

Felipe Ventura Vargas

**Estudando o Universo em Larga Escala
Função de Correlação de dois Pontos:
Uma Análise de suas limitações e
potencialidades**

Brasil

2016

Felipe Ventura Vargas

**Estudando o Universo em Larga Escala
Função de Correlação de dois Pontos:
Uma Análise de suas limitações e potencialidades**

O presente trabalho de conclusão de curso, requisito para a graduação no curso de Física Bacharelado, tem por objetivo analisar uma ferramenta muito utilizada na Cosmologia, a Função de Correlação de Dois Pontos. Esta é uma primeira etapa para um estudo sobre a distribuição de matéria no universo.

Universidade Federal do Rio Grande – FURG

Faculdade de Física - Bacharelado

Programa de Graduação

Orientador: Evelise Gausmann

Brasil

2016

Felipe Ventura Vargas

Estudando o Universo em Larga Escala

Função de Correlação de dois Pontos:

Uma Análise de suas limitações e potencialidades/ Felipe Ventura Vargas. – Brasil,
2016-

62 p. : il. (algumas color.) ; 30 cm.

Orientador: Evelise Gausmann

Trabalho de Conclusão de Curso – Universidade Federal do Rio Grande – FURG

Faculdade de Física - Bacharelado

Programa de Graduação, 2016.

1. Função de Correlação. 2. Distribuição de Matéria. 3. Distribuição Estatística.
I. Evelise Gausmann. II. Universidade Federal do Rio Grande. III. Faculdade de
Física Bacharelado. IV. Função de Correlação de dois Pontos: Uma Análise de
suas limitações e potencialidades

Felipe Ventura Vargas

**Estudando o Universo em Larga Escala
Função de Correlação de dois Pontos:
Uma Análise de suas limitações e potencialidades**

O presente trabalho de conclusão de curso, requisito para a graduação no curso de Física Bacharelado, tem por objetivo analisar uma ferramenta muito utilizada na Cosmologia, a Função de Correlação de Dois Pontos. Esta é uma primeira etapa para um estudo sobre a distribuição de matéria no universo.

Trabalho aprovado. Brasil, dezembro de 2016:

Evelise Gausmann
Orientadora

Professor
Fabrício Ferrari

Professora
Dinalva Sales

Brasil
2016

Dedico este trabalho aos curiosos

Agradecimentos

"Se cheguei até aqui foi porque me apoiei no ombro dos gigantes"

Começo meus agradecimentos com uma frase famosa de Sir Isaac Newton. Não sei na verdade se tal frase foi dita por Newton e resolvi não pesquisar, pois não importa. Ouvi essa frase no primeiro ano de faculdade em uma palestra quando ainda estudava Engenharia Mecânica. Talvez na época tenha levado levemente o significado da frase, mas hoje, com certeza, seu significado pesa para mim. Não porque tenha caminhado no ombro de gigantes da ciência ou por ter ido longe como Newton. Mas porque, se cheguei a algum lugar, muita gente contribuiu para isso. Incontáveis pessoas influenciaram a minha vida de maneira que talvez não imaginem e nem mesmo eu saiba. Amigos, familiares, professores, todos tiveram um papel crucial na minha formação com profissional e como ser humano.

Inúmeras pessoas influenciaram a minha vida e nem todas elas receberão o merecido agradecimento nesse trabalho e algumas delas não estão mais entre nós para recebê-lo. Gostaria primeiramente de agradecer aos meus professores. Agradeço ao professor Fabrício Ferrari, por ter me ajudado na tomada de decisão de troca de curso ao mostrar seu trabalho e falar da profissão do Físico. Agradeço também aos choques de realidade que, embora nada sutis, me ajudaram a colocar o pé no chão.

Agradeço à minha orientadora, Evelise Gausmann, que me resgatou naquele momento em que sentimos que deveríamos jogar tudo pro alto e vender arte na praia. A paciência, cobrança e orientação foram essenciais para que tudo tenha se concretizado. Sem ela eu tenho certeza de que não estaria terminando este trabalho.

Outros professores contribuíram de maneira menos direta, mas também importante. Agradeço ao professor Magno Collares pelas ajudas administrativas para o meu intercâmbio, pelas ajudas na volta, pelos conselhos profissionais e pessoais. Agradeço à professora Rosângela Menegotto Costa, minha primeira orientadora dentro do curso de Física, que me ajudou na minha adaptação à mudança. Agradeço a todos os outros professores do curso de Física, que me passaram conhecimento, exemplo e mostraram os caminhos para ser um profissional das mais diversas formas. Espero um dia ser como vocês.

Também minha família e meus amigos tiveram um papel crucial para que eu estivesse hoje aqui. Agradeço aos meus pais, Gilberto Santos Vargas e Maria Albertina Ventura Vargas, pelo apoio moral e material dos meus estudos. Agradeço à minha irmã, Stefani Ventura Vargas, pela contínua preocupação e apoio incondicional à minha decisão de seguir essa caminhada. Agradeço aos meus amigos Leonardo Ferreira, Helena Vasena, André D. Oliveira pelo apoio moral e pela tentativa de ajudar a manter minha sanidade

nos momentos conturbados (não sei se funcionou). Agradeço à minha namorada, Gabriela Siston, pela compreensão na ausência, pelo apoio nos momentos necessários e por estar ao meu lado quando preciso.

Também gostaria de agradecer àqueles que já não mais podem receber os agradecimentos. Gostaria de agradecer às pessoas que faleceram e não pude me despedir por estar nessa caminhada. Agradeço ao vô Américo e a Dona Alfa, que faleceram durante meu intercâmbio na Irlanda. Agradeço também à amiga Shani Sevilla, pelo apoio quando precisei, pela minha gata Mileva. Uma pena não poderem estar neste momento.

Aos que não foram citados mas fizeram parte da minha vida, agradeço de coração, pois tenham certeza, todos vocês fazem parte de quem sou hoje.

*“Não são os sentidos que percebem o infinito;
não é pelos sentidos que chegamos a esta conclusão,
porque o infinito não pode ser objeto dos sentidos. [...] Ora, se conhecemos por experiência que eles nos enganam
com respeito à superfície do globo no qual nos encontramos,
muito mais devemos suspeitar deles quando querem
referir-se ao côncavo céu estrelado.”*

*Giordano Bruno, em
Sobre o Infinito, o Universo e os Mundos*

Resumo

A Função de Correlação de Dois Pontos é um dos principais métodos estatísticos utilizados dentro da cosmologia, popularizada em sua forma espacial por [Groth e Peebles \(1977\)](#). É uma ferramenta muito importante, ao comparar distribuições amostrais com distribuições geradas por um Processo Pontual de Poisson Homogêneo, identifica escalas nas quais há homogeneidade ou aglomeração. Por ter tal importância, a Função de Correlação de dois Pontos deve ter suas limitações e potencialidades exploradas. Estudos e testes com diferentes distribuições sintéticas foram realizados. Os resultados mostraram que a Função de Correlação de Dois Pontos consegue identificar escalas características nas quais distribuições podem ser considerada homogêneas, concentradas ou apresentam vazios. Um limite para a resolução das oscilações que representam escalas características foi encontrado, caso o afastamento entre picos das oscilações (e, portanto, distância entre distribuições) for inferior a quatro vezes o desvio padrão das distribuições, as oscilações são sobrepostas - informação que não se perde no histograma de distâncias que forma a função. Também, um resultado encontrado na literatura foi confirmado. Foi mostrado que diferentes distribuições podem gerar Funções de Correlação de dois Pontos similares.

Palavras-chaves: Função de Correlação de Dois Pontos. Distribuições de distâncias. Distribuições de pontos. Escalas Características.

Abstract

The Two-Point Correlation Function is one of the main statistical methods used in cosmology, popularized in its spatial form by [Groth e Peebles \(1977\)](#). It is a very important tool when comparing sample distributions with distributions generated by a Poisson Homogeneous Point Process, identifying scales in which there is homogeneity or agglomeration. Because of such importance, the Two-Point Correlation Function must have its limitations and potentialities explored. Studies and tests with different synthetic distributions were performed. The results showed that the Two-Point Correlation Function can identify characteristic scales in which distributions can be considered homogeneous, concentrated or voids. A limit for the resolution of the oscillations representing characteristic scales was found. If the distance between peaks of the oscillations (and therefore distance between distributions) is less than four times the standard deviation of the distributions, the oscillations are superimposed - information that is not lost in the distance histogram that forms the function. Also, a result found in the literature was confirmed: it has been shown that different distributions can generate similar Two-Point Correlation Functions.

Key-words: Two-Point Correlation Function. Distributions of distances. Distributions of points. Scales Features.

Lista de ilustrações

Figura 1 – “ <i>The Origin of the Milky Way</i> ”, por Jacopo Tintoretto	3
Figura 2 – Distribuição de Aglomerados Globulares projetadas no Plano Galático.	5
Figura 3 – Comparação de Distâncias no Sistema Galático	6
Figura 4 – Nebulosas classificadas por Hubble	7
Figura 5 – Lei de Hubble	8
Figura 6 – Dados de Sloan Digital Sky Survey (SDSS)	9
Figura 7 – Densidade de Probabilidade Gaussiana	16
Figura 8 – Exemplo de Dados Distribuição Gaussiana 3D	17
Figura 9 – Projeção Bidimensional de Dados Sintéticos	17
Figura 10 – Exemplo de Distribuição Beta	19
Figura 11 – Exemplo Tridimensional de Nuvem Beta	19
Figura 12 – Exemplo de Distribuição	20
Figura 13 – Exemplo de Distribuição projetada no plano: 10 mil pontos	21
Figura 14 – Exemplo de Distribuição Frequências de Distâncias	22
Figura 15 – Número de pontos: Probabilidade de Poisson	23
Figura 16 – Amostra gerada por um Processo Pontual de Poisson Homogêneo	24
Figura 17 – Significado dos picos: 8 Nuvens	30
Figura 18 – Significado dos picos: 8 Nuvens projetadas	30
Figura 19 – Histograma distribuição de estudo dos picos	31
Figura 20 – Correlação distribuição de estudo dos picos	31
Figura 21 – Visualização das distribuições Iguais	32
Figura 22 – Correlações médias de Nuvens Iguais: Ampliada	33
Figura 23 – Correlação de Nuvens Iguai	33
Figura 24 – Nuvens centradas em $(2,2,2)$ e $(5,5,5)$	34
Figura 25 – Nuvens centradas em $(3,3,3)$ e $(6,6,6)$	35
Figura 26 – Função de Correlação Suavizada: Nuvens 2 e 5, 3 e 6	35
Figura 27 – Nuvens: Gaussiana e Caixa homogênea	36
Figura 28 – Projeção 2D de Nuvens: Gaussiana e Caixa homogênea	37
Figura 29 – Nuvens: Gaussiana e Caixa Beta	37
Figura 30 – Projeção 2D de Nuvens: Gaussiana e Caixa Beta	38
Figura 31 – Nuvens Gaussianas	38
Figura 32 – Projeção 2D de Nuvens Gaussianas	39
Figura 33 – Comparação Correlação Distribuições Distintas	40
Figura 34 – Comparação Correlação Distribuições: Segundo Pico	41
Figura 35 – Comparação Correlação Distribuições: Com desvios no terceiro pico	41
Figura 36 – Distribuições centradas em $x = 30$ u.m. e $x = 70$ u.m.	43

Figura 37 – Distribuições centradas em $x = 35$ u.m. e $x = 65$ u.m.	43
Figura 38 – Distribuições centradas em $x = 40$ u.m. e $x = 60$ u.m.	44
Figura 39 – Distribuições centradas em $x = 45$ u.m. e $x = 55$ u.m.	44
Figura 40 – Histograma de nuvens em $x = 30$ u.m. e $x = 70$ u.m.	45
Figura 41 – Histograma de nuvens em $x = 35$ u.m. e $x = 65$ u.m.	45
Figura 42 – Histograma de nuvens em $x = 40$ u.m. e $x = 60$ u.m.	46
Figura 43 – Histograma de nuvens em $x = 45$ u.m. e $x = 55$ u.m.	46
Figura 44 – Correlação de nuvens em $x = 30$ u.m. e $x = 70$ u.m.	47
Figura 45 – Correlação de nuvens em $x = 35$ u.m. e $x = 65$ u.m.	47
Figura 46 – Correlação de nuvens em $x = 40$ u.m. e $x = 60$ u.m.	48
Figura 47 – Correlação de nuvens em $x = 45$ u.m. e $x = 55$ u.m.	48
Figura 49 – Histograma de 8 Nuvens	49
Figura 50 – FC2P de 8 Nuvens	50
Figura 51 – Histograma 27 Nuvens	51
Figura 52 – FC2P de 27 Nuvens	52
Figura 53 – Histograma de 64 Nuvens	53
Figura 54 – FC2P de 64 Nuvens.	54
Figura 55 – Histograma renormalizado e comparado com histograma homogêneo: 8 nuvens	55
Figura 56 – Histograma renormalizado e comparado com histograma homogêneo: 27 nuvens	56
Figura 57 – Histograma renormalizado e comparado com histograma homogêneo: 64 nuvens	56
Figura 58 – Histogramas renormalizados	56
Figura 59 – Composição de FC2P de 8 nuvens, 27 nuvens e 64 nuvens	57

Lista de tabelas

Tabela 1 – Distâncias entre pares de pontos	21
Tabela 2 – Picos de distribuição 8 Nuvens	31
Tabela 3 – Picos de distribuições diferentes	40

Sumário

I	INTRODUÇÃO	1
1	O ESTUDO DO UNIVERSO	3
	Introdução	3
1.1	Do mito à ciência	3
1.2	Contextualização	4
1.3	Justificativa	10
1.4	Objetivo	11
1.4.1	Geral	11
1.4.2	Específico	11
II	METODOLOGIA	13
2	METODOLOGIA	15
	Metodologia	15
2.1	Dados Sintéticos	15
2.1.1	Distribuição Gaussiana	16
2.1.2	Distribuição Beta	18
2.2	Função de Correlação de Dois Pontos	19
2.2.1	Distribuição de Frequência de Distâncias	20
2.2.2	Processo Pontual de Poisson Homogêneo	23
2.2.3	Implementação	24
III	RESULTADOS	27
3	APRESENTAÇÃO DOS DADOS	29
3.1	Primeiros Estudos	29
3.1.1	Estudo dos picos	29
3.1.1.1	Nuvens Gaussianas	29
3.1.1.2	Histograma e Correlação	30
3.1.1.3	Variações no primeiro pico	32
3.1.2	Teste de Centralidade	34
3.2	Testes com Distribuições	36
3.2.1	Teste Diferentes Distribuições	36

3.2.1.1	Funções de Correlação de Dois Pontos	40
3.2.2	Testes Gaussianas: Sobreposição de Picos	43
3.2.2.1	Distribuições Gaussianas: Distância variável	43
3.2.2.2	Histogramas	45
3.2.2.3	Funções de Correlação	47
3.2.3	Teste de Distribuições Gaussianas	49
3.2.3.1	8 nuvens, Histograma e Correlação	49
3.2.3.2	27 nuvens, Histograma e Correlação	51
3.2.3.3	64 nuvens, Histograma e Correlação	53
3.2.3.4	Comparação das nuvens	55
4	CONCLUSÃO	59
	Conclusão	59
	REFERÊNCIAS	61

Parte I

Introdução

1 O Estudo do Universo

“Já que o homem é instrumento do seu próprio conhecimento, é necessário estudá-lo como um instrumento antes de possamos avaliar o valor do que os seus sentidos nos dizem com respeito ao mundo”
Bertrand Russel, em *Delineamentos da Filosofia*

1.1 Do mito à ciência

A busca por um entendimento da existência do Universo é algo que está nos estudos da ciência desde os primórdios da sociedade. Povos antigos utilizavam os céus para prever o futuro, associando a posição dos astros às estações do ano, épocas da colheita e muitos outros fatores que auxiliavam na manutenção física e espiritual das comunidades. O pintor renascentista, Jacopo Tintoretto, representa em sua obra *“The Origin of the Milky Way”* (Figura 1), exposta na Galeria Nacional de Londres, a origem da Via Láctea segundo os gregos. Segundo a mitologia grega a deusa Hera acidentalmente esparrama seu leite nos céus ao ser mordida por Hércules durante a amamentação, criando a Vialáctea.



Figura 1 – *“The Origin of the Milky Way”*, por Jacopo Tintoretto

A busca evoluiu nos últimos séculos, do mito à ciência. Filósofos e pensadores influenciaram à crítica do misticismo que permeou a humanidade por quase toda sua existência. Giordano Bruno, um dos primeiros pensadores a criticar a Igreja questiona os conceitos do espaço em si e do tamanho do Unvierso cristão. Na tradução da obra *“Sobre*

o *Infinito, o Universo e os Mundos*” (Bruno; Galilei; Lobo A.,), Giordano Bruno afirma no diálogo de um de seus personagens:

“O mundo será alguma coisa que não se encontra. Se você afirma. [...] que fora do mundo há um ente intelectual e divino, de sorte que Deus venha a ser lugar de todas as coisas, você mesmo se encontrará em muita dificuldade para fazer entender como uma coisa incorpórea, inteligível e sem dimensões possa ser o lugar duma coisa dimensionada. Porque, se você afirma que contém como uma forma e da mesma maneira como a alma contém o corpo, você não responde à questão do estar fora nem à pergunta daquilo que se encontra além e fora do Universo. E se você quer explicar dizendo que onde está e onde não existe coisa alguma tampouco existe lugar, nem o além nem o extra, não me satisfará; porque são palavras e desculpas que não se podem realmente pensar. Com efeito, é absolutamente impossível que com qualquer juízo ou fantasia [...] você possa levar-me a afirmar, com real intenção, que exista tal superfície, tal limite, tal extremidade além da qual não exista nem corpo nem vácuo; mesmo que ali estivesse Deus, porque a divindade não tem por função encher o vácuo. E por conseqüência, não tem, de modo algum, a função de terminar o corpo; porque tudo o que se diz terminar ou é forma exterior ou é corpo continente. E de qualquer forma que você o quisesse afirmar, você estaria prejudicando a dignidade da natureza divina e universal. Além daquele convexo, ela não estaria num lugar, nem em parte alguma, e conseqüentemente não existiria.”

Os avanços tecnológicos que permitiram uma instrumentação capaz de visualizar além dos limites do nosso conhecimento expandiram o tamanho do nosso universo corroborando os críticos dos dogmas religiosos. Nossa visão de uma Terra plana, centro de tudo, foi modificada, chegamos a um ponto de compreensão de um universo infinito e abstrato que nos libertou da doutrina da igreja. A ciência deve então tentar descrever o universo com base na realidade e não na crença.

