**Руководство по использованию программы для анализа доверительных интервалов с расчетом по методу Кокса (Cox).**

Скрипт разработан для планирования исследований бентоса. По уже имеющейся выборке можно спрогнозировать, какое количество проб требуется собрать в новом месте для достижения результата с определенной точность. Метод расчета доверительного интервала по Cox предполагает, что данные в выборке имеют логнормальное распределение.

**Описание программы**

Программа анализирует зависимость точности измерений (ширины доверительных интервалов) от объема данных для трех уровней значимости (70%, 90%, 95%). Основной результат - расчет количества проб, необходимых для достижения заданной точности.

**Требования**

* R (версия 4.0+)
* Пакеты R:
  + ggplot2
  + dplyr

Установка пакетов:

R

install.packages(c("ggplot2", "dplyr"))

**Инструкция по запуску**

1. **Подготовка данных**:
   * Создайте файл sample.csv в рабочей директории
   * Данные должны быть в одном столбце с заголовком "x"
   * Пример формата:

x

1.5

2.3

0.8

...

1. **Запуск программы**:
   * Скопируйте код в файл analysis\_script.R
   * Выполните в R:

R

source("analysis\_script.R")

**Алгоритм работы**

1. **Загрузка данных**:
   * Чтение файла sample.csv
   * Расчет среднего значения
2. **Подготовка данных**:
   * Создание 1000 перестановок исходных данных
   * Логарифмирование значений (с обработкой нулевых значений)
3. **Расчет доверительных интервалов**:
   * Использование метода Кокса для расчета интервалов
   * Три уровня значимости: 70%, 90%, 95%
4. **Анализ результатов**:
   * Расчет средних значений ширины интервалов
   * Степенная регрессия для моделирования зависимости
   * Визуализация результатов
5. **Сохранение результатов**:
   * Уравнения регрессии (nls\_regression\_equations)
   * Таблица с предсказанным количеством проб (predicted\_number\_of\_samples.txt)
   * График зависимости (ggplot2)

**Интерпретация результатов**

**Основные выходные файлы**

1. predicted\_number\_of\_samples.txt:
   * Таблица с количеством проб для достижения заданной точности
   * Пример:

text

Инт.% 70% 90% 95%

20 50 85 110

30 30 55 70

...

* + Где:
    - "Инт.%" - требуемая точность (ширина интервала в % от среднего)
    - Остальные столбцы - необходимое количество проб для каждого уровня значимости

1. nls\_regression\_equations:
   * Уравнения степенной регрессии для каждого уровня значимости
   * Пример:

text

mean\_interval70: y = 120.5 \* x^-0.85 (R² = 0.98, AIC = 1500)

...

1. **График зависимости**:
   * Визуализация зависимости количества проб от ширины интервала
   * Точки - реальные данные, линии - предсказания модели

**Настройка параметров**

1. **Изменение уровней значимости**:
   * В коде измените z-значения в вызовах calculate\_interval():

R

interval70 <- calculate\_interval(log\_data, 1.04, data) # 70%

interval90 <- calculate\_interval(log\_data, 1.64, data) # 90%

interval95 <- calculate\_interval(log\_data, 1.96, data) # 95%

1. **Изменение диапазона точности**:
   * В конце скрипта измените значения для предсказания:

R

x\_values <- c(20, 30, 40, 50, 60) # Требуемые значения точности

**Пример вывода**

Уравнения регрессии:

mean\_interval70: y = 120.5 \* x^-0.85 (R² = 0.98, AIC = 1500)

mean\_interval90: y = 200.3 \* x^-0.92 (R² = 0.97, AIC = 1600)

mean\_interval95: y = 250.1 \* x^-0.95 (R² = 0.96, AIC = 1700)

Предсказанное количество проб:

Инт.% 70% 90% 95%

20 50 85 110

30 30 55 70

40 20 40 50

50 15 30 40

60 10 25 30

**Заключение**

Программа позволяет:

1. Определить необходимое количество проб для достижения заданной точности
2. Сравнить поведение доверительных интервалов при разных уровнях значимости
3. Визуализировать зависимость точности от объема данных