

**HOMOJEN POISSON SÜRECİNDE
DEĞİŞME NOKTASINA İLİŞKİN
İSTATİSTİKSEL ÇIKARSAMA**

**STATISTICAL INFERENCE
REGARDING THE CHANGE POINT
IN HOMOGENEOUS POISSON PROCESS**

AYTEN YİĞİTER

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

İSTATİSTİK Anabilim Dalı için öngördüğü

DOKTORA TEZİ

olarak hazırlanmıştır.

2006

Fen Bilimleri Enstitüsü Müdürlüğü'ne,

Bu çalışma jürimiz tarafından **İSTATİSTİK ANABİLİM DALI 'nda DOKTORA TEZİ** olarak kabul edilmiştir.

Başkan :.....(İmza).....
(Prof. Dr. Fetih YILDIRIM)

Üye (Danışman) :.....(İmza).....
(Prof. Dr. Ceyhan İNAL)

Üye (Var ise Eş Danışman) :.....(İmza).....
(Prof. Dr. Öniz TOKTAMIŞ))

Üye :.....(İmza).....
(Doç Dr. Salih ÇELEBİOĞLU)

Üye :.....(İmza).....
(Doç Dr. Gül ERGÜN)

ONAY

Bu tez/...../..... tarihinde Enstitü Yönetim Kurulunca belirlenen yukarıdaki jüri üyeleri tarafından tarihinde kabul edilmiştir.

...../...../.....

Prof.Dr. Ahmet R. ÖZDURAL
FEN BİLİMLERİ ENSTİTÜSÜ MÜDÜRÜ

HOMOJEN POISSON SÜRECİNDE DEĞİŞME NOKTASINA İLİŞKİN İSTATİSTİKSEL ÇIKARSAMA

Ayten Yiğiter

ÖZ

Bu çalışmanın amacı, homojen Poisson sürecinde değişme noktasının ve sürecin parametrelerinin tahmin edicilerini elde etmek; değişimin olup olmadığını test etmek için bir test istatistiği önermektir.

Bu amaçla, İkinci Bölüm’de değişme noktası sorununa ilişkin genel bilgi verilmiş; yapılan çalışmalar açıklanmış ve Poisson süreçlerinde değişme noktası sorununu ele alan çalışmalar incelenmiştir. Üçüncü Bölüm’de, homojen Poisson sürecinde en çok olabilirlik yönteminden yararlanılarak değişme noktasının ve parametrelerin elde edilen tahmin edicileri, test istatistiği ve uygulamalar verilmiştir. Son olarak Dördüncü Bölüm’de, elde edilen sonuçlar tartışılmıştır.

Anahtar Kelimeler: Değişme noktası, Poisson süreçleri, en çok olabilirlik yöntemi, olabilirlik oranı.

Danışman: Prof. Dr. Ceyhan İNAL, Hacettepe Üniversitesi, İstatistik Bölümü, Olasılık Teorisi ve Olasılıksal Süreçler Anabilim Dalı

STATISTICAL INFERENCE REGARDING THE CHANGE POINT IN HOMOGENEOUS POISSON PROCESS

Ayten Yiğiter

ABSTRACT

The aim of this study is to obtain the estimators of the change point and related parameters of the homogeneous Poisson process, and to propose a test statistic in order to test a change in the process.

For this purpose, general information about the change point problem and previous studies regarding to the change point are explained and the previous studies which are corresponded to the change point problem in Poisson processes are investigated in Chapter II. The estimators of the change point and the parameters of the homogeneous Poisson process which are obtained by using the maximum likelihood method, the test statistic and applications are given in Chapter III. Finally, the results are discussed in Chapter IV.

Keywords: Change point, Poisson processes, maximum likelihood method, likelihood ratio.

Advisor: Prof. Dr. Ceyhan İNAL, Hacettepe University, Department of Statistics, Probability Theory and Stochastic Processes Section

TEŞEKKÜR

Kendisinden çok şey öğrendiğim danışmanım Sayın Prof. Dr. Ceyhan İnal'a, bu çalışmanın her aşamasında gösterdiği özen, özveri ve sabır için sonsuz teşekkür ederim.

Sayın Prof. Dr. Öniz Toktamış ve Sayın Doç. Dr. Gül Ergün'e katkıları, yardımları için teşekkür ederim.

Sayın Dr. Murat Büyükyazıcı, Sayın Dr. Nazan Danacıođlu ve tüm öteki araştırma görevlisi arkadaşlarıma verdikleri destek için teşekkür ederim.

Bölüm başkanımız ve bölüm çalışanlarına sağladıkları elverişli çalışma ortamı için teşekkür ederim.

En değerli zenginliğim olan AİLEM'e desteklerini esirgemedikleri için teşekkür ederim.

İÇİNDEKİLER

	<u>Sayfa</u>
ÖZ	i
ABSTRACT	ii
TEŞEKKÜR	iii
İÇİNDEKİLER.....	iv
ÇİZELGELER DİZİNİ.....	vi
ŞEKİLLER DİZİNİ.....	ix
1. GİRİŞ	1
2. GENEL BİLGİLER.....	3
2.1. Değişme Noktası Sorunu	3
2.2. Olasılık Dağılımlarında Değişme Noktasının Tahmini ve Hipotez Testleri.....	4
2.2.1. En Çok Olabilirlik Yöntemi ve Rasgele Yürüyüşler	4
2.2.1.1. Üstel dağılımlı raslantı değişkenleri dizisinde değişme noktasının tahmini ve tahmin edicinin asimtotik dağılımı.....	8
2.2.1.2. Poisson dağılımlı raslantı değişkenleri dizisinde değişme noktasının tahmini ve tahmin edicinin asimtotik dağılımı.....	10
2.2.2. Hipotez Testi.....	17
2.3. Homojen Poisson Sürecinde Değişme Noktası Sorunu	19
2.3.1. Poisson Sürecinde Değişme Noktasının ve Parametrelerin Tahminine İlişkin Çalışmalar.....	20
2.3.2. Poisson Sürecinde Değişme Noktasına İlişkin Hipotez Testi.....	24
2.4. İngiltere'deki Kömür Madeni Kazaları Veri Kümesi	26
3. HOMOJEN POISSON SÜRECİNDE DEĞİŞME NOKTASININ TAHMİNİ ve HİPOTEZ TESTİ.....	29
3.1. Nokta Tahmini.....	29
3.1.1. λ_1 ve λ_2 'nin Bilindiği Durumda τ 'nun Tahmini.....	30
3.1.2. λ_1 'in Bilindiği, λ_2 'nin Bilinmediği Durumda τ ve λ_2 'nin Tahmini.....	31
3.1.3. λ_1 ve λ_2 'nin Bilinmediği Durumda Tahminler.....	33
3.1.4. Uygulama	33
3.1.4.1. λ_1 ve λ_2 'nin bilindiği durumda τ 'nun tahmini için örnekler	34

	<u>Sayfa</u>
3.1.4.2. λ_1 'in bilindiđi, λ_2 'nin bilinmediđi durumda, τ ve λ_2 'nin tahminleri için örnekler	39
3.1.4.3. λ_1 ve λ_2 'nin bilinmediđi durumda τ ve λ_1, λ_2 'nin tahminleri için örnekler	43
3.1.4.4. Kömür madeni kazaları veri kümesine ilişkin tahminler	46
3.2. Poisson Sürecinde Deđişme Noktasına İlişkin Test İstatistiđi	48
3.2.1. Olabilirlik Oranına Dayalı Test İstatistiđi	48
3.2.2. Uygulama	51
3.2.3. Kömür Madeni Kazaları Veri Kümesine İlişkin Hipotezlerin Testi	57
3.2.4. Hava Trafiđi Veri Kümesine İlişkin Hipotez Testi	59
4. SONUÇ ve TARTIŞMA	62
KAYNAKLAR.....	65

ÇİZELGELER DİZİNİ

	<u>Sayfa</u>
Çizelge 2.1. Maden kazaları arasında geçen günler (yukarıdan aşağıya doğru) (Jarett, 1979)	26
Çizelge 2.2. İngiltere'deki kömür madeni kazaları arasında geçen günlerin sayısını içeren veri kümesi için değişme noktasının nokta tahminleri ve aralık tahminleri (West and Ogden, 1997).....	27
Çizelge 3.1. $\lambda_1 = 0.5$, $\lambda_2 = 1.5$, $v = 10$, $n = 30$ ve $T = 38.76919$ için tahminler	34
Çizelge 3.2. $\lambda_1 = 0.5$, $\lambda_2 = 1.5$, $v = 15$, $n = 30$ ve $T = 38.79692$ için tahminler	34
Çizelge 3.3. $\lambda_1 = 0.5$, $\lambda_2 = 1.5$, $v = 35$, $n = 100$ ve $T = 113.15480$ için tahminler	35
Çizelge 3.4. $\lambda_1 = 0.5$, $\lambda_2 = 1.5$, $v = 50$, $n=100$ ve $T = 114.94556$ için tahminler	35
Çizelge 3.5. $\lambda_1 = 1.5$, $\lambda_2 = 0.5$, $v = 15$, $n = 30$ ve $T = 43.47320$ için tahminler	35
Çizelge 3.6. $\lambda_1 = 1.5$, $\lambda_2 = 0.5$, $v = 20$, $n = 30$ ve $T = 39.31683$ için tahminler	35
Çizelge 3.7. $\lambda_1 = 1.5$, $\lambda_2 = 0.5$, $v = 50$, $n = 100$ ve $T = 153.80751$ için tahminler	36
Çizelge 3.8. $\lambda_1 = 1.5$, $\lambda_2 = 0.5$, $v = 65$, $n = 100$ ve $T = 119.28360$ için tahminler	36
Çizelge 3.9. $\lambda_1 = 1$, $\lambda_2 = 5$, $v = 10$, $n = 30$ ve $T = 15.70841$ için tahminler ...	36
Çizelge 3.10. $\lambda_1 = 1$, $\lambda_2 = 5$, $v = 15$, $n = 30$ ve $T = 14.39305$ için tahminler ...	36
Çizelge 3.11. $\lambda_1 = 1$, $\lambda_2 = 5$, $v = 35$, $n = 100$ ve $T = 48.52171$ için tahminler ..	37
Çizelge 3.12. $\lambda_1 = 1$, $\lambda_2 = 5$, $v = 50$, $n = 100$ ve $T = 56.57775$ için tahminler ..	37
Çizelge 3.13. $\lambda_1 = 5$, $\lambda_2 = 1$, $v = 15$, $n = 30$ ve $T = 21.78410$ için tahminler	37
Çizelge 3.14. $\lambda_1 = 5$, $\lambda_2 = 1$, $v = 20$, $n = 30$ ve $T = 10.80383$ için tahminler	37
Çizelge 3.15. $\lambda_1 = 5$, $\lambda_2 = 1$, $v = 50$, $n = 100$ ve $T = 57.87960$ için tahminler ..	38
Çizelge 3.16. $\lambda_1 = 5$, $\lambda_2 = 1$, $v = 65$, $n = 100$ ve $T = 45.55770$ için tahminler .	38
Çizelge 3.17. $\lambda_1 = 1.5$, $\lambda_2 = 0.5$, $v = 35$, $n = 100$ ve $T = 148.53118$ için tahminler	38

ÇİZELGELER DİZİNİ (devam ediyor)

Sayfa

Çizelge 3.18. $\lambda_1 = 0.5, \lambda_2 = 1.5, v = 20, n = 30$ ve $T = 40.31612$ için tahminler	39
Çizelge 3.19. $\lambda_1=0.5, v = 10, n = 30$ ve $T = 38.76919$ için tahminler	39
Çizelge 3.20. $\lambda_1 = 0.5, v = 15, n = 30$ ve $T = 38.79692$ için tahminler	39
Çizelge 3.21. $\lambda_1 = 0.5, v = 35, n = 100$ ve $T = 113.15480$ için tahminler	40
Çizelge 3.22. $\lambda_1 = 0.5, v = 50, n = 100$ ve $T = 114.94556$ için tahminler	40
Çizelge 3.23. $\lambda_1 = 1.5, v = 15, n = 30$ ve $T = 43.47320$ için tahminler	40
Çizelge 3.24. $\lambda_1 = 1.5, v = 20, n = 30$ ve $T = 39.31683$ için tahminler.....	40
Çizelge 3.25. $\lambda_1 = 1.5, v = 50, n = 100$ ve $T = 153.80751$ için tahminler	40
Çizelge 3.26. $\lambda_1 = 1.5, v = 65, n = 100$ ve $T = 119.28360$ için tahminler	41
Çizelge 3.27. $\lambda_1 = 1, v = 10, n = 30$ ve $T = 15.70841$ için tahminler	41
Çizelge 3.28. $\lambda_1 = 1, v = 15, n = 30$ ve $T = 14.39305$ için tahminler	41
Çizelge 3.29. $\lambda_1 = 1, v = 35, n = 100$ ve $T = 48.52171$ için tahminler	41
Çizelge 3.30. $\lambda_1 = 1, v = 50, n = 100$ ve $T = 56.57775$ için tahminler	41
Çizelge 3.31. $\lambda_1 = 5, v = 15, n = 30$ ve $T = 21.78410$ için tahminler	42
Çizelge 3.32. $\lambda_1 = 5, v = 20, n = 30$ ve $T = 10.80383$ için tahminler	42
Çizelge 3.33. $\lambda_1 = 5, v = 50, n = 100$ ve $T = 57.87960$ için tahminler	42
Çizelge 3.34. $\lambda_1 = 5, v = 65, n = 100$ ve $T = 45.55770$ için tahminler	42
Çizelge 3.35. $\lambda_1 = 1.5, v = 35, n = 100$ ve $T = 148.53118$ için tahminler	43
Çizelge 3.36. $\lambda_1 = 0.5, v = 20, n = 30$ ve $T = 40.31612$ için tahminler	43
Çizelge 3.37. $v = 10, n = 30$ ve $T = 38.76919$ için tahminler.....	43
Çizelge 3.38. $v = 15, n = 30$ ve $T = 38.79692$ için tahminler.....	43
Çizelge 3.39. $v = 35, n = 100$ ve $T = 113.15480$ için tahminler.....	44
Çizelge 3.40. $v = 50, n = 100$ ve $T = 114.94556$ için tahminler.....	44
Çizelge 3.41. $v = 15, n = 30$ ve $T = 43.47320$ için tahminler.....	44
Çizelge 3.42. $v = 20, n = 30$ ve $T = 39.31683$ için tahminler.....	44
Çizelge 3.43. $v = 50, n = 100$ ve $T = 153.80751$ için tahminler.....	44
Çizelge 3.44. $v = 65, n = 100$ ve $T = 119.28360$ için tahminler.....	45
Çizelge 3.45. $v = 10, n = 30$ ve $T = 15.70841$ için tahminler.....	45
Çizelge 3.46. $v = 15, n = 30$ ve $T = 14.39305$ için tahminler.....	45

ÇİZELGELER DİZİNİ (devam ediyor)

	<u>Sayfa</u>
Çizelge 3.47. $v = 35$, $n = 100$ ve $T = 48.52171$ için tahminler.....	45
Çizelge 3.48. $v = 50$, $n = 100$ ve $T = 56.57775$ için tahminler.....	45
Çizelge 3.49. $v = 15$, $n = 30$ ve $T = 21.78410$ için tahminler.....	46
Çizelge 3.50. $v = 20$, $n = 30$ ve $T = 10.80383$ için tahminler.....	46
Çizelge 3.51. $v = 50$, $n = 100$ ve $T = 57.87960$ için tahminler.....	46
Çizelge 3.52. $v = 65$, $n = 100$ ve $T = 45.55770$ için tahminler.....	46
Çizelge 3.53. $\lambda_1 = 0.0085$ (gün), $\lambda_2 = 0.0025$ (gün), $n = 190$ ve $T = 40549$ (gün) için tahmin.....	47
Çizelge 3.54. $\lambda_1 = 0.0085$ (gün), $n = 190$ ve $T = 40549$ (gün) için tahminler ..	47
Çizelge 3.55. $n = 190$ ve $T = 40549$ (gün) için tahminler ..	48
Çizelge 3.56. Eşitlik (3.20)'de verilen hipotezlerin testi	51
Çizelge 3.57. Eşitlik (3.21)'de verilen hipotezlerin testi	52
Çizelge 3.58. Eşitlik (3.22)'de verilen hipotezlerin testi	53
Çizelge 3.59. Eşitlik (3.23)'te verilen hipotezlerin testi	53
Çizelge 3.60. Eşitlik (3.24)'te verilen hipotezlerin testi	54
Çizelge 3.61. Eşitlik (3.25)'te verilen hipotezlerin testi	55
Çizelge 3.62. $\lambda_1=0.1$, $\lambda_2=0.2$, $v = 15$, $n = 30$ ve $T = 209.97127$ için tahminler..	55
Çizelge 3.63. Eşitlik (3.26)'da verilen hipotezlerin testi	56
Çizelge 3.64. $\lambda_1=5.5$, $\lambda_2=5$, $v = 50$, $n = 100$ ve $T = 20.11312$ için tahminler ...	56
Çizelge 3.65. Eşitlik (3.27)'de verilen hipotezlerin testi	57
Çizelge 3.66. Eşitlik (3.28)'de verilen hipotezlerin testi	58
Çizelge 3.67. Eşitlik (3.29)'da verilen hipotezlerin testi	58
Çizelge 3.68. Eşitlik (3.30)'da verilen hipotezlerin testi	59
Çizelge 3.69. Uçakların iniş zamanları (yukarıdan aşağıya doğru) (Hsu, 1979).....	60
Çizelge 3.70. Hava trafiği veri kümesi için λ_1 ve λ_2 bilinmediği durumda τ ve λ_1 , λ_2 için tahminler.....	61
Çizelge 3.71. Eşitlik (3.31)'de verilen hipotezlerin testi	61

ŞEKİLLER DİZİNİ

	<u>Sayfa</u>
Şekil 2.1 Değişme noktasına göre rasgele yürüyüşlerin gösterimi	7
Şekil 2.2 Poisson sürecinde birim zaman aralıklarında ortaya çıkan olaylar ve değişme noktası	23
Şekil 3.1 Maden kazaları arasında geçen günler	47
Şekil 3.2 New York bölgesinde 30 Nisan 1969'da uçakların iniş zamanları arasında geçen zamanların (saniye olarak) uçakların iniş sırasına göre saçılımı	60

1. GİRİŞ

İstatistikte bilinmeyen parametreye ilişkin tahmin, genellikle alınan örneklemin aynı dağılımlı bağımsız raslantı değişkenleri olduğu varsayılarak elde edilir. Bir örnekleme oluşturan raslantı değişkenleri dizisinde, dağılımın değiştiği durumla da karşılaşılabilir. İşte bu sorun, değişme noktası (change-point) sorunu olarak adlandırılır.

Olasılıksal (stokastik) süreçlerin uygulama alanlarında da, zaman içinde süreci etkileyen farklı etkenler nedeniyle, değişme noktası sorunu ile çok sık karşılaşılır. Olasılıksal süreçler içinde Poisson süreçleri uygulama alanları çok geniş olduğu için önemlidir. Poisson süreçlerinde değişme noktası sorunu Akman ve Raftery (1986), Raftery ve Akman (1986), Loader (1992), West ve Ogden (1997) tarafından ele alınmıştır. Bu çalışmalar incelendiğinde değişme noktasının tahmin edicisinin esnek varsayımlar altında elde edilebildiği, ayrıca tahmin edicilerden tahminlerin elde edilmesi uzun hesaplamaları gerektirdikleri; süreç boyunca değişme olup olmadığının testi için işlevsel yapıda bir test fonksiyonuna ulaşamadığı görülmüştür.

Bu çalışmada, homojen Poisson sürecinde değişme noktasının ve sürecin parametrelerinin varsayımlara dayanmayan, uygulama kolaylığı sağlayan tahmin edicilerinin; değişmenin olup olmadığını test etmek için test istatistiğinin elde edilmesi amaçlanmıştır.

Bilindiği gibi, homojen Poisson sürecinde zaman içinde gözlenen olayların sayısı Poisson dağılımına sahip raslantı değişkenleridir ve olaylar arasında geçen zamanların dağılımı üstel dağılımdır. Bu nedenle Poisson dağılımlı ya da üstel dağılımlı raslantı değişkenleri dizisi içinde dağılımın parametresinin bilinmeyen bir zaman noktasında değiştiği durumu inceleyen çalışmalar bu çalışma için önemlidir: Worsley (1986); Haccou, Meelis and Geer (1988); Jandhyala and Fotopoulos (1999, 2001); Boudjellaba, MacGibbon and Sawyer (2001); Fotopoulos and Jandhyala (2001).

Bu alıřmanın İkinci Bölümü'nde, deęiřme noktası sorununa iliřkin genel bilgi verilmiř, raslantı deęiřkenleri dizisindeki deęiřme noktasının tahmini için kullanılan yöntemler ve yapılan alıřmalar açıklanmıř; Poisson süreçlerinde deęiřme noktası sorununu ele alan alıřmalar incelenmiřtir. Üçüncü Bölüm'de, homojen Poisson sürecinde deęiřme noktasının ve ilgilenilen parametrelerin en çok olabilirlik yönteminden yararlanılarak elde edilen tahmin edicileri ve test istatistięi verilmiřtir. Elde edilen kuramsal sonuçların başarıları çeřitli uygulamalarla karşılařtırmalı olarak sunulmuř; ayrıca elde edilen kuramsal sonuçların literatürde kullanılan veri kümelerindeki uygulamaları verilmiřtir. Son olarak Dördüncü Bölüm'de, yapılan alıřma ve elde edilen sonuçlar tartiřılmıřtır.

