

KEŞFEDİCİ VERİ ANALİZİ TEKNİKLERİNİN SÜREÇ YETENEK ANALİZİ ÇALIŞMALARINDA UYGULANMASI

Cenk ÖZLER*

ÖZET

Keşfedici veri analizi (KVA), bir veri yığınının yapısını iyi anlamak üzere, çeşitli tekniklerin kullanıldığı bir veri analizi yaklaşımıdır. Bu çalışmada, verilerin yapıları hakkında hızlı ve kolay bir şekilde bilgi sağlayan ve en sık kullanılan keşfedici veri analizi teknikleri kısaca tanıtılmıştır. Ardından, simetrik dağılım göstermeyen verileri simetrik hale dönüştürme amacı ile uygulanacak kuvvet dönüşümünün belirlenmesi ile ilgili bir yöntem olan “simetri için dönüşüm” yöntemi anlatılmıştır. Son olarak bir süreç yetenek çalışması uygulamasında karşılaşılmış olan simetrik olmayan veriler üzerinde simetri için dönüşüm yönteminin uygulaması gerçekleştirilmiş ve dönüştürülmüş veriler üzerinden süreç yetenek çalışması gerçekleştirilmiştir. Süreç yetenek çalışmalarında kuvvet dönüşümünün uygulanmasının avantaj ve dezavantajları tartışılmıştır.

Anahtar Kelimeler: Keşfedici Veri Analizi, Kuvvet Dönüşümü, Süreç Yetenek Analizi

Applying Exploratory Data Analysis Techniques in Process Capability Analysis

ABSTRACT

Exploratory data analysis (EDA) is an approach for data analysis that employs a variety of techniques to maximize our understanding of a data set and uncover underlying structures. This study gives basic information to most popular EDA techniques which provide quick and easy insight into data structures. Then transformation for symmetry technique which is used for transforming non-symmetric data by a power transformation is introduced. An application of the technique is presented for non-symmetric data observed in a process capability study. After the transformation, symmetric data is used for capability calculations. Pros and cons of using power transformations for process capability analyses are also discussed.

Key Words: Exploratory Data Analysis, Power Transformation, Process Capability Analyses

1. GİRİŞ

Klasik istatistiksel teknikler, oldukça katı olan bazı varsayımlar (verilerin simetrik olması, normal dağılım göstermesi, sapan gözlem olmaması vb.) geçerli olduğunda en iyi olacak şekilde tasarlanmışlardır. Bununla beraber, uygulamalarda karşılaşılan ve klasik analizlerin gerektirdiği varsayımların sağlanmadığı durumlarda, bu tekniklerin uygulamaları oldukça yanıltıcı sonuçlar verebilir. Bu amaçla son yıllarda geliştirilen sağlam (robust) ve keşfedici yöntemler, istatistiksel analizlerin etkinliğini arttırmıştır.

Keşfedici veri analizi (KVA) teknikleri, verilerin örüntü yapıları hakkında hızlı ve kolay bir şekilde bilgi sahibi olmamıza yardımcı olmaktadır. KVA tekniklerinin kullanılması sonucu verilerin yapısı, üzerinde yapılacak analizler (regresyon analizi, varyans analizi, istatistiksel süreç kontrolü, süreç yetenek analizi vb.) için gereken varsayımları sağlamıyorsa veriler yeniden ifade

* Yrd. Doç. Dr., DEÜ İİBF Ekonometri Bölümü İstatistik Anabilim Dalı.

edilerek, diğeri bir deyişle dönüştürülerek tekrar KVA'ne tabi tutulabilir. KVA sonucunda veriler, analizlerin gerektirdiği varsayımları sağlamıyorsa verilerin dağılışındaki dalgalanmalara daha dirençli olan sağlam teknikler kullanılabilir. KVA teknikleri için Tukey (1977) ve Hoaglin vd. (1983)'e, sağlam teknikler için ise yine Hoaglin vd. (1983) ve Huber (1981)'e başvurulabilir.

Günümüzde istatistik tekniklerin en yaygın olarak uygulandığı sektörlerin başında otomotiv parçası üreten kuruluşlar gelmektedir. Bu kuruluşlar, müşterisine sattığı her parçanın seri üretiminden önce, üretim süreçlerinin müşterilerinin şartlarını karşıladığını kanıtlayabilmek için *Süreç Yetenek Analizlerini* gerçekleştirmek zorundadır. Süreç yetenek analizleri ise ön seri üretimden elde edilen ölçümlerin dağılışılarının normal dağılışı olmasını gerektirmektedir. Ardından bu verilerden hesaplanan süreç yetenek indeksleri kullanılarak üretim sürecinin müşteri şartlarını karşılamada yeterli olup olmadığına karar verilmektedir. Literatürde, verilerin normal dağılışı göstermediği durumlar için bazı alternatif indeksler önerilmektedir (bkz. Kotz ve Johnson, 1993). Ancak bu indekslerin kullanımı pratikte pek kolay değildir ve bu sebeple kuruluşlardaki kullanımı da yaygınlaşmamıştır.

