14. Основы теории вероятностей

Теорией вероятностей называется математическая наука, изучающая оценивание, сопоставление массовых случайных событий и закономерности в множествах таких событий.

Теория вероятностей как предмет изучения имеет математические модели реальной действительности. Однако модели теории вероятностей обладают спецификой, заключающейся в том, что в них учтен фактор случайности.

Методы теории вероятностей не отменяют и не упраздняют случайности, непредсказуемости исхода отдельного опыта, но дают возможность предсказать, с каким-то приближением, средний суммарный результат массы однородных случайных явлений. Чем большее количество однородных случайных явлений фигурирует в задаче, тем отчетливее выявляются присущие им специфические законы, тем с большей уверенностью и точностью можно осуществлять научный прогноз.

Цель вероятностных (статистических) методов – в том, чтобы, минуя слишком сложное (и зачастую практически невозможное) исследование отдельного случайного явления, обратиться непосредственно к законам, управляющим массами таких явлений. Изучение этих законов позволяет не только осуществлять прогноз в области случайных явлений, но и целенаправленно влиять на ход этих явлений, контролировать их, ограничивать сферу действия случайности, сужать ее влияние на практику.

Математические законы теории вероятностей – это отражение реальных статистических законов, объективно существующих закономерностей в массовых случайных явлениях природы. К изучению этих явлений теория вероятностей применяет математический метод и по своему методу является одним из разделов математики, столь же точным и строгим, как другие математические науки.

14.1. Случайные события

Определения и свойства вероятности

Для количественного сравнения между собой событий по степени возможности их появления вводится определенная мера, которая называется вероятностью события. Существуют различные подходы к определению вероятности события, каждое из них имеет свою область применения.

Будем говорить, что события в данном испытании образуют *полную группу*, если в результате испытания обязательно появится хотя бы одно из них. Рассмотрим полную группу равновозможных и несовместных случайных событий. Событие такой группы называется *благоприятствующим событию A*, если появление этого события всегда влечет за собой появление события *A*.

Классическое определение относится к равновозможным и несовместным исходам испытания, образующим полную группу. *Классической вероятностью* события *A* называется отношение числа *m* исходов, благоприятствующих событию *A*, к общему числу *n* всех возможных исходов испытания, если эти исходы равновозможны, несовместны и образуют полную группу: $P(A)=\frac{m}{n}$.

Вероятность обладает следующими свойствами:

1) Каждому случайному событию *A* ставится в соответствие неотрицательное действительное число, которое является вероятностью события *A*: $P(A)=p$.

2) Вероятность достоверного события равна единице: $P(\Omega)=1$.

3) Вероятность невозможного события равна нулю: $P(\varnothing)=1$.

4) Вероятность случайного события есть положительное число, заключенное между нулем и единицей: $0<P(A)<1$.

Геометрическая вероятность применяется в задачах, сводящихся к случайному бросанию точки на конечный участок прямой, плоскости или пространства.

Пусть дана область G, в которой содержится другая областьg.

Назовем мерой этих областей ($mes\ G$, $mes\ g$) их длину (в случае, когда G и g – отрезки или дуги), площадь (когда G и g – плоские фигуры), объем (когда G и g – пространственные фигуры).

Пусть в область G наудачу бросается точка, причем попадания в любую точку области – равновозможные события. Как оценить числом «шансы» попадания этой точки в область g (событие A)? Будем предполагать, что вероятность попадания точки в любую часть области g пропорциональна мере этой части и не зависит от ее формы и расположения в области G.

Геометрической вероятностью появления случайной точки внутри некоторой области g называется отношение меры этой области g к мере всей области G, в которой может появиться данная точка: $P(A) = \dfrac{mes\ g}{mes\ G}$.

Относительной частотой события A называют отношение числа m исходов, в которых появилось событие A, к общему числу n проведенных испытаний: $P(A) = \dfrac{m}{n}$. Принципиальное отличие классической вероятности от относительной частоты состоит в том, что классическую вероятность вычисляют, не производя опыта, тогда как относительная частота – непосредственный результат опыта. Многочисленные наблюдения показывают, что для значительного большинства случайных событий существует такое число p, свое для каждого события, что относительная частота при достаточно большом числе испытаний мало отличается от этого числа p (кроме редких случаев). Такое число p и называют статистической вероятностью события. То есть с*татистическая вероятность* приближенно равна относительной частоте (или числу, близкому к ней) при достаточно большом числе испытаний.

В качестве еще одного свойства вероятности приведем следующее утверждение (теорему сложения вероятностей). Вероятность появления хотя бы одного из двух событий равна сумме вероятностей этих событий без вероятности их произведения: $P(A+B) = P(A) + P(B) - P(A \cdot B)$. В случае *несовместных* событий это равенство приобретает вид:

$P(A+B)=P(A)+P(B)$. Это утверждение следует из комбинаторного правила сложения. Кроме того, для решения задач бывает необходимо использовать комбинаторное *правило произведения*: если один объект *A* можно выбрать *k* способами, а другой объект *B* после этого (в том числе, если выбор *B* осуществляется независимо от *A*) можно выбрать *l* способами, то пары объектов *A* и *B* (в указанном порядке) можно выдрать $k \cdot l$ способами. Это правило распространяется на любое число объектов.

Пример. Игральный кубик брошен один раз. Какова вероятность выпадения четного числа очков?

Решение. Возможны шесть исходов испытания: выпадение 1, 2, 3, 4, 5, 6 очков ($n=6$). Эти исходы равновозможны, несовместны и образуют полную группу. Событию *A* (выпадению четного числа очков) благоприятствуют $m=3$ случаев (выпадение 2, 4, 6 очков). Следовательно, $P(A)=\frac{3}{6}=\frac{1}{2}$.

Пример. Бросают два игральных кубика. Какова вероятность того, что сумма выпавших очков равна 4?

Решение

Подсчитаем все возможные исходы испытания. И для одного и для другого кубика в отдельности таких исходов 6. Так как каждый из 6 исходов для одного кубика может сочетаться с любым из 6 исходов для другого, то число всех возможных исходов равно: $n=6\cdot 6=36$.

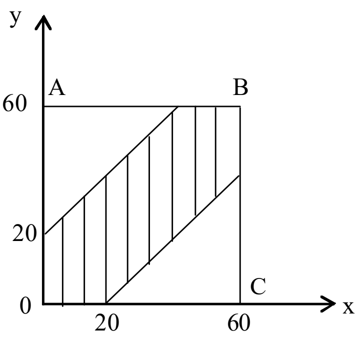
Подсчитаем теперь благоприятные варианты, т.е. варианты, где сумма очков равна 4. Таких вариантов три: (1,3), (3,1), (2,2). Следовательно, $P(A)=\frac{3}{36}=\frac{1}{12}$.

Пример. (*Задача о встрече*). Два лица *A* и *B* договорились встретится в определенном месте между 12 и 13 часами. Пришедший первым ждет другого 20 минут, после чего уходит. Найти вероятность встречи *A* и *B*, если каждый из них наудачу выбирает момент своего прихода (между 12 и 13 часами) и моменты прихода независимы.

Решение

Обозначим моменты прихода лиц A и B соответственно x (мин), и y (мин), взяв за начало отсчета 12 часов. Так как A и B должны были встретится в течение 60 минут, то всевозможные исходы определяются системой неравенств: $0 \le x \le 60$, $0 \le y \le 60$.

Благоприятными исходами будут те значения из них, для которых абсолютная величина разности между моментами x и y приходов лиц A и B не превышает 20 минут: $|x - y| \le 20$ – дополнительное неравенство.