2. GENEL BİLGİLER

2.1. Değişme Noktası Sorunu

Değişme noktası sorunu ile ilgili çalışmaların 1950'lere dayandığı bilinmektedir. Ancak son çeyrek asırda karmaşık sistemlerin hayatımızın her alanına girmesiyle, bu sorunla ilgili çalışmalar özellikle aşağıda belirtilen alanlarda daha da artmıştır:

- Hata(fault) bulma ve kontrol etme(monitoring),
- Endüstriyel süreçler,
- Karmaşık sistemler (uçaklar, gemiler, roketler, nükleer güç santralleri, vb.),
- Kalite kontrolü,
- Doğal afetler (depremler, tsunami, vb.),
- Biyotıp (biomedicine).

Yukarda belirtilen alanlardaki ortak sorun, ilgilenilen değişkenlere ilişkin bir ya da birden çok beklenmedik (abrupt) değişimin saptanmasıdır. Burada asıl güçlük, doğrudan gözlenemeyen ve öteki düzensizliklerle birlikte ölçülmüş öz değişimleri saptamaktır. Ancak var olan ölçümler, çevreye bağlı düzensizliklerin ve yapıdaki değişmelerin birleşik bilgisini içerir (Basseville and Nikiforov, 1993). Bu nedenle değişme noktası sorununu çözümüleme yöntemleri ve bu yöntemlerin özellikleri uygulama alanına göre farklılaşmaktadır.

Literatürde değişme noktası sorunu, kapsamlı olarak öncelikle ikiterimli (binom) ve normal dağılıma sahip rastlantı değişkenleri için incelenmiştir (Hinkley, 1970, 1971; Hinkley and Hinkley 1970; Sen and Srivastava, 1975; Hsu, 1979; Worsley, 1979, 1983; Chen and Gupta, 1999 vb.). Bu çalışmaları Poisson, üstel ve gamma dağılımlı raslantı değişkenlerinde değişme noktası sorununu inceleyen çalışmalar izlemiştir (Worsley, 1986; Siegmund, 1988; Jandhyala and Fotopoulos, 1999, 2001; Boudjellaba, MacGibbon and Sawyer, 2001 vb.). Chen ve Gupta (2001)'nin çalışmalarında, değişme noktası sorunu ile ilgili yaygın olarak kullanılan yöntemleri aşağıdaki gibi özetlemiştir:

- En Çok Olabilirlik Yöntemi (Sıklıkçı Yaklaşım),
- Bayesci Çözümleme,
- Birikimli Toplama (CUSUM:Cumulative Sum) Dayalı Yöntem,
- Bilgi (Information) Kuramına Dayalı Yöntem,
- Hiyerarşik Alt Hipotezlerle Olabilirlik Oran Testine (LRT with a Hierarchy of Sub Hypotheses) Dayalı Yöntem,
- Dalgacık Yaklaşımına (Wavelet Approximation) Dayalı Yöntem,
- Parametrik Olmayan Yöntemler.

Yukarıda sıralanan yöntemler içinde en çok olabilirlik yöntemi, en çok kullanılan yöntemdir.

2.2 Olasılık Dağılımlarında Değişme Noktasının Tahmini ve Hipotez Testleri

Bu çalışmada homojen Poisson sürecine ilişkin değişme noktası sorununun çözümlenmesinde en çok olabilirlik yönteminden yararlanıldığı için, Kesim 2.2.1'de bu yöntem açıklanacaktır.

2.2.1. En Çok Olabilirlik Yöntemi ve Rasgele Yürüyüşler

Bir X_1, X_2, \dots, X_n raslantı değişkenleri dizisi verilsin. Bu dizi içinde bilinmeyen bir noktada, dağılımın değiştiğini varsayalım:

$$\begin{aligned} F_1(x_i; \theta_1), & \quad i = 1, 2, \dots, v \\ F_2(x_i; \theta_2), & \quad i = v+1, v+2, \dots, n. \end{aligned} \quad (2.1)$$

Burada v , dağılımın değiştiği noktayı gösterir. θ_1 ve θ_2 parametreleri biliniyor ya da bilinmiyor olabilir.

Değişme noktası v parametresi için şu iki varsayım vardır:

- Parametre uzayı $1, 2, \dots, n-1$ değerlerinden oluşur,
- $F_1(\cdot)$ ve $F_2(\cdot)$ dağılım fonksiyonları v parametresine bağlı değildir.

Eşitlik (2.1)'de verilen dağılım fonksiyonları v parametresini içermediği için, v parametresi, dolaylı parametre ya da indis parametresi olarak adlandırılmaktadır.

Yukarıda verilen tanımlar altında v için, X_1, X_2, \dots, X_n raslantı değişkenleri dizisinin $\log(\ln)$ olabilirlik fonksiyonu,

$$\begin{aligned} \log L(v) &= \log \left(\prod_{i=1}^v f_1(x_i; \theta_1) \prod_{i=v+1}^n f_2(x_i; \theta_2) \right) \\ &= \sum_{i=1}^v \log f_1(x_i; \theta_1) + \sum_{i=v+1}^n \log f_2(x_i; \theta_2) \end{aligned} \quad (2.2)$$

biçimindedir. Burada $f_1(x_i; \theta_1)$ ve $f_2(x_i; \theta_2)$, sırasıyla $F_1(x_i; \theta_1)$ ve $F_2(x_i; \theta_2)$ dağılım fonksiyonlarına karşılık gelen olasılık yoğunluk fonksiyonlarıdır. v bir indis parametresi olduğu için, \log olabilirlik fonksiyonunun en büyük olduğu noktanın bulunması, yani v 'ye göre türevinin alınması olanaksızdır. Bu nedenle v parametresine ilişkin çıkarsama, \log olabilirlik fonksiyonu üzerinde dönüşümler yapılarak elde edilir. Bu dönüşümlerden biri, \log olabilirlik fonksiyonunun rasgele yürüyüş olarak ifade edilmesidir.

Hinkley (1970)'in yaptığı çalışmada, raslantı değişkenleri dizisi içinde tek değişme noktasına ilişkin çıkarsama için \log olabilirlik fonksiyonu rasgele yürüyüş biçiminde ifade edilmiş; normal dağılımlı raslantı değişkenleri için değişme noktası ele alınmış; değişme noktasının tahmin edicisi, asimtotik dağılımı ve hipotez testlerine ilişkin çıkarsamalar elde edilmiştir. Hinkley ve Hinkley (1970)'in yaptıkları çalışmada, ikiterimli raslantı değişkenleri için de değişme noktasının tahmin edicisi, asimtotik dağılımı ve hipotez testlerine ilişkin çıkarsamaların elde edilmesinde aynı yöntem kullanılmıştır. Şimdi bu kesimde bu yöntem açıklanacaktır.

X_1, X_2, \dots, X_n raslantı değişkenleri dizisine ilişkin Eşitlik (2.2)'de verilen \log

olabilirlik fonksiyonuna $\sum_{i=1}^v \log f_2(x_i; \theta_2)$ ifadesi eklenip çıkarılırsa,

$$\log L(v) = \sum_{i=1}^v [\log f_1(x_i; \theta_1) - \log f_2(x_i; \theta_2)] + \sum_{i=1}^n \log f_2(x_i; \theta_2) \quad (2.3)$$

elde edilir. Burada,

$$Y_i = \log \left[\frac{f_1(X_i; \theta_1)}{f_2(X_i; \theta_2)} \right], \quad i = 1, 2, \dots, n-1 \quad (2.4)$$

olmak üzere, $Y_i = y_i$ için, log olabilirlik fonksiyonu,

$$\log L(v) = \sum_{i=1}^v y_i + \sum_{i=1}^n \log f_2(x_i; \theta_2)$$

biçiminde yazılabilir. İlgilendiğimiz parametre v olduğu için, log olabilirlik fonksiyonunun v 'yü içeren ilk parçası önemlidir; ikinci parçası ihmal edilebilir. v 'nün en çok olabilirlik tahmin edicisi \hat{v} , $\sum_{i=1}^j Y_i$ ifadesini en büyük yapan j değeridir. Buna göre,

$$\hat{v} = \operatorname{argmax}_{1 \leq j \leq n-1} \sum_{i=1}^j Y_i$$

biçimindedir. X_i raslantı değişkenlerinin bağımsız olmaları nedeniyle Y_i 'ler de bağımsızdır. Burada, $\sum_{i=1}^j Y_i - \sum_{i=1}^v Y_i$, $j = 1, 2, \dots, n-1$, farkları alınır, değişme noktası v 'den önceki ve sonraki gözlemlere ilişkin log olabilirlik fonksiyonları iki bağımsız rasgele yürüyüş olarak tanımlanır:

$$W = (0, -\sum_{i=1}^j Y_{v-i+1}), \quad j = 1, 2, \dots, v-1,$$

$$W^* = (0, \sum_{i=0}^j Y_{v+i}), \quad j = 1, 2, \dots, n-v-1.$$

Burada,

$$W_j = -Y_{v-j+1}, \quad j = 1, 2, \dots, v-1,$$

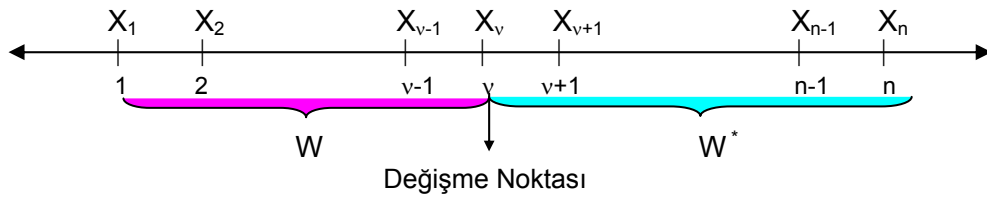
$$W_j^* = Y_{v+j}, \quad j = 1, 2, \dots, n-v-1$$

ile gösterilirse,

$$W = (0, W_1, W_1 + W_2, \dots, \sum_{j=1}^{v-1} W_j), \quad (2.5)$$

$$W^* = (0, W_1^*, W_1^* + W_2^*, \dots, \sum_{j=1}^{n-v-1} W_j^*) \quad (2.6)$$

olur ve W , sol yan rasgele yürüyüş; W^* , sağ yan rasgele yürüyüş olarak adlandırılır (Şekil 2.1).



Şekil 2.1 Değişme noktasına göre rasgele yürüyüşlerin gösterimi

Bu rasgele yürüyüşler aynı dağılımlı bağımsız artışlara sahiptirler. Şekil 2.1'de görüldüğü gibi W , değişme noktası v 'den daha küçük tam sayılar için log olabilirlik fonksiyonunun; W^* , değişme noktası v 'den daha büyük tam sayılar için log olabilirlik fonksiyonunun bilgisini içerir. Bu iki rasgele yürüyüşü en büyük yapan değerlerden büyük olanına karşılık gelen j ($j = 1, 2, \dots, n-1$) değeri, v 'nün \hat{v} tahmini olur. Eğer her iki rasgele yürüyüşte de en büyük değere karşılık gelen adım sayısı sıfır ise, $\hat{v} = v$ olur. Eşitlik (2.5)'te $\sum_{j=1}^k W_j$, $k = 1, 2, \dots, v-1$, toplamlarından W rasgele

yürüyüşünü en büyük yapan toplamı M ile, Eşitlik (2.6)'da $\sum_{j=1}^k W_j^*$,

$k = 1, 2, \dots, n-v-1$, toplamlarından W^* rasgele yürüyüşünü en büyük yapan toplamı M^* ile gösterirsek,

$$I = \inf(k: M = \sum_{j=1}^k W_j), \quad M = 0 \text{ ise, } I = 0,$$

$$I^* = \inf(k: M^* = \sum_{j=1}^k W_j^*), \quad M^* = 0 \text{ ise, } I^* = 0$$

biçiminde tanımlanan I ve I^* indisleri sırasıyla W ve W^* rasgele yürüyüşlerini en büyük yapan değerlere karşılık gelen adım sayılarıdır. Buna göre v 'nün \hat{v} tahmin edicisi $1, 2, \dots, n-1$ değerlerini alan bir raslantı değişkenidir. \hat{v} , $\max W$ ve $\max W^*$ 'in değerlerine göre v 'den küçük, büyük ya da v 'ye eşit değerler alabilir:

- $\max W > \max W^* \geq 0$ ise, \hat{v} , v 'nün değerinden k adım küçüktür;
- $\max W > \max W^* \geq 0$ ise, \hat{v} , v 'nün değerine eşittir;
- $\max W^* > \max W \geq 0$ ise, \hat{v} , v 'nün değerinden k adım büyüktür.

Buna göre \hat{v} 'nin, v 'den küçük, büyük ya da v 'ye eşit olması olasılıkları,

$$\begin{aligned}
 P(\hat{v} \text{ 'nin } v \text{ 'den küçük olması}) &= P(I = k, M > M^*, M > 0), & k = 1, 2, \dots \\
 P(\hat{v} \text{ 'nin } v \text{ 'ye eşit olması}) &= P(M = 0) P(M^* = 0), & (2.7) \\
 P(\hat{v} \text{ 'nin } v \text{ 'den büyük olması}) &= P(I^* = k, M^* > M, M^* > 0), & k = 1, 2, \dots
 \end{aligned}$$

olarak tanımlanır.

Bu yöntemin kullanımı, raslantı değişkenleri dizisinin dağılım fonksiyonuna göre farklı çözümlenmeleri gerektirir. Raslantı değişkenleri dizisi, kesikli bir dağılım fonksiyonuna sahip ise, $f(\cdot)$ ifadesi $p(\cdot)$ ile yer değiştirir. Kesim 2.2.1.1 ve Kesim 2.2.1.2'de bu yöntemin üstel dağılıma ve Poisson dağılımına uygulamaları verilecektir.

2.2.1.1. Üstel dağılımlı raslantı değişkenleri dizisinde değişme noktasının tahmini ve tahmin edicinin asimtotik dağılımı

Üstel dağılımlı raslantı değişkenleri dizisinde değişme noktasının tahmini, Jandhyala ve Fotopoulos (1999) tarafından ele alınmıştır. Hinkley (1970)'in çalışmasının bir uzantısı olan bu çalışmada, üstel ve normal dağılımda değişme noktalarının en çok olabilirlik tahmin edicilerine ilişkin olasılıkların elde edilmesi için bir algoritma geliştirilmiş; aynı araştırmacıların 2001'deki ikinci çalışmasında da üstel dağılımlı raslantı değişkenleri dizisinde değişme noktasının asimtotik dağılımı kuramsal olarak elde edilmiştir. Bu kesimde bu çalışmalar kısaca açıklanacaktır.

X_1, X_2, \dots, X_n , üstel dağılıma sahip raslantı değişkenleri dizisi olsun. Bu raslantı değişkenlerinin dağılım fonksiyonu aşağıdaki gibidir:

$$F_1(x_i; \lambda_1) = 1 - e^{-\lambda_1 x_i}, \quad i = 1, 2, \dots, v$$

$$F_2(x_i; \lambda_2) = 1 - e^{-\lambda_2 x_i}, \quad i = v+1, v+2, \dots, n.$$

Burada v , değişme noktası olmak üzere, bu raslantı değişkenleri dizisinin log olabilirlik fonksiyonu,

$$\begin{aligned} \log L(v) &= \log \left(\prod_{i=1}^v \lambda_1 e^{-\lambda_1 x_i} \prod_{i=v+1}^n \lambda_2 e^{-\lambda_2 x_i} \right) \\ &= \sum_{i=1}^v \log(\lambda_1 e^{-\lambda_1 x_i}) + \sum_{i=v+1}^n \log(\lambda_2 e^{-\lambda_2 x_i}) \end{aligned} \quad (2.8)$$

biçimindedir. Eşitlik (2.8)'de $\sum_{i=1}^v \log(\lambda_2 e^{-\lambda_2 x_i})$ ifadesi eklenip çıkarılırsa,

$$\log L(v) = \sum_{i=1}^v [\log(\lambda_1 / \lambda_2) - (\lambda_1 - \lambda_2) x_i] + \sum_{i=1}^n (\log \lambda_2 - \lambda_2 x_i) \quad (2.9)$$

elde edilir (Jandhyala and Fotopoulos, 1999). Bu ifadenin sağ yanında yer alan ikinci toplam, v değişme noktasını içermediği için ihmal edilebilir. $Y_i = \log(\lambda_1 / \lambda_2) - (\lambda_1 - \lambda_2) X_i$ olmak üzere, v 'nün en çok olabilirlik tahmin edicisi \hat{v} , $\sum_{i=1}^j Y_i$ ifadesini en büyük yapan j değeridir. Buna göre,

$$\hat{v} = \operatorname{argmax}_{1 \leq j \leq n-1} \sum_{i=1}^j Y_i$$

biçimindedir. Burada, $\sum_{i=1}^j Y_i - \sum_{i=1}^{j-1} Y_i$, $j = 1, 2, \dots, n-1$, farkları alınır, Eşitlik (2.5) ve

Eşitlik (2.6)'daki iki bağımsız rasgele yürüyüşte,

$$W_j = -Y_{v-j+1} = (\lambda_1 - \lambda_2) X_{v-j+1} - \log(\lambda_1 / \lambda_2), \quad j = 1, 2, \dots, v-1$$

$$W_j^* = Y_{v+j} = \log(\lambda_1 / \lambda_2) - (\lambda_1 - \lambda_2) X_{v+j}, \quad j = 1, 2, \dots, n-v-1$$

olur. Eşitlik (2.7)'deki olasılıkları hesaplamak rasgele yürüyüşleri oluşturan raslantı değişkenlerinin sürekli olmaları nedeniyle oldukça güçtür. Bu olasılıklar $\min\{v, n-v\} \rightarrow \infty$ için elde edilmiştir (Jandhyala and Fotopoulos, 1999, 2001).

2.2.1.2. Poisson dağılımlı raslantı değişkenleri dizisinde değişme noktasının tahmini ve tahmin edicinin asimtotik dağılımı

Boudjellaba, MacGibbon ve Sawyer (2001), Hinkley (1970)'in çalışmasını Poisson dağılımına uyarlamışlar; değişme noktasının en çok olabilirlik tahmin edicisini, tahmin edicinin asimtotik dağılımını elde etmişlerdir. Bu kesimde kısaca bu çalışma açıklanacaktır.

Poisson dağılımına sahip kitleden rasgele bir örneklem X_1, X_2, \dots, X_n olsun. Bu raslantı değişkenleri dizisinde, dağılımın parametresi olan λ_1 'in bilinmeyen bir noktada değiştiği ve λ_2 değerini aldığı varsayalım. Bu durumda,

$$P(X_i = x) = e^{-\lambda_1} \frac{\lambda_1^x}{x!}, \quad i = 1, 2, \dots, v$$

$$= e^{-\lambda_2} \frac{\lambda_2^x}{x!}, \quad i = v+1, v+2, \dots, n.$$

olarak tanımlanır. Burada λ_1 ve λ_2 bilinen, v bilinmeyen parametrelerdir. v değişme noktası için log olabilirlik fonksiyonu,

$$\log L(v) = \log \left(\prod_{i=1}^v e^{-\lambda_1} \frac{\lambda_1^{x_i}}{x_i!} \prod_{i=v+1}^n e^{-\lambda_2} \frac{\lambda_2^{x_i}}{x_i!} \right)$$

$$= \sum_{i=1}^v \log \left(e^{-\lambda_1} \frac{\lambda_1^{x_i}}{x_i!} \right) + \sum_{i=v+1}^n \log \left(e^{-\lambda_2} \frac{\lambda_2^{x_i}}{x_i!} \right)$$

biçimindedir. Bu ifade, $\sum_{i=1}^v \log \left(e^{-\lambda_1} \frac{\lambda_1^{x_i}}{x_i!} \right)$ ifadesi eklenip çıkarıldıktan sonra yeniden düzenlenirse,

$$\begin{aligned}
\log L(v) &= \sum_{i=1}^v x_i \log(\lambda_1 / \lambda_2) - \sum_{i=1}^v (\lambda_1 - \lambda_2) + \left(\sum_{i=1}^n \log \left(e^{-\lambda_2} \frac{\lambda_2^{x_i}}{x_i!} \right) \right) \\
&= \log(\lambda_1 / \lambda_2) \sum_{i=1}^v (x_i - \delta) + \left(\sum_{i=1}^n \log \left(e^{-\lambda_2} \frac{\lambda_2^{x_i}}{x_i!} \right) \right)
\end{aligned} \tag{2.10}$$

elde edilir (Boudjellaba, MacGibbon and Sawyer, 2001). Burada $\delta = \frac{\lambda_1 - \lambda_2}{\log(\lambda_1 / \lambda_2)}$

biçimindedir. $Y_i = X_i - \delta$ olmak üzere, v 'nün en çok olabilirlik tahmin edicisi \hat{v} , $\sum_{i=1}^j Y_i$ ifadesini en büyük yapan j değeridir. Bu durumda,

$$\hat{v} = \operatorname{argmax}_{1 \leq j \leq n-1} \sum_{i=1}^j Y_i \tag{2.11}$$

biçimindedir. Burada, $\sum_{i=1}^j Y_i - \sum_{i=1}^{j-1} Y_i$, $j = 1, 2, \dots, n-1$, farkları alınır, Eşitlik (2.5) ve

Eşitlik (2.6)'daki iki bağımsız rasgele yürüyüşte,

$$W_j = -Y_{v-j+1} = -X_{v-j+1} + \delta, \quad j = 1, 2, \dots, v-1$$

$$W_j^* = Y_{v+j} = X_{v+j} - \delta, \quad j = 1, 2, \dots, n-v-1$$

tanımlanır. Burada δ , kesirli bir sayıdır; bu nedenle rasgele yürüyüşlerin en büyük değerleri birden çok olabilir. Bunu gidermek için iki yol vardır. Birinci öneride δ 'nın kesirli olan bu değeri, bu değerden büyük olan tamsayıların en küçüğü ile değiştirilir; ikincisinde ise, δ 'nın kesirli olan bu değeri, bu değerden küçük olan tamsayıların en büyüğü ile değiştirilir (Boudjellaba, MacGibbon ve Sawyer, 2001).

