T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

EPİDEMİYOLOJİDE NOKTA SÜREÇLERİ UYGULAMASI

ISSA RAMADANE

YÜKSEK LİSANS TEZİ
MATEMATİK MÜHENDİSLİĞİ ANABİLİM DALI
MATEMATİK MÜHENDİSLİĞİ PROGRAMI

DANIŞMAN
DOÇ. DR. REŞAT KÖŞKER

İSTANBUL, 2016

T.C.

YILDIZ TEKNİK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

EPİDEMİYOLOJİDE NOKTA SÜREÇLERİ UYGULAMASI

Issa RAMADANE tarafından hazırlanan tez çalışması 08.04.2016tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Matematik Mühendisliği Anabilim Dalı'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Tez Danışmanı

Doç. Dr. Reşat KÖŞKER

Yıldız Teknik Üniversitesi

Jüri Üyeleri

Doç. Dr. Reşat KÖŞKER

Yıldız Teknik Üniversitesi _____

Doç. Dr. Ülkü BABUŞÇU YEŞİL

Yıldız Teknik Üniversitesi _____

Doç. Dr. Ahmet KIRIŞ

İstanbul Teknik Üniversitesi _____

ÖNSÖZ

Bu çalışmada, Orta Afrika Cumhuriyeti'ndeki sıtma hastalığı ile ilgili ulaşabildiğimiz verileri kullanarak, bu hastalığın söz konusu coğrafyada yayılımının tahminini yapmaya çalıştık. Bunun için çeşitli nokta süreçleri analizlerini kullandık ve bunlarla elde ettiğimiz sonuçları karşılaştırdık. Ulaşabildiğimiz verilerin işlenmesinden elde edilen sonuçlar, hastalığın incelenen coğrafyada bir süre daha sorun olmaya devam edeceği, dolayısıyla otoritelerin halk sağlığı ile ilgili bu yaklaşıma dayalı politikalar üretmeleri gerektiğini ortaya çıkarmıştır.

Bu tez boyunca, tavsiye ve rehberliği devam eden tez danışmanım Doç. Dr Reşat KÖŞKER'e içten şükranlarımı sunarım. Ayrıca, Yıldız Teknik Üniversitesi, Matematik Mühendisliği Bölümü öğretim üyelerine de yüksek lisans boyunca desteklerinden dolayı teşekkür ederim. Burada geçirdiğim zaman boyunca beni kendi vatanımda gibi hissettiren Türkiye Cumhuriyeti Yurtdışı Türkler ve Akraba toplulukları Başkanlığı (YTB)'na her türlü yardım ve desteklerinden dolayı şükran duyduğumu ifade etmek isterim. Orta Afrika İstatistik Enstitüsü, Sosyal ve Ekonomik Çalışmalar (CAISSES)'ne bu çalışmada kullanılan veriyi sağladıkları için minnettarım.

Sınıf arkadaşlarımın araştırma boyunca yardımları ve cesaretlendirmeleri teşvik edici olmuştur. Özellikle arkadaşım Yunus ATALAN'a katkı ve yardımlarından dolayı teşekkürü bir borç bilirim.

Özellikle, tüm eğitim hayatım boyunca gösterdikleri sabır, hem maddi hem de manevi destek için babam Ramadane MISKINE, annem MEREM Delmas ve kardeşim AMANGABAKO Mahamoud Delmas'a teşekkür ederim.

Ayrıca, Bayan BABOUR Reine'ye bana verdiği moral ve Türkiye'deki akademik yolculuğumdaki manevi destek için minnettarım.

Amcalarım OUMAR Miskine Simon, Dr. AGUID Soumouck, SELEMANE Sylvain ve ailelerini, finansal ve moral destekleri için burada anmadan geçemeyeceğim.

Mays, 2016

Issa RAMADANE

İÇİNDEKİLER

Sayfa

SİMGE LİSTESİ....viii

KISALTMA LİSTESİ....ix

ŞEKİL LİSTESİ.... x

ÇİZELGE LİSTESİ....xi

ÖZET....xii

ABSTRACT....xiv

BÖLÜM 1

GİRİŞ.... 1

1.1 Literatür İncelemesi.... 1
1.2 Tezin Amacı.... 5
1.3 Hipotez.... 7
1.4 Tezin Organizasyonu.... 8

BÖLÜM 2

NOKTA SÜREÇLERİ TEORİSİNİN TEMEL NOTASYONLARI.... 10

2. Giriş.... 10

2.1 Nokta Süreçleri.... 10
2.2 Zamansal nokta süreci.... 12
2.3 Yoğünlük Fonksiyonu.... 15
2.4 Bir Nokta sürecin En Çok Olabilirlik Süreci.... 16
2.5 Nokta Süreci için Rezidü Analizi.... 18
2.6 Dispersiyon ve Otokorelasyon Analizi.... 19
2.7 Tamamen Rastlantısal Dağılım.... 21
2.8 Temel Nokta süreçleri.... 22
2.8.1 Bernoulli ve Binomial süreçleri.... 23
2.8.2 Poisson Süreci.... 23

2.8.3 Yenileme Süreci 25
2.8.4 Cox Süreçleri 26
2.8.5 Hawkes Süreçleri 27

BÖLÜM 3

POSTERİOR DAĞILIM ÖRNEKLEMİ 30

3. Giriş 30

3.1 Stokastik Simülasyon algoritmaları 31
3.1.1 Kabul-Red Yöntemi 32
3.2 Nokta Süreçleri ve Markov zinciri Monte Carlo (MCMC) 34
3.2.1 Markov zinciri 35
3.2.2 MCMC Metodunun Verimliliği ve Yakınsaklık Hızı 36
3.3 Metropolis-Hastings Algoritmaları 38
3.3.1 Bağımsız Metropolis-Hasting Algoritması 39
3.3.2 Rasgele yürüyen Metropolis-Hasting Algoritması 40
3.3.3 Her seferinde bir değişken Metropolis-Hastings algoritması 41
3.4 Gibbs Örnekleme Algoritması 41
3.5 Hibrid Gibbs Örnekleme 43
3.6 Tersinir atlama Markov Zinciri Monte Carlo (RJMCMC) 43
3.7 Sonuç 44

BÖLÜM 4

ZAMAN-MEKANSAL NOKTA SÜREÇLERİ 46

4. Giriş 46

4.1 Zaman-Mekansal İlişkili Nokta süreçleri 47
4.2 Dağılım Özellikleri 48
4.3 Zaman-Mekansal Nokta Süreçleri için Bayesian Yaklaşım 49
4.4 İncelenen Süreçlerin İstatistiği 50
4.4.1 Zaman-Mekansal Cox Süreci 51
4.4.2 Dirichlet süreci 55
4.4.2.1 Posteriör Dağılımı 57
4.4.2.2 Bayesyen Çerçevesinde Karışık Modeller 57

BÖLÜM 5

EPİDEMİYOLOJİ İÇİN NOKTA SÜREÇLERİ 60

5. Giriş 60

5.1 Epidemiyoloji 61
5.1.1 Mekansal Epidemiyoloji 63
5.1.2 Olayların Zaman-mekansal Dağılımı 65
5.1.3 Riskin Mekansal Değişimi 66
5.2 Uygulama: Orta Afrika Cumhuriyetinde Sıtma Hastalığı Dağılımı Çalışması 68

5.3 Analiz ve Tartışma 80
BÖLÜM 6
SONUÇ VE ÖNERİLER 82
KAYNAKLAR 83
ÖZGEÇMİŞ 94

SİMGE LİSTESİ

$\mathbb{N}$	Doğal sayılar kümesi
$\mathbb{R}$	Reel saylar kümesi
N	Nokta süreçleri
F	filtrasyon
$\mathcal{B}$	Borel σ-cebri
δ	Dirac ölçümü
λ	Nokta süreçlerin yoğunluğu
T	Zamansal
s	Mekânsal
θ	Yoğun parametresi
λ_θ	Yoğunluk süreci
λ_L	Lebesgue ölçümü
$\mathcal{L}_T$	Log-olabilirlik
I_m	Moran indeksi
$\Lambda(t)$	Poisson yoğunluğu
(χ, d)	Metrik uzayı
$S \times T$	Mekân-zamansal
$(\chi, \mathcal{B}_d, \nu)$	Ölçüm uzayı
$(\Omega, \mathcal{B}, \mathbb{P})$	Olasılık uzayı

KISALTMA LİSTESİ

MCMC	Markov Chain Monte Carlo
RGPH	Recensement général de la population et d'habitat (genel nüfus ve ikamet edenler)
CAR	Central African Republic (Orta Afrika Cumhuriyeti)
GIS	geographic information system
CRD	completely random distribution
CSR	complete spatial randomness
PPP	Poisson point process
HPPP	homogeneous Poisson point process
HTPP	homogeneous temporal Poisson process
RJMCMC	Reversible jump Markov Chain Monte Carlo
SSA	Stochastic simulation algorithms
CDF	cumulative distribution function
DPM	Dirichlet process mixture
SMR	standard morbidity ratio
WHO	World Health Organization
CAISSES	Central African Institute of Statistics, Social and Economic Studies

ŞEKİL LİSTESİ

Sayfa

Şekil 1. 1 Bir nokta süreci olarak 1974 den 1983 e İngiltere de lancashire deki vaka ile bağlantılı gırtlak kanseri vakası haritası (DIGGLE 1990)[40].....4
Şekil 1. 2 Mekansal bir sürecin gözlemlerine bir örnek: The Piazza Road datası [Higdon et al 1999.] her bir nokta Piazza yolunun etrafında yapılan bir dioksin konsantrasyonu ölçümünün lokasyonu göstermektedir [40].....5
Şekil 2. 1 Bir sayma sürecinin mümkün gerçekleşmesi ve karşılık gelen örnek Fonksiyonu..... 14
Şekil 2. 2 Geçici nokta sürecin bekleyiş zamanlarının grafik gösterimi(44) 15
Şekil 2. 3 Farklı hipotezler altında Moran indeksi Im değerler. 20
Şekil 2. 4 Dallanma sürecinin modellenmesi 29
Şekil 3. 1 Kabul-red açıklaması..... 33
Şekil 4. 1 Hiyerarşik modelin döngüsel grafiği 59
Şekil 5. 1 Kümeleme metoduyla oluşan dağılımların nokta süreçleri örnekleri (a) rastgele dağılım, (b) kümelenmiş dağılım (c) düzgün dağılım.....65
Şekil 5. 2 Orta Afrika Cumhuriyeti ve komşuların haritası..... 68
Şekil 5. 3 Orta Afrika'da sıtma ölümünün demografik yoğunluğu..... 69
Şekil 5. 4 Nokta süreçleri dağılımı grafikleri, (a) Orta Afrika Cumhuriyetindeki 12 yıllık, sıtma hastalığının rastgele zaman-mekansal dağılımı (b) Poisson nokta süreci dağılımı.....73
Şekil 5. 5 Sıtma hastalığının, farklı nokta süreçleri ile (a) rastgele mekansal dağılımının, (b) Poisson nokta süreci, (c) Beta dağılımı; (d) teta posteror dağılımı; (e) Binom ve Poisson dağılımının yoğunluk fonksiyonları ile, (f) Normal dağılımı.....77
Şekil 5. 6 Dirichlet süreciyle Gibbs-Metropolis-Hastings hybrid işlemleriyle elde edilmiş MCMC sonuçları.....77
Şekil 5. 7 (a) sıtma insidansı grafiği, (b) ölüm insidansı grafiği 79

ÇİZELGE LİSTESİ

Sayfa

Çizelge 5. 1 Orta Afrika Cumhuriyetine ait sıtma verisinin özeti.....70
Çizelge 5. 2 2011 ve 2012 yıllarına ait olayların dağılımı, bu çalışma başkent Bangui yaıl yayılmış.....77
Çizelge 5. 3 Hastalığın insidansı (hastalıktan her yıl etkilenen yeni hasta sayısı).....78
Çizelge 5. 4 İstatistiksel bilgileri.....80

ÖZET

EPİDEMİYOLOJİDE NOKTA SÜREÇLERİ UYGULAMASI

Issa RAMADANE

Matematik Mühendisliği Anabilim Dalı

Yüksek Lisans Tezi

Tez Danışmanı: Doç. Dr. Reşat KÖŞKER

Nokta süreçleri mekânsal ve zaman-mekânsal nokta desenlerini ve çeşitli olayların dağılımını tarif eden araçlar olarak sıklıkla kullanılır. Bu tezde, mekân-zamansal nokta süreçleri ile ilgilendik ve Dirichlet süreci ve çeşitli uzantıları gibi yeni araçlar dahil olmak üzere, çeşitli zaman-mekânsal nokta süreçlerinin Bayesian istatiksel analiziner genel bir bakış verdik. Bu modeller mekânsal epidemiyolojik dağılımlarda uygulanmıştır. Burada, son yıllarda (2002-2013) Orta Afrika Cumhuriyeti'nde sıtma hastalığının yayılması durumları üzerinde yoğunlaştık. Bu tezde altı bölüm vardır. İlk iki bölüm, bu modellerle ilgili bilgileri toparlamak amacıyla teorik nokta süreçlerine ayrılmıştır. Burada ele alınan nokta süreçleri, olayların dağılımının analizinde önemli rol oynayan Binomial süreci, Poisson süreci, Cox süreci ve Hawkes süreci gibi temel nokta süreçleridir.

Üçüncü ve dördüncü bölümlerde, Bayesian istatistik tartışılmıştır. Bu bölümlerde, Shot Noise yoğunluk ölçümünün genelleşmiş hali olan zaman-mekânsal Cox sürecinin olaylarına yol açan odak bölgesindeki olayların oluşumunu ele aldık. Gizli Poisson süreci, Dirichlet sürecine göre dağılan ölçüm fonksiyonunun yoğunluğuna katkıda bulunur. Çeşitli dönemlerdeki gözlemler boyunca oluşan olayların mekânsal konumunu dikkate alan veriden dolayı, simülasyon modelleri için tanıtılan MCMC ve parametrik olmayan yöntemlerle istatistiksel çıkarımların genel görünüşünü verdik.

Bölüm 5'de, zaman-mekânsal epidemiyolojik çerçevede epidemiyoloji için özel nokta süreçleri üzerinde durulmuştur. Orta Afrika Cumhuriyeti'nde sıtma hastalığının coğrafi yayılımı için bu süreçleri uyguladık. Her bir eleman için olayların en fazla biri ile ilgili kesikli değişkeni olan nokta süreçlerini analiz ederiz. Önceki bölümlerde tanıtılan

modeller test edilerek Orta Afrika Cumhuriyeti'ndeki sıtma hastalığının risk karakterizasyonu belirlenir.

Altıncı bölümde, elde dilen sonuçlar tartışılmıştır ve yorumlanmıştır.

Anahtar Kelimeler: Nokta süreçleri, Poisson süreci, Dirichlet süreci, örneklem, zaman-mekânsal, MCMC yöntemleri, epidemiyoloji, sıtma hastalığı

ABSTRACT

THE APPLICATION OF POINT PROCESSES IN EPIDEMIOLOGY

Issa RAMADANE

Department of Matematical Engineering

MSc. Thesis

Adviser: Assoc. Dr. Reşat KÖŞKER

The point processes are often used as tools for describing spatial or spatiotemporal point patterns. They are also used to describe various phenomena concerning distribution of events. In this thesis, we are interested in spatiotemporal point processes and aim to give an overview of Bayesian statistical analysis for point processes as well as for recent tools like the Dirichlet process and its diverse extensions. These models are applied in spatial epidemiological distributions. We focus on situations where the malaria disease has spread in Central African Republic in recent years (2002-2013). The thesis contain of six chapters. The first two chapters are devoted to theoretical point processes, thus aiming to develop knowledge of those models. The point processes therein treated are fundamental point processes such as Binomial processes, Poisson processes, Cox processes and Hawkes processes, which play an important role in phenomenal distribution.

In the third and fourth chapters, we discuss Bayesian statistics. In those chapters, we consider occurrences of events in a focus area leading to the realization of a spatiotemporal Cox process, which is a generalization of the Shot Noise intensity measure. A hidden Poisson process contributes to the intensity measure function that is distributed according to a Dirichlet process. As for data concerning spatial locations of occurrences during several periods of observation, we give an overview of statistical

inference with non-parametric and MCMC methods, which are introduced for simulation models.

The fifth chapter is devoted to specific point processes for epidemiology in a spatial epidemiological framework. We apply such processes to the geographical distribution of the malaria disease in Central Africa Republic. We analyse the case where the point process support is discrete with at most one occurrence for each element. The models treated in previous chapters are tested, yielding a risk characterisation of malaria disease in Central African Republic.

In sixth chapter; we discuss and interpret the obtained results.

Keywords: Point processes, Poisson process, Dirichlet process, sampling, spatiotemporal, MCMC methods, epidemiology, malaria desease

BÖLÜM 1

GİRİŞ

1.1 Literatür İncelemesi

Mekânsal istatistikler yirminci yüzyılın ikinci yarısında istatistiksel metodolojinin geliştirilmesi için en verimli alanlardan biri olmuştur. Aslında birçok problem stokastik doğaya sahiptir ve uygun istatistiksel yöntemlere göre modellenmeye ihtiyaç duyar. Bu tezde zamansal ya da mekânsal olarak yayılmış olayların oluşumunun istatistiksel çalışmasıyla ilgileneceğiz.

Bir nokta süreci, verilen bir uzayda nokta konumlarını modelleyen raslantısal bir süreçtir. Zamansal, mekansal ve zaman-mekansal nokta süreçleri, olayların oluşum sıralarının modeli yüzünden sismoloji (depremlerin merkez üssü ve şiddeti) [5, 6, 7 ve 8], ekoloji (bitki çeşitlerinin dağılımı) [9], epidemiyoloji (hastalıkların felaketlerin nokta konumları) [10 ve 11], nöro-kimyasal [12 ve 13] gibi çeşitli alanlarda bir süredir kullanılmaktadır. Sonrasında birçok yazar matematiksel olarak çalışmalar yapmıştır, örneğin sismik olayların konumlarına, var oluş zamanlarına ve şiddetlerine göre yayılım analizlerinde matematiksel yaklaşımlar uygulanmıştır[4, 5, 6, 7, 8 ve 14]. Eğer sadece uzay ve zamanda temsil edilen merkez üstlerinin dağılımlarını ele alacak olursak işaretsiz zaman-mekânsal nokta sürecine sahip oluruz. Hastalıkların etkileri üzerine çalışmalarda ayrıca oransal modellemede sahibiz görülmektedir [15 ve 16]. Karr[17], DAley ve Vere-Jones[3] tarafından nokta süreçlerine sağlanan katkıları göz önüne alalım. Aslında zaman-mekânsal nokta süreçlerinin gerçekleşmesi, rastgele her bir noktanın bir olayın yer veya zamanını belirttiği bir nokta kümesidir. Aşağıdaki örnekleri inceleyelim:

1) Hastalığı taşıyan sivrisineğin bir kişiye ulaşma zamanı nokta süreçleri ile gösterilebilir. Nokta süreçlerin gerçekleşmesine başka bir örnek olarak, şekil 1.1[40]'de görüldüğü gibi her bir nokta "bir olay" ı temsil etmektedir(bu durumda kanser oluşumu olarak).Burada ilgi çekici olan durumların daha çok kümelenme ya da yığılmış olmanın mümkün olan herhangi kirlilik kaynaklarının konumunun kümelenmeyle beklenenden daha çok ilgili olup olmamasıdır. Bu geri planı değerlendirmek için, hastalık kontrol haritası(gösterilmedi) durumların yayılımlarında beklenen varyasyon temsilinde sıklıkla kullanılır, bu genelde nüfus yoğunluğu ile ilgili bir fonksiyondur.

2) Bitkilerin yerleri mekânsal nokta süreçleri ile gösterilebilir. Şekil 1.2[40]'deki diğer bir örnek, küçük bitki pozisyonlarının bir alanda mekansal nokta süreçlerinin gerçekleşmesi, konumların sayısında mekansal süreçler altındaki gözlemlerden oluşan veri kümesinin bir örneğidir. Bu tipteki verilerin analizinin temel amacı, her bir ölçümün sadece rastgele hata bileşeninin varlığı ile gözlemlendiğinin kabul edilmesiyle beraber ilgi alanındaki bütün konumlarda mekânsal süreçlerin değerini belirlemektir. Fakat biz ayrıca yukarıda bazı yönlerden öntanımlı sınır yâda olay, ortalamanın süreçlerinin gerçekleştiği bölgeleri belirlemeyle de ilgileniyoruz.

3) Bir periyot boyunca bir bölgede bir depremin merkez üssünün yeri ve oluş zamanı, zaman-mekansal nokta süreçlerinin oluşumudur.

Bir nokta süreci aslında bir metrik uzayda model noktalarının konumlarının stokastik sayma ölçüsüdür[3]. Bu $\lambda(.)$ koşullu yoğunluk süreci ile tek yolun bazı şartlar altında karakterize edilmesidir[3,17,18,19 ve 28]. Modelleme sadece aracılığı ile yapılır. Poisson süreçleri genelde korelasyon olmayan uzay yada zamanda homojen problemlerin çalışılması için önerilen basit bir model örnektir. Genelde $\lambda(.)$ zaman ve mekândan bağımsız fakat öncesindeki olaylara bağlıdır.

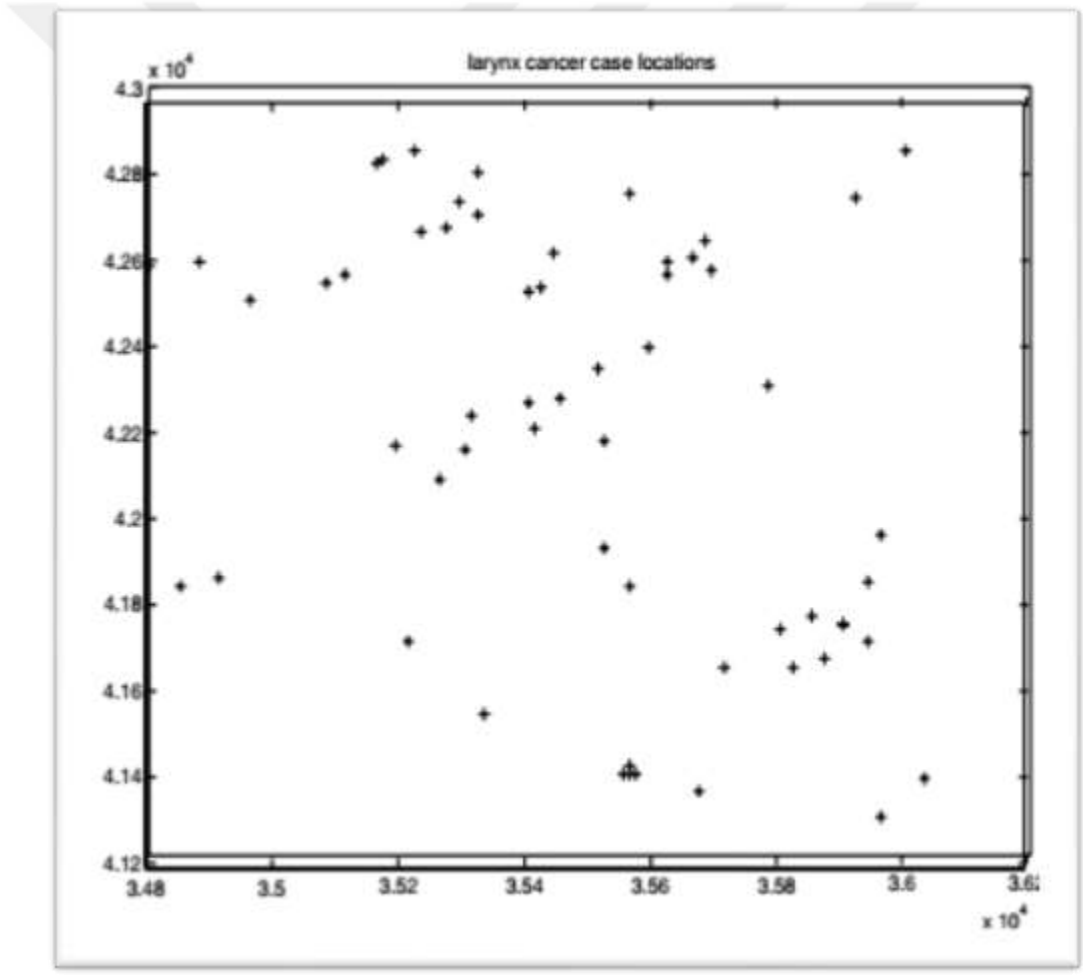
Otokorelasyona sahip zamanda ya da uzayda çeşitli yoğunluklardaki birçok olay vardır. Dolayısıyla biz olayların karakteristiğini göz önüne alarak onları ilgilenilen uygun bir modele göre dikkate almalıyız. Örneğin bazı yazarlar Cox süreçleri [21, 22 ve 23] ve diğerleri Hawke süreçleri [6, 24, 25 ve 27] gibi hafızanın süreçleri ile ilgilenmektedir.

Coğrafi bölgelerde olayların oluşumunun zamansal serilerinin oluşması durumunda zaman-mekânsal nokta süreçleri uygun sayma, tanımlama, çalışılmış olayların görünümüne ve nihai açıklaması için oldukça uygundur.

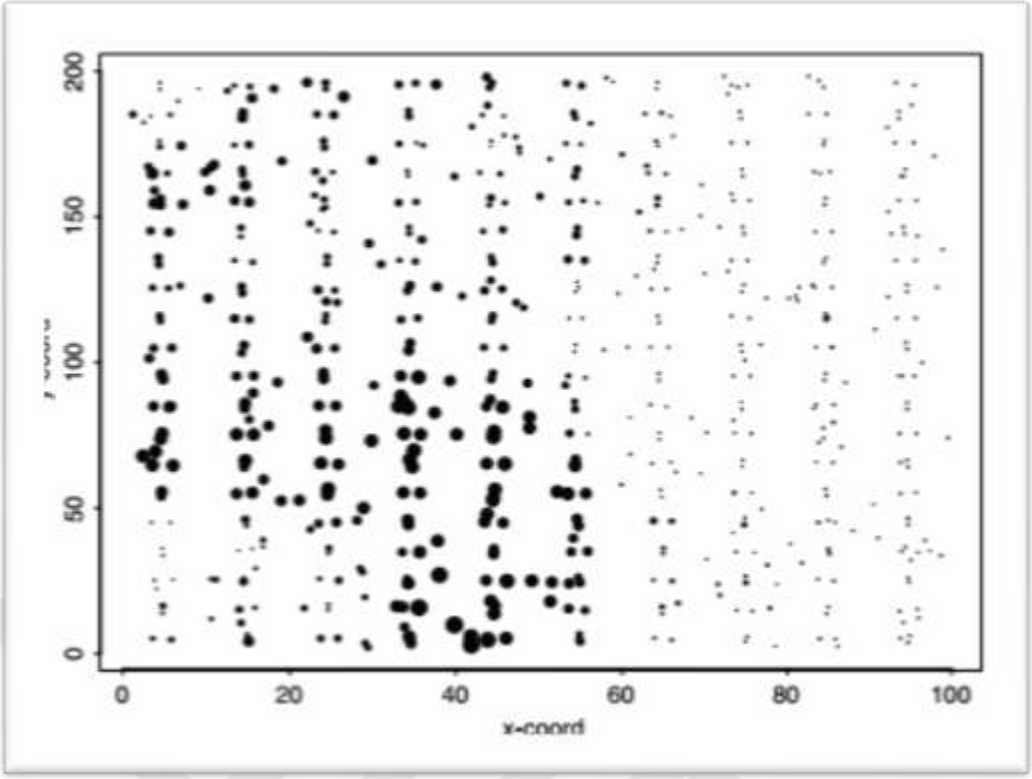
Genel olarak, $\lambda(.)$nın önceki olaylara da bağlı olduğuna dikkat edilmelidir. $(\mathcal{H}_t)_{t\geq 0}$ bir filtrasyon olsun, t zamanı hariç tutulduğunda, tarihsel sürecin şartına bağlı olarak $\mathcal{H}_{t-}$ in komşuluğunda sezgisel olarak beklenen olayları belirleyen zamansal ve mekansal noktanın sürecini, onun koşulları olasılık yoğunluğu $\lambda(x,t/\mathcal{H}_{t-})$aracılığıyla tanımlayabiliriz. Orta değişkenlerin genelinde yoğunluk süreci boyutsal örnekleme planı vasıtasıyla gözlemlenebilir. Fakat gözlemlenen orta değişken şartına bağlı olarak aşırı yayılım yinede önerilebilir, bu bilinmeyen veya beklenmeyen orta değişkenlerden kaynaklı rastgele etkiler modellerinin varlığı anlamına gelmektedir. Brix'in çalışmalarına devamla [27], Brix, Chadoeuf [28] ve Möller'in Cox Shot Noise süreci üzerine, Brix ve Diggle [22]'ın Cox log-Gaussian süreci üzerine çalışmalar vardır. Gözlemlenen nokta sürecinin yoğunluğuna etki eden çevresel faktörlerin arasındaki olay korelâsyonlarını açıklama açısından Dirichlet sürecinin [30], Shot Noise sürecinin genelleştirilmiş bir modeline katkısını araştırmak oldukça önemlidir. İntegrasyonun çevre değişkenleri üzerine bilginin nokta süreci modelinde öncülüğü bizi Bayesyen yaklaşımına götürür [33].

Programlama gücü Bayesyen analizi için oldukça etkili araçların çıkmasına müsaade edene kadar, 80 yıl boyunca bu olasılıklara olan nesnel yaklaşım, Bayesyen ve Bayesyen olmayanlar arasında oldukça canlı bir tartışma konusu olmuştur. Bayesyen çıkarımsal yöntemi, çevresel değişkenlerin olası etkilerini, kullanılan verilerde dahil olmak üzere bir hiyerarşik düzen vasıtasıyla dikkate almak için uygulanabilir. Sıtma hastalığının veri dağılımimkânlarını, sağlık göstergesinin uzaysal analiz değişkenini ve patojen yayılımındaki riski, bunun yoğunluğunu, şiddetini ve sayısının evrimini, enfeksiyonun zamansal ve mekânsal yayılımını ve boyutsal çeşitliliğinin gözlemlenmesini ortaya koyarız. Green ve Richardson[34] tarafından önerilmesi üzerine biz, modellemede gizli katkının sayısının değişken olduğu konusuna açıklık getirmek için bir tersinebilir sıçrama MCMC (Markov Chain Monte Carlo) modeli geliştiriyoruz.

Bu çalışma sadece verinin bir coğrafi bölge içerisinde bir nokta modeli olarak gözlemlendiği olayların gerçekleştiği durumlara odaklanmaktadır. Detaydaki gelişim ile bağlantılı olan avantajlar, böylesi bir koleksiyon boyunca daha fazla kaynak düzenlendikçe bu tür veri setlerin ulaşılabilirliliğinin artmasına yol açmıştır. Hali hazırda teorik olarak iyi kurulmuş olsa da mekânsallığın uygulamalı istatistiksel modellemesi ve özellikle de zamansal mekânsal nokta modelleri hala hem uygun bir metodolojinin bulunmayışından hem de sıklıkla yüksek hesaplama gerekliliklerin varlığından dolayı oldukça yeni bir alan konumundadır. Fakat buna rağmen metodolojinin ve ulaşılabilir pratik envanterin evrilmeye ve gelişmeye devam etmesi önemlidir. Böylelikle epidemiyologistlar ve diğer araştırmacılar verilerinden azami fayda sağlayabilirler.



Şekil 1. 1 Bir nokta süreci olarak 1974 den 1983 e İngiltere de Lancashire deki vaka ile bağlantılı gırtlak kanseri vakası haritası (DIGGLE 1990)[40]



Şekil 1. 2Mekânsal bir sürecin gözlemlerine bir örnek: The Piazza Road verisi [Higdon et al 1999.] her bir nokta Piazza yolunun etrafında yapılan bir dioksin konsantrasyonu ölçümünün lokasyonu göstermektedir [40].

1.2 Tezin Amacı

Epidemiyoloji alanı, belirli bir insan nüfusu içerisindeki risk faktörleri çalışmaları ve hastalık vakaları, bunların çoğalmaları ve yayılmaları konusu ile ilgilenmektedir. Fakat aynı zamanda sıkı sistem genişlemesiyle hayvanlar veya bitkilerle de ilgilenir. Epidemiyolojide sıklıkla araç gereçler ve istatistiksel modeller, daha sonra da belirleyici ve olasılıksal dinamik modelleri kullanmıştır. Bu yaklaşımlar nokta süreci ve veri analizinde, aynı zamanda modeller üzerinden patojen faktörlerin difüzyonunu önceden tahmin veya kontrol etmek için kullanılmıştır. Enfeksiyon problemlerinin yaygın olması bağlamında ve yeni ortaya çıkmakta olan hastalıklarda matematiksel ve istatistiksel modeller güçlü analiz ve önsezi gereçleri sunmaktadır. Bunun aksine matematik modelleme yaklaşımının biyolojik olgular oluşturduğu açık problemlerden beslenebilir. Bu nedenle matematik, biyoloji ve hatta epidemiyoloji bir sürü açık sorunun ve henüz keşfedilmemiş alanların olduğu muazzam bir çalışma alanı olarak gereklidir.

Diğer taraftan uygulamalı matematik çalışmaları konusu epidomiyoloji ve genellikle biyoloji etki alanına ya çok nadiren ya da oldukça kısmi olarak ulaşmaktadır.

Son zamanlarda matematik alanındaki uzman dersler ve/veya epidemiyoloji alanındaki istatistikler bazı epidemiyoloji problemlerinin çözümüne oldukça büyük katkılarda bulunmaktadır. Diğer taraftan Orta Afrika Cumhuriyetinde yaşanan problemlere yönelik uygulamalı matematik çalışmaları azdır. Hatta epidemiyolojiye yönelik olanlardan bile daha azdır. Bu tezin amacı matematiksel gereçlerin önemini göstermek ve bunları sahra altı Afrika ülkelerinin en az gelişmişlerinden biri olan Orta Afrika Cumhuriyetinin en büyük halk sağlığı problemlerinden birini teşkil eden sıtma hastalığı gibi epidemiyolojik problemlerin çözümünde kullanmaktır. Coğrafi konumundan dolayı Orta Afrika Cumhuriyeti sivrisineklerin ve sıtma taşıyıcı faktörlerin gelişimi için oldukça elverişli bir tropik iklime sahiptir.

Bu epidemik hastalığın yok edilmesi veya en azından pandemik seviyeye indirgemek için genel anlamda tüm alanlarda ve özelde halk sağlığı alanındaki gelişme faktörleri çabalarına rağmen sıtmanın her geçen gün daha fazla insana yayılması durdurulamamaktadır. 3895139 sakini ile Orta Afrika Cumhuriyeti (genel nüfus ve ikamet edenler, (GCPR 2003)) genel sağlık ve özelde de sıtma, çoğunlukla hastalık vakaları ve korunmasız nüfusun ölümü yüzünden Orta Afrika Cumhuriueti'nin nüfusunda azalış gösterecek gibi görünmemektedir [2].

Dünyadaki hastalık oranı ve ölüm oranının ilk sebebi olarak, sahra altı Afrika vakaların %75 ni sıtma olarak kaydetmektedir. Sahraaltı Afrikasında her yıl ölümlerin %90 yani 2.7 – 3 milyon ölüm sıtma yüzünden gerçekleşmektedir. Buradaher 30 saniyede bir çocuk ölmektedir [1]. Orta Afrika Cumhuriyetinde patoloji konsültasyonun ve hastane bakımının %40 nı ölümlerin %45 ni teşkil etmekte, medikal kaynaklara göre [1], vaziyeti daha vahim bir hale getiren birincil faktörlerin arasında şunlar yer almaktadır: fakirlik, sağlıksız çevresel koşullara bölgenin istikrarsızlığının eklenmesi, insanların sağlıksız çevrelerde yaşamaya zorlanması. Böylece insanlar bu hastalığın giderek artmasına sebebiyet veren sivrisinek ısırıklarına maruz kalmaktadırlar. Amaç bu musibetin modellenmesine ve analizine matematiksel olarak katkıda bulunmaktır.

