Галкин С.В, Панов В.Ф., Петрухина О.С.

Краткий курс
основ теории вероятности
и математической статистики

конспект лекций

(часть 1)

Москва 2007

Лекция 1

Вероятность.

В теории вероятностей рассматриваются такие явления или опыты, конкретный исход которых не определяется однозначно условиями опыта (случаен), но по результатам большого числа экспериментов в среднем может быть предсказан (свойство статистической устойчивости).

Элементарным событием (элементарным исходом) называется любое событие - исход опыта, которое нельзя представить в виде объединения других событий. Так как исход опыта случаен, то и любое элементарное событие случайно, далее будем говорить просто о событиях, не подчеркивая их случайность.

Пространством элементарных событий Ω (исходов) называется множество всех элементарных событий (исходов) $(\omega_1,..\omega_n...)$, если в результате опыта обязательно наступает какой-либо из элементарных исходов и только один. Пространство элементарных событий может содержать конечное, счетное и даже бесконечное множество элементарных событий.

Случайным событием (событием) называется подмножество пространства элементарных событий. **Элементами события** являются элементарные события, образующие это событие.

Пример. Бросается одна монета, события: герб $(\omega_1 = \Gamma)$ или решка $(\omega_2 = P)$. $\Omega = (\Gamma, P)$.

Пример. Бросаются две монеты: $\Omega = ((\Gamma, \Gamma), (\Gamma, P), (P, \Gamma), (P, P))$.

Пример. Капля дождя падает на прямоугольную площадку. $\Omega = ((x, y), a < x < b, c < y < d)$..

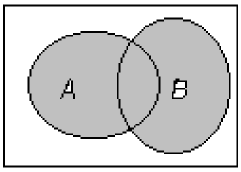
Достоверное событие – событие, которое всегда происходит в результате данного опыта, оно содержит все элементарные события и обозначается Ω.

Невозможное событие – «событие», которое не может произойти в результате данного опыта, оно не содержит элементарных событий и обозначается $\varnothing$.

Действия над событиями.

События определены как множества, поэтому действия над ними аналогичны действиям над множествами и хорошо иллюстрируются диаграммами Венна. Пространство Ω будем обозначать прямоугольником, элементарное событие – точкой прямоугольника, а каждое событие – подмножеством точек этого прямоугольника. Результат операции над событиями будем заштриховывать.

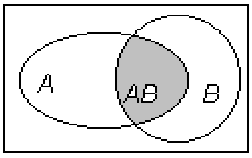
Пример. Пусть выбираются карты из колоды карт. Событие A – выбор червонной карты, событие B – выбор десятки.



Суммой двух событий A и B называется событие $C = A + B$ (или $C = A \cup B$), состоящее из элементарных событий, принадлежащих A или B.

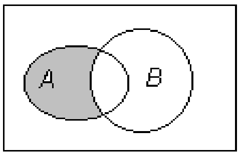
<u>Пример</u>.

$C = A + B$ – выбор любой червонной карты или любой десятки.



Произведением двух событий A и B называется событие $D = A\,B$ (или $C = A \cap B$), состоящее из элементарных событий, принадлежащих и A, и B.

<u>Пример</u>. $A\,B$ – выбор десятки червей.



Разностью двух событий A и B называется событие $A \setminus B$, состоящее из элементарных событий, принадлежащих А и не принадлежащих B.

<u>Пример</u>. $A \setminus B$ –выбор любой червонной карты, кроме десятки.

Классификация событий

Событие, состоящее из всех элементарных событий, не содержащихся в A, обозначим $\overline{A}$ и будем называть ***противоположным*** событием $\overline{A} = \Omega \setminus A$.

Пример. A –выбор червонной карты; $\overline{A}$ –выбор любой карты другой масти.

Два события A и B будем называть ***совместными***, если имеется хотя бы одно общее элементарное событие, т.е если $A\,B \neq \varnothing$ и ***несовместными***, если $A\,B = \varnothing$.

Пример. A – выбор червонной карты и B – выбор десятки – совместные события, так как $A\,B$ (выбор червонной десятки) $\neq \varnothing$.

Пример. Бросается игральный кубик. A – выпадение четного числа очков, $A = \{2, 4, 6\}$. B – выпадение нечетного числа очков, $B = \{1, 3, 5\}$. Очевидно, что A и B несовместны.

Полная группа событий – *это совокупность n событий A_1, A_2, ..., A_n, одно из которых обязательно произойдет, т.е.* $\sum_{i=1}^{n} A_i = \Omega$.

Свойства операций над событиями

Если $A \subseteq B$, то $A + B = B$, $A\,B = A$.

Отсюда следует **1**. $\overline{\Omega} = \varnothing$, **2**. $A + A = A$, **3**. $A\,A = A$, **4**. $A + \Omega = \Omega$, **5**. $A + \varnothing = A$,

6. $A\,\Omega = A$, **7**. $A\varnothing = \varnothing$, **8**. $\overline{\overline{A}} = A$, **9**. $A + \overline{A} = \Omega$, **10**. $A\overline{A} = \varnothing$.

Коммутативность операций:

$A + B = B + A$; $A\,B = B\,A$.

Ассоциативность операций:

$A + (B + C) = (A + B) + C$, $A(B\,C) = (A\,B)C = A\,B\,C$.

Дистрибутивность $A\,(B + C) = A\,B + A\,C$, $A + (B\,C) = (A + B)(A + C)$.

Пример. Вычислим $(A + B)(A + C) = A\,A + B\,A + A\,C + B\,C = A + B\,C$.

В самом деле, $B\,A \subset A$, $A\,C \subset A$, $A\,A = A$, тогда $A\,A + B\,A = A$, $A + A\,C = A$.

Правило двойственности (теорема де Моргана)

$\overline{\sum_{k=1}^{\infty} A_k} = \prod_{k=1}^{\infty} \overline{A_k}$, $\overline{\prod_{k=1}^{\infty} A_k} = \sum_{k=1}^{\infty} \overline{A_k}$ *Пример.* $\overline{A + B} = \overline{A}\,\overline{B}$ $\overline{AB} = \overline{A} + \overline{B}$

Алгебра событий

Пусть Ω - пространство элементарных событий. **Алгеброй событий** называется такая система случайных событий S, что 1) $S \supset \Omega$, 2) $\forall\ A,\ B \subset S \Rightarrow A + B \subset S$, $A\,B \subset$S, $A \setminus B \subset S$.

Следствие $\varnothing = \Omega \setminus \Omega \subset S$

Пусть Ω содержит конечное число элементов, $\Omega = \{\omega_1, \ldots \omega_n\}$. Тогда алгебру S можно построить как множество всех подмножеств Ω.

$S = \{\varnothing,\ \{\omega_1\},\ \ldots\ \{\omega_n\},\ \{\omega_1, \omega_2\}, \ldots \{\omega_1, \omega_n\}, \ldots \{\omega_{n-1}, \omega_n\}, \ldots \{\omega_1, \ldots \omega_n\}\}$, в ней всего 2^n элементов.

Аналогично стоится алгебра для счетного числа событий.

События должны выбираться из алгебры событий.

Для бесконечного (не счетного) числа событий класс событий должен быть сужен. Вводится σ- алгебра событий.

Сигма-алгеброй (Σ-алгеброй) событий называется непустая система подмножеств пространства элементарных событий, такая, что

1) $A \subset \Sigma \Rightarrow \overline{A} \subset \Sigma$,

2) $A_1, A_2, \ldots, A_n \ldots \subset \Sigma \Rightarrow (A_1 + A_2 + \ldots + A_n + \ldots) \subset \Sigma$, $(A_1 A_2 .. A_n \ldots) \ldots \subset \Sigma$.

Любая сигма-алгебра событий является алгеброй событий, но не наоборот.

Вероятность

Классическое определение вероятности события

Случаями *называются равновозможные, несовместные события, составляющие полную группу.*

Пусть пространство элементарных событий Ω содержит конечное число случаев.

Пусть N – общее число случаев в Ω, а N_A – число случаев, образующих событие A (или, как говорят, благоприятствующих событию A).

Определение. ***Вероятностью события*** A называется отношение $P(A) = \dfrac{N_A}{N}$.

Это - ***классическое определение вероятности***.

Примеры: 1. Бросание игральной кости. $\Omega = \{\omega_1, \omega_2, ...\omega_6\}, \quad N = 6.$.

A – количество очков кратно трем, $A = \{\omega_3, \omega_6\}$, $N_A = 2$.

$$P(A) = \frac{2}{6} = \frac{1}{3}.$$

2. В урне a белых и b черных шаров. Опыт – вынимается один шар.

A – шар черный.

$$P(A) = \frac{b}{a+b}.$$

Исходя из классического определения вероятности события, легко доказать свойства вероятности:

1) $P(\Omega) = 1\,(N_A = N)$; 2) $0 \le P(A) \le 1$ ($0 \le N_A \le N$); 3) Если $A\,B = \varnothing$, то $P(A+B) = P(A) + P(B)$ $(N_{A+B} = N_A + N_B)$

и их следствия

4) $P(\varnothing) = 0$ ($N_{\varnothing} = 0$); 5) $P(\overline{A}) = 1 - P(A)$ ($A + \overline{A} = \Omega$, $A\overline{A} = \varnothing$, $P(A) + P(\overline{A}) = 1$);

6) Если $A \subseteq B$, то $P(A) \le P(B)$ $(N_A \le N_B)$.

Для определения общего числа равновозможных исходов и числа благоприятствующих исходов используется **основной принцип комбинаторики:** *пусть некоторая операция P представляет собой последовательность n операций P_k (k=1, ...n), каждая из которых может быть выполнена m_k способами. Тогда операция P может быть выполнена $m_1 \cdot m_2 \cdot ... \cdot m_n$ способами.*

Пусть мы делаем выборку поочередно m элементов (например, шаров из урны) из n элементов. Мы можем возвращать очередной шар (в урну), тогда при каждом очередном выборе мы будем иметь все те же n шаров. Такая выборка называется выборкой **с возвращением**. Мы можем и не возвращать шар, тогда при каждом выборе мы будем выбирать из все меньшего числа шаров. Такая выборка называется выборкой **без возвращения.** С другой стороны, если учитывать **порядок** появления шаров, то выборка называется упорядоченной или **размещением из n по m (**шаров**).** Если порядок шаров не учитывать (важно, какие шары выбраны, но не важно, в каком порядке), то такая выборка называется **неупорядоченной** или **сочетанием из n по m** (шаров**).** Выясним, сколькими способами можно произвести ту или иную выборку

	Сочетания	Размещения
Без возвращения	$C_n^m = \dfrac{n(n-1)\ldots(n-m+1)}{m!} = \dfrac{n!}{m!\,(n-m)!}$	$A_n^m = n(n-1)\ldots(n-m+1)$
С возвращением	$\overline{C_n^m} = C_{n+m-1}^m$	$\overline{A_n^m} = n^m$

Формулы для размещений легко получаются из принципа комбинаторики. Для того, чтобы перейти от размещений (без возвращений) к сочетаниям (без возвращений), нужно упорядочить выборки, т.е. исключить те, которые отличаются только порядком элементов. Выборки, отличающиеся только порядком элементов, называются **перестановками**. Число перестановок из m элементов равно $P_m = A_m^m = m!$. Поэтому $C_m^n = \dfrac{A_n^m}{P_m}$.

Доказательство формулы для сочетаний с возвращением приведено в учебнике /1/.

Пример. Производится выборка двух шаров ($m = 2$) из урны, в которой находится $n = 3$ шара.

1) Размещения с возвращением:

(1,1) (1,2) (1,3) (2,1) (2,2) (2,3) (3,1) (3,2) (3,3), $\overline{A_3^2} = 3^2 = 9$.

2) Размещения (без возвращения): (1,2) (1,3) (2,1) (2,3) (3,1) (3,2), $A_3^2 = 6$.

3) Сочетания с возвращением: (1,1) (1,2) (1,3) (2,2) (2,3) (3,3), $C_{2+3-1}^2 = C_4^2 = 6$.