1.2 Contextualização

As ideias da busca pela comprovação empírica de teorias afloraram do pensamento aristotélico e atravessaram os tempos para gerar a práxis exigida pelos rígidos critérios científicos. Em um dos primeiros estudos do entendimento das estruturas existentes no universo “*On the construction of the Heavens*”, Herschel (1785) comenta: “*But first let me mention that, if we would hope to make any progress in an investigation of this delicate nature, we ought to avoid two opposite extremes, of which I can hardly say which is the most dangerous. If we indulge a fanciful imagination and build worlds of our own, we must not wonder at our going wide from the path of truth and nature; but there will vanish like the Cartesian vortices, that soon gave way when better theories were offered. On the other hand, if we add observation to observation, without attempting to draw not only*

certain conclusions, but also conjectural views from them, we offend against the very end for which only observations ought to be made. I will endeavour to keep a proper medium; but if I should deviate from that, I could with not fall into latter error."¹ Este trabalho foi o primeiro a identificar o que Herschel chamou de "nebulosas", dando origem ao estudo das galáxias e ao que posteriormente se tornou o Catálogo Messier, publicado em 1771.

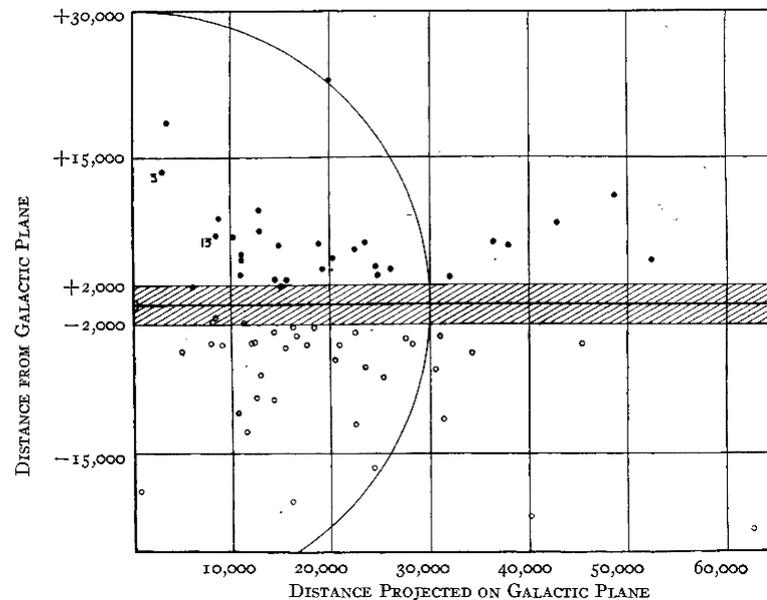


Figura 2 – Distribuição de Aglomerados Globulares projetadas no Plano Galáctico. A unidade de distância é Parsec. Retirado de [Shapley \(1918\)](#).

Muitos anos depois outras estruturas foram estudadas por [Shapley \(1918\)](#). O trabalho de Shapley analisou aglomerados globulares² (Figura 2) e, através do tamanho angular, brilho das estrelas e período de Cefeidas³, estimou o centro e determinou a verdadeira escala de tamanho de nossa Galáxia (Figura 3), montando a primeira estrutura de um sistema galáctico.

O estudo das velas padrão⁴ permitiu que Edwin Hubble fizesse suas grandes contribuições para a ciência. [Hubble \(1926\)](#) faz uma análise estatística de 400 nebulosas

¹ [TRADUÇÃO LIVRE] “Mas, antes, deixe-me mencionar que, se quisermos avançar em uma investigação dessa delicada natureza, devemos evitar dois extremos opostos, dos quais não posso dizer qual é o mais perigoso. Se nos entregarmos a uma imaginação fantasiosa e construirmos mundos nossos, não devemos nos admirar de que estejamos longe do caminho da verdade e da natureza. Mas desaparecerão como os vórtices cartesianos, que logo cederam quando melhores teorias foram oferecidas. Por outro lado, se acrescentarmos a observação à observação, sem tentar extrair não apenas certas conclusões, mas também visões conjecturais delas, ofendemos o próprio fim pelo qual apenas devem ser feitas observações. Eu me esforçarei para manter um meio apropriado. Mas caso deva desviar de um deles, eu evitaria a todo custo o último erro comentado.”

² Aglomerados Globulares são Aglomerados Estelares fechados. Uma discussão mais detalhada pode ser encontrada no capítulo 22.2, página 233 de [Kepler e Saraiva \(2011\)](#)

³ Cefeidas são estrelas cujas luminosidades variam com periódicas bem definidas. Ao saber a luminosidade de uma estrela é possível saber a distância em que se encontra. Uma discussão mais detalhada relacionando a magnitude e a distância de estrelas pode ser encontrado no capítulo 2.1, página 42 de [Souza \(2004\)](#)

⁴ Velas padrão são objetos astronômicos cuja a distância pode ser muito bem definida e podem ser

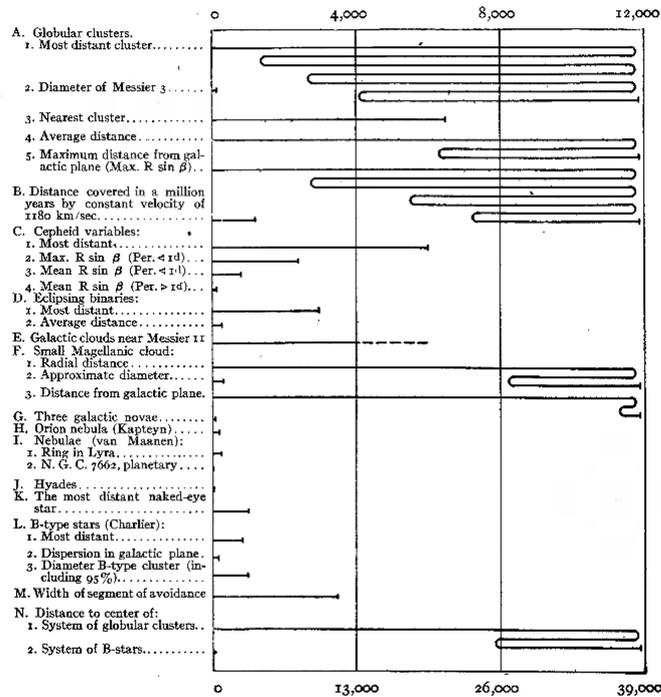


Figura 3 – Comparação de Distância no Sistema Galático. As distâncias na parte superior da figura estão em Parsecs enquanto as distâncias na parte inferior estão em anos-luz. Retirado de [Shapley \(1918\)](#).

extragaláticas e determina que 97 % dos objetos chamados de "*nebulosas*" encontradas por [Herschel \(1785\)](#) apresentam forma em torno de um núcleo variando de elípticas a espirais. Hubble então faz a classificação dessas galáxias que variam de "*primordiais*" à "*tardias*"⁵. São galáxias como conhecemos nos dias atuais.

Neste artigo Hubble também corrobora a ideia, inicialmente utilizada por [Shapley \(1918\)](#), de que a magnitude aparente dos objetos é uma medida de distância. Hubble também descobre que a massa luminosa das nebulosas é da ordem de $2,6 \times 10^8$ massas solares⁶. Em [Hubble \(1929\)](#) é desenvolvida a famosa lei que relaciona distância e velocidade das galáxias⁷. Em [Hubble e Humason \(1931\)](#) são discutidos métodos para determinar

utilizados para determinar a distância de objetos próximos. As estrelas Cefeidas utilizadas por Shapley são um exemplo de velas padrão. O período de pulsação e a magnitude bem determinada das Cefeidas faz com que tenham propriedades físicas bem previsíveis, fazendo com que se possa facilmente calcular a distância em que estão. Outro exemplo de vela padrão são as Supernova do tipo IA, cuja a luminosidades permitiu que a taxa de expansão do Universo fosse mais precisamente medida para distâncias muito maiores. Uma discussão mais aprofundada sobre Velas Padrão pode ser encontrado no capítulo 7 de [Ryden \(2003\)](#)

⁵ As palavras utilizadas por Hubble para classificar as galáxias podem dar uma ideia enganosa de que as galáxias foram classificadas em ordem temporal. Em uma nota de rodapé em seu artigo, Hubble explica que os termos utilizadas servem para dar uma ideia de evolução no grau de complexidade das estruturas.

⁶ Uma massa solar é $19,891 \times 10^{30}$ Kg.

⁷ A lei de Hubble foi um importante marco do estudo da astrofísica. Devido a falta de recursos instrumentais na época, o valor que relacionava distância e velocidade encontrado por Hubble era linear e muito diferente de valores atuais. Na época Hubble teve acesso a dados que correspondiam a uma fração muito menor do universo que conseguimos observar hoje em dia, mas de qualquer maneira

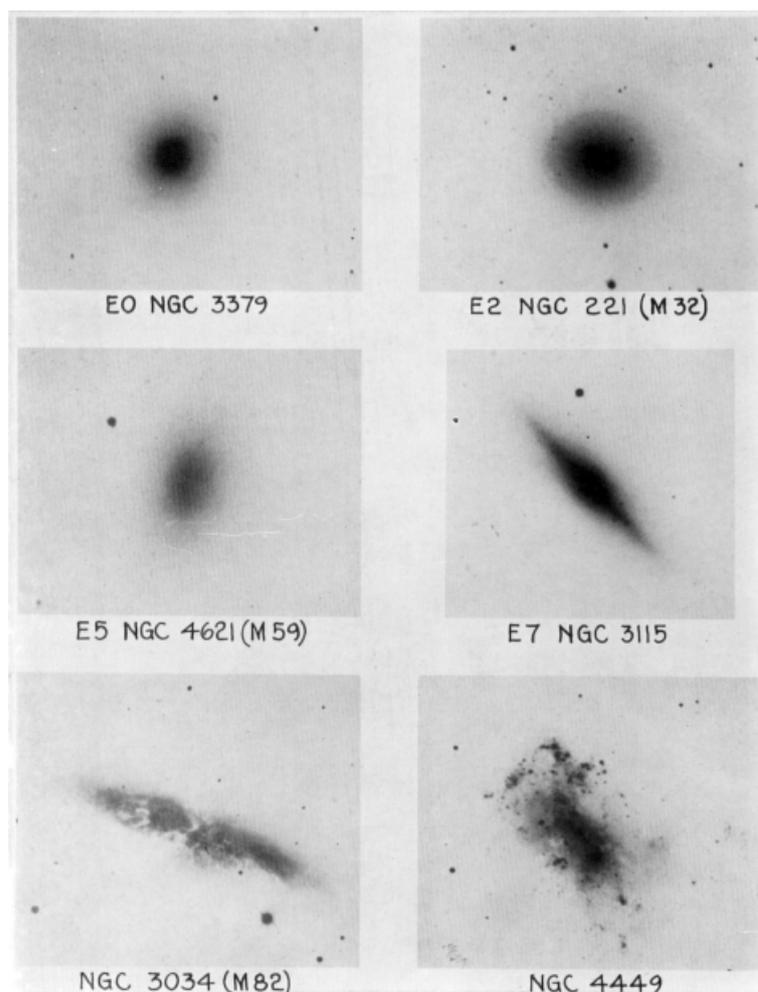


Figura 4 – Nebulosas Classificadas por Hubble (1926). Hubble classifica essas galáxias entre “*primordiais*” e “*tardias*” para organizá-las em ordem de complexidade. Retirado de Hubble (1926)

distâncias de nebulosas extragaláticas, também com base nas ideias discutidas por Shapley (1918). Então em Hubble (1936) é estudado a relação entre o redshift⁸ e a distribuição de nebulosas.

Estrelas se organizam em galáxias. Seria possíveis que tais estruturas massivas se organizassem em algo ainda maior? Tal questão sobre a distribuição de matéria no Universo é importante, pois ela modifica a teoria e as previsões de como deve ser o Universo. de Vaucouleurs (1953), através do estudo da velocidade radial de galáxias, encontra evidências

o fato de ter estabelecido essa relação foi muito importante para a área.

⁸ Redshift é o desvio para o vermelho causado pela influência da velocidade relativa entre observador e fonte, ou devido a influência gravitacional que muda o comprimento de onda da luz. Este efeito é assemelhado ao Efeito Doppler sonoro que observamos ao ver uma ambulância passar. Quando a ambulância está se aproximando seu som parece mais agudo, já que a fonte está se aproximando dando a impressão de que o comprimento de onda é menor do que o emitido. Mas no momento em que a ambulância começa a se afastar, temos a impressão contrária, seu som parece mais grave, já que a fonte está se afastando e o comprimento de onda parece maior do que deveria. Maiores explicações podem ser encontradas no capítulo 21.6 de Kepler e Saraiva (2011)

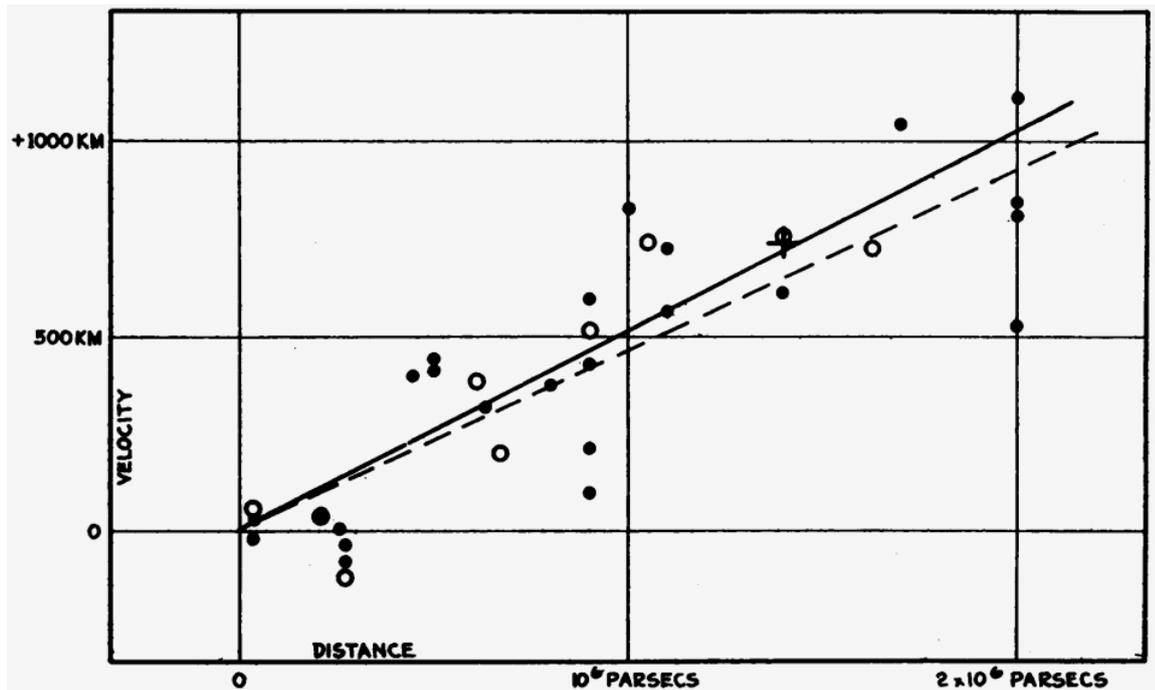


Figura 5 – Relação velocidade/distância entre nebulosas extragalácticas estabelecida por Hubble (1929). Adaptado de Hubble (1929).

de que existam estruturas que chamou de "*supergaláxias*"⁹, uma organização ainda maior que galáxias onde estas estavam reunidas. A cada descoberta sobre o que é e como eram as grandes estruturas do Universo ampliou-se ainda mais o nosso entendimento das escalas de tais estruturas, com isso novas questões surgiram. A distribuição de matéria no Universo começa a apontar uma tendência a se aglomerar, tornando-se menos homogênea do que se imaginava, contrariando o Princípio Cosmológico¹⁰

O estudo de estruturas em larga escala continuou e cada vez mais evidências apontavam uma falta de homogeneidade na distribuição de matéria do Universo. Shane e Wirtanen (1954) fez uma contagem de nebulosas extragalácticas e apontou uma tendência de que aglomerados ocorreriam em grupos. Afirmou que a distribuição dos aglomerados é uma extensão da distribuição das galáxias e isto aparentava ser uma tendência natural dos aglomerados.

Com evidências para a existência de aglomerados de galáxias, surge a questão se tais estruturas seriam uma extensão das galáxias, ou apenas galáxias reunidas. Caso sejam, não é possível dizer que o Universo é homogêneo, pois galáxias são

⁹ Posteriormente conhecidas como aglomerados, superaglomerados e o grupo local.

¹⁰ O princípio cosmológico influencia a ciência há séculos. Na obra "*Philosophiæ Naturalis Principia Mathematica*" (Newton; et, al., 1686), Newton acreditava em um universo perfeito com a Terra em movimento orbital ao redor do Sol, cercada por um espaço infinito e uniforme. Na modernidade, o Princípio Cosmológico prevê um universo homogêneo e isotrópico em escalas suficientemente grandes, garantindo que qualquer observador em qualquer ponto do Universo observe o mesmo universo. Uma explicação didática e mais elaborada sobre o Princípio Cosmológico pode ser encontrada no capítulo 2.2 de Ryden (2003).

