

# Статистика в НМД 2

# Логика проверки статистических гипотез

## Статистические критерии

**Статистический критерий** (statistical test) – статистический метод принятия решения о том, стоит ли отвергнуть нулевую гипотезу в пользу альтернативной или нет, соответствующий особенностям выборки.

- **Эмпирическое значение критерия** (или расчётное) рассчитывается определённым статистическим способом.
- **Критическое значение критерия** - известное (например, заданное таблично) эталонное значение соответствующее определённому уровню значимости при различных степенях свободы.

# Процедура проверки статистической гипотезы

Сформулировать нулевую и альтернативную гипотезы.

Выбрать статистический тест (критерий).

Выбрать требуемый уровень значимости ( $\alpha=0.05, 0.01, 0.025, \dots$ ).

Вычислить эмпирическое значение критерия по тесту.

Сравнить с критическим значением критерия по тесту при заданном критическом уровне значимости.

Принять решение (для большинства тестов приемлемо правило: если вычисленное значение больше, чем критическое, нулевая гипотеза отклоняется).

# Классификация статистических методов (критериев)

## По количеству анализируемых признаков

**Одномерные** (анализ каждого признака в отдельности);

**Двухмерные** (одновременный анализ двух признаков, например анализ связей изучаемых признаков - ассоциации или корреляции);

**Многофакторные** (анализ трех признаков и более одновременно, например многофакторный дисперсионный анализ, многофакторный регрессионный анализ, дискриминантный анализ).

# По статистическим принципам, лежащим в основе методов

**Параметрические** - применяются для анализа нормально распределенных количественных признаков (**метрические шкалы – интервальные или относительные**). Требуют вычисления средних величин и показателей рассеивания.

**Непараметрические** - применяются в остальных случаях:

- для анализа количественных признаков независимо от вида их распределения (не требуют, чтобы данные подчинялись какому-то определённому типу распределения);
- для анализа маленьких по размеру выборок;
- для анализа качественных признаков (неметрические шкалы – номинальные или порядковые).

# По возможности учета имеющихся априори предположений

**Односторонние тесты** - учитывают исходное (априорное) предположение о том, что в одной из групп распределение признака смещено в определенную сторону (в сторону увеличения либо уменьшения) по отношению к другой. Другими словами, направление эффекта задано.

**Двусторонние тесты** - используются в отсутствие исходного предположения о том, что в одной из групп распределение признака смещено в определенную сторону (в сторону уменьшения или увеличения) по отношению к другой. Другими словами - направление эффекта заранее не известно.

## По зависимости или независимости сопоставляемых выборок

**Тесты для независимых выборок.** Используются в случае, если при формировании выборок объекты исследования набирались в группы независимо друг от друга.

**Тесты для зависимых выборок.** Используются для анализа динамики показателей, полученных в одной группе.

## Статистические методы сравнения 2-х выборок

Вид тестовых шкал	Вид данных	Зависимые выборки		Независимые выборки	
		Параметрический	Непараметрический	Параметрический	Непараметрический
Номинальная (наименований)	Качественные: номинальные, категориальные	-	Критерий Мак-Немара	-	Критерий хи-квадрат Пирсона (критерий согласия)
Порядковая (ранговая, ординальная)	Качественные: порядковые, ранговые, ординальные	-	Критерий знаков Вилкоксона	-	U-критерий Манна-Уитни Критерий ранговых сумм Вилкоксона Критерий Колмогорова-Смирнова
Интервальная (разностей)	Количественные: дискретные, непрерывные	t-критерий Стьюдента (независимые выборки)	Критерий знаков Вилкоксона	t-критерий Стьюдента (независимые выборки)	U-критерий Манна-Уитни Критерий ранговых сумм Вилкоксона Критерий Колмогорова-Смирнова
Абсолютная (отношений)					

