Лекция №1

§ 1. Понятие случайного события

Определение. Событие А называется случайным, если при осуществлении определенной совокупности условий оно может либо произойти, либо не произойти. Осуществление определенной совокупности условий называется испытанием или экспериментом.

Примеры: подбрасывание монеты, игральной кости.

Определение. События называются равновозможными, если есть основания считать, что ни одно из них не является более предпочтительным, чем другое.

Примеры: подбрасывание фальшивой монеты со смещенным центром тяжести.

Определение. Случайные события называются несовместными (или взаимоисключающими), если в одном и том же испытании появление одного из них исключает появление других.

Пример: выпадение пятёрки и тройки на верхней грани игральной кости - два несовместных события. Выпадение числа больше трёх {4,5,6} и выпадение четного числа {2,4,6} - события совместные.

Определение. Вся совокупность несовместных исходов экспериментов называется пространством элементарных событий Ω. Исходы ω_i, входящие в эту совокупность, называются элементарными событиями.

$\Omega = \{ \omega_1, \omega_2, \ldots \omega_n\}$

Пример: подбрасывание игральной кости Ω ={1,2,3,4,5,6}

§2. Математическая модель испытания

Пространство элементарных событий может быть построено не единственным способом, выбор математической модели зависит от условий поставленной задачи.

Пример 1: подбрасывание игральной кости

Ω ={1,2,3,4,5,6}, Ω = {чётное число, нечётное число}.

Пример 2: однократное подбрасывание монеты

Ω={герб, цифра}, Ω={герб, цифра, ребро}

Пример 3: двукратное подбрасывание монеты.

Ω_1={ГГ,ЦЦ,ГЦ,ЦГ}

Ω_2={Г-0 раз, Г-1 раз, Г-2 раза}

Можно провести аналогию между понятиями теории вероятностей и теории множеств.

ω_i ↔ элемент множества

Ω ↔ пространство

А ↔ подмножество

Определение. Случайным событием называется любое подмножество пространства элементарных исходов.

Пример 4: подбрасывание игральной кости, $\Omega=\{1,2,3,4,5,6\}$

A={четное число}={2,3,4}

B= {число очков>4} ={5,6}

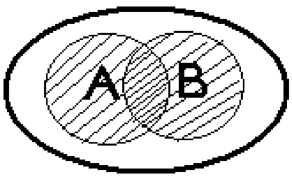
Определение. Событие называется невозможным, если оно не может произойти в данном эксперименте. Событие называется достоверным, если в данном эксперименте оно происходит всегда.

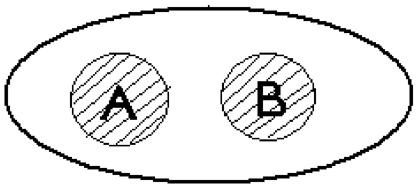
§3. Операции над событиями. Алгебра случайных событий

1. Событие, заключающееся в том, что произошло хотя бы одно из случайных событий A или B, называется суммой событий.

$$C = A+B = A \cup B$$

$A = \{2, 4, 6\}$, $B = \{5, 6\}$, $C = \{2, 4, 5,6\}$



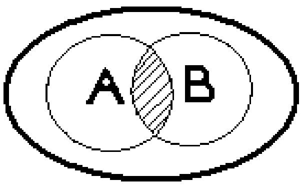


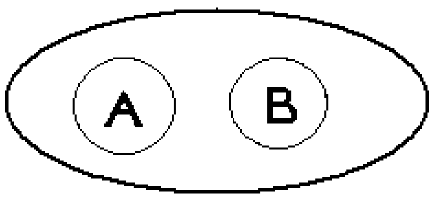
← если $AB = \varnothing$

2. Событие, заключающееся в том, что произошло и событие A, и событие B одновременно, называется произведением событий A и B.

$$C = AB = A \cap B$$

$A = \{2, 4, 6\}$, $B = \{5, 6\}$, $C = \{6\}$



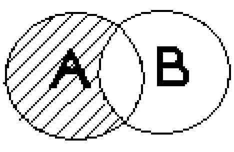


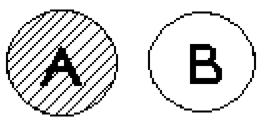
← если $AB = \varnothing$

3. Событие, заключающееся в том, что A произошло, а B - нет, называется разностью событий A и B.

$$C = A \setminus B = A - B$$

$A = \{2, 4, 6\}$, $B = \{5, 6\}$, $C = \{2,4\}$



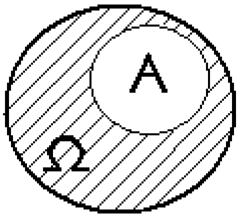


← если $AB = \varnothing$

4. Событие противоположно событию A, если оно содержит все исходы, не принадлежащие A.

$$\bar{A} = \Omega \setminus A$$

$$A = \{2, 4, 6\}, \quad \bar{A} = \{1, 3, 5\}$$



5. Тождественные события (A=B) – события, которые содержат одни и те же элементарные исходы.

Говорят, что A влечет за собой B , если при наступлении A произойдет B.

$B = \{5,6\}$, $A = \{5\}$

Свойства введенных операций.

1. Коммутативность: $A + B = B + A$; $AB = BA$.
2. Ассоциативность : $(A + B) + C = A + (B+C)$; $(AB)C = A(BC)$.
3. Дистрибутивность умножения относительно сложения: $(A + B)C = AC + BC$.

Некоторые полезные соотношения.

а) $A + A = A$ $\qquad A + \Omega = \Omega$

$AA = A$ $\qquad A\Omega = A$

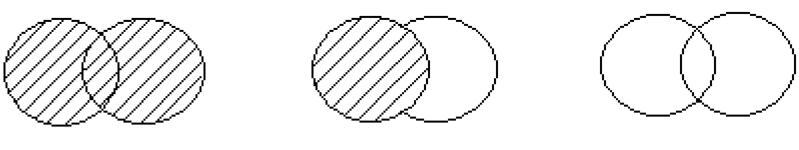
$A\bar{A} = \varnothing$ $\qquad A + \bar{A} = \Omega$

б) правила де Моргана

$$\overline{A + B} = \bar{A} \cap \bar{B}$$

$$\overline{AB} = \bar{A} + \bar{B}$$

в) $A + B = A + (B - AB) = AB + \bar{A}B + \bar{B}A = A + B\bar{A}$



г) $A - B = A\bar{B}$

Определение. Пусть Ω – пространство элементарных событий, U – множество всех подмножеств Ω, включая невозможные события и достоверное событие. Множество U называют алгеброй событий, если оно замкнуто относительно операции сложения и умножения.

Замечание: для испытаний с конечным числом исходов n U всегда является алгеброй и содержит 2^n элементов.

§4. Классическое определение вероятности

Определение. Пусть $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ – пространство элементарных равновозможных событий, u- алгебра событий ($N(U) = 2^n$) .

Тогда:

а) $P(\omega_i) = 1/n$, $i = 1, 2, \ldots, n$

б) $P(A) = m/n$, где $A = \{\omega_{i1}, \omega_{i2}, \ldots, \omega_{im}\}$

Вероятность события А – есть отношение числа исходов, благоприятствующих А, к общему числу исходов n.

Пример 1: подбрасывание игральной кости

$\Omega = \{1, 2, 3, 4, 5, 6\}$, $n = 6$

$A = \{2, 4, 6\}$, $m = 3$, $P(A) = m/n = 3/6 = ½$

$B = \{5, 6\}$, $m = 2$, $P(B) = 2/6 = 1/3$

Пример 2: вытаскивание карты из колоды.

$\Omega = \{1, \ldots, 36\}$, $n = 36$

A = {дама}, $m = 4$, $P(A) = 4/36 = 1/9$

Пример 3: двукратное подбрасывание монеты.

Ω_1 = {гг, цц, гц, цг}

Ω_2 = {г – 0 раз, г - 1 раз, г – 2 раза}

A = {герб выпал один раз}

Для вычисления $P(A)$ выбираем Ω_1 , т.к. Ω_2 содержит неравновозможные исходы.

$n = 4$, $m = 2$, $P(A) = 2/4 = 1/2$

Лекция №2.

§5. Геометрическая вероятность. Статистическая вероятность. Субъективная вероятность.

Геометрическая вероятность.

Чтобы преодолеть недостаток классического определения вероятности, состоящий в том, что оно неприменимо к испытаниям с бесконечным числом исходов, вводят понятие *геометрической вероятности.*

Определение. Геометрической вероятностью P(A) наступления некоторого события A в испытании называют отношение G_A/G_Ω, где G_Ω – геометрическая мера, выражающая общее число всех равновозможных исходов данного испытания, а G_A – мера, выражающая количество благоприятствующих событию исходов. На практике в качестве такой геометрической меры чаще всего выступает длина или площадь, реже – объём.

Пример 1. Метровую ленту случайным образом разрезают ножницами. Найти вероятность того, что длина обрезка составит не менее 80 см.

Пример 2. Задача о встрече. Двое друзей условились встретиться в определенном месте между 12 и 13 часами. Первый ждет другого 20 мин. и уходит. Чему равна вероятность того, что встреча состоится?

Относительная частота события и статистическая вероятность.

Определение. Относительной частотой W(A) события A называют отношение числа испытаний m, в которых данное событие появилось, к общему числу фактически проведённых испытаний n

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

Относительная частота рассчитывается исключительно ПОСЛЕ опытов на основе фактически полученных данных.

В том случае, если серии испытаний проводятся в неизменных условиях, то относительная частота обнаруживает свойство устойчивости, то есть колеблется около определённого значения.

Пример.

Количество бросков монеты, n	Число появлений орла, m	Относительная частота, $\omega = \frac{m}{n}$
4040	2048	0,5069
12000	6019	0,5016
24000	12012	0,5005

Если в одинаковых (примерно одинаковых) условиях проведено достаточно много испытаний, то за статистическую вероятность события принимают относительную частоту данного события либо близкое число.

Субъективная вероятность.

Субъективная вероятность - численное значение вероятности, полученное на основе заключения эксперта. Субъективная вероятность ассоциируется с определенным типом поведения человека в процессе принятия решений и используется не для вычислений частот, а для предсказания поведения во время принятия решений.

§6. Аксиоматическое определение вероятности события

Определение. Пусть Ω - пространство элементарных событий, U-алгебра событий, A - событие, принадлежащее алгебре событий. Вероятностью P(A) события A называется числовая функция, определенная для любого A из U и удовлетворяющая следующим условиям (аксиомам):

1) P(A) всегда неотрицательна ($P(A)\geq 0$)

2)$P(\Omega) =1$

3)$P(\Sigma A_к) = \Sigma P(A_к)$ - для несовместных событий $A_1, A_2,\ldots A_n$ ($A_i A_j= \varnothing$).

Свойства вероятности:

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

Доказательство:

$A+\bar{A}=\Omega$; $P(A+\bar{A})=P(\Omega)$; $A\bar{A}= \varnothing$; $P(A)+P(\bar{A})=1 \rightarrow P(\bar{A})=1-P(A)$

2.Формула сложения вероятностей (вероятность суммы двух совместных событий)

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

Доказательство:

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

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

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

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

3.Монотонность вероятности

Пусть событие A влечет за собой наступление B ($A\in B$) тогда $P(B)\geq P(A)$

Доказательство:

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

$P(B)\geq P(A)$

Тройка (Ω,U,P) образует *вероятностное пространство*.

§7. Условные вероятности

Определение. Условной вероятностью $P(B/A)=P_A(B)$ называют вероятность события B, вычисленную в предположении, что событие A уже произошло.

Пример. В урне 3 белых и 3 черных шара, из урны дважды вынимают по одному шару, не возвращая их обратно. Найти вероятность извлечь черный шар при втором испытании, если при первом испытании был извлечен белый шар.

Решение. Пусть A={при первом испытании извлечен белый шар}, B ={при втором испытании извлечен черный шар}

Способ 1: Очевидно, что после извлечения белого шара в урне осталось 5 шаров, из которых 3 черных. Тогда $P(B/A)=3/5$

Способ 2: $P(AB)=N_{AB}/N=3*3/A_6^2=9/30$; $P(A)=1/2$

$P(AB)= N_{AB}/N = N_{AB}/N_A *N_A/N=P(B/A)*P(A)$

$P(B/A)=P(AB)/P(A)=9/30:1/2=3/5$

$P(B/A)=P(AB)/P(A)$ – формула для вычисления условной вероятности.

§8. Вероятность произведения событий.

Вероятность произведения событий A и B вычисляется по формуле

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

которую называют *формулой умножения вероятностей*.

Если число рассматриваемых событий больше двух, то вероятность произведения событий следует вычислять, последовательно применяя формулу умножения вероятностей. Например, для трех событий

$P(ABC)=P(AB)*P(C/AB)=P(A)*P(B/A)*P(C/AB)$

Определение. События A и B называются *независимыми*, если $P(AB)=P(A)*P(B)$. Для независимых событий $P(B/A)=P(B)$

Замечание. Независимые события всегда совместны. Обратное утверждение, вообще говоря,

Определение. События $A_1, A_2,\dots A_n$ называются *независимыми в совокупности*, если для любого $1\le k\le n$ и любого набора различных меж собой индексов $i_1,i_2\dots i_k$ имеет место равенство:$P(A_{i_1}\cap A_{i_2}\dots\cap A_{i_k})=P(A_{i_1})P(A_{i_2})\dots P(A_{i_k})$

Замечание. Если события независимы в совокупности, то они попарно независимы, т.е. любые два события независимы. Обратное утверждение, вообще говоря, неверно: из попарной независимости не вытекает независимость в совокупности.

Пример 1. Производится бросание двух игральных костей.

A = {на первой кости выпало четное число очков}

B= {на второй кости выпало нечетное число очков }

C = {сумма очков четна}

P(ABC)=0; P (A)=P(B)=P(C)=1/2; P (AC)=P(AB)=P(CB)=1/4

События попарно независимы, но не являются независимыми в совокупности.

Пример 2. Из 100 студентов, находящихся в аудитории, 50 человек знают английский язык, 40 французский язык и 35 немецкий. Английский и французский языки знают 20 студентов, английский и немецкий - 8 студентов, французский и немецкий - 10 студентов. Все три языка знают 5 человек. Один из студентов вышел из аудитории.

E = {вышедший знает английский язык}

F = {вышедший знает французский язык}

D = {вышедший знает немецкий язык}

Указать все пары независимых событий? Указать, являются ли события E, F, D независимыми в совокупности?

Решение.

а) P(EF) =P(E)*P(F/E)= (50/100)*(20/50) =0,2

P(E)=50/100 =0,5

P(F)=40/100 =0,4

P(E)*P(F)=0,5*0,4=0,2

P(EF)=P(E)*P(F)=0,2; следовательно E и F – независимы.

б) P(E)=0,5

P(D)=35/100 =0,35

P(E)*P(D)=0,175

P(ED)=P(E)*P(D/E)=0,5*8/50=0,08

P(ED) $\neq$ P(E)*P(D); следовательно E и D являются зависимыми.

В) P(F)=0,4

P(D)=0,35

P(F)*P(D)=0,14

P(FD)=P(F)*P(D/F)=0,4*0,25=0,1

P(FD) $\neq$ P(F)*P(D) ; следовательно F и D являются зависимыми.

Так как среди рассматриваемых событий есть две пары зависимых, то E, F, D зависимы в совокупности.

§8. Формула полной вероятности

Теорема. Пусть выполняются следующие условия:

- события $H_1, H_2, \ldots H_n$ – попарно несовместные события, т.е. $H_i H_j = \varnothing$ для любых $i \neq j$;

- события обладают конечными вероятностями $P(H_i) > 0$;

- событие А наступает только вследствие наступления одного из событий H_i,

т.е. $A \in H_1 + H_2 + \ldots + H_n$.

Тогда вероятность события А вычисляется по формуле

$$P(A)= P(A/H_1)P(H_1)+ P(A/H_2)P(H_2)+\ldots+ P(A/H_n)P(H_n)$$

Доказательство.

Событие А может наступить, если наступит одно из несовместных событий $H_1, H_2, \ldots, H_n$.

$A= AH_1 + AH_2 + \ldots + AH_n$ - разложение А по несовместным событиям.

$P(A)= P(AH_1 + AH_2 + \ldots + AH_n)= P(AH_1) + P(AH_2) + \ldots + P(AH_n)= P(A/H_1)P(H_1)+ P(A/H_2)P(H_2)+\ldots+ P(A/H_n)P(H_n)$

Замечания.

1.Если события $H_1, H_2, \ldots H_n$ – попарно несовместные события и $H_1+H_2+\ldots+H_n = \Omega$, то множество событий $H_1, H_2, \ldots, H_n$ называется *полной группой событий*.

2.Часто события $H_1, H_2, \ldots, H_n$ называют *гипотезами*.

3.Формула полной вероятности работает в том случае, если множество гипотез H_i счетно .

Пример 1.

На фабрике, изготавливающей болты, 1-ая машина производит – 25%, 2-ая – 35%, 3-я – 40% всей продукции. Брак в продукции составляет 5%, 4% и 2% соответственно. Какова вероятность того, что случайно выбранный болт оказался дефектным?

Решение.

A={случайно выбранный болт оказался дефектным}

H_i={болт изготовлен на i-ой машине}.

Тогда

$P(H_1)=0{,}25$; $P(H_2)=0{,}35$; $P(H_3)=0{,}40$ $(P(H_1)+P(H_2)+P(H_3)=1)$

$P(A/H_1)=0{,}05$; $P(A/H_2)=0{,}04$; $P(A/H_3)=0{,}02$

По формуле полной вероятности

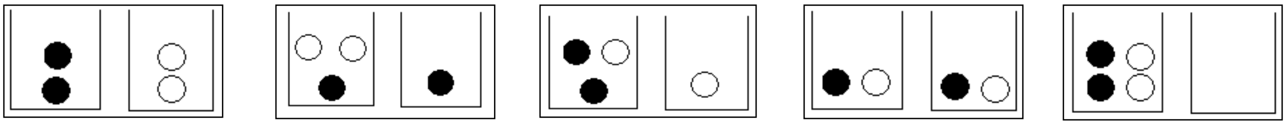
$P(A) = 0{,}25*0{,}05 + 0{,}35*0{,}04 + 0{,}4*0{,}02 = 0{,}0345$

Пример 2. Древняя легенда гласит. Один султан прогневался на звездочета и решил его казнить. Чтобы быть справедливым, он приказал звездочету разложить 2 шарика белого цвета и 2 шарика черного цвета по двум урнам. Звездочету нужно наугад выбрать урну и вытащить из нее один шарик: белый шарик – жизнь, черный – смерть. Каким образом звездочету нужно разложить шарики, чтобы вероятность сохранить жизнь была бы максимальной?

Решение.

Варианты расположения шариков:

1) 2) 3) 4) 5)



H_1={ выбрана 1 урна} $P(H_1) = 1/2$

H_2={ выбрана 2 урна} $P(H_2) = 1/2$

A={звездочет вытащил белый шарик}

1)$P(A)=1/2*1 + 1/2*0 = 1/2$

2)$P(A)=1/2*2/3 + 1/2*0 = 1/3$

3)$P(A)=1/2*1/3 + 1/2*1 = 2/3$

4)$P(A)=1/2*1/2 + 1/2*1/2 = 1/2$

5)$P(A)=1/2*1/2 + 0*1/2 = 1/4$

Наибольшая вероятность для звездочета вытащить белый шарик – при варианте №3.

§9. Формула Байеса

С формулой полной вероятности тесно связана формула Байеса. Если до опыта вероятности гипотез были $P(H_i)$, а в результате опыта появилось событие А, то с учетом этого события "новые", т.е. условные вероятности гипотез вычисляются по формуле Байеса.

Вероятность совместного наступления событий A и H_k может быть выражена через условные вероятности двумя способами

$$P(AH_k) = P(A)P(H_k/A) = P(H_2) P(A/H_k)$$

Следовательно

$$P(H_k/A) = P(H_k)P(A/H_k) /P(A)$$ - формула Байеса, где

$P(H_k)$ -априорная вероятность гипотезы H_k

$P(H_k/A)$ - вероятность гипотезы H_k при наступлении события A (апостериорная вероятность);

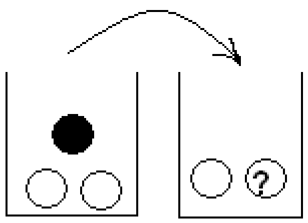
$P(A/H_k)$ -вероятность наступления события A при истинности гипотезы H_k;

$P(A)$ -полная вероятность наступления события A.

Формула Байеса позволяет «переставить причину и следствие»: по известному факту наступления события вычислить вероятность того, что оно было вызвано данной причиной.

Пример 1а.

В урне 3 шара: 2 белых и 1 черный. Из урны наудачу берут один шар и откладывают в сторону. После этого выбирается еще 1 шар. Какова вероятность во второй попытке выбрать белый шар?



Вероятность во второй попытке выбрать белый шар зависит от того, какие шары останутся в урне

H_1={отложен белый шар}

H_2={отложен черный шар}

A = {вытащили при второй попытке белый шар}

P(A) - ?

Вероятности гипотез : $P(H_1) = 2/3$, $P(H_2) = 1/3$

$P(A/H_2) = 1/2$

$P(A/H_2) = 1$

По формуле полной вероятности

$P(A) = P(H_1)* P(A/H_1) + P(H_2)* P(A/H_2) = 2/3*1/2 + 1/3*1 = 2/3$

Пример 1б.

Какова вероятность, что был отложен белый шар, если во 2 попытке был вынут белый?

$P(H_1/A)$ - ?

$P(AH_1) = P(A)P(H_1/A) = P(H_1) P(A/H_2)$

$P(H_1/A) = P(H_1) P(A/H_1) /P(A) = 2/3*1/2 : 2/3 = 1/2$

Пример 2:

3 стрелка произвели залп, после чего в мишени обнаружена 1 пробоина. Найти вероятность того, что мишень поразил 3 стрелок, если известно, что вероятности попадания в мишень первым, вторым и третьим стрелком соответственно равны: 0,6; 0,5; 0,4.

Решение.

Предупреждение: нельзя в качестве гипотез рассматривать события, связанные с попаданием отдельными стрелками в мишень. Эти события – совместны! Сумма их вероятностей (0,5+0,6+0,4=1,5) больше единицы.

В качестве гипотезы примем H_1={мишень поразил 3 стрелок}, именно об этом событии идет речь в вопросе задачи. Несовместное и противоположное с H_1 событие H_2={3 стрелок не попал} будет второй гипотезой ($P(H_1)+P(H_2)=1$). Результатом испытания является событие A = {в мишени обнаружена 1 пробоина}. Необходимо вычислить $P(H_1/A)$.

$P(H_1) = 0,4$

$P(H_2) = 1-0,4 = 0,6$

$P(A/H_1) = (1-0,6)*(1-0,5) = 0,4*0,5 = 0,2$ – мишень поражена третьим стрелком, первый и второй - не попали.

$P(A/H_2) = 0,6*0,5 + 0,4*0,5 = 0,5$– мишень поражена вторым или первый стрелком, третий - не попал.

$P(A) = 0,4*0,2 + 0,6*0,5 = 0,38$ – полная вероятность (вероятность того, что в мишени обнаружена 1 пробоина)

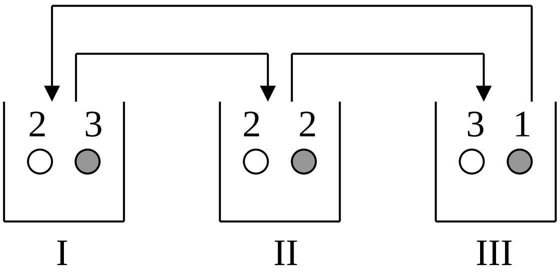
$P(H_1/A) = 0,4*0,2 / 0,38 = 0,211$

Лекция №3

§ 10. Последовательность испытаний

а) последовательность зависимых испытаний

Пример.



А$_к$= {при k–ом перекладывании был переложен белый шар} k =1,2,3

Ω={ $A_1A_2A_3$; $A_1A_2\overline{A}_3$ …}

С = {3 раза извлекают белый шар}={ $A_1A_2A_3$ }

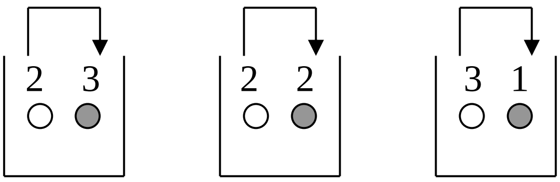
P(A_1)= $\frac{2}{5}$; P($\overline{A}_1$)= $\frac{3}{5}$

P(A_2)= $\frac{2}{5}\cdot\frac{3}{5}+\frac{3}{5}\cdot\frac{2}{5}=\frac{12}{25}$; P($\overline{A}_2$)= $\frac{13}{25}$

P(A_3)= $\frac{12}{25}\cdot\frac{4}{5}+\frac{13}{25}\cdot\frac{3}{5}=\frac{87}{125}$; P($\overline{A}_2$)= $\frac{38}{125}$

P(C) = P($A_1A_2A_3$)=P(A_1)P(A_2/A_1)P(A_3/A_1A_2)= $\frac{2}{5}\cdot\frac{3}{5}\cdot\frac{4}{5}=\frac{24}{125}$

б) последовательность независимых испытаний



А$_к$={при *k*–ом перекладывании был переложен белый шар}, *k* =1,2,3

Ω ={ $A_1A_2A_3$; $A_1A_2\overline{A}_3$ …}

С = {3 раза извлекают белый шар}={ $A_1A_2A_3$ }

P(C) = P($A_1A_2A_3$)=P(A_1)P(A_2)P(A_3)= $\frac{2}{5}\cdot\frac{2}{4}\cdot\frac{3}{4}=\frac{3}{20}$

в) цепь Маркова (случайный процесс с дискретными состояниями и дискретным временем)

Предположим, что исходом каждого испытания может быть одно событие из полной группы несовместимых событий $A_1A_2....A_m$, причем условная вероятность $p_{ij}^{(k)}$ появления какого-то события A_j при *(k+1)*-ом испытании зависит только от того, какое событие A_i появилось при

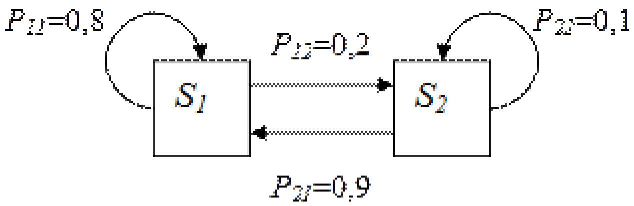
k-ом испытании, и не зависит от того, какие события появились при более ранних испытаниях. Такая последовательность событий называется *простой цепью Маркова*. Если условная вероятность перехода от события к событию обусловлена только этими событиями, но не зависит от номера испытания, то соответствующая простая цепь Маркова называется *однородной* и может быть описана с помощью матрицы переходных вероятностей.

Если для однородной Марковской цепи заданы начальное распределение вероятностей и матрица переходных вероятностей $\|p_{ij}\|$, то вероятности состояний системы $P_j(k)$ ($i, j = 1, 2, \dots, n$) определяются по рекуррентной формуле:

$$P_j(k) = \sum_{i=1}^{n} P_i(k-1)\,p_{ij} \qquad (*)$$

Пример: Рассмотрим процесс функционирования системы - автомобиль. Пусть автомобиль (система) в течение одной смены (суток) может находиться в одном из двух состояний: исправном (S1) и неисправном (S2).

Граф состояний автомобиля



В результате проведения массовых наблюдений за работой автомобиля составлена матрица вероятностей перехода

$$\|p_{ij}\| = \begin{Vmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{Vmatrix}$$

p_{11} = 0,8 – вероятность того, что автомобиль останется в исправном состоянии;

p_{12}=0,2 – вероятность перехода автомобиля из состояния «исправен» в состояние «неисправен»;

p_{21}=0,9 – вероятность перехода автомобиля из состояния «неисправен» в состояние «исправен»;

p_{22} =0,1 – вероятность того, что автомобиль останется в состоянии «неисправен».

Вектор начальных вероятностей состояний автомобиля задан, т.е. $P_1(0) = 0$ и $P_2(0)=1$.

Требуется определить вероятности состояний автомобиля через трое суток.

