

Санкт-Петербургский Политехнический университет Петра  
Великого

Отчет по лабораторной работе №2

**Анализ выбросов в распределениях**

Студент:	Швачко Никита Андреевич
Преподаватель:	Баженов Александр Николаевич
Группа:	5030102/20202

Санкт-Петербург 2025

## 1 Формулировка задания и его формализация

Для анализа двухмерных нормальных распределений и смесей нормальных распределений необходимо:

1. Сгенерировать выборки размерами 20, 60 и 100 элементов для нормального двумерного распределения  $N(x, y, 0, 0, 1, 1, \rho)$  с коэффициентами корреляции  $\rho = 0.0, 0.5, 0.9$ .
2. Каждая выборка генерируется 1000 раз, после чего вычисляются следующие статистики:
  - Среднее значение для коэффициентов корреляции Пирсона, Спирмена и квадратного коэффициента корреляции.
  - Среднее значение квадрата этих коэффициентов.
  - Дисперсия этих коэффициентов.
3. Повторить все вычисления для смеси нормальных распределений:

$$f(x, y) = 0.9N(x, y, 0, 0, 1, 1, 0.9) + 0.1N(x, y, 0, 0, 10, 10, -0.9)$$

4. Изобразить сгенерированные точки на плоскости и нарисовать эллипс равновероятности для каждого случая.

## 2 Сгенерированные точки и эллипсы равновероятности

Для каждой выборки была выполнена генерация точек и построение эллипсов равновероятности. Ниже представлены графики сгенерированных точек для различных коэффициентов корреляции.

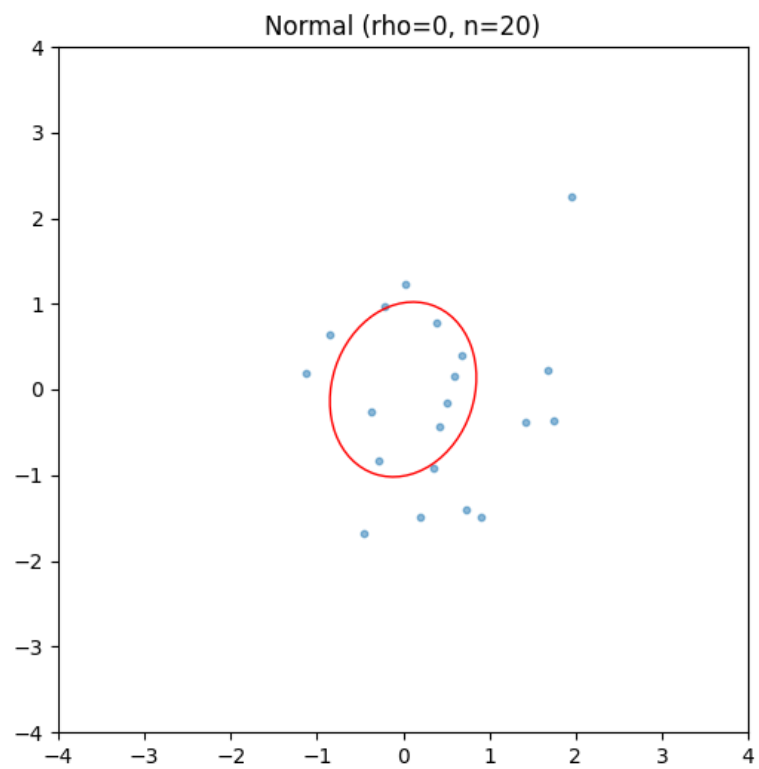


Рис. 1: Генерация точек для  $\rho = 0.0$  и эллипс равновероятности для  $n = 20$

