ABMYO
Dergisi. 26, (2012) (29-40)

Uç değerler teorisi ve bir bankacılık uygulaması

Özge EREN[1]Orcan ALPAR[2]

Özet:

Bu çalışmada, az sıklıkta gerçekleşen ancak etkisi kuvvetli olan olayların ortaya çıkabilme ihtimalleri Uç Değerler Teorisi ile analiz edilmiştir. Bu teori, uç ve nadir olaylar ile ilişkilendirilen ihtimallerin hesaplanması için kullanılan bir araçtır. Çok çeşitli alanlarda kendine uygulama alanı bulan bu analiz, son yıllarda özellikle finansal alanlarda yoğunlukla kullanılmaktadır. Bu çalışmada ise, Uç Değerler Teorisi'nin genel mantığı hipotetik veriler üzerinden gösterilmiştir. Sadece teorik çerçeveyi göstermek yerine, bir banka için belirli bir zaman periyodunda gerçekleşen sahtekarlık olayları sonucunda ortaya çıkan kayıp değerleri ile ilgili örnek çalışma oluşturulmuştur.

Anahtar Kelimeler: Uç Değerler Teorisi; Operasyonel Risk; Hill tahmincisi

Extreme value theory and an application on a bank

Abstract:

Extreme value theory (EVT) is a tool used to consider probabilities associated with extreme and rare events. This paper is related with extreme events which usually appear with low frequently but high effects. It has been applied in various fields such as hydrology, as well as in studies of pollution, meteorology but in recent years, it is mainly used in the financial areas. General concepts of EVT is presented by a hypothetical case study. It is about fraud events in a Bank during specific time. This study aims to identify and fit the appropriate extreme value distribution to loss data by using the method of Hill estimation technique.

Keywords: Extreme Value Theory; Operational Risk; Hill Estimation Method

1. Uç Değerler Teorisine Giriş

Uç olaylar çoğunlukla beklenmeyen, olağandışı, nadir olarak gözlemlenebilen olaylardır. Ancak birçoğu önemli sonuçlar doğurabilecek niteliktedir. Uç olayların ortaya çıkabilme ihtimalinin hesabı birçok disiplin için oldukça önemlidir. Ancak ortaya çıkışlarındaki az sıklık ve genellikle karmaşık yapıları onlar üzerindeki çalışma zorluğunu da beraberinde getirmektedir. Uç olaylar sonucu ortaya çıkan uç değerler ise, bir veri seti içerisindeki en büyük ya da en küçük değerlerdir. Hemen hemen her veri seti içerisinde uç değerleri barındırır.

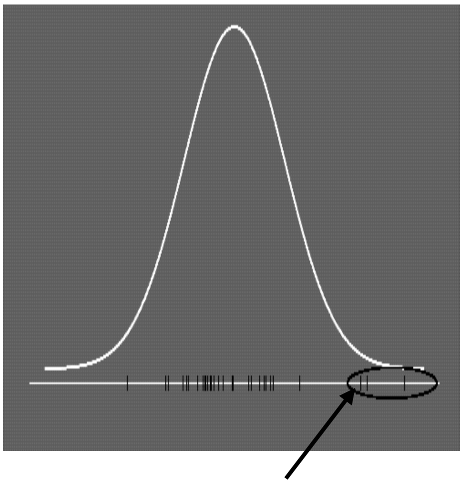
[1] Dr., İstanbul Aydın Üniversitesi Öğretim Görevlisi
[2] Dr., İstanbul Üniversitesi İşletme Fakültesi Sayısal Yöntemler Bilim Dalı

Standart istatistik analizlerde, genellikle verilerin genel eğilimine odaklanılarak verileri temsil eden bir ortalama hesaplanır. Bu tip analizlerde, uç değerler özellikle devre dışı bırakılarak elimine edilmektedir (Regresyon hesaplamalarında). [3] Bu açıdan bakıldığında uç değerler, genel olarak iki nedenden ötürü analiz edilir ya veri grubundan ayrıştırmak için ya da temel birer kaynak olarak düşünülüp üzerinde çalışmak içindir. [4]

Eğer araştırmacı tarafından bu değerler analiz edilmek istenirse, bu değerlerin analizi için kullanılacak yöntemlerden biri Uç Değerler Teorisi'dir (UDT). Bu teori Şekil1'de görüldüğü gibi dağılımın merkezindeki gözlem değerlerinden ziyade, dağılımın kuyruğundaki verilerle ilgilenir. Bu teori dağılımın kuyruk davranışlarını üzerinde çalışmak için oldukça sağlam bir yapıdadır. [5]

