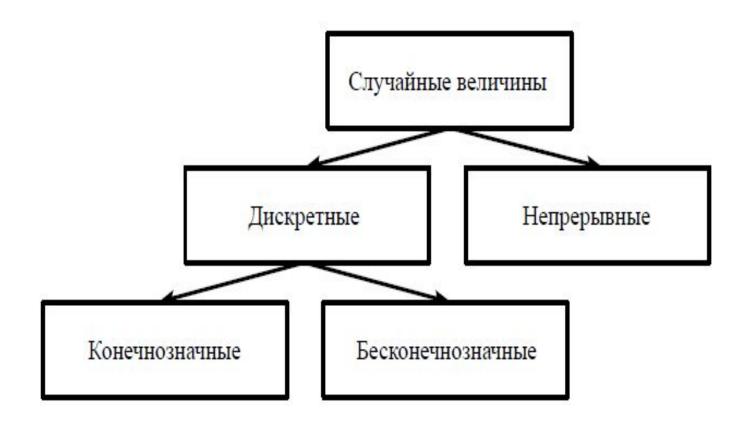
# Случайные величины

- Определение 3.1. Случайной величиной называется величина, которая в результате опыта принимает то или иное случайное значение.
- Обычно случайная величина обозначается большой буквой, а её возможные значения такой же маленькой (может быть, с индексом) или числами. Вот некоторые примеры.
  - Число на верхней грани кости при её бросании X. Возможные значения X: 1, 2, ..., 6. Эти числа можно обозначить  $x_1, x_2, ..., x_6$ . Всего у данной величины 6 возможных значений.
  - Число покупателей в магазине в течение дня X. Возможные значения этой величины: 0, 1, 2, .... Здесь верхний предел неизвестен. В теоретических исследованиях удобно считать, что возможные значения X все целые неотрицательные числа (бесконечное множество значений  $x_0, x_1, x_2, ..., x_n, ...$ ).
  - Время работы изделия до отказа T. Здесь возможные значения неотрицательные действительные числа. Мы не знаем максимального значения, поэтому также считаем, что  $t \in [0; \infty)$ .

- **Определение 3.2.** Случайная величина называется *дискретной*, если множество её возможных значений конечно или является бесконечным счётным множеством.
- Определение 3.3. Случайная величина называется *непрерывной*, если множество её возможных значений целиком заполняет некоторый промежуток или систему промежутков.
- **Определение 3.4.** Дискретная случайная величина называется *конечнозначной*, если множество её возможных значений конечно.
- **Определение 3.5.** Дискретная случайная величина называется *бесконечнозначной*, если множество её возможных значений является бесконечным счётным множеством.



- Другая классификация, независимая от приведенной на рис. 3.1 – это разделение случайных величин на одномерные (скалярные) и многомерные (векторные).
- Все те величины, которые мы выше рассматривали это одномерные, или скалярные случайные величины.
- Определение 3.6. Многомерной случайной величиной, или системой случайных величин, или случайным вектором называется совокупность нескольких рассматриваемых совместно случайных величин.

• Обозначив  $p_k = P(X = x_k)$ , получим основное правило, которому подчиняются вероятности принятия дискретной случайной величиной её возможных значений:

$$\sum_{k=1}^{n} p_k = 1 \qquad (3.1)$$

• для конечнозначной величины и

$$\sum_{k=1}^{\infty} p_k = 1 \quad (3.2)$$

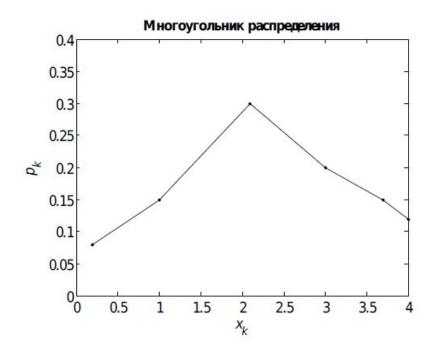
• для бесконечнозначной.

