УЧЕБНО ПОМАГАЛО ПО СТАТИСТИКА

Юлияна Костадинова Бонева

СОФИЯ
2013

Анотация Учебното помагало е предназначено за студентите от специалност Устройство и управление на земи и имоти от УАСГ. То съдържа решени задачи, дадени в края на всяка тема, както и нерешени задачи за упражнение.

Целта на помагалото е да помогне на студентите да овладеят основните задачи от теорията на вероятностите и математическата статистика. В последната глава са показани примерни решения с MATLAB на част от задачите. Като приложение са дадени основни статистически таблици.

Съдържание

Глава 1

КОМБИНАТОРИКА

Комбинаториката е дял от математиката, който изучава избора и наредбата на обекти, състоящи се от краен брой елементи, а също така методите за броене на общото количество начини по които това може да се направи. При решаването на задачи от комбинаториката се налага да определим броя на елементите на някакво крайно множество. За тази цел се използват две основни правила.

Теорема 1.1. *(Правило на произведението) Ако един елемент A на едно множество може да се избере по* m_a *начина, а друг елемент B по* m_b *начина, то изборът на двойката елементи A и B може да стане по* $m_a.m_b$ *начина.*

Пример 1.1. *В група по танци има 10 мъже и 6 жени. Колко двойки може да съставим така, че всички мъже да танцуват с всички жени?*

За да съставим двойка, е необходимо първо да се избере мъж, а след това жена или обратно. Изборът на мъж може да стане по 10 начина, а изборът на жена може да стане по 6 начина. Тогава според правилото на произведението общото количество на различните смесени двойки е 10.6=60.

Теорема 1.2. *(Правило на сумата) Ако някакъв елемент A на едно множество може да бъде избран по* m_a *начина, а друг елемент B по* m_b

начина, то изборът на А или В може да бъде осъществен по $m_a + m_b$ начина.

Пример 1.2. *Ако от град А до град В има две различни шосета, три железопътни линии и маршрутът се обслужва от две различни авиокомпании, по колко различни начина можем да стигнем от А до В?*

От град А до град В можем да стигнем по 2+3+2=7 различни начина.

Нека е дадено множество $A = \{a_1, a_2, \dots a_N\}$, състоящо се от краен брой N различни елементи. Това множество ще наричаме *генерална съвкупност*. Всяка k-торка от елементи на генералната съвкупност се нарича *извадка* с обем k.

Нека имаме генерална съвкупност, която съдържа N елемента. От нея се прави извадка с обем k, $k \leq N$, като последователно се изваждат k елемента и се записват техните номера. Ако след всяко изваждане елементът не се връща в генералната съвкупност, казваме, че *извадката е без връщане*. Ако след всяко изваждане елементът се връща, *извадката е с връщане*. Една извадка се нарича *наредена*, когато елементите се подреждат по реда на тяхното изтегляне. Например (ABC) и (BCA) са различни, ако те са наредени. При ненаредените извадки редът на появяване на елементите е без значение. В този случай (ABC) и (BCA) са една и съща извадка.

1.1 Пермутации

Пермутации без повторение са наредени извадки без връщане от елементите на генералната съвкупност, съдържаща N елемента, като всяка извадка е с обем N. Две пермутации се различават една от друга по реда на елементите, участващи в тях. Ако с P_N се означи броят на пермутациите от N елемента, то

$$P_N = N(N-1)(N-2)\dots 2.1 = N!.$$

Пермутации с повторение имаме тогава, когато от всички N елемента n_1 се повтарят, други n_2 се повтарят и т.н. n_k се повтарят. Броят на всички пермутации в този случай е $\widetilde{P}_N$, където

$$\widetilde{P}_N = \frac{N!}{n_1!n_2!\dots n_k!}.$$

Задачи:

1. По колко начина 5 деца могат да седнат на една пейка?

<u>Решение</u>: Броят на всички различни начина е равен на броя на пермутациите от 5 елемента: $P_5 = 5! = 1.2.3.4.5 = 120$.

2. Колко различни седемцифрени числа могат да се съставят от две единици, две двойки и три тройки?

<u>Решение</u>: Използваме формулата за пермутация с повторение. При $n_1 = 2$, $n_2 = 2$ и $n_3 = 3$, се получава $\widetilde{P}_7 = \frac{7!}{2!2!3!} = 210$.

3. Колко различни шестцифрени числа кратни на 5 без повторение на цифрите могат да се получат от цифрите 1, 2, 3, 4, 5 и 6?

<u>Решение</u>: За да е кратно на 5, числото трябва да завършва на 5. На останалите 5 позиции имаме да разположим 5 цифри - 1, 2, 3, 4 и 6. Броят на числата е $5! = 120$.

4. Енциклопедия се състои от 8 тома, номерирани от 1 до 8. По колко възможни начина могат да се подредят на етажерка така, че томовете да не следват един след друг по реда на своите номера? (Отг. 40319)

5. На рафт от библиотека са поставени 3 поредици от приключенски романи, всяка от които съдържа по 5 книги, номерирани от 1 до 5. Трите поредици са една до друга, но отделните номерирани книги не са подредени според номерацията си. Да се намери по колко начина могат да бъдат подредени тези книги. (Отг. $(5!)^3.3!$)

6. Петя и Петър посещават курсове по народни танци в група от 20 ученици. Ръководителят ги нарежда по случаен начин.

а) Колко са различните нареждания в редица, при които между Петя и

Петър има точно 5 ученици?
б) Когато танцът е кръгово хоро, да се намери колко са подредбите, при които между Петя и Петър по дъгата от Петя към Петър в посока обратна на часовниковата стрелка има точно 5 други ученици?
Решение: а) Всички възможни подреждания на учениците без Петя и Петър са (20-2)! на брой. Независимо от това как са разположени 5-мата ученици (които са между Петя и Петър) и 20-5-1 ученици (които не са между Петя и Петър), Петя може да заеме мястото си в редицата по (20-5-1) начина, Петър също. Следователно двамата могат да заемат мястото си в редицата по 2.14 начина. На всяко от тези подреждания съответстват 18! начина за подреждане на другите ученици. Тогава броят на различните подреждания в редица, при които между Петя и Петър има точно 5-ма ученици, е 2.14.18!.
б) Когато танцът е кръгово хоро и броенето на учениците е в посока обратна на часовниковата стрелка, така че между тях да има 5 ученика, мястото на Петя и Петър е без значение. Тогава броят на нарежданията е (20-2)!=18!.

7. Да се намери броят на различните пермутации на буквите в думата СТАТИСТИКА. (Отг. 75600)

8. Номер на мобилен телефон се състои от 10 различни цифри и започва с 08. Колко са възможностите за останалите цифри от номера? (Отг. 40320)

9. При игра на бридж между четирима играчи се разпределя колода от 52 карти, по 13 на всеки. Да се намери по колко различни начина могат да получат карти играчите?
Решение: Търсим броя на различните извадки със състав (13,13,13,13) с обем 52. Трябва да се намери броят на пермутациите с повторения $\widetilde{P}_{52}(13, 13, 13, 13)$. Така за броя на търсените начини получаваме

$$\widetilde{P}_{52}(13, 13, 13, 13) = \frac{52!}{(13!)^4}.$$

10. Нека имаме 8 топки, от които 5 бели, 1 черна, 1 червена и 1 синя.

Да се намери по колко начина можем да образуваме редица, съдържаща 6 топки. (Отг. 178)

1.2 Вариации

Вариации без повторение са наредени извадки без връщане. Ако генералната съвкупност има обем N, а извадката е с обем k, казваме, че имаме вариации от N елемента k-ти клас. Броят на вариациите от N елемента k-ти клас се бележи с V_N^k и се дава с формулата

$$V_N^k = N(N-1)(N-2)\dots(N-k+1) = \frac{n!}{(n-k)!}.$$

Вариации от N елемента k-ти клас с повторения са наредени извадки с обем k. Броят им се намира по формулата

$$\widetilde{V}_N^k = N^k.$$

Задачи:

11. Изчислете броя на всички различни трицифрени числа, които могат да се съставят от цифрите:

а) от 1 до 9;

б) от 0 до 9. (В числата да няма повтарящи се цифри)

Решение: а) Броят на всички трицифрени числа, (всяко от което е съставено от различни цифри), които могат да се съставят с цифрите от 1 до 9 е $V_9^3 = 9.8.7 = 504$.

б) Броят на всички различни трицифрени числа е V_{10}^3, намален с броя на трицифрените числа, чиято първа цифра е 0: $V_{10}^3 - V_9^2 = 648$.

12. Студентска група има 5 изпита, които студентите трябва да положат през изпитната сесия, която продължава 20 дни. По колко начина може да се направи разписание за изпитите на групата? (Отг. 29649)

13. Асансьор с 4 - ма пътници може да спре на 8 етажа. По колко начина пътниците могат да слязат по етажите, ако е известно, че всеки от четиримата пътници слиза на един от осемте етажа? (Отг. 1680)

14. Да се намери колко числа, по-малки от 10000, могат да се запишат с помощта на цифрите 1 и 2.

Решение: Числото 10000 има 5 цифри. Числата, които са по-малки от 10000 имат 4 цифри и са записани само с цифрите 1 и 2, могат да имат:
4-цифри - те са на брой $\widetilde{V}_2^4 = 2^4 = 16$;
3-цифри - те са на брой $\widetilde{V}_2^3 = 2^3 = 8$;
2-цифри - те са на брой $\widetilde{V}_2^2 = 2^2 = 4$.
Следователно търсеният брой е 28.

15. В състезание участват 24 ученици. Да се намери по колко различни начина могат да бъдат заети първите 3 места. (Отг 12144)

16. Автомобилни номера включват една, две или три букви и четири цифри. Да се намери броят на такива номера, ако се използват 5 букви от българската азбука и десетте цифри. (Отг. 1550000)

17. Колко е броят на различните номера на мобилни телефони от вида: 0888******, които завършват на едноцифрено просто число? (Отг. 4.10^5)

18. Да се решат уравненията
а) $V_n^5 = 18V_{n-2}^4$. (Отг. 9 и 10)
б) $V_n^4 = 12V_n^2$. (Отг. 6)

19. Дете подрежда по случаен начин точно 5 фигурки от картон в редица, от ляво на дясно. Три от фигурките са квадратни, а две са с формата на кръг. Квадратните се различават една от друга по дължината на страната си, а тези с формата на кръг са с различни радиуси. Намерете броя на начините за подреждане на петте фугурки, ако се започва и завършва с квадратна. (ДЗИ 2012) (Отг. 36)

1.3 Комбинации

Комбинации без повторение от N елемента k-ти клас са крайните ненаредени множества, съдържащи по k елемента, избрани от N елемента. Следователно комбинациите се различават помежду си не по местата на избраните елементи, а само по самите елементи. Броят на всевъзможните комбинации от N елемента k-ти клас се бележи с C_N^k и се изчислява по формулата

$$C_N^k = \frac{V_N^k}{P_k} = \frac{N(N-1)(N-2)\dots(N-k+1)}{1.2.\dots k}.$$

Комбинации с повторения от N елемента k-ти клас е всяка ненаредена извадка с връщане с обем k от една генерална съвкупност от N елемента. Две комбинации с повторение са еднакви, ако се състоят от равен брой елементи a_1, равен брой елементи a_2 и т.н. Например комбинациите с повторение $(1,2,3,1,2),(3,2,1,1,2),(1,1,2,3,2)$ от 5-ти клас на елементите $\{1,2,3\}$ са еднакви. Нека с $\widetilde{C}_N^k$ означим броя на различните комбинации с повторение от N елемента, k-ти клас. За този брой имаме формулата

$$\widetilde{C}_N^k = C_{N+k-1}^{N-1} = C_{N+k-1}^k.$$

Задачи:

20. Студентска група се състои от 12 български и 8 чуждестранни студенти. По колко възможни начина могат да бъдат избрани 5 студенти, трима от които да са българи?

<u>Решение</u>: Изчисляват се броят на комбинациите от 12 елемента трети клас и броят на комбинациите от 8 елемента втори клас. Получените комбинации се умножават (използва се правилото на умножението): $C_{12}^3.C_8^2 = \frac{12.11.10}{1.2.3}.\frac{8.7}{1.2} = 220.28 = 7160.$

21. Партида съдържа 9 изделия. От нея се избират 3 изделия с връщане. Какъв е броят на извадките?

<u>Решение</u>: В случая имаме извадки с обем 3 от генералната съвкупност, съдържаща 9 изделия, с връщане на извадено изделие. Броят на всички извадки е равен на броя на комбинациите с повторения: $\widetilde{C}_9^3 = C_{11}^3 = \frac{11.10.9}{1.2.3} = 165.$

22. По колко начина могат да се образуват от 20 спортисти 2 отбора с по 10 спортисти? (Отг. 184756)

23. По колко начина могат да се разделят 30 ябълки и 10 круши между 4 деца?

Решение: Ако 30 - те ябълки считаме за еднакви, можем да разгледаме разпределението им на 4 места като комбинация с повторение на 30 елемента, 4-ти клас. Следователно броят на начините е $\widetilde{C}_{30}^4 = C_{30+4-1}^{4-1} = C_{33}^3 = 5456$. Аналогично, 10 круши могат да се разпределят между 4 деца по $\widetilde{C}_{10}^4 = C_{4+10-1}^{4-1} = C_{13}^3 = 286$. начина. Но на всяко едно разделяне на ябълките между децата отговарят 286 разделяния на крушите и приложим правилото за умножение, получаваме $5456.286 = 1560416$ начина.

24. В книжарница продават марки от 10 вида, като от всеки вид има поне 8 марки. Да се намери по колко начина могат да се купят 8 марки. (Отг. 24310)

25. Колко прави могат да се построят през 7 точки, от които три лежат на една права?(Отг. 19)

26. Колко диагонала има правилен 10-ъгълник? (Отг. 35)

27. Събрание от 80 души избира комисия с председател, секретар и трима членове за извършване на определена дейност. По колко начина може да се направи това? (Отг. $V_{80}^2.C_{78}^3 = 480800320$)

28. Колко окръжности са определени от 10 точки, ако никои 3 от тях не лежат на една права, но 4 от тях лежат на една окръжност? (Отг. 117)

29. Да се докаже, че за всеки две естествени числа n и k такива, че $n \geq k$, е в сила $C_n^k = C_n^{n-k}$.

Глава 2

ВЕРОЯТНОСТ. ОСНОВНИ ФОРМУЛИ И ТЕОРЕМИ.

2.1 Събития и действия с тях. Класическо определение за вероятност.

Основно понятие в теорията на вероятностите е понятието *събитие*. Събитието е резултат от провеждането на опит, изходът от който не можем да предвидим. Под влияние на различни фактори можем да имаме различни изходи. Всички възможни изходи при даден опит образуват *пространството от елементарни събития* при този опит (ще означаваме с S). Всеки един от елементите на това пространство наричаме *елементарно събитие*. Подмножество на множеството S наричаме *случайно събитие*. Празното множество се нарича *невъзможно събитие*. То не може да настъпи при даден опит. Събитието S наричаме *сигурно събитие*. За всяко събитие A се определя неговото *противоположно събитие* $\overline{A}$, което се състои от всички елементи на S, несъставляващи A. Образуването на $\overline{A}$ от A е операция допълнение. Нека G е съвкупността от всички събития - подмножества на S. Под *обединение* или сума на две събития A и B се разбира събитие C, което се състои от елементарните събития, влизащи поне в едно от събитията A или B. Обединението се означава с $A \cup B$. *Сечение* (произведение) на две събития A и B е събитие D, състоящо се от всички елементарни събития, общи за A и B. Сечението се означава с $A \cap B$. Когато сечението на две множества е празното множество, събитията се наричат несъвместими. За операциите допълнение, обединение и сечение са в сила свойствата:

1. $A \cup B = B \cup A$; $A \cap B = B \cap A$.
2. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
$A \cap (B \cap C) = (A \cap B) \cup (A \cap C)$.
3. За всяко $A \in G$, $\ A \cup \emptyset = A$.
4. За всяко $A \in G$, $\ A \cap S = A$.
5. $A \cup \overline{A} = S$, $A \cap \overline{A} = \emptyset$.

На всеки от елементарните изходи от S може да се съпостави число P, което се нарича *вероятност* на елементарното събитие. Нека $S = \{w_1, w_2, \dots w_n\}$ е пространството с краен брой елементарни събития при даден експеримент и на всяко елементарно събитие е съпоставено число $P(w_i)$. Числата $P(w_i)$, $\ i = 1, 2, \dots n$, се наричат вероятности на изходите $w_1, w_2, \dots w_n$, ако

$$P(w_i) \geq 0; i = 1, 2, \dots n.$$

$$P(w_1) + P(w_2) + \dots P(w_n) = 1.$$

Формула за класическа вероятност: $P(A) = \frac{m}{n}$.
Числата m и n се наричат съответно брой на благоприятните за настъпване на събитието A елементарни изходи и брой на всички възможни изходи при дадения опит. Вероятностите на събитията в класическия модел имат свойствата:
1. $0 \leq P(A) \leq 1$, за всяко $A \in G$,
2. $P(S) = 1$,
3. $P(A) + P(\overline{A}) = 1$ за всяко $A \in G$.

Постановка на задачата при класическа вероятност.
Провежда се случаен опит, в който броят на елементарните изходи е

краен и всички изходи са еднакво възможни. Да се намери вероятността на случайно събитие А, свързано с този опит.

План за решение.

1. Определяме пространството от елементарни събития.
2. Проверяваме дали всички елементарни изходи са равновъзможни.
3. Преброяваме елементарните изходи, които водят до настъпването на събитието А (благоприятни изходи m). Това се прави непосредствено или с помощта на комбинаторика.
4. Преброяваме общия брой елементарни изходи n (броя на елементите на пространството от елементарни събития).
5. Съгласно формулата за класическа вероятност пресмятаме

$$P(A) = \frac{m}{n}.$$

Пример 2.1. *Опит - стрелба по мишена. Случайно събитие - попадение.*
Опит - хвърляне на монета. Случайно събитие - падане на герб.
Опит - хвърляне на правилен зар. Случайно събитие - падане на нечетен брой точки. Благоприятните събития са 1, 3, 5.

Пример 2.2. *Монета се хвърля 3 пъти и получените резултати се записват.*
а) Определете пространството на елементарните събития;
б) Определете елементарните събития, които са съставящи на следните събития
б1) A= най-малко два пъти се е паднало лице;
б2) B= първите три хвърляния са герб;
б3) C= последното хвърляне е лице;
в) Посочете елементарните събития, съставящи следните събития $A \cup B$ *и* $A \cap B$.

Решение: а) Пространството от елементарни събития ще се състои от 2.2.2=8 елементарни събития. Нека L е появата на ези, а G - на тура в едно отделно хвърляне. Тогава $S = LLL, LLG, LGG, LGL, GGG, GGL, GLL, GLG.$

б) Съответните събития ще влючват следните елементарни събития:
б1) $\{LLL, LLG, LGL, GLL\}$
б2) $\{GGG\}$
б3) $\{LLL, LGL, GLL, GGL\}$.

в) $A \cup B = \{LLL, LLG, LGL, GLL\} \cup \{GGG\} = \{LLL, LLG, LGL, GLL, GGG\}$. $A \cap B = \{LLL, LLG, LGL, GLL\} \cap \{GGG\} = \emptyset$.

Задачи:

1. Опишете пространствата от елементарни събития за всеки от следващите опити:
а) произвеждаме изстрел по мишена, представляваща 10 концентрични кръга, номерирани от 1 до 10;
б) провежда се футболен мач между два отбора.

2. Колко елементарни събития има всяко от следващите случайни събития:
а) сумата на две едноцифрени числа да е 12;
б) произволно избрана дата от календара да е 13-ти;
в) случайно избрана дума от множеството А={ромб, квадрат, хипотенуза, перпендикуляр, куб, пирамида} да съдържа не по- малко от две гласни.

3. На десет картончета са написани числата $1, 2, 3, \ldots 10$. Случайно е взето едно картонче. В кои от следващите отговори са дадени всички възможни изходи от опита:
а) {четно, нечетно}
б) {просто; 4; 6; 8; 9; 10}
в) {четно; 1; 3; 5}
г) {не повече от три; не по - малко от четири}.

4. Посочете кои от следващите събития са 1) случайни 2) сигурни 3) невъзможни:
а) изваждане на цветна топка от кутия, в която има 3 сини и 5 червени

топки;
б) получаване на 25 точки от 4 изпита, ако за всеки изпит максималният брой точки е 5;
в) при хвърляне на зар да се паднат повече от 6 точки;
г) при хвърляне на зар да се паднат 4 точки;
д) при теглене на карта от колода от 52 карти да се паднат две аса.

5. Кои от следните двойки събития са несъвместими:
а) случайно избрано естествено число между 1 и 100 включително: се дели на 10; се дели на 11;
б) попадение; пропуск при еднократен изстрел;
в) печеливш; губещ в шахматна партия;
г) случайно избрано естествено число между 1 и 25 включително е; четно; кратно на три?

