Öneri, C.2, S.12.
Haziran 1999, ss.119-126.

ÖMÜR SÜRESİ VERİLERİNİN ANALİZİNDE KULLANILAN REGRESYON MODELLERİ

Dr. Fatma ACAR

Uludağ Üniversitesi İ.İ.B.F. Ekonometri Bölümü, Yardımcı Doçent

I. GİRİŞ

Bilindiği gibi regresyon analizi değişkenler arasındaki sebep-sonuç ilişkisini ortaya koymak amacı ile yapılmaktadır. Regresyon analizinde parametrelerin tahmin edilmesi amacıyla en çok kullanılan ve en bilinen teknik olağan en küçük kareler tekniğidir. Ancak bu tekniğin kullanımı birçok varsayımı gerekli kılmaktadır[1]. Bu varsayımların en önemlilerinden biri, hata terimlerinin normal dağılıma sahip olmasıdır. Ancak söz konusu varsayımın gerçekleşmesi için öncelikle hem bağımlı hem de bağımsız değişkenlerin sürekli olmaları gerekmektedir. Uygulamada ise zaman zaman bağımlı, zaman zaman da bağımsız değişkenlerin nitel yapıda oldukları görülmektedir.

Bağımsız değişkenlerin nitel yapıda olması durumunda, kukla değişken kavramı ortaya çıkmakta ve bu değişkenlerin özelliklerine bağlı olarak (olağan en küçük kareler tekniğini kullanmak mümkün olmadığından) farklı yaklaşımlar benimsenmektedir[2].

Bağımlı değişkenlerin sürekli olmaması durumunda ise karşımıza sınırlı bağımlı değişkenli modeller çıkmaktadır ki; bunlar logit, probit ve poisson türü modellerdir[3]. Sınırlı bağımlı değişkenli modellerde, açıklayıcı değişkenlerdeki değişmenin bağımlı değişkeni etkilediği olasılık düzeyi ortaya konmaktadır.

Oysa probit modellerin bir uzantısı durumundaki tobit modellerde incelenen örneklemde bağımlı değişkenin bazı değerleri ile ilgili bilgi bulunmazken, yalnızca bağımsız değişkene ilişkin bilgiler elde edilebilmektedir[4]. Bu tür örneklemlere **"cencored örneklem"** adı verilirken, hem bağımlı hem de bağımsız değişkene ilişkin bilgi kaybı söz konusu olduğunda ise **"truncated örneklem"** durumu ile karşılaşılmaktadır[5].

Yatay kesit verilerinde yukarıda sözü edilen bilgi kayıpları söz konusu olduğunda tobit modeller kullanılırken, örneklem **"ömür süresi"** verilerine ilişkin ise cencored ya da truncated deyimleri farklı bir anlam kazanmakta ve söz konusu verilerin analizinde kullanılan kullanılan regresyon modelleri de değişmektedir.

II. ÖMÜR SÜRESİ VERİLERİNİN YAPISI

Bir birimin işlevsel hale geldiği sürenin başlangıcından ilk bozulma zamanına kadar geçen süre biçiminde tanımlanabilecek olan ömür süresine ilişkin verilerin yapısı diğer verilerden büyük ölçüde farklılık göstermektedir. Ömür süresi verilerine ilişkin örneklemde; bazı birimlerin ömür süresi bitmiş olduğu halde, bazıları hala ömrünü tamamlamamış durumda bulunabilir. İncelenen dönem itibarıyla henüz tamamlanmamış veri görünümünde bulunan bu tür veriler **"cencored"** olarak nitelendirilmektedir. Bu tür verilere ilişkin regresyon analizinde amaç; verilerin ömür süresi ile bu süreyi etkileyen faktörler arasında bir sebep-sonuç ilişkisi kurmaktır. Ömür süresi verilerinin elde edilmesi sürecindeki farklılıklar nedeni ile, dört farklı veri tipinden sözetmek mümkündür.

I. Tip Veri : Ömür süresine ilişkin olarak yapılan testlerde zaman deterministik, tamamlanan birim sayısı (s) stokastik faktör olarak alınmışsa I. Tip durdurulmuş veri söz konusu olur[6]. Bu tür testlerde birimler t=0 döneminde gözleme alınmakta ve test t_0 döneminde sona ermektedir. $t_0 - 0$ dönemi belirlidir. Bu dönem içinde analize alınan n birimden yalnızca r tanesi konusunda bilgi sahibi olunurken $n-r$ birimin ömür süresi henüz sona ermediğinden onlarla ilgili tam bilgi yoktur.