- Определение 3.7. Законом распределения дискретной случайной величины называется любое правило, по которому всем возможным значениям  $x_k$  случайной величины X ставятся в соответствие вероятности их появления  $p_k$ .
- Фактически закон распределения это функциональная зависимость  $p_k$  от  $x_k$ . Её можно задавать разными способами. Во-первых, это может быть таблица. Она называется *рядом распределения*.
- **Пример 3.1.** В таблице 3.1 показан ряд распределения случайной величины X суммы чисел на верхних гранях двух костей при их бросании. Возможные значения этой величины целые числа от 2 до 12.
- Таблица 3.1. Пример ряда распределения дискретной случайной величины

x <sub>k</sub> ¤	2¤	<i>3</i> ¤	<b>4</b> ¤	<b>5</b> ¤	6¤	7¤	8a	<b>9</b> ¤	10¤	11¤	12¤
p <sub>k</sub> a	$\frac{1}{36}^{\square}$	$\frac{2}{36}^{\square}$	$\frac{3}{36}^{\square}$	$\frac{4}{36}^{\square}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	1 36

- **Пример 3.2.** Построим с помощью MATLAB многоугольник распределения. Предполагаем, что ряд распределения (таблица вида 3.1) задан.
- Результат на рисунке.

```
x=[0.2\ 1\ 2.1\ 3\ 3.7\ 4]; % задали x(i) p=[0.08\ 0.15\ 0.3\ 0.2\ 0.15\ 0.12]; % задали p(i) p=p/sum(p); % нормировали до единичной суммы plot(x,p,'k-',x,p,'k.') % многоугольник распределения ylim([o\ 0.4]) set(get(gcf,'CurrentAxes'),... 'FontName','Times New Roman Cyr','FontSize',10) % шрифт title('\bfMногоугольник распределения') % заголовок xlabel('\itx\_k') % метка оси OX ylabel('\itp\_k') % метка оси OY
```



• Пример 3.3 аналитического закона распределения для конечнозначной величины: вероятность появления различного количества гербов при 3 бросаниях монеты:

$$p_{k} = \begin{cases} \frac{c_{3}^{k}}{8}; & k \in [0, 1, 2, 3] \\ 0; & k \notin [0, 1, 2, 3] \end{cases}$$

• В этом распределении участвуют биномиальные коэффициенты, и оно называется биномиальным.

- Следующая характеристика случайной величины это функция распределения. Её другие названия: интегральная функция распределения, интегральный закон распределения. В англоязычной литературе применяется термин the cumulative distribution function. Эта характеристика годится и для дискретных, и для непрерывных величин.
- Определение 3.8. Функцией распределения случайной величины X называется вероятность принятия ею значений, меньших конкретного числа x, рассматриваемая как функция x:

$$F(x)=P(X < x)$$
. (3.5)

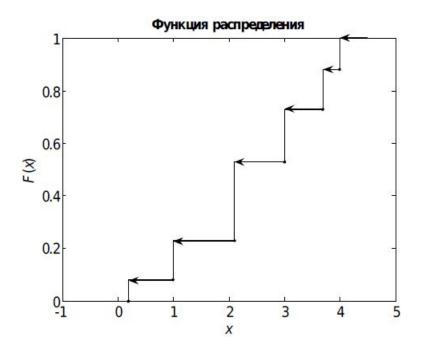
• Функция распределения является вероятностью, поэтому она безразмерная.

- Свойство 3.1. Функция распределения это вероятность некоторого события, поэтому.
- Свойство 3.2.  $F(-\infty) = P(X < -\infty) = 0$ .
- Свойство 3.3.  $F(\infty) = P(X < \infty) = 1$ .
- **Свойство 3.4.** Функция распределения неубывающая:  $x_2 > x_1$ :  $F(x_2) \ge F(x_1)$ .
- **Свойство 3.5.** Вероятность попадания величины в полуинтервал равна разности значений функции распределения на его концах:  $P(x_1 \ X < x_2) = F(x_2) F(x_1)$ .

