



EGE ÜNİVERSİTESİ

YÜKSEK LİSANS TEZİ

**SÜREKLİ DEĞERLİ DİZİLERDE
TEKRAR İSTATİSTİKLERİ**

Ash Ayşe BİLİR

Tez Danışmanı : Yrd. Doç. Dr. Sevcan DEMİR ATALAY

İstatistik Anabilim Dalı

Bilim Dalı Kodu : 406.02.01

Sunuş Tarihi : 27.06.2012

Bornova-İZMİR

2012

E. Ü. FEN BİLİMLERİ ENSTİTÜSÜ

EGE ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ

(YÜKSEK LİSANS TEZİ)

**SÜREKLİ DEĞERLİ DİZİLERDE
TEKRAR İSTATİSTİKLERİ**

Aslı Ayşe BİLİR

Tez Danışmanı : Yrd. Doç. Dr. Sevcan DEMİR ATALAY

İstatistik Anabilim Dalı

Bilim Dalı Kodu : 406.02.01

Sunuş Tarihi : 27.06.2012

Bornova-İZMİR

Aslı Ayşe BİLİR tarafından yüksek lisans tezi olarak sunulan "Sürekli Değerli Dizilerde Tekrar İstatistikleri" başlıklı bu çalışma E.Ü. Lisansüstü Eğitim ve Öğretim Yönetmeliği ile E.Ü. Fen Bilimleri Enstitüsü Eğitim ve Öğretim Yönergesi'nin ilgili hükümleri uyarınca tarafımızdan değerlendirilerek savunmaya değer bulunmuş ve 27.06.2012 tarihinde yapılan tez savunma sınavında aday oybirliği/oyçokluğu ile başarılı bulunmuştur.

Jüri Üyeleri:

İmza

Jüri Başkanı

: Yrd. Doç. Dr. Sevgen DEMİR
ATALAY

Raportör Üye

: Yrd. Doç. Dr. Ali MERT

Üye

: Yrd. Doç. Dr. Burcu ADIVAR

S. Demir
Ali Mert
Burcu Adıvar

ÖZET**SÜREKLİ DEĞERLİ DİZİLERDE
TEKRAR İSTATİSTİKLERİ**

BİLİR, Aslı Ayşe

Yüksek Lisans Tezi, İstatistik Anabilim Dalı
Tez Danışmanı: Yrd. Doç. Dr. Sevcan DEMİR ATALAY
Haziran 2012, 84 sayfa

Bu tez çalışmasında, literatürde sürekli değerli dizilerin sabit bir kriter değere göre iki değerli hale dönüştürülmesi işlemine alternatif bir yöntem sunulmuştur. Bulanık küme teorisine dayanan bu yöntem ile karar vericinin tercihinine göre değişebilen esnek bir modelleme imkanı sağlanmıştır. Bu bağlamda, “klasik model” ve “önerilen model” olarak adlandırılan bu iki yöntemin karşılaştırılması amacıyla hidroloji alanında bir uygulama çalışması yapılmış ve önerilen modele uygun olarak elde edilen tekrar istatistiklerinin olasılık dağılımları kullanılarak ileriye dönük tahminlerde bulunulmuştur.

Anahtar sözcükler: Tekrar istatistikleri, iki değerli diziler, sürekli değerli diziler, bulanık küme teorisi, klasik model, önerilen model, kriter değer, aşma olasılığı

ABSTRACT

RUN STATISTICS IN

CONTINUOUS VALUED SEQUENCES

BİLİR, Aslı Ayşe

MSc in Statistics

Supervisor: Assoc. Prof. Dr. Sevcan DEMİR ATALAY

June 2012, 84 pages

In this thesis, an alternative method is presented to the procedure of transforming continuous valued sequences into binary sequences in accordance with a constant critical value. With this method based on the fuzzy sets theory a flexible modelling which can alter according to the decision maker's choice is provided. Within this context aiming comparison of the two models called as "classical model" and "proposed model", an applicational study in the field of hydrology has been carried out and using probability distributions of run statistics retrieved in accordance with the proposed model, predictions have been made.

Keywords: Run statistics, binary sequences, continuous valued sequences, fuzzy set theory, classical model, proposed model, critical value, exceedance probability

TEŞEKKÜR

Lisans eğitimimden sonra uzun bir ara vererek başladığım ve iş hayatı ile birlikte yürütmeye çalıştığım yüksek lisans eğitimimde, öncelikle ders aşamasını başarılı bir şekilde tamamladığım Ondokuz Mayıs Üniversitesi (OMÜ) İstatistik Bölümü Öğretim Üyesi Doç. Dr. Sayın V. Rezan USLU ve Doç Dr. Sayın Erol EĞRİOĞLU hocalarıma, yardımını hiçbir zaman esirgemeyerek eski bilgilerimi hatırlamamı sağlayan Hacettepe Üniversitesi İstatistik Bölümü Öğretim Üyesi Doç. Dr. Sayın Serpil AKTAŞ ALTUNAY hocama;

Ege Üniversitesi'ne yatay geçiş yapmamda en büyük desteği sağlayan Türkiye İstatistik Kurumu (TÜİK) İzmir Bölge Müdürü Sayın Rıdvan YAKA'ya ve tüm çalışma arkadaşlarıma, yüksek lisans eğitimim boyunca verdiği değerli katkılardan dolayı Ege Üniversitesi İstatistik Bölüm Başkanı Prof. Dr. Sayın Şanslı ŞENOL hocama;

Tezin her aşamasında bana yardımcı olan ve büyük bir titizlikle çalışmamı sağlayan danışmanım Ege Üniversitesi İstatistik Bölümü Öğretim Üyesi Yrd. Doç. Dr. Sayın Sevcan DEMİR ATALAY hocama, bu süreçte her konuda desteğini gördüğüm Ege Üniversitesi İstatistik Bölüm Başkan Yrd., Yrd. Doç. Dr. Sayın Ali MERT hocama sonsuz teşekkür ve saygılarımı sunarım.

Bugünlere gelmemde büyük emekleri olan annem Güner ÖZGAN, babam Necdet ÖZGAN ve abim Osman ÖZGAN'a, ders çalışmama ve tez yazmama ortam hazırlayarak manevi desteklerini hiçbir zaman esirgemeyen annelerim Güner ÖZGAN ve Kerime BİLİR'e, sabır ve yardımlarından dolayı sevgili eşim İsmail BİLİR'e ve tüm annelerin ders çalıştığını düşünen ve bu durumun bir an önce bitmesini sabırsızlıkla bekleyen canım kızım Ceylan Nilsu BİLİR'e sonsuz teşekkür ve sevgilerimi sunarım.

İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET	v
ABSTRACTvii
TEŞEKKÜR	ix
ŞEKİLLER DİZİNİ	xiii
TABLolar DİZİNİ.....	xiv
1.GİRİŞ.....	1
2.GENEL BİLGİLER.....	3
2.1 Tanım ve Kavramlar	4
2.2 Bazı Tekrar İstatistiklerinin Olasılık Dağılımları	9
3.TEKRAR İSTATİSTİKLERİNİN UYGULAMA ALANLARI.....	14
3.1 Rasgelelik Testleri	14
3.2 Hidroloji.....	18
3.3 İstatistiksel Süreç Kontrolü	20
3.4 Moleküler Biyoloji	21
3.5 Güvenilirlik.....	22

İÇİNDEKİLER (devam)

	<u>Sayfa</u>
4.ÖNERİLEN MODEL	24
5.UYGULAMA	30
5.1 Değişken ve Havza Seçimi	32
5.2 Modelleme	34
5.3 Önerilen Modelin Havzalara Uygulanması	39
5.3.1 Küçük Menderes havzası	39
5.3.2 Gediz havzası	44
5.3.3 Kızılırmak havzası	50
5.3.4 Sakarya havzası	55
5.4 İleriye Dönük Tahminlerin Yapılması	61
5.5 Havzaların Karşılaştırılması	71
6.SONUÇ	75
KAYNAKLAR DİZİNİ	77
ÖZGEÇMİŞ	84

ŞEKİLLER DİZİNİ

<u>Şekil</u>	<u>Sayfa</u>
4.1. Üçgensel bulanık sayı.....	25
4.2 Risk düzeyi ile rasgele değişken arasındaki ilişki	26
5.1 MYAMA ve taşkın risk düzeyi arasındaki doğrusal ilişki	34
5.2 Log-logistik (3P) dağılım grafiği ($\alpha=3,2$, $\beta=223,5$, $\gamma=-28,1$)	40
5.3 Log-logistik (3P) dağılım grafiği ($\alpha=3,2$, $\beta=477,5$, $\gamma=-7,6$)	46
5.4 Log-logistik (3P) dağılım grafiği ($\alpha=2,8$, $\beta=367,3$, $\gamma=335,4$).....	51
5.5 Log-logistik (3P) dağılım grafiği ($\alpha=3,8$, $\beta=364,1$, $\gamma=181,1$).....	56
5.6 a) $G_{n,k}$, b) S_n , c) L_n , d) T_k tekrar istatistiklerine göre taşkın tahmin olasılıkları ile kriter değerler arasındaki ilişki.....	65
5.7 e) $G_{n,k}$, f) S_n , g) L_n , h) T_k tekrar istatistiklerine göre taşkın tahmin olasılıkları ile kriter değerler arasındaki ilişki.....	70

TABLolar DİZİNİ

<u>Tablo</u>	<u>Sayfa</u>
3.1. İzmir İline düşen yıllık toplam yağış miktarları (mm).....	15
3.2 R_n nin olasılık dağılımı ve p olasılık değerleri.....	16
5.1 Tanımlayıcı istatistikler.....	40
5.2 Run testi.....	41
5.3 Düzeylere göre MYAMA değerlerinin iki değerli hale çevrilmesi.....	42
5.4 Taşkın tahmin ve kaynakların etkin kullanım oranları.....	43
5.5 Tanımlayıcı istatistikler.....	45
5.6 Run testi.....	46
5.7 Düzeylere göre MYAMA değerlerinin iki değerli hale çevrilmesi.....	48
5.8 Taşkın tahmin ve kaynakların etkin kullanım oranları.....	49
5.9 Tanımlayıcı istatistikler.....	51
5.10 Run testi.....	52
5.11 Düzeylere göre MYAMA değerlerinin iki değerli hale çevrilmesi.....	53
5.12 Taşkın tahmin ve kaynakların etkin kullanım oranları.....	54
5.13 Tanımlayıcı istatistikler.....	56
5.14 Run testi.....	57
5.15 Düzeylere göre MYAMA değerlerinin iki değerli hale çevrilmesi.....	58

TABLolar DİZİNİ (devam)

<u>Tablo</u>	<u>Sayfa</u>
5.16 Taşkın tahmin ve kaynakların etkin kullanım oranları.....	60
5.17 Aşma olasılıkları.....	61
5.18 1991-1997 yılları arası ileriye dönük taşkın tahmin oranları	63
5.19 Aşma olasılıkları.....	66
5.20 1991-1997 yılları arası ileriye dönük taşkın tahmin oranları	68
5.21 Havzaların karşılaştırılması	73

1. GİRİŞ

İstatistik biliminde, belirli bir araştırmaya dayalı olarak değişkenlerin tanımlanması, bu değişkenlerin alacağı sayısal değerlerin (veri) belirlenmesi ve sayısal değerlerden oluşan diziler üzerinde istatistiksel yöntemler kullanılarak tahmin ve çıkarımların yapılması çok önemlidir.

Değişkenler, gözlenme biçimlerine göre “nitel” veya “nicel” olarak gruplandırılırken, ölçülen değerlerin matematiksel durumuna göre “kesikli” veya “sürekli” olarak ikiye ayrılırlar (Saraçbaşı ve Kutsal, 1987).

Kesikli veya sürekli verilerden oluşan bir dizi, içerdiği verinin yapısına ve elde edilmiş biçimine göre doğrusal (linear) olabildiği gibi dairesel (circular) olarak tanımlanabilir (Ghent and Zar, 1992). Örneğin; günlük ortalama sıcaklık, aylık toplam yağış miktarı, yıllık maksimum akım değeri gibi ardıl verilerin oluşturduğu diziler doğrusal dizilerdir. Dakikalık periyotlara bölünmüş 24 saatlik bir zaman diliminde t günde n sayıda bebeğin doğum saatlerinin frekans dağılımı, saatlik periyotlara bölünmüş 24 saatlik bir zaman diliminde yağışların çok olduğu t günde toplam yağış miktarlarının saatlere göre dağılımı, yaz aylarını içeren bir zaman diliminde t yılda maksimum sıcaklıkların aylara göre dağılımı, mevsimlik periyotlara bölünmüş 1 yıllık zaman diliminde deprem, güneş tutulması, gel-git gibi doğa olaylarının t yılda mevsimlere göre frekans dağılımı, hücre kesitleri gibi doğada dairesel olarak bulunan sistemlerdeki nitel özelliklerin incelenmesi gibi başlangıç ve bitiş noktası arasında bir ilişkinin (döngünün) bulunduğu diziler dairesel dizilerdir.

$\xi_1, \xi_2, \dots, \xi_n$ doğrusal ya da dairesel olarak tanımlanabilen iki değerli rasgele bir dizinin elemanlarını gösterebilir. Başarı-başarısız, sağlam-bozuk ve sulak-kurak gibi tanımlamaların “1” ve “0” elemanları ile temsil edilebildiği bu dizide, üst üste gelen aynı tür elemanların diğer türden elemanlarla sonlanması ile oluşan alt dizilere tekrar (run) adı verilmektedir.

İki değerli rasgele dizilere dayalı olarak tanımlanan tekrar istatistikleri ve bu istatistiklerin dağılımları, rasgelelik testleri, hidroloji, istatistiksel kalite kontrol süreçleri, moleküler biyoloji ve güvenilirlik gibi çeşitli sistemlerin karar verme süreçlerinde etkin bir yöntem olarak kullanılmaktadır.

Bu tez çalışmasının ikinci bölümünü oluşturan genel bilgiler başlığı altında öncelikle tekrar istatistikleri ile ilgili literatürde yer alan çalışmalar özetlenmiş ve tekrar istatistiklerinin tanımları verilerek bazı varsayımlar altında dağılım özellikleri incelenmiştir. Ardından, belirli kriter değerler kullanılarak iki değerli hale çevrilebilen sürekli değerli diziler üzerinde tekrar istatistiklerinin tanımlanabileceği gösterilmiştir. Bölümün sonunda ise bazı tekrar istatistiklerinin olasılık dağılımları verilmiştir. Çalışmanın üçüncü bölümünde, tekrar istatistiklerinin uygulama alanları incelenmiş ve literatüre parametrik olmayan testlerden “run” testi olarak geçen rasgelelik testi için bir örnek sunulmuştur. Dördüncü bölümde, sürekli değerli dizilerin iki değerli hale çevrilmesi ile ilgili klasik modele alternatif bir model önerilmiştir. Klasik modelde kullanılan sabit eşik değerinin yerine, karar vericinin tercihine göre değişebilen bir kriter değer kullanılması gerektiği vurgulanarak, bulanık küme mantığına göre oluşturulan “önerilen model” ile ilgili tanımlamalar yapılmıştır. Beşinci bölümde, önerilen modelin örneklenmesi amacıyla tekrar istatistiklerinin uygulama alanlarından birisi olan hidroloji tercih edilmiştir. Hidrolojide ekstrem doğa olayları hayati önem taşıdığı için doğal afet olarak nitelendirilen sel ve taşkın olayları incelenmiş ve bu olaylar üzerinde en çok etkisi olan değişken belirlenerek o değişkene ait veri kümesi kullanılmıştır. Ardından gerekli hesaplamalar yapılarak ilgili veri kümesi üzerinde tekrar istatistikleri tanımlanmış ve tekrar istatistiklerinin olasılık dağılımları kullanılarak ileriye dönük tahminler yapılmıştır.

2. GENEL BİLGİLER

Tekrar istatistikleri ile ilgili ilk çalışma, De Moivre'nin 1738 yılına ait "Doctrine of Chance" adlı yapıtında yer almaktadır. Geçmiş bu kadar eskiye dayansa da, tekrar istatistikleri ile ilgili sistematik ve stokastik modelleme çalışmalarına 1940'larda ağırlık verilmiştir.

Mood (1940), Wald and Wolfowitz (1940), Mosteller (1941), Wolfowitz (1943), tekrar istatistikleri teorisinin temelini oluşturan toplam tekrar sayısı (R_n) ve toplam başarı tekrarlarının sayısına ($R_n^{(1)}$) dayalı klasik tekrar istatistikleri üzerine çalışmışlardır.

Philippou and Makri (1986), Hirano (1986), Godbole (1990a), k uzunluklu başarı tekrarlarının sayısını veren $E_{n,k}$ tekrar istatistiği üzerine, Muselli (1996), k ve daha büyük uzunluklu başarı tekrarlarının sayısını veren $G_{n,k}$ tekrar istatistiği üzerine, Aki and Hirano (1988), Hirano (1986), Philippou and Makri (1986), Godbole (1992b), üst üste düşmeyen, ardıl k uzunluklu başarı tekrarlarının sayısını veren $N_{n,k}$ tekrar istatistiği üzerine, Ling (1988, 1989), Hirano et al. (1991), Godbole (1992b), Chryssaphinou et al. (1993), Koutras and Alexandrou (1997), Godbole et al. (1997), üst üste düşen, ardıl k uzunluklu başarı tekrarlarının sayısını veren $M_{n,k}$ tekrar istatistiği üzerine, Philippou and Muwafi (1982), k uzunluklu bir başarı tekrarı elde edinceye kadar yapılan deneme sayısını veren T_k tekrar istatistiği üzerine, Philippou and Makri (1986), Godbole (1990a), Muselli (1996), Fu (1986), Lou (1996), Fu and Lou (2003), Vaggelatos (2003), Eryılmaz (2005b), en uzun başarı tekrarının uzunluğunu veren L_n tekrar istatistiği üzerine çalışmışlar ve literatüre kazanımlar sağlamışlardır.

Şimdiye kadar çeşitli kaynaklarda ayrıntılı olarak yer alan tekrar istatistikleri ile ilgili özellikle, Balakrishnan and Koutras (2002) ile Fu and Lou (2003) tarafından hazırlanan kaynak kitaplar geniş bir içerik sunmaktadır.

Bu bölümde, tekrar istatistikleri ile ilgili genel tanım ve kavramlar verilmiş ve tekrar istatistiklerinin dağılım özellikleri bazı varsayımlar altında incelenmiştir. Ayrıca, sürekli değerlerden oluşan rasgele bir dizinin uygun bir kriter değer belirlenerek, iki değerli rasgele bir diziye dönüştürülmesinden söz edilmiş ve bazı tekrar istatistiklerinin olasılık dağılımları verilmiştir.

2.1 Tanım ve Kavramlar

Literatürde “run” olarak bilinen tekrar kavramı üst üste gelen aynı tür elemanların oluşturduğu dizinin diğer elemanlarla sonlanması ile oluşan bir alt dizi olarak tanımlanır. Temel olarak, “1” ile başarının “0” ile başarısızlığın temsil edildiği iki değerli rasgele dizilerde tekrarlar, başarı ya da başarısızlıkların ardıl olarak gelmesiyle oluşan alt dizilerdir (Balakrishnan and Koutras, 2002; Fu and Lou, 2003).

Örneğin, 100111011 şeklindeki ardıl denemeler dizisinde, ilk olarak “1” uzunluğunda bir başarı tekrarı, “2” uzunluğunda başarısızlık tekrarı, “3” uzunluğunda başarı tekrarı, “1” uzunluğunda başarısızlık tekrarı ve son olarak ta “2” uzunluğunda başarı tekrarı elde edilmiştir. Uzunluk, tekrar içerisindeki deneme sayısını göstermektedir.

$\xi_1, \xi_2, \dots, \xi_n$ iki değerli rasgele bir dizinin elemanlarını göstermek üzere, bu diziye dayalı olarak tanımlanan ve en çok kullanılan tekrar istatistikleri aşağıda verilmektedir.

- R_n , n deneme sonundaki toplam tekrar sayısı ($R_n = R_n^{(1)} + R_n^{(0)}$) (Mood, 1940; Wald and Wolfowitz, 1940; Mosteller, 1941; Wolfowitz, 1943),

- $R_n^{(1)}$, n deneme sonundaki toplam başarı tekrarlarının sayısı,

- $R_n^{(0)}$, n deneme sonundaki toplam başarısızlık tekrarlarının sayısı,

- L_n , n deneme sonundaki en uzun başarı tekrarının uzunluğu (Philippou and Makri, 1986; Godbole, 1990a; Muselli, 1996; Fu, 1986; Lou, 1996; Fu and Lou, 2003; Vaggelatou, 2003; Eryılmaz, 2005b),

- $E_{n,k}$, n deneme sonundaki k uzunluklu başarı tekrarlarının sayısı (Philippou and Makri, 1986; Hirano, 1986; Godbole, 1990a),

- $G_{n,k}$, n deneme sonundaki k ve daha büyük uzunluklu başarı tekrarlarının sayısı (Muselli, 1996),

- $N_{n,k}$, n deneme sonundaki üst üste düşmeyen, ardıl k uzunluklu başarı tekrarlarının sayısı (Aki and Hirano, 1988; Hirano, 1986; Philippou and Makri, 1986; Godbole, 1992b),

- $M_{n,k}$, n deneme sonundaki üst üste düşen, ardıl k uzunluklu başarı tekrarlarının sayısı (Ling, 1988, 1989; Hirano et al., 1991; Godbole, 1992b; Chryssaphinou et al., 1993; Koutras and Alexandrou, 1997; Godbole et al., 1997),

- T_k , k uzunluklu bir başarı tekrarı elde edinceye kadar yapılan deneme sayısı (Philippou and Muwafi, 1982),

- S_n , n deneme sonundaki toplam başarı sayısı (Balakrishnan and Koutras, 2002; Fu and Lou, 2003)

olmak üzere; 110100111001111 şeklinde tanımlanan $n=15$ elemanlı iki değerli bir dizi için,

- 15 deneme sonundaki toplam tekrar sayısı, $R_{15} = 7$
- 15 deneme sonundaki toplam başarı tekrarlarının sayısı, $R_{15}^{(1)} = 4$
- 15 deneme sonundaki toplam başarısızlık tekrarlarının sayısı, $R_{15}^{(2)} = 3$
- 15 deneme sonundaki en uzun başarı tekrarının uzunluğu, $L_{15} = 4$
- 15 deneme sonundaki “2” uzunluklu başarı tekrarlarının sayısı, $E_{15,2} = 1$
- 15 deneme sonundaki “2” ve daha büyük uzunluklu başarı tekrarlarının sayısı, $G_{15,2} = 3$
- 15 deneme sonundaki üst üste düşmeyen, ardıl 2 uzunluklu başarı tekrarlarının sayısı, $N_{15,2} = 4$

1 1 0 1 0 0 1 1 1 0 0 1 1 1 1

- 15 deneme sonundaki üst üste düşen, ardıl “2” uzunluklu başarı tekrarlarının sayısı, $M_{15,2} = 6$

$$\underline{110100111001111}$$

- “2” uzunluklu bir başarı tekrarı elde edinceye kadar yapılan deneme sayısı, $T_2 = 2$

- “3” uzunluklu bir başarı tekrarı elde edinceye kadar yapılan deneme sayısı, $T_3 = 9$

- 15 deneme sonundaki toplam başarı sayısı, $S_{15} = 10$ olarak hesaplanmaktadır.

Literatürde tekrar istatistiklerinin dağılım özellikleri, bazı varsayımlar altında incelenmiştir. Buna göre, $\xi_1, \xi_2, \dots, \xi_n$ olarak tanımlanan iki değerli dizi elemanları,

- birbirinden bağımsız ve aynı Bernoulli dağılımına sahip olmaları,
 $P(\xi_i = 1) = p, P(\xi_i = 0) = q, p + q = 1 \quad i=1,2,\dots,n$

- birbirinden bağımsız, ancak farklı dağılıma sahip olmaları,
 $P(\xi_1 = 1) = p_i, P(\xi_1 = 0) = q_i, p_i + q_i = 1 \quad i=1,2,\dots,n$

- homojen Markov bağımlı olup,

$$P(\xi_{i+1} = 0 \mid \xi_i = 0) = P_{00}$$

$$P(\xi_{i+1} = 1 \mid \xi_i = 0) = P_{01}$$

$$P(\xi_{i+1} = 0 \mid \xi_i = 1) = P_{10}$$

$$P(\xi_{i+1} = 1 \mid \xi_i = 1) = P_{11} \quad i=1,2,\dots,n$$

geçiş olasılıklarına sahip olmaları,

- homojen olmayan Markov bağımlı olup,

$$P(\xi_{i+1} = 0 \mid \xi_i = 0) = P_{00}^{i+1}$$

$$P(\xi_{i+1} = 1 \mid \xi_i = 0) = P_{01}^{i+1}$$

$$P(\xi_{i+1} = 0 \mid \xi_i = 1) = P_{10}^{i+1}$$

$$P(\xi_{i+1} = 1 \mid \xi_i = 1) = P_{11}^{i+1} \quad i=1,2,\dots,n$$

geçiş olasılıklarına sahip olmaları,

- simetrik bağımlı (exchangeable) olmaları (Eryılmaz ve Demir, 2007)

varsayımları altında incelenmiştir.

Literatürde, iki değerli diziler üzerinde tanımlanan tekrar istatistikleri, sürekli değerli dizilerin analizinde de kullanılmıştır. Sürekli değerli diziler üzerinde tekrar istatistiklerinin tanımlanabilmesi amacıyla verilerin belirli kriter değerler kullanılarak iki değerli hale dönüştürülmesi yoluna gidilmiştir. Örneğin, X_1, X_2, \dots, X_n sürekli değerlerden oluşan rasgele bir dizi olsun. Dizi elemanları, ilgilenilen sistemin eşik değeri (trancation level) olarak adlandırılan bir x_0 kriter değeri ile karşılaştırılarak iki değerli rasgele bir diziye dönüştürülebilir. Birçok çalışmada ortalama, ortanca ve tepe değeri gibi veri kümesini özetleyen konum ölçüleri, x_0 kriter değeri olarak kullanılmıştır (Brownlee, 1965; Sokal and Rohlf, 1969; Ghent and Zar, 1992; O'Brien and Dyck, 1985; Toros, 1993).

X_1, X_2, \dots, X_n sürekli değerlerden oluşan rasgele bir dizi ve x_0 ilgilenilen sisteme ilişkin bir eşik değeri olmak üzere, elemanları $\xi_1, \xi_2, \dots, \xi_n$ olan ξ_i ($i = 1, 2, \dots, n$) rasgele değişkeninin alacağı değerler ve buna ilişkin olasılıklar aşağıdaki gibi tanımlanabilir.

$$\xi_i = \begin{cases} 1, & X_i > x_0 \\ 0, & X_i \leq x_0 \end{cases}$$

$$P(\xi_i = 1) = P(X_i > x_0) \quad (2.1)$$

$$P(\xi_i = 0) = P(X_i \leq x_0)$$

(2.1) eşitliğinde verilen $p = P(\xi_i = 1) = P(X_i > x_0)$ olasılık ifadesi “aşma olasılığı” olarak adlandırılmaktadır.

Çoğunlukla “ortanca” değeri olarak tercih edilen x_0 kriter değerinden büyük olan X_i değerleri için ξ_i ($i = 1, 2, \dots, n$) rasgele değişkeni “1” değerini alırken, x_0 kriter değerinden küçük veya eşit olan X_i değerleri için de “0” değerini alır. İlgilenilen sisteme göre “1” değeri; başarı, sağlam, sulak gibi tanımlamaları ifade ederken, “0” değeri başarısız, bozuk, kurak gibi “1” değerindeki tanımlamaların karşıt anlamlarını ifade eder.

O halde; elemanları X_1, X_2, \dots, X_n olan sürekli değerli bir dizi, belirlenen bir kriter değer sayesinde, elemanları $\xi_1, \xi_2, \dots, \xi_n$ olan iki değerli rasgele bir diziye dönüştürülerek tekrar istatistiklerinin kullanımına uygun hale getirilebilir.