Rastsal değişkenlerin toplamı modellendiği zaman merkezi limit teoreminin oynadığı rolün aynısı, yine rastsal değişkenlerin uç değerleri modellendiği zaman UDT tarafından oynanır. Her iki durumda da teori bize limit dağılımlarının ne olması gerektiğini söyler. [6]

Şekil 1 Dağılımların Uç Değerleri



1.1 Uç Değerler Teorisinin Gelişimi

Uç olayların ihtimallerini hesaplamak, birçok bilim dalı için çok önemlidir. Uç değerler teorisi ilk olarak Fréchet tarafından 1927 yılında oluşturulmaya başlanmış ve bir yıl sonra Fisher & Tippet [7] tarafından 1928 yılında yayımlanan makale ile geliştirilmiştir. Gnedenko tarafından 1943 yılında yapılan çalışmalara ek olarak, EJ Gumbel (1958)[8] teoriyi bir kitapla pekiştirmiştir.

---

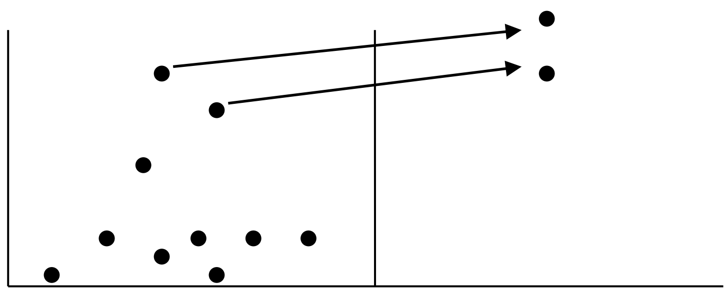
[3] Holt, Tom. "Extreme Value Theory", http://www.cru.uea.ac.uk/cru/projects/mice/ MICE_methods.pdf, Erişim:12 Temmuz 2011
[4] Kalaycı, Şeref . "SPSS Uygulamalı Çok değişkenli İstatistik Teknikleri"
[5] Gençay,R.; Selçuk F.(2004) "Extreme value theory and Value-at-Risk:Relative performance in emerging markets" International Journal of Forecasting
[6] Önalan, Ömer. İktisadi ve İdari Bilimler Fakültesi Dergisi Cilt:XVIII Sayı:1 İstanbul syf 427

[7] Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proceeding of Cambridge Philosophical Society
[8] Gumbel, E. (1958), Statistics of Extremes, Colombia University Press, New York

UDT birçok alanda kendine uygulama bulmuş olan bir teoridir. Başlıca uygulama alanları hidroloji (Smith, 1989; Davison - Smith, 1990; Coles - Tawn, 1996), sigortacılık (Beirlant, 1994; Mikosch, 1997; McNeil, 1997; RootzNen-Tajvidi, 1997) finans (Danielsson de Vries, 1997; McNeil, 1998, 1999; Embrechts., 1998) telekomunikasyon ve uç olaylarla ilgili diğer tüm alanlardır.

UDT olasılık dağılımlarının kuyruklarına odaklanarak nadir olaylara teorik bir çerçeve sunar. Bu teorinin cazip bir tarafı ön varsayım yapılmadan uygulanmasıdır. [9] Bilimsel olarak UDT geçmiş uç olaylar ile geleceği öngörmeye çalışan bir teoridir. Bu bakımdan geçmiş veriler önemlidir.

Şekil 2 Geçmiş Dönem Verileri Uç Değerler Teorisi ile Öngörme



Uç değerler istatistiksel olarak iki şekilde analiz edilmektedir Bunlardan biri uç olayların sayılarını ve diğer ise uç olayların büyüklüğünü ölçmektedir .[10]

Nadir olayların sayımı prosedürü ile, belirli zaman aralığı veya bir dizi deney sonucunda oluşan nadir olayların sayısı sayılmaktadır. İstatitistiksel olarak aşım değerleri kesikli dağılımlar ile modellenirler. Ortaya çıkan problemin özelliğine göre Binom, Poisson, Geometrik ve hipergeometrik dağılımlar kullanılabilir.

