Снежана Георгиева Гочева-Илиева

ПРИЛОЖНА МАТЕМАТИКА

http://fmi.uni-plovdiv.bg/gocheva/PM-lekcii.pdf

Пловдив, 2013

УНИВЕРСИТЕТСКО ИЗДАТЕЛСТВО
„ПАИСИЙ ХИЛЕНДАРСКИ“

Рецензент: Проф. д-р Николай Кюркчиев

ISBN 978-954-423-847-6

Този учебник е предназначен за обучение на студентите от специалност Бизнесинформационни технологии на Факултета по математика и информатика при Пловдивския университет „Паисий Хилендарски”.

Учебникът съдържа курс лекции по дисциплината „Приложна математика” за редовно и задочно обучение, 1ви курс БИТ, четен от автора през последните 4 години. Разработените теми съответстват на учебната програма. Всяка тема съдържа теоретична част и решени примери. Част от примерите са решени с помощта на специализираните софтуерни пакети *Mathematica* на фирмата Wolfram Research (www.wolfram.com) и *SPSS* на фирмата IBM (http://www-01.ibm.com/software/analytics/spss/). За последните могат да се ползват лицензираните копия на ФМИ или демонстрационните продукти на производителите, както и интернет онлайн машината http://www.wolframalpha.com/.

Всякакви бележки, коментари и запитвания можете да изпращате на електронен адрес: snow@uni-plovdiv.bg

Авторът

Съдържание:

1. Безкрайни числови редици – видове, свойства, сходимост, изчисляване границата на числова редица

Определение 1.1. Числова редица наричаме последователност от номерирани числа:

$$a_1, a_2, \ldots, a_n, \ldots$$

Означения: a_1 - 1-ви член, a_2 - 2-ри член, …, a_n - n-ти член, …;

n - пореден номер (индекс) на съответния член на редицата;

$\{a_n\}_{n=1}^{\infty}$ - общо означение на редицата.

Пример 1.1. $a_1 = 1,\ a_2 = \dfrac{1}{2},\ a_3 = \dfrac{1}{3}, \ldots, a_{100} = \dfrac{1}{100} \quad \to \quad \boxed{a_n = \dfrac{1}{n}}$

Пример 1.2. $1, \frac{1}{4}, \frac{1}{9}, \frac{1}{16}, \ldots, \frac{1}{n^2}, \ldots$, $a_n = \frac{1}{n^2}$ или $\left\{ \frac{1}{n^2} \right\}_{n=1}^{\infty}$.

Пример 1.3. $1, 3, 5, 7, 9, 11, \ldots$, $a_n = 2n - 1$ или $\left\{ 2n - 1 \right\}_{n=1}^{\infty}$.

Пример 1.4. $1, -1, 1, -1, 1, -1, \ldots$, $a_n = (-1)^{n+1}$ или $\left\{ (-1)^{n+1} \right\}_{n=1}^{\infty}$.

Пример 1.5. $\sqrt{3}, \sqrt{4}, \sqrt{5}, \ldots, \sqrt{n+2}, \ldots$, или $\left\{ \sqrt{n+2} \right\}_{n=1}^{\infty}$.

Пример 1.6. $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \ldots, \frac{n}{n+1}, \ldots, a_n = \frac{n}{n+1}$ или $\left\{ \frac{n}{n+1} \right\}_{n=1}^{\infty}$.

Начини на задаване на числова редица:

- Аналитично (с формула): $b_n = \frac{2n+3}{n^2+1}$
- Описателно (с думи): "Всички членове са равни на 3." Т.е. редицата е: 3, 3, 3, 3, 3, 3, 3…
- Рекурентно: Членовете на редицата се изчисляват от едно или няколко предишни: $a_1 = 1,\ a_2 = 1,\ \ldots,\ a_{n+1} = a_n + a_{n-1}, \ldots.$ Тази редица е прочута в математиката и се нарича редица на Фибоначи. Изчисляваме началните й членове и получаваме: 1, 1, 2, 3, 5, 8, 13, 21, 34, …

Намирането на аналитична формула обаче е трудно. Формулата е:

$$a_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n + \left(\frac{1-\sqrt{5}}{2}\right)^n\right].$$

Фибоначи е италиански математик от Пиза от 13 век. Той е известен не само като най-големия математик на Средновековието, но и с това, че е въвел използването на арабските цифри и десетичната бройна система в Европа и е положил началото на съвременното счетоводство и финанси.

Аритметични действия с числови редици – поелементно (почленно):

- Сума на редици: $z_n = x_n + y_n$, $n = 1, 2, 3, \ldots$
- Разлика: $z_n = x_n - y_n$, $n = 1, 2, 3, \ldots$
- Произведение:, $z_n = x_n . y_n$, $n = 1, 2, 3, \ldots$
- Деление: $z_n = x_n / y_n$, $n = 1, 2, 3, \ldots$
- Умножение с число: $z_n = \lambda x_n$, $n = 1, 2, 3, \ldots$, λ е число.

Свойства на числовите редици

Определение 1.2. Редицата е нарастваща (монотонно растяща), когато всеки следващ член е по-голям или равен на предишния:

$$x_1 \le x_2 \le x_3 \le \ldots \le x_n \le x_{n+1} \le \ldots$$

Пример 1.7. 0, 2, 4, 6, 8,… е строго нарастваща редица.

Пример 1.8. $\frac{1}{10}, \frac{1}{100}, \frac{1}{1000}, \frac{1}{10000}, \ldots$ е строго намаляваща редица.

Пример 1.9. Редицата $1, 0, 2, 0, 3, 0, \ldots$ е нито нарастваща, нито намаляваща.

Пример 1.10. Аритметичната прогресия $1,\ 0.8,\ 0.6,\ 0.4, \ldots$ е намаляваща с разлика $d = -0.2$.

Определение 1.3. Числова редица наричаме ограничена, ако нейните членове не стават безкрайно големи или безкрайно малки, т.е. съществува положителна константа $M > 0$:

$$|a_n| \leq M \text{, за всяко } n.$$

В противен случай редицата е неограничена.

Лесно се вижда, че редиците от примерите 1.2, 1.2, 1.4, 1.6, 1.8 са ограничени, а редиците от примерите 1.3, 1.5, 1.7, 1.9, 1.10 са неограничени.

Каква според вас е редицата на Фибоначи?

Граница на числова редица

Определение 1.4. Редица, чиито членове при безкрайно нарастване на n стават все по-близки до нула, че нарича безкрайно малка редица.

Такава безкрайно малка редица е например т.н. хармонична редица (виж Пример 1.1.):

$$1, \frac{1}{2}, \frac{1}{3}, ..., \frac{1}{n}, ... \; .$$

В този случай казваме, че редицата е сходяща към нула и нулата е нейната граница. Това се записва така:

$\lim\limits_{n\to\infty}\frac{1}{n}=0$ (четем "Лимес от едно върху n, при n клонящо към безкрайност, е равен на нула").

Безкрайно малки редици са тези от Пр. 1.1, 1.2, 1.8.

Аналогично, в общия случай имаме следното

Определение 1.5. Ако при безкрайно нарастване на индекса n членовете на редицата се стремят по стойност към някакво крайно число a, казваме, че редицата е сходяща и нейната граница е равна на a и бележим с $\lim\limits_{n\to\infty} a_n = a$ (четем "Лимес от a_n, при n клонящо към безкрайност, е равен на a").

Определение 1.6. Ако при безкрайно нарастване на индекса n има безкрайно големи (малки) членове на редицата, тя е разходяща.

Намирането на граница на дадена числова редица може да се окаже трудна задача. Ние ще използваме за простите примери по-долу дадените теореми, а за по-сложните ще използваме компютър.

Теорема 1. Ако една редица е сходяща, тя е ограничена.

Теорема 2. Ако редицата е монотонна и ограничена, тя е сходяща.

Теорема 3. Ако редицата $\{a_n\}_{n=1}^{\infty}$ е сходяща, то са сходящи и редиците: $\{\lambda.a_n\}_{n=1}^{\infty}$, $\{|a_n|\}_{n=1}^{\infty}$, $\{a_n+\lambda\}_{n=1}^{\infty}$ (λ е константа).

Теорема 4. Ако две редици $\{a_n\}_{n=1}^{\infty}$ и $\{b_n\}_{n=1}^{\infty}$ са сходящи, т.е. съществуват границите им $\lim_{n\to\infty} a_n = a$ и $\lim_{n\to\infty} b_n = b$, то тогава са сходящи тяхната сума, разлика, произведение и частно (ако $b_n \neq 0$, $b \neq 0$, при което:

А) $\lim\limits_{n\to\infty}(a_n + b_n) = a + b,$ Б) $\lim\limits_{n\to\infty}(a_n - b_n) = a - b,$

В) $\lim\limits_{n\to\infty}(a_n . b_n) = a.b,$ Г) $\lim\limits_{n\to\infty}(a_n / b_n) = a / b.$

Теорема 5. (за сравнение). Нека две редици $\{a_n\}_{n=1}^{\infty}$ и $\{b_n\}_{n=1}^{\infty}$ са сходящи и съществуват границите им $\lim\limits_{n\to\infty} a_n = a$ и $\lim\limits_{n\to\infty} b_n = b$ и за всяко n за членовете на трета редица $\{c_n\}_{n=1}^{\infty}$ е изпълнено неравенството: $a_n \leq c_n \leq b_n$. Тогава редицата $\{c_n\}_{n=1}^{\infty}$ е също сходяща и нейната граница се намира между тези на другите две, т.е. $a \leq c \leq b$.

Определение 1.7. Ако от дадена редица се премахнат някои или безкраен брой членове, получаваме т.н. подредица на дадената.

Теорема 6. Ако една редица е сходяща към граница a, то всяка нейна подредица също е сходяща и то към същата граница a.

Пример 1.11. Редицата $\frac{1}{2}, \frac{1}{4}, \frac{1}{6} \ldots, \frac{1}{2n}, \ldots$ е подредица на 1.1 и също е сходяща към 0.

Намиране граници на числови редици

Като използвме горните теореми и свойства на сходящите редици и някои известни граници, можем да намираме границите на някои редици. Ще илюстрираме това с помощта на следните примери:

Пример 1.12. Да се намери границата на редиците:

А) $\lim\limits_{n\to\infty}\dfrac{n}{2n+1}$; Б) $\lim\limits_{n\to\infty}\dfrac{n^2-3n+1}{3n^2-2n}$;

В) $\lim\limits_{n\to\infty}\dfrac{n}{\sqrt{7n+1}}$; Г) $\lim\limits_{n\to\infty}\dfrac{1}{\sqrt{n+1}-\sqrt{n-1}}$.

Решение:

А) Числителят и знаменателят клонят едновременно към безкрайност, затова особеността е в безкрайност. Ще се опитаме да съкратим тази особеност. Затова изнасяме безкрайното *n* пред скоби в числителя и знаменателя и го съкращаваме:

$$\lim_{n\to\infty}\frac{n}{2n+1}=\lim_{n\to\infty}\frac{n}{n(2+\frac{1}{n})}=\lim_{n\to\infty}\frac{1}{(2+\frac{1}{n})}$$

$$=\frac{1}{(2+\lim\limits_{n\to\infty}\frac{1}{n})}=\frac{1}{(2+0)}=\frac{1}{2}$$

Б) Аналогично изнасяме пред скоби “безкрайните” множители за да ги съкратим:

$$\lim_{n\to\infty}\frac{n^2-3n+1}{3n^2-2n}=\lim_{n\to\infty}\frac{n^2(1-\frac{3}{n}+\frac{1}{n^2})}{n^2(3-\frac{2}{n})}$$

$$=\lim_{n\to\infty}\frac{(1-\frac{3}{n}+\frac{1}{n^2})}{(3-\frac{2}{n})}=\frac{1-0-0}{3-0}=\frac{1}{3}.$$

в) $$\lim_{n\to\infty}\frac{n}{\sqrt{7n+1}}=\lim_{n\to\infty}\frac{n}{\sqrt{n(7+\frac{1}{n})}}=\lim_{n\to\infty}\frac{n}{\sqrt{n}\sqrt{(7+\frac{1}{n})}}$$

$$=\frac{\lim\limits_{n\to\infty}\sqrt{n}}{\sqrt{(7+\lim\limits_{n\to\infty}\frac{1}{n})}}=\frac{\lim\limits_{n\to\infty}\sqrt{n}}{\sqrt{7+0}}=\frac{1}{\sqrt{7}}.\lim_{n\to\infty}\sqrt{n}\to\infty\ ;$$

$$\text{г) } \lim_{n\to\infty}\frac{1}{\sqrt{n+1}-\sqrt{n-1}}=\lim_{n\to\infty}\frac{\sqrt{n+1}+\sqrt{n-1}}{\left(\sqrt{n+1}-\sqrt{n-1}\right)\left(\sqrt{n+1}+\sqrt{n-1}\right)}$$

$$=\lim_{n\to\infty}\frac{\sqrt{n+1}+\sqrt{n-1}}{\left(\sqrt{n+1}\right)^2-\left(\sqrt{n-1}\right)^2}=\lim_{n\to\infty}\frac{\sqrt{n+1}+\sqrt{n-1}}{n+1-(n-1)}$$

$$=\lim_{n\to\infty}\frac{\sqrt{n}}{2}\left(\sqrt{1+\frac{1}{n}}+\sqrt{1-\frac{1}{n}}\right)=\lim_{n\to\infty}\frac{\sqrt{n}}{2}\left(\sqrt{1+\lim_{n\to\infty}\frac{1}{n}}+\sqrt{1-\lim_{n\to\infty}\frac{1}{n}}\right)$$

$$=\lim_{n\to\infty}\frac{\sqrt{n}}{2}\left(\sqrt{1+0}+\sqrt{1-0}\right)=\lim_{n\to\infty}\frac{\sqrt{n}}{2}(1+1)=\lim_{n\to\infty}\sqrt{n}=\infty\,.$$

Решение с *Mathematica*: Разпечатваме първите 20 члена за проследим началните стойности и поведението на редицата. Виждаме, че е нарастваща. Границата изчисляваме с вградената функция **Limit[]**:

$$\text{Table}\left[\frac{1}{\sqrt{n+1}-\sqrt{n-1}},\ \{n,\ 20\}\right]\ //\ \text{N}$$

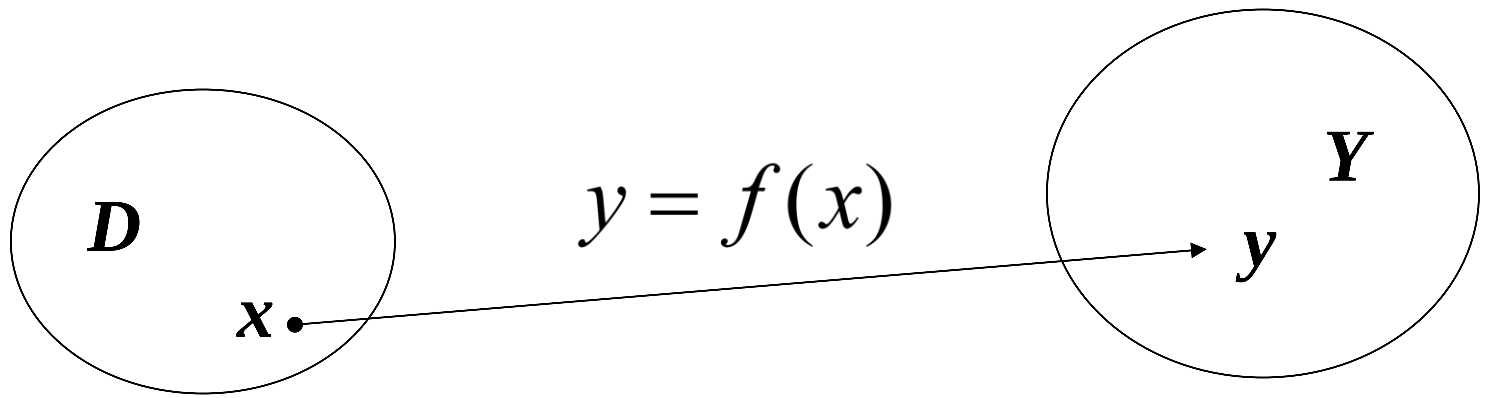
```
{0.707107, 1.36603, 1.70711, 1.98406, 2.22474,
 2.44091, 2.63896, 2.82288, 2.99535, 3.15831,
 3.31319, 3.46109, 3.60288, 3.73927, 3.87083,
 3.99804, 4.12132, 4.241, 4.35739, 4.47074}
```

$$\text{Limit}\left[\frac{1}{\sqrt{n+1}-\sqrt{n-1}},\ n\to\infty\right]$$

$$\infty$$

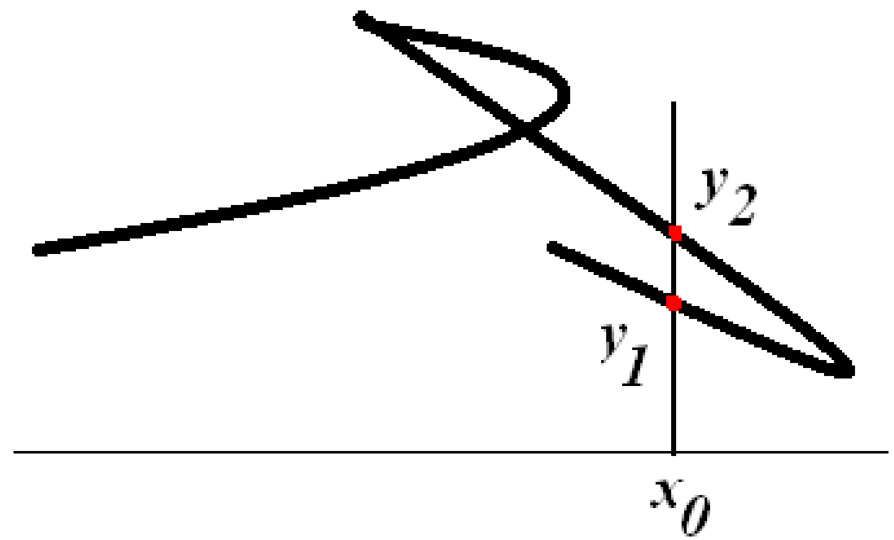
2. Функции – определение, свойства, граница и непрекъснатост на функции. Обратни тригонометрични функции (аркус-функции)

Определение 2.1. Функцията $y = f(x)$ е правило, по което на всеки елемент $x \in D$ (четем "хикс принадлежи на де голямо"), от едно числово множество $\boldsymbol{D}$ се съпоставя точно един елемент $y \in Y$ на друго числово множество $\boldsymbol{Y}$. Използва се още означението $f : D \to Y$.



Означения: D - дефиниционно множество на функцията, Y - множество на значенията. Променливата x се нарича независима величина (аргумент) на функцията, а променливата y се нарича зависима променлива.

Пример 2.1 Следната графика НЕ е графика на функция, защото на един и същ аргумент x съответстват повече от една стойност, в случая за x_0 има 2 стойности: y_1 и y_2.



Начини на задаване на функции:

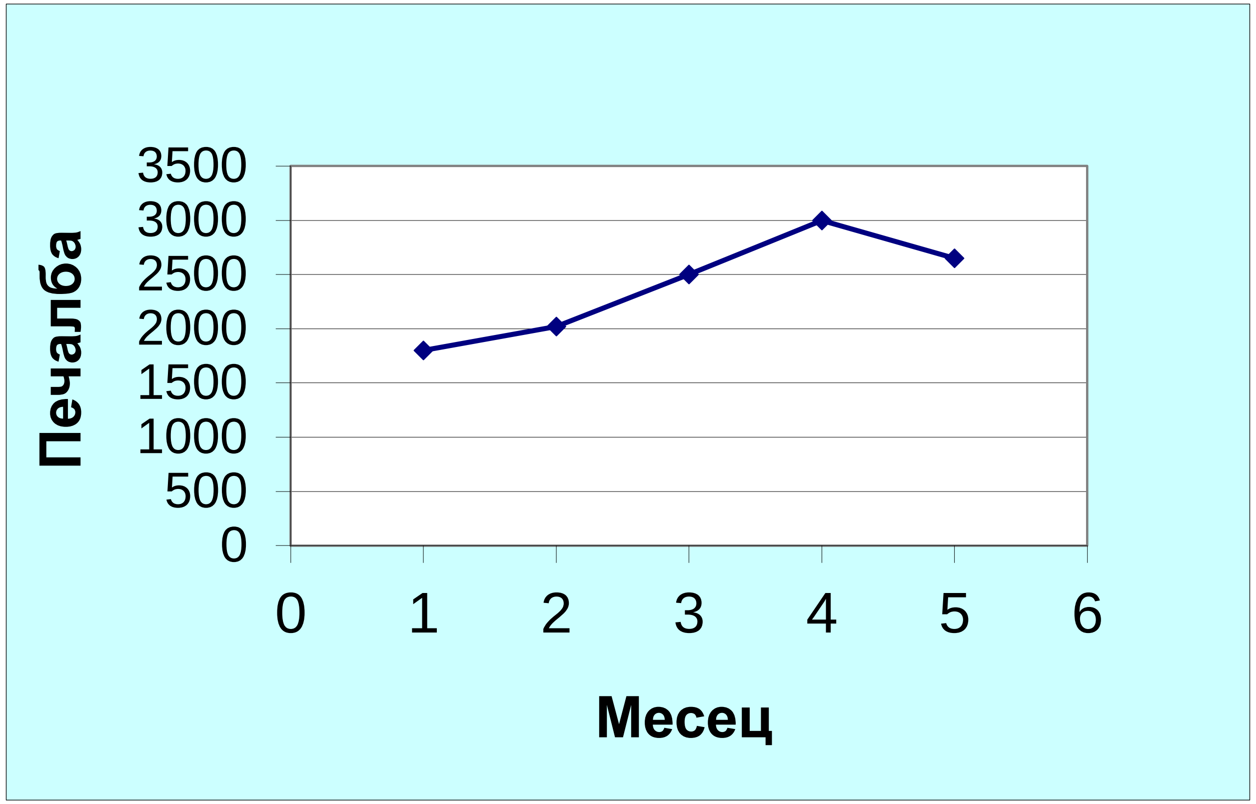
- Аналитично (с формула): напр. $y = \dfrac{x+3}{x^2-1}$ (**)

Очевидно точките $x = \pm 1$ не са от дефиниционната област на тази функция.

- Таблично задаване:

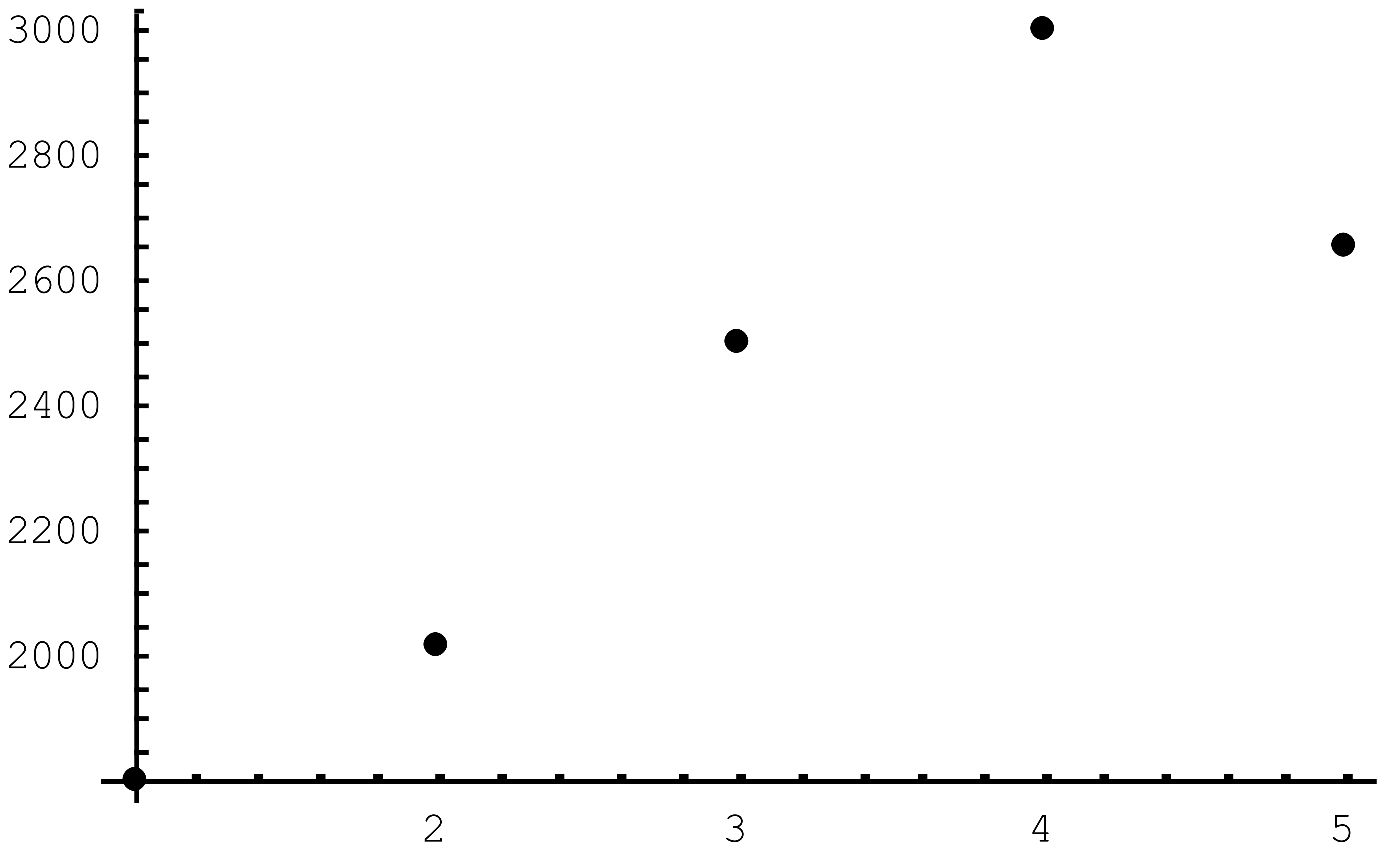
месеци	януари	февруари	март	април	май
***x*-**	1	2	3	4	5
***y* - печалба**	1800	2020	2500	3000	2650

- Графично: а) за горния пример с таблично зададена функция

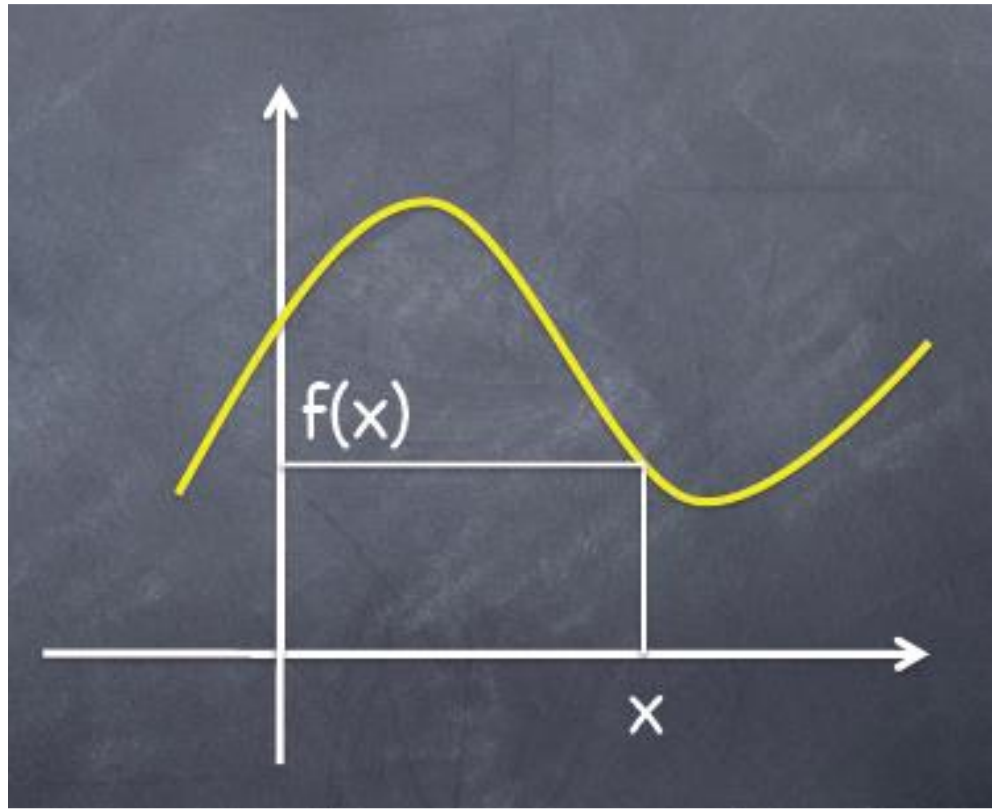


- Същото със система *Mathematica*:

```
xx = {1, 2, 3, 4, 5}
yy = {1800, 2020, 2500, 3000, 2650}
data = Table[{xx[[i]], yy[[i]]}, {i, 1, 5}]
ListPlot[data, PlotStyle → PointSize[0.02]]
```

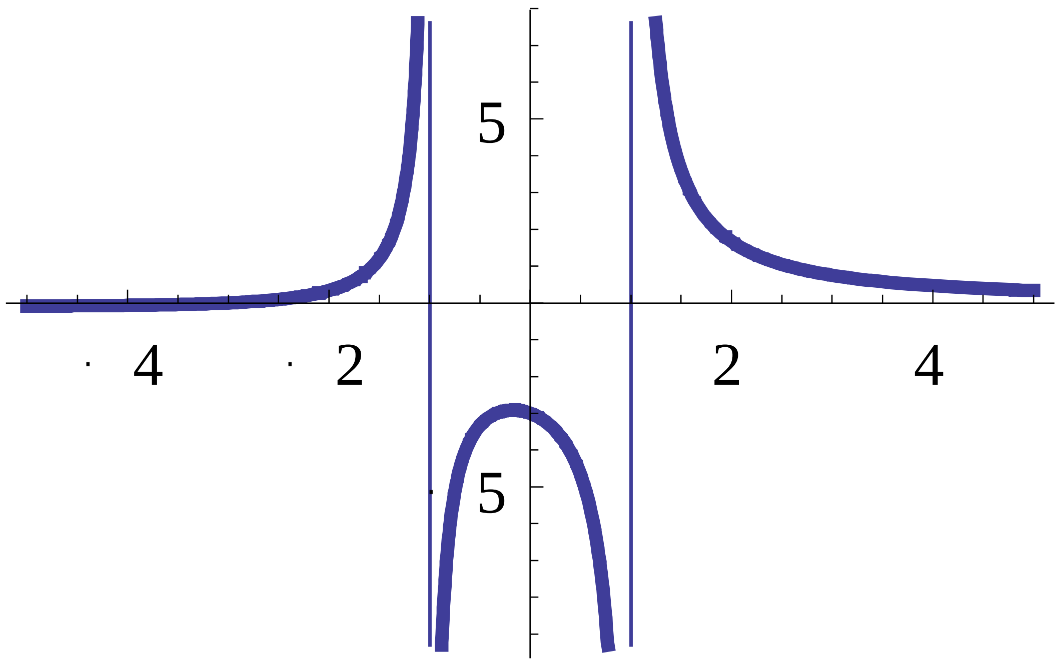


б) Функция, зададена за x в някакъв интервал. Тогава всяка точка с координати (x, y) се чертае с една точка, като y се изчислява (измерва) от формулата или правилото $y = f(x)$.



Пример 2.2 Със система Математика да начертаем функцията (**):

$$y = \frac{x+3}{x^2-1}; \text{ Plot}[y, \{x, -5, 5\}]$$



Получаваме дефиниционна област:

$$D = (-\infty, -1) \cup (-1, 1) \cup (1, \infty)$$

Леонард Ойлер е един от най-известните математици. Живял е през 18 век. Има голям принос в изучаването на функциите, важна част от математическия анализ.

Определяне на дефиниционната област (ДО) на функция

За да се определи ДО на дадена функция трябва да се има в предвид:

а) къде функционалната стойност не може да се изчисли и

б) дали е зададен отнапред някакъв интервал за ДО.

В първия случай, от ДО трябва да се изключат:

- точките, в които знаменателят става 0, тъй като на нула не се дели;
- областите, за които не съществува квадратен корен, т.е. при наличие на израз от вида $\sqrt{A}$, трябва да се намери относно x в кои интервали е валидно неравенството $A<0$; същото важи и за корен четвърти, шести и т.н.
- областите, за които не съществува логаритъм, тъй като $Log(B)$ съществува само при $B>0$ и др.

- да се изключат случаите, при които задачата няма реален смисъл, напр. времето винаги е неотрицателна величина, температурата при някой химичен процес не може да е по-ниска от $5^{o}C$ и др.

Пример 2.3 Да се определят ДО на функциите:

а) $f(x)=\dfrac{\sqrt{2x-3}}{x-1}$, б) $f(x)=\ln\left(\dfrac{x}{x+1}\right)$, където $\ln(.)$ е натурален логоритъм.

Решение:
а) При $x=1$ знаменателят се анулира. Освен това трябва подкоренната величина да е неотрицателна, т.е. $2x-3\geq 0$. Получаваме условията: $x\geq\dfrac{3}{2}, x\neq 1$, откъдето ДО=$x\geq\dfrac{3}{2}$, т.е. $D=[\dfrac{3}{2},\infty)$.

б) Логаритъмът е определен само при положителен аргумент, т.е. трябва $\dfrac{x}{x+1} > 0$. Или което същото, $x(x+1) > 0$. Това е квадратно неравенство. Уравнението $x(x+1) = 0$ има корени $x_1 = 0,\ x_2 = -1$. Произведението е положително в интервалите $(-\infty, -1) \cup (0, \infty)$. Точката -1 трябва да се изключи, защото знаменателят не може да е нула.

Окончателно отговорът е: $D = (-\infty, -1) \cup (0, \infty)$.

Свойства на функциите

Определение 2.2. Една функция е ограничена отгоре, ако за всяко $x \in D$ стойностите на функцията не стават безкрайно големи, т.е. съществува положителна константа $M > 0$:

$$f(x) \leq M, \text{ за всяко } x \in D.$$

Определение 2.3. Една функция е ограничена отдолу, ако за всяко $x \in D$ стойностите на функцията не стават безкрайно малки (напр. минус безкрайност), т.е. съществува константа m:

$$f(x) \geq m, \text{ за всяко } x \in D.$$

Определение 2.4. Една функция е ограничена, ако за всяко $x \in D$ стойностите на функцията не стават нито безкрайно големи нито безкрайно малки, т.е. съществува положителна константа $M > 0$:

$$\left| f(x) \right| \leq M \text{, за всяко } x \in D.$$

В противен случай функцията е неограничена.

Пример 2.4. Очевидно функцията от Пример 2.2. е неограничена.

Граница на функция

Пример 2.5 (основен). По аналогия с редиците, да разгледаме най-напред функцията

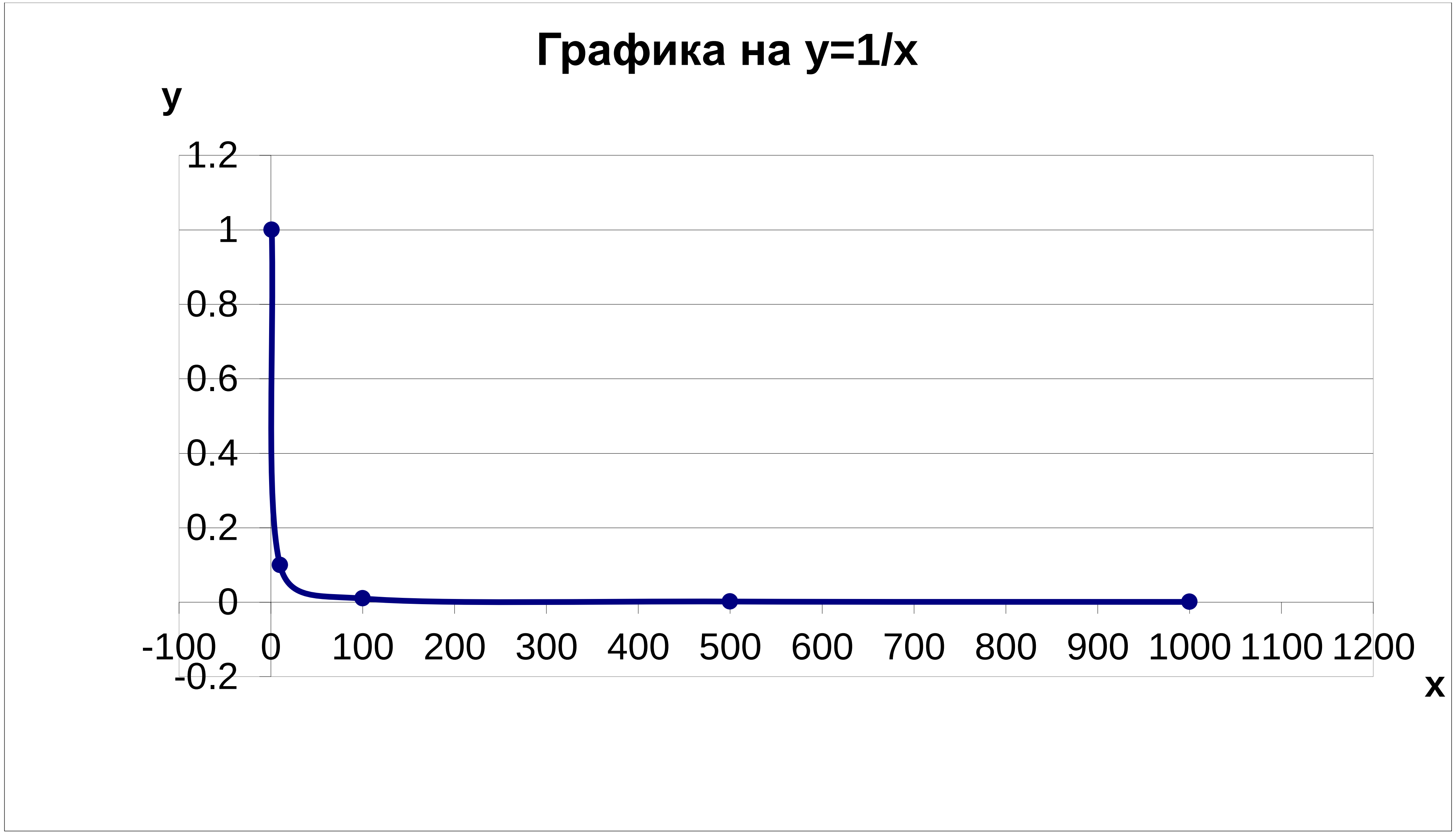
$$y(x) = \frac{1}{x}$$

при $x \to \infty$. Очевидно когато знаменателят расте (става много голям), тази дроб ще има стойности много близки до нулата, както и да избираме стойностите за x (освен нула). В този случай казваме, че функцията е сходяща към нула и нулата е нейната граница. Това се записва така:

$$\lim_{x \to \infty} \frac{1}{x} = 0$$

(четем "Лимес от едно върху x, при x клонящо към безкрайност, е равен на нула"). Това се вижда и от следната таблица с няколко различни x:

x	
1	1
10	0,1
100	0,01
00	0,002
1000	0,001



Определение 2.5. Функцията $f(x)$ <u>има граница</u> $\lim\limits_{x\to x_0} f(x)=L$, при x клонящо към дадена (крайна) т. x_0, когато за всяко положително число ε съществува положително число δ, такова че за всяко x, за което $0<\left|x-x_0\right|<\delta$, функцията $f(x)$ е дефинирана и $\left|f(x)-L\right|<\varepsilon$.

Да припомним, че неравенството $0<\left|x-x_0\right|<\delta$ означава, че $x_0-\delta<x<x_0+\delta$, т.е. $x\in(x_0-\delta,x_0+\delta)$. Тогава определението 2.5. означава, че когато за всяко x, което е в достатъчно малък интервал около x_0 съответната стойност на функцията $f(x)$ е достатъчно близо до L (в малък интервал около L), функцията $f(x)$ има граница L.

Определение 2.6. Функцията $f(x)$ <u>има граница</u> $\lim\limits_{x\to\infty} f(x) = L$, при x клонящо към $+\infty$, когато за всяко положително число ε съществува положително число N, такова че за $x > N$ функцията $f(x)$ е дефинирана и $\left|f(x) - L\right| < \varepsilon$.

Аналогично се определя границата на функция при x клонящо към $-\infty$.

Когато граница не съществува, а функцията расте (намалява) към безкрайност (минус безкрайност), можем да пишем условно сходимост към $\pm\infty$.

Например: $\lim\limits_{x\to\pm 0} \dfrac{1}{x} = \pm\infty$.

Аритметични действия за граници на функции

Те са напълно аналогични на действията със сходящи редици.

Ако съществуват границите: $\lim\limits_{x\to x_0} f(x) = L$ и $\lim\limits_{x\to x_0} g(x) = M$, а λ е реално число, то:

А) $\lim\limits_{x\to x_0}\left[f(x) \pm g(x)\right] = \lim\limits_{x\to x_0} f(x) \pm \lim\limits_{x\to x_0} g(x) = L \pm M$ -сума или разлика

Б) $\lim\limits_{x\to x_0} \lambda f(x) = \lambda \lim\limits_{x\to x_0} f(x) = \lambda L$ - умножение с константа

В) $\lim\limits_{x\to x_0} f(x).g(x) = \lim\limits_{x\to x_0} f(x).\lim\limits_{x\to x_0} g(x) = L.M$ - умножение

Г) $\lim\limits_{x\to x_0} \dfrac{f(x)}{g(x)} = \lim\limits_{x\to x_0} f(x) / \lim\limits_{x\to x_0} g(x) = \dfrac{L}{M},\ g(x) \neq 0, M \neq 0$ - деление

Тук x_0 може да бъде крайно число, $+\infty$ или $-\infty$.

Някои забележителни граници на функции

$$\lim_{x\to 0}\frac{\sin x}{x}=1,\qquad \lim_{x\to\pm\infty}\left(1+\frac{1}{x}\right)^{x}=e,\qquad \lim_{x\to\infty}\left(1+\frac{k}{x}\right)^{x}=e^{k}$$

$$\lim_{x\to 0}\frac{\arcsin x}{x}=1,\qquad \lim_{x\to 0}\frac{\operatorname{arctg} x}{x}=1,\qquad \lim_{x\to 0}\frac{\ln(1+x)}{x}=1,$$

$$\lim_{x\to 0}\frac{e^{x}-1}{x}=1.$$

В изчисленията тези граници могат да се ползват наготово.

<u>**Забележка**.</u> Определения на обратните функции *arcsin*, *arctg* и други са дадени в края на лекцията.

Пример 2.6 Намерете границите на функциите:

А) $\lim\limits_{x\to 1}\dfrac{(2x^2+7)}{x+1}=\dfrac{\lim\limits_{x\to 1}(2x^2+7)}{\lim\limits_{x\to 1}(x+1)}=\dfrac{2.1^2+7}{1+1}=\dfrac{9}{2}$, няма никаква особеност.

Заместваме директно $x=1$ и изчисляваме по правилата за граници.

Б) $\lim\limits_{x\to 1}\dfrac{(2x^2+3)}{x-1}=\dfrac{\lim\limits_{x\to 1}(2x^2+3)}{\lim\limits_{x\to 1}(x-1)}=\dfrac{2.1^2+3}{\lim\limits_{x\to 1}(x-1)}=\left[\dfrac{5}{0}\right]$- Границата не съществува, защото числителят е ограничен (равен на 5), а знаменателят клони към нула, следователно дробта може да става $\pm\infty$ при $x\to 1$.

в) $\lim\limits_{x \to 1} \dfrac{(2x^2 + x - 3)}{x - 1} = \left[\dfrac{0}{0}\right]$, т.е. получаваме нула и в числителя и в знаменателя при $x = 1$. Значи $x = 1$ е корен както на числителя, така и на знаменателя. За да решим задачата ще изолираме множител $\left(x - 1\right)$, с цел да го съкратим за $x \neq 1$. Наистина решаваме квадратното уравнение за числителя и получаваме корени $x_1 = 1, x_2 = \dfrac{-3}{2}$. Оттам след разлагането на числителя на дроби, намираме:

$$\lim_{x \to 1} \frac{(2x^2 + x - 3)}{x - 1} = \lim_{x \to 1} \frac{(2x + 3)(x - 1)}{x - 1} = \lim_{x \to 1} (2x + 3) = 2.1 + 3 = 5.$$

Тук сме съкратили “особеността” $\left(x - 1\right)$.

Г) $$\lim_{x\to\pm\infty}\frac{3x^2+2x+1}{x^2-x+5}=[\frac{\infty}{\infty}]=\lim_{x\to\pm\infty}\frac{x^2\left(3+\frac{2}{x}+\frac{1}{x^2}\right)}{x^2\left(1-\frac{1}{x}+\frac{5}{x^2}\right)}=\lim_{x\to\pm\infty}\frac{3}{1}=3.$$

Особеността е в безкрайност. Затова изваждаме като множител x^2, което дава особеността, за да я съкратим. В скобите при $x\to\pm\infty$ използваме основната граница $\lim\limits_{x\to\pm\infty}\frac{1}{x}=0$.

Д) $\lim\limits_{x \to +\infty} x\left(\sqrt{x^2+1}-x\right)=[\infty].[0].$

Ще рационализираме втория множител. Получаваме:

$$=\lim_{x \to +\infty} \frac{x\left(\sqrt{x^2+1}-x\right).\left(\sqrt{x^2+1}+x\right)}{\left(\sqrt{x^2+1}+x\right)}=\lim_{x \to +\infty} \frac{x\left(x^2+1-x^2\right)}{\left(\sqrt{x^2+1}+x\right)}$$

$$=\lim_{x \to +\infty} \frac{x}{x\left(\sqrt{1+\dfrac{1}{x^2}}+1\right)}=\frac{1}{1+1}=\frac{1}{2}$$

.

Тук в последния ред отново е изнесен пред скоби множителят, който дава особеността ($x \to \infty$), за да го съкратим.

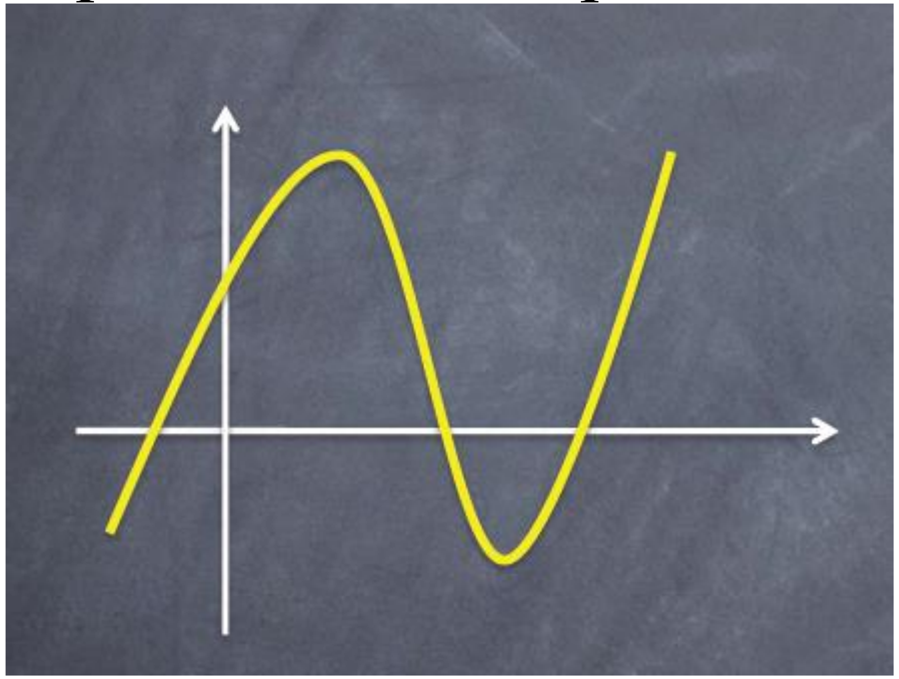
Е) Пример за използване на забележителни граници

$$\lim_{x\to 0}\frac{\sin 3x}{\sin 4x}=\lim_{x\to 0}\frac{\sin 3x}{3.4.x}.\frac{3.4.x}{\sin 4x}=\frac{1}{4}\lim_{x\to 0}\frac{\sin 3x}{3x}.3\lim_{x\to 0}\frac{4x}{\sin 4x}.$$

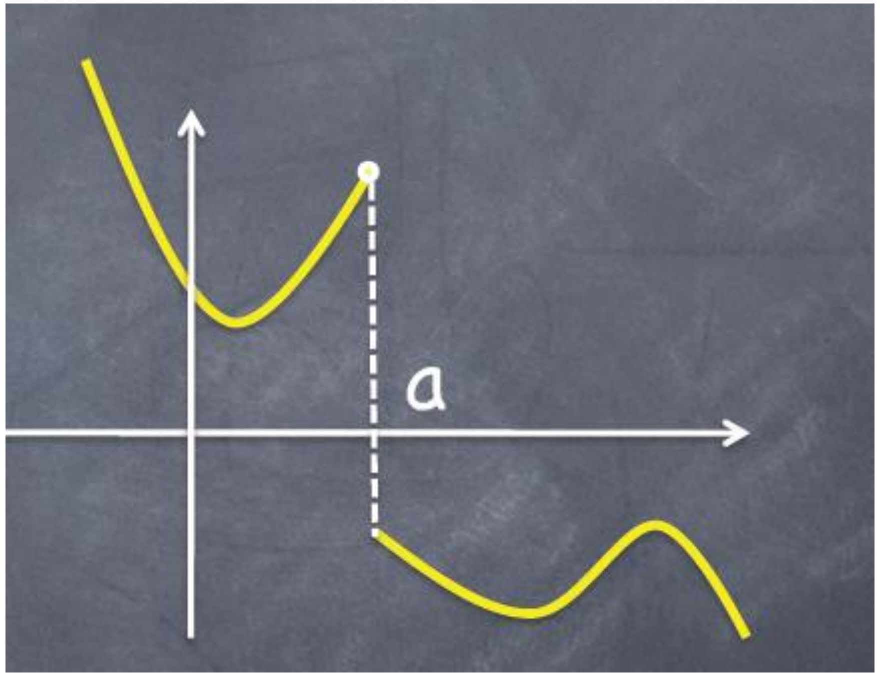
$$=\frac{3}{4}.1.1=\frac{3}{4}$$

Определение 2.7. Функцията $f\left(x\right)$ е <u>непрекъсната</u> в крайна т. $x_0 \in D$, когато съществува границата $\lim\limits_{x \to x_0} f(x)$ и $\lim\limits_{x \to x_0} f(x) = f(x_0)$.

Непрекъснатост и прекъснатост на функция --→ примерни графики



Непрекъсната функция във всяка точка от D.



Прекъсната функция в точката ***x=a,*** в останалите точки е непрекъсната.

Пример 2.7 Изследвайте за непрекъснатост функцията $f(x)$ в точката $x_0 = 0$:

$$f(x) = \begin{cases} x\sin\left(\dfrac{1}{x}\right), & x \neq 0 \\ 0 & x = 0 \end{cases}.$$

Решение:
Дефиниционната област е: $(-\infty, \infty)$.

Случай а) нека $x_0 \neq 0$. Тогава имаме $f(x) = x\sin\left(\dfrac{1}{x}\right)$. Границата може да се намери по теоремите за действия с граници на функции:

$$\lim_{x\to x_0} f(x) = \lim_{x\to x_0}\left(x\sin\frac{1}{x}\right) = \lim_{x\to x_0} x.\lim_{x\to x_0}\sin\frac{1}{x} = x_0\sin\left(\lim_{x\to x_0}\frac{1}{x}\right).$$

$$= x_0\sin\frac{1}{x_0} = f(x_0)$$

Всички пресмятания са възможни, защото $x_0 \neq 0$ и делението е възможно. Напр. ако $x_0 = 0$, с калкулатор или компютър изчисляваме

```
Sin[1.]
0.841471
```

Ако $x_0 = 1^0$ в градуси, а не в радиани, тогава:

```
dd = 1. Degree
dd * Sin[1/dd]
0.0174533
0.0118599
```

Следователно : $\lim\limits_{x\to 1^0}(x\sin\frac{1}{x}) = 1^0 \sin\frac{1}{1^0} = 0.0118599.$

<u>Случай б)</u> нека $x_0 = 0$. Тогава имаме $f(0) = 0$. Ще установим най-напред, че съществува граница при $x \to 0$ и след това тя трябва да е равна на 0, за да имаме непрекъснатост по горното определение 2.5.