Eryılmaz (2007), Eryılmaz ve Fu (2008), X_1, X_2, \dots, X_n rasgele değişkenler dizisinin sürekli değerler alması durumunu göz önüne almış ve ε komşuluk tekrarları olarak adlandırılan bir tanım ortaya koymuşlardır.

Eryılmaz'ın (2007) çalışmasında, X_1, X_2, \dots, X_n sürekli bir F dağılımına sahip bağımsız ve aynı dağılımlı rasgele değişkenler dizisi için ξ_k rasgele değişkeni (2.2) eşitliğinde verildiği gibi tanımlanarak, ε – komşuluklu R_n^ε tekrar istatistiğinin dağılımı elde edilmiştir.

$$\xi_k = \begin{cases} 1, & X_k \notin (X_{k-1} - \varepsilon, X_{k-1} + \varepsilon) \\ 0, & \text{diğer durumlarda} \end{cases}$$

$$R_n^\varepsilon = 1 + \sum_{k=2}^n \xi_k \quad (2.2)$$

$$R_n^\varepsilon = 1 + \xi_2 + \xi_3 + \dots + \xi_n$$

Örneğin, 3,2, 4,1, 4,2, 3,9, 1,7, 1,6 gibi sürekli değerlerden oluşan 6 elemanlı bir dizide, $\varepsilon = 0.35$ olarak alındığında R_n^ε nin değeri (2.2) eşitliğinden yararlanılarak hesaplanabilir.

İkinci değerden başlamak üzere bütün değerler, bir önceki değer ε kadar altı ve üstü ile kıyaslanarak aralığın dışında kalan değerler için ξ_k rasgele değişkeni “1” değerini alır. Diğer bir ifade ile dizideki ikinci değerden başlamak üzere bütün değerler bir önceki değer ile kıyaslanır ve elde edilen fark ε değeri ile karşılaştırılır. ε dan büyük olan fark değerleri “1” ile, küçük olan fark değerleri “0” ile gösterilir ve bu değerler ξ_2, \dots, ξ_n rasgele değişkenler dizisini oluşturur. Böylece, ξ_k değerlerine bağlı olarak R_n^ε elde edilmiş olur. O halde, $\varepsilon = 0.35$ olarak alındığında R_6^ε tekrar istatistiği “3” olarak hesaplanmaktadır.

2.2 Bazı Tekrar İstatistiklerinin Olasılık Dağılımları

Bu bölümde, literatürde en çok kullanılan tekrar istatistiklerinin olasılık dağılımları verilmiştir. Özellikle, kullanımı çok eski yıllara dayanan ve tekrar istatistikleri teorisinin temelini oluşturan R_n tekrar istatistiği, ayrıntılı olarak incelenmiştir.

R_n tekrar istatistiği:

Toplam tekrar sayısını veren R_n tekrar istatistiği, “1” ile ifade edilen tekrarların sayısı $R_n^{(1)}$ ile “0” ile ifade edilen tekrarların sayısı $R_n^{(0)}$ nin toplamından oluşmaktadır. ($R_n = R_n^{(1)} + R_n^{(0)}$)

“1” ve “0” değerlerini alabilen $\xi_1, \xi_2, \dots, \xi_n$ dizisi üzerinde tanımlanan R_n tekrar istatistiği, I_k ($k=2,3,\dots,n$) indikatör rasgele değişkenlerinin toplamı olarak elde edilebilmektedir.

$$I_k = \begin{cases} 1, & X_k \neq X_{k-1} \\ 0, & X_k = X_{k-1} \end{cases}$$

$$R_n = 1 + \sum_{k=2}^n I_k \quad (2.3)$$

$$R_n = I + I_2 + I_3 + \dots + I_n$$

Örneğin, 1101001110 şeklinde tanımlanan $n=10$ elemanlı iki değerli rasgele bir dizi için (2.3) eşitliği ile verilen tanımlamalar kullanıldığında, R_n tekrar istatistiği 6 olarak hesaplanmaktadır. Burada I_k indikatörü, ikinci sayıdan başlanmak üzere, her sayının bir önceki sayı ile kıyaslanması ve sayılar aynı ise “0”, farklı ise “1” değerini alması ile elde edilmektedir.

R_n tekrar istatistiğinin olasılık dağılımı, $\xi_1, \xi_2, \dots, \xi_n$ şeklinde tanımlanan dizi elemanlarının birbirinden bağımsız ve aynı Bernoulli dağılımına sahip olmaları varsayımı altında, $R_n^{(1)}$ ve $R_n^{(0)}$ nin ortak olasılık dağılımından elde edilmektedir.

Teorem 2.1 $R_n^{(1)}$ ve $R_n^{(0)}$ in ortak olasılık dağılımı, Mood (1940) tarafından aşağıdaki gibi tanımlanmıştır.

$$f_{R_n^{(1)}, R_n^{(0)}}(r_1, r_2) = \frac{c \binom{n_1 - 1}{r_1 - 1} \binom{n_2 - 1}{r_2 - 1}}{\binom{n_1 + n_2}{n_1}} \quad (2.4)$$

(2.4) eşitliğinde, $r_1 = 1, 2, \dots, n_1$, $r_2 = 1, 2, \dots, n_2$ olmak üzere, $r_1 = r_2$ veya $r_1 = r_2 \pm 1$ olarak ifade edilir. Eğer, $r_1 = r_2$ ise $c = 2$, $r_1 = r_2 \pm 1$ ise $c = 1$ değerini almaktadır.

Burada, $\xi_1, \xi_2, \dots, \xi_n$ dizisi içerisinde “1” türünden elemanların sayısı $S_n^{(1)}$ ve “0” türünden elemanların sayısı $S_n^{(0)}$ olmak üzere,

$$f_{R_n^{(1)}, R_n^{(0)}}(r_1, r_2) = P(R_n^{(1)} = r_1, R_n^{(0)} = r_2 | S_n^{(1)} = n_1, S_n^{(0)} = n_2)$$

şeklinde de ifade edilebilir.

İspat 2.1 $R_n^{(1)}$ ve $R_n^{(0)}$ nin ortak olasılık dağılımı, nesnelerin diziliş mantığından elde edilmektedir. Şöyleki; n sayıda aynı tür nesne (ayrıt edilemez), r hücreye hiçbir hücrenin boş kalmaması koşulu ile $\binom{n-1}{r-1}$ farklı şekilde dağıtılır.

Nesnelerin A ile, hücre ayırıcının B ile gösterildiği bir dizinin, r hücreye bölünebilmesi için $r - 1$ tane hücre ayırıcına ihtiyaç vardır. x_i , i . hücredeki nesne sayısını göstermektedir ($x_1 + x_2 + \dots + x_r = n$). Örneğin, $n=3$ ve $r=2$ olsun. Bu durumda, 3 aynı tür nesne, iki hücreye hiçbir hücrenin boş kalmaması koşulu ile $\binom{n-1}{r-1} = \binom{3-1}{2-1} = \binom{2}{1} = 2$ farklı şekilde dağıtılır.

A B A A → Birinci hücrede 1, ikinci hücrede 2 nesne,

$$(x_1=1, x_2=2, x_1 + x_2 = 3)$$

A A B A → Birinci hücrede 2, ikinci hücrede 1 nesne

$$(x_1=2, x_2=1, x_1 + x_2 = 3)$$

olduđuna gre, iki tr nesneye sahip iki deęerli dizilerde, 1.tr nesnelere tekrarı ile bařlayan toplam dzen sayısı, $\binom{n_1-1}{r_1-1} \binom{n_2-1}{r_2-1}$ farklı řekilde, 2.tr nesnelere tekrarı ile bařlayan toplam dzen sayısı, $\binom{n_2-1}{r_2-1} \binom{n_1-1}{r_1-1}$ farklı řekilde bulunur.

r_1 ve r_2 arasında $r_1 = r_2$ ya da $r_1 = r_2 \pm 1$ řeklinde bir iliřki vardır. Eęer, $r_1 = r_2 + 1$ ise dizi mutlaka 1. tr nesne ile, $r_1 = r_2 - 1$ ise dizi mutlaka 2. tr nesne ile bařlar. Eęer, $r_1 = r_2$ ise dizi her iki tr nesne ile de bařlayabilir.

Ortak daęılımın paydasını oluřturan $\binom{n_1+n_2}{n_1}$ ifadesi, $n = n_1 + n_2$ elemanlı bir dizi ierisinde n_1 adet "1" deęerinin ve n_2 adet "0" deęerinin yer aldıęı farklı diziliřlerin sayısını gstermektedir. O halde, $R_n^{(1)}$ ve $R_n^{(2)}$ nin ortak olasılık daęılımını (2.4) eřitlięindeki gibi elde edilebilir.

Teorem 2.2 $R_n^{(1)}$ ve $R_n^{(0)}$ nin marjinal olasılık daęılımları, Mood (1940) tarafından ařaęıdaki gibi tanımlanmıřtır.

$$f_{R_n^{(1)}}(r_1) = P(R_n^{(1)} = r_1) = \frac{\binom{n_1-1}{r_1-1} \binom{n_2+1}{r_1}}{\binom{n_1+n_2}{n_1}} \quad r_1 = 1, 2, \dots, n_1 \quad (2.5)$$

$$f_{R_n^{(2)}}(r_2) = P(R_n^{(0)} = r_1) = \frac{\binom{n_1+1}{r_2} \binom{n_2-1}{r_2-1}}{\binom{n_1+n_2}{n_1}} \quad r_2 = 1, 2, \dots, n_2 \quad (2.6)$$

İspat 2.2 $R_n^{(1)}$ in marjinal olasılık daęılımını, $R_n^{(1)}$ ve $R_n^{(0)}$ nin ortak olasılık daęılımınının r_2 zerinden toplamını alınarak bulunmaktadır.

$$f_{R_n^{(1)}}(r_1) = \sum_{r_2} f_{R_n^{(1)}, R_n^{(2)}}(r_1, r_2)$$

$$f_{R_n^{(1)}}(r_1) = \sum_{r_2} \frac{c \binom{n_1-1}{r_1-1} \binom{n_2-1}{r_2-1}}{\binom{n_1+n_2}{n_1}}$$

r_2 nin tanım aralıęında alacaęı deęerler r_1 , r_1-1 ve r_1+1 dir. Eęer, $r_2 = r_1$ ise $c=2$ deęerini, $r_2 = r_1-1$ ya da r_1+1 ise $c=1$ deęerini almaktadır. r_2 nin alacaęı  durum ve bu durumlarda c 'nin alacaęı deęerler sırasıyla yerine yazıldıęında, (2.5) eřitlięinde yer alan $R_n^{(1)}$ in marjinal olasılık daęılımını elde edilmiř olur. $R_n^{(2)}$

nin marjinal olasılık dağılımı ise $R_n^{(1)}$ ve $R_n^{(2)}$ nin ortak olasılık dağılımının r_1 üzerinden toplamı alınarak bulunmaktadır.

Teorem 2.3 R_n tekrar istatistiğinin olasılık dağılımı Mood (1940) tarafından aşağıdaki gibi tanımlanmıştır.

$$f_{R_n}(r) = \begin{cases} \frac{2 \binom{n_1-1}{\frac{r}{2}-1} \binom{n_2-1}{\frac{r}{2}-1}}{\binom{n_1+n_2}{n_1}} & , r \text{ çift ise} \\ \frac{\binom{n_1-1}{\frac{r-1}{2}} \binom{n_2-1}{\frac{r-3}{2}} + \binom{n_1-1}{\frac{r-3}{2}} \binom{n_2-1}{\frac{r-1}{2}}}{\binom{n_1+n_2}{n_1}} & , r \text{ tek ise} \end{cases} \quad (2.7)$$

$$r = 2, \dots, n_1 + n_2$$

İspat 2.3 R_n tekrar istatistiğinin olasılık dağılımı, $R_n^{(1)}$ ve $R_n^{(2)}$ nin ortak olasılık dağılımı kullanılarak bulunmaktadır.

$R_n^{(1)}$ ve $R_n^{(0)}$ nin ortak olasılık dağılımında, öncelikle, toplam tekrar sayısı olan r nin çift ya da tek olmasına göre c , r_1 ve r_2 nin alacağı değerler hesaplanmaktadır. r cinsinden elde edilen r_1 ve r_2 değerleri, $R_n^{(1)}$ ve $R_n^{(0)}$ nin ortak olasılık dağılımında yerine yazılarak R_n tekrar istatistiğinin olasılık dağılımı,

r çift ise,

$$f_{R_n}(r) = \frac{2 \binom{n_1-1}{\frac{r}{2}-1} \binom{n_2-1}{\frac{r}{2}-1}}{\binom{n_1+n_2}{n_1}}$$

r tek ise,

$$f_{R_n}(r) = \frac{\binom{n_1-1}{\frac{r+1}{2}-1} \binom{n_2-1}{\frac{r-1}{2}-1} + \binom{n_1-1}{\frac{r-1}{2}-1} \binom{n_2-1}{\frac{r+1}{2}-1}}{\binom{n_1+n_2}{n_1}}$$

şeklinde elde edilir.

R_n tekrar istatistiği ile ilgili dağılımlar verildikten sonra, $\xi_1, \xi_2, \dots, \xi_n$ şeklinde tanımlanan dizi elemanlarının birbirinden bağımsız ve aynı dağılıma

sahip olmaları varsayımı altında, uygulamalarda sıkça kullanılan $G_{n,k}$, S_n , L_n ve T_k tekrar istatistiklerinin de olasılık dağılımları verilmiştir.

n deneme sonundaki k ve daha büyük uzunluklu başarı tekrarı sayısını gösteren $G_{n,k}$ tekrar istatistiğinin olasılık dağılımı (2.8) eşitliği ile elde edilmektedir (Muselli, 1996).

$$f_{G_{n,k}}(x) = P(G_{n,k} = x)$$

$$= \sum_{l=kx}^a \sum_{j=0}^b (-1)^j \binom{n-l+1}{x} \binom{n-l+1-x}{j} \binom{n-k(x+j)}{n-1} p^l (1-p)^{n-l} \quad (2.8)$$

$$a = \min(n, n-x+1), \quad b = \min\left(\left\lfloor \frac{l-kx}{k} \right\rfloor, n-l-x+1\right)$$

n deneme sonundaki toplam başarı sayısını gösteren S_n tekrar istatistiğinin olasılık dağılımı (2.9) eşitliği ile elde edilmektedir (Balakrishnan and Koutras, 2002; Fu and Lou, 2003).

$$f_{S_n}(n_1) = P(S_n = n_1) = \binom{n}{n_1} p^{n_1} (1-p)^{n-n_1} \quad (2.9)$$

n deneme sonunda en uzun başarı tekrarının uzunluğunu gösteren L_n tekrar istatistiğinin olasılık dağılımı (2.10) eşitliği ile elde edilmektedir (Philippou and Makri, 1986; Godbole, 1990a; Muselli, 1996; Fu, 1986; Lou, 1996; Fu and Lou, 2003; Vaggelatos, 2003; Eryılmaz, 2005b).

$$f_{L_n}(k) = P(L_n < k)$$

$$= \sum_{l=0}^n \sum_{j=0}^{\min\left(\left\lfloor \frac{l}{k} \right\rfloor, n-l+1\right)} (-1)^j \binom{n-l+1}{j} \binom{n-jk}{n-l} p^l (1-p)^{n-l} \quad (2.10)$$

k uzunluklu bir başarı tekrarı elde edinceye kadar yapılan deneme sayısını gösteren T_k tekrar istatistiğinin olasılık dağılımı (2.11) eşitliği ile elde edilmektedir (Philippou and Muwafi, 1982).

$$x \leq 2k \text{ için} \quad f(x) = \begin{cases} 0, & 0 \leq x < k \\ p^k, & x = k \\ qp^k & k < x \leq 2k \end{cases} \quad (2.11)$$

$$x > 2k \text{ için} \quad f(x) = qp^k \left[1 - \sum_{i=0}^{x-k-1} f(i) \right]$$

3. TEKRAR İSTATİSTİKLERİNİN UYGULAMA ALANLARI

İki değerli rasgele dizilere dayalı olarak tanımlanan tekrar istatistikleri, rasgelelik testleri, hidroloji, istatistiksel kalite kontrol süreçleri, moleküler biyoloji ve güvenilirlik gibi birçok alanda karar verme süreçlerinde etkin bir yöntem olarak kullanılmaktadır.

Bu bölümde, tekrar istatistiklerinin uygulama alanları alt başlıklar altında incelenecektir.

3.1 Rasgelelik testleri

Tekrar istatistikleri ilk olarak elde edilen gözlemler dizisinin rasgeleliğini test etmek amacıyla kullanılmış ve literatüre parametrik olmayan testlerden “run” testi olarak geçmiştir.

İki farklı değer alabilen bir dizinin elemanlarının rasgele/homojen olduğu şeklindeki H_0 yokluk hipotezinin, elemanların rasgele/homojen olmadığı alternatifine karşı testi için tekrar istatistiklerini kullanmak mümkün olmaktadır. (Mood, 1940; Wald and Wolfowitz, 1940; Swed and Eisenhart, 1943). Bu amaçla ilgili hipotezler literatürde aşağıdaki gibi verilmiştir:

H_0 : Dizi elemanları rasgeledir/homojendir.

H_1 : Dizi elemanları rasgele değildir/homojen değildir.

Rasgelelik testlerinde en çok kullanılan tekrar istatistiği toplam tekrar sayısı olan R_n dir. Wald and Wolfowitz (1940), H_0 hipotezini test etmek için toplam tekrar sayısının olasılık dağılımını kullanmıştır. Bu test, literatüre “Wald-Wolfowitz Run testi” olarak geçmiştir. Swed and Eisenhart (1943) ise toplam tekrar sayısının olasılık dağılımını kullanarak, 20 den küçük tekrar sayıları için dağılım tablosu oluşturmuştur. Bu test, literatüre “Swed-Eisenhart Run testi” olarak geçmiştir (Lou, 1996).

Toplam tekrar sayısına dayalı rasgelelik testlerinde, R_n nin olasılık dağılımı kullanılarak p olasılık değerleri hesaplanır ve α değeri ile karşılaştırılarak H_0 yokluk hipotezi test edilir. Çok az ya da çok fazla sayıda tekrar, rasgeleliğin azlığını gösterir ve H_0 yokluk hipotezinin reddedilmesi anlamına gelir.

Güncel bir veri kümesi üzerinden örnek verilecek olursa, Devlet Meteoroloji İşleri Genel Müdürlüğü'nden (DMİ) elde edilen 1970-2010 yılları arası (41 yıllık) İzmir iline düşen toplam yağış miktarları (mm) Tablo 3.1'de verilmiştir. Sürekli bir rasgele değişken olan yıllık toplam yağış miktarları, ortanca değeri (672.3) kriter olarak kullanılarak iki değerli hale dönüştürülmüştür.

Tablo 3.1 İzmir İline düşen yıllık toplam yağış miktarları (mm)

Yıl	Yağış Miktarı (mm)		Yıl	Yağış Miktarı (mm)	
1970	627.2	0	1990	542.4	0
1971	725.3	1	1991	465.3	0
1972	339.3	0	1992	361.5	0
1973	624.2	0	1993	672.3	0
1974	700.0	1	1994	635.8	0
1975	714.0	1	1995	790.0	1
1976	872.7	1	1996	803.8	1
1977	468.0	0	1997	711.4	1
1978	895.5	1	1998	1086.1	1
1979	789.5	1	1999	652.5	0
1980	780.9	1	2000	554.0	0
1981	1056.2	1	2001	928.5	1
1982	661.1	0	2002	650.4	0
1983	549.4	0	2003	642.3	0
1984	685.1	1	2004	503.6	0
1985	526.8	0	2005	829.0	1
1986	706.3	1	2006	745.3	1
1987	746.7	1	2007	487.2	0
1988	627.5	0	2008	427.3	0
1989	519.4	0	2009	1071.9	1
			2010	1016.4	1

Tablo 3.1'de, örneklem büyüklüğü (n) 41 olan iki değerli dizide, "0" ile ifade edilen değerlerin sayısı $n_1=21$, "1" ile ifade edilen değerlerin sayısı $n_2 =20$ olarak elde edilmiştir.

Tablo 3.2'de, n_1 , n_2 ve toplam tekrar sayısı r nin tüm mümkün değerleri, (2.7) eşitliği ile belirtilen R_n tekrar istatistiğinin olasılık dağılımında yerine yazılarak $f_R(r)$ değerleri elde edilmiştir. Bu değerler kullanılarak $p = P(R \leq r)$ olasılık değerleri hesaplanmıştır. p olasılık değerleri α değeri ile karşılaştırılarak %5 hata payı ve %95 güvenlilikle verilerin rasgele olup olmadığına karar verilmiştir.

Tablo 3.2 R_n nin olasılık dağılımı ve p olasılık değerleri

n	n_1	n_2	r	$f_R(r)$	$p = P(R \leq r)$
41	21	20	2	0.0000000	0.0000000
41	21	20	3	0.0000000	0.0000000
41	21	20	4	0.0000000	0.0000000
41	21	20	5	0.0000000	0.0000000
41	21	20	6	0.0000002	0.0000002
41	21	20	7	0.0000014	0.0000016
41	21	20	8	0.0000082	0.0000099
41	21	20	9	0.0000339	0.0000437
41	21	20	10	0.0001396	0.0001833
41	21	20	11	0.0004326	0.0006159
41	21	20	12	0.0013397	0.0019556
41	21	20	13	0.0032377	0.0051933
41	21	20	14	0.0078151	0.0130084
41	21	20	15	0.0150720	0.0280804
41	21	20	16	0.0290276	0.0571080
41	21	20	17	0.0453556	0.1024635
41	21	20	18	0.0707547	0.1732182
41	21	20	19	0.0904087	0.2636270
41	21	20	20	0.1153039	0.3789309
41	21	20	21	0.1210691	0.5000000
41	21	20	22	0.1268343	0.5000000
41	21	20	23	0.1095387	0.3731657
41	21	20	24	0.0943396	0.2636270
41	21	20	25	0.0668239	0.1692874
41	21	20	26	0.0471698	0.1024635
41	21	20	27	0.0272133	0.0552938
41	21	20	28	0.0156302	0.0280804
41	21	20	29	0.0072569	0.0124502
41	21	20	30	0.0033493	0.0051933
41	21	20	31	0.0012281	0.0018440
41	21	20	32	0.0004466	0.0006159
41	21	20	33	0.0001256	0.0001693
41	21	20	34	0.0000349	0.0000437
41	21	20	35	0.0000072	0.0000088
41	21	20	36	0.0000014	0.0000016
41	21	20	37	0.0000002	0.0000002
41	21	20	38	0.0000000	0.0000000
41	21	20	39	0.0000000	0.0000000
41	21	20	40	0.0000000	0.0000000
41	21	20	41	0.0000000	0.0000000

Tablo 3.2 den elde edilen sonuçlara göre; ilgilenilen dizide maksimum 41 değerini alabilen toplam tekrar sayısı için 16'tan küçük toplam tekrar sayıları “çok az tekrar”; 27'den büyük toplam tekrar sayıları ise “çok fazla tekrar” anlamına

gelmekte ve her iki durum da rasgeleliğin azlığını göstermektedir. Dolayısıyla, toplam tekrar sayısı 16'tan küçük ve 27'den büyük ise "Dizi elemanları %5 hata payı ile rasgele değildir.", 16 ile 27 arasında ise "Dizi elemanları %5 hata payı ile rasgeledir." sonucu ortaya çıkmaktadır.

Sonuç olarak, Tablo 3.1'de yer alan iki değerli dizi için (2.3) eşitliğinden elde edilen toplam tekrar sayısı 18 olarak ve p olasılık değeri ise 0,1732182 ($p > \alpha = 0.05$) olarak hesaplandığından yokluk hipotezi reddedilemez ve İzmir İli'ne düşen yıllık toplam yağış miktarlarının %5 hata payı ve %95 güvenlilikle rasgele olduğu söylenebilir.

Yokluk hipotezinin test edilmesi için toplam tekrar sayısı kullanılabilirdiği gibi tekrar uzunlukları da kullanılabilir. Tekrar uzunluklarına dayalı rasgelelik testlerinde ise R_n nin olasılık dağılımına bakılmadan tekrar uzunluklarının varyanslarının ağırlıklandırılmış lineer kombinasyonu hesaplanır. Hesaplanan değer, $f = \vartheta_1 + \vartheta_2$ serbestlik derecesine sahip $\chi^2_{1-\alpha, f}$ tablo değeri ile karşılaştırılarak H_0 yokluk hipotezi test edilir (O'Brien and Dyck, 1985). Örneğin; "1" ile başarının "0" ile başarısızlığın temsil edildiği iki değerli bir dizide,

n_1 : Başarı sayısı

r_1 : Başarı tekrarlarının sayısı

s_1^2 : Başarı tekrar uzunluklarının varyansı

$$s_1^2 = \sum_{i=1}^{n_1} (x_i - \bar{x})^2$$

$$c_1 = (r_1^2 - 1)(r_1 + 2)(r_1 + 3)/(2r_1(n_1 - r_1 - 1)(n_1 + 1))$$

$$\vartheta_1 = c_1 n_1 (n_1 - r_1) / (r_1 (r_1 + 1))$$

n_2 : Başarısızlık sayısı

r_2 : Başarısızlık tekrarlarının sayısı

s_2^2 : Başarısızlık tekrar uzunluklarının varyansı

$$s_2^2 = \sum_{i=1}^{n_2} (x_i - \bar{x})^2$$

$$c_2 = (r_2^2 - 1)(r_2 + 2)(r_2 + 3)/(2r_2(n_2 - r_2 - 1)(n_2 + 1))$$

$$\vartheta_2 = c_2 n_2 (n_2 - r_2) / (r_2 (r_2 + 1))$$

olmak üzere, başarı ve başarısızlık tekrar uzunluklarının varyanslarının lineer kombinasyonu,

$$\chi^2_f = c_1 s_1^2 + c_2 s_2^2 \quad (\text{O'Brien and Dyck, 1985})$$

şeklinde ifade edilir. Burada, $c_1 s_1^2$ ve $c_2 s_2^2$ değerleri bağımsız ve ki-kare dağılımına sahiptir. Hesaplanan χ^2_f değeri, $f = \vartheta_1 + \vartheta_2$ serbestlik derecesine sahip $\chi^2_{1-\alpha, f}$ tablo değeri ile karşılaştırılarak H_0 hipotezinin kabul ya da red olduğuna karar verilir.

3.2 Hidroloji

Suyun yeryüzü, yeraltı ve atmosferdeki dağılımını, fiziksel ve kimyasal özelliklerini, çevreyle ve canlılarla ilişkilerini inceleyen temel ve uygulamalı bir bilim dalı olan hidroloji, tekrar istatistiklerinin en çok kullanıldığı alanlardan biridir (Salazar and Yevjevich (1975), Şen (1976), Şen (1990), Eryılmaz (2005b)).

Belirli bir bölgedeki kuraklık veya aşırı yağışlar gibi hidrolojik karakterli ekstrem doğa olayları ile ilgili kriter değerler belirlenerek tekrar istatistikleri yardımıyla çeşitli hidrolojik çıkarımlarda ve tahminlerde bulunmak mümkün olmaktadır.

Hidrolojide ekstrem doğa olayları hayati önem taşımaktadır. Özellikle, belli bir hidrolojik havzada sağanak yağışların olması ya da küresel ısınmadan dolayı kar-buz erimelerinin artması; o hidrolojik havzada yer alan akarsu debisinin normalin üstünde değerler almasına ve akarsu yatağının bu debiyi taşıyamamasından dolayı suyun taşarak (taşkın) canlılara, çevredeki arazilere ve yerleşim yerlerine zarar vermesine (sel) neden olabilmektedir.

Uşkay ve Aksu (2002), doğal afet olarak nitelendirilen sel ve taşkın olaylarını şöyle tanımlamıştır:

“Bir akarsuyun muhtelif nedenlerle yatağından taşarak, çevresindeki arazilere, yerleşim yerlerine, altyapı tesislerine ve canlılara zarar vermek suretiyle, etki bölgesinde normal sosyo-ekonomik faaliyeti kesintiye uğratacak ölçüde bir akış büyüklüğü oluşturması olayıdır.”

