

**ТЕОРИЯ ВЕРОЯТНОСТЕЙ
И МАТЕМАТИЧЕСКАЯ
СТАТИСТИКА**

КОНСПЕКТ ЛЕКЦИЙ

Ю. В. Щербакова

Данная книга представляет собой полный конспект лекций по курсу «Теория вероятности и математическая статистика». Предназначена для студентов высших учебных заведений. Предложенный материал поможет подготовиться к сдаче зачета или экзамена. Книга соответствует образовательному стандарту.

ЛЕКЦИЯ №1. Случайные события

1. Основные понятия теории вероятностей

1. Испытания и события. Виды случайных событий.

Случайным называется событие, если при реализации установленной совокупности условий S оно может либо произойти, либо не произойти. При этом событие будет рассматриваться как результат испытания. Например, человек стреляет по мишени, которая разделена на четыре части. Тогда выстрел — это испытание, а попадание в конкретную область мишени есть событие.

Если появление одного из событий исключает появление других событий в одном и том же испытании, то события называются несовместными. Например, была подброшена монета, при этом появление герба исключает появление надписи. Значит события «выпадение герба» и «выпадение надписи» — совместные.

Если в результате испытания появится хотя бы одно из событий, то можно говорить, что несколько событий образуют полную группу. Следовательно, если события, которые образуют полную группу, попарно несовместны, то при испытании появится одно и только одно из этих событий.

Например, человек старался выстрелить по цели. Непременно случится одно из событий: попадание или промах. Значит можно сказать, что эти два несовместных события образуют полную группу.

Равновозможными называют события, если есть повод полагать, что ни одно из них не является более возможным, чем другое. Например, при бросании монеты выпадение «герба» и выпадение надписи являются равновозможными событиями. Ведь монета правильной цилиндрической формы изготовлена из однородного материала, а присутствие чеканки не оказывает влияния на выпадение той или иной стороны монеты.

2. Определение вероятности.

Вероятность события A — это отношение числа благоприятствующих этому событию исходов к общему числу всех равново-

можных несовместных элементарных исходов, которые образуют полную группу:

$$P(A) = m / n,$$

где m — число элементарных исходов, которые благоприятствуют A ;
 n — число всех возможных элементарных исходов испытания.

При этом предполагается, что элементарные исходы несовместны, равновозможны и образуют полную группу. Следовательно, можно записать следующие три *свойства*.

1. Вероятность достоверного события равна единице. Следовательно, если событие достоверно, то каждый элементарный исход испытания благоприятствует событию, тогда $m = n$, и

$$P(A) = m / n = n / n = 1.$$

2. Вероятность невозможного события равна нулю. Следовательно, если событие невозможно, то ни один из элементарных исходов испытания не благоприятствует событию, тогда $m = 0$, и

$$P(A) = m / n = 0 / n = 0.$$

3. Вероятность случайного события есть положительное число, заключенное между нулем и единицей. Следовательно, случайному событию благоприятствует лишь часть из общего числа элементарных исходов испытания, тогда $0 < m < n$, стало быть, $0 < m / n < 1$, и

$$0 < P(A) < 1 \text{ и } 0 \leq P(A) \leq 1.$$

3. Основные формулы комбинаторики.

Перестановки — это комбинации, которые состоят из одних и тех же n различных элементов и отличаются только порядком их расположения. При этом число всех возможных перестановок равно

$$P_n = n!,$$

где $n! = 1 \times 2 \times 3 \times \dots \times n$.

Размещения — это комбинации, которые составлены из n различных элементов по m элементов, отличающиеся составом элементов или их порядком. Число всех возможных размещений равно:

$$A_n^m = n(n-1)(n-2)\dots(n-m+1).$$

Сочетания — это комбинации, которые составлены из n различных элементов по m элементов, отличающиеся хотя бы одним элементом. Число сочетаний равно C :

$$C_n^m = n!/(m!(n-m)!).$$

Следует заметить, что числа размещений, перестановок и сочетаний связаны равенством:

$$A_n^m = P_m C_n^m.$$

Правило суммы. Если определенный объект A может быть избран из совокупности объектов m способами, а второй объект B может быть избран n способами, то в таком случае брать или A , или B можно $m + n$ способами.

Правило произведения. В том случае если объект A можно выбрать из совокупности объектов m способами и после каждого такого выбора объект B можно выбрать n способами, то пара объектов (A, B) в заданном порядке может быть выбрана mn способами.

4. Относительная частота. Устойчивость относительной частоты. Статистическая вероятность.

Относительная частота события — это отношение числа испытаний, в которых событие появилось, к общему числу фактически произведенных испытаний:

$$W(A) = m / n,$$

где m — число появлений события; n — общее число испытаний.

Следует отметить, что для определения вероятности не обязательно, чтобы испытания производились в действительности. Для определения относительной частоты необходимо, чтобы испытания были произведены фактически.

Значит, можно сказать, что вероятность вычисляют до опыта, а относительную частоту — после опыта.

Например, сделали 25 выстрелов по цели, при этом было зарегистрировано 15 попаданий. Значит, относительная частота поражения цели равна:

$$W(A) = 15/25.$$

По словам математика В. Е. Гмурмана, в различных опытах относительная частота изменяется мало (тем меньше, чем больше произведено испытаний), колеблясь около некоторого постоянного числа. Этим постоянным числом является вероятность появления события.

За статистическую вероятность события принимают относительную частоту или число, близкое к ней.

В том случае, если событие достоверно, то $m = n$, а относительная частота равна:

$$m/n = n/n = 1.$$

Значит, статистическая вероятность достоверного события равна единице.

Если событие невозможно, то $m = 0$, а относительная частота равна:

$$0/n = 0.$$

Следовательно, статистическая вероятность невозможного события равняется нулю.

Для любого события $0 \leq m \leq n$ относительная частота равна:

$$0 \leq m/n \leq 1.$$

Другими словами, статистическая вероятность любого события лежит в пределах между нулем и единицей.

Для наличия статистической вероятности события A необходимо:

- 1) чтобы была возможность производить ограниченное число испытаний, в каждом из которых событие A наступает или не наступает;
- 2) чтобы существовала устойчивость относительных частот появления события A в различных сериях довольно большого числа испытаний.

5. Геометрические вероятности.

Недостаток классического определения вероятности состоит в том, что оно неприменимо к испытаниям с бесконечным числом исходов. Для того чтобы его преодолеть, вводят геометрические вероятности — вероятности попадания точки в область (например, отрезок, часть плоскости).

Допустим, отрезок l есть часть отрезка L , на котором наудачу поставлена точка. Это значит, что могут выполняться следующие предположения: точка может оказаться в любой точке отрезка L , вероятность попадания точки на отрезок l пропорциональна длине этого отрезка и не зависит от его расположения относительно отрезка L . Тогда вероятность попадания точки на отрезок l равна:

$$P = \text{Длина } l / \text{Длина } L.$$

Например, на отрезок OA длины L числовой оси O_x наудачу поставлена точка $B(x)$, и необходимо найти вероятность того, что меньший из отрезков OB и BA имеет длину, большую $L/3$. Пусть вероятность попадания точки на отрезок пропорциональна длине отрезка и не зависит от его расположения на числовой оси, тогда можно разбить отрезок OA точками C и D на три равные части. Решение будет верно, если точка $B(x)$ попадет на отрезок CD длины $L/3$. В этом случае вероятность равна:

$$P = (L/3) / L = 1/3.$$

Допустим, плоская фигура k является частью плоской фигуры K , на которую наудачу поставлена точка. Это значит, что могут выполняться следующие предположения: поставленная точка может оказаться в любой точке фигуры K , при этом вероятность попадания брошенной точки на фигуру k пропорциональна площади этой фигуры и не зависит ни от ее расположения относительно K , ни от формы k . Тогда вероятность нахождения точки в фигуре k равна:

$$P = \text{Площадь } k / \text{Площадь } K.$$

2. Теорема сложения вероятностей

1. Теорема сложения вероятностей несовместных событий.

Событие, включающее появление события A , или события B , или обоих этих событий, называется *суммой $A + B$ двух событий A и B* . Если два события A и B являются несовместными, то $A + B$ есть событие, включающее появление одного из этих событий.

Событие, заключающееся в появлении хотя бы одного из событий, называется *суммой нескольких событий*. К примеру, событие $A + B + C$ заключается в появлении одного из событий: A , B , C , A и B , A и C , B и C , A и B и C .

Пусть события A и B являются несовместными, и известны вероятности этих событий. С помощью теоремы сложения можно узнать, как найти вероятность того, что наступит либо событие A , либо событие B .

Теорема. Вероятность появления одного из двух несовместных событий, неважно какого, равняется сумме вероятностей этих событий:

$$P(A + B) = P(A) + P(B).$$

Доказательство. $m_1 + m_2$ есть число элементарных исходов, которое благоприятствует наступлению или события A , или события B . Поэтому

$$P(A + B) = (m_1 + m_2) / n = m_1 / n + m_2 / n,$$

где n — общее число допустимых элементарных исходов испытания;

m_1 — число исходов, которые благоприятствуют событию A ;

m_2 — число исходов, которые благоприятствуют событию B .

Причем, если

$$m_1 / n = P(A) \text{ и } m_2 / n = P(B),$$

то

$$P(A + B) = P(A) + P(B).$$

Следствие. Вероятность появления одного из нескольких попарно несовместных событий, неважно какого, равняется сумме вероятностей этих событий:

$$P(A_1 + A_2 + \dots + A_n) = P(A_1) + P(A_2) + \dots + P(A_n).$$

Доказательство. Пусть даны три попарно несовместных события A , B и C . Значит, появление одного из них равнозначно наступлению одного из двух событий, $A + B$ и C , поэтому:

$$P(A + B + C) = P[(A + B) + C] = P(A + B) + P(C) = P(A) + P(B) + P(C).$$

2. Полная группа событий. Противоположные события.

Теорема. Сумма вероятностей событий A_1, A_2, \dots, A_n , которые являются полной группой, равняется единице:

$$P(A_1) + P(A_2) + \dots + P(A_n) = 1.$$

Доказательство. Поскольку появление одного из событий полной группы достоверно, и вероятность достоверного события равняется единице, то

$$P(A_1 + A_2 + \dots + A_n) = 1.$$

Если два произвольные события полной группы являются несовместными, в этом случае пользуются теоремой сложения:

$$P(A_1 + A_2 + \dots + A_n) = P(A_1) + P(A_2) + \dots + P(A_n).$$

Следовательно, приравняв последние два равенства, получим:

$$P(A_1) + P(A_2) + \dots + P(A_n) = 1.$$

Например, директор школы принимает конверты с контрольными работами из городов A , B и C . Вероятность получения пакета из города A равна $0,6$, из города B — $0,1$. События «конверт получен из города A », «конверт получен из города B », «конверт получен из города C » есть полная группа. Тогда сумма вероятностей этих событий равна единице, т. е. $0,6 + 0,1 + p = 1$, тогда вероятность того, что очередной конверт будет получен из города C равна $p = 1 - 0,7 = 0,3$.

Два события, которые единственно возможны и образуют полную группу, называются *противоположными*. A — одно из двух противоположных событий, а другое \bar{A} .

Например, при стрельбе попадание и промах — противоположные события, где A — попадание, то \bar{A} — промах.

Теорема. Сумма вероятностей событий, которые противоположны, равна единице:

$$P(A) + P(\bar{A}) = 1.$$

Доказательство. Противоположные события образуют полную группу, а сумма вероятностей событий, которые образуют полную группу, равна единице.

Если p — вероятность одного из двух противоположных событий, то q — вероятность другого события. Следовательно,

$$P + q = 1.$$

Если требуется решить задачу на отыскание вероятности события A , то иногда лучше сначала найти вероятность события \bar{A} , а потом вычислить искомую вероятность по формуле

$$P(A) = 1 - P(\bar{A}).$$

3. Принцип практической невозможности маловероятных событий.

Принцип практической невозможности маловероятных событий заключается в следующем: если случайное событие очень маловероятно, то практически можно полагать, что в единичном испытании это событие не наступит.

Уровень значимости — это достаточно малая вероятность, при которой можно полагать, что событие практически невозможно. Как правило, на практике принимают уровни значимости, пределы которых равны 0,01 и 0,05, причем уровень значимости 0,01 называется однопроцентным.

Также с помощью проанализированного принципа можно делать предсказания не только о событиях, которые имеют малую вероятность, но и о событиях, вероятность которых примерно равна единице. В самом деле, если у события A вероятность близка к нулю, то вероятность противоположного события \bar{A} близка к единице.

Однако если событие A не появилось, то противоположное событие \bar{A} наступило. Следовательно, если вероятность случайного события очень близка к единице, то можно полагать, что в единичном испытании это событие наступит.

3. Теорема умножения вероятностей

1. Произведение событий. Условная вероятность.

Событие AB , состоящее в совместном появлении этих событий, называют *произведением двух событий A и B* . К примеру, если A — пригодная деталь, B — окрашенная деталь, то AB — деталь пригодная и окрашенная. Событие, при котором совместно появляются все несколько событий, называется *произведением нескольких событий*. Так, пусть A , B , C — представляет собой появление решки соответственно в первом, втором и третьем подбрасывании монеты, тогда ABC — будет являться выпадением решки во всех трех испытаниях.

Если случайное событие представлено как событие, которое при осуществлении совокупности условий S может произойти или не произойти, и если при вычислении вероятности события, кроме условий S , никаких других ограничений нет, то такая вероятность называется безусловной. Если же налагаются и другие дополнительные условия, то в таком случае вероятность события будет условной. Например, нередко подсчитывают вероятность события B при дополнительном условии, что совершилось событие A .

Вероятность события B , подсчитанная в предположении, что событие A уже наступило, называется условной вероятностью и обозначается $P_A(B)$.

Условная вероятность события B при условии, что событие A уже наступило:

$$P_A(B) = P(AB) / P(A), \text{ если } P(A) > 0.$$

2. Теорема умножения вероятностей.

Допустим известны вероятности $P(A)$ и $P_A(B)$ двух событий A и B . Для нахождения вероятности того, что появится и событие A , и событие B можно воспользоваться теоремой умножения.

Теорема. Вероятность совместного появления двух событий равняется произведению вероятности одного из них на условную вероятность другого, подсчитанную в догадке, что первое событие уже наступило:

$$P(AB) = P(A)P_A(B).$$

Доказательство. Если условная вероятность равна

$$P_A(B) = P(AB) / P(A),$$

то

$$P(AB) = P(A) P_A(B).$$

Замечание. Используя формулу $P(AB) = P(A)P_A(B)$ для события BA , получим:

$$P(BA) = P(B)P_B(A).$$

Так как событие BA такое же, как и событие AB , то по-другому ее можно записать следующим образом:

$$P(AB) = P(B)P_B(A).$$

Следовательно, равенство справедливо:

$$P(A)PA(B) = P(B)PB(A).$$

Следствие. Вероятность совместного появления нескольких событий равняется произведению вероятности одного из них на условные вероятности всех остальных. При этом вероятность каждого последующего события подсчитывается в предположении, что все предыдущие события уже появились:

$$P(A_1, A_2, A_3, \dots, A_n) = P(A_1)P_{A_1}(A_2)P_{A_1, A_2}(A_3) \dots = P_{A_1, A_2, \dots, A_{n-1}}(A_n).$$

Здесь $P_{A_1, A_2, \dots, A_{n-1}}(A_n)$ — это вероятность события A_n , найденная в предположении, что события A_1, A_2, \dots, A_{n-1} наступили.

Тогда можно записать

$$P(ABC) = P(A)P_A(B)P_{AB}(C).$$

3. Независимые события. Теорема умножения для независимых событий.

Положим, что вероятность события B не зависит от появления события A .

Событие B называется независимым от события A в том случае, если появление события A не меняет вероятности события B , другими словами, если условная вероятность события B равняется его безусловной вероятности:

$$P_A(B) = P(B).$$

Подставив $P(AB) = P(A)P_A(B)$ в $P(A)P_A(B) = P(B)P_B(A)$, получим:

$$P(A)P(B) = P(B)P_B(A).$$

Тогда

$$P_B(A) = P(A).$$

Другими словами, условная вероятность события A в предположении, что наступило событие B , равняется его безусловной вероятности, т. е. событие A не зависит от события B .

Если событие B не зависит от события A , то и событие A не зависит от события B . Это значит, что *свойство независимости событий взаимно*.

Теорема умножения $P(AB) = P(A)P_A(B)$ для независимых событий выглядит следующим образом:

$$P(AB) = P(A)P(B).$$

Это означает, что вероятность совместного появления двух независимых событий равна произведению вероятностей этих событий. Последним равенством пользуются для определения независимых событий.

Два события называются *независимыми* тогда, когда вероятность совмещения равна произведению вероятностей этих событий. Иначе события называют *зависимыми*.

Несколько событий называют *попарно независимыми*, если независимы каждые два из них.

Несколько событий называют *независимыми*, если, во-первых, независимы каждые два из них и, во-вторых, независимы каждое событие и все вероятные произведения других.

Следствие. Вероятность совместного появления нескольких независимых в совокупности событий, равна произведению вероятностей этих событий:

$$P(A_1, A_2, \dots, A_n) = P(A_1)P(A_2)\dots P(A_n).$$

Доказательство. Пусть даны три события A, B и C , совмещение которых равносильно совмещению событий A, B и C , поэтому

$$P(ABC) = P(AB \times C).$$

Поскольку события A, B и C в совокупности независимы, то независимы и события AB и C , а также A и B . Тогда на основании теоремы умножения для двух независимых событий получается:

$$P(AB \times C) = P(AB)P(C) \text{ и } P(AB) = P(A)P(B).$$

Тогда будем иметь:

$$P(ABC) = P(A)P(B)P(C).$$

4. Вероятность появления хотя бы одного события.

Допустим, в результате испытания могут появиться n независимых в совокупности событий или некоторые из них. При этом вероятности появления каждого из этих событий даны. Для нахождения вероятности того, что наступит хотя бы одно из этих событий, воспользуемся следующей теоремой.

Теорема. Вероятность появления хотя бы одного из событий A_1, A_2, \dots, A_n , которые независимы в совокупности, равняется разности между единицей и произведением вероятностей противоположных событий $\bar{A}_1, \bar{A}_2, \dots, \bar{A}_n$:

$$P(A) = 1 - q_1, q_2, \dots, q_n.$$

Доказательство. Событие, стоящее в появлении хотя бы одного из событий A_1, A_2, \dots, A_n обозначим через A . События A и

$$P(A) = 1 - P(\bar{A}_1, \bar{A}_2, \dots, \bar{A}_n) = 1 - P(\bar{A}_1), P(\bar{A}_2), \dots, P(\bar{A}_n) = 1 - q_1, q_2, \dots, q_n.$$

$$A_1, A_2, \dots, A_n$$

(ни одно из событий не наступило) противоположны. Поэтому сумма их вероятностей равна единице:

$$P(A) + P(\bar{A}_1, \bar{A}_2, \dots, \bar{A}_n) = 1.$$

Тогда

Частный случай. Пусть события A_1, A_2, \dots, A_n имеют одинаковую вероятность q . Тогда вероятность появления хотя бы одного из этих событий:

$$P(A) = 1 - q^n.$$

4. Следствия теорем сложения и умножения

1. Теорема сложения вероятностей совместных событий.

Совместными называют два события в том случае, если при одном и том же испытании появление одного из них не исключает появления другого. Например, если A — это появление пяти очков при бросании игральной кости, а B — появление нечетного числа очков, то события A и B можно назвать совместными.

Допустим, даны события A и B . Известно, что они совместны. Также даны вероятности этих событий и вероятность их совместного появления. Для того чтобы найти вероятность события $A + B$ (появится хотя бы одно из событий A и B), можно воспользоваться теоремой сложения вероятностей совместных событий.

Теорема. Вероятность того, что появится хотя бы одно из совместных событий, равняется сумме вероятностей этих событий

минус вероятность их совместного появления:

$$P(A + B) = P(A) + P(B) - P(AB).$$

Доказательство. По условию задачи события A и B совместны, следовательно, событие $A + B$ наступит в том случае, если наступит одно из трех несовместных событий: $A\bar{B}$, $\bar{A}B$ или AB . По теореме сложения вероятностей совместных событий:

$$P(A + B) = P(A\bar{B}) + P(\bar{A}B) + P(AB).$$

Если наступит одно из двух совместных событий: AB или $A\bar{B}$, то можно сказать, что событие A произойдет. Применяя теорему сложения вероятностей несовместных событий, получим:

$$P(A) = P(A\bar{B}) + P(AB),$$

а отсюда

$$P(A\bar{B}) = P(A) - P(AB).$$

Подобным образом будем иметь:

$$P(B) = P(\bar{A}B) + P(AB),$$

тогда

$$P(\bar{A}B) = P(B) - P(AB).$$

Следовательно, получится:

$$P(A + B) = P(A) + P(B) - P(AB).$$

Замечание 1. Используя полученную формулу, нужно обратить внимание на то, что события A и B могут быть как независимыми, так и зависимыми. Тогда для независимых событий запишется:

$$P(A + B) = P(A) + P(B) - P(A)P(B),$$

а для зависимых:

$$P(A + B) = P(A) + P(B) - P(A)P_A(B).$$

Замечание 2. Пусть события A и B несовместны, тогда их совмещение — это невозможное событие, а значит, $P(AB) = 0$. Следовательно, для несовместных событий можно записать:

$$P(A + B) = P(A) + P(B).$$

Формула $P(A + B) = P(A) + P(B) - P(A)P_A(B)$ справедлива и для совместных и для несовместных событий.

2. Формула полной вероятности.

Допустим событие A может наступать при условии появления одного из несовместных событий B_1, B_2, \dots, B_n . Эти события образуют полную группу. Также даны вероятности этих событий и условные вероятности $P_{B_1}(A), P_{B_2}(A), \dots, P_{B_n}(A)$ события A . Для нахождения вероятности события A можно воспользоваться следующей теоремой.

Теорема. Пусть несовместные события B_1, B_2, \dots, B_n образуют полную группу. Тогда вероятность события A , которое может наступить только при условии появления одного из этих несовместных событий, равняется сумме произведений вероятностей каждого из этих событий на соответствующую условную вероятность события A :

$$P(A) = P(B_1)P_{B_1}(A) + P(B_2)P_{B_2}(A) + \dots + P(B_n)P_{B_n}(A).$$

Эта формула есть формула полной вероятности.

Доказательство. Есть вероятность, что (по условию) событие A может наступить в том случае, если наступит одно из несовместных событий B_1, B_2, \dots, B_n , т. е. появление события A — это наступление одного, не важно какого, из несовместных событий B_1A, B_2A, \dots, B_nA . Учитывая теорему сложения, получим:

$$P(A) = P(B_1A) + P(B_2A) + \dots + P(B_nA).$$

В соответствии с теоремой умножения вероятностей зависимых событий будет:

$$P(B_1A) = P(B_1)P_{B_1}(A); P(B_2A) = P(B_2)P_{B_2}(A); \dots; P(B_nA) = P(B_n)P_{B_n}(A).$$

Формула полной вероятности:

$$P(A) = P(B_1)P_{B_1}(A) + P(B_2)P_{B_2}(A) + \dots + P(B_n)P_{B_n}(A).$$

3. Вероятность гипотез. Формулы Байеса.

Пусть несовместные события B_1, B_2, \dots, B_n образуют полную группу. Допустим событие A может наступить при условии появления одного из упомянутых выше событий. Так как заблаговременно неизвестно, какое из этих событий наступит, то их можно назвать гипотезами. Тогда вероятность появления события A равна:

$$P(A) = P(B_1)P_{B_1}(A) + P(B_2)P_{B_2}(A) + \dots + P(B_n)P_{B_n}(A).$$

Найдем условные вероятности $P_A(B_1), P_A(B_2), \dots, P_A(B_n)$.

В соответствии с теоремой умножения условная вероятность $P_A(B_1)$ равна:

$$P(AB_1) = P(A)P_A(B_1) = P(B_1)P_{B_1}(A),$$

Тогда

$$P_A(B_1) = \frac{P(B_1)P_{B_1}(A)}{P(A)}.$$

После преобразований получим

$$P_A(B_1) = \frac{P(B_1)P_{B_1}(A)}{P(B_1)P_{B_1}(A) + P(B_2)P_{B_2}(A) + \dots + P(B_n)P_{B_n}(A)}.$$

Условная вероятность любой гипотезы B_i ($i = 1, 2, \dots, n$) рассчитывается следующим образом:

$$P_A(B_i) = \frac{P(B_i)P_{B_i}(A)}{P(B_1)P_{B_1}(A) + P(B_2)P_{B_2}(A) + \dots + P(B_n)P_{B_n}(A)}.$$

Эти формулы носят название формул Байеса, благодаря которым можно оценить вероятности гипотез после того, как становится известным результат испытания, в результате которого появилось событие A .

5. Повторение испытаний

1. Формула Бернулли.

Если осуществляется несколько испытаний, к тому же вероятность события A в каждом испытании не зависит от исходов других испытаний, то такие испытания носят название *независимых относительно события A* .

Событие A в различных независимых испытаниях может иметь или различные вероятности, или одну и ту же вероятность.

Под *сложным событием* понимают совмещение нескольких отдельных *простых* событий.

Допустим, делается n независимых испытаний. В каждом из них событие A может появиться или не появиться. Будем думать, что во всяком испытании вероятность события A одна и та же, равная p . Значит, вероятность того, что событие A не наступит в каждом испытании также постоянна, причем равна она $q = 1 - p$.

Пусть необходимо подсчитать вероятность того, что при n испытаниях событие A произойдет ровно k раз, а не осуществится $n - k$ раз. Следует иметь в виду, что не требуется, чтобы событие A повторилось ровно k раз в определенной последовательности. К примеру, если события A появилось 3 раза в четырех испытаниях, то допустимы следующие сложные события: $AAA\bar{A}$, $AA\bar{A}A$, $A\bar{A}AA$, $\bar{A}AAA$. Таким образом, $AAA\bar{A}$ соответственно обозначает, что в первом, втором и третьем испытаниях событие A появилось, а в четвертом испытании оно не наступило. Другими словами, появилось противоположное событие \bar{A} .

Здесь под $P_n(k)$ понимают искомую вероятность. К примеру, $P_7(2)$ обозначает вероятность того, что в семи испытаниях событие появится ровно 2 раза, причем не наступит 5 раз. Искомую вероятность можно найти благодаря формуле Бернулли.

Вывод формулы Бернулли. Вероятность одного сложного события, заключающаяся в том, что в n испытаниях событие A наступит k раз и не наступит $n - k$ раз, равна $p^k q^{n-k}$. Это следует из теоремы умножения вероятностей независимых событий. Символ C_n^k означает, что таких сложных событий может быть столько, сколько можно составить сочетаний из n элементов по k элементам. Поскольку эти сложные события являются несовместными, то в соответствии с теоремой сложения вероятностей

несовместных событий искомая вероятность равняется сумме вероятностей всех возможных сложных событий. Так как вероятности всех этих сложных событий не отличаются, то вероятность появления k раз события A в n испытаниях равняется произведению вероятности одного сложного события и их числа:

$$P_n(k) = C_n^k p^k q^{n-k} \text{ или } P_n(k) = \frac{n!}{k!(n-k)!} p^k q^{n-k}.$$

Это есть формула Бернулли.

2. Локальная теорема Лапласа.

Использование формулы Бернулли при больших значениях n довольно проблематично, поскольку формула требует выполнения действий над гигантскими числами. При подсчете такого рода величин накапливаются погрешности, и окончательный результат может значительно отличаться от действительного.

Чтобы найти интересующую вероятность, не обязательно пользоваться формулой Бернулли. Это можно сделать в соответствии с локальной теоремой Лапласа, которая дает асимптотическую формулу. Данная формула позволяет приблизительно найти вероятность появления события ровно k раз в n испытаниях при условии, что число испытаний довольно велико. Эту теорему также называют теоремой Муавра — Лапласа.

Локальная теорема Лапласа. Если вероятность p появления события A в каждом испытании неизменна и не равна нулю и единице, то в таком случае вероятность $P_n(k)$ того, что событие A появится в n испытаниях ровно k раз, приблизительно равняется значению функции:

$$y = \frac{1}{\sqrt{npq}} \times \frac{1}{\sqrt{2\pi}} e^{-x^2/2} = \frac{1}{\sqrt{npq}} \times \varphi(x)$$

при $x = (k - np) / \sqrt{npq}$.

Таким образом, вероятность того, что событие A появится в n независимых испытаниях ровно k раз, приблизительно рав-

на:

$$P_n(k) \approx \frac{1}{\sqrt{npq}} \times \varphi(x),$$

где $x = (k - np) / \sqrt{npq}$.

3. Интегральная теорема Лапласа.

Пусть производится n испытаний, причем в каждом из них вероятность появления события A неизменна и равна p ($0 < p < 1$). Для того чтобы найти вероятность $P_n(k_1, k_2)$ того, что событие A появится в n испытаниях не менее k_1 и не более k_2 раз, воспользуемся интегральной теоремой Лапласа.

Теорема. Если вероятность p наступления события A в каждом испытании неизменна и не равна нулю и единице, то вероятность $P_n(k_1, k_2)$ того, что событие A появится в n испытаниях от k_1 до k_2 раз, приблизительно равна:

$$P_n(k_1, k_2) \cong \frac{1}{\sqrt{2\pi}} \int_{x'}^{x''} e^{-z^2/2} dz,$$

где $x' = (k_1 - np) / \sqrt{npq}$ и $x'' = (k_2 - np) / \sqrt{npq}$.

Преобразуем это соотношение следующим образом:

$$\begin{aligned} P_n(k_1, k_2) &\cong \frac{1}{\sqrt{2\pi}} \int_{x'}^0 e^{-z^2/2} dz + \frac{1}{\sqrt{2\pi}} \int_0^{x''} e^{-z^2/2} dz = \\ &= \frac{1}{\sqrt{2\pi}} \int_0^{x'} e^{-z^2/2} dz - \frac{1}{\sqrt{2\pi}} \int_{x'}^0 e^{-z^2/2} dz = \Phi(x'') - \Phi(x'). \end{aligned}$$

Следовательно, вероятность появления события A в n независимых испытаниях от k_1 до k_2 раз равна:

$$P_n(k_1, k_2) \cong \Phi(x'') - \Phi(x'),$$

где $x' = (k_1 - np) / \sqrt{npq}$ и $x'' = (k_2 - np) / \sqrt{npq}$.

4. Вероятность отклонения относительной частоты от постоянной вероятности в независимых испытаниях.

