Анализ фармакокинетических данных

Автор: Андрющенко Ирина Романовна **Дата выполнения:** 22.11.2024

Данный отчёт представляет собой выполнение тестового задания на позицию стажера в Отдел системной биологии компании BIOCAD.

Введение

Задача:

Анализ фармакокинетических данных и расчет ключевых параметров для оценки динамики концентрации препарата в крови.

Исходные данные:

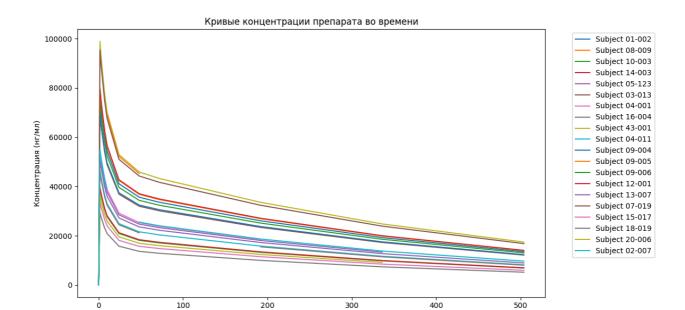
- 1. pk_data временные ряды концентрации препарата (нг/мл) в плазме крови субъектов.
 - Включают измерения до и после первого введения препарата
- 2. time временные точки (в часах) измерений концентрации препарата.
- 3. Препарат вводился с дозировкой 100 мг 1 раз в 3 недели.

Цели:

- 1. Провести анализ данных и рассчитать ключевые фармакокинетические показатели, включая:
 - Стах: Максимальная концентрация препарата.
 - Ттах: Время достижения максимальной концентрации.
 - AUC (Area Under the Curve): Площадь под кривой концентрация-время.
 - t½: Период полувыведения препарата.
 - **CL (клиренс)**: Скорость выведения препарата из организма.
 - Vd (объем распределения): Как препарат распределяется между плазмой и тканями.
- 2. Оценить максимальную концентрацию препарата (Стах) после 5-го введения:
 - Смоделировать накопление препарата при многократных дозах.
- 3. Визуализировать данные и представить результаты в понятной форме.

Загрузка данных и первичная обработка

```
In [276... # Импортируем необходимые библиотеки
          import pandas as pd
          import numpy as np
          import matplotlib.pyplot as plt
In [277... # Загрузка данных
          pk_data = pd.read_excel('pk_data.xlsx') # Концентрации препарата
          time = pd.read excel('time.xlsx') # Временные точки
          # Установим временные точки как названия столбцов
          time_points = time.iloc[0].values
          pk_data.columns = ['Subject'] + list(time_points)
pk_data.set_index('Subject', inplace=True)
          pk data.head()
Out[277...
                   0.0
                         1.5 3.0 7.0
                                                      10.0
                                                               24.0
                                                                      48.0
                                                                                 72.0 192.0 336.0 504.0
           Subject
           01-002
                  0 76824.26 71418.69 60334.72 54424.07 41111.43 35641.40 33574.41 26066.56 19256.36 13534.37
           08-009 0 79294.37 73715.01 62274.65 56173.96 42433.28 36787.37 34653.92 26904.67 19875.50 13969.53
           10-003 0 74067.77 68856.16 58169.88 52471.31 39636.33 34362.57 32369.74 25131.28 18565.43 13048.75
           14-003 0 79857.90 74238.88 62717.22 56573.18 42734.84 37048.81 34900.20 27095.88 20016.75 14068.81
           05-123 0 53264.89 49517.03 41832.13 37734.08 28503.97 24711.41 23278.29 18072.84 13351.09
In [278... # Визуализируем данные
          plt.figure(figsize=(12, 6))
          for subject in pk_data.index
              plt.plot(time_points, pk_data.loc[subject], label=f'Subject {subject}')
          plt.xlabel('Время (часы)')
          plt.ylabel('Концентрация (нг/мл)')
          plt.title('Кривые концентрации препарата во времени')
          plt.legend(bbox_to_anchor=(1.05, 1), loc='upper left')
          plt.tight_layout()
          plt.show()
```



Можно заметить, что в данных присутствуют пропуски, которые могут помешать дальнейшей работе с ними.