II. Tip Veri : Tamamlanan birim sayısının (r) deterministik, buna karşılık test süresinin stokastik olduğu durumlarda II. tip durdurulmuş veri söz konusu olmaktadır[7]. Bu tür testlerde test t=0 döneminde başlatılır ve n birimden r tanesi tamamlandığında sona erer. Dolayısıyla t_0 belli değildir.

III. Tip Veri : Bu tür veri tipi ilk iki türün bileşiminden oluşmaktadır. Hem test dönemi hem de tamamlanmış birim sayısı başlangıçta belirlenmektedir. Test tamamlanma dönemi ya da tamamlanma sayısından birinin oluştuğu dönemde sona ermektedir.

IV. Tip Veri : Bazı durumlarda teste başlama süresi tüm birimler için aynı olmayabilir. Bu durumdaki veriler için farklı başlangıç zamanları bulunmaktadır.

Ancak, testin sona erdirilme süresi t_0 belirlidir. Bu süre içinde tamamlanan birim sayısı ise stokastiktir[8].

III. ÖMÜR SÜRESİ VERİLERİNİN FONKSİYONLARI VE DAĞILIMLARI

III.1 Ömür Süresi Verilerinin Fonksiyonları

Ömür süresi verilerine ilişkin olarak üç temel fonksiyondan hareket etmekte fayda vardır. Bunlar; Hazard fonksiyonu, Survival fonksiyon ve Olasılık yoğunluk fonksiyonudur. Bu fonksiyonlar ömür süresini açıklamada kullanılan birbirleri ile ilişkili fonksiyonlardır. $t = 0$ döneminde işlevsel hale gelen birimin bozulma zamanını gösteren T bu fonksiyonların açıklanmasında önemli bir yere sahiptir. Birimin bozulma zamanını etkileyen çok değişik faktörler bulunduğundan T'nin belirlenmesi mümkün olmamaktadır. Ancak, birimin belli bir dönemde bozulma olasılığı ortaya konulabilir. Bu doğrultuda bozulma zamanının birikimli dağılım fonksiyonu (F(t)) ömür süresinin t'ye eşit veya altında olma olasılığını göstermektedir.

$$F(t) = P(T \leq t) = \int_0^t f(x)dx \qquad 0 \leq t < \infty$$

biçiminde tanımlanan bu fonksiyonda F(t) birimin $[0,t]$ aralığında bozulma olasılığını vermektedir. Söz konusu birikimli dağılım fonksiyonundan hareketle bir birimin ömür (survival) fonksiyonu S(t) ise;

$$S(t) = 1 - F(t) = P(S_t) \qquad t > 0 \text{ için}$$

biçiminde tanımlanabilir.

T döneminde işlevsel olan bir birimin anlık bozulma oranı ise Hazard fonksiyonu (h(t)) yardımıyla elde edilebilir ve ;

$$h(t) = \frac{F'(t)}{1 - F(t)}$$

biçiminde gösterilebilir. Söz konusu fonksiyonlar arasında aşağıdaki ilişki ortaya konulabilir[8].

Tablo 1 : Ömür Süresi Verilerinin Fonksiyonları Arasındaki İlişkiler

	$F(t)$	$R(t)$	$h(t)$
$F(t)$	-	$1-R(t)$	$1-e^{-\int_0^t h(x)dx}$
$R(t)$	$1-F(t)$	-	$h(x).e^{-\int_0^t h(x)dx}$
$h(t)$	$\dfrac{F'(t)}{1-F(t)}$	$-\dfrac{R'(t)}{1-R(t)}$	-