- **Пример 3.2 (продолжение).** Найти функцию распределения случайной величины, многоугольник распределения которой мы построили.
- **Решение.** Для вычисления функции распределения применяем определение (3.5) и правило (3.6).
- График функции распределения дискретной величины представляет ступенчатую ломаную. Нарисуем её средствами MATLAB.

```
F=cumsum(p); % значения функции распределения x1=[x(1)-0.5 x x(end)+0.5]; % добавки слева и справа F1=[o F 1]; figure; % новая фигура stairs(x1,F1,'k-'); % ломаная xl=xlim; % границы рисунка yl=ylim; hold on
```

```
plot(x,F1(1:end-2),'k.') % добавили точки
hh=get(gca);
hp=hh.Position; % положение осей на фигуре
for i=1:length(F),
xi=x1(2+i:-1:1+i); % координаты стрелок
Fi=[F(i) F(i)];
xi=(xi-xl(1))/(xl(2)-xl(1))*hp(3)+hp(1); % нормализуем
Fi=(Fi-yl(1))/(yl(2)-yl(1))*hp(4)+hp(2); % стрелки
annotation('arrow',xi,Fi); % добавляем стрелки
end
hold off
set(get(gcf,'CurrentAxes'),...
'FontName','Times New Roman Cyr','FontSize',10) % wpupm
title(' bf \Phi y + \kappa u u s pacnpe de ne + u s') % заголовок
xlabel('\itx') % метка оси ОХ
ylabel('\itF\rm(\itx\rm)') % метка оси ОҮ
```



- Определение 3.9. Случайная величина X называется непрерывной, если её функция распределения F(x) непрерывна.
- Все остальные свойства F(x) остаются в силе: это функция неубывающая, меняется от о при  $x \to -$  до 1 при  $x \to +$ . Но при вычислении вероятности попадания в промежуток мы можем свободно добавлять или отбрасывать концы промежутка, т. к. в непрерывной величине учёт или неучёт конечного (или даже счётного) числа точек не влияет на вероятность:

$$P(x_1 \le X \le x_2) = P(x_1 < X \le x_2) = P(x_1 \le X < x_2) = P(x_1 < X < x_2) = P(x_1 < X < x_2) = F(x_2) - F(x_1).$$
 (3.8)

• Определение 3.10. Плотностью распределения непрерывной случайной величины X называется предел отношения вероятности её попадания в малый интервал шириной  $\Delta x$  вблизи точки x к ширине интервала  $\Delta x$  при  $\Delta x \to 0$ :  $f(x) = \lim_{\Delta x \to 0} \frac{P(x \le X, x + \Delta x)}{\Delta x}. \tag{3.9}$ 

- Другие названия: плотность вероятностей, дифференциальная функция распределения, дифференциальный закон распределения. Английское название: the partial distribution function.
- Свойство 3.6. Плотность распределения есть производная от функции распределения:

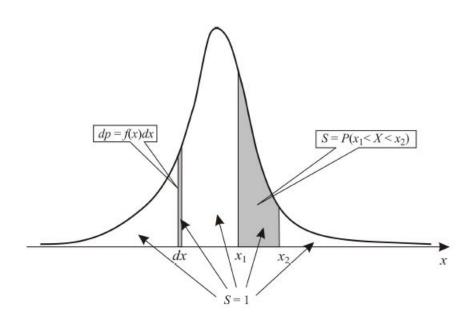
 $(3.1(f(x)) = \frac{dF(x)}{dx})$ 

- **Свойство 3.7.** Так как F(x) неубывающая, то  $f(x) \ge 0$ .
- Свойство 3.8 обратное к свойству 3.6:

- (3.11)  $F_x(x) = \int_{-\infty}^x f_x(t) dt$  Здесь аргумент плотности распределения обозначен через  $\iota$ , т. к. x верхний предел интегрирования.
- Свойство 3.9. Из предыдущего свойства имеем:

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

Это условие называется условием нормировки плотности распределения.



• **Свойство 3.10.** Вероятность попадания непрерывной величины в интервал (или отрезок, или полуинтервал) равна интегралу от плотности распределения по этому интервалу:

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x) dx.$$

• **Пример 3.5.** Плотность распределения случайной величины задана с точностью до неизвестного множителя k:

$$f(x) = \begin{cases} k(x-1); & x \in [1; 3] \\ 0; & k \notin [1; 3] \end{cases}$$
 (3.14)

Требуется вычислить этот множитель k, найти функцию распределения и построить графики f(x) и F(x).