Bu çalışmada öncelikle literatürde geçen ve yaygın olarak kullanılan bazı KVA teknikleri açıklanmıştır. Ardından Hoaglin vd. (1983)'te verilen, verileri simetrik hale getirmek için gerçekleştirilen bir veri dönüşüm yöntemi açıklanmıştır. Daha sonra süreç yetenek analizi amacı ile toplanmış verilerin KVA teknikleri ile analizi gerçekleştirilmiş ve analiz sonuçlarına bağlı olarak simetriyi sağlamak için veriler dönüştürülerek süreç yetenek analizi yapmak için uygun hale getirilmiştir. Son olarak, veri dönüşümünün avantajları ve dezavantajları tartışılmıştır.

2. KEŞFEDİCİ VERİ ANALİZİ TEKNİKLERİ

2.1 Yaygın Olarak Kullanılan Keşfedici Teknikler

Verilerin yapısının anlaşılmasında en yaygın olarak kullanılan görsel özetler, *gövde ve yaprak gösterimi* ve *kutu grafikleridir*. Verilerin sayısal olarak özetlenmesi için ise genellikle sıra istatistiklerine dayanan *harf değerleri* (5 sayı özeti, 7 sayı özeti vb.) kullanılmaktadır. Bu özetler uygulayıcılara,

- veri yığınının ne kadar simetrik olduğu,
- sayıların yayılmasının ne olduğu,
- diğerlerinden uzakta olan verilerin varlığı,
- verilerin çoğunluğunun nerede toplandığı,
- verilerin aralarında boşluklar olup olmadığı

hakkında bilgi sunarlar.

İncelemek istediğimiz veriler,

24,6 29,9 21,7 18,7 18,2 17,7 16,7 19,9 16,1 17,6 25,5

olsunlar. Bu veriler için gövde ve yaprak gösterimi Tablo 1'deki gibidir. Gövde ve yaprak gösterimi bir veri değerindeki sayıdaki rakamların gövde ve yaprak olmak üzere ikiye ayrılması ile oluşturulmaktadır. İlk veri değeri 24,6'daki ilk rakam 2 gövde olarak, ikinci rakam 4 ise yaprak olarak ayrılmıştır. Tablo 1'de en soldaki sütun, satırın sonundaki verilerin *derinliğini* vermektedir. Derinlik, veriler küçükten büyüğe ve büyükten küçüğe sıralandıklarında sahip oldukları iki sıra numarasından küçük olanıdır. Medyanın bulunduğu satırda ise derinlik yerine parantez içerisinde o satırın sıklığı verilir. İkinci sütunda veri değerleri için gövdeler, sağındaki satır boyunca da yaprak değerleri verilmektedir. Verilerin gövde ve yaprak olarak nasıl ayrılacağına ise oluşturulacak ideal satır sayısını belirleyerek karar vermek gerekmektedir. Hoaglin vd. (1983) satır sayısının yaklaşık olarak, $L = 2\sqrt{n}$ veya $L = [10 \times \log_{10}n]$ şeklinde belirlenebileceğini söylemiştir. Buradaki köşeli parantez, içindekinin *tam değerini* ifade etmektedir.

Aynı veriler için kutu grafiğini oluşturursak Şekil 1'deki gösterimi elde edebiliriz. Kutu grafiğindeki kutunun ortasındaki çizgi ortancayı (medyanı) göstermektedir. Kutunun kenarlarından aşağıda olanı alt dördüncü (lower fourth) yada 1. kartil (F_L), yukarıda olanı ise üst dördüncü (upper fourth) yada 3. kartildir (F_U). Kutunun altından ve üzerinden çizilen dikey doğruların uzunluğu ise *sapan değer* (outlier) olmayan son verinin değerine kadardır. Eğer verilerde sapan değer varsa bu doğruların uzağında ayrıca gösterilir. Verilerdeki sapan değerlerin hesaplanması için dirençli bir değişkenlik ölçüsü olan *dördüncü-yayılmayı* (fourth-spread = d_F),

$$d_F = F_U - F_L$$

kullanılmaktadır. Buradan sapan değerler için *sınır noktaları* (outlier cutoffs), $F_L - 1,5d_F$ ve $F_U + 1,5d_F$ şeklinde belirlenebilir. Yani F_L 'nin $1,5d_F$ altında olan veya F_U 'nin $1,5d_F$ üzerinde olan gözlemler diğerlerinden uzakta olan gözlemler olarak tanımlanırlar. Şekil 1'e baktığımızda, verilerin dağılımının sağa çarpık olduğunu ve sapan gözlemin bulunmadığını söyleyebiliriz.