Eşitlik (2.11)'de verilen v 'nün tahmin edicisi \hat{v} 'nin dağılımı analitik yapıda olmadığı için elde edilememektedir. Bu çalışmada olabilirlik fonksiyonun v 'ye göre iki bağımsız rasgele yürüyüş olarak tanımlanmış; \hat{v} 'nin v 'den küçük, \hat{v} 'nin v 'ye eşit ve \hat{v} 'nin v 'den büyük olması olasılıkları Teorem 2.1'de verilmiştir.

Teorem 2.1: Eşitlik (2.7)'deki olasılıklar, $v \rightarrow \infty$, $n-v \rightarrow \infty$ varsayımı altında aşağıdaki biçimde elde edilir (Boudjellaba, MacGibbon and Sawyer, 2001):

$$P(\hat{v} \text{ 'nin } v \text{ 'den küçük olması}) = \sum_{\ell=0}^{[k\delta]} p_{\ell,k} \left\{ p_{0,0}^* + \sum_{j=1}^{\infty} \sum_{i=\left[\frac{j+\ell-\delta k}{\delta}\right]+1}^{\left[\frac{j}{\delta}\right]} p_{j,i}^* \right\}, \quad k > 0,$$

$$P(\hat{v} \text{ 'nin } v \text{ 'ye eşit olması}) = p_{0,0} p_{0,0}^*,$$

$$P(\hat{v} \text{ 'nin } v \text{ 'den büyük olması}) = \sum_{\ell=[k\delta]+1}^{\infty} p_{\ell,k}^* \left\{ p_{0,0} + \sum_{j=0}^{\infty} \sum_{i=\left[\frac{j}{\delta}\right]+1}^{\left[\frac{j+\ell-\delta k}{\delta}\right]} p_{j,i} \right\}, \quad k > 0$$

Burada,

$$p_{0,0} = \exp\left(-\sum_{k=1}^{\infty} \left(\frac{1}{k}\right) \sum_{i=0}^{[k\delta]} e^{-k\lambda_1} \frac{k^i \lambda_1^i}{i!}\right),$$

$$p_{0,0}^* = \exp\left(-\sum_{k=1}^{\infty} \left(\frac{1}{k}\right) \sum_{j=[k\delta]+1}^{\infty} e^{-k\lambda_2} \frac{k^j \lambda_2^j}{j!}\right)$$

ve $p_{\ell,k}$ ile $q_{\ell,k}$,

$$\begin{aligned} p_{\ell,k} &= q_{\ell,k}, & -\ell + \delta k &\geq 0 \\ &= 0, & -\ell + \delta k &< 0, \end{aligned}$$

$$q_{\ell,k} = \sum_{\ell-\delta(k-1) \leq j \leq \ell} e^{-\lambda_1} \frac{\lambda_1^j}{j!} q_{\ell-j,k-1}, \quad \ell \geq 1, k \geq 2$$

$$= e^{-\lambda_1} \frac{\lambda_1^\ell}{\ell!} p_{0,0}, \quad \ell \geq 0, k = 1$$

$$= e^{-k\lambda_1} p_{0,0}, \quad \ell = 0, k \geq 1$$

$$= 0, \quad \text{öteki değerler}$$

biçimindedir. Ayrıca $p_{\ell,k}^*$ ve $q_{\ell,k}^*$,

$$p_{\ell,k}^* = q_{\ell,k}^*, \quad \ell - \delta k \geq 0$$

$$= 0, \quad \ell - \delta k < 0,$$

$$\begin{aligned}
q_{\ell,k}^* &= \sum_{0 \leq j \leq \ell - \delta(k-1)} e^{-\lambda_2} \frac{\lambda_2^j}{j!} q_{\ell-j,k-1}^*, & \ell \geq 1, k \geq 2 \\
&= e^{-\lambda_2} \frac{\lambda_2^\ell}{\ell!} p_{0,0}^*, & \ell \geq 1, k = 1 \\
&= 0, & \text{öteki değerler}
\end{aligned}$$

olarak tanımlanır. Burada indislerde yer alan $[\cdot]$ ifadesi tam değer fonksiyonunu göstermektedir.

Tanıt: Eşitlik (2.7)'deki olasılıklar, sağ ve sol rasgele yürüyüşleri en büyük yapan değerlerin bulunmasına dayanmaktadır. Buna göre,

$$\begin{aligned}
\psi &= \max_{j \geq 1} \{W_1, W_1 + W_2, W_1 + W_2 + W_3, \dots\}, \\
\psi^* &= \max_{j \geq 1} \{W_1^*, W_1^* + W_2^*, W_1^* + W_2^* + W_3^*, \dots\}
\end{aligned}$$

ve $M = \max(0, \psi)$ ile $M^* = \max(0, \psi^*)$ olmak üzere, $p_{\ell,k}$ ile $q_{\ell,k}$ ifadeleri,

$$\begin{aligned}
p_{\ell,k} &= P(M = -\ell + \delta k), \\
q_{\ell,k} &= P(\psi = -\ell + \delta k)
\end{aligned}$$

olasılıklarını gösterebiliriz. Burada,

$$\psi = -\ell + \delta k = \bigcup_{j \in \Psi_{\ell,k}} (\{W_1 = -j + \delta\} \cap \{\tilde{\psi} = -(\ell - j) + \delta(k-1)\})$$

ve

$$\tilde{\psi} = \max_{j \geq 1} \{W_2, W_2 + W_3, W_2 + W_3 + W_4, \dots\}$$

olarak tanımlanır. $\ell \geq 1, k \geq 2$ için $q_{\ell,k}$,

$$\begin{aligned}
q_{\ell,k} &= P(\psi = -\ell + \delta k) = P\left(\psi = \max_{j \geq 1} \left\{ \sum_{i=1}^j W_i \right\} = -\ell + \delta k\right) \\
&= P\left(\bigcup_{j \in \Psi_{\ell,k}} (\{W_1 = -j + \delta\} \cap \{\tilde{\psi} = -(\ell - j) + \delta(k-1)\})\right)
\end{aligned}$$

$$\begin{aligned}
&= \sum_{\ell-\delta(k-1) \leq j \leq \ell} P(W_1 = -j + \delta) q_{\ell-j, k-1} \\
&= \sum_{\ell-\delta(k-1) \leq j \leq \ell} P(-X_v + \delta = -j + \delta) q_{\ell-j, k-1} \\
&= \sum_{\ell-\delta(k-1) \leq j \leq \ell} P(X_v = j) q_{\ell-j, k-1} \\
&= \sum_{\ell-\delta(k-1) \leq j \leq \ell} e^{-\lambda_1} \frac{\lambda_1^j}{j!} q_{\ell-j, k-1}
\end{aligned}$$

elde edilir. $\tilde{M} = \max(0, \tilde{\psi})$ olmak üzere, $\ell \geq 0, k = 1$ için,

$$\begin{aligned}
q_{\ell, 1} &= P(\psi = -\ell + \delta) = P\left(\psi = \max_{j \geq 1} \left\{ \sum_{i=1}^j W_i \right\} = -\ell + \delta\right) \\
&= P(W_1 = -\ell + \delta) P(\tilde{M} = 0)
\end{aligned}$$

olur. Buna göre,

$$\begin{aligned}
q_{\ell, 1} &= P(-X_v + \delta = -\ell + \delta) P(\tilde{M} = 0) \\
&= P(X_v = \ell) p_{0,0} \\
&= e^{-\lambda_1} \frac{\lambda_1^\ell}{\ell!} p_{0,0}
\end{aligned}$$

bulunur. $\ell = 0, k \geq 1$ için,

$$\begin{aligned}
q_{0, k} &= P(\psi = \delta k) = P\left(\psi = \max_{j \geq 1} \left\{ \sum_{i=1}^j W_i \right\} = \delta k\right) \\
&= P(W_1 + W_2 + \dots + W_k = \delta k) P(\tilde{M} = 0) \\
&= P(-X_v - X_{v-1} - \dots - X_{v-k+1} + \delta k = \delta k) p_{0,0} \\
&= P(X_v + X_{v-1} + \dots + X_{v-k+1} = 0) p_{0,0} \\
&= e^{-k\lambda_1} p_{0,0}
\end{aligned}$$

elde edilir. $p_{\ell, k}^*$ ile $q_{\ell, k}^*$,

$$p_{\ell,k}^* = P(M^* = \ell - \delta k)$$

$$q_{\ell,k}^* = P(\psi^* = \ell - \delta k)$$

olasılıkları benzer biçimde bulunabilir. $\ell \geq 1, k \geq 2$ için,

$$\begin{aligned} q_{\ell,k}^* &= P(\psi^* = \ell - \delta k) = P\left(\psi^* = \max_{j \geq 1} \left\{ \sum_{i=1}^j W_i^* \right\} = \ell - \delta k\right) \\ &= P\left(\bigcup_{j \in \Psi_{\ell,k}^*} \left(\{W_1^* = j - \delta\} \cap \{\tilde{\psi}^* = (\ell - j) - \delta(k-1)\} \right)\right) \\ &= \sum_{0 \leq j \leq \ell - \delta(k-1)} P(W_1^* = j - \delta) q_{\ell-j,k-1}^* \\ &= \sum_{0 \leq j \leq \ell - \delta(k-1)} P(X_{v+1} - \delta = j - \delta) q_{\ell-j,k-1}^* \\ &= \sum_{0 \leq j \leq \ell - \delta(k-1)} P(X_{v+1} = j) q_{\ell-j,k-1}^* \\ &= \sum_{0 \leq j \leq \ell - \delta(k-1)} e^{-\lambda_2} \frac{\lambda_2^j}{j!} q_{\ell-j,k-1}^* \end{aligned}$$

elde edilir. Burada, $\tilde{\psi}^* = \max_{j \geq 1} \{W_2^*, W_2^* + W_3^*, W_2^* + W_3^* + W_4^*, \dots\}$ olarak tanımlanır.

$\tilde{M}^* = \max(0, \tilde{\psi}^*)$ olmak üzere, $\ell \geq 1, k = 1$ için,

$$\begin{aligned} q_{\ell,1}^* &= P(\psi^* = \ell - \delta) = P\left(\psi^* = \max_{j \geq 1} \left\{ \sum_{i=1}^j W_i^* \right\} = \ell - \delta\right) \\ &= P(W_1^* = \ell - \delta) P(\tilde{M}^* = 0) \\ &= P(W_1^* = \ell - \delta) p_{0,0}^* \\ &= P(X_{v+1} - \delta = \ell - \delta) p_{0,0}^* \\ &= P(X_{v+1} = \ell) p_{0,0}^* \\ &= e^{-\lambda_2} \frac{\lambda_2^\ell}{\ell!} p_{0,0}^* \end{aligned}$$

bulunur.

Bu ifadelerde yer alan $p_{0,0}$, olasılığı,

$$p_{0,0} = \exp \left\{ - \sum_{k=1}^{\infty} \left(\frac{1}{k} \right) P \left(\sum_{j=1}^k W_j > 0 \right) \right\}$$

biçimindedir (Feller, 1966). Burada $P \left(\sum_{j=1}^k W_j > 0 \right)$ olasılığı,

$$\begin{aligned} P \left(\sum_{j=1}^k W_j > 0 \right) &= P \left(\sum_{j=1}^k (-X_j + \delta) > 0 \right) \\ &= P \left(\sum_{j=1}^k X_j < \delta k \right) \\ &= \sum_{j=1}^{[\delta k]} e^{-k\lambda_1} \frac{(k\lambda_1)^j}{j!} \end{aligned}$$

olarak bulunur. $p^*_{0,0}$ olasılığı da benzer biçimde aşağıda verildiği gibidir (Feller, 1966):

$$p^*_{0,0} = \exp \left\{ - \sum_{k=1}^{\infty} \left(\frac{1}{k} \right) P \left(\sum_{j=1}^k W_j^* > 0 \right) \right\}.$$

Burada $P \left(\sum_{j=1}^k W_j^* > 0 \right)$ olasılığı aşağıdaki gibi bulunur:

$$\begin{aligned} P \left(\sum_{j=1}^k W_j^* > 0 \right) &= P \left(\sum_{j=1}^k (X_j - \delta) > 0 \right) \\ &= P \left(\sum_{j=1}^k X_j > \delta k \right) \\ &= \sum_{j=[\delta k]+1}^{\infty} e^{-k\lambda_1} \frac{(k\lambda_1)^j}{j!}. \end{aligned}$$

\hat{v} 'nin v 'den küçük olması, $M = -l + \delta k$, $M^* = j - \delta i$ ve $-l + \delta k > j - \delta i$ ($-l + \delta k > 0$ ve $j - \delta i \geq 0$) olduğunu vurgular. Buna göre, \hat{v} 'nin v 'den küçük bir değer olması olasılığı,

$$P(\hat{v}'\text{'nin } v\text{'den küçük olması}) = \sum_{\ell=0}^{[k\delta]} p_{\ell,k} \left\{ p_{0,0}^* + \sum_{j=1}^{\infty} \sum_{i=\left[\frac{j+\ell-\delta k}{\delta}\right]+1}^{\left[\frac{j}{\delta}\right]} p_{j,i}^* \right\}, \quad k > 0$$

olarak elde edilir. \hat{v}' 'nin v 'den büyük olması ise, $M^* = \ell - \delta k$, $M = -j + \delta i$ ve $\ell - \delta k > -j + \delta i$ ($\ell - \delta k > 0$ ve $-j + \delta i \geq 0$) olduğunu vurgular. Buna göre, \hat{v}' 'nin v 'den büyük olması olasılığı,

$$P(\hat{v}'\text{'nin } v\text{'den büyük olması}) = \sum_{\ell=[k\delta]+1}^{\infty} p_{\ell,k}^* \left\{ p_{0,0} + \sum_{j=0}^{\infty} \sum_{i=\left[\frac{j}{\delta}\right]+1}^{\left[\frac{j+\ell-\delta k}{\delta}\right]} p_{j,i} \right\}, \quad k > 0$$

biçiminde elde edilir. \hat{v}' 'nin v 'ye eşit olması ise, $M = 0$ ve $M^* = 0$ demektir. \hat{v}' 'nin v 'ye eşit olması olasılığı, $P(M = 0) = p_{0,0}$ ve $P(M^* = 0) = p_{0,0}^*$ olasılıklarının çarpımına eşit,

$$P(\hat{v}'\text{'nin } v\text{'ye eşit olması}) = p_{0,0} p_{0,0}^*$$

olur (Boudjellaba, MacGibbon and Sawyer, 2001).

2.2.2. Hipotez Testi

Bir X_1, X_2, \dots, X_n raslantı değişkenleri dizisi verilsin. Bu raslantı değişkenlerinin dağılım fonksiyonları Eşitlik (2.1)'deki gibi tanımlansın. Bu dizi içinde bilinmeyen bir noktada, dağılımın parametresinde değişme olup olmadığına ilişkin hipotezler,

H_0 : θ parametresi X_1, X_2, \dots, X_n dizisi için aynıdır.

H_1 : θ parametresi X_v 'den ($1 \leq v < n$) sonra değişir.

biçimindedir. Burada H_0 yokluk hipotezi, “dağılımın parametresinde değişme yoktur”, H_1 seçenek hipotezi, “dağılımın parametresinde değişme vardır” olarak ifade edilir. Bu hipotezler değişme noktası bilindiğinde, bağımsız iki örneklemin aynı kitleden gelip gelmediğine ilişkin hipotezler olarak düşünülebilir. Ancak, v bilinmediği için bu hipotezlerin testi bilinen testlerden daha farklıdır.

μ beklenen değeri ve σ^2 varyansı olan normal dağılımlı X_1, X_2, \dots, X_n raslantı değişkenleri dizisinde,

H_0 : μ parametresi X_1, X_2, \dots, X_n dizisi için aynıdır.

H_1 : μ parametresi X_v 'den ($1 \leq v < n$) sonra değişir.

hipotezlerini test etmek için Hinkley (1971), kümülatif toplama dayalı test istatistiğini; Sen ve Srivastava (1975), Worsley (1979), James, James ve Siegmund (1987), Gombay ve Horvath (1996) olabilirlik oranına dayalı test istatistiğini elde etmişlerdir. Chen ve Gupta (1999), normal dağılıma ilişkin çalışmalarında, X_1, X_2, \dots, X_n raslantı değişkenleri dizisi için,

H_0 : μ parametresi X_1, X_2, \dots, X_n dizisi için ve

σ^2 parametresi X_1, X_2, \dots, X_n dizisi için aynıdır.

H_1 : μ parametresi X_v 'den ($1 \leq v < n$) sonra ve

σ^2 parametresi X_v 'den ($1 \leq v < n$) sonra değişir.

hipotezlerini, Schwarz bilgi kriterlerini elde ederek test etmişlerdir.

Beklenen değeri $\mu = \frac{1}{\lambda}$ olan üstel dağılımlı X_1, X_2, \dots, X_n raslantı değişkenleri dizisinde,

H_0 : μ parametresi X_1, X_2, \dots, X_n dizisi için aynıdır.

H_1 : μ parametresi X_v 'den ($1 \leq v < n$) sonra değişir.

hipotezlerini test etmek için Worsley (1986) ve Siegmund (1988), olabilirlik oranına dayalı test istatistiğini elde etmişlerdir.

Parametreleri m ve λ olan gamma dağılımlı X_1, X_2, \dots, X_n raslantı değişkenleri dizisinde,

H_0 : λ parametresi X_1, X_2, \dots, X_n dizisi için aynıdır.

H_1 : λ parametresi X_v 'den ($1 \leq v < n$) sonra değişir.

hipotezlerini test etmek için Hsu (1979), olabilirlik oranına dayalı test istatistiğini elde etmiştir. Ramanayake (2004), parametreleri m ve λ olan gamma dağılımlı X_1, X_2, \dots, X_n raslantı değişkenleri dizisinde,

H_0 : m parametresi X_1, X_2, \dots, X_n dizisi için aynıdır.

H_1 : m parametresi X_v 'den ($1 \leq v < n$) sonra değişir.

hipotezlerinin testi için olabilirlik oranına dayalı test istatistiğini elde etmiştir.

Parametreleri n ve p olan ikiterimli dağılımlı X_1, X_2, \dots, X_n raslantı değişkenleri dizisinde,

H_0 : p parametresi X_1, X_2, \dots, X_n dizisi için aynıdır.

H_1 : p parametresi X_v 'den ($1 \leq v < n$) sonra değişir.

hipotezlerini test etmek için Worsley (1983), kümülatif toplama ve olabilirlik oranına dayalı test istatistiklerini elde etmiştir.

Parametresi λ olan Poisson dağılımlı X_1, X_2, \dots, X_n raslantı değişkenleri dizisinde,

H_0 : λ parametresi X_1, X_2, \dots, X_n dizisi için aynıdır.

H_1 : λ parametresi X_v 'den ($1 \leq v < n$) sonra değişir.

hipotezlerini test etmek için Boudjellaba, MacGibbon ve Sawyer (2001), kümülatif toplama ve olabilirlik oranına dayalı test istatistiklerini elde etmişlerdir.

2.3. Homojen Poisson Sürecinde Değişme Noktası Sorunu

Poisson süreçleri (homojen, homojen olmayan ve birleşik) endüstri, biyoloji, tıp, ekoloji, aktüerya, demografi vb. birçok uygulama alanına sahip olması nedeniyle olasılıksal süreçlerin en temel süreçlerindedir.

$\{N_t, t \geq 0\}$, λ ortalama oranı ile homojen Poisson süreci olsun. λ , birim zamanda ortaya çıkması beklenen olay sayısıdır. Homojen Poisson sürecinde $(0, t]$ zaman aralığında ortaya çıkan olay sayısı N_t 'nin olasılık fonksiyonu aşağıdaki gibidir:

$$P(N_t = i) = e^{-\lambda t} \frac{(\lambda t)^i}{i!}, \quad i = 0, 1, 2, \dots$$

$$= 0, \quad \text{öteki değerler.}$$

$\{N_t, t \geq 0\}$ homojen Poisson sürecinde $S_0 = 0$ olmak üzere, olayların ortaya çıkış zamanları S_1, S_2, \dots ve olaylar arasında geçen zamanlar (waiting times) $T_1 = S_1 - S_0, T_2 = S_2 - S_1, \dots$ olsun. T_1, T_2, \dots 'ler üstel dağılıma sahip bağımsız raslantı değişkenleridir ve bu raslantı değişkenlerinin olasılık yoğunluk fonksiyonu,

$$f_T(t) = \lambda e^{-\lambda t}, \quad t > 0$$

$$= 0, \quad t \leq 0$$

biçimindedir.

Homojen Poisson sürecinde λ 'nın λ_1 olan değerinin bilinmeyen bir zaman noktasından sonra λ_2 ($\lambda_1 \neq \lambda_2$) olduğu durumu düşünelim ve parametrenin değiştiği zaman noktasını da τ ile gösterelim. Raslantı değişkenleri dizisinde değişme noktası kesikli bir parametre olarak alınmasına karşın, burada τ sürekli bir parametredir. τ 'nin tahmini ile birlikte λ_1 ve λ_2 'nin tahminleri de önemlidir.