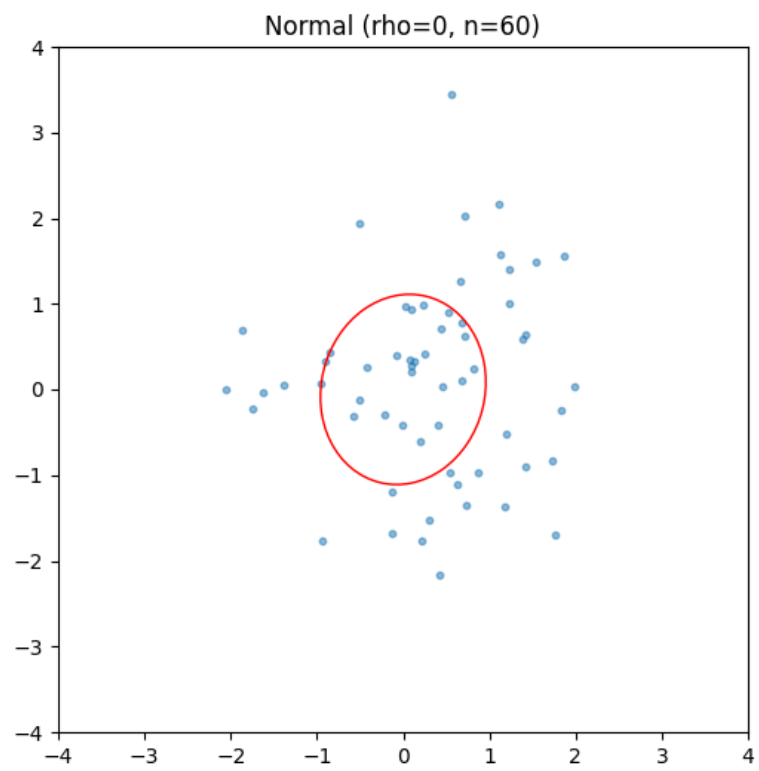


Рис. 2: Генерация точек для  $\rho = 0.0$  и эллипс равновероятности для  $n = 60$

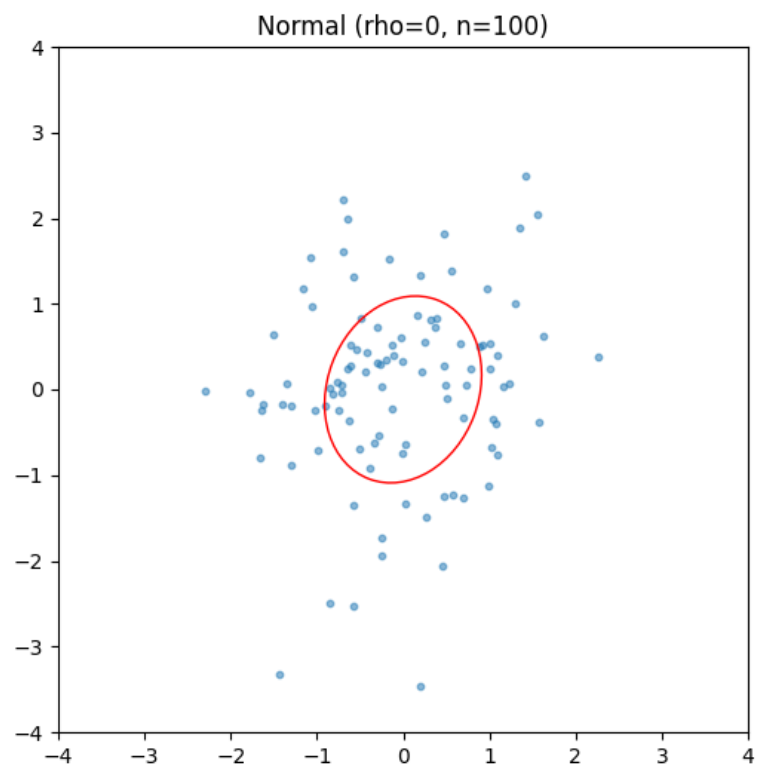


Рис. 3: Генерация точек для  $\rho = 0.0$  и эллипс равновероятности для  $n = 100$