Tablo 1: Örnek veriler için gövde ve yaprak gösterimi MINITAB çıktısı

Stem-and-leaf of C1 N = 11

Leaf Unit = 1,0

4 1 6677

(3) 1 889

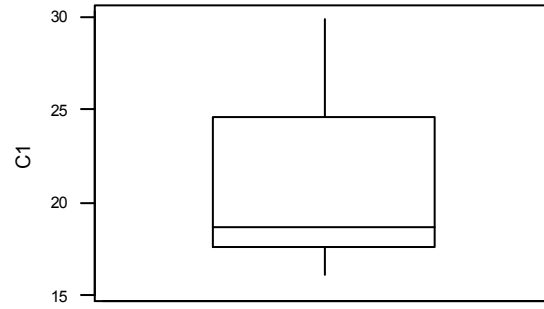
4 2 1

3 2

3 2 45

1 2

1 2 9



Şekil 1: Örnek verileri için kutu grafiği MINITAB çıktısı

Verilerin sayısal olarak özetlenmesinde kullanılan bir yaklaşım ise *harf değerleri gösterimidir*. Harf değerleri gösteriminin en basit hali *5 sayı özetidir* ve Tablo 2'deki gibi gösterilir. Böyle bir tabloda kutu grafiğindeki bilgilerin yanında medyan ve dördüncülerin derinlikleri ile, verilerdeki en büyük ve en küçük değerleri de görmek mümkündür. Verilerin sayısı uygun olduğunda, bu tablo içerisine *sekizinciler*, *onaltıncılar* vs. ilave edilerek *7 sayı özet*, *9 sayı özet* vs. elde edilebilir. Tablo 2'nin en solunda kullanılan medyan, dördüncü, sekizinci vs. harflerle Tablo 3'teki gibi sembolize edilmektedir.

Tablo 2: 5 sayı özeti

n			
M	Medyanın derinliği	medyan	
H	Dördüncünün derinliği	Alt dördüncü	Üst dördüncü
1		Alt uç değer	Üst uç değer

Tablo 3: Özet sıra istatistiklerinin sembolleri

Sembol	Kuyruk alanı	Sıra İstatistiğinin Adı
M	1/2	Medyan
H veya F	1/4	Dördüncü
E	1/8	Sekizinci
D	1/16	Onaltıncı
C	1/32	Otuzikinci
B	1/64	Atmışdördüncü
A	1/128	Yüzyirmisekizinci
Z	1/256	İkiyüzellialtıncı
Y	1/512	Beşyüzonikinci
X	1/1024	Binyirmidördüncü

Yukarıdaki örnek verileri için harf değer gösteriminin MINITAB çıktısı Tablo 4 ile verilmiştir. MINITAB, Tablo 2'deki verilerin yanında harf özetlerinin ortalamalarını (Mid) ve farklarını (Spread) da vermektedir. Mid değerlerine bakıldığında bir artış eğilimi olduğu görülmektedir ki, bu da verilerin dağılımının sağa çarpık olduğunu göstermektedir. Diğer taraftan, Mid değerlerinin azalma eğilimi göstermesi verilerin sola çarpık olduğunu, herhangi bir trend göstermemesi ise simetrik olduğuna işaret etmektedir.