Nadir gerçekleşen olayların dağılımı genellikle Poisson dağılımı ile gösterilir. Bu yüzden bu dağılım nadir olayların dağılımı olarak kabul edilir. Az sıklıkta gerçekleşen olayların ortaya çıkma sıklığını ölçer ancak olayların büyüklüğünü ölçmez. Bu sebeple bu dağılım bir sıklık dağılımıdır ve nadir olayların genel eğilimini ortaya koymaz.

x bir tesadüfi değişken olmak üzere, poisson dağılımı olasılık fonksiyonu aşağıdaki gibidir;

$$p(x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad \textbf{(1)}$$

[9] Coles, Stuart. “An introduction to statistical modelling of etreme values” syf 45
[10] Robert R. Kinnison, Applied Extreme Value Statistics, Battelle Press, 1985

μ (lamda) belirli bir zaman veya alan aralığında olayın gerçekleşme sayısının ortalamasını gösteren 0'dan büyük bir sayısıdır.

e ise doğal logaritma tabanıdır (e= 2,71828) [11]

1.2 Uç Değerler Teorisindeki Temel Yaklaşımlar

Uç Değerler Teorisi'nde iki temel yaklaşım bulunmaktadır. İlki maksimum ve minimum gerçekleşmelerin dağılımlarının modellenmesidir. Belirli koşullar altında, serilerin maksimum değerlerinin dağılımı Gumbel (Fisher-Tippett tip I), Frechet (Fisher-Tippett tip II) ya da Weibull (Fisher-Tippet tip III) dağılımına yakınsar. Fisher-Tippet'ın sonuçlarına dayanarak, normal koşullar altında, uç değerler dağılımları için sınırlı bir dağılım formu mevcuttur (Fisher - Tippet,1927).

Bu üç dağılım tek bir dağılım adı altında Genelleştirilmiş Uç Değerler (GUV) dağılımı adı altında birleştirilmiş (unified) bir dağılımdır.[12]

Fisher-Tippet-Gnedenko teorisi ile, oldukça geniş olan n tane gözlemin maksimum gözlemlerinden, bağımsızca özdeş bir şekilde dağılan kayıp verilerin ($x_1, x_2, x_3............, x_n$), genel koşullar altında yaklaşık olarak GUV dağılımı şeklinde dağılır. Bu dağılımının kümülatif olasılık dağılımı fonksiyonu aşağıdaki gibidir;

$$F(x) = \exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\} \qquad \xi \neq 0 \qquad \textbf{(2)}$$

$$\exp\left\{-\left[\exp\left(\frac{x-\mu}{\sigma}\right)\right]\right\} \qquad \xi = 0 \qquad \textbf{(3)}$$

Yukarıdaki kümülatif fonksiyonlarda μ (ortalama) yer parametresi, σ (standart sapma) skala parametresi ve ξ (biçim) ise kuyruk indeks parametresidir. GUV dağılımı 3 türlü forma sahiptir.

$\xi > 0$ ise dağılım Frechet dağılımı
$\xi < 0$ ise dağılım Weibull dağılımı
$\xi = 0$ ise dağılım Gumbel dağılımı olarak belirtilir.

Diğer yaklaşım ise belirli bir eşik değerinin üzerindeki zarar değerinin hangi olasılıklarla ortaya çıkabileceğinin hesaplanması olan Eşik Aşım Teoremi'dir. Bu teoremde yalnızca belli bir eşiği aşan verilere bakılır.

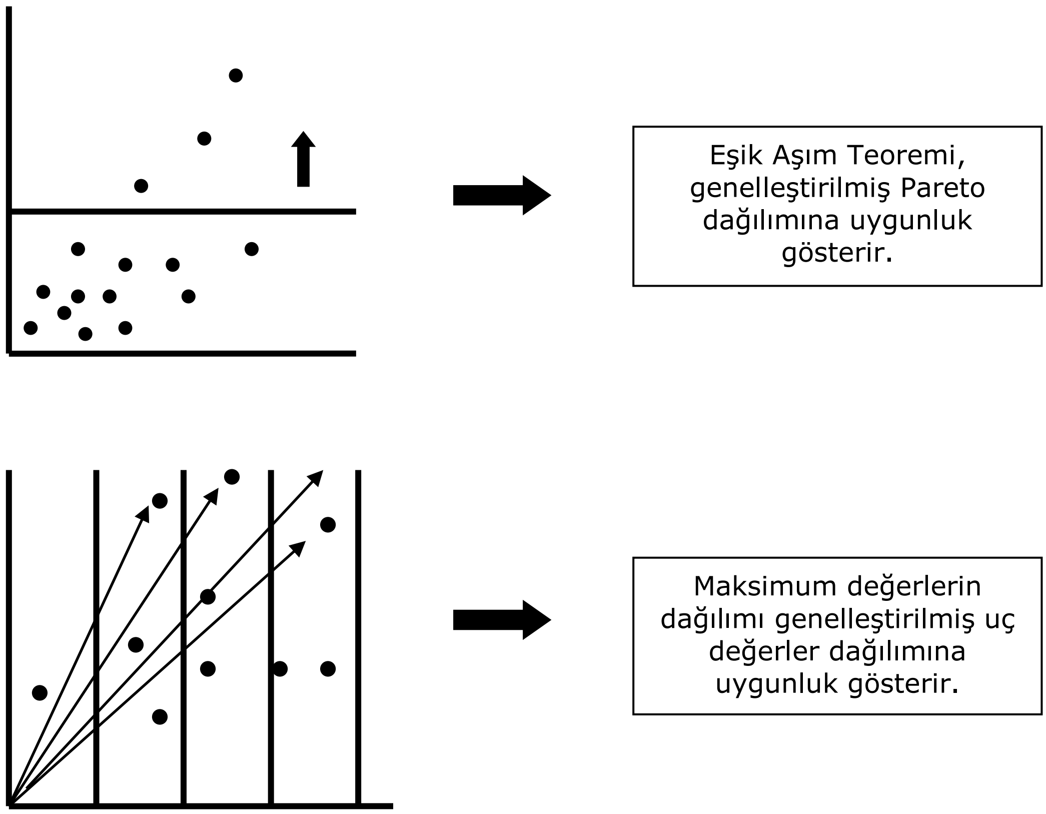
[11] Orhunbilge, Neyran (2000). "Tanımsal İstatistik Olasılık ve Olasılık Dağılımları" syf:203 İstanbul
[12] Reiss, R.D. and M. Thomas, (1997), *Statistical Analysis of Extreme Values* Birkhauser Verlag, Boston, MA