По смыслу задачи все исходы равновозможные, однако исходов бесконечное множество, поэтому применять здесь классическое определение вероятности нельзя. Воспользуемся геометрической вероятностью.

Системе неравенств соответствует квадрат $OABC$. Его площадь $S_G = 60^2 = 3600$.

Дополнительное неравенство (2) равносильно системе $-20 \le x - y \le 20$, или $\left[\begin{array}{l} y \le x + 20, \\ y \ge x - 20. \end{array}\right.$ Эта система неравенств выполняется для тех точек квадрата $OABC$, которые лежат не выше прямой $y = x + 20$ и не ниже прямой $y = x - 20$ (на рисунке эта область заштрихована). Площадь заштрихованной фигуры легко вычислить, отняв от площади квадрата $OABC$ площади двух равных прямоугольных треугольников с катетами, равными 40: $S_g = 3600 - 2 \cdot \frac{1}{2} \cdot 40 \cdot 40 = 2000$.

По геометрическому определению вероятности окончательно найдем:

$$P(A) = \frac{S_g}{S_G} = \frac{2000}{3600} = \frac{5}{9}.$$

Применение комбинаторики к вычислению вероятности

Для решения целого ряда задач необходимо знать некоторые сведения из комбинаторики – раздела математики, который изучает различные комбинации из элементов заданного множества.

Даны n различных элементов. Составим из них (не повторяясь в использовании элементов) всевозможные группы по m элементов в каждой (не путать с n и m в понятии классической вероятности!).

Сочетаниями называются такие наборы, которые отличаются одна от другой хотя бы одним элементом (порядок элементов роли не играет). Число сочетаний из n элементов по m $C_n^m = \dfrac{n!}{m!\,(n-m)!}$, где $n! = 1\cdot 2\cdot 3\cdot \ldots \cdot n$ и принято считать, что $0! = 1$ и $1! = 1$.

Наборы, которые отличаются одна от другой как самими элементами, так и их порядком, называются *размещениями.* Число размещений из из n элементов по m $A_n^m = \dfrac{n!}{(n-m)!}$.

Наборы, которые отличаются одна от другой только порядком элементов ($m = n$), называются *перестановками.* Их число $P_n = n!$

Аналогичные комбинаторные величины известны и для «схемы с возвращениями», когда элементы исходного множества (или их дубликаты) можно использовать многократно:

– Количество «перестановок с повторениями из n_1, n_2, …, n_m элементов (наборы по (n_1+n_2+…+n_m) элементов в каждой, состоящие из одних и тех же m исходных элементов, повторяющихся n_1, n_2, …, n_m раз, и отличающиеся только порядком их расположения: $\overline{P}_n = \dfrac{P_{n_1+n_2+\ldots+n_m}}{P_{n_1}\cdot P_{n_2}\cdot\ldots\cdot P_{n_m}} = \dfrac{(n_1+n_2+\ldots+n_m)!}{n_1!\cdot n_2!\cdot\ldots\cdot n_m!}$;

– Количество «размещений с повторениями из n элементов по m» $\overline{A}_n^m = \underbrace{n\cdot n\cdot\ldots\cdot n}_{m\,раз} = n^m$

– количество «сочетаний с повторениями из n элементов по m» $\overline{C}_n^m = C_{n+m-1}^m$

Пример. В урне 10 белых и 15 черных шаров, которые тщательно перемешаны (в дальнейшем это всегда будет подразумеваться). Наудачу из урны извлекают одновременно три шара. Чему равна вероятность того, что все шары черные?

Решение

Как правило, вначале лучше считать все возможные исходы испытания. При этом цвет шаров роли не играет. В данной задаче из 25 шаров берем 3 шара. Так как порядок шаров не имеет значения, то это сочетания. Их число найдем по формуле: $n = C_{25}^{3} = \frac{25!}{3!\cdot 22!} = \frac{23\cdot 24\cdot 25}{1\cdot 2\cdot 3} = 2300$.

Благоприятный исход испытания – все три шара черные. Их можно извлечь из 15 черных шаров, поэтому $m = C_{15}^{3} = \frac{15!}{3!\cdot 12!} = \frac{13\cdot 14\cdot 15}{1\cdot 2\cdot 3} = 455$.

Искомая вероятность $p = \frac{455}{2300} = \frac{91}{460}$.

Пример. В партии из 10 деталей имеется 7 стандартных. Найти вероятность того, что среди одновременно извлеченных наудачу шести деталей ровно 4 детали стандартные.

Решение

$n = C_{10}^{6}$. При подсчете m часто допускают ошибку, считая, что $m = C_{7}^{4}$ (4 стандартные детали можно взять лишь из 7 стандартных). При этом не учитывается, что из партии деталей берут не 4, а 6 деталей, т.е. одновременно с четырьмя стандартными деталями надо взять две нестандартные из трех нестандартных, что можно сделать C_{3}^{2} способами. По правилу произведения $m = C_{7}^{4}\cdot C_{3}^{2}$. Итак, $p = \frac{C_{7}^{4}\cdot C_{3}^{2}}{C_{10}^{6}} = \frac{1}{2}$.

Пример. На полке помещены 7 книг разных авторов и трехтомник одного автора. Сколькими способами можно расставить эти книги на полке так, чтобы книги одного автора стояли рядом?

Решение. Представим себе 3 книги одного автора как одну книгу. Тогда получаем 8 книг, которые можно расставить на полке P_8 способами. Учитывая при этом, что 3 книги одного автора можно переставлять друг с другом P_3

способами, и пользуясь правилом умножения, находим общее число всех способов расстановки книг на полке при указанном условии: $P_3 \cdot P_8 = 3! \cdot 8! = 241920$.

Пример. Из цифр 0, 1, 2, …, 9 наудачу выбирают 4 цифры и приставляют одна к другой. Какова вероятность того, что образуется четырехзначное число?

Решение

Число всех возможных исходов - это число всех различных групп, содержащих 4 цифры, взятые из 10 цифр. Так как числа могут отличаться друг от друга, как самими цифрами, так и их порядком, то такие группы являются размещениями. По формуле их число $n = A_{10}^4 = 10 \cdot 9 \cdot 8 \cdot 7$.

Посчитаем благоприятные исходы, т.е. количество четырехзначных чисел. Заметим, что такое число образует любая группа из 4-х цифр, если только нуль не стоит на первом месте. Например, группа 0127 не является четырехзначным числом, т.е. не является благоприятным исходом. Так как к нулю можно присоединять любые 3 цифры, взятые из оставшихся девяти, то число неблагоприятных исходов равно $A_9^3 = 9 \cdot 8 \cdot 7$.

Следовательно, $m = A_{10}^4 - A_9^3$ и $p = \dfrac{A_{10}^4 - A_9^3}{A_{10}^4} = 1 - \dfrac{9 \cdot 8 \cdot 7}{10 \cdot 9 \cdot 8 \cdot 7} = \dfrac{9}{10}$.

Условная вероятность Независимые события. Формула полной вероятности

Вероятностью события B при условии, что наступило событие A (условной вероятностью), называется число $P(B/A) = \dfrac{P(A \cdot B)}{P(A)}$, $\left(P(A) \neq 0\right)$.

Нетрудно показать, что условные вероятности обладают следующими свойствами «обычных» (безусловных) вероятностей.

1. $0 \le P(B/A) \le 1$.

2. Если $C = A + B$, причем события A и B несовместные, то для любого события D $P(C/D) = P(A/D) + P(B/D)$.

3. $P(\overline{B}/A) = 1 - P(B/A)$

На практике часто условную вероятность можно найти непосредственно. Тогда определение позволяет найти вероятность произведения двух событий.

