lambda t}.$$

Определение 6.4. **Показательным законом надежности** называют функцию надежности, определяемую равенством

$$R(t) = e^{-\lambda t}, \qquad (6.9)$$

где λ – интенсивность отказов.

Пример. Пусть время безотказной работы элемента распределено по показательному закону с плотностью распределения

$f(t) = 0,1\, e^{-0,1t}$ при $t \ge 0$.

Найти вероятность того, что элемент проработает безотказно в течение 10 часов.

Решение. Так как $\lambda = 0,1$, $R(10) = e^{-0,1\cdot 10} = e^{-1} = 0,368$.

Лекция 7.

Основные числовые характеристики дискретных и непрерывных случайных величин: математическое ожидание, дисперсия и среднее квадратическое отклонение. Их свойства и примеры.

Закон распределения (функция распределения и ряд распределения или плотность вероятности) полностью описывают поведение случайной величины. Но в ряде задач достаточно знать некоторые числовые характеристики исследуемой величины (например, ее среднее значение и возможное отклонение от него), чтобы ответить на поставленный вопрос. Рассмотрим основные числовые характеристики дискретных случайных величин.

Математическое ожидание.

Определение 7.1. **Математическим ожиданием** дискретной случайной величины называется сумма произведений ее возможных значений на соответствующие им вероятности:

$$M(X) = x_1p_1 + x_2p_2 + \ldots + x_np_n \,. \qquad (7.1)$$

Если число возможных значений случайной величины бесконечно, то

$$M(X) = \sum_{i=1}^{\infty} x_i p_i \,,$$

если полученный ряд сходится абсолютно.

Замечание 1. Математическое ожидание называют иногда *взвешенным средним*, так как оно приближенно равно среднему арифметическому наблюдаемых значений случайной величины при большом числе опытов.

Замечание 2. Из определения математического ожидания следует, что его значение не меньше наименьшего возможного значения случайной величины и не больше наибольшего.

Замечание 3. Математическое ожидание дискретной случайной величины есть *неслучайная* (постоянная) величина. В дальнейшем увидим, что это же справедливо и для непрерывных случайных величин.

Пример 1. Найдем математическое ожидание случайной величины X – числа стандартных деталей среди трех, отобранных из партии в 10 деталей, среди которых 2 бракованных. Составим ряд распределения для X. Из условия задачи следует, что X может принимать значения 1, 2, 3. $p(1) = \dfrac{C_8^1 \cdot C_2^2}{C_{10}^3} = \dfrac{1}{15}$, $p(2) = \dfrac{C_8^2 \cdot C_2^1}{C_{10}^3} = \dfrac{7}{15}$, $p(3) = \dfrac{C_8^3}{C_{10}^3} = \dfrac{7}{15}$.

Тогда

$$M(X) = 1 \cdot \frac{1}{15} + 2 \cdot \frac{7}{15} + 3 \cdot \frac{7}{15} = 2{,}4.$$

Пример 2. Определим математическое ожидание случайной величины X – числа бросков монеты до первого появления герба. Эта величина может принимать бесконечное число значений (множество возможных значений есть множество натуральных чисел). Ряд ее распределения имеет вид:

X	1	2	…	n	…
p	0,5	$(0,5)^2$	…	$(0,5)^n$	…

Тогда

$$M(X)=\sum_{n=1}^{\infty}\frac{n}{2^n}=\frac{1}{2}+2\cdot\left(\frac{1}{2}\right)^2+3\cdot\left(\frac{1}{2}\right)^3+\ldots+n\cdot\left(\frac{1}{2}\right)^n+\ldots=\sum_{n=1}^{\infty}\frac{1}{2^n}+\frac{1}{2}\sum_{n=1}^{\infty}\frac{1}{2^n}+\ldots+$$

$$+\frac{1}{2^n}\sum_{n=1}^{\infty}\frac{1}{2^n}+\ldots=1\cdot\left(1+\frac{1}{2}+\frac{1}{4}+\ldots+\frac{1}{2^n}+\ldots\right)=1\cdot 2=2$$

(при вычислении дважды использовалась формула суммы бесконечно убывающей геометрической прогрессии: $S=\frac{b_1}{1-q}$, откуда $\frac{1}{2}+\frac{1}{4}+\ldots+\frac{1}{2^n}+\ldots=1,\ 1+\frac{1}{2}+\frac{1}{4}+\ldots+\frac{1}{2^n}+\ldots=2$).

Свойства математического ожидания.

1) Математическое ожидание постоянной равно самой постоянной:

$$M(C) = C. \tag{7.2}$$

Доказательство. Если рассматривать C как дискретную случайную величину, принимающую только одно значение C с вероятностью $p = 1$, то $M(C) = C\cdot 1 = C$.

2) Постоянный множитель можно выносит за знак математического ожидания:

$$M(CX) = C\,M(X). \tag{7.3}$$

Доказательство. Если случайная величина X задана рядом распределения

x_i	x_1	x_2	…	x_n
p_i	p_1	p_2	…	p_n

то ряд распределения для CX имеет вид:

Cx_i	Cx_1	Cx_2	…	Cx_n
p_i	p_1	p_2	…	p_n

Тогда

$M(CX) = Cx_1p_1 + Cx_2p_2 + \ldots + Cx_np_n = C(x_1p_1 + x_2p_2 + \ldots + x_np_n) = CM(X)$.

Определение 7.2. Две случайные величины называются **независимыми**, если закон распределения одной из них не зависит от того, какие значения приняла другая. В противном случае случайные величины **зависимы**.

Определение 7.3. Назовем **произведением независимых случайных величин *X* и *Y*** случайную величину XY, возможные значения которой равны произведениям всех возможных значений X на все возможные значения Y, а соответствующие им вероятности равны произведениям вероятностей сомножителей.

3) Математическое ожидание произведения двух независимых случайных величин равно произведению их математических ожиданий:

$$M(XY) = M(X)M(Y). \qquad (7.4)$$

Доказательство. Для упрощения вычислений ограничимся случаем, когда X и Y принимают только по два возможных значения:

x_i	x_1	x_2
p_i	p_1	p_2

y_i	y_1	y_2
g_i	g_1	g_2

Тогда ряд распределения для XY выглядит так:

XY	x_1y_1	x_2y_1	x_1y_2	x_2y_2
p	p_1g_1	$p_2\,g_1$	p_1g_2	p_2g_2

Следовательно,

$M(XY) = x_1y_1 \cdot p_1g_1 + x_2y_1 \cdot p_2g_1 + x_1y_2 \cdot p_1g_2 + x_2y_2 \cdot p_2g_2 = y_1g_1(x_1p_1 + x_2p_2) +$
$+ y_2g_2(x_1p_1 + x_2p_2) = (y_1g_1 + y_2g_2)(x_1p_1 + x_2p_2) = M(X) \cdot M(Y).$

Замечание 1. Аналогично можно доказать это свойство для большего количества возможных значений сомножителей.

Замечание 2. Свойство 3 справедливо для произведения любого числа независимых случайных величин, что доказывается методом математической индукции.

Определение 7.4. Определим **сумму случайных величин X и Y** как случайную величину $X + Y$, возможные значения которой равны суммам каждого возможного значения X с каждым возможным значением Y; вероятности таких сумм равны произведениям вероятностей слагаемых (для зависимых случайных величин – произведениям вероятности одного слагаемого на условную вероятность второго).

4) Математическое ожидание суммы двух случайных величин (зависимых или независимых) равно сумме математических ожиданий слагаемых:

$$M(X + Y) = M(X) + M(Y). \qquad (7.5)$$

Доказательство.

Вновь рассмотрим случайные величины, заданные рядами распределения, приведенными при доказательстве свойства 3. Тогда возможными значениями $X + Y$ являются $x_1 + y_1$, $x_1 + y_2$, $x_2 + y_1$, $x_2 + y_2$.

Обозначим их вероятности соответственно, как p_{11}, p_{12}, p_{21} и p_{22}.

Найдем $M(X + Y) = (x_1 + y_1)p_{11} + (x_1 + y_2)p_{12} + (x_2 + y_1)p_{21} + (x_2 + y_2)p_{22} =$
$= x_1(p_{11} + p_{12}) + x_2(p_{21} + p_{22}) + y_1(p_{11} + p_{21}) + y_2(p_{12} + p_{22}).$

Докажем, что $p_{11} + p_{22} = p_1$. Действительно событие, состоящее в том, что $X + Y$ примет значения $x_1 + y_1$ или $x_1 + y_2$ и вероятность которого равна $p_{11} + p_{22}$, совпадает с событием, заключающемся в том, что $X = x_1$ (его вероятность p_1).

Аналогично доказывается, что $p_{21} + p_{22} = p_2$, $p_{11} + p_{21} = g_1$, $p_{12} + p_{22} = g_2$. Значит,

$$M(X + Y) = x_1p_1 + x_2p_2 + y_1g_1 + y_2g_2 = M(X) + M(Y).$$

Замечание. Из свойства 4 следует, что сумма любого числа случайных величин равна сумме математических ожиданий слагаемых.

Пример. Найти математическое ожидание суммы числа очков, выпавших при броске пяти игральных костей.

Найдем математическое ожидание числа очков, выпавших при броске одной кости:

$M(X_1) = (1 + 2 + 3 + 4 + 5 + 6)/6=3,5$

Тому же числу равно математическое ожидание числа очков, выпавших на любой кости. Следовательно, по свойству 4 $M(X)$=5/6

Дисперсия.

Для того, чтобы иметь представление о поведении случайной величины, недостаточно знать только ее математическое ожидание. Рассмотрим две случайные величины: X и Y, заданные рядами распределения вида

X	49	50	51
p	0,1	0,8	0,1

Y	0	100
p	0,5	0,5

Найдем $M(X) = 49\cdot0,1 + 50\cdot0,8 + 51\cdot0,1 = 50$, $M(Y) = 0\cdot0,5 + 100\cdot0,5 = 50$. Как видно, математические ожидания обеих величин равны, но если для X $M(X)$ хорошо описывает поведение случайной величины, являясь ее наиболее вероятным возможным значением (причем остальные значения ненамного отличаются от 50), то значения Y существенно отстоят от $M(Y)$. Следовательно, наряду с математическим ожиданием желательно знать, насколько значения случайной величины отклоняются от него. Для характеристики этого показателя служит дисперсия.

