

УДК 524.42.2

## МОДЕЛИРОВАНИЕ ПРОСТРАНСТВЕННОГО РАСПРЕДЕЛЕНИЯ ГАЛАКТИК И КЛАСТЕРИЗАЦИЯ С ИСПОЛЬЗОВАНИЕМ ТОЧЕЧНЫХ ПРОЦЕССОВ

Скробот Д.А., студент

Белорусский государственный университет информатики и радиоэлектроники<sup>1</sup>  
г. Минск, Республика Беларусь

Анисимов В.Я. – канд. физ.-мат. наук, доцент

**Аннотация.** Рассматривается моделирование пространственного распределения галактик с использованием пуассоновского и кластерного процесса Неймана–Скотта. Для анализа применяются периодические граничные условия, K-функция Рипли, расстояния до ближайшего соседа и кластеризация DBSCAN.

**Ключевые слова.** Галактики, пространственное распределение, точечные процессы, пуассоновский процесс, процесс Неймана–Скотта, кластеризация, K-функция Рипли, статистический анализ, моделирование, кластерные структуры, 3D-визуализация, Space Engine, крупномасштабная структура Вселенной.

Пространственное распределение галактик отражает эволюцию крупномасштабной структуры Вселенной и требует статистических моделей, способных описывать как равномерные, так и кластеризованные конфигурации точек [1]. В работе рассматриваются две базовые модели: пуассоновский точечный процесс (полностью случайное размещение) и кластерный процесс Неймана–Скотта, в котором сначала формируются центры кластеров, а затем вокруг них генерируются дочерние точки с гауссовским разбросом.

Для минимизации краевых эффектов наблюдаемый объем задается в виде куба с периодическими граничными условиями. Такая модель эквивалентна трехмерному тору и позволяет корректно оценивать межточечные расстояния в окрестности границ. В реализации использованы параметры: размер ячейки  $L = 100$ , число точек пуассоновского процесса  $N = 10000$ , число родительских точек  $N_0 = 100$ , среднее число потомков  $m = 100$ , дисперсия кластера  $\sigma = 3$ . Процедура эксперимента включает генерацию реализаций, применение периодических граничных условий, оценку K-функции на сетке радиусов и расчет распределения расстояний до ближайшего соседа. Для кластеризации используется DBSCAN, а для визуального анализа формируются 3D-сцены и каталоги Space Engine. Реализация на Python обеспечивает воспроизводимость за счет фиксированного генератора случайных чисел и возможность параметрического анализа при изменении  $\sigma$ ,  $N_0$  и  $m$ . При вычислении расстояний используется периодическая метрика:

$$d(x_i, x_j) = \sqrt{\sum_{k=0}^2 \min(|\Delta_k|, L - |\Delta_k|)^2}, \quad (1)$$

где  $\Delta_k$  – разность координат по  $k$ -й оси.

Схема периодических граничных условий показана на рисунке 1.