Пусть делается n независимых испытаний, в каждом из которых вероятность появления события A неизменна и равняется p ($0 < p < 1$).

Найдем вероятность того, что отклонение относительной частоты m/n от постоянной вероятности p по абсолютной величине не превышает заданного числа $\varepsilon > 0$, т. е.

$$P(|m/n - p| \leq \varepsilon).$$

Преобразуем и получим:

$$-\varepsilon \leq m/n - p \leq \varepsilon \text{ или } -\varepsilon \leq (m - np)/n \leq \varepsilon$$

Умножая на $\sqrt{n/pq}$, получим:

$$-\varepsilon\sqrt{n/(np)} \leq (m - np)/\sqrt{npq} \leq \varepsilon\sqrt{n/(pq)}.$$

Следует воспользоваться интегральной теоремой Лапласа, причем

$$x' = -\varepsilon\sqrt{n/(pq)} \text{ и } x'' = \varepsilon\sqrt{n/(pq)},$$

тогда

$$\begin{aligned} P(-\varepsilon\sqrt{n/(pq)} \leq (m - np)/\sqrt{npq} \leq \varepsilon\sqrt{n/(pq)}) &\approx \\ \approx \frac{1}{\sqrt{2\pi}} \int_{-\varepsilon\sqrt{n/(pq)}}^{\varepsilon\sqrt{n/(pq)}} e^{-z^2/2} dz &= \frac{2}{\sqrt{2\pi}} \int_0^{\varepsilon\sqrt{n/(pq)}} e^{-z^2/2} dz = 2\phi(\varepsilon\sqrt{n/(pq)}). \end{aligned}$$

Тогда окончательно получим:

$$P(|m/n - p| \leq \varepsilon \approx 2\phi(\varepsilon\sqrt{n/(pq)}).$$

Таким образом, вероятность осуществления неравенства $m/n - p \leq \varepsilon$ приблизительно равняется значению удвоенной функции Лапласа:

$$2\phi(x) \text{ при } x = \varepsilon\sqrt{n/(pq)}.$$

ЛЕКЦИЯ №2. Случайные величины

1. Виды случайных величин. Задание дискретной случайной величины

1. Случайная величина. Схема Бернулли.

Случайная величина — это величина, которая в результате испытания примет одно и только одно возможное значение, заранее неизвестное и зависящее от случайных причин, которые наперед не могут быть учтены.

Например, число родившихся девочек из двухсот новорожденных есть случайная величина, имеющая следующие возможные значения: 0, 1, 2, ..., 200.

Случайные величины обозначают прописными буквами X, Y, Z , а их возможные значения — строчными x, y, z . К примеру, если случайная величина X обладает четырьмя возможными значениями, то они будут обозначены так: x_1, x_2, x_3, x_4 .

Дискретные и непрерывные случайные величины.

Случайная величина, принимающая отдельные, изолированные возможные значения с определенными вероятностями, называется дискретной, *прерывной* случайной величиной. Следует знать, что число возможных значений дискретной случайной величины может быть конечным или бесконечным.

Непрерывная случайная величина — это та величина, которая может принимать все значения из определенного конечного или бесконечного промежутка. Ясно, что число возможных значений непрерывной случайной величины бесконечно много.

Схема Бернулли. Пусть A — случайное событие по отношению к некоторому опыту σ . Пространство элементарных событий, связанное с опытом σ , состоит только из двух элементов A и \bar{A} . Обозначим вероятности этих элементов соответственно через p и q ($p + q = 1$).

Допустим, что опыт σ неизменных условиях повторяется определенное число раз, например 3. Условимся трехкратное осуществление σ рассматривать как некий новый опыт S .

Если интересоваться только наступлением или ненаступлением A , то следует принять, что пространство элементарных событий, отвечающее опыту S , состоит из всевозможных строк длиной 3: (A, A, A) , (A, A, \bar{A}) , (A, \bar{A}, A) , (A, \bar{A}, \bar{A}) , (\bar{A}, A, A) , (\bar{A}, A, \bar{A}) , (\bar{A}, \bar{A}, A) , $(\bar{A}, \bar{A}, \bar{A})$, которые можно составить из A и \bar{A} . Каждая из указанных строк означает ту или иную последовательность появлений или неоявлений события A в трех опытах σ . Например, строка (A, \bar{A}, \bar{A}) означает, что в первом из опытов наступило A , а во втором и третьем — \bar{A} . Любая комбинация исходов трех опытов представляет собой тройку независимых событий. Но в таком случае элементарному событию (A, \bar{A}, \bar{A}) естественно приписать вероятность, равную $p \times q \times q$, событию $(\bar{A}, \bar{A}, \bar{A})$ — вероятность $q \times q \times q$ и т. д.

Мы приходим к следующей вероятностной схеме для опыта S (т. е. для трехкратного осуществления σ). Пространство Q элементарных событий есть множество из 23 строк. С каждой строкой сопоставляется в качестве вероятности число $p^k q^l$, где показатели степеней k и l определяют, сколько раз символы A и \bar{A} входят в выражение для данной строки.

Вероятностные схемы такого рода называются *схемами Бернулли*. В общем случае схема Бернулли определяется заданием двух чисел: натурального числа n и произвольного p , $0 \leq p \leq 1$. Пространство Q состоит из всевозможных строк длиной n , составленных из символов A и \bar{A} (число таких строк равно 2^n). Каждой строке приписывается вероятность $p^k (1-p)^l$, где k и l — число вхождений символов A и \bar{A} в данную строку.

2. Закон распределения вероятностей дискретной случайной величины.

Случайные величины могут обладать одинаковыми наборами возможных значений, а их вероятности — различными.

Закон распределения дискретной случайной величины — это есть соответствие между возможными значениями и их вероятностями. Задается он таблично, аналитически и графически.

Если закон распределения дискретной случайной величины задан в табличном виде, то первая строка таблицы заключает возможные значения, а вторая — их вероятности:

X	x_1	x_2	\dots	x_n
P	p_1	p_2	\dots	p_n

Поскольку случайная величина в одном испытании принимает одно и только одно возможное значение, можно сказать, что события $X = x_1, X = x_2, \dots, X = x_n$ образуют полную группу. Это означает, что сумма вероятностей этих событий (сумма вероятностей второй строки таблицы) равняется единице:

$$p_1 + p_2 + \dots + p_n = 1.$$

Следовательно, если множество возможных значений X бесконечно (счетно), то ряд $p_1 + p_2 + \dots$ сходится и его сумма равна единице.

Закон распределения дискретной случайной величины для наглядности изображают графически. С этой целью в прямоугольной системе координат строят точки (x_i, p_i) , после чего соединяют их отрезками прямых, а в результате полученная фигура называется многоугольником распределения.

3. Биномиальное распределение.

Допустим, происходит n независимых испытаний. Событие A в каждом из них может появиться или не появиться. Вероятность наступления события во всех испытаниях неизменна и равна p . Тогда вероятность того, что оно не появится, будет $q = 1 - p$.

Пусть дискретная случайная величина X — это число появлений события A в этих испытаниях. Для нахождения закона распределения величины X нужно определить возможные значения X , а также их вероятности. Ясно, что в n испытаниях событие A может либо не появиться, либо появиться 1 раз, либо 2 раза, ..., либо n раз. Следовательно, возможные значения X будут следующими: $x_1 = 0, x_2 = 1, x_3 = 2, \dots, x_{n+1} = n$. Для нахождения вероятности этих возможных значений можно использовать формулу Бернулли:

$$P_n(k) = C_n^k p^k q^{n-k},$$

где $k = 0, 1, 2, \dots, n$.

Данная формула и есть аналитическое выражение искомого закона распределения.

Биномиальное распределение вероятностей — это распределение, которое определяется формулой Бернулли. Этот закон получил название биномиального потому, что правая часть равенства может быть рассмотрена как общий член разложения бинома Ньютона:

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

Значит, первый член разложения p^n обуславливает вероятность наступления рассматриваемого события n раз в n независимых испытаниях. Второй член $np^{n-1}q$ обуславливает вероятность наступления события $n-1$ раз. Соответственно последний член q^n определяет вероятность того, что событие не появится ни разу. Биномиальный закон также можно записать в виде таблицы:

$$X \quad n \quad n-1 \quad \dots \quad k \quad \dots \quad 0 ;$$

$$P \quad p^n \quad np^{n-1}q \quad \dots \quad C_n^k p^k q^{n-k} \quad \dots \quad q^n .$$

4. Распределение Пуассона.

Допустим, делается n независимых испытаний, причем в каждом из них вероятность появления события A равняется p . Чтобы определить вероятности k появлений события в этих испытаниях, применяют формулу Бернулли. В том случае, когда n велико, применяют асимптотическую формулу Лапласа, но если вероятность события мала ($p \leq 0,1$), то этой формулой нельзя пользоваться. В этом случае используют асимптотическую формулу Пуассона.

Вычислим вероятность того, что при очень большом числе испытаний, в каждом из которых вероятность события очень мала, событие наступит ровно k раз. Следует отметить, что произведение np остается постоянным и равным $np = \lambda$. Среднее число появлений события при различных значениях n , остается постоянным.

Чтобы найти интересующую нас вероятность, можно воспользоваться формулой Бернулли:

$$P_n(k) = \frac{n(n-1)(n-2)\dots[n-(k-1)]}{k!} p^k (1-p)^{n-k}.$$

Поскольку $np = \lambda$, то $p = \lambda/n$, следовательно,

$$P_n(k) = \frac{n(n-1)(n-2)\dots[n-(k-1)]}{k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}.$$

Так как n очень велико, вместо $P_n(k)$ следует найти

Так как произведение np неизменно, то при $n \rightarrow \infty$ вероятность $p \rightarrow 0$. Следовательно,

$$\begin{aligned} P_n(k) &\cong \lim_{n \rightarrow \infty} \frac{n(n-1)(n-2)\dots[n-(k-1)]}{k!} \times \frac{\lambda^k}{n^k} \left(1 - \frac{\lambda}{n}\right)^{n-k} = \\ &= \frac{\lambda^k}{k!} \lim_{n \rightarrow \infty} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) \left(1 - \frac{\lambda}{n}\right)^{n-k} = \\ &= \frac{\lambda^k}{k!} \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n \times \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^{-k} = \frac{\lambda^k}{k!} \times e^{-\lambda} \times 1. \end{aligned}$$

Следовательно эта формула выражает закон распределения Пуассона вероятностей массовых (n велико) и редких (p мало) событий:

$$P_n(k) = \lambda^k e^{-\lambda} / k!.$$

5. Простейший поток событий.

Последовательность событий, наступающая в случайные моменты времени, называется *потоком событий*. Выделяют следующие свойства потоков.

Свойство стационарности: вероятность появления k событий на произвольном промежутке времени зависит только от числа k и от длительности t промежутка и не зависит от начала его отсчета. При этом предполагается, что разные промежутки времени являются непересекающимися.

Если поток имеет свойство стационарности, то вероятность появления k событий за промежуток времени длительности t есть функция, которая зависит лишь от k и t .

Свойство отсутствия последствия: вероятность появления k событий на любом промежутке времени не зависит от того, появлялись или не появлялись события в моменты времени, которые предшествовали началу рассматриваемого промежутка, т. е. условная вероятность появления k событий на произвольном

промежутке времени, подсчитанная при любых предположениях о том, что делалось до начала рассматриваемого промежутка, равняется безусловной вероятности. Следовательно, можно сказать, что предыстория потока не оказывает влияния на вероятность появления событий в ближайшем будущем.

Если потоку свойственно отсутствие последействия, то наблюдается взаимная независимость появлений того или иного числа событий в непересекающиеся промежутки времени.

Свойство ординарности: появление двух и более событий за небольшой промежуток времени невозможно, т. е. вероятность появления более одного события пренебрежимо мала по сравнению с вероятностью появления лишь одного события.

Если потоку свойственна ординарность, то за бесконечно малый промежуток времени может появиться не более одного события.

Поток событий, который обладает свойствами стационарности, отсутствия последействия и ординарности, называется *простейшим (пуассоновским) потоком*.

Замечание. Если поток есть сумма крайне большого числа независимых стационарных потоков, влияние каждого из которых на суммарный поток очень мало, то суммарный поток (при условии его ординарности) близок к простейшему.

Среднее число событий, появляющихся в единицу времени, называется *интенсивностью потока A* .

В случае если известна постоянная интенсивность потока, то вероятность появления k событий простейшего потока за время t определяется формулой Пуассона, которая отражает все свойства простейшего потока:

$$P_n(k) = (\lambda t)^k e^{-\lambda t} / k!.$$

В соответствии с формулой, вероятность появления k событий за время t , при данной интенсивности есть функция k и t , а это и характеризует свойство стационарности.

Пусть $k = 0$ и $k = 1$, тогда будем искать соответственно вероятности того, что события появились и, что появилось одно событие:

$$P_t(0) = e^{-\lambda t}, P_t(1) = \lambda t e^{-\lambda t}.$$

Вероятность появления более одного события определяется

исходя из следующего:

$$P_t(k > 1) = 1 - [P_t(0) + P_t(1)] = 1 - [e^{-\lambda t} + \lambda t e^{-\lambda t}].$$

Применяя разложение

$$e^{-\lambda t} = 1 - \lambda t + (\lambda t)^2 / 2! - \dots$$

после преобразований будем иметь:

$$P_t(k > 1) = (\lambda t)^2 / 2! + \dots$$

После сравнения $P_t(t)$ и $P_t(k > 1)$ приходим к выводу, что при малых значениях t вероятность появления более одного события пренебрежимо мала по сравнению с вероятностью наступления одного события. Данный факт обуславливает свойство ординарности. Формула Пуассона есть математическая модель простейшего потока событий.

6. Геометрическое распределение.

Пусть совершаются независимые испытания. Вероятность появления события A в каждом из них равняется p ($0 < p < 1$). Значит, вероятность того, что оно не появится, будет $q = 1 - p$. При появлении события A , испытания заканчиваются. Следовательно, если событие A появилось в k -м испытании, то в предыдущих $k-1$ испытаниях оно не появлялось.

Пусть X — дискретная случайная величина есть число испытаний, которые нужно осуществить до первого появления события A . Понятно, что возможные значения X и есть натуральные числа: $x_1 = 1, x_2 = 2, \dots$

Допустим в первых $k-1$ испытаниях событие A не наступило, а появилось в k -м испытании, тогда по теореме умножения вероятностей независимых событий вероятность этого «сложного события» найдется как

$$P(X = k) = q^{k-1}p.$$

Если допустить, что в формуле $k = 1, 2, \dots$, то получится геометрическая прогрессия с первым членом p и знаменателем q ($0 < q < 1$).

$$p, qp, q^2p, \dots, q^{k-1}p, \dots$$

Распределение $P(X = k) = q^{k-1}p$ по той же причине называется *геометрическим*.

Ряд $p, qp, q^2p, \dots, q^{k-1}p, \dots$ сходится, а его сумма равняется единице:

$$p/(1-q) = p/p = 1.$$

7. Гипергеометрическое распределение.

Допустим, в партия состоит из N изделий. В ней имеется M стандартных изделий, т. е. $M < N$. Из партии случайно извлекают n изделий (каждое изделие может быть отобрано с одинаковой вероятностью). При этом отобранное изделие перед отбором следующего не возвращается в партию, в связи с этим формула Бернулли здесь нельзя использовать. Пусть случайная величина X есть число m стандартных изделий среди n отобранных. Тогда возможные значения X будут следующими: $0, 1, 2, \dots, \min(M, n)$.

Для нахождения вероятности того, что $X = m$ (среди n отобранных изделий ровно m стандартных), применим классическое определение вероятности.

Совокупное число возможных элементарных исходов испытания равняется числу способов, которыми можно извлечь n изделий из N , другими словами, числу сочетаний C_N^n .

Отыщем число исходов, которые благоприятствуют событию $X = m$ (среди взятых n изделий ровно m стандартных). При этом m стандартных изделий можно отобрать из M стандартных изделий C_M^m способами, причем остальные $n-m$ изделий должны быть нестандартными. Взять $n-m$ нестандартных изделий из $N-m$ нестандартных изделий можно

$$C_{N-M}^{n-m}$$

способами. Таким образом, число благоприятствующих исходов по правилу умножения равно

$$C_M^m C_{N-M}^{n-m}.$$

Вероятность, которую нужно было найти, равняется отношению числа исходов, которые благоприятствуют событию $X = m$,

к числу всех элементарных исходов:

$$P(X = m) = \frac{C_M^n C_{N-M}^{n-m}}{C_N^n}.$$

Эта формула представляет собой распределение вероятностей, называемое *гипергеометрическим*.

Поскольку m — это случайная величина, можно сказать, что гипергеометрическое распределение определяется тремя параметрами: N , M , n .

Например, дано 50 изделий, среди которых 20 цветных. Для нахождения вероятности, что среди извлеченных 5 изделий окажется ровно 3 цветных, отметим, что $N = 50$, $M = 20$, $n = 5$, $m = 3$. Следовательно, искомая вероятность равна:

$$P(X = 3) = C_{20}^3 C_{30}^2 / C_{50}^5 = 0,234.$$

2. Математическое ожидание дискретной случайной

1. Числовые характеристики дискретных случайных величин. Математическое ожидание дискретной случайной величины.

Числовыми характеристиками случайной величины называются числа, описывающие случайную величину суммарно. Одной из важных числовых характеристик является математическое ожидание.

Математическое ожидание приблизительно равно среднему значению случайной величины.

При решении ряда задач достаточно, чтобы было известно математическое ожидание.

К примеру, если известно, что математическое ожидание числа попаданий первого стрелка больше, чем у второго, то первый стрелок в среднем попадает чаще, чем второй.

Математическое ожидание дискретной случайной величины.

Сумма произведений всех возможных значений дискретной случайной величины на их вероятности, называется математическим ожиданием дискретной случайной величины.

Допустим, случайная величина X может принимать лишь значения x_1, x_2, \dots, x_n , вероятности которых соответственно равняются p_1, p_2, \dots, p_n .

В этом случае математическое ожидание $M(X)$ случайной ве-

личины X обуславливается следующим равенством:

$$M(X) = x_1 p_1 + x_2 p_2 + \dots + x_n p_n.$$

Когда дискретная случайная величина X принимает счетное множество возможных значений, то в этом случае можно записать

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

При этом математическое ожидание существует при условии, что ряд в правой части равенства сходится абсолютно. Таким образом, математическое ожидание числа появлений события в одном испытании равняется вероятности этого события.

Допустим сделано n испытаний, в которых случайная величина X приняла m_1 раз значение x_1 . Соответственно m_2 раз приняла значение x_2 , ..., m_k раз значение x_k . При этом $m_1 + m_2 + \dots + m_k = n$. Следовательно, сумма всех значений, которые приняли X , равняется

$$x_1 m_1 + x_2 m_2 + \dots + x_k m_k.$$

Для отыскания среднего арифметического \bar{X} всех значений, которые были приняты случайной величиной, нужно разделить вычисленную сумму на общее число испытаний:

$$\bar{X} = (x_1 m_1 + x_2 m_2 + \dots + x_k m_k) / n$$

или по-другому

$$\bar{X} = (x_1 m_1 + x_2 m_2 + \dots + x_k m_k) / n$$

Поскольку отношение m_1 / n — это относительная частота W_1 значения x_1 , а m_2 / n — это относительная частота W_2 значения x_2 и так далее, то можно записать данное соотношение иначе:

$$\bar{X} = x_1 W_1 + x_2 W_2 + \dots + x_k W_k.$$

Положим, что число испытаний довольно большое, тогда относительная частота приблизительно равняется вероятности появления события:

$$W_1 \cong p_1, W_2 \cong p_2, \dots, W_k \cong p_k.$$

Учтя это, после преобразований получим:

$$\bar{X} \cong x_1 p_1 + x_2 p_2 + \dots + x_k p_k.$$

При этом правая часть этого приближенного равенства и есть $M(X)$, тогда

$$\bar{X} \cong M(X).$$

Это означает, что математическое ожидание приблизительно равно среднему арифметическому наблюдаемых значений случайной величины.

2. Свойства математического ожидания.

Свойство 1. Математическое ожидание постоянной величины и есть сама постоянная:

$$M(C) = C.$$

Доказательство. Пусть постоянная C является дискретной случайной величиной, имеющей одно возможное значение C и принимающей его с вероятностью $p = 1$. Таким образом,

$$M(C) = C \times 1 = C.$$

Свойство 2. Постоянный множитель допускается выносить за знак математического ожидания:

$$M(CX) = CM(X).$$

Доказательство. Допустим, случайная величина X задана законом распределения вероятностей:

$$\begin{array}{l} X \quad x_1 \quad x_2 \dots x_n; \\ P \quad p_1 \quad p_2 \dots p_n. \end{array}$$

Закон распределения случайной величины CX запишется следующим образом:

$$\begin{array}{cccc} CX & Cx_1 & Cx_2 \dots Cx_n; \\ P & p_1 & p_2 \dots p_n. \end{array}$$

При этом математическое ожидание случайной величины CX :

$$\begin{aligned} M(CX) &= Cx_1p_1 + Cx_2p_2 + \dots + Cx_np_n = \\ &= C(x_1p_1 + x_2p_2 + \dots + x_np_n) = CM(X). \end{aligned}$$

Значит, $M(CX) = CM(X)$.

Свойство 3. Математическое ожидание произведения двух независимых случайных величин равняется произведению их математических ожиданий:

$$M(XY) = M(X)M(Y).$$

Доказательство. Допустим, X и Y — это независимые случайные величины, которые заданы определенными законами распределения вероятностей:

$$\begin{array}{cccc} X & x_1, x_2 & Y & y_1, y_2; \\ P & p_1, p_2 & g & g_1, g_2. \end{array}$$

Для того чтобы составить все значения, которые может принимать случайная величина XY , необходимо перемножить все возможные значения X на каждое возможное значение Y . В результате получится x_1y_1 , x_2y_1 , x_1y_2 и x_2y_2 . Закон распределения XY запишется следующим образом:

$$\begin{array}{cccc} XY & x_1y_1 & x_2y_1 & x_1y_2 & x_2y_2; \\ P & p_1g_1 & p_2g_1 & p_1g_2 & p_2g_2. \end{array}$$

Для простоты предположили, что все возможные значения произведения различны. Однако если это не так, то доказательство проводится аналогично.

Сумма произведений всех возможных значений на их вероятности и есть математическое ожидание:

$$M(XY) = x_1y_1 \times p_1g_1 + x_2y_1 \times p_2g_1 + x_1y_2 \times p_1g_2 + x_2y_2 \times p_2g_2$$

или иначе

$$M(XY) = y_1 g_1 \times (x_1 p_1 + x_2 p_2) + y_2 g_2 (x_1 p_1 + x_2 p_2) = \\ = (x_1 p_1 + x_2 p_2)(y_1 g_1 + y_2 g_2) = M(X) \times M(Y)$$

Следовательно,

$$M(XY) = M(X)M(Y).$$

Следствие. Математическое ожидание произведения нескольких взаимно независимых случайных величин равняется произведению их математических ожиданий. К примеру, для трех случайных величин можно записать:

$$M(XYZ) = M(XY \times Z) = M(XY)M(Z) = M(X)M(Y)M(Z).$$

Свойство 4. Математическое ожидание суммы двух случайных величин есть сумма математических ожиданий слагаемых:

$$M(X + Y) = M(X) + M(Y).$$

Доказательство. Допустим X и Y — случайные величины, которые заданы законами распределения:

$$\begin{array}{cccc} X & x_1 & x_2 & Y & y_1 & y_2; \\ p & p_1 & p_2 & g & g_1 & g_2. \end{array}$$

Для того чтобы составить все возможные значения величины $X + Y$, нужно к каждому возможному значению X прибавить каждое возможное значение Y . В результате этого получится $x_1 + y_1$, $x_1 + y_2$, $x_2 + y_1$ и $x_2 + y_2$. Для простоты допустим, что эти возможные значения различны, и обозначим их вероятности соответственно p_{11} , p_{12} , p_{21} и p_{22} .

Математическое ожидание величины $X + Y$ есть сумма произведений возможных значений на их вероятности:

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

или

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

Необходимо доказать, что $p_{11} + p_{12} = p_1$. Поскольку событие, заключающееся в том, что X примет значение x_1 (вероятность его p_1), то ему следует событие, заключающееся в том, что $X + Y$ примет значение $x_1 + y_1$ или $x_1 + y_2$ (вероятность $p_{11} + p_{12}$), и обратно. Поэтому $p_{11} + p_{12}$. Таким же образом доказываются равенства:

$$p_{21} + p_{22} = p_2, p_{11} + p_{21} = g_1 p_{12} + p_{22} = g_2.$$

Подставляя правые части этих равенств в предыдущее соотношение, будем иметь:

$$M(X + Y) = (x_1 p_1 + x_2 p_2) + (y_1 g_1 + y_2 g_2)$$

или

$$M(X + Y) = M(X) + M(Y).$$

Следствие. Математическое ожидание суммы нескольких случайных величин равняется сумме математических ожиданий слагаемых, к примеру, для трех слагаемых будет:

$$\begin{aligned} M(X + Y + Z) &= M[(X + Y) + Z] = \\ &= M(X + Y) + M(X) + M(Y) + M(Z). \end{aligned}$$

3. Математическое ожидание числа появлений события в независимых испытаниях.

Допустим, осуществляется n независимых испытаний. В каждом из этих испытаний вероятность появления события A неизменна и равняется p . Для того чтобы найти среднее число появлений события A в этих испытаниях, можно воспользоваться следующей теоремой.

Теорема. Математическое ожидание $M(X)$ числа появлений события A среди n независимых испытаний равняется числу испытаний, умноженному на вероятность появления события в каждом испытании:

$$M(X) = np.$$

Доказательство. Пусть случайная величина X есть число наступления события A среди n независимых испытаний.

Общее число X появлений события A среди этих испытаний складывается из чисел появлений события в отдельных испытаниях, вследствие этого, если X_1 есть число появлений события в первом испытании, X_2 — во втором, ..., X_n — в n -м, общее число появлений события будет равно $X = X_1 + X_2 + \dots + X_n$.

Исходя из третьего свойства математического ожидания имеем:

$$M(X) = M(X_1) + M(X_2) + \dots + M(X_n).$$

Все слагаемые правой части равенства являются математическими ожиданиями числа появлений события в одном испытании, т. е. $M(X_1)$ — в первом, $M(X_2)$ — во втором и т. д. Поскольку математическое ожидание числа появлений события в одном испытании равняется вероятности события, то можно записать следующее равенство:

$$M(X_1) = M(X_2) = \dots = M(X_n) = p.$$

В этом случае окончательно получим:

$$M(X) = np.$$

3. Дисперсия дискретной случайной величины

1. Отклонение случайной величины от ее математического ожидания.

Целесообразность введения числовой характеристики рассеяния случайной величины.

Пусть даны дискретные случайные величины X и Y , которые заданы следующими законами распределения:

$$\begin{array}{cc} X & -0,01 \quad 0,01 \\ p & 0,5 \quad 0,5 \end{array} \quad \begin{array}{cc} Y & -100 \quad 100 \\ p & 0,5 \quad 0,5 \end{array} .$$

Вычислим математические ожидания этих величин:

$$M(X) = -0,01 \times 0,5 + 0,01 \times 0,5 = 0;$$

$$M(Y) = -0,01 \times 0,5 + 100 \times 0,5 = 0.$$

В этом случае математические ожидания равны, а возможные значения неодинаковы. При этом X принадлежат возможные значения, близкие к математическому ожиданию, а Y — те, которые далеки от своего математического ожидания.

Следовательно, зная только математическое ожидание случайной величины, еще нельзя сказать, какие значения она может принимать, и будут ли они близки к математическому ожиданию, т. е. математическое ожидание не может характеризовать целиком случайную величину.

В связи с этим вместе с математическим ожиданием включают и другие числовые характеристики. К примеру, для того чтобы оценить близость расположения возможных значений случайной величины к ее математическому ожиданию, можно воспользоваться числовой характеристикой — дисперсией.

Отклонение случайной величины от ее математического ожидания.

Допустим, X есть случайная величина, а $M(X)$ является ее математическим ожиданием, тогда $X - M(X)$ будет разностью между случайной величиной и ее математическим ожиданием, которое называют *отклонением*.

Имея закон распределения X :

$$\begin{array}{cccc} X & x_1 & x_2 & \dots & x_n; \\ p & p_1 & p_2 & \dots & p_n. \end{array}$$

Можно записать закон распределения отклонения. Если случайная величина приняла значение x_1 , то отклонение примет значение $x_1 - M(X)$. Вероятность этого события будет p_1 , а значит, вероятность того, что отклонение примет значение $x_1 - M(X)$, тоже будет равняться p_1 . Таким же образом дело будет обстоять и для других возможных значений отклонения.

Следовательно, отклонение обладает следующим законом распределения:

$$\begin{array}{cccccc} X - M(X) & x_1 - M(X) & x_2 - M(X) & \dots & x_n - M(X); \\ p & p_1 & p_2 & \dots & p_n. \end{array}$$

Теорема. Математическое ожидание отклонения равняется нулю:

$$M[X - M(X)] = 0.$$

Доказательство. Мы знаем, что математическое ожидание разности равняется разности математических ожиданий; математическое ожидание постоянной равняется самой постоянной. На основании этих свойств математического ожидания и знания, что $M(X)$ — постоянная величина, будем иметь:

$$\begin{aligned} M[X - M(X)] &= M(X) - M[M(X)] = \\ &= M(X) - M(X) = 0. \end{aligned}$$

2. Дисперсия дискретной случайной величины.

Нередко бывает необходимым найти рассеяние возможных значений случайной величины около ее среднего значения. К примеру, в артиллерии немаловажно знать, насколько близко друг к другу расположатся снаряды около цели, в которую нужно попасть.

Целесообразным считается в таких случаях заменить возможные отклонения их абсолютными значениями или их квадратами. Однако если возможные отклонения заменяют их абсолютными значениями, то приходится иметь дело с абсолютными величинами. В свою очередь, это может привести к серьезным затруднениям, поэтому зачастую подсчитывают среднее значение квадрата отклонения, которое и называется дисперсией.

Дисперсия (рассеяние) дискретной случайной величины — это математическое ожидание квадрата отклонения случайной величины от ее математического ожидания:

$$D(X) = M[X - M(X)]^2.$$

Допустим, дан закон распределения случайной величины, который выглядит следующим образом:

$$\begin{array}{ccccccc} X & x_1 & x_2 & \dots & x_n; \\ p & p_1 & p_2 & \dots & p_n \end{array}$$

В этом случае квадрат отклонения задан следующим законом

распределения:

$$\begin{array}{ccccccc} [X - M(X)]^2 & [x_1 - M(X)]^2 & [x_2 - M(X)]^2 & \dots & [x_n - M(X)]^2; \\ p & p_1 & p_2 & \dots & p_n. \end{array}$$

На основании определения дисперсии, можно записать:

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

Следовательно, для того чтобы найти дисперсию, хватит того, чтобы была найдена сумма произведений возможных значений квадрата отклонения на их вероятности.