Öncelikli Amaçlar:

1) Bu hastalığın ortaya çıkmasına, yayılmasına ve gösterdiği dirence yol açan süreci belirlemek ve ölçmek;

2) Patojen-sistemi oluşturan nüfusu karakterize etmek;

3) Nüfuslar ve bunların mesafe, yayılma hızı üzerindeki etkileri arasındaki etkileşimi ortaya koyan nispi süreci çalışmayı mümkün kılan Peter Diggle (1990)'ın öne sürdüğü nokta süreci ile modellemek;

4) Enfeksiyonun kökeni o geçiş yollarını ve ortaya çıkışını analiz etmek ve pandemiğin risklerini tahmin etmek.

Metodoloji:

Bu hedeflere ulaşmak için aşağıdaki yöntem kullanılacaktır:

1) Epidemiyojik veri (sıtmanın istatistiksel verisi) Orta Afrika Cumhuriyeti'nin, halk sağlığıarşivinden ve Dünya Sağlık Örgütü'nden toplanmıştır. Kütüphane belgelerinin yanında internetteki araştırmalar bu çalışmayı güçlendirmek için kullanılmıştır.

2) Teorik kısmın tanıtımı bizim bu alandaki öncülerimizin kullandığı gereçlere hakim olmamıza müsaade etmektedir.

Toplanan verilerin analizi aşağıdaki farklı metotlar ve yazılımlar tarafından gerçekleştirilmiştir:

1) Epidemiyolojik veri için Microsoft Office Excel kullanılacaktır.

2) Yazlım R paketi spatstat [35] grafiksel sunum, analiz ve zamansal ve mekânsal nokta sürecinin simülasyonu için uygulanacaktır.

1.3 Hipotez

Burada analiz edilecek hastalıkların zamansal ve zaman-mekânsal yönlerine istatistiksel vurgu yapan epidemiyolojik analizlerin bazı kavramları ifade edilecektir. Bu istatiktiksel bakış açısını daha ileriye taşıyabilmek için bazı varsayıma dayalı sorular ifade edilecektir.

Nokta süreç verisini içeren uygulamanın analizinde doğal başlangıç yeri şu soruyu sormaktır. “Tamamen rastgele fenomenleri mi gözlemliyoruz?” Diğer bir deyişle, gözlemlenen noktalarda ilgilendiğimiz yapıların olup olmadığına bakılır, yani mekânsal rastgeleliğin tamamında olması mümkün olmayanlar ilgi çekicidir.

Nokta sürecinin uzayda noktaların bir rasgele dağılımı olduğunu söylüyoruz. Bu problemi sonuçlandırmak amacıyla geçici ve mekânsal veya geçici nokta sürecini elde etmek için zemin eksenini veririz. Bu noktalar bazen olaylar denilen t_i momentleridir. Poisson nokta sürecinin bilinen yoğunluğu tarafından üretilen noktalar kümesini ölçen Ripley [36] istatistiği kullanılır. Bu, teorik olarak beklenenden çok büyük olan mekânsal yapıları karakterize eder ve çevresel heterojenliği [34] hesaba katan gerçek verilerin gerekliliğini sağlar. Aynı zamanda yerel türlerin varlığını ve değişken olabilen hastalıkların yoğunluğunu destekler.Modelleme seçiminde, korelasyon fonksiyon çifti homojen olmayan zaman ve mekan için genişleyecektir. Parametrik olmayan Bayesiyen yaklaşımı gizli Morkov alanı ve epidemiyolojideki uygulamasını sunmamıza izin verir.

1.4 Tezin Organizasyonu

Bu tez, sonuçlar bölümüyle beraber altı bölümden oluşmaktadır. İlk iki bölümde 3 ve 4 bölümlerde önerilen modelleri ve yöntemleri anlamamızı sorgulayan nokta süreçlerinin temel rotasyonları anlatılacaktır.

Bölüm 1 genel girişe ayrılmıştır.

Bölüm 2 bazı temel nokta süreçlerinin girişi ile devam etmiştir; bu bölümde temel rotasyonlar ve sonrasında bu çalışma için gerekli kalıplar ifade edilmiştir.

Bölüm 3’de Bayesiyen yaklaşım ve Monte Carlo metotlarının bazı kavramları verilmiştir. MCMC (Markov Chain Monte Carlo) prosedürleri ve örnekleri sonsal dağılıma yakınsamak ve hesaplamak için ifade edilmiş ve tartışılmıştır.

Bölüm 4’de Cox sürecinin genelleştirilmiş shot Noise bir zaman-mekânsal gelişimi ifade edilmiştir. Gizlenmiş süreçlerin zaman-mekânsal yoğunluğu üzerindeki gizli süreç katkısının bir Dirichlet süreci ile ilişkili Poisson süreci olduğu durumunu değerlendirdik. Gözlem tarihlerine mekânsal olayların tam harita tipinin verisi için olabilirliğin yazılması Bayesyen çıkarımı yazmamıza izin verir.

Bölüm5'de ifade edilen zaman-mekânsal nokta süreç modellerinin tamamen mekânsal epidemiyoloji yaklaşımı üzerinde durur. Mekânsal örneklem birimlerinin sayma sürecinin dağılımsal özellikleri ifade edilmiştir.

Aynı zamanda nokta süreçlerinin epidemiyolojik ve stokastik modellerdeki uygulaması ifade edilmiştir.

BÖLÜM 2

NOKTA SÜREÇLERİ TEORİSİNİN TEMEL NOTASYONLARI

2. Giriş

Nokta süreçleri teorisinin başlangıcı çok eskiye dayanmaktadır. Çünkü detaylı bilgi için geçmişte insanlar gökyüzünün bir bölgesindeki yıldızları sayıyorlardı, sel baskını, depremler, kuyruklu yıldızın ortaya çıkması gibi olayları kaydediyorlardı.

Daha yakın zamanlarda, nokta süreçleri yaşam tablosu, sayma problemleri, fiziksel parçacıklar, nüfus süreçleri, haberleşme mühendisliği gibi alanlarda kullanılmaktadır.

Bir nokta süreci temel teşkil eden stokastik mekanizmaya, fonksiyona veya gözlemlediğimiz sonlu nokta şeklini üreten kavrama göre kullanılır.

Süreç hakkında yaptığımız kabuller ele alınan problemin sonuçlarının çıkarımında en az modelin seçimi kadar etkilidir. Bu nedenle, epidemiyolojideki vakaların uygulaması göz önünde bulundurularak sürecin temel teorisinin üstesinden gelinmeye çalışılacaktır.

Bu bölümde, nokta süreçleri teorisine kısa bir giriş yapılacaktır ve teorinintemel notasyonları ifade edilecektir. Teori hakkında [37] ve kitaplar [3, 39 ve 41]'e bakılabilir.

2.1 Nokta Süreçleri

Bu bölümde, nokta süreçleri hakkında (özellikle Bernoulli süreci, Poisson süreci, Cox süreci ve Hawkes süreci) genel bilgi verilecektir.

Nokta süreçleri, ölçüm teorisini kullanır. Bir nokta süreci, tamsayı değerlerindeki stokastik ölçüm olarak görülebilir. (X, d) metrik uzayı üzerinde olayların dağılımı olarak

kullanılır. Bir yıldaki pandemik tarihlerin listesi gibi sabit sayıda olma gerekliliği olmayan noktaların sayısını rastgele değişken olarak alır. Bu durumda, çalışmamız (X, d) ayrılabilir ve tam metrik uzayındadır ve $\forall \omega \in \Omega, \mu_\omega(.)$ olmak üzere $\mu: \Omega \times \mathcal{B}(X) \to \mathbb{N}$ rastgele değişkeni ele alırız. $X \subseteq \mathbb{R}$, $\mathbb{N}$ değerlerini bir ölçüm kümesi olsun. Metrik uzayda, metrik olarak olaylar arasındaki yakınlığı dikkate alacağız.

Tanım 2.1.1: $(\Omega, \mathcal{B}, \mathbb{P})$ olasılık uzayı olsun. Olasılık uzayında N nokta süreci, $\mathcal{M}_o$ da noktaların rastgele dağılımıdır. Tamsayı değerli X üzerinde ölçüm kümesi aşağıdaki gibi tanımlanır:

$$N: \Omega \to \mathcal{M}_0$$

$$\omega \mapsto \sum \delta_{X_i}(\omega) \tag{2.1}$$

$$N = \sum_i \delta_{X_i} \tag{2.2}$$

Burada $X_i: \Omega \to X$ rastgele değişken, $(X_i)_{i \in I}(X, d)$ metrik uzayında tamsayı değerli rastgele değişkenler ailesi, δ ise Dirac ölçümüdür. X üzerindeki nokta sürecinden, doğal olarak X^k üzerindeki nokta sürecine ulaşılır. Buna göre, k negatif olmayan tamsayı olmak üzere,

$$N = \sum_{i_1, \ldots, i_k \; distincts \; \in \{1, \ldots, n\}} \delta\left(X_{i_1}, \ldots, X_{i_k}\right) \tag{2.3}$$

$\mathcal{M}^{(k)}(.) = E[N^{(k)}(.)]$ bu süreçle ilgili olan X^k üzerindeki ölçümü göstermektedir. Örneğin, $k = 2$ olması, pandemik çiftlerinin yoğunluğu anlamına gelecektir. Olayların bağımsız olduğu kabulü altında, bu ölçüm aşikâr olur. Genelde, nokta süreçlerinin kapsamındaki noktaların ilişkisini gösterir [38].

Literatürdeki araştırmalar, genelde zaman üzerindeki olaylarla ilgilidir [3 ve 39]. Kronolojik gelişim, uzayın filtrelenmesini gerektirir, böylece $(\Omega, \mathcal{B}, \mathbb{P})$ olasılık uzayı, $(\mathcal{F}_t)_{t \geq 0} \mathfrak{B}$ nın aşağıdakini sağlayan σ-cebri olmak üzere $(\Omega, \mathcal{B}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})$ yazabiliriz.

$$\forall t, \; s < t \Rightarrow \mathcal{F}_s \subset \mathcal{F}_t \tag{2.4}$$

$$\mathcal{F}_{t^-} = \bigcap_{s<t} \mathcal{F}_s \tag{2.5}$$

Bir σ-cebri, ölçülen olaylarınkümesi olarak tanımlanır. Olasılık anlamında bu, olayların ayrıklaştırılması veya örneğin "t zamanında cevaplanmış sorular" demektir. Filtrasyon, ölçülebilir olaylar kümesindeki değişimi gösterir ve bilgi kaybı veya kazanımı sağlar [43].

Tanım 2.1.2: Bir filtrasyon, ölçülebilir uzayda σ-cebrinde dizi artışıdır. Tahmin edilebilir süreç, prior zamanda bilinen stokastik süreçtir. Ele alınan dizilerde kapalı olan en küçük sınıfı oluşturur ve soldan sürekli süreçlerin hepsini içerir. $(\Omega, \mathcal{B}, \mathbb{P})$ üzerinde verilen $F = (\mathcal{F}_t)$ filtrasyonda, eğer bütün $t \in \mathbb{R}_+$ ler için $(X_t)\mathcal{F}_t$- ölçülebilirse ölçülebilir fonksiyonların bir (X_t) süreci F-tahmin edilebilirdir.

Tanım 2.1.3: X noktalar kümesi ve (X, d) metrik uzay olsun. Eğer, $x_n \in (X, d)$ olmak üzere herhangi $X = (x_1, \dots x_n, \dots)$ rastgele deneyin noktaları ise X e yapılandırılmıştır denir. Eğer (X, d) uzayının sınırlı herhangi Borel $\mathcal{B}$ kümesinde noktaların sayısı sonlu ise yapılandırma yerel olarak sonludur. N^{lf} yerel sonlu yapılandırma ailesidir. Dolayısıyla yapılandırma noktaların birikim süreci değildir: $\forall x \in X,\ \forall r > o\ \cap B(x, r)$ sonludur.

Yukarıdakilerden, dışmerkez dağılımlarının çoklu noktalar gösterebileceğini söyleyebiliriz. Yani, iki veya daha çok dışmerkez aynı mekânda yerleşmiş olabilir. Bu durumda çoklu noktaların yapılandırılmasından bahsederiz. Fakat uygulamaların çoğunda, noktaların yapılandırılması, fiziksel olarak imkânsız olduğundan, iki veya daha çok nokta aynı mekânda olamaz. Örneğin, aynı merkezi paylaşan iki hücre olamaz. Dolayısıyla basit noktaların yapılandırılması sözkonusudur[44].

Çakışık noktalar yoksa$X$ üzerinde N nokta süreci basit nokta süreci olarak adlandırılır. Bu durumda neredeyse her $x \in X$ için $N(\{x\}) \in \{0.1\}$ dir. Başka bir değişle, bir N nokta süreci eğer aynı zamanda en çok bir olay görülürse basit olarak adlandırılır:

$$\forall x \in X, N(\{x\}) \in \{0,1\} \text{veya} \mathbb{P}(N \in N_X) = 1 \qquad (2.6)$$

2.2 Zamansal nokta süreci

Geçici nokta süreci belirli olayın gerçekleşmesi için $\{\tau_i\}_{i \in I}$zamanından oluşan stokastik veya rastgele süreçtir. Bu süreç sürekli zamanda meydana gelen iki sonuçlu olayın zaman dizisinden oluşan bir süreçtir [3]. Zaman noktalarının sonlu bir sınırlanmış kümesinde tanımlanmışlardır. Zamanda her bir noktada sayısız değer alabilen sürekli değerli sürecin tersine, bir nokta süreci bu zamanda olayın meydana gelmesi ve gelmemesi olmak üzere sadece iki mümkün değer alabilir. Bu süreç farklı bölgelerde kullanılır, birçok uygulama alanı vardır; uygun bir nokta süreç modeli geçici homojensizliğe uygulanabilir. Örneğin, epidemiyolojide bir hastalığın ortaya çıkış sıklığı

hem risk altında olan popülasyonun uzaysal dağılımını hem de mevsimsel etkiler gibi sistematik geçici varyasyonu yansıtır. Bu alt bölüm ile aşağıdaki tanım verilebilir;

Tanım 2.2.1 Aşağıdaki geçici nokta süreci göz önüne alınsın

$\sum_{i=1}^{\infty} \delta_{X_i}$ (2.7)

Burada X_i pozitif keyfi değişkenler öyle ki (monoton artanlık) $X_1 \leq \cdots \leq X_i \leq \cdots$

0 zamanından t zamanını içerecek kadar zamanda meydana gelen olayların sayısı $N(t)$ ise bir rastgele süreç $\{N(t), t \in [0,\infty)\}$'e sayma süreci denir. Bir sayma süreci için, [s,t) aralığında meydana gelen olayların sayısı$N(t) - N(s)$olmak üzere, aşağıdakilerin varlığını kabul edelim.

$$1. N(0) = 0$$

$2. N(t) \in \{0, 1, 2, \ldots\} \forall t \in [0,\infty)$

3. $0 \leq s < t$ için, $N(t) - N(s)$ (2.8)

Sayma süreci, varışları modeller. Her bir olayın meydana gelmesini bir varış olarak adlandırırız. Örneğin, $N(t)$bir şehirde$t$ zamanına kadar görülen hastalık durumlarının sayısı ise, her bir durum yine varış olarak isimlendirilir. Şekil2.1 bir sayma sürecinin mümkün gerçekleşmesini ve karşılık gelen örnek fonksiyonunu gösterir.N ile ilişkilendirilen $N(t)_{t\geq 0}$ sayma süreci,

$\boldsymbol{N} = \sum_{i=1}^{\infty} 1_{\{X_i \leq t\}}$ (2.9)

olarak tanımlanır.

(N_t)F-uyarlanmış için bir filtreleme $F = (\mathcal{F}_t)_{t\geq 0}$ olsun.F ile alakalı bir N nokta sürecinin dengeleyicisi aşağıdaki koşulları sağlayan bir tek $A = (A_t)$sürecidir:

1) A, F-öngörülebilirdir.
2) $M_t = N_t - A_t$ Martingale merkezlidir

Martingale aşağıdaki özelliklere sahip $X(t)$ rastgele süreçtir;

Her $T > t$ için $E[X(T)/\mathcal{F}_t]$=$X(t)$ ve $E[|X(t)|] < \infty$ (2.10)

Burada $\mathcal{F}_t, t$zamanındaki filtrasyondur. Eğer bir öngörülebilir pozitif (λ_t) süreci

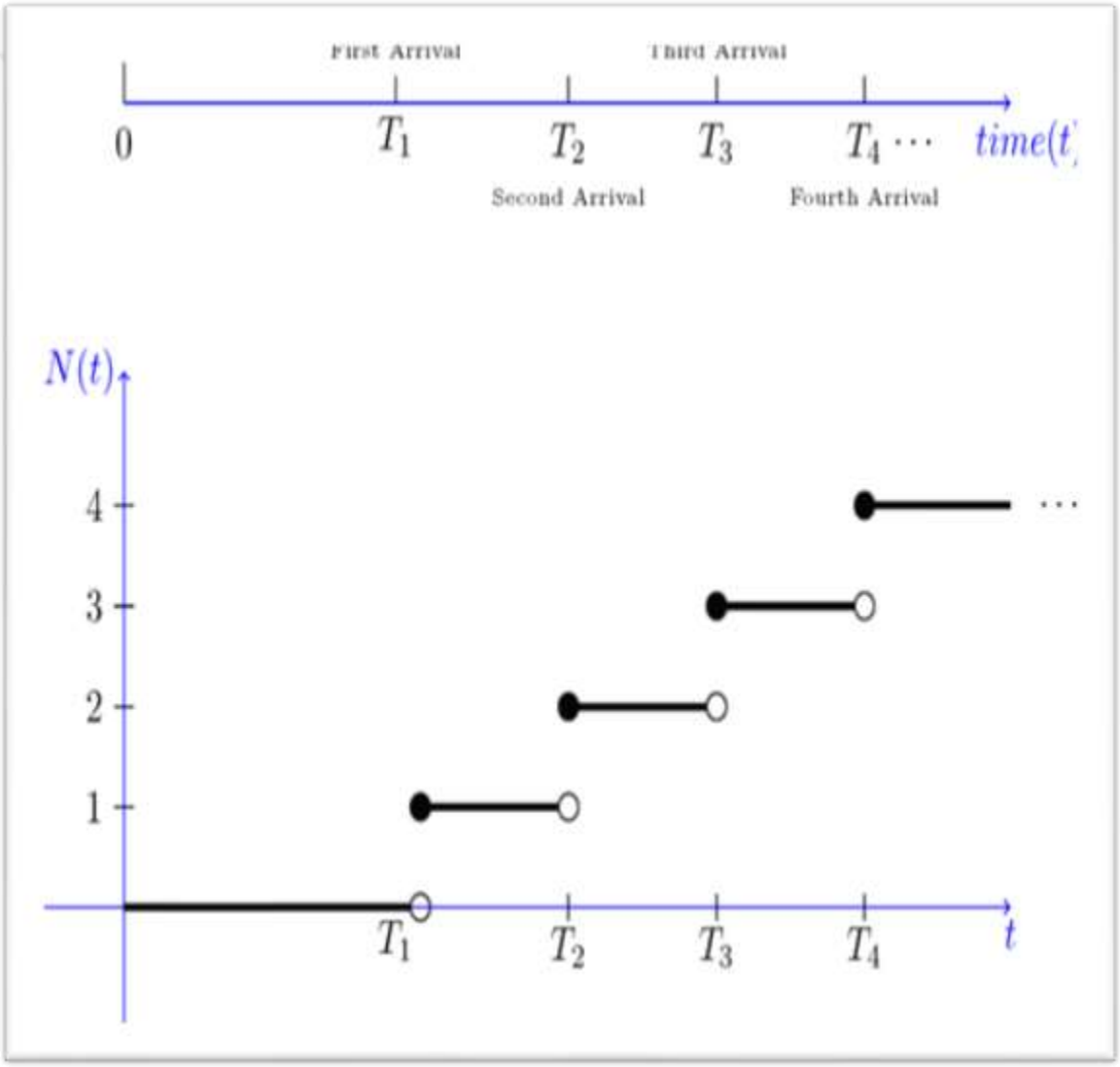
$dA_t = E(dN_t/\mathcal{F}_{t-})\lambda_t dt$olacak şekilde varsa

$$A_t = \int_0^t \lambda_s \, ds \qquad (2.11)$$

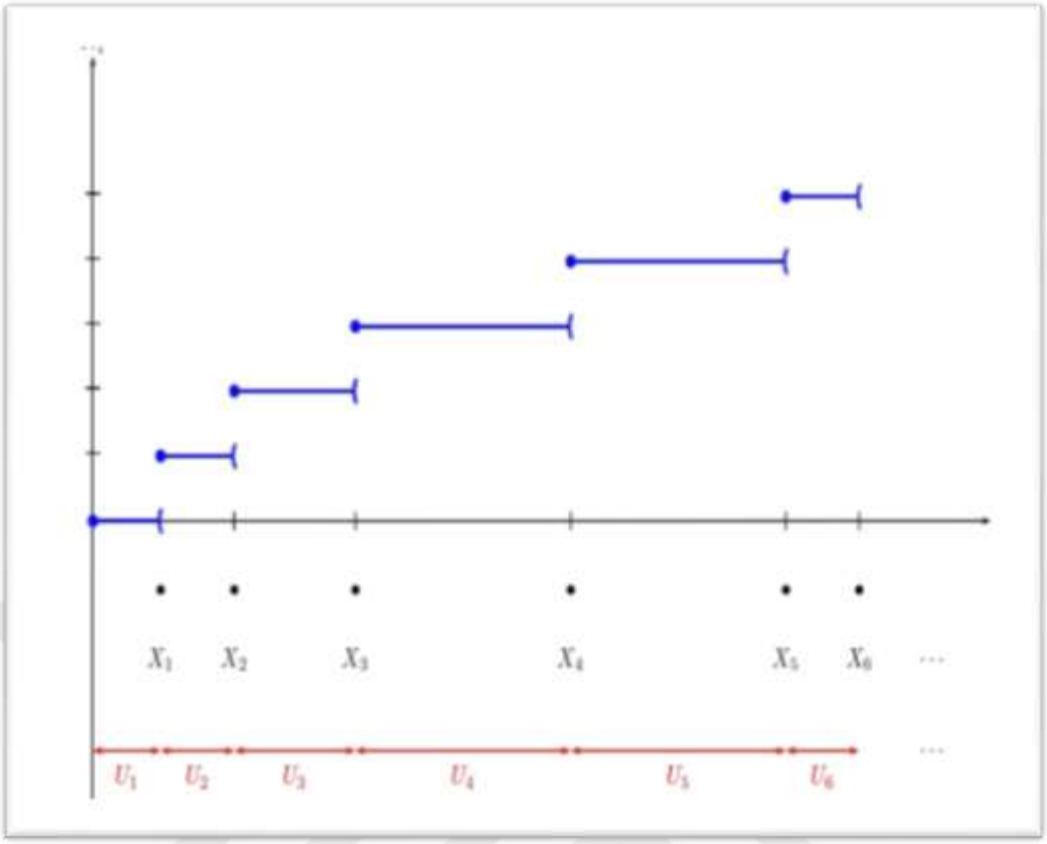
yazabiliriz.

λ= (λ_t)N' nin stokastik F-yoğunluğu olarak tanımlanır.

Not edelim ki bir nokta sürecini işaretlerle ilişkilendirebiliriz (2.8). Bu işaretler nitel veya nicel olabilir. Geçici nokta sürecin özelliklerinden biri şudur ki süreç, X_i meydana gelme tarihleri ile ayrıca $U_i = X_i - X_{i-1}$. Bekleme zamanlarıyla belirtilebilir. Süreç basit olduğunda, 1 birim sıçrayışla artan basamak fonksiyonu, sayma süreç temsili gözlenir, bu geçici nokta sürecin gerçeklenmesini mükemmel bir şekilde belirtir. (Şekil 2.2)



Şekil 2. 1 Bir Sayma Sürecinin Mümkün Gerçekleşmesi Ve Karşılık Gelen Örnek Fonksiyonu



Şekil 2. 2 Geçici nokta sürecin bekleyiş zamanlarının grafik gösterimi[44]

2.3 Yoğunluk Fonksiyonu

Bir nokta sürecinin ana yapısının temel birinci mertebeden özelliği yoğunluk fonksiyonu olarak bilinir. Belirli bir nokta sürecin yoğunluk fonksiyonu gözlenen noktaların görünüş ve yapısını dikte etmenin anahtarıdır. Bu nedenle, sonucun büyük bir kısmı yoğunluğun tahmini ile alakalı mekânsal nokta süreçlere dayanır. $N(\mathcal{B})$,$\mathcal{B}$ penceresi dâhilindeki olay veya noktaların sayısını göstersin ve δ_x, $x$merkezli küçük bir dairesel bölge olsun. Yoğunluk fonksiyonu$\lambda(.)$, aşağıdaki gibi verilir;

$$\lambda(x) = \lim_{|\delta_x|\to 0} \left\{\frac{E[N(\delta_x)]}{|\delta_x|}\right\}, x \in X \qquad (2.12)$$

$\lambda(x)$ (bir önceki bölümde bahsedildiği gibi) x konumunda birim alana düşen noktaların tahmini sayısı olarak yorumlanır. Eğer $|\delta_x|$ sonsuz küçük bir alan belirtiyorsa $(x)|\delta_x|$ ile uzayın bu alt kümesinde bir nokta bulunması olasılığını gösterir. Ayrıca

$$E[N(X)] = \int_X \lambda(x)dx \qquad (2.13)$$

şeklindedir. (2.8)'den görülür ki (2.7)'nin $N(X)^{-1}$ ile çarpılması X üzerinde tanımlı sürekli olasılık yoğunluk fonksiyonunu sağlar. Bu vurgularla nokta süreç modelinde geleneksel teoriden fazlası ile desteklenen bir kabul: sadeliktir (bu terim [52]'de Illian vd tarafından kullanılmıştır.) Ayrıca, Diggle [53] tarafından sıralı nokta süreç olarak refere edilen bir nokta süreç, belirli bir x lokasyonunda bir noktadan daha fazla nokta gözlemleme olasılığı sıfır olan nokta süreçtir. Böyle bir nokta süreci belirli bir x yerinde bir noktadan fazla gözlem olasılığının sıfır olmasıyla ilgilidir. Pratikte, bu ihmal edilebilirdir. Örneğin yoğunluk fonksiyonuna rağmen kaydedilen olayların kaba yerleşimi ile ilgili veri kümesinin boyutu bilgisini de ele alan standart nokta süreci analizidir. Yoğunluk fonksiyonu, örneğin "eşit oyun alanlarında" farklı iki noktayı karşılaştırmak için kullanışlı olur.

2.4 Bir Nokta Sürecin En Çok Olabilirlik Süreci

Bu bölümde gösterilecektir, bir nokta sürecin olabilirliği stokastik bir süreçtir modelde içerilen parametre tahmindir. Bu kavram [45, 39 ve 46]'nın gibi birçok yazar tarafından çalışılmıştır.Belki de, maksimum en çok olabilirlik tahmini birçok nokta süreç modeli için zorlu olmasına rağmen, en çok olabilirlik tabanlı metotlar model-fitting nokta şablonları için en çok kullanılan yaklaşımlardır. Böyle problemlerin üstesinden gelmek için, genellikle "log-en çok olabilirlik" fonksiyonu göz önüne alınır ve en çok olabilirliği hesaplamak için Mont Carlo similasyonu gerekebilir.

Bazı durumlarda, bir T final zaman noktasında sürecin dağılım bilgisinden süreçler, bir sürecin [0,T] sonlu zaman aralığındaki zaman dinamiklerini tanımlamak için kullanılabilir.

Bunu göstermek için, kolay gözlenen nokta süreçlerin yoğunluğunun $\theta \in \Theta(\Theta \subset \mathbb{R}^d)$.Parametresine bağlı olduğunu kabul edelim.$(X, \mathcal{B}, \nu)$, f ölçü uzayında $(x_i, t_i)_{i\in[1,n]}$ şeklinde gözlenen veriler için, $X \subset \mathbb{R}^2$ uzaysal bölgesinde [0,T] zaman periyotlu, N basit uzay geçici nokta sürecin gerçekleşmesine karşılık gelir. O halde, Θ' nın fonksiyonu ile ilgili olarak en çok olabilirlik,

$$L_T(\theta, (x_i,\ t_i)) = \left[\prod_{0<t_j\leq T} \lambda_\theta(x_j, t_j/\mathcal{H}_{t-})\right].\exp(-\int_0^T\int_X \lambda_\theta(x, t/\mathcal{H}_{t-})\nu(dx)dt) \quad (2.14)$$

Önde gelen log-en çok olabilirlik

$$\mathcal{L}_T\big(\theta,(x_i,t_i)\big)=\sum_{0<t_j\le T} log\lambda_\theta(x_j,t_j/\mathcal{H}_{t-})-\int_0^T\int_X\lambda_\theta(x,t/\mathcal{H}_{t-})\nu(dx)dt \qquad (2.15)$$

Burada $\lambda_\theta(x,t/\mathcal{H}_{t-})$ sürecin (x,t)noktasındaki yoğunluğudur ve $(\mathcal{H}_{t-})_{t\ge0}$ herhangi bir tarihteki tarihsel gözlemi sağlayan filtrasyondur. θ bir parametre, $\mathbb{R}^d$ (d∈$\mathbb{N}^*$), uzayındaki bir vektör sıklıkla istatistiksel çıkarımdır.

θ elemanları için maksimum olabilirlik tahmin araçları log-olabilirliği maksimum hale getirerek elde edilir. (2.14) ile verilen tahmin araçları genellikle tutarlılık, tarafsızlık, asimptotik normalite ve etkililik gibi ilginç özelliklere sahiptirler [6,47 ve 48]. Not edilmelidir ki, $t_i$meydana gelme tarihi tam olarak gözlemlenmezse, log-en çok olabilirlik kavramı özdeş değildir. Bu kavram ileriki bölümlerde daha ayrıntılı görülecektir.

Olabilirlik tekniği ile bazı modellerin yeterlilik testinin olasılığı da gösterilmiştir. Ağ modeller durumunda, olabilirlik oran metodu modelin daha genel bir sınıfında belirli bir modeli kabul eden bir testin inşası ile sonuçlanmıştır [9, 49 ve 50].

Daha sonra, Diggle vd [54] mekânsal geçici nokta süreçlerin kısmi olabilirliğini, yaşam çözümlemelerinde orantılı risk modelleri için kısmi olabilirliğin adaptasyonu olarak öne sürmüştür. Bu yaklaşımla, gözlenen olay zamanlarında iyileştirme yapılması ve gözlenen zaman-sıralı olaylar için olabilirliğin göz önüne alınması ile geçici entegrasyonlar önlenir.i, zaman noktasında gözlenen i.inci olay için olasılık

$$p_i=\frac{\lambda(t_i,s_i/\mathcal{H}_{t-})}{\int_{\mathcal{R}(t_i)}\lambda(t_i,s/\mathcal{H}_{t-})ds} \qquad (2.16)$$

dır. Burada $\mathcal{R}(t_i)t_i$zamanındaki risk kümesini gösterir. Uzaysal kesikli süreçlerde $|W|<\infty$ ve $\mathcal{R}(t_i)=\{s_i: t_j>t_i, j\in\{1,\dots,N\}\}\cup\{s\in W\}$, $w\subset\mathbb{R}^2$ gözlenen bölgelerdir.

Aslında integral bir toplamdır (uygulama için bkz. Diggle [55]), oysaki uzaysal sürekli süreçte, olaylar W bölgesinde herhangi bir yerde oluşabilir, buradan $\mathcal{R}(t_i)=W$'dir.

Kısmi olabilirlik $i=1,\dots,N$ için bütün $p_i$olasılıklarının çarpımıdır ve buradan kısmi log-en çok olabilirlik aşağıdaki gibi yazılır;

$$\mathcal{L}_p(\Theta,x,T,W)=\sum_{i=1}^{N}\log(p_i) \qquad (2.17)$$

Yaşam analizlerinde, kısmi benzerliğe dayalı tahmin genellikle kabul edilir. Çünkü maksimum kısmi log-en çok olabilirlik tahmin araçları asimptotik normalite gibi tam maksimum olabilirlik tahmin araçlarının sahip olduğu genel asimptotik özelliklerin hepsine sahiptir. Bunun yanında etkinlik kaybı olabilir ve mekânsal nokta süreç modelleri için tutarlı tahminleri veren metotun hangi şartlara sahip olması gerektiği henüz tam olarak belirlenememiştir [56]. Örneğin orantılı risk modellerinde (2.17)'dekip_i kesrinde iptal edilen temel risk bazı parametreler kısmi olabilirlik için tanımsız olabilir. Bu çalışmada göstereceğimiz diğer bir önemli kavram rezidü analizidir [57,5].

2.5 Nokta Süreci için Rezidü Analizi

Çoğunlukla herhangi bir istatistik unsurla elde edilen rezidü analizi,olası problemleri tanımlamak ve bunların niteliğini değerlendirmek için bize genel bir yaklaşım sağlar. Sistematik bir bakış açısının her elde edilişinde, bunun verideki tüm sistematik unsurları yakalamada başarılı olduğunu söyleyemeyiz.

Rezidü sayesinde bağımlı değişkenin gözlemlenen değeri ile bir model altındaki tahmini değeri arasında fark, problemdeki modern uygunluğun niteliğini belirlemede kullanıldığını anlarız. Nokta süreçleri çerçevesinde, bir nokta sürecinin rezidü kavramı için Baddeley'e bakılabilir [58]. Bu makalede, geçici nokta süreçleri için rezidü kavramının zamansal nokta süreçlerine bir genelleştirilmesi verilir. Aslında, doğadaki düzen bunun varlığını zor kılmasına ragmen güncel yaşam analizinde yaygın şekilde kullanılan son şey budur.

Geçici bir nokta sürecinin olay zamanı $(t_i)_{i\in[1,N]}$ olsun, geçici bir nokta süreci için rezidü analizi prensibi t_i'deki birikimli yoğunluğun tahmini hesaplamasını içerir ve $\hat{\Lambda}(t_i)$ ile gösterilir. $\tau_i = \hat{\Lambda}(t_i)$ yazarak, gözönüne alınan nokta süreci modelinin $\hat{\Lambda}$ birikimli yoğunluğuna sahip olduğu varsayım altında, 1 yoğunlukla bir Poisson sürecinin gerçekleşmesine sahip oluruz.

Eğer $(N_t)_{t\geq 0}$ bir $(\lambda(t))$ şartlı yoğunluk süreciyle tanımlanan geçici nokta süreciyle ilişkili süreç hesabını ifade ederse, o zaman rezidü süreci, $I(t) = N_t - \int_0^t \lambda(s)ds$ ile $(I(t))_{t\geq 0}$

yenilik süreci sıfır beklentilisinin bir Martingale'si olmasından oluşturulabilir, böylece model doğrudur. O zaman rezidü sürecitir.

$$R(t) = N_t - \hat{\Lambda}(t) \tag{2.18}$$

şeklindedir. Modeldeki verinin yeterliliğinin kalitesi 0 mertebeli rezidüye yaklaşılmasına göre belirlenir. Ayrıca nokta sürecinin rezidü analizinin çok boyutlu bir süreçte ve zaman-mekânsal süreçte incelendiği de belirtilmelidir [59, 60 ve 61].

2.6 Dispersiyonve Otokorelasyon Analizi

Bir(X, d)metrik uzayı üzerinde bir N nokta sürecinin gerçekleşmesine istatistiksel analizi, gözleme bağlıdır. Aslında tekniksel kısıtlamalara ve örnekleme maliyetlerine göre bazı durumlarla da karşılaşabiliriz. X'deki yoldan dolayı olayın konumunun elde tutulması, olaydaki ve örnekteki noktalar arasındaki mesafe ölçümleri veya örnekteki olayların hesaplanması gibidir [62, 15, 64 ve 65].