# Сравнение 2-х выборок по 1-му признаку

## Параметрические критерии – количественные нормально распределённые данные

### Критерий Стьюдента (t-критерий Стьюдента)

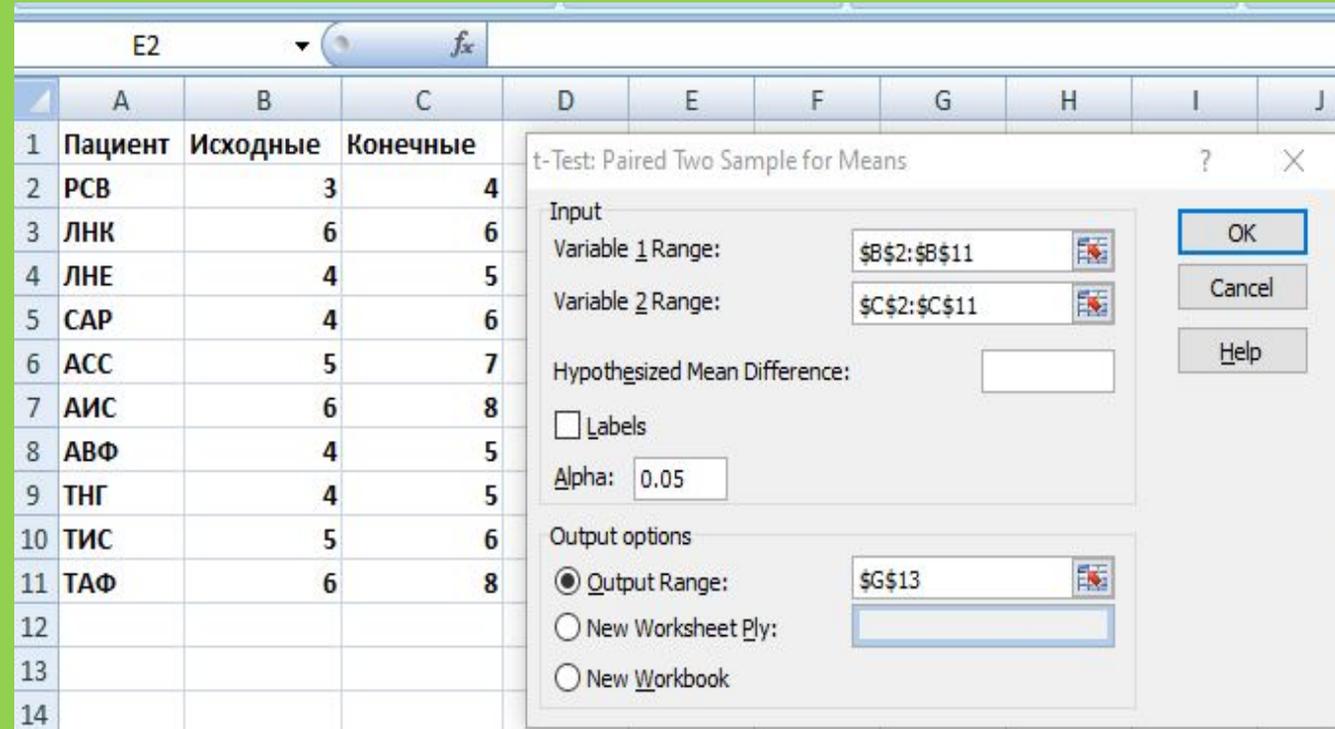
- t-критерий Стьюдента (псевдоним У. Госсета). Позволяет проверить гипотезу о статистической значимости разности двух средних арифметических в 2-х зависимых или независимых выборках.
- **Нулевая гипотеза** – отсутствие различий средних арифметических значений переменной в двух выборках. В нулевой гипотезе используемая статистика соответствует **t-распределению**.
  - **Характеристики t-распределения:** близко к нормальному, но кривая более приплюснута, с более длинными «хвостами»;
- **Требования к выборкам (математические допущения)**
  - Нормальность распределения переменных в обеих выборках.
  - Одинаковость дисперсии (стандартного отклонения).
  - Допускается сравнение не более 2-х групп.
  - Группы могут большие ( $n > 30$ ), малые или не равными по размеру. Однако, в маленьких выборках трудно установить характер распределения.

# Оценка статистической значимости различий между среднегрупповыми значениями с помощью t-критерия Стьюдента в MS Excel

- В MS Excel имеется три варианта t-критерия Стьюдента:
  - парный двухвыборочный тест для средних (paired two sample for means) для зависимых выборок,
  - двухвыборочный с одинаковыми дисперсиями (two-sample assuming equal variance) для независимых выборок,
  - двухвыборочный с неравными дисперсиями (two-sample assuming unequal variance) для независимых выборок.