Определение 7.5. **Дисперсией (рассеянием)** случайной величины называется математическое ожидание квадрата ее отклонения от ее математического ожидания:

$$D(X) = M\,(X - M(X))^2. \qquad (7.6)$$

Пример.

Найдем дисперсию случайной величины X (числа стандартных деталей среди отобранных) в примере 1 данной лекции. Вычислим значения квадрата отклонения каждого возможного значения от математического ожидания:

$(1 - 2,4)^2 = 1,96$; $(2 - 2,4)^2 = 0,16$; $(3 - 2,4)^2 = 0,36$. Следовательно,

$$D(X) = 1{,}96 \cdot \frac{1}{15} + 0{,}16 \cdot \frac{7}{15} + 0{,}36 \cdot \frac{7}{15} = \frac{28}{75} \approx 0{,}373.$$

Замечание 1. В определении дисперсии оценивается не само отклонение от среднего, а его квадрат. Это сделано для того, чтобы отклонения разных знаков не компенсировали друг друга.
Замечание 2. Из определения дисперсии следует, что эта величина принимает только неотрицательные значения.
Замечание 3. Существует более удобная для расчетов формула для вычисления дисперсии, справедливость которой доказывается в следующей теореме:

Теорема 7.1.

$$D(X) = M(X^2) - M^2(X). \quad (7.7)$$

Доказательство.
Используя то, что $M(X)$ – постоянная величина, и свойства математического ожидания, преобразуем формулу (7.6) к виду:
$D(X) = M(X - M(X))^2 = M(X^2 - 2X \cdot M(X) + M^2(X)) = M(X^2) - 2M(X) \cdot M(X) + M^2(X) = M(X^2) - 2M^2(X) + M^2(X) = M(X^2) - M^2(X)$, что и требовалось доказать.

Пример. Вычислим дисперсии случайных величин X и Y, рассмотренных в начале этого раздела.
$M(X) = (49^2 \cdot 0{,}1 + 50^2 \cdot 0{,}8 + 51^2 \cdot 0{,}1) - 50^2 = 2500{,}2 - 2500 = 0{,}2.$
$M(Y) = (0^2 \cdot 0{,}5 + 100^2 \cdot 0{,}5) - 50^2 = 5000 - 2500 = 2500.$
Итак, дисперсия второй случайной величины в несколько тысяч раз больше дисперсии первой. Таким образом, даже не зная законов распределения этих величин, по известным значениям дисперсии мы можем утверждать, что X мало отклоняется от своего математического ожидания, в то время как для Y это отклонение весьма существенно.

Свойства дисперсии.

1) Дисперсия постоянной величины C равна нулю:

$$D(C) = 0. \quad (7.8)$$

Доказательство. $D(C) = M((C - M(C))^2) = M((C - C)^2) = M(0) = 0.$

2) Постоянный множитель можно выносить за знак дисперсии, возведя его в квадрат:

$$D(CX) = C^2 D(X). \quad (7.9)$$

Доказательство.
$D(CX) = M((CX - M(CX))^2) = M((CX - CM(X))^2) = M(C^2(X - M(X))^2) =$
$= C^2 D(X).$

3) Дисперсия суммы двух независимых случайных величин равна сумме их дисперсий:

$$D(X + Y) = D(X) + D(Y). \quad (7.10)$$

Доказательство.
$D(X+Y) = M(X^2+2XY+Y^2) - (M(X)+M(Y))^2 = M(X^2) + 2M(X)M(Y) +$
$+ M(Y^2) - M^2(X) - 2M(X)M(Y) - M^2(Y) = (M(X^2) - M^2(X)) + (M(Y^2) - M^2(Y)) =$
$D(X) + D(Y)$.

Следствие 1. Дисперсия суммы нескольких взаимно независимых случайных величин равна сумме их дисперсий.

Следствие 2. Дисперсия суммы постоянной и случайной величин равна дисперсии случайной величины.

4) Дисперсия разности двух независимых случайных величин равна сумме их дисперсий:

$$D(X-Y) = D(X) + D(Y). \qquad (7.11)$$

Доказательство.
$D(X-Y) = D(X) + D(-Y) = D(X) + (-1)^2 D(Y) = D(X) + D(Y)$.

Дисперсия дает среднее значение квадрата отклонения случайной величины от среднего; для оценки самого отклонения служит величина, называемая средним квадратическим отклонением.

Определение 7.6. **Средним квадратическим отклонением** σ случайной величины X называется квадратный корень из дисперсии:

$$\sigma = \sqrt{D(X)}. \qquad (7.12)$$

Пример. В предыдущем примере средние квадратические отклонения X и Y равны соответственно

$\sigma_x = \sqrt{0{,}2} \approx 0{,}447; \quad \sigma_y = \sqrt{2500} = 50.$

Числовые характеристики непрерывных случайных величин.

Распространим определения числовых характеристик случайных величин на непрерывные случайные величины, для которых плотность распределения служит в некотором роде аналогом понятия вероятности.

Определение 7.7. **Математическим ожиданием** непрерывной случайной величины называется

$$M(X) = \int_{-\infty}^{+\infty} x f(x)dx. \qquad (7.13)$$

Замечание 1. Общее определение дисперсии сохраняется для непрерывной случайной величины таким же, как и для дискретной (опр. 7.5), а формула для ее вычисления имеет вид:

$$D(X) = \int_{-\infty}^{+\infty} x^2 f(x)dx - M^2(X). \qquad (7.14)$$

Среднее квадратическое отклонение вычисляется по формуле (7.12).

Замечание 2. Если все возможные значения непрерывной случайной величины не выходят за пределы интервала $[a, b]$, то интегралы в формулах (7.13) и (7.14) вычисляются в этих пределах.

Пример. Плотность распределения случайной величины X имеет вид:

$$f(x) = \begin{cases} 0, & x < 2 \\ -\dfrac{3}{4}(x^2 - 6x + 8), & 2 \le x \le 4 \\ 0, & x > 4. \end{cases}$$

Найти $M(X)$, $D(X)$, σ.

$$M(X) = -\frac{3}{4}\int_2^4 x(x^2 - 6x + 8)dx = -\frac{3}{4}\left(\frac{x^4}{4} - 2x^3 + 4x^2\right)\Bigg|_2^4 = 3;$$

Решение.

$$D(X) = -\frac{3}{4}\int_2^4 x^2(x^2 - 6x + 8)dx - 9 = -\frac{3}{4}\left(\frac{x^5}{5} - \frac{3x^4}{2} + \frac{8x^3}{3}\right)\Bigg|_2^4 - 9 = 9{,}2 - 9 = 0{,}2; \quad \sigma = \sqrt{0{,}2} \approx 0{,}447.$$

Числовые характеристики случайных величин, имеющих некоторые стандартные законы распределения.

1. Биномиальное распределение.

Для дискретной случайной величины X, представляющей собой число появлений события A в серии из n независимых испытаний (см. лекцию 6), $M(X)$ можно найти, используя свойство 4 математического ожидания. Пусть X_1 – число появлений A в первом испытании, X_2 – во втором и т.д. При этом каждая из случайных величин X_i задается рядом распределения вида

X_i	0	1
p_i	q	p

Следовательно, $M(X_i) = p$.

Тогда

$$M(X) = \sum_{i=1}^{n} M(X_i) = \sum_{i=1}^{n} p = np.$$

Аналогичным образом вычислим дисперсию:

$D(X_i) = 0^2 \cdot q + 1^2 \cdot p - p^2 = p - p^2 = p(1 - p)$, откуда по свойству 4 дисперсии

$$D(X) = \sum_{i=1}^{n} D(X_i) = np(1 - p) = npq.$$

2. Закон Пуассона.

Если $p(X = m) = \dfrac{a^m}{m!}e^{-a}$, то

$M(X) = \displaystyle\sum_{m=1}^{\infty} m\frac{a^m}{m!}e^{-a} = ae^{-a}\sum_{m=1}^{\infty}\frac{a^{m-1}}{(m-1)!} = ae^{-a}e^{a} = a$ (использовалось разложение в ряд Тейлора функции e^x).

Для определения дисперсии найдем вначале

$M(X^2) = \displaystyle\sum_{m=1}^{\infty} m^2\frac{a^m}{m!}e^{-a} = a\sum_{m=1}^{\infty} m\frac{a^{m-1}}{(m-1)!}e^{-a} =$

$$= a\sum_{m=1}^{\infty}((m-1)+1)\frac{a^{m-1}}{(m-1)!}e^{-a} = a\left(\sum_{m=1}^{\infty}(m-1)\frac{a^{m-1}}{(m-1)!}e^{-a} + \sum_{m=1}^{\infty}\frac{a^{m-1}}{(m-1)!}e^{-a}\right) = a(a+1).$$

Поэтому $D(X) = a^2 + a - a^2 = a$.

Замечание. Таким образом, обнаружено интересное свойство распределения Пуассона: математическое ожидание равно дисперсии (и равно единственному параметру a, определяющему распределение).

3. Равномерное распределение.

Для равномерно распределенной на отрезке $[a, b]$ непрерывной случайной величины

$$M(X) = \int_a^b x\frac{1}{b-a}dx = \frac{x^2}{2(b-a)}\bigg|_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2},$$

то есть математическое ожидание равномерно распределенной случайной величины равно абсциссе середины отрезка $[a, b]$.

Дисперсия

$$D(X) = \int_a^b x^2\frac{1}{b-a}dx - \frac{(a+b)^2}{4} = \frac{b^3-a^3}{3(b-a)} - \frac{(a+b)^2}{4} = \frac{a^2+ab+b^2}{3} - \frac{a^2+2ab+b^2}{4} =$$

$$= \frac{(b-a)^2}{12}.$$

4. Нормальное распределение.