Eşik aşım modeli, genellikle uç değerleri ölçmek için kullanılan bir diğer önemli dağılım olan Pareto dağılımını kullanmaktadır. Sahip olduğu 3 parametre $G_{\mu,\varepsilon,\varphi}$ ile bu dağılım aşağıdaki gibi tanımlanmaktadır ;

$$G_{\mu,\varepsilon.\psi}(y) = \begin{cases} 1 - e^{-z} & \varepsilon = 0 \\ 1 - (1 + \varepsilon z)^{-1/\varepsilon} & \varepsilon \neq 0 \end{cases} \tag{4}$$

$$\begin{array}{ll} z \geq 0 & \varepsilon \geq 0 \\ 0 \leq z \leq -1/\varepsilon & \varepsilon < 0 \end{array} \tag{5}$$

Şekil 3 İki Yaklaşımdaki Farklılıklar



2. İstatistik Dağılımların Uygunluk Sınamaları

Bir örnek kütlenin hangi populasyon dağılımına uygun düştüğü kararını verebilmek için genellikle uygunluk istatistiklerinden faydalanılır.

Çalışmada kullanılan verilerin hangi dağılıma uygun olduğunu Olasılık-Olasılık testi (P-P plot) ile ölçmeye çalışırken, veri setlerinin belirli bir dağılıma ne kadar uyup uymadığının sınanması yapılmaktadır. Eğer seçilen dağılımın doğru bir model olduğu tespit edilirse noktaların yaklaşık olarak bir doğru üzerinde olması gerekmektedir.

2.1 Parametre Tahmini

Bir örnek gurubun olasılık dağılımının parametrelerinin tahmin edilmesi (mühendislikte ve çeşitli bilim dallarında) olayların gelecek dönem davranışlarını öngörebilmek için oldukça önemlidir. Parametrelerin tahmin edilmesi için istatistikte kullanılan birçok teknik mevcuttur. Moment Ağırlıklandırılmış Olasılıklar Metodu, Maksimum Benzerlikler Metodu, Hill Tahmincisi bu yöntemlerden bazılarıdır. Bu çalışmada Hill tahmincisini kullanılarak biçim parametresi tahmin edilmiştir. Hill tahmincisi için iki temel metot kullanabilinir.

Hill tahmincisini kullanarak çeşitli verilerin kuyruk indeks parametresi ξ bulunabilir.[13]

Metot 1

$$\xi_\kappa = \left(\frac{1}{k-1} \sum_{i=1}^{k-1} \ln(x_i) \right) - \ln(x_k) \tag{6}$$

Metot 2

$$\xi_\kappa = \left(\frac{1}{k} \sum_{i=1}^{k} \ln(x_i) \right) - \ln(x_k) \tag{7}$$

3. Araştırma Bulguları

Örnek çalışmada, bir bankanın 3 yıl boyunca karşılaştığı sahtekarlık olayları neticesinde maruz kaldığı kayıp değerleri ile çalışılmıştır. Bu tip kayıplar bankaların riskliliğini arttıran unsurlardır. Burada örnek olarak ele alınan kayıplar bir banka için operasyonel risk olarak tanımlanır. Operasyonel risk ise en temel anlamıyla" başarısız ve yetersiz içsel süreçlerden, çalışan personel ile kullanılan sistemlerden veya bir takım dışsal olaylardan kaynaklanan, dolaylı ya da doğrudan ortaya çıkan risk türü olarak tanımlanmaktadır. [14]

Her bir değer, bir aylık kümülatif toplam kayıp değeridir. UDT teorisi, finansal kurumlar açısından operasyonel risk unsuru olarak kabul edilebilecek olan uç olayları modellemede de kullanılabilmektedir. Tüm hesaplamalar GUD yaklaşımı ile yapılmıştır.