4) Сочетания (без возвращения): (1,2) (1,3) (2,3), $C_3^2 = 3$.

Пример. **Задача о выборке «бракованных» деталей**.

В партии из N одинаковых деталей M бракованных. Выбирается (не возвращая) n деталей. Какова вероятность того, что среди них окажется ровно m бракованных?

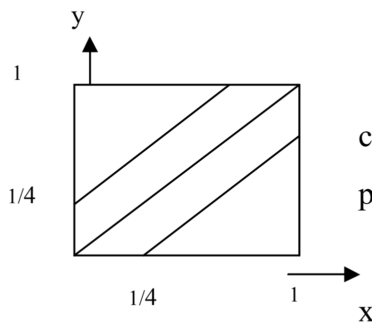
Общее количество случаев (сочетания из N деталей по n) равно C_N^n. Мы выбираем m бракованных деталей из M бракованных, но и одновременно выбираем (n - m) деталей без брака из N - M деталей без брака. Тогда, по основному принципу комбинаторики, такому выбору благоприятствует $C_M^m C_{N-M}^{n-m}$ случаев. Поэтому искомая вероятность равна $p = \dfrac{C_M^m C_{N-M}^{n-m}}{C_N^n}$.

Геометрическая вероятность

Формула классической вероятности применяется только в схеме случаев, что встречается довольно редко. Отношение $P(A)=\dfrac{N_A}{N}$ представляет собой «долю» благоприятных исходов среди всех возможных исходов. Аналогичным образом подсчитывают вероятность события в некоторых более сложных случаях, когда имеется *бесконечное* число *равновозможных* исходов.

Пример. **Задача о встрече.** Два студента договорились встретиться от 10 до 11 часов на определенном месте, причем первый студент, пришедший на место, ждет товарища 15 минут и уходит. Какова вероятность встречи?

Выберем начало системы координат в точке (10, 10). Отложим по осям системы координат x- время прихода первого студента, y – время прихода второго студента.



Тогда множество $|x-y|<1/4,\ 0<x<1,\ 0<y<1$ содержит точки (события) встречи студентов. Его мера (площадь) равна 1- $(3/4)^2$ = 7/16. Так как мера (площадь) Ω =1, то $P(A)$ = 7/16.

Статистическая вероятность

Формулы классической вероятности и геометрической вероятности справедливы только для случая *равновозможных* исходов. В действительности мы на практике имеем дело с *неравновозможными* исходами. В этих случаях можно определить вероятность случайного события, используя понятие ***частоты события***. Пусть надо определить вероятность того, что в испытании произойдет событие A. Для этого *в одинаковых условиях* проводятся испытания, в каждом из которых возможны два исхода: A и $\overline{A}$. ***Частотой*** события A будем называть отношение числа N_A испытаний, в которых зафиксировано событие A к общему числу N испытаний.

Статистической вероятностью события *A называется* $P(A)=\lim_{N\to\infty}\dfrac{N_A}{N}$

Заметим, что по классическому, геометрическому и статистическому определениям для вероятности события *P(A)* выполнены три основных свойства:

1) $P(A)\geq 0$, 2) $P(\Omega)=1$, 3) $P(A_1+\ldots+A_n)=P(A_1)+\ldots+P(A_n)$ если $A_1,\ldots,A_n$ попарно несовместны. Однако в этих определениях элементарные события предполагаются равновозможными.

А.Н. Колмогоров отказался от предположения равновозможности элементарных событий и распространил третье свойство на счетное число событий из сигма-алгебры событий. Это дало ему возможность дать общее аксиоматическое определение вероятности события.

Аксиоматическое определение вероятности (по А.Н.Колмогорову)

Вероятностью $P(A)$ называется числовая функция, заданная на сигма – алгебре событий, удовлетворяющая трем аксиомам:

1) неотрицательность $P(A) \geq 0, \forall A \subset \Sigma$ - сигма – алгебре событий на Ω,

2) нормировка $P(\Omega) = 1$,

3) расширенная аксиома сложения: для любых попарно несовместных событий $A_1, \ldots, A_n$ выполнено $P(A_1 + \ldots + A_n + \ldots) = P(A_1) + \ldots + P(A_n) + \ldots + \ldots$(счетная аддитивность).

Итак, по А.Н. Колмогорову *вероятность (вероятностная мера) - это числовая неотрицательная нормированная счетно - аддитивная функция (множества – события), заданная на сигма – алгебре событий.*

Если Ω состоит из конечного или счетного числа событий, то в качестве сигма – алгебры Σ может рассматриваться алгебра S событий. Тогда по аксиоме 3 вероятность любого события A равна сумме вероятностей элементарных событий, составляющих A.

Вероятностным пространством называется тройка (Ω, Σ, P).

Свойства вероятности

1) $P(\overline{A}) = 1 - P(A)$. В самом деле, $A + \overline{A} = \Omega$, $A, \overline{A}$ несовместны. По аксиоме 3 $P(A + \overline{A}) = P(A) + P(\overline{A}) = P(\Omega) = 1$.

2) $P(\varnothing) = 0$. Так как $\forall A \quad A + \varnothing = A$, по аксиоме 3 $P(A + \varnothing) = P(A) + P(\varnothing) = P(A)$ $\Rightarrow P(\varnothing) = 0$.

3) Если $A \subset B$, то $P(A) \leq P(B)$. Так как $B = A + B \setminus A$, по аксиоме 3 $P(B) = P(A) + P(B \setminus A)$, но по аксиоме 1 $P(B \setminus A) \geq 0$.

Пример. Из урны с четырьмя шарами с номерами 1, 2, 3, 4 три раза наугад вынимают шар и записывают его номер: а) возвращая шары, б) не возвращая шары. Какова вероятность 1) получить комбинацию 111, 2) из номеров шаров составить возрастающую последовательность?

В случае а) имеем размещения с возвращением, $N = 4^3$, 1) N_A=1, $P = \frac{1}{4^3}$, 2) N_A = C_4^3, так как возрастающую последовательность можно составить всегда из неповторяющихся номеров. $P = C_4^3/4^3$.

В случае б) $N = C_4^3$,1) Так как номера шаров не повторяются, то N_A =0 $P = 0$,

2) P = 1, так как $N = N_A = C_4^3$.

Пример. Пять человек садятся в поезд метро, состоящий из пяти вагонов. Какова вероятность того, что они окажутся в разных вагонах?

Общее число элементарных событий равно числу размещений с повторением из пяти элементов по пять $N = 5^5$. Число элементарных событий, составляющих A, равно 5! Поэтому $P = 5!/5^5$.

Лекция 2

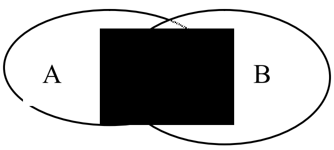
Условная вероятность

Вероятность события A при условии, что событие В наступило, будем называть ***условной*** и обозначать $P(A/\text{B})$.

Если никаких дополнительных условий не накладывается, то вероятность называется ***безусловной***. Это – обычная, определенная выше вероятность.

Рассмотрим пример. Пусть в данной аудитории присутствует N студентов. Среди них N_A –любящих математику, N_B – любящих физику, N_{AB} – любящих и математику, и физику. Лектор случайно выбирает одного студента. Введем следующие события:

A – случайно выбранный студент любит математику, *B* – физику, *AB* – и математику, и физику. На диаграммах Венна это выглядит так.



Тогда вероятности этих событий равны $P(A) = \frac{N_A}{N} = \frac{mesA}{mes\Omega}$, $P(B) = \frac{N_B}{N} = \frac{mesB}{mes\Omega}$, . $P(AB) = \frac{N_{AB}}{N} = \frac{mesAB}{mes\Omega}$. Это - безусловные вероятности (mes – мера множества - площадь области).

Предположим теперь, что мы захотели узнать вероятность того, что случайно выбранный любитель физики любит еще и математику. В этом случае количество всех возможных исходов N_B (выбираем только любителей физики), а количество благоприятных исходов – N_{AB}. Получим

$$P(A/B) = \frac{N_{AB}}{N_B} = \frac{mesAB}{mesB} = \frac{mes\Omega \cdot P(AB)}{mes\Omega \cdot P(B)} = \frac{P(AB)}{P(B)}$$

Мы рассмотрели частный случай. Введем в общем случае следующее формальное определение.

В общем случае будем называть ***условной вероятностью события А при условии В***

$$P(A/B)=\frac{P(AB)}{P(B)} \text{ (если } P(B)\neq 0).$$

Формула вероятности произведения событий

(теорема умножения вероятностей). Независимые события.

Из формулы условной вероятности следует

теорема умножения вероятностей *P(AB) = P(B)·P(A/B) = P(A)·P(B/A).*

Вероятность совместного наступления двух событий (вероятность произведения этих событий) равна произведению вероятности одного из них на условную вероятность другого.

Теорема умножения вероятностей может быть обобщена на случай произвольного числа событий.

$$P(\prod_{i=1}^{n}A_i)=P(A_1)\cdot P(A_2/A_1)\cdot P(A_3/A_1A_2)\cdot\ldots\cdot P(A_n/\prod_{i=1}^{n-1}A_i).$$

Событие А будем называть ***независимым*** от события *B*, если *P(A/B) = P(A)*, т.е. если условная вероятность равна безусловной.

Два события называются независимыми, если наступление одного из них не изменяет вероятность другого. В противном случае события называются зависимыми.

События $A_1, A_2, \ldots, A_n$ называются независимыми в совокупности, если для любых k из них вероятность их произведения равна произведению их вероятностей. В частности, выполнено соотношение $P(\prod_{i=1}^{n}A_i)=\prod_{i=1}^{n}P(A_i)$.

Можно показать, что из попарной независимости не вытекает независимость в совокупности.

Пример 1. Наугад вытаскивается одна карта из тщательно перетасованной колоды в 36 карт.

A – вытащили даму; $P(A)=4/36=1/9$;

1) Дополнительная информация: произошло событие B – вытащили карту бубновой масти,

$P(B)=1/4, P\left(AB\right)=1/36.$ $P(A/B)=\frac{P(AB)}{P(B)}=\frac{1/36}{1/4}=1/9;$ $P(A)=P(A/B)\Rightarrow A$ и B – независимы.

2) Дополнительная информация: произошло событие C – вытащили «картинку» (валет или дама, или король), $P\left(C\right)=12/36$, $P\left(AC\right)=4/36$.

$$P(A/C)=\frac{P(AC)}{P(C)}=\frac{4/36}{12/36}=1/3\neq 1/9;\quad P(A)\neq P(A/C)\Rightarrow A \text{ и } C \text{ зависимы.}$$

$$P(A/C)=\frac{P(AC)}{P(C)}=\frac{4/36}{12/36}=1/3\neq 1/9;\quad P(A)\neq P(A/C)\Rightarrow A \text{ и } C \text{ зависимы.}$$

Пример 2. На плотной бумаге написано слово «стипендия». Разрезав надпись на карточки «с», «т», «и», «п», «е», «н», «д», «и», «я» и перемешав их, вытаскиваем наугад шесть букв. Какова вероятность того, что мы вытащим последовательно «п», «е», «н», «с», «и», «я»?

P(«пенсия») = P(п) P(е/п) P(н/пе) P(с/пен) P(и/пенс) P(я/пенси) =

= 1/9 1/8 1/7 1/6 2/5 1/4 = 1/30240.

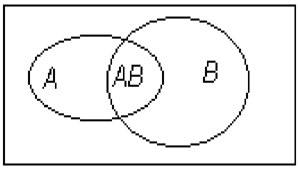
Решая эту задачу методами комбинаторики, получим

$$\text{P(«пенсия»)} = \frac{P_2}{A_9^6}=\frac{2!}{9\cdot 8\cdot 7\cdot 6\cdot 5\cdot 4}=\frac{1}{30240}.$$

Формула вероятности суммы совместных событий

(теорема сложения вероятностей).

