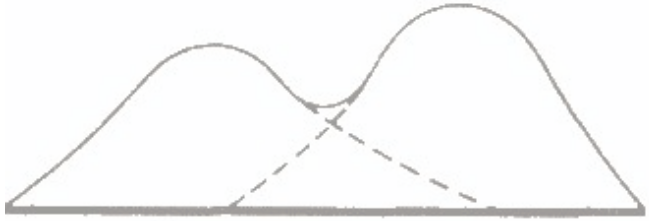
Описательные характеристики распределения тестовых результатов

1. **Меры среднего положения** (меры центральной тенденции).
   - Мода, медиана, среднее
2. **Меры вариации** (рассеяния, изменчивости данных).
   - Размах, дисперсия, стандартное отклонение, коэффициент вариации
3. **Меры формы** (меры симметрии и островершинности кривой распределения).
   - Коэффициенты асимметрии, эксцесса
4. Квантили.
   - Квартили, процентили

Меры среднего положения

- Это обобщающие показатели, характеризующие типичное значение, присущее большинству единиц совокупности, позволяющие выявлять закономерности.

- ***Мода, М –*** это наиболее часто встречающееся значение признака в исследуемой совокупности.
  - **Унимодальное распределение,**
  - ***Бимодальное распределение,***
  - ***Мультимодальное распределение.***



Бимодальное распределение

В данной совокупности имеются две относительно самостоятельные группы.

- ***Медиана, Md*** – это значение, которое делит упорядоченную совокупность данных пополам, так что одна половина значений больше медианы, а другая – меньше.

Меры среднего положения

- ***Среднее,*** $\bar{x}$, – это значение, которое отражает типичное значение для исследуемой совокупности в данных условиях:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n},$$

где n – объем совокупности, x_i – i-е значение совокупности.

– Чтобы средняя величина была действительно обобщающей характеристикой, улавливающей закономерность, она должна применяться к достаточно однородной совокупности.

– Средняя величина рассчитывается только для количественных признаков.

Соотношение мер среднего положения

- У симметричных унимодальных распределений:

$$\bar{x} = M = Md$$

- У унимодальных распределений с правой асимметрией:

$$M < Md < \bar{x}$$

- У унимодальных распределений с левой асимметрией:

$$M > Md > \bar{x}$$

Выбор меры центральной тенденции в зависимости от типа измерительной шкалы

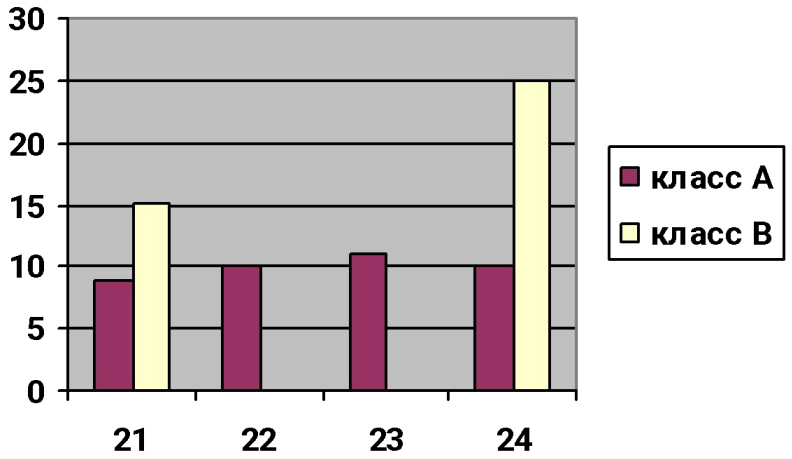
Тип шкалы	Меры центральной тенденции
Номинальная	Мода
Ранговая	Мода, медиана
Интервальная	Мода, медиана, среднее
Отношений	Мода, медиана, среднее

Меры изменчивости

- ***Меры изменчивости*** – показатели, измеряющие вариацию (разброс) значений совокупности. К мерам изменчивости относятся: размах, дисперсия, стандартное отклонение, коэффициент вариации, и др.

