T.C.
MARMARA ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
EKONOMETRİ ANABİLİM DALI
İSTATİSTİK BİLİM DALI

GENELLEŞTİRİLMİŞ DOĞRUSAL MODELLER İLE SİGORTA ŞİRKETLERİNDE HASAR REZERVİNİN TAHMİNİ

Yüksek Lisans Tezi

YUSUF ARSLAN

İstanbul, 2018

T.C.
MARMARA ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
EKONOMETRİ ANABİLİM DALI
İSTATİSTİK BİLİM DALI

GENELLEŞTİRİLMİŞ DOĞRUSAL MODELLER İLE SİGORTA ŞİRKETLERİNDE HASAR REZERVİNİN TAHMİNİ

Yüksek Lisans Tezi

YUSUF ARSLAN

Danışman: PROF. DR. DİLEK ALTAŞ

İstanbul, 2018

T.C.
MARMARA ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ MÜDÜRLÜĞÜ

TEZ ONAY BELGESİ

EKONOMETRİ Anabilim Dalı İSTATİSