



2016.04.01.STAT.02

## UNSHARED AND SHARED FRAILTY MODELS

Nihal ATA TUTKUN \*

Assoc. Prof. Dr., Hacettepe University, Ankara

Diren YEĞEN †

Republic of Turkey, Ministry of Culture and Tourism

Received: 28 December 2016

Accepted: 27 April 2016

### Abstract

The Cox regression model which is commonly used in survival analysis is established under the proportional hazards assumption. However cases in which the data shows heterogeneity come across in studies. In this case, heterogeneity should be explained in order to make the interpretations more effective which were obtained depending on the model. Frailty models are one of the survival analysis methods which were developed for explaining heterogeneity.

In this study, frailty models are examined theoretically and were applied to the lung cancer data. The unshared frailty model has been used to explain the difference between general risk and momentary risk of individuals in the data set. As for comparing the momentary risk between individuals with various levels of explanatory variables with other individuals, shared frailty models have been used.

*Keywords: Survival analysis, Cox regression, nonproportional hazards, parametric regression models, frailty models*

## PAYLAŞILMAMIŞ VE PAYLAŞILMIŞ ZAYIFLIK MODELLERİ

### Özet

Yaşam çözümlemesinde sıklıkla kullanılan Cox regresyon modeli orantılı tehlikeler varsayımı altında kurulmaktadır. Ancak çalışmalarda verinin heterojen özellik gösterdiği durumlar ile karşılaşmaktadır. Bu durumda modele bağlı olarak elde edilen yorumların daha etkin olabilmesi için heterojenliğin açıklanması gerekmektedir. Zayıflık modelleri heterojenliğin açıklanması için geliştirilmiş bir yaşam çözümlemesi yöntemidir.

Bu çalışmada, zayıflık modelleri teorik açıdan incelenmiş ve akciğer kanseri verisi kullanılarak bir uygulama yapılmıştır. Veri kümesindeki bireylerin taşıdığı genel risk ile herhangi bir bireyin anlık riski arasındaki farklılığı açıklamada paylaşılmamış zayıflık modeli kullanılmıştır. Açıklayıcı değişkenlerin çeşitli düzeylerine sahip bireylerin veri kümesindeki diğer bireylere göre anlık riskinin karşılaştırılmasında ise paylaşılmış zayıflık modelleri kullanılmıştır.

*Anahtar Kelimeler : Yaşam çözümlemesi, orantısız tehlikeler, parametrik regresyon modelleri, zayıflık modelleri*

\* [nihalata@hacettepe.edu.tr](mailto:nihalata@hacettepe.edu.tr) (Corresponding author)

† [dy11163@hotmail.com](mailto:dy11163@hotmail.com)

## 1. GİRİŞ

Yaşam çözümlemesi mühendislik, tıp, biyoloji ve demografi gibi bilim dallarında kullanılan temel bir araştırma yöntemidir. Yaşam çözümlemesi, tıbbi ve demografik çalışmalarda incelenen ölümlülük kavramının bir karşılığı olarak ortaya çıkmıştır. İlgilenilen olayın ortaya çıkma süresine yani başarısızlık süresi verilerine dayanan bu araştırma yöntemi Cox (1974)'un geliştirdiği Orantılı Tehlikeler Modeli yaklaşımıyla beraber geniş bir uygulama alanına yayılmıştır.

Yaşam çözümlemesinde, bağımlı değişken olarak ele alınan yaşam süresinin açıklayıcı değişkenler tarafından etkilenebileceği göz önünde bulundurulduğunda, regresyon modellerinin yaşam çözümlemesinde önemli bir yere sahip olduğu görülmektedir. Bu modellerden biri de zayıflık (frailty) modelidir. Zayıflık modeli özellikle tıp, biyoloji ve genetik çalışmalarının da içinde bulunduğu çeşitli alanlarda kullanılmaktadır. Orantılı tehlikeler modelinde aynı değişken değerine sahip olan birimlerin aynı yaşam süresine sahip olacağı varsayılmaktadır. Ancak bu durum, aynı tedavi, yaş ve cinsiyet grubundaki gözlenen tüm bireylerin aynı gözlenen yaşam sürelerine sahip olduğu anlamına gelmemektedir. Bazı çalışmalarda, ölçülen açıklayıcı değişkenler dışında yaşam süresini önemli derecede etkileyen ancak gözlenemeyen başka faktörler de olabilir. Bu durum, birimlerin heterojenliği olarak belirtilmektedir. Zayıflık modelinin temeli, birimler arasındaki heterojenliği açıklamak için ölçülemeyen rasgele etkiyi modele dahil etmektir.

Bu çalışmanın amacı, zayıflık modellerinin yapısını ve türlerini incelemektir. İkinci bölümde, zayıflık modellerinde kullanılan yöntemler verilmiştir. Üçüncü bölümde ise akciğer kanseri verilerine klasik yaşam çözümlemesi yöntemlerinin yanı sıra zayıflık modelleri uygulanmış ve elde edilen sonuçlar incelenmiştir.

## 2. Zayıflık Modelleri

Yaşam çözümlemesinde zamana bağlı verilerin bağımsız ve aynı dağılımdan geldiği yani kitlenin homojen olduğu varsayılmaktadır. Fakat gözlemler incelenince birimlerin aynı dağılımdan gelmediği, bağımsız olmadığı aksine kendine özgü özellikleri olduğu söylenebilir. Bu farklılık ile kitle heterojen bir nitelik kazanmaktadır. Bu heterojenliği değerlendirmek zordur, ancak önemlidir. Heterojenlik başlıca şu iki nedenden kaynaklanabilir:

Gözlenebilen risk faktörlerinden kaynaklanan değişkenlik,

Bilinmeyen açıklayıcı değişkenlerden kaynaklanan heterojenlik.

Heterojenlik orantısız veya azalan tehlikeler gibi bazı beklenmedik sonuçları açıklamaktadır. Eğer bazı birimler yüksek başarısızlık riski taşıyorsa, diğerleri daha az riskli bir grup oluşturmaya eğilimlidirler.

Gözlenemeyen zayıflığı göz önüne almadan tahmin edilen bireysel tehlike oranı, zaman geçtikçe tehlike fonksiyonunda yapısal bir bozulmaya yol açmaktadır. Kitlenin farklı riskler taşıyan birimlerin karışımı olduğu varsayılırsa karma modeller kullanılabilir. Gözlenemeyen riskler zayıflık olarak tanımlanmaktadır. Bu zayıflık değeri bilinemediğinden tehlike fonksiyonuna çarpımsal olarak dahil edilmektedir. Çünkü birimler ile kitle arasındaki ilişkinin yapısı, zayıflığın birimler arasındaki dağılımına bağlıdır. Zayıflık terimi farklı dağılım türlerine uyabilir. Hougaard (1984, 1995), Clayton (1978), Yashin vd. (1995), Oakes (1967), Congdon (1995) çalışmalarında zayıflık terimi için en çok kullanılan dağılımlar, Gamma ve ters-Gauss dağılımlarıdır. Zayıflık dağılımının varyansı ise incelenen kitledeki heterojenlik derecesini belirlemektedir.

Zayıflık modeli ilk kez Vaupel (1979) tarafından mortalite çalışmalarında uygulanmıştır. Lancaster (1979) işsizlik sürelerinin modellenmesinde zayıflık modelini kullanmıştır. Zayıflık modeli ile ilgili çalışmalar birçok araştırmacının ilgisini çekmiştir (McGilchrist ve Aisbett, 1991; Guo ve Rodriquez, 1992, Yashin and Iachine, 1995; Babiker ve Cuzick, 1994; Hougaard, 2000; Wienke ve diğ., 2000, 2004; O'Quigley ve Stare, 2002, 2004). Bu konuyla ilgili çalışmalar özellikle biyoloji ve genetik çalışmalarının da içinde bulunduğu çeşitli bilim dallarında kullanılmıştır.