Для вычисления математического ожидания нормально распределенной случайной величины воспользуемся тем, что *интеграл Пуассона* $\int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}}dz = \sqrt{2\pi}$.

$$M(X) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty} xe^{-\frac{(x-a)^2}{2\sigma^2}}dx = (z = \frac{x-a}{\sigma}) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}(\sigma z + a)e^{-\frac{z^2}{2}}dz =$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\sigma z e^{-\frac{z^2}{2}}dz + \frac{a}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{z^2}{2}}dz = 0 + \frac{a}{\sqrt{2\pi}}\sqrt{2\pi} = a$$

(первое слагаемое равно 0, так как подынтегральная функция нечетна, а пределы интегрирования симметричны относительно нуля).

$$D(X) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}(x-a)^2 e^{-\frac{(x-a)^2}{2\sigma^2}}dx = \frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} z\cdot ze^{-\frac{z^2}{2}}dz = (u = z, dv = ze^{-\frac{z^2}{2}}) =$$

$$= \frac{\sigma^2}{\sqrt{2\pi}}\left(-z\cdot e^{-\frac{z^2}{2}}\bigg|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty}e^{-\frac{z^2}{2}}dz\right) = \frac{\sigma^2}{\sqrt{2\pi}}\left(-0+\sqrt{2\pi}\right) = \sigma^2.$$

Следовательно, параметры нормального распределения (a и σ) равны соответственно математическому ожиданию и среднему квадратическому отклонению исследуемой случайной величины.

Лекция 8.

Случайные векторы (системы нескольких случайных величин). Закон распределения вероятностей дискретной двумерной случайной величины. Функция распределения и плотность распределения двумерной случайной величины, их свойства. Вероятность попадания случайной точки в произвольную область. Отыскание плотностей вероятности составляющих двумерной случайной величины. Равномерное распределение на плоскости.

Наряду с одномерными случайными величинами, возможные значения которых определяются одним числом, теория вероятностей рассматривает и многомерные случайные величины. Каждое возможное значение такой величины представляет собой упорядоченный набор нескольких чисел. Геометрической иллюстрацией этого понятия служат точки n-мерного пространства, каждая координата которых является случайной величиной (дискретной или непрерывной), или n-мерные векторы. Поэтому многомерные случайные величины называют еще случайными векторами.

Двумерные случайные величины.

1. Дискретные двумерные случайные величины.

Закон распределения дискретной двумерной случайной величины (X, Y) имеет вид таблицы с двойным входом, задающей перечень возможных значений каждой компоненты и вероятности $p(x_i, y_j)$, с которыми величина принимает значение (x_i, y_j):

Y	X					
	x_1	x_2	…	x_i	…	x_n
y_1	$p(x_1, y_1)$	$p(x_2, y_1)$	…	$p(x_i, y_1)$	…	$p(x_n, y_1)$
…	…	…	…	…	…	…
y_j	$p(x_1, y_j)$	$p(x_2, y_j)$	…	$p(x_i, y_j)$	…	$p(x_n, y_j)$
…	…	…	…	…	…	…
y_m	$p(x_1, y_m)$	$p(x_2, y_m)$	…	$p(x_i, y_m)$	…	$p(x_n, y_m)$

При этом сумма вероятностей, стоящих во всех клетках таблицы, равна 1. Зная закон распределения двумерной случайной величины, можно найти законы распределения ее составляющих. Действительно, событие $X = x_1$ представляется собой сумму несовместных событий

$(X = x_1, Y = y_1), (X = x_1, Y = y_2),\ldots, (X = x_1, Y = y_m)$, поэтому

$p(X = x_1) = p(x_1, y_1) + p(x_1, y_2) + \ldots + p(x_1, y_m)$ (в правой части находится сумма вероятностей, стоящих в столбце, соответствующем $X = x_1$). Так же можно найти вероятности остальных возможных значений X.

Для определения вероятностей возможных значений Y нужно сложить вероятности, стоящие в строке таблицы, соответствующей $Y = y_j$.

Пример 1. Дан закон распределения двумерной случайной величины:

Y	X		
	-2	3	6
-0,8	0,1	0,3	0,1
-0,5	0,15	0,25	0,1

Найти законы распределения составляющих.

Решение. Складывая стоящие в таблице вероятности «по столбцам», получим ряд распределения для X:

X	-2	3	6
p	0,25	0,55	0,2

Складывая те же вероятности «по строкам», найдем ряд распределения для Y:

Y	-0,8	-0,5
p	0,5	0,5

2. Непрерывные двумерные случайные величины.

Определение 8.1. **Функцией распределения $F(x, y)$** двумерной случайной величины (X, Y) называется вероятность того, что $X < x$, а $Y < y$:

$$F(x, y) = p(X < x, Y < y). \qquad (8.1)$$

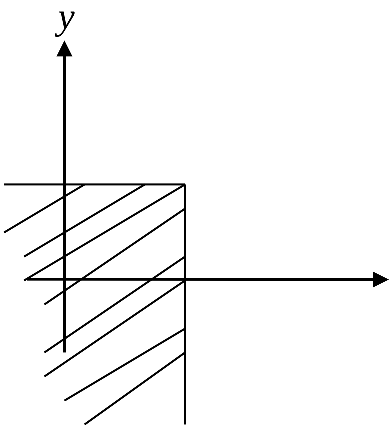


Рис.1.

Это означает, что точка (X, Y) попадет в область, заштрихованную на рис. 1, если вершина прямого угла располагается в точке (x, y).

Замечание. Определение функции распределения справедливо как для непрерывной, так и для дискретной двумерной случайной величины.

Свойства функции распределения.

1) $0 \le F(x, y) \le 1$ (так как $F(x, y)$ является вероятностью).

2) $F(x, y)$ есть неубывающая функция по каждому аргументу:

$$F(x_2, y) \geq F(x_1, y), \text{ если } x_2 > x_1;$$

$$F(x, y_2) \geq F(x, y_1), \text{ если } y_2 > y_1.$$

Доказательство.

$F(x_2, y) = p(X < x_2, Y < y) = p(X < x_1, Y < y) + p(x_1 \leq X < x_2, Y < y) \geq$
$\geq p(X < x_1, Y < y) = F(x_1, y)$.

Аналогично доказывается и второе утверждение.

3) Имеют место предельные соотношения:

a) $F(-\infty, y) = 0$; b) $F(x, -\infty) = 0$; c) $F(-\infty, -\infty) = 0$; d) $F(\infty, \infty) = 1$.

Доказательство. События a), b) и c) невозможны (так как невозможно событие $X<-\infty$ или $Y<-\infty$), а событие d) достоверно, откуда следует справедливость приведенных равенств.

4) При $y = \infty$ функция распределения двумерной случайной величины становится функцией распределения составляющей X:

$$F(x, \infty) = F_1(x).$$

При $x = \infty$ функция распределения двумерной случайной величины становится функцией распределения составляющей Y :

$$F(\infty, y) = F_2(y).$$

Доказательство. Так как событие $Y < \infty$ достоверно, то
$F(x, \infty) = p(X < x) = F_1(x)$.

Аналогично доказывается второе утверждение.

Определение 8.2. **Плотностью совместного распределения вероятностей (двумерной плотностью вероятности)** непрерывной двумерной случайной величины называется смешанная частная производная 2-го порядка от функции распределения:

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}. \qquad (8.2)$$

Замечание. Двумерная плотность вероятности представляет собой предел отношения вероятности попадания случайной точки в прямоугольник со сторонами Δx и Δy к площади этого прямоугольника при $\Delta x \to 0,\ \Delta y \to 0$.

Свойства двумерной плотности вероятности.

1) $f(x, y) \geq 0$ (см. предыдущее замечание: вероятность попадания точки в прямоугольник неотрицательна, площадь этого прямоугольника положительна, следовательно, предел их отношения неотрицателен).

2)

$$F(x, y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f(x, y)dxdy$$

(следует из определения двумерной плотности вероятности).

3)

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy = 1$$

(поскольку это вероятность того, что точка попадет на плоскость Оxy, то есть достоверного события).

Вероятность попадания случайной точки в произвольную область.

Пусть в плоскости Оxy задана произвольная область D. Найдем вероятность того, что точка, координаты которой представляют собой систему двух случайных величин (двумерную случайную величину) с плотностью распределения $f(x, y)$, попадет в область D. Разобьем эту область прямыми, параллельными осям координат, на прямоугольники со сторонами Δx и Δy.

Вероятность попадания в каждый такой прямоугольник равна $f(\xi_i,\eta_i)\Delta x\Delta y$, где (ξ_i,η_i) - координаты точки, принадлежащей прямоугольнику.

Тогда вероятность попадания точки в область D есть предел интегральной суммы $\sum_{i=1}^{n} f(\xi_i,\eta_i)\Delta x\Delta y$, то есть

$$p((X,Y)\subset D) = \iint_D f(x,y)dxdy. \tag{8.3}$$

Отыскание плотностей вероятности составляющих двумерной случайной величины.

Выше было сказано, как найти функцию распределения каждой составляющей, зная двумерную функцию распределения. Тогда по определению плотности распределения

$$f_1(x) = \frac{dF_1(x)}{dx} = \frac{dF(x,\infty)}{dx} = \frac{d\left(\int_{-\infty}^{x}\int_{-\infty}^{\infty} f(x,y)\right)}{dx} = \int_{-\infty}^{\infty} f(x,y)dy. \tag{8.4}$$

Аналогично находится

$$f_2(y) = \int_{-\infty}^{\infty} f(x,y)dx. \tag{8.4$'$}$$

Условные законы распределения составляющих дискретной двумерной случайной величины.

Рассмотрим дискретную двумерную случайную величину и найдем закон распределения составляющей X при условии, что Y примет определенное значение (например, $Y = y_1$). Для этого воспользуемся формулой Байеса, считая гипотезами события $X = x_1$, $X = x_2,\dots$, $X = x_n$, а событием A – событие $Y = y_1$. При такой постановке задачи нам требуется найти условные вероятности гипотез при условии, что A произошло. Следовательно,

$$p(x_i / y_1) = \frac{p(x_i, y_1)}{p(y_1)}.$$

Таким же образом можно найти вероятности возможных значений X при условии, что Y принимает любое другое свое возможное значение:

$$p(x_i / y_j) = \frac{p(x_i, y_j)}{p(y_j)}. \qquad (8.5)$$

Аналогично находят условные законы распределения составляющей Y:

$$p(y_j / x_i) = \frac{p(x_i, y_j)}{p(x_i)}. \qquad (8.5`)$$

Пример. Найдем закон распределения X при условии Y = -0,8 и закон распределения Y при условии X = 3 для случайной величины, рассмотренной в примере 1.

$$p(x_1 / y_1) = \frac{0{,}1}{0{,}5} = \frac{1}{5} = 0{,}2; \quad p(x_2 / y_1) = \frac{0{,}3}{0{,}5} = \frac{3}{5} = 0{,}6; \quad p(x_3 / y_1) = \frac{0{,}1}{0{,}5} = \frac{1}{5} = 0{,}2.$$

$$p(y_1 / x_2) = \frac{0{,}3}{0{,}55} = \frac{6}{11}; \quad p(y_2 / x_2) = \frac{0{,}25}{0{,}55} = \frac{5}{11}.$$

Условные законы распределения составляющих дискретной двумерной случайной величины.