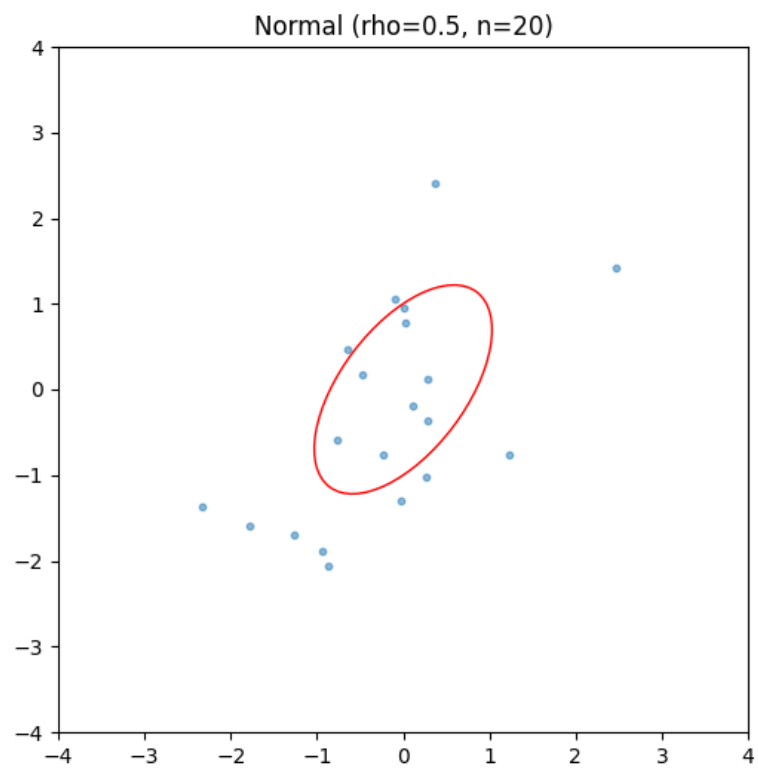


Рис. 4: Генерация точек для  $\rho = 0.5$  и эллипс равновероятности для  $n = 20$

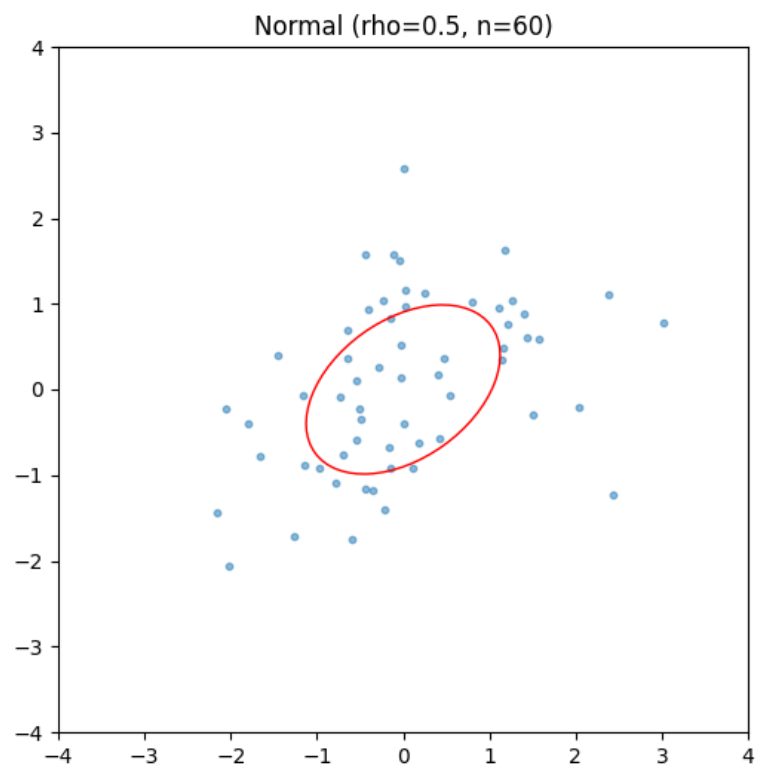


Рис. 5: Генерация точек для  $\rho = 0.5$  и эллипс равновероятности для  $n = 60$