6. Посочете неверните твърдения:
а) Всеки от възможните резултати при даден опит се нарича елементарно събитие;
б) При всяко изпитване се сбъдва поне едно елементарно събитие;
в) Елементарните събития при хвърляне на един зар са 6;
г) Всяко събитие се състои поне от едно елементарно събитие;
д) Събитието падане на четен брой точки при хвърляне на един зар се състои от три елементарни събития.

7. Един съд съдържа 9 бели и 11 черни топки. Извадена е по случаен начин една топка. Нека събитието A е изваждане на бяла топка, а събитието B - изваждане на черна топка. Посочете верните твърдения:
а) Събитията A и B са противоположни;
б) Събитията A и B са несъвместими;
в) Събитията A и B са елементарни;
г) За сбъдването на събитието A има 9 благоприятни случая.

8. Хвърляме 2 правилни различни зарчета. Да се намери вероятността, точките които се падат на едното зарче да са два пъти повече от точките, които се падат на другото.

Решение: а) Определяме пространството от елементарни събития.

$$(1,1)(1,2)(1,3)(1,4)(1,5)(1,6)$$
$$(2,1)(2,2)(2,3)(2,4)(2,5)(2,6)$$
$$(3,1)(3,2)(3,3)(3,4)(3,5)(3,6)$$
$$(4,1)(4,2)(4,3)(4,4)(4,5)(4,6)$$
$$(5,1)(5,2)(5,3)(5,4)(5,5)(5,6)$$
$$(6,1)(6,2)(6,3)(6,4)(6,5)(6,6)$$

б) Проверяваме елементарните изходи за равна възможност за поява. Тъй като зарчетата са правилни, всички елементарни изходи са равновъзможни.

в) Преброяваме изходите, които водят до настъпването на събитието А (благоприятни).

За събитието А благоприятни са тези изходи (n_1, n_2), за които $n_1 = 2n_2$. Това са $(1,2), (2,4), (3,6), (2,1), (4,2), (6,3)$. Така $m = 6$.

г) Преброяваме общия брой елементарни изходи $n = 36$.

д) Съгласно класическото определение за вероятност

$$P(A) = \frac{m}{n} = \frac{6}{36}.$$

9. Хвърляме две правилни различни зарчета. Да се намери вероятността на събитието А - сумата от получените точки е четно число.

Решение: Ще използваме таблицата от зад. 8. Отделните изходи, които са равновероятни, са 36 на брой. Събитието А настъпва при 18 от тях. За вероятността на събитието А се получава $P(A) = \frac{18}{36} = \frac{1}{2}$.

10. Телефонен номер се състои от четири цифри. Да се намери вероятността всичките му цифри да са различни.

Решение: Всички възможни изходи са равни на броя на всички четирицифрени числа. Благоприятни изходи са тези, които могат да се получат като наредени четворки, получени от цифрите $0, 1, 2, \ldots 9$. Тогава $m = V_{10}^4 = 5040$ и $P(A) = \frac{5040}{10000} = 0.504$.

11. Партида съдържа 15 изделия, 5 от които са дефектни. Избираме по случаен начин 4 изделия. Да се намери вероятността на събитието:

а) Едно от изделията е дефектно, трите са качествени.

б) Всички избрани изделия са дефектни.

Решение: а) Четири изделия могат да се изберат от 15 по C_{15}^4 начина. Едно от изделията ще бъде дефектно, когато се избере измежду 5 - те дефектни, което може да стане по C_5^1 начина, а трите качествени се избират между 10 - те по C_{10}^3 начина. Тогава вероятността на събитието е $\frac{C_5^1.C_{10}^5}{C_{15}^4}$.

б) Благоприятните случаи за събитието са C_5^4 на брой, а общите са C_{15}^4 на брой. Така вероятността е $\frac{C_5^4}{C_{15}^5}$.

12. На рафт в библиотека са поставени по случаен начин една до друга 4 различни книги по математика и 3 различни книги по химия. Да се намери вероятността книгите по един и същ предмет да са една до друга.

Решение: Броят на всички възможни начини за поставяне на книгите по математика и химия са $(4+3)! = 7!$. За да намерим броя на благоприятните случаи за събитието A={книгите по един и същ предмет да са една до друга} умножаваме броя на пермутациите от 4 елемента с броя на пермутациите от 3 елемента. Имайки предвид, че книгите по математика могат да са както в дясно, така и в ляво на книгите по химия, получаваме, че броят на благоприятните изходи за събитието A е 2.3!.4!. Следователно $P(A) = \frac{2.3!.4!}{7!}$.

13. Цифрите 1, 2, 3, 4, 5 и 7 са записани на картончета и поставени в кутия. По случаен начин се изваждат едно след друго три картончета и се поставят едно до друго в реда на изтеглянето. Да се намери вероятността полученото число да е четно. (Отг. $\frac{1}{3}$)

14. От колода от 52 карти се вади случайно една. Да се намери вероятността извадената карта да е асо. (Отг.≈ 0.077)

15. На отделни картончета са написани буквите А, Е, В, О, Р, Я. Избират се 4 картончета по случаен начин и се подреждат в произволен

ред. Да се намери вероятността да се получи думата ВЯРА. (Отг.$\frac{1}{360}$).

16. От колода съдържаща 52 карти, по случаен начин се изваждат 2 карти. Да се намери вероятността да са осмица и дама. (Отг. $\approx 0,0264$)

17. Да се намери вероятността при игра на белот играч да получи кварта, т.е. 4 поредни карти от един вид (купи, кари, пики, спатии).
Решение: При игра на белот участват 32 карти - 8 купи, 8 кари, 8 пики и 8 спатии. Общият брой на начините за раздаване на 32 карти на 4 човека по 8 карти е равен на C_{32}^{8}. Нека получената кварта е от купи. Тъй като броят на купите е 8, са възможни 5 кварти от купи. На всяка една от тези 5 кварти съответстват C_{28}^{4} разпределения по 4 на неучастващите в квартата 28 карти. Така получаваме, че за петте кварти от купите са възможни $5.C_{28}^{4}$ начина за разпределение на картите. Общо картите са 4 вида, затова броя на благоприятните изходи за събитието е равен на $4.5.C_{28}^{4}$. Така търсената вероятност е $P = \frac{4.5.C_{28}^{4}}{C_{32}^{8}}$.

18. Да се намери вероятността при игра на белот играч да получи терца, т.е. три последователни карти от един и същи вид. (Отг. ≈ 0.271)

19. От 20 ученици (15 момчета и 5 момичета) след жребии определят група от 4 - ма души за участие в олимпиада. Каква е вероятността групата да се състои от 2 момичета и две момчета. (Отг. $\approx 0,2167$)

20. Да се намери вероятността в играта "6 от 49" на Спортния тотализатор в един попълнен фиш да са познати а) 6 числа б) 3 числа. (Отг. а)$7,1511.10^{-8}$, б) $\approx 0,0176$)

21. В група от 20 души 5 - ма не могат да плуват. Да се намери вероятността две случайно посочени лица от групата да не могат да плуват. (Отг. $\frac{1}{19}$)

22. Десет пътници се качват в три вагона. Да се намери вероятността в един от вагоните да се качат 6 - ма, в другия трима, а в третия - един човек. (Отг. $\frac{C_{10}^{6}.C_{4}^{3}.C_{1}^{1}}{3^{10}} \approx 0,0142$)

23. В шахматен турнир участват 20 състезатели, които по жребии се разделят на две равни по брой групи. Да се намери вероятността двамата най - силни играчи да попаднат в различни групи. (Отг. $\approx 0,526$)

2.2 Условна вероятност. Теорема за умножение на вероятности.

Дефиниция 2.1. *Условна вероятност на събитието A при условие, че се е сбъднало събитието B, се нарича отношението*

$$P(A/B) = \frac{P(A \cap B)}{P(B)}.$$

Дефиниция 2.2. *Събитията $A_1, A_2, \dots A_n$ са независими, когато*

$$P(A_i \cap A_j) = P(A_i).P(A_j), \;\; \forall A_i, A_j, \; i \neq j.$$

Събитията $A_1, A_2, \dots A_n$ са независими в съвкупност, ако за всяко множество от индекси $\{i_1, i_2, \dots i_m\}$, т.ч. $1 \leq i_1 < i_2 < \dots < i_m \leq n$, $2 \leq m \leq n$ е изпълнено

$$P(A_{i_1} \cap A_{i_2} \cap \dots A_{i_m}) = \prod_{k=1}^{m} P(A_{i_k}).$$

Теорема 2.1. *(Теорема за умножение на вероятности) Нека*

$$P(\bigcap_{k=1}^{m} A_k) > 0, \;\; 1 \leq m \leq n, \;\; k = 1, 2, \dots n.$$

Тогава

$$P(\bigcap_{k=1}^{m} A_k) = P(A_1). \prod_{m=1}^{n-1} P(A_{m+1}/\bigcap_{k=1}^{m} A_k).$$

В случая на две събития имаме

$$P(A \cap B) = P(A).P(B/A) = P(B).P(A/B).$$

Пример 2.3. *1. В партида, която съдържа 20 изделия, 15 са стандартни. По случаен начин се изважда едно изделие, което не се връща в партидата, след което отново се изважда едно изделие. Да се пресметне вероятността при второто изваждане изделието да се окаже стандартно, ако първия път изваденото изделие е също стандартно.*

Решение: Когато втори път изваждаме по случаен начин изделие, в партидата има 19 изделия. Множеството на елементарните събития съдържа $n = 19$ елемента, а благоприятните са 14, защото преди това е извадено стандартно изделие. Събитията B={ първия път да се извади стандартно изделие}, и A={ втория път да се извади стандартно изделие} са зависими и $P(A/B) = \frac{14}{19} \approx 0,74$.

Пример 2.4. *Хвърляме правилно зарче. Да се намери вероятността, броят точки които се падат, да е по-голям от три (събитие A), ако е известно, че се е паднала страна с четен брой точки (събитие B).*

Решение: На събитие B съответстват 2, 4 и 6 точки, а на събитие A - 4, 5 и 6 точки. Тогава $A \cap B = \{4, 6\}$ и $P(A/B) = \frac{P(A\cap B)}{P(B)} = \frac{2}{6} : \frac{3}{6} = \frac{2}{3}$.

Задачи:

24. От колода от 52 карти изваждаме по случаен начин 4 карти. Да се намери вероятността извадените карти да са от различен цвят.

Решение: Нека A_1 е събитието първата извадена карта да е от произволен цвят, A_2 - събитието цветът на втората карта да не съвпада с цвета на първата, A_3 - събитието цветът на третата карта да е различен от тези на първите две карти и A_4 - последната карта да е от последния останал цвят. Събитията са зависими, следователно за вероятността на събитието A се получава

$$P(A) = P(A_1).P(A_2/A_1).P(A_3/A_1 \cap A_2).P(A_4/A_1 \cap A_2 \cap A_3).$$

$$P(A_1) = 1,\ P(A_2/A_1) = \frac{39}{51},\ \ P(A_3/A_1 \cap A_2) = \frac{26}{50},$$

$$P(A_4/A_1 \cap A_2 \cap A_3) = \frac{13}{49}; P(A) \approx 0,1.$$

25. Един съд съдържа 5 бели, 4 червени и 1 черна топка. Вземаме една след друга 2 топки без връщане. Да се намери вероятността първата извадена топка да е бяла, а втората - черна.

Решение: Разглеждаме събитията A={първата взета топка е бяла} и B={втората топка е черна}. Тогава $P(A \cap B) = P(A).P(B/A)$. $P(A) = \frac{5}{10}$, $P(B/A) = \frac{1}{9}$. Така търсената вероятност е $P(A \cap B) = \frac{1}{18}$.

26. Спортист участва в състезание. Вероятността да преодолее първото препятствие е $p_1 = 0,9$, да преодолее второто препятствие е $p_2 = 0,9$ и вероятността да преодолее третото препятствие е $p_3 = 0,85$. Намерете вероятността, спортистът да преодолее всичките три препятствия.

Решение: Трите събития са независими. Затова $P = p_1.p_2.p_3 = 0,612$.

27. От цифрите 1, 2, 3, 4, 5 се избира една цифра, а след това от останалите още една. Да се намерят вероятностите на събитията:

а) Първата цифра да е четна, втората - нечетна;

б) И двете цифри да са четни. (Отг. а)0,3, б)0,1)

28. Да се намери вероятността на събитието в случайно избрана година 29 февруари да се окаже петък. (Отг. $\frac{1}{28}$)

29. В една партида от 16 изделия има 10 стандартни и 6 нестандартни. Последователно се вземат по случаен начин 3 изделия за проверка. Да се намери вероятността първите две да са стандартни, а третото - нестандартно. (Отг. $\frac{9}{56}$)

30. В пакет има 20 химикалки, от които 12 са червени, 5 са сини и останалите - черни. От пакета е взета произволна химикалка и после е върната в него. След това отново по случаен начин е взета химикалка. Да се намери вероятността нито една от двете взети химикалки да не е червена. (Отг. $\frac{4}{25}$)

2.3 Теорема за събиране на вероятности

Теорема 2.2. *Вероятността за настъпване на поне едно от събитията $A_1, A_2, \dots A_n$ е*

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n) -$$

$$-P(A_1 \cap A_2) - P(A_1 \cap A_3) - \dots - P(A_{n-1} \cap A_n) + P(A_1 \cap A_2 \cap A_3) + \dots + (-1)^n P(A_1 \cap A_2 \cap \dots \cap A_n).$$

В частния случай за две събития се получава

$$P(A \cup B) = P(A_1) + P(A_2) - P(A \cup B).$$

Когато събитията $A_i, \quad i = 1, 2 \dots n$ са несъвместими, то:

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

Вероятността на сума от съвместими събития може да се пресметне и като се използва формулата за вероятност на противоположно събитие. Противоположно на събитието да настъпи поне едно A_i е да не настъпи нито едно A_i. Следователно:

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = 1 - P(\overline{A_1} \cap \overline{A_2} \cap \dots \cap \overline{A_n}).$$

Задачи:

31. Двама студенти се явяват на изпит. Вероятността да си вземе изпита е 0,6 за първия студент и 0,8 за втория. Да се намери вероятността поне един студент да си вземе изпита.

Решение: Нека А и В са събитията съответно първият студент да вземе изпита и вторият студент да вземе изпита. Събитията са съвместими. Тогава

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0,6 + 0,8 - 0,6.0,8 = 0,92.$$

32. Трима стрелци стрелят в цел. Вероятността за попадение е 0,5 за първия, 0,7 за втория и 0,9 за третия. Да се намери вероятността поне един от тях да улучи целта.

Решение: Нека събитията А, В и С са съответно първият, вторият и третият стрелец улучват целта. Събитията са съвместими. Следователно според теоремата за събиране на вероятности ще имаме:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) +$$

$$P(A \cap B \cap C) = 0,685.$$

Задачата може да се реши и с обратна вероятност.

$$P(A \cup B \cup C) = 1 - P(A \cap B \cap C) = 0,685.$$

33. Да се намери вероятността случайно взето двуцифрено число да се окаже кратно на 2 или на 5. (Отг. 0,6)

34. Кутия съдържа 10 червени, 8 сини и 8 зелени молива. По случаен начин се избират два молива. Каква е вероятността избраните моливи да са от различен цвят?
Решение: С А, В и С ще означим събитията
А={ червен и син молив},
В={ червен и зелен молив},
С={ син и зелен молив}. Събитията са несъвместими. Затова

$$P(A \cup B \cup C) = P(A) + P(B) + P(C).$$

Според формулата за класическа вероятност ще имаме:

$$P(A) = \frac{C_{10}^1 . C_8^1}{C_{26}^2}, \ \ P(B) = \frac{C_{10}^1 . C_8^1}{C_{26}^2}, \ \ P(C) = \frac{C_8^1 . C_8^1}{C_{26}^2}.$$

Търсената вероятност е $P \approx 0,692$.

35. Конспект съдържа 30 въпроса, от които студент знае 25. Явявайки се на изпит, той тегли два въпроса. Да се намери вероятността студентът да си вземе изпита, ако за това е достатъчно да знае и двата паднали му се въпроса или един от тях и един допълнително зададен му въпрос.

Решение: Студентът ще си вземе изпита, ако се случи точно едно от двете събития: А={знае и двата паднали му се въпроса} или В={ знае един от тях и един допълнителен}. За да му се паднат два въпроса, които знае, трябва да ги избере измежду 25 - те. Затова $P(A) = \frac{C_{25}^2}{C_{30}^2}$. Събитието В се състои от две събития: B_1 ={да изтегли един от 25-те и един от 5-те, които не знае} и B_2 ={при дотълнителното теглене да изтегли въпрос, който знае.}. Така $P(B_1) = \frac{C_{25}^1}{C_5^1}$ и $P(B_2) = \frac{24}{28}$. Според теоремата за умножение на вероятности $P(B) = P(B_1).P(B_2)$, а според теоремата за събиране на несъвместими събития

$$P(A \cup B) = P(A) + P(B) = 0,94.$$

36. От партида, съдържаща 30 качествени и 10 дефектни изделия избирзме по случаен начин 5 изделия. Да се намери вероятността на събитието поне 4 от извадените изделия да са качествени.

Решение: Събитието поне 4 от изделията да са качествени е сума от несъвместимите събития A - точно 4 от изделията са качествени и B - точно 5 от изделията са качествени.

$$P(A) = \frac{C_{30}^4 . C_{10}^1}{C_{40}^5}, \quad P(B) = \frac{C_{30}^5}{C_{40}^5}, \quad P(A \cup B) = 0,63.$$

37. Две урни съдържат съответно: първата - 5 бели и 3 червени топки, а втората - 7 бели и 4 червени топки. Ако от всяка урна се извадят наведнъж по две топки, да се намери вероятността да се извадят две бели топки поне от едната урна. (Отг. ≈ 0.6026)

38. От група от 10 мъже и 12 жени трябва да се изберат по случаен начин 3 - ма души за делегати на конференция. Да се намери каква е вероятността поне 2 жени да бъдат изпратени на конференцията. (Отг. $\frac{4}{7}$)

39. Трима изследователи, независимо един от друг, правят измерване на физична величина. Вероятността да допусне грешка при изследването е 0,2 за първия, 0,25 за втория и 0,3 за третия изследовател. Да се намери вероятността поне един от изследователите да допусне грешка, ако всеки от тях е направил по едно измерване. (Отг. 0,58)

40. В кашон от 50 ел. крушки 3 са дефектни. Намерете вероятността от 2 произволно взети крушки и двете да са дефектни. (Отг. 0,002)

41. От колода от 36 карти произволно теглим 3. Каква е вероятността сред тях да се окаже поне една дама? (Отг. 0,3053)

42. Да се намери вероятността да се свържем с даден мобилен номер, ако вероятността, в това време да има сигнал “заето”, е 0,8, а вероятността, абонатът да не отговаря е, 0,4. (Отг. 0,12)

43. Две урни съдържат съответно първата - 6 сини и 7 червени топки, а втората - 8 сини и 5 червени топки. Ако от всяка урна се извадят наведнъж по две топки, каква е вероятността да се извадят две сини

топки поне от едната урна? (Отг. $\frac{489}{1014}$)

44. Хвърлят се два правилни различни зара. Да се намери вероятността поне на един от заровете да се паднат четен брой точки. (Отг. 0,75)

2.4 Формула за пълната вероятност и формула на Бейс

Дефиниция 2.3. *Събитията $H_1, H_2, \dots H_n$ образуват пълна група събития, ако са несъвместими и при всеки опит настъпва точно едно от тях, т.е.*

$$a) \bigcup_{k=1}^{n} H_k = S, \;\; b) H_i \cap H_j = \emptyset, \;\; i \neq j, \;\; c) P(H_k) > 0, k = 1, 2, \dots n.$$

Ако $H_1, H_2, \dots H_k$ образуват пълна група събития, то за всяко събитие

$$A = A \cap S = A \cap (\bigcap_{k=1}^{n} H_k) = \bigcap_{k=1}^{n} (A \cap H_k),$$

и затова

$$P(A) = P(\bigcup_{k=1}^{n} (A \cap H_k)) = \sum_{k=1}^{n} P(A \cap H_k) = \sum_{k=1}^{n} P(H_k).P(A/H_k).$$

Формулата

$$P(A) = \sum_{k=1}^{n} P(H_k).P(A/H_k)$$

се нарича формула за пълната вероятност на случайни събития. Когато е известно, че събитието А е настъпило, условните вероятности на събитията H_i се пресмятат с помоща на формулата на Бейс:

$$P(H_i/A) = \frac{P(H_i).P(A/H_i)}{P(A)}.$$

Можем да използваме формулата за пълната вероятност и формулата на Бейс, ако събитието А може да настъпи само съвместно с едно от образуващите пълна група събития $H_1, H_2, \dots H_n$.

Задачи:

45. Три завода произвеждат еднородна продукция. 70%, 80% и 90% от производството съответно на първия, втория и третия завод е от стандартни изделия. Количествата изделия от трите завода, постъпили в един склад, са в отношение 4: 3: 3. По случаен начин от склада е избрано едно изделие. Каква е вероятността то да бъде стандартно?