Ниже приведена теорема, которой удобно пользоваться для нахождения дисперсии.

Теорема. Дисперсия равняется разности между математическим ожиданием квадрата случайной величины X и квадратом ее математического ожидания:

$$D(X) = M(X^2) - [M(X)]^2.$$

Доказательство. Математическое ожидание $M(X)$ — это постоянная величина, поэтому, $2M(X)$ и $M^2(X)$ будут также постоянными величинами. Зная это и применяя свойства математического ожидания, которые говорят, что постоянный множитель можно вынести за знак математического ожидания, а математическое ожидание суммы равно сумме математических ожиданий слагаемых, можно упростить формулу, определяющую дисперсию:

$$\begin{aligned} D(X) &= M[X - M(X)]^2 = M[X^2 - 2XM(X) + \\ &+ M^2(X)] = M(X^2) - 2M(X)M(X) + M^2(X) = \\ &= M(X^2) - 2M^2(X) + M^2(X) = M(X^2) - M^2(X). \end{aligned}$$

Значит окончательно будем иметь

$$D(X) = M(X^2) - [M(X)]^2.$$

3. Свойства дисперсии.

Свойство 1. Дисперсия постоянной величины C равняется нулю:

$$D(C) = 0.$$

Доказательство. Исходя из определения дисперсии, запишем

$$D(C) = M\{[C - M(C)]^2\}.$$

Известно, что математическое ожидание постоянной равняется самой постоянной. Тогда применяя это свойство математического ожидания, будем иметь

$$D(C) = M\{(C - C)^2\} = M(0),$$

а значит

$$D(C) = 0.$$

Это ясно, поскольку постоянная величина не меняется и не рассеивается.

Свойство 2. Чтобы вынести постоянный множитель за знак дисперсии, необходимо возвести его в квадрат:

$$D(CX) = C^2 D(X).$$

Доказательство. На основании определения дисперсии можно записать, что

$$D(CX) = M\{[CX - M(CX)]^2\}.$$

Известно, что постоянный множитель можно выносить за знак математического ожидания. На основании этого свойства математического ожидания получится:

$$\begin{aligned} D(CX) &= M\{[CX - M(CX)]^2\} = M\{C^2[X - M(X)]^2\} = \\ &= C^2 M\{[X - M(X)]^2\} = C^2 D(X), \end{aligned}$$

а значит, запишем следующее:

$$D(CX) = C^2 D(X).$$

Это очевидно, потому что при $|C| > 1$ величина CX имеет по абсолютной величине возможные значения большие, чем величина X , поэтому эти значения рассеяны вокруг математического ожидания $M(CX)$ сильнее, чем возможные значения X вокруг $M(X)$. Другими словами, $D(CX) > D(X)$, и наоборот, если $0 < |C| < 1$, то $D(CX) < D(X)$.

Свойство 3. Дисперсия суммы двух независимых случайных величин равняется сумме дисперсий этих величин, т. е.

$$D(X + Y) = D(X) + D(Y).$$

Доказательство. На основании формулы для вычисления дисперсии запишем:

$$D(X + Y) = M[(X + Y)^2] - [M(X + Y)]^2.$$

Далее нужно раскрыть скобки и воспользоваться свойствами математического ожидания суммы нескольких величин и произведения двух независимых случайных величин. Вследствие этого получится:

$$\begin{aligned} D(X + Y) &= M[X^2 + 2XY + Y^2] - [M(X) + M(Y)]^2 = \\ &= M(X^2) + 2M(X)M(Y) + M(Y^2) - M^2(X) - \\ &- 2M(X)M(Y) - M^2(Y) = \{M(X^2) - [M(X)]^2\} + \\ &+ \{M(Y^2) - [M(Y)]^2\} = D(X) + D(Y). \end{aligned}$$

А значит,

$$D(X + Y) = D(X) + D(Y).$$

Следствие 1. Дисперсия суммы нескольких взаимно зависимых случайных величин равняется сумме дисперсий этих величин. К примеру, для трех слагаемых запишется:

$$\begin{aligned} D(X + Y + Z) &= D[X + (Y + Z)] = D(X) + D(Y + Z) = D(X) + \\ &+ D(Y) + D(Z). \end{aligned}$$

При наличии произвольного числа слагаемых доказательство нужно провести методом математической индукции.

Следствие 2. Дисперсия суммы постоянной и случайной величин равняется дисперсии случайной величины:

$$D(C + X) = D(X).$$

Доказательство. Поскольку величины C и X являются независимыми, то по третьему свойству будем иметь:

$$D(C + X) = D(C) + D(X).$$

Исходя из первого свойства $D(C) = 0$, значит,

$$D(C + X) = D(X).$$

Это очевидно при условии, что величины X и $X + C$ отличаются только началом отсчета, а значит, они рассеяны вокруг своих математических ожиданий в равной мере.

Свойство 4. Дисперсия разности двух независимых случайных величин равняется сумме их дисперсий, т. е.

$$D(X - Y) = D(X) + D(Y).$$

Доказательство. На основании третьего свойства будем иметь:

$$D(X - Y) = D(X) + D(-Y).$$

Исходя из второго свойства, запишем

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

или

$$D(X - Y) = D(X) + D(Y).$$

4. Дисперсия числа появлений события в независимых испытаниях.

Допустим, осуществляется n независимых испытаний. В каждом из этих испытаний вероятность появления события A неизменна.

Для нахождения дисперсии числа появлений события в этих испытаниях следует воспользоваться следующей теоремой.

Теорема. Дисперсия числа появлений события A в n независимых испытаниях, в каждом из которых вероятность появления события неизменна, равняется числу испытаний, умноженному на вероятности появления и не появления события в одном испытании:

$$D(X) = npq.$$

Доказательство. Пусть случайная величина X есть число появлений события A в n независимых испытаниях. Следовательно, общее число появлений события в этих испытаниях равняется сумме появлений события в отдельных испытаниях:

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

Здесь X_1 есть число наступлений события в первом испытании, X_2 — во втором, ..., X_n — в n -м. Величины X_1, X_2, \dots, X_n являются взаимно независимыми, поскольку исход каждого испытания не зависит от исходов остальных. На основании этого можно воспользоваться следствием 1:

$$D(X) = D(X_1) + D(X_2) + \dots + D(X_n).$$

Подсчитаем дисперсию X_1 по формуле

$$D(X_1) = M(X_1^2) - [M(X_1)]^2.$$

Так как величина X_1 есть число появлений события A в первом испытании, то $M(X_1) = p$.

Теперь вычислим математическое ожидание величины X_1^2 , которая может принимать только два значения: 1^2 с вероятностью p и 0^2 с вероятностью q :

$$M(X_1^2) = 1^2 \times p + 0^2 \times q = p.$$

Теперь подставим найденные результаты в соотношение для $D_1(X)$, вследствие чего запишем:

$$D(X_1) = p - p^2 = p(1-p) = pq.$$

Отсюда видно, что дисперсия каждой из остальных случайных величин также равняется pq . Поэтому, заменив каждое слагаемое правой части выражения для $D(X)$ на pq , окончательно запишем:

$$D(X) = npq.$$

5. Среднее квадратическое отклонение. Среднее квадратическое отклонение суммы взаимно независимых случайных величин.

Чтобы оценить рассеяние возможных значений случайной величины около ее среднего значения, кроме дисперсии, пользуются и другими характеристиками.

Среднее квадратическое отклонение случайной величины X есть квадратный корень из дисперсии:

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

Следует отметить, что дисперсия обладает размерностью, равной квадрату размерности случайной величины. Поэтому в случае, если желательно, чтобы оценка рассеяния имела размерность случайной величины, подсчитывают не дисперсию, а среднее отклонение, т. е. если X выражается в метрах, то $\sigma(X)$ — также в метрах, а $D(X)$ — в квадратных метрах.

Среднее квадратическое отклонение суммы взаимно независимых случайных величин.

Допустим, средние квадратические отклонения нескольких взаимно независимых случайных величин являются известными. Среднее квадратическое отклонение суммы этих величин можно подсчитать благодаря следующей теореме.

Теорема. Среднее квадратическое отклонение суммы конечного числа взаимно независимых случайных величин равняется квадратному корню из суммы квадратов средних квадратических отклонений этих величин:

$$\sigma(X_1 + X_2 + \dots + X_n) = \sqrt{\sigma^2(X_1) + \sigma^2(X_2) + \dots + \sigma^2(X_n)}.$$

Доказательство. Пусть X есть сумма рассматриваемых взаимно независимых величин, тогда запишем:

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

Дисперсия суммы нескольких взаимно независимых случайных величин будет равняться сумме дисперсий слагаемых. На основании этого запишем:

$$D(X) = D(X_1) + D(X_2) + \dots + D(X_n).$$

Исходя из этого равенства можно заключить, что:

$$\sqrt{D(X)} = \sqrt{D(X_1) + D(X_2) + \dots + D(X_n)}.$$

а следовательно,

$$\sigma(X) = \sqrt{\sigma^2(X_1) + \sigma^2(X_2) + \dots + \sigma^2(X_n)}.$$

6. Одинаково распределенные взаимно независимые случайные величины

Так как по закону распределения можно подсчитать числовые характеристики случайной величины, то можно сделать вывод, что если несколько случайных величин обладают одинаковыми распределениями, то их числовые характеристики равны.

Пусть дано n взаимно независимых случайных величин X_1, X_2, \dots, X_n , имеющих равные распределения, а значит, и равные характеристики.

Среднее арифметическое рассматриваемых случайных величин равно:

$$\bar{X} = (X_1 + X_2 + \dots + X_n) / n.$$

Ниже приведены три положения, показывающие связь между числовыми характеристиками среднего арифметического \bar{X} и соответствующими характеристиками каждой отдельной величины.

Положение 1. Математическое ожидание среднего арифметического одинаково распределенных взаимно независимых случайных величин равняется математическому ожиданию a каждой из величин:

$$M(\bar{X}) = a.$$

Доказательство. Известно, что постоянный множитель можно вынести за знак математического ожидания и математическое ожидание суммы равно сумме математических ожиданий слагаемых. Применяя эти свойства математического ожидания, можно записать:

$$\begin{aligned} M(\bar{X}) &= M\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right) = \\ &= \frac{M(X_1) + M(X_2) + \dots + M(X_n)}{n}. \end{aligned}$$

Поскольку математическое ожидание каждой из величин равно a , то будет:

$$M(\bar{X}) = na / n.$$

Положение 2. Дисперсия среднего арифметического n одинаково распределенных взаимно независимых случайных величин в n раз меньше дисперсии D каждой из величин:

$$D(\bar{X}) = D / n.$$

Доказательство. Известно, что постоянный множитель можно вынести за знак дисперсии, возведя его в квадрат, а дисперсия суммы независимых величин равна сумме дисперсий слагаемых. Применяя свойства дисперсии, будем иметь:

$$\begin{aligned} D(\bar{X}) &= D\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right) = \\ &= \frac{D(X_1) + D(X_2) + \dots + D(X_n)}{n^2}. \end{aligned}$$

Так как по условию дисперсия каждой из величин равняется D , то будем иметь:

$$D(\bar{X}) = D / n^2 = D / n.$$

Положение 3. Среднее квадратическое отклонение среднего арифметического n одинаково распределенных взаимно независимых случайных величин в \sqrt{n} раз меньше среднего квадратического отклонения σ каждой из величин:

$$\sigma(\bar{X}) = \sigma / \sqrt{n} .$$

Доказательство. Поскольку $D(\bar{X}) = D/n$, то среднее квадратическое отклонение \bar{X} равняется следующему:

$$\sigma(\bar{X}) = \sqrt{D(\bar{X})} = \sqrt{D/n} = \sqrt{D} / \sqrt{n} = \sigma / \sqrt{n} .$$

Известно, что дисперсия и среднее квадратическое отклонение являются мерами рассеяния случайной величины. Поэтому среднее арифметическое достаточно большого числа взаимно независимых случайных величин имеет значительно меньшее рассеяние, чем каждая величина отдельно.

7. Начальные и центральные теоретические моменты.

Пусть дискретная случайная величина X задана следующим законом распределения:

$$\begin{array}{l} X \quad 1 \quad 2 \quad 5 \quad 100 ; \\ p \quad 0,6 \quad 0,2 \quad 0,19 \quad 0,01 . \end{array}$$

Вычислим математическое ожидание X :

$$M(X) = 1 \times 0,6 + 2 \times 0,2 + 5 \times 0,19 + 100 \times 0,01 = 2,95 .$$

Закон распределения X^2 будет выглядеть следующим образом:

$$\begin{array}{l} X^2 \quad 1 \quad 4 \quad 25 \quad 10000 ; \\ p \quad 0,6 \quad 0,2 \quad 0,19 \quad 0,01 . \end{array}$$

Подсчитаем математическое ожидание X^2 :

$$M(X^2) = 1 \times 0,6 + 4 \times 0,2 + 25 \times 0,19 + \\ + 10000 \times 0,01 = 106,15.$$

Очевидно, что $M(X^2)$ существенно больше $M(X)$, так как после возведения в квадрат возможное значение величины X^2 , соответствующее значению $x = 100$ величины X , стало равным 10 000, другими словами, оно существенно увеличилось.

Благодаря тому, что мы перешли от $M(X)$ к $M(X^2)$, есть воз-

можность лучше учесть влияние на математическое ожидание того возможного значения, которое велико и имеет малую вероятность.

Начальный момент порядка k случайной величины X — это математическое ожидание величины X^k :

$$\nu_k = M(X^k).$$

Например,

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

Применив это, формулу для вычисления дисперсии $D(X) = M(X^2) - [M(X)]^2$ запишем следующим образом:

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

Рационально рассматривать, кроме моментов случайной величины X , можно также и моменты отклонения $X - M(X)$.

Центральный момент порядка k случайной величины X — это математическое ожидание величины $(X - M(X))^k$:

$$\mu_k = M[(X - M(X))^k].$$

Например,

$$\mu_1 = M[X - M(X)] = 0;$$

$$\mu_2 = M[(X - M(X))^2] = D(X).$$

Сравнивая последние два равенства, будем иметь:

$$\mu_2 = \nu_2 - \nu_1^2.$$

Используя определение центрального момента и пользуясь свойствами математического ожидания, можно получить следующие формулы:

$$\mu_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3;$$

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

Моменты более высоких порядков используются редко.

4. Закон больших чисел

1. Неравенство Чебышева. Теорема Чебышева.

Для дискретных и непрерывных случайных величин справедливо неравенство Чебышева.

Пусть дискретная случайная величина X задана следующей таблицей распределения:

X	x_1	x_2	...	x_n
P	p_1	p_2	...	p_n

Требуется оценить вероятность того, что отклонение случайной величины от ее математического ожидания не превосходит по абсолютной величине положительного числа ε . Если ε довольно мало, то мы оценим вероятность того, что X примет значения, довольно близкие к своему математическому ожиданию. П. Л. Чебышев доказал неравенство, благодаря которому можно дать интересующую нас оценку.

Неравенство Чебышева. Вероятность того, что отклонение случайной величины X от ее математического ожидания по абсолютной величине меньше положительного числа ε , не меньше, чем $1 - D(X)/\varepsilon^2$:

$$P(|X - M(X)| < \varepsilon) \geq 1 - D(X)/\varepsilon^2.$$

Теорема Чебышева. Если X_1, X_2, \dots, X_n , — попарно независимые случайные величины, причем дисперсии их равномерно ограничены (не превышают постоянного числа C), то, как бы мало ни было положительное число ε , вероятность неравенства

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

будет как угодно близка к единице, если число случайных тин достаточно велико, т. е. в условиях теоремы

$$\lim_{n \rightarrow \infty} P_0 \left| \frac{X_1 + X_2 + \dots + X_n}{n} - \frac{M(X_1) + M(X_2) + \dots + M(X_n)}{n} \right| < \varepsilon_P = 1.$$

Теорема Чебышева гласит: если рассматривается достаточно большое число независимых случайных величин, которые имеют ограниченные дисперсии, то почти достоверным можно считать событие, заключающееся в том, что отклонение среднего арифметического случайных величин от среднего арифметического их математических ожиданий будет по абсолютной величине сколь угодно малым.

Доказательство. Случайная величина, являясь средним арифметическим случайных величин, запишется следующим образом:

$$\bar{X} = (X_1 + X_2 + \dots + X_n) / n.$$

Вычислим математическое ожидание \bar{X} . Известно, что постоянный множитель можно вынести за знак математического ожидания и математическое ожидание суммы равно сумме математических ожиданий слагаемых. Применяя эти свойства математического ожидания, получим:

$$M_0 \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{M(X_1) + M(X_2) + \dots + M(X_n)}{n}.$$

Относительно величины \bar{X} , по неравенству Чебышева, будем

ИМЕТЬ:

$$P_{\circ} \left| \frac{X_1 + X_2 + \dots + X_n}{n} - M_{\circ} \frac{X_1 + X_2 + \dots + X_n}{n} \right| < \varepsilon \geq$$

$$\geq 1 - \frac{D_{\circ} \frac{X_1 + X_2 + \dots + X_n}{n}}{\varepsilon^2}$$

или

$$P_{\circ} \left| \frac{X_1 + X_2 + \dots + X_n}{n} - \frac{M(X_1) + M(X_2) + \dots + M(X_n)}{n} \right| < \varepsilon \geq$$

$$\geq 1 - \frac{D_{\circ} \frac{X_1 + X_2 + \dots + X_n}{n}}{\varepsilon^2}.$$

Известно, что постоянный множитель можно вынести за знак дисперсии, возведя его квадрат, и дисперсия суммы независимых случайных величин равна сумме дисперсий слагаемых. Применяя эти свойства дисперсии, будем иметь:

$$D_{\circ} \frac{(X_1 + X_2 + \dots + X_n)}{n} = \frac{D(X_1) + D(X_2) + \dots + D(X_n)}{n^2}.$$

Дисперсии всех случайных величин (по условию) ограничены постоянным числом C , т. е. имеют место неравенства: $D(X_1) \leq C$, $D(X_2) \leq C$, ..., $D(X_n) \leq C$. Следовательно, можно записать, что

$$(D(X_1) + D(X_2) + \dots + D(X_n)) / n^2 \leq (C + C + \dots + C) / n^2 =$$

$$= nC / n^2 = C / n,$$

а значит,

$$D \left(\frac{X_1 + X_2 + \dots + X_n}{n} \right) \leq \frac{C}{n}.$$

Подставляя правую часть последнего неравенства в предпоследнее (при этом последнее может быть лишь усилено), будем иметь:

$$P_0 \left| \frac{X_1 + X_2 + \dots + X_n}{n} - \frac{M(X_1) + M(X_2) + \dots + M(X_n)}{n} \right| < \varepsilon \geq \geq 1 - \frac{C}{n\varepsilon^2}.$$

На основании последнего, переходя к пределу, будет записано следующее:

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{X_1 + X_2 + \dots + X_n}{n} - \frac{M(X_1) + M(X_2) + \dots + M(X_n)}{n} \right| < \varepsilon \right) \geq 1.$$

Следует отметить, что вероятность не может превышать единицу, тогда

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

Что и требовалось доказать.

Обозначим через a математическое ожидание каждой из случайных величин. Здесь среднее арифметическое математических ожиданий также равно a . Если X_1, X_2, \dots, X_n , являются попарно независимыми случайными величинами, которые имеют одно и то же математическое ожидание a , а также если дисперсии этих величин равномерно ограничены, то, как бы мало ни было число $\varepsilon > 0$, вероятность неравенства

$$\left| \frac{X_1 + X_2 + \dots + X_n}{n} - a \right| < \varepsilon$$

будет как угодно близка к единице, если число случайных величин достаточно велико.

Это есть формулировка теоремы Чебышева для частного случая.

Таким образом, в условиях теоремы будет иметь место ра-

венство:

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

Сущность теоремы Чебышева заключается в следующем: хотя отдельные независимые случайные величины могут принимать значения, далекие от своих математических ожиданий, среднее арифметическое достаточно большого числа случайных величин с большой вероятностью принимает значения, близкие к определенному постоянному числу, а именно к числу $(M(X_1) + M(X_2) + \dots + M(X_n)) / n$ (или к числу a в частном случае). Другими словами, отдельные случайные величины могут иметь существенный разброс, а их среднее арифметическое незначительно рассеяно.

Следовательно, нельзя быть уверенным в том, какое возможное значение примет каждая из случайных величин, однако можно предвидеть, какое значение примет их среднее арифметическое.

Значит, среднее арифметическое достаточно большого числа независимых случайных величин, дисперсии которых равномерно ограничены, теряет характер случайной величины. Это связано с тем, что отклонения каждой из величин от своих математических ожиданий могут быть как положительными, так и отрицательными. В среднем арифметическом же они будут взаимно погашаться.

Теорема Чебышева справедлива как для дискретных, так и для непрерывных случайных величин.

2. Теорема Бернулли.

Допустим, осуществляется n независимых испытаний. В каждом из этих испытаний вероятность появления события A равняется p . Чтобы выяснить возможность предугадывать, какова приблизительно будет относительная частота появлений события, следует воспользоваться теоремой Бернулли.

Теорема Бернулли. Если в каждом из n независимых испытаний вероятность p появления события A постоянна, то как угодно близка к единице вероятность того, что отклонение относительной частоты от вероятности p по абсолютной величине будет сколь угодно малым, если число испытаний достаточно велико.

Иначе говоря, если ε — сколь угодно малое положительное число, то при соблюдении условий теоремы можно записать следующее неравенство:

$$\lim_{n \rightarrow \infty} P(|m/n - p| < \varepsilon) = 1.$$

Доказательство теоремы вытекает из независимости испытаний Бернулли и теоремы Чебышева.

5. Функция распределения вероятностей случайной величины

1. Определение функции распределения.

Известно, что дискретная случайная величина может быть задана набором всех ее возможных значений и их вероятностей. Данный способ задания не является общим. Он неприменим, например, для непрерывных случайных величин.

Пусть случайная величина X имеет возможные значения, которые сплошь заполняют интервал (a, b) . В этом случае нельзя составить перечень всех возможных значений X . Поэтому рационально дать общий способ задания любых типов случайных величин.

Для этого вводят функции распределения вероятностей случайной величины.

Если x — действительное число, то вероятность события, заключающегося в том, что X примет значение, меньшее x . Другими словами вероятность события $X < x$ следует обозначить через $F(x)$. Поскольку x изменяется, то изменяется $F(x)$. Следовательно, $F(x)$ есть функция от x .

Функция распределения — это функция $F(x)$, которая определяет вероятность того, что случайная величина X в результате испытания примет значение, меньшее x :

$$F(x) = P(X < x).$$

Геометрически это равенство можно объяснить следующим образом: $F(x)$ есть вероятность того, что случайная величина примет значение, которое изображается на числовой оси точкой, лежащей левее точки x .

По-другому функцию распределения называют интегральной функцией.

Случайная величина называется непрерывной, если ее функция распределения есть непрерывная, кусочно-дифференцируемая функция с непрерывной производной.

2. Свойства функции распределения.

Свойство 1. Значения функции распределения принадлежат отрезку $[0, 1]$:

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

Доказательство. Свойство происходит из определения функции распределения как вероятности, т. е. вероятность всегда есть неотрицательное число, не превышающее единицы.

Свойство 2. $F(x)$ — неубывающая функция:

$$F(x_2) \geq F(x_1),$$

если $x_2 > x_1$.

Доказательство. Допустим $x_2 > x_1$, тогда событие, состоящее в том, что X примет значение, меньшее x_2 , можно разделить на следующие два несовместных события:

- 1) X примет значение, меньшее x_1 , с вероятностью $P(X < x_1)$;
- 2) X примет значение, удовлетворяющее неравенству $x_1 \leq X < x_2$ с вероятностью $P(x_1 \leq X < x_2)$.

Тогда на основании теоремы сложения имеем

$$P(X < x_2) = P(X < x_1) + P(x_1 \leq X < x_2),$$

а значит,

$$P(X < x_2) - P(X < x_1) = P(x_1 \leq X < x_2)$$

или

$$F(x_2) - F(x_1) = P(x_1 \leq X < x_2).$$

Так как любая вероятность является положительным числом, то $F(x_2) - F(x_1) \geq 0$ или $F(x_2) \geq F(x_1)$, что и требовалось доказать.

Следствие 1. Вероятность того, что случайная величина примет значение, заключенное в интервале (a, b) , равняется приращению функции распределения на этом интервале:

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

Данное следствие основано на том, что

$$F(x_2) - F(x_1) = P(x_1 \leq X < x_2),$$

если положить, что $x_2 = b$ и $x_1 = a$.

Следствие 2. Вероятность того, что непрерывная случайная величина X примет одно определенное значение, равна нулю.

В самом деле, положив в формуле

$$P(a \leq X < b) = F(b) - F(a) \quad a = x_1, \quad b = x_1 + \Delta x,$$

будем иметь следующее неравенство:

$$P(x_1 \leq X < x_1 + \Delta x) = F(x_1 + \Delta x) - F(x_1).$$

Если Δx устремить к нулю, то функция $F(x)$ — непрерывна, поскольку X есть непрерывная случайная величина. Поэтому в точке x_1 разность $F(x_1 + \Delta x) - F(x_1)$ также стремится к нулю. Это означает, что $P(X = x_1) = 0$. На основании этого не трудно убедиться в справедливости равенств:

$$P(a \leq X < b) = P(a < X < b) = P(a < X \leq b) = P(a \leq X \leq b).$$

К примеру, равенство $P(a < X \leq b) = P(a < X < b)$ доказывается следующим образом:

$$P(a < X \leq b) = P(a < X < b) = P(X = b) = P(a < X < b).$$

Заметим, что в результате испытания случайная величина неизбежно примет одно из возможных значений.

Свойство 3. Если возможные значения случайной величины принадлежат интервалу (a, b) , то:

- 1) $F(x) = 0$ при $x \leq a$;
- 2) $F(x) = 1$ при $x \geq b$.

Доказательство.

1. Допустим $x_1 \leq a$, тогда событие $X < x_1$ невозможно, поскольку по условию величина X не принимает значений, меньших x_1 . Значит, его вероятность равна нулю.

2. Допустим $x_2 \geq b$, тогда событие $X < x_2$ достоверно поскольку все возможные значения X меньше x_2 . Значит, его вероятность равна единице.

Следствие. Если возможные значения непрерывной случайной величины размещены на всей оси x , то достоверны следующие предельные соотношения:

$$\lim_{x \rightarrow -\infty} F(x) = 0, \quad \lim_{x \rightarrow \infty} F(x) = 1.$$

Благодаря свойствам функции распределения можно представить, как выглядит график функции распределения непрерывной случайной величины.

По первому свойству: график находится в полосе, которая ограничена прямыми $y = 0$, $y = 1$.

По второму свойству: при возрастании x в интервале (a, b) , в котором заключены все возможные значения случайной величины, график «поднимается вверх».

По третьему свойству: при $x \leq a$ ординаты графика равны нулю; при $x \geq b$ ординаты графика равны единице.

Следует заметить, что график функции распределения непрерывной случайной величины имеет ступенчатый вид.

6. Плотность распределения вероятностей непрерывной случайной величины

1. Определение плотности распределения. Вероятность попадания непрерывной.

Определение плотности распределения.

Способ задания непрерывной случайной величины с помощью функции распределения не является единственным. Также непрерывную случайную величину можно задать, используя другую функцию — плотность распределения или плотность вероятности. По-другому ее называют дифференциальной функцией.

Плотность распределения вероятностей непрерывной случайной величины X есть функция $f(x)$ — первая производная от функции распределения $F(x)$:

$$f(x) = F'(x).$$

Следовательно, функция распределения является первообразной для плотности распределения. Однако для описания распределения вероятностей дискретной случайной величины плотность распределения нельзя применять.

Вероятность попадания непрерывной случайной величины в заданный интервал.

Если известна плотность распределения, то можно подсчитать вероятность того, что непрерывная случайная величина примет значение, которое принадлежит заданному интервалу. Этот подсчет основан на следующей теореме.

Теорема. Вероятность того, что непрерывная случайная величина X примет значение, принадлежащее интервалу (a, b) , равняется определенному интегралу от плотности распределения, взятому в пределах от a до b :

$$P(a < X < b) = \int_a^b f(x)dx.$$

Доказательство. Известно, что

$$P(a \leq X < b) = F(b) - F(a),$$

тогда по формуле Ньютона — Лейбница,

$$F(b) - F(a) = \int_a^b F'(x)dx = \int_a^b f(x)dx.$$

Следовательно,

$$P(a \leq X < b) = \int_a^b f(x)dx.$$

Поскольку $P(a \leq X < b) = P(a < X < b)$, то окончательно запишем:

$$P(a < X < b) = \int_a^b f(x)dx.$$

Геометрически полученный результат можно объяснить следующим образом: вероятность того, что непрерывная случайная величина примет значение, которое принадлежит интервалу (a, b) ,

равняется площади криволинейной трапеции, ограниченной осью Ox , кривой распределения $f(x)$ и прямыми $x = a$ и $x = b$.

2. Нахождение функции распределения по известной плотности распределения.

Если известна плотность распределения $f(x)$, то можно найти функцию распределения $F(x)$ по следующей формуле:

$$F(x) = \int_{-\infty}^x f(x) dx.$$

В самом деле, $F(x)$ есть вероятность того, что случайная величина примет значение, меньшее x , т. е.

$$F(x) = P(X < x).$$

Ясно, что неравенство $X < x$ можно написать в виде двойного неравенства $-\infty < X < x$, а значит,

$$F(x) = P(-\infty < X < x).$$

Если $a = -\infty$, $b = x$ в формуле $P(a < X < b) = \int_a^b f(x) dx$, то получится

$$P(-\infty < X < x) = \int_{-\infty}^x f(x) dx.$$

Поскольку $F(x) = P(-\infty < X < x)$, то заменив $P(-\infty < X < x)$ на $F(x)$, окончательно получим:

$$F(x) = \int_{-\infty}^x f(x) dx.$$

Следовательно, если известна плотность распределения, то можно отыскать функцию распределения. Соответственно по известной функции распределения может быть найдена плотность распределения:

$$F(x) = F'(x).$$

3. Свойства плотности распределения.