Cenk ÖZLER

Harf değerleri gösterimi ve grafiksel özetler ile ilgili diğer bir konu da, harf özetlerinin derinliklerinin bulunmasıdır. Medyanın derinliği, gözlem sayısı n tek ise $(n + 1) / 2$, n çift ise $n / 2$ 'dir. Derinlik, buçuklu çıktığında, kendinden küçük ve büyük derinliğe sahip iki verinin ortalaması hesaplanarak ilgili harf özeti bulunur. Örneğin, medyanın derinliği 5,5 ise, 5. ve 6. gözlemlerin ortalaması alınarak medyan bulunur. Dördüncüler için derinlik,

$$\text{Dördüncünün derinliği} = ([\text{Medyanın derinliği}] + 1) / 2$$

şeklinde hesaplanabilir. Burada köşeli parantez [.], içerisindeki ifadenin tam değerinin alınacağı anlamına gelmektedir. Benzer şekilde sekizincinin derinliği,

$$\text{Sekizincinin derinliği} = ([\text{Dördüncünün derinliği}] + 1) / 2$$

şeklinde hesaplanabilir. Buradan tüm harf değerlerinin derinliklerini bulmak için bir genelleme yapılırsa herhangi bir harf değeri için derinlik,

$$\text{Harf değerinin derinliği} = ([\text{Bir önceki harf değerinin derinliği}] + 1) / 2$$

eşitliğinden hesaplanabilir.

Tablo 4. Örnek veriler için harf değerleri gösterimi MINITAB çıktısı.

Letter Value Display: C1

	Depth	Lower	Upper	Mid	Spread
N=	11				
M	6,0	18,700	18,700		
H	3,5	17,650	23,150	20,400	5,500
E	2,0	16,700	25,500	21,100	8,800
1	1	16,100	29,900	23,000	13,800

Bu bölümde bahsettiğimiz harf değerleri kullanılarak, normal dağılış gösteren veriler için standart sapmanın bir tahmini de kolay bir şekilde bulunabilmektedir. Normal dağılış gösteren veriler için standart sapmanın tahmini olan $F\text{-yalancisigma}$ ($F\text{-pseudosigma}$),

$$F\text{-yalancisigma} = \frac{d_F}{1,349}$$

eşitliğinden hesaplanabilir (Hoaglin vd. 1983).

2.2 Verilerin Simetri Sağlama Amacı İle Dönüştürülmesi

Bir veri yığıını x_1, x_2, \dots, x_n üzerinde yapılan bir kuvvet dönüşümü sonrasında dönüştürülmüş değerler y_1, y_2, \dots, y_n olsun. Hoaglin vd. (1983),

verilerin simetrisini iyileştirmeye yönelik olarak aşağıdaki kuvvet dönüşümünün kullanılabilceğini belirtmiştir:

$$y = \begin{cases} x^p & (p > 0) \\ \log x & (p = 0) \\ -x^p & (p < 0) \end{cases}$$

Hoaglin vd. (1983), kuvvet değeri p 'nin seçimi için, 2. bölümde bahsedilen harf ortalamalarının trendine dayanan bir yöntem tanıtmıştır. Bu yöntemde göre harf değerleri gösterimindeki tüm harf değerleri için,

$$\frac{x_U + x_L}{2} - M = (1 - p) \frac{(x_U - M)^2 + (M - x_L)^2}{4M}$$

eşitliğinden p değerleri hesaplanıp, p 'lerin medyanı uygun kuvvet değeri olarak seçilmektedir.

3. SÜREÇ YETENEK İNDEKSLERİ

Süreç yetenek analizleri, bir kuruluşun üretim sürecinin müşteri ve / veya mühendislik şartlarına uygun ürün üretme yeteneğine sahip olup olmadığını belirlemeye yönelik olarak, özellikle seri üretim öncesindeki ön seri üretim aşamasında gerçekleştirilebilmektedir. Süreç yeteneğinin tespiti amacı ile kullanılan başlıca indeksler şunlardır:

$$\hat{C}_p = \frac{\bar{U}SL - ASL}{6s}, \quad \hat{C}_{pk} = \min \left\{ \frac{\bar{U}SL - \bar{x}}{3s}, \frac{\bar{x} - ASL}{3s} \right\} \quad \text{ve}$$
$$\hat{C}_{pm} = \frac{\bar{U}SL - ASL}{6\sqrt{s^2 + (\bar{x} - T)^2}}$$

Burada $\bar{U}SL$, incelenen kalite karakteristiği için üst spesifikasyon limitini; ASL , alt spesifikasyon limitini; $\bar{x} = \frac{\sum x}{n}$, verilerin aritmetik ortalamasını;

$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$, verilerin örnek standart sapmasını ve T , kalite karakteristiği için hedef değeri göstermektedir. İndeksler ile ilgili detaylar için Montgomery (2001), Kotz ve Johnson (1993) ve Kotz ve Lovelace (1998)'e başvurulabilir.