Теорема умножения вероятностей. Вероятность совместного появления двух событий равна произведению вероятности одного из них на условную вероятность другого, вычисленную в предположении, что первое событие произошло: $P(A\cdot B)=P(A)\cdot P(B/A)=P(B)\cdot P(A/B)$.

Два события *A* и *B* называются *независимыми*, если вероятность одного из них не зависит от того, произошло или не произошло другое, т.е. $P(B/A)=P(B)$ и $P(A/B)=P(A)$.

Очевидно, в случае независимых событий *A* и *B* теорема умножения вероятностей примет вид $P(A\cdot B)=P(A)\cdot P(B)$, т.е. вероятность совместного появления двух независимых событий равна произведению вероятностей этих событий.

Простым следствием теорем сложения и умножения вероятностей является формула полной вероятности. Пусть событие *A* может наступить вместе с одним из событий B_1, $B_2,\ldots,B_n$, образующих полную группу несовместных событий. Тогда вероятность события *A* равна сумме произведений вероятностей каждого из событий B_1, $B_2,\ldots,B_n$, на соответствующую условную вероятность события $P(A/B_i)$: $P(A)=P(B_1)P(A/B_1)+P(B_2)P(A/B_2)+...+P(B_n)P(A/B_n)$. Эта формула называется *формулой полной вероятности*.

Допустим, что произведено испытание, в результате которого появилось событие *A*. Тогда вероятности гипотез $P(B_1)$, $P(B_2),\ldots,$ $P(B_n)$, которые были вычислены до опыта (априорные вероятности), могут измениться в соответствии с результатом опыта. Поэтому чтобы найти эти послеопытные (апостериорные) значения вероятностей применяют *формулу Байеса*: $P(B_i/A)=\dfrac{P(B_i)\cdot P(A/B_i)}{P(A)}$, $i=1,2,...,n$, где знаменатель $P(A)$ есть полная вероятность.

Пример. Известно, что в семье двое детей и хотя бы один из них мальчик (случай близнецов исключается). Найти вероятность того, что второй ребенок – мальчик, если вероятность рождения мальчика равна 0,5.

Решение.

Обозначим мальчика буквой М, девочку буквой Д и условимся первой буквой обозначать старшего ребенка.

Пусть событие A – «один (из двух) ребенок мальчик», событие B – «второй ребенок мальчик». Без дополнительного условия, что в семье есть мальчик, следовало бы рассмотреть четыре возможности, однако ситуацию ДД мы исключаем и имеем три варианта: ММ, МД, ДМ. Важно, что эти случаи равновозможны. Из них благоприятны к B два исхода ММ и ДМ, поэтому $P(B/A)=\frac{2}{3}$.

Пример. Имеются буквы разрезной азбуки А, А, Н, С, А, Н. Их наудачу один раз приставляют одна к другой слева направо. Какова вероятность того, что образуется слово «АНАНАС»?

Решение. Событие «слово АНАНАС» представим в виде комбинации (в данном случае, в виде совместного появления, т.е. произведения) следующих событий: «на первом месте буква А», «на втором месте буква Н», и т.д., «на шестом месте буква С». Поэтому $p=\frac{3}{6}\cdot\frac{2}{5}\cdot\frac{2}{4}\cdot\frac{1}{3}\cdot\frac{1}{2}\cdot 1=\frac{1}{60}$.

Пример. В урне 5 белых и 7 черных шаров. Один за другим вынимают два шара (без возврата). Какова вероятность того, что оба шара одинакового цвета?

Решение

Событие «оба шара одинакового цвета» означает, что оба шара белые или оба шара черные. В свою очередь событие «оба шара белые» раскладывается на элементарные события «первый шар белый» и «второй шар белый». Аналогично раскладывается событие «оба шара черные». Эти рассуждения удобно записать в виде условной схемы: $Б_1Б_2+Ч_1Ч_2$.

После этого ответ получится сразу на основании теорем сложения и умножения вероятностей: $p=\frac{5}{12}\cdot\frac{4}{11}+\frac{7}{12}\cdot\frac{6}{11}=\frac{31}{66}$.

Пример. Студент c знает не все экзаменационные билеты. Что для него выгоднее: отвечать первым или вторым?

Решение

Будем называть «хорошими» те билеты, которые студент знает, и «плохими» те билеты, которые он не знает. Допустим, что всего n билетов, из которых k «хороших» ($k<n$). Если студент c отвечает первым, то вероятность того, что он возьмет «хороший» билет (событие A) равна $P(A)=\frac{k}{n}$.

Допустим, что студент берет билет вторым. Первый студент может вынуть как «хороший» для нашего студента билет, так и «плохой», поэтому событие A представится в виде следующей комбинации: $A = XX + ПX$, где Х взят «хороший» билет, П – взят «плохой» билет. По теоремам сложения и умножения вероятностей найдем: $P(A) = \frac{k}{n} \cdot \frac{k-1}{n-1} + \frac{n-k}{n} \cdot \frac{k}{n-1} = \frac{k}{n(n-1)}(k-1+n-k) = \frac{k}{n}$.

Вероятность события A не изменилась, т. е. шансы вытянуть «хороший» билет остались прежними.

Пример. В первом ящике содержится 5 деталей, из которых 3 детали стандартные, во втором – 6 деталей, из которых 5 стандартные. Из наудачу взятого ящика наудачу берут деталь. Какова вероятность того, что эта деталь стандартная?

Решение

Пусть событие A – «вынута стандартная деталь». Возможны две гипотезы: событие B_1 – «деталь из первого ящика», B_2 – «деталь из второго ящика». Так как всего два ящика, то $P(B_1) = P(B_2) = \frac{1}{2}$.

По формуле полной вероятности получаем $p = \frac{1}{2} \cdot \frac{3}{5} + \frac{1}{2} \cdot \frac{5}{6} = \frac{43}{60}$.

Пример. Партия деталей изготовлена двумя рабочими. Первый рабочий изготовил 2/3 партии, второй – 1/3. Вероятность брака для первого рабочего 1% а для второго 10%. Для контроля взяли одну деталь, и она оказалась бракованной. Найти вероятность того, что эта деталь изготовлена вторым рабочим.

Решение

Событие «контрольная деталь бракованная» (событие A) может произойти лишь вместе с одним из двух несовместных событий: «деталь изготовлена первым рабочим» (событие B_1,) и «деталь изготовлена вторым рабочим» (событие B_2). Эти события являются гипотезами. Их вероятности до испытания равны: $\mathrm{P}(B_1) = \frac{2}{3}$, $\mathrm{P}(B_2) = \frac{1}{3}$.

Кроме того, известны условные вероятности: $\mathrm{P}\left(A/B_1\right)=0.01$, $\mathrm{P}\left(A/B_2\right)=0.1$.

По формуле полной вероятности найдем: $\mathrm{P}(A)=\frac{2}{3}\cdot 0.01+\frac{1}{3}\cdot 0.1=\frac{1}{3}\cdot 0.12$.

Надо переоценить вторую гипотезу B_2 учетом результата испытания, т.е. найти $P\left(B_2/A\right)$. По формуле Байеса $P\left(B_2/A\right)=\frac{1/3\cdot 0.1}{1/3\cdot 0.12}=\frac{5}{6}$.

Как видим, вероятность того, что контрольная бракованная деталь изготовлена вторым рабочим, велика, несмотря на то, что доля деталей, изготовленных им, в два раза меньше доли первого рабочего. Такой результат объясняется тем, что брак значительной части партии был по вине второго рабочего.