Свойство 1. Плотность распределения является неотрицательной функцией:

$$f(x) \geq 0.$$

Доказательство. Так как функция распределения есть неубывающая функция, значит, ее производная $F'(x) = f(x)$ есть функция положительная.

Геометрически это свойство означает, что точки, которые принадлежат графику плотности распределения, расположены либо над осью Ox , либо на этой оси.

Иначе график плотности распределения называется кривой распределения.

Свойство 2. Несобственный интеграл от плотности распределения в пределах от $-\infty$ до ∞ равен единице:

$$\int_{-\infty}^{\infty} f(x)dx = 1.$$

Доказательство. Несобственный интеграл

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

показывает вероятность события, заключающегося в том, что случайная величина будет иметь значение, которое принадлежит интервалу $(-\infty, \infty)$. Соответственно, такое событие является достоверным, а значит, вероятность его равна единице.

Из геометрических соображений можно сказать, что вся площадь криволинейной трапеции, ограниченной осью Ox и кривой распределения, равна единице.

Например, если все возможные значения случайной величины относятся к интервалу (a, b) , то

$$\int_a^b f(x)dx = 1.$$

4. Вероятностный смысл плотности распределения.

Обозначим через $F(x)$ функцию распределения непрерывной случайной величины X . На основании определения плотности распределения $f(x) = F'(x)$ или

$$f(x) = \lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x}.$$

Разность $F(x + \Delta x) - F(x)$ обуславливает вероятность того, что X будет иметь значение, которое принадлежит интервалу $(x, x + \Delta x)$. Следовательно, предел отношения вероятности того, что непрерывная случайная величина примет значение, принадлежащее интервалу $(x, x + \Delta x)$ к длине этого интервала (при $\Delta x \rightarrow 0$) равен значению плотности распределения в точке x .

Рационально рассматривать значение функции $f(x)$ в точке x как плотность вероятности в этой точке.

Таким образом, функция $f(x)$ находит плотность распределения вероятности для каждой точки x .

Мы знаем, что приращение функции приблизительно равно дифференциалу функции:

$$F(x + \Delta x) - F(x) \cong dF(x)$$

или

$$F(x + \Delta x) - F(x) \cong F'(x)dx.$$

Поскольку $F'(x) = f(x)$ и $dx = \Delta x$, то

$$F(x + \Delta x) - F(x) \cong f(x)\Delta x.$$

Запишем вероятностный смысл этого равенства: вероятность того, что случайная величина примет значение, которое принадлежит интервалу

$$(x, x + \Delta x),$$

приблизительно равна произведению плотности вероятности в точке x на длину интервала Δx .

Из геометрических соображений можно сказать: вероятность того, что случайная величина будет иметь значение, которое принадлежит интервалу

$$(x, x + \Delta x),$$

примерно равна площади прямоугольника с основанием Δx и высотой $f(x)$.

Тогда площадь прямоугольника равная произведению $f(x)\Delta x$, только приблизительно равняется площади криволинейной трапеции — истинной вероятности, определяемой определенным

интегралом

$$\int_x^{x+\Delta x} f(x)dx .$$

5. Закон равномерного распределения вероятностей.

На практике случается сталкиваться с различными распределениями непрерывных случайных величин. Плотности распределений непрерывных случайных величин по-другому называют *законами распределений*. Наиболее часто встречающиеся — это законы равномерного, нормального и показательного распределений.

Равномерным называется такое распределение вероятностей, если на интервале, которому принадлежат все возможные значения случайной величины, плотность распределения сохраняет постоянное значение.

Замечание. Если R — это непрерывная случайная величина, которая распределена равномерно в интервале $(0, 1)$, а r — ее возможные значения, то вероятность попадания величины R (в результате испытания) в интервал (c, d) , принадлежащий интервалу $(0, 1)$, равна его длине:

$$P(c < R < d) = d - c.$$

В самом деле, плотность рассматриваемого равномерного распределения равна:

$$f(r) = 1 / (1 - 0) = 1,$$

а значит, вероятность попадания случайной величины R в интервал (c, d) будет:

$$P(c < R < d) = \int_c^d f(r)dr = \int_c^d 1 \times dr = d - c.$$

7. Нормальное распределение

1. Числовые характеристики непрерывных случайных величин.

Пусть дана непрерывная случайная величина X , которая задана плотностью распределения $f(x)$. Пусть все возможные значения X относятся к отрезку $[a, b]$. Разобьем этот отрезок на n частичных отрезков длиной $\Delta x_1, \Delta x_2, \dots, \Delta x_n$, а затем выберем в каждом из них произвольную точку x_i ($i = 1, 2, \dots, n$).

Для нахождения математического ожидания непрерывной величины необходимо составить сумму произведений возможных значений x_i на вероятности попадания их в интервал Δx_i :

$$\sum x_i f(x_i) \Delta x_i.$$

Следует помнить, что произведение $f(x)\Delta x$ приблизительно равняется вероятности попадания X в интервал Δx .

Затем переходят к пределу при стремлении к нулю длины наибольшего из частичных отрезков, в результате чего получают определенный интеграл:

$$\int_a^b x f(x) dx.$$

Определенный интеграл

$$M(X) = \int_a^b x f(x) dx$$

называется математическим ожиданием непрерывной случайной величины X , у которой возможные значения принадлежат отрезку $[a, b]$.

Если возможные значения принадлежат всей оси Ox , то будет иметь место следующее выражение:

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

Пусть несобственный интеграл сходится абсолютно, другими словами, существует интеграл

$$\int_{-\infty}^{\infty} |x| f(x) dx.$$

В случае невыполнения этого требования значение интеграла зависело бы от скорости стремления нижнего предела к $-\infty$, а верхнего — к $+\infty$.

Дисперсия непрерывной величины определяется аналогично

дисперсии дискретной величины.

Дисперсия непрерывной случайной величины — это математическое ожидание квадрата ее отклонения.

В том случае если возможные значения X принадлежат отрезку $[a, b]$, то

$$D(X) = \int_a^b [x - M(X)]^2 f(x) dx,$$

а если возможные значения принадлежат всей оси x , то

$$D(X) = \int_{-\infty}^{\infty} [x - M(X)]^2 f(x) dx.$$

Среднее квадратическое отклонение непрерывной случайной величины можно найти, применяя равенство

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

Замечание 1. Свойства математического ожидания и дисперсии дискретных величин такие, как и для непрерывных величин.

Замечание 2. Для вычисления дисперсии существуют более удобные формулы:

$$D(X) = \int_a^b x^2 f(x) dx - [M(X)]^2;$$

$$D(X) = \int_{-\infty}^{\infty} x^2 f(x) dx - [M(X)]^2.$$

2. Нормальное распределение.

Распределение вероятностей непрерывной случайной величины, которое описывается плотностью, называется нормальным:

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

Нормальное распределение определяется двумя параметрами: a и σ , зная которые, можно задать нормальное распределение. Вероятностный смысл этих параметров таков: a — математическое ожидание, σ — среднее квадратическое отклонение нормального распределения S .

1. На основании определения математического ожидания непрерывной случайной величины, можно записать:

$$M(x) = \int_{-\infty}^{\infty} xf(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} xe^{-(x-a)^2/2\sigma^2} dx.$$

Если $z = (x-a)/\sigma$, то $dx = \sigma dz$, причем новые пределы интегрирования равны старым. В этом случае получится:

$$\begin{aligned} M(x) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma z + a) e^{-z^2/2} dz = \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \sigma z e^{-z^2/2} dz + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} dz. \end{aligned}$$

Так как под знаком интеграла стоит нечетная функция, то первое из слагаемых равно нулю. При этом пределы интегрирования симметричны относительно начала координат. Второе из слагаемых равно a — интеграл Пуассона:

$$\int_{-\infty}^{\infty} e^{-z^2/2} dz = \sqrt{2\pi}.$$

Математическое ожидание нормального распределения равно параметру a :

$$M(X) = a.$$

2. На основании определения дисперсии непрерывной случайной величины, а также зная, что $M(X) = a$, получится:

$$D(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-a)^2 e^{-(x-a)^2/2\sigma^2} dx.$$

Если $z = (x - a) / \sigma$, то $x - a = \sigma z$, $dx = \sigma dz$, причем новые пределы интегрирования равны старым.

Тогда можно записать

$$D(x) = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z \times z e^{-z^2/2} dx.$$

Пусть $u = z$, $dv = z e^{-z^2/2} dz$, тогда интегрируя по частям, получим:

$$D(x) = \sigma^2.$$

Таким образом,

$$\sigma(X) = \sqrt{D(X)} = \sqrt{\sigma^2} = \sigma,$$

т. е. среднее квадратическое отклонение нормального распределения равно параметру σ .

Замечание 1. Нормальное распределение с произвольными параметрами a и σ ($\sigma > 0$) называют *общим*.

Нормальное распределение с параметрами $a = 0$ и $\sigma = 1$ называют *нормированным*. Так, например, если X является нормальной величиной с параметрами a и σ , то $U = (X - a) / \sigma$ есть нормированная нормальная величина. При этом $M(U) = 0$, $\sigma(U) = 1$.

Плотность нормированного распределения запишется следующим образом:

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

Замечание 2. Функцию $F(x)$ общего нормального распределения можно записать так:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-(z-a)^2/(2\sigma^2)} dz,$$

а функцию нормированного распределения, которая табулирована, так:

$$F_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{(-z)^2/2} dz,$$

причем

$$F(x) = F_0((x-a)/\sigma).$$

Замечание 3. Применяя функцию Лапласа

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{(-z)^2/2} dz,$$

можно вычислить вероятность попадания нормированной нормальной величины X в интервал $(0, x)$. Следовательно,

$$P(0 < X < x) = \int_0^x \varphi(x) dx = \phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{(-z)^2/2} dz = \phi(x).$$

Замечание 4. Известно, что

$$\int_{-\infty}^{\infty} \varphi(x) dx = 1 \quad \text{— во } \varphi(x),$$

стало быть, в силу симметрии $\varphi(x)$ относительно нуля

$$\int_{-\infty}^0 \varphi(x) dx = 0,5,$$

а значит, и $P(-\infty < X < 0) = 0,5$, можно получить:

$$F_0(x) = 0,5 + \phi(x).$$

В самом деле,

$$\begin{aligned} F_0(x) &= P(-\infty < X < x) = \\ &= P(-\infty < X < 0) + P(0 < X < x) = 0,5 + \phi(x). \end{aligned}$$

3. Нормальная кривая.

Нормальной кривой (кривой Гаусса), называется график плотности нормального распределения. Предлагается методами дифференциального исчисления исследовать функцию

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-a)^2/(2\sigma^2)}.$$

1. Данная функция определена на всей оси x .
2. Функция принимает положительные значения при всех значениях x , т. е. нормальная кривая расположена над осью Ox .
3. Предел функции при неограниченном возрастании $|x|$ равен нулю:

$$\lim_{|x| \rightarrow \infty} y = 0,$$

т. е. ось Ox является горизонтальной асимптотой графика.

4. Для исследования функции на экстремум, нужно отыскать первую производную:

$$y' = -\frac{x-a}{\sigma^3\sqrt{2\pi}} e^{-(x-a)^2/(2\sigma^2)}.$$

Очевидно, что $y' = 0$ при $x = a$, $y' > 0$ при $x < a$, $y' < 0$ при $x > a$.

Таким образом, при $x = a$ функция имеет максимум, который равен $1/(\sigma\sqrt{2\pi})$.

5. Разность $x-a$ находится в аналитическом выражении функции в квадрате. Другими словами график функции симметричен относительно прямой $x = a$.

6. Для исследования функции на точки перегиба следует найти вторую производную:

$$y'' = -\frac{1}{\sigma^3\sqrt{2\pi}} e^{-(x-a)^2/(2\sigma^2)} \left[1 - \frac{(x-a)^2}{\sigma^2} \right].$$

Очевидно, что при $x = a + \sigma$ и $x = a - \sigma$ вторая производная равна нулю, а при переходе через эти точки она меняет знак.

При этом в обеих этих точках значение функции равно. Следовательно, точки графика $(a - \sigma, 1/(\sigma\sqrt{2\pi}e))$ и $(a + \sigma, 1/(\sigma\sqrt{2\pi}e))$ представляют собой точками перегиба.

4. Влияние параметров нормального распределения на форму нормальной кривой.

Пусть необходимо выяснить, как влияют на форму и расположение нормальной кривой значения параметров a и σ . Поскольку графики функций $f(x)$ и $f(x-a)$ обладают одинаковой формой, то, сдвинув график $f(x)$ в положительном направлении оси x на a единиц масштаба при $a > 0$ или в отрицательном направлении при $a < 0$, будем иметь график $f(x-a)$. Значит, изменение величины параметра a (математического ожидания) не изменяет формы нормальной кривой, а только приводит к ее сдвигу вдоль оси Ox вправо, если a возрастает, и влево, если a убывает.

И наоборот, если изменяется параметр σ (среднее квадратическое отклонение), причем максимум дифференциальной функции нормального распределения равен $1/(\sigma\sqrt{2\pi})$. Поэтому с возрастанием σ максимальная ордината нормальной кривой убывает, а сама кривая делается более пологой, т. е. сжимается к оси Ox . Если σ убывает, то нормальная кривая имеет более острую вершину и растягивается в положительном направлении оси Oy .

Следует отметить, что при любых значениях параметров a и σ площадь, которая ограничена нормальной кривой и осью x , по-прежнему будет равна единице.

Нормированной называют нормальную кривую

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

при $a = 0$ и $\sigma = 1$.

5. Вероятность попадания в заданный интервал нормальной случайной величины.

Поскольку случайная величина X задана плотностью распределения $f(x)$, то вероятность того, что X примет значение, которое принадлежит интервалу (α, β) , следующая:

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} f(x) dx.$$

Если случайная величина X распределена по нормальному закону, то вероятность того, что X примет значение, принадлежащее интервалу (α, β) , равняется:

$$P(\alpha < X < \beta) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-(x-a)^2/(2\sigma^2)} dx.$$

Здесь необходимо преобразование этой формулы таким образом, чтобы можно было пользоваться готовыми таблицами. Если $z = (x-a) / \sigma$, то $x = \sigma z + a$, $dx = \sigma dz$. Следует найти новые пределы интегрирования. Тогда если $x = \alpha$, то $z = (\alpha - a) / \sigma$; если $x = \beta$, то $z = (\beta - a) / \sigma$.

Таким образом, имеем:

$$\begin{aligned} P(\alpha < X < \beta) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{(\alpha-a)/\sigma}^{(\beta-a)/\sigma} e^{-z^2/2} (\sigma dz) = \\ &= \frac{1}{\sqrt{2\pi}} \int_{(\alpha-a)/\sigma}^0 e^{-z^2/2} dz + \frac{1}{\sqrt{2\pi}} \int_0^{(\beta-a)/\sigma} e^{-z^2/2} dz = \\ &= \frac{1}{\sqrt{2\pi}} \int_0^{(\beta-a)/\sigma} e^{-z^2/2} dz - \frac{1}{\sqrt{2\pi}} \int_0^{(\alpha-a)/\sigma} e^{-z^2/2} dz. \end{aligned}$$

Применяя функцию Лапласа

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz,$$

получим:

$$P(\alpha < X < \beta) = \phi\left(\frac{\beta-a}{\sigma}\right) - \phi\left(\frac{\alpha-a}{\sigma}\right).$$

6. Вычисление вероятности заданного отклонения.

Нередко нужно найти вероятность того, что отклонение нормально распределенной случайной величины X по абсолютной величине меньше заданного положительного числа δ .

Другими словами, необходимо найти вероятность осуществления неравенства $|X-a| < \delta$.

Для этого следует заменить это неравенство равносильным ему двойным неравенством:

$$-\delta < X - a < \delta \text{ или } a - \delta < X < \delta + a.$$

Используя формулу

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

получим:

$$\begin{aligned} P(|X - a| < \delta) &= P(a - \delta < X < \delta + a) = \\ &= \Phi\left[\frac{(a + \delta) - a}{\sigma}\right] - \Phi\left[\frac{(a - \delta) - a}{\sigma}\right] = \Phi\left(\frac{\delta}{\sigma}\right) - \Phi\left(\frac{-\delta}{\sigma}\right). \end{aligned}$$

Поскольку

$$\Phi(-\delta/\sigma) = -\Phi(\delta/\sigma),$$

а следовательно, функция Лапласа — нечетная, окончательно получим:

$$P(|X - a| < \delta) = 2\Phi(\delta/\sigma).$$

В частности, при $a = 0$ будет

$$P(|X| < \delta) = 2\Phi(\delta/\sigma).$$

Если две случайные величины нормально распределены и $a = 0$, то вероятность принять значение, которое принадлежит интервалу $(-\delta, \delta)$, больше у той величины, которая имеет меньшее значение σ . Этот факт полностью соответствует вероятностному смыслу параметра σ , которое является средним квадратическим отклонением. Оно характеризует рассеяние случайной величины вокруг ее математического ожидания.

Замечание. События, заключающиеся в осуществлении неравенств $|X - a| < \delta$ и $|X - a| \geq \delta$, являются противоположными.

В связи с этим если вероятность осуществления неравенства равна p , то вероятность неравенства $|X - a| < \delta$ равна $1 - p$.

7. Правило трех сигм.

Преобразуем формулу $P(|X - a| < \delta) = 2\phi(\delta/\sigma)$, допустив $\delta = \sigma t$. Тогда получится:

$$P(|X - a| < \sigma t) = 1 - \phi(t).$$

Пусть $t = 3$, значит, $\sigma t = 3\sigma$. В этом случае

$$P(|x - a| < 3\sigma) = 2\phi(3) = 2 \times 0,49865 = 0,9973.$$

Другими словами, вероятность того, что отклонение по абсолютной величине будет меньше утроенного среднего квадратического отклонения, равна 0,9973.

Вероятность того, что абсолютная величина отклонения превысит утроенное среднее квадратическое отклонение, крайне незначительна, а именно равна 0,0027. Это значит, что только в 0,27 % случаев это может произойти. На основании принципа невозможности маловероятных событий такого рода события можно считать практически невозможными. Отсюда следует правило трех сигм: если случайная величина распределена нормально, то абсолютная величина ее отклонения от математического ожидания не превышает утроенного среднего квадратического отклонения.

Практически правило трех сигм применяют следующим образом: если неизвестно распределение изучаемой случайной величины, но условие, указанное в приведенном правиле, выполняется, то можно полагать, что изучаемая величина распределена нормально. В противном же случае она не распределена нормально.

8. Понятие о теореме Ляпунова. Формулировка центральной предельной теоремы.

Центральная предельная теорема (теорема Ляпунова): если случайная величина X представляет собой сумму очень большого числа взаимно независимых случайных величин, влияние каждой из

которых на всю сумму ничтожно мало, то X имеет распределение, близкое к нормальному.

Сформулируем центральную предельную теорему, устанавливая условия, при которых сумма большого числа независимых слагаемых имеет распределение, близкое к нормальному.

Пусть X_1, X_2, \dots, X_n есть последовательность независимых случайных величин, каждая из которых имеет конечные математическое ожидание и дисперсию:

$$M(X_k) = a_k, D(X_k) = b_k^2.$$

Обозначим

$$S_n = X_1 + X_2 + \dots + X_n, \quad A_n = \sum_{k=1}^n a_k, \quad B_n^2 = \sum_{k=1}^n b_k^2,$$

а функция распределения нормированной суммы будет равна:

$$F_n(x) = P\left(\frac{S_n - A_n}{B_n} < x\right).$$

К последовательности X_1, X_2, \dots применима центральная предельная теорема в том случае, если при любом x функция распределения нормированной суммы при $n \rightarrow \infty$ стремится к нормальной функции распределения:

$$\lim_{n \rightarrow \infty} P\left[\frac{S_n - A_n}{B_n} < x\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-z^2/2} dz.$$

Так, например, если все случайные величины X_1, X_2, \dots одинаково распределены, то к этой последовательности применима центральная предельная теорема, если дисперсии всех величин $X_i (i = 1, 2, \dots)$ конечны и отличны от нуля. Если для $\delta > 0$ при $\delta \rightarrow \infty$ отношение Ляпунова

$$L_n = C_n / B_n^{2+\delta},$$

где

$$C_n = \sum_{k=1}^n M|X_k - a_k|^{2+\delta},$$

стремится к нулю (условие Ляпунова), то в этом случае к последовательности X_1, X_2, \dots применима центральная предельная теорема.

Сущность условия Ляпунова: необходимо чтобы каждое слагаемое суммы $(S_n - A_n)/B_n$ оказывало на сумму ничтожное влияние.

Замечание. Характеристическая функция случайной величины X — это функция $\varphi(t) = M[e^{itX}]$.

Для дискретной случайной величины X с возможными значениями x_k и их вероятностями p_k характеристическая функция будет выглядеть следующим образом:

$$\varphi(t) = \sum_k e^{itx_k} p_k.$$

Для непрерывной случайной величины X с плотностью распределения $f(x)$ характеристическая функция будет выглядеть следующим образом:

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

Характеристическая функция суммы независимых случайных величин равна произведению характеристических функций слагаемых.

9. Оценка отклонения теоретического распределения от нормального. Асимметрия и эксцесс

Распределение относительных частот называют *эмпирическим*. Его изучением занимается математическая статистика.

Распределение вероятностей называют *теоретическим*. Его изучением занимается теория вероятностей.

При изучении распределений, отличных от нормального, требуется количественно оценить это различие. Для этого вводят специальные характеристики, в частности асимметрию и эксцесс. Эти характеристики для нормального распределения равны нулю. В том случае если для изучаемого распределения асимметрия и эксцесс обладают небольшими значениями, то можно предположить близость этого распределения к нормальному.

В противном случае, большие значения асимметрии и эксцесса указывают на значительное отклонение от нормального.

Для симметричного распределения график такого распределения симметричен относительно прямой $x = M(X)$, причем каждый центральный момент нечетного порядка равен нулю. Центральные моменты нечетного порядка для несимметричных распределений не равны нулю. Вследствие этого любой из этих моментов (кроме момента первого порядка, который равен нулю для любого распределения) может предназначаться для оценки асимметрии. Лучше всего выбрать простейший из них, т. е. момент третьего порядка μ_3 . Все же принять этот момент для оценки асимметрии неудобно, поскольку его величина зависит от единиц, в которых измеряется случайная величина. Чтобы устранить этот недостаток, μ_3 делят на σ^3 и, таким образом, получают безразмерную характеристику.

Отношение центрального момента третьего порядка к кубу среднего квадратического отклонения называют асимметрией теоретического распределения, т. е.

$$A_s = \mu_3 / \sigma^3.$$

Когда «длинная часть» кривой распределения размещена справа от математического ожидания, говорят, что асимметрия положительна. Когда «длинная часть» кривой находится слева от математического ожидания, то имеет место отрицательная асимметрия. Точка максимума дифференциальной функции называется модой. Относительно нее по расположению кривой определяют знак асимметрии: если «длинная часть» кривой расположена правее моды, то асимметрия положительна, если слева — отрицательна.

Эксцесс применяют в случае, когда нужно оценить «крутость», т. е. больший или меньший подъем кривой теоретического распределения по сравнению с нормальной кривой.

Характеристика, которая определяется равенством

$$E_k = (\mu_4 / \sigma^4) - 3,$$

называют эксцессом теоретического распределения.

Для нормального распределения $\mu_4 / \sigma^4 = 3$, а это означает, что эксцесс равен нулю. Следовательно, если эксцесс некоторого распределения не равен нулю, то кривая этого распределения другая, чем нормальная кривая. В случае, если эксцесс положительный, кривой характерна более высокая и острая вершина, чем нормальная кривая. Когда эксцесс отрицательный, то сравниваемая кривая обладает более низкой и плоской вершиной, чем нормальная кривая. В таких случаях предполагается, что нормальное и теоретическое распределения обладают одинаковыми математическими ожиданиями и дисперсиями.

10. Функция одного случайного аргумента и ее распределение.

Если каждому возможному значению случайной величины X соответствует одно возможное значение случайной величины Y , то Y носит название функции случайного аргумента X , т. е.

$$Y = \varphi(X).$$

Пути нахождения распределения функции по известному распределению дискретного и непрерывного аргумента.

1. Положим, что аргумент X является дискретной случайной величиной.

Если всевозможным значениям аргумента соответствуют всевозможные значения функции, то вероятности соответствующих значений X и Y равны между собой.

Если различным возможным значениям X отвечают значения Y , причем среди этих значений Y есть равные между собой, то нужно складывать вероятности повторяющихся значений Y .

2. Положим, что аргумент X является непрерывной случайной величиной. Нужно определить распределение функции $Y = \varphi(X)$, зная плотность распределения случайного аргумента X . Если $y = \varphi(x)$ представляет собой дифференцируемую строго возрастающую или строго убывающую функцию, обратная функция которой $x = \psi(y)$, то плотность распределения $g(y)$ случайной величины Y находится с помощью равенства:

$$g(y) = f[\psi(y)]|\psi'(y)|.$$

Замечание. Применяя эту формулу, можно доказать, что линейная функция $Y = AX + B$ нормально распределенного аргумента X также распределена нормально. При этом для нахождения математического ожидания Y нужно в выражение функции подставить вместо аргумента X его математическое ожидание a :

$$M(Y) = Aa + B.$$

Для нахождения среднего квадратического отклонения Y , следует среднее квадратическое отклонение аргумента X умножить на модуль коэффициента при X :

$$\sigma(Y) = |A|\sigma(X).$$

11. Математическое ожидание функции одного случайного аргумента.

Функция случайного аргумента X задана законом $Y = \varphi(X)$. Пусть необходимо определить математическое ожидание этой функции, если известен закон распределения аргумента.

1. Пусть аргумент X является дискретной случайной величиной с возможными значениями x_1, x_2, \dots, x_n , вероятности которых соответственно равны p_1, p_2, \dots, p_n .

Ясно, что Y также является дискретной случайной величиной с возможными значениями $y_1 = \varphi(x_1), y_2 = \varphi(x_2), \dots, y_n = \varphi(x_n)$. Известно, что событие «величина X приняла значение x_i означает, что событие «величина Y приняла значение $\varphi(x_i)$. В связи с этим можно говорить, что вероятности возможных значений Y соответственно равны p_1, p_2, \dots, p_n . Таким образом, математическое ожидание функции равно:

$$M[\varphi(X)] = \sum_{i=1}^n \varphi(x_i) p_i.$$

2. Допустим, аргумент X является непрерывной случайной величиной, которая задана плотностью распределения $f(x)$. Чтобы найти математическое ожидание функции $Y = \varphi(X)$, для начала можно отыскать плотность распределения $g(y)$ величины Y , а потом применить формулу:

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

В том случае, если отыскать функцию $g(y)$ затруднительно, можно непосредственно вычислить математическое ожидание функции $\varphi(X)$ по формуле:

$$M[\varphi(X)] = \int_{-\infty}^{\infty} \varphi(x)f(x)dx.$$

Например, если возможные значения X лежат в интервале (a, b) , то

$$M[\varphi(X)] = \int_a^b \varphi(x)f(x)dx.$$

12. Функция двух случайных аргументов. Распределение суммы независимых слагаемых. Устойчивость нормального распределения.

Пусть каждой паре возможных значений случайных величин X и Y соответствует одно возможное значение случайной величины Z . В этом случае Z носит название функции двух случайных аргументов X и Y :

$$Z = \varphi(X, Y).$$

Довольно часто нужным бывает отыскание распределения функции $Z = X + Y$ по известным распределениям слагаемых. К примеру, если X есть погрешность показаний измерительного прибора, распределенная нормально, а Y является погрешностью округления показаний до ближайшего деления шкалы, которая распределена равномерно, то возникает потребность в нахождении закона распределения суммы погрешностей $Z = X + Y$.

1. Положим, что X и Y являются дискретными независимыми случайными величинами. Для составления закона распределения функции $Z = X + Y$ необходимо найти все возможные значения Z , а также их вероятности.

2. Положим, что X и Y есть непрерывные случайные величины. Если X и Y также являются независимыми, то плотность распределения $g(z)$ суммы $Z = X + Y$ (при условии, что плотность хотя бы

одного из аргументов задана на интервале $(-\infty, \infty)$ одной формулой) можно найти, используя или равенство:

$$g(z) = \int_{-\infty}^{\infty} f_1(x)f_2(z-x)dx$$

или равносильное равенство:

$$g(z) = \int_{-\infty}^{\infty} f_1(z-y)f_2(y)dy.$$

Когда возможные значения аргументов не являются отрицательными, то $g(z)$ можно найти по формуле:

$$g(z) = \int_0^z f_1(x)f_2(z-x)dx.$$

Можно также использовать равносильное равенство:

$$g(z) = \int_0^z f_1(z-y)f_2(y)dy.$$

Композиция — это плотность распределения суммы независимых случайных величин.

Устойчивый закон распределения вероятностей — это закон, при котором композиция законов есть тот же закон, отличающийся параметрами. Нормальному закону присуще свойство устойчивости. Свойство устойчивости говорит: композиция нормальных законов также имеет нормальное распределение. Другими словами, математическое ожидание и дисперсия этой композиции равны соответственно суммам математических ожиданий и дисперсий слагаемых.

К примеру, пусть X и Y есть независимые случайные величины, распределенные нормально с математическими ожиданиями и дисперсией, соответственно равными $a_1 = 3$, $a_2 = 4$, $D_1 = 1$, $D_2 = 0,5$. Тогда композиция этих величин (т. е. плотность вероятности суммы $Z = X + Y$) также распределена нормально. При этом математическое ожидание и дисперсия композиции равны $a = 3 + 4 = 7$, $D = 1 + 0,5 = 1,5$.

13. Распределение «хи квадрат». Распределение F Фишера - Снедекора. Распределение Стьюдента.

Распределение «хи квадрат». Пусть X_i ($i = 1, 2, \dots, n$) являются нормальными независимыми случайными величинами.

При этом математическое ожидание каждой из этих величин равняется нулю, а среднее квадратическое отклонение равно единице. Тогда сумма квадратов этих величин равна:

$$\chi^2 = \sum_{i=1}^n X_i^2.$$

Она распределена по закону χ^2 («хи квадрат»), имея $k = n$ степени свободы. В том случае, если эти величины связаны одним линейным соотношением, например $\sum X_i = nX$, число степеней свободы $k = n - 1$.

Плотность такого распределения можно записать следующим образом:

$$f(x) = \begin{cases} 0 & \text{при } x \leq 0 \\ \frac{1}{2^{\frac{k}{2}} \Gamma(k/2)} e^{-x/2} x^{(k/2)-1} & \text{при } x > 0 \end{cases}$$

Здесь $\Gamma = \int_0^{\infty} t^{x-1} e^{-t} dt$ представляет собой гамма-функцию, а в конкретном случае

$$\Gamma(n + 1) = n!.$$

Таким образом, распределение «хи квадрат» определяется одним параметром, которое есть число степеней свободы k .