III.2 Ömür Süresi Verilerinin Dağılımları

Aynı özelliğe sahip verilerin istatistiksel analizini kolaylaştırmak ve hızlandırmak amacıyla önceden geliştirilmiş teorik modeller[9] biçiminde tanımlanabilecek olan dağılımlar, rassal değişkenin özelliğine bağlı olarak kesikli ve sürekli olmak üzere iki başlık altında incelenmektedir. Dağılımların önemi; verilerin yapısına uygun dağılımın belirlenmesi halinde, söz konusu dağılımın özelliklerini kullanarak analizi hızlandırmak ve kolaylaştırmalarından kaynaklanmaktadır. Ömür süresi verileri belli bir zaman dönemine ilişkin oldukları için dağılımlarının da sürekli olması kaçınılmazdır. Bu konuda yapılan çalışmalardan üstel ve gompertz[10], quadratik gompertz ve weibull türü dağılımların yaygınlıkla kullanılan ömür dağılımları olduğu görülmüştür. Verilerin bu dağılımlardan herhangi birine uygun olduğunun saptanması halinde regresyon modeli parametrik hale dönüşebilmektedir. Aksi halde, daha esnek yapıdaki diğer modellerden yararlanmak mümkün olmaktadır.

IV. ÖMÜR SÜRESİ VERİLERİNE İLİŞKİN REGRESYON MODELLERİ

Ömür süresi verilerine ilişkin analiz yapılırken regresyon modelinin seçiminde ömür verilerinin yapısı önem kazanmaktadır. Söz konusu verilerin dağılımı konusunda bilgi varsa parametrik regresyon modellerinden uygun olanını, aksi halde ise parametrik olmayan regresyon modellerinden birinin seçilmesi gerekecektir. Bu doğrultuda söz konusu regresyon modellerini parametrik olmayan regresyon ve parametrik regresyon modelleri olmak üzere iki başlık altında toplamakta fayda vardır.

IV.1 Parametrik Olmayan Regresyon Modelleri

IV.1.1 Cox'un Oransal Hazard Regresyonu

Oransal Hazard modeli Cox (1972) tarafından ortaya atılmış ömür verilerinin yapısı hakkında herhangi bir varsayım gerektirmeyen son derece esnek bir regresyon modelidir. Modelin temelinde Hazard oranı veya ömür (survival) zamanının bağımlı değişken olarak alınması ve bazı faktörlere bağlı olarak değiştiği mantığı bulunmaktadır. Orantılı Hazard regresyonu

$$h(t,z) = h_0(t)\exp(b_i z_i)$$

biçimindedir. Burada;

$h(t,z)$: Hazard oranı

$h_0(t)$: Temel hazard fonksiyonu

b_i : Katsayılar vektörü

z_i : Bağımsız değişkenler vektörü

olarak yer almaktadır[11].

Cox'un orantılı hazard modeli hazard fonksiyonu ile ilgili herhangi bir varsayım gerektirmediği için oldukça esnek olmakla birlikte, kendine özgü iki varsayımı bulunmaktadır. Bunlardan biri değişkenlerin log-doğrusal fonksiyonu ile hazard fonksiyonu arasındaki ilişkinin çarpımsal olmasıdır. Diğeri ise, açıklayıcı değişkenler ile hazard fonksiyonu arasındaki ilişkinin log-doğrusal olmasıdır. Bu varsayımlardan ilki oransallık varsayımı adını alır ve hazard oranının zamandan etkilenmediğini anlatır. Ancak uygulamada bu varsayımı gerçekleştirmek her zaman mümkün değildir. Bu nedenle Cox'un zamana bağımlı oransal hazard modeli bu sakıncayı ortadan kaldırmaktadır.

Cox'un oransal hazard modelinde parametre tahmini amacıyla;

$$L = \Pi \exp(b;s_i) / \sum \exp(bz_j)^{d(i)}$$

kısmi olabilirliğini maksimize eden Newton-Raphson yöntemi kullanılmaktadır[12]. Formülde

d_i : t_i zamanındaki kayıp gözlem sayısını

s_i : değişkenler toplamını

z_j : t_i zamanında risk setindeki j durumu için bağımsız değişken vektörünü gösterir.

Parametrelerin test edilmesi amacı ile

$$W = \beta - 1 / Var(\beta).\beta$$

formülü kullanılarak Wald istatistiği hesaplanır. β : tahmin edilen parametre $Var(\beta)$: parametrelerin varyansıdır[13]. Modelin uyum iyiliği ise

$$\chi^2 = 2(L_0 - L_1)$$

formülü ile elde edilen test istatistiğinden hareketle test edilir. Burada

L_1 : Bütün açıklayıcı değişkenler için log-olabilirlik fonksiyonu

L_0 :Bütün açıklayıcı değişkenleri sıfıra yaklaştıran olabilirlik fonksiyonudur.