Повторные независимые испытания

До сих пор мы рассматривали вероятности событий, возникающих, в основном, в результате единичных испытаний. На практике часто встречается такая схема событий, при которой испытания повторяются, причем нас интересует не результат каждого испытания, в котором может появиться или не появиться событие *A*, а общее число появлений события *A* в серии испытаний. Например, пусть из партии поступивших изделий на контроль выдается *n* изделий, причем вся партия будет забракована, если среди проверяемых изделий окажется *k* бракованных. Возникает задача нахождения вероятности того, что событие *A* (брак) появится *k* раз из *n* испытаний.

Определение. Повторение независимых относительно события *A* испытаний, в каждом из которых вероятность появления события *A* одна и та же, называют *схемой Бернулли.*

Вероятность $P_n\left(k\right)$ того, что в серии из *n* независимых испытаний событие *A* появится ровно *k* раз (безразлично, в какой последовательности), если вероятность появления события *A* в одном испытании равна *p*, определяется по *формуле Бернулли* $P_n\left(k\right)=C_n^k p^k q^{n-k}$.

Вычисление вероятности $P_n\left(k\right)$ при больших *n* (без применения современных средств) представляет значительные технические трудности. Так, при $n=100$, $k=30$, $p=0.2$ получим: $P_{100}\left(30\right)=C_{100}^{30}\left(0.2\right)^{30}\cdot\left(0.8\right)^{70}$.

Оказывается, что при достаточно больших *n* можно найти приближенную замену формуле Бернулли, а именно $P_n\left(k\right) \approx \frac{1}{\sqrt{npq}} \cdot \varphi\left(x\right)$ при $x = \frac{k - np}{\sqrt{npq}}$, где $\varphi\left(x\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. Точность значений вероятностей $P_n\left(k\right)$ возрастает с увеличением *n*. Это утверждение и составляет в практическом аспекте содержание *локальной теоремы Лапласа.*

Для функции $\varphi(x)$ составлены таблицы ее приближенных значений. При пользовании таблицами следует иметь в виду, что $\varphi(x)$ – четная функция, т. е. $\varphi(-x) = \varphi(x)$, а также то, что при $x \geq 4$ $\varphi(x) \approx 0$. Формула Бернулли и локальная теорема Лапласа отвечают на один и тот же вопрос: какова вероятность того, что событие *A* в *n* независимых испытаниях произойдет *k* раз, если вероятность события *A* в каждом испытании равна *p*.

Пусть надо найти вероятность $P_n\left(k_1 \leq k \leq k_2\right)$ того, что в *n* независимых испытаниях событие *A* произойдет от k_1 до k_2 раз (например, от 10 до 20 раз). *Интегральная теорема Лапласа* позволяет без сложных вычислений приближенно найти вероятность того, что в условиях схемы Бернулли событие *A* в *n* испытаниях произойдет от k_1 до k_2 раз, если только *n достаточно велико*. Основное утверждение интегральной теоремы Лапласа состоит в следующем: $P_n\left(k_1 \leq k \leq k_2\right) \approx \Phi\left(x_2\right) - \Phi\left(x_1\right)$, где $x_1 = \frac{k_1 - np}{\sqrt{npq}}$, $x_2 = \frac{k_2 - np}{\sqrt{npq}}$.

Функция $\Phi(x)$ называется *функцией Лапласа*, или *интегралом вероятностей*. Она имеет вид: $\Phi\left(x\right) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$. Так как неопределенный интеграл $\int_0^x e^{-\frac{t^2}{2}} dt$ не может быть выражен в виде конечной комбинации элементарных функций, то для функции $\Phi(x)$ составлены таблицы. При пользовании таблицами следует учесть, что $\Phi(x)$ - нечетная функция, т. е. $\Phi\left(-x\right) = -\Phi\left(x\right)$, а при $x > 5$ можно принять $\Phi(x) \approx 0.5$.

Пример. Стрелок стреляет по мишени 3 раза. Вероятность попадания в мишень для данного стрелка равна 0.6. Какова вероятность ровно двух попаданий в мишень?

Решение. Стрелок должен из $n = 3$ выстрелов попасть $k = 2$ раз. По формуле Бернулли находим искомую вероятность $P_3(2) = C_3^2 (0.6)^2 \cdot (0.4)^{3-2} = 0.432$.

Пример. Монету бросают 4040 раз. Какова вероятность того, что герб выпадет ровно 2048 раз?

Решение

По условию $n = 4040$, $k = 2048$. Вероятность выпадения герба $p = 0.5$. Так как n велико, используем локальную теорему Лапласа. Находим: $x = \dfrac{k - np}{\sqrt{npq}} = \dfrac{2048 - 4040 \cdot 0.5}{\sqrt{4040 \cdot 0.5 \cdot 0.5}} \approx \dfrac{28}{37.78} \approx 0.74$. По таблице $\varphi(x)$ находим

$\varphi(0.74) = 0.3034$. Искомая вероятность: $P_{4040}(2048) \approx \dfrac{1}{37.38} \cdot 0.3034 \approx 0.008$.

Таким образом, событие, о котором идет речь, является практически невозможным.

Пример. В страховой компании застраховано от пожара 10000 человек. На основании статистических данных установлено, что вероятность пожара в течении трех лет равна 0.006. Страховой взнос составляет 12 тыс. руб. в год, а если за это время произойдет пожар, компания выплачивает пострадавшему 1000 тыс. руб. Найти вероятность того, что:

а) по истечении трех лет работы страховая компания потерпит убыток;

б) страховая компания получит прибыль не менее 40000 тыс. руб.

Решение

а) 10000 человек внесли в кассу компании 120000 тыс. руб., следовательно, страховая компания потерпит убыток, если за страховой суммой обратятся более 120 человек. Для нахождения вероятности этого события используем интегральную теорему Лапласа. Исходные данные: $n = 10000$, $p = 0.006$, $k_1 = 121$, $k_2 = 10000$.

Вычисляем $x_1 = \dfrac{121 - 10000 \cdot 0.006}{\sqrt{59.64}} \approx 7.9$, $x_2 = \dfrac{10000 - 10000 \cdot 0.006}{\sqrt{59.64}} \approx 1287.6$.

Так как $x_1 > 5$ и $x_2 > 5$, то $\Phi(x_1) \approx \Phi(x_2) \approx 0.5$. Следовательно, $P_{10000}(121 \le k \le 10000) = 0.5 - 0.5 = 0$.

б) Для получения прибыли не менее 40000 тыс. грн. страховая компания должна выдать не более (120000-40000) тыс. = 80000 тыс. руб. Это произойдет, если за страховой суммой обратятся не более 80 человек. Вероятность этого события найдем по интегральной теореме Лапласа при $k_1 = 0$, $k_2 = 80$.

Вычисляем $x_1 = \dfrac{0 - 10000 \cdot 0.006}{\sqrt{59.64}} \approx -7.77$, $x_2 = \dfrac{80 - 10000 \cdot 0.006}{\sqrt{59.64}} \approx 2.59$.

В таблице $\Phi(x)$ есть значения $\Phi(2.58) = 0.4951$ и $\Phi(2.60) = 0.4953$. Искомая вероятность равна $P_{10000}(0 \le k \le 80) \approx \Phi(2.59) - \Phi(-7.77) =$

$$= \Phi(2.59) + \Phi(7.77) = 0.4952 + 0.5000 = 0.9952.$$

Как видим, при данных условиях практически достоверно, что страховая компания за год не только не потерпит убыток, но и получит прибыль не менее 40000 тыс. руб.