Используя матрицу переходных вероятностей и формулу (*), определим вероятности состояний $P_i(k)$ после первого шага (после первых суток):

$P_1(1) = P_1(0) \times p_{11} + P_2(0) \times p_{21} = 0 \cdot 0{,}8 + 1 \cdot 0{,}9 = 0{,}9$;

$P_2(1) = P_1(0) \times p_{12} + P_2(0) \times p_{22} = 0 \cdot 0{,}2 + 1 \cdot 0{,}1 = 0{,}2$

Вероятности состояний после второго шага (после вторых суток) таковы:

$P_1(2) = P_1(1) \times p_{11} + P_2(1) \times p_{21} = 0{,}9 \times 0{,}8 + 0{,}1 \times 0{,}9 = 0{,}81;$

$P_2(2) = P_1(1) \times p_{12} + P_2(1) \times p_{22} = 0{,}9 \times 0{,}2 + 0{,}1 \times 0{,}1 = 0{,}19$

Вероятности состояний после третьего шага (после третьих суток) равны

$P_1(3) = P_1(2) \times p_{11} + P_2(2) \times p_{21} = 0{,}81 \times 0{,}8 + 0{,}19 \times 0{,}9 = 0{,}819$

$P_2(3) = P_1(2) \times p_{12} + P_2(2) \times p_{22} = 0{,}81 \times 0{,}2 + 0{,}19 \times 0{,}1 = 0{,}181$

Таким образом, после третьих суток автомобиль будет находиться в исправном состоянии с вероятностью 0,819 и в состоянии «неисправен» с вероятностью 0,181

г) *поток событий (случайный процесс с дискретными состояниями и непрерывным временем)*

Часто встречаются ситуации, когда переход системы из состояния в состояние происходит не в фиксированные, а в случайные моменты времени, которые указать заранее невозможно. Например, любая деталь или агрегат могут выходить из строя в любой, непредсказуемый заранее момент времени. Когда событий много и они следуют друг за другом, то они образуют *поток*.

Поток событий — это последовательность *однородных* событий, наступающих одно за другим в случайные промежутки времени.

Интенсивность потока λ — это среднее число событий в единицу времени. Интенсивность потока можно рассчитать экспериментально по формуле: $\lambda = N/T_н$, где N — число событий, произошедших за время наблюдения $T_н$.

Если интервал между событиями равен константе или определен какой-либо формулой в виде: $t_j = f(t_{j-1})$, то поток называется детерминированным. Иначе поток называется случайным.

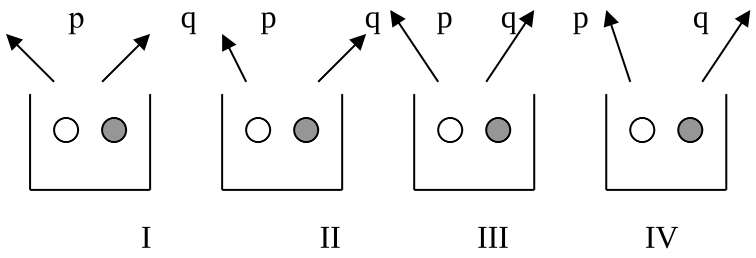
Случайные потоки бывают:

- *ординарные*: вероятность одновременного появления двух и более событий равна нулю;
- *стационарные*: частота появления событий $\lambda(t) = const$;
- *без последействия*: вероятность появления случайного события не зависит от момента совершения предыдущих событий.

Поток, который обладает свойствами стационарности, отсутствия последействия и ординарности, называют *пуассоновским (или простейшим)*. Название “пуассоновский” связано с тем, что для этого потока число событий, попадающих на любой фиксированный интервал времени, будет распределено по закону Пуассона.

§ 11. Схема Бернулли. Формула Бернулли

Пример 1. В урне расположены шары 2-х цветов – белых и черных. Вероятность извлечь белый шар равна *p*, а черный – *q=1-p*. Проведено 4 последовательных испытания по схеме с возвращением. Найти вероятность того, что белый шар появился 2 раза.



$A_к$ = {при *k*–ом испытаниях - белый шар }

B ={белый шар появился 2 раза}

B = { $A_1A_2\overline{A}_3\overline{A}_4, A_1\overline{A}_2A_3\overline{A}_4, A_1\overline{A}_2\overline{A}_3A_4, \overline{A}_1A_2A_3\overline{A}_4, \overline{A}_1A_2\overline{A}_3A_4, \overline{A}_1\overline{A}_2A_3A_4$ }

P($A_1A_2\overline{A}_3\overline{A}_4$)= $p \cdot p \cdot q \cdot q = p^2 \cdot q^2 = P(A_1\overline{A}_2A_3\overline{A}_4) = ...$

P(B) = P($A_1A_2\overline{A}_3\overline{A}_4$)+P($A_1\overline{A}_2A_3\overline{A}_4) + ... = N(B) \cdot p^2q^2$

N(B) = C_4^2

P(B)= $C_4^2 \cdot p^2 \cdot q^2$

Схема Бернулли:

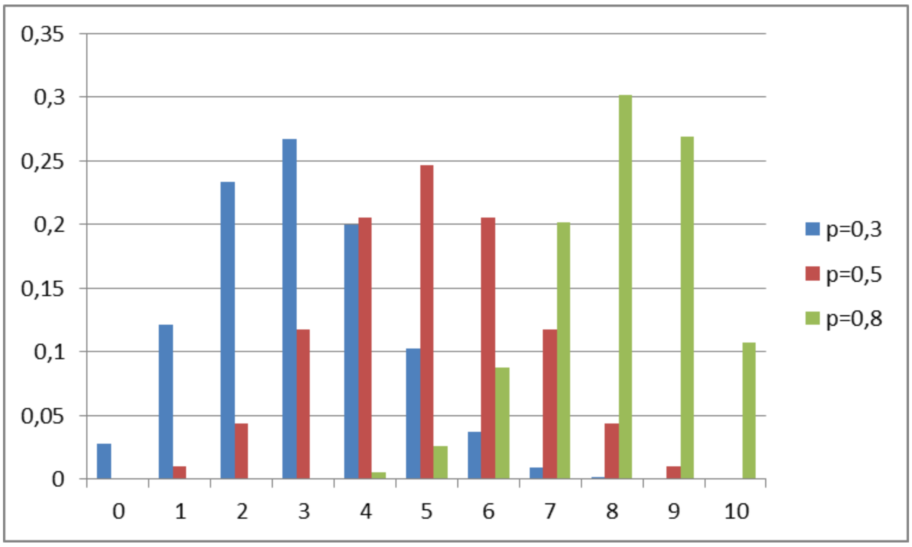
Пусть проведено n независимых испытаний, каждое из которых заканчивается либо «успехом» с вероятностью *p*, либо «неудачей» с вероятностью *q=1-p*. Тогда вероятность появления *m* «успехов» вычисляется по формуле Бернулли

$$P_n(m) = C_n^m \cdot p^m \cdot q^{n-m}$$

Вероятность появления $m_1 \le m \le m_2$ «успехов» считается по формуле Бернулли для отрезка

$$P_n(m_1 \le m \le m_2) = \sum_{m=m_1}^{m_2} C_n^m \cdot p^m \cdot q^{n-m}$$

Зависимость вероятности $P_n(m)$ от числа «успехов» m; n=10.



Наиболее вероятное число появления «успехов» $m^* = np$.

Пример 2. Монету бросают 4 раза. Найти вероятность того, что орел выпадет 2 раза.

Решение.

n=4, m=2. Вероятность выпадения орла $p=\frac{1}{2}$, вероятность выпадения решки $q=\frac{1}{2}$

$$P_4(2)=C_4^2(\frac{1}{2})^2\cdot(\frac{1}{2})^2=\frac{3}{8}$$

Пример 3. Монета бросается 5 раз. Найти вероятность того, что орел выпадет менее 2 раз.

Решение.

n=5, $0\le m\le 1$, p=$\frac{1}{2}$, q=$\frac{1}{2}$

$$P_5(0\le m\le 1)=P_5(0)+P_5(1)=C_5^0\cdot(\frac{1}{2})^0\cdot(\frac{1}{2})^5+C_5^1\cdot(\frac{1}{2})^1\cdot(\frac{1}{2})^4=\frac{3}{16}$$

Пример 4. Переход от схемы последовательных независимых испытаний с двумя исходами *(биномиальной схемы)* к *полиномиальной схеме.*

§12. Предельная теорема Пуассона

Теорема. Пусть число испытаний достаточно велико ($n\to\infty$), а вероятность появления интересующего нас события в каждом испытании достаточно мала ($p\to 0$), тогда вероятность

$$P_n(m)=\frac{\lambda^m}{m!}e^{-\lambda}\text{ , где }\lambda=np\ \in(0;\infty),\ m=0,1,2\ldots n$$

Замечание 1: на практике формулу Пуассона используют обычно когда $n>100$, а $npq\le 9$.

Замечание 2: поскольку $p\to 0$, то теорему называют законом редких явлений.

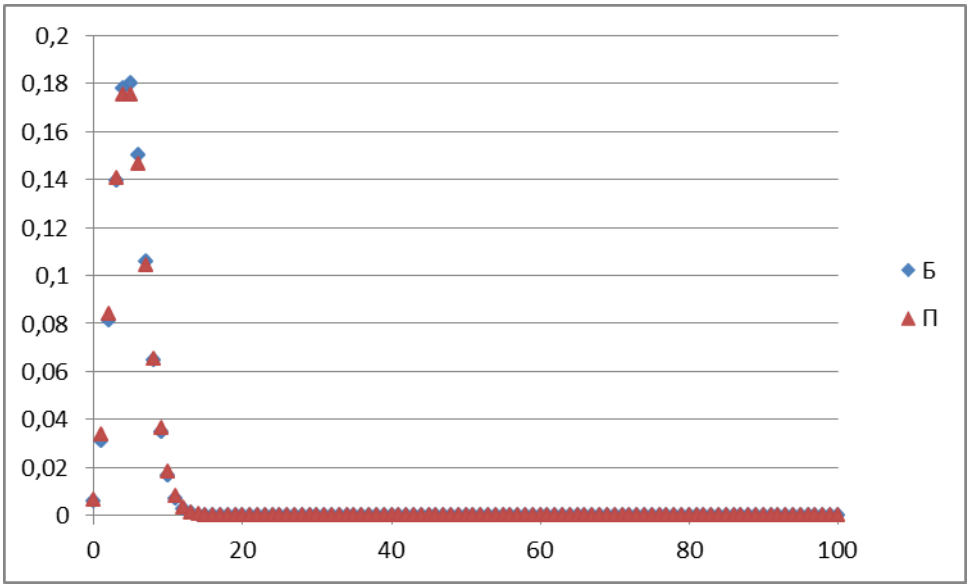
Доказательство.

$$P_n(m)= C_n^m \cdot p^m \cdot q^{n-m} = \frac{n!}{m!(n-m)!} p^m \cdot (1-p)^m = \frac{n!}{m!(n-m)!} \cdot (\frac{\lambda}{n})^m \cdot (1-\frac{\lambda}{n})^{n-m} = =$$

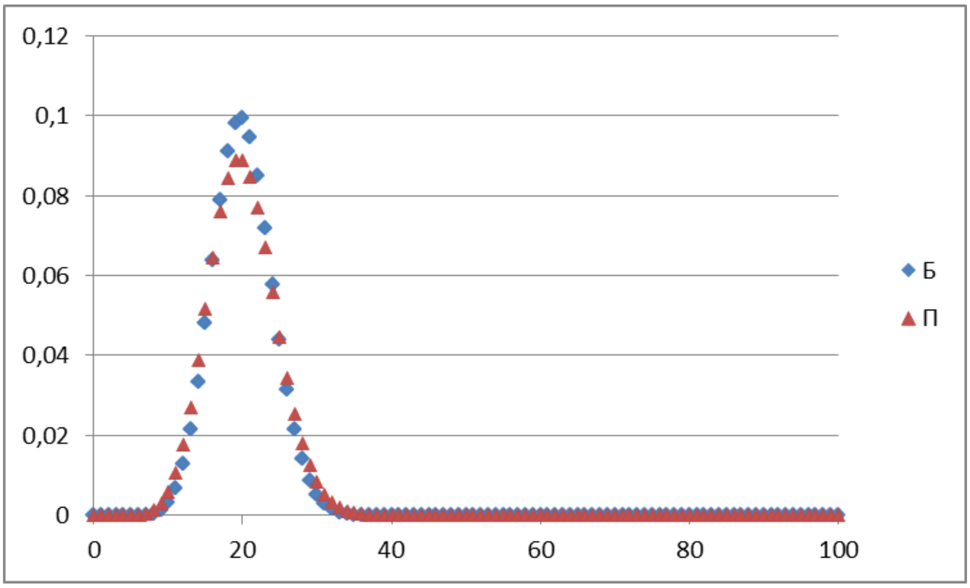
$$\frac{n(n-1)(n-2)...(n-m+1)}{m!} \cdot \frac{\lambda^m}{n^m} (1-\frac{\lambda}{n})^{n-m} \to \frac{\lambda^m}{m!} \cdot e^{-\lambda} \quad \text{при } n\to\infty$$

Сравнение вероятностей $P_n(m)$, вычисленных по формулам Бернулли (Б) и Пуассона (П).

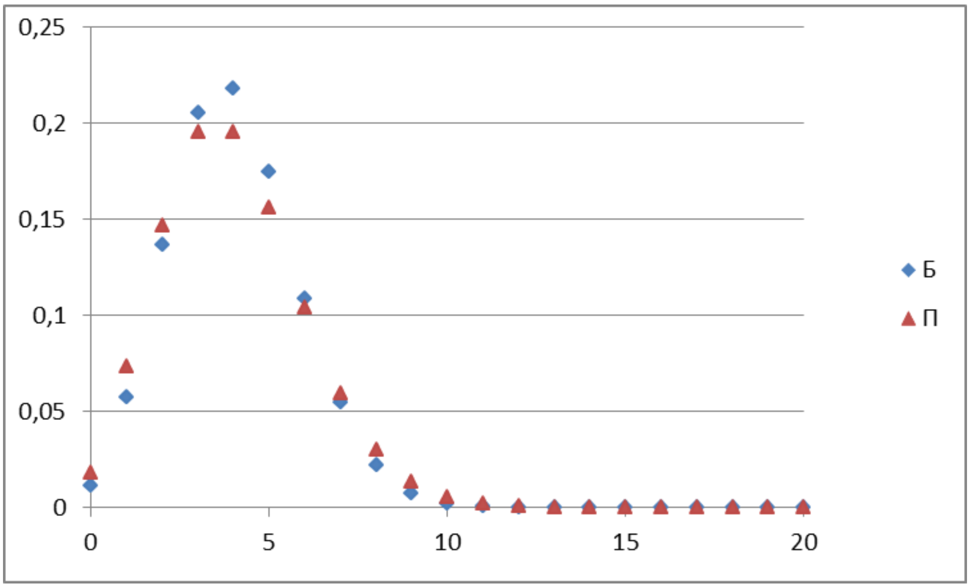
n=100, p=0,05; np=5, npq=4,75



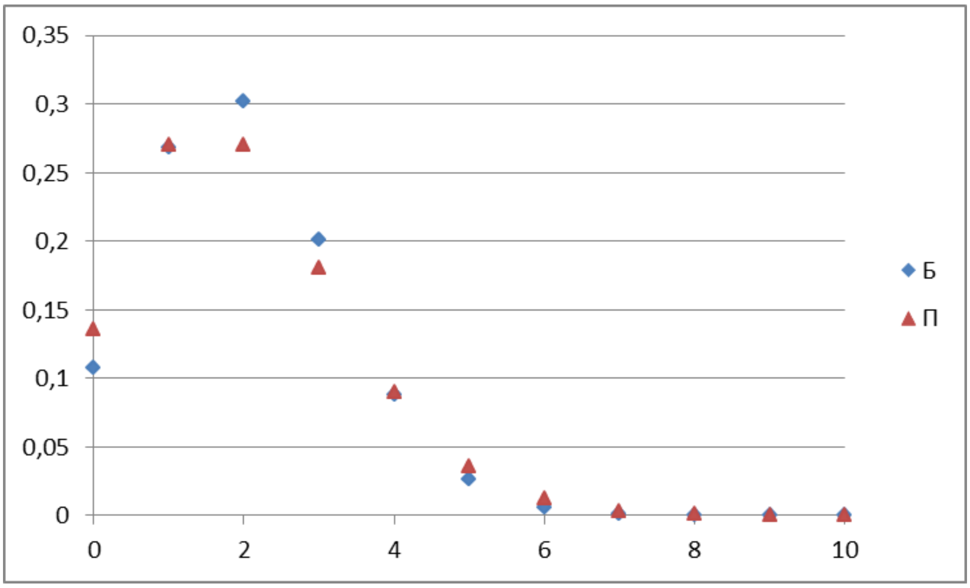
n=100, p=0,2; np=20, npq=16



n=20, p=0,2; np=4, npq=3,2



n=10, p=0,2; np=2, npq=1,6



Пример 1. На факультете насчитывается 500 студентов. Какова вероятность того, что 1 сентября является днем рождения одновременно для 3-х студентов.

n=500; p=1/365; q=364/365; m=3;

$$P_{500}(3) = C_{500}^{3}\cdot\left(\frac{1}{365}\right)^{3}\cdot\left(\frac{364}{365}\right)^{497} = \{n=500>100;\ npq=500\cdot 1/365\cdot 364/365<9;\ \ \lambda=np=500/365\approx 1{,}37\}\approx$$

$$\approx \frac{1{,}37^{3}}{3!}\cdot e^{-1{,}37}\approx 0{,}109$$

Замечание. Для *простейшего потока* вероятность появления m событий за время τ равна:

$$P(m) = \frac{(\lambda\tau)^{m}}{m!}\cdot e^{-\lambda\tau}$$

Вероятность того, что за время τ не появится ни одного события (m = 0) равна $P(0) = e^{-\lambda\tau}$.

Вероятность появления хотя бы одного события $P(m>0) = 1 - e^{-\lambda\tau}$

Пример 2. Среднее число заказов такси, поступающих на диспетчерский пункт в одну минуту, равно трем. Найти вероятность того, что за 2 мин поступит: а) четыре вызова; б) не менее четырех вызовов.

Решение.

$\lambda = 3;\ \tau = 2$

а) m = 4; $P(4) = \frac{(6)^4}{4!} \cdot e^{-6} \approx 0{,}135$

б) m≥4;   P(m≥4) =1- P(3)-P(2)-P(1)-P(0)≈0,8475

§13. Локальная теорема Муавра-Лапласа

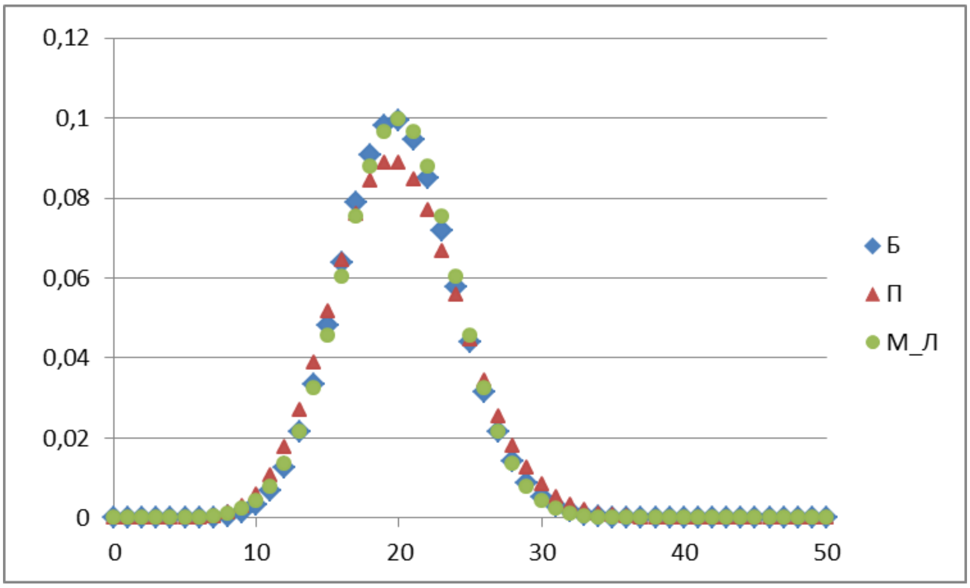
Теорема. Пусть число испытаний достаточно велико (n→∞), а вероятность появления интересующего нас события в каждом испытании постоянна и отлична от 0 и 1, тогда

$$P_n(m) = \frac{1}{\sqrt{npq}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \text{ где } x = \frac{m - np}{\sqrt{npq}}$$

Замечание: на практике формула Муавра-Лапласа когда *npq>9*

Сравнение вероятностей $P_n(m)$, вычисленных по формулам Бернулли (Б), Пуассона (П) и Муавра-Лапласа (М_Л).

n=100, p=0,2; np=20, npq=16



Пример.

Найти вероятность того, что событие *A* наступит 80 раз в 400 испытаниях, если вероятность появления этого события в каждом испытании равна 0,2.

n=400, m=80, p=0,2; (400>100, npq=64>10)

$$P_{400}(80)=C_{400}^{80}(0{,}2)^{80}(0{,}8)^{320}=\frac{1}{\sqrt{64}}e^{-\frac{(80-80)^2}{2\cdot 64}}\frac{1}{\sqrt{2\pi}}\approx 0{,}04986$$

§14. Интегральная теорема Муавра-Лапласа.

Теорема. Пусть число испытаний достаточно велико ($n\to\infty$), а вероятность появления, интересующего нас события, в каждом отдельном испытании постоянна и отлична от 0 и 1, тогда

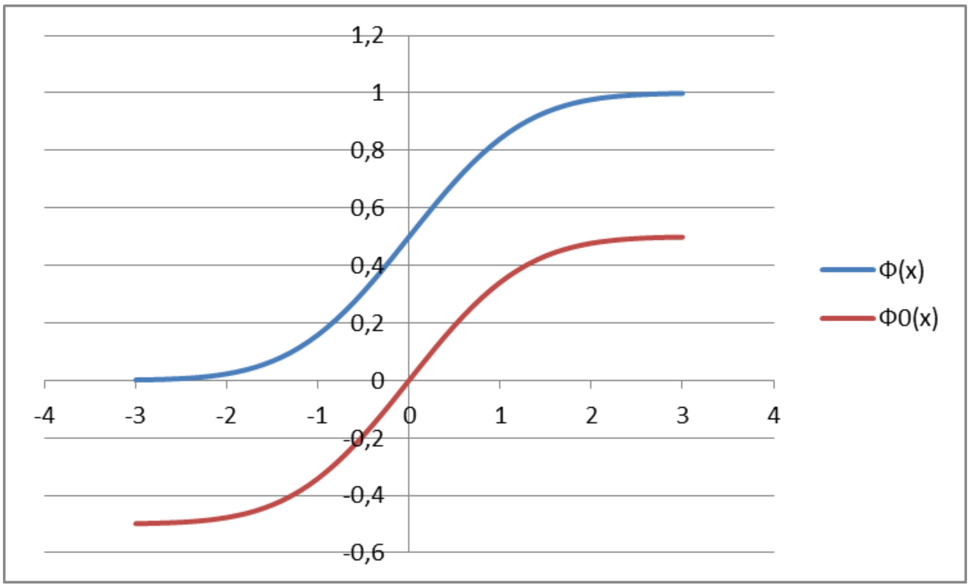
$$P_n(m_1\le m\le m_2)=\frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}dx$$

$$a=\frac{m_1-np}{\sqrt{npq}}\quad b=\frac{m_2-np}{\sqrt{npq}}$$

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

$\Phi_0(x)=\frac{1}{\sqrt{2\pi}}\int_{0}^{x} e^{-\frac{t^2}{2}}dt$ - нормированная функция Лапласа

$$P(m_1\le m\le m_2)=\frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}dx=\frac{1}{\sqrt{2\pi}}\left(\int_0^b e^{-\frac{x^2}{2}}dx-\int_0^a e^{-\frac{x^2}{2}}dx\right)=\Phi_0(b)-\Phi_0(a)=\Phi(b)-\Phi(a)$$



Пример 1. Найти вероятность того, что при 100 подбрасываниях монеты, число выпадений герба будет в пределах от 40 до 60.

n=100, p=0,5, 40≤m≤60, npq=25>10, a=-2, b=2

$$P_{100}(40\le m\le 60)=\Phi_0(2)-\Phi_0(-2)=0{,}9544$$

Пример 2. Вероятность отклонения относительной частоты появления события в серии независимых испытаний от постоянной вероятности появления этого события в отдельном испытании.

$P\left(\left|\frac{m}{n}-p\right|\le\varepsilon\right)$-?

Решение.

1) преобразуем неравенство

$$\left|\frac{m}{n}-p\right|\le\varepsilon\Leftrightarrow-\varepsilon\le\frac{m}{n}-p\le\varepsilon$$

$$P(-\varepsilon\sqrt{\frac{n}{pq}}\le\frac{m-np}{\sqrt{npq}}\le\varepsilon\sqrt{\frac{n}{pq}})-?$$

2) используем интегральную формулу Муавра-Лапласа

$$P(m_1\le m\le m_2)=P\left(a\le\frac{m-np}{\sqrt{npq}}\le b\right)$$

$$\frac{m_1-np}{\sqrt{npq}}\le\frac{m-np}{\sqrt{npq}}\le\frac{m_2-np}{\sqrt{npq}}$$

$$P\left(-\varepsilon\sqrt{\frac{n}{pq}}\le\frac{m-np}{\sqrt{npq}}\le\varepsilon\sqrt{\frac{n}{pq}}\right)=\frac{1}{\sqrt{2\pi}}\int_{-\varepsilon\sqrt{\frac{n}{pq}}}^{\varepsilon\sqrt{\frac{n}{pq}}}e^{-\frac{x^2}{2}}dx=2\Phi_0\left(\varepsilon\sqrt{\frac{n}{pq}}\right)\to1 \quad \text{при } n\to\infty$$

Примерный вариант контрольной работы №1.

1. В партии, состоящей из 1000 изделий, четыре изделия имеют дефекты. Для контроля отбираются 100 изделий. Найти вероятность того, что среди отобранных изделий не окажется бракованных. (0,656)
2. Жюри состоит из трёх судей, выносящих решение независимо друг от друга: двое из них, каждый с вероятностью 0,8, принимают правильное решение, а третий для вынесения решения подбрасывает монету. Окончательное решение принимается большинством голосов. Найти вероятность вынесения правильного решения. (0,8)
3. Статистика запросов кредитов в банке такова: 10% — государственные органы, 20% — другие банки, остальные — физические лица. Вероятности того, что взятый кредит не будет возвращён, составляют 0,01, 0,05 и 0,2 соответственно. Определить, какая доля кредитов в среднем не возвращается.(0,151)
4. Объемы товара, которые магазин получает товар от трёх поставщиков, относятся как 11:4:5. Продукция, поступающая от первого поставщика, содержит 5% брака,

поступающая от второго поставщика — 6% брака, а поступающая от третьего поставщика — 8% брака. Покупатель оставил в книге пожеланий покупателя жалобу о низком качестве приобретённого товара. Найти вероятность того, что плохой товар, вызвавший нарекания покупателя, поступил от второго поставщика. (0,2)

5. Известно, что из числа зрителей определённой телепрограммы 70% смотрят и рекламные блоки. Группу, состоящую из трех наугад выбранных телезрителей, опрашивают относительно содержания рекламного блока. Рассчитать вероятность того, что в группе 2 человека смотрят рекламные блоки.(0,441)

Лекция №4

§ 15. Понятие случайной величины и её функции распределения вероятностей

Пример1а. Игральная кость подбрасывается 1 раз. Если выпадает четное число, игрок выигрывает 1 руб., если нечетное – проигрывает 1 руб. Какова вероятность того, что игрок выиграет?

$\Omega = \{1, 2, 3, 4, 5, 6\}$, $A = \{2, 4, 6\}$, $P(A) = 3/6=1/2$

ω_i	1	2	3	4	5	6
X	-1	+1	-1	+1	-1	+1

X	-1	+1
P	1/2	1/2

Определение. Случайной величиной X называется действительная *числовая* функция $X = X(\omega)$, определенная на множестве (пространстве) элементарных событий Ω и такая, что для любого $x\in\mathbb{R}$ множество тех ω, для которых $X(\omega) < x$, принадлежит алгебре событий данного эксперимента.