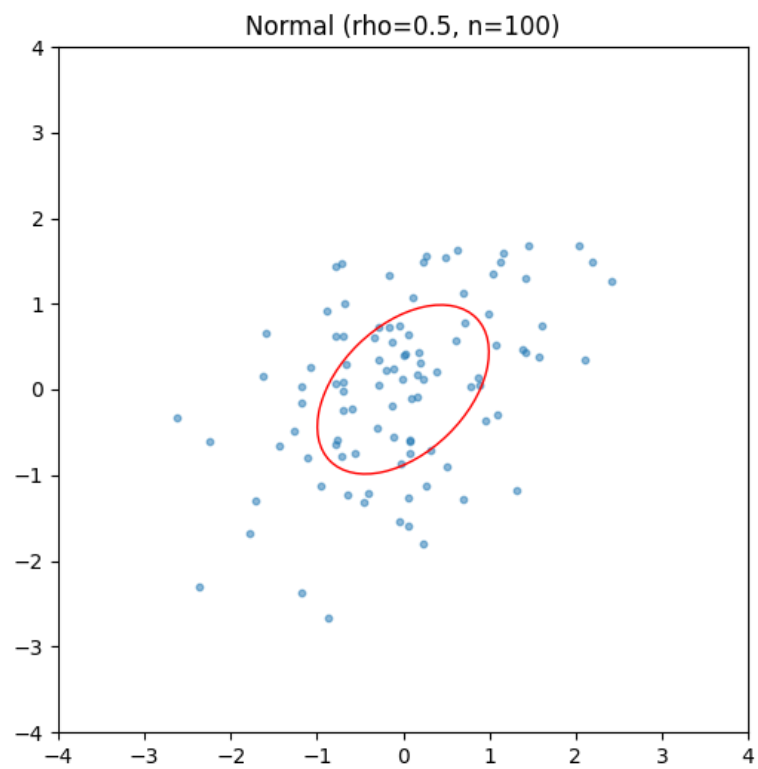


Рис. 6: Генерация точек для  $\rho = 0.5$  и эллипс равновероятности для  $n = 100$



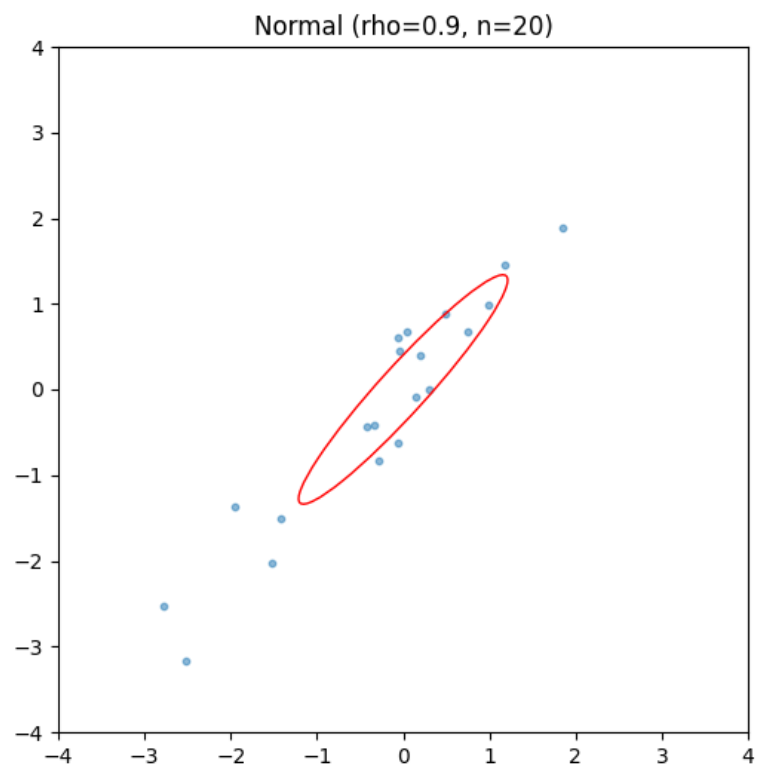


Рис. 7: Генерация точек для  $\rho = 0.9$  и эллипс равновероятности для  $n = 20$