.Пусть мы имеем два совместных события A и B. Преобразуем их сумму в сумму несовместных событий:



$$A+B=A+B\setminus A\Rightarrow P(A+B)=P(A)+P(B\setminus A),$$

$$B=B\setminus A+AB\Rightarrow P(B)=P(B\setminus A)+P(AB)\Rightarrow P(B\setminus A)=P(B)-P(AB).$$

Подставляя второе выражение в первое, получим

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

Эта формула составляет суть **теоремы сложения вероятностей:** *вероятность суммы событий равна сумме вероятностей этих событий за вычетом вероятности произведения событий.*

Пример. По мишени один раз стреляют два стрелка. Вероятность попадания первого стрелка в мишень p_1 = 0,7, второго – p_2 = 0,8. Какова вероятность того, что кто-нибудь из них попадет в мишень?

$A = A_1 + A_2$, A попадание в мишень; A_1 – попал первый стрелок; A_2 – попал второй стрелок.

$P(A) = P(A_1 + A_2) = P(A_1) + P(A_2) - P(A_1A_2) = P(A_1) + P(A_2) - P(A_1)P(A_2) = 0.7 + 0{,}8 - 0{,}7\cdot 0{,}8 = 0{,}94$.

Получим **вероятность суммы трех совместных событий**.

$$P(A+B+C)=P(A+B)+P(C)-P((A+B)C)=P(A)+P(B)-P(AB)+P(C)-P(AC)-$$
$$-P(BC)+P(ABC), \textit{так как } ACBC=ABC$$

Получена формула

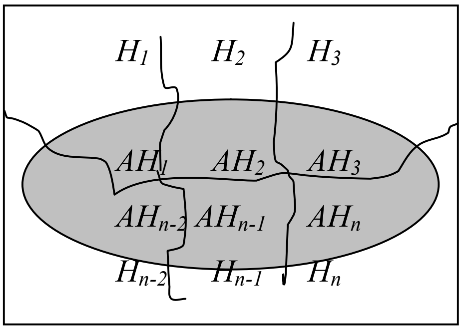
P(A + B + C) = P(A) + P(B) + P(C) – P(AB) – P(AC) – P(BC) + P(ABC).

Обобщая полученный результат на сумму n совместных событий, получим формулу

$$P(\sum_{i-1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i A_j) + \sum_{i<j<k} P(A_i A_j A_k) + \ldots + (-1)^{n+1} P(\prod_{i=1}^{n} A_i).$$

Формула полной вероятности.

Пусть требуется определить вероятность события A, которое может произойти в сочетании с одним из событий $H_1, H_2, \ldots, H_n$, образующих полную группу несовместных событий ($\sum_{i=1}^{n} H_i = \Omega,\ H_i H_j = \varnothing,\quad i \neq j$). Эти события будем называть гипотезами.



$$A = A\Omega = A(H_1 + H_2 + \cdots + H_n) = \sum_{i=1}^{n} AH_i$$

В соответствии со свойством 3) вероятности и теоремой умножения вероятностей

$$P(A) = P(\sum_{i=1}^{n} AH_i) = \sum_{i=1}^{n} P(AH_i) = \sum_{i=1}^{n} P(H_i) P(A / H_i)$$

$$P(A) = \sum_{i=1}^{n} P(H_i) P(A / H_i)$$

Пример. Из n экзаменационных билетов студент знает m («хорошие билеты»*)*. Что лучше: брать на экзамене билет первым или вторым?

Решение. Введем событие A – студент взял «хороший» билет.

Студент берет билет первым. В этом случае $P(A) = \frac{m}{n}.$

Студент берет билет вторым. Введем две гипотезы:

H_1 – первый студент взял «хороший» билет, $H_2 = \overline{H_1}$.

$$P(A) = P(H_1) P(A / H_1) + P(H_2) P(A / H_2) = \frac{m}{n} \cdot \frac{m-1}{n-1} + (1 - \frac{m}{n}) \cdot \frac{m}{n-1} = \frac{m}{n}.$$

Вывод: безразлично, брать билет первым или вторым.

Формула Байеса (теорема гипотез).

В соответствии с теоремой умножения вероятностей

$P(AH_i) = P(H_i)\cdot P(A/H_i) = P(A)\cdot P(H_i/A)$.

В это равенство подставим значение $P(A)$, вычисленное по формуле полной вероятности и найдем $P(H_i/A)$.

$$P(A)=\sum_{i=1}^{n}P(H_i)P(A/H_i) \qquad P(H_i/A) = P(H_i)\frac{P(A/H_i)}{P(A)}=\frac{P(H_i)P(A/H_i)}{\sum_{i=1}^{n}P(H_i)P(A/H_i)}$$

Это следствие из теоремы умножения и формулы полной вероятности называется ***формулой Байеса*** или теоремой гипотез.

Вероятности гипотез $P(H_i)$, входящие в формулу полной вероятности, называют ***априорными***, т.е. «доопытными». Пусть опыт произведен и его результат известен, т.е. мы знаем, произошло или не произошло событие А. Получившийся результат мог произойти при осуществлении какой-то одной гипотезы H_i. Дополнительная информация об исходе опыта перераспределяет вероятности гипотез. Эти перераспределенные вероятности гипотез $P(H_i/A)$ называют ***апостериорными*** , т.е. «послеопытными».

Пример В одной из корзин 1 камешек и 4 кусочка хлеба, во второй – 4 камешка и 1 кусочек хлеба. Мышка наугад выбирает корзину, бежит к ней и вытаскивает кусочек хлеба - событие А (предполагается, что он затем вновь возвращается в корзину). Какова вероятность события А? Каковы вероятности того, что второй раз мышка побежит к первой корзине, ко второй корзине? Какова вероятность того, что она второй раз вытащит кусочек хлеба?

Рассмотрим гипотезы

H_1 – мышка бежит к первой корзине,

H_2 – мышка бежит ко второй корзине.

$P(H_1) = 1/2 = P(H_2)$ (априорные вероятности)

$$P(A/H_1)=4/5, \quad P(A/H_2)=1/5,. \; P(A)=4/5\;1/2+1/5\,1/2=1/2$$

$$P(H_1/A) = \frac{P(H_1)P(A/H_1)}{\sum_{i=1}^{2}P(H_i)P(A/H_i)}=\frac{1/2\;4/5}{1/2}=4/5$$

$$P(H_2/A) = \frac{P(H_2)P(A/H_2)}{\sum_{i=1}^{2}P(H_i)P(A/H_i)}=\frac{1/2\;1/5}{1/2}=1/5$$ (апостериорные вероятности).

При втором подходе $P(A)=4/5\;4/5+1/5\,1/5=17/25>1/2$.

Мышка обучилась, второй раз она выберет первую корзину с большей вероятностью и добьется большего успеха.

Заметим, что это – один из основных принципов обучения кибернетических систем.

Лекция 3

Случайные величины

Случайная величина – это величина (число), которая в результате опыта может принимать то или иное значение.

Более строго, **случайная величина – это числовая функция случайного события.** $X = X(\omega), \omega \in \Omega \in S$. Здесь S - алгебра событий.

Случайная величина называется **дискретной**, если множество ее значений конечно или счетно. Например, число очков на грани брошенной кости, число бросков монеты до появления герба – дискретные случайные величины.

Случайная величина называется **непрерывной,** если ее значения заполняют некоторый интервал, возможно, бесконечный. Здесь надо рассматривать Σ - сигма - алгебру событий. Например, расстояние от центра мишени при стрельбе, время до отказа прибора, ошибка измерения – непрерывные случайные величины.

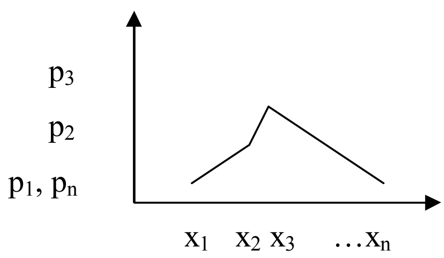
Рассмотрим *дискретную случайную величину*, принимающую значения $x_1,..x_n$. Имеем полную группу (иначе, не все значения учтены) несовместных событий $X = x_1,..,X = x_n$. Вероятности этих событий равны соответственно $p_1,..p_n$ $(p_1 + ... + p_n = 1)$. Будем говорить, что дискретная случайная величина X принимает значения $x_1,..x_n$ с вероятностями $p_1,..p_n$.

Законом распределения дискретной случайной величины называется любое соотношение, устанавливающее зависимость между ее значениями $x_1,..x_n$ и вероятностями $p_1,..p_n$, с которыми эти значения достигаются.

Основные формы закона распределения дискретной случайной величины: **ряд распределения** – таблица

x_i	x_1	…..	x_n
p_i	p_1	…..	p_n

многоугольник распределения



Можно задать закон распределения в виде **аналитической зависимости**, связывающей значения $x_1,..x_n$ и вероятности $p_1,..p_n$.

Рассмотрим *непрерывную случайную величину.* Для непрерывной случайной величины $P(X = x) = 0$, поэтому рассматривают события $X < x$ и вероятности этих событий.

Функцией распределения непрерывной случайной величины $F(x)$ называется вероятность события $X < x$. $F(x) = P(X < x)$.

Свойства функции распределения.

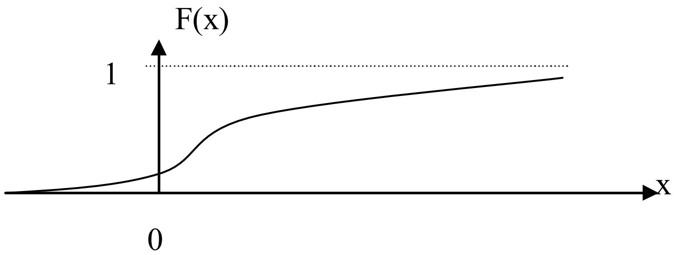
1) $0 \le F(x) \le 1$ по аксиомам вероятности,

2) $F(x_1) \le F(x_2)$, если $x_1 < x_2$, т.е. **функция распределения – неубывающая функция.** В самом деле, $X < x_1 \Rightarrow X < x_2$, следовательно, $F(x_1) = P(X < x_1) \le P(X < x_2) = F(x_2)$.

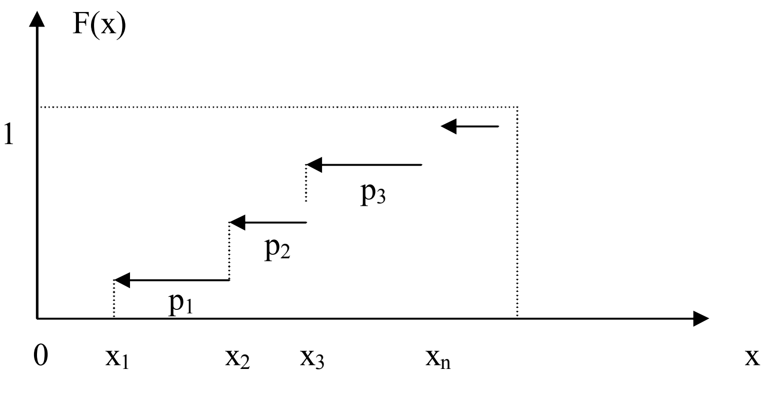
3) $\lim_{x\to-\infty} F(x) = 0,\quad \lim_{x\to+\infty} F(x) = 1$ В самом деле, событие $X < -\infty$ - невозможное, и его вероятность нулевая. Событие $X < +\infty$ - достоверное, и его вероятность равна 1.

4) $P(a \le X < b) = F(b) - F(a)$. Так как события $A = (X < a)$ *и* $B = (a \le X \le b)$ несовместны и событие $C = (X < b)$ есть сумма этих событий, то $F(b) = P(X < b) = P(C) = P(A) + P(B) = F(a) + P(a \le X \le b)$.

График функции распределения имеет, примерно, следующий вид



Функцию распределения можно определить и для дискретной случайной величины. Ее график будет графиком ступенчатой функции со скачками p_i в точках x_i, непрерывной слева в этих точках. $P(X = x_i) = p_i = F(x_i + 0) - F(x_i)$



Для непрерывной случайной величины вводится плотность распределения вероятностей.