Elde edilen χ^2 değeri ile anlamlı sonuç elde edilmesi için $\chi^2 > \chi^2_k$ olmak zorundadır. Aksi halde ele alınan değişkenler hazard fonksiyonunu açıklamak konusunda yeterli olmayacaktır.

IV.1.2 Cox'un Zamana Bağımlı Oransal Hazard Regresyonu

Yukarıda da değinildiği gibi açıklayıcı değişkenler için sözü edilen oransallık varsayımının her zaman geçerli olmadığı ve bazı açıklayıcı değişkenlerin değerlerinin zamandan etkilendiği görülmektedir. Örneğin, yaş, eğitim durumu, gelir, vb. gibi faktörlerin zamandan etkilendiği görülmektedir. Bu durumda açıklayıcı değişkenler zamanın bir fonksiyonu olarak karşımıza çıkmaktadır. Zamanın etkisini katmak amacıyla

$$h(t,z) : h_0(t).\exp\{b_1 z + b_2 [z \log(t) - 5{,}4]\}$$

biçiminde bir model kullanma gereği ortaya çıkmaktadır[14].

Yukarıdaki koşullu zamana bağımlı hazard modelinde hazard oranı; temel hazardın $h_0(t)$, değişkenlerin (z_i) ve zamanın (t) logaritmanın z.katının fonksiyonudur. Formüldeki 5,4 sabiti ölçekleme sabitidir. Bir başka deyişle, yukarıdaki model zaman ve bağımsız değişkenlerin bir fonksiyonu olarak karşımıza çıkmaktadır.

Cox'un zamana bağımlı oransal hazard modeli için parametre tahminleri, parametrelerin anlamlılık testleri ve modelin uyum iyiliği için yukarıdaki yaklaşımlar burada da geçerlidir. Ancak kısmi olabilirlik fonksiyonu zamanı da içerecek biçimde[15]

$$L = \Pi \exp(b' ; s_i(t)) / \sum \exp(b' z_j(t))^{d(i)}$$

şeklinde oluşturulmalıdır.

$s_i(t_i)$: t_i zamanında kaybolan zamana bağımlı bağımsız değişkenlerin toplamı

$z_j(t_i)$: t_i zamanında risk setindeki j durumu için zamana bağımlı bağımsız değişken vektörü.

d_i : t_i zamanında kayıp durum sayısı.

Yukarıdaki kısmi olabilirlik fonksiyonu Newton-Rapson yöntemi kullanılarak minimize edilmektedir.

IV.2 Parametrik Regresyon Modelleri

Yukarıda sözü edilen regresyon modelleri, ömür süresinin dağılımına ilişkin herhangi bir varsayım gerektirmemektedirler. Oysa ömür süresinin sürekli dağılımlardan herhangi birine uyduğu belirlendiğinde ona uygun bir regresyon modeli oluşturmak mümkün olabilir. Bu doğrultuda bu bölümde üstel, log-normal ve normal regresyon modelleri üzerinde durulacaktır.

IV.2.1 Üstel Regresyon

Yukarıda da değinildiği gibi ömür verileri daha çok üstel ve gompertz dağılıma uymaktadırlar. Bu nedenle de bu verilere uygun bir üstel regresyon modeli oluşturmak son derece doğaldır. Bu konuda yapılmış ilginç çalışmalar bulunmaktadır. Üstel regresyon modeli:

$$S(z) = \exp(a + b_i z_i)$$

biçimindedir. Burada;

$S(z)$: Ömür süresi

b_i : Parametre vektörü

z_i : Bağımsız değişkenler vektörüdür.

Modelin parametrelerini tahmin etmek amacıyla :

$$L(z) = \exp(a + b_i . z_i)$$

eşitliğinden maksimum olabilirlik tahminleri için Newton-Raphson yöntemi kullanılmaktadır[16].

Modelin uyum iyiliği ;

$$\chi^2 = 2(L_0 - L_1)$$

ile test edilmektedir.

Modelin üstellik varsayımını kontrol etmek için diğer bir yol ise standart üstel sıra istatistiklerin (α) hesaplanıp, bunlara ilişkin grafiğin çizilmesidir. Eğer tüm noktalar doğru bir çizgi üzerinde toplanmışsa, modelin üstel bir model olduğuna karar verilir.

