ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

# SANSÜRLÜ GÖZLEMLERDE YENİLEME FONKSİYONUNUN TAHMİNİ

Çiğdem DANIŞ

İSTATİSTİK ANABİLİM DALI

ANKARA
2008

# ÖZET

Yüksek Lisans Tezi

SANSÜRLÜ GÖZLEMLERDE YENİLEME FONKSİYONUNUN TAHMİNİ

Çiğdem DANIŞ

Ankara Üniversitesi
Fen Fakültesi
İstatistik Bölümü

Danışman: Yrd. Doç. Dr. Halil AYDOĞDU

Stokastik modellemede oldukça sık kullanılan yenileme süreçlerinde tahmin problemi ortaya çıkmaktadır. Yenileme fonksiyonunun tahmini günümüzde hala incelenmektedir. Bu çalışmada sağdan rasgele sansürlenmiş örneklem durumunda yenileme fonksiyonu için bazı parametrik ve parametrik olmayan tahmin ediciler ele alınır ve bunların bazı istatistiksel özellikleri incelenir. Ayrıca bu tahmin edicilerin değerlerinin nasıl hesaplanacağı üzerinde durulur.

**2008, 83 sayfa**

**Anahtar Kelimeler:** Yenileme süreci, yenileme fonksiyonu, sağdan rasgele sansürleme

## ABSTRACT

Master Thesis

ESTIMATION OF THE RENEWAL FUNCTION FOR CENSORED OBSERVATIONS

Çiğdem DANIŞ

Ankara University
Graduate School of Natural and Applied Sciences
Department Statistics

Supervisor: Asst. Prof. Dr. Halil AYDOĞDU

Renewal processes commonly used in stochastic modelling and estimation problem is encountered in these processes. The estimation of renewal function is still examined. In this study, some parametric and nonparametric estimators of the renewal function for random right censored samples are considered and their statistical properties are investigated. Moreover, it is also examined how the values of these estimators are computed.

**2008, 83 sayfa**

**Key Words**: Renewal process, renewal function, right censoring, Kaplan-Meier estimator

## TEŞEKKÜR

Bana araştırma olanağı sağlayan ve çalışmamın her safhasında benden yardımlarını esirgemeyen, önerileri ile beni yönlendiren danışman hocam Sayın Yrd. Doç. Dr. Halil AYDOĞDU'ya ve çalışmam boyunca hoşgörü ve desteklerini esirgemeyen eşime, aileme teşekkürlerimi sunarım.

Çiğdem DANIŞ
Ankara, Ocak 2008

## İÇİNDEKİLER

## SİMGELER DİZİNİ

| | |
|---|---|
| $*$ | Stieltjes konvolüsyon Operatörü |
| $\mu$ | $F$ dağılımının Ortalaması |
| $\sigma^2$ | $F$ dağılımının Varyansı |
| $I$ | İndikatör fonksiyonu |
| $\hat{\alpha}$ | $\alpha$ 'nın tahmin edicisi |
| $\hat{\beta}$ | $\beta$ 'nın tahmin edicisi |
| $\hat{\theta}$ | $\theta$ 'nın tahmin edicisi |
| $\mathbb{R}$ | Reel Sayılar |
| $\mathbb{N}$ | Doğal Sayılar |
| $F^{k*}(t)$ | $F$ dağılımının kendisiyle olan $k$ kez konvolüsyonu |
| $[\![...]\!]$ | Tam değer fonksiyonu |
| $\approx$ | Yaklaşık olarak eşit |
| $M(t)$ | Yenileme Fonksiyonu |
| $\hat{M}_n(t)$ | $M(t)$ 'nin tahmin edicisi |
| $\hat{M}_{KM}(t)$ | $M(t)$ 'nin parametrik olmayan Kaplan-Meier tahmin edicisi |
| $M_a(t)$ | $M(t)$ 'nin asimptotik tahmin edicisi |

## ÇİZELGELER DİZİNİ

# 1. GİRİŞ

Bir yenileme süreci zamanın bir fonksiyonu olarak gerçekleşen olayların (yenilemelerin) sayısını sayan bir sayma sürecidir, burada ardışık yenilemeler arası geçen zaman süreleri birbirlerinden bağımsız, aynı $F$ dağılımlı rasgele değişkenlerdir. Yenileme süreçleri güvenilirlik analizi, trafik akışı, envanter, risk teorisi, garanti analizi ve uygulamalı istatistiğin bir çok sahasında stokastik modellemede kullanılan güçlü araçlardır.

Yenileme süreçleri ile ilgili uygulamalarda genellikle sürecin ortalama değer fonksiyonu (yenileme fonksiyonu) bilgisine ihtiyaç duyulur. Yenileme fonksiyonu için $F$ dağılım fonksiyonuna bağlı açık ifadeler vardır ve bu fonksiyon prensipte bu ifadelerin birinden hesaplanabilir. Fakat, $F$ dağılım fonksiyonu tam olarak bilinse bile birkaç dağılım dışında yenileme fonksiyonu analitik olarak elde edilemez.

Uygulamada genellikle $F$ dağılım fonksiyonu ya bilinmiyordur, ya da şekilsel olarak bilinirken dağılımın parametreleri bilinmiyordur. Bu durumda yenileme fonksiyonu verilerden tahmin edilmek zorundadır. Veriler $F$ dağılımlı kitleden tam örnekleme ile gelebileceği gibi sansürlü olarak da gelebilir. Örneğin, garanti analizi için gözlemlenen veriler genelde sağdan birinci tip sansürlenmişlerdir. Yenileme teorisinde ise çoğu kez sağdan rasgele sansürlenmiş veriler ile karşılaşılır. Sağdan birinci tip sansürleme sağdan rasgele sansürlemenin özel bir halidir. Gözlemlerin sansürlü olması durumunda yenileme fonksiyonunun parametrik ve parametrik olmayan tahmini üzerinde çalışılması gereken önemli bir problemdir.

Tam örneklem durumunda yenileme fonksiyonu değerinin tahmini literatürde değişik yazarlar tarafından incelenmiş ve hala incelenmeye devam edilmektedir. Bu çalışmada sağdan rasgele sansürlenmiş örneklem durumunda yenileme fonksiyonu için parametrik ve parametrik olmayan tahmin ediciler önerilmekte ve bu tahmin edicilerin bazı istatistiksel özellikleri araştırılmaktadır. Bu çalışma aşağıdaki biçimde düzenlenmiştir.

İkinci bölümde bazı temel kavramlar verilmiştir. Konvolüsyon kavramı, Laplace-Stieltjes ve Laplace dönüşümü, yenileme denklemi ve çözümü üzerinde durulmuştur.

Üçüncü bölümde ilk olarak yenileme süreçleri üzerinde durularak bazı asimptotik sonuçlar verilmiştir. Daha sonra yenileme sürecin ortalama değer fonksiyonu tanıtılarak bazı özellikleri üzerinde durulmuş ve yenileme fonksiyonu ile ilgili bazı limit teoremler verilmiştir.

Dördüncü bölümde karşılaşılabilecek bazı dağılımlar için yenileme fonksiyonu analitik olarak elde edilmiştir. Ayrıca, genelde bu fonksiyon için analitik ifadeler mevcut olmadığından bu fonksiyonun RS(Rieman-Stieltjes) yöntemi ile sayısal olarak elde edilmesi üzerinde durulmuştur.

Beşinci bölümde sansürlü veriler hakkında bilgi verilerek, yenileme süreçlerinde önemli bir yer tutan sağdan rasgele sansürleme çeşidi üzerinde durulmuştur. Daha sonra $F$ 'nin şekilsel olarak bilindiği fakat parametrelerinin bilinmediği durumda sağdan rasgele sansürlü örnekleme dayalı olarak bilinmeyen parametrelerin en çok olabilirlik tahmin edicileri üstel, Weibull, log-normal ve gamma dağılımları için bulunmuştur. Üstel dağılım hariç bu tahmin edicilerin bulunmasında kullanılan denklem sistemlerinin analitik çözümü olmadığından sayısal bir yöntem olan Newton-Raphson yöntemi kullanılmıştır. Ardından $F$ 'nin hiç bilinmediği durumunda sağdan rasgele sansürlü örnekleme dayalı olarak $F$ 'nin değerinin tahmini için ömür tablo yöntemi ve Kaplan-Meier tahmin edicisi verilmiştir.

Altıncı bölümde ise sağdan rasgele sansürlü verilerde yenileme fonksiyonunun değerinin tahmini için parametrik ve parametrik olmayan tahmin ediciler önerilmiş ve bu tahmin edicilerin nasıl hesaplanacağı problemi üzerinde durulmuştur. Ayrıca önerilen tahmin edicilerin bazı istatistiksel özellikleri incelenmiştir. Daha sonra bu tahmin edicilerin bir örnek üzerinde değerlendirilmesi yapılmıştır.

## 2. TEMEL KAVRAMLAR

Bu bölümde çalışmada gerekli olan bazı temel bilgiler verilmektedir. Konvolüsyon kavramı, Laplace-Stieltjes ve Laplace dönüşümü hatırlatıldıktan sonra yenileme denklemi ve çözümü üzerinde durulmaktadır.

### 2.1 Konvolüsyon İşlemi

F ve G iki dağılım fonksiyonu olsun.

$$F*G(t) = \int_{-\infty}^{\infty} G(t-x)dF(x), \quad t \in \mathbb{R}$$

ile tanımlanan $F*G$ fonksiyonuna $F$ ile $G$'nin konvolüsyonu denir. $F$ ve $G$ dağılım fonksiyonu iken $F*G$ de bir dağılım fonksiyonudur.

#### 2.1.1 Konvolüsyon işleminin özellikleri

Konvolüsyon işlemi değişme ve birleşme özelliğine sahip olduğu gibi aynı zamanda bir etkisiz elemana sahiptir. Bu özellikler aşağıda gösterilmektedir.

$F$ ve $G$ herhangi iki dağılım fonksiyonu olmak üzere,

$$\begin{aligned} F*G(t) &= \int_{-\infty}^{\infty} G(t-x)dF(x), \\ &= G(t-x)F(x)\Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} F(x)dG(t-x) \\ &= G(-\infty)F(\infty) - G(\infty)F(-\infty) - \int_{-\infty}^{+\infty} F(x)dG(t-x) \\ &= \int_{-\infty}^{+\infty} F(t-y)dG(y) \\ &= G*F(t) \end{aligned}$$

dir. Böylece * konvolüsyon işleminin değişme özelliğine sahip olduğu görülmüş olur.

Herhangi $F$ , $G$ ve $H$ dağılım fonksiyonları için birleşme özelliği olduğunu gösterelim.

$$
\begin{aligned}
F*(G*H)(t) &= \int_{-\infty}^{+\infty} (G*H)(t-x)dF(x) \\
&= \int_{-\infty}^{+\infty} (H*G)(t-x)dF(x) \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} G(t-x-y)dH(y)dF(x) \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} G(t-x-y)dF(x)dH(y) \\
&= \int_{-\infty}^{+\infty} (F*G)(t-y)dH(y) \\
&= H*(F*G)(t) \\
&= (F*G)*H(t)
\end{aligned}
$$

olduğundan

$$F*(G*H) = (F*G)*H$$

dir. Yani * konvolüsyon işlemi birleşme özelliğine sahiptir.

$$
F^{0*}(t) = \begin{cases} 0, & t < 0 \\ 1, & t \geq 0 \end{cases}
$$

olsun. Bu dağılım sıfır noktasında yoğunlaşmış dağılım ya da sıfır noktasındaki Dirac dağılımı olarak bilinir. Dirac dağılımı konvolüsyon işleminin etkisiz elemanıdır. Çünkü,

$F$ herhangi bir dağılım fonksiyonu olmak üzere,

$$
\begin{aligned}
F^{0*}*F(t) &= \int_{-\infty}^{+\infty} F(t-x)dF^{0*}(x) \\
&= \int_{-\infty}^{0^-} F(t-x)dF^{0*}(x) + \int_{0^-}^{0^+} F(t-x)dF^{0*}(t) + \int_{0^+}^{\infty} F(t-x)dF^{0*}(t) \\
&= F(t-0)(F^{0*}(0^+) - F^{0*}(0^-)) \\
&= F(t)
\end{aligned}
$$

dir.

**Teorem 2.1** $X$ ve $Y$ birbirlerinden bağımsız sırasıyla $F$ ve $G$ dağılım fonksiyonlarına sahip iki rasgele değişken olsun. Bu durumda $X+Y$ rasgele değişkeninin dağılım fonksiyonu,

$$F_{X+Y}(t) = F*G(t), \quad t \in \mathbb{R}$$

dir.

**İspat**

$$\begin{aligned}
F_{X+Y}(t) &= P(X+Y \le t) \\
&= \int_{-\infty}^{+\infty} P(X+Y \le t \mid Y=y)dG(y) \\
&= \int_{-\infty}^{+\infty} P(X \le t-y \mid Y=y)dG(y) \\
&= \int_{-\infty}^{+\infty} P(X \le t-y)dG(y) \\
&= \int_{-\infty}^{+\infty} F(t-y)dG(y) \\
&= F*G(t)
\end{aligned}$$

**Teorem 2.2** $F$ mutlak sürekli bir dağılım fonksiyonu olmak üzere $G$ herhangi bir dağılım fonksiyonu olsun. Bu durumda $F*G$ dağılım fonksiyonu mutlak süreklidir (Feller 1971).

$F$ ve $G$ dağılım fonksiyonları sırasıyla $f$ ve $g$ olasılık yoğunluk fonksiyonlarına sahip olsun. Bu durumda

$$f*g(t) = \int_{-\infty}^{\infty} g(t-x)f(x)dx, \qquad t \in \mathbb{R}$$

olmak üzere $F*G$ dağılım fonksiyonu $f*g$ olasılık yoğunluk fonksiyonuna sahiptir.

$X_1,\ldots,X_n$ birbirlerinden bağımsız aynı $F$ dağılım fonksiyonuna sahip $n$ tane rasgele değişken olsun. $F^{n*} = F*...*F$ ($n$ kere) olmak üzere $S_n = X_1 + ... + X_n$ rasgele değişkeninin dağılım fonksiyonu $P(S_n \leq t) = F^{n^*}(t)$ dir. Ayrıca $F$ dağılım fonksiyonu $f$ yoğunluk fonksiyonuna sahip ise $f^{n^*} = f * ... * f$ ($n$ kere) olmak üzere $F^{n^*}$ dağılım fonksiyonu da $f^{n^*}$ yoğunluk fonksiyonuna sahiptir (Feller 1971).

## 2.2 Laplace-Stieltjes ve Laplace Dönüşümü

$\mathbb{R}$ ' de tanımlı, sınırlı ve azalmayan bir $F$ fonksiyonunun Laplace-Stieltjes dönüşümü,

$$F_{LS}(t) = \int_{-\infty}^{\infty} e^{-xt} dF(x), \qquad -\infty < t < \infty$$

ile tanımlanır. $F$ ve $G$, $\mathbb{R}$ de sınırlı, azalmayan ve $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to -\infty} G(x) = 0$ şartını sağlayan iki fonksiyon olsun. Bu durumda,

$$F_{LS}(t) = \int_{-\infty}^{\infty} e^{-xt} dF(x)$$

$$G_{LS}(t) = \int_{-\infty}^{\infty} e^{-xt} dG(x)$$

olmak üzere

$$\left(F * G\right)_{LS}(t) = F_{LS}(t) G_{LS}(t)$$

dir (Kawata 1972).

Tek değişkenli bir F fonksiyonunun Laplace dönüşümü aşağıdaki integralin mevcut olması durumunda

$$F_L(t) = \int_{-\infty}^{\infty} e^{-tx} F(x) dx$$

ile verilir.

## 2.3 Yenileme Denklemi ve Çözümü

$A$ bilinmeyen bir fonksiyon olmak üzere dağılım fonksiyonu özelliklerine sahip bir $F$ fonksiyonu ile $a$ fonksiyonu bilinsin. Bu durumda

$$A(t) = a(t) + \int_0^t A(t-x)dF(x), \qquad t \geq 0 \tag{2.1}$$

tipindeki bir integral denkleme 'Yenileme Denklemi' denir (Karlin ve Taylor 1975). $F$ bir dağılım fonksiyonu olsun. $F$ bir $f$ yoğunluk fonksiyonuna sahipse (2.1) denklemi

$$A(t) = a(t) + \int_0^t A(t-x)f(x)dx, \qquad t \geq 0$$

$$= a(t) + \int_0^t f(t-x)A(x)dx, \qquad t \geq 0$$

integral denklemine dönüşür.

**Teorem 2.3** $a$ sınırlı bir fonksiyon olsun. Bu durumda (2.1) yenileme denkleminin sonlu aralıklar üzerinde sınırlı bir tek $A$ çözümü vardır ve bu çözüm olan fonksiyon

$$A(t) = a(t) + \int_0^t a(t-x)d(\sum_{n=1}^{\infty} F^{n^*}(x)), \qquad t \geq 0 \tag{2.2}$$

dır (Karlin ve Taylor 1975).

## 3. YENİLEME TEORİSİ

Bir yenileme süreci, zamanın bir fonksiyonu olarak gerçekleşen olayların sayısını sayan bir stokastik süreçtir; bu süreçte ardışık yenilemeler arasında geçen zaman süreleri birbirlerinden bağımsız ve aynı dağılımlı rasgele değişkenlerdir.

Yenileme süreçleri, üreticiler için yer değiştirme problemlerinin analiziyle başlamış olsa bile yenileme süreçleri ile ilgili teori (yenileme teorisi) uygulamalı olasılık problemlerinin geniş bir sahasında çok sayıda uygulamaya sahiptir. Örneğin, envanter teorisinde bir yenileme süreci, talep noktaları arasındaki ardışık zaman sürelerini modellemek için kullanılır. Güvenilirlik problemlerinde ise bozulan bir makinenin ardışık onarımları ya da yer değiştirmelerini modellemek için kullanılır. Buna ilave olarak yenileme teorisinin işçi sayısının planlanması problemlerinde güçlü bir araç olduğu kanıtlanmıştır. Buradaki yenileme süreci, atama (tayin) ile verilen bir işten ayrılışların dizisini modellemek için kullanılır.

### 3.1 Yenileme Süreçleri

$\{X_n, \quad n=1,2,...\}$ negatif olmayan, bağımsız ve aynı $F$ dağılım fonksiyonuna sahip rasgele değişkenlerin dizisi olsun. Aşikar durumlardan kaçınmak için kabul edelim ki $F(0) = P(X_n = 0) < 1$, yani $X_n$ bir olasılıkla sıfıra eşit olmasın.

$X_1$: İlk yenileme yapılıncaya kadar geçen zaman süresi

$X_2$: İlk yenileme yapıldıktan sonra, ikinci yenilemeye kadar geçen zaman süresi

.

.

.

$X_n$: $(n-1)$. yenileme yapıldıktan sonra, $n$.yenilemeye kadar geçen zaman süresi olarak ifade edilsin. $S_0 = 0$, $S_n = X_1 + ... + X_n$ , $n \geq 1$ olmak üzere $S_n$ rasgele değişkeni $n$. yenileme yapılıncaya kadar geçen zaman süresidir. Her $t \geq 0$ için $N(t)$,

$$N(t) = \sup\{n : S_n \le t\}$$

ile tanımlansın. $N(t)$ sadece $t$ zamanına kadar yani $[0,t]$ zaman aralığındaki yenilemelerin sayısıdır. Bu şekilde tanımlanan $N(t)$ yenileme rasgele değişken ve $\{N(t),\, t \ge 0\}$ stokastik süreci de bir 'yenileme süreci' olarak adlandırılır (Ross 1983).

$F(0) < 1$ olduğundan $F$ dağılımının ortalaması $\mu$ sıfırdan farklı olup güçlü büyük sayılar yasasının göz önüne alınmasıyla $S_n$ en fazla $n$'nin sonlu bir sayıdaki değerleri için $t$'ye eşit ya da $t$'den küçük olabilir. Bu nedenle $N(t)$ sonlu olmak zorundadır ve $N(t)$'nin yukarıdaki tanımında sup yerine maks kullanılabilir. $N(t)$ sonlu bir rasgele değişken olmasına rağmen bir olasılıkla $\lim_{t\to\infty} N(t) = \infty$'dur. Çünkü

$$\begin{aligned} P(N(\infty) < \infty) &= P(En\ az\ bir\ n\ için\ X_n = \infty) \\ &= P\Big(\bigcup_{n=1}^{\infty}(X_n = \infty)\Big) \\ &\le \sum_{n=1}^{\infty} P(X_n = \infty) \\ &= 0 \end{aligned}$$

dır. Her sabit $t \ge 0$ için $N(t)$ yenileme rasgele değişkeninin olasılık dağılımı,

$$\begin{aligned} P(N(t) = n) &= P(S_n \le t\ ,\ S_{n+1} > t) \\ &= P(S_{n+1} > t) - P(S_{n+1} > t\ ,\ S_n > t) \\ &= P(S_{n+1} > t) - P(S_n > t) \\ &= F^{n*}(t) - F^{(n+1)*}(t), \quad n = 0,1,... \end{aligned}$$

ile verilir.

**Teorem 3.1** $\{N(t),\ t \ge 0\}$ yenileme süreci her mertebeden sonlu momentlere sahiptir, yani her $t \ge 0$ $k \ge 0$ için $E(N^k(t)) < \infty$.

**İspat** $P(X_i = 0) < 1$ olduğundan dolayı $P(X_i > \alpha) = \beta > 0$ olacak biçimde bir $\alpha > 0$ sayısı bulunabilir. $i = 1, 2, ...$ olmak üzere

$$X_i^{'} = \begin{cases} \alpha, & X_i > \alpha \\ 0, & X_i \le \alpha \end{cases}$$

rasgele değişkenleri göz önüne alınsın.

$$f_{X_i^{'}} = \begin{cases} F(\alpha), & x = 0 \\ 1 - F(\alpha), & x = \alpha \\ 0, & d.y, \end{cases} \qquad ya\ da \quad F_{X_i^{'}} = \begin{cases} 0, & x < 0 \\ 1 - \beta, & 0 \le x < \alpha \\ 1, & x \ge \alpha \end{cases}$$

olduğundan $X_i^{'}$'ler aynı dağılımlıdırlar. Ayrıca $X_i^{'}$'lerin bağımsız oldukları da açıktır. O halde $\{X_i^{'}, i = 1, 2, ...\}$ negatif olmayan birbirlerinden bağımsız aynı dağılımlı rasgele değişkenlerin bir dizisidir. $\{N^{'}(t), t \ge 0\}$ bu dizi üzerine kurulu diğer bir yenileme süreci olsun. Bu süreçte yenilemeler yalnızca $t = n\alpha, \quad n = 0, 1, ...$ zamanlarında gerçekleşebilirler.

$Y_1$, sıfırdan $\alpha$ zamanına ilk geçinceye kadar gerçekleşen yenilemelerin sayısı,

$Y_2$, $\alpha$ dan $2\alpha$ zamanına ilk geçinceye kadar gerçekleşen yenilemelerin sayısı,

.

.

.