14.2. Случайные величины

Одно из самых важных понятий теории вероятностей – понятие случайной величины. Случайной величиной называется переменная величина, которая в результате испытания принимает то или иное значение в зависимости от случайных обстоятельств. Случайные величины обозначают заглавными буквами латинского алфавита $X, Y, Z, \ldots$, а их возможные значения – соответствующими малыми буквами $x, y, z, \ldots$ с индексами. Очевидно, что для задания случайной величины недостаточно знать ее значения. Необходимо также знать вероятность появления этих значений. Таким образом, для задания случайной величины должно быть установлено соотношение, связывающее ее значения и соответствующие вероятности.

Законом распределения случайной величины называется всякое соотношение, устанавливающее связь между возможными значениями случайной величины и их вероятностями. Закон распределения можно задать табличным, графическим и аналитическим способами.

Случайные величины можно задавать с помощью интегральной функции распределения. *Интегральной функцией распределения* (или *функцией распределения*) называется функция $F(x)$, которая определяет для каждого значения x вероятность того, что случайная величина X примет значение, меньшее числа x $F(x) = P(X < x)$. Свойства функции распределения:

1. Значения функции распределения заключены в интервале $[0; 1]$.
2. Функции распределения является неубывающей функцией.
3. Вероятность того, что случайная величина X примет какое-либо значение из $(a; b)$, равна приращению функции распределения на этом интервале $P(a < X < b) = F(b) - F(a)$.

Дискретные случайные величины

Дискретной называют такую случайную величину, которая принимает отдельные, изолированные друг от друга значения из некоторого промежутка. Для дискретных случайных величин простейшей формой закона распределения является перечень всех возможных значений соответствующих вероятностей. Такую форму закона распределения называют рядом распределения (или рядом распределения вероятностей). Его обычно оформляют в виде таблицы:

X	x_1	x_2	…	x_n
p	p_1	p_1	…	p_n

В этой таблице значения x_i принято располагать в порядке возрастания, среди них, не должно быть одинаковых.

Так как случайная величина X обязательно примет или значение x_1 или значение x_2, или … значение x_n, то сумма событий $X = x_1,\ X = x_2, ..., X = x_n$ есть событие достоверное. Учитывая, что эти события несовместны, приходим к выводу, что в ряде распределения $p_1 + p_2 + ... + p_n = 1$. Если дискретная случайная величина принимает бесконечное множество значений, то должно выполняться условие: $p_1 + p_2 + ... + p_n + ... = 1$. Данное равенство представляет собой условие нормировки и используется для контроля результата в задачах.

Пример. Пусть случайная величина X – число очков, выпадающих при одном бросании игрального кубика. Построить ряд распределения этой случайной величины.

Решение

Очевидно, ряд распределения X имеет следующий вид:

X	1	2	3	4	5	6
p	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Контроль: $\frac{1}{6}+\frac{1}{6}+...+\frac{1}{6}=\frac{1}{6}\cdot 6=1$.

Пример. Вероятности попаданий при трех выстрелах соответственно равны $p_1 = 0.6,\ p_2 = 0.7,\ p_3 = 0.8$. Стрелок, имеющий 3 патрона, стреляет до первого попадания в мишень. Найти ряд распределения числа выстрелов.

Решение

Случайная величина X (число выстрелов) может принять лишь значения 1, 2, 3. Для составления ряда распределения необходимо найти вероятности этих значений.

Если $X=1$, то стрелок попал с первого раза, т.е. $P(X=1)=0.6$. Если $X=2$, то это означает, что первый раз стрелок не попал в мишень, а второй раз – попал, т. е. $P(X=2)=q_1 p_2=0.4\cdot 0.7=0.28$.

Аналогично, $\mathrm{P}(X=3)=q_1 q_2 p_3+q_1 q_2 q_3=0.096+0.4\cdot 0.3\cdot 0.2=0.12$.

Контроль*:* 0.6+0.28+0.12=1.

Ряд распределения имеет вид:

X	1	2	3
p	0.6	0.28	0.12

Непрерывные случайные величины

Непрерывной называется случайная величина, которая может принимать любое значение из некоторого интервала. Непрерывные случайные величины можно задавать с помощью дифференциальной функции (функции плотности распределения). *Плотностью вероятности* $f(x)$ называют предел*:* $f(x)=\lim\limits_{\Delta x\to 0}\dfrac{F(x+\Delta x)-F(x)}{\Delta x}$. Кривой распределения непрерывной случайной величины называют график ее плотности вероятности. Рассмотрим основные свойства плотности вероятности.

1. Плотность вероятности есть неотрицательная функция: $f(x)\geq 0$.

2. Интеграл с бесконечными пределами от плотности вероятности равен единице: $\int\limits_{-\infty}^{+\infty} f(x)\,dx=1$.

3. Если все значения случайной величины лежат в интервале (a,b), то $\int\limits_a^b f(x)\,dx=1$.

4. Вероятность того, что непрерывная случайная величина примет какое-либо значение из интервала (a,b) равна определенному интегралу от плотности распределения на этом интервале $P(a<X<b)=\int\limits_a^b f(x)dx$.

5. Связь между интегральной и дифференциальной функциями $f(x)=F'(x)$, $F(x)=\int\limits_{-\infty}^{x} f(x)dx$.

Пример. Дана плотность вероятности случайной величины *X:*

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

Найти: параметр *a*; функцию распределения $F(x)$; вероятность того, что *X* примет значение из интервала $\left(\frac{1}{2};2\right)$.

Решение

1. Определим параметр *a*. Так как функция $f(x)$ на различных участках задана разными выражениями, то $\int\limits_{-\infty}^{+\infty} f(x)dx$ разобьем на три интеграла: $\int\limits_{-\infty}^{+\infty} f(x)dx=\int\limits_{-\infty}^{0} 0dx+\int\limits_0^1 axdx+\int\limits_1^{+\infty} 0dx=1$. Отсюда $a\left.\frac{x^2}{2}\right|_0^1=1$, $a\cdot\frac{1}{2}=1$, $a=2$.

2. Найдем функцию $F(x)$. Так как вид функции $f(x)$ зависит от интервала изменения аргумента *x*, то рассмотрим три случая

а) $x\le 0$, $F(x)=\int\limits_{-\infty}^{0} 0dx=0.$

б) $0<x\le 1$, $F(x)=\int\limits_{-\infty}^{x} f(x)dx=\int\limits_{-\infty}^{0} 0dx+\int\limits_0^x 2xdx=x^2.$

в) $x>1$, $F(x)=\int\limits_{-\infty}^{x} f(x)dx=\int\limits_{-\infty}^{0} 0dx+\int\limits_0^1 2xdx+\int\limits_1^x 0dx=1.$

Итак, $F(x)=\begin{cases}0 & \text{при } x\le 0,\\ x^2 & \text{при } 0<x\le 1,\\ 1 & \text{при } x>1.\end{cases}$ Значения $F(x)$ при $x\le 0$ и $x>1$, как видим, соответствуют свойствам этой функции.

3. Вероятность попадания в указанный интервал находим по формуле $P\left(\frac{1}{2}<X<2\right)=\int\limits_{1/2}^{2} f(x)dx.$ Так как на различных участках интервала $\left(\frac{1}{2};2\right)$ функция $f(x)$ имеет различный вид, то разобьем интеграл на два интеграла: $P\left(\frac{1}{2}<X<2\right)=\int\limits_{1/2}^{1} 2xdx+\int\limits_{1}^{2} 0dx=\frac{3}{4}.$ Также можно определить указанную вероятность по формуле $P\left(\frac{1}{2}<X<2\right)=F(2)-F\left(\frac{1}{2}\right)=1-\left(\frac{1}{2}\right)^2=\frac{3}{4}.$

Пример. Дана плотность вероятности случайной величины X: $f(x)=\begin{cases}0 & \text{при } x\le 0,\\ x^2 & \text{при } 0<x\le 1,\\ 0 & \text{при } x>1.\end{cases}$ Найти значение функции распределения $F(x)$ при $x>1$.