<u>Решение</u>: Нека означим с А събитието - избраното изделие да бъде стандартно, а с H_1, H_2, H_3 хипотезите - избраното изделие да бъде от първия, втория и третия завод, съответно. Във формулата за пълната вероятност заместваме

$$P(H_1) = \frac{4}{10},\ \ P(H_2) = \frac{3}{10},\ \ P(H_3) = \frac{3}{10};$$
$$P(A/H_1) = 0,7,\ \ P(A/H_2) = 0,8,\ \ P(A/H_3) = 0,9.$$

За пълната вероятност получаваме

$$P(A) = 0,79.$$

46. В един университет съотношението между редовни и задочни студенти е 3:2. Известно е, че 8% от редовните и 6% от задочните студенти са пълни отличници. По случаен начин е избран един студент и той се оказал отличник. Каква е вероятността студентът да е задочник?

<u>Решение</u>: Нека събитието А е студентът е отличник. Събитието А е настъпило. Затова ще използваме формулата на Бейс. Събитието А може да настъпи съвместно с H_1 - студентът е редовен или с H_2 - студентът е задочник. Затова търсената вероятност е $P(H_2/A)$. Имаме

$$P(H_1) = \frac{3}{5},\ \ P(H_2) = \frac{2}{5},\ \ P(A/H_1) = \frac{8}{100},\ \ P(A/H_2) = \frac{6}{100}.$$

Заместваме във формулата на Бейс и получаваме:

$$P(H_2/A) = \frac{P(H_2).P(A/H_2)}{P(H_1/A).P(H_1) + P(H_2/A).P(H_2)} \approx 0,33.$$

47. Три оръдия стрелят по три цели. Всяко оръдие си избира целта по случаен начин, независимо от другите. Вероятността всяко оръдие да улучи избраната цел е P. Да се намери вероятността да бъдат поразени точно две цели.

<u>Решение</u>: Нека хипотезите са: H_1 - всяко оръдие стреля в различна цел, H_2 - всияки оръдия стрелят по една цел и H_3 - две от оръдията стрелят в една цел, а третото в друга.

$$P(H_1) = 1.\frac{2}{3}.\frac{1}{3} = \frac{2}{9};$$

$$P(H_2) = 1.\frac{1}{3}.\frac{1}{3} = \frac{1}{9};$$

$$P(H_3) = \frac{3.C_3^2.2}{3^3} = \frac{2}{3}, (P(H_3) = 1 - P(H_1) - P(H_2)).$$

$$P(A/H_1) = 3P^2(1-P);, \;\; P(A/H_2) = 0, \;\; P(A/H_3) = (1-(1-P)^2)P = P^2(2-P);$$

$$P(A) = 2P^2 - \frac{3}{4}P^3.$$

48. В три партиди електрически крушки има по 20, 30 и 50 броя. Вероятностите те да работят безотказно една година са съответно $p_1 = 0,7, \;\; p_2 = 0,8, \;\; p_3 = 0,9$. Да се намери вероятността произволно взета крушка от наличните 100 крушки да не се повреди. (Отг. 0,83)

49. От колода от 52 карти са взети две произволни карти без връщане. Да се намери вероятността втората изтеглена карта да е дама. (Отг. $\frac{17}{221}$)

50. От урна, съдържаща 3 бели и 2 черни топчета са преместени 2 случайно взети топчета във втора урна, съдържаща 4 бели и 4 черни топчета. Да се намери вероятността случайно взето топче от втората урна да е бяло. (Отг. 0,52)

51. В урна се намират 4 червени и 5 бели топчета. Извадени са последователно две топчета без връщане. Ако втората извадена топка е бяла, да се намери вероятността и първата да е бяла. (Отг. 0,5)

52. На картончета са написани буквите от думата КОМБИНАТОРИКА, но две от картончетата са се изгубили. Да се намери вероятността

на картонче, изтеглено след това, да се окаже написана гласна буква. (Отг. $\frac{6}{13}$)

53. Студент не успял да се подготви по всичките n въпроса от конспекта за даден изпит. Да се намери дали вероятността да изтегли въпрос, който не знае, е по - малка, когато тегли първи или когато тегли 2-ри? (Отг. една и съща в двата случая)

54. Петър отишъл на изпит, но се бил подготвил по 25 от всичките 30 въпроса. Преди него се явява друг студент и тегли въпрос. Да се намери вероятността Петър да изтегли въпрос, който знае. (Отг.$\frac{5}{6}$)

55. В един университет 30% от мъжете и 20% от жените следват история. 45% от студентите са жени. Ако случайно избран студент учи история, каква е вероятността да е жена? (Отг. 0,353)

56. На изпит се явява група от 10 студенти. Конспектът съдържа 20 въпроса. Трима от студентите могат да отговорят по всички въпроси, четирима от студентите знаят 16 въпроса, двама от тях знаят 10 въпроса и един знае само 5 въпроса. Случайно избран студент отговорил и на трите зададени му въпроса. Да се намери вероятността студентът да е знаел само 16 въпроса. (Отг. 0,38)

57. В автобус пътуват n пътници. На следващата спирка всеки от тях може да слезе с вероятност r. На спирката с вероятност P няма да се качи нов пътник и с вероятност $1 - P$ ще се качи точно един пътник. Да се намери вероятността след потеглянето на автобуса в него да има точно n пътника. (Отг. $P(1-r)^n + (1-P)nr(1-r)^{n-1}$)

58. Рибар лови риба на три места, които посещава с равни вероятности. Ако той хвърли въдицата на първото място, рибата кълве с вероятност 0,4, на второто - с вероятност 0,5 и на третото - с вероятност 0,3. Известно е, че рибарят е хвърлил въдицата три пъти, а рибата е клъвнала само веднъж. Да се намери вероятността рибарят да е ловил на първото място. (Отг. 0,346)

59. В търговска фирма постъпват телевизори в отношение 3:4:5. Телевизорите постъпили от 1-я, 2-я и 3-я доставчик не се нуждаят от ремонт

в течение на гаранционния срок с вероятности съответно 98%, 88% и 92% от случаите. Намерете вероятността, произволно постъпил телевизор във фирмата да не се нуждае от ремонт. (Отг. 0,91)

60. В 9 от 10 еднакви урни има по 2 бели и 2 черни топки, а в десетата - 5 бели и 1 черна топка. От случайно избрана урна е изтеглена бяла топка. Да се намери вероятността тази топка да е извадена от десетата урна. (Отг. $\frac{5}{32}$)

2.5 Вероятност при повтаряне на опитите. Формула на Бернули

Дефиниция 2.4. *Ще казваме, че една редица от опити се провежда по схемата на Бернули, ако са изпълнени следните условия: 1) При всеки опит настъпва единият от два единствено възможни изхода А и $\overline{A}$; 2) Опитите са независими; 3) $P(A)$ има една и съща стойност при всички опити.*

Събитията А и $\overline{A}$ ще наричаме успех и неуспех, а техните вероятности ще означаваме с $P(A) = p$ и $P(\overline{A}) = q$, $p + q = 1$. Нека да разгледаме следната задача. Провеждат се n опита по схемата на Бернули. Търсим вероятността $P_n(k)$ че в n опита събитието А ще се появи точно k пъти. Задачата се решава с помощта на формулата на Бернули

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

Най - вероятният брой k_0 появявания на събитието А при n независими опита удовлетворява неравенствата

$$np - q \leq k_0 \leq np + p.$$

Задачи:

61. Вероятностите за раждане на момче и на момиче са 0,5. Да се пресметне вероятността в семейство с 6 деца да има точно 2 момчета.

Решение: Нека $A =$\{в семейството има точно 2 момчета\}. Провеждаме опити по схемата на Бернули с $n = 6,\ p = 0,5,\ q = 0,4,\ k = 2.$

$$P(A) = P_2(6) = C_6^2.(0,5)^2.(0,5)^4 = 15.(\frac{1}{2})^2.(\frac{1}{2})^4 = \frac{15}{64} \approx 0,234.$$

62. Предприятие купува 4 машини, за всяка от които с вероятност 0,1 ще бъде необходима резервна част от даден вид. Да се намери вероятността да бъдат необходими най-много две резервни части от този вид.
Решение: Означаваме с А събитието, чиято вероятност търсим. За да бъдат необходими най - много две резервни части, това означава, че ще са необходими 0 или 1 или 2 части. Тогава

$$P(A) = P_4(0) + P_4(1) + P_4(2),$$

$$P_4(0) = C_4^0.(0,1)^0.(0,9)^4 = (\frac{9}{10})^4,$$

$$P_4(1) = C_4^1.(0,1)^1.(0,9)^3 = \frac{4.9^3}{10^4},$$

$$P_4(2) = C_4^2.(0,1)^2.(0,9)^2 = \frac{6.9^2}{10^4}.$$

$$P(A) = (\frac{9}{10})^4 + \frac{4.9^3}{10^4} + \frac{6.9^2}{10^4} = 0,9963.$$

63. В цех има 6 мотора. За всеки от тях вероятносттс да бъде включен в даден момент е 0,6. Да се намери вероятността в даден момент да бъдат включени:
а) точно 5 мотора;
б) най-много 5 мотора;
в) поне един мотор. (Отг. а) 0,187 б) 0,953 в)0,97)

64. При игра на бридж (52 карти) единият от четиримата играчи три поредни раздавания не е получавал нито едно асо. Има ли основание да се оплаква, че не му върви?
Решение: Нека означим с А събитието играчът да не получи асо при 1 раздаване. $P(A) = \frac{C_{48}^{13}}{C_{52}^{13}} \approx 0,3038$. В схемата на Бернули имаме

$$P_3(0) = C_3^0.(0,3038)^3 \approx 0,028.$$

Тъй като получената вероятност е малка, играчът има право да се оплаква.

65. Колко Бернулиеви опита трябва да бъдат проведени, за да бъде равен на 51 най-вероятният брой на успехите, ако вероятността за успех е 0,64?

Решение: Имаме

$$k_0 = 51, \ \ p = 0,64, \ \ q = 1 - 0,64 = 0,36.$$

Тогава

$$n.0,64 - 0,36 \leq 51 \leq n.0,64 + 0,64.$$

Така получаваме, че

$$78,68 \leq n \leq 80,25.$$

Следователно трябва да се проведат 79 или 80 опита.

66. Най - вероятният брой доброкачествени изделия в партида от 90 изделия е равен на 82. Каква е вероятността едно изделие в партидата да бъде доброкачествено? (Отг. $\frac{82}{91} \leq \frac{83}{91}$)

67. В студио има 3 телевизионни камери. За всяка от тях вероятността да бъде включена в даден момент е 0,5. Да се намери вероятността в даден момент да бъде включена поне една камера. (Отг. 0,875)

68. Стрелец стреля в цел до първо попадение. Да се намери вероятността да му остане поне един патрон, ако той е получил 10 патрона и вероятността за попадение при всеки изстрел е постоянна и е равна на 0,2. (Отг. 0,376)

69. Да се намери вероятността за настъпване на събитието A при всеки опит, ако най-вероятният му брой появявания при 160 опита е 44. (Отг. $\frac{44}{161} \leq p \leq \frac{45}{161}$)

2.6 Геометрична вероятност

Когато пространството от елементарните изходи при даден опит е неизброимо и може да се представи във вид на част от права, равнина или

пространство, имаме задача от геометрична вероятност. При геометричния модел вероятностите се изчисляват като отношение на мерки (дължини на отсечки, лица на равнинни фигури, обеми на тела). Нека имаме множество от геометрични обекти X и за всяко подмножество $X_1 \subset X$ е дефинирана функция $\mu(X_1)$, удовлетворяваща условията $\mu(X_1) \geq 0$ и $\mu(X_{11} + X_{12}) = \mu(X_{11}) + \mu(X_{12})$ за $X_{11}, X_{12} \subset X_1, \;\; X_{11} \cap X_{12} = \emptyset$.
Нека X_1 има крайна мярка, т.е. $\mu(X_1) < \infty$. За произволно $A \subset X_1$ дефинираме функция, която има свойствата на класическата вероятност. За така дефинираната функция $P(A)$ е в сила формулата

$$P(A) = \frac{\mu(A)}{\mu(X_1)}.$$

Задачи:
70. Пръчка с дължина x се счупва на две части. Да се намери вероятността по-късото парче да има дължина, по-голяма от $\frac{x}{4}$.
Решение: Всички възможни изходи запълват интервала $(0, x)$, защото пръчката може да бъде счупена в произволна своя точка. Разделяме пръчката на 4 равни части с помощта на точките A, B и C. Благоприятни за събитието са тези изходи, при които точката на счупване се намира между A и B или между B и C (средата е в точка B). Дължината на отсечката е мярка на множеството от всички възможни изходи, а сумата от дължините на AB и BC - мярката на благоприятните за събитието изходи. За търсената вероятност се получава

$$P = \frac{\frac{x}{4} + \frac{x}{4}}{x} = \frac{1}{2}.$$

71. В кръг с радиус R е вписан квадрат. Каква е вероятността случайно избрана точка от кръга да се окаже вътре в квадрата?
Решение: Нека M е събитието - произволна точка от кръга k е в квадрата ABCD. Тогава $P(M) = \frac{S_{ABCD}}{S_k}$. Но $AB = R.\sqrt{2}$ и $S_{ABCD} = 2R^2$, следователно $P(M) = \frac{2R^2}{\pi R^2} = \frac{2}{\pi}$.

72. (Задача за срещата) Двама студенти се уговорили да се срещнат на опрелено място между 12 и 13 часа. Дошлият пръв чака в продължение на 15 минути и си тръгва. Да се намери вероятността студентите да

се срещнат, ако всеки от тях случайно избира времето си за идване.
Решение: Нека означим с x момента (между 12 и 13 часа), в който първият студент пристига на определеното място, а с y-момента на пристигане на втория студент. Необходимо и достатъчно условие да се състои срещата, е да бъде изпълнено неравенството $|x-y| \leq 15$, т.е. $-15 \leq x-y \leq 15$. Тогава, ако разглеждаме x и y като координати на точка от равнината Oxy, то моментът на идването на първия и втория между 12 и 13 часа ще определя точка (x, y) от квадрат със страна 60. Така множеството от елементарните събития ще се илюстрира с точките от квадрата $0 \leq x \leq 60,\ \ 0 \leq y \leq 60$. Частта от този квадрат, съдържаща точките от него, в които може да се състои срещата, се определя от неравенствата $-15 \leq x-y \leq 15,\ 0 \leq x \leq 60,\ 0 \leq y \leq 60$. Благоприятните изходи ще бъдат заключени между правите :

$$y = x + 15,\ \ y = x - 15.$$

Търсената вероятност е

$$P = \frac{60^2 - 45^2}{60^2} = \frac{7}{16}.$$

73. Кръгла мишена с диаметър 20 см. е разделена на три части с помощта на две концентрични окръжности с диаметри 15 см. и 10 см. Ако е известно, че мишената е улучена, да се намери вероятността да е поразен вътрешния кръгов венец. (Отг. $\frac{125\pi}{625\pi}$)

74. Минно поле в море е организирано така, че мините са разположени по права линия на всеки 100 метра. Каква е вероятността кораб с широчина 20 метра, пресичащ под прав ъгъл правата, по която се намират мините, да се взриви? (Отг. $\approx 0,2$)

75. В провоъгълен паралелепипед с ръбове 4, 6 и 10 см. е избрана случайно точка M. Намерете вероятността точката да се окаже вътре в даден куб с ръб 3 см., намиращ се вътре в паралелепипеда. (Отг. $\approx 0,113$)

Глава 3

Случайни величини

3.1 Дискретни случайни величини и техните характеристики

Дефиниция 3.1. *Числова величина, която в резултат на даден опит може да приема различни стойности, се нарича случайна величина.*

Случайната величина X се разглежда като числова функция $X(w) : S \to \mathbb{R}$, определена върху множеството S от елементарните изходи на даден експеримент.

В практиката има случаи, когато дадена величина, която ни интересува, може да приема различи стойности. Например броят на слънчевите дни в даден месец, броят на точките при хвърляне на зар, колко пътни произшествия ще има през дадена седмица и т.н. Всички тези примери показват така наречените случайни величини.

Ако множеството от стойности $x_1, x_2, \dots$ на една случайна величина е крайно или безкрайно, но изброимо (могат да се номерират), тя се нарича дискретна случайна величина. Най - често такива величини са резултат на броене.

За да бъде определена напълно една дискретна случайна величина, е необходимо да се знаят всички възможни стойности, които тя може да приеме. Освен възможните стойности трябва да се знаят и вероятностите, с които тя приема тези стойности: $P(X = X_1) = p_1, P(X = X_2) = p_2, \dots P(X = X_k) = p_k$.

Стойностите, които една дискретна случайна величина може да приема и вероятностите с които тя приема тези стойности се представят в таблица наречена *закон за разпределение*:

X	x_1	x_2	x_3	$\dots$	x_k
P	p_1	p_2	p_3	$\dots$	p_k

(3.1)

Очевидно $p_1 + p_2 + \dots p_k = 1$, защото събитията $(X = x_1), (X = x_2), \dots (X = x_k)$ образуват пълна група.

Функция на разпределение на случайната величина X се нарича функция $F(X)$, която е равна на вероятността, случайната величина X да приеме стойност по-малка от стойността на променливата x, т.е

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

Функцията $F(X)$ е неотрицателна, ограничена отгоре и монотонно растяща, непрекъсната в точките x_i и имаща скок, равен на

$$P_i = P(X = x_i) = P(x_i \leq x \leq X_{i+1}) = F(x_{i+1}) - F(x_i), \;\; i = 1, 2, \dots k.$$

Математическото очакване е важна числова характеристика на случайната величина. Това е постоянно число, около което се групират стойностите на случайната величина. То дава приблизителна представа за съответната случайна величина.

Нека е зададена дискретна случайна величина с разпределение зададено с (3.1).

Дефиниция 3.2. *Математическо очакване (средна стойност) на случайната величина X се нарича сумата от произведенията на всичките й възможни стойности и съответните им вероятности, т.е.*

$$EX = \overline{X} = \sum_{i=1}^{k} x_i p_i.$$

Пример 3.1. *Да са намери математическото очакване на случайната величина X - брой точки при хвърляне на един зар.*
Разпределението на случайната величина е

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

За математическото очакване се получава

$$EX = 1.\frac{1}{6} + 2.\frac{1}{6} + 3.\frac{1}{6} + 4.\frac{1}{6} + 5.\frac{1}{6} + 6.\frac{1}{6} = 3,5.$$

Математическото очакване притежава следните по-важни свойства:
1. $EC = C$ (C - неслучайна величина)
2. $E(X \pm Y) = EX \pm EY$ (това свойство може да се докаже и за произволен брой събираеми)
3. $E(kX) = k.E(X)$ (k - константа)
4. $E(X.Y) = EX.EY$, ако X и Y са независими случайни величини.

Пример 3.2. *Дадени са две независими случайни величини с разпределения*

X	1	2
P	$0,4$	$0,6$

и

X	0	1	2
P	$0,2$	$0,3$	$0,5$

Да се намери $E(2X + 3Y)$.
Като се използват свойства 2 и 3, получаваме

$$E(2X + 3Y) = E(2X) + E(3Y) = 2E(X) + 3E(Y).$$

Намираме $EX = 1.0,4 + 2.0,6 = 1,6$ *и* $EY = 0.0,2 + 1.0,3 + 2.0,5 = 1,5$.
Заместваме и получаваме $E(2X + 3Y) = 7,1$.

Само математическото очакване не може да характеризира в достатъчна степен една случайна величина. Може да се случи две случайни величини да имат едни и същи математически очаквания, но разпределенията им съществено да се различават. Освен това трябва да имаме числова характеристика, по която да съдим за отклонението на случайната величина от нейното математическо очакване. Оценка за това разсейване се получава чрез друга важна характеристика, наречена *дисперсия*.

Дефиниция 3.3. *Дисперсия на случайната величина X се нарича математическото очакване на квадрата на отклоненията на случайната величина от нейното математическо очакване:*

$$DX = E(X - EX)^2.$$

Освен със символа DX, дисперсията се означава и със σ_x^2. Квадратният корен от дисперсията се нарича *стандартно отклонение*: $\sigma = \sqrt{DX}$. Като се използват свойствата на математическото очакване, може да се изведе формула, с която практически се работи по - лесно, а именно

$$DX = EX^2 - (EX)^2.$$

Пример 3.3. *Дадена е случайна величина с разпределение*

X	1	2	3
P	0,4	0,2	0,4

Да се намери DX.

Намираме $EX = 1.0,4 + 2.0,2 + 3.0,4 = 2$, *след което изчисляваме* $EX^2 = 1^2.0,4 + 2^2.0,2 + 3^2.0,4 = 4,8$. *Тогава* $DX = 4,8 - 2^2 = 0,8$. *Стандартното отклонение е* $\sigma_x = \sqrt{0,8} = 0,894$.

Забележка. *Квадрат на случайната величина X се нарича случайната величина* X^2, *стойностите на която са* $x_1^2, x_2^2, \dots x_k^2$ *със същите вероятности на X:*$p_1, p_2, \dots p_k$. *Трябва да имаме предвид, че* $X.X \neq X^2$ *и* $X.X \neq 2X$.

Някои от по - важните свойства да дисперсията са:

1. $DC = 0$ (C - неслучайна величина)

2. $D(X \pm Y) = DX + DY$
3. $D(kX) = k^2.DX$ (k - константа).