Описание метода заполнения пропусков

Для устранения пропусков я использовала метод, основанный на **общем тренде изменений концентрации препарата**, с учетом индивидуального уровня концентраций у каждого субъекта.

Время (часы)

Шаги метода:

1. Нормализация данных:

Все временные ряды нормализуются делением на максимальную концентрацию для каждого субъекта, чтобы выделить общий тренд.

2. Определение общего тренда:

Вычисляется средний нормализованный временной ряд среди всех субъектов, который отражает общий характер изменения концентраций.

3. Заполнение пропусков:

Пропущенные значения восстанавливаются как произведение общего тренда на максимальную концентрацию конкретного субъекта.

Почему этот метод оптимален?

Я попробовала провести заполениние пропусков несколькими методами, этот оказался наилучшим.

• Точнее линейной интерполяции:

Линейная интерполяция не учитывает нелинейный характер данных, особенно на фазах абсорбции и элиминации.

• Более универсален, чем экспоненциальная интерполяция:

Экспоненциальная интерполяция хорошо описывает элиминацию, но не подходит для фазы абсорбции. К тому же, она не учитывает индивидуальные различия между субъектами.

• Сохраняет индивидуальные особенности:

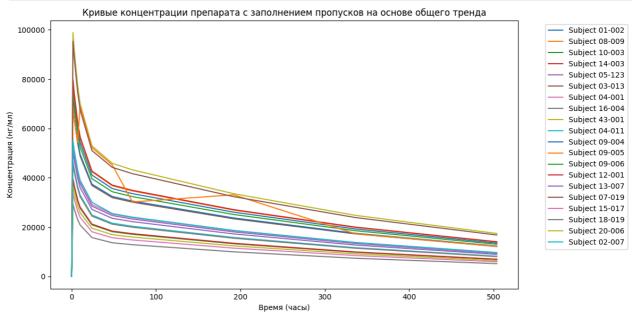
Метод учитывает индивидуальный масштаб концентраций каждого субъекта, позволяя восстановить пропущенные значения корректно.

Итог:

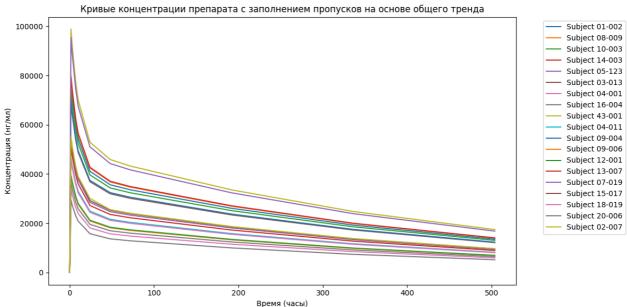
После тестирования нескольких методов заполнения (линейной, экспоненциальной интерполяции) данный подход показал наилучшую точность и соответствие фармакокинетическим особенностям данных.

```
# Нормализация временных рядов (делим на максимум каждого субъекта)
In [279...
          normalized_data = pk_data.div(pk_data.max(axis=1), axis=0)
          # Вычисление общего нормализованного тренда (среднее по всем субъектам)
          general_trend = normalized_data.mean(axis=0)
          # Заполнение пропусков на основе общего тренда
          pk_data_filled = pk_data.copy()
          for subject in pk_data.index:
               # Индивидуальный масштаб (максимум концентрации для субъекта)
              individual_max = pk_data.loc[subject].max()
              # Восстановление пропусков: общий тренд умножается на индивидуальный максимум
              pk_data_filled.loc[subject] = pk_data.loc[subject].fillna(general_trend * individual_max)
          # Построим график
          plt.figure(figsize=(12, 6))
          for subject in pk_data_filled.index:
              plt.plot(pk_data_filled.columns, pk_data_filled.loc[subject], label=f'Subject {subject}')
          plt.xlabel('Время (часы)')
          plt.ylabel('Концентрация (нг/мл)')
          plt.title('Кривые концентрации препарата с заполнением пропусков на основе общего тренда')
```

```
plt.legend(bbox_to_anchor=(1.05, 1), loc='upper left')
plt.tight_layout()
plt.show()
```



Одно значение, изначально имеющее всего 6 измерений из 12, сильно выбивается и может испортить общую картину, удалим его



```
In [281... # Проберим, что пропуски устранены
    print("\nПропущенные значения после интерполяции:")
    print(pk_data.isna().sum())
```