Плотностью распределения (вероятностей) называется *производная функции распределения* $p(x) = F'(x)$.

Ясно, что $F(x) = \int\limits_{-\infty}^{x} p(x)dx$, *т.к.* $F(-\infty) = 0$, $P(a \le X < b) = F(b) - F(a) = \int\limits_{a}^{b} p(x)dx$.

Часто функцию распределения называют *интегральным законом распределения, а плотность распределения – дифференциальным законом распределения.* Так как $P(x < X < x + \Delta x) = \int\limits_{x}^{x+\Delta x} p(x)dx \approx p(x)\Delta x$, то p(x)dx называется *элементом вероятности.*

Свойства плотности распределения

1) $p(x) \ge 0$, так как функция распределения – неубывающая функция,

2) $\int\limits_{-\infty}^{+\infty} p(x)dx = 1$ (*условие нормировки)* , так как $F(+\infty) = 1$.

Числовые характеристики случайных величин

Начальный момент s-го порядка.

Для дискретных случайных величин $\alpha_s = \sum\limits_{i=1}^{n} x_i^s p_i$.

Для непрерывных случайных величин $\alpha_s = \int\limits_{-\infty}^{+\infty} x^s p(x)dx$.

Математическим ожиданием случайной величины называется ее первый начальный момент $m_x = M(x) = \alpha_1(X)$.

Для дискретных случайных величин $m_x = x_1 p_1 + ... + x_n p_n$. Если на числовой оси расположить точки $x_1,..x_n$ с массами $p_1,..p_n$, то m_x - абсцисса центра тяжести системы точек. Аналогично, для непрерывных случайных величин $m_x = \int\limits_{-\infty}^{+\infty} xp(x)dx$ имеет смысл абсциссы центра тяжести кривой распределения.

Математическое ожидание функции случайной величины вычисляют по формулам

$M(f(X)) = \sum\limits_{i=1}^{n} f(x_i)p_i$ для дискретной случайной величины,

$M(f(X)) = \int\limits_{-\infty}^{+\infty} f(x)p(x)dx$ для непрерывной случайной величины.

Свойства математического ожидания.

1) *M (C) = C*.

Для дискретных случайных величин: если *X = C* с вероятностью *p = 1*, то *M (C) = x p =C*.

Для непрерывных случайных величин: $M(C)=\int_{-\infty}^{+\infty}Cp(x)dx = C\int_{-\infty}^{+\infty}p(x)dx = C$ по условию нормировки для плотности вероятностей.

2) *M(CX) = C M(X)*. В самом деле, константу можно вынести из суммы в дискретном случае и из под интеграла в непрерывном случае.
3) *M(X+Y) = M(X) + M(Y)*. (без доказательства).
4) *M(|X|) ≥ |M(X)|* (без доказательства).

Центрированной случайной величиной называется $\overset{0}{X} = X - m_x = X - M(X)$.

Центральный момент s-го порядка.

Для дискретной случайной величины $\mu_s = \sum_{i=1}^{n}(x_i - m_i)^s p_i$.

Для непрерывной случайной величины $\mu_s = \int_{-\infty}^{+\infty}(x - m_x)^s p(x)dx$. $\mu_s = M((\overset{0}{X})^s)$.

Дисперсией называется второй центральный момент случайной величины. $D_x = D(X) = M((\overset{0}{X})^2) = M((X - m_x)^2)$.

По свойствам математического ожидания получим

$D_x = M(X^2) - 2(m_x)^2 + (m_x)^2 = M(X^2) - (m_x)^2$**.** Эта формула часто применяется. Дисперсия – это характеристика рассеяния, она характеризует концентрацию кривой распределения (графика плотности распределения) около математического ожидания. Если на числовой оси расположить точки x_i с массами p_i, то дисперсия – это момент инерции системы материальных точек относительно центра тяжести m_x.

Для дискретных случайных величин $D_x = \sum_{i=1}^{n}(x_i - m_x)^2 p_i$.

Для непрерывных случайных величин $D_x = \int_{-\infty}^{+\infty}(x - m_x)^2 p(x)dx$.

Свойства дисперсии.

1) $D_x \ge 0$ (под интегралом стоит квадрат функции).

2) $D_C = 0$ ($m_C = C$, $m_{C^2} = C^2$)

3) $D_{Cx} = C^2 D_x$ (выведите сами, вынося C^2 из под знака суммы или из-под интеграла).

Средним квадратическим отклонением называется $\sigma(x) = \sqrt{D_x}$.

Кроме этих основных числовых характеристик используются **коэффициент асимметрии** $s_k = \frac{\mu_3}{\sigma^3}$, **эксцесс** – мера островершинности распределения $Ex = \frac{\mu_4}{\sigma^4} - 3$, **среднее арифметическое отклонение** $M(|\overset{0}{X}|)$, **мода** – наиболее вероятное значение для дискретных величин или значение, где плотность максимальна для непрерывных величин, **медиана** *Me(X)* – абсцисса точки, в которой площадь, ограниченная кривой плотности распределения, делится пополам (точка, в которой *F(x) = ½).*

Пример. Стрелок делает один выстрел и с вероятностью p попадает в мишень. Пусть *X* – количество попаданий в мишень. Это – дискретная случайная величина, принимающая два значения $x_1 = 0$ *и* $x_2 = 1$ с вероятностями $q = 1\text{-}p$, p соответственно. Построим ряд распределения X

x_i	0	1
p_i	q	p

Функция распределения равна $F(x) = \begin{cases} 0, & -\infty < x \le 0 \\ q, & 0 < x \le 1 \\ 1 & 1 < x < +\infty \end{cases}$,

Математическое ожидание равно $M(X) = m_x = 0\ q + 1\ p = p$.

Если составить ряд распределения для случайной величины X^2, то мы получим ту же таблицу (так как $0^2 = 0$ и $1^2 = 1$). Поэтому $M(X^2) = p$, а дисперсию можно вычислить по формуле $D(X) = M(X^2) - (m_x)^2 = p - p^2 = p(1\text{-}p) = pq$.

Распределение называется **равномерным на отрезке [*a,b*],** если плотность случайной величины X постоянна на отрезке [a,b] p(x) = p и равна нулю вне этого отрезка.

Из условия нормировки для плотности вероятности следует

$1 = \int_a^b p dx = p(b-a)$. Отсюда следует, что $p = \frac{1}{b-a}$ - **плотность равномерного распределения. Функция распределения величины, распределенной равномерно на отрезке [a,b],** равна

$F(x) = \int_{-\infty}^{x} p(x)dx = \begin{cases} 0, & x \le a \\ \frac{x-a}{b-a} & a < x \le b \\ 1 & x > b \end{cases}$. Вычислим математическое ожидание и дисперсию величины, распределенной равномерно на отрезке $[a,b]$.

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

$$D_x=\int_a^b\left(x-\frac{a+b}{2}\right)^2 p\,dx=p\int_a^b\left(x^2-(a+b)x+\frac{(a+b)^2}{4}\right)dx=p\left(\frac{b^3-a^3}{3}-\frac{(a+b)(b^2-a^2)}{2}+\frac{(a+b)^2(b-a)}{4}\right)=$$

$$\left[\frac{a^2+ab+b^2}{3}-\frac{a^2+2ab+b^2}{2}+\frac{a^2+2ab+b^2}{4}\right]=$$

$$\frac{1}{12}\left(4a^2+4ab+4b^2-6a^2-12ab-6b^2+3a^2+6ab+3b^2\right)=$$

$$=\frac{1}{12}\left(b^2-2ab+a^2\right)=\frac{(b-a)^2}{12}$$

Лекция 4

Повторные испытания

Пусть производится n опытов (испытаний), в каждом из которых может наступить один из N исходов. Если результаты одного испытания не зависят от результатов других испытаний, то такие испытания называются **независимыми.**

Например, стрелок делает n выстрелов в мишень, в которой имеется N областей попадания: десятка, девятка и т.д.

Возможны две ситуации: условия проведения испытаний не меняются (ситуация A) или меняются от испытания к испытанию (ситуация B).

Рассмотрим ситуацию А.

Пусть число исходов равно двум (N = 2). Схема независимых испытаний с двумя исходами называется **схемой Бернулли**.

Два исхода соответствуют в приведенном примере попаданию (успеху) или не попаданию в мишень, причем в каждом выстреле вероятность попадания равна p, а вероятность промаха равна $q = 1 - p$. Обозначим вероятность попасть m раз из n выстрелов $P(m,n)$. $P(0,n)=q^n$, так как в каждом опыте стрелок промахивается. Вероятность попасть один раз равна $P(1,n)=npq^{n-1}$, так как стрелок может попасть при первом, втором, … n ом выстреле. $P(2,n)=C_n^2p^2q^{n-2}$,так как два попадания (порядок не важен) должны быть размещены (выборки без возвращения) среди n выстрелов. Аналогично

$P(m,n)=P_n(m)=C_n^m p^m q^{n-m}$ **- формула Бернулли.**

Распределение $P_n(m)=C_n^m p^m q^{n-m}$ называют **биномиальным.**

В самом деле, $P_n(m)$ – коэффициенты при z^m в разложении по степеням z **производящей функции** $\varphi(z)=(q+pz)^n$.

Из формулы Бернулли вытекают *два следствия*:

1) Вероятность появления успеха в n испытаниях не менее m_1 раз и не более m_2 раз равна $P(m_1 \le m \le m_2) = \sum_{m=m_1}^{m_2} C_n^m p^m q^{n-m}$,

2) Вероятность хотя бы одного успеха в n испытаниях равна $(m_1 = 1, m_2 = n)$
$$R(1,n) = P(m \ge 1) = \sum_{m=1}^{n} C_n^m p^m q^{n-m} = 1 - C_n^0 p^0 q^n = 1 - q^n .$$

Если X имеет биномиальное распределение, то $\boldsymbol{M_x = np, D_x = npq}$.

Пусть в ситуации А число исходов равно N*, а их вероятности равны* $p_1,\dots,p_N$. Вычислим вероятность того, что после n испытаний i –й исход наступит m_i раз $(m_1 + \ldots + m_N = n)$
$$P(m_1 \ldots . m_N, n) = C_n^{m_1} C_{n-m_1}^{m_2} \ldots C_{n-m_1-m_2-\ldots-m_{N-1}}^{m_N} p_1^{m_1} p_2^{m_2} \ldots p_N^{m_N} .$$

Заметим, что
$$C_n^{m_1} C_{n-m_1}^{m_2} \ldots C_{n-m_1-m_2-\ldots-m_{N-1}}^{m_N} = \frac{n!}{m_1!(n-m_1)!} \frac{(n-m_1)!}{m_2!(n-m_1-m_2)!} \cdots \frac{(n-m_1-\ldots-m_{N-1})!}{m_N!(n-m_1-\ldots m_N)!} =$$
$$\frac{n!}{m_1! m_2! \ldots m_N!},$$
так как $n = m_1 + \ldots + m_N$.

Поэтому $P(m_1,\ldots\ldots,m_N, n) = \dfrac{n!}{m_1! m_2! \ldots m_N!} p_1^{m_1} p_2^{m_2} \ldots p_N^{m_N}$. Это – **полиномиальное распределение.**

$P(m_1,\ldots,m_N,n)$ - это коэффициенты при $z_1^{m_1} \ldots z_N^{m_N}$ в разложении по степеням $z_1 \ldots z_n$ **производящей функции** $\varphi(z_1 \ldots z_N) = (z_1 p_1 + \ldots + z_N p_N)^n$.

Рассмотрим ситуацию В. Здесь вероятность того или иного исхода зависит от номера испытания, так как условия испытаний различны. $P(m_1,\ldots m_N,n)$ - это коэффициенты при $z_1^{m_1} \ldots z_N^{m_N}$ в разложении по степеням $z_1 \ldots z_n$ **производящей функции** $\varphi(z_1 \ldots z_N) = \prod_{i=1}^{n} (z_1 p_{1i} + \ldots + z_N p_{Ni})$ при N исходах.