Ако се опитаме директно да заместим с 0, ще получим $0.\sin\frac{1}{0} = 0.\sin(\infty)$. Тук обаче забелязваме, че синусът както и да се променя е ограничен, тъй като не може да стане по абсолютна стойност по-голям от 1, за реални x. Или $\left|\sin(t)\right| \le 1$ за всяко реално $t \in (-\infty, \infty)$.

Тогава решаваме задачата елементарно така. От неравенството

$$0 \le \left| f(x) \right| = \left| x\sin\frac{1}{x} \right| \le |x|.\left|\sin\frac{1}{x}\right| \le |x|.1$$

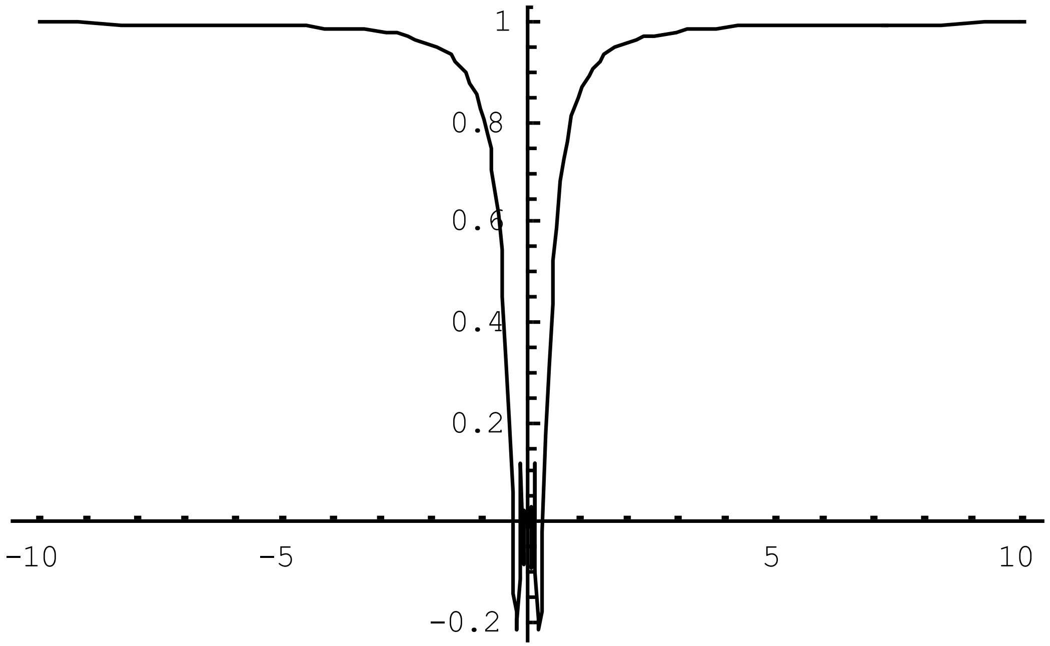
следва, че при $x \to 0$ с прилагане на теоремата за сравнение на граници (лекция 1, слайд 10-11): $0 \le \lim\limits_{x\to 0} \left| f(x) \right| \le \lim\limits_{x\to 0} |x| .1 = 0$.

Следователно $\lim\limits_{x\to 0} f(x) = 0 = f(0)$. Тъй като границата съществува и е равна на 0, колкото е $f(0)$, заключаваме, че функцията е непрекъсната в точката 0.

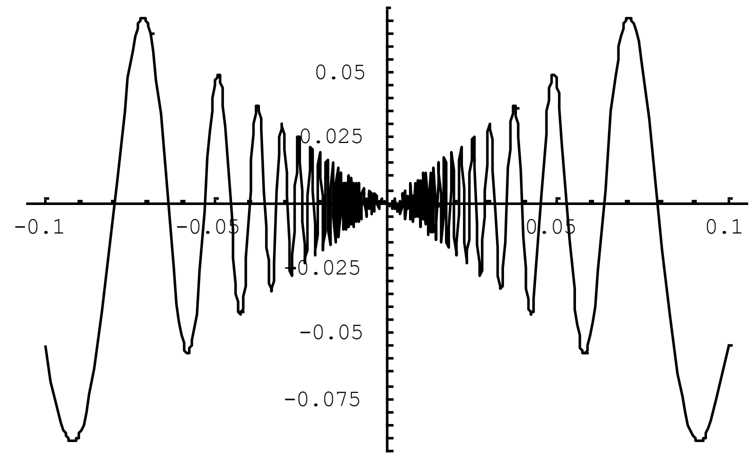
<u>От случаите а) и б) следва, че функцията $f(x)$ е непрекъсната за всяко x от дефиниционната си област.</u>

С помощта на система *Mathematica* нека начертаем функцията, за да видим как изглежда: отначало в интервала [-10; 10], а след това и близо около нулата – в интервала [-0.1; 0.1]. Очевидно функцията става 0 близо до $x = 0$ осцилирайки около 0 и цялата графика е с поведение на непрекъсната функция.

Plot[x * Sin[1/x], {x, -5, 5}]



Plot[x * Sin[1/x], {x, -0.1, 0.1}]



Освен това, директно намирайки границата, виждаме, че тя е 0:

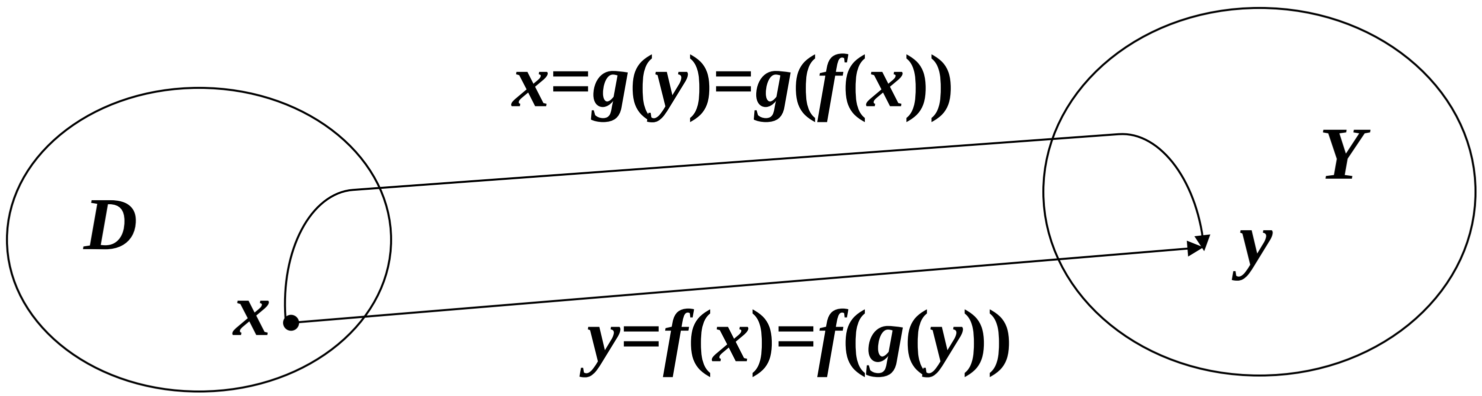
Limit[x * Sin[1/x], x → 0]

0

Обратна функция

Има случаи на функции $y=f(x)$, при които е възможно и задаване на функция $g(y)$, съпоставяща на всяко y първоначалното x. Това е възможно, ако на различни $x \in D$ съответстват различни $y \in Y$.

Определение 2.8. Функцията $g(y) = x$ задаваща правило, по което на всеки елемент $y \in Y$ съответства $x \in D$, се нарича обратна на $f(x)$.



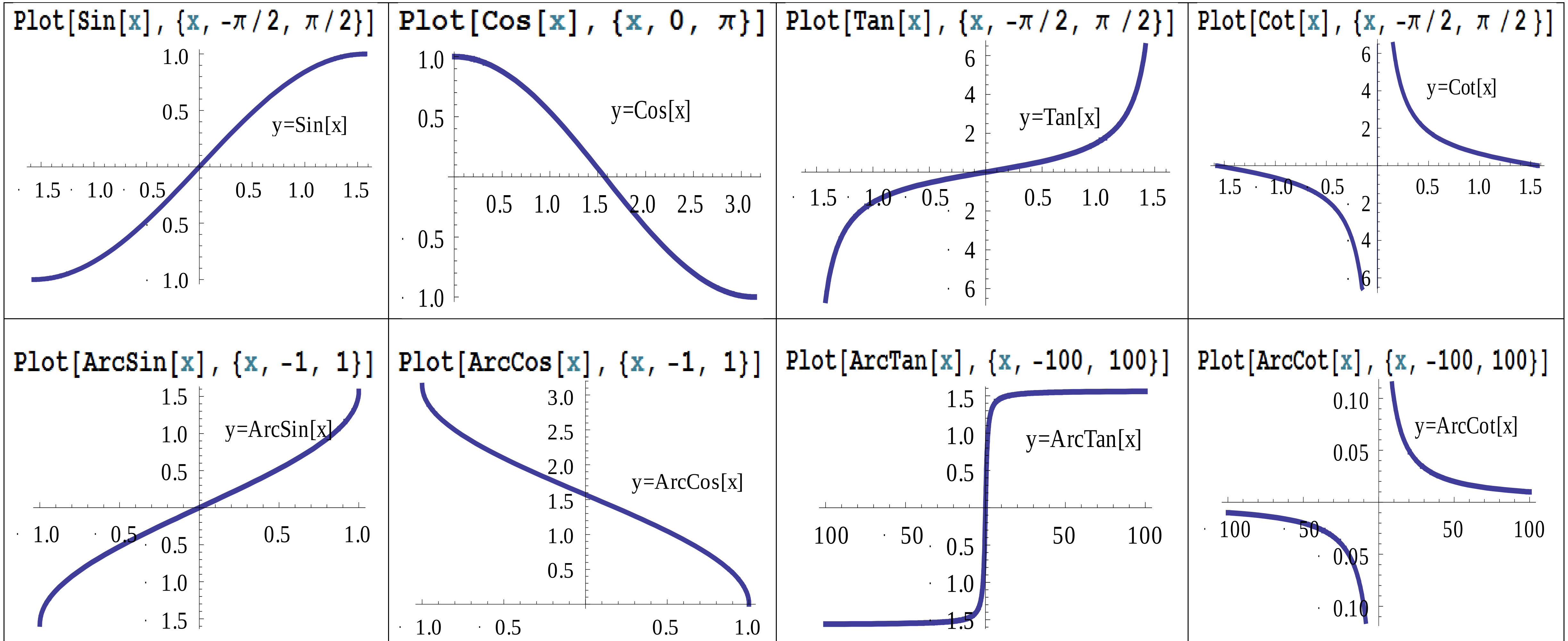
Пример 2.8 Нека дефиниционното множество съдържа само 5 числа: $D=\{-2,-1,0,1,2\}$ и да дефинираме функцията $y=f(x)=x^2$, $x\in D$. За множеството на функционалните стойности $y\in Y$ получаваме: $Y=\{0,1,4\}$. Оттук веднага се вижда, че не може да дефинираме обратна функция, за която дефиниционно множество е Y, а стойностите са $x\in D$. Напр. за точката 4 не съществува функция, която да му съпоставя точно едно $x\in D$, тъй като има 2 такива стойности: -2 и 2 (виж Определение 2.1).

Пример 2.9 Ако сега изберем $D=\{0,1,2\}$, за същата функция $y=f(x)=x^2$, $x\in D$ ще имаме $Y=\{0,1,4\}$, и можем да дефинираме обратна функция $x=g(y)=\sqrt{y}$ за всяко $y\in Y$.

Обратни тригонометрични функции (аркус функции)

Функция	Аркус функция	Зависимости
$y=\sin x,\ x\in[-\frac{\pi}{2},\frac{\pi}{2}]$ $y\in[-1,1]$	$y=\arcsin x,\ x\in[-1,1]$ $y\in[-\pi/2,\pi/2]$	$\sin(\arcsin x)=x$, $\arcsin(\sin x)=x$
$y=\cos x,\ x\in[0,\pi]$, $y\in[-1,1]$	$y=\arccos x,\ x\in[-1,1]$, $y\in[0,\pi]$	$\cos(\arccos x)=x$, $\arccos(\cos x)=x$
$y=\operatorname{tg} x,\ x\in(-\frac{\pi}{2},\frac{\pi}{2})$, $y\in(-\infty,\infty)$	$y=\operatorname{arc\,tg} x,\ x\in(-\infty,\infty)$, $y\in(-\frac{\pi}{2},\frac{\pi}{2})$	$\operatorname{tg}(\operatorname{arc\,tg} x)=x$, $\operatorname{arc\,tg} x(\operatorname{tg} x)=x$
$y=\operatorname{cotg} x,\ x\in(-\frac{\pi}{2},\frac{\pi}{2})$ $x\neq 0,\quad y\in(-\infty,\infty)$	$y=\operatorname{arc\,cotg} x,\ x\in(-\infty,\infty)$, $x\neq 0,\quad y\in(-\frac{\pi}{2},\frac{\pi}{2}),\ y\neq 0$	$\operatorname{cotg}(\operatorname{arc\,cotg} x)=x$, $\operatorname{arc\,cotg} x(\operatorname{cotg} x)=x$

Графики на основните тригонометрични функции и техните аркус функции, получени с Mathematica, са показани на следните фигури:



АРКУС СИНУС	АРКУС КОСИНУС	АРКУС ТАНГЕНС	АРКУС КОТАНГЕНС

Забележка. За случая на ArcCot виж дискусията: http://www.intmath.com/blog/which-is-the-correct-graph-of-arccot-x/6009

3. Производна на функция. Пресмятане на производна. Производни от по-висок ред. Понятие за частна производна на функция на много променливи. Диференциал на функция. Приложение на производните

3.1. Производна на функция

Определение 3.1. Дадена е функцията $y = f(x), \quad x \in D$.
Производна на функцията $f(x)$ в точката $x_0 \in D$ се нарича границата

$$y'(x_0) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x},$$ където $\Delta x = x - x_0,\ \Delta y = y(x_0 + \Delta x) - y(x_0)$,

ако тази граница съществува.

Забележка! Производната е локално свойство на функцията в определена точка. Когато производната съществува във всяка точка от дефиниционната си област или нейно подмножество D_1, казваме, че функцията има производна (или е диференцируема) в D_1.

А) Геометричен смисъл на производната в дадена точка $x_0 \in D$:

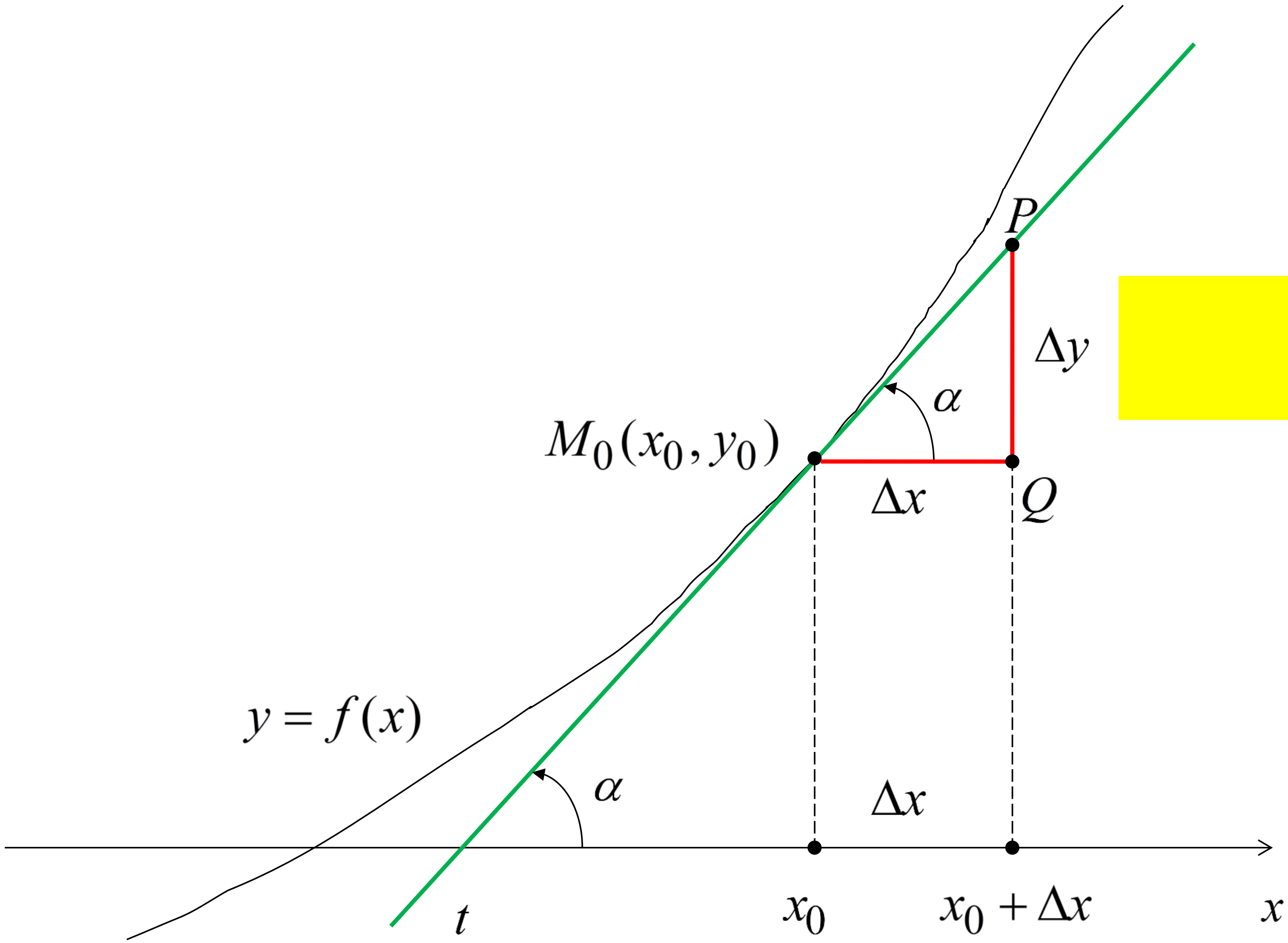
Като се използва определението за тангенс в правоъгълния триъгълник M_0PQ, веднага се вижда, че

$$y'(x_0) \approx tg\,\alpha = \frac{\Delta y}{\Delta x},$$

където α е ъгълът, който сключва допирателната t към графиката на $f(x)$ в допирната точка M_0 с положителната посока на абсцисната ос Ox – виж Фиг. 3.1.

След граничния преход ще имаме равенството:

$$y'(x_0) = tg\,\alpha$$



Фиг.3.1. Геометричен смисъл на производната.

В случая на Фиг. 3.1 стойностите на функцията се увеличават (нарастват) при увеличаване (нарастване) на аргумента x, при което ъгъл α е в интервала $(0, 90^0)$ и тангенсът е положителен. Тогава, съгласно дефиницията имаме:

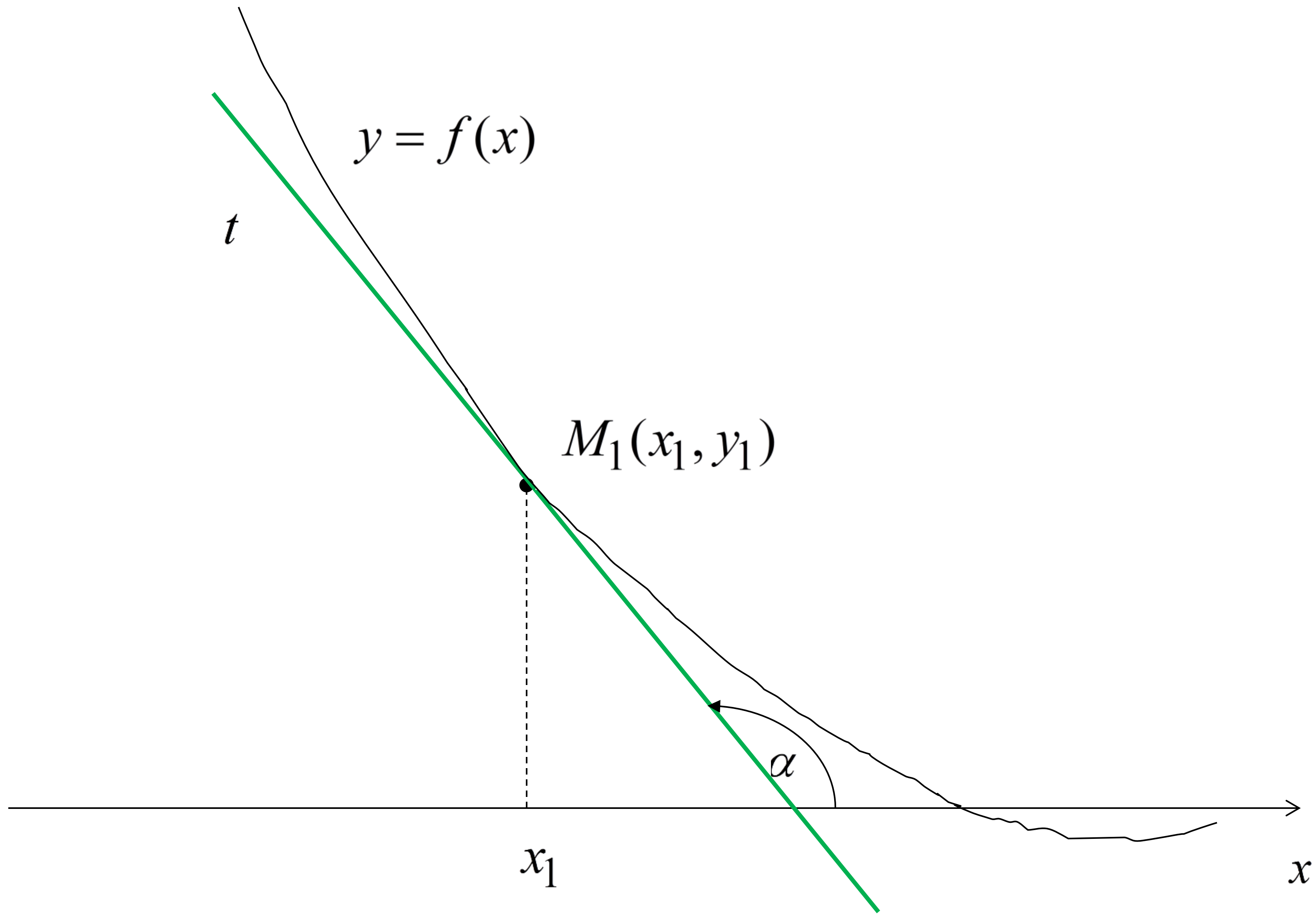
$y'(x) \approx tg\alpha > 0$ - функцията е растяща в околност на точката x_0.

Аналогично при обратен наклон на допирателната – виж Фиг. 3.2 – тангенсът е отрицателен и функцията намалява, т.е.:

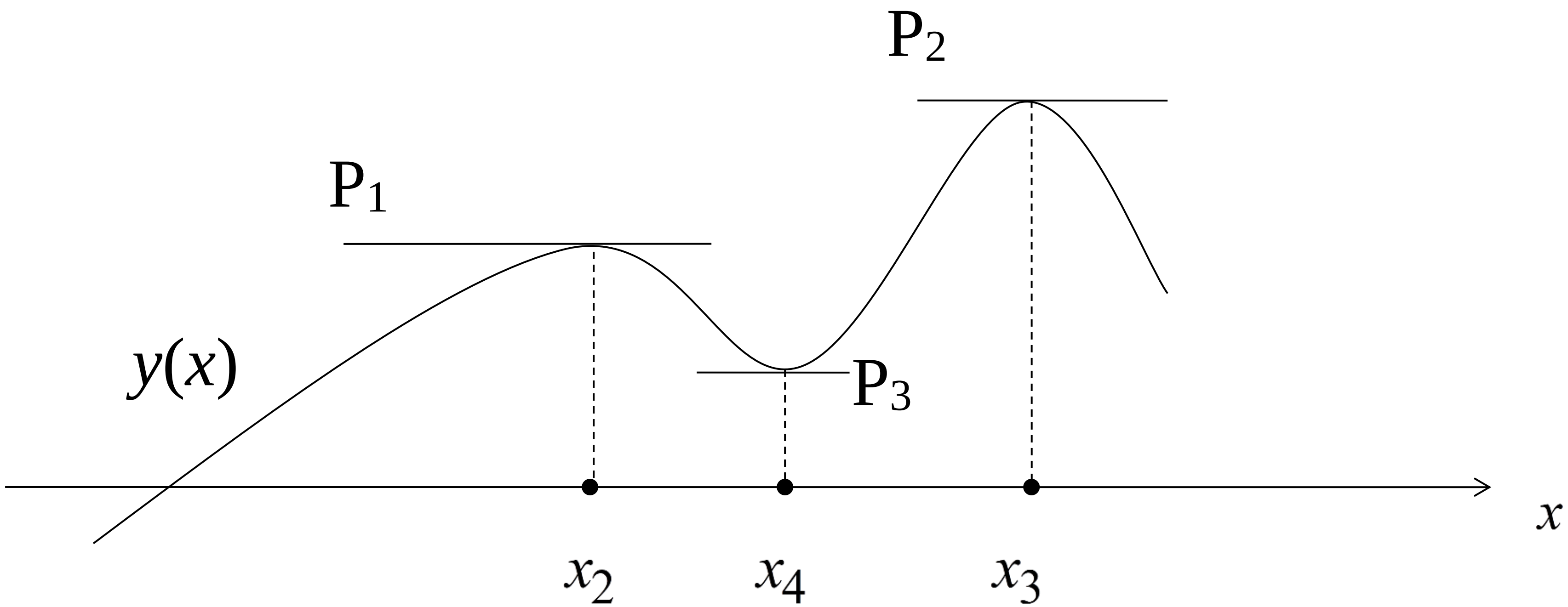
$y'(x) \approx tg\alpha < 0$ - функцията е намаляваща в околност на точката x_1.

Ако в една точка производната съществува и $\alpha = 0$, то има условия функцията да не намалява и да не расте и да има евентуален екстремум.

$y'(x_2) = tg\alpha = 0$- възможно е функцията да има локален максимум или минимум в околност на точката (виж Фиг. 3.3).

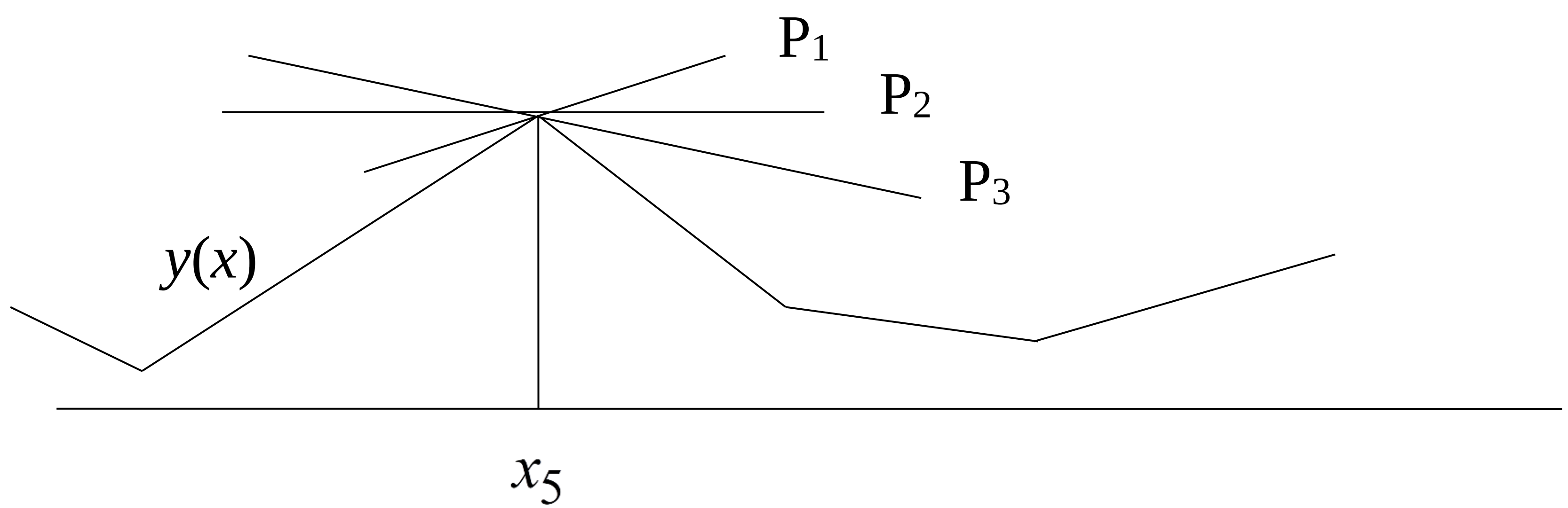


Фиг.3.2. Намаляваща функция в околност на т. x_1 (случай $y'(x) < 0$).



Фиг. 3.3. Локален максимум и минимум на функция в околност на вътрешна точка в *D*. Показаната функция има локален максимум в т. x_2 и x_3 и локален минимум в т. x_4 (необходимо условие за това е ъглите на допирателните с *Ox* да са нули, т.е. допирателните да са успоредни на *Ox*).

В случая на ръб в графиката на функцията $y(x)$ за някоя точка $x-x^*$, в тази точка ще има много допирателни P1, P2, … Може строго да се покаже, че в такъв случай, в точката x^* не съществува границата $y'(x^*) = \lim\limits_{\Delta x \to 0} \dfrac{\Delta y}{\Delta x}$, $\Delta x = x - x^*$ и следователно НЯМА производна в такава точка. (Виж Фиг. 3.4. Има ли други такива точки на фигурата?).



Фиг. 3.4. В т. x_5 не съществува производна $y'(x_5)$.

Б) Физичен смисъл на производната

y' - е скоростта на изменение на функцията $y = f(x)$ спрямо аргумента x. Положителната скорост означава нарастване, отрицателната – намаляване. y' показва тенденцията на изменение на функцията и какво може да се очаква в близкото бъдеще (или минало).

Еквивалентни означения на производната

1. $y'(x)$ - четем: “игрек прим от хикс”
2. y'_x - четем: “игрек прим спрямо хикс” или “производна на игрек спрямо хикс”
3. $\dfrac{dy}{dx}$ - четем: ”де игрек делено на де хикс”. Означава също диференциал на y относно променливата x.

Основни правила за диференциране на функции

(виж също Таблица на производните във файл Formuli 1)

1) $(c)' = 0$ - производна на константа

2) $(cu)' = cu'$ - производна на произведение на константа с функция

3) $(u \pm v)' = u' \pm v'$ - производна на сума (разлика) на две функции u и v

4) $(u.v)' = u'.v + u.v'$ - производна на произведение на две функции

5) $\left(\dfrac{u}{v}\right)' = \dfrac{u'.v - u.v'}{v^2}$ - производна на частно на две функции

6) $\left(u(v(x))\right)' = u'_v.v'_x$ - производна на функция от функция (сложна функция)

3.2. Пресмятане на производна на функция

Пример 3.1 Извод на правило 1). От определение 3.1 за $y(x)=c$

имаме: $$y'(x)=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\lim_{\Delta x\to 0}\frac{c-c}{\Delta x}=0.$$

Пример 3.2 Извод на правило 2) От определение 3.1 за $y(x)=cu(x)$:

$$\left(c.u(x)\right)'=y'(x)=\lim_{\Delta x\to 0}\frac{y(x_0+\Delta x)-y(x_0)}{\Delta x}=\lim_{\Delta x\to 0}\frac{cu(x_0+\Delta x)-cu(x_0)}{\Delta x}=$$

$$c.\lim_{\Delta x\to 0}\frac{u(x_0+\Delta x)-u(x_0)}{\Delta x}=c.\lim_{\Delta x\to 0}\frac{\Delta u}{\Delta x}=c.u'(x)$$

Пример 3.3 Да се намери производната на функцията $y(x)=x$

$$(x)'=\lim_{\Delta x\to 0}\frac{y(x)-y(x_0)}{\Delta x}=\lim_{\Delta x\to 0}\frac{x-x_0}{\Delta x}=\lim_{\Delta x\to 0}\frac{x_0+\Delta x-x_0}{\Delta x}=\lim_{\Delta x\to 0}\frac{\Delta x}{\Delta x}=1.$$

Пример 3.4 Да се намери производната на функцията $y(x) = x^2$

$$\left(x^2\right)' = \lim_{\Delta x\to 0}\frac{y(x)-y(x_0)}{\Delta x} = \lim_{\Delta x\to 0}\frac{x^2-x_0^2}{\Delta x} = \lim_{\Delta x\to 0}\frac{\left(x-x_0\right)\left(x+x_0\right)}{\Delta x}.$$

$$\lim_{\Delta x\to 0}\frac{\Delta x(x+x-\Delta x)}{\Delta x} = \lim_{\Delta x\to 0}(2x-\Delta x) = 2x - \lim_{\Delta x\to 0}\Delta x = 2x-0 = 2x$$

където е заместено $x_0 = x - \Delta x$, съгласно Определение 3.1.

Пример 3.5 С помощта на таблицата на производните да се пресметнат първите производни на следните функции:

а) $y = 5x^2 + 7x - 4,25$

б) $y = \sqrt{3x}$

в) $y = x\sin x$

г) $y = \dfrac{\sin x}{x^2}$

д) $y = x^2 e^x \cos x$

е) $y = 4e^{-3x}$

ж) $y = \sin^4\left(\dfrac{2x+3}{1-x}\right)$

з) $y = \ln(x^3+1)$

Решение:

а) Сума на функции:

$$y' = \left(5x^2\right)' + (7x)' - (4,25)' = 5\left(x^2\right)' + 7(x)' - 0 = 5.2x + 7.1 - 0 = 10x + 7$$

Б) Умножение на функция с константа:

$$y' = \left(\sqrt{3x}\right)' = \left(\sqrt{3}.\sqrt{x}\right)' = \sqrt{3}\left(x^{\frac{1}{2}}\right)' = \sqrt{3}.\frac{1}{2}.x^{\left(\frac{1}{2}-1\right)} = \frac{\sqrt{3}}{2}x^{-\frac{1}{2}} = \frac{\sqrt{3}}{2\sqrt{x}}$$

б) Произведение на функции:

$$y' = (x\sin x)' = (x)'.\sin x + x.(\sin x)' = 1.\sin x + x.\cos x = \sin x + x\cos x$$

г) Частно на функции:

$$y' = \left(\frac{\sin x}{x^2}\right)' = \frac{x^2(\sin x)' - \left(x^2\right)'.\sin x}{x^4} = \frac{x^2\cos x - 2x\sin x}{x^4}$$

д) Произведение от три функции:

$$y' = \left(x^2 e^x \cos x\right)' = \left(x^2 e^x\right)' \cos x + x^2 e^x (\cos x)' = \left(\left(x^2\right)' e^x + x^2\left(e^x\right)'\right)\cos x +$$

$$x^2 e^x(-\sin x) = \left(2xe^x + x^2 e^x\right)\cos x - x^2 e^x \sin x = e^x\left(x^2\cos x + 2x\cos x - x^2\sin x\right)$$

е) Сложна функция (функция от функция):

$$y' = \left(4e^{-3x}\right)' = 4\left(e^u\right)'_u.(-3x)'_x = -4e^{-3x}.(-3).1 = -12e^{-3x}$$

ж) Функция от функция от функция: $y = u^4,\ u = \sin(v),\ v = \dfrac{2x+3}{1-x}$

$$y' = \left[\sin^4\left(\frac{2x+3}{1-x}\right)\right]' = \left[(u)^4\right]'_u .(\sin v)'_v .\left(\frac{2x+3}{1-x}\right)'_x =$$

$$4u^3.\cos(v).\frac{(2x+3)'.(1-x)-(2x+3).(1-x)'}{(1-x)^2} =$$

$$4\sin^3\left(\frac{2x+3}{1-x}\right).\cos\left(\frac{2x+3}{1-x}\right).\frac{5}{(1-x)^2}.$$

з) Сложна функция: $y = \ln(u),\ u = x^3 + 1.$

$$y' = \left(\ln(x^3+1)\right)' = \left(\ln(u)\right)'_u .(u)'_x = \frac{1}{u}.(x^3+1)' = \frac{1}{(x^3+1)}.3x^2 = \frac{3x^2}{(x^3+1)}$$

Пример 3.6 Да се решат задачите от Пример 3.5 с помощта на ***Mathematica***.

Решение:

	Вход	Изход
а)	`y1 = 5 x^2 + 7 x - 4.25;` `∂_x y1`	$7 + 10\,x$
б)	`y2 = Sqrt[3 x];` `∂_x y2`	$\frac{\sqrt{3}}{2\sqrt{x}}$
в)	`y3 = x*Sin[x];` `∂_x y3`	x Cos[x] + Sin[x]
г)	`y4 = Sin[x]/x^2;` `∂_x y4`	$\frac{\mathrm{Cos}[x]}{x^2} - \frac{2\,\mathrm{Sin}[x]}{x^3}$

д)	y5 = $x^2 * e^x *$ Cos[x]; ∂_x y5	$2\, e^x\, x\, \text{Cos}[x] + e^x\, x^2\, \text{Cos}[x] - e^x\, x^2\, \text{Sin}[x]$
е)	y6 = $4\, e^{-3\, x}$; ∂_x y6	$-12\, e^{-3\, x}$
ж)	y8 = $\left(\text{Sin}\left[\frac{2\, x + 3}{1 - x}\right]\right)^4$; ∂_x y8	$4 \left(\frac{2}{1 - x} + \frac{3 + 2\, x}{(1 - x)^2}\right) \text{Cos}\left[\frac{3 + 2\, x}{1 - x}\right] \text{Sin}\left[\frac{3 + 2\, x}{1 - x}\right]^3$
з)	y7 = Log$[x^3 + 1]$; ∂_x y7	$\frac{3\, x^2}{1 + x^3}$

<u>Забележка.</u> За възможностите на *Mathematica* за решаване на задачи от диференциалното и интегрално смятане, препоръчваме да се разгледа помощната част на продукта: Help, *tutorial* / *CalculusOverview*

3.3. Производни от по-висок ред

Втора производна на функция y'' - намира се като се диференцира първата производна y'.

Трета производна y''' - диференцираме втората производна y'' и т.н.

Пример 3.7 Да се пресметнат вторите производни на следните функции:

а) $y = 2x - \cos(3x)$, б) $y = \dfrac{4-x}{3x^2+2}$.

Решение:

а) Намираме : $y'(x) = 2 + 3\sin(3x)$. Диференцираме полученото:

$$y'' = \left(y'(x)\right)' = 9\cos(3x).$$

С помощта на система Mathematica смятаме така:

За y' набираме `∂x (2 x - Cos[3 x])`;

получаваме: `2 + 3 Sin[3 x]`.

За y'' набираме `∂x,x (2 x - Cos[3 x])` и получаваме: `9 Cos[3 x]`.

б) Първата производна е:

$$y' = \left(\frac{4-x}{3x^2+2}\right)' = \frac{(4-x)'\left(3x^2+2\right)-\left(3x^2+2\right)'(4-x)}{\left(3x^2+2\right)^2} =$$

$$\frac{-\left(3x^2+2\right)-6x(4-x)}{\left(3x^2+2\right)^2} = \frac{-3x^2-2-24x+6x^2}{\left(3x^2+2\right)^2} = \frac{3x^2-24x-2}{\left(3x^2+2\right)^2}$$

Втората производна е:

$$y'' = \frac{3x^2 - 24x - 2}{\left(3x^2+2\right)^2} = \frac{(6x-24)\left(3x^2+2\right)^2 - 2\left(3x^2+2\right).6x.\left(3x^2-24x-2\right)}{\left(3x^2+2\right)^4} =$$

$$= \frac{(6x-24)\left(3x^2+2\right) - 12x\left(3x^2-24x-2\right)}{\left(3x^2+2\right)^3} = \frac{-18x^3+216x^2+36x-48}{\left(3x^2+2\right)^3}.$$

С помощта на система *Mathematica* смятаме така:

За y' набираме:

```
∂x (4 - x)/(3 x^2 + 2)
Together[%]
```

получаваме: $\dfrac{-2-24\,x+3\,x^2}{\left(2+3\,x^2\right)^2}$.

За y'' набираме:

$$\partial_{x,x}\frac{4-x}{3x^2+2};$$

```
Together[%]
```

и получаваме: $$-\frac{6\left(8-6x-36x^2+3x^3\right)}{\left(2+3x^2\right)^3}.$$

Задача. Намерете на ръка третата производна на функцията от Пример 3.7-а).

Задача. Намерете с компютър втора и трета производна на всички функции от Пример 3.5.

3.4. Понятие за частна производна на функция на много променливи

Нека е дадена функция на няколко променливи:

Напр. функция на две независими променливи: $y = f(x, y)$

Функция на три независими променливи: $y = f(x, y, z)$ и т.н.

Определение 3.2. Ако е дадена функцията $f(x, y)$, $(x, y) \in D \subset R^2$, то частна производна относно x на функцията $f(x, y)$ в точката $(x_0, y_0) \in D$ се нарича границата

$$\frac{\partial f}{\partial x}(x_0, y_0) = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x},$$ където $\Delta x = x - x_0$,

ако тази граница съществува. Т.е. частната производна е обикновена производна при фиксирани стойности на другите независими променливи, в случая при фиксирано $y = y_0 \in D$.

Аналогично се дефинира частна производна относно y:

$$\frac{\partial f}{\partial y}(x_0, y_0) = \lim_{\Delta y \to 0} \frac{f(x_0, y+\Delta y) - f(x_0, y_0)}{\Delta y}, \text{ при фиксирано } x = x_0 \in D.$$

Втори частни производни са производни от първите относно една независима променлива и съответно фиксирани останалите променливи (ако съответните граници съществуват).

Например, втора частна производна само по x е:

$$\frac{\partial^2 f}{\partial x^2}(x_0, y_0) = \frac{\partial f}{\partial x}\left(\frac{\partial f}{\partial x}\right)(x_0, y_0) = \lim_{\Delta x \to 0} \frac{\frac{\partial f}{\partial x}(x_0 + \Delta x, y_0) - \frac{\partial f}{\partial x}(x_0, y_0)}{\Delta x}$$

Например, втора частна производна (смесена производна) по y спрямо първа производна по x е:

$$\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) = \frac{\partial f}{\partial y}\left(\frac{\partial f}{\partial x}\right)(x_0, y_0) = \lim_{\Delta y \to 0} \frac{\frac{\partial f}{\partial x}(x_0, y_0 + \Delta y) - \frac{\partial f}{\partial x}(x_0, y_0)}{\Delta y}$$

Пример 3.8 Да се пресметнат всички първи частни производни на следните функции:

а) $f = x^5 - 3xy + 8\sqrt{y}$ б) $f = \dfrac{3 + x}{4y^2 + 1}$

<u>Решение:</u>

а) $\dfrac{\partial f}{\partial x} = \left(x^5 - 3xy + 8\sqrt{y}\right)_x' = 5x^4 - 3y,$ (y се счита за константа).

$\dfrac{\partial f}{\partial y} = \left(x^5 - 3xy + 8\sqrt{y}\right)_y' = -3x + \dfrac{4}{\sqrt{y}},$ (x е константа).

С помощта на система Mathematica смятаме така:

За $\dfrac{\partial f}{\partial x}$ набираме $\partial_x \left(x^5 - 3\,x * y + 8\sqrt{y}\right)$

получаваме: $5\,x^4 - 3\,y$

За $\dfrac{\partial f}{\partial y}$ набираме $\partial_y \left(x^5 - 3\,x * y + 8\sqrt{y}\right)$ и получаваме: $-3\,x + \dfrac{4}{\sqrt{y}}$

б) $\dfrac{\partial f}{\partial x} = \left(\dfrac{3+x}{4y^2+1}\right)'_x = 1$, ($y$ се приема за константа).

$$\frac{\partial f}{\partial y} = \left(\frac{3+x}{4y^2+1}\right)'_y = -\frac{8y(3+x)}{\left(4y^2+1\right)^2},$$ (x се приема за константа).

Задача. Намерете на ръка всички втори частни производни на функциите от Пример 3.8.

Задача. Намерете с компютър всички частни производни от първи и втори ред на функциите:

а) $f(x,y)=\cos(x+3y)$,

б) $f(x,y)=\dfrac{1}{1+\arctan(xy-2)}$

в) $f(x,y)=\dfrac{x+4y}{x^2+y^2}$,

г) $f(x,y,z)=x^3+3x^2y+yz^3$

3.5. Диференциал на функция

Понятието „диференциал” е въведено от Лайбниц и Бернули за изразяване на безкрайно малки величини.

Определение 3.3 Нека $y(x)$ е диференцируема функция в някакъв интервал. Формално означаваме

$y'(x) = \frac{dy}{dx}$. (четем „игрек прим равно на де игрек делено на де хикс”)

Оттук получаваме диференциала на y:

$$dy = y'(x)dx$$

(четем „де игрек е равно на игрек прим по де хикс”). Диференциалът е функция както на x, така и на dx.

Готфрид Вилхелм Лайбниц
Gottfried Wilhelm Leibniz (1646-1716)

Лайбниц е един от най-големите универсални учени в света. Всъщност, списъкът на неговите значителни научни приноси е толкова голям, колкото и неговите ранообразни дейности. Като инженер той е изчислявал елементи на машини, часовници и даже машини за минното дело. Като библиотекар е създал съвременния каталог. Като математик е създал основите на топологията, но е един и от основателите на диференциалното и интегралното смятане, заедно с Нютон. Въведените от него означения са станали стандарт. В логиката е работил върху двоичните системи, и много други. Като физик е постигнал значителни успехи в механиката и по-специално в теорията на моментите. Има приноси и в лингвистиката, историята, естетиката и политическата теория.

(Internet Encyclopedia of Philosophy)

Правила за пресмятане на диференциали

Те са аналогични на тези за намиране на производни, съгласно Определение 3.3.

1) $d(c) = 0$ - диференциал на константа
2) $d(cu) = c.du$ - диференциал на произведение на константа с функция
3) $d(u \pm v) = du \pm dv$- диференциал на сума(разлика) на две функции
4) $d(u.v) = du.v + u.dv$ - диференциал на произведение на две функции
5) $d\left(\dfrac{u}{v}\right) = \dfrac{v.du - u.dv}{v^2}$ - диференциал на частно на две функции
6) $d\big(u(v(x))\big) = du.dv$ - диференциал на функция от функция (сложна функция)

Пример 3.9 Да се пресметнат диференциалите на следните функции:

а) $y = 5x^3 - 2x$ б) $y = tg\left(x^2 - 7x + 2\right)$

Решение:

а) $dy = y'dx = \left(5x^3 - 2x\right)' dx = \left(15x^2 - 2\right)dx$

б) $dy = y'dx = \left(tg\left(x^2 - 7x + 2\right)\right)' dx = \left(\dfrac{(2x-7)}{\cos^2(x^2 - 7x + 2)}\right)dx$

3.7. Приложение на производните за намиране на граници - правило на Бернули-Лопитал

Намирането на граници на функции при наличие на особености от вида $\left[\frac{0}{0}\right]$ и $\left[\frac{\infty}{\infty}\right]$ при изпълнение на някои условия, може да се осъществи с известното правило на Бернули-Лопитал. Правилото е предложено от Бернули и публикувано за първи път от Лопитал през 1696 година.

Условията за прилагане са:
а) Функциите $f(x), g(x)$ са определени и диференцируеми в някаква околност на т. a;
б) За границите знаем, че $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = 0$ или $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = \infty$.

б) съществува границата $\lim\limits_{x\to a}\dfrac{f'(x)}{g'(x)}$ за $g'(x)\neq 0$ в същата околност;

Тогава съществува границата:

$$\lim_{x\to a}\frac{f(x)}{g(x)}=\lim_{x\to a}\frac{f'(x)}{g'(x)}\qquad \text{(Правило на Бернули-Лопитал)}$$

Забележка. Възможно е многократно прилагане на правилото.

Пример 3.10 Намерете границите на функциите:

А) $y=\dfrac{x^2-2}{3x^2+4}$ при $x\to\infty$; Б) $y=\dfrac{x^2-4}{x^2+x-2}$ при $x\to -2$;

Решение: А) След директно заместване получаваме особеност $\left[\dfrac{\infty}{\infty}\right]$.

Задачата може да се реши с изнасяне на най-големия множител x^2 и съкращаване, както научихме в Лекция 2.

С правилото на Бернули-Лопитал пресмятаме производните в числителя и знаменателя и получаваме:

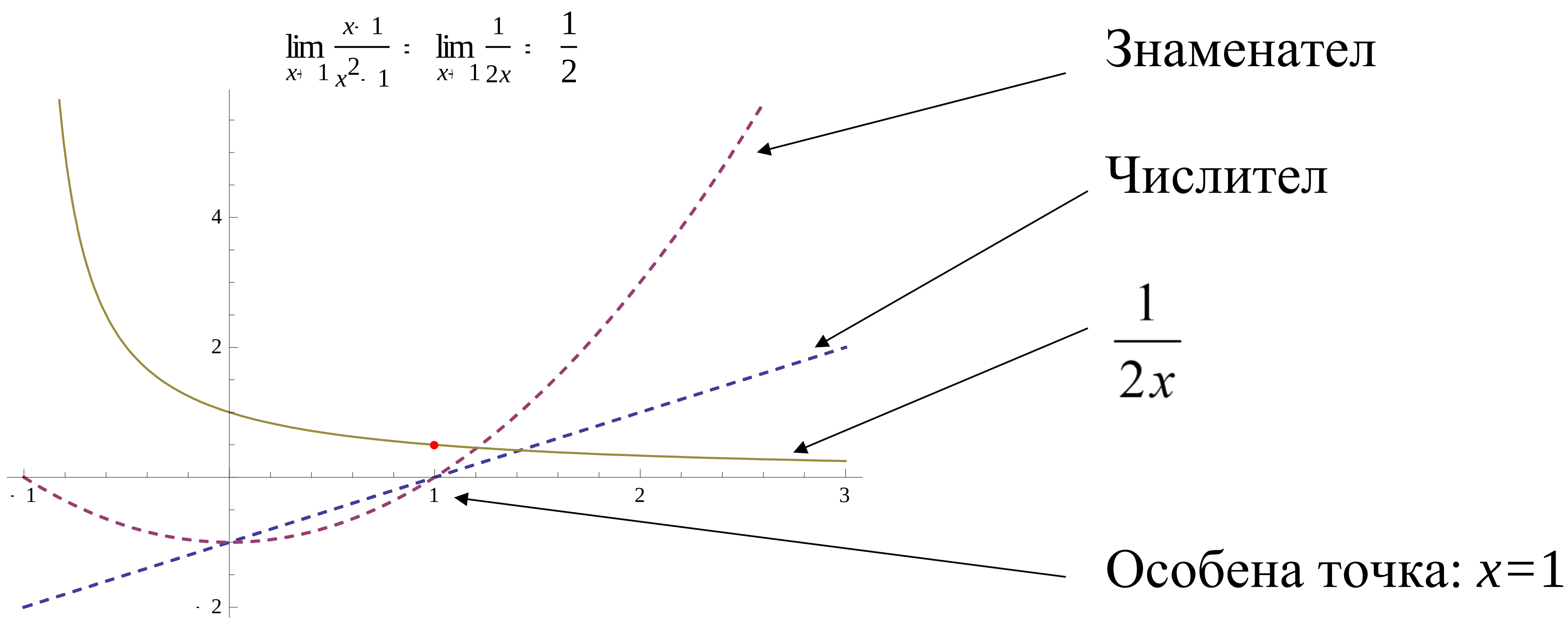
$$\lim_{x\to\infty}\frac{x^2-2}{3x^2+4}=\left[\frac{\infty}{\infty}\right]=\lim_{x\to\infty}\frac{\left(x^2-2\right)'}{\left(3x^2+4\right)'}=\lim_{x\to\infty}\frac{2x}{6x}=\frac{1}{3}\ .$$

Б) Аналогично: $\lim\limits_{x\to-2}\dfrac{x^2-4}{x^2+x-2}=\left[\dfrac{0}{0}\right]=\lim\limits_{x\to-2}\dfrac{\left(x^2-4\right)'}{\left(x^2+x-2\right)'}=$

$$\lim_{x\to-2}\frac{2x}{2x+1}=\frac{2(-2)}{2(-2)+1}=\frac{-4}{-3}=\frac{4}{3}.$$

Интересни примери с Mathematica може да намерите на:

http://demonstrations.wolfram.com/LHospitalsRuleFor00Forms/



3.8. Приложение на производните за намиране на екстремуми на функции на една променлива

В 3.1 бяха въведени интуитивно понятията за нарастване, намаляване и локален екстремум (максимум и минимум) на функция. Поставя се задачата как да намираме точките от дефиниционната област D на една функция $f(x)$, в която тя достига локален екстремум.