# Алгоритм вычисления

1. Из пакета «Анализ данных» выбрать необходимый тест, например парный двухвыборочный тест для средних (paired two sample for means).
2. В диалоговом окне парного двухвыборочного теста установить необходимые параметры:
  - Интервал переменной 1 (Variable 1 range) и интервал переменной 2 (Variable 2 range), выбрать ячейки двух цифровых рядов с результатами измерений.
  - Альфа (Alpha) – задает уровень статистической значимости (в большинстве случаев достаточно уровня 0,05).
  - Выходной диапазон (Output range) – выбрать первую ячейку выходного диапазона на том же листе.
  - Нажать [OK]. Результаты статистического анализа появятся в выходном диапазоне.



# Результаты анализа различий с помощью парного двухвыборочного t-теста Стьюдента

t-Test: Paired Two Sample for Means			
Парный двухвыборочный тест для средних			
		Variable 1	Variable 2
Среднее	Mean	4.7	6
Дисперсия	Variance	1.1222222	1.7777778
Наблюдения	Observations	10	10
Корреляция Пирсона	Pearson Correlation	0.8653106	
Гипотетическая разность средних	Hypothesized Mean Difference	0	
Степени свободы	df	9	
t-статистика	t Stat	-6.0907767	
P(T<=t) одностороннее	P(T<=t) one-tail	9.063E-05	
t критическое одностороннее	t Critical one-tail	1.8331129	
P(T<=t) двухстороннее	P(T<=t) two-tail	0.0001813	
t критическое двухстороннее	t Critical two-tail	2.2621572	

Примечание.  
 Значение 9.063E-05 для одностороннего P. E-05 означает, что цифры перед E-05 нужно умножить на 10 в степени минус 5, т.е. 0,00009063. Округляется до 3-го знака после запятой и записывается как p=0,000

# Непараметрический аналог t-критерия Стьюдента для зависимых выборок - критерий знаковых рангов Вилкоксона (Wilcoxon Signed-Rank Test)

## Алгоритм проведения

- Открыть веб страницу <http://www.socscistatistics.com/tests/signedranks/>

### Wilcoxon Signed-Rank Test Calculator

Treatment 1	Treatment 2
4	3
5	3
6	5
3	2
5	4
8	7
1	1
3	2
9	8

Significance Level:

0.01

0.05

1 or 2-tailed hypothesis?:

One-tailed

Two-tailed

No calculation performed yet.

Calculate

Reset

Вилкоксон тест *непараметрических* тестов, призванных оценить разницу между двумя процедурами или условиями, где образцы коррелируют. В частности это подходит для оценки данных от дизайна повторил мер в ситуации, когда не выполняются предварительные условия для зависимых выборок  $t$  тест. Так например, он может использоваться для оценки данных из эксперимента, который смотрит на чтение способности детей до и после того, как они проходят период интенсивной подготовки.

#### *Требования к*

- Соответствующие данные
- Зависимой переменной является непрерывным - другими словами, должна в принципе, можно было различать между значениями в  $n$ -го десятичного
- Для максимальной точности не должно быть никаких связей, хотя этот тест - как и другие - способ обработки связей

#### *Null гипотеза*

Null гипотеза утверждает, что *Медиан* двух образцов идентичны.

Take me to the calculator!



Translate URL

English (Auto-Detected)



Russian



Views

Translated 100%

Mouse over text to see original

Введите парные лечебные значения в текстовые поля ниже, либо одно очко в строке или в разделенном запятыми списке. Помните, потому что Вилкоксона подписали ранг тест предназначен для пары значений, вам нужно иметь одинаковое количество баллов в обеих условиях лечения.

Лечение 1



Лечение 2



- Применяется для сравнения двух зависимых групп по одному признаку.
- При размере выборки  $n > 20$  статистика критерия имеет нормальное распределение. Поэтому можно пользоваться  $z$  значением для оценки гипотезы. При небольшой выборке (меньше 10) пользуются  $W$  значениями.
- Если различия в паре данных равно 0 (например, у участника исследования нет изменения параметра), данная пара исключается из анализа, что уменьшает размер выборки. .
- **Сущность теста:** оценивается разность измерений в каждой паре. Разности располагаются в порядке модульного значения и ранжируются.
- **Требования к сравниваемым выборкам.**
  - Непрерывные данные (continuous).
- **Нулевая гипотеза** – отсутствие статистически значимых различий между медианами сравниваемых выборок.
- **Оценка**
  - Сравняется рассчитанное и критическое значения критерия  $W$ .
  - Если  $W_{рас} > W_{кр}$  – различия между выборками считаются статистически значимыми.
  - Если  $W_{рас} < W_{кр}$  – различия между выборками считаются статистически не значимыми.