Bu metodların bir özet sunumu Perry tarafından yapılmıştır [66] ve ekoloji ölçümündeki bu çalışmadaki uygulama için kıyaslamalı bir çalışmadır. Ayrıca belirtelim ki, bir nokta dağılımının bu farklı analiz prosedürleri genellikle bir sabitin birinci ve ikinci mertebeli özelliklerine ve bazen de izotropik nokta süreçlerine bağlıdır [67]. Bu son salınımlar farklı mekânsal otokorelasyon ölçümlerine sahiptirler. Verilen bir mekânsal birimli istatistiksel değişken için benzerliklerin derecesini hesaplamak kronolojik seri ile müküdür [68]. Coğrafi bilgi sistemindeki (GIS) otokorelasyon mekânı, bir objenin etraftaki diğer objelerle benzerlik derecesini anlamaya yardım eder. Mekansal otokorelasyonu ölçmek için kullanılan çoğu unsurlardan en önemlisi Moran'ın indeksi dir (aynı değişkenli farklı değerler arasındaki korelasyon). Bu indeks, bireylerin mekânsal konumlarına uydurulan ağırlıkların hesaplanan otokorelasyon indeksinden varyans tahmininde farklıdır (burada varyans tahmininden kasıt aşağıdaki kesrin paydasıdır):

$$I_m = \frac{n \sum_i \sum_j W_{ij}(x_i - \bar{x})(x_j - \bar{x})}{(\sum_i \sum_j W_{ij}) \sum_k (x_k - \bar{x})^2} \tag{2.19}$$

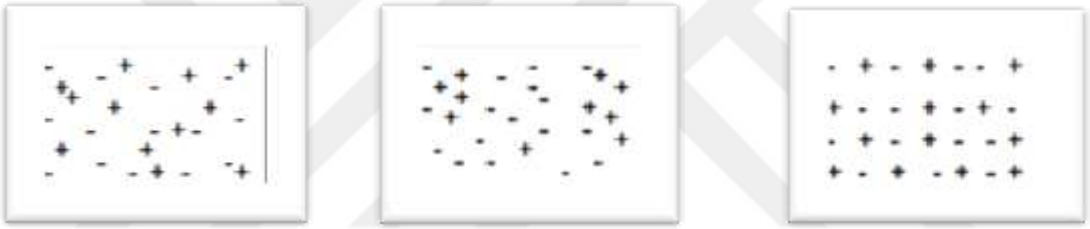
Burada $\bar{x}$, W tanım kümesi üzerindeki değer anlamına gelir ve W_{ij} matriks ağırlıkları i ve $j$mekânları ile ilişkilidir, $\sum_i \sum_j (W_{ij})$ bu ağırlıkların toplamıdır ve x_i, i birimdeki

gözlenen değerdir. Örneğin, mekânsal ilişkiler basit komşuluk ilişkileriyle tanımlanabilir: Eğer i ve j, 0'ın dışındaki birer komşu iseler, o zaman W_{ij}, i ve j arasındaki uzaklığın tersi olabilir.

x değişkeni sayı veya gruplanan veriler için çalışılan bölgedeki olayların oranı alabilir. Böyle bir istatistiksel indeksin doğruluğu farklı hipotez davranışları için yeterlidir.

Böylece I'nin ortalama değeri,

1) x_i ile gösterlen makansal rastgele dağılımı için 0 değerine yakın olur
2) Pozitif mekânsal otokorelasyon, haritadaki (izlerin mekânsal toplamı) benzer küme değerlerinin bir arada olmasıdır.
3) Negatif mekânsal otokorelasyon haritadaki (izlerin mekânsal dağılımı veya düzensizliği) benzemeyen küme değerlerin bir arada olmasıdır.



Rastlantısal dağılım ($I_m = 0$) Kümeleme ($I_m > 0$) Dispersiyon (düzenlilik $I_m < 0$)

Şekil 2. 3 Farklı hipotezler altında Moran'ın indeksi I_m değerler.

Moran'ın I faydalısı, kesikli veya sürekli değişken çeşitli uygun izler için, mekânsal dağılım üzerinde elde edilen bilginin niteliği düşük ve çok ayırt edici olmadığından oldukça kısaltılmıştır. Örneğin değişiklikte mekansal olmayan ve mekansal değişkenleri paylaşan ve korelasyon içeriğinin uzunluğunu birleştiren variogram veya korelogram sağlandıkça I faydalısı daha az olur [69]. Henüz variogram veya korelogram klasik olarak yoğunluğun hesaplanmasında kullanılmaz; fakat mekânsal örneklerin yapılmasından önce tanımlayıcı bir unsur olarak oldukça kullanılmıştır.

İstatistikler, birbirinden bağımsız olan gözlemlere dayandığından, bu durum, bu mekânsal otokorelasyonun neden önemli olduğunun ana sebeplerinden biridir. Eğer bir grafikte otokorelasyon varsa bu durum gözlemlerin birbirinden bağımsız olmasını ihlal eder.

Potansiyel uygulama ekolojinin küreselleşmesini, dağılımını ve hastalık ile yayılmış olay veya dağılımı yayılmayı analiz eder mi? Bu gibi problemleri mekânsal otokorelasyonun kullanılmasıyla daha iyi anlaşılabilir.

Belirtelim ki, Moran'ın hesaplama indeksi kullanımdaki grup bilgisi durumunda belli başlı populasyon dağılımının potansiyel heterojenliğini gözönüne almaz 100.000 kişi için 100 kişi veya 100 kişi için 1 durumun oranı aynı istatistiksel değere sahip olacak ve böylece ne aynı biyolojik hassasiyet ne de aynı biyolojik değer var olacaktır.

Son olarak Moran'ın hesaplama indeksi,prior ağırlıkların ve komşuluk civrındaki sayıların seçilmesi anlamına gelir ki bu seçim her zaman kolay değildir.

2.7 Tamamen Rastlantısal Dağılım

Varsayılan bir tamamen rastlantısal dağılım(CRD), bir homojen Poisson nokta sürecine neden olan bir süreçtir ve 2'li prensibe dayanır:

1) Bulunan bireyler bakımından çalışılan bölgedeki her bir yerleşime göre tek düzeliğin bir prensibi olarak benzer özelliklere sahiptir.
2) Yerleşimdeki bir olayın doğruluğu altında bağımsızlık prensibi diğer yerleşimdeki olayların doğruluğunu etkilenmez.

Diğer bir deyişle, çalışılan bölgenin mükemmel şekilde homojen olduğunu ve bulunan kişiler arasında etkileşimin olmadığını varsayarız. Bu şartlar altında bu saf rastlantı bir Poisson olasılılık dağılımı ile formalize edilebilir. Belirtelim ki, noktaların mekânsal konfigurasyonu yönleri, benzer şekildeki birim uzaylı nokta hesaplamasının tahmini bir RCD altında kıyaslanan ölçü ve şekil bakımından nokta hesaplamasının değişkenlik derecesidir. Son dağılım, 2.3 başlığıyla açıklayacağımız homojen Poisson sürecinin gerçekleşmesine karşılık gelir.

Sayıma ait olasılık dağılımı, tahmin varyansa eşit olan bir Poisson dağılımının bir CRD olması ile ilişkilidir. Bu CRD farkının belirteç unsuru olarak kullanılan ortalama varyansın oranını açıklar. Kabul edelim ki $N_1, N_2, \dots, N_n, n$ mekan biriminde gerçekleşlenen etkilerin hesaplanmış noktaları olsun. Gözlemlenen dispersyon indeks:

$$\hat{I}_d = \frac{\sum_{i=1}^{n}(N_i - \bar{N})^2}{(n-1)\bar{N}} \qquad (2.20)$$

'dir, burada, $\overline{N}$, N_i'in aritmetik ortalamasıdır.

Bu indeksin teorik ve pratik ilginçliği Taylor tarafından vurgulanmıştır [70]. Özellikle ekolojide bireyler arasındaki; dispersiyon derecesinin bir ölçüsü olarak homojen Poisson süreci varsayımı altında kullanılır. TeorikI_d indeks dispersiyonu dağılımdaki yoğunluk çeşidi veya heterojenlikteki çevresel dağılımdır. Bu olaydaki indekse aşırı dağılım noktaları veya aşırı dağılımlar noktası denir. Örneğin, bu olay, bireyler karşılıklı etkileştiğinde veya çevresel değişkenli ortak dağılım gösterildiğinde ekoloji ortaya çıkar. Buna rağmen I_d iyice 1'den çok küçük olduğundan ona alt dağılım veya alt- dağılım bireyleri denir. Ekolojide bu olay, kişisel davranışlarda ve bölgesel davranışlarda rekabet varsa ortaya çıkar; böylece bireyler çok ender bir oraya gelir.

Kabalık indeksli olan dispersiyon deceresini ölçen başka bir indeksin varlığını da gözönüne alalım[70]. $\hat{I}_d$, bir tamamen rastlantısal dağılım için$n-1$ serbestlik dereceli ki-kare dağılımlı $(n-1)\hat{I}_d$'in asimptotik dağılımına dayanan alt dağılımlı veya aşırı dağılım hipotezine karşı CRD hipotez testlerini inşa etmede kullanılır [71].

2.8 Temel Nokta süreçleri

Bu bölümde, en önemli süreçlerden biri olan Poisson nokta süreçlerini (ppp) ele alacağız. Daha sonra Bernoulli ve iki terimli süreçlerinin, yenileme süreci, Cox süreci ve Hawkes süreci gibi nokta süreçlerinin tuturlılığı tanıtılacaktır.

Poisson nokta süreci, nokta süreçleri teorisinde önemli bir rol oynar. Bunlar, noktalar arasında tam mekânsal rastlantılılığa (CRS) sahiptir. Aslında çoğu mekânsal nokta modelleri noktalar arasındaki etkileşimin bazı derecelerin yok pratikte kullanışsızdırlar. Buna karşın, istatistiksel çalışmalarda referans noktalar olarak ve daha çok yapısal nokta süreç modelleri için yapısal blok olarak hizmet ederler. Yerel olarak integrallenebilen, her sınırlı $\mathcal{B} \subseteq S$ için $\int_{\mathcal{B}} \lambda(x)dx < \infty$, $\lambda: S \to [0,\infty)$ yoğunluklu $S \subseteq \mathbb{R}^d$ için $(S, \mathcal{B}, \nu)$ ölçü uzayını gözönüne alalım.

Ayrıca $\mu(\mathcal{B}) = \int_{\mathcal{B}} \lambda(x)dx$ yoğunluk ölçüsünü tanımlayalım. Bu ölçü, her sınırlı $\mathcal{B} \subseteq S$ ve her $x \in S$ için $\mu(\{x\}) = 0$ dağılımı için yerel olarak $\mu(\mathcal{B}) < \infty$ sonludur. Şimdi Bernoulli ve Binomial (iki terimli) süreçleri tanımlanacaktır.

2.8.1 Bernoulli ve Binomial süreçleri

Herhangi bir homojen Poisson sürecinin önemli bir özelliği, uzayın sınırlı bir alt kümesinde noktaların sayılarının verilmesidir; bu noktalar alt küme üzerinde bağımsız ve eşit olarak dağılmıştır. Bu şartlı özellik homojen Poisson süreciyle ilgili diğer sonuçları ve onların kullanımlarının tanımlanmasını ispatlamada ve simülasyonda da önemli rol aynar.

Bir homojen Poisson sürecinin bu tür şartlanmasından kaynaklanan N nokta süreci, Bernoulli sürecinin özel bir çeşididir [72]. Bazı kitaplarda Bernoulli süreçleri, örnekolarak süreçler diye de adlandırılır [73]. Bu yüzden S üzerinde eşit olarak dağılmış tek bir nokta içeren bir N nokta süreci Bernoulli süreci diye adlandırılır. $X{\sim}\mathcal{U}(X)$ ile $N = \delta_X$ diye yazılır.

Bir Binomial süreci (iki terimli) nokta süreci n bağımsız Bernoulli süreçlerinin üst üste çalışmasıdır. n noktalı Binomial nokta süreci, S üzerindeki eşit dağılıma göre bağımsız olarak dağılmış n noktanın bir birleştirilmiş sürecidir. Bu ise$S = \sum_{i=1}^{n} \delta_{X_i}$şeklinde yazılır; burada X_i,$\mathcal{U}(X)$'e göre bir bağımsız ve özdeş dağılım ($i.i.d$)'dir

$$\forall A \in \mathcal{B}, N(A){\sim}\mathfrak{B}(n, \frac{\nu(A)}{\nu(S)}) \qquad (2.21)$$

2.8.2 Poisson Süreci

Rastgele bir dağılımda, varyans düzenli bir dağılım için olandan daha büyüktür ama kümelenmiş dağılım için olandan daha az önemlidir.

Gerçekten, rastgele bir dağılımda bireylerin yeniden gruplandığı bölge alanlarını ve diğer alanların boş olduklarını gözlemleyebiliriz. Buna rağmen, kümelenmiş alanlar kümelenmiş bir dağılım için olandan daha az ve alanlar düzenli bir dağılımdan daha boş olacaktır. Poisson dağılımı, varyansın deneye eşit olduğu bu dağılımın modellenmesine imkân verir.

Bu bölümde, epidemiyolojik problem modeline olanak sağlayan teorik bir yapı geliştireceğiz. Poisson nokta süreci (PPP) rastgele nokta modeli için en basit ve en önemli süreçtir. PPP, sıfır model ve nokta örneklere gerçekçi model inşası için bir başlangıç noktası gibi merkezi bir rol oynar. Homojen Poisson nokta süreci genellikle

nokta süreç modellemesinin teorisine esas olarak düşünülür. Homojen PPP, CSR nin ideal teorik gösterimidir ve gözlem oluşumu için en basit stokastik mekanizmadır [53].

Poisson nokta süreci, Nnin yoğunluk ölçümü olarak adlandırılan ve bu nokta sürecinin yoğunluk ölçümü olan $\mathbb{R}^d$ üzerindeki ayrıntılı ölçüm tarafından karakterize edilir [74]. Poisson nokta süreçleri çoğu çalışmada kaynak olarak kullanılır [3].

Tanım 2.8.1$\mu(.),(S,\mathcal{B},v)$ölçülü uzayı üzerinde ayrıntılı, yerel, sonlu bir ölçü ve bütün sınırlı $\mathcal{B} \subseteq S$ Borel kümeleri için $\mu(S) > 0$ ve $\mu(A) < \infty$olsun.

1) $\forall A \in \mathcal{B}$, $N(A)$keyfi $A \in \mathcal{B}$için parametreli bir $P(\mu(A))$ Poisson dağılımına sahiptir. Yani$P(N(A) = k) = \frac{\mu(A)^k}{k!}\exp(-\mu(A)$'dır.
2) Her bir $k \in \mathbb{N}$ için $N(A_1)$, $N(A_2), \dots, N(A_k)$ bağımsız rastgele değişkenlerdir ve $\mu(A_i) < \infty$, iA_1, $A_2, \dots, A_k \in \mathcal{B}$ kili ayrık kümelerdir.

Şartlarını sağlayan bir N nokta sürecine $\mu(.)$yoğunluğu ile bir Poisson nokta süreci denir. Eğer $\rho > 0$ yoğunluk fonksiyonu var ve sabit ise Nnokta sürecine ρ yoğunluğu ile homojen Poisson nokta süreci (HPPP) denir ve $\mu(.) = \rho v(.)$'dır.

HPPP hipotezi çoğu kez uzaysal dağılımla ilgili istatistiksel analiz veya yürütülen çalışmada bir ilgisi olan geçici olaylar boyunca test edilir.

Sezgisel bir bakış açısıyla, bir S uzayı üzerinden bir HPPP, aşağıdaki iki nokta ile birlikte karakterize edilmiş olayların bir CRD si tanımlanabilir:

1) Homojenlik: Snin bütün noktaları oluşum pozisyonu için aynı şansa sahiptir.
2) Bağımsızlık: Farklı olaylar bağımsız bir şekilde birinden diğerine dönüştürülebilir.

Bazı durumlarda, Poisson süreci zamanda oluşan rastgele gözlemlerin bir serisi için bir modeldir. Bu durumda aşağıdaki tanımı verebiliriz:

Tanım 2.8.2$\lambda(t)$yoğunluk fonksiyonlu bir Poisson süreci bir sayma sürecidir, $N(t)$ $[0,t]$aralığında meydana gelen olayların sayısını göstermek üzere t$(N(t))_{t\geq 0}$ bağımsız artışlara ve $\Lambda(t)$ parametreli bir Poisson dağılımına sahiptir. $P\big(\Lambda(t)\big)$herhangi $t \geq 0$ değeri için Poisson dağılımını gösterir, $\Lambda(t)$,

$$\Lambda(t) = \int_0^t \lambda(u)du \qquad (2.22)$$

olarak tanımlanan birikimli yoğunluk fonksiyonudur.

Bu tanımda $\lambda(t) \geq 0,\ \forall t \geq 0$ ve$\int_0^t \lambda(u)du < \infty$ olduğu kabul edilir. Bu yüzden, bu tanımdan $\Lambda(t) = E[N(t)]$elde edilir. $\Lambda(t)$'nin her yerde sürekli ve herhangi sonlu aralık üzerinde sonlu sayıda noktalarda türevli, Λ'nin türevlenebilir olduğu herhangi bir noktada $\lambda(t) = \Lambda'(t)$ olduğunu kabul edelim.Eğer $\lambda,\ \lambda(t) = \rho > 0$ şeklinde bir sabit fonksiyon ise $N(t)$ HPP dir.

$(T_i)_{i\in\mathbb{N}^*},\ N_t = \#\{T_i, T_i \leq t\}$ geçici bir nokta süreci olsun. Bu durumda$(N_t)_{t\geq 0}$, $(T_i)_{i\in\mathbb{N}^*}$süreci ile bağlantılı bir sayma sürecidir. Bu sürece homojen geçici Poisson süreci denir (HTPP). Eğer $N(t),\ \lambda t,\ \forall t \geq 0$ parametreli bir Poisson dağılımı ise bu parametreye Poisson sürecinin hızı denir.

λyoğunluklu bir HTPP basit bir tanımlamaya sahiptir: Art arda gelen olaylar arasındaki zaman aralıklarıλ ters ortalamalı bağımsız katlanarak dağıtılmış rastgele değişkenlerdir. Bu tanımlama HTPP için bir simulasyon algoritması kurmada çok kullanışlıdır.

2.8.3 Yenileme Süreci

Yenileme süreci, tekrarlı bağımsız denemelerde model oluşumu incelemesi için teorik birsistem sağlar. Yenileme süreci çoğu çalışmanın konusu olmuştur; [19, 75 ve 76] da bahsedidiğin üzere birçok uygulama alanında bir modeldir ve önemli teorik problemler için bir kaynaktır. Yenileme süreci ardaşık noktalar arasındaki aralığın bağımsız ve aynı şekilde pozitif rastgele değişkenlere dağıtılmış olma özelliğine sahiptir. Bu süreç, zamanda rastgele meydana gelen olaylar için stokastik modele idealleştirilir. Bu geçici olaylar jenerik olarak yenilemeler veya gelişmeler gibi işaret edilir.

Tanım 2.8.3Eğer bekleme zamanları $V_n = T_n - T_{n-1},\ n \geq 1$ bağımsız ve aynı şekilde dağıtılmış ise ($P(V_n = \infty)$olasılığının sağlanmasıyla birlikte) geçici nokta süreci $T = (T_n)_{n\geq 1}$, $T_n$yenilenme süreleri ile bir yenileme sürecidir. (2.7) terimi, geri dönülen zaman süreci $X'i, X_t = t - T_t$(en son yenilemeden beri geçen zaman)eşitliği ile belirleyerek sayma sürecini tanımlar. Xhomojen Markov süresidir ((F_t^X)filtrasyonuna göre). Eğer $V_n = T_n - T_{n-1},\ n \geq 1$, $\varepsilon(a)$ile belirtilen üstel bir $\frac{1}{a}$ dağılımına sahip ise $(T_n)_{n\geq 1}$, $\rho = a$yoğunluklu bir HPP dir.

2.8.4 Cox Süreçleri

Poisson sürecinin doğal bir genişlemesi negatif olmayan bir $\Lambda(x), x \in S$süreci tarafından sürdürülen, Λüzerinde koşullu bir Cox sürecidir, Cox süreci yoğunluk fonksiyonlu bir Poisson sürecidir [3, 77 ve 78].

Cox süreci, çevresel olarak sürdürülen nokta süreci olayları için doğal bir modeldir, esasen noktalar arasındaki etkileşim tarafından yürütülen olaylar için daha az doğaldır. Bir epidemiyolojik içerikteki bu iki durum örnekleri sırasıyla bulaşıcı olmayan veya bulaşıcı bir hastalık durumlarının uzaysal dağılımı olabilir. Bulaşıcı olmayan bir hastalıkta, olayların gözlenen uzaysal modeli, hassas bireylerin maruz kaldıkları uzaysal değişimden gözlenen veya gözlenmeyen risk faktörlerinin bir bileşimine sonuç verir. Koşullu etkilenmede, olaylar bağımsız bir şekilde meydana gelir. Tersine, bulaşıcı bir hastalıkta gözlenen model yakındaki bireye hastalık geçirdiği bulaşıcı durumların en azından kısmi sonucudur. Bu olgusal farka rağmen, heterojen bir çevrede bağımsız değişkenin stokastik olarak gösterildiği işlemler ve homojen bir çevrede stokastik etkileşimler arasında deneysel olarak ayırım yapmak zor hatta imkânsızdır [79].

İstatistiksel etkinin üç noktası fark edilmelidir. Birincisi, Λ boyunca rastgele çevresel bir heterojenlik modellenebilir, eğer Λ durağan ise Cox süreci de durağandır. İkincisi, N_X'in sadece bir gerçeklenişi var olduğunda N Cox süreci onun karşılığı olan N/Λ Poisson sürecinden ayırt edilemez [29]. Üçüncüsü, çarpım yoğunlukları çözümlenebilir iken en çok olabilirlik genel bilinmezdir.

Burada, eğer$N$, μ'e bağlı olarak homojen olmayan bir Poisson nokta süreci ise stokastik μ yoğunluklu N nokta süreci Cox süreci veya ikili stokastik Poisson süreci olarak adlandırılır. Homojen olmayan bir Poisson süreci boyunca, nokta topluluğunun bazı derecesinin var olduğu CRS durumunu bozmak için özellikle oluşturulmuş örneklere benzer nokta modelleri gözlemlenebilir. Buna rağmen, Poisson sürecinin ana yönü bir sabitin varlığıdır.

$\{x, y\} \in S \subseteq \mathbb{R}^d$olduğu durumda $\Lambda(x)$, ortalama ve kovaryans fonksiyonları ile negatif olmayan, de-boyutlu rastgele bir alan olsun. Sırasıyla

$$\mu(x) = E[\Lambda(x)]\text{ve}C(x, y) = cov(\Lambda(x), \Lambda(y)) \quad (2.23)$$

Bu durumda $\Lambda(x) = \lambda(x)$ üzerinde şartlı, gözlenen olaylar yukarıdaki $\lambda(x)$ fonksiyonu gibi yoğunluk fonksiyonu ile homojen olmayan bir Poisson sürecidir. $d = 2'$de amaçlarımız için, ilgi alanını $\{x, y\} \in S \subseteq \mathbb{R}^2$ şeklindeki coğrafi bölgeye sınırladığımızda bu sonsuz süreç sonlu hale gelir.

Eğer esas stokastik süreç ikinci mertebeden durağan yani $\mu \equiv \mu(x)$ ve, $C(x, y) \equiv C\|x - y\|$, $\|.\|$ Öklid uzaklığı ise bu durumda oluşan nokta süreci de durağan olacaktır. Bu herhangi ardışık nokta modellerinin CSR göstermesi anlamına gelmez. Yoğunluk fonksiyonlarını oluşturan kesin anlaşılması zor ilişki yapılı bir Cox sürecini düşünelim. Stokastik sürecin bir gerçeklenişinin verilmesiyle, sonraki nokta süreci homojen olmayan Poisson olacak, bu yüzden (koşullu olarak) durağan olmayacak. Cox süreci C tarafından tanımlanan noktaların yerlerinin belirlemesini bırakır, modelin görünüşünü kontrol eder.

Bu (koşullu) Poisson süreci ve ikili stokastik süreç arasındaki önemli ayrımın daha iyi anlaşılmasını sağlar. Eğer belli bir nokta modelinde gözlemler boyunca bağlılığın (ilişkinin) herhangi istatistiksel delilini ortaya çıkarabilsek, bu durumda verinin bir Poissondan çıktığı bir önsel kabul ihlal edilmiş olabilirdi. Buna rağmen, (sonlu) Poisson ve Cox sürecini şeklen ayırt etmek aşırı zordur. Aslında bunu yapmak imkânsızdıröyleyse sınırlayıcı kabuller ortalama yoğunluk fonksiyonu değişiminin oranı ile ilgili örnek olusturmazsa sadece basit bir gerçekleniş verilir. Bu yüzden, geçici Cox süreçleri belleksiz süreçlerdir, çünkü stokastik yoğunluklu süreçler geçmişe referans yapmaz.

2.8.5 Hawkes Süreçleri

Hawkess süreçleri, 70 lerin başında [24] Hawkes tarafından tanıtılan çok değişkenli nokta süreçlerinin bir sınıfıdır, stokastik yoğunluk vektörü ile tanımlanır. Hawkes, kendiliğinden çalışan bir süreç tanıtarak kendiliğinden çalışan terimin işaret ettiği lineer gösterimi vurgular. Bir Hawkes süreci zaman değişim yoğunluğunu veya kısmen süreç geçmişi tarafından belirlenen bir sürecin olay oluşum oranını modeller. Diğer taraftan basit bir süreç olayların tarihini hesaba katmaz. Hawkes süreci şimdiye kadar birkaç isim vermek gerekirse deprembilimi, nörolojik bilimi, epidemiyoloji, sigortacılık ve finans gibi uygulama çeşitlerinde kullanılmıştır [80].

N nokta sürecini dikkate alalım. Bu durumda eğer $s < t < u$için$Coc\{N(s,t), N(t,u)\} > 0$ise N ye kendiliğinden çalışan süreç denilebilir. Eğer bu kovaryans değeri negatif ise N kendiliğinden düzeltendir. Bu yüzden kendiliğindendüzelen bir süreçte noktaların engelleyici etkisi var iken kendiliğinden çalışan bir nokta sürecinde noktaların oluşumu diğer noktaların daha iyi meydana gelmesine neden olur. Tanımdan, bir Poisson sürecinin ne kendiliğinden çalışan ne de kendiliğinden düzelten bir süreç olmadığı görülür.

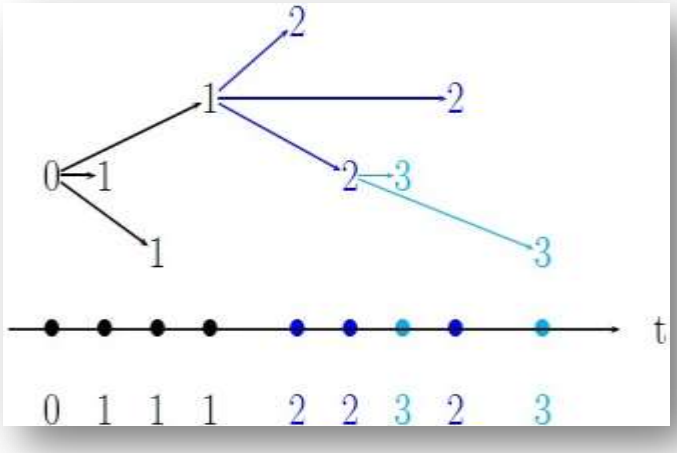
Kendiliğinden çalışan nokta süreç modelleri geçici olarak kümelenen olayları modellemek için sismolojide sık sık kullanılır. Yaygın bir şekilde kullanılan bir örnek koşullu yoğunluğu

$$\lambda(t) = \mu(t) + \sum_{T_n<t} \nu(t - T_n) \qquad (2.24)$$

Hawkes sürecidir. Burada $\mu(t)$belirleyici birikim oranını gösterir ve ν kümelenme yoğunluğunu kontrol eder. $\mu(t)$ nin sabit olduğu bir örnek sunulmuştur [81]. Hawkes modelleri depremlerin artçı sarsıntıları, artçı sarsıntıların da artçı sarsıntıları olabilir düşüncesini kapsayan, bazen salgın tür artçı deprem dizisi (ETAS) modeli olarak adlandırıldıkları sismolojide yaygın şekilde kullanılırlar. Deprem artçı sarsıntılarının modellenmesinde sıklıkla kullanılan ν kümelenme yoğunluğunun bir şekli Omori-Utsu kanunudur [6].

$$\nu(t) = K(t + c)^{-P} \qquad (2.25)$$

Bu zamanla kümelenme davranışındaki güç yasası bozulmasına karşılık gelir. λ yoğunluklu durağan bir Poisson süreci düşündüğümüz durumda grup için merkezde toplanır. Grubun 1 den küçük yoğunluklu Poisson sürecine göre bir ebeveynin bir nesil ürettiği sonlu bir dallanma sürecinin bağımsız gerçeklenişlerini oluşturduğunu düşünürsek bu durumda kendiliğinden çalışan jenerasyon grupları ile tanımlanır.



Şekil 2. 4 Dallanma sürecinin modellenmesi

BÖLÜM 3

POSTERİOR DAĞILIM ÖRNEKLEMİ

3. Giriş

Bu bölümün amacı, gözlenen rastgele olaylardan vebu olayları üreten dağılımdan yararlanarak gelecek olayı tahmin etmekle ilgili kısa özet vermektir. Nokta süreci üzerinde gözlenebilen olaylarla ilgili model parametrelerinin posterior dağılımını elde etmek ilginçtir. Zaman-mekânsal nokta süreci çalışmasından, olayların zamanı veya yerleri veya her ikisini de elde ederiz.

Eğer literatürdeki nokta süreçleri örneklemleri ile ilgili algoritmalara bakarsak, bunları aşağıdaki gibi verebiliriz [82, 83, 84 ve 85].

1) Doğum ve ölüm algoritmaları;
2) Tam simülasyon algoritmaları;
3) Tersinir atlama Markov zinciri Monte Carlo (RJMCMC) yöntemi;
4) Kesikli zaman örneklemi,
5) Sürekli zaman örneklemi,
6) Metropolis-Hastings algoritmaları;

Bu bölümde, üçüncü ve son algoritmalara yoğunlaşacağız.

θ, Θ parametre uzayının bir elemanı olsun. Bayesian yaklaşımı, frekans yaklaşımından Θ uzayı ile ilgili ölçüm olasılığı ile ayrılır. Bu ölçüm, θ nın prior dağılımı olarak adlandırılır ve bu parametre ile ilgili kullanılabilir bilgiler verir. Bayesian istatistiksel

yöntemler, bilgisayarların gelişimiyle 90'larda geliştirilmişlerdir [33 ve 86]. Bayesian çıkarımının amacı, gözlemlerin koşullu posterior dağılımını elde etmektir. [33] kaynağında Robert, prior dağılımın seçiminin Bayesyen analiz çerçevesinde önemli olduğunu ve aydınlatıcı olmayan dağılımların seçimiyle bu seçimin etkisinin azaldığını göstermiştir. Fikir, sağlanan veri ile karşılaştırılan prior dağılımla elde edilen bilginin minimizasyonudur. [33 ve 86] kaynakları prior dağılımın seçimiyle karşılaştırılan Bayesyen yaklaşımının örtenliğini tartışmışlardır. [33]'e göre bu bir dezavantaj değildir. Bayesyen süreci verini, alt bölüm 3.2 de verilecek olan Markov zinciri ile Monte Carlo (MCMC) yöntemlerini kullanır. Biz, Monte Carlo Markov zinciri yöntemine dayanan en etkili hesaplama algoritmalarını kullanacağız. Bu yöntemler, Bayesyen istatistiğin son yıllardaki büyük gelişmeleriyle pek çok problemin çözümünde kullanılmış [34, 87 ve 88]. Nümerik değer üretmeye izin veren yönteme Monte Carlo yöntemi diyeceğiz. Bu, rastgele örneklemler kullanarak matematiksel fonksiyonların istatisitiksel gelişimi anlamında taşır. Diğer bir değişle, nümerik simülasyon tekniğidir. Bu teknik, finans, proje yönetimi, enerji, üretim, mühendislik, araştırma ve geliştirme gibi geniş alanlarda kullanılmaktadır.

3.1 Stokastik Simülasyon algoritmaları

Stokastik simülasyon algoritmaları (SSA), Markov Chain'in nümerik simülasyonudur. Karmaşık tahminleri çözmeyi amaçlar. Deterministiğe benzerliklerine karşın, bu rastgele yöntemler büyük boyutlu uzaylarda araştırmalara izin verir. Temelde, Stokastik etkiler, bir sistemdeki niceller sönümleri mümkün olacak kadar küçük olduklarında kritik olurlar. SSA, sürekli zamanda sonlu simüle edilmiş yörüngeler kurma işlemleridir. Örneğin epidemiyolojide ortaya çıkan zooteknik enfeksiyon, bundan ilk etkilenen insandan daha çok insana geçmezse kötü uygulanmış olur. İnsanla hayvan temasa devam ederse, daha büyük epidemiğe yol açan olaylara sebep olacaktır [89].

Eğer $X_i(t),\ \ i, (i = 1, \dots, n)t$ zamandaki popülasyondaki bireylerin sayısı ve $X(t)$ her biri negatif olmayan tamsayı olmak üzere $i \geq 1 X(t) = (X_1(t), \dots, X_n(t))$ vektörü ile tam tanımlı ise SSA rastgele değişkenler veya konum vektörü $X(t)$'yi başlangıçta (t_0 zamanında) $X(t_0) = X_0$ olarak verilen sistem için tahmin eder. Ani Q matris oranı veya popülasyonda en az bir ani olay olarak (örneğin, doğum, ölüm, hareket, kazalar,

enfeksiyon…) görülen konumlar arası geçişler, zaman üzerinde değişimlere sebep olur. Genelde, bu matris son derece seyrek olur, dolayısıyla bunu direkt kullanmayacağız. Bunun yerine her bir j geçiş için, λ_j oranını ve $\mu_j = (\mu_{j1}, \dots, \mu_{jn})$ (i tamsayı) vektörünü tanımlayacağız. Bu, eğer geçiş $X(t) + \mu_j$ yolunu takip ederse yerdeki değişikliği yansıtır. Keyfi bir geçiş sayısı, tamsayı kümesinden olan j'yi içerebilir. Çünkü homojen-zaman süreci modelliyoruz ve λ_j geçiş oranı, belirli $X(t)$ konumuna bağlı olmasına rağmen t zamanına bağlı değildir. Negatif değer alarak herhangi bir i konumuna yol açan geçiş 0 oranına sahip olmalıdır. ($\{\mu_j\}$) geçişler kümesi ve geçiş oranlarının hesabı ($\lambda(X)$) konumunu verir.

Sezgisel olarak, algoritma bütün geçiş oranları yaklaşık sabit kaldığı zaman periyodunu belirler ve bütün i konum değişkenleri yaklaşık 1 olasılık değeri almak suretiyle sıfırdan büyük kalırlar. Simülasyon, τ uzunluklu period üzerinde sıçrayabilir ve her bir geçiş için, $y_j \sim poisson(\tau\lambda_j)$ olmak üzere $X(t+\tau) = X(t) + \sum_j y_j \mu_j$ boyunca ortaya çıkan geçişlerin Poisson dağılım sayısının etkisini toplar.

Geçiş oranlarının sık sık değiştiği modellerde ve sıçramanın üzerinde periyod zamanları seçildiğinde doğru şekilde etkilerin ayarlanmasında zorluklar artar. Kabul-red algoritmaları olan önemli bir stokastik simülasyon örneği aşağıda verilecektir.