IV.2.2 Normal ve Log-Normal Regresyon

Bu modelde ömür süresinin ya da logaritmalarının normal dağılımdan geldiği varsayılmakta ve model

$$f_i = a + z_i\, b_i$$

biçiminde tanımlanmaktadır. Burada;

f_i : Ömür süresi

b_i : Parametre vektörü

z_i : Bağımsız değişkenler vektörüdür.

Eğer Log-normal regresyon modeli oluşturulmak istenirse f_i'nin yerine logaritmaları konulmaktadır. Yukarıdaki model için maksimum olabilirlik tahminlerini elde etmek amacıyla **"beklenen maksimizasyon"** yöntemi kullanılmaktadır[17]. Modelin uyum iyiliği yine χ^2 istatistiği ile test edilmektedir.

V. UYGULAMA

100 kişi üzerinde yapılan bir araştırmada evlilik sürelerini etkileyen faktörler olarak yaş, kişisel gelir ve çocuk sayısı alınmıştır. Evlilik süreleri ise bağımlı değişken olarak modele katılmıştır. Araştırma kapsamında incelenen kişilerin evliliklerinin başlangıç tarihi farklı ancak incelemenin tarihi 30.05.1999 olduğundan veriler IV. Tip veri görünümündedir. Sözü edilen üç faktörün evliliklerin ömür süresine etkisini ortaya koymak amacıyla elde edilen veriler, yukarıda kuramsal çerçevesi açıklanmaya çalışılan regresyon modellerinin tümüne uygulanarak sonuçlar karşılaştırılmış ve verilerin yapısına en uygun model bulunmaya çalışılmıştır.

V.1.Tanımlayıcı İstatistikler

Tablo 2' de de görülebileceği gibi incelenen örneklemdeki kişilerin gelir ortalaması 178750000 TL olurken, standart sapma 151833000 TL dir. Minimum gelir 0.000 TL, maksimum gelir ise 650000000 TL dir. Örneklemdeki minimum çoçuk sayısı 0,maksimum çocuk sayısı 7 ve bunlara ilişkin ortalama 2, standart sapma ise 1 çocuktur.Yaş değişkeninin ortalaması 43 ve standart sapma 10 olurken, minimum yaş25 ve maksimum yaş 69 olmaktadır. Evli kalınan ortalama sürenin 4487 gün ve standart sapmasının 4480 olduğu görülmektedir. Evlilikler için minimum süre 143 gün,maksimum süre ise 18259 gündür.

Tablo 2 : Tanımlayıcı İstatistikler

Değişkenler	Ortalama	Standart sapma	Minimum	Maksimum
Gelir	178750000	151833000	0,00000	650000000
Çocuk	2	1	0,00000	7
Yaş	43	10	25	69
Gün Sayısı	4487	4480	143	18259

V.2.Parametrik Olmayan Regresyon Modelleri

Tablo 3 ve 4' deki sonuçlara bakıldığında her iki modelinde verilere uygun oldukları görülmektedir. Çünkü modelin genel anlamlılığının test edildiği χ^2 sonuçları sırasıyla 99,8169 ve 86,3880 dir. Bu sonuçlara ilişkin gözlenen anlamlılık düzeyleri (p) ise her iki model için de 0.0000 dır. Bu nedenle % 5 anlamlılık düzeyinde gelir,çocuk sayısı ve yaş değişkenlerinin evliliklerin ömür sürelerini açıklamak amacıyla kullanılabileceğini söylemek mümkündür.

Daha önce de belirtildiği gibi modelin genel anlamlılığı χ^2 istatistiği ile test edilirken parametrelerin anlamlılığını test etmek amacıyla Wald istatistiği kullanılmaktadır. Bu doğrultuda parametrelere ilişkin test sonuçlarına bakıldığında tablo 3 ve 4 de de görülebileceği gibi % 5 anlamlılık düzeyinde yalnızca çocuk sayısı değişkeni parametresinin (p=0,0000) anlamlı bir şekilde sıfırdan farklı olduğu görülmektedir.