Лечение 1

1  
2  
3  
4  
4  
5  
5  
6  
6  
7  
7

Лечение 2

2  
3  
3  
5  
6  
6  
7  
8  
9  
9  
10

Уровень значимости:

0.01

0.05

1 или 2-хвост гипотеза?:

Одностороннее

Двустороннее

Не забудьте выбрать уровень значимости и является ли ваша гипотеза один или два хвост.

Calculate

Reset

Treatment 1	Treatment 2	Sign	Abs	R	Sign R
1	2	-1	1	2.5	-2.5
2	3	-1	1	2.5	-2.5
3	3	n/a	0	n/a	n/a
4	5	-1	1	2.5	-2.5
4	6	-1	2	6.5	-6.5
5	6	-1	1	2.5	-2.5
5	7	-1	2	6.5	-6.5
6	8	-1	2	6.5	-6.5
6	9	-1	3	9.5	-9.5
7	9	-1	2	6.5	-6.5

Significance Level:

0.01

0.05

1 or 2-tailed hypothesis?:

One-tailed

Two-tailed

#### Result Details

*W-value: 0*

*Mean Difference: 1.7*

*Sum of pos. ranks: 0*

*Sum of neg. ranks: 55*

*Z-value: -2.8031*

*Mean (W): 27.5*

*Standard Deviation (W): 9.81*

*Sample Size (N): 10*

*Result 1 - Z-value*

The Z-value is -2.8031. The p-value is 0.00512. The result is significant at  $p \leq 0.05$ .

*Result 2 - W-value*

The W-value is 0. The critical value of W for  $N = 10$  at  $p \leq 0.05$  is 8. Therefore, the result is significant at  $p \leq 0.05$ .

1. В поля «Лечение 1» («Treatment 1») и «Лечение 2» («Treatment 2») ввести данные из двух выборок.
2. Выбрать уровень значимости (Significance level) – 0,05
3. Выбрать вид гипотезы – 1- или 2-хвостовую (1 or 2-tailed hypothesis) - 2-хвостовую.
4. Нажать «Рассчитать» (calculate).
5. Оценить статистическую значимость по результатам анализа

*Результат 1 - Z-значение*

Значение  $z$  равно -2.5205.

Примечание:  $N(8)$  (количество пар) не является достаточно большим для формирования нормального распределения Wilcoxon  $W$  статистики. Поэтому не возможно рассчитать точное значение  $p$ .

*Результат 2 - W-значение*

Значение  $W$  критерия равно 0. Критическое значение  $W$  при  $N = 8$  ( $p < .05$ ) равно 1.

**Результат является статистически значимым при  $p < .05$ .**

**Статистические методы  
исследования зависимостей:  
корреляционный, дисперсионный и  
регрессионный анализ**

# Функциональная и статистическая зависимость

- **Функциональная зависимость** (взаимосвязь) - каждому значению одной переменной соответствует строго определенное значение другой. Например, в функции  $y = 2 * x$  каждому значению  $x$  соответствует в два раза большее значение  $y$ .
- **Статистическая зависимость** – каждому значению одного параметра может соответствовать несколько значений другого; с изменением одного признака изменяется и другой.

**Для определения статистической зависимости применяют:**

- **корреляционный и дисперсный анализ** - для установления факта наличия/отсутствия зависимости между переменными (характера и силы зависимости).
- **регрессионный анализ** - для нахождения количественной зависимости между переменными.

# Корреляционный анализ. Виды корреляции

**Корреляционный анализ** – применение статистических методов для исследования взаимосвязи между переменными, т.е. насколько согласованно они меняются. **Основной задачей** корреляционного анализа является определение тесноты (силы) и направленности статистической зависимости изучаемых показателей. Следует помнить, что корреляция не есть причинность.