Zayıflık modelleri paylaşılmış ve paylaşılmamış zayıflık modelleri olmak üzere başlıca iki bölüme incelenmektedir.

### 2.1. Paylaşılmamış Zayıflık Modeli

Paylaşılmamış zayıflık modelleri başarısızlık süresi ilişkisiz olan birimleri incelemektedir. Bu yaklaşımda kitlenin, belli gözlenebilir değişkenler ile açıklandığı ve homojen olduğu varsayımı dikkate alınmaktadır. Örneğin, tıbbi bir çalışmada bireylerin farklı ilaçlara ya da uygulanan tedaviye farklı tepkiler verdiği gözlenebilir. Heterojenlik açıklanması zor fakat bir o kadar da gerekli bir durumdur. Yapılan çalışmalarda, zayıflık modelleri bu heterojenliği açıklamak adına kullanılmaktadır. Bu yaklaşımların temeli, bireylerin farklı zayıflıklara sahip olduğu ve en zayıf olanın, daha az zayıf olandan daha erken başarısızlığa uğrayacağıdır.

En çok uygulanan zayıflık modeli orantılı tehlikeler modelini öngörmektedir. Bu da rasgele etki üzerinde kullanılmıştır. Birimlere ait tehlike fonksiyonu; gözlenemeyen, zamandan bağımsız rasgele değişken olan  $Z$ 'ye bağlıdır. Bu rasgele değişken temel tehlike fonksiyonuna çarpımsal olarak etki eder ve

$$h(t, Z) = Zh_0(t) \quad (1)$$

biçimindedir. Burada  $Z$  kitle içinde değişkenlik gösteren rasgele değişken olarak ele alınır. Zayıflık modelleri  $E(Z)=1$  ve  $V(Z)=\sigma^2$  olacak biçimde standartlaştırılmıştır. Burada  $Z$ 'nin varyansı temel riskin içindeki heterojenliğin ölçüsü olarak yorumlanabilir. Eğer  $Z=1$  olursa zayıflık dağılımını içeren model orantılı tehlikeler modeline dönüşmektedir.

Eşitlik 1'de verilen çarpımsal zayıflık modeline göre, zayıflık zamandan bağımsız ve temel tehlike fonksiyonuna çarpımsal olarak etki etmektedir. Bu durum da yaşam çözümlenmesindeki gözlenemeyen heterojenliğe neden olmaktadır. Eşitlik 1'e bilinen açıklayıcı değişkenler eklenirse:

$$h(t, Z, X) = Zh_0(t) \exp(\beta^T X) \quad (2)$$

elde edilir. Burada  $X=(X_1, \dots, X_k)$  açıklayıcı değişkenler vektörü ve  $\beta=(\beta_1, \dots, \beta_k)$  regresyon parametreleri vektörüdür.  $S(t/Z)$ , zayıflık ( $Z$ ) kısıtı altında birimlerin yaşam fonksiyonunu göstermektedir ve

$$S(t/Z)$$

$$= \exp\left(-\int_0^t h(s, Z) ds\right) = \exp\left(-Z \int_0^t h_0(s) ds\right) = \exp(-ZH_0(t))$$

biçiminde elde edilir. Burada  $H_0(t) = \int_0^t h_0(s) ds$  temel

bikimli tehlike fonksiyonudur.

Kitlenin yaşam fonksiyonu, yaşam fonksiyonlarının ortalamasıdır. Yaşam fonksiyonu ve olasılık yoğunluk fonksiyonu zayıflık fonksiyonunun ortalaması ve varyansı ile nitelendirilir. Olabilirlik fonksiyonu ise  $(T_i, \Delta_i, X_i, Z_i)$  ve  $(i = 1, 2, \dots, n)$  için

$$\prod_{i=1}^n (Z_i h_0(t_i) \exp(\beta^T X_i))^{\delta_i} \exp(-Z_i h_0(t_i) \exp(\beta^T X_i))$$

biçiminde verilmektedir.

Zayıflığın her bir birim için sürecin tamamı boyunca belirlendiği varsayılmaktadır. Fakat kitle bileşenleri süreç boyunca değişikliğe uğrar. Zayıf birimler erken başarısızlığa eğilimlidir ve ilk başarısız olacak olanlardır. Buna bağlı olarak risk altındaki kitlenin zayıflık dağılımı süreç ilerledikçe değişmektedir.

Bu nedenle kitleye ait ortalama zayıflık  $\int_0^{\infty} z$

$f(z \setminus T > t) dz$ , zamanla azalacaktır. Aşağıdaki teorem bu yapıyı açıklamaktadır.

**Teorem:** Eşitlik 1 ile verilen model gözönüne alınırsa;

$$h(t) = \frac{f(t)}{S(t)} \text{ eşitliğinden yararlanarak,}$$

$h(t) = E(h(t, Z) | T > t)$  elde edilebilir. Daha açık biçimiyle;

$$h(t) = \int_0^{\infty} h(t, Z) f(z | T > t) dz = h_0(t) \int_0^{\infty} z f(z | T > t) dz$$

olur. Burada,  $f(z | T > t)$   $t$  anında hayatta kalanlara ait zayıflığın yoğunluk fonksiyonudur (Wienke, 2011).

## 2.2. Zayıflık Terimi için Kullanılan Dağılımlar

Uygulamalarda en çok rastlanan zayıflık terimi dağılımları; Gamma ve Ters-Gauss dağılımlarıdır.

### 2.2.1. Gamma Zayıflık Modeli

Gamma dağılımı uygulamalarda sıklıkla karşılaşılan dağılımlardan biridir. Gamma dağılımı yaşam verilerine iyi uyum sağlayan bir dağılımdır. Gamma dağılımının Laplace dönüşümü aracılığıyla, birikimli yoğunluk fonksiyonu, tehlike fonksiyonu rahatça açıklanabilmektedir. Gamma dağılımı parametreleri sayesinde  $\Gamma(k, \lambda)$ ,  $k=1$  olduğunda üstel dağılıma,  $k$  çok büyük seçildiğinde normal dağılıma benzer olması bu dağılımı kullanışlı biçime getirmektedir. Vaupel, Manton ve Stallard (1979) geliştirdikleri yaklaşımla zayıflık terimi  $Z \sim \Gamma(k, \lambda)$  olduğunda modelin yapısını ortaya koymuştur.  $k$  konum,  $\lambda$  ölçek parametresi olmak üzere  $Z$ 'nin marjinal yoğunluk fonksiyonu,

$$f(z) = \frac{\lambda^k z^{k-1} e^{-\lambda z}}{\Gamma(k)}; \quad z > 0, k > 0, \lambda > 0$$

olarak tanımlanmaktadır. Laplace dönüşümü,

$$L(u) = \frac{1}{\Gamma(k)} \lambda^k \int e^{-uz} z^{k-1} e^{-\lambda z} dz$$

$$= \frac{\lambda^k}{(\lambda + u)^k} \frac{1}{\Gamma(k)} (\lambda + u)^k \int z^{k-1} e^{-(\lambda+u)z} dz$$

biçimindedir.  $k$  ve  $(\lambda+u)$  parametreleriyle gamma dağılımının Laplace dönüşümünün birinci ve ikinci türevi,

$$L'(u) = -\frac{k}{\lambda} \left(1 + \frac{u}{\lambda}\right)^{-k-1},$$

$$L''(u) = \frac{k(k+1)}{\lambda^2} \left(1 + \frac{u}{\lambda}\right)^{-k-2}$$

olarak verilmektedir.  $u=0$  noktasında bu eşitlikler yardımıyla beklenen değer ve varyans,

$$E(Z) = \frac{k}{\lambda}, \quad V(Z) = \frac{k(k+1)}{\lambda^2} - \frac{k^2}{\lambda^2} = \frac{k}{\lambda^2}$$

biçimde elde edilmektedir.  $k=\lambda$  kısıtı altında bu dağılım

$$E(Z)=1 \text{ ve } \sigma^2 = \frac{1}{\lambda} \text{ olmaktadır.}$$

$Z \sim \Gamma\left(\frac{1}{\sigma^2}, \frac{1}{\sigma^2}\right)$  olan gamma dağılımının olasılık yoğunluk fonksiyonu;

$$f(z) = \frac{1}{\Gamma\left(\frac{1}{\sigma^2}\right)} \left(\frac{1}{\sigma^2}\right)^{\frac{1}{\sigma^2}} z^{\frac{1}{\sigma^2}-1} \exp\left(-\frac{z}{\sigma^2}\right)$$

olmakta ve koşulsuz yaşam fonksiyonu, koşulsuz olasılık yoğunluk fonksiyonu ve tehlike fonksiyonu ise sırasıyla aşağıda verilmektedir.