- ***Размах вариации*** (***R***), отражает пределы изменчивости значений совокупности.
- Представляет собой разность между максимальным (x_{max}) и минимальным значением (x_{min}) совокупности:

$$R = x_{\max} - x_{\min}.$$



- Размах не учитывает всех значений в выборке и определяется только двумя значениями.

Меры изменчивости

- ***Дисперсия,*** δ^2 **или** s^2 – это значение, которое отражает внутреннюю изменчивость значений исследуемой совокупности:

$$\delta^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1},$$

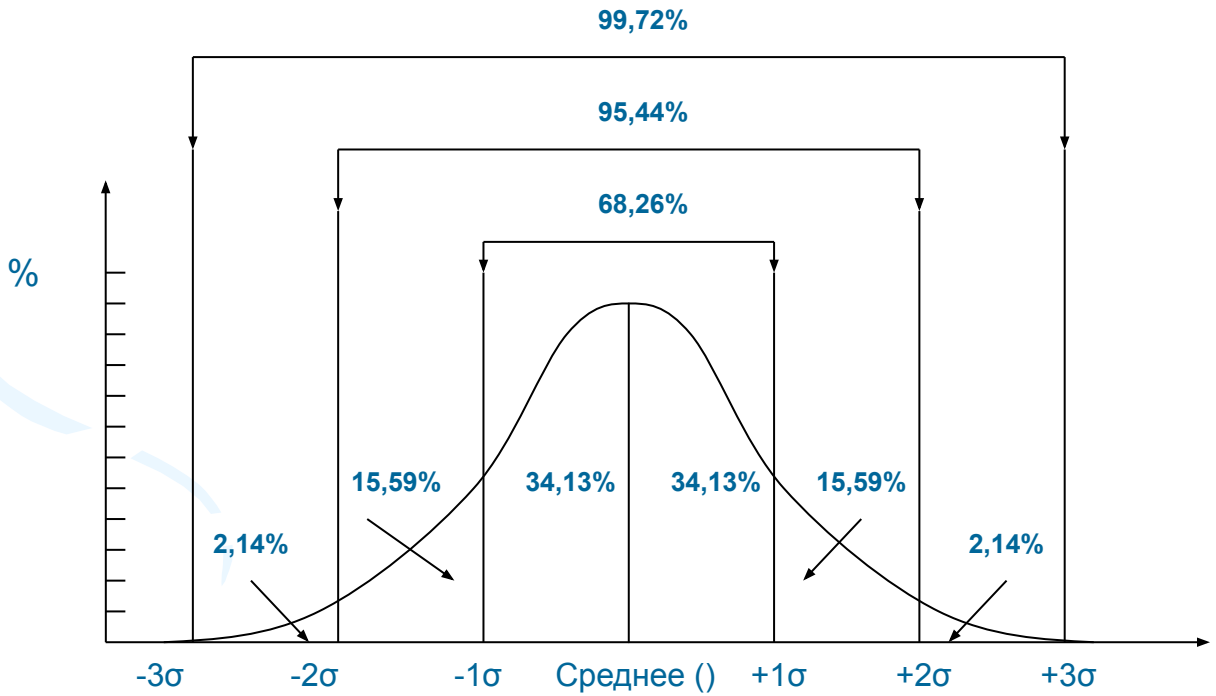
- где n – объем совокупности,
  x_i – i-е значение, $\bar{x}$ – среднее значение.

- ***Стандартное отклонение,*** δ **или** s, показывает насколько в среднем отклоняется каждое значение (x_i) от среднего $\bar{x}$:

$$\delta = \sqrt{\delta^2},$$

Меры изменчивости

- **<u>В случае нормальности</u>** распределения исследуемой совокупности значений справедливо следующее (правило «Трех сигм»):
- 70% значений лежит между $\bar{x} - \sigma$ и $\bar{x} + \sigma$,
- 95% значений лежит между $\bar{x} - 2\sigma$ и $\bar{x} + 2\sigma$,
- 99% значений лежит между $\bar{x} - 3\sigma$ и $\bar{x} + 3\sigma$.