3.1.1 Kabul-Red Yöntemi

Kabul-red yöntemi, keyfi olasılık dağılımından, düzgün dağılım ve ilgili dağılımdan rastgele örneklem bileşenleri olarak verilen rastgele örneklem üretmek için bir algoritmadır. Bu yöntemin, üretilen rastgele sayılarınters birikimli dağılım fonksiyonu metodu (CDF) üzerinde en önemli avantajı ve CDF'ne de terslerinin hesabını gerektirmesidir. Dolayısıyla, çoğu durumlarda daha hızlı çalışır. Bu metod, genellikle dağılımı bilinmeyen verileri simüle etmek için kullanılır. Bunu yapmak için örneklem, dağılım bilinmese bile, bir dağılımdan oluşturulur ve bu, örneklenmiş değerin kabul ya da reddi için kesin kriterdir.

Direkt olarak örneklemenin zor veya imkânsız olduğu $f(x)$ dağılımından örneklemek istediğimizi varsayalım. Ayrıca güvenilir örnekleme yöntemlerinden (düzgün dağılım gibi) tasarı g(x) dağılımımızın olduğunu kabul edelim.

Bu durumda

$C.\mathrm{g}(x) \geq f(x)$, $\forall x$ (3.1)

şeklinde C sabiti varsa, $C, \frac{f(x)}{C.\mathrm{g}(x)}$ oranı 1'e yakın olduğunda $C.\mathrm{g}(x)$ den kabul edilen örneklem, $f(x)$ dağılımından elde edileni sağlar. Diğer taraftan, eğer oran 1'e yaklaşmıyorsa örneği reddedeceğiz.

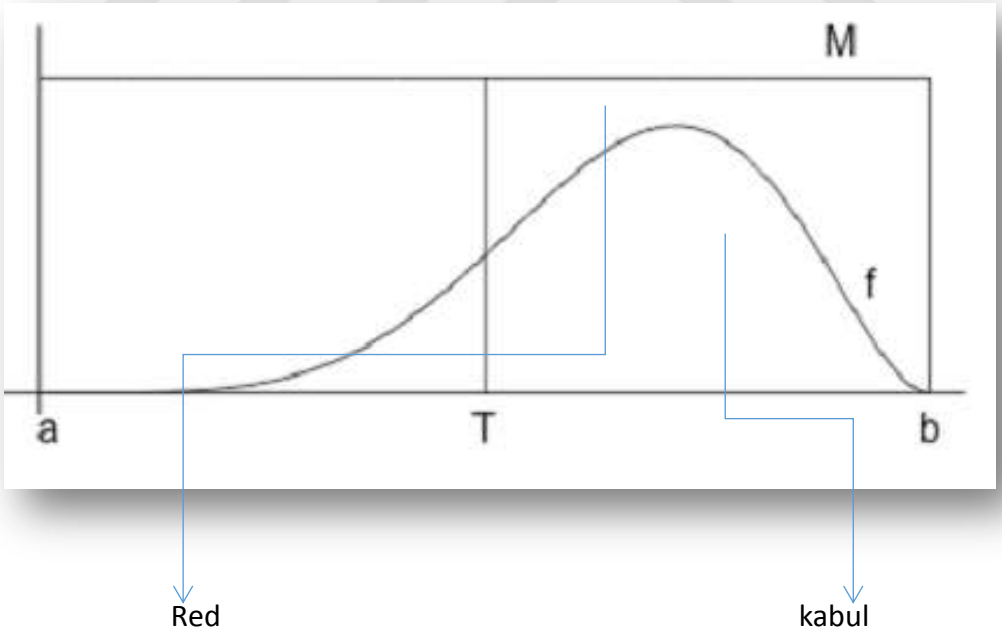
Algoritma A: Kabul-red algoritması

Adım 1: $f(t) \leq C.g(t) = M(t), m(t) = g(t)$ olmak üzere m yoğunluklu T üret

Adım 2: [0, 1]'de düzgün olan ve T'den bağımsız U üret $U \leq \frac{f(T)}{C.\mathrm{g}(T)} \rightarrow kabul$, ve $X = T$ al

Adım 3: değilse $\rightarrow red$, ***adım1***'egeri dön

$M(X)$, [a, b]'de $M(X) \geq f(X)$ koşulunu sağlayan bir fonksiyon olmak üzere, kabul-red yönteminin bir temsili Şekil 3.1 de verilmiştir.



Şekil 3. 1 Kabul-red açıklaması

Örnek 3.1.1. Cauchy dağılımından $\mathcal{N}(0,1)$ normal dağılımı üretilmesihedef dağılım

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$$

Öerilen dağılım.

$$g(x) = \frac{1}{\pi}.\frac{1}{(1+x^2)}$$

M seçimi

$$\frac{f(x)}{g(x)} = \sqrt{\frac{\pi}{2}}(1+x^2)e^{-\frac{x^2}{2}} \leq \sqrt{\frac{2\pi}{e}} \approx 1.52$$

± 1'de değer kabul edilebilir olasılık değerini alır: $\frac{1}{M} \approx 0.66$

3.2 Nokta Süreçleri ve Markov zinciri Monte Carlo (MCMC)

Nokta süreçleri, bazı olayları modellemeye yarar. Örneğin, olayların oluşma tarihlerini modellerken gözlem zamanı sürekli şeklinde düşünülür, fakat bir cihazın bozulma zamanı, yağmur dizisinin tarihi, rastgele tarih dizileri gibi olayların modellenmesinde rastgele değişkenleri kesikli olan Poisson süreci ve Cox süreci gibi klasik modeller kullanılır. Bu modeller epidemiyolojide enfeksiyon ve bulaşma süreçlerinde de uygundurlar.Buradaki amaç, simülasyon için bir algoritma, analiz için istatistiksel araçlar ve iki ve üç boyutlu grafik gösterim araçları önermektir.

Nokta süreçleri çalışmasında asıl zorluk, normalizasyon sabiti, $Z(\theta)$ hesaplanamaz olmasıdır. Bu zorluğun üzerinden gelmek üzere, parametre tahmini ve simülasyon için Markov Chain Monte Carlo (MCMC) tekniklerini kullanmak mümkündür. Şimdi, yukarıdakileri, sınırlı bölgede nokta süreci ile simülüsyon yapmak üzere uygulayacağız. Bu süreç birim yoğunluğun $\pi_1(dx)$ ölçümlü Poisson süreci ile ilgili $f(x)$ yoğunluk fonksiyonuna sahiptir [90 ve 91]. $f(x)\pi_1(dx)$ olacak şekilde sınırlı Markov Chain değerleriyle noktalar deseni oluşturacağız ve Markov Chain yöntemini nokta ekleme ve silme ile yayarken sınırlayacağız.

Nokta süreçlerinin simülasyonunda kullanılan temel tanımları vereceğiz. Sonra, etki ve yakınsaklık hızını tanıtacağız.

x gözlem vektörleri, p gözlem ortak yoğunluğu, $p_1\theta$'nın yoğunluğu, p_2 verilen θ için x'in yoğunluğu olsun. Bu durumda

$$p(x,\ \theta) = p_1(\theta)p_2(x/\theta) \qquad (3.2)$$

olur. Buradan, posterior p_2 dağılımı, ortak dağılım p ile orantılıdır. Monte Carlo yöntemleri, analitik olarak elde edilemeyen posterior dağılımlarda yaklaşık olarak iyi sonuçlar geliştirirler.

3.2.1 Markov zinciri

Markov zinciri sayılamaz çoklukta durum uzaylarına sahip MCMC algoritmalarında kullanılır. Bu türden Markov zincirleri teorisi, sayılabilir durum uzaylarının Markov zincirlerine benzerdir, ama notasyonda bazı farklılıklar vardır. Bazı tanımlar aşağıda kısaca verilmiştir.

Tanım 3.2.1: $(\Omega, \mathcal{B}, \mathbb{P})$ olasılık uzayı ,$(E, \mathcal{E})$ ölçülebilir bir uzay ve $\{X_t, t \in \mathbb{N}\}$ durum uzayı olarak adlandırılan, sayılabilir E kümesi üzerinde değerler alan kesikli-zaman stokastik süreci olsun. Ω çıktılar kümesidir. Eğer,

$$\forall t \in \mathbb{N}^*, \forall A \in \mathcal{E}, \mathbb{P}(X_t \in A / X_{t-1}, \dots, X_0) = \mathbb{P}(X_t \in \mathbb{N} / X_{t-1}) \qquad (3.3)$$

sağlanırsa.$\{X_t\}$'ye Markov zinciri denir. Eğer,$\mathbb{P}(X_t = i / X_{t-1} = j), t$'den bağımsızsa Markov zincirine zamansal-homojendir denir.

Örnek 3.2.1 Epidomiyolojide bir örnek olarak varsayalım ki, enfekte olmuş her birey iyileşmeden ya da tedavi edilmeden önce tüm hassas bireylerle temas şansına sahip olsun. Bu durumda enfekte olmuş ve hassas bireylerin sayısı Markov zinciriyle modellenebilir.

Tanım 3.2.2 ölçülebilir bir$(E, \mathcal{E})$ uzayı üzerindeki bir Markov zinciri için

a) Bir σ –olasılık ölçüsü ϕ 'ye Markov zinciri için indirgenemez ölçü denir ancak ve ancak

$$\forall (x, A) \in E \times \mathcal{E} \phi(A) > 0 \text{ için } \exists n \in \mathbb{N}^* \text{ öyleki,}$$

$$\phi(X_n \in A / X_0 = x) > 0$$

Diğer bir deyişle, sadece bir tane iletişim sınıfı varsa tüm durumlar birbiriyle ilişkilidir.

b) Bütün durumları periyodik olmayan yani tüm periyotları 1'e eşit olan veya $\forall x \in E, d(x) = 1$ sağlayan Markov zincirine periyodik olmayan, aksi takdirde

yani $d(x) = g.c.d\{m \geq 1, \mathbb{P}^{(m)}(x,x) \geq 0\}$ olduğu durumda ise Markov zincirine periyodiktir denir.

Yani A'nın periyodu, bu sınıfın durumunun tüm tekrar sayılarının en büyük ortak bölenidir. Markov zincirinin dağılımı içindeki yakınsaklığı garanti etmek için periyodik olmama gerekli koşuldur.

Bir Markov zinciri, sabit dağılımı $\pi(.)$ varsayarak periyodik olmayan ve π- indirgenemez ise $\{X_t, t \in \mathbb{N}\}$ zinciri bir ergodiktir. Ergodiklik kavramı Markov zincirinin başlangıç noktası X_0 ne olursa olsun hedef dağılıma ulaşabilmeyi herzaman garanti eder.

Unutulmamalıdır ki, MCMC metodu ilgili dağılıma kernel $\mathbb{P}$ yakınsaması geçişiyle bir $\{X_t, t \in \mathbb{N}\}$ Markov zincirinin simüle edilmesine açıkca izin verir. MCMC algoritmaları enstrümantal dağılımının simülasyonu için, kabul-ret metodu ve ağırlıklı örnekleme olarak bir çekicilik oluşturur. Bu enstrümantal dağılım bir geçiş kernel ya da bir şartlı dağılım ile karakterize edilebilir.

3.2.2 MCMC Metodunun Verimliliği ve Yakınsaklık Hızı

MCMC yönteminde, örnekleyici verimleri simülasyon uzay boyutunda kovaryans yapısı, hedef dağılımı ve multimodalitesinin arasında değişir. Unutulmamalıdır ki, bazı Markov zincirleri diğerlerinden yakınsaklık hızı, asimptotik varyans veya karıştırılmış hız açısından daha iyi çalıştığından MCMC yöntemiyle daha verimli zincirler bulunabilir. Bazı kompleks problem simülasyonu ile başa çıkmak için MCMC algoritmasının uygulaması olarak adlandırılan iyi parametreleri seçmek zorundayız: verilen bir sabit ölçü ile Markov zincirinin tekliği yoktur ve MCMC kullanıcısı verilen ailede daha iyi çekirdek geçişiyle yüzleşir, her MCMC algoritması uygun iterasyon sahiptir. Unutulmamalıdır ki, bir çekirde krasgele mekanizmaya gore seçilmiştir. Bu algoritmalar Markov zincirinin gerçekleşmesi üzerine çalışmaz, onların davranış çalışması sadece Markov zincir teorisinin sonuçlarını rapor etmez ve birçok yönden bu algoritmaların anlaşılması hala açık bir problemdir. Sonuç olarak son on yıl içinde MCMC algoritmaları adaptif olarak adlandırılan süreçlere evrildi.

Adaptif MCMC algoritmaları zincirin son iterasyonundan bilgiler kullanarak Markov zincirini anında geliştirmeye çalışır. Bu, pratikte verimliliği anlamlı bir şekilde geliştirir

[97, 98, 99, 101, 103 ve 104]. Ne yazık ki çoğu adaptif MCMC algoritması artık Markovian değildir, bu nedenle algoritmanın hedef dağılıma yakınsaması çok daha zordur ve bazen durumlarda hata verir [102]. Bu soru geniş bir şekilde araştırılmaktadır ve araştırmacılar çeşitli şartlar altında adaptif MCMC algoritmalarının ergodisini ispatlamıştır [93, 94, 96, 97, 98, 99, 100 ve 104]. Ancak bu sonuçlar, Pratik uygulamalarda doğrudan doğrulamanın sanal olarak imkânsız olduğu "çevreleme" ya da "eş zamanlı polinom sürüklenme koşulları" gibi tüm koşulları gerektirir ve kolayca sınanabilir varsayımları olmayan geniş uygulanabilir yakınsaklık sonuçları yoktur. Bunun anlamı, pratikte adaptif MCMC algoritmaları kullanıldığında asimptotik yakınsaklığın bile genellikle garantisi yoktur ve kullanıcı basitçe [102]'de verilen anormal davranışın ortaya çıkmamasını umut etmelidir.

Ayrıca, MCMC çıktılarının bazı istatistiksel testleri, yakınsaklık için daha yüksek güven seviyesi elde edilerek geliştirildi. En yaygın olanları Gelman-Rubin testi [105], Geweke testi [106] ve Raftery-Lewis testidir [107].

Gelman-Rubin testleri, zincirler arasında ve her bir zincirin kendi içinde varyansını karşılaştırarak, çoklu simüle edilmiş MCMC zincirlerinin analizine dayanır. Buradaki fikir, parametre uzayındaki çoklu-modellemeyi bulmaya çalışmaktır. Durum uzayında, farklı yerler için zincirlerle başlayarak, yerel maksimumu bulma ve orada kalma mümkündür. Ancak birkaç zincir dikkate alınarak yakınsaklık vurgulanabilir.

Geweke testleri yakınsaklık hatasını belirlemek için Markov zincirinin ilk ve son parçaları arasındaki değerleri karşılaştırır. Bu test zincirin değişmezliğini yönetmeye imkân tanır. Bu gözlemler, $n = N_a + N_b$ iterasyon sayısı olmak üzere N_a ilk gözlemlerin ortalaması ile N_b sonraki gözlemlerin ortalaması arasındaki sapmaya dayanmaktadır.

Raftery-Lewis yaklaşımı ise, uygun inceleme aralığı ve atılacak yazılmış örneklerin sayısı ile birlikte her yakınsama için gereken iterasyon sayısını tahmin eder. Her iki miktarın ayrı bir tahmini verilen modelde her bir değişken için elde edilebilir. Raftery-Lewis yaklaşımı yakınsaklığı belirlemek için kriter olarak kullanıcının özelleştirilmiş dağılım tahmininin doğruluğunu kullanır.

3.3 Metropolis-Hastings Algoritmaları

MCMC metodlarının herhangi bir ailesinden daha genel bir metod, hedef yoğunluğu üzerindeki koşulları azaltması sebebiyle tartışmasız olarak Metropolis-Hasting Algoritmasıdır. Bu algoritma ilk olarak Metropolis ve Al (1953) tarafından yayınlandı ve 1970'de Hasting genellendi. Verilen π hedef yoğunluğu, simüle etmesi kolay olan $q(y/x)$ensturmental koşullu yoğunluğu ile ilgilidir. Ek olarak π 'nin tüm desteklerinin bir keşfine yol açan dağılıma sahip $q(\cdot/x)$ ve bir bağımsız sabit x'e kadar bilinen $\{\pi(y)|q(y/x)\}$ oranı şeklinde sadece teorik gerektirmelerde q neredeyse rasgele olabilir. Muhtemel vektör olan$x^{(0)}$ ile başlayalım. Metropolis-Hasting Algoritması π hedef yoğunluğu ve q koşullu yoğunlukla ilgili olarak aşağıdaki geçiş çekirdegi ile bir $\{X^{(t)}\}$ Markov zinciri üretir [108].

Algoritma B: Metropolis-Hasting

Verilen $x^{(t)}$ için

Oluştur $Y_t \sim q(y/x^{(t)})$.

Hesapla,$\rho(x,y) = min\left\{\frac{\pi(y)}{\pi(x)}.\frac{q(x/y)}{q(y/x)}, 1\right\}$ olmak üzere

$$X^{(t+1)} = \begin{cases} Y_t\ ,\ \rho(x^{(t)}, Y_t)\ olasılığı\ ile \\ x_t,\ 1-\rho(x^{(t)}, Y_t) olasılığı\ ile' \end{cases}$$

Bu adımları 0'dan n'ye kadar tekrar ederek muhtemelen h_n ile birlikte $E_f[h(x)]$ 'I tahmin edecek örnek üretiriz. Ancak bu durumun geçerliliğinin ikna edici olduğunu söylemek için ilk olarak hedef yoğunluğa yakınsaklığın ve sabit dağılımın varlığını bize garanti eden özelliğe sahip algoritma ile üretilen zinciri doğrulamak gerekir. Seçilmiş $q$dağılım tippine göre Metropolis-Hasting Algoritmasının, alt bazı kategorileri vardır. İlk olarak x_t bağımsız olmalıdır yani $q(x,y) = q(y)$. Bu versiyonun zorluğu uzay içerisinde işlem yapabilecek bir eğrinin bulunmasıdır. Ayrıca global olan yoğunluğu arıyoruz, özellikle büyük boyutlar olması durumunda genellikle hedef dağılım hakkında olan bilgimiz ve komplekslilik ile bloke ediliriz. Bu nedenle desteğin işlenmesi için daha çok yerel yaklaşıma dönülmesi sıkça tercih edilebilir.

Bu amaç için, rasgele yürüyen bir Markov zinciri oluşturmak adına yani, q enstürmental yoğunluğu, zincirin mevcut değeri üzerinde merkez olacak ve simetrik olacak şekilde algoritma üretiriz. ε_t, q'ya göre dağıtılmış rasgele pertürbasyon olmak üzere ve $x^{(t)}$ bağımsız olmak üzere amaçlanan değer Y_t, $x^{(t)} + \varepsilon_t$'den elde edilir. Bir başka deyişle q, $q\left(\left|y - x^{(t)}\right|\right)$ formunda olacaktır.

Prensip: Metropolis-Hastings algoritması, bir enstürmantal dağılıma göre üretilen Y rasgele değişkenlerin bir kabul-ret mekanizması tarafından kullanılır. Bu bölümde Metropolis-Hastings algoritmasının genellikle üç çeşidini kullanacağız:

3.3.1 Bağımsız Metropolis-Hasting Algoritması

Genel Metropolis-Hastings algoritmasının bu özel durumu $X^{(t)}$'nin bir bağımsız öneri dağılımını kullanır. Bir başka deyişle, bu özel durum, üretilen y adayı $x^{(t)}$'nin mevcut durumunun bağımsız olmak üzere$q(y)$ ile de gösterilen $q(x^{(t)}, y)$ enstürmantal yoğunluğu kullanmaya dayanır. Markov zinciri aşağıdaki algoritmayla oluşturulmuştur:

Algorithm C: Bağımsız Metropolis-Hastings algoritması

Verilen $x^{(t)}$ için

$Y_t \sim q(y/x^{(t)})$=g(y) üret.

Hesapla, $X_{t+1} = \begin{cases} Y_t : min\left\{1, \frac{\pi(Y_t)g(x^{(t)})}{\pi(x^{(t)})g(Y_t)}\right\} olasılığı\ ile \\ x^{(t)}\ diğer\ durumlarda. \end{cases}$

Bu durumda kabul olasılığının hedef ve önerilen dağılım arasındaki oran olarak $\rho(x, y) = \min\{\pi(y)/q(y) \,|\, \pi(x)/q(x)\}$ (aslında oranların oranı) şeklinde yazılabildiği gözlemlenir. Bu Kabul-ret örnekleme metoduna benzerdir, fakat aday reddedilmişse ve eğer önerilen dağılım her $x \in Supp(\pi)$ için $\pi(x) \leq C.q(x)$ eşitsizliğini sağlarsa mevcut değer tutulması ile yani $X^{(t+1)} = X^{(t)}$ ile algoritma farklıdır. Dolayısıyla q, $Supp(\pi)$ üzerinde kesin pozitif ve $Supp(\pi)$'de $Supp(q)$ tarafından içerilmelidir. Ek olarak bu algoritmayı optimize etmek istersek üretilmiş adayın kabul oranını arttırmak için q yoğunluğunu π'ye yakın seçmemiz gerekir.

3.3.2 Rasgele yürüyen Metropolis-Hasting Algoritması

Doğal olarak, Metropolis-Hastings'in Pratik inşası için yaklaşım önerisi, Markov zincirinin mevcut değerinin komşuluğunun yerel işletilmesini gözönüne almak, bir önceki simüle edilmiş değeri hesaba katarak bir sonraki değeri hesaplamasıdır. Burada rastgele yürüyen diziyi göz önüne alırız, dolayısıyla$Y_t = X^{(t)} + \varepsilon_t, \quad \varepsilon_t \sim g(.)$

şeklinde, Y_t,$X^{(t+1)}$'in önceki $X^{(t)}$etrafında bir pertürbasyon ile tipik olarak üretilmesidir.

Burada $E(\varepsilon_t) = 0$'dır. $x^{(t)}$ mevcut durumunun bağımsız yoğunluğu q olmak üzere $\left(x^{(t)}, y\right) = q(y - x^{(t)})$ formunda alınan enstürmantal yoğunluğun kullanılmasına dayalı algoritmayı göstermek istiyoruz. Yani aday y, $y = x^{(t)} + \varepsilon_t$ ile üretilir. Burada ε_t rastgele pertübasyon ya da$X^{(t)}$'nin bağımsız q(.) dağılımının rasgele yürüyüş artırımıdır, örneğin düzgün dağılımın ya da normal dağılımın tek boyutlu yapılardaki anlamı $Y_t \sim \mathcal{U}(X^{(t)} - \delta,\ X^{(t)} + \delta$ veya $Y_t \sim \mathcal{N}(X^{(t)}, \tau^2)$ dir.

Enstürmantal q(.) yoğunlukları genellikle orjinde merkezlenmiştir ve düzgün ya da normal yoğunluklar olarak kolaylıkla simüle edilebilir. Unutulmamalıdır ki, eğer q genellikle düşünüldüğü gibi simetrik ise, $q(\varepsilon_t) = q(-\varepsilon_t)$'tir ve dolayısıyla algoritma aşağıdaki gibi verilebilir:

Algorithm. D: Rastgele yürüyen Metropolis-Hastings Algoritması

Verilen $x^{(t)}$

1. $Y_t \sim q(y/x^{(t)})$=g(*y*-$x^{(t)}$) üret.
2. Hesapla,$X_{t+1} = \begin{cases} Y_t, & min\left\{1, \frac{\pi(Y_t)}{\pi(x^{(t)})}\right\}\ olasılığı\ ile \\ x^{(t)}, & diğer\ durumlarda. \end{cases}$

3.3.3 Her seferinde bir değişken Metropolis-Hastings algoritması

Verinin boyutu büyük olduğunda, hedef yoğunluğu π'ye yaklaşım sağlayan q çok boyutlu yoğunluğunu bulmak zor olur. $d \geq 1$ olacak şekilde veri seçilirse ideal yaklaşım elde edilir. Bu durumda, her şey hedef dağılım şekline bağlı olur. Veriyi $\left(x_1^{(t)}, \ldots, x_d^{(t)}\right)$ şeklinde d bileşene ayırmak avantajlı olabilir. Böylece, her $x^{(t)}$ bileşeni başarılı şekilde

güncellendiğinde yeni veri üretilir, buradan da q_i d boyutlu yoğunluğuna ulaşılabilir. $x^{(t+1)}$ yeni vektörü elde etmek için, Metropolis-Hastings algoritmalarının her biri, farklıπ yoğunluğunun koşullu dağılımları olan bir Markov zinciri gösterir. Algoritma aşağıdaki gibi işler:

Algoritma E: her seferinde bir değişken (one variable at time)

1.y , $x_i^{(t)}$'nin boyutu ($x^{(t)}$ nin boyutu değil) olmak üzere $y \sim q(x^{(t)}, .)$ simülasyonu

$x_y^{(t)}$, y ile yer değiştirmiş $x_i^{(t)}$ yerini göstermek üzere

2.Hastings oranını hesaplama

$$R = \frac{\pi(x_y^{(t)}) q_i(x_y^{(t)}, x_i^{(t)}))}{\pi(x^{(t)}) q_i(x^{(t)}, y))}$$

3.(Metropolis red) $x_y^{(t)}$ 'ye geç ($min\{1, R\}$ olasılığı ile)

.Diğer durumda $x^{(t)}$'de kal

Etkili bir örnekleyici oluşturmak için, bu temel ve tüm değişkenleri değiştirecek güncel adımları birleştirmeliyiz. Bu güncelleme, uygulanabilir bir konumda başlamışsa yine uygulanabilir konumda kalır.

Bir MCMC yöntemi, önceki yöntemlere göre esnekliğinin yanında önemli iki avantaja sahiptir. İlki, enstrümantal yoğunluğunda kısıtlama olmaması, ikincisi ise uygulaması süresince enstrümantal fonksiyonun seçiminde kullanıcıya çok fazla esneklik vermesidir. Üstelik çeşitli adımlar için zincir aynı konumda kaldığından, red için örnekleme yönteminde farkettiğimiz adayların kaybı minimum olur. Bu işlem, enstrümantal fonksiyon hedef yoğunluğa çok yakın olduğunda zaman açısından önemli kazanımlara sağlar. Bu algoritmanın, aynı zamanda dezavantajları da vardır: üretilmiş örnek bağımsız değildir, dolayısıyla zincirin bekleme periyodunu belirlememiz gerekir ve hesaplama süresi, olayların uzun veya kısa olmalarına bağlıdır.

3.4 Gibbs Örnekleme Algoritması

Gibbs örnekleme algoritması, bu işler için uygun MCMC algoritmalarının en basitlerinden biridir. Gibbs örneklemede fikir, kalan değişkenlerin son değerlerinde sabitlenerek her bir değişkenin kendi koşullu dağılımıyla posterior örneklerin

üretilmesine dayanır. Bu teknik, önceki bölümde gösterilen Metropolis-Hastings algoritmasının özel durumudur ve ilginçtir. Ancak, q enstrümantal dağılım ve π hedef dağılım arasındaki bağlantının olmaması yöntemin yakınsamasını olumsuz etkileyebilir. Gibbs örneklemin kullanışlılığı, gerçekte π dağılımına dayanmasındandır. Bu yöntem, Gelman (1984) tarafından görüntü resterasyonunda tanıtılmış ve Tanner ve Wong (1987) ve Gelfand (1990) tarafından istatistik uygulamalar için geliştirilmiştir. Bu algoritma $x \in \mathbb{R}^d$ olmak üzere $\pi(x)$ ortak dağılımının simüle edilmesinde kullanılmış ve $\pi(x_i/x_1, x_2, \dots, x_{i-1}, x_{1+1}, \dots, x_d)$ koşullu dağılımlarının kolayca simüle edilebildiği görülmüştür.

Örneğin, $X_1, X_2\ ve\ X_3$ rastgele değişkenlerini ele alalım ve bu değişkenlerin $x_1^{(0)}, x_2^{(0)} ve\ x_3^{(0)}$ (genellikle q dağılımından örneklenmiş değerler) başlangıç değerleriyle işleme başlayalım. i iterasyonunda, $x_1^{(i)} \sim \pi(X_1 = x_1/X_2 = x_2^{(i-1)}, X_3 = x_3^{(i-1)})$, $x_2 \sim \pi(X_2 = x_2/X_1 = x_1^{(i)}, X_3 = x_3^{(i-1)})$ ve $x_3 \sim \pi(X_3 = x_3/X_1 = x_1^{(i)}, X_2 = x_2^{(i)})$ örneklerini oluştururuz. Bu işlem, yakınsaklık sağlanana kadar devam eder (örneklem değerleri, sanki gerçek posterior dağılımdan elde edilmiş gibi aynı dağılıma sahiptirler). Gibbs örnekleyicisinin detayları algoritma E ile verileecektir;

Algoritma F: Gibbs örnekleyici

Başla $x^{(0)} \sim q(x)$

İterasyon için: $i = 1,2, \dots, do$

$$x_1^{(i)} \sim \pi(X_1 = x_1/X_2 = x_2^{(i-1)}, X_3 = x_3^{(i-1)}, \dots X_d = x_d^{(i-1)})$$

$$x_2^{(i)} \sim \pi(X_2 = x_2/X_1 = x_1^{(i)}, X_3 = x_3^{(i-1)}, \dots X_d = x_d^{(i-1)})$$

$$\vdots$$

$$x_d^{(i)} \sim \pi(X_d = x_d/X_1 = x_1^{(i)}, X_2 = x_2^{(i)}, \dots X_{d-1} = x_{d-1}^{(i)})$$

$$end\ for$$

Üretilen zincirin $\pi(x)$ ortak dağılıma yaklaştığını görürüz. Aksi halde, Gibbs örnekleyici, simülasyon anlamında dağılım hakkında daha gelişmiş bilgiler verir. Bu durumda aşağıdaki koşullu dağılımı simüle edebiliriz.

Burada verilen algoritma. E rastgele değerlerle başlar ve bu algoritma ile simüle edilmiş örneklemin evvelki iterasyonları, gerçek posterior dağılımı gösterme gerekliliği yoktur.

3.5 Hibrid Gibbs Örnekleme

Gibbs örnekleyicisinin yakınsaması yavaş olabilir, çünkü iterasyonla tek bir bileşen simüle edilir. Aslında, çoğu problemler için, X'in bir veyadaha çok elementlerinin koşullu dağılımları kapalı formda kullanışlı değildir. Bu durumda, hybrid MCMC algoritması zinciri zamanın bir alt kümesi olarak uygun koşullu dağılımlardan örneklem kullanmak üzere direct çizim (Gibbs adımı) veya Metropolis Hashting (Metropolis adımı) ile günceller. Böylece, bir hibrid algoritma Gibbs örnekleyici ile M-H algoritmasını farklı modeller için MCMC algoritmalarında önemli esneklik kazandırmak için birleştirir. Örneğin, d=3 için bir hybrid algoritması aşağıdaki güncelleştirmeler dizisini takip edebilir [109].

1) Bu koşullu dağılım kapalı formda kullanışlı olduğundan $X_1^{(t+1)}/(x_2^{(t)}, x_3^{(t)})$ Gibbs adımı ile güncelle
2) Bu koşullu dağılım kapalı formda kullanışlı olmadığından veya kapalı formda örneklem oluşturmak zor olduğundan $(X_2^{(t+1)}, X_3^{(t+1)})/(x_1^{(t+1)})$ M-H algoritması ile güncelle. Burada $X_2\ ve\ X_3$'e engel olmak, bu elemanlar yüksek korelasyonlu olduğundan, tavsiye edilebilir.
3) Bu koşullu dağılım kapalı formda kullanışlı olmadığından $X_2^{(t+1)}/(x_1^{(t+1)}, x_3^{(t)})$ zincirin rastgele yürüyüşü adımı ile güncelle
4) $X_3^{(t+1)}/(x_1^{(t+1)}, x_2^{(t+1)})$ Gibbs adımı ile güncelle

Daha yüksek boyutlar durumunda, hybrid algoritması Gibbs ve Metropolis adımlarının karışımlarını içerir. Hibrid algoritması, MCMC yöntemlerinin uygulanabilirliğini, sadece hesaplanabilen veya örneklenebilen posterior koşulların ürettiği model gerektirdiğinden, önemli ölçüde arttırmaktadır. Üstelik, bölüm 3.3.1'de görülen g yoğunluğunu düşünürsek, π g'nin marjinal dağılımı olduğu da dikkate alırsak, g'nin koşulu dağılımlarını simüle etmek π'nınkileri simüle etmekten bazen daha kolaydır.

3.6 Tersinir Atlama Markov Zincirli Monte Carlo (RJMCMC)

Tersinir sıçrama, Green [91], tarafından tanıtılmış, sonra Green ve Richardon [34] ve Ehlers ve Brooks [110] gibi araştırmacılar tarafından geliştirilmiştir. Markov zinciri

iterasyonları arasında parametre uzayının boyutu değişebiliyorsa MCMC simülasyonun genel çerçevesi sağlar. RJMCMC M-H algoritmasının daha genel konum uzayı üzerine genişlemesi olarak da görülebilir. Algoritmalar, Bayesyen modellerin istatistiksel yapılarının en iyi kullanışlarını mümkün kılarlar, güçlüdürler ve hesaplama bakımından etkindirler. Yapay ve gerçek veri kümeleri kullanılarak çok çeşitli problemler üzerinde iyi performans gösterirler. RJMCMC yöntemlerinin amacı, gözlenen veriyi daha iyi açıklayan ilgili modellerin posterior dağılımını tahmin etmek ve parametre tahminlerini ağırlıklandırmaktır.

Bu modelin uygulaması için, x gözlenmiş veri için, $k \in \mathcal{K}$. Parametresi ile indeksli $(\mathcal{M}_k)k \in \mathbb{N}^*$ aday modelin sayılabilir sonlu birleşiminin mevcut olduğu kabul edilir. k indeksi, değişkeni gösteren bir yardımcı mod olarak düşünülebilir. Öyle ki, $\mathcal{M}_{k'}k = j$ olduğu yerdeki modeli gösterir. Her bir $\mathcal{M}_k$ modeli, $\theta^{(k)} \in \Theta^{n_k} \subset \mathbb{R}^{n_k}$ bilinmeyen parametrelerin n_k boyutlu vektörüne sahiptir. Burada n_k, $k \in \mathcal{K}$ farklı modeller için farklı değerler alabilir. Verilen gözlenmiş veri $x\left(k, \theta^{(k)}\right)$'nın ortak posterior dağılımı, $L(x/k, \theta^{(k)})$ olabilirlik ile $\mathcal{M}_k$ modeli altında $\theta^{(k)}$'nın prior dağılımından kurulmuş olan $\mathrm{p}(k, \theta^{(k)}) \propto \mathrm{p}(\theta^{(k)}/k)\mathrm{p}(k)$ ortak prior dağılımının çarpımı olarak elde edilebilir. Buradan, ortak posterior dağılım:

$$\pi(k, \theta^{(k)}/x) = \frac{L(x/k,\theta^{(k)})\mathrm{p}(\theta^{(k)}/k)\mathrm{p}(k)}{\sum_{j\in\mathcal{K}}\int_{\theta^{(k)}\in\Theta^{n_j}} L(x/j,\theta^{(j)'})\mathrm{p}(\theta^{(j)'}/j)\mathrm{p}(j)d\theta^{(j)'}} \qquad (3.4)$$

olur. Tersinir sıçrama algoritması (3.4) ortak posterior dağılımını, Markov zincirinin $k \in \mathcal{K}$ nın s konumu, konum uzayı üzerinde değişebilen boyutlarda $\left(k, \theta^{(k)}\right)$ formunda olduklarından $\Theta = \bigcup_{k\in\mathcal{K}}(\{k\} \times \mathbb{R}^{n_k})$ konum uzayı üzerinde MCMC örnekleyicisinin hedef dağılımı olarak kullanır. Bu algoritma, S uzayında nokta sürecinin dağılımı gibi olan $\pi(.)$ ele alındığında nokta süreçlerinin modellenmesinde kullanılabilir. S uzayı doğum ve ölüm modellerini simüle etmeye izin veren, $\nu(.)$ yoğunluklu $\mu(.)$ Poisson dağılımı ve bir yoğunluk olarak tanımlanır.