$Y_n$, $(n-1)\alpha$ dan $n\alpha$ zamanına ilk geçinceye kadar gerçekleşen yenilemelerin sayısı olmak üzere

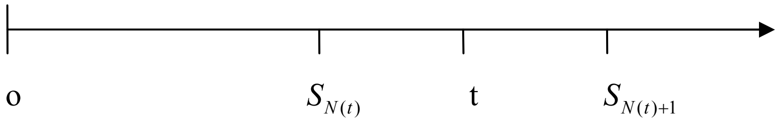
$$Y_1 + Y_2 + ... + Y_n = 1 + N^{'}(t), \qquad (n-1)\alpha \le t < n\alpha \qquad ; \quad n = 1, 2, ...$$

elde edilir. Her i için $X_i^{'} \le X_i$ olduğundan $N(t) \le N^{'}(t)$ dir. $Y_i$'lerin birbirlerinden bağımsız $\beta$ başarı olasılıklı geometrik dağılıma sahip rasgele değişkenler oldukları açıktır. Bu durumda $1 + N^{'}(t)$ negatif binom dağılımına sahip bir rasgele değişkendir. Negatif binom dağılımı her mertebeden sonlu momentlere sahip olup $N(t) < 1 + N^{'}(t)$ olduğundan $N(t)$ her mertebeden sonlu momentlere sahiptir.

### 3.1.1 N(t) Rasgele değişkeniyle ilgili bazı asimptotik sonuçlar

**Teorem 3.2** $P(\lim_{t\to\infty}\frac{N(t)}{t}=\frac{1}{\mu})=1$.

**İspat**




$$S_{N(t)} \leq t < S_{N(t)+1} \tag{3.1}$$

olduğunu biliyoruz. Güçlü büyük sayılar yasasından bir olasılıkla

$$\frac{S_{N(t)}}{N(t)} \xrightarrow[t\to\infty]{} \mu \qquad \text{ve} \qquad \frac{S_{N(t)+1}}{N(t)+1} \xrightarrow[t\to\infty]{} \mu$$

yazılabilir. (3.1) eşitsizliğinden

$$\frac{S_{N(t)}}{N(t)} \leq \frac{t}{N(t)} \leq \frac{S_{N(t)+1}}{N(t)+1}\frac{N(t)+1}{N(t)}$$

olur. $t\to\infty$ ile limit alınırsa, bir olasılık ile

$$\mu \leq \lim_{t\to\infty}\frac{t}{N(t)} \leq \mu$$

elde edilir. Böylece,

$$P(\lim_{t\to\infty}\frac{N(t)}{t}=\frac{1}{\mu})=1$$

sonucuna ulaşılır.

Teorem 3.1'in kullanılmasıyla uzun süre çalışmakta olan bir yenileme sürecinde birim zamanda yapılan yenilemelerin sayısı yaklaşık olarak $1/\mu$ olduğu sonucuna varılır.

**Teorem 3.3** (Yenileme süreçleri için merkezi limit teoremi) $\mu$ ve $\sigma^2$ sırasıyla $X_n$ rasgele değişkeninin, yani bir yenileme aralığının uzunluğu sırasıyla beklenen değeri ve varyansı olmak üzere sonlu olsunlar. Bu durumda

$$\lim_{t\to\infty} P\left(\frac{N(t)-(\frac{t}{\mu})}{\sigma\sqrt{\frac{t}{\mu^3}}} < x\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-x^2/2}dx$$

dır (Karlin ve Taylor 1975).

**İspat** $x$ sabit ve $\lim\limits_{\substack{t\to\infty\\ n\to\infty}} \frac{t-n\mu}{\sigma\sqrt{n}} = -x$ olacak biçimde $n$ ve $t$ sonsuza götürülsün. Bu durumda $\Phi$ standart normal dağılımın dağılım fonksiyonu olmak üzere merkezi limit teoreminden

$$\begin{aligned}\lim_{\substack{t\to\infty\\ n\to\infty}} P\left(S_n > t\right) &= \lim_{\substack{t\to\infty\\ n\to\infty}} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} > -x\right)\\ &= 1-\Phi(-x)\\ &= \Phi(x)\end{aligned}$$

$P(S_n > t) = P\left\{ N(t) < n \right\}$ olduğundan

$$\begin{aligned}\Phi(x) &= \lim_{\substack{t\to\infty\\ n\to\infty}} P(S_n > t)\\ &= \lim_{\substack{t\to\infty\\ n\to\infty}} P\left(N(t) < n\right)\end{aligned}$$

$$\lim_{\substack{t\to\infty\\ n\to\infty}} P\left(\frac{N(t)-t/\mu}{\sigma\sqrt{t/\mu^3}} < \frac{n-t/\mu}{\sigma\sqrt{t/\mu^3}}\right)$$

yazılabilir.

$$\begin{aligned}\frac{n-t/\mu}{\sigma\sqrt{t/\mu^3}} &= \frac{n\mu - t}{\sigma\sqrt{t/\mu}}\\ &= -\frac{(t-n\mu)}{\sigma\sqrt{n}}\frac{\sqrt{n\mu}}{\sqrt{t}}\end{aligned}$$

olup $\lim\limits_{\substack{t\to\infty\\ n\to\infty}} \frac{\sqrt{n\mu}}{\sqrt{t}} = 1$ olacağından

$$\lim_{\substack{t\to\infty \\ n\to\infty}} \frac{n - t/\mu}{\sigma\sqrt{t/\mu^3}} = x$$

elde edilir. Bu durumda

$$\Phi(x) = \lim_{\substack{t\to\infty \\ n\to\infty}} p\left(\frac{N(t) - t/\mu}{\sigma\sqrt{t/\mu^3}} < \frac{n - t/\mu}{\sigma\sqrt{t/\mu^3}}\right)$$

$$= \lim_{t\to\infty} p\left(\frac{N(t) - t/\mu}{\sigma\sqrt{t/\mu^3}} < x\right)$$

olur.

Yukarıdaki teoremden $t$ büyükken $N(t)$ rasgele değişkeninin asimptotik dağılımı yaklaşık olarak $\frac{t}{\mu}$ ortalama ve $\frac{\sigma^2 t}{\mu^3}$ varyansı ile asimptotik normal dağılımlıdır, yani

$$N(t) \sim AN\left(\frac{t}{\mu}, \frac{\sigma^2 t}{\mu^3}\right)$$

dir.

## 3.2 Yenileme fonksiyonu

$\{ N(t),\ t \geq 0 \}$ bir yenileme süreci olmak üzere,

$$M(t) = E(N(t)), \qquad t \geq 0$$

ile verilen $M$ ortalama değer fonksiyonuna yenileme fonksiyonu denir (Karlin ve Taylor 1975). Burada $M(t)$, $[0,t]$ zaman aralığında yapılan yenilemelerin ortalama sayısıdır.

$$I_k = \begin{cases} 1, & S_k \leq t \\ 0, & S_k > t \end{cases}$$

olsun.

$$N(t) = \sum_{k=1}^{\infty} I_k$$

olup

$$
\begin{aligned}
E(N(t)) &= E\left(\sum_{k=1}^{\infty} I_k\right) \\
&= \sum_{k=1}^{\infty} E(I_k) \\
&= \sum_{k=1}^{\infty} P(S_k \leq t) \\
&= \sum_{k=1}^{\infty} F^{k*}(t)
\end{aligned}
$$

elde edilir. O halde

$$
M(t) = \sum_{k=1}^{\infty} F^{k*}(t) \ , \qquad t \geq 0 \tag{3.2}
$$

dir.

(3.2) ifadesinin kullanılmasıyla $M$ ’nin sağdan sürekli ve azalmayan bir fonksiyon olduğu kolaylıkla elde edilir. Bununla birlikte $\lim_{t\to\infty} M(t) = \infty$ olup $M$ yenileme fonksiyonu $t \to \infty$ için bire yakınsamaması dışında bir dağılım fonksiyonunun tüm özelliklerine sahiptir.

$A$ bilinmeyen ve $a$ bilinen bir fonksiyon olmak üzere (2.1) ile verilen

$$
A(t) = a(t) + \int_0^t A(t-x)dF(x), \qquad t \geq 0
$$

integral denklemini göz önüne alalım. Teorem 3.3’den dolayı bu integral denklemin çözüm fonksiyonunun $M$ yenileme fonksiyonuna bağlı olarak

$$
A(t) = a(t) + \int_0^t a(t-x)dM(x), \qquad t \geq 0
$$

şeklinde ortaya çıkar.

(3.2) ifadesinin kullanılmasıyla $M$ yenileme fonksiyonu için bir integral denklem elde edilebilir. $F^{(k+1)*}(t)$ yerine matematiksel olarak denk olan $\int_0^t F^{k*}(t-x)dF(x)$ integralinin alınmasıyla

$$M(t) = F(t) + \int_0^t M(t-x)dF(x)$$
$$= F(t) + F*M(t), \quad t \geq 0 \tag{3.3}$$

bulunur. Bu integral denklem bir yenileme denklemidir. (3.3) denklemi

$$M(t) = F(t) + \int_o^t F(t-x)dM(x), \qquad t \geq 0 \tag{3.4}$$

olarak yazılabilir. Çünkü $M*F = F*M$ dir. $F$ dağılım fonksiyonu $f$ olasılık yoğunluk fonksiyonuna sahip ise (3.3) denklemi

$$M(t) = F(t) + \int_0^t M(t-x)f(x)dx, \qquad t \geq 0 \tag{3.5}$$

biçiminde yazılabilir.

$F$ ve $M$ fonksiyonlarının sırasıyla Laplace-Stieltjes dönüşümleri

$$F_{LS}(t) = \int_0^\infty e^{-tx}dF(x)$$

ve

$$M_{LS}(t) = \int_0^\infty e^{-tx}dM(x)$$

olmak üzere (3.3) yenileme denkleminden, Laplace-Stieltjes dönüşümünün kullanılmasıyla

$$M_{LS}(t) = F_{LS}(t) + F_{LS}(t)M_{LS}(t)$$

elde edilir. O halde

$$M_{lS}(t) = \frac{F_{LS}(t)}{1-F_{LS}(t)}, \qquad t > 0 \tag{3.6}$$

ve

$$F_{LS}(t) = \frac{M_{LS}(t)}{1+M_{LS}(t)}, \qquad t > 0 \tag{3.7}$$

dir. Ayrıca $F$ dağılım fonksiyonu $f$ olasılık yoğunluk fonksiyonuna sahip iken $f$ ve $M$ fonksiyonlarının sırasıyla Laplace dönüşümleri

$$f_L(t) = \int_0^\infty e^{-tx}f(x)dx$$

$$M_L(t) = \int_0^\infty e^{-tx} M(x)dx$$

olmak üzere

$$M_L(t) = \frac{f_L(t)}{t(1-f_L(t))} \tag{3.8}$$

olduğu kolaylıkla elde edilebilir (Cox 1962).

Bir fonksiyonun Laplace-Stieltjes dönüşümünün o fonksiyonu tek olarak belirlemesinden (Kawata 1972) sırasıyla (3.6) ve (3.7) eşitliklerinin de kullanılmasıyla $F$ dağılım fonksiyonu $M$ yenileme fonksiyonunu ve $M$ yenileme fonksiyonu da $F$ dağılım fonksiyonu tek olarak belirler.

Şimdi $M$ 'nin türev fonksiyonunu bulalım. $F$ dağılım fonksiyonu $f$ olasılık yoğunluk fonksiyonuna sahip ise

$$\frac{dM(t)}{dt} = \frac{d}{dt}\sum_{k=1}^{\infty} F^{k*}(t) = \sum_{k=1}^{\infty}\frac{d}{dt}F^{k*}(t) = \sum_{k=1}^{\infty} f^{k*}(t)$$

dir. Bu durumda

$$m(t) = \frac{dM(t)}{dt}$$

ile tanımlı $m$ fonksiyonuna 'Yenileme Yoğunluğu' denir (Ross 1983).

$$M(t) = \int_0^t m(x)dx$$

olduğu açıktır. (3.5) integral denkleminde $t$ ye göre türev alınmasıyla

$$m(t) = f(t) + \int_0^t m(t-x)f(x)dx$$

lineer ikinci çeşit Volterra integral denklemi elde edilir. $\Delta t > 0$ olmak üzere

$$M(t+\Delta t) - M(t) = \sum_{k=1}^{\infty} P(t < S_k \le t + \Delta t)$$

olduğundan, $m(t)\Delta t$ dar $(t\,,\, t+\Delta t)$ aralığında bir yenileme olması olasılığı olarak ifade edilebilir. Ayrıca $m(t)$ $t$ civarındaki dar bir aralıkta beklenen yenilemelerin ortalama sayısını tarif eder.

### 3.2.1 Yenileme fonksiyonu ile ilgili bazı limit teoremler

İlk olarak uzun süre çalışmakta olan bir yenileme sürecinde birim zamanda yapılan yenilemelerin beklenen sayısının yaklaşık $1/\mu$ olduğunu ifade eden 'elemanter yenileme teoremini' verelim.

**Teorem 3.4** (Karlin ve Taylor 1975) $\{N(t),\ t \geq 0\}$ her $n$ için $E(X_n) = \mu < \infty$ olacak biçimde bir yenileme süreci olsun. Bu durumda,

$$\lim_{t \to \infty} \frac{M(t)}{t} = \frac{1}{\mu}$$

dir.

$M(t)$'nin bunun dışındaki asimptotik özelliklerinin bulunması için hazırlayıcı bir tanımı aşağıda verilmektedir.

**Tanım 3.1** Bir $X$ rasgele değişkeni $\lambda > 0$ olmak üzere bir olasılıkla $\{k\lambda : k = 0, \pm 1, ...\}$ kümesinden değerler alır ise $X$ rasgele değişkenine ve onun dağılım fonksiyonuna aritmetik denir. Bu özelliğe sahip en büyük $\lambda$ sayısı da dağılımın gereni olarak adlandırılır (Grimmett and Stirzaker 1992).

Bu çalışmada yenileme aralıklarının aritmetik olmadığı durum göz önüne alınır. Aritmetik durumda çoğu kez benzer sonuçlar geçerlidir.

Şimdi anahtar yenileme teoremi olarak bilinen teoremi verelim. Yenileme teorisindeki birtakım limit sonuçlara doğrudan bu teorem yardımıyla ulaşılabilir.

**Teorem 3.5** $a$, $(0, \infty)$ aralığı üzerinde tanımlı ve

i. Her $t > 0$ için $a(t) \geq 0$

ii. $\int_0^\infty a(t)dt < \infty$

iii. $a$ artmayan

özelliklerine sahip bir fonksiyon olsun. Bu durumda $F$ aritmetik olmayan bir dağılım fonksiyonu ise

$$\lim_{t\to\infty}\int_0^t a(t-x)dM(x) = \frac{1}{\mu}\int_0^\infty a(x)dx$$

dir (Smith 1958).

**Sonuç 3.1** Bir yenileme sürecinde ardışık yenilemeler arası geçen süreler aritmetik olmayan bir $F$ dağılım fonksiyonuna ve sonlu $\mu$ ortalamasına sahip olsun. Bu durumda $a$ fonksiyonu yukarıdaki üç özelliğe sahip ise (2.1) de verilen

$$A(t) = a(t) + \int_0^t A(t-x)dF(x)$$

yenileme denklemi için

$$\lim_{t\to\infty} A(t) = \frac{1}{\mu}\int_0^\infty a(x)dx$$

elde edilir.

**İspat** (2.1) yenileme denkleminin çözüm fonksiyonu Teorem 2.3 den

$$A(t) = a(t) + \int_0^t a(t-x)dM(x)$$

ile verilir.Bu durumda $t\to\infty$ ile limiti alınıp Teorem 3.5 kullanılırsa

$$\lim_{t\to\infty} A(t) = \lim_{t\to\infty} a(t) + \frac{1}{\mu}\int_0^\infty a(x)dx$$

elde edilir. Ayrıca $a$ 'nın sahip olduğu özelliklerden $\lim_{t\to\infty} a(t) = 0$ olacağından

$$\lim_{t\to\infty} A(t) = \frac{1}{\mu}\int_0^\infty a(x)dx$$

olacaktır.

Şimdi yenileme teoremi olarak bilinen **'**Blackwell teoremini' verelim.

**Teorem 3.6** Herhangi bir $\{ N(t),\ t \geq 0\}$ yenileme sürecinde ardışık yenilemeler arası geçen süreler aritmetik olmayan $F$ dağılım fonksiyonu ve sonlu $\mu$ ortalamasına sahip ise herhangi bir $h > 0$ için

$$\lim_{t\to\infty}\left[M(t) - M(t-h)\right] = \frac{h}{\mu}$$

dir (Smith 1958).

**İspat** $h > 0$ olmak üzere

$$a(t) = \begin{cases} \dfrac{1}{h}, & 0 < t \leq h \\ 0, & t > h \end{cases}$$

alınmasıyla, Teorem 3.5 kullanılarak istenilen sonuca ulaşılır

Bu teoremde uzun süredir çalışan bir yenileme sürecinde $h$ uzunluğundaki bir aralıkta yapılan yenilemelerin beklenen sayısının yaklaşık $h/\mu$ olduğu ifade edilir.

Şimdi $M$ yenileme fonksiyonunun asimptotik açılımında ikinci bir terim belirleyen aşağıdaki teoremi verelim.

**Teorem 3.7** $F$ sonlu $\mu_2$ ikinci momente sahip aritmetik olmayan bir dağılım fonksiyonu olsun. Bu durumda

$$\lim_{t\to\infty}\left(M(t) - \frac{t}{\mu}\right) = \frac{\mu_2}{2\mu^2} - 1 \qquad (3.\ 10)$$

dir (Karlin ve Taylor 1975).

Yukarıdaki teoremde $M(t)$ için verilen asimptotik ifade $t$'nin büyük değerleri için $M(t)$'nin hesaplanmasında ve aynı zamanda $F$ bilinmediğinde $M(t)$'nin tahmininde çok faydalıdır. $\mu$ ve $\sigma^2$ sırasıyla $F$ dağılımının beklenen değerini ve varyansını göstermek üzere $c^2 = \dfrac{\sigma^2}{\mu^2}$ diyelim, burada $c$, $F$ dağılımının değişim katsayısıdır. Bu durumda (3.10) asimptotik ifadesinden, yeterince büyük $t$ için

$$M(t) \approx \frac{t}{\mu} + \frac{1}{2}(c^2 - 1)$$

yazılabilir. $F$, $\mu$ ortalamalı üstel dağılım fonksiyonu, yani $c = 1$ iken yukarıdaki asimptotik açılım her $t \geq 0$ için $M(t)$'nin kendisine eşittir. $c^2$ çok büyük ya da çok küçük değilse bu asimptotik açılım $t$'nin bazı değerleri için pratikte yeterince doğrudur. Sayısal incelemeler göstermiştir ki $M(t)$ için $\dfrac{t}{\mu} + \dfrac{1}{2}(c^2 - 1)$ açılımı pratikte $t \geq t_0$ için kullanılabilir, burada

$$t_0 = \begin{cases} \dfrac{3}{2} c^2 \mu, & c^2 > 1 \\ \mu, & 0.2 < c^2 \leq 1 \\ \dfrac{\mu}{2c}, & 0 < c^2 \leq 0.2 \end{cases}$$

dir (Tijms 1994). $c^2 \to 0$ ile açılım kötüleşir. Yaklaşımın göreli hatası $t \geq t_0$ için tipik olarak %5'in altındadır ve çoğu kez %2'den daha küçüktür.

# 4. YENİLEME FONKSİYONUNUN HESABI

Bir $\{ N(t),\ t \geq 0\}$ yenileme sürecinde yenilemeler arası geçen zaman süreleri dağılım fonksiyonu $F$ bilindiğinde sürecin $M$ yenileme fonksiyonu görünüşte (3.2), (3.3) ve (3.6) ifadelerinin birinden elde edilebilir. Fakat genelde iki parametreli üstel, düzgün, hiper üstel ve gamma dağılımı dışında $M$ yenileme fonksiyonu bu denklemlerden analitik olarak elde edilemez. Bu durumda $M$ sayısal olarak elde edilebilir. Literatürde Laplace ve Ters Laplace dönüşümlerinin hesabına, kuvvet serileri açılımına, kübik Spline yaklaşımına ve yenileme integral denkleminin sayısal hesabına dayalı bazı yöntemler vardır (Baxter 1981, Baxter ve diğerleri 1982, Xie 1989). Kolay olarak programlanabilmesi hemen hemen tüm durumlarda basitliği ve yakınsaklığı ile iyi sonuçlar veren ve diğer bilinen yöntemlerle karşılaştırıldığında uygulanabilirliği daha fazla olan bir yöntem Xie'nin RS(Rieman-Stieltjes) yöntemidir (Xie 1989). Bu kısımda yukarıda belirtilen dağılımlar için $M$ yenileme fonksiyonun analitik ifadeleri verilir. Ayrıca bu fonksiyonun sayısal hesabı için RS yöntemi üzerinde durulur.

## 4.1 Yenileme Fonksiyonunun Analitik Olarak İncelenmesi

Bu kısımda iki parametreli üstel, düzgün, hiper üstel ve gamma dağılım fonksiyonlarına dayalı yenileme süreçlerin yenileme fonksiyonlarının analitik olarak elde edilmesi üzerinde durulmaktadır.

### 4.1.1 İki parametreli üstel dağılım

$X$ rasgele değişkenin olasılık yoğunluk fonksiyonu

$$f(x) = \begin{cases} \theta_1 e^{-\theta_1 (x-\theta_2)}, & x > \theta_2;\ \theta_1, \theta_2 > 0 \\ 0 \qquad\qquad , & dy \end{cases}$$

olsun. $n = 1, 2, ...$ için $X_n$'ler yukarıda verilen iki parametreli üstel dağılıma sahip birbirinden bağımsız rasgele değişkenler olmak üzere $\{ N(t),\ t \geq 0\}$ yenileme süreci göz önüne alınsın. $n \geq 1$ için $S_n = X_1 + ... + X_n$ rasgele değişkeninin olasılık yoğunluk fonksiyonu ardışık olarak hesaplanan konvolüsyon integrallerinden

$$f^{n*}(x) = \begin{cases} \dfrac{\theta_1^n (x - n\theta_2)^{n-1}}{(n-1)!} e^{-\theta_1 (x - n\theta_2)}, & x > n\theta_2 \\ 0 \qquad\qquad\qquad\qquad , & x \leq n\theta_2 \end{cases}$$

bulunur. Bu ifade yardımıyla $S_n$ 'nin dağılım fonksiyonu

$$F^{n*}(x) = \begin{cases} 0 & , \quad x > n\theta_2 \\ 1 - \sum_{i=0}^{n-1} \dfrac{\left[\theta_1 (x - n\theta_2)\right]^i}{i!} e^{-\theta_1 (x - n\theta_2)} & , \quad x \leq n\theta_2 \end{cases}$$

elde edilir. $F^{n*}(x)$ 'in analitik ifadesine bağlı olarak bu yenileme süreciyle ilgili fonksiyonlar analitik olarak elde edilebilir. $N(t)$ yenileme rasgele değişkeninin olasılık fonksiyonu,

$$P(N(t) = n) = \begin{cases} 0, & t \leq n\theta_2 \\ e^{-\theta_1 (t - (n+1)\theta_2)} \sum_{i=0}^{n} \dfrac{\left[\theta_1 (t - (n+1)\theta_2)\right]^i}{i!} - e^{-\theta_1 (t - n\theta_2)} - \sum_{i=0}^{n-1} \dfrac{\left[\theta_1 (t - n\theta_2)\right]^i}{i!}, & (n+1)\theta_2 \leq t \\ 1 - \sum_{i=0}^{n-1} \dfrac{\left[\theta_1 (t - n\theta_2)\right]^i}{i!} e^{-\theta_1 (t - n\theta_2)}, & n\theta_2 \leq t \leq (n+1)\theta \end{cases}$$

dır (Barlow ve Proschan 1965). (3.2) konvolüsyon serisi yardımıyla $M(t)$ yenileme fonksiyonunun elde edilmesi üzerinde duralım. $r = \left[\left|t / \theta_2\right|\right]$ olmak üzere,

$$\begin{aligned} M(t) &= \sum_{n=1}^{\infty} F^{n*}(t) \\ &= \sum_{n=1}^{r} F^{n*}(t) \\ &= r - \sum_{i=0}^{r-j} e^{-\theta_1 (t - (r-j)\theta_2)} \frac{\left[\theta_1 (t - (r - j + 1)\theta_2)\right]^i}{i!} \\ &= r - \sum_{i=0}^{r} e^{-\theta_1 (t - s\theta_2)} \sum_{i=0}^{s-1} \frac{\left[\theta_1 (t - s\theta_2)\right]^i}{i!} \end{aligned}$$

dir.