Hidrolojik olaylar çok sayıda değişkenden etkilendiği için tek bir değişkeni göz önünde bulundurarak doğal afetlerin gerçekleşeceğini söylemek doğru bir yaklaşım değildir. Örneğin, belli bir hidrolojik havzaya bir yıl içerisinde düşen toplam yağış miktarının çok fazla olması, o havza içerisinde sel ve taşkın gibi doğal afetlerin kesinlikle olacağı anlamına gelmemektedir. Önceki yıllarda yaşanan bir kuraklık söz konusu ise ilgili yılda çok fazla yağışın olması herhangi bir soruna yol açmayacaktır. Çünkü, yağın yağmurun büyük bir kısmının buharlaşması ve toprak tarafından emilmesi; kanallara yönelerek akarsu ve derelere akacak olan su miktarının azalmasına yol açacaktır.

Ayrıca, bir bölgeye düşen yağışın şiddeti ve süresi, yağış miktarından çok daha önemli bir bulgudur. Dolayısıyla; toprak cinsi, bitki örtüsü, zemin özelliği ve yüzey eğimi gibi hidrolojik havzanın coğrafi özelliklerini barındıran değişkenler ile yağış miktarı/şiddeti/süresi, sıcaklık, nem ve rutubet gibi iklim özelliklerini barındıran değişkenler, hidrolojik karakterli ekstrem olayların gerçekleşmesinde büyük rol oynamaktadır.

Taşkınlar, kıyı taşkınları, göl taşkınları, buz hareketi taşkınları ve sel taşkınları olarak çeşitlilik gösterir. Özellikle, aşırı yağışlar ve kar-buz erimeleri gibi birçok iklimsel faktörden etkilenen akarsu taşkınları, sel olayının en yaygın örneğidir (Uşkay ve Aksu, 2002).

Diğer değişkenlerin sabit olduğu varsayımı altında; bir değişkene ait geçmişte elde edilen veri kümesi kullanılarak tekrar istatistikleri yardımıyla mevcut durum belirlenebilir. Daha sonra, bu verilerin hangi dağılıma uyduğu tespit edilerek dağılım parametreleri ile birlikte olasılıklar hesaplanabilir. Elde edilen olasılıklar, tekrar istatistiklerinin dağılımları içerisinde kullanılarak ileriye yönelik tahminler ve planlamalar yapılabilir.

Örneğin, belirli bir periyotta (10 yıl) bir akarsudan geçen akım değerleri incelendiğinde akarsu debisinin üzerine çıkan değerlerin toplam sayısı hidrolojik açıdan önemli bir bilgi olabilir. İki değerli bir dizide “1” ile temsil edilen değerlerin toplam sayısı, S_n tekrar istatistiği ile ifade edilebilmektedir.

3.3 İstatistiksel süreç kontrolü

İstatistiksel süreç kontrolü, bir sürecin denetlenmesi ve kontrol altına alınmasında kullanılan bir kalite kontrol metodudur. Literatürde farklı tanımlamaları yapılan kalite, Garvin'e (1984) göre, ürün açısından bakıldığında, ürünün kullanım ömrü gibi fiyatlandırılmamış özelliklerin bütünü, tüketici açısından bakıldığında, ürünlerin tüketicilerin ihtiyaçlarını karşılayabilme kapasitesi, imalatçı açısından bakıldığında, imal edilen ürünün beklenen tasarıma uygunluk derecesi, “değer” açısından bakıldığında ise ürünün makul bir fiyattaki performans derecesi olarak tanımlanabilir (Kibritçioğlu, 1998).

İstatistiksel süreç kontrolünde “istatistik”, bir bütünün tamamının kontrol edilmesi yerine belli oranda örnekler çekilerek bütün hakkında tahminde bulunulmasını sağlayan araçları ifade eder. “Süreç”, bir ürün ya da hizmetin önceden belirtilen özellikleri sağlaması için kullanılan makine, yöntem, malzeme ve insan gücünün tamamını içerir. “Kontrol” ise süreçteki verilerin ölçümü ve analizinde istatistiksel yöntemlerin kullanılması anlamına gelmektedir (Durman ve Pakdil, 2012).

İstatistiksel süreç kontrolünde tekrar istatistiklerinin kullanılması ile ilgili ilk çalışma, Mosteller (1941) ve Wolfowitz (1943) tarafından yapılmıştır. Daha sonra, Weiler (1953), Kitagawa and Seguchi (1956, 1957), Moore (1958), Woodall (1983, 1984), Champ and Woodall (1987), Walker et al. (1991), Schmuelli and Cohen (2000), Chakraborti (2000) bu konudaki çalışmaları sürdürmüşlerdir.

İstatistiksel süreç kontrol aşamaları, veri toplama, sistem yeterlilik analizi, süreç yeterlilik analizi ve kalite kontrol kartlarından oluşmaktadır (Durman ve Pakdil, 2012).

Kalite Kontrol kartları, sürecin zamana göre değişimini gösteren, alt ve üst kontrol limitleri olan bir grafiktir. Bu grafik sayesinde sürecin, istatistiksel olarak kontrol altında olup olmadığı tespit edilebilir.

Bir kontrol kartının performansına, belirli özelliklerin tekrar uzunluklarına dayalı olarak karar vermek mümkündür. Performans ölçütü olarak tekrar uzunluğunun ortalaması (ARL-Average Run Length) kullanılmaktadır. Alt kontrol limiti (AKS) ve üst kontrol limiti (ÜKL) arasında değer alması beklenen alt grup örneklem ortalamasının (\bar{X}_l) AKS veya ÜKS dışına çıkması, sistemin sinyal vermesini sağlayarak sürecin kontrol altında olmadığını göstermektedir.

\bar{X} kontrol kartının alt ve üst limitleri $\mu_0 \pm a_n \sigma_0$ (μ ve σ^2 biliniyorsa) ile hesaplanmaktadır. Burada, $a_n = n^{-1/2} Z_{\alpha/2}$ olarak ifade edilir. O halde, sinyal verme olayının gerçekleşip gerçekleşmemesine bağlı olarak iki değerli dizi elemanları,

$$\xi_i = \begin{cases} 1, & \bar{X}_i \in (\mu_0 - a_n \sigma_0, \mu_0 + a_n \sigma_0) \\ 0, & \bar{X}_i \notin (\mu_0 - a_n \sigma_0, \mu_0 + a_n \sigma_0) \end{cases}$$

şeklinde tanımlanır.

Dolayısıyla, elemanları $\xi_1, \xi_2, \dots, \xi_n$ olan iki değerli denemelere dayalı bir dizide, k ardıl "1" (sinyal verme) değeri oluşuncaya kadar geçmesi gereken zamanı gösteren T_k tekrar istatistiği,

$$T_k = \min \left(t: \sum_{j=t-k+1}^t \xi_j = k \right)$$

şeklinde tanımlanır ve performans ölçütü olarak kullanılan ortalama tekrar uzunluğu $ARL = E(T_k)$ ile ifade edilir.

3.4 Moleküler biyoloji

Moleküler biyoloji, canlılardaki olayları moleküler düzeyde araştıran bir biyoloji dalıdır. Özellikle hayati önem taşıyan nükleik asitler yani DNA'lar moleküler biyolojinin en önemli araştırma konularından birisidir. Tekrar istatistikleri de moleküler biyolojide modelleme ve sıralı eşleşmeler ile ilgili çalışmalarda kullanılmaktadır (Chvátal and Sankoff (1975), Waterman (1982, 2000), Arratia ve Waterman (1985), Mott et al., (1990), Goldstein and Waterman (1992), Fu et al. (1999)).

DNA zinciri, nükleotid adı verilen dört farklı yapıdan oluşmaktadır. Adenin(A), Guanin(G), Timin(T) ve Sitozin(S) olarak adlandırılan nükleotidler özel bir sıra ile dizilerek DNA zincirini oluşturmaktadır. İki DNA zincirinin karşılıklı eşleşmesi ile DNA sarmalı meydana gelmektedir (Clark, 2007). DNA nükleotidleri zaman içerisinde mutasyona uğradıkları için DNA zincirinin elemanları arasında belirli bir sıra olduğunu söylemek mümkün değildir. Dolayısıyla, DNA zincirinin elemanları rasgele değişkenlerdir.

Moleküler biyolojide, “iki DNA zinciri arasında benzerlik yoktur” şeklindeki H_0 yokluk hipotezinin, “zincirler arası benzerlik olduğu” alternatifine karşı testi için tekrar istatistiklerini kullanmak mümkün olmaktadır. Örneğin, iki bireyin DNA dizilişleri,

$$1. \text{ birey için } v_1^1, v_2^1, \dots, v_n^1 \in \{A, G, C, T\}$$

$$2. \text{ birey için } v_1^2, v_2^2, \dots, v_n^2 \in \{A, G, C, T\}$$

şeklinde gösterilsin. Bu iki diziliş arasında yapılan eşleşmeye bağlı olarak aşağıdaki gibi tanımlanan

$$\xi_i = \begin{cases} 1, & v_i^1 = v_i^2 \\ 0, & v_i^1 \neq v_i^2 \end{cases} \quad i = 1, 2, \dots, n$$

değişkeni, elemanları $\xi_1, \xi_2, \dots, \xi_n$ olan iki değerli rasgele bir diziyi oluşturmaktadır. Dolayısıyla bu rasgele dizi üzerinde, n deneme sonunda en uzun başarı tekrarının uzunluğunu gösteren L_n tekrar istatistiği yardımıyla yokluk hipotezi test edilebilir. Verilen bir α değeri için, $P(L_n \geq m | H_0) \leq \alpha$ koşulunu gerçekleştiren m değeri için, $L_n \geq m$ ise yokluk hipotezi reddedilir.

3.5 Güvenilirlik

Ardıl n -den- k -çıkışlı F sistemi en önemli güvenilirlik sistemlerinden birisidir. Bu sistemle ilgili ilk çalışma Chiang ve Niu (1981) tarafından yapılmıştır. Daha sonra, Papastavridis and Koutras (1994), Chao et al. (1995), Koutras et al. (1995, 1996), Aki and Hirano (1996), Malinowski and Preuss (1996), Preuss (1997), Dao and Silio (1998), Aki (1999), Boland and Papastavridis (1999), Lam and Zhang (1999, 2000), Muselli (2000a), Lam and Ng (2001) bu konudaki çalışmalarını sürdürmüşlerdir.

Tekrar istatistiklerinden biri olan en uzun tekrar istatistiği ile yakından ilişkili ve güvenilirlik sistemlerinden biri olan ardıl “n–den–k–çıkışlı F sistemi” terimi ilk olarak Chiang ve Niu (1981) tarafından kullanılmış, ardından bu konu ile ilgili çok sayıda çalışma yapılmıştır.

Ardıl n–den–k–çıkışlı F sistemi, Z bileşen kümesi ve $\phi: Z \rightarrow \{0,1\}$ yapı fonksiyonundan oluşan $\Sigma = (Z, \phi)$ sistemindeki bileşenlerin, herhangi bir anda çalışır (1) ve çalışmaz (0) durumlarında olduğu varsayılarak, k ya da daha fazla bileşenin ardıl olarak çalışması durumunda $\phi(X) = 1$ yapı fonksiyonu ile tanımlanan bir sistemdir. Bu sistemin güvenilirliğinin ölçülmesi için tekrar istatistikleri kullanılmaktadır.

Birbirinden bağımsız, p olasılığı ile çalışır, $q = 1 - p$ olasılığı ile çalışmaz durumda olan bileşenlerin oluşturduğu ardıl n–den–k–çıkışlı F sisteminin güvenilirliği tekrar istatistikleri kullanılarak ölçülebilir.

n deneme sonunda en uzun başarı tekrarının uzunluğunu gösteren L_n tekrar istatistiği yardımıyla ardıl n–den–k–çıkışlı bir sistemin güvenilirliği, $E\phi(X) = P\{L_n \geq k\}$ olarak tanımlanmaktadır.

4. ÖNERİLEN MODEL

Literatürde var olan klasik modele göre, elemanları X_1, X_2, \dots, X_n olan sürekli değerli bir dizi, x_0 kriter değeri (eşik değeri) kullanılarak, elemanları $\xi_1, \xi_2, \dots, \xi_n$ olan iki değerli rasgele bir diziyeye dönüştürülebilir. Kriter değer olarak genellikle ortalama veya ortanca gibi veri kümesini özetleyen konum ölçülerinin kullanıldığı modelde, elemanları $\xi_1, \xi_2, \dots, \xi_n$ olan ξ_i ($i = 1, 2, \dots, n$) rasgele değişkeni,

$$\xi_i = \begin{cases} 1, & X_i > x_0 \\ 0, & X_i \leq x_0 \end{cases}$$

şeklinde tanımlanmaktadır.

Bu bakış açısına göre, x_0 değerinden çok az büyük bir X_i değeri için ξ_i ($i = 1, 2, \dots, n$) rasgele değişkeni “1” değerini alırken, x_0 değerinden çok az küçük bir X_i değeri için de “0” değerini almaktadır. Oysaki, her iki X_i değeri arasında mutlak değer olarak çok fazla bir fark bulunmamasına rağmen farklı sonuçların ortaya çıkması, özellikle ileriye dönük tahminlerde yanlış yönlendirmelere neden olabilir.

Bu nedenle, x_0 gibi sabit bir eşik değerinin belirlenmesi yerine, “0.0” ile “1.0” arasında değer alabilen bir γ değişkeni tanımlanarak, karar vericinin tercihine göre değişebilen bir kriter değerinin belirlenmesi modele esneklik kazandıracaktır.

γ değişkeninin “0.0” ile “1.0” arasında alacağı her bir değer, “risk düzeyi” olarak adlandırılmaktadır. Düzeyler arası aralık genişliğinin 0.1 olarak belirlendiği bir modelde, γ değişkeni onbir farklı değer almaktadır. Her bir risk düzeyine bağlı olarak kriter değerleri belirlenerek, iki değerli rasgele diziler elde edilmekte ve karar vericiye sistemle ilgili tercih yapabilmesi için farklı alternatifler sunulmaktadır.

Eğer, karar verici, ilgili sistemin yaratacağı sonuçların kritik olduğunu düşünen pesimist bir gözlemci ise kriter değeri, riskin daha az olduğu düzeylerde (0.0’a yakın) belirlemeyi tercih edecektir. Yine, karar verici eldeki kısıtlı kaynakları göz önüne alarak maliyeti düşünen ve risk alabilen bir gözlemci ise kriter değeri, riskin daha yüksek olduğu düzeylerde (1.0’e yakın) belirleyecektir.

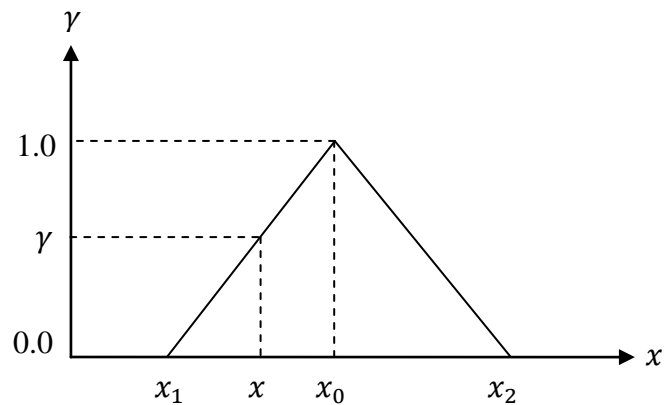
0.0 ile 1.0 arasında değer alabilen γ değişkenine göre tanımlanan model, esasında bulanık küme mantığına dayanmaktadır. Dolayısıyla, önerilen model ile klasik modelin karşılaştırılabilmesi için öncelikle geleneksel küme ve bulanık küme arasındaki farkın çok iyi anlaşılması gerekmektedir.

Geleneksel küme teorisinde, x değeri kümenin elemanı ise “1” değerini, değilse “0” değerini almaktadır. Örneğin, boyu 1.85 cm’den uzun olanların “uzun boylu” olarak tanımlandığı bir kümede, boyu 1.84 cm olan bir kişi “kısa boylu” olmamasına rağmen “uzun boylu” kümesinin elemanı olarak tanımlanmaz. Bu sorun, kümeye üye olan ve olmayanlar arasında keskin bir ayırım yapılmasından kaynaklanmaktadır (Eğrioğlu, 2010). Diğer bir ifade ile, küme üyeliği arasındaki geçiş; “0”dan “1”e, “1”den 0’a olmak üzere esnek olmayan bir durumu içermektedir.

Oysaki, bulanık küme teorisinde, kümenin elemanı olmayan değerlerden, kümenin tam elemanı olan değerlere doğru “esnek” ve “dereceli” bir geçişe izin verilmektedir (Özkan, 2003). Dolayısıyla, boyu 1.84 cm olan bir kişinin “kısa boylu” olmadığı ve uzun boylular kümesine belirli bir üyelik derecesi (örneğin 0.9) ile kısmi üye olduğu söylenebilir.

Belirli bir üyelik derecesi ile bulanık kümenin elemanı olan nesnelere, bulanık sayılarla ifade edilir. Bulanık sayılar, bulanık kümelerin özel bir alt kümesidir (Özkan, 2003). Uygulamalarda en sık kullanılan bulanık sayı türü üçgensel bulanık sayılardır. Üçgensel bir bulanık sayı, Şekil 4.1’deki gibi gösterilebilir.

Şekil 4.1 Üçgensel bulanık sayı



Şekil 4.1’de x_1 değerinden x_0 değerine kadar monoton artan bir fonksiyon, x_0 değerinden x_2 değerine kadar monoton azalan bir fonksiyon söz konusudur. x_0 noktası bulanık sayının merkezi olmak üzere, x_1 ve x_2 değerleri 0.0 üyelik derecesine (üye değil), x_0 değeri ise 1.0 üyelik derecesine (tam üye) sahiptir. x_1 ve x_2 değerlerinin arasında kalan bir x değeri ise kısmi üyelik derecesine sahiptir.

Üçgensel bulanık sayının üyelik fonksiyonu ise aşağıdaki gibi tanımlanmaktadır.

$$\gamma = \begin{cases} 0, & x < x_1 \\ \frac{x - x_1}{x_0 - x_1}, & x_1 \leq x \leq x_0 \\ \frac{x_2 - x}{x_2 - x_0}, & x_0 \leq x \leq x_2 \\ 0, & x \geq x_2 \end{cases} \quad (4.1)$$

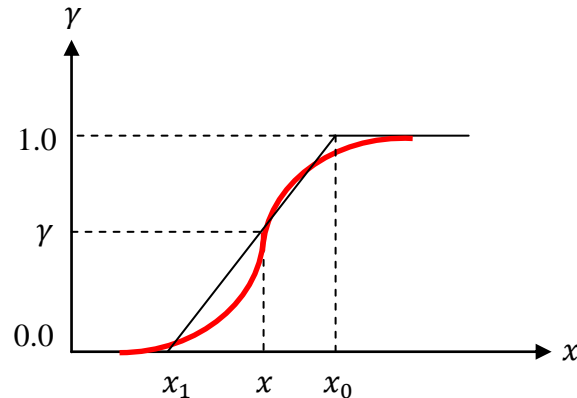
(4.1) eşitliğindeki üyelik fonksiyonu kullanılarak, fonksiyonun artan kısmında tanımlanan bir x değeri,

$$\gamma = \frac{x - x_1}{x_0 - x_1} \text{ ise } x = x_1 + (x_0 - x_1)\gamma \quad (4.2)$$

eşitliği ile hesaplanabilir.

O halde; bulanık küme teorisinde “üyelik derecesi” olarak tanımlanan γ değişkeni, önerilen modelde “risk düzeyi” olarak tanımlandığında, risk düzeyi ile ilgilenilen rasgele değişkenin aldığı değerler (x) arasındaki grafiksel ilişki Şekil 4.2’deki gibi gösterilebilir. İki değişken arasındaki ilişki eğrisel olabildiği gibi doğrusal da olabilir. Bu çalışmada kolaylık sağlaması açısından doğrusal bir ilişki göz önüne alınacaktır.

Şekil 4.2 Risk düzeyi ile rasgele değişken arasındaki ilişki



Şekil 4.2’de, x_1 değeri riskin en az olduğu 0.0 düzeyine, x_0 değeri ise riskin en yüksek olduğu 1.0 düzeyine karşılık gelmektedir. x_0 değeri en üst risk düzeyini temsil ettiği için bu değerden sonra gelebilecek her değer için grafik, x eksenine paralel olarak ilerleyecektir. Dolayısıyla, üçgensel bulanık sayıdaki gibi azalan bir fonksiyon söz konusu değildir. x_1 ile x_0 arasında belirlenen bir x değeri için göze alınan (yüklenilen) risk, x_0 değerine doğru artarken, x_1 değerine doğru azalmaktadır.

Risk düzeyinin üyelik fonksiyonu aşağıdaki gibi tanımlanabilir.

$$\gamma = \begin{cases} 0, & x < x_1 \\ \frac{x - x_1}{x_0 - x_1}, & x_1 \leq x < x_0 \\ 1, & x \geq x_0 \end{cases} \quad (4.3)$$

(4.3) eşitliğindeki üyelik fonksiyonu kullanılarak, her bir risk düzeyine göre x kriter değeri;

$$\gamma = \frac{x - x_1}{x_0 - x_1} \text{ ise } x = x_1 + (x_0 - x_1)\gamma \quad (4.4)$$

eşitliği ile hesaplanabilir.

O halde, sürekli değerli bir diziyi iki değerli hale dönüştürmek için kullanılacak olan kriter değer, klasik modelde ortalama ya da ortanca gibi sabit bir x_0 değeri iken, önerilen modelde γ değişkenine bağlı olarak değişen $x_1 + (x_0 - x_1)\gamma$ değeri olacaktır.

$\gamma = 0$ olması durumunda kriter değer x_1 değerine, $\gamma = 1$ olması durumunda ise kriter değer x_0 değerine (klasik durum) karşılık gelmektedir. x_0 ve x_1 değerleri, uzman görüşü yanında veri kümesinin özelliğine göre belirlenmelidir.

γ değişkenine bağlı olarak tanımlanan modelde, X_1, X_2, \dots, X_n sürekli değerlerden oluşan rasgele bir dizi ve $x_1 + (x_0 - x_1)\gamma$ kriter değer olmak üzere, elemanları $\xi_1^\gamma, \xi_2^\gamma, \dots, \xi_n^\gamma$ olan $\xi_i^\gamma (i = 1, 2, \dots, n)$ rasgele değişkenini,

$$\xi_i^\gamma = \begin{cases} 1, & x_i > x_1 + (x_0 - x_1)\gamma \\ 0, & x_i \leq x_1 + (x_0 - x_1)\gamma \end{cases} \quad (4.5)$$

fonksiyonu ile tanımlamak mümkün olacak ve ξ_i^γ rasgele değişkeninin alacağı değerlere ilişkin olasılıklar,

$$P(\xi_i^\gamma = 1) = P(x_i > x_1 + (x_0 - x_1)\gamma) \quad (4.6)$$

$$P(\xi_i^\gamma = 0) = P(x_i \leq x_1 + (x_0 - x_1)\gamma)$$

eşitlikleri ile elde edilebilecektir.

(4.6) eşitliğinde verilen $p^\gamma = P(\xi_i^\gamma = 1) = P(X_i > x_1 + (x_0 - x_1)\gamma)$ olasılık ifadesi “aşma olasılığı” olarak adlandırılmakta ve γ değişkenine bağlı olarak değişen kriter değerlerin üzerinde yer alan gözlemlerin oranını göstermektedir.

Son olarak; bulanık mantığın, önerilen modele getirdiği esneklik ve avantajların daha iyi anlaşılabilmesi bakımından, klasik ve önerilen modeller için “bulanık risk kümesi” oluşturulmuştur.

Klasik model için \check{A}_1 , önerilen model için \check{A}_2 ile gösterilen bulanık risk kümeleri, riski çok olan değerlerin kümesi olarak düşünölsün. Bu durumda, \check{A}_1 ve \check{A}_2 bulanık kümeleri aşağıdaki gibi tanımlanabilir.

$$\check{A}_1 = \left\{ \frac{1.0}{x_0} \right\} \quad (4.7)$$

$$\check{A}_2 = \left\{ \frac{0.0}{x_1} + \frac{0.1}{x_1 + 0.1(x_0 - x_1)} + \dots + \frac{0.9}{x_1 + 0.9(x_0 - x_1)} + \frac{1.0}{x_0} \right\}$$

Klasik modelde, veri kümesi tek bir kriter değeri ile karşılaştırıldığı için en yüksek risk değerine sahip olan x_0 kriter değeri, \check{A}_1 risk kümesinin “kesin üyesi”dir.

Önerilen modelde ise en düşük risk değerine sahip olan x_1 kriter değeri, \check{A}_2 risk kümesinin üyesi değil iken, en yüksek risk değerine sahip olan x_0 kriter değeri \check{A}_2 risk kümesinin “kesin üyesi”dir.

$x_1 + 0,1(x_0 - x_1)$ değeri ise 0.1 üyelik derecesi ile \check{A}_2 risk kümesinin “kısmi üyesi”dir. Diğer kriter değerler de, bulanık kümenin γ üyelik derecesi ile kısmi üyeleridir.

Önerilen modelde, bulanıklık, yalnızca risk düzeyi için uygulanmış ve böylelikle her bir düzeye karşılık gelen kriter değerler derecelendirilmiştir. Belli bir “üyelik derecesi” ile risk kümesinin üyesi olan kriter değerler sayesinde, sürekli değerli dizi, yine iki değerli hale dönüştürülmüştür. Dolayısıyla, veri kümesinin bulanık hale çevrilmesi gibi bir durum söz konusu değildir.

Ancak; klasik modelde, “sorun var” ifadesinin “1” ile, “sorun yok” ifadesinin ise “0” ile temsil edildiği iki değerli rasgele bir dizi için keskin ifadeler söz konusuyken, önerilen modelde, özellikle “1” değeri için “fazla sorun olmaz”, “biraz sorun olabilir”, “sorun olma olasılığı yüksek, önlem alınmalı” gibi ifadeler kullanılarak veri kümesinin özelliğine göre sözel bir derecelendirme yapılabilir.

Bu tarz bir yaklaşım, iki değerli rasgele diziler üzerinde tanımlanabilen tekrar istatistiklerinin yorumlanmasında büyük bir kolaylık sağlayacaktır.

5. UYGULAMA

Bu bölümde, bir önceki bölümde açıklanan ve “önerilen model” olarak ifade edilen modeli örneklendirmek amacıyla **hidroloji** alanında bir uygulama yapılmıştır. Öncelikle, hidrolojik olaylar ve bu olaylara neden olan değişkenler ile ilgili bir ön bilgi verilerek adım adım uygulama aşamaları açıklanmıştır.

Türkiye’de toplam 26 hidrolojik havza (su toplama havzası) bulunmaktadır. Bu havzalarda yer alan meteoroloji istasyonları ve akım gözlem istasyonları (AGİ) tarafından sıcaklık ve yağış gibi **meteorolojik veriler** elde edilirken; barajlardaki su seviyesi, dere ve akarsu yataklarından saniyede geçen akım miktarları gibi **hidrolojik veriler** de derlenebilmektedir.

Hidrolojik olaylar çok sayıda değişkenden etkilenmektedir. Toprak cinsi, bitki örtüsü, zemin özelliği, yüzey eğimi gibi hidrolojik havzanın coğrafi özelliklerini barındıran değişkenler ile yağış miktarı/şiddeti/süresi, kar/buz erimesi, sıcaklık, nem, rutubet gibi iklim özelliklerini barındıran değişkenler, sel ve taşkın olaylarının gerçekleşmesinde büyük rol oynamaktadır.