Нека $f(x)$ е определена в интервала $(a; b)$ и $x_0 \in (a; b)$.

Определение 3.4 Казваме,че функцията $f(x)$ има <u>локален максимум</u> $f(x_0)$ в точката x_0, ако съществува околност I на точката x_0, в която

$f(x) \le f(x_0)$ за $\forall x \in I$ (четем „за всяко x от I ”).

Функцията има строг максимум, ако $f(x) < f(x_0)$, за $x \neq x_0$.

Аналогично $f(x)$ има <u>локален минимум</u> $f(x_0)$ в точката x_0, ако съществува околност I на точката x_0, в която $f(x) \geq f(x_0)$ за $\forall x \in I$

$f(x) \le f(x_0)$ за $\forall x \in I$. Минимумът е строг, ако $f(x) > f(x_0)$, за $x \neq x_0$.

Когато функцията изпълнява някое от условията в Определение 3.4, ще казваме, че тя има локален екстремум в т. x_0.

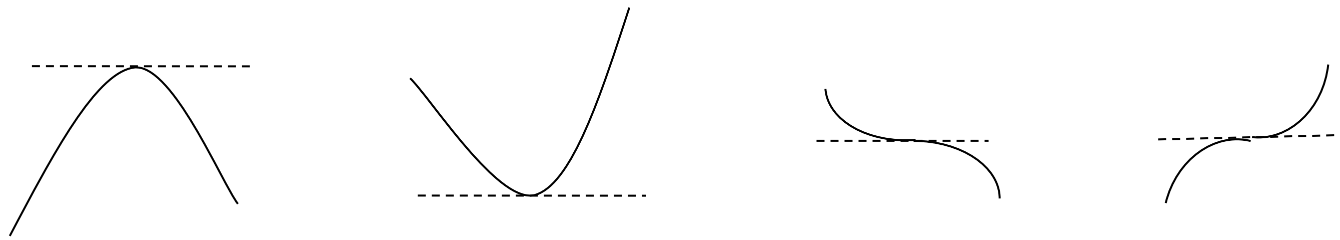
Сега ще се спрем на случая, когато функцията е диференцируема и ще покажем как производните се използват за намиране на екстремуми.

Малка теорема на Ферма (необходимо условие за екстремум)

Ако функцията $f(x)$ има екстремум в точка $x_0 \in (a; b)$ и е диференцируема в тази точка, то $f'(x_0) = 0$.

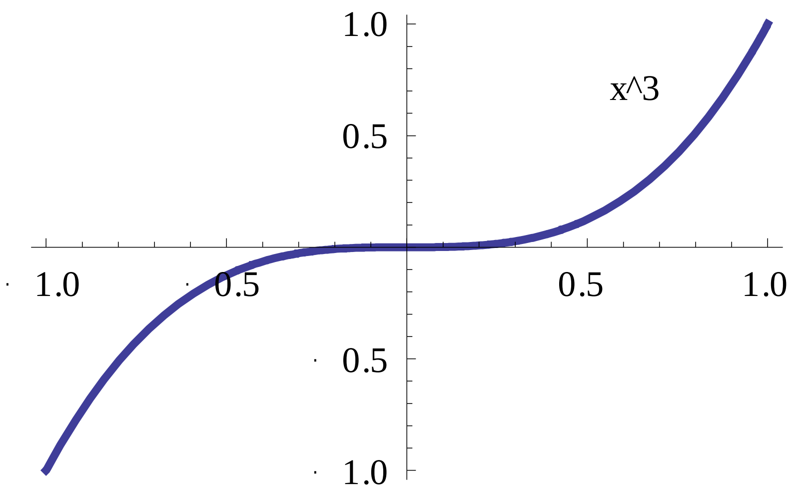
Такава точка $x_0 \in (a; b)$ се нарича стационарна точка.

При геометричното тълкуване на производната бяха дадени примери, в които се достига минимум и максимум на функция (Фиг. 3.3). Случаите, в които се изпълнява необходимото условие на малката теорема на Ферма за анулиране на f', геометрично са следните:



Фиг. 3.5 Случаи, в които $f'(x_0) = 0$: в първите две точки има съответно максимум и минимум, а в останалите две точки – няма.

Пример 3.11 Функцията $f(x) = x^3$ има първа производна $f' = 3x^2$, която се анулира в т. x_0=0. Но f няма екстремум, както се вижда от графиката й на Фиг. 3.6, получена с Mathematica.



Локален екстремум обаче може да се достига и в точки, в които функцията няма производни (виж Фиг. 3.4)!

Определение 3.5 Точките от $(a; b)$, в които $f(x)$ не е диференцируема или е диференцуема с изпълнено необходимо условие $f'(x_0) = 0$ се наричат критични точки на функцията $f(x)$.

От тези разсъждения следва, че условието $f'(x_0) = 0$ не гарантира съществуване на екстремум в т. $x = x_0$, т.е. това условие не е достатъчно. За намиране на екстремум се изследват знаците на първата производна в отворен интервал, съдържащ дадена критична точка и поведението на функцията (растяща, намаляваща) – виж подтемата „Геометричен смисъл на производната” и Фиг. 3.1.

Правило с използване на производна:

1) Ако в стационарна точка x_k производната $f'(x_k)$ си сменя знака, то:

- $f(x)$ има максимум, ако смяната е от (+) към (−)
- $f(x)$ има минимум, ако смяната е от (−) към (+)

2) Ако в стационарна точка x_k производната $f'(x_k)$ не си сменя знака, то функцията $f(x)$ няма екстремум в точката x_k.

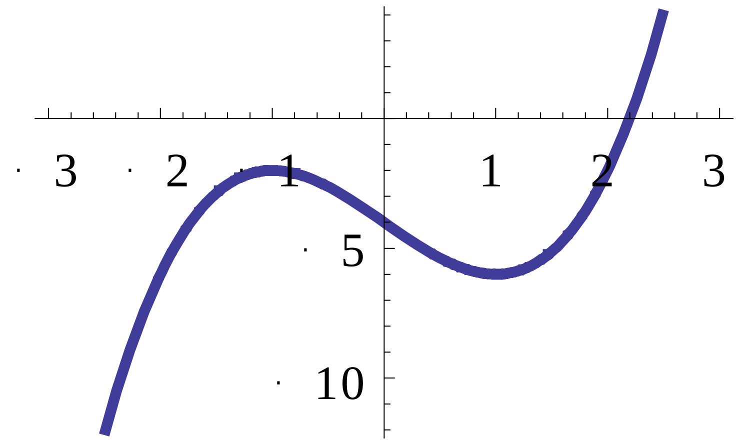
Пример 3.12 Да се намерят екстремумите на функцията. $f(x) = x^3 - 3x - 4$.

Решение: Функцията е дефинирана за всяко $x \in (-\infty, \infty)$. Намираме производната $f'(x) = 3x^2 - 3$. Решаваме уравнението $f'(x) = 0$.

$f'(x) = 3x^2 - 3 = 0$, $x_{1,2} = \pm 1$. Тези две точки са стационарните, в които може да има или да няма екстремум. Записваме таблицата и определяме знаците на първата производна в получените три подинтервала $(-\infty, x_1)$, (x_1, x_2), (x_2, ∞):

x	$-\infty$	$x_1 = -1$		$x_2 = 1$	$+\infty$
$f'(x)$	$+$	0	$-$	0	$+$
$f(x)$	$\nearrow$		$\searrow$		$\nearrow$
		max		min	

Отговор: Локален максимум в x_1, равен на $f(x_1) = f(-1) = -2$ и локален минимум в x_2, равен на $f(x_2) = f(1) = -6$.
Графиката е показана на фигурата.



3.9. Примери за приложения на производните в икономиката

Пример 3.13 Установено е, че фирма „Донев ООД" след t години е натрупала приходи по закона

$$f(t) = 1{,}7t^2 + 2{,}4 \text{ млн. лв}$$

Да се намерят приходите в следващ момент $t + \Delta t$ и да се изчисли $f(t+\Delta t)$ при $t=3$ години и $\Delta t=1/2$ години.

Решение: От дефиницията за производна имаме приближеното равенство

$$f'(t) \approx \frac{f(t+\Delta t) - f(t)}{\Delta t}$$

Тогава $f(t+\Delta t) \approx f(t) + f'(t)\Delta t$ (*)

Изчисляваме: $f'(t) = \left(1{,}7t^2 + 2{,}4\right)' = 3{,}4t$.

Заместваме в (*) с по-горе дадените стойности:

$$f(t+\Delta t) \approx f(3+0,5) \approx f(3)+0,5f'(3)$$

$$=1,7.3^2+2.4+0,5(3.3,4)=17,7+5,1=22,8$$

Отговор: $f(3,5) \approx 22,8$

Коментар: Ако в момента $t=3$ приходите са били $f(3) \approx 17,7$, то след половин година се очаква да бъдат 22,8.

Други приложения в икономиката

Нека q е обемът на произведената продукция, а $c(q)$ е функцията на разходите за производство. Да означим с Δq допълнително количество на произвежданата продукция при текущо ниво q.

Тогава допълнителните разходи са

$$\Delta c = c(q+\Delta q) - c(q)$$

Относителните разходи са:

$$\frac{\Delta c}{\Delta q} = \frac{c(q+\Delta q) - c(q)}{\Delta q}$$

При $\Delta q \to 0$ получаваме моментната скорост на изменение на разходите, наречени <u>маргинални разходи</u>

$$c'(q) = \lim_{\Delta q \to 0} \frac{\Delta c}{\Delta q} = \lim_{\Delta q \to 0} \frac{c(q+\Delta q) - c(q)}{\Delta q} \quad \text{или} \quad c'(q) = \frac{dc}{dq}$$

Прието е маргиналният разход да се бележи с Mc, т.е.:
$Mc = c'(q)$.

Пример 3.14 Известно е, че функцията на разходите има вида:
$c(q) = 100q - \frac{0,2}{3}q^3$. Да се намерят маргиналните разходи при q = 20 единици обем на производство.

Решение: $Mc = c'(q) = 100 - 0,2q^2$. При q = 20,
$Mc = 100 - 0,2.20^2 = 20$.

Това означава, че при обем на производството 20 единици, производствените разходи за следващата 21-ва единица са 20.

4. Неопределен интеграл – свойства, пресмятане

В предната лекция и упражненията разгледахме понятието производна на функция. Нека е дадена функция $g(x)$, дефинирана в някакъв отворен интервал (a, b) и нека за всяка точка x от (a, b) съществува производната $g'(x)$. Казваме още, че $g(x)$ е диференцируема в (a, b).

Задачата при дадена функция $g(x)$ да се намери нейната производна функция $g'(x)$ ще запишем формално така:

$g(x) \xrightarrow[\text{диференциране}]{} g'(x)$, където $'$ е знак за диференциране.

Ако тук използваме означението за диференциал на $g(x)$ (виж по-подробно лекция 3), ще имаме

$$dg(x) = g'(x)dx \tag{1}$$

В много задачи се налага да се реши обратната задача или задачата за антидиференциране, която се формулира така:
Ако е известна някаква функция $f(x)$, дефинирана и непрекъсната в (a, b) да се намери диференцируема в (a, b) функция $F(x)$, чиято производна да е равна на $f(x)$. Математическото действие се нарича интегриране, или формално

$$f(x) \underset{\textit{интегриране}}{\longrightarrow} F(x), \quad \text{така че}$$

$$F'(x) = f(x) \text{ или } dF(x) = f(x)dx. \tag{2}$$

Лесно се съобразява, че ако последните равенства са верни за някаква функция $F(x)$, те са верни и за $F(x) + C$, тъй като производната на константа C е винаги нула. Следователно задачата за интегрирането има безброй решения.

4.1. Определение и свойства на неопределения интеграл

Определение 4.1 Една функция $F(x)$ се нарича примитивна функция или неопределен интеграл на $f(x)$, ако

$$F'(x) = f(x) \tag{2}$$

и с помощта на знаците за диференциал и интеграл $f(x)$ $\int$ се бележи с

$$\int f(x)\,dx = F(x) + C \tag{3}$$

(четем „интеграл от еф де хикс е равно на еф голямо от хикс плюс це”).

Важно е да се подчертае, че на практика неопределеният интеграл е фамилия от функции $F(x) + C$, с отчитането на произволна константа C за всяка примитивна $F(x)$!

Тук дадената функция $f(x)$ между знака за интеграл и диференциал се нарича подинтегрална функция, а независимата променлива x - интеграционна променлива.

Пример 4.1 Да се изчисли интеграл от 1.

<u>Решение:</u> В случая $f(x) = 1$. Записваме задачата като $\int 1dx = ?$ или опростено като $\int dx$. Търсим функция, чиято производна е равна на 1. Тъй като $x' = 1$ веднага намираме, че $F(x) = x + C$. Следователно:

$$\int dx = x + C \tag{4}$$

В частност това означава, че знаците за интегриране и диференциране като противоположни действия едно до друго „се унищожават” и остава само интеграционната променлива x с добавена константа.

Пример 4.2 Да се изчисли интеграл от cos(x).

Решение: От таблицата на производните знаем, че производната на $F(x) = \sin(x)$ е равна на $f(x) = \cos(x)$. Следователно по Определение 4.1:

$$\int \cos(x)dx = \sin(x) + C \tag{5}$$

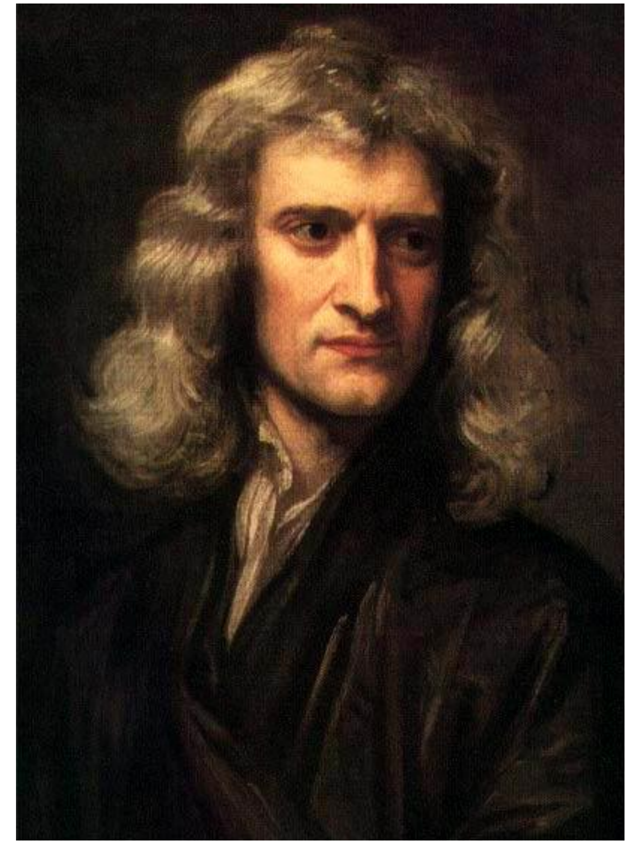
Пример 4.3 Да се изчисли интегралът $\int e^x dx$.

Решение: От таблицата на производните знаем, че производната на e^x е равна на e^x. Следователно по Определение 4.1:

$$\int e^x dx = e^x + C \tag{6}$$

Диференциалното и интегрално смятане имат големи приложения във всички области на науката и практиката. Някои техни елементи са били използвани още от древността от Архимед, но получават съществено развитие в трудовете на Исак Нютон и Готфрид Лайбниц през втората половина на 17 век.

Сър **Исак Нютон** е английски физик, математик, астроном и философ, постигнал изключителни успехи във всички области, в които е работил. Най-известният му труд е „Математически принципи на натурфилософията”, публикуван през 1687, в който е изложил закона за всемирното привличане и три закона на механиката, с които поставя основите на класическата механика. Работейки над проблемите на физиката, Исак Нютон поставя началото на математическия анализ (диференциалното и интегрално смятане), който е в основата на развитието на науката до наши дни. Разработил е също теорията за разлагане на светлината и много други математически и физически теории.



Исак Нютон (1643 – 1727)

Основните математически постижения на разностранния немски учен **Готфрид Лайбниц** са в областта на математическия анализ и на формализирането на математиката. В лайпцигските "Acta eruditorum" публикува разработеното от него диференциално смятане през 1684 г., а интегралното смятане - през 1686г.

Първата работа съдържа знаците за диференциране, правила за диференциране, твърдения за екстремумите и инфлексните точки. Втората работа съдържа знака за интеграл.

Лайбниц разработва анализа независимо от Исак Нютон, но въпреки това между тях се разгаря продължителен и безсмислен спор за приоритет.

Готфрид Лайбниц (1646 – 1716)

СВОЙСТВА НА НЕОПРЕДЕЛЕНИЯ ИНТЕГРАЛ

Свойство 1^0. $\boxed{\left(\int f(x)\,dx\right)' = f(x)}$

Доказателство: Следва веднага от (2) и (3), тъй като

$$\left(\int f(x)\,dx\right)' = \left(F(x)+C\right)' = F'(x)+0 = f(x).$$

Свойство 2^0 . Интеграл от сума е сумата от интегралите:

$$\boxed{\int\left[f(x)+g(x)\right]dx = \int f(x)dx + \int g(x)\,dx}$$

Доказателство: Да означим с $F(x)$ и $G(x)$ примитивните на $f(x)$ и $g(x)$.
От свойствата на производните:

$\left(F(x)+G(x)\right)' = f(x)+g(x)$. Освен това по определението

$$\int\left[f(x)+g(x)\right]dx = F(x)+G(x)+C = \left(F(x)+c_1\right)+\left(G(x)+c_2\right)$$

$$= \int f(x)\,dx + \int g(x)\,dx$$

където $C = c_1 + c_2$.

Свойство 3^0 . Изнасяне на константа: $\boxed{\int \lambda f(x)\,dx = \lambda \int f(x)dx}$

Доказателство: Да означим с $F(x)$ примитивната на $f(x)$. От свойствата на производните:

$\left(\lambda F(x)\right)' = \lambda F(x)$, но $\int \lambda f(x)\,dx = \lambda F(x) + C = \lambda(F(x)+c_\lambda) = \lambda \int f(x)\,dx$

Свойство 4^0 . $\boxed{\int g'(x)dx = g(x)+C}$

Автоматично следва от Определение 4.1.

Свойство (**4^0**) се нарича „внасяне под знака на диференциала" и се използва често за смяна на интеграционната променлива x с нова променлива $u = g(x)$.

В (2) имахме $dF(x) = f(x)dx$. Когато преобразуваме израза $f(x)dx$ във вида $dF(x)$ се казва, че функцията $f(x)$ се ***внася под знака на диференциала*** d и обратно, когато преобразуваме израза $dF(x)$ във вида $f(x)dx$ се казва, че функцията $F(x)$ се ***изнася пред знака на диференциала*** d.

4.2. Пресмятане на неопределен интеграл

По същество намирането на примитивната функция става чрез основната таблица на интегралите, свойствата и преобразования.
Ще разгледаме по-подробно 3 основни техники: непосредствено интегриране, интегриране чрез полагане и интегриране по части.

Техника 1 Непосредствено интегриране (Свеждане до табличен интеграл)

1) $\int x^2 dx = \dfrac{x^3}{3} + C$

2) $\int \dfrac{dt}{\cos^2 t} = tg\, t + C$

3) $\int\left(\frac{5}{x}+\sin x\right)dx=\int\frac{5}{x}dx+\int\sin x dx=5\int\frac{dx}{x}+\int\sin x dx=\ln|x|-\cos x+C$

4) $\int\frac{(x+1)^3}{x}dx=\int\frac{x^3+3x^2+3x+1}{x}dx=\int x^2dx+3\int xdx+3\int dx+\int\frac{1}{x}dx$

$$=\frac{x^3}{3}+\frac{3x^2}{2}+3x+\ln|x|+C$$

5) $\int\frac{1}{\sqrt{x}}dx=\int x^{-\frac{1}{2}}dx=\frac{x^{-\frac{1}{2}+1}}{-\frac{1}{2}+1}+C=\frac{\sqrt{x}}{\frac{1}{2}}+C=2\sqrt{x}+C$

Пресмятане на интегралите 1)-5) с *Mathematica*:

	Задаваме (Input)	Out
1)	$\int x^2 \, dx$	$\frac{x^3}{3}$
2)	$\int \frac{1}{\mathrm{Cos}[t]^2} \, dt$	Tan[t]
3)	$\int (5\,x + \mathrm{Sin}[x]) \, dx$	$\frac{5\,x^2}{2} - \mathrm{Cos}[x]$
4)	$\int \frac{(x+1)^3}{x} \, dx$	$3\,x + \frac{3\,x^2}{2} + \frac{x^3}{3} + \mathrm{Log}[x]$
5)	$\int \frac{1}{\sqrt{x}} \, dx$	$2\,\sqrt{x}$

Забележка. Константата *C* не се добавя от системата.

Техника 2 Интегриране чрез полагане

Когато в подинтегралната функция има сложен израз, който "ни пречи", можем да се опитаме да направим смяна на интеграционната променлива, като направим подходящо полагане.

Нека в неопределения интеграл $\int f(x)dx$ да сменим интеграционната променлива x с нова променлива t, чрез полагането

$\boxed{x = \varphi(t)}$,

където $\varphi(t)$ е непрекъсната функция с непрекъсната производна в разглеждания интервал и има обратна функция $\boxed{t = \psi(x)}$, т.е. $\boxed{\varphi(\psi(x)) = x}$

За да преминем към новата променлива t пресмятаме и диференциала

$\boxed{dx = d\varphi(t) = \varphi'(t)dt}$.

Заместваме всичко в изходния интеграл и получаваме:

$$\int f(x)dx = \int f(\varphi(t)).\varphi'(t)dt \qquad (7)$$

6) $\int (x+1)^{100} dx$, полагаме $\left\| \begin{array}{l} x+1=t \\ dx = t'.dt = 1.dt \\ d(x+1) = dx + d(1) = dx \end{array} \right.$, $\boxed{d(x+\lambda) = dx}$

Тогава $\int (x+1)^{100} dx = \int t^{100} dt = \dfrac{t^{101}}{101} + C = \dfrac{(x+1)^{101}}{101} + C$

7) $\int \cos(3x)dx = \dfrac{1}{3}\int \cos(3x)d(3x)$;

Полагаме $t = 3x$, $\boxed{d(\lambda x) = \lambda.dx}$, $d(3t) = 3dt$ и получаваме

$$\frac{1}{3}\int \cos t dt = \frac{1}{3}\sin t + C = \frac{1}{3}\sin 3x + C$$

8) $\int (5x-7)^7\,dx = \frac{1}{5}\int (5x-7)^7\,d(5x-7) = \frac{1}{5}\frac{(5x-7)^8}{8} + C$

9) $\int \frac{1}{2+x^2}\,dx = \int \frac{1}{x^2+\left(\sqrt{2}\right)^2}\,dx = \frac{1}{\sqrt{2}}\arctan\frac{x}{\sqrt{2}} + C$

10) $\int \frac{1}{2-x^2}\,dx = -\int \frac{1}{\left(x^2-2\right)}\,dx = -\int \frac{1}{x^2-\left(\sqrt{2}\right)^2}\,dx = -\frac{1}{2\sqrt{2}}\ln\left|\frac{x-\sqrt{2}}{x+\sqrt{2}}\right| + C$

11) $\int \frac{1}{x\ln x}\,dx$

Полагаме $t=\ln x \Rightarrow dt=(\ln x)'.dx=\dfrac{1}{x}dx$.

Преминаваме от променлива x към t:

$$\int \frac{1}{x\ln x}dx=\int \frac{1}{t}dt=\ln|t|+C=\ln\left|\ln|x|\right|+C$$

12) $\displaystyle\int \cos^2 x\,dx=\int \frac{1+\cos 2x}{2}dx=\int \frac{1}{2}dx+\int \frac{\cos 2x}{2}dx=\frac{1}{2}\int dx+\frac{1}{2}\int \cos 2x dx=$

$$=\frac{1}{2}x+\frac{1}{4}\int \cos 2x d2x=\frac{1}{2}x+\frac{1}{4}\sin 2x+C$$

Упътване: Намаляваме степента с формулата: $\cos^2 x=\dfrac{1+\cos 2x}{2}$

13) $I = \int \cos^3 x\, dx = \int \cos^2 x . \cos x\, dx = \int \cos^2 x\, d\left(\sin x\right)$

Използваме $\cos^2 x + \sin^2 x = 1$

$$I = \int (1 - \sin^2 x) d \sin x = \int 1. d \sin x - \int \sin^2 x . d \sin x = \sin x - \frac{\sin^3 x}{3} + C$$

14) $\int \frac{dx}{x^2(1-x^2)} = \int \frac{1+x^2-x^2}{x^2(1-x^2)} dx = \int \frac{1-x^2}{x^2(1-x^2)} dx + \int \frac{x^2}{x^2(1-x^2)} dx = \int \frac{1}{x^2} dx + \int \frac{1}{1-x^2} dx$

$$= \int x^{-2} dx - \int \frac{1}{x^2-1} dx = \frac{x^{-2+1}}{-2+1} - \frac{1}{2} \ln \left| \frac{x-1}{x+1} \right| = -\frac{1}{x} - \frac{1}{2} \ln \left| \frac{x-1}{x+1} \right| + C$$

15) $\int \frac{x^2+2}{x^2+1} dx = \int \frac{x^2+1+1}{x^2+1} dx = \int \frac{x^2+1}{x^2+1} dx + \int \frac{dx}{x^2+1} = x + \arctan x + C$

16) $\int \frac{2}{x^2(x^2+1)}dx = 2\int \frac{x^2+1-x^2}{x^2(x^2+1)}dx = 2\int \frac{x^2+1}{x^2(x^2+1)}dx - 2\int \frac{x^2}{x^2(x^2+1)}$

$= 2\int \frac{1}{x^2}dx - 2\int \frac{1}{x^2+1}dx = \ldots$

17) $\int \frac{1}{x-3}dx = \int (x-3)^{-1} d(x-3) = \ln|x-3| + C$, с внасяне $d(x+\lambda) = dx$.

18) $\int \frac{1}{\sqrt{1-25x^2}}dx = \frac{1}{5}\int \frac{1}{\sqrt{1-(5x)^2}}d(5x) = \frac{1}{5}\arcsin 5x + C$

19) $\int \frac{1}{9+4x^2}dx = \int \frac{1}{3^2+(2x)^2}dx = \frac{1}{2}\int \frac{1}{3^2+(2x)^2}d(2x) = \frac{1}{6}\arctan\frac{2x}{3} + C$

20) $\int \frac{1}{e^{3x-2}}dx = \frac{1}{-3}\int e^{-3x+2}d(-3x+2) = -\frac{1}{3}e^{-3x+2} + C = -\frac{1}{3e^{3x-2}} + C$

защото $\int e^t dt = e^t + C$.

21) $\int \sin 3x dx = \frac{1}{3}\int \sin 3x d(3x) = -\frac{\cos 3x}{3} + C$

22) $$\int \frac{1}{\sqrt{5x^2-4}}dx = \frac{1}{\sqrt{5}}\int \frac{1}{\sqrt{\left(\frac{x\sqrt{5}}{2}\right)^2-1}}d\left(\frac{x\sqrt{5}}{2}\right) = \frac{1}{\sqrt{5}}\ln\left|\frac{x\sqrt{5}}{2}+\frac{1}{2}\sqrt{5x^2-4}\right|+C$$

23) $$\int \frac{1}{\sqrt{x+2}-\sqrt{x-3}}dx = \int \frac{1}{\left(\sqrt{x+2}-\sqrt{x-3}\right)}.\frac{\left(\sqrt{x+2}+\sqrt{x-3}\right)}{\left(\sqrt{x+2}+\sqrt{x-3}\right)}dx$$

$$= \int \frac{\sqrt{x+2}+\sqrt{x-3}}{5}dx == \frac{1}{5}\int \sqrt{x+2}d(x+2)+\frac{1}{5}\int \sqrt{x-3}d(x-3)$$

$$= \frac{2}{15}(x+2)^{\frac{3}{2}}+\frac{2}{15}(x-3)^{\frac{3}{2}}+C$$

24) $\displaystyle\int \frac{1}{(1+x^2)\arctan x}\,dx = \int \frac{1}{\arctan x}\,d(\arctan x) = \ln\left|\arctan x\right| + C$

25) $\displaystyle I = \int \frac{5x-3}{7+2x^2}\,dx = 5\int \frac{x}{7+2x^2}\,dx - 3\int \frac{dx}{7+2x^2} = I_1 - I_2$

$$I_1 = 5\int \frac{x}{7+2x^2}\,dx = \frac{5}{2}\int \frac{1}{7+2x^2}\,dx^2 =$$

$$\frac{5}{4}\int \frac{1}{7+2x^2}\,d\left(2x^2+7\right) = \frac{5}{4}\ln(7+2x^2) + C$$

,

$$I_2 = 3\int \frac{dx}{7+2x^2} = \frac{3}{\sqrt{2}}\int \frac{1}{\left(\sqrt{7}\right)^2 + \left(x\sqrt{2}\right)^2}\,d\sqrt{2}x = \frac{3}{\sqrt{2}\sqrt{7}}\arctan\left(\frac{\sqrt{2}x}{\sqrt{7}}\right) + C$$

$$I = \ln(7+2x^2) - \frac{3}{\sqrt{2}\sqrt{7}}\arctan(\frac{\sqrt{2}x}{\sqrt{7}}) + C$$

26) $\int(\cos^5 x.\sin x)dx = \int \cos^5 x d(-\cos x) = -\frac{\cos^6 x}{6} + C$

с полагане: $\cos x = t$

$d(\cos x) = -\sin x\, dx = dt$ -> $\boxed{-\sin x.dx = dt}$

27) $\int \frac{x^2}{\cos^2(1+2x^3)}dx = \frac{1}{6}\int \frac{d(2x^3+1)}{\cos^2(1+2x^3)} = \frac{1}{6}\tan(2x^3+1) + C$

Интегриране с Mathematica: Използваме вградената функция :

$$I = \int f(x)dx$$

или

Integrate[f [x], x]

Домашна работа към техники 1 и 2 (на ръка и с Mathematica):

а) $\int \sin^3 x\, dx$,   б) $\int \sin^3 5x\, dx$

Техника 3. Интегриране по части

Нека функциите $u(x), v(x)$ имат непрекъснати първи производни в някакъв интервал D. Тогава от формулата за диференциране на произведение имаме:

$$(uv)' = u'v + uv'.$$

Интегрираме двете части в посочения интервал и получаваме

$$\int (uv)'\,dx = \int (u'v + uv')\,dx \Rightarrow uv = \int u'v\,dx + \int uv'\,dx = \int u\,dv + \int v\,du$$

- след внасяне под знака на диференциала. Оттук следва:

Формула за интегриране по части: $\int u\,dv = u.v - \int v\,du$

28) $\int xe^{2x}dx = \int xd\left(\frac{1}{2}e^{2x}\right) = \frac{x}{2}e^{2x} - \int \frac{1}{2}e^{2x}dx = \frac{x}{2}e^{2x} - \frac{1}{4}e^{2x} + C$

$$= \left(\frac{x}{2} - \frac{1}{4}\right)e^{2x} + C$$

29) $I = \int \arctan x.dx = x.\arctan x - \int x.d(\arctan x)$

Тук сме интегрирали веднъж по части като $u = \arctan x, \quad v = x$

По-нататък използваме, че

$$\boxed{d(\arctan x) = (\arctan x)'.dx = \frac{1}{1+x^2}dx}$$

Така получаваме:

$$I = x.\arctan x - \int x.d(\arctan x) = x.\arctan x - \int \frac{x}{1+x^2}dx =$$

$$x.\arctan x - \frac{1}{2}\int \frac{d(x^2)}{1+x^2} = x.\arctan x - \frac{1}{2}\ln\left(1+x^2\right) + C$$

30) $\int x.\arctan x \; dx = \frac{1}{2}\int \arctan x \; d(x^2) =$

(внасяме x подзнака на доференциала):

$$= \frac{1}{2}x^2.\arctan x - \frac{1}{2}\int x^2.d(\arctan x) =$$

$$\frac{1}{2}x^2.\arctan x-\frac{1}{2}\int\frac{x^2}{1+x^2}dx=$$

$$\frac{x^2\arctan x}{2}-\frac{1}{2}\int\frac{1+x^2}{1+x^2}dx-\int\frac{-1}{1+x^2}dx=\ldots$$

31) $I=\int e^x\sin x.dx=-\int e^x d(\cos x)=-e^x\cos x+\int\cos x.d(e^x)=$

$-e^x\cos x+\int e^x\cos x.dx=-e^x\cos x+e^x\cos x-\int\sin x.d(e^x)=$

$-e^x\cos x+e^x\cos x-\int e^x\sin x.dx$

т.е. получихме отдясно търсения интеграл $\int e^x\sin x\,dx$.

Или уравнението относно интеграла $I = e^x(\sin x - \cos x) - I$.

Оттук: $2I = e^x(\sin x - \cos x) \Rightarrow I = \dfrac{1}{2}e^x(\sin x - \cos x) + C$

<u>Домашна работа:</u> а) $\int e^{2x}\sin x\,dx$, б) $\int x^2 \arctan x\,dx$. Упътване към б): Внася се “по-леката” функция (виж по-нататък).

Тежест на функциите при интегриране по части

най-лека	e^x
леки	$\sin x$, $\cos x$
тежки	$1, x, x^2, \ldots, x^n$
най-тежки	$\arctan x$, $\arcsin x$,...

Допълнителни задачи:

32) $\int \operatorname{tg} x\, dx = \int \frac{\sin x}{\cos x} dx = -\int \frac{1}{\cos x} d(\cos x) = -\ln|\cos x| + C$

33) $\int \frac{1}{\sin x} dx = \int \frac{1}{2\sin\frac{x}{2}\cos\frac{x}{2}} dx = \frac{1}{2}\int \frac{\cos\frac{x}{2}}{2\sin\frac{x}{2}\cos^2\frac{x}{2}} dx =$

$$\frac{1}{2}.2\int \frac{1}{\cos^2\frac{x}{2}}.\frac{1}{\tan\frac{x}{2}} d\left(\frac{x}{2}\right) = \int \frac{1}{\tan\frac{x}{2}} d\left(\tan\frac{x}{2}\right) = \ln\left|\tan\frac{x}{2}\right| + C$$

5. Определен интеграл – свойства, пресмятане

Определеният интеграл от дадена реална функция $f(x)$ в интервал $[a, b]$ е число. Геометрично това число е равно на лице, обем и пр. В двумерния случай това е лицето на фигурата, заключена между графиката на $f(x)$, абсцисната ос и вертикалните линии през краищата на интервала $[a, b]$, като частта под Ox се изважда.

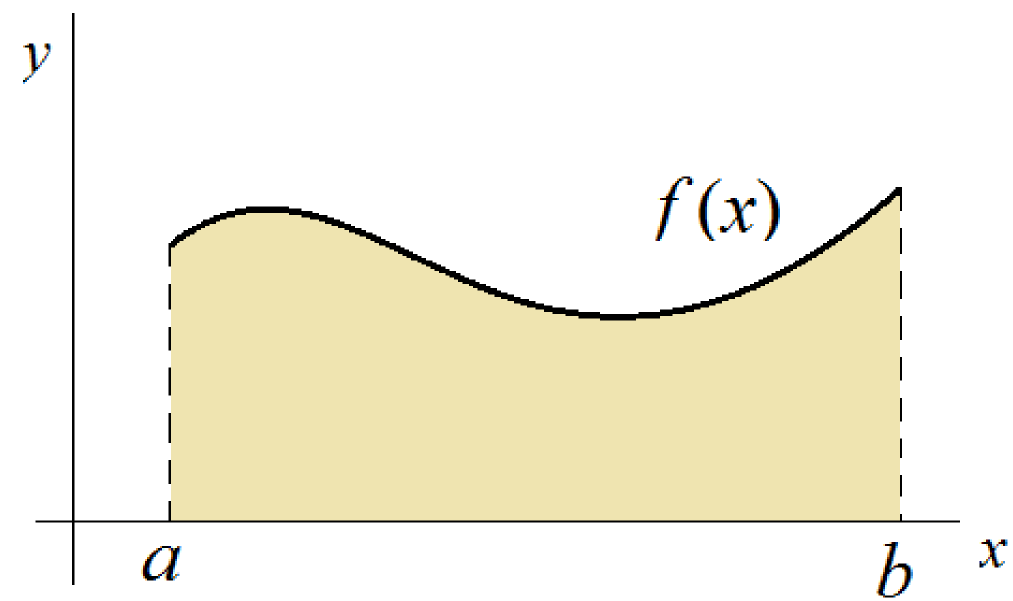
Записано по-точно, определеният интеграл

$$I = \int_a^b f(x)dx$$

е число, равно на лицето на фигурата, заключено между

$\{x = a, x = b, y = 0,$ и $f(x)\}$

Четем „интеграл от „а” до „бе” от еф от хикс де хикс” .



Горното неформално определение е въведено от големия немски математик Георг Фридрих Бернхард Риман през 1854 г. То се формализира по следните два начина:

5.1. Определение на риманов интеграл чрез интегрални суми

Нека реалната функция $f(x)$ е дефинирана в интервал $[a, b]$. Избираме цяло число $n > 0$ и разделяме интервала $[a, b]$ на n части с произволни различни точки x_k, $k = 0,1,\ldots,n,$ така че:

$$a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$$

Така получаваме n подинтервала $[x_{k-1}, x_k]$ с дължини $\Delta x_k = x_k - x_{k-1}$. Нека най-голямата дължина от всичките интервалчетата е $d = \max(\Delta x_k)$. Във всеки подинтервал избираме произволна точка $\xi_k \in [x_{k-1}, x_k]$.

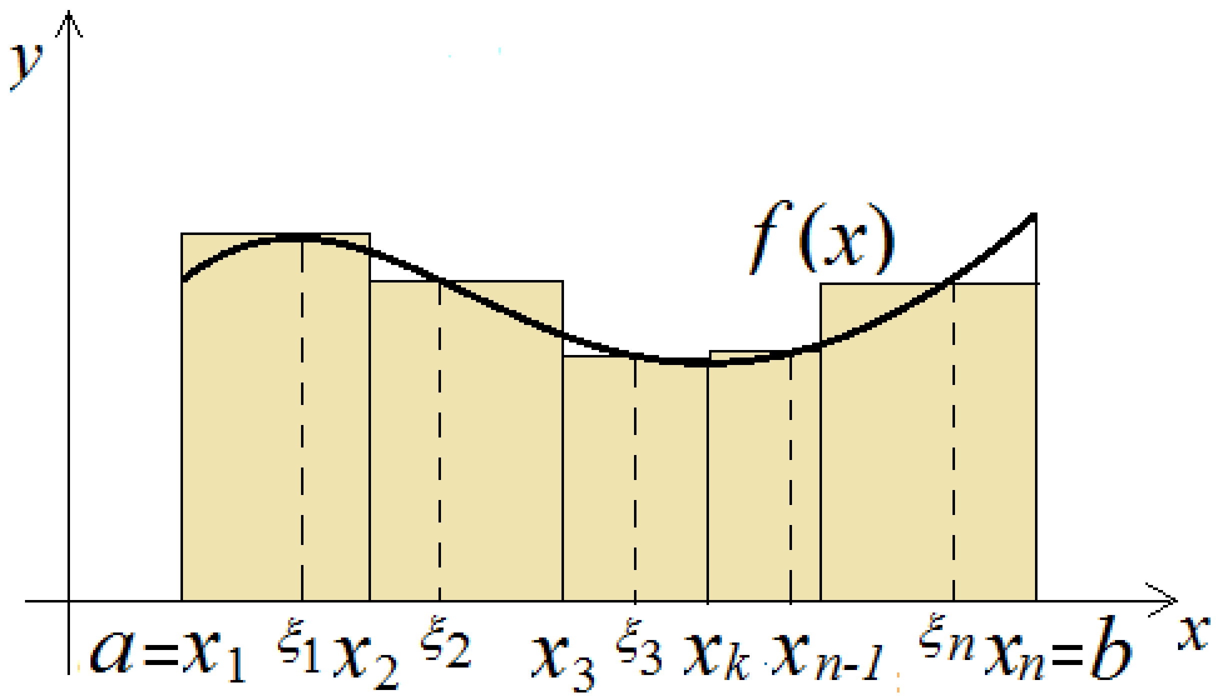
Интегрална сума се нарича сумата

$$S_x = \sum_{k=1}^{n} f(\xi_k)\Delta x_k .$$

Определение 5.1 Ако при произволно разделяне на подинтервали така че $d \to 0$ и независимо от избора на ξ_k съществува граница на интегралните суми, то тази граница се нарича определен (риманов) интеграл на функцията $f(x)$ в интервала $[a, b]$, т.е.

$$I = \int_a^b f(x)dx = \lim_{d \to 0} S_x .$$

Казваме, че в този случай $f(x)$ е интегруема в $[a, b]$.



$$I = \int_a^b f(x)dx \approx S_x = \sum_{k=1}^{n} f(\xi_k)\Delta x_k$$

5.2 Определение на риманов интеграл чрез суми на Дарбу

Нека реалната функция $f(x)$ е дефинирана в интервал $[a, b]$. Отново ако $n > 0$ е някое цяло число, разделяме интервала $[a, b]$ на n части с произволни различни точки $x_k,\ k = 0, 1, \ldots, n,$ такива че:

$$a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$$

и получаваме n подинтервала $[x_{k-1}, x_k]$ с дължини $\Delta x_k = x_k - x_{k-1}$.

Във всеки подинтервал допускаме, че съществуват точна горна граница на функцията и точна долна граница на функцията, които означаваме с

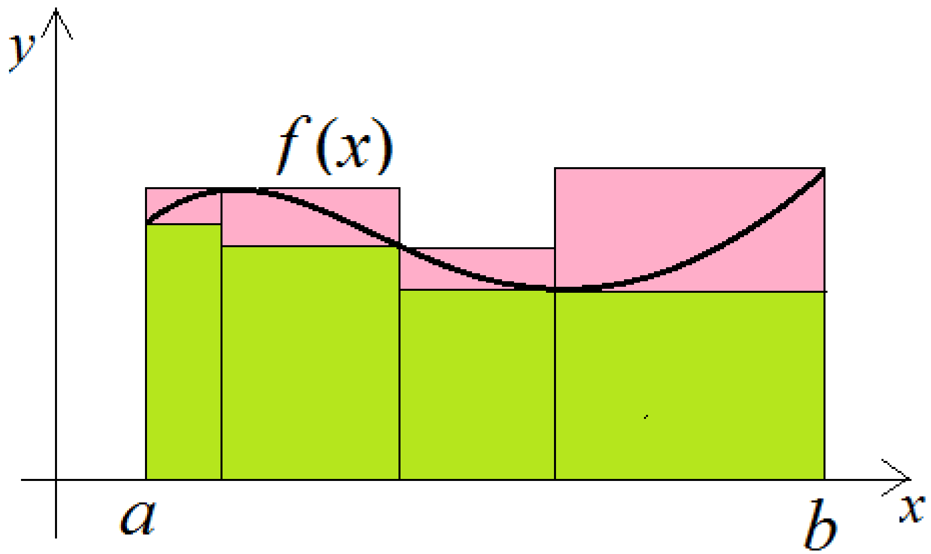
$$M_k = \sup_{x \in [x_{k-1}, x_k]} f(x) \quad \text{и} \quad m_k = \inf_{x \in [x_{k-1}, x_k]} f(x).$$

Долна сума на Дарбу се нарича сумата $s_n = \sum_{k=1}^{n} m_k . \Delta x_k$.

Горна сума на Дарбу е сумата: $S_n = \sum_{k=1}^{n} M_k . \Delta x_k$.

Имаме $s_n \leq S_n$.

Сумите на Дарбу в случая на 4 интервала от долната фигура са: долна сума на Дарбу е сумата от лицата на зелените квадратчета, а горна сума на Дарбу е сумата от зелените и розовите квадратчеата.



Определение 5.2 Ако при $n \to \infty$, $d \to 0$, съществуват граници на малката и голямата суми на Дарбу и границите им са равни, то казваме че тази граница е определен риманов интеграл на функцията $f(x)$ в интервала $[a, b]$, т.е.

$$I = \int_a^b f(x)dx = \lim_{n\to\infty} S_n = \lim_{n\to\infty} s_n$$

Казваме, че в този случай $f(x)$ е интегруема по Риман в $[a, b]$.

5.3. СВОЙСТВА НА ОПРЕДЕЛЕНИЯ ИНТЕГРАЛ

Свойство 1^0. Умножение с константа $\lambda \neq 0$

$$\lambda \int_a^b f(x)dx = \int_a^b \lambda f(x)dx = \int_a^b f(x)d(\lambda x)$$

Свойство 2^0 . Интеграл от сума/разлика на функции

$$\int_a^b [f(x) \pm g(x)]dx = \int_a^b f(x)dx \pm \int_a^b g(x)dx$$

Свойство 3^0. Адитивност (сума по области)

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx, \quad a \leq c \leq b.$$

Свойство 4^0. Размяна на границите

$$\int_a^b f(x)dx = -\int_b^a f(x)dx$$

Свойство 5^0. Запазва положителния си знак

Ако $f(x)$ е интегруема в $[a, b]$ и нека $f(x) \geq 0$ за $x \in [a,b]$. Тогава

$$\int_a^b f(x)dx \geq 0.$$

Свойство 6^0. Непрекъснатост и прекъснатост на функцията

В сила са следните твърдения, които привеждаме без доказателства.

А) Всяка непрекъсната в $[a, b]$ функция $f(x)$ е интегруема по Риман.

Б) Прекъснатите функции може да са, а може и да не са интегруеми.

В) Функцията е интегруема по Риман и ако е ограничена и множеството от точките, където е ограничена са краен брой и/или изброимо множество.

Свойство 7^0. Смяна на променливите в определения интеграл

Нека функцията $f(x)$ е непрекъсната в отворения интервал (a, b). Нека също така, функцията $\varphi(t)$ има непрекъсната производна в отворения интервал (α, β) при което $\varphi(\alpha)=a,\ \varphi(\beta)=b.$

Тогава

$$\int_a^b f(x)dx = \int_\alpha^\beta f(\varphi(t))d\varphi(t).$$

5.4 Основна теорема на интегралното смятане за връзка между неопределен и определен интеграл (Теорема на Нютон-Лайбниц)

Ако $f(x)$ е дефинирана в $[a, b]$ и има примитивна $F(x)$, т.е.

$\int f(x)dx = F(x) + C$, то: $\int_a^b f(x)dx = F(b) - F(a) = F(x)\Big|_a^b$,

където $F(x)$ е една която и да е примитивна на $f(x)$.

За основната теорема на интегралното смятане прочетете повече на:

http://demonstrations.wolfram.com/TheFundamentalTheoremOfCalculus/

Примери:

34) $\int_0^1 x^3 dx = \left.\frac{x^4}{4}\right|_0^1 = \frac{1^4}{4} - \frac{0^4}{4} = \frac{1}{4}$

35) $\int_0^{\frac{\pi}{2}} \cos x dx = \sin x\Big|_0^{\frac{\pi}{2}} = \sin\left(\frac{\pi}{2}\right) - \sin 0 = 1 - 0 = 1$

36) Намерете лицето на фигурата, ограничено от графиките на функциите:

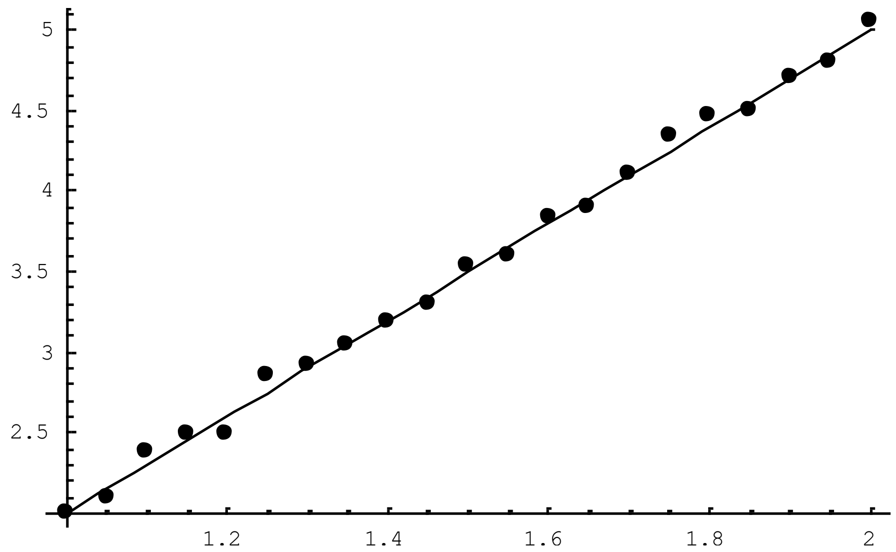
$f(x) = 2x - 1,\ \ g(x) = 2x^2 + x + 1$.

6. Приближение на таблични данни. Интерполационен полином на Лагранж. Метод на най-малките квадрати (МНМК)

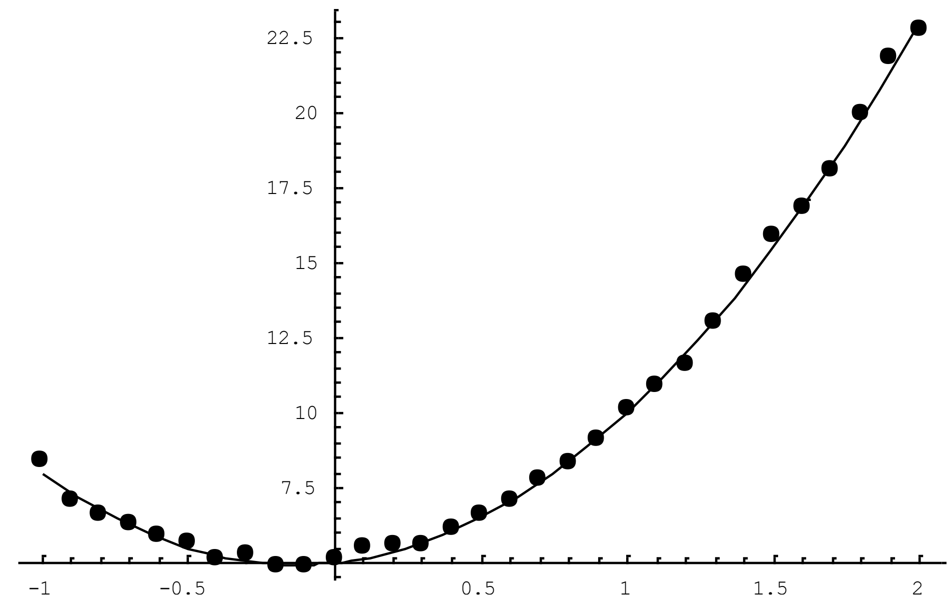
6.1. Основни етапи при приближаване на функции, зададени таблично

Приближаването на функции е основна задача от практиката. В общия случай функцията е известна в дискретно (крайно) множество от точки и търсим нейна стойност в някаква друга точка. Естествено, ние не я знаем. Затова стандартният подход при приближаване на финкции е:

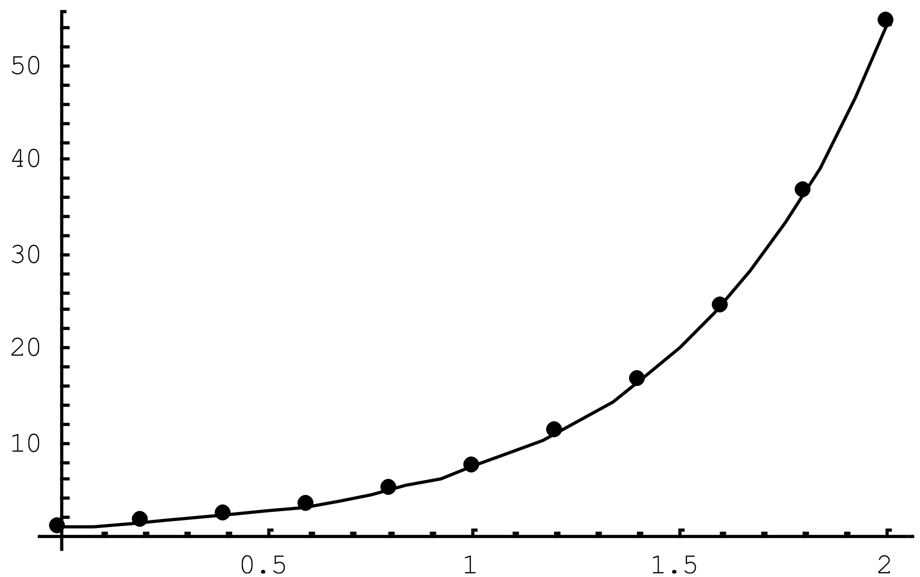
1. Начертаваме графиката на функцията по известните точки $(x_i, y_i),\ \ i = 0,1,\ldots,n$.
2. Опитваме се да “познаем” вида (класа) на търсената функция по графиката; тя може да прилича, например на: полином от някаква степен (права линия – полином от първа степен, парабола – полином от втора степен и т.н.), тригонометрична функция, експоненциална, логаритмична и т.н. Примери:



$$y \approx ax + b$$



$$y \approx ax^2 + bx + c$$



$$y \approx e^{\lambda x}$$

3. Според броя на точките и класа на функцията и други характеристики на приближаваната функция използваме различен метод за нейното приближаване.
4. Получаваме някаква формула, “близка” до данните.
5. Използваме получената приближаваща формула, за да пресметнем стойности на функцията в точки, в които не разполагаме с данни.