Определение 8.3. **Условной плотностью φ(*x*/*y*)** распределения составляющих X при данном значении $Y = y$ называется

$$\varphi(x / y) = \frac{f(x, y)}{f_2(y)} = \frac{f(x, y)}{\int\limits_{-\infty}^{\infty} f(x, y)dx}. \qquad (8.6)$$

Аналогично определяется условная плотность вероятности Y при $X = x$:

$$\psi(y / x) = \frac{f(x, y)}{f_1(x)} = \frac{f(x, y)}{\int\limits_{-\infty}^{\infty} f(x, y)dy} \qquad (8.6`)$$

Равномерное распределение на плоскости.

Система двух случайных величин называется **равномерно распределенной на плоскости**, если ее плотность вероятности $f(x, y)$ = const внутри некоторой области и равна 0 вне ее.

Пусть данная область – прямоугольник вида $a \le x \le b, \ c \le y \le d$.

Тогда из свойств $f(x, y)$ следует, что

$$f(x, y) = \begin{cases} \dfrac{1}{S_{пр}} = \dfrac{1}{(b-a)(d-c)} & \textit{внутри прямоугольника,} \\ 0 & \textit{вне его.} \end{cases}$$

Найдем двумерную функцию распределения:

$$F(x,y)=\frac{1}{(b-a)(d-c)}\int_c^y\int_a^x dxdy=\frac{(x-a)(y-c)}{(b-a)(d-c)}$$

при $a < x < b$, $c < y < d$, $F(x, y) = 0$, при $x \le a$ или $y \le c$, $F(x, y) = 1$ при $x \ge b$, $y \ge d$.

Функции распределения составляющих, вычисленные по формулам, приведенным в свойстве 4 функции распределения, имеют вид: $F_1(x)=\frac{x-a}{b-a}$, $F_2(y)=\frac{y-c}{d-c}$.

Лекция 9.

Некоторые числовые характеристики одномерных случайных величин: начальные и центральные моменты, мода, медиана, квантиль, коэффициенты асимметрии и эксцесса. Числовые характеристики двумерных случайных величин: начальные и центральные моменты. Корреляционный момент и коэффициент корреляции. Коррелированность и зависимость случайных величин.

Определение 9.1. **Начальным моментом порядка *k*** случайной величины X называется математическое ожидание величины X^k:

$$\nu_k = M(X^k). \tag{9.1}$$

В частности, $\nu_1 = M(X)$, $\nu_2 = M(X^2)$.

Следовательно, дисперсия $D(X) = \nu_2 - \nu_1^2$.

Определение 9.2. **Центральным моментом порядка *k*** случайной величины X называется математическое ожидание величины $(X - M(X))^k$:

$$\mu_k = M((X - M(X))^k). \tag{9.2}$$

В частности, $\mu_1 = M(X - M(X)) = 0$, $\mu_2 = M((X - M(X))^2) = D(X)$.

Можно получить соотношения, связывающие начальные и центральные моменты:

$$\mu_2=\nu_2-\nu_1^2,\quad \mu_3=\nu_3-3\nu_2\nu_1+2\nu_1^2,\quad \mu_4=\nu_4-4\nu_3\nu_1+6\nu_2\nu_1^2-3\nu_1^4.$$

Мода и медиана.

Такая характеристика случайной величины, как математическое ожидание, называется иногда *характеристикой положения*, так как она дает представление о положении случайной величины на числовой оси. Другими характеристиками положения являются мода и медиана.

Определение 9.3. **Модой *М*** дискретной случайной величины называется ее наиболее вероятное значение, **модой *М*** непрерывной случайной величины – значение, в котором плотность вероятности максимальна.

Пример 1.

Если ряд распределения дискретной случайной величины X имеет вид:

X	1	2	3	4
p	0,1	0,7	0,15	0,05

то $M = 2$.

Пример 2.

Для непрерывной случайной величины, заданной плотностью распределения $f(x)=\dfrac{1}{\pi(1+x^2)}$, модой является абсцисса точки максимума: $M = 0$.

Замечание 1. Если кривая распределения имеет больше одного максимума, распределение называется **полимодальным**, если эта кривая не имеет максимума, но имеет минимум – **антимодальным**.

Замечание 2. В общем случае мода и математическое ожидание не совпадают. Но, если распределение является симметричным и модальным (то есть кривая распределения симметрична относительно прямой $x = M$) и имеет математическое ожидание, оно совпадает с модой.

Определение 9.4. **Медианой *Me*** непрерывной случайной величины называют такое ее значение, для которого

$$p(X < Me) = p(X > Me). \tag{9.3}$$

Графически прямая $x = Me$ делит площадь фигуры, ограниченной кривой распределения, на две равные части.

Замечание. Для симметричного модального распределения медиана совпадает с математическим ожиданием и модой.

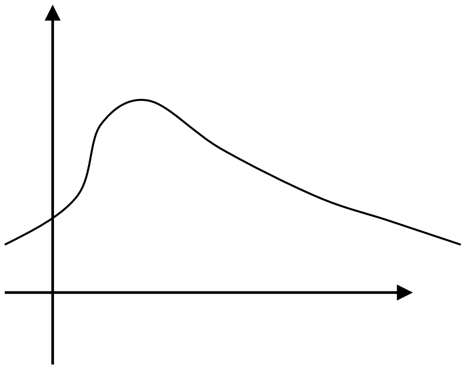
Определение 9.5. Для случайной величины X с функцией распределения $F(X)$ **квантилью порядка *p*** $(0 < p < 1)$ называется число K_p такое, что $F(K_p) \le p$, $F(K_p + 0) \ge p$. В частности, если $F(X)$ строго монотонна, K_p: $F(K_p) = p$.

Асимметрия и эксцесс.

Если распределение не является симметричным, можно оценить асимметрию кривой распределения с помощью центрального момента 3-го порядка. Действительно, для симметричного распределения все нечетные центральные моменты равны 0 (как интегралы от нечетных функций в симметричных пределах), поэтому выбран нечетный момент наименьшего порядка, не тождественно равный 0. Чтобы получить безразмерную характеристику, его делят на σ^3 (так как μ_3 имеет размерность куба случайной величины).

Определение 9.6. **Коэффициентом асимметрии** случайной величины называется

$$S_k = \frac{\mu_3}{\sigma^3}\,. \tag{9.4}$$



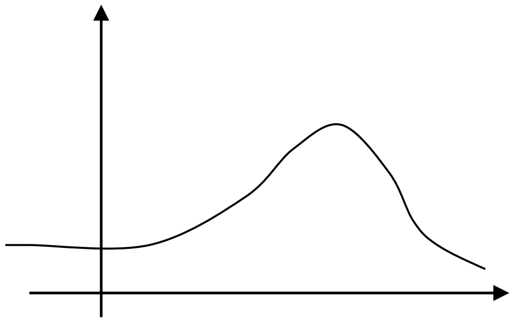


Рис.1. Рис.2.

В частности, для кривой, изображенной на рис.1, $S_k > 0$, а на рис.2 $S_k < 0$. Для оценки поведения кривой распределения вблизи точки максимума (для определения того, насколько «крутой» будет его вершина) применяется центральный момент 4-го порядка.

Определение 9.7. **Эксцессом** случайной величины называется величина

$$Ex = \frac{\mu_4}{\sigma^4} - 3. \tag{9.5}$$

Замечание. Можно показать, что для нормального распределения $\frac{\mu_4}{\sigma^4} = 3$, и, соответственно, $Ex = 0$. Для кривых с более острой вершиной $Ex > 0$, в случае более плоской вершины $Ex < 0$.

Числовые характеристики двумерных случайных величин.

Такие характеристики, как начальные и центральные моменты, можно ввести и для системы двух случайных величин.

Определение 9.8. **Начальным моментом порядка *k*, *s*** двумерной случайной величины (X, Y) называется математическое ожидание произведения X^k на Y^s:

$$\alpha_{k,s} = M(X^k Y^s). \tag{9.6}$$

Для дискретных случайных величин

$$\alpha_{k,s} = \sum_i \sum_j x_i^k y_j^s p_{ij},$$

для непрерывных случайных величин

$$\alpha_{k,s} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^k y^s f(x,y)\,dxdy.$$

Определение 9.9. **Центральным моментом порядка *k*, *s*** двумерной случайной величины (X, Y) называется математическое ожидание произведения $(X - M(X))^k$ на $(Y - M(Y))^s$:

$$\mu_{k,s} = M((X - M(X))^k (Y - M(Y))^s). \tag{9.7}$$

Для дискретных случайных величин

$$\mu_{k,s} = \sum_i \sum_j (x_i - M(X))^k (y_j - M(Y))^s p_{ij},$$

для непрерывных случайных величин

$$\mu_{k,s} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - M(X))^k (y - M(Y))^s f(x, y)dxdy.$$

При этом $M(X) = \alpha_{1,0}$, $M(Y) = \alpha_{0,1}$, $D(X) = \mu_{2,0}$, $D(Y) = \mu_{0,2}$.

Корреляционный момент и коэффициент корреляции.