Özellikle; aşırı ve şiddetli yağışlar, kar-buz erimeleri gibi iklimsel faktörlerin etkisiyle ortaya çıkan sel ve taşkın olayları, telafisi mümkün olmayan sorunlara yol açabilmektedir. Akarsu yatağının normalin üzerinde değer alan debiyi taşıyamamasından dolayı suyun taşması olarak adlandırılan **akarsu taşkınları**, sel olayının en yaygın örneğidir (Uşkay ve Aksu, 2002).

Sel ve taşkın olayları rasgele karakterli hidrolojik olaylardır (Onüçyıldız, 1999). Çünkü, akarsu kesitinden geçen akım miktarları, bir ya da birden çok değişkenin etkisiyle akarsu yatağının taşıyamayacağı seviyeye ulaşabilmektedir. Bu değişkenlerin birbirini nasıl etkilediğinin ya da hangi periyotta hangi debi değeri ile taşkına neden olduğunun önceden bilinmesi de mümkün değildir. Örneğin, bir akarsu üzerinde, önceki yıllarda bir sorun yaşanmamışken, bir yıl içerisinde -farklı değişkenlerin etkisiyle- birden çok taşkın meydana gelebilir. Bu nedenle, taşkınlar rasgele olaylar olarak kabul edilirler (Büyükkaracıgan ve Kahya, 2007, 2009).

Akarsu taşkınlarının önceden tahmin edilebilmesi, olası bir felaketin önlenmesi bakımından çok önemlidir. Ancak, bu tahminin yapılabilmesi için akarsu debisini etkileyen rasgele değişkenlerin hepsinin bir arada incelenmesi

mümkün değildir. Bu nedenle, diğer değişkenlerin sabit tutulduğu varsayımı altında belli bir periyotta tek bir değişkene ait veri kümesi incelenerek istatistiksel yöntemlerle ileriye dönük tahminlerde bulunulabilir.

Hidrolojik olaylar, dünyanın güneş etrafında bir yıllık sürede (yıllık dolanım) dönmesiyle tekrarlanmaktadır. Bu yüzden kullanılacak olan rasgele değişkenin birim zamanı, “bir yıllık periyot” olarak alınabilir (Büyükkaracıoğlu ve Kahya, 2007, 2009).

Akarsu taşkınlarının en çok yağışlara bağlı olarak gerçekleştiği düşünülürse; toplam/maksimum yağış miktarları yıllık periyotta incelenebilir. Ancak, akarsu yatağında, yağış dışında başka bir değişkenden dolayı taşkın meydana gelirse, yağış miktarları taşkın olayını açıklamada tek başına yetersiz kalacaktır. Aynı durum diğer değişkenler için de geçerlidir.

Taşkına neden olan değişkenlerin yerine, olayın kaynağına inilerek, akarsu en kesitinden geçen akım miktarlarının incelenmesi düşünülebilir. Yıllık periyotta gözlemlenen maksimum akım değerlerinin -hangi değişkenin etkisiyle olursa olsun- akarsu yatağının taşıyamayacağı seviyeye ulaşması, ilgili yılda taşkın meydana geleceğini göstermektedir. Bu nedenle, yıllık anlık maksimum akım değerleri, hidroloji alanında önemli bir rasgele değişkendir.

Yıllık toplam/maksimum yağış verileri gibi iklimsel veriler ile akarsu en kesitinde ölçülebilen yıllık anlık maksimum akım değerleri, sürekli değerler alan rasgele değişkenlerdir. Bu değişkenlere ait veri kümeleri üzerinde kriter değerler kullanılarak iki değerli diziler elde edilmesi ve bu diziler üzerinde tekrar istatistiklerinin tanımlanarak hem mevcut veri kümesinin analizinin yapılması, hem de ileriye dönük tahminlerde bulunulması mümkün olmaktadır.

Tez çalışmasının uygulama aşamasında ilk olarak kullanılacak olan rasgele değişken ve bu değişkenin elde edileceği havzalar belirlenmiştir. Daha sonra seçilen değişkene ait veri kümesi üzerinde her bir havza için modelleme yapılarak, klasik model ve önerilen model arasındaki fark gösterilmeye çalışılmıştır. Son olarak da, tekrar istatistiklerinin olasılık dağılımları kullanılarak ileriye dönük tahminler yapılmış, akabinde havzaların birbiri ile kıyaslanmasına olanak sağlanmıştır.

5.1 Değişken ve Havza Seçimi

Değişken belirleme sürecinde, öncelikle -diğer değişkenlerin sabit olduğu varsayımı altında- sel ve taşkın olaylarının en çok yağış miktarından etkilendiği göz önüne alınmış ve Yıllık Toplam Yağış (mm) verileri incelenmiştir.

Ancak, yağışın şiddeti ve süresinin, yağış miktarından çok daha önemli bir bulgu olduğu anlaşılmış ve çok fazla yağın yağmurun -önceki yıllarda bir kuraklık söz konusu olduğunda- her zaman taşkına neden olmadığı görülmüştür.

Uzman görüşü doğrultusunda, tek bir değişkene bağlı kalarak sel ve taşkın olayları hakkında sağlıklı bir karar verilemeyeceği sonucuna ulaşılmış ve akarsu yataklarından (özellikle anakol üzerindeki) saniyede geçen su miktarlarının (m^3/s) incelenmesinin -olayın kaynağından çözülebilmesi bakımından- daha doğru bir yaklaşım olduğuna karar verilmiştir.

Bu nedenle, DSI'ye ait Akım Gözlem İstasyonları (AGİ) tarafından elde edilen ve bir yıl içerisinde akarsu yatağından geçen maksimum su miktarını veren Yıllık Anlık Maksimum Akım (YAMA) değerleri (m^3/s), incelenmesi gereken “**değişken**” olarak belirlenmiştir.

YAMA değerleri gibi hidrolojik zaman serilerinde istatistiksel analizlerin yapılabilmesi için verilerin homojen ve birbirinden bağımsız yani rasgele olması gerekir (Toros, 1993; Onüçyıldız, 1999; Büyükkaracığan ve Kahya, 2007, 2009; Em vd., 2007).

Homojenlik, veri kümesindeki tüm gözlemlerin aynı kitleden gelmesi yani aynı özellikleri göstermesi anlamındadır. Diğer bir ifade ile, YAMA değerlerinin elde edildiği akarsu yatağının ve üzerindeki akım gözlem istasyonunun yapısının bozulmaması gerekir. Örneğin, akarsu ya da dere yatağının etrafında, istasyon tarafından ölçülecek su miktarını direkt olarak etkileyecek, -şehirleşme başta olmak üzere- baraj, köprü gibi herhangi bir yapılaşma olmamasına ve arazi kullanım şeklinin değişmemesine dikkat edilmelidir. Ayrıca, Haan'a (1977) göre mevcut verilerin gelecek yıllardaki akımları temsil edebilmesi için tahmin yapılmak istenen yıllarda da veri kümesinin homojenlik durumunun muhafaza ediliyor olması gerekmektedir (Onüçyıldız, 1999; Büyükkaracığan ve Kahya, 2007, 2009).

Bağımsızlık özelliği ise doğası gereği rasgele karaktere sahip olan taşkın olaylarının sürekli olarak tek bir değişkenin etkisinde kalmaması anlamında kullanılır. Zaten aksi bir durum, veri kümesinde bağımlılık sorununu ortaya çıkaracaktır.

Bu bilgiler doğrultusunda, konunun uzmanı ile birlikte, homojenlik ve bağımsızlık özelliğine sahip bir akarsuyun anakol üzerindeki Akım Gözlem İstasyonuna ait YAMA değerlerinin incelenmesine karar verilmiştir.

Önerilen modele uygun şekilde kriter değerler belirlenerek, YAMA değerlerinden oluşan sürekli değerli dizi, iki değerli diziye dönüştürülmüş ve tekrar istatistikleri kullanılarak bazı hesaplamalar yapılmıştır. Modelden elde edilen sonuçlar, akarsuya ait gerçek taşkın verileri ile karşılaştırılmıştır. Ancak, Taşkın Yıllıkları incelendiğinde, akarsuda çok az sayıda taşkın meydana geldiği görülmüştür. Çok sayıda taşkın olması beklenen anakol üzerinde bu denli az taşkın yaşanması, yan kollardaki taşkınların daha da az olacağı, hatta hiç olmayacağı anlamına gelmektedir. Zaten, taşkın yaşanan bir havzada, -taşkın neden olan akarsuyun yatağının sedde alınması gibi- bir takım önlemlerin alınması, o akarsu üzerinde uzun bir süre taşkın olmayacağının bir göstergesidir. Dolayısıyla, modelden elde edilen sonuçlar, tekrar istatistikleri kullanılarak ileriye dönük tahmin ve çıkarımlar yapmak için yeterli olmamıştır.

Bu nedenle, tek bir AGİ'den elde edilen YAMA değerleri yerine, havzayı temsil edebilecek akarsular üzerindeki istasyonlardan elde edilen YAMA değerlerinin maksimumlarının kullanılmasına karar verilmiştir. Maksimum YAMA değerleri, MYAMA (m^3/s) değişkeni ile gösterilmiştir.

Aynı havza içerisinde (aynı kitlede) yer alan akarsulardaki akım miktarları, havzanın genel özelliklerinden (toprak cinsi, su tutma özelliği, yağış miktarı, vb.) etkileneceği için -dış bir etken olmadığı sürece- homojen bir yapı gösterecektir. Dolayısıyla, MYAMA değerlerinin de bu özelliğe sahip olması beklenir.

Bu bağlamda, konunun uzmanı ile birlikte, sel ve taşkın olaylarının yoğun olarak yaşandığı **Kızılırmak** ve **Sakarya** havzaları ile sel ve taşkın olaylarının nadir olarak görüldüğü **Küçük Menderes** ve **Gediz** havzaları, modellemede kullanılacak havzalar olarak belirlenmiştir. Bu havzalarda yer alan AGİ'ler içerisinde homojenlik özelliği bozulmayan ve gelecekte de aynı özelliği

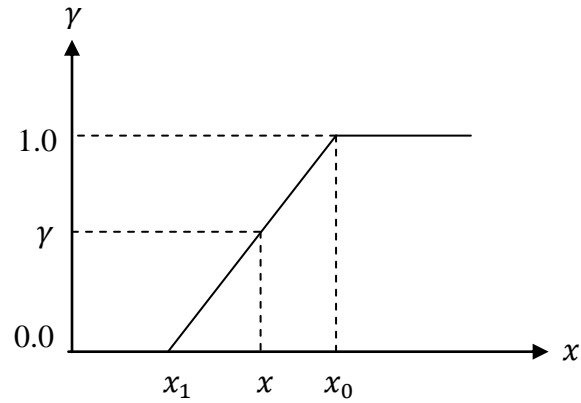
taşıyabilecek nitelikteki istasyonlara ait verilerin kullanılmasına da özen gösterilmiştir.

Belirlenen havzalar için MYAMA değerlerinden oluşan dört ayrı veri kümesi oluşturulmuş ve bu veri kümelerinin de homojen bir yapıya sahip olduğu “Run” testi kullanılarak, Bölüm 5.3.1 - Bölüm 5.3.4’te gösterilmiştir.

5.2 Modelleme

X sürekli rasgele değişkeni, bir havzaya ait MYAMA(m^3/s) değerlerini; “0.0” ile “1.0” arasında değer alan γ değişkeni ise havzadaki taşkın risk düzeyini göstermek üzere, X (MYAMA) ve γ (taşkın risk düzeyi) değişkenleri arasındaki doğrusal ilişki aşağıdaki şekilde tanımlanmıştır.

Şekil 5.1 MYAMA ve taşkın risk düzeyi arasındaki doğrusal ilişki



Şekil 5.1’de, ilgilenilen havzadaki maksimum eşik değeri olan x_0 değeri taşkın riskinin en çok olduğu “1.0” düzeyine, minimum eşik değeri olan x_1 değeri ise taşkın riskinin en az olduğu “0.0” düzeyine karşılık gelmektedir. Minimum eşik değerinin altındaki bir MYAMA değeri için havzada taşkın olma ihtimali yok denecek kadar az iken, maksimum eşik değerinin üstündeki bir MYAMA değeri için havzada taşkın olma ihtimali oldukça yüksektir. x_1 ile x_0 arasındaki bir MYAMA değeri için havzada taşkın olma riski x_0 değerine doğru artarken, x_1 değerine doğru azalmaktadır.

x_0 ve x_1 değerleri, uzman görüşü ile birlikte, havzanın coğrafi konumu, iklimi, yağış miktarı, su tutma özelliği ve taşkın olma sıklığı gibi birçok faktör göz önüne alınarak belirlenmiştir. Özellikle, havzanın geneline uygun modelin oluşturulabilmesi için MYAMA değerleri, gerçek taşkın verileri ile

karşılaştırılmıştır. Bunun için öncelikle, taşkın yıllıkları incelenerek, ilgilenilen havzada hangi yıllarda taşkın olayının gerçekleştiği tespit edilmiş ve tespit edilen yıllar veri kümesi üzerinde işaretlendikten sonra MYAMA değerleri küçükten büyüğe doğru sıralanarak taşkın olan yılların veri kümesinin hangi yüzdelik dilimlerinde yoğunlaştığına bakılmıştır. x_0 ve x_1 değerleri de bu yüzdelik dilimler arasından seçilmiştir. İlgilenilen sistemin yapısına göre eşik değerlerinin doğru olarak tespit edilebilmesi, ileriye dönük tahminlerin gerçekçi ve tutarlı olması bakımından çok önemlidir. Dolayısıyla, ilgilenilen sistemi etkileyen faktörlerin göz ardı edildiği ve her sistem için sabit bir eşik değerinin önerildiği klasik model bu konuda yetersiz kalmaktadır.

Maksimum ve minimum eşik değerleri belirlendikten sonra, γ değişkeninin alacağı her bir değer için kriter değerler $x_1 + (x_0 - x_1)\gamma$ ifadesi ile hesaplanmıştır. (4.5) eşitliğinden yararlanılarak, kriter değerlerden büyük olan MYAMA değerleri için ξ_i^γ ($i = 1, 2, \dots, n$) rasgele değişkeninin “1” değerini, küçük ve eşit olan MYAMA değerleri için de “0” değerini aldığı bir tanımlama yapılmış ve bir tablo üzerinde gösterilmiştir. Böylelikle, MYAMA değerlerinden oluşan sürekli değerli dizi, “0” ve “1” değerlerinden oluşan iki değerli rasgele bir dizi haline dönüştürülmüş olur.

Gerçek taşkın yıllarının da tablo üzerinde belirtilmesiyle birlikte iki değerli rasgele diziler için aşağıda yer alan dört farklı durum ortaya çıkmıştır.

- “1” ile belirtilen yıllar, gerçekte taşkın olan (“1”) yıllara karşılık geliyorsa, ilgili yıl “**önlem alınan ve taşkın olan yıl**” ya da “**taşkın yılı**” olarak ifade edilmiştir. Bu durum, modellemenin başarısını göstermektedir.
- “1” ile belirtilen yıllar, gerçekte taşkın olmayan (“0”) yıllara karşılık geliyorsa, ilgili yıl “**önlem alındığı halde taşkın olmayan yıl**” olarak ifade edilmiştir.
- “0” ile belirtilen yıllar, gerçekte taşkın olmayan (“0”) yıllara karşılık geliyorsa, ilgili yıl “**önlem alınmayan ve taşkın olmayan yıl**” ya da “**sorunsuz yıl**” olarak ifade edilmiştir. Bu durum da yine modellemenin başarısını göstermektedir.
- “0” ile belirtilen yıllar, gerçekte taşkın olan (“1”) yıllara karşılık geliyorsa, ilgili yıl “**önlem alınmadığı halde taşkın olan yıl**” ya da “**kaçırılan yıl**”

olarak ifade edilmiştir. Bu durum, modellemenin başarısız olduğunu göstermektedir.

Gerçek taşkın verileri ilgili tabloya dahil edildiğinde, “taşkın yılı” veya “önlem alındığı halde taşkın olmayan yıl” olarak yorumlanabilen “1” değeri için genel olarak “**önlem alınan yıl**” ifadesi de kullanılmıştır. Aynı şekilde, sorunsuz yıl” veya “kaçırılan yıl” olarak yorumlanabilen “0” değeri için “**önlem alınmayan yıl**” genellemesi yapılmıştır.

Her bir risk düzeyi için kriter değerlerin toplam kaç yılda geçildiğini gösteren S_n tekrar istatistiği, “1” (önlem alınan yıl) ile ifade edilen yılların toplamına eşittir.

Risk düzeyleri bazında toplam “taşkın yılı”nın; havzadaki toplam taşkın yılına oranı “**taşkın tahmin oranı**”nı verirken, S_n tekrar istatistiğine oranı “**kaynakların etkin kullanım oranı**”nı vermektedir. Havza genelinde ne kadar çok önlem alınırsa, o kadar çok taşkın yakalanmakta ve buna bağlı olarak da taşkın tahmin oranı yüksek değerler almaktadır.

Taşkın tahmin oranının yüksek olduğu risk düzeylerinde ise genellikle kaynakların etkin kullanım oranı düşük çıkmaktadır. Çünkü, önlem alınan yıl sayısının artırılması, önlem alındığı halde taşkın olmayan yılların sayısını da artıracığı için kaynakların etkin kullanım oranını düşürecektir. Ancak bu durum, her ne kadar kaynakların israfı gibi görülse de, olası bir felaketin önlenmesi bakımından ileriye dönük bir yatırım olarak düşünülmelidir.

Dolayısıyla, havza genelinde yapılacak ileriye dönük tahminlerde, en etkin risk düzeyinin belirlenebilmesi için taşkın tahmin ve kaynakların etkin kullanım oranları birlikte değerlendirilmiştir. Karar vericinin tercihinine göre değişiklik gösterebilen en etkin risk düzeyi için “**tercih edilen risk düzeyi**” ifadesi kullanılmıştır.

Eğer, karar verici, yaşanacak olası bir taşkın can ve mal kaybına neden olabileceğini düşünüyorsa, öncelikle en yüksek taşkın tahmin oranına sahip düzeyleri belirleyerek, bu düzeyler arasından kaynakların etkin kullanım oranının en yüksek olduğu düzeyi , “tercih edilen risk düzeyi” olarak seçmelidir.

Yine, karar verici eldeki kısıtlı kaynakları da göz önüne alarak, yaşanacak olası bir taşkının çok fazla bir zarara neden olmayacağını düşünüyorsa, “tercih edilen risk düzeyi”ni kaynakların etkin kullanım oranına öncelik vererek belirlemelidir.

Gerekli hesaplamalar yapılarak tercih edilen risk düzeyinin belirlendiği bir havzada, ileriye dönük tahminlerin yapılabilmesi için tekrar istatistiklerinin olasılık dağılımları kullanılmıştır.

Tekrar istatistiklerinin olasılık dağılımları içinde kullanılan ve “**aşma olasılığı**” olarak adlandırılan $p(\text{önlem}) = p^\gamma = P(\xi_i^\gamma = 1) = P(X_i > x_1 + (x_0 - x_1)\gamma)$ değerleri, γ değişkenine bağlı olarak değişen kriter değerlerin üzerinde yer alan gözlemlerin oranını göstermektedir. Diğer bir ifadeyle, önlem alınan yılların, tüm yıllara oranını ifade etmektedir. $p(\text{önlem})$ değerleri, X (MYAMA) rasgele değişkeninin uyduğu dağılımın olasılık dağılımı kullanılarak hesaplanmıştır.

Literatürde, MYAMA gibi hidrolojik verilerin, en çok iki veya üç parametrelili log-normal, log-pearson ve log-logistic dağılımlarına uyduğu görülmektedir (Onüçyıldız, 1999; Büyükkaracıgan ve Kahya, 2007, 2009). Parametre sayısı arttıkça, dağılımın veri kümesine uygunluğu da artmaktadır. Ancak, parametre sayısının çok olduğu dağılımların tercih edilebilmesi için örneklem büyüklüğünün yeterli sayıda olması gerekir.

Uygulamada belirlenen havzalar için seçilen örneklemelere ait MYAMA değerleri için en uygun dağılımın, Kolmogrov-Smirnov testine göre, Wakeby dağılımı olduğu görülmüştür. Wakeby dağılımı 5 parametrelili bir dağılımdır (Choi et al., 2011; Onüçyıldız, 1999). Ancak, havzalardan (kitle) çekilen örneklemelerin büyüklüğü ($n_i = 30, i = 1,2,3,4$) yeterli düzeyde olmadığı için bu dağılım tercih edilmemiştir.

Kolmogrov-Smirnov testine göre Wakeby dağılımdan sonra veri kümeleri için en uygun dağılımın, üç parametrelili log-logistic(3P) dağılımı olduğu görülmüştür.

Log-logistic(3P) dağılımının olasılık yoğunluk fonksiyonu,

$$f(x) = \frac{\alpha}{\beta} \left(\frac{x - \gamma}{\beta} \right)^{\alpha-1} \left(1 + \left(\frac{x - \gamma}{\beta} \right)^\alpha \right)^{-2} \quad (7)$$

eşitliği ile ifade edilmekte olup; α ($\alpha > 0$) şekil (*shape*) parametresini, β ($\beta > 0$) ölçek (*scale*) parametresini, γ ($\gamma > 0$, $\gamma \leq x$) ise yer (*location*) parametresini ifade etmektedir (Choi et al., 2011).

Şekil parametresi, dağılımın basık ya da sivri olma özelliğini gösterir. Ölçek parametresi, dağılımın hem basık-sivri olma özelliğini, hem de x eksenindeki konumunu ölçeklendirir. Yer parametresi ise dağılımın şekil ve ölçek parametreleri ile belirlenen şeklini muhafaza ederek, x eksenindeki konumunu belirler.

X (MYAMA) rasgele değişkeninin sahip olduğu log-logistic(3P) dağılımının olasılık yoğunluk fonksiyonu kullanılarak, aşma olasılıkları

$$p(\text{önlem}) = p^\gamma = P(X_i > X_1 + (X_0 - X_1)\gamma) = \int_{X_1 + (X_0 - X_1)\gamma}^{\infty} f(x) dx \quad (8)$$

eşitliği ile hesaplanmıştır.

Hesaplanan $p(\text{önlem})$ değerleri, ileriye dönük tahminler için tekrar istatistiklerinin olasılık dağılımları içinde kullanılmış, ancak bazı sorunlarla karşılaşmıştır. Şöyleki, $p(\text{önlem})$ değerleri, “taşkın” olan yıllar ile birlikte “önlem alındığı halde taşkın olmayan” yılları da içerdiği için havzadaki gerçek taşkın oranının oldukça üzerinde değerler almaktadır. Bu yüksek oranların direkt olarak olasılık dağılımlarının içinde kullanılması, ileriye dönük tahminlerde hatalı sonuçlar elde edilmesine neden olacaktır. Bu yüzden, havzalardaki gerçek oran göz önünde bulundurularak, her düzeydeki $p(\text{önlem})$ değerlerinin, önlem alındığı halde taşkın olmayan yılların getirdiği etkiden arındırılması gerekmektedir.

Zaman serileri analizinde “*mevsimsel düzeltme*” olarak sıkça rastlanılan bu durum için $p(\text{önlem})$ değerleri, her bir risk düzeyindeki “taşkın yılı” olarak ifade edilen yılların oranı $p(\text{taşkın})$ ile çarpılarak, $p(\text{düzeltilmiş})$ değerleri elde edilmiştir. Böylelikle, $p(\text{önlem})$ değerleri üzerinde “düzeysel düzeltme” yapılarak, önlem alınan yıllar içindeki taşkın olan yılların oranı elde edilmiştir.

O halde, tekrar istatistiklerinin olasılık dağılımlarında $p(\text{önlem})$ değerleri yerine, $p(\text{düzeltilmiş})$ değerleri kullanılarak ileriye dönük tahmin olasılıkları hesaplanmıştır. Bu olasılıklar, her düzeyde birbirine yakın sonuçlar verse bile, taşkın tahminlerinde mutlaka “tercih edilen risk düzeyi”nde yer alan olasılıklar kullanılmalıdır. Çünkü, havzanın yapısı, karar vericinin tutumu, taşkın tahmin ve

kaynakların etkin kullanım oranları gibi birçok faktör göz önüne alınarak belirlenen “tercih edilen risk düzeyi”, diğer risk düzeylerine göre çok daha doğru ve tutarlı sonuçlar ortaya çıkaracaktır.

Bu bölümde modelleme ve ileriye dönük tahminler ile ilgili ayrıntılı olarak verilen tanım ve kavramlar, Bölüm 5.3 ve Bölüm 5.4’te belirlenen havzalara uygulanmıştır.

5.3 Önerilen Modelin Havzalara Uygulanması

Bu bölümde, dört havzadan elde edilen veri kümeleri için öncelikle tanımlayıcı istatistikler verilerek, X rasgele değişkeninin sahip olduğu dağılımın parametreleri belirlenmiş ve MYAMA değerlerinin homojenliği test edilmiştir.

Daha sonra, havza özelliklerine göre minimum ve maksimum eşik değerleri elde edilerek kriter değerler hesaplanmış ve sürekli değerler iki değerli hale dönüştürülmüştür. Elde edilen iki değerli diziler üzerinde S_n tekrar istatistiğinin tanımlanması ile birlikte taşkın tahmin ve kaynakların etkin kullanım oranları hesaplanarak, “tercih edilen risk düzeyi” belirlenmiştir.

5.3.1 Küçük Menderes havzası

1961-1990 yılları arasında (30 yıllık) Küçük Menderes havzasında yer alan 9 adet AGİ’den derlenen YAMA (m^3/s) değerlerinin yıl bazında maksimumları alınarak, MYAMA (m^3/s) değerleri elde edilmiştir.

X_1, X_2, \dots, X_{30} , MYAMA değerlerinden oluşan rasgele bir dizi olmak üzere, bu değerlere ait tanımlayıcı istatistikler Tablo 5.1’de yer almaktadır.

Tablo 5.1’den görüleceği üzere; ortalaması $229.4 m^3/s$ olan MYAMA değerlerinin %77’si **77.6 – 381.2** (229.4 ± 151.8) arasında yer almaktadır. Diğer bir ifade ile MYAMA değerleri, ortalama etrafında $\pm 151.8 m^3/s$ olarak yaygınlık göstermektedir.

Çarpıklık katsayısının (1.6) sıfırdan büyük olması, dağılımın normal dağılıma göre sağa çarpık; basıklık katsayısının (2.9) sıfırdan büyük olması ise dağılımın normal dağılıma göre sivri olduğu anlamına gelmektedir.

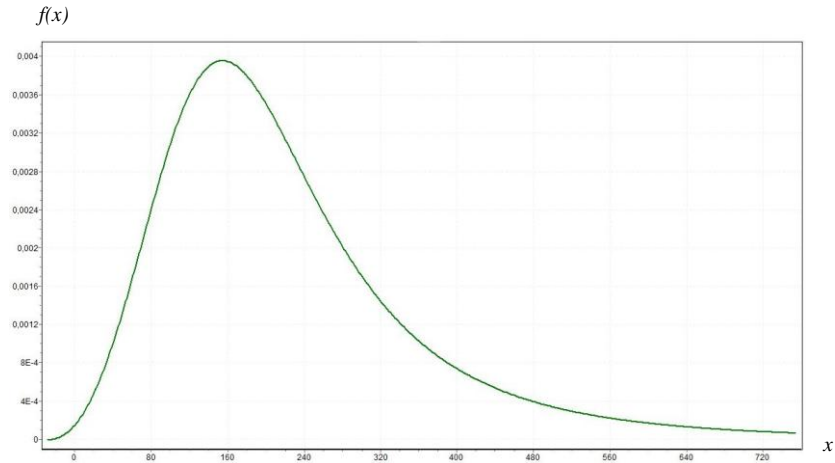
Yüzdelerle bakıldığı zaman, MYAMA değerlerinin %10'unun 64.1, %25'inin 132.3, %50'sinin 188.5 (ortanca), %75'inin 292.5 ve %90'nın da 450.1 değerlerinin altında yer aldığı görülmektedir.