2.3.1. Poisson Sürecinde Değişme Noktasının ve Parametrelerin Tahminine İlişkin Çalışmalar

Homojen Poisson sürecinde değişme noktası sorununun çözümüne ilişkin Akman ve Raftery (1986)'in yaptıkları iki çalışma vardır. Bu çalışmalardan birincisinde $\{N_t, t \geq 0\}$ homojen Poisson sürecinde $(0, T]$ zaman aralığında n olay gözlemlendiği;

$\tau \rightarrow \infty, T \rightarrow \infty$ için, $\frac{\tau}{T} = \theta$ 'nin bir değişmez, $0 < a \leq \theta \leq b < 1$ olduğu varsayımı altında

τ ve λ_1, λ_2 parametreleri hakkındaki çıkarımlara,

$$X(s; c, d) = \{s'(1-s')\}^{1/2} \left\{ \frac{N_{sT} - N_{cT}}{s'} - \frac{N_{dT} - N_{sT}}{1-s'} \right\}, \quad 0 \leq c \leq s \leq d \leq 1$$

sürecine dayandırılmıştır. Burada, $s' = \frac{s-c}{d-c}$ olarak tanımlanmış ve tahmin ediciler,

$$\hat{\theta} = \inf \left\{ s : X(s; 0, 1) = \sup_{a \leq u \leq b} X(u; 0, 1) \right\}, \quad (2.12)$$

$$\hat{\tau} = \hat{\theta}T,$$

$$\hat{\lambda}_1 = \frac{N_{\hat{\tau}}}{\hat{\tau}},$$

$$\hat{\lambda}_2 = \frac{n - N_{\hat{\tau}}}{T - \hat{\tau}}$$

biçiminde bulunmuştur (Akman and Raftery, 1986). İkinci çalışma ise Bayesci çıkarsamaya dayalıdır. τ , λ_1 ve λ_2 'ye ilişkin önsel olasılık yoğunluk fonksiyonları verildiğinde, sonsal olasılık yoğunluk fonksiyonları Bayes teoreminin uygulanmasıyla elde edilmiştir. Raftery ve Akman (1986)'ın bu çalışmasında, homojen Poisson sürecinde $(0, T]$ zaman aralığında n olayın gözlemlendiği varsayılmış; olaylar arasında geçen T_1, T_2, \dots, T_n zamanlarının üstel dağılımlı olduğu düşünülerek olabilirlik fonksiyonu,

$$L(\tau, \lambda_1, \lambda_2) = \lambda_1 e^{-\lambda_1 t_1} \lambda_1 e^{-\lambda_1 t_2} \dots \lambda_1 e^{-\lambda_1 t_{N_\tau}} \lambda_2 e^{-\lambda_2 t_{N_\tau+1}} \dots \lambda_2 e^{-\lambda_2 t_n}$$

biçiminde alınmıştır ve $s_0 = 0$, $t_1 = s_1 - s_0, \dots, t_n = s_n - s_{n-1}$, $s_n = T$ olduğu için,

$$\begin{aligned} L(\tau, \lambda_1, \lambda_2) &= \lambda_1 e^{-\lambda_1 (s_1 - s_0)} \lambda_1 e^{-\lambda_1 (s_2 - s_1)} \dots \lambda_1 e^{-\lambda_1 (\tau - s_{N_\tau})} \\ &\quad \times \lambda_2 e^{-\lambda_2 (s_{N_\tau+1} - s_{N_\tau})} \dots \lambda_2 e^{-\lambda_2 (s_n - s_{n-1})} \\ &= \lambda_1^{N_\tau} e^{-\lambda_1 \tau} \lambda_2^{n - N_\tau} e^{-\lambda_2 (T - \tau)} \end{aligned} \quad (2.13)$$

elde edilmiştir. Burada τ , bir olayın ortaya çıkma zamanına karşılık gelmektedir. λ_1 ve λ_2 için önsel olasılık yoğunluk fonksiyonları, $i = 1, 2$ için,

$$f(\lambda_i) = \frac{a_i^{b_i+1}}{\Gamma(b_i+1)} \lambda_i^{b_i} e^{-a_i \lambda_i}, \quad \lambda_i > 0 \quad (2.14)$$

$$= 0, \quad \lambda_i \leq 0$$

ve τ 'nin önsel olasılık yoğunluk fonksiyonu $f(\tau)$ biçiminde alınmış; sonsal olasılık yoğunluk fonksiyonu, $t = (t_1, t_2, \dots, t_n)$ olmak üzere,

$$f(\tau, \lambda_1, \lambda_2 / t) = \frac{a_1^{b_1+1} a_2^{b_2+1}}{\Gamma(b_1+1) \Gamma(b_2+1)} \lambda_1^{N_\tau+b_1} e^{-\lambda_1(\tau+a_1)} \lambda_2^{n-N_\tau+b_2} e^{-\lambda_2(T-\tau+a_2)} f(\tau)$$

biçiminde elde edilmiştir. Buradan τ , λ_1 ve λ_2 'in bileşen (marjinal) sonsal olasılık yoğunluk fonksiyonları bulunmuştur.

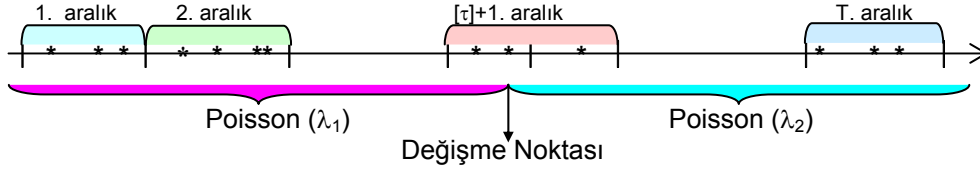
$\{N_t, t \geq 0\}$ homojen Poisson sürecinde S_1, S_2, \dots, S_n gözlem zamanlarında ortaya çıkan olay sayıları sırasıyla $N_{S_1}, N_{S_2}, \dots, N_{S_n}$ olsun. Bu raslantı değişkenleri içinde dağılımın parametresi olan λ_1 'in bilinmeyen bir zaman noktasında değiştiğinin incelenmesi $N_{S_1}, N_{S_2}, \dots, N_{S_n}$ raslantı değişkenlerinin bağımlı olmaları nedeniyle güçtür. Ancak bu güçlük, West ve Ogden (1997)'in çalışmasında farklı bir yaklaşımla aşılmaya çalışılmıştır. Bu çalışmada, $(0, T]$ zaman aralığı birim aralıklara (periyotlara) ayrılmış ve bu aralıkların her birinde ortaya çıkan olay sayıları sırasıyla X_1, X_2, \dots, X_T olarak alınmıştır. Değişme noktası τ 'nin $([\tau]+1)$ 'inci aralıkta ortaya çıktığı varsayımı altında, $\rho(\tau) = \tau - [\tau]$ olmak üzere, X_1, X_2, \dots, X_T raslantı değişkenlerinin dağılımı,

$$X_i \sim \text{Poisson}(\lambda_1), \quad i = 1, 2, \dots, [\tau]$$

$$X_i \sim \text{Poisson}(\rho(\tau) \lambda_1 + (1 - \rho(\tau)) \lambda_2), \quad i = [\tau]+1$$

$$X_i \sim \text{Poisson}(\lambda_2), \quad i = [\tau]+2, [\tau]+3, \dots, T$$

biçiminde tanımlanmıştır. Burada $[.]$ ifadesi, tam değer fonksiyonunu gösterir. Yukarıda tanımlanan aralıklar ve bu aralıklarda ortaya çıkan olayların gösterimi Şekil 2.2'de verildiği gibidir:



Şekil 2.2 Poisson sürecinde birim zaman aralıklarında ortaya çıkan olaylar ve değişme noktası

τ 'nın bilindiği varsayımı altında, λ_1 ve λ_2 parametrelerinin en çok olabilirlik tahmin edicileri,

$$\begin{aligned} \log L(\lambda_1, \lambda_2, \tau) = & -\tau\lambda_1 + \sum_{i=1}^{[\tau]} x_i \log(\lambda_1) - (T - (\tau + 1))\lambda_2 + \sum_{i=[\tau]+2}^T x_i \log(\lambda_2) \\ & - \rho(\tau)\lambda_1 - (1 - \rho(\tau))\lambda_2 + x_{[\tau]+1} \log(\rho(\tau)\lambda_1 + (1 - \rho(\tau))\lambda_2) - \log \sum_{i=1}^T x_i! \end{aligned}$$

log olabilirlik fonksiyonundan aşağıdaki gibi elde edilmiştir:

$$\begin{aligned} \hat{\lambda}_1 = & \frac{1}{\tau} \left\{ \sum_{i=1}^{[\tau]} X_i + \frac{\rho(\tau)\hat{\lambda}_1}{\rho(\tau)\hat{\lambda}_1 + (1 - \rho(\tau))\hat{\lambda}_2} X_{[\tau]+1} \right\}, \\ \hat{\lambda}_2 = & \frac{1}{T - \tau} \left\{ \sum_{i=[\tau]+2}^T X_i + \frac{(1 - \rho(\tau))\hat{\lambda}_2}{\rho(\tau)\hat{\lambda}_1 + (1 - \rho(\tau))\hat{\lambda}_2} X_{[\tau]+1} \right\}. \end{aligned} \quad (2.15)$$

Çalışmada ayrıca i^* 'ın bilindiği ve $\tau \in [i^*, i^* + 1)$ olduğu varsayımı altında; λ_1, λ_2 ve τ 'nın en çok olabilirlik tahmin edicileri, $Y_i = X_1 + X_2 + \dots + X_i$ olmak üzere,

$$\hat{\lambda}_1 = \frac{Y_{i^*}}{i^*}, \quad \hat{\lambda}_2 = \frac{Y_T - Y_{i^*+1}}{T - (i^* + 1)}, \quad \hat{\tau} = i^* + \frac{X_{i^*+1} - \hat{\lambda}_2}{\hat{\lambda}_1 - \hat{\lambda}_2} \quad (2.16)$$

biçiminde bulunmuştur (West and Ogden, 1997).

Loader (1992)'ın yaptığı çalışmada, homojen olmayan Poisson sürecinde değişme noktasının olup olmadığına ilişkin hipotezler ele alınmış; önem düzeyi ve testin gücünün hesaplanmasında büyük sapma (large deviation) yöntemleri kullanılmıştır. Bu çalışmada, $(0, T]$ gözlem zamanında ortalama oran $\lambda(t)$,

$$\begin{aligned}\lambda(t) &= e^{a+bt}, & 0 < t \leq \tau \\ &= e^{a+\delta+bt}, & \tau < t \leq T\end{aligned}\quad (2.17)$$

olarak alınmıştır. Burada a , b , δ ve τ bilinmeyen parametrelerdir. Eşitlik (2.17), log-doğrusal-değişme noktası (log-linear and change point) modeli olarak adlandırılır. Eşitlik (2.17)'de verilen $\lambda(t)$ ortalama oranı ile homojen olmayan Poisson sürecinde olayların S_1, S_2, \dots, S_n zamanlarında gözlemlendiği ve $N_T = n$ olduğu varsayılarak olabilirlik fonksiyonu,

$$L(\tau, a, \delta, b) = \exp \left\{ an + \delta(n - N_\tau) + b \sum_{i=1}^n s_i - \frac{e^a}{b} (e^{b\tau} - 1 + e^\delta (e^{bT} - e^{b\tau})) \right\}$$

biçiminde alınmış ve değişmez τ için, a , b , δ 'nin en çok olabilirlik tahmin edicileri,

$$\begin{aligned}\hat{a} &= \log \left(\frac{\hat{b} N_\tau}{e^{\hat{b}\tau} - 1} \right), \\ \hat{\delta} &= \log \left(\frac{(n - N_\tau)(e^{\hat{b}\tau} - 1)}{N_\tau (e^{\hat{b}T} - e^{\hat{b}\tau})} \right), \\ \sum_{i=1}^n S_i &= N_\tau \frac{\tau e^{\hat{b}\tau}}{e^{\hat{b}\tau} - 1} + (n - N_\tau) \frac{T e^{\hat{b}T} - \tau e^{\hat{b}\tau}}{e^{\hat{b}T} - e^{\hat{b}\tau}} - \frac{n}{\hat{b}}\end{aligned}\quad (2.18)$$

olarak elde edilmiştir. Burada \hat{b} için verilen çözüm kapalı biçimde olduğundan sayısal olarak elde edilebileceği, ancak olabilirlik fonksiyonunun süreksiz olması nedeniyle τ 'nin en çok olabilirlik tahmin edicisine ulaşamayacağı belirtilmiştir (Loader, 1992).

2.3.2. Poisson Sürecinde Değişme Noktasına İlişkin Hipotez Testi

Raftery ve Akman (1986)'nin yaptıkları çalışmada, homojen Poisson sürecinde değişme noktası için, $H_0 : \lambda_1 = \lambda_2$ hipotezinin doğru olduğu varsayımı altındaki model M_0 ile, $H_1 : \lambda_1 \neq \lambda_2$ hipotezin doğru olduğu varsayımı altındaki model M_1 'in karşılaştırıldığı Bayes faktörü,

$$B_{0,1} = \frac{f(t/M_0)}{f(t/M_1)}$$

biçiminde vermişlerdir. Eşitlik (2.14)'te verilen önsel olasılık yoğunluk fonksiyonları kullanılarak,

$$f(t/M_0) = \int_0^{\infty} f(t/\lambda_1) f(\lambda_1) d\lambda_1 = \frac{\Gamma(n+b_1+1)}{\Gamma(b_1+1)} \frac{a_1^{b_1+1}}{(a_1+T)^{n+b_1+1}}$$

ve

$$\begin{aligned} f(t/M_1) &= \int_0^T \int_0^{\infty} \int_0^{\infty} f(t/\lambda_1, \lambda_2, \tau) f(\lambda_1, \lambda_2, \tau) d\lambda_1 d\lambda_2 d\tau \\ &= \frac{a_1^{b_1+1} a_2^{b_2+1}}{\Gamma(b_1+1) \Gamma(b_2+1)} \sum_{i=0}^n \Gamma(i+b_1+1) \Gamma(n-i+b_2+1) \\ &\quad \times \int_{S_i}^{S_{i+1}} (\tau+a_1)^{-(i+b_1+1)} (T-\tau+a_2)^{-(n-i+b_2+1)} f(\tau) d\tau \end{aligned}$$

bulunmuştur (Rafferty and Akman, 1986). H_0 hipotezinin reddedilmesi için, Bayes faktörünün logaritmasının sıfıra yakın olması gerektiğini belirtmişlerdir.

Loader (1992)'in çalışmasında, Eşitlik (2.17)'de verilen $\lambda(t)$ ortalama oranı ile homojen olmayan Poisson sürecinde,

$$H_0 : \delta = 0$$

$$H_1 : \delta \neq 0$$

hipotezlerinin testi için \hat{b} , b 'nin Eşitlik (2.18)'de elde edilen tahmini ve \hat{b}_0 , b 'nin H_0 hipotezi altında elde edilen tahmini olmak üzere, log olabilirlik oranı istatistiği,

$$\Lambda(\tau) = N_{\tau} \log \left(\frac{\hat{b} N_{\tau} (e^{\hat{b}_0 T} - 1)}{\hat{b}_0 n (e^{\hat{b} \tau} - 1)} \right) + (n - N_{\tau}) \log \left(\frac{\hat{b} (n - N_{\tau}) (e^{\hat{b}_0 T} - 1)}{\hat{b}_0 n (e^{\hat{b} T} - e^{\hat{b} \tau})} \right) + (\hat{b} - \hat{b}_0) \sum_{i=1}^n S_i \quad (2.19)$$

biçiminde verilmiş; τ bilinmediği için olabilirlik oranı testinin Eşitlik (2.19)'un en büyüklenmesine dayanacağı belirtilmiştir.

2.4. İngiltere'deki Kömür Madeni Kazaları Veri Kümesi

Literatürde değişme noktasını ele alan çalışmaların önemli bir kesiminde İngiltere'de 15 Mart 1852 ve 22 Mart 1962 yılları arasında 10 kişiden daha çok kişinin öldüğü 191 kömür madeni kazasının yer aldığı gerçek veri kümesi kullanılmıştır. Bu veriler ilk olarak Magnuire, Pearson ve Wynn (1952)'in çalışmalarında kullanılmış; daha sonra Jarett (1979) tarafından düzeltilmiştir (West and Ogden, 1997). Kazalar arasında geçen zamanlar (gün olarak) Çizelge 2.1'de verilmiştir.

Çizelge 2.1. Maden kazaları arasında geçen günler (yukarıdan aşağıya doğru) (Jarett, 1979)

157	65	53	93	127	176	22	1205	1643	312
123	186	17	24	218	55	61	644	54	536
2	23	538	91	2	93	78	467	326	145
124	92	187	143	0	59	99	871	1312	75
12	197	34	16	378	315	326	48	348	364
4	431	101	27	36	59	275	123	745	37
10	16	41	144	15	61	54	456	217	19
216	154	139	45	31	1	217	498	120	156
80	95	42	6	215	13	113	49	275	47
12	25	1	208	11	189	32	131	20	129
33	19	250	29	137	345	388	182	66	1630
66	78	80	112	4	20	151	255	292	29
232	202	3	43	15	81	361	194	4	217
826	36	324	193	72	286	312	224	368	7
40	110	56	134	96	114	354	566	307	18
12	276	31	420	124	108	307	462	336	1358
29	16	96	95	50	188	275	228	19	2366
190	88	70	125	120	233	78	806	329	952
97	225	41	34	203	28	17	517	330	632

West ve Ogden (1997)'in yaptıkları çalışmada, yukarıdaki veri kümesi kullanılarak Eşitlik (2.16)'daki i^* 'in çeşitli değerleri için τ 'nun nokta ve aralık tahminleri elde edilmiş; ayrıca daha önceki çalışmalarda bu veri kümesi kullanılarak τ için elde edilen nokta ve aralık tahminleri bir çizelgede verilmiştir (Çizelge 2.2). Çizelge 2.2'de Worsley'e ait değerler, 1986'da üstel ailedeki değişme noktasına ilişkin yaptığı çalışmada elde edilen kuramsal sonuçların bu veri kümesine uygulanmasıyla elde edilen değerlerdir.

Çizelge 2.2. İngiltere'deki kömür madeni kazaları arasında geçen günlerin sayısını içeren veri kümesi için değişme noktasının nokta tahminleri ve aralık tahminleri (West and Ogden, 1997)

Yöntem	τ 'nun tahmini	Alt Sınır	Üst Sınır
Loader (1992)	10 Mart 1890	8.08 yıl	
Akman ve Raftery (1986)	10 Mart 1890	6 Ekim 1886	17 Aralık 1898
Rafter ve Akman (1986)	27 Ağustos 1890	15 Mayıs 1887	3 Ağustos 1895
Worsley (1986)	1890	1884	1895

Kesim 2.3'te açıklanan çalışmalarda, homojen Poisson süreci için τ değişme noktası ile λ_1 ve λ_2 parametrelerinin Eşitlik (2.12)'de verilen tahmin edicileri $\frac{\tau}{T}$ 'nin, $0 < a \leq \frac{\tau}{T} \leq b < 1$ aralığında olduğu varsayımı altında elde edilmiş; ancak burada a ve b değerlerinin belirlenmesi bir sorundur. Ayrıca τ değişme noktasına ilişkin elde edilen tahmin ediciden tahminin elde edilmesi kolay değildir. West ve Ogden (1997)'in çalışmasında λ_1 ve λ_2 parametrelerinin tahminleri ancak τ bilindiğinde bulunabilir. Yine bu çalışmada τ 'nun i^* ile i^*+1 aralığında olduğu biliniyorsa, τ , λ_1 ve λ_2 parametrelerine ilişkin tahminler Eşitlik (2.16)'dan bulunabilir. Ancak i^* 'ın belirlenmesi her zaman mümkün olmayabilir. Ayrıca birden çok olası i^* için birden çok $\hat{\tau}$, $\hat{\lambda}_1$ ve $\hat{\lambda}_2$ tahminleri elde edilir. Eşitlik (2.13)'te verilen olabilirlik fonksiyonu N_τ 'nin da fonksiyonudur. τ bilinmediği için N_τ 'da bilinemez. τ , λ_1 ve λ_2 parametrelerine ilişkin tahminlerin elde edilmesinde N_τ 'nin bulunması bir sorundur. Homojen olmayan Poisson sürecinde, süreç boyunca ortalama oran fonksiyonunda değişme olup olmadığının testi için elde edilen test fonksiyonu olabilirlik oranına dayalıdır (Eşitlik (2.19)). Bu test fonksiyonu kullanılarak α yanılma düzeyinde H_0 hipotezinin reddedildiği bölgenin sınırlarının bulunması güçtür.

Bu çalışmada, yukarıda belirttiğimiz güçlükleri aşmak için homojen Poisson sürecinde τ değişme noktası ile λ_1 ve λ_2 parametrelerinin uygulama kolaylığı sağlayan tahmin edicilerinin; değişmenin olup olmadığını test etmek için, test

istatistiğinin elde edilmesi amaçlanmıştır. Bu amaçla Üçüncü Bölüm' de, λ_1 ve λ_2 bilindiğinde τ 'nın; λ_1 bilindiğinde τ ve λ_2 'nin; λ_1 ve λ_2 bilinmediğinde τ , λ_1 ve λ_2 'nin en çok olabilirlik yöntemine dayalı tahmin edicilerinin ve olabilirlik oranına dayalı test istatistiğinin nasıl elde edildikleri açıklanacak; yapılan uygulamalar verilecektir.

