

## Seyahat Süresi Değişkenliğinin ve Güvenilirliğinin İstatistiksel Analizi

Sibel SERTOK<sup>1</sup>, Berk ANBAROĞLU<sup>2</sup>, Özgür YENİAY<sup>\*3</sup>

<sup>1,3</sup>Hacettepe Üniversitesi, Fen Fakültesi, İstatistik Bölümü, 06800, Ankara, Türkiye

<sup>2</sup>Hacettepe Üniversitesi, Mühendislik Fakültesi, Geomatik Mühendisliği Bölümü, 06800, Ankara, Türkiye

(Alınış / Received: 13.03.2020, Kabul / Accepted: 22.05.2020, Online Yayınlanma / Published Online: 20.08.2020)

### Anahtar Kelimeler

Büyük veri,  
Seyahat süresi dağılımı,  
Seyahat süresi güvenilirliği,  
Veri tabanı yönetim  
sistemleri

**Özet:** Şehir içi ulaşım ağlarında bir noktadan başka bir noktaya yapılan yolculukların seyahat sürelerinin istatistiksel olarak modellenmesi ve bunların güvenilirliğinin analiz edilmesi büyük önem arz etmektedir. Gerçekleştirilen kimi çalışmalarda seyahat süresini modellemek için teorik olarak uygun olmayan dağılımlar tercih edilmiştir. Bunun yanında birçok çalışma haftanın günlerini hafta içi ve hafta sonu olmak üzere iki farklı sınıfta değerlendirmektedir. Gelişen sensör teknolojileri ile elde edilen büyük ulaşım verisi üzerinden seyahat sürelerinin istatistiksel dağılımı ve güvenilirliği gibi şehir hayatına etki eden faktörler analiz edilebilmektedir. Bu çalışmada New York şehrine ait açık ulaşım verisi analiz edilerek, birçok çalışmanın varsaydığı hafta içi-hafta sonu ayrımının gerçekçi olmayabileceği ve bir gün içindeki farklı zaman dilimlerinde istatistiksel sonuçların farklı çıkabileceği tespit edilmiştir. John F. Kennedy ve LaGuardia havalimanları arasında 07:00-19:00 arasında gerçekleşen taksi yolculuklarının seyahat süresi değişkenliği ve dağılımları incelendiğinde, seyahat sürelerinin Log-Lojistik (3P) dağılımına uyduğu tespit edilmiştir. Bu dağılım Salı ve Çarşamba günleri baskın dağılım değildir ve bu günlerde seyahat süresi güvenilirliğinin düşük olduğu zaman dilimlerinin bulunduğu tespit edilmiştir.

## Analysis of Travel Time Variability and Reliability in Big Taxi Data

### Keywords

Big data,  
Travel time distribution,  
Travel time reliability,  
Database management  
systems

**Abstract:** Many cities are dealing with traffic problems. One of emerging of these problems is that travel times are longer than estimated values. Variability in travel times leads to disruptions in supply chain and delayed passengers, amongst others. Variability of travel time is a phenomenon dependent on time and space. Determining this variability is important for transportation planning in traffic-intensive times and regions. Such a quest requires big geospatial data. The taxi dataset of New York City has been openly distributed since 2009, which describes each taxi trip with attributes such as the pickup and dropoff location and time, total cost of the trip and number of passengers. The aim of this article is to determine the travel time variability and distributions between John F. Kennedy and LaGuardia airports using the openly available taxi dataset involving approximately 140 million trips that occurred in 2015. The journeys, which take place every day of the week and between 07:00 and 19:00, are analyzed at 15 minute intervals. Log-Normal, Log-Normal (3P), Log-Logistics, Log-Logistics (3P), Weibull, Gamma and Burr distributions were examined. 95% of travel time, buffer time index and planned time index were used as the travel time reliability measurements. According to the results obtained, travel times are consistent with Log-Logistics (3P) distribution. This distribution is not the dominant distribution on Tuesdays and Wednesdays and it has been found that there are times when travel time reliability is low.

### 1. Giriş

Dünya nüfusunun artmasıyla trafik problemleri de çoğalmıştır. Hem gelişmiş hem de gelişmekte olan

ülkelerin başlıca sorunlarından biri de trafiğin etkin bir şekilde yönetilememesidir. Bir şehirde yaşayan kişi sayısının ve dolayısıyla ekonomik faaliyetlerin artmasıyla, trafik ağlarındaki yoğunluk da

artmaktadır. Böylece, bir noktadan başka bir noktaya yapılan yolculukların süresi de etkilenmektedir. Bir başka deyişle, seyahat süresi mekâna ve zamana bağlı bir olgu olarak değişkenlik göstermektedir. Bu değişkenliğin belirlenmesi ve analiz edilmesi de araştırmacıların ilgisini çekmektedir. Seyahat süresi dağılımının belirlenmesiyle, trafiğin etkin bir şekilde yönetimi kapsamında önemli bir adım atılmış olacaktır.

Bazı araştırmacılar seyahat sürelerini ifade etmek için Normal dağılım gibi simetrik dağılımların daha uygun olacağını belirtmektedirler. Örneğin; Taylor Fransa'daki metro ve otobüslerin seyahat süresi dağılımını normal dağılım olarak tespit etmiştir [1]. Bir diğer örnekte ise Avustralya'nın Melbourne şehrindeki otobüs seyahat süreleri analiz edilmiştir ve normal dağılımın seyahat sürelerini modellemek için değerlendirilebileceği tespit edilmiştir [2]. Ancak, Normal dağılımın simetrik yapısı dolayısıyla eksi değerler de içerebileceği ve bunun gerçekte bağdaşmayacağı görülmektedir

Kimi araştırmacılar da seyahat sürelerinin pozitif değerli olması gerektiği yönüne dikkat çekerek, Log-Normal dağılım gibi çarpık dağılımları incelemişlerdir. Bu araştırmacıardan Kieu ve arkadaşları seyahat süresi dağılımına en uygun dağılımın Log-Normal dağılım olduğunu iddia etmişlerdir [3]. Bu iddialarını da merkezi limit teoremine dayandırmışlardır. Başka bir makalede Beijing'de dört farklı yol tipiyle yaptıkları çalışmada da Log-Normal dağılımın en iyi dağılım olduğu bulunmuştur [4].

Diğer bir çarpık dağılım olan Log-Lojistik dağılımı da seyahat süresi dağılımı için uygun bir dağılımdır. Log-Normal dağılımı karşılaştırıldığında aslında her iki dağılımın da şekli birbirine benzer ancak Log-Lojistik dağılımın analiz edilmesinin daha kolay olduğu belirtilmektedir [5]. Şili Santiago'daki metro otobüs ve arabaların seyahat süresi değişkenliğini tahmin etmek için Log-Lojistik dağılımından yararlanılmıştır [6].

Diğer dağılımlar kadar çok literatürde karşılığı olmasa da yeni bir istatistiksel dağılım olan Burr dağılımı da seyahat süresini modellemek için önerilmektedir. Taylor ve arkadaşı bu dağılımın esnek bir şekle ve uzun üst kuyrukları tanımlama yeteneğine sahip olduğunu belirtmişlerdir [7]. Bu da seyahat süresi güvenilirliğini çok daha detaylı ölçmeyi sağlayabilmektedir.