Tablo 5.1 Tanımlayıcı İstatistikler

İstatistikler	
MYAMA (m ³ /s)	
N	30
Ortalama	229.4
Ortanca	188.5
Tepe Değeri	180
St. Sapma	151.8
Varyans	23041.9
Çarpıklık Kat.	1.6
Basıklık Kat.	2.9
Dağ. Genişliği	652.7
Minimum	40.3
Maksimum	693
Toplam	6881.2
Yüzdeler	
(P10)	64.1
(P25)	132.3
(P50)	188.5
(P75)	292.5
(P90)	450.1

Küçük Menderes havzasında, 30 yıla ait MYAMA değerleri Kolmogrov-Smirnov testine göre $\alpha=3.2$, $\beta=223.5$ ve $\gamma=-28.1$ parametreleri ile log-logistic (3P) dağılımına uymaktadır. Dağılım grafiği Şekil 5.2'de gösterilmiştir.

Şekil 5.2 Log-logistic (3P) dağılım grafiği ($\alpha=3.2$, $\beta=223.5$, $\gamma=-28.1$)



Şekil 5.2’de MYAMA değerlerinin normal dağılıma göre daha sivri ve sağa çarpık olduğu görülmektedir.

MYAMA değerlerinin homojen olup olmadığının belirlenmesi için Run testi kullanılmıştır. Hata payının 0.05, test değerinin “ortanca” olarak alındığı Run testi ile iki değerli dizi elemanlarının homojen olduğu şeklindeki H_0 yokluk hipotezi, elemanların homojen olmadığı alternatifine karşı test edilmiştir. Bu testin sonuçları Tablo 5.2’de verilmiştir.

Tablo 5.2 Run testi

Run Testi	
MYAMA (m ³ /s)	
Test Değeri (Ortanca)	188.5
MYAMA < Test Değeri	15
MYAMA >= Test Değeri	15
Toplam MYAMA sayısı	30
Run sayısı	13
Z	-0.929
Asymp. Sig. (2-yönlü)	0.353

Tablo 5.2’den elde edilen sonuçlara göre; 0.353 olarak hesaplanan p olasılık değeri $\alpha = 0.05$ ’ten büyük olduğu için yokluk hipotezi reddedilemez. Dolayısıyla, Küçük Menderes havzasından elde edilen MYAMA değerlerinin %5 hata (yanılma) payı ve %95 güvenilrlikle homojen olduğu söylenebilir.

Sel ve taşkın olaylarının nadir olarak görüldüğü Küçük Menderes havzasında taşkın yıllıkları incelendiğinde 1961-1990 yılları arasında yalnızca 5 kez taşkın meydana geldiği görülmüştür. Dolayısıyla, 30 yılda havzada taşkın olma oranı 5/30 yani %16.6’dır.

MYAMA değerleri küçükten büyüğe doğru sıralandığı zaman, gerçekte taşkın olan yılların çoğunluğunun, veri kümesinin %50’sinin üzerindeki yıllara denk geldiği görülmüştür. Bu durumda, maksimum eşik değeri (x_0) ve minimum eşik değerinin (x_1), veri kümesinin %50’sinin (188.5) üzerindeki değerlerden seçilmesine karar verilmiştir.

Tablo 5.3’te maksimum eşik değeri (x_0), 292.5 ($P75$) ve minimum eşik değeri (x_1), 188.5 ($P50$, ortanca) alınarak, γ parametresinin alacağı her bir değer için kriter değerler hesaplanmıştır. Kriter değerlerden büyük olan MYAMA

değerleri için ξ_i^y ($i = 1, 2, \dots, n$) rasgele değişkeninin “1” değerini, küçük ve eşit olan MYAMA değerleri için de “0” değerini aldığı tabloda, her bir risk düzeyi için S_n tekrar istatistiği hesaplanmıştır.

Ayrıca, risk düzeyleri için hesaplanan kriter değerler ile birlikte klasik modelde kullanılan “ortalama” değeri de ayrı bir sütun olarak tabloya eklenmiştir. “Ortanca” değeri ise modelde 0.0 düzeyindeki kriter değere eşit olduğu için ayrı olarak gösterilmemiştir.

Tablo 5.3 Düzeylere göre MYAMA değerlerinin iki değerli hale çevrilmesi

	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	Ort.
Su Yılı	188.5	198.9	209.3	219.7	230.1	240.5	250.9	261.3	271.7	282.1	292.5	229.4
1961	0	0	0	0	0	0	0	0	0	0	0	0
1962	1	1	1	1	0	0	0	0	0	0	0	0
1963	0	0	0	0	0	0	0	0	0	0	0	0
1964	0	0	0	0	0	0	0	0	0	0	0	0
1965	1	1	1	1	1	1	1	1	1	1	1	1
1966	1	1	0	0	0	0	0	0	0	0	0	0
1967	0	0	0	0	0	0	0	0	0	0	0	0
1968	0	0	0	0	0	0	0	0	0	0	0	0
1969	0	0	0	0	0	0	0	0	0	0	0	0
1970	1	0	0	0	0	0	0	0	0	0	0	0
1971	1	1	1	1	1	0	0	0	0	0	0	1
1972	1	1	1	1	1	1	1	1	1	1	1	1
1973	0	0	0	0	0	0	0	0	0	0	0	0
1974	1	1	1	1	1	0	0	0	0	0	0	1
1975	1	1	1	1	1	1	1	1	1	1	1	1
1976	0	0	0	0	0	0	0	0	0	0	0	0
1977	1	1	1	1	0	0	0	0	0	0	0	1
1978	1	1	1	1	1	1	1	1	1	1	1	1
1979	1	1	1	1	1	1	1	1	1	0	0	1
1980	1	1	1	1	1	1	1	1	1	1	0	1
1981	1	1	1	1	1	1	1	1	1	1	1	1
1982	1	1	1	1	1	1	1	1	1	1	1	1
1983	0	0	0	0	0	0	0	0	0	0	0	0
1984	1	1	1	1	1	1	1	1	1	1	1	1
1985	0	0	0	0	0	0	0	0	0	0	0	0
1986	0	0	0	0	0	0	0	0	0	0	0	0
1987	0	0	0	0	0	0	0	0	0	0	0	0
1988	0	0	0	0	0	0	0	0	0	0	0	0
1989	0	0	0	0	0	0	0	0	0	0	0	0
1990	0	0	0	0	0	0	0	0	0	0	0	0
S_n	15	14	13	13	11	9	9	9	9	7	7	12

Küçük Menderes havzasında 1965, 1974, 1976, 1978 ve 1981 yıllarında taşkın meydana gelmiştir. Tablo 5.3'te, gerçekte taşkın meydana gelen yıllar, şerit halinde (satur) boyanarak gösterilmiştir. Bu yıllara denk gelen ve “taşkın yılı” olarak tahmin edilen “1” değerleri de koyu olarak belirtilmiştir. Örneğin, 0.0 düzeyi için 1965, 1974, 1978 ve 1981 yılları “taşkın yılı”, 1962, 1966, 1970-1972, 1975, 1977,1979, 1980, 1982 ve 1984 yılları “önlem alındığı halde taşkın olmayan yılı”, 1976 yılı “felaket yılı” ve kalan diğer yıllar da “sorunsuz yılı” olarak ifade edilmektedir.

Tablo 5.3'ten görüleceği üzere; gerçekte taşkın meydana gelen yıllar (1976 yılı hariç), **0.0-0.4** düzeyleri arası ve “**ortalama değer**” için de “**taşkın yılı**” olarak tahmin edilmiştir. Ancak, **0.5-1.0** düzeyleri arasında, önlem alınan yıl sayısının azalmasıyla, modelin taşkın yakalama özelliği de azaldığı için “**kaçırılan yıl**” sayısında artış gözlenmiştir.

Dolayısıyla; S_n tekrar istatistiğine göre, önlem alınan yıl sayısının azalmasıyla (maliyet düşer, risk artar), gerçek taşkınların tahmin edilebilme olanağının azaldığı; önlem alınan yıl sayısının artmasıyla (maliyet artar, risk azalır) ise gerçek taşkınların tahmin edilebilme olanağının arttığı söylenebilir. Örneğin; 0.4 düzeyinde 11 kez önlem alındığında 4 taşkın yılı, 0.5 düzeyinde ise önlem sayısı 9'a düşürüldüğünde 3 taşkın yılı doğru olarak tahmin edilebilmiştir.

Havzadaki toplam taşkın yılı ile her bir düzeydeki toplam taşkın yılı ve S_n tekrar istatistiği kullanılarak, taşkın tahmin ve kaynakların etkin kullanım oranları hesaplanmış ve Tablo 5.4'de gösterilmiştir.

Tablo 5.4 Taşkın tahmin ve kaynakların etkin kullanım oranları

Düzye	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	Ort.
Kriter Değer	188.5	198.9	209.3	219.7	230.1	240.5	250.9	261.3	271.7	282.1	292.5	229.4
S_n	15	14	13	13	11	9	9	9	9	7	7	12
Taşkın tahmin oranı	0.800	0.800	0.800	0.800	0.800	0.600	0.600	0.600	0.600	0.600	0.600	0.800
Kaynakların etkin kullanım oranı	0.267	0.286	0.308	0.308	0.364	0.333	0.333	0.333	0.333	0.429	0.429	0.333

Tablo 5.4'te **0.0-0.4** düzeyleri arası ve “**ortalama değer**” için taşkın tahmin oranı %80 olarak belirlenmiştir. **0.5-1.0** düzeyleri arasında ise taşkın tahmin oranı %60'a düşmektedir.

Taşkın tahmin oranının yüksek olduğu 0.0-0.4 düzeyleri arası ve “ortalama değer” içerisinde, kaynakların etkin kullanım oranı en yüksek olan %36.4 ile 0.4 düzeyidir. 0.9 ve 1.0 düzeylerinde her ne kadar kaynaklar diğer düzeylere göre daha etkin kullanılıyor gibi görünse de, bu düzeyler taşkınları tahmin etmede yetersiz kalmaktadır.

Dolayısıyla; karar verici tarafından Küçük Menderes havzasında yaşanacak olası bir taşkınım yerleşim yerlerine veya ekili arazilere zarar vereceği düşünülüyorsa, tercih edilen risk düzeyi **0.4** olarak belirlenmelidir. Ayrıca, karar verici, 0.4 düzeyinin yanısıra, taşkın tahmin oranı aynı, fakat kaynakların etkin kullanım oranı düşük olan 0.0-0.3 düzeyleri arası ve ortalama değer gibi farklı alternatiflere de sahiptir. Ancak, herhangi bir maliyet kısıtı yoksa bu düzeylerin seçilmesi önerilmektedir.

Sonuç olarak, taşkın olaylarının az sayıda yaşandığı Küçük Menderes havzası için klasik modelde sabit eşik değeri olarak kullanılan ortalama ve ortanca değerleri, önerilen modelde de “tercih edilen risk düzeyi” olarak seçilebilir.

5.3.2 Gediz havzası

1961-1990 yılları arasında (30 yıllık) Gediz havzasında yer alan 18 adet AGİ'den derlenen YAMA (m^3/s) değerlerinin maksimumları alınarak, MYAMA (m^3/s) değerleri elde edilmiştir.

X_1, X_2, \dots, X_{30} , MYAMA değerlerinden oluşan rasgele bir dizi olmak üzere bu değerlere ait tanımlayıcı istatistikler, Tablo 5.5'de yer almaktadır.

Tablo 5.5'den görüleceği üzere; ortalaması $538.2 m^3/s$ olan MYAMA değerlerinin %77'si **223.8 – 852.6** (538.2 ± 314.4) arasında yer almaktadır. Diğer bir ifade ile MYAMA değerleri, ortalama etrafında $\pm 314.4 m^3/s$ olarak yaygınlık göstermiştir.

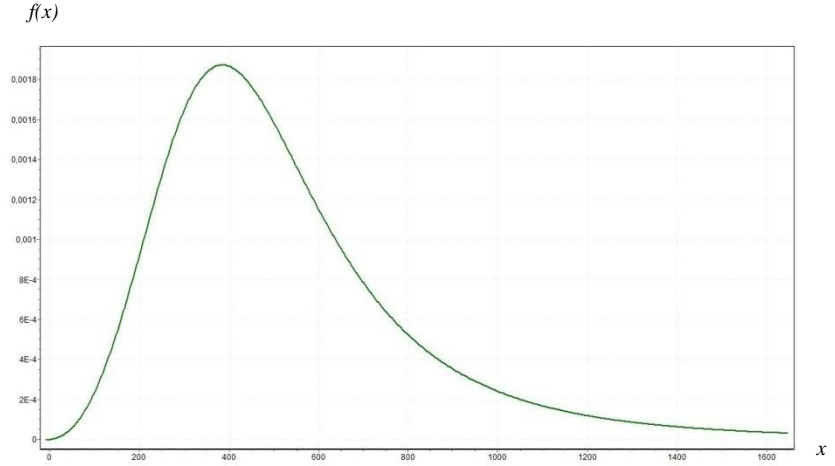
Çarpıklık katsayısının (1.6) sıfırdan büyük olması, dağılımın normal dağılıma göre sağa çarpık; basıklık katsayısının (3.3) sıfırdan büyük olması ise dağılımın normal dağılıma göre sivri olduğu anlamına gelmektedir.

Yüzdelerle bakıldığı zaman, MYAMA değerlerinin %10'unun 189, %25'inin 316, %50'sinin 482.5 (ortanca), %75'inin 654 ve %90'ının da 943.4 değerlerinin altında yer aldığı görülmektedir.

Tablo 5.5 Tanımlayıcı İstatistikler

İstatistikler	
MYAMA (m ³ /s)	
N	30
Ortalama	538.2
Ortanca	482.5
Tepe Değeri	615
St. Sapma	314.4
Varyans	98826.7
Çarpıklık Kat.	1.6
Basıklık Kat.	3.3
Dağ. Genişliği	1420
Minimum	135
Maksimum	1555
Toplam	16147
Yüzdeler	
(P10) 10	189.0
(P25) 25	316.0
(P50) 50	482.5
(P75) 75	654.0
(P90) 90	943.4

Gediz havzasında, 30 yıla ait MYAMA değerlerinin Kolmogrov-Smirnov testine göre $\alpha=3.2$, $\beta=477.5$ ve $\gamma=-7.6$ parametreleri ile log-logistic(3P) dağılımına uymaktadır. Dağılım grafiği Şekil 5.3'te gösterilmiştir.

Şekil 5.3 Log-logistic (3P) dağılım grafiği ($\alpha=3.2$, $\beta=477.5$ ve $\gamma=-7.6$)

Şekil 5.3'te MYAMA değerlerinin normal dağılıma göre daha sivri ve sağa çarpık olduğu görülmektedir.

MYAMA değerlerinin homojen olup olmadığının belirlenmesi için Run testi kullanılmıştır. Hata payının 0.05, test değerinin “ortanca” olarak alındığı Run testi ile iki değerli dizi elemanlarının homojen olduğu şeklindeki H_0 yokluk hipotezi, elemanların homojen olmadığı alternatifine karşı test edilmiştir. Bu testin sonuçları Tablo 5.6'da verilmiştir.

Tablo 5.6 Run testi

Run Testi	
MYAMA (m ³ /s)	
Test Değeri (Ortanca)	482.5
MYAMA < Test Değeri	15
MYAMA >= Test Değeri	15
Toplam MYAMA sayısı	30
Run sayısı	17
Z	0.186
Asymp. Sig. (2-yönlü)	0.853

Tablo 5.6'dan elde edilen sonuçlara göre, 0.853 olarak hesaplanan p olasılık değeri $\alpha = 0.05$ 'ten büyük olduğu için yokluk hipotezi reddedilemez. Dolayısıyla, Gediz havzasından elde edilen MYAMA değerlerinin %5 hata (yanılma) payı ve %95 güvenilirlikle homojen olduğu söylenebilir.

Sel ve taşkın olaylarının Küçük Menderes havzasına göre biraz daha fazla görüldüğü Gediz havzasında taşkın yıllıkları incelendiğinde 1961-1990 yılları

arasında 7 kez taşkın meydana geldiği görülmüştür. Dolayısıyla, 30 yılda havzada taşkın olma oranı $7/30$ yani %23.3'dür.

MYAMA değerleri küçükten büyüğe doğru sıralandığı zaman, taşkın olan yılların çoğunluğunun veri kümesinin %50'sinin üzerindeki yıllara denk geldiği görülmüştür. Bu durumda, maksimum eşik değeri (x_0) ile minimum eşik değerinin (x_1), veri kümesinin %50'sinin (482.5) üzerindeki değerlerden seçilmesine karar verilmiştir.

Tablo 5.7'de maksimum eşik değeri (x_0), 654 ($P75$) ve minimum eşik değeri (x_1), 482.5 ($P50$, ortanca) alınarak, γ parametresinin alacağı her bir değer için kritik değerler hesaplanmıştır. Kriter değerlerden büyük olan MYAMA değerleri için ξ_i^γ ($i = 1, 2, \dots, n$) rasgele değişkeninin "1" değerini, küçük ve eşit olan MYAMA değerleri için de "0" değerini aldığı tabloda, her bir düzey için S_n tekrar istatistiği hesaplanmıştır.

Ayrıca, risk düzeyleri için hesaplanan kriter değerler ile birlikte klasik modelde kullanılan "ortalama" değeri de ayrı bir sütun olarak tabloya eklenmiştir. "Ortanca" değeri ise modelde 0.0 düzeyindeki kriter değere eşit olduğu için ayrı olarak gösterilmemiştir.

Gediz havzasında 1961, 1963, 1965, 1966, 1971, 1974 ve 1981 yıllarında taşkın meydana gelmiştir. Tablo 5.7'de gerçekte taşkın meydana gelen yıllar, şerit halinde (satur) boyanarak gösterilmiştir. Bu yıllara denk gelen ve "taşkın yılı" olarak tahmin edilen "1" değerleri de koyu olarak belirtilmiştir. Örneğin, 0.0 düzeyi için 1963, 1966, 1971, 1974 ve 1981 yılları "taşkın yılı", 1962, 1968-1970, 1976, 1979, 1980, 1982, 1984 ve 1987 yılları "önlem alındığı halde taşkın olmayan yıl", 1961 ve 1965 yılları "felaket yılı" ve kalan diğer yıllar da "sorunsuz yıl" olarak ifade edilmektedir.

Tablo 5.7 Düzeylere göre MYAMA değerlerinin iki değerli hale çevrilmesi

	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	Ort.
Su Yılı	482.5	499.7	516.8	534.0	551.1	568.3	585.4	602.6	619.7	636.9	654.0	538.2
1961	0	0	0	0	0	0	0	0	0	0	0	0
1962	1	1	1	1	1	1	1	0	0	0	0	1
1963	1	1	1	1	1	1	1	1	1	1	1	1
1964	0	0	0	0	0	0	0	0	0	0	0	0
1965	0	0	0	0	0	0	0	0	0	0	0	0
1966	1	1	1	1	1	1	1	1	1	1	1	1
1967	0	0	0	0	0	0	0	0	0	0	0	0
1968	1	1	1	1	1	1	1	1	1	1	1	1
1969	1	0	0	0	0	0	0	0	0	0	0	0
1970	1	1	1	1	1	1	1	1	1	1	0	1
1971	1	0	0	0	0	0	0	0	0	0	0	0
1972	0	0	0	0	0	0	0	0	0	0	0	0
1973	0	0	0	0	0	0	0	0	0	0	0	0
1974	1	1	1	1	1	1	1	1	0	0	0	1
1975	0	0	0	0	0	0	0	0	0	0	0	0
1976	1	1	1	1	1	1	1	1	1	1	1	1
1977	0	0	0	0	0	0	0	0	0	0	0	0
1978	0	0	0	0	0	0	0	0	0	0	0	0
1979	1	1	0	0	0	0	0	0	0	0	0	0
1980	1	1	1	1	1	1	1	1	0	0	0	1
1981	1	1	1	1	1	1	1	1	1	1	1	1
1982	1	1	1	1	1	1	1	1	1	1	1	1
1983	0	0	0	0	0	0	0	0	0	0	0	0
1984	1	1	1	1	1	1	1	1	1	1	1	1
1985	0	0	0	0	0	0	0	0	0	0	0	0
1986	0	0	0	0	0	0	0	0	0	0	0	0
1987	1	1	1	1	1	0	0	0	0	0	0	1
1988	0	0	0	0	0	0	0	0	0	0	0	0
1989	0	0	0	0	0	0	0	0	0	0	0	0
1990	0	0	0	0	0	0	0	0	0	0	0	0
S_n	15	13	12	12	12	11	11	10	8	8	7	12

Tablo 5.7'den görüleceği üzere, gerçekte taşkın meydana gelen yıllar (1961, 1965 ve 1971 yılları hariç), **0.0-0.7** düzeyleri arası ve “**ortalama değer**” için de “**taşkın yılı**” olarak tahmin edilmiştir. 1971 yılındaki taşkın ise yalnızca **0.0** düzeyinde tahmin edilmiştir. **0.8-1.0** düzeyleri arasında, önlem alınan yıl sayısının azalmasıyla “kaçırılan yıl” sayısında artış gözlenmiştir.

S_n tekrar istatistiğine göre, 0.7 düzeyinde 10 kez önlem alındığında 4 taşkın yılı, 0.8 düzeyinde ise önlem sayısı 8'e düşürüldüğünde 3 taşkın yılı doğru olarak tahmin edilebilmiştir. Havzadaki 5 taşkın yılı ise yalnızca en fazla önlemin

alındığı 0.0 düzeyinde tahmin edilebilmiştir. 0.1 düzeyinde, önlem sayısı iki azaltıldığında bile yakalanan taşkın sayısı 4'e düşmektedir.

Havzadaki toplam taşkın yılı ile her bir düzeydeki toplam taşkın yılı ve S_n tekrar istatistiği kullanılarak, taşkın tahmin ve kaynakların etkin kullanım oranları hesaplanmış ve Tablo 5.8'de gösterilmiştir.

Tablo 5.8 Taşkın tahmin ve kaynakların etkin kullanım oranları

Düzye	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	Ort.
Kriter Değer	482.5	499.7	516.8	534.0	551.1	568.3	585.4	602.6	619.7	636.9	654.0	538.2
S_n	15	13	12	12	12	11	11	10	8	8	7	12
Taşkın tahmin oranı	0.714	0.571	0.571	0.571	0.571	0.571	0.571	0.571	0.429	0.429	0.429	0.571
Kaynakların etkin kullanım oranı	0.333	0.308	0.333	0.333	0.333	0.364	0.364	0.400	0.375	0.375	0.429	0.333

Tablo 5.8'de taşkın tahmin oranı, **0.0** düzeyi için %71.4, **0.1-0.7** düzeyleri arası ve **ortalama** değer için %57.1, **0.8-1.0** düzeyleri arası %42.9 olarak hesaplanmıştır.

Taşkın tahmin oranının en yüksek olduğu ve aynı oran içerisinde, kaynakların etkin kullanım oranının en yüksek olduğu düzey 0.0 düzeyidir. 0.7 ve 1.0 düzeylerinde her ne kadar kaynaklar diğer düzeylere göre daha etkin kullanılıyor gibi gözükse de, bu düzeyler taşkınları tahmin etmede yetersiz kalmaktadır.

Dolayısıyla, karar verici tarafından Gediz havzasında yaşanacak olası bir taşkın yerleşim yerlerine veya ekili arazilere zarar vereceği düşünülüyorsa, tercih edilen risk düzeyi **0.0** (ortanca) olarak belirlenmelidir. Ancak, bölgede meydana gelebilecek olası bir taşkın çevreye çok fazla zarar vermeyeceği, verse bile can ve mal kaybına neden olmayacağı düşünülüyorsa, tercih edilen risk düzeyi, 0.1-0.7 düzeyleri arasından ya da ortalama değer olarak seçilebilir.

Küçük Menderes havzasında olduğu gibi, Gediz havzasında da taşkın sayısının az olması, gerçekte yaşanan tüm taşkınların model ile tahmin edilebilme olasılığını düşürmektedir. Bu nedenle, taşkın tahmin oranları, sel ve taşkın olaylarının yoğun olarak yaşandığı havzalarda daha yüksek değerler almaktadır.

Sonuç olarak, taşkın olaylarının az sayıda yaşandığı Gediz havzası için klasik modelde sabit eşik değeri olarak kullanılan ortalama ve ortanca değerleri, önerilen modelde de “tercih edilen risk düzeyi” olarak seçilebilir.

5.3.3 Kızılırmak havzası

1961-1990 yılları arasında (30 yıllık) Kızılırmak havzasında yer alan 26 adet AGİ'den derlenen YAMA (m^3/s) değerlerinin maksimumları alınarak, MYAMA (m^3/s) değerleri elde edilmiştir.

X_1, X_2, \dots, X_{30} , MYAMA değerlerinden oluşan rasgele bir dizi olmak üzere, bu değerlere ait tanımlayıcı istatistikler, Tablo 5.9'de yer almaktadır.

Tablo 5.9'dan görüleceği üzere, ortalaması $772.2 m^3/s$ olan MYAMA değerlerinin %77'si **488.2 – 1056.2** (772.2 ± 284.0) arasında yer almaktadır. Diğer bir ifadeyle, MYAMA değerleri, ortalama etrafında $\pm 284.0 m^3/s$ olarak yaygınlık göstermiştir.

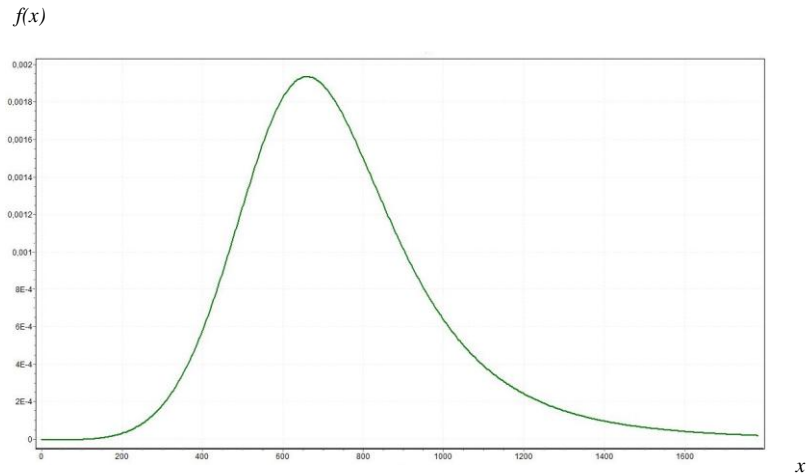
Çarpıklık katsayısının (1.5) sıfırdan büyük olması, dağılımın normal dağılıma göre sağa çarpık; basıklık katsayısının (2.7) sıfırdan büyük olması ise dağılımın normal dağılıma göre sivri olduğu anlamına gelmektedir.

Yüzdeliğe bakıldığı zaman, MYAMA değerlerinin %10'unun 473.3, %25'inin 568.8, %50'sinin 757.0 (ortanca), %75'inin 834.8 ve %90'ının da 1210.8 değerlerinin altında yer aldığı görülmektedir.

Tablo 5.9 Tanımlayıcı İstatistikler

İstatistikler	
MYAMA (m ³ /s)	
N	30
Ortalama	772.2
Ortanca	757
Tepe Değeri	591
St. Sapma	284.0
Varyans	80630.9
Çarpıklık Kat.	1.5
Basıklık Kat.	2.7
Dağ. Genişliği	1249
Minimum	424
Maksimum	1673
Toplam	23166
Yüzdeler	
(P10) 10	473.3
(P25) 25	568.8
(P50) 50	757.0
(P75) 75	834.8
(P90) 90	1210.8

Kızılırmak havzasında, 30 yıla ait MYAMA değerleri Kolmogrov-Smirnov testine göre $\alpha=2.8$, $\beta=367.3$ ve $\gamma=335.4$ parametreleri ile log-logistic(3P) dağılımına uymaktadır. Dağılım grafiği Şekil 5.4’de gösterilmiştir.

Şekil 5.4 Log-logistic(3P) dağılım grafiği ($\alpha=2.8$, $\beta=367.3$ ve $\gamma=335.4$)

Şekil 5.4’de MYAMA değerlerinin normal dağılıma göre daha sivri ve biraz sağa çarpık olduğu görülmektedir.