Определение 9.10. **Корреляционным моментом** системы двух случайных величин называется второй смешанный центральный момент:

$$K_{xy} = \mu_{1,1} = M((X - M(X))(Y - M(Y))). \qquad (9.8)$$

Для дискретных случайных величин

$$K_{xy} = \sum_i \sum_j (x_i - M(X))(y_j - M(Y))p_{ij},$$

для непрерывных случайных величин

$$K_{xy} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - M(X))(y - M(Y))f(x, y)dxdy.$$

Безразмерной характеристикой коррелированности двух случайных величин является **коэффициент корреляции**

$$r_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y} \qquad (9.9)$$

Корреляционный момент описывает связь между составляющими двумерной случайной величины. Действительно, убедимся, что для независимых X и Y $K_{xy} = 0$. В этом случае $f(x,y) = =f_1(x)f_2(y)$, тогда

$$K_{xy} = \int_{-\infty}^{\infty}(x - M(X))f_1(x)dx\int_{-\infty}^{\infty}(y - M(Y))f_2(y)dy = \mu_1(x)\mu_2(y) = 0.$$

Итак, две независимые случайные величины являются и некоррелированными. Однако понятия коррелированности и зависимости не эквивалентны, а именно, величины могут быть зависимыми, но при этом некоррелированными. Дело в том, что коэффициент корреляции характеризует не всякую зависимость, а только линейную. В частности, если $Y = aX + b$, то $r_{xy} = \pm 1$. Найдем возможные значения коэффициента корреляции.

Теорема 9.1. $|r_{xy}| \le 1.$

Доказательство. Докажем сначала, что $|K_{xy}| \le \sigma_x \sigma_y.$

Действительно, если рассмотреть случайную величину

$$Z_1 = \sigma_y X - \sigma_x Y$$

и найти ее дисперсию, то получим:

$$D(Z_1) = 2\sigma_x^2\sigma_y^2 - 2\sigma_x\sigma_y K_{xy}.$$

Так как дисперсия всегда неотрицательна, то

$$2\sigma_x^2\sigma_y^2 - 2\sigma_x\sigma_y K_{xy} \ge 0,$$ откуда $|K_{xy}| \le \sigma_x \sigma_y.$

Отсюда $\left|\frac{K_{xy}}{\sigma_x \sigma_y}\right| = |r_{xy}| \le 0,$ что и требовалось доказать.

Лекция 10.

Функции от случайных величин. Функция одного случайного аргумента, ее распределение и математическое ожидание. Функция двух случайных аргументов. Распределение суммы независимых слагаемых. Устойчивость нормального распределения.

В предыдущих лекциях рассматривались некоторые законы распределения случайных величин. При решении задач часто удобно бывает представить исследуемую случайную величину как функцию других случайных величин с известными законами распределения, что помогает установить и закон распределения заданной случайной величины.

Определение 10.1. Если каждому возможному значению случайной величины X соответствует одно возможное значение случайной величины Y, то Y называют **функцией случайного аргумента X:** $Y = \varphi(X)$. Выясним, как найти закон распределения функции по известному закону распределения аргумента.

1) Пусть аргумент X – дискретная случайная величина, причем различным значениям X соответствуют различные значения Y. Тогда вероятности соответствующих значений X и Y равны.

Пример 1. Ряд распределения для X имеет вид:

X	5	6	7	8
p	0,1	0,2	0,3	0,4

Найдем закон распределения функции $Y = 2X^2 - 3$:

Y	47	69	95	125
p	0,1	0,2	0,3	0,4

(при вычислении значений Y в формулу, задающую функцию, подставляются возможные значения X).

2) Если разным значениям X могут соответствовать одинаковые значения Y, то вероятности значений аргумента, при которых функция принимает одно и то же значение, складываются.

Пример 2. Ряд распределения для X имеет вид:

X	0	1	2	3
p	0,1	0,2	0,3	0,4

Найдем закон распределения функции $Y = X^2 - 2X$:

Y	-1	0	3
p	0,2	0,4	0,4

(так как $Y = 0$ при $X = 0$ и $X = 2$, то $p\ (Y = 0) = p\ (X = 0) + p\ (X = 2) = 0{,}1 + 0{,}3 = 0{,}4$).

3) Если X – непрерывная случайная величина, $Y = \varphi(X)$, $\varphi(x)$ – монотонная и дифференцируемая функция, а $\psi(y)$ – функция, обратная к $\varphi(x)$, то плотность распределения $g(y)$ случайно функции Y равна:

$$g(y) = f(\psi(y))\,|\,\psi'(y)\,|. \qquad (10.1)$$

Пример.

$$f(x)=\frac{1}{\pi(1+x^2)},\ \ Y=x^3$$
.

Тогда

$$\psi(y)=\sqrt[3]{y},\ \ g(y)=\frac{1}{\pi(1+y^{\frac{2}{3}})}\cdot\left(\frac{1}{3}y^{-\frac{2}{3}}\right)=\frac{1}{3\pi y^{\frac{2}{3}}(1+y^{\frac{2}{3}})}$$

Математическое ожидание функции одного случайного аргумента.

Пусть $Y = \varphi(X)$ – функция случайного аргумента X, и требуется найти ее математическое ожидание, зная закон распределения X.

1) Если X – дискретная случайная величина, то

$$M(Y)=M(\varphi(x))=\sum_{i=1}^{n}\varphi(x_i)p_i. \tag{10.2}$$

Пример 3. Найдем M(Y) для примера 1:

$M(Y) = 47·0,1 + 69·0,2 + 95·0,3 + 125·0,4 = 97$.

2) Если X – непрерывная случайная величина, то $M(Y)$ можно искать по-разному. Если известна плотность распределения $g(y)$, то

$$M(Y)=\int_{-\infty}^{\infty}yg(y)dy. \tag{10.3}$$

Если же $g(y)$ найти сложно, то можно использовать известную плотность распределения $f(x)$:

$$M(Y)=\int_{-\infty}^{\infty}\varphi(x)f(x)dx. \tag{10.4}$$

В частности, если все значения X принадлежат промежутку (a, b), то

$$M(Y)=\int_{a}^{b}\varphi(x)f(x)dx. \tag{10.4`}$$

Функция двух случайных величин. Распределение суммы независимых слагаемых.

Определение 10.2. Если каждой паре возможных значений случайных величин X и Y соответствует одно возможное значение случайной величины Z, то Z называют **функцией двух случайных аргументов** X и Y: $Z = \varphi(X, Y)$.

Рассмотрим в качестве такой функции сумму $X + Y$. В некоторых случаях можно найти ее закон распределения, зная законы распределения слагаемых.

1) Если X и Y – дискретные *независимые* случайные величины, то для определения закона распределения $Z = X + Y$ нужно найти все возможные значения Z и соответствующие им вероятности.

Пример 4.

Рассмотрим дискретные случайные величины X и Y, законы распределения которых имеют вид:

X	-2	1	3
p	0,3	0,4	0,3

Y	0	1	2
p	0,2	0,5	0,3

Найдем возможные значения Z:

-2 + 0 = -2 (p = 0,3·0,2 = 0,06),

-2 + 1 = -1 (p = 0,3·0,5 = 0,15),

-2 + 2 = 0 (p = 0,3·0,3 = 0,09),

1 + 0 = 1 (p = 0,4·0,2 = 0,08),

1 + 1 = 2 (p = 0,4·0,5 = 0,2),

1 + 2 = 3 (p = 0,4·0,3 = 0,12),

3 + 0 = 3 (p = 0,3·0,2 = 0,06),

3 + 1 = 4 (p = 0,3·0,5 = 0,15),

3 + 2 = 5 (p = 0,3·0,3 = 0,09).

Сложив вероятности повторившегося дважды значения Z = 3, составим ряд распределения для Z:

Z	-2	-1	0	1	2	3	4	5
p	0,06	0,15	0,09	0,08	0,2	0,18	0,15	0,09

3) Если X и Y – непрерывные *независимые* случайные величины, то, если плотность вероятности хотя бы одного из аргументов задана на (-∞, ∞) одной формулой, то плотность суммы $g(z)$ можно найти по формулам

$$g(z) = \int_{-\infty}^{\infty} f_1(x) f_2(z-x)dx = \int_{-\infty}^{\infty} f_1(z-y) f_2(y)dy, \qquad (10.5)$$

где $f_1(x)$, $f_2(y)$ – плотности распределения слагаемых. Если возможные значения аргументов неотрицательны, то

$$g(z) = \int_{0}^{z} f_1(x) f_2(z-x)dx = \int_{0}^{z} f_1(z-y) f_2(y)dy. \qquad (10.6)$$

Замечание. Плотность распределения суммы двух независимых случайных величин называют **композицией**.

Устойчивость нормального распределения.

Определение 10.3. Закон распределения вероятностей называется **устойчивым**, если композиция таких законов есть тот же закон (возможно, отличающийся другими значениями параметров).

В частности, свойством устойчивости обладает нормальный закон распределения: композиция нормальных законов тоже имеет нормальное распределение, причем ее математическое ожидание и дисперсия равны суммам соответствующих характеристик слагаемых.

Лекция 11.

Нормальный закон распределения на плоскости. Линейная регрессия. Линейная корреляция.

Определение 11.1. **Нормальным законом распределения на плоскости** называют распределение вероятностей двумерной случайной величины (X, Y), если

$$f(x,y)=\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r_{xy}^2}}e^{-\frac{1}{2(1-r_{xy}^2)}\left(\frac{(x-a_1)^2}{\sigma_x^2}+\frac{(y-a_2)^2}{\sigma_y^2}-2r_{xy}\frac{x-a_1}{\sigma_x}\frac{y-a_2}{\sigma_y}\right)} \qquad (11.1)$$

Таким образом, нормальный закон на плоскости определяется 5 параметрами: a_1, a_2, σ_x, σ_y, r_{xy}, где a_1, a_2 – математические ожидания, σ_x, σ_y – средние квадратические отклонения, r_{xy} – коэффициент корреляции X и Y. Предположим, что $r_{xy}=0$, то есть X и Y некоррелированы. Тогда из (11.1) получим:

$$f(x,y)=\frac{1}{2\pi\sigma_x\sigma_y}\cdot e^{-0{,}5\left(\frac{(x-a_1)^2}{\sigma_x^2}+\frac{(y-a_2)^2}{\sigma_y^2}\right)}=\frac{1}{\sigma_x\sqrt{2\pi}}e^{-\frac{(x-a_1)^2}{2\sigma_x^2}}\cdot\frac{1}{\sigma_y\sqrt{2\pi}}e^{-\frac{(y-a_2)^2}{2\sigma_y^2}}=f_1(x)f_2(y).$$

Следовательно, из некоррелированности составляющих нормально распределенной двумерной случайной величины следует их независимость, то есть для них понятия независимости и некоррелированности равносильны.