Процентное распределение наблюдений под нормальной кривой

Меры изменчивости

- Для сопоставления изменчивости признаков используют ***коэффициент вариации***

$$Cv = \frac{\sigma}{\overline{x}} \cdot 100\% .$$

- При симметричных распределениях коэффициент вариации Cv не превышает 50%. При сильно асимметричных рядах распределения коэффициент вариации Cv может принимать значение 100% и даже выше.

Варьирование считается

- слабым, если $Cv < 10\%$,
- средним, если $10\% < Cv \leq 33\%$,
- значительным, при $Cv > 33\ \%$.

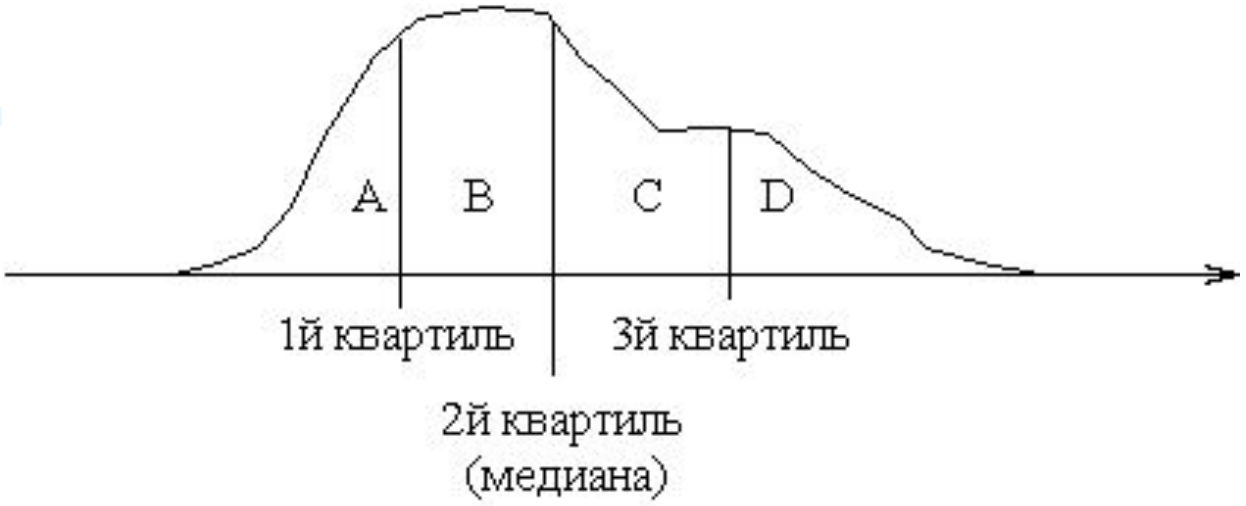
- Если значение коэффициента вариации Cv не превышает 33%, то изучаемая совокупность считается однородной.

Квантили

Квантиль – делит совокупность на заданное число равных частей по количеству значений.

Число частей может быть различным, отсюда и разные квантили – **квартили**, **децили**, **перцентили**.

Квартиль делит совокупность на четыре равные части, по 25% значений а каждой части.



Квантили

- ***Дециль*** делит совокупность на 10 равных частей, по 10% значений в каждой части.

- ***Перцентиль*** делит совокупность на 100 равных частей.