3.7 Sonuç

Nokta süreçlerini simüle etmeye yarayan MCMC yöntemleriyle bazı örneklem algoritmalarını tanıtan posterior örneklemin teorik açıklamalarıyla bu bölümü

bitiriyoruz. Bu modellerinkoşullarının etkinliği ve yakınsaklık hızları Marko zinciri ve Gibbs süreçleri kanalıyla incelenmiş ve bu algoritmalar arasındaki etkileşim bu süreçlerde dağılımları ve parametreleri hakkında prior bilgi olarak tanıtılmıştır. Markov zincirinin yapısı, nokta süreci simülasyonları için anlamlıdır ve bu gösterilmiştir. Green ve RJMCMC algoritmalarına göre esnekliği nedeniyle Metropolis-Hastings bu tezin ilerleyen bölümlerinde kullanılacaktır.

Sonuç olarak, literatürdeki bazı güncel modeller uygulamalarıyla birlikte gösterilmişlerdir. Bu çalışmalar, çeşitli epidemiyolojik olaylarda nokta süreçleri kanalıyla avantajlar sağlarlar. Bu bölüm, bu tezde tanıtılan parametre tahmini ve nokta süreçlerinin karışımı üzerinde yoğunlaşacak olan bir sonraki bölüm için iyi bir hazırlık olmuştur.

BÖLÜM 4

ZAMAN-MEKÂNSAL NOKTA SÜREÇLERİ

4. Giriş

Bir Zaman-Mekânsal nokta süreci, ele alınan olayın mekânını ve zamanını gösteren rastgele noktalar topluluğudur. Bu bölümde, Shot-Noise yoğunluk ölçümünün genelleşmiş hali olan zaman-mekânsal Cox süreçlerini kurmak için çalışılan bölgedeki olayları ele alacağız. [27, 28, 29 ve 31]. Otokorelasyon gösteren ve yoğunluğunun zaman ve/veya mekânsal olarak değiştiği pek çok olay vardır [9]. Biz de zaman ve/veya mekânda oluşan doğal olaylarla ilgileneceğiz. Uygulamada, parametrik olmayan Bayesian analiz çerçevesinde, süreçlerin yoğunluğu, sınırlı çekirdekli karışık Dirichlet süreçleri şeklinde parametrik olmayan süreç olarak modellenir.

Dirichlet süreci, kendi olasılık dağılım kümesi üzerinde tanımlı bir olasılık dağılımıdır [114]. Başka bir deyişle, bu süreç olasılık dağılım kümeleri üzerinde tanımlı bir dağılımdır. Bu genellikle, rastgele değişkenlerin dağılımı hakkında ilk bilgileri tanımlamak için Bayesian çıkarsamada kullanılır. Öyle ki, rastgele değişkenler diğer özel dağılımlara göre dağılırlar. Bu, bilinmeyen olasılık dağılımının Bayesian tahmin çerçevesinde tanımını yapmayı sağlar [30, 111, 112 ve 113]. Simülasyon için, MCMC kullanarak yaklaşık çıkarsama yapmak yaygındır. Bu hesaplamalar, örneklemlerden Monte Carlo tahminleri ile yapılır. Bu örneklemler sonsal (posterior) dağılımı invaryantlarını koruyan Markov zinciri simülasyonu ile elde edilirler. Belirtmek gerekir ki, MCMC hareketleri, çoğu parametrik olmayan Bayesian modeller için sıradan

bilgisayarlarla simüle edilebilir. Saklı işlemler (görünmeyen, gözden kaçan olaylar) ile ilgili olayların yoğunluğunun zaman-mekânsal nokta süreci ilginçtir.

Gerçekte, zaman-mekânsal nokta süreci olayı, sayılabilir $\{(s_i, t_i)_{i\in\mathbb{N}^*}\}$ noktalar kümesinden oluşur. Burada $s_i \in \mathbb{R}^2 i.$ olayın yerini gösterirken $t_i \in T \subset \mathbb{R}^+ i.$ olayın zamanını verir.

4.1 Zaman-Mekânsal İlişkili Nokta süreçleri

Uygulamada, analiz için kullanılabilir veri, gözlenen $S \times T$ sonlu ve kısıtlı zaman-mekansal bölgenin oluşturduğu $(s_i,\ t_i),\ i = 1, \dots, n$ noktalardır. Burada S tipik bir poligon, T bir kapalı küme, n ise Dmekânsal bölgesindeki olaylar sayısıdır ($D \subseteq S$). Zaman-mekânsal nokta süreci davranışı, tipik olarak gözlem yapılan zaman ve yerde verilen olayın x yeri ve t zamannda olayın yoğunluğunun beklenen değerini gösteren $\lambda(t, x)$ yapısında bir fonksiyonla modellenir.

Bu konuda çalışan bazı yazarlar, çalışmalarını tek değişkenli (zamanı dikkate almayan) zaman-mekânsal nokta süreçleri ile yapmışlardır [115]. Bir zaman-mekânsal nokta süreci, bir ve iki boyutlu değişkenlerle ifade edilir:

Eğer $\lambda(s,t) = \lambda(t)$ ve $\lambda_2\big((s,t),(s',t)\big) = \lambda_2(s - s', t)$ ise mekânda

Eğer $\lambda(s,t) = \lambda(s)$ ve $\lambda_2\big((s,t),(s,t')\big) = \lambda_2(s, t - t')$ ise zamanda.

Eğer f $\lambda(s,t) = \lambda$ ve $\lambda_2\big((s,t),(s',t')\big) = \lambda_2(s - s', t - t')$ ise hem zamanda hem de mekânda.Eğer, (u, v) uzay-mekânsal vektörlerin farkı $u = \|s - s'\|,\ v = |t - t'|$ olmak üzere $\lambda_2\big((s,t),(s',t')\big) = \lambda_2(u, v)$ ise değişkenin birinin sabit olduğu zaman-mekânsal nokta süreci aynı zamanda izotropiktir.

Bir zaman-mekânsal nokta süreci, ağırlıklı yoğunluğa sahip iki değişkenlidir ve yoğunluk fonksiyonu sınırlı ise izotropiktir. Aynı zamanda, korelasyon fonksiyon çiftleri sadece zaman-mekansal (u, v) vektörlerin farkına bağlıdır. İki değişkenli ağırlıklı yoğunluk sadece mekânsal nokta süreçleri için tanımlanmıştır [116]. Gabriel ve Diggle [54] zaman-mekânsal durum için düzgün açılımı verdiler. Dolayısıyla, burada zaman-mekânsal nokta sürecinin yoğunluğunun çarpanlarına ayrılabildiği durum ele alınacaktır.

Bir zaman-mekânsal nokta süreci, eğer yoğunluğu herhangi $(s,t) \in S \times T$ için $\lambda(s,t) = m(s)\mu(t)$şeklinde çarpanlarına ayrılabiliyorsa birinci dereceden ayrılabilirdir. Sabit bir zaman-mekânsal nokta süreci, eğer $\gamma(u,v) = \lambda_2(u,v) - \lambda^2$ yoğunluğu $\gamma(u,v) = \gamma_s(u)\gamma_t(v)$şeklinde çarpanlarına ayrılabiliyorsa ikinci derece ayrılabilirdir.

Genelde, ikinci derece ayrılabilirlik, zaman ve mekân bileşenlerinin bağımsızlığını ifade eder. Üstelik Poisson süreci, yalnız ve yalnız birinci dereceden ayrılabilir ise bağımsız bileşenlere sahiptir.

$\mathcal{B}_d, d$ uzaklığı gösteren Borel kümesi olmak üzere $(\chi, \mathcal{B}_d, v)$ ölçüm uzayını ele alalım. $A \in \mathcal{B}_d$ve$Y$ gözlenen olay ile ilgili süreç olsun. Aşağıda $Y(A)$ ile A keyfi bölgesindeki olay sayısını gösterecektir. (Y gözlenen dağılımdır.) Bundan sonra aşağıdaki ifadeyi ele alacağız [44]:

$$\lambda(x,t) = \sum_{i:a_i,x_i,t_i} a_i \phi(d((x,t),(x_i,t_i))) \quad (4.1)$$

Burada, ϕ zaman-mekânsal çekirdeği, (a_i, x_i, t_i) ise $\mathcal{N}$ değerinin $(\chi \times \mathbb{R}_+^* \times \mathbb{R}_+)$ den olduğu bir Poisson sürecidir. Poisson sürecinin saklı süreç durumunda μ yoğunluklu kararlı olduğu kabul edilir. $x_i i$.Saklı olayın uzaydaki konumunu, $t_i i$.Saklı olayın zamanını , a_i ise i.Saklı olayın katkısını gösterir. $a_i \sim \Gamma(b,1)$ nin Poisson dağılımının eşlenik dağılımı olduğunu vurgulamak gerekir. Zamanın gözlenen yoğunluğu

$$\lambda(.,t) = \int_\chi \lambda(x,t) v(dx) \quad (4.2)$$

Şeklinde tanımlanır. Burada λ gözlenen sürecin zaman-mekânsal yoğunluğudur.

4.2 Dağılım Özellikleri

Y, λ Poisson yoğunluklu mümkün yolların Cox sürecini göstermek üzere $Y(B \times [0,t]), B \times [0,t]$ de gözlenen olayların sayısı olsun [19]. Bu durumda, $(\chi, \mathcal{B}, v)$ ölçüm uzayında sürecin koşullu yoğunluğu, rastgele $\Lambda(.)$ ölçümünden elde edilen rastgele bir alandır. Süreç $Y(B \times [0,t]) \sim Pois(\int_B [\int_0^t \lambda(x,u) v(dx)] du)$Poisson dağılımı ile işler, sonra aşağıdaki iki merkezi koşulu tanımlar:

$$E[Y(B \times [0,t])/\lambda] = \int_B [\int_0^t \lambda(x,u) v(dx)] du \quad (4.3)$$

$$Var[Y(B \times [0,t])/\lambda] = \int_B [\int_0^t \lambda(x,u) v(dx)] du \quad (4.4)$$

Bu eşit 2'li merkezi momentli homojen olmayan Poisson süreci ile aynıdır. Beklenen değerin iterasyonu ile koşullu olmayan momentler aşağıdaki gibi belirlenir:

$$E[Y(B \times [0,t])] = E\left\{\int_B [\int_0^t \lambda(x,u)v(dx)]du\right\} \quad (4.5)$$

$$Var[Y(B \times [0,t])] = E\{Var[Y(B \times [0,t])/\lambda]\} + Var\{E[Y(B \times [0,t])/\lambda]\} \quad (4.6)$$

(4.5) ifadesini (4.6) da kullanırsak

$$Var[Y(B \times [0,t])] = E\left\{\int_B . \int_0^t \lambda(x,u)v(dx)du\right\} + Var\left\{\int_B . \int_0^t \lambda(x,u)v(dx)du\right\} (4.7)$$

elde edilir. Zaman ve mekânda yerel olaylar çalışılırken, farklı formlar gözlemlenebilir [9 ve 10]. Rastgele ölçüm $\Lambda = \lambda$ olarak verilirse Y homojen olmayan Poisson süreci ile ilgili olur. Yukarıdaki özellikler bu durumda da geçerlidir.

4.3 Zaman-Mekânsal Nokta Süreçleri için Bayesyen Yaklaşım

Öncelikle farklı gözlem tipleri için zaman-mekânsal nokta süreçlerinin istatistik analiz yöntemlerinden bahsedeceğiz. Uzay ve zamanda meydana gelen olaylar olduğunda farklı formlarda gözlemlere sahip olabiliriz [9 ve 10]. Bazı olayları tanımlamak için, gözlenen verinin zamansal analizine ihtiyaç duyarız. Zamansal analiz, modeller, trendler, mevsimsellikler, aykırılıklar ve patlamalar gibi gözlemler dizisi olarak gösterilebilen doğal olayları tanımlamayı amaçlar. Bir zaman serisi, zamanda sıralanabilen gözlemler veya olaylar dizisidir. Zaman serisi verisi, eğer zamanın her anında gözlem varsa sürekli, eğer düzenli veya düzenli olmayan aralıklarda gözlem varsa kesikli olabilir. Bu epidemiyolojik çalışma, zaman-mekânsal verinin analizinde Bayesian katkının önemli olduğunu göstermiştir.

Ele alınan durumlarda, ekolojik süreçleri ve iklim değişimlerini daha anlaşılır kılmak için ve ayrıca halk sağlığı verileri ve saklı uzayda bulaşıcı hastalıkların yayılma haritası ile ilgili yeni teknolojiler geliştirmek için, mekanda, zamanda ve zaman-mekansalda elde edilen verilerin hesaba katılarak istatistiksel analizi çok önemlidir. Zaman-mekânsal veri için istatistik, istatistiksel hesaplama yanı sıra hiyerarşik ve özellikle Bayes yaklaşımı son gelişmeleri içeren sistematik nicel tekniklerdir. Genelde, gözlenen veri aşağıdaki formlardan birine sahiptir:

1) Verilen herhangi bir olay için bu olayın gerçekleştiği zaman ve yere sahip olduğumuz $X \times [0,T]$ sürekli zamanında tam bir harita. Genellikle, bu gözlem tipi elde edilmesi zor olduğu için azdır.

2) $t_1 < t_2 < \cdots < t_k$ zamanları üzerinde aşamalı gözlemlenen verili k kapsamlı harita. Gerçekte, kendi yoğunluğu tarafından sağlanan nokta sürecinin olasılığı mevcut olmadığından pratikte sürekli zaman gözlemleri çok nadirdir. Dolayısıyla, her mekânsal haritada, iki gözlem verisiyle sınırlı birikimli yoğunluk değerine sahibiz.

3) Öncekilerden daha az bilgi veren üçüncü tip ise, X in X_i alt uzayında sayılan X in parçalarında oluşan $t_1 < t_2 < \cdots < t_k$ gözlem zamanlı ve kesin yeri belli olmayan olaylardır.

Uç değer analizlerinin uygulamalarında mekânsal Poisson yoğunlukları modellenirken, Kottas ve Sanso [23] homojen olmayan mekânsal Poisson sürecinin gözlem kümesinden elde edildiği kabul edilen noktaların uzaysal konfigürasyonunun analizi için bir metod önermişlerdir. Bu metod, doğrudan Poisson yoğunluğu ile ilgili olan sınırlı bölgede tanımlanan yoğunluk fonksiyonunun modellemesine dayanır. Kottas ve Sansö, karma dağılım için iki değişkenli Beta dağılımını ve prior Dirichlet sürecini kullanarak parametrik olmayan dağılımın karışık bir modelini geliştirmişlerdir. Benzer çalışmalar, Green and Richardson'un [34] çalışmasında da vardır. Bu çalışmada, n jeografik bölgelerde sayılmış epidemiyolojik mekânsal veriyi analiz etmek için hiyerarşik Bayesian yarı-parametrik model ele alınmıştır.

Bu çaalışmada, epidemiyolojik gözetimi Bayesian çerçevesinde simüle etmek için Dirichlet süreci ile ilgili olan zamansal Cox süreci üzerinde oğunlaşılmıştır.

4.4 İncelenen Süreçlerin İstatistiği

Dirichlet süreci ile ilgili zaman-mekânsal Cox sürecinin istatistiksel içeriğinde, çalışılan bölgede görülen epidemiyolojik olaylar ele alınmıştır. Bölüm 4.1 de üzerinde durulduğu gibi, bu olaylar genelleştirimiş Shot Noise yoğunluk ölçümü ile yönlenmiş zaman-mekansal Cox sürecini meydana getirir[44]. Model, Gama dağılımı üzerinde merkezlenmiş Dirichlet sürecine uygun olarak gizli bir Poisson süreciyle ilgili bir ölçü

yoğunluğuna karşılık gelir. Ardaşık gözlem tarihlerinin çeşitli çiftleri arasında meydana gelen olayların mekânsal bölgelerini içeren veri için, genelleşmiş MCMC motodları kullanılarak ilgilenilen parametreler için Bayesian çıkarımsal kavram geliştirilir. Bu kavram, ilgilenilen parametrelerin simülasyonu için Larissa Audrey VALMY [44] tarafından tanımlanmıştır.

Gözlenen veriler ile ilgili olayların zaman ve mekânsal çalışmaları ile ilgilenirken, bu analiz çeşitli şekillerde yapılabilir [15, 16, 36, 117, 118, 119 ve 120]. 4.2 de gözlenen veri ile ilgili üç durum gösterilmiştir. Burada, zamanın düzenli olmayan aralıklarında X gözlem alanının tam haritasını kullanmakla ilgileneceğiz. Olayların yeri gözlemlenir olayların zamanı ise sadece kapalı bir aralıktadır.

Ayrıca, araştırılan olayların dağılımı genelleşmiş Shot Noise Cox sürecini takip eder [31]. Üstelik bu model olayların gerçekleştiği çevrelere bağımlı ve birinin diğerlerinden bağımsız olarak dağıldığını söyler. Zaman-mekânsal yoğunluğun gözlenen olaylarla ilgili rastgele katkının düzgün toplamı olduğunu da ileri sürer. Basit nokta sürecinin bu tipi, sismoloji [6 ve 121], ekoloji [9, 78 ve 122], nörofizyoloji [12 ve 123], hidroloji [124], görüntü analizleri ve epidemiyoloji [23, 125 ve 126] ve orman araştırmaları [28 ve 50] gibi pekçok alanda kullanılmıştır. Olaylar tamamen rastgele, fakat kümeleşmiş olarak gözlemlenmiştir ve mekansal korelasyon ile net pozitif zaman mevcuttur. Neyman-Scott ve diğerleri [31] gruplarda kullanılan Poisson süreçleri de kullanmışlardır. Model

Dirichlet sürecine dayanan bu süreçler için parametrik olmayan Bayesyen yaklaşım, gözlenen nokta sürecinin yoğunluğuna gizli çevresel katkının etkisi olarak modellenir. Bu bölümde nokta süreci ile ilgili model, ilgilendiğimiz verinin bu olayların meydana geldiği tarihler konusundaki kısıtlı bilgi ile bilinen mekânda meydana gelen olaylar olduğunu vurgular. Takip eden alt bölümlerde, zaman-mekânsal Cox süreci ve Dirichlet süreci özetlenmiştir.

4.4.1 Zaman-Mekânsal Cox Süreci

Bir zaman-mekânsal Cox süreci, olayın olasılık yoğunluk fonksiyonu $\Lambda(.)$ üzerinde zaman-mekânsal Poisson süreci koşulu olarak tanımlanır. Bölüm 2'de Cox sürecinin kendi yoğunluğunun değişimiyle Poisson süreci olan bir stokastik süreç

olduğuaçıklamıştı. Poisson süreci yoğunluğunun değişmediği durumda zaman-mekânsal nokta süreci, bölüm 4'de ifade ettiğimiz gibi, sabitir. Bütün bunlardan dolayı, bu modeller için kullanılan notasyon ve kavramları açıklamakla başlayalım. $X \subseteq \mathbb{R}^2$ sınırlı düzlemsel bölgede gözlemlerin N nokta süreci rastgele yoğunluk fonksiyonu $\Lambda(s), s \in X$ tarafından yönetilen Cox süreci olduğunu kabul edelim. Zaman-mekânsal Cox süreci ile ilgili diğer süreçlerin kavramları aşağıda açıklanmıştır.

$(\Omega, \mathcal{B}, \mathrm{P})$bir olasılık uzayı ve$\left(X \times \mathbb{R}_+, \mathcal{B}_X \otimes \mathcal{B}_{\mathbb{R}_+}, \nu \otimes \lambda_L\right)\lambda_L$Lebesgue ölçümü olmak üzere ölçüm uzayı olsun. Yukarıda $t_1 < t_2 < \cdots < t_J, j = 1, \ldots, J$ zamanlarında gözlenen verinin tam harita tipi olduğu verilmişti. Bu durumda, $M_j = \left\{x_1, \ldots, x_{n_j}\right\} t_j$ zamanına kadar X in ürettiği olayların mekansal kümesi eleman sayısı n_j olmak üzere, t_j zamanında haritadaki gözlemi gösterir. Bu kümeler, $\left(t_j\right) j \in \{1, \ldots, J\}$ zamanları arasında ardaşık olarak gözlenen N uzay-mekansal nokta süreci olayları olarak düşünelim. Bu tarihlerin, kesikli zaman durumunda belirli düzenli aralıklar halinde veriler olmadığından, sürekli zaman olarak düşüneceğimizi belirtelim. M_J haritadaki gözlemin sonuncusudur ve dolayısıyla önceki J gözlem haritalarını içermektedir. Ayrıca, mekansal yerlerin zamanı tamdır, oysa $\left\{T_i,\ i = \{1, \ldots, J\}\right\}$ olaylarının tarihleri bilgisi sembolik olarak $T_i \in \left]t_{j-1}, t_j\right[$ şeklinde verilmiş, olduğundan bunlar tam değildir.

Bu açıklamalarda, Λ ölçüm yoğunluğu kesin düzgün koşullarda olduğunda, bir uzay-zamansal Cox süreci modeli oluşturmanın tek yolu vardır. Bunun yoğunluk süreci aşağıdaki gibi kurulur: eğer X üzerinde yayılma ν ve $X \times \mathbb{R}_+$ de indeksli λ bir pozitif ölçülebilir süreç mevcut ise

$$\forall (B, t) \in \mathcal{B}_X \times \mathbb{R}_+, \Lambda(B \times [0, t]) = \int_B [\int_0^t \lambda(x, t) \nu(dx)] dt \qquad (4.8)$$

Burada λ, N sürecin yönlenmiş yoğunluğu olarak adlandırılır. $\lambda(x, t)\nu(dx))dt$ fonksiyonu, $\nu(dx)$ yüzeyinin x in sonsuz küçük dx bölgesi civarında ve sonsuz küçük dt zaman aralığında görülen olayları açıklar. Belirli düzenlilik koşulları altında, λ nn N koşullu sürecinin olasılığını yazabiliriz [3 ve 17]. $\left\{\left(x_j, t_j\right),\ j = \left\{1, \ldots, n_j\right\}\right\}$ verisi uzay-mekanda bütün olayların zamanları ve yerlerini içerir. θ model parametrelerinin vektörü olmak üzere λ_θ yönlenmiş yoğunluğunun oluşturduğu N_C Cox süreci ile

ilgileneceğiz. $t_0 = 0$ da M_j/M_{j-1} haritasına λ_θ koşullu olasılığının ifadesi aşağıdaki gibidir:

$$\mathcal{L}_{M_j/M_{j-1}} = e^{-\Lambda_\theta(X\times[t_{j-1},t_j]} \prod_{k=n_{j-1}+1}^{n_j} \int_{t_{j-1}}^{t_j} \lambda_\theta(x_k,t)dt \qquad (4.9)$$

Buradaki M_j/M_{j-1}, M_j/M_{j-1} de üretilmiş olayların haritası ile ilgilidir.

λ_θ yoğunluk süreci ile koşullu global olabilirliği (likelihood), (4.9)'dan elde edilebilir. Buradan,

$$L(x_1,\ldots..,x_n,\theta) = e^{-\Lambda_\theta(X\times[0,T_j]} \prod_{j=1}^{J} \prod_{k=n_{j-1}+1}^{n_j} \lambda_\theta(x_k,t)dt \qquad (4.10)$$

bulunur. Bu sonuçlar, yoğunluk süreci rastgele olmadığından her uygun zaman-mekânsal nokta sürecinin homojen olmayan Poisson sürecinin olabilirliği olarak açıklanabileceğini gösterir.

Bu bölümün kalan kısımlarında, yoğunluk sürecinin,λ_θ sürecinin bir dağılımı olduğunu kabul edeceğiz. [34]'te mekânsal heterojenliği dikkate alan çeşitli yaklaşımlar önerilmiştir. Burada, yarı-parametrik Bayesian yaklaşımı gizli Markov alanı modelleri ile gösterilmiştir ve epidemiyoloji ve görüntü problemlerin uygulanmıştır. Mekânsal Cox sürecine bakarsak, örneğin [128]'de, bazı yönlenmiş ölçüm modelleri karşılaştırılmıştır, [22, 129 ve 130] kaynaklarında ise tanıtılan Log-Gaussian Cox süreci bu sınıfı zaman-mekânsal süreci içerecek şekilde geliştirilmiştir. Log-Gaussian Cox süreci metodolojisinin gelişimi [131]'de, sayılabilir ve çok değişken olan kümelenmiş verileri de kapsaya tartışılmıştır. Shot Noise Cox nokta süreçleri, mekânsal ve zaman-mekânsal kümeleşmiş süreçler için önemli modellerdir. [22, 129 ve 130] kaynaklarında, Log-Gaussian ve Shot Noise süreçleri, verilen yoğunluk süreci modelinin genelleştirilmesine vurgu yapılmıştır. Shot Noise sürecinin genişlemesiyle yoğunluk süreci için karışık model üretilebilmiştir ve çalışılan uzayın bölgesi mevcut olduğunda, farklı yoğunluk modelleri uygulandığından ve bu bölgenin tanımlanmasına ihtiyaç olmadığından, bu genişleme ilginçtir.

Benzer çalışmalar, pozitif otokorelasyonlu Cox sürecini yoğunluğu üzerinde de yapılmıştır. Bazı süreçleri gözden geçirdikten sonra, şimdi Lévy tabanında Cox sürecini

esas alan genelleşmiş Shot Noise sürecine öncülük eden zaman-mekânsal Cox süreci ile ilgileneceğiz [31 ve 131]. Shot Noise Cox süreci, Cox süreci için modelleme çerçevesinde doğal engeller gösterecektir. N_l, $(X \times \mathbb{R}_+^2, \mathcal{B}_X \times \mathcal{B}_{\mathbb{R}_+^2})$ esasına dayanan negatif olmayan Lévy ile ilgili bir nokta sürecini gösterir. Genelleşmiş zaman-mekânsal Shot Noise Cox süreci, kendi yoğunluk süreci ile aşağıdaki gibi tanımlanır:

$$\forall (x,t) \in A_t \subseteq X \times \mathbb{R}_+, \lambda_\theta = \int_B a K_{(\sigma_1,\sigma_2)}\big((x,t),(y,s)\big) d(N_l(y,s,a) \qquad (4.11)$$

Burada,$K_{(\sigma_1,\sigma_2)}$,$A_t(x)$ nın bütün (y,t) noktaları için $A_t(x)$ olasılık yoğunluğu üzerinde $K_{(\sigma_1,\sigma_2)}(.,(y,s))$olan çekirdektir.

(4.11) yoğunluk süreci ile ilgili açıklamalardan, aşağıda verilen bir başka yoğunluk süreci ifadesine ulaşılır:

$$\forall (x,t) \in A_t,\ \lambda_\theta(x,t) = \sum_{j=1}^{n} a_j K_{(\sigma_1,\sigma_2)}((x,t),\big(y_j,s_j\big)) \qquad (4.12)$$

Burada, n katkı rastgele sayısı olmak üzere, $(y_{j,}, s_j)$, $X \times \mathbb{R}_+$ de N_1 nokta sürecinin gerçekleşmesi, a_j, $\mathbb{R}_+$ üzerinde G olasılık dağılımı ile ilgili ve $(y_{j,}, s_j)$ den bağımsız olan özdeş işaretlerin dağılımıdır. Böylece, eğer N_1nokta süreci μ skaler yoğunluklu homojen Poisson süreci ile ilgili ve X üzerindeki N nokta süreci N_1 nokta süreci için koşul ise,$\lambda_\theta$yoğunluk süreci standart Shot Noise süreci olur. Bağımsız Poissson süreçleri yoğunluğunun süperpozisyonu $a_j K_{(\sigma_1,\sigma_2)}(.,\big(y_j,s_j\big)$ ile verilir. İlk meydana gelen dağılımların tanımı, ilk olayın kendisine komşu ikinci olayı ürettiği bu süreçlerle yapılabilir. Pratikte sismolojide artçı depremler, ekolojide açıkdeniz bitkileri ve epidemiyolojide etkilenmeyen temaslar olabilir. Buna benzer bir olay için [28] kaynağında burada anlatıldığı gibi modellenmiş Poisson süreci tartışılmıştır.

Genelleştirilmiş Shot Noise Cox süreci, herbir N_1 nokta sürecinde çekirdek fonksiyonuyla tanımlanır [31]. Bu fonksiyon rastgele bant genişliğini dikkate alır. Diğer taraftan, çoğu uygulamada, N_1süreci Poisson sürecini gösterirse, bu diğer modellerin daha uygun olacaklarını kanıtlar. Örneğin, yaşayan bazı böceklerin türiçi rekabet nedeniyle bölgesel davranışları olduğunda itici Markov modelleri kullanılır.

4.4.2 Dirichlet süreci

Parametrik olmayan Bayesian istatistik üzerinde yoğunlaşmanın geliştirğdiği perspektiften, Ferguson, stokastik olasılık ölçümü olan Dirichlet süreci kavramını tanıtmıştır. Bu süreçle ilişkili olasılık dağılımı (Dirichlet dağılımı) tanımlamakla başlayacağız. Bu çok değişkenli Beta dağılımının genelleşmişidir. Beta dağılımı parametreleri α, β yı açıklayabiliriz. Bunlar, $\alpha - 1$başarılı (p olasılığıyla) ve $\beta - 1$ başarısız $(1 - p$ olasılıkla) gözlemden sonra $p$parametreli Binom dağılımının olabilirliğini gösterirler. Beta dağılımının, Binom dağılımının eşleniği olduğunu söyleriz. Yani, rastgele değişkenler, biri veya diğerlerinin özel dağılımları ile ilgili dağılırlar.

Bu durumda, Dirichlet dağılımı çok-değerli dağılımın eşleniği olarak tanımlanır. $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_k) \in \mathbb{R}^k$, $k \geq 2$ için yoğunluk parametresi adı verilen bir vektör olmak üzere. k dereceli ve αparametreli Dirichlet dağılımı aşağıdaki formda tanımlanır:

$$f(x_1, \dots, x_k, \alpha) = \frac{1}{B(\alpha)} \prod_{i=1}^{k} x_i^{\alpha_i - 1} \tag{4.13}$$

Çalışma bölgesinde x_i lerin toplamı 1'e eşittir, diğer durumlarda yoğunluk 0'dır. Sabitin normalizasyonu, Gamma fonksiyonuyla (çarpanların genellemesi) açıklanan çok-değişimli Beta fonksiyonudur:

$$B(\alpha) = \frac{\prod_{i=1}^{k} \Gamma(\alpha_i)}{\Gamma \sum_{i=1}^{k} \alpha_i} \tag{4.14}$$

$Dir(\alpha_1, \dots, \alpha_k)$,$\alpha_i$ parametresi ile ilgili yoğunluktur, dolayısıyla k boyutlu bir fonksiyon vardır.

Dirichlet dağılımının bu kavramı kullanılarak Dirichlet süreçleri de tanımlanabilir. Bir Dirichlet süreci, olasılık dağılımının meydana geldiği stokastik süreçlerin bir ailesidir. Genellikle Bayesian çıkarımlarda rastgele değişkenlerin dağılımları ile ilgili prior bilgilerinin tanımlanmasında kullanılır. Bu rastgele değişkenler, biri veya diğerlerinin özel dağılımları cinsinden dağılırlar. Detaylar için [30, 133 ve 134] makalelerinden faydalanılabilir. Bir Dirichlet sürecinin gerçekleştiği olayların 1 olasılıkla kesikli olması önemli bir özelliktir. G_0bir olasılık ölçümü, αise pozitif reel sayı (yoğunlaşma parametresi) olsun. G nin, ölçülebilir uzayın herhangi $(A_1, \dots, A_k)$ parçalanışı için

(X, H)ölçülebilir uzayında stokastik ölçüm olasılığı olduğunu düşünelim. G'yi aşağıdaki gibi tanımlayalım:

$$G(A_1), \dots, G(A_k) \sim Dir(\alpha G_0(A_1), \dots., \alpha G_0(A_k)) \quad (4.15)$$

(4.15) ifadesi, basit olarak $G \sim DP(\alpha, G_0)$ şeklinde yazılabilir. Bunun anlamı, bir kümenin ölçümü, herhangi A Borel kümesinin beklenen değeri $E[G(A)] = G_0(A)$ verdiğinden, G'nin beklenen değeri veya temel ölçüsü olarak adlandırılan G_0 başlangıç ölçümü ile orantılı olan parametrelerinin Dirichlet dağılımına göre kendi üzerine tasviri olduğudur. $\alpha > 0$ skaleri, aşağıdaki şekilde G_0 civarında G'nin değişebilirliğini karakterize eden kesin bir parametredir.

A ve B, $A \cap B = \emptyset$ koşulunu sağlayan iki Borel kümesi olmak üzere. Bu durumda:

$$Var\big(G(A)\big) = (G_0(A) \times \big(1 - G_0(a)\big))/(\alpha + 1) \quad (4.16)$$

$$Cov\big(G(A), G(B)\big) = (G_0(A) \times G_0(B))/(\alpha + 1) \quad (4.17)$$

Eğer stokastik ölçüm olasılığı $G \sim DP(\alpha, G_0)$ ise, [113] makalesinde tanımlanan G Strick-Breaking gösterilimi aşağıdaki formdadır:

$$V_k \overset{i.i.d}{\sim} Beta(1, \alpha)$$

$$\theta_k \overset{i.i.d}{\sim} G_0$$

$$P_k = V_k \prod_{j=1}^{k-1} (1 - V_j)$$

$$G(.) = \sum_{k=1}^{\infty} P_k \delta_{\theta_k}(.) \quad (4.18)$$

Burada$\delta_{\theta_k}(.)$,$\theta_k$de Dirac ölçümüdür. Üstelik (4.18) denklemlerinde verilen $P = (P_k)$ dizilerindeki P_k'lar$0 \le P_k \le 1$ve $\sum_{k=1}^{\infty} P_k = 1$ koşullarını sağlayan θ_k bağımsız rastgele değişkenlerin olasılığıdır.

Bir Dirichlet süreci, aynı zamanda n adet gözlemden sonraki sürecin posteriör dağılımıdır. Bu süreç, gerçekleşen olaylardanG_0'ların tahmininde çok kullanışlıdır.

4.4.2.1 Posteriör Dağılımı

Parametrik olmayan modellerin pratikteki kapsamı çok geniştir ve sağlamlık gerekliliklerini sağladıklarından bu modellerin uygulamalı çalışmalarda kullanımı gittikçe artmaktadır, .

Parametrik olmayan Bayesian genişlemesi yaklaşımının pratik kullanımı zordur. Bu zorluklar simülasyon teknikleri ile giderilmektedir. Bu teknikler, Dirichlet sürecinin posteriör gösterilimi ile üretilirler ve kümeleme çalışmalarında hızlı sonuç verdiklerinden dolayı kullanımları hızla yaygınlaşmaktadır.

Dirichlet sürecinin stokastik ölçüm olasılığı olarak kullanımının temel motivasyonu, kendisi de bir Dirichlet süreci olan güncel posterior dağılımı basitleştirmesidir [134]. Aşağıdaki modeli ele alalım:

$$y_1, \dots, y_n/G \overset{i,i,d}{\sim} G \text{ve} G/\alpha, G_0 \sim DP(\alpha, G_0) \tag{4.19}$$

Bu durumda, G'nin posteriör dağılımı

$$G/y_1, \dots, y_n \sim DP(\alpha^*, G_0^*), \alpha^* = \alpha + n, G_0^* = \frac{\alpha}{\alpha+n} G_0 + \frac{1}{\alpha+n} \sum_{i=1}^{n} \delta_{y_i} \tag{4.20}$$

G dağılımının marjinelliği ile elde edilen tahmin dağılımı, aşağıdaki Polya gösterilimi ile verilir.

$$y_{n+1}/y_1, \dots, y_n \sim \frac{\alpha}{\alpha+n} G_0 + \frac{1}{\alpha+n} \sum_{i=1}^{n} \delta_{y_i} \tag{4.21}$$

Böylece, örneklenmiş olan $y_1,\dots,y_n$ gizli değişkenler koşullu (4.21) denkleminden şu sonuç çıkarılabilir: yeni değerin olasılığının $\frac{\alpha}{\alpha+n}$ olduğu önceki değerden elde edilmiş olasılık $\frac{1}{\alpha+n}$'dir. Bu model, Dirichlet sürecinin olaylarını açık olarak görmemizi sağlar. Bu olaylar, sonsuz sayılabilir karışık dağılım olarak görülebilir [135]. Dirichlet sürecinin bir olayı, kesiklidir ve bu özellik Dirichlet süreci tarafından modellenen karışık modellerde sınıflandırmada önemli rol oynar.