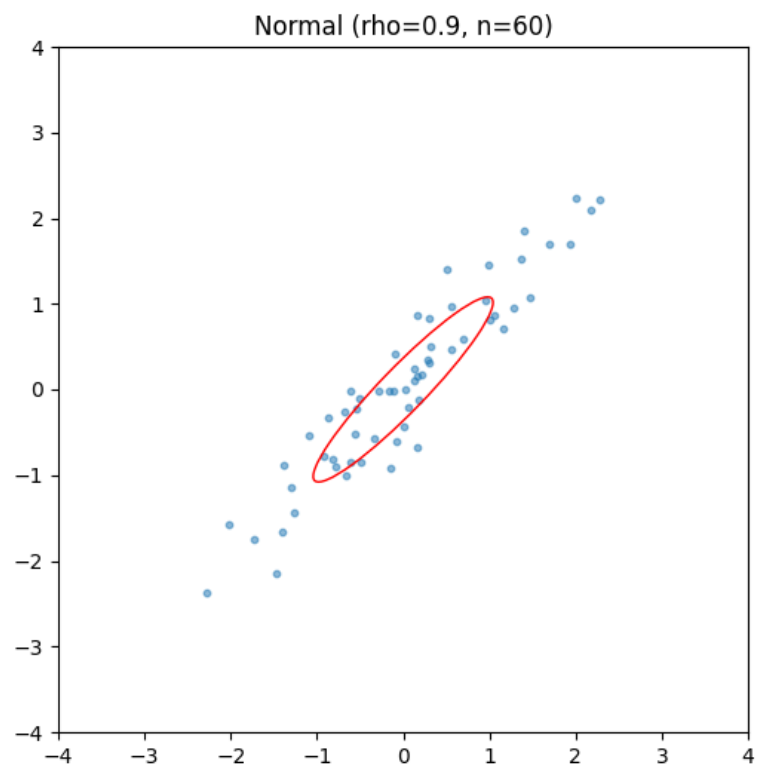


Рис. 8: Генерация точек для  $\rho = 0.9$  и эллипс равновероятности для  $n = 60$

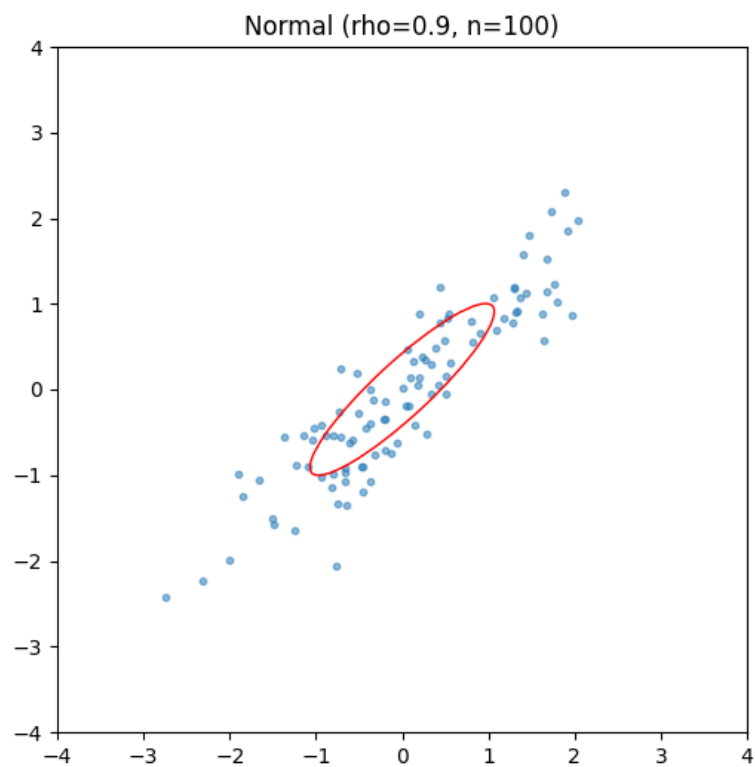


Рис. 9: Генерация точек для  $\rho = 0.9$  и эллипс равновероятности для  $n = 100$

### 3 Результаты для смеси нормальных распределений

Также были выполнены вычисления для смеси нормальных распределений. Графики точек и эллипсов равновероятности для смеси представлены ниже.