### 4.1.2 Düzgün dağılım

$X$ rasgele değişkeninin olasılık yoğunluk fonksiyonu,

$$f(x)=\begin{cases}\dfrac{1}{\theta}, & 0<x<\theta \\ 0, & d.y\end{cases}$$

olsun. $n=1,2,...$ için $X_n$'ler yukarıda verilen düzgün dağılıma sahip bağımsız rasgele değişkenler olmak üzere $\{N(t),\ t\geq 0\}$ yenileme süreci göz önüne alınsın. $n\geq 1$ için $S_n=X_1+...+X_n$ rasgele değişkeninin olasılık yoğunluk fonksiyonu ardışık olarak hesaplanan konvolüsyon integrallerinden

$$f^{n^*}(x)=\begin{cases}\dfrac{x^{n-1}}{(n-1)!\theta^n} & , \ 0<x\leq\theta \\ \dfrac{1}{(n-1)!\theta^n}\left[x^{n-1}-\binom{n}{1}(x-\theta)^{n-1}\right] & , \ \theta\leq x\leq 2\theta \\ \dfrac{1}{(n-1)!\theta^n}\left[x^{n-1}-\binom{n}{1}(x-\theta)^{n-1}+\binom{n}{2}(x-2\theta)^{n-1}\right] & , \ 2\theta\leq x\leq 3\theta \\ . \\ . \\ . \\ \dfrac{1}{(n-1)!\theta^n}\left[x^{n-1}-\binom{n}{1}(x-\theta)^{n-1}+\binom{n}{2}(x-2\theta)^{n-1}-...+(-1)^{n-1}(x-(n-1)\theta)^{n-1}\right] & , \ (n-1)\theta\leq x\leq n\theta \\ 0 & , \ dy\end{cases}$$

$$=\begin{cases}\dfrac{1}{(n-1)!\theta^n}\displaystyle\sum_{j=0}^{k}(-1)^j\binom{n}{j}(x-j\theta)^{n-1}, & k\theta\leq x\leq(k+1)\theta\ ;\ k=0,1,...,n-1 \\ 0 & , \quad d.y\end{cases}$$

elde edilir. Bu ifade yardımıyla $S_n$ 'nin dağılım fonksiyonu

$$F^{n^*}(x)=\begin{cases}0 & , \ x<0 \\ \dfrac{1}{n!\theta^n}\displaystyle\sum_{j=0}^{k}(-1)^j\binom{n}{j}(x-j\theta)^{n}, & k\theta\leq x\leq(k+1)\theta\ ;\ k=0,1,...,n-1 \\ 1 & , \ x\geq n\theta\end{cases}$$

bulunur. Böylece $N(t)$ yenileme rasgele değişkenin olasılık fonksiyonu için bir kapalı form

$$P\left(N(t)=n\right)=\begin{cases}0 & ,(n+1)\theta\le t\\ \left((n+1)\theta-t\right)\sum_{j=0}^{k}\dfrac{(-1)^j}{(n-j+1)}\dbinom{n}{j}(t-j\theta)^n & ,k\theta\le t\le(k+1)\theta\ ;k=0,1,...,n-1\\ 1-\dfrac{1}{(n+1)!\theta^{n+1}}\sum_{j=0}^{n}(-1)^j\dbinom{n+1}{j}(t-j\theta)^{n+1} & ,n\theta\le t\le(n+1)\theta\end{cases}$$

olarak bulunur. Ayrıca $M(t)$ konvolüsyon serisi yardımıyla aşağıdaki gibi bulunur.

$$M(t)=\sum_{n=1}^{\infty}F^{n*}(t)$$

$$=\begin{cases}-1+e^{t/\theta} & ,0\le t\le\theta\\ -1+e^{t/\theta}-\dfrac{t-\theta}{\theta}e^{\frac{t-\theta}{\theta}} & ,\theta\le t\le 2\theta\\ -1+e^{t/\theta}-\dfrac{t-\theta}{\theta}e^{\frac{t-\theta}{\theta}}+\dfrac{(t-2\theta)^2}{2!\theta^2}e^{\frac{t-2\theta}{\theta}}, & ,2\theta\le t\le 3\theta\\ . & \\ . & \\ . & \\ -1+e^{t/\theta}-\dfrac{t-\theta}{\theta}e^{\frac{t-\theta}{\theta}}+...+(-1)^n\dfrac{(t-n\theta)^2}{2!\theta^2}e^{\frac{t-n\theta}{\theta}} & ,n\theta\le t\le(n+1)\theta\\ . & \\ . & \\ . & \end{cases}$$

$$=-1+\sum_{j=1}^{k}\frac{(-1)^j}{j!}\left(\frac{t-j\theta}{\theta}\right)^j e^{\frac{t-j\theta}{\theta}} \qquad ,k\theta\le t\le(k+1)\theta\ ;\ k=0,1,...$$

dır.

### 4.1.3 Hiper üstel dağılım

$X$ rasgele değişkeninin olasılık yoğunluk fonksiyonu

$$f(x)=\begin{cases}\dfrac{\theta_1\theta_2}{\theta_2-\theta_1}\left(e^{-\theta_1 x}-e^{-\theta_2 x}\right), & x>0;\ \theta_2>\theta_1>0\\ 0 & , \quad d.y\end{cases}$$

olsun. $n=1,2,...$ için $X_n$'ler yukarıda verilen hiper üstel dağılıma sahip rasgele değişkenler olmak üzere $\{N(t),\ t\geq 0\}$ yenileme süreci göz önüne alınsın. Bu sürecin ortalama değer fonksiyonu (3.2) konvolüsyon serisi ya da (3.3) integral denklemi dışında kolaylıkla (3.6) ya da (3.8) den elde edilebilecek Laplace-Stieltjes ya da Laplace dönüşümünün ters dönüşümünün alınmasıyla elde edilebilir. Gerçekten

$$f_L(t)=\frac{\theta_1\theta_2}{\left(t+\theta_1\right)\left(t+\theta_2\right)}$$

olup (3.8) den

$$M_L\left(t\right)=\frac{\theta_1\theta_2}{t^2\left(t+\theta_1+\theta_2\right)}$$

elde edilir. $M_L$ nin ters Laplace dönüşümünün alınmasıyla

$$M(t)=\frac{\theta_1\theta_2}{\left(\theta_1+\theta_2\right)^2}\left(e^{-(\theta_1+\theta_2)t}-1+\left(\theta_1+\theta_2\right)t\right)\ , \quad t\geq 0$$

bulunur.

### 4.1.4 Gamma dağılımı

$\{N(t),\ t\geq 0\}$ yenilemeler arası geçen zaman süreleri $\alpha,\ \beta>0$ parametreli gamma dağılımına sahip bir yenileme süreci olsun. Bu durumda $\left[0,t\right]$ deki her bir yenileme süresinin olasılık yoğunluk ve dağılım fonksiyonu sırasıyla

$$f(x)=\frac{1}{\Gamma(\alpha)\beta^\alpha}x^{\alpha-1}e^{-x/\beta}, \qquad x\geq 0$$

ve

$$F(x)=\frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^x x^{\alpha-1}e^{-x/\beta}dx, \qquad x\geq 0$$

dır. $\alpha>0$ reel sayısının doğal sayı olması durumunda $F$ için kapalı bir formun

$$F(x) = 1 - \sum_{k=0}^{\alpha-1} \frac{(x/\beta)^k}{k!} e^{-x/\beta}, \qquad x \geq 0$$

olduğu bilinmektedir. $n \geq 1$ için $S_n$ rasgele değişkeni $n\alpha$ ve $\beta$ parametreli gamma dağılımına sahip olduğundan

$$F^{n^*}(t) = 1 - \sum_{k=0}^{n\alpha-1} \frac{(t/\beta)^k}{k!} e^{-t/\beta}$$

ve

$$P\left(N(t) = n\right) = \sum_{k=n\alpha}^{(n+1)\alpha-1} \frac{(t/\beta)^k}{k!} e^{-t/\beta}$$

dır.

Şimdi $N(t)$ nin olasılık üreten fonksiyonu yardımıyla $M(t)$' nin bir kapalı form ifadesini elde edelim. $N(t)$' nin olasılık üreten fonksiyonunu $\psi$ ile gösterelim.

$$G(z,t) = \sum_{n=1}^{\infty} z^{n-1} F^{n^*}(t)$$

tanımlanmak üzere

$$\begin{aligned} \psi(z,t) &= \sum_{n=0}^{\infty} z^n P(N(t) = n) \\ &= 1 + (z-1)G(z,t) \end{aligned}$$

dir. $y^\alpha = z$ dönüşümü yapılmasıyla

$$\begin{aligned} G(z,t) &= \sum_{n=1}^{\infty} z^{n-1} F^{n^*}(t) \\ &= \sum_{n=1}^{\infty} z^{n-1} \left[ \int_0^t \frac{(x/\beta)^{n\alpha-1}}{\beta(n\alpha-1)!} e^{-x/\beta} dx \right] \\ &= \int_0^t \sum_{n=1}^{\infty} \frac{z^{n-1} (x/\beta)^{n\alpha-1}}{\beta(n\alpha-1)!} e^{-x/\beta} dx \\ &= y^{1-\alpha} \int_0^t \frac{e^{-x/\beta}}{\beta} \left( \sum_{n=1}^{\infty} \left( \frac{xy}{\beta} \right)^{n\alpha-1} / (n\alpha-1)! \right) dx \end{aligned}$$

bulunur. Herhangi bir $u$ reel sayısı ve $\alpha$ tamsayısı için $i^2 = -1$ olmak üzere $c = e^{2\pi i/\alpha}$ alındığında

$$\sum_{n=1}^{\infty} \frac{u^{n\alpha-1}}{(n\alpha-1)!} = \frac{1}{\alpha} \sum_{r=0}^{\alpha-1} c^r e^{uc^r}$$

olduğu bilinmektedir (Parzen 1962). Böylece

$$\begin{aligned} G(z,t) &= y^{1-\alpha} \frac{1}{\alpha} \int_0^t \frac{e^{-x/\beta}}{\beta} \left( \sum_{r=0}^{\alpha-1} c^r e^{xyc^r/\beta} \right) dx \\ &= y^{1-\alpha} \frac{1}{\alpha\beta} \sum_{r=0}^{\alpha-1} c^r \int_0^t e^{x(yc^r-1)/\beta} dx \\ &= \frac{1}{\alpha z} \sum_{r=0}^{\alpha-1} \frac{z^{1/\alpha} c^r}{1 - z^{1/\alpha} c^r} \left( 1 - e^{-t\left(1-z^{1/\alpha} c^r\right)/\beta} \right) \end{aligned}$$

bulunduğundan $N(t)$ 'in olasılık üreten fonksiyonu

$$\begin{aligned} P(z,t) &= 1 + (z-1) G(z,t) \\ &= 1 + \frac{z-1}{\alpha z} \frac{z^{1/\alpha}}{1 - z^{1/\alpha}} \left( 1 - e^{-t(1-z^{1/\alpha})/\beta} \right) + \frac{z-1}{\alpha z} \sum_{r=1}^{\alpha-1} \frac{z^{1/\alpha} c^r}{1 - z^{1/\alpha} c^r} \left( 1 - e^{-t(1-z^{1/\alpha} c^r)/\beta} \right) \end{aligned}$$

olarak bulunur. Buradan

$$h(z,t) = \frac{z-1}{\alpha} \frac{z^{1/\alpha}}{1 - z^{1/\alpha}} (1 - e^{-t(1-z^{1/\alpha})/\beta})$$

ve

$$g(z,t) = \sum_{r=1}^{\alpha-1} \frac{z^{1/\alpha} c^r}{1 - z^{1/\alpha} c^r} \left( 1 - e^{-t(1-z^{1/\alpha} c^r)/\beta} \right)$$

gösterimiyle

$$\frac{d\psi(z,t)}{dz} = \frac{dh(z,t)}{dz} + \frac{1}{\alpha} \left( \frac{1}{z^2} g(z,t) + \frac{z-1}{z} \frac{dg(z,t)}{dz} \right)$$

yazılabilir.

$$g(z,t)|_{z=1} = \sum_{r=1}^{\alpha-1} \frac{c^r}{1 - c^r} \left( 1 - e^{-t(1-c^r)/\beta} \right)$$

$$\frac{dg(z,t)}{dz}|_{z=1} = \sum_{r=1}^{\alpha-1} \left[ \frac{1}{\alpha} \frac{c^r}{1 - c^r} \left( 1 + \frac{c^r}{1 - c^r} \right) \left( 1 - e^{-t(1-c^r)/\beta} \right) - \frac{t}{\alpha\beta} e^{-t(1-c^r)/\beta} \right]$$

$$\frac{dh(z,t)}{dz}|_{z=1} = \frac{1}{\alpha} \frac{d}{dz} \left( \frac{z^{1/\alpha} - z^{1/\alpha-1}}{1 - z^{1/\alpha}} \right) \left( 1 - e^{-t(1-z^{1/\alpha})/\beta} \right) + \frac{1}{\alpha} \frac{d}{dz} \left( 1 - e^{-t(1-z^{1/\alpha})/\beta} \right) \left( \frac{z^{1/\alpha} - z^{1/\alpha-1}}{1 - z^{1/\alpha}} \right)$$

dır. $z \to 1$ için

$$\left(1-e^{-t(1-z^{1/\alpha})/\beta}\right)\to 0, \qquad \frac{d}{dz}\left(1-e^{-t(1-z^{1/\alpha})/\beta}\right)\to -\frac{t}{\alpha\beta}$$

olduğundan

$$\lim_{z\to 1}\frac{dh(z,t)}{dz}=\frac{t}{\alpha\beta}$$

bulunur. Bu durumda

$$\lim_{z\to 1}\frac{dP(z,t)}{dz}=\frac{t}{\alpha\beta}+\frac{1}{\alpha}\sum_{r=1}^{\alpha-1}\frac{c^r}{1-c^r}\left(1-e^{-t(1-c^r)/\beta}\right)$$

dır. Böylece bu sürecin yenileme fonksiyonu $\alpha$ bir doğal sayı olmak üzere,

$$M(t)=\frac{t}{\alpha\beta}+\frac{1}{\alpha}\sum_{r=1}^{\alpha-1}\frac{c^r}{1-c^r}\left(1-e^{-t(1-c^r)/\beta}\right) \tag{4.1}$$

ile kapalı bir formda elde edilir.

$\alpha=1,\ \alpha=2,\ \alpha=3\ ve\ \alpha=4$ için sırasıyla (4.1) den

$$M(t)=\frac{t}{\beta}$$

$$M(t)=\frac{t}{2\beta}-\frac{1}{4}+\frac{1}{4}e^{-2t/\beta}$$

$$M(t)=\frac{t}{3\beta}-\frac{1}{3}+\frac{1}{3}e^{-3t/2\beta}(\cos(\frac{\sqrt{3}t}{2\beta})+\frac{1}{\sqrt{3}}\sin(\frac{\sqrt{3}t}{2\beta}))$$

$$M(t)=\frac{t}{4\beta}-\frac{3}{8}+\frac{1}{8}e^{-2t/\beta}+\frac{1}{4}e^{-t/\beta}(\cos(\frac{t}{\beta})+\sin(\frac{t}{\beta}))$$

ifadeleri ile bulunur.

## 4.2 M(t) nin RS Yöntemi ile Sayısal Hesabı

Xie (1989), $M$ yenileme fonksiyonunun (3.4) integral denkleminden sayısal olarak çözümü için RS(Rieman-Stieltjes) olarak adlandırılan bir yöntem vermiştir. RS yöntem kolay olarak programlanabilir, hemen hemen tüm durumlarda basitliği ve yakınsaklığı ile iyi sonuçlar verir ve diğer bilinen yöntemlerle karşılaştırıldığında uygulanabilirliği daha fazladır. Bu yöntem özellikle $f$ olasılık yoğunluk fonksiyonu bilinmediği ve aykırı noktalara sahip iken faydalıdır. Burada, $M(t)$ yenileme fonksiyonunun sayısal hesabı için verilen RS yöntemi üzerinde durulmaktadır.

Riemann-Stieltjes integralin tanımının ışığı altında $\int_a^b g(x)dh(x)$ integrali, $[a,b]$ aralığının bir parçalanması $\Delta = \{x_0, x_1, ..., x_n\}$ olmak üzere,

$$\int_a^b g(x)dh(x) \approx \sum_{i=1}^{n} g((x_i + x_{i-1})/2)(h(x_i) - h(x_{i-1})) \tag{4.2}$$

şeklinde yazılabilir. Riemann-Stieltjes integralinin sayısal hesaplanabilmesi için kullanılabilecek bu formülde parçalanmanın normu küçüldükçe yaklaşımın daha iyi olacağı açıktır.

Şimdi (3.4) denklemini, yani

$$M(t_i) = F(t) + \int_o^{t_i} F(t-x)dM(x), \qquad t \geq 0$$

integral denklemini göz önüne alalım. $t$ verilmiş bir değer ve $\{t_0, t_1, ..., t_n\}$ $0 = t_0 < t_1 < ... < t_n = t$ şartını sağlayan bir parçalanma olsun. Bu durumda

$$M(t) = F(t_i) + \int_o^{t} F(t_i - x)dM(x), \qquad t \geq 0$$

olup (4.2) ifadesinin kullanılmasıyla

$$M(t_i) \approx F(t_i) + \sum_{j=1}^{i} F(t_i - (t_j + t_{j-1})/2)(M(t_j) - M(t_{j-1}))$$

bulunur. Böylece $T_i = \sum_{j=1}^{i-1} F(t_i - (t_j + t_{j-1})/2)(M(t_j) - M(t_{j-1}))$ alındığında $M(t_i)$ ardışık olarak, $\tilde{M}(t_0) = 0$ olmak üzere

$$\tilde{M}(t_i) = \frac{F(t_i) + T_i - F(t_i - (t_i + t_{i-1})/2)\,\tilde{M}(t_{j-1})}{1 - F(t_i - (t_i + t_{i-1})/2}, \qquad i = 1, 2, ..., n$$

ile yaklaşık olarak hesaplanabilir (Xie 1989a).

$F$ dağılım fonksiyonu yerine $f$ olasılık yoğunluk fonksiyonu verildiğinde ve $F$'nin kapalı bir formda ifadesi yok ise $F$ dağılım fonksiyonu

$$F(t_i) = F(t_{i-1}) + f((t_i + t_{i-1})/2)\frac{t}{n}, \qquad i = 1, 2, ..., n$$

formülü ile kolaylıkla hesaplanabilir.

(3.3) ve (3.4) denklemleri teorik olarak denktirler. Literatürde en yaygın olanı (3.3) denklemi olmasına rağmen RS yöntemi (3.4) denkleminin çözümüne kısıtlanmıştır. Eşit olmayan adım uzunlukları kullanılırsa (3.4) denklemi ardışık çözüm için daha basit görünür. Aynı zamanda (3.3) yerine (3.4)'ün kullanılmasının temel avantajı $F(t)$'nin büyük $t$'ler için hemen hemen sabit olmasıdır ve böylece $F(t_i) - F(t_{i-1})$ deki yuvarlatma hatası nispeten büyük olacaktır.

## 5. SANSÜRLÜ VERİLER

Sansürleme, zaman ve maliyet gibi bir takım sınırlamalar nedeniyle, örneklemdeki birimlerden elde edilen gözlemlerin analize dahil edilememesi veya elde edilemeyen bilgilerin göz ardı edilmesidir (Topçu 2007). Örneğin, sabit ya da rasgele takip süreli çalışmalarda, belirlenen süre sonunda ilgilenilen olayın (bozulma gibi) hala gerçekleşmemiş olmasıdır. Bu tür durumlarda karşımıza çıkan gözlemlere 'sansürlü veriler' adı verilir

Literatürde farklı sansürleme çeşitleri verilmiştir (Lawless 2003). Bunlardan bazıları aşağıdadır:

- Sağdan Sansürleme
    - I. Tip Sansürleme (zaman sansürleme)
    - Rasgele Sansürleme (zaman rasgele sansürleme)
    - II. Tip Sansürleme (parça sansürleme)
- Soldan Sansürleme
- İkili Sansürleme
- Aralık Sansürleme

Bu çalışmada yukarıdaki sansürleme çeşitlerinden yalnızca sağdan rasgele sansürleme ile ilgilenmekteyiz. Bu sansürleme durumu aşağıdaki kısımda ayrıntılı bir şekilde ele alınmıştır.

### 5.1 Sağdan Rasgele Sansürleme

$i = 1, 2, ..., n$ için $Y_i$ i.parçanın ömrünü ve $T_i$ i.parça ile ilgili sansürleme rasgele değişkenini göstersin. Burada hem $Y_i$ ($i = 1, 2, ..., n$)'ler hem de $T_i$ ($i = 1, 2, ..., n$)'ler bağımsız ve aynı dağılımlı rasgele değişkenlerdir. Ayrıca $Y_i$ ($i = 1, 2, ..., n$) ve $T_i$ ($i = 1, 2, ..., n$) rasgele değişkenleri bağımsızdır.

$i=1,2,...,n$ olmak üzere i.bileşenin $T_i$ rasgele değişkeni ile sansürlenmesi altında *i.* bileşen için sansürlü gözlemimiz $X_i = \min(Y_i, T_i)$ dir. Bu şekilde oluşturulan $X_1, \ldots, X_n$ rasgele değişkenlerine 'sağdan rasgele sansürlenmiş $n$ birimlik örneklem' denir.

Yenileme süreçleri ile ilgili uygulamalarda çoğu kez bu tip gözlemler ile karşılaşılır. Örneğin, yeni çıkan bir ürün farklı zamanlarda satılır ve üretici bu satılan ürünlerin bozulma sürelerini 12 ay gözler ise, satılan ürünlerin ömürleri karşımıza sağdan rasgele sansürlenmiş veriler olarak çıkar.