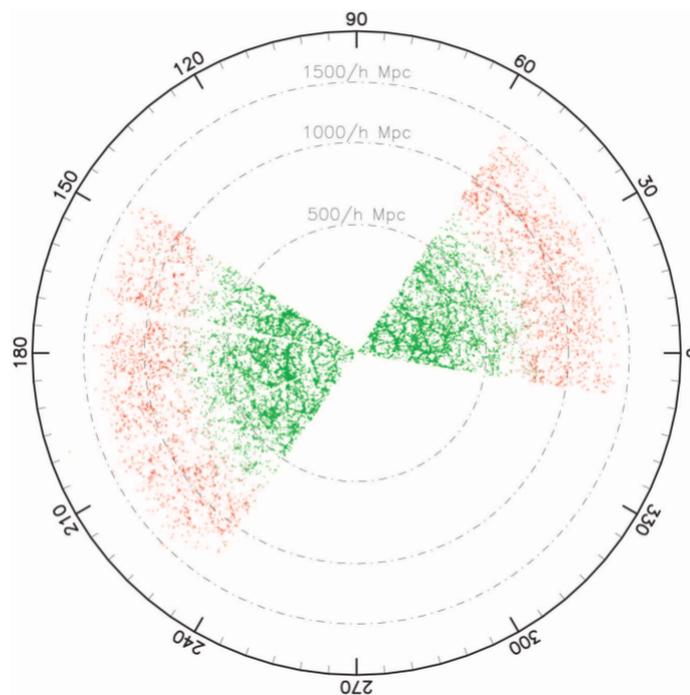


Figura 6 – Dados do Sloan Digital Sky Survey (SDSS). Amostra de distribuição de galáxias dos dois hemisférios Celestes. Até 2001 o levantamento havia detectado 14 milhões de objetos. Adaptado [Stoughton e et, al. \(2002\)](#).

nucleadas e isto requer que uma densidade maior de matéria esteja próxima ao núcleo. Se os aglomerados de aglomerados também forem nucleados e esse padrão se repetir para escalas muito maiores então teremos pontos de grande densidade de massa e regiões sem massa alguma em todas as escalas especiais. Hoje se sabe que o padrão de homogeneidade e isotropia está em escalas acima de 100 Mpc, mas na metade do século XX, muitas evidências colocavam em cheque tais propriedades. Tais super aglomerados também foram estudados em [Abell \(1958\)](#) onde um catálogo é preparado com 2712 aglomerados ricos de galáxias, 1682 aglomerados que respeitam critérios estatísticos de amostras homogêneas são selecionados levando a conclusões importantes: a densidade espacial do centro dos aglomerados não varia com a distância, inferindo que tais aglomerados não eram uma extensão da distribuição de galáxias, mas sim uma forma mais homogênea de disposição, não nucleada. [Abell \(1958\)](#) também afirma que existe um padrão significativo não aleatório de distribuição dos aglomerados, formando o que chamou de aglomerados de segunda ordem.

A partir de então todas as evidências começaram a apontar em favor do Princípio Cosmológico. Um dos maiores levantamentos recentemente feitos é o *Sloan Digital Sky Survey (SDSS)* ([Stoughton; et, al., 2002](#)). A Figura 6 a distribuição de 14 milhões de objetos nos dois hemisférios celestes, apontando um grande grau de homogeneidade na distribuição de matéria no universo.

1.3 Justificativa

A discussão sobre a homogeneidade do universo se estendeu até que surveys em escalas muito grandes foram realizados. Mas foi necessário desenvolver ferramentas para quantificar tal homogeneidade e por fim entender se o Princípio Cosmológico é uma abordagem científica válida. A Função de Correlação é um dos principais métodos estatísticos utilizados dentro da cosmologia, popularizada na área por [Groth e Peebles \(1977\)](#).

Utilizando a análise de campos estocásticos, partindo da hipótese mais simples, que é imaginar um campo uniforme com flutuações e então comparar qualquer distribuição com tal campo. A função nos diz o quanto uma distribuição qualquer difere de uma distribuição aleatória e uniforme, gerada por um Processo Pontual de Poisson Homogêneo. O Processo Pontual de Poisson Homogêneo é um método utilizado para gerar dados de distribuições homogêneas que sejam estatisticamente independentes. Mais informações serão dadas em capítulos mais avançados. A distribuição gerada por esse processo é homogênea, então caso o comportamento da função de correlação tenda a zero para grandes escalas espaciais, poderíamos dizer que, em tal escala, o Princípio Cosmológico poderia ser aplicado.

A Função é uma interessante ferramenta de análise de distribuição estatística que pode nos dar informações sobre distribuições de pontos quaisquer. Utilizando apenas a localização espacial de objetos, pode ser aplicada em diversas escalas. Em uma das primeiras aplicações na cosmologia, [Totsuji e Kihara \(1969\)](#) mostra, para dados reais, que a Função de Correlação pode ser aproximada por uma lei de potências. Quer dizer que para cada escala de distância, o grau de homogeneidade varia exponencialmente. Existem distâncias características que podem ser identificadas pela função, apontando a transição da escala em que há aglomeração para uma escala em que a distribuição começa a ser homogênea.

Como visto em [Einasto, Saar e Klypin \(1986\)](#), onde função é aplicada para amostras de dados observacionais de distribuições de galáxias e vazios no universo, há um comprimento de correlação, uma escala limite até onde se pode dizer que há baixa homogeneidade. Existem duas definições para o comprimento de correlação. Para a Física Estatística, é uma distância na qual a “memória espacial” de flutuações locais persistem, ou seja, onde o efeito de pequenas perturbações no sistema é sentido. Já para a estatística cosmológica, o comprimento de correlação representa onde a amplitude de flutuações de densidade começa a se tornar pequena demais em comparação com a densidade média dos objetos ao todo. Um bom material para o entendimento é *“Statistical Physics for Cosmical Structures”* ([Gabrielli et al., 2005](#)).

Já em [Einasto et al. \(1997\)](#), são analisados dados de distribuições sintéticas que estudam como a Função de Correlação de dois pontos varia em relação à geometria da

distribuição gerada. Em largas escalas a função pode mostrar qual a geometria de uma distribuição de super aglomerados e em pequenas escalas a função varia com o formato e o tamanho da distribuição.

1.4 Objetivo

1.4.1 Geral

Assim como nossa sociedade têm expandido a tecnologia, também nossa visão do universo vem se expandindo. Com a utilização da interferometria, telescópios espaciais e outras tecnologias, a quantidade de dados observacionais cresce em uma escala em que é humanamente impossível processar. Entender como a matéria é distribuída no universo é um problema complexo que exige uma abordagem estatística devido a grande quantidade de objetos a serem analisados. A cosmologia busca construir uma teoria física coerente para explicar diversas observações do universo em larga escala.

Este trabalho é o início de um estudo com o objetivo de entender como se comporta e como está distribuída a matéria no universo como um todo.

1.4.2 Específico

Como mencionado anteriormente, a Função de Correlação é um dos métodos mais influentes no estudo estatístico da distribuição de matéria no universo. Sendo utilizada desde a metade do século XX para determinar o grau de homogeneidade de distribuições e quantificar até que ponto o Princípio Cosmológico pode ser aplicado. Através de análise de dados numéricos de modelos com distribuições sintéticas, o objetivo do trabalho é estudar a influência de variações de parâmetros estatísticos e de arranjos de distribuições na Função de Correlação de Dois Pontos. Com esse estudo teremos um panorama mais claro de quais as características da Função de Correlação de dois pontos serão realmente úteis ao extrair informações de dados reais, assim como as limitações desta ferramenta.

Parte II

Metodologia

2 Metodologia

“When inhomogeneties are considered (if at all) they are treated as unimportant fluctuations amenable to first order variational treatment. Mathematical complexity is certainly an understandable justification, and economy or simplicity of hypotheses is a valid principle of scientific methodology: but submission of all assumptions to the test of empirical evidence is an even more compelling law of science.”
Gerard de Vaucouleurs (1970)

2.1 Dados Sintéticos

A análise da Função de Correlação de Dois Pontos (FC2P) será feita com dados sintéticos controlados, ou seja, distribuições cujas características são controladas e conhecidas. As principais distribuições utilizadas neste trabalho são Distribuições Gaussianas, Distribuições Beta e Distribuições Homogêneas. Na seções [2.1.1](#) e [2.1.2](#) serão explicados como são geradas nuvens de pontos seguindo as distribuições Gaussianas e distribuições Betas respectivamente. A distribuição homogênea é obtida através de um Processo Pontual de Poisson Homogêneo, que será melhor explicado na seção [2.2.2](#). A vantagem da utilização deste método é que todas as características das distribuições analisadas com a função já estarão definidas, servindo como base de comparação e para futuras análises.

As distribuições foram geradas em Python, onde é possível determinar o tipo de distribuição, quantidade de pontos, afastamento entre pontos, centro (valor médio) para nuvens de pontos, afastamento entre nuvens e todas as outras possíveis características da nuvem. Assim, em um ambiente controlado, teremos certeza de que os resultados encontrados representam características das amostras. Como a FC2P pode ser utilizada em qualquer escala de distâncias e estamos usando dados sintéticos, a unidade de espaço utilizada será arbitrária ou, *u.m.* (unidade de medida).

2.1.1 Distribuição Gaussiana

A Figura 8 representa duas distribuições sintéticas geradas por uma densidade de probabilidade Gaussiana:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (2.1)$$

onde μ representa a média da distribuição e σ o desvio padrão, uma medida da dispersão dos pontos em torno da média. A Figura 7 é um exemplo de distribuição de probabilidade Gaussiana. As nuvens gaussianas (Figura 8) foram feitas gerando distribuições de pontos nas três dimensões (x, y, z) seguindo a lei descrita pela equação 2.1. É possível ver que em uma distribuição Gaussiana teremos uma probabilidade maior de encontrar pontos perto do centro das nuvens (ponto médio) do que nas extremidades. O desvio padrão nos dirá o quão concentrada ou espalhada (distante do centro) estarão os pontos de uma nuvem.

As nuvens da Figura 8 foram feitas cada uma com dez mil pontos, centradas nos pontos $(2,2,2)$ *u.m.* e $(8,8,8)$ *u.m.* e ambas com desvio padrão de $\sigma = 0.5$ *u.m.* Na Figura 9 temos uma projeção bidimensional das mesmas nuvens. Projeções bidimensionais serão utilizadas eventualmente pois podem facilitar a visualização.

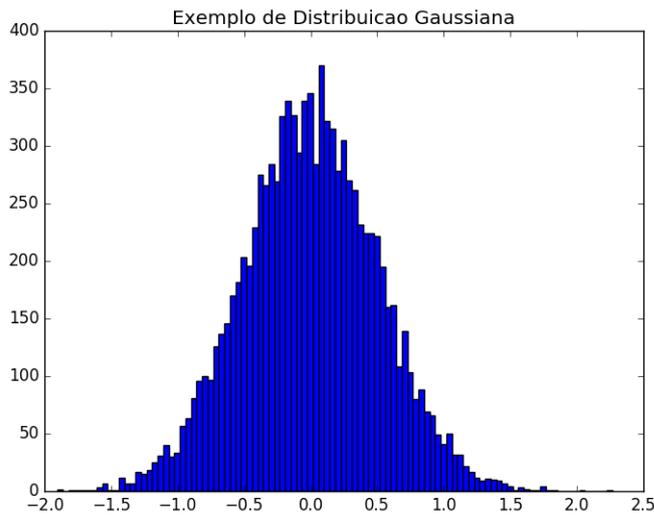


Figura 7 – Densidade de Probabilidade Gaussiana

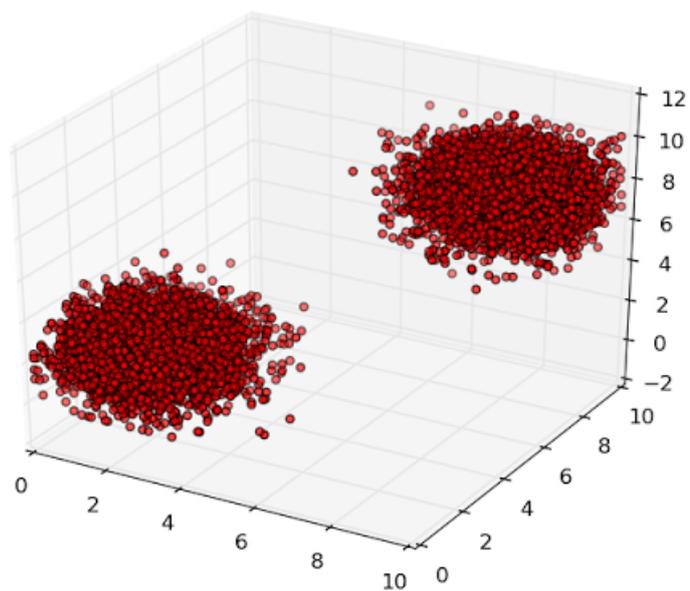


Figura 8 – Exemplo de Dado Distribuição Gaussiana Tridimensional.

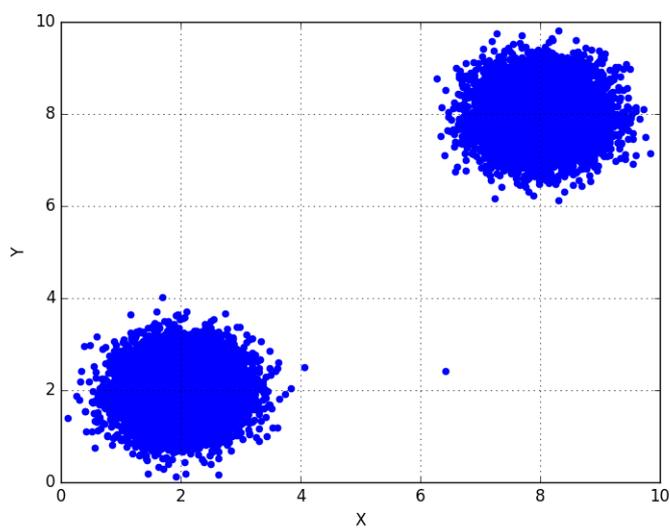


Figura 9 – Projeção Bidimensional de Dados Sintéticos

2.1.2 Distribuição Beta

A distribuição Beta segue a lei de distribuição:

$$p(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} \quad (2.2)$$

para $0 < x < 1$ e coeficientes arbitrários $\alpha, \beta > 0$. No modelo, os parâmetros α e β definem a forma da distribuição. Se $\alpha = \beta$, a distribuição é simétrica, se $\alpha > \beta$, a assimetria é negativa e, no caso de $\alpha < \beta$, sua assimetria é positiva. Já o desvio padrão σ está relacionado com os coeficientes α e β na forma:

$$\sigma(\alpha, \beta) = \sqrt{\frac{\alpha\beta}{(\alpha + \beta + 1)(\alpha + \beta)^2}} \quad (2.3)$$

onde $B(\alpha, \beta)$ é uma função de normalização:

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx \quad (2.4)$$

Na Figura 10 temos um exemplo de distribuição Beta com $\alpha = \beta = 0.3$. Neste caso é possível ver que a probabilidade de termos pontos nos extremos é maior do que ter pontos no centro da distribuição, se comparado com uma distribuição Gaussiana, como mostrado na Figura 11. A nuvem Beta (Figura 11) foi feita gerando distribuições de pontos nas três dimensões (x, y, z) seguindo a lei descrita pela equação 2.2. A distribuição Beta varia de acordo com a combinação de α e β escolhida, podendo variar do exemplo mostrado na Figura 10 até uma distribuição aproximadamente Gaussiana. Embora a distribuição Beta não tenha significado físico, sua forma nos dá uma ideia de uma distribuição com uma densidade maior de ponto nas extremidade e uma densidade menor de pontos (ou um vazio) no centro. A Função de Correlação de Dois Pontos de vazios pode ser útil em análises de distribuições reais, portanto, se bem conhecida, pode nos dar informações importantes.

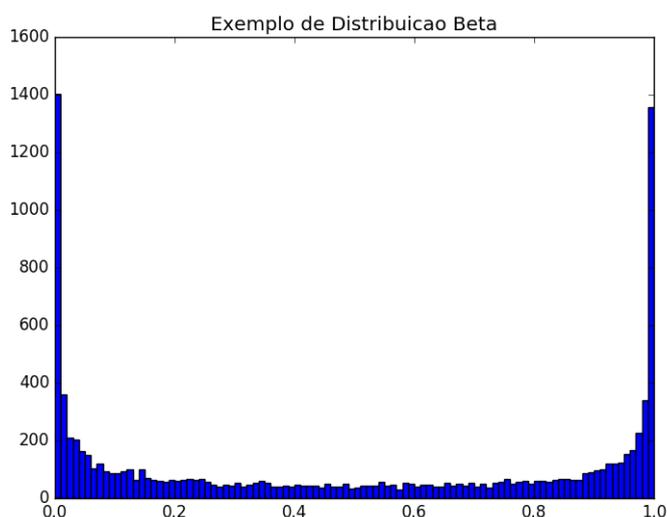


Figura 10 – Exemplo de Distribuição Beta

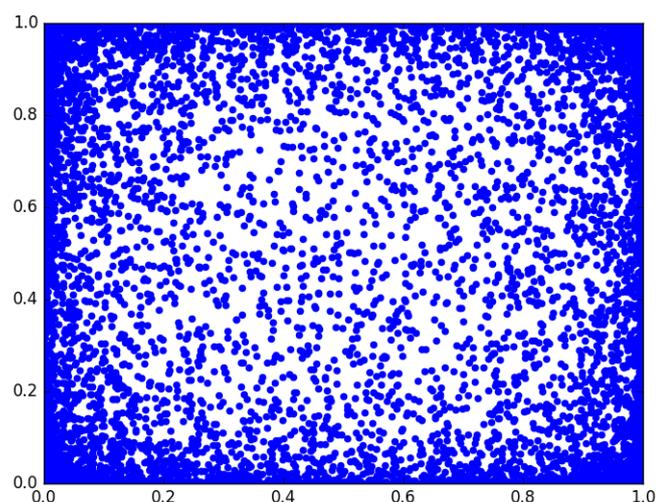


Figura 11 – Exemplo Tridimensional de Nuvem Beta, com 10 mil pontos, projetada Bidimensionalmente. Neste caso podemos ver uma concentração maior de pontos nas bordas, o que pode ser utilizado para estudar propriedades de vazios.

2.2 Função de Correlação de Dois Pontos

A FC2P é uma formulação específica de Função de Correlação em que são comparadas distribuições de frequências de distâncias entre dois pontos de dois tipos de amostra. Uma delas é uma distribuição homogênea aleatória gerada por um Processo Pontual de Poisson Homogêneo e a outra é a amostra utilizada para análise, podendo ser sintética ou de dados reais de galáxias. Valores positivos da FC2P representam uma aglomeração (amostra positivamente correlacionada), caso a função seja negativa há vazios (amostra anti-correlacionada) e se a função for nula a quantidade dessas distâncias é igual a de

uma distribuição homogênea (amostra correlacionada). Mais detalhes serão dados na seção 2.2.3 sobre a implementação da função.

Para entendermos melhor o significado dos valores da FC2P primeiro deveremos entender como são feitas as distribuições de frequência de distâncias e posteriormente como e o que é um Processo Pontual de Poisson.

2.2.1 Distribuição de Frequência de Distâncias

Para entender como é feita a distribuição de frequências vamos começar de maneira simples. Imaginemos quatro pontos dispostos no vértice de um quadrado de aresta igual a 1 *u.m.*, como mostrado na Figura 12. Ao montar a distribuição de frequência de distâncias, devemos contabilizar todas as distâncias entre pontos sem repetição (por exemplo: distâncias do ponto 1 ao ponto 2 e do ponto 2 ao ponto 1 serão consideradas a mesma, sendo contabilizada apenas uma vez). No espaço euclidiano, teremos que a distância entre dois pontos é $d = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}$ onde x, y, z são as coordenadas dos pontos e os índices i e j representam os pontos dos quais a distância está sendo calculada.