При двух исходах $P(m,n)$ - это коэффициент при z^m в разложении производящей функции
$$\varphi(z) = \prod_{i=1}^{n} (q_i + p_i z), \text{ где } q_i + p_i = 1, i = \overline{1,n} .$$

Примеры.

1) Какова вероятность с пяти раз вытащить из колоды в 36 карт а)три туза, б)не менее одного туза?

а) $P(3,5)=C_5^3\left(\frac{1}{9}\right)^3\left(\frac{8}{9}\right)^2$, б) $R(1,5)=1-P(0,5)=1-\left(\frac{8}{9}\right)^5$.

2) Мишень для опытного стрелка содержит три области попадания: 10, 9, пусто. Вероятность попасть при одном выстреле в десятку – 0,2, в девятку – 0,7, в «пусто» – 0,1. Какова вероятность в серии из 10 выстрелов попасть в «пусто» два раза, в девятку 4 раза, в десятку 4 раза?

$$P(2,4,4,10)=\frac{10!}{2!\,4!\,4!}(0{,}1)^2(0{,}7)^4(0{,}2)^4.$$

3) Производится три выстрела в мишень. При первом выстреле вероятность попасть в мишень равна 0,5, не попасть 0,5. При втором выстреле – соответственно 0,4 и 0,6, при третьем выстреле 0,3 и 0,7. Какова вероятность два раза попасть в мишень?

$$\varphi(z)=(0{,}5+0{,}5z)(0{,}6+0{,}4z)(0{,}7+0{,}3z)=0{,}21+0{,}44z+0{,}29z^2+0{,}06z^3.$$

Вероятность не попасть ни разу 0,21, один раз – 0,44, два раза – 0,29, три раза – 0,06.

Распределения, связанные с повторными испытаниями

Геометрическое распределение

Рассмотрим схему Бернулли. Обозначим X – число испытаний до первого успеха, если вероятность успеха в одном испытании p. Если первое испытание успешно, то $X = 0$. Следовательно, $P(X=0)=p$. Если $X = 1$, т.е. первое испытание неудачно, а второе успешно, то по теореме умножения $P(X=1)=qp$. Аналогично, если $X = n$, то все испытания до n-ого неудачны и $P(X=n)=q^n p$. Составим ряд распределения случайной величины X

X	0	1	2	…	n	…
p	p	qp	$q^2 p$	…	$q^n p$	…

Случайная величина с таким рядом распределения имеет **геометрическое распределение.**

Проверим условие нормировки $p+qp+q^2p+...+q^np+...=p(1+q+q^2+...)=\frac{p}{1-q}=\frac{p}{p}=1$.

Гипергеометрическое распределение

Рассмотрим схему испытаний, обобщающую задачу о выборке бракованных деталей и похожую на ситуацию A с N исходами. Пусть имеется n элементов, разделенных на группы: n_1 элементов первого типа, n_2 – второго типа и т.д., n_N – N-ого типа. Какова вероятность, выбрав m элементов, получить среди них m_1 элементов из первой группы, m_2 – из второй и т.д. m_N - из N-ой?

Ее легко вычислить по классическому определению вероятностей с учетом теоремы умножения:

$$P(m_1, m_2, ..m_N, n) = \frac{C_{n_1}^{m_1} C_{n_2}^{m_2} ... C_{n_N}^{m_N}}{C_n^m}.$$

В частности, при $N = 2$ ($m_2 = m - m_1$, $n_2 = n - n_1$) (задача о бракованных деталях)

$$P(m_1, n) = \frac{C_{n_1}^{m_1} C_{n-n_1}^{m-m_1}}{C_n^m}.$$

Формула Пуассона и распределение Пуассона

Пусть число испытаний n велико, вероятность успеха p мала и $\lambda = np$ мало. Тогда вероятность наступления m успехов в n испытаниях можно приближенно определить по **формуле Пуассона**: $P(m,n) \approx \frac{\lambda^m}{m!} e^{-\lambda}$.

Заметим, что по формуле Пуассона можно считать вероятность неуспеха, если q мало, приняв $\lambda = nq$.

Случайная величина с рядом распределения m, $P(m,n) \approx \frac{\lambda^m}{m!} e^{-\lambda}$ имеет **распределение Пуассона**. Чем больше n, тем формула Пуассона точнее. Для грубых расчетов формулу применяют при n =10, $\lambda = 0 - 2$, при n = 100 $\lambda = 0 - 3$. При инженерных расчетах формулу применяют при n = 20, $\lambda = 0 - 3$, n =100, $\lambda = 0 - 7$. При точных расчетах формулу применяют при n = 100, $\lambda = 0 - 7$, n =1000, $\lambda = 0 - 15$.

Вычислим математическое ожидание и дисперсию случайной величины, имеющей распределение Пуассона.

$$M(X) = \sum_{m=0}^{\infty} m \frac{\lambda^m}{m!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{m=1}^{\infty} \frac{\lambda^{m-1}}{(m-1)!} = \lambda e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda e^{-\lambda} e^{\lambda} = \lambda,$$

$$M(X(X-1)) = \sum_{m=0}^{\infty} m(m-1) \frac{\lambda^m}{m!} e^{-\lambda} = \lambda^2 \sum_{m=2}^{\infty} \frac{\lambda^{m-2}}{(m-2)!} e^{-\lambda} = \lambda^2$$

$$M(X(X-1)) = M(X^2) - M(X) = \lambda^2, \quad M(X^2) = \lambda^2 + \lambda,$$

$$D(X) = M(X^2) - (M(X))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Лекция 5

Экспоненциальное и нормальное распределения

Экспоненциальное распределение.

Непрерывная случайная величина имеет экспоненциальное распределение, если ее плотность распределения задается формулой

$p(x)=\begin{cases}0, & x<0\\ \lambda e^{-\lambda x}, & x\ge 0\end{cases}$, $\lambda>0$ - параметр экспоненциального распределения.

Для случайной величины, имеющей экспоненциальное распределение, $M(X)=\dfrac{1}{\lambda}$, $D(X)=\dfrac{1}{\lambda^2}$, $F(x)=\begin{cases}0, & x<0\\ 1-e^{-\lambda x}, & x\ge 0\end{cases}$.

Если времена между последовательными наступлениями некоторого события – независимые, экспоненциально распределенные случайные величины с параметром λ, то число наступлений этого события за время t имеет пуассоновское распределение с параметром λt. Геометрическое распределение является дискретным аналогом экспоненциального распределения.

Нормальное распределение (распределение Гаусса).

Непрерывная случайная величина имеет **нормальное распределение** (распределена нормально или по Гауссу), если ее плотность имеет вид

$$p(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(x-a)^2}{2\sigma^2}},\ x\in R.$$

Вычислим математическое ожидание и дисперсию нормально распределенной случайной величины.

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

$$\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}ye^{-\frac{y^2}{2}}dy+a\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{y^2}{2}}dy=0+a\frac{1}{\sqrt{2\pi}}\sqrt{2\pi}=a.$$

Вычислите аналогично $D(X)=\sigma^2$, среднеквадратическое отклонение $=\sigma$.

Обозначим плотность стандартного нормального распределения (при $M(X)=a=0,\quad D(X)=\sigma^2=1$) $\qquad \varphi(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$,

обозначим функцию распределения стандартного нормального распределения

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{x^2}{2}}dx=\frac{1}{2}+\Phi_0(x),$$

где $\Phi_0(x)=\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_0^x e^{-\frac{x^2}{2}}dx$ - **интеграл Лапласа.** Значения $\Phi_0(x)$ можно найти в стандартных таблицах.

Вычислим вероятность попадания нормально распределенной случайной величины на отрезок $[a,b]$.

$$P\{a<X<b\}=\frac{1}{\sigma\sqrt{2\pi}}\int_a^b e^{-\frac{(y-m)^2}{2\sigma^2}}dy \xrightarrow[x=(y-m)/\sigma]{} \frac{1}{\sqrt{2\pi}}\int_{(a-m)/\sigma}^{(b-m)/\sigma} e^{-\frac{x^2}{2}}dx=$$

$=\Phi_0\left(\frac{b-m}{\sigma}\right)-\Phi_0\left(\frac{a-m}{\sigma}\right)$. При вычислении вероятности полезно учитывать нечетность функции $\Phi_0(x)$:

$$\Phi_0(-x)=\frac{1}{\sqrt{2\pi}}\int_0^{-x} e^{-\frac{x^2}{2}}dx =_{y=-x}= -\int_0^{y} e^{-\frac{y^2}{2}}dy=-\Phi_0(x).$$

Локальная и интегральная формулы Муавра – Лапласа

Если в схеме Бернулли число испытаний n велико, причем p и $q = 1 - p$ велики, то для всех m справедливы: **локальная формула Муавра – Лапласа**

$$P(m,n)\sqrt{npq}\approx\varphi(x),\quad x=\frac{m-np}{\sqrt{npq}}$$

и **интегральная формула Муавра – Лапласа**

$$P(m_1\le m\le m_2)=P\left(\frac{m_1-np}{\sqrt{npq}}\le\frac{m-np}{\sqrt{npq}}\le\frac{m_2-np}{\sqrt{npq}}\right)\approx\Phi(x_2)-\Phi(x_1)=\Phi_0(x_2)-\Phi_0(x_1),$$

$$x_1=\frac{m_1-np}{\sqrt{npq}},\ x_2=\frac{m_2-np}{\sqrt{npq}}.$$

Это означает, что при большом числе испытаний распределение числа успехов становится нормальным.

Иногда приходится оценивать вероятность отклонения частоты события от вероятности. Покажем, как можно использовать для этого интегральную формулу Муавра – Лапласа.

Заметим, что $\frac{m-np}{\sqrt{npq}}=\left(\frac{m}{n}-p\right)\sqrt{\frac{n}{pq}}$. Запишем интегральную формулу Муавра – Лапласа

$P\left(a<\frac{m-np}{\sqrt{npq}}<b\right)\approx\frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{t^2}{2}}dt$ в виде

$P\left(a\sqrt{\frac{n}{pq}}<\left(\frac{m}{n}-p\right)\sqrt{\frac{n}{pq}}<b\sqrt{\frac{n}{pq}}\right)\approx\frac{1}{\sqrt{2\pi}}\int_{a\sqrt{\frac{n}{pq}}}^{b\sqrt{\frac{n}{pq}}} e^{-\frac{t^2}{2}}dt$. Поэтому

$P\left(a<\left(\frac{m}{n}-p\right)<b\right)\approx\Phi_0\left(b\sqrt{\frac{n}{pq}}\right)-\Phi_0\left(a\sqrt{\frac{n}{pq}}\right)$. Если интервал симметричен, $-\varepsilon=a,\quad b=\varepsilon$,

то по нечетности $\Phi_0(x)$ $\quad P\left(\left|\frac{m}{n}-p\right|<\varepsilon\right)\approx2\Phi_0\left(\varepsilon\sqrt{\frac{n}{pq}}\right)$.

Примеры.

1) (3.42 /1/) Телефонная станция обслуживает 1000 абонентов. Вероятность поступления каждого вызова за минуту 0,0005. Какова вероятность, что за минуту поступит не менее двух вызовов? Здесь n = 1000, p = 0,0005, $\lambda = np$ =0.5. $P \approx 1 - P(0,\ 0{,}5) - P(1,\ 0{,}5) = 1 - 0{,}606 - 0{,}303 = 0{,}091$ (по таблице $P(m, \lambda)$ /1/).

2) (3.43 /1/) Известно, что 20% автомобилей нарушают скоростной режим. Какова вероятность того, что из 1000 автомобилей 210 нарушат правила? Здесь надо пользоваться локальной формулой Муавра-Лапласа при n = 1000, p = 0,2, m = 300. $\sqrt{npq} = \sqrt{1000 \cdot 0{,}2 \cdot 0{,}8} = 12{,}65, \quad x = \frac{210 - 200}{12{,}65} = 0{,}79, \quad \varphi(0{,}79) = 0{,}292, \ P \approx \frac{0{,}292}{12{,}65} \approx 0{,}02$.