$i=1,2,...,n$ için $Y_i$ rasgele değişkeninin dağılım fonksiyonu $F$ ve olasılık yoğunluk fonksiyonu $f$ ile gösterelim. $F$ şekilsel olarak bilinirken bazı parametreleri bilinmesin. $\theta_1, \theta_2, ..., \theta_r$ $F$ 'in bilinmeyen parametrelerini göstersin. Bu durumda $X_1, \ldots, X_n$ sağdan rasgele sansürlü örneklemine dayalı olarak $\theta_1, \theta_2, ..., \theta_r$ parametrelerinin tahminleri yardımıyla $F$ 'nın bir noktadaki tahminini yapabiliriz.

## 5.2 Sağdan Rasgele Sansürlü Örneklem Durumunda Parametre Tahmini

Bu kısımda $F$ fonksiyonel olarak bilinirken, bilinmeyen $\theta_1, \theta_2, ..., \theta_r$ parametrelerinin sağdan rasgele sansürlenmiş örneklemine dayalı olarak ençok olabilirlik tahmin edicilerinin elde edilmesi üzerinde durulacaktır.

$T_i$ $(i=1,2,...,n)$ rasgele değişkeninin dağılım fonksiyonu $G$ ve olasılık yoğunluk fonksiyonunu $g$ ile gösterelim. $i=1,2,...,n$ için,

$$I_i = \begin{cases} 1, & Y_i \le T_i \\ 0, & Y_i > T_i \end{cases}$$

olmak üzere

$$\begin{aligned} X_i &= \min(Y_i, T_i) \\ &= Y_i I_i + T_i (1 - I_i) \end{aligned}$$

yazılabilir. ($I_i = 1$) olayı $i$. bileşen için sansürlemenin yapılmadığını gösterirken ($I_i = 0$) olayı sansürlemenin yapıldığını gösterir. Şimdi ilk olarak sağdan rasgele sansürlü $X_1,\ldots,X_n$ örneklemi durumunda olabilirlik fonksiyonunun oluşturulması için $X_i$ ve $I_i$ $(i = 1,2,...,n)$ rasgele değişkenlerinin ortak dağılımlarını bulalım.

$$\begin{aligned} X_i = x_i, I_i = 1 &\Leftrightarrow X_i = x_i \quad , \; Y_i \le T_i \\ &\Leftrightarrow X_i = Y_i = x_i, \; X_i \le T_i \\ &\Leftrightarrow Y_i = x_i \quad , \; x_i \le T_i \end{aligned}$$

olduğundan

$$f_{X_i,I_i}(x_i,1) = f(x_i)(1 - G(x_i^-))$$

ve

$$\begin{aligned} X_i = x_i, I_i = 0 &\Leftrightarrow X_i = x_i \quad , \; Y_i > T_i \\ &\Leftrightarrow X_i = x_i \quad , \; Y_i > T_i \quad , X_i = T_i \\ &\Leftrightarrow Y_i = x_i \quad , \; Y_i > x_i \end{aligned}$$

olduğundan

$$f_{X_i,I_i}(x_i,0) = g(x_i)(1 - F(x_i))$$

dir. Böylece $\delta_i \in \{0,1\}$ için

$$f_{X_i,I_i}(x_i,\delta_i) = \left[f(x_i)\right]^{\delta_i}\left[1 - F(x_i)\right]^{1-\delta_i}\left[f_{T_i}(x_i)\right]^{1-\delta_i}\left[1 - F_{T_i}(x_i)\right]^{\delta_i}$$

yazılabilir. Bu durumda sağdan rasgele sansürlü $X_1,\ldots,X_n$ örneklemi için $L(\theta_1,\theta_2,...\theta_r)$ olabilirlik fonksiyonu aşağıdaki gibi tanımlanır (Binshan, 1988).

$$L(\theta_1,\theta_2,...\theta_r) = \prod_{i=1}^{n}[f(x_i)]^{\delta_i}\left[1 - F(X_i)\right]^{1-\delta_i}\prod_{i=1}^{n}[f_T(x_i)]^{1-\delta_i}\left[1 - F(x_i^-)\right]^{\delta_i}$$

$$= \prod_{i=1}^{n}[f(x_i)]^{\delta_i}\left[1 - F(X_i)\right]^{1-\delta_i}\prod_{i=1}^{n}[g(x_i)]^{1-\delta_i}\left[1 - G(x_i)\right]^{\delta_i}$$

olur. Bu $L$ fonksiyonunu maksimum yapan $\theta_1,\theta_2,...,\theta_r$ değerlerine parametrelerin en çok olabilirlik tahminleri denir ve bu tahminler sırasıyla $\hat{\theta}_1,\hat{\theta}_2...,\hat{\theta}_r$ ile gösterilir. $G$ dağılımı $\theta_1,\theta_2,...,\theta_r$ parametrelerini içermez ise olabilirlik fonksiyonundaki ikinci çarpım en çok olabilirlik tahmin edicilerin bulunma işlemini etkilemeyecektir

En çok olabilirlik yöntemi ile uygun bir şekilde çalışmak için gözlemlerin kümesini $D$ ve $C$ gibi iki alt kümeye ayıralım. $D$ gözlenmiş bozulma zamanlarını ve $C$ sansürlenmiş gözlemlerin indislerinin kümesini göstersin. $D$ kümesinin boş küme olmadığı varsayımı altında $C$ ve $D$ kümeleri yardımıyla olabilirlik fonksiyonu,

$$L(\theta_1,...,\theta_r) = \prod_{i \in D} f(x_i) \prod_{i \in C} (1 - F(x_i)) \prod_{i \in C} g(x_i) \prod_{i \in D} (1 - G(x_i))$$

biçiminde yazılabilir. Çoğu durumda $\hat{\theta}_1, \hat{\theta}_2 ..., \hat{\theta}_r$ en çok olabilirlik tahmin edicileri

$$\frac{\partial \log L(\theta)}{\partial \theta_i} = 0, \quad i = 1, 2, ..., r$$

denklem sisteminin çözümüyle elde edilir. Yaygın kullanılan birkaç $F$ dağılımı dışında bu denklem sisteminin analitik çözümü yoktur. Bu durumda sayısal bir yöntem ile çözümün gerekliliği ortaya çıkmaktadır. Literatürde en çok kullanılan denklem sistemini çözme yöntemlerinden birisi Newton-Raphson yöntemidir. Bu çalışmada ele alınacak dağılımların parametrelerinin tahminlerini elde etmek için aşağıda tarif edilen bu yöntem kullanılmıştır.

### 5.2.1 Newton-Rapson Yöntemi

$i = 1, 2, ..., r$ için $\theta_i^0$ $\theta_i$ parametresinin gerçek değeri,

$$U_i(\theta_1, \ldots, \theta_r) = \frac{\partial \log L(\theta_1, \ldots, \theta_r)}{\partial \theta_i} \quad i = 1, 2, ..., r$$

ve

$$V(\theta_1, \ldots, \theta_r) = \begin{bmatrix} \frac{\partial^2 \log L(\theta_1, \ldots, \theta_r)}{\partial \theta_1^2} & \frac{\partial^2 \log L(\theta_1, \ldots, \theta_r)}{\partial \theta_1 \partial \theta_2} & \cdots & \frac{\partial^2 \log L(\theta_1, \ldots, \theta_r)}{\partial \theta_1 \partial \theta_r} \\ \frac{\partial^2 \log L(\theta_1, \ldots, \theta_r)}{\partial \theta_2 \partial \theta_1} & \frac{\partial^2 \log L(\theta_1, \ldots, \theta_r)}{\partial \theta_2^2} & \cdots & \frac{\partial^2 \log L(\theta_1, \ldots, \theta_r)}{\partial \theta_2 \partial \theta_r} \\ . & & & \\ . & & & \\ . & & & \\ \frac{\partial^2 \log L(\theta_1, \ldots, \theta_r)}{\partial \theta_r \partial \theta_1} & \frac{\partial^2 \log L(\theta_1, \ldots, \theta_r)}{\partial \theta_r \partial \theta_2} & \cdots & \frac{\partial^2 \log L(\theta_1, \ldots, \theta_r)}{\partial \theta_r^2} \end{bmatrix}_{rxr}$$

olsun.

$$U(\theta_1,\ldots,\theta_r) = \begin{bmatrix} U_1(\theta_1,\ldots,\theta_r) \\ . \\ . \\ . \\ U_r(\theta_1,\ldots,\theta_r) \end{bmatrix}$$ diyelim. $i = 1,2,...,r$ için $\theta_i^*$ $\theta_i$ ile $\theta_i^0$ arasında

olmak üzere $U(\theta_1,\ldots,\theta_r)$ 'nın $(\theta_1^0,...,\theta_r^0)$ etrafında Taylor serisine açılması ile denklem sistemi

$$U(\theta_1,\ldots,\theta_r) = U(\theta_1^0,...,\theta_r^0) + V(\theta_1^0,...,\theta_r^0)\begin{bmatrix} \theta_1^* - \theta_1^0 \\ . \\ . \\ . \\ \theta_r^* - \theta_r^0 \end{bmatrix} \qquad (5.1)$$

yazılabilir (Gertsbakh 1989). (5.1)'de $\theta_i = \hat{\theta}_i$ $(i = 1,2,...,r)$ alınsın, burada $\hat{\theta}_i$ $\theta_i$ 'nin en çok olabilirlik tahmin edicisidir. Bu durumda $U(\hat{\theta}_1,\ldots,\hat{\theta}_r) = 0$ olup (5.1) den

$$\begin{bmatrix} \theta_1^* \\ . \\ . \\ . \\ \theta_r^* \end{bmatrix} = \begin{bmatrix} \theta_1^0 \\ . \\ . \\ . \\ \theta_r^0 \end{bmatrix} - V^{-1}(\theta_1^0,...,\theta_r^0)U(\theta_1^0,...,\theta_r^0)$$

bulunur. Yukarıdaki ifade yardımıyla $i = 1,2,...,r$ için $\theta_i^{(1)}$ $\hat{\theta}_i$'nin başlangıç değeri olmak üzere,

$$\begin{bmatrix} \theta_1^{(m+1)} \\ . \\ . \\ . \\ \theta_r^{(m+1)} \end{bmatrix} = \begin{bmatrix} \theta_1^{(m)} \\ . \\ . \\ . \\ \theta_r^{(m)} \end{bmatrix} - V^{-1}(\theta_1^{(m)},...,\theta_r^{(m)})U(\theta_1^{(m)},...,\theta_r^{(m)}), \qquad m = 1,2,...$$

iterasyon işlemleri göz önüne alınır. $i = 1,2,...,r$ için $\theta_i^{(m)}$, $\theta_i^{(m+1)}$ sayısına yeterince yakınsa iterasyon durdurulur. $m \to \infty$ iken $i = 1,2,...,r$ için $\theta_i^{(m)} \to \hat{\theta}_i$, $i = 1,2,...,r$ dir.

Aşağıda bu çalışmada kullanılmakta olan bazı dağılımların bilinmeyen parametreleri için sağdan rasgele sansürlü örnekleme durumunda en çok olabilirlik tahmin edicilerinin elde edilmesi üzerinde durulmaktadır.

### 5.2.2 Üstel dağılım

$F$ , $\theta$ parametreli üstel dağılım fonksiyonu olsun; yani $F(x) = 1 - e^{-x/\theta}, \quad x \geq 0,\ \theta > 0$ olasılık yoğunluk fonksiyonuda $f(x) = \frac{1}{\theta} e^{-x/\theta}, \quad x \geq 0,\ \theta > 0$ dır. Sağdan rasgele sansürlenmiş $X_i = \min(Y_i, T_i) \quad (i = 1, 2, ..., n)$ örneklemini göz önüne alalım. Burada $Y_i$'lerin dağılımının $\theta$ parametreli üstel ve $T_i$' lerin dağılımının bir $G$ dağılımı olduğunu hatırlatalım. Kabul edelim ki $G$ dağılımı $\theta$ parametresini içermesin. Bu durumda $L(\theta)$ olabilirlik fonksiyonu

$$\begin{aligned} L(\theta) &= \prod_{i \in D} f(x_i) \prod_{i \in C} \left[1 - F(x_i)\right] \\ &= \prod_{i \in D} \frac{1}{\theta} e^{-x/\theta} \prod_{i \in C} \left[e^{-x/\theta}\right] \end{aligned}$$

olup

$$\begin{aligned} \ln L(\theta) &= \sum_{i \in D} \ln \frac{1}{\theta} - \sum_{i \in D} \frac{1}{\theta} x_i - \sum_{i \in C} \frac{1}{\theta} x_i \\ &= S(D) \ln \frac{1}{\theta} - \sum_{i=1}^{n} \frac{1}{\theta} x_i \end{aligned}$$

olur. Bu durumda

$$\frac{d \ln L(\theta)}{d\theta} = -\frac{S(D)}{\theta} + \sum_{i=1}^{n} \frac{1}{\theta^2} x_i = 0$$

denkleminin çözümünden $\theta$'nın en çok olabilirlik tahmin edicisi

$$\hat{\theta} = \frac{\sum_{i=1}^{n} X_i}{S(D)} \tag{5.2}$$

olarak elde edilir.

### 5.2.3 Weibull dağılım

$F$ , $\alpha$ şekil ve $\beta$ ölçek parametreli Weibull dağılım fonksiyonu olsun,

$$F(x) = 1 - e^{-(x/\beta)^{\alpha}} \qquad , x > 0$$

olup bu dağılımın olasılık yoğunluk fonksiyonu,

$$f(x) = \frac{\alpha}{\beta^{\alpha}} x^{\alpha-1} e^{-(x/\beta)^{\alpha}} \qquad , \quad x > 0; \quad \alpha, \beta > 0$$

dir.

$X_1, \ldots, X_n$; $F$ dağılımından sağdan rasgele sansürlenmiş $n$ birimlik örneklem olmak üzere, sansürleme rasgele değişkeninin dağılım fonksiyonu olan $G$'nin $\alpha$ ve $\beta$ parametrelerini içermediğini kabul edelim. Bu durumda

$$L(\alpha, \beta) = \prod_{i \in D} \left( \frac{\alpha}{\beta^{\alpha}} x_i^{\alpha-1} e^{-(x_i/\beta)^{\alpha}} \right) \prod_{i \in C} \left( e^{-(x_i/\beta)^{\alpha}} \right)$$

ve

$$\begin{aligned}
\ln L(\alpha, \beta) &= \sum_{i \in D} \ln \left( \frac{\alpha}{\beta^{\alpha}} x_i^{\alpha-1} e^{-(x_i/\beta)^{\alpha}} \right) \sum_{i \in C} \ln \left( e^{-(x_i/\beta)^{\alpha}} \right) \\
&= \sum_{i \in D} \left( \ln \alpha - \alpha \ln \beta + (\alpha - 1) \ln x_i - \frac{x_i^{\alpha}}{\beta^{\alpha}} \right) - \sum_{i \in C} \frac{x_i^{\alpha}}{\beta^{\alpha}} \\
&= S(D) \ln \alpha - S(D) \alpha \ln \beta + (\alpha - 1) \sum_{i \in D} \ln x_i - \frac{1}{\beta^{\alpha}} \sum_{i \in D} x_i^{\alpha} - \frac{1}{\beta^{\alpha}} \sum_{i \in C} x_i^{\alpha} \\
&= S(D) \ln \alpha - S(D) \alpha \ln \beta + (\alpha - 1) \sum_{i \in D} \ln x_i - \frac{1}{\beta^{\alpha}} \sum_{i=1}^{n} x_i^{\alpha}
\end{aligned}$$

olup

$$\frac{\partial \ln L(\alpha, \beta)}{\partial \alpha} = \frac{S(D)}{\alpha} - S(D) \ln \beta + \sum_{i \in D} \ln x_i - \frac{1}{\beta^{\alpha}} \sum_{i=1}^{n} x_i^{\alpha} \ln x_i + \frac{\ln \beta}{\beta^{\alpha}} \sum_{i=1}^{n} x_i^{\alpha} = 0 \quad (5.3)$$

ve

$$\frac{\partial \ln L(\alpha, \beta)}{\partial \beta} = -\frac{S(D)\alpha}{\beta} + \frac{\alpha}{\beta^{\alpha+1}} \sum_{i=1}^{n} x_i^{\alpha} = 0 \quad (5.4)$$

denklemleri elde edilir. (5.4) denkleminden

$$\beta = \left( \frac{\sum_{i=1}^{n} x_i^{\alpha}}{S(D)} \right)^{1/\alpha} \tag{5.5}$$

olur. Bu ifadenin (5.3) de yerine konulmasıyla, $\alpha$ ‘ya göre

$$\frac{1}{\alpha} + \frac{\sum_{i=1}^{n} \ln x_i}{S(D)} - \frac{\sum_{i=1}^{n} x_i^{\alpha} \ln x_i}{\sum_{i=1}^{n} x_i^{\alpha}} = 0 \tag{5.6}$$

lineer olmayan denklemi elde edilir. Bu denklemi Matlab paket programında Newton Raphson yöntemi ile çözülebilir. Burada Newton Raphson yönteminin verilişindeki notasyonlara bağlı kalınmak üzere,

$$U(\alpha) = \frac{1}{\alpha} + \frac{\sum_{i=1}^{n} \ln x_i}{S(D)} - \frac{\sum_{i=1}^{n} x_i^{\alpha} \ln x_i}{\sum_{i=1}^{n} x_i^{\alpha}}$$

ve

$$\begin{aligned} V(\alpha) &= -\frac{1}{\alpha^2} - \frac{\left( \sum_{i=1}^{n} x_i^{\alpha} \left(\ln x_i\right)^2 \sum_{i=1}^{n} x_i^{\alpha} - \left( \sum_{i=1}^{n} x_i^{\alpha} \ln x_i \right)^2 \right)}{\left( \sum_{i=1}^{n} x_i^{\alpha} \right)^2} \\ &= -\frac{1}{\alpha^2} - \frac{\sum_{i=1}^{n} x_i^{\alpha} \left(\ln x_i\right)^2}{\sum_{i=1}^{n} x_i^{\alpha}} + \left( \frac{\sum_{i=1}^{n} x_i^{\alpha} \ln x_i}{\sum_{i=1}^{n} x_i^{\alpha}} \right)^2 \end{aligned}$$

dir. Bu durumda Newton Raphson yönteminden $\alpha(1)$ başlangıç değeri ile,

$$\alpha(m+1)=\alpha(m)-\frac{U(\alpha(m))}{V(\alpha(m))}$$

$$=\alpha(m)-\frac{\dfrac{1}{\alpha(m)}+\dfrac{\sum_{i=1}^{n}\ln x_i}{S(D)}-\dfrac{\sum_{i=1}^{n}x_i^{\alpha(m)}\ln x_i}{\sum_{i=1}^{n}x_i^{\alpha(m)}}}{-\dfrac{1}{\alpha(m)^2}-\dfrac{\sum_{i=1}^{n}x_i^{\alpha(m)}\left(\ln x_i\right)^2}{\sum_{i=1}^{n}x_i^{\alpha(m)}}+\left(\dfrac{\sum_{i=1}^{n}x_i^{\alpha(m)}\ln x_i}{\sum_{i=1}^{n}x_i^{\alpha(m)}}\right)^2}, \qquad m=1,2,... \tag{5.7}$$

iterasyon işlemlerine ulaşılır. $\alpha\left(m+1\right)$ yeterince $\alpha\left(m\right)$ sayısına yeterince yakınsa iterasyon durdurularak $\alpha$ 'nın en çok olabilirlik tahmin değeri olan $\hat{\alpha}=\alpha(m+1)$ olarak bulunmuş olur.

**Örnek 5.2.3.1** $\alpha$ ve $\beta$ parametreli Weibull dağılımı için sağdan rasgele sansürlenmiş gözlem değerleri

$x_1=7.92$ $x_2=4.55$ $x_3=6.77$ $x_4=6.77$ $x_5^*=7.51$ $x_6^*=7.08$ $x_7^*=6.59$ $x_8^*=6.08$
$x_9^*=5.15$ $x_{10}^*=1.95$ $x_{11}^*=4.08$ $x_{12}^*=7.45$ $x_{13}^*=5.23$ $x_{14}^*=5.23$

olarak verilsin. Burada * ile belirtilen sayılar sansürlenmiş gözlem değerleridir. (5.6) deki iterasyon işlemleri aşağıda EK 1'de verilen Matlab programında yaptırılmasıyla $\alpha$ 'nın en çok olabilirlik tahmin değeri $\hat{\alpha}=7.9272$ olarak bulunur. (5.4) denkleminden de $\beta$ 'nın en çok olabilirlik tahmin değeri $\hat{\beta}=7.8533$ olur.

### 5.2.4 Log-normal dağılım

$F$, $\mu$ ve $\sigma$ parametreli log-normal dağılım fonksiyonu olsun. $\Phi$ standart normal dağılımın dağılım fonksiyonu olmak üzere,

$$F(x)=\Phi\left(\frac{\log x-\mu}{\sigma}\right) \quad , \qquad x>0\ ;\ \mu\in\mathbb{R}\ ,\sigma>0$$

dir. Ayrıca bu dağılımın f olasılık yoğunluk fonksiyonu,

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma\, x} e^{-\frac{1}{2}\left(\frac{\log x - \mu}{\sigma}\right)^2}, \quad x > 0 \; ; \mu \in \mathbb{R} \; , \sigma > 0$$

dır.