$$S(t) = L(H_0(t)) = \frac{1}{(1 + \sigma^2 H_0(t))^{\frac{1}{\sigma^2}}}$$

$$f(t) = \frac{h_0(t)}{(1 + \sigma^2 H_0(t))^{\frac{1}{\sigma^2}+1}},$$

$$h(t) = \frac{h_0(t)}{1 + \sigma^2 H_0(t)}.$$

Yaşayan bireyler/birimler üzerinden zayıflık dağılımının olasılık yoğunluk fonksiyonu,

$$f(z|X, T > t) = \frac{S(t|X, z)f(z)}{S(t|X)}$$

$$= \exp(-zH_0(t)e^{\beta^T X}) z^{\frac{1}{\sigma^2}-1} \exp\left(-z\left(\frac{1}{\sigma^2} + H_0(t)e^{\beta^T X}\right)\right)$$

$$= \frac{\left(\frac{1}{\sigma^2} + H_0(t)e^{\beta^T X}\right)^{\frac{1}{\sigma^2}}}{\Gamma\left(\frac{1}{\sigma^2}\right)} z^{\frac{1}{\sigma^2}-1} \exp\left(-z\left(\frac{1}{\sigma^2} + H_0(t)e^{\beta^T X}\right)\right)$$

olarak tanımlanmaktadır. Benzer biçimde  $t$  anında başarısızlığa uğrayanlar için olasılık yoğunluk fonksiyonu aşağıdaki biçimde verilmektedir:

$$f(z|X, T = t) = \frac{f(t|X, z)f(z)}{f(t|X)}$$

$$= \frac{zh_0(t) \exp(-zH_0(t)e^{\beta^T X}) z^{\frac{1}{\sigma^2}-1} \exp\left(-\frac{z}{\sigma^2}\right)}{\Gamma\left(\frac{1}{\sigma^2}\right) \sigma^{\frac{2}{\sigma^2}} h_0(t) (1 + \sigma^2 H_0(t)e^{\beta^T X})^{-\frac{1}{\sigma^2}-1}}$$

$$= \frac{\left(\frac{1}{\sigma^2} + H_0(t)e^{\beta^T X}\right)^{\frac{1}{\sigma^2}+1}}{\Gamma\left(\frac{1}{\sigma^2} + 1\right)} z^{\frac{1}{\sigma^2}+1-1} \exp\left(-z\left(\frac{1}{\sigma^2} + H_0(t)e^{\beta^T X}\right)\right)$$

Yukarıdaki eşitlikten yararlanarak  $t$  anındaki ölümlerin ortalaması;

$$E(Z|X, T = t) = \frac{1 + \sigma^2}{1 + \sigma^2 H_0(t)e^{\beta^T X}}$$

ve  $t$  anından sonraki ölümlerin ortalaması;

$$E(Z|X, T > t) = \frac{1}{1 + \sigma^2 H_0(t)e^{\beta^T X}}$$

olarak verilmektedir.  $t$  anında başarısız olan birimler için zayıflık daha yüksek bir ortalama vermektedir. Bu durum yüksek risk taşıyan birimlerin daha erken başarısızlığa

uğrayacağıının bir işaretidir. t anında başarısız olan birimlerin varyansı,

$$V(Z|X, T = t) = \frac{\sigma^2(1 + \sigma^2)}{(1 + \sigma^2 H_0(t) e^{\beta^T X})^2}$$

iken yaşayanlar için varyans;

$$V(Z|X, T > t) = \frac{\sigma^2}{(1 + \sigma^2 H_0(t) e^{\beta^T X})^2}$$

olarak tanımlanmaktadır. Bu iki eşitlikten de görüldüğü gibi zayıflığın varyansı zamanla azalmaktadır (Gutierrez, 2002; Wienke, 2011).

2.2.2. Ters-Gauss Zayıflık Modeli

Ters-Gauss zayıflık modeli, gamma zayıflık modelinin bir alternatifi olarak önerilmiştir. Bu zayıflık modeli Hougaard (1974) tarafından geliştirilmiş Klein (1992), Keiding, Andersen ve Klein (1997), Price ve Manatunga (2001), Economou ve Caroni (2005), Kheiri (2007), Duchateau ve Janssen (2007) tarafından da incelenmiştir.

$\mu > 0$  ve  $\lambda > 0$  olmak üzere ters-Gauss dağılımının olasılık yoğunluk fonksiyonu;

$$f(z) = \frac{\sqrt{\lambda}}{\sqrt{2\pi z^3}} \exp\left(-\frac{\lambda}{2\mu^2 z} (z - \mu)^2\right)$$

olarak verilmektedir. Ters-Gauss dağılımın Laplace dönüşümü ise aşağıdaki biçimdedir:

$$L(u) = E(e^{-uz}) = \exp\left(-\frac{\lambda\sqrt{1 + \frac{2\mu^2 u}{\lambda}}}{\mu} + \frac{\lambda}{\mu}\right) = \exp\left(\frac{\lambda}{\mu} \left(1 - \sqrt{1 + \frac{2\mu^2 u}{\lambda}}\right)\right)$$

Laplace dönüşümlerinin birinci ve ikinci türevleri yardımı ile beklenen değer ve varyans hesaplanabilmektedir. Birinci ve ikinci türevler, sırasıyla,

$$L'(u) = -\frac{\mu}{\sqrt{1 + \frac{2\mu^2 u}{\lambda}}} \exp\left(\frac{\lambda}{\mu} \left(1 - \sqrt{1 + \frac{2\mu^2 u}{\lambda}}\right)\right),$$

$$L''(u) = \frac{\mu^3}{\lambda \left(1 + \frac{2\mu^2 u}{\lambda}\right)^{\frac{3}{2}}} \exp\left(\frac{\lambda}{\mu} \left(1 - \sqrt{1 + \frac{2\mu^2 u}{\lambda}}\right)\right) + \frac{\mu^2}{1 + \frac{2\mu^2 u}{\lambda}} \exp\left(\frac{\lambda}{\mu} \left(1 - \sqrt{1 + \frac{2\mu^2 u}{\lambda}}\right)\right)$$

olarak verilmektedir. Beklenen değer ve varyans ise  $u=0$  noktasındaki türevlerin hesaplanması ile aşağıdaki biçimde ifade edilmektedir:

$$E(Z) = -L'(0) = \mu,$$

$$V(Z) = L''(0) - (L'(0))^2 = \frac{\mu^3}{\lambda}$$

Eğer  $E(Z) = \mu = 1$  ve  $V(Z) = \sigma^2 = \frac{1}{\lambda}$  alınırsa

Laplace dönüşümü,  $L(u) = e^{\frac{1}{\sigma^2}(1 - \sqrt{1 + 2\sigma^2 u})}$  olmaktadır.