Bir ürünün kalite karakteristiğine ait yalnızca tek bir spesifikasyon limitinin olması durumunda kullanılan indeksler ise şunlardır:

$$\hat{C}_{pL} = \frac{\bar{x} - ASL}{3s} \quad (\text{yalnızca alt spesifikasyon limiti var ise}),$$

$$\hat{C}_{pU} = \frac{\bar{USL} - \bar{x}}{3s} \quad (\text{yalnızca üst spesifikasyon limiti var ise}).$$

Yukarıda verilen tüm süreç yetenek indekslerinin değerinin mümkün olduğunca yüksek olması istenen bir durumdur. İndeks değerlerinin genellikle 1,33 veya 1,67'den büyük olması istenir. Ancak bu indekslerin kullanılarak karar verilebilmesi için ölçülen değerlerin dağılışının normal olması gerekmektedir.

4. SÜREÇ YETENEK ANALİZİ ÇALIŞMASINDA KEŞFEDİCİ VERİ ANALİZİ TEKNİKLERİNİN KULLANILMASI: BİR ÖRNEK UYGULAMA

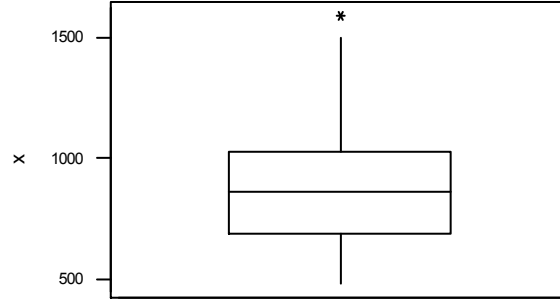
Bu bölümde metal parçalara ısıl işlem yapılan bir atölyede, metal parçaların sertlik değerlerinin müşteri şartlarına uygunluğunu doğrulamak için yapılan bir süreç yetenek çalışması anlatılmıştır. Uygulama yapılan atölye yönetiminin isteği doğrultusunda, bu bölümde gerçek veriler kullanılmayıp, gerçek verilerin üzerine doğrusal dönüşüm uygulanmıştır ve dönüştürülmüş veriler kullanılmıştır. Sertlik için yalnızca alt spesifikasyon limiti mevcuttur ve $ASL = 450$ 'dir. Süreç yeteneğini anlamak için müşteri isteği doğrultusunda 100 parça üretilmiş ve sertlik testine tabi tutulmuştur. Elde edilen ölçümlerin sonuçları Tablo 5 ile verilmiştir. Bu vaka için uygun olan süreç yetenek indeksi \hat{C}_{pL} 'dir.

Süreç yetenek indekslerinin kullanılmasından önce Tablo 5'teki veriler için gövde ve yaprak gösterimi Tablo 6 ile, kutu grafiği ise Şekil 2 ile verilmiştir. Bu şekillere bakıldığında verilerin dağılışının bir miktar sağa çarpık olduğu ve bir adet de sapan değer olduğu görülmektedir. Bu durumda verilerin bu hali ile süreç yetenek indekslerinin hesaplanarak yorumlarda bulunulması doğru değildir.

Sertlik verileri için harf değerleri, Tablo 7 ile verilmiştir. Harf değerlerinin ortalamalarına bakıldığında yukarıdan aşağıya doğru bir artış görülmektedir. Bu da verilerin sağa çarpık olduğuna işarettir.

Tablo 5: 100 Metal Parça İçin Sertlik Değerleri

619	961	902	1146	656
622	1043	678	1147	765
941	1116	998	1048	835
1198	1209	642	908	1460
765	1272	886	926	887
854	670	654	596	782
736	1591	1004	812	850
685	686	921	774	893
871	663	730	914	1118
645	851	999	675	945
904	1417	895	874	612
1305	1215	1157	504	715
747	1158	1187	857	656
1445	1227	941	620	563
1143	851	916	1501	1186
800	972	754	1227	1245
741	525	938	679	1003
821	944	1036	825	845
481	776	592	671	788
662	1008	587	635	754



Şekil 2: Sertlik Verileri İçin Kutu Grafiği

Tablo 6: Sertlik Verileri İçin Gövde ve Yaprak Gösterimi(MINITAB Sonuçları)

Stem-and-Leaf Display: x

Stem-and-leaf of x N = 100

Leaf Unit = 10

```
1  4 8
3  5 02
7  5 6899
14 6 1122344
26 6 555667777788
31 7 13344
39 7 55667788
45 8 012234
(11) 8 55555778899
44 9 000112234444
32 9 6799
28 10 000344
```

Kesfedici Veri Analizi Tekniklerinin Süreç Yetenek Analizi Çalışmalarında Uygulanması