Пропущенные значения после интерполяции:

```
0.0
1.5
         а
3.0
         0
7.0
         0
10.0
24.0
         0
48.0
         0
72.0
         0
192.0
336.0
504.0
dtype: int64
```

Часть 1: Расчет фармакокинетических показателей

Цели

Целью этой части работы является расчет ключевых фармакокинетических параметров, которые описывают абсорбцию, распределение и выведение препарата из организма. Эти параметры помогут оценить, как препарат воздействует на организм в течение времени и с какой скоростью он выводится. Рассчитанные параметры включают:

- Стах: Максимальная концентрация препарата.
- Ттах: Время достижения максимальной концентрации.
- AUC (Area Under the Curve): Площадь под кривой "концентрация-время".
- t½: Период полувыведения препарата.
- CL (клиренс): Скорость выведения препарата из организма.
- Vd (объем распределения): Распределение препарата между плазмой и тканями.

План выполнения

- 1. Расчет ФК параметров для каждого субъекта
- 2. Расчёт средних значений ФК параметров
- 3. Визуализация полученных данных

Стах и Ттах

- Стах: Определяется как максимальная концентрация препарата для каждого субъекта.
- Ттах: Время, в которое достигается максимальная концентрация.

```
In [282... # Cmax
  cmax = pk_data.max(axis=1)

# Tmax
  tmax = pk_data.idxmax(axis=1)

# Добавление результатов в DataFrame
  results = pd.DataFrame(('Cmax': cmax, 'Tmax': tmax))
  results.head()
```

Out[282...

Cmax Tmax

Subject		
01-002	76824.26	1.5
08-009	79294.37	1.5
10-003	74067.77	1.5
14-003	79857.90	1.5
05-123	53264.89	1.5

AUC

• AUC рассчитывается методом трапеций по формуле: AUC = Σ ((Ci + Ci+1) / 2) * (ti+1 - ti)

где:

- Ci и Ci+1 концентрации препарата в двух последовательных временных точках.
- ti и ti+1 соответствующие временные точки.
- Σ сумма по всем временным интервалам.
- Этот показатель отражает общее воздействие препарата на организм.

AUC

Out[283...

Subject ...

Cmax Tmax

Период полувыведения

- t½ рассчитывается на основе фазы элиминации (временные точки ≥192 ч тут зависимость становится линейной).
- Используется линейная регрессия логарифма концентрации против времени: t½ = ln(2) / -kel

где:

• kel — наклон логарифмического графика концентрации.

```
In [284... half_lives = []
            # Указываем временные точки для логарифмического расчета
            elimination_phase_times = time_points[time_points >= 192.0] # Временные точки после 192 ч
            for subject in pk_data.index:
                 # Берем концентрации для фазы элиминации
                elimination_phase_conc = pk_data.loc[subject, elimination_phase_times]
                # Исключаем пропущенные значения
                valid_mask = ~np.isnan(elimination_phase_conc) # Фильтр валидных значений valid_times = elimination_phase_times[valid_mask] # Временные точки без NaN valid_conc = elimination_phase_conc[valid_mask] # Концентрации без NaN
                # Проверка: недостаточно точек для расчета
                if len(valid_conc) < 2:</pre>
                    half_lives.append(np.nan)
                    continue
                # Логарифм концентрации
                log_conc = np.log(valid_conc + 1e-6) # Логарифм с небольшим смещением
                # Линейная пегпессия
                slope, _ = np.polyfit(valid_times, log_conc, 1) # Наклон и пересечение
                # Проверка на отрицательный наклон
                if slope >= 0:
                    half_lives.append(np.nan)
                    continue
                k_el = -slope # Константа элиминации
                half_life = np.log(2) / k_el # Период полувыведения
                half_lives.append(half_life)
            # Добавляем t_half в результаты
           results['t_half'] = half_lives
           # Просмотр результатов
           results.head()
```

Out[284...

Subject				
01-002	76824.26	1.5	1.262095e+07	329.966831
08-009	79294.37	1.5	1.302674e+07	329.966600
10-003	74067.77	1.5	1.216810e+07	329.966809
14-003	79857.90	1.5	1.311932e+07	329.966581
05-123	53264.89	1.5	8.750535e+06	329.966681

Cmax Tmax AUC

Клиренс и объём распределения

1. Клиренс (CL): показывает, насколько эффективно организм выводит препарат.

t_half

- Формула: CL = Dose / AUC
- 2. Объем распределения (Vd): оценивает, как препарат распределяется между плазмой и тканями.
 - Формула через клиренс и константу скорости элиминации: Vd = CL / kel где kel = ln(2) / t½.