$X_1, \ldots, X_n$ yukarıdaki $F$ dağılımından sağdan rasgele sansürlenmiş $n$ birimlik bir örneklem olsun. Sansürleme rasgele değişkeninin dağılım fonksiyonu olan $G$ 'nin $\mu$ ve $\sigma$ parametrelerini içermediğini kabul edelim. Bu durumda,

$$y_i = \frac{\log x_i - \mu}{\sigma} \quad , \qquad i = 1, 2, ..., n$$

olmak üzere,

$$\begin{aligned}
\ln L(\mu,\sigma) &= \sum_{i\in D} \ln f(x_i) + \sum_{i\in C} \ln\left(1 - F(x_i)\right) \\
&= \sum_{i\in D} \ln \frac{1}{\sqrt{2\pi}\,\sigma\, x_i} e^{-\frac{1}{2}\left(\frac{\log x - \mu}{\sigma}\right)^2} + \sum_{i\in C} \ln\left[1 - \Phi(y_i)\right] \\
&= \sum_{i\in D} \left[ -\frac{y_i^2}{2} - \ln\sigma - \ln\sqrt{2\pi} \right] + \sum_{i\in C} \ln\left[1 - \Phi(y_i)\right] \\
&= -\frac{1}{2}\sum_{i\in D} y_i^2 - S(D)\log\sigma - S(D)\ln\sqrt{2\pi} + \sum_{i\in C} \ln\left[1 - \Phi(y_i)\right]
\end{aligned}$$

olur. Bu dağılımın bilinmeyen $\mu$ ve $\sigma$ parametrelerinin en çok olabilirlik tahmin edicilerini bulabilmek için yukarıdaki denklemin $\mu$ ve $\sigma$ 'ya göre türevleri alınarak aşağıdaki lineer olmayan denklemlere ulaşılır.

$$\frac{\partial \ln L(\mu,\sigma)}{\partial \mu} = \frac{1}{\sigma}\sum_{i\in D} y_i + \frac{1}{\sigma}\sum_{i\in C} \frac{\varphi(y_i)}{1 - \phi(y_i)} = 0$$

ve

$$\frac{\partial \ln L(\mu,\sigma)}{\partial \sigma} = -\frac{S(D)}{\sigma} + \frac{1}{\sigma}\sum_{i\in D} y_i^2 + \frac{1}{\sigma}\sum_{i\in C} \frac{y_i\varphi(y_i)}{1 - \phi(y_i)} = 0$$

Burada $\varphi$ standart normal dağılımın olasılık yoğunluk fonksiyonudur. Bu iki denklemin ortak analitik çözümü yoktur. $\mu$ ve $\sigma$ 'nın en çok olabilirlik tahminleri Newton

Raphson yöntemi ile hesaplanabilir. Bu yöntemin çalıştırılmasında gerekli matematiksel ifadeler aşağıdaki gibi çıkartılır.

$$\frac{\partial^2 \ln L(\mu,\sigma)}{\partial\mu\,\partial\sigma} = -\frac{1}{\sigma^2}\sum_{i\in D} y_i + \frac{1}{\sigma}\sum_{i\in D}\frac{-y_i}{\sigma} - \frac{1}{\sigma^2}\sum_{i\in C}\frac{\varphi(y_i)}{1-\phi(y_i)} + \frac{1}{\sigma}\sum_{i\in C}\frac{\left(1-\phi(y_i)\right)\varphi(y_i)\left(-y_i\right)\frac{\left(-y_i\right)}{\sigma} - \varphi(y_i)\frac{y_i}{\sigma}\varphi(y_i)}{\left[1-\phi(y_i)\right]^2}$$

$$= -\frac{2}{\sigma^2}\sum_{i\in D} y_i - \frac{1}{\sigma^2}\sum_{i\in C}\frac{\varphi(y_i)}{1-\phi(y_i)} + \frac{1}{\sigma}\sum_{i\in C}\frac{\left(1-\phi(y_i)\right)\varphi(y_i)y_i^2 - y_i\varphi^2(y_i)}{\left[1-\phi(y_i)\right]^2}$$

$$\frac{\partial^2 \ln L(\mu,\sigma)}{\partial\sigma\,\partial\mu} = \frac{1}{\sigma}\sum_{i\in D} 2y_i\left(-\frac{1}{\sigma}\right) + \frac{1}{\sigma}\sum_{i\in C}\frac{\left(1-\phi(y_i)\right)\left\{-\frac{1}{\sigma}\varphi(y_i) + y_i\varphi(y_i)(-y_i)\left(-\frac{1}{\sigma}\right)\right\} - \frac{y_i\varphi(y_i)\varphi(y_i)}{\sigma}}{\left[1-\phi(y_i)\right]^2}$$

$$= -\frac{2}{\sigma^2}\sum_{i\in D} y_i - \frac{1}{\sigma^2}\sum_{i\in C}\frac{y_i^2\varphi(y_i) - \varphi(y_i)}{1-\phi(y_i)} - \frac{1}{\sigma^2}\sum_{i\in C}\frac{y_i\varphi^2(y_i)}{\left[1-\phi(y_i)\right]^2}$$

$$\frac{\partial^2 \ln L(\mu,\sigma)}{\partial\sigma^2} = \frac{S(D)}{\sigma^2} - \frac{1}{\sigma^2}\sum_{i\in D} y_i^2 + \frac{1}{\sigma}\sum_{i\in D} 2y_i(-1)\frac{y_i}{\sigma} - \frac{1}{\sigma^2}\sum_{i\in C}\frac{y_i\varphi(y_i)}{1-\phi(y_i)} + \frac{1}{\sigma}\sum_{i\in C}\frac{\left(1-\phi(y_i)\right)\left\{-\frac{y_i\varphi(y_i)}{\sigma} + \frac{y_i^3\varphi(y_i)}{\sigma}\right\} - \frac{y_i\varphi(y_i)y_i\varphi(y_i)}{\sigma}}{\left[1-\phi(y_i)\right]^2}$$

$$= \frac{S(D)}{\sigma^2} - \frac{3}{\sigma^2}\sum_{i\in D} y_i^2 - \frac{1}{\sigma^2}\sum_{i\in C}\frac{y_i\varphi(y_i)}{1-\phi(y_i)} + \sum_{i\in C}\frac{y_i\varphi(y_i)}{1-\phi(y_i)}$$

ve

$$\frac{\partial^2 \ln L(\mu,\sigma^2)}{\partial\mu^2} = \frac{1}{\sigma}\sum_{i\in D}\left(-\frac{1}{\sigma}\right) + \frac{1}{\sigma}\sum_{i\in C}\frac{\left(1-\phi(y_i)\right)\varphi(y_i)\left(\frac{y_i}{\sigma}\right) - \frac{\varphi^2(y_i)}{\sigma}}{\left[1-\phi(y_i)\right]^2}$$

.

$$= -\frac{1}{\sigma^2}S(D) - \frac{1}{\sigma^2}\sum_{i\in C}\frac{y_i\varphi(y_i)\left(1-\phi(y_i)\right) - \varphi^2(y_i)}{\left[1-\phi(y_i)\right]^2}$$

Böylece

$$U(\mu,\sigma) = \begin{bmatrix} \dfrac{\partial \ln L(\mu,\sigma)}{\partial\mu} \\ \dfrac{\partial \ln L(\mu,\sigma)}{\partial\sigma} \end{bmatrix}$$

ve

$$V(\mu,\sigma)=\begin{bmatrix} \dfrac{\partial^2 \ln L(\mu,\sigma)}{\partial\mu^2} & \dfrac{\partial^2 \ln L(\mu,\sigma)}{\partial\mu\,\partial\sigma} \\ \dfrac{\partial^2 \ln L(\mu,\sigma)}{\partial\sigma\,\partial\mu} & \dfrac{\partial^2 \ln L(\mu,\sigma)}{\partial\sigma^2} \end{bmatrix}$$

olmak üzere Newton Raphson yönteminden $\mu(1)$ ve $\sigma(1)$ başlangıç değerleri ile,

$$\begin{bmatrix} \mu(m+1) \\ \sigma(m+1) \end{bmatrix}=\begin{bmatrix} \mu(m) \\ \sigma(m) \end{bmatrix}-V^{-1}\left(\mu(m),\sigma(m)\right)U\left(\mu(m),\sigma(m)\right), \quad m=1,2,...$$

iterasyon işlemlerine ulaşılır. Sırasıyla $\mu(m)$ ve $\sigma(m)$ sırasıyla $\mu(m+1)$ ve $\sigma(m+1)$ sayılarına yeterince yakınsa iterasyon durdurularak $\mu$ ile $\sigma$'nın ençok olabilirlik tahmin değerlerine ulaşılır.

**Örnek 5.2.4.1** $F$, $\mu$ ve $\sigma$ parametreli log-normal dağılım fonksiyonu olsun. Sansürleme rasgele değişkeninin dağılımı (0,8) aralığındaki düzgün dağılım alınarak simülasyon yoluyla $\mu=0$ ve $\sigma=1$ olan log-normal dağılımdan $n=50$ birimlik sağdan rasgele sansürlenmiş örneklem çekildiğinde bu sansürlü gözlemlere bağlı olarak Newton-Raphson yöntemi yardımıyla $\mu$ ve $\sigma$'nın en çok olabilirlik tahmin edicileri $\hat{\mu}=0.0737$ ve $\hat{\sigma}=0.9959$ olarak hesaplanmıştır. Buna ilişkin olarak yazılan Matlab programı EK 2' de verilmiştir.

### 5.2.5 Gamma dağılımı

$F$, $\alpha$ şekil ve $\beta$ ölçek parametreli gamma dağılımı,

$$f(x:\alpha,\beta)=\frac{1}{\Gamma(\alpha)\beta^{\alpha}}x^{\alpha-1}e^{-x/\beta}, \qquad x>0\ ,\ \alpha>0\ ,\ \beta>0$$

olasılık yoğunluk fonksiyonu ile verilir. $\alpha$ bir doğal sayı ise gamma dağılımının $F$ dağılım fonksiyonu aşağıda verilen analitik ifadeye sahiptir. $\alpha$ 'nın doğal sayı olmaması durumunda $F$ 'nın analitik ifadesi yoktur.

$$F(x) = 1 - \sum_{k=0}^{\alpha-1} \frac{(x/\beta)^k}{k!} e^{-x/\beta} \qquad , \quad x > 0; \alpha \in \{1, 2, ...\}, \beta > 0$$

$X_1, \dots, X_n$ $\alpha$ ve $\beta$ parametreli bir gamma dağılımından sağdan rasgele sansürlenmiş $n$ birimlik bir örneklem olmak üzere, sansürleme rasgele değişkeninin dağılım fonksiyonu olan $G$ 'nin $\alpha$ ve $\beta$ parametrelerini içermediğini kabul edelim. Bu durumda

$$\begin{aligned} L(\alpha,\beta) &= \prod_{i\in D} f(x_i) \prod_{i\in C} \left(1 - F(x_i)\right) \\ &= \prod_{i\in D} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x_i^{\alpha-1} e^{-x_i/\beta} \prod_{i\in C} \left(1 - F(x_i)\right) \end{aligned}$$

olup

$$\begin{aligned} \ln L(\alpha,\beta) &= \sum_{i\in D} \left[ (\alpha-1)\ln x_i - \frac{x_i}{\beta} - \ln\Gamma(\alpha) - \alpha \ln\beta \right] + \prod_{i\in C} \left(1 - F(x_i)\right) \\ &= (\alpha-1)\sum_{i\in D} \ln x_i - \frac{1}{\beta}\sum_{i\in D} x_i - S(D)\ln\Gamma(\alpha) - S(D)\alpha\ln\beta + \prod_{i\in C} \left(1 - F(x_i)\right) \end{aligned}$$

bulunur. $\alpha$ ve $\beta$ 'nın ençok olabilirlik tahmin edicilerini bulabilmek için bu denklemin $\alpha$ ve $\beta$ ' ya göre türevleri alınarak aşağıdaki iki denkleme ulaşılır.

$$\begin{aligned} \frac{\partial \ln L(\alpha,\beta)}{\partial\alpha} &= \sum_{i\in D} \ln x_i - \frac{S(D)d\ln\Gamma(\alpha)}{d\alpha} - S(D)\ln\beta + \sum_{i\in C} \frac{\frac{d}{d\alpha}\left[1 - F(x_i)\right]}{1 - F(x_i)} \\ &= \sum_{i\in D} \ln x_i - \frac{S(D)d\ln\Gamma(\alpha)}{d\alpha} - S(D)\ln\beta + \sum_{i\in C} \frac{\frac{d\ln\Gamma(\alpha)}{d\alpha}\int_0^{x_i/\beta} \frac{e^{-t}t^{\alpha-1}}{\Gamma(\alpha)}dt - \frac{1}{\Gamma(\alpha)}\int_0^{x_i/\beta} t^{\alpha-1}e^{-t}\ln t dt}{1 - F(x_i)} \end{aligned}$$

$$= \sum_{i\in D} \ln x_i - \frac{S(D)d\ln\Gamma(\alpha)}{d\alpha} - S(D)\ln\beta + \sum_{i\in C} \frac{\frac{d\ln\Gamma(\alpha)}{d\alpha}\int_0^{x_i/\beta} \frac{e^{-t}t^{\alpha-1}}{\Gamma(\alpha)}dt - \frac{1}{\Gamma(\alpha)}\int_0^{x_i/\beta} t^{\alpha-1}e^{-t}\ln t dt}{1 - F(\frac{x_i}{\beta}; \alpha, \beta = 1)} \quad (5.7)$$

ve

$$
\begin{aligned}
\frac{\partial \ln L(\alpha,\beta)}{\partial\beta} &= \frac{1}{\beta^2}\sum_{i\in D}\ln x_i - \frac{S(D)\alpha}{\beta} + \sum_{i\in C}\frac{\frac{d}{d\beta}\left[1-F(x_i)\right]}{1-F(x_i)} \\
&= \frac{1}{\beta^2}\sum_{i\in D}\ln x_i - \frac{S(D)\alpha}{\beta} + \sum_{i\in C}\frac{\frac{x_i^{\alpha}e^{-x_i/\beta}}{\Gamma(\alpha)\beta^{\alpha+1}}}{1-F(x_i)} \\
&= \frac{1}{\beta^2}\sum_{i\in D}\ln x_i - \frac{S(D)\alpha}{\beta} + \alpha\sum_{i\in C}\frac{f(x_i;\alpha+1,\beta)}{1-F(x_i)}
\end{aligned}
\tag{5.8}
$$

olduğundan denklem sistemimiz

$$
\sum_{i\in D}\ln x_i - \frac{S(D)d\ln\Gamma(\alpha)}{d\alpha} - S(D)\ln\beta + \sum_{i\in C}\frac{\frac{d\ln\Gamma(\alpha)}{d\alpha}\int_0^{x_i/\beta}\frac{e^{-t}t^{\alpha-1}}{\Gamma(\alpha)}dt - \frac{1}{\Gamma(\alpha)}\int_0^{x_i/\beta}t^{\alpha-1}e^{-t}\ln t dt}{1-F(\frac{x_i}{\beta};\alpha,\beta=1)} = 0
$$

ve

$$
\frac{1}{\beta^2}\sum_{i\in D}\ln x_i - \frac{S(D)\alpha}{\beta} + \alpha\sum_{i\in C}\frac{f(x_i;\alpha+1,\beta)}{1-F(x_i)} = 0
$$

olur. Bu iki denklemin ortak analitik çözümü yoktur. $\alpha$ ve $\beta$'nın en çok olabilirlik tahminleri Newton-Raphson yöntemi ile hesaplanabilir. Bu yöntemin çalıştırılmasında gerekli matematiksel ifadeler aşağıdaki çıkartılır.

$$
\begin{aligned}
\frac{\partial^2 \ln L(\alpha,\beta)}{\partial\alpha\,\partial\beta} &= \frac{-S(D)}{\beta} + \sum_{i\in C}\frac{\left[1-F(x_i)\right]\frac{d\ln\Gamma(\alpha)}{d\alpha}\left(-\frac{x_i}{\beta^2}\right)\frac{e^{-x_i/\beta}\left(\frac{x_i}{\beta}\right)^{\alpha-1}}{\Gamma(\alpha)}}{\left[1-F(x_i)\right]^2} + \sum_{i\in C}\frac{\left[1-F(x_i)\right]\frac{1}{\Gamma(\alpha)}\frac{x_i}{\beta^2}\left(\frac{x_i}{\beta}\right)^{\alpha-1}e^{-x_i/\beta}\ln\left(\frac{x_i}{\beta}\right)}{\left[1-F(x_i)\right]^2} \\
&\quad - \frac{\frac{1}{\Gamma(\alpha)\beta^{\alpha+1}}x_i^{\alpha}e^{-x_i/\beta}\left[\frac{d\ln\Gamma(\alpha)}{d\alpha}\int_0^{x_i/\beta}\frac{e^{-t}t^{\alpha-1}}{\Gamma(\alpha)}dt - \frac{1}{\Gamma(\alpha)}\int_0^{x_i/\beta}t^{\alpha-1}e^{-t}\ln t dt\right]}{\left[1-F(x_i)\right]^2} \\
&= \frac{-S(D)}{\beta} - \alpha g(\alpha)\sum_{i\in C}\frac{f(x_i,\alpha+1,\beta)}{1-F(x_i)} + \alpha\sum_{i\in C}\frac{f(x_i,\alpha+1,\beta)}{1-F(x_i)}\ln\left(x_i/\beta\right) \\
&\quad - \alpha g(\alpha)\sum_{i\in C}\frac{f(x_i,\alpha+1,\beta)F(x_i)}{\left[1-F(x_i)\right]^2} + \alpha\sum_{i\in C}\frac{f(x_i,\alpha+1,\beta)}{\left[1-F(x_i)\right]^2}
\end{aligned}
$$

$$\frac{\partial^2 \ln L(\alpha,\beta)}{\partial \beta^2} = -\frac{S(D)\alpha}{\beta^2} + \frac{2}{\beta^3}\sum_{i\in D} x_i + \frac{\sum_{i\in C}\frac{x_i^{\alpha}}{\Gamma(\alpha)}\frac{e^{-x_i/\beta}}{\beta^{\alpha+2}}\left[\frac{x_i}{\beta}-(\alpha+1)\right]\left[1-F(x_i)\right] - \frac{x_i^{\alpha}e^{-x_i/\beta}}{\Gamma(\alpha)\beta^{\alpha+1}}\frac{x_i^{\alpha}e^{-x_i/\beta}}{\Gamma(\alpha)\beta^{\alpha+1}}}{\left[1-F(x_i)\right]^2}$$

$$= -\frac{S(D)\alpha}{\beta^2} - \frac{2}{\beta^3}\sum_{i\in D} x_i + \sum_{i\in C}\frac{\frac{x_i^{\alpha}e^{-x_i/\beta}\left[\frac{x_i}{\beta}-(\alpha+1)\right]}{\Gamma(\alpha)\beta^{\alpha+2}}}{\left[1-F(x_i)\right]} - \sum_{i\in C}\frac{x_i^{2\alpha}e^{-2x_i/\beta}}{\Gamma(\alpha)\beta^{\alpha+1}\left[1-F(x_i)\right]^2}$$

$$= -\frac{S(D)\alpha}{\beta^2} - \frac{2}{\beta^3}\sum_{i\in D} x_i + \frac{\alpha}{\beta}\sum_{i\in C}\frac{f(x_i;\alpha+1,\beta)\left[\frac{x_i}{\beta}-(\alpha+1)\right]}{\left[1-F(x_i)\right]} - \alpha^2\sum_{i\in C}\frac{f^2(x_i;\alpha+1,\beta)}{\left[1-F(x_i)\right]^2}$$

ve

$$\frac{\partial^2 \ln L(\alpha,\beta)}{\partial \alpha^2} = -S(D)\frac{d}{d\alpha}g(\alpha) + \frac{d}{d\alpha}g(\alpha)\sum_{i\in C}\frac{F(x_i)}{1-F(x_i)} + g(\alpha)\sum_{i\in C}\frac{d}{d\alpha}\left[\frac{F(x_i)}{1-F(x_i)}\right] - \sum_{i\in C}\frac{d}{d\alpha}\left[\frac{I_i}{1-F(x_i)}\right]$$

$$= -S(D)\frac{d}{d\alpha}g(\alpha) + \frac{d}{d\alpha}g(\alpha)\sum_{i\in C}\frac{F(x_i)}{1-F(x_i)} + g(\alpha)\sum_{i\in C}\frac{I_i - g(\alpha)F(x_i)}{\left[1-F(x_i)\right]^2} - \sum_{i\in C}\frac{K_i}{1-F(x_i)} - \sum_{i\in C}\frac{I_i\left(I_i - g(\alpha)\right)}{\left[1-F(x_i)\right]^2}$$

burada

$$I_i = \int_0^{x_i/\beta}\frac{t^{\alpha-1}e^{-t}}{\Gamma(\alpha)}\ln t\,dt \qquad i=1,2,\dots,n,$$

$$K_i = \int_0^{x_i/\beta}\frac{t^{\alpha-1}e^{-t}(\ln t)^2}{\Gamma(\alpha)}dt \qquad i=1,2,\dots,n$$

ve

$$g(\alpha) = \frac{d\left(\ln\Gamma(\alpha)\right)}{d\alpha} = \frac{\frac{d\Gamma(\alpha)}{d\alpha}}{\Gamma(\alpha)}$$

şeklinde tanımlanır.

$$U(\alpha,\beta) = \begin{bmatrix}\dfrac{\partial \ln L(\alpha,\beta)}{\partial\alpha}\\[2ex] \dfrac{\partial \ln L(\alpha,\beta)}{\partial\beta}\end{bmatrix}$$

ve

$$V(\alpha,\beta) = \begin{bmatrix}\dfrac{\partial^2 \ln L(\alpha,\beta)}{\partial\alpha^2} & \dfrac{\partial^2 \ln L(\alpha,\beta)}{\partial\alpha\,\partial\beta}\\[2ex] \dfrac{\partial^2 \ln L(\mu,\sigma^2)}{\partial\beta\,\partial\alpha} & \dfrac{\partial^2 \ln L(\mu,\sigma^2)}{\partial\beta^2}\end{bmatrix}$$

olmak üzere Newton Raphson yönteminden $\alpha(1)$ ve $\beta(1)$ başlangıç değerleri ile,

$$\begin{bmatrix} \alpha(m+1) \\ \beta(m+1) \end{bmatrix} = \begin{bmatrix} \alpha(m) \\ \beta(m) \end{bmatrix} - V^{-1}\left(\alpha(m),\beta(m)\right) U\left(\alpha(m),\beta(m)\right), \quad m = 1,2,...$$

iterasyon işlemlerine ulaşılır. $\alpha(m)$ ve $\beta(m)$ sırasıyla $\alpha(m+1)$ ve $\beta(m+1)$ sayısına yeterince yakınsa iterasyon durdurulur. $\alpha$ ile $\beta$ ‘nın en çok olabilirlik tahmin değerleri olan $\hat{\alpha} = \alpha(m+1)$ ve $\hat{\beta} = \beta(m+1)$ sayılarına ulaşılır.

**Örnek 5.2.5.1** $F$ $\alpha$ ve $\beta$ parametreli bir gamma dağılım fonksiyonu olsun. Sansürleme rasgele değişkeninin dağılımı (0,8) aralığındaki düzgün dağılım alınarak, simülasyon yoluyla $\alpha = 2$ ve $\beta = 1$ olan gamma dağılımdan $n = 50$ birimlik sağdan rasgele sansürlenmiş örneklem çekildiğinde bu sansürlü gözlemlere bağlı olarak Newton Raphson yöntemi yardımıyla $\alpha$ ve $\beta$’nın en çok tahmin edicileri $\hat{\alpha} = 1.8329$ ve $\hat{\beta} = 0.9816$ olarak bulunmuştur.Buna ilişkin olarak yazılan Matlab programı EK3’de verilmiştir.

### 5.3 Ömür Tablo Yöntemleri ve Çarpım-Limit (Kaplan-Meier) Tahmini

Yenilemeler arası geçen zaman süreleri dağılım fonksiyonu $F$ olan bir yenileme süreci için $F$’in bilinmediğini kabul edelim. $F$ dağılımından sağdan rasgele sansürlenmiş $n$ birimlik bir örneklem göz önüne alınsın, burada gözlemler yalnızca hangi aralıklardan bileşenlerin öldüğü (bozulduğu) ya da sansürlendiği bilinmek üzere sınıflandırılsın ve tam olarak ömürleri ve sansürleme zamanları bilinmesin. Zaman ekseni $a_0 = 0$, $a_k = t$, $a_{k+1} = \infty$ ve $t$ gözlem üzerindeki bir üst limit olmak üzere $k+1$ tane, $I_j = \left[a_{j-1}, a_j\right)$ $j = 1,2,...,k+1$ aralıklarına bölünsün. Bundan dolayı gözlemler $k+1$ tane aralığın her birisine düşen ömürlerin ve sansürleme zamanlarının sayılarından oluşur. Son aralık olan $I_{k+1}$ yalnızca ömür zamanlarının bulunduğu aralık olarak ele alınabilecektir, çünkü

$t$ zamanına kadar bozulmayan bütün bileşenler $I_{k+1}$'deki bir zamanda bozulmak zorundadır.