Началните моменти се дефинират като математическо очакване на целите степени на случайната величина X

$$m_k = \sum X^k = \sum x_i^k.p_i, \quad k \in \mathbb{N}.$$

Величината $(X - EX)$ се нарича отклонение. *Централните моменти* се дефинират като начални моменти на отклонението

$$\mu_k = E(X - EX)^k = \sum (x_i - EX)^k.p_i.$$

Всъщност, вторият централен момент е точно дисперсията на случайната величина.

Задачи:

1. Дадени са независимите случайни величини X и Y със следните закони на разпределение:

X	-2	0	1	2	3
P	$0,1$	$0,2$	$0,2$	$0,3$	$0,2$

и

Y	-3	-1	1
P	$0,3$	$0,1$	$0,6$

Да се определят:

а) функцията на разпределение $F(X)$;

б) законът и функцията на разпределение на случайната величина $Z_1 = 3.X - 1$;

в) законът и функцията на разпределение на величината $Z_2 = 2.X^2$;

г) да се начертаят графиките на функциите на разпределение на величините X, Z_1 и Z_2.

д) да се определят математическото очакване и дисперсията на случайните величини $X, Y, Z_3 = 5.X - 3.Y + 2$;

е) да се определят математическото очакване и дисперсията на случайната величина Z_2.

<u>Решение</u>: а) Функцията на разпределение на случайната величина X има вида:

при $-\infty < x \leq -1, \quad F(X) = 0,$
при $-2 < x \leq 0, \quad F(X) = 0,1$
при $0 < x \leq 1, \quad F(X) = 0,1 + 0,2 = 0,3,$
при $1 < x \leq 2, \quad F(X) = 0,3 + 0,2 = 0,5,$
при $2 < x \leq 3, \quad F(X) = 0,5 + 0,3 = 0,8,$
при $3 < x \leq \infty, \quad F(X) = 0,8 + 0,2 = 1;$

б) стойностите на случайната величина $Z_1 = 3X - 1$ се получават от стойностите на X, като се умножи всяка стойност на X с 3 и от получения резултат се извади единица, а вероятностите са същите като вероятностите на X. Така получаваме :

$Z_1 = 3X - 1$	-7	-1	2	5	8
P	$0,1$	$0,2$	$0,2$	$0,3$	$0,2$

.

От така получения закон пресмятаме функцията на разпределение:

$$F(X) = \begin{cases} 0 & x \leq -7, \\ 0,1 & -7 < x \leq -1, \\ 0,3 & -1 < x \leq 2, \\ 0,5 & 2 < x \leq 5, \\ 0,8 & 5 < x \leq 8, \\ 1 & x > 8; \end{cases}$$

в) аналогично на б) получаваме

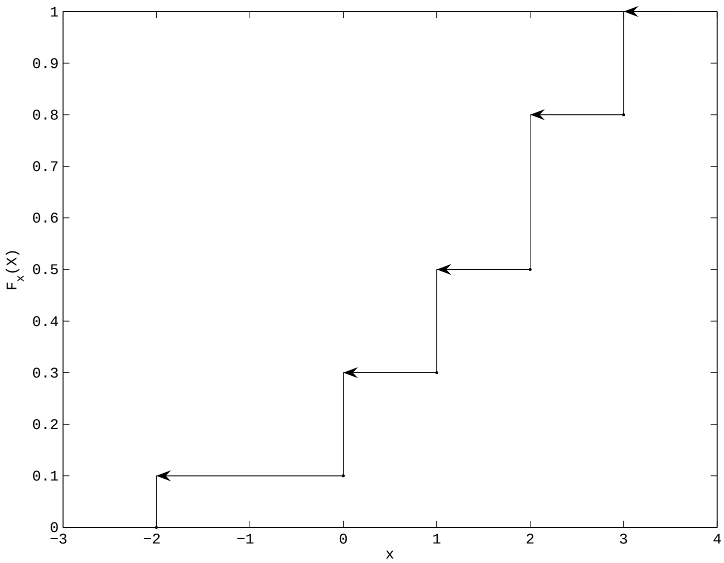
$Z_2 = 2X^2$	8	0	2	8	18
P	$0,1$	$0,2$	$0,2$	$0,3$	$0,2$

Вижда се, че стойността на $Z_2 = 8$ се приема два пъти и затова ще преработим таблицата като съберем двете вероятности за $Z_2 = 8$.

$Z_2 = 2X^2$	0	2	8	18
P	$0,2$	$0,2$	$0,5$	$0,2$

;

г) Графиката на функцията на разпределение има вида:



Фиг.1

д)

$$EX = 2.0,1 + 0.0,2 + 1.0,2 + 2.0,1 + 3.0,2 = 1,2,$$

$$DX = 2^2.0,1 + 0^2.0,2 + 1^2.0,2 + 2^2.0,3 + 3^2.0,2 - (1,2)^2 = 2,16,$$

$$EY = 3.0,3 + (-1).0,1 + 1.0,6 = -0,4,$$

$$DY = (-3)^2.0,3 + (-1)^2.0,1 + 1^2.0,6 - (-0,4)^2 = 3,24,$$

$$EZ_3 = E(5X - 3Y + 2) = 5EX - 3EY + 2,$$

$$EZ_3 = 5.1,2 - 3.(-0,4) + 2 = 9,2,$$

$$DZ_3 = D(5X - 3Y + 2) = 25DX + 9DY + 0 = 83,16.$$

е) използвайки в), получаваме:

$$E(Z_3) = E(2X^2) = 0.0,2 + 2.0,2 + 8.0,4 + 18.0,2 = 7,2.$$

2. Случайната величина е подчинена на закона:

X	1	3	4	6	7
P	$0,1$	$0,1$	$0,3$	$0,4$	$0,1$

Да се намери функцията на разпределение и да се построи графиката й. Да се изчислят EX и DX. (Отг. $EX = 4,7,\ DX = 3,41$)

3. В урна има 4 топки с номера 1, 2, 3 и 4. Извадени са по случаен начин две топки. Да се намери законът на разпределение на случайната величина X, равна на сумата от номерата на извадените топки.
Решение: Възможните стойности на случайната величина са 3, 4, 5, 6 и 7. X приема стойност 7, ако сме изтеглили топки с номера 1 и 2. Затова $P(X = 3) = \frac{1}{6}$. За всички останали случаи, освен за $P(X = 5)$ вероятността е същата. За $P(X = 5)$ имаме два случая - топки с номера 1, 4 и топки с номера 2 и 3. Затова $P(X = 5) = \frac{2}{6}$. Тогава законът за разпределение е

X	3	4	5	6	7
P	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

4. Стрелец, който има 4 патрона, стреля по цел докато улучи или докато изразходва всичките си патрони. Вероятността за улучване на целта е една и съща при всеки изстрел и е равна на 0,3. Да се определят: законът на разпределение, математическото очакване и дисперсията на случайната величина X - брой направени изстрели. (Отг. $EX = 2,533,\ \ DX = 1,535$)

5. Продавач на сладолед може да направи дневно оборот от 60 лв. при слънчево време и 20 лв. при дъждовен ден. Съставете таблица на разпределението на дневния оборот X, ако вероятността за слънчев ден е 0,65. Намерете EX+DX. (Отг. 410)

6. Дискретната случайна величина X приема само три стойности: $x_1 = 2$ с вероятност 0,3, x_2 с вероятност p_2 и x_3 с вероятност $0,2$. Да се определят x_2 и p_2, ако е известно, че $EX = 3,6$. (Отг. $x_2 = 4,\ \ p_2 = 0,5$)

7. Броят на счупените буркани в един пакет при транспортирането им с автомобил е случайна величина, която има разпределение

X	0	1	2	3
P	$0,5$	$0,3$	$0,1$	$0,1$

При транспортирането с влак същата случайна величина има разпределение

Y	0	1	2
P	$0,8$	$0,1$	$0,1$

Посочете верните твърдения:
а) $EX = EY$,
б)$E(10X + 10Y) = 11$,
в) $EX > EY$,
г) $EX < EY$,
д) $DX > DY$,
е) $D(X + Y) = 1,37$,
ж) $D(X - Y) = 1,37$. (0тг. б, в, д, е, ж)

8. В партида от 10 изделия има едно нестандартно. За да го отстранят проверяват изделията едно след друго по случаен начин. Да се намери законът на разпределението, математическото очакване и дисперсията на случайната величина X - броя на проверените изделия. (Отг. EX=5,5, DX=8,25)

3.2 Непрекъснати случайни величини и техните характеристики

Ръстът на жените над 15 годишна възраст, времето за безотказна работа на даден уред, времето за обслужване на клиенти в магазин са примери на непрекъснати случайни величини. Стойностите, които една непрекъсната случайна величина приема, плътно изпълват даден интервал. Най - често те са резултат на измерване. За количествена характеристика на

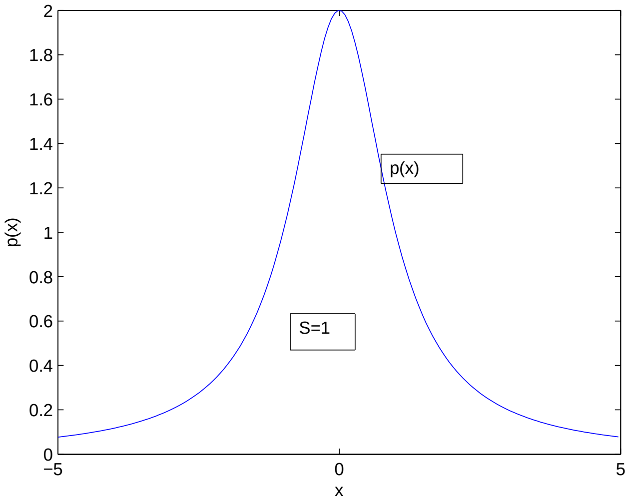
разпределението на случайната величина X се използва вероятността на събитието $(X < x)$, т.е. функцията на разпределение (интегрална функция) $F(X) = P(X < x)$. $F(X)$ притежава следните свойства:
1. $F(X)$ е ненамаляваща функция.
2. $F(X)$ е непрекъсната отляво, т.е. $\lim_{x\to x_0-} F(x) = F(x_0)$.
3. $\lim_{x\to -\infty} F(x) = 0, \quad \lim_{x\to\infty} F(x) = 1$.
Всяка функция, която има горните свойства се явява функция на разпределение на някаква случайна величина X. Диференциалната функция $p(x)$ или още плътност на вероятностите на непрекъснатата случайна величина X се нарича производната на $F(x)$.

$$p(x) = F'(x) = \lim_{\Delta x\to 0} \frac{F(x+\Delta x) - F(x)}{\Delta x}.$$

Някои от по-важните свойства на плътността са:
1. $p(x) \geq 0$, защото $p(x)$ е производна на растяща функция;
2. Вероятностите случайната величина X да има стойности в интервала $[a,b]$ е равна на определения интеграл от плътността на разпределение в този интервал, т.е.

$$P(a < X < b) = F(b) - F(a) = \int_a^b p(x)dx$$

3. $\int_{-\infty}^{\infty} p(x)dx = 1$. От това свойство можем да направим извод, че лицето на криволинейния трапец, ограничено от графиката на плътността $p(x)$ и оста Ox, е равно на единица



Фиг.2

4. Ако е известна плътността $p(x)$ на една непрекъсната случайна величина, то интегралната функция $F(x)$ се намира от $F(x) = \int_{-\infty}^{\infty} p(x)dx$.

Математическото очакване и дисперсия се дефинират както следва:

$$EX = \int_{-\infty}^{\infty} x.p(x)dx,$$

$$DX = \sigma^2 = \int_{-\infty}^{\infty} (x - EX)^2.p(x)dx = \int_{-\infty}^{\infty} x^2.p(x)dx - (EX)^2.$$

Числото M_e се нарича медиана, ако $F(M_e) = \frac{1}{2}$. Мода се нарича числото M_o, в което $p(x)$ има локален максимум. Началните и централните моменти се изчисляват по формулите:

$$m_k = \int_{-\infty}^{+\infty} x^k.p(x)dx, \;\; \mu_k = \int_{-\infty}^{+\infty} (x - EX)^k.p(x)dx.$$

Задачи:

1. Една непрекъсната случайна величина има плътност на вероятностите:

$$p(x) = \begin{cases} C(4 - 2x) & x \in (0, 2) \\ 0, & x \notin (0, 2) \end{cases}$$

Да се намери константата С и интегралната функция на разпределение $F(x)$.

<u>Решение</u>: За намиране на С използваме свойство 3: $\int_{-\infty}^{\infty} p(x)dx = 1$. Интервалът $(-\infty, +\infty)$ се разделя от точките 0 и 2 на три подинтервала - $(-\infty; 0), (0, 2), (0, +\infty)$. Тогава $\int_{-\infty}^{+\infty} p(x)dx = \int_{-\infty}^{0} 0dx + \int_{0}^{2} C(4-2x)dx + \int_{2}^{+\infty} 0dx = C\int_{0}^{2}(4-2xdx = C\left(\int_{0}^{2} 4dx - \int_{0}^{2} 2xdx\right) = C(8-4) = 4C$. Понеже $\int_{-\infty}^{\infty} p(x)dx = 1$, то от равенството $4C = 1$ следва, че $C = \frac{1}{4}$. Тогава

$$p(x) = \begin{cases} \frac{1}{4}(4-2x) & x \in (0,2) \\ 0, & x \notin (0,2) \end{cases}$$

За намиране на интегралната функция $F(x)$ използваме свойство 4.

а) Нека $x \in (-\infty; 0)$.Тогава

$$F(x) = \int_{-\infty}^{x} p(x)dx = \int_{-\infty} x0dx = 0.$$

б) Нека $x \in (0; 2)$.Тогава

$$F(x) = \int_{-\infty}^{x} p(x)dx = \int_{-\infty} x0dx + \int_{0}^{x} \frac{1}{4}(4-2x)dx = x - \frac{x^2}{4}.$$

в) Нека $x \in (2; +\infty)$.Тогава

$$F(x) = \int_{-\infty}^{x} p(x)dx = \int_{-\infty} x0dx + \int_{0}^{2} \frac{1}{4}(4-2x)dx + \int_{2}^{x} 0dx = 1.$$

Окончателно за $F(x)$ получаваме

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

2. Нека А е положителна константа. Диференциалната функция на разпределение на случайната величина X е:

$$p(x) = \begin{cases} 0, & -\infty < x < -A, \\ A^2 - x^2, & -A \le x \le 0, \\ (x-A)^2, & 0 < x \le A \\ 0, & x > A. \end{cases}$$

Да се определят: A, F(x), EX, DX, $P(-\frac{1}{2} < x < \frac{1}{2})$, $P(-1 < x < 0)$, $P(0 < x < 1)$.

<u>Решение</u>: За намирането на коефицента A отново ще използваме, че $\int_{-\infty}^{\infty} f(x)dx = 1$.
$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{-A} 0dx + \int_{-A}^{0} (A^2 - x^2)dx + \int_{0}^{A} (x - A)^2 dx + \int_{A}^{\infty} 0dx = 0 + A^2 \int_{-\infty}^{0} dx - \int_{-A}^{0} x^2 dx + \int_{0}^{A} (x_A)^2 d(x - A) + 0 = A^2 x|_{-A}^{0} - \frac{x^3}{3}{}_{-A}^{0} + \frac{(x-A)}{3}|_{0}^{A} = 0 + A^3 - 0 - \frac{A^3}{3} + 0 + \frac{A^3}{3} = 1$. Следователно $A^3 = 1$ и $A = 1$. При тази стойност на A $p(x)$ има вида

$$p(x) = \begin{cases} 0, & -\infty < x < -1, \\ 1 - x^2, & -1 \leq x \leq 0, \\ (x-1)^2, & 0 < x \leq 1 \\ 0, & x > 1. \end{cases}$$

За интегралната функция $F(x)$получаваме:
При $-\infty < x < -1, \quad F(x) = \int_{-\infty}^{x} 0dt = 0.$
При $-1 \leq x \leq 0, \quad F(x) = \int_{-\infty}^{-1} 0dt + \int_{-1}^{x} (1 - t^2)dt = \frac{-x^3+3x+2}{3}$
При $0 < x \leq 1, \quad F(x) = \int_{-\infty}^{-1} 0dt + \int_{-1}^{0} (1 - t^2)dt + \int_{0}^{x} (t-1)^2 dt = \frac{(x-1)^3}{3} + 1.$

При $x > 1, \quad F(x) = \int_{-\infty}^{-1} 0dt + \int_{-1}^{0} (1 - t^2)dt + \int_{0}^{1} (t-1)^2 dt + \int_{1}^{x} 0dt = 1.$
Следователно интегралната функция има вида

$$F(x) = \begin{cases} 0, & -\infty < x < -1, \\ \frac{-x^3+3x+2}{3}, & -1 \leq x \leq 0, \\ \frac{(x-1)^3}{3} + 1, & 0 < x \leq 1 \\ 1, & x > 1. \end{cases}$$

За намирането на математическото очакване представяме интервала чрез сума от четири подинтервала, защото плътността се задава с различни изрази за различните интервали.
$EX = \int_{-\infty}^{-1} 0dx + \int_{-1}^{0} x(1 - x^2)dx + \int_{0}^{1} x(x-1)^2 dx + \int -1^{\infty} 0dx = 0 + \int_{-1}^{0} (x - x^3)dx + \int_{0}^{1} (x^3 - 2x^2 + x)dx + 0 = -\frac{1}{6}.$
Аналогично постъпваме и при изчисляването на дисперсията

$DX = \int_{-\infty}^{\infty} x^2 p(x)dx - (EX)^2 = \frac{5}{36}$.
За изчисляване на търсените вероятности използваме стойностите на интегралната функция като избираме съответната формула според аргумента:

$$P(-\frac{1}{2} < x < \frac{1}{2}) = F(\frac{1}{2}) - F(-\frac{1}{2}) = \frac{3}{4},$$
$$P(-1 < x < 0) = F(0) - F(-1) = \frac{2}{3},$$
$$P(0 < x < 1) = F(1) - F(0) = \frac{1}{3}.$$

3. Да се определи диференциалната функция на разпределение $p(x)$, ако интегралната функция на разпределение на непрекъснатата случайна величина X е:

$$F(x) = \begin{cases} 0, & x \leq 0, \\ \frac{x^2}{1+x^2}, & x > 0. \end{cases}$$

Решение: При $x \leq 0, \;\; p(x) = 0$.
При $x > 0, \;\; p(x) = (\frac{x^2}{1+x^2})' = \frac{2x}{(1+x^2)^2}$.Тогава

$$p(x) = \begin{cases} 0, & x \leq 0, \\ \frac{2x}{(1+x^2)^2}, & x > 0. \end{cases}$$

4. Случайната величина X има плътност:
а) $p(x) = \begin{cases} 0, & x < 0, \\ Acos(x), & 0 \leq x \leq \frac{\pi}{2}, \\ 0, & x > \frac{\pi}{2}; \end{cases}$
б) $p(x) = \begin{cases} 0, & x \leq 0, \\ A.e^{-x}, & x > 0, \end{cases}$
Да се определят коефицентът А и интегралната функция на разпределение.

Отг. а) А=1, $F(x) = \begin{cases} 0, & x \leq 0, \\ \sin(x), & 0 \leq x \leq \frac{\pi}{2}, \\ 1, & x > \frac{\pi}{2}; \end{cases}$

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

5. Да се определят A, EX и DX за непрекъснатата случайна величина X, имаща плътност на разпределение:

а) $p(x) = \begin{cases} \frac{Ax}{1+x^2}, & 0 \leq x \leq \sqrt{e^2-1}, \\ 0, & x \notin (0, \sqrt{e^2-1}); \end{cases}$ (Отг. A=1, EX=1,34, DX=0,399)

б) $p(x) = \begin{cases} A + 3x^2, & 0 \leq x \leq 1, \\ 0, & x \notin (0,1); \end{cases}$ (Отг. A=0, EX=0,75, DX=0,0375)

в) $p(x) = \begin{cases} 0, & x < -1, \\ (x-1)^2, & -1 \leq x \leq 0, \\ A(1-x), & 0 \leq x \leq 1, \\ 0, & x > 1; \end{cases}$ (Отг. A=6/5, EX=0,1, DX=0,13)

6. Случайната величина X е зададена с функция на разпределение:

$F(x) = \begin{cases} 0, & x \leq 1, \\ x - 1, & 1 < x \leq 2, \\ 1, & x > 2 \end{cases}$ Изчислете $P(0,5 < X < 1,5)$. (Отг. 0,5)

Глава 4

Основни закони на разпределение на случайни величини

4.1 Дискретни разпределения

Ще разгледаме две основни дискретни разпределения: биномно и Поасоново.

В практиката много често се налага един опит да се повтаря много пъти, без да се нарушава комплексът от условия. Нека са проведени n еднакви, независими опита по схемата на Бернули. Нека вероятността за успех е p и за неуспех *q=1-p*.