Bu dönüşüm sayesinde yaşam fonksiyonu ve tehlike fonksiyonu elde edilebilmektedir. Yaşam fonksiyonu ve tehlike fonksiyonu sırasıyla,

$$S(t) = e^{\frac{1}{\sigma^2}(1 - \sqrt{1 + 2\sigma^2 H_0(t)})}$$

$$h(t) = \frac{h_0(t)}{(1 + 2\sigma^2 H_0(t))^{\frac{1}{2}}}$$

biçimindedir. t anından itibaren yaşayan birimlere ait zayıflığın olasılık yoğunluk fonksiyonu,

$$f(z|X, T > t) = \frac{S(t|X, z)f(z)}{s(t|X)} = \frac{1}{\sqrt{2\pi^2 z^3}} \exp\left(-\frac{(z - (1 + 2\sigma^2 H_0(t) e^{\beta^T X})^{-\frac{1}{2}})^2}{2\sigma^2 z} \right) \frac{1}{1 + 2\sigma^2 H_0(t) e^{\beta^T X}}$$

olarak tanımlanmaktadır. Bu birimler için beklenen değer

$$E(Z|X, T > t) = \frac{1}{\sqrt{1 + \sigma^2 H_0(t) e^{\beta^T X}}} \text{ ve varyans ise}$$

$$V(Z|X, T > t) = \frac{\sigma^2}{(1 + \sigma^2 H_0(t) e^{\beta^T X})} \text{ biçimindedir}$$

(Wienke, 2011).

### 2.3. Paylaşılmış Zayıflık Modeli

Paylaşılmış zayıflık modeli ile ilgili ilk çalışmalar Clayton (1978) ve Clayton and Cuzick (1985) tarafından yapılmıştır. Hougaard (1986) Weibull bireysel tehlike fonksiyonu ile paylaşılmış zayıflık modelini, Whitmore ve Lee (1991) üstel bireysel tehlike fonksiyonu ile ters-Gauss paylaşılmış zayıflık modelini ve Sahu (1997) ise Gibbs örneklemesini kullanarak Bayesci paylaşılmış zayıflık modelini incelemişlerdir. Xue ve Brookmeyer (1996) paylaşılmış zayıflık modelinin kısıtları üzerine çalışma yapmıştır. Ibrahim ve diğ. (2001) parametrik modeller gibi yarı parametrik paylaşılmış zayıflık modellerine de Bayesci yaklaşım uygulamıştır. Klein ve Moeschberger (2003) yarı parametrik paylaşılmış zayıflık modeline EM algoritmasını uygulamışlardır.

Modelin paylaşılmış zayıflık modeli olarak adlandırılmasının nedeni gruptaki birimlerin aynı zayıflığı paylaşmasıdır (Clayton, 1978). Bu birimlerin başarısızlık süreleri koşullu bağımsızdır. Bu koşul, zayıflık (Z) üzerinden şekillenmektedir. Her birimin sahip olduğu temel tehlike fonksiyonu  $Zh_0(t)$  biçimindedir. Birimlerin başarısızlık süreleri arasındaki bağımlılık durumu, Z değerinin birimler için ortak olmasından kaynaklanmaktadır. Bu yaklaşımda zayıflık terimi olayların oluş zamanları arasındaki ilişkiyi modellemede kullanılmaktadır. Bu yaklaşıma göre, paylaşılmış zayıflık modeli, benzer kümelerde bulunan birimler için ortak bir zayıflık terimine sahiptir. Bu zayıflık (Z) rasgele dağılmaktadır. Yani, bu model koşullu bağımsız bir modeldir ve zayıflık grup içindeki tüm birimler için ortaktır. Dolayısıyla olayların oluş zamanları arasında bir bağımlılık yaratır.

Bu yaklaşımdaki tehlike modeli aynı paylaşılmamış zayıflık modelindeki gibidir. En önemli fark paylaşılmış zayıflık modelinde, zayıflık birimlerin grup içerisinde paylaştığı ilişkili riski ifade etmektedir. Bu yüzden zayıflık birimler yerine birimlerin oluşturduğu grubu ifade etmektedir.

Paylaşılmış zayıflık yaklaşımına göre, gruptaki tüm başarısızlık sürelerini verilen zayıflıktan koşullu bağımsızdır. Zayıflık teriminin değeri zaman boyunca sürekli ve kümedeki birimler için ortaktır. Bu durum kümedeki başarısızlık süreleri üzerindeki bağımlılığın nedenidir. Bu bağımlılık paylaşılmış zayıflık modellerinde hep pozitif değerler almaktadır.

Yaşam çözümlemesinde paylaşılmış zayıflık modeli, n adet küme ve i ( $1 \leq i \leq n$ ) olmak üzere kümesinin  $n_i$  adet gözleme sahip olduğu varsayımı altında; gözlemlerin  $Z_i$  ( $1 \leq i \leq n$ ) zayıflıklarına sahip olduğu ifade edilmektedir. ( $1 \leq i \leq n, 1 \leq j \leq n_i$ ) olmak üzere  $X_{ij}$  ise i'inci kümedeki j'inci gözlemin başarısızlık süresi olan  $T_{ij}$  hakkında bilgi içeren açıklayıcı değişkenler vektörünü ifade etmektedir. Her bir i ( $1 \leq i \leq n$ ) kümesindeki yaşam süreleri,  $Z_i$  üzerinden koşullu olarak tehlike fonksiyonlarından bağımsız olarak varsayılmaktadır.

Paylaşılmış zayıflık modeli,

$$h(t / X_{ij}, Z_i) = Z_i h_0(t) \exp(\beta^T X_{ij}) \quad (3)$$

biçiminde verilmektedir. Burada  $h_0(t)$  temel tehlike fonksiyonunu ve  $\beta$  tahmin edilebilen parametreler vektörünü göstermektedir. Zayıflıklar ( $Z_i$ ), birbirinden bağımsız ve aynı dağılımlı ve  $\theta$  zayıflık dağılımının parametresi olmak üzere aynı  $f(z, \theta)$  olasılık yoğunluk fonksiyonuna sahiptir.

Çok değişkenli yaşam fonksiyonu zayıflık ( $Z_i$ ) üzerinden koşullu olmak üzere, i. gruptaki birimler için,

$$\begin{aligned} S(t_{i1}, \dots, t_{in_i} / X_i, Z_i) \\ = S(t_{i1} / X_{i1}, Z_i) \dots S(t_{in_i} / X_{in_i}, Z_i) \\ = \exp\left(-Z_i \sum_{j=1}^{n_i} H_0(t_{ij}) \exp(\beta^T X_{ij})\right) \end{aligned} \quad (4)$$

biçimindedir.  $H_0(t) = \int_0^t h_0(s) ds$  ve

$X_i = (X_{i1}, \dots, X_{in_i})$  açıklayıcı değişkenler matrisi olmak üzere Eşitlik 4'ün  $Z_i$  üzerinden beklenen değeri marjinal yaşam fonksiyonunu vermektedir:

$$\begin{aligned} S(t_{i1}, \dots, t_{in_i} / X_i, Z_i) = \\ = E \exp\left(-Z_i \sum_{j=1}^{n_i} H_0(t_{ij}) \exp(\beta^T X_{ij})\right) \\ = L\left(\sum_{j=1}^{n_i} H_0(t_{ij}) \exp(\beta^T X_{ij})\right). \end{aligned}$$

Burada L zayıflık değişkeninin Laplace dönüşümünü ifade etmektedir. Böylece, çok değişkenli yaşam fonksiyonu temel tehlike fonksiyonunun dağılım fonksiyonunda hesaplanan zayıflık dağılımının Laplace dönüşümü ile açıklanabilir.