Случайные величины

Дискретные (ДСВ) Непрерывные (НСВ)

Определение. Функцией распределения вероятностей случайной величины X называется функция $F_X(x) = P(X < x)$

Пример1б. Условие примера 1а. Построить функцию распределения выигрыша (проигрыша) игрока.

$$F_X(x) = \begin{cases} 0, & x \le -1 \\ \frac{1}{2}, & -1 < x \le 1 \\ 1, & x > 1 \end{cases}$$

Свойства функции распределения:

1. $0 \le F_X(x) \le 1 \quad \forall x \in R$
2. $F_X(x_1) \le F_X(x_2) \quad \forall\, x_1 \le x_2$ (функция распределения – неубывающая функция)

   $F_X(x_2) = P(X < x_2) = P(X < x_1) + P(x_1 \le X \le x_2)$

   $P(x_1 \le X \le x_2) \ge 0 \;\; \Rightarrow \;\; P(X < x_2) \ge P(X < x_1)$

   Следствие: $P(x_1 \le X < x_2) = F_X(x_2) - F_X(x_1)$
3. $F_X(-\infty) = 0, \; F_X(+\infty) = 1$
4. Функция распределения F(x) непрерывна слева.

§16. Дискретные случайные величины (ДСВ)

Определение. Случайная величина X называется *дискретной случайной величиной*, если все её значения можно пронумеровать, т.е. $X = \{x_i\}$, $(i=1, 2, 3\dots)$

Определение. Законом распределения дискретной случайной величины (распределением) называется соответствие между ее возможными значениями и их вероятностями, то есть совокупность $\{x_i, p_i\}$, $(i=1, 2, 3\dots)$

а) Закон распределения может быть задан в виде таблицы - *ряда распределения ДСВ.*

Определение. *Рядом распределения ДСВ* называется таблица, первая строка которой содержит все возможные значения ДСВ, расположенные в порядке возрастания, а вторая строка соответствующие этим значениям вероятности.

x_i							
p_i							

Пример 1. Из урны, содержащей 4 белых и 6 черных шаров, извлекается 3 шара. Написать закон распределения (ряд распределения) для числа белых шаров среди извлеченных.

X= {число белых шаров среди извлеченных}

x_i	0	1	2	3
p_i	1/6	1/2	3/10	1/30

P(X=0) $= \frac{C_4^0 C_6^3}{C_{10}^3} = \frac{1}{6}$; P(X=1) $= \frac{C_4^1 C_6^2}{C_{10}^3} = \frac{1}{2}$; P(X=2) $= \frac{C_4^2 C_6^1}{C_{10}^3} = \frac{3}{10}$; P(X=3) $= \frac{C_4^3}{C_{10}^3} = \frac{1}{30}$

$\sum p_i=1$

$F_X(x) = P(X<x) = \sum_{x_i<x} p_i = \sum_{x_i<x} p\ (X = x_i)$

Алгоритм построения графика функции распределения ДСВ.

1.Строим декартовую систему координат.

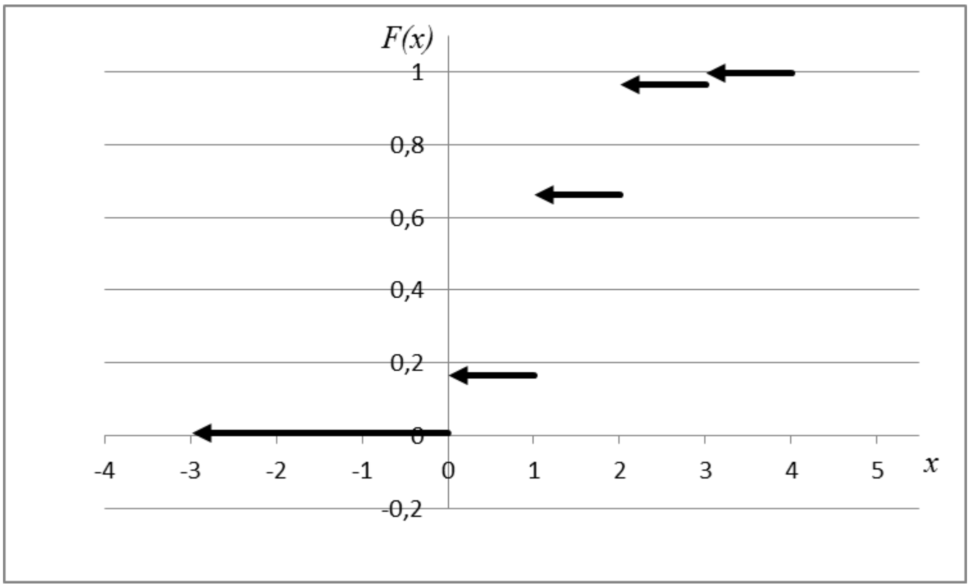
2.На оси x отмечаем возможные значения случайной величины x_i.

3.Слева от минимального значения (x_1) $F_X(x)=0$, справа от максимального - $F_X(x)=1$.

4.В точках x_i, соответствующих возможным значениям случайной величины, F(x) претерпевает скачок, величина которого равна вероятности этого значения p_i

5.Для дискретной случайной величины график F(x)-«лесенка».

6.Устраняем неоднозначность функции в точках разрыва (ставим стрелки).



$$F_X(x) = \begin{cases} 0, & x \le 0 \\ \frac{1}{6}, & -0 < x \le 1 \\ \frac{2}{3}, & 1 < x \le 2 \\ \frac{29}{30}, & 2 < x \le 3 \\ 1, & x > 3 \end{cases}$$

б) Закон распределения может быть задан *в виде формулы*.

Пример 2. Из урны, содержащей 4 белых и 6 черных шаров, извлекается 3 шара по схеме выборки с возвращением. Записать закон распределения для числа X белых шаров среди извлеченных.

Решение.

X= {число белых шаров среди извлеченных}

n=3, m-количество успехов (количество белых шаров в выборке)

X= m, где m =0, 1, 2, 3.

$P(X=m) = P_3(m) = C_3^m\ (0{,}4)^m (0{,}6)^{3-m}$

в) Закон распределения может быть задан *с помощью названия распределения*.

- Биномиальное распределение.

ДСВ X называется распределенной по биномиальному закону B (n, p) (имеет биномиальное распределение с параметрами n, p), если ее возможные значения X=0, 1, 2… n,

а вероятности вычисляются по формуле:

$P(X=m) = C_n^m p^m q^{n-m}$.

В примере 2 случайная величина X имеет биномиальное распределение с параметрами (3;0,4).

x_i	0	1	2	3
p_i	0,216	0,432	0,288	0,064

- Распределение Пуассона (закон редких явлений).

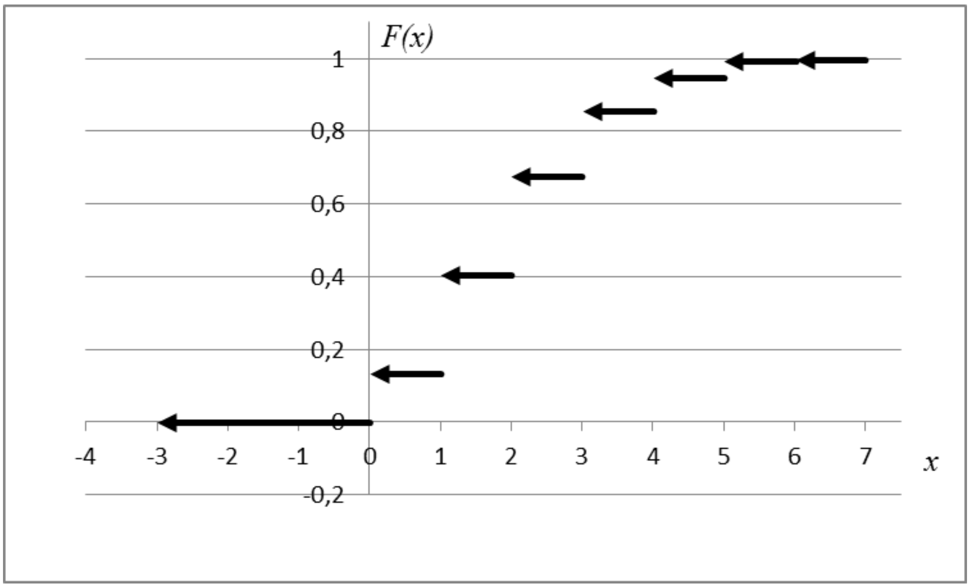
ДСВ X называется распределенной по закону Пуассона (имеет распределение Пуассона) с параметром λ >0, если ее возможные значения X=0,1,2…, а вероятности вычисляются по формуле:

$$P(X=m) = \frac{\lambda^m e^{-\lambda}}{m!}$$

Пример 3. ДСВ X имеет распределение Пуассона с параметром λ=2. Построить ряд распределения вероятностей X.

x_i	0	1	…	m	…
p_i	0,135	0,271	…	$\frac{2^m e^{-2}}{m!}$	…

$F_X(x)$-?



- Геометрическое распределение.

ДСВ Х называется распределенной по геометрическому закону (имеет геометрическое распределение) с параметром p, если она принимает значения Х=1,2,3…, а вероятности вычисляются по формуле:

P(X=m) = $(1-p)^{m-1}p$

Пример 4. Производится стрельба по мишени до первого попадания. Вероятность попадания при каждом выстреле равна 0,8. Считая боезапас неограниченным, найти закон распределения числа израсходованных патронов. (Геометрическое распределение с параметром p=0,8).

x_i	1	2	…	m	…
p_i	0,8	0,16	…	$(0,2)^{m-1}0,8$	…

- Гипергеометрическое распределение.

ДСВ Х называется распределенной по гипергеометрическому закону (имеет гипергеометрическое распределение) с параметрами n_1, n_2, n, если она принимает значения Х= $m_1, m_1+1,...m_2$, а вероятности вычисляются по формуле:

$$P\ (X=m) = \frac{C_{n_1}^{m} C_{n_2}^{n-m}}{C_{n_1+n_2}^{n}}$$

$m = m_1, m_1+1...m_2$, $\quad m_1 = \max\ (0; n-n_2)$, $\quad m_2 = \min\ (n_1; n)$

См. пример 1. Гипергеометрическое распределение с параметрами $n_1 = 4, n_2 = 6, n = 3$.

§17. Непрерывные случайные величины (НСВ)

Определение. Непрерывной случайной величиной X называется такая случайная величина, которая может принимать любые числовые значения в заданном интервале и для которой существует предел:

$\rho_X(x) = \lim\limits_{\Delta x \to 0} \dfrac{P(x \le X < x + \Delta x)}{\Delta x}$, называемый плотностью распределения вероятностей

Свойства плотности распределения вероятностей.

a) $\rho_X(x) \ge 0$

b) $\rho_X(x) = \dfrac{dF_X(x)}{dx}$, $F_X(x) = P(X < x) = \int\limits_{-\infty}^{x} \rho_X(t)dt$

c) $\int\limits_{-\infty}^{\infty} \rho_X(x)dx = 1$ - условие нормировки

d) $P(\alpha \le X < \beta) = F(\beta) - F(\alpha) = \int\limits_{\alpha}^{\beta} \rho_X(x)dx$

- Случайная величина X называется равномерно распределенной на отрезке $[a;\ b]$ (имеющей равномерное распределение с параметрами *a,b*), если её плотность распределения вероятностей имеет вид:

$$\rho_X(x) = \begin{cases} \dfrac{1}{b-a}, & при\, x \in [a;b] \\ 0, & при\, x \notin [a;b] \end{cases}$$

$F_X(x)$ -?

а) $x \le a$

$$F(x) = \int\limits_{-\infty}^{x} \rho(t)dt = 0$$

б) $a < x \le b$

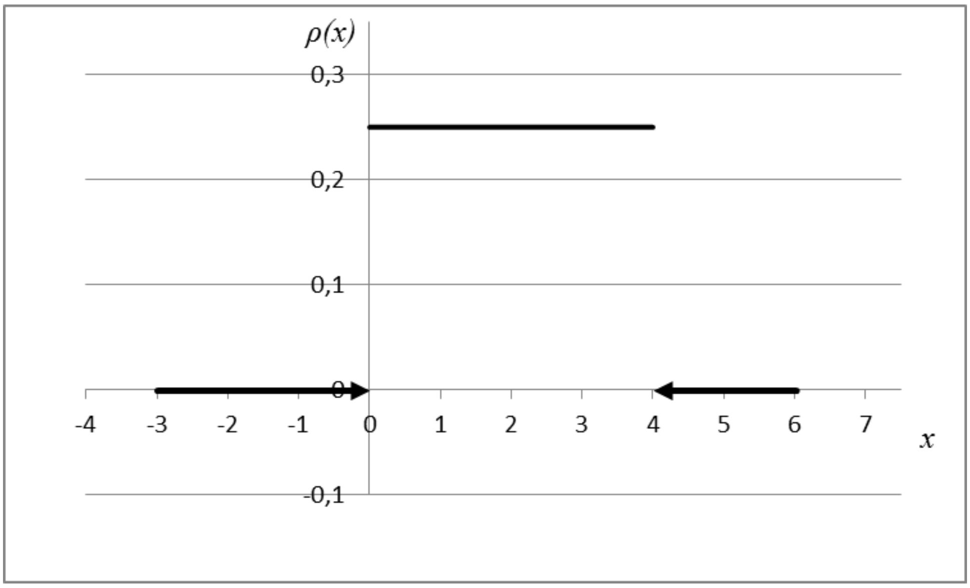
$$F(x) = \int\limits_{-\infty}^{x} \rho(t)dt = \int\limits_{-\infty}^{a} \rho(t)dt + \int\limits_{a}^{x} \rho(t)dt = \int\limits_{-\infty}^{a} 0dt + \int\limits_{a}^{x} \frac{dt}{b-a} = \frac{x-a}{b-a}$$

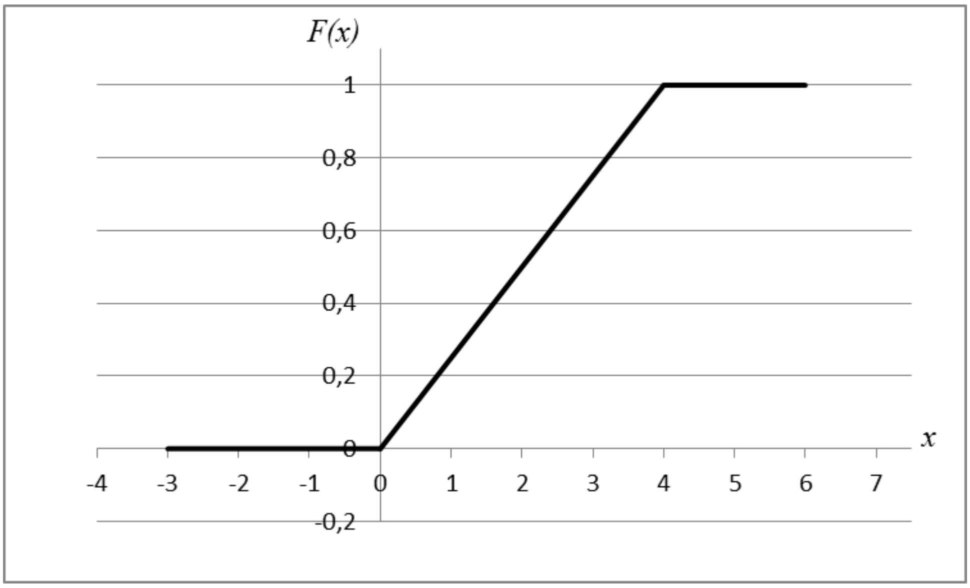
в) $x > b$

$$F(x) = \int\limits_{-\infty}^{x} \rho(t)dt = \int\limits_{-\infty}^{b} \rho(t)dt + \int\limits_{b}^{x} \rho(t)dt = F(b) + \int\limits_{b}^{x} 0dt = F(b) = \frac{b-a}{b-a} = 1$$

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

Графики $\rho_X(x)$ *и* $F_X(x)$*; a=0, b=4*





- Случайная величина X называется распределенной по показательному (экспоненциальному) закону с параметром $\lambda > 0$ (имеющей показательное распределение с параметром $\lambda > 0$), если её плотность распределения вероятности имеет вид:

$$\rho_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

$F_X(x)$ -?

а) $x \leq 0$

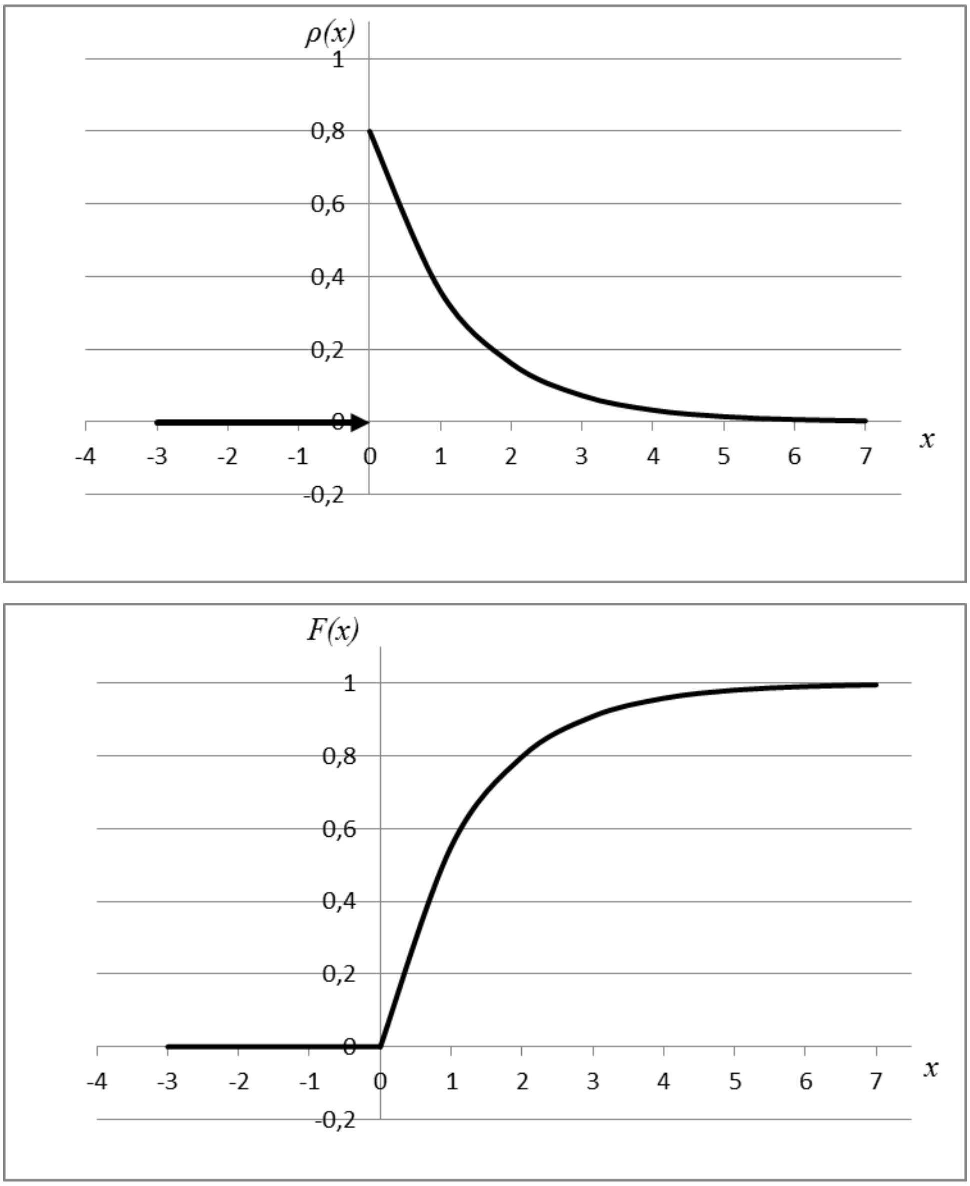
$F(x)=0$

б) $x>0$

$$F(x)=\int_{-\infty}^{x}\rho(t)dt=\int_{-\infty}^{0}0dt+\int_{0}^{x}\lambda e^{-\lambda t}dt=-\frac{\lambda e^{-\lambda t}}{\lambda}\Bigg|_{0}^{x}=-e^{-\lambda t}\Bigg|_{0}^{x}=-e^{-\lambda x}-(-e^{-\lambda 0})=1-e^{-\lambda x}$$

$$F(x)=\begin{cases}0, & x\le 0\\ 1-e^{-\lambda x}, & x>0\end{cases}$$

Графики $\rho_X(x)$ *и* $F_X(x)$; $\lambda=0,8$



- Случайная величина X называется распределенной по нормальному закону с параметрами a,σ (имеющей нормальное распределение с параметрами a,σ), если её плотность распределения вероятностей имеет вид:

$$\rho_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-a)^2}{2\sigma^2}} \qquad X \rightarrow N(a;\sigma)$$

Если $a = 0, \sigma = 1$, т.е. $X \rightarrow N(0;1)$, то говорят, что НСВ X имеет стандартное нормальное распределение.

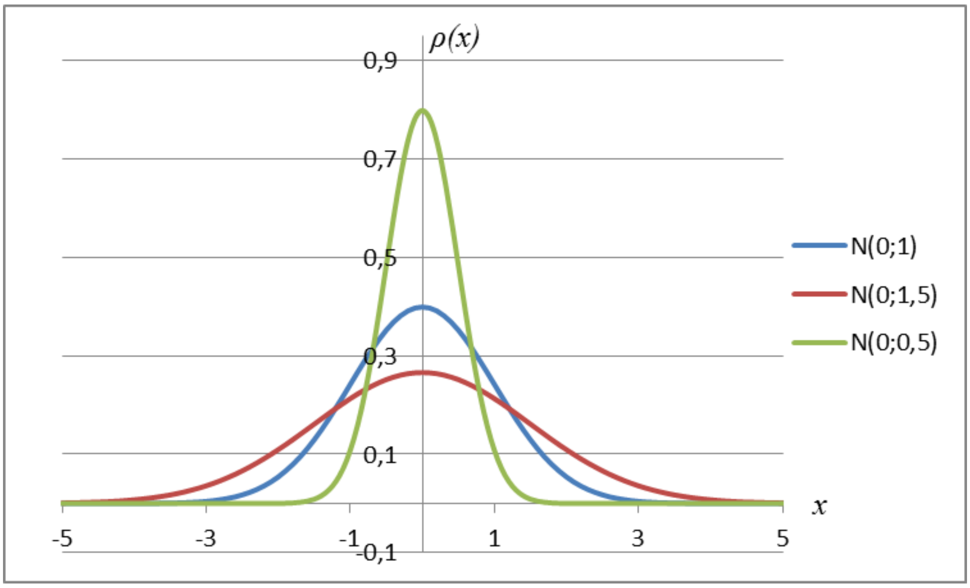
Пусть $X \rightarrow N(a;\sigma)$. $F_X(x)$-?

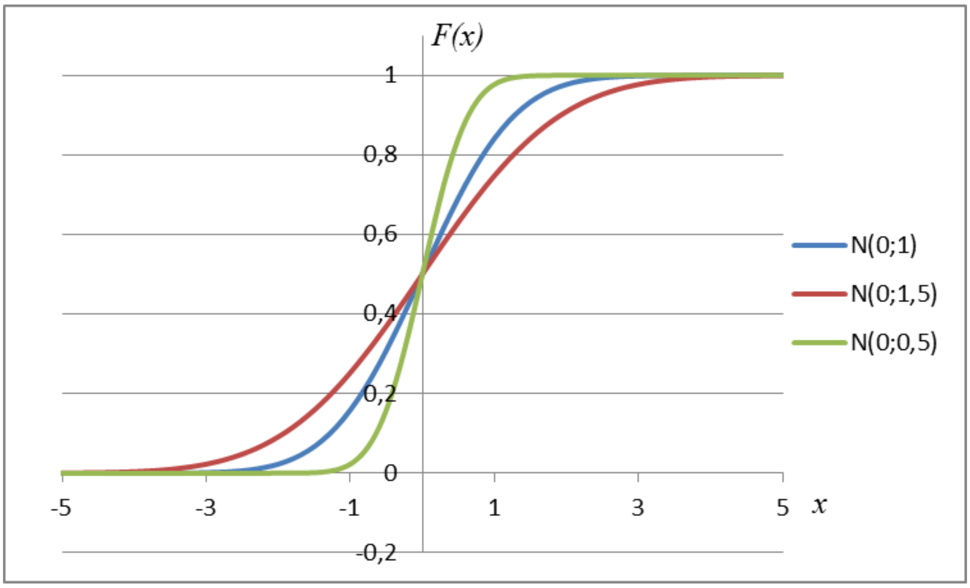
$$F(x) = \int_{-\infty}^{x} \rho(t)dt = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{\frac{-(t-a)^2}{2\sigma^2}} dt = \left\{ \begin{matrix} z = \dfrac{t-a}{\sigma} \\ dz = \dfrac{dt}{\sigma} \end{matrix} \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-a}{\sigma}} e^{-\frac{z^2}{2}} dz = \Phi\left(\frac{x-a}{\sigma}\right) = \frac{1}{2} + \Phi_0\left(\frac{x-a}{\sigma}\right)$$

$$F(x) = \Phi\left(\frac{x-a}{\sigma}\right) = \frac{1}{2} + \Phi_0\left(\frac{x-a}{\sigma}\right)$$

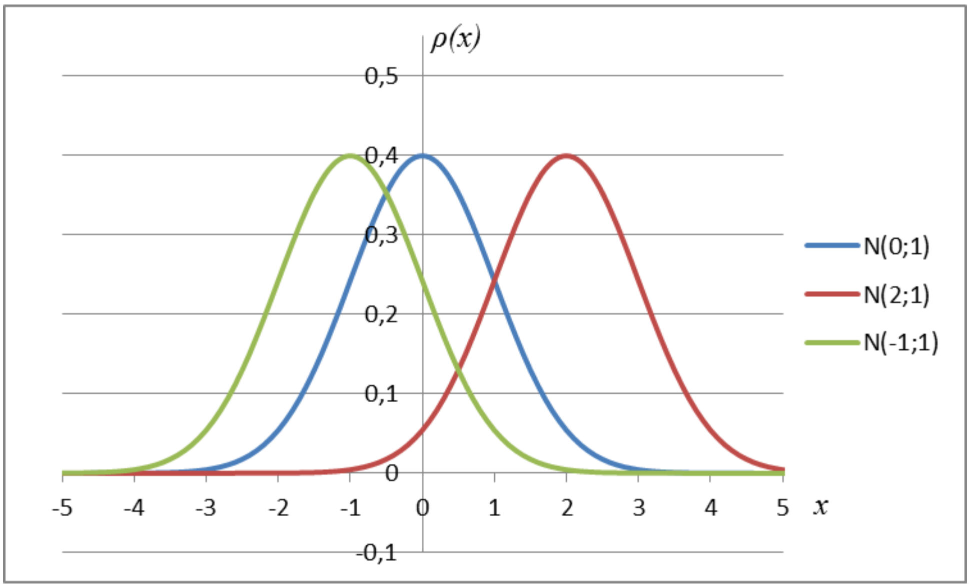
Если $X \rightarrow N(0;1)$, то $F_X(x) = \Phi(x) = \frac{1}{2} + \Phi_0(x)$

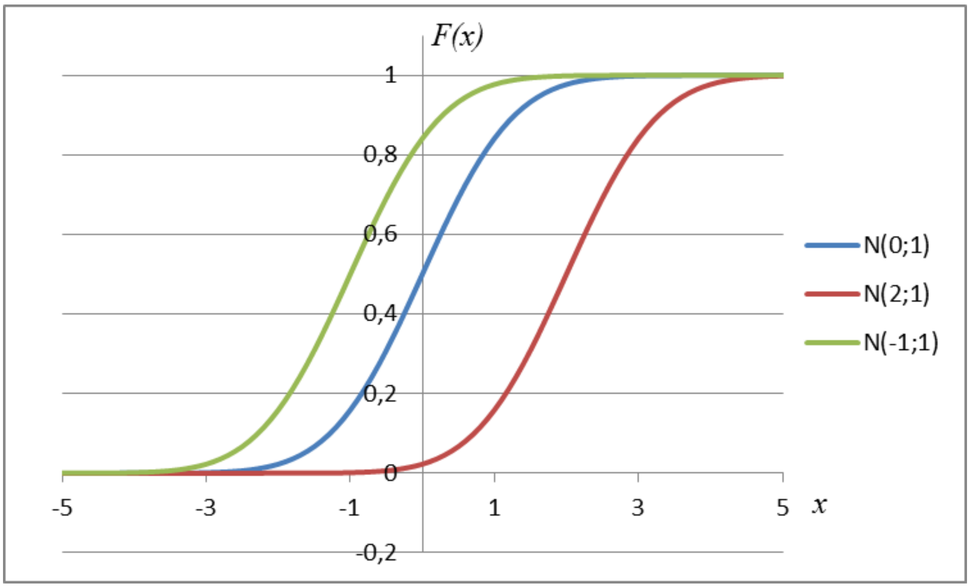
Влияние параметра σ:





Влияние параметра a:





Вероятность попадания значений X в заданный интервал (отрезок, промежуток)

$$P(\alpha \le x \le \beta) = F(\beta) - F(\alpha) = \int_{\alpha}^{\beta} \rho(x)dx$$

$$P(\alpha \le x \le \beta) = \Phi_0(\frac{\beta-\alpha}{\sigma}) - \Phi_0(\frac{\alpha-a}{\sigma}) = \Phi(\frac{\beta-a}{\sigma}) \ -\Phi(\frac{\alpha-a}{\sigma})$$

§18. Правило трех сигм

Пусть случайная величина X распределена по нормальному закону X→ $N(a;\sigma)$

$$P(\alpha \le x \le \beta) = F_X(\beta) - F_X(\alpha)$$

$$\alpha = a - \delta$$

$$\beta = a + \delta$$

$$P(\alpha \le x \le \beta) = \Phi_0(\frac{\beta-\alpha}{\sigma}) - \Phi_0(\frac{\alpha-a}{\sigma}) = \Phi_0(\frac{a+\delta-a}{\sigma}) \ -\Phi_0(\frac{a-\delta-a}{\sigma}) = 2\Phi_0(\frac{\delta}{\sigma})$$

Если $\delta = 3\sigma$, то

$$P(a-3\sigma \le x \le a+3\sigma) = 2\Phi_0(3) \approx 0{,}9973$$

Правило трёх сигм — практически все значения нормально распределённой случайной величины лежат в интервале $(a-3\sigma; a+3\sigma)$.