Линейная регрессия.

Пусть составляющие X и Y двумерной случайной величины (X, Y) зависимы. Будем считать, что одну из них можно приближенно представить как линейную функцию другой, например

$$Y\approx g(X)=\alpha+\beta X, \qquad (11.2)$$

и определим параметры α и β с помощью метода наименьших квадратов.

Определение 11.2. Функция $g(X)=\alpha+\beta X$ называется **наилучшим приближением** Y в смысле метода наименьших квадратов, если математическое ожидание $M(Y-g(X))^2$ принимает наименьшее возможное значение; функцию $g(X)$ называют **среднеквадратической регрессией** Y на X.

Теорема 11.1. Линейная средняя квадратическая регрессия Y на X имеет вид:

$$g(X)=m_y+r\frac{\sigma_y}{\sigma_x}(X-m_x), \qquad (11.3)$$

где $m_x=M(X), m_y=M(Y), \sigma_x=\sqrt{D(X)}, \sigma_y=\sqrt{D(Y)}, r=\frac{K_{xy}}{\sigma_x\sigma_y}$ - коэффициент корреляции X и Y.

Доказательство. Рассмотрим функцию

$$F(\alpha,\beta)=M(Y-\alpha-\beta X)^2 \qquad (11.4)$$

и преобразуем ее, учитывая соотношения

$M(X-m_x)=M(Y-m_y)=0,\qquad M((X-m_x)(Y-m_y))=K_{xy}=r\sigma_x\sigma_y$:

$$F(\alpha,\beta)=\sigma_y^2+\beta^2\sigma_x^2-2r\sigma_x\sigma_y\beta+(m_y-\alpha-\beta m_x)^2.$$

Найдем стационарные точки полученной функции, решив систему

$$\begin{cases}\dfrac{\partial F}{\partial\alpha}=-2(m_y-\alpha-\phi m_x)=0,\\ \dfrac{\partial F}{\partial\beta}=2\beta\sigma_x^2-2r\sigma_x\sigma_y=0.\end{cases}$$

Решением системы будет

$$\beta=r\frac{\sigma_y}{\sigma_x},\alpha=m_y-r\frac{\sigma_y}{\sigma_x}m_x.$$

Можно проверить, что при этих значениях функция $F(\alpha,\ \beta)$ имеет минимум, что доказывает утверждение теоремы.

Определение 11.3. Коэффициент $\beta=r\dfrac{\sigma_y}{\sigma_x}$ называется **коэффициентом регрессии *Y* на *X***, а прямая

$$y-m_y=r\frac{\sigma_y}{\sigma_x}(x-m_x)\quad -\qquad(11.5)$$

- прямой среднеквадратической регрессии *Y* на *X*.

Подставив координаты стационарной точки в равенство (11.4), можно найти минимальное значение функции $F(\alpha,\beta)$, равное $\sigma_y^2(1-r^2)$.

Эта величина называется **остаточной дисперсией** *Y* относительно *X* и характеризует величину ошибки, допускаемой при замене *Y* на $g(X)=\alpha+\beta X$.

При $r=\pm1$ остаточная дисперсия равна 0, то есть равенство (11.2) является не приближенным, а точным.

Следовательно, при $r=\pm1$ *Y* и *X* связаны *линейной* функциональной зависимостью.

Аналогично можно получить прямую среднеквадратической регрессии *X* на *Y*:

$$x-m_x=r\frac{\sigma_x}{\sigma_y}(y-m_y)\qquad(11.6)$$

и остаточную дисперсию *X* относительно *Y*.

При $r=\pm1$ обе прямые регрессии совпадают. Решив систему из уравнений (11.5) и (11.6), можно найти точку пересечения прямых регрессии – точку с координатами $(m_x,\ m_y)$, называемую **центром совместного распределения величин *X* и *Y*.**

Линейная корреляция.

Для двумерной случайной величины (*X, Y*) можно ввести так называемое **условное математическое ожидание *Y*** при $X=x$. Для дискретной случайной величины оно определяется как

$$M(Y \mid X = x) = \sum_{j=1}^{m} y_j p(y_j / x), \tag{11.7}$$

для непрерывной случайной величины –

$$M(Y \mid X = x) = \int_{-\infty}^{\infty} y\psi(y/x)dy \, . \tag{11.8}$$

Определение 11.4. **Функцией регрессии *Y* на *X*** называется условное математическое ожидание

$$M(Y / x) = f(x).$$

Аналогично определяется условное математическое ожидание *X* и функция регрессии *X* на *Y*.

Определение 11.5. Если обе функции регрессии *X* на *Y* и *Y* на *X* линейны, то говорят, что *X* и *Y* связаны **линейной корреляционной зависимостью**. При этом графики линейных функций регрессии являются прямыми линиями, причем можно доказать, что эти линии совпадают с прямыми среднеквадратической регрессии.

Теорема 11.2. Если двумерная случайная величина (*X, Y*) распределена нормально, то *X* и *Y* связаны линейной корреляционной зависимостью.

Доказательство. Найдем условный закон распределения *Y* при *X* = *x*

$$\left(\psi(y/x) = \frac{f(x,y)}{f_1(x)} \right),$$

используя формулу двумерной плотности вероятности нормального распределения (11.1) и формулу плотности вероятности *X*:

$$f_1(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e - \frac{(x-a_1)^2}{2\sigma_x^2} \, . \tag{11.9}$$

Сделаем замену $u = \dfrac{x-a_1}{\sigma_x}, \ v = \dfrac{y-a_2}{\sigma_y}$. Тогда

$$\psi(y/x) = \frac{1}{\sqrt{2\pi}\sigma_y} e^{-\frac{(v-ru)^2}{2(1-r^2)}} = \frac{1}{\sigma_y \sqrt{1-r^2}\sqrt{2\pi}} e^{-\frac{\left(y-\left(a_2+r\frac{\sigma_y}{\sigma_x}(x-a_1)\right)\right)^2}{2\sigma_y^2(1-r^2)}} .$$

Полученное распределение является нормальным, а его математическое ожидание $M(Y/x) = a_2 + r\dfrac{\sigma_y}{\sigma_x}(x-a_1)$ есть функция регрессии *Y* на *X* (см. определение 11.4)).

Аналогично можно получить функцию регрессии *X* на *Y*:

$$M(X/y) = a_1 + r\frac{\sigma_x}{\sigma_y}(y-a_2) \, .$$

Обе функции регрессии линейны, поэтому корреляция между X и Y линейна, что и требовалось доказать. При этом уравнения прямых регрессии имеют вид

$$y - a_2 = r\frac{\sigma_y}{\sigma_x}(x - a_1), \qquad x - a_1 = r\frac{\sigma_x}{\sigma_y}(y - a_2),$$

то есть совпадают с уравнениями прямых среднеквадратической регрессии (см. формулы (11.5), (11.6)).

Лекция 12.

Распределения «хи-квадрат», Стьюдента и Фишера. Связь этих распределений с нормальным распределением.

Рассмотрим некоторые распределения, связанные с нормальным и широко применяющиеся в математической статистике.

Распределение «хи-квадрат».

Пусть имеется несколько нормированных нормально распределенных случайных величин: $X_1, X_2, \ldots, X_n$ ($a_i = 0$, $\sigma_i = 1$). Тогда сумма их квадратов

$$\chi^2 = \sum_{i=1}^{n} X_i^2 \qquad (12.1)$$

является случайной величиной, распределенной по так называемому **закону «хи-квадрат»** с $k = n$ степенями свободы; если же слагаемые связаны каким-либо соотношением (например, $\sum X_i = n\bar{X}$), то число степеней свободы $k = n - 1$.

Плотность этого распределения

$$f(x) = \begin{cases} 0, & x \le 0; \\ \dfrac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} e^{-\frac{x}{2}} x^{\frac{k}{2}-1}, & x > 0. \end{cases} \qquad (12.2)$$

Здесь $\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$ - гамма-функция; в частности, $\Gamma(n + 1) = n!$.

Следовательно, распределение «хи-квадрат» определяется одним параметром – числом степеней свободы k.

Замечание 1. С увеличением числа степеней свободы распределение «хи-квадрат» постепенно приближается к нормальному.

Замечание 2. С помощью распределения «хи-квадрат» определяются многие другие распределения, встречающиеся на практике, например, распределение случайной величины $\sqrt{\chi^2}$ - длины случайного вектора (X_1, $X_2, \ldots$, X_n), координаты которого независимы и распределены по нормальному закону.

Распределение Стьюдента.

Рассмотрим две независимые случайные величины: Z, имеющую нормальное распределение и нормированную (то есть $M(Z) = 0$, $\sigma(Z) = 1$), и V, распределенную по закону «хи-квадрат» с k степенями свободы.

Тогда величина

$$T = \frac{Z}{\sqrt{\frac{V}{k}}} \tag{12.3}$$

имеет распределение, называемое ***t* – распределением или распределением Стьюдента** с k степенями свободы.

С возрастанием числа степеней свободы распределение Стьюдента быстро приближается к нормальному.

Распределение *F* Фишера – Снедекора.

Рассмотрим две независимые случайные величины U и V, распределенные по закону «хи-квадрат» со степенями свободы k_1 и k_2 и образуем из них новую величину

$$F = \frac{U / k_1}{V / k_2}. \tag{12.4}$$

Ее распределение называют **распределением *F* Фишера – Снедекора** со степенями свободы k_1 и k_2. Плотность его распределения имеет вид

$$f(x) = \begin{cases} 0, \quad x \le 0; \\ C_0 \dfrac{x^{\frac{k_1-2}{2}}}{(k_2 + k_1 x)^{\frac{k_1+k_2}{2}}}, \quad x > 0, \end{cases} \tag{12.5}$$

где $C_0 = \dfrac{\Gamma\left(\dfrac{k_1+k_2}{2}\right) k_1^{\frac{k_1}{2}} k_2^{\frac{k_2}{2}}}{\Gamma\left(\dfrac{k_1}{2}\right)\Gamma\left(\dfrac{k_2}{2}\right)}$.