$n_j = a_{j-1}$ zamanında risk altında (çalışan ve sansürlenmemiş) bileşen sayısı

$d_j = I_j$'de bozulanların sayısı

$w_j = I_j$'de geri çekilenlerin (sansürlenenlerin) parça sayısı

olsun. $I_j$'nin başında yaşayan bileşen sayısı $n_j$ olduğundan

$$n_1 = n \text{ ve } n_j = n_{j-1} - d_{j-1} - w_{j-1}, \quad j = 1, 2, ..., k+1$$

dir.

$F$ dağılım fonksiyonuna sahip bir rasgele değişkeni $Y$ ile gösterelim.

$$\bar{F}(a_j) = P(Y > a_j), \qquad j = 1, 2, ..., k+1$$

ve

$$p_j = P(Y > a_j \,|\, Y > a_{j-1}) \text{ ve } q_j = 1 - p_j \;, \; j = 1, 2, ..., k+1$$

olmak üzere

$$\bar{F}(a_j) = P(\text{bir bileşenin } I_j \text{ ötesinde çalışması})$$

ve

$$p_j = P(\text{bir bileşenin } I_j \text{ ötesinde çalışması} \mid I_{j-1} \text{ ötesinde çalışması})$$

$$q_j = P(\text{bir bileşenin } I_j \text{ ötesinde bozulması} \mid I_{j-1} \text{ ötesinde çalışması})$$

yazılabilir.

$$\bar{F}(a_0) = 1$$

$$\bar{F}(a_1) = P(Y > a_1)$$

$$\bar{F}(a_2) = P(Y > a_2) = P(Y > a_2 \,|\, Y > a_1)P(Y > a_1)$$

$$\bar{F}(a_3) = P(Y > a_3) = P(Y > a_3 \,|\, Y > a_2)P(Y > a_2 \,|\, Y > a_1)P(Y > a_1)$$

.

.

.

$$\bar{F}(a_k) = P(Y > a_k) = P(Y > a_k \,|\, Y > a_{k-1})P(Y > a_{k-1} \,|\, Y > a_{k-2})...P(Y > a_1)$$

oluğundan

$$\bar{F}(a_j) = p_1 p_2 ... p_j \ , \qquad j = 1, 2, ..., k+1 \tag{5.9}$$

olur. Yukarıdaki ifade ile gözlem biriminin $I_j$'yi geçme çalışma (yaşam) olasılığı herbir aralığın başlangıcına kadar çalıştığı verildiğinde $I_j$'ye kadar olan aralıkları geçmenin koşullu çalışma olasılıklarının çarpımı olarak verilir. Bu sonuç tahmin problemine yaklaşım için bir temel oluşturur.

Şimdi $\bar{F}\left(a_j\right)$'lerin tahmini problemini ele alalım. Eğer veri sansürlenmemiş ise bunun gerçekleştirilmesinde hiçbir sıkıntı olmayacaktır. Bu durumda açık bir tahmin, gerçekten $\bar{F}\left(a_j\right)$ için ençok olabilirlik tahmin edicisi $\frac{n_{j+1}}{n}$'dir, bu $a_j$'de çalışan bileşenlerin oranıdır. Eğer aralıklar geri çekilmeleri, yani sansürleme zamanlarını içerirse böyle olmayacaktır, çünkü $n_{j+1}$'in $a_j$ zamanında hala çalışan bileşenlerin sayısı olması gerekli değildir. Muhtemelen bazı sansürlenmiş bileşenler $a_j$'de aynı zamanda hala çalışacağından $\frac{n_{j+1}}{n}$ çoğu durumda $\bar{F}\left(a_j\right)$'yi alttan tahmin etme eğiliminde olacaktır. Aşağıda verilecek olan ömür tablosu (life table) yöntemi ile bu problem çözülebilecektir.

Bu yöntemde temel fikir $\bar{F}\left(a_j\right)$'nin bir tahminini elde etmekte (5.9) ifadesini kullanmaktır. Böylelikle sansürleme olsa bile gözlemlere dayanarak $\bar{F}\left(a_j\right)$'lerin anlamlı tahminlerini vermek genellikle mümkün olmaktadır.

Eğer $I_j$ aralığında sansürleme yok ise $q_j$'nin anlamlı bir tahmini $\hat{q}_j = \frac{d_j}{n_j}$ dir, çünkü $q_j$ bileşenin $I_j$'nin başında çalıştığı verildiğinde o parçanın $I_j$'de ölmesinin koşullu olasılığıdır. Fakat $w_j > 0$ ise $\frac{d_j}{n_j}$'nin $q_j$'yi alttan tahmin etmesi beklenebilecektir, çünkü $I_j$'de sansürlenmiş bileşenlerden bazılarının $I_j$'nin bitiminden önce ölmüş olabilirler. Ayrıca daha önce sansürlenmiş olupta $I_j$ ’de ölenler zaten $n_j$ ’nin içinde yoktur. Bundan dolayı sansürlü gözlemler için bir düzenleme yapmak arzulanır. En sık kullanılan yöntem $q_j$ sayısını standart ömür tablo tahmini olarak adlandırılan

$$\hat{q}_j = \frac{d_j}{n_j^{'}} \tag{5.10}$$

ile tahmin etmektir, burada $n_j' = n_j - \frac{w_j}{2}$ dir (Lawless 2003). (5.10) ifadesi $n_j > 0$ olmasını gerektirir. $n_j = 0$ olduğunda uygunluk nedenleri için $\hat{q}_j = 1$ ile tanımlanır. $I_j$ aralığı için risk altındaki bileşenlerin etkili bir sayısı olarak düşünülebilir; bu bir anlamda sansürlü bir bileşenin aralığın yarısı için risk altında çalıştığını kabul eder. Bu düzenleme keyfidir, fakat çoğu durumda akla uygundur. Bazı durumlarda $q_j$ ’nin diğer tahmin edicileri tercih edilebilir. Örneğin, $I_j$ ’deki tüm sansürlemeler $I_j$ ’nin sonunda sağda gerçekleşmişse $\hat{q}_j = \frac{d_j}{n_j}$ tahmini uygun olur, halbuki bütün sansürlemeler $I_j$ ’nin başlangıcında gerçekleşmişse $\hat{q}_j = \frac{d_j}{n_j - w_j}$ uygun olur. $\hat{q}_j$ ve $\hat{p}_j = 1 - \hat{q}_j$ tahminleri hesaplandıktan sonra $\bar{F}\left(a_j\right)$ (5.9) yardımıyla

$$\hat{\bar{F}}(a_j) = \hat{p}_1 ... \hat{p}_j \ , \quad j = 1, 2, ..., k+1 \tag{5.11}$$

ile tahmin edilebilir.

Ömür tablosunun kendisi veriyi, $\hat{q}_j$ ve $\hat{\bar{F}}(a_j)$ tahminlerini gösteren bir tablodur. Bu tablo genellikle herbir aralık için $n_j$, $d_j$, $w_j$, $\hat{q}_j$ ve $\hat{\bar{F}}(a_j)$ değerlerini veren sütunları bulundurur. Bazen $n_j'$, $\hat{p}_j$ ve arasıra dağılımın diğer karakteristiklerinin tahminlerini veren ilave sütunları bulundurur. Bütün $w_j = 0$ olduğu özel durum için $\hat{\bar{F}}(a_j)$ sansürsüz durum için önceden bahsedilen $\frac{n_{j+1}}{n}$ tahminine indirgenir.

**Örnek 5.3.1** 300 araba aküsü 300 çalışan arabaya yerleştirilmiş ve 4 yıl boyunca yılda bir kez kontrol edilmiştir. Sansürlemeler (geri çekilmeler) kazaların ve gözetimden arabanın çıkartılması sonucunda ortaya çıkmıştır. Bu 300 aküye ilişkin olarak aralık bazında $d_i$, $w_i$ ve $n_i$ değerleri Çizelge 5.1'te verilmiştir. (Gertsbakh 1989). Bu değerler yardımıyla $n_i'$, $\hat{p}_i$, $\hat{\bar{F}}(a_j)$ ve $\hat{F}(a_j)$ değerleri hesaplanarak elde edilen ömür tablosu Çizelge 5.2'de verilmiştir

Çizelge 5.1 Araba akülerine ilişkin veriler

| i | Zaman aralığı | $d_i$ | $w_i$ | $n_i$ |
|---|---|---|---|---|
| 1 | $[0,1)$ | 20 | 72 | 300 |
| 2 | $[1,2)$ | 15 | 50 | 208 |
| 3 | $[2,3)$ | 18 | 35 | 143 |
| 4 | $[3,4)$ | 17 | 23 | 90 |

Çizelge 5.2 Araba akülerine ilişkin ömür tablosu

| i | Zaman aralığı | $d_i$ | $w_i$ | $n_i$ | $n_i'$ | $\hat{p}_i$ | $\hat{\bar{F}}(a_i)$ | $\hat{F}(a_i)$ |
|---|---|---|---|---|---|---|---|---|
| 1 | $[0,1)$ | 20 | 72 | 300 | 264 | 0.924 | 0.924 | 0.076 |
| 2 | $[1,2)$ | 15 | 50 | 208 | 183 | 0.918 | 0.848 | 0.152 |
| 3 | $[2,3)$ | 18 | 35 | 143 | 125.5 | 0.857 | 0.727 | 0.273 |
| 4 | $[3,4)$ | 17 | 23 | 90 | 78.5 | 0.783 | 0.569 | 0.431 |

Ömür tablosu yönteminde $F$ dağılımlı gelen sağdan rasgele sansürlü örneklem durumunda bileşenlerin bozulma ve sansürleme zamanları bilinmeyip gözlemler aralıklar ile verilmişti. Şimdi kabul edelim ki $n$ birimlik örneklemdeki bozulma zamanları ve bozulma zamanları gözlenemeyen bileşenlerin sansürleme zamanları bilinsin. Bu durumda amacımız verilen bir $t \geq 0$ için $F(t)$ ya da denk olarak $\bar{F}(t)$'yi tahmin etmektir. Tam örneklem durumunda $\bar{F}(t)$'nin tahmin edicisinin

$$\hat{\bar{F}}(t)=\frac{\text{ömürleri } t\text{'ye eşit ya da } t\text{'den büyük gözlemlerin sayısı}}{n} \tag{5.12}$$

ile verildiği iyi bilimektedir. Örneklemin sansürlü olması durumunda ömürleri $t$'ye eşit ve $t$'den büyük olan gözlemlerin sayısı genellikle tam olarak bilinmeyeceğinden (5.12) tahmin edicisinin yeniden düzenlenmesi gerekecektir. Aşağıda tarif edilecek olan düzenleme ile $\bar{F}(t)$'nin çarpım-limit ya da Kaplan-Meier olarak adlandırılan tahminine ulaşılır.

Kabul edelim ki $n$ tane bileşenden $k$ tanesi farklı $t_1 < t_2 < ... < t_k$ zamanlarında bozulsun. $t_j$ $(j=1,2,...,k)$ zamanında birden fazla bozulmaya izin verilsin ve $d_j$ $t_j$'deki bozulmaların sayısı olsun. $t_1, t_2, \ldots, t_k$ bozulma zamanlarına ilaveten aynı zamanda ömürleri gözlenemeyen bileşenler için sansürleme zamanları vardır. Bu durumda $\bar{F}(t)$'nin Kaplan-Meier tahmini

$$\hat{\bar{F}}_{KM}(t)=\prod_{j:t_j<t}\frac{n_j-d_j}{n_j} \tag{5.13}$$

ile tanımlanır (Lawless 2002), burada $n_j$ $t_j$ zamanında risk altında bulunan, yani $t_j$ zamanına kadar henüz bozulmamış ve sansürlenmemiş bileşenlerin sayısıdır.

Eğer bir $t_j$'ye eşit olarak kaydedilmiş bir sansürleme zamanı var ise bu sansürleme zamanının $t_j$'den çok küçük miktarda büyük olduğu düşünülerek, $t_j$'de bozulan bileşenler için olduğu gibi $n_j$'nin içinde yer alacağı kabul edilmektedir. Bu akla uygundur, çünkü bir $T$ zamanında sansürlenen bileşen elbette hemen $T$'yi geçtiğinde yaşayacaktır. Örneklemdeki en büyük gözlem zamanının bir sansürleme zamanı olması durumunda Kaplan-Meier tahmini yalnızca bu son gözleme kadar tanımlı olarak alınır (Lawless 2003).

(5.13)'ün elde edilmesindeki neden esasen ömür tablolarındaki $\hat{\bar{F}}(a_j)=\hat{p}_1\hat{p}_2...\hat{p}_j$ tahmini için olan ile aynıdır; yani, $\hat{\bar{F}}(t)$ bir çarpım olarak kurulur ve bu çarpımdaki herbir terim bir bileşenin $t_j$'ye kadar çalıştığı verildiğinde $t_j$ zamanını çalışarak geçirmesinin koşullu olasılığının bir tahmini olarak düşünülebilir. (5.13) tahmini gerçekte standart ömür tablo işleminin bir limit alma durumudur, ömür tablosundaki aralıkların sayısı sonsuza götürüldüğünde ve son aralık hariç tüm aralıkların uzunlukları sıfıra yaklaştırıldığında elde edilir. Bu ilişki aşağıda incelenir. Hiçbir sansürleme olmadığı zaman $n_1 = n$ ve $n_j = n_{j-1} - d_{j-1}$ $(j = 2,...,k)$ olur ve (5.13) tahmini (5.12) tahminine indirgenir. Hem sansürlü hem de sansürsüz durumunda $\hat{\bar{F}}(t)$ $t = 0$ da 1'e eşit ve herbir $t_j$ zamanından sonra hemen bir $(n_j - d_j)/n_j$ çarpımı ile düşen bir basamak fonksiyonudur. Sansürleme zamanlarında tahmin değişmez; fakat sansürleme zamanlarının etkisi $n_j$'nin değerlerinde ve böylece $\hat{\bar{F}}(t)$'nin basamaklarının büyüklüklerinde anlaşılır.

**Örnek 5.3** Belli bir cihaz için bozulma zamanları (ay olarak) aşağıdaki Çizelge 5.3'de verilmiştir. Burada bozulma zamanları için sansürlemenin rasgele ve bağımsız olduğu varsayılmaktadır.

Çizelge 5.3. Cihazın bozulma zamanları

| Sansürlenmemiş Gözlemler | Sansürlenmiş Gözlemler |
|---|---|
| 36.3 41.7 43.9 49.9 50.1 50.8 51.9 58.9 60.7 | 26.8 29.6 33.4 35.0 40.0 41.9 |
| 52.1 52.3 52.3 52.4 52.6 52.7 53.1 59.0 | 42.5 |
| 53.6 53.6 53.9 53.9 54.1 54.6 54.8 59.1 | |
| 54.8 55.1 55.4 55.9 56.0 56.1 56.5 59.6 | |
| 56.9 57.1 57.1 57.3 57.7 57.8 58.1 60.4 | |

Bazı $t$'ler için $\bar{F}(t)$'nin Kaplan-Meier tahmin edicileri Çizelge 5.4'de gösterilmektedir. Bu tahmin değerlerini hesaplatan Matlab programı EK 4'de verilmiştir.

Çizelge 5.4 Kaplan-Meier tahmin değerleri

| t | $\hat{F}_{KM}(t)$ | t | $\hat{F}_{KM}(t)$ |
|---|---|---|---|
| 36.3 | 0.0227 | 55.1 | 0.5597 |
| 41.7 | 0.0460 | 55.4 | 0.5842 |
| 43.9 | 0.0705 | 55.9 | 0.6086 |
| 49.9 | 0.0949 | 56.0 | 0.6331 |
| 50.1 | 0.1194 | 56.1 | 0.6575 |
| 50.8 | 0.1438 | 56.5 | 0.6820 |
| 51.9 | 0.1683 | 56.9 | 0.7065 |
| 52.1 | 0.1928 | 57.1 | 0.7554 |
| 52.3 | 0.2417 | 57.3 | 0.7798 |
| 52.4 | 0.2662 | 57.7 | 0.8043 |
| 52.6 | 0.2906 | 57.8 | 0.8288 |
| 52.7 | 0.3151 | 58.1 | 0.8532 |
| 53.1 | 0.3395 | 58.9 | 0.8777 |
| 53.6 | 0.3885 | 59.0 | 0.9022 |
| 53.9 | 0.4374 | 59.1 | 0.9266 |
| 54.1 | 0.4618 | 59.6 | 0.9511 |
| 54.6 | 0.4863 | 60.4 | 0.9755 |
| 54.8 | 0.5352 | 60.7 | 1.0000 |

# 6. SANSÜRLÜ VERİLERDE YENİLEME FONKSİYONUNUN TAHMİNİ

$\{N(t),\, t \geq 0\}$ yenilemeler arası geçen zaman süreleri dağılım fonksiyonu $F$ olan bir yenileme süreci olsun. Yenileme süreçleri ile ilgili uygulamalarda genellikle bu sürecin ortalama değer fonksiyonu olan $M$ yenileme fonksiyonu bilgisine ihtiyaç duyulur. Eğer $F$ tam olarak biliniyor ise yaygın olarak kullanılan birkaç $F$ dağılımı için $M$ analitik olarak hesaplanabilir. Çoğu durumda $M$ 'nin analitik ifadesi olmayıp sayısal olarak hesaplanması gerekir. Bu fonksiyon Xie'nin RS(Rieman-Stieltjes) yöntemi yardımıyla sayısal olarak kolaylıkla hesaplanabilir. Pratikte $F$ dağılım fonksiyonunun şekli bilinirken bazı parametreleri bilinmez ya da $F$ hiç bilinmez. Bu tür durumlarda $F$ dağılımından alınan $n$ birimlik örnekleme dayalı olarak $M$ 'nin değerlerinin tahmin edilmesi gerekmektedir. Bu değerler $F$ 'nin yukarıda bahsedilen durumuna bağlı olarak parametrik ya da parametrik olmayarak tahmin edilirler.

Aşağıdaki kısımda $F$ dağılımından gelen örneklemin sağdan rasgele sansürlenmiş olması durumunda $M(t)$ 'nin parametrik tahmini üzerinde durulur.

## 6.1 $M(t)$ için Bir Parametrik Tahmin Edici

$\theta_1, \theta_2, ..., \theta_r$ $F$ dağılımının bilinmeyen parametreleri olsun ve $F = F(\theta_1, \theta_2, ..., \theta_r)$ yazılsın. $X_1, \ldots, X_n$ $F$ dağılımından $n$ birimlik sağdan rasgele sansürlenmiş örneklem olmak üzere bu örnekleme dayalı $\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_r$ sırasıyla $\theta_1, \theta_2, ..., \theta_r$ bilinmeyen parametrelerinin tahmin edicileri olsunlar. $\hat{F}_n = F(\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_r)$ $F$ 'nin bir tahmin edicisidir. $F_n^{k*} = \underbrace{\hat{F}_n * ... * \hat{F}_n}_{k\ kez}$ olmak üzere $M(t)$ 'nin (3.2) konvolüsyon serisi ifadesinde $F^{k*}(t)$ yerine tahmin edicisi olan $\hat{F}_n^{k*}(t)$ alınmasıyla

$$\hat{M}_n(t) = \sum_{k=1}^{\infty} \hat{F}_n^{k*}(t), \qquad t \geq 0 \tag{6.1}$$

ile tanımlanan $\hat{M}_n(t)$ yenileme fonksiyonunun $t$ noktasındaki değerinin bir tahmin edicisi olarak göz önüne alınabilir.

Eğer $F$ dağılım fonksiyonunun parametrelerine bağlı olarak $M$ yenileme fonksiyonu için bir analitik ifade var ise bu ifadede dağılımın parametreleri yerine tahmin edicilerinin konulmasıyla elde edilen tahmin edicinin $\hat{M}_n(t)$ olduğu açıktır.

**Örnek 6.1** $F$, $\theta$ parametreli üstel dağılım fonksiyonu olsun; yani $F(x) = 1 - e^{-x/\theta}$, $\quad x \geq 0,\ \theta > 0$. Bu dağılım üzerine kurulu yenileme süreci $1/\theta$ oranlı Poisson süreci olup

$$M(t) = \frac{t}{\theta}, \qquad t \geq 0$$

dır. $X_1, \ldots, X_n$ $F$ dağılımından sağdan rasgele sansürlenmiş $n$ birimlik örneklem olsun. (5.2) den $\theta$'nın en çok olabilirlik tahmin edicisin $\hat{\theta} = \dfrac{\sum_{i=1}^{n} X_i}{S(D)}$ olduğunu biliyoruz. Bu durumda $M(t)$'nin parametrik tahmin edicisi

$$\hat{M}_n(t) = \frac{S(D)}{\sum_{i=1}^{n} X_i} t$$

olur. $\hat{\theta}$ ençok olabilirlik tahmin edicisi $\theta$ için tutarlı olduğundan $\hat{M}_n(t)$ tahmin edicisinin $M(t)$ için tutarlı olduğu açıktır.

Şimdi kabul edelim ki $G$ sansürleme dağılım fonksiyonu $\theta_1$ parametreli üstel olsun. Bu durumda herhangi $i \in \{1, 2, ..., n\}$ için

$$
\begin{aligned}
P(X_i > x) &= P(\min(Y_i, T_i) > x) \\
&= P(Y_i > x, T_i > x) \\
&= P(Y_i > x)P(T_i > x) \\
&= e^{-x/\theta} e^{-x/\theta_1} \\
&= e^{-x(\frac{1}{\theta}+\frac{1}{\theta_1})}, \qquad x > 0
\end{aligned}
$$

olduğundan $X_i$ rasgele değişkeni $\dfrac{\theta\theta_1}{\theta+\theta_1}$ ortalaması ile üstel dağılımlı olur. Böylelikle,

$\sum_{i=1}^{n} X_i$ rasgele değişkeni $\alpha = n$ ve $\beta = \dfrac{\theta\theta_1}{\theta+\theta_1}$ parametreli gamma dağılımı olup,

$$
\begin{aligned}
E(\hat{M}_n(t)) &= S(D)\,t \int_0^\infty \frac{1}{y(n-1)!}\left(\frac{\theta+\theta_1}{\theta\theta_1}\right)^n y^{n-1} e^{-y\frac{\theta+\theta_1}{\theta\theta_1}} dy \\
&= \frac{S(D)\,t}{(n-1)!}\left(\frac{\theta+\theta_1}{\theta\theta_1}\right)^n \int_0^\infty y^{n-1-1} e^{-y\frac{\theta+\theta_1}{\theta\theta_1}} dy \\
&= \frac{S(D)\,t}{(n-1)!}\left(\frac{\theta+\theta_1}{\theta\theta_1}\right)^n (n-2)!\left(\frac{\theta\theta_1}{\theta+\theta_1}\right)^{n-1} \\
&= \frac{S(D)\,t}{(n-1)}\left(\frac{\theta+\theta_1}{\theta\theta_1}\right), \qquad n > 1
\end{aligned}
$$

elde edilir. $\hat{\theta} = \dfrac{\sum_{i=1}^{n} X_i}{S(D)}$ tahmin edicisi $\theta$ için asimptotik yansız olduğundan

$$
\begin{aligned}
E(\hat{\theta}) = \frac{E\left(\sum_{i=1}^{n} X_i\right)}{S(D)} &= \frac{1}{S(D)} \int_0^\infty y \frac{1}{\Gamma(n)}\left(\frac{\theta+\theta_1}{\theta\theta_1}\right)^n y^{n-1} e^{-y\left(\frac{\theta+\theta_1}{\theta\theta_1}\right)} dy \\
&= \frac{1}{S(D)(n-1)!}\left(\frac{\theta+\theta_1}{\theta\theta_1}\right)^n \int_0^\infty y^{n+1-1} e^{-y\left(\frac{\theta+\theta_1}{\theta\theta_1}\right)} dy \\
&= \frac{1}{S(D)(n-1)!}\left(\frac{\theta+\theta_1}{\theta\theta_1}\right)^n n!\left(\frac{\theta+\theta_1}{\theta\theta_1}\right)^{n+1} \\
&= \frac{n}{S(D)} \frac{\theta\theta_1}{\theta+\theta_1} \xrightarrow[n\to\infty]{} \theta
\end{aligned}
$$

olup $\quad \lim\limits_{n\to\infty} \dfrac{n}{S(D)} = \dfrac{\theta+\theta_1}{\theta_1}$ olduğu

bulunur. Bu durumda

$$\lim_{n\to\infty} E(\hat{M}_n(t)) = \lim_{n\to\infty} \frac{S(D)t}{(n-1)} \frac{\theta+\theta_1}{\theta\theta_1}$$
$$= \frac{\theta_1}{\theta+\theta_1} t \frac{\theta+\theta_1}{\theta\theta_1}$$
$$= \frac{t}{\theta}$$
$$= M(t)$$

dir. Böylece $\hat{M}_n(t)$'nin $M(t)$ için asimptotik yansız olduğu gösterilmiş olur.