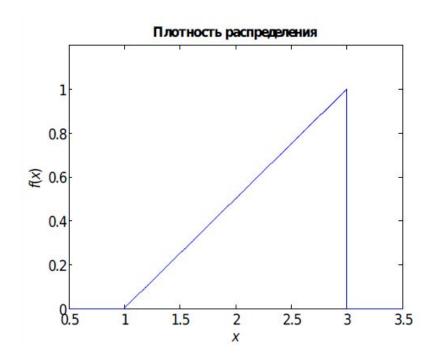
• **Решение** выполняем с помощью MATLAB. Множитель k находим из условия нормировки (3.12). Вычисляем интеграл от f(x) в заданных пределах, приравниваем его 1, и из этого уравнения находим k:

```
syms\ x\ k\ \% описали символические переменные x1=1;\ \% границы интервала x2=3;\ f=k^*(x-1);\ \% плотность распределения fprintf(f(x)=\%s\n',char(f)) I1=int(f,x,x1,x2);\ \% считаем интеграл ks=solve(I1-1,k);\ \% решаем уравнение I1=1 fprintf(f(x)=\%s\n',char(ks)) f(x)=k^*(x-1) Множитель k=f(x)
```

• Подставляем полученное значение k в выражение для f(x) и строим её график. Он показан на рисунке.

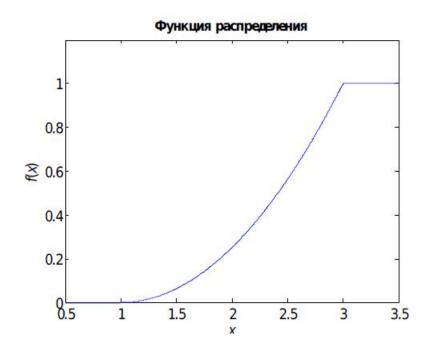
```
fs=subs(f,k,ks); % подставили решение disp('\Piлотность распределения:') 120 fprintf(['f(x)=\%s; \%d<=x<=\%d; \land nf(x)=o'... 'вне этого отрезка.\land n'], char(fs), x1, x2) xp1=x1-0.25*(x2-x1); % границы рисунка xp2=x2+0.25*(x2-x1); xp=linspace(xp1,xp2,1000); % абсциссы для графика fp=subs(fs,x,xp).*(xp>=x1).*(xp<=x2); % ординаты
```

```
plot(xp,fp) % рисуем график ylim([o\ 1.2*max(fp)]); % границы по вертикали set(get(gcf,'CurrentAxes'),... 'FontName','Times New Roman Cyr','FontSize',10) % шрифт title('\bf\Piлотность распределения') % заголовок xlabel('\itx') % метка оси OX ylabel('\itf\rm(\itx\rm)') % метка оси OY \Piлотность распределения: f(x)=1/2*x-1/2;\ 1<=x<=3; f(x)=0 вне этого отрезка.
```



- Для вычисления функции распределения используем свойство 3.8.
- Как видим, на 1-м и 3-м участках функция распределения равна соответственно 0 и 1, и требуется вычислить её лишь на 2-м. Вычисляем и строим график (рисунок).

```
F=int(fs,x,x1,x); % ф-ция распределения на среднем участке disp('\Phi yнкция распределения:') fprintf(['F(x)=o; x<\%d; \nF(x)=\%s; \%d<=x<=\%d; \n'... 'F(x)=1; x>\%d. \n'],x1,char(F),x1,x2,x2) Fp=subs(F,x,xp).*(xp>=x1).*(xp<=x2)+... ones(size(xp)).*(xp>x2); % opduhamы figure; % новая фигура plot(xp,Fp) % pucyem paghuk
```



• Определение 3.11. Математическим ожиданием или средним случайной величины X называется сумма произведений всех возможных значений величины X на вероятности их появления. При этом слово "сумма" должно пониматься обобщённо: для дискретной конечнозначной величины — это обычная сумма:

$$m_x = \sum_{k=1}^n x_k p_k;$$
 (3.15)

- ...
- а для непрерывной интеграл:

$$m_x = \sum_{k=1}^n x_k p_k;$$
 (3.15)

- **Свойство 3.11.** МО детерминированной величины C (константы) равно ей самой.
- Пример 3.6. Рассмотрим непрерывную величину с плотностью распределения

 $f(x) = \frac{1}{\pi(x^2 + 1)}.$ (3.18)

• Функция распределения вычисляется по формуле (3.11):

$$F(x) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{dt}{t^2 + 1} = \frac{1}{\pi} \left( arctg \ x + \frac{\pi}{2} \right)$$

• Но попытка вычислить МО по формуле (3.17) приводит к интегралу:

$$m_{x} = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{t \, dt}{t^2 + 1},$$

который расходится. Поэтому у этой величины математического ожидания нет.