Seyahat süresi güvenilirliği ile seyahat süresi dağılımı birbiriyle ilişkili iki konudur. Seyahat süresi güvenilirliğinin araştırılması için, öncelikle seyahat süresi dağılımının belirlenmesi faydalıdır [4]. Seyahat süresi güvenilirliğinin birden fazla tanımı vardır. Örneğin toplu ulaşım araçlarının planlanan zamanda hedeflerine ulaşması durumu olarak ya da bir yoldaki

trafik sıkışıklığının belirsizlik düzeyi olarak tanımlanabilir. Başka bir deyişle seyahat süresi güvenilirliğini, seyahat süresinin zaman içinde ne kadar değiştiğidir. Bu değişimin birçok nedeni olabilir. Örneğin toplu taşımadaki otobüs hatlarının güvenilirliği ile ilgilenilecek olunsaydı, güvenilirliği etkileyen temel faktörler arasında hat güzergâhlarına ait trafik koşulları, hattın uzunluğu ve yolcu nitelikleri (engelli veya bebek arabalı yolcular) gibi değişkenler bulunacak idi [8]. Bunların dışında, hava durumu, trafik kazaları, yol çalışmaları, konserler ve spor maçları gibi olaylar da seyahat süresi güvenilirliğini etkileyebilmektedir. Bu olaylar her gün olmadıkları için güvenilir olmayan seyahat süresi üretmeye de elverişlidirler. Seyahat eden insanlar da bu tür belirsizliklerin seyahat sürelerine etkilerini bildikleri zaman, seyahat seçimlerini nerede, ne zaman ve nasıl yapacaklarını daha etkin bir şekilde belirleyebilirler.

New York Şehrindeki taksi veri setini kullanarak seyahat süresi güvenilirliği, çarpıklık, varyasyon katsayısı ve seyahat süresinin enliği yöntemlerini kullanarak belirlenmiştir [9]. Başka bir makalede ise Japonya'nın Hanshin Otoyolu verilerini kullanarak seyahat süresi güvenilirliği araştırılmıştır. Kullandıkları güvenilirlik ölçümleri ise seyahat süresinin 95'lik yüzdeleri, standart sapma, varyasyon katsayısı ve tampon zaman indeksidir [10].

Bir diğer çalışmada ise farklı seyahat süresi güvenilirlik ölçüm yöntemleri karşılaştırılmıştır. Bu yöntemler arasında %95 seyahat süreleri, standart sapma, varyasyon katsayısı, varyasyon yüzdesi, tampon indeksi, planlama zaman indeksi, seyahat süresi indeksi, çarpıklık katsayısı ve kötü yolculukların gecikme zamanlarını gösteren sefalet (misery) indeksidir. İncelemeler sonucu varyasyon katsayısının seyahat süresi güvenilirliğini standart sapmaya göre daha iyi temsil ettiği tespit edilmiştir [11].

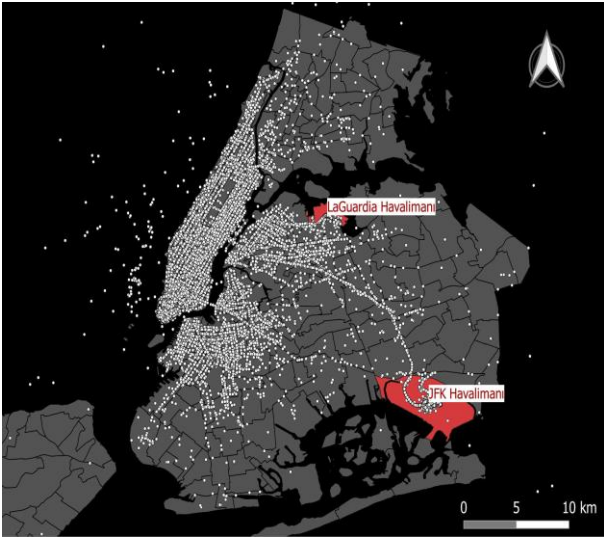
Seyahat süresi güvenilirliğini ölçen genel ölçümlerin dışında, seyahat sürelerinin çarpıklık ve genişliğini kullanarak farklı bir ölçümde önerilmiştir. Bu ölçüm güvenilir olmayan indeks (unreliability index) olarak adlandırılmıştır. Ancak güvenilir olmayan indeks kullanabilmek için rota uzunluğunun bilinmesi gerekmektedir [12].

Seyahat süresinin güvenilirliğinin analizi için büyük ulaşım verisine ihtiyaç duyulmaktadır. Gelişen GPS teknolojisinin kullanımı bu ihtiyacı karşılama yönünde öne çıkmaktadır; çünkü ucuzlayan GPS maliyetleri ile birçok aracın anlık konumu ve hızı gibi veriler kolaylıkla toplanabilmektedir. Her gün yüz binlerce taksi yolculuğunun yapıldığı bir şehirde, GPS alıcısı yerleştirilmiş her aracın (ör. taksi) sabit aralıklarla (ör. dakikada bir) izlenmesi çok maliyetli bir işlemdir. Dolayısıyla, araştırmacılar daha kısıtlı

veri setleri üzerinden de çalışabilecek yöntemler üzerine yoğunlaşmışlardır. Bu tür veri setlerinde yolculukların başlangıç ve bitiş noktalarına ek olarak yolcu sayısı, yolculuğun uzunluğu ve başlangıç/bitiş zamanı gibi öznelikler de bulunmaktadır. Açık bir şekilde sunulan New York Şehrine ait taksi veri seti de bu alanda bilinen en değerli veri setlerinden biridir [13].

## 2. Materyal ve Metot

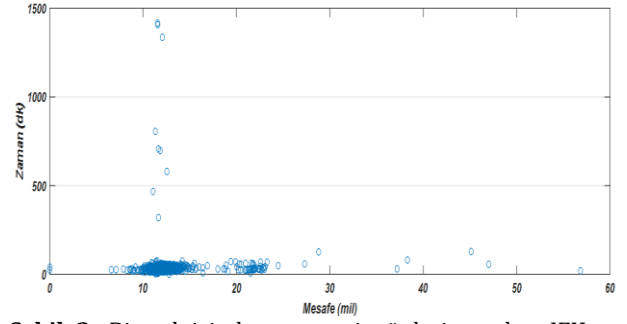
Bu çalışma da New York şehrindeki iki havalimanı olan John F. Kennedy (JFK) ve La Guardia arasındaki seyahat süresi güvenilirliği ve dağılımı araştırılmıştır. İki havalimanının harita üzerindeki görüntüsü Şekil 1'de gösterilmiştir.



Şekil 1. New York'taki iki havalimanı: JFK ve LaGuardia ve 23 Aralık 2015 günü yapılan yolculukların başlangıç noktaları

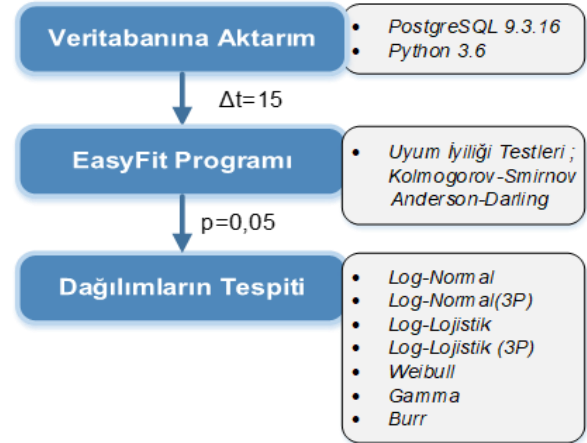
Bu iki havalimanının seçilmesinin nedeni aynı bölge (Queens) içinde bulunmaları ve aralarında tren ulaşımının olmamasıdır.