Следует отметить, что с увеличением числа степеней свободы распределение медленно приближается к нормальному.

Распределение F Фишера-Снедекора. Пусть U и V есть независимые случайные величины, которые распределены по закону χ^2 со степенями свободы k_1 и k_2 . В этом случае величина

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

обладает распределением, которое называется распределением F Фишера — Снедекора со степенями свободы k_1 и k_2 . Иногда его обозначают через V^2 .

Плотность этого распределения можно записать следующим образом:

$$f(x) = \begin{cases} 0 & \text{при } x \leq 0 \\ C_0 \frac{x^{(k_1-2)/2}}{(k_2 + k_1 x)^{(k_1+k_2)/2}} & \text{при } x > 0 \end{cases}$$

где

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

Отсюда можно сделать вывод, что распределение F определяется двумя параметрами — числами степеней свободы.

Распределение Стьюдента. Если Z есть нормальная случайная величина, причем $M(Z) = 0$, $\sigma(Z) = 1$, а V — независимая от Z величина, распределенная по закону χ^2 с k степенями свободы, то величина

$$T = \frac{Z}{\sqrt{V/k}}$$

имеет распределение, которое называют t -распределением (или распределением Стьюдента) с k степенями свободы.

Таким образом, отношение нормированной нормальной величины к квадратному корню из независимой случайной величины, распределенной по закону «хи квадрат» с k степенями свободы, деленной на k , распределено по закону Стьюдента с k степенями свободы. При этом чем больше число степеней свободы, тем распределение Стьюдента быстрее приближается к нормальному.

8. Показательное распределение

1. Определение показательного распределения. Вероятность попадания в заданный интервал показательного распределенной случайной величины.

Распределение вероятностей непрерывной случайной величины X , которое описывается плотностью

$$f(x) = \begin{cases} 0 & \text{при } x < 0 \\ \lambda e^{-\lambda x} & \text{при } x \geq 0, \end{cases}$$

называют *показательным (экспоненциальным) распределением*.

Здесь λ представляет собой постоянную положительную величину, которой определяется показательное распределение. В этом его преимущество по сравнению с распределениями, зависящими от большого числа параметров. Как правило, параметры неизвестны, и приходится находить их приближенные значения. Однако легче оценить один параметр, чем два или три. Например, непрерывной случайной величиной, распределенной по показательному закону, может быть время между появлениями двух последовательных событий простейшего потока.

Пусть требуется вычислить функцию распределения показательного закона:

$$F(x) = \int_{-\infty}^x f(x) dx = \int_{-\infty}^0 0 dx + \lambda \int_0^x e^{-\lambda x} dx = 1 - e^{-\lambda x}.$$

Следовательно,

$$F(x) = \begin{cases} 0 & \text{при } x < 0 \\ \lambda e^{-\lambda x} & \text{при } x \geq 0. \end{cases}$$

Здесь показательный закон был определен при помощи плотности распределения; однако его можно определить и при помощи функции распределения.

Вероятность попадания в заданный интервал показательно распределенной случайной величины.

Пусть непрерывная случайная величина X распределена по показательному закону, заданному функцией распределения

$$F(x) = 1 - e^{-\lambda x} \quad (x \geq 0).$$

Вычислим вероятность попадания в интервал (a, b) этой величины.

Зная, что

$$P(a < X < b) = F(b) - F(a),$$

и учтя, что $F(a) = 1 - e^{-\lambda a}$, $F(b) = 1 - e^{-\lambda b}$, запишем:

$$P(a < X < b) = e^{-\lambda a} - e^{-\lambda b}.$$

Причем значения функции e^{-x} находят по таблице.

2. Числовые характеристики показательного распределения.

Пусть дана непрерывная случайная величина X , которая распределена по показательному закону

$$f(x) = \begin{cases} 0 & \text{при } x < 0 \\ \lambda e^{-\lambda x} & \text{при } x \geq 0. \end{cases}$$

Для того чтобы найти математическое ожидание

$$M(X) = \int_0^{\infty} x f(x) dx = \lambda \int_0^{\infty} x e^{-\lambda x} dx,$$

следует проинтегрировать по частям. В результате получится

$$M(X) = 1/\lambda.$$

Следовательно, математическое ожидание показательного распределения равняется обратной величине параметра X . Для того чтобы найти дисперсию

$$D(x) = \int_0^{\infty} x^2 f(x) dx - [M(x)]^2 = \lambda \int_0^{\infty} x^2 e^{-\lambda x} dx - 1/\lambda^2,$$

следует проинтегрировать по частям. Тогда получим:

$$\lambda \int_0^{\infty} x^2 e^{-\lambda x} dx = 2/\lambda^2.$$

Значит, можно записать, что

$$D(X) = 1/\lambda^2.$$

Для нахождения среднего квадратического отклонения нужно сначала извлечь квадратный корень из дисперсии:

$$\sigma(X) = 1/\lambda.$$

Тогда, сравнив $M(X) = 1/\lambda$ и $\sigma(X) = 1/\lambda$, можно сказать, что

$$M(X) = \sigma(X) = 1/\lambda.$$

Это означает, что математическое ожидание и среднее квадратическое отклонение показательного распределения равны между собой.

Замечание 1. Если на практике исследуется показательное распределенная случайная величина, к тому же параметр λ неизвестен, причем математическое ожидание тоже неизвестно, то считают его приближенное значение, в качестве которой принимают выборочную среднюю \bar{x} . В этом случае приближенное значение параметра λ считают посредством следующего равенства:

$$\lambda^* = 1/\bar{x}.$$

Замечание 2. Если есть повод думать, что изучаемая на практике случайная величина обладает показательным распределением, то, чтобы проверить эту гипотезу, считают оценку математического ожидания и среднего квадратического отклонения. Другими словами, ищут выборочную среднюю и выборочное среднее квадратическое отклонение. Поскольку математическое ожидание и среднее квадратическое отклонение показательного распределения равны между собой, их оценки должны различаться несущественно. В случае если оценки не будут почти отличаться одна к другой, то данные наблюдений подтверждают гипотезу о показательном распределении изучаемой величины. И наоборот, если оценки значительно отличаются, то гипотезу следует отвергнуть.

Показательное распределение весьма обширно применяется в приложениях (например, в теории надежности одним из основных понятий которой является функция надежности).

3. Функция надежности. Показательный закон надежности.

Допустим элемент (некоторое устройство) начинает работать в момент времени $t_0 = 0$, а по окончании времени длительностью t элемент перестает работать. Пусть T — непрерывная случайная величина, представляющая собой длительность времени безотказной работы элемента.

Если элемент работал исправно (до наступления отказа) время, которое меньше t , то за время длительностью t произойдет отказ.

Значит, функция распределения $F(t) = P(T < t)$ обуславливает вероятность отказа за время длительностью t . Таким образом, вероятность исправной работы за это же время t , т. е. вероятность противоположного события $T > t$, будет равняться:

$$R(t) = P(T > t) = 1 - F(t).$$

Функция надежности $R(t)$ — это функция, которая обуславливает вероятность исправной работы элемента за время длительностью t .

$$R(t) = P(T > t).$$

Показательный закон надежности.

Нередко длительность времени исправной работы элемента имеет показательное распределение. Функция этого распределения записывается следующим образом:

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

Следовательно, так как $R(t) = P(T > t) = 1 - F(t)$, то функция надежности в случае показательного распределения времени исправной работы элемента будет записана так:

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

Показательный закон надежности — это функция надежности, которая определяется следующим равенством:

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

где λ обозначает интенсивность отказов.

Благодаря этой формуле можно найти вероятность исправной работы элемента на интервале времени длительностью t , если время безотказной работы имеет показанное распределение.

Замечание. Если неисправности элементов в случайные моменты времени образуют простейший поток, то вероятность того, что за время длительностью t не будет ни одной неисправности, будет выглядеть следующим образом:

$$P_1(t) = e^{-\lambda t}.$$

Эта формула согласуется с равенством $R(t) = e^{-\lambda t}$, так как λ , в обеих формулах имеет один и тот же смысл — постоянная интенсивность отказов.

4. Характеристическое свойство показательного закона надежности.

Показательный закон надежности довольно прост и удобен для решения практических задач. При этом весьма многие формулы теории надежности существенно упрощаются, поскольку этот закон имеет важное свойство: вероятность безотказной работы элемента за время t не зависит от времени предыдущей работы до начала рассматриваемого интервала, а зависит только от длительности времени t (при заданной интенсивности отказов λ).

Чтобы доказать свойство сначала, введем обозначения событий: A — исправная работа элемента на интервале $(0, t_0)$ длительностью t_0 ; B — исправная работа на интервале $(t_0, t_0 + t)$ длительностью t . Значит, AB представляет собой исправную работу на интервале $(0, t_0 + t)$ длительностью $t_0 + t$.

Подсчитать вероятности этих событий можно по формуле $R(t) = e^{-\lambda t}$:

$$P(A) = e^{-\lambda t_0}, P(B) = e^{-\lambda t}, P(AB) = e^{-\lambda(t_0+t)} = e^{-\lambda t_0} e^{-\lambda t}.$$

Условная вероятность того, что элемент будет работать исправно на интервале $(t_0, t_0 + t)$ при условии, что он уже проработал исправно на предыдущем интервале $(0, t_0)$, определяется так:

$$P_A(B) = \frac{P(AB)}{P(A)} = \frac{e^{-\lambda t_0} e^{-\lambda t}}{e^{-\lambda t_0}} = e^{-\lambda t}.$$

В данной формуле нет t_0 , а есть только t , а это говорит о следующем: время работы на предыдущем интервале не влияет на величину вероятности безотказной работы на последующем интервале, а зависит только от длины последующего интервала.

Свойство доказано.

Полученный результат можно выразить немного по-другому. При сравнении вероятности $P(B) = e^{-\lambda t}$ и $P_A(B) = e^{-\lambda t}$ можно сказать, что условная вероятность исправной работы элемента на интервале длительностью t , подсчитанная в предположении, что

элемент проработал исправно на предыдущем интервале, равняется безусловной вероятности.

Таким образом, в случае показательного закона надежности исправная работа элемента в прошлом не влияет на величину вероятности его безотказной работы в ближайшем будущем.

Замечание. Рассматриваемым свойством обладает лишь показательное распределение, следовательно, если изучаемая случайная величина этим свойством обладает, то она распределена по показательному закону. К примеру, при допущении, что метеориты распределены равномерно в пространстве и во времени, вероятность столкновения метеорита с космическим кораблем зависит от того, сталкивались или нет метеориты в корабль до начала рассматриваемого интервала времени. Значит, случайные моменты времени столкновения метеоритов с космическим кораблем распределены по показательному закону.

9. Система двух случайных величин

1. Понятие о системе нескольких случайных величин.

Случайными величинами, возможные значения которых определялись одним числом, называются одномерные величины. К примеру, количество очков, которое может выпасть при бросании игральной кости, является дискретной одномерной величиной. При этом расстояние от орудия до места падения снаряда является непрерывной одномерной случайной величиной.

Величины, у которых возможные значения определяются двумя, тремя, ..., n числами, называются соответственно двумерными, трехмерными, ..., n -мерными.

Пусть (X, Y) есть двумерная случайная величина, каждую из которых называют составляющей (компонентой). Обе эти величины X и Y , которые рассматриваются одновременно, создадут систему двух случайных величин. Также n -мерную величину можно рассматривать как систему n случайных величин. К примеру, трехмерная величина (X, Y, Z) обуславливает систему трех случайных величин X, Y и Z .

Пример. Автоматическая машина делает стальные плитки. Если длина X и ширина Y являются контролируемыми размерами, то будет иметь место двумерная случайная величина (X, Y) . В том случае если контролируется и высота Z , то будем иметь трехмерную величину (X, Y, Z) .

Также двумерная случайная величина (X, Y) геометрически объясняется или как случайная точка $M(X, Y)$ на плоскости (т. е. как точка со случайными координатами), или как случайный вектор \overline{OM} . Аналогично трехмерная случайная величина геометрически истолковывается как точка $M(X, Y, Z)$ в трехмерном пространстве или как вектор \overline{OM} .

Рациональным считается различать дискретные (составляющие этих величин дискретны) и непрерывные (составляющие этих величин непрерывны) многомерные случайные величины.

2. Закон распределения вероятностей дискретной двумерной случайной величины.

Закон распределения дискретной двумерной случайной величины — это перечень возможных значений этой величины, т. е. пар чисел (x_i, y_j) и их вероятностей $p(x_i, y_j)$ ($i = 1, 2, \dots, n; j = 1, 2, \dots, m$). Как правило, закон распределения задается таблицей с двойным входом (см. табл. 1).

В первой строке таблицы имеются все возможные значения составляющей X , а в первом столбце — все возможные значения составляющей Y . В ячейке, которая находится на пересечении столбца x_i и строки y_j , показана вероятность $p(x_i, y_j)$ того, что двумерная случайная величина примет значение (x_i, y_j) .

Поскольку события $(X = x_i, Y = y_j)$ ($i = 1, 2, \dots, n; j = 1, 2, \dots, m$) представляют собой полную группу, то сумма вероятностей, находящихся во всех клетках таблицы, равна единице.

Закон распределения заданный таблицей с двойным входом.

Таблица 1

Y	X			
	x_1	x_2	x_j	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	$p(x_1, y_m)$	$p(x_2, y_m)$	$p(x_i, y_m)$	$p(x_n, y_m)$

Если известен закон распределения двумерной дискретной случайной величины, имеется возможность найти законы распределения каждой из составляющих. В самом деле, например, события $(X = x_1, Y = y_1), (X = x_1, Y = y_2), \dots, (X = x_1, Y = y_m)$

несовместны, значит, вероятность $P(x_1)$ того, что X примет значение x_1 , по теореме сложения будет следующей:

$$P(x_1) = p(x_1, y_1) + p(x_1, y_2) + \dots + p(x_1, y_m).$$

Следовательно, вероятность того, что X примет значение x_1 , равняется сумме вероятностей столбца x_1 . В общем, для нахождения вероятности $P(X = x_i)$ необходимо просуммировать вероятности столбца x_i . Таким же образом, если сложить вероятности строки y_j , получится вероятность $P(Y = y_j)$.

3. Функция распределения двумерной случайной величины. Свойства функции распределения двумерной случайной величины.

Пусть дана двумерная случайная величина (X, Y) , безразлично, дискретная она или непрерывная.

Допустим, x, y есть пара действительных чисел, тогда вероятность события, что X примет значение, которое меньше x , причем Y примет значение, которое меньше y , обозначится $F(x, y)$. Если x и y будут изменяться, то будет изменяться и $F(x, y)$. Другими словами, $F(x, y)$ — это функция от x и y .

Функция распределения двумерной случайной величины (X, Y) — это функция $F(x, y)$, которая определяет для каждой пары чисел x, y вероятность того, что X примет значение, меньшее x , причем Y примет значение, меньшее y :

$$F(x, y) = P(X < x, Y < y).$$

Это равенство геометрически можно объяснить следующим образом: $F(x, y)$ представляет собой вероятность того, что случайная точка (X, Y) попадет в бесконечный квадрант с вершиной (x, y) , который расположен левее и ниже этой вершины.

Свойства функции распределения двумерной случайной величины.

Свойство 1. Значения функции распределения удовлетворяют двойному неравенству:

$$0 \leq F(x, y) \leq 1.$$

Доказательство. Данное свойство получается из определения функции распределения как вероятности: вероятность — всегда неотрицательное число, которое не превышает единицу.

Свойство 2. $F(x, y)$ является неубывающей функцией по каждому аргументу, другими словами.

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

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

Доказательство. Докажем, что $F(x, y)$ есть неубывающая функция по аргументу x . Событие, заключающееся в том, что составляющая X примет значение, меньшее x_2 , причем составляющая $Y < y$, можно разделить на следующие два несовместных события:

- 1) X примет значение, меньшее x_1 , причем $Y < y$ с вероятностью $P(X < x_1, Y < y)$;
- 2) X примет значение, которое удовлетворяет неравенству, причем $Y < y$ с вероятностью $P(x_1 \leq X < x_2, Y < y)$.

На основании теоремы сложения будем иметь:

$$P(X < x_2, Y < y) = P(X < x_1, Y < y) + P(x_1 \leq X < x_2, Y < y),$$

а отсюда

$$P(X < x_2, Y < y) - P(X < x_1, Y < y) = P(x_1 \leq X < x_2, Y < y)$$

или

$$F(x_2, y) - F(x_1, y) = P(x_1 \leq X < x_2, Y < y).$$

Всякая вероятность представляет собой неотрицательное число, а значит,

$$F(x_2, y) - F(x_1, y) \geq 0 \quad \text{или } F(x_2, y) \geq F(x_1, y).$$

Свойство доказано.

Более ясным свойство становится тогда, когда есть возможность воспользоваться геометрическим объяснением функции распределения как вероятности попадания случайной точки в бесконечный квадрант с вершиной (x, y) . Следовательно, при возрастании x правая граница этого квадранта перемещается вправо, причем вероятность попадания случайной точки в новый квадрант не может уменьшиться.

Точно так же можно доказать, что $F(x, y)$ — это неубывающая функция по аргументу y .

Свойство 3. Имеют место предельные соотношения:

- 1) $F(-\infty, y) = 0$;
- 2) $F(x, -\infty) = 0$;
- 3) $F(\infty, -\infty) = 0$;
- 4) $F(\infty, \infty) = 1$.

Доказательство.

1. Вероятность события $X < -\infty$ и $Y < -\infty$ обозначим через $F(-\infty, y)$.

Однако такое событие невозможно, так как невозможно событие $X < -\infty$, а значит, вероятность этого события равна нулю.

Наиболее понятным становится свойство, если прибегнуть к геометрической интерпретации: при $x \rightarrow -\infty$ правая граница бесконечного квадранта неограниченно сдвигается влево, причем вероятность попадания случайной точки в квадрант стремится к нулю.

2. Поскольку событие $Y < -\infty$ невозможно, значит, $F(x, -\infty) = 0$.

3. Так как событие $X < -\infty$ и $Y < -\infty$ невозможно, то $F(-\infty, -\infty) = 0$.

4. Событие $X < \infty$ и $Y < \infty$ достоверно, а значит, вероятность этого события $F(\infty, \infty) = 1$.

Данное свойство будет наглядно ясным, если учесть, что при $x \rightarrow \infty$ и $y \rightarrow \infty$ бесконечный квадрант превращается во всю плоскость xOy . Таким образом, попадание случайной точки (X, Y) в эту плоскость является достоверным событием.

Свойство 4.

1. Функция распределения системы при $y = \infty$ делается функцией распределения составляющей X :

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

2. Функция распределения системы при $x = \infty$ будет функцией распределения составляющей Y :

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

Доказательство.

1. Поскольку событие $Y < \infty$ достоверно, то $F(x, \infty)$ определяет вероятность события $X < x$. Другими словами оно представляет собой функцию распределения составляющей X .
2. Доказательство производится аналогично.

4. Вероятность попадания случайной точки в полуполосу.

Если применить функцию распределения системы случайных величин X и Y , то можно без труда подсчитать вероятность того, что в результате испытания случайная точка попадает в полуполосу $x_1 < X < x_2$ и $Y < y$, а также в полуполосу $X < x$ и $y_1 < Y < y_2$.

После того как вычли из вероятности попадания случайной точки квадрант с вершиной (x_2, y) вероятность попадания точки в квадрант с вершиной (x_1, y) , получилось:

$$P(x_1 \leq X < x_2, Y < y) = F(x_2, y) - F(x_1, y).$$

Точно так же получится:

$$P(X < x_1, y_1 \leq Y < y_2) = F(x, y_2) - F(x, y_1).$$

Следовательно, вероятность попадания случайной точки в полуполосу равняется приращению функции распределения по одному из аргументов.

5. Вероятность попадания случайной точки в прямоугольник.

Пусть имеется прямоугольник $ABCD$, у которого стороны параллельны координатным осям, а уравнения сторон записываются в виде:

$$X = x_1, X = x_2, Y = y_1 \text{ и } Y = y_2.$$

Для нахождения вероятности попадания случайной точки (X, Y) в этот прямоугольник нужно из вероятности попадания случайной точки в полуполосу AB с вертикальной штриховкой (эта вероятность равна $F(x_2, y_2) - F(x_1, y_2)$) отнять вероятность попадания точки в полуполосу CD с горизонтальной штриховкой (эта

вероятность равна $F(x_2, y_1) - F(x_1, y_1)$):

$$\begin{aligned} P(x_1 \leq X < x_2, y_1 \leq Y < y_2) &= \\ &= [F(x_2, y_2) - F(x_1, y_2)] - [F(x_2, y_1) - F(x_1, y_1)]. \end{aligned}$$

6. Плотность совместного распределения вероятностей непрерывной двумерной случайной величины (двухмерная плотность вероятности).

С помощью функции распределения задавалась двумерная случайная величина, а непрерывную двумерную величину можно задать, применяя плотность распределения. Допустим, функция распределения $F(x, y)$ всюду непрерывна и имеет всюду (за исключением конечного числа кривых) непрерывную частную производную второго порядка.

Плотность совместного распределения вероятностей $F(x, y)$ двумерной непрерывной случайной величины (X, Y) есть вторая смешанная частная производная от функции распределения:

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

Геометрически эту функцию можно представить в виде поверхности, называемой поверхностью распределения.

7. Нахождение функции распределения системы по известной плотности распределения.

Если известна плотность совместного распределения $f(x, y)$, то можно вычислить функцию распределения $F(x, y)$ по следующей формуле:

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

Это непосредственно вытекает из определения плотности распределения двумерной непрерывной случайной величины (X, Y) .

8. Вероятностный смысл двумерной плотности вероятности.

Вероятность того, что случайная точка (X, Y) попадет в прямоугольник $ABCD$, будет равняться:

$$\begin{aligned} P(x_1 < X < x_2, y_1 < Y < y_2) &= \\ &= [F(x_2, y_2) - F(x_1, y_2)] - [F(x_2, y_1) - F(x_1, y_1)]. \end{aligned}$$

Обозначим левую часть равенства как P_{ABCD} , а также применим к правой части теорему Лагранжа. В результате получится:

$$P_{ABCD} = F''_{xy}(\xi, \eta) \Delta x \Delta y,$$

где

$$x_1 < \xi < x_2, \quad \Delta x = x_2 - x_1, \quad y_1 < \eta < y_2, \quad \Delta y = y_2 - y_1.$$

Из этого следует:

$$F''_{xy}(\xi, \eta) = \frac{P_{ABCD}}{\Delta x \Delta y}$$

или

$$f(\xi, \eta) = \frac{P_{ABCD}}{\Delta x \Delta y}.$$

Поскольку $\Delta x \Delta y$ равняется площади прямоугольника $ABCD$, сделаем вывод: $f(\xi, \eta)$ представляет собой отношение вероятности попадания случайной точки в прямоугольник $ABCD$ к площади этого прямоугольника.

Далее перейдем в равенстве

$$f(\xi, \eta) = \frac{P_{ABCD}}{\Delta x \Delta y}$$

к пределу при $\Delta x \rightarrow 0$ $\Delta y \rightarrow 0$, тогда, $\xi \rightarrow x$, $\eta \rightarrow y$, а значит,

$$f(\xi, \eta) \rightarrow f(x, y).$$

Таким образом, функция $f(x, y)$ может быть рассмотрена как предел отношения вероятности попадания случайной точки в прямоугольник (со сторонами Δx и Δy) к площади этого прямоугольника, когда обе стороны прямоугольника стремятся к нулю.

9. Вероятность попадания случайной точки в произвольную область.

Перепишем соотношение $f(\xi, \eta) = \frac{P_{ABCD}}{\Delta x \Delta y}$ следующим образом:

$$f(\xi, \eta) \Delta x \Delta y = P_{ABCD}.$$

Из этого можно сделать вывод: произведение $f(\xi, \eta) \Delta x \Delta y$ представляет собой вероятность попадания случайной точки в прямоугольник со сторонами Δx и Δy .

Если в плоскости xOy задана произвольная область D , то можно ввести обозначение события, которое состоит в попадании случайной точки в эту область: $(X, Y) \subset D$.

Теперь разобьем область D на n элементарных областей прямыми, которые параллельны оси Oy и находятся на расстоянии одна от другой, а также прямыми, которые параллельны оси Ox и сходятся на расстоянии Δx одна от другой. Здесь для простоты предполагается, что эти прямые пересекают контур области не более чем в двух точках. Поскольку события, заключающиеся в попадании случайной точки в элементарные области, несовместны, то вероятность попадания в область D приближенно (сумма элементарных областей приближенно равна области D) равна сумме вероятностей попаданий точки в элементарные области:

$$P((X, Y) \subset D) \approx \sum_{i=1}^n f(\xi_i, \eta_i) \Delta x \Delta y.$$

После перехода к пределу при $\Delta x \rightarrow 0$ и $\Delta y \rightarrow 0$, получится:

$$P((X, Y) \subset D) = \iint_{(D)} f(x, y) dx dy.$$

Следовательно, для вычисления вероятности попадания случайной точки (X, Y) в область D , достаточно найти двойной интеграл по области D от функции $f(x, y)$.

Равенство

$$P((X, Y) \subset D) = \iint_{(D)} f(x, y) dx dy$$

геометрически можно объяснить следующим образом: вероятность попадания случайной точки (X, Y) в область D равна объему тела, ограниченного сверху поверхностью $z = f(x, y)$, основанием которого служит проекция этой поверхности на плоскость xOy .

Замечание. Подынтегральное выражение $f(x, y)dxdy$ есть элемент вероятности. В свою очередь, элемент вероятности находит вероятность попадания случайной точки в элементарный прямоугольник со сторонами dx и dy .

10. Свойства двумерной плотности вероятности.

Свойство 1. Двумерная плотность вероятности неотрицательна:

$$f(x, y) \geq 0.$$

Доказательство. Вероятность попадания случайной точки в прямоугольник со сторонами Δx и Δy представляет собой неотрицательное число, при этом площадь этого прямоугольника есть положительное число. Таким образом, отношение этих двух чисел, а следовательно, и их предел (при $\Delta x \rightarrow 0$ и $\Delta y \rightarrow 0$), равный $f(x, y)$, есть неотрицательное число:

$$f(x, y) \geq 0.$$

Это свойство непосредственно вытекает из того, что $F(x, y)$ есть неубывающая функция своих аргументов.

Свойство 2. Двойной несобственный интеграл с бесконечными пределами от двумерной плотности равняется единице:

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

Доказательство. Бесконечные пределы интегрирования свидетельствуют о том, что область интегрирования есть вся плоскость xOy . Так как событие, заключающееся в том, что случайная точка попадет при испытании на плоскость xOy , достоверно, то вероятность этого события равна единице, т. е.

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

11. Отыскание плотностей вероятности составляющих двумерной случайной величины.

Допустим, мы знаем, чему равна плотность совместного распределения вероятностей системы двух случайных величин.

Для нахождения плотности распределения каждой из составляющих сначала нужно найти плотность распределения составляющей X . Пусть $F_1(x)$ есть функция распределения составляющей X . На основании определения плотности распределения одномерной случайной величины можно записать:

$$f_1(x) = \frac{dF_1(x)}{dx}.$$

Учтем, что

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

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

тогда можно найти

$$F_1(x) = \int_{-\infty}^x \int_{-\infty}^{\infty} f(x, y) dx dy.$$

Если продифференцировать обе части этого равенства по x , то получится

$$\frac{dF_1}{dx} = \int_{-\infty}^{\infty} f(x, y) dy$$

или

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy.$$

Таким же образом находится плотность распределения составляющей Y :

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

Следовательно, плотность распределения одной из составляющих равна несобственному интегралу с бесконечными пределами от плотности совместного распределения системы.

При этом переменная интегрирования соответствует другой составляющей.

12. Условные законы распределения составляющих системы дискретных случайных величин.

Если события A и B зависимы, то условная вероятность события B отличается от его безусловной вероятности, тогда

$$P(A|B) = P(AB)/P(A).$$

Подобное положение имеет место и для случайных величин. Для того чтобы объяснить зависимость между составляющими двухмерной случайной величины, вводят понятие условного распределения.

Пусть дана дискретная двухмерная случайная величина (X, Y) , а возможные значения составляющих таковы: $x_1, x_2, \dots, x_n; y_1, y_2, \dots, y_m$.

Предположим, что в результате испытания величина Y приняла значение $Y = y_1$, при этом X примет одно из своих возможных значений: x_1 , или x_2 , или x_n . Пусть $p(x_i|y_1)$ есть условная вероятность того, что X примет, например, значение x_1 , при условии, что $Y = y_1$. Эта вероятность не будет равняться безусловной вероятности $p(x_1)$.

Условные вероятности составляющей в общем случае обозначим следующим образом:

$$p(x_i|y_j) \quad (i = 1, 2, \dots, n; j = 1, 2, \dots, m).$$

Условное распределение составляющей X при $Y = y_j$ есть совокупность условных вероятностей

$$p(x_1|y_j), p(x_2|y_j), \dots, p(x_n|y_j),$$

которые находятся в предположении, что событие $Y = y_j$, (j имеет одно и то же значение при всех значениях X) уже наступило. Точно так же находится условное распределение составляющей Y .

Если известен закон распределения двухмерной дискретной случайной величины, то можно применить формулу $P_A(B) = P(AB)/P(A)$, тем самым вычислить условные законы распределения составляющих.

К примеру, условный закон распределения X в предположении, что событие $Y = y_1$ уже произошло, можно найти по формуле:

$$p(x_i|y_1) = \frac{p(x_i, y_1)}{p(y_1)} \quad (i = 1, 2, \dots, n).$$

Условные законы распределения составляющей X в общем случае обуславливаются соотношением:

$$p(x_i|y_j) = \frac{p(x_i, y_j)}{p(y_j)}.$$

Подобным образом находят условные законы распределения составляющей Y :

$$p(y_j|x_i) = \frac{p(x_i, y_j)}{p(x_i)}.$$

Замечание. Сумма вероятностей условного распределения равняется единице. В самом деле, поскольку при фиксированном y_j имеем

$$\sum_{i=1}^n p(x_i, y_j) = P(y_j), \quad \text{то}$$

$$\sum_{i=1}^n p(x_i|y_j) = \sum_{i=1}^n p(x_i, y_j) / p(y_j) = p(y_j) / p(y_j) = 1.$$

Таким же образом доказывается, что при фиксированном x_i

$$\sum_{j=1}^m p(y_j|x_i) = 1.$$

Данное свойство условных распределений применяют для контроля выделений.