Birleşik yaşam fonksiyonu ise yaşam fonksiyonlarından Eşitlik 5 ile elde edilebilir. Buradaki

gruplar arasında bağımsızlık olduğu varsayımı bahsi geçen çıkarımı elde etmede kritik noktadır.

$$S(t_{11}, \dots, t_{m_n} / X_1, \dots, X_n) = \prod_{i=1}^n L \left( \sum_{j=1}^{n_i} H_0(t_{ij}) \exp(\beta^T X_{ij}) \right) \quad (5)$$

Tek değişkenli koşulsuz yaşam fonksiyonları Laplace dönüşümünün bir sonucu olarak elde edilebilmektedir:

$$\begin{aligned} S(t_{ij} / X_{ij}) &= ES(t_{ij} / X_{ij}, Z_i) \\ &= E \exp(-Z_i H_0(t_{ij}) \exp(\beta^T X_{ij})) \\ &= L(H_0(t_{ij}) \exp(\beta^T X_{ij})). \end{aligned}$$

$L^{-1}$  Laplace dönüşümünün ( $L$ ) tersi olmak üzere,

$$H_0(t_{ij}) \exp(\beta^T X_{ij}) = L^{-1}(S(t_{ij} / X_{ij}))$$

ve  $i$ . grup için koşulsuz yaşam fonksiyonu,

$$\begin{aligned} S(t_{i1}, \dots, t_{in_i} / X_i) \\ = L(L^{-1}(S(t_{i1}/X_{i1})) + \dots + L^{-1}(S(t_{in_i}/X_{in_i}))) \end{aligned}$$

biçimindedir.  $n$  grup sayısını göstermek üzere her bir gruptaki birim sayısı  $n_i$  ( $j=1, \dots, n_i, i=1, \dots, n$ ) olmak üzere  $i$ . gruptaki  $j$ . birim için başlangıç zamanı, bitiş zamanı ve başarısız ya da durdurulmuş olması ( $t_{0ij}, t_{ij}, d_{ij}$ ) ile gösterilsin. Bu durumda paylaşılmış zayıflık modelinin olabilirlik fonksiyonu,

$$\begin{aligned} L_{ij}(\alpha_i) &= \frac{S_{ij}(t_{ij} / \alpha_i)}{S_{ij}(t_{0ij} / \alpha_i)} [h_{ij}(t_{ij} / \alpha_i)]^{d_{ij}} \\ &= \left[ \frac{S_{ij}(t_{ij})}{S_{ij}(t_{0ij})} \right]^{\alpha_i} [\alpha_i h_{ij}(t_{ij})]^{d_{ij}} \end{aligned}$$

biçimindedir.  $D_i = \sum_{j=1}^{n_i} d_{ij}$  olursa,  $i$ . grup için olabilirlik fonksiyonu

$$L_i(\alpha_i) = \alpha_i^{D_i} \prod_{j=1}^{n_i} \left[ \frac{S_{ij}(t_{ij})}{S_{ij}(t_{0ij})} \right]^{\alpha_i} [h_{ij}(t_{ij})]^{d_{ij}}$$

ifadesinde  $\alpha_i$  ye göre integral alınarak,

$$L_i = \int_0^{\infty} L_i(\alpha_i) g(\alpha_i) d\alpha_i \text{ biçiminde hesaplanır.}$$

Zayıflık terimi Gamma dağılımına sahip ise aşağıdaki gibi hesaplanır:

$$L_i = \left[ \prod_{j=1}^{n_i} [h_{ij}(t_{ij})]^{d_{ij}} \right] \frac{\Gamma(1/\theta + D_i)}{\Gamma(1/\theta)} \theta^{\rho} \left[ 1 - \theta \sum_{j=1}^{n_i} \ln \frac{S_{ij}(t_{ij})}{S_{ij}(t_{0ij})} \right].$$

Paylaşılmış gamma zayıflık modeli (gözlenen açıklayıcı değişkenler modele dahil edilirse),

$$\hat{Z}_i = \frac{1/\sigma^2 + \sum_{j=1}^{n_i} \delta_{ij}}{1/\sigma^2 + \sum_{j=1}^{n_i} e^{\beta X_{ij}} H(t_{ij})}, \quad i=1, \dots, n$$

biçiminde yazılır (Gutierrez, 2002; Wienke, 2011).

### 3. UYGULAMA

#### 3.1. Veri

Bu çalışmada, Ata (2005)'de orantısız tehlikeler için yaşam modellerinin incelendiği çalışmanın verileri ele alınmıştır. İbn-i Sina Hastanesi Göğüs Cerrahisi Bölümü'nde tedavi gören 236 akciğer kanseri hastasına ait veriler kullanılmıştır. Hastalar, ameliyat olduktan sonra hastalıklarının ilk nüks etmesine kadar geçen süre (min=1 ay, max=93 ay) boyunca izlenmiştir. Akciğer kanseri hastalarının yaşam sürelerini etkileyen faktörler yaşam çözümlemesi yöntemleri kullanılarak belirlenmeye çalışılmıştır.

Çalışmada, hastaların ameliyat olduğu tarihten hastalığın ilk nüksetmesine kadar geçen süre (ay olarak) yaşam süresi olarak alınmıştır. Hastalığın nüksetmesi başarısızlık olarak ifade edilmiştir. Hastalığı nüksetmeyen hastalar durdurulmuş olarak tanımlanmıştır. Hastaların izlenme süresi sona erdiğinde 236 hastadan 94'ünde (%39.8) başarısızlık ve 142'sinde (%60.2) durdurma gözlenmiştir. Uygulamada yaş (YS), sigara tüketimi (paket yıl olarak, ST), genişletilmiş rezeksiyon (extended resection, ER), tümörün boyutu (mm olarak, BY), invazyon (İV) ve patolojik evre (PE) değişkenleri gruplandırılarak çözümlemeye alınmıştır. Bu değişkenler ve değişkenlerin düzeyleri Tablo 1.'de verilmiştir.

**Tablo 1.** Kullanılan Değişkenler ve Düzeyleri

Değişken	Değişken Düzeyleri	n	%	Başarı sız	D
<b>Yaş</b>	1. <=39	13	5.5	5	8
	2. 40-49	38	16.1	12	26
	3. 50-59	76	32.2	33	43
	4. 60-69	81	34.3	32	49
	5. >=70	28	11.9	12	16
<b>Sigara Tüketimi</b>	1. <=5	16	6.8	4	12
	2. 6-30	62	26.3	23	39
	3. 31-60	123	52.1	48	75
	4. >=61	35	14.8	19	16
<b>Genişletilmiş Rezeksiyon</b>	0. Yok	190	80.5	71	119
	1. Var	46	19.5	23	23
<b>Boyut</b>	1. <=30	73	30.9	25	48
	2. 31-40	46	19.5	12	34
	3. 41-50	41	17.4	18	23
	4. >=50	76	32.2	39	37
<b>İnvazyon</b>	0. Yok	136	57.6	50	86
	1. Var	100	42.4	44	56
<b>Patolojik Evre</b>	1. Evre I	102	43.2	28	74
	2. Evre II	61	25.8	24	37
	3. Evre III	60	25.4	30	30
	4. Evre IV	13	5.5	12	1