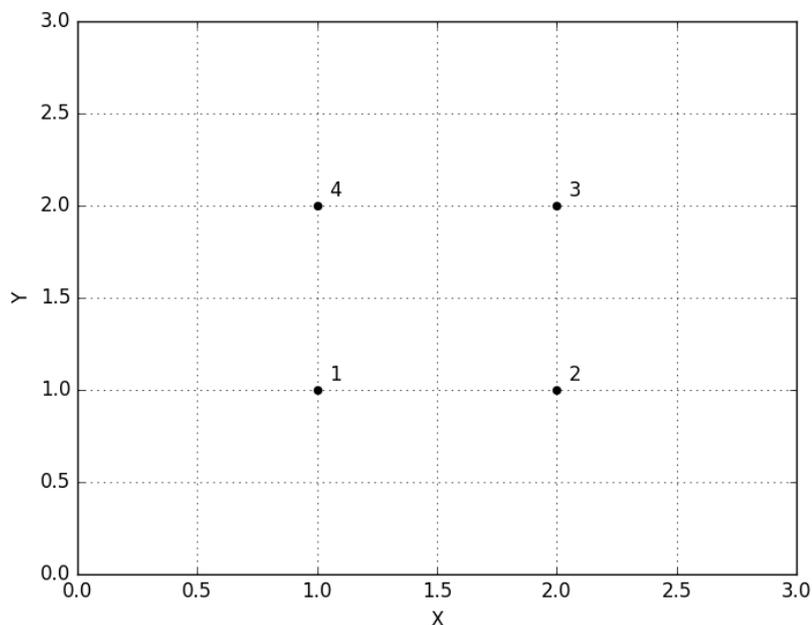


Figura 12 – Exemplo de Distribuição de Pontos.

Assim, calculando as distâncias entre todos os pontos sem repetição, podemos construir a Tabela 1. Agora temos a frequência com que as distâncias se repetem. Podemos ver pela tabela que há quatro distâncias de 1 *u.m.* (nas arestas) e duas distâncias de $\sqrt{2} \approx 1,41$ *u.m.* (entre os vértices), então a contagem de distâncias para esse caso será de 4 distâncias para a distância 1 *u.m.* e de 2 distâncias para a distância 1,41 *u.m.*.

Evidentemente este é um caso simples, o caso em que estamos interessados envolve uma quantidade de pontos muito maior.

Pares	Distâncias (<i>u.m.</i>)
1 e 2	1
1 e 3	$\sqrt{2}$
1 e 4	1
2 e 3	1
2 e 4	$\sqrt{2}$
3 e 4	1

Tabela 1 – Distâncias entre pares de pontos

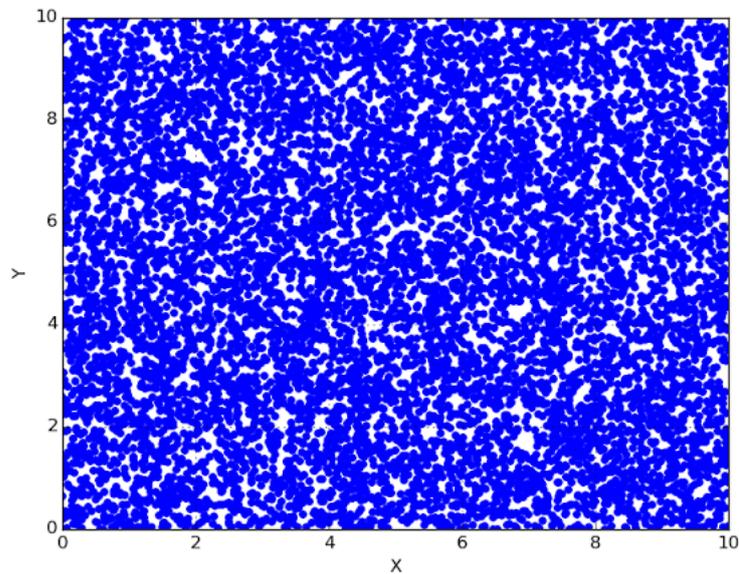


Figura 13 – Exemplo de Distribuição de 10 mil Pontos projetada no plano.

A Figura 13, com dez mil pontos, representa algo mais próximo do que queremos. A mesma ideia aplicada para a Figura 12 será utilizada aqui, calculando todas as distâncias entre todos os pontos sem repetição. O resultado é mostrado na Figura 14, onde a abscissa representa as distâncias e a ordenada o número de vezes que essa distância aparece. A Figura 14 também é conhecida como um histograma, pois associa uma variável a uma frequência de classe - neste caso distância e frequência de distâncias.

Como é possível ver nesta figura, para distâncias muito pequenas temos uma baixa frequência. A medida que as distâncias aumentam, aumentam também a sua frequência, até que a partir de determinada distância, o centro da distribuição, os valores voltam a diminuir. Isso acontece porque para distâncias pequenas há poucos vizinhos, essa quantidade vai aumentando e com ela aumenta a contagem de distâncias. Mas ao se aproximar de distâncias maiores ainda, devido ao fato da amostra ser finita, a contagem de distâncias começa a cair

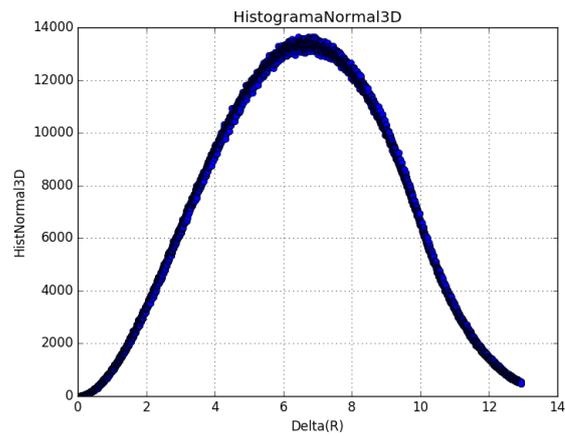


Figura 14 – Exemplo de Distribuição de Frequências de Distâncias para 10 mil Pontos.

novamente. As distâncias muito pequenas são tão frequentes quanto distâncias próximas ao valor máximo. Esta é uma explicação genérica para a contagem de distâncias. Os valores máximos e de inflexão dependem da distribuição analisada.

2.2.2 Processo Pontual de Poisson Homogêneo

O Processo Pontual de Poisson Homogêneo é um processo que garante que a probabilidade de encontrar um ponto independe de outros pontos gerados anteriormente na distribuição. Teremos agora uma explicação do algoritmo que forma o Processo Pontual de Poisson Homogêneo baseado na seção 3.2.3 de [Martínez \(2001\)](#). Neste trabalho focaremos em sua aplicação discreta para gerar os dados sintéticos.

A formação do Processo Pontual consiste na geração de regiões de volume onde pontos sejam dispostos aleatoriamente e o número de pontos em cada região se dá de maneira probabilística. No caso do Processo Pontual de Poisson a distribuição de probabilidade estatística é a de Poisson dada por:

$$p(k, \lambda) = \frac{\lambda e^{-\lambda}}{k!} \quad (2.5)$$

onde k é o número de eventos esperados que ocorrem no intervalo λ . No caso homogêneo, o intervalo λ é mantido constante. Um número de pontos, $p(k, \lambda) = N$, é gerado para cada região. Em seguida N pontos são dispostos aleatoriamente em uma região, ou “caixa”. É calculado um novo N para cada caixa, repetindo o processo várias vezes. Diversas caixas são geradas da mesma forma como mostrado na Figura 15.

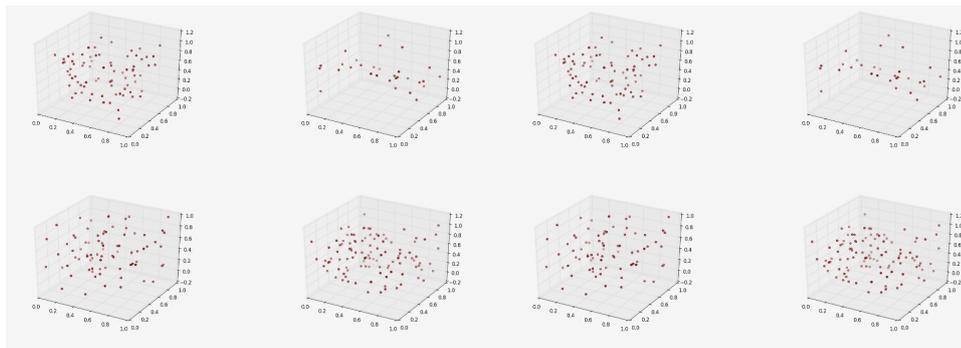


Figura 15 – Caixas com número de pontos gerado por uma distribuição de Poisson

O somatório das caixas gera uma distribuição Homogênea e a probabilidade de seus pontos são independentes. Quanto mais caixas forem formadas para a distribuição, melhor essa característica é garantida. Se pegarmos mil caixas de dez pontos (como as da Figura 15) e somarmos, podemos ter algo como na Figura 16. Este é então o método utilizado neste trabalho para gerar distribuições homogêneas. A contagem de distâncias de tal distribuição se dá como a mostrada pela Figura 14.

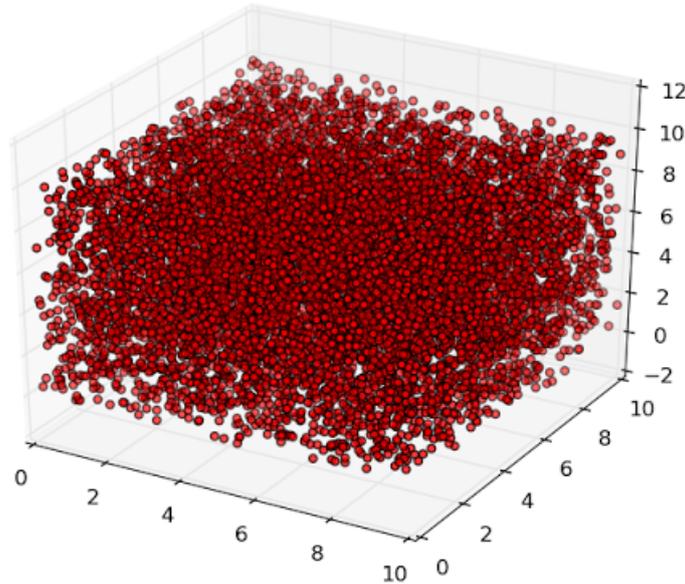


Figura 16 – Amostra gerada por um Processo Pontual de Poisson Homogêneo

2.2.3 Implementação

A FC2P apresenta diversas fórmulas para sua representação discreta. Como é visto em [Hamilton \(1993\)](#), existem formas mais complexas da função que corrigem efeitos causados pelas bordas das amostras mas que, para fins do estudo do comportamento da função, não influenciam tão fortemente em suas principais características. Portanto, a fórmula discreta aplicada será a utilizada em [Einasto, Saar e Klypin \(1986\)](#),

$$\xi(r) = \frac{\langle DD(r) \rangle n_r}{\langle RR(r) \rangle n_d} - 1 \quad (2.6)$$

onde $\langle DD(r) \rangle$ é contagem de distâncias em um intervalo $r \pm dr$ entre pares de pontos da distribuição estudada e $\langle RR(r) \rangle$ é a contagem de distâncias para uma distribuição gerada por um Processo Pontual de Poisson Homogêneo. A razão $\frac{n_r}{n_d}$ é um coeficiente de normalização para que as amostras não precisem ter necessariamente o mesmo número de pontos. Os valores n_r e n_d representam, respectivamente, o número de distâncias na amostra homogênea e na amostra estudada.

Ao dividir a distribuição de frequências (ou contagem de distâncias) de uma amostra pela distribuição de frequências gerada por uma distribuição homogênea de pontos teremos valores que nos indicam quantas vezes mais determinadas distâncias estão presentes na amostra em relação à amostra homogênea. Ao diminuir uma unidade, contagens de distâncias que são iguais se tornam nulas na FC2P (distâncias não correlacionadas). Caso haja mais distâncias na amostra do que na distribuição homogênea o valor da função de

correlação será positivo (positivamente correlacionadas), caso o número de distâncias seja na amostra seja menor teremos valores negativos na função (anti correlacionadas) o que nos indicará a presença de vazios ou regiões de baixa densidade.

Em [Einasto et al. \(1997\)](#) os autores estudam o comportamento da FC2P variando distribuições de pontos. Neste artigo, é possível ver que os valores da FC2P estão em escala logarítmica, já que os gráficos apresentados variam muito relação às escalas de distâncias, começando por um pico muito grande depois caindo rapidamente a zero e oscilando em torno desse valor.

Como há uma maior variação no primeiro pico da FC2P, normalmente a função, nesta região, é aproximada por uma lei de potências. A distância na qual o primeiro pico da FC2P chega a zero é chamada de comprimento de correlação que, segundo [Gabrielli et al. \(2005\)](#), é uma escala na qual as flutuações dos valores de distâncias se tornam pequenas amplitudes em relação à densidade média da distribuição de distâncias de uma distribuição homogênea de pontos. O comprimento de correlação é uma distância característica na qual as amostras tendem à homogeneidade, principalmente para dados reais. A FC2P normalmente apresenta um comportamento oscilatório tendo picos e vales. O centro dos picos representa distâncias características¹ da FC2P. Tais distâncias e a relação com os picos serão melhor entendidos na seção 3.1 e uma análise mais representativa das distribuições dos picos será feita na seção 3.1.1. Utilizando a Equação 2.6 é implementado o cálculo discreto da FC2P na linguagem de programação Python.

¹ Tais distâncias normalmente são associadas ao tamanho das distribuições e ao afastamento entre distribuições

Parte III

Resultados

3 Apresentação dos Dados

3.1 Primeiros Estudos

3.1.1 Estudo dos picos

Primeiramente geramos uma amostra de nuvens gaussianas centralizadas nas arestas de um cubo. Esta distribuição simples nos ajudará a FC2P resultante e sua relação com as distribuições.

3.1.1.1 Nuvens Gaussianas

Foram geradas oito distribuições de pontos que seguem uma lei gaussiana (Equação 2.1). O centro de cada distribuição é o vértice de um cubo de aresta 50 *u.m.* com coordenadas dos centros (x, y, z) : (10,10,10), (10,10,60), (10,60,60), (60,60,60), (10,60,10), (60,10,10), (60,60,10) e (60,10,60). O desvio padrão, σ , de todas as distribuições é 5 *u.m.*

As Figuras 17 e 18 nos mostram como as distribuições estão dispostas. A Figura 18 aponta de maneira mais didática as principais distâncias envolvidas na distribuição, excluindo apenas a distância que cruza o cubo, pois a figura é apenas uma projeção bidimensional das distribuições. De maneira análoga ao explicado na seção 2.2.1 as distâncias entre todos os pontos devem ser contadas. Uma análise superficial da figura nos dá uma ideia de quais devem ser os picos esperados, pois estes estão associados à distâncias características das distribuições. O diâmetro de cada nuvem está em torno de 10 *u.m.* (já que $\sigma = 5$ *u.m.*), logo as principais distâncias envolvidas devem ser:

- 10 *u.m.*, escala interna das nuvens;
- 50 *u.m.*, distâncias (entre pontos das nuvens) que não cruzam a face do cubo nem o cubo;
- $50\sqrt{2}$ *u.m.*, distâncias (entre pontos das nuvens) diagonais da face do cubo;
- $50\sqrt{3}$ *u.m.*, distâncias (entre pontos das nuvens) nas diagonais do cubo;

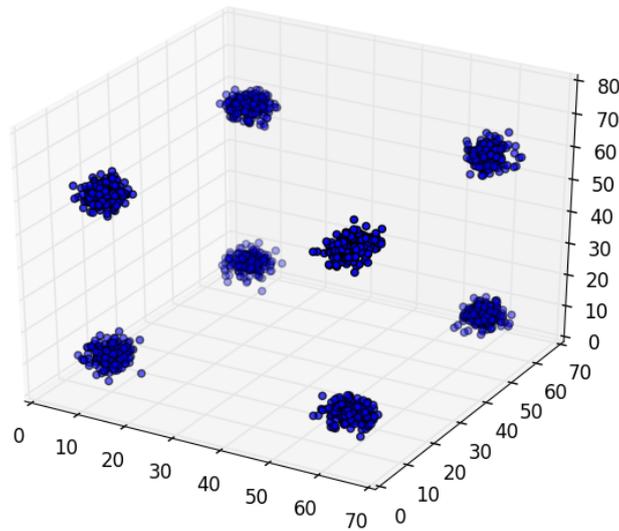


Figura 17 – Significado dos picos: 8 Nuvens

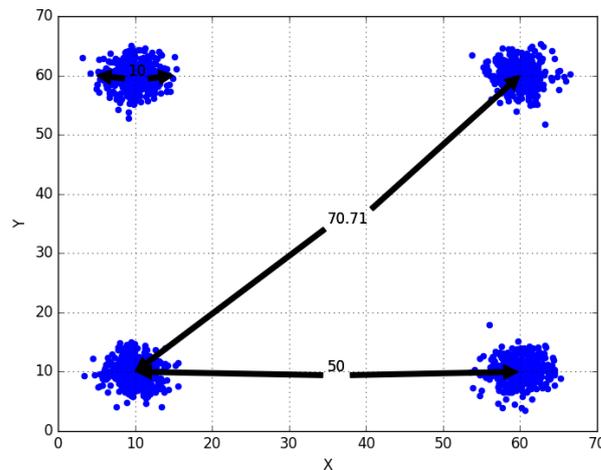


Figura 18 – Significado dos picos: 8 Nuvens projetadas em duas dimensões

3.1.1.2 Histograma e Correlação

As Figuras 19 e 20 representam, respectivamente, o histograma de contagem de distâncias para a distribuição de oito nuvens e a FC2P calculada para esta mesma distribuição. A Tabela 2 nos dá as posições dos picos das duas figuras.