Лекция №5

§19. Математическое ожидание дискретной случайной величины

Определение. *Математическим ожиданием дискретной случайной величины* называется сумма произведений всех ее возможных значений на их вероятности, т.е.

$$m_x = M[X] = MX = \sum_{i=1}^{n} x_i p_i$$

Вероятностный смысл математического ожидания: оно дает представление о среднем значении случайной величины.

Пример 1. Найти математическое ожидание ДСВ X:

x_i	1	2	5
p_i	0,3	0,5	0,2

M[X]=1·0,3+2·0,5+5·0,2=2,3

Пример 2. Вероятность попадания в цель при одном выстреле равна 0,25. Стрельба ведется до первого попадания. Предполагая боезапас неограниченным, а результаты независимыми, найти среднее число израсходованных патронов. (Геометрическое распределение с параметром $p=0{,}25$)

Решение.

X - число израсходованных патронов

$p=0{,}25$, $q=1-p=0{,}75$

x_i	1	2	3	…	k	…
p_i	p	$q\cdot p$	$q^2\cdot p$	…	$q^{k-1}\cdot p$	…

$$MX=\sum_{m=1}^{\infty}mq^{m-1}p=p\sum_{m=1}^{\infty}\left(q^m\right)'=p\left(\sum_{m=1}^{\infty}q^m\right)'=p\left(\frac{q}{1-q}\right)'=p\frac{1}{(1-q)^2}=\frac{1}{p}$$

$$MX=\frac{1}{p},\ p=0{,}25\Rightarrow\ MX=\frac{1}{0{,}25}=4$$

Пример 2. Найти математическое ожидание случайной величины, имеющей распределение Пуассона.

Решение.

$$P(X=m)=\frac{\lambda^m}{m!}e^{-\lambda}$$

$$MX=\sum_{m=0}^{\infty}m\frac{\lambda^m}{m!}e^{-\lambda}=\sum_{m=1}^{\infty}m\frac{\lambda^m}{m!}e^{-\lambda}=\{k=m-1\}=\sum_{k=0}^{\infty}(k+1)\frac{\lambda^{k+1}}{(k+1)!}e^{-\lambda}=\lambda e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}=\lambda e^{-\lambda}e^{\lambda}=\lambda$$

$$MX=\lambda$$

§20. Свойства математического ожидания дискретной случайной величины

1. Математическое ожидание постоянной величины равно самой этой величине: $M[C]=C.$

Доказательство:

x_i	C
p_i	1

$$MX=M[C]=\sum_{i=1}^{1}C\cdot 1=C$$

2. Постоянный множитель можно вынести за знак математического ожидания: $M[CX]=CM[X].$

Доказательство:

$$M[CX]=\sum_{i=1}^{n}(Cx_i)p_i=C\sum_{i=1}^{n}x_i p_i=CM[X]$$

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

$$M[XY]=M[X]\cdot M[Y].$$

4. Математическое ожидание суммы двух случайных величин равно сумме математических ожиданий:

$$M[X+Y]=M[X]+M[Y].$$

Пример. Найти математическое ожидание случайной величины Х, имеющей биномиальный закон распределения с параметрами n и p.

1) $X=X_1+X_2+...+X_n$

X_i – число успехов в *i*-ом испытании (индикатор).

X_i	0	1
P_i	q	p

$q=1-p$, $\quad MX_i=0\cdot q+1\cdot p=p$

2) $MX=M[X_1+X_2+...+X_n]=\sum_{i=1}^{n}MX_i=np$, $\quad MX=np$

Для *гипергеометрическкo* распределения: $MX=n\dfrac{n_1}{n_1+n_2}$

§21. Математическое ожидание непрерывной случайной величины

Определение. *Математическим ожиданием непрерывной случайной величины X*, возможные значения которой принадлежит отрезку [*a*,*b*], называется определенный интеграл вида

$$MX=\int_a^b x\rho(x)dx$$

(если $a\to\infty$ *и* $b\to\infty$, то интеграл считается абсолютно сходящимся).

Для математического ожидания НСВ справедливы те же свойства, что и для ДСВ.

а) Математическое ожидание СВ, равномерно распределенной на отрезке [*a*,*b*]

$$MX=\int_a^b x\frac{1}{b-a}dx=\frac{1}{b-a}\left.\frac{x^2}{2}\right|_a^b=\frac{b^2-a^2}{2(b-a)}=\frac{b+a}{2}$$

б) Математическое ожидание СВ, распределенной по показательному закону с параметром λ.

$$MX=\int_0^\infty x\lambda e^{-\lambda x}dx=\{y=\lambda x,\quad dy=\lambda dx\}=\int_0^\infty \frac{y}{\lambda}\lambda e^{-y}\frac{1}{\lambda}dy=\frac{1}{\lambda}\int_0^\infty ye^{-y}dy=$$

$$\left\{\begin{array}{l}\int_a^в UdV=UV\Big|_a^в-\int_a^в VdU\\ \int_a^в U'Vdx=UV\Big|_a^в-\int_a^в V'Udx\end{array}\right\}=\left\{\begin{array}{ll}V=y & V'=1\\ U'=e^{-y} & U=-e^{-y}\end{array}\right\}=\frac{1}{\lambda}\left[-ye^{-y}\Big|_0^\infty-\int_0^\infty(-e^{-y})dy\right]=$$

$$\frac{1}{\lambda}\left[0+\int_0^\infty e^{-y}dy\right]=\frac{1}{\lambda}\int_0^\infty e^{-y}dy=\frac{1}{\lambda}(-e^{-y})\Big|_0^\infty=\frac{1}{\lambda}$$

в) Математическое ожидание СВ, имеющей нормальное распределение $N(a,\sigma)$

$$M[x]=\int_{-\infty}^{\infty}x\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-a)^2}{2\sigma^2}}dx=\int_{-\infty}^{\infty}(x-a)e^{-\frac{(x-a)^2}{2\sigma^2}}\frac{1}{\sqrt{2\pi}\sigma}dx+\frac{1}{\sqrt{2\pi}\sigma}a\int_{-\infty}^{\infty}e^{-\frac{(x-a)^2}{2\sigma^2}}dx=$$

$$=\int_{-\infty}^{\infty}te^{-\frac{t^2}{2\sigma^2}}\frac{1}{\sqrt{2\pi}\sigma}dt+\frac{1}{\sqrt{2\pi}}a\int_{-\infty}^{\infty}e^{-\frac{y^2}{2}}dy=a\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{y^2}{2}}dy=a\Phi(\infty)=a$$

§22. Дисперсия случайной величины

Определение. *Дисперсией случайной величины* называют математическое ожидание квадрата отклонения случайной величины от ее математического ожидания.

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

Замечание. Случайная величина $\overset{o}{X}=X-MX$ называется отклонением. $M[\overset{o}{X}]=0$.

Теорема. $D[X]=M[X^2]-(M[X])^2$

Доказательство.

$DX=M[X^2-2X\cdot M[X]+(M[X])^2]=M[X^2]-2M[X]\cdot M[X]+(M[X])^2=M[X^2]-(M[X])^2$

Пример 1. Найти дисперсию X:

x_i	1	2	5
p_i	0,3	0,5	0,2

$M[X]=1\cdot 0{,}3+2\cdot 0{,}5+5\cdot 0{,}2=2{,}3$

$DX=M[(X-MX)^2]$ — 1 способ

$DX=M[X^2]-(M[X])^2$ — 2 способ

а) 1 способ

X-MX	-1,3	-0,3	2,7
P	0,3	0,5	0,2

$(X\text{-}MX)^2$	1,69	0,09	7,29
P	0,3	0,5	0,2

$M[(X\text{-}MX)^2]=(-1,3)^2\cdot 0,3+(-0,3)^2\cdot 0,5+(2,7)^2\cdot 0,2=2,01$

б) 2 способ:

X^2	1	4	25
P	0,3	0,5	0,2

$M[X^2]=1\cdot 0,3+4\cdot 0,5+25\cdot 0,2=7,3$

$(M[X])^2=(2,3)^2=5,29$

$DX=M[X^2]-(M[X])^2=7,3-5,29=2,01$

Определение. Величина $\sigma= (D[X])^{1/2}$ называется средним квадратичным (квадратическим) отклонением.

§23.Свойства дисперсии

1) $D[X]\geq 0$

Доказательство.

Т.к. дисперсия есть математическое ожидание случайной величины, значения которой неотрицательны $D[X]=M[(X-M[X])^2]$, то $D[X]$ не может быть меньше 0.

2) $D[C]=0$, C-const

Доказательство.

$D[C]=M[(C-MC)^2]=M[0]=0$

3) $D[CX]=C^2D[X]$

Доказательство.

$D[CX]=M[(CX-M[CX])^2]=M[(CX)^2]-(M[CX])^2=C^2M[X^2]-C^2(MX)^2=C^2D[X]$

4) $D[X+Y]=D[X]+D[Y]$; (справедливо только для независимых случайных величин).

Доказательство.

$D[X+Y]=M[(X+Y)^2]-(M[X+Y])^2=M[X^2+2XY+Y^2]-(MX+MY)^2=$

$=M[X^2]+2MX\cdot MY+M[Y^2]-(MX)^2-2MX\cdot MY-(MY)^2= D[X]+D[Y]$

Замечание:

$D[X-Y]=D[X]+D[Y]$

$D[X-Y]=D[X+(-1)\cdot Y]=DX+D[-Y]=DX+(-1)^2\cdot DY=DX+DY$

Пример 1. Дисперсия случайной величины, распределенной по *равномерному закону*.

$$MX=\frac{a+b}{2}$$

$$M[X^2]=\int_a^b x^2\frac{1}{b-a}dx=\frac{1}{b-a}\frac{x^3}{3}\Big|_a^b=\frac{b^3-a^3}{b-a}=\frac{b^2+ab+a^2}{3}$$

$$DX=M[X^2]-(M[X])^2=\frac{b^2+ab+a^2}{3}-\frac{(a+b)^2}{4}=\frac{4b^2+4ab+4a^2-3a^2-6ab-3b^2}{12}=\frac{(b-a)^2}{12}$$

Пример 2. Дисперсия случайной величины, распределенной по *показательному закону*:

$$DX=\int_0^\infty x^2\lambda e^{-\lambda x}dx-\frac{1}{\lambda^2}=\frac{1}{\lambda^2}\left[\int_0^\infty t^2e^{-t}dt-1\right]=\frac{1}{\lambda^2}\left[t^2(-e^{-t})\Big|_0^\infty+\int_0^\infty e^{-t}2tdt-1\right]=\frac{1}{\lambda^2}(0+2-1)=\frac{1}{\lambda^2}$$

Пример 3. Дисперсия случайной величины, распределенной по *нормальному закону:*

$$DX=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}(x-a)^2e^{-\frac{(x-a)^2}{2\sigma^2}}dx=\left\{\frac{x-a}{\sigma}=z\right\}=\sigma^2\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}z^2e^{-\frac{z^2}{2}}dz=\begin{Bmatrix}U=z & U'=1\\ V'=ze^{\frac{-z^2}{2}} & V=e^{-\frac{z^2}{2}}\end{Bmatrix}=$$

$$=\sigma\frac{1}{\sqrt{2\pi}}\left\{-ze^{-\frac{z^2}{2}}\Big|_{-\infty}^{\infty}+\int e^{-\frac{z^2}{2}}dz\right\}=\sigma^2\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{z^2}{2}}dz=\sigma^2$$

$$DX=\sigma^2$$

Пример 4. Дисперсия случайной величины, распределенной по *биноминальному закону*:

X-число успехов в серии из n независимых испытаний.

$X=X_1+X_2+\ldots+X_n$

X_i- индикатор

X_i	0	1
P	q	p

$$MX_i=1\cdot p+0\cdot q=p$$

$$DX_i=M[X_i^2]-(MX_i)^2=(0\cdot q+1\cdot p)-p^2=p-p^2=pq$$

$$D[X]=D[X_1+X_2+\ldots+X_n]=\sum_{i=1}^n D[X_i]=npq$$

Для *геометрического* распределения: $DX=\dfrac{q}{p^2}$

Для распределения *Пуассона*: $DX=MX=\lambda$

Для *гипергеометрического* распределения: $DX=\dfrac{n_1+n_2-n}{n_1+n_2-1}n\dfrac{n_1}{n_1+n_2}\dfrac{n_2}{n_1+n_2}$

§24. Оценка отклонения теоретического распределения от нормального. Асимметрия и эксцесс

Определение. *Начальным моментом* порядка k случайной величины X называют математическое ожидание случайной величины X^k .

$\nu_k = M[X^k]$, k= 1, 2, 3 ….

ν_1= MX, ν_2=M[X^2], DX= $\nu_2 - \nu_1^2$

Определение. Центральным моментом порядка k случайной величины X называют математическое ожидание случайной величины $(X\text{-}MX)^k$

$\mu_k = M[(X\text{-}MX)^k]$, k= 1, 2, 3 ….

μ_1=0, μ_2= DX

Пусть ρ(x) – функция, график которой симметричен относительно x =MX, тогда любой центральный момент нечетного порядка будет равен 0. Тогда очевидно, что μ_k (k=2n+1) может служить мерой асимметричности распределения.

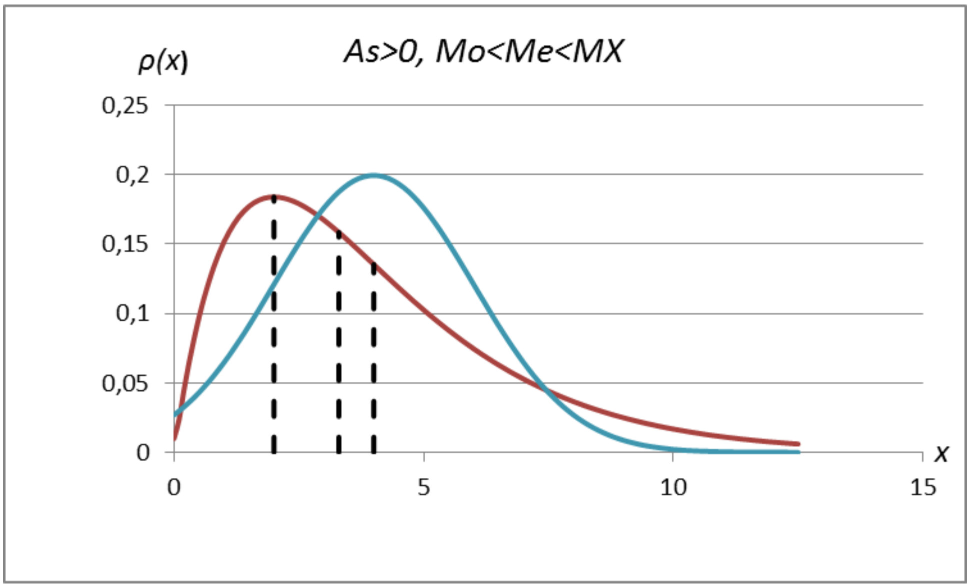
Определение. Асимметрией теоретического распределения называется отношение центрального момента третьего порядка к кубу среднеквадратичного отклонения:

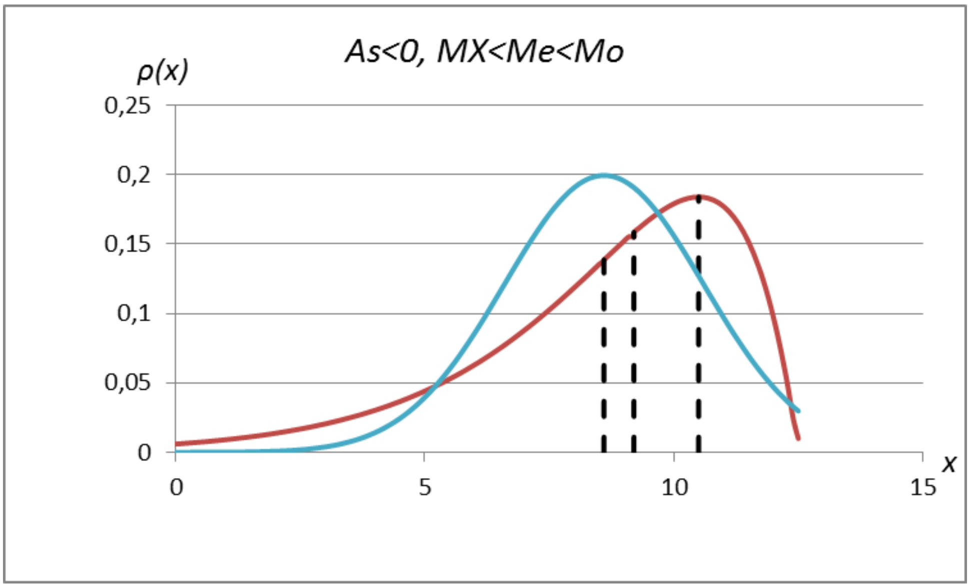
$$A_s = \frac{\mu_3}{\sigma^3}$$

Для нормального распределения A_s=0.

A_s>0 (правосторонняя асимметрия), если «длинная» часть кривой ρ(x) расположена справа от MX. A_s<0 (левосторонняя асимметрия), если «длинная» часть кривой ρ(x) - слева от MX. Вместо MX на практике используют моду распределения. *Мода* (*Мо*) – это значение случайной величины, имеющее наибольшую вероятность (для дискретных СВ) или плотность вероятности (для непрерывных СВ).

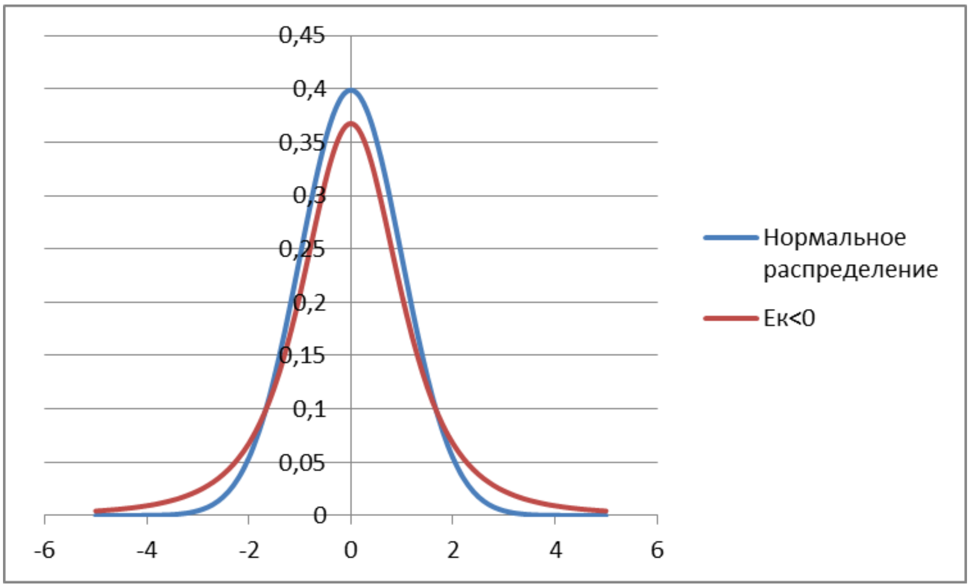
*Медианой непрерывной случайной величины X (**Me**)* называется такое ее значение x, для которого $F(x) = \frac{1}{2}$.





Определение. Эксцессом называют характеристику $E_к = \frac{\mu_4}{\sigma^4} - 3$

Для нормального распределения $\frac{\mu_4}{\sigma^4} = 3$, $E_к = 0$



Лекция №6.

Закон больших чисел(§25, §26)

Математические теоремы, в которых выясняются закономерности поведения суммы большого числа случайных величин, а также условия их (закономерностей) возникновения, получили общее название *закона больших чисел.*

§25. Неравенства Маркова и Чебышева

а) *неравенство Маркова*: для любой случайной величины X и $\varepsilon > 0$

$$P(|X| \geq \varepsilon) \leq \frac{M[|X|]}{\varepsilon}$$

Замечание 1: в основном неравенство Маркова формулируют для неотрицательных случайных величин в виде $P(X \geq \varepsilon) \leq \dfrac{M[X]}{\varepsilon}$

Замечание 2: неравенство Маркова иногда называют 1-ым неравенством Чебышева.

Доказательство:

Пусть X – непрерывная неотрицательная случайная величина

$$P(X > \varepsilon) = \int_{\varepsilon}^{\infty} \rho(x)dx \leq \int_{\varepsilon}^{\infty} \frac{x}{\varepsilon}\rho(x)dx \leq \frac{1}{\varepsilon}\int_{-\infty}^{\infty} x\rho(x)dx = \frac{M[X]}{\varepsilon}$$

Следствия:

1. $P(|X| \geq \varepsilon) = P(X^2 \geq \varepsilon^2) \leq \dfrac{M[X^2]}{\varepsilon^2} = \dfrac{\nu^2}{\varepsilon^2}$
2. $P(X < \varepsilon) \geq 1 - \dfrac{M[X]}{\varepsilon}$

Пример 1. Наблюдениями установлено, что средняя оценка студентов на контрольной работе по математике – 35 баллов.

а) оценить вероятность того, что наугад выбранный студент напишет контрольную на «отлично».

$$P(X \geq 45) \leq \frac{35}{45} = \frac{7}{9}$$

б) оценить вероятность того, что наугад выбранный студент не напишет контрольную на положительную оценку.

$$P(X < 25) \geq 1 - \frac{35}{25}$$

б) *неравенство Чебышева*: для любой случайной величины X и $\varepsilon > 0$

$$P(|X - a| \geq \varepsilon) \leq \frac{D[X]}{\varepsilon^2}$$, где $a = M[X]$

Доказательство:

Пусть X – непрерывная случайная величина

$$P\left(\left|X - a\right| \geq \varepsilon\right) = \int\limits_{|x-a|\geq\varepsilon} \rho_x(x)dx$$

$$P(|X - a| \geq \varepsilon) = \left\{\begin{matrix} |x - a| \geq \varepsilon \\ \begin{bmatrix} x \leq a - \varepsilon \\ x \geq a + \varepsilon \end{bmatrix} \end{matrix}\right\} \leq \int_{-\infty}^{a-\varepsilon} \frac{(x-a)^2}{\varepsilon^2}\rho(x)dx + \int_{a+\varepsilon}^{\infty} \frac{(x-a)^2}{\varepsilon^2}\rho(x) \leq \int\limits_{-\infty}^{\infty} \frac{(x-a)^2}{\varepsilon^2}\rho(x)dx =$$

$$= \frac{1}{\varepsilon^2}\int\limits_{-\infty}^{\infty}(x-a)^2\rho(x)dx = \frac{D[X]}{\varepsilon^2}$$

Следствие:

$$P(|X - a| < \varepsilon) \geq 1 - \frac{D[X]}{\varepsilon^2}$$

Пример 2. Устройство состоит из 10 независимо работающих элементов. Вероятность отказа каждого элемента за время T равна 0,05. С помощью неравенства Чебышева оценить вероятность того, что абсолютная величина разности между числом отказавших элементов и средним числом (математическим ожиданием) отказов за время T окажется: а) меньше двух; б) не меньше двух.

$M[X]$=10*0,05=0,5 $D[X] = 10*0,05*0,95 = 0,475$

а) $P(|X - 0{,}5| < 2) \geq 1 - \frac{0{,}475}{4} = 0{,}88$

Точное решение (по формуле Бернулли) $P(-1{,}5 < X < 2{,}5) = P(0 \leq X \leq 2) = 0{,}988$

б) $P(|X - 0{,}5| \geq 2) \leq \frac{0{,}475}{4} = 0{,}12$

Точное решение $P(|X-0{,}5|\geq 2)=1-0{,}988=0{,}012$

Замечание:

Доказательство неравенств для дискретных случайных величин проводится аналогично.

§26. Теорема Чебышева (Закон больших чисел в форме Чебышева)

Теорема.

Пусть

- $X_1, X_2, \ldots X_n$ – попарно-независимые случайные величины с какими угодно распределениями.
- $M[X_k]=a_k$, $D[X_k]=\sigma_k^2$, причем $D[X_k]\leq H$ $k=1, 2\ldots n$
- $\overline{X_n}=\frac{1}{n}(X_1+X_2+\ldots+X_n)$

Тогда

$\lim\limits_{n\to\infty}P(|\overline{X_n}-\overline{a})|>\varepsilon)=0$, где $\overline{a}=M[\overline{X_n}]$

Доказательство.

$$M[\overline{X_n}]=M[\frac{1}{n}(X_1+X_2+\ldots+X_n)]=\frac{1}{n}M[X_1+X_2+\ldots+X_n]=\frac{1}{n}(a_1+a_2+\ldots+a_n)$$

$$D[\overline{X_n}]=D[\frac{1}{n}(X_1+X_2+\ldots+X_n)]=\frac{1}{n^2}D[X_1+X_2+\ldots+X_n]=\frac{1}{n^2}(\sigma_1^2+\sigma_2^2+\ldots+\sigma_n^2)$$

Неравенство Чебышева для случайной величины $\overline{X_n}$ имеет вид

$$P(|\overline{X_n}-\overline{a}|\geq\varepsilon)\leq\frac{D[\overline{X_n}]}{\varepsilon^2}$$

по условию $D[X_k]\leq H\Rightarrow$

$$P(|\overline{X_n}-\overline{a}|\geq\varepsilon)\leq\frac{1}{n^2}\frac{nH}{\varepsilon^2}=\frac{H}{n\varepsilon^2}\Rightarrow$$

$$\lim_{n\to\infty}P(|\overline{X_n}-\overline{a}|\geq\varepsilon)=0$$

Пример 1. За значение некоторой величины принимают среднеарифметическое достаточно большого числа ее измерений. Предполагая, что среднеквадратичное отклонение возможных результатов каждого измерения не превосходит 5 мм, оценить вероятность того, что при 1000 измерений неизвестной величины отклонение принятого значения от истинного по абсолютной величине не превзойдет 0,5 мм.

$$P(|\overline{X_n} - \bar{a}| < \varepsilon) \geq 1 - \frac{H}{n\varepsilon^2}$$

$$P(|\overline{X_n} - \bar{a}| < 0{,}5) \geq 1 - \frac{25}{1000 * 0{,}25} = 0{,}9$$

Следствия:

1. $\lim\limits_{n\to\infty} P(|\overline{X_n} - \bar{a}| < \varepsilon) = 1$

2. Закон больших чисел в форме Бернулли

(m_n – число успехов в серии из n независимых испытаний, p-вероятность успеха в каждом отдельном испытании)

$$\lim_{n\to\infty} P(|\frac{m_n}{n} - p| \geq \varepsilon) = 0$$

т.к. $D[X_k] = npq \quad \to \quad P(|\frac{m_n}{n} - p| \geq \varepsilon) \leq \frac{pq}{n\varepsilon^2}$

Пример 2. Вероятность того, что изделие является качественным, равна 0,9. Сколько следует проверить изделий, чтобы с вероятностью не меньшей 0,95 можно было утверждать, что абсолютная величина отклонения доли качественных изделий от 0,9 не превысит 0,01?

$$P(|\frac{m_n}{n} - p| < \varepsilon) \geq 1 - \frac{pq}{n\varepsilon^2}$$

$$P(|\frac{m_n}{n} - 0{,}9| < 0{,}01) \geq 1 - \frac{0{,}9 * 0{,}1}{n * 0{,}0001} = 1 - \frac{900}{n} \geq 0{,}95$$

$$n \geq 18000$$

3. Закон больших чисел в форме Пуассона

(m_n – число успехов в серии из n независимых испытаний, p_i-вероятность успеха в i-ом отдельном испытании)

$$\lim_{n\to\infty} P(|\frac{m_n}{n} - \frac{1}{n}\sum_{i=1}^{n} p_i| \geq \varepsilon) = 0$$

(при доказательстве используется неравенство $P(|\frac{m_n}{n} - \frac{1}{n}\sum_{i=1}^{n} p_i| \geq \varepsilon) \leq \frac{1}{4n\varepsilon^2}$)

§27. Понятие о центральной предельной теореме (ЦПТ)

ЦПТ Ляпунова устанавливает общие достаточные условия, при которых суммы независимых случайных величин имеют асимптотически нормальное распределение.