22 10
22 11 11444
17 11 55889
12 12 01224
7 12 7
6 13 0
5 13
5 14 14
3 14 6
2 15 0

HI 159;

Tablo 7. Sertlik verileri için Harf Değerleri (MINITAB Sonuçları)

Letter Value Display: x

	Depth	Lower	Upper	Mid	Spread
N=	100				
M	50,5	864,000	864,000		
H	25,5	685,500	1022,000	853,750	336,500
E	13,0	642,000	1198,000	920,000	556,000
D	7,0	596,000	1272,000	934,000	676,000
C	4,0	563,000	1445,000	1004,000	882,000
B	2,5	514,500	1480,500	997,500	966,000
A	1,5	492,500	1546,000	1019,250	1053,500
1		481,000	1591,000	1036,000	1110,000

Verilerin simetrisini iyileştirecek bir kuvvet dönüşümünü gerçekleştirmek için her bir harf değeri için bölüm 2.2’de verilen hesaplamalar gerçekleştirilmiş ve Tablo 8’deki sonuçlar elde edilmiştir. p değerlerine baktığımızda, p ’lerin medyanı 0,08193’tür. Bu değer sıfıra oldukça yakın bir değer olduğu için verilerin logaritma dönüşümü uygun dönüşüm olarak görülebilir.

Tablo 8: Simetri İçin Uygun Dönüşüm Hesaplamaları

$\frac{x_U + x_L}{2}$	$\frac{(x_U - M)^2 + (M - x_L)^2}{4M}$	$1 - p$	p
-10,25 1,62337	16,443	-0,62337	
56,00	46,539	1,20328	-0,20328
70,00	68,949	1,01524	-0,01524
140,00	123,889	1,13004	-0,13004
133,25	145,141	0,91807	0,08193
155,25	174,519	0,88959	0,11041
172,00	195,376	0,88036	0,11964

Verilerin logaritması alındığında elde edilen yeni veriler Tablo 9 ile verilmiştir. Dönüştürülmüş veriler için gövde ve yaprak gösterimi Tablo 10 ile, kutu grafiği Şekil 3 ile ve harf değerleri Tablo 11 ile verilmiştir. Bu sonuçlara bakıldığında logaritma dönüşümünün verileri simetrik hale getirdiği görülmektedir.

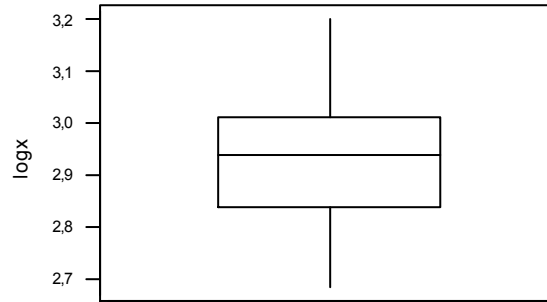
Veriler üzerinde gerçekleştirilen logaritma dönüşümünün simetriyi iyileştirdiği görülmektedir. Bu durumun verilerin normal dağılışa da yaklaşmasına katkıda bulunup bulunmadığını doğrulamak amacı ile hem ham veriler hem de dönüştürülmüş veriler için normal olasılık grafikleri Şekil 4 ve Şekil 5 ile verilmiştir. Sonuç olarak $\log(x)$ verilerinin normal dağılışa uyduğu görülmektedir.

Tablo 9. 100 Metal Parça İçin $\log(\text{Sertlik})$ Değerleri

2,79169	2,98272	2,95521	3,05918	2,8169
2,79379	3,01828	2,83123	3,05956	2,88366
2,97359	3,04766	2,99913	3,02036	2,92169
3,07846	3,08243	2,80754	2,95809	3,16435
2,88366	3,10449	2,94743	2,96661	2,94792
2,93146	2,82607	2,81558	2,77525	2,89321
2,86688	3,20167	3,00173	2,90956	2,92942
2,83569	2,83632	2,96426	2,88874	2,95085

Kesfedici Veri Analizi Tekniklerinin Süreç Yetenek Analizi Çalışmalarında Uygulanması

2,94002	2,82151	2,86332	2,96095	3,04844
2,80956	2,92993	2,99957	2,82930	2,97543
2,95617	3,15137	2,95182	2,94151	2,78675
3,11561	3,08458	3,06333	2,70243	2,85431
2,87332	3,06371	3,07445	2,93298	2,81690
3,15987	3,08884	2,97359	2,79239	2,75051
3,05805	2,92993	2,96190	3,17638	3,07408
2,90309	2,98767	2,87737	3,08884	3,09517
2,86982	2,72016	2,97220	2,83187	3,00130
2,91434	2,97497	3,01536	2,91645	2,92686
2,68215	2,88986	2,77232	2,82672	2,89653
2,82086	3,00346	2,76864	2,80277	2,87737