## Виды корреляции

### • По направлению

- **Прямая (положительная)** - с увеличением (уменьшением) одного признака в основном увеличиваются (уменьшаются) значения другого.
- **Обратная (отрицательная)** - с увеличением (уменьшением) одного признака в основном уменьшаются (увеличиваются) значения другого.

### • По форме

- **Линейная** - изменение одной переменной на одну единицу всегда приводит к изменению другой переменной на одну и ту же величину (график представляет прямую линию).
- **Нелинейная** – любая другая.

# Корреляционное поле (диаграмма рассеивания)

- Графическое представление данных в прямоугольной система координат, при котором каждой паре переменных соответствует одна точка (зона разброса).
- Используется для предварительного визуального анализа корреляционной связи.
- Позволяет оценить наличие корреляции (группировка точек вдоль одной линии), направление корреляции и её силу (по плотности точек).



# Коэффициент корреляции

**Коэффициент корреляции** - количественная мера взаимосвязи (совместной изменчивости) двух переменных.

## Признаки (характеристики) коэффициентов линейной корреляции ( $r$ )

- На основании коэффициентов корреляции можно судить только о прямолинейной корреляционной взаимосвязи между признаками.
- Значения коэффициентов корреляции не может быть меньше -1 и больше +1.
  - $R = 0$  - связь между признаками  $x$ ,  $y$  *отсутствует*.
  - Если значения коэффициентов корреляции отрицательные - связь между признаками  $X$  и  $Y$  *обратная*.
  - Если значения коэффициентов корреляции положительные - связь между признаками  $X$  и  $Y$  *прямая* (положительная).
  - $r = 1.0$  (функциональная взаимосвязь, так как значению одного показателя соответствует только одно значение другого показателя и поэтому никакой вариации на диаграмме рассеяния не наблюдается);

- **Оценка коэффициента корреляции по шкале Чертока:**
  - 0,9- 1 очень сильная статистическая связь;
  - 0,9-0,7 сильная;
  - 0,7-0,5 средняя;
  - 0,5-0,3 умеренная;
  - Менее 0,3 – слабая.
  - Менее 0,1 –связь практически отсутствует.

# Коэффициент детерминации

- Является квадратом коэффициента корреляции зависимой и независимой переменных.
- Показывает, в какой степени изменчивость переменной отклика обусловлена (детерминирована) влиянием другой переменной.
- Обладает важным преимуществом по сравнению с коэффициентом корреляции. Допускается его усреднение для нескольких выборок. В отличие от коэффициента корреляции коэффициент детерминации линейно возрастает с увеличением силы связи.

# Коэффициент корреляции Пирсона (Pearson Correlation Coefficient)

- **Назначение:** используется для оценки силы и направления линейной связи между 2 метрическими переменными в одной выборке. Связь устанавливается между абсолютными значениями признаков.
- **Требования к выборке**
  - Интервальные или абсолютные шкалы измерений.
  - Нормальное распределение (или близкое к нормальному).
  - Линейность ассоциации.
  - Отсутствие выбросов.

# Коэффициент ранговой корреляции Спирмена (Spearman's Rho)

- **Назначение** – непараметрический тест, используемый для оценки силы линейной ассоциации между двумя переменными. При этом связь устанавливается не между самими переменными, а между рангами.
- **Требования к переменным**
  - Ранговая шкала.
  - Данные представляются в виде связанных пар.
  - Ассоциация между данными должна быть монотоническая – переменные увеличиваются или одна увеличивается, а другая уменьшается.

# Регрессионный анализ

- Линейная регрессия сходна, но не идентична линейной корреляции. Регрессионный анализ проводится, если корреляционный анализ выявил взаимосвязь между переменными.
- **Регрессионный анализ** позволяет предсказать или оценить значение (зависимой) переменной отклика по известным значениям одной или нескольких (независимых) предикторных переменных.
- **Виды регрессионного анализа**
  - **Простой** – используется одна предикторная переменная.
  - **Множественная** – несколько предикторных переменных.
  - **Логический** – переменная отклика является бинарной (двоичной) категориальной.
  - **Линейный** - переменная отклика непрерывна и линейно связана с независимой (независимыми) переменными. Как линейный так и логический анализ может быть простым или множественным.
  - **Нелинейный** – между переменными, которые связаны нелинейно и не могут быть трансформированы в линейную.