- **Пример 3.2 (продолжение).** Найти МО дискретной конечнозначной величины, для которой мы уже построили многоугольник и функцию распределения.
- **Решение.** Вычисляем по формуле (3.15). mx = sum(x.\*p); % считаем MO  $fprintf('Mamemamuческое ожидание <math>Mx = \%8.5f.\n',mx);$  **Математическое ожидание Mx = 2.43100.**
- Пример 3.5 (продолжение). Найти МО непрерывной величины, для которой мы построили ранее плотность и функцию распределения.
- Решение. Вычисляем по формуле (3.17).

```
mx=int(x*fs,x,x1,x2); % вычисляем MO fprintf('Mamemamuческое ожидание <math>Mx=\%s=\%8.5f.\n',... char(mx),eval(mx))
```

Математическое ожидание Mx=7/3=2.33333.

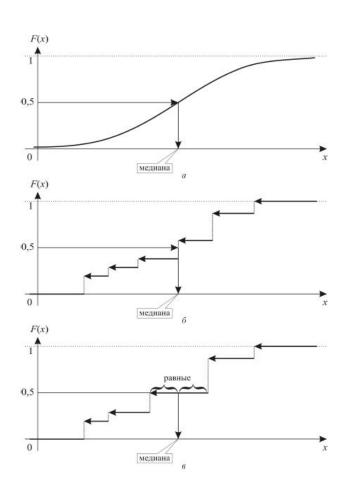
- **Определение 3.12.** *Мода* случайной величины это наиболее вероятное её значение (т. е. которое чаще всего встречается).
- В общем случае они не совпадают с математическими ожиданиями. В зависимости от вида многоугольника распределения или графика f(x) распределения бывают унимодальными (один максимум одна мода) и полимодальными (несколько мод). Бывают также распределения, у которых все возможные значения являются модами. Они называются равномерными.
- Например, случайная величина X число на верхней грани игральной кости при её бросании имеет 6 возможных значений (1, 2, 3, 4, 5 и 6) с одинаковыми вероятностями 1/6. Все они являются модами.

- **Пример 3.2 (продолжение).** Найти моду дискретной конечнозначной величины.
- **Решение.** Определяем то значение  $x_k$ , для которого  $p_k$  максимальна. [pmax,ipmax]=max(p); % pmax u номер точки modx=x(ipmax); % мода распределения  $fprintf('Moda=\%5.2f.\n',modx);$  Moda=2.10.
- **Пример 3.5 (продолжение).** Найти моду непрерывной величины. **Решение.** Определяем то значение x, для которого f(x) максимальна.

```
[fmax,ifmax]=max(fp); % максимальная f(x) modx=xp(ifmax); % мода распределения fprintf('Moda=\%5.2f.\n',modx); Moda=3.00.
```

- Определение 3.13. *Медиана* случайной величины это такое значение  $x_m$ , при котором функция распределения  $F(x_m) = 0.5$ .
- **Пример 3.2 (продолжение).** Найти медиану дискретной конечнозначной величины.
- **Решение.** Для дискретной величины медиана может приходиться на горизонтальный участок графика F(x) или на точку разрыва. Проверяем: если F(x) = 0.5 приходится на точку разрыва, то это и есть медиана. Если же F(x) = 0.5 достигается в двух точках, берём середину соединяющего их отрезка.

```
imed=find(F==0.5); % есть ли точки, где F(x)=0.5 if isempty(imed), % нет таких точек imed=min(find(F>0.5)); % номер точки разрыва через 0.5 medx=x(imed); % медиана else % есть такие точки medx=mean(x(imed:imed+1)); % середина отрезка c F(x)=0.5 end fprintf('Meduaha = %5.2f.\n',medx); Meduaha = 2.10.
```



- Пример 3.5 (продолжение). Найти медиану непрерывной величины.
- **Решение.** У нас есть аналитическое выражение для F(x). Приравниваем его 0,5 и решаем полученное уравнение.

```
medx=eval(solve(F-0.5)); % ищем медиану medx=medx(find((medx>=x1)&(medx<=x2))); % нужное решение fprintf('Meduaha=%8.5f.\n',medx); Meduaha=2.41421.
```