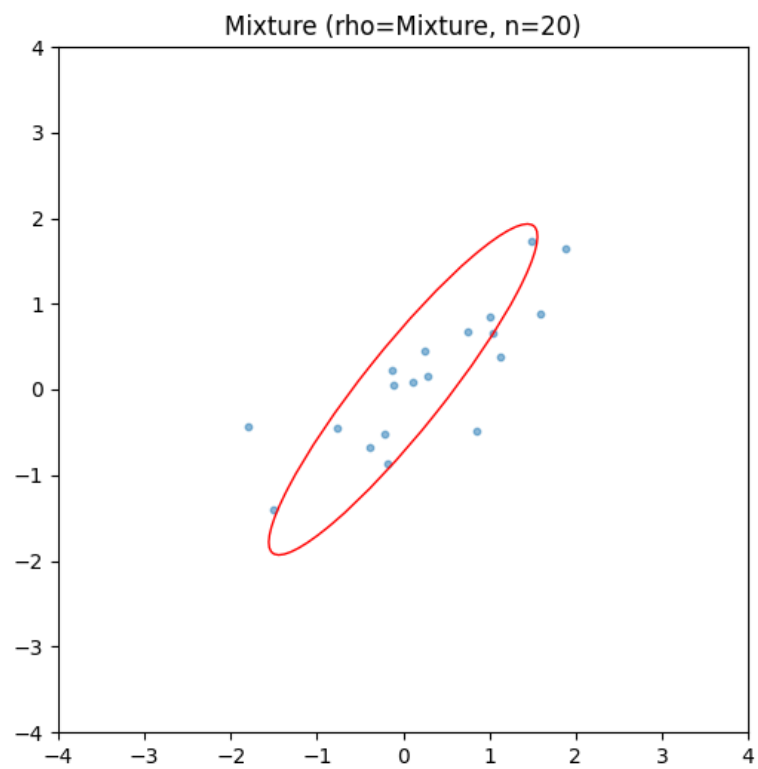


Рис. 10: Генерация точек для смеси с эллипсом равновероятности для  $n = 20$

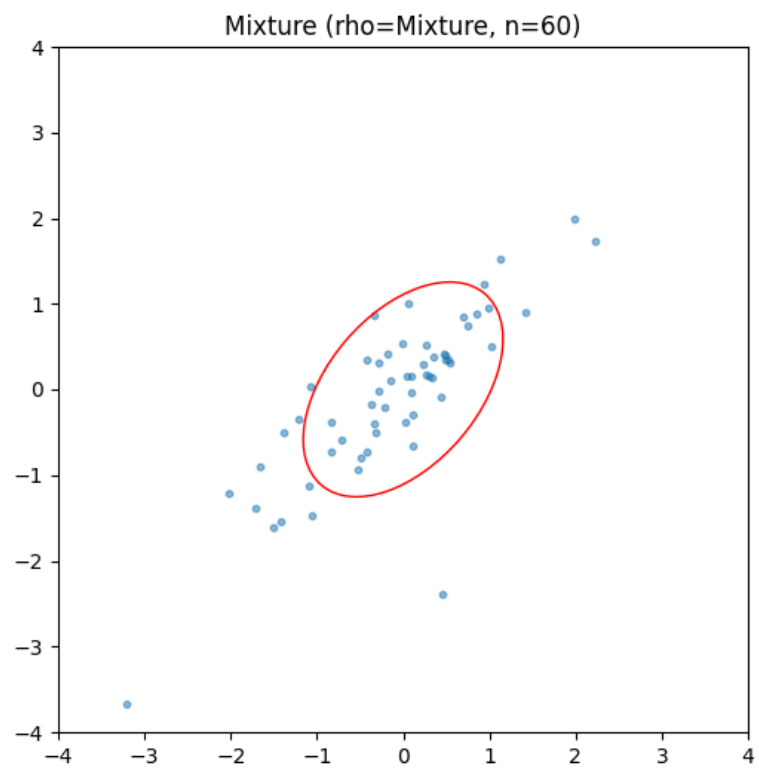


Рис. 11: Генерация точек для смеси с эллипсом равновероятности для  $n = 60$

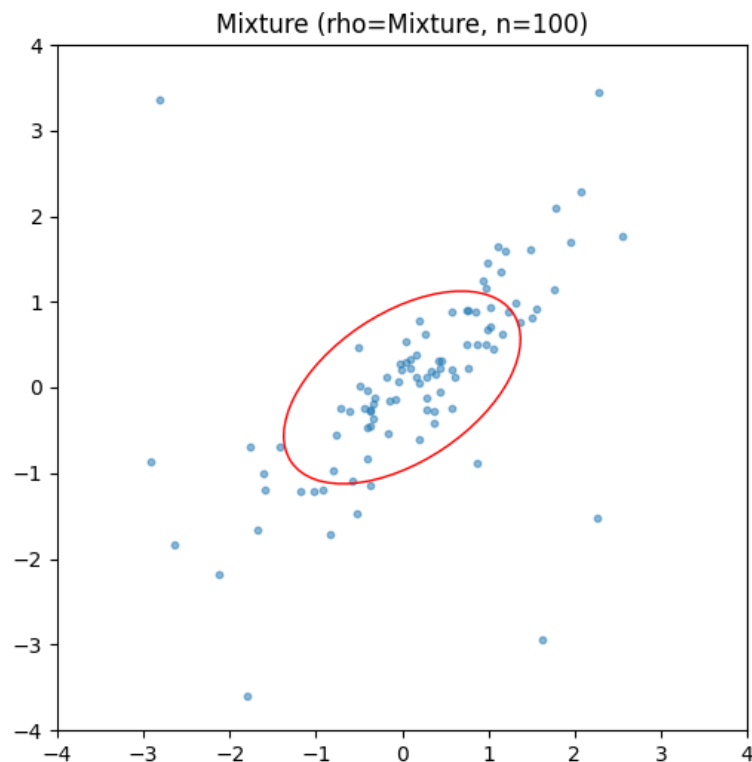


Рис. 12: Генерация точек для смеси с эллипсом равновероятности для  $n = 100$

## 4 Результаты вычислений для коэффициентов корреляции

В результате вычислений для каждого коэффициента корреляции были получены следующие статистики (среднее значение, среднее квадрата и дисперсия) для выборок размером 20, 60 и 100.

## 5 Выводы

- Коэффициент корреляции Пирсона, Спирмена и квадрантный коэффициент дают схожие результаты для нормального распределения с высоким коэффициентом корреляции.
- Для смеси нормальных распределений результаты отличаются из-за смешанных компонент с различной корреляцией.
- Графики и эллипсы равновероятности демонстрируют изменение зависимости между переменными в зависимости от значения коэффициента корреляции.

Таблица 1: Средние значения и дисперсии коэффициентов корреляции для различных выборок

Размер выборки	Коэффициент Пирсона	Коэффициент Спирмена	Квадрантный коэффициент
('Normal', 0, 20)	0.12343	0.05969	-0.09170
('Normal', 0, 60)	0.07463	0.11577	0.19617
('Normal', 0, 100)	0.14622	0.13461	0.08948
('Normal', 0.5, 20)	0.56445	0.57348	0.30410
('Normal', 0.5, 60)	0.40189	0.44387	0.29707
('Normal', 0.5, 100)	0.44021	0.40305	0.21928
('Normal', 0.9, 20)	0.95226	0.92174	0.70520
('Normal', 0.9, 60)	0.93982	0.92567	0.79920
('Normal', 0.9, 100)	0.90720	0.89664	0.84114
('Mixture', ", 20)	0.89881	0.83134	0.69410
('Mixture', ", 60)	0.50792	0.70468	0.53360
('Mixture', ", 100)	0.51057	0.71010	0.65790