4.4.2.2 Bayesyen Çerçevesinde Karışık Modeller

Son yıllarda, parametrik olmayan Bayes metotlarına olan ilgi, esneklikleri, verimlilikleri ve posterior hesaplamaları için kolay algoritmalar sunmaları dolayısıyla artmıştır. Son

araştırmalarda, karışık Dirichlet süreci (DPM) ni esas alan prior dağılımın belirlenmesi sorusu üzerinde yoğunlaşılmaktadır [30].

Bir karışımdaki Dirichlet süreci karıştırma yasası (mixing law) ve karşık yasası (mixed law) olarak görülebilir. Bu alt bölümde, parametrik olmayan karışık bileşenlerin prior dağılımı olan sürecin en önemli uygulamalarından biri verilmektedir. Karışık modellerin yayılma modelleri ile iyi bilinen ilişkileri vardır [136, 137 ve 138]. Bu çalışmada parametrik olmayan Bayesian analizleri tartışacağız ve genelleşmiş Dirichlet sürecinin farklı değişkenlerini karşılaştıracağız. y_i gözlemi aşağıdaki şekilde tanımlansın:

$$y_i/\theta_i \sim K(./\theta_i),\ \theta_i/G \sim G \qquad (4.22)$$

Burada, $K(./\theta_i)$, gözlemi θ_i parametre koşullu $y_i$dağılımının karışık yasasını gösterir. θ parametresi için sayılamaz sonsuz Θdurum uzayı için karıştırıcılarla bir genelleme yapmak, G θ nın dağılımı olmak üzere mümkündür ve bu karıştırma yasasıdır. Karışık sürecin yazılımı aşağıdaki gibi $G_0 \sim DP(\alpha, G_0)$ile formüle edilebilir [135]:

Karışık yasası

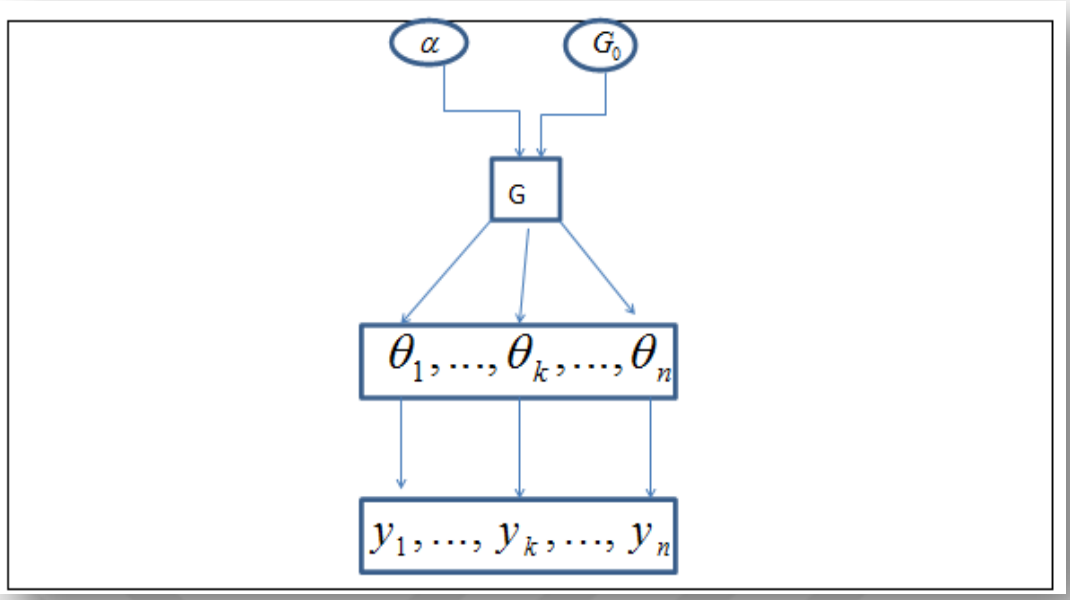
$$f(.) = \int_\theta K(./\theta_i)dG(\theta) \qquad (4.23)$$

Karışık Karıştırma yasası

Bayesian çerçevesinde, G karıştırma yasası, bazen Pprior dağılımın tahmin edilemeyenidir. y_i (4.23) denklemiyle verilen karışık dağılımın bir olayı olsun. G için prior dağılım seçerek, Dirichlet süreci, hiyerarşik formda bilinen modellere göre tahmin probleminin yoğunluğu ile yeniden formüle edilebilir. Dolayısıyla

$$\begin{cases} y_i/\theta_i \sim K(./\theta_i) \\ \theta_i/G \sim G \\ G \sim DP(\alpha, G_0) \end{cases} \qquad (4.22)$$

yazılabilir. Burada $G_0 \sim DP(\alpha, G_0)$,αve G_0 parametreli Dirichlet sürecidir ve prior dağılım $G \sim P$ ifadesi ile stokastik ölçüm olasılığına bağlıdır. Bu süreci Şekil 4.1 de verilen grafik ile özetleyebiliriz.



Şekil 4. 1 Hiyerarşik modelin döngüsel grafiği

[139] kaynağında gösterildiği gibi, Dirichlet süreçleri ile standart uzaysal Gamma süreci arasında bağlılık vardır. Bu durum, modelin posterior dağılımına esas teşkil ettiği için hesaplamaları daha da karmaşık hale getirmektedir. Bu bölümde, ilgilendiğimiz kesikli Dirichlet süreci verilmiştir. Bu süreç, farklı zaman ve yerde oluşan olayların ürettiği katkı değerlerinin yoğunluğunun özdeş oluğu yerlerdeki durumları tanımlamamıza imkân verir. Shot Noise Cox süreçlerini içeren rastgele katkılar, genelleşmiş modellede olduğu gibi bağımsızlardır ve özdeş dağılmışlardır [29]. Sıklıkla kullanılan dağılım Gamma dağılımıdır. Cox süreci modeli ile tartıştığımız bütün bu süreçler, Dirichlet süreci ile ilgili genelleşmiş Shot Noise tipli süreç yoğunluğuna sahiplerdir. Böylece, Poisson süreci ile ilgili gizli olaylara sebep olan zaman ve mekânı içeren modeller kurabiliriz. Bu olayların her bir katkısı Gamma beklenen dağılımın Dirichlet sürecini takip eder. Bu durumda, hastalıkların ortaya çıkması, hastalıkların taşınması ve bitki yerleşimi gibi olayların dağılımı, her nokta tek bir olayı göstermek üzere uzaysal nokta süreci olarak ele alınabilir. Ele alınan nokta süreçleri teorileri, epidemiyolojide uygulama yapmak üzere bir sonraki bölümde kullanılacaktır.

BÖLÜM 5

EPİDEMİYOLOJİ İÇİN NOKTA SÜREÇLERİ

5. Giriş

Şimdiye kadar nokta süreçleri teorisini inceledik. Bunların bazıları, (örneğin hastalık veya risk yayılımının tahmini gibi olaylar.) [3 ve 19] tarafından verilen zaman nokta süreçlerini, bazıları [53, 117 ve 119] tarafından verilen mekânsal nokta süreçlerini ve diğerleri de potansiyel model olarak, [55 ve 54]tarafından verilen ve birçok araştırmacının da kullandığı zaman-mekânsal nokta süreçleri ile ilgiliydi. Burada, önceki bölümde tartışılan mekânsal istatistik modellerine dayanan ve epidemiyolojide kullanılan nokta süreçleri çeşitleri ile ilgileneceğiz. Çalışılan bölge, bize ilişki içinde olduğu bölgelerde de incelenen olayla ilgili küresel anlamda bilgi sağlar. Gözlenen özel yayılma, bu mekanizmanın coğrafi bölgelerde sebep olduğu etkileri anlamamıza yardımcı olur.

Bu bölümde, Orta Afrika Cumhuriyeti (Central African Republic - CAR)'de sıtma hastalığı ile ilgili epidemiyoloji kontrol çalışmasına dayanan bu süreçlerin uygulaması yapılmıştır. CAR'ın bazı bölgelerindeki çevresel etkilerin sağlık risklerini ve uyarı sistemini (bulaşıcı hastalık taşıma riski olan hastanın belirlenmesi durumu) dikkate alarak, sağlık problemlerini çözme araştırmaları, epidemiyoloji çalışmalarını en iyi tanımlayan stokastik modellere dayanır. Bu stokastik modeller, hastalıktan her yıl etkilenen yeni hasta sayısını içeren özel haritalara dayanan coğrafi bilgilerin istatistiksel gelişimini sağlar. Ele alınan modeller, genelde sağlık göstergelerindeki mekânsal değişimlerinin analizi ve patoloji problemlerinde de kullanılırlar. Çalışmamızda,

vektörün (hastalığın taşındığı vektör) varlığı, hastalığın geçişi ve vektörle temas olasılığı faktörleri dikkate alınacaktır.

Epidemiyolojik problemlerin çözümü araştırmaları, genelde kesin olmayan değişkenliklerin sebeplerinin araştırılmasına bağlıdır. Bu değişkenlikler, risk faktörünün oluşumunun sonuçları, tedavinin başarısı veya hastalığın teşhisinin olasılıklar cinsinden ifade edilememesine sebep olur. Çalışmamızda, hastalığın oluşum olasılıkları veya CAR' da sıtma hastalığından her yıl etkilenen yeni hasta sayısı 2002-2013 yılları arasında ele alınmış ve sıtma için zaman-mekânsal dağılım veya bu olasılıkların zaman-mekânsal gelişimi analiz edilmiştir. Bu dağılım, değişimlerin, müdahalelerin veya diğer faktörlerin tanımlanmasına bağlılığının nasıl olduğunu değerlendirmeye izin verir. Ayrıca, hastalık geçişlerinin azalması veya durumunu korumasının katkısını da verir. Çalışma, sıtma salgını gelişiminin gözlemini ve olasılık araçları ile bu olayın istatistiksel modellemesini kullanmayı gerektirir. Bu ana araçlar, önceki bölümde verdiğimiz nokta süreçleri teorilerine dayanır. Bu teoriler ise ele alınan hastalığın etkisini ve ele alınan bölgedeki yerini ölçer.

İstatistiksel epidemiyolojinin, mekânsal istatistikleri içeren çok geniş uygulama alanları vardır ve günümüzde pek çok araştırmacının çalışma konusudur. Ayrıca, uzaydaki yerlerin önemli olduğu ve veri işlemine izin veren araçları önerir. Bu bölümde, kullanılan olasılık ve istatistik araçları, daha çok nokta süreçleri için ele alınmıştır. 5.1.1 ve 5.1.2 alt başlıklarını içeren 5.1'de uzaysal epidemiyoloji kavramlarına dayanan bazı epidemiyolojik kavramları özetleyeceğiz. 5.2 de ise çalışılan süreçler, CAR' da sıtma hastalığının zaman-mekânsal uygulamasını yapacağız. Fakat kullanılan analiz yöntemlerin hala mekânsal epidemiyoloji metotlarına dayanmakta olduğunu belirtelim.

5.1 Epidemiyoloji

Epidemiyoloji, hem koruyucu tıbbın temel bilimi ve hem de halk sağlığı politikası bilgilerinin kaynağıdır. Bitki, havyan veya insanlardaki hastalık frekansının değişimini tarif eder ve bu değişimleri inceler. Amacı hastalığın sebeplerini anlamak ve tedavi ve koruma metotları geliştirmektir. Epidemiyolojik çalışmaların büyük bir kısmında, epidemiyolojisitler sadece hastalık frekansıyla değil aynı zamanda ele alınan gruptaki

frekanslardaki farklılıklarla da ilgilenirler. Başka bir değişle, maruz kalma ve hastalanma arasındaki ilişkiyi çalışırlar. Ele alınan uzayda, hastalık kazayla oluşmaz, insanların bazı aktiviteleri buna sebep olur. Gerçekte, sebeplerini bilirsek pek çok hastalığı önleyebiliriz. İstatistiksel metotların epidemiyolojik problemlerin çözümlenmesine önemli katkıları vardır. Ayrıca hastalığa sebep olan lezyonlar ve erken ölümleri önlemek için sağlık politikalarının oluşturulmasında etkin olan pek çok faktörlerin belirlenmesi ve analizinde önemli rol oynarlar.

Bu çalışmada, epidemiyolojinin genel ilgilerinden farklı olarak gözlemler ve sıtma hastalığının uzaysal görünüşlerinin analizi ile sınırlıyız. Nihai amaç, varsayılan risk faktörleri ve hastalık arasındaki ilişkiyi belirlemek ve uzayda ve zamanda hastalığın yayılımının genişlemesine sebep olan faktörleri belirlemektir. Meydana gelen olayların çalışılmasında mekânsal epidemiyoloji ve çevreyle ilişkileri merkez alınmıştır. Bu yaklaşım, global mekanizmayı anlamamıza yardımcı olacaktır. Basit bir gözlemden mevcut durumu ve bunlara sebep olan çoklu faktörleri tanımlamanın zor olduğunu belirtelim. Çünkü değişkenleri çok fazladır ve her bir risk faktörü rastgele karakter gösterir. Bu değişkenler, global olayların mekânsal dağılımında yer alırlar. Bu, sadece gereklilik koşulları gerektirir.

Sağlık ile ilgili olayların çoğunun, hastalığın pozitif veya negatif olmasını sağlayan çok çeşitli sebepleri olabilir. Pozitif olmasını sağlayan sebeplerden bazıları verilen bir bölgede hastalığın geçişi veya vektör oluşumu vb. pasif olmasını sağlayan sebepler ise yaşam koşulları, hastalık yapıcı (patojen) veya vektör temasının olasılığı, risk faktörüne maruz kalma, beslenme alışkanlıkları vb. dir. Her yıl hastalık sayısındaki değişim, kısmen zaman ve uzayda yer alan bu faktörlerin sonucudur. Bu durum, sağlık olaylarının uzayda olduğu gibi zamanda da sadece rastgele oluşumlardan neden farklı olduğunu açıklar. Gözlenen olayın sebeplerinin değişkenliğini dikkate alırsak, epidemiyolojistler genelde bazı sağlık problemlerini açıklamada ciddi problemler yaşayabilirler. Bu değişkenliklere sebep olan faktörlerin epidemiyolojik araştırmasında, istatistikçi analitik strateji ile araya girer. Çalışmamızda, nokta süreci modelleri kullanmanın avantajı, farklı faktörlerin mekânsal değişkenliklerini kavramak için zaman ve mekânsal dağılımlara genel bir bakış sağlamasıdır. Epidemiyolojik tanıları geliştirmek için yerleri dikkate almakta rastgele değişkenlik kadar önemlidir. Şimdi, zaman-

mekânsal ve mekânsal nokta süreçlerinin kullanımında genel bakışa olanak tanıyan mekânsal epidemiyolojiyi özetleyeceğiz.

5.1.1 Mekânsal Epidemiyoloji

Mekânsal epidemiyoloji, yıllardan beri merak konusudur. Özellikle, verilen hastalıkların mekânsal dinamiği ve dağılımlarını araştırır. 19. Yüzyılda Londra'da John Snow tarafından yapılan kolera epidemik analizleri bir örnektir. Hastalık durumları veya çevresel durumlarından söz etmek için, olayların yerlerini bilmek gerekir. Coğrafi plan veya hastalık bulaşma yerinin belirleyen hastalık insidans haritası, hastalığın geçiş modu hakkında bilgiyi ve özellikle oluşumunu önleme veya ele alınan coğrafi alanda kökten temizlemeyi sağlar. Belirli bir zamandan beri, bulaşıcı veya parazitik hastalıkların mekânsal dağılımının özelliklerini araştırmak için metodlar geliştirilmektedir. Coğrafi analizler metodları, hastalığın mekânsal yayılması analizine uygulanabilirler. İki analitik yöntemden bahsedilebilir:

1) Harita üzerinde olayların düzgün tanımlanmasına dayanan tanımlayıcı yöntemler (haritalama yöntemi): Haritanın bir özetidir, harita üzerinde okumayı kolaylaştırır, komşu verinin değerine dayanan istatistiksel serinin istikrarsız değerlerini elemeyi sağlayan matematik araçları vardır, uzay veya mekânsal heterojenlikteki yerleriyle değişkenin korelasyonunun tahmini olan mekânsal otokorelasyonun ölçümünü sağlar.
2) İyi sınırlanmış coğrafi bölgede çalışılan olay veya hastalığın kümeleşmesini tanımlamayı amaçlayan zaman-mekânsal kümeleme analiz metodları. Bu metodlar, sağlıktaki geçişlerin etkisini incelemek için kullanılır.

Coğrafi analizler yöntemleri, mekânsal olayların yerleşiminde önemli rol oynarlar ve epidemiyolojiste yaygın olaylar, hastalıkların yayınla veya sönümlenmesinin anlaşılmasında yol gösteren yöntemini sağlamlaştırmasını sağlar. Ayrıca, verilen uzay ve zamanda yinelenen iç patojenlerin oluşumunu açıklar ve istatistiksel modellerle simüle edilir. Hastalık durumları uzay ve zamanda gruplaştırılarak zaman-mekânsal nokta süreci kurulur. Bu oluşum "kümeleme" olarak adlandırılır ve basit olarak çalışılan değişkenin mekânsal organizasyonudur. Diğer değişken ile (risk faktörü) ilgisi olabilir

de. Bu özel mekânsal organizasyon (desenleme), birbirlerinin yakınındaki hastalıkların birleşimi olarak tanımlanır. Yakınlık, coğrafi uzaklıklar dikkate alınarak belirlenir.

Bazı durumlarda, çalışılan coğrafi bölgede, özel düzenleme olmadan hastalığın rastgele dağılması durumunda kümeleme oluşumu olmadan da mekânsal oluşum kullanılabilir. Hastalıkların bu coğrafi dağılımı, düzgün (tamamen mekansal rastgelelikle) olabilir, fakat popülasyon riskinin mekansal dağılımına ve i.mekândaki çalışmanın mekansal dağılımına güçlü şekilde bağlıdır. Kümelemede mekânsal heterojinite, tipik olarak heterojen Poissson dağılımı ile gösterilir. Yani, popülasyon riski dağılımının bağımlılığı (sabit risk hipotezleri) sıfır hipotezleri ile ilgilidir: i.mekânsal olayların beklenen değerinin yayılımı (n_i etkili) $P_0 = \frac{N_X}{N}$ile $Y_i \rightsquigarrow \mathcal{P}ois(\lambda_i = n_i p_0)$ olur. Burada N_X ve N sırasıyla, gözlenen durumların toplam sayısı ve çalışılan bölgedeki insanların toplam sayısıdır.

C,mekânsal birimde kümeleme oluşumu, alternatif hipotezler şu dağılımlarla gösterilebilir: $Y_c \rightsquigarrow \mathcal{P}ois(\lambda_c = n_c p_c)$, burada $P_C \gg P_0$ özel birimin, özel riskini göstermektedir. Şekil 5.1 kümeleme analiz metodlarının mümkün dağılımı için nokta süreçleri oluşumunu göstermektedir.

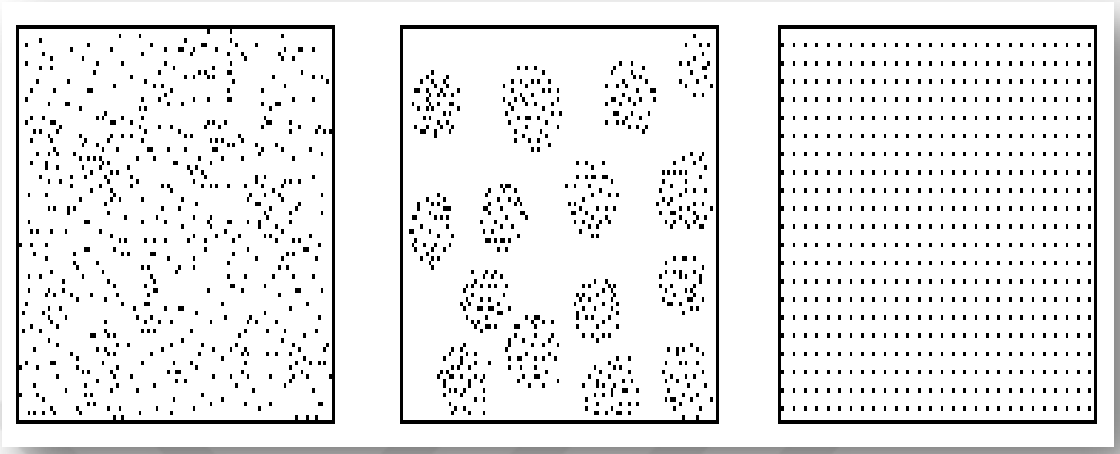
Eğer çalışılan popülasyonda fazlalık varsa ve bu fazlalık biliniyorsa, bu sıradışı kümelemenin açıklayabildiği bir veya çeşitli hipotezler varsa kümeleme analizi yapmak işe yarar.

Çalışmaların çoğunda, kümeleme analizleri iki ana yöntemle yapılır: global ve yerel tespit metodları [140].

Genel tespit metodları bölgede ortaya çıkan kümeleşmiş hastalıkları belirlemekle ilgilidirler. Veri genellikle mekânsal bir birimde kümeleşir (örneğin bölgede sıtma hastalığından ölümler). Kümeleme yerleri hakkında prior kabuller yapmaksızın veri kümesinin analizi kullanılır.

Ele alınan tespit metodları, bir iletim kaynağı ya da kirlenme ve aşırı şüpheli hastalık vakaları arasında bir ilişkiyi analiz etmek için kullanılırlar. Test edilen hipotezler, verinin ilk gözlemleriyle sağlanırlar. Dolayısıyla, ilk sonuçların açıklanması kolay değildir. Sıtma hastalığı durumunda, klinik olarak çok detay vermeksizin yıllara göre hasta

olmayanlarla hasta olanların dağılımlarını karşılaştıracağız. Fakat istatistiksel gözlemler ve sahip olduğumuz bütün veri bölgesel değil globaldir. Veri uygun olmadığından, yukarıda açıklanan yöntemlerin bazılarının kullanımı imkânsızdır.



(a) (b) (c)

Şekil 5. 1Kümeleme metoduyla oluşan dağılımların nokta süreçleri örnekleri (a) rastgele dağılım, (b) kümelenmiş dağılım (c) düzgün dağılım.

5.1.2 Olayların Zaman-mekânsal Dağılımı

Nokta süreci metodlarında ele alınan hipotezlerin anlaşılması için, hastalık olaylarının dağılımında mekânsalveri analizi yapmak gerekir. Dağılım profillerinin farklı sınıflamaları için rastgeleliği dikkate almak lazımdır. Çünkü, bu durumların dağılımı referans popülasyon gerektirir. Bu dağılım, diğer hastalıkların ekstrapolasyonuyla, örneklemeyle ve derecesinin değişimiyle bilinebilir. Genelde, risk faktörleri gibi başka bileşenlerle zaman ve uzayda heterojen olarak dağılır. x uzaydaki koordinatı, t zamanı, z,$X \times T \times Z$ kümesi üzerinde standart ölçümle ilgili vektör değişkeni göstermek üzere $\lambda_0(x,t,z)$ yoğunluk fonksiyonuyla tanımlanabilir. Hastalığın doğal bulaşımı gibi epidemiyolojik olayın karmaşık araştırmasının hipotezi için, bireyler arasındaki temas yapısının kesinliğiωve ω' üç değişkenli fonksiyonlar olmak üzere $\rho_0(\omega,\omega')$ değerini belirleyen x,t,z değişkenleri için gereklidir. $\rho_0$fonksiyonu, verilen $\omega=(x,t,z)$için $\omega'=(x',t',z')$değerini belirleyen olasılık dağılımını gösterir.

Yan yana temelli bir nokta sürecinin, $\lambda_1(x,t,z)$ bireysel yoğunluğu ve/veya $\rho_1(\omega,\omega')$etkileşim yapısı vardır. "Ele alınan nokta süreci temel popülasyon (yoğunluğu

ve/veya etkileşimi) ile benzer yapıya sahip mi ?" "Farklılıkların gözlemlendiği veya istatistiki olarak aykırılık olan $X \times T \times Z$ bölgesi belirlenebilir mi ?" soruları sorulabilir. Bu soruların cevapları parametrik olmayan testlerin kullanımı ile mümkündür ve bu testler yerellikteki anomaliden önce global anomalinin varlığını sınar.

Böüm 2'nin 2.1.5 alt bölümünde, yayılım ve korelasyon analizlerinin mekansal dağılımlarda önemli rol oynadığını söylemiştik. Örneğin, düzgün dağılım birbirinden aynı uzaklıktaki yaklaşık yerleşmiş durumları gösterir. Eğer birim yüzeyde olayların sayısını rastgele değişken olarak alırsak davranışı Binom dağılımı olarak modellenebilir. Böyle bir dağılımın varyansı, beklenen değerinden daha küçüktür. Önceki metottan, olayların meydana geldiği uzaya bağlı olan hastalıkların değişiminin riskini kolayca göstermek mümkündür. Dolayısıyla bu olay mekânsaldır.

5.1.3 Riskin Mekânsal Değişimi

Epidemiyolojide, sağlık çalışmalarındaki mekânsal değişim göstergesi ve çevresel maruz kalma faktörü (riskleri) çalışılan olayların tanımlanarak etiyolojik hipotezler önerilmesine izin verir [141] . İstatistiki olarak, riskin mekânsal değişimi yoğunluk fonksiyonuna bağlıdır. $x \in X \subseteq S$hastalığın mekansal pozisyonu ise bu olayın olasılık yoğunluk fonksiyonu ile kontroller orantılı değildir.$r(x)$, $x$pozisyonundaki hasta kişinin olasılığını göstersin. Bu durumda, J. P. Diggle [142] tarafından tanıtılan geleneğe göre, "kontrol fonksiyonunun var olan durumun fonksiyonu olmamalı" koşulunu ileri sürer. Dolayısıyla kontrol ve mevcut durumlar için olan fonksiyonlar aşağıdaki gibi verilir:

$$\lambda_1(x) = r(x)\lambda(x) \tag{5.1}$$

$$\lambda_2(x) = C(1 - r(x))\lambda(x) \tag{5.2}$$

Burada, $\lambda(x)$ele alınan popülasyonun yoğunluğu, C çalışılan hipotezlere göre değer alan bir sabittir. Buradan, ancak ve ancak $r(x)$ sabitse $\lambda_1(x)$ ve$\lambda_2(x)$ fonksiyonlarının orantılı olacağı görülür. Bu nedenle, $r(x)$fonksiyonu risk yüzeyi adını alır.

Mekânsal kümelemede mekânsal risk değişimi, bölüm 5.1.1 de görüldüğü gibi, coğrafi bölgenin hastalık olaylarının olaylar arasında ilişki kurmanın zor olduğu coğrafi bölgede hastalar olayları bağımsız olarak meydana gelen şeklindeki hipotezlerin bağımsızlığı altında tanımını gösteriyor. Gerçekte, heterojen çevrede bağımsız bir olay ile homojen

çevredeki bir olay arasında ampirik ayırımı kurmak zor veya imkansızdır. Eğer $r(x) = r$ sıfır hipotezi ise, o halde mekânsal risk değişimi yoktur. Böylece, kontrol ve olayların yoğunluk fonksiyonları rastgele dağılırlar (mekânsal kümeleme değil). Çünkümekânsal kümeleme ve riskteki değişim aynı sıfır hipotezine alternatif oluşturur.

N_{X_i}birimmekanda gözlenen olayların sayısı, $Y_i$olayların beklenen değeri ve $\theta_i i$. birimin göreceli riski olsun. Maksimum göreli risk tahmincisinin kararsızlığı Poisson dağılımı ile modellenebilir: $N_{X_i} \sim Pois(Y_i, \theta_i)$

θ_i nin maksimum olabilirlik tahmincisi ve varyansı sırasıyla,:

$$\hat{\theta} = \frac{N_{X_i}}{Y_i}, Var(\hat{\theta}) = \frac{N_{X_i}}{{Y_i}^2} \quad (5.3)$$

olarak verilir. Küçük birimler veya değişmeyen risk tahmininden büyük standart ölüm oranı (Standard morbidity ratio-SMR) ile ilgili varyans olabilen küçük değerli birimler gözlemleriz. SMR nin değişkenliği farklıdır ve mekânsal birimlere bağlıdır. Ek olarak, yoğunluk yüzeyi koşulu ve coğrafi bölgede olayların bağımsız olarak meydana geldiği kabulü altında, olayları gösteren harita homojen olmayan Poisson sürecinde oluşur. Kontrol yoğunluk fonksiyonu dizayndan bağımsız olarak meydana gelir.$\lambda_2(x)$ yoğunluk yüzeyi üzerindeki koşul ve kontrol tasvirleri ikinci bağımsız Poisson sürecini oluşturur. Böylece, (5.1) ve (5.2) den aşağıdaki yazılır:

$$\lambda_1/\lambda_2 = C^{-1} r(x)/\{1 - r(x)\} \quad (5.4)$$

Bu, $r(x)/\{1 - r(x)\}$hasta olasılığıyla sabitin çarpımı ve risk yüzeyini, parametrik olmayan şekilde tahmin edilir.(yani iki yoğunluk fonksiyonlarının tahmini) . Nokta süreci modelleri kullanılarak gelişen, ekolojik olaylar ve/veya çevreyle ilgili risk faktörlerinin gelişiminin zaman ve uzaydaki hasta dağılımının analizinde kullanılan mekansal epidemiyolojini bu kısa giriş bilgisinden sonra, bunu sıtma hastalığına uygulayacağız. Kullanılan verinin kesikli olabildiğini gözlemledik. Bu çalışmayı yürütmek için, bu hastalığın CAR'da rastgele dağıldığını ve zaman ve uzayda açıkça ortaya çıktığını kabul ediyoruz.

5.2 Uygulama: Orta Afrika Cumhuriyetinde Sıtma Hastalığı Dağılımı Çalışması

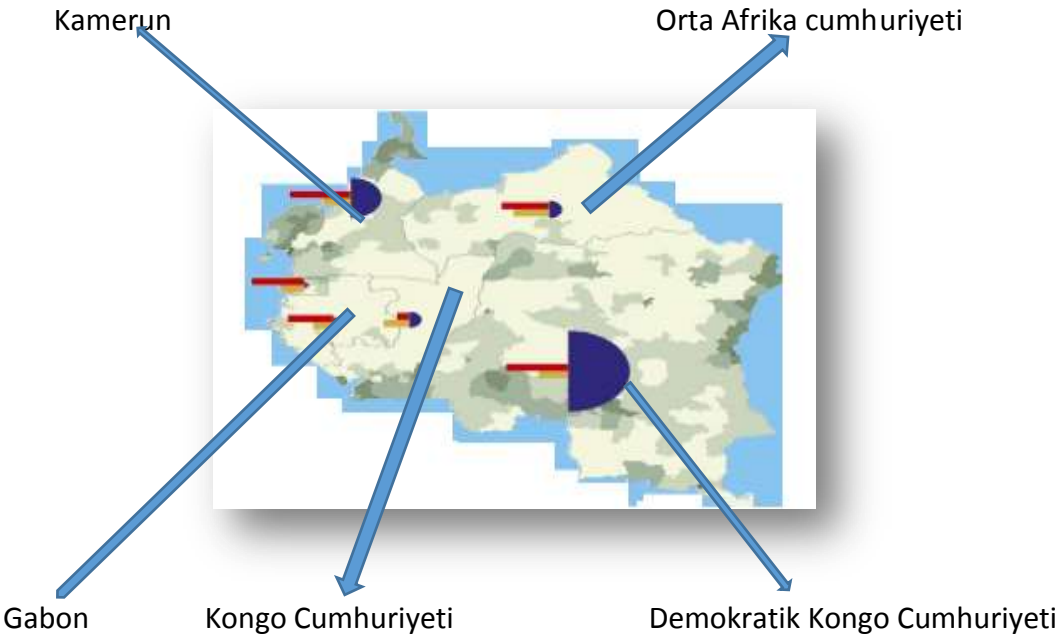
Sıtma, Afrikanın Sahra bölgesinin alt kısımlarında yaygın tropical bir hastalıktır. Bu bölgede plasmodium falciparum, plasmodium vivax, plasmodium ovale, plasmodium malaria ve knowlesi gibi bu hastalığa sebep olan pek çok patojen vardır. Dünya Sağlık Örgütü (WHO)'ya göre sıtma hastalığının Orta Afrika'daki en çok görülen türü plasmodium malariae'dir.

Bu bölümde, mekânsal ve uzay-mekânsal nokta süreci yöntemleri kullanılarak CAR'da sıtma hastalığınındağılımının analizi ile ilgileneceğiz. Afrika kıtasının tam ortasında yer alan Orta Afrika Cumhuriyeti, 623.000 Km^2 yüzölçümüne sahiptir. Doğusunda Sudan (175km) ve Güvey Sudan (990km), kuzeyinde Çad Cumhuriyeti (1197km) , batısında Kamerun (797km) ve Kongo Cumhuriyeti (467km), güneyinde Demokratik Kongo Cumhuriyeti (1577 km) vardır. Ülkenin denize kıyısı yoktur ve 2015 tahminlerine göre nüfusu 5.391.539 dir. Dolayısıyla kilometre başına 8.6 kişi düşecek şekilde nüfus yoğunluğunun az olduğu bir ülkedir. [http://www.statistiques-mondiales.com/centrafrique.htm]. Orta Afrika Cumhuriyetinin, coğrafi yeri ise 7.00 kuzey enlemi ve 21.00 Doğu boylamıdır. Sıtma hastalığının yayılmasında etkili olan Anofel sivrisineğinin üremesine uygun tropik bir iklime sahiptir. Hastalıktan etkilenmiş dişi anofel sivrisineği bu hastalığı ısırmak sureti ile insanlara bulaştırmaktadır.



Şekil 5. 2 Orta Afrika Cumhuriyeti ve komşuların haritası

Afrika kıtasının alt bölgesinde olmasından dolayı, Orta Afrika cumhuriyeti Kamerun, Kongo Cumhuriyeti, Demokratik Kongo Cumhuriyeti ve Gabon Cumhuriyeti gibi ülkelerle benzer iklim şartlarına sahiptir. Bu ülkeler de dişi Anofel sivrisineğinden kaynaklanan benzer epidemik problemler yaşarlar. Sözü edilen bölge, Mart ayından Ekim ayına kadar süren yağış sezonunda (Orta Afrika Cumhuriyeti'nin güneyinde) sıtma hastalığının geçiş sezonuyla hiper endemik bölgedir. Bu sezonda, ortalama sıcaklık 15-32 ° C aralığındadır ve hastalık geçişleri yüksek seviyededir. Bu geçişler, kuru dönemin ortasında (Ocak ayı gibi) minimum seviyeye düşer. Doğal koşulların yanı sıra (iklim gibi), nüfusun sosyokültürel aktiviteleri (nehrin kıyılarında tarım yapmak, hijyenik olmayan koşullarda bina malzemeleri üretmek) gibi çevresel koşullar da hastalık için uygun koşullar oluşturmaktadır. 2014'de Dünya Sağlık Örgütü WHO' nun yayınladığı rapora göre, sıtma hastalığı dünyayı tehdit etmeye devam etmektedir. Dünya çapında, özellikle bulaşma riskinin yüksek olduğu ve parazit etkilenmesine duyarlı olan Orta Afrika bölgesi dahil olmak üzere 3.3 milyar nüfus toplamına sahip 97 ülkede 1.2 milyar kişi bu hastalığa yakalanma riski altındadır (Bir yıl boyunca sıtma hastalığından etkilenme şansı 1/1000 den büyüktür.) Dolayısıyla, CAR'ın halk sağlığı açısından sıtma önemli bir kaygı sebebidir. Sosyal yardım kuruluşlarıyla beraber, Orta Afrika Cumhuriyeti hükümeti kıtada ölümcül hastalık olan sıtma hastalığının kökünü kurutmak için olağan üstü çaba harcamaktadır.