Güneydeki havalimanı olan JFK'den başlayan ve LaGuardia'da biten yolculukların süresi ve uzunluğu Şekil 2'deki gibi grafiğe döküldüğünde, yolculuk sürelerindeki değişkenlik daha iyi anlaşılmaktadır. Bu iki havalimanı arasında aktarmalı uçuşu olan yolcuların trafik düzensizliğinden sıkıntı çekmemeleri için farklı gün ve saatlerdeki seyahat süresi güvenilirliği üç farklı yöntem kullanılarak tespit edilmiştir. Amerika Birleşik Devletleri Ulaştırma Bakanlığının önerdiği bu yöntemler, i) seyahat sürelerinin 95'lik yüzdeleri, ii) tampon zaman indeksi (buffer time index) ve iii) planlanan zaman indeksidir (planning time index). Haftanın yedi günü 15 dakikalık zaman dilimlerine bölünmüş olup, bu süreçlerde gerçekleştirilen yolculukların belirtilen üç yöntem ile seyahat süresi güvenilirliği analiz edilmiştir. Sonrasında ise, seyahat süresi güvenilirliği ile doğrudan ilişkili bir kavram olan seyahat süresi dağılımı tespit edilmiştir.



Şekil 2. Bir yıl içinde pazartesi günleri yapılan JFK ve LaGuardia arasında yapılan yolculukların zaman ve mesafe ilişkisinin gösterimi

Makale kapsamında önerilen metodoloji Şekil 3'te gösterilmiştir. İlk olarak seyahat süresi dağılımı ve güvenilirlik ölçümü için gerekli verinin hazırlanması için gerekli olan adımlar tanıtılacaktır. Böylece, haftanın her günü için belirli bir bölgeden, başka bir bölgeye ve belirli bir zaman aralığında yapılan yolculukların süresi elde edilebilecektir. İkinci olarak da elde edilen yolculuk sürelerinin hangi dağılıma uyduğunun tespit edilmesi gerekmektedir. Bu kapsamda incelenen dağılımlar Log-Normal, Log Normal (3P), Log-Lojistik, Log Lojistik (3P), Weibull, Gamma ve Burr gibi sadece pozitif değer alabilen dağılımlardır. Seyahat sürelerinin hangi dağılıma uyduğu EastFit programı ile belirlenmiştir. Son olarak da seyahat süresi güvenilirliği yöntemlerinden, Amerika Birleşik Devletleri Ulaştırma Bakanlığının önerdiği üç ölçüm kullanılmıştır. Bunlar seyahat sürelerinin 95'lik yüzdeleri, tampon zaman indeksi ve planlanan zaman indeksidir.



Şekil 3. Seyahat sürelerinin dağılımının tespiti

### 2.1. Veri hazırlığı

İlk aşamada büyük taksi veri setinin bir veri tabanı yönetim sistemine aktarılması gerekmektedir. Bu kapsamda tercih edilen yazılım PostgreSQL olmuştur.

Bunun temel nedeni de üzerine eklenti olarak kurulan ve konumsal sorgulamaya izin veren PostGIS kütüphanesidir. Böylece, enlem ve boylamı verilen yolculuk başlangıç ve bitiş noktalarının hangi

bölgelerden başladığı tespit edilebilir. Nitekim, bu makalenin temel aldığı veri setinde (New York şehrinde 2015 yılına ait sarı taksi veri seti) kaydedilen her yolculuğunun başladığı ve bittiği enlem-boylam verisi bulunmaktadır. Bunun yanında taksi bölgelerini içeren geometrik veri katmanı da bulunmaktadır. Böylece, PostGIS ile birlikte Kod 1’de belirtilen sorgu gerçekleştirilerek, haftanın herhangi bir günü için (örneğin salı), belirli bir zaman aralığında (örneğin 09:15 - 09:30 arasında), JFK havalimanından başlayıp, LaGuardia havalimanına gitmiş tüm taksi yolculuklarının yolculuk süreleri elde edilmektedir.

Şekil 4’de belirtilen sorguda taksi yolculuklarının verisi ‘yolculuklar’ adlı tabloda tutulmaktadır ve taksi bölgelerini içeren tablo da ‘bolge’ adıyla saklanmaktadır. Her bir taksi yolculuğunun başlangıç ve bitiş noktası ‘k\_baslangic’ ve ‘k\_bitis’ adlı özniteliklerde saklanmaktadır. Öncelikle bir bölge içinde kalan nokta sorgusu ile (st\_contains), ‘JFK’ adlı taksi bölgesinden başlayan ve ‘LaGuardia’ adlı taksi bölgesinden tamamlanan yolculuklar tespit edilmektedir. Bunlar içinde de salı günü (son satırdaki 2 ifadesinden anlaşıldığı üzere) 09:15 - 09:30 arasında başlayan tüm yolculuklar belirlenmektedir. Bu yolculukların da yolculuk süresi ilk satırda belirtildiği gibi, yolculuğun bitiş zamanından (t\_bitis) yolculuğun başlangıç zamanı (t\_baslangic) zamanı çıkarılarak bulunabilmektedir.

```
select t_bitis - t_baslangic
from yolculuklar
full join bolge z1 on st_contains(b1.geom, k_baslangic)
full join bolge z2 on st_contains(b2.geom, k_bitis)
where t_baslangic >= '2015-01-01' and t_baslangic < '2016-01-01'
and (extract(hour from t_baslangic) between 9 and 9)
and (extract(minute from t_baslangic) between 15 and 30)
and b1.ad = 'JFK' and b2.ad = 'LaGuardia'
and extract(dow from t_baslangic) in 2
```

**Şekil 4.** Salı günleri 09:15-09:30 arası yapılan yolculuk sürelerinin bulunması

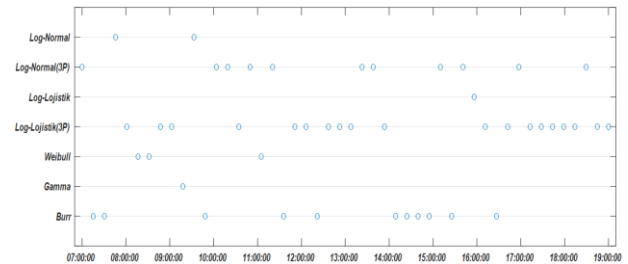
Bu sorgunun birçok zaman aralığında ve farklı günlerde çalıştırılması için ilgili değişkenleri manuel olarak değiştirilmesi maliyetli bir işlem olacaktır. Dolayısıyla, Python programlama dilinde bir betik yazılarak, bu değerlerin geliştirilen yazılım aracılığıyla otomatik olarak elde edilmesi sağlanmıştır. Böylece, haftanın her günü için ayrı ayrı olarak, 07:00 ile 19:00 arasında JFK havalimanından LaGuardia havalimanına yapılan yolculuk süreleri 15 dakikalık zaman aralıklarında incelenmiştir.

## 2.2. Seyahat süresi dağılımları

Bu bölüm de iki bölge arasındaki on beş dakikalık zaman dilimlerinin her birinin dağılımı bulunmaya çalışılmıştır. Seyahat süresi dağılımını bulmak değişkenliğin istatistiksel olarak yapısını ve düzenini anlamamızı sağlar.