$\hat{M}_n(t)$ tahmin edicisinin tutarlılığı aşağıdaki teoremde verilmiştir.

**Teorem 6.1** $F = F(\theta_1, \theta_2, ..., \theta_r)$, $f = f(\theta_1, \theta_2, ..., \theta_r)$ olasılık yoğunluk fonksiyonu ile mutlak sürekli bir dağılım fonksiyonu ve $f$ herbir $\theta_i$, $i = 1, 2, ..., r$ parametresine göre sürekli olsun. $X_1, \ldots, X_n$ $F$ dağılımından sağdan rasgele sansürlenmiş $n$ birimlik bir örneklem olmak üzere bu örnekleme dayalı $\hat{\theta}_1, \ldots, \hat{\theta}_r$ sırasıyla $\theta_1, \ldots, \theta_r$ parametrelerinin güçlü tutarlı tahmin edicileri ise her sabit $t \geq 0$ için $\hat{M}_n(t)$ , $M(t)$'nin güçlü tutarlı bir tahmin edicisidir, yani

$$\hat{M}_n(t) \xrightarrow[n\to\infty]{h.h.h.y} M(t)$$

dir.

**İspat** $\hat{f}_n = f(\hat{\theta}_1, ..., \hat{\theta}_r)$ diyelim. $f$ olasılık yoğunluk fonksiyonunun $\theta_i$ parametresine göre sürekli ve $\hat{\theta}_i \xrightarrow[n\to\infty]{h.h.h.y} \theta_i$, $i = 1, 2, ..., r$ olmasından dolayı her sabit $t$ için

$$\hat{f}_n(t) \xrightarrow[n\to\infty]{h.h.h.y} f(t)$$

olur. Bu durumda Scheffe teoreminden $\hat{F}_n(t) \xrightarrow[n\to\infty]{h.h.h.y} F(t)$ bulunur ve sonucunda tümevarım metodu ile $k = 1, 2, ...$ için $\hat{F}_n^{k*}(t) \xrightarrow[n\to\infty]{h.h.h.y} F^{k*}(t)$ ifadesine ulaşılır. $\overline{A}$ $\hat{F}_n^{k*}(t)$'nin $F^{k*}(t)$'ye yakınsamadığı noktaların kümesi olmak üzere $P(\overline{A}) = 0$ dır. $F(0) < 1$ ve $F$ sağdan sürekli olduğundan $F(a) < 1$ olacak şekilde en az bir $a > 0$

sayısı vardır. $r \in \mathbb{N}$ sayısı $t \le ra$ olacak biçimde seçilsin. $(S_r \le t) \Rightarrow (S_r \le ra)$ olacağından $P(S_r \le t) \Rightarrow P(S_r \le ra)$ dir. Ayrıca $(Y_1 + \ldots + Y_n \le ra)$ olayı $(Y_1 > a, \ldots, Y_r > a)$ olayının gerçekleşmemesini gerektirdiğinden ve rasgele değişkenlerin bağımsız olmalarından

$$P(S_r \le ra) \le 1 - P(Y_1 > a, \ldots, Y_r > a)$$
$$= 1 - [F(a)]^r, \quad [F(a)]^r > 0$$

elde edilir. Böylece

$$F^{r*}(t) \le 1 - [1 - F(t/r)]^r, \quad [1 - F(t/r)]^r > 0 \tag{6.2}$$

yazılabilir. Bu durumda $t > 0$ için $F^{r*}(t) < 1$ olacak biçimde bir $r \ge 1$ vardır. Her $k \ge 1$ için $\hat{F}_n^{k*}(t)$ $F^{k*}(t)$'ye $A$ üzerinden yakınsak olduğundan (6.2) deki $r$ için $\delta = 1 - F^{r*}(t)$ olmak üzere her sabit $\omega \in A$ için bir $n_0(\omega)$ vardır öyle ki her $n \ge n_0(w)$ için $\hat{F}_n^{k*}(t) < 1 - \delta/2$ dir. Burada $1 - \delta/2 < 1$ olduğu açıktır.

$$g_{n,n_o(\omega)} = \begin{cases} 1 - \delta/2, & n \ge n_o(\omega) \\ 1 \quad , & n < n_o(\omega) \end{cases}$$
$$= (1 - \delta/2) I(n \ge n_o(\omega)) + 1.I(n < n_o(\omega))$$

olsun. Bu durumda $k \ge r$ için

$$\hat{F}_n^{k*}(t) \le F^{(k-r)*}(t) F^{r*}(t)$$
$$\le \left(F^{r*}(t)\right)^{[\![k/r]\!]}$$

olduğunun göz önüne alınmasıyla

$$\hat{F}_n^{k*}(t) \le \left(g_{n,n_0(\omega)}\right)^{[\![k/r]\!]}$$

dir. $n > n_0(\omega)$ için

$$\sum_{k=1}^{\infty} \hat{F}_n^{k*}(t) \le \sum_{k=1}^{r} \hat{F}^{k*}(t) + \sum_{k=r+1}^{\infty} \hat{F}^{k*}(t)$$
$$\le \sum_{k=1}^{r} \left(F^{r*}(t)\right)^{[\![k/r]\!]} + \sum_{k=r+1}^{\infty} \left(g_{n,n_0(\omega)}\right)^{[\![k/r]\!]} < \infty$$

olduğundan Lebesgue baskın yakınsaklık teoreminden her $\omega \in A$ için

$$\lim_{n\to\infty}\sum_{k=1}^{\infty}\hat{F}_n^{k*}(t)=\sum_{k=1}^{\infty}\lim_{n\to\infty}\hat{F}_n^{k*}(t)$$
$$=\sum_{k=1}^{\infty}F^{k*}(t)$$

elde edilir. Böylece $P(A)=1$ olduğundan her sabit $t$ için $\hat{M}_n(t)$ $M(t)$'ye hemen hemen her yerde yakınsaktır. Böylelikle ispat tamamlanır.

$\hat{M}_n(t)$'nin asimptotik yansızlığı aşağıdaki teorem ile verilmektedir.

**Teorem 6.2** Teorem 6.1'in şartları altında herhangi bir $t_0 \geq 0$ için $F(t_0)<1$ ise $t \leq t_0$ için

$$\lim_{n\to\infty}E(\hat{M}_n(t))=M(t),$$

yani $t \leq t_0$ için $\hat{M}_n(t)$ $M(t)$ için asimptotik yansızdır.

**İspat** $F(t_0)<1$ olduğundan $F\left(t_0\right)\leq\frac{c}{1+c}$ olacak biçimde bir $c>0$ sabiti vardır. Bu durumda

$$\begin{aligned}
M\left(t_0\right) &= \sum_{k=1}^{\infty}F^{k*}\left(t_0\right)\\
&\leq \sum_{k=1}^{\infty}F^{k}\left(t_0\right)\\
&= \sum_{k=1}^{\infty}\left(\frac{c}{1+c}\right)^k\\
&= \frac{c}{1+c}\left(1+\frac{c}{1+c}+\left(\frac{c}{1+c}\right)^2+\dots\right.\\
&= \frac{c}{1+c}\frac{1}{1-\frac{c}{1+c}}\\
&= c
\end{aligned}$$

olur. Böylece $F(t_0)<1$ olduğundan $\hat{F}(t_0)<1$ olup $t\leq t_0$ olmak üzere bir $c$ pozitif sabiti için

$$\hat{M}_n(t) \leq c$$

yazılabilir. Sonuçta sınırlı yakınsaklık teoreminden $t \leq t_0$ için

$$\lim_{n\to\infty} E\left[\hat{M}_n(t)\right] = M(t)$$

elde edilir. Bu da ispatı tamamlar.

$M(t)$ için analitik bir ifade olmaksızın $\hat{M}_n(t)$ tahmin edicisinin hesabı için (6.1) konvolüsyon serisi ifadesi genelde kullanışsızdır. $M(t)$ yenileme fonksiyonu için (3.4) integral denkleminden $\hat{M}_n(t)$,

$$\hat{M}_n(t) = \hat{F}_n(t) + \int_0^t \hat{M}_n(t-x)d\hat{F}_n(x) \tag{6.3}$$

biçiminde yazılabilir. Bu integral denklemin hesaplanması tahmin edicinin konvolüsyon serisi ifadesine göre daha kolaydır ve Xie'nin kısım (4.2) de anlatılan RS yöntemi ile kolaylıkla hesaplanabilir.

### 6.2 $M(t)$ için Bir Parametrik Olmayan Tahmin Edici

Bu kısımda $F$ dağılım fonksiyonu bilinmediğinde bu dağılımdan alınan sağdan rasgele sansürlenmiş örnekleme dayalı olarak $F(t)$'nin Kısm 5.3 de verilen Kaplan-Meier tahmin edicisi $\hat{F}_{KM}(t)$ yardımıyla $M(t)$'nin parametrik olmayan tahmini üzerinde durulmaktadır.

$X_1,\ldots,X_n$ $F$ dağılımından sağdan rasgele sansürlenmiş $n$ birimlik bir örneklem olsun. Bu örnekleme dayalı olarak $F(t)$'nin tahmini için (5.13) de verilen

$$\hat{\bar{F}}_{KM}(t) = \prod_{j:t_j<t} \frac{n_j - d_j}{n_j}$$

Kaplan-Meier tahmin edicisini göz önüne alalım. $M(t)$'nin (3.2) konvolüsyon serisi ifadesinde $F(t)$ yerine bir tahmin edicisi olan $\hat{F}_{KM}(t)$ alınmasıyla tanımlanan

$$\hat{M}_{KM}(t) = \sum_{k=1}^{\infty} \hat{F}_{KM}^{k*}(t) \tag{6.4}$$

$M(t)$'nin parametrik olmayan bir tahmin tahmin edicisi olarak göz önüne alınabilir. Burada $\hat{F}_{KM}^{k*}(t)$, $\hat{F}_{KM}(t)$'nin kendi kendisiyle olan $k$ kez Stieltjes konvolüsyonudur.

$(X_1,\ldots,X_n)$ örnekleminin her sabit değeri için $\hat{M}_{KM}(t)$ tahmin edicisinin değeri (6.4) denkleminden kolaylıkla elde edilemez. $X_1,\ldots,X_n$ örnekleminin her sabit değeri için (3.4) yenileme denkleminden $\hat{M}_{KM}(t)$

$$\hat{M}_{KM}(t) = \hat{F}_{KM}(t) + \int_0^t \hat{M}_{KM}(t-x)d\hat{F}_{KM}(x) \tag{6.5}$$

biçiminde yazılabilir. Böylece $\hat{M}_{KM}(t)$ bu denklemin çözülmesiyle elde edilebilir. Scheneider, Lin ve O'Cinneide bu (6.5) gibi bir integral denkleminin sayısal çözümü için bir yöntem önermişlerdir. Şimdi önerilen bu yöntemi kısaca verelim. $(0,t)$ aralığı kesiklendirilerek $X_i$, $i = 1,2,...,n$ ler uygun bir $h$ tamsayısı ile çarpılarak yeniden ölçeklendirilir ve en yakın tamsayıya yuvarlatılır. Böylece sonuçta yenileme fonksiyonunun hesabı için $\hat{F}_{KM}$ fonksiyonuna bir aritmetik dağılım fonksiyonu ile yaklaşılır. RS-yöntem yaklaşımına benzer olarak (6.5) deki integralin bir sonlu toplamla yer değiştirmesiyle $\hat{M}_{KM}(t)$ ardışık olarak hesaplabilir. $t = k$ ve $F_a$ yukarıda bahsedilen aritmetik dağılım fonksiyonu olmak üzere $\hat{M}_{KM}$

$$\hat{M}_{KM}^{a}(i) = F_a(i) + \sum_{j=1}^{i} \hat{M}_{KM}^{a}(i)(F_a(j) - F_a(j-1)), \quad i = 1,2,...,k \tag{6.6}$$

toplamından ardışık olarak hesaplanır. $i = 1,2,...,k$ olmak üzere $\hat{M}_{KM}(i/h) = \hat{M}_{KM}^{a}(i)$ dir.

## 6.3 $\hat{M}_{KM}(t)$ için (3.11) Asimptotik İfadesine Bağlı Bir Tahmin Edici

$M(t)$ için (3.10)'da verilen

$$\lim_{t\to\infty}\left(M(t)-\frac{t}{\mu}\right)=\frac{\mu_2}{2\mu^2}-1$$

ifadesini göz önüne alalım. $X_1,\ldots,X_n$ $F$ dağılımından sağdan rasgele sansürlenmiş $n$ birimlik bir örneklem olsun. $\hat{\mu}$ ve $\hat{\mu}_2$ bu örnekleme dayalı $F$'in sırasıyla $\mu$ ortalamasının ve $\mu_2$ ikinci momentinin tahmin edicisi olmak üzere yukarıdaki limit ifadesine bağlı olarak yeterince büyük $t$ için

$$\hat{M}_a(t)=\frac{t}{\hat{\mu}}+\frac{\hat{\mu}_2}{2\hat{\mu}^2}-1 \tag{6.7}$$

ile tanımlanan $\hat{M}_a(t)$ $M(t)$ için bir tahmin edici olarak göz önüne alınabilir. $F$ dağılım fonksiyonu biliniyor fakat parametreleri bilinmiyor ise $\mu$ ve $\mu_2$ parametrik olarak tahmin edilir. Eğer $F$ dağılım fonksiyonu bilinmiyor ise $\hat{F}_{KM}(t)$ Kaplan-Meier tahmin edicisine bağlı olarak $\mu$ ve $\mu_2$ sırasıyla

$$\hat{\mu}=\int_0^\infty t\,d\hat{F}_{KM}(t)$$

ve

$$\hat{\mu}_2=\int_0^\infty t^2\,d\hat{F}_{KM}(t)$$

ile elde edilebilir.

## 6.4 Bir Örnek

$M(t)$ yenileme fonksiyonunun sansürlü verilerde tahmini için tanımlanan $\hat{M}(t)$, $\hat{M}_{KM}(t)$ ve $\hat{M}_a(t)$ tahmin edicilerinin nasıl hesaplandığını göstermek ve birbirleriyle karşılaştırmak amacıyla aşağıda verilen veriler kullanılacaktır.

Bir üretici firma tarafından yeni bir ürün piyasaya sürülmektedir. İlk ürünün satılma zamanından itibaren 12 ay boyunca satılan ürünlerin bu süredeki ömürleri (ay olarak) gözlenmektedir. Başlangıçtan itibaren rasgele farklı zamanlarda satılan ürünlere ilişkin sağdan rasgele sansürlü gözlemler aşağıda verilmektedir.

Çizelge 6.1 Başlangıçtan itibaren rasgele farklı zamanlarda satılan ürünlere ilişkin sağdan rasgele sansürlü gözlemler

| i | Satılma Zamanı | i.ürün için ömür( $X_i$ ) | $\delta_i$ |
|---|---|---|---|
| 1 | 0.00 | 7.92 | 1 |
| 2 | 1.09 | 4.55 | 1 |
| 3 | 2.45 | 6.77 | 1 |
| 4 | 3.12 | 6.77 | 1 |
| 5 | 4.49 | 7.51 | 0 |
| 6 | 4.92 | 7.08 | 0 |
| 7 | 5.41 | 6.59 | 0 |
| 8 | 5.92 | 6.08 | 0 |
| 9 | 6.85 | 5.15 | 0 |
| 10 | 10.05 | 1.95 | 0 |
| 11 | 7.92 | 4.08 | 0 |
| 12 | 4.55 | 7.45 | 0 |
| 13 | 6.77 | 5.23 | 0 |
| 14 | 6.77 | 5.23 | 0 |

Yukarıdaki tabloda bulunan $\delta_i$ ’nin 1 olması $i.$ürünün 12 aydan önce bozulduğunu (sansürsüz) gösterirken 0 olması 12. ayda hala çalıştığını (sansürlü) gösterir.

Firma 12 aydan sonra en azından garanti masrafını tahmin etmek için $t$ zamanına kadar olan bozulmalarının sayısı olan $M_n(t)$ ’yi tahmin etmek isteyebilir.

İlk olarak $\hat{M}_n(t)$ parametrik tahmin edicisinin nasıl hesaplanacağını göstermek amacıyla $F$ dağılım fonksiyonunu

$$f(x) = \frac{\alpha}{\beta^\alpha} x^{\alpha-1} e^{-(x/\beta)^\alpha} \qquad , \quad x > 0; \quad \alpha, \beta > 0$$

olasılık yoğunluk fonksiyonu ile bir Weibull dağılım fonksiyonu olarak ele alalım. Bu gözlemlere dayalı Weibull dağılımının $\alpha$ ve $\beta$ parametrelerinin en çok olabilirlik tahmin değerleri Newton Raphson yöntemi yardımıyla sırasıyla $\hat{\alpha} = 7.9272$ ve $\hat{\beta} = 7.8533$ olarak bulunur. Daha sonra $t = 12, 24$ için $\hat{M}_{14}(t)$, (6.3) integral denkleminin RS yöntemi yardımıyla sayısal olarak çözdürülmesiyle $\hat{M}_{14}(12) = 1.0476$ ve $\hat{M}_{14}(24) = 2.8386$ elde edilir.

$\alpha$ ve $\beta$'nın en çok olabilirlik tahmin değerlerinin ve $\hat{M}_{14}(12)$ ile $\hat{M}_{14}(24)$'ün (6.3) integral denkleminden hesaplanmasına ilişkin Matlab programı EK 5'de verilmiştir.

Şimdi $F$ dağılım fonksiyonunun bilinmediğini kabul edelim. Bu durumda $M(t)$'yi tahmin etmek için $\hat{M}_{KM}(t)$ ya da $M_a(t)$ parametrik olmayan tahmin edicileri kullanılabilir. Veriler 100 ile çarpılarak yeniden ölçeklendirilir. (6.6) ifadesi yardımıyla $\hat{M}_{KM}(12) = 1.0579$ ve $\hat{M}_{KM}(24) = 3.0216$ tahmin değerleri elde edilir. Bu hesaplanmaları yapan Matlab programı EK 6'da verilmiştir.

$\hat{F}_{KM}(t)$ Kaplan-Meier tahmin edicisine bağlı olarak

$$\hat{\mu} = \int_0^\infty t d\hat{F}_{KM}(t) = 7.2878$$

ve

$$\hat{\mu}_2 = \int_0^\infty t^2 d\hat{F}_{KM}(t) = 54.0625$$

bulunduğundan $M(t)$ için (6.7)'de verilen $\hat{M}_a(t)$ tahmini

$$\hat{M}_a(t) = \frac{t}{7.2878} - 0.49105$$

olarak elde edilir. Bu durumda $\hat{M}_a(12) = 1.1555$ ve $\hat{M}_a(24) = 2.8021$ olur. Bu tahmin değerlerini hesaplatan Matlab programı EK 7'de verilmiştir.

# 7. SONUÇ

Bu çalışmada yenileme süreçlerine temel olan bazı kavramlar tanıtılarak, yenileme teorisi üzerinde durulmuştur. Daha sonra yenileme fonksiyonunun hesabı analitik olarak elde edilebilen bazı dağılımlar için yapılmış, analitik olarak elde edilemeyenler için ise Xie'nin RS yöntemi verilmiştir. Tezin konusu doğrultusunda sansürlü veriler tanıtılıp, yenileme süreçlerinde sık kullanılan sansürleme çeşidi olan sağdan rasgele sansürleme türü anlatılmıştır. Sağdan rasgele sansürlü örnekleme durumunda dağılım şekilsel olarak bilinip fakat parametreleri bilinmediği durumda parametre tahmini en çok olabilirlik tahmin yöntemi ile yapılmıştır. Daha sonra $F$'nin hiç bilinmediği durumda $F$'nin değerinin Kaplan-Meier tahmini yapılmıştır.

Sonuç olarak bu çalışmanın amacı doğrultusunda sağdan rasgele sansürlü verilerde ilk olarak $F$'in şekilsel olarak bilinip fakat parametreleri bilinmediği durumda yenileme fonksiyonunun parametrik bir tahmin edicisi elde edilip bu tahmin edicinin tutarlılık ve yansızlığı incelenmiştir. Ardından $F$'nin hiç bilinmediği durumdaki yenileme fonksiyonunun Kaplan-Meier tahmin edicisi elde edilmiştir. Son olarak ise yenileme fonksiyonunun asimptotik bir tahmin edicisi elde edilmiştir. Elde edilen bu tahmin edicileri bir örnek üzerinde uygulamasını yaparak karşılaştırılma imkanı sağlanmıştır.

## KAYNAKLAR


Aydoğdu, H. 1997. Yenileme Süreçlerinde Tahmin. Doktora Tezi. Ankara Üniversitesi , 153s, Ankara.

Barlow, E. R. and Proschan, F. 1965. Mathematical Theory of Reliability. John Willey& Sons, 290. New York.

Baxter, L. A. 1981. Some Remarks on Numerical Convolution. Communications in Statistics, Ser. B(10), 281-288

Baxter, L. A., Scheuer, E. M., McConalogue, D. J. and Blinschke, W. R. 1982. On the Tabulation of the Renewal Function. Technometrics, Volume 24, 151-158.

Binshan, L. 1988. Estimation of the Renawal Function. The İnterdepartmental Program in Business Administration. 84. Lousiana.

Cox, D. R. 1962. Renewal Theory. Methen and Co, London.

Feller, E. W. 1971. An Introduction to Probability Theory and Its Applications. John Wiley & Sons, Volume II, New York.

Frees, E. W. 1986a. Warranty and Renewal Function Estimation. Nov. Res. Logist. Q.,33,361;372

Gertsbakh, I. B. 1989. Statistical Reliability Theory. Marcel Dekker, 331, New York.

Frees, E. W. 1986b. Nonparemetrik Renawal Function Estimation. Ann. Statist., Volume 14 (4), 1366-1378.

Grimmett, G. R. and Stirzaker, D. R. 1992. Probability and Random Processes. Oxford University Pres Inc, New York.