Дискретната случайна величина има *биномно разпределение*, ако приема стойности $0, 1, 2, \ldots, n$ със съответни вероятности $P(X = k) = C_n^k p^k q^{n-k}$ и

$$EX = np, \;\; DX = npq, \;\; p + q = 1, \;\; p \geq 0, \;\; q \geq 0.$$

Ако броят на опитите е много голям, а вероятността за появяване на събитие А е малко число, така, че np е число от порядъка на няколко единици, намирането на вероятността $P(X = k)$ - при n опита събитието A да се сбъдне k пъти, се изчислява по формулата на Поасон: $P(X = k) = \frac{\lambda^k.e^{-\lambda}}{k!}$. Дискретната случайна величина има Поасоново разпределение, ако приема стойности $0, 1, \ldots$ като $P(X = k) = \frac{\lambda^k.e^{-\lambda}}{k!}$. Може да се докаже, че $EX = \lambda$ и $DX = \lambda$. Фактът, че математическото очакване и дисперсията при *Поасоновото разпределение* са равни, се използва в статистиката за доказване, че дадена случайна величина има такова разпределение.

Задачи:

1. Вероятността за попадение в цел при един изстрел е 0,4. Проведени са 5 изстрела. Съставете таблица на разпределението на броя на попаденията в целта. Намерете математическото очакване и дисперсията на случайната величина.

Решение: Случайната величина има биномно разпределение. Затова ще пресметнем вероятностите с формулата на Бернули. Понеже са проведени 5 изстрела, възможните стойности на случайната величина са 0, 1, 2, 3, 4 и 5. Съответните вероятности са:

$P(X = 0) = C_5^0(0,4)^0(0,6)^5 = \frac{243}{3125}$,
$P(X = 1) = C_5^1(0,4)^1(0,6)^4 = \frac{810}{3125}$,
$P(X = 2) = C_5^2(0,4)^2(0,6)^3 = \frac{1080}{3125}$,
$P(X = 3) = C_5^3(0,4)^3(0,6)^2 = \frac{720}{3125}$,
$P(X = 4) = C_5^4(0,4)^4(0,6)^1 = \frac{240}{3125}$,
$P(X = 5) = C_5^5(0,4)^5(0,6)^0 = \frac{32}{3125}$.

Тогава таблицата на разпределение е

k	0	1	2	3	4	5
$P_5(k)$	$\frac{243}{3125}$	$\frac{810}{3125}$	$\frac{1080}{3125}$	$\frac{720}{3125}$	$\frac{240}{3125}$	$\frac{32}{3125}$

За биномното разпределение $EK = np = 5.0,4 = 2$ и $DX = npq = 2.0,4.0,6 = 1,2$.

2. В магазин са доставени 2000 изделия. Преди транспортирането всички изделия са били качествени. Вероятността за повреда на отделно изделие при транспортиране е 0,001. Да се намери вероятността при

транспортиране да се повредят точно 3 изделия.
Решение: Тъй като броят на изделията (опити) е голям, а вероятността за повреда е малка, това разпределение е Поасоново. Тогава $\lambda = 2000.0,001 = 2.$ и $P = \frac{2^3.e^{-2}}{3!} \approx 0,18$.

3. Каква е вероятността случайна величина с Поасоново разпределение да приеме стойност не по-малка от 1, ако са направени 1000 независими опита, а вероятността при всеки опит за сбъдване на дадено събитие е $p = 0,002$? (Отг. $1 - \frac{1}{e^2}$)

4. В съд има 2 бели и 6 черни топки. По случаен начин вадим 3 пъти по една топка като всеки път връщаме изваденатa топка. Да се напише разпределението на броя на извадените бели топки. Да се намери математическото очакване и дисперсията.

5. Да се намери дисперсията на дискретна случайна величина X-броя на появяването на събитието A в два независими опита, ако вероятностите за поява на събитието са равни и EX=0,9. (Отг. 0,495)

6. Потокът от кораби, които пристигат в дадено пристанище, е разпределен по закона на Поасон със средна стойност 0,5 кораба в час. Намерете вероятността за едно денонощие да пристигнат 15 кораба. (Отг.$\approx 0,07$)

7. Коректура съдържа средно 500 печатни грешки на 1000 страници. Намерете вероятността на дадена страница да има не по-малко от 2 печатни грешки.
Решение: Ще извършим изчисленията с помощта на закона на Поасон. Знаем, че $EX = \lambda = 0.5$. Ще пресметнем търсената вероятност с помощта на вероятността на обратното събитие (да има 0 или 1 грашка). Така $P(k \leq 2) = 1 - P(k = 0) - P(k = 1) = 1 - (0,5)^0.e^{-0,5} - (0,5)^1.e^{-0,5} = 0,090.$

4.2 Непрекъснати разпределения

Непрекъснатата случайна величина X има *равномерно разпределение* в интервала $[a, b]$, ако плътността има вида :

$$p(x) = \begin{cases} 0, & x \notin [a, b], \\ \frac{1}{b-a}, & x \in [a, b]. \end{cases}$$

Интегралната функция има вида

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

$$EX = \frac{a+b}{2}, \ \ DX = \frac{(b-a)^2}{12}.$$

Пример 4.1. *Ако непрекъснатата случайна величина е равномерно разпределена в интервала* $[3, 5]$*, използвайки горните резултати, получаваме* $p(x) = \begin{cases} 0, & x < 3, \\ \frac{1}{2}, & 3 \leq x \leq 5, \\ 0, & x > 5; \end{cases}$ $F(x) = \begin{cases} 0, & x < 3, \\ \frac{x-3}{2}, & 3 \leq x \leq 5, \\ 1, & x > 5; \end{cases}$
$EX = 4, \ \ DX = \frac{1}{3}$.

Ако плътността на разпределение на една непрекъсната случайна величина има вида $p(x) = \begin{cases} 0, & x < 0, \\ \lambda e^{-\lambda x}, & x \geq 0, \ \ \lambda > 0, \end{cases}$ казваме, че X има *показателно разпределение.*

Времето между постъпване на два последователни кораба в едно пристанище, времето на обслужване на клиент в магазин или времетраенето на безотказна работа на някакво изделие имат разпределения, близки до показателното.

Интегралната функция на показателното разпределение е

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

Може да се покаже, че $EX = \frac{1}{\lambda}, \ \ DX = \frac{1}{\lambda^2}$.

Пример 4.2. *Времето T между две аварийни спирания на автоматична линия е разпределено по показтелния закон с математическо* *очакване 4 денонощия. Намерете вероятността за 5 денонощия да не настъпи нито едно аварийно спиране на линията.*

Решение: От равенството $ET = \frac{1}{\lambda} = 4$ следва, че $\lambda = \frac{1}{4}$. Тогава интегралната функция на разпределение е $F(t) = 1 - e^{-\frac{1}{4}t}$. Търсената вероятност е $P(T > 5) = e^{-5/4}$.

4.3 Нормално разпределение

Ще казваме, че непрекъснатата случайна величина X има нормално разпределение, ако има плътност на вероятностите

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}, \ \ -\infty < x < \infty, \ \ a > 0, \ \ \sigma > 0.$$

Нормалното разпределение има следните числови характеристики:

$$EX = a, \ \ DX = \sigma^2.$$

В математиката функцията $\Phi(x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$ се нарича функция на Лаплас. Нейните стойности са зададени в таблица. Чрез тази функция се изразява интегралната функция $F(x)$ на нормалното разпределение

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

В практическите приложения трябва да се има предвид, че функцията на Лаплас е нечетна, т.е. $\Phi(-t) = -\Phi(t)$.

Пример 4.3. *Непрекъснатата случайна величина има нормално разпределение и $EX = a = 3$ и $\sigma = 2$. Да се определи вероятността $P(1 < X < 5)$, т.е. вероятността случайната величина да попадне в интервала $(1, 5)$.*

Решение: $P(1 < X < 5) = F(5) - F(1) = \frac{1}{2} + \frac{1}{2}\Phi\left(\frac{x-a}{\sigma}\right) = \frac{1}{2} + \frac{1}{2}\Phi(1) - \frac{1}{2} - \frac{1}{2}\Phi(-1) = \Phi(1)$. От таблицата за стойностите на $\Phi(t)$ получаваме $\Phi(t) = 0,68$. Тогава $P(1 < X < 5) = 0,68$.

Пример 4.4. *(Правило на 3-те сигми). Нека непрекъснатата случайна величина има нормално разпределение с параметри a і σ. Да се покаже, че повече от 99% от стойностите на случайната величина попадат в интервала $(a-3\sigma; a+3\sigma)$.*

Решение: При нормално разпределение

$$P(x_1 < X < x_2) = \frac{1}{2}[\Phi(\frac{x_1 - a}{\sigma}) - \Phi(\frac{x_2 - a}{\sigma})].$$

За нашия пример $x_1 = a - 3\sigma$ и $x_2 = a + 3\sigma$. Тогава $P(a - 3\sigma < X < a+3\sigma) = \frac{1}{2}[\Phi(3) - \Phi(-3)] = \Phi(3)$. От таблицата намираме $\Phi(3) = 0,99721$. Правилото указва, че с вероятност близка до единица, стойностите на нормално разпределена случайна величина се намират в интервала $(a - 3\sigma, a + 3\sigma)$. В икономиката и техниката много величини се явяват случайни величини с нормален закон на разпределение. Това се обяснява от факта, че тези величини се получават в резултат на сумиране на много случайни величини и съгласно централната гранична теорема имат закон на разпределение, близък до нормалния.

Задачи:

8. Случайната величина X има нормално разпределение с $EX = a$ и $DX = \sigma^2$. Да се определи $P(|x - a| < \epsilon)$.

Решение:

$$P(|x-a| < \epsilon) = P(-\epsilon < x - a < \epsilon) = P(-\epsilon + a < x < \epsilon + a) =$$

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

9. Автомат пакетира ориз. Теглото на пакетите е случайна величина, разпределена по нормален закон с $EX = 1000$gr. и $\sigma = 2,04$ gr.Да се определи в какви граници се гарантира теглото на пакетите с вероятност 0,95.

Решение: По условие $P(1000 - \epsilon < X < 1000 + \epsilon) = 0,95$. Споед задача (8), $0,95 = \Phi(\frac{\epsilon}{2,04})$. Тогава $\frac{\epsilon}{2,04} = 1,96$. Така получаваме, че $\epsilon \approx 4$. Следователно гарантираме с вероятност 0,95, че теглото на пакетите ще бъде в граници от 996 до 1004гр.

10. Една компания произвежда електрически крушки. Известно е, че продължителността на живота на крушките е нормално разпределена случайна величина със средна 800 часа и стандартно отклонение 40 часа. Ако случайно от една партида се изберат 100 крушки, какъв ще бъде относителният дял на крушките с очаквана продължителност на живот между 778 и 834 часа?
<u>Решение</u>: Нека продължителността на живот на крушките е означена с X. Вероятността за произволно избрана крушка X да е между 778 и 834 часа е $P(778 < X < 834) = F(834) - F(778) = \frac{1}{2} + \frac{1}{2}\Phi\left(\frac{834-800}{40}\right) - \frac{1}{2} - \frac{1}{2}\Phi\left(\frac{778-800}{40}\right) = \frac{1}{2}[\Phi(0,85) + \Phi(0,55)] = 0,511$.Следователно, 51,11% от произведените крушки ще светят между 778 и 834 часа.

11. Истинското тегло на 10 - килограмови пакети с брашно, които се опаковат машинно е нормално разпределена величина с очакване 10 кг. и дисперсия 0,1 кг. Каква е вероятността един случайно избран пакет да тежи не по-малко от 9кг. и 875 гр.?
<u>Решение</u>: Нека X е истинското тегло на пакетите. Търсим вероятността $P(X \geq 9,875) = 1 - P(X < 9,875) = 1 - F(9,875) = 1 - [\frac{1}{2} + \frac{1}{2}\Phi(\frac{9,875-10}{0,1})] = \frac{1}{2}(1 + \Phi(1,25)) = 0,894$.

12. Средният резултат от тест по математика е 82 точки, а стандартното отклонение е 5 точки. Комисията поставя отлична оценка на 12% от най-високите резултати. Разпределението на резултатите е нормално. Каква ще е долната граница на резултата, за който ще се постави отлична оценка? (Отг. 88т.)

13. Нека окачествяването на яйцата се извършва по следния начин: ако яйцето мине през отвор с диаметър 35мм, то се счита за маломерно и в противен случай нормално. Диаметърът на яйцето е случайна величина, разпределена по нормален закон с математическо очакване 43мм и стандартно отклонение 8,25мм. Да се определи процентът на нормалните яйца. (Отг. 83,4%)

14. Матетамическото очакване и стандартното отклонение на нормално разпределена случайна величина X са съответно 10 и 2. Намерете

вероятността в резултат на опит X да приеме стойност от интервала $(12, 14)$. (Отг. 0,136)

15. Провежда се измерване на диаметър на вал без систематични грешки. Случайните грешки на измерването X са подчинени на нормален закон със стандартно отклонение $\sigma = 10$мм. Намерете вероятността измерването да се проведе с грешка, не по-голяма по абсолятна стойност от 15мм.

Решение: Математическото очакване на случайните грешки е нула. Следователно търсим $P(|x| < \epsilon)$. Според зад. (8), $P(|x| < \epsilon) = \Phi(\frac{\epsilon}{\sigma})$. В случая имаме $\epsilon = 15$ и $\sigma = 10$.Следователно $P(x < 15) = \Phi(1, 4) = 0, 8664$.

16. Детайл се произвежда от автомат и се счита годен, ако отклонението на размера от стандарта не превишава 10мм. Случайните отклонения от размера е нормално разпределена случайна величина със $\sigma = 5$мм и $a = 0$. Колко процента годни детайла произвежда автоматът? (Отг. 95%)

17. Случайната величина X е разпределена по нормален закон със $\sigma = 5$мм. Намерете дължината на интервала, симетричен относно математическото очакване, в който с вероятност 0,9973 ще попадне случайната величина. (Отг. 30мм.)

Глава 5

Закон за големите числа

Приемането на определена стойност на една случайна величина при даден опит зависи от много случайни причини, отчитането на които практически е невъзможно. Отдавна е било забелязано, че средното аритметично на голям брой еднородни случайни величини се колебае незначително около средната стойност на техните математически очаквания, когато случайните величини отговарят на някои условия. Тези условия съставят най-важното съдържание на закона за големите числа. Към този закон се отнасят теоремите на Бернули, Поасон, Чебишев, Марков, Колмогоров и др. Нека $X_1, X_2, \ldots, X_n$ са еднакво разпределени независими случайни величини, всяка от които има крайно математическо очакване, равно на a. В този случай, както бе посочено, се твърди, че

$$\frac{X_1 + X_2 + \ldots + X_n}{n} \to a,$$

при $n \to \infty$.

Нека да разгледаме следния пример. Да допуснем, че при едни и същи условия се правят без систематична грешка измервания на една величина. Случайните грешки, допуснати в отделните измервания, са независими и поради постоянните условия, еднакво разпределени. Въз основа на закона за големите числа, средното аритметично на резултатите от измерванията с голяма вероятност ще бъде произолно близко до измерваната величина. Естествено е да се постави въпросът за размерите на възможните отклонения на средното аритметично от тази постоянна величина и за закона на разпределение на тези отклонения. Отговор на този въпрос дава

Теорема 5.1. *(Централна гранична теорема) Ако случайните величини $X_1, X_2, \ldots, X_n$ са еднакво разпределени и имат крайна дисперсия, при $n \to \infty$ е всила следното твърдение:*

$$P\left\{\frac{\sum_{k=1}^{n} X_k - na}{\sigma\sqrt{n}} < x\right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt,$$

където $a = EX_k$ и $\sigma^2 = DX_k$ за $k = 1, 2, \ldots, n$.

По- нататък ще разгледаме някои от групата теореми, които установяват и обясняват устойчивостта на средните стойности на голям клас случайни величини и се обединяват под името **Закон за големите числа.**

Теорема 5.2. *(Неравенство на Чебишев) Ако случайната величина X има крайна дисперсия и $\epsilon > 0$ е произволно число, то в сила е неравенството*

$$P(|X - EX| < \epsilon) \geq 1 - \frac{DX}{\epsilon^2}.$$

Неравенството оценява вероятността за отклоняване на случайната величина от математическото й очакване. Трябва да се има предвид, че тези оценки са доста груби. Главното приложение на неравенството е, че с негова помощ се доказва теорема, наричана закон за големите числа. Смисълът на теоремата е, че да се предскаже стойността, която ще приеме една от много случайни величини е невъзможно, но средно аритметичното от стойностите им е относително постоянно.

Теорема 5.3. *(Теорема на Чебишев) Ако случайните величини $X_1, X_2, \ldots, X_n, \ldots$ са независими помежду си и дисперсиите им са ограничени от едно и също число C, т.е. $DX_1 \leq C, \ldots, DX_n \leq C$, то за всяко $\epsilon > 0$*

$$P\left(\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \frac{EX_1 + EX_2 + \ldots + EX_n}{n}\right| < \epsilon\right) \geq 1 - \frac{C}{n\epsilon^2}.$$

Оттук следва, че

$$\lim_{n\to\infty} P\left(\left|\frac{X_1+X_2+\ldots+X_n}{n}-\frac{EX_1+EX_2+\ldots+EX_n}{n}\right|<\epsilon\right)=1.$$

Задачи:

1. Вероятността за настъпване на събитието А във всеки от 1300 опита е 0,3. Нека X е броят на настъпване на събитието. Използвайки неравенството на Чебишев, оценете вероятността за това, отклонението на X от математическото очакване да бъде по-голямо от 30.

Решение: Величината X има биномно разпределение с $EX = 1300.0,3 = 390$ и дисперсия $DX = 1300.0,3.0,7 = 273$. Прилагаме неравенството за $\epsilon = 30 : P(|X-390| < 30) \geq 1 - \frac{273}{30^2} = 0,7$.

2. Вероятността за настъпване на събитието А при всеки опит е 0,25. Оценете вероятността, броят на появяванията на събитието А да бъде в граници от 150 до 250, ако са извършени 800 независими опита. (Отг. $P \geq 0,94$)

3. Математическото очакване на скоростта на вятъра на дадена височина е 0,25км/час, а стандартното отклонение е 4,5км/час. Какви скорости на вятъра, с вероятност по-голяма от 0,9, можем да очакваме на тази височина? (Отг. между 10,8 и 39,2)

4. Приложима ли е теоремата на Чебишев към редицата от независими случайни величини $\{X_n\}$, ако :

X_n	$-n^2$	0	n^2
$P(X_n)$	2^{-n}	$1-2^{1-n}$	2^{-n}

Решение: $EX_n = 0,\ DX_n = n^4.2^{1-n}$. Временно ще предполагаме, че n се изменя непрекъснато (т.е. ще означим n с x) и търсим максимума на функцията $\phi(x) = \frac{x^4}{2^{x-1}}$. Критичните точки на първата производна са: $x_1 = 0$ и $x_2 = \frac{4}{ln(2)} \approx 5,8$. Тъй като n приема само цели положителни стойности, не използваме x_1. В точката x_2 функцията има максимум. Тогава $DX_n = \frac{n^4}{2^{n-1}}$. При $n = 5$ имаме $DX_5 = \frac{625}{16}$, а при $n = 6$, дисперсията е $DX_6 = 40,5$. Следователно възможно най-голямата дисперсия е 40,5, т.е. дисперсията на случайните величини X_n е равномерно ограничена от 40,5. Тъй като предположенията, съдържащи се в теоремата на Чебишев са изпълнени, към разглежданата редица от независими случайни величини е приложима теоремата на Чебишев.

5. Да се изясни дали е приложима теоремата на Чебишев, за редицата от независими случайни величини $X_1, X_2, \ldots, X_n, \ldots$, ako X_n има следното разпределение :

X_n	$-5n$	0	$5n$
$P(X_n)$	$1/3n^2$	$1-2/3n^2$	$1/3n^2$

6. Колко независими измервания на дадена величина трябва да бъдат направени, така че с вероятност, не по - малка от 0,98, да можем да твърдим, че средното аритметично на резултатите от измерванията се отличава от истинската стойност по модул с по-малко от 0,01, ако дисперсията на отделно измерване не надминава 1?
Решение: Нека x_i е резултатът от $i-$ тото измерване. Да означим с а истинската стойност на неслучайната величина - средната величина. Тогава $EX_i = a$, а $DX_i \leq 1$. Трябва да се намери онази стойност на n, за която

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} -a\right| < 0,01\right) \geq 0,98.$$

Последното неравенство ще бъде изпълнено, ако $1 - \frac{C}{n\epsilon^2} = 1 - \frac{1}{n.0{,}01^2} \geq 0,98$. Тогава $n \geq 500000$.

7. Колко пъти трябва да се измери дадена величина, истинската стойност на която е равна на а, така че с вероятност не по - малка от 0,95 да можем да твърдим, че средното аритметично на резултата от тези измервания ще се отличава от а по модул с по-малко от 2, ако стандартното отклонение на всяко измерване е по-малко от 10? (Отг. 500)

Глава 6

Елементи на математическата статистика

Статистиката е наука за методите на количествения анализ на масовите явления. За разлика от теорията на вероятностите, която на базата на една генерална съвкупност, изучава вида на разпределението на различни случайни величини и техните числови характеристики, математическата статистика решава обратната задача. Едни от основните задачи на статистиката са:

а) В резултат на n независими опита на случайна величина X имаме получени значения $x_1, x_2, \ldots, x_n$. Трябва да се определи макар и приближено неизвестната функция на разпределение на случайната величина.