Таким образом, распределение Фишера определяется двумя параметрами – числами степеней свободы.

Лекция 13.

Закон больших чисел. Неравенство Чебышева. Теоремы Чебышева и Бернулли.

Изучение статистических закономерностей позволило установить, что при некоторых условиях суммарное поведение большого количества случайных величин почти утрачивает случайный характер и становится закономерным (иначе говоря, случайные отклонения от некоторого среднего поведения взаимно погашаются). В частности, если влияние на сумму отдельных слагаемых является равномерно малым, закон распределения суммы приближается к нормальному. Математическая формулировка этого утверждения дается в группе теорем, называемой **законом больших чисел**.

Неравенство Чебышева.

Неравенство Чебышева, используемое для доказательства дальнейших теорем, справедливо как для непрерывных, так и для дискретных случайных величин. Докажем его для дискретных случайных величин.

Теорема 13.1(неравенство Чебышева).

$p(\,|X - M(X)| < \varepsilon\,) \geq D(X) / \varepsilon^2.$ (13.1)

Доказательство. Пусть X задается рядом распределения

X	x_1	x_2	…	x_n
p	p_1	p_2	…	p_n

Так как события $|X - M(X)| < \varepsilon$ и $|X - M(X)| \geq \varepsilon$ противоположны, то

$p\,(\,|X - M(X)| < \varepsilon\,) + p\,(\,|X - M(X)| \geq \varepsilon\,) = 1$, следовательно,

$p\,(\,|X - M(X)| < \varepsilon\,) = 1 - p\,(\,|X - M(X)| \geq \varepsilon\,)$.

Найдем $p\,(\,|X - M(X)| \geq \varepsilon\,)$.

$D(X) = (x_1 - M(X))^2p_1 + (x_2 - M(X))^2p_2 + \ldots + (x_n - M(X))^2p_n$.

Исключим из этой суммы те слагаемые, для которых $|X - M(X)| < \varepsilon$. При этом сумма может только уменьшиться, так как все входящие в нее слагаемые неотрицательны. Для определенности будем считать, что отброшены первые k слагаемых. Тогда

$D(X) \geq (x_{k+1} - M(X))^2p_{k+1} + (x_{k+2} - M(X))^2p_{k+2} + \ldots + (x_n - M(X))^2p_n \geq \varepsilon^2\,(p_{k+1} + p_{k+2} + \ldots + p_n)$.

Отметим, что $p_{k+1} + p_{k+2} + \ldots + p_n$ есть вероятность того, что $|X - M(X)| \geq \varepsilon$, так как это сумма вероятностей всех возможных значений X, для которых это неравенство справедливо. Следовательно, $D(X) \geq \varepsilon^2\, p(|X - M(X)| \geq \varepsilon)$, или $p\,(|X - M(X)| \geq \varepsilon) \leq D(X) / \varepsilon^2$. Тогда вероятность противоположного события $p(\,|X - M(X)| < \varepsilon\,) \geq D(X) / \varepsilon^2$, что и требовалось доказать.

Теоремы Чебышева и Бернулли.

Теорема 13.2 (теорема Чебышева). Если X_1, X_2,…, X_n – попарно независимые случайные величины, дисперсии которых равномерно ограничены ($D(X_i) \leq C$), то для сколь угодно малого числа ε вероятность неравенства

$$\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \frac{M(X_1) + M(X_2) + \ldots + M(X_n)}{n}\right| < \varepsilon$$

будет сколь угодно близка к 1, если число случайных величин достаточно велико.

Замечание. Иначе говоря, при выполнении этих условий

$$\lim_{n\to\infty} p\left(\left|\frac{X_1+X_2+\ldots+X_n}{n}-\frac{M(X_1)+M(X_2)+\ldots+M(X_n)}{n}\right|<\varepsilon\right)=1.$$

Доказательство. Рассмотрим новую случайную величину

$$\overline{X}=\frac{X_1+X_2+\ldots+X_n}{n}$$

и найдем ее математическое ожидание.
Используя свойства математического ожидания, получим, что

$$M\left(\frac{X_1+X_2+\ldots+X_n}{n}\right)=\frac{M(X_1)+M(X_2)+\ldots+M(X_n)}{n}.$$

Применим к $\overline{X}$ неравенство Чебышева:

$$p\left(\left|\frac{X_1+X_2+\ldots+X_n}{n}-\frac{M(X_1)+M(X_2)+\ldots+M(X_n)}{n}\right|<\varepsilon\right)\ge 1-\frac{D\left(\frac{X_1+X_2+\ldots+X_n}{n}\right)}{\varepsilon^2}.$$

Так как рассматриваемые случайные величины независимы, то, учитывая условие теоремы, имеем:

$$D\left(\frac{X_1+X_2+\ldots+X_n}{n}\right)=\frac{D(X_1)+D(X_2)+\ldots+D(X_n)}{n^2}\le\frac{Cn}{n^2}=\frac{C}{n}.$$

Используя этот результат, представим предыдущее неравенство в виде:

$$p\left(\left|\frac{X_1+X_2+\ldots+X_n}{n}-\frac{M(X_1)+M(X_2)+\ldots+M(X_n)}{n}\right|<\varepsilon\right)\ge 1-\frac{C}{n\varepsilon^2}.$$

Перейдем к пределу при $n\to\infty$:

$$\lim_{n\to\infty} p\left(\left|\frac{X_1+X_2+\ldots+X_n}{n}-\frac{M(X_1)+M(X_2)+\ldots+M(X_n)}{n}\right|<\varepsilon\right)\ge 1.$$

Поскольку вероятность не может быть больше 1, можно утверждать, что

$$\lim_{n\to\infty} p\left(\left|\frac{X_1+X_2+\ldots+X_n}{n}-\frac{M(X_1)+M(X_2)+\ldots+M(X_n)}{n}\right|<\varepsilon\right)=1.$$

Теорема доказана.

Следствие.
Если X_1, X_2, …, X_n – попарно независимые случайные величины с равномерно ограниченными дисперсиями, имеющие одинаковое математическое ожидание, равное a, то для любого сколь угодно малого $\varepsilon > 0$ вероятность неравенства $\left|\frac{X_1+X_2+\ldots+X_n}{n}-a\right|<\varepsilon$ будет как угодно близка к 1, если число случайных величин достаточно велико.
Иначе говоря,

$$\lim_{n\to\infty} p\left(\left|\frac{X_1+X_2+\ldots+X_n}{n}-a\right|<\varepsilon\right)=1.$$

Вывод: среднее арифметическое достаточно большого числа случайных величин принимает значения, близкие к сумме их математических ожиданий, то есть утрачивает характер случайной величины. Например, если проводится серия измерений какой-либо физической величины, причем:

а) результат каждого измерения не зависит от результатов остальных, то есть все результаты представляют собой попарно независимые случайные величины;

б) измерения производятся без систематических ошибок (их математические ожидания равны между собой и равны истинному значению *а* измеряемой величины);

в) обеспечена определенная точность измерений, следовательно, дисперсии рассматриваемых случайных величин равномерно ограничены;

то при достаточно большом числе измерений их среднее арифметическое окажется сколь угодно близким к истинному значению измеряемой величины.

Теорема Бернулли.

Теорема 13.3 (теорема Бернулли). Если в каждом из n независимых опытов вероятность p появления события A постоянна, то при достаточно большом числе испытаний вероятность того, что модуль отклонения относительной частоты появлений A в n опытах от p будет сколь угодно малым, как угодно близка к 1:

$$\lim_{n\to\infty} p\left(\left|\frac{m}{n}-p\right|<\varepsilon\right)=1. \tag{13.2}$$

Доказательство. Введем случайные величины X_1, X_2, …, X_n, где X_i – число появлений A в i-м опыте. При этом X_i могут принимать только два значения: 1(с вероятностью p) и 0 (с вероятностью $q = 1 – p$). Кроме того, рассматриваемые случайные величины попарно независимы и их дисперсии равномерно ограничены (так как $D(X_i) = pq$, $p + q = 1$, откуда $pq \le ¼$). Следовательно, к ним можно применить теорему Чебышева при $M_i = p$:

$$\lim_{n\to\infty} p\left(\left|\frac{X_1+X_2+\ldots+X_n}{n}-p\right|<\varepsilon\right)=1.$$

Но $\dfrac{X_1+X_2+\ldots+X_n}{n}=\dfrac{m}{n}$, так как X_i принимает значение, равное 1, при появлении A в данном опыте, и значение, равное 0, если A не произошло. Таким образом,

$$\lim_{n\to\infty} p\left(\left|\frac{m}{n}-p\right|<\varepsilon\right)=1,$$

что и требовалось доказать.

Замечание. Из теоремы Бернулли *не следует*, что $\lim\limits_{n\to\infty} \frac{m}{n} = p.$
Речь идет лишь о *вероятности* того, что разность относительной частоты и вероятности по модулю может стать сколь угодно малой. Разница заключается в следующем: при обычной сходимости, рассматриваемой в математическом анализе, для всех n, начиная с некоторого значения, неравенство $\left|\frac{m}{n} - p\right| < \varepsilon$ выполняется всегда; в нашем случае могут найтись такие значения n, при которых это неравенство неверно. Этот вид сходимости называют **сходимостью по вероятности**.

Лекция 14.

Центральная предельная теорема Ляпунова. Предельная теорема Муавра-Лапласа.

Закон больших чисел не исследует вид предельного закона распределения суммы случайных величин. Этот вопрос рассмотрен в группе теорем, называемых **центральной предельной теоремой.** Они утверждают, что закон распределения суммы случайных величин, каждая из которых может иметь различные распределения, приближается к нормальному при достаточно большом числе слагаемых. Этим объясняется важность нормального закона для практических приложений.

Характеристические функции.

Для доказательства центральной предельной теоремы используется метод характеристических функций.