Tablo 3: Cox'un Oransal Hazard Regresyon Modeli Sonuçları

χ^2 = 99,8169		s.d.= 3		İterasyon = 17		p=0,0000
Değişken	Parametre	Standart hata	t- değeri	Üstel parametre	Wald istatistiği	p
Gelir	0,00000	0,00000	1,92129	1,00000	3,69137	0,054703
Çocuk	-1,46545	,210579	-6,95916	0,230973	48,42995	0,00000
Yaş	0,01172	0,017170	0,68241	1,011786	0,46568	0,494986

Tablo 4:Cox'un Zamana Bağımlı Oransal Hazard regresyon Modeli Sonuçları

χ^2 = 86,3880		s.d.= 3		İterasyon = 17		p=0,0000
Değişken	Parametre	Standart hata	t- değeri	Üstel parametre	Wald istatistiği	p
Gelir	0,00000	0,00000	0,68656	1,00000	0,47136	0,494367
Çocuk	-0,684681	0,103705	-6,60222	0,504251	43,58925	0,00000
Yaş	0,010495	0,009149	0,1,14718	',010551	1,31602	0,494986

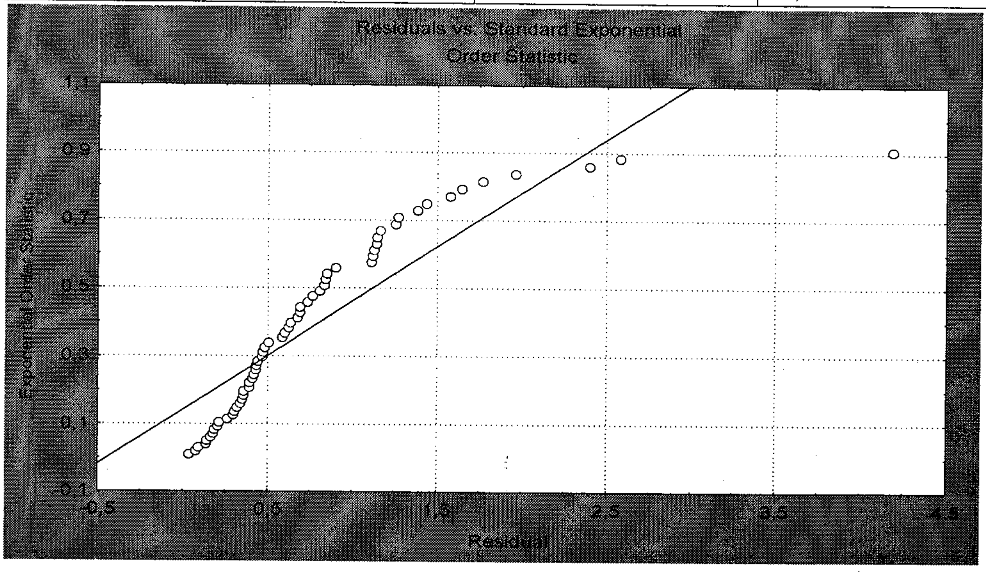
V.3.Parametrik Regresyon Modelleri

Tablo 5, 6 ve 7 ' de sonuçları görülen üstel, log-normal ve normal regresyon modellerinden elde edilen χ^2 istatistikleri (123,749-113,750 ve 138.711) ve gözlenen anlamlılık düzeylerine (p= 0,0000) bakıldığında tüm modellerin verilere uygun olduğu görülmektedir. Ancak verilerin grafikleri incelendiğinde üstel modelin grafiğinde artıkların doğru üzerinde toplandığını söyleyebilmek mümkün değildir. log-normal ve normal regresyon modellerinin normal olasılık grafikleri incelendiğinde ise özellikle log-normal regresyon modelinin grafiğinde artıkların normal dağılıma uygun olduğu görülmektedir. Bu doğrultuda, parametrik regresyon modelleri arasında verilere en uygun modelin log-normal model olduğu söylenebilir.

Parametrelerin anlamlılığına bakıldığında ise üç modelde de % 5 anlamlılık düzeyinde gelir ve yaş değişkenlerine ilişkin parametrelerin anlamlı bir şekilde sıfırdan farklı olduklarını söyleyebilmek mümkün değildir. Burada da yalnızca çocuk sayısına ilişkin parametre, anlamlı bir şekilde sıfırdan farklıdır.