Easy Fit programı yardımıyla önceki kısımda elde edilen yolculuk sürelerinin en çok hangi dağılıma uyum sağladığı tespit edilmiştir. Bu programın kullanılmasının sebebi diğer programlara göre daha geniş dağılım seçeneklerine sahip olması ve bu dağılımların içinde seyahat süreleri için ihtiyaç duyulan dağılımları barındırmasıdır.

Verinin hangi dağılıma uyduğunu görmek için örneklem büyüklüğüne göre uyum iyiliği testlerinden Kolmogorov-Smirnov ve Anderson-Darling testlerinin %5 önemlilik derecesi baz alınmıştır. Kolmogorov-Smirnov testi örneklem büyüklüğünün 35’ten büyük olduğu durumlarda kullanılması daha uygun olduğu için kullanılmıştır. Anderson-Darling testi de yirmi ile elli arası örneklem büyüklüğüne uygun olduğu için kullanılmıştır [14]. Haftanın her günü için ve belirlenen zaman aralıklarında elde edilen yolculuk sayıları (örneklem büyüklüğü) 22 ile 138 arasında değişmektedir ve toplamda 336 aralıkta (48 zaman dilimi \* 7 gün), sadece 5 aralıkta 30’un altındadır. Her bir veri seti için her dağılımın olasılık değerine bakılmıştır. Eğer olasılık değeri 0.05’ten küçükse veri o dağılıma uymuyor, 0.05’ten büyükse de veri yukarıdaki dağılımlardan birine uyuyor olarak kabul edilmiştir. Seyahat sürelerinin dağılımının nasıl tespit edileceğini belirten çizim Şekil 3’te gösterilmiştir. Bütün günlerin her zaman aralığı için dağılımı bulunduktan sonra yüzdelenmiştir. Örneğin pazartesi günü on dokuz tane Log-Lojistik (3P) dağılımına uyan zaman dilimi varsa,  $\frac{19}{48} * 100 = \%39.58$  ile pazartesi gününün Log-Lojistik ile modellenebileceği kabul edilmiştir. Pazartesi günü için incelenen tüm zaman aralıklarında hangi dağılımın baskın dağılım olduğu Şekil 5’da gösterilmiştir.



**Şekil 1.** Pazartesi günü için bütün dağılımların zamansal dağılımı

### 2.2.1. Log-Normal dağılım

Normal dağılımın dönüştürülmesinden elde edilen Log-Normal dağılım sağa çarpık bir dağılımdır. Yani yalnızca pozitif değerler alır. İki parametrelili olan bu dağılım ekonomiden biyolojiye değişik alanlarda kullanılabilir [15]. Log-Normal dağılımda x değişken,  $\mu$  ölçüm parametresi ve  $\sigma$  şekil parametresidir ve olasılık yoğunluk fonksiyonu ile kümülatif yoğunluk fonksiyonu sırasıyla denklem 1 ve 2’de belirtilmiştir.

Olasılık Yoğunluk Fonksiyonu:

$$f(x; \mu, \sigma) = \frac{e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}}}{x\sigma\sqrt{2\pi}} \quad (1)$$

Kümülatif Yoğunluk Fonksiyonu:

$$F(x; \mu, \sigma) = \Phi\left(\frac{\ln x - \mu}{\sigma}\right) \quad (2)$$

$\Phi$ : Standart normal bir değişkenin kümülatif dağılım fonksiyonudur.

### 2.2.2. Log-Normal (3P) dağılım

Log-Normal dağılımın üç parametrelidir. Dağılımda  $x$  değişken,  $\mu$  (ölçüm parametresi) ve  $\sigma$  (şekil parametresi) ve  $\gamma$  (konum parametresi) bulunmaktadır. Log-Normal (3P)'nin olasılık yoğunluk fonksiyonu ile kümülatif yoğunluk fonksiyonu sırasıyla denklem 3 ve 4'de belirtilmiştir

Olasılık Yoğunluk Fonksiyonu:

$$f(x; \mu, \sigma, \gamma) = \frac{e^{-\frac{(\ln(x-\gamma)-\mu)^2}{2\sigma^2}}}{(x-\gamma)\sigma\sqrt{2\pi}} \quad (3)$$

Kümülatif Yoğunluk Fonksiyonu

$$F(x; \mu, \sigma, \gamma) = \Phi\left(\frac{\ln(x-\gamma) - \mu}{\sigma}\right) \quad (4)$$

### 2.2.3. Log-Lojistik dağılım

Bu dağılım Weibull dağılımına alternatif bir dağılımdır. İki parametrelidir. Log-Lojistik dağılımı sürekli ve çarpık bir dağılımdır. Yukarıda da bahsedildiği gibi lojistik dağılımdan dönüştürülmüştür. Bu dağılım seyahat süresi dağılımı dışında hidroloji, ekonomi ve aktüerya alanlarında da kullanılmaktadır [16]. Dağılımda ki  $x$  değişken,  $\mu$  şekil parametresi,  $\sigma$  ölçüm parametresidir. Log-Lojistik dağılımının olasılık yoğunluk fonksiyonu ile kümülatif yoğunluk fonksiyonu sırasıyla denklem 5 ve 6'da belirtilmiştir.

Olasılık Yoğunluk Fonksiyonu

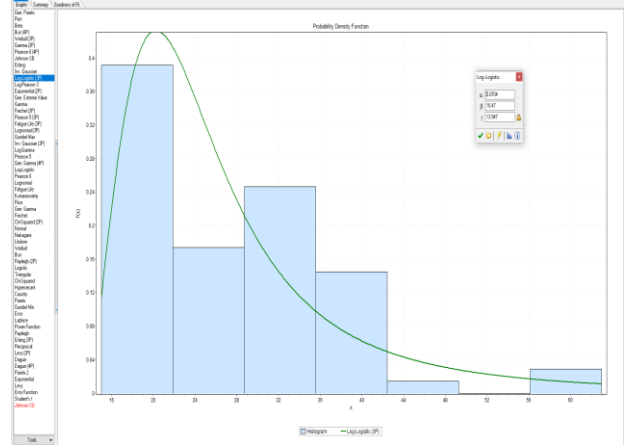
$$f(x; \mu, \sigma) = \frac{\left(\frac{\mu}{\sigma}\right) \left(\frac{x}{\sigma}\right)^{\mu-1}}{\left(1 + \left(\frac{x}{\sigma}\right)^\mu\right)^2} \quad (5)$$

Kümülatif Yoğunluk Fonksiyonu

$$F(x; \mu, \sigma) = \frac{1}{1 + \left(\frac{x}{\sigma}\right)^\mu} \quad (6)$$

### 2.2.4. Log-Lojistik (3P) dağılım

Log-Normal (3P)'de olduğu gibi bu dağılımda da konum parametresi ( $\gamma$ ) bulunmaktadır. Örnek olarak pazartesi 12:30-12:45 arası da bu dağılıma uymaktadır. Bu dağılımın Easy-fit gösterimi Şekil 6'de gösterilmiştir. Dağılımda  $\alpha$  ölçüm parametresi,  $\beta$  şekil parametresi,  $\gamma$  ise konum parametresidir. Log-Lojistik (3P)'nin olasılık yoğunluk fonksiyonu ile kümülatif yoğunluk fonksiyonu sırasıyla denklem 7 ve 8'de belirtilmiştir.