Теорема. Если случайные величины $X_1, X_2, \ldots X_n$ независимы, одинаково распределены и имеют конечную дисперсию, то при $n \to \infty$ равномерно по всей числовой оси $x \in (-\infty;+\infty)$ вероятность того, что

$$P\left(\frac{X_1+X_2+...+X_n-an}{\sqrt{n}\sigma}<x\right) \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt$$

где $a=M[X_n]$, $\sigma=\sqrt{D[X_n]}$

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

$$P\left(\frac{\overline{X_n}-a}{\sigma/\sqrt{n}}<x\right)\to\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt$$

$$P(Y<x)\to\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt$$

$$\frac{\overline{X_n}-a}{\sigma/\sqrt{n}}\to N(0;1)$$

$$\overline{X_n}\to N\left(a;\sigma/\sqrt{n}\right)$$

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. Примером такой случайной величины может служить случайная ошибка прямого измерения.

§28. Распределение функции одного случайного аргумента

Определение. Пусть X – случайная величина. Если каждому возможному значению X соответствует одно возможное значение случайной величины Y, то Y – функция случайного аргумента X ($Y=\varphi(X)$)

1) X – дискретная случайная величина.

Пример 1. $Y = X^2$

x_i	- 2	2	3
p_i	0,4	0,5	0,1

y_i	4	9
p_i	0,9	0,1

Вероятности соответствующих значений X и Y равны, если двум значениям X соответствует одно значение Y, то соответствующие вероятности складываются.

2) X – непрерывная случайная величина с плотностью распределения вероятностей $\rho_X(x)$

$Y=\varphi(X)$

Если *y=φ(x)* – дифференцируемая, *строго убывающая или строго возрастающая функция*, обратная которой $x=\varphi^{-1}(y)=\psi(y)$. Тогда плотность распределения случайной величины Y вычисляется по формуле:

$$\rho_Y(y)=\rho_X(\psi(y))|\psi'(y)|$$

Пример 2. $X\to N(a,\sigma)$, $Y=AX+B$, $\rho_Y(y)=?$

Плотность случайной величины *X*	$\rho_X(x)$	$\rho_X(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-(x-a)^2}{2\sigma^2}}$
Функциональная зависимость между случайными величинами *X* и *Y*.	*y=φ(x)*	$y=Ax+B$
Обратная функция	$x=\varphi^{-1}(y)=\psi(y)$	$x=\frac{y-B}{A}$
Модуль производной обратной функции	$\lvert\psi_i'(y)\rvert$	$\frac{1}{\lvert A\rvert}$
Плотность случайной величины Y	$\rho_Y(y)=\rho_X(\psi(y))\lvert\psi'(y)\rvert$	$\frac{1}{\sqrt{2\pi}\sigma_Y}e^{-\frac{(y-a_Y)^2}{2\sigma_Y^2}}$

$$\rho_Y(y)=\frac{1}{\sqrt{2\pi}\sigma}e^{-(\frac{y-B}{A}-a)^2/2\sigma^2}\left|\frac{1}{A}\right|=\frac{1}{\sqrt{2\pi}\sigma|A|}e^{-\frac{(y-B-aA)^2}{2\sigma^2A^2}}=\frac{1}{\sqrt{2\pi}\sigma_Y}e^{-\frac{(y-a_Y)^2}{2\sigma_Y^2}}$$

$MY=a_Y=B+aA$

$DY=\sigma_Y^2=\sigma^2|A|^2$

Замечание 1. Если функция *y=φ(x)* не монотонна на интервале возможных значений *X*, то этот интервал следует разбить на интервалы монотонности, найти плотность для каждого интервала, а затем результаты просуммировать.

$$\rho_Y(y)=\sum_{i=1}^{k}\rho_X(\psi_i(y))|\psi_i'(y)|$$

Пример 3. Случайная величина *X* распределена равномерно в интервале $\left(-\frac{\pi}{2};\frac{\pi}{2}\right)$. Найти плотность распределения случайной величины $Y=\cos(X)$.

Функция $y=\cos x$ немонотонная в интервале $\left(-\frac{\pi}{2};\frac{\pi}{2}\right)$, ее значения лежат в интервале (0,1).

Плотность случайной величины *X*	$\rho_X(x)$	$\rho_X(x)=\begin{cases}\frac{1}{\pi}, & при\, x\in\left(\frac{-\pi}{2};\frac{\pi}{2}\right)\\ 0, & при\, x\notin\left(\frac{-\pi}{2};\frac{\pi}{2}\right)\end{cases}$
Функциональная зависимость между случайными величинами *X* и *Y*.	*y=φ(x)*	$y=\cos x$
Обратная функция	$x=\begin{cases}\psi_1(y)\\ \psi_2(y)\end{cases}$	$\begin{cases}\psi_1(y)=\arccos(y)\\ \psi_2(y)=-\arccos(y)\end{cases}$
Модуль производной обратной функции	$\lvert\psi_i'(y)\rvert$	$\frac{1}{\sqrt{1-y^2}}$
Плотность случайной величины Y	$\rho_Y(y)=\sum_{i=1}^{k}\rho_X(\psi_i(y))\lvert\psi_i'(y)\rvert$	$\begin{cases}\frac{2}{\pi\sqrt{1-y^2}}, & при\, y\in(0;1)\\ 0, & при\, y\notin(0;1)\end{cases}$

$$\rho_Y(y)=\sum_{i=1}^{k}\rho_X(\psi_i(y))\lvert\psi_i'(y)\rvert=\frac{1}{\pi}\frac{1}{\sqrt{1-y^2}}+\frac{1}{\pi}\frac{1}{\sqrt{1-y^2}}=\frac{1}{\pi}\frac{2}{\sqrt{1-y^2}}$$

Замечание 2. Если X – дискретная случайная величина, *Y =φ(X)*, то

$$M[Y(X)]=M[Y]=\sum_{i=1}^{n}\varphi(x_i)\,p_i \qquad D[Y(X)]=D[Y]=\sum_{i=1}^{n}\varphi^2(x_i)\,p_i-(\sum_{i=1}^{n}\varphi(x_i)\,p_i)^2$$

Если *X* – непрерывная случайная величина, *Y =φ(X*), то

$$M[Y]=\int_{-\infty}^{+\infty}\varphi(x)\rho_X(x)dx \qquad D[Y]=\int_{-\infty}^{+\infty}\varphi^2(x)\rho_X(x)dx-(\int_{-\infty}^{+\infty}\varphi(x)\rho_X(x)dx)^2$$

§29. Распределение функции двух случайных аргументов

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

Рассмотрим далее функцию *Z=X+Y*. Плотность распределения суммы независимых случайных величин называют композицией.

1. Пусть *X* и *Y* – дискретные независимые случайные величины. Для того, чтобы составить закон распределения функции *Z=X+Y*, надо найти все возможные значения *Z* и их вероятности. Т.к. *X* и *Y* независимые случайные величины, то $z_i=x_i+y_i$, $p(z_i)=p(x_i)*p(y_i)$. Если $z_i=z_j$, то их вероятности складываются.

Пример 1.

x_i	1	3

p_i	0,1	0,9

y_i	0	1	2
p_i	0,2	0,3	0,5

z_i	1	2	3	4	5
p_i	0,02	0,03	0,23	0,27	0,45

2. Пусть *X* и *Y* – непрерывные случайные величины. Если *X* и *Y* независимы, то плотность распределения $\rho_Z(z)$ суммы Z=X+Y (при условии, что плотность хотя бы одного из аргументов задана на интервале (-∞;∞) одной формулой) может быть найдена с помощью формулы:

$$\rho_Z(z) = \int_{-\infty}^{\infty} \rho_X(x)\rho_Y(z-x)dx = \int_{-\infty}^{\infty} \rho_X(z-y)\rho_Y(y)dy$$

Если возможные значения аргументов неотрицательны, то

$$\rho_Z(z) = \int_0^z \rho_X(x)\rho_Y(z-x)dx = \int_0^z \rho_X(z-y)\rho_Y(y)dy$$

Определение. Закон распределения вероятностей называют устойчивым, если композиция таких законов есть тот же закон (возможно отличающийся параметрами).
MZ=MX+MY; DZ=DX+DY.
Пример 2.
Доказать, что нормальный закон распределения – устойчивый.

$$\rho_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} e^{\frac{-(x-a)^2}{2\sigma_X{}^2}} \qquad \rho_Y(x) = \frac{1}{\sqrt{2\pi}\sigma_Y} e^{\frac{-(x-b)^2}{2\sigma_Y{}^2}}$$

$$\rho_Z(z) = \int_{-\infty}^{\infty} \rho_X(x)\rho_Y(z-x)dx = \frac{1}{\sqrt{2\pi}\sigma_X}\frac{1}{\sqrt{2\pi}\sigma_Y}\int_{-\infty}^{\infty} e^{\frac{-(x-a)^2}{2\sigma_X^2}} e^{\frac{-(z-x-b)^2}{2\sigma_Y^2}} dx = \frac{1}{\sqrt{2\pi(\sigma_Y^2+\sigma_X^2)}} e^{\frac{-(z-a-b)^2}{2(\sigma_Y^2+\sigma_X^2)}}$$

Пример 3.

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

$$\rho_Z(z) = \int_0^z \rho_X(x)\rho_Y(z-x)dx = \int_0^z \lambda e^{-\lambda x}\lambda e^{-\lambda(z-x)}dx = \lambda^2 z e^{-\lambda z}$$

$$\rho_Z(z) = \begin{cases} \lambda^2 z e^{-\lambda z}, & z \ge 0 \\ 0, & z < 0 \end{cases}$$

Замечание Можно показать, что, если $X_1, X_2, \dots X_n$ независимые случайные величины, имеющие показательное распределение с параметром λ>0, то $Z= X_1+X_2+ \dots +X_n$ имеет гамма-распределение (Г-распределение) с параметрами λ>0 и η=n

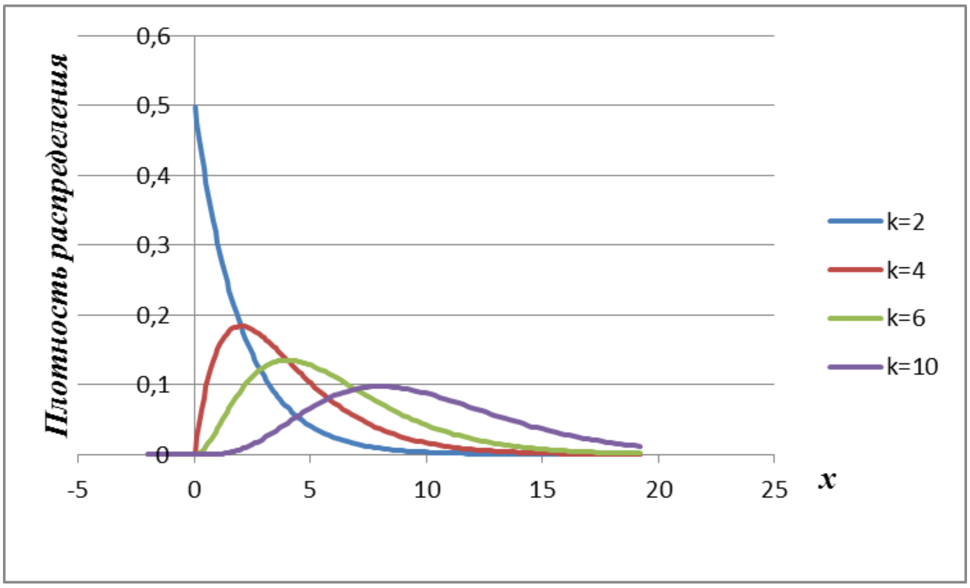
$$\rho_Z(z)=\begin{cases}\dfrac{\lambda^n}{(n-1)!}z^{n-1}e^{-\lambda z}, & z\ge 0\\ 0, & z<0\end{cases}$$

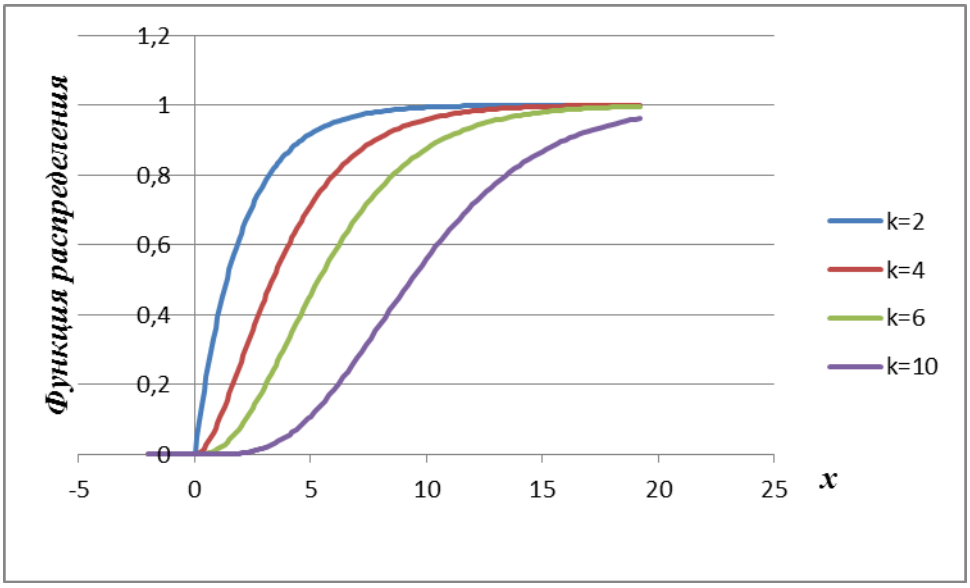
§30. Распределения χ^2 (хи-квадрат), Стьюдента, Фишера-Снедекора.

1) Распределение χ^2.

Определение. Пусть X_i *(i=1,2,…n)* - независимые случайные величины, имеющие стандартное нормальное распределение *N(0;1)*; тогда случайная величина $\chi^2=\sum_{i=1}^{n}X_i^2$ имеет закон распределения $\chi^2(k)$ с *k=n* степенями свободы.

Замечание. Если величины X_i связаны дополнительными соотношениями, например, $\dfrac{\sum X_i}{n}=\overline{X}$, то количество степеней свободы *k=n-1*.





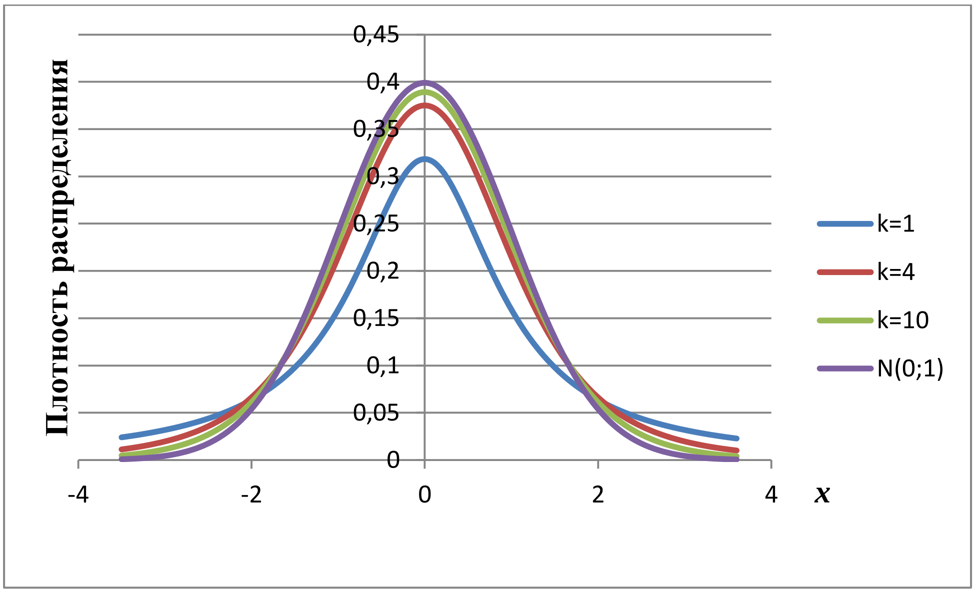
С увеличением k распределение медленно приближается к нормальному распределению.

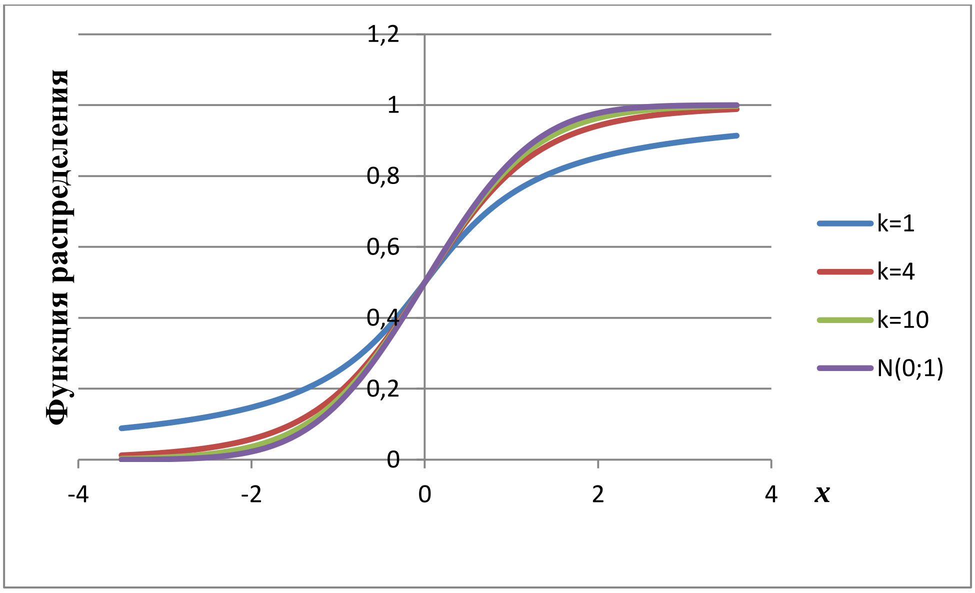
$M[\chi^2] = k$

$D[\chi^2] = 2k$

2) Распределение Стьюдента.

Пусть Z - случайная величина со стандартным нормальным распределением $N(0;1)$, V- независимая от Z величина, имеющая распределение χ^2 с k степенями свободы. Тогда случайная величина $T = \dfrac{Z}{\sqrt{V/k}}$ имеет распределение Стьюдента (t-распределение) с k степенями свободы.





$MT = 0 \quad (k > 1)$

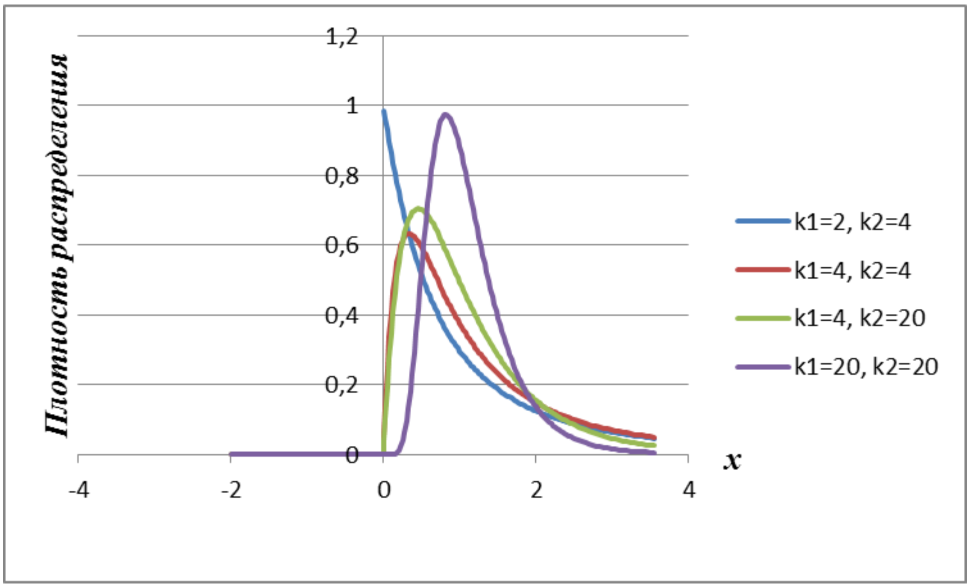
$$DT = \frac{k}{k-2} \quad (k > 2)$$

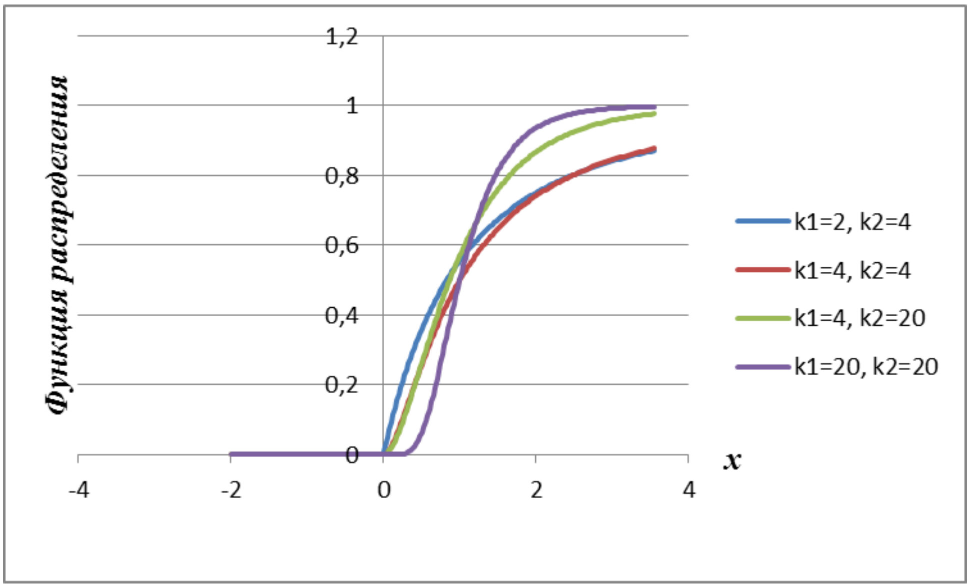
$$E_k = \frac{3(k-2)}{k-4}$$

При увеличении количества степеней свободы *k* распределение Стьюдента быстро стремится к стандартному нормальному распределению. На практике распределение Стьюдента заменяют нормальным при количестве степеней свободы *k*>30.

3) Распределение Фишера-Снедекора

Пусть *U*,*V*- независимые случайные величины, распределённые по закону χ^2 с k_1, k_2 степенями свободы соответственно, тогда $F = \frac{U/k_1}{V/k_2}$ имеет распределение Фишера-Снедекора с количеством степеней свободы k_1, k_2 $(F(k_1, k_2))$.





$$MF = \frac{k_2}{k_2 - 2}$$

$$DF = \frac{2k_2^2(k_1 + k_2 - 2)}{k_1(k_2 - 2)^2(k_2 - 4)} \qquad k_2 > 4$$

Распределение Фишера медленно стремится к нормальному распределению при увеличении числа степеней свободы k_1, k_2 .

Типовой вариант контрольной работы №2 по теме «Случайные величины»

1. Наибольшее давление в камере порохового ракетного двигателя распределено нормально со среднеквадратичным отклонением 10 кг/см^2. Если давление больше 105 кг/см^2, то двигатель возвращается на завод для регулирования. В партии двигателей 5% были возвращены на завод. Найти среднее значение наибольшего давления в камере порохового ракетного двигателя.

Ответ: a=88,55

2. Случайная величина имеет плотность распределения $\rho(x)=\dfrac{c}{x^2+6x+11}$.

Найти величину параметра *c*.

Ответ: $c=\dfrac{\sqrt{2}}{\pi}$

3. Найти дисперсию числа выпадений герба при трех бросаниях симметричной монеты.

Ответ: DX=3/4

4. Мишень состоит из круга и двух колец с номерами “2” и “3”. Попадание в круг дает 3 очка, в кольцо “2”- 2 очка, в кольцо “3”- 1 очко. Вероятности попадания в круг, кольца “2”и “3” соответственно равны 0,2; 0,3; 0,5. Построить ряд распределения для случайной суммы выбитых очков в результате двух попаданий.

Ответ:

x_i	2	3	4	5	6
p_i	0,25	0,3	0,29	0,12	0,04

5. Случайная величина X имеет плотность распределения $\rho(x)=\begin{cases} 0, x<0 \\ \sqrt{\dfrac{2}{\pi}}x^2 e^{-\frac{x^2}{2}}, x\ge 0 \end{cases}$.

Найти математическое ожидание X.

Ответ: $MX=\dfrac{4}{\sqrt{2\pi}}$

6. Случайная величина X имеет стандартное нормальное распределение. Найти закон распределения случайной величины $Y=X^3$.

Ответ: $\rho(y)=\dfrac{1}{3\sqrt{2\pi}}y^{-\frac{2}{3}}e^{-\frac{\sqrt[3]{y^2}}{2}}$

Лекция №8.

§31. Системы случайных величин. Двумерные случайные величины.

Понятие о системе случайных величин.

Многомерная случайная величина $\boldsymbol{X}=(X_1,X_2,\dots,X_k)$ — это совокупность *k* случайных величин X_i, или случайный вектор в *k*-мерном пространстве с координатами ($X_1, X_2, \dots, X_k$).

Определение. Закон распределения вероятностей многомерной случайной величины задаётся её функцией распределения

$F_{\boldsymbol{X}}(\boldsymbol{x})=P\left(X_1<x_1 \cap X_2<x_2 \cap \dots \cap X_k<x_k\right)$,

которая является детерминированной неотрицательной числовой функцией многих переменных $\boldsymbol{x}=(x_1,x_2,\dots,x_k)$ и принимает значения на отрезке [0; 1] .

Свойства функции распределения многомерной случайной величины.

1^0. $F_{\boldsymbol{X}}(\boldsymbol{x})=0$, если среди x_j хотя бы одна компонента равна $-\infty$.

2^0. $F_{\boldsymbol{X}}(\boldsymbol{x})=1$, если все компоненты равны ∞.

3^0. $F_{\boldsymbol{X}}(\boldsymbol{x})$ не убывает по каждому аргументу.

4^0. $F_X(x)$ непрерывна слева по каждому аргументу.

5^0. С помощью $F_X(x)$ можно вычислить вероятность попадания случайной точки в k-мерный параллелепипед.

Многомерные случайные величины, так же, как и одномерные, могут быть *дискретными* (когда наборы возможных значений образуют конечное или счётное множество) или *непрерывными* (когда множество наборов возможных значений несчётно).

Всюду ниже в данном параграфе будут рассматриваться *двумерные случайные величины*.

Закон распределения вероятностей дискретной двумерной случайной величины.