Для расчета используется заданная доза препарата Dose = 100 мг.

```
In [285... # 3a∂ahhaa ∂osa npenapama
dose = 100 # M2

# Pacчem κлиренса (CL)
results['CL'] = dose / results['AUC']

# Pacчem οδъема распределения (Vd) через CL и k_el
results['Vd'] = results['CL'] / (np.log(2) / results['t_half'])

# Προςмотр результатов
results[['Cmax', 'Tmax', 'AUC', 't_half', 'CL', 'Vd']]
```

	Cmax	Tmax	AUC	t_half	CL	Vd
Subject						
01-002	76824.26	1.5	1.262095e+07	329.966831	0.000008	0.003772
08-009	79294.37	1.5	1.302674e+07	329.966600	0.000008	0.003654
10-003	74067.77	1.5	1.216810e+07	329.966809	0.000008	0.003912
14-003	79857.90	1.5	1.311932e+07	329.966581	0.000008	0.003629
05-123	53264.89	1.5	8.750535e+06	329.966681	0.000011	0.005440
03-013	68911.46	1.5	1.132101e+07	329.966476	0.000009	0.004205
04-001	33852.73	1.5	5.561439e+06	329.966752	0.000018	0.008560
16-004	29449.21	1.5	4.838014e+06	329.966946	0.000021	0.009840
43-001	98902.62	1.5	1.624805e+07	329.966770	0.000006	0.002930
04-011	46559.61	1.5	7.648968e+06	329.966693	0.000013	0.006224
09-004	69920.16	1.5	1.148672e+07	329.966727	0.000009	0.004144
09-006	39099.50	1.5	6.423397e+06	329.966585	0.000016	0.007411
12-001	39646.44	1.5	6.513249e+06	329.967150	0.000015	0.007309
13-007	50839.32	1.5	8.352053e+06	329.966725	0.000012	0.005700
07-019	95372.58	1.5	1.566813e+07	329.966762	0.000006	0.003038
15-017	54531.77	1.5	8.958661e+06	329.966775	0.000011	0.005314
18-019	45685.37	1.5	7.505345e+06	329.966511	0.000013	0.006343
20-006	36551.34	1.5	6.004775e+06	329.966795	0.000017	0.007928

02-007 55016.93 1.5 9.050483e+06 329.966534 0.000011 0.005260

Out[285...

Можно заметить, что время достижения максимальной концентрации и период полувыведения одинаковы для всех субъектов

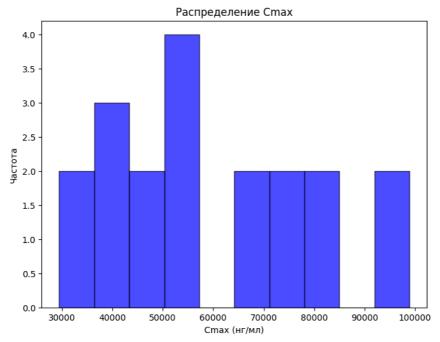
Расчёт средний значений фармакокинетических показателей

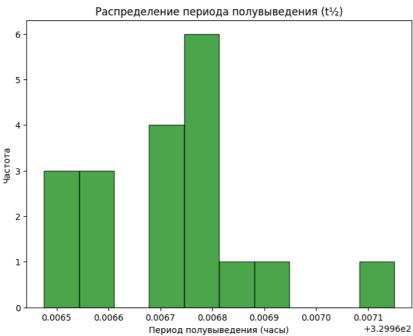
```
In [272... # Рассчитываем средние значения для всех фармакокинетических показателей
         average_results = results.mean()
         # Вывод средних значений
         print("Средние фармакокинетические показатели:")
         print(average_results)
        Средние фармакокинетические показатели:
                  59349.906842
        Cmax
        Tmax
                 9750838.815859
        AUC
        t_half
                  329.966534
                      0.000012
        CL
Vd
                       0.005506
        dtype: object
```