3. HOMOJEN POISSON SÜRECİNDE DEĞİŞME NOKTASININ TAHMİNİ ve HİPOTEZ TESTİ

Poisson sürecinde değişme noktasını ele alan çalışmalar incelendiğinde, değişme noktasının tahmin edicisi için kullanışlı bir çözüme ulaşılamadığı görülmektedir. Bu çalışmada Kesim 2.3'te açıklanan çalışmalardan yararlanılarak, ancak farklı bir yaklaşımla homojen Poisson sürecinde τ değişme noktasının, λ_1 ve λ_2 parametrelerinin tahmin edicileri elde edilmiş; değişme olup olmadığına ilişkin hipotez testi için test istatistiği önerilmiştir.

$\{N_t, t \geq 0\}$ homojen Poisson sürecinde değişme noktası τ , $(0, T]$ zaman aralığında sürekli bir parametre olarak ele alınmış; $(0, \tau]$ zaman aralığında birim zamanda ortaya çıkması beklenen olay sayısı λ_1 ve $(\tau, T]$ zaman aralığında birim zamanda ortaya çıkması beklenen olay sayısı λ_2 olmak üzere, tahmin ediciler aşağıda verilen üç durum için elde edilmiştir:

1. λ_1 ve λ_2 'nin bilindiği durum;
2. λ_1 'in bilindiği, λ_2 'nin bilinmediği durum;
3. λ_1 ve λ_2 'nin bilinmediği durum.

3.1. Nokta Tahmini

Bu çalışmada, τ , λ_1 ve λ_2 'nin tahmin edicilerini elde etmek için $(0, T]$ zaman aralığında n olayın gözleendiği, n 'inci olay ortaya çıkıncaya dek geçen S_n zamanı için $S_n=T$ ve herhangi bir v 'üncü ($v=1, 2, \dots, n-1$) olay ortaya çıkıncaya dek geçen S_v zamanı için $S_v=\tau$ olduğu varsayılmıştır. Bu varsayımlar altında Eşitlik (2.13)'te verilen olabilirlik fonksiyonundan yararlanılabileceği düşünülmüştür. Ancak bu fonksiyonun N_τ raslantı değişkenine bağlı olması τ , λ_1 ve λ_2 'nin tahmin edicilerinin elde edilmesinde güçlük yaratmaktadır. Bu güçlüğü aşılmasında $(0, T]$ zaman aralığında ortaya çıkan olay sayısı bilindiğinde, N_τ 'nin aşağıda verilen koşullu dağılımından yararlanılmıştır:

$$P(N_\tau = k / N_T = n) = \binom{n}{k} \left(\frac{\lambda_1 \tau}{\lambda_1 \tau + \lambda_2 (T - \tau)} \right)^k \left(\frac{\lambda_2 T}{\lambda_1 \tau + \lambda_2 (T - \tau)} \right)^{n-k}, \quad k=0,1,\dots,n \quad (3.1)$$

= 0, öteki değerler.

Eşitlik (3.1)'den N_τ 'nin koşullu dağılımının, $\left(n, \frac{\lambda_1 \tau}{\lambda_1 \tau + \lambda_2 (T - \tau)} \right)$ parametreleri ile ikiterimli dağılım olduğu görülmektedir. N_τ 'nin beklenen değeri Eşitlik (3.2)'de verilmiştir:

$$E(N_\tau / N_T = n) = \frac{n \lambda_1 \tau}{\lambda_1 \tau + \lambda_2 (T - \tau)}. \quad (3.2)$$

3.1.1. λ_1 ve λ_2 'nin Bilindiği Durumda τ 'nin Tahmini

λ_1 ve λ_2 bilindiğinde, Eşitlik (2.13)'te verilen olabirlik fonksiyonu τ ve N_τ 'nin fonksiyonudur:

$$L_1(\tau, N_\tau) = \lambda_1^{N_\tau} e^{-\lambda_1 \tau} \lambda_2^{n-N_\tau} e^{-\lambda_2 (T-\tau)}.$$

Bu fonksiyonun logaritmasının beklenen değeri, Eşitlik (3.1) ve Eşitlik (3.2)'ye göre bulunursa,

$$E[\log L_1(\tau, N_\tau)] = \frac{n \lambda_1 \tau}{\lambda_1 \tau + \lambda_2 (T - \tau)} \log \lambda_1 + \left(n - \frac{n \lambda_1 \tau}{\lambda_1 \tau + \lambda_2 (T - \tau)} \right) \log \lambda_2 - \lambda_1 \tau - \lambda_2 (T - \tau) \quad (3.3)$$

elde edilir. Görüldüğü gibi, bu fonksiyon yalnızca τ 'nin bir fonksiyonudur. Eşitlik (3.3)'teki fonksiyonun τ 'ya göre türevi alınır ve sıfıra eşitlenirse,

$$n \lambda_1 \lambda_2 T \log \left(\frac{\lambda_1}{\lambda_2} \right) - (\lambda_1 - \lambda_2) [(\lambda_1 - \lambda_2) \tau + \lambda_2 T]^2 = 0, \quad (3.4)$$

$$(\lambda_1 - \lambda_2)^3 \tau^2 + 2(\lambda_1 - \lambda_2)^2 \lambda_2 T \tau + (\lambda_1 - \lambda_2) (\lambda_2 T)^2 - n \lambda_1 \lambda_2 T \log \left(\frac{\lambda_1}{\lambda_2} \right) = 0$$

eşitliğine ulaşılır. Bu ifade τ 'ya göre çözülmüşse, τ 'nın en çok olabilirlik tahmin edicisi aşağıdaki gibi bulunur:

$$\hat{\tau}_1 = \frac{-(\lambda_1 - \lambda_2)^2 \lambda_2 T + \sqrt{n \lambda_1 \lambda_2 (\lambda_1 - \lambda_2)^3 \left[\log \left(\frac{\lambda_1}{\lambda_2} \right) \right] T}}{(\lambda_1 - \lambda_2)^3}, \quad (3.5a)$$

$$\hat{\tau}_2 = \frac{-(\lambda_1 - \lambda_2)^2 \lambda_2 T - \sqrt{n \lambda_1 \lambda_2 (\lambda_1 - \lambda_2)^3 \left[\log \left(\frac{\lambda_1}{\lambda_2} \right) \right] T}}{(\lambda_1 - \lambda_2)^3}. \quad (3.5b)$$

Eşitlik (3.5a) ve (3.5b)'den elde edilen tahminlerden $(0, T]$ aralığında olanı τ 'nın en çok olabilirlik tahmini olacaktır. Eşitlik (3.5a)'dan elde edilen tahminin, her zaman $(0, T]$ aralığının içinde, Eşitlik (3.5b)'den elde edilen tahminin de her zaman $(0, T]$ aralığının dışında olduğu benzetim çalışmaları sonucunda görülmüştür.

Eşitlik (3.3)'te verilen beklenen değer, $L_1(\tau, N_\tau)$ olabilirlik fonksiyonunun N_τ 'ya göre tanımlanan bir beklenen değeridir. Bu eşitliğin en büyüklenmesiyle değişme noktasının tahmin edicisine ulaşılmıştır. Burada tahmin edicilerin elde edilmesinde olabilirlik fonksiyonun beklenen değerinin en büyüklenmesi, son yıllarda uygulama kolaylığı yönünden tercih edilen ve özellikle tam olmayan veri yapılarında parametre tahminlerinde başarılı sonuçlar veren EM (Expected Maximization:EM) algoritmasına bir benzerlik göstermektedir. Ancak bu çalışmada tahminlerin elde edilmesinde kullanılan beklenen değer en büyüklenmesi yaklaşımı, EM algoritması gibi yineleme içeren yöntemlerden farklıdır.

3.1.2. λ_1 'in Bilindiği, λ_2 'nin Bilinmediği Durumda τ ve λ_2 'nin Tahmini

Bu kesimde, önce λ_2 'nin tahmin edicisi daha sonra da τ 'nın tahmin edicisi elde edilecektir. λ_1 bilindiğinde Eşitlik (2.13)'te verilen olabilirlik fonksiyonu τ, N_τ ve λ_2 'nin fonksiyonu olur:

$$L_2(\tau, N_\tau, \lambda_2) = \lambda_1^{N_\tau} e^{-\lambda_1 \tau} \lambda_2^{n - N_\tau} e^{-\lambda_2 (T - \tau)}. \quad (3.6)$$

Yukarıdaki fonksiyonun λ_1 'i içeren kesimi, τ 'dan önce elde edilen örnekleme ve λ_2 'yi içeren kesimi, τ 'dan sonra elde edilen örnekleme ilişkin olabilirlik fonksiyonları olarak düşünülebilir. λ_2 'nin tahmin edicisini elde etmek için λ_2 'yi içeren kesimini göz önüne almak yeterlidir:

$$L_3(\tau, N_\tau, \lambda_2) = \lambda_2^{n-N_\tau} e^{-\lambda_2(T-\tau)}. \quad (3.7)$$

Bu fonksiyonun logaritmasının beklenen değeri, önceki örnekleme ortaya çıkan olay sayısı N_τ 'nin beklenen değerinin $\lambda_1\tau$ 'ya eşit olduğu dikkate alınarak aşağıdaki gibi bulunur:

$$E[\log L_3(\tau, N_\tau, \lambda_2)] = (n - \lambda_1\tau) \log(\lambda_2) - \lambda_2(T - \tau). \quad (3.8)$$

Eşitlik (3.8)'deki ifadenin λ_2 'ye göre türevi alınıp, sıfıra eşitlenirse,

$$\begin{aligned} \frac{n - \lambda_1\tau}{\lambda_2} - (T - \tau) &= 0, \\ -\lambda_2(T - \tau) - n + \lambda_1\tau &= 0, \\ \lambda_2 &= \frac{n - \lambda_1\tau}{T - \tau} \end{aligned} \quad (3.9)$$

elde edilir.

λ_1 bilindiğine ve λ_2 Eşitlik (2.9)'da verildiği gibi τ 'nin fonksiyonu olarak bulunduğuna göre, τ 'nin tahmin edicisini elde etmek için Eşitlik (2.13)'teki olabilirlik fonksiyonunu dolayısıyla Eşitlik (3.3) ve Eşitlik (3.4)'ü kullanmak gerekir. Eşitlik (3.4)'te λ_2 yerine Eşitlik (3.9)'daki değeri konulursa, τ 'nin tahmin edicisi,

$$\lambda_1 T(n - \lambda_1 \hat{\tau}) \log\left(\frac{\lambda_1 T - \lambda_1 \hat{\tau}}{n - \lambda_1 \hat{\tau}}\right) - n\lambda_1 T + n^2 = 0 \quad (3.10)$$

elde edilir. Buna göre λ_2 'nin tahmin edicisi,

$$\hat{\lambda}_2 = \frac{n - \lambda_1 \hat{\tau}}{T - \hat{\tau}} \quad (3.11)$$

olur. τ 'nın Eşitlik (3.10)'da verilen tahmin edicisi kapalı biçimde olmakla birlikte, sayısal değeri matematiksel paket programların yardımıyla hesaplanabilmektedir.

3.1.3. λ_1 ve λ_2 'nin Bilinmediği Durumda Tahminler

λ_1 ve λ_2 'nin bilinmediği durumda τ , λ_1 ve λ_2 'nin tahmin edicilerinin analitik olarak bulunması oldukça güçtür. Bu durumda, τ 'nın örneklem verilerine göre alabileceği değerlerin s_1, s_2, \dots, s_n olacağı düşünülerek,

$$\hat{\lambda}_1 = \frac{N_{s_i}}{s_i} = \frac{i}{s_i}, \quad \hat{\lambda}_2 = \frac{n - N_{s_i}}{T - s_i} = \frac{n - i}{T - s_i}, \quad i = 1, 2, \dots, n \quad (3.12)$$

tahmin edicilerinden $(\hat{\lambda}_{1_1}, \hat{\lambda}_{2_1}), (\hat{\lambda}_{1_2}, \hat{\lambda}_{2_2}), \dots, (\hat{\lambda}_{1_n}, \hat{\lambda}_{2_n})$ tahminleri elde edilir. $i = 1, 2, \dots, n$ için, Eşitlik (3.3)'teki fonksiyonu en büyük yapan $(s_i, \hat{\lambda}_{1_i}, \hat{\lambda}_{2_i})$ değerleri, $\tau, \lambda_1, \lambda_2$ 'nin $\hat{\tau}, \hat{\lambda}_1, \hat{\lambda}_2$ en çok olabilirlik tahminleri olacaktır.

3.1.4. Uygulama

Bu kesimde, Kesim 3.1.1, Kesim 3.1.2 ve Kesim 3.1.3'te elde edilen tahmin edicilerin gerçek değerlere yakın tahminler verdiklerini göstermek için, Poisson sürecinde olaylar arasında geçen zamanın üstel dağılımlı olduğu göz önüne alınmıştır. Çeşitli λ_1, λ_2 ($\lambda_1 \neq \lambda_2$) değerleri için λ_1 parametresine sahip üstel dağılımdan v birimlik, λ_2 parametresine sahip üstel dağılımdan $n-v$ birimlik olmak üzere toplam n birimlik $T_1, \dots, T_v, \dots, T_n$ örneklemeleri seçilmiştir. T_1 'in birinci olay, $T_1 + \dots + T_v$ 'nin v 'üncü olay, $\dots, T_1 + \dots + T_n$ 'nin n 'inci olay ortaya çıkıncaya dek geçen zamanlar olduğu düşünülerek, $\tau = S_v = T_1 + \dots + T_v$ ve $T = S_n = T_1 + \dots + T_n$ olarak alınmış; Eşitlik (3.5a), Eşitlik (3.10), Eşitlik (3.11) ve Eşitlik (3.12)'den τ ve λ_1, λ_2 'nin tahminleri elde edilmiş; gerçek değerlerle karşılaştırılmıştır.

Poisson süreçlerinde değişme noktasına ilişkin önceki çalışmalarda değişmenin gözlenebilmesi ve saptanabilmesi için, $\frac{\tau}{T}$ oranının, $0 < a \leq \frac{\tau}{T} \leq b < 1$ aralığında olması gerektiği belirtilmiş; ancak a ve b 'nin değerleri verilmemiştir. Bu çalışmada ise, Kesim 3.1.1, Kesim 3.1.2 ve Kesim 3.1.3'te verilen her üç durum için,

$n=30,100$ örneklem büyüklüklerinde üstel dağılımdan λ_1 ve λ_2 'nin ($\lambda \neq \lambda_2$) rasgele seçilen 0.01, 0.05, 0.1, 0.5, 1, 2, 5, 8, 10, 20 değerlerine ve çeşitli ν değerlerine göre 500 yinleme içeren benzetim çalışması yapılmış; değişimin gözlenebilmesi ve saptanabilmesi için $\frac{\tau}{T}$ oranının genel olarak $\frac{1}{6} < \frac{\tau}{T} < \frac{5}{6}$ aralığında olması gerektiği bulunmuştur.

3.1.4.1. λ_1 ve λ_2 'nin bilindiği durumda τ 'nun tahmini için örnekler

Burada Kesim 3.1.1'de elde edilen tahmin edicinin gerçek değere yakın tahmin verip vermediğini göstermek için, $n = 30, 100$ örneklem büyüklüklerine ve çeşitli λ_1, λ_2 ($\lambda_1 \neq \lambda_2$) değerlerine göre üstel dağılımdan rasgele örneklemeler çekilmiş; bu örneklemeler için, Eşitlik (3.5a)'dan elde edilen tahminler Çizelge 3.1-Çizelge 3.16'da verilmiştir. Ayrıca Eşitlik (3.5a)'dan elde edilen tahminler, Kesim 2.2.1'de açıklandığı gibi, ν 'üncü olaydan önceki ve ν 'üncü olaydan sonraki rasgele yürüyüşleri en büyük yapan değere karşılık gelen τ_w değerleri ile karşılaştırılmış; uyumlu oldukları görülmüştür.

Çizelge 3.1. $\lambda_1 = 0.5, \lambda_2 = 1.5, \nu = 10, n = 30$ ve $T = 38.76919$ için tahminler

$\tau = 26.58168, \frac{\tau}{T} = 0.68564$	
Eşitlik (3.5a)'ya Göre	Rasgele Yürüyüşler
$\hat{\tau}$	τ_w
27.19693	26.58168

Çizelge 3.2. $\lambda_1 = 0.5, \lambda_2 = 1.5, \nu = 15, n = 30$ ve $T = 38.79692$ için tahminler

$\tau = 28.91163, \frac{\tau}{T} = 0.74520$	
Eşitlik (3.5a)'ya Göre	Rasgele Yürüyüşler
$\hat{\tau}$	τ_w
27.22746	28.91168

Çizelge 3.3. $\lambda_1 = 0.5, \lambda_2 = 1.5, \nu = 35, n = 100$ ve $T = 113.15480$ için tahminler

$\tau = 74.15410, \frac{\tau}{T} = 0.65533$	
Eşitlik (3.5a)'ya Göre	Rasgele Yürüyüşler
$\hat{\tau}$	τ_w
73.17396	74.15410

Çizelge 3.4. $\lambda_1 = 0.5, \lambda_2 = 1.5, \nu = 50, n=100$ ve $T = 114.94556$ için tahminler

$\tau = 80.20288, \frac{\tau}{T} = 0.69775$	
Eşitlik (3.5a)'ya Göre	Rasgele Yürüyüşler
$\hat{\tau}$	τ_w
75.09904	71.75337

Çizelge 3.5. $\lambda_1 = 1.5, \lambda_2 = 0.5, \nu = 15, n = 30$ ve $T = 43.47320$ için tahminler

$\tau = 12.94673, \frac{\tau}{T} = 0.29781$	
Eşitlik (3.5a)'ya Göre	Rasgele Yürüyüşler
$\hat{\tau}$	τ_w
11.04456	12.94673

Çizelge 3.6. $\lambda_1 = 1.5, \lambda_2 = 0.5, \nu = 20, n = 30$ ve $T = 39.31683$ için tahminler

$\tau = 11.44214, \frac{\tau}{T} = 0.29102$	
Eşitlik (3.5a)'ya Göre	Rasgele Yürüyüşler
$\hat{\tau}$	τ_w
11.51632	11.44214

Çizelge 3.7. $\lambda_1 = 1.5, \lambda_2 = 0.5, \nu = 50, n = 100$ ve $T = 153.80751$ için tahminler

$\tau = 35.71465, \frac{\tau}{T} = 0.23220$	
Eşitlik (3.5a)'ya Göre	Rasgele Yürüyüşler
$\hat{\tau}$	τ_w
35.67116	36.60768

Çizelge 3.8. $\lambda_1 = 1.5, \lambda_2 = 0.5, \nu = 65, n = 100$ ve $T = 119.28360$ için tahminler

$\tau = 43.34456, \frac{\tau}{T} = 0.36337$	
Eşitlik (3.5a)'ya Göre	Rasgele Yürüyüşler
$\hat{\tau}$	τ_w
39.49690	43.34456

Çizelge 3.9. $\lambda_1 = 1, \lambda_2 = 5, \nu = 10, n = 30$ ve $T = 15.70841$ için tahminler

$\tau = 12.37520, \frac{\tau}{T} = 0.78810$	
Eşitlik (3.5a)'ya Göre	Rasgele Yürüyüşler
$\hat{\tau}$	τ_w
11.93785	12.37520

Çizelge 3.10. $\lambda_1 = 1, \lambda_2 = 5, \nu = 15, n = 30$ ve $T = 14.39305$ için tahminler

$\tau = 11.63100, \frac{\tau}{T} = 0.80810$	
Eşitlik (3.5a)'ya Göre	Rasgele Yürüyüşler
$\hat{\tau}$	τ_w
10.62298	11.63100

Çizelge 3.11. $\lambda_1 = 1, \lambda_2 = 5, v = 35, n = 100$ ve $T = 48.52171$ için tahminler

$\tau = 37.94226, \frac{\tau}{T} = 0.78196$	
Eşitlik (3.5a)'ya Göre	Rasgele Yürüyüşler
$\hat{\tau}$	τ_w
35.95196	37.94226

Çizelge 3.12. $\lambda_1 = 1, \lambda_2 = 5, v = 50, n = 100$ ve $T = 56.57775$ için tahminler

$\tau = 45.71201, \frac{\tau}{T} = 0.80795$	
Eşitlik (3.5a)'ya Göre	Rasgele Yürüyüşler
$\hat{\tau}$	τ_w
44.05023	45.71201

Çizelge 3.13. $\lambda_1 = 5, \lambda_2 = 1, v = 15, n = 30$ ve $T = 21.78410$ için tahminler

$\tau = 3.74306, \frac{\tau}{T} = 0.17183$	
Eşitlik (3.5a)'ya Göre	Rasgele Yürüyüşler
$\hat{\tau}$	τ_w
3.61887	3.29616

Çizelge 3.14. $\lambda_1 = 5, \lambda_2 = 1, v = 20, n = 30$ ve $T = 10.80383$ için tahminler

$\tau = 3.92683, \frac{\tau}{T} = 0.36347$	
Eşitlik (3.5a)'ya Göre	Rasgele Yürüyüşler
$\hat{\tau}$	τ_w
3.68288	4.26041

Çizelge 3.15. $\lambda_1 = 5, \lambda_2 = 1, v = 50, n = 100$ ve $T = 57.87960$ için tahminler

$\tau = 10.60068, \frac{\tau}{T} = 0.18315$	
Eşitlik (3.5a)'ya Göre	Rasgele Yürüyüşler
$\hat{\tau}$	τ_w
12.50718	10.75329

Çizelge 3.16. $\lambda_1 = 5, \lambda_2 = 1, v = 65, n = 100$ ve $T = 45.55770$ için tahminler

$\tau = 12.48924, \frac{\tau}{T} = 0.27414$	
Eşitlik (3.5a)'ya Göre	Rasgele Yürüyüşler
$\hat{\tau}$	τ_w
12.54445	13.86144

Yapılan çok sayıda uygulama sonunda $\frac{\tau}{T}$ oranının $\frac{1}{6}$ 'dan küçük ya da $\frac{5}{6}$ 'dan büyük olduğu durumlarda Eşitlik (3.5a)'dan elde edilen tahminlerin gerçek değerlerden uzaklaştıkları gözlenmiştir (Çizelge 3.17-Çizelge 3.18).