Двумерные дискретные случайные величины удобно задавать с помощью таблиц распределения

	x_1	x_2	…	x_n
y_1	p_{11}	p_{12}	…	p_{1n}
y_2	p_{21}	p_{22}	…	p_{2n}
…	…	…	…	…
y_m	p_{m1}	p_{m2}	…	p_{mn}

В такой таблице заголовки столбцов x_j соответствуют всем возможным значениям первой компоненты X, а названия строк y_i - всем возможным значениям второй компоненты Y. При этом в клетку, находящуюся в i -й строке и в j -м столбце, записывается значение вероятности

$$p_{ij} = P\left(X = x_j \cap Y = y_i\right)$$

Естественно, $\sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij} = 1$

Функция распределения двумерной дискретной случайной величины равна

$$F_{X,Y}(x,y) = \sum_{x_j<x}\sum_{y_i<y} p_{ij}$$

Вероятность попадания двумерной случайной величины (X,Y) в открытый прямоугольник равна

$$P\{(a_1 < X < b_1) \cap (a_2 < Y < b_2)\} = F_{X,Y}(b_1,b_2) - F_{X,Y}(a_1,b_2) - F_{X,Y}(b_1,a_2) + F_{X,Y}(a_1,a_2)$$

Законы распределения каждой из компонент такой двумерной случайной величины (так называемые *маргинальные* или *частные* законы распределения) восстанавливаются по таблице распределения при помощи формул

$$p(x_j) = P\left(X = x_j\right) = \sum_{i=1}^{m} p_{ij} \qquad p(y_i) = P\left(Y = y_i\right) = \sum_{j=1}^{n} p_{ij}$$

	x_1	x_2	…	x_n	
y_1	p_{11}	p_{12}	…	p_{1n}	$p(y_1)$
y_2	p_{21}	p_{22}	…	p_{2n}	$p(y_2)$
…	…	…	…	…	…
y_m	p_{m1}	p_{m2}	…	p_{mn}	$p(y_m)$
	$p(x_1)$	$p(x_2)$	…	$p(x_n)$	

Пример 1. Найти частные законы распределения компонент двумерной случайной величины, заданной таблицей распределения. Вычислить значение функции совместного распределения в точке (*3;4,5*).

Y \ X	1	3	5
2	0,1	0,3	0,2
4	0,06	0,18	0,16

Решение.

$$F_{XY}(3;4{,}5)=\sum_{x_j<3}\sum_{y_i<4{,}5}p_{ij}=0{,}1+0{,}06=0{,}16$$

	1	3	5	
2	0,1	0,3	0,2	$p(y_1)=0{,}6$
4	0,06	0,18	0,16	$p(y_2)=0{,}4$
	$p(x_1)=0{,}16$	$p(x_2)=0{,}48$	$p(x_3)=0{,}36$	

X	1	3	5
p	0,16	0,48	0,36

Y	2	4
p	0,6	0,4

Закон распределения вероятностей непрерывной двумерной случайной величины.

Двумерная случайная величина называется *непрерывной*, если её функция распределения может быть представлена в виде

$$F_{XY}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} \rho_{XY}(t_1,t_2)dt_1dt_2$$

Функция $\rho_{XY}(x,y)$ называется плотностью распределения двумерной случайной величины ($X;Y$).

Свойства плотности распределения двумерной случайной величины.

1^0. $\rho_{XY}(x,y) \ge 0$

2^0. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \rho_{XY}(t_1,t_2)dt_1dt_2 = 1$

3^0. $\rho_{XY}(x,y) = \dfrac{\partial^2 F_{XY}(x,y)}{\partial x \partial y}$

4^0. $\int_G \rho_{XY}(t_1,t_2)dt_1dt_2$ - вероятность попадания значений случайной величины в область G.

Если непрерывная двумерная случайная величина ($X;Y$) имеет плотность $\rho_{XY}(x,y)$, то одномерные случайные величины X и Y также являются непрерывными, и их плотности можно рассчитать по формулам

$$\rho_X(x) = \int_{-\infty}^{\infty} \rho_{XY}(x,y)dy \qquad \rho_Y(y) = \int_{-\infty}^{\infty} \rho_{XY}(x,y)dx$$

Условные законы распределения вероятностей для двумерной случайной величины.

Рассмотрим двумерную случайную величину ($X;Y$). Пусть случайная величина Y приняла значение $Y= y_i$. Тогда функция распределения случайной величины X называется условной функцией распределения случайной величины X при условии $Y= y_i$ и обозначается $F_X(x/Y = y_i)$.

Для условной функции распределения справедливы те же свойства, что и для безусловной.

А). *Дискретная двумерная СВ.*

	x_1	x_2	…	x_n	
y_1	p_{11}	p_{12}	…	p_{1n}	$p(y_1)$
y_2	p_{21}	p_{22}	…	p_{2n}	$p(y_2)$
…	…	…	…	…	…
y_m	p_{m1}	p_{m2}	…	p_{mn}	$p(y_m)$

	$p(x_1)$	$p(x_2)$	...	$p(x_n)$	

Определение. Условным распределением случайной величины *X* при *Y*= y_i называется совокупность условных вероятностей $p(x_1/y_i), p(x_2/y_i)...p(x_n/y_i)$, вычисленных в предположении, что событие $\{\omega : Y = y_i\}$ уже наступило.

$$p(x_j/y_i) = \frac{p(x_j, y_i)}{p(y_i)} = \frac{p_{ji}}{p(y_i)}$$

Аналогично находят и условные законы распределения случайной величины *Y*.

$$p(y_i/x_j) = \frac{p(x_j, y_i)}{p(x_j)} = \frac{p_{ji}}{p(x_j)}$$

Замечание: $p(x_1/y_i) + p(x_2/y_i) + ... + p(x_n/y_i) = 1$

Пример 2. Найти условный закон распределения составляющей *X* двумерной случайной величины (*X;Y*), заданной таблицей распределения, при *Y=4* .

Y \ *X*	*1*	*3*	*5*
2	*0,1*	*0,3*	*0,2*
4	*0,06*	*0,18*	*0,16*

Решение.

Из примера 1:

	1	*3*	*5*	
2	*0,1*	*0,3*	*0,2*	$p(y_1) = 0{,}6$
4	*0,06*	*0,18*	*0,16*	$p(y_2) = 0{,}4$
	$p(x_1) = 0{,}16$	$p(x_2) = 0{,}48$	$p(x_3) = 0{,}36$	

X	*1*	*3*	*5*
$p(x_j/Y=4)$	*3/20*	*9/20*	*8/20*

Б). *Непрерывная двумерная СВ.*

Определение. Условной плотностью распределения вероятностей случайной величины *X* при данном значении *Y*= *y* называется отношение плотности распределения двумерной случайной величины (*X;Y*) к плотности распределения случайной величины *Y*

$$\rho_X(x/y)=\frac{\rho_{XY}(x,y)}{\rho_Y(y)}=\frac{\rho_{XY}(x,y)}{\int\limits_{-\infty}^{\infty}\rho_{XY}(x,y)dx}$$

Аналогично

$$\rho_Y(y/x)=\frac{\rho_{XY}(x,y)}{\rho_X(x)}=\frac{\rho_{XY}(x,y)}{\int\limits_{-\infty}^{\infty}\rho_{XY}(x,y)dy}$$

В). *Условное математическое ожидание.*

Дискретная СВ:

$$M(Y/X=x)=\sum_{i=1}^{m}y_i p(y_i/x)$$

Непрерывная СВ:

$$M(Y/X=x)=\int\limits_{-\infty}^{\infty}y\rho(y/x)dy$$

Пример 3. Найти математические ожидания составляющих *X* и *Y* двумерной случайной величины (*X;Y*), заданной таблицей распределения, и условное математическое ожидание *X* при *Y=4* .

Y \ X	*1*	*3*	*5*
2	*0,1*	*0,3*	*0,2*
4	*0,06*	*0,18*	*0,16*

Решение.

Из примеров 1,2:

X	*1*	*3*	*5*
p	*0,16*	*0,48*	*0,36*

MX=0,16+1,44+1,8=3,4

Y	*2*	*4*
p	*0,6*	*0,4*

MY =1,2+1,6=2,8

X	*1*	*3*	*5*
$p(x_j/Y=4)$	*3/20*	*9/20*	*8/20*

M(X/ Y =4)=0,15+1,35+2=3,5

Г). *Независимость случайных величин.*

Две случайные величины называются независимыми, если закон распределения одной из них не зависит от того, какие возможные значения приняла другая величина. Следовательно, условные распределения независимых величин равны их безусловным распределениям.

Теорема. Для того, чтобы случайные величины X и Y были независимыми, необходимо и достаточно, чтобы

$$F_{XY}(x,y)=F_X(x)F_Y(y)$$

Следствие. $\rho_{XY}(x,y)=\rho_X(x)\rho_Y(y)$

Числовые характеристики системы двух случайных величин.

А). MX, MY, DX, DY

Пример 4. Из примеров 1,2,3 для двумерной случайной величины ($X;Y$), заданной таблицей распределения

Y \ X	1	3	5
2	0,1	0,3	0,2
4	0,06	0,18	0,16

X	1	3	5
p	0,16	0,48	0,36

MX=0,16+1,44+1,8=3,4

DX=1*0,16+9*0,48+25*0,36-3,4^2=1,92

Y	2	4
p	0,6	0,4

MY =1,2+1,6=2,8

DY=4*0,6+16*0,4-2,8^2=0,96

Б). *Корреляционный момент* μ_{XY}

Определение. Ковариацией или корреляционным моментом μ_{XY} двух случайных величин X и Y называют математическое ожидание произведения отклонений этих величин.

$$\operatorname{cov}(X,Y)=\mu_{XY}=M\left[(X-MX)(Y-MY)\right]=M\left[XY\right]-M[X]M[Y]$$

Случайные величины X и Y называют коррелированными, если μ_{XY} отличен от нуля. Если $\mu_{XY} = 0$, то X и Y некоррелированные величины.

Теорема. Корреляционный момент μ_{XY} двух независимых случайных величин равен нулю.

Замечание: обратное утверждение, вообще говоря, неверно.

Некоторые полезные формулы.

1. $M[XY] = M[X]M[Y] + \mu_{XY}$
2. $D[X+Y] = D[X] + D[Y] + 2\mu_{XY}$

В). *Коэффициент корреляции (ковариации)*

Определение. Коэффициентом корреляции двух случайных величин X и Y называют отношение их корреляционного момента к произведению среднеквадратичных отклонений.

$$r_{XY} = \frac{\mu_{XY}}{\sigma_X \sigma_Y}$$

Свойства коэффициента корреляции:

1^0. $0 \le |r_{XY}| \le 1$

2^0. Для независимых случайных величин $r_{XY} = 0$.

3^0. Если $|r_{XY}| = 1$, то между X и Y существует функциональная линейная зависимость

Коэффициент корреляции отражает наличие *линейной* связи между X и Y, чем ближе этот коэффициент по модулю к 1, тем сильнее *линейная* связь. Отрицательные (положительные) значения r_{XY}, указывают на то, что при возрастании значений одной случайной величины значения другой *имеют тенденцию* к убыванию (возрастанию).

Пример 4. Из примеров 1,2,3 для двумерной случайной величины ($X;Y$), заданной таблицей распределения

Y \ *X*	*1*	*3*	*5*
2	*0,1*	*0,3*	*0,2*
4	*0,06*	*0,18*	*0,16*

MX=0,16+1,44+1,8=3,4 DX=1*0,16+9*0,48+25*0,36-3,4^2=1,92

MY =1,2+1,6=2,8 DY=4*0,6+16*0,4-2,8^2=0,96

$$\mu_{XY} = M[XY] - M[X]M[Y] =$$
$$= 1*2*0{,}1 + 3*2*0{,}3 + 5*2*0{,}2 + 1*4*0{,}06 + 3*4*0{,}18 + 5*4*0{,}16 - 3{,}4*2{,}8 = 0{,}08$$

$$r_{XY}=\frac{\mu_{XY}}{\sigma_X\sigma_Y}=\frac{0{,}08}{\sqrt{1{,}92*0{,}96}}=0{,}059$$

Нормальный закон распределения на плоскости

$$\rho_{XY}(x,y)=\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-r^2}}\exp\left\{-\frac{1}{2}Q(x,y)\right\}$$

$$Q(x_1,x_2)=\frac{1}{1-r^2}\left[\frac{(x-a_X)^2}{\sigma_X^2}-2r\frac{x-a_X}{\sigma_X}\frac{y-a_Y}{\sigma_Y}+\frac{(y-a_Y)^2}{\sigma_Y^2}\right]$$

$r=r_{XY}$, $MX=a_X$, $MY=a_Y$, $DX=\sigma_X^2$, $DY=\sigma_Y^2$

Линейная регрессия

Коэффициент корреляции отражает *наличие* связи между X и Y, но не определяет вида зависимости. Можно показать, что зависимость поведения «в среднем» случайной величины Y от значения случайной величины X характеризуется функцией $\varphi(x)=M[Y/X=x]$.

Определение. Функция $\varphi(x)=M[Y/X=x]$ называется *функцией регрессии* или просто *регрессией* случайной величины Y на случайную величину X, а ее график – *линией регрессии*.

Линейная регрессия Y на X имеет вид

$$y=\phi(x)=Ax+B=MY+r_{XY}\frac{\sigma_Y}{\sigma_X}(x-MX)$$

Линейная регрессия X на Y имеет вид

$$x=\psi(y)=Cy+D=MX+r_{XY}\frac{\sigma_X}{\sigma_Y}(y-MY)$$

При построении этих прямых на плоскости они пересекаются в точке ($MX; MY$), которая называется центром совместного распределения.

$$A=r_{XY}\frac{\sigma_Y}{\sigma_X}\qquad B=MY-r_{XY}\frac{\sigma_Y}{\sigma_X}MX,$$

$$C=r_{XY}\frac{\sigma_X}{\sigma_Y}\qquad D=MX-r_{XY}\frac{\sigma_X}{\sigma_Y}MY.$$

Коэффициенты A и B (C и D) определяются *методом наименьших квадратов* из условия минимума функции $\Phi(A,B)=M[(Y-(AX+B))^2]$

$\Phi(A,B)=M[(Y-MY+MY-B-A(X-MX)-AMX)^2]$ = *(выделение желтым: добавили и отняли)*=

$=M[(Y-MY)^2]+\underline{2M[(Y-MY)(MY-AMX-B-A(X-MX))]}+\underline{M[(MY-AMX-B-A(X-MX))^2]}=$

$=\sigma_Y^2+\underline{2M[(Y-MY)(MY-AMX-B)]}\ \underline{-2MA[(Y-MY)(X-MX)]}+$

$+\underline{M[(MY-AMX-B)^2-2A[(X-MX)(MY-AMX-B)+A^2(X-MX)^2]}=$

$=\sigma_Y^2$-$2A\,\mu_{XY}$ + $M[\,(MY\text{-}AMX\text{-}B)^2]$-$2AM[(X\text{-}MX)(\,MY\text{-}AMX\text{-}B)]$ +$A^2M(X\text{-}MX)^2]$=

(выделение красным: равно нулю)$=\sigma_Y^2$-$2A\,r_{XY}\,\sigma_Y\,\sigma_X$+ $A^2\,\sigma_X^2$+$(MY\text{-}AMX\text{-}B)^2$

$\Phi\,(A,\,B) = \sigma_Y^2\text{-}2A\,r_{XY}\,\sigma_Y\,\sigma_X + A^2\,\sigma_X^2 + (\,MY\text{-}AMX\text{-}B)^2$

$$\begin{cases}\dfrac{\partial\Phi}{\partial A}=0\\ \dfrac{\partial\Phi}{\partial B}=0\end{cases}\quad\to\quad\begin{cases}2A\sigma_X^2-2r_{XY}\sigma_X\sigma_Y=0\\ -2(MY-B-AMX)=0\end{cases}\quad\to\quad\begin{cases}A=r_{XY}\dfrac{\sigma_Y}{\sigma_X}\\ B=MY-r_{XY}\dfrac{\sigma_Y}{\sigma_X}MX\end{cases}$$

$\Phi_{min}\,(A,\,B) = \sigma_Y^2(1\text{-}\,r_{XY}^{\ 2})$ – остаточная дисперсия

Лекция №9.

Задачи и основные понятия математической статистики.

Первичная обработка выборки. Точечные и интервальные оценки.

§32. Понятие генеральной совокупности. Выборка из генеральной совокупности.

Основными задачами математической статистики являются:

1. Разработка методов получения (сбора) информации.
2. Построение методов обработки полученной информации.

Определение. Под *генеральной совокупностью* понимается случайный количественный признак X, присущий рассматриваемому явлению или каждому элементу исследуемого множества.

Определение. *Выборкой* объема n из генеральной совокупности *X* с функцией распределения $F_X(x)$ называется последовательность $x^{(1)}, x^{(2)}, \ldots x^{(n)}$ наблюдаемых значений случайной величины X, соответствующим n независимым повторениям эксперимента.

Выборка должна быть *репрезентативной*, т.е. наиболее полно и адекватно представлять свойства исследуемого объекта.

1. Выборка должна быть достаточно большого объема
2. Выборка должна представлять все группы исследуемого объекта.
3. Выборка должна быть случайной.

Пример. Дана выборка объема n =15.

5, 3, 7, 10, 5, 5, 2, 10, 7, 2, 7, 7, 4, 2, 4.

$x^{(1)}, x^{(2)}, x^{(3)}, x^{(4)}, x^{(5)}, x^{(6)}, x^{(7)}, x^{(8)} \ldots\ldots x^{(15)}$

$x^{(i)}$- варианта.

Определение. Наблюдаемые значения $x^{(i)}$, записанные в порядке возрастания называются *вариационным рядом*.

2, 2, 2, 3, 4, 4, 5, 5, 5, 7, 7, 7, 7, 10, 10.

Определение. *Статистический ряд* - таблица, первая строка которой - перечень вариант, вторая строка - перечень соответствующих им частот или относительных частот.

$x^{(i)}$	2	3	4	5	7	10
n_i	3	1	2	3	4	2

$\Sigma n_i = n$

n_i- частота появления значения $x^{(i)}$ в выборке.

Статистический ряд относительных частот:

$x^{(i)}$	2	3	4	5	7	10
n_i/n	3/15	1/15	2/15	3/15	4/15	2/15

n_i/n - относительная частота (*частость* w_i)

$\Sigma\, w_i = \Sigma n_i/n = 1$

Определение. *Размах выборки* W - разность между максимальным и минимальным значением элементов выборки.

W=10-2=8

Для *большого объема данных или в случае непрерывного признака X* используют группированные выборки, для которых строят интервальный статистический ряд.

h=W/k, h- шаг, W- размах выборки, k- число интервалов разбиения (для выборок большого объема можно, например, выбрать $k \approx \log_2 n + 1$).

W=8, k =4, h=8/4=2

$x^{(i)}$	$2 \le x_i < 4$	$4 \le x_i < 6$	$6 \le x_i < 8$	$8 \le x_i \le 10$
n_i	4	5	4	2

Интервальный ряд относительных частот

$x^{(i)}$	$2 \le x_i < 4$	$4 \le x_i < 6$	$6 \le x_i < 8$	$8 \le x_i \le 10$
w_i	4/15	5/15	4/15	2/15

Если в статистическом распределении вместо частот (относительных частот) указать накопленные частоты (относительные накопленные частоты), то такой ряд называют *кумулятивным*.

Накопленной частотой называется число значений признака X, меньших заданного значения x: $H(x) = n(X < x)$, то есть, число вариант x_j в выборке, отвечающих условию $x_j < x$.

Дискретный кумулятивный ряд:

$x^{(i)}$	2	3	4	5	7	10	>10
$H(x^{(i)})$	0	3	4	6	9	13	15

Интервальный кумулятивный ряд:

$x^{(i)}$	$x_i < 2$	$2 \le x_i < 4$	$4 \le x_i < 6$	$6 \le x_i < 8$	$8 \le x_i \le 10$
$H(x^{(i)})$	0	4	9	13	15

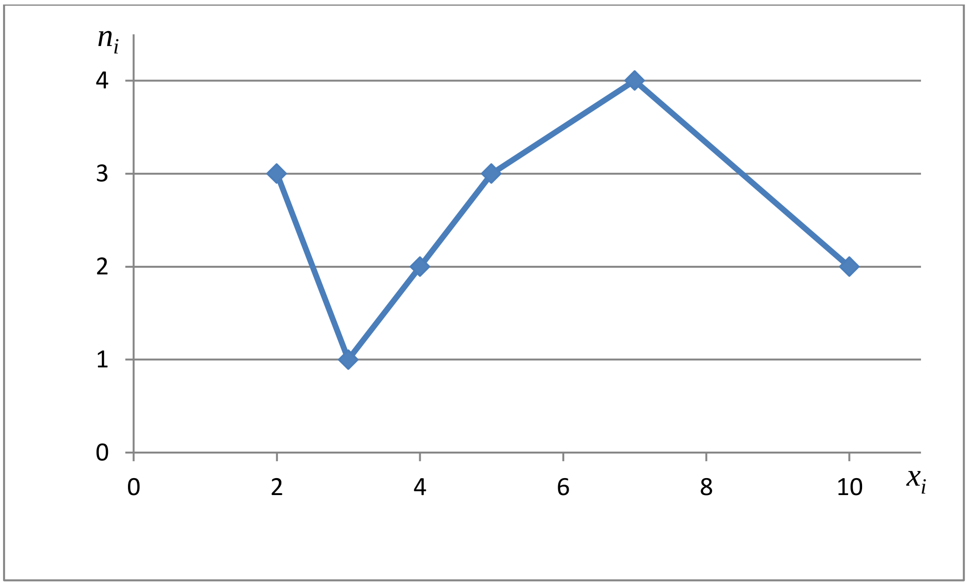
Аналогично строятся кумулятивные ряды относительных частот.

§33. Графическое представление выборки.

1. Полигон частот (для малых выборок).

Полигон частот - ломаная, отрезки которой соединяют точки с координатами $x^{(i)}$, n_i.

$x^{(i)}$	2	3	4	5	7	10
n_i	3	1	2	3	4	2

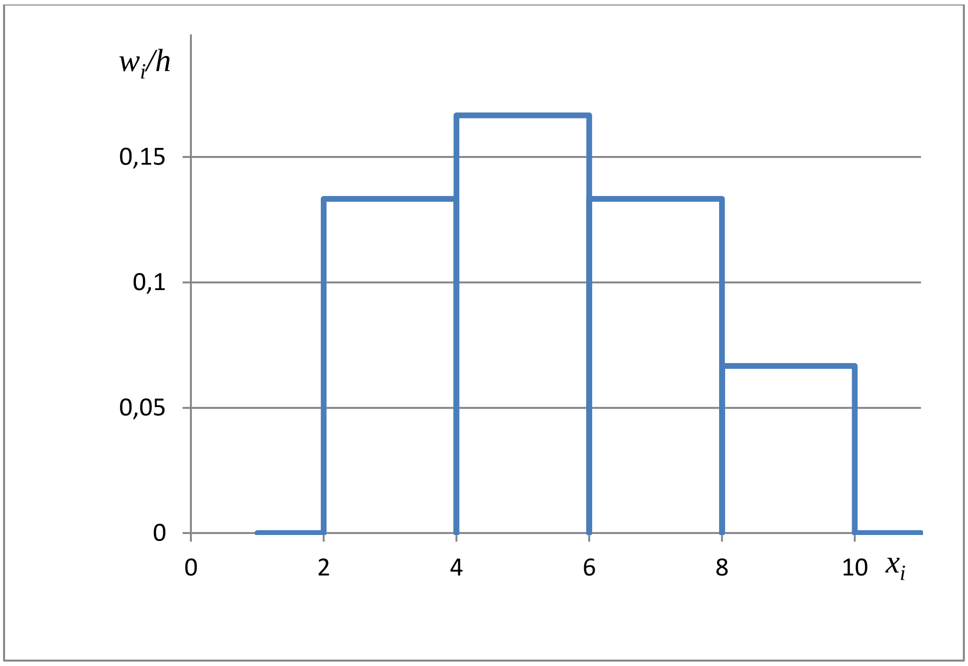


2. Гистограмма частот (для группированных выборок).

Гистограмма частот - ступенчатая фигура, состоящая из прямоугольников, основанием которых служат частичные интервалы, длиной h, а высоты равны отношению w_i / h.

$x^{(i)}$	$2 \le x_i < 4$	$4 \le x_i < 6$	$6 \le x_i < 8$	$8 \le x_i \le 10$
w_i	4/15	5/15	4/15	2/15

Гистограмма относительных частот



Площадь гистограммы относительных частот равна единице; она дает представление о возможном распределении (плотности) непрерывной генеральной совокупности.

3. Эмпирическая функция распределения.

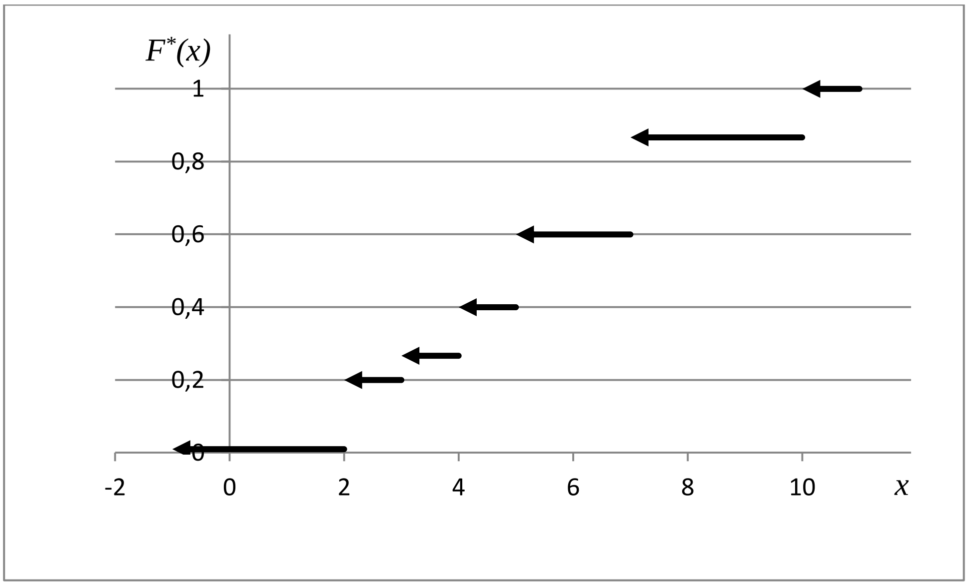
Эмпирическая функция распределения $F^*_X(x)=\Sigma n_i/n$ определяет для любого x долю вариант $x^{(i)}$ в выборке, для которых справедливо условие $x^{(i)}<x$.

- $F^*_X(x)\leq 1$
- $F^*_X(x)$- неубывающая функция.
- $F^*_X(x)=0$ при $x\leq x^{(i)}_{\min}$
- $F^*_X(x)=1$ при $x> x^{(i)}_{\max}$

$x^{(i)}$	2	3	4	5	7	10
n_i/n	3/15	1/15	2/15	3/15	4/15	2/15

x	$x<2$	$2\leq x<3$	$3\leq x<4$	$4\leq x<5$	$5\leq x<7$	$7\leq x<10$	$x>10$
$F^*_X(x)$	0	3/15	4/15	6/15	9/15	13/15	1

$F^*_X(x)$ - накопленная частость.



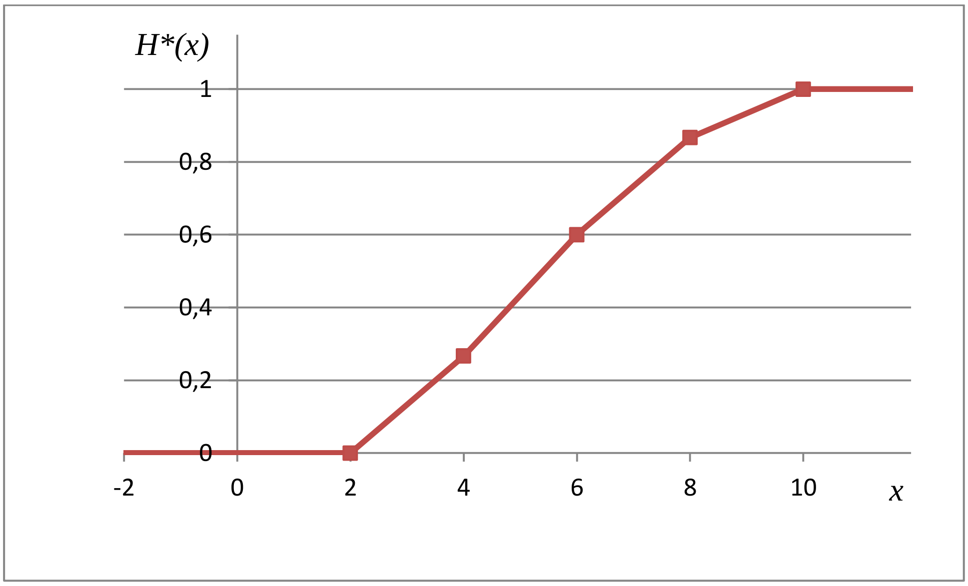
3. Кумулята (для непрерывного признака).