- Например,

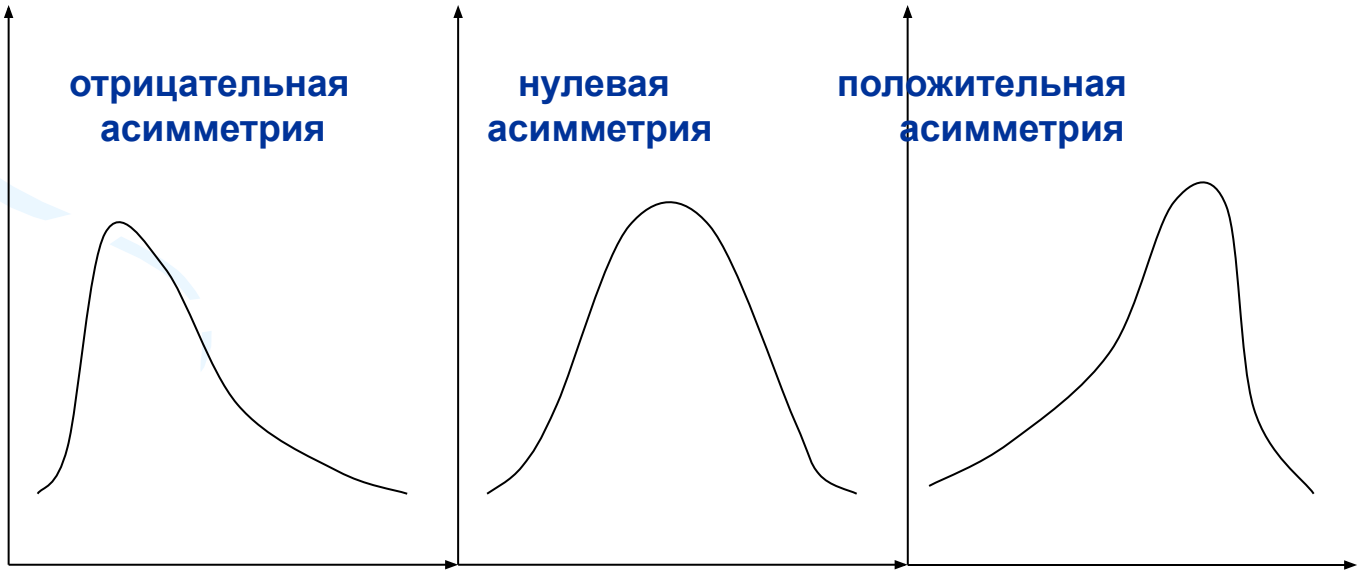
$$D_1 = C_{10},\ Q_1 = C_{25},\ Md = C_{50},\ Q_3 = C_{75}$$ и т.д.

- **<u>Необходимость расчета квантилей</u>** вызвана теми же причинами, что и расчет медианы: низкая чувствительность к случайным резким отклонениям значений признака.

Меры формы

Асимметрия (As) – степень отклонения распределения от симметричного распределения, характерного для нормальной кривой.

- Асимметрия *As* принимает значения в диапазоне от –3 до +3.
  - As = 0, распределение симметрично;
  - As < 0, левосторонняя асимметрия,
  - As > 0, правосторонняя асимметрия.

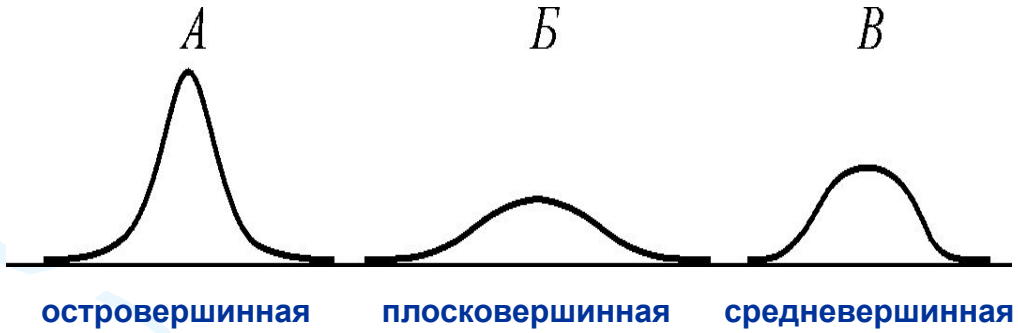


Меры формы

Эксцесс (Ex) – степень островершинности кривой распределения.

Эксцесс (Ex) принимает значения в диапазоне от –3 до +3.