Çizelge 3.17. $\lambda_1 = 1.5, \lambda_2 = 0.5, v = 35, n = 100$ ve $T = 148.53118$ için tahminler

$\tau = 24.27752, \frac{\tau}{T} = 0.16365$	
Eşitlik (3.5a)'ya Göre	Rasgele Yürüyüşler
$\hat{\tau}$	τ_w
36.36155	26.08110

Çizelge 3.18. $\lambda_1 = 0.5$, $\lambda_2 = 1.5$, $\nu = 20$, $n = 30$ ve $T = 40.31612$ için tahminler

$\tau = 34.71044, \frac{\tau}{T} = 0.86096$	
Eşitlik (3.5a)'ya Göre	Rasgele Yürüyüşler
$\hat{\tau}$	τ_w
28.90576	34.71044

3.1.4.2. λ_1 'in bilindiği, λ_2 'nin bilinmediği durumda, τ ve λ_2 'nin tahminleri için örnekler

τ ve λ_2 için Kesim 3.1.2'de elde edilen tahmin edicilerin gerçek değerlere yakın sonuç verip vermediğini göstermek için $n = 30$, 100 örneklem büyüklüklerine ve çeşitli λ_1 , λ_2 ($\lambda_1 \neq \lambda_2$) değerlerine göre üstel dağılımdan rasgele örneklemeler çekilmiş; bu örneklemeler için, τ ve λ_2 'nin Eşitlik (3.10) ve Eşitlik (3.11)'den elde edilen tahminleri Çizelge 3.19-Çizelge 3.34'te verilmiştir.

Çizelge 3.19. $\lambda_1=0.5$, $\nu = 10$, $n = 30$ ve $T = 38.76919$ için tahminler

$\tau = 26.58168, \lambda_2=1.5, \frac{\tau}{T} = 0.68564$	
$\hat{\tau}$	$\hat{\lambda}_2$
25.33077	1.28993

Çizelge 3.20. $\lambda_1 = 0.5$ $\nu = 15$, $n = 30$ ve $T = 38.79692$ için tahminler

$\tau = 28.91163, \lambda_2 = 1.5, \frac{\tau}{T} = 0.74520$	
$\hat{\tau}$	$\hat{\lambda}_2$
25.33887	1.41247

Çizelge 3.21. $\lambda_1 = 0.5$, $\nu = 35$, $n = 100$ ve $T = 113.15480$ için tahminler

$\tau = 74.15410, \lambda_2 = 1.5, \frac{\tau}{T} = 0.65533$	
$\hat{\tau}$	$\hat{\lambda}_2$
79.34701	1.78439

Çizelge 3.22. $\lambda_1 = 0.5$, $\nu = 50$, $n = 100$ ve $T = 114.94556$ için tahminler

$\tau = 80.20288, \lambda_2 = 1.5, \frac{\tau}{T} = 0.69775$	
$\hat{\tau}$	$\hat{\lambda}_2$
79.95562	1.71541

Çizelge 3.23. $\lambda_1 = 1.5$, $\nu = 15$, $n = 30$ ve $T = 43.47320$ için tahminler

$\tau = 12.94673, \lambda_2 = 0.5, \frac{\tau}{T} = 0.29781$	
$\hat{\tau}$	$\hat{\lambda}_2$
12.25070	0.37229

Çizelge 3.24. $\lambda_1 = 1.5$, $\nu = 20$, $n = 30$ ve $T = 39.31683$ için tahminler

$\tau = 11.44214, \lambda_2 = 0.5, \frac{\tau}{T} = 0.29102$	
$\hat{\tau}$	$\hat{\lambda}_2$
11.99664	0.43942

Çizelge 3.25. $\lambda_1 = 1.5$, $\nu = 50$, $n = 100$ ve $T = 153.80751$ için tahminler

$\tau = 35.71465, \lambda_2 = 0.5, \frac{\tau}{T} = 0.23220$	
$\hat{\tau}$	$\hat{\lambda}_2$
41.32133	0.35762

Çizelge 3.26. $\lambda_1 = 1.5$, $\nu = 65$, $n = 100$ ve $T = 119.28360$ için tahminler

$\tau = 43.34456, \lambda_2 = 0.5, \frac{\tau}{T} = 0.36337$	
$\hat{\tau}$	$\hat{\lambda}_2$
39.16325	0.51491

Çizelge 3.27. $\lambda_1 = 1$, $\nu = 10$, $n = 30$ ve $T = 15.70841$ için tahminler

$\tau = 12.37520, \lambda_2 = 5, \frac{\tau}{T} = 0.78810$	
$\hat{\tau}$	$\hat{\lambda}_2$
11.44813	4.35460

Çizelge 3.28. $\lambda_1 = 1$, $\nu = 15$, $n = 30$ ve $T = 14.39305$ için tahminler

$\tau = 11.63100, \lambda_2 = 5, \frac{\tau}{T} = 0.80810$	
$\hat{\tau}$	$\hat{\lambda}_2$
10.92906	5.50549

Çizelge 3.29. $\lambda_1 = 1$, $\nu = 35$, $n = 100$ ve $T = 48.52171$ için tahminler

$\tau = 37.94226, \lambda_2 = 5, \frac{\tau}{T} = 0.78196$	
$\hat{\tau}$	$\hat{\lambda}_2$
36.65474	5.33796

Çizelge 3.30. $\lambda_1 = 1$, $\nu = 50$, $n = 100$ ve $T = 56.57775$ için tahminler

$\tau = 45.71201, \lambda_2 = 5, \frac{\tau}{T} = 0.80795$	
$\hat{\tau}$	$\hat{\lambda}_2$
39.67362	3.56874

Çizelge 3.31. $\lambda_1 = 5$, $\nu = 15$, $n = 30$ ve $T = 21.78410$ için tahminler

$\tau = 3.74306$, $\lambda_2 = 1$, $\frac{\tau}{T} = 0.17183$	
$\hat{\tau}$	$\hat{\lambda}_2$
4.01759	0.55790

Çizelge 3.32. $\lambda_1 = 5$, $\nu = 20$, $n = 30$ ve $T = 10.80383$ için tahminler

$\tau = 3.92683$, $\lambda_2 = 1$, $\frac{\tau}{T} = 0.36347$	
$\hat{\tau}$	$\hat{\lambda}_2$
3.52978	1.69796

Çizelge 3.33. $\lambda_1 = 5$, $\nu = 50$, $n = 100$ ve $T = 57.87960$ için tahminler

$\tau = 10.60068$, $\lambda_2 = 1$, $\frac{\tau}{T} = 0.18315$	
$\hat{\tau}$	$\hat{\lambda}_2$
12.91822	0.78754

Çizelge 3.34. $\lambda_1 = 5$, $\nu = 65$, $n = 100$ ve $T = 45.55770$ için tahminler

$\tau = 12.48924$, $\lambda_2 = 1$, $\frac{\tau}{T} = 0.27414$	
$\hat{\tau}$	$\hat{\lambda}_2$
12.36531	1.15004

Yapılan çok sayıda uygulama sonunda $\frac{\tau}{T}$ oranının $\frac{1}{6}$ 'dan küçük ya da $\frac{5}{6}$ 'dan büyük olduğu durumlarda Eşitlik (3.9) ve Eşitlik (3.10)'dan elde edilen tahminlerin gerçek değerlerinden uzaklaştıkları gözlenmektedir (Çizelge 3.35-Çizelge 3.36).

Çizelge 3.35. $\lambda_1 = 1.5$, $\nu = 35$, $n = 100$ ve $T = 148.53118$ için tahminler

$\tau = 24.27752, \lambda_2 = 0.5, \frac{\tau}{T} = 0.16365$	
$\hat{\tau}$	$\hat{\lambda}_2$
41.03829	0.35762

Çizelge 3.36. $\lambda_1 = 0.5$, $\nu = 20$, $n = 30$ ve $T = 40.31612$ için tahminler

$\tau = 34.71044, \lambda_2 = 1.5, \frac{\tau}{T} = 0.86096$	
$\hat{\tau}$	$\hat{\lambda}_2$
25.77205	1.17669

3.1.4.3. λ_1 ve λ_2 'nin bilinmediği durumda τ ve λ_1, λ_2 'nin tahminleri için örnekler

Çeşitli λ_1, λ_2 ($\lambda_1 \neq \lambda_2$) değerlerine ve örneklem büyüklükleri 30, 100'e göre üstel dağılımdan çekilen rasgele örneklem için, τ ve λ_1, λ_2 'nin Eşitlik (3.12)'den elde edilen tahminleri Çizelge 3.37-Çizelge 3.52'de verilmiştir.

Çizelge 3.37. $\nu = 10$, $n = 30$ ve $T = 38.76919$ için tahminler

$\tau = 26.58168, \lambda_1=0.5, \lambda_2=1.5, \frac{\tau}{T} = 0.68564$		
$\hat{\tau}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
26.58168	0.37620	1.64020

Çizelge 3.38. $\nu = 15$, $n = 30$ ve $T = 38.79692$ için tahminler

$\tau = 28.91163, \lambda_1 = 0.5, \lambda_2 = 1.5, \frac{\tau}{T} = 0.74520$		
$\hat{\tau}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
28.91163	0.51882	1.51741

Çizelge 3.39. $\nu = 35$, $n = 100$ ve $T = 113.15480$ için tahminler

$\tau = 74.15410, \lambda_1 = 0.5, \lambda_2 = 1.5, \frac{\tau}{T} = 0.65533$		
$\hat{\tau}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
74.15410	0.47199	1.66664

Çizelge 3.40. $\nu = 50$, $n = 100$ ve $T = 114.94556$ için tahminler

$\tau = 80.20288, \lambda_1 = 0.5, \lambda_2 = 1.5, \frac{\tau}{T} = 0.69775$		
$\hat{\tau}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
71.75337	0.54353	1.41229

Çizelge 3.41. $\nu = 15$, $n = 30$ ve $T = 43.47320$ için tahminler

$\tau = 12.94673, \lambda_1 = 1.5, \lambda_2 = 0.5, \frac{\tau}{T} = 0.29781$		
$\hat{\tau}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
12.94673	1.15859	0.49138

Çizelge 3.42. $\nu = 20$, $n = 30$ ve $T = 39.31683$ için tahminler

$\tau = 11.44214, \lambda_1 = 1.5, \lambda_2 = 0.5, \frac{\tau}{T} = 0.29102$		
$\hat{\tau}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
11.44214	1.74729	0.35875

Çizelge 3.43. $\nu = 50$, $n = 100$ ve $T = 153.80751$ için tahminler

$\tau = 35.71465, \lambda_1 = 1.5, \lambda_2 = 0.5, \frac{\tau}{T} = 0.23220$		
$\hat{\tau}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
36.60768	1.39315	0.41809

Çizelge 3.44. $v = 65$, $n = 100$ ve $T = 119.28360$ için tahminler

$\tau = 43.34456, \lambda_1 = 1.5, \lambda_2 = 0.5, \frac{\tau}{T} = 0.36337$		
$\hat{\tau}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
44.44490	1.43498	0.45431

Çizelge 3.45. $v = 10$, $n = 30$ ve $T = 15.70841$ için tahminler

$\tau = 12.37520, \lambda_1 = 1, \lambda_2 = 5, \frac{\tau}{T} = 0.78810$		
$\hat{\tau}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
12.37520	0.80807	6.00023

Çizelge 3.46. $v = 15$, $n = 30$ ve $T = 14.39305$ için tahminler

$\tau = 11.63100, \lambda_1 = 1, \lambda_2 = 5, \frac{\tau}{T} = 0.80810$		
$\hat{\tau}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
11.63100	1.28966	5.43075

Çizelge 3.47. $v = 35$, $n = 100$ ve $T = 48.52171$ için tahminler

$\tau = 37.94226, \lambda_1 = 1, \lambda_2 = 5, \frac{\tau}{T} = 0.78196$		
$\hat{\tau}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
37.94226	0.92245	6.14399

Çizelge 3.48. $v = 50$, $n = 100$ ve $T = 56.57775$ için tahminler

$\tau = 45.71201, \lambda_1 = 1, \lambda_2 = 5, \frac{\tau}{T} = 0.80795$		
$\hat{\tau}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
45.71201	1.09380	4.60162

Çizelge 3.49. $v = 15$, $n = 30$ ve $T = 21.78410$ için tahminler

$\tau = 3.74306, \lambda_1 = 5, \lambda_2 = 1, \frac{\tau}{T} = 0.17183$		
$\hat{\tau}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
3.74306	4.00741	0.83144

Çizelge 3.50. $v = 20$, $n = 30$ ve $T = 10.80383$ için tahminler

$\tau = 3.92683, \lambda_1 = 5, \lambda_2 = 1, \frac{\tau}{T} = 0.36347$		
$\hat{\tau}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
4.26041	4.92911	1.37543

Çizelge 3.51. $v = 50$, $n = 100$ ve $T = 57.87960$ için tahminler

$\tau = 10.60068, \lambda_1 = 5, \lambda_2 = 1, \frac{\tau}{T} = 0.18315$		
$\hat{\tau}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
10.75329	4.74274	1.03976

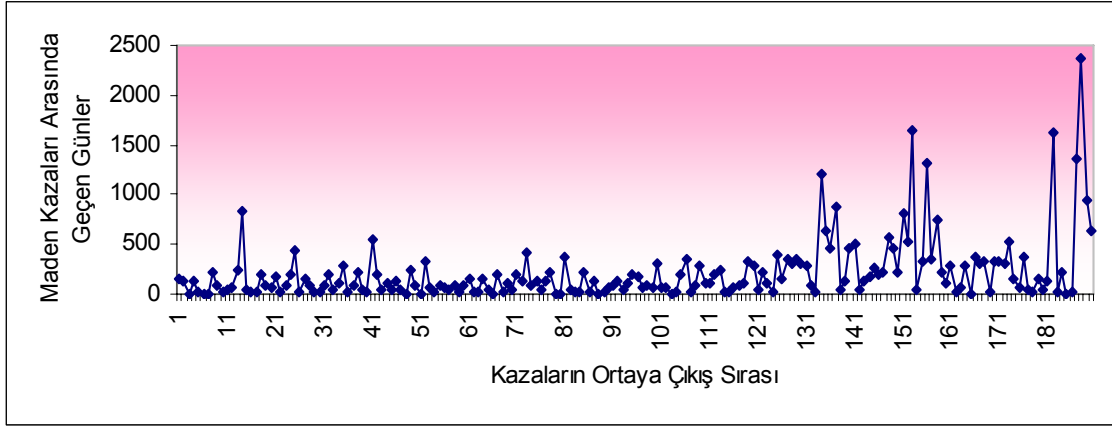
Çizelge 3.52. $v = 65$, $n = 100$ ve $T = 45.55770$ için tahminler

$\tau = 12.48924, \lambda_1 = 5, \lambda_2 = 1, \frac{\tau}{T} = 0.27414$		
$\hat{\tau}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
13.86144	4.97784	0.97803

3.1.4.4. Kömür madeni kazaları veri kümesine ilişkin tahminler

Kesim 2.4'te verilen İngiltere'deki kömür madeni kazalarına ilişkin veri kümesi, değişme noktası sorunu üzerinde çalışan istatistikçilerin çoğunun kullandığı veri kümesidir. Poisson süreçlerinde değişme noktasına ilişkin önceki çalışmalarda elde edilen tahmin ediciler Kesim 2.4'te belirtildiği gibi bazı esnek varsayımlara dayanmaktadır. Bu durum, tez çalışmasında elde edilen tahmin edicilerin önceki çalışmalarda elde edilen tahmin edicilerle karşılaştırılması için önemli bir engeldir. Bu nedenle tez çalışmasında bu veri kümesi için elde edilen tahminler önceki

çalışmalarda elde edilen tahminlerle karşılaştırılmıştır. İngiltere'deki kömür madeni kazaları arasında geçen günlerin, kazaların ortaya çıkış sırasına göre saçılımı Şekil 3.1'de verildiği gibidir. Şekil 3.1'den görüldüğü gibi, kazalar arasında geçen günlerin sayısı bir zaman noktasından sonra artmaktadır. Önceki çalışmalarda bu veri kümesi için $\lambda_1 = 0.0085$ (gün) ($\lambda_1 = 3.1025$ (yıl)) ve $\lambda_2=0.0025$ (gün) ($\lambda_2=0.9125$ (yıl)) olarak bulunmuştur. Bu veri kümesi kullanılarak Kesim 3.1.1, Kesim 3.1.2 ve Kesim 3.1.3'te elde edilen tahmin edicilerden, λ_1 ve λ_2 bilindiği durumda, τ 'nun tahmini Çizelge 3.53'te , yalnızca λ_1 bilindiği durumda, τ ve λ_2 'nin tahminleri Çizelge 3.54'te; λ_1 ve λ_2 bilinmediği durumda τ , λ_1 ve λ_2 'nin tahminleri Çizelge 3.55'te verilmiştir.



Şekil 3.1 Maden kazaları arasında geçen günler

Çizelge 3.53. $\lambda_1 = 0.0085$ (gün), $\lambda_2 = 0.0025$ (gün), $n = 190$ ve $T = 40549$ (gün) için tahmin

Tahmin	
$\hat{\tau}$	13560.40574 (gün), 18 Nisan 1888

Çizelge 3.54. $\lambda_1 = 0.0085$ (gün), $n = 190$ ve $T = 40549$ (gün) için tahminler

Tahmin	
$\hat{\tau}$	13173.05981, 28 Mayıs 1887
$\hat{\lambda}_2$	0.0028 (gün), 1.022 (yıl)

Çizelge 3.55. $n = 190$ ve $T = 40549$ (gün) için tahminler

Tahmin	
$\hat{\tau}$	14240 (gün), 11 Mart 1890
$\hat{\lambda}_1$	0.008708 (gün), 3.18 (yıl)
$\hat{\lambda}_2$	0.002509 (gün), 0.92 (yıl)

Kesim 2.2.1'de verilen Çizelge 2.2'deki tahminlerle Çizelge 3.53 - Çizelge 3.55'te verilen tahminlerin uyumlu oldukları görülmektedir. Buna göre, uygulaması güç olan tahmin edicileri kullanmak yerine Kesim 3.1'de elde ettiğimiz ve uygulama kolaylığı sağlayan tahmin edicileri kullanmak çok daha uygun olacaktır. Ayrıca λ_1 ve λ_2 bilindiği durumda, τ 'nın Eşitlik (3.5a)'da verilen tahmin edicisinin; λ_1 bilindiği durumda, τ ve λ_2 'nin Eşitlik (3.10), Eşitlik (3.11)'de verilen tahmin edicilerinin analitik yapıda olmaları, tahmin edicilerin özelliklerinin araştırılmasına da olanak sağlayacaktır.

3.2. Poisson Sürecinde Değişme Noktasına İlişkin Test İstatistiği

Bu kesimde homojen Poisson sürecinde değişmenin olup olmadığına ilişkin hipotezi test etmek için test istatistiğinin nasıl elde edildiği açıklanacak ve uygulamalar verilecektir.

3.2.1. Olabilirlik Oranına Dayalı Test İstatistiği

Bu çalışmada, tahmin ediciler olabilirlik fonksiyonundan elde edildikleri için, test istatistiği de olabilirlik oranına dayalı olarak elde edilmiştir.