$x^{(i)}$	$x_i<2$	$2\le x_i<4$	$4\le x_i<6$	$6\le x_i<8$	$8\le x_i\le 10$	>10
$H(x^{(i)})$	0	4	9	13	15	15

а) наносим точки с координатами $(x\,;H^*(x))$:

x	2	4	6	8	10
$H^*(x)$	0	4/15	9/15	13/15	1

б) соединяем их отрезками



Часто кумуляту обозначают так же, как и эмпирическую функцию распределения, $F^*_X(x)$.

§34. Точечные оценки

Важной задачей математической статистики является задача оценивания (приближенного определения) по выборочным данным параметров закона распределения признака X генеральной совокупности. Статистические оценки могут быть точечными и интервальными. Точечной оценкой называют оценку, которая определяется одним числом.

Пусть $x^{(1)}, x^{(2)}, \ldots x^{(n)}$ - выборка объема n из генеральной совокупности с функцией распределения $F(x^{(i)},\theta)$ с неизвестным параметром θ. Произвольная функция элементов выборки $X^{(1)}, X^{(2)}, \ldots X^{(n)}$ называется статистикой.

Значение статистики $\theta_n(x^{(1)}, x^{(2)}, \dots x^{(n)})$ является *точечной оценкой параметра* θ, если оно приблизительно равно оцениваемому параметру и может заменить его с достаточной степенью точности в статистических расчетах.

Оценка θ_n называется *несмещенной*, если математическое ожидание оценки равно оцениваемому параметру, т.е. $M[\theta_n] = \theta$

Оценка θ_n называется *состоятельной*, если для любого $\varepsilon > 0$ при $n \to \infty$ $P(|\theta_n - \theta| < \varepsilon) \to 1$

Оценка θ_n называется *эффективной*, если при фиксированном n она имеет наименьшую дисперсию.

Точечные оценки числовых характеристик распределения (метод моментов).

Пусть $X: x_1, x_2, \dots, x_N$ - генеральная совокупность объема N с функцией распределения

$F_X(x)$, $MX=a$, $DX = \sigma^2$

$x^{(1)}, x^{(2)}, \dots x^{(n)}$ - выборка объема n.

Генеральной средней $x_Г$ называется среднее арифметическое значений элементов генеральной совокупности.

$$x_Г = \frac{1}{N}\sum x_i = a$$

Выборочной средней x_B называется среднее арифметическое элементов выборки.

$$x_B = \frac{1}{n}\sum x^{(i)}$$

Теорема: x_B является несмещенной оценкой $x_Г$.

Доказательство: $M[X_B] = M[\sum_{i=1}^{n}(\frac{1}{n}X^{(i)})] = \frac{1}{n}M[\sum_{i=1}^{n}X^{(i)}] = \frac{1}{n}na = a$

Замечание. $D[X_B] = D[\sum_{i=1}^{n}(\frac{1}{n}X^{(i)})] = \frac{1}{n^2}D[\sum_{i=1}^{n}X^{(i)}] = \frac{1}{n^2}n\sigma^2 = \frac{\sigma}{n}$

Например: если $X \to N(a, \sigma)$, то $X_B \to N(a, \frac{\sigma}{\sqrt{n}})$

Генеральной дисперсией $D_Г$ называют величину $D_Г = \frac{1}{N}\sum(x_i - a)^2$, $D_Г = \sigma^2$

Выборочной дисперсией D_B называют величину $D_B = \frac{1}{n}\sum(x^{(i)} - x_B)^2$

Теорема: выборочная дисперсия D_B является смещенной оценкой генеральной дисперсии.

$s^2 = \frac{1}{n-1}\sum(x^{(i)} - x_B)^2$ - *исправленная выборочная дисперсия*, или несмещенная оценка дисперсии генеральной совокупности.

Если математическое ожидание генеральной совокупности известно, то в качестве несмещенной оценки генеральной совокупности используется

$$s_0^2 = \frac{1}{n}\sum (x^{(i)} - a)^2$$

Заключение (основные формулы):

1) Оценка математического ожидания генеральной совокупности

$$x_B = \frac{1}{n}\sum x^{(i)}$$

2) Точечные оценки дисперсии генеральной совокупности

$D_B = \frac{1}{n}\sum (x^{(i)} - x_B)^2$ - смещенная оценка дисперсии генеральной совокупности

$s^2 = \frac{1}{n-1}\sum (x^{(i)} - x_B)^2$ - несмещенная оценка дисперсии генеральной совокупности

$s_0^2 = \frac{1}{n}\sum (x^{(i)} - a)^2$ - несмещенная оценка дисперсии генеральной совокупности, если известно ее математическое ожидание.

Замечания.

А). На практике в качестве характеристик среднего значения генеральной совокупности также рассматривают моду и медиану распределения. По выборке медиану оценивают по формулам:

$x_{\text{ме}} = \dfrac{x^{(j)} + x^{(j+1)}}{2}$, если $n = 2j$ – четное; $x_{\text{ме}} = x^{(j+1)}$, если $n = 2j+1$ – нечетное.

Мода – наиболее часто встречающееся в выборке значение признака X.

Б). В качестве характеристик вариации рассматривают также выборочное среднеквадратичное отклонение $\sigma_B = \sqrt{D_B}$ и *коэффициент вариации* $v = \dfrac{\sigma_{\text{в}}}{\bar{x}_{\text{в}}} \cdot 100\,\%$.

Коэффициент вариации применяют для сравнения вариации признаков сильно отличающихся по величине, или имеющих разные единицы измерения (разные наименования).

<u>Метод наибольшего (максимального) правдоподобия.</u>

Метод наибольшего правдоподобия – это метод оценки неизвестных параметров распределения, в основе которого – поиск максимального значения функции правдоподобия.

Достоинства:

1. Может использоваться в случае, когда теоретические моменты распределения отсутствуют.

2. Оценки в основном состоятельны и эффективны.

3. Оценки распределены асимптотически нормально.

4. Наиболее полно используются данные о выборке (особенно полезны в случае малых выборок).

Недостатки:

1. Оценки могут быть смещенными.

2. Сложность вычислений.

3. Не всегда совпадают с оценками по методу моментов.

1. Дискретная случайная величина X

Пусть $x_1, x_2, \ldots x_n$ -выборка объема n из генеральной совокупности *X с известной* функцией распределения $F(x,\theta)$, θ – неизвестный параметр, который нужно определить по выборке, $p(x_i, \theta)$ – вероятность того, что в результате испытания X примет значение x_i.

Определение. Функцией правдоподобия дискретной случайной величины называется функция аргумента θ

$$L(x_1, x_2, \ldots x_n, \theta) = p(x_1, \theta)\, p(x_2, \theta) \ldots p(x_n, \theta)$$

В качестве точечной оценки θ принимается $\theta^*(x_1, x_2, \ldots x_n)$,при которой L принимает для данной выборки наибольшее значение.

θ^*-оценка наибольшего правдоподобия.

Определение. Логарифмической функцией правдоподобия называют функцию $\ln(L)$.

Алгоритм построения θ^*:

А) осуществляем выборку $x_1, x_2, \ldots x_n$, вычисляем вероятности $p(x_1, \theta), p(x_2, \theta), \ldots, p(x_n, \theta)$ и строим L и $\ln(L)$.

Б) $$\left.\frac{d(\ln L)}{d\theta}\right|_{\theta = \theta^*} = 0$$

В) $$\left.\frac{d^2(\ln L)}{d\theta^2}\right|_{\theta = \theta^*} < 0$$

Пример. Найти методом максимального правдоподобия оценку параметра λ распределения Пуассона.

Решение:

А) θ=λ

$p(x_i, \lambda) = \dfrac{\lambda^{x_i}}{x_i!} e^{-\lambda}$, где n – число испытаний? x_i-число появлений события в i – ом опыте

$$L = p(x_1;\lambda)\,p(x_2;\lambda)...p(x_n;\lambda) = \frac{\lambda^{x_1}}{x_1!}e^{-\lambda} \cdot \frac{\lambda^{x_2}}{x_2!}e^{-\lambda} ... \cdot \frac{\lambda^{x_n}}{x_n!}e^{-\lambda}$$

$$\ln L = \ln\left(\frac{\lambda^{\sum x_i}}{x_1!x_2!...x_n!}e^{-n\lambda}\right) = \sum x_i \ln\lambda - \ln(x_1!x_2!...x_n!) - n\lambda$$

Б) $\dfrac{d(\ln L)}{d\lambda}\Big|_{\lambda=\lambda^*} = 0 \quad \Rightarrow \quad \dfrac{\sum x_i}{\lambda^*} - n = 0 \quad \Rightarrow \quad \lambda^* = \dfrac{\sum x_i}{n} = x_B$

В) $\dfrac{d^2(\ln L)}{d\lambda^2}\Big|_{\lambda=\lambda^*} = -\dfrac{\sum x_i}{(\lambda^*)^2} < 0$

2. Непрерывная случайная величина X

Пусть $x_1, x_2, ... x_n$ -выборка объема n из генеральной совокупности *X с известной* функцией распределения *F(x,θ), θ* – неизвестный параметр, который нужно определить по выборке, $\rho(x_i,\theta)$ – плотность распределения вероятностей X в точке x_i.

Определение. Функцией правдоподобия непрерывной случайной величины X называют

$$L(x_1, x_2, ...x_n, \theta) = \rho(x_1,\theta)\rho(x_2,\theta)...\rho(x_n,\theta)$$

Пример. Найти методом максимального правдоподобия оценку параметра λ показательного распределения:

А) θ=λ

$$\rho(x,\lambda) = \begin{cases} \lambda e^{-\lambda x}, & x \in (0;\infty) \\ 0, & x \notin (0;\infty) \end{cases}$$

$$L = \lambda^n e^{-\lambda \sum x_i}$$

$$\ln L = n\ln\lambda - \lambda\sum x_i$$

Б) $\dfrac{d(\ln L)}{d\lambda}\Big|_{\lambda=\lambda^*} = \dfrac{n}{\lambda^*} - \sum x_i = 0 \quad \Rightarrow \quad \lambda^* = \dfrac{1}{x_B}$

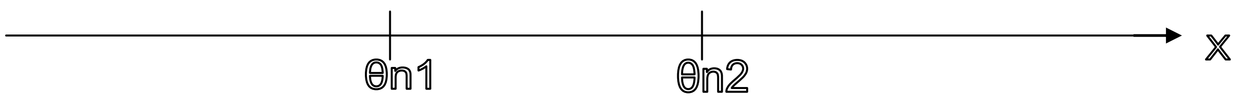
В) $\dfrac{d^2(\ln L)}{d\lambda^2}\Big|_{\lambda=\lambda^*} = -\dfrac{n}{(\lambda^*)^2} < 0$

§35. Доверительные интервалы для параметров генеральной совокупности.

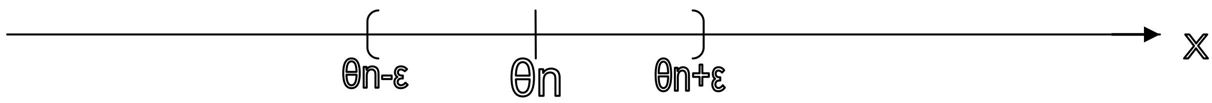
Пусть X: $x_1, x_2, ..., x_N$ - генеральная совокупность с функцией распределения *F(x,θ)* , зависящей от параметра θ.

X₁: $x_1^{(1)}, x_1^{(2)}, ..., x_1^{(n)}$ θ_{n1} – оценка θ, полученная по выборке X_1

X₂: $x_2^{(1)}, x_2^{(2)}, ..., x_2^{(n)}$ θ_{n2}– оценка θ, полученная по выборке X_2



Задача - по выборке объема n построить интервал, которому с вероятностью (1-α) принадлежит истинное значение параметра θ.



$P(|\theta - \theta_n| < \varepsilon) = 1 - \alpha$, $\varepsilon > 0$

α – уровень значимости, *p=1- α* – доверительная вероятность, $(\theta_n - \varepsilon; \theta_n + \varepsilon)$ - доверительный интервал.

Определение. Доверительным интервалом для параметра *θ* генеральной совокупности *X* с функцией распределения *F(x,θ)* называется интервал $(\theta_n - \varepsilon; \theta_n + \varepsilon)$, в который истинное значение параметра *θ* попадает с вероятностью *p=1- α*.

Пример 1. *X* имеет <u>нормальное распределение</u> с параметрами (m,σ). Найти доверительный интервал для математического ожидания m по результатам n наблюдений при условии, что σ^2 известна, а доверительная вероятность равна *p=1- α*.

$\theta \to m$ $\qquad\qquad$ $\theta_n \to x_B$ $\qquad\qquad$ $x_B = \dfrac{1}{n}\sum_{i=1}^{n} x^{(i)}$

$$P(|m - x_B| < \varepsilon) = 1 - \alpha$$
$$X \to N(m, \sigma)$$
$$X^{(i)} \to N(m, \sigma)$$
$$X_B \to N(m, \frac{\sigma}{\sqrt{n}})$$
$$Y = \frac{X_B - m}{\frac{\sigma}{\sqrt{n}}} \to N(0,1)$$

$$P(|m - x_B| < \varepsilon) = P(-\varepsilon < m - x_B < \varepsilon) = P(-\frac{\sqrt{n}}{\sigma}\varepsilon < \frac{m - x_B}{\sigma/\sqrt{n}} < \frac{\sqrt{n}}{\sigma}\varepsilon) = 1 - \alpha$$

$$P(\alpha < Y < \beta) = F(\beta) - F(\alpha) = \Phi_0(\frac{\beta - m_Y}{\sigma_Y}) - \Phi_0(\frac{\alpha - m_Y}{\sigma_Y}) = \Phi_0(\beta) - \Phi_0(\alpha)$$

$$P(-\frac{\varepsilon\sqrt{n}}{\sigma} < Y < \frac{\varepsilon\sqrt{n}}{\sigma}) = \Phi_0(\frac{\varepsilon\sqrt{n}}{\sigma}) - \Phi_0(-\frac{\varepsilon\sqrt{n}}{\sigma}) = 2\Phi_0(\frac{\varepsilon\sqrt{n}}{\sigma}) = 2\Phi(\frac{\varepsilon\sqrt{n}}{\sigma}) - 1 = 1 - \alpha$$
$$\Phi(\frac{\varepsilon\sqrt{n}}{\sigma}) = 1 - \frac{\alpha}{2}$$

Определение. Квантилью порядка *p* непрерывного теоретического распределения случайной величины *X* называется действительное число x_p, удовлетворяющее уравнению:

$P(X < x_p) = p$ <u>ИЛИ</u> $F_X(x_p) = p$

($F_X(x)$ - функция распределения *X*)

Введем обозначения t=1- α/2, u_t - квантиль порядка *t* стандартного нормального распределения. Тогда

$$\frac{\varepsilon\sqrt{n}}{\sigma} = u_t$$
$$\varepsilon = \frac{\sigma}{\sqrt{n}}u_t$$
$$x_B - \varepsilon < m < x_B + \varepsilon$$
$$x_B - \frac{\sigma}{\sqrt{n}}u_t < m < x_B + \frac{\sigma}{\sqrt{n}}u_t$$

Пример 2. *X* имеет нормальное распределение с параметрами (m,σ). Найти доверительный интервал для σ по результатам n наблюдений при условии, что m известно, а доверительная вероятность равна 1-α.

$$P(\sigma_1 < \sigma < \sigma_2) = 1 - \alpha$$
$$\sigma_1 - ?; \sigma_2 - ?$$

$$s_0^2 = \frac{1}{n}\sum_{i=1}^{n}(x^{(i)} - m)^2 = \frac{1}{n}\sum_{i=1}^{n}(\frac{x^{(i)} - m}{\sigma})^2\sigma^2 = \frac{\sigma^2}{n}\sum_{i=1}^{n}(\frac{x^{(i)} - m}{\sigma})^2$$

$$X^{(i)} \to N(m,\sigma)$$

$$X^{(i)} - m \to N(0,\sigma)$$

$$\frac{X^{(i)} - m}{\sigma} \to N(0,1)$$

$\chi^2(n) = Z_1^2 + Z_2^2 + ... + Z_n^2$, где $Z_i \to N(0,1)$

$$\Rightarrow \frac{s_0^2 n}{\sigma^2} \to \chi^2(n)$$

$$P(\varepsilon_1 < \chi^2(n) < \varepsilon_2) = 1 - \alpha$$

$$P(\chi^2(n) < \varepsilon_1) = \frac{\alpha}{2} \qquad P(\chi^2(n) > \varepsilon_2) = \frac{\alpha}{2}$$

$$\varepsilon_1 = \chi^2_{\frac{\alpha}{2}}(n) \qquad \varepsilon_2 = \chi^2_{1-\frac{\alpha}{2}}(n)$$

$$P(\chi^2_{\frac{\alpha}{2}}(n) < \chi^2(n) < \chi^2_{1-\frac{\alpha}{2}}(n)) = 1 - \alpha$$

$$\chi^2(n) = \frac{s_0^2 n}{\sigma^2}$$

$$P(\frac{s_0^2 n}{\chi^2_{1-\frac{\alpha}{2}}(n)} < \sigma^2 < \frac{s_0^2 n}{\chi^2_{\frac{\alpha}{2}}(n)}) = 1 - \alpha$$

$$\sqrt{\frac{s_0^2 n}{\chi^2_{1-\frac{\alpha}{2}}(n)}} < \sigma < \sqrt{\frac{s_0^2 n}{\chi^2_{\frac{\alpha}{2}}(n)}}$$

Примечание.

1. Доверительный интервал для математического ожидания в случае, если дисперсия генеральной совокупности *неизвестна*:

$$x_B - \frac{s}{\sqrt{n}} t_{1-\frac{\alpha}{2}}(n-1) < m < x_B + \frac{s}{\sqrt{n}} t_{1-\frac{\alpha}{2}}(n-1)$$

2. Доверительный интервал для дисперсии при *неизвестном* математическом ожидании:

$$\frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}$$

§36. Статистические гипотезы.

Определение. Статистической гипотезой называется гипотеза о виде неизвестного распределения или о параметрах известного распределения. Выдвинутую гипотезу называется основной (нулевой) и обозначается H_0 .Противоречащую ей называется конкурирующей (альтернативной) и обозначается H_1.

Гипотеза называется простой, если она содержит только одно предположение. Сложная гипотеза состоит из простых.

α -ошибка первого рода (вероятность отвергнуть правильную гипотезу).

β -ошибка второго рода (вероятность принять неверную нулевую гипотезу).

При увеличении α возрастает и β . И наоборот. Единственный способ одновременного уменьшения α и β - увеличение количества испытаний. Величина (1- β) носит название мощности критерия.

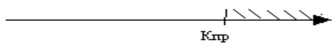
§37. Статистический критерий проверки основной гипотезы H_0

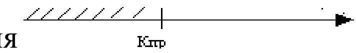
Определение. Для проверки основной гипотезы H_0 используют специально подобранную случайную величину, точное или приближенное распределение которое известно. Эта случайная величина называется статистическим критерием или просто критерием К.

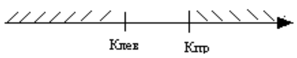
К =К ($x^{(1)}, x^{(2)}, \ldots\ x^{(n)}$)

Определение. Областью принятия гипотезы H_0 называется совокупность значений критерия, при которых гипотезу принимают. Критической областью называется совокупность значений критерия, при которых нулевую гипотезу отвергают.

Точки, отделяющие область принятия гипотезы от критической области, называется критическими точками. Критическая область может быть:

1. правосторонняя 

2. левосторонняя 

3. двусторонняя 

Выбор одного из этих случаев определяется видом конкурирующей гипотезы.

1. $P(K > K_{пр}) = \alpha$
2. $P(K < K_{лев}) = \alpha$
3. $P(K > K_{пр}) = \dfrac{\alpha}{2};\quad P(K < K_{лев}) = \dfrac{\alpha}{2}$

Основные шаги при проверке статистических гипотез:

1) выдвигаем H_0
2) выдвигаем H_1
3) задаем α - уровень значимости
4) строим статистический критерий
5) строим критическую область

6) считаем наблюдаемое значение критерия и сравниваем с критическими точками

7) если наблюдаемое значение попадает в область принятия гипотезы, то нет причины отвергать H_0; если наблюдаемое значение попадает в критическую область, то H_0 отвергается на заданном уровне значимости α.

Лекция №11.

§38. Проверка гипотезы о виде распределения случайной величины.

Критерий согласия Пирсона.

Пусть X – генеральная совокупность (случайная величина) с неизвестной функцией распределения $F_X(x)$.

$x^{(1)}, x^{(2)}, \ldots.\ x^{(n)}$ – выборка (результаты измерений; $n \geq 50$)

1) H_0: выборка сделана из генеральной совокупности X с функцией распределения $F_X(x)$

2) H_1: функция распределения X отлична от $F_X(x)$

3) задаем α - уровень значимости

4) строим статистический критерий

Разобьем область, которой принадлежат результаты измерений, на r промежутков $S_i[a_{i-1};a_i)$ одинаковой длины ($i=1, 2, \ldots, r$; $a_0=-\infty$, $a_1, \ldots, a_{(r-1)}$, $a_r=+\infty$)

Пусть

n_i – количество значений X из числа наблюдаемых, которые принадлежат S_i; $\sum n_i = n$;

p_i – вероятность того, что значения X принадлежат промежутку S_i, вычисленная в предположении справедливости нулевой гипотезы;

$$p_i = P(X \in S_i) = P(a_{i-1} \leq X < a_i) = F_X(a_i) - F_X(a_{i-1}) = \Delta F_X(a_i) \quad (\textstyle\sum p_i = 1);$$

$p_i^* = \dfrac{n_i}{n}$ - относительная частота попадания значений X из числа наблюдаемых в промежуток S_i.

Тогда

$$p_i = P(X \in S_i) = P(a_{i-1} \leq X < a_i) = F_X(a_i) - F_X(a_{i-1}) = \Delta F_X(a_i) \quad (\textstyle\sum p_i = 1);$$

$$p_i^* = F_x^*(a_i) - F_x^*(a_{i-1}) = \Delta F_x^*(a_i) \qquad (\textstyle\sum p_i^* = 1);$$

$F_x^*(x)$ - эмпирическая функция распределения.

За меру отклонения истинной функции распределения $F_X(x)$ от эмпирической $F_x^*(x)$ возьмем следующую случайную величину $\chi^2 = \sum_{i=1}^{r} \dfrac{(n_i - np_i)^2}{np_i} = \sum_{i=1}^{r} (p_i^* - p_i)^2 \dfrac{n_i}{n}$

Теорема Пирсона:

Какова бы ни была $F_X(x)$, при $n\to\infty$ распределение величины χ^2 стремится к распределению хи-квадрат с $k=r-1$ степенями свободы.

Замечание. Если в процессе проверки гипотезы приходится производить оценку параметров распределения, то количество степеней свободы χ^2 уменьшается на количество неизвестных параметров.

χ^2 – статистический критерий для проверки нулевой гипотезы

5) строим критическую область

Критическая область - правосторонняя

$P(\chi^2 < \chi^2_{\text{кр.}}) = 1-\alpha$

$\chi^2_{\text{кр.}} = \chi^2_{1-\alpha}(r-1)$ – квантиль распределения хи-квадрат с $k=r-1$ степенями свободы.

6) считаем наблюдаемое значение критерия $\chi^2_{\text{набл.}}$ и сравниваем его с критическим $\chi^2_{\text{кр.}}$

7) Вывод: если $\chi^2_{\text{набл.}} < \chi^2_{\text{кр.}}$, то нет оснований отвергнуть нулевую гипотезу на выбранном уровне значимости α, в противном случае H_0 отвергается на заданном уровне значимости α.

Пример.

По выборке объема $n=50$ проверить гипотезу о нормальном распределении генеральной совокупности.

Номер интервала	Границы интервалов	Частота
1	10-12	2
2	12-14	4
3	14-16	7
4	16-18	11
5	18-20	15
6	20-22	8
7	22-24	3

В качестве параметров нормального распределения выберем их точечные оценки: $x_B =$ 17,76; s^2 = 8,798; s = 2,966

1) H_0: генеральная совокупность X имеет нормальное распределение с параметрами a=17,76 и σ= 2,966;

2) H_1: распределение X отлично от $N(17{,}76;\ 2{,}966)$;

3) α =0,1

4) строим статистический критерий

i	S_i	n_i	p_i	np_i	n_k	np_k	$n_k - np_k$	$(n_k - np_k)^2/np_k$
1	-∞;12	2	0,0026	0,13	2+4=6	5,1	0,9	0,159
2	12-14	4	0,0994	4,97				
3	14-16	7	0,1756	8,78	7	8,78	-1,78	0,361
4	16-18	11	0,2543	12,715	11	12,715	-1,715	0,231
5	18-20	15	0,2445	12,225	15	12,225	2,775	0,630
6	20-22	8	0,1472	7,36	8+3=11	11,18	-0,18	0,003
7	22;+∞	3	0,0764	3,82				
	сумма	*50*	*1*	*50*	*50*	*50*	*0*	***1,384***

$$p_1 = P(X \in S_1) = P(-\infty \le X < 12) = \Phi_0(\frac{12-17,76}{2,966}) - \Phi_0(\frac{-\infty-17,76}{2,966}) = 0,0026$$

$$p_2 = P(X \in S_2) = P(12 \le X < 14) = \Phi_0(\frac{14-17,76}{2,966}) - \Phi_0(\frac{12-17,76}{2,966}) = 0,0994$$

…………………

Замечание: если для какого-либо интервала не выполняется условие $np_i \ge 5$, то этот интервал объединяется с соседним интервалом.

1,384 – наблюдаемое значение статистического критерия χ^2

5) строим критическую область

$\chi^2_{кр.} = \chi^2_{1-\alpha}(k)$ - квантиль распределения хи-квадрат с $k=r-l-1$ степенями свободы порядка 1-α.

l - число неизвестных параметров распределения, которые пришлось оценивать.

l =2, r=5

$\chi^2_{кр.} = \ \chi^2_{0,9}(2) = 4,61$

6) χ^2 набл. =1,384 χ^2 кр.=4,61

1,384 < 4,61

7) Вывод: нет оснований отвергнуть нулевую гипотезу на уровне значимости α =0,1.

<u>Критерий Колмогорова – Смирнова</u>

в классическом виде является более мощным, чем критерий Пирсона; используется для проверки гипотезы о соответствии эмпирического любому теоретическому непрерывному распределению $F_X(x)$ *с заранее известными параметрами;* применим для негруппированных данных или для группированных в случае малой ширины интервала.

$x^{(1)}, x^{(2)}, \ldots. x^{(n)}$ – выборка (результаты измерений; $n \geq 35$)

1) H_0: выборка сделана из генеральной совокупности X с функцией распределения $F_X(x)$

2) H_1: функция распределения X отлична от $F_X(x)$

3) α - уровень значимости

4) строим статистический критерий

$$D_n = \sup_x \left| F_X(x) - F_X^*(x) \right|$$

K= $\sqrt{n} D_n$ - статистика критерия Колмогорова

5) строим критическую область

Критическая область – правосторонняя; критические значения $K_{кр.} = \lambda_{1-\alpha}$ составляют: $\lambda_{0.9}$ = 1.22; $\lambda_{0.95}$ = 1.36; $\lambda_{0.99}$ = 1.63

6) сравниваем наблюдаемое значение критерия $K_{набл.}$ с критическим $K_{кр.}$

7) Вывод: если $K_{набл.} < K_{кр.}$, то нет оснований отвергнуть нулевую гипотезу на выбранном уровне значимости *α*, в противном случае H_0 отвергается на заданном уровне значимости α.