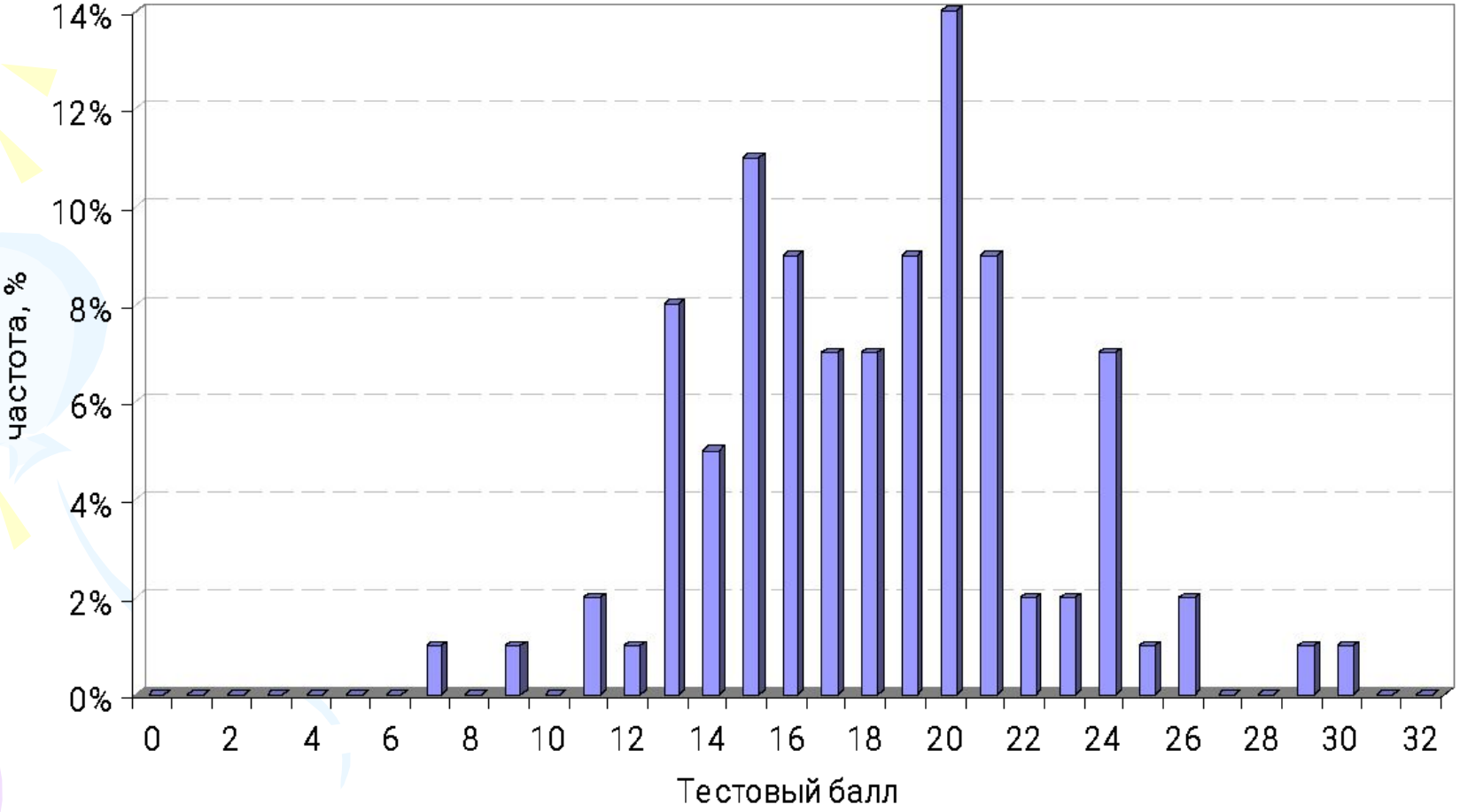
- ***Ex*** = 0, распределение средневершинно;
- ***Ex*** < 0, плосковершинная кривая,
- ***Ex*** > 0, островершинная кривая.



Понятие «эксцесс» применимо лишь к унимодальным распределениям. Если две моды, то говорят об эксцессе кривой в окрестности каждой моды.

Меры асимметрии и эксцесса можно использовать для сравнения различных распределений.

Гистограмма распределения результатов тестирования



План апробации тестовых заданий

1.
2.
3.
4.
5.
6.
7.
8.
9.
10.

Спасибо за внимание!

Ваши вопросы
и пожелания