MYAMA değerlerinin homojen olup olmadığının belirlenmesi için Run testi kullanılmıştır. Hata payının 0.05 ve test değerinin “ortanca” olarak alındığı Run testi ile iki değerli dizi elemanlarının homojen olduğu şeklindeki H_0 yokluk hipotezi, elemanların homojen olmadığı alternatifine karşı test edilmiştir. Bu testin sonuçları Tablo 5.10’da verilmiştir.

Tablo 5.10 Run testi

Run Testi	
MYAMA (m ³ /s)	
Test Değeri (Ortanca)	757.0
MYAMA < Test Değeri	15
MYAMA >= Test Değeri	15
Toplam MYAMA sayısı	30
Run sayısı	13
Z	-0.929
Asymp. Sig. (2-yönlü)	0.353

Tablo 5.10’den elde edilen sonuçlara göre, 0.353 olarak hesaplanan p olasılık değeri $\alpha = 0.05$ ’ten büyük olduğu için yokluk hipotezi reddedilemez. Dolayısıyla, Kızılırmak havzasından elde edilen MYAMA değerlerinin %5 hata (yanılma) payı ve %95 güvenilirlikle homojen olduğu söylenebilir.

Sel ve taşkın olaylarının yoğun olarak yaşandığı Kızılırmak havzasında taşkın yıllıkları incelendiğinde 1961-1990 yılları arasında 17 kez taşkın meydana geldiği görülmüştür. Dolayısıyla, 30 yılda havzada taşkın olma oranı 17/30 yani %56.6’dır.

MYAMA değerleri küçükten büyüğe doğru sıralandığı zaman, taşkın olan yılların çoğunluğunun YAMA değerlerinin %10’unun üzerindeki yıllara denk geldiği görülmüştür. Bu durumda, maksimum eşik değeri (x_0) ve minimum eşik değerinin (x_1), veri kümesinin %10’unun (473.3) üzerindeki değerlerden seçilmesine karar verilmiştir.

Tablo 5.11’de maksimum eşik değeri (x_0) 568.8 (P_{25}) ve minimum eşik değeri (x_1) 473.3 (P_{10}) alınarak, γ parametresinin alacağı her bir değer için kriter değerler hesaplanmıştır. Kriter değerlerden büyük olan MYAMA değerleri için ξ_i^{γ} ($i = 1, 2, \dots, n$) rasgele değişkeninin “1” değerini, küçük ve eşit olan MYAMA değerleri için de “0” değerini aldığı tabloda, her bir düzey için S_n tekrar istatistiği hesaplanmıştır.

Ayrıca, risk düzeyleri için hesaplanan kriter değerler ile birlikte klasik modelde kullanılan “ortalama” ve “ortanca” değerleri de ayrı bir sütun olarak tabloya eklenmiştir. Kızılırmak havzasında kriter değerler veri kümesinin 10. ve 25. yüzdilik dilimleri arasından seçildiği için “ortanca” değeri, Küçük Menderes ve Gediz havzalarında olduğu gibi 0.0 düzeyine karşılık gelmemektedir.

Tablo 5.11 Düzeylere göre MYAMA değerlerinin iki değerli hale çevrilmesi

	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	Ort.	Ortc.
Su Yılı	473.3	482.9	492.4	502	511.5	521.1	530.6	540.2	549.7	559.3	568.8	772.2	757
1961	1	1	1	1	1	1	1	1	1	1	1	0	0
1962	0	0	0	0	0	0	0	0	0	0	0	0	0
1963	1	1	1	1	1	1	1	1	1	1	1	1	1
1964	1	1	1	1	1	1	1	1	1	1	1	0	0
1965	1	1	1	1	1	1	1	1	1	1	1	1	1
1966	1	1	1	1	1	1	1	1	1	1	1	1	1
1967	1	1	1	1	1	1	1	1	1	1	1	1	1
1968	1	1	1	1	1	1	1	1	1	1	1	1	1
1969	1	1	1	1	1	1	1	1	1	1	1	1	1
1970	1	1	1	1	1	1	1	1	1	1	1	1	1
1971	1	1	1	1	1	1	1	0	0	0	0	0	0
1972	1	1	1	1	1	1	1	1	1	1	1	1	1
1973	1	1	1	1	1	1	1	1	1	1	1	0	0
1974	0	0	0	0	0	0	0	0	0	0	0	0	0
1975	1	1	1	1	1	1	1	1	1	1	1	1	1
1976	1	1	1	1	1	1	1	1	1	1	1	1	1
1977	1	1	1	1	1	1	1	1	1	1	1	0	0
1978	1	1	1	1	1	1	1	1	1	1	1	0	0
1979	1	1	1	1	1	1	1	1	0	0	0	0	0
1980	1	1	1	1	1	1	1	1	1	1	1	1	1
1981	1	1	1	1	1	1	1	1	1	1	1	1	1
1982	1	1	1	1	1	1	1	1	1	1	1	1	1
1983	1	1	1	1	1	1	1	1	1	1	1	0	1
1984	1	1	1	1	1	1	1	1	1	1	1	0	0
1985	1	1	1	1	1	1	1	1	1	1	1	0	0
1986	0	0	0	0	0	0	0	0	0	0	0	0	0
1987	1	1	1	1	1	1	1	1	1	1	1	0	0
1988	1	1	1	1	1	1	1	1	1	1	1	1	1
1989	1	1	1	1	0	0	0	0	0	0	0	0	0
1990	1	1	1	1	1	1	1	0	0	0	0	0	0
S _n	27	27	27	27	26	26	26	24	23	23	23	14	15

Kızılırmak havzasında 1962-1964, 1966-1969, 1971-1977, 1979-1981 yılları arasında taşkın meydana gelmiştir. Tablo 5.11’de gerçekte taşkın meydana gelen yıllar, şerit halinde (satur) boyanarak gösterilmiştir. Bu yıllara denk gelen ve “taşkın yılı” olarak tahmin edilen “1” değerleri de koyu olarak belirtilmiştir.

Örneğin, 0.0 düzeyi için 1963, 1964, 1966-1969, 1971-1973, 1975-1977, 1979-1981 yılları “taşkın yılı”, 1961, 1965, 1970, 1978, 1982-1985, 1987-1990 yılları “önlem alındığı halde taşkın olmayan yıl”, 1962 ve 1974 yılları “felaket yılı” ve 1986 yılı ise “sorunsuz yıl” olarak ifade edilmektedir.

Tablo 5.11’den görüleceği üzere, gerçekte taşkın meydana gelen yıllar (1962 ve 1974 yılları hariç), **0.0-0.6** düzeyler arası için de “taşkın yılı” olarak tahmin edilmiştir. **0.7** düzeyinden itibaren tahmin edilen taşkınların sayısı azalmakta olup, ortalama ve ortanca düzeylerinde en düşük değeri almaktadır. Örneğin, S_n tekrar istatistiğine göre, 0.6 düzeyinde 26 kez önlem alındığında 15 taşkın yılı, 0.7 düzeyinde önlem sayısı 24’e düşürüldüğünde 14 taşkın yılı doğru olarak tahmin edilebilmiştir.

Ortalama ve ortanca değerlerinde ise önlem sayısının 14’e düşürülmesiyle birlikte yakalanan taşkın sayısı 10’a düşerken, kaçırılan yıl sayısı 7’ye çıkmaktadır.

Havzadaki toplam taşkın yılı ile her bir düzeydeki toplam taşkın yılı ve S_n tekrar istatistiği kullanılarak, taşkın tahmin ve kaynakların etkin kullanım oranları hesaplanmış ve Tablo 5.12’de gösterilmiştir.

Tablo 5.12 Taşkın tahmin ve kaynakların etkin kullanım oranları

Düzyey	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	Ort.	Ortc.
Kriter Değer	473.3	482.9	492.4	502	511.5	521.1	530.6	540.2	549.7	559.3	568.8	772.2	757
S_n	27	27	27	27	26	26	26	24	23	23	23	14	15
Taşkın tahmin oranı	0.882	0.882	0.882	0.882	0.882	0.882	0.882	0.824	0.765	0.765	0.765	0.588	0.588
Kaynakların etkin kullanım oranı	0.556	0.556	0.556	0.556	0.577	0.577	0.577	0.583	0.565	0.565	0.565	0.714	0.667

Tablo 5.12’de taşkın tahmin oranı, **0.0-0.6** düzeyleri için %88.2, **0.7** düzeyi için %82.4, **0.8-1.0** düzeyleri arası %76.5, **ortalama** ve **ortanca** değerleri için %58.8 olarak hesaplanmıştır.

Taşkın tahmin oranının en yüksek olduğu ve aynı oran içerisinde kaynakların etkin kullanım oranının en yüksek olduğu düzeyler 0.4-0.6 düzeyleri arasındadır. Ortalama ve ortanca değerleri için her ne kadar kaynaklar diğer düzeylere göre daha etkin kullanılıyor gibi görünse de bu değerler, taşkınları tahmin etmede yetersiz kalmaktadır.

Dolayısıyla, karar verici tarafından Kızılırmak havzasında yaşanacak olası bir taşkının yerleşim yerlerine veya ekili arazilere zarar vereceği düşünülüyorsa, tercih edilen risk düzeyi 0.4-0.6 düzeyleri arasından belirlenmelidir. Eğer, Kızılırmak havzası için kaynakların etkin kullanım oranındaki %2.1'lik farkın önemsiz olduğu düşünülürse, tercih edilen risk düzeyi, 0.0-0.3 düzeyleri arasından da seçilebilir. Ancak, Kızılırmak havzasında, Gediz ve Küçük Menderes havzalarında olduğu gibi düzey belirleme konusunda çok fazla esnek davranılmamalıdır. Özellikle, taşkın tahmin oranının en düşük olduğu ortalama ve ortanca değerleri, karar verici tarafından kesinlikle tercih edilmemelidir. Aksi bir durum, telafisi mümkün olmayan felaketlere yol açabilir.

Sonuç olarak, klasik modelde sabit eşik değeri olarak kullanılan ortalama ve ortanca değerleri, sel-taşkın olaylarının yoğun olarak yaşandığı Kızılırmak havzası için kesinlikle önerilen düzeyler değildir.

5.3.4 Sakarya havzası

1961-1990 yılları arasında (30 yıllık) Sakarya havzasında yer alan 45 adet AGİ'den derlenen YAMA (m^3/s) değerlerinin yıl bazında maksimumları alınarak, MYAMA (m^3/s) değerleri elde edilmiştir.

X_1, X_2, \dots, X_{30} , MYAMA değerlerinden oluşan rasgele bir dizi olmak üzere, bu değerlere ait tanımlayıcı istatistikler, Tablo 5.13'de yer almaktadır.

Tablo 5.13'den görüleceği üzere, ortalaması $587.1 m^3/s$ olan MYAMA değerlerinin %77'si **394.4 – 779.8** (587.1 ± 192.7) arasında yer almaktadır. Yani, MYAMA değerleri, ortalama etrafında $\pm 192.7 m^3/s$ olarak yaygınlık göstermiştir. Çarpıklık katsayısının (0.9) sıfırdan büyük olması, dağılımın normal dağılıma göre sağa çarpık; basıklık katsayısının (0.2) sıfırdan büyük olması ise dağılımın normal dağılıma göre biraz daha sivri olduğu anlamına gelmektedir.

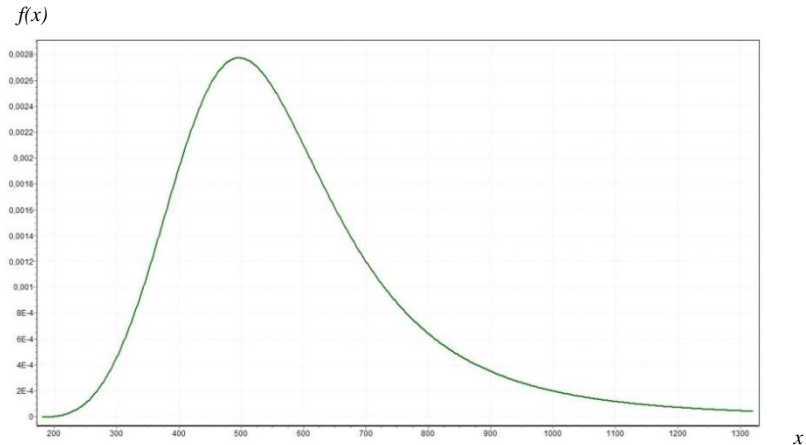
Yüzdelerle bakıldığı zaman, MYAMA değerlerinin %10'unun 396.1, %25'inin 444.8, %50'sinin 526 (ortanca), %75'inin 663.3 ve %90'ının da 973.1 değerlerinin altında yer aldığı görülmektedir.

Tablo 5.13 Tanımlayıcı İstatistikler

İstatistikler	
MYAMA (m ³ /s)	
N	30
Ortalama	587.1
Ortanca	526
Tepe Değeri	498
St. Sapma	192.7
Varyans	37147.7
Çarpıklık Kat.	0.9
Basıklık Kat.	0.2
Dağ. Genişliği	727
Minimum	279
Maksimum	1.006
Toplam	17612
Yüzdeler	
(P10) 10	396.1
(P25) 25	444.8
(P50) 50	526.0
(P75) 75	663.3
(P90) 90	973.1

Sakarya havzasında, 30 yıla ait MYAMA değerleri Kolmogrov-Smirnov testine göre $\alpha=3.8$, $\beta=364.1$ ve $\gamma=181.1$ parametreleri ile log-logistic(3P) dağılımına uymaktadır. Dağılım grafiği Şekil 5.5'te gösterilmiştir.

Şekil 5.5 Log-logistic(3P) dağılım grafiği ($\alpha=3.8$, $\beta=364.1$, $\gamma=181.1$)



Şekil 5.5'te MYAMA değerlerinin normal dağılıma göre daha sivri ve sağa çarpık olduğu görülmektedir.

MYAMA değerlerinin homojen olup olmadığının belirlenmesi için Run testi kullanılmıştır. Hata payının 0.05 ve test değerinin “ortanca” olarak alındığı Run testi ile iki değerli dizi elemanlarının homojen olduğu şeklindeki H_0 yokluk hipotezi, elemanların homojen olmadığı alternatifine karşı test edilmiştir. Bu testin sonuçları Tablo 5.14'de verilmiştir.

Tablo 5.14 Run testi

Run Testi	
MYAMA (m ³ /s)	
Test Değeri (Ortanca)	526.0
MYAMA < Test Değeri	15
MYAMA >= Test Değeri	15
Toplam MYAMA sayısı	30
Run sayısı	14
Z	-0.557
Asymp. Sig. (2-yönlü)	0.577

Tablo 5.14'den elde edilen sonuçlara göre, 0.577 olarak hesaplanan p olasılık değeri $\alpha = 0.05$ 'ten büyük olduğu için yokluk hipotezi reddedilemez. Dolayısıyla, Sakarya havzasından elde edilen MYAMA değerlerinin %5 hata (yanılma) payı ve %95 güvenlilikle homojen olduğu söylenebilir.

Sel ve taşkın olaylarının oldukça yoğun olarak yaşandığı Sakarya havzasında taşkın yıllıkları incelendiğinde 1961-1990 yılları arasında 25 kez taşkın meydana geldiği görülmüştür. Dolayısıyla, 30 yılda havzada taşkın olma oranı 25/30 yani %83.3'dür.

MYAMA değerleri küçükten büyüğe doğru sıralandığı zaman, taşkın olan yılların çoğunluğunun veri kümesinin %10'unun üzerindeki yıllara denk geldiği görülmüştür. Bu durumda, maksimum eşik değeri (x_0) ve minimum eşik değerinin (x_1), veri kümesinin %10'unun (396.1) üzerindeki değerlerden seçilmesine karar verilmiştir.

Tablo 5.15'te maksimum eşik değeri (x_0) 444.8 (P_{25}) ve minimum eşik değeri (x_1) 396.1 (P_{10}) alınarak, γ parametresinin alacağı her bir değer için kriter değerler hesaplanmıştır. Kriter değerlerden büyük olan MYAMA değerleri için

ξ_i^y ($i = 1, 2, \dots, n$) rasgele deęişkeninin “1” deęerini, küçük ve eşit olan MYAMA deęerleri için de “0” deęerini aldığı tabloda, her bir düzey için S_n tekrar istatistięi hesaplanmıştır.

Ayrıca, risk düzeyleri için hesaplanan kriter deęerler ile birlikte klasik modelde kullanılan “ortalama” ve “ortanca” deęerleri de ayrı bir sütun olarak tabloya eklenmiştir. Sakarya havzasında kriter deęerler veri kümesinin 10. ve 25. yüzdilik dilimleri arasından seçildięi için “ortanca” deęeri, Küçük Menderes ve Gediz havzalarında olduęu gibi 0.0 düzeyine karşılık gelmemektedir.

Tablo 5.15 Düzeylere göre MYAMA deęerlerinin iki deęerli hale çevrilmesi

	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	Ort.	Ortc.
Su Yılı	396.1	401	405.8	410.7	415.6	420.5	425.3	430.2	435.1	439.9	444.8	587.1	526
1961	0	0	0	0	0	0	0	0	0	0	0	0	0
1962	1	1	1	1	1	1	1	1	1	1	1	0	1
1963	1	1	1	1	1	1	1	1	1	1	1	1	1
1964	1	1	1	1	1	1	1	1	1	1	1	1	1
1965	1	1	1	1	1	1	1	1	1	1	1	1	1
1966	1	1	1	1	1	1	1	1	1	1	1	1	1
1967	0	0	0	0	0	0	0	0	0	0	0	0	0
1968	1	1	1	1	1	1	1	1	1	1	1	1	1
1969	1	1	1	1	1	1	1	1	1	1	1	1	1
1970	1	1	1	1	1	1	1	1	1	1	1	1	1
1971	1	1	1	1	1	1	1	1	1	1	1	0	0
1972	1	1	1	1	1	1	1	1	1	1	1	1	1
1973	1	1	1	1	1	1	1	1	1	1	1	0	0
1974	1	1	1	1	1	1	1	1	1	1	1	0	0
1975	1	1	1	1	1	1	1	1	1	1	1	0	0
1976	0	0	0	0	0	0	0	0	0	0	0	0	0
1977	1	1	1	1	1	1	1	1	1	1	1	0	0
1978	1	1	1	1	1	1	1	1	1	1	1	0	0
1979	1	1	1	1	1	1	1	1	1	1	1	0	0
1980	1	1	1	1	1	1	1	1	1	1	1	0	1
1981	1	1	1	1	1	1	1	1	1	1	1	1	1
1982	1	1	1	1	1	1	1	1	1	1	1	1	1
1983	1	1	1	1	1	1	0	0	0	0	0	0	0
1984	1	1	1	1	1	1	1	1	1	1	1	1	1
1985	1	1	0	0	0	0	0	0	0	0	0	0	0
1986	1	1	1	1	1	1	1	1	1	1	0	0	0
1987	1	1	1	1	1	1	1	1	1	1	1	1	1
1988	1	1	1	1	1	1	1	1	1	1	1	0	0
1989	1	0	0	0	0	0	0	0	0	0	0	0	0
1990	1	1	1	1	1	1	1	1	1	1	1	1	1
S_n	27	26	25	25	25	25	24	24	24	24	23	13	15

Sakarya havzasında 1962-1965, 1967-1968, 1970-1986, 1988-1989 yılları arasında taşkın meydana gelmiştir. Tablo 5.15'te, gerçekte taşkın meydana gelen yıllar, şerit halinde (satır) boyanarak gösterilmiştir. Bu yıllara denk gelen ve “taşkın yılı” olarak tahmin edilen “1” değerleri de koyu olarak belirtilmiştir. Örneğin, 0.0 düzeyi için 1962-1965, 1968, 1970-1975, 1977-1986, 1988 ve 1989 yılları “taşkın yılı”, 1966, 1969, 1987 ve 1990 yılları “önlem alındığı halde taşkın olmayan yıl”, 1967 ve 1976 yılları “felaket yılı” ve 1961 yılı sorunsuz yıl olarak ifade edilmektedir.

Tablodan açıkça görüleceği üzere, gerçekte taşkın meydana gelen yıllar (1967, 1976, 1985 ve 1989 yılları hariç), **0.0-0.5** düzeyleri arası için de “taşkın yılı” olarak tahmin edilmiştir. 1985 yılındaki taşkın **0.0-0.1** düzeylerinde, 1989 yılındaki taşkın ise yalnızca **0.0** düzeyinde tahmin edilmiştir. 0.6 düzeyinden itibaren tahmin edilen taşkınların sayısı azalmakta olup, ortalama ve ortanca düzeylerinde en düşük değeri almaktadır. Örneğin, S_n tekrar istatistiğine göre, 0.0 düzeyinde 27 kez önlem alındığında 23 taşkın yılı, 0.1 düzeyinde önlem sayısı 26'ya düşürüldüğünde 22 taşkın yılı doğru olarak tahmin edilebilmiştir.

Aynı şekilde, 0.5 düzeyinde 25 kez önlem alındığında 21 taşkın yılı, 0.6 düzeyinde önlem sayısı 24'e düşürüldüğünde 20 taşkın yılı doğru olarak tahmin edilebilmiştir.

Ortanca ve ortalama değerlerinde ise sırasıyla önlem sayısının 15 ve 13'e düşürülmesiyle birlikte yakalanan taşkın sayısı ortanca değeri için 11'e, ortalama değeri için 9'a düşerken, kaçırılan yıl sayısı ortanca değeri için 14'e, ortalama değeri için 16'ya çıkmaktadır.

Sakarya havzasındaki toplam taşkın yılı ile her bir düzeydeki toplam taşkın yılı ve S_n tekrar istatistiği kullanılarak, taşkın tahmin ve kaynakların etkin kullanım oranları hesaplanmış ve Tablo 5.16'da gösterilmiştir.

Tablo 5.16 Taşkın tahmin ve kaynakların etkin kullanım oranları

Düzy	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	Ort.	Ortc.
Kriter Değer	396.1	401	405.8	410.7	415.6	420.5	425.3	430.2	435.1	439.9	444.8	587.1	526
S _n	27	26	25	25	25	25	24	24	24	24	23	13	15
Taşkın tahmin oranı	0.920	0.880	0.840	0.840	0.840	0.840	0.800	0.800	0.800	0.800	0.760	0.360	0.440
Kaynakların etkin kullanım oranı	0.852	0.846	0.840	0.840	0.840	0.840	0.833	0.833	0.833	0.833	0.826	0.692	0.733

Tablo 5.16’da taşkın tahmin oranı, **0.0** düzeyi için %92, **0.1** düzeyi için %88, **0.2-0.5** düzeyleri arası %84, **0.6-0.9** düzeyleri arası %80, **1.0** düzeyi için %76, ortanca değeri için %44 ve ortalama değeri için %36 olarak hesaplanmıştır.

Taşkın tahmin oranının en yüksek olduğu ve aynı oran içerisinde kaynakların etkin kullanım oranının en yüksek olduğu düzey yalnızca 0.0 düzeyidir. Tabloda kaynakların etkin kullanım oranlarına bakıldığında, en etkin düzeyin 0.0 olduğu görülmektedir.

Dolayısıyla, karar verici tarafından Sakarya havzasında yaşanacak olası bir taşkın yerleşim yerlerine veya ekili arazilere zarar vereceği düşünülüyorsa, tercih edilen risk düzeyi **0.0** olarak belirlenmelidir. Eğer, Sakarya havzası ile ilgili alınan ya da ileriye dönük alınacak olan önlemlerin yeterli olduğu düşünülüyorsa, tercih edilen risk düzeyi, 0.1-0.5 düzeyleri arasından da seçilebilir.

Ancak, Sakarya havzası, Kızılırmak havzasına göre çok daha fazla sel-taşkın olaylarının yaşandığı bir havza olduğu için düzey belirleme konusunda çok fazla esnek davranılmamalı, hatta 0.1 düzeyinin üzerine çıkılmamalıdır. Özellikle, taşkın tahmin ve kaynakların etkin kullanım oranlarının en düşük olduğu ortalama ve ortanca değerleri hiçbir koşulda tercih edilmemelidir.

Sonuç olarak, klasik modelde sabit eşik değeri olarak kullanılan ortalama ve ortanca değerleri, sel-taşkın olaylarının yoğun olarak yaşandığı Sakarya havzası için kesinlikle önerilen düzeyler değildir.

5.4 İleriye Dönük Tahminlerin Yapılması

Mevcut veri ve bulgulardan yararlanılarak ileriye dönük tahminler yapmak için sel-taşkın olaylarının az sayıda yaşandığı havzalardan Küçük Menderes havzası ve yoğun olarak yaşandığı havzalardan Kızılırmak havzası örnek olarak seçilmiştir.

Seçilen havzalar için log-logistic(3P) dağılımının ve tekrar istatistiklerinin olasılık dağılımları kullanılarak kriter değerler için aşma olasılıkları ve ileriye dönük tahmin değerleri ayrı ayrı hesaplanmıştır.

Küçük Menderes havzası:

X (MYAMA) rasgele değişkeninin sahip olduğu üç parametrelili ($\alpha=3.2$, $\beta=223.5$ ve $\gamma=-28.1$) log-logistic(3P) dağılımının olasılık yoğunluk fonksiyonu kullanılarak, her düzeydeki kriter değerler için aşma olasılıkları hesaplanmış ve Tablo 5.17'de gösterilmiştir.

Tablo 5.17 Aşma olasılıkları

Düzye	$p(\text{önlem}) = P(X_i > X_1 + (X_0 - X_1)\gamma)$
0.0	$P(X > 188.5) = 0.525$
0.1	$P(X > 198.9) = 0.488$
0.2	$P(X > 209.3) = 0.452$
0.3	$P(X > 219.7) = 0.418$
0.4	$P(X > 230.1) = 0.387$
0.5	$P(X > 240.5) = 0.357$
0.6	$P(X > 250.9) = 0.330$
0.7	$P(X > 261.3) = 0.304$
0.8	$P(X > 271.7) = 0.281$
0.9	$P(X > 282.1) = 0.259$
1.0	$P(X > 292.5) = 0.240$
Ortalama	$P(X > 229.4) = 0.389$

Tablo 5.17’de, 0.0 düzeyinde 188.5 olan kriter değeri için aşma olasılığı yani önlem alınan yılların oranı %52.5 olarak hesaplanmıştır. Kriter değerler büyüdükçe, önlem alınan yıl sayısı azaldığından, $p(\text{önlem})$ oranları daha düşük değerler almaktadır.

Küçük Menderes havzası için kriter değerler, veri kümesinin %50-75’i arasından belirlendiğinden ortalama değeri için aşma olasılığı da diğer düzeylerdeki oranlara yakın değeri almaktadır.

Gerçek taşkın oranının %16.6 olduğu Küçük Menderes havzasında, her bir düzey için hesaplanan aşma olasılıklarının, gerçek oranın oldukça üzerinde yer aldığı görülmektedir. Çünkü, aşma olasılıkları olarak ifade edilen $p(\text{önlem})$ değerleri, yalnızca taşkın olan yılları değil, önlem alındığı halde taşkın olmayan yılları da içermektedir. Dolayısıyla, $p(\text{önlem})$ değerlerinin, önlem alındığı halde taşkın olmayan yılların getirdiği etkiden arındırılması için her bir düzeyde “taşkın yılı” olarak ifade edilen yılların, toplam yıla oranını veren $p(\text{taşkın})$ değerleri hesaplanmıştır. Hesaplanan bu değerler, $p(\text{önlem})$ ile çarpılarak $p(\text{düzeltilmiş})$ oranları elde edilmiştir.

$p(\text{düzeltilmiş})$ oranları, (2.8), (2.9), (2.10) ve (2.11) eşitlikleri ile verilen $G_{n,k}$, S_n , L_n ve T_k tekrar istatistiklerinin olasılık dağılımları içinde kullanılarak 1991-1997 yılları (7 yıllık) arası ileriye dönük taşkın tahmin olasılıkları hesaplanmıştır.