Решение. Значение функции распределения $F(x)$ при $x>1$ определяется по формуле $F(x)=\int\limits_{-\infty}^{x} f(x)dx=\int\limits_{-\infty}^{0} 0dx+\int\limits_{0}^{1} x^2dx+\int\limits_{1}^{x} 0dx=\frac{1}{3}.$ Однако по свойствам функции распределения при $x>1$ должно быть $F(x)=1$. Ошибка кроется в самом условии примера. Плотность вероятности (как и функция распределения) не может быть задана произвольно, должны выполняться все ее свойства. В данном случае не выполняется свойство $\int\limits_{-\infty}^{+\infty} f(x)dx=1.$ Следовательно, данная в условии плотность вероятности не корректна и задача «не имеет решения», то есть правильный непротиворечивый ответ не может быть найден.

Основные числовые характеристики случайных величин

Закон распределения дает полную вероятностную характеристику случайной величины. Однако при решении многих задач нет необходимости характеризовать случайную величину исчерпывающим образом, а удобнее пользоваться некоторыми количественными показателями, которые давали бы в сжатой форме достаточную суммарную информацию о случайной величине. Иногда закон распределения случайной величины вообще неизвестен (или его трудно найти), а поэтому ее полное изучение невозможно, однако знание этих количественных показателей (числовых характеристик) все же позволяет решать многие вопросы, относящиеся к случайной величине.

В теории вероятностей и математической статистике применяются следующие числовые характеристики.

Математическое ожидание

Для дискретной случайной величины математическое ожидание определяется по формуле: $M(X)=\sum_{i=1}^{n} x_i p_i$. Для непрерывной случайной величины $M(X)=\int_{-\infty}^{+\infty} x\cdot f(x)\,dx$. Свойства математического ожидания:

1. Математическое ожидание постоянной величины равно самой постоянной: $M\left(C\right)=C$.

2. Математическое ожидание суммы двух случайных величин равно сумме математических ожиданий слагаемых: $M\left(X+Y\right)=M\left(X\right)+M\left(Y\right)$.

3. Математическое ожидание произведения двух независимых случайных величин равно произведению их математических ожиданий: $M\left(X\cdot Y\right)=M\left(X\right)\cdot M\left(Y\right)$.

4. Постоянный множитель можно выносить за знак математического ожидания: $M\left(C\cdot X\right)=C\cdot M\left(X\right)$.

Дисперсия

Для дискретной случайной величины дисперсия определяется по формуле $D(X)=\sum_{i=1}^{n} x_i^{\,2} p_i-\left(\sum_{i=1}^{n} x_i p_i\right)^2$. Для непрерывной случайной величины $D(X)=\int_{-\infty}^{+\infty} x^2\cdot f(x)\,dx-\left(\int_{-\infty}^{+\infty} x\cdot f(x)\,dx\right)^2$. Для практического вычисления дисперсии удобно пользоваться следующей формулой: $D(X) = M(X^2) - M^2(X)$, т.е. дисперсия равна разности между математическим ожиданием квадрата случайной величины и квадратом ее математического ожидания. Свойства дисперсии:

1. Дисперсия постоянной величины равна нулю: $D(C) = 0$.

2. Постоянный множитель можно вынести за знак дисперсии, возведя его в квадрат: $D(C \cdot X) = C^2 \cdot D(X)$.

3. Дисперсия суммы двух независимых случайных величин равна сумме дисперсий слагаемых: $D(X+Y) = D(X) + D(Y)$.

4. Дисперсия суммы постоянной величины и случайной равна дисперсии случайной величины: $D(C+X) = D(X)$.

Среднее квадратическое отклонение

Среднее квадратическое отклонение определяется по формуле $\sigma(X) = \sqrt{D(X)}$.

Пример. Найти математическое ожидание суммы числа очков, которые могут выпасть при одном бросании пяти игральных кубиков.

Решение

Составить ряд распределения суммы числа выпадающих очков в данном случае технически довольно затруднительно, но легко составить ряд распределения числа очков, которые могут выпасть при бросании одного игрального кубика. После нахождения математического ожидания этой случайной величины достаточно воспользоваться свойствами математического ожидания.

Ряд распределения числа очков, выпавших при бросании одного игрального кубика, имеет вид:

X	1	2	3	4	5	6
p	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Контроль: $\sum p_i = 1$.

Математическое ожидание этой случайной величины находим по формуле

$$M(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3.5.$$

Искомое математическое ожидание равно $3.5 + 3.5 + \ldots + 3.5 = 17.5$.

Пример. Случайная величина имеет следующую плотность вероятности: $f(x) = \frac{\lambda}{2} e^{-\lambda|x|}$ $(\lambda > 0, -\infty < x < \infty)$. Найти ее математическое ожидание.

Решение. Так как данная случайная величина является непрерывной, то $M\left(X\right)=\dfrac{\lambda}{2}\int\limits_{-\infty}^{\infty}xe^{-\lambda|x|}dx$. Функция $xe^{-\lambda|x|}$ нечетная, а промежуток интегрирования симметричен, поэтому $M(X)=0$.

Пример. Найти дисперсию случайной величины X, заданной рядом распределения:

X	1	2	3
p	0.3	0.2	0.5

Решение. Найдем вначале $M(X)$: $M(X)=1\cdot 0.3+2\cdot 0.2+3\cdot 0.5=2.2$. Для нахождения $M(X^2)$ составим аналогичную сумму с квадратами всех значений *X*. Тогда $M(X^2)=1\cdot 0.3+4\cdot 0.2+9\cdot 0.5=5.6$. Далее вычисляем дисперсию по формуле $D(X)=5.6-(2.2)^2=0.76$.

Пример. Плотность вероятности случайной величины *X* имеет вид:
$$f(x)=\begin{cases}0 & \text{при } x\le -a,\\ \dfrac{1}{a}\left(1-\dfrac{|x|}{a}\right) & \text{при } -a<x<a,\\ 0 & \text{при } x\ge a.\end{cases}$$
Найти дисперсию и среднее квадратическое отклонение этой случайной величины.

Решение

Найдем вначале математическое ожидание. Учитывая, что вне интервала $(-a,a)$ $f(x)=0$, имеем: $M(X)=\dfrac{1}{a}\int\limits_{-a}^{a}x\left(1-\dfrac{|x|}{a}\right)dx$. Подынтегральная функция нечетная, а отрезок интегрирования симметричен относительно точки $x=0$, поэтому $M(X)=0$. Далее определяем дисперсию
$$D(X)=\int\limits_{-\infty}^{+\infty}(x-0)^2\cdot f(x)dx=\int\limits_{-\infty}^{-a}0\,dx+\int\limits_{-a}^{a}x^2\frac{1}{a}\left(1-\frac{|x|}{a}\right)dx+\int\limits_{a}^{+\infty}0\,dx==\frac{1}{a}\int\limits_{-a}^{a}x^2\left(1-\frac{|x|}{a}\right)dx.$$
Подынтегральная функция четная, отрезок интегрирования симметричен относительно $x=0$. С учетом того, что при $x>0$ $|x|=x$, получим:
$$D(X)=\frac{2}{a}\int\limits_{0}^{a}\left(x^2-\frac{x^3}{a}\right)dx=\frac{a^2}{6},\quad \sigma(X)=\sqrt{D(X)}=\frac{a}{\sqrt{6}}.$$

Основные законы распределения случайных величин

Каждый закон распределения определяется плотностью вероятности, функцией распределения, числовыми характеристиками и вероятностью попадания на интервал.