Şekil 5. 3 Orta Afrika'da sıtma ölümünün demografik yoğunluğu

Kullanılan veri, http://knoema.com/WHOWMS2013/who-world-malaria-statistics-2013 adresinden elde edilmiştir ve WHO tarafından sağlanan verilerdir. Bu verinin kaynağı, WHO, Orta Afrika İstatistik Enstitüsü, Sosyal ve Ekonomik Çalışmalar (CAISSES)'dir. Bu veri, 2002-2013 yıllarına ait Orta Afrika Cumhuriyetinin sağlık merkezleri ve farklı hastanelerinden elde edilmiş 12 yıllık verilerdir. Veriye göre, ele alınan ülkenin yaklaşık %5 'I sıtma hastasıdır ve çoğu 5 yaşından küçük çocuklar olmak üzere ortalama nüfusun % 0.5 'i sıtma hastalığından ölmektedir.

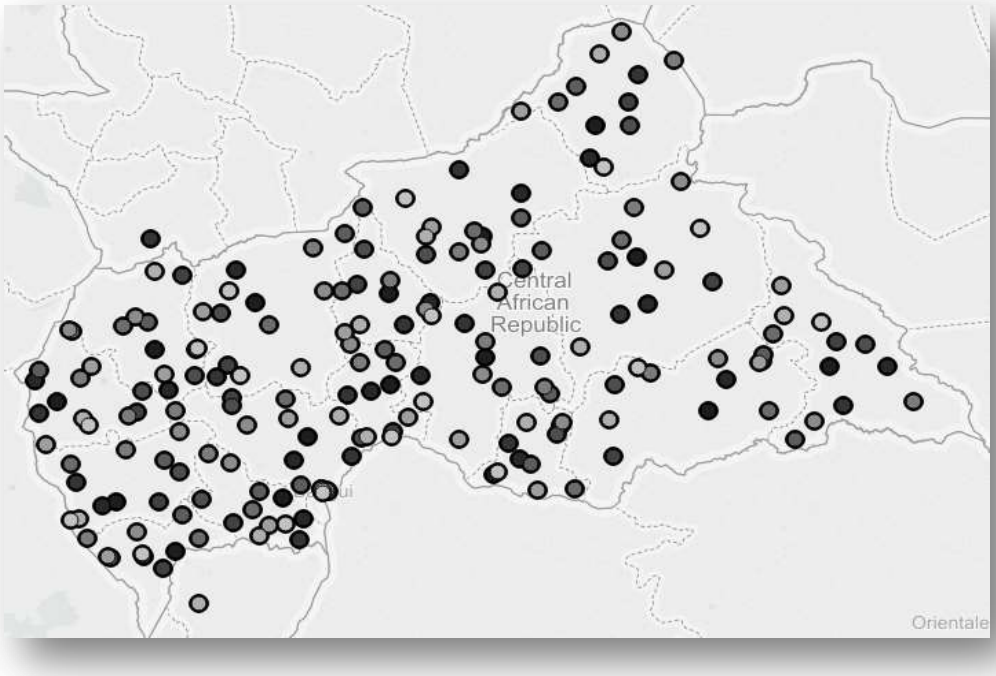
Çizelge 5. 1Orta Afrika Cumhuriyetine ait sıtma verisinin özeti

No	Yılları	Sıtma sayısı	Ölüm sayısı	%S	%Ö	Nüfus sayısı
1	2013	491074	1026	9.51	0.21	5166510
2	2012	459999	1442	9.95	0.313	4623604
3	2011	221980	858	4.92	0.39	4510833
4	2010	66484	526	1.51	0.8	4400813
5	2009	175210	667	4.10	0.4	4310296
6	2008	152260	456	3.61	0.3	4221642
7	2007	119477	578	2.90	0.484	4134811
8	2006	114403	865	2.83	0.724	4049766
9	2005	131856	668	3.33	0.51	3966479
10	2004	129367	859	3.32	0.664	3893595
11	2003	78094	417	2.01	0.534	3895139
12	2002	43093	535	1.14	1.242	3767248

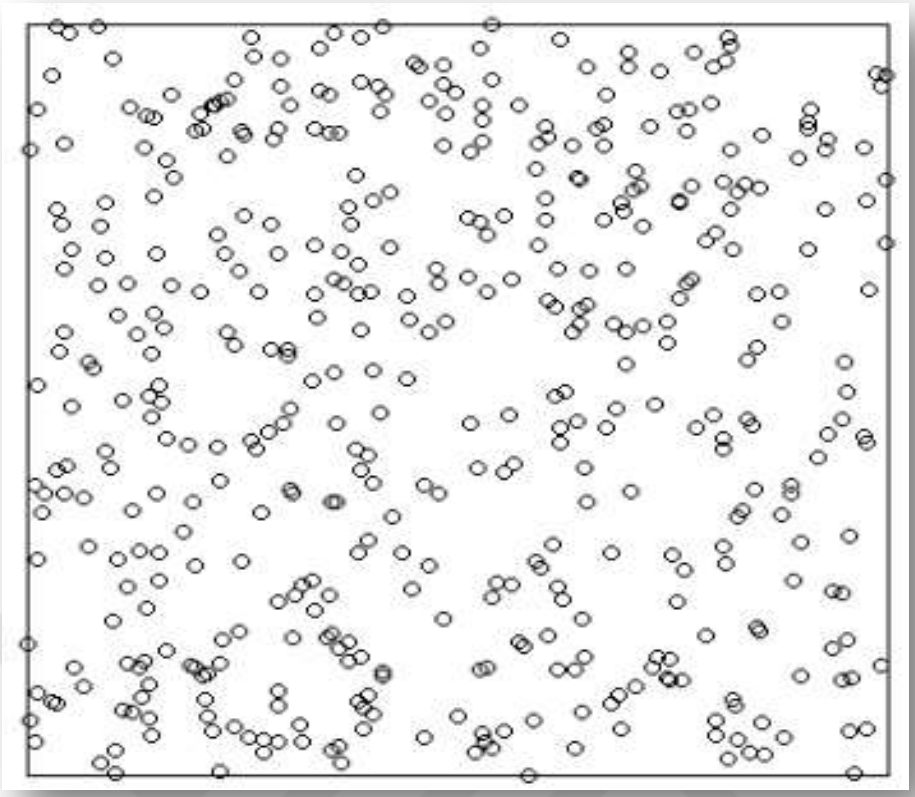
Yukarıda tartışılan nokta süreci ve uzaysal dağılım yöntemleri kullanılarak, ele alınan bölge üzerindeki dağılımını rastgele gruplama şeklinde olduğu görülmektedir. Dolayısıyla sıfır hipotezi olarak bölüm 2.1.5 'e göre MORAN endeksi $I_m = 0$ alınır. Ayrıca, sıtma hastalığının dağılımının nufus yoğunluğu, iklimsel ve çevresel faktörlere

de bağlı olduğunu tekrar vurgulayalım. Epidemiyoloji alanı üzerindeki tüm sayısal veriler, kümeleme dağılımını kurmak için gereklidir. Elimizdeki veri, saha başına veya bölge başına olayların dağılımını göstermemektedir, dağılım tüm ülke için verilmiştir. Bunlara ek olarak, verinin yapısından dolayı, öncelikli ve önemli bölgeleri içeren mekânsal dağılım göstergesi yoktur. Fakat son 12 yıl boyunca bu hastalığın insidansını (her yıl yeni hasta olma olayı) Orta Afrika Cumhuriyeti ülkesindeki demografiye ve politik ve askeri krizlere dayanan verilerden, görmek mümkündür.

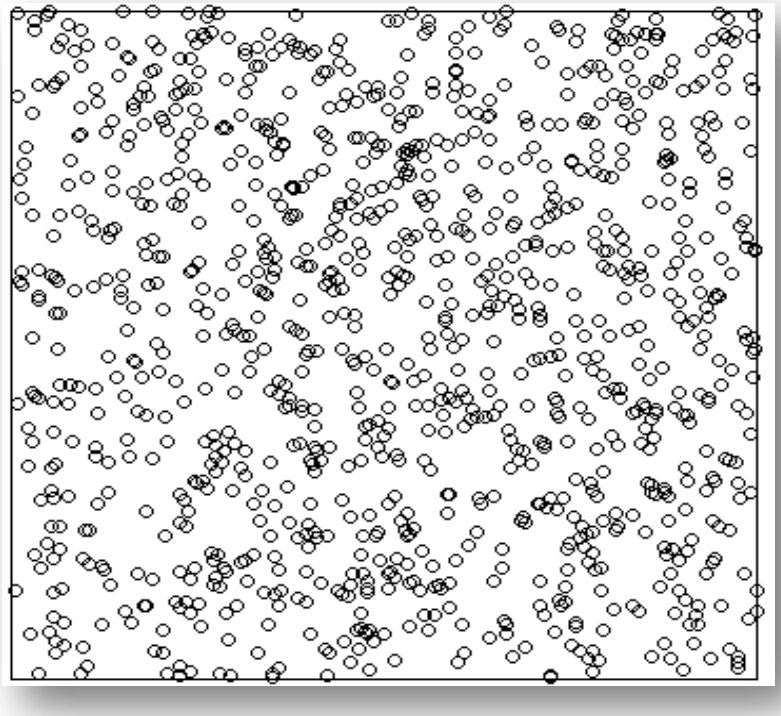
3.1.2 sürümüne sahip R istatistik yazılımının spatsat, mcmc ve Rcmdr paketlerini kullandık. (The R foundation for statistical computing, X86_64-w64-mingw32/X64, 64-bit, 2014-10-31). Bunlar http://CRAN.R-project.org web sayfasından, dağılımı harita üzerinde göstermeye yarayan araçlar olan Tableau public 9.2 yazılımı ise https://public.tableau.com/s/ web adresinden ücretsiz elde edilebilmektedir. Şekilleri üretmek ve simüle etmek için, windows işletim sisteminde $W = X \times Y \subseteq \mathbb{R}^2$mekânsalgösterilimini ve$W' = X \times Y \times [0,12] \subseteq \mathbb{R}^2 \times T$üzerinde de grafik olarak zaman-mekânsal gösterilimi kullandık. Böylece elimizdeki sıtma hastalığı verilerinin kümeleri ile aşağıdaki mekânsal ve zaman-mekânsal sonuçlar elde edilmiştir:



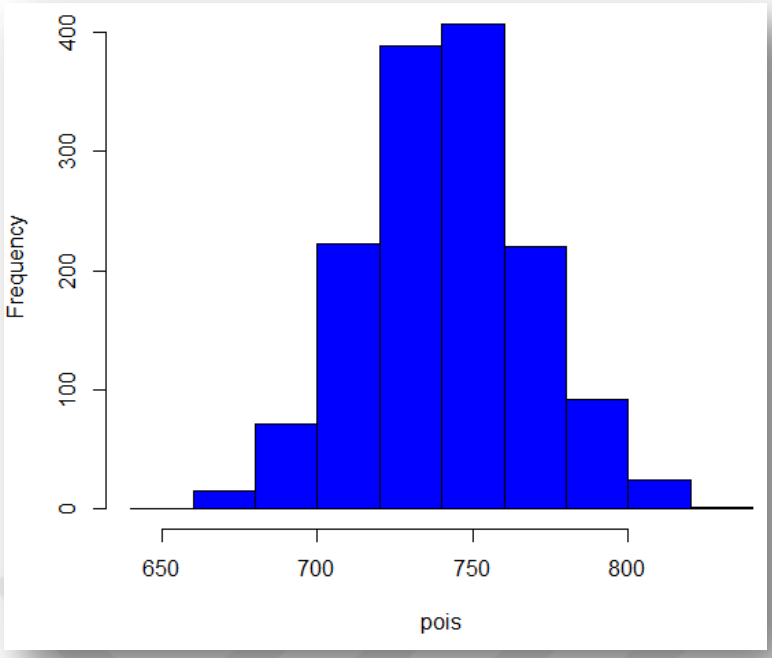
(a) Hastalığın ülke coğrafyandaki 12 yıllık dağılımı (bir nokta 12 hastayı gösterir).



(b) $\lambda = 742$, parametreli Poisson dağılımı (her nokta 12 ölüm göstermektedir)

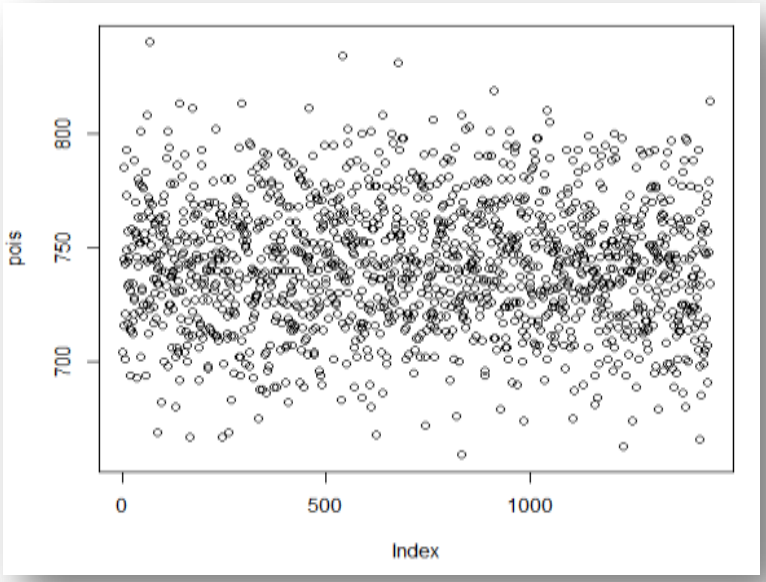


(c) $\lambda = 181941$, parametreli Poisson dağılımı (her nokta 12 hastayı göstermektedir)

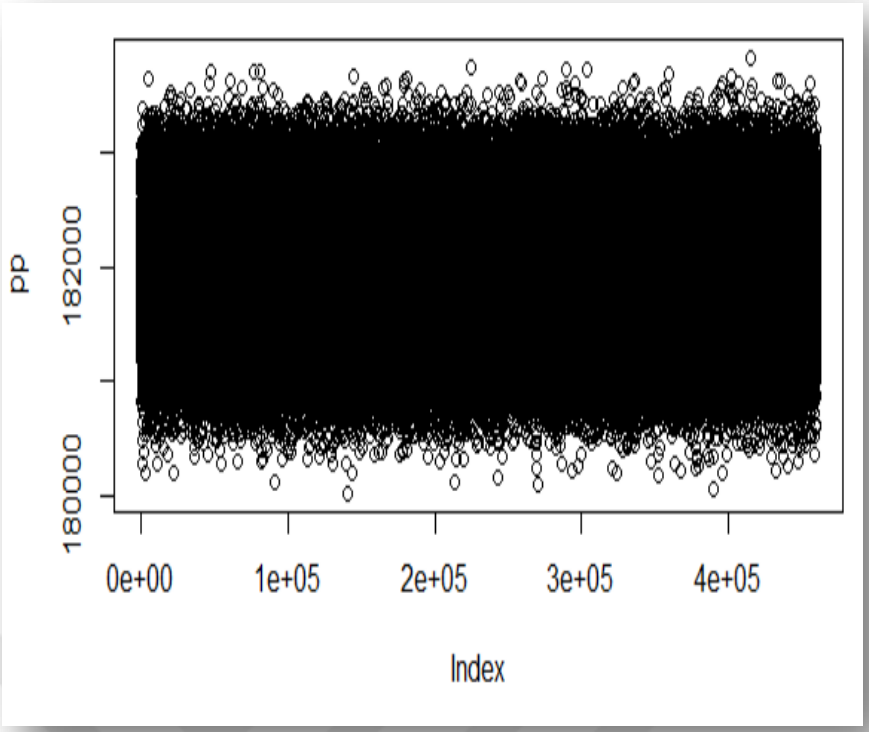


(d) Sıtma hastalığın histogramı ($\lambda = 181941$ Poisson süreci gerçekleşmesi)

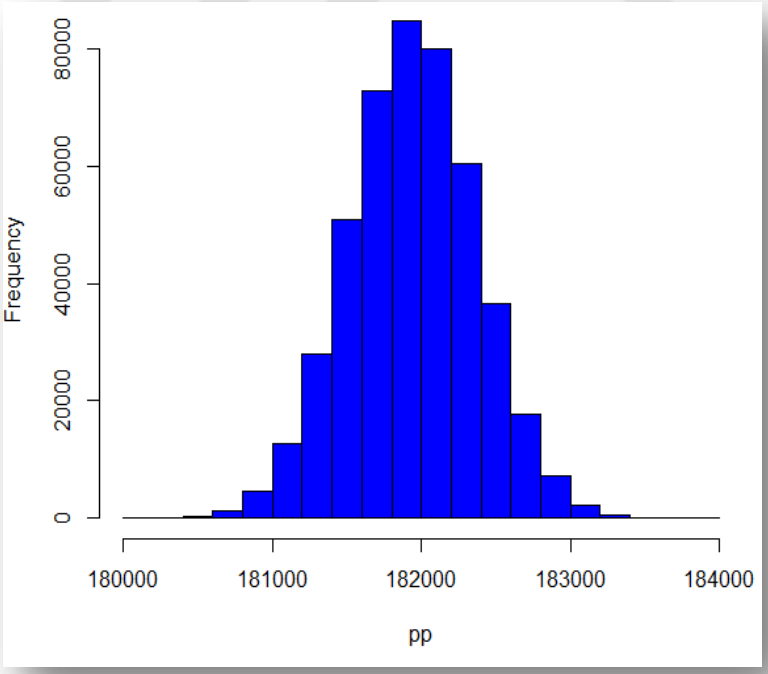
Şekil 5. 4 Nokta süreçleri dağılımı grafikleri, (a) Orta Afrika Cumhuriyetindeki 12 yıllık, sıtma hastalığının rastgele zaman-mekânsal dağılımı (b) Poisson nokta süreci dağılımı ölüm gerçekleşmesi, (c) Poisson norkta süreci sıtma hastalığın gerçekleşmesi, (d) Sıtma hastalığın Poisson histogramı



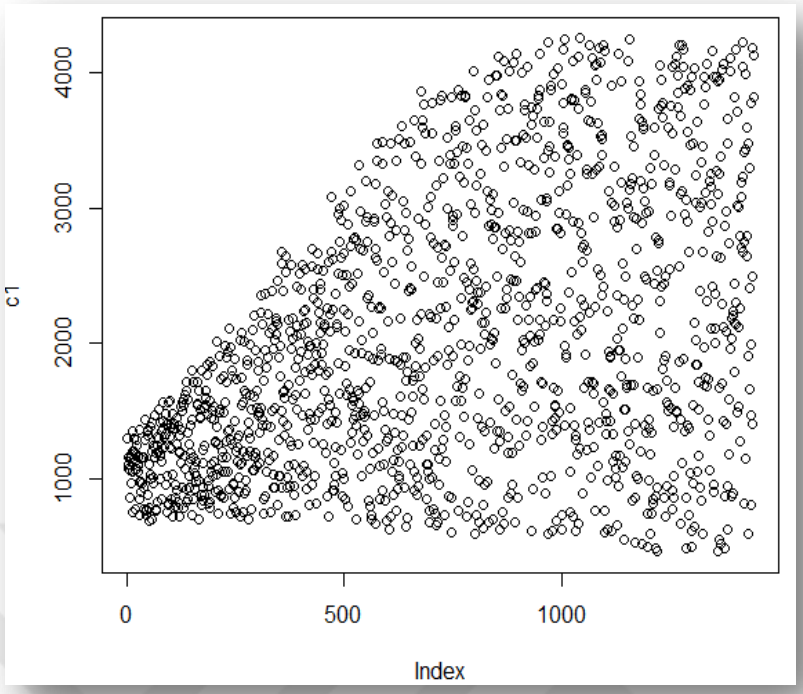
(a) $\lambda = 742 \times 20$=14840, (Poisson süreci), sonraki 20 yıl ölüm tahmininin dağılımı



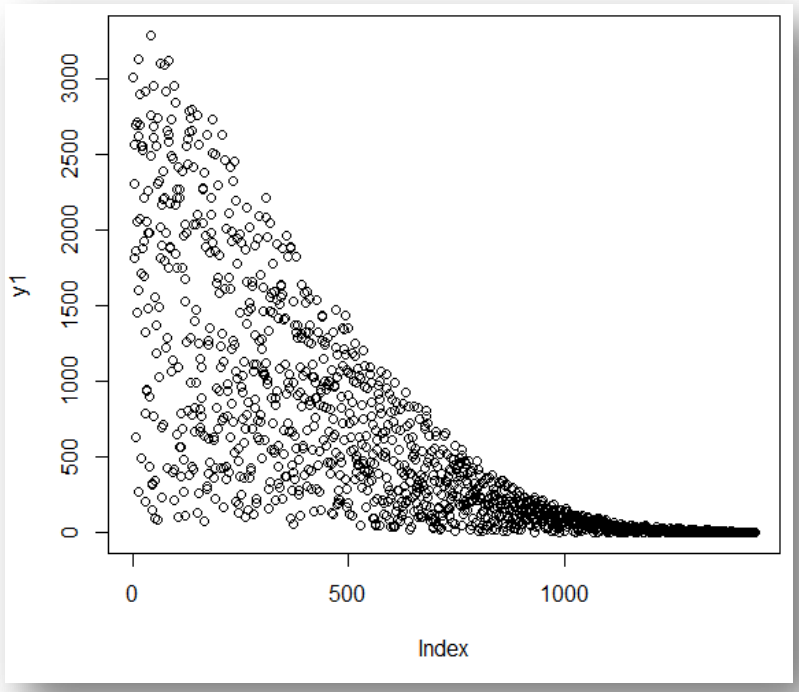
(b) $\lambda = 181941 \times 20 = 3638820$, parametreli Poisson dağılımı, sonraki 20 yıl sıtma hastalığı tahmininin dağılımı



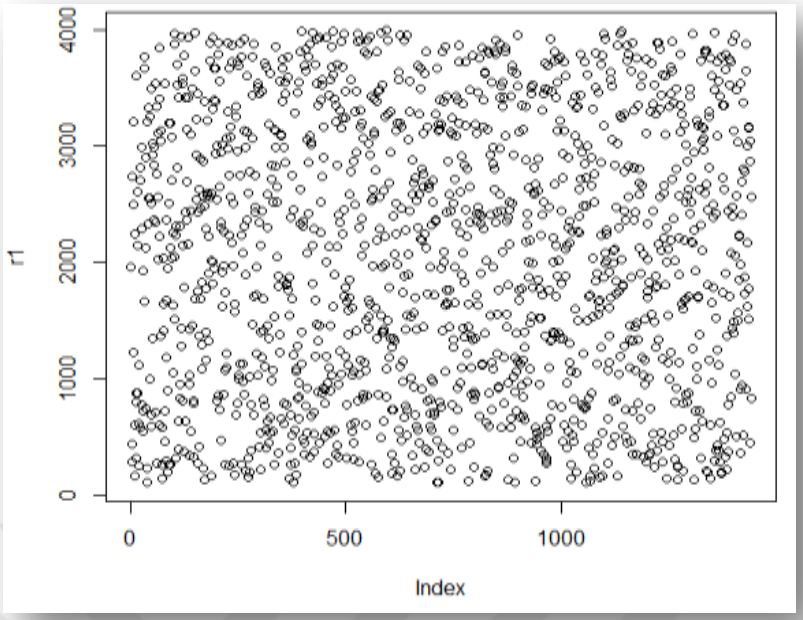
(c) $\lambda = 3638820$, Sıtma hastalığı Poisson histogramı, sonraki 20 yıl tahmininin histogramı



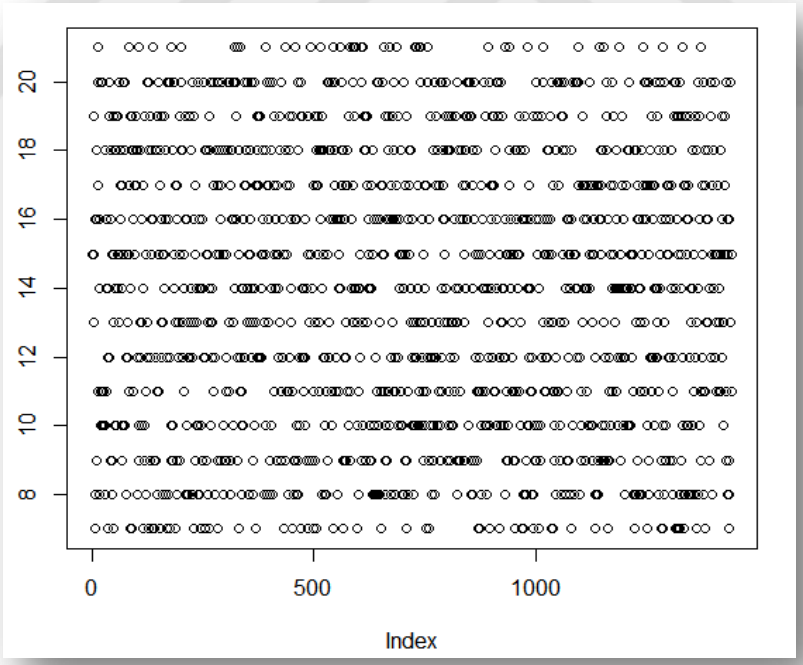
(d) $\lambda = 181941$,parametreli sağa çarpık ile sıtma hastalığı homojen olmayan Beta dağılımı



(e) $\lambda = 742$, sola çarpık ile ölüm homojen olmayan Teta posterior dağılımı

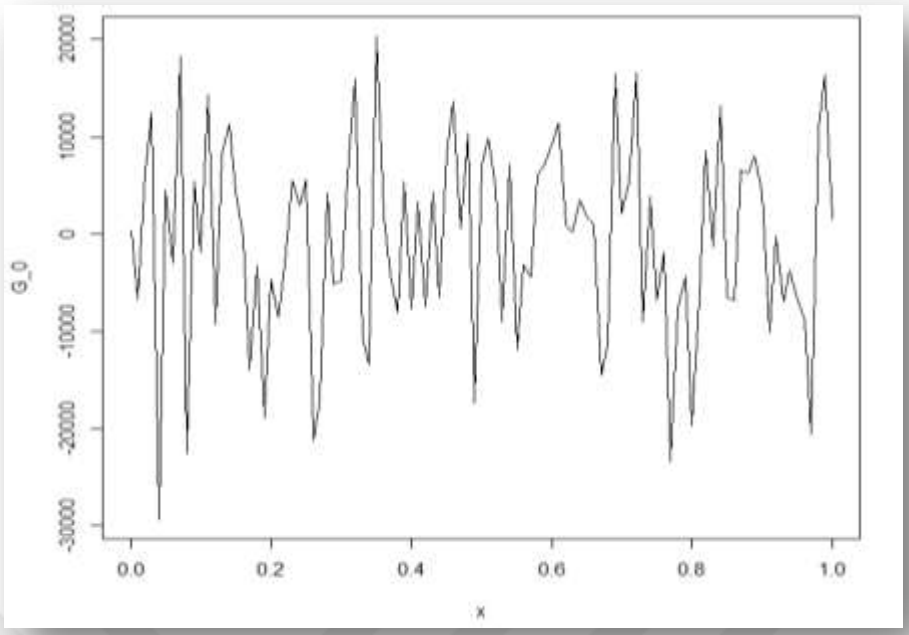


(f)) $\lambda = 181941$, parametreli homojen Poisson dispersiyonu

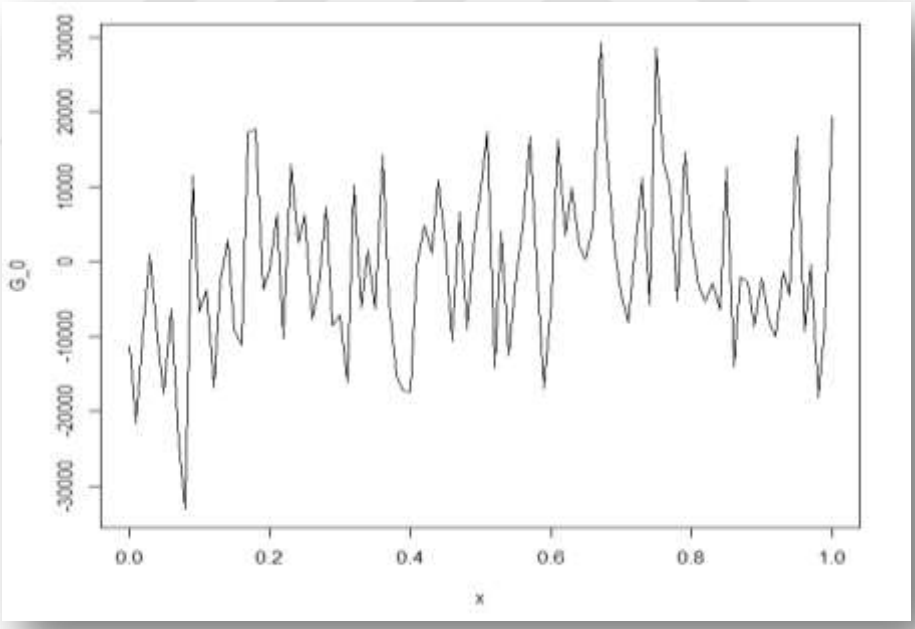


(g) $\lambda = 181941$, normal dağılımıyla sıtma hastalığın dispersiyonu

Şekil 5. 5 Sıtma hastalığının, farklı nokta süreçleri ile (a)rastgele mekânsal dağılımın, (b) Poisson nokta süreci, (c) Sıtma hastalığı Poisson histogramı(d) Beta dağılımı; (e) teta posteror dağılımı; (f) Binom ve Poisson dağılımının yoğunluk fonksiyonları ile (g) Normal dağılımı



(a) $\lambda = 742$, Dirichlet sürciyle ölüm gürütüsü



(b) $\lambda = 181941$, Dirichlet süreciyle sıtma hastalığın gürütüsü

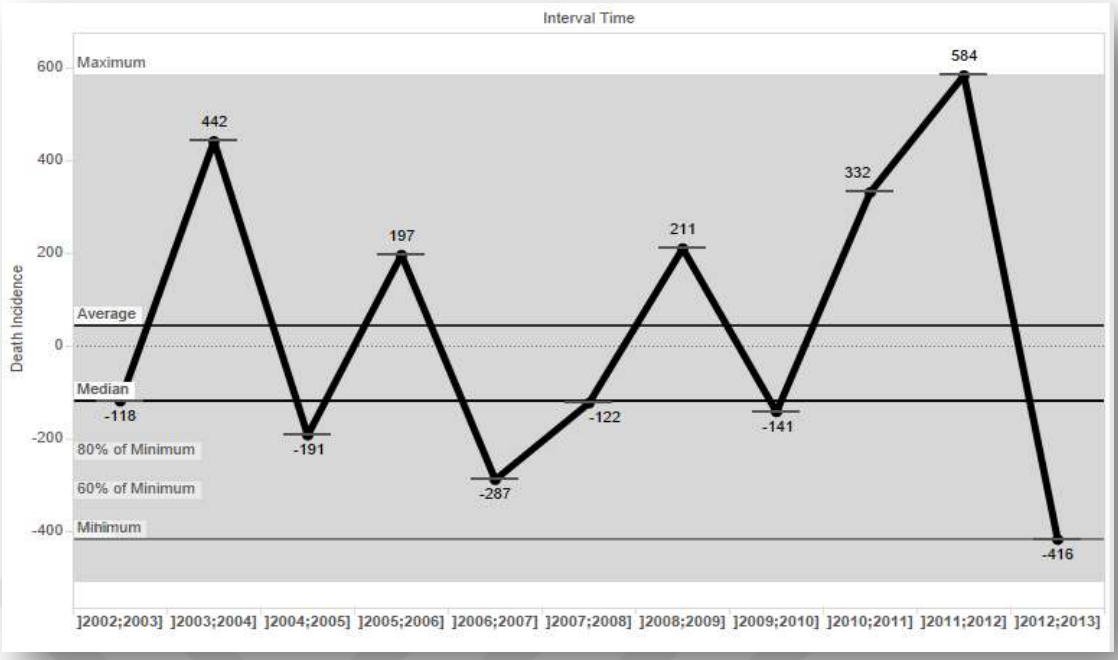
Şekil 5. 6 Dirichlet süreciyle Gibbs-Metropolis-Hastings hybrid işlemleriyle elde edilmiş MCMC sonuçlar

Çizelge 5. 22011 ve 2012 yıllarına ait olayların dağılımı, bu çalışma başkent Bangui'de yapılmış

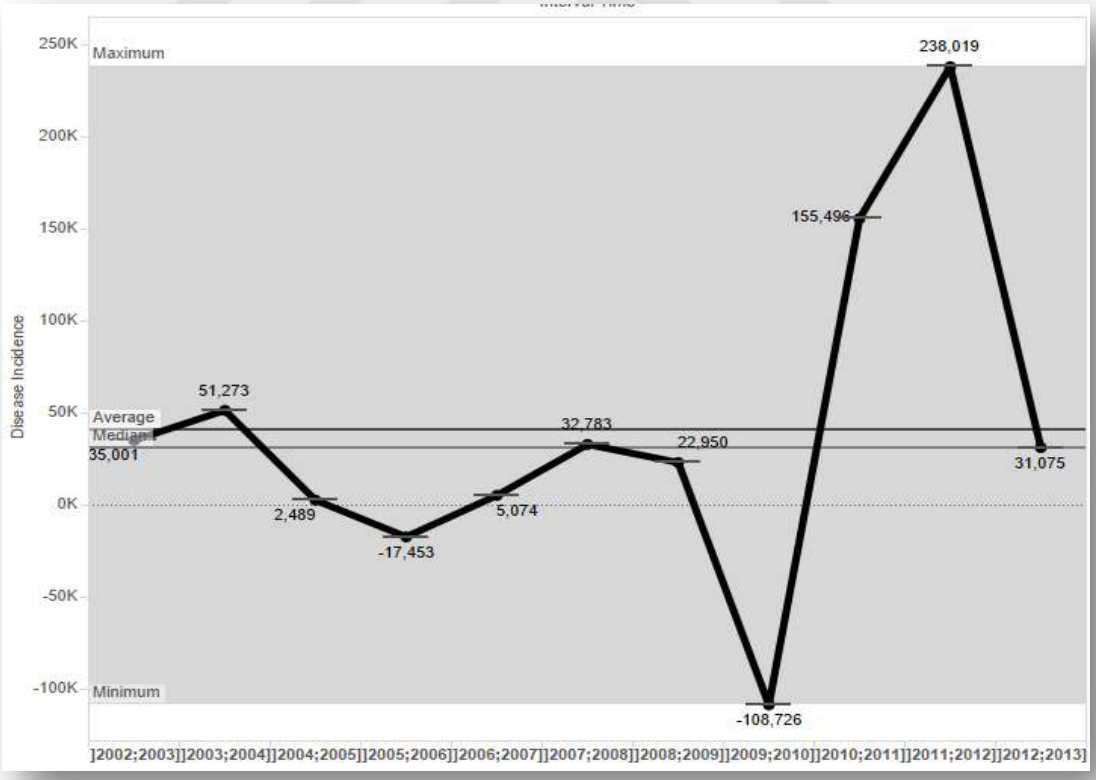
Durumları (A=yaş)	2011	2012	%/2011	%/2012
Basit(A<5 yaş)	101178	175950	45.58	38.25
Ciddi(A<5 yaş)	29020	75362	13.073	16.383
Basit(A>5 yaş	76894	157673	34.64	34.277
Ciddi(A>5 yaş)	14888	51014	6.707	11.09
Bölge Dışından hasta(A<5 yaş)	48181	199149	56.82	54.98
Bölge dışından hasta(A>5 yaş)	36615	150359	43.18	45.02
Yatan hastaa(A<5 yaş)	7870	30719	61.13	63.24
Yatan hasta(A>5 yaş)	5004	17857	38.87	36.76
Doğru tedavi (A<5 yaş)	42755	164681	56.8	54.35
Doğru tedavi(A>5 yaş)	32518	138320	43.2	45.65
Terapetik başarısızlık (A<5 yaş)	15793	6923	70.31	47.33
Terapetik başarısızlık (A>5 yaş)	6670	7.704	29.69	52.67
Ölüm(A<5 yaş)	725	1248	84.6	86.6
Ölüm(A>5 yaş)	133	194	15.4	13.4

Çizelge 5. 3hastalığın insidansı (hastalıktan her yıl etkilenen yeni hasta sayısı)

Zaman aralığı	Hasta insidansı	Ölüm insidansı
]2002;2003]	35001	-118
]2003;2004]	51273	442
]2004;2005]	2489	-191
]2005;2006]	-17453	197
]2006;2007]	5074	-287
]2007;2008]	32783	-122
]2008;2009]	22950	211
]2009;2010]	-108726	-141
]2010;2011]	155496	332
]2011;2012]	238019	584
]2012;2013]	31075	-416



(a) Ölüm insidans grafiği



(b) Sıtma hastalığı insidans grafiği

Şekil 5. 3(a) sıtma insidansı grafiği, (b) ölüm insidansı grafiği

Çizelge 5. 4 İstatistiksel bilgiler

Durumlar	Ortalama	Medyan	standart sapma	min	max	1st Qu	3rd Qu
Sıtma hastası	181941	130611.5	145395.3	43093	491074	105326	186903
Ölüm	742	667.5	290.4305	417	1442	532.8	860.5
Hasta insidansı	40726	31075	90083.67	-108726	238019	3782	43137
Ölüm insidansı	51.1818	-118	319.2857	-416	584	-166	271.5

5.3 Analiz ve Tartışma

Sıtma hastalığın geçişinin zaman ve uzayda yayılması nokta süreçleri olarak ele alınabilir. Bu zamansal ve mekânsal olayların modellemesi, bu çalışma alanında bir pandemik olduğunu ortaya koymaktadır. Bu model, hemen hemen tüm alanlarda kullanılan güçlü matematiksel araçlar sağlar ve olayları istatistiksel amaçlı analiz edebilir. Aynı zamanda, halk sağlığı problemlerinde yol gösterici bir araçtır. Resmi makamlara dikkatli olmaları yolunda işaret veya uyarı sistemi olarak da gösterilebilir.