Podemos observar no primeiro histograma (Figura 19) que o primeiro pico está muito próximo de 5 u.m. que é o desvio padrão das nuvens. O segundo pico está próximo de 50 u.m. , que é a distância entre as nuvens. O terceiro pico está próximo de $50\sqrt{2} \text{ u.m.}$ que é a distância entre nuvens na diagonal da face do cubo. O quarto pico está próximo de $50\sqrt{3} \text{ u.m.}$, que é a distância na diagonal atravessando o cubo. Na Figura 20, o segundo

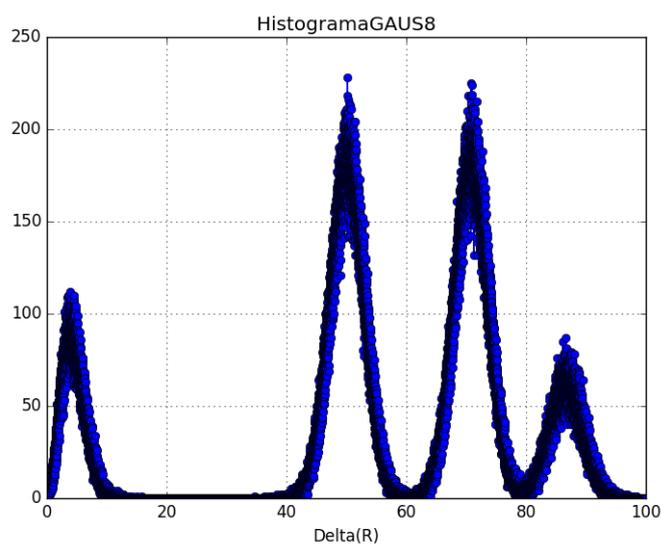


Figura 19 – Histograma distribuição de estudo dos picos

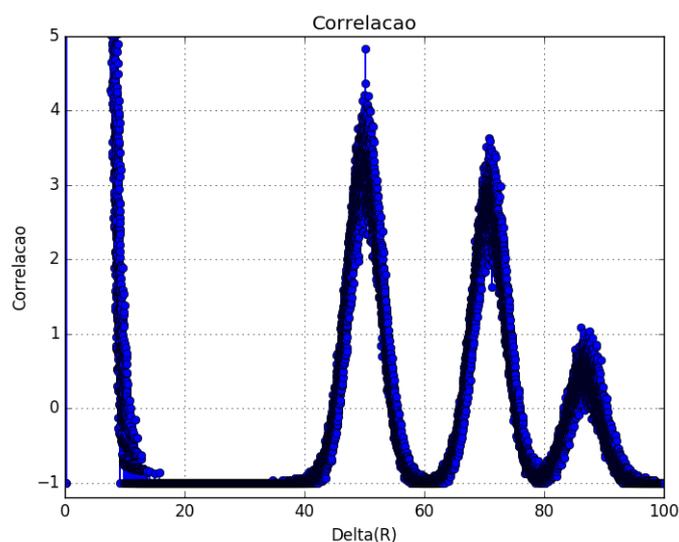


Figura 20 – Correlação distribuição de estudo dos picos

pico se aproxima do diâmetro das nuvens, o terceiro, quarto e quinto picos seguem a mesma lógica do histograma da Figura 19.

Picos (<i>u.m.</i>)	P1	P2	P3	P4	P5
Histograma	4,78	50,04	71,74	87,67	
Correlação	0,25	9,15	50,04	71,74	87,67

Tabela 2 – Picos de distribuição 8 Nuvens: Histograma e Correlação

Pela definição da FC2P (Vide seção 2.2.3) os picos do histograma e da FC2P não devem divergir. No entanto o primeiro pico está em um valor um pouco diferente. Como discutiremos adiante, o primeiro pico é caótico, pois pequenas variações nas condições

iniciais afetam de forma significativa seu valor, podendo ser completamente diferente para distribuições estatisticamente iguais. Tal característica deve ser melhor estudada na próxima seção.

3.1.1.3 Variações no primeiro pico

Para entender melhor o primeiro pico da FC2P foram feitas dez distribuições de pontos estatisticamente iguais. As distribuições são compostas por duas nuvens com mil pontos cada, centradas em 25 *u.m.* e 75 *u.m.*, com desvio padrão, σ , igual a 5 *u.m.*. A Figura 21 dá uma representação visual de como cada distribuição se dá espacialmente projetada no plano. A distribuição homogênea gerada tem vinte mil pontos (dez vezes mais pontos que a amostra testada). A FC2P para cada uma das distribuições foi calculada. Em seguida uma FC2P média foi calculada¹, afim de diminuir erros causados por variação estatística. Embora as distribuições sejam estatisticamente iguais a disposição dos pontos é aleatória, logo pequenos erros estatísticos devem surgir.

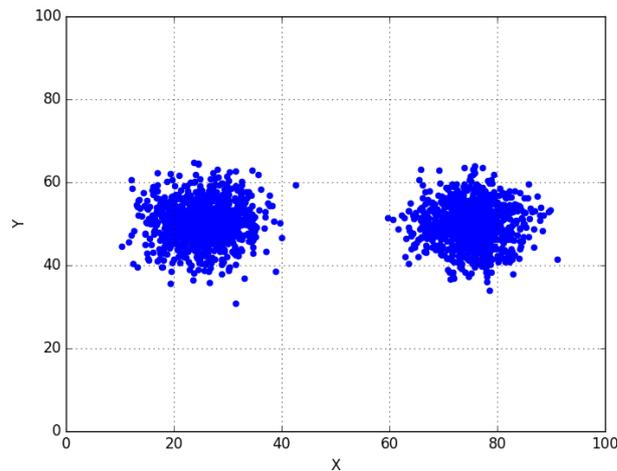


Figura 21 – Visualização das distribuições Iguais. As nuvens foram projetadas bidimensionalmente para melhor visualização.

O processo de calcular a média das FC2P foi repetido três vezes, de maneira que três FC2P média foram calculadas e estão dispostas nas Figuras 22 e 23. As FC2P foram suavizadas, dividindo as curvas em 51 segmentos interpolados por polinômios de terceiro grau. A suavização afeta um pouco a amplitude dos gráficos, mas não a posição dos picos, característica que estamos interessados nesse trabalho. Portanto podemos seguir com a análise da posição dos picos sem preocupações sobre erros associados à suavização.

É possível ver que mesmo se tratando de FC2P médias suavizadas, o primeiro pico para cada uma delas é completamente diferente. Pequenas variações estatísticas geram grandes variações no primeiro pico da função.

¹ A lógica para diminuir o erro estatístico é análoga à utilizada por Einasto et al. (1997), onde as FC2P são a média de 10 FC2P.

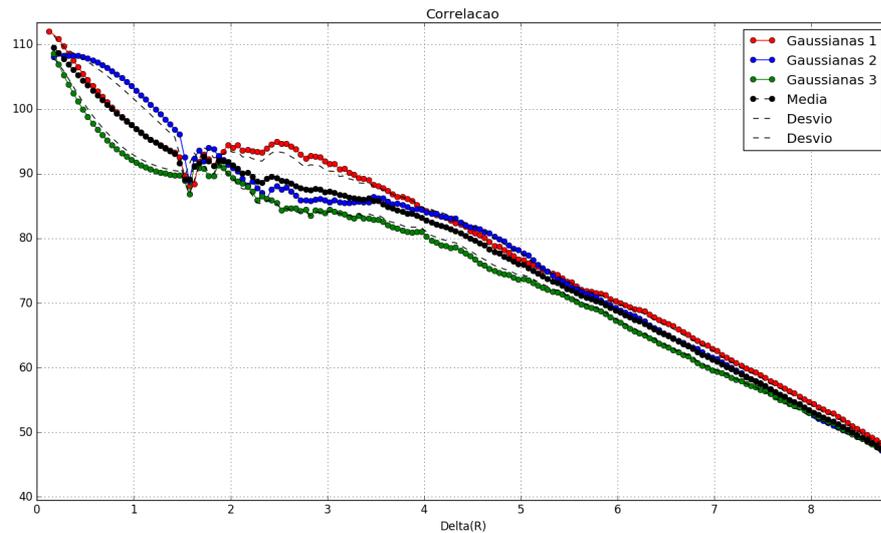


Figura 22 – Correlações médias com filtro de suavidade. As curvas coloridas representam a média de 10 FC2P, a curva preta é a média das 3 curvas médias. As linhas tracejadas são calculadas a partir do desvio padrão de 30 FC2P de distribuições estatisticamente iguais. A variação do primeiro pico é grande mesmo em uma função suavizada e pode ficar fora do desvio padrão das curvas.

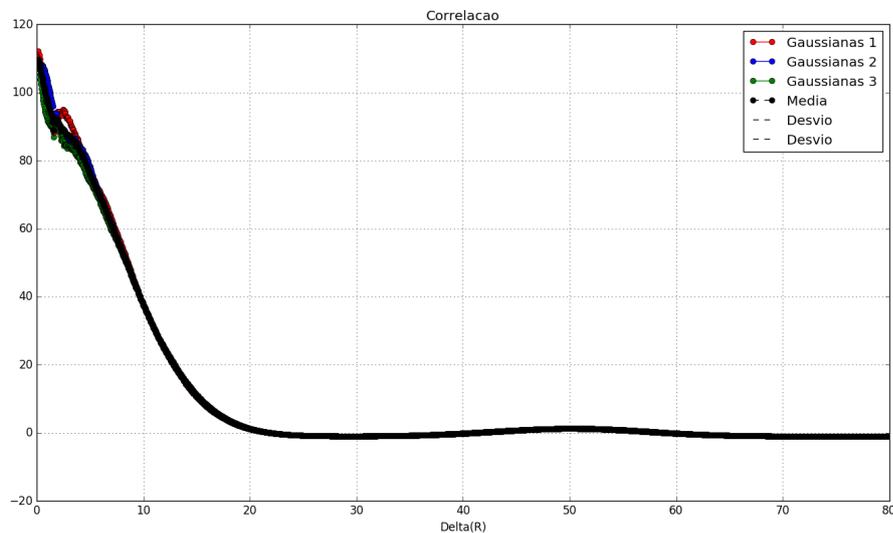


Figura 23 – Correlações médias. O primeiro pico se mostra completamente instável, mas a partir de certa escala de distância as curvas da FC2P se tornam indistinguíveis.

A função de correlação divide o histograma de uma amostra pelo histograma de uma distribuição homogênea e uniforme. O histograma de uma distribuição homogênea e uniforme tende a zero nas extremidades (similar à Figura 14), o que deixa a razão muito sensível nessas regiões, aplicando flutuações da ordem do erro causado pelo fato da distribuição ser aleatória. Distribuições iguais, centradas na mesma posição com o mesmo desvio padrão apresentam variação maior nas extremidades do que no resto da função

(Figuras 22 e 23). O primeiro pico não deve ser levado em consideração ao se comparar distribuições diferentes já que diverge mesmo para distribuições iguais. De fato, a literatura não utiliza o valor do primeiro pico. Para analisar esta primeira oscilação é comum que seja feita uma aproximação por uma lei de potências. Com tal lei é encontrado o comprimento de correlação (onde o primeiro pico começa a tender a zero), que seria uma distância na qual a distribuição, para dados reais, passa a ter um comportamento mais homogêneo. No entanto as análises deste trabalho se dão em dados sintéticos e a aproximação por uma lei de potências não tem tanta representatividade.

3.1.2 Teste de Centralidade

Como explicado no capítulo 2, a Função de Correlação de dois pontos depende das distâncias relativas entre os pontos comparadas com uma distribuição homogênea, então é de se esperar que ao mudar a localização das distribuições sem alterar a distância relativa entre as distribuições, não alteramos a função de correlação. Diversos testes foram realizados variando a centralidade. Os testes tiveram os mesmos resultados que podem ser representados com as distribuições geradas a seguir.

Foram geradas duas distribuições gaussianas com mil pontos cada e desvio padrão, $\sigma = 0,5 \text{ u.m.}$. As distribuições têm todos os parâmetros iguais, exceto a posição, pois uma está centrada nas coordenadas (2,2,2) e a outra está centrada nas coordenadas (5,5,5) (Figura 24).

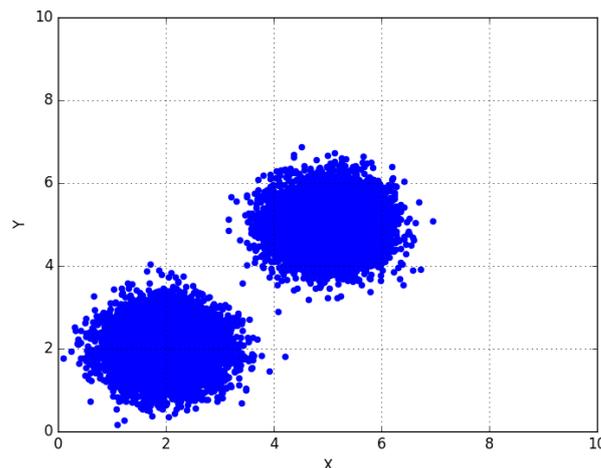


Figura 24 – As nuvens estão centradas em (2,2,2) e (5,5,5) e contém mil pontos cada, $\sigma = 0,5 \text{ u.m.}$ em relação ao centro de cada nuvem.

Ao mover as distribuições para um novo centro, (3,3,3) e (6,6,6) (Figura 25) mantemos a distância relativa entre elas constante, equivalente a $3\sqrt{3} \text{ u.m.}$ Caso a FC2P seja calculada para as mesmas nuvens, apenas mudando a centralização, é obtido, como esperado, exatamente a mesma FC2P, pois a função utiliza apenas a distância relativa

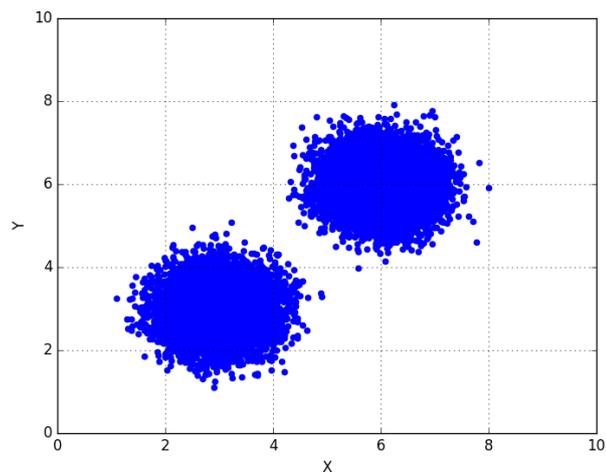


Figura 25 – As nuvens estão centradas em $(3,3,3)$ e $(6,6,6)$ e contém mil pontos cada, $\sigma = 0,5$ u.m. em relação ao centro de cada nuvem.

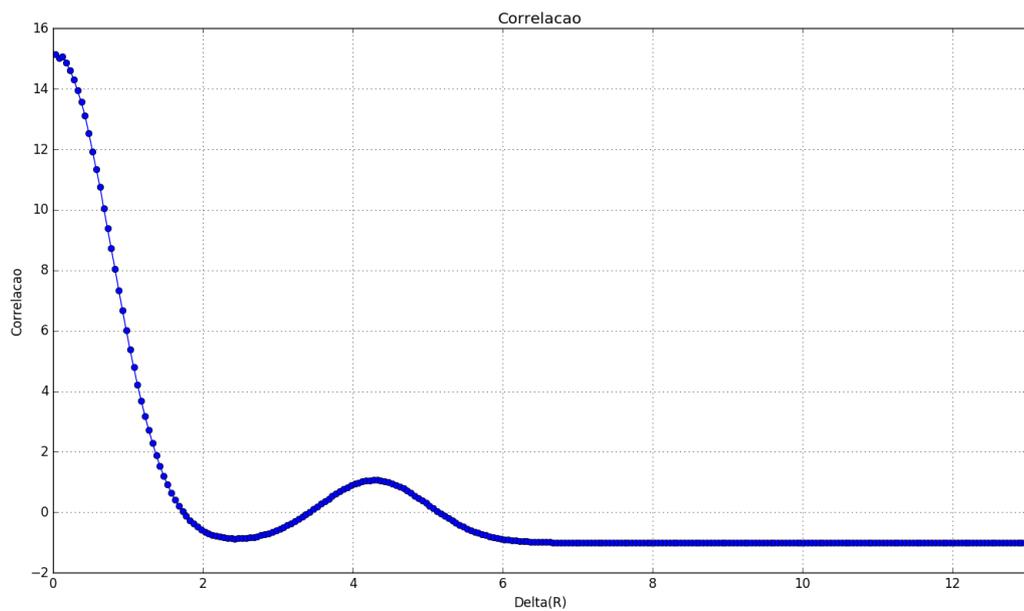


Figura 26 – Função de Correlação de dois pontos suavizada encontrada para as nuvens centradas em 2 e 5 e nuvens centradas em 3 e 6. A amostra é a mesma, a única diferença é o centro deslocado.

entre as nuvens. A FC2P suavizada por 51 polinômios de terceiro grau pode ser vista na Figura 26.

3.2 Testes com Distribuições

Para determinar as principais características das distribuições que influenciam a Função de Correlação de dois pontos foram realizados testes com diferentes distribuições.

3.2.1 Teste Diferentes Distribuições

Outra característica interessante de ser estudada é a influência da função de distribuição utilizada para gerar as nuvens. Foram feitos testes com diferentes distribuições de pontos em cada nuvem. Em cada teste foi alterado a lei de distribuição de uma das nuvens.

Foram mantidos constantes o número de pontos, o centro das nuvens e o desvio padrão (σ) de cada nuvem. A parte relevante em manter o centro das nuvens constantes é que a distância entre os centros se manteve constante e não o centro em si, como estudado na seção 3.1.2. Foram selecionados alguns dos testes que representam os resultados para serem mostrados nesse trabalho. Os testes apresentados foram feitos com a comparação entre uma nuvem Gaussiana, Equação 2.1, e uma caixa Homogênea, gerada pelo processo explicado na seção 2.2.2. As nuvens foram reescaladas de forma que o seu desvio padrão total fosse de $\sigma = 5 \text{ u.m}$, e foram representadas na Figura 27.