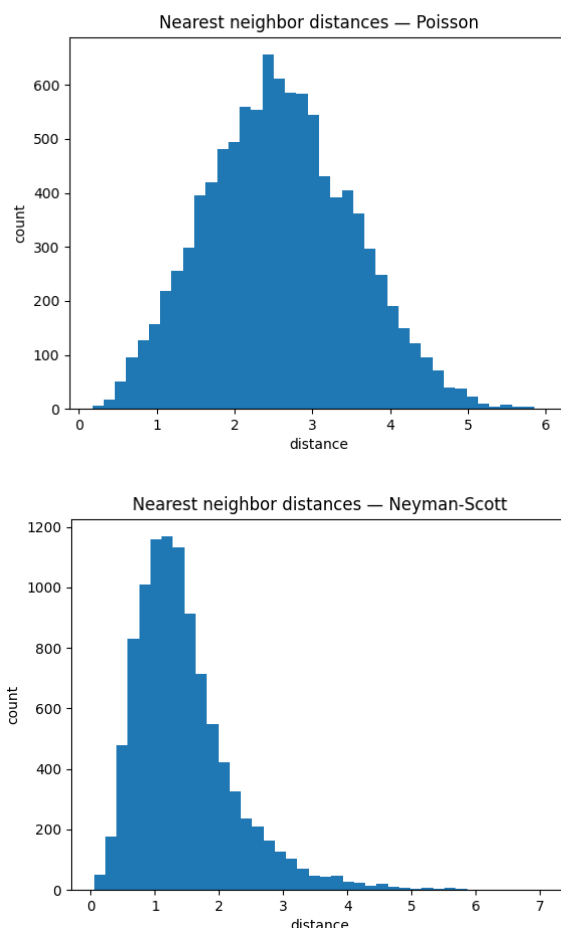


Рисунок 1 – Схематическое представление периодического объема.

Кластеризация оценивается с помощью К-функции Рипли [2]:

$$K(r) = \frac{1}{\lambda} E[N_r], \quad (2)$$

где  $\lambda$  – интенсивность процесса;  $N_r$  – число точек на расстоянии не более  $r$  от произвольной точки. Для полностью случайного распределения в трехмерном случае  $K_p(r) = \frac{4}{3}\pi r^3$ . Дополнительно анализируется распределение расстояний до ближайшего соседа, чувствительное к локальному уплотнению. Для выделения скоплений применяется DBSCAN с параметрами  $\varepsilon = 2.5$  и  $min\_samples = 10$ , что соотносится с масштабом кластеров  $\sigma$  [3]. Построение статистик выполняется на полном наборе попарных расстояний с учетом периодичности, что обеспечивает корректное сравнение с моделью полностью случайного распределения. Величины  $K(r)$  вычисляются для последовательности радиусов, что позволяет отслеживать переход от локальных структур к крупномасштабной однородности. Параметры DBSCAN подбираются так, чтобы отделять плотные ядра кластеров и сохранять разреженный фон, который интерпретируется как шум или межкластерное пространство.

Полученные результаты показывают, что пуассоновский процесс дает однородное заполнение объема, а эмпирическая  $K(r)$  согласуется с теоретической кривой  $K_p(r)$ . Для процесса Неймана–Скотта наблюдается существенное превышение  $K(r)$  над  $K_p(r)$  при малых  $r$ , что указывает на наличие кластеров. Гистограммы расстояний до ближайшего соседа смещаются в область меньших расстояний, подтверждая локальное уплотнение. Алгоритм DBSCAN выделяет компактные группы и шумовые точки, сохраняя естественную неоднородность распределения. Примеры реализаций пуассоновского и кластерного процессов приведены на рисунке 2, а сравнение К-функций – на рисунке 3.