Şekil 5.3, ele alınan bölge boyunca dağılmış hastalığı, Poisson süreci ile göstermektedir. Bu rastgele yeniden ortaya çıkma, sıtmanın CAR'ın her coğrafi bölgesini etkilediğini gösterir. Bu yaygınlık, geçiş elemanı olan (patojen) anofel sivrisineklerinin hemen hemen hergün bu bölgede bulunmalarıyla açıklanabilir. Harita üzerinde her bir nokta verilen periyotta ve coğrafi alanda virüslü bireyi temsil eder. Dolayısıyla, sözü edilen haritadan CAR'ın güney kısmında bu hastalığın yoğunluğunun yüksek olduğu görülür. Bu, Poisson sürecinin yoğunluğuyla ilgili bölgenin populasyon yoğunluğunun da fazla olmasına bağlanabilir. Poisson yoğunluğu çok geniş olduğunda, uzayda noktaların bir doygunluğu vardır. Bu durumda hangi noktaların hasta bireyleri gösterdiğini belirlemek zordur. Şekil 5.4 (a) ele alınan uzaydaki noktaların yoğunluğunu, dolayısıyla az önce sözü edilen zorluğun mevcudiyetini göstermektedir. Bu zorluk, Possion sürecinin bağımsızlık kabulünden dolayı ortaya çıkmaktadır. Yine de, verilerimiz diğer kümeleme yöntemlerine uygun olmadığından, nokta süreçlerinin temel süreçlerinden olan

Poisson süreci kullanılmıştır. Diğer taraftan, yoğunluğu çok olduğunda noktalarının göstterilimindeki yavaşlık kanıtlanmıştır. Sıtma yayılımının uygulamasında kullanmadığımız (verilerimizin uygun olmamasından dolayı) Cox süreci gibi diğer süreçler bu durumda işe yarayabilir.

Çalışmamızda, 5000 iterasyonlu Dirichlet süreci simüle edilmiştir. Bu süreç, t zamanına bağlı olan sıtma hastalığının frekansını göstermektedir. Hastalık, pandemik döneminde büyük ölçüde artıp ve diğer zamanlarda azalmasından dolayı frekansın zamana bağlı olduğunu söyleyebiliriz. Şekil 5.5, bu hastalığın frekansının rezonansını göstermektedir.

Çizelge 5.2, 2011-2012 dönemindeki olayların dağılımını göstermektedir. Bu dönem, ülkede yaşayan nüfusun sağlıklı olmayan çevresel şartlarda bulunmalarını zorunlu kılan krizin başlangıcı olan duyarlı bir dönemdir. Sivrisineklerden korunmak için elverişli koşullarda olmadıklarından dolayı, popülasyon bu hastalığın yayılmasına sebep olan sivrisineklerin ısırmalarına maruz kalmışlar ve yüksek oranda çocuk ölümleri meydana gelmiştir. Çizelge 5.2'de çocukların daha zayıf olmalarından kaynaklı hastalık ve ölüm oranları görülmektedir. Örneğin, 2011'de sıtma hastalığından ölenlerin %84.6 sı 5 yaş altı çocuklardır. Bu sayı 2012'de %86,6 olarak bu hastalıktan ölümler için riskin en büyük değerine ulaşmıştır.

Çizelge 5.3 ve Şekil 5.6 sıtma kaynaklı hastalık ve ölümün insidansı göstermektedir. Buradan, azalmış sayılar eğrisi zaman zaman önemli ölçüde arttığı kolayca gözlenebilir. Her yıl ortaya çıkan yeni olaylar olarak hastalığın insidansını tanımlarsak, hükümet ve sosyal ortaklarının uğraşılarına rağmen bu hastalığın hala bir endişe kaynağı olarak kaldığını görürüz. Çalışmamızın konusu olan 12 yıl (2002-2013) verilerini kullanırsak, sıtma nedeniyle teyit edilmiş 447,981 yeni olay ve kaydedilmiş ölüm olaylarının 1,025 olduğu görülür. Bu da, sözkonusu hastalığın halk sağlığı açısından hala kalıcı bir tehdit olmaya devam ettiğini söylemektedir.

Bu nedenlerle, sağlık otoritelerinin, söz konusu problemlerin önleme ve araştırmasında nokta süreçleri gibi matematiksel araçlarını dikkate almaları gerekmektedir.

BÖLÜM 6

SONUÇ VE ÖNERİLER

Zaman ve uzayda ortaya çıkan problem ve olguların çalışılması için bunların gerçekleştikleri konum bilgisi gereklidir. Uzay ve zamanda ortaya çıkan fenemonolojik problem ve durumlar ile çalışmak için konumlarını bilmek gerekir. Dolayısıyla, gözlenen olayların uzayda oluştuğu konumların belirlenmesi çalışmanın ilk adımıdır. Çalışılan olayların etkili tahmini ve tarifini yapabilmek için olayların oluştuğu zaman ve pozisyon çok önemlidir. Epidemiyolojik olaylar, uzay ve zamanda oluşan sismik ve çevresel olaylar gibi doğal olayların bilimsel araştırma yöntemlerinin sık sık özellikleri olayların zamanının gözlemin çok önemli kılar. Bu yüzden, çalışmamızda nokta süreçleri yukarıda sözü edilen olaylara uyarlanmıştır.

İlk olarak, nokta süreçlerinin teorik yaklaşımaları üzerinde durulmuştur. Bu yaklaşımlar, hala sorular barındıran bilimsel alanların çözüme ulaştırılmasında araçlardır. Bu araçlar problere ya çözüm üretir ya da yeni problemler ortaya çıkarabilirler. Son on yılda, nokta süreçleri çalışmaları artış göstermiştir. Bu çalışmaların bir kısmı bu süreçler için teoriler geliştirirken, diğer kısmı da bunları çeşitli alanlarda uygulamaktadır.İlk iki bölümde, nokta süreçlerinin teorisine katkıda bulunan Poisson süreci, Bernoulli süreci, Cox süreci and Hawkes süreci gibi temel modeller ele alınmıştır.

Bölüm 3'de, Marko Chain Monte Carlo algoritmaları (MCMC), Metropolis-Hastings ve Gibbs Hibrid algoritmaları gibi temel algoritmalar simülasyon yapmak ve kullanılan yöntemlerin performansını ölçmek üzere tanıtılmıştır. Özellikle MCMC yöntemi dikkat çekicidir. Epidemiyoloji konusu olan problemimize uygun yöntem elde etmek için sıtma hastalığı simülasyonunda kullanılan başlıca yöntemlerdendir. Süreçle ilgili yaklaşımın olasılığı ise bir sonraki bölümde ele alınmıştır.

Bayesian parametric olmayan yaklaşım için, zaman-mekânsal nokta süreçleri olan Cox ve Dirichlet'nin birleşimi önerilir. Bu model Bayesian yaklaşımda istatistiksel araçların kullanımında bilgi sahibi olmamızı sağlar. Buna ilaveten, Dirichlet ve Cox karışık süreci parametrik olmayan yaklaşımın posteror dağılımını bulmak için yararlıdır ve kümeleme araştırmalarında geniş kullanıma sahiptir. Ancak, bizim çalışmada verilerin özelliklerinden dolayı, sıtma hastalığı dağılımında doğru şekilde kullanılamamaktadır. Bunun sebebi, verilerimizin bazı yöntemlerin uygulanması için uygun olmayacak şekilde global olmasıdır. Sıtmanın Dirichlet sürecinin yayılımına bağlı hesaplanan çevresel değişkenler çalışılmıştır. Bu yöntem, verininmekânsal pozisyonuna yoğunlaşır. Bu da mekânsal epidemiyoloji uygulamasında yararlıdır.

Bölüm 5, problemimize uyarlanan nokta süreçleriyle epidemiyolojik istatistiksel yaklaşım çalışılmıştır. CAR'da politik krizler döneminde sağlık problemlerinin yoğunluğu, sıtma hastalığının çalışma konusu yapılmasında motive edici faktor olmuştur. Sözü edilen coğrafyada birçok kişiyi olumsuz etkileyen ve hala problem olmaya devam eden bu olayı modellemek ve geleceği ile ilgili doğru tahminlere ulaşmak sağlık politikaları için belirleyici olacaktır. Dolayısıyla, sonuçlar otoritelere olayın sönümlenmesi ve nihayetinde tamamen ortadan kaldırılması için işaretler vermektedir. Çalışmamızda kullandığımız veriler, daha detaylı (hastalıkların yerel bölgelerdeki durumu, saha çalışması gerektiren durumlarda yerinde gözlem sonuçları vs.) ve sayı olarak daha fazla olması durumunda nokta süreçlerinin çok daha etkin sonuçlar vereceği aşikârdır. Bu çalışma, nokta süreçlerinin farklı problemlere nasıl uygulanması gerektiği anlamında genel görüş kazandırdığından gelecek çalışmalar için ufuk açıcı olmuştur. Örneğin, büyük şehirlerdeki zaman–mekânsal kriminoloji, belirli hastalıkların yayılımı gibi olayları karakterize eden ve tanımlayan yeni stokastik modeller geliştirmek mümkündür. Bu konuda, yazılımları etkin kullanabilmek onlar hakkkında sahip olunan bilgi ile sınırlı olduğundan, öncelikle bu bilginin genişletilmesinde fayda vardır.

KAYNAKLAR

[1] Statistiques sanitaires mondiales, publications de l'Organisation mondiale disponibles sur le site Web de l'OMS (www. who. int), ISBN 978 92 4 256458 7, publiées dans le bulletin d'information de 2013

[2] Document de stratégie de réduction de la pauvreté (DSRP 2008-2010), publié Septembre 2007 par Ministère de l'Economie, du Plan et de la Coopération internationale www.minplan-rca.org

[3] D. Daley and D. Vere-Jones, (2003). An Introduction to the Theory of Point Processes, volume1: elementary theory and methods. Springer, 2nd Ed., New York.

[4] M.N. M. Van Lieshout, (2000). Markov Point Processes and their Applications, Imperial College Press, London.

[5] Y. Ogata, (1988). "Statistical models for earthquake occurrences and residual analysis for pointprocesses", Journal of American Statistical Association, 83:9-27

[6] Y. Ogata, (1988). Space-time point-process models for earthquake occurrences, Annals of the Institute of Statistical Mathematics, 50:379-402

[7] J. Zhuang, Y. Ogata, and D. Vere-Jones, (2005). Diagnostic Analysis of Spacetime Branching Processes for Earthquakes, A. Baddeley, P. Gregori, J. Mateu, R. Stoica and D. Stoyan, Chap 15:275-290.

[8] Holden L., Sannan S., Bungun H., (2003). "A stochastic marked point process model for earthquakes", Natural Hazards and Earth System Sciences, 3:95-101

[9] J.vaillant, (1992). "échantillonnage et étude statistique de populations en milieu hétérogène", Revue de Statistique Appliquée, 40(1):15-26; J. Vaillant, (1992). "Testing hypotheses about occupancy rates for clustered populations", Biometrical, Journal, 34(2):735-744.

[10] J. Vaillant, (1991). "Negative binomal distributions of individuals and spatio-temporal cox processes", Scandinavian Journal of Statistics, 18:235-248

[11] J. Vaillant, G. Puggioni, L. Waller, and J.-H. Daugrois, (2011). "A spatiotemporal analysis of the spread of sugar cane yellow leaf virus", Journal of Time Series Analysis, 32:396-406

[12] W. Truccolo, U. Eden, J. Donoghue, and E. Brown, (2005). “A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects”, Journal of Neurophysiology, 93:1074

[13] V. Benes and B. z. Frcalova, (2008) “Modelling and simulation of a neurophysiological experiment by spatio-temporal point processes”, Image Analysis & Stereology, 27:47-52.

[14] Kogan, (2004). “Short-term properties of earthquake catalogs and models of earthquake source”, Bulletin of the Seismological Society of America, 94:1207-1228.

[15] Diggle P. J., (2005). “A partial likelihood for spatio-temporal point processes”, The Berkeley Electronic Press, 75.

[16] Kottas A., Duan J. A., Gelfand A. E., (2007). “Modeling disease incidence data with spatial and spatio-temporal dirichlet process mixtures”, Biometrical Journal, 49:1-14.

[17] A. F. Karr, (1991). Point Processes and Their Statistical Inference, Probability: pure and applied ; 2, New-York, 2nd edition ed.

[18] D. Daley and D. Vere-Jones, (2003). An Introduction to the Theory of Point Processes, Springer, 2nd Ed., New York.

[19] D. R. Cox and V. Isham, (1980). Point processes, Monographs on Applied Probability and Statistics, Chapman & Hall, London.

[20] M. Tanner, (1996). Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions, New-York.

[21] P. Brémaud, (1981) Point Processes and Queues, Martingales Dynamics, Springer Verlag.

[22] A. Brix and P. Diggle, (2001). “Spatio-temporal prediction for log-gaussian cox processes”, Journal of the Royal Statistical Society B, 63:823-841.

[23] A. Kottas and B. Sanso, (2007). Bayesian mixture modeling for spatial poisson process intensities, with applications to extreme value analysis, Statistical Planning and Inference 137:3151-3163.

[24] A. G. Hawkes, (1971). Spectra of some self-exciting and mutually exciting point processes, Biometrika, 58:83-90.

[25] L. Puri and P. Tuan, (1986). Maximum likelihood estimation for stationary point processes, Proceedings of the National Academy of Sciences, 83:541-5

[26] A. G. Hawkes, D. Oakes, (1974). “A cluster representation of a self-exciting process, Journal of Applied Probability, 11:493-503.

[27] A. Brix, (1999). Generalized gamma measures and shot-noise cox processes, Advanced in Applied Probability, 31:929-953.

[28] A. Brix and J. Chadoeuf, (2002). “Spatio-temporal modeling of weeds and shotnoise cox processes”, Biometrical Journal, 44:83-99.

[29] J. Møller, (2003). Shot noise cox processes, Advanced in Applied Probability, 35:614-640.

[30] T. Ferguson, (1973). A bayesian analysis of some nonparametric problems, The Annals of Statistics, 1:209-230.

[31] J. Møller and G. Torrisi, (2005). Generalised shot noise cox processes, Advanced in Applied Probability, 37:48-74

[32] J. A. Duan, M. Guindani, and A. E. Gelfand, (2007). Generalized spatial dirichlet process models, Biometrika, 94:809-825.

[33] C. Robert, (2005). Le choix bayesien, principes et pratique, Springer, Paris.

[34] P. Green and S. Richardson,(2002). "Hidden markov models and disease mapping", American Statistical Association, 97:1055-1070.

[35] Gabriel MW, Woods LW, Poppenga R, Sweitzer RA, Thompson C, Matthews SM, et al. (2012). "Anticoagulant Rodenticides on our Public and Community Lands: Spatial Distribution of Exposure and Possoning of a rare forest caarnivore". Plos ONE 7(7):e40163.dois:10.1371/journal.pone.0040163.

[36] Ripley, B. D. (1981) Spatial Statistics. Wiley, 252pp. ISBN 0-471-08367-4

[37] Philippe Robert, (2000). Réseaux et files d'attente: méthodes probabilistes, ISBN 978-3-540-67872-4.

[38] Y.V. Fyodorov, (2005).Introduction to the Random Matrix Theory: Gaussian Unitary Ensembleand Beyond, in Recent Perspectives in Random Matrix Theory and Number Theory, London Mathematical Society Lecture Note Series 322

[39] P. K. Andersen, O. Borgan, R. D. Gill, and N. Keiding, (1993). Statistical Models Based on Counting Processes, Springer series in statistics, Springer.

[40] Adrew. B. Lawson, David. G. T. Denison, (2002). spatial cluster Modeling, ISBN: 1-58488-266-2.

[41] D.J. Daley, Vere Jones, (1998). an introduction to the theory of point process, (1998), ISBN 978-1-4757-2001-3, springer New York.

[42] J. Møller, (2005). "Residual analysis for spatial point processes", J. R. Statist. Soc. B, 67, Part 5: 617–666.

[43] en.wikipedia.org/wiki/filtration-mathematics. 02.06.2015

[44] Larissa Audey VALEMY, (2012). Model hierarchique et processus ponctuel spatio-temporel – application en epidemiologie et en sismiologie, l'Université des Antilles et de la Guyane.

[45] Jacobsen, M, (2006). Point Process Theory And Applications. Marked Point and Piecewise Deterministic Processes, Birkhäuser Verlag GmbH, Boston.

[46] Frederic Paik Schoenberg, David R. Brillinger & Peter Guttorp, (2002). Point processes, spatial–temporal. John Wiley & Sons, Ltd, Chichester, Volume 3, 1573–1577.

[47] S. Rathbun, (1996). Asymptotic properties of the maximum likelihood estimator for spatio-temporal point processes, Journal of Statistical Planning and Inference, 51:55-74.

[48] J. L. Jensen and N. V. Petersen, (1999). "Asymptotic normality of the maximum likelihood estimator in state space models, Annals of Statistics, 27:514-535.

[49] L. Paninski, (2004). Maximum likelihood estimation of cascade point-process neural encoding models, Computation in Neural Systems, 15:243-262.

[50] J. Møller and C. Diaz-Avalos, (2010). "Structured spatio-temporal shot-noise cox point process models, with a view to modelling forest res", Scandinavian Journal of Statistics, 37:2-25.

[51] Andrew B. Lawson, (2006). Statistical Methods in Spatial Epidemiology, ISBN: 978-0-470-01484-4

[52] Janine B. Illian, Sigrunn H. Sørbye and Havard Rue, (2013). A toolbox for fitting spatial point process models using integrated nested LAPLACE approximation (INLA).

[53] P. J. DIGGLE, (2003). Statistical Analysis of Spatial Point Pattern, (second edition). London: EdwardArnold.

[54] GABRIEL, E. and P. J. DIGGLE, (2009). Second-order analysis of inhomogeneous spatiotemporal point process data. Statistica Neerlandica 63:43–51.

[55] P. J. DIGGLE, (2006). "Spatio-temporal point processes, partial likelihood, foot-and-mouth" Statistical Methods in Medical Research, 15:325–336.

[56] P. J. DIGGLE, GOMEZ-RUBIO, V. P.E. BROWN, A. G. CHETWYND and S. GOODING, (2007). Second-order analysis of inhomogeneous spatial point processes using casecontrol Data, Biometrics, 63:550–557

[57] D.R. Cox, and P. A. W. Lewis, (1972). Multivariate point processes. Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability 401—448.

[58] A. Baddeley, R. Turner, J. Møller, and M. Hazelton, (2005). "Residual analysis for spatial point processes", Journal of the Royal Statistical Society. Series B (Statistical Methodology), 67:617-666.

[59] F. P. Schoenberg, (2003). "Multidimensional residual analysis of point process models for earthquake occurrences", Journal of the American Statistical Association, 98:789-795.

[60] J. Zhuang, (2006). "Second-order residual analysis of spatio-temporal point processes and applications in model evaluation", Journal of the Royal Statistical Society, 64:635-653.

[61] Y. Ogata, K. Katsura, and M. Tanemura, (2003). "Modelling heterogeneous spacetime occurrences of earthquakes and its residual analysis", Journal of the Royal Statistical Society. Series C (Applied Statistics), 52:499-509.

[62] P. J. Diggle, (1983). Statistical Analysis of Spatial Point Patterns, Academic Press, London.

[63] P. J. Diggle and R. K. Milne,(1983). “Negative binomial quadrat counts and point Processes”, Scandinavian Journal of Statistics. Theory and Applications, 10.

[64] J. Vaillant and S. Derridj, (1989). Estimation of european corn borer egg masses density by sampling of runs, Researches on Population Ecology, 31.

[65] R. M. Fewster, J. L. Laake, and S. T. Buckland, (2005). Line transect sampling in small and large regions, Biometrics, 61:856-859

[66] G. Perry, B. Miller, and N. Enright, (2006). A comparison of methods for the statistical analysis of spatial point patterns in plant ecology, Plant Ecology, 187(1):59-82.

[67] G. Perry, B. Miller, and N. Enright, (2006). A comparison of methods for the statistical analysis of spatial point patterns in plant ecology, Plant Ecology, 187(2).59-82.

[68] C. Gouriéroux and A. Monfort, (1995). Séries temporelles et modèles dynamiques, Second Edition, Collection Économie et Statistiques Avancées. Série École Nationale de la Statistique et de l'Administration et du Centre d'Études des Programmes Économiques, Economica.

[69] Arnaud, M. et Emery, X. (2000). Estimation et interpolation spatiale. Hermes Science Publications, Paris, 20.

[70] L. Taylor, (1984). Assessing and interpreting the spatial distributions of insect populations, Annual Review of Entomology, 29:321-357.

[71] A. Rogers, (1993). Statistical Analysis of Spatial Dispersion the quadrat method, Pion Limited., London.

[72] Kingman, J. F. C. (1993). Poisson Processes. Oxford University Press. ISBN 0-19-853693-3.

[73] D. N. Shanbhag and Calyampudi Radhakrishna Rao (2001), Handbook of Statistics 19: Stochastic Processes: Theory and Methods. North-Holland, P.606, ISBN:0-444-500146

[74] Adrian, R., Wilhelm, S. & Gerten, D. (2006) Life-history traits of lake plankton species may govern their phenological response to climate warming. Global Change Biology, 12:652–661.

[75] Cox, D. R. (1972). The statistical analysis of dependicies in point processes. In: Stochastic Point Processes (ed. P. A. W. Lewis), Wiley, New York, 55–66.

[76] D. R. Cox, P. A. W. Lewis. (1966), the statistical analysis of series of events, Baldwin City, KS, U.S.A. ISBN 13: 9780412218002

[77] Cox, D. R. (1955). "Some Statistical Methods Connected with Series of Events". Journal of the Royal Statistical Society 17 (2): 129–164. doi:10.2307/2983950.

[78] Matern, B. (1971). Doubly stochastic Poisson processes in the plane. In: Statistical Ecology (eds. G. P. Patil, E. C. Pielou and W. E. Waters), volume 1,Pennsylvania State University Press, University Park, 195–213.

[79] Bartlett, M. S. (1964). The spectral analysis of two-dimensional point processes. Biometrika 51:299–311.

[80] Emmanuel Bacry. (2015) .Hawkes Processes in Finance, Market Microstructure and Liquidity, World Scientic Publishing Company, UMR 7641, 91128 Palaiseau, France. 06/30 Vol. 1, No. 1- 1550005

[81] A. G. Hawkes, (1971). Spectra of some self-exciting and mutually exciting point processes, Biometrika, 58 (1971)83-90.

[82] A.J Baddeley and M.M. Van Lieshout. (1992). ICM for object recognition. In Yadolah Dodge and Joe Whittaker edition, Coputational Statistiacs Volume 2:271-286 Phisica/springer Heidelberg New York.

[83] Kendall, W. S. And Møller, J. (2000). Perfect simulation using dominating processes on ordered spaces, with application to locally stable point processes. Advances in Applied Probability 32, 844–865

[84] van Lieshout, M. N. M. and R. S. Stoica (2003a). The Candy model revisited: properties and inference. Stat. Neerlandica 57, 1–30. van Lieshout, M. N. M. and R. S. Stoica (2003b). Perfect simulation for marked point processes. Research report PNA-0306, CWI

[85] van Lieshout, M. N. M. and R. S. Stoica (2006). Perfect simulation for marked point processes. Computational Statistics and Data Analysis 51, 679–698.

[86] J. Droesbeke, J. Fine, and G. Saporta, (2002). Méthodes Bayésiennes en Statistique, Technip.

[87] R. Neal, (2000). “Markov chain sampling methods for dirichlet process mixture models”, Journal of Computational and Graphical Statistics, 9:249-265

[88] J. Albert, (2009). Bayesian Computation with R, Springer, New York, 2nd ed.

[89] J.O. Lloyd-Smith, D. George, K.M. Pepin, V.E.Pitzer, J.R. Pulliam, A.P. Dobson, P.J. Hudson and B.T. Grenfell, (2013). Epidemic dynamics at the human – animal interf.

[90] Geyer, C. J. and Møller J. (1994). “Simulation and likelihood inference for spatial point processes”, Scandanavian Journal of Statistics, 21:359–373.

[91] P. Green, (1995). Reversible jump markov chain monte carlo computation and bayesian model determination, Biometrika, 82:71-732.

[92] J.S. Rosenthal, (2011). Optimal distributions and adaptive MCMC. In S. Book, A. Gelman, G.L. Jones and X. Meng editors Handbook of Marko Chain Monte Carlo 93-112 Taylor and Francis.

[93] C. Adrieu and Y.F. Atchadé, (2007). On the efficiency of adaptive MCMC algorithms. Electronic communications in probability 12 (33): 336-349.

[94] C. Adrieu and E. Moulines, (2006). On the ergodicity properties of some adaptives Marko Chain Monte Carlo algorithms. The annal of applied probability 16(3):1462-1505.

[95] Y.F. Atchadé and J.S. Rosenthal, (2005). On adaptive Marko Chain Monte Carlo algorithms Bernouilli. 11(3):815-828.

[96] G. Fort. , E. Moulines and P. Priouret. (2011). Convergence of adaptive and interacting Marko Chain Monte Carlo algorithms. The annals of Statistics 39 (6):3262-3289.

[97] P. Giordani and R. Kohn, (2010). Adaptive independent Metropolis Hastings by fast estimation of mixtures of normals. Journal of computational and graphical Statistics 19(2):243-259.

[98] H. Haario, M. Laine, A. Mira and E. Saksman DRAM. (2006). Efficient adaptive MCMC statistic and computating 16(4):339-354.

[99] H. Haario, E. Saksman and J. Tamminen, (2001). An adaptive Metropolis algorithms. Bernoulli 7(2). 223-242.

[100] G.O. Roberts and J.S. Rosenthal, (2007). “Coupling and ergodicity of adaptive Marko Chain Monte Carlo algorithms”. Journal of applied probability , 44(2):458-475.

[101] G.O. Roberts and J.S. Rosenthal, (2009). “Example of adaptives MCMC” . Journal of computational and graphical Statistics, 18(2):349-367.

[102] J.S. Rosenthal, (2004). adaptives MCMC Java Applet. On the web at: URL: http//probability.cal/jeff/java/adapt.htm.

[103] E. Turno, N.Bochkina, A.M. Hein and S.Richardson, (2007). BGX: a Bioconductor package for Bayesian integrated analysis offymetrix Genechips. BMC bioinformatics. 8(1):439-448.

[104] M. Vihola, (2012). Robust adaptive Metropolis algorithm with coerceptance rate. Statistics and computing 22(5): 997-1008.

[105] A. Gelman and D. B. Rubin, (1992). A single series from the Gibbs provides of security. In Bayesian Statistics 4 (J. M. Bernado, J. O. Berger, A. P. David and A. F. M. Smith, eds).625-632 Oxford Univ press.

[106] J.Geweke, (1992). Evaluating the accuracy of sampling based approaches to calculation of posterior moment. . In Bayesian Statistics 4 (J. M. Bernado, J. O. Berger, A. P. David and A. F. M. Smith, eds). 169-193 Oxford Univ press.

[107] A. E. Raftery and Lewis, (1992). How many iteration in Gibbs sampler? . In Bayesian Statistics 4 (J. M. Bernado, J. O. Berger, A. P. David and A. F. M. Smith, eds). 763-773.

[108] S. Geman and D. Geman, (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration og images. IEEE. Transactions on pattern analysis and machine intelligence. 6721-741.

[109] H. Geolf Givens and A. J. Hoeting, (2013). Weley series in computational statistics: 978-0-470-53331-4

[110] R. S. Ehlers and S. P. Brooks, (2008) “Adaptative proposal construction for reversible jumps mcmc”, Scandinavian Journal, 35:677-690.

[111] D. Blackwell and J. Macqueen, (1973). Ferguson distributions via polya urn schemes, Annals of Statistics, 353-355.

[112] J. Sethuraman, (1994). A constructive denition of dirichlet priors, Statistica Sinica, 4 :639-650,

[113] Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei, (2006). Hierarchical dirichlet processes, Journal of the American Statistical Association, 101:1566-1581.

[114] Taylor, B.M. and Diggle, P.J, (2013). Corrigendum: Spatio-temporal Prediction for logGaussian Cox Processes Journal of the Royal Statistical Society B 75, 601.

[115] DIGGLE, P.J. (2013). Statistical Analysis of Spatial and Spatio-Temporal Point Patterns (3rd edition) Boca Raton: CRC Press

[116] Baddeley A, Turner R (2005). “spatstat: An R Package for Analyzing Spatial Point Patterns.” Journal of Statistical Software, 12(6), 1–42. URL http://www.jstatsoft.org/v12/i06/.

[117] A. Baddeley, J. Møller and R. , Waagepetersen, (2003). Non- and semi-parametric estimation of interaction inhomogeneous point patterns statistica. Nerlandica 54:329-350.

[118] J. Møller and R. Waagepetersen, (2007). Modern statistics for spatial point processes, Scandinavian Journal of Statistics, 34:643-684.

[119] N. Cressie,(1993). Statistics for Spatial Data, John Wileyand Sons Inc. New York.

[120] L. Waller, (2005). Bayesian thinking in spatial statistics, Handbook of Statistics Handbook of Statistics, 25:589622.

[121] Møller, J. and Waagepetersen, R. P. (2003), Statistical Inference and Simulation for Spatial Point Processes, Chapman and Hall/CRC, Florida

[122] J. Vaillant, (1994). Tests of occupancy rates for clustered populations: Case where the cluster sizes are known, Biometrical Journal, 36: 33-42.

[123] L. Knopoff, (2000). The magnitude distribution of declustered earthquakes in southern california, Proceedings of the National Academy of Sciences of the United States of America, 97:11880-11884.

[124] P. Cowpertwaita, P. O'Connella, A. Metcalfeb, and J. Mawdsley, (1996). Stochastic point process modelling of rainfall. 1. single-site tting and validation, Journal of Hydrology, 175: 17-46.

[125] L. Waller, (2005). Bayesian thinking in spatial statistics, Handbook of Statistics Handbook of Statistics, 25:589-622.

[126] Rosychuk, R.J.; Huston, C. and Prasad, N.G.N. (2006). Spatial event cluster detection using a compound Poisson distribution. Biometrics, 62, 465-470.

[127] D. Stoyan and A. Penttinen, (2000). Recent applications of point process methods in forestry statistics, Statistical Science, 15:6178.

[128] J. Møller and R. Waagepetersen, (2007). Modern statistics for spatial point processes, Scandinavian Journal of Statistics, 34:643-684.

[129] Møller, J. Syversveen, A. R. and Waagepetersen, R. P. (1998). Log gaussian cox processes. Scand. J. Statist. 25, 451–482.

[130] Diggle, P.J., Guan, Y., Hart, A.C., Paize, F. and Stanton, M. (2010). "Estimating individual-level risk in spatial epidemiology using spatially aggregated information on the population at risk", Journal of the American Statistical Association, 105(492):1394–1402.

[131] Diggle, P.J., Kaimi, I. and Abellana, R. (2010). Partial-likelihood analysis of spatio-temporal point-process data. Biometrics, 66(2):347–354.

[132] G. Hellmund, (2009). Levy based Cox point processes doctora tezi, departement of Mathematic sciences, university of Aorhus.

[133] J. Sethuraman, (1994). A constructive denition of dirichlet priors, Statistica Sinica, 4:639650

[134] David M. Blei and Michael I. Jordan, (2006). Variational inference for Dirichlet process mixtures. Bayesian Analysis 1(1).

[135] F. Chimard, (2010). Mélanges de processus ponctuels spatio-temporels et approche bayésienne semi-paramétrique, PhD thesis, Université des Antilles et de la Guyane

[136] P. Muller, F. Quintana and G. Rosner, (2004). "A Method for Combining Inference Across Related Nonparametric Bayesian Models," Journal of the Royal Statistical Society, 66:735–749

[137] R.Neal, (1992). Bayesian Mixture Modeling, in Proceedings of the Workshop on Maximum Entropy and Bayesian Methods of Statistical Analysis, vol. 11:197–211.

[138] Y. Chung and D.B. Dunson, (20015). Nonparametric Bayes Conditional Distribution Modeling with Variable Selection. 655 Huntington Ave. SPH2, 4th Floor, BostonMA 02115.

[139] Rao, V.A. and Teh, Y.W. (2009). Spatial normalized Gamma processes. Advances in Neural Information Processing Systems 22 (NIPS.

[140] M. C. Jackson, L. Huang, J. Luo, M. Hachey and E. Feuer. 2009. "Comparison of tests for spatial heterogeneity on data with global clustering patterns and outliers". Internationa Journal of Health Geographics 8:55.

[141] Elliott P, Wakefield JC, Best NG, Briggs DJ. (2000). Spatial epidemiology: methods and applications. In: Elliott P, Wakefield JC, Best NG, Briggs DJ, (dir.).

Spatial epidemiology: methods and applications. Oxford: Oxford University Press; 3-14

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı	:İssa RAMADANE
Doğum Tarihi ve Yeri	: 25/02/1983 Birao-Orta Afrika Cumhuriyeti
Yabancı Dili	:Fransızca, Türkçe, İngilizce
E-posta	:ramadanis70@yahoo.fr

ÖĞRENİM DURUMU

Derece	**Alan**	**Okul/Üniversite**	**Mezuniyet Yılı**
Yüksek lisans	Matematik Müh.	YTÜ	2016
Lisans	Matematik	Bangui Üniversitesi	2011
Lise	Seri C	Lycée D'Etat des Rapides	2005

İŞ TECRÜBESİ

Yıl	**Firma/Kurum**	**Görevi**
2010	Lise	Matematik Öğretmeni
2005	Acabef	Danışman