б) Нека е известна функцията на разпределение $F(x)$ на някаква случайна величина. По резултатите от n независими опита : $x_1, x_2, \ldots, x_n$ трябва да се оцени видът на разпределение и точността на тази оценка.

в) На основание на някакви съображения се предлага хипотеза за вида на разпределение или за параметрите на разпределение на някаква случайна величина. Задачата е да се провери съвместими ли са резултатите от наблюдението с така издигнатата хипотеза.

6.1 Понятие за вариационен ред

Резултатите от наблюденията над един признак или от извършване на даден опит се нанасят по редове и колони, т.е. във формата на таблица. Например да предположим, че са получени следните резултати по даден предмет на 20 студенти:
3, 4, 2, 5, 4, 6, 6, 3, 5, 4, 5, 4, 4, 5, 3, 2, 4, 3, 5, 5. Успехът варира между числата 2, 3, 4, 5 и 6, които в статистиката се наричат варианти. Ако подредим тези данни по големина, ще получим следния ред:
2, 2, 3, 3, 3, 3, 4, 4, 4, 4, 4 ,4, 5, 5, 5, 5, 5, 5, 6, 6. Броят на равните стойности на признака се нарича честота и се бележи с f_i. В нашия пример $f_1 = 2$ е честота на оценката 2, $f_2 = 4$ е честота на оценката 3 и т.н. Оценките в нарастващ ред и съответните им честоти се нанасят в таблица, наречена вариационен ред.

Успех на студентите	2	3	4	5	6
Честота	2	4	6	6	2

В други случаи признакът, който се разглежда, може да приеме много повече различни стойности. В тези случаи използваме интервален вариационен ред.

Пример 6.1. *Нека разгледаме признака "Брой на селскостопанските ферми в даден регион, състоящ се от 60 селища". Получени са следните данни:*
3, 5, 7, 17, 31, 5, 7, 23, 21, 13, 15, 22, 6, 14, 20, 28, 11, 2, 25, 20, 15, 17, 4, 7, 8, 4, 13, 17, 23, 24, 5, 6, 23, 18, 10, 12, 24, 11, 5, 7, 23, 25, 3, 0, 3, 5, 17, 22, 24, 3, 15, 21, 2, 6, 18, 27, 30, 31, 5, 10.

За посочените данни групировката се извършва по следния начин. Определя се най-малката стойност на признака - $x_{min} = 0$ и най-голямата стойност - $x_{max} = 31$. Разликата $\delta = x_{max} - x_{min} = 31$ се нарича размах на признака. Размахът е интервал, в който попадат всички стойности на признака. Той се разделя на групи. Ширината на всяка група се намира като размахът се раздели на броя на интервалите. Избираме 8 интервала. Тогава ширината на всеки интервал е 4. Интервалният ред на разпределение на броя ферми е даден в следващата таблица.

брой ферми	$0 \div 3$	$4 \div 7$	$8 \div 11$	$12 \div 15$	$16 \div 19$	$20 \div 23$	$24 \div 27$	$28 \div 31$
Честота	7	15	5	7	6	10	6	4

Средната стойност на вариационен ред е важна негова характеристика. Тя показва числото, около което се групират стойностите на изучавания признак на данните от определена извадка. Средната стойност се изчислява по формулата:

$$\overline{X} = \frac{x_1 f_1 + x_2 f_2 + \ldots + x_k f_k}{f_1 + f_2 + \ldots + f_k} = \frac{\sum_{i=1}^{k} x_i f_i}{\sum_{i=1}^{k} f_i}.$$

Ако статистическият ред е интервален, вместо $x_1, x_2 \ldots, x_k$ заместваме със средите на интервалите.

Така средният успех на студентите е

$$\overline{X} = \frac{2.2 + 3.4 + 4.6 + 5.6 + 6.2}{2 + 4 + 6 + 6 + 2} = \frac{82}{20} = 4,1.$$

По същия начин средният брой на фермите е

$$\overline{X} = \frac{2.7 + 6.15 + 10.5 + 14.7 + 18.6 + 22.10 + 26.6 + 30.4}{7 + 15 + 5 + 7 + 6 + 10 + 6 + 4} = \frac{856}{60} \approx 14,27.$$

Мода се нарича числовата стойност на признака с най-голяма честота. Ако модата е една, разпределението на признака се нарича унимодално, при две моди - бимодално. При повече моди имаме мултимодално разпределение. Когато статистическото разпределение е дискретно, ане интервално, модата се определя непосредствено. При интервално статистическо разпределение може непосредствено да се установи само кой е модалния интервал. В този случай модата се пресмята по формулата

$$M_0 = x_0^{'} + \frac{(n_0 - n_{0-1})h}{(n_0 - n_{0-1}) + (n_0 - n_{0+1})},$$

където
x_0' е левият край на модалния интервал,
n_0 е честотата в модалния интервал,
n_{0-1} е честотата в предмодалния интервал,
n_{0+1} е честотата в следмодалния интервал,
h е ширината на модалния интервал.

Медиана (Me) на един признак е стойността на този признак, при който 50% от наблюденията имат стойности, по-големи от Me, а 50% са по-малки от Me. Ако редът е дискретен, Me е
- средния член на реда, получен от първоначалния като стойностите са подредени във възходящ ред (при n - нечетно)
- средноаритметичното на двата централни члена на реда, получен от първоначалния като стойностите са подредени във възходящ ред (при n - четно).

В общия случай, медианата при интервално статистическо разпределение се пресмята по формулата

$$Me = x_m^{'} + \frac{\frac{n}{2} - (n_1 + n_2 + \ldots + n_{m-1})}{n_m}.h,$$

където
m е номерът на медианната група,
$x_m^{'}$ е левият край на медианния интервал,
h е ширината на медианния интервал.

Пример 6.2. *В таблицата*

Разход	$(14;16]$	$(16;18]$	$(18;20]$	$(20;22]$	$(22;24]$	$(24;26]$	$(26;28]$
Брой	4	5	7	11	15	10	3

са дадени отговорите на 55 човека, попитани за техния разход (в лв.) за храна в даден ден. Намерете Mo и Me.

От таблицата виждаме, че модалната група е n_5, съответстваща на модалния интервал $(22;24]$, т.к. $n_5 = 15$ е най - голямата честота. Тогава: $x_0^{'} = 22$, $n_0 = 15$, $n_{0-1} = 11$, $n_{0+1} = 10$. От формулата за Мо получаваме

$$Mo = 22 + \frac{(15-11).2}{(15-11)+(15-10)} = 22 + \frac{8}{9} \approx 22,9$$

Тъй като $n = 55$ е нечетно число, централният член е 28 - я, който се намира в интервала $(22; 24]$. Тогава $m = 5$, $x_m^{'} = 22$, $h = 2$ и за медианата получаваме

$$Me = 22 + \frac{27,5-27}{15}.2 = 22 + \frac{1}{15} \approx 22,07.$$

В статистиката голям интерес представлява групирането на дадения признак около неговата средна стойност, т.е. отклоненията на признака от неговата средна стойност - $(x_i - \overline{x})$. В теоретичните изследвания за разсейването се съди по така наречената дисперсия, която се определя по формулата

$$\sigma_x^2 = \frac{\sum(x_i-\overline{x})^2 f_i}{\sum f_i}.$$

За опростяване на изчисленията, може да се използва формулата

$$\sigma_x^2 = \overline{x}^2 - (\overline{x})^2.$$

Пример 6.3. *Нека имаме разпределение на даден признак x, зададен в интервален ред.*

Интервал	*Среда*(x_i^*)	*Честота*(f_i)	$x_i^* f_i$	$x*^2$	$x_i *^2 f_i$
$5 \div 15$	10	2	20	100	200
$15 \div 25$	20	6	120	400	2400
$25 \div 35$	30	10	300	900	9000
$35 \div 45$	40	4	160	1600	6400

Да се намери дисперсията.

За намиране на дисперсията пресмятаме последователно

$$\overline{x} = \frac{20+120+300+160}{22} = \frac{600}{22} = 27,3,$$

$$\overline{x}^2 = \frac{200 + 2400 + 9000 + 6400}{22} = \frac{18000}{22} \approx 818,$$
$$\sigma^2 \approx 818 - (27,3)^2 \approx 72,71.$$

Задачи:

1. Получени са резултати от наблюденията над интервалите от време между последователно пристигане на плавателни съдове в дадено пристанище. За даден срок в пристанището са пристигнали 185 съда.

x_i	$0 \div 4$	$4 \div 8$	$8 \div 12$	$12 \div 16$	$16 \div 20$	$20 \div 24$	$24 \div 28$	$28 \div 32$
f_i	67	43	30	18	11	7	5	4

Пресметнете средната стойност, модата, медианата и дисперсията на така получения интервален вариационен ред. (Отг. $\overline{x}$= 8,36, σ^2= 50,49, Mo=0,17, Me=46,38)

2. Да се изчислят стандартното отклонение и дисперсията при определяне на наличие на даден химичен елемент в 5 паралелни проби в % - ти 20,32; 20,55; 20,42; 20,48; 20,50. (Отг. σ_x^2=0,062, σ_x=0,249)

3. Измерени са дължините на 50 изделия. Получените данни са групирани в следната таблица:

Дължина x	$6,5$	$8,5$	$10,5$	$12,5$	$14,5$	$16,5$
Честота	2	4	18	15	8	3

Да се изчислят средната стойност, дисперсията и стандартното отклонение. (Отг. $\overline{x}$=11,78, σ_x^2= 5,24, σ_x=2,29)

4. Дадени са данните 20, 17, 19, 28, 25, 29, 29, 32, 41, 32, 23, 44, 26, 21, 43, 45, 18, 46, 30, 31, 25, 47, 29, 22, 36, 37. Да се групират в 5 групи и да се намери средната стойност, модата и медианата. (Отг. $\overline{x}$=30,92, Mo=30,71, Me=31,14)

6.2 Точкови и интервални оценки

При обработката на статистически данни често възниква въпросът за оценка на неизвестни параметри в разпределението на случайни величини въз основа на наблюдения върху стойности на тези случайни величини. Тези оценки могат да бъдат точкови и интервални. При точковите оценки параметърът се оценява само с една стойност, а при интервалните с две стойности - горна и долна доверителна граница.

6.2.1 Точкови оценки

Нека $x_1, x_2, \ldots, x_n$ са резултатите от n независими наблюдения над случайната величина X. Нека $F(x, \theta)$ е функция на разпределение на случайната величина, а θ е неизвестен параметър, който трябва да оценим по резултатите на n независими наблюдения. Точкова оценка на параметъра θ се нарича функцията $\widetilde{\theta} = \widetilde{\theta}(x_1, x_2, \ldots, x_n)$, която зависи само от резултатите на независимите наблюдения и от известни величини, но не зависи от неизвестния параметър θ. Точковата оценка $\widetilde{\theta}$ е случайна величина, т.к. се изчислява по елементите на извадка с обем n от разпределението на случайната величина X. На случайната величина $\widetilde{\theta}$ съответстват математическо очакване $E\widetilde{\theta}$, дисперсия $D\widetilde{\theta}$ и т.н. Точковите оценки на неизвестните параметри трябва да притежават три важни свойства:

а) **Неизместени оценки**. Оценката $\widetilde{\theta}(x_1, x_2, \ldots, x_n)$ на параметъра θ се нарича неизместена, ако $E\widetilde{\theta} = \theta$. Неизместената оценка дава възможност да се оцени неизвестният параметър без систематична грешка.

б) **Състоятелна оценка**. Точковата оценка се нарича състоятелна, ако при неограничено увеличаване на броя на наблюденията клони по вероятност към оценявания параметър, т.е. ако при всяко $\epsilon > 0$ е изпълнено $P(|\widetilde{\theta} - \theta| < \epsilon) \to 1$ при $n \to \infty$.

в) **Ефективна оценка**. Неизместена оценка с минимална дисперсия се нарича ефективна оценка.

Теорема 6.1. *Средното аритметично*

$$\overline{x} = \frac{1}{n}(x_1 + x_2 + \ldots + x_n)$$

е неизместена оценка на математическото очакване EX на наблюдаваната случайна величина X.

Нека DX=D е дисперсията на величината X. Оценка на параметъра D е величината

$$D_n = S_x = \frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{x})^2,$$

която се нарича изместена статистическа дисперсия. Оказва се, че D_n е състоятелна оценка за D, но е изместена относно D, т.е. $E(D_n) \neq D$. Може да се покаже, че $E(D_n) = \frac{n-1}{n}D$. Следователно, неизместената оценка за дисперсията е $\overline{D} = \widetilde{S}_x^2 = \frac{n}{n-1}D_n$.

Пример 6.4. *Да се намерят точковите оценки на математическото очакване и дисперсията на неизвестния признак X по дадена извадка:*

x_i	$0,03$	$0,05$	$0,07$
f_i	2	5	3

Неизместената оценка за EX е средната стойност на извадката

$$\overline{x} = \frac{0,03.2 + 0,05.5 + 0,07.3}{10} = 0,052.$$

За дисперсията има две оценки:
- Отместена оценка $D = S_x^2 = \overline{X}^2 - (\overline{X})^2 = 0,0029$.
- Неотместена оценка $\overline{D} = \widetilde{S}_x^2 = \frac{n}{n-1}S_x^2 = \frac{10}{9}.0,0029 = 0,0032$.

Един от стандартните методи за намиране на точкови оценки е методът на максималното правдоподобие. При този метод се предполага, че е известен видът на закона за разпределение на изучаваната случайна величина, а са неизвестни някои или всички параметри, които участват в закона на разпределението.

Нека резултатите от n независими наблюдения на непрекъснатата случайна величина X са $x_1, x_2, \ldots, x_n$ и нека диференциалната функция на X - $p(x, \theta)$ зависи от параметъра θ. Функцията

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

се нарича функция на правдоподобие. Същността на метода се състои в това, че за оценка на параметъра θ трябва да се вземе тази стойност на аргумента, за която функцията L има максимум.

Друг метод, който се използва за намиране на точкови оценки е методът на моментите. Той се състои в следното. Първо, от извадката изчисляваме необходимите извадъчни моменти, които оценяват съответните моменти на разпределението на величината. Търсените оценки на параметрите се определят, като се реши системата уравнения, отразяващи зависимостта между тези параметри и моментите на разпределението.

Задачи:

1. Известно е, че X е равномерно разпределена в интервала $[a, b]$. Да се намерят оценки на параметрите a и в по дадена извадка по метода на моментите.

X_i	$0 \div 4$	$4 \div 8$	$8 \div 12$	$12 \div 16$
f_i	5	3	7	5

.

<u>Решение</u>: За равномерното разпределение имаме (заменяме интервалите със средите им).

X_i^*	2	6	10	14
f_i	5	3	7	5

.

За оценките $\overline{X}$ и $\widetilde{S_x^2}$ на EX и DX получаваме:

$$\overline{X} = \frac{1}{20}(2.5 + 6.3 + 10.7 + 14.5) = 8,4.$$

$$S_x^2 = \overline{X} - (\overline{X}^2) = \frac{1}{20}(4.5 + 36.3 + 100.7 + 196.5) - (8,4)^2 = 19,84.$$

Следователно $\widetilde{S_x^2} = \frac{20}{19}.19,84 = \frac{512}{475}$. Така за оценките $\widetilde{a}$ и $\widetilde{b}$ получаваме системата

$$\left| \begin{array}{l} \frac{\widetilde{a}+\widetilde{b}}{2} = \frac{42}{5} \\ \frac{(\widetilde{b}-\widetilde{a})^2}{12} = \frac{512}{475} \end{array} \right.$$

Решението на системата е $\widetilde{a} \approx 0,56$ и $\widetilde{b} \approx 16,24$.

2. Случайната величина X (времето на работа на елемент) има показателно разпределение, зададено с диференциална функция

$$p(x) = \lambda e^{-\lambda x}.$$

В таблицата е дадено емпиричното разпределение на средното време на работа на 200 елемента

x_i	$2,5$	$7,5$	$12,5$	$17,5$	$22,5$	$27,5$
f_i	133	45	15	4	2	1

Намерете с метода на моментите точкова оценка на неизвестния параметър λ на показателното разпределение.

Решение: Известно е, че математическото очакване на тази случайна величина е $EX = \frac{1}{\lambda}$. Пресмятаме извадковото математическо очакване

$$\overline{X} = \frac{1}{200}(2,5.133 + 7,5.45 + 12,5.15 + 17,5.4 + 22,5.2 + 27,5.1) = 5.$$

Следователно $\widetilde{\lambda} = 0,2$.

3. Случайната величина X (брой на нестандартните изделия в партида) е разпределена по закона на Поасон. В следващата таблица е дадено разпределението на нестандартните изделия в една партида. Намерете с метода на моментите точкова оценка на неизвестния параметър на разпределението на Поасон. (Отг. 0,5)

x_i	0	1	2	3	4
f_i	132	43	20	3	2

4. Над случайната величина X с равномерно разпределение в интервала [a, b] са направени 200 независими наблюдения, в резултат на които събитието A е настъпвало в различни моменти от време. В следващата таблица е дадено емпиричното разпределение на настъпването на събитието. Намерете по метода на моментите точкови оценки за параметрите a и b. (Отг. $\widetilde{a} = 2,26$, $\widetilde{b} = 22,36$)

$x_{i-1} - x_i$	$2 \div 4$	$4 \div 6$	$6 \div 8$	$8 \div 10$	$10 \div 12$	$12 \div 14$	$14 \div 16$	$16 \div 18$	$18 \div 20$	$20 \div 22$
m_i	21	16	15	26	22	14	21	22	18	25

5. Като използвате метода на моментите, намерете точкови оценки на параметрите а и σ^2 на нормалното разпределение, зададено с диференциалната функция на разпределение

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

(Отг. $\widetilde{a} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \widetilde{S}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \widetilde{a}))^2$.)

6.2.2 Интервални оценки

Възниква въпросът за оценка на параметъра θ не с число (точкова оценка), а с интервал (θ_1, θ_2), така че вероятността за покриване на θ от този интервал да не е по-малка от дадено $p = 1 - \alpha$, т.е.

$$P(\theta_1 < \theta < \theta_2), \ \ p \in (0, 1).$$

Тук p се нарича доверителна вероятност, θ_1 и θ_2 - доверителни граници, интервалът (θ_1, θ_2) - доверителен интервал, α- ниво на значимост. Ако изберем α да е малко, тогава $p = 1 - \alpha$ е близо до 1 и търсената вероятност се доближава до достоверно събитие.

6.2.3 Доверителен интервал за математическото очакване и дисперсията на нормално разпределение

Нека над случайната величина X, разпределена по нормален закон с неизвестно математическо очакване а и известна дисперсия σ^2, са направени n независими наблюдения. За намирането на доверителен интервал за математическото очакване първо намираме точкова оценка за

$\overline{X} = \frac{1}{n}\sum_{k=1}^{n} x_k$. Доверителният интервал

$$(\overline{X} - d,\ \ \overline{X} + d)$$

с доверителна вероятност $1 - \alpha$ покрива неизвестното математическо очакване а. Известно е, че вероятността, оценката на а да лежи в интервала е стойността на функцията $\Phi\left(\frac{d}{\sigma_{\overline{x}}}\right)$.

Пример 6.5. *Искаме да намерим интервална оценка на средния ръст на студентите от един университет с доверителна вероятност 0,95. Имаме извадка с обем* $n = 100$*. Нека средният ръст на извадката е* $\overline{x} = 1,73$*, а стойността на извадковата дисперсия е* $S^2 = 0,00245$.

Използваме равенството

$$P(\overline{x} - d < a < \overline{x} + d) = \Phi\left(\frac{d}{\sigma_{\overline{x}}}\right).$$

Заместваме лявата страна на равенството с $95\% = 0,95$. Получаваме

$$0,95 = \Phi\left(\frac{d}{\sigma_{\overline{x}}}\right).$$

От таблица.... получаваме $\frac{d}{\sigma_{\overline{x}}} = 1,96$. За намиране на d , трябва да знаем $\sigma_{\overline{x}}$. Използваме, че $\sigma_{\overline{x}^2} = \frac{\sigma^2}{100}$, като за оценка на σ^2 използваме $S^2 = 0,00245$. Тогава $\sigma_{\overline{x}} = 0,0049$ и $d = 1,96.0,0049 = 0,0097 \approx 0,01$. Така доверителният интервал е $(1,73 - 0,01; 1,73 + 0,01)$ или $(1,72m.; 1,74m.)$. Полученият резултат показва, че с вероятност 0,95 може да се твърди, че средният ръст на студентите ще се различава от средния ръст от извадката с $\pm 0,01$.