- Определение 3.14. Моментом (начальным моментом) m-го порядка случайной величины называется сумма произведений всех возможных значений величины X в m-й степени на вероятности их появления.
- Момент величины X обозначается  $M(X^m)$  или  $a_m$ .
- Вычислим несколько первых начальных моментов для **примеров 3.2 и 3.5** (**продолжение**).

```
disp('Начальные моменты:');
for i=1:5,
    alpha=sum(x.^i.*p); % момент i-го порядка
    fprintf('Alpha(%d)=%12.5f\n',i,alpha);
end

Начальные моменты:
Alpha(1)= 2.43100
Alpha(2)= 7.24970
Alpha(3)= 23.60689
Alpha(4)= 81.01697
Alpha(5)= 287.89826
```

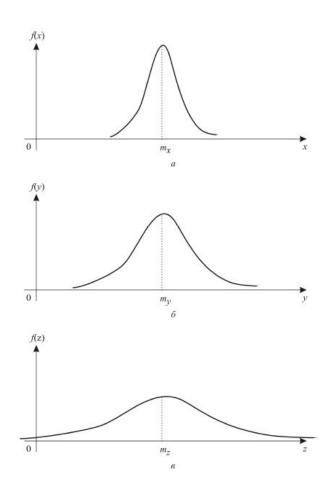
```
disp('Начальные моменты:');
for i=1:5,
  alpha=int(x^i*fs,x,x1,x2); % момент i-го порядка
 fprintf(Alpha(%d)=%s=%12.5f\n',i,...
    char(alpha),eval(alpha));
end
Начальные моменты:
Alpha(1)=7/3=2.33333
Alpha(2)=17/3= 5.66667
Alpha(3)=71/5= 14.20000
Alpha(4)=547/15= 36.46667
Alpha(5)=2005/21= 95.47619
```

- Определение 3.15. Центральным моментом m-го порядка случайной величины называется сумма произведений всех возможных значений величины X, из которых вычтено её MO, в m-й степени на вероятности их появления.
- Формулы для вычисления центральных моментов отличаются от формул (3.21)-(3.23) для начальных моментов только тем, что из значений X вычитается MO.
- Свойство 3.12. Первый центральный момент всегда равен о.

• Формулы вида (3.24) имеют место и здесь. Вот несколько первых центральных моментов для примеров 3.2 и 3.5 (продолжения). disp('Центральные моменты:'); for i=1:5,  $mu=sum((x-mx).^i.*p);$  % момент i-го порядка  $fprintf('Mu(\%d)=\%12.5f\n',i,mu);$ end Центральные моменты: Mu(1) = 0.00000Mu(2) = 1.33994Mu(3) = -0.53191Mu(4) = 3.75172Mu(5) = -3.67639

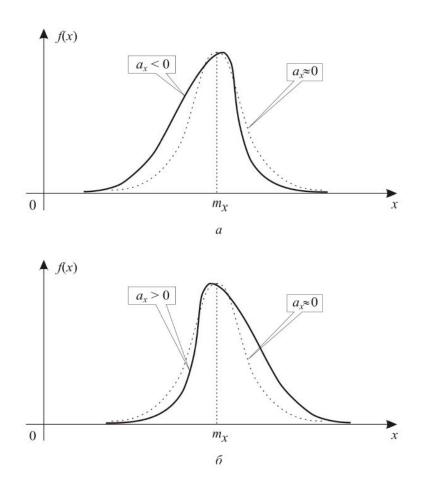
```
disp('Центральные моменты:');
for i=1:5,
    mu=int((x-mx)^i*fs,x,x1,x2); % момент i-го порядка
    fprintf('Mu(%d)=%s=%12.5f\n',i,char(mu),eval(mu));
end
Центральные моменты:
Mu(1)=0= 0.00000
Mu(2)=2/9= 0.22222
Mu(3)=-8/135=-0.05926
Mu(4)=16/135= 0.11852
Mu(5)=-128/1701=-0.07525
```