Şekil 2. Easy-fit programından bir görüntü (Pazartesi 12:30-12:45 zaman dilimine ait)

Olasılık Yoğunluk Fonksiyonu

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

Kümülatif Yoğunluk Fonksiyonu

$$F(x; \alpha, \beta, \gamma) = \left(1 + \frac{\beta}{x-\gamma}\right)^\alpha \quad (8)$$

### 2.2.5. Weibull dağılım

Sürekli bir dağılım olan Weibull dağılımı pozitif değerler almaktadır. Tıp ve jeofizik alanlarında kullanılmaktadır [17]. Bu dağılımda  $x$  değişken,  $k$  şekil parametresi,  $\lambda$  ölçek parametresidir. Weibull dağılımının olasılık yoğunluk fonksiyonu ile kümülatif yoğunluk fonksiyonu sırasıyla denklem 9 ve 10'da belirtilmiştir.

Olasılık Yoğunluk Fonksiyonu:

$$f(x; k, \lambda) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k} \quad (9)$$

Kümülatif Yoğunluk Fonksiyonu:

$$F(x; k, \lambda) = 1 - e^{-\left(\frac{x}{\lambda}\right)^k} \quad (10)$$

### 2.2.6. Gamma dağılımı

Gamma dağılımı pozitif değerlerin analizinde kullanılan sürekli bir dağılımdır. Sağlık, fen bilimleri gibi alanlarda kullanılmaktadır [18]. Bu dağılımda  $x$  değişken,  $k$  şekil parametresi,  $\theta$  ölçek parametresidir. Gamma dağılımının olasılık yoğunluk fonksiyonu ile kümülatif yoğunluk fonksiyonu sırasıyla denklem 11 ve 12'de belirtilmiştir.

Olasılık Yoğunluk Fonksiyonu:

$$f(x; k, \theta) = x^{k-1} \frac{e^{-x/\theta}}{\theta^k \Gamma(k)} \quad (11)$$

Kümülatif Yoğunluk Fonksiyonu:

$$F(x; k, \theta) = \frac{\Gamma(x(k))}{\Gamma(k)} = \frac{\int_0^x t^{k-1} e^{-t} dt}{\int_0^\infty t^{k-1} e^{-t} dt} \quad (12)$$

### 2.2.7. Burr dağılımı

Negatif olmayan değerler için sürekli bir dağılımdır. Rüzgâr enerji potansiyeli tahmin etmede [19] ve aktüerya alanında kullanılmaktadır [20]. Bu dağılımda  $x$  değişken,  $c$  ve  $k$  parametreleri de şekil parametresidir. Burr dağılımının olasılık yoğunluk fonksiyonu ile kümülatif yoğunluk fonksiyonu sırasıyla denklem 13 ve 14'de belirtilmiştir.

Olasılık Yoğunluk Fonksiyonu:

$$f(x; c, k) = ck \frac{x^{c-1}}{(1+x^c)^{k+1}} \quad (13)$$

Kümülatif Yoğunluk Fonksiyonu:

$$F(x; c, k) = 1 - (1+x^c - k) \quad (14)$$

### 2.3. Seyahat süresi güvenirligi

Farklı seyahat süresi güvenirligi ölçüm yöntemleri bulunmaktadır. Bu makale kapsamında seyahat sürelerinin 95'lik yüzdeleri (95th percentile), tampon zaman indeksi (buffer time index) ve planlanan zaman indeksi (planning time index) değerlendirilmiştir.

İlk güvenirlilik ölçüm yöntemi olarak 95'lik yüzdendir. Yapılan tüm yolculuk süreleri tek bir kümede toplansa, bu değer kümedeki bütün değerlerin %95'inden daha büyüktür. Bir başka deyişle, bu değer üstünde yolculuk süresinin gözlemlenmesinin olasılığı %5'dir [10].

Tampon zaman indeksi, seyahat edenlerin gitmek istedikleri yere zamanında varmaları için ortalamaya eklenmesi gereken ek zamanı verir. Bu ek zaman beklenmeyen gecikmeleri ifade etmektedir. Tapon

zaman indeksinin hesaplanması denklem 15'de belirtildiği gibidir.

$$\frac{\sum_{n=1}^N (SSw, x, 95th - \overline{SSw, x})}{SSw, x} \quad (15)$$

Denklemden belirtilen SS seyahat süresini,  $w$  haftanın günlerini,  $x$  ise zaman dilimlerini belirtmektedir.  $SSw, x$  ise ilgili gün ( $w$ ) ve zaman dilimindeki ( $x$ ) ortalama seyahat süresini, göstermektedir. Bu indeksin süreye çevrilmesi için ortalama değer ile çarpılması gerekmektedir. Örneğin, ortalama seyahat süresinin 20 dakika olduğu bir yolculukta tampon zaman indeksi 0.78 çıktığında, ortalama değeri eklenmesi gereken süre 15.6 olur.

Planlama zaman indeksi, tampon süresi dâhil edildiğinde planlanması gereken toplam seyahat süresini belirtir. Tampon zaman indeksi, gerekli ek zamanı gösterirken, planlanan zaman indeksi gereken toplam seyahat süresini içermektedir. Ayrıca planlanan zaman indeksi hem beklenmeyen gecikmeleri hem de tipik gecikmeleri içermektedir. Bu indeksin hesaplanabilmesi için serbest akış süresinin hesaplanması gerekmektedir. Planlama zaman indeksinin hesaplanması denklem 16'da gösterildiği gibi yapılmaktadır.

$$\frac{\sum_{n=1}^N SSw, x, 95th}{serbest akış seyahat süresi} \quad (16)$$

Bu indeksin de süreye çevrilebilmesi için serbest akış seyahat sürelerinin planlanan zaman indeksiyle çarpılması gerekmektedir. Örneğin serbest akış zamanı 16 dakika, planlanan zaman indeksi de 2.68 çıkarsa planlanması gereken süre 43 dakika olmaktadır.

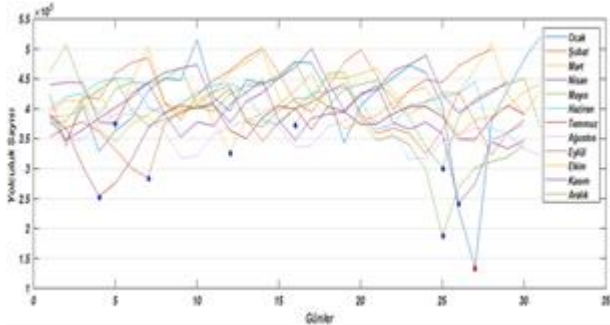
### 3. Bulgular

Bu çalışmada seyahat süresi dağılımının ve seyahat süresi güvenirliginin tespiti için New York şehrinin 2015 yılına ait sarı taksi verisi analiz edilmiştir. Bu veriye internet üzerinden açık bir şekilde erişmek mümkündür [21]. Her iki yolculuk için id numarası, yolculuğun başladığı/tamamlandığı konum/saat, toplam tutar, kaç kişinin yolculuk ettiği ve ne kadar yol gidildiği gibi özniteliklerin olduğu veride toplamda on dokuz öznitelik bulunmaktadır. Veri setinde 2015 yılına ait toplamda yaklaşık 144 milyon yolculuk analiz edilmiştir. Bu büyük veri Python aracılığıyla PostgreSQL veri tabanından haftanın her gününü 07:00-19:00 arasında 15 dakika zaman aralıklarında dosyalayacak şekilde çekilmiştir.