§38. *Проверка гипотез о параметрах известного распределения генеральной совокупности*

Проверка гипотез о параметрах нормально распределенной генеральной совокупности

$X \to N(m;\sigma)$ $x^{(1)}, x^{(2)}, \ldots, x^{(n)}$ – выборка

H_0	Предположени относительно других параметров	Статистический критерий K	Распределение K при условии справедливости H_o	Область принятия гипотезы H_o для двустороннего критерия	Область принятия гипотезы H_o для правостороннего критерия
$\sigma^2 = \sigma_0^2$	m-известно	$K = \frac{nS_0^2}{\sigma_0^2}$ $S_0^2 = \frac{1}{n}\sum (X^{(i)} - m)^2$	$\chi^2(n)$	$\chi^2_{\frac{\alpha}{2}}(n) < K < \chi^2_{1-\frac{\alpha}{2}}(n)$ $H_1: \sigma^2 \neq \sigma_0^2$	$K < \chi^2_{1-\alpha}(n)$ $H_1: \sigma^2 > \sigma_0^2$
	m-неизвестно	$K = \frac{(n-1)S^2}{\sigma_0^2}$ $S^2 = \frac{1}{n-1}\sum (x^{(i)} - X_B)^2$	$\chi^2(n-1)$	$\chi^2_{\frac{\alpha}{2}}(n-1) < K < \chi^2_{1-\frac{\alpha}{2}}(n-1)$ $H_1: \sigma^2 \neq \sigma_0^2$	$K < \chi^2_{1-\alpha}(n-1)$ $H_1: \sigma^2 > \sigma_0^2$
$\sigma_1^2 = \sigma_2^2$	m_1,m_2 - известн	$s_{01}^2 > s_{02}^2$ $K = \frac{S_{01}^2}{S_{02}^2}$	*F*(n_1;n_2)	$K < F_{1-\alpha/2}(n_1;n_2)$ $H_1: \sigma_1^2 \neq \sigma_2^2$	$K < F_{1-\alpha}(n_1;n_2)$ $H_1: \sigma_1^2 > \sigma_2^2$
	m_1,m_2- неизвестны	$s_1^2 > s_2^2$ $K = \frac{S_1^2}{S_2^2}$	*F*(n_1-1;n_2-1)	$K < F_{1-\alpha/2}(n_1-1;n_2-1)$	$K < F_{1-\alpha}(n_1-1;n_2-1)$ $H_1: \sigma_1^2 > \sigma_2^2$

H_0	Предположени относительно других параметров	Статистический критерий K	Распределение K при условии справедливости H_o	Область принятия гипотезы H_o для двустороннего критерия	Область принятия гипотезы H_o для правостороннего критерия
$m = m_0$	σ^2 - известна	$K = \dfrac{X_B - m_0}{\sigma / \sqrt{n}}$	$N(0;1)$	$\lvert K \rvert < u_{1-\alpha/2}$ $H_1: m \neq m_0$	$K < u_{1-\alpha}$ $H_1: m > m_0$
	σ^2 -неизвестн	$K = \dfrac{X_B - m_0}{S / \sqrt{n}}$	$T(n-1)$	$\lvert K \rvert < t_{1-\alpha/2}(n-1)$ $H_1: m \neq m_0$	$K < t_{1-\alpha}(n-1)$ $H_1: m > m_0$
$m_1 = m_2$	σ_1^2 и σ_2^2 известны	$K = \dfrac{X_{B1} - X_{B2}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$	$N(0;1)$	$\lvert K \rvert < u_{1-\alpha/2}$ $H_1: m_2 \neq m_1$	$K < u_{1-\alpha}$ $H_1: m_1 > m_2$
	σ_1^2 и σ_2^2 - неизвестны А) выборки независимы, справедлива гипотеза $\sigma_1^2 = \sigma_2^2$	$K = \dfrac{X_{B1} - X_{B2}}{S\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$, где $S = \left(\dfrac{(n_1 - 1)S_1^2}{n_1 + n_2 - 2} + \dfrac{(n_2 - 1)S_2^2}{n_1 + n_2 - 2} \right)^{1/2}$	$T(n_1 + n_2 - 2)$	$\lvert K \rvert < t_{1-\alpha/2}(n_1 + n_2 - 2)$ $H_1: m_2 \neq m_1$	$K < t_{1-\alpha}(n_1 + n_2 - 2)$ $H_1: m_1 > m_2$

	σ_1^2 и σ_2^2 - неизвестны Б) выборки зависимы и $\sigma_1^2=\sigma_2^2$	$K=\dfrac{\overline{d}}{S_d/\sqrt{n}}$, где $d^{(i)}=X_1^{(i)}-X_2^{(i)}$ $\overline{d}=\dfrac{1}{n}\sum d^{(i)}$ $S_d=\dfrac{1}{n-1}\sum(d^{(i)}-\overline{d})^2$	$N(0;1)$	$\lvert K\rvert < u_{1-\alpha/2}$ $H_1: m_2\neq m_1$	$K < u_{1-\alpha}$ $H_1: m_1>m_2$
	σ_1^2 и σ_2^2 - неизвестны В) выборки независимы, справедлива гипотеза $\sigma_1^2\neq\sigma_2^2$	$K=\dfrac{X_{B1}-X_{B2}}{\sqrt{\dfrac{S_1^2}{n_1}+\dfrac{S_2^2}{n_2}}}$	$\approx T(k)$ $k=\dfrac{\left(\dfrac{S_1^2}{n_1}+\dfrac{S_2^2}{n_2}\right)^2}{\left(\dfrac{S_1}{n_1}\right)^2\dfrac{1}{n_1-1}+\left(\dfrac{S_2}{n_2}\right)^2\dfrac{1}{n_2-1}}$	$\lvert K\rvert < t_{1-\alpha/2}(k)$ $H_1: m_2\neq m_1$	$K < t_{1-\alpha}(k)$ $H_1: m_1>m_2$

Пример. На завод поступила партия станков. По результатам исследования 13 станков найдена исправленная выборочная дисперсия размера изготовления станками деталей s^2=14,6. Требуется ли дополнительная наладка станка, если допустимая характеристика рассеяния контролируемого размера деталей σ_0^2=12? (α=0,01-уровень значимости)

Решение.

1). $H_0 : \sigma^2 = \sigma_0^2$

2). $H_1 : \sigma^2 > \sigma_0^2$

3). $\alpha=0{,}01$

4). $K = \dfrac{(n-1)S^2}{\sigma_0^2} = \chi^2\,(n-1)$

5). $\chi^2_{кр}=\chi^2_{1-\alpha}(12)= \chi^2_{0,99}(12)=26{,}2$

6). $\chi^2_{набл.}= \dfrac{12*14{,}6}{12} = 14{,}6$

7). $\chi^2_{набл.}<26{,}2 \Rightarrow$ нет основания отвергнуть H_0 на уровне значимости 0,01, следовательно подналадка не нужна.

*<u>Проверка гипотез о параметре **p** биномиального распределения.</u>*

*(сравнение относительной частоты с гипотетической вероятностью **<u>p</u>** появления события в отдельном испытании)*

$X \to B(n;p)$; $x^{(1)}, x^{(2)}, \ldots. x^{(n)}$ – выборка;

x_k	x_1	x_2
w_k	m/n	(n-m)/n

H_0	Статистический критерий K	Распределение K при условии справедливости H_o	Область принятия гипотезы H_o для двустороннего критерия	Область принятия гипотезы H_o для правостороннего критерия
$p = p_0$	$K = \dfrac{(\frac{m}{n} - p_0)\sqrt{n}}{\sqrt{p_0 q_0}}$, где $q_0 = 1 - p_0$	$\approx N(0;1)$	$\lvert K\rvert < u_{1-\alpha/2}$ $H_1: p \neq p_0$	$K < u_{1-\alpha}$ $H_1: p > p_0$
$p_1 = p_2$	$K = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}}$, где $\hat{p}_1 = \dfrac{m_1}{n_1}$; $\hat{p}_2 = \dfrac{m_2}{n_2}$	$\approx N(0;1)$	$\lvert K\rvert < u_{1-\alpha/2}$ $H_1: p_1 \neq p_2$	$K < u_{1-\alpha}$ $H_1: p_1 > p_2$

В электронных таблицах Excel для проверки гипотез о параметрах нормально распределенных генеральных совокупностей по результатам экспериментов есть специальные тесты, упрощающие процедуру вычислений.

Двухвыборочный Z-тест для средних служит для проверки гипотезы о различии между средними математическими ожиданиями двух нормальных распределений с известными дисперсиями.

Двухвыборочный t-тест с одинаковыми (различными) дисперсиями используется для проверки гипотезы о равенстве средних двух нормально распределенных генеральных совокупностей.

Парный двухвыборочный t-тест используется для проверки гипотезы о равенстве средних в том случае, если обе выборки сделаны из одной и той же генеральной совокупности (например, в разные моменты времени, до и после какого-либо воздействия).

Двухвыборочный f – тест для дисперсий служит для проверки гипотез о равенстве дисперсий двух нормальных распре

Лекция №12.

§39. Понятие о регрессионном анализе. Линейная выборочная регрессия. Метод наименьших квадратов (МНК). (см. лекцию 8)

Основные задачи регрессионного анализа:

- ✓ Вычисление выборочных коэффициентов регрессии
- ✓ Проверка значимости коэффициентов регрессии
- ✓ Проверка адекватности модели
- ✓ Выбор лучшей регрессии
- ✓ Вычисление стандартных ошибок, анализ остатков

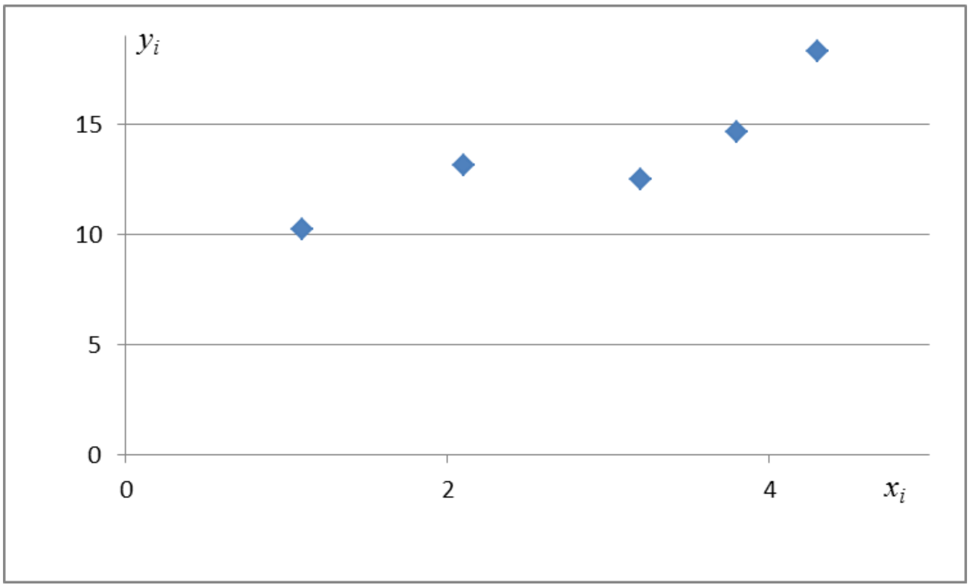
Построение простой регрессии по экспериментальным данным.

Предположим, что случайные величины X и Y связаны линейной корреляционной зависимостью $y = \phi(x) = Ax + B$, для отыскания которой проведено n независимых измерений X и Y.

$x_1,\ x_2,\ \ldots\ x_n$

$y_1,\ y_2, \ldots\ y_n$

Диаграмма рассеяния (разброса, рассеивания)



(x_i, y_i) – координаты экспериментальных точек.

Выборочное уравнение прямой линии регрессии Y на X имеет вид

$y = \hat{A}x + \hat{B}$.

Задача: подобрать $\hat{A}$ и $\hat{B}$ таким образом, чтобы экспериментальные точки как можно ближе лежали к прямой $y = \hat{A}x + \hat{B}$.

Для того, что бы провести прямую $y = \hat{A}x + \hat{B}$, воспользуемся МНК. Потребуем, чтобы $\sum_{i=1}^{n} \delta_i^2 \to \min$, где $\delta_i = \hat{y}_i - y_i$, $\hat{y}_i = \hat{A}x_i + \hat{B}$.

Постулаты регрессионного анализа, которые должны выполняться при использовании МНК.

1. Y и δ подчинены нормальному закону распределения.
2. Дисперсия Y постоянна и не зависит от номера измерения.
3. Результаты наблюдений y_i в разных точках независимы.
4. Входные переменные x_i независимы, неслучайны и измеряются без ошибок.

Введем функцию ошибок $\Phi(\hat{A}, \hat{B}) = \sum_{i=1}^{n} \delta_i^2 = \sum_{i=1}^{n} \left(y_i - \left(\hat{A}x_i + \hat{B}\right)\right)^2$ и найдем ее минимальное значение.

$$\begin{cases} \dfrac{\partial \Phi(\hat{A}, \hat{B})}{\partial \hat{A}} = \sum_{i=1}^{n} 2\left(y_i - \hat{A}x_i - \hat{B}\right)\left(-x_i\right) = 0, \\ \dfrac{\partial \Phi(\hat{A}, \hat{B})}{\partial \hat{B}} = \sum_{i=1}^{n} -2\left(y_i - \hat{A}x_i - \hat{B}\right) = 0. \end{cases}$$

$$\begin{cases} \sum_{i=1}^{n} \left(y_i - \hat{A}x_i - \hat{B}\right)x_i = 0, \\ \sum_{i=1}^{n} \left(y_i - \hat{A}x_i - \hat{B}\right) = 0. \end{cases}$$

$$\begin{cases} \hat{A}\cdot\left(\sum_{i=1}^{n} x_i^2\right) + \hat{B}\cdot\left(\sum_{i=1}^{n} x_i\right) = \sum_{i=1}^{n} x_i y_i, \\ \quad \hat{A}\cdot\left(\sum x_i\right) + \hat{B}\cdot n = \sum_{i=1}^{n} y_i\,. \end{cases}$$

Решив систему, получим искомые значения $\hat{A}$ и $\hat{B}$

$$\begin{cases} a_{11}\cdot\hat{A} + a_{12}\cdot\hat{B} = b_1, \\ a_{21}\cdot\hat{A} + a_{22}\hat{B} = b_2. \end{cases}$$

$$\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\cdot a_{22} - a_{12}\cdot a_{21} = n\cdot\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2$$

$$\Delta_{\hat{A}} = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix} = b_1\cdot a_{22} - a_2\cdot b_{12} = n\cdot\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \cdot \sum_{i=1}^{n} y_i$$

$$\Delta_{\hat{B}} = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix} = a_{11}b_2 - a_{21}b_1 = \sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i$$

$$\hat{A} = \frac{\Delta_{\hat{A}}}{\Delta} = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i}^{n} x_i \sum y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

$$\hat{B} = \frac{\Delta_{\hat{B}}}{\Delta} = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}.$$

$\hat{A}$ и $\hat{B}$ является несмещенными оценками истинных значений коэффициентов A и B.

$$\hat{A} = \frac{\frac{1}{n-1}\left(\sum_{i=1}^{n} x_i y_i - n x_B y_B\right)}{\frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n(x_B)^2\right)} = \frac{K_{xy}}{s_x^2} \quad \text{или} \quad \hat{A} = \frac{\frac{1}{n}\left(\sum_{i=1}^{n} x_i y_i - n x_B y_B\right)}{\frac{1}{n}\left(\sum_{i=1}^{n} x_i^2 - n(x_B)^2\right)} = \frac{\hat{\mu}_{xy}}{D_x^B}$$, где

$K_{xy} = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i y_i - n x_B y_B\right)$-несмещенная оценка корреляционного момента (ковариации),

$s_x^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n(x_B)^2\right)$- несмещенная оценка дисперсии X,

$\hat{\mu}_{xy} = \frac{1}{n}\left(\sum_{i=1}^{n} x_i y_i - n x_B y_B\right)$- выборочная ковариация,

$D_x^B = \frac{1}{n}\left(\sum_{i=1}^{n} x_i^2 - n(x_B)^2\right)$- выборочная дисперсия X.

$\hat{B} = y_B - \hat{A}x_B$

$\hat{r}_{xy} = \dfrac{K_{xy}}{s_x s_x}$ - выборочный коэффициент корреляции

Коэффициент детерминации

$$\underbrace{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}_{SSe} = \sum_{i=1}^{n}(y_i - y_B - (\hat{y}_i - \bar{y}))^2 = \underbrace{\sum_{i=1}^{n}(y_i - y_B)^2}_{SS} - \underbrace{\sum_{i=1}^{n}(\hat{y}_i - y_B)^2}_{SSp};$$

$y_i -$ *наблюдаемое экспериментальное значение СВ* Y *при* $x = x_i$

$\hat{y}_i -$ *предсказанное значение* Y , *удовлетворяющее уравнению регрессии*

$y_B -$ *средневыборочное значение* Y

$SS = SS_p + SS_e$

$R^2 = \dfrac{SS_p}{SS}$ - коэффициент детерминации, доля изменчивости Y, объясняемая рассматриваемой регрессионной моделью. Для парной линейной регрессии $R^2 = \hat{r}_{xy}^2$.

Коэффициент детерминации принимает значения от 0 до 1. Чем ближе значение коэффициента к 1, тем сильнее зависимость. При оценке регрессионных моделей это используется для доказательства адекватности модели (качества регрессии). Для приемлемых моделей предполагается, что коэффициент детерминации должен быть хотя бы не меньше 0,5 (в этом случае коэффициент множественной корреляции превышает по модулю 0,7). Модели с коэффициентом детерминации выше 0,8 можно признать достаточно хорошими (коэффициент корреляции превышает 0,9). Подтверждение адекватности модели проводится на основе дисперсионного анализа путем проверки гипотезы о значимости коэффициента детерминации.

$H_0 : R^2 = 0 \Rightarrow$ *регрессия незначима*

$H_1 : R^2 > 0 \Rightarrow$ *регрессия значима*

α – уровень значимости

$F = \dfrac{R^2}{1 - R^2}(n-2) \to F(1, n-2)$ - статистический критерий

Критическая область – правосторонняя; $F_{крит} = F_{1-\alpha}(1, n-2)$

Если $F_{крит} < F_{набл}$, то нулевая гипотеза отвергается на заданном уровне значимости, следовательно, коэффициент детерминации значим, следовательно, регрессия адекватна.

§40. Мощность статистического критерия. Функция мощности.

Верная гипотеза \ Принимаемая Гипотеза	H_0	H_1
H_0	$1-\alpha$	α
H_1	β	$1-\beta$

Определение. *Мощностью критерия* $1-\beta$ называют вероятность попадания критерия в критическую область при условии, что справедлива конкурирующая гипотеза.

Задача: построить критическую область таким образом, чтобы мощность критерия была максимальной.

Определение. *Наилучшей критической областью* (НКО) называют критическую область, которая обеспечивает минимальную ошибку второго рода β.

Пример 1: По паспортным данным автомобиля расход топлива на 100 километров составляет 10 литров. В результате измерения конструкции двигателя ожидается, что расход топлива уменьшится. Для проверки были проведены испытания 25 автомобилей с модернизированным двигателем; выборочная средняя расхода топлива по результатам испытаний составила 9,3 литра. Предполагая, что выборка получена из нормально распределенной генеральной совокупности с математическим ожиданием m и дисперсией σ^2=4л^2, проверить гипотезу, утверждающую, что изменение конструкции двигателя не повлияло на расход топлива.

1) H_0: m=10 л

2) H_1: m<10 л

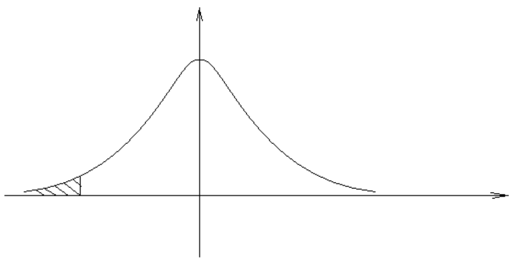
3) Уровень значимости $\alpha = 0{,}05$

4) Статистический критерий

$$K = \frac{X_B - m}{\sqrt{\frac{\sigma^2}{n}}} \to N(0;1)$$

5) Критическая область - левосторонняя

$$K_{кр} = u_{0,05} = -u_{0,95} = -1{,}645$$

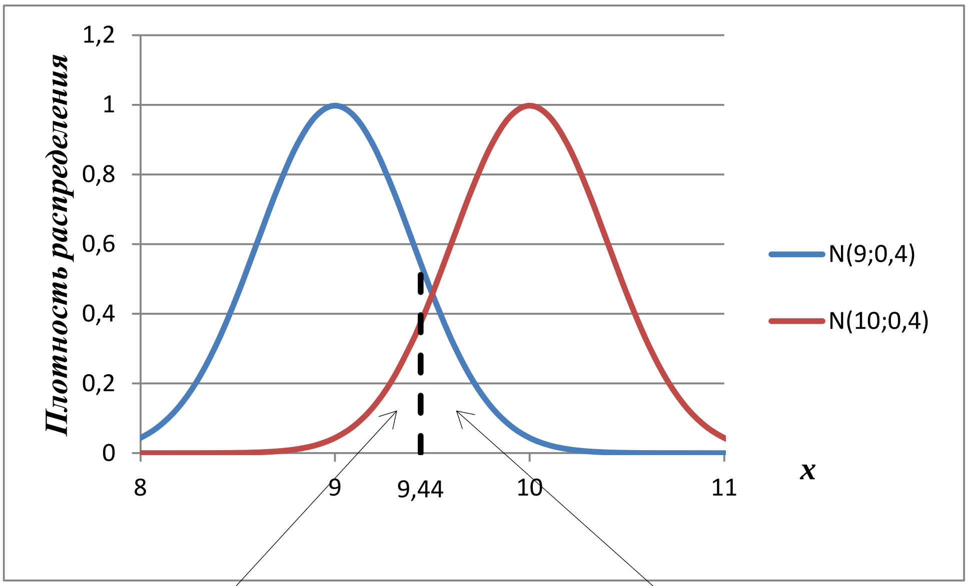


6) $K_{набл} = \dfrac{9{,}3-10}{\sqrt{\dfrac{4}{25}}} = -1{,}75$

7) $K_{набл} < K_{кр}$ следовательно H_0 отвергается на уровне значимости $\alpha = 0{,}05$

Пример 2: В условиях примера 1 предположим, что наряду с H_0 : m=10 л рассматривается конкурирующая гипотеза H_1: m=9 л, а критическая область задана неравенством $X_B < 9{,}44$ Найти вероятность ошибок I рода и II рода.

$X_B \to N(m; \dfrac{\sigma}{\sqrt{n}})$



$S_1 = \alpha$ $S_2 = \beta$

$$\alpha = P(X_B < 9{,}44 / H_0 : m = 10) = \Phi\left(\frac{9{,}44-10}{\sqrt{\frac{4}{25}}}\right) = \Phi(-1{,}4) = 0{,}08$$

(≈8% автомобилей имеют меньший расход топлива)

$$\beta = P(X_B > 9{,}44 / H_1 : m = 9) = 1 - \Phi\left(\frac{9{,}44-9}{\sqrt{\frac{4}{25}}}\right) = 0{,}136$$

(≈13,6% автомобилей, имеющих расход топлива 9л на 100 км, классифицируются как автомобили, имеющие расход 10 литров).

Определение. Пусть проверяется H0: $\theta = \theta_0$, а V_K - критическая область критерия с заданным уровнем значимости α. Функцией мощности критерия $M(V_K, \theta)$ называется вероятность отклонения Hо как функция параметра θ, т.е.

$$M(V_K, \theta) = P(K \in V_K|_\theta)$$

$M(V_K, \theta_0) = \alpha$ - ошибка 1-ого рода

$M(V_K, \theta_1) = P(K \in V_K) = 1 - \beta$ - мощность критерия

Пример 3: построить график функции мощности из примера 2 для n=25 и n=100

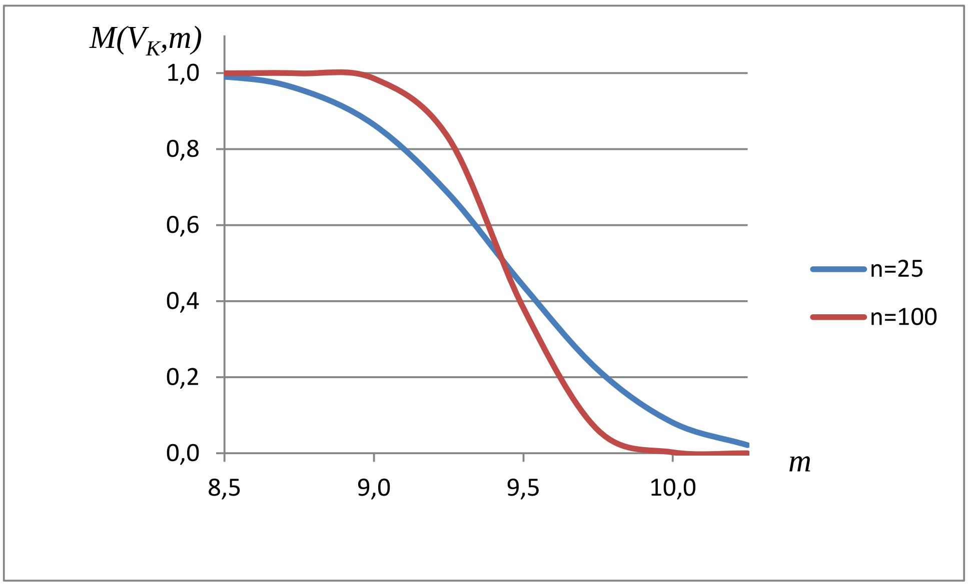
$\theta \to m$

$V_k : X_B < 9{,}44$ попадает в критическую область.

n=25 $\quad M(V_K, m) = P(X_B < 9{,}44|_m) = \Phi\left(\frac{9{,}44-m}{\sqrt{\frac{4}{25}}}\right)$

n=100 $\quad M(V_K, m) = P(X_B < 9{,}44|_m) = \Phi\left(\frac{9{,}44-m}{\sqrt{\frac{4}{100}}}\right)$

m	n=25	n=100
8,00	1,000	1,000
8,50	0,991	1,000
8,75	0,958	1,000
9,00	0,864	0,986
9,25	0,683	0,829
9,50	0,440	0,382
9,75	0,219	0,061
10,00	0,081	0,003
10,25	0,021	0,000



Пример 4. Какой минимальный объем выборки следует взять в условии примера 2 для того, чтобы обеспечить $\alpha = 0{,}01$ и $\beta \le 0{,}1$. $V_K = ?$

$$\alpha = P(X_B < K_{кр} / H_0 : m = 10) = \Phi\left(\frac{K_{кр} - 10}{\sqrt{\frac{4}{n}}}\right) = 0{,}01$$

$$\beta = P(X_B > K_{кр} / H_1 : m = 9) = 1 - \Phi\left(\frac{K_{кр} - 9}{\sqrt{\frac{4}{n}}}\right) \le 0{,}1$$

$$\begin{cases} \dfrac{K_{кр}-10}{\sqrt{\dfrac{4}{n}}} = u_{0,01} = -2{,}326 \\ \dfrac{K_{кр}-9}{\sqrt{\dfrac{4}{n}}} \geq u_{0,9} = 1{,}282 \end{cases} \Rightarrow n \geq 53, \quad K_{кр}\big|_{n=53} \geq 9{,}361$$

Лемма Неймана-Пирсона.

При проверке простой гипотезы H_0 против простой альтернативной гипотезы H_1 наилучшая критическая область (НКО) критерия заданного уровня значимости α состоит из точек выборочного пространства (выборок объема n), для которых справедливо неравенство:

$$\frac{L(X_1, X_2,..., X_n / H_0)}{L(X_1, X_2,..., X_n / H_1)} < C_\alpha$$

C_α - константа, зависящая от α;

$X_1, X_2,..., X_n$ - элементы выборки;

L- функция правдоподобия при условии, что соответствующая гипотеза верна.

Пример 5.

Случайная величина X имеет нормальное распределение с параметрами m и σ, σ известно. Найти НКО для проверки H_0: $m=m_0$, против H_1: $m=m_1$, причем $m_0<m_1$

Решение:

$$\frac{L(X_1, X_2,..., X_n / H_0)}{L(X_1, X_2,..., X_n / H_1)} = \frac{\dfrac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{\sum (X_i - m_0)^2}{2\sigma^2}}}{\dfrac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{\sum (X_i - m_1)^2}{2\sigma^2}}} = a \cdot e^{-\frac{X_B(m_1-m_0)}{\sigma^2/n}} < C_\alpha,$$

где $a = e^{-\frac{(m_0{}^2 - m_1{}^2)}{2\sigma^2/n}}$;

Ошибка первого рода: $P(X_B > K_{кр} / H_0) = \alpha$

$$P(X_B < K_{кр} / H_0) = \Phi\left(\frac{K_{кр} - m_0}{\sigma/\sqrt{n}}\right) = 1 - \alpha$$

$$\frac{K_{кр} - m_0}{\sigma/\sqrt{n}} = u_{1-\alpha}$$

НКО: $X_B > m_0 + u_{1-\alpha}\dfrac{\sigma}{\sqrt{n}}$.