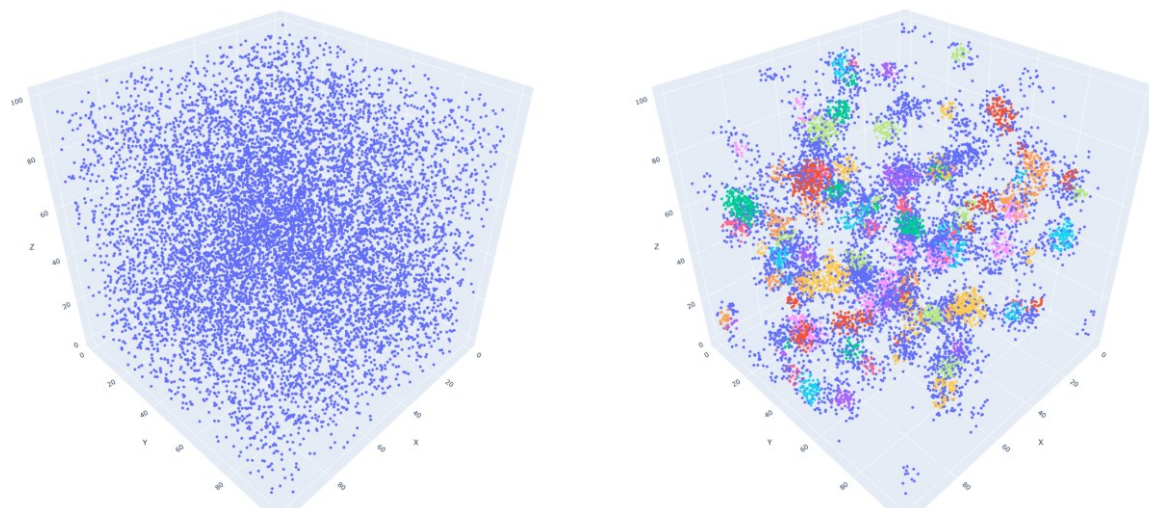


Рисунок 2 – Примеры 3D-реализаций точечных процессов.

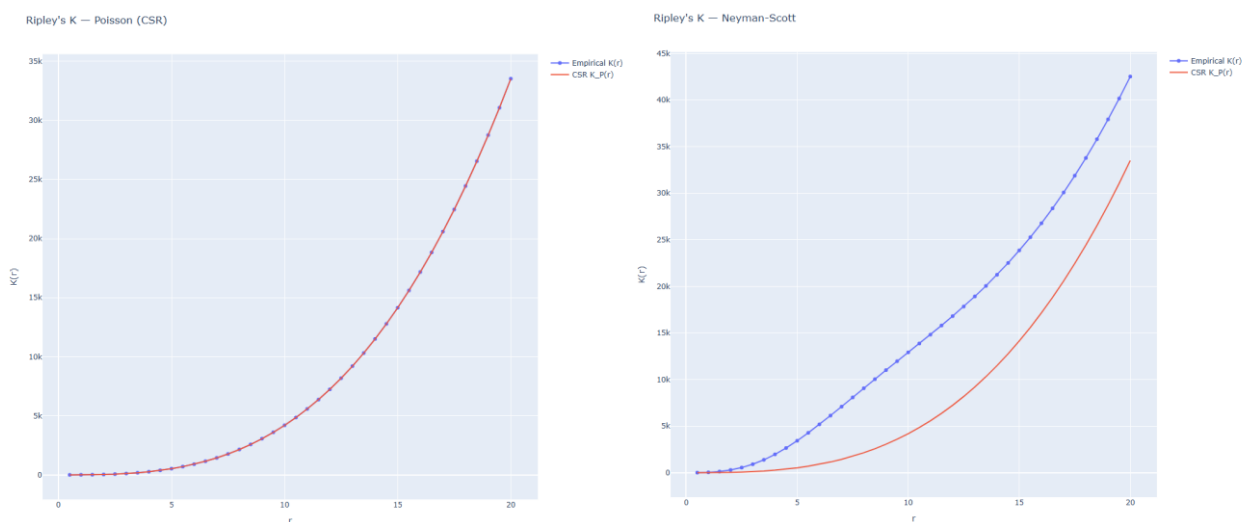


Рисунок 3 – K-функции для двух типов распределений.

Для визуальной проверки реалистичности структуры сформирован каталог для Space Engine, содержащий координаты (RA, Dec, Dist), типы галактик, абсолютные величины и ориентацию. В каталог также включены радиусы и кватернионы ориентации, что позволяет корректно отображать объекты в сцене. Визуализация в 3D подтверждает наличие скоплений и разреженных областей, что качественно согласуется с ожидаемой космической структурой. Итоговые визуализации из Space Engine приведены на рисунке 4.