Taksi veri setinin indirilebileceği web sayfasından New York taksi bölgelerini içeren shapefile dosyası da değerlendirilmiştir. Bu dosyaya göre New York 263 bölgeye ayrılmıştır. En çok yolculukların yapıldığı bölge ise Manhattan bölgesidir. Analizde

kullanılan veriler New York'un en önemli hava limanlarından John F. Kennedy ve LaGuardia havalimanları arasındaki yolculuk sürelerine odaklanmıştır. Ancak, önerilen metodoloji geliştirilebilir ve herhangi iki bölge arasında da aynı analiz tekrar edilebilir. İncelene zaman aralığı 15 dakikalık zaman dilimlerine bölüldüğünden, bir günde 48 zaman dilimi ortaya çıkmaktadır. Bunu bir matris olarak düşünürsek 48x7 boyutlu bir matris oluşmaktadır. Bir başka deyişle, toplamda 336 test yapılmıştır.

Veriler detaylı incelendiğinde her ay belirli gün veya günlerde artış ya da azalışlar olmaktadır. Yolculuk sayılarındaki değişim Şekil 7'te gösterilmiştir. New York için özel olan günlerde örneğin Ocak ayının 19'u Martin Luther King'in doğum gününde, 17 Mart St.Patrick gününde, 5 Nisan Paskalya Bayramında, 4 Temmuz Bağımsızlık günü gibi özel günlerde taksi kullanımını düşürmüştür. 14 Şubat Sevgililer gününde ve 17 Ekim Sweetest gününde ise taksi kullanımını artırmıştır. Yıl içerisinde en fazla düşüş 27 Ocak günü, en fazla artışta 31 Ocak günü olmuştur. Ayrıca hafta içi hafta sonuna göre ve kış aylarında yaz aylarına göre daha fazla yolculuk yapıldığı görülmüştür. Yukarıda belirtildiği gibi, taksi kullanımında düşüş olan özel günler daha sonra serbest akış zamanını hesaplamak için kullanılacaktır.

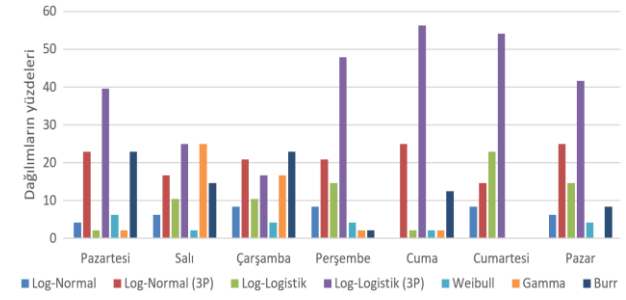


Şekil 3. Yıl içerisinde yapılan yolculuk sayısının aylara göre değişimi ve serbest akış süresi hesaplanması için kullanılan günler

### 3.1. Seyahat süresi dağılımı

Haftanın her günü için tüm zaman aralıkları için seyahat süresinin baskın dağılımı belirlenmiştir. Elde edilen sonuçlar Şekil 8'te sunulmuştur ve buna göre pazartesi, perşembe, cuma, cumartesi ve pazar günleri Log-Logistik (3P) dağılımına uyum sağlamaktadır. Salı günü yarı yarıya Log-Logistik (3P) ve Gamma dağılımına uymaktadır. Çarşamba günü ise en iyi Burr dağılımına uyum sağladığı söylenebilir. Bu duruma göre salı ve çarşamba günlerinin birbirinden farklı olduğu belirtilebilir. Ayrıca dağılımlar arasında en az Gamma dağılımının uyum sağladığı da gözlemlenmektedir. Buradan da görülmektedir ki haftanın her gününün birbirinden bağımsız düşünülmesi daha fayda olacaktır. Böylece hafta içi yolculukları ve hafta sonu gibi genelleştirmeler

yerine, her günün ayrı modellenmesinin daha doğru bir yaklaşım olacağı belirmektedir.



Şekil 4. Dağılımların yüzdelik gösterimi

### 3.2. Seyahat süresi dağılımı

Bu kısımda seyahat süresi güvenilirlik ölçümlerinden, seyahat süresi indeksi, tampon indeksi ve planlanan zaman indeksinin etkinliği araştırılmıştır.

#### 3.2.1. Serbest akış seyahat süresi hesaplanması

New York'ta bazı özel günlerde taksi kullanım sayısı azalmaktadır. Bu günler de serbest-akış zamanı ölçülebilir. Serbest akış seyahat zamanı (free-flow travel time), yolların daha boş olduğu zamanlardaki seyahat sürelerini temsil etmektedir. Bu ölçümün doğru olması için birden fazla gün kullanılmıştır. Bu günler:

- 19 Ocak: Martin Luther King'in doğum günü
- 17 Mart: St.Patrick günü
- 5 Nisan: Paskalya
- 25 Mayıs: Anma günü
- 4 Temmuz: Bağımsızlık günü
- 7 Eylül: İşçi Bayramı
- 12 Ekim: Kolomb Günü
- 26 Kasım: Şükran günü
- 25 Aralık: Noel

Bu günler Şekil 7'te mavi noktayla işaretlenmiştir. Mavi noktaların dışında ki kırmızı nokta ise 27 Ocak gününü göstermektedir. Nitekim o günde de düşüş yaşanmıştır. Ancak bu düşüş o günlerde gerçekleşen kar fırtınasından kaynaklanmaktadır. Bir başka deyişle diğer özel günler gibi her sene aynı gün düşüş yaşanmayabilir. Onun için 27 Ocak günü serbest akış zamanı hesaplamaya katılmamıştır.

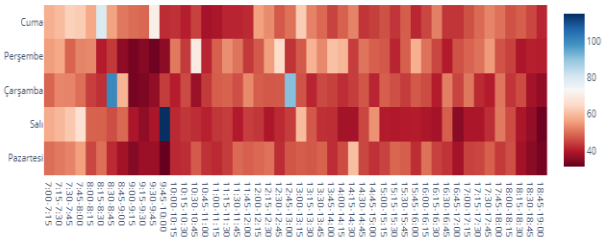
Serbest akış zamanını ölçmek için yukarıda bahsedilen günleri diğer günlerdeki gibi 15 dakikalık aralıklarla 7:00- 19:00 arasında olmak üzere toplanmıştır ve her bir zaman aralığı için ortalama seyahat süresi bulunmuştur. Bu sürelerle serbest akış seyahat süreleri denilebilir.

#### 3.2.2. Serbest akış seyahat süresi hesaplanması

Ölçümlerin sonuçları sıcaklık haritası (heatmap) olarak gösterilmiştir. Sıcaklık haritasına göre mavi ve koyu mavi renkli olanlar seyahat süresi

güvenirliğinin düşük olduğu zaman aralıklarını göstermektedir. Şekil 9'de gösterildiği gibi salı günü 9:45-10:00, çarşamba 8:30-8:45, 12:45-13:00 ve cuma 8:15-8:30 da yaşanabilecek en kötü seyahat süreleri fazla olan zaman dilimleridir. Bu zaman dilimleri daha detaylı incelendiğinde salı günü 9:45-10:00 arası 95'lik yüzdeyi 46 yolculuktan iki yolculuk aşmıştır ve bu iki yolculuğun aynı günlerde olmadığı görülmüştür.