Tablo 1'deki verileri kullanarak uç değerler teorisinin sahtekârlık olayları sonucu ortaya çıkan kayıp değerlerinin gelecek dönem tahmin hesaplanmalarında nasıl kullanılacağı gösterilecektir. Hesaplamalar için gözlemlenen veriler $\{x_1, x_2, x_3, \ldots, x_n\}$ büyükten küçüğe doğru sıralanmıştır.

---

[13] Da Costa Lewis,Nigel Operational Risk with Excel and VBA: Applied Statistical Methods for Risk Management s 201

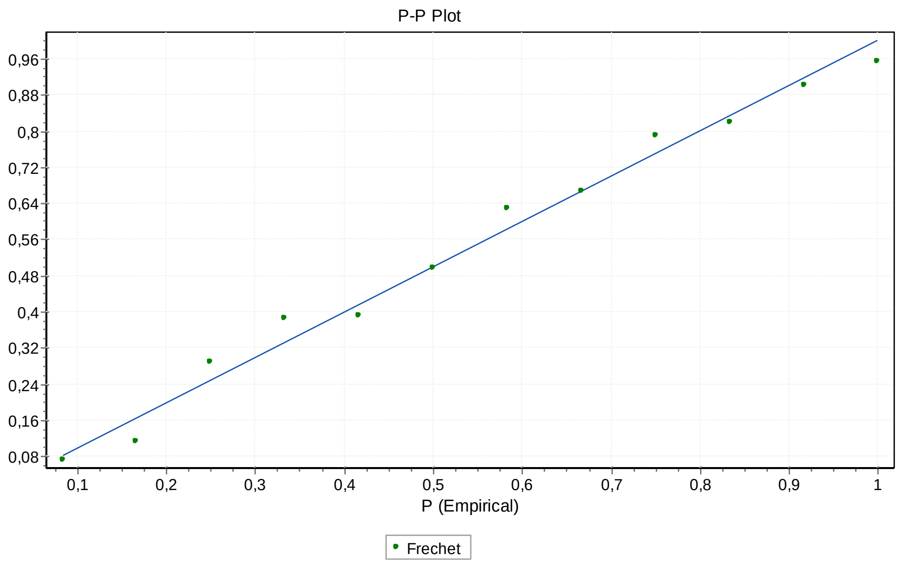
[14] Boyacıoğlu, Melek Acar." Operasyonel Risk ve Yönetimi Bankacılar Dergisi Sayı 43, 2002

Tablo 1

Zaman / Zarar Miktarı (.000$)	T_1	T_2	T_3
1	652	904	3002
2	380	820	850
3	250	691	704
4	222	520	532
5	155	440	500
6	142	300	444
7	108	290	300
8	89	270	260
9	88	160	150
10	74	95	123
11	51	80	80
12	45	78	78

Tablo 2'de görüldüğü gibi her bir T yılı için yapılan uygunluk testi sınamalarında yaklaşık olarak her bir senenin dağılımın yapılan olasılık-olasılık testleri sonucunda, önemli bir uç değer dağılımı olan Frechet dağılımına yakınsadığı görülmektedir. [15]Ayrıca bu dağılımın her üç yıl içinde uygun olduğu elde edilen biçim parametresi değerlerinden de anlaşılabilmektedir.