3) (3.44 /1/) Монету подбрасывают 10000 раз. Найти вероятность того, что частота выпадения герба будет отличаться от 0,5 не более, чем на 2%. Здесь надо пользоваться интегральной формулой Муавра-Лапласа при n = 10000, p = 1/2, m_1 = 400, m_2 = 600. Тогда $\sqrt{npq} = \sqrt{10000 \cdot 0{,}5 \cdot 0{,}5} = 50, \quad x_1 = -2, x_2 = 2, \quad P \approx \Phi_0(2) - \Phi_0(-2) = 2 \cdot 0{,}47725 = 0{,}9545.$

Другие распределения, часто используемые в инженерных расчетах

Распределение Вейбулла. Это распределение с плотностью

$$p(x) = \begin{cases} 0, & x < 0 \\ \alpha\, \beta\, x^{\beta-1} e^{-\alpha\, x^{\beta}}, & x \geq 0 \quad (\alpha > 0, \beta > 0) \end{cases}.$$

и функцией распределения

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

Если $\beta = 1$, то распределение Вейбулла превращается в *экспоненциальное*, а при $\beta = 2$ - в **распределение Релея.**

Достаточно близкую к распределению Вейбулла плотность имеет **гамма – распределение:**

$$p(x) = \begin{cases} 0, & x < 0 \\ \dfrac{\lambda^{\gamma} x^{\gamma-1}}{\Gamma(\gamma)} e^{-\lambda x}, & x \geq 0 \quad (\lambda > 0, \gamma > 0) \end{cases}.$$

Здесь $\Gamma(\gamma) = \int\limits_0^{+\infty} x^{\gamma-1} e^{-x} dx$ - гамма-функция.

Если $\gamma = k$ - целое число, то гамма-распределение превращается в **распределение Эрланга** порядка k. Если k – нечетное число, $\gamma = \frac{k}{2}, \lambda = \frac{1}{2}$, то гамма-распределение превращается в **распределение** $\aleph^2$ **(хи-квадрат) распределение** с k степенями свободы. При $\gamma = 1$ (так как $\Gamma(\gamma + 1) = \gamma\Gamma(\gamma)$, $\Gamma(n) = (n-1)!$) гамма-распределение переходит в экспоненциальное. Для всех рассмотренных распределений составлены таблицы, по которым можно определять значения функций распределения.

Лекция 6.

Двумерные случайные величины

Совокупность двух случайных величин *(X,Y),* заданных на вероятностном пространстве (Ω, S, P) (или $(\Omega, \Sigma, \mathrm{P})$) , называют **двумерной случайной величиной** или **двумерным случайным вектором**; X, Y называют координатами случайного вектора.

Это определение можно обобщить и на совокупность n случайных величин.

Функцией распределения случайного вектора (X,Y) или **совместной функцией распределения** случайных величин X, Y называется

$$F(x,y) = P\{X < x, Y < y\}.$$

Свойства функции распределения.

1 $0 \le F(x,y) \le 1$ (Это – свойство вероятности, а $F(x,y)$ - вероятность).

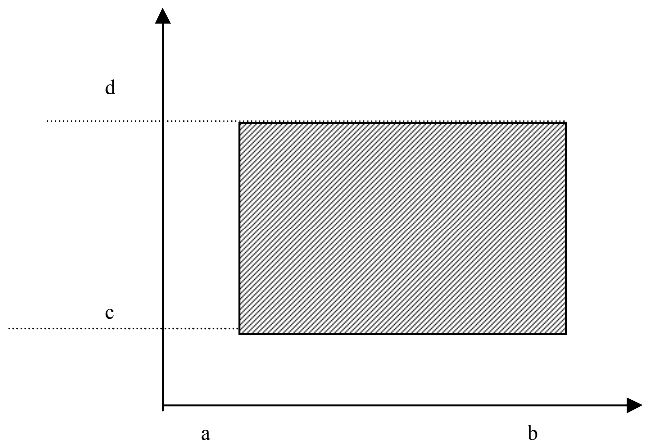
2 $F(x,y)$ - неубывающая функция по каждому из своих аргументов. (В самом деле, если $x_1 < x_2$, то событие $(X < x_1, Y < y)$ включено в событие $(X < x_2, Y < y)$, следовательно, $F(x_1,y) \le F(x_2,y)$).

3 $F(-\infty,y) = F(x,-\infty) = 0$ (события $(X < -\infty, Y < y)$, $(X < x, Y < -\infty)$ - невозможные, поэтому их вероятность равна нулю).

4 $F(+\infty,+\infty) = 1$ (событие $(X < +\infty, Y < +\infty)$ достоверно).

5 $P(a \le X \le b, c \le Y \le d) = F(b,d) - F(b,c) - F(a,d) + F(a,c)$

Геометрически, $F(b,d)$- площадь полосы левее и ниже точки (b,d), Вычитая из нее $F(b,c)$ и $F(a,d)$, мы *два раза* вычтем площадь полосы левее и ниже точки (a,c). Для того, чтобы получить площадь прямоугольника – левую часть равенства, надо вычитать эту площадь один раз, поэтому надо добавить ее, т.е. $F(a,c)$ в правую часть равенства.



6. $F(x,y)$непрерывна слева по каждому из аргументов

7. $F(x,+\infty)=F_X(x), F(+\infty, y)=F_Y(y)$. Так как событие $(X<+\infty)$ достоверно, то пересечение событий $(X<+\infty)$ и $(Y<y)$ есть событие $(Y<y)$. Поэтому первое равенство справедливо. Аналогично доказывается справедливость второго равенства.

Двумерная случайная величина *(X , Y)* **дискретна**, если *X, Y* - дискретные случайные величины. Для нее составляется таблица распределения – аналог ряда распределения для одномерной случайной величины. $p_{i,j} = P(X = x_i, Y = y_j)$,

X	Y				
	y_1	y_2	…..	y_m	P_X
x_1	p_{11}	p_{12}	…	p_{1m}	p_{X1}
x_2	p_{21}	p_{22}	…	p_{2m}	p_{X2}
…….	…	…	…	…	…
x_n	p_{n1}	p_{n2}	…	p_{nm}	p_{Xn}
P_Y	p_{Y1}	p_{Y2}	…	p_{Ym}	

$p_{nm} = P(X = x_n, Y = y_m)$, $p_{Ym} = P(Y = y_m) = p_{1m} + p_{2m} + \ldots + p_{nm}$, $p_{Xn} = p_{n1} + p_{n2} + \ldots + p_{nm}$.

График функции распределения для двумерной случайной величины напоминает «лестницу», уровень ступеней которой изменяется скачком на p_{ij} при переходе через точку (x_i, y_j) в положительном направлении по оси OX и по оси OY. Если зафиксировать $x = x_i$, то при увеличении у эти скачки будут на $p_{i1}, p_{i2}, \ldots p_{im}$ (от нуля до p_{Xi}). Если зафиксировать $y = y_j$, то при увеличении x скачки будут на $p_{1j}, p_{2j}, \ldots p_{nj}$ (от нуля до p_{Yj}). Нижние ступени (при $x \le x_1$ *и* $y \le y_1$) находятся на нулевом уровне, самая верхняя ступень (при $x > x_n$, $y > y_m$) находится на уровне 1. Если зафиксировать $x > x_n$ то при увеличении у эти скачки будут на $p_{Y1}, p_{Y2}, \ldots p_{Ym}$ (от нуля до 1). Если зафиксировать $y > y_m$, то при увеличении x скачки будут на $p_{X1}, p_{X2}, \ldots p_{Xn}$ (от нуля до 1).

***Пример*.** Проводятся два выстрела в мишень. При каждом выстреле вероятность попадания p, вероятность промаха $q = 1 - p$. Случайная величина X_i – число попаданий при i – том выстреле. Найдем закон распределения случайного вектора $(X_1, X_2) = (X, Y)$.

X	Y		
	$y_1=0$	$y_2=1$	P_X
$x_1=0$	q^2	qp	$p_{X1}=q$
$x_2=1$	pq	p^2	$p_{X2}=p$
P_Y	$p_{Y1}=q$	$p_{Y2}=p$	

Построим функцию распределения

$$F(x)=\begin{cases} 0, & x\le 0, y\le 0 \\ q^2, & (0<x\le 1, 0<y\le 1) \\ q, & (0<x\le 1, y>1) \\ q, & (x>1, 0<y\le 1) \\ 1, & (x>1, y>1) \end{cases}.$$ В самом деле, при $(x\le 0, y\le 0)$ – событие $\{X<x, Y<y\}$ - невозможное, при $(x>1,\ y>1)$ событие $\{X<x, Y<y\}$ – достоверное.

При $(0<x\le 1, 0<y\le 1)$ событие $\{X<x, Y<y\}$ представляет собой событие $\{X=0, Y=0\}$. Поэтому при $(0<x\le 1, 0<y\le 1)$ $F(x) = P\{X=0, Y=0\} = q^2$.

При $(0<x\le 1, y>1)$ событие $\{X<x, Y<y\}$ представляет собой объединение несовместных событий $\{X=0, Y=0\}$ и $\{X=0, Y=1\}$. Поэтому при $(0<x\le 1, y>1)$ $F(x) = P\{X=0, Y=0\} + P\{X=0, Y=1\} = q^2 + pq = q(p+q) = q$. Аналогично, в случае $(x>1, 0<y\le 1)$ $F(x) = P\{X=0, Y=0\} + P\{X=1, Y=0\} = q^2 + pq = q(p+q) = q$.

Двумерная случайная величина **непрерывна**, если *X, Y* - непрерывные случайные величины и ее функцию распределения можно представить в виде сходящегося несобственного интеграла от плотности распределения.

$$F(x,y)=P(X<x, Y<y)=\int_{-\infty}^{x}\int_{-\infty}^{y} p(x,y)dxdy.$$

Двойной интеграл можно записать в виде повторных (внешний по *x*, внутренний по *y* и наоборот). Если предполагать непрерывность плотности по x и y, то, дифференцируя по переменным верхним пределам, получим

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

Свойства плотности

1. $p(x,y)\ge 0$ (функция распределения – неубывающая функция).
2. $P(a<X<b, c<Y<d)=\int_a^b\int_c^d p(x,y)dxdy$ (по свойству 5 функции распределения).

   Справедливо обобщение $P(X\in D)=\iint_D p(x,y)dxdy$.
3. $P(x<X<x+\Delta x, y<Y<y+\Delta y)\approx p(x,y)\Delta x\Delta y$.
4. $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x,y)dxdy=1$ (по свойству 4 функции распределения).
5. $P(X=x, Y=y)=0$.
6. $p_X(x)=\int_{-\infty}^{+\infty} p(x,y)dy$, $p_Y(y)=\int_{-\infty}^{+\infty} p(x,y)dx$ (Свойство 7 функции распределения).

Независимость случайных величин.

Случайные величины *X, Y* называются **независимыми**, если $F(x,y)=F_X(x)F_Y(y)$, где $F_X(x)$, $F_Y(y)$ - функции распределения случайных величин *X, Y*.

Если случайные величины непрерывны, то, дифференцируя это соотношение по *x, y*, получим $p(x,y)=\dfrac{\partial^2 F(x,y)}{\partial x \partial y}=F_X'(x)F_Y'(y)=p_X(x)p_Y(y)$.

Соотношение $p(x,y)=p_X(x)p_Y(y)$ поэтому можно считать **определением независимости непрерывных случайных величин.**

Для **дискретных** случайных величин определение **независимости** можно записать в виде $p_{ij}=P(X=x_i, Y=y_j)=P(X=x_i)P(Y=y_j)=p_{X\,i}p_{Y\,j}$.

Математическое ожидание.