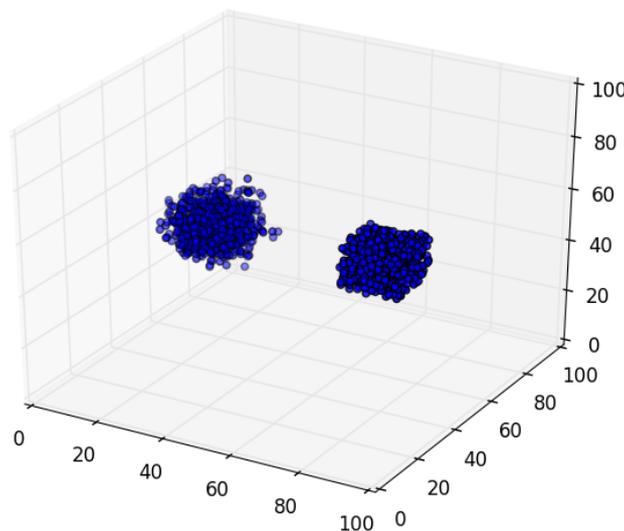


Figura 27 – Nuvens: Gaussiana e Caixa Homogênea

Uma nuvem gaussiana e uma caixa formada por uma lei de distribuição Beta - de maneira análoga à caixa Homogênea, foi gerada uma distribuição Beta (Equação 2.2 com $\alpha = \beta = 0.3$ e então a distribuição foi reescalada para que seu desvio padrão total fosse

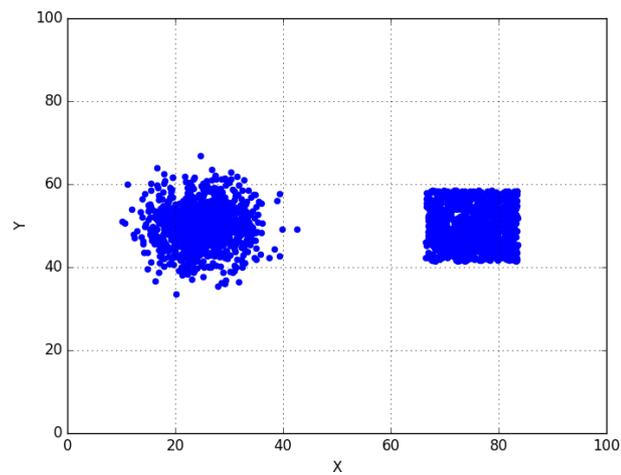


Figura 28 – Projeção 2D de Nuvens: Gaussiana e Caixa Homogênea

$\sigma = 5 \text{ u.m}$ - representadas pela Figura 29 e duas nuvens Gaussianas representadas pela Figura 31. O desvio padrão de todas as nuvens é 5 u.m . e a distância entre o centro das nuvens é 50 u.m .

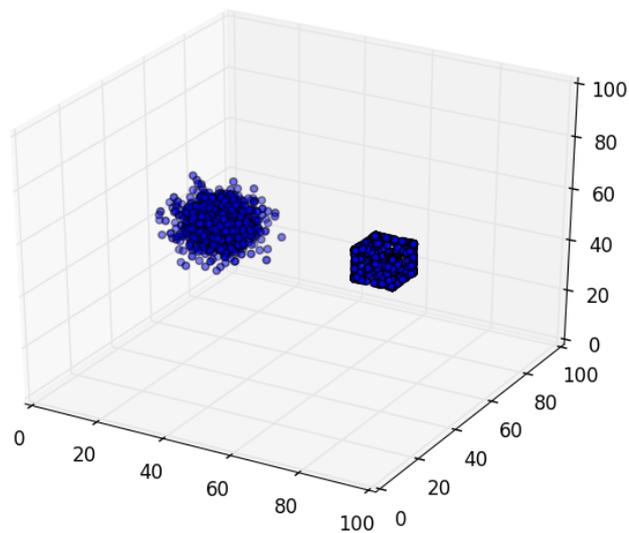


Figura 29 – Nuvens: Gaussiana e Caixa Beta

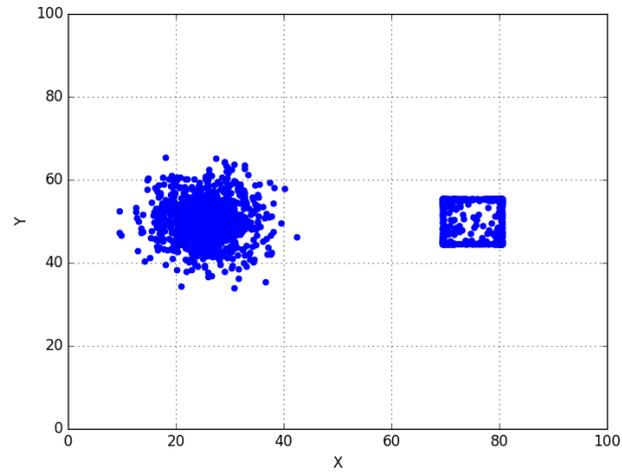


Figura 30 – Projecção 2D de Nuvens: Gaussiana e Caixa Beta

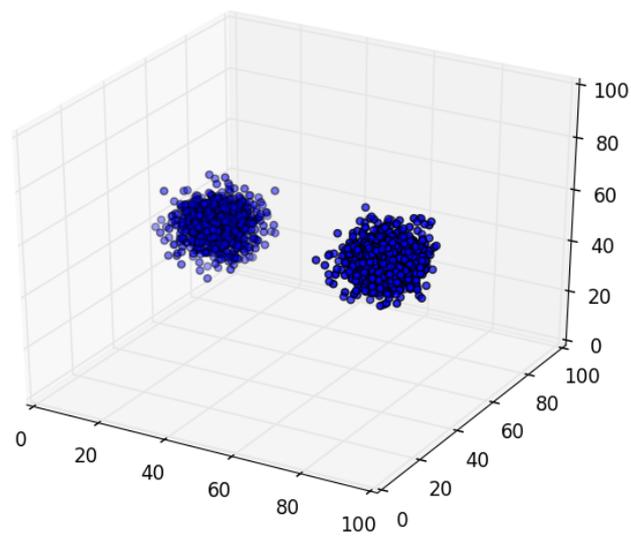


Figura 31 – Nuvens Gaussianas

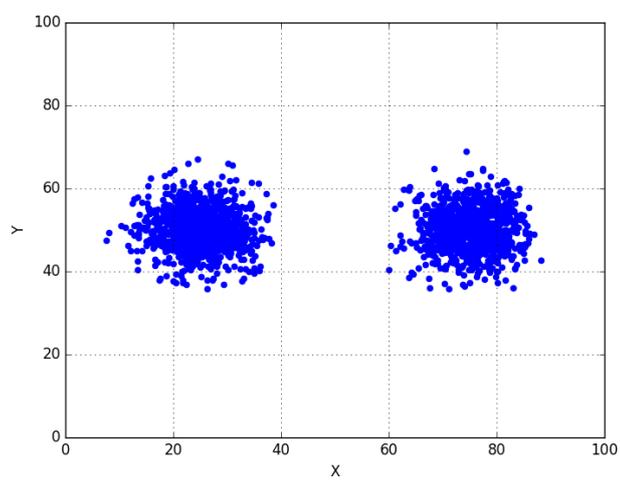


Figura 32 – Projeção 2D de Nuvens: Gaussianas

3.2.1.1 Funções de Correlação de Dois Pontos

A FC2P para cada tipo de distribuição foi calculada a partir da média de 10 FC2P², ou seja, as curvas de correlação são na verdade a média de várias. São esperados dois picos principais em escalas diferentes. Um pico correspondente às distâncias internas de cada nuvem e um pico correspondente às distâncias entre as nuvens. As FC2Ps encontradas foram as ilustradas nas Figuras 33 e 34.

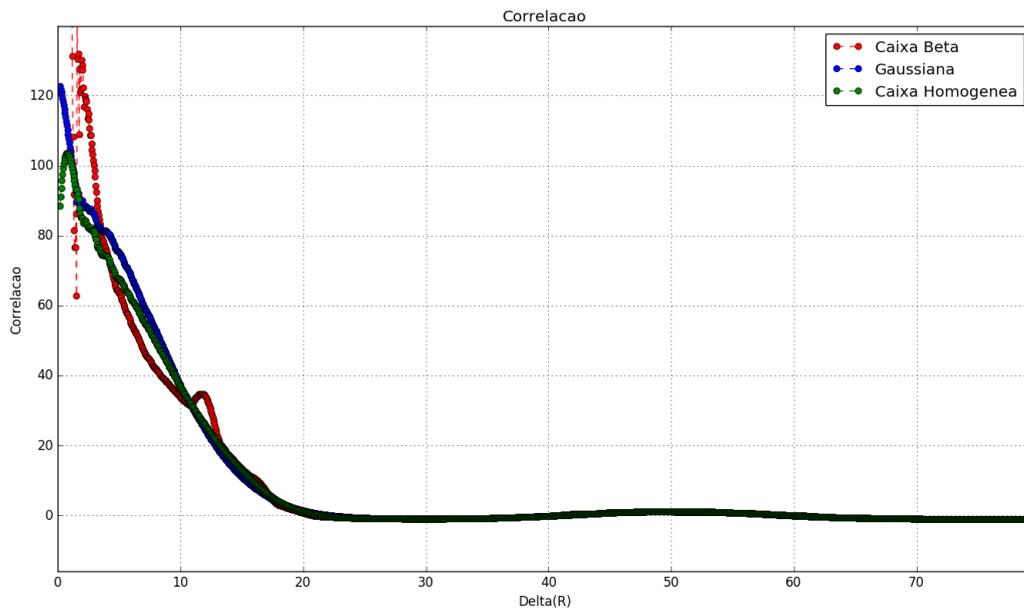


Figura 33 – Comparação Correlação Distribuições Distintas

A Tabela 3 indica a posição dos principais picos de cada distribuição. A função de correlação de duas gaussianas apresentou três picos principais. O que foi muito similar ao encontrado na função de correlação aplicada às distribuições compostas por uma gaussiana e um quadrado homogêneo. A escala de distâncias do terceiro pico é correspondente ao afastamento entre as nuvens. Na Figura 33 podemos ver que a primeira oscilação é

Picos (u.m.)	P1	P2	P3	P4	P5	P6	P7	P8	P9
Gaussianas	0,675	3,875	50,125						
Homogênea	0,775	3,725	49,625						
Beta	1,925	11,675	42,975	45,625	48,825	51,575	52,225	53,525	56,075

Tabela 3 – Picos de distribuição diferentes: Duas nuvens gaussianas, nuvem gaussiana e nuvem cúbica homogênea, nuvem gaussiana e nuvem beta

completamente diferente para as três curvas, no entanto esta região não serve para nos garantir que a FC2P consegue diferenciar distribuições diferentes. Como estudado na seção

² Análogo ao feito em Einasto et al. (1997).

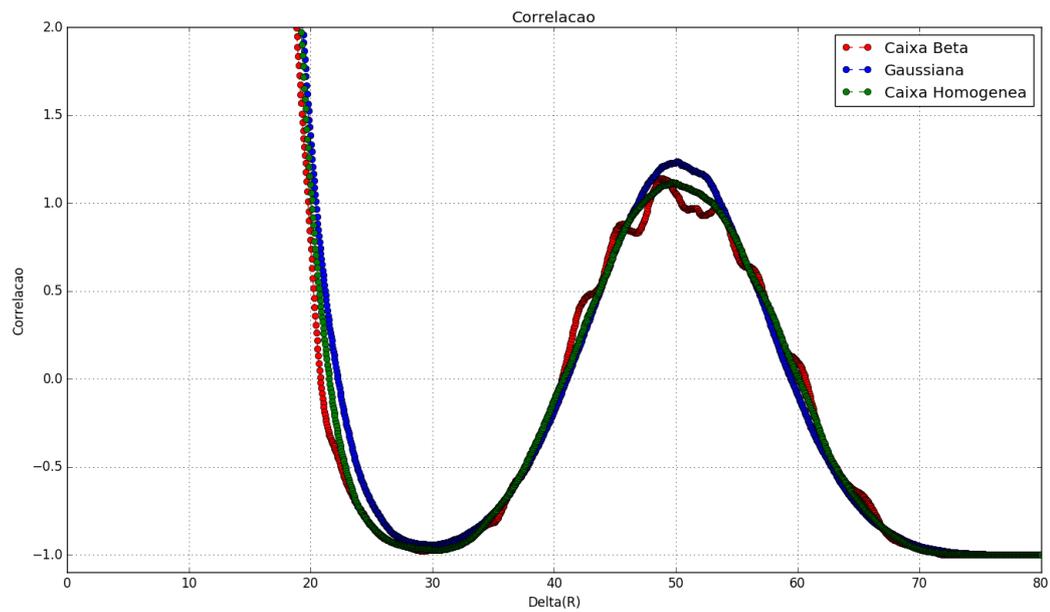


Figura 34 – Comparação Correlação Distribuições: Escala do Segundo pico

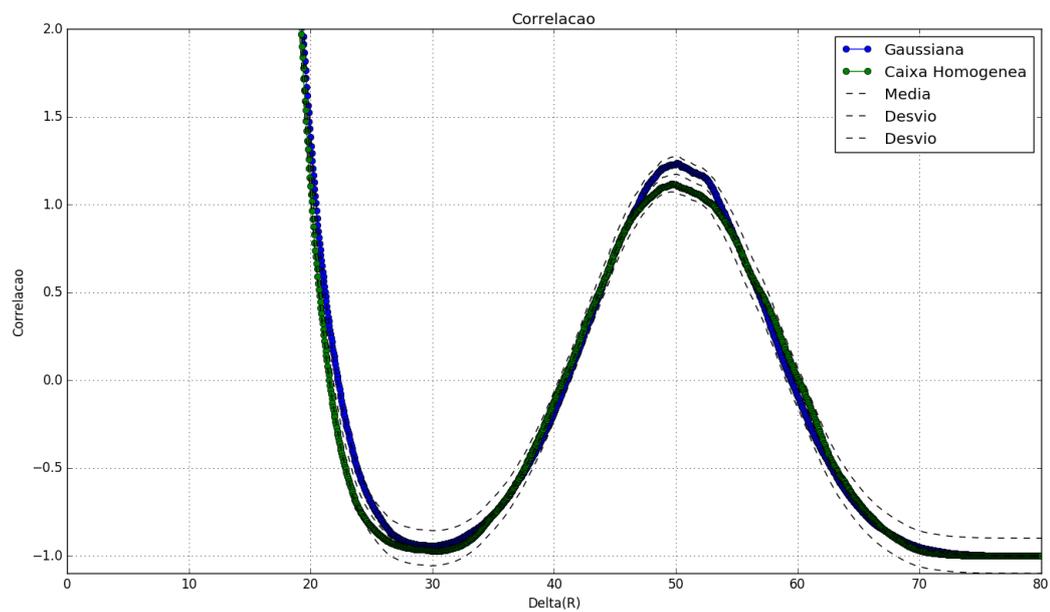


Figura 35 – Comparação Correlação Distribuições: Com desvios no terceiro pico

3.1.1.3 a primeira oscilação da FC2P é extremamente instável e varia muito mesmo para distribuições iguais.

Em regiões mais estáveis a análise é mais confiável. Podemos ver na Figura 34 que o comportamento das FC2P é similar na escala do segundo pico, mas a FC2P da distribuição beta apresenta oscilações. No entanto as outras duas amostras são extremamente similares. Podemos ver na Figura 35 duas curvas coloridas e uma curva pontilhada feita pela média da

FC2P gerada por dez distribuições com a caixa homogêna. As linhas tracejadas superiores e inferiores são calculadas a partir do desvio padrão da média de tais curvas. É possível ver que ambas as curvas estão dentro dos limites do desvio padrão da curva média tracejada. Embora as distribuições sejam visualmente diferentes, a FC2P não consegue diferenciar leis de distribuição diferentes. Este resultado não é novo, como pode ser visto na seção 8.2.7 de [Starck e Murtagh \(2005\)](#), onde distribuições foram geradas por processos totalmente diferentes e apresentaram FC2P muito similares.

Como visto nas Figuras [33](#) e [34](#) o teste com a distribuição beta foi o mais diferente dos três resultados mostrados. Este teste apresentou nove picos (Tabela [3](#)). O segundo pico de Beta apresentado na tabela, está bem saliente na Figura [33](#) e está em 11,675. Este pico é da ordem de escala do afastamento dos pontos da distribuição Beta. Ao criar o vazio no meio da distribuição, foram criadas mais distâncias da ordem de grandeza do vazio, o que aumentou a contagem na FC2P. As oscilações próximas ao terceiro pico das outras duas funções foram criadas pelo mesmo motivo, ao diminuir as distâncias no centro e aumentar nas bordas as regiões correspondentes à distância entre as nuvens e essas distâncias modificadas foram afetadas. É possível criar distribuições que apresentam FC2P similares, no entanto se as distribuições apresentarem variações de densidades de pontos grandes o suficiente (como o vazio criado na caixa de distribuição Beta) estas podem ser detectadas pela FC2P.

3.2.2 Testes Gaussianas: Sobreposição de Picos

Uma característica interessante de ser observada é até onde a FC2P consegue distinguir a separação das distribuições. Para tal análise foram geradas duas distribuições gaussianas com mil pontos cada, desvio padrão $\sigma = 5 \text{ u.m.}$, centradas em pontos diferentes (de maneira análoga às distribuições anteriores), mantendo o desvio padrão constante. As nuvens foram aproximadas e em cada passo de aproximação foi feito o cálculo da Função de Correlação, até o limite que não era mais possível identificar o pico referente ao afastamento entre as nuvens. O centro das distribuições têm coordenadas z e y em 50 u.m. e apenas a coordenada x varia.

3.2.2.1 Distribuições Gaussianas: Distância variável

Uma representação das distribuições é ilustrada abaixo para mostrar como as nuvens foram aproximadas. As Figuras 36, 37, 38 e 39 têm a distância entre os centros das nuvens reduzida em passo de 10 u.m.

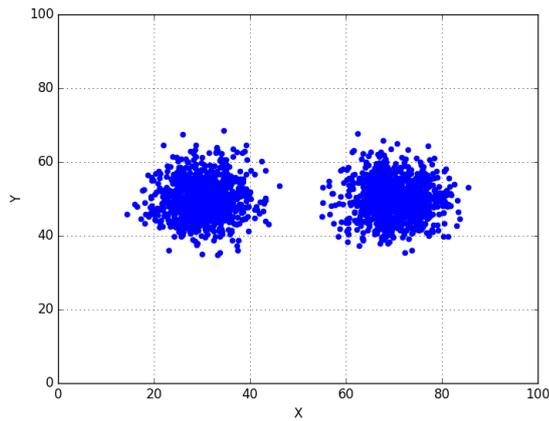


Figura 36 – Distribuições centradas em $x = 30 \text{ u.m.}$ e $x = 70 \text{ u.m.}$

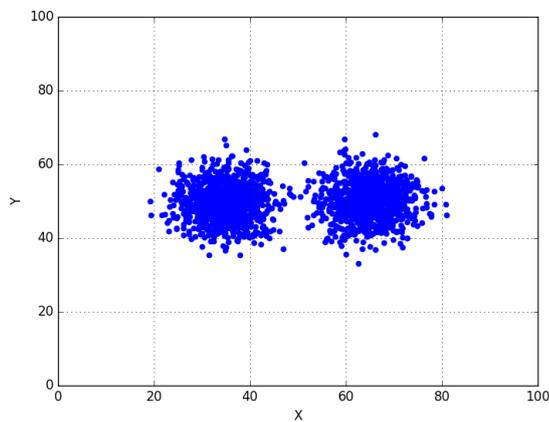


Figura 37 – Distribuições centradas em $x = 35 \text{ u.m.}$ e $x = 65 \text{ u.m.}$

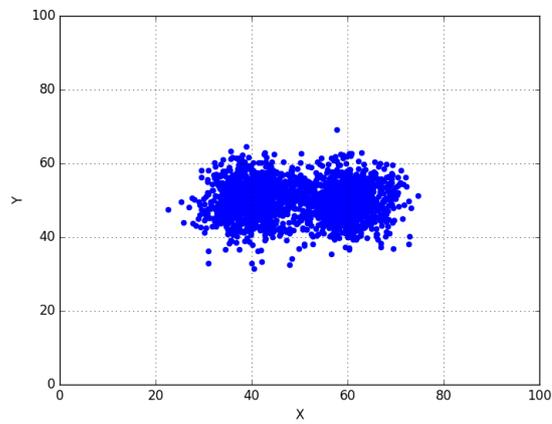


Figura 38 – Distribuições centradas em $x = 40$ u.m. e $x = 60$ u.m.