- **Определение 3.16.** *Дисперсией* случайной величины называется её 2-й центральный момент.
- Обозначения дисперсии: D(X),  $D_x$ , . Английский термин the variance. Размерность дисперсии равна квадрату размерности X.
- **Свойство 3.13.** Для детерминированной величины C (константы) дисперсия равна 0.
- Свойство 3.14. Дисперсия случайной величины положительна.
- **Свойство 3.15.** Дисперсия равна второму начальному моменту минус квадрат первого начального момента. Или: дисперсия равна МО квадрата случайной величины минус квадрат её МО.
- **Определение 3.17.** *Среднеквадратичным отклонением (СКО)* случайной величины называется квадратный корень из её дисперсии.
- Другие названия: стандартное отклонение, стандарт. Англоязычные термины:  $standard\ deviation$ , standard. Обозначение:  $s_{_{v}}$ .



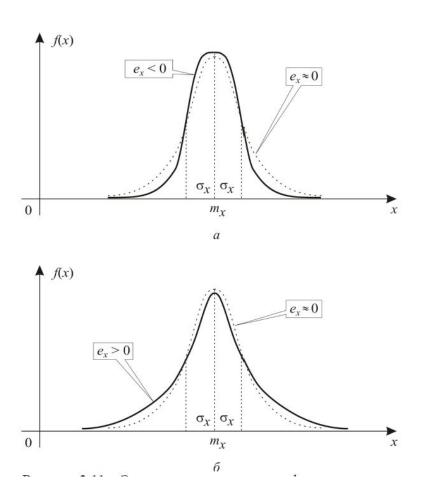
- Определение 3.18. Асимметрией случайной величины называется отношение 3-го центрального момента к кубу СКО.
- Английский термин: the skewness. Обозначения: A(X) или  $a_x$ . Асимметрия является безразмерной величиной (она специально так введена).
- Формула для её вычисления:

(3.3 
$$a_x = \frac{\mu_3}{\sigma_x^3}$$
.



- **Определение 3.19.** *Эксцессом* случайной величины называется отношение 4-го центрального момента к 4-й степени СКО (квадрату дисперсии), из которого вычтено число 3.
- В англоязычной литературе используется термин *the kurtosis*. Обозначается эксцесс: E(X) или  $e_x$ . Он является безразмерной величиной и вычисляется по формуле:

$$e_{x}=\frac{\mu_{4}}{\sigma_{x}^{4}}-3.$$



• Закончим примеры 3.2 и 3.5: посчитаем дисперсию, СКО, асимметрию и эксцесс.

```
Dx=sum((x-mx).^2.*p); % ducnepcus
    Sx=Dx^{0.5}; % CKO
  Ax = sum((x-mx).^3.*p)/Sx^3; % асимметрия
    Ex=sum((x-mx).^4.*p)/Dx^2-3; % эксцесс
fprintf('Дисперсия Dx=\%8.5f; \n',Dx);
fprintf('Cpeднеквадратичное \mathfrak{S}_{\overline{m}} = \mathfrak{S}_{\overline{k}} =
fprintf('Acumмempuя Ax=\%8.5f; \n',Ax);
fprintf('9\kappa cuecc\ Ex=\%8.5f.\n',Ex);
  Дисперсия Dx= 1.33994;
    Среднеквадратичное отклонение Sx= 1.15756;
  Aсимметрия Ax=-0.34294;
    Эксцесс Ex = -0.91042.
```

```
Dx = simple(int((x-mx)^2*fs,x,x1,x2)); \% \partial ucnepcus
Sx = simple(Dx^0.5); \% CKO
Ax = simple(int((x-mx)^3*fs,x,x1,x2)/Sx^3); \% acummempus
Ex=simple(int((x-mx)^4*fs,x,x1,x2)/Dx^2-3); % эксцесс
fprintf('Дисперсия Dx=%s=%8.5f; \n', char(Dx), eval(Dx));
fprintf('Cpedheквadpamuчное omклонение Sx=%s=%8.5f; \n',...
char(Sx),eval(Sx));
fprintf(')xcuecc Ex=%s=%8.5f. \n', char(Ex), eval(Ex));
Дисперсия Dx=2/9=0.22222;
Среднеквадратичное отклонение Sx=1/3*2^{(1/2)}=0.47140;
Асимметрия Ax=-2/5*2^{(1/2)}=-0.56569;
9\kappa cuecc Ex = -3/5 = -0.60000.
```