Методи за приближаване:

Интерполиране,

Метод на най-малките квадрати,

Интерполиране със сплайни и др.

6.2 Задача на интерполирането

Постановка на задачата. Нека функцията $y = f(x)$ е дефинирана в някакъв интервал и е известна таблица от стойностите ѝ:

x_i	x_0	x_1	...	x_n
y_i	y_0	y_1	...	y_n

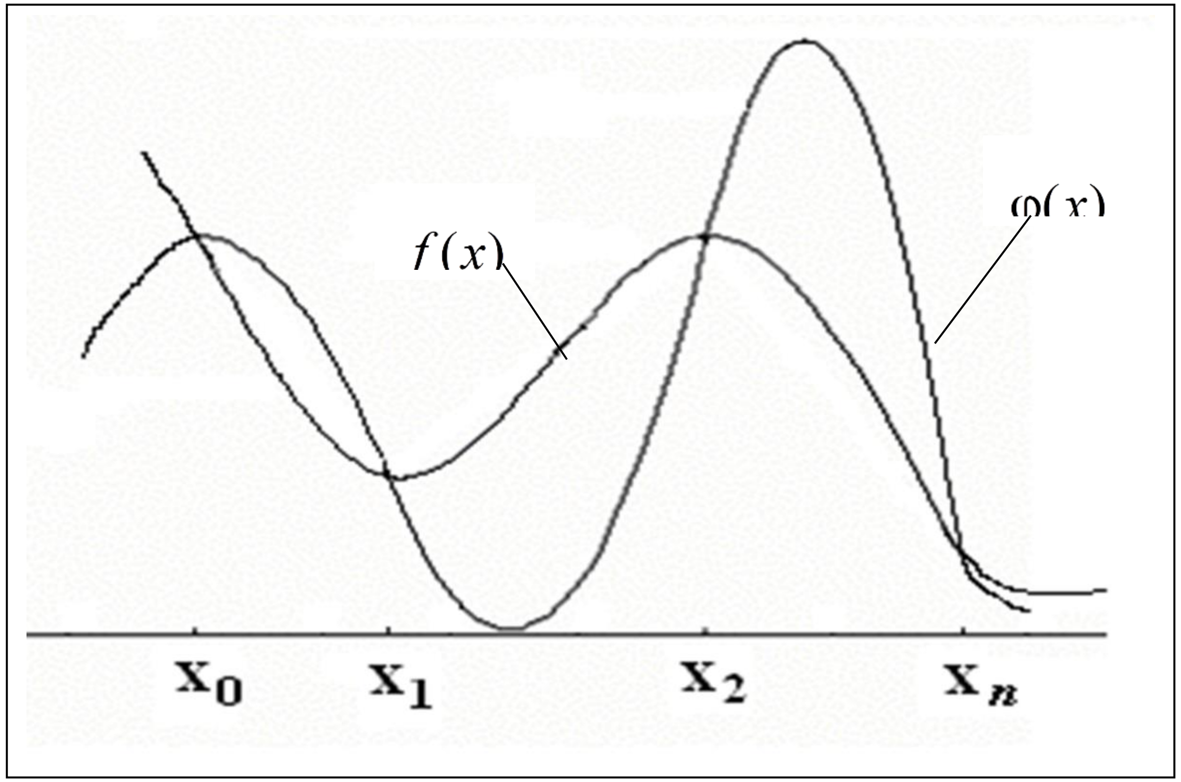
Търси се приближаваща фунция $\varphi(x)$, такава че

$$\varphi(x_i) = y_i,\ i = 0,1,\dots,n. \qquad (1)$$

Тази задача се решава при задаване класа на функциите $\varphi(x)$.

Определение. 6.1. Точките $x_0, x_1, \dots, x_n$ се наричат възли на интерполирането.

Графично условието (1) означава, че приближаващата функция $\varphi(x)$ минава през точките (x_i, y_i), тъй като има в x_i същите стойности като $f(x)$ - виж следващата фигура.



Примери.

1) Ако изберем системата от функции $1, x, x^2, \dots, x^n$ и разгледаме всички линейни комбинации по тази система, ще получим класа на

полиномите от n-та степен $P_n(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$.
Интерполационният полином ще се определя от условието
$P_n(x_i) = y_i,\ i = 0,1,\ldots,n$.

2) Ако изберем системата функции

$$\frac{1}{2}, \cos(x), \sin(x), \cos(2x), \sin(2x), \ldots, \cos(nx), \sin(nx), \ldots$$

ще получим класа на тригонометричните полиноми от n-та степен
$T_n(x) = a_0 + a_1 \cos(x) + b_1 \sin(x) + \ldots + a_n \cos(nx) + b_n \sin(nx)$ и т.н.

Според вида на всяка конкретна функция $f(x)$ най-напред се определя класът от приближаващи функции, а след това и съответната интерполираща функция (полином) $\varphi(x)$ по условието (1).

6.3. Интерполационен полином на Лагранж

Дадена е таблицата със стойности на функцията $y = f(x)$:

x_i	x_0	x_1	...	x_n
y_i	y_0	y_1	...	y_n

Търси се приближаваща функция във вид на полином $P_n(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$, така че

$$P_n(x_i) = y_i,\; i = 0,1,\ldots,n. \tag{2}$$

Без ограничение по-нататък ще считаме, че възлите са подредени във възходящ ред, т.е. $x_0 < x_1 < \ldots < x_n$.

Съществуването на такъв полином се осигурява със следния интерполационен полином на Лагранж, който има вида:

$$L_n(x) = \sum_{i=0}^{n} y_i F_i(x), \tag{3}$$

където $F_i(x)=\prod_{\substack{j=0\\ j\neq i}}^{n}\frac{(x-x_j)}{(x_i-x_j)}$ са коефициентите на Лагранж.

Те имат свойството: $F_i(x_i)=1,\ F_i(x_j)=0,\ j\neq i$.

Формулата добива вида: $$L_n(x)=\sum_{i=0}^{n} y_i \prod_{\substack{j=0\\ j\neq i}}^{n}\frac{(x-x_j)}{(x_i-x_j)} \qquad (4)$$

Формулата (3) или (4) се записва подробно така:

$$L_n(x) = y_0 \frac{(x-x_1)\dots(x-x_n)}{(x_0-x_1)\dots(x_0-x_n)} + y_1 \frac{(x-x_0)(x-x_2)\dots(x-x_n)}{(x_1-x_0)(x_1-x_2)\dots(x_1-x_n)} + \dots$$

$$+ y_k \frac{(x-x_0)\dots(x-x_{k-1})(x-x_{k+1})\dots(x-x_n)}{(x_k-x_0)\dots(x_k-x_{k-1})(x_k-x_{k+1})\dots(x_k-x_n)} + \qquad (5)$$

$$\dots + y_n \frac{(x-x_0)\dots(x-x_{n-1})}{(x_n-x_0)\dots(x_n-x_{n-1})}$$

Проверяваме, че полиномът удовлетворява условията (2).

$\boxed{\text{При } x = x_0}$:

$$L_n(x_0) = y_0 \frac{(x_0 - x_1)\dots(x_0 - x_n)}{(x_0 - x_1)\dots(x_0 - x_n)} + y_1 \frac{(x_0 - x_0)(x_0 - x_2)\dots(x_0 - x_n)}{(x_1 - x_0)(x_1 - x_2)\dots(x_1 - x_n)} + \dots$$

$$+ y_n \frac{(x_0 - x_0)\dots(x_0 - x_{n-1})}{(x_n - x_0)\dots(x_n - x_{n-1})} = y_0.1 + 0 + \dots + 0 = y_0$$

$\boxed{\text{При } x = x_1}$:

$$L_n(x_1) = y_0 \frac{(x_1 - x_1)\dots(x_1 - x_n)}{(x_0 - x_1)\dots(x_0 - x_n)} + y_1 \frac{(x_1 - x_0)(x_1 - x_2)\dots(x_0 - x_n)}{(x_1 - x_0)(x_1 - x_2)\dots(x_1 - x_n)} + \dots$$

$$+ y_n \frac{(x_1 - x_0)(x_1 - x_1)\dots(x_1 - x_{n-1})}{(x_n - x_0)\dots(x_n - x_{n-1})} = y_0.0 + y_1.1 + \dots + 0 = y_1$$

и т.н. за всички възли x_i.

В сила е следната

Теорема за съществуване и единственост на интерполационния полином на Лагранж

Ако всички интерполационни възли са различни, т.е. $x_i \neq x_j,\ i \neq j$, то интерполационният полином на Лагранж е единствен.

Доказателство. Съществуването бе показано по-горе с проверката на условията за интерполиране. Но твърдението може да се докаже и независимо от този факт. Заместваме всеки от възлите в условията (2):

$$x = x_0: \quad P_n(x_0) = a_0 + a_1 x_0 + a_2 x_0^2 + \ldots + a_n x_0^n = y_0$$

$$x = x_1: \quad P_n(x_1) = a_0 + a_1 x_1 + a_2 x_1^2 + \ldots + a_n x_1^n = y_1 \qquad (6)$$

....

$$x = x_n: \quad P_n(x_n) = a_0 + a_1 x_n + a_2 x_n^2 + \ldots + a_n x_n^n = y_n$$

Тук коефициентите на полинома са n+1 неизвестни $a_0, a_1, \ldots, a_n$.

От линейната алгебра е известно, че необходимото и достатъчно условие за съществуване и единственост на решение на линейната система с ненулева дясна част (което е общия за нас случай) е детерминантата на системата да е различна от нула. Така получаваме следната детерминанта, известна като детерминанта на Вандермонд

$$\det = \begin{vmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \dots & \dots & & & \\ & & & & \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{vmatrix} = \prod_{i>j,\, j=0}^{n} (x_i - x_j) \neq 0.$$

Очевидно, когато всички възли са ралични, $x_i \neq x_j,\ i \neq j$, стойността на тази детерминанта е различна от нула, с което теоремата е доказана.

Без доказателство ще приведем следната

Теорема за оценка на грешката от интерполиране

Нека функцията $f(x)$ е дефинирана и непрекъсната в интервала $[a,b]$ и освен това съществуват и са непрекъснати в $[a,b]$ и производните ѝ до n+1-ви ред: $f'(x), f''(x),..., f^{(n+1)}(x)$. Нека по стойностите на функцията във възлите $x_0 < x_1 < ... < x_n$, $x_i \in [a,b]$ е построен интерполационнният полином на Лагранж. Тогава за всяка т. x от дефи- ниционната област е в сила следната оценка за грешката в тази точка:

$$f(x) - L_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)(x-x_1)...(x-x_n), \quad \xi \in (a,b) \qquad (7)$$

или като оценка на абсолютната грешка:

$$\left| f(x) - L_n(x) \right| \le \frac{M_{n+1}}{(n+1)!}\left|(x-x_0)(x-x_1)...(x-x_n)\right|, \quad M_{n+1} = \max_{[a,b]}\left| f^{(n+1)}(\xi) \right| \qquad (8)$$

Особености при интерполиране с алгебрични полиноми

1. От оценката на грешката (7)-(8) лесно се съобразява, че грешката е малка, когато възлите са близко един до друг, така, че произведението $\left|(x-x_0)(x-x_1)...(x-x_n)\right|$ да бъде число близко до нула.
2. По същите съображения е добре така да избираме възлите, че точката x, в която търсим стойност да бъде по възможност между възлите на интерполиране. Когато x е извън интервала на възлите, казваме, че се прави екстраполиране. В някои случаи е възможна само екстраполация, но грешката е по-голяма.
3. Теоретично е доказано, че при степен $n > 6$ интерполационният полином обикновено все по-силно „скача“ около функцията. Затова се интерполира с неголеми n.

Пример 6.1 Дадена е таблица на фунцията $y = f(x) = \sqrt{x+3}$:

x_i	1.0	1.2	1.4	1.6	1.8	2.0
y_i	2.	2.049	2.098	2.145	2.191	2.236

А) Да се намери приближена стойност в точката $x' = 1.65$ с полином на Лагранж от втора степен.

Б) Да се оцени грешката от интерполирането.

<u>*Решение.*</u> За да построим полинома от втора степен (n=2) са необходими 3 възела. Съгласно горните забележки, като знаем, че $x' = 1.65$, да изберем възли на интерполирането $x_0 = 1.6, x_1 = 1.8, x_2 = 2$, така че x' да е вътре в интервала $[1.65, 2]$. Това ще намали грешката. Формулата на полинома от втора степен, съгласно (5) е:

$$L_2(x) = y_0 \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} + y_1 \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} + y_2 \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} \qquad (9)$$

Като заместим $x = x'$ в (9), приближената стойност е:

$$L_2(x') = y_0 \frac{(x' - x_1)(x' - x_2)}{(x_0 - x_1)(x_0 - x_2)} + y_1 \frac{(x' - x_0)(x' - x_2)}{(x_1 - x_0)(x_1 - x_2)} + y_2 \frac{(x' - x_0)(x' - x_1)}{(x_2 - x_0)(x_2 - x_1)}$$

Заместваме тук нашите данни и получаваме:

$$L_2(1.65) = 2.145 \frac{(1.65 - 1.8)(1.65 - 2)}{(1.6 - 1.8)(1.6 - 2)} + 2.191 \frac{(1.65 - 1.6)(1.65 - 2)}{(1.8 - 1.6)(1.8 - 2)}$$

$$+ 2.236 \frac{(1.65 - 1.6)(1.65 - 1.8)}{(2 - 1.6)(2 - 1.8)}$$

$$L_2(1.65) = 2.145\frac{(-0.15)(-0.35)}{(-0.2)(-0.4)} + 2.191\frac{(0.05)(-0.35)}{(0.2)(-0.2)} + 2.236\frac{(0.05)(-0.15)}{(0.4)(0.2)}$$

$$L_2(1.65) = 2.145 * 0.6562 + 2.191 * 0.4375 + 2.236 * (-0.09375) \approx 2.15659.$$

Нашите данни са с 3 знака след десетичната точка (неотстранима грешка = 0.001), затова закръгляме засега резултата на 2.157.

Код на *Mathematica* за изчисляване по формулата (5):

```
x =.; xx = 1.65;
xi = {1.6, 1.8, 2}; yi = {2.145, 2.191, 2.236};
```

$$L2[x\_] := yi_{[[1]]}\frac{(x - xi_{[[2]]})(x - xi_{[[3]]})}{(xi_{[[1]]} - xi_{[[2]]})(xi_{[[1]]} - xi_{[[3]]})} + yi_{[[2]]}\frac{(x - xi_{[[1]]})(x - xi_{[[3]]})}{(xi_{[[2]]} - xi_{[[1]]})(xi_{[[2]]} - xi_{[[3]]})} + yi_{[[3]]}\frac{(x - xi_{[[1]]})(x - xi_{[[2]]})}{(xi_{[[3]]} - xi_{[[1]]})(xi_{[[3]]} - xi_{[[2]]})}$$

```
Expand[L2[x]]
Print[" Приближена стойност L2(xx)=", L2[xx]]
```

$$\text{Print}\left[\text{" Точна стойност f(xx)=", } \sqrt{xx + 3}\right]$$

$$1.741 + 0.2725\,x - 0.0125\,x^2$$

```
Приближена стойност L2(xx)=2.15659
Точна стойност f(xx)=2.15639
```

```
(* Проверка за правилност на получения полином *)
L2[1.6]
L2[1.8]
L2[2]
  2.145
  2.191
  2.236
---------------
```

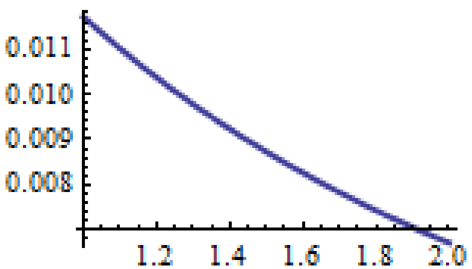
Наистина като заместим 1.6, 1.8, 2 получаваме точно първоначалните данни от таблицата на функцията. Значи това е търсеният полином на Лагранж.

<u>Оценка на грешката:</u> В случая полиномът е от втора степен, затова по формула (8) пресмятаме третата производна и правим графика на третата производна, за да видим къде е максимумът по абсолютна стойност:

```
f[x_] := √(x + 3)    (* Дефинираме функцията*)
∂x,x,x f[x]   (* Намираме трета производна*)
```

Получаваме: $\dfrac{3}{8\,(3+x)^{5/2}}$.

```
f3[x_] := Abs[3/(8 (3 + x)^(5/2))]
(* Дефинираме трета производна - абс. стойност*)
Plot[f3[x], {x, 1, 2}]
```



От графиката определяме, че максимумът по абсолютна стойност на третата производна е в точката 1. Заместваме по формулата (8):

```
M3 = f3[1.]
```

$$greshka = M3\,\frac{\text{Abs}[(xx - xi_{[[1]]})\ (xx - xi_{[[2]]})\ (xx - xi_{[[2]]})]}{3\,!}$$

```
0.0117188
```

2.19727×10^{-6}

Получихме теоретична грешка на метода от порядъка 0.0000022. Но тъй като неотстранимата ни грешка беше ε= 0.001, то се взима по-лошата грешка, т.е. 0.001. Така крайният резултат на приближената стойност с интерполационния полином на Лагранж е вярна само с три знака, или: $f(1.65)=\sqrt{1.65+3}\approx 2.157$. Като сравним с точната стойност, която в случая знаем и е 2.15639, виждаме, че това отговаря на истината.

Освен това, като прегледаме първоначалните данни виждаме, че точката $x'=1.65$ е между точките 1.6 и 1.8, в които стойностите на функцията са съответно :
за 1.6 → 2.145, за 1.8 → 2.191,
а ние получихме приближено за 1.65→ 2.157, т.е. между тях. Което прилича на истина, дори и да не знаехме точната стойност! Това е един от критериите за правилност на резултата в общия случай.

В този пример успяхме да намерим грешката, като използвахме формула (8). В реалния случай обаче, формула на функцията не е известна, а само нейни значения в определен брой точки. Дори обикновено не се знае дали съществуват някакви производни и до какъв ред. Затова теоретичното оценяване на грешката има ограничено приложение и е съществено при прилагане на формулите, избор на точките на интерполиране и др. (Виж особености при интерполирането по-горе).

За упражнение:
1. Решете същата задача с полином от трета степен.
2. Постройте полином от втора степен с възли 1.4, 1.6, 1.8 и намерете приближената стойност на функцията в същата точка 1.65.

6.4 Метод на най-малките квадрати (МНМК) с алгебрични полиноми

Отново разглеждаме случая, когато търсим да приближим реална функция на една реална променлива.

<u>Изходна постановка.</u> Нека функцията $y = f(x)$ е дефинирана в някакъв интервал, в който е известна таблица от стойностите й

x_i	x_1	x_2	...	x_N
y_i	y_1	y_2	...	y_N

Тук точките $x_1, x_2, \ldots, x_N$ не е задължително да са различни.

Основно предположение: броят на данните N е много по-голям от степента m на приближаващия полином

$$P_m(x) = c_0 + c_1 x + \ldots + c_m x^m \qquad (10)$$

$P_m(x) = c_0 + c_1 x + \ldots + c_m x^m$

т.е. за разлика от Лагранж, сега $m << N$.

Означаваме разликата (остатъка) между стойностите на функцията y_i и полинома $P_m(x_i)$ за всяка точка x_i с

$$r_i = P_m(x_i) - y_i = c_0 + c_1 x_i + c_2 x_i^2 + \ldots + c_m x_i^m - y_i, \; i = 1, \ldots, N. \qquad (11)$$

Критерият за намиране на коефициентите c_j в (9) по МНМК е сумата от квадратите на всички остатъци да бъде възможно най-малка. Стигаме до задачата: Измежду всички полиноми от дадена степен m да се намери този, за който

$$\sum_{i=1}^{N} r_i^2 = \min \qquad (12)$$

Минимумът тук зависи от $c_0, c_1, \ldots, c_m$. По-подробно от (11) и (12) трябва да намерим минимума на функцията с променливи c_j:

$$\Phi(c_0, c_1, \ldots, c_m) = \sum_{i=1}^{N} r_i^2 = \sum_{i=1}^{N} (c_0 + c_1 x_i + c_2 x_i^2 + \ldots + c_m x_i^m - y_i)^2. \qquad (13)$$

Ще използваме без доказателство следната теорема от математическия анализ:

Необходимото условие за екстремум на диференцируема функция на много променливи в околност на дадена точка е всички частни производни да са равни на нула, т.е.

$$\left|\begin{array}{l} \dfrac{\partial \Phi(c_0, c_1, \dots, c_m)}{\partial c_0} = 0 \\ \dfrac{\partial \Phi(c_0, c_1, \dots, c_m)}{\partial c_1} = 0 \\ \dots \\ \dfrac{\partial \Phi(c_0, c_1, \dots, c_m)}{\partial c_m} = 0 \end{array}\right. . \qquad (14)$$

За първото уравнение в (14) диференцираме (13) спрямо c_0 при фиксирани останали коефициенти c_j (тук x_i, y_i са дадени). Имаме:

$$\frac{\partial \Phi}{\partial c_0} = 2\sum_{i=1}^{N}(c_0 + c_1x_i + c_2x_i^2 + \ldots + c_mx_i^m - y_i) = 0.$$

За второто уравнение аналогично:

$$\frac{\partial \Phi}{\partial c_1} = 2\sum_{i=1}^{N}(c_0 + c_1x_i + c_2x_i^2 + \ldots + c_mx_i^m - y_i).x_i = 0, \ldots$$

$$\frac{\partial \Phi}{\partial c_m} = 2\sum_{i=1}^{N}(c_0 + c_1x_i + c_2x_i^2 + \ldots + c_mx_i^m - y_i).x_i^m = 0.$$

След приведение пред неизвестните стигаме до системата

$$\left|\begin{array}{l}
Nc_0 + \left(\sum_{i=1}^{N} x_i\right)c_1 + \left(\sum_{i=1}^{N} x_i^2\right)c_2 + \ldots + \left(\sum_{i=1}^{N} x_i^m\right)c_m = \sum_{i=1}^{N} y_i \\
\left(\sum_{i=1}^{N} x_i\right)c_0 + \left(\sum_{i=1}^{N} x_i^2\right)c_1 + \ldots \qquad + \left(\sum_{i=1}^{N} x_i^{m+1}\right)c_m = \sum_{i=1}^{N} x_i y_i \\
\ldots \\
\left(\sum_{i=1}^{N} x_i^m\right)c_0 + \left(\sum_{i=1}^{N} x_i^{m+1}\right)c_1 + \ldots \qquad + \left(\sum_{i=1}^{N} x_i^{2m}\right)c_m = \sum_{i=1}^{N} x_i^m y_i \quad .
\end{array}\right. \qquad (15)$$

Очевидно това е линейна система за намиране на $c_0, c_1, \ldots, c_m$.

Може да се покаже, че тази система има единствено решение. При немного големи m тя може да се решава с детерминанти или по метода на

Гаус. Обикновено МНМК с алгебрични полиноми от вида (10) се прилага за $m \le 5$.

Определение. Полиномът $P_m^*(x) = c_0^* + c_1^* x + \ldots + c_m^* x^m$, чиито коефициенти $c_0^*, c_1^*, \ldots, c_m^*$ са решения на (15) се нарича полином на най-добро приближение по МНМК към функцията $f(x)$.

Определение. Грешка на приближението по МНМК или средноквадратична грешка се нарича отклонението на полинома на най-добро приближение по МНМК

$$\delta_m = \sqrt{\sum_{i=1}^{N} r_i^2} = \sqrt{\sum_{i=1}^{N} (c_0^* + c_1^* x_i + c_2^* x_i^2 + \ldots + c_m^* x_i^m - y_i)^2}. \tag{16}$$

Как се избира най-добрата степен m за полинома по МНМК?

Това в общия случай е труден проблем. Обикновено постъпват така:

Нека данните имат неотстранима грешка ε. Започва се с $m=1$, изчислява се δ_1. Ако $\delta_1 \approx \varepsilon$, значи решението е добро. Ако ср.кв. грешка не е приблизително равна на ε, се продължава с $m=2$ и δ_2 и т.н., докато се намери оптималното m, за което $\delta_m \approx \varepsilon$. Ако има резки скоци и не може да се опредли m, то се избират други базисни функции и т.н.

Подробен пример за МНМК демонстрираме от:

http://fmi.uni-plovdiv.bg/evlm/DBbg/database/numan/leastsquares_BG/index.html

6.5 Приближаване на експериментални данни

Често при провеждане на експерименти се натрупва голямо количество таблици от данни, за която експериментаторът желае да установи аналитичен закон (формула), която възможно най-добре да описва данните. Обикновено по характер тези данни са получени с голяма неотстранима грешка, най-вече зависеща от точността на прибора, методиката на измерване и други фактори.

Да предположим, че е избрана система от базисни функции $f(x) \approx c_0\varphi_0(x) + c_1\varphi_1(x) + \ldots + c_m\varphi_m(x)$ и се търси приближение от вида

$$f(x) \approx c_0\varphi_0(x) + c_1\varphi_1(x) + \ldots + c_m\varphi_m(x). \qquad (17)$$

Този тип приближение е линейна комбинация на избраните m+1 на брой базисни функции $\varphi_j(x), j = 0,1,\ldots m$ и се нарича обобщен полином.

Очевидно (17) е линейна функция спрямо базисните функции. Коефициентите $c_j, j = 0,1,\dots,m$ са неизвестни.

Примери за базисни функции:

1) Най-често това са едночлените $\varphi_j(x) = x^j,\ j = 0,1,\dots m$.

Така приближаващата функция като тяхна линейна комбинация е алгебричен полином $P_m(x)$ от степен m, както разгледахме по-горе

$$f(x) \approx P_m(x) = c_0 + c_1 x + \dots + c_m x^m$$

2) Ако данните наподобяват тригонометрична функция, можем да изберем базисни полиноми $\varphi_j(x) = \sin(jx),\ j = 0,1,\dots m$. Тогава

$$f(x) \approx T_m(x) = c_0 + c_1 \sin(x) + c_2 \sin(2x) + \dots + c_m \sin(mx).$$

3) Друг вариант е експоненциалната апроксимация

$$f(x) \approx E_m(x) = c_0 e^{\lambda_0 x} + c_1 e^{\lambda_1 x} + \dots + c_m e^{\lambda_m x},$$ където λ_j са зададени.

Пример 6.2 На долната таблица са дадени данните от измерване на генерираната изходна мощност Pout при различни стойности на захранващата входна електрическа мощност PIN на медицински лазер.

А) Намерете полином от първа степен, който най-добре приближава тези експериментални данни и оценете грешката. Единствен ли е полиномът?

Б) Намерете полином от втора степен за същите данни и сравнете с резултата от А).

В) Прогнозирайте каква мощност може да очакваме при *PIN*=5.5 KW?

PIN, KW	Pout, W
2	40
2.3	42
2.5	50
2.5	60

2.5	70
3	68
3.3	100
3.5	73
3.5	102
4	96
4	104
4.5	100
4.5	112
5	120
5	120

Решение. Най-напред определяме броя на данните, тук той е N=15. След това от условието определяме, че независимата променлива е x=PIN, а функцията е y=y(x)=$Pout$. Забелязваме освен това, че за едни и същи x са измерени различни y. Това е допустимо в МНМК.

Цялото решение на задачата със система Mathematica демонстрираме директно от сайта:

http://fmi.uni-plovdiv.bg/evlm/DBbg/database/numan/MNMK_CuBr_laser/2leastsquares_CuBr.html

7. Елементи на комбинаториката. Случайни събития, вероятност, събиране и умножение на вероятности. Условна вероятност. Формула на пълната вероятност. Формула на Бейс

7.1 КОМБИНАТОРИКА

Комбинаториката е наука, която изследва множества с краен брой елементи от гледна точка на техни подмножества, групи, наредби и др., получени по определен критерий.

Пермутация от n елемента – това е наредба от всичките n елемента на дадено множество, като се отчита позицията им.
Броят на всички възможни
Пермутации от n елемента: $P_n = n! = n(n-1)(n-2)...2.1$ (7.1)

0!=1 (по дефиниция)

3!=3.2.1=6

5!=5.4.3.2.1=120

10!=10.9.8.7.6.5.4.3.2.1

Mathematica: n!

Пример 7.1 Фармацевтична фирма предлага 4 вида нови лекарства. Дизайнер иска да подреди по една опаковка от всяко лекарство на витрината, като ги поставя едно до друго. По колко различни начина може да стане това?

Решение: Тук n=4. Да означим лекарствата с буквите А, Б, В и Г. Тогава всевъзможните наредби се получават по следния начин: за първи елемент можем да изберем А, Б, В или Г, т.е. общо 4 различни възможности. За втори елемент при избран първи има 3 възможности, след това за трети елемент при избрани първи два – по 2 възможности и за последния елемент – по 1 възможност. Така получаваме всичко P_4=4.3.2.1=24 различни наредби. Те са:

АБВГ, АБГВ, АВБГ, АВГБ, АГБВ, АГВБ,
БАВГ, БАГВ, БВАГ, БВГА, БГАВ, БГВА,
ВАБГ, ВАГБ, ВБАГ, ВБГА, ВГАБ, ВГБА,
ГАБВ, ГАВБ, ГБАВ, ГБВА, ГВАБ, ГВБА.

Вариация от n елемента от k-ти клас без повторение – това е наредба от по k елемента от общо n, като се отчита мястото на елемента и всеки елемент може да участва само веднъж.

Броят на всички вариации без повторение е:

$$V_n^k = \frac{n!}{(n-k)!} = n(n-1)...(n-k+1). \tag{7.2}$$

Пример 7.2 Нека по дадени А, Б, В и Г да образуваме всички възможни наредби от по две букви, без повторение.

Решение: В случая n=4, k=2 и имаме случай на вариации от 4 елемента от 2-ри клас без повторение. Процедираме както в Пример 7.1, но спираме след избора на втората буква. Получаваме общо V_4^2 =4.3=12 различни наредби:

АБ, АВ, АГ, БА, БВ, БГ, ВА, ВБ, ВГ, ГА, ГБ, ГВ.

<u>Вариация от n елемента от k-ти клас с повторение</u> е вариация, в която всеки избор се извършва от цялото множество елементи. Броят на всички възможни вариации с повторение е:

$$\tilde{V}_n^k = n^k. \qquad (7.3)$$

Пример 7.3 Да се напишат вариациите с повторение за пример 7.2.
Решение: Сега всеки избран символ може да участва отново. Ще получим общ брой вариации с повторение: $\tilde{V}_4^2$ =4.4=4^2=16:
АБ, АВ, АГ, БА, БВ, БГ, ВА, ВБ, ВГ, ГА, ГБ, ГВ, АА, ББ, ВВ, ГГ.

Пример 7.4 Колко различни номера в регистрационните табелки на автомобилите могат да се изпишат, ако номерът е четирицифрен?
Решение: Броят на цифрите е десет: 0,1,2, ..., 9, т.е. n=10. От тях трябва да избираме по 4, т.е. k=4, като повторението на цифрите се допуска. От формула (7.3): $\tilde{V}_{10}^4$ =10^4=10000 различни номера.

<u>Комбинация от n елемента от k-ти клас без повторение</u> – така наричаме вариация без повторение, в която мястото на елемента не се отчита.

Броят на всички комбинации от n елемента от k-ти клас е:

$$C_n^k = \frac{n!}{k!(n-k)!} = \frac{n(n-1)...(n-k+1)}{k!}. \tag{7.4}$$

Пример 7.5 От група от десет деца трябва да се изберат три за участие в новогодишна томбола. По колко различни начина може да се сформира тройката деца?

Решение: Множеството от децата е $A = \{a_1, a_2, ..., a_{10}\}$ с n=10 елемента. От него избираме групи по k=3 деца, като в групите няма наредба между децата. Така получаваме комбинации без повторение, които по формула (7.4) са на брой: $C_{10}^3 = \frac{10.9.8}{1.2.3} = 120.$

Комбинация от n елемента от k-ти клас с повторение – така наричаме комбинация, в която елементите могат да се повтарят. Броят на всички комбинации от *n* елемента от *k*-ти клас е:

$$\tilde{C}_n^k = C_{n+k-1}^k. \tag{7.5}$$

Пример 7.6 За група от 10 деца се провежда томбола, в която се раздават 3 награди. Всяко дете може да получи произволен брой награди. Колко възможни разпределения на наградите са възможни?

Решение: Имаме n=10, k=3. Тъй като няма ограничение по колко награди може да получи едно дете, то става въпрос за комбинации с повторение, чийто брой е: $\tilde{C}_{10}^3 = C_{10+3-1}^3 = C_{12}^3 = \dfrac{12.11.10}{1.2.3} = 220.$

<u>Принцип на събирането</u>

Нека са дадени две множества A и B, съответно с n и m елемента, които нямат общи елементи. Тогава броят на всички възможности за избор на един елемент независимо от кое от тези множества е n+m.

Пример 7.7 Студентите от първи курс специалност фармация са общо 60, от втори курс – 45, от трети – 41, а от четвърти курс - 50. От трите курса трябва да се излъчи един студент за отговорник на специалността. По колко различни начина може да се избере отговорник?

Решение: Тук имаме четири множества и изборът трябва да е от обединението на всичките множества. Общият брой студенти е:
$n = 60+45+41+50=196$ студенти, от които се избира един. Това може да стане по 196 начина.

Принцип на умножението

Нека са дадени две множества A и B, съответно с n и m елемента, които нямат общи елементи. Тогава броят на всички възможни избора от по 2 елемента, от които първият е от A, а вторият от B е равен на n.m. Полученото множество от наредени двойки елементи се нарича <u>декартово произведение на A и B</u> и се бележи с $A\times B$**.**

Пример 7.8 В един ресторант се предлагат 3 вида супи, 8 вида основни ястия, 4 вида салати и 5 вида десерти. Колко различни на брой менюта могат са се сервират, състоящи се от супа, ястие, салата и десерт?

Решение: Броят на супите е n_1=3, на основните ястия - n_2=8, на салатите - n_3=4 и на десертите - n_4=5. За всеки вид супа от общо 3 могат да се предложат по 8 основни ястия, т.е. 3.8=24 възможности от по 2 ястия. За всяко от тях по 4 салати – 24.4=96 и по 5 десерта – общо $n_1n_2n_3n_4$=3.8.4.5=480 различни менюта.

Пример 7.9 В партида от 100 изделия има 20 дефектни. По колко различни начина могат да се извадят 10 изделия, точно 3 от които са дефектни.

Решение: Дадено е, че дефектните изделия са 20, тогава недефектните са 80. Всички възможни избори на 3 дефектни от 20 са $n = C_{20}^{3}$, останалите 10-3=7 недефектни се избират от 80 по $m = C_{80}^{7}$ начина. На всяка тройка дефектни съответстват 7 недефектни. Следователно всички избори са $n.m = C_{20}^{3}.C_{80}^{7}$.

7.2. СЛУЧАЙНО СЪБИТИЕ

Теорията на вероятностите е наука, която изследва закономерностите на случайните явления.

Под ***експеримент*** или ***опит*** се разбира изпълнението на определен комплекс от условия, при които настъпва едно или друго явление и се регистрира някакъв факт. Условията на опита трябва да предоставят: а) възможност (реална или абстрактна) за неограничен брой повторения на комплекса условия, б) предварително да е известно множеството на възможните резултати (изходи) от провеждания опит, в) при всеки опит да може да се реализира един и само един от всичките възможни изходи, който се нарича ***елементарен изход***.

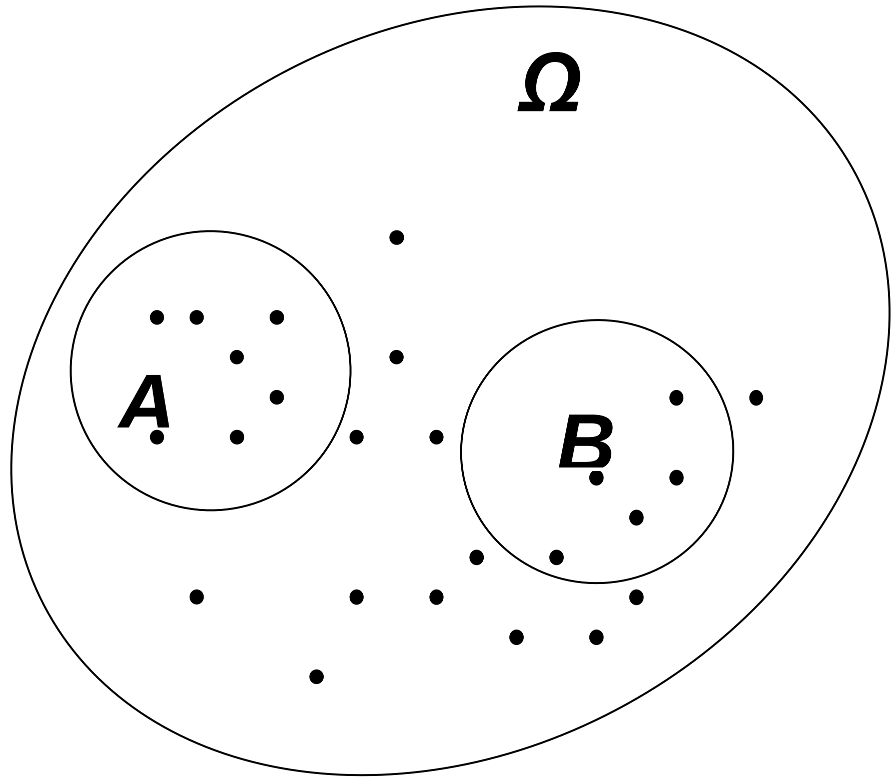
Пространство на елементарните изходи Ω се нарича съвкупността от всички елементарни изходи на даден опит. Всеки елементарен изход от опита се бележи с ω, т.е. $\omega \in \Omega$. Броят на елементарните изходи може да е краен, безкраен, или да принадлежи на някакъв интервал.

Пример 7.10 Опитът се състои в еднократно подхвърляне на симетричен зар. Да се определи пространството на елементарните изходи Ω.
Решение: Възможните елементарни изходи са да се паднат 1, 2, 3, 4, 5 или 6 точки.

Ако означим с $\omega_k = k$ паднали се точки, то
$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$.

Пример 7.11 Подхвърля се многократно симетрична монета, докато се падне герб за първи път. Да се определи Ω.
Решение: Да означим с *Г* падането на герб, а с *Л* – падането на лице. Нека означим първото падане на *Г* при *k*-то хвърляне след k=1 лица *Л* с ω_k. Тогава получаваме безкрайното пространство
Ω={ *Г, ЛГ, ЛЛГ, ЛЛЛГ, ..., ЛЛ...ЛГ, ...*}=
$=\{\omega_1, \omega_2, \omega_3, \omega_4, \ldots, \omega_k, \ldots\}$.

Случайно събитие е всяко подмножество на Ω. Ще го ще означаваме с големи латински букви: *A*, *B*, *C*, ... с или без индекс.
Събитието е произволен факт, който може или не може да се осъществи като резултат от дадения опит. Казваме, че събитието *A* настъпва, когато се е осъществил елементарен изход от подмножеството *A*. Такъв елементарен изход се нарича ***благоприятен***.



Фиг.7.1. Съвкупността Ω от всички елементарни изходи ω_k (изобразени като точки). Сл.с. A и B са подмножества на Ω.

Пример 7.12 Хвърля се еднократно симетричен зар, т.е. пространството от всички събития е $\Omega=\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$, където ω_k е елементарното събитие: "падат се k точки". Случайното събитие "падат се нечетен брой точки" е събитието $A=\{\omega_1, \omega_3, \omega_5\}$, което очевидно е подмножество на Ω. Благоприятни за A са събитията ω_1 или ω_3 или ω_5, когато се паднат. Друг пример е събитието B "падат с повече от 4 точки", т.е. $B=\{\omega_5, \omega_6\}$, с благоприятни изходи ω_5, ω_6.

<u>Допустимо събитие</u> е това, което непременно настъпва при всеки опит.
<u>Невъзможно събитие</u> е, такова, което никога не настъпва.
Например събитието $C=\Omega$ е достоверно, а събитието $D=\varnothing$ е невъзможно.

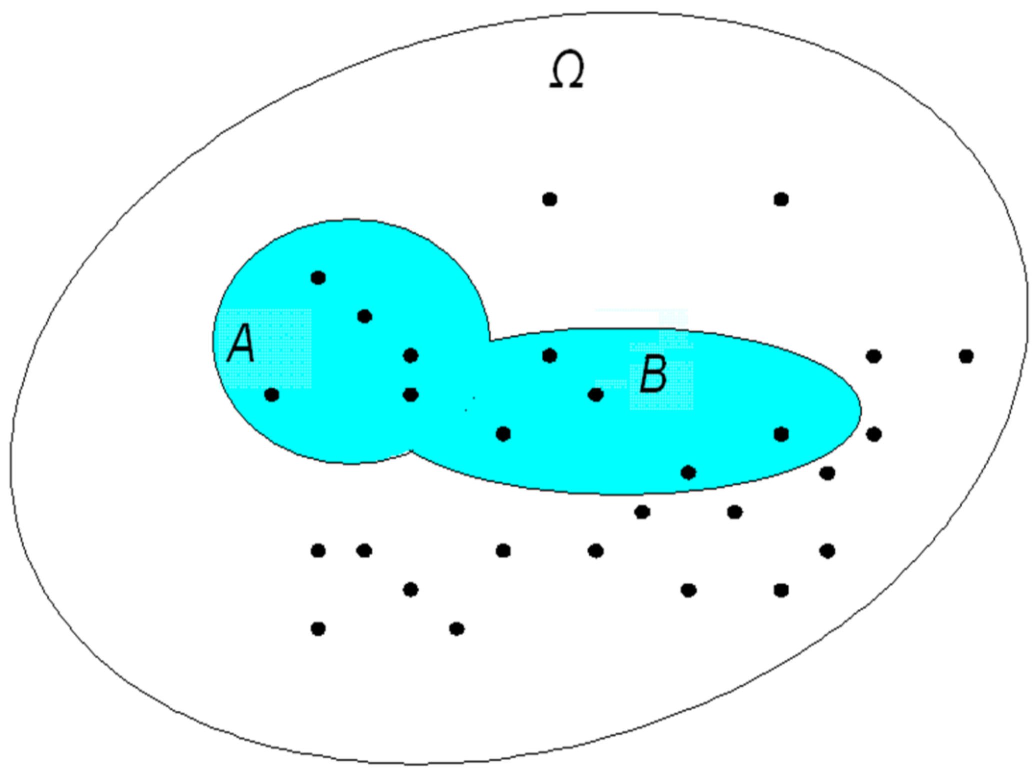
Различават се ***съставни*** (разложими) и ***елементарни*** (неразложими) събития.
От пример 7.3 събитията A и B са съставни, защото имат съответно три и два елементарни изхода.

Противоположно събитие на A е събитието, $\overline{A} = \Omega \backslash A$ (“не A”).
Пример 7.13 Хвърля се симетричен зар. Ако е дадено събитието $A=\{\omega_1, \omega_3\}$, то $\overline{A} =\{\omega_2, \omega_4, \omega_5, \omega_6\}$ е неговото противоположно събитие.

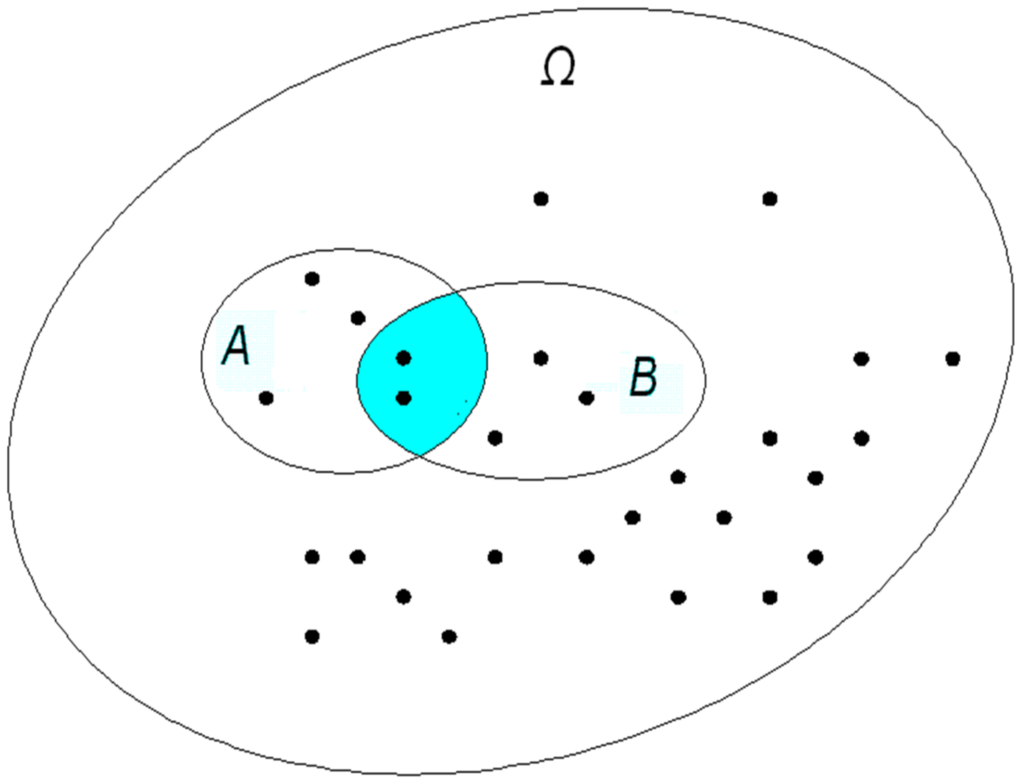
Съотношения между събития

Нека Ω е пространството на елементарните събития на даден опит, а A, B, C, ... са случайни събития, т.е. подмножества на Ω. Тогава съотношенията между събитията се описват като съотношения между множества. Привеждаме основните операции и съотношения между множества:

1) $A \subset B$ означава, че A е подмножество на B, или събитието A влече събитието B.
2) $A \bigcup B$ е обединение на A и B. Това събитие настъпва ако настъпва поне едно от събитията A или B.
3) $A \bigcap B$ е сечение на A и B, т.е. това събитие настъпва, когато и двете събития настъпват. Понякога сечението се означава като AB.
4) $\overline{A} = \Omega \backslash A$ е допълнение на A. Означава, че не настъпва A. Очевидно $A \bigcup \overline{A} = \Omega$.
5) Ако $A \subset B$ и $B \subset A$, то казваме, че A е еквивалентно на B и бележим с $A=B$.
6) $C=A \backslash B$ се нарича разлика на A и B. Това е събитие, състоящо се от сбъдване на А и несбъдване на В, т.е. $C = A \bigcap \overline{B}$.



Фиг. 7.2а. Обединението $A \cup B$ е съвкупността всички елементарни изходи *на* A и B, взети заедно.



Фиг. 7.2б. Сечението $A \cap B$ (потъмнената средна част) е подмножеството от общите елементарни изходи на A и B.

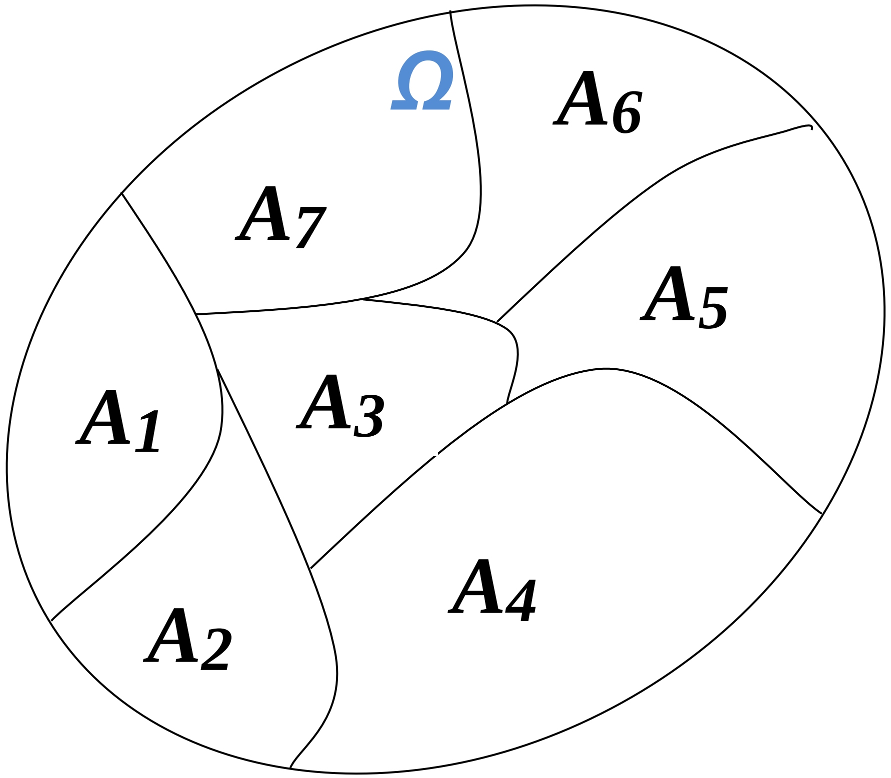
Две събития A и B се наричат ***несъвместими***, когато $A \bigcap B = \varnothing$, т.е. съвместната им реализация е невъзможна.

Пример 7.14 Събитията A и $\overline{A}$ са несъвместими, защото нямат общи елементи, т.е. $A \bigcap \overline{A} = \varnothing$.

Пример 7.15 Хвърля се симетричен зар. Събитието A: "пада се нечетен брой точки" и събитието B: "пада се шестица" са несъвместими. В случая $A=\{\omega_1, \omega_3, \omega_5\}$ и $B=\{\omega_6\}$.

Пълна група от събития. Казваме, че събитията $A_1, A_2, ..., A_p$ образуват пълна група от събития, ако те са две по две несъвместими и изчерпват всички възможни изходи на опита, т.е. ако

$A_i \bigcap A_j = \varnothing$ при $i \neq j$ и $\bigcup_{i=1}^{p} A_i = \Omega$ (Виж пример на Фиг. 7.3).