D:Durdurulmuş

**Tablo 2.** Log-lojistik HBS Modeli Çözümlemesinin Sonuçları

Değişken ve düzeyleri		$\hat{\beta}$	S.H.	Alt sınır – Üst sınır	p değeri
<b>Kesişimterimi</b>		5.3139	0.5773	4.1825 – 6.4454	< 0.0001
<b>Yaş</b>	1. <=39				
	2. 40-49	0.1174	0.4483	-0.7613 – 0.9962	<b>0.7934</b>
	3. 50-59	-0.2057	0.4139	-1.0168 – 0.6054	<b>0.6191</b>
	4. 60-69	-0.0162	0.4198	-0.8390 – 0.8066	<b>0.9692</b>
	5. >=70	-0.3570	0.4870	-1.3115 – 0.5975	<b>0.4636</b>
<b>Sigara Tüketimi</b>	1. <=5				
	2. 6-30	-0.5232	0.4564	-1.4177 – 0.3714	<b>0.2517</b>
	3. 31-60	-0.415	0.4330	-1.2501 – 0.4471	<b>0.3538</b>
	4. >=61	-1.0662	0.4642	-1.9760 – -0.1563	<b>0.0216</b>
<b>Genişletilmiş Rezeksiyon</b>	1. Yok				
	2. Var	-0.2812	0.2552	-0.7815 – 0.2190	<b>0.2705</b>
<b>Boyut</b>	1. <=30				
	2. 31-40	0.3124	0.3041	-0.2837 – 0.9084	<b>0.3043</b>
	3. 41-50	-0.4470	0.3060	-1.0469 – 0.1528	<b>0.1441</b>
	4. >=50	-0.7138	0.2627	-1.2286 – 0.1989	<b>0.0066</b>
<b>İnvazyon</b>	1. Yok				
	2. Var	-0.1851	0.2549	-0.6846 – 0.3144	<b>0.4676</b>
<b>Patolojik Evre</b>	1. Evre I				
	2. Evre II	-0.1737	0.2893	-0.7408 – 0.3933	<b>0.5482</b>
	3. Evre III	-0.6372	0.2634	-1.1534 – 0.1211	<b>0.0155</b>
	4. Evre IV	-1.5326	0.3570	-2.2322 – 0.8330	< 0.0001
<b>Ölçek (<math>\sigma</math>)</b>		<b>0.6669</b>	<b>0.0586</b>	<b>0.5614 – 0.7922</b>	

Yaşam çözümlemesinde kullanılan regresyon tipi modellerde değişken düzeylerinden biri referans kategorisi olarak alınmakta ve değişken düzeylerinin yorumlanması buna göre yapılmaktadır. Başarısızlık

Bu çalışmada ise hızlandırılmış başarısızlık zamanı modelleri ve zayıflık modelleri aynı veri kümesi için incelenmiş ve Ata (2005)'in çalışmasında bulunan sonuçlar ile birlikte değerlendirilmiştir.

### 3.2. Hızlandırılmış Başarısızlık Süresi Modeli Sonuçları

Hızlandırılmış başarısızlık süresi (HBS) modeli gibi parametrik yöntemler yaşam süresi bilinen bir dağılıma uygunluk gösteriyorsa, parametrik olmayan ya da yarı parametrik yöntemlere göre daha iyi sonuçlar vermektedir. Akciğer kanseri veri kümesi için HBS modelleri kapsamında üstel, Weibull, log-lojistik, log-normal, Gompertz, Gamma HBS modelleri ile çözümleme yapılmıştır. Modellerin anlamlılığını test etmek için olabilirlik oranı (LR) test istatistiği kullanılmış ve tüm modellerin istatistiksel olarak anlamlı olduğu görülmüştür ( $p < 0.05$ ). İncelenen modeller için AIC değerleri sırasıyla Üstel (AIC=478.250), Weibull (AIC=475.778), Log-lojistik (AIC=467.979), Log-normal (AIC=471.586), Gompertz (AIC=480.214), Gamma (AIC=468.593) olarak elde edilmiştir. Log-lojistik HBS modeli akciğer kanseri verisi için kullanıldığında elde edilen sonuçlar Tablo 2'de verilmiştir.

süresini etkileyen faktörleri belirlemek için Tablo 2 incelendiğinde, sigara tüketimi değişkeninin 4. düzeyinin, boyut değişkeninin 4. düzeyinin ve patolojik evre değişkeninin 3. ile 4. düzeylerinin %95 güven düzeyinde önemli olduğu görülmüştür ( $p < 0.05$ ).



Parametrik regresyon modellerini yorumlarken hızlandırma faktörünün yaşam süresinin azalmasını hızlandırıp hızlandırmadığı dikkate alınmaktadır. Buna göre Tablo 2.'deki bilgilerden yararlanarak sigara tüketimi 61'den fazla olan hastalara ait hızlandırma faktörü yaklaşık 0.35 olup, sigara tüketimi 61'den fazla olan hastaların yaşam süresi sigara tüketimi 5'ten az olan hastaların yaşam süresinden 2.9 kat daha kısadır biçiminde yorumlanabilir. Benzer biçimde tümör boyutu 50 mm'den büyük olan hastaların ortalama yaşam süresi tümör boyutu 30 mm'den küçük olan hastalara göre yaklaşık 2 kat, patolojik evresi Evre-IV olan hastaların ortalama yaşam süresi patolojik evresi Evre-I olan hastalara göre yaklaşık 4.6 kat daha kısadır.

**Tablo 3.** Gamma Zayıflık Terimi İçeren Log-Lojistik HBS Modeli Çözümlemesi

Değişken ve düzeyleri		$\hat{\beta}$	Hızlandırma Faktörü	p değeri
<b>Kesişim terimi</b>		5.0954	-	< 0.0001
<b>Yaş</b>	1. <=39	-	-	-
	2. 40-49	-0.1674	0,846	0.159
	3. 50-59	-0.2928	0,746	0.479
	4. 60-69	-0.7039	0,495	0.783
	5. >=70	-0.1135	0,893	0.718
<b>Sigara Tüketimi</b>	1. <=5	-	-	-
	2. 6-30	-0.4679	0,626	0.301
	3. 31-60	-0.3169	0,728	0.453
	4. 5>=61	-0.9642	0,381	0.035
<b>Genişletilmiş Rezeksiyon</b>	1. Yok	-	-	0,052
	2. Var	-0.06006	0,548	
<b>Boyut</b>	1. <=30	-	-	-
	2. 31-40	0.3354	1,398	0.264
	3. 41-50	-0.4412	0,643	0.156
	4. >=50	-0.7572	0,469	0.004
<b>İnvazyon</b>	1. Yok	-	-	-
	2. Var	-0.02770	0,973	0.921
<b>Patolojik Evre</b>	1. Evre I	-	-	-
	2. Evre II	-0.1950	0,823	0.515
	3. Evre III	-0.6418	0,526	0.016
	4. Evre IV	-1.3255	0,266	0.001
Ölçek ( $\sigma$ )		0.5322		
Zayıflık terimi ( $\theta$ )		0.9616		0.061
-2LogL		435.1918		

**Tablo 4.** Ters-Gauss Zayıflık Terimi İçeren Log-Lojistik HBS Modeli Çözümlemesi

Değişken ve düzeyleri		$\hat{\beta}$	Hızlandırma Faktörü	p değeri
<b>Kesişim terimi</b>		4.9835	-	< 0.0001
<b>Yaş</b>	1. <=39	-	-	-
	2. 40-49	-0.1849	0.831	0.690
	3. 50-59	-0.6946	0.499	0.471
	4. 60-69	-0.1173	0.890	0.779
	5. >=70	-0.3026	0.739	0.164
<b>Sigara Tüketimi</b>	1. <=5	-	-	-
	2. 6-30	-0.4651	0.628	0.303
	3. 31-60	-0.9474	0.388	0.459
	4. 5>=61	-0.3127	0.731	0.041
<b>Genişletilmiş Rezeksiyon</b>	1. Yok	-	-	-
	2. Var	-0.6297	0.533	0.053
<b>Boyut</b>	1. <=30	-	-	-
	2. 31-40	0.3308	1.392	0.272
	3. 41-50	-0.4462	0.640	0.149
	4. >=50	-0.753	0.471	0.004
<b>İnvazyon</b>	1. Yok	-	-	-
	2. Var	-0.0518	0.950	0.851

Değişken ve düzeyleri	$\hat{\beta}$	Hızlandırma Faktörü	p değeri
Patolojik Evre	1. Evre I	-	-
	2. Evre II	-0.1714	0.842
	3. Evre III	-0.6041	0.546
	4. Evre IV	-1.3511	0.259
Ölçek ( $\sigma$ )	0.4954		
Zayıflık terimi ( $\theta$ )	2.0351		0.055
-2LogL	435.0314		