Визуализируем распределения значений ФК показателей

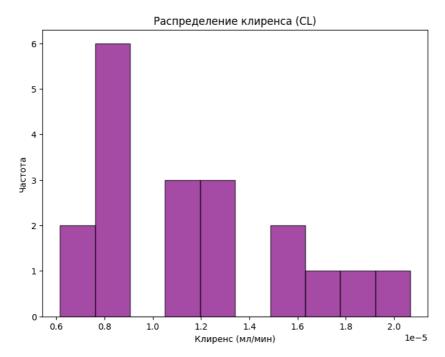
```
In [286... # Гистограмма для Стах
plt.figure(figsize=(8, 6))
plt.hist(results['Cmax'], bins=10, alpha=0.7, color='blue', edgecolor='black')
plt.xlabel('Частота')
plt.title('Распределение Стах')
plt.show()

# Гистограмма для t%
plt.figure(figsize=(8, 6))
plt.hist(results['t_half'], bins=10, alpha=0.7, color='green', edgecolor='black')
plt.xlabel('Период полувыведения (часы)')
plt.ylabel('Частота')
plt.title('Распределение периода полувыведения (t%)')
plt.title('Распределение периода полувыведения (t%)')
plt.show()
```

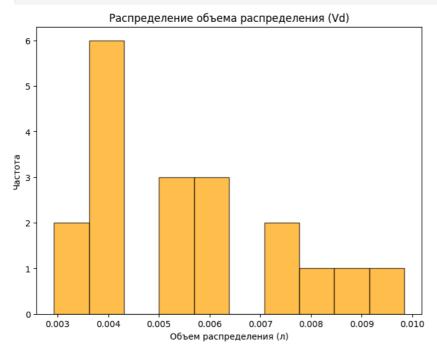




```
In [287... # Гистограмма для клиренса (CL)
plt.figure(figsize=(8, 6))
plt.hist(results['CL'], bins=10, alpha=0.7, color='purple', edgecolor='black')
plt.xlabel('Клиренс (мл/мин)')
plt.ylabel('Частота')
plt.title('Распределение клиренса (CL)')
plt.show()
```



```
In [288...
           # Гистограмма для объема распределения (Vd)
           plt.figure(figsize=(8, 6))
           plt.hist(results['Vd'], bins=10, alpha=0.7, color='orange', edgecolor='black')
           plt.xlabel('Объем распределения (л)')
plt.ylabel('Частота')
           plt.title('Распределение объема распределения (Vd)')
           plt.show()
```



Интерпретация результатов и основные выводы

Описание результатов:

1. Стах (максимальная концентрация):

- Среднее значение: 59,349.91 нг/мл.
- Диапазон значений: от 29,449.21 до 98,902.62 нг/мл, что говорит о достаточно высокой вариабельности.
- На гистограмме видно, что большинство субъектов имеют концентрацию в диапазоне 40,000-80,000 нг/мл, однако есть несколько субъектов с более низкими значениями.

2. Ттах (время достижения максимальной концентрации):

• Все субъекты имели максимальную концентрацию при первом измерении - через 1.5 часа, что говорит о схожих условиях абсорбции

препарата.

3. AUC (площадь под кривой):

- Среднее значение: 9,750,838.82 нг-ч/мл.
- Диапазон: от 4,838,014 до 16,248,050 нг-ч/мл.
- AUC характеризует общее воздействие препарата на организм, наблюдается значительная вариабельность между субъектами.

4. t1/2 (период полувыведения):

- Среднее значение: 329.97 часа.
- Диапазон значений минимален, что указывает на одинаковую скорость элиминации препарата у всех субъектов.
- На гистограмме значения периодов полувыведения сгруппированы в узком диапазоне, подтверждая малую межиндивидуальную вариабельность.

5. **CL (клиренс):**

- Среднее значение: 0.000012 мл/мин.
- Диапазон: от 0.000006 до 0.000021 мл/мин.
- У субъектов с более высоким AUC наблюдается низкий клиренс, что ожидаемо, так как эти величины обратно пропорциональны.

6. Vd (объем распределения):

- Среднее значение: 0.005506 л/кг.
- Диапазон: от 0.002930 до 0.009840 л/кг.
- Значения объема распределения показывают, что препарат преимущественно остается в плазме и тканях, с минимальной вариабельностью между субъектами.