13. Условные законы распределения составляющих системы непрерывных случайных величин.

Пусть (X, Y) есть непрерывная двумерная случайная величина.

Условная плотность распределения составляющих X при данном значении $Y = y$ есть отношение плотности совместного распределения $f(x, y)$ системы (X, Y) к плотности распределения $f_2(y)$ составляющей Y :

$$\varphi(x|y) = f(x, y) / f_2(y).$$

Условная плотность $\varphi(x|y)$ отличается от безусловной плотности $f_1(x)$ тем, что функция $\varphi(x|y)$ дает распределение X при условии, что составляющая Y приняла значение $Y = y$; а функция $f_1(x)$ дает распределение X независимо от того, какие из возможных значений приняла составляющая Y .

Таким же образом вычисляется условная плотность составляющей Y при данном значении $X = x$:

$$\psi(y|x) = f(x, y) / f_1(x).$$

Когда известна плотность совместного распределения $f(x, y)$, тогда условные плотности составляющих могут быть подсчитаны в силу $PA(B) = P(AB)/P(A)$ и $\psi(y|x) = f(x, y) / f_1(x)$. по формулам:

$$\varphi(x|y) = f(x, y) / \int_{-\infty}^{\infty} f_2(x, y) dx;$$

$$\psi(y|x) = f(x, y) / \int_{-\infty}^{\infty} f_2(x, y) dy.$$

Перепишем формулы $PA(B) = P(AB)/P(A)$ и $\psi(y|x) = f(x, y) / f_1(x)$ следующим образом:

$$f(x, y) = f_2(y)\varphi(x|y) \quad , \quad f(x, y) = f_1(x)\psi(y|x).$$

Отсюда сделаем вывод: умножая закон распределения одной из составляющих на условный закон распределения другой составляющей, можно определить закон распределения системы случайных величин.

Условные плотности обладают следующими свойствами:

$$\varphi(x|y) \geq 0, \quad \int_{-\infty}^{\infty} \varphi(x|y) dx = 1;$$

$$\psi(y|x) \geq 0, \quad \int_{-\infty}^{\infty} \psi(y|x) dy = 1.$$

14. Условное математическое ожидание.

Условное математическое ожидание является существенной характеристикой условного распределения вероятностей.

Условное математическое ожидание дискретной случайной величины Y при $X = x$ (x — определенное возможное значение X) есть произведение возможных значений Y на их условные вероятности:

$$M(Y|X = x) = \sum_{j=1}^m y_j p(y_j|x).$$

Для непрерывных величин можно записать:

$$M(Y|X = x) = \int_{-\infty}^{\infty} y \psi(y|x) dy,$$

где $\psi(y|x)$ — условная плотность случайной величины Y при $X = x$.

Условное математическое ожидание $M(Y|x)$ представляет собой функцию от x :

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

Ее называют *функцией регрессии* Y на X .

Подобным образом можно найти условное математическое ожидание случайной величины X и функция регрессии X на Y :

$$M(X|y) = \varphi(y).$$

15. Зависимые и независимые случайные величины.

Теорема. Для того чтобы случайные величины X и Y были независимыми, необходимо и достаточно, чтобы функция распределения системы (X, Y) была равна произведению функций распределения составляющих:

$$F(x, y) = F_1(x)F_2(y).$$

Доказательство.

1. Необходимость. Если X и Y — независимые случайные величины, то события $X < x$ и $Y < y$ независимы. Значит, вероятность совмещения этих событий равна произведению их вероятностей:

$$P(X < x, Y < y) = P(X < x)P(Y < y),$$

или

$$F(x, y) = F_1(x)F_2(y).$$

2. Достаточность. Допустим, $F(x, y) = F_1(x)F_2(y)$.

Тогда

$$P(X < x, Y < y) = P(X < x)P(Y < y).$$

Это означает, что вероятность совмещения событий $X < x$ и $Y < y$ равна произведению вероятностей этих событий. Таким образом, случайные величины X и Y являются независимыми.

Следствие. Для того чтобы непрерывные случайные величины X и Y были независимыми, необходимо и достаточно, чтобы плотность совместного распределения системы (X, Y) была равна произведению плотностей распределения составляющих:

$$f(x, y) = f_1(x)f_2(y).$$

Доказательство.

1. Необходимость. Допустим, X и Y являются независимыми непрерывными случайными величинами (на основании предыдущей теоремы), т. е.

$$F(x, y) = F_1(x)F_2(y).$$

Продифференцируем это равенство по x , а потом по y . В результате получим:

$$\frac{\partial^2 F}{\partial x \partial y} = \frac{\partial F_1}{\partial x} \frac{\partial F_2}{\partial y},$$

или по определению плотностей распределения двумерной

и одномерной величин можно записать:

$$f(x, y) = f_1(x)f_2(y).$$

2. Достаточность. Предположим, что

$$f(x, y) = f_1(x)f_2(y),$$

тогда, интегрируя это равенство по x и по y , получим:

$$\int_{-\infty}^y \int_{-\infty}^x F(x, y) dx dy = \int_{-\infty}^x f_1(x) dx \int_{-\infty}^y f_2(y) dy,$$

или

$$F(x, y) = F_1(x)F_2(y).$$

Тогда, на основании предыдущей теоремы, можно сделать вывод, что X и Y независимы.

Замечание. Поскольку приведенные выше условия являются необходимыми и достаточными, то можно записать следующие определения независимых случайных величин:

- 1) две случайные величины называются независимыми, если функция распределения системы этих величин равняется произведению функций распределения составляющих;
- 2) две непрерывные случайные величины называются независимыми, если плотность совместного распределения системы этих величин равна произведению плотностей распределения составляющих.

16. Числовые характеристики систем двух случайных величин. Корреляционный момент. Коэффициент корреляции.

Чтобы описать систему двух случайных величин, кроме математических ожиданий и дисперсий составляющих, применяют и другие характеристики, к которым относятся корреляционный момент и коэффициент корреляции.

Корреляционный момент случайных величин X и Y — это математическое ожидание произведения отклонений этих величин:

$$\mu_{xy} = M \{ [X - M(X)] [Y - M(Y)] \}.$$

Чтобы найти корреляционный момент дискретных величин, используют следующую формулу:

$$\mu_{xy} = \sum_{i=1}^n \sum_{j=1}^m [x_i - M(X)] [y_j - M(Y)] p(x_i, y_j),$$

а для непрерывных величин применяют другую формулу:

$$\mu_{xy} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x - M(X)] [y - M(Y)] f(x, y) dx dy.$$

Корреляционный момент характеризует связь между величинами X и Y . При этом он равен нулю, если X и Y независимы. Таким образом, если корреляционный момент равен нулю, то X и Y являются независимыми случайными величинами.

Коэффициент корреляции r_{xy} случайных величин X и Y — это отношение корреляционного момента к произведению средних квадратических отклонений этих величин:

$$r_{xy} = \mu_{xy} / (\sigma_x \sigma_y).$$

Коэффициент корреляции r_{xy} равен корреляционному моменту нормированных величин X' и Y' :

$$\begin{aligned} r_{xy} &= \frac{M \{ [X - M(X)] [Y - M(Y)] \}}{\sigma_x \sigma_y} = \\ &= M \left[\frac{X - M(X)}{\sigma_x} \frac{Y - M(Y)}{\sigma_y} \right] = M(X'Y') = \mu_{x'y'}. \end{aligned}$$

Теорема 1. Абсолютная величина корреляционного момента двух случайных величин X и Y меньше либо равна среднему геометрическому их дисперсий:

$$|\mu_{xy}| \leq \sqrt{D_x D_y}.$$

Доказательство. Пусть $Z_1 = \sigma_y X - \sigma_x Y$ — случайная величина. Определим ее дисперсию:

$$D(Z_1) = M[Z_1 - m_{z_1}]^2.$$

После преобразований получим:

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

Любая дисперсия является неотрицательной величиной, а значит,

$$2\sigma_x^2\sigma_y^2 - 2\sigma_x\sigma_y\mu_{xy} \geq 0,$$

тогда

$$\mu_{xy} \leq \sigma_x\sigma_y.$$

Вводим случайную величину $Z_2 = \sigma_y X + \sigma_x Y$, тогда найдем

$$\mu_{xy} \geq -\sigma_x\sigma_y.$$

В совокупности $\mu_{xy} \leq \sigma_x\sigma_y$ и $\mu_{xy} \geq -\sigma_x\sigma_y$ дадут

$$-\sigma_x\sigma_y \leq \mu_{xy} \leq \sigma_x\sigma_y$$

или

$$|\mu_{xy}| \leq \sigma_x\sigma_y.$$

Итак,

$$\mu_{xy} \leq \sqrt{D_x D_y}.$$

Теорема 2. Абсолютная величина коэффициента корреляции имеет значение, не превышающее единицы:

$$|r_{xy}| \leq 1.$$

Доказательство. Разделим обе части двойного неравенства

$$-\sigma_x\sigma_y \leq \mu_{xy} \leq \sigma_x\sigma_y$$

на произведение положительных чисел $\sigma_x \sigma_y$, в результате получим:

$$-1 \leq r_{xy} \leq 1.$$

Следовательно,

$$|r_{xy}| \leq 1.$$

17. Коррелированность и зависимость случайных величин.

Коррелированными называют две случайные величины X и Y , если их корреляционный момент не равен нулю. При этом X и Y будут называться *некоррелированными* величинами, если их корреляционный момент равен нулю.

Две коррелированные величины являются зависимыми. В самом деле, предположив обратное, мы должны сделать вывод, что $\mu_{xy} = 0$, а это противоречит условию, поскольку для коррелированных величин $\mu_{xy} \neq 0$.

Обратное предположение не всегда справедливо, т. е. если две величины зависимы, то они могут быть как коррелированными, так и некоррелированными. Иначе говоря, корреляционный момент двух зависимых величин может равняться нулю, а может и не равняться.

К примеру, две зависимые величины могут быть некоррелированными, тогда из коррелированности двух случайных величин следует их зависимость, однако из зависимости еще не вытекает коррелированность. Поскольку две величины независимы, то из этого следует их некоррелированность. Однако из некоррелированности еще нельзя заключить о независимости этих величин.

Из некоррелированности нормально распределенных величин следует их независимость.

18. Нормальный закон распределения на плоскости.

Зачастую на практике встречаются двумерные случайные величины, которые обладают нормальным распределением.

Нормальный закон распределения на плоскости — это распределение вероятностей двумерной случайной величины:

(X, Y) , если

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

Как видно, нормальный закон на плоскости определяется пятью параметрами (a_1 , a_2 , σ_x , σ_y и r_{xy}), которые имеют следующий вероятностный смысл: a_1 , a_2 — математические ожидания, σ_x , σ_y — средние квадратические отклонения, r_{xy} — коэффициент корреляции величин X и Y .

Если составляющие двухмерной формально распределенной случайной величины некоррелированы, то они будут независимыми. Допустим X и Y некоррелированы, тогда если $r_{xy} = 0$, получим:

$$\begin{aligned} f(x, y) &= \frac{1}{2\pi\sigma_x\sigma_y} e^{-0,5\left[\frac{(x-a_1)^2}{\sigma_x^2} + \frac{(y-a_2)^2}{\sigma_y^2}\right]} = \\ &= \frac{1}{\sigma_x\sqrt{2\pi}} \times e^{-\left[\frac{(x-a_1)^2}{2\sigma_x^2}\right]} \times \frac{1}{\sigma_y\sqrt{2\pi}} \times e^{-\left[\frac{(y-a_2)^2}{2\sigma_y^2}\right]} = \\ &= f_1(x)f_2(y). \end{aligned}$$

Следовательно, если составляющие нормально распределенной случайной величины некоррелированы, то плотность совместного распределения системы равна произведению плотностей распределения составляющих. Из этого вытекает независимость составляющих. Обратное утверждение справедливо.

Для нормально распределенных составляющих двухмерной случайной величины понятия независимости и некоррелированности равносильны.

19. Линейная регрессия. Прямые линии среднеквадратической регрессии.

Пусть имеется двухмерная случайная величина (X, Y) , где X и Y — зависимые случайные величины, причем одна из величин является функцией другой.

Ограничимся приближенным представлением величины Y в виде линейной функции величины X :

$$Y \cong g(x) = \alpha X + \beta,$$

где α и β — параметры, которые нужно определить.

Это можно сделать, применив метод наименьших квадратов.

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

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

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

где $m_x = M(X)$, $m_y = M(Y)$, $\sigma_x = \sqrt{D(X)}$, $\sigma_y = \sqrt{D(Y)}$, $r = \mu / (\sigma_x \sigma_y)$ — коэффициент корреляции величин X и Y .

Доказательство. Рассмотрим функцию независимых двух аргументов α и β :

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

Зная, что

$$M(X - m_x) = M(Y - m_y) = 0,$$

$$M[(X - m_x) \times M(Y - m_y)] = \mu_{xy} r \sigma_x \sigma_y,$$

и выполнив преобразования, получим:

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

Для исследования функции $F(\alpha, \beta)$ на экстремум будем приравнять к нулю частные производные.

$$\begin{cases} \frac{dF}{d\alpha} = -2(m_y - \alpha - \beta m_x) = 0; \\ \frac{dF}{d\beta} = 2\beta\sigma_x^2 - 2r\sigma_x\sigma_y = 0. \end{cases}$$

Тогда

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

Рассматриваемая функция при этих значениях α и β принимает наименьшее значение. Значит, линейную среднюю квадратическую регрессию Y и X можно записать следующим образом:

$$g(X) = \alpha + \beta X = m_y - r \frac{\sigma_y}{\sigma_x} X$$

или

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

Здесь коэффициент $\beta = r \frac{\sigma_y}{\sigma_x}$ есть коэффициент регрессии Y на X , а прямая

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

представляет собой *прямую среднеквадратической регрессии* Y на X .

Подставим найденные значения a и b в соотношение $F(\alpha, \beta) = M[Y - \alpha - \beta X]^2$. В результате получим минимальное значение функции $F(a, b)$, которое равно $\sigma_y^2(1 - r^2)$. Величина $\sigma_y^2(1 - r^2)$ есть остаточная дисперсия случайной величины Y относительно случайной величины X , которая характеризует величину ошибки. Эти ошибки, в свою очередь, допускают при замене Y линейной функцией $g(X) = \alpha + \beta X$. При $r = \pm 1$ остаточная дисперсия равна нулю, т. е. при этих крайних значениях коэффициента корреляции не возникает ошибки при представлении Y в виде линейной функции от X .

Значит, если коэффициент корреляции $r = \pm 1$, то Y и X являются линейно зависимыми.

Таким же образом можно получить прямую среднеквадратической регрессии X на Y :

$$x - m_x = r \frac{\sigma_y}{\sigma_x} (Y - m_y),$$

где $r \frac{\sigma_x}{\sigma_y}$ коэффициент регрессии X на Y , и остаточную дисперсию $\sigma_x^2(1-r^2)$ величины X относительно Y .
если $r = \pm 1$, то обе прямые регрессии

$$y - m_y = r \frac{\sigma_y}{\sigma_x} (x - m_x) \text{ и } x - m_x = r \frac{\sigma_x}{\sigma_y} (Y - m_y) \text{ совпадают.}$$

Из этих уравнений вытекает: обе прямые регрессии проходят через точку (m_x, m_y) — центр совместного распределения величин X и Y .

20. Линейная корреляция. Нормальная корреляция.

Пусть дана двумерная случайная величина (X, Y) . Когда обе функции регрессии Y на X и X на Y линейны, тогда можно говорить, что X и Y связаны линейной корреляционными линиями. При этом они совпадают с прямыми среднеквадратической регрессии.

Теорема. Пусть двумерная случайная величина (X, Y) распределена нормально, тогда X и Y связаны линейной корреляционной зависимостью.

Доказательство. Двухмерная плотность вероятности имеет вид:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} e^{-\frac{(u^2+v^2-2rv)/(2(1-r^2))}{1-r^2}},$$

где $u = (x - a_1)/\sigma_x$, $v = (y - a_2)/\sigma_y$.

Тогда плотность вероятности составляющей X имеет следующий вид:

$$f_1(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-u^2/2}.$$

Для нахождения функции регрессии $M(Y/x)$ сначала вычислим условный закон распределения величины Y при $X = x$:

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

Если подставить

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} e^{-(u^2+v^2-2uv)/(2(1-r^2))}$$

и $u = (x - a_1)/\sigma_x$, $v = (y - a_2)/\sigma_y$ в правую часть этой формулы, затем выполнить преобразования, получим:

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

Заменяя u и v по формулам $u = (x - a_1)/\sigma_x$, $v = (y - a_2)/\sigma_y$, окончательно получим:

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

Условное распределение нормально с математическим ожиданием, т. е. функцией регрессии Y на X

$$M(Y/x) = a_2 + r\frac{\sigma_y}{\sigma_x}(x - a_1).$$

и дисперсией $\sigma_y^2(1-r^2)$.

Функция регрессии X и Y получается аналогичным способом:

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

Поскольку обе функции регрессии линейны, то корреляция между величинами X и Y линейная. Теорема доказана.

На основании вероятностного смысла параметров двумерного нормального распределения приходим к уравнению прямых регрессии:

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

которые совпадают с уравнениями прямых среднеквадратической регрессии.

ЛЕКЦИЯ №3. Случайные функции

1. Случайные функции

1. Определение случайной функции.

Основные задачи.

Выделяют два основных вида задач, решение которых требует использования теории случайных функций: анализа и синтеза.

Прямая задача (анализ). Здесь заданы параметры некоторого устройства и вероятностные характеристики (математические ожидания, корреляционные функции, законы распределения) поступающей на его вход функции (сигнала, процесса). Необходимо найти характеристики на выходе устройства (по ним судят о качестве работы устройства).

Обратная задача (синтез). В этом случае заданы вероятностные характеристики входной и выходной функций. Необходимо спроектировать оптимальное устройство (найти его параметры), которое будет осуществлять преобразование заданной входной функции в такую выходную функцию, которая имеет заданные характеристики. При решении этой задачи необходимо, кроме аппарата случайных функций, привлечение и других дисциплин.

Определение случайной функции.

Случайная функция — это функция неслучайного аргумента t , которая при каждом фиксированном значении аргумента является случайной величиной.

При этом случайные функции аргумента t обозначают прописными буквами $X(t)$, $Y(t)$ и т. д.

К примеру, если U является случайной величиной, то функция $X(t) = t^2 U$ тоже случайная. В самом деле, при каждом фиксированном значении аргумента эта функция является случайной величиной: при $t_1 = 3$ получим случайную величину $X_1 = 9U$, при $t_2 = 4$ — случайную величину $X_2 = 16U$ и т. д.

Сечение случайной функции есть случайная величина, соответствующая фиксированному значению аргумента случайной функции.

К примеру, для случайной функции $X(t) = t^2 U$ при значениях аргумента $t_1 = 3$ и $t_2 = 4$ были получены соответственно случайные величины $X_1 = 9U$ и $X_2 = 16U$, являющиеся сечениями заданной случайной функции.

Значит, случайную функцию можно рассматривать как совокупность случайных величин $\{X(t)\}$, которые зависят от параметра t .

Реализация (траектория выборочной функцией) случайной функции $X(t)$ — это неслучайная функция аргумента t , которой может быть равна случайная функция в результате испытания.

Следовательно, если в опыте наблюдают случайную функцию, то в действительности наблюдают одну из возможных ее реализаций. Значит, при повторении опыта будет наблюдаться другая реализация.

Реализации функции $X(t)$ пишут строчными буквами $x_1(t)$, $x_2(t)$ и так далее, а индекс свидетельствует о номере испытания. Например, пусть $X(t) = U \sin t$, а U есть непрерывная случайная величина, которая в первом испытании приняла возможное значение $u_1 = 3$, а во втором испытании $u_2 = 4,6$. Тогда реализациями $X(t)$ будут соответственно неслучайные функции $x_1(t) = 3 \sin t$ и $x_2(t) = 4,6 \sin t$.

Следовательно, случайную функцию можно рассматривать как совокупность ее возможных реализаций.

Случайный (стохастический) процесс есть случайная функция аргумента t , и истолковывается он как время, т. е. если самолет должен лететь с определенной постоянной скоростью, то на самом деле из-за воздействия случайных факторов (колебания температуры, изменения силы ветра и др.), предугадать влияние которых заранее невозможно, скорость изменяется. Тогда скорость самолета будет случайной функцией от непрерывно изменяющегося аргумента (времени), а значит, скорость есть случайный процесс.

Если аргумент случайной функции изменяется дискретно, то соответствующие ему значения случайной функции (случайные величины) образуют *случайную последовательность*.

Следует отметить, что аргументом случайной функции может быть не только время. Допустим, если измеряется диаметр ткацкой нити вдоль ее длины, то вследствие воздействия случайных факторов диаметр нити изменяется. Здесь диаметр есть случайная функция от непрерывно изменяющегося аргумента (длины нити).

Задание случайной функции аналитически не предоставляется возможным. В отдельных случаях, если известен вид случайной функции, а определяющие ее параметры являются случайными величинами, то задать ее аналитически можно.

К примеру, случайными могут быть функции:

$X(t) = \sin \Omega t$, где Ω — случайная величина; $X(t) = U \sin t$, где U — случайная величина; $X(t) = U \sin \Omega t$, где Ω и U — случайные величины.

2. Математическое ожидание случайной функции и его свойства.

Рассмотрим случайную функцию $X(t)$. При фиксированном значении аргумента, например при $t = t_1$, получим сечение — случайную величину $X(t_1)$ с математическим ожиданием $M[X(t_1)]$. (Полагаем, что математическое ожидание любого сечения существует.) Таким образом, каждое фиксированное значение аргумента определяет сечение — случайную величину, а каждой случайной величине соответствует ее математическое ожидание. Отсюда следует, что каждому фиксированному значению аргумента t соответствует определенное математическое ожидание; это означает, что математическое ожидание случайной функции есть функция (неслучайная) от аргумента t ; ее обозначают через $m_x(t)$. В частном случае функция $m_x(t)$ может сохранять постоянное значение при всех допустимых значениях аргумента. Дадим теперь определение математического ожидания.

Математическим ожиданием случайной функции $X(t)$ называют неслучайную функцию $m_x(t)$, значение которой при каждом фиксированном значении аргумента t равно математическому ожиданию сечения, соответствующего этому же фиксированному значению аргумента:

$$m_x(t) = M[X(t)].$$

Геометрически математическое ожидание случайной функции можно истолковать как среднюю кривую, около которой расположены другие кривые — реализации; при фиксированном значении аргумента математическое ожидание есть среднее значение сечения (средняя ордината), вокруг которого расположены его возможные значения (ординаты).

Свойства математического ожидания случайной функции. Используя свойства математического ожидания случайной величины, легко получить свойства математического ожидания случайной функции.

Свойство 1. Математическое ожидание неслучайной функции равно самой неслучайной функции:

$$M[\varphi(t)] = \varphi(t).$$

Свойство 2. Неслучайный множитель $\varphi(t)$ можно выносить за знак математического ожидания:

$$M[\varphi(t)X(t)] = \varphi(t)M[X(t)] = \varphi(t)m_x(t).$$

Свойство 3. Математическое ожидание суммы двух случайных функций равно сумме математических ожиданий слагаемых:

$$M[X(t) + Y(t)] = m_x(t) + m_y(t).$$

Следствие. Для того чтобы найти математическое ожидание суммы случайной и неслучайной функций, достаточно к математическому ожиданию случайной функции прибавить неслучайную функцию:

$$M[X(t) + \varphi(t)] = m_x(t) + \varphi(t).$$

3. Дисперсия случайной функции и ее свойства.

Рассмотрим случайную функцию $X(t)$. При фиксированном значении аргумента, например при $t = t_1$ получим сечение — случайную величину $X(t_1)$ с дисперсией $D[X(t_1)] \geq 0$. (Предполагается, что дисперсия любого сечения существует). Таким образом, каждое фиксированное значение аргумента определяет сечение — случайную величину, а каждой случайной величине соответствует ее дисперсия. Отсюда следует, что каждому фиксированному значению аргумента t соответствует определенная дисперсия; это означает, что дисперсия случайной функции есть функция (неслучайная, причем неотрицательная) от аргумента t ; ее обозначают через $D_x(t)$. В частном случае $D_x(t)$ может сохранять постоянное значение при всех допустимых значениях аргумента.

Дисперсией случайной функции $X(t)$ называют неслучайную неотрицательную функцию $D_x(t)$, значение которой при каждом фиксированном значении аргумента t равно дисперсии сечения, соответствующего этому же фиксированному значению аргумента:

$$D_x(t) = D[X(t)].$$

Дисперсия характеризует степень рассеяния возможных реализаций (кривых) вокруг математического ожидания случайной функции (средней кривой).

При фиксированном значении аргумента дисперсия характеризует степень рассеяния возможных значений (ординат) сечения вокруг математического ожидания сечения (средней ординаты).

Часто вместо дисперсии рассматривают среднее квадратическое отклонение случайной функции, которое определяют по аналогии со средним квадратическим отклонением случайной величины.

Средним квадратическим отклонением случайной функции называют квадратный корень из дисперсии:

$$\sigma_x(t) = \sqrt{D_x(t)} .$$

Свойства дисперсии случайной функции.

Используя свойства дисперсии случайной величины, легко получить свойства дисперсии случайной функции.

Свойство 1. Дисперсия неслучайной функции $\varphi(t)$ равна нулю:

$$D[\varphi(t)] = 0.$$

Свойство 2. Дисперсия суммы случайной функции $X(t)$ и неслучайной функции $\varphi(t)$ равна дисперсии случайной функции:

$$D[X(t) + \varphi(t)] = D_x(t).$$

Свойство 3. Дисперсия произведения случайной функции $X(t)$ на неслучайную функцию $\varphi(t)$ равна произведению квадрата неслучайного множителя на дисперсию случайной функции:

$$D[X(t)\varphi(t)] = \varphi^2(t)D_x(t).$$

4. Корреляционная функция случайной функции и ее свойства.

Корреляционной функцией случайной функции $X(t)$ называют неслучайную функцию $K_x(t_1, t_2)$ двух независимых аргументов t_1 и t_2 , значение которой при каждой паре фиксированных значений аргументов равно корреляционному моменту сечений, соответ-

ствующих этим же фиксированным значениям аргументов:

$$K_x(t_1, t_2) = M[\dot{X}(t_1) \dot{X}(t_2)].$$

Свойства корреляционной функции.

Свойство 1. При перестановке аргументов корреляционная функция не изменяется (свойство симметрии):

$$K_x(t_1, t_2) = K_x(t_2, t_1).$$

Доказательство. По определению корреляционной функции,

$$K_x(t_1, t_2) = M[\dot{X}(t_1) \dot{X}(t_2)].$$

$$K_x(t_2, t_1) = M[\dot{X}(t_2) \dot{X}(t_1)].$$

Правые части этих равенств равны (математическое ожидание произведения не зависит от порядка сомножителей), следовательно, равны и левые части. Итак,

$$K_x(t_1, t_2) = K_x(t_2, t_1).$$

Свойство 2. Прибавление к случайной функции $X(t)$ неслучайного слагаемого $\varphi(t)$ не изменяет ее корреляционной функции.

Если

$$Y(t) = X(t) + \varphi(t),$$

то

$$K_y(t_1, t_2) = K_x(t_1, t_2).$$

Доказательство.

$$\dot{Y}(t) = \dot{X}(t).$$

Отсюда $\dot{Y}(t_1) = \dot{X}(t_1)$ и $\dot{Y}(t_2) = \dot{X}(t_2)$. Следовательно,

$$M[\dot{Y}(t_1) \dot{Y}(t_2)] = M[\dot{X}(t_1) \dot{X}(t_2)].$$

Итак,

$$K_y(t_1, t_2) = K_x(t_1, t_2)].$$

Свойство 3. При умножении случайной функции $X(t)$ на неслучайный множитель $\varphi(t)$ ее корреляционная функция умножается на произведение $\varphi(t_1)\varphi(t_2)$.

Если

$$Y(t) = X(t)\varphi(t),$$

то

$$K_y(t_1, t_2) = K_x(t_1, t_2)]\varphi(t_1)\varphi(t_2).$$

Доказательство.

$$\dot{Y}(t) = \dot{X}(t).$$

Следовательно,

$$K_y(t_1, t_2) = M[\dot{Y}(t_1)\dot{Y}(t_2)] = M\{[\dot{X}(t_1)\varphi(t_1)][\dot{X}(t_2)\varphi(t_2)]\}.$$

Вынесем неслучайные множители за знак математического ожидания:

$$K_y(t_1, t_2) = \varphi(t_1)\varphi(t_2)M[\dot{X}(t_1)\dot{X}(t_2)] = \varphi(t_1)\varphi(t_2)K_x(t_1, t_2).$$

Итак,

$$K_y(t_1, t_2) = K_x(t_1, t_2)]\varphi(t_1)\varphi(t_2).$$

Свойство 4. Абсолютная величина корреляционной функции не превышает среднего геометрического дисперсий соответствующих сечений:

$$|K_x(t_1, t_2)| \leq \sqrt{D_x(t_1)D_x(t_2)}.$$

Доказательство. Известно, что для модуля корреляционного момента двух случайных величин справедливо неравенство:

$$|\mu_{xy}| \leq \sqrt{D_x D_y}.$$

При фиксированных значениях аргументов t_1 и t_2 значение корреляционной функции равно корреляционному моменту соответствующих сечений — случайных величин $X(t_1)$ и $X(t_2)$. Поэтому неравенство можно записать так:

$$|K_x(t_1, t_2)| \leq \sqrt{D_x(t_1) D_x(t_2)}.$$

5. Нормированная корреляционная функция.

Нормированной корреляционной функцией случайной функции $X(t)$ называют неслучайную функцию двух независимых переменных t_1 и t_2 , значение которой при каждой паре фиксированных значений аргументов равно коэффициенту корреляции сечений, соответствующих этим же фиксированным значениям аргументов:

$$\rho_x(t_1, t_2) = \frac{K_x(t_1, t_2)}{\sigma_x(t_1) \sigma_x(t_2)}.$$

Учитывая, что

$$\sigma_x(t_1) = \sqrt{D_x(t_1)} = \sqrt{K_x(t_1, t_1)} \quad \text{и} \quad \sigma_x(t_2) = \sqrt{K_x(t_2, t_2)},$$

получим

$$\rho_x(t_1, t_2) = \frac{K_x(t_1, t_2)}{\sqrt{K_x(t_1, t_1)} \sqrt{K_x(t_2, t_2)}}.$$

Таким образом, зная корреляционную функцию, можно найти нормированную корреляционную функцию.