Ако стандартното отклонение не е известно, намирането на доверителен интервал при доверителна вероятност γ за математическото очакване се търси, като спазваме следните стъпки:
- Намираме точкова оценка за EX: $\overline{x} = \frac{1}{n}\sum_{k=1}^{n} x_k$.
- Намираме $D_n = \frac{1}{n}\sum(X_k - \overline{x})^2$, и $\overline{D} = \frac{n}{n-1}D_n$
- Определяме $s = \sqrt{\overline{D}}$

- От таблица определяме $x_\alpha = t_{\frac{1+\gamma}{2}}(n-1)$ на разпределението на Стюдънт с n-1 степени на свобода и пресмятаме точността

$$\Delta = \frac{x_\alpha}{\sqrt{n}}\overline{s}.$$

-Доверителният интервал за математическото очакване е $(\overline{x} - \Delta; \overline{x} + \Delta)$ и оценяваната величина попада в интервала с доверителна вероятност γ.

По подобен начин можем да търсим доверителен интервал за дисперсията DX на нормално разпределена случайна величина X с доверителна вероятност γ по извадка с обем n:
- Ако EX е известно, то

$$\left(\frac{n.D_n}{\chi^2_{\frac{1+\gamma}{2}}(n)} < DX < \frac{n.D_n}{\chi^2_{\frac{1-\gamma}{2}}(n)}\right).$$

-Ако EX не е известно, то

$$\left(\frac{(n-1).\overline{D}}{\chi^2_{\frac{1+\gamma}{2}}(n-1)} < DX < \frac{(n-1).\overline{D}}{\chi^2_{\frac{1-\gamma}{2}}(n-1)}\right).$$

Тук $\chi^2_{\frac{1+\gamma}{2}}(n)$ и $\chi^2_{\frac{1-\gamma}{2}}(n)$ са квантили на χ^2 разпределение с n степени на свобода, а $\chi^2_{\frac{1+\gamma}{2}}(n-1)$ и $\chi^2_{\frac{1-\gamma}{2}}(n-1)$ са квантили на χ^2 разпределение с (n-1) степени на свобода.

Задачи:

1. Клиент поръчва на фирма изработката на даден вид изделие. Известно е, че теглото на изделията е нормално разпределена величина. Направени са измервания на 9 случайно избрани изделия, при които резултатите са:

X_i	103	104	105	107	108
f_i	1	3	2	2	1

Намерете:

а) Доверителен интервал за средното тегло с доверителна вероятност $\gamma = 0,98$

б) Доверителен интервал за отклонението в теглото с $\gamma = 0,95$.

Решение: а) $\gamma = 0,98$; $\overline{X} = 105,2$; $\overline{D}_n = 2,945$; $s = \sqrt{2,945} = 1,716$. $x_\alpha = t_{0,99}(8) = 2,896$. Следователно $\delta = \frac{2,896}{\sqrt{9}}.1,716 = 1,656$. Доверителният интервал е $(103,5; 106,8)$.
б) $\chi^2_{\frac{1+0,95}{2}}(8) = \chi_{0,975}(8) = 17,5.$, $\chi^2_{\frac{1-0,95}{2}}(8) = \chi_{0,025}(8) = 2,18.$. Тогава интервалът за дисперсията е $1,346 < DX < 10,807$ и $1,1602 < \sigma < 3,2874$.

2. В офис е поставена кафе машина, която е регулирана така, че количеството кафе, което подава, е нормално разпределена със $\sigma = 10$ гр. Направена е случайна извадка от 50 чаши и е установено, че средната стойност е 150 гр. на чаша. Намерете 95% интервал за средната стойност.
Решение: От условието имаме, че $\overline{X} = 150; \sigma = 10$. Тогава $0,95 = \Phi(\frac{d}{\sigma_{\overline{x}}}) = 1,96$. Следователно $\frac{d}{\sigma_{\overline{x}}} = 1.96$, $\sigma_{\overline{x}} = \frac{\sigma^2}{50} = \frac{100}{50} = 2$. Така $d = \sqrt{2}.1,96$. За доверителния интервал за средната стойност получаваме $(147,228; 152,772)$.

3. Измерва се концентрацията на химичен елемент във въздуха на 10 различни места, като е известно, че разпределението е нормално. Резултатите са

$$6,5; 7,2; 7,6; 6,8; 7,0; 7,9; 6,6; 6,1; 6,5; 7,0.$$

а) Ако дисперсията в концентрацията е 0,49, да се определи доверителния интервал за средната концентрация с доверителна вероятност 0,95.
б) Да се определи доверителният интервал за средната концентрация с доверителна вероятност 0,90.
в) Да се определи доверителният интервал за дисперсията с доверителна вероятност 0,98.
Решение: а) Обемът на извадката е $n = 10$. Числовите характеристики на извадката са:
$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = 6,92. D_n = \frac{1}{10}[(6,5-6,92)^2 + \ldots + (7-6,92)^2] = 0,2656$, $\overline{D} = \frac{10}{9}.0,2656 = 0,2951$. $\sigma^2_{\overline{x}} = \frac{0,49}{10} = 0,049$. Тогава $\frac{d}{\sqrt{0,049}} = 1,96$, $d = 0,433$. За доверителния интервал получаваме $(6,4861; 7,3539)$.
б) Използваме формулите за доверителен интервал при неизвестна дисперсия. $s = \sqrt{\overline{D}} = 0,5432$. $x_\alpha = t_{\frac{1+0,9}{2}}(9) = t_0,95(9) = 1,833$, $\delta = \frac{1,833.0,5432}{\sqrt{10}} = 0,3149$. За доверителния интервал се получава (6,605; 7,2349).
в) Използваме формулите за доверителен интервал за дисперсия при не-

известно средно:

$$\left(\frac{(n-1).\overline{D}}{\chi^2_{\frac{1+\gamma}{2}}(n-1)} < DX < \frac{(n-1).\overline{D}}{\chi^2_{\frac{1-\gamma}{2}}(n-1)} \right),$$

$\chi^2_{0,99}(9) = 21,7$, $\chi^2_{0,01}(9) = 2,09$. Така получаваме $0,1226 < DX < 1,2721$.

4. От партида еднотипни съпротивления са избрани по случаен начин за контрол 10 броя. Измерванията са дали следните отклонения от номиналната стойност (в килоома): +1, +3, -2, +2, +4, +2, +5, +3, -2, +4. Намерете доверителен интервал за оценка на математическото очакване на съпротивлението при предположение, че съпротивлението е нормална случайна величина. Изберете доверителна вероятност 0,95. (Отг. (0,28; 3,72))

5. Чрез оценяването на случайно избрани изделия от партида по няколко контролни показателя се оценява нейното качество. Получените оценки от 8 изделия са: 5,30; 4,85; 5,95; 4,60; 5,30; 4,40. Известно е, че оценките са нормално разпределени.
а) Да се намери доверителен интервал за средната оценка с доверителна вероятност 0,90, ако отклонението в оценяването е 1,3. (Отг. (4,3916; 5,9084))
б) Да се намери доверителен интервал за дисперсията с доверителна вероятност 0,99. (Отг. (0,1225; 1,2492))

6.3 Корелационен и регресионен анализ

Целта на регресионния анализ е да установи силата на връзката между зависима променлива и една или повече независими променливи. При еднофакторния регресионен анализ имаме 2 количествени променливи X и Y. Задачата е да се определи дали има зависимост между двете променливи и ако има, да се установи силата на тази връзка чрез коефицента на корелация (r_{xy}). Нека приемем, че разглеждаме линейна зависимост и я търсим във вида $y = a + bx$. На практика не знаем r_{xy}, но можем да го оценим чрез извадката, като пресметнем статистическия коефицент на корелация $\overline{r}_{xy}$. За целта пресмятаме величините:

$$S_x = \sum_{i=1}^{n} x_k, \quad S_y = \sum_{i=1}^{n} y_i,$$

$$S_{xx} = \sum_{i=1}^{n} x_i^2, \quad S_{yy} = \sum_{i=1}^{n} y_i^2, \quad S_{xy} = \sum_{i=1}^{n} x_i y_i,$$

след което определяме оценките за EX, DX,EY, DY, K_{XY} и r_{XY}:

$$\overline{x} = \frac{S_x}{n}, \quad \overline{D}_X = \frac{nS_{xx} - S_x^2}{n^2},$$

$$\overline{y} = \frac{S_y}{n}, \quad \overline{D}_y = \frac{nS_{yy} - S_y^2}{n^2},$$

$$\overline{K}_{XY} = \frac{nS_{xy} - S_x S_y}{n^2}, \quad \overline{r}_{XY} = \frac{\overline{K}_{XY}}{\sqrt{\overline{D}_X \overline{D}_Y}}.$$

Коефицентите a и b се определят с формулите:

$$b = \frac{\overline{K}_{XY}}{\overline{D}_X}, \quad a = \overline{y} - a.\overline{x}.$$

Пример 6.6. *Да се намери линията на регресия и да се определи коефицентът на корелация, ако X и Y приемат следните стойности:*

x_i	1	2	3	4	5	6
y_i	10	3	9	2	6	8

Имаме, че

$$n = 6, \quad \overline{x} = \frac{1+2+3+4+5+6}{6} = 3,5, \quad \overline{y} = \frac{10+3+9+2+6+8}{6} = 6,3,$$

$$\overline{D}_X = \frac{6.(1+4+9+16+25+36) - (3,5)^2}{36} = 14,82,$$

$$\overline{D}_Y = \frac{6.(100+9+81+4+36+64) - (6,3)^2}{36} = 47,9,$$

$$\overline{K}_{XY} = \frac{6.(10+6+27+8+30+48) - 21.38}{36} = \frac{129-798}{36} = -18,59,$$

$$\overline{r}_{XY} = \frac{-18,59}{\sqrt{14,82.47,9}} \approx -0,7.$$

За коефицентите на линията на регресия имаме

$$b = \frac{-18,59}{14,82} = -1,25, \quad a = 6,3 - (-1,25).3,5 = 10,675.$$

За уравнението на линията на регресия получаваме $y = 10,675 - 1,25x$.

Задачи:

1. Да се намери коефицентът на корелация и линията на регресия за данните зададени в следната таблица:

x_i	21	22	24	30	22	28	24	26	28	21
y_i	170	180	200	300	220	320	150	180	200	150

Решение: Последователно пресмятаме характеристиките на данните:

$$S_x \sum x_i = 246; S_y = \sum y_i = 2070; S_{xx} = 6146; S_{yy} = 459500; S_{xy} = 52160$$

$$D_X = 9,44; D_Y = 3101; K_{XY} = 123,8$$

Тогава следва, че $\overline{R}_{xy} = 0,7236$. Пресмятаме и коефицентите a и b за линията на регресия. $a = -115,61, \quad b = 13,1144$.Следователно уравнението на правата линия има следния аналитичен вид

$$Y = -115,61 + 13,1144X.$$

Коефицентът на корелация е по-голям от 0,7 затова връзката между двете променливи е силна.

2. Намерете коефицента на корелация за данните

x_i	-4	-3	-2	-1	0	1	2	3	4	5
y_i	2	3	4	1	2	3	3	2	1	0

(Отг. -0,54).

3. Намерете уравнението на линейна регресия за зависимост на у от х за данните

x_i	-4	-3	-2	-1	0	1	2	3	4	5
y_i	-2	-1	0	1	2	3	3	5	6	7

(Отг. Y=1,91+0,98X).

4. Намерете коефицента на корелация на данните, зададени със следната таблица:

v/u	-2	-1	0	1	2
-2	4	6	–	–	–
-1	–	8	10	–	–
0	–	–	32	3	9
1	–	–	4	12	6
2	–	–	–	1	5

(Отг. $\overline{R}_{xy} = 0,7647$)

5. Нека имаме резултатите на 5 студента, постигнати на тест по математика(X) и тест за интелигентност(У)

Тест математика	11	12	10	14	16	18
Тест за интелигентност	60	62	59	70	61	73

Намерете уравнението на линейна регресия и предвидете, ако студент получи 15 точки на теста по математика, колко точки ще получи на теста за интелигентност. (Отг. $y = 45,267 + 1,4.x$. При получени 15 точки на теста по математика, предвидимо ще получи около 66 точки на теста за интелигентност)

6.4 Проверка на хипотези

При решаването на много задачи е необходимо да се прави предположение за вида на законите на разпределение на разглежданите случайни величини или за съотношението между техните числови характеристики. Такива предположения е прието да се наричат статистически *хипотези*. Различаваме *прости* и *сложни* статистически хипотези. Простата хипотеза, за разлика от сложната, напълно определя теоретичната функция на разпределение на случайната величина. Например "вероятността за настъпване на дадено събитие е $\frac{1}{2}$", "законът на разпределение на случайната величина е нормален с параметри 0 и 1" са прости хипотези, а хипотезите "вероятността за настъпване на дадено събитие е между 0,2 и 0,5", "случайната величина не е разпределена по нормален закон" са сложни. Хипотезата, която се проверява е прието да се нарича нулева и се бележи с Ho. Заедно с Ho се разглежда и *алтернативна* хипотеза H1, която е отрицание на Ho. За проверка на статистически хипотези се използва съставена извадкова характеристика (статистика) $\widetilde{\theta}_n(x_1, x_2 \dots x_n)$, получена от извадака $X_1, X_2 \dots X_n$, на която точното или приблизително разпределение е известно. По тази статистика определяме критично значение $\theta_{\text{кр}}$ - такова, че ако Ho е вярна, то вероятността $P(\widetilde{\theta}_n > \theta_{\text{кр}} = \alpha$ е малка (тук α е вероятността да отхвърлим Ho, когато тя е вярна и се нарича ниво на значимост). Затова, ако $\widetilde{\theta}_n > \theta_{\text{кр}}$, то Ho се отхвърля. Правилото по което Ho се отхвърля или приема се нарича статистически тест.

В настоящото ръководство се разглеждат задачи за проверка на хипотези относно вида на разпредление на случайни величини. За проверката на хипотези за вида на разпределение ще разгледаме критерия χ^2 (критерий на Пирсън). Критерият се състои в следното. Множеството от възможните стойности на случайната величина X се разбива на s непресичащи се части $J_1, J_2, \dots J_s$. Нека $p_1, p_2, \dots p_s$са вероятностите случайната величина X да приеме стойност в интервалите $J_1, J_2, \dots J_s$. Нека $m_1, m_2, \dots m_s$ е броят на емпиричните стойности, попадащи в интервалите. За мярка на различието на разпределението на извадката $F_n(x)$ от предполагаемото разпределение $F(x)$ се приема случайната величина

$$\chi_q^2 = \sum_{i=1}^{s} \frac{(m_i - np_i)^2}{np_i},$$

която при $n \to \infty$ асимптотически е разпределена по закона χ^2 с k степени на свобода. Броят на степените на свобода се определя по формулата $k = s - r - 1$, където s е броят на групите на извадката, r е броят на параметрите на предполагаемото разпределение, които се оценяват по данните на извадката. При зададено ниво на съгласие α, за да проверим нулевата хипотеза H_0, трябва по таблицата за χ^2 разпределение да намерим критичната точка $\chi^2(\alpha, k)$. Ако $\chi_q^2 < \chi^2(\alpha, k)$, няма основание да отхвърлим хипотезата.

Използването на този критерий е свързано с някои ограничения. Обемът на извадката трябва да бъде най-малко 50 и теоретичните честоти трябва да бъдат не по-малко от 5-8 в група.

Задачи:

1. Радиоактивно вещество е наблюдавано в продължение на 2 608 равни интервала от време. За всеки от интервалите е регистриран броят на частиците, които са попаднали в брояча.

i	0	1	2	3	4	5	6	7	8	9	10
m_i	57	203	383	525	532	408	273	139	45	27	16

.

Като използвате χ^2 критерия, проверете хипотезата за съгласуваност на наблюдаваните честоти със закона на разпределение на Поасон. За ниво на значимост α приемете 0,05.

Решение: Въз основа на данните от таблицата изчисляваме оценката на параметъра $\widetilde{\lambda}$ на закона на разпределение на Поасон по формулата

$$\widetilde{\lambda} = \frac{1}{n} \sum_{i=0}^{11} i m_i,$$

където $n = 2608$. Тогава $\widetilde{\lambda} = 3,87$. Спомагателните изчисления за прилагане на критерия са поместени в следната таблица.

i	$p_i = \frac{e^{\widetilde{\lambda}}.\widetilde{\lambda}^k}{k!}$	np_i	$m_i - np_i$	$(m_i - np_i)^2$	$\frac{(m_i-np_i)^2}{np_i}$
0	$0,021$	$54,8$	$2,2$	$4,84$	$0,088$
1	$0,081$	$211,2$	$-8,2$	$67,24$	$0,318$
2	$0,156$	$406,8$	$-23,8$	$566,44$	$1,392$
3	$0,201$	$524,2$	$0,8$	$0,64$	$0,001$
4	$0,195$	$508,6$	$23,4$	$547,56$	$1,007$
5	$0,151$	$393,8$	$14,2$	$201,64$	$0,512$
6	$0,097$	$253,0$	$20,0$	$400,00$	$1,581$
7	$0,054$	$140,8$	$-1,8$	$3,24$	$0,023$
8	$0,026$	$67,8$	$-22,8$	$519,84$	$7,667$
9	$0,011$	$28,7$	$-1,7$	$2,89$	$0,101$
10	$0,007$	$18,3$	$-2,3$	$5,29$	$0,289$

Сумираме числата от последната колона на таблицата и получаваме $\chi_q = 12,979$. Имаме 11 интервала и един параметър, затова $k = 11 - 1 - 1 = 9$. По таблицата определяме $\chi^2(0,05;9) = 16,9$. Тъй като $\chi_q < \chi^2(0,05;9)$, няма основание да отхвърлим хипотезата за съгласуваност на наблюдаваните честоти със закона на разпределение на Поасон.

2. Цифрите 0, 1, 2, 3, $\ldots 9$ измежду първите 800 десетични знака на числото π се наблюдават съответно 74, 92, 83, 79, 80, 73, 77, 75, 76, 91 пъти. Проверете хипотезата за съгласуваност на тези данни със закона за равномерно разпределение при ниво на значимост 10%.

Решение: Пред вид предположението за равномерно разпределение на цифрите, съществува постоянна вероятност $p_i = 0,1$ за наблюдаване на произволна от тях. Тогава $np_i = 800.0,1 = 80$. Стойността на критерия χ_q^2 е

$$\chi_q^2 = [(74-80)^2+(92-80)^2+(83-80)^2+(79-80)^2+((80-80)^2+(73-80)^2+$$

$$+(77-80)^2+(75-80)^2+(76-80)^2+(91-80)^2] : 80 = 5,125.$$

Степените на свобода са $k = 10 - 1 = 9$. По таблицата намираме критичната точка $\chi^2(0,10;9) = 14,7$. Тъй като $\chi_q^2 < \chi^2(0,10;9)$ няма основание да отхвърлим хипотезата за съгласуваност на наблюдаваните честоти със закона за равномерно разпределение.

3. В резултат на регистрацията на времето на пристигане на 800 посетители на един магазин е получено емпиричното разпределение отразено в следната таблица

$x_i - x_{i+1}$	$0-1$	$1-2$	$2-3$	$3-4$	$4-5$	$5-6$	$6-7$	$7-8$
m_i	259	167	109	74	70	47	40	34

.

При ниво на значимост 0,01 проверете хипотезата, че времето на пристигане на посетителите на магазина е разпределено по показателен закон.
Решение: За оценка на параметъра λ на предполагаемото показателно разпределение е необходимо средното време на посещение.

$$\overline{x} = (0,5.259+1,5.167+2,5.109+3,5.74+4,5.70+5,5.47+6,5.40+7,5.34) : 800 = 2,5.$$

Така параметърът $\lambda = \frac{1}{\overline{x}} = 0,4$. Следователно диференциалната функция на разпределение е

$$f(x) = 0,4e^{-0,4x}(x \geq 0).$$

Намираме вероятността случайната величина X да приеме стойност във всеки интервал:

$$p_1 = P(0 < X < 1) = e^{-0,4.0} - e^{-0,4.1} = 0,3297,$$

$$p_2 = P(1 < X < 2) = e^{-0,4.1} - e^{-0,4.2} = 0,2210,$$

$$p_3 = P(2 < X < 3) = e^{-0,4.2} - e^{-0,4.3} = 0,1481,$$

$$p_4 = P(3 < X < 4) = e^{-0,4.3} - e^{-0,4.4} = 0,0993,$$

$$p_5 = P(4 < X < 5) = e^{-0,4.4} - e^{-0,4.5} = 0,0666,$$

$$p_6 = P(5 < X < 6) = e^{-0,4.5} - e^{-0,4.6} = 0,0446,$$

$$p_7 = P(6 < X < 7) = e^{-0,4.6} - e^{-0,4.7} = 0,0299,$$

$$p_8 = P(7 < X < 8) = e^{-0,4.7} - e^{-0,4.8} = 0,0200.$$

Спомагателните изчисления са поместени в следващата таблица

m_i	p_i	np_i	$(m_i - np_i)^2$	$\frac{(m_i-np_i)^2}{np_i}$
259	$0,3297$	$263,76$	$22,6576$	$0,086$
167	$0,2210$	$176,8$	$96,04$	$0,543$
109	$0,1481$	$118,48$	$89,8704$	$0,759$
74	$0,0993$	$79,44$	$29,5936$	$0,373$
70	$0,0666$	$53,28$	$279,5584$	$5,247$
47	$0,0446$	$35,68$	$128,1424$	$3,591$
40	$0,0299$	$23,92$	$258,5664$	$10,81$
34	$0,02$	16	324	$20,25$

Сега $k = 8 - 2 = 6$; $\chi_q^2 = 41,659$; $\chi^2(0,01;6) = 16,8$. Следователно имаме основание да отхвърлим разглежданата хипотеза.