Şekil 3. log(Sertlik) Verileri İçin Kutu Grafiği

Tablo 10: log(Sertlik) Verileri İçin Gövde ve Yaprak Gösterimi(MINITAB Sonuçları)

Stem-and-Leaf Display: logx

Stem-and-leaf of logx2 N = 100

Leaf Unit = 0,010

```
1 26 8
2 27 0
3 27 2
4 27 5
7 27 677
11 27 8999
17 28 000111
26 28 222223333
27 28 5
33 28 666777
39 28 888899
43 29 0011
50 29 2222233
50 29 444455555
41 29 666677777
32 29 8899
28 30 00011
23 30 2
22 30 44555
17 30 66777
12 30 88889
7 31 01
5 31
5 31 55
```

3 31 67

1 31

1 32 0

Tablo 11: log(Sertlik) verileri için Harf Değerleri (MINITAB Sonuçları)

Letter Value Display: logx

	Depth	Lower	Upper	Mid	Spread
N=	100				
M	50,5	2,936	2,936		
H	25,5	2,836	3,009	2,923	0,173
E	13,0	2,808	3,078	2,943	0,271
D	7,0	2,775	3,104	2,940	0,329
C	4,0	2,751	3,160	2,955	0,409
B	2,5	2,711	3,170	2,941	0,459
A	1,5	2,692	3,189	2,941	0,497
1		2,682	3,202	2,942	0,520

Dönüştürülmüş veriler üzerinden süreç yetenek analizini gerçekleştirmek için *ASL*'nin de logaritmasını almak gereklidir. Ayrıca bu veriler için standart sapma, *F-yalancısigma* kullanılarak tahminlenebilir. Dönüştürülmüş veriler için süreç yetenek analizi sonuçları aşağıdaki gibi elde edilebilir:

$$ASL_{\log x} = \log(450) = 2,65321$$

$$F\text{-yalancısigma} = (3,009 - 2,836) / 1,349 = 0,12824$$

$$\bar{x} = 2,93671$$

$$\hat{C}_{pL} = (2,93671 - 2,65321) / (3 \times 0,12824) = 0,7369$$

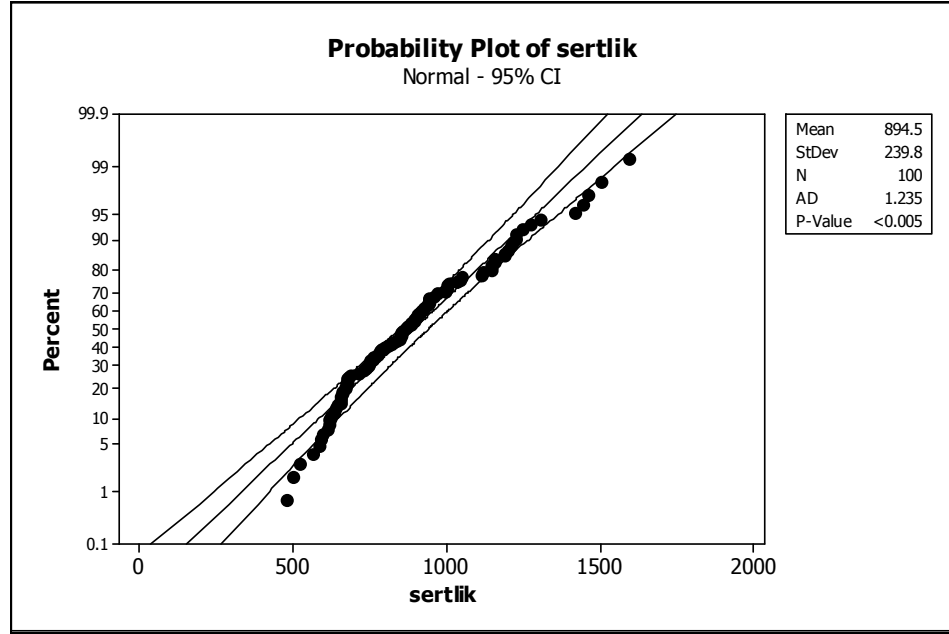
\hat{C}_{pL} değerine bakıldığında süreç yeteneğinin zayıf olduğu sonucuna varılabilir.