Homojen Poisson sürecinde sürecin parametresi olan λ 'nın, zaman içinde değişip değişmediğine ilişkin H_0 yokluk hipotezi ve H_1 seçenek hipotezi aşağıdaki gibidir:

$$\begin{aligned} H_0 : \lambda = \lambda_1, \quad 0 < t \leq T & \quad H_0: (0, T] \text{ aralığında } \lambda \text{ da değişme yoktur.} \\ H_1 : \lambda = \lambda_1, \quad 0 < t \leq \tau & \quad \text{ya da} \quad H_1: \lambda, \tau \text{ dan sonra değişir.} \\ \quad : \lambda = \lambda_2, \quad \tau < t \leq T & \end{aligned}$$

Yukarıda verilen hipotezlerin testi için, $(0, T]$ zaman aralığında n olayın gözlemlendiği ve $S_v = \tau$, $v = 1, 2, \dots, n-1$, $S_n = T$ olduğu varsayımı altında olabilirlik oranı,

$$\Lambda = \frac{\lambda_1^n e^{-\lambda_1 T}}{\lambda_1^{N_\tau} e^{-\lambda_1 \tau} \lambda_2^{n-N_\tau} e^{-\lambda_2 (T-\tau)}} \quad (3.13)$$

biçiminde yazılabilir. Bu oranda paydaki fonksiyon, H_0 hipotezi altında; paydadaki fonksiyon, H_1 hipotezi altında olabilirlik fonksiyonlarıdır. Eşitlik (3.13)'ün her iki yanının logaritması alınırsa,

$$\log \Lambda = (n - N_\tau) \log \left(\frac{\lambda_1}{\lambda_2} \right) - (\lambda_1 - \lambda_2)(T - \tau)$$

bulunur. H_0 doğru olduğu varsayımı altında, N_τ 'nin beklenen değeri $\lambda_1 \tau$ olduğuna göre,

$$E[\log \Lambda] = \tau \left[(\lambda_1 - \lambda_2) - \lambda_1 \log \left(\frac{\lambda_1}{\lambda_2} \right) \right] + n \log \left(\frac{\lambda_1}{\lambda_2} \right) - (\lambda_1 - \lambda_2)T$$

elde edilir. $X = E[\log \Lambda]$ alınırsa,

$$X = aT + b \quad (3.14)$$

yazılabilir. Burada,

$$a = -(\lambda_1 - \lambda_2), \quad b = \tau \left[(\lambda_1 - \lambda_2) - \lambda_1 \log \left(\frac{\lambda_1}{\lambda_2} \right) \right] + n \log \left(\frac{\lambda_1}{\lambda_2} \right)$$

biçimindedir ve X , olabilirlik oranı test istatistiğidir. Bu istatistiğin H_0 doğru olduğu varsayımı altında T 'nin olasılık yoğunluk fonksiyonunun n ve λ_1 parametreleri ile,

$$f_T(t) = \frac{\lambda_1^n}{\Gamma(n)} e^{-\lambda_1 t} t^{n-1}, \quad t > 0$$

$$= 0, \quad t \leq 0$$

gamma olasılık yoğunluk fonksiyonu olduğu düşünülürse ve Eşitlik (3.14)'e göre,

$t = \frac{x-b}{a}$ alınırsa, $\lambda_1 < \lambda_2$ için,

$$f_X(x) = \frac{\lambda_1}{\Gamma(n)} \left(\frac{x-b}{a} \right)^{n-1} e^{-\lambda_1 \left(\frac{x-b}{a} \right)} \left(\frac{1}{a} \right), \quad b < x < \infty \quad (3.15)$$

=0, öteki değerler

ve $\lambda_1 > \lambda_2$ için,

$$f_X(x) = \frac{\lambda_1}{\Gamma(n)} \left(\frac{x-b}{a} \right)^{n-1} e^{-\lambda_1 \left(\frac{x-b}{a} \right)} \left(-\frac{1}{a} \right), \quad -\infty < x < b \quad (3.16)$$

=0, öteki değerler

elde edilir. Burada $\lambda_1 < \lambda_2$ için $a > 0$, $\lambda_1 > \lambda_2$ için $a < 0$ olduğu düşünülmüştür.

Olabilirlik oranı test istatistiği üç durum için ele alınmıştır: λ_1 ve λ_2 'nin bilindiği durum; λ_1 'in bilinip, λ_2 'nin bilinmediği durum; λ_1 ve λ_2 'nin bilinmediği durum.

Birinci durumda, τ yerine; ikinci durumda, τ ve λ_2 yerine; üçüncü durumda, τ ile λ_1 ve λ_2 'nin yerine en çok olabilirlik tahminleri alınır. Buna göre, Eşitlik (3.15) ve Eşitlik (3.16)'daki olasılık yoğunluk fonksiyonlardaki a ve b değerleri, λ_1 , λ_2 'nin her ikisi de bilindiğinde,

$$a = -(\lambda_1 - \lambda_2), \quad b = \hat{\tau} \left[(\lambda_1 - \lambda_2) - \lambda_1 \log \left(\frac{\lambda_1}{\lambda_2} \right) \right] + n \log \left(\frac{\lambda_1}{\lambda_2} \right) \quad (3.17)$$

ve λ_1 bilinip, λ_2 bilinmediğinde,

$$a = -(\lambda_1 - \hat{\lambda}_2), \quad b = \hat{\tau} \left[(\lambda_1 - \hat{\lambda}_2) - \lambda_1 \log \left(\frac{\lambda_1}{\hat{\lambda}_2} \right) \right] + n \log \left(\frac{\lambda_1}{\hat{\lambda}_2} \right) \quad (3.18)$$

ve λ_1 , λ_2 'nin her ikisi de bilinmediğinde,

$$a = -(\hat{\lambda}_1 - \hat{\lambda}_2), \quad b = \hat{\tau} \left[(\hat{\lambda}_1 - \hat{\lambda}_2) - \lambda_1 \log \left(\frac{\hat{\lambda}_1}{\hat{\lambda}_2} \right) \right] + n \log \left(\frac{\hat{\lambda}_1}{\hat{\lambda}_2} \right) \quad (3.19)$$

biçiminde olur.

H_0 hipotezinin reddedildiği bölgenin sınırları (kritik değerler),

$$P(X < C_0) = P(X > C_1) = \alpha / 2$$

eşitliği yardımıyla bulunur. Bu kritik değerler, örneklemden elde edilen $X_{\text{Hesap}} = aT + b$ değeri ile karşılaştırılır. Eğer,

$$X_{\text{Hesap}} < C_0 \text{ ya da } X_{\text{Hesap}} > C_1$$

ise, H_0 hipotezi reddedilir ve H_1 hipotezi kabul edilir.

3.2.2. Uygulama

Bu kesimde Kesim 3.2.1'de verilen test istatistiğinin geçerliliğini göstermek için, Kesim 3.1.4'te verilen örneklerden yararlanılmıştır.

Çizelge 3.1'de verilen örnek için hipotezler,

$$\begin{aligned} H_0 : \lambda = \lambda_1 = 0.5, & \quad 0 < t \leq 38.76919 \\ H_1 : \lambda = \lambda_1 = 0.5, & \quad 0 < t \leq \tau \\ & \quad \lambda = \lambda_2 = 1.5, \quad \tau < t \leq 38.76919 \end{aligned} \quad (3.20)$$

biçimindedir. $\lambda_1 < \lambda_2$ olduğu için, Eşitlik (3.15)'te bulunan olabilirlik oranı test istatistiğinin olasılık yoğunluk fonksiyonundaki a ve b değişmezleri, Eşitlik (3.17)'den sırasıyla 1 ve -45.21586 elde edilmiş; bu değerler kullanılarak X raslantı değişkeninin tanım aralığı $-45.21586 < x < \infty$ bulunmuştur. Buna göre, $\alpha=0.05$ yanılma düzeyinde kritik bölgenin sınırları C_0 , C_1 ve X_{Hesap} , Çizelge 3.56'da verildiği gibi elde edilmiştir.

Çizelge 3.56. Eşitlik (3.20)'de verilen hipotezlerin testi

C_0	C_1	X_{Hesap}
- 4.73410	38.08178	-6.44667

$X_{\text{Hesap}} < - 4.73410$ olduğu için, H_0 hipotezi reddedilir.

Çizelge 3.8'de verilen örnek için hipotezler,

$$\begin{aligned}
 H_0 : \lambda = \lambda_1 = 1.5, & \quad 0 < t \leq 119.28360 \\
 H_1 : \lambda = \lambda_1 = 1.5, & \quad 0 < t \leq \tau \\
 & \quad \lambda = \lambda_2 = 0.5, \quad \tau < t \leq 119.28360
 \end{aligned}
 \tag{3.21}$$

biçimindedir. $\lambda_1 > \lambda_2$ olduğu için, Eşitlik (3.16)'da bulunan olabilirlik oranı test istatistiğinin olasılık yoğunluk fonksiyonundaki a ve b değişmezleri, Eşitlik (3.17)'den sırasıyla -1 ve 84.27046 elde edilmiş; bu değerler kullanılarak X raslantı değişkeninin tanım aralığı $-\infty < x < 84.27046$ bulunmuştur. Buna göre, $\alpha=0.05$ yanılma düzeyinde kritik bölgenin sınırları C_0 , C_1 ve X_{Hesap} , Çizelge 3.57'de verildiği gibi elde edilmiştir.

Çizelge 3.57. Eşitlik (3.21)'de verilen hipotezlerin testi

C_0	C_1	X_{Hesap}
3.91779	30.02781	-35.01314

$X_{\text{Hesap}} < 3.91779$ olduğu için, H_0 hipotezi reddedilir.

Çizelge 3.27'de verilen örnek için hipotezler,

$$\begin{aligned}
 H_0 : \lambda = \lambda_1 = 1, & \quad 0 < t \leq 15.70841 \\
 H_1 : \lambda = \lambda_1 = 1, & \quad 0 < t \leq \tau \\
 & \quad : \lambda = \lambda_2, \quad \tau < t \leq 15.70841
 \end{aligned}$$

ya da

$$\begin{aligned}
 H_0 : \lambda = \lambda_1 = 1, & \quad 0 < t \leq 15.70841 \\
 H_1 : \lambda, & \tau \text{ 'dan sonra değişir.}
 \end{aligned}
 \tag{3.22}$$

biçimindedir. $\lambda_1 < \hat{\lambda}_2$ olduğu için, Eşitlik (3.15)'te bulunan olabilirlik oranı test istatistiğinin olasılık yoğunluk fonksiyonundaki a ve b değişmezleri, Eşitlik (3.18)'den sırasıyla 3.35460 ve -65.69802 elde edilmiş; bu değerler kullanılarak X raslantı değişkeninin tanım aralığı $-65.69802 < x < \infty$ bulunmuştur. Buna göre,

$\alpha=0.05$ yanılma düzeyinde kritik bölgenin sınırları C_0 , C_1 ve X_{Hesap} , Çizelge 3.58'de verildiği gibi elde edilmiştir.

Çizelge 3.58. Eşitlik (3.22)'de verilen hipotezlerin testi

C_0	C_1	X_{Hesap}
2.20203	74.01711	-13.00258

$X_{\text{Hesap}} < 2.20203$ olduğu için, H_0 hipotezi reddedilir.

Çizelge 3.34'te verilen örnek için hipotezler,

$$\begin{aligned} H_0 : \lambda = \lambda_1 = 5, & \quad 0 < t \leq 45.55770 \\ H_1 : \lambda = \lambda_1 = 5, & \quad 0 < t \leq \tau \\ & \quad \lambda = \lambda_2, \quad \tau < t \leq 45.55770 \end{aligned}$$

ya da

$$H_0 : \lambda = \lambda_1 = 5, \quad 0 < t \leq 45.55770$$

$$H_1: \lambda, \tau \text{ dan sonra değişir.} \quad (3.23)$$

biçimindedir. $\lambda_1 > \hat{\lambda}_2$ olduğu için, Eşitlik (3.16)'da bulunan olabilirlik oranı test istatistiğinin olasılık yoğunluk fonksiyonundaki a ve b değişmezleri, Eşitlik (3.18)'den sırasıyla -3.84996 ve 103.70722 elde edilmiş; bu değerler kullanılarak X raslantı değişkeninin tanım aralığı $-\infty < x < 103.70722$ bulunmuştur. Buna göre, $\alpha = 0.05$ yanılma düzeyinde kritik bölgenin sınırları C_0 , C_1 ve X_{Hesap} , Çizelge 3.59'da verildiği gibi elde edilmiştir.

Çizelge 3.59. Eşitlik (3.23)'te verilen hipotezlerin testi

C_0	C_1	X_{Hesap}
10.90086	41.05761	-71.68810

$X_{\text{Hesap}} < 10.90086$ olduğu için, H_0 hipotezi reddedilir.

Çizelge 3.39'da verilen örnek için hipotezler,

$$\begin{aligned}
H_0 : \lambda = \lambda_1, & \quad 0 < t \leq 113.15480 \\
H_1 : \lambda = \lambda_1, & \quad 0 < t \leq \tau \\
& \quad \lambda = \lambda_2, \quad \tau < t \leq 113.15480
\end{aligned}$$

ya da

H_0 : (0, 113.15480] aralığında λ 'da deęişme yoktur.

H_1 : λ , τ 'dan sonra deęişir. (3.24)

biçimindedir. $\hat{\lambda}_1 < \hat{\lambda}_2$ olduęu için, Eşitlik (3.15)'te bulunan olabilirlik oranı test istatistięinin olasılık yoğunluk fonksiyonundaki a ve b deęişmezleri, Eşitlik (3.19)'dan sırasıyla 1.19465 ve -170.59267 elde edilmiş; bu deęerler kullanılarak X raslantı deęişkeninin tanım aralığı $-170.59267 < x < \infty$ bulunmuştur. Buna göre, $\alpha = 0.05$ yanılma düzeyinde kritik bölgenin sınırları C_0 , C_1 ve X_{Hesap} , Çizelge 3.60'da verildięi gibi elde edilmiştir.

Çizelge 3.60. Eşitlik (3.24)'te verilen hipotezlerin testi

C_0	C_1	X_{Hesap}
35.34703	134.47730	-35.41228

$X_{\text{Hesap}} < 35.34703$ olduęu için, H_0 hipotezi reddedilir.

Çizelge 3.42'de verilen örnek için hipotezler,

$$\begin{aligned}
H_0 : \lambda = \lambda_1, & \quad 0 < t \leq 39.31683 \\
H_1 : \lambda = \lambda_1, & \quad 0 < t \leq \tau \\
& \quad \lambda = \lambda_2, \quad \tau < t \leq 39.31683
\end{aligned}$$

ya da

H_0 : (0, 39.31683] aralığında λ 'da deęişme yoktur.

H_1 : λ , τ 'dan sonra deęişir. (3.25)

biçimindedir. $\hat{\lambda}_1 > \hat{\lambda}_2$ olduęu için, Eşitlik (3.16)'da bulunan olabilirlik oranı test istatistięinin olasılık yoğunluk fonksiyonundaki a ve b deęişmezleri, Eşitlik (3.19)'dan sırasıyla -1.38854 ve 31.73132 elde edilmiş; bu deęerler kullanılarak X raslantı deęişkeninin tanım aralığı $-\infty < x < 31.73132$ bulunmuştur. Buna göre,

$\alpha=0.05$ yanılma düzeyinde kritik bölgenin sınırları C_0 , C_1 ve X_{Hesap} , Çizelge 3.61'de verildiği gibi elde edilmiştir.

Çizelge 3.61. Eşitlik (3.25)'te verilen hipotezlerin testi

C_0	C_1	X_{Hesap}
-1.36624	15.64626	-22.86167

$X_{\text{Hesap}} < -1.36624$ olduğu için, H_0 hipotezi reddedilir.

Verilen hipotezlerin testinde H_0 hipotezinin reddedilmesi λ_1 ve λ_2 arasında farkın büyüklüğü ile doğru orantılıdır. Bu nedenle λ_1 ve λ_2 arasındaki farkın küçük olması durumunda test fonksiyonunun davranışı göstermek için, iki örneklem çekilmiş; bu örneklem için sonuçlar verilmiştir.

$n = 30$ örneklem büyüklüğünde $\lambda_1 = 0.1$ ve $\lambda_2 = 0.2$ 'ye göre üstel dağılımdan çekilen rasgele örnekleme ilişkin tahminler Eşitlik (3.12)'den elde edilmiş ve Çizelge 3.62'de verilmiştir.

Çizelge 3.62. $\lambda_1=0.1$, $\lambda_2=0.2$, $v = 15$, $n = 30$ ve $T = 209.97127$ için tahminler

$\tau = 133.69140, \frac{\tau}{T} = 0.63671$		
\hat{t}	$\hat{\lambda}_1$	$\hat{\lambda}_2$
128.62770	0.10107	0.20899

Çizelge 3.62'ye göre,

$$\begin{aligned}
 H_0 : \lambda = \lambda_1, & \quad 0 < t \leq 209.97127 \\
 H_1 : \lambda = \lambda_1, & \quad 0 < t \leq \tau \\
 & \quad \lambda = \lambda_2, \quad \tau < t \leq 209.97127
 \end{aligned}$$

ya da

$H_0: (0, 209.97127]$ aralığında λ 'da değişme yoktur.

$H_1: \lambda, \tau$ 'dan sonra değişir. (3.26)

hipotezlerini test etmek isteyelim. Çizelge 3.62'de verilen tahminlere göre, $\hat{\lambda}_1 < \hat{\lambda}_2$ olduğu için, Eşitlik (3.15)'te bulunan olabilirlik oranı test istatistiğinin olasılık yoğunluk fonksiyonundaki a ve b değişmezleri, Eşitlik (3.19)'dan sırasıyla 0.10792 ve -26.23125 elde edilmiş; bu değerler kullanılarak X raslantı değişkeninin tanım aralığı $-26.23125 < x < \infty$ bulunmuştur. Buna göre, $\alpha = 0.05$ yanılma düzeyinde kritik bölgenin sınırları C_0 , C_1 ve X_{Hesap} , Çizelge 3.63'te verildiği gibi elde edilmiştir.

Çizelge 3.63. Eşitlik (3.26)'da verilen hipotezlerin testi

C_0	C_1	X_{Hesap}
-4.61855	18.24031	-3.57115

$X_{\text{Hesap}} > -4.61855$ olduğu için H_0 hipotezi kabul edilir.

$n = 100$ örneklem büyüklüğünde $\lambda_1 = 5.5$ ve $\lambda_2 = 5$ 'e göre üstel dağılımdan çekilen rasgele örnekleme ilişkin tahminler Eşitlik (3.12)'den elde edilmiş ve Çizelge 3.62'de verilmiştir.

Çizelge 3.64. $\lambda_1=5.5$, $\lambda_2=5$, $v = 50$, $n = 100$ ve $T = 20.11312$ için tahminler

$\tau = 9.12867, \frac{\tau}{T} = 0.453871$		
\hat{t}	$\hat{\lambda}_1$	$\hat{\lambda}_2$
9.69006	5.67592	4.31375

Çizelge 3.64'e göre,

$$\begin{aligned}
 H_0 : \lambda &= \lambda_1, & 0 < t \leq 20.11312 \\
 H_1 : \lambda &= \lambda_1, & 0 < t \leq \tau \\
 & \lambda = \lambda_2, & \tau < t \leq 20.11312
 \end{aligned}$$

ya da

H_0 : $(0, 20.11312]$ aralığında λ 'da değişme yoktur.

H_1 : λ , τ 'dan sonra değişir. (3.27)

hipotezlerini test etmek isteyelim. Çizelge 3.64'te verilen tahminlere göre, $\hat{\lambda}_1 > \hat{\lambda}_2$ olduğu için, Eşitlik (3.16)'da bulunan olabilirlik oranı test istatistiğinin olasılık yoğunluk fonksiyonundaki a ve b değişmezleri, Eşitlik (3.19)'dan sırasıyla -1.36217 ve 25.54864 elde edilmiş; değerler kullanılarak X raslantı değişkeninin tanım aralığı $-\infty < x < 25.54864$ bulunmuştur. Buna göre, $\alpha = 0.05$ yanılma düzeyinde kritik bölgenin sınırları C_0 , C_1 ve X_{Hesap} , Çizelge 3.65'te verildiği gibi elde edilmiştir.

Çizelge 3.65. Eşitlik (3.27)'de verilen hipotezlerin testi

C_0	C_1	X_{Hesap}
-3.37725	6.02201	-1.84885

$X_{\text{Hesap}} > -3.37725$ olduğu için H_0 hipotezi kabul edilir.

3.2.3. Kömür Madeni Kazaları Veri Kümesine İlişkin Hipotezlerin Testi

Kesim 2.4'te verilen kömür madeni kazası veri kümesi kullanılarak, λ_1 , λ_2 'nin her ikisinin de bilindiği; λ_1 'in bilinip, λ_2 'nin bilinmediği; λ_1 ve λ_2 'nin her ikisinin de bilinmediği üç durum için hipotezler tanımlanmış ve bu hipotezlerin testi için test istatistiğinin α düzeyinde kritik bölgenin sınırları elde edilmiştir (Çizelge 3.66, Çizelge 3.67, Çizelge 3.68). Bu hipotezlerin testinde Kesim 3.1.4.4'te elde edilen tahminler kullanılmıştır.

Kömür madeni kazaları veri kümesi için λ_1 , λ_2 'nin her ikisinin de bilindiği durumda hipotezler,

$$\begin{aligned}
 H_0 : \lambda = \lambda_1 = 0.0085, & \quad 0 < t \leq 40549 \\
 H_1 : \lambda = \lambda_1 = 0.0085, & \quad 0 < t \leq \tau \\
 & \quad \lambda = \lambda_2 = 0.0025, \quad \tau < t \leq 40549
 \end{aligned} \tag{3.28}$$

biçimindedir. Çizelge 3.53'teki τ 'nun tahmini kullanılarak, Eşitlik (3.16)'da bulunan olabilirlik oranı test istatistiğinin olasılık yoğunluk fonksiyonundaki a ve b değişmezleri, Eşitlik (3.17)'den sırasıyla -0.00600 ve 172.82319 elde edilmiş; bu değerler kullanılarak X raslantı değişkeninin tanım aralığı $-\infty < x < 172.82319$

bulunmuştur. Buna göre, $\alpha = 0.05$ yanılma düzeyinde kritik bölgenin sınırları C_0 , C_1 ve X_{Hesap} , Çizelge 3.66'da verildiği gibi elde edilmiştir.

Çizelge 3.66. Eşitlik (3.28)'de verilen hipotezlerin testi

C_0	C_1	X_{Hesap}
18.97572	57.09869	-70.47081

$X_{\text{Hesap}} < 18.97572$ olduğu için, H_0 hipotezi reddedilir.