Tablo 5.18’de, Küçük Menderes havzası için Bölüm 5.3.1’de hesaplanan taşkın tahmin oranları (TTO) ve kaynakların etkin kullanım oranları (KEKO) ile $p(\text{önlem})$, $p(\text{taşkın})$, $p(\text{düzeltilmiş})$ ve ileriye dönük taşkın tahmin olasılıkları yer almaktadır.

Tekrar istatistiklerinin olasılık dağılımlarında kullanılan parametreler (n , n_1 , k , x), havzanın durumuna ve tahmin yapılacak yıl sayısına göre belirlenmiştir. Bu değerler karar vericinin tercihinin göre farklılık gösterebilir.

Tablo 5.18 1991-1997 yılları arası ileriye dönük taşkın tahmin olasılıkları

Düzyey:	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	Ort.
Kriter Değer	188.5	198.9	209.3	219.7	230.1	240.5	250.9	261.3	271.7	282.1	292.5	229.4
TTO	0.800	0.800	0.800	0.800	0.800	0.600	0.600	0.600	0.600	0.600	0.600	0.800
KEKO	0.267	0.286	0.308	0.308	0.364	0.333	0.333	0.333	0.333	0.429	0.429	0.333
p(önlem)	0.525	0.488	0.452	0.418	0.387	0.357	0.330	0.304	0.281	0.259	0.240	0.389
p(taşkın)	0.133	0.133	0.133	0.133	0.133	0.100	0.100	0.100	0.100	0.100	0.100	0.133
p(düz.)	0.070	0.065	0.060	0.056	0.052	0.036	0.033	0.030	0.028	0.026	0.024	0.052
$P(S_n \leq n_1)$ $n=7,$ $n_1=2$	0.990	0.990	0.994	0.994	0.996	0.998	0.999	0.999	0.999	0.999	1.000	0.996
$P(G_{n,k}=x)$ $n=7,$ $k=2,$ $x=1$	0.044	0.035	0.035	0.027	0.027	0.020	0.020	0.014	0.014	0.009	0.009	0.027
$P(L_n < k)$ $n=7,$ $k=2$	0.955	0.964	0.964	0.972	0.972	0.980	0.980	0.986	0.986	0.991	0.991	0.972
$P(T_k=x)$ $x=2,$ $k=1$	0.065	0.065	0.056	0.056	0.048	0.038	0.029	0.029	0.029	0.029	0.020	0.048

Tercih edilen risk düzeyinin **0.4** olarak belirlendiği Küçük Menderes havzası için Tablo 5.18’de yer alan ileriye dönük taşkın tahmin olasılıkları yorumlandığında,

“ n deneme sonundaki toplam başarı sayısı” anlamına gelen S_n tekrar istatistiği ile 7 yıl sonunda meydana gelebilecek toplam taşkın sayısının 2 ve 2’den küçük olma olasılığının %99.6 olduğu görülmektedir.

“ n deneme sonundaki “ k ” ve daha büyük uzunluklu başarı tekrarlarının sayısı” anlamına gelen $G_{n,k}$ tekrar istatistiği ile 7 yıl sonunda, ardarda (ardıl) meydana gelebilecek 2 ve 2’den büyük taşkınların oluşturduğu alt dizilerin sayısının “bir” olma olasılığının %2.7 olduğu görülmektedir.

“ n deneme sonundaki en uzun başarı tekrarının uzunluğu” anlamına gelen L_n tekrar istatistiği ile 7 yıl sonunda, ardarda meydana gelebilecek taşkınların oluşturduğu en uzun alt dizinin eleman sayısının 2'den küçük olma olasılığının %97.2 olduğu görülmektedir.

“ k uzunluklu bir başarı tekrarı elde edinceye kadar yapılan deneme sayısı” anlamına gelen T_k tekrar istatistiği ile 2. yılın sonunda bir kez taşkın olma olasılığının %4.8 olduğu görülmektedir.

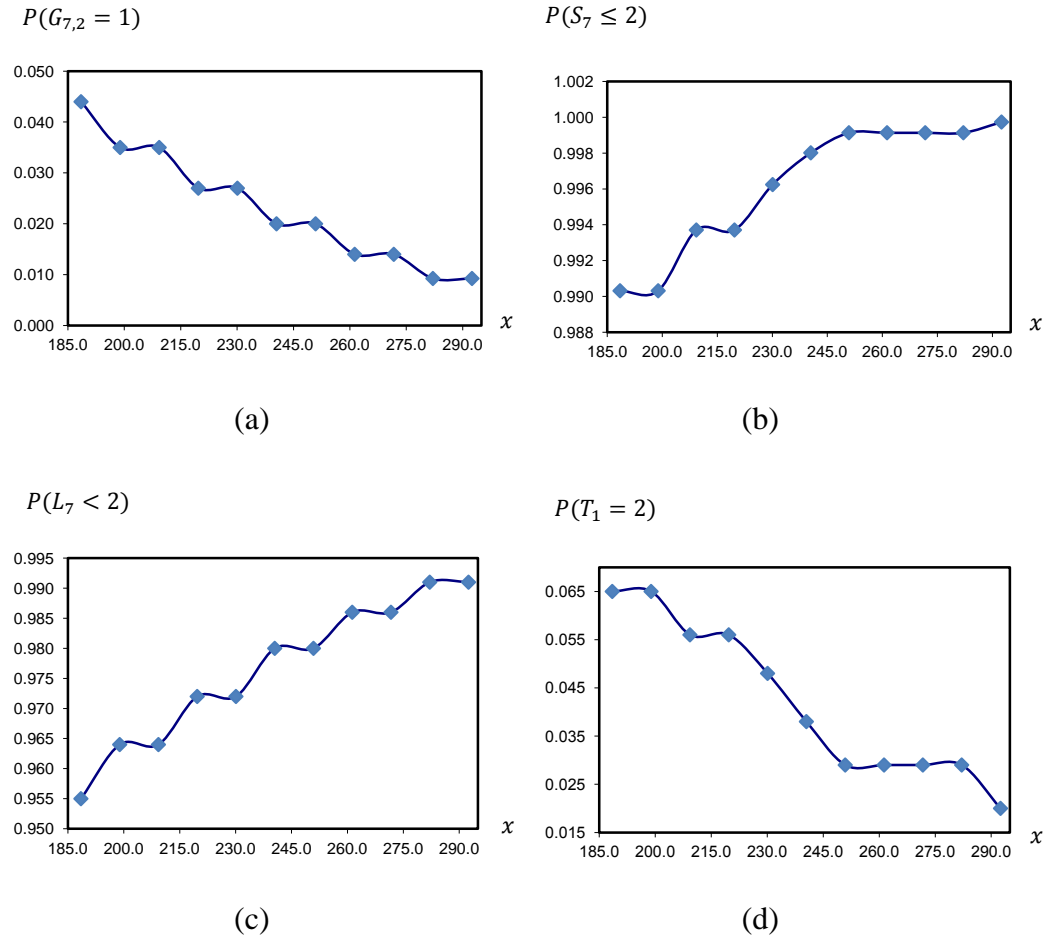
O halde, yapılan tahminler sonucunda 1991-1997 yılları arasında Küçük Menderes havzasında; S_n tekrar istatistiğine göre 2'den fazla taşkın olmayacağı, $G_{n,k}$ ve L_n tekrar istatistiklerine göre 2 taşkın olsa bile ardarda yıllarda meydana gelmeyeceği ve T_k tekrar istatistiğine göre olabilecek ilk taşkın 1991 ve 1992 yıllarında olmayacağı söylenebilir.

Taşkın yıllıkları incelendiğinde; 1961-1990 yılları arasında 5 kez taşkın meydana gelen havzada, 1991-1997 yılları arasında yalnızca 1 kez (1995 yılı) taşkın meydana geldiği görülmüştür. Dolayısıyla, yapılan tahminler gerçek verileri doğrular niteliktedir.

$G_{n,k}$, S_n , L_n ve T_k tekrar istatistiklerinin olasılık dağılımlarından elde edilen taşkın tahmin olasılıkları ile kriter değerler arasındaki grafiksel ilişki ise Şekil 5.6'da verilmiştir.

Şekil 5.6'da, x eksenleri hesaplanan kriter değerleri, y eksenleri (a)'da $P(G_{7,2} = 1)$ olasılık değerlerini, (b)'de $P(S_7 \leq 2)$ olasılık değerlerini, (c)'de $P(L_7 < 2)$ olasılık değerlerini ve (d)'de ise $P(T_1 = 2)$ olasılık değerlerini göstermektedir.

Şekil 5.6 a) $G_{n,k}$, b) S_n , c) L_n , d) T_k tekrar istatistiklerine göre taşkın tahmin olasılıkları ile kriter değerler arasındaki ilişki



(a) - (d) grafiklerinde, en düşük risk düzeyine (0.0) karşılık gelen kriter değerde (188.5), %4.4 değerini alan $P(G_{7,2} = 1)$ ve %6.5 değerini alan $P(T_1 = 2)$ olasılık değerleri, risk düzeyi arttıkça (kriter değerler büyüdükçe) azalma göstermektedir.

(b) - (c) grafiklerinde, en düşük risk düzeyine (0.0) karşılık gelen kriter değerde (188.5), %99 değerini alan $P(S_7 \leq 2)$ ve %95.5 değerini alan $P(L_7 < 2)$ olasılık değerleri ise risk düzeyi arttıkça (kriter değerler büyüdükçe) artış göstermektedir.

(b) grafiğinde, $P(S_7 \leq 2)$ olasılık değerinin, risk düzeyi arttıkça artış göstermesinin nedeni, S_n tekrar istatistiğinin 2'den küçük eşit olma olasılığının araştırılmasından kaynaklanmaktadır. Risk düzeyinin artmasıyla, yakalanan taşkın sayısında bir azalma olacağından, taşkın sayısının 2'den küçük olma olasılığı daima artacaktır. Eğer, S_n tekrar istatistiğinin 2'ye eşit olma olasılığı

araştırılıyorsa, risk düzeyi arttıkça yakalanan taşkın sayısı azalacağından olasılık değerinin de azalması beklenirdi.

Kızılırmak havzası:

X (MYAMA) raslantı değişkeninin sahip olduğu üç parametrelili ($\alpha=2.8$, $\beta=367.3$ ve $\gamma=335.4$) log-logistic(3P) dağılımının olasılık yoğunluk fonksiyonu kullanılarak, her düzeydeki kriter değerler için aşma olasılıkları hesaplanmış ve Tablo 5.19'da gösterilmiştir.

Tablo 5.19'da, 0.0 düzeyi için 473.3 olan kriter değer için aşma olasılığı yani önlem alınan yılların oranı %94 olarak hesaplanmıştır. Kriter değerler büyüdükçe, önlem alınan yıl sayısı azaldığından, $p(\text{önlem})$ oranları daha düşük değerler olacaktır.

Kızılırmak havzası için kriter değerler, veri kümesinin %10-25'i arasından belirlendiği için ortalama ve ortanca değerlerinin aşma olasılıkları, diğer düzeylerdeki oranlara göre çok düşüktür.

Tablo 5.19 Aşma olasılıkları

Düzye	$p(\text{önlem}) = P(X_i > X_1 + (X_0 - X_1)\gamma)$
0.0	$P(X > 473.3) = 0.940$
0.1	$P(X > 482.9) = 0.928$
0.2	$P(X > 492.4) = 0.915$
0.3	$P(X > 502.0) = 0.901$
0.4	$P(X > 511.5) = 0.887$
0.5	$P(X > 521.1) = 0.871$
0.6	$P(X > 530.6) = 0.854$
0.7	$P(X > 540.2) = 0.837$
0.8	$P(X > 549.7) = 0.819$
0.9	$P(X > 559.3) = 0.800$
1.0	$P(X > 568.8) = 0.780$
Ortalama	$P(X > 772.2) = 0.380$
Ortanca	$P(X > 757.0) = 0.400$

Gerçek taşkın oranının %56.6 olduğu Kızılırmak havzasında, ortalama ve ortanca değerleri dışındaki düzeyler için hesaplanan aşma olasılıklarının, gerçek oranın oldukça üzerinde yer aldığı görülmektedir. Çünkü, aşma olasılıkları olarak ifade edilen $p(\text{önlem})$ değerleri, yalnızca taşkın olan yılları değil, önlem alındığı halde taşkın olmayan yılları da içermektedir.

Dolayısıyla, $p(\text{önlem})$ değerlerinin, önlem alındığı halde taşkın olmayan yılların getirdiği etkiden arındırılması için her bir düzeyde “taşkın yılı” olarak ifade edilen yılların, toplam yıla oranını veren $p(\text{taşkın})$ değerleri hesaplanmıştır. Hesaplanan bu değerler, $p(\text{önlem})$ ile çarpılarak $p(\text{düzeltilmiş})$ oranları elde edilmiştir.

$p(\text{düzeltilmiş})$ oranları (2.8), (2.9), (2.10) ve (2.11) eşitlikleri ile verilen $G_{n,k}$, S_n , L_n ve T_k tekrar istatistiklerinin olasılık dağılımları içinde kullanılarak, 1991-1997 yılları (7 yıllık) arası ileriye dönük taşkın tahmin oranları hesaplanmıştır.

Tablo 5.20’de, Kızılırmak havzası için Bölüm 5.3.3’de hesaplanan taşkın tahmin oranları (TTO) ve kaynakların etkin kullanım oranları (KEKO) ile $p(\text{önlem})$, $p(\text{taşkın})$, $p(\text{düzeltilmiş})$ ve ileriye dönük taşkın tahmin oranları yer almaktadır.

Tekrar istatistiklerinin olasılık dağılımlarında kullanılan parametreler (n , n_1 , k , x), havzanın durumuna ve tahmin yapılacak yıl sayısına göre belirlenmiştir. Bu değerler karar vericinin tercihine göre farklılık gösterebilir. Ayrıca, parametrelere farklı değerler verilerek, aynı tekrar istatistiği için birden fazla hesaplamalar yapılmıştır.

Tablo 5.20 1991-1997 yılları arası ileriye dönük taşkın tahmin oranları

Düzyey:	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	Ort.	Ortc.
Kriter Değer:	473.3	482.9	492.4	502.0	511.5	521.1	530.6	540.2	549.7	559.3	568.8	772.2	757.0
TTO	0.882	0.882	0.882	0.882	0.882	0.882	0.882	0.824	0.765	0.765	0.765	0.588	0.588
KEKO	0.556	0.556	0.556	0.556	0.577	0.577	0.577	0.583	0.565	0.565	0.565	0.714	0.667
p(önlem)	0.940	0.928	0.915	0.901	0.887	0.871	0.854	0.837	0.819	0.800	0.780	0.380	0.400
p(taşkın)	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.467	0.433	0.433	0.433	0.333	0.333
p(düz.)	0.470	0.464	0.458	0.451	0.443	0.435	0.427	0.391	0.355	0.347	0.338	0.127	0.133
$P(S_n \leq n_1)$ $n=7,$ $n_1=3$	0.565	0.587	0.587	0.608	0.629	0.629	0.650	0.729	0.800	0.800	0.816	0.993	0.993
$P(S_n \leq n_1)$ $n=7,$ $n_1=4$	0.820	0.834	0.834	0.847	0.860	0.860	0.872	0.913	0.944	0.944	0.951	0.999	0.999
$P(S_n \leq n_1)$ $n=7,$ $n_1=5$	0.955	0.960	0.960	0.964	0.968	0.968	0.972	0.984	0.991	0.991	0.992	1.000	1.000
$P(G_{n,k}=x)$ $n=7,$ $k=2,$ $x=1$	0.569	0.560	0.560	0.550	0.539	0.539	0.528	0.480	0.424	0.424	0.409	0.088	0.088
$P(G_{n,k}=x)$ $n=7,$ $k=2,$ $x=2$	0.119	0.112	0.112	0.105	0.098	0.098	0.092	0.068	0.048	0.048	0.044	0.001	0.001
$P(L_n < k)$ $n=7,$ $k=2$	0.312	0.329	0.329	0.345	0.363	0.363	0.380	0.452	0.528	0.528	0.547	0.911	0.911
$P(L_n < k)$ $n=7,$ $k=3$	0.682	0.698	0.698	0.713	0.728	0.728	0.743	0.798	0.847	0.847	0.858	0.990	0.990
$P(T_k=x)$ $x=4,$ $k=2$	0.117	0.114	0.114	0.111	0.108	0.108	0.105	0.093	0.080	0.080	0.076	0.015	0.015

Tercih edilen risk düzeyinin 0.4-0.6 arası olarak belirlendiği Kızılırmak havzası için Tablo 5.20'de yer alan ileriye dönük taşkın tahmin oranları yorumlandığında;

“*n deneme sonundaki toplam başarı sayısı*” anlamına gelen S_n tekrar istatistiği ile 7 yıl sonunda meydana gelebilecek toplam taşkın sayısının 3 ve 3'den küçük olma olasılığının %62'nin üzerinde, 4 ve 4'den küçük olma olasılığının %85'in üzerinde, 5 ve 5'ten küçük olma olasılığının ise %96'nın üzerinde olduğu görülmektedir.

“*n deneme sonundaki “k” ve daha büyük uzunluklu başarı tekrarlarının sayısı*” anlamına gelen $G_{n,k}$ tekrar istatistiği ile 7 yıl sonunda, ardarda meydana gelebilecek 2 ve 2'den büyük taşkınların oluşturduğu alt dizilerin sayısının 1 olma olasılığının %52'nin üzerinde, 2 olma olasılığının ise %10'nun altında olduğu görülmektedir.

“*n deneme sonundaki en uzun başarı tekrarının uzunluğu*” anlamına gelen L_n tekrar istatistiği ile 7 yıl sonunda, ardarda meydana gelebilecek taşkınların oluşturduğu en uzun alt dizinin eleman sayısının 2'den küçük olma olasılığının %40'ın altında, 3'den küçük olma olasılığının ise %72'nin üzerinde olduğu görülmektedir.

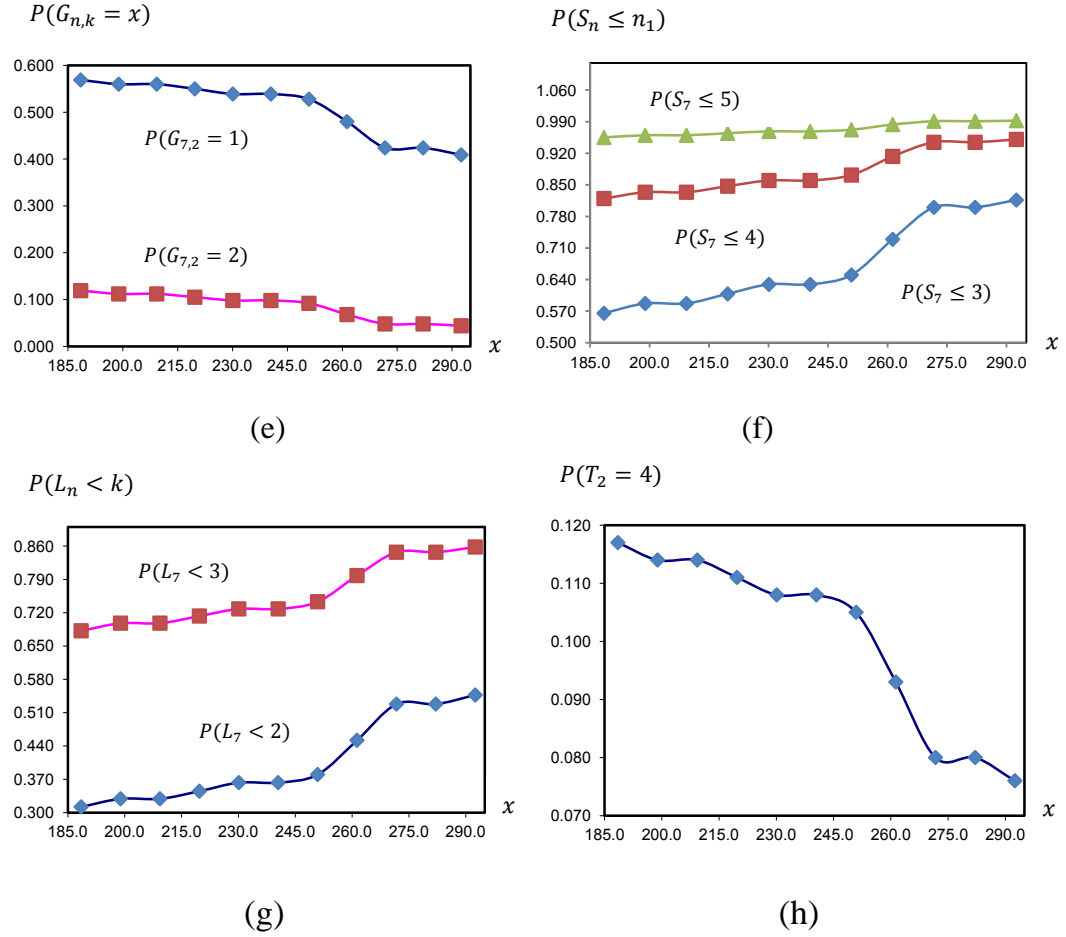
“*k uzunluklu bir başarı tekrarı elde edinceye kadar yapılan deneme sayısı*” anlamına gelen T_k tekrar istatistiği ile 4. yılın sonunda ardarda iki kez taşkın olma olasılığının %11'in altında olduğu görülmektedir.

O halde, yapılan tahminler sonucunda 1991-1997 yılları arasında Kızılırmak havzasında; S_n tekrar istatistiğine göre 5'den fazla taşkın olmayacağı, $G_{n,k}$ ve L_n tekrar istatistiklerine göre en fazla iki yıl ardarda taşkın meydana geleceği ve bu durumun 7 yılda yalnızca 1 kez gerçekleşeceği, T_k tekrar istatistiğine göre de iki yıl ardarda yaşanacak olası bir taşkın 1993-1994 yıllarında olmayacağı söylenebilir.

Taşkın yıllıkları incelendiğinde; 1961-1990 yılları arasında 17 kez taşkın meydana gelen havzada, 1991-1997 yılları arasında 3 kez (1992, 1995 ve 1996 yılları) taşkın meydana geldiği görülmüştür. Dolayısıyla, yapılan tahminler gerçek verileri doğrular niteliktedir.

$G_{n,k}$, S_n , L_n ve T_k tekrar istatistiklerinin olasılık dağılımlarından elde edilen taşkın tahmin olasılıkları ile kriter değerler arasındaki grafiksel ilişki ise Şekil 5.7’de verilmiştir.

Şekil 5.7 e) $G_{n,k}$, f) S_n , g) L_n , h) T_k tekrar istatistiklerine göre taşkın tahmin olasılıkları ile kriter değerler arasındaki ilişki



Şekil 5.7’de, x eksenleri hesaplanan kriter değerleri, y eksenleri (e)’de $P(G_{7,2} = 1)$, $P(G_{7,2} = 2)$ olasılık değerlerini, (f)’de $P(S_7 \leq 3)$, $P(S_7 \leq 4)$, $P(S_7 \leq 5)$ olasılık değerlerini, (g)’de $P(L_7 < 2)$, $P(L_7 < 3)$ olasılık değerlerini ve (h)’de ise $P(T_2 = 4)$ olasılık değerlerini göstermektedir.

Her bir grafik kendi içerisinde incelendiğinde,

(e) grafiğinde, 7 yıl sonunda ardarda meydana gelebilecek 2 ve 2’den büyük taşkınların oluşturduğu alt dizilerin sayısının 1 olması olasılığının ($P(G_{7,2} = 1)$), iki olması olasılığına ($P(G_{7,2} = 2)$) göre daha yüksek değerler aldığı ve risk

düzeyi arttıkça (kriter değerler büyüdükçe) her iki olasılık değerinde de bir azalma olduğu görülmektedir.

(f) grafiğinde, en düşük risk düzeyine (0.0) karşılık gelen kriter değerinde (188.5), 7 yıl sonunda meydana gelebilecek toplam taşkın sayısının 3 ve 3'ten küçük olma olasılığının ($P(S_7 \leq 3)$) %56.5, 4 ve 4'ten küçük olma olasılığının ($P(S_7 \leq 4)$) %82, 5 ve 5'ten küçük olma olasılığının ($P(S_7 \leq 5)$) %95.5 değerini aldığı görülmektedir. Toplam taşkın sayısının (n_1) artmasıyla, $P(S_n \leq n_1)$ olasılığı daha yüksek değerler almaktadır. Risk düzeyinin artmasıyla (kriter değerler büyüdükçe) her üç olasılık değerinde de bir artış gözlenmektedir. Ancak, n_1 değerinin artması durumunda, olasılık değerlerindeki değişim daha az olmaktadır.

(g) grafiğinde, 7 yıl sonunda ardarda meydana gelebilecek taşkınların oluşturduğu en uzun alt dizinin eleman sayısının 3'ten küçük olma ($P(L_7 = 3)$) olasılığının, 2'den küçük olma ($P(L_7 = 2)$) olasılığına göre daha yüksek değerler aldığı ve risk düzeyi arttıkça (kriter değerler büyüdükçe) her iki olasılık değerinde de bir artış olduğu görülmektedir.

(h) grafiğinde, en düşük risk düzeyine (0.0) karşılık gelen kriter değerinde (188.5), %11.7 değerini alan $P(T_2 = 4)$ olasılık değerleri, risk düzeyi arttıkça (kriter değerler büyüdükçe) azalma göstermektedir.

5.5 Havzaların Karşılaştırılması

Önceki bölümlerde, Küçük Menderes ve Gediz havzaları için veri kümesinin %50-75'i arasında kalan değerler ile Kızılırmak ve Sakarya havzaları için veri kümesinin %10-25'i arasında kalan değerler, "kriter değer" olarak seçilmiş ve risk düzeyi bazında bir modelleme yapılmıştı.

Yapılan modellemede, her havza ve risk düzeyi için taşkın tahmin oranları ve kaynakların etkin kullanım oranları ayrı ayrı hesaplanarak, tercih edilen risk düzeyleri belirlenmişti.

Örnek olarak seçilen Küçük Menderes ve Kızılırmak havzaları için aşma olasılıkları hesaplanmış ve tekrar istatistiklerinin olasılık dağılımları kullanılarak ileriye dönük tahminler yapılmıştı.

Bu bölümde ise, her havza ve risk düzeyi için hesaplanan taşkın tahmin oranları (TTO), kaynakların etkin kullanım oranları (KEKO) ile $p(\text{önlem})$, $p(\text{taşkın})$ ve $p(\text{düzeltilmiş})$ oranları, Tablo 5.21’de bir arada gösterilerek, havzaların birbiri ile kıyaslanmasına olanak sağlanmıştır.

Tablo 5.21’de kullanılan değişkenler aşağıda kısaca özetlenmiştir.

Havza Taşkın Oranı (HTO): Taşkın yıllıklarından elde edilen gerçek taşkın oranı.

Taşkın Tahmin Oranı (TTO): Risk düzeyi bazında toplam “taşkın yılı”nın, havzadaki toplam “taşkın yılı”na oranı.

Kaynakların Etkin Kullanım Oranı (KEKO): Risk düzeyi bazında toplam “taşkın yılı”nın, önlem alınan yılların toplamına (S_n) oranı.

$p(\text{önlem})$: Risk düzeyi bazında önlem alınan yılların toplamının, toplam yıl sayısına oranı. (Log-logistic dağılımının olasılık fonksiyonundan hesaplanır.)

$p(\text{taşkın})$: Risk düzeyi bazında toplam “taşkın yılı”nın, toplam yıl sayısına oranı.

$p(\text{düzeltilmiş})$: Risk düzeyi bazında önlem alınan yıllar içinde taşkın olan yılların oranı. ($p(\text{önlem})$ ile $p(\text{taşkın})$ değerlerinin çarpılması ile hesaplanır.)

Tablo 5.21’de, “tercih edilen risk düzeyi” olarak, Küçük Menderes havzası için “0.4” düzeyi, Gediz havzası için “0.0” düzeyi, Kızılırmak havzası için “0.4-0.6” düzeyleri ve Sakarya havzası için “0.0” düzeyi kullanılmıştır.