Фиг.7.3. Събитията $A_1, A_2, \dots, A_7$ образуват пълна група събития.

Закони за съотношения между събития

Нека A, B и C са произволни събития. В сила са следните закони:

1) $A \bigcup B = B \bigcup A,\ \ A \bigcap B = B \bigcap A$ - комутативност

2) $A \bigcup (B \bigcup C) = (A \bigcup B) \bigcup C, \quad A \bigcap (B \bigcap C) = (A \bigcap B) \bigcap C$ - асоциативност

3) $A \bigcap (B \bigcup C) = (A \bigcap B) \bigcup (A \bigcap C) \quad A \bigcup (B \bigcap C) = (A \bigcup B) \bigcap (A \bigcup C)$- дистрибутивност

4) $\overline{A \bigcup B} = \bar{A} \bigcap \bar{B}, \quad \overline{A \bigcap B} = \bar{A} \bigcup \bar{B}$ - закони на де Морган.

Пример 7.16 Опитът се състои с проверка за доброкачественост на тройки изделия. Нека A е събитието “поне едно от трите проверявани изделия е недоброкачествено”, а B е събитието “и трите изделия са доброкачествени”. Какво означават събитията: а) $A\bigcup B$; б) $A\bigcap B$; в) $\bar{B}$; г) $\bar{B}$?

Решение: Имаме два вида изделия – “недоброкачествени” и “доброкачествени”, които ще означим съответно с {Н, Д}. От тях избираме всевъзможните наредени тройки с повторение, които са общо $\tilde{V}_2^3=2^3=8$ на брой. По-точно получаваме

Ω={НДД, ННД, НДН, ННН, ДНН, ДДН, ДНД, ДДД}.

Събитието A={НДД, ННД, НДН, ННН, ДНН, ДДН, ДНД},

а събитието B={ДДД}.

Тогава: а) $A\bigcup B=\Omega$; б) $A\bigcap B=\varnothing$; в) $\bar{B}=B$; г) $\bar{B}=A$.

7.3. ВЕРОЯТНОСТ НА СЛУЧАЙНИ СЪБИТИЯ

Вероятност на случайно събитие е число, изразяващо обективната възможност за настъпването му. Вероятността на събитието A се бележи с $P(A)$.

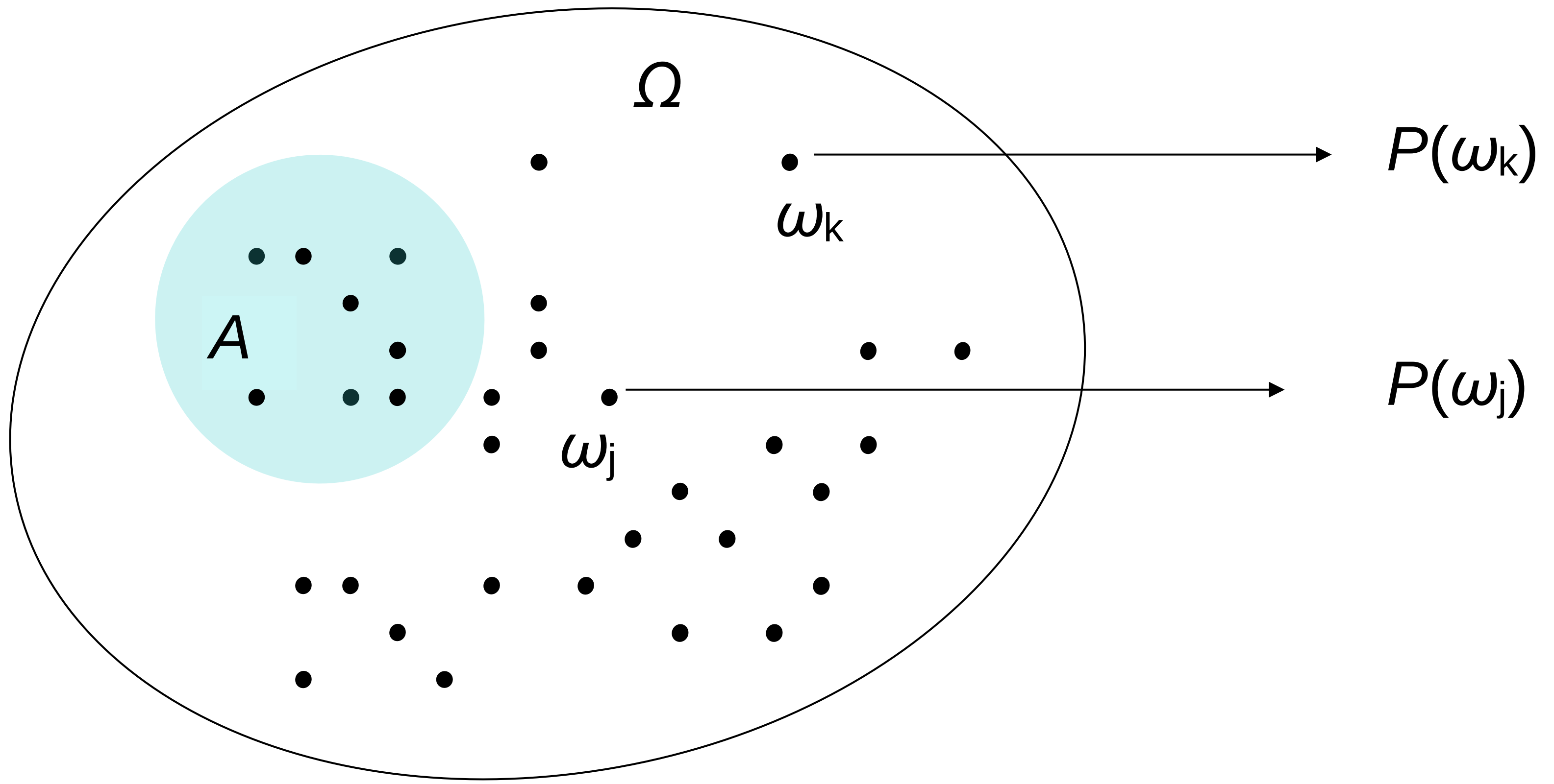
<u>Определение за вероятност</u> Нека $\Omega=\{\omega_1, \omega_2, \dots, \omega_k, \dots\}$ е пространството от всички (краен или безкраен брой) елементарни изходи от даден опит и на всеки елементарен изход ω_k, ($k=1,2,\dots$) е съпоставено число $P(\omega_k)$, така че са изпълнени условията:

$$P(\omega_k) \geq 0 \quad \text{- неотрицателност,} \tag{7.6}$$

$$P(\omega_1) + P(\omega_2) + \dots + P(\omega_k) + \dots = 1 \quad \text{- нормираност.} \tag{7.7}$$

Тогава ***<u>вероятност на случайното събитие A</u>*** е сумата на вероятностите на елементарни изходи от Ω, които са благоприятни за A, т.е.

$$P(A) = \sum_{\omega_k \in A} P(\omega_k). \tag{7.8}$$



Фиг.7.4. На всеки елементарен изход ω_k се съпоставя реално число $P(\omega_k)$, наречено негова вероятност, така че да са изпълнени условията (7.6)-(7.7). **<u>Вероятност на събитието *A*</u> е сумата от вероятностите на неговите елементарни изходи.**

Вероятностен модел (Ω,P) на даден опит е зададен, ако за него са известни:

а) пространството на елементарните му изходи $\Omega=\{\omega_1, \dots, \omega_k, \dots\}$,

б) вероятностите $P(\omega_k)$, $k=1,2,\dots$, за които са изпълнени условията (7.6) - (7.7).

Основни свойства на вероятността:

$$0 \le P(A) \le 1 \text{ за всяко случайно събитие } A, \tag{7.9}$$

$$P(\Omega) = 1, \tag{7.10}$$

Ако A и B са несъвместими, т.е. $A \cap B = \varnothing$, то

$$P(A \cup B) = P(A) + P(B). \tag{7.11}$$

Като следствия от тези свойства веднага се доказват следните равенства:

$$P(A) = 1 - P(\bar{A}), \tag{7.12}$$

$$P(\varnothing) = 0. \tag{7.13}$$

Пример 7.17 Хвърля се несиметричен зар с вероятностен модел:

$$\begin{pmatrix} \Omega \\ P \end{pmatrix} = \begin{pmatrix} \omega_1 & \omega_2 & \omega_3 & \omega_4 & \omega_5 & \omega_6 \\ \frac{1}{8} & \frac{1}{4} & \frac{1}{6} & \frac{1}{8} & \frac{1}{4} & \frac{1}{12} \end{pmatrix},$$

където ω_k означава, че са се паднали k точки.

Намерете: а) вероятността да се паднат 6 точки; б) вероятността да се падне четен брой точки; в) вероятността да не се падне 1.

Решение: а) $P(\omega_6) = \frac{1}{12}$; б) $A = \{\omega_2, \omega_4, \omega_6\}$, $P(A) = P(\omega_2) + P(\omega_4) + P(\omega_6)$ $= \frac{1}{4} + \frac{1}{8} + \frac{1}{12} = \frac{11}{24}$; в) $B = \{ \omega_1\}$, $P(B) = 1 - P(\overline{B}) = 1 - \frac{1}{8} = \frac{7}{8}$.

7.4. КЛАСИЧЕСКА ВЕРОЯТНОСТ

Определение

$$p(A) = \frac{m}{n} = \frac{\text{брой благоприятни изходи за } A}{\text{общ брой възможни изходи от опита}}$$

$0 \leq p(A) \leq 1$ за всяко случайно събитие A

$p(\Omega) = 1, \qquad p(\emptyset) = 0$

$p(\overline{A}) = 1 - p(A).$

Пример 7.18 Хвърлят се едновременно два зара. Коя вероятност е по-голяма: да се падне сума на точките равна на 7 или да се падне сума от 8 точки?

<u>Решение:</u> Тук $\Omega=\{\omega_k=(i, j)\}$, където i са точките от първи зар, а j – от втори зар. При всеки опит се пресмята сумата $i+j$. Всеки зар може да има точки от 1 до 6, независимо от точките на другия зар. Тогава всевъзможните елементарни изходи са $n=\tilde{V}_6^2=6^2=36$. Означаваме с A събитието "да се падне сума от 7 точки", а с В – събитието "да се падне сума от 8 точки". Броят на благоприятните изходи на A са $m_1=6$: A={(1,6), (6,1), (2,5), (5,2), (3,4), (4,3)}. Броят на благоприятните изходи на B са $m_2=5$: B={(2,6), (6,2), (3,5), (5,3), (4,4)}. Тогава $P(A)=\dfrac{m_1}{n}=\dfrac{6}{36}$, а $P(B)=\dfrac{m_2}{n}=\dfrac{5}{36}$. Получихме, че вероятността да се падне сбор 7 е по-голяма от вероятността да се падне сбор 8.

Пример 7.19 В кутия има 6 билета от лотария, от които 4 печелят. Да се намери вероятността при изваждане на 2 билета по случаен начин: а) и двата билета да печелят; б) единият да печели, а другият – не.

Решение: а) Броят на всички възможни изходи е $n = C_6^2 = \dfrac{6.5}{1.2} = 15.$

Изваждането на 2 печеливши билета от 4 възможни може да стане по $m = C_4^2 = \dfrac{4.3}{1.2} = 6$ начина. Търсената вероятност е $p = \dfrac{C_4^2}{C_6^2} = \dfrac{6}{15} = 0.4.$

б) Едновременно изваждане на печеливш и непечеливш билет може да стане по $m_1.m_2 = C_4^1.C_2^1 = 4.2 = 8$ различни начина. Сега вероятността е

$$p = \frac{C_4^1.C_2^1}{C_6^2} = \frac{8}{15} \approx 0.533.$$

Този тип вероятност се нарича ***хипергеометрична*** и е специален вид класическа вероятност.

7.5. СТАТИСТИЧЕСКА ВЕРОЯТНОСТ

В практиката можем да измерваме **относителната честота на събитие**, т.е. колко пъти се е осъществило дадено събитие спрямо броя фактически проведените опити. Това е опитна експериментална характеристика, където m е броят на осъществяване на събитие A, а n е общият брой опити.

Вероятността в този случай се определя като границата $P(A)=\lim_{n\to\infty}\frac{m}{n}$. В общия случай взимаме приблизително $P(A)\approx\frac{m}{n}$, когато n е достатъчно голямо.

Пример 7.20 В автопарк е установено, че от постъпили 1560 автомобила за 1 година са се продали 451. Да се намери честотата (вероятността) на събитието A=”Продава се една лека кола”.

Решение: $P(A)\approx 451/1560\approx 0.289$.

7.6. ПРЕСМЯТАНЕ НА ВЕРОЯТНОСТИ

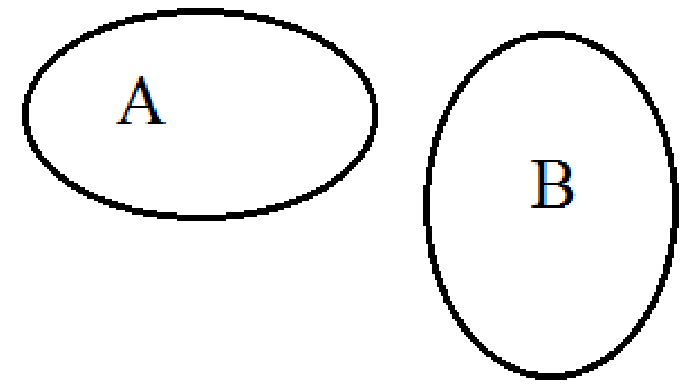
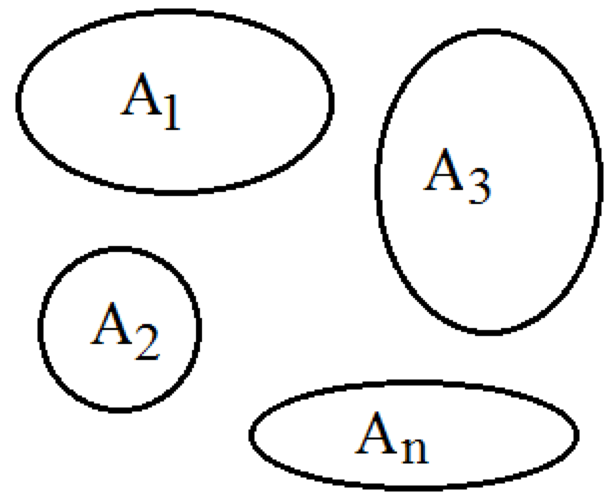
СЪБИРАНЕ НА ВЕРОЯТНОСТИ

Сума на две събития е събитието, при което се сбъдва поне едно от тези събития (*A* или *B*).

Ако събитията A и B са ***несъвместими***, т.е. $A \bigcap B = \emptyset$, то $A + B$ означава или настъпване на *A* или настъпване на *B*.

Когато събитията са ***съвместими***, т.е. $A \bigcap B \neq \emptyset$, то $A + B$ означава или настъпване на *A* или настъпване на *B* или настъпване и на двете.

Теорема за събиране на вероятности на събития:

а) Ако събитията A и B са несъвместими, т.е. $A \bigcap B = \emptyset$, то $p(A+B) = p(A) + p(B)$, защото $p(AB) = 0$	
б) При повече от две по две несъвместими събития $A_1, A_2, ..., A_n$: $p(\sum_{k=1}^{n} A_k) = \sum_{k=1}^{n} p(A_k)$	

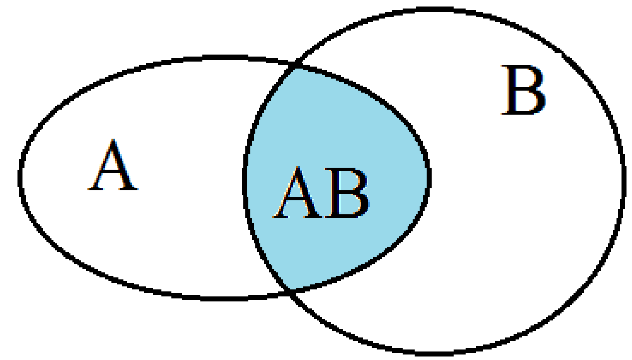
Пример 7.21 Студентите от първи курс на специалност БИТ са разпределени в три групи съответно: първа група 32 студента, втора група 30 и трета група 35 студенти. Избира се случайно един студент от курса за групов отговорник. Да се намери вероятността той да е от първа или втора група.

Решение: Общият брой студенти е 32+30+35=97. Означаваме с A събитието “студентът е от първа група”, с B – събитието “студентът е от втора група”. Очевидно $P(A) = \dfrac{32}{97}$, а $P(B) = \dfrac{30}{97}$. Тъй като един и същ студент не може да бъде едновременно и в двете групи, събитията A и B са несъвместими. Тогава търсената вероятност по теоремата за събиране на вероятности е $P(A+B) = \dfrac{32}{97} + \dfrac{30}{97} = \dfrac{62}{97} \approx 0.639$.

Теорема за събиране

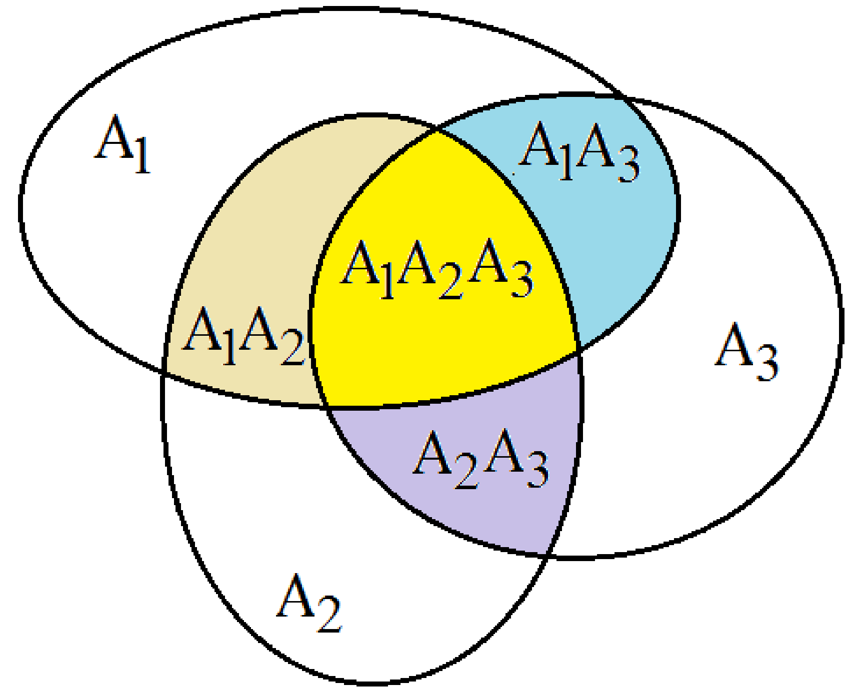
в) при две произволни събития

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



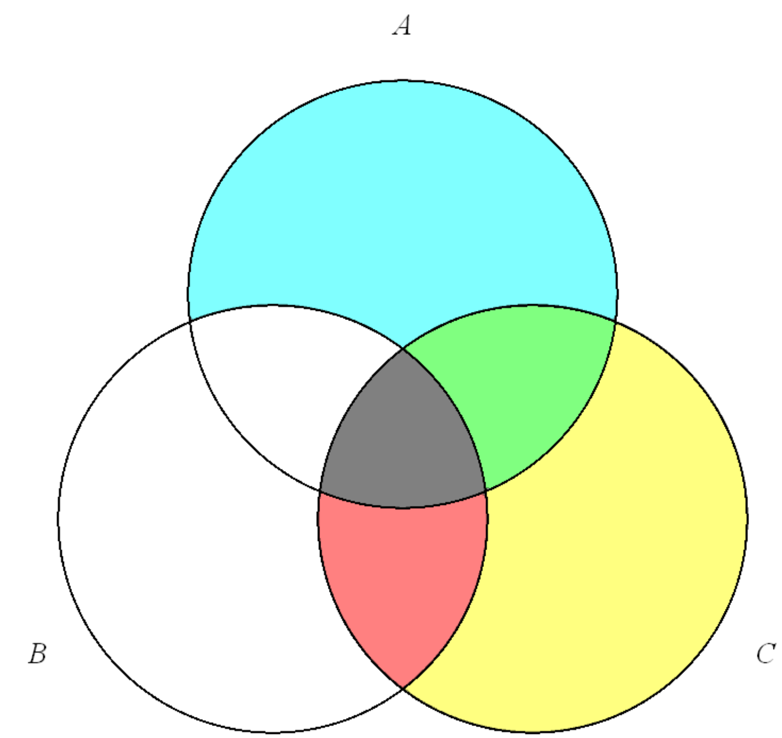
г) при три произволни събития:

$$p(A_1+A_2+A_3)=p(A_1)+p(A_2)+p(A_3)$$
$$-p(A_1A_2)-p(A_1A_3)-p(A_2A_3)$$
$$+p(A_1A_2A_3)$$



За представянето на всевъзможните 127 непразни множества и сечения, получени при пресичане на 3 множества може да разгледате и демонстрацията на т.н.
диаграма на Вен (Venn diagram)

http://demonstrations.wolfram.com/VennDiagrams/



д) При повече събития $A_1, A_2, ..., A_n$:

$$p(\sum_{k=1}^{n} A_k) = \sum_{k=1}^{n} p(A_k) - \sum_{k<l} p(A_k A_l) + \sum_{k<l<s} p(A_k A_l A_s) + ... + (-1)^{n-1} p(A_1 A_2 ... A_n)$$

в) частен случай: При $A_1, A_2, ..., A_n$, които са независими в съвкупност:

$$p(\sum_{k=1}^{n} A_k) = \sum_{k=1}^{n} p(A_k) - \sum_{k<l} p(A_k) p(A_l) + \sum_{k<l<s} p(A_k) p(A_l) p(A_s) - ...$$

$$+ (-1)^{n-1} p(A_1) p(A_2) ... p(A_n)$$

Пример 7.22 В голяма група студенти 25% говорят испански, 50% говорят английски, 20% говорят и двата езика. Да се намери вероятността произволно избран студент да говори поне един от двата езика.

Решение: Събитието A= "студентът говори испански" има вероятност $P(A)$=0.25. Събитието B= "студентът говори английски" има вероятност

$P(B)$=0.5. Събитието $A.B$= “студентът говори и двата езика” има вероятност $P(\mathrm{A}B)$=0.2. По теоремата за събиране на вероятности:

$$p(A+B)=p(A)+p(B)-p(AB)=0.25+0.5-0.2=0.55 \, .$$

7.7. Условна вероятност

Условна вероятност:

$$p(A/B)=\frac{p(AB)}{p(B)}, \quad p(B)\neq 0.$$

Това е вероятността да настъпи събитието *A*, при условие, че е настъпило събитието *B*.

Произведение на събития – това е събитие, при което тези събития настъпват едновременно (напр. *A* и *B*).

Теореми за умножение на вероятности:

а) Ако $p(A)>0, \ p(B)>0$, то $p(AB)=p(A).p(B/A)=p(B).p(A/B)$

б) При повече събития:

$$p(A_1A_2\dots A_n)=p(A_1)p(A_2/A_1)p(A_3/A_1A_2)\dots\ p(A_n/A_1A_2\dots A_{n-1})$$

<u>Независими събития</u>, ако $p(AB) = p(A)\,p(B)$. При независими събития $p(A/B) = p(A)$. Това означава, че сбъдването на едното събитие не променя вероятността на сбъдване на другото събитие.

Пример 7.23 Хвърлят се последователно 2 монети. Означаваме събитията: *A*=”Пада се герб на първата монета”; *B*=”Пада се поне един герб”; *C*=”Пада се поне едно лице”; *D*=”Пада се герб на втората монета”. Зависими ли са събитията: а) *A* и *D*; б) *B* и *C*?

Решение: *n*=4, Ω={(ГГ), (ГЛ), (ЛГ), (ЛЛ)}, *A*={(ГГ), (ГЛ }, *B*={(ГГ), (ГЛ), (ЛГ)}, *C*={(ГЛ), (ЛГ), (ЛЛ)} и *D*={(ГГ), (ЛГ)}.

а) $A \cap D$={(ГГ)}. Имаме: $p(A \cap D) = \frac{1}{4}$, $p(A) = \frac{2}{4} = \frac{1}{2}$, $p(D) = \frac{2}{4} = \frac{1}{2}$,

$$p(A/D) = \frac{p(AD)}{p(D)} = \frac{1/4}{1/2} = \frac{1}{2} = p(A)$$

. Следва, че събитията *A* и *D* са независими.

б) $B\cap C$={(ГЛ), (ЛГ)}. Имаме: $p(B\cap C)=\frac{2}{4}=\frac{1}{2}$, $p(B)=\frac{3}{4}$, $p(C)=\frac{3}{4}$, $p(B/C)=\frac{p(BC)}{p(C)}=\frac{1/2}{3/4}=\frac{2}{3}\neq p(B)$. Следва, че събитията B и C са зависими.

Пример 7.24 В кутия има 6 бели и 9 черни топки. Последователно без връщане се изваждат 2 топки. Да се намери вероятността първата топка да е бяла, а втората да е черна.

Решение: Означаваме с A събитието "първата топка е бяла", с B - събитието "втората топка е черна". Вероятността на едновременното им сбъдване е вероятността на произведението $p(AB)$. Изчисляваме $p(A)=\frac{6}{15}=\frac{2}{5}$. След сбъдване на A, в кутията остават 5 бели и 9 черни топки. Сега вероятността за B е $p(B/A)=\dfrac{9}{14}$. По теоремата за умножение на вероятностите $p(AB)=p(A)P(B/A)=\dfrac{2}{5}.\dfrac{9}{14}=\dfrac{9}{35}\approx 0.257$.

Пример 7.25 Студент е научил 20 от 25 въпроса за изпит. Каква е вероятността да отговори на 2 случайно зададени му въпроса от конспекта?

Решение 1: Означаваме A=”Отговорил на първи въпрос”, B=”Отговорил на втория въпрос”. $p(AB)=p(A)\,p(B/A)=\dfrac{20}{25}.\dfrac{19}{24}\approx 0.633$.

Решение 2: С преброяване на благоприятните към всички комбинации имаме:

$$p(AB) = \frac{C_{20}^2}{C_{25}^2} \approx 0.633$$
.

Пример 7.26 Вероятността в ракова клетка мишена да попадне радионуклеид е 0.7, а да попадне втори вид радионуклеид – е 0.8. Каква е вероятността в клетката мишена да попадне поне един от радионуклеидите, ако едновременно се приложат двата препарата?

<u>Решение:</u> Двете събития могат да настъпят независимо едно от друго. По теоремата за събиране на вероятности имаме:

$$p(A_1 + A_2) = p(A_1) + p(A_2) - p(A_1A_2) = 0.7 + 0.8 - 0.7 * 0.8 = 0.94$$
.

Пример 7.27 Към пример 3.10 нека има и трети радионуклеид, който попада в клетката мишена с вероятност 0.6. Намерете вероятността в клетката да попадне поне един от радионуклеидите, ако едновременно се приложат трите препарата.

Решение: Трите събития могат да настъпят независимо едно от друго. По теоремата за събиране на вероятности в общия случай имаме:

$$p\left(\bigcup_{k=1}^{3} A_k\right) = \sum_{k=1}^{3} p(A_k) - \sum_{k<l} p(A_k A_l) + \sum_{k<l<s} p(A_k A_l A_s) =$$

$$0.7 + 0.8 + 0.6 - 0.7 * 0.8 - 0.8 * 0.6 - 0.7 * 0.6 + 0.7 * 0.8 * 0.6 = 0.976 .$$

7.8. ПЪЛНА ВЕРОЯТНОСТ И ФОРМУЛА НА БЕЙС

Нека се търси вероятността на събитие A, зависеща от сбъдването на краен брой събития (хипотези) $H_1, H_2, ..., H_n$, от които може да се сбъдне само едно, т.е. те образуват пълна група несъвместими събития:

$\bigcup_{k=1}^{n} H_k = \Omega$ и $H_i \cap H_j = \varnothing$ за $i \neq j$.

Оттук следва, че $\sum_{k=1}^{n} p(H_k) = 1$.

Формула за пълната вероятност:

Ако вероятностите $p(H_k)$ и условните вероятности $p(A/H_k)$, $k=1,2,...,n$ са предварително известни, то

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

Пример 7.28 В три еднакви партиди изделия са открити съответно: 20%, 15% и 5% дефектни изделия. Намерете вероятността за откриване на дефектно изделие при случайно изваждане на изделие от произволна партида.

Решение: Означаваме *A*=”извадeното изделие е дефектно”. Хипотезите са: H_1 =”Изделието е взето от първа партида”, H_2 =”Изделието е взето от втора партида”, H_3 =”Изделието е взето от трета партида”.

Вероятностите на хипотезите са еднакви: $p(H_1) = p(H_2) = p(H_3) = \frac{1}{3}$.

Получаваме

$$p(A) = \sum_{k=1}^{3} p(H_k)\, p(A/H_k) = \frac{1}{3}.\frac{20}{100} + \frac{1}{3}.\frac{15}{100} + \frac{1}{3}.\frac{5}{100} = 0.0667 + 0.05 + 0.01667$$

$$\approx 0.1333$$

За изчисляване на пълната вероятност често се съставя таблица:

k	$p(H_k)$	$p(A/H_k)$	Произведения	Стойност
1	1/3	0,2	(1/3)*0,20 =	0,0667
2	1/3	0,15	(1/3)*0,15 =	0,0500
3	1/3	0,05	(1/3)*0,05 =	0,0167
			Сума =	0,1333

Формула на Бейс:

При условие, че се е сбъднало събитието A, условните вероятности на хипотезите $H_k,\ k=1,2,\dots,n$ могат да се намерят по следната формула:

$$p(H_k/A)=\frac{p(A/H_k)p(H_k)}{\sum\limits_{j=1}^{n}p(A/H_j)p(H_j)}. \tag{7.14}$$

Пример 7.29 При данните от пример 3.12 след като е извадено дефектно изделие, каква е вероятността то да е било взето съответно от първа, втора или трета партида?

Решение: По формулата на Бейс знаменателят е равен на пълната вероятност $p(A)\approx 0.133$. Последователно получаваме:

Вероятността дефектното изделие да е взето от първа партида е:

$$p(H_1 / A) = \frac{p(A/H_1)\,p(H_1)}{\sum_{j=1}^{3} p(A/H_j)\,p(H_j)} = \frac{0.2 * \frac{1}{3}}{0.133} \approx 0.5$$

.

Вероятността дефектното изделие да е взето от втора партида е:

$$p(H_2 / A) = \frac{p(A/H_2)\,p(H_2)}{\sum_{j=1}^{3} p(A/H_j)\,p(H_j)} = \frac{0.15 * \frac{1}{3}}{0.133} \approx 0.375$$

.

Вероятността дефектното изделие да е взето от трета партида е:

$$p(H_3 / A) = \frac{p(A/H_3)\,p(H_3)}{\sum_{j=1}^{3} p(A/H_j)\,p(H_j)} = \frac{0.05 * \frac{1}{3}}{0.133} \approx 0.125$$

8. Случайни величини и разпределения. Функция на разпределение

Определение 8.1. **Случайната величина** (или случайна променлива) X в зависимост от изхода на даден опит взима само една стойност от своето дефиниционно множество D_X, с някаква вероятност P. Напр. ако X приема стойност x_1 с вероятност 0.2, то записваме $P(X=x_1)=0{,}2$. Четем "Вероятността X да е равно на x_1 е 0,2"

Сл.в. X се нарича **дискретна сл.в.**, ако приема по случаен начин стойности от D_X, което се състои от краен или безкраен брой стойности, такива, че могат да се номерират (изброимо множество).

Сл.в. X се нарича **непрекъсната сл.в.**, ако приема по случаен начин стойности от Dx , което е краен или безкраен интервал от числовата ос.
Сл.в. се означават с големи латински букви X, Y, Z ...или гръцките букви ξ, η, ..., а техните стойности – с малки латински букви x, y, z…

Закон за разпределение на сл.в. – това е описание на връзката между всяка възможна стойност на сл.в. и съответната й вероятност. При $x \notin D_X$ вероятността в общия случай се счита за 0.

8.1. Дискретни сл.в. и разпределения

За дискретните сл.в. законът на разпределение е обикновено таблица, формула или графика (полигон, начупена линия).

Ето примерна таблица на разпределение на дискретната сл.в. ξ (четем „кси”) с крайно дефиниционно множество $D_X = \{ x_1, x_2, \ldots, x_n \}$ и съответните вероятности. Тъй като ξ може да приема всички стойности от D_X, то сумата от вероятностите е 1:

ξ	x_1	x_2	…	x_n
$P(\xi=x_i)$	p_1	p_2	…	p_n

$$\sum_{i=1}^{n} p_i = 1$$,

В по-общия случай дискретната сл.в. може да има и изброим брой стойности, т.е. нейната таблица на разпределение има вида:

ξ	x_1	x_2	…	x_n	…
$P(\xi=x_i)$	p_1	p_2	…	p_n	…

$$\sum_{i=1}^{\infty} p_i = 1, \quad 0 \le p_i \le 1$$

Пример 8.1 Проверете дали следните таблици задават закони за разпределение на сл.в.:

а)

ξ	2	3	4	5
$P(\xi=x_i)$	0.2	0.3	0.1	0.4

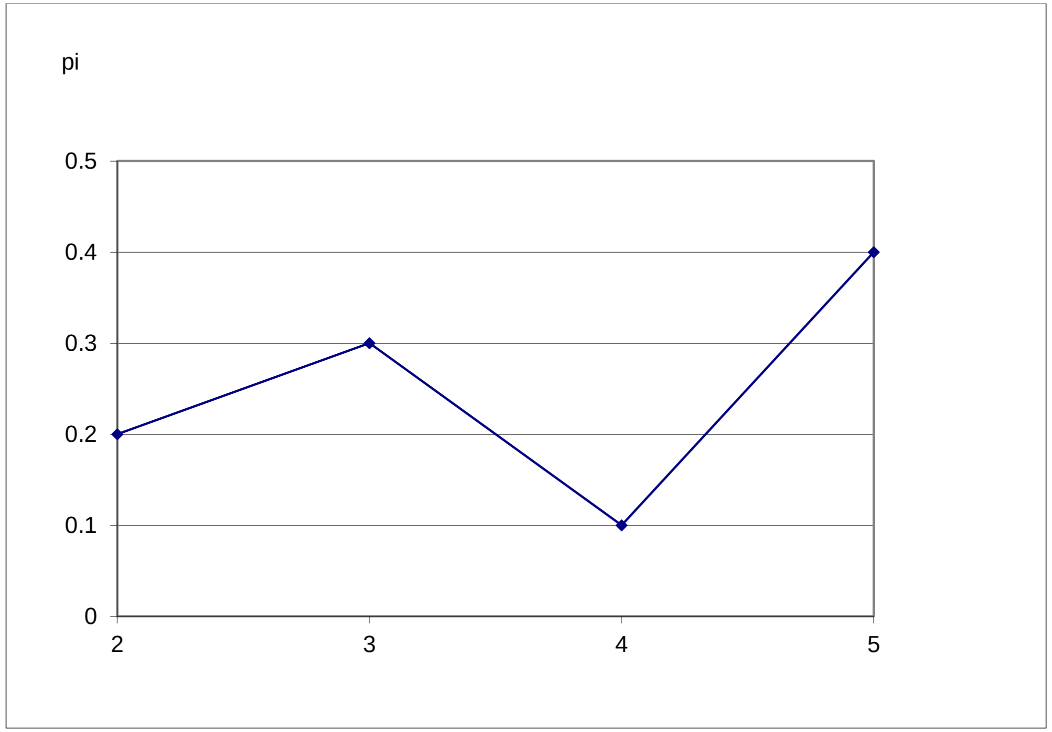
б)

ξ	0	1	2	3
$P(\xi=x_i)$	0.2	0.2	0.4	0.3

Решение: а) Тук $\sum_{i=1}^{4} p_i = 0.2+0.3+0.1+0.4=1$, това е закон за разпределение.

б) $\sum_{i=1}^{4} p_i = 0.2+0.2+0.4+0.3 \neq 1$, т.е. това не е закон за разпределение.

Пример 8.2 Представете графично закона на разпределение от Пример 8.1.



Пример 8.3 В лотария се разиграват 100 билета. От тях има една награда от 50 лв и 10 награди от 1 лв. Представете таблично закона на разпределение на сл.в., описваща стойността на възможната печалба на участник с един лотариен билет.

Решение: Дефиниционното множество са възможните стойности на печалбата, т.е. $D_X = \{ 50, 1, 0 \}$. Вероятността да се спечелят 50 лв е 1 на 100, т.е. 0.01; вероятността да се спечели 1 лв е 0.1, останалото е за 0 лв. Имаме таблицата:

ξ	0	1	50
$P(\xi=x_i)$	0.89	0.1	0.01

Проверка: Сумата на вероятностите е: 0.89 + 0.1 + 0.01 = 1.

8.2. Функция на разпределение (ф.р.) на сл.в.

Определение. Функцията $F(x)$, която показва за всяка стойност на $x \in (-\infty, \infty)$ вероятността сл.в. ξ да приеме стойност по-малка от x, се нарича **функция на разпределение на сл.в.** ξ и се бележи с

$$F(x) = P(\xi < x) = \sum_{x_i < x} p_i .$$

От това определение следва, че ф.р. $F(x)$ на сл.в. ξ с нарастване на x акумулира вероятностите и за дискретна сл.в. се изчислява така:

$$F(x) = \begin{cases} 0, & x < x_1 \\ p_1, & x_1 \le x \le x_2 \\ p_1 + p_2, & x_2 \le x \le x_3 \\ p_1 + p_2 + p_3, & x_3 \le x \le x_4 \\ \ldots \end{cases}$$

От това представяне могат да се получат следните

<u>**Свойства на ф.р.**</u>

1) $F(x) = P(\xi < x) \ge 0$
2) $F(x)$ е монотоннорастяща функция
3) $P(a \le \xi \le b) = F(b) - F(a)$

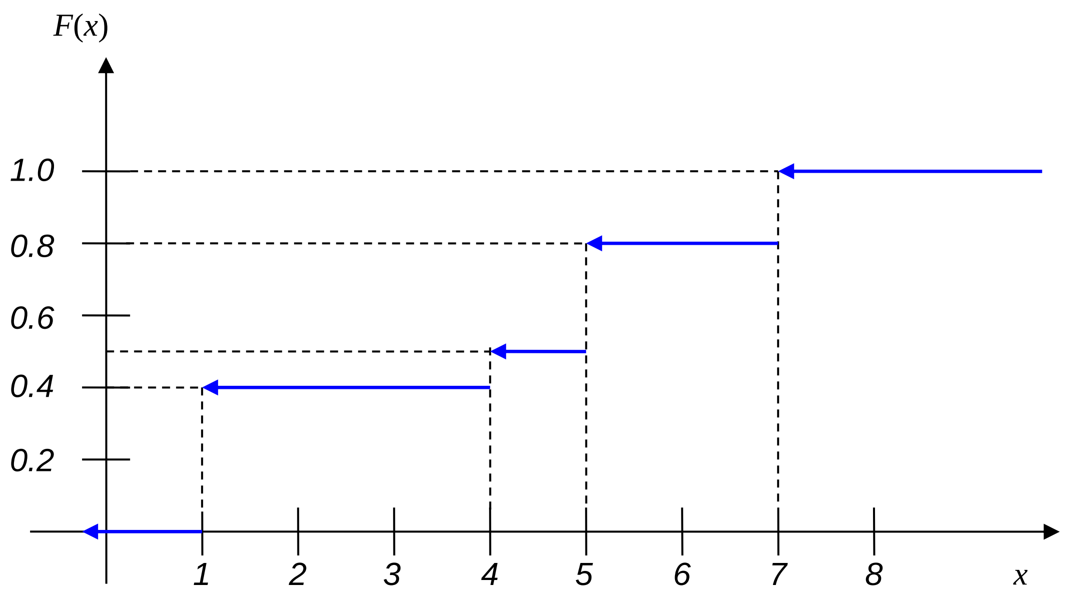
Пример 8.4 Даден е законът на разпределение на сл.в. ξ с долната таблица. Да се намери ф.р. и да се начертае нейната графика.

ξ	1	4	5	7
$P(\xi=x_i)$	0.4	0.1	0.3	0.2

Решение: Имаме $D_X=\{1, 4, 5, 7\}$. За всяка стойност на $x\in D_X \subset(-\infty,\infty)$ последователно изчисляваме $F(x)$:

1. При $x<1$, $F(x)=P(\xi<1)=0$.
2. При $1\le x\le 4$, (напр. за $x=2$) имаме $F(x)=P(\xi<4)=P(\xi=1)=0.4$.
3. При, $4\le x\le 5$ (напр. за $x=4.33$) имаме
   $F(x)=p(\xi<5)=P(\xi=1)+P(\xi=4)=p_1+p_2=0.4+0.1=0.5$
4. При $5\le x\le 7$, $F(x)=p(\xi<7)=P(\xi=1)+P(\xi=4)+P(\xi=5)=0.5+0.3=0.8$
5. При $7\le x$, $F(x)=P(\xi=1)+P(\xi=4)+P(\xi=5)+P(\xi=7)=0.8+0.2=1$

Графиката на получената ф.р. е:



Виждаме, че ф.р. на дискретна сл.в. е дефинирана за цялата реална ос. Ф.р. е прекъсната в точките от дефиниционното множество, като големините на отскоците са равни на стойностите на съответните вероятности в тези точки.

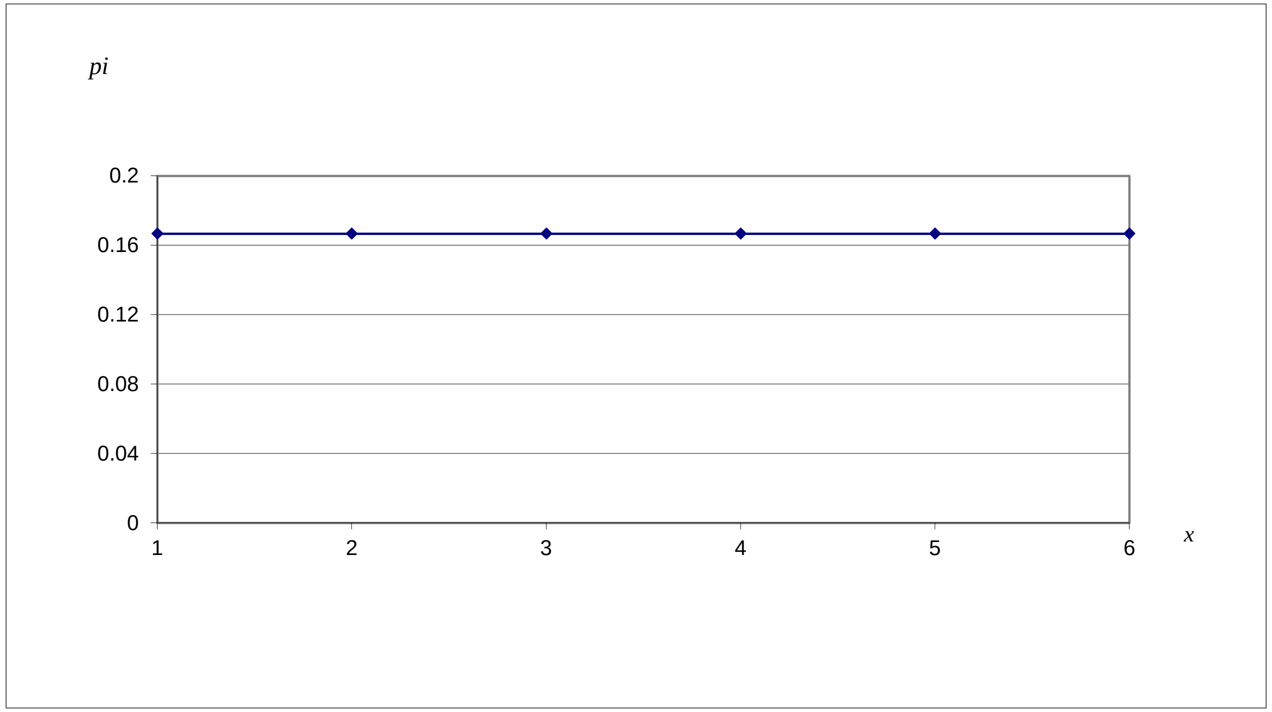
Равномерно разпределение

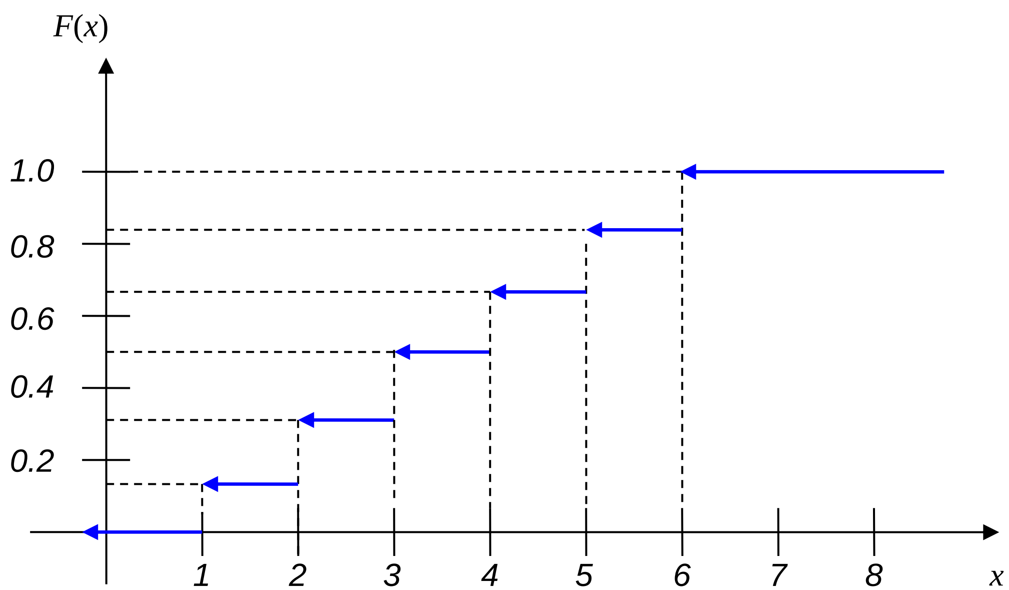
Това е случаят на равни вероятности.

Пример 8.5 Хвърляме симетричен зар. Нека сл.в. ξ е броят паднали се точки. Намерете закона на разпределение. Постройте ф.р.

Решение: Имаме 6 възможности, т.е. $D_X=\{1, 2, 3, 4, 5, 6\}$. Законът на разпределение е:

ξ	1	2	3	4	5	6
$P(\xi=x_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$


pi
0.2
0.16
0.12
0.08
0.04
0
1
2
3
4
5
6
x


F(x)
1.0
0.8
0.6
0.4
0.2
1
2
3
4
5
6
7
8
x

Геометрично разпределение

Провежда се серия от независими опити, при всеки от които събитието A настъпва с една и съща вероятност p. Опитите продължават до първата поява на A.

Пример 8.6 Стрелец стреля, докато улучи целта. При всеки изстрел вероятността да улучи е 0.6. Да се намери законът на разпределение на случайната величина ξ – брой на изстреляните напразно патрони.

Решение: Щом успехът е с вероятност p=0.6, то неуспехът (не улучва) е с вероятност q=1-0.6=0.4. Ако ξ =0, то няма напразно изстреляни патрони, т.е. стрелецът уцелва от първия път – вероятността $P[\xi=0]=p=0.6$. При $\xi=1$ стрелецът има един напразен изстрел и улучва с втория патрон, т.е. $P[\xi=1]=p.q=0.6*0.4=0.24$. Аналогично $P[\xi=2]=p.q^2=0.6*0.4^2$ $=0.24*0.4=0.096$, и т.н. Получаваме таблица на разпределението:

ξ	0	1	2	3	4	5	…
p_i	p	$p.q$	$p.q^2$	$p.q^3$	$p.q^4$	$p.q^5$	…
p_i	0.6	0.24	0.096	0.0384	0.01536	0.0061	

Биномно разпределение – схема на Бернули

Редица от n независими опита, при които събитието A настъпва с една и съща вероятност $P(A)=p$ (и следователно $P(\bar{A}) = 1 - p = q$) се нарича схема на Бернули и се бележи с (n,p). С υ се означава броят на настъпване на събитието A (или брой "успехи").

Вероятността да настъпят точно k "успехи", т.е. точно k настъпвания на събитието A се бележи с $P_n(k)$. Изпълнени са равенствата:

$$P_n(k) = C_n^k p^k q^{n-k}, \; k = 0,1,\dots,n$$

$$P_n(k_1 \le \upsilon \le k_2) = \sum_{k=k_1}^{k_2} P_n(k)$$ - вероятността за брой “успехи” от k_1 до k_2, $P(\upsilon \ge 1) = 1 - q^n$ - вероятност за поне един “успех”.

Пример 8.7 Монета се хвърля 4 пъти. Да се намери разпределението на сл.в. ξ – брой паднали се гербове. Да се направи графика на разпределението и на функцията на разпределение.

Решение: Вероятностите се пресмятат по формулата на Бернули.

Тук n=4, Събитието е A=”Пада се герб”, т.е. p=q=0.5.

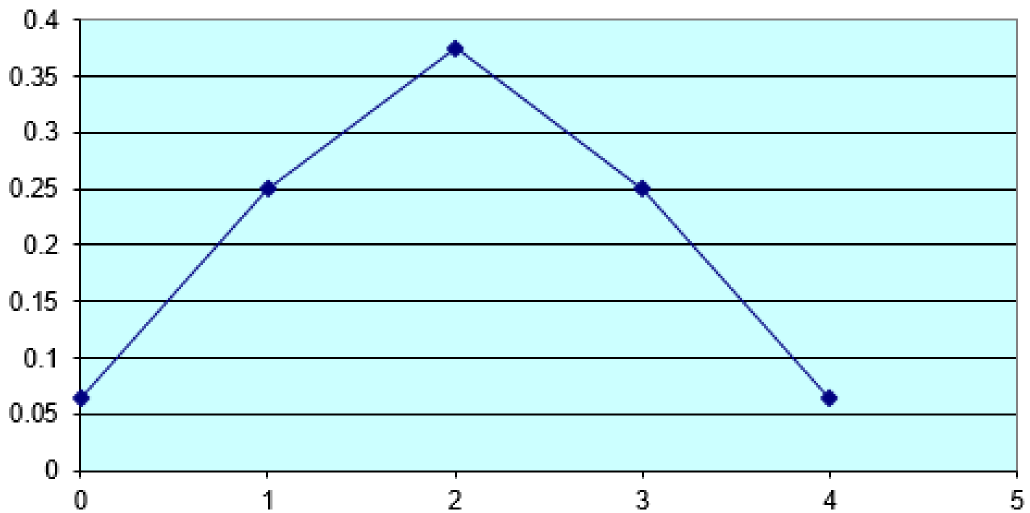
$P_4(0) = C_4^0 q^4 = (0.5)^4 = 0.0625$, $P_4(1) = C_4^1 pq^3 = 4(0.5)^4 = 0.25$,

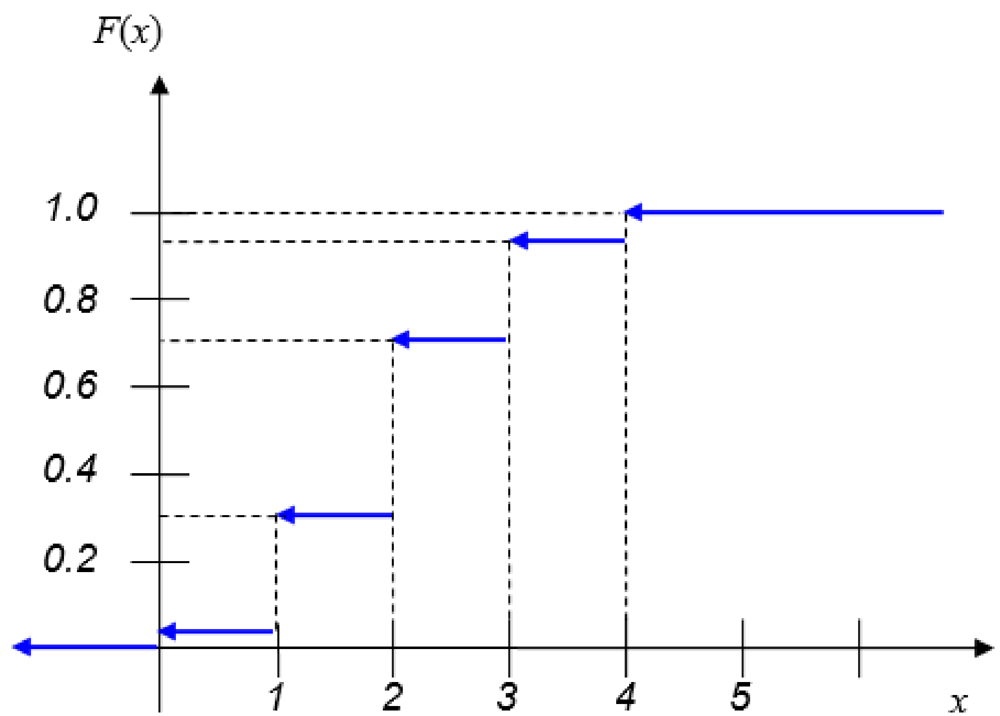
$P_4(2) = C_4^2 p^2 q^2 = 6(0.5)^4 = 0.375$, $P_4(3) = C_4^3 p^3 q = 4(0.5)^4 = 0.25$,

$P_4(4) = C_4^4 p^4 = (0.5)^4 = 0.0625$

ξ	0	1	2	3	4
p_i	0.0625	0.25	0.375	0.25	0.0625

Биномно разпределение





Задача. Известно е, че 10% от лекарствата на фармацевтична фирма не покриват европейските стандарти. Проверяват се случайно взети 8 опаковки лекарства на фирмата. Да се намери разпределението на сл.в. ξ – брой стандартни лекарства от избраните.