Определение 14.1. **Характеристической функцией** случайной величины X называется функция

$$g(t) = M(e^{itX}) \tag{14.1}$$

Таким образом, $g(t)$ представляет собой математическое ожидание некоторой комплексной случайной величины $U = e^{itX}$, связанной с величиной X. В частности, если X – дискретная случайная величина, заданная рядом распределения, то

$$g(t) = \sum_{k=1}^{n} e^{itx_k} p_k \quad . \tag{14.2}$$

Для непрерывной случайной величины с плотностью распределения $f(x)$

$$g(t) = \int_{-\infty}^{+\infty} e^{itx} f(x)dx. \tag{14.3}$$

Пример 1. Пусть X – число выпадений 6 очков при одном броске игральной кости. Тогда по формуле (14.2)

$$g(t) = e^{it\cdot 0}\cdot\frac{5}{6} + e^{it\cdot 1}\cdot\frac{1}{6} = \frac{5+e^{it}}{6}.$$

Пример 2. Найдем характеристическую функцию для нормированной непрерывной случайной величины, распределенной по нормальному закону $\left(f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \right)$.
По формуле (14.3)

$$g(t) = \int_{-\infty}^{+\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{itx - \frac{x^2}{2}} dx = e^{-\frac{t^2}{2}}$$

(использовалась формула $\int_{-\infty}^{+\infty} e^{-Ax^2 \pm 2Bx - C} dx = \sqrt{\frac{\pi}{A}} e^{-\frac{AC - B^2}{A}}$ и то, что $i^2 = -1$).

Свойства характеристических функций.

1. Функцию $f(x)$ можно найти по известной функции $g(t)$ по формуле

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} g(t) dt. \tag{14.4}$$

(преобразование (14.3) называется *преобразованием Фурье*, а преобразование (14.4) – *обратным преобразованием Фурье*).
2. Если случайные величины X и Y связаны соотношением $Y = aX$, то их характеристические функции связаны соотношением

$$g_y(t) = g_x(at). \tag{14.5}$$

3. Характеристическая функция суммы независимых случайных величин равна произведению характеристических функций слагаемых: для $Y = \sum_{k=1}^{n} X_k$

$$g_y(t) = g_{x_1}(t) \cdot g_{x_2}(t) \cdot \ldots \cdot g_{x_n}(t) \tag{14.6}$$

Теорема 14.1 (центральная предельная теорема для одинаково распределенных слагаемых). Если $X_1, X_2, \ldots, X_n, \ldots$ - независимые случайные величины с одинаковым законом распределения, математическим ожиданием m и дисперсией σ^2, то при неограниченном увеличении n закон распределения суммы $Y_n = \sum_{k=1}^{n} X_k$ неограниченно приближается к нормальному.
Доказательство.
Докажем теорему для непрерывных случайных величин $X_1, X_2, \ldots, X_n$ (доказательство для дискретных величин аналогично). Согласно условию теоремы, характеристические функции слагаемых одинаковы:

$$g_x(t) = \int_{-\infty}^{+\infty} e^{itx} f(x) dx.$$

Тогда по свойству 3 характеристическая функция суммы Y_n будет

$$g_{y_n}(t) = g_x^n(t).$$

Разложим функцию $g_x(t)$ в ряд Маклорена:

$$g_x(t) = g_x(0) + g_x'(0)t + \left(\frac{g_x''(0)}{2} + \alpha(t)\right)t^2$$, где $\alpha(t) \to 0$ при $t \to 0$.

Найдем

$$g_x(0) = \int_{-\infty}^{+\infty} f(x)dx = 1,\quad g_x'(0) = \int_{-\infty}^{+\infty} ixe^{itx} f(x)dx\Big|_{t=0} = i\int_{-\infty}^{+\infty} xe^{itx} f(x)dx\Big|_{t=0} = i\int_{-\infty}^{+\infty} xf(x)dx = im.$$

Если предположить, что $m = 0$ (перенести начало отсчета в точку m), то $g_x'(0) = 0$.

$$g_x''(0) = -\int_{-\infty}^{+\infty} x^2 e^{itx} f(x)dx\Big|_{t=0} = -\int_{-\infty}^{+\infty} x^2 f(x)dx = -\sigma^2$$

(так как $m = 0$). Подставив полученные результаты в формулу Маклорена, найдем, что

$$g_x(t) = 1 - \left(\frac{\sigma^2}{2} - \alpha(t)\right)t^2.$$

Рассмотрим новую случайную величину $Z_n = \frac{Y_n}{\sigma\sqrt{n}}$, отличающуюся от Y_n тем, что ее дисперсия при любом n равна 0. Так как Y_n и Z_n связаны линейной зависимостью, достаточно доказать, что Z_n распределена по нормальному закону, или, что то же самое, что ее характеристическая функция приближается к характеристической функции нормального закона (см. пример 2). По свойству характеристических функций

$$g_{z_n}(t) = g_{y_n}\left(\frac{t}{\sigma\sqrt{n}}\right) = \left(g_x\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n = \left(1 - \left(\frac{\sigma^2}{2} - \alpha\left(\frac{t}{\sigma\sqrt{n}}\right)\right)\frac{t^2}{n\sigma^2}\right)^n.$$

Прологарифмируем полученное выражение:

$$\ln g_{z_n}(t) = n\ln(1-k),$$ где $$k = \left(\frac{\sigma^2}{2} - \alpha\left(\frac{t}{\sigma\sqrt{n}}\right)\right)\frac{t^2}{n\sigma^2},\quad \lim_{n\to\infty} k = 0.$$

Разложим $\ln(1-k)$ в ряд при $n \to \infty$, ограничившись двумя членами разложения, тогда $\ln(1 - k) \approx - k$. Отсюда

$$\lim_{n\to\infty} \ln g_{z_n}(t) = \lim_{n\to\infty} n\cdot(-k) = \lim_{n\to\infty}\left(-\frac{t^2}{2} + \alpha\left(\frac{t}{\sigma\sqrt{n}}\right)\frac{t^2}{\sigma^2}\right) = -\frac{t^2}{2} + \lim_{n\to\infty}\frac{t^2}{\sigma^2}\alpha\left(\frac{t}{\sigma\sqrt{n}}\right),$$

где последний предел равен 0, так как $\alpha(t) \to 0$ при $t \to 0$. Следовательно,

$\lim_{n\to\infty} \ln g_{z_n}(t) = -\frac{t^2}{2}$, то есть $\lim_{n\to\infty} g_{z_n}(t) = e^{-\frac{t^2}{2}}$ - характеристическая функция нормального распределения.

Итак, при неограниченном увеличении числа слагаемых характеристическая функция величины Z_n неограниченно приближается к

характеристической функции нормального закона; следовательно, закон распределения Z_n (и Y_n) неограниченно приближается к нормальному.
Теорема доказана.
А.М.Ляпунов доказал центральную предельную теорему для условий более общего вида:

Теорема 14.2 (теорема Ляпунова). Если случайная величина X представляет собой сумму очень большого числа взаимно независимых случайных величин, для которых выполнено условие:

$$\lim_{n\to\infty}\frac{\sum_{k=1}^{n} b_k}{\left(\sum_{k=1}^{n} D_k\right)^{\frac{3}{2}}} \qquad (14.7)$$

где b_k – третий абсолютный центральный момент величины $X_к$, а D_k – ее дисперсия, то X имеет распределение, близкое к нормальному (условие Ляпунова означает, что влияние каждого слагаемого на сумму ничтожно мало).

Практически можно использовать центральную предельную теорему при достаточно небольшом количестве слагаемых, так как вероятностные расчеты требуют сравнительно малой точности. Опыт показывает, что для суммы даже десяти и менее слагаемых закон их распределения можно заменить нормальным.
Частным случаем центральной предельной теоремы для дискретных случайных величин является теорема Муавра-Лапласа.
Теорема 14.3 (теорема Муавра-Лапласа). Если производится n независимых опытов, в каждом из которых событие A появляется с вероятностью p, то справедливо соотношение:

$$p\left(\alpha < \frac{Y - np}{\sqrt{npq}} < \beta\right) = \Phi(\beta) - \Phi(\alpha), \qquad (14.8)$$

где Y – число появлений события A в n опытах, $q = 1 – p$.
Доказательство.
Будем считать, что $Y = \sum_{i=1}^{n} X_i$, где X_i – число появлений события A в i-м опыте. Тогда случайную величину $Z = \frac{Y - m_y}{\sigma_y}$ (см. теорему 14.1) можно считать распределенной по нормальному закону и нормированной, следовательно, вероятность ее попадания в интервал (α, β) можно найти по формуле

$$p(\alpha < Z < \beta) = \Phi(\beta) - \Phi(\alpha).$$

Поскольку Y имеет биномиальное распределение,

$$m_y = np,\ \ D_y = npq,\ \ \sigma_y = \sqrt{npq}.$$

Тогда $Z = \dfrac{Y - np}{\sqrt{npq}}$. Подставляя это выражение в предыдущую формулу, получим равенство (14.8).

Следствие.

В условиях теоремы Муавра-Лапласа вероятность $p_n(k)$ того, что событие A появится в n опытах ровно k раз, при большом количестве опытов можно найти по формуле:

$$p_n(k) \approx \frac{1}{\sqrt{npq}} \cdot \varphi(x), \tag{14.9}$$

где $x = \dfrac{k - np}{\sqrt{npq}}$, а $\varphi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ (значения этой функции приводятся в специальных таблицах).

Пример 3. Найти вероятность того, что при 100 бросках монеты число выпадений герба окажется в пределах от 40 до 60.

Применим формулу (14.8), учитывая, что n = 0,5. Тогда np = 100·0,5 = 50, $\sqrt{npq} = \sqrt{100 \cdot 0{,}5 \cdot (1 - 0{,}5)} = 5.$

Тогда, если $40 < Y < 60,\ -2 < \dfrac{Y - 50}{5} < 2.$ Следовательно,

$$p(40 < Y < 60) = p\left(-2 < \frac{Y - 50}{5} < 2\right) = \Phi(2) - \Phi(-2) = 0{,}9772 - 0{,}0228 = 0{,}9544.$$

Пример 4. В условиях предыдущего примера найти вероятность того, что выпадет 45 гербов.

Найдем $x = \dfrac{45 - 50}{5} = -1$, тогда

$$p_{100}(45) \approx \frac{1}{5} \cdot \varphi(-1) = \frac{1}{5} \cdot \varphi(1) = \frac{1}{5} \cdot 0{,}2420 = 0{,}0484.$$