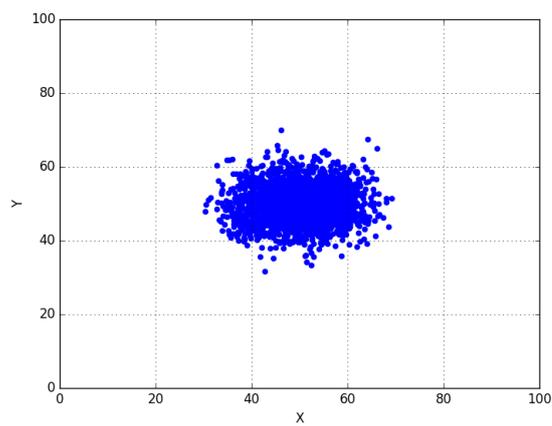


Figura 39 – Distribuições centradas em $x = 45$ u.m. e $x = 55$ u.m.

3.2.2.2 Histogramas

Os histogramas das distribuições também mostram os picos das distâncias características. No histograma da Figura 42 ainda é possível identificar o segundo pico referente ao afastamento das nuvens, já na Figura 43 os picos estão totalmente sobrepostos.

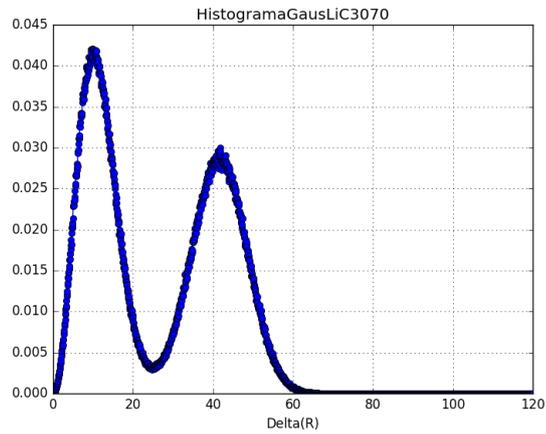


Figura 40 – Histograma de nuvens em $x = 30 \text{ u.m.}$ e $x = 70 \text{ u.m.}$

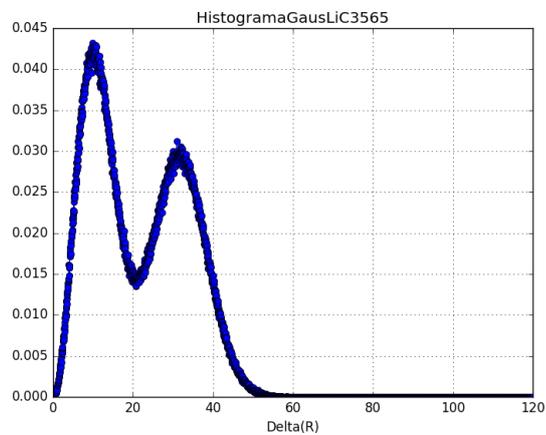


Figura 41 – Histograma de nuvens em $x = 35 \text{ u.m.}$ e $x = 65 \text{ u.m.}$

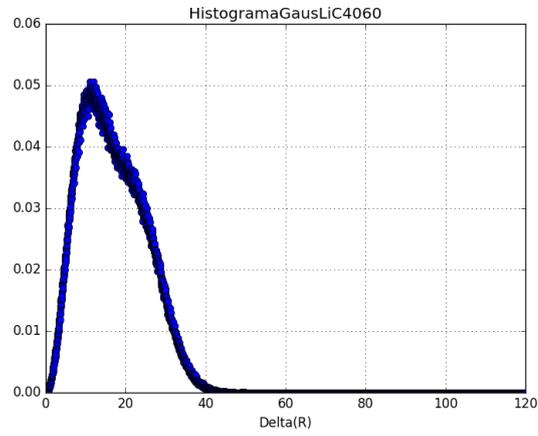


Figura 42 – Histograma de nuvens em $x = 40$ u.m. e $x = 60$ u.m.

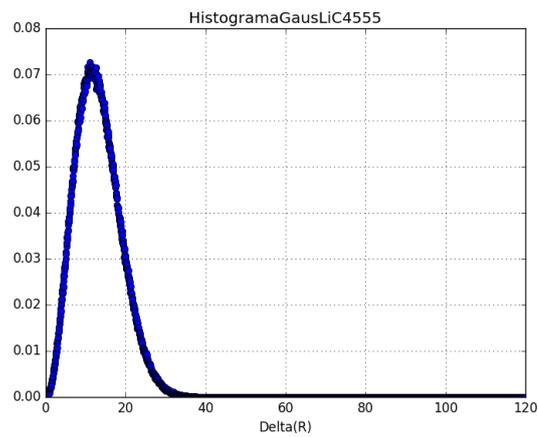
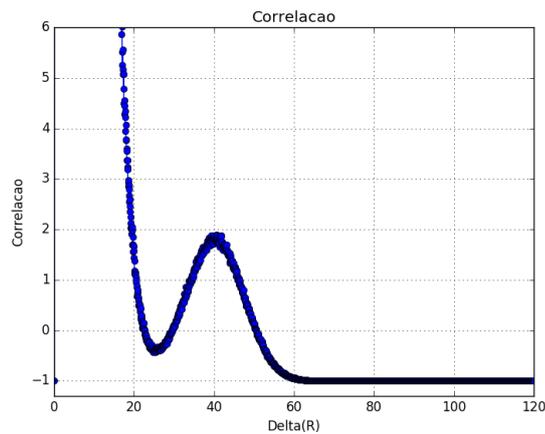
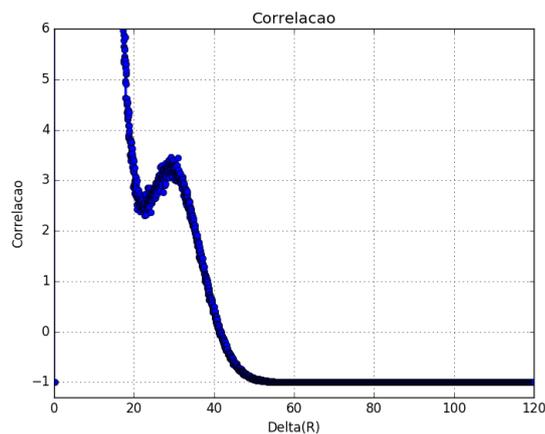


Figura 43 – Histograma de nuvens em $x = 45$ u.m. e $x = 55$ u.m.

3.2.2.3 Funções de Correlação

No caso da FC2P observamos que na Figura 46 a informação sobre o segundo pico já esta totalmente perdida, quando no histograma da Figura 42 ainda podemos observá-lo. Ou seja, a FC2P perde informação em relação ao histograma que a forma. O limite no qual a função de correlação não perde informação sobre picos consecutivos é equivalente a quatro desvios padrões da distribuição. No caso representado, as nuvens com um desvio padrão de $\sigma = 5 \text{ u.m.}$, tiveram os picos sobrepostos na FC2P quando a distância entre as nuvens chegou a 20 u.m. , quatro vezes o desvio padrão de cada nuvem.

Figura 44 – Correlação de nuvens em $x = 30 \text{ u.m.}$ e $x = 70 \text{ u.m.}$ Figura 45 – Correlação de nuvens em $x = 35 \text{ u.m.}$ e $x = 65 \text{ u.m.}$

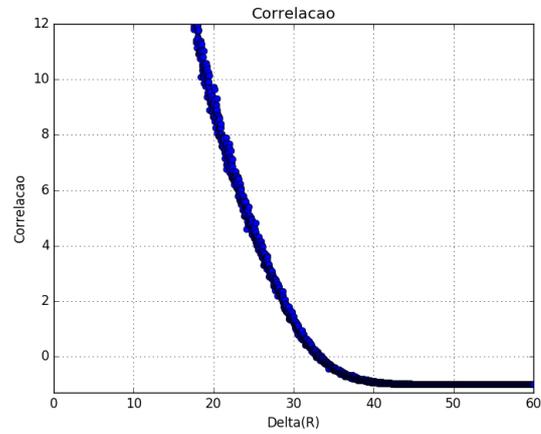


Figura 46 – Histograma de nuvens em $x = 40 \text{ u.m.}$ e $x = 60 \text{ u.m.}$

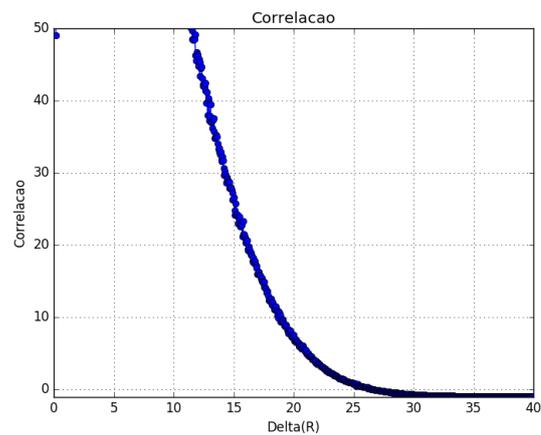


Figura 47 – Correlação de nuvens em em $x = 45 \text{ u.m.}$ e $x = 55 \text{ u.m.}$

3.2.3 Teste de Distribuições Gaussianas

Para melhor entender como a FC2P varia foi feito um teste onde a densidade de nuvens aumenta. As nuvens se aproximam e o número de nuvens é aumentado em cada etapa. Cada nuvem tem quinhentos pontos, o desvio padrão das nuvens é constante $\sigma = 5$ *u.m.*

3.2.3.1 8 nuvens, Histograma e Correlação

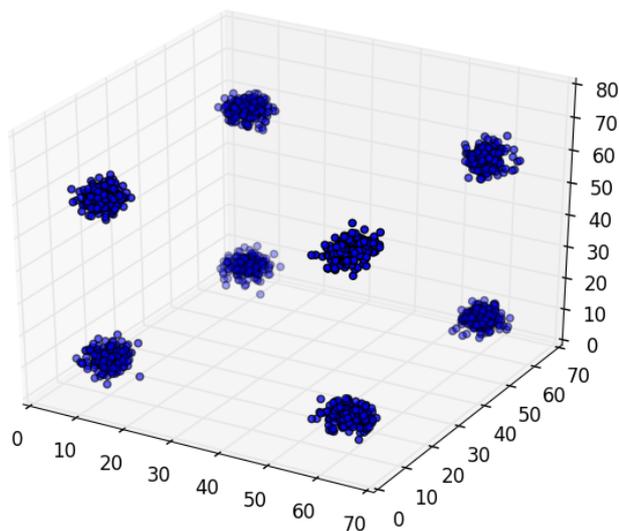


Figura 48

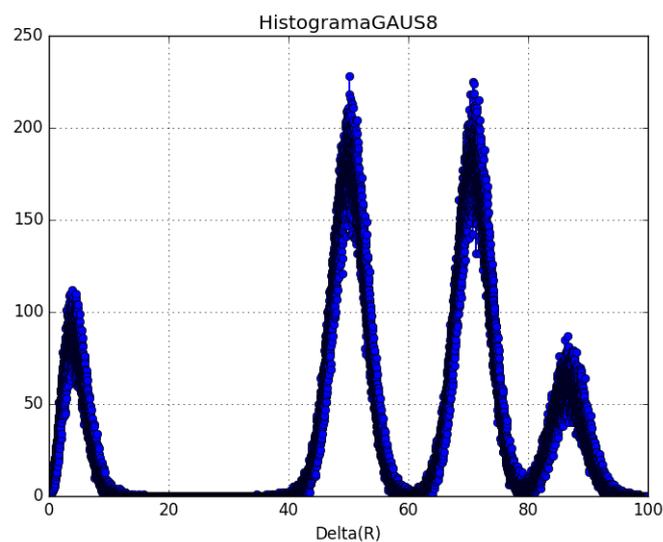


Figura 49 – Histograma de 8 Nuvens. Picos (*u.m.*): P1 4,78; P2 50,04; P3 71,74; P4 87,67

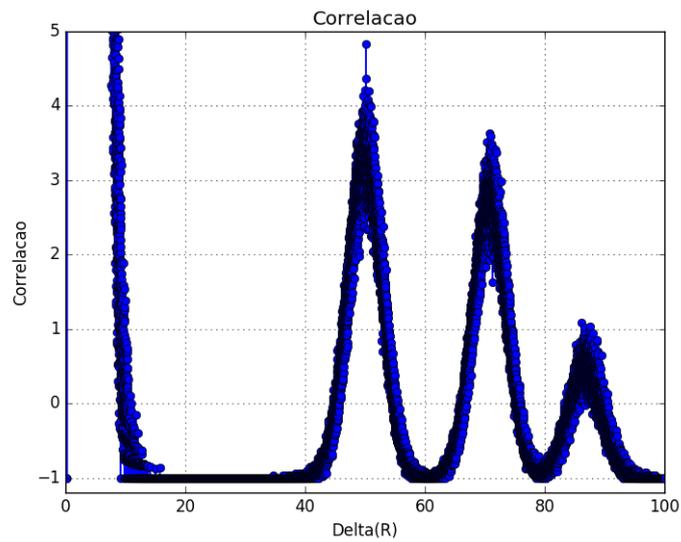


Figura 50 – FC2P de 8 nuvens. Picos (*u.m.*): P1 0,25; P2 9,15; P3 50,06; P4 71,74; P5 87,67;

3.2.3.2 27 nuvens, Histograma e Correlação

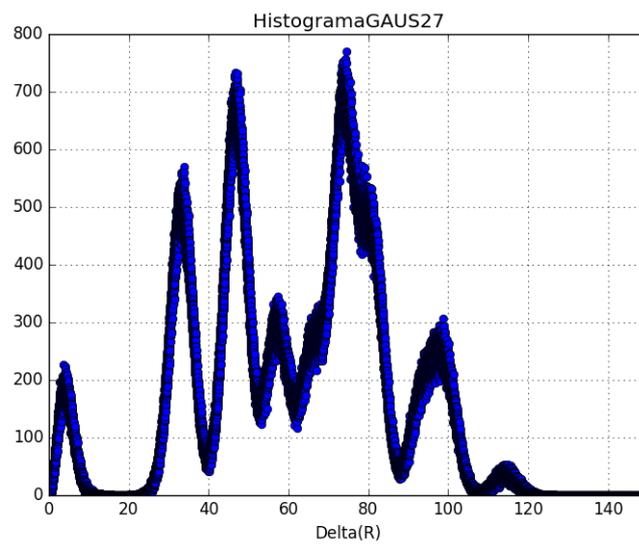
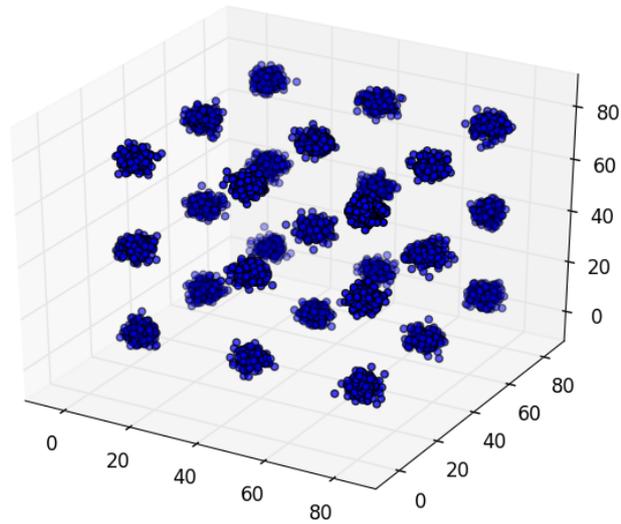


Figura 51 – Histograma de 27 Nuvens. Picos (*u.m.*): P1 4,2; P2 32,13; P3 45,37; P4 56,54; P5 75,67; P6 79,77; P7 98,69; P8 99,62; P9 112,46

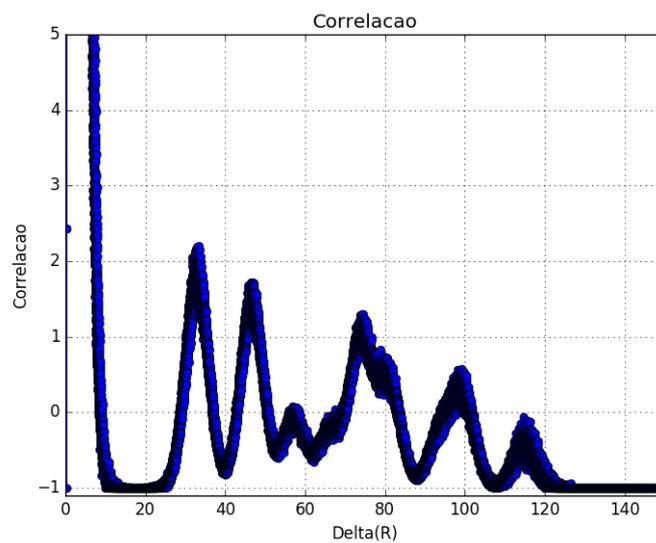


Figura 52 – FC2P de 27 Nuvens. Picos (*u.m.*): P1 1,44; P2 32,14; P3 47,27; P4: 56,54; P5 72,25; P6 75,67; P7 79,86; P8 99,62; P9 117,32