Tabloda, tercih edilen risk düzeyleri şerit halinde (sütun) boyanarak gösterilmiş ve bu alanlarda yer alan olasılık değerleri de koyu olarak belirtilmiştir. Karar vericinin tercih edebileceği alternatif risk düzeyleri ise yalnızca şerit halinde (sütun) boyanarak gösterilmiştir.

Tablo 5.21 Havzaların karşılaştırılması

	Oranlar	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	Ort.	Ortc.
K. Menderes (HTO: 0.166)	TTO	0.800	0.800	0.800	0.800	0.800	0.600	0.600	0.600	0.600	0.600	0.600	0.800	0.800
	KEKO	0.267	0.286	0.308	0.308	0.364	0.333	0.333	0.333	0.333	0.429	0.429	0.333	0.267
	P (önlem)	0.525	0.488	0.452	0.418	0.387	0.357	0.330	0.304	0.281	0.259	0.240	0.389	0.525
	P (taşkın)	0.133	0.133	0.133	0.133	0.133	0.100	0.100	0.100	0.100	0.100	0.100	0.133	0.133
	P (düz.)	0.070	0.065	0.060	0.056	0.052	0.036	0.033	0.030	0.028	0.026	0.024	0.052	0.070
Gediz (HTO: 0.233)	TTO	0.714	0.571	0.571	0.571	0.571	0.571	0.571	0.571	0.429	0.429	0.429	0.571	0.714
	KEKO	0.333	0.308	0.333	0.333	0.333	0.364	0.364	0.400	0.375	0.375	0.429	0.333	0.333
	P (önlem)	0.479	0.452	0.426	0.401	0.377	0.354	0.333	0.313	0.295	0.277	0.260	0.395	0.479
	P (taşkın)	0.167	0.133	0.133	0.133	0.133	0.133	0.133	0.133	0.100	0.100	0.100	0.133	0.167
	P (düz.)	0.080	0.060	0.057	0.053	0.050	0.047	0.044	0.042	0.029	0.028	0.026	0.053	0.080
Kızılırmak (HTO: 0.566)	TTO	0.882	0.882	0.882	0.882	0.882	0.882	0.882	0.824	0.765	0.765	0.765	0.588	0.588
	KEKO	0.556	0.556	0.556	0.556	0.577	0.577	0.577	0.583	0.565	0.565	0.565	0.714	0.667
	P (önlem)	0.940	0.928	0.915	0.901	0.887	0.871	0.854	0.837	0.819	0.800	0.780	0.380	0.400
	P (taşkın)	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.467	0.433	0.433	0.433	0.333	0.333
	P (düz.)	0.470	0.464	0.458	0.451	0.443	0.435	0.427	0.391	0.355	0.347	0.338	0.127	0.133
Sakarya (HTO: 0.833)	TTO	0.920	0.880	0.840	0.840	0.840	0.840	0.800	0.800	0.800	0.800	0.760	0.360	0.440
	KEKO	0.852	0.846	0.840	0.840	0.840	0.840	0.833	0.833	0.833	0.833	0.826	0.692	0.733
	P (önlem)	0.881	0.872	0.862	0.852	0.842	0.831	0.820	0.809	0.797	0.785	0.773	0.398	0.551
	P (taşkın)	0.767	0.733	0.700	0.700	0.700	0.700	0.667	0.667	0.667	0.667	0.633	0.300	0.367
	P (düz.)	0.675	0.639	0.604	0.597	0.589	0.582	0.547	0.539	0.531	0.524	0.490	0.119	0.202

“Tercih edilen risk düzeyi”ne göre taşkın tahmin oranı, Küçük Menderes havzası için %80, Gediz havzası için %71.4, Kızılırmak havzası için %88.2 ve Sakarya havzası için %92 olarak hesaplanmıştır.

Küçük Menderes ve Gediz havzalarında taşkın sayısının diğer iki havzaya göre daha az olması, gerçekte yaşanan tüm taşkınların tahmin edilebilme

olasılığını düşürmektedir. Bu nedenle, sel ve taşkın olaylarının yoğun olarak yaşandığı havzalarda taşkın tahmin oranı daha yüksek çıkmaktadır.

Küçük Menderes ve Gediz havzalarında diğer düzeylere göre yakın değerler alan ve “tercih edilen risk düzeyi” olarak ta kullanılabilen “ortalama” ve “ortanca” değerleri, Kızılırmak ve Sakarya havzalarında kesinlikle kullanılmamalıdır. Çünkü, bu tür havzalarda kriter değerler, veri kümesinin %10-25’i arasından belirlendiği için ortalama ve ortanca, bu değerlerin çok üzerinde kalmaktadır. Dolayısıyla, tercih edilen risk düzeyinde %85’in üzerine çıkabilen taşkın tahmin oranları, ortalama ve ortanca değerleri için %60’ın altına düşmektedir.

Tüm havza ve risk düzeylerinde, $p(\text{önlem})$ oranları, taşkın olan yıllar ile önlem alındığı halde taşkın olmayan yılların birlikte değerlendirilmesinden dolayı, HTO’dan yüksek değerler almıştır. $p(\text{önlem})$ değerlerinin bu haliyle ileriye dönük tahminlerde kullanılması -her geçen yıl daha çok taşkın yaşanacağı anlamına geleceğinden- doğru olmayacaktır. Oysaki, yaşanan her bir taşkın sonrası, yetkililer tarafından mevcut önlemler artırılarak ya da yeni önlemler alınarak, havzada taşkın olma oranı düşürülmeye çalışılmaktadır.

Özellikle, DSİ tarafından, sel ve taşkın olaylarının önlenmesi veya etkilerinin azaltılabilmesi için yapısal önlemler içeren projeli faaliyetler ile taşkın olayının her aşamasında iyileştirme çalışmaları yapılmaktadır. Bu çalışmalar çerçevesinde, taşkınların önlenmesi amacıyla barajlar ve taşkından korunma tesisleri inşa edilmekte, akarsu yataklarında ise düzenlenme ve ıslah çalışmaları yapılmaktadır. (Uşkay ve Aksu, 2002)

Alınan önlemler nedeniyle, ileriye dönük taşkın tahminleri için mevcut HTO’dan çok daha düşük oranlara ihtiyaç duyulmaktadır. $p(\text{düzeltilmiş})$ oranları, tüm havzalarda HTO’nun %10-15 altında değerler almaktadır.

Dolayısıyla, ileriye dönük tahminler için $p(\text{düzeltilmiş})$ oranlarının kullanılması, -havza genelinde alınan önlemlerin de göz ardı edilmemesi bakımından- çok daha doğru ve tutarlı sonuçlar ortaya koyacaktır.

6. SONUÇ

Literatürde “run” olarak bilinen tekrar kavramı üst üste gelen aynı tür elemanların oluşturduğu dizinin, diğer elemanlarla sonlanması ile oluşan bir alt dizi olarak tanımlanır.

Başarı-başarısız, sağlam-bozuk gibi tanımlamaların “1” ve “0” elemanlarıyla gösterildiği iki değerli rasgele dizilerle; hidroloji, moleküler biyoloji, istatistiksel kalite kontrol, güvenilirlik ve rasgelelik testleri gibi çeşitli sistemlerde karşılaşılmakta ve bu dizilere dayalı olarak tanımlanan tekrar istatistikleri, ilgili sisteme ilişkin karar verme süreçlerinde kullanılmaktadır.

Literatürde, iki değerli diziler üzerinde tanımlanan tekrar istatistikleri, sürekli değerli dizilerin analizinde de kullanılmıştır. Sürekli değerli diziler üzerinde tekrar istatistiklerinin tanımlanabilmesi için dizi elemanlarının belirli kriter değerler kullanılarak iki değerli hale dönüştürülmesi gerekmektedir.

Örneğin, X_1, X_2, \dots, X_n sürekli değerlerden oluşan rasgele bir dizi olsun. Literatürde var olan klasik modelde, dizi elemanları, ilgilenilen sistemin eşik değeri (trancation level) olarak adlandırılan bir x_0 kriter değeri ile karşılaştırılarak iki değerli rasgele bir diziye dönüştürülebilir. Birçok çalışmada, ortalama, ortanca ve tepe değeri gibi veri kümesini özetleyen konum ölçüleri, x_0 kriter değeri olarak kullanılmıştır.

Çoğunlukla “ortanca” değeri olarak tercih edilen x_0 kriter değerinden büyük olan X_i değerleri için ξ_i ($i = 1, 2, \dots, n$) rasgele değişkeni “1” değerini alırken, x_0 kriter değerinden küçük veya eşit olan X_i değerleri için de “0” değerini alır. İlgilenilen sisteme göre “1” değeri; başarı, sağlam, sulak gibi tanımlamaları ifade ederken, “0” değeri başarısız, bozuk, kurak gibi “1” değerindeki tanımlamaların karşıt anlamlılarını ifade eder.

Ancak, tek bir kriter değerinin, “eşik değeri” olarak kullanılmasının bazı sakıncaları bulunmaktadır. Örneğin, klasik modelde, x_0 değerinden çok az büyük bir X_i değeri için ξ_i ($i = 1, 2, \dots, n$) rasgele değişkeni “1” değerini alırken, x_0 değerinden çok az küçük bir X_i değeri için ξ_i ($i = 1, 2, \dots, n$) rasgele değişkeni “0” değerini almaktadır. Oysaki, her iki X_i değeri arasında mutlak değer olarak çok fazla bir fark bulunmamasına rağmen farklı sonuçların ortaya çıkması, özellikle ileriye dönük tahminlerde yanlış yönlendirmelere neden olabilmektedir.

Bu nedenle, x_0 gibi sabit bir eşik değerinin belirlenmesi yerine, “0.0” ile “1.0” arasında değer alabilen bir γ değişkeni tanımlanarak, karar vericinin tercihinine göre değişebilen bir kriter değerinin belirlenmesi doğru bir yaklaşım olacaktır.

Bu yaklaşımla klasik modele alternatif olarak sunulan ve “önerilen model” olarak adlandırılan modelde, sürekli değerli diziyi iki değerli hale dönüştürmek için x_0 yerine, “0.0” ile “1.0” arasında değer alabilen γ değişkenine bağlı olarak değişen $x_1 + (x_0 - x_1) \gamma$ değerinin kriter değeri olarak kullanılması modele esneklik kazandırmaktadır.

Risk düzeyi olarak adlandırılan γ değişkeninin alacağı onbir farklı değer (düzeyler arası aralık genişliği 0.1 olarak alındığında) için hesaplanan kriter değerleri arasında, araştırılan sistemin özelliklerine göre “tercih edilen risk düzeyi”nin seçilmesi, hem modeli daha kullanışlı hale getirmekte, hem de karar vericiye sistemle ilgili tercih yapabilmesi için farklı alternatifler sunmaktadır.

Hidroloji alanından bir uygulama yapılan bu tez çalışmasında, gerçek veriler kullanıldığında, “klasik model”in hem mevcut veriler üzerinde, hem de ileriye dönük tahmin süreçlerinde yetersiz kaldığı ve bu süreçte, kriter değerlerinin birçok faktör göz önünde bulundurularak sistemin özelliklerine göre belirlendiği “önerilen model”in daha avantajlı olduğu açıkça görülmektedir.

Sonuç olarak, karar vericiye sistemle ilgili tercih yapabilmesi için farklı alternatifler sunan ve tahmin süreçlerinde de doğru ve tutarlı sonuçların elde edilmesini sağlayan “önerilen model”, ilgilenilen değişkenin özelliklerine göre tanımlanarak, tekrar istatistiklerinin diğer uygulama alanlarında da rahatlıkla kullanılabilir.

KAYNAKLAR DİZİNİ

- Aki, S.**, 1999, Distributions of runs and consecutive systems on directed trees, *Annals of the Institute of Statistical Mathematics*, 51, 1 – 15.
- Aki, S. and Hirano, K.**, 1988, Some characteristics of the binomial distribution of order k and related distributions, *In Statistical Theory and Data Analysis II*, 211 – 222, Elsevier Science Publishers, B.V., North – Holland.
- Aki, S. and Hirano, K.**, 1996, Lifetime distribution and estimation problems of consecutive – k – out – of – n : F systems, *Annals of the Institute of Statistical Mathematics*, 48, 185 – 199.
- Arratia, R. and Waterman, M.S.**, 1985, Critical phenomena in sequence matching, *Annals of Probability*, 13, 1236 – 1249.
- Balakrishnan, N. and Koutras, M.V.**, 2002, *Runs and Scans with Applications*. John Wiley&Sons.
- Bayazıt, M., Avcı, İ. ve Şen, Z.**, 2009, *Hidroloji Uygulamaları*, Birsen Yayınevi, İstanbul, 286s.
- Boland, P.J. and Papastavridis, S.**, 1999, Consecutive k – out – of – n : F systems with cycle k , *Statistics & Probability Letters*, 44, 155– 160.
- Brownlee, K. A.**, 1965, *Statistical theory and methodology in science and engineering*, 2nd ed. John Wiley&Sons., New York, 570 p.
- Büyükkaracığan, N. ve Kahya, E.**, 2007, Taşkın frekans analizinde kullanılan değişik dağılımların Konya havzası yıllık pik akım serilerine uygulanıp karşılaştırılması, Konferans Bildirisi, Uluslararası Küresel İklim Değişikliği ve Çevresel Etkileri Konferansı.
- Büyükkaracığan, N. ve Kahya, E.**, 2009, Konya havzası akarsuları yıllık pik akım serilerinin taşkın frekans analizi, *Teknik-Online dergi*, 8(3):246-261.
- Chakraborti, S.**, 2000, Run length, average run length and false alarm rate of Shewhart \bar{X} – bar chart: Exact derivations by conditioning, *Commun. Statist. – Simula.*, 29(1), 61 – 81.
- Champ, C.W. and Woodall, W.H.**, 1987, Exact results for Shewhart control charts with supplementary runs rules, *Technometrics*, 29, Vol.4, 393 – 399.
- Chao, M.T., Fu, J.C. and Koutras, M.V.**, 1995, A survey of the reliability studies of consecutive – k – out – of – n : F systems and its related systems, *IEEE Transactions on Reliability*, 44, 120 – 127.

KAYNAKLAR DİZİNİ (devam)

- Chiang, D.T. and Niu, S.**,1981, Reliability of consecutive – k – out – of – n: F system, *IEEE Transactions on Reliability*, 30, 87 – 89.
- Choi, J. S., Hong S., Chi, S. B., Lee, H. B., Park, C. K., Kim, H. W., Yeu, T. K. and Lee, T. H.**, 2011, Probability distribution for the shear strength of seafloor sediment in the KR5 area for the development of manganese nodule miner, *Ocean Engineering*, 29, Vol.38, 2033 – 2041.
- Chvátal, V. and Sankoff, D.**, 1975, Longest common subsequences of two random sequences, *Journal of Applied Probability*, 12, 306 – 315.
- Chryssaphinou, O., Papastavridis, S. and Tsapelas,, T.**, 1993, On the number of overlapping success runs in a sequence of independent Bernoulli trials, *Applications of Fibonacci Numbers*, 103–112, Kluwer Academic Publishers, Amsterdam, The Netherlands.
- Clark, N.**, 2007, “Liquid crystal phases of tiny DNA molecules point up new scenario for first life on Earth”, http://www.eurekalert.org/pub_releases/2007-11/uoca-lcp112007.php (Erişim tarihi: 7 Mart 2012)
- Dao, H.M. and Silio, C.B., Jr.**, 1998, Ring network with a constrained number of consecutively – bypassed stations, *IEEE Transactions on Reliability*, 47, 35 – 43.
- De Moivre, A.**, 1738, *The Doctrine of Chances: A Method of Calculating the Probability of Events in Play* (second edition).
- Durman, B. ve Pakdil, F.**, 2012, “İstatistikî proses kontrol uygulamaları için bir sistem tasarımı”, <http://www.ekonometridernegi.org/bildiriler/o18s3.pdf> (Erişim tarihi: 02 Şubat 2012)
- Eğrioğlu, E.**, 2010, Bulanık İstatistik Ders Notları, Ondokuz Mayıs Üniversitesi.
- Em, A., Hamidi, N. ve Toprak, Z. F.**, 2007, GAP bölgesinde toplam yağışların değişimi ve homojenlik analizi, Kongre Bildirisi, 1. Türkiye İklim Değişikliği Kongresi.
- Eryılmaz, S.**, 2005b, The longest run statistic associated with exchangeable binary variables, *Turkish Journal of Engineering and Environmental Sciences*, 29, 105 – 112.
- Eryılmaz, S.**, 2007, Extension of runs to the continuous valued sequences, *Statistics & Probability Letters*, Vol 77, 4, 383-388.

KAYNAKLAR DİZİNİ (devam)

- Eryılmaz, S. ve Demir, S.**, 2007, Success runs in a sequence of exchangeable binary trials. *Journal of Statistical Planning and Inference*, 137, 2954–2963.
- Eryılmaz, S. ve Fu, J.**, 2008, Runs in continuous valued sequences, *Statistics & Probability Letters*, vol. 78, 759-765.
- Fu, J. C.**, 1986, Reliability of consecutive- k -out-of- n :F systems with $(k-1)$ -step Markov Dependence, *IEEE Trans. Reliability*, vol. R-35, 602-606.
- Fu, J.C. and Lou, W.Y.W.**, 2003, Distribution theory of runs and patterns and its applications. A finite markov chain imbedding technique, World Scientific Pub., USA.
- Fu, J.C., Lou, W.Y.W. and Chen, S.C.**, 1999, On the probability of pattern matching in nonaligned DNA sequences: a finite Markov chain imbedding approach, *In Scan Statistics and Applications*, 287 – 302, Birkhäuser, Boston.
- Ghent, A. W., Zar, J. H.**, 1992, Runs of two kinds of elements on a circle a redevelopment, with corrections, from the perspective of biological research, *American Midland Naturalist*, 128:377-396.
- Gibbons, J. D. and Chakraborti, S.**, 2003, Nonparametric Statistical Inference, Marcel Dekker Inc., Newyork, 645p.
- Godbole, A.P.**, 1990a, Specific formulae for some success run distributions, *Statistics & Probability Letters*, 10, 119 – 124.
- Godbole, A. P.**, 1992b, The exact and asymptotic distribution of overlapping success runs, *Commun. Statist.–Theory Meth.*, 21(4), 953-967.
- Godbole, A.P., Papastavridis, S.G. and Weishaar, R.S.**, 1997, Formulae and recursions for the joint distribution of success runs of several lengths, *Ann. Inst. Statist. Math.*, 49, 141-153.
- Goldstein, L. and Waterman, M.S.**, 1992, Poisson, compound Poisson and process approximations for testing statistical significance in sequence comparisons, *Bulletin of Mathematical Biology*, 54, 785 – 812.
- Hirano, K.**, 1986, Some properties of the distributions of order k , *In Fibonacci Numbers and Their Applications*, 43–53, D.Reidel Publishing Company, Dordrecht, The Neterhlands.

KAYNAKLAR DİZİNİ (devam)

- Hirano, K., Aki, S., Kashiwagi, N. and Kuboki, H.,** 1991, On Ling's binomial and negative binomial distributions of order k , *Statistics & Probability Letters*, 11, 503-509.
- Kibritçioğlu, A.,** 1998, "Firma ve ürün kalitesi nedir ve neden önemlidir?", <http://129.3.20.41/eps/it/papers/0509/0509011.pdf> (Erişim tarihi: 10 Nisan 2012)
- Kitagawa, T. and Seguchi, T.,** 1956, The combined use of runs in statistical quality controls, *Bulletin of Mathematical Statistics*, 7, 25– 45.
- Kitagawa, T. and Seguchi, T.,** 1957, The combined use of runs in statistical quality controls II, *Bulletin of Mathematical Statistics*, 7, 53 – 72.
- Koutras, M. V. and Alexandrou, V. A.,** 1997, Non-parametric randomness tests based on success runs of fixed length, *Statistics & Probability Letters*, 32, 393-404.
- Koutras, M. V., Papastavridis, S.G. and Petakos, K.I.,** 1995, Bounds and limit theorems for coherent reliability structures, *In Recent Advances in Life – Testing and Reliability*, 267 – 292, CRC Press, Boca Raton, Florida.
- Koutras, M. V., Papastavridis, S.G. and Petakos, K.I.,** 1996, Bounds for coherent reliability structures, *Statistics & Probability Letters*, 26, 285 – 292.
- Lam, Y. and Ng, H.K.T.,** 2001, A general model for consecutive- k -out-of- n : F system with exponential distribution and $(k - 1) -$ step Markov dependence, *European Journal of Operational Research*, 129, 663 – 682.
- Lam, Y. and Zhang, Y.L.,** 1999, Analysis of repairable consecutive-2-out-of- n : F systems with Markov dependence, *International Journal of Systems Science*, 30, 799 – 809.
- Lam, Y. and Zhang, Y.L.,** 2000, Repairable consecutive- k -out-of- n : F system with Markov dependence, *Naval Research Logistics*, 47, 18 – 39.
- Ling, K.D.,** 1988, On Binomial distributions of order k , *Statistics & Probability Letters*, 6, 247 – 250.
- Ling, K. D.,** 1989, A new class of negative binomial distributions of order k , *Statistics & Probability Letters*, 7, 371-376.

KAYNAKLAR DİZİNİ (devam)

- Lou, W. Y. W.**, 1996, On runs and longest run tests: a method of finite Markov chain imbedding, *J. Amer. Statist. Assoc.*, 91, 1595-1601.
- Malinowski, J. and Preuss, W.**, 1996, A recursive algorithm evaluating the exact reliability of a circular consecutive $k - \text{within} - m - \text{out} - \text{of} - n : F$ system, *Microelectronics and Reliability*, 36, 1389 – 1394.
- Mood, A. M.**, 1940, The distribution theory of runs, *Annals of the Institute of Statistical Mathematics*, 11, 367 – 392.
- Moore, P:T.**, 1958, Some properties of runs in quality control procedures, *Biometrika*, 45, 89 – 95.
- Mosteller, F.**, 1941, Note on an application of runs to quality control charts, *Annals of Mathematical Statistics*, 12, 228 – 232.
- Mott, R.F., Kirkwood, T.B.L. and Curnow, R.N.**, 1990, An accurate approximation to the distribution of the length of the longest matching word between two random DNA sequences, *Bulletin of Mathematical Biology*, 52, 773 – 784.
- Muselli, M.**, 1996, Simple expressions for success run distributions in Bernoulli trials, *Statistics & Probability Letters*, 31, 121 – 128.
- Muselli, M.**, 2000a, New improved bounds for reliability of consecutive- k -out-of- n : F systems, *Journal of Applied Probability*, 37, 1164 – 1170.
- O'Brien, P. C. and Dyck, P. J.**, 1985, A runs test based on run lengths, *Biometrics*, 41:237-244.
- Onüçyıldız, M.**, 1999, Türkiye Akarsularında Taşkın Debilerinin Belirlenmesi, Doktora Tezi, Selçuk Üniversitesi Fen Bilimleri Enstitüsü, 140s (yayımlanmamış).
- Özkan, M.**, 2003, Bulanık Hedek Programlama, Ekin Kitapevi, Bursa, 288s.
- Papastavridis, S.G. and Koutras, M.V.**, 1994, Consecutive – k – out-of – n systems, *In New Trends in System Reliability Evaluation*, 228–248, Elsevier Science Publishers, Amsterdam, The Netherlands.
- Philippou, A.N. and Makri, F.S.**, 1986, Success runs and longest runs, *Statistics & Probability Letters*, 4, 211 – 215.

KAYNAKLAR DİZİNİ (devam)

- Philippou, A.N. and Muwafi, A.A.**, 1982, Waiting for the kth consecutive success and the Fibonacci sequence of order k, *The Fibonacci Quarterly*, 20, 28 – 32.
- Preuss, W.**, 1997, On the reliability of generalized consecutive systems, *Nonlinear Analysis – Theory, Methods & Applications*, 30, 5425 – 5429.
- Salazar, P.G. and Yevjevich, V.**, 1975, Analysis of drought characteristics by the theory of runs, *Hydrol. Pap.*,80, Colorado State University, Fort Collins.
- Saraçbaşı, T. ve Kutsal, A.**, 1987, Betimsel İstatistik, Hacettepe Üniversitesi Fen Fakültesi Yayınları, Ankara, 129s.
- Schmuelli, G. and Cohen, A.**, 2000, Run – related probability functions applied to the sampling inspection, *Technometrics*, 42, 188 – 202.
- Sokal, R. R. and Rohlf, F. J.**, 1969, *Biometry*, W. H. Freeman&Co., San Francisco, 776p.
- Swed, F.S. and Eisenhart, C.**, 1943, Tables for testing randomness of grouping in a sequence of alternatives, *Annals of Mathematical Statistics*, 14, 66 – 87.
- Şen, Z.**, 1976, Wet and dry periods of annual flow series, *J. Hydraul. Eng., ASCE, Prof. Pap. 12457, 102(HY10)*, 1503 – 1514.
- Şen, Z.**, 1990, Critical drought analysis by second order Markov chain, *Journal of Hydrology*, in press.
- Toros, H.**, 1993, Klimatolojik Serilerden Türkiye İkliminde Trend Analizi, Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, 72s (yayımlanmamış).
- Uşıkay, S. ve Aksu, S.**, 2002, Ülkemizde taşkınlar, nedenleri, zararları ve alınması gereken önlemler, *Türkiye Mühendislik Haberleri*, 47(420-421-422):133-136.
- Vaggelatou, E.** (2003) On the length of the longest run in a multi state markov chain, *Statistics and Probability Letters*, 62, 211 – 221.
- Wald, A. and Wolfowitz, J.**, 1940, On a test whether two samples are from the same population, *Annals of Mathematical Statistics*, 11, 147 – 162.

KAYNAKLAR DİZİNİ (devam)

- Walker, E., Philpot, J.W. and Clement, J.,** 1991, False signal rates for the Shewhart control chart with supplementary runs tests, *Journal of Quality Technology*, 23, 247 – 252.
- Waterman, M.S.,** 1982, Frequencies of restriction sites, *Nucleic Acids Research*, 11, 8951 – 8956.
- Waterman, M.S.,** 2000, Introduction to Computational Biology, CRC Press, Boca Raton, Florida.
- Weiler, H.,** 1953, The use of runs to control the mean in quality control, *Journal of the American Statistical Association*, 48, 816 – 825.
- Wendy Lou, W. Y.,** 1996, On runs and longest run tests: A method of finite Markov chain imbedding, *American Statistical Association*, 91:1595-1601.
- Wolfowitz, J.,** 1943, On the theory of runs with some applications to quality control, *Annals of Mathematical Statistics*, 14, 280 – 288.
- Woodall, W.H., 1983,** The distribution of the run length of one –sided CUSUM procedures for continuous random variables, *Technometrics*, 26, 295 – 301.
- Woodall, W.H., 1984,** On the Markov chain approach to the two–sided distribution of CUSUM procedure, *Technometrics*, 26, 41– 46.

ÖZGEÇMİŞ

1975 yılında Ankara’da doğan Aslı Ayşe BİLİR, ilk, orta ve lise öğrenimini Ankara’da tamamlamıştır. 1997 yılında Hacettepe Üniversitesi Fen Fakültesi İstatistik Bölümü’nden mezun olmuştur. Mezuniyet sonrası, kısa bir süre Megar Araştırma Danışmanlık şirketinde İstatistikçi olarak görev yapmıştır. 1998 yılında Devlet İstatistik Enstitüsü (DİE)’nin açtığı “Uzman Yardımcılığı” sınavını kazanarak memuriyet hayatına başlamıştır. 2004 yılında “Devlet İstatistik Enstitüsünde Elektronik Veri Güvenliği” konulu tezini başarı ile tamamlayarak Uzmanlık kadrosunu almıştır.

Türkiye İstatistik Kurumu (TÜİK) Başkanlığı Merkez Teşkilatında 7 yıl, TÜİK Samsun Bölge Müdürlüğü’nde 6 yıl görev yaptıktan sonra 2010 yılında TÜİK İzmir Bölge Müdürlüğü’ne atanan Aslı Ayşe BİLİR, halen daha TÜİK İzmir Bölge Müdürlüğü’nde “TÜİK Uzmanı” kadrosunda görev yapmaktadır.