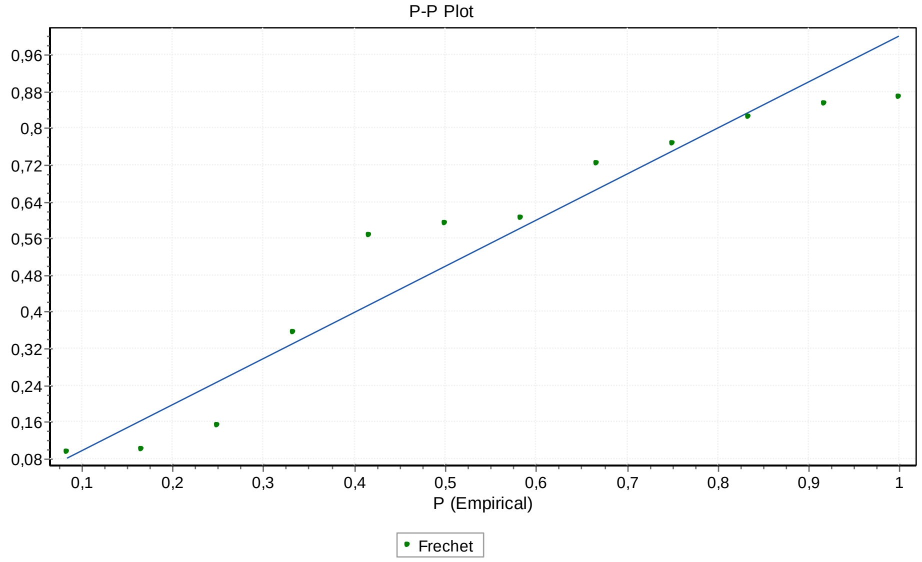
Tablo 2 Her bir sene için P-P Plot Sınamaları

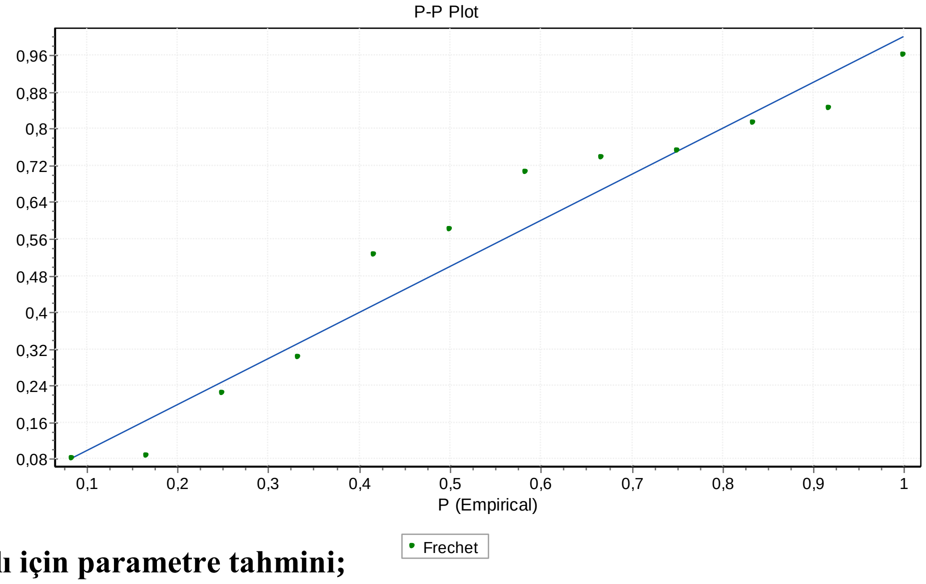


[15] Easy Fit 5.5 sürümü ile elde edilmiştir.

Tablo 3

T_1 **T**





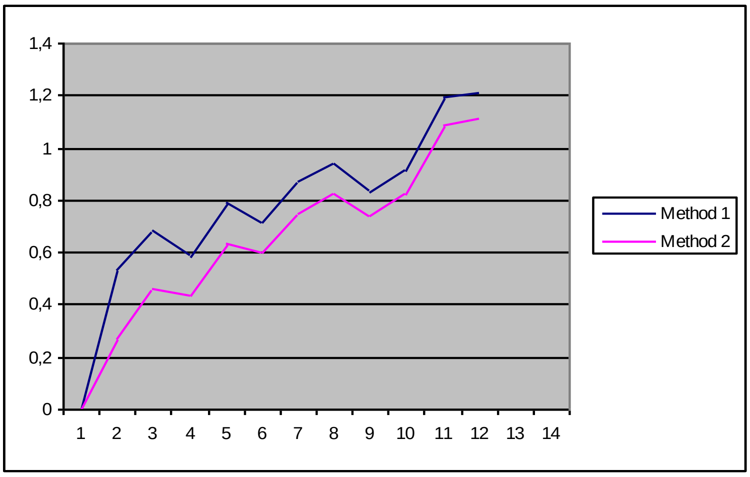
T_1 yılı için parametre tahmini;

Parametre μ (aritmetik ortalama) ve σ (standart sapma) örnek kütle üzerinden hesaplanabilmektedir. Uç değerlerin dağılımlarının en önemli parametresi biçim parametresi olan ξ dir. Eğer veriler uç dağılımlara tam olarak uygunluk gösterirse, biçim parametresi daha anlamlı olacaktır. Kayıp verileri büyükten küçüğe sıralayarak $x_1 > x_2 > x_3 > \ldots\ldots\ldots\ldots x_n$, kuyruk index parametresi ξ 'yi Hill tahmincisi ile tahmin edilmiştir. İki metot kullanılarak bulunulan biçim parametresi ve 2 teknik arasındaki farklılıklar Tablo 3'de gösterilmiştir.