Tablo 5: Üstel Regresyon Modeli Sonuçları

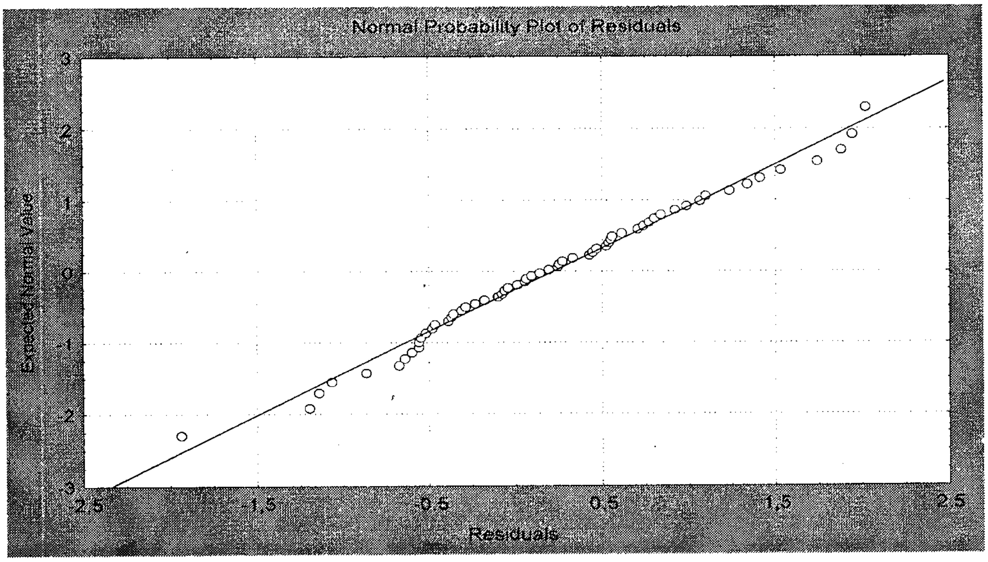
χ^2 = 123,749	s.d.= 3	İterasyon = 7	p=0,0000
Değişken	Parametre	Standart hata	t- değeri
Gelir	0,00000	0,00000	-2,09042
Çocuk	1,223651	0,154414	7,924475
Yaş	-0,017667	0,015652	-1,12884
Sabit	8,1000020	0,801726	10,10323



Grafik 1: Üstel Sıra İstatistiklerine Karşı Artıkların Grafiği

Tablo 6: Log-Normal Regresyon Modeli Sonuçları

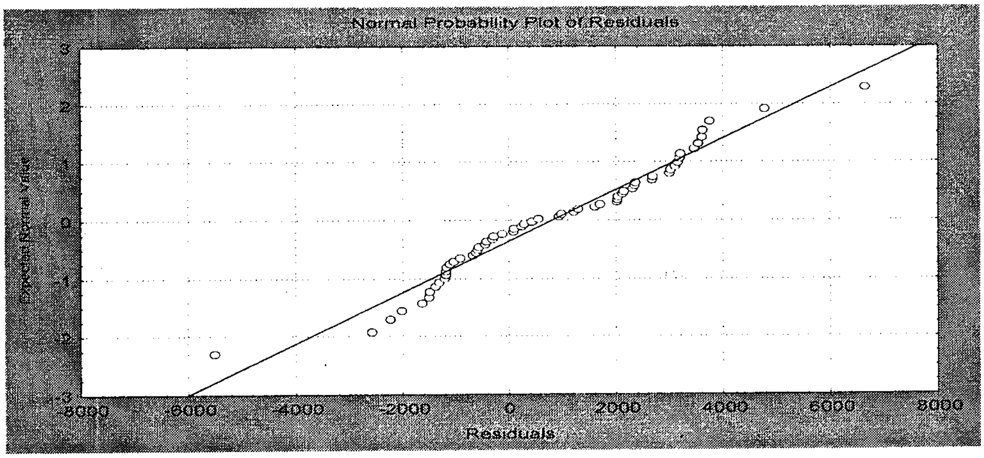
$\chi^2 = 113,750$	s.d.= 3	İterasyon = 38	P=0,0000
Değişken	Parametre	Standart hata	t- değeri
Gelir	0,00000	0,00000	-2,41459
Çocuk	0,986624	0,109805	8,98520
Yaş	-0,014277	0,011068	-1,28995
Sabit	7,81958	0,595923	13,12180