Нормированная корреляционная функция имеет те же свойства, что и корреляционная функция, причем свойство 4 заменяется на следующее: абсолютная величина нормированной корреляционной функции не превышает единицы:

$$|\rho_x(t_1, t_2)| \leq 1.$$

Это свойство следует из того, что при фиксированных значениях аргументов значение нормированной корреляционной

функции равно коэффициенту корреляции двух случайных величин — соответствующих сечений, а абсолютная величина коэффициента корреляции не превышает единицы:

$$\rho_x(t_1, t_2) = \frac{K_x(t_1, t_2)}{\sigma_x(t_1)\sigma_x(t_2)} \quad \text{или} \quad \rho_x(t_1, t_2) = \frac{K_x(t_1, t_2)}{\sqrt{K_x(t_1, t_1)}\sqrt{K_x(t_2, t_2)}},$$

что при равных значениях аргументов нормированная корреляционная функция равна единице: $\rho_x(t, t) = 1$.

Очевидно, нормированная корреляционная функция имеет тот же вероятностный смысл, что и коэффициент корреляции: чем ближе модуль этой функции к единице, тем линейная связь между сечениями сильнее; чем ближе модуль этой функции к нулю, тем эта связь слабее.

6. Взаимная корреляционная функция и ее свойства.

Взаимной корреляционной функцией двух случайных функций $X(t)$ и $Y(t)$ называют неслучайную функцию $R_{xy}(t_1, t_2)$ двух независимых аргументов t_1 и t_2 , значение которой при каждой паре фиксированных значений аргументов равно корреляционному моменту сечений обеих функций, соответствующих этим же фиксированным значениям аргументов:

$$R_{xy}(t_1, t_2) = M[\dot{X}(t_1)\dot{Y}(t_2)].$$

Коррелированными называют две случайные функции, если их взаимная корреляционная функция не равна тождественно нулю.

Некоррелированными называют две случайные функции, взаимная корреляционная функция которых тождественно равна нулю.

Свойства взаимной корреляционной функции.

Свойство 1. При одновременной перестановке индексов и аргументов взаимная корреляционная функция не изменяется:

$$R_{xy}(t_1, t_2) = R_{yx}(t_2, t_1).$$

Свойство 2. Прибавление к случайным функциям $X(t)$ и $Y(t)$ неслучайных слагаемых, соответственно $\varphi(t)$ и $\psi(t)$, не изме-

няет их взаимной корреляционной функции.

Если

$$X_1(t) = X(t) + \varphi(t) \text{ и } Y_1(t) = Y(t) + \psi(t),$$

то

$$R_{x_1y_1}(t_1, t_2) = R_{yx}(t_1, t_2).$$

Свойство 3. При умножении случайных функций $X(t)$ и $Y(t)$ на неслучайные множители, соответственно $\varphi(t)$ и $\psi(t)$, взаимная корреляционная функция умножается на произведение $\varphi(t_1)\psi(t_2)$.

Если

$$X_1(t) = X(t)\varphi(t), \quad Y_1(t) = Y(t)\psi(t),$$

то

$$R_{x_1y_1}(t_1, t_2) = R_{xy}(t_1, t_2)\varphi(t_1)\psi(t_2).$$

Свойство 4. Абсолютная величина взаимной корреляционной функции двух случайных функций не превышает среднего геометрического их дисперсий:

$$|R_{xy}(t_1, t_2)| \leq \sqrt{D_x(t_1)D_y(t_2)}.$$

Доказательства этих свойств аналогичны доказательствам свойств корреляционной функции.

7. Нормированная взаимная корреляционная функция.

Наряду с взаимной корреляционной функцией для оценки степени зависимости сечений двух случайных функций пользуются характеристикой — нормированной взаимной корреляционной функцией.

Нормированной взаимной корреляционной функцией двух случайных функций $X(t)$ и $Y(t)$ называют неслучайную функцию двух независимых аргументов t_1 и t_2 :

$$\rho_{xy}(t_1, t_2) = \frac{R_{xy}(t_1, t_2)}{\sqrt{K_x(t_1, t_1)}\sqrt{K_x(t_2, t_2)}} = \frac{R_{xy}(t_1, t_2)}{\sqrt{D_x(t_1)}\sqrt{D_x(t_2)}}.$$

Нормированная взаимная корреляционная функция имеет те же свойства, что и взаимная корреляционная функция, причем свойство 4 заменяется следующим свойством: абсолютная величина нормированной взаимной корреляционной функции не превышает единицы:

$$|\rho_{xy}(t_1, t_2)| \leq 1.$$

8. Характеристики суммы случайных функций.

Пусть $X(t)$ и $Y(t)$ — случайные функции. Найдем характеристики суммы этих функций по известным характеристикам слагаемых.

Теорема 1. Математическое ожидание суммы двух случайных функций равно сумме математических ожиданий слагаемых.

Если

$$Z(t) = X(t) + Y(t),$$

то

$$m_z(t) = m_x(t) + m_y(t).$$

Теорема 2. Корреляционная функция суммы двух коррелированных случайных функций равна сумме корреляционных функций слагаемых и взаимной корреляционной функции, которая прибавляется дважды (с разным порядком следования аргументов).

Если

$$Z(t) = X(t) + Y(t),$$

то

$$K_z(t_1, t_2) = K_x(t_1, t_2) + K_y(t_1, t_2) + R_{xy}(t_1, t_2) + R_{xy}(t_2, t_1).$$

Доказательство. По определению корреляционной функции,

$$K_z(t_1, t_2) = M[\dot{Z}(t_1)\dot{Z}(t_2)].$$

В силу замечания 1 на стр 85

$$\dot{Z}(t) = \dot{X}(t) + \dot{Y}(t).$$

Следовательно,

$$\dot{Z}(t_1)\dot{Z}(t_2) = [\dot{X}(t_1) + \dot{Y}(t_1)][\dot{X}(t_2) + \dot{Y}(t_2)].$$

Выполнив умножение, приравняем математические ожидания обеих частей равенства:

$$M[\dot{Z}(t_1)\dot{Z}(t_2)] = M[\dot{X}(t_1)\dot{X}(t_2)] + M[\dot{Y}(t_1)\dot{Y}(t_2)] + M[\dot{X}(t_1)\dot{Y}(t_2)] + M[\dot{Y}(t_1)\dot{X}(t_2)].$$

По определению корреляционной и взаимной корреляционной функций, имеем:

$$K_z(t_1, t_2) = K_x(t_1, t_2) + K_y(t_1, t_2) + R_{xy}(t_1, t_2) + R_{yx}(t_1, t_2).$$

Учитывая, что $R_{yx}(t_1, t_2) = R_{xy}(t_2, t_1)$, окончательно получим

$$K_z(t_1, t_2) = K_x(t_1, t_2) + K_y(t_1, t_2) + R_{xy}(t_1, t_2) + R_{xy}(t_2, t_1).$$

Методом математической индукции теорему можно обобщить на n попарно коррелированных случайных функций.

Если

$$Z(t) = \sum_{i=1}^n X_i(t),$$

то

$$K_z(t_1, t_2) = \sum_{i=1}^n K_x(t_1, t_2) + \sum_{i \neq j} R_{x_i y_j}(t_1, t_2),$$

где индексы i, j второго слагаемого есть размещения чисел $1, 2, \dots, n$, взятых по два.

Итак, дисперсия суммы двух некоррелированных случайных функций равна сумме дисперсий слагаемых.

9. Производная случайной функции и ее характеристики.

При изучении случайных величин встречалось понятие сходимости по вероятности. Для изучения случайных функций необходимо ввести среднеквадратичную сходимость.

Говорят, что последовательность случайных величин X_1, X_2, \dots, X_n , сходится в среднеквадратичном к случайной величине X , если математическое ожидание квадрата разности $X_n - X$ стремится к нулю при $n \rightarrow \infty$:

$$M[(X_n - X)^2] = 0.$$

Случайную величину X называют *пределом в среднеквадратичном* последовательности случайных величин X_1, X_2, \dots, X_n , и пишут:

$$X = \lim X_n.$$

Заметим, что из среднеквадратичной сходимости следует сходимость по вероятности; обратное утверждение, вообще говоря, неверно.

Случайную функцию $X(t)$ называют *дифференцируемой*, если существует такая функция $X'(t)$ (ее называют производной), что

$$\lim_{\Delta t \rightarrow 0} M \left[\frac{X(t + \Delta t) - X(t)}{\Delta t} - X'(t) \right]^2 = 0.$$

Итак, *производной случайной функции* $X(t)$ называют среднеквадратичный предел отношения приращения функции к приращению аргумента Δt при $\Delta t \rightarrow 0$:

$$X'(t) = \lim_{\Delta t \rightarrow 0} \frac{X(t + \Delta t) - X(t)}{\Delta t}.$$

Пусть известны характеристики случайной функции. Как найти характеристики ее производной? Ответ на этот вопрос дают теоремы, приведенные ниже, причем рассматриваются только среднеквадратично дифференцируемые случайные функции.

Теорема 1. Математическое ожидание производной $X'(t) = x$ от случайной функции $X(t)$ равно производной от ее математического ожидания:

$$m_x'(t) = m_x'(t).$$

Доказательство. По определению производной,

$$X'(t) = \lim_{\Delta t \rightarrow 0} \frac{X(t + \Delta t) - X(t)}{\Delta t}.$$

Приравняем математические ожидания обеих частей равенства, а затем изменим порядок нахождения математического ожидания и предела (законность изменения порядка этих операций примем без доказательства):

$$M[X'(t)] = \lim_{\Delta t \rightarrow 0} M \left[\frac{X(t + \Delta t) - X(t)}{\Delta t} \right].$$

Используя свойства математического ожидания, получим:

$$M[X'(t)] = \lim_{\Delta t \rightarrow 0} \frac{m_x(t + \Delta t) - m_x(t)}{\Delta t} = m'_x(t).$$

Итак, $m_x^\bullet(t) = m'_x(t)$.

Теорема 2. Корреляционная функция производной от случайной функции $X(t)$ равна второй смешанной производной от ее корреляционной функции:

$$K_x^\bullet(t_1, t_2) = \frac{\partial^2 K_x(t_1, t_2)}{\partial t_1 \partial t_2}.$$

Доказательство. По определению корреляционной функции,

$$K_x^\bullet(t_1, t_2) = M \left[\dot{X}'(t_1) \dot{X}'(t_2) \right].$$

Представим произведение производных как вторую смешанную частную производную:

$$\dot{X}'(t_1) \dot{X}'(t_2) = \frac{\partial^2 [\dot{X}(t_1) \dot{X}(t_2)]}{\partial t_1 \partial t_2}.$$

Следовательно,

$$K_x^\bullet(t_1, t_2) = M \left\{ \frac{\partial^2 [\dot{X}(t_1) \dot{X}(t_2)]}{\partial t_1 \partial t_2} \right\}.$$

Изменив порядок операций нахождения математического ожидания и дифференцирования, окончательно получим:

$$K_{\dot{x}}(t_1, t_2) = \frac{\partial^2 M[\dot{X}(t_1)\dot{X}(t_2)]}{\partial t_1 \partial t_2} = \frac{\partial^2 K_x(t_1, t_2)}{\partial t_1 \partial t_2}.$$

Итак,

$$K_{\dot{x}}(t_1, t_2) = \frac{\partial^2 K_x(t_1, t_2)}{\partial t_1 \partial t_2}.$$

Теорема 3. Взаимная корреляционная функция случайной функции $X(t)$ и ее производной $\dot{X}(t) = \dot{x}$ равна частной производной от корреляционной функции по соответствующему аргументу (если индекс \dot{x} при R записан на первом (втором) месте, то дифференцируют по первому (второму) аргументу):

$$1) R_{\dot{x}x}(t_1, t_2) = \frac{\partial K_x(t_1, t_2)}{\partial t_2};$$

$$2) R_{xx\dot{x}}(t_1, t_2) = \frac{\partial K_x(t_1, t_2)}{\partial t_1}.$$

10. Интеграл от случайной функции и его характеристики.

Интегралом от случайной функции $X(t)$ по отрезку $[0, t]$ называют предел в среднеквадратическом интегральной суммы при стремлении к нулю частичного интервала Δs_i максимальной длины (переменная интегрирования обозначена через s , чтобы отличить ее от предела интегрирования t):

$$Y(t) = \lim_{\Delta s_i \rightarrow 0} \sum X(s_i) \Delta s_i = \int_0^t X(s) ds.$$

Пусть известны характеристики случайной функции. Как найти характеристики интеграла от случайной функции? Ответ на этот вопрос дают теоремы, приведенные ниже.

Теорема 1. Математическое ожидание интеграла от случайной функции равно интегралу от ее математического ожидания.

Если

$$Y(t) = \int_0^t X(s)ds,$$

то

$$m_y(t) = \int_0^t m_x(s)ds.$$

Доказательство. По определению интеграла,

$$Y(t) = \lim_{\Delta s_i \rightarrow 0} \sum X(s_i)\Delta s_i.$$

Приравняем математические ожидания обеих частей равенства:

$$M[Y(t)] = M[\lim_{\Delta s_i \rightarrow 0} \sum X(s_i)\Delta s_i].$$

Изменим порядок нахождения математического ожидания и предела (законность изменения порядка этих операций примем без доказательства):

$$M[Y(t)] = \lim_{\Delta s_i \rightarrow 0} [M \sum X(s_i)\Delta s_i].$$

Воспользуемся теоремой сложения математических ожиданий:

$$M[Y(t)] = \lim_{\Delta s_i \rightarrow 0} \sum m_x(s_i)\Delta s_i.$$

Учитывая, что $\sum m_x(s_i)\Delta s_i$ — интегральная сумма функции $m_x(s)$, окончательно получим:

$$m_y(t) = \int_0^t m_x(s)ds.$$

Теорема 2. Корреляционная функция интеграла от случайной функции $X(t)$ равна двойному интегралу от ее корреляционной функции.

Если

$$Y(t) = \int_0^t X(s) ds,$$

то

$$K_y(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} K_x(s_1, s_2) ds_1 ds_2.$$

Доказательство. По определению корреляционной функции,

$$K_y(t_1, t_2) = M[\overset{\circ}{Y}(t_1)\overset{\circ}{Y}(t_2)].$$

Центрированная случайная функция:

$$\begin{aligned} \overset{\circ}{Y}(t) &= Y(t) - m_y(t) = \int_0^t X(s) ds - \int_0^t m_x(s) ds = \\ &= \int_0^t [X(s) - m_x(s)] ds \end{aligned}$$

или

$$\overset{\circ}{Y}(t) = \int_0^t \overset{\circ}{X}(s) ds.$$

Поскольку под знаком определенного интеграла переменную интегрирования можно обозначать любой буквой, обозначим переменную интегрирования в одном интеграле через s_1 , а в другом — через s_2 (чтобы отличить переменные интегрирования и пределы интегрирования):

$$\overset{\circ}{Y}(t_1) = \int_0^{t_1} \overset{\circ}{X}(s_1) ds_1, \quad \overset{\circ}{Y}(t_2) = \int_0^{t_2} \overset{\circ}{X}(s_2) ds_2.$$

Следовательно,

$$\begin{aligned} \overset{\circ}{Y}(t_1)\overset{\circ}{Y}(t_2) &= \int_0^{t_1} \overset{\circ}{X}(s_1)ds_1 \int_0^{t_2} \overset{\circ}{X}(s_2)ds_2 = \\ &= \int_0^{t_1} \int_0^{t_2} \overset{\circ}{X}(s_1)\overset{\circ}{X}(s_2)ds_1ds_2. \end{aligned}$$

Приравняем математические ожидания обеих частей равенства:

$$M[\overset{\circ}{Y}(t_1)\overset{\circ}{Y}(t_2)] = M\left[\int_0^{t_1} \int_0^{t_2} \overset{\circ}{X}(s_1)\overset{\circ}{X}(s_2)ds_1ds_2\right].$$

Изменив порядок операций нахождения математического ожидания и интегрирования, окончательно получим:

$$K_y(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} K_x(s_1, s_2)ds_1ds_2.$$

Теорема 3. Взаимная корреляционная функция случайной функции $X(t)$ и интеграла

$$Y(t) = \int_0^t X(s)ds$$

равна интегралу от корреляционной функции случайной функции $X(t)$:

$$1) \quad R_{xy}(t_1, t_2) = \int_0^{t_2} K_x(t_1, s)ds;$$

$$2) \quad R_{xy}(t_1, t_2) = \int_0^{t_1} K_x(s, t_2)ds.$$

Доказательство.

1. По определению взаимной корреляционной функции,

$$R_{xy}(t_1, t_2) = M[\overset{\circ}{X}(t_1)\overset{\circ}{Y}(t_2)].$$

В силу соотношения

$$\overset{o}{Y}(t) = \int_0^t \overset{o}{X}(s) ds \quad \text{центрированная функция}$$

$$\overset{o}{Y}(t) = \int_0^t \overset{o}{X}(s) ds,$$

следовательно,

$$\overset{o}{Y}(t_2) = \int_0^{t_2} \overset{o}{X}(s) ds.$$

Подставим правую часть этого равенства

$$\text{в } R_{xy}(t_1, t_2) = M[\overset{o}{X}(t_1)\overset{o}{Y}(t_2)] :$$

$$R_{xy}(t_1, t_2) = M \left[\overset{o}{X}(t_1) \int_0^{t_2} \overset{o}{X}(s) ds \right] = M \left[\int_0^{t_2} \overset{o}{X}(t_1) \overset{o}{X}(s) ds \right].$$

Операции нахождения математического ожидания и интегрирования можно менять местами, поэтому

$$R_{xy}(t_1, t_2) = M \int_0^{t_2} \overset{o}{X}(t_1) \overset{o}{X}(s) ds,$$

и окончательно

$$R_{xy}(t_1, t_2) = \int_0^{t_2} K_x(t_1, s) ds.$$

2. Доказывается аналогично.

11. Комплексные случайные величины и их числовые характеристики.

Комплексной случайной величиной называют величину $Z = X + Y_i$, где X и Y — действительные случайные величины.

Сопряженной случайной величине $Z = X + Y_i$ называют случайную величину $\bar{Z} = X - Y_i$

Обобщим определения математического ожидания и дисперсии на комплексные случайные величины так, чтобы, в частности, при $Y = 0$ эти характеристики совпали с ранее введенными

характеристиками действительных случайных величин, т. е. чтобы выполнялись требования:

$$\begin{aligned} m_z &= m_x; \\ D_z &= D_x. \end{aligned}$$

Математическим ожиданием комплексной случайной величины $Z = X + Y_i$ называют комплексное число

$$m_z = m_x + m_y i.$$

В частности, при $y = 0$ получим $m_z = m_x$, т. е. требование $m_z = m_x$ выполняется.

Дисперсией комплексной случайной величины Z :

$$D_z = M \left[\left| \overset{\circ}{Z} \right|^2 \right].$$

В частности, при $Y = 0$ получим $D_z = M \left[\left| \overset{\circ}{X} \right|^2 \right] = D_x$, т. е.

требование $D_z = D_x$ выполняется.

Учитывая, что математическое ожидание суммы равно сумме математических ожиданий слагаемых, имеем:

$$\begin{aligned} D_z &= M \left[\left| \overset{\circ}{X} \right|^2 \right] = M[(\overset{\circ}{X})^2 + (\overset{\circ}{Y})^2] = \\ &= M[(\overset{\circ}{X})^2] + M[(\overset{\circ}{Y})^2] = D_x + D_y. \end{aligned}$$

Итак, дисперсия комплексной случайной величины равна сумме дисперсий ее действительной и мнимой частей:

$$D_z = D_x + D_y.$$

Известно, что корреляционный момент двух равных случайных величин $X_1 = X_2 = X$ равен дисперсии D_x — положительному

действительному числу. Обобщим определение корреляционного момента так, чтобы, в частности, корреляционный момент двух равных комплексных случайных величин $Z_1 = Z_2 = Z$ был равен дисперсии D_z — положительному действительному числу, т. е. чтобы выполнялось требование

$$\mu_{zz} = D_z.$$

Корреляционным моментом двух комплексных случайных величин называют математическое ожидание произведения отклонения одной из величин на сопряженное отклонение другой:

$$\mu_{z_1 z_2} = M[(Z_1 - m_{z_1})\overline{(Z_2 - m_{z_2})}] = M[\overset{o}{Z}_1 \overset{o}{Z}_2].$$

В частности, при $Z_1 = Z_2 = Z$, учитывая, что произведение сопряженных комплексных чисел равно квадрату их модуля, получим:

$$\mu_{zz} = M[\overset{o}{Z} \overline{\overset{o}{Z}}] = M\left[\left|\overset{o}{Z}\right|^2\right] = D_z,$$

т. е. требование $\mu_{zz} = D_z$ выполняется.

Корреляционный момент комплексных случайных величин $Z_1 = X_1 + Y_1 i$ и $Z_2 = X_2 + Y_2 i$ выражается через корреляционные моменты действительных и мнимых частей этих величин следующей формулой:

$$\mu_{z_1 z_2} = \mu_{x_1 y_1} + \mu_{x_2 y_2} + (\mu_{x_2 y_1} - \mu_{x_1 y_2})i.$$

2. Стационарные случайные функции

1. Определение стационарной случайной функции.

Как правило, среди случайных функций выделяют класс функций, математические ожидания которых сохраняют одно и то же постоянное значение при всех значениях аргумента t и корреляционные функции которых зависят только от разности аргументов $t_2 - t_1$.

Вполне очевидно, что для таких функций начало отсчета аргумента может быть выбрано произвольно.

Стационарная функция — это случайная функция $X(t)$, математическое ожидание которой постоянно при всех значениях аргумента t и корреляционная функция которой зависит только от разности аргументов $t_2 - t_1$. Следовательно, можно сделать вывод:

- 1) корреляционная функция стационарной случайной функции является функцией одного аргумента $\tau = t_2 - t_1$, т. е.

$$K_x(t_1, t_2) = k_x(t_2 - t_1) = k_x(\tau);$$

- 2) дисперсия стационарной случайной функции постоянна при всех значениях аргумента t и равна значению ее корреляционной функции в начале координат ($\tau = 0$):

$$D_x(t) = K_x(t, t) = k_x(t - t) = k_x(0).$$

2. Свойства корреляционной функции стационарной случайной функции.

Свойство 1. Корреляционная функция стационарной случайной функции является четной функцией:

$$k_x(\tau) = k_x(-\tau).$$

Доказательство. При перестановке аргументов корреляционная функция любой случайной функции не изменяется. В частности, для стационарной функции

$$k_x(t_2 - t_1) = k_x(t_1 - t_2).$$

Пусть $\tau = t_2 - t_1$, тогда получится

$$k_x(\tau) = k_x(-\tau).$$

Свойство 2. Абсолютная величина корреляционной функции стационарной случайной функции не может быть больше ее значения в начале координат:

$$|k_x(\tau)| \leq k_x(0).$$

Доказательство. Для любой случайной функции

$$|K_x(t_1, t_2)| \leq \sqrt{D_x(t_1)D_x(t_2)}.$$

Тогда для стационарной функции

$$K_x(t_1, t_2) = k_x(\tau) \text{ и } D_x(t_1) = D_x(t_2) = k_x(0).$$

Таким образом,

$$|k_x(\tau)| \leq \sqrt{k_x(0)k_x(0)} = k_x(0).$$

3. Нормированная корреляционная функция стационарной случайной функции.

Нормированную корреляционную функцию можно определить так:

$$\rho_x(t_1, t_2) = \frac{K_x(t_1, t_2)}{\sigma_x(t_1)\sigma_x(t_2)}.$$

Для стационарной функции числитель и знаменатель этой дроби записываются в виде

$$K_x(t_1, t_2) = k_x(\tau), \quad \sigma_x(t) = \sqrt{D_x(t)} = \sqrt{k_x(0)}.$$

Значит, для стационарной функции правая часть данной формулы равна $k_x(\tau)/k_x(0)$ и является функцией одного аргумента τ . Тогда и левая часть той формулы есть функция от τ .

Нормированная корреляционная функция стационарной случайной функции — это неслучайная функция аргумента τ :

$$\rho_x(\tau) = k_x(\tau)/k_x(0).$$

Абсолютная величина нормированной корреляционной функции стационарной случайной функции не должна быть больше единицы.

Поскольку абсолютная величина частного равняется частному абсолютных величин, то можно прийти к выводу:

$$|\rho_x(\tau)| = |k_x(\tau)/k_x(0)| = |k_x(\tau)|/|k_x(0)|.$$

Так как $|k_x(\tau)| \leq |k_x(0)|$, то окончательно получим:

$$|\rho_x(\tau)| \leq 1.$$

Замечание. При $\tau = 0$ нормированная корреляционная функция равна единице:

$$\rho_x(0) = k_x(0) / k_x(0) = 1.$$

4. Стационарно связанные случайные функции.

Две случайные функции $X(t)$ и $Y(t)$ называют *стационарно связанными*, если их взаимная корреляционная функция зависит только от разности аргументов $\tau = t_2 - t_1$:

$$R_{xy}(t_1, t_2) = r_{xy}(\tau).$$

Взаимная корреляционная функция стационарно связанных случайных функций имеет следующее свойство:

$$r_{xy}(\tau) = r_{yx}(-\tau).$$

Данное равенство вытекает из свойства 1 взаимной корреляционной функции, которое гласит: при одновременной перестановке индексов и аргументов взаимная корреляционная функция не изменяется:

$$r_{xy}(t_2 - t_1) = r_{yx}(t_1 - t_2) \text{ или } r_{xy}(\tau) = r_{yx}(-\tau).$$

Это свойство геометрически можно объяснить следующим образом: график кривой $r_{yx}(-\tau)$ симметричен графику кривой $r_{xy}(\tau)$ относительно оси ординат.

Следует отметить, что если каждая из двух случайных функций стационарна, то еще нельзя говорить о том, что их взаимная корреляционная функция зависит лишь от разности аргументов.

Стационарные и стационарно связанные функции — это две стационарные случайные функции $X(t)$ и $Y(t)$, у которых взаимная корреляционная функция зависит лишь от разности аргументов $\tau = t_2 - t_1$.

5. Корреляционная функция производной стационарной случайной функции.

Теорема. Корреляционная функция производной $X'(t) = \dot{X}(t)$ дифференцируемой стационарной случайной функции $X(t)$ равняется второй производной от ее корреляционной функции, взятой со знаком минус:

$$k_x(\tau) = -k_x''(\tau).$$

Доказательство. Корреляционная функция производной любой дифференцируемой случайной функции равняется второй смешанной производной от ее корреляционной функции:

$$K_x(t_1, t_2) = \frac{\partial^2 K_x(t_1, t_2)}{\partial t_1 \partial t_2}.$$

Поскольку $X(t)$ есть стационарная функция, то ее корреляционная функция зависит только от разности аргументов:

$$K_x(t_1, t_2) = k_x(\tau).$$

Так как $\tau = t_2 - t_1$, то

$$\frac{\partial \tau}{\partial t_1} = -1 \quad \text{и} \quad \frac{\partial \tau}{\partial t_2} = 1.$$

Принимая во внимание эти равенства, получим:

$$\begin{aligned} K_x(t_1, t_2) &= \frac{\partial^2 k_x(\tau)}{\partial t_1 \partial t_2} = \frac{\partial}{\partial t_1} \left[\frac{\partial k_x(\tau)}{\partial t_2} \right] = \frac{\partial}{\partial t_1} \left[\frac{dk_x(\tau)}{d\tau} \frac{\partial \tau}{\partial t_2} \right] = \\ &= \frac{d^2 k_x(\tau)}{d\tau^2} \frac{\partial \tau}{\partial t_1} = k_x''(\tau) \times (-1) = -k_x''(\tau). \end{aligned}$$

Следовательно, искомая корреляционная функция зависит лишь от τ , значит $K_x(t_1, t_2) = k_x(\tau)$.

Таким образом,

$$k_x(\tau) = -k_x''(\tau).$$

6. Взаимная корреляционная функция стационарной случайной функции и ее производной.

Теорема. Взаимная корреляционная функция дифференцируемой стационарной случайной функции

$$X'(t) = \dot{x}$$

и ее производной равняется первой производной от корреляционной функции $k_x(\tau)$, которая берется со своим (противоположным) знаком, если индекс \dot{x} стоит на втором (первом) по порядку месте:

$$r_{xx}(\tau) = k_x'(\tau), \quad r_{\dot{x}x}(\tau) = -k_x'(\tau).$$

При этом $\tau = t_2 - t_1$.

Доказательство.

1. На основании определения взаимной корреляционной функции,

$$R_{xx}(\tau) = M[\dot{X}(t_1) \dot{X}'(t_2)] = M \left\{ \frac{\partial[\dot{X}(t_1) \dot{X}(t_2)]}{\partial t_2} \right\}.$$

В данном случае можно переставить операции нахождения математического ожидания и дифференцирования, поэтому

$$R_{xx}(\tau) = \frac{\partial M[\dot{X}(t_1) \dot{X}(t_2)]}{\partial t_2} = \frac{\partial[K_x(t_1, t_2)]}{\partial t_2}.$$

Поскольку $X(t)$ является стационарной функцией, то ее корреляционная функция зависит лишь от разности аргументов: $K_x(t_1, t_2) = k_x(\tau)$, $\tau = t_2 - t_1$ и, значит,

$$\frac{\partial \tau}{\partial t_2} = 1.$$

Следовательно,

$$R_{xx}(t_1, t_2) = \frac{\partial k_x(\tau)}{\partial t_2} = \frac{dk_x(\tau)}{d\tau} \frac{\partial \tau}{\partial t_2} = k'_x(\tau) \times 1 = k'_x(\tau).$$

В данном равенстве его правая часть зависит лишь от τ , а значит, и левая часть есть функция от τ . Если обозначить ее через $r_{xx}(\tau)$, то получится

$$r_{xx}(\tau) = k'_x(\tau).$$

2. Здесь доказательство проводится аналогично.

Замечание. Так как взаимная корреляционная функция $r_{xx}(\tau)$ зависит лишь от τ , то стационарная случайная функция и ее производная стационарно связаны.

7. Корреляционная функция интеграла от стационарной случайной функции.