Seyahatlerin 95'lik yüzdeleri



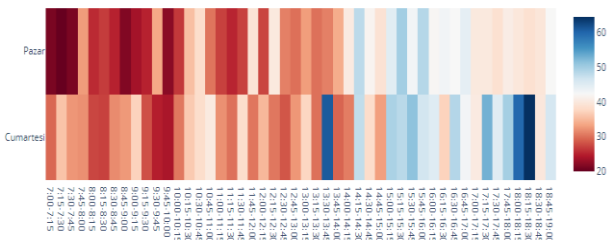
Şekil 5. Hafta içi için seyahatlerin 95'lik yüzdeleri

Çarşamba 8:30-8:45 zaman dilimi incelendiğinde ise 65 yolculuktan üç yolculuk 95'lik yüzdeyi aşmaktadır ve bu yolculukların ikisinin aynı ay içinde olduğu ama aynı gün içerisinde olmadığı görülmüştür.

Aynı ay içerisinde olan yolculuklardan biri 2 Aralık diğer ise 20 Aralık tarihindedir. Çarşamba 12:45-13:00 aralığında ise toplam 74 yolculuktan üçü 95'lik yüzdeyi aşmaktadır ve bu yolculuklardan ikisi aynı gün (28 Ekim) öğle saatlerinde çok yakın saatlerde gerçekleşmektedir. Hafta içi bir diğer incelemeye değer zaman aralığı ise cuma 8:15-8:30 arasındadır. Bu aralıkta 89 yolculuktan dördü 95'lik dilimi aşmaktadır ve bu dört yolculuktan ikisi 8 Mayıs sabah saatlerinde gerçekleşmiştir. Daha detaylı bir inceleme ile belirtilen zaman aralığında trafik kazası gibi beklenmedik bir trafik olayının gerçekleşip gerçekleşmediği incelenebilir

Hafta sonu (Şekil 10) ise cumartesi ve pazar günlerinin öğleden sonraki zamanlarında kötü bir seyahat süresiyle karşılaşılması olasıdır.

Seyahatlerin 95'lik yüzdeleri



Şekil 6. Hafta sonu için seyahatlerin 95'lik yüzdeleri

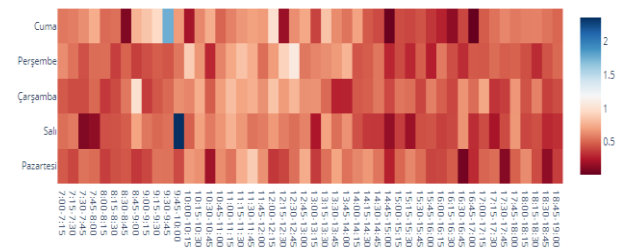
Cumartesi 18:15-18:30 arasında hafta sonu için en fazla olan süre bu zaman dilimindedir. Bu zaman dilimi de ayrıntılı incelendiğinde yirmi iki yolculuktan bir tanesi 95'lik dilimi aşmaktadır. Hafta

sonu ikinci en kötü seyahat süresi ise cumartesi günü 13:30-13:45 arasındadır. Bu aralıktaki 75 yolculuğun dördü 95'lik zaman dilimini aşmaktadır ve bu dört yolculuktan ikisi aynı güne denk gelmektedir. 16 Mayıs tarihine denk gelen iki yolculuk öğle saatlerinde aralarında yaklaşık on dakikalık farkla hareket etmektedirler. Aynı günlere denk gelen yolculukların hepsine bakıldığında denk geldiği günlerle ilgili herhangi özel bir şey bulunmamaktadır. Bu gecikmeler kazalardan, yol çalışmalarından ya da başka bir nedenden dolayı olabilir.

### 3.2.3. Tampon zaman indeksi

Tampon zaman indeksinin hafta içi değerleri Şekil 11'da gösterilmiştir.

Tampon Zaman İndeksi

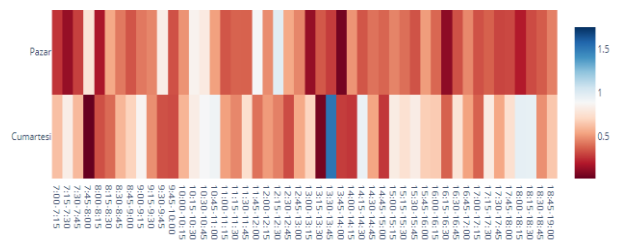


Şekil 11. Hafta içi için seyahatlerin tampon zaman indeksi

Bu değerlere göre hafta içi ortalamaya eklenmesi gereken en çok zaman salı günü 9:45-10:00 arasındadır. Orana göre gezi ortalama değerinden yaklaşık iki kat daha uzun sürebileceğini gösteriyor. Bu zaman diliminden sonra ikinci olarak cuma günü 9:30-10:00 arası gelmektedir. Çarşamba ve perşembe günleri öğle saatleri için de eklenmesi gereken zamanın fazla olduğu söylenebilir. Ortalamaya en az eklenmesi gereken zaman dilimi de cuma günü 16:45-17:00 saat dilimidir. Genel olarak da sabah saatleri akşam saatlerine göre daha az eklenmesi gereken zaman olduğu görülmüştür.

Şekil 12'da gösterildiği üzere cumartesi günü 13:30-13:45 arası ortalama en fazla eklenmesi gereken zaman oluşmaktadır. Bu değer hafta içi eklenecek değer neredeyse yarısı kadardır. En az eklenmesi gereken zaman ise yine cumartesi günü 7:45-8:00 arasındadır

Tampon Zaman İndeksi

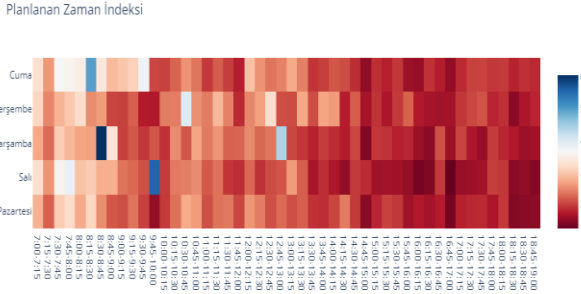


Şekil 12. Hafta sonu için seyahatlerin tampon zaman indeksi



### 3.2.4. Planlanan zaman indeksi

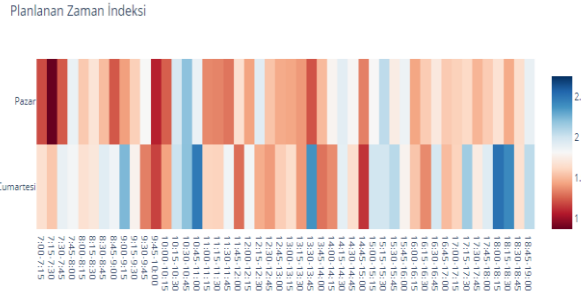
Hafta içi planlanan zaman indeksinin değerleri Şekil 13'de sunulmaktadır. Sonuçlara göre planlanan zaman indeksinin en fazla olduğu zaman dilimi çarşamba günü 8:30-8:45 arasındadır. Tampon zaman indeksinin en fazla olduğu gün ve zaman dilimi planlanan zamanda da yüksektir.



Şekil 13. Hafta içi için planlanan zaman indeksi

Bu da güvenilirlik ölçümlerinin birbirleriyle ilişkili olduklarını göstermektedir. Hafta içi en düşük değer ise salı günü 16:45-17:00 arasındadır.