Gamma zayıflık terimi ve Ters-Gauss zayıflık terimi ile kurulan modeller %10 anlamlılık düzeyinde değerlendirilirse bu modellere ilişkin zayıflık terimlerinin anlamlı olduğu görülmektedir. Bu da zayıflık terimi içeren modellerin kullanımının daha uygun olduğunu göstermektedir. Zayıflık terimlerinin yorumlanmasında zayıflık teriminin heterojenliğin bir ölçüsü olduğu bilgisi altında Pankratz, Andrade ve Thernau (2005) yaklaşımına göre zayıflık teriminin karekökünün üstel ifadesi  $\exp(\sqrt{\theta})$  rastgele etkinin yani zayıflığın etkisini göstermektedir. Log-lojistik dağılıma sahip Gamma zayıflık terimi ile kurulan model için  $\exp(\sqrt{\theta}) = \exp(\sqrt{.9616}) = 2.67$  olarak elde edilmiştir. Bu değer bir hastanın, çalışma sürecinde seçilen bir noktada, tüm çalışma grubuna ait riskten 2.67 kata kadar daha fazla ya da az risk taşıyabileceğini ifade etmektedir. Benzer biçimde log-lojistik dağılıma sahip ters-Gauss zayıflık terimi ile kurulan model için  $\exp(\sqrt{\theta}) = \exp(\sqrt{2.0351}) = 4.16$  olarak elde edilmiştir. Bu değer bir hastanın, çalışma sürecinde seçilen bir noktada, tüm çalışma grubuna ait riskten 4.16 kata kadar daha fazla ya

da az risk taşıyabileceğini ifade etmektedir. Bu iki model karşılaştırılken zayıflık teriminin anlamlılığı ve -2LogL beraber dikkate alınırsa, log-lojistik dağılıma sahip ters-Gauss zayıflık terimi ile kurulan modelin daha iyi olduğu söylenebilmektedir.

#### 3.4. Paylaşılmış Zayıflık Modeli Sonuçları

Açıklayıcı değişkenlerin aynı düzeylere sahip bireyler bir grup olarak ele alınmış ve aynı zayıflık terimini paylaşacakları düşünülerek paylaşılmış zayıflık modelleri incelenmiştir. Bu modellerin, yaş, sigara tüketimi, genişletilmiş rezeksiyon, invazyon değişkenleri için anlamlı olmadığı görülmüştür. Boyut değişkeni ve patolojik evre için ise paylaşılmış zayıflık modelleri anlamlı bulunmuştur ( $p < 0.05$ ). Bu değişkenlere ilişkin bilgiler Tablo 5 ve Tablo 6'da verilmiştir.

**Tablo 5.** Boyut Değişkeni için Zayıflık Terimi İçeren Log-Lojistik Hbs Modeli Çözümlemesi

Açıklayıcı Değişken		Boyut					
		Gamma			Ters Gauss		
		Parametre tahmini	Hızlandırma Faktörü	p-değeri	Parametre tahmini	Hızlandırma Faktörü	p-değeri
Yaş	1. <=39						
	2. 40-49	0.1294	1.138	0.777	0.1274	1.136	0.781
	3. 50-59	-0.1575	0.854	0.708	-0.1569	0.855	0.709
	4. 60-69	0.0030	1.003	0.995	0.0037	1.003	0.993
	5. >=70	-0.2071	0.813	0.675	-0.2080	0.812	0.674
Sigara Tüketimi	1. <=5						
	2. 6-30	-0.6108	0.543	0.012	-0.6062	0.545	0.208
	3. 31-60	-0.4620	0.630	0.314	-0.4559	0.634	0.321
	4. >=61	-1.2160	0.300	0.204	-1.2126	0.297	0.012
Genişletilmiş Rezeksiyon	1. Yok						
	2. Var	-0.2888	0.750	0.285	-0.2920	0.747	0.283
İnvazyon	1. Yok						
	2. Var	-0.3060	0.736	0.229	-0.3018	0.739	0.236
Patolojik Evre	1. Evre I						
	2. Evre II	-0.1822	0.833	0.536	-0.1811	0.834	0.538
	3. Evre III	-0.6045	0.546	0.023	-0.6034	0.547	0.023
	4. Evre IV	-1.5297	0.217	0.000	-1.5330	0.216	0.000
Theta ( $\theta$ )		0.1152		0.027	0.1307		0.027
-2LogL		447.891			447.876		

**Tablo 6.** Patolojik Evre Değişkeni için Zayıflık Terimi İçeren Log-Lojistik Hbs Modeli Çözümlemesi

Açıklayıcı Değişken		Patolojik Evre					
		Gamma			Ters Gauss		
		Parametre tahmini	Hızlandırma Faktörü	p-değeri	Parametre tahmini	Hızlandırma Faktörü	p-değeri
<b>Yaş</b>	1. <=39						
	2. 40-49	-0.0416	0.959	0.930	-0.2151	0.806	0.664
	3. 50-59	-0.3232	0.724	0.444	-0.4052	0.667	0.367
	4. 60-69	-0.0886	0.915	0.835	-0.1272	0.881	0.768
	5. >=70	-0.4773	0.620	0.327	-0.5904	0.554	0.236
<b>Sigara Tüketimi</b>	1. <=5						
	2. 6-30	-0.4514	0.637	0.323	-0.3736	0.688	0.404
	3. 31-60	-0.4083	0.665	0.340	-0.3636	0.695	0.380
	4. 5>=61	-1.027	0.358	0.029	-0.9470	0.388	0.043
<b>Genişletilmiş Rezeksiyon</b>	1. Yok						
	2. Var	-0.3607	0.697	0.264	-0.6004	0.549	0.191
<b>Boyut</b>	1. <=30						
	2. 31-40	0.4217	1.525	0.165	0.4812	1.618	0.118
	3. 41-50	-0.4182	0.658	0.181	-0.3998	0.670	0.186
	4. >=50	-0.6330	0.531	0.021	-0.6380	0.528	0.021
<b>İnvazyon</b>	1. Yok						
	2. Var	-0.0865	0.917	0.747	-0.0044	0.996	0.987
<b>Theta (Θ)</b>		0.514		<0.0001	2.104		<0.0001
<b>-2LogL</b>		446.972			445.441		

Tablo 5’de boyut değişkeni bir küme olarak değerlendirilmiş olup kitle ikiye ayrılmıştır, boyut değişkeni için referans grubu olarak boyut 4 ( $\geq 50$ ) belirlenmiştir. Kümedeki tüm bireylerin aynı zayıflığı paylaştığı varsayımı altında paylaşılmış zayıflık modeli kurulmuş ve model 0.05 anlamlılık düzeyinde önemli bulunmuştur. Gamma ve ters Gauss zayıflık terimlerini içeren her iki model için de zayıflık teriminin etkisi yaklaşık olarak  $\exp(\sqrt{\theta}) = 1.4$  elde edilmiştir. Buna göre tümörün boyutu  $\geq 50$  olan bireylerin çalışma grubundaki diğer bireylere göre 1.4 kata kadar daha fazla ya da daha az risk taşıdığı söylenebilmektedir.

Benzer biçimde; Tablo 6’da patolojik evre değişkeni bir küme olarak değerlendirilmiş olup kitle, patolojik evre değişkeni için referans grubu Evre-IV olmak üzere ikiye ayrılmıştır. Paylaşılmış zayıflık modeli, zayıflık terimi Gamma ve ters Gauss olmak üzere kurulmuş ve model istatistiksel açıdan anlamlı bulunmuştur. Zayıflık terimi Gamma olan paylaşılmış zayıflık modeli için yaklaşık olarak  $\exp(\sqrt{\theta}) = 2$  elde edilmiştir. Bu değer patolojik evresi Evre-IV olan bireylerin çalışma grubundaki bireylere göre 2 kata kadar daha fazla ya da daha az risk taşıdığını ifade etmektedir. Bir diğer model olan, ters Gauss zayıflık terimli paylaşılmış zayıflık modelinde ise  $\exp(\sqrt{\theta}) = 4.25$  olarak elde edilmiştir. Bu model dikkate alınarak, patolojik evresi Evre-IV olan bireylerin çalışma grubundaki diğer bireylere göre 4.25 kata kadar daha fazla ya da daha az risk taşıdığını ifade etmektedir.