Tablo 3

Aylık	Kayıplar (.000$)	lnx	ξ (Metot 1)	ξ (Metot 2)
1	652	6,476972	0	0
2	380	5,940171	0,5368011	0,26840056
3	250	5,521461	0,6871109	0,45807393
4	222	5,393628	0,5859073	0,43943047
5	155	5,043425	0,7896329	0,63170632
6	142	4,955827	0,7193044	0,59942032
7	108	4,682131	0,8731161	0,74838527
8	89	4,488636	0,9418801	0,82414511
9	88	4,477337	0,8354447	0,74261748
10	74	4,304065	0,9158892	0,82430028
11	51	3,931826	1,1965397	1,0877634
12	45	3,806662	1,2129265	1,11184933

Şekil 4 T_1 yılı için ξ değeri farklı 2 metotla



μ =187,6667

σ =174,97913 ξ =0,7745461 (Metot 1)

ξ =0,64467437(Metot 2)

Daha önce bahsedildiği gibi GUD dağılımları üç forma sahiptir. Eğer ξ > 0, dağılım yapısı Frechet dağılımına uygun düşmektedir. Çalışmada da böyle olduğu hesaplanmıştır.

Örneğin T_1 yılı için 380.000 USD (Kurum için önemli olabilecek bir eşik değer olduğu düşünülsün) üzerinde bir kayıpla gelecek dönemlerde karşılaşma olasılığı %32 olarak hesaplanmıştır. Böyle bir olasılığın ne kadar bir zamanda oluşabileceği ise 1/.32 $\cong 3$ yıldır. Bu hesaplamalar yapılır çünkü risk yöneticileri kendileri için önemli olan belirli bir eşik değerinin üzerindeki bir zarar değerinin hangi olasılıklarla ortaya çıkabileceğini bilmek ister.

$P(X_j > 380) = 1 - G(380)$ $P(X_j > 380) = 1 - 0,68 = 0,32$

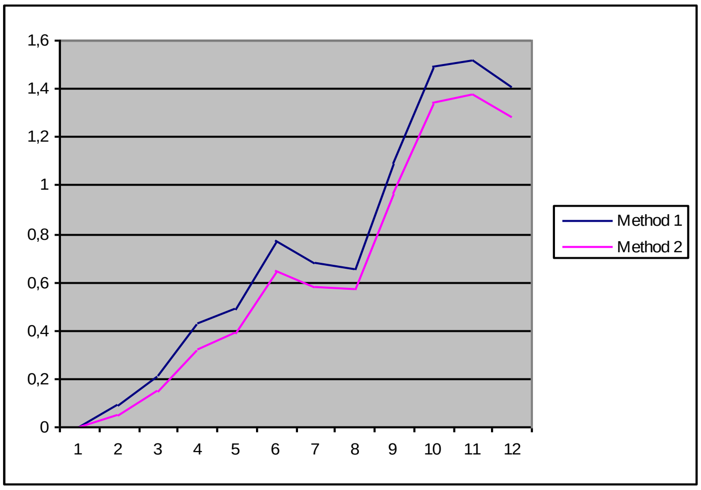
Bu hesaplamalar her bir yıl için ayrı ayrı yapıldığında ise aşağıdaki sonuçlara ulaşılır;

Tablo 4

T_2 **yılı için**

Aylık	**Kayıplar (.000$)**	**lnx**	**ξ (Metot 1)**	**ξ (Metot 2)**
1	900	6,802395	0	0
2	820	6,709304	0,09309	0,046545
3	690	6,536692	0,219158	0,146105
4	520	6,253829	0,428968	0,321726
5	440	6,086775	0,48878	0,391024
6	300	5,703782	0,774016	0,645014
7	290	5,669881	0,678915	0,581927
8	270	5,598422	0,653386	0,571713
9	160	5,075174	1,094961	0,973299
10	95	4,553877	1,494596	1,345136
11	80	4,382027	1,516986	1,379079
12	78	4,356709	1,404396	1,287363