Grafik 2: Log-Normal Regresyonda Artıkların Normal Olasılık Grafiği

Tablo 7: Normal Regresyon Modeli Sonuçları

$\chi^2 = 138,411$	s.d.= 3	İterasyon = 54	p=0,0000
Değişken	Parametre	Standart hata	t- değeri
Gelir	0,00000	0,00000	-1,50404
Çocuk	3780,383	328,744	11,49947
Yaş	11,875	34,109	0,34814
Sabit	487,804	1835,448	0,26567



Grafik 3: Normal Regresyonda Artıkların Normal Olasılık Grafiği

VI. SONUÇ

Bu çalışmada ömür süresi verilerinin analizine ilişkin regresyon modellerinin kuramsal çerçevesi üzerinde durulmuş ve daha sonra modeller evliliklerin ömür süresi verilerine uygulanmıştır. Elde edilen sonuçlar; tüm modellerde gelir, çocuk sayısı ve yaş değişkenlerinin evlilik süresini açıklamada kullanılabileceği görülmüştür. Çünkü, modelin genel anlamlılığını test eden χ^2 test istatistiklerine bakıldığında tüm modeller için sonuçlar anlamlı görünmektedir. Bu arada parametrik regresyon modelleri içinde en uygun modelin log-normal model olduğu söylenebilir.

Parametrelerin anlamlılık testi sonuçlarına bakıldığında ise, tüm modellerde gelir ve yaş değişkenleri anlamlı bulunmazken yalnızca çocuk sayısı değişkenine ilişkin parametre anlamlı bir şekilde sıfırdan farklı bulunmuştur.

KAYNAKLAR

[1]-SERPER.Ö. Uygulamalı İstatistik II. Genişletilmiş 3. Baskı. Filiz Kitabevi. İstanbul. 1996.

[2]-İŞYAR.Y. Ekonometrik Modeller. Uludağ Üniversitesi Basımevi.1994.

[3]-SEVÜKTEKİN.M. "Niteliksel Tercih Modelleri: Probit. Logit ve Alternatif Bir Çözüm Olarak Poisson Modeli Tahminlerinin Bir Karşılaştırması " II. Ulusal Ekonometri ve İstatistik Sempozyumu Bildirileri. Kasım 1995.

[4]-GUJARATİ.D.N. Basıc Econometrics. Third Edition. Literatür Yayıncılık. 1995

[5]-NELSON.F.D. " A Test For Misspecification in the Censored Normal Model" Econometrica, Vol.49. N.5,sept.1981.

[6]-SIRVANCI.M., YANG.G." Estimation of the Weibull Parameters Under Type I Censoring" JASA,V.79,N.385,March 1984.

[7]-ŞAHİNOĞLU.M, VURAL.C, "Censored Data Analysis in Survival for Studies Softward Quality Testing " I. Ulusal Ekonometri ve İstatistik Sempozyumu Bildirileri. Ekim 1994.

[8]-ŞENTÜRK.A. Ömür Verileri Analizi ve Bir Uygulama. Basılmamış Doktora Tezi. Bursa.1998.

[9]-AYTAÇ,M. Matematiksel İstatistik.Genişletilmiş 2. Baskı. Ezgi Kitabevi Yayınları. Bursa.1999.

[10]-EFRON.B. " Logistic Regression. Survival Analysis and the Kaplan- Meier Curve. JASA, Vol.83. N.402. June 1988.

[11]-EFRON.B." The Efficiency. of Cox's Likeliood Function for Censored Data " JASA. Vol.72. N. 358.Sept 1977.

[12]-BRESLOW.N.E. " Covariance Analysis of Censored Survival Data". Biometrics. 30.1974.

[13]-LİN.D.Y. WEİ.L.J." The Robust İnference for the Cox Proportional Hazard Models " JASA. Vol.84. N.408. 1989.

[14]-STATISTICA. Help İndex.

[15]-LAWLES. J.F. Statistical Models and Methods for Lifetime Data. John Wiley& Sons Inc.. New York. 1982.

[16]-LAGAKOS.S.W.KUHNS.M.H."Maximum Likelihood Estimation for Censored Exponentional Survival Data with Covariates Applied Statistics ".27.1978.

[17]-COX.D.R, OAKOS.D. Analysis of Survival data.Chapman and Hall Ldt.. London.1984.