4. Изпитва се чуствителността на 50 телевизора. Данните от изпитването са поместени в следната таблица, като в първата колона са дадени интервалите на чуствителност в микроволтове, а във втората - броят телевизори, чуствителността на които е попаднала в дадения интервал.

Интервали	n_i	Интервали	n_i	Интервали	n_i
$75 - 125$	1	$275 - 325$	8	$475 - 525$	2
$125 - 175$	2	$325 - 375$	8	$525 - 575$	2
$175 - 225$	4	$375 - 425$	6	$575 - 625$	1
$225 - 275$	9	$425 - 475$	5	$625 - 675$	2

Проверете хипотезата, това разпределение е нормално с параметри $\widetilde{m_x}$ и $\widetilde{\sigma_x}$ при ниво на значимост 0,05.

Решение: Определяме параметрите:

$\widetilde{m_x} = \frac{1}{50}(100.1+150.2+200.4+250.9+300.8+350.8+400.6+450.5+500.2+550.2+600.1+650.2) = 346.$

$$\widetilde{\sigma_x} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \widetilde{m_x})^2} \approx 125,4.$$

Следователно, плътността на вероятностите има вида

$$f(x) = \frac{1}{125,4\sqrt{2\pi}} e^{-\frac{(x-346)^2}{2(125,4)^2}},$$

а функцията на разпределение $F(x) = \frac{1}{2}\left[\Phi\left(\frac{x-346}{125,4}\right)+1\right]$. Обединяваме всички данни в 7 групи, защото във всяка група трябва да има поне 5 наблюдения.

Интервали	< 225	$225-275$	$275-325$	$325-375$	$375-425$	$425-475$	> 475
m_i	7	9	8	8	6	5	7

Следва да намерим вероятностите

$$p_i = \frac{1}{2}\left[\Phi\left(\frac{x_{i+1}-346}{125,4}\right) - \Phi\left(\frac{x_i-346}{125,4}\right)\right],$$

където $x_i,\ \ x_{i+1}$ са границите на $i-$ тата група.

$$p_1 = 0,1686; p_2 = 0,1192; p_3 = 0,1459; p_4 = 0,1576;$$

$$p_5 = 0,1443; p_6 = 0,0722; p_7 = 0,1922.$$

Определяме χ^2

m_i	p_i	np_i	$(m_i - np_i)^2$	$\frac{(m_i-np_i)^2}{np_i}$
7	$0,1686$	$8,43$	$2,045$	$0,24$
9	$0,1192$	$5,96$	$9,241$	$1,55$
8	$0,1459$	$7,295$	$0,497$	$0,07$
8	$0,1576$	$7,88$	$0,0144$	$0,002$
6	$0,1443$	$7,215$	$1,476$	$0,2$
5	$0,0722$	$3,61$	$1,932$	$0,54$
7	$0,1922$	$9,61$	$6,812$	$0,71$

Тогава $\chi^2 \approx 3,31$, $\ k = 7-2-1 = 4$. От таблицата намираме $\chi_q^2(0,05,4) = 9,49$. Следователно можем да считаме хипотезата за правдоподобна.

5. Контролира се размерът X на даден вид детайли, изработвани на машина. Проверени са 68 детайла. С ниво на значимост $\alpha = 0,1$ да се

провери хипотезата за нормално разпределение на контролирания размер.

$x_{i-1} - x_i$	$(3,4)$	$(4,5)$	$(5,6)$	$(6,7)$	$(7,8)$
m_i	5	15	23	19	6

(Отг. $\chi^2 \approx 0,3917, \quad \chi_q^2 = 4,61.$)

6. В резултат на проверката на 200 елемента за продължителност на безотказна работа е получено емпирично разпределение

$x_{i-1} - x_i$	$(0,5)$	$(5,10)$	$(10,15)$	$(15,20)$	$(20,25)$	$(25,30)$
m_i	133	45	15	4	2	1

Проверете хипотезата за съгласуваност на времето на безотказна работа на елементите със закона на показателно разпределение при ниво на значимост 0,05.
Упътване: Обединете честотите в последните 3 групи. (Отг. $\chi^2 \approx 1,53, \quad \chi_q^2 = 6,0.$)

Глава 7

Вероятности и статистика с MATLAB

Системата MATLAB [1] съдържа множество от функции и процедури, необходими за сложни числени операции, моделиране на технически и физични системи и тяхната визуализация. Системата напълно поддържа операции с вектори, матрици, многомерни масиви, масиви от клетки и масиви от записи. Тя може да се използва като мощен калкулатор, поддържащ не само обичайните аритметични действия, но и сложни операции като намиране на обратна матрица, решаване на системи уравнения, векторно смятане и др.

MATLAB има вграден програмен език от високо ниво, което дава възможност не само за работа в режим на калкулатор, но и в програмен режим. Това позволява всеки потребител сам да разработи решението на желаната от него задача. Системата дава възможност за обръщение към програми, написани на FORTRAN, C, C++. Някой процедури в MATLAB използват в качество на параметър имена на функции. В този случай е необходимо да се създаде т.нар. m–файл. При обръщението към тези процедури трябва да се укаже името на m-файла, в който е изчислена стойността на функцията за известна стойност на аргумента.

[1]MATLAB © е търговска марка на Math Wotks, Inc.

7.1 Комбинаторика

За работа с пермутации в MATLAB съществуват следните функции

- **factorial(n)**-пресмята произведението на целите числа от 1 до n
- **randperm(n)** - връща случайна пермутация на числата от 1 до n
- **perms(v)** - създава матрица с n! реда и n стълба със всички възможни пермутации на елементите на вектора v

```
>> p80=factorial(80)
p80 =

  7.1569e+118

>> p9=randperm(9)

p9 =

     6     7     1     2     5     3     9     4     8

>> allp=perms(1:3)

allp =

     3     2     1
     3     1     2
     2     3     1
     2     1     3
     1     2     3
     1     3     2
```

В стандартния MATLAB няма вградена функция за пресмятане на вариации. Може да се използва $factorial(n)/factorial(n-k)$. Пълният брой комбинации се от N елемента k-ти клас се изчислява с помощта на командата *nchoosek* със следния синтаксис

- **nchoosek(N,k)** - връща броя на комбинациите
- **nchoosek(v,k)** - за векторът v с дължина N връща матрица с C_N^k реда и k стълба. Всеки ред е една комбинация.

Пример 7.1. *Колко са начините за разделяне на студентска група от 25 човека на две подгрупи, във всяка от които има от 10 до 15 човека?*

Разделяме групата на 2 части - това означава да изберем една от частите. Избираме 10 човека от 25 по C_{25}^{10} начина, 11 - по C_{25}^{11}начина и т.н. Общия брой начини е

$$C_{25}^{10} + C_{25}^{11} + C_{25}^{12} + C_{25}^{13} + C_{25}^{14} + C_{25}^{15}.$$

Пресмятаме с помощта на MATLAB:

```
>> C=0;
>> for k=10:15,
C=C+nchoosek(25,k);
end
 >> fprintf('Общ брой случай=%d.\n',C)


Общ брой случай=25852920.
```

7.2 Вероятности

Пример 7.2. *Хвърляме 6 зарчета. Да се намери вероятността сумата от точките да е между 20 и 30.*

```
 >> nn=7; % брой зарчета
>> n=6^nn  % общ брой случай

n =

      279936

 >> x=sum(fullfact(6*ones(1,nn)),2);
 % сумира елементите на редовете на матрицата с всевъзможните
  пермутации на цифрите от 1 до 10

>> m=length(find((x>=20)&(x<=30)))  % брой благоприятни случай

m =

      215259

>> p=m/n  % пресмята търсената вероятност

p =

    0.7690
```

Пример 7.3. *Хвърляме монета 50 пъти. Каква е вероятността броя на падането на лице да е между 20 и 25?*

```
>> m=50; %брой опити
>> n=2^m %брой благоприятни случаи

n =
```

```
  1.1259e+015

>> m1=[];  % цикъл за пресмятане на комбинациите
>> for k=20:25
m1=[m1 nchoosek(m,k)];
end
>> m1=sum(m1) % брой благоприятни случаи

m1 =

  5.5921e+014

>> p=m1/n %търсената вероятност

p =

    0.4967
```

7.3 Случайни величини

Пример 7.4. *Дадена е дискретна случайна величина със закон на разпределение:*

x	$0,2$	1	2	3	$3,5$	4
P	$0,2$	$0,15$	$0,3$	$0,2$	$0,1$	$0,05$

. *Да се намерят (начертаят):*

а) полигона на честотите

б) графиката на функцията на разпределение

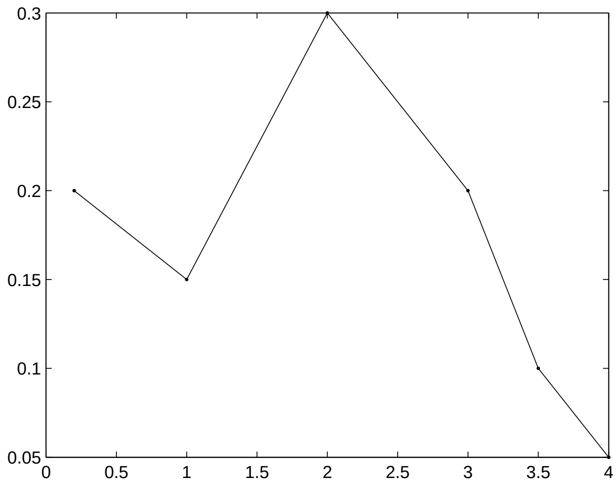
в) математическото очакване, модата и дисперсията на X.

Решение: а)

```
  >> x=[0.2 1 2 3 3.5 4];      %x(i)
>> p=[0.2 0.15 0.3 0.2 0.1 0.05];       %p(i)
```

```
>> plot(x,p,'k-',x,p,'k.')   %чертае полигона
```

Получаваме следната графика



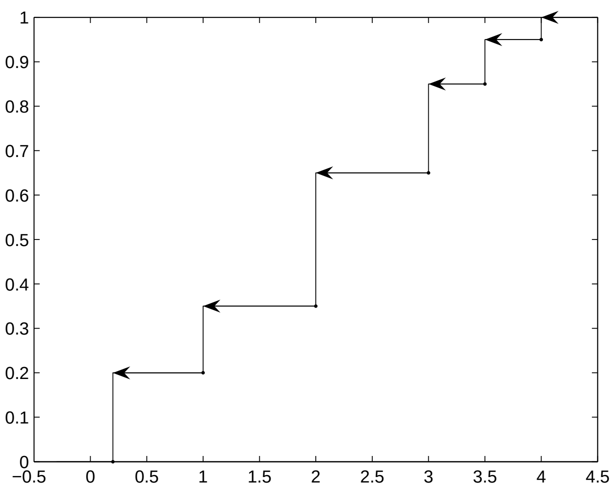
Фиг.3

б)

```
>> F=cumsum(p);         % стойностите на функцията на разпределение
>> x1=[x(1)-0.5 x x(end)+0.5];
>> F1=[0 F 1];
>> figure;   % нова графика
>> stairs(x1,F1,'k-');
>> xl=xlim;   % граници на графиката
>> yl=ylim;
>> hold on
>> plot(x,F1(1:end-2),'k.')  %добавяме точки
>> hh=get(gca);
>> hp=hh.Position;
>> for i=1:length(F),
xi=x1(2+i:-1:1+i);
```

```
Fi=[F(i) F(i)];
xi=(xi-xl(1))/(xl(2)-xl(1))*hp(3)+hp(1);
Fi=(Fi-yl(1))/(yl(2)-yl(1))*hp(4)+hp(2);
annotation('arrow',xi,Fi); % добавяме стрелки
end
>> hold off
```

Изпълняваме командите и получаваме графиката



Фиг.4

в)

```
   >> mx=sum(x.*p) % математическото очакване

mx =

    1.9400
>> [pmax,ipmax]=max(p); %максималната верочтност и номера на x
>> modx=x(ipmax)% пресмята модата
```

```
modx =

     2

>> dx=sum((x-mx).^2.*p)  % пресмята дисперсията

dx =

    1.4194
```

Пример 7.5. *Непрекъсната случайна величина X има плътност на вероятностите*

$$p(x) = \begin{cases} 0, & x < 0, \\ C(2x+1), & 0 \leq x \leq 2, \\ 0, & x > 2. \end{cases}$$

Да се определят: A, F(x), EX и DX.

```
>> syms x c  % дефинираме символни променливи
>> x1=0;
>> x2=2;
>> f=c*(2*x+1);
>> I1=int(f,x,x1,x2);
>> cs=solve(I1-1)  %решаваме уравнението I1=1 за намиране на C

cs =

1/6

>> fs=subs(f,c,cs)
```

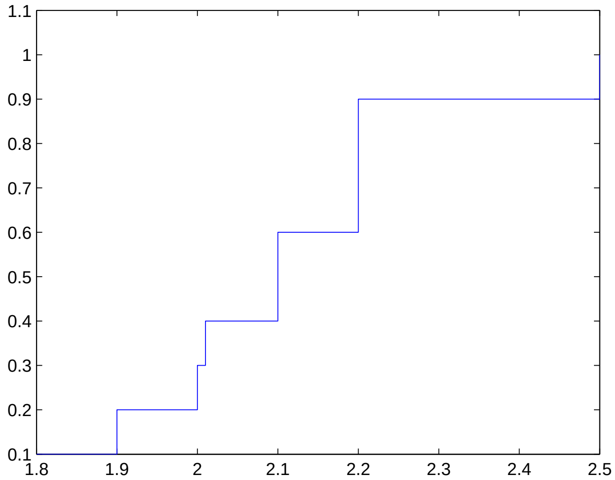
```
fs =

x/3 + 1/6

>> ex=int(x*fs,x,x1,x2) %пресмятаме математическото очакване

ex =

11/9

 >> dx=simple(int((x-ex)^2*fs,x,x1,x2))  %пресмятаме дисперсията

dx =

23/81
```

Пример 7.6. *За оценка на качеството на изделия, произвеждани от даден завод, се прави извадка с обем n=10, като се измерва теглото на изделията. Случайната величина X е теглото на едно изделие.*
а) Да се намери вариационният ред на извадката;
б) Да се построи талица на статистическия ред на извадката;
в) Да се начертае полигона на честотите.

```
>> x=[2.1 2.2 2 1.9 1.8 2.1 2.5 2.01 2.2 2.2] %извадка
x =
  2.1   2.2    2.0    1.9   1.8    2.1    2.5    2.01    2.2    2.2

>> broj=length(x); %обем на извадката
>> y=sort(x) %вариационния ред на извадката

y =
```

```
    1.8   1.9   2.0    2.01    2.1    2.1    2.2    2.2    2.2    2.5
..........................................
% построяване на статистическия ред

>> xx=y(1);
>> for i=2:broj
if y(i)~=y(i-1)
xx=[xx,y(i)];
end
end
>> xx  %елементи без повторения

xx =

    1.8    1.9    2.0    2.01    2.1    2.2    2.5

>> k=length(xx) %брой елементи без повторения
k =

     7

...................................
% абсолютни честоти
>> n=zeros(1,k);
>> for i=1:k,for j=1:broj
if y(j)==xx(i)
n(i)=n(i)+1;
end
end
end
```

```
>> n  %абсолютни честоти

n =

     1     1     1     1     2     3     1

>> l=n/broj %относителни честоти

w =

    0.1    0.1    0.1    0.1    0.2    0.3    0.1

>> ll=cumsum(l) %комулирани честоти
ll =

    0.1    0.2    0.3    0.4    0.6    0.9    1.0
................................
%полигон на честотите

>> stairs(xx,ll)
```



Фиг.5

Пример 7.7. *Нека разполагаме с 300 броя данни за сл.величина X - дължина на детайл, произвеждан от автомат, с нормален закон на разпределение. Да се групират данните при n=30.*
а) да се намерят извадковото средно и извадковата дисперсия;
б) да се начертае хистограмата на честотите;
в) да се намерят доверителни интервали с вероятност $\gamma = 0,9$ *за математическото очакване EX и дисперсията DX на генералната съвкупност.*

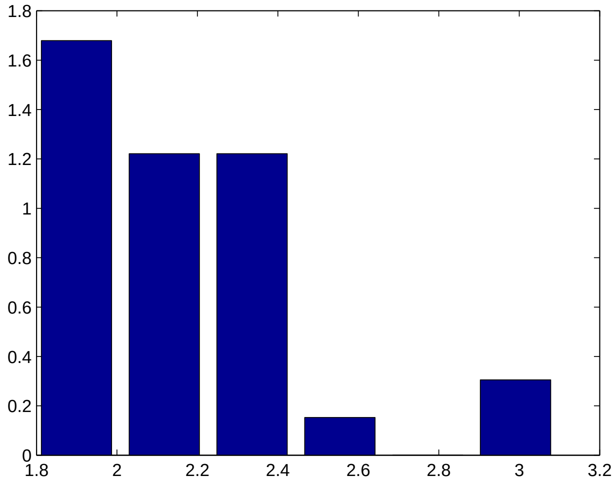
В командния ред на MATLAB се въвеждат командите:

```
>>danni=[2.2  2.3  2.01  2  2.31  2.33  2.12  3  3.1  2.23  1.98  2  1.99
 1.89  2.09  2.12   1.79  2  2.12  1.90  2  2.3
  2.15  1.88  2.3  2.1  2.5  2  2.4  2.3]
>> mean(danni) %средната стойност

ans =
```

```
    2.1803

>> std(danni) %стандартно отклонение

ans =

    0.2917

>> std(danni)^2 % извадкова дисперсия

ans =

    0.0851

>> min(danni);max(danni)
>> y=sort(danni)

y =

  Columns 1 through 10

    1.7900    1.8800    1.8900    1.9000    1.9800    1.9900
    2.0000    2.0000    2.0000    2.0000

  Columns 11 through 20

    2.0000    2.0100    2.0900    2.1000    2.1200    2.1200
    2.1200    2.1500    2.2000    2.2300

  Columns 21 through 30
```

```
    2.3000    2.3000    2.3000    2.3000    2.3100    2.3300
    2.4000    2.5000    3.0000    3.1000

>> m=6 % брой интервали

m =

     6

>> delta=y(30)-y(1) %размах на признака

delta =

    1.3100

>> d=delta/m %ширина на интервала

d =

    0.2183

>> i=0:m;g=y(1)+i*d %граници на интервалите

g =

    1.7900    2.0083    2.2267    2.4450    2.6633    2.8817    3.1000

>> [p,mint]=hist(y,m) % честота в интервали среда на интервала

p =
```

```
    11     8     8     1     0     2


mint =

    1.8992    2.1175    2.3358    2.5542    2.7725    2.9908

 >> w=p/30 % относителни честоти

w =

    0.3667    0.2667    0.2667    0.0333         0    0.0667

>> G=cumsum(w)% натрупани относителни честоти

G =

    0.3667    0.6333    0.9000    0.9333    0.9333    1.0000

>> F=w/d;
>> bar(xstar, F)%чертае хистограмата на вариационния ред
```



Фиг.6

```
.................................
%Намиране на доверителните интервали за EX и DX съответно
>> delta=1.699*std(danni)/sqrt(30)

delta =

    0.0905

>> left1=mean(danni)-delta
left1 =

    2.0899

>> right1=mean(danni)+delta

right1 =
```

```
    2.2708

>> left2=(29)*(std(danni)^2)/42.6

left2 =

    0.0579

>> right2=29*(std(danni)^2)/13.1

right2 =

    0.1883
```

Библиография

[1] Б. Димитров, Н. Янев. *Вероятности и Статистика.* София, 2007, ISBN 978-954-849-332.

[2] Р. Стайков. *Статистика.* Варна, 1999, ISBN 954-715-067-7.

[3] Й. Йорданов. *Matlab - Преобразувания, Изчисления, Визуализация, Част II.* Техника, София, 2006, ISBN 954-03-0661-2.

[4] Й. Стоянов, Й. Миразчийски, Ц. Игнатов, М. Танушев. *Ръководство по Теория на вероятностите.* София, 2001, ISBN 954-8495-23-6.

[5] К. Калинов. *Ръководство за решаване на задачи по Теория на вероятности и статистика.* София, 2004.

[6] D. Duffy. *Advanced Engineering Mathematics with MATLAB.* Chapman & Hall /CRC Press, 2003, ISBN 1-58488-349-9.

[7] В. Мещеряков. *Задачи по статистике и регрессионному анализу с MATLAB.* М.: Диалог, 2009, ISBN 978-5-86404-228-1.

[8] D. Higham, N. Higham. *MATLAB Guide.* SIAM, 2005, ISBN 0-89871-578-4.

[9] С. Иглин. *Теотия вероятностей и математическая статистика на базе MATLAB.*, Харков, ХГУ, 2006, ISBN 966-593-466-X.