Математическим ожиданием функции двумерной случайной величины называется

$M(f(X,Y))=\sum\limits_{i,j} f(x_i,y_j)p_{ij}$ в дискретном случае,

$M(f(X,Y))=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} f(x,y)p(x,y)dxdy$ в непрерывном случае.

Свойства математического ожидания.

1. $M(C)=C$ $\left(M(C)=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} Cp(x,y)dxdy=C\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} p(x,y)dxdy=C\right.$ по условию нормировки)

2. $M(X+Y)=M(X)+M(Y)$, $M(CX)=CM(X)$, $M(CY)=CM(Y)$.

$$M(X+Y)=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}(x+y)p(x,y)dxdy=\int\limits_{-\infty}^{+\infty}x\left(\int\limits_{-\infty}^{+\infty}p(x,y)dy\right)dx+\int\limits_{-\infty}^{+\infty}y\left(\int\limits_{-\infty}^{+\infty}p(x,y)dx\right)dy=$$

$$\int\limits_{-\infty}^{+\infty}xp_X(x)dx+\int\limits_{-\infty}^{+\infty}yp_Y(y)dy=M(X)+M(Y).$$

$$M(CX)=\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}Cxp(x,y)dxdy=C\int\limits_{-\infty}^{\infty}x(\int\limits_{-\infty}^{\infty}p(x,y)dy)dx=C\int\limits_{-\infty}^{\infty}xp_X(x)dx=CM(X).$$

3 $M(XY)=M(X)M(Y)$ **для независимых** случайных величин.

$$M(XY)=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}(xy)p(x,y)dxdy=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}(xy)p_X(x)p_Y(y)dxdy=\int\limits_{-\infty}^{+\infty}xp_X(x)dx\int\limits_{-\infty}^{+\infty}yp_Y(y)dy=M(X)M(Y).$$

Ковариация (корреляционный момент).

Ковариацией случайных величин X, Y называют $\operatorname{cov}(X, Y) = M\big((X - M(X))(Y - M(Y))\big)$.

Свойства ковариации.

1. $\operatorname{cov}(X, Y) = M(XY) - M(X)M(Y)$.

$\operatorname{cov}(X, Y) = M\big((X - M(X))(Y - M(Y))\big) = M\big(XY - M(X)Y - XM(Y) + M(X)M(Y)\big) =$
$M(XY) - M(X)M(Y) - M(X)M(Y) + M(X)M(Y) = M(XY) - M(X)M(Y)$.

2. $\operatorname{cov}(X, X) = D(X)$.

По свойству 1 $\operatorname{cov}(X, X) = M(X^2) - (M(X))^2 = D(X)$.

3. Если *X, Y* независимы, то $\operatorname{cov}(X, Y) = 0$, (обратное неверно).

Если случайные величины независимы, то $M(XY) = M(X)M(Y)$, тогда по свойству 1 $\operatorname{cov}(X, Y) = 0$.

Случайные величины *X, Y* называются **некоррелированными,** если $\operatorname{cov}(X, Y) = 0$, **из некоррелированности не следует независимость**, **из независимости следует некоррелированность.**

4. $\operatorname{cov}((aX + b), (cY + d)) = ac\operatorname{cov}(X, Y)$.

По свойству 1

$\operatorname{cov}((aX + b), (cY + d)) = M((aX + b)(cY + d)) - M(aX + b)M(cY + d) =$

$acM(XY) + bcM(Y) + adM(X) + bd - acM(X)M(Y) - bcM(Y) - daM(X) - bd =$
$ac(M(XY) - M(X)M(Y)) = ac\operatorname{cov}(X, Y)$.

5. $|\operatorname{cov}(X, Y)| \le \sqrt{D(X)D(Y)}$ **.**

Рассмотрим случайную величину $z = aX + Y$. $D(z) = D(aX + Y) =$

$M[(aX + Y) - M(aX + Y)]^2 = M[a(X - M(X)) + (Y - M(Y))]^2 =$

$M\left[a^2(X - M(X))^2 + 2a(X - M(X))(Y - M(Y)) + (Y - M(Y))^2\right] =$

$a^2 D(X) + 2a\operatorname{cov}(X, Y) + D(Y)$.

Заметим, что отсюда следует **свойство дисперсии** (при a *=1*)

$D(X + Y) = D(X) + 2\operatorname{cov}(X, Y) + D(Y)$.

Так как $D(z) \ge 0$, то $a^2 D(X) + 2a\operatorname{cov}(X, Y) + D(Y) \ge 0$. Это возможно только, если дискриминант этого квадратного трехчлена относительно a меньше или равен нулю. Выпишем это требование к дискриминанту:

$(\operatorname{cov}(X, Y))^2 - D(X)D(Y) \le 0$. Отсюда следует свойство 5.

6. Для того, чтобы случайные величины были **линейно зависимы** ($Y = aX + b$), **необходимо и достаточно, чтобы** $\left|\operatorname{cov}(X,Y)\right| = \sqrt{D(X)D(Y)}$ **.**

Необходимость. Пусть $Y = aX + b$. Тогда

$D(Y) = M(((aX+b)-(aM(X)+b))^2) = M(a^2(X-M(X))^2) = a^2 D(X)$.

$\operatorname{cov}(X,Y) = M((X-M(X))(Y-M(Y))) = M((X-M(X))(aX+b-aM(X)-b) =$
$aM((X-M(X))^2) = aD(X)$. $\left|\operatorname{cov}(X,Y)\right| = |a|D(X) = \sqrt{D(X)}\sqrt{a^2D(X)} = \sqrt{D(X)}\sqrt{D(Y)}$.

Достаточность. Пусть $\left|\operatorname{cov}(X,Y)\right| = \sqrt{D(X)D(Y)}$. Тогда (доказательство свойства 5) $D(z) = D(aX+Y) = 0$, следовательно, z - детерминированная величина, т.е. $aX + Y = const$, поэтому величины X, Y – линейно зависимы.

Коэффициентом корреляции называется $\rho = \dfrac{\operatorname{cov}(X,Y)}{\sqrt{D(X)D(Y)}}$.

Свойства коэффициента корреляции.

1. $\rho(X,X) = 1$.
2. Если X, Y – независимы, то $\rho(X,Y) = 0$.
3. $\rho(aX+b, cY+d) = sign(ac)\rho(X,Y)$.
4. $-1 \le \rho(X,Y) \le 1$.
5. $\left|\rho(X,Y)\right| = 1$ тогда и только тогда, когда X,Y линейно зависимы.

Двумерное равномерное распределение.

Случайный вектор (X, Y) равномерно распределен в области D (площадь D равна S), если его плотность распределения задана так: $p(x,y) = 0$, если $(x,y) \notin D$, $p(x,y) = 1/S$, если $(x, y) \in D$.

Пример. Случайный вектор (X,Y) равномерно распределен в прямоугольнике $0 \le x \le a$, $0 \le y \le b$.

$M(X) = \dfrac{1}{ab}\int_0^a dx \int_0^b x dy = \dfrac{a}{2}$, аналогично $M(Y) = \dfrac{b}{2}$.

$M(X^2) = \dfrac{1}{ab}\int_0^a dx \int_0^b x^2 dy = \dfrac{a^2}{3}$, аналогично $M(Y^2) = \dfrac{b^2}{3}$.

$D(X) = M(X^2) - (M(X))^2 = \dfrac{a^2}{3} - \dfrac{a^2}{4} = \dfrac{a^2}{12}$, аналогично $D(Y) = \dfrac{b^2}{12}$.

$M(XY) = \dfrac{1}{ab}\int_0^a dx \int_0^b xy dy = \dfrac{1}{ab}\dfrac{a^2}{2}\dfrac{b^2}{2} = \dfrac{ab}{4}$.

Следовательно, случайные величины X, Y не коррелированны.

Двумерное нормальное распределение.

Двумерная случайная величина (X,Y) распределена нормально со средними значениями m_1, m_2, дисперсиями σ_1^2, σ_2^2 и коэффициентом корреляции ρ, если ее плотность задана

$$p(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{\frac{-1}{2(1-\rho^2)}\left(\frac{(x-m_1)^2}{\sigma_1^2}-2\rho\frac{(x-m_1)(y-m_2)}{\sigma_1\sigma_2}+\frac{(y-m_2)^2}{\sigma_2^2}\right)\right\}.$$

Задача линейного прогноза.

Заданы характеристики $m_1, m_2, \sigma_1, \sigma_2, \rho$ случайного вектора (X_1, X_2). Вводится случайная величина – оценка $\widetilde{X}=aX_1+b$ - линейный прогноз X_2. Вычислить a, b, чтобы линейный прогноз был наилучшим среднеквадратическим (в смысле минимума погрешности оценки: $M((\widetilde{X}-X_2)^2)\to\min$).

$$M((\widetilde{X}-X_2)^2)=D(\widetilde{X}-X_2)+(M(\widetilde{X}-X_2))^2=D(\widetilde{X})-2\operatorname{cov}(\widetilde{X},X_2)+D(X_2)+((M(\widetilde{X})-M(X_2))^2=$$
$$a^2\sigma_1^2-2a\rho\sigma_1\sigma_2+\sigma_2^2+(am_1+b-m_2)^2.$$

За счет выбора b можно лишь минимизировать последнее слагаемое, сделав его нулем: $\underline{b=m_2-am_1}$. Теперь остается обеспечить минимум квадратного трехчлена от a (найти вершину параболы): $\underline{a=-\frac{-2\rho\sigma_1\sigma_2}{2\sigma_1^2}=\rho\frac{\sigma_2}{\sigma_1}}$. Подставляя это значение, найдем

$\underline{b=m_2-\rho\frac{\sigma_2}{\sigma_1}m_1}$. Вычислим погрешность оценки при этих значениях параметров

$$M((\widetilde{X}-X_2)^2)=(\rho\frac{\sigma_2}{\sigma_1})^2\sigma_1^2-2\rho^2\frac{\sigma_2}{\sigma_1}\sigma_1\sigma_2+\sigma_2^2=\sigma_2^2(1-\rho^2).$$

При линейной зависимости X_1, X_2 $(|\rho|=1)$ оценка точна, погрешность равна нулю.

Чем меньше коэффициент корреляции, тем грубее оценка. В крайнем случае, при отсутствии корреляции ($\rho=0$) $a=0, b=m_2$, $\widetilde{X}=m_2$, $M((\widetilde{X}-X_2)^2)=\sigma_2^2$.

Лекция 7.

Законы больших чисел и центральная предельная теорема.

Неравенства Чебышева.

Первое неравенство Чебышева. Пусть случайная величина $X\geq 0$ и существует ее математическое ожидание $M(X)$. Тогда для любого $\varepsilon>0$ выполнено первое неравенство Чебышева $P(X\geq\varepsilon)\leq\frac{M(X)}{\varepsilon}$.

Доказательство. В дискретном случае $M(X)=\sum_{k=1}^{n} x_k P(X=x_k) \geq \sum_{(x_k \geq \varepsilon)} x_k P_k \geq \varepsilon \sum_{(x_k \geq \varepsilon} P_k = \varepsilon\, P(X \geq \varepsilon)$.

В непрерывном случае $M(X)=\int_{-\infty}^{+\infty} xp(x)dx \geq \int_{(x \geq \varepsilon)} xp(x)dx \geq \varepsilon \int_{(x \geq \varepsilon)} xp(x)dx \geq \varepsilon\, P(x \geq \varepsilon)$.

Второе неравенство Чебышева. Пусть существуют математическое ожидание и дисперсия случайной величины $M(X)=m, D(X)=D$. Тогда для любого $\alpha > 0$ выполнено второе неравенство Чебышева $P(|X-m| \geq \alpha) \leq \dfrac{D}{\alpha^2}$

Доказательство проведем для непрерывного случая**,** для дискретного случая оно доказывается аналогично.

$$D=\int_{-\infty}^{+\infty}(x-m)^2 p(x)dx=\int_{-\infty}^{+\infty}|x-m|^2 p(x)dx \geq \int_{|x-m| \geq \alpha} |x-m|^2 p(x)dx \geq \alpha^2 \int_{|x-m| \geq \alpha} p(x)dx \geq \alpha^2 P(|X-m| \geq \alpha)$$

Последовательность случайных величин *сходится по вероятности* к числу a $\left(\{X_n\} \xrightarrow{P} a\right)$, если $\forall \varepsilon > 0, \delta > 0\ \exists N(\varepsilon,\delta) : \forall n > N \Rightarrow P(|X_n - a| < \varepsilon) > 1-\delta$.