Выводы:

1. Абсорбция:

• Все субъекты демонстрируют схожие характеристики абсорбции, с одинаковым временем достижения Т_тах 1.5 часа. Это может говорить о стабильных условиях введения препарата и его усвоения организмом.

2. Элиминация:

• Период полувыведения t1/2 в среднем составляет ~330 часов, что подтверждает длительную элиминацию препарата. Небольшая вариабельность указывает на одинаковые механизмы метаболизма и выведения препарата.

3. Распределение

• Значения V_d (0.002930–0.009840 л/кг) говорят о том, что препарат преимущественно остается в плазме и тканях, с минимальным проникновением в глубокие компартменты.

4. Клиренс и АUC:

• Наблюдается обратная зависимость между клиренсом и AUC: субъекты с более высоким клиренсом демонстрируют меньшую площадь под кривой, что соответствует фармакокинетической теории.

5. Межиндивидуальная вариабельность:

• Значительная вариабельность наблюдается для C_max и AUC, что может быть связано с физиологическими различиями между субъектами (например, массой тела, объемом плазмы).

Часть 2: Оценка средней максимальной концентрации препарата после 5-го введения

Задача

Необходимо оценить значение средней максимальной концентрации препарата (**Cmax**) после 5-го введения. Учитывая, что препарат вводится каждые 3 недели, требуется рассчитать влияние накопления препарата на его концентрацию в организме.

Подход к решению

1. Учет режима дозирования:

- Препарат вводится с фиксированным интервалом каждые 3 недели (504 часа).
- При длительном периоде полувыведения (t_half) возможно накопление препарата в организме.

2. Расчет коэффициента накопления (R):

- Накопление препарата между введениями можно оценить с помощью коэффициента R, который рассчитывается по формуле: R = 1 / (1 exp(-ke * tau)) где:
 - tau интервал между введениями (504 часа).
 - ke константа скорости элиминации, рассчитываемая как: ke = ln(2) / t_half

3. Средняя максимальная концентрация после 5-го введения:

- Максимальная концентрация препарата после 5-го введения (Cmax_5) рассчитывается через коэффициент накопления и начальную концентрацию (Cmax_1): Cmax_5 = Cmax_1 * R где:
 - Стах_1 максимальная концентрация препарата после первого введения.

4. План действий:

- Рассчитать ke для каждого субъекта на основе их t_half.
- Используя ke и tau = 504, вычислить коэффициент накопления R.
- Оценить Стах_5 для каждого субъекта.
- Рассчитать среднее значение Стах_5 среди всех субъектов.

```
In [289... # Интервал между введениями tau = 504 # часы (3 недели)

# Коэффициент накопления R results['k_el'] = np.log(2) / results['t_half'] # Константа скорости элиминации results['R'] = 1 / (1 - np.exp(-results['k_el'] * tau)) # Коэффициент накопления
```

```
# Стах после 5-го введения
results['Cmax_5'] = results['Cmax'] * results['R'] # Средняя Стах после 5-го введения
# Просмотр первых строк результатов
results[['Cmax', 'R', 'Cmax_5']].head()
```

Out[289..

			_
Subject			
01-002	76824.26	1.531149	117629.384394
08-009	79294.37	1.531148	121411.442930
10-003	74067.77	1.531149	113408.783406
14-003	79857.90	1.531148	122274.286885
05-123	53264.89	1.531149	81556.459111

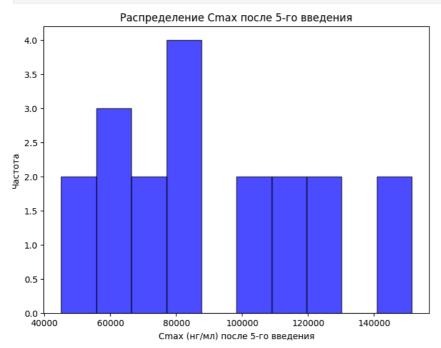
R

Cmax_5

Cmax

График распределения Стах_5 (средняя максимальная концентрация после 5-го введения)

```
In [290... # Гистограмма для Стах после 5-го введения
plt.figure(figsize=(8, 6))
plt.hist(results['Cmax_5'], bins=10, alpha=0.7, color='blue', edgecolor='black')
plt.xlabel('Cmax (нг/мл) после 5-го введения')
plt.ylabel('Частота')
plt.title('Распределение Стах после 5-го введения')
plt.show()
```