Упътване. Използвайте формулите за биномно разпределение.

Числови характеристики на дискретни случайни величини

Математическо очакване на дискретна сл.в. ξ	$E(\xi)=\sum_{i=1}^{n} x_i p_i = x_1 p_1 + x_2 p_2 + \ldots + x_n p_n$
Дисперсия на дискретна сл.в. ξ	$D(\xi)=\sum_{i=1}^{n}\left(x_i - E(\xi)\right)^2 p_i$ или $D(\xi)=\sum_{i=1}^{n} x_i^2 . p_i - \left(E(\xi)\right)^2$
Стандартно отклонение (средно квадратично отклонение)	$\sigma_\xi = \sqrt{D(\xi)}$
Мода	$Mo = x_j$, за което се получава максималната вероятност

Някои основни дискретни разпределения и числовите им характеристики

Разпределение	Формула за изчисляване на вероятността $p_k = P\{\xi = k\}$	Математическо очакване $E(\xi)$	Дисперсия $D(\xi)$
Равномерно	$p_k = \frac{1}{n}$, $k = 1,2, \ldots, n$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$
Геометрично $0<p<1$, $q=1-p$	$q^k p$, $k = 0, 1, \ldots$	$\frac{q}{p}$	$\frac{q}{p^2}$
Биномно $0<p<1$, $q=1-p$	$P_n(k) = C_n^k p^k q^{n-k}$, $k = 0, 1, \ldots, n$	$n.p$	$n.p.q$

Пример 8.8 За пример 8.6 да се изчислят математическото очакване E, дисперсията D и модата Mo.

Решение: От таблицата на разпределението или директно по горните формули:

ξ	0	1	2	3	4	5	…
p_i	p	$p.q$	$p.q^2$	$p.q^3$	$p.q^4$	$p.q^5$	…
p_i	0.6	0.24	0.096	0.0384	0.01536	0.0061	

$Mo=0$, $E(\xi)=q/p=0.4/0.6=0.667$, $D(\xi)=q/p^2=0.4/0.6^2=1.11$.

Пример 8.9 За пример 8.7 да се изчислят модата, математическото очакване и дисперсията.

Решение: Вероятностите се пресмятат по формулата на Бернули.

Тук $n=4$, $p=q=0.5$. Тогава $E(\xi)=n.p = 4.\ 0{,}5=2$; $D(\xi)=n.p.q=2.\ 0{,}5=1$;
$Mo=2$.

Непрекъснати сл.в. и разпределения

Непрекъснатата сл.в. е функция, определена за всяко $-\infty < x < \infty$. Тя има безброй много стойности, разположени плътно в интервал от реалната права или върху цялата права. Примери за непрекъснати сл.в. са: времето за излитане на самолет, промяната на напрежението на апарат за 1 час и др.

За непрекъсната сл.в. не се разглежда стойността на вероятността й във фиксирана точка, там тя е равна на нула. За непрекъсната сл.в. се задава вероятността, която тя приема в някакъв интервал от числовата права. Вместо таблица, вероятностите се определят с използване на т.н. вероятностна плътност на разпределение.

<u>Определение 8.2</u> Вероятностна плътност на разпределение на непрекъснататата сл.в. ξ се нарича реалната функция $f(x),\ x \in (-\infty, \infty)$, със следните свойства:

а) $f(x) \geq 0$, б) $\int\limits_{-\infty}^{\infty} f(x)\,dx = 1$, в) $P(\alpha < \xi < \beta) = \int\limits_{\alpha}^{\beta} f(x)\,dx$.

Тук свойството б) означава, че лицето на фигурата, заключено между графиката на $f(x)$ и оста Ox е равно на 1.

<u>Определение</u>. <u>Функция на разпределение (ф.р.) на непрекъсната сл.в. :</u>

$$F(x) = P(\xi < x) = \int_{-\infty}^{x} f(x)\,dx, \qquad -\infty < x < \infty$$

Оттук след диференциране получаваме връзката между ф.р. и плътността на разпределението: $f(x) = F'(x)$

Като използваме по-горното свойство в) и свойствата на определения интеграл, получаваме, че вероятността сл.в. ξ да приема стойност в който и да е интервал $(\alpha;\beta)$, $(\alpha;\beta]$, $[\alpha;\beta]$, или $[\alpha;\beta)$ е:

$$P(\alpha < \xi < \beta) = F(\beta) - F(\alpha)$$

<u>Свойства на ф.р.</u>

1) $F(x) = P(\xi < x)$, за всяко $x \in (-\infty, \infty)$
2) $F(x)$ е монотоннорастяща функция
3) $\lim\limits_{x\to-\infty} F(x) = 0$ и $\lim\limits_{x\to\infty} F(x) = 1$
4) $P(\alpha \le \xi \le \beta) = F(\beta) - F(\alpha)$

Математическо очакване на непр. сл.в. :

$$E(\xi)=\int_{-\infty}^{\infty} x\, f(x)\, dx$$

Дисперсия:

$$D(\xi)=\int_{-\infty}^{\infty} (x-E(\xi))^2\, f(x)\, dx$$

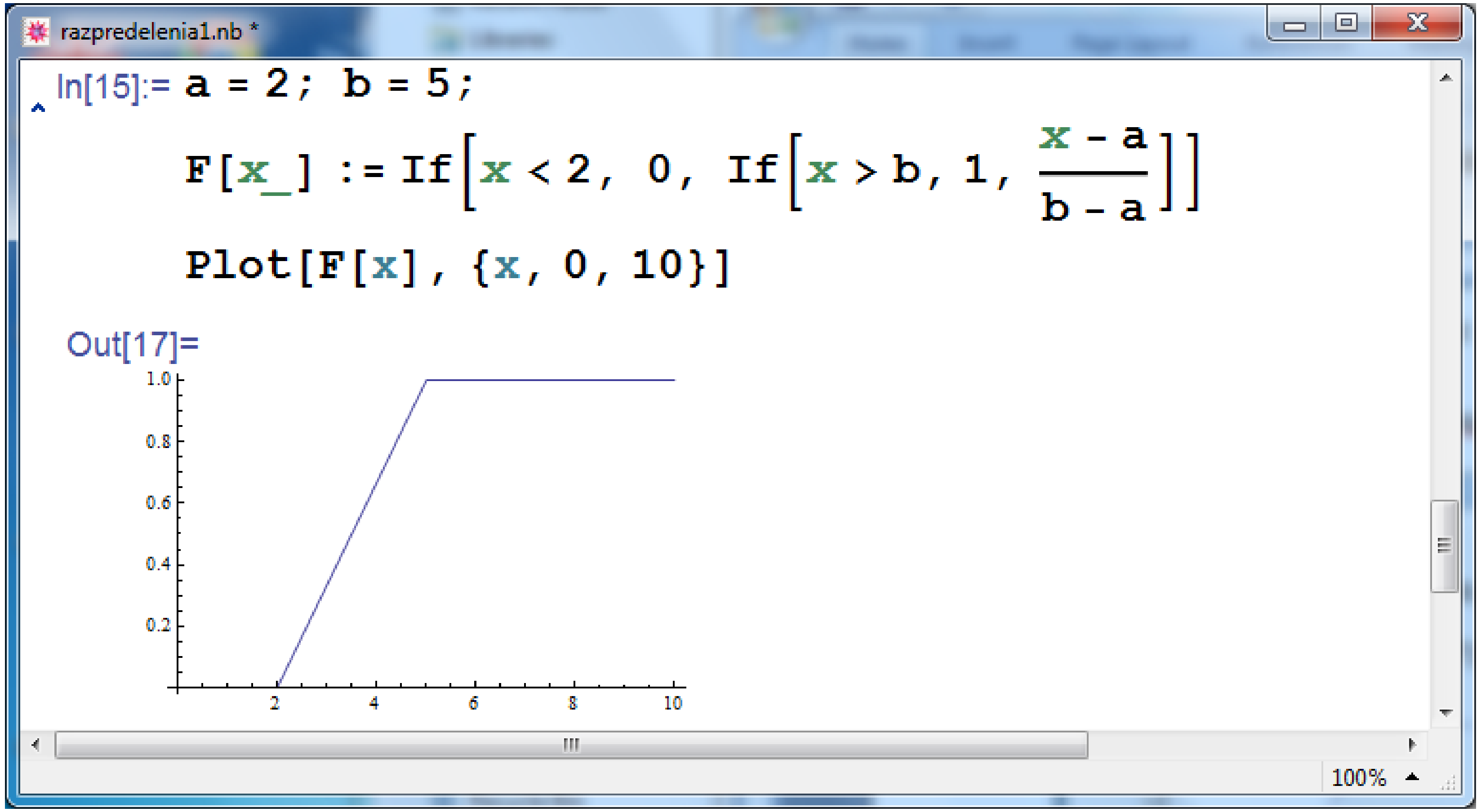
α-квантил на разпределение: Нека $0<\alpha<1$, а x_α е число, което е решение на уравнението $F(x_\alpha)=\alpha$. Тогава x_α се нарича α-квантил на разпределението на сл.в. ξ. Ясно е, че $F(x_\alpha)=P(\xi<x_\alpha)=\alpha$.

<u>Равномерно разпределение</u>. Има ф.р.:

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

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

Пример 8.10 Да се изобрази функцията на равномерното разпределение при a=2, b=5.



Нормално разпределение (закон на Гаус) ~ N(m,σ):

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

, σ>0, m – параметри,

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-m)^2}{2\sigma^2}} dt$$

, $E(\xi) = m = Mo = Md, \ D(\xi) = \sigma^2$.

<u>Стандартно нормално разпределение ~ N(0,1):</u>

Задава се с вероятностна плътност на разпределение:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty < x < \infty$$

.

Ф.р. на стандартното нормално разпределение съответно е:

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

.

Тук константата $\frac{1}{\sqrt{2\pi}}$ се слага, за да се гарантира, че площта под графиката на $\varphi(x)$ е равна на единица.

Това лесно се проверява и с Mathematica:

In[37]:= $\varphi[t\_] := \frac{1}{\sqrt{2\,\pi}}\, e^{-\frac{t^2}{2}}$

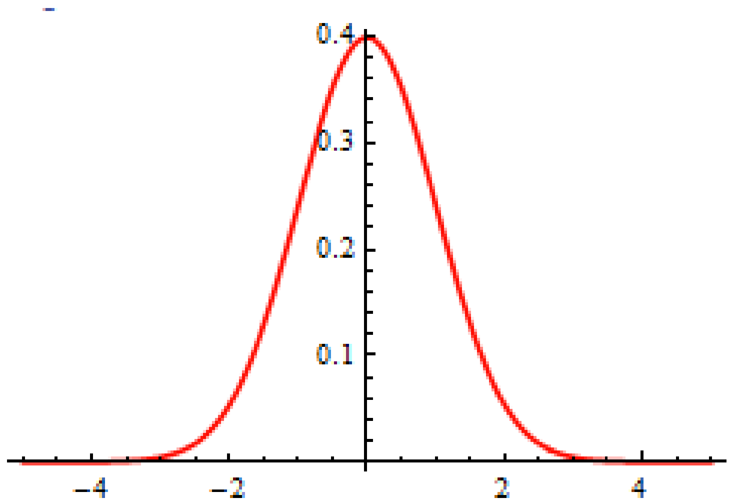
$\int_{-\infty}^{\infty} \varphi[t]\, dt$

```
Plot[φ[t], {t, -5, 5}]
```

Out[38]=

1

Out[39]=

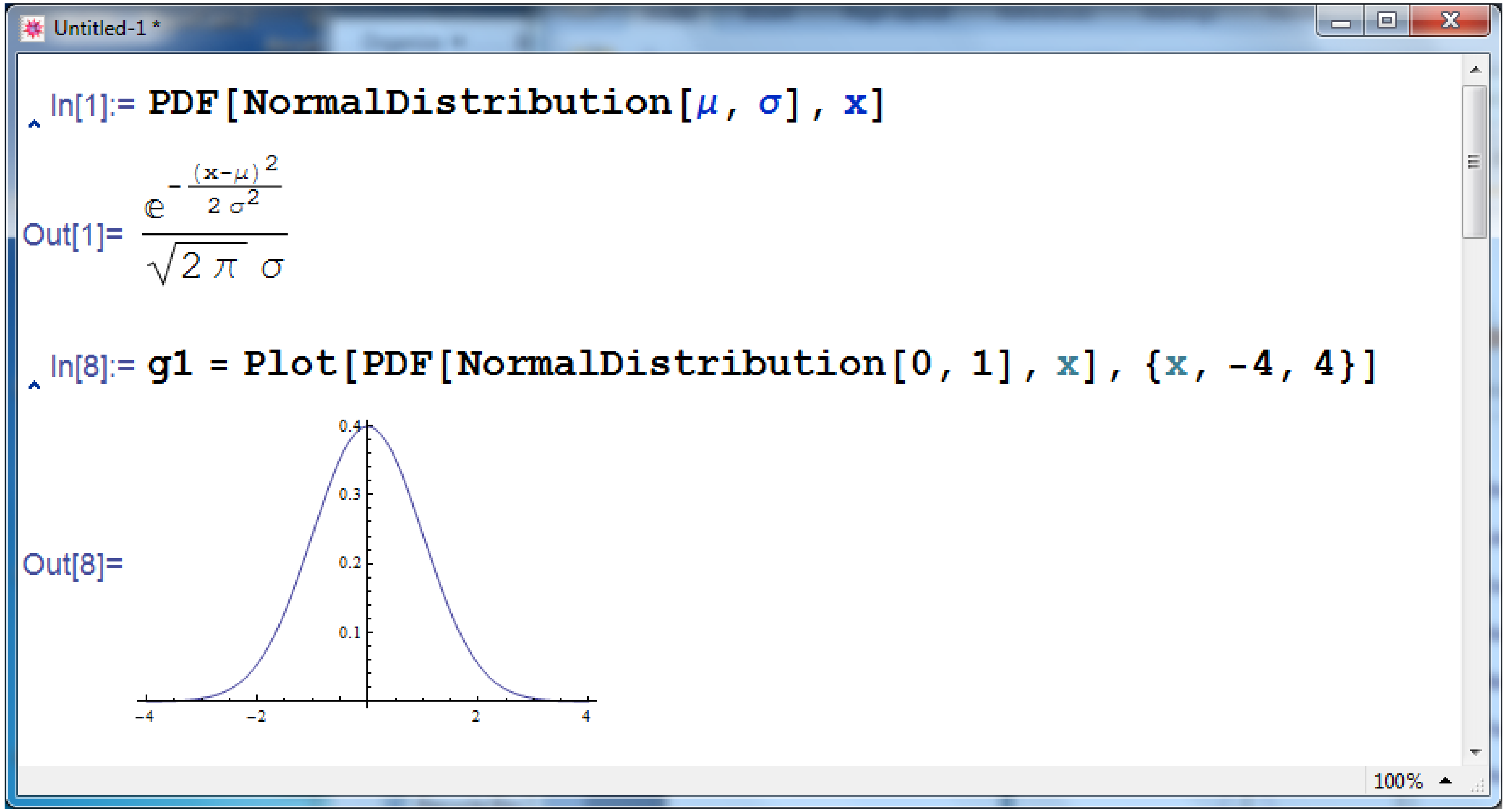


Получената графика на стандартното нормално разпределение има характерна форма на **камбана**, с максимална стойност в 0, симетрична спрямо нулата и извън интервала (-3,3) е почти нула:

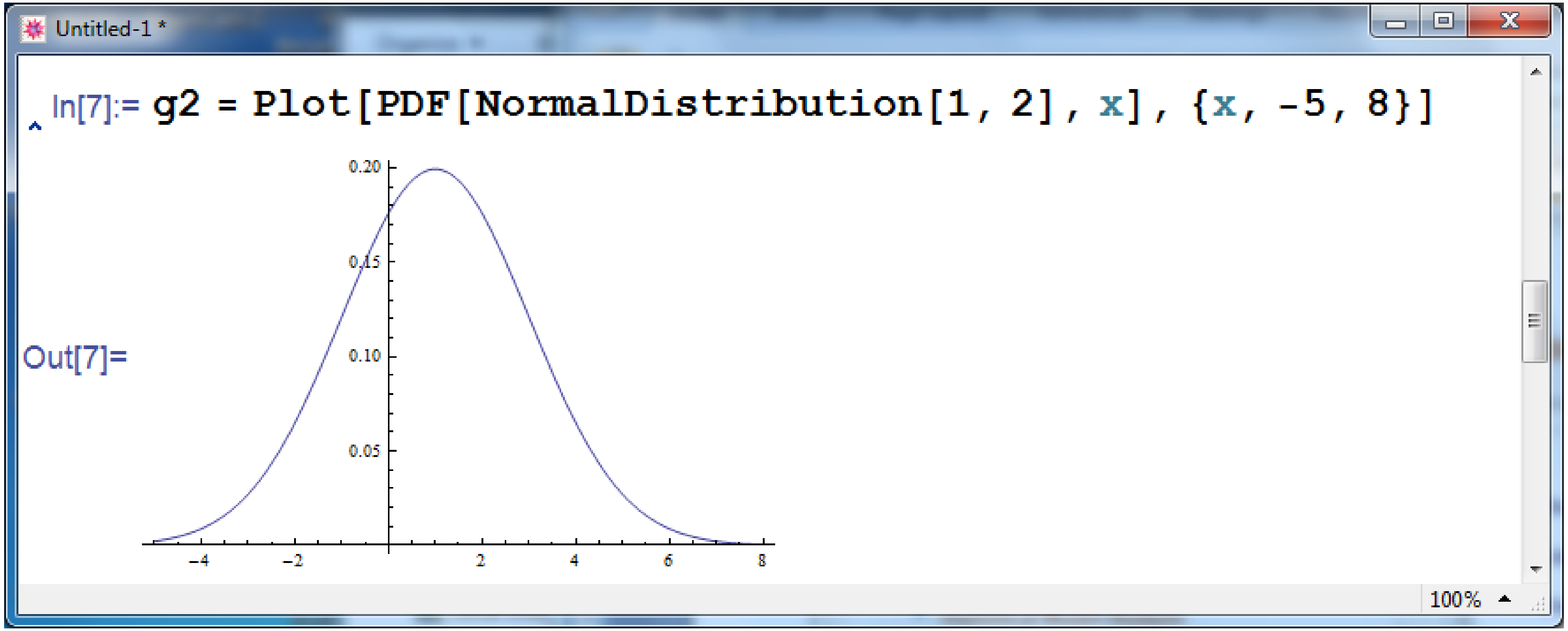
```
In[42]:= φ[0.]

Out[42]=
    0.398942
```

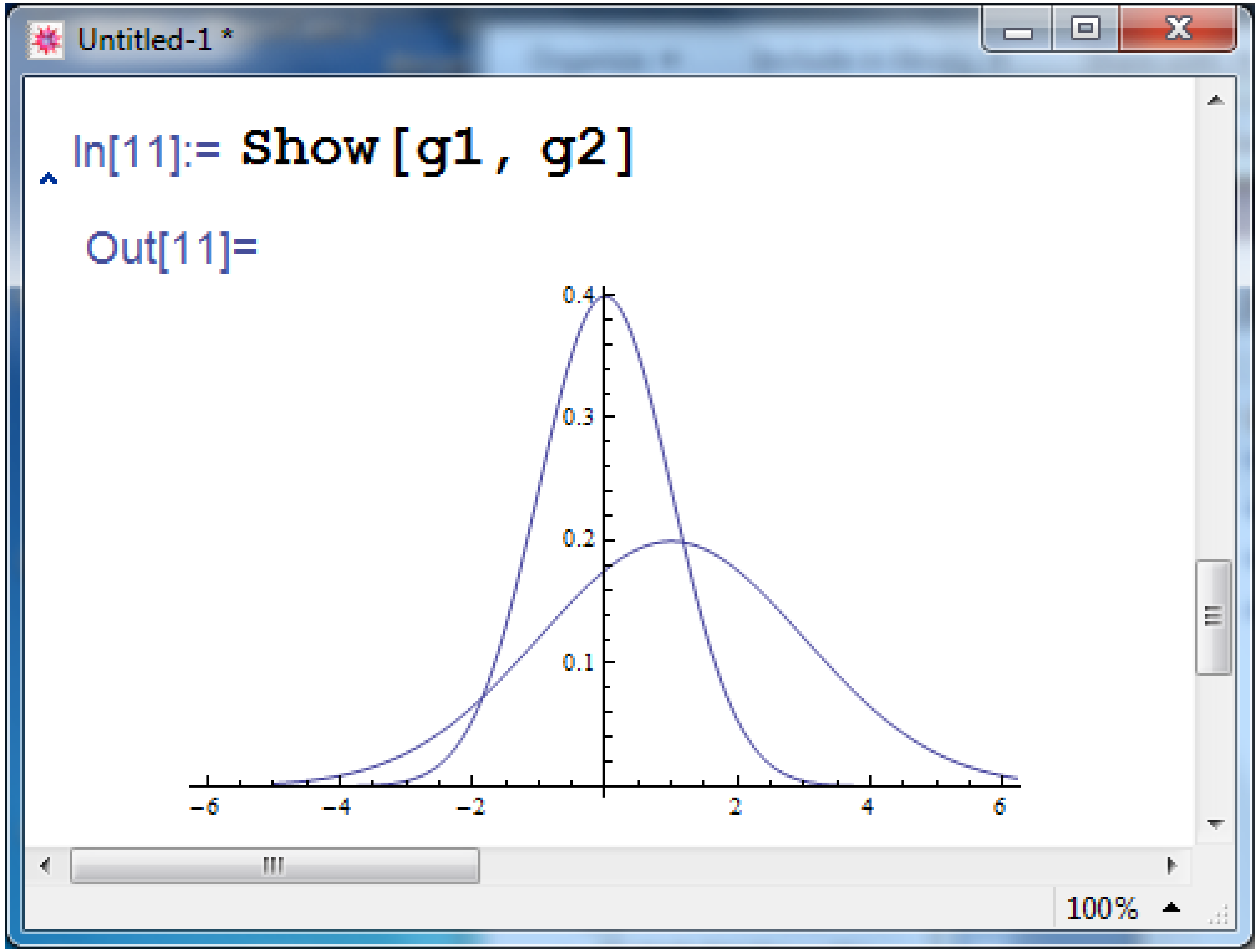
Друг вариант е да се ползват вградените функции на Mathematica, напр.:



За N(1,2) графиката е:



Сравнение на двете графики получаваме с:



Математическо очакване и дисперсия на стандартното нормално разпределение:

$$E(\xi) = Mo = Md = 0, \; D(\xi) = 1.$$

Свойства на стандартното нормално разпределение:

За x>0: $\varphi(-x) = \varphi(x)$,

$$\Phi(-x) = 1 - \Phi(x),$$

$$P(0 < \xi < x) = P(-x < \xi < 0) = \Phi(x) - \frac{1}{2},$$

$$P(-x < \xi < x) = P(|\xi| < x) = 2\Phi(x) - 1,$$

$$P(\alpha < \xi < \beta) = \Phi(\beta) - \Phi(\alpha).$$

Нормиране: $F(x) = \Phi(\frac{x-m}{\sigma})$

Вероятността нормално разпределената сл.в. ξ с разпределение N(m,σ) да принадлежи на интервала (α,β) е:

$$P(\alpha < \xi < \beta) = F(\beta) - F(\alpha) = \Phi(\frac{\beta - m}{\sigma}) - \Phi(\frac{\alpha - m}{\sigma}).$$

Пример 8.11 Непрекъснатата сл.в. ξ има функция на разпределение $F(x)$, зададена с формулата:

$$F(x) = \begin{cases} 0 & \text{при } x \le 0 \\ \frac{2}{3}x - \frac{1}{9}x^2 & \text{при } 0 < x \le 3 \\ 1 & \text{при } x \ge 3 \end{cases}$$

а) Да се намери вероятността сл.в. ξ да приеме стойност в интервала [2; 4).

б) Да се определи плътността на разпределение.

Решение: а) По дефиниция $P(2 \leq \xi < 4) = P(2 < \xi < 4) =$

$$F(4) - F(2) = 1 - (\frac{2}{3}.2 - \frac{1}{9}.2^2) = 1 - \frac{8}{9} = \frac{1}{9}.$$

б) От връзката между ф.р. и плътността имаме:

$$f(x) = F'(x) = \begin{cases} 0 & \text{при } x \leq 0 \\ \frac{2}{3} - \frac{2}{9}x & \text{при } 0 < x \leq 3 \\ 1 & \text{при } x > 3 \end{cases}$$

Пример 8.12 Непрекъснатата сл.в. ξ има нормално разпределение с параметри m=6 и σ=0,4. Да се намери:

а) Вероятността $P(\xi < 1,6)$

б) Вероятността $P(1,3 < \xi < 2,3)$

в) α-квантилът при α=0,95.

Решение: а) $P(\xi < 1,6) = F(1,6) = \Phi(\frac{1,6-m}{\sigma}) = \Phi(\frac{1,6-6}{0,4})$

$$= \Phi(\frac{-4,4}{0,4}) = \Phi(-1,1) = 1 - \Phi(1,1) = 1 - 0,86433 = 0,13567$$

.

Обърнете внимание, че тук преминаването от $F(x)$ към стандартното нормално разпределение $\Phi(x)$ става чрез формулата за нормиране.

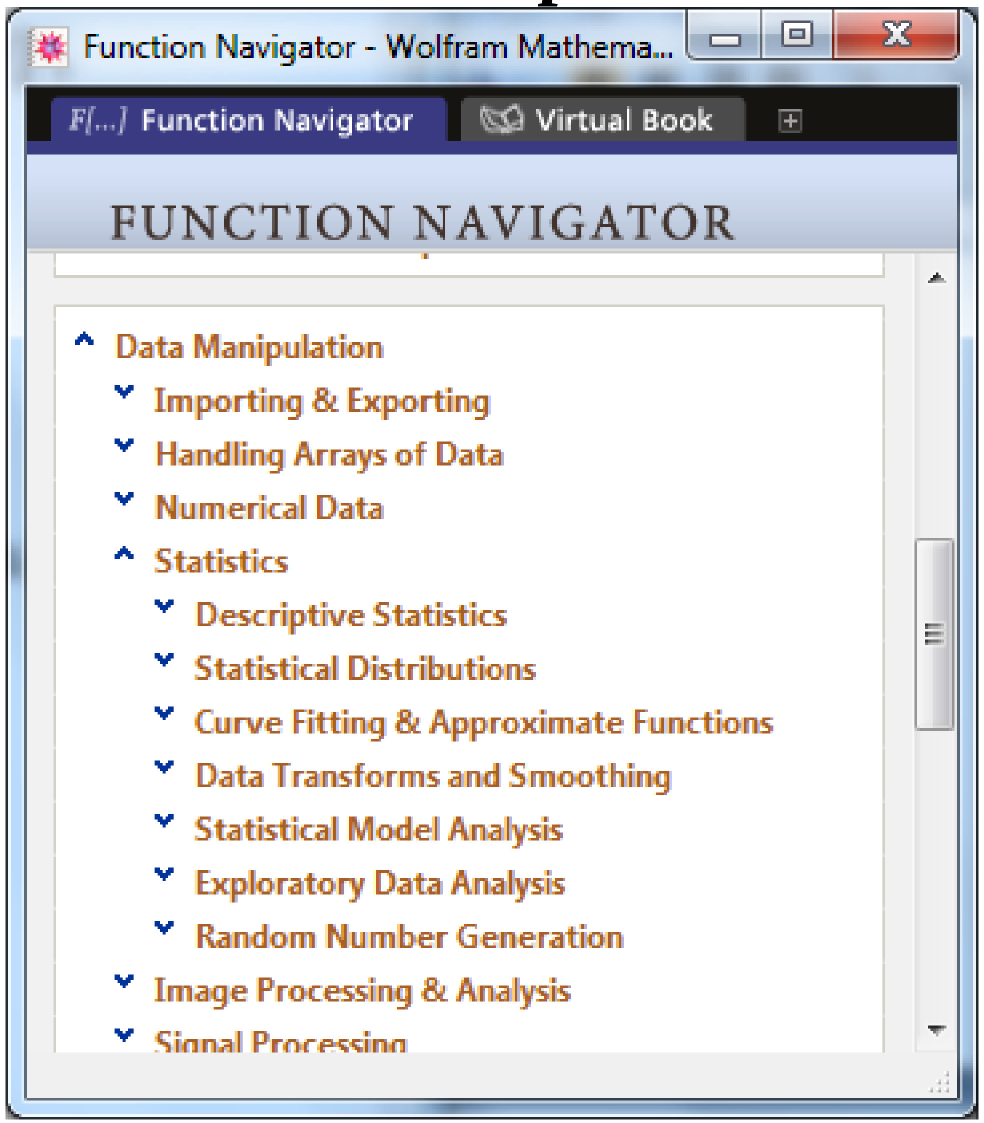
в) В този случай е зададена вероятността α=0,95 и търсим за кое x_α тя се достига. Т.е. трябва да решим уравнението $F(x_\alpha) = 0{,}95$. От нормирането имаме:

$$F(x_\alpha) = \Phi(\frac{x_\alpha - m}{\sigma}) = \Phi(\frac{x_\alpha - 6}{0{,}4}) = 0{,}95.$$

От таблицата на стандартното нормално разпределение намираме, че вероятност 0,95 се достига приблизително при аргумент 1,65. Оттук се получава

$\frac{x_\alpha - 6}{0{,}4} = 1{,}65$, откъдето $x_\alpha = 6 + 0{,}4.1{,}65 = 6{,}66$.

Подробно описание на възможностите на система Математика за статистическа обработка на данни може да се намери в:
***Mathematica Help*/Function navigator/Data manipulation/Statistics**



9. Елементи на статистиката

9.1. Основни понятия на статистиката

Статистиката е дял от математиката за изучаване на случайни явления на базата на обработка на емпирични данни с цел извличане на заключения за тях.

Определение. Генерална съвкупност (ГС) (популация) е крайно или безкрайно, конкретно или принципиално множество от статистически единици (обекти, случаи), които имат поне един общ признак.

Извадка е част (подмножество) от ГС, която подлежи на изследване. Стремежът е да се изследва такава извадка, която дава вярна престава за състоянието на изучаваното явление в ГС.

Първична обработка на данни

Вариационен ред е подреден по големина списък на данните от извадката.

Пример 9.1 Изходни данни:

12, 23, 14, 23, 11, 12, 25, 8, 9, 12, 11, 14, 25, 12, 9.

Вариационен ред:

8, 9, 9, 11, 11, 12, 12, 12, 12, 14, 14, 23, 23, 25, 25.

Статистическият ред или честотната таблица има вида:

x_i	x_1	x_2	x_3	…	x_k
n_i	n_1	n_2	n_3	…	n_k

където на първия ред се записват във възходящ ред само различните стойности на елементите от извадката, а във втория – тяхната честота, т.е. колко пъти се срещат в извадката. Общият брой данни (наблюдения) n се нарича **обем на извадката**.

Пример 9.2 От горните данни статистическият ред е:

x_i	8	9	11	12	14	23	25
n_i	1	2	2	4	2	2	2

Относителна честота на даден елемент от статистическия ред е: $v_i = \dfrac{n_i}{n}$.

x_i	x_1	x_2	x_3	…	x_k
$v_i = \dfrac{n_i}{n}$	$\dfrac{n_1}{n}$	$\dfrac{n_2}{n}$	$\dfrac{n_3}{n}$	$\dots$	$\dfrac{n_k}{n}$

Пример 9.3 На следващата таблица, в третия ред са дадени съответните честоти, брой данни n=15.

x_i	8	9	11	12	14	23	25
n_i	1	2	2	4	2	2	2
$v_i = \dfrac{n_i}{n}$	$\dfrac{1}{15}$	$\dfrac{2}{15}$	$\dfrac{2}{15}$	$\dfrac{4}{15}$	$\dfrac{2}{15}$	$\dfrac{2}{15}$	$\dfrac{2}{15}$

<u>Групиран статистически ред.</u> Използва се, когато данните са много. Тогава вариационният ред се разделя на интервали, обикновено с равни дължини, като във всеки интервал трябва да има данни. Броят на подинтервалите k може да се намери от неравенството $2^k>n$. Определят се средите на получените интервали, да ги означим с m_i. Построява се таблица на групирания статистически ред, като на първия ред се записват стойностите на m_i, а на втория – честотите, т.е. броят на данните, които се намират в дадения подинтервал.

M_i	m_1	m_2	m_3	...	m_k
f_i	f_1	f_2	f_3	...	f_k

Пример 9.4 Нека необработените данни от наблюдение са:
68 71 77 83 79 72 74 57 67 69 50 60 70 66 76 70 84 59 75 94 65 72 85 79 71
83 84 74 82 97 77 73 78 93 95 78 81 79 90 83 80 84 91 101 86 93 92 102 80 69

Размерът на извадката е n=50. Приблизително 2^6>50, затова избираме брой подинтервали k=6. Дължините на интервалите нека са равни : $h=(x_{max}-x_{min})/k=(102-50)/6=8{,}7\approx 9$. По-добре е да вземем тук 10 вместо 9, т.е. h=10. Групираният статистически ред е даден в долната таблица.

интервал	45-55	55-65	65-75	75-85	85-95	95-105
m_i	50	60	70	80	90	100
f_i	1	4	15	19	8	3
v_i	$\frac{1}{50}$	$\frac{4}{50}$	$\frac{15}{50}$	$\frac{19}{50}$	$\frac{8}{50}$	$\frac{3}{50}$

Точкови и интервални оценки (характеристики на извадка)

$\bar{x}$ - **Средноаритметична стойност** на извадката. Тя е ефективна оценка на математическото очакване на генералната съвкупност (ГС):

За вариационен ред: $$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

За статистически ред: $$\bar{x} = \frac{x_1 n_1 + x_2 n_2 + \ldots + x_k n_k}{n}$$

За групиран статистически ред: $$\bar{x} = \frac{m_1 f_1 + m_2 f_2 + \ldots + m_k f_k}{n}$$

$\bar{x}_w$ - **Средноаритметична стойност с тегло w_i за наблюдението x_i:**

$$\bar{x}_w = \frac{x_1 w_1 + x_2 w_2 + \ldots + x_n w_n}{w_1 + w_2 + \ldots + w_n}$$

Md - Медиана на извадката. Тя е оценка на медианата на ГС. Нейната стойност се явява среда на извадката, при което броят на елементите по-малки или равни на Md е равен на броя на елементите по-големи или равни на Md.

За вариационен и статистически ред: Md = средния елемент на извадката при нечетно n

Md = средно аритметично от двата средни елемента при четно n

За групирани данни: $$Md = (m_{md} - \frac{h}{2}) + h\frac{\frac{n}{2} - \sum_{j=1}^{md-1} f_j}{f_{md}},$$

	$\frac{n}{2}$ - определя мястото на медианната група $i= md$, т.е. интервалът на медианата, където натрупаната честота е $\geq \frac{n}{2}$. h – ширина на медианната група със среда m_{md}, f_{md} – честота в медианната група, $\sum_{j=1}^{md-1} f_j$ - натрупаната честота в предходната група.
Mo - Мода на извадката. Тя е най-често срещаното значение в извадката.	
От статистическия ред:	**Mo** = x_i, за което честотата n_i е най-голяма.
За групирани данни:	$Mo = (m_{Mo} - \frac{h}{2}) + h \frac{f_{Mo} - f_{Mo-1}}{(f_{Mo} - f_{Mo-1}) + (f_{Mo} - f_{Mo+1})}$,

Индексът $i=Mo$ е мястото на модалната група, т.е. интервалът, в който честотата f_i е най-голяма,
h – ширина на модалната група със среда m_{Mo},
f_{Mo}, f_{Mo-1}, f_{Mo+1} – честоти в модалната, предишната и следващата групи.

s^2 - Дисперсия на извадката (извадкова дисперсия). Тя е ефективна оценка на дисперсията на ГС:

За вариационен ред: $$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

За статистически ред: $$s^2 = \frac{1}{n_1 + n_2 + \ldots + n_k - 1}\sum_{i=1}^{k} n_i (x_i - \bar{x})^2$$

За групиран статистически ред:	$$s^2 = \frac{1}{f_1 + f_2 + \ldots + f_k - 1} \sum_{i=1}^{k} f_i (m_i - \bar{x})^2$$
Стандартното отклонение на извадката е ефективна оценка на стандартното отклонение на ГС. $$s = \sqrt{s^2}$$	

Забележка. Примери за точкови и интервални оценки ще бъдат разгледани на упражнения.

Изследване на зависимости

Корелация

Нека x е представителна извадка на ГС X, а y – представителна извадка на ГС Y.

Корелацията е вид евентуална зависимост между двете генерални съвкупности X и Y. За да установим дали има корелационна зависимост трябва да пресметнем и оценим корелационния коефициент ρ. Негово приближение е т.н. извадков коефициент на корелация r. Формулите за пресмятане на извадковия коефициент на корелация r са:

$$SSx = \sum (x_i - \bar{x})^2 = \sum x_i^2 - \frac{(\sum x_i)^2}{n}, \quad SSy = \sum (y_i - \bar{y})^2 = \sum y_i^2 - \frac{(\sum y_i)^2}{n},$$

$$SSxy = \sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i\, y_i - \frac{(\sum x_i)(\sum y_i)}{n},$$

откъдето:

$$r = \frac{SSxy}{\sqrt{SSx}\sqrt{SSy}}.$$

Корелационният коефициент е число между -1 и 1, т.е. $-1 \le r \le 1$. Колкото по-голяма е неговата стойност по абсолютна стойност до 1, толкова по-силна е зависимостта между *x* и *y*. Когато $r > 0$, наличието на корелационна зависимост означава, че с нарастване на *x* зависимата променлива *y* също расте. Когато $r < 0$ с нарастване на *x* променливата *y* намалява.

Практическо правило: Ако $|r|>0,85$, то може да се предполага, че между x и y има зависимост, но това не е достатъчно.
След определянето на корелационния коефициент r се прави проверка за неговата значимост.

При предположение за нормално разпределени ГС: $X \sim N(m_1,\sigma_1)$, $Y \sim N(m_2,\sigma_2)$ се разглеждат хипотезите:

H_0: $\rho=0$, т.е. няма корелационна зависимост между X и Y;

H_1: $\rho \neq 0$, т.е. има корелационна зависимост между X и Y.

За проверката на тези хипотези се прилага t-статистиката по разпределението на Стюдънт. Това става по следния начин:

1. Най-напред като се използва r се пресмята

$$t=\frac{r}{\sqrt{\dfrac{1-r^2}{n-2}}}.$$

2. След това се намира $t_{\text{критично}}$ - квантил на t-разпределението на Стюдънт с n-2 степени на свобода и порядък $q = 1 - \frac{\alpha}{2}$, където α е избрано ниво на значимост, напр. α= 5%. Намирането на $t_{\text{кр.}}$ може да стане по таблици или с функцията на Excel TINV(α;n-2). Степените на свобода се определят, като от размера на извадката n се извади броят на участващите променливи. В нашия случай имаме брой променливи равен на 2 (x и y).
3. Прилага се решаващото правило:

Решаващо правило: Ако $|t| > t_{\text{кр.}}$, то се отхвърля основната хипотеза H_0 и се приема, че X и Y са корелационно зависими. Ако $|t| < t_{\text{кр.}}$, то се приема, че X и Y са корелационно независими.

Пример 1. Дадени са данни от две извадки x и y, представители на две нормално разпределени ГС. Да се изследва има ли зависимост между тях и да се оцени значимостта на извадковия корелационен коефициент със степен на доверие α= 5%. При наличие на зависимост да се построи уравнението на линейна регресия по метода на най-малките квадрати.

x	5	4	7	10	11	8	9	11	8	9
y	30	27	38	48	59	54	42	63	52	47

Решение: Построяваме таблицата, изчисляваме сумите и заместваме във формулите за r.

i	x_i	y_i	x_i^2	y_i^2	x_i*y_i
1	5	30	25	900	150
2	4	27	16	729	108
3	7	38	49	1444	266

4	10	48	100	2304	480
5	11	59	121	3481	649
6	8	54	64	2916	432
7	9	42	81	1764	378
8	11	63	121	3969	693
9	8	52	64	2704	416
10	9	47	81	2209	423
Суми	82	460	722	22420	3995

По формулите: $n = 10$,

$$SSx = \sum x_i^2 - \frac{(\sum x_i)^2}{n} = 722 - \frac{82.82}{10} = 49{,}6$$

$$SSy = \sum y_i^2 - \frac{(\sum y_i)^2}{n} = 22420 - \frac{460.460}{10} = 1260$$

,

$$SSxy = \sum x_i \, y_i - \frac{(\sum x_i)(\sum y_i)}{n} = 3995 - \frac{82.460}{10} = 223,$$

$$r = \frac{SSxy}{\sqrt{SSx}\sqrt{SSy}} = \frac{223}{\sqrt{49,6}\sqrt{1260}} = 0,892029.$$

Тъй като |r|>0,85 заключаваме, че има корелационна зависимост, но това не е достатъчно. Необходимо е да направим проверка за значимостта на получения извадков корелационен коефициент *r*. При предположение за нормално разпределени ГС: $X \sim N(m_1,\sigma_1)$, $Y \sim N(m_2,\sigma_2)$ пресмятаме *t* по формулите:

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0,892029}{\sqrt{\frac{1-0,892029^2}{8}}} \approx 5,5822.$$

Така имаме: *r* = 0,892029, t = 5,5822

За да направим оценката на хипотезите от таблицата на t разпределението на Стюдънт при даденото ниво на значимост α= 5%, за *n*-2=8 степени на свобода и порядък на квантила $q = 1 - \frac{\alpha}{2}$ =0,975, намираме $\boldsymbol{t_{кр.} = 2{,}306}$.

Тъй като $t > t_{кр.}$, заключаваме, че съществува корелационна зависимост между *x* и *y*.

Регресия с метода на най-малките квадрати

Ако с t-статистиката се установи наличието на зависимост, тя може да се определи явно като формула $y = \hat{y}(x)$ с различни методи за приближаване на функции. Един от най-използваните е методът на най-малките квадрати (МНМК). Това се нарича регресия, а графиката на получената зависимост – регресионна крива.

Линейна регресия. Когато графиката на данните прилича на отсечка, приближаваща функция се търси като полином от първа степен във вида $\hat{y}(x) = a_0 + a_1 x$. Коефициентите a_0, a_1 се определят от линейната система

$$\left| \begin{aligned} n a_0 \quad + \sum_{i=1}^{n} x_i a_1 &= \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i a_0 + \sum_{i=1}^{n} x_i^2 a_1 &= \sum_{i=1}^{n} x_i y_i \end{aligned} \right. .$$

<u>Квадратична регресия</u>. Търси се приближаваща функция като полином от втора степен:

$\hat{y}(x) = a_0 + a_1 x + a_2 x^2$. Коефициентите a_0, a_1, a_2 се намират от линейната система

$$\begin{cases} n a_0 + \sum_{i=1}^{n} x_i a_1 + \sum_{i=1}^{n} x_i^2 a_2 = \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i a_0 + \sum_{i=1}^{n} x_i^2 a_1 + \sum_{i=1}^{n} x_i^3 a_2 = \sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} x_i^2 a_0 + \sum_{i=1}^{n} x_i^3 a_1 + \sum_{i=1}^{n} x_i^4 a_2 = \sum_{i=1}^{n} x_i^2 y_i \end{cases}.$$

Формулата за грешката на приближенията по МНМК е в средноквадратичен смисъл: $R = \sqrt{\sum_{i=1}^{n} (y_i - \hat{y}(x_i))^2}$.

<u>Пример 2.</u> По данните от Пример 1, (Корелация) да се определи уравнението на:

а) линейната регресия, б) квадратична регресия.

Решение:

а) Като използваме намерените суми

$\sum_{i=1}^{n} x_i = 82, \quad \sum_{i=1}^{n} x_i^2 = 722, \quad \sum_{i=1}^{n} y_i = 460, \quad \sum_{i=1}^{n} x_i y_i = 3995$ и заместим в системата за линейна регресия, получаваме системата:

$$\left| \begin{array}{l} 10a_0 \quad + \quad 82a_1 = 460 \\ 82a_0 + 722a_1 = 3996 \end{array} \right. .$$

Можем да решим системата например по метода на Крамер.

Пресмятаме детерминантите

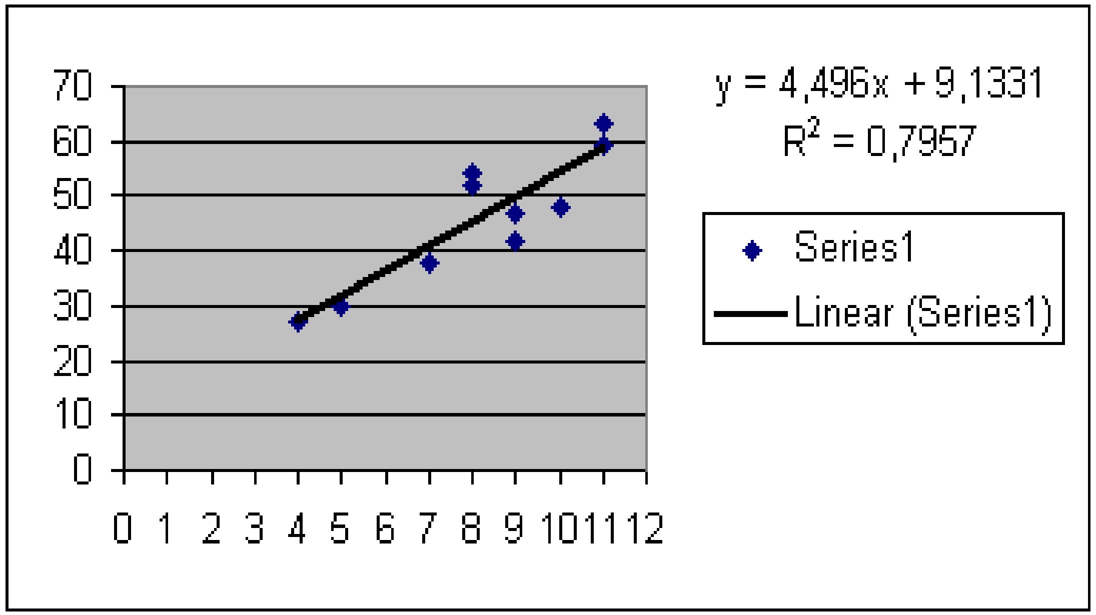
$\Delta = \begin{vmatrix} 10 & 82 \\ 82 & 722 \end{vmatrix} = 10.722 - 82.82 = 496$, $\quad \Delta 1 = \begin{vmatrix} 460 & 82 \\ 3995 & 722 \end{vmatrix} = 460.722 - 82.3995 = 4530$,

$$\Delta 2 = \begin{vmatrix} 10 & 460 \\ 82 & 3995 \end{vmatrix} = 10.3995 - 82.460 = 2230 .$$

Тогава $a_0 = \dfrac{\Delta_1}{\Delta} = \dfrac{4530}{496} = 9{,}1331$, $a_1 = \dfrac{\Delta_2}{\Delta} = \dfrac{2230}{496} = 4{,}4960$.

Уравнението на линейна регресия е: $\hat{y}(x) = 9{,}1331 + 4{,}4960x$.

На графиката е построена линейната зависимост по МНМК за данните от примера.



Б) За да намерим квадратичната крива на регресия пресмятаме още три стълба към таблицата:

	x_i^3	x_i^4	$x_i^2 y_i$
	125	625	750
	64	256	432
	343	2401	1862
	1000	10000	4800
	1331	14641	7139
	512	4096	3456
	729	6561	3402
	1331	14641	7623
	512	4096	3328
	729	6561	3807
суми	6676	63878	36599

Съставяме системата:

$$\begin{cases} 10a_0 + 82a_1 + 722a_2 = 460 \\ 82a_0 + 722a_1 + 6676a_2 = 3996 \\ 722a_0 + 6676a_1 + 63878a_2 = 36599 \end{cases}$$

Нейното решение е: a_0=6,8924, a_1=5,1434 и a_2=-0,0425. Следователно уравнението на приближаващата крива е:

$$\hat{y}(x) = 6{,}8924 + 5{,}1434x - 0{,}0425x^2 .$$

Многомерна линейна регресия

Разгледаните случаи на регресия са частен случай на линейна регресия с една зависима и една независима проеменлива. Линейността се разлежда спрямо коефициентите на регресия, по-горе означени с $a_0, a_1, a_2 \ldots$.

Накратко ще приведем необходимите теоретични постановки за многомерен РА.

Задачата е:

Нека са известни измервания (данни) за някаква променлива величина *y*, зависеща от *p* на брой независими величини $x_1, x_2, \ldots, x_p$. Всяка променлива величина $y, x_1, x_2, \ldots, x_p$ има по *n* на брой измерени стойности (наблюдения). Търсим линейна зависимост от вида

$$y = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_p x_p , \qquad (1)$$

където $b_0, b_1, b_2 ..., b_p$ са неизвестните коефициенти на регресия. Най-стандартният метод за тяхното определяне е методът на най-малките квадрати, рагледан в лекция 6 и в примерите по-горе. Получените стойности на коефициентите $b_0, b_1, b_2 ..., b_p$ в (1) се наричат оценки на регресията, а y - предсказана стойност или оценка на зависимата изходна променлива *y*.

Разликите $\varepsilon_i = y_i - y_i, \quad i = 1, 2, ..., n$ са грешките (остатъци, отклонения, резидиуми) на регресията.

За прилагането както на едномерен, така и многомерен РА се изисква изпълнението на следните основни условия:

- Изходните данни трябва да имат случаен характер, да бъдат от интервален или относителен тип. Категорийните данни трябва да бъдат предварително трансформирани в двоични.
- Наблюденията трябва да бъдат независими.

- Зависимата променлива трябва да има нормално разпределение или близко до нормалното.
- При линейна регресия се предполага, че общата зависимост има линеен или близък до линеен характер.