#### 4. SONUÇ

Yaşam çözümlemesinde gözlenemeyen değişkenlerden kaynaklanan heterojenliğin açıklaması zayıflık modelleri ile yapılabilmektedir. Yaşam

çözümlemesi kullanılarak elde edilen bulguların tutarlılığı için heterojenliğin bireylere ya da gruplara etkisinin incelenmesi gerekmektedir. Heterojenliğin yüksek olduğu durumlarda çalışma grubuna dair çözümleme yapılırken zayıflık modelleri göz önünde bulundurulmalıdır.

Akciğer kanseri tedavisi gören 236 bireye ait veri kümesine yaşam çözümlemesi yöntemleri uygulanmıştır. Zayıflık terimi içeren modeller ayrıntılı olarak incelenmiştir. Yaşam süresinin log-lojistik dağılıma uyum sağladığı gözlenmiş olup çözümlenmeler bu dağılım üzerinden çeşitlendirilmiştir. Paylaşılmamış zayıflık modeli ve log-lojistik HBS modeli karşılaştırıldığında, paylaşılmamış zayıflık modelinin veriyi yorumlamada daha iyi olduğu görülmüştür. Paylaşılmış zayıflık modelleri ise verinin açıklayıcı değişkenlere göre gruplandırılması durumunda grupların birbirinden ne oranda az ya da fazla risk taşıdığını görebilmek ve yorumlayabilmek için kullanılmıştır.

Ülkemizde, özellikle kanser verileri yaşam çözümlemesi yaklaşımı ile incelenirken verinin homojen olduğu varsayımı yapılarak orantılı tehlikeler modelinden yararlanılmaktadır. Ancak, gerçekte incelenen verinin homojen olmadığı durumlarda karşılaşılmaktadır. Yaşam verisinin homojen olmaması, çözümleme sonucunda ulaşılan değerlendirmelerin etkinliğinin azalmasına ve verinin yeterince açıklanamamasına neden olmaktadır. Bu nedenle, yaşam çözümlemesi verilerinde heterojenliğin açıklanabilmesi için zayıflık modellerinin göz önüne alınması gerekmektedir.

## References

- Ata, N. (2005). Yaşam Çözümlemesinde Orantısız Hazard Modeli, Yüksek Lisans Tezi, Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, Ankara, 2005.
- Babiker A., Cuzick J. (1994). A simple frailty model for family studies with covariates. *Statistics in Medicine*, 13, 1679-1692.
- Clayton, D. (1978). A model for association in bivariate life tables and its applications in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, 65, 141–151.
- Clayton, D., Cuzick, J. (1995). Multivariate generalisations of the proportional hazards model (with discussion). *Journal of the Royal Statistical Society, Series A*, 148, 82–117.
- Congdon, P. (1995). Modelling Frailty In Area Mortality. *Statistics in Medicine*, 14, 1859-1874.
- Cox, D.R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society, Series B*, 34, 187-220.
- Duchateau, L., Janssen, P. (2007). *The Frailty Model*, Springer, New York.
- Economou, P., Caroni, C. (2005). Graphical tests for the assumption of gamma and inverse Gaussian frailty distributions. *Lifetime Data Analysis* 11, 565–582.
- Guo, G., Rodriguez, G. (1992). Estimating a multivariate proportional hazards model for clustered data using the EM algorithm with an application to child survival in Guatemala. *Journal of the American Statistical Association*, 87, 969-976.
- Gutierrez, R.G. (2002). Parametric frailty and shared frailty survival models. *The Stata Journal*, 2, 22–44.
- Hougaard, P. (1984). Life table methods for heterogeneous populations. *Biometrika*, 71, 75–83.
- Hougaard, P. (1986). Survival models for heterogeneous populations derived from stable distributions. *Biometrika* 73, 387–396.
- Hougaard, P. (1995). Frailty models for survival data. *Lifetime Data Analysis*, 1, 255–273.
- Hougaard, P. (2000). *Analysis of Multivariate Survival Data*, Springer, New York.
- Ibrahim, J. G., Chen, M., Sinha, D. (2001), *Bayesian Survival Analysis*, Springer, New York.
- Keiding, N., Andersen, P., Klein, J. (1997). The role of frailty models and accelerated failure time models in describing heterogeneity due to omitted covariates. *Statistics in Medicine* 16, 215–224.
- Kheiri, S., Kimber, A., Meshkani, M.R. (2007). Bayesian analysis of an inverse Gaussian correlated frailty model, *Computational Statistics and Data Analysis* 51, 5317–5326.
- Klein, J.P. (1992). Semiparametric estimation of random effects using the Cox model based on the EM algorithm. *Biometrics* 48, 795–806.
- Klein, J. P. and Moeschberger, M. L. (2003). *Survival Analysis: Techniques for Censored and Truncated Data*, 2nd edition, Springer, New York.
- Lanchester, T. (1979). Econometric methods for the duration of unemployment. *Econometrica*, 47, 939–956.
- McGilchrist, C.A., Aisbett, C.W. (1991). Regression with frailty in survival analysis. *Biometrics* 47, 461–466.
- Oakes, D. (1982). A concordance test for independence in the presence of censoring. *Biometrics* 38, 451–455.
- O’Quigley J., Stare, J. (2002). Proportional hazards models with frailties and random effects. *Statistics in Medicine*, 21, 3219-3233.
- O’Quigley J., Stare, J. (2004). Fit and frailties in proportional hazards regression. *Statistics in Medicine*, 21, 3219-3233.
- Pankratz, V. S., de Andrade, M. and Thernau, T. M. (2005). Random effects Cox proportional hazard model: general variance components methods for time-to-event data. *Genetic Epidemiology*, 28, 97-109.
- Price, D.L., Manatunga, A.K. (2001). Modelling survival data with a cured fraction using frailty models. *Statistics in Medicine* 20, 1515–1527.
- Sahu, S.K., Dey, D.K., Aslanidou, H., Sinha, D. (1997). A weibull regression model with gamma frailties for multivariate survival data. *Lifetime Data Analysis*, 3, 123–137.
- Vaupel, J.W., Manton, K., Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16, 439–454, 1979.
- Whitmore, G.A., Lee, M.L.T. (1991). A multivariate survival distribution generated by an inverse gaussian mixture of exponentials, *Technometrics*, 33, 39–50.
- Wienke A. (2011). *Frailty Models in Survival Analysis*, Chapman&Hall, Florida.
- Wienke, A., Christensen, K., Holm, N., Yashin, A. (2000). Heritability of death from respiratory diseases: an analysis of Danish twin survival data using a correlated frailty model. In: *Medical Infobahn for Europe*. A. Hasman et al. (eds.), IOS Press, Amsterdam, 407–411.
- Wienke, A., Yashin, I.A., Zdravkovic, S., Pedersen, N.L., Marenberg, M.E., De Faire, U. (2004). Genetic influences on CHD-Death and the impact of known risk factors: Comparison of two frailty models. *Behavior Genetics*, 34, 585-592.
- Xue, X., Brookmeyer, R. (1996). Bivariate frailty model for the analysis of multivariate survival time. *Lifetime Data Analysis* 2, 277-289.
- Yashin, A.I., Iachine, I.A. (1995). Genetic analysis of durations: correlated frailty model applied to survival of danish twins. *Genetic Epidemiology*, 12, 529 – 538.
- Yashin, A.I., Vaupel, J.W., Iachine, I.A. (1995). Correlated individual frailty: An advantageous approach to survival analysis of bivariate data. *Mathematical Population Studies*, 5, 145 – 159.