Şekil 5 T_2 yılı için ξ değeri farklı 2 metotla

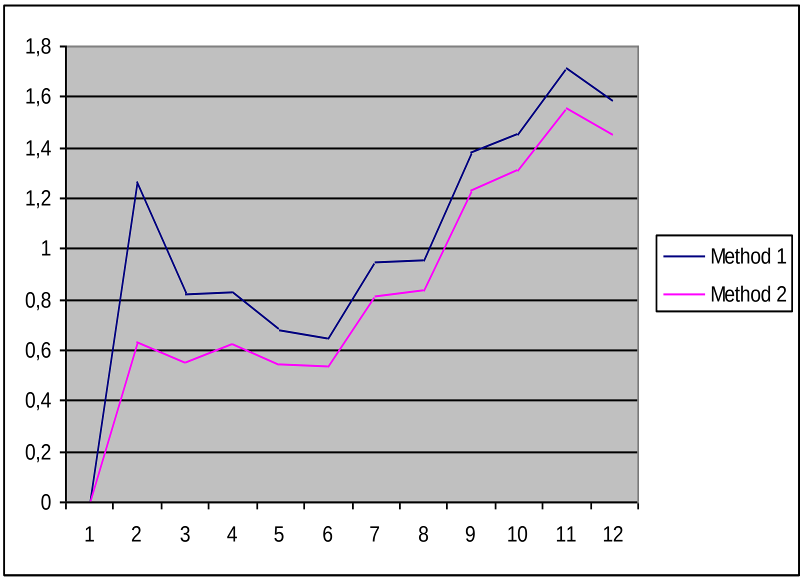


Tablo 5

T_3 **yılı için**

Aylık	Kayıplar	lnx	ξ (Metot 1)	ξ (Metot 2)
1	3000	8,006368	0	0
2	850	6,745236	1,261131	0,630566
3	700	6,55108	0,824722	0,549814
4	530	6,272877	0,828018	0,621013
5	500	6,214608	0,679282	0,543426
6	450	6,109248	0,648786	0,540655
7	300	5,703782	0,94612	0,81096
8	260	5,560682	0,954061	0,834803
9	150	5,010635	1,38485	1,230978
10	120	4,787492	1,454121	1,308709
11	80	4,382027	1,714174	1,55834
12	78	4,356709	1,583658	1,451686

Şekil 6 T_3 yılı için ξ değeri farklı 2 metotla



4. Sonuç

Finansal kurumlar için geçmiş zarar verilerini kullanarak gelecekte belirli bir eşik değerinin üzerindeki bir zarar değerinin, hangi olasılıklarla ortaya çıkabileceğini hesaplanması oldukça önemlidir. Çünkü bu sayede finansal kurumlar, düşük olasılıklarla oluşabilecek muhtemel zararları karşılıklar ayırarak, kendilerini sigorta ettirerek veya teknoloji üzerine çeşitli yatırımlar yaparak etkisini azaltmaya çalışırlar.

Yapılan çalışma özellikle ortalama üstünde kalan zarar verilerinin hangi olasılıklarla ortaya çıkabileceğini hipotetik bir örnek üzerinden hesaplayarak, kurumların nasıl bir risk ile karşı karşıya kaldıklarını ortaya koymaktadır.

Kurumlar bu şekilde, hesaplanılan olasılıklar ile ileriye yönelik tahmin yürüterek kendilerini gelecek dönemlere hazırlarlar.

Kaynaklar

Boyacıoğlu, Melek Acar “Operasyonel Risk ve Yönetimi” Bankacılar Dergisi Sayı 43, 2002
Coles, Stuart. “An introduction to statistical modelling of etreme values” syf 45

Da Costa Lewis,Nigel Operational Risk with Excel and VBA: Applied Statistical Methods for Risk Management . syf 201

Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proceeding of Cambridge Philosophical Society

Gençay,R. And Selçuk F.(2004) “Extreme value theory and Value-at-Risk:Relative performance in emerging markets” International Journal of Forecasting

Gumbel, E. (1958), Statistics of Extremes, Colombia University Press, New York

Holt, Tom. “Extreme Value Theory”, http://www.cru.uea.ac.uk/cru/projects/mice/ MICE_methods.pdf, (Erişim:12 Temmuz 2009

Kalaycı, Şeref . “SPSS Uygulamalı Çok değişkenli İstatistik Teknikleri"

Orhunbilge, Neyran (2000). “Tanımsal İstatistik Olasılık ve Olasılık Dağılımları” syf:203 İstanbul

Önalan, Ömer. İktisadi ve İdari Bilimler Fakültesi Dergisi Cilt:XVIII Sayı:1 İstanbul Syf 427

Reiss, R.D. and M. Thomas, (1997), *Statistical Analysis of Extreme Values* Birkhauser Verlag, Boston, MA

Robert R. Kinnison, Applied Extreme Value Statistics, Battelle Press, 1985