Съществуват голям брой статистически тестове за валидност на регресионния модел. Когато те са изпълнени, регресионният модел се счита за статистически значим и пригоден за практическо използване и прогнози. Ще приведем само най-основните изисквания за валидност на РА (при ниво 95%):

- Нивото на значимост на целия модел, оценяван с F статистика (в таблица ANOVA) трябва да бъде < 0.05
- Коефициентът на детерминация R^2 и коефициентът на многомерна корелация R (по абсолютна стойност) трябва да са близки до 1 и желателно да са по-големи от 0.5

- Изследването за мултиколинеарност между независимите променливи да е приемливо (за SPSS тестът е VIF < 10)
- Да се направи анализ на остатъците от регресията: разпределението на остатъците да е нормално (напр. с хистограма, P-P, Q-Q тест, тест на Колмогоров-Смирнов, тест на Шапиро-Уилкс), разстоянията на Махалонобис и Кук да са малки, да се изследват отдалечените точки (outliers) и др.
- Препоръчва се и провеждане на процедура на крос-валидация. Тя се състои в разделяне на дадената извадка на две непресичащи се подмножества, наречени тренировъчно и тестово. С тренировъчното множество се построява и изследва нов модел. С този модел се предсказват стойностите на зависимата променлива за наблюденията от тестовото множество. Ако резултатите са близки до модела с тренировъчното подможество и с изходния модел, се приема че изходния модел е валиден.

Подробно решени примери с SPSS са разгледани по-нататък в учебния курс, както и на упражнения.

10. Факторен анализ

Описание на метода

Приложението на факторния анализ (ФА) за изследване на данни е отдавна наложен и широко използван статистически подход в психологията, социалната сфера, икономиката, естествените науки и други области.

ФА е статистическа техника, предназначена за преобразуване на множество от корелиращи данни в ново множество с некорелиралищи изкуствени променливи или фактори, които обясняват възможно по-голяма част от общата вариация на изходните данни. С тази техника се постига редуциране на броя на началните променливи, чрез групирането на тези, които корелират помежду си в общ фактор и разделянето на некорелиращите в различни фактори. Математически това се постига с намаляване размерността на първоначалното пространство чрез установяване на нов базис от променливи (фактори). Факторите могат да бъдат ортогонални или наклонени.

Нека $x_1, x_2, ..., x_p$ са началните променливи, всяка с n наблюдения. Означаваме с $\mathbf{X} \in \mathbb{R}^{p \times n}$ съответната матрица от данни. Във ФА тази матрица се моделира като линейна комбинация на k ($k<p$) на брой n-мерни фактори, $F_1, F_2, ..., F_k$ плюс грешки $E_1, E_2, ..., E_p$:

$$\mathbf{X} = \mathbf{LF} + \mathbf{E}, \tag{1}$$

където $\mathbf{F} \in \mathbb{R}^{k \times n}$ е матрицата на факторните стойности, $\mathbf{L} \in \mathbb{R}^{p \times k}$ е матрица на факторните тегла, $\mathbf{E} \in \mathbb{R}^{p \times n}$ е матрицата на грешките.

Основна задача във ФА е определянето на броя на факторите k, който зависи съществено от силата на корелационните зависимости между данните. За по-силно разграничаване на принадлежността на изходните променливи към един или друг фактор се извършва и допълнително преобразуване на факторите с т. нар. въртене на факторите.

ФА позволява за всеки от факторите да се пресметнат факторните му стойности (factor scores), с което факторите се получават като нови латентни променливи. Тези променливи могат да бъдат използвани в последващи статистически анализи вместо първоначалните променливи. Факторните променливи са в стандартизиран вид със z – стойности, със средно нула и стандартно отклонение единица.

Разработени са голям брой методи за извличане на факторите: метод на главните елементи, метод на най-малките квадрати, алфа-факторинг и др. За въртене на факторите се използва най-често т.н. варимакс (Varimax) метод, но има и други методи – Quartimax, Equamax, както и такива, при които факторите са наклонени, напр. методите Oblimin, Promax и др.

Ще отбележим по-специално, че най-разпространения тип ФА е т.н. изследователски (exploratory) ФА, който се използва в този труд. Процедурите на изследователския ФА (определяне на броя на факторите, извличане на факторите, метод на въртене и др.) не са достатъчно формализирани и получаваното решение не е единствено. Счита се, че намереният факторен модел е добър, когато има достатъчно ясна практическа интерпретация в рамките на разглежданата предметна област.

На ФА е посветена огромна по количество литература, вкл. с използване на специализиран софтуер.

Основни изисквания за приложение на ФА

За получаване на адекватен факторен модел се изисква:

- Изходните данни да имат случаен характер.
- Да бъдат от интервален или относителен тип. Категорийни и номинални данни не се допускат. Наблюденията трябва да бъдат независими.
- Препоръчваният брой наблюдения е поне 50 (виж и по-долу).

- Във ФА участват корелиращи помежду си променливи. Променлива, която не корелира с останалите трябва да бъде предварително изключена от ФА (случай на unique variable). При последващи анализи такава променлива се използва заедно с получените фактори вместо изходните предиктори.
- В идеалния случай данните е добре да имат многомерно нормално разпределение. Последното се проверя с т.н. условие за адекватност на факторния модел с предварителна проверка на данните с теста на Kaizer-Mayer-Olkin или (KMO тест), чиято стойност трябва да е >0.5, както и установяване на сферичност на облака от данни с теста на Bartlett. Може да се проверяват и стойностите на MSA (Measure of Sampling Adequacy) за всяка променлива, които също е препоръчително да са над 0.5.

ФА по метода на главните елементи (МГЕ)

Ще разгледаме само най-разпространения метод на извличане на факторите по МГЕ (PCA – Principal Component Analysis).

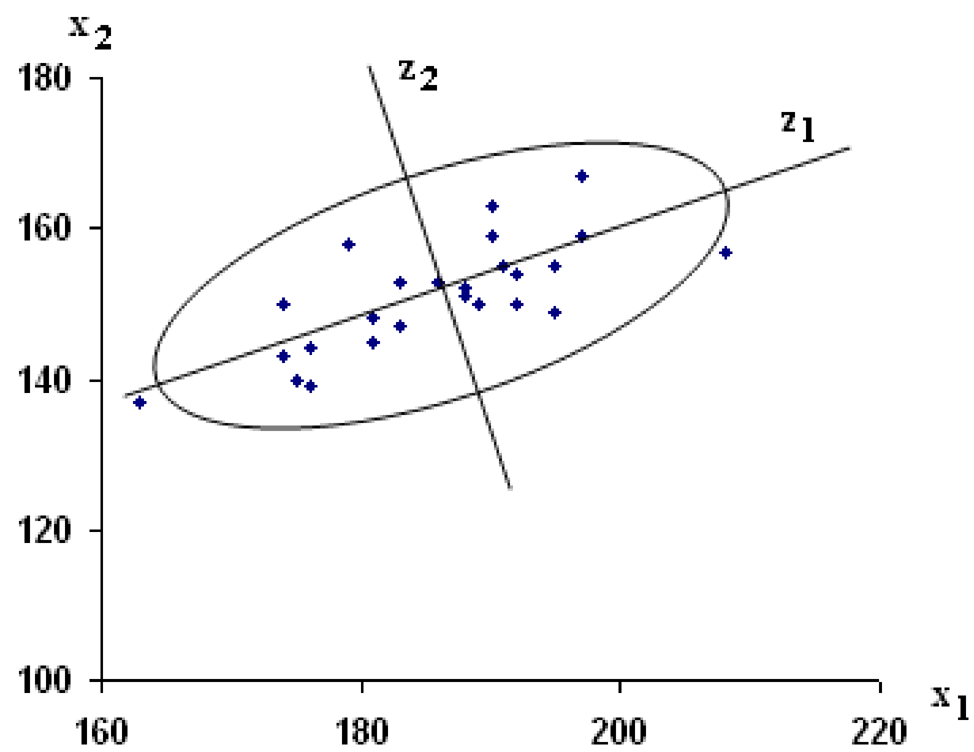
В този случай по МГЕ най-напред се търси корелационната матрица

$$\mathbf{R} = \begin{pmatrix} 1 & r(x_1, x_2) & \dots & r(x_1, x_p) \\ r(x_2, x_1) & 1 & \dots & r(x_2, x_p) \\ \dots & \dots & \dots & \dots \\ r(x_p, x_1) & r(x_p, x_2) & \dots & 1 \end{pmatrix}, \qquad (2)$$

където $r(x_i, x_j)$ е корелационният коефициент за променливите x_i, x_j.

Сумата от собствените стойности на тази матрица е точно равна на броя на участващите променливи p. Относителната стойност на всяко собствено значение λ_α изразява доколко съответният й фактор F_α участва в обясняването на общата вариация на началните данни. Във ФА броят на факторите най-често се избира да бъде равен на броя на собствените стойности на корелационната матрица, които са по-големи от единица, което е известно като правило на Кайзер. Но има примери, от които се вижда, че фактори, съответстващи на собствени стойности доста по-малки от единица също могат да оказват съществено влияние върху модела. Това налага при недостатъчно добри резултати от модела да се включват и допълнителни фактори, не отговарящи на условието на Кайзер.

На Фиг. 10.1 е даден пример, илюстриращ метода на главните компоненти, с който изходната ортогонална система на две променливи x_1, x_2 е трансформирана в нова ортогонална система z_1, z_2, относно която данните имат по-добро представяне. В частност се вижда, че координатните оси се завъртат и се минимизира общата сума на квадратите на разстоянията от точките до осите в новата координатна система спрямо същото за старата координатна система.



Фиг. 10.1. Трансформация на променливите x_1, x_2 към z_1, z_2 с метода на главните компоненти.

При определен фиксиран брой на факторите, те се извличат по метода на главните компоненти с разлагането (1) и се получава начално решение. След това се извършва процедура на въртене и окончателно се намират факторите като “завъртяно решение”.

Основен момент във ФА е получаването на матрицата на факторните тегла $\mathbf{L}^{rot} = (l_{ij})$ (factor loadings) на завъртяното решение, наречена ротационна матрица. С факторните тегла се изразява връзката между факторите и изходните величини: Те са коефициентите на регресия на изходните величини върху групата от фактори, което се изразява от приближените равенства (във вид на линейни комбинации)

$$\left| \begin{array}{l} F_1 \approx l_{11}x_1 + l_{12}x_2 + \ldots + l_{1p}x_p \\ F_2 \approx l_{21}x_1 + l_{22}x_2 + \ldots + l_{2p}x_p \\ \ldots \\ F_k \approx l_{k1}x_1 + l_{k2}x_2 + \ldots + l_{kp}x_p \end{array} \right. . \qquad (3)$$

Променлива, умножена по факторно тегло, чиято абсолютна стойност е голяма (напр. над 0.5) може да бъде групирана към съответния фактор, а когато абсолютната стойност на теглото е под избраната граница, участието на променливата се пренебрегва. Успешният факторен модел изисква дадена променлива да се групира по този начин само към един фактор.

Зависимост между размера на извадката и факторното тегло

Няма установено правило за това, какъв трябва да бъде размерът на извадката, за да се провежда ФА. Битува мнението, че за всяка променлива трябва да има поне 10 или дори 20 наблюдения. Например, при 8 променливи по това правило извадката трябва да е с обем поне 80.

Друг основен момент е съотношението между обема на извадката и големината на факторното тегло, за да се приеме, че факторното тегло е значимо и променливи с факторни тегла над посоченото се групират със съответния фактор. Обикновено се приема, че границата за значимост е 0.5. По-прецизни резултати при ниво на значимост $\alpha = 0.05$ са приведени в табл. 1. От таблицата следва, че например при извадка с размер n=100, само променливи с факторни тегла над 0.55 трябва да се считат за статистически значими за съответния фактор.

Табл. 10.1. Зависимост между обема на извадката и нивото на значимост на факторното тегло.

Размер на извадката, необходим за значимост (ниво 0.05)	**Факторните тегла трябва да са не по-малки от**
350	0.30
250	0.35
200	0.40
150	0.50
100	0.55
85	0.60
70	0.65
60	0.70
50	0.75

Алгоритъм на процедурата на ФА

- Нормализиране на всички променливи чрез трансформация към *z*-стойности
- Изчисляване на корелационната матрица и съответните статистики
- Проверка за адекватност на ФА
- Извличане на факторите по на метода на главните елементи или друга техника за изчисляване на комуналите (натрупваните вариации) и разпределение на общата вариация
- Избор на броя на факторите
- Получаване на начално факторно решение
- Въртене на факторите и получаване на ротационната матрица
- Преценка за правилно групиране на променливите по фактори
- Изчисляване и запомняне на факторните стойности за по-нататъшни статистически анализи
- Интерпретация на факторите и резултатите от ФА.

Валидация на резултатите от ФА

Макар ФА да не е добре формализирана техника, отделните му етапи, описани по-горе трябва да удовлетворяват известен брой тестове и изисквания, които служат за валидация на получавания факторен модел. Основните от тях са:

- Използваните променливи да корелират помежду си и със зависимите променливи, т.е. бивариантните корелационни коефициенти трябва да са високи. В противен случай ФА не се препоръчва
- Статистическите тестове за адекватност на ФА за дадената извадка да са изпълнени: КМО тестът трябва да е >0.5, Бартлет тестът за сферичност, включващ и χ^2 теста за многомерно нормално разпределение трябва да е статистически значим, т.е. Sig. < 0.05.
- Валидиране на избора на броя на факторите по някакъв метод, напр. с изчисляване на репродуцираните корелационни остатъци. За целта

по пресметнатите фактори и факторни тегла се получават приближенията на променливите по формула (1), изчислява се корелационната им матрица, наречена репродуцирана корелационна матрица и тя се сравнява с първоначалната корелационна матрица. Резидиумите трябва да са достатъчно малки, в рамките на 0.05.

- Групирането във факторите в ротационната матрица трябва да бъде коректно, т.е. дадена променлива може да участва само в един фактор (да корелира силно с него), а с другите фактори да има слаба корелация (виж Табл. 10.1).

Кога да използваме ФА

ФА дава възможност за построяване на модели на базата на експерименталните данни, когато прякото прилагане на регресионни техники е невъзможно или силно затруднено.

С ФА се решават следните видове задачи:

- Класификация на изследваните независими величини чрез групирането им във фактори на базата на взаимната им корелация
- Отхвърляне на несъстоятелните входни величини, т.е. тези, които нямат реално влияние върху изходните променливи
- Получаване на адекватни на данните мултифакторни модели, които обясняват в много голям процент (желателно над 70%) изменчивостта на приближаваните данни
- Получаване на факторни променливи, които са взаимно независими и подходящи за по-нататъшни статистически обработки от типа на параметрични и непараметрични регресионни анализи.

Факторен анализ с SPSS

Видяхме, че ФА е статистическа техника за преобразуване на множество от m корелиращи променливи в по-малък брой множество от k некорелиращи променливи (фактори), които описват възможно по-голяма част от изменчивостта на началните данни.
ФА може да се прилага за класификации, за генериране на хипотези относно причинно-следствени връзки или да подготви данните за следващи статистически обработки.

ФА се използва широко в социологията, икономиката, психологията, физиката, инженерните науки и др. области.

Като цяло. ФА се базира на редуцираната корелационна матрица на изходните данни. След получаване на матрицата **F** (извличане на факторните променливи – начално решение), се извършва т.н. въртене на факторите. Според метода на въртене полученото завъртяно решение представлява графично ортогонални или наклонени направления в общото количество от данни.

Ще отбележим, че основната цел на ФА е да може получените фактори да бъдат подходящо интерпретирани и да съответстват на смисъла на изследваните данни, като ги групират по подходящ начин. Математически ФА не е строго формализирана процедура и резултатите от него са приложими именно според това дали резултатите имат реална интерпретация.

SPSS статистики за ФА

ФА изисква да се изчисляват:

- Корелационната матрица на променливите и техните нива на значимост (Significance levels, Sig.)
- Детерминантата на корелационната матрица
- Евентуално: обратна и репродуцирана корелационна матрица, огледална матрица (inverse, reproduced correlation matrix, anti-image)
- KMO (Kaiser-Meyer-Olkin) тест за адекватност и Bartlett's test of sphericity
- Незавъртяно начално решение (unrotated solution, communalities, and eigenvalues – собствени стойности)
- Завъртяно решение с факторните тегла (rotated solution with factor loadings)
- Факторни променливи (factor scores)

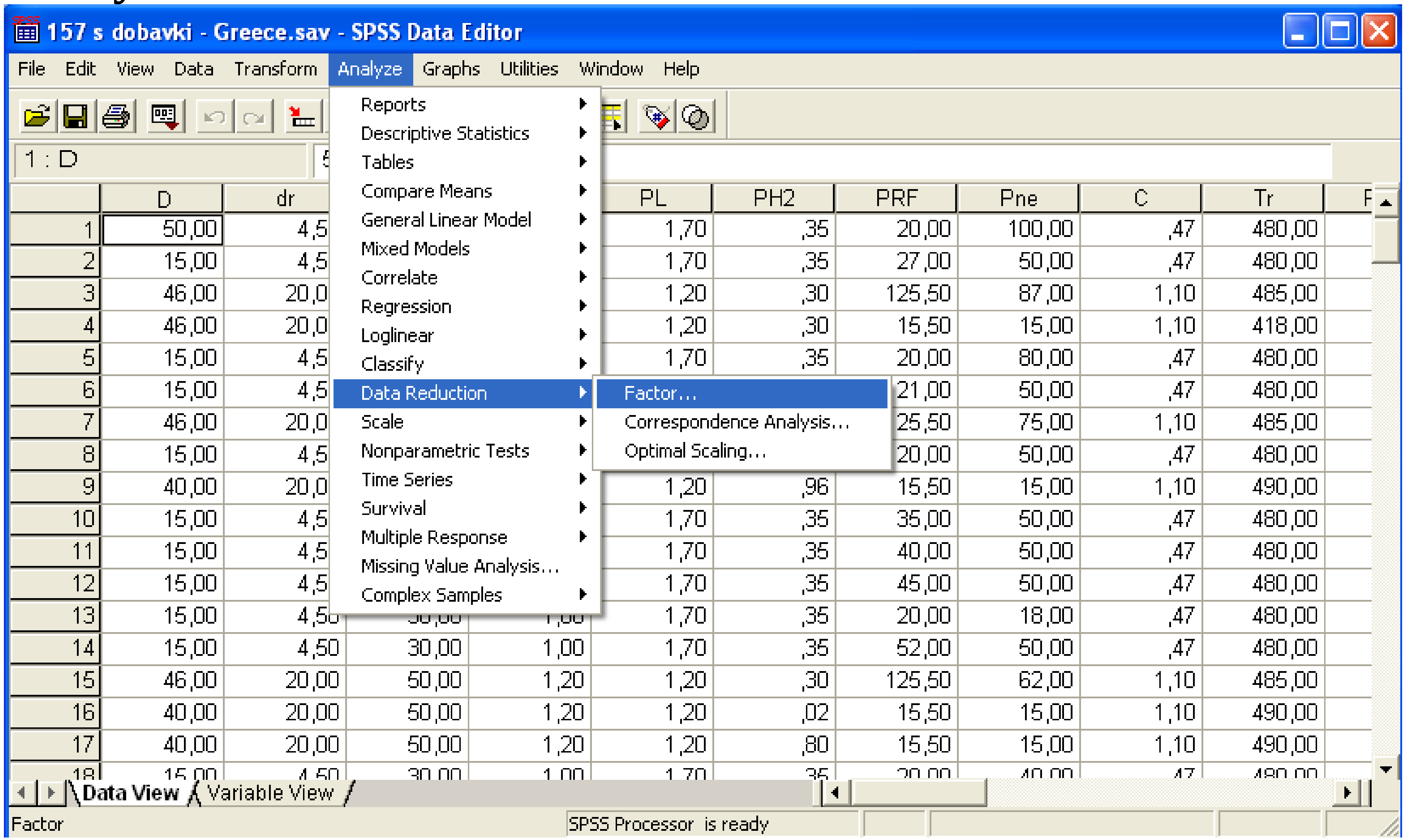
Основни процедури на ФА с SPSS

- Изчисляване на корелационната матрица
- Проверка на тестовете за адекватност на ФА
- Прилагане на Метод на главните елементи (PCA -Principal Component Analysis) (PCA) и (или) друга техника за изчисляване на натрупаните вариации и обща вариация на данните по фактори
- Избор на броя фактори
- Извличане на факторите
- Въртене на факторите и получаване на факторните тегла
- Изчисляване и съхраняване на факторните променливи за следващи анализи

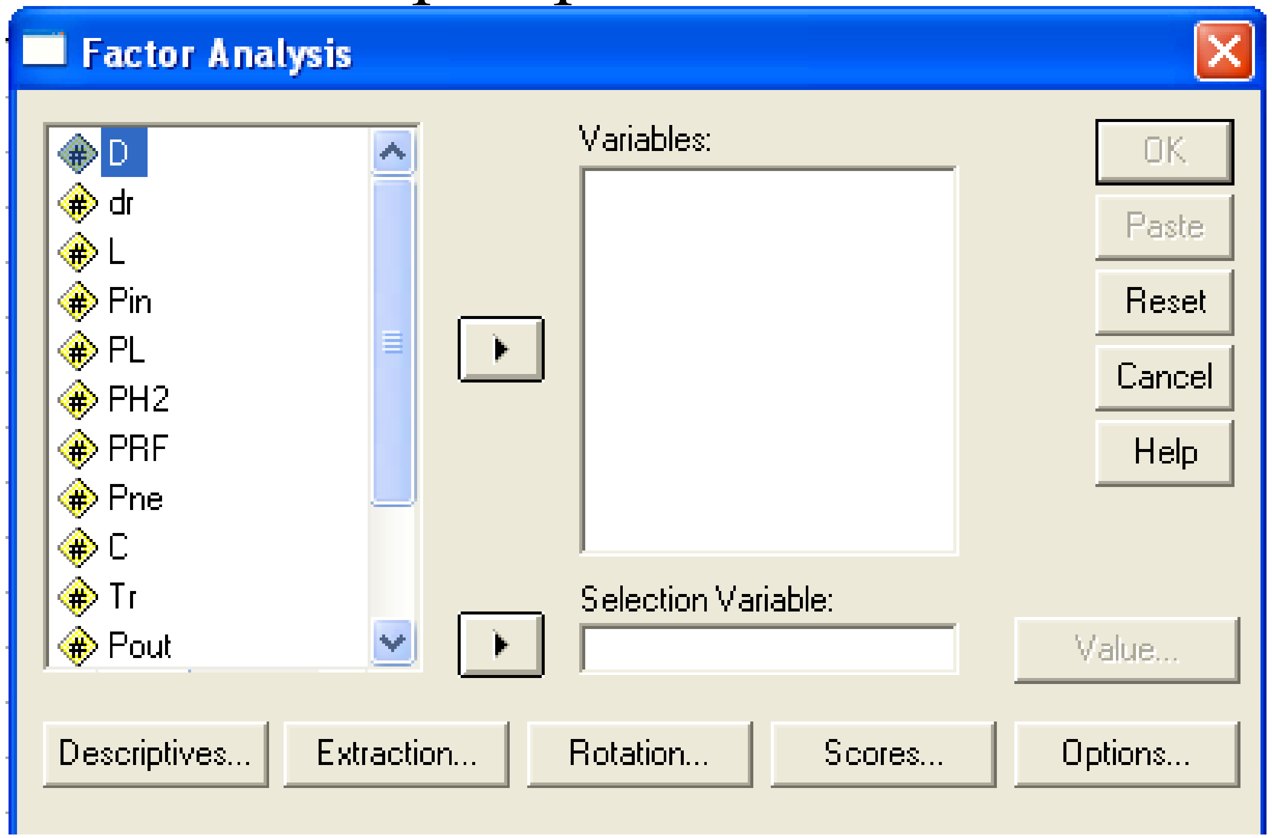
Провеждане на ФА с SPSS

1) **Въвеждаме данните.** Те автоматично се преобразуват в стандартизирани (обезмерени) *z*-променливи (със средна =0 и стандартно отклонение = 1).
2) **От главното меню на SPSS избираме:**

Analyze/Data reduction/Factor



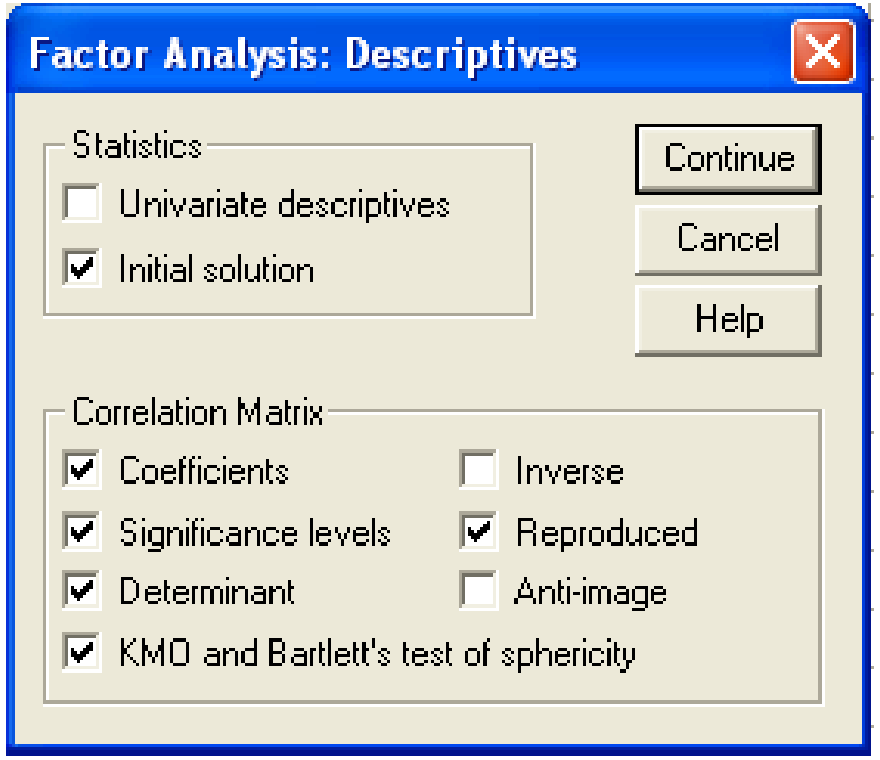
Появява се прозорецът на ФА:



3) Преместваме желаните променливи отляво вдясно в областта Variables. В началния стадий може да включим всички променливи, вкл. зависимите, за да проверим силата на зависимост между тях. При следващи анализи зависимите променливи и т.н. единични променливи се пропускат.

4) Избираме опции от менюто 

+ Continue.



Изчислява се корелационната матрица със всички двойки корелационни коефициенти между избраните променливи и останалите статистики, заявени в менюто, вкл. начално полученото решение за факторните променливи (Initial solution).

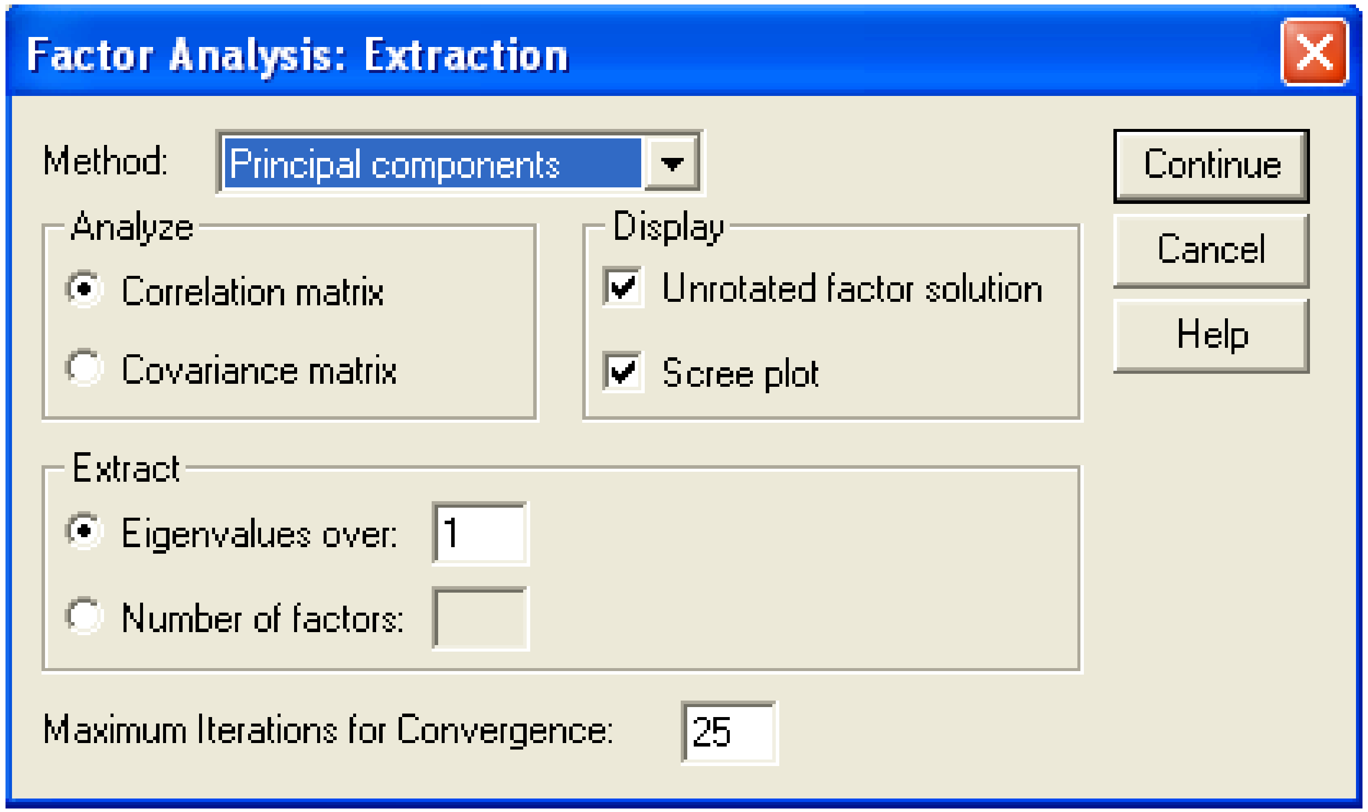
Целта ни е да определим кои променливи корелират силно една с друга и техните корелации са статистически значими (т.е. корелационните коефициенти да са поне >0,3 и нагоре, със съответен коефициент на значимост Sig. <0,05). Ако има променлива, която корелира силно със зависимите променливи, но не корелира с останалите независими променливи, тя се нарича единична променлива. Всички единични променливи се отстраняват по-нататък от ФА и се добавят при друг тип анализи (регресионен, диспресионен и др.).

ФА е адекватен (смислен), когато едновременно:

- Има големи корелационни коефициенти.
- KMO тестът за адекватност на извадката е >0,5.
- Bartlett's test за сферичност на облака данни има Sig. <0,05.

5) От прозореца Factor Analysis отваряме 

. Избираме метода на извличане на факторите и кои резултати да покажем. Натискаме Continue.



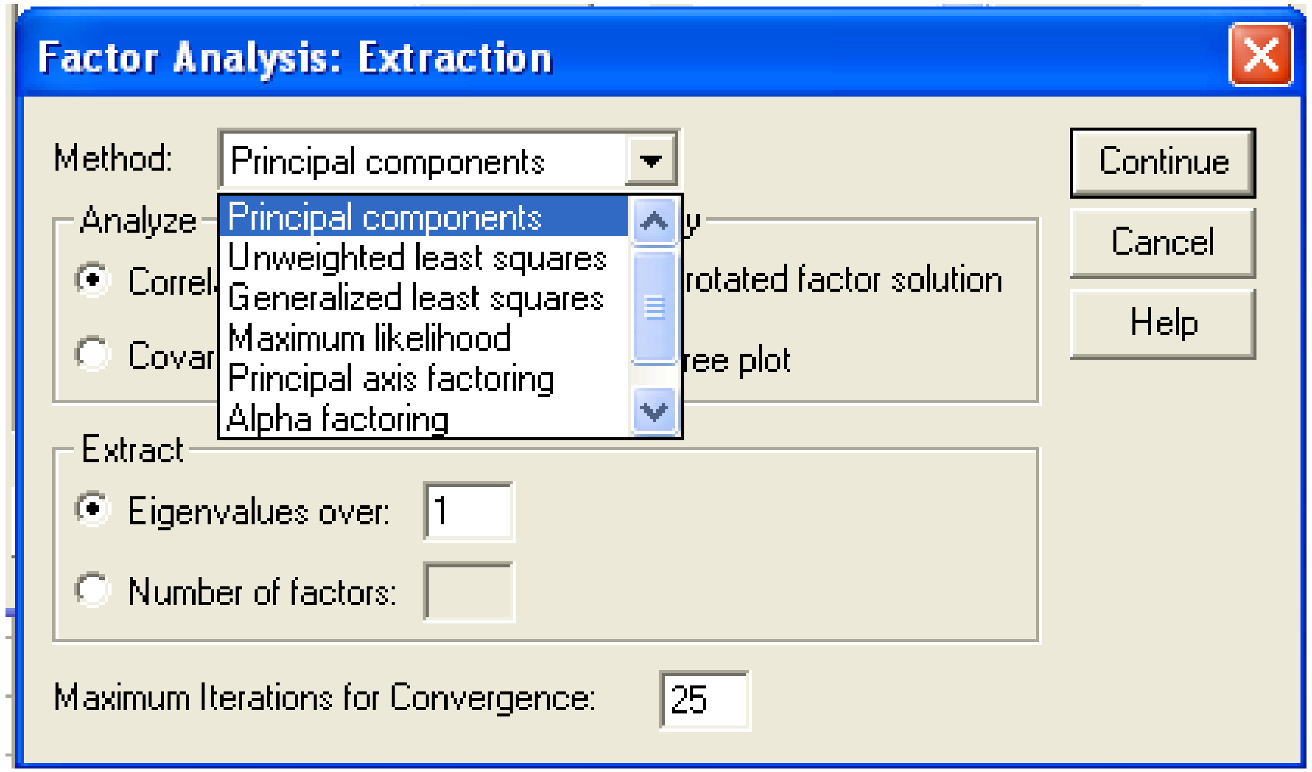
SPSS предлага 7 различни метода за извличане на фактори:

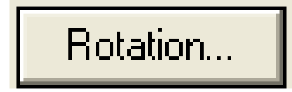
- ♦ Principal component analysis (по подразбиране)- метод на главните елементи
- ♦ Unweighted least squares (обикновен метод на най-малките квадрати)
- ♦ Generalized least squares (обобщен МНМК)
- ♦ Maximum likelihood
- ♦ Principal axis factoring
- ♦ Alpha factoring
- ♦ Image factoring

Стандартно се работи с Principal components. Това е метод, който преобразува всичките n променливи в n нови променливи, според тяхното относително тегло в общата вариация на облака данни.

Следователно, сумата от техните вариации по този метод е винаги равна на 1.

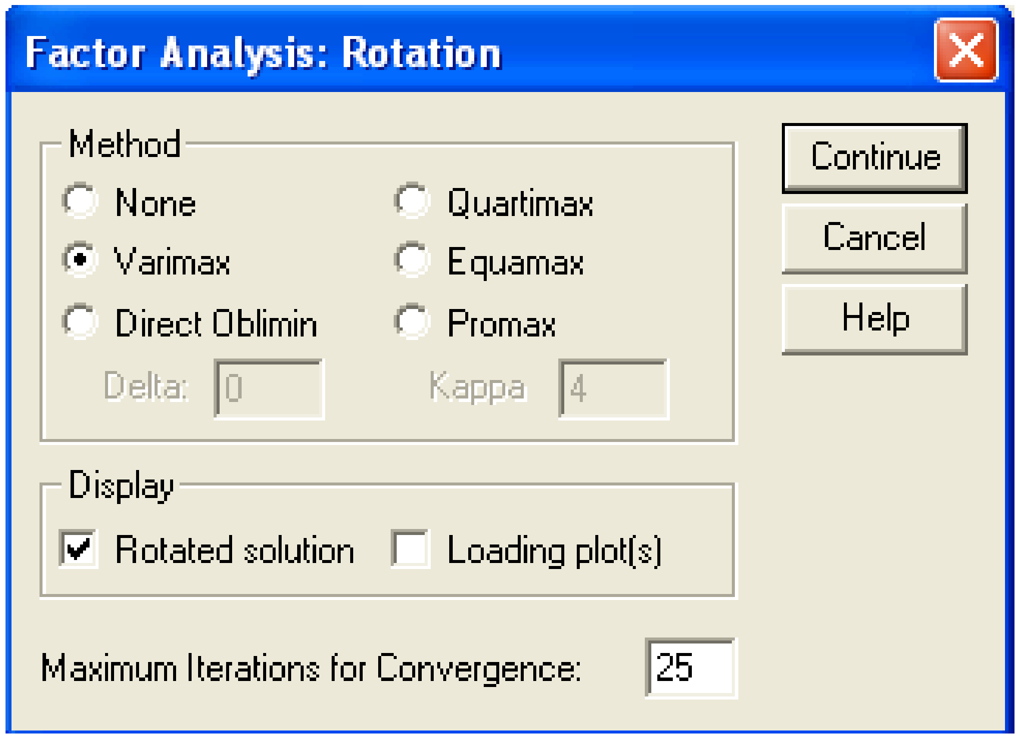
Методът изчислява абсолютните стойности на т.н. собствени стойности на корелационната матрица. За ФА обикновено се избират тези фактори, които имат с.ст. >= 1, но ако моделът не е достатъчно добър, се избират и с по-малки с.ст.



6) Метод на въртене – от бутона 

и Continue.

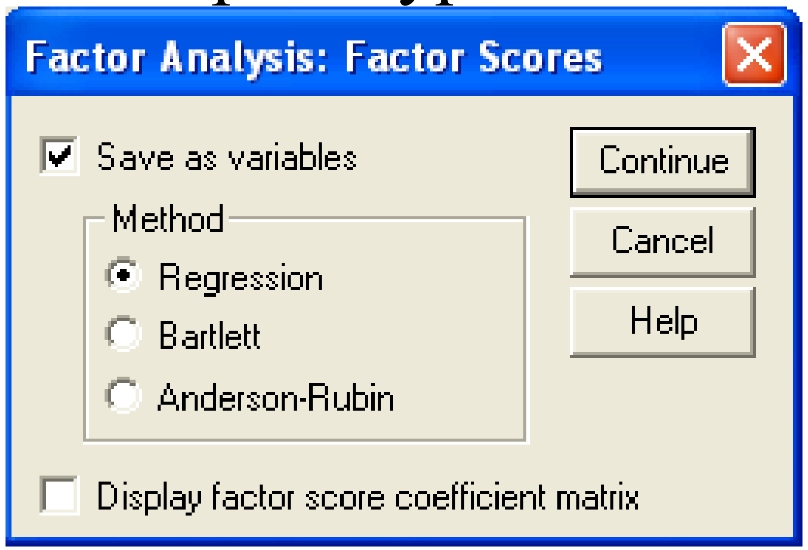
В SPSS има 6 метода на въртене. Най- стандартният е Varimax метод. Математически въртенето е получаване на нов базис с взаимноортогонални или наклонени променливи.



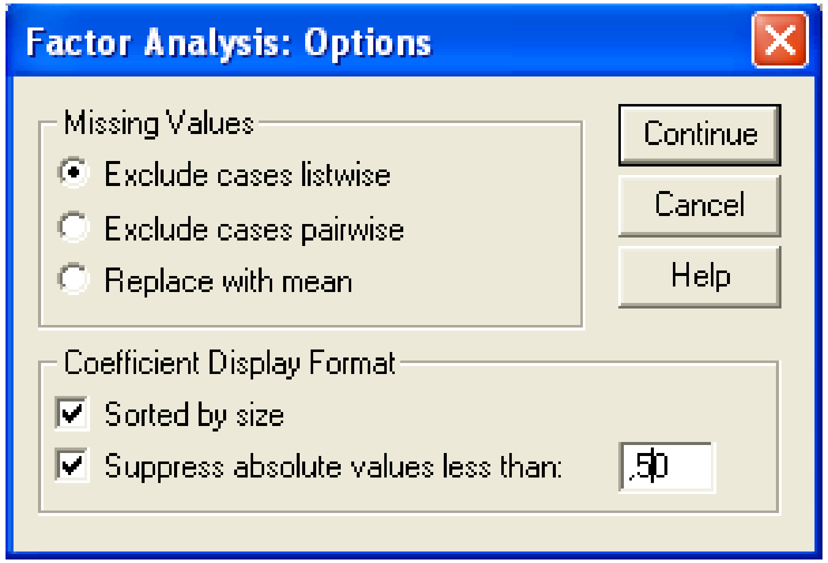
7) Изчисляване и запомняне на факторните стойности (променливи) - 

.

Тази процедура се използва, когато са успешни предишните дотук.



8) В прозореца Factor analysis/Options можем да изберем метод за обработване на липсващи данни и формата на компонентите на ФА.



9) Накрая в основния прозорец Factor analysis избираме OK, за да стартираме анализа

Резултатите се появяват в отделен прозорец и могат да се съхранят при желание като отделен файл. Факторните променливи се добавят към изходните данни.

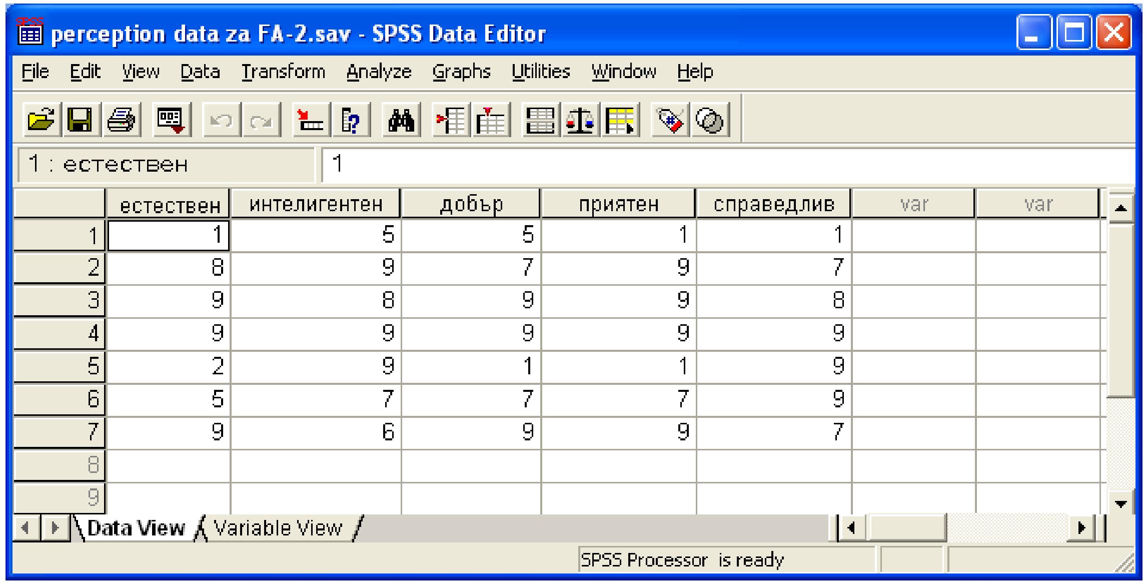
Пример 10.1 за ФА

Данни: Момиче на 12 години отговаря по 9 точкова възходяща скала (от 1 до 9) за възприятията си от 7 свои познати. Класирането е по следните 5 описания: “естествен”, “интелигентен”, “добър”, “приятен” и “справедлив”. Да се направи групиране на данните с ФА.

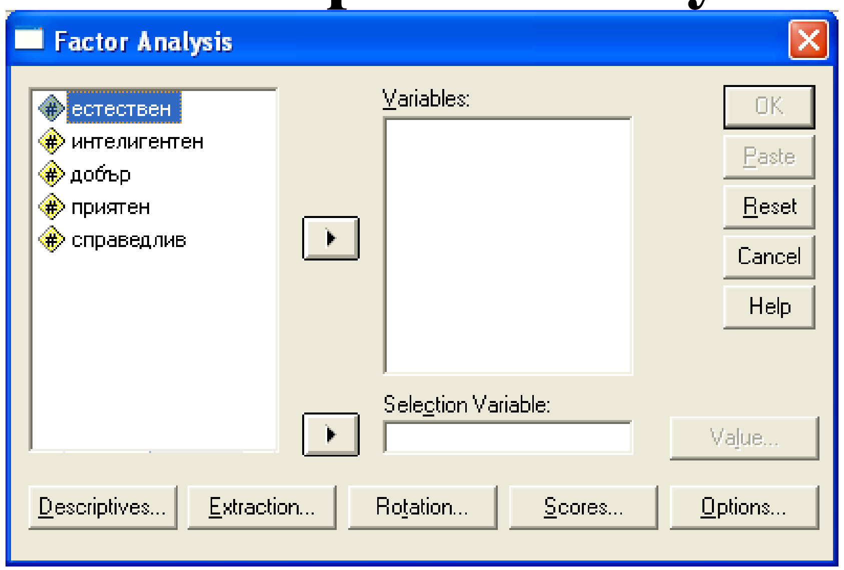
Табл. 10.2:

	естествен	**интелигентен**	**добър**	**приятен**	**справедлив**
съученичка 1	**1**	**5**	**5**	**1**	**1**
сестра	**8**	**9**	**7**	**9**	**7**
съученичка 2	**9**	**8**	**9**	**9**	**8**
баща	**9**	**9**	**9**	**9**	**9**
учител	**2**	**9**	**1**	**1**	**9**
съученик	**5**	**7**	**7**	**7**	**9**
съученичка 3	**9**	**6**	**9**	**9**	**7**

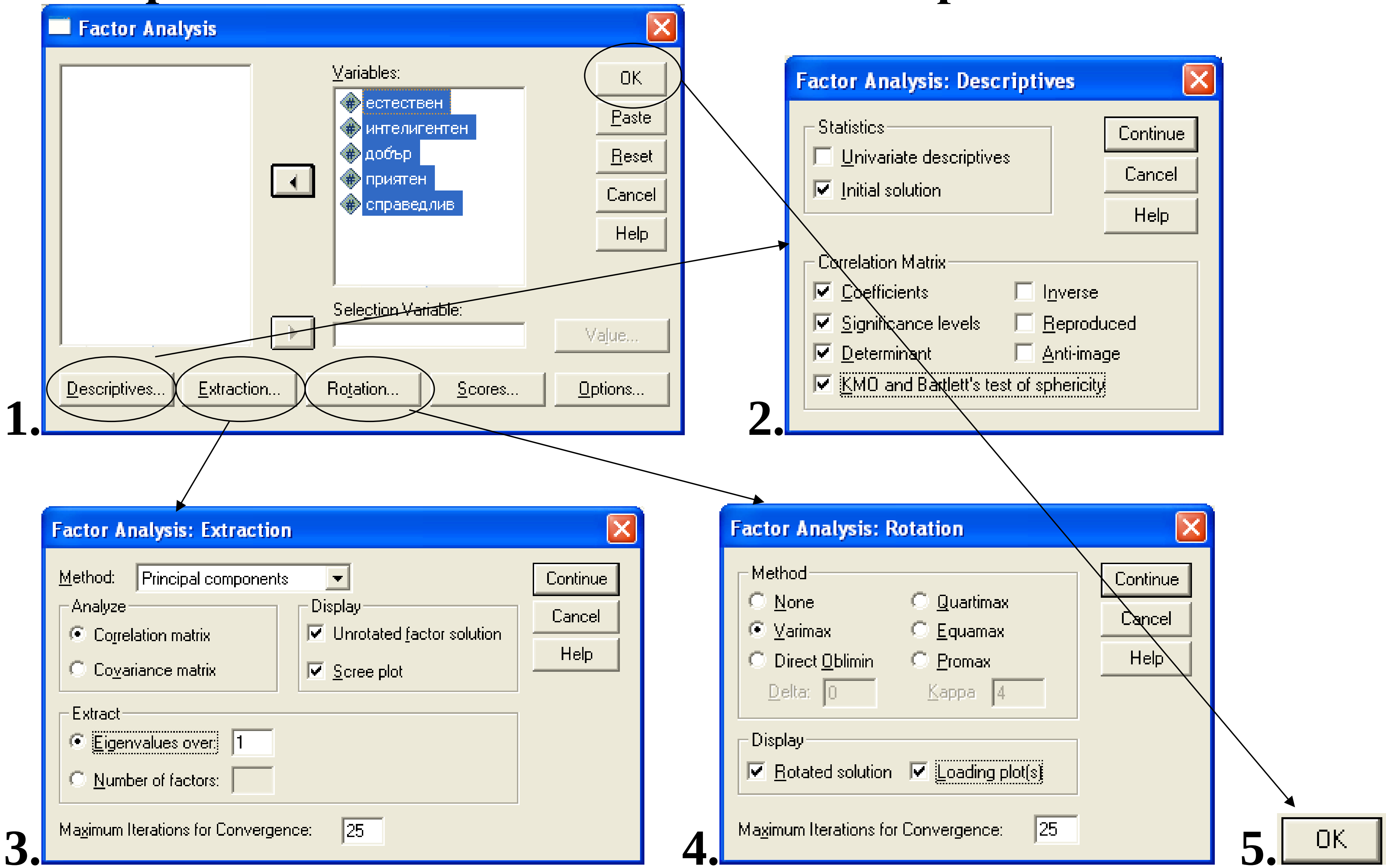
П.1. Въвеждаме данните:



П.2. Избираме: Analyze/Data reduction/Factor



П.3. Пренасяме в полето Variables и отваряме подменютата:



Резултати:
а) корелационна матрица

Correlation Matrix[a]

		естествен	интелигентен	добър	приятен	справедлив
Correlation	естествен	1,000	,338	,854	,969	,484
	интелигентен	,338	1,000	-,101	,280	,737
	добър	,854	-,101	1,000	,886	,125
	приятен	,969	,280	,886	1,000	,472
	справедлив	,484	,737	,125	,472	1,000
Sig. (1-tailed)	естествен		,229	,007	,000	,136
	интелигентен	,229		,415	,271	,029
	добър	,007	,415		,004	,394
	приятен	,000	,271	,004		,142
	справедлив	,136	,029	,394	,142	

a. Determinant = ,001

Матрицата е симетрична. Детерминантата е 0.001, не е 0 и формално ФА може да се проведе. Гледаме само корелационните коефициенти > 0,5. Съответните им нива на значимост от долната половина на таблицата в случая имат нива на значимост Sig. <0,05. Това показва, че тези корелационни зависимости са статистически значими и трябва да участват в анализа. Останалите са незначими. В частност за тази извадка най-голям е корелационният коефициент между “приятен” и “естествен” (0.969) и той е значим.

б) тестове за адекватност на ФА:

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,689
Bartlett's Test of Sphericity	Approx. Chi-Square	23,320
	df	10
	Sig.	,010

Тук KMO=0,689 >0,5, следователно данните са подходящи за ФА и моделът е адекватен. Бартлет тестът има ниво на значимост Sig.=0.010<0.05, което означава, че облакът от данни е сферичен.

в) компоненти от МГЕ (PCA)

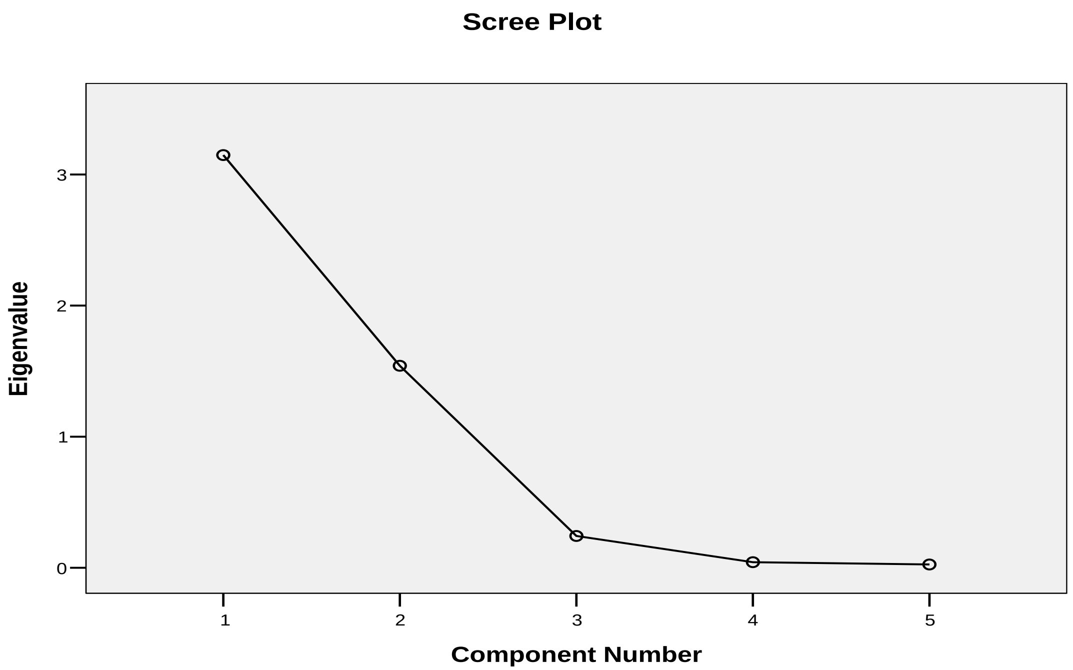
Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3,148	62,966	62,966	3,148	62,966	62,966
2	1,541	30,814	93,780	1,541	30,814	93,780
3	,243	4,860	98,639			
4	,043	,851	99,491			
5	,025	,509	100,000			

Extraction Method: Principal Component Analysis.