λ_1 'in bilinip, λ_2 'nin bilinmediği durumda hipotezler,

$$\begin{aligned} H_0 : \lambda = \lambda_1 = 0.0085, & \quad 0 < t \leq 40549 \\ H_1 : \lambda = \lambda_1 = 0.0085, & \quad 0 < t \leq \tau \\ & \quad \lambda = \lambda_2, \quad \tau < t \leq 40549 \end{aligned}$$

ya da

$$H_0 : \lambda = \lambda_1 = 0.0085, \quad 0 < t \leq 40549$$

$$H_1 : \lambda, \tau \text{ 'dan sonra değişir.} \quad (3.29)$$

biçimindedir. Çizelge 3.54'teki τ ve λ_2 'nin tahminleri kullanılarak, Eşitlik (3.16)'da bulunan olabilirlik oranı test istatistiğinin olasılık yoğunluk fonksiyonundaki a ve b değişmezleri, Eşitlik (3.18)'den sırasıyla -0.00570 ve 161.73348 elde edilmiş; bu değerler kullanılarak X raslantı değişkeninin tanım aralığı $-\infty < x < 161.73348$ bulunmuştur. Buna göre, $\alpha = 0.05$ yanılma düzeyinde kritik bölgenin sınırları C_0 , C_1 ve X_{Hesap} , Çizelge 3.67'de verildiği gibi elde edilmiştir.

Çizelge 3.67. Eşitlik (3.29)'da verilen hipotezlerin testi

C_0	C_1	X_{Hesap}
15.57838	51.79521	-69.39582

$X_{\text{Hesap}} < 15.57838$ olduğu için, H_0 hipotezi reddedilir.

λ_1, λ_2 her ikisinin de bilinmediği durumda hipotezler,

$$\begin{aligned}
H_0 : \lambda = \lambda_1, & \quad 0 < t \leq 40549 \\
H_1 : \lambda = \lambda_1, & \quad 0 < t \leq \tau \\
& \quad : \lambda = \lambda_2, \quad \tau < t \leq 40549
\end{aligned}$$

ya da

H_0 : (0, 40549] aralığında λ 'da deęişme yoktur.

H_1 : λ , τ 'dan sonra deęişir. (3.30)

biçimindedir. Çizelge 3.55'teki τ ve λ_1 , λ_2 'nin tahminleri kullanılarak, Eşitlik (3.16)'da bulunan olabilirlik oranı test istatistiğinin olasılık yoğunluk fonksiyonundaki a ve b deęişmezleri, Eşitlik (3.19)'dan sırasıyla -0.00620 ve 170.39899 elde edilmiş; bu deęerler kullanılarak X raslantı deęişkeninin tanım aralığı $-\infty < x < 170.39899$ bulunmuştur. Buna göre, $\alpha = 0.05$ yanılma düzeyinde kritik bölgenin sınırları C_0 , C_1 ve X_{Hesap} , Çizelge 3.68'de verildiği gibi elde edilmiştir.

Çizelge 3.68. Eşitlik (3.30)'da verilen hipotezlerin testi

C_0	C_1	X_{Hesap}
15.24561	53.69218	-80.96426

$X_{\text{Hesap}} < 15.2456$ olduğu için, H_0 hipotezi reddedilir.

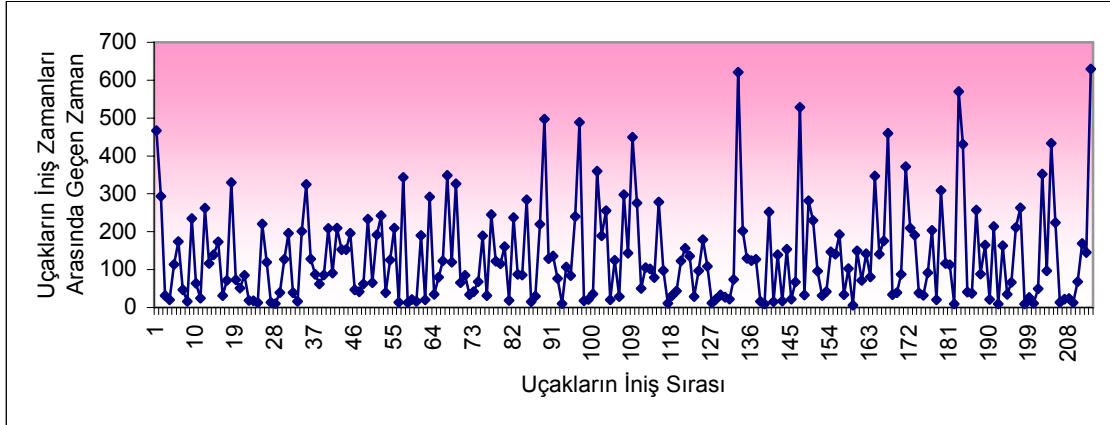
3.2.4. Hava Trafiği Veri Kümesine İlişkin Hipotez Testi

Literatürde deęişme noktasını incelemek için kullanılan gerçek veri kümelerinden biri de hava trafiği ile ilgilidir. Bu veri kümesi Hsu'nun 1979'taki ve Ramanayake'nin 2004'teki çalışmalarında gamma dağılımına sahip bağımsız raslantı deęişkenlerinin parametresinde deęişme olup olmadığına ilişkin hipotez testinde kullanılmıştır. Bu veriler Birleşik Havacılık İdaresi (Federal Aviation Administration) tarafından toplanmıştır. Veriler, 30 Nisan 1969 tarihinde 16:00-24:00 GMT(Greenwich Mean Time) saatleri arasında, New York bölgesindeki hava kontrol sektörlerinden elde edilmiştir. Gözlemin başlangıcından itibaren 213 uçağın inişi kaydedilmiştir. Çizelge 3.69'daki deęerler gözlemin başlangıcından itibaren saniye cinsinden iniş zamanlarını göstermektedir.

Çizelge 3.69. Uçakların iniş zamanları (yukarıdan aşağıya doğru) (Hsu, 1979)

467	2837	5202	7755	10624	13710	16515	19150	22095	25498	27971
761	2855	5355	7835	10639	14008	16537	19432	22182	25712	28116
792	2868	5551	7958	10669	14151	16570	19662	22554	25721	28746
812	3089	5598	8307	10889	14601	16597	19758	22764	25884	
926	3209	5640	8427	11386	14877	16619	19789	22955	25919	
1100	3223	5702	8754	11515	14927	16693	19831	22993	25985	
1147	3233	5935	8819	11651	15032	17314	19978	23025	26196	
1163	3272	6000	8904	11727	15134	17516	20119	23117	26459	
1398	3399	6192	8938	11737	15213	17646	20312	23321	26468	
1462	3595	6435	8980	11844	15491	17770	20346	23341	26494	
1487	3634	6474	9048	11928	15589	17897	20449	23650	26505	
1749	3650	6600	9237	12168	15600	17913	20455	23766	26554	
1865	3851	6810	9268	12657	15631	17922	20604	23879	26906	
2004	4176	6824	9513	12675	15674	18174	20675	23888	27003	
2177	4304	7168	9635	12696	15797	18189	20817	24458	27437	
2208	4391	7181	9750	12732	15953	18328	20898	24889	27661	
2279	4453	7202	9910	13092	16089	18345	21245	24930	27675	
2609	4539	7218	9929	13281	16118	18499	21386	24967	27697	
2682	4748	7408	10167	13536	16215	18521	21562	25224	27721	
2733	4839	7428	10254	13556	16394	18588	22022	25312	27734	
2818	5049	7720	10340	13681	16503	19117	22056	25477	27802	

Çizelge 3.69'daki verilerin saçılımı Şekil 3.2'de verilmiştir.



Şekil 3.2 New York bölgesinde 30 Nisan 1969'da uçakların iniş zamanları arasında geçen zamanların (saniye olarak) uçakların iniş sırasına göre saçılımı

Bu çalışmada bu veri kümesi kullanılarak Eşitlik (3.12)'den elde edilen tahminler Çizelge (3.70)'te verilmiştir.

Çizelge 3.70. Hava trafiği veri kümesi için λ_1 ve λ_2 bilinmediği durumda τ ile λ_1 ve λ_2 için tahminler

n = 213, T= 28746 (saniye)	
Tahmin	
$\hat{\tau}$	28116 (saniye)
$\hat{\lambda}_1$	0.00754 (saniye)
$\hat{\lambda}_2$	0.00159 (saniye)

Bu veri kümesi için hipotezler,

$$H_0 : \lambda = \lambda_1, \quad 0 < t \leq 28746$$

$$H_1 : \lambda = \lambda_1, \quad 0 < t \leq \tau$$

$$: \lambda = \lambda_2, \quad \tau < t \leq 28746$$

ya da

$$H_0: (0, 28746] \text{ aralığında } \lambda' \text{ da de\u0131işme yoktur.}$$

$$H_1: \lambda, \tau' \text{ dan sonra de\u0131işir.} \quad (3.31)$$

biçimindedir. Çizelge 3.70'teki τ ile λ_1 ve λ_2 'nin tahminleri kullanılarak, Eşitlik (3.16)'da bulunan olabilirlik oranı test istatistiğinin olasılık yoğunluk fonksiyonundaki a ve b de\u0131işmezleri, Eşitlik (3.19)'dan sırasıyla -0.00595 ve 168.85503 elde edilmiş; bu de\u0131erler kullanılarak X raslantı de\u0131işkeninin tanım aralığı $-\infty < x < 168.85503$ bulunmuştur. Buna göre, $\alpha = 0.05$ yanılma düzeyinde kritik bölgenin sınırları C_0 , C_1 ve X_{Hesap} , Çizelge 3.71'de verildiği gibi elde edilmiştir.

Çizelge 3.71. Eşitlik (3.31)'de verilen hipotezlerin testi

C_0	C_1	X_{Hesap}
-22.53913	22.58773	-2.18367

$X_{\text{Hesap}} > -22.53913$ olduğu için H_0 hipotezi kabul edilir. Hsu'nun 1979'taki ve Ramanayake'nin 2004'teki çalışmalarında da uçakların inişleri arasında geçen zamanın dağılımında zaman içinde de\u0131işme olmadığı sonucuna varılmıştır.

4. SONUÇ ve TARTIŞMA

Bu çalışmada homojen Poisson sürecinde değişme noktasına ve parametrelere ilişkin tahmin edicilerin ve değişmenin olup olmadığını testi ve test istatistiğinin elde edilmesi amaçlanmıştır. Bu amaçla bazı varsayımlar altında tahmin ediciler analitik olarak elde edilmiştir. Değişmenin olup olmadığını testi için olabilirlik oranı istatistiğinin H_0 hipotezinin doğru olduğu varsayımı altında dağılımı bulunmuştur. Bu istatistiğin dağılımından yararlanılarak test istatistiğinin işlevselliği çeşitli uygulamalarla gösterilmiştir.

Literatürde değişme noktası ile ilgili çalışmalarda $\frac{\tau}{T}$ oranının bir aralık içinde olması gerektiği belirtilmiş; ancak bu aralığın sınırları verilmemiştir. Bu çalışmada, $\frac{\tau}{T}$ oranının $\frac{1}{6} < \frac{\tau}{T} < \frac{5}{6}$ aralığında olması gerektiği saptanmıştır. $\frac{\tau}{T}$ oranının bu aralığın sınırlarına yakın ya da bu sınırların dışında olduğunda Kesim 3.1.1 ve Kesim 3.1.2'de verilen tahmin edicilerden elde edilen tahminlerin gerçek değerlerden uzaklaştıkları görülmüştür (Çizelge 3.17, Çizelge 3.18, Çizelge 3.35 ve Çizelge 3.36).

Kesim 3.1.1, Kesim 3.1.2 ve Kesim 3.1.3'te önerilen tahmin edicilerden elde edilen tahminlerin gerçek değerlere yakın oldukları Çizelge 3.1- Çizelge 3.16, Çizelge 3.19- Çizelge 3.34, Çizelge 3.37- Çizelge 3.52'de görülmektedir. Ayrıca çalışmada önerilen tahmin edicilerden İngiltere'deki kömür madeni kazalarına ilişkin gerçek veri kümesi için hesaplanan tahminler Çizelge 3.53, Çizelge 3.54 ve Çizelge 3.55'te verilmiş; bu tahminlerin, Poisson süreçlerinde ve üstel dağılımlı raslantı değişkenleri dizisinde değişme noktası sorunu ile ilgili önceki çalışmalarda elde edilen ve Çizelge 2.2'de verilen tahminlerle uyumlu oldukları görülmüştür.

Kesim 3.2.1'de, olabilirlik oranına dayalı test istatistiğinin dağılımı elde edilmiş; test istatistiğinin davranışı λ_1, λ_2 arasındaki farkın yeterince büyük olduğu, yani değişmenin önemli olduğu ve λ_1, λ_2 arasındaki farkın çok küçük olduğu, yani değişmenin önemsiz olduğu durumlar için incelenmiştir. Yapılan uygulamalarda test fonksiyonunun H_0 hipotezinin reddine ya da kabulüne ilişkin doğru sonuçlar

vermekte olduđu Çizelge 3.56-Çizelge 3.61, Çizelge 3.63, Çizelge 3.65'ten görölmektedir. Ayrıca test istatistiğinin İngiltere'deki kömür madeni kazaları ile ilgili veri kümesindeki sonuçları Çizelge 3.66-Çizelge 3.68'de verilmiş; kömür madeni kazaları arasında geçen zamanın ortalamasında bir zaman noktasından sonra artma olduđu, bir başka deyişle birim zamanda (bir günde) ortaya çıkması beklenen kaza sayısında azalma olduđu sonucuna varılmıştır. Literatürde kullanılan başka bir veri kümesi de hava trafiğine ilişkin veri kümesidir. Elde edilen test istatistiğinin bu veri kümesinde verdiğı sonuç hava alanına inen uçakların sayılarının dağılımında değışmenin olmadığı yönündedir. Çizelge 3.71'de verilen sonuçların, önceki çalışmalarda elde edilen sonuçlarla uyumlu oldukları görölmektedir.

Kesim 3.1.1, Kesim 3.1.2 ve Kesim 3.1.3'te önerilen tahmin edicilerin homojen Poisson sürecinde değışme noktası sorununa ilişkin çalışmalarda elde edilen tahmin edicilerle karşılaştırılmak için tahmin edicilerin dağılımının bulunması gerekmektedir. Ancak hem tez çalışmasında elde edilen tahmin edicilerin hem de önceki çalışmalara da elde edilen tahmin edicilerin dağılımlarına ulaşılammıştır. Bu nedenle kuramsal dayanaklı bir karşılaştırma yapılamamış; bu karşılaştırma önceki çalışmalarda kullanılan gerçek veri kümeleri kullanılarak yapılmıştır.

λ_1 ve λ_2 bilindiğinde Eşitlik (3.5a)'da elde edilen τ 'nun tahmin edicisinden örneklem büyüklüğü ve gözlem zamanının uzunluğu kullanılarak tahminlerin elde edilmesi çok kolaydır. Ancak, τ 'nun tahmin edicisinin dağılımının bulunması kolay değildir. τ 'nun tahmin edicisinin dağılımı bulabilmek için öncelikle T 'nin dağılımını bulunması gerekir. T , n 'inci olay ortaya çıkıncaya dek geçen, (N_τ, λ_1) parametreleriyle gamma dağılımlı raslantı değışkeni ile $(n-N_\tau, \lambda_2)$ parametreleriyle gamma dağılımlı raslantı değışkeninin toplamıdır. N_τ bilinmediği için T 'nin dağılımının bulunması oldukça güçtür. Ancak bu olanaksızdır anlamına gelmez.

Bu çalışmada, λ_1, λ_2 bilindiğinde τ 'nun; λ_1 bilindiğinde τ ve λ_2 'nin; λ_1, λ_2 bilinmediğinde, τ, λ_1 ve λ_2 'nin tahmin edicilerine ulaşılmıştır. Bu tahmin edicilerin özellikleri daha sonraki çalışmalarda incelenebilir. Tahmin edicilerin dağılımları bulunabilirse, aralık tahminleri elde edilebilir. Bu çalışmada, test istatistiğinin H_0

hipotezinin doğru olduđu varsayımı altında dağılımı bulunmuştur. Test istatistiğinin H_1 hipotezinin doğru olduđu varsayımı altında dağılımı elde edilerek test fonksiyonunun gücü ile ilgili çalışmalar öteki bir çalışma alanı olarak önerilebilir.

Üçüncü Bölüm'de elde edilen kuramsal sonuçların geçerliliği ve yapılan uygulamalar benzetim tabanlı çalışmalara dayanmaktadır. Geniş ölçekli benzetim çalışmalarına ve ülkemize ait gerçek veri kümelerine, elde edilen kuramsal sonuçların uygulanması, hem kuramsal sonuçlarla ilgili daha ayrıntılı bilginin elde edilmesinin hem de ülkemize ait mevcut bir sorununun çözümüne katkısının araştırılmasını sağlayacaktır. Bu nedenle elde edilen kuramsal sonuçların bu amaçlar çerçevesinde ileriki çalışmalarda kullanılması önerilebilir.

Son olarak, bu çalışmada homojen Poisson sürecinde değişme noktası sorunu incelenmiştir. Bu çalışmada elde edilen sonuçlar, homojen olmayan Poisson süreci ve birleşik Poisson sürecinde değişme noktası sorununun incelenmesinde yararlı olacaktır.

KAYNAKLAR

- Akman, V. E. and Raftery, A. E., 1986, Asymptotic inference for a change-point Poisson process, *The Annals of Statistics*, 14, 4, pp. 1583-1590.
- Basseville, M. and Nikiforov, I. V., 1993, *Detection of Abrupt Change: Theory and Application*, Prentice-Hall, New York, 447p.
- Boudjellaba, H., MacGibbon, B. and Sawyer, P., 2001, On exact inference for change in a Poisson sequence, *Communications in Statistics A: Theory and Methods*, 30(3), pp. 407-434.
- Chen, J. and Gupta, A. K., 1999, Change point analysis of a gaussian model, *Statistical Papers*, 40, pp. 323-333.
- Chen, J. and Gupta, A. K., 2001, On change point detection and estimation, *Communications in Statistics B: Simulation and Computation*, 30(3), pp. 665-697.
- Feller, W., 1966, *An Introduction to Probability Theory and Its Application*, John Wiley and Sons Inc., New York, Volume II, 669p.
- Fotopoulos, S. and Jandhyala, V., 2001, Maximum likelihood estimation of a change-point for exponentially distributed random variables, *Statistics & Probability Letters*, 51, pp. 423-429.
- Gombay, E. and Horváth, L., 1996, On the rate of approximations for maximum likelihood tests in change-point models, *Journal of Multivariate Analysis*, 56, pp. 120-152.
- Haccou, P., Meelis, E. and Geer, S., 1988, The likelihood ratio test for the change point problem for exponentially distributed random variables, *Stochastic Process and Their Applications*, 27, pp. 121-139.

- Hinkley, D. V., 1970, Inference about the change-point in a sequence of random variables, *Biometrika*, 57, 1, pp. 1-17.
- Hinkley, D. V. and Hinkley, E. A., 1970, Inference about the change-point in a sequence of binomial variables, *Biometrika*, 57, 3, pp. 477-488.
- Hinkley, D. V., 1971, Inference about the change-point from cumulative sum tests, *Biometrika*, 58, 3, pp. 509-523.
- Hsu, D. A., 1979, Detecting shifts of parameter in gamma sequences with applications to stock prices and air traffic flow analysis, *Journal of the American Statistical Association*, 74, pp. 31-40.
- James, B., James, K. L. and Siegmund, D., 1987, Tests for a change-point, *Biometrika*, 74, 1, pp. 71-83.
- Jandhyala, V. and Fotopoulos, S., 1999, Capturing the distributional behaviour of the maximum likelihood estimator of a change point, *Biometrika*, 86, 1, pp. 129-140.
- Jandhyala, V. and Fotopoulos, S., 2001, Rate of convergence of the maximum likelihood estimate of a change-point, *Sankhyā A: The Indian Journal of Statistics*, 63, 2, pp. 277-285.
- Jarett, R. G., 1979, A note on the intervals between coal-mining disasters, *Biometrika*, 66, pp. 191-193.
- Loader, C. R., 1992, A log-linear model for a Poisson process change point, *The Annals of Statistics*, 20, 3, pp. 1391-1411.
- Ramanayake, A., 2004, Tests for a change point in the shape parameter of gamma random variables, *Communications in Statistics A: Theory and Methods*, 33, 4, pp. 821-833.

- Raftery, A. E. and Akman, V. E., 1986, Bayesian analysis of a Poisson process with a change point, *Biometrika*, 73, pp. 85-89.
- Sen, A. and Srivastava, M. S., 1975, On tests for detecting change in mean, *The Annals of Statistics*, 3, 1, pp. 98-108.
- Siegmund, D., 1988, Confidence sets in change-point problems, *International Statistical Review*, 56, 1, pp. 31-48.
- West, R. W. and Ogden, R. T., 1997, Continuous-time estimation of a change-point in a Poisson process, *Journal of Statistical Computation and Simulation*, 56, pp. 293-302.
- Worsley, K. J., 1979, On the likelihood ratio test for a shift in location of normal populations, *Journal of the American Statistical Association*, 74, pp. 365-367.
- Worsley, K. J., 1983, The power of likelihood ratio and cumulative sum tests for a change in a binomial probability, *Biometrika*, 70, 2, pp. 455-464.
- Worsley, K. J., 1986, Confidence region and test for a change-point in a sequence of exponential family random variables, *Biometrika*, 73, 1, pp. 91-104.

ÖZGEÇMİŞ

Adı Soyadı : Ayten YİĞİTER

Doğum Yeri : Şereflikoçhisar

Doğum Yılı : 11. 09. 1972

Eğitim ve Akademik Durumu

Lise :1985-1989, Ankara Lisesi

Lisans :1991-1995, Hacettepe Üniversitesi Fen Fakültesi İstatistik Bölümü

Yüksek Lisans :1997-2000, Hacettepe Üniversitesi Fen Fakültesi İstatistik Bölümü

Yabancı Dil : İngilizce

İş Tecrübesi

1998-2006 Hacettepe Üniversitesi Fen Fakültesi İstatistik Bölümü'nde Araştırma Görevlisi