Şekil 14'de hafta sonu değerleri için baktığımızda da cumartesi günü 18:00-18:15 arası en fazla, pazar da 7:15-7:30 arası en az değeri göstermektedir.



Şekil 14. Hafta sonu için planlanan zaman indeksi

## 4. Tartışma ve Sonuç

Gelişen sensör teknolojisi sayesinde, şehir içi ulaşım ağlarının daha etkin bir şekilde yönetilmesinin önü açılmaktadır. Elde edilen büyük ulaşım veri setleri sayesinde daha etkin bir şekilde seyahat süresi dağılımı ve güvenilirliği tespit edilebilmektedir. Böylece daha etkin bir ulaşım planlaması yapılabilir ve trafik sıkışıklığına bağlı gecikmelerin önüne geçilebilir. New York şehrine ait yaklaşık 144 milyon yolculuğun analiz edilmesi sonucunda seyahat sürelerinin Log-Lojistik(3P) dağılımı ile modellenebileceği görülmektedir. Bunun yanında haftanın her bir günün aynı dağılmadığı, çarşamba ve salı günlerinin farklı olduğu da görülmektedir. Seyahat süresi güvenilirliklerine bakıldığında da salı ve çarşamba günlerinde eklenmesi gereken sürenin diğer hafta içi günlerden daha fazla olduğu görülmektedir. Bu da birçok mevcut çalışmada olan hafta-içi günlerin aynı değerlendirilebileceği varsayımıyla çelişmektedir.

Tampon zaman indeksinde salı günü ortalamaya eklenmesi gereken zaman en fazladır. Bu planlanan zaman indeksi ölçümünde de belirgin olarak aynı olduğu görülmüştür. Sabah saatleri dışında hafta içi pazartesi ve salı günleri hemen hemen aynı planlanan zamanları takip etmektedirler. Hafta sonu cumartesi günü 13:30-13:45 arasında da ortalamaya eklenmesi gereken zaman en fazladır. Planlanan zaman indeksi de bunu desteklemektedir. Buna göre tampon zaman indeksi ölçümleri ile planlanan zaman indeksi ölçümlerinin birbirleriyle tutarlı oldukları da söylenebilir.

Genel olarak seyahat süreleri ölçümlerinin birbirleriyle ilişkili olduğu söylenebilir; ancak kimi durumlarda da ölçümler arasında tutarsızlıklar da olabilmektedir. Elde edilen sonuçlara göre hafta içi güvenilirliğin fazla olduğu günler perşembe ve pazartesi günleridir, hafta sonu ise pazar günü cumartesi gününe göre daha fazla güvenilirliğe sahiptir.

Gelecek çalışma olarak ise, sadece başlangıç ve bitiş noktası değil de tüm yolculuğun kaydedildiği veri setlerinin incelenmesi ile, yol bazında seyahat süresi güvenilirliği çalışmaları gerçekleştirilebilir.

## Teşekkür

Bu çalışma TÜBİTAK 118Y282 numaralı proje tarafından desteklenmiştir.

## Kaynakça

- [1] Taylor, M. A. P. 1982. Travel Time Variability-The Case of Two Public Modes. *Transportation Science*, 16, 507-521.
- [2] Mazloumi, E., Currie, G., Rose, G. 2010. Using GPS Data to Gain Insight into Public Transport Travel Time Variability. *Journal of Transportation Engineering*, 136, 623-631.
- [3] Kieu, L.-M., Bhaskar, A., Chung, E. 2015. Public Transport Travel-Time Variability Definitions and Monitoring. *Journal of Transportation Engineering*, 141, 04014068.
- [4] Chen, P., Tong, R., Lu, G., Wang, Y. 2018. Exploring Travel Time Distribution and Variability Patterns Using Probe Vehicle Data: Case Study in Beijing. *Journal of Advanced Transportation*. 1-13.
- [5] Singh, V. P., Guo, H., Yu, F. X. 1993. Parameter estimation for 3-parameter log-logistic distribution (LLD3) by Pome. *Stochastic Hydrology and Hydraulics*, 7, 163-177.
- [6] Durán-Hormazábal, E., Tirachini, A. 2016. Estimation of travel time variability for cars, buses, metro and door-to-door public transport trips in Santiago, Chile. *Research in Transportation Economics*, 59, 26-39.

- [7] Taylor, M. A. P., Susilawati. 2012. Modelling Travel Time Reliability with the Burr Distribution. *Procedia- Social and Behavioral Sciences*, 54, 75-83.
- [8] Özuysal, M., Çalışkanelli, S. P., Tanyel, S. 2016. A discussion on the dwell time values at bus stops. *Pamukkale University Journal of Engineering Sciences*, 22, 460-467.
- [9] Yazici, M. A., Kamga, C., Mouskos, K. C. 2012. Analysis of Travel Time Reliability in New York City Based on Day-of-Week and Time-of-Day Periods. *Transportation Research Record: Journal of the Transportation Research Board*, 2308, 83-95.
- [10] Higatani, A., Kitazawa, T., Tanabe, J., Suga, Y., Sekhar, R., Asakura, Y. 2009. Empirical Analysis of Travel Time Reliability Measures in Hanshin Expressway Network. *Journal of Intelligent Transportation Systems*, 13, 28-38.
- [11] Pu, W. 2011. Analytic Relationships between Travel Time Reliability Measures. *Transportation Research Record: Journal of the Transportation Research Board*, 2254, 122-130.
- [12] Van Lint, J. W. C., Van Zuylen, H. J., Tu, H. 2008. Travel time unreliability on freeways: Why measures based on variance tell only half the story. *Transportation Research Part A: Policy and Practice*, 42, 258-277.
- [13] Hochmair, H. H. 2016. Spatiotemporal Pattern Analysis of Taxi Trips in New York City. *Transportation Research Record: Journal of the Transportation Research Board*, 2542, 45-56.
- [14] Demir, E., Saatçioğlu, Ö., İmrol, F. 2016. Uluslararası Dergilerde Yayımlanan Eğitim Araştırmalarının Normallik Varsayımları Açısından İncelenmesi, 19s.
- [15] Fowlkes, E. B., Srivastava, V. K., Giles, D. E. A., Walker-Smith, M., Basford, K. E. Lognormal Distributions, 408s.
- [16] Lima, S. R., Cordeiro, G. M., Lima, S. R., Cordeiro, G. M. 2017. The Extended Log-Logistic Distribution: Properties and Application. *Anais da Academia Brasileira de Ciências*, 89, 3-17.
- [17] Lai, C.-D., Murthy, D. N., Xie, M. 2006. Weibull Distributions and Their Applications. In *Springer Handbook of Engineering Statistics*, ed. Pham, H., pp 63-78. Springer London, London. doi:10.1007/978-1-84628-288-1\_3.
- [18] Demirci Biçer, H., Biçer, C. 2017. Gamma ve Weibull Dağılımları Arasında Kullback-Leibler Uzaklığına Dayalı Ayrım. *Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, 17, 500-505.
- [19] Mert, İ., Karakuş, C. Burr Dağılımı Kullanılarak Rüzgâr Enerjisi Potansiyeli Tahmini, 8s.
- [20] Das, J., Nath, D. 2016. Burr Distribution as an Actuarial Risk Model and the Computation of Some of Its Actuarial Quantities Related to the Probability of Ruin. *Journal of Mathematical Finance*, 6, 213-231.
- [21] About TLC- TLC. Available from <https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page> (Erişim Tarihi: 30.05.2020).