Законы больших чисел.

Законы больших чисел могут быть записаны в разных формах, но суть их состоит в том, что при большом числе случайных явлений их средний результат практически перестает быть случайным.

Теорема Чебышева

(Закон больших чисел в форме Чебышева)

При достаточно большом количестве *независимых* опытов среднее арифметическое наблюденных значений случайной величины сходится по вероятности к ее математическому ожиданию. $\dfrac{\sum_{k=1}^{n} X_k}{n} \underset{n\to\infty}{\xrightarrow{P}} m$.

Доказательство. Рассмотрим $Y=\dfrac{\sum_{k=1}^{n} X_k}{n}$, $M(Y)=\dfrac{mn}{n}=m$, $D(Y)=\dfrac{nD(X)}{n^2}=\dfrac{D}{n}$.

Тогда по второму неравенству Чебышева $P(|Y-m| \geq \varepsilon) \leq \dfrac{D(Y)}{\varepsilon^2}=\dfrac{D}{n\varepsilon^2}$.

Если выбрать $N = E(\frac{D}{\delta\varepsilon^2})$, ($E(\,)$ - целая часть), то при $n>N$ будет $\frac{D}{n\varepsilon^2} < \frac{D}{N\varepsilon^2} = \delta$, следовательно, при $n>N$ будет выполнено неравенство $P\{|Y-m| \geq \varepsilon\} < \delta$, поэтому при тех же значениях n будет $P\{|Y-m| < \varepsilon\} \geq 1-\delta$.

Следовательно, $\frac{\sum_{k=1}^{n} X_k}{n} \xrightarrow[n\to\infty]{P} m$. Теорема Чебышева доказана.

Обобщенная теорема Чебышева.

Пусть X_1, X_n – независимые случайные величины с математическими ожиданиями $m_1,\ldots m_n$ и дисперсиями $D_1\ldots, D_n$. Пусть дисперсии *ограничены в совокупности* ($D_k < L$, k-1,2…n). Тогда среднее арифметическое наблюденных значений случайных величин сходится по вероятности к среднему арифметическому их математических ожиданий.

$$\frac{\sum_k X_k}{n} \xrightarrow{P} \frac{\sum_k m_k}{n}$$

Доказательство. Рассмотрим, $Y = \frac{\sum_{k=1}^{n} X_k}{n}$. $M(Y) = \frac{\sum_{k=1}^{n} m_k}{n}$,

Оценим $D(Y) = \frac{D\left(\sum_{k=1}^{n} X_k\right)}{n^2} = \frac{1}{n^2} M\left(\sum X_k - M(\sum X_k)\right)^2 = \frac{1}{n^2} M\left(\sum\left(X_k - M(X_k)\right)\right)^2 \leq$

$$\frac{1}{n^2} M\left(\sum\left(X_k - M(X_k)\right)^2\right) + \frac{2}{n^2}\sum_{p<k} (X_p - M(X_p))(X_k - M(X_k))) =$$

(случайные величины независимы, следовательно, и не коррелированны)

$\frac{1}{n^2}\sum M\left(X_k - M\left(X_k\right)\right)^2 = \frac{1}{n^2}\sum D_k \leq \frac{Ln}{n^2} = \frac{L}{n} \xrightarrow[n\to\infty]{} 0$. Отсюда по второму неравенству Чебышева следует утверждение теоремы (доказательство сходимости по вероятности проводится как в предыдущей теореме).

Теорема Маркова.

Пусть X_1, X_n – *зависимые* случайные величины с математическими ожиданиями $m_1,\ldots m_n$ и дисперсиями $D_1\ldots, D_n$. Пусть $\frac{D\left(\sum_{k=1}^{n} X_k\right)}{n^2} \xrightarrow[n\to\infty]{} 0$. Тогда среднее арифметическое наблюденных значений случайных величин сходится по вероятности к среднему арифметическому их математических ожиданий.

Доказательство. Доказательство сходимости по вероятности проводится как в теореме Чебышева.

Теорема Бернулли.

При неограниченном увеличении числа опытов – независимых испытаний частота события сходится по вероятности к вероятности события.

Доказательство проводится аналогично теореме Чебышева.

Предельные теоремы.

Центральная предельная теорема – это любая теорема, ставящая условия, при которых функция распределения суммы индивидуально малых случайных величин с ростом числа слагаемых сходится к нормальной функции распределения.

Центральная предельная теорема подтверждает следующее: если исход случайного эксперимента определяется большим числом случайных факторов, влияние каждого из которых пренебрежимо мало, то такой эксперимент хорошо аппроксимируется нормальным распределением с параметрами $m, \quad D.$, подобранными соответствующим образом.

Теорема Ляпунова.

Пусть X_k – независимые случайные величины, имеющие математические ожидания $M(X_k) = m_k$ и дисперсии $D(X_k) = D_k$. Обозначим $I = \dfrac{\sum_{k=1}^{n}(X_k - m_k)}{\sqrt{\sum_{k=1}^{n} D_k}}$. Если можно подобрать такое $\delta > 0$, что $\dfrac{\sum_{k=1}^{n} M\left|X_k - m_k\right|^{2+\delta}}{\left(\sqrt{\sum_{k=1}^{n} D_k}\right)^{2+\delta}} \xrightarrow[n\to\infty]{} 0$, то при $n \to \infty$ $F_I(x) \to \Phi(x) = \dfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$ равномерно по x.

Теорема Леви – Линдеберга.

Пусть X_k – независимые *одинаково распределенные* случайные величины, имеющие математические ожидания $M(X_k) = m$ и дисперсии $D(X_k) = \sigma^2$. Обозначим $I = \dfrac{\sum_{k=1}^{n}(X_k - m)}{\sigma\sqrt{n}}$.

Тогда при $n \to \infty$ $F_I(x) \to \Phi(x) = \dfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2} dt$ равномерно по x.

Замечание. В теореме Леви – Линдеберга (ее чаще всего и называют центральной предельной теоремой) $\sqrt{\sum_{k=1}^{n} D_k} = \sqrt{n\sigma^2} = \sigma\sqrt{n}$, условие $\dfrac{\sum_{k=1}^{n} M|X_k - m_k|^{2+\delta}}{\left(\sqrt{\sum_{k=1}^{n} D_k}\right)^{2+\delta}} \xrightarrow[n\to\infty]{} 0$

выполнено, оно превращается в $\dfrac{1}{n^{\delta-1}} \xrightarrow[n\to\infty]{} 0$ (проверьте сами) из-за требования «одинаковости распределений», т.е. равенства вкладов случайных величин в случайную величину I.

Если рассматривать схему Бернулли, то из теоремы Леви – Линдеберга следует интегральная теорема Муавра – Лапласа.

Интегральная теорема Муавра – Лапласа.

Пусть производится n независимых испытаний, в каждом из которых с вероятностью p может появиться событие А. Обозначим X_k - число появлений события в k –ом испытании ($M(X_k) = p, D(X_k) = pq$). Обозначим $X = \sum_{k=1}^{n} X_k$ - общее число появлений события в n испытаниях ($M(X) = np, D(X) = npq$). Обозначим $I = \dfrac{X - np}{\sqrt{npq}}$. Тогда при $n \to \infty$

$F_I(x) \to \Phi(x) = \dfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2} dt$ равномерно по x.

Отсюда следует практическое правило вычисления

$$P\left(a < \frac{\sum_{k=1}^{n} X_k - np}{\sqrt{npq}} < b \right) \approx \Phi(b) - \Phi(a) = \Phi_0(b) - \Phi_0(a), \text{ где } \forall a, b,$$

$\Phi_0(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{1}{2}t^2}dt$. Так как $\forall a, b,$ то заменим a на $\frac{m_1-np}{\sqrt{npq}}$, b на $\frac{m_2-np}{\sqrt{npq}}$, $\sum_{k=1}^{n}X_k=m.$ Выведем формулу для вероятности нахождения числа успехов в заданном интервале, приведенную в лекции 5:

$$P\left(m_1<\sum_{k=1}^{n}X_k<m_2\right)=P\left(\frac{m_1-np}{\sqrt{npq}}<\frac{m-np}{\sqrt{npq}}<\frac{m_2-np}{\sqrt{npq}}\right)\approx\Phi_0\left(\frac{m_2-np}{\sqrt{npq}}\right)-\Phi_0\left(\frac{m_1-np}{\sqrt{npq}}\right).$$

Заменим a на $a\sqrt{\frac{n}{pq}}$, b на $b\sqrt{\frac{n}{pq}}$ в силу произвольности a, b.

Тогда $P\left(a\sqrt{\frac{n}{pq}}<\frac{m-np}{\sqrt{npq}}<b\sqrt{\frac{n}{pq}}\right)\approx\Phi_0\left(b\sqrt{\frac{n}{pq}}\right)-\Phi_0\left(a\sqrt{\frac{n}{pq}}\right)$.

Но $P\left(a\sqrt{\frac{n}{pq}}<\frac{m-np}{\sqrt{npq}}<b\sqrt{\frac{n}{pq}}\right)=P\left(a<\frac{m-np}{\sqrt{npq}}\frac{\sqrt{pq}}{\sqrt{n}}<b\right)=P\left(a<\left(\frac{m}{n}-p\right)<b\right)$.

Поэтому справедлива формула для вычисления отклонения частоты от вероятности

$$P\left(a<\left(\frac{m}{n}-p\right)<b\right)\approx\Phi_0\left(b\sqrt{\frac{n}{pq}}\right)-\Phi_0\left(a\sqrt{\frac{n}{pq}}\right),$$

Если интервал симметричный, т.е. $a=-\varepsilon, b=\varepsilon$, то по нечетности функции Лапласа получим

$$P\left(\left|\frac{m}{n}-p\right|<\varepsilon\right)\approx2\Phi_0\left(\varepsilon\sqrt{\frac{n}{pq}}\right).$$

Пример. Вероятность появления события $p = 0{,}8$. Сделано $n = 100$ независимых испытаний. Найти вероятность того, что событие произойдет не менее 75 и не более 90 раз.

$$P(75<X<90)=\Phi_0\left(\frac{90-100\;0{,}8}{\sqrt{100\;0{,}8\;0{,}2}}\right)-\Phi_0\left(\frac{75-100\;0{,}8}{\sqrt{100\;0{,}8\;0{,}2}}\right)=\Phi_0(2{,}5)-\Phi_0(-1{,}25)$$
$$=0{,}4938+0{,}3944=0{,}9882$$

Пример. Бюффон бросил монету 4040 раз и получил герб 2048 раз. Найти вероятность отклонения частоты появления герба от вероятности.

$$\left|\frac{m}{n}-p\right|=\left|\frac{2048}{4040}-\frac{1}{2}\right|\approx0{,}007=\varepsilon\quad P\approx2\Phi_0\left(0{,}007\sqrt{\frac{4040}{0{,}25}}\right)=2\Phi_0(0{,}89)=0{,}626$$

Содержание.

Литература.

1. Теория вероятностей (Математика в техническом университете вып. XVI) / Под ред. В.С. Зарубина и А.П. Крищенко. М.: МГТУ им. Н.Э. Баумана, 2001, 455с.
2. Вентцель Е.С. Теория вероятностей: Учеб. М.: Наука, 1969, 564с.
3. Гмурман В.Е. Теория вероятностей и математическая статистика. М: Высш. Шк., 1972, 477с.
4. Вентцель Е.С., Овчаров Л.А. Теория вероятностей: Задачник. М.: Наука, 1973. 172с.
5. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике. М : Высш. шк., 1975. 334с.