Теорема. Корреляционная функция интеграла от стационарной случайной функции равняется:

$$Y(t) = \int_0^t X(s) ds.$$

Доказательство. Известно, что корреляционная функция интеграла

$$Y(t) = \int_0^t X(s) ds$$

от случайной функции $X(t)$ равняется двойному интегралу от ее корреляционной функции:

$$K_y(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} K_x(s_1, s_2) ds_1 ds_2.$$

Поскольку корреляционная функция стационарной случайной функции зависит только от разности аргументов, т. е.

$$K_x(s_1, s_2) = k_x(s_2 - s_1)$$

то получится

$$K_y(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} k_x(s_2 - s_1) ds_1 ds_2.$$

Следствие. Дисперсия интеграла

$$Y(t) = \int_0^t X(s) ds$$

от стационарной случайной функции равна:

$$D_y(t) = 2 \int_0^t (t - \tau) k_x(\tau) d\tau.$$

Если во второй по счету формуле этого вопроса $t_1 = t_2 = t$, то получится:

$$\begin{aligned} K_y(t, t) &= \int_0^t (t - \tau) k_x(\tau) d\tau - \\ &- \int_0^0 (t - t - \tau) d\tau + \int_0^t (t - \tau) k_x(\tau) d\tau. \end{aligned}$$

Приведем подобные члены и получим:

$$D_y(t) = 2 \int_0^t (t - \tau) k_x(\tau) d\tau.$$

8. Определение характеристик эргодических стационарных случайных функций из опыта.

Среди ряда стационарных случайных функций можно выделить отдельный класс функций. Его оценка характеристик путем усреднения множества реализаций равнозначна усреднению по времени только одной реализации достаточно большой длительности.

Эргодической называют стационарную случайную функцию $X(t)$, если ее характеристики, найденные усреднением множества реализаций, совпадают с соответствующими характеристиками,

которые, в свою очередь, были получены усреднением по времени одной реализации $x(t)$. Эта реализация наблюдалась на интервале $(0, T)$ достаточно большой длительности.

Достаточное условие эргодичности стационарной случайной функции $X(t)$ относительно математического ожидания заключается в следующем: ее корреляционная при $\tau \rightarrow \infty$ функция $k_x(\tau)$ стремится к нулю:

$$\lim_{\tau \rightarrow \infty} k_x(\tau) = 0.$$

Достаточное условие эргодичности стационарной случайной функции $X(t)$ относительно корреляционной функции заключается в следующем: при $\tau \rightarrow \infty$ корреляционная функция $k_x(\tau)$ стремится к нулю:

$$\lim_{\tau \rightarrow \infty} k_x(\tau) = 0$$

где

$$Y(t, \tau) = \overset{\circ}{X}(t) \overset{\circ}{X}(t + \tau).$$

Для оценки математического ожидания эргодической стационарной случайной функции $X(t)$ по наблюдавшейся на интервале $(0, T)$ реализации $x(t)$ принимается среднее по времени ее значение:

$$m_x^* = \frac{1}{T} \int_0^T x(t) dt.$$

Корреляционная функция стационарной случайной функции:

$$k_x(\tau) = M[\overset{\circ}{X}(t) \overset{\circ}{X}(t + \tau)].$$

Следовательно, оценить $k_x(\tau)$ означает оценить математическое ожидание функции

$$\overset{\circ}{X}(t) \overset{\circ}{X}(t + \tau),$$

поэтому можно воспользоваться соотношением для m_x^* . При этом нужно помнить, что функция $\overset{\circ}{X}(t + \tau)$ определена при $t + \tau \leq T$, а значит, $t \leq T - \tau$.

Таким образом, в качестве оценки корреляционной функции эргодической стационарной случайной функции принимают:

$$k_x^*(\tau) = \frac{1}{T-\tau} \int_0^{T-\tau} x(t)x(t+\tau)dt$$

или

$$k_x^*(\tau) = \frac{1}{T-\tau} \int_0^{T-\tau} x(t)x(t+\tau)dt - [m_x^*]^2.$$

3. Элементы математической статистики

1. Вариационный ряд.

Пусть изучается некоторая случайная величина x . Для этого проводится ряд независимых опытов или *наблюдений*, в каждом из которых величина x принимает то или иное значение.

Совокупность полученных значений

$$x_1, x_2, \dots, x_n$$

величины x (где n — число опытов) является произведенной выборкой. Эта совокупность называется *статистическим рядом*, который играет роль исходного числового материала, подлежащего дальнейшей обработке и анализу.

Первый этап обработки ряда x_1, x_2, \dots, x_n — составление *вариационного ряда*. Его получают следующим образом: среди чисел x_1, x_2, \dots, x_n отбирают все различные и располагают их в порядке возрастания:

$$\alpha_1, \alpha_2, \dots, \alpha_m,$$

где $\alpha_1 < \alpha_2 < \dots < \alpha_m$.

Пример 1. По журналу посещаемости собраны данные о числе пропущенных занятий по физике (за один семестр) у 25 студентов II курса, в результате чего получены значения:

2, 5, 0, 1, 6, 3, 0, 1, 5, 4, 0, 3, 3, 2, 1, 4, 0, 0, 2, 3, 6, 0, 3, 0, 1.

Вариационный ряд состоит из шести различных чисел:

0, 1, 2, 3, 4, 5, 6.

Следующий этап обработки ряда x_1, x_2, \dots, x_n — составление *эмпирического закона распределения*. Форма его записи зависит от характера изучаемой случайной величины x .

Пусть x — *дискретная* случайная величина, тогда наиболее естественной формой эмпирического закона распределения является *таблица частот* (табл. 2.), которая показывает, с какой частотой наблюдалось то или иное значение. Таблица частот имеет вид:

Таблица частот

Таблица 2

α_1	α_2	...	α_m
μ_1	μ_2	...	μ_m

В первой строке представлены числа вариационного ряда $\alpha_1, \alpha_2, \dots, \alpha_m$, а во второй — их частоты, т. е. числа:

$$\mu_i = k_i / n,$$

где n — число всех опытов;

k_i — число опытов, в которых наступало событие $x = \alpha_i$.

Поэтому с увеличением числа n опытов таблица частот будет все более приближаться к истинному закону распределения величины x .

Для примера 1 запишем таблицу частот. Просматривая исходный статистический ряд

2, 5, 0, 1, 6, 3, 0, 1, 5, 4, 0, 3, 3, 2, 1, 4, 0, 0, 2, 3, 6, 0, 3, 0, 1,

находим частоту появления каждого из чисел 0, 1, 2, 3, 4, 5, 6 (чисел вариационного ряда). Число 0 встречается 7 раз, значит, его частота равна $7/25$. Аналогично запишутся остальные частоты в таблице частот (табл. 3):

Таблица частот

Таблица №3

0	1	2	3	4	5	6
7/25	4/25	3/25	5/25	2/25	2/25	2/25

Рассмотрим теперь другой крайний случай, когда величина x является *непрерывной* случайной величиной. В этом случае пользоваться таблицей частот в прежнем ее виде уже не имеет смысла, так как она мало показательна для данного распределения. Характерной чертой непрерывного распределения является тот факт, что вероятность каждого отдельного значения равна нулю. Значит, в ряду x_1, x_2, \dots, x_n не будет повторений, и таблица частот примет вид (табл. 4):

Таблица частот

Таблица 4

x_1	x_2	...	x_n
$1/n$	$1/n$...	$1/n$

Такая таблица дает малую информацию о распределении значений величины x .

Эмпирический закон распределения задают в этом случае по-другому и записывают с помощью *интервальной таблицы частот* (табл. 5):

Интервальная таблица частот

Таблица №5

$c_1; c_2$	$c_1; c_2$...	$c_j; c_{j+1}$
μ_1	μ_2	...	μ_j

Эта запись означает, что весь диапазон изменения величины x разбит на интервалы (границами i -го интервала являются c_i и c_{i+1}); число $\tilde{\mu}_i$ ($i = 1, 2, \dots, l$) есть частота попадания в j -й интервал, т. е.

$$\tilde{\mu}_i = \frac{\tilde{k}_i}{n},$$

где \tilde{k}_i — количество чисел исходном ряду x_1, x_2, \dots, x_n , приходящихся на i -й интервал.

На практике число интервалов выбирается обычно в пределах одного-двух десятков.

2. Оценка неизвестной вероятности по частоте.

Пусть в результате n -кратного повторения данного опыта наступило k раз событие A , тогда частота наступления A оказалась равной:

$$\mu = k/n.$$

Необходимо по известной частоте μ оценить неизвестную вероятность p события A .

Пусть случайная величина x есть число наступлений события A в одном опыте, закон распределения которой имеет вид (табл. 6):

Закон распределения случайной величины X

Таблица № 6

Значения x	0	1
Вероятности	$1 - p$	p

При этом p — неизвестное число, выступающее в качестве параметра. Тогда

$$M[x] = 0 \times (1-p) + 1 \times p = p,$$

$$D[x] = (0-p)^2 (1-p) + (1-p)^2 p = p(1-p).$$

Отсюда видно, что параметр p является математическим ожиданием величины x .

Точечной оценкой для математического ожидания является эмпирическое среднее:

$$\tilde{p} = \frac{1}{n} \sum_{i=1}^n x_i.$$

В данном случае $\sum_{i=1}^n x_i$ — число наступлений события A в n опытах, тогда \tilde{p} — это частота наступления события A в n опытах, т. е. μ .

Величину \tilde{p} при больших значениях n можно считать распределенной приближенно нормально, что вытекает из центральной

предельной теоремы, примененной к сумме одинаково распределенных независимых случайных величин x_1, x_2, \dots, x_n .

$$\sum_{i=1}^n x_i$$

Тогда в качестве доверительного интервала для искомой вероятности p можно принять интервал

$$|\tilde{p} - p| < t_\alpha \frac{\sigma[x]}{\sqrt{n}}.$$

Возводя в квадрат обе части этого неравенства и учитывая

$$\frac{\sigma[x]}{\sqrt{n}} = \sqrt{\frac{p(1-p)}{n}},$$

можно записать доверительный интервал для p :

$$(\mu - p) < \frac{t_\alpha}{n} p(1-p),$$

где t_α определяется из условия $2\phi(t_\alpha) = \alpha$. Относительно p соотношение x_1, x_2, \dots, x_n есть квадратное неравенство, при решении которого можно выяснить, что доверительный интервал для p есть

$$p_1 < p < p_2,$$

где p_1 и p_2 — корни квадратного уравнения, которое получается из x_1, x_2, \dots, x_n путем замены знака $<$ на знак равенства. Выражения для p_1 и p_2 имеют громоздкий вид:

$$p_1 = \frac{\mu + \frac{1}{2} \frac{t_\alpha^2}{n} - t_\alpha \sqrt{\frac{\mu(1-\mu)}{n} + \frac{1}{4} \left(\frac{t_\alpha^2}{n}\right)^2}}{1 + \frac{t_\alpha^2}{n}};$$

$$p_2 = \frac{\mu + \frac{1}{2} \frac{t_\alpha^2}{n} + t_\alpha \sqrt{\frac{\mu(1-\mu)}{n} + \frac{1}{4} \left(\frac{t_\alpha^2}{n}\right)^2}}{1 + \frac{t_\alpha^2}{n}}.$$

При больших n эти выражения заменяют более простыми:

$$p_1 \approx \frac{\mu + \frac{1}{2} \frac{t_\alpha^2}{n} - t_\alpha \sqrt{\frac{\mu(1-\mu)}{n}}}{1 + \frac{t_\alpha^2}{n}};$$

$$p_2 \approx \frac{\mu + \frac{1}{2} \frac{t_\alpha^2}{n} + t_\alpha \sqrt{\frac{\mu(1-\mu)}{n}}}{1 + \frac{t_\alpha^2}{n}}.$$

3. Корреляция.

Пусть в некотором опыте наблюдаются две случайные величины x и y .

Поскольку x и y обусловлены одним и тем же опытом, то это означает, что x и y *скоррелированы* (согласованы) друг с другом.

Одной из характеристик корреляции является корреляционный момент

$$K[x, y] = M[x, y] = M[(x - m_x)(y - m_y)],$$

где m_x и m_y — математические ожидания величин x и y .

Справедлива формула:

$$K[x, y] = M[xy] - m_x m_y.$$

Для получения этой формулы надо записать

$$(x - m_x)(y - m_y) = xy - m_x y - m_y x + m_x m_y$$

и приравнять друг к другу математические ожидания левой и правой частей.

Если величины x и y независимы, то их корреляционный момент равен нулю, поэтому неравенство нулю величины $K[x, y]$ указывает на наличие связи между x и y .

Некоррелированными называются случайные величины x и y , для которых корреляционный момент равен нулю. Таким образом, из независимости величин x и y следует их некоррелированность, обратное неверно.

Корреляционный момент зависит от выбора единиц измерения для x и y . Если при измерении x и y в килограммах было получено значение $K = 20 \text{ кг}^2$, то, приняв за единицу измерения 1 г, получим для корреляционного момента значение $K = 20 \times 10^6 \text{ г}^2$. Это затрудняет сравнение корреляционных моментов различных систем случайных величин, для преодоления такого затруднения вводят другую характеристику связи между x и y — *коэффициент корреляции*.

Определение. Коэффициент корреляции случайных величин x и y — это число,

$$r[x, y] = \frac{K[x, y]}{\sigma[x]\sigma[y]}$$

— отношение корреляционного момента к произведению средних квадратичных отклонений величин x и y .

Теорема. Коэффициент корреляции всегда заключен между -1 и 1 :

$$-1 \leq r \leq 1.$$

Если $r = 1$, величины x и y связаны линейной зависимостью:

$$y = ax + b \quad (a, \text{ и } b = \text{const}),$$

причем $a > 0$; при $r = -1$ между величинами x и y имеет место линейная зависимость с $a < 0$.

Доказательство. Рассмотрим математическое ожидание случайной величины

$$(y + kx)^2,$$

где $\overset{\circ}{y} = y - m_y$;

$\overset{\circ}{x} = x - m_x$,

t — любое действительное число.

Тогда:

$$\begin{aligned} M[(\overset{\circ}{y} + t\overset{\circ}{x})^2] &= M[\overset{\circ}{y}^2 + 2t\overset{\circ}{y}\overset{\circ}{x} + \overset{\circ}{x}^2] = \\ &= M[\overset{\circ}{y}^2] + 2tM[\overset{\circ}{y}\overset{\circ}{x}] + M[\overset{\circ}{x}^2] = \\ &= D[y] + 2tK[x, y] + D[x]. \end{aligned}$$

Отсюда получили равенство вида:

$$M[(\overset{\circ}{y} + t\overset{\circ}{x})^2] = \alpha t^2 + 2\beta t + \gamma,$$

где $\alpha = D[x]$;

$\beta = K[x, y]$;

$\gamma = D[y]$.

Квадратный трехчлен, стоящий в правой части этого равенства, при любом значении t не-отрицателен. Отсюда вытекает, что дискриминант этого трехчлена, т. е. выражение

$$\beta^2 - \alpha\gamma,$$

есть число неположительное.

Значит,

$$K^2[x, y] - D[x]D[y] \leq 0$$

или

$$\frac{K^2[x, y]}{D[x]D[y]} \leq 1.$$

Мы пришли к неравенству $r^2 \leq 1$, означающему, что величина r заключена в промежутке от -1 до $+1$.

Предположим теперь, что $r^2 = 1$, т. е. r равно -1 или 1 . В этом случае дискриминант указанного выше квадратного трехчлена равен нулю. Отсюда вытекает, что трехчлен имеет действительный корень, т. е. при некотором действительном значении $t = -a$ выражение $\alpha t^2 + 2\beta t + \gamma$ равно нулю. Но тогда мы должны иметь:

$$M[(\overset{\circ}{y} - a \overset{\circ}{x})^2] = 0,$$

а это, в свою очередь, означает:

$$\overset{\circ}{y} - a \overset{\circ}{x} = 0 \text{ или } y = ax + b.$$

Значит, можно сказать, что близость величины r^2 к единице есть признак того, что зависимость между x и y близка к линейной. Если при этом $r > 0$, то с возрастанием x возрастает в среднем и y , тогда говорят о *положительной корреляции* между величинами x и y ; если же $a < 0$, то при возрастании x величина y в среднем убывает *отрицательная корреляция*.

При фиксированном значении величины x величина y все еще остается случайной, но с законом распределения, зависящей от выбранного значения x .

Пусть величина x приняла некоторое значение x_i . При этом условии вероятность любого из возможных значений y_j величины y будет:

$$r_j = p(y = y_j / x = x_i) = \frac{p(x = x_i, y = y_j)}{p(x = x_i)}.$$

Таким образом, при фиксированном значении x распределения (табл 7):

Условный закон распределения величины y

Таблица 7

y_1	y_2	...
r_1	r_2	...

(нетрудно видеть, что $r_1 + r_2 + \dots = 1$). Этот закон называют *условным законом распределения величины y* ; роль условия играет $x = x_i$.

Располагая условным законом распределения для y , можно, разумеется, найти условное математическое ожидание y . Так как оно зависит от выбранного значения $x = x_i$, обозначим его $m_y(x_i)$. Следовательно,

$$m_y(x_i) = y_1 r_1 + y_2 r_2 + \dots$$

Условное математическое ожидание играет роль во многих вопросах. В частности, с его помощью можно сформулировать необходимое и достаточное условие, при котором y является некоторой функцией от x . С этой целью заметим, что $m_y(x_i)$ есть на самом деле *случайная величина* (поскольку каждое значение x_i принимается с определенной вероятностью). Рассмотрим дисперсию этой величины, т. е. число

$$M[(m_y(x) - m_y)^2],$$

и сравним ее с дисперсией самой величины y . Оказывается, всегда справедливо неравенство:

$$M[(m_y(x) - m_y)^2] \leq D[y],$$

причем равенство достигается в том и только в том случае, когда случайная величина y есть функция от случайной величины x . Отметим, что число

$$\eta_{y,x}^2 = \frac{M[(m_y(x) - m_y)^2]}{D[y]}$$

называют *корреляционным отношением* y к x . Из сказанного выше следует, что всегда

$$0 \leq \eta_{y,x}^2 \leq 1,$$

причем равенство $\eta_{y,x}^2 = 1$ является критерием того, что y есть функция от x .

4. Метод наименьших квадратов.

Предположим, что зависимость между величинами x и y близка к линейной (коэффициент корреляции r близок к $+1$ или -1). В этом случае естественно ставить вопрос о функции

$$y = ax + b,$$

которая *наилучшим образом* выражала бы зависимость y от x . Для нахождения такой функции можно воспользоваться методом наименьших квадратов.

Пусть над системой (x, y) произведено n независимых опытов, в результате чего получены экспериментальные точки:

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n).$$

Требуется найти такую прямую $y = ax + b$, чтобы сумма квадратов отклонений экспериментальных точек от этой прямой, т. е. выражение

$$\sum_{i=1}^n [y_i - (ax_i + b)]^2,$$

обращалась в минимум.

Обозначим последнее выражение через $\phi(a, b)$:

$$\phi(a, b) = \sum_{i=1}^n [y_i - (ax_i + b)]^2,$$

Чтобы найти значения переменных a и b , обращающие это выражение в минимум, нужно приравнять к нулю производные по a и b :

$$\frac{\partial \phi}{\partial a} = -2 \sum_{i=1}^n [y_i - (ax_i + b)] x_i = 0;$$

$$\frac{\partial \phi}{\partial b} = -2 \sum_{i=1}^n [y_i - (ax_i + b)] = 0.$$

После преобразований получится:

$$\sum x_i y_i - a \sum x_i^2 - b \sum x_i = 0;$$

$$\sum y_i - a \sum x_i - nb = 0,$$

(для упрощения записей мы пишем \sum вместо $\sum_{i=1}^n$). Это система

двух линейных уравнений с двумя неизвестными a и b . Разделив обе части каждого уравнения на n , получаем:

$$\tilde{m}_{xy} - a\tilde{m}_{x^2} - b\tilde{m}_x = 0 \text{ и } \tilde{m}_y - a\tilde{m}_x - b = 0.$$

Решив эту систему, находим значения неизвестных параметров a и b :

$$a = \frac{\tilde{m}_{xy} - \tilde{m}_x \tilde{m}_y}{\tilde{m}_x^2 - \tilde{m}_x^2} \text{ и } b = \tilde{m}_y - a \tilde{m}_x .$$

Таким образом, искомая линейная зависимость y от x имеет вид:

$$y = ax + (\tilde{m}_y - a \tilde{m}_x) \text{ или } y - \tilde{m}_y = a(x - \tilde{m}_x).$$

Заметим, что выражение $\tilde{m}_{xy} - \tilde{m}_x \tilde{m}_y$, стоящее в числителе формулы

$$a = \frac{\tilde{m}_{xy} - \tilde{m}_x \tilde{m}_y}{\tilde{m}_x^2 - \tilde{m}_x^2},$$

есть эмпирический корреляционный момент $\tilde{K}_{x,y}$, знаменатель же $\tilde{m}_x^2 - \tilde{m}_x^2$ может быть заменен на \tilde{D}_x . Значит, для параметра a имеем выражение

$$a = \frac{\tilde{K}_{x,y}}{\tilde{D}_x} = \frac{\tilde{K}_{x,y}}{\tilde{\sigma}_x^2} = \frac{\tilde{K}_{x,y}}{\tilde{\sigma}_x \tilde{\sigma}_x} \times \frac{\tilde{\sigma}_y}{\tilde{\sigma}_x} = \tilde{r} \times \frac{\tilde{\sigma}_y}{\tilde{\sigma}_x},$$

где $\tilde{r} = \frac{\tilde{K}_{x,y}}{\tilde{\sigma}_x \tilde{\sigma}_y}$ — эмпирический коэффициент корреляции.

Здесь был рассмотрен метод наименьших квадратов применительно лишь к одному частному случаю, когда зависимость между величинами x и y является приблизительно линейной. Рассмотрим теперь более общий случай, когда рассматриваемая зависимость *близка к некоторой функциональной зависимости*. В этом случае тоже возникает задача о сглаживании ее методом наименьших квадратов.

Допустим, что тип разыскиваемой функциональной зависимости y от x предписан заранее из каких-либо соображений точнее, эта зависимость должна принадлежать семейству:

$$y = \varphi(x; a, b, \dots),$$

где φ — данная функция, в выражение которой входят некоторые параметры a, b, \dots . Необходимо подобрать значения этих параметров так, чтобы кривая $y = \varphi(x; a, b, \dots)$ наименее уклонялась от экспериментальных точек.

Решение этой задачи методом наименьших квадратов заключается в отыскании таких значений параметров, для которых выражение

$$\phi(a, b, \dots) = \sum [y_i - \varphi(x_i; a, b, \dots)]^2$$

принимает наименьшее значение. Эту задачу можно свести к решению системы уравнений:

$$\frac{\partial \phi}{\partial a} = 0, \quad \frac{\partial \phi}{\partial b} = 0, \quad \dots$$

При этом число уравнений совпадает с числом неизвестных параметров.

Решить эту систему в общем виде нельзя; для этого нужно задаться конкретным видом функции φ .

В ряде случаев функцию φ задают в виде многочлена:

$$a_0 + a_1x + a_2x^2 + \dots + a_kx^k,$$

где роль параметров играют коэффициенты a_0, a_1, \dots, a_k . В некоторых случаях φ выбирается как комбинация показательных функций:

$$a_1e^{\alpha_1x} + a_2e^{\alpha_2x} + \dots + a_ke^{\alpha_kx},$$

где какие-то из чисел a_i, α_i могут быть заданы заранее, в то время как другие неизвестны. Возможны и другие формы задания функции φ .

Содержание

ЛЕКЦИЯ №1. Случайные события.	3
1. Основные понятия теории вероятностей	3
1. Испытания и события. Виды случайных событий.	3
2. Определение вероятности.	3
3. Основные формулы комбинаторики.	4
4. Относительная частота. Устойчивость относительной частоты. Статистическая вероятность.	5
5. Геометрические вероятности.	7
2. Теорема сложения вероятностей	8
1. Теорема сложения вероятностей несовместных событий.	8
2. Полная группа событий. Противоположные события.	9
3. Принцип практической невозможности маловероятных событий.	10
3. Теорема умножения вероятностей	11
1. Произведение событий. Условная вероятность.	11
2. Теорема умножения вероятностей.	11
3. Независимые события. Теорема умножения для независимых событий.	13
4. Вероятность появления хотя бы одного события.	14
4. Следствия теорем сложения и умножения	15
1. Теорема сложения вероятностей совместных событий.	15
2. Формула полной вероятности.	17
3. Вероятность гипотез. Формулы Байеса.	18
5. Повторение испытаний	19
1. Формула Бернулли.	19
2. Локальная теорема Лапласа.	20
3. Интегральная теорема Лапласа.	21
4. Вероятность отклонения относительной частоты от постоянной вероятности в независимых испытаниях.	22
ЛЕКЦИЯ №2. Случайные величины	23
1. Виды случайных величин. Задание дискретной случайной величины	23
1. Случайная величина. Схема Бернулли.	23
2. Закон распределения вероятностей дискретной случайной величины.	24
3. Биномиальное распределение.	25
4. Распределение Пуассона.	26
5. Простейший поток событий.	27
6. Геометрическое распределение.	29
7. Гипергеометрическое распределение.	30
2. Математическое ожидание дискретной случайной	31
1. Числовые характеристики дискретных случайных величин. Математическое ожидание дискретной случайной величины.	31
2. Свойства математического ожидания.	33
3. Математическое ожидание числа появлений события в независимых испытаниях.	36

3. Дисперсия дискретной случайной величины	37
1. Отклонение случайной величины от ее математического ожидания	37
2. Дисперсия дискретной случайной величины	39
3. Свойства дисперсии	41
4. Дисперсия числа появлений события в независимых испытаниях	43
5. Среднее квадратическое отклонение. Среднее квадратическое отклонение суммы взаимно независимых случайных величин	45
6. Одинаково распределенные взаимно независимые случайные величины	46
7. Начальные и центральные теоретические моменты	48
4. Закон больших чисел	50
1. Неравенство Чебышева. Теорема Чебышева	50
2. Теорема Бернулли	54
5. Функция распределения вероятностей случайной величины	55
1. Определение функции распределения	55
2. Свойства функции распределения	56
6. Плотность распределения вероятностей непрерывной случайной величины	58
1. Определение плотности распределения. Вероятность попадания непрерывной	58
2. Нахождение функции распределения по известной плотности распределения	60
3. Свойства плотности распределения	60
4. Вероятностный смысл плотности распределения	61
5. Закон равномерного распределения вероятностей	63
7. Нормальное распределение	63
1. Числовые характеристики непрерывных случайных величин	63
2. Нормальное распределение	65
3. Нормальная кривая	69
4. Влияние параметров нормального распределения на форму нормальной кривой	70
5. Вероятность попадания в заданный интервал нормальной случайной величины	70
6. Вычисление вероятности заданного отклонения	71
7. Правило трех сигм	73
8. Понятие о теореме Ляпунова. Формулировка центральной предельной теоремы	73
9. Оценка отклонения теоретического распределения от нормального. Асимметрия и эксцесс	75
10. Функция одного случайного аргумента и ее распределение	77
11. Математическое ожидание функции одного случайного аргумента	78
12. Функция двух случайных аргументов. Распределение суммы независимых слагаемых. Устойчивость нормального распределения	79
13. Распределение «хи квадрат». Распределение F Фишера - Снедекора. Распределение Стьюдента	80
8. Показательное распределение	82

1. Определение показательного распределения. Вероятность попадания в заданный интервал показательно распределенной случайной величины.	82
2. Числовые характеристики показательного распределения.	84
3. Функция надежности. Показательный закон надежности.	85
4. Характеристическое свойство показательного закона надежности.	87
9. Система двух случайных величин	88
1. Понятие о системе нескольких случайных величин.	88
2. Закон распределения вероятностей дискретной двумерной случайной величины.	89
3. Функция распределения двумерной случайной величины. Свойства функции распределения двумерной случайной величины.	90
4. Вероятность попадания случайной точки в полуполосу.	93
5. Вероятность попадания случайной точки в прямоугольник.	93
6. Плотность совместного распределения вероятностей непрерывной двумерной случайной величины (двухмерная плотность вероятности).	94
7. Нахождение функции распределения системы по известной плотности распределения.	94
8. Вероятностный смысл двумерной плотности вероятности.	94
9. Вероятность попадания случайной точки в произвольную область.	95
10. Свойства двумерной плотности вероятности.	97
11. Отыскание плотностей вероятности составляющих двумерной случайной величины.	97
12. Условные законы распределения составляющих системы дискретных случайных величин.	99
13. Условные законы распределения составляющих системы непрерывных случайных величин.	100
14. Условное математическое ожидание.	102
15. Зависимые и независимые случайные величины.	102
16. Числовые характеристики систем двух случайных величин. Корреляционный момент. Коэффициент корреляции.	104
17. Коррелированность и зависимость случайных величин.	107
18. Нормальный закон распределения на плоскости.	107
19. Линейная регрессия. Прямые линии среднеквадратической регрессии.	108
20. Линейная корреляция. Нормальная корреляция.	111
ЛЕКЦИЯ №3. Случайные функции	114
1. Случайные функции	114
1. Определение случайной функции.	114
2. Математическое ожидание случайной функции и его свойства.	116
3. Дисперсия случайной функции и ее свойства.	117
4. Корреляционная функция случайной функции	

и ее свойства.	118
5. Нормированная корреляционная функция.	121
6. Взаимная корреляционная функция и ее свойства.	122
7. Нормированная взаимная корреляционная функция.	123
8. Характеристики суммы случайных функций.	124
9. Производная случайной функции и ее характеристики.	125
10. Интеграл от случайной функции и его характеристики.	128
11. Комплексные случайные величины и их числовые характеристики.	132
2. Стационарные случайные функции	134
1. Определение стационарной случайной функции.	134
2. Свойства корреляционной функции стационарной случайной функции.	135
3. Нормированная корреляционная функция стационарной случайной функции.	136
4. Стационарно связанные случайные функции.	137
5. Корреляционная функция производной стационарной случайной функции.	138
6. Взаимная корреляционная функция стационарной случайной функции и ее производной.	139
7. Корреляционная функция интеграла от стационарной случайной функции.	140
8. Определение характеристик эргодических стационарных случайных функций из опыта.	141
3. Элементы математической статистики	143
1. Вариационный ряд.	143
2. Оценка неизвестной вероятности по частоте.	146
3. Корреляция.	148
Содержание	156

Щербакова Ю. В.

КОНСПЕКТ ЛЕКЦИЙ
ТЕОРИЯ ВЕРОЯТНОСТЕЙ
И МАТЕМАТИЧЕСКАЯ
СТАТИСТИКА

Завредакцией: *И. С. Дозорова*
Корректор: *Г. А. Серикова*
Технический редактор: *Т. И. Федорова*

Формат: 84 × 108/32
Гарнитура: «Ньютон»