5. SONUÇ

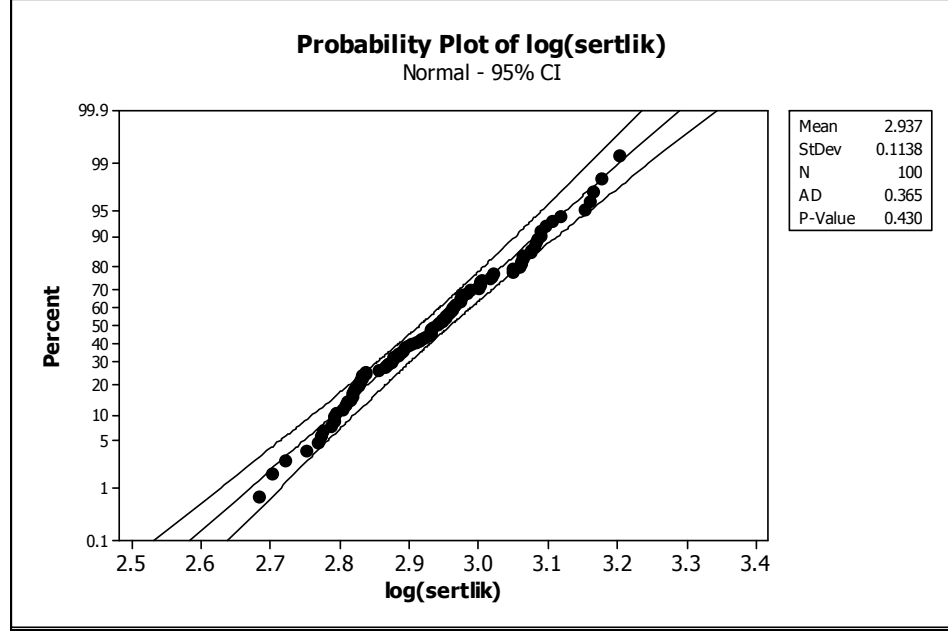
Süreç yetenek analizlerinde, önceki bölümlerde de bahsedildiği gibi, verilerin dağılışının normal olması indekslerin geçerli sonuçlar vermesi bakımından gereklidir. Hoaglin vd. tarafından önerilen simetri için veri dönüşümü tekniği, verilerin dağılışının simetrik olmadığı durumlarda simetriyi

Cenk ÖZLER

iyileştirmektedir. Simetrisinin iyileşmesi ise, verilerin dağılımını normale de yaklaştırmaktadır. Normal olasılık grafikleri yardımıyla bu durum görülebilir.



Şekil 4: Sertlik verileri için Normal Olasılık Grafiği



Şekil 5: log(Sertlik) verileri için Normal Olasılık Grafiği

Veri dönüşümü, simetriyi iyileştirmekle beraber, bir dezavantajı da beraberinde getirmektedir. Kalite karakteristiğine ait spesifikasyonlar için hedef değer T^n 'nin söz konusu olduğu ve süreç yetenek çalışmasında dikkate alınması istendiği durumlarda, \hat{C}_p , \hat{C}_{pk} ve \hat{C}_{pm} indekslerinin kullanılması mümkün olmayacaktır. Bunun nedeni, bu indekslerin, hedef değer T^n 'nin, ASL ve $ÜSL$ 'nin tam ortasında olduğu durumlarda kullanılabiliyor olmasıdır. Verilerin üzerine kuvvet dönüşümü uygulanmasıyla yeni toleranslar yeni hedef değer etrafında asimetrik duruma düşecektir. Bu durumda bu indeksler yerine asimetrik toleranslar için önerilen indekslerin kullanılması daha uygundur. Asimetrik toleranslar için kullanılan indeksler için Kotz ve Johnson (1993)'e başvurulabilir.

KAYNAKÇA

- Huber, P. (1981), *Robust Statistics*, Wiley, Chichester.
- Hoaglin, D. C., Mosteller, F., ve Tukey, J. W. (1983), *Understanding Robust and Exploratory Data Analysis*, Wiley, New York.
- Kotz, S. and Johnson, N. L. (1993), *Process Capability Indices*, Chapman & Hall, London.
- Kotz, S. ve Lovelace, C. R. (1998), *Process Capability Indices in Theory and Practice*, Arnold, London.

Cenk ÖZLER

- Montgomery, D. (2001), *Introduction to Statistical Quality Control*, 4. Baskı, Wiley, New York.
- Tukey, J.W. (1977), *Exploratory Data Analysis*, Addison-Wesley, Massachusetts.