Два фактора ще обяснят общо 93,780% от цялата извадка. Третата компонента ще осигури до 98,639% общо. Затова в първи вариант избираме 2 фактора.

г) графика на с.ст.



Тази графика нагледно показва колко фактори е добре да вземем. Спира се там, където вече става полегата, т.е. два или три фактора стигат.

д) незавъртяно, начално решение за 2 фактора

Component Matrix[a]

	Component	
	1	2
естествен	,972	-,155
интелигентен	,461	,827
добър	,803	-,577
приятен	,970	-,204
справедлив	,637	,677

Extraction Method: Principal Component Analysis.

a. 2 components extracted.

Няма ясно разделяне на променливите по фактори (компоненти) . Напр. променливата "справедлив" им тегла 0.637 в първи фактор и 0.677 във втори. Така тя не може ясно да се групира в никой от двата фактора. Продължаваме анализа с въртене.

е) завъртяно решение с Варимакс метод

Rotated Component Matrix[a]

	Component	
	1	2
естествен	,931	,321
интелигентен	,017	,947
добър	,980	-,131
приятен	,952	,276
справедлив	,244	,897

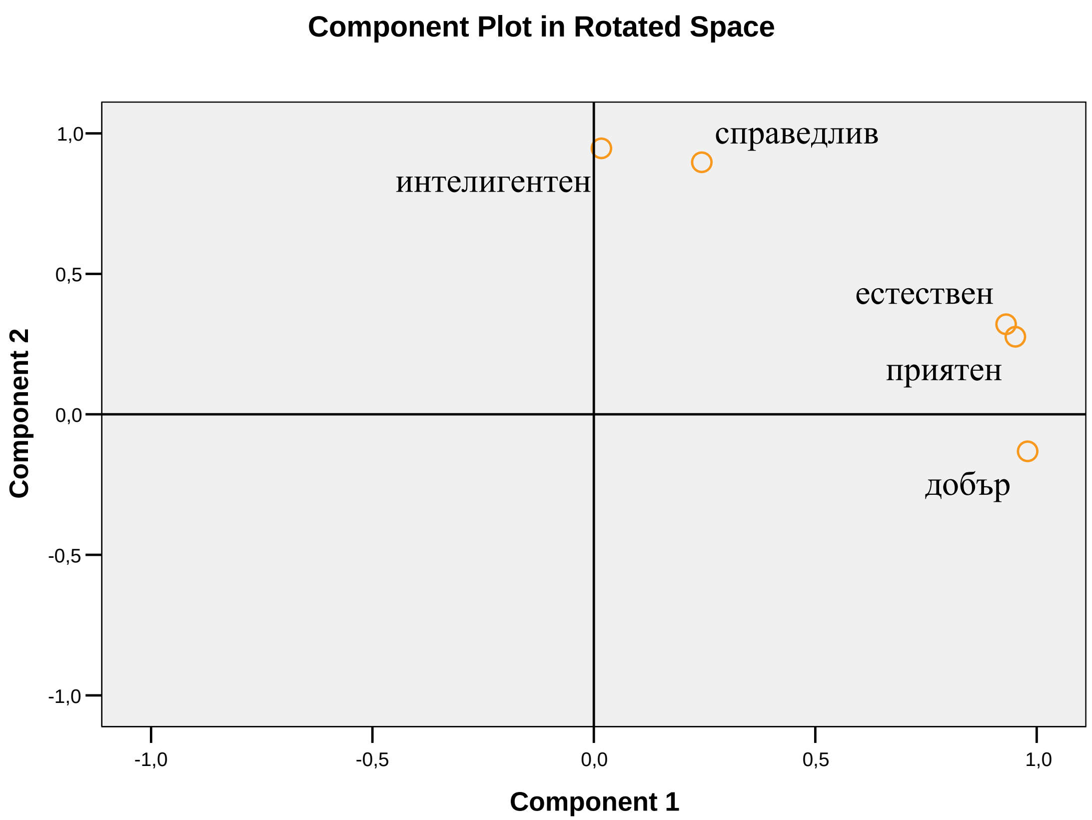
Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

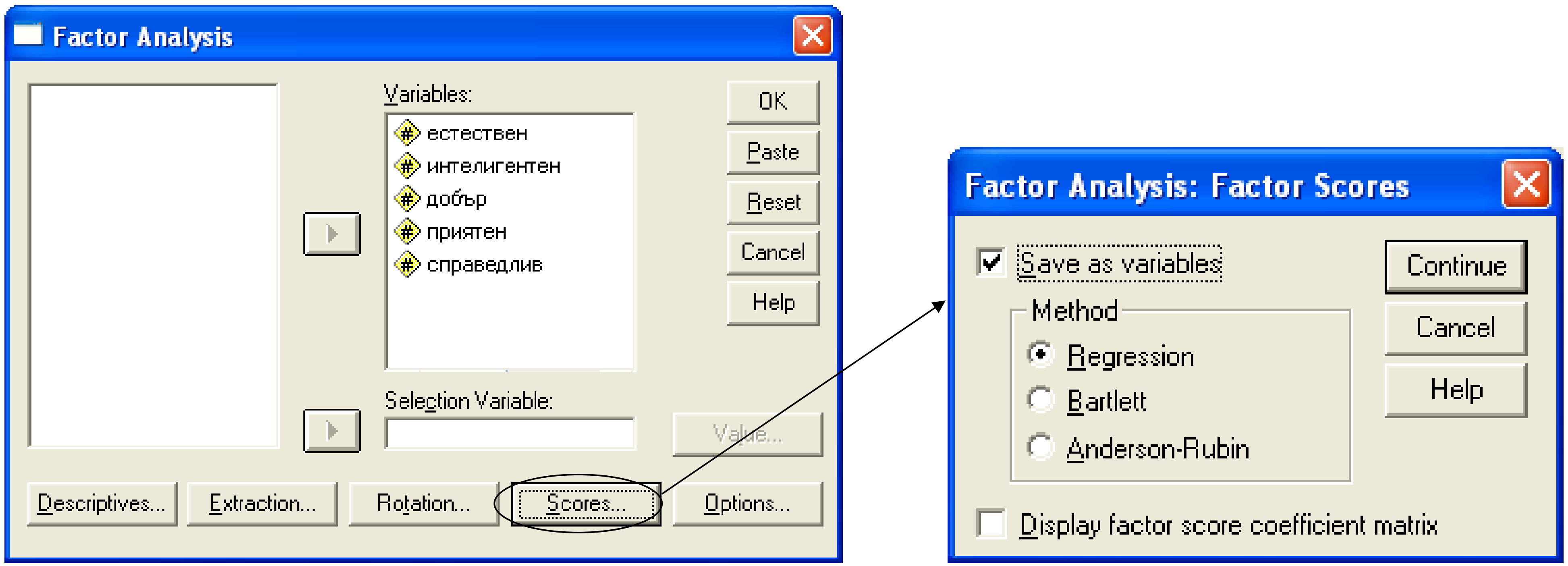
В първата компонента (фактор) ще се групират само променливите с тегла над 0,5 – това са 1, 3 и 4- ред – “естествен”, “добър” и “приятен”. Този първи фактор ще наречем “приятелски”.

Вторият фактор е съставен от групиране на 2 и 5 ред – “интелигентен” и “справедлив”. Можем да го наречем “възвишен”.

ж) графика на завъртяното решение в новото двумерно пространство от 2 фактора:

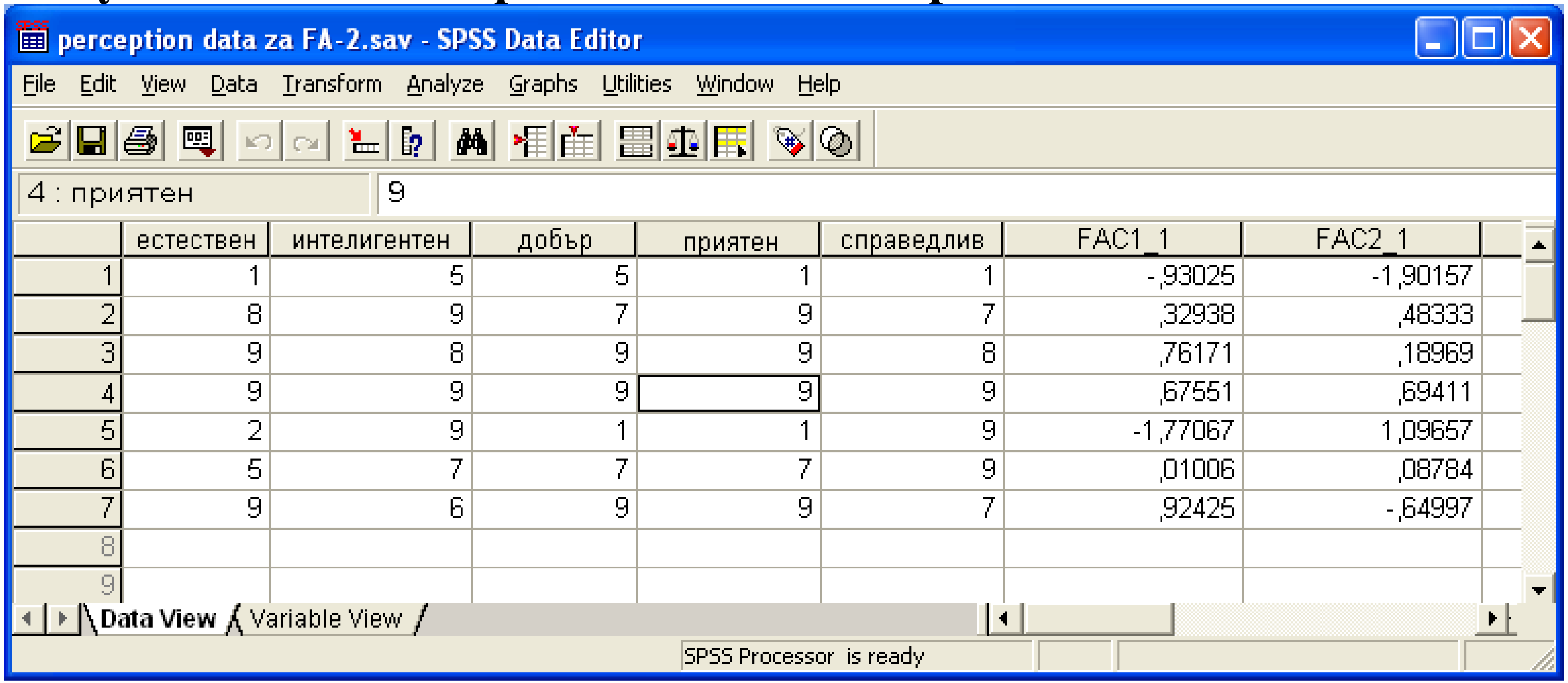


П.5. Запомняне на факторните променливи:
От менюто Factor Analysis избираме бутона Scores…



и потвърждаваме с Continue и OK.

Получаваме 2 нови променливи към първоначалните 5:



	естествен	интелигентен	добър	приятен	справедлив	FAC1_1	FAC2_1
1	1	5	5	1	1	-,93025	-1,90157
2	8	9	7	9	7	,32938	,48333
3	9	8	9	9	8	,76171	,18969
4	9	9	9	9	9	,67551	,69411
5	2	9	1	1	9	-1,77067	1,09657
6	5	7	7	7	9	,01006	,08784
7	9	6	9	9	7	,92425	-,64997
8							
9							

Те описват нашите данни 93,78%, така че могат да се използват вместо петте променливи. При това са почти ортогонални и не корелират помежду си. Проверете последното твърдение.

Задача. Направете ФА за следните данни. Проверете адекватност на модела.

Данни: За 30 търговски марки японски вина “Сейшу” се изследва връзката между променливите:

y1 – вкус,

y2 – мирис,

и променливите:

x1 – pH,

x2 – киселинност 1,

x3 – киселинност 3,

x4 – съдържание на оризова ракия,

x5 – горена захар,

x6 – общо количество захар,

x7 – алкохол,

x8 – формил-азот

Табл. 10.2. Сейшу измервания

y1	*y2*	*x1*	*x2*	*x3*	*x4*	*x5*	*x6*	*x7*	*x8*
1,0	,8	4,05	1,68	,85	3,0	3,97	5,00	16,90	122,0
,1	,2	3,81	1,39	,30	,6	3,62	4,52	15,80	62,0
,5	,0	4,20	1,63	,92	-2,3	3,48	4,46	15,80	139,0
,7	,7	4,35	1,43	,97	-1,6	3,45	3,98	15,40	150,0
-,1	-1,1	4,35	1,53	,87	-2,0	3,67	4,22	15,40	138,0
,4	,5	4,05	1,84	,95	-2,5	3,61	5,00	16,78	123,0
,2	-,3	4,20	1,61	1,09	-1,7	3,25	4,15	15,81	172,0
,3	-,1	4,32	1,43	,93	-5,0	4,16	5,45	16,78	144,0
,7	,4	4,21	1,74	,95	-1,5	3,40	4,25	16,62	153,0
,5	-,1	4,17	1,72	,92	-1,2	3,62	4,31	16,70	121,0
-,1	,1	4,45	1,78	1,19	-2,0	3,09	3,92	16,50	176,0
,5	-,5	4,45	1,48	,86	-2,0	3,32	4,09	15,40	128,0
,5	,8	4,25	1,53	,83	-3,0	3,48	4,54	15,55	126,0
,6	,2	4,25	1,49	,86	2,0	3,13	3,45	15,60	128,0
,0	-,5	4,05	1,48	,30	,0	3,67	4,52	15,38	99,0

-,2	-,2	4,22	1,64	,90	-2,2	3,59	4,49	16,37	122,8
,0	-,2	4,10	1,55	,85	1,8	3,02	3,62	15,31	114,0
,2	,2	4,28	1,52	,75	-4,8	3,64	4,93	15,77	125,0
-,1	-,2	4,32	1,54	,83	-2,0	3,17	4,62	16,60	119,0
,6	,1	4,12	1,68	,84	-2,1	3,72	4,83	16,93	111,0
,8	,5	4,30	1,50	,92	-1,5	2,98	3,92	15,10	68,0
,5	,2	4,55	1,50	1,14	,9	2,60	3,45	15,70	197,0
,4	,7	4,15	1,62	,78	-7,0	4,11	5,55	15,50	106,0
,6	-,3	4,15	1,32	,31	,8	3,56	4,42	15,40	49,5
-,7	-,3	4,25	1,77	1,12	,5	2,84	4,15	16,65	164,0
-,2	,0	3,95	1,36	,25	1,0	3,67	4,52	15,98	29,5
,3	-,1	4,35	1,42	,96	-2,5	3,40	4,12	15,30	131,0
,1	,4	4,15	1,17	1,06	-4,5	3,89	5,00	16,79	168,2
,4	,5	4,16	1,61	,91	-2,1	3,93	4,35	15,70	118,0
-,6	-,3	3,85	1,32	,30	,7	3,61	4,29	15,71	48,0

11. Многомерен клъстерен анализ

1) Обща характеристика

Клъстерният анализ (КА) е общо название на множество изчислителни процедури, използвани при създаването на класификации на обекти. Той позволява да се открият нови зависимости и свойства в дадено множество от данни, които не биха могли да се установят с други известни теоретични и експериментални методи.

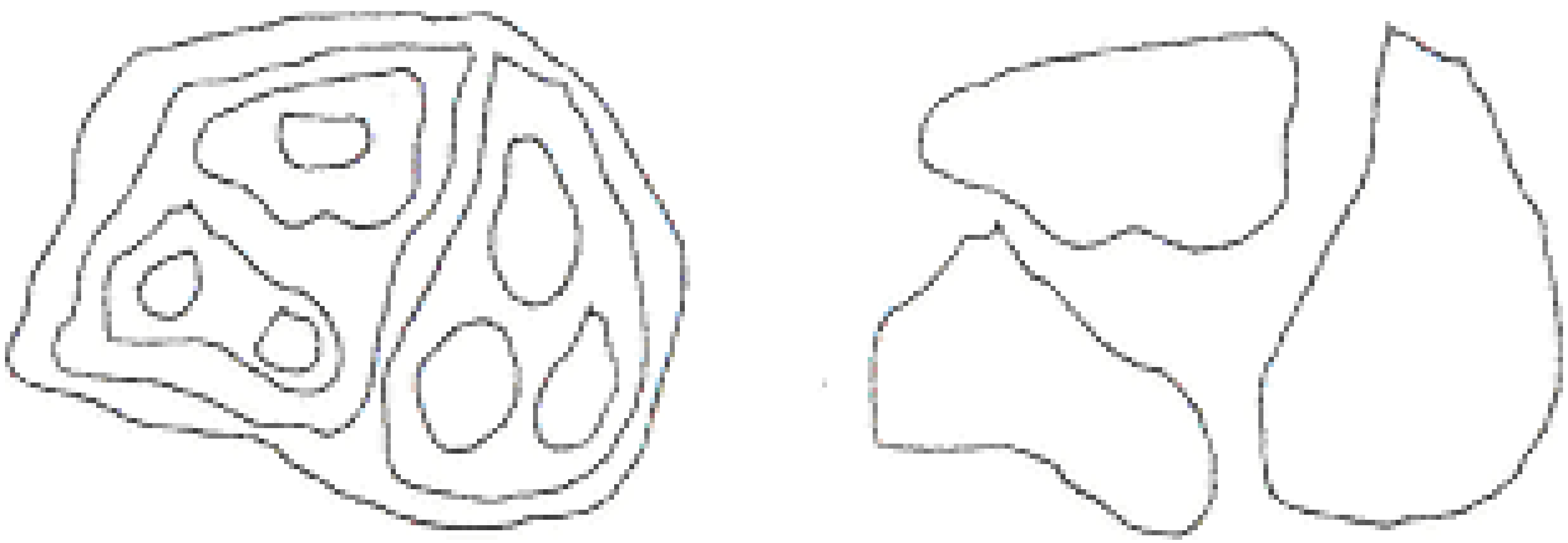
Класифицираните обекти могат да бъдат както наблюдения (случаи), така и променливи. В специализираната статистическа литература не съществува дефиниция на понятието клъстер, която може да претендира за универсалност. Най-общо може да се каже, че клъстерът е подсъвкупност (група) от обекти, които са сходни (близки) в определена степен и в съответствие с определен критерий.

Като редуцираща техника КА има общи елементи с дискриминантния и факторния анализ. Например при дискриминантния анализ също се класифицират единици, но в предварително определени групи. Различен е и познавателният подход при двата анализа. При дискриминантния анализ се изследват различията между известните групи и възможностите за разработване на ефективно правило за класификация на единиците, а при КА се изследват възможностите за формиране на логически обосновани и практически обясними еднородни групи (клъстери) от единици. При ФА се класифицират променливи величини в общи фактори, а не единици.

Целта на КА е да се намери оптимално групиране на наблюденията (или на променливите), при което елементите от даден клъстер са подобни, но клъстерите един с друг ясно се отличават.

Обикновено в КА броят на групите предварително е неизвестен и се определя от изследователя в процеса на изследването.

Съществуват различни видове КА, най-разпространените са: йерархични методи и КА по метода на К-средните (разделящи или нейерархични). В случая на интервални данни и сравнително малък размер на извадката (под 500), по-подходящи са йерархични методи, на които ще се спрем по-подробно.



Фиг. 11.1. а) Йерархично клъстеризиране б) Нейерархично клъстеризиране

Формираните групи (клъстери) трябва да бъдат еднородни (хомогенни) вътре и разнородни (хетерогенни) помежду си по зададени характеристики. Количественото оценяване на понятието „сходство” е свързано с понятието „метрика”. При този подход за сходство събитията се представят като точки в координатното пространство. Установените сходства и различия между точките се определят в съответствие с разстоянията между тях.

Два обекта са идентични, ако разстоянието между тях е нула. Колкото по-голямо е разстоянието между два обекта, толкова по-голямо е несходството между тях.

Нека таблицата от данни

$$\mathbf{X} = (x_{ij}),\ i = 1,\dots,n,\ j = 1,\dots p \qquad (1)$$

се записва във вида:

$$\mathbf{X} = \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix} = \left(x_{(1)}, x_{(2)}, \ldots, x_{(p)} \right) \quad ,$$

където x_i' е векторът, съставен от i -тите съответни наблюдения за всяка променлива $x_{(j)}$, $x_{(j)}$ е вектор-стълб на j-тата променлива. Можем да групираме по редове (по наблюдения) или по стълбове (по променливи).

2) Основни изисквания за приложение на КА

За КА няма специални дискриминиращи изисквания към данните, техните разпределения и взаимозависимости. За сметка на това, не съществува механизъм за различаване на подходящите и неподходящите за изследването променливи. Ако някои съществени променливи бъдат игнорирани, резултатите могат да са неадекватни. Другият проблем е подходящият избор на оптималния брой клъстери, което е една неформализирана процедура.

При правилно определяне на решението, в случая на КА по променливи, много автори считат, че резултатите трябва формално да са много близки до тези от ФА за същите данни.

3) Мерки за сходство и различие в КА

Съществуват голям брой разстояния, за измерване на близостта (сходството), които се използват в КА:

1) Евклидово разстояние

Най-често използваната мярка за разстояние между интервални данни е обикновеното евклидово разстояние между два вектора

$x_i = (x_{i1}, x_{i2}, \ldots, x_{ip})'$ и $x_j = (x_{j1}, x_{j2}, \ldots, x_{jp})'$:

$$d(x_i, x_j) = \sqrt{\sum_{m=1}^{p} (x_{im} - x_{jm})^2}, \quad i, j = 1, \ldots, n. \qquad (2)$$

2) Квадрат на евклидовото разстояние

$$d(x_i, x_j) = \sum_{m=1}^{p} (x_{im} - x_{jm})^2, \quad i, j = 1, \ldots, n. \qquad (3)$$

3) Чебишево разстояние

$$d(x_i, x_j) = \max \left| x_{im} - x_{jm} \right|, \quad i, j = 1, \ldots, n. \qquad (4)$$

4) Разстояние на Минковски от ред *r*

$$d(x_i, x_j) = \left(\sum_{m=1}^{p} \left| x_{im} - x_{jm} \right|^r \right)^{\frac{1}{r}}, \quad i, j = 1, \dots, n. \tag{5}$$

Така, в случай на n наблюдения трябва да пресметнем матрицата на разстоянията $D = d(x_i, x_j)$, която е симетрична и с нулеви елементи по главния диагонал.

Например, нека имаме три обекта (точки в двумерното пространство), с измервания за две променливи (x_1, x_2): (2, 5), (4, 2), (7, 9). Тук $p = 2, n = 3$. Използвайки обикновеното евклидово разстояние пресмятаме разстоянията: $d_{12} = \sqrt{(2-4)^2 + (5-2)^2} = \sqrt{13} \approx 3,6$,
$d_{13} = \sqrt{(2-7)^2 + (5-9)^2} = \sqrt{41} \approx 6,4$ и
$d_{23} = \sqrt{(4-7)^2 + (2-9)^2} = \sqrt{58} \approx 7,6.$

За матрицата на разстоянията получаваме

$$D_1 = d(x_1, x_2) = \begin{pmatrix} 0 & 3,6 & 6,4 \\ 3,6 & 0 & 7,6 \\ 6,4 & 7,6 & 0 \end{pmatrix}.$$

Очевидно е обаче, че разстоянието зависи от мерната единица. Например, ако умножим първата променлива x_1 по 10 (да кажем смяна на метър с дециметри), матрицата ще стане

$$D_2 = d(x_1, x_2) = \begin{pmatrix} 0 & 20,22 & 50,16 \\ 20,22 & 0 & 30,81 \\ 50,16 & 30,81 & 0 \end{pmatrix}.$$

Сега най-голямото разстояние е d_{13}, вместо предишното d_{23}.

Този проблем се решава обикновено чрез предварително стандартизиране (образмеряване) на данните чрез z - трансформация

$$z_i = \frac{x_i - \overline{x}}{s}, \; i = 1, \ldots, n$$

или по друг обезмеряващ начин.

Може да се определят и други видове разстояния. Например, широко известна мярка се явява манхатъновото разстояние или „разстояние на градските квартали” (city-block) и др.

Формиране на клъстерите

След като се избере измерител на подобие (или различие) между обектите и се пресметне матрицата на разстоянията, се преминава към втория етап на КА - образуването на клъстерите. Множеството методи, които са разработени в статистическата литература за тази цел, се

разделят на две групи: методи за йерархична (вертикална) клъстеризация и методи за нейерархична (хоризонтална) клъстеризация.
Методите за йерархична клъстеризация са разделени в две подгрупи: агломеративни и разделящи методи. При агломеративните методи се извършват последователни сливания на единици и клъстери. Започва се с n клъстера, представляващи отделните единици, и след последователни сливания се стига до един клъстер, обединяващ всички единици. При разделящите методи подходът е обратен - започва се с един клъстер, който обединява всички единици, и след последователни разделяния се завършва до *p* клъстера, като всяка единица формира отделен клъстер.

От разработените клъстерни методи най-често се използват йерархичните агломеративни методи.

Най-разпространеният способ за представяне на резултатите на тези методи се явява дендрограмата (дървовидна диаграма), която графически изобразява йерархичната структура, породена от матрицата на сходство и правилата за обединение на клъстерите. Съществуват различни стратегии на обединение на обектите в клъстери и след това на самите клъстери. При образуване на клъстери във вид на “верига” най-често се използват методите на междугруповото свързване (Between-groups linkage) и на най-близкия съсед (Nearest neighbor).

Ще опишем най-разпространените методи за групиране на клъстерите в йерархичните агломеративни методи.

1) Междугрупово свързване

Разстоянието между два клъстера A и B се дефинира като средната стойност на $n_A.n_B$ на брой разстояния между n_A точки от A и n_B точки от B:

$$D(A,B) = \frac{1}{n_A.n_B} \sum_{i=1}^{n_A} \sum_{j=1}^{n_B} d(x_i, x_j), \qquad (6)$$

където сумата се взима по всички x_i от A и всички x_j от B. Тук $d(x_i, x_j)$ е избраното разстояние между векторите x_i и x_j. На всяка стъпка се обединяват двата клъстера с най-малкото разстояние.

2) Вътрешногрупово свързване

Този метод е подобен на междугруповото свързване, но се изчисляват всевъзможните разстояния между всички точки на двата клъстера A и B, т.е. включитено и разстоянията между точките от един и същ клъстер. Формулата е:

$$D(A,B) = \frac{1}{(n_A + n_B).(n_A + n_B - 1)} \sum_{i,j} d(x_i, x_j), \qquad (7)$$

където сумата е върху всички точки x_i и x_j от A и B, а $d(x_i, x_j)$ е избраното разстояние. На всяка стъпка се обединяват двата клъстера с най-малкото разстояние.

3) Най-близкия съсед

При този метод разстоянието между два клъстера A и B е равно на минимума от всички разстояния между точките на A и точките на B:

$$D(A,B) = \min\left\{d(x_i, x_j),\ x_i \in A,\ x_j \in B\right\}. \tag{8}$$

4) Най-далечния съсед

Тук се пресмятат всички разстояния по формулата:

$$D(A,B) = \max\left\{d(x_i, x_j),\ x_i \in A,\ x_j \in B\right\}. \tag{9}$$

Обединяват се в нов клъстер тези два клъстера, чието разстояние е минимално.

5) Центроиден метод

За всеки клъстер *A* се изчислява средният вектор $\bar{y}_A = \left(\sum_{i=1}^{n_A} y_i \right) / n_A$,

където n_A е броят вектори (точки) на *A*. След това се изчисляват всички евклидови разстояния между средните вектори на всеки два клъстера *A* и *B*:

$$D(A, B) = (\bar{y}_A, \bar{y}_B). \qquad (10)$$

Обединяват се двата клъстера с минимално разстояние. Центроидът на новия клъстер *AB* се намира по формулата

$$\bar{y}_{AB} = \frac{n_A \bar{y}_A + n_B \bar{y}_B}{n_A + n_B}. \qquad (11)$$

6) Метод на Уард (Ward's method)

http://sites.stat.psu.edu/~ajw13/stat505/fa06/19_cluster/09_cluster_wards.html

Определяне броя на клъстерите

Това не е добре формализирана процедура. Първият начин е при разчитане на дендрограмата да се построят пресечни линии там, където разстоянието между един формиран клъстер и друг е по-голямо, например по-голям от поне 20% от максималната скала.

Друг начин е полученото решение да се сравни с решението от друг клъстеризационен метод и/или разстояние.

4) Валидация на резултатите от КА

Няма установени количествени оценки за валидация на резултатите. Препоръчват се следните методи:

- проверка за стабилност на избрания клъстерен модел чрез сравняване с резултати от няколко метода
- “крос”-сравнения, което се проверява чрез случайно разделяне на извадката на две непресичащи се подмножества, провеждане на КА за всяко от тях и сравнение на моделите

- проверка за състоятелност на променлива. Когато някоя променлива (или наблюдение при КА на наблюденията) се класифицира в различни клъстери при различни методи, тя следва да се изключи от модела като несъстоятелна или да остане под наблюдение при следващи статистически анализи.

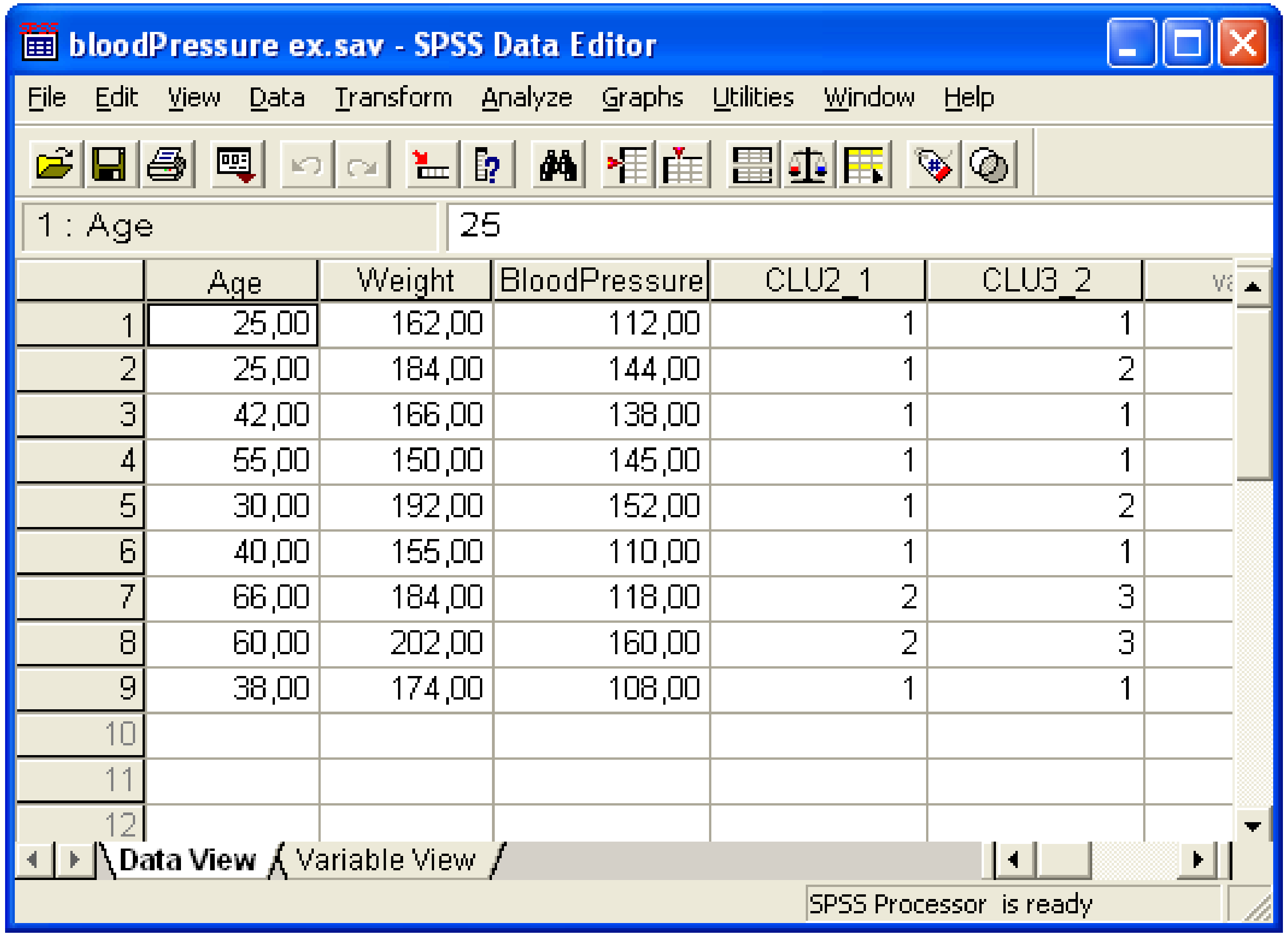
Как се провежда КА със SPSS? ---> Повторете примерите в ЕВЛМ:

http://www.fmi-plovdiv.org/evlm/DBbg/database/studentbook/SPSS_CA_2.pdf
- Клъстерен анализ с SPSS: Клъстерен анализ на K-средните

http://www.fmi-plovdiv.org/evlm/DBbg/database/studentbook/SPSS_CA_3.pdf -

Клъстерен анализ с SPSS: йерархичен клъстерен анализ

Пример 11.1. Данни за възраст, тегло и кръвно налягане на 9 пациента. Да се направи опит с КА за класифициране на пациентите.



	Age	Weight	BloodPressure	CLU2_1	CLU3_2
1	25,00	162,00	112,00	1	1
2	25,00	184,00	144,00	1	2
3	42,00	166,00	138,00	1	1
4	55,00	150,00	145,00	1	1
5	30,00	192,00	152,00	1	2
6	40,00	155,00	110,00	1	1
7	66,00	184,00	118,00	2	3
8	60,00	202,00	160,00	2	3
9	38,00	174,00	108,00	1	1
10					
11					
12					

Провеждаме клъстерен анализ по наблюдения (9 cases). Задължително обезразмеряваме.
На най-първия етап, когато всички наблюдения са единични клъстери таблицата с разстоянията между всеки 2 от тях е следната:

Proximity Matrix

	Squared Euclidean Distance								
Case	1:Case 1	2:Case 2	3:Case 3	4:Case 4	5:Case 5	6:Case 6	7:Case 7	8:Case 8	9:Case 9
1:Case 1	,000	4,148	3,025	7,180	7,061	1,164	9,099	16,415	1,257
2:Case 2	4,148	,000	2,425	7,752	,480	6,638	9,126	7,114	4,325
3:Case 3	3,025	2,425	,000	1,705	3,334	2,379	4,606	6,875	2,537
4:Case 4	7,180	7,752	1,705	,000	8,641	4,149	6,135	9,497	6,591
5:Case 5	7,061	,480	3,334	8,641	,000	9,334	8,837	4,465	6,197
6:Case 6	1,164	6,638	2,379	4,149	9,334	,000	5,892	15,247	1,206
7:Case 7	9,099	9,126	4,606	6,135	8,837	5,892	,000	5,642	4,042
8:Case 8	16,415	7,114	6,875	9,497	4,465	15,247	5,642	,000	11,481
9:Case 9	1,257	4,325	2,537	6,591	6,197	1,206	4,042	11,481	,000

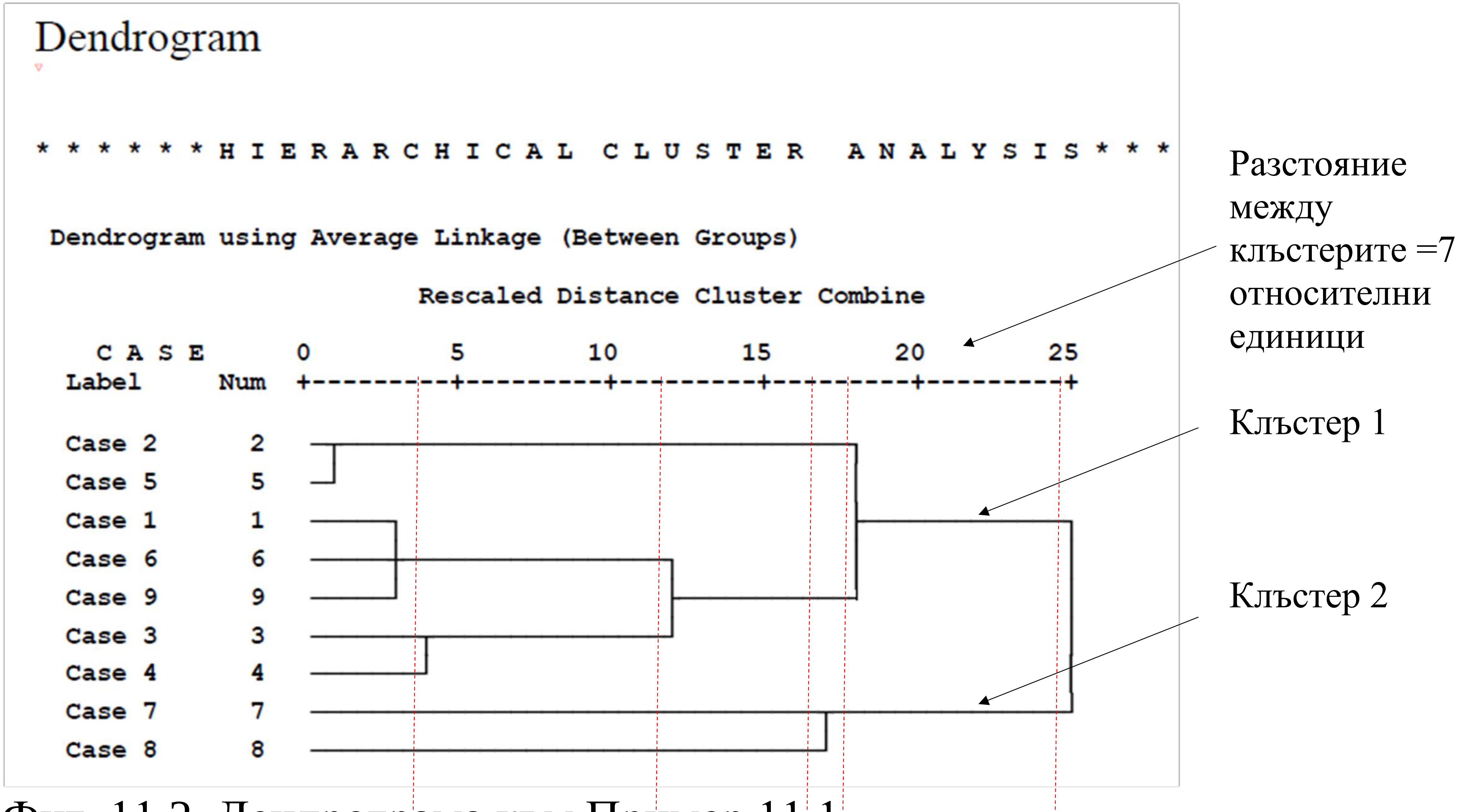
This is a dissimilarity matrix

Най-късото разстояние е между случаи 2 и 5 = 0.480. Те са комбинирани в клъстер с номер 2 (взима се по-малкият номер от 2 и 5). Това е етап 1 (Stage 1) от формирането на клъстерите (виж Agglomeration Schedule).

След това най-близките са 1 и 6 – етап 2, формира се клъстер с номер 1, и т.н. Разстоянието на 7 и 8 етап е по-голямо, отколкото между 6 и 7 етап. Затова е по-правдоподобно да изберем 2 клъстера за крайно решение. Това най-ясно е вижда на дендрограмата от следващата страница.

Agglomeration Schedule

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	2	5	,480	0	0	7
2	1	6	1,164	0	0	3
3	1	9	1,231	2	0	5
4	3	4	1,705	0	0	5
5	1	3	4,310	3	4	7
6	7	8	5,642	0	0	8
7	1	2	5,985	5	1	8
8	1	7	8,488	7	6	0



Фиг. 11.2. Дендрограма към Пример 11.1.

В SPSS стандартно разстоянията са приведени към относителни (Rescaled distance), спрямо най-голямото, което се приема за 25 относителни единици. За улеснение сме построили червени вертикални пунктирани линии на дендрограмата от Фиг. 11.1. Разстоянието между тях е разстоянието между клъстерите в относителни единици. Най-дясната червена вертикална линия пресича 2 хоризонтални – за 2-клъстерното решение. Клъстеризирането в 2 групи е статистически адекватно, тъй като двата клъстера са ясно раздалечени с голямото разстояние от 3-клъстерното решение, с над 5 относителни единици (в случая 7 единици). Изиксват се минимум 5.

Заключение:

Оптимално решение – 2 клъстера. При 2 клъстера групите са:

1ви клъстер – добро здравословно състояние, пациенти 1, 2, 3, 4, 5, 6, 9;

2ри клъстер – лошо здравословно състояние, пациенти 7, 8.

Пример 11.2. Приложение на клъстерен анализ за изследване основните характеристики на лазер с пари на меден бромид на базата на реални експериментални данни.

Решение: Описание на данните

Разглеждаме 7 основни характеристики на CuBr лазер:

Входни характеристики (независими променливи):

D– вътрешен диаметър на лазерната тръба,

dr– вътрешен диаметър на вътрешните пръстени, сложени в тръбата,

L – разстояние между електродите,

Pin – средна входна електрическа мощност,

PL – входна електрическа мощност на единица дължина, с 25% загуби,

PH2 – налягане на водорода от газовата смес.

Изходна характеристика (зависима променлива):

Eff– изходна лазерна ефективност

Използваме следната таблица с данни:

Табл. 11.1. Експериментални данни за CuBr лазер.

D, mm	dr, mm	L, cm	Pin, KW	PH2, Torr	PL, KW/cm	EFF, %
15	4.5	30	1	0.3	1.7	0.7
40	20	50	1.2	0.3	1.2	1.6
50	30	100	2.5	0.55	1.2	2.1
50	30	140	2.3	0.55	0.8	2
40	40	50	1.2	0.6	1.2	2
40	40	120	2.7	0.6	1.1	2.2
58	58	200	3.3	0.5	0.8	3

Ще проведем йерархичен КА по променливи (variables).

Клъстерният модел ще покаже групирането на променливите и структурата в групите. Така експериментаторът ще знае в какво съотношение на характеристиките е най-добре да планира бъдещ експеримент.

1. Набираме данните в SPSS.
2. Стартираме процедурата с: **Analize/ Classify/ Hierarchical cluster**.
3. От основния прозорец внимаваме да маркираме **Variables**, вместо Cases.
4. В прозореца Statistics избираме **Proximity matrix** и **Cluster membership** от 2 до 5.
5. В Plots избираме **Dendrogram**.
6. В Method... избираме: **Between-groups Linkage**, distance: **Squared Euclidean distance** и Standardize: **Z-scores**.
7. В основния прозорец избираме OK.

Получаваме:

Табл. 11.2. Матрица на разстоянията (Proximity matrix) на независимите променливи.

Variable	D	dr	L	Pin	P_L	P_{H2}
D	.000	2.330	2.402	2.845	23.004	5.123
dr	2.330	.000	2.584	2.921	21.417	3.638
L	2.402	2.584	.000	0.653	22.169	6.239
Pin	2.845	2.921	0.653	.000	20.800	5.310
P_L	23.004	21.417	22.169	20.800	.000	18.780
P_{H2}	5.123	3.638	6.239	5.310	18.780	.000

Табл. 11.2 е началното състояние, когато всяка променлива се разглежда като отделен клъстер. Най-малко е разстоянието между L и Pin, което е 0.653. Затова тези променливи ще се групират най-напред в нов клъстер

(виж по-надолу Фиг. 11.2). Най-голямо е разстоянието между PL и D, 23.004. Изобщо PL има най-големи разстояния в сравнение с другите.

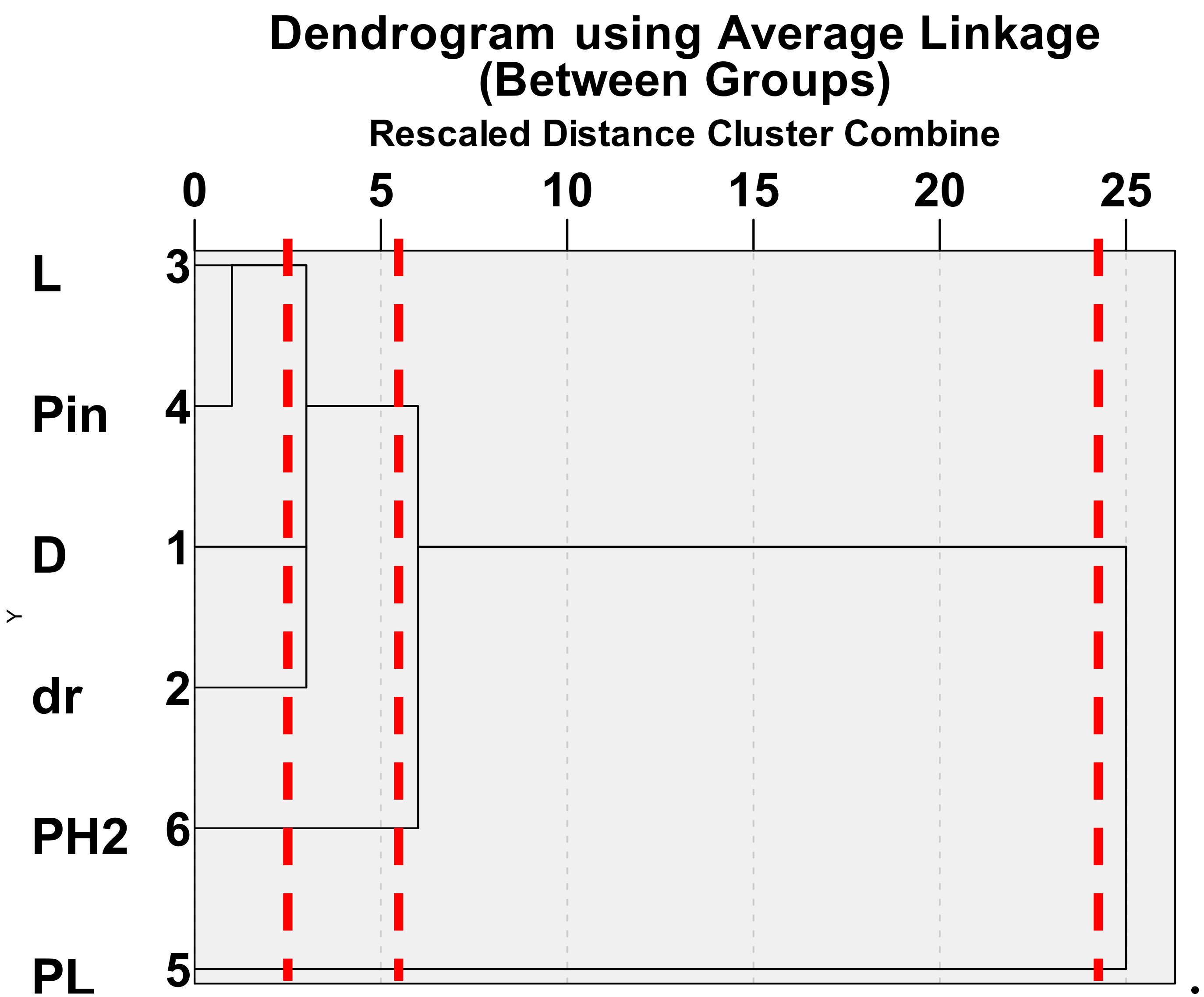
Крайното разпределение по клъстери с избрания метод на междугрупово свързане и квадрат на евклидовото разстояние е дадено на Табл. 11.3.

Табл. 11.3. Разпределение по клъстери.

Cluster Membership

Case	5 Clusters	4 Clusters	3 Clusters	2 Clusters
D	1	1	1	1
dr	2	1	1	1
L	3	2	1	1
Pin	3	2	1	1
PL	4	3	2	2
PH2	5	4	3	1

За да разберем кой е най-добрият клъстерен модел от дадените в Табл. 11.3, както в Пример 11.1 изследваме получената дендрограма:



Фиг. 11.3. Дендрограма на клъстерите с метода на междугруповото свързване и квадрат на евклидово разстояние.

За улеснение на Фиг. 11.3 сме нарисували червени вертикални линии, които отразяват различните клъстерни решения. Броят на пресечените хоризонтални линии е броят на клъстерите за дадената вертикална права. Приема се, че голям отскок е при поне 5 относителни единици разстояние между 2 съседни линии. От дендрограмата се вижда, че най-голям отскок с почти 18 относителни единици имаме между най-дясната и средната червена линия. Оттук определяме, че най-доброто групиране за нашите данни е решението с 2 клъстера.

От табл. 11.3 това решение е:
1-ви клъстер= {L, Pin, D, dr, PH2}
2-ри клъстер={PL}

Променливата PH2 последна се групира към първия клъстер. При по-нататъшни изследвания с по-големи извадки се установява, че тя ще се отдалечи достатъчно, за да формира отделен клъстер.

Задача. Изследвайте близостта на 6-те променливи с изходната зависима променлива Eff, като определите в коя група попада с помощта на КА. Какви заключения можем да предложим?

Коментар. Разгледаните примери включват много малки по размер извадки и затова получените резултати не отразяват адекватно свойствата на генералната съвкупност, от която са взети. Трябва да се разглеждат само като учебни примери.

Използвана литература

Основна:

[1] Кулев Г., Математически анализ 1, Пловдивско унив. изд., Пловдив, 1999.
[2] Ангелов Ст., М. Аролска, Г. Векова и др., Ръководство за решаване на задачи по математически анализ, I част. Пловдивско унив. изд., Пловдив, 2002.
[3] Въндев Д. Л., Записки по приложна статистика 1, СУ "Св. Кл. Охридски", С., 2003. https://www.matematika.bg/visha-matematika/uchebnici/applstat1.pdf
[4] Въндев Д. Л., Записки по приложна статистика 2, СУ "Св. Кл. Охридски", С., 2003.
[5] Дончев Д., М. Дилчева, В. Кинова, Практическо ръководство по статистика, изд. АВТОСпектър, Пловдив, 2002 и 2ро изд. 2009.
[6] Гочева-Илиева, С. Г. Въведение в система Mathematica. Габрово: ЕксПрес, 2009.
[7] Field A., Dicovering Statistics Using SPSS, 3rd ed. London: Sage Publications, 2009.
[8] Кулина Хр. Н., М. В. Такев, Разработка на учебни материали по многомерен статистически анализ на данни от областта на маркетинга с помощта на софтуерния пакет SPSS, Научна сесия на СУБ-Пловдив, 2013. Научни трудове на СУБ-Пловдив, Серия В, Техника и технологии, том XI, стр. 52-59, ISSN 1311-9192.
[9] www.wolfram.com
[10] www.wolframalpha.com

Допълнителна литература:

[1] Izenman A. J., Modern Multivariate Statistical Techniques Regression, Classification, and Manifold Learning, New York: Springer, 2008.

[2] Jolliffe I.T. "A note on the Use of Principal Components in Regression". Journal of the Royal Statistical Society, Series C (Applied Statistics), vol. 31 (3), pp. 300–303(1982). doi:10.2307/2348005

[3] Ross S. M., Introduction to probability and statistics for engineers and scientists, 3 rd ed., Elsevier Academic Press, Amsterdam, New York, 2004.