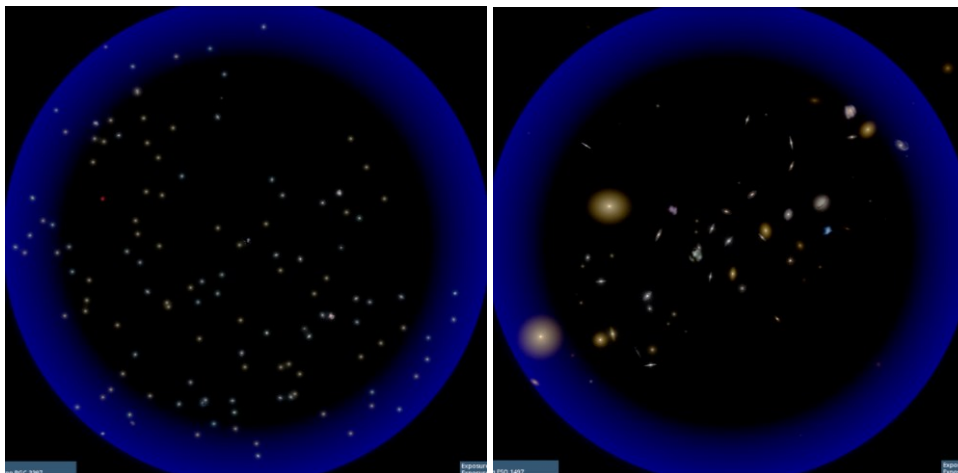


Рисунок 4 – Результат визуализации пуассоновского и кластерного распределений в Space Engine.

Практическая полезность работы заключается в том, что синтетические каталоги позволяют тестировать алгоритмы кластеризации, оценивать чувствительность статистик к параметрам модели, подбирать масштабы кластеров и проверять корректность обработки наблюдательных данных. Подобные модели применимы при подготовке вычислительных экспериментов, валидации программных средств визуализации и обучении методам пространственного анализа. Кроме того, полученный инструментарий можно использовать как основу для калибровки по реальным каталогам и моделирования эффектов наблюдательных ошибок. Практический результат состоит в наличии воспроизводимого конвейера, который объединяет статистический анализ, кластеризацию и визуальную проверку, что удобно для сравнительных исследований и настройки параметров алгоритмов.

Ограничения текущего подхода связаны с упрощенной физической моделью: не учитываются эффекты эволюции во времени, выборка по яркости, систематические искажения расстояний и влияние наблюдательных ошибок. В дальнейшем возможно расширение модели за счет анизотропных кластеров, вариативной интенсивности и сопоставления с реальными обзорными данными.

Таким образом, применение точечных процессов и K-функции Рипли позволяет количественно отличать случайные и кластерные распределения. Процесс Неймана–Скотта адекватно воспроизводит неоднородность галактических скоплений и может использоваться как базовая модель для дальнейшей калибровки по наблюдательным данным. Дополнительные статистические оценки локальных расстояний повышают интерпретируемость результатов и позволяют сравнивать модели между собой.

**Список использованных источников:**

1. Illian, J. *Statistical Analysis and Modelling of Spatial Point Patterns* / J. Illian, A. Penttinen, H. Stoyan, D. Stoyan. – Chichester : Wiley, 2008. – 534 p.
2. Ripley, B. D. *Modelling spatial patterns* // *Journal of the Royal Statistical Society. Series B*, 1977. – Vol. 39, No. 2. – P. 172–212.
3. Ester, M. *A density-based algorithm for discovering clusters in large spatial databases with noise* / M. Ester, H.-P. Kriegel, J. Sander, X. Xu // *Proceedings of KDD-96*. – 1996. – P. 226–231.

UDC 524.42.2

## MODELING THE SPATIAL DISTRIBUTION OF GALAXIES AND CLUSTERING USING POINT PROCESSES

*Skrobot D.A., student*

*Belarusian State University of Informatics and Radioelectronics  
Minsk, Republic of Belarus*

*Anisimov V.Y. – PhD in Physics and Mathematics*

**Annotation.** The modeling of the spatial distribution of galaxies is examined using Poisson and Neyman–Scott cluster processes. Periodic boundary conditions, Ripley's K-function, nearest-neighbor distances, and DBSCAN clustering are employed for the analysis.

**Keywords.** Galaxies, spatial distribution, point processes, Poisson process, Neyman–Scott process, clustering, Ripley's K-function, statistical analysis, modeling, cluster structures, 3D visualization, Space Engine, large-scale structure of the Universe.