Kawata, T. 1972. Fourier Analysis in Probability Theory. Acedemic Pres,New York.

Kaplan, E. L. and Meier, P. Nonparametric Estimation From Incomplete Observations. Journal of the American Statistical Association, Volume 53 (282), 457-481.

Karlin, S. and Taylor, H. M. 1975. A First Course in Stochastic Processes. Second edition. Acedemic Press, 557, New York.

Lawless, J. F. 2003. Statistical Models and Methods for Lifetime Data. John Wiley & Sons, 630, Canada.

Parzen, E. 1962. Stochastic Processes. Holden-Day, San Fransisco.

Ross, M. S. 1983. Stochastic Processes. John Wiley & Sons, 510, New York.

Schneider, H., Lin, B. S. and O'Cinneide, C. 1990. Comparison of Nonparemetrik Estimators for the Renewal Function . J. Roy. Statist. Soc. Ser. C39,55;61.

Smith, W. L. 1958. Renewal Theory and Its Ramifications. J. Roy. Statis. Soc. B, Volume 20, 243-302.

Tijms, 1994. Stochastic Models: An Algorithmic Approach. John Wiley & Sons

Topçu, Ç. 2007. Greenwood ve Kaplan-Meier Metodu Yardımı ile Varyans Tahmini. Yüksek Lisans Tezi. Ankara Üniversitesi, 48s, Ankara.

Xie, M. 1986a. On the Solution of Renewal-Type Integral Equation. Commun. Statist. Simula. Volume18 (1), 281-293.

Xie, M. 1986b. Some Results on the Renewal Equations. Commun. Statist. Theory Meth. Volume18 (3), 1159-171.

# EKLER

**EK 1** Weibull Dağılımı için Sağdan Rasgele Sansürlenmiş Gözlemlerin En Çok Olabilirlik Tahmin Edicilerini Hesaplatan Matlab Programı

**EK 2** Log-Normal Dağılımı için Sağdan Rasgele Sansürlenmiş Gözlemlerin En Çok Olabilirlik Tahmin Edicilerini Simulasyon Yoluyla Hesaplatan Matlab Programı

**EK 3** Gamma Dağılımı için Sağdan Rasgele Sansürlenmiş Gözlemlerin En Çok Olabilirlik Tahmin Edicilerini Simulasyon Yoluyla Hesaplatan Matlab Programı

**EK 4** Verilen Cihazın Bozulma Zamanlarına İlişkin Kaplan-Meier Tahmin Edicileri

**EK 5** Weibull Dağılımı Sahip Sağdan Rasgele Sansürlenmiş Gözlemlere Dayalı Yenileme Fonksiyonunu Hesaplatan Matlab Programı

**EK 6** $F$ Dağılım Fonksiyonu Hakkında Hiçbir Bilgimizin Olmadığı Durumda Yenileme Fonksiyonunun Tahminini Hesaplatan Matlab Programı

**EK 7** Yenileme Fonksiyonunun Asimptotik Hesabını Yapan Matlab Programı

**EK 1 Weibull Dağılımı için Sağdan Rasgele Sansürlenmiş Gözlemlerin En Çok Olabilirlik Tahmin Edicilerini Hesaplatan Matlab Programı**

```
clear
%alfa ve beta parametreli Weibull dağılımının
%rasgele sansürlenmiş gözlemler ile parametrelerinin ençok olabilirlik
%tahmin edicilerinin Newton-Raphson yöntemi yardımıyla hesabı
%c değeri sanürlenmemiş gözlemlerin sayısı ve d değeri bu gözlemlerin ortalamasıdır
x=[7.92 4.55 6.77 6.77 7.51 7.08 6.59 6.08 5.15 1.95 4.08 7.45 5.23 5.23];
c=4;
T=0;
for i=1:c
T=T+log(x(i));
end
d=T/c;
a(1)=3;
for j=1:5000
T1=sum(x.^a(j)); T2=sum((x.^a(j)).*log(x)); T3=sum((x.^a(j)).*(log(x).^2));
R(j)=(1/a(j))+d-(T2/T1);
S(j)=(-1/(a(j)^2))-(T3/T1)+((T2/T1)^2);
a(j+1)=a(j)-(R(j)/S(j));
% Newton Raphson yönteminde iterasyonların durdurulmasında yeterince yakınlığın
ölçüsü 10^-5 olarak alınır.
if abs(a(j+1)-a(j))<.00001
ALFA=a(j+1);
break
end
end
BETA=(T1/c)^(1/ALFA);
ALFA
BETA
```

## EK 2 Log-Normal Dağılımı için Sağdan Rasgele Sansürlenmiş Gözlemlerin En Çok Olabilirlik Tahmin Edicilerini Simulasyon Yoluyla Hesaplatan Matlab Programı

```
clear
%Simülasyon-lognormal(mü=0,sigma=1, b.d.=1.65)
%Sansürleme rasgele değişkeninin dağılımı için (0,8) aralığındaki düzgün
%dağılım kullanılır
%n gözlem sayısı
%mü ve sigma parametreli lognormal dağılımının birinci tip durdurulmuş
%(rasgele sansürleme) gözlemler ile parametrelerinin ençok olabilirlik
%tahmin edicilerinin Newton-Raphson yöntemi yardımıyla hesabı
n=50;
y=lognrnd(0,1,n,1);
t=8*rand(n,1);
x=min(y,t);
a=(y<t);
b=sum(a);
c=x.*a;
c=sort(c);
for i=1:b
sy(i)=c(n-b+i);
end
% sy sansürsüz gözlemler vektörü ve b sansürsüz gözlem sayısı
a1=1-(y<t);
b1=sum(a1);
% eğer b1 sıfır ya da denk olarak b n örnek hacmine eşit ise programı durdur
c1=x.*a1;
c1=sort(c1);
for i=1:b1
sv(i)=c1(n-b1+i);
end
% sv sansürlü gözlemler vektörü ve b1 sansürlü gözlem sayısı
%[y t x a]
%b, b1
%sy
%sv
%Aşağıda Newton-Raphson yöntemiyle mü ve sigma parametrelerinin ençok
%olabilirlik tahmin değerleri elde edilir
%Başlangıç değerleri
m(1)=0; s(1)=1
for i=1:10000
pv=[m(i);s(i)];
ysy=(log(sy)-m(i))./s(i);
ysv=(log(sv)-m(i))./s(i);
u(i)=(sum(ysy)+sum(normpdf(ysv)./(1-normcdf(ysv))))/s(i);
v(i)=(-b+sum(ysy.^2)+sum(ysv.*normpdf(ysv)./(1-normcdf(ysv))))/s(i);
```

```
U=[u(i);v(i)];
 % U  kolon vektörü olabilirlik fonksiyonunun logaritmasının herbir parametresine
 göre  kismi türevleri
t1=sum(ysv.*normpdf(ysv)./(1-normcdf(ysv)));
t2=sum((normpdf(ysv).^2)./((1-normcdf(ysv)).^2));
t3=sum(ysy.^2);
t4=sum(ysv.*normpdf(ysv).*(-2+(ysv.^2))./(1-normcdf(ysv)));
t5=sum((ysv.^2).*(normpdf(ysv).^2)./((1-normcdf(ysv)).^2));
t6=sum(ysy);
t7=sum(normpdf(ysv)./(1-normcdf(ysv)));
t8=sum((ysv.^2).*normpdf(ysv)./(1-normcdf(ysv)));
t9=sum(ysv.*(normpdf(ysv).^2)./((1-normcdf(ysv)).^2));
A(1,1)=(-b+t1-t2)/((s(i)^2));
A(1,2)=((-2*t6)-t7+t8-t9)/((s(i)^2));
A(2,1)=A(1,2);
A(2,2)=(b-(3*t3)+t4-t5)/((s(i)^2));
% Newton Raphson yönteminde iterasyonların durdurulmasında yeterince yakınlığın
ölçüsü 10^-5  olarak alınır.
C=pv-inv(A)*U;
m(i+1)=C(1,1);
s(i+1)=C(2,1);
if abs([m(i+1);s(i+1)]-[m(i);s(i)])<.00001
p1=m(i+1);
p2=s(i+1);
break
end
end
p1
p2
%p1 mü ve p2 sigmadır
```

**EK 3 Gamma Dağılımı için Sağdan Rasgele Sansürlenmiş Gözlemlerin En Çok Olabilirlik Tahmin Edicilerini Simulasyon Yoluyla Hesaplatan Matlab Programı**

```
clear
%Simülasyon- Gamma(alfa=2,beta=1, beklenen değer=2)
%Sansürleme rasgele değişkeninin dağılımı (0,8) aralığındaki düzgün
%dağılımdır
%n gözlem sayısı
%alfa ve beta parametreli gamma dağılımının birinci tip durdurulmuş
%(rasgele sansürleme) gözlemler ile parametrelerinin ençok olabilirlik
%tahmin edicilerinin Newton-Raphson yöntemi yardımıyla hesabı
%M(t) nin hesabı için RS-yönteminde  0.005 eşit adım uzunluğu kullanılmıştır
alfa=2; beta=1;
n=50;
y=gamrnd(alfa,beta,n,1);
t=8*rand(n,1);
x=min(y,t);
a=(y<t);
b=sum(a);
c=x.*a;
c=sort(c);
for i=1:b
   sy(i)=c(n-b+i);
end
% sy sansürsüz gözlemler vektörü ve b sansürsüz gözlem sayısı
a1=1-(y<t);
b1=sum(a1);
% Eğer b1 sıfır ya da denk olarak b n örnek hacmine eşit ise programı durdur
c1=x.*a1;
c1=sort(c1);
for i=1:b1
sv(i)=c1(n-b1+i);
end
% sv sansürlü gözlemler vektörü ve b1 sansürlü gözlem sayısı
%[y t x a]
%b, b1
%sy
%sv
%Aşağıda Newton-Raphson yöntemiyle alfa ve beta parametrelerinin ençok
%olabilirlik tahmin değerleri elde edilir
%Başlangıç değerleri
%af(1)=2; be(1)=1;
af(1)=(mean(sy)/std(sy,1))^2
be(1)=(std(sy,1)^2)/mean(sy)
for i=1:10000
pv=[af(i);be(i)];
```

```
for j=1:b1
xx=10^(-15):0.001:sv(j)/be(i);
yy=exp(-xx).*(log(xx)).*(xx.^(af(i)-1));
yz=exp(-xx).*((log(xx)).^2).*(xx.^(af(i)-1));
k1(j)=trapz(xx,yy)/gamma(af(i));
k2(j)=trapz(xx,yz)/gamma(af(i));
end
t1=sum(log(sy));
t2=sum(gamcdf(sv,af(i),be(i))./(1-gamcdf(sv,af(i),be(i))));
t3=sum(k1./(1-gamcdf(sv,af(i),be(i))));
t4=sum(sy);
t5=sum(gampdf(sv,af(i)+1,be(i))./(1-gamcdf(sv,af(i),be(i))));
u(i)=(-b*(psi(af(i))+log(be(i))))+t1+(psi(af(i))*t2)-t3;
v(i)=(-b*af(i)/be(i))+(t4/(be(i)^2))+af(i)*t5;
U=[u(i);v(i)];
% U  kolon vektörü olabilirlik fonksiyonunun logaritmasının herbir parametresine
göre kismi türevleri
t6=sum(log(sv/be(i)).*gampdf(sv,af(i)+1,be(i))./(1-gamcdf(sv,af(i),be(i))));
t7=sum(gamcdf(sv,af(i),be(i)).*gampdf(sv,af(i)+1,be(i))./((1gamcdf(sv,af(i),be(i))).^2));
t8=sum(k1.*gampdf(sv,af(i)+1,be(i))./((1-gamcdf(sv,af(i),be(i))).^2));
t9=sum((-1-af(i)+(sv/be(i))).*gampdf(sv,af(i)+1,be(i))./(1-gamcdf(sv,af(i),be(i))));
t10=sum((gampdf(sv,af(i)+1,be(i))./(1-gamcdf(sv,af(i),be(i))).^2));
t11=sum((k1-(psi(af(i))*gamcdf(sv,af(i),be(i))))./((1-gamcdf(sv,af(i),be(i))).^2));
t12=sum(k2./(1-gamcdf(sv,af(i),be(i))));
t13=sum((k1-psi(af(i))).*k1./((1-gamcdf(sv,af(i),be(i))).^2));
A(1,1)=((-b+t2)*psi(1,af(i)))+(psi(af(i))*t11)-t12-t13;
A(1,2)=(-b/be(i))-(af(i)*psi(af(i))*t5)+(af(i)*t6)-(af(i)*psi(af(i))*t7)+(af(i)*t8);
A(2,1)=A(1,2);
A(2,2)=(b*af(i)/(be(i)^2))-(2*t4/(be(i)^3))+(af(i)*t9/be(i))-((af(i)^2)*t10);
 C=pv-inv(A)*U;
 af(i+1)=C(1,1);
 be(i+1)=C(2,1);
 if af(i+1)<0
 af(i+1)=-af(i+1)
 end
 if abs([af(i+1);be(i+1)]-[af(i);be(i)])<.00001
 p1=af(i+1);
 p2=be(i+1);
 break
 end
 end
 p1
 p2
 %p1 alfa ve p2 betadır
```

**EK 4 Verilen Cihazın Bozulma Zamanlarına İlişkin Kaplan-Meier Tahmin Edicileri**

```
clear
% x gözlem vektörü
x=[36.3 41.7 43.9 49.9 50.1 50.8 51.9 52.1 52.3 52.3 52.4 52.6 52.7 53.1 53.6 53.6 53.9
53.9 54.1 54.6 54.8 54.8 55.1 55.4 55.9 56 56.1 56.5 56.9 57.1 57.1 57.3 57.7 57.8 58.1
58.9 59 59.1 59.6 60.4 60.7 26.8 29.6 33.4 35 40 41.9 42.5];
x=sort(x);
% s sansürsüz gözlem vektörü
s=[36.3 41.7 43.9 49.9 50.1 50.8 51.9 52.1 52.3 52.3 52.4 52.6 52.7 53.1 53.6 53.6 53.9
53.9 54.1 54.6 54.8 54.8 55.1 55.4 55.9 56 56.1 56.5 56.9 57.1 57.1 57.3 57.7 57.8 58.1
58.9 59 59.1 59.6 60.4 60.7];
s=sort(s);
% aşağıda farklı sansürsüz gözlemler küçükten büyüğe sıralanır ve herbir gözlemin
% tekrar sayısı elde edilir
v=histc(s,s);
v1=max(v);
l=length(v);
r=sum(v>0);
%v1 en fazla tekrar sayısı
u=0;
for j=1:r
for i=u+1:l
if  v(i)>0
a(j)=v(i);
u=i;
break
end
end
end
for i=1:l
for j=1:v1
if v(i)==j
v(i)=v(i)/j;
break
end
end
end
h=sort(v.*s);
g=sum(h==0);
for i=1:r
s1(i)=h(g+i);
end
 %[s1' a']
% s1 farklı sansürsüz gözlemler vektörü ve a bu gözlemlerin tekrar
% sayılarını veren vektördür
```

```
%aşağıda Kaplan-Meier tahmin edicisinin z noktasındaki değeri elde edilir
z=41.7;
n(1)=length(x); d(1)=0; T=1; k=sum(s1<=z);
for i=2:k+1
n(i)=sum(x>=s1(i-1)); d(i)=a(i-1);
T=T*(1-(d(i)/n(i)));
end
 T
```

**EK 5 Weibull Dağılımı Sahip Sağdan Rasgele Sansürlenmiş Gözlemlere Dayalı Yenileme Fonksiyonunu Hesaplatan Matlab Programı**

```
clear
%alfa ve beta parametreli Weibull dağılımının birinci tip durdurulmuş
%(rasgele sansürleme) gözlemler ile parametrelerinin ençok olabilirlik
%tahmin edicilerinin Newton-Raphson yöntemi yardımıyla hesabı
%c değeri sanürlenmemiş gözlemlerin sayısı ve d değeri bu gözlemlerin ortalamasıdır
x=[7.92 4.55 6.77 6.77 7.51 7.08 6.59 6.08 5.15 1.95 4.08 7.45 5.23 5.23];
c=4;
T=0;
for i=1:c
T=T+log(x(i));
end
d=T/c;
a(1)=3;
for j=1:5000
T1=sum(x.^a(j));T2=sum((x.^a(j)).*log(x));
T3=sum((x.^a(j)).*(log(x).^2));
R(j)=(1/a(j))+d-(T2/T1);
S(j)=(-1/(a(j)^2))-(T3/T1)+((T2/T1)^2);
a(j+1)=a(j)-(R(j)/S(j));
if abs(a(j+1)-a(j))<.00001
ALFA=a(j+1);
break
end
end
BETA=(T1/c)^(1/ALFA);
ALFA
BETA
z=24;
k=z/.005;
 H(1)=0;
 F1=1-exp(-((.0025/BETA)^ALFA));
 for I=2:k+1
 F(I)=1-exp(-((I-.5)*.005/BETA)^ALFA);
 H(I)=1-exp(-((I-1)*.005/BETA)^ALFA);
 end
 for I=2:k+1
 for J=2:I
 H(I)=H(I)+F(I-J+1)*(H(J)-H(J-1));
 end
 H(I)=(H(I)-F1*H(I-1))/(1-F1);
 end
 for I=1:k
 H(I)=H(I+1);
 end
```

```
 for i=1:k/20
 M(i)=H(20*i);
%Yenileme fonksiyonunun tahmini
 end
Z=0.1:0.1:z;
[Z' M']
```

**EK 6 $F$ Dağılım Fonksiyonu Hakkında Hiçbir Bilgimizin Olmadığı Durumda Yenileme Fonksiyonunun Tahminini Hesaplatan Matlab Programı**

```
clear
% x gözlem vektörü
x=[7.92 4.55 6.77 6.77 7.51 7.08 6.59 6.08 5.15 1.95 4.08 7.45 5.23 5.23];
x=sort(x);
% s sansürsüz gözlem vektörü
s=[7.92 4.55 6.77 6.77];
s=sort(s);
% aşağıda farklı sansürsüz gözlemler küçükten büyüğe sıralanır ve herbir gözlemin
% tekrar sayısı elde edilir
% hs ölçeklendirme sayısı
hs=100;
x=round(hs*x);
s=round(hs*s);
v=histc(s,s);
v1=max(v);
l=length(v);
r=sum(v>0);
%v1 en fazla tekrar sayısı
u=0;
for j=1:r
for i=u+1:l
if  v(i)>0
a(j)=v(i);
u=i;
break
end
end
end
for i=1:l
for j=1:v1
if v(i)==j
v(i)=v(i)/j;
break
end
end
end
h=sort(v.*s);
g=sum(h==0);
for i=1:r
s1(i)=h(g+i);
end
%[s1' a']
% s1 farklı sansürsüz gözlemler vektörü ve a bu gözlemlerin tekrar
% sayılarını veren vektördür
```

```
%aşağıda Kaplan-Meier tahmin edicisinin z noktasındaki değeri elde edilir
n(1)=length(x); d(1)=0;
for j=1:length(s1);
T=1; k=sum(s1<=s1(j));
for i=2:k+1
n(i)=sum(x>=s1(i-1)); d(i)=a(i-1);
T=T*(1-(d(i)/n(i)));
end
KM(j)=T;
End
F=1-KM;
us=24*hs;
Fa(1)=0;
for i=2:s1(1)
Fa(i)=0;
end
s1(r+1)=us+1;
for j=1:r
for i=s1(j)+1:s1(j+1)
Fa(i)=F(j);
end
end
M(1)=Fa(2);
for i=2:us
tt=0;
for j=1:i-1
tt=tt+(M(i-j)*(Fa(j+1)-Fa(j)));
end
M(i)=tt+Fa(i+1);
end
M(hs*12), M(hs*24)
```

**EK 7 Yenileme Fonksiyonunun Asimptotik Hesabını Yapan Matlab Programı**

```
clear
% x gözlem vektörü
x=[7.92 4.55 6.77 6.77 7.51 7.08 6.59 6.08 5.15 1.95 4.08 7.45 5.23 5.23];
x=sort(x);
% s sansürsüz gözlem vektörü
s=[7.92 4.55 6.77 6.77];
s=sort(s);
% aşağıda farklı sansürsüz gözlemler küçükten büyüğe sıralanır ve herbir gözlemin
% tekrar sayısı elde edilir
v=histc(s,s);
v1=max(v);
l=length(v);
r=sum(v>0);
%v1 en fazla tekrar sayısı
u=0;
for j=1:r
for i=u+1:l
if  v(i)>0
a(j)=v(i);
u=i;
break
end
end
end
for i=1:l
for j=1:v1
if v(i)==j
v(i)=v(i)/j;
break
end
end
end
h=sort(v.*s);
g=sum(h==0);
for i=1:r
s1(i)=h(g+i);
end
%[s1' a']
% s1 farklı sansürsüz gözlemler vektörü ve a bu gözlemlerin tekrar
% sayılarını veren vektördür
%aşağıda Kaplan-Meier tahmin edicisinin z noktasındaki değeri elde edilir
n(1)=length(x); d(1)=0;
for j=1:length(s1);
T=1; k=sum(s1<=s1(j));
for i=2:k+1
```

```
n(i)=sum(x>=s1(i-1)); d(i)=a(i-1);
T=T*(1-(d(i)/n(i)));
end
KM(j)=T;
end
ft=1-KM;
ft
plot(s1,ft,'.')
%plot(s1,ft)
%z=input ('zaman=');
%T=1; k=sum(s1<=z);
%for i=2:k+1
%n(i)=sum(x>=s1(i-1)); d(i)=a(i-1);
%T=T*(1-(d(i)/n(i)));
%end
%1-T
% bdt ve vt sırasıyla K-M tahmin edicisine bağlı olarak dağılımın
% beklenen değer ve varyansının tahminidir
b(1)=s1(1)*ft(1);
for i=2:length(s1)
b(i)=s1(i)*(ft(i)-ft(i-1));
end
bdt=sum(b)
%b1(1)=s1(1)*1;
%for i=2:length(s1)
%b1(i)=(s1(i)-s1(i-1))*KM(i-1);
%end
%sum(b1)
b2(1)=(s1(1)^2)*ft(1);
for i=2:length(s1)
b2(i)=(s1(i)^2)*(ft(i)-ft(i-1));
end
bdt2=sum(b2)
vt=sum(b2)-(bdt^2);
vt, sqrt(vt)
```

# ÖZGEÇMİŞ

Adı Soyadı :Çiğdem DANIŞ
Doğum Yeri :Malatya
Doğum Tarihi :05.08.1979
Medeni Hali :Evli
Yabancı Dili :İngilizce

Eğitim Durumu(Kurum ve Yıl)

Lise : Kanuni Lisesi (1997)
Lisans : Hacettepe Üniversitesi Fen Fakültesi
İstatistik Bölümü (2004)
Yüksek Lisans: Ankara Üniversitesi Fen Bilimleri Enstitüsü
İstatistik Anabilim Dalı (Şubat 2006-Şubat 2008)

Çalıştığı Kurum / Kurumlar ve Yıl

M.E.B İstiklal İlköğretim Okulu (2004)
Özel Tümay Dersanesi(2004)
Yüzüncü Yıl Üniversitesi Fen Fakültesi İstatistik Bölümü (2005)
Ankara Üniversitesi Fen Fakültesi İstatistik Bölümü(2006-)