3.2.3.3 64 nuvens, Histograma e Correlação

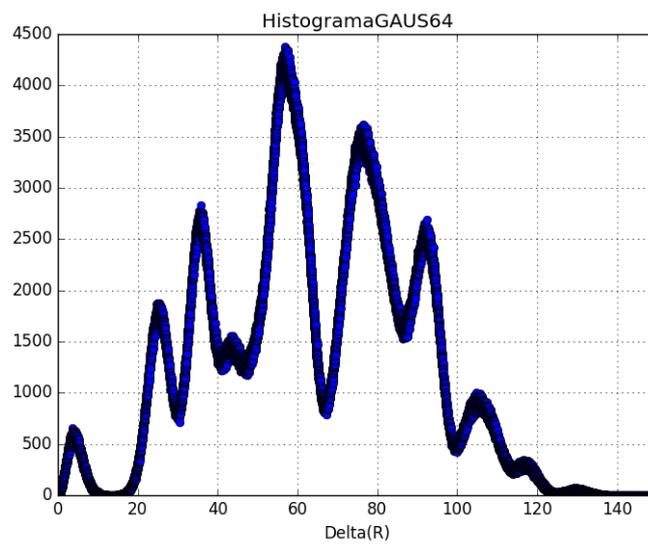
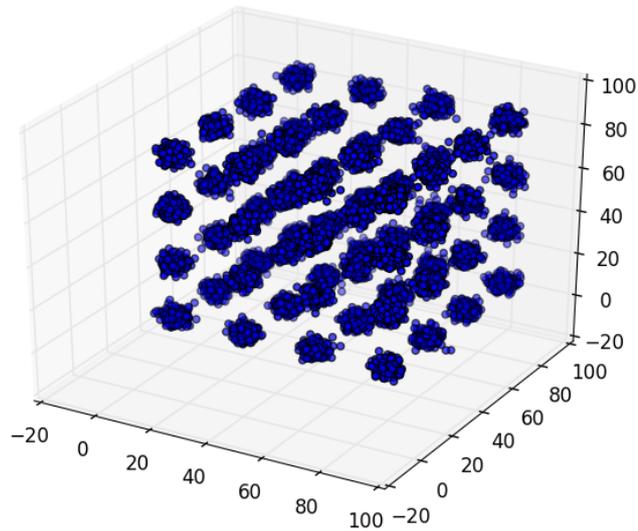


Figura 53 – Histograma de 64 Nuvens. Picos (*u.m.*): P1 4,51; P2 26,23; P3 36,13; P4 42,42; P5 57,05; P6 75,82; P7 91,76 P8 106,82 P9 117,53 P10: 130,59

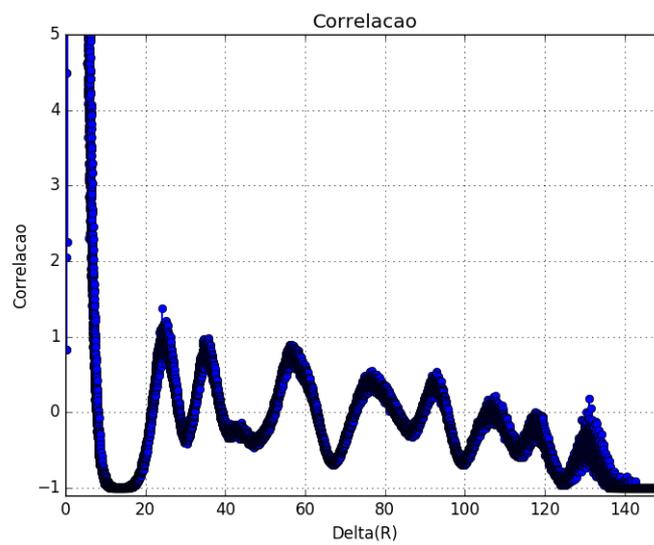


Figura 54 – FC2P de 64 Nuvens. Picos (*u.m.*): P1 1,67; P2 25,52; P3 33,29; P4 56,59; P5 75,04; P6 85,84; P7 92,62; P8 106,64; P9 117,53; P10 130,99

3.2.3.4 Comparação das nuvens

As Figuras 55, 56 e 57 mostram uma comparação entre o histograma das distribuições de 8, 27 e 64 nuvens, com um histograma gerado por uma distribuição homogênea gerado por um Processo Pontual de Poisson Homogêneo. É possível perceber que ao aumentar o número de nuvens e diminuir as distâncias entre as nuvens os picos começam a se aproximar e diminuir a amplitude. Nas figuras abaixo temos os histogramas renormalizados e comparados com a distribuição normal homogênea gerada pelo processo pontual de poisson. Na Figura 58 podemos ver que ao aumentar o número de nuvens e diminuir as distâncias entre elas, as distâncias começam a se distribuir de maneira mais homogênea, diminuindo a amplitude dos picos e os aproximando. Os histogramas das amostras começam a tender para o histograma homogêneo. Na Figura 59 podemos ver que a FC2P também começa a tender a zero quando é aumentada a densidade de nuvens. No limite em que a distância entre os picos é da ordem de quatro desvios padrão das nuvens, a FC2P se degenera totalmente, pois todos os picos ficam sobrepostos. Este resultado já é esperado, pois ao gerar diversas distribuições aleatórias e deixá-las próximas temos um processo muito similar ao Processo Pontual de Poisson Homogêneo explicado na seção 2.2.2.

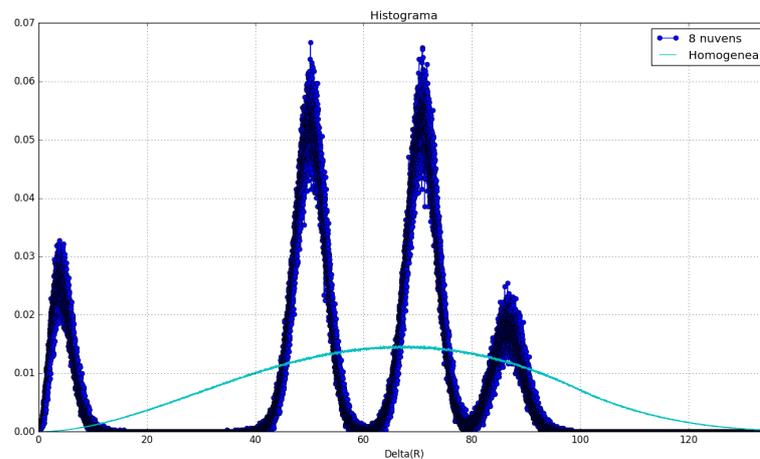


Figura 55 – Histograma renormalizado e comparado com histograma homogêneo: 8 nuvens

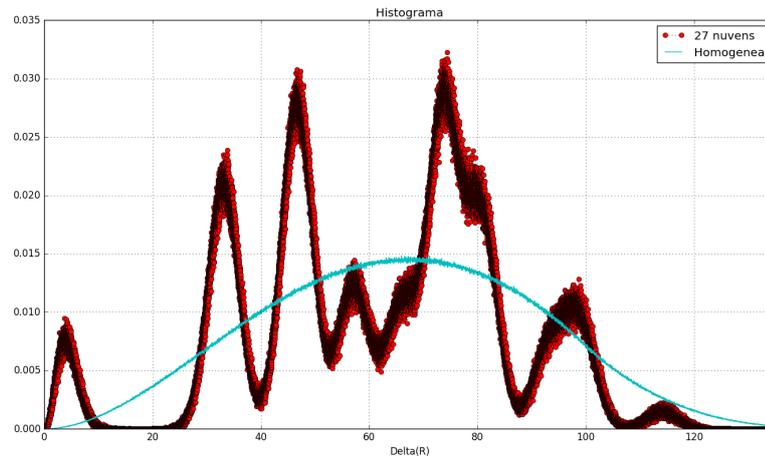


Figura 56 – Histograma renormalizado e comparado com histograma homogêneo: 27 nuvens

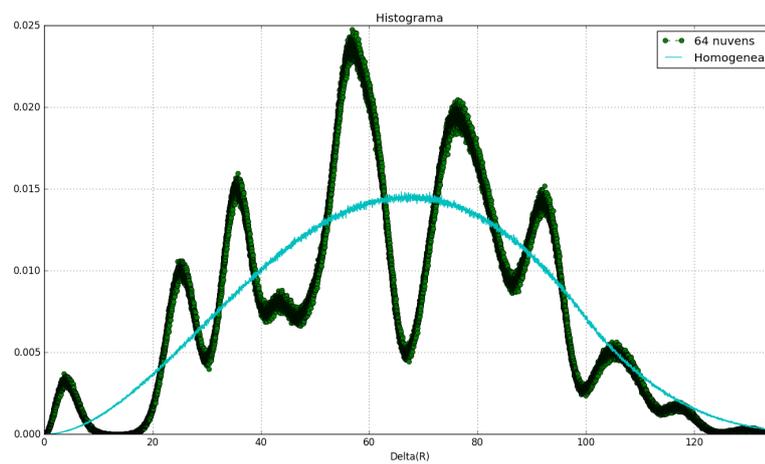


Figura 57 – Histograma renormalizado e comparado com histograma homogêneo: 64 nuvens

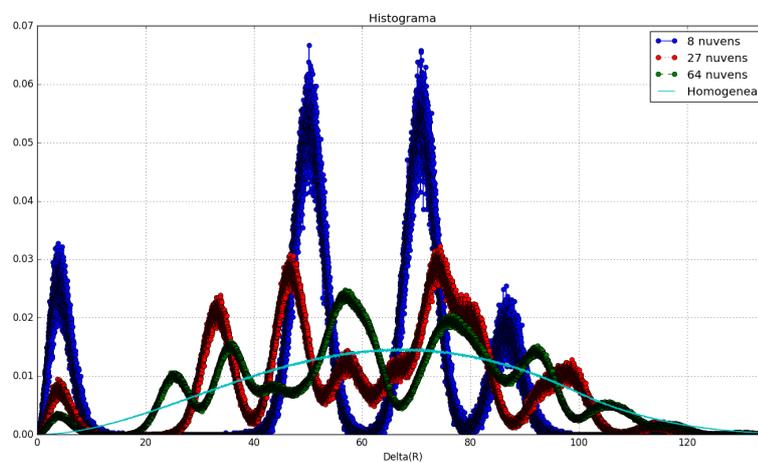


Figura 58 – Histogramas renormalizados

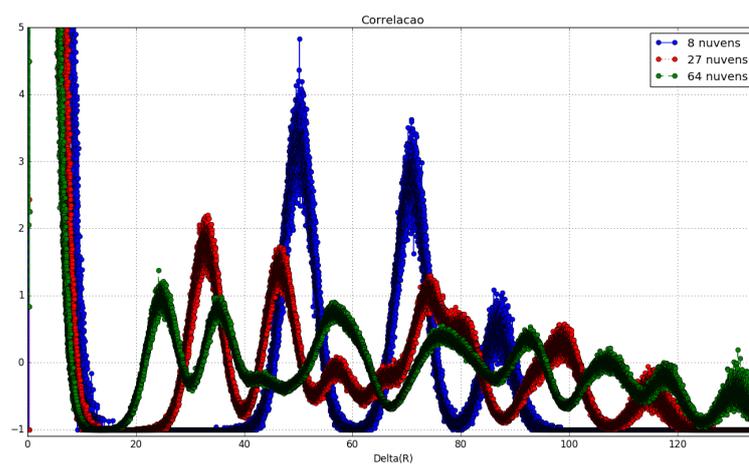


Figura 59 – Composição de FC2P de 8 nuvens, 27 nuvens e 64 nuvens

4 Conclusão

A Função de Correlação aplicada em sua forma espacial é uma ferramenta estatística que começou a ser utilizada na metade do século XX para tentar determinar de forma mais acurada o grau de homogeneidade do universo ao comparar distâncias de distribuições com distâncias de uma distribuição homogênea e, assim, verificar se e em que escala o Princípio Cosmológico poderia ser aplicado. Tal ferramenta vem sendo utilizada até hoje e tem uma influência muito grande no estudo de distribuições em grandes escalas, por isso um estudo de suas limitações e capacidades é importante.

Os primeiros estudos com picos mostraram que a Função de Correlação de Dois Pontos (FC2P) consegue determinar escalas características de aglomeração através da posição dos picos. Os testes com distribuições mostraram que, utilizando apenas as distâncias relativas, a FC2P independe da posição das distribuições. Testes com diferentes nuvens mostraram que a FC2P é degenerada, pois não consegue diferenciar certas distribuições. Mas caso haja uma anomalia maior na distribuição de pontos, como o vazio gerada no centro da caixa de distribuição Beta, é possível identificar variações, que não determinam o tipo de distribuição analisado, sempre tornando possível que se crie uma distribuição de pontos que gere a mesma FC2P. Tal estudo mostra que a função de correlação é degenerada, pois diferentes distribuições podem produzir estatisticamente os mesmos resultados, o que está de acordo com a literatura, como pode ser visto na seção 8.2.7 de [Starck e Murtagh \(2005\)](#).

Analisando apenas um tipo de distribuição (no caso a gaussiana) foi possível estudar características da função que independem da distribuição escolhida. Como mencionado anteriormente, a FC2P é degenerada, então características que não dependem do tipo de distribuição devem ser mais explorado para que a ferramenta forneça resultados relevantes. Os testes com Gaussianas nos mostraram que há uma sobreposição de picos na FC2P quando a escala das distâncias é da ordem de quatro vezes o desvio padrão, σ da distribuição. No caso mostrado, quando as nuvens chegaram a uma distância próxima equivalente a quatro vezes o desvio padrão da distribuição de pontos de cada nuvem, foi possível detectar apenas um pico na função de correlação. Este resultado é importante pois, no histograma de distribuição de distâncias das amostras é possível ainda identificar os diferentes picos que representavam as duas distribuições, enquanto a função de correlação identificava apenas um pico que poderia ser referente a uma mesma distribuição com um desvio padrão maior. Ou seja, a função de correlação perde informações em relação ao histograma. Ao variar a densidade de nuvens é possível observar como os histogramas e, também, as FC2Ps se aproximam do esperado por uma distribuição homogênea quando as nuvens começam a ficar mais numerosas e próximas. Tal resultado é esperado, pois para um número muito

grande de nuvens teríamos um processo análogo ao que gera a distribuição homogênea, o Processo Pontual de Poisson Homogêneo.

Tais estudos nos mostram que a FC2P não falha ao efetuar a função para qual foi originalmente implementada: encontrar a escala para qual uma distribuição pode ser considerada homogênea. Mas ao tentar detectar estruturas mais finas e de maneira mais precisa tem suas limitações, embora ainda consiga identificar escalas de distâncias características das distribuições - se respeitados os limites nos quais os picos se sobrepõem - podendo servir para parametrizar distâncias características em métodos não paramétricos mais precisos.

Referências

Abell, G. O. The Distribution of Rich Clusters of Galaxies. *apjs*, v. 3, p. 211, maio 1958. Citado na página 9.

Bruno, G.; Galilei, G.; Lobo A., i. . . p. . A. y. . . s. . O. e. . . v. . . *Sobre o infinito, o universo e os mundos / Giordano Bruno ; [tradução de Helda Barraco e Nestor Deola]. O ensaiador / Galileu Galilei ; [tradução e notas de Helda Barraco]. A cidade do sol / Tommaso Campanella ; [tradução e notas de Aristides Lobo].* [S.l.: s.n.]. Citado na página 4.

de Vaucouleurs, G. Evidence for a local supergalaxy. *aj*, v. 58, p. 30, fev. 1953. Citado na página 7.

Einasto, J. et al. The supercluster-void network - III. The correlation function as a geometrical statistic. *mnras*, v. 289, p. 813–823, ago. 1997. Citado 4 vezes nas páginas 10, 25, 32 e 40.

Einasto, J.; Saar, E.; Klypin, A. A. Structure of superclusters and supercluster formation. V - Spatial correlation and voids. *mnras*, v. 219, p. 457–478, abr. 1986. Citado 2 vezes nas páginas 10 e 24.

Gabrielli, A. et al. *Statistical Physics for Cosmic Structures*. [S.l.]: Springer, 2005. ISBN 3-540-40745-6. Citado 2 vezes nas páginas 10 e 25.

Groth, E. J.; Peebles, P. J. E. Statistical analysis of catalogs of extragalactic objects. VII - Two- and three-point correlation functions for the high-resolution Shane-Wirtanen catalog of galaxies. *apj*, v. 217, p. 385–405, out. 1977. Citado 3 vezes nas páginas 11, 13 e 10.

Hamilton, A. J. S. Toward Better Ways to Measure the Galaxy Correlation Function. *apj*, v. 417, p. 19, nov. 1993. Citado na página 24.

Herschel, W. On the construction of the heavens. by william herschel, esq. f. r. s. *Philosophical Transactions of the Royal Society of London*, The Royal Society, v. 75, p. 213–266, 1785. ISSN 02610523. Disponível em: <<http://www.jstor.org/stable/106755>>. Citado 2 vezes nas páginas 4 e 6.

Hubble, E. A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Science*, v. 15, p. 168–173, mar. 1929. Citado 2 vezes nas páginas 6 e 8.

Hubble, E. Effects of Red Shifts on the Distribution of Nebulae. *apj*, v. 84, p. 517, dez. 1936. Citado na página 7.

Hubble, E.; Humason, M. L. The Velocity-Distance Relation among Extra-Galactic Nebulae. *apj*, v. 74, p. 43, jul. 1931. Citado na página 6.

Hubble, E. P. Extragalactic nebulae. *apj*, v. 64, dez. 1926. Citado 2 vezes nas páginas 5 e 7.

- Kepler, S. O.; Saraiva, M. *Astronomia e Astrofísica*. [S.l.]: Departamento de Astronomia - Instituto de Física, Universidade Federal do Rio Grande do Sul, 2011. Citado 2 vezes nas páginas 5 e 7.
- Martínez, V. J. *Statistical of the Galaxy Distribution*. [S.l.]: CHAPMAN HALL/CRC, 2001. Citado na página 23.
- Newton, I.; et, al. *Philosophiae Naturalis Principia Mathematica*. 1. ed. [S.l.: s.n.], 1686. Citado na página 8.
- Ryden, B. *Introduction to Cosmology*. [S.l.]: Addison Wesley, The Ohio State University, 2003. Citado 2 vezes nas páginas 6 e 8.
- Shane, C. D.; Wirtanen, C. A. The distribution of extragalactic nebulae. *aj*, v. 59, p. 285–304, set. 1954. Citado na página 8.
- Shapley, H. Globular Clusters and the Structure of the Galactic System. *pasj*, v. 30, p. 42, fev. 1918. Citado 3 vezes nas páginas 5, 6 e 7.
- Souza, R. E. *Introdução à Cosmologia*. [S.l.]: Universidade de São Paulo, 2004. ISBN 85-314-0843-1. Citado na página 5.
- Starck, J.; Murtagh, F. *Astronomical Image and Data Analysis*. Second. [S.l.]: Springer, 2005. Citado 2 vezes nas páginas 42 e 59.
- Stoughton, C.; et, al. Sloan Digital Sky Survey: Early Data Release. , v. 123, p. 485–548, jan. 2002. Citado na página 9.
- Totsuji, H.; Kihara, T. The Correlation Function for the Distribution of Galaxies. *pasj*, v. 21, p. 221, 1969. Citado na página 10.