Биномиальный закон распределения

Дискретная случайная величина X имеет *биномиальный закон распределения*, если она принимает значения 0, 1, 2, ..., m, ..., n с вероятностями $P(X=m)=C_n^m p^m q^{n-m}$, где $0<p<1$, $q=1-p$, $m=0, 1, 2, ..., n$.

Биномиальный закон распределения представляет собой закон распределения числа $X=m$ наступлений события A в n независимых испытаниях, в каждом из которых оно может произойти с одной и той же вероятностью p.

Ряд распределения биномиального закона имеет вид:

x_i	0	1	2	…	m	…	n
p_i	q^n	$C_n^1 pq^{n-1}$	$C_n^2 p^2 q^{n-1}$	…	$C_n^m p^m q^{n-m}$	…	p^n

Математическое ожидание случайной величины X, распределённой по биномиальному закону, $M(X)=np$, а её дисперсия $D(X)=npq$.

Закон распределения Пуассона

Дискретная случайная величина X имеет *закон распределения Пуассона*, если она принимает значения 0, 1, 2, ..., m, ... (бесконечное, но счётное множество значений) с вероятностями $P(X=m)=\dfrac{\lambda^m e^{-\lambda}}{m!}$, где $m=0,1,2,...$

Ряд распределения закона Пуассона имеет вид:

x_i	0	1	2	…	m	…
p_i	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\dfrac{\lambda^2 e^{-\lambda}}{2!}$	…	$\dfrac{\lambda^m e^{-\lambda}}{m!}$	…

Математическое ожидание и дисперсия случайной величины X, распределённой по закону Пуассона, совпадают и равны значению параметра λ этого закона, т. е. $M(X)=\lambda$ и $D(X)=\lambda$. При условии $p\to 0$, $n\to\infty$, $np\to\lambda=const$ закон распределения Пуассона является предельным случаем биномиального закона.

Геометрическое распределение

Дискретная случайная величина X имеет *геометрическое распределение*, если она принимает значения 1, 2, ..., m, ... (бесконечное, но счётное множество значений) с вероятностями $P(X = m) = pq^{m-1}$, где $0 < p < 1$, $q = 1 - p$, $m = 1, 2, ...$

Ряд геометрического распределения имеет вид:

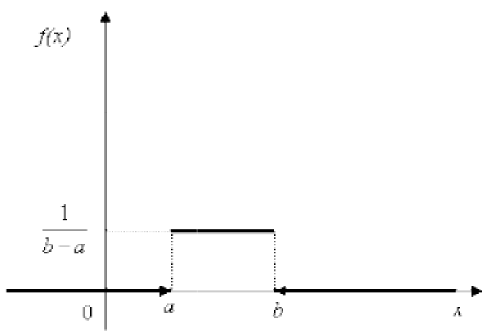
x_i	1	2	3	…	M	…
p_i	p	pq	pq^2	…	pq^{m-1}	…

Очевидно, что вероятности p_i образуют геометрическую прогрессию с первым членом p и знаменателем q (отсюда и название "геометрическое распределение"). Случайная величина $X = m$, имеющая геометрическое распределение, представляет собой число m испытаний, проведённых по схеме Бернулли, с вероятностью p наступления события в каждом испытании до первого положительного исхода. Математическое ожидание случайной величины X, имеющей геометрическое распределение с параметром p, $M(X) = \frac{1}{p}$, а её дисперсия $D(X) = \frac{q}{p^2}$, где $q = 1 - p$.

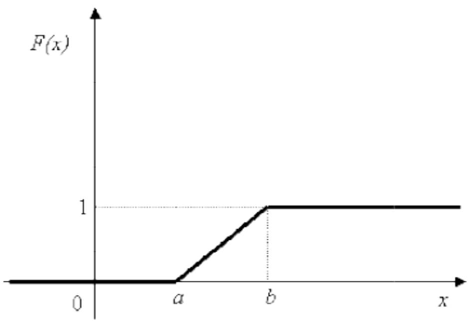
Равномерный закон распределения

Непрерывная случайная величина X имеет *равномерный закон распределения*, на отрезке $[a, b]$, если её плотность вероятности $f(x)$ постоянна на этом отрезке и равна нулю вне его, т.е.

$$f(x) = \begin{cases} 0 & при\ \ a < x, \\ \dfrac{1}{b-a} & при\ \ a \le x \le b, \\ 0 & при\ \ x > b. \end{cases}$$



Кривая распределения $f(x)$ приведена на рисунке справа.



Функция распределения случайной величины X, распределённой по равномерному закону, $F(x) = \begin{cases} 0 & \text{при } x < a, \\ \dfrac{x-a}{b-a} & \text{при } a \le x \le b, \\ 1 & \text{при } x > b, \end{cases}$ график этой функции представлен на рисунке слева.

Её математическое ожидание $M(X)=\dfrac{a+b}{2}$, а её дисперсия $D(X)=\dfrac{(b-a)^2}{12}$.

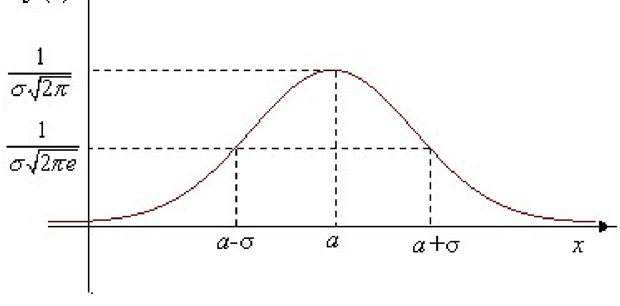
Равномерный закон распределения используется при анализе ошибок округления при проведении числовых расчётов (например, ошибка округления числа до целого распределена равномерно на отрезке [–0,5; 0,5]), в ряде задач массового обслуживания, при статистическом моделировании наблюдений, подчинённых заданному распределению. Так, случайная величина *X*, распределённая равномерно на отрезке [0; 1], называемая *случайным числом* от 0 до 1, служит исходным материалом для получения случайных величин с любым законом распределения.

Показательный (экспоненциальный) закон распределения

Непрерывная случайная величина *X* имеет *показательный (экспоненциальный) закон распределения* с параметром λ, если её плотность вероятности $f(x)$ имеет вид: $f(x)=\begin{cases}0 & \text{при } x<0,\\ \lambda e^{-\lambda x} & \text{при } x\ge 0.\end{cases}$ Функция распределения случайной величины *X*, распределённой по показательному закону, есть $F(x)=\begin{cases}0 & \text{при } x<0,\\ 1-e^{-\lambda x} & \text{при } x\ge 0,\end{cases}$ её математическое ожидание $M(X)=\dfrac{1}{\lambda}$, а её дисперсия $D(X)=\dfrac{1}{\lambda^2}$.

Показательный закон распределения играет большую роль в теории массового обслуживания и теории надёжности. Так, например, интервал времени *T* между двумя соседними событиями в простейшем потоке событий имеет показательное распределение с параметром λ – интенсивностью потока.