# Простая линейная регрессия

С помощью регрессионного анализа определяются параметры прямой, которая наилучшим способом предсказывает значение одной переменной на основании значения другой по **уравнению регрессии (регрессия y на x)**:

$$y = a + bx,$$

$$y = a + bx,$$

$y$  - зависимая или предикторная переменная;

$x$  – независимая, объясняющая;

$a$  – точка пересечения прямой с осью ординат, является начальной ординатой и даёт значение  $y$  при  $x=0$ , свободный член;

$b$  задаёт наклон линии регрессии – среднее изменение  $y$  при изменении  $x$  на единицу (угловой коэффициент линейной регрессии).



- **ANOVA** в отличие от множественного регрессионного анализа (работает с непрерывными предикторными переменными) использует категориальные предикторные переменные.
- **ANCOVA** – ковариационный анализ включает как непрерывные, так и категориальные предикторные переменные.

# Методы статистической обработки данных в зависимости от их типа и от задач

	Методы	
<b>Задачи</b>	<b>Параметрические (для количественных нормально распределенных признаков)</b>	<b>Непараметрические (для количественных признаков, независимо от вида распределения, и для порядковых и номинальных качественных признаков)</b>
Выполнение описательной статистики	Вычисление средних значений, сигмы и т.д.	Вычисление медиан, интерквартильных интервалов, пропорций
Сравнение двух независимых групп по одному признаку	<u>t-критерий Стьюдента</u> для независимых выборок (непарный)	<u>Критерии Манна-Уитни (U-тест)</u> , X-критерий Ван дер Вардена, Колмогорова-Смирнова, Вальда- Вольфовица, $\chi^2$ (критерий хи-квадрат)
Сравнение двух зависимых групп по одному признаку	<u>t-критерий Стьюдента</u> для зависимых выборок	<u>Критерий Вилкоксона</u> , критерий знаков, критерий МакНемара
Сравнение трех независимых групп и более по одному признаку	ANOVA	ANOVA по Краскелу-Уоллису, медианный критерий, критерий $\chi^2$
Сравнение трех и более зависимых групп по одному признаку	Критерий Кокрана	ANOVA по Фридмену, критерий Кокрана
Анализ взаимосвязи двух признаков	<u>Корреляционный анализ по Пирсону</u>	Критерий $\chi^2$ , <u>корреляционный анализ рангов по Спирмену</u> и др.
Одновременный анализ трех признаков и более	Регрессионный, дискриминантный, факторный, кластерный анализ	Логистический регрессионный анализ, анализ древовидных диаграмм и др.

Статистическая ошибка второго рода

- А. характеризует мощность теста.
- Б. является более критичной, чем ошибка первого рода.
- В. состоит в ошибке не обнаружить различия или связи, которые на самом деле существуют.
- Г. связана с нулевой гипотезой.
- Д. Верно всё вышеперечисленное.
- Е. **Верно А, В.**

Статистически значимыми являются результаты

- А. если эмпирическое р-значение больше критического уровня значимости.
- Б. если достигнутый уровень значимости превышает 95%.
- В. **если нулевая гипотеза отвергается.**
- Г. если доверительная вероятность равна критическому уровню значимости.
- Д. Верно А, В.
- Е. Верно В, Г.

Какой статистический критерий следует выбрать, если направление эффекта неизвестно?

- А. параметрический.
- Б. непараметрический.
- В. односторонний.
- Г. **двусторонний.**
- Д. одномерный.
- Е. многофакторный.

Критерий знаков Вилкоксона применяют для данных полученных в

- А. номинальной шкале.
- Б. порядковой шкале.
- В. интервальной шкале.
- Г. абсолютной шкале.
- Д. Верно В, Г.
- Е. **Верно Б, В, Г.**

Корреляционный анализ

- А. позволяет определить причины изменения переменных.
- Б. определяет тесноту (силы) статистической зависимости показателей.
- В. определяет направленность статистической зависимости показателей.
- Г. является методом описательной статистики.
- Д. Верно всё вышеперечисленное.
- Е. **Верно Б, В.**

Коэффициент корреляции

- А. является квадратом коэффициента детерминации.
- Б. показывает, в какой степени дисперсия переменной отклика обусловлена влиянием другой переменной.
- В. линейно возрастает с увеличением силы связи между переменными.
- Г. **не может быть больше плюс 1.**
- Д. Верно В, Г.
- Е. Верно Б, В.