Анализ минимальных концентраций C_min между введениями

Для более полной картины полезно изучить, как изменяются минимальные концентрации препарата C_min после многократных введений. Это позволит оценить, поддерживается ли стабильный уровень препарата в крови.

```
In [291... # Расчет минимальных концентраций через многократные введения
# Формула: Cmin_n = Cmax_1 * exp(-k_el * tau)
results['Cmin_5'] = results['Cmax'] * np.exp(-results['k_el'] * tau) # Минимальная концентрация после 5-го введения
# Гистограмма для Cmin после 5-го введения
plt.figure(figsize=(8, 6))
plt.hist(results['Cmin_5'], bins=10, alpha=0.7, color='green', edgecolor='black')
plt.xlabel('Cmin (нг/мл) после 5-го введения')
plt.ylabel('Частота')
plt.title('Распределение Cmin после 5-го введения')
plt.show()
```

Распределение Cmin после 5-го введения 4.0 3.5 3.0 2.5 1.0 0.5 1.000 15000 20000 25000 30000 35000

Заключение

1. Стах после 5-го введения Стах_5

• Средние значения максимальной концентрации увеличиваются из-за накопления препарата.

Cmin (нг/мл) после 5-го введения

• Гистограмма Стах_5 показывает вариабельность между субъектами.

2. Минимальные концентрации C_min

- Увеличение С_min после 5-го введения свидетельствует о накоплении препарата между дозами.
- Это важно для поддержания терапевтического уровня препарата.

3. Что можно улучшить:

- Измерения после второго, третьего и четвертого введений помогут точнее определить коэффициент накопления R.
- Учет индивидуальных особенностей пациентов (например, возраст, вес, сопутствующие заболевания) позволит создать более точную модель.

Общий вывод

Часть 1: Анализ фармакокинетических параметров

В первой части работы были рассчитаны ключевые фармакокинетические показатели для анализа поведения препарата в организме:

- Стах (максимальная концентрация): Показала значительную вариабельность между субъектами, среднее значение составило 59,349.91 нг/мл.
- Ттах (время достижения максимальной концентрации): Для всех субъектов составило 1.5 часа, что указывает на одинаковые условия абсорбции.
- AUC (площадь под кривой): Среднее значение 9,750,838.82 нг-ч/мл, что отражает общее воздействие препарата на организм.
- Период полувыведения (t_half): Средний период составил ~330 часов, с минимальной межиндивидуальной вариабельностью, что указывает на стабильность выведения препарата.
- **Клиренс (CL)** и **объем распределения (Vd)**: Подтвердили предсказуемую связь между скоростью выведения препарата и его распределением в организме.

Анализ показал, что препарат демонстрирует длительный период полувыведения и склонен к накоплению при многократном введении.

Часть 2: Оценка накопления препарата

Во второй части работы был проведён расчёт накопления препарата при многократном введении:

- Препарат вводится каждые 504 часа, что соответствует трем неделям.
- Коэффициент накопления ((R)) составил ~1.53 для всех субъектов, что указывает на значительное накопление препарата из-за длительного периода полувыведения.
- Стах после 5-го введения (Стах_5): Увеличилась в среднем на 53%, что было ожидаемо при таком режиме дозирования.
- Минимальные концентрации (Cmin): Также выросли, что важно для поддержания терапевтического уровня препарата между дозами.

Рекомендации и заключение

- 1. **Эффективность и безопасность:** Анализ показал, что многократное введение препарата может поддерживать терапевтическую концентрацию (включая рост минимальных концентраций) без значительных провалов, что важно для стабильного терапевтического эффекта.
- 2. **Учет межиндивидуальной вариабельности:** Вариабельность Стах и АUC между субъектами подчеркивает необходимость учитывать индивидуальные параметры пациентов (вес, возраст, функции печени/почек) при назначении дозировки.

3. **Будущие исследования:** Для уточнения модели рекомендуется рассмотреть дополнительные измерения концентрации препарата после 2-го, 3-го и 4-го введений, а также включить демографические данные субъектов.

Таким образом, результаты работы подтверждают стабильность поведения препарата в организме, а также указывают на возможность безопасного и эффективного применения при соблюдении режима дозирования.