Нормальный закон распределения



Непрерывная случайная величина *X* имеет *нормальный закон распределения* с параметрами *a* и σ, если её плотность вероятности $f(x)$ имеет вид: $f(x)=\dfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}$. Кривая нормального распределения $f(x)$ (нормальная кривая или кривая Гаусса) приведена на рисунке справа. Математическое ожидание случайной величины *X*, распределённой по нормальному закону, равно параметру *a* этого закона, а её дисперсия – квадрату параметра σ, т.е. $M(X) = a$, $D(X) = \sigma^2$. Нормальный закон распределения случайной величины с параметрами $a = 0$ и $\sigma = 1$ называется *стандартным* или *нормированным*, а соответствующая нормальная кривая – *стандартной* или *нормированной*.

Функция распределения $F(x)$ случайной величины *X*, распределённой по нормальному закону, выражается через приведенную (нечетную) функцию Лапласа $\Phi(x)$ по формуле: $F(x) = \frac{1}{2} + \Phi\left(\frac{x-a}{\sigma}\right)$.

Перечислим наиболее важные свойства случайной величины, распределённой по нормальному закону.

1. Вероятность попадания случайной величины в интервал (x_1, x_2), равна $P(x_1 < X < x_2) = \Phi(t_2) - \Phi(t_1)$, где $t_1 = \frac{x_1 - a}{\sigma}$, $t_2 = \frac{x_2 - a}{\sigma}$.

2. Вероятность того, что отклонение случайной величины *X*, распределённой по нормальному закону, от математического ожидания *a* не превысит по абсолютной величине величину $\Delta > 0$ равна $P(|X - a| \le \Delta) = 2\Phi(t)$, где $t = \frac{\Delta}{\sigma}$.

3. «Правило трёх сигм»: если случайная величина *X* распределена нормально (с параметрами *a* и σ), то практически достоверно, что абсолютная величина её отклонения от математического ожидания не превосходит утроенного среднего квадратического отклонения, т.е. $P(|X - a| \le 3\sigma) \approx 1$.

Нормальный закон распределения наиболее часто встречается на практике. Главная особенность, выделяющая его среди других законов, состоит в том, что он является *предельным* законом, к которому приближаются другие законы при весьма часто встречающихся типичных условиях.

Пример. В рекламных целях торговая фирма вкладывает в каждую пятую единицу товара приз. Найти закон распределения числа призов, полученных при четырёх сделанных покупках.

Решение

Вероятность того, что в случайно сделанной покупке окажется денежный приз, равна $p = 1/5 = 0.2$. Случайная величина X – число покупок, в которые вложен денежный приз, имеет биномиальный закон распределения с параметрами $n = 4$ и $p = 0.2$. Ряд распределения X имеет вид (Значения $p_i = P(X = m)$, ($m = 0, 1, 2, 3, 4$) вычислены по формуле $p(X = m) = C_4^m \cdot 0.2^m \cdot 0.8^{4-m}$):

x_i	0	1	2	3	4
p_i	0.4096	0.4096	0.1536	0.0256	0.0016

Пример. Вероятность поражения цели равна 0,6. Производится стрельба по мишени до первого попадания (число патронов не ограничено). Требуется составить ряд распределения числа сделанных выстрелов, найти математическое ожидание и дисперсию этой случайной величины. Определить вероятность того, что для поражения цели потребуется не более трёх патронов.

Решение

Случайная величина X (число сделанных выстрелов) имеет геометрическое распределение с параметром $p = 0.6$. Ряд распределения X имеет вид:

x_i	1	2	3	...	m	...
p_i	0.6	0.24	0.096	...	$0.6 \cdot 0.4^m$	...

По формулам $M(X) = \dfrac{1}{p} = \dfrac{1}{0.6} = \dfrac{5}{3}$, $D(X) = \dfrac{q}{p^2} = \dfrac{0.4}{0.6^2} = \dfrac{10}{9}$.

Вероятность того, что для поражения цели потребуется не более трёх патронов равна $P(X \le 3) = P(X = 1) + P(X = 2) + P(X = 3) =$ $= 0.6 + 0.24 + 0.096 = 0.936$.

Пример. Автобус ходит регулярно с интервалом 10 минут. Пассажир приходит на остановку в случайный момент времени. Найти математическое ожидание и среднее квадратическое отклонение случайной величины X –времени ожидания автобуса. Какова вероятность того, что ждать пассажиру придётся не больше двух минут?

Решение

Случайная величина X (время ожидания автобуса на отрезке $[0;\, 10]$, мин. имеет равномерный закон распределения $f(x) = 1/10 = 0.1$. Поэтому средне время ожидания автобуса $M(X) = \dfrac{0+10}{2} = 5$ мин. Дисперсия времени ожидания автобуса $D(X) = \dfrac{(10-0)^2}{12} = \dfrac{100}{12} \approx 8.34$ мин^2, отсюда среднее квадратическое отклонение $\sigma(X) = \sqrt{D(X)} = \sqrt{\dfrac{100}{12}} = \dfrac{5\sqrt{3}}{3} \approx 2.89$. Вероятность события A – «пассажиру придется ждать не больше двух минут» $P(A) = P(X \le 2) = \int\limits_0^2 0.1\,dx = 0.2$.

Дополнительные числовые характеристики случайных величин

В некоторых задачах теории вероятностей требуется вычислить начальные и центральные моменты случайных величин. *Начальные моменты порядка* k определяются по формулам $M_k = \sum\limits_{i=1}^{n} x_i p_i$ и $M_k = \int\limits_{-\infty}^{\infty} x^k f(x)\,dx$ (для дискретной или непрерывной случайной величины).

Центральные моменты порядка k определяются по формулам $m_k = \sum\limits_{i=1}^{n} \left(x_i - M(X)\right)^k p_i$ и $m_k = \int\limits_{-\infty}^{\infty} \left(x - M(X)\right)^k f(x)\,dx$.

Для их вычисления удобно использовать следующие формулы, связывающие центральные моменты с начальными:

$$m_2 = M_2 - (M_1)^2;\; m_3 = M_3 - 3M_2M_1 + 2(M_1)^3,$$

$$m_4 = M_4 - 4M_3M_1 + 6M_2(M_1)^2 - 3(M_1)^4.$$

Заметим, что M_1 является математическим ожиданием, а m_2 – дисперсией случайной величины.

Моменты более высокого порядка используются редко, а по введенным моментам вычисляются такие характеристики, как *асимметрия* $As(X) = \dfrac{m_3}{\left(\sigma(X)\right)^3}$, *эксцесс* $Ex(X) = \dfrac{m_4}{\left(\sigma(X)\right)^4} - 3$. Для равномерного и нормального законов распределения асимметрия, характеризующая симметричность распределения, равна нулю. Эксцесс, характеризующий островершинность распределения, равен нулю для нормального закона.

Коэффициент вариации случайной величины вычисляется по формуле $V(X) = \dfrac{\sigma(X)}{M(X)}$. Для показательного распределения он равен единице.

Модой $M_o(X)$ непрерывной случайной величины X называется ее наиболее вероятное значение (значение, для которого плотность распределения $f(x)$ достигает максимума).

Медианой $M_e(X)$ непрерывной случайной величины X называется такое ее значение, для которого вероятность того, что случайная величина X примет значение, меньшее или большее медианы $M_e(X)$, равна 0,5: $P\left(X < M_e(X)\right) = \int\limits_{-\infty}^{M_e(x)} f(x)dx = P\left(X > M_e(X)\right) = \int\limits_{M_e(x)}^{\infty} f(x)dx = \dfrac{1}{2}$.

Квантилью уровня q называется такое значение x_q случайной величины X, при котором функция ее распределения принимает значение, равное q: $F(x_q) = P\left(X < x_q\right) = q$.

Квантили $x_{0,25}$ и $x_{0,75}$ получили название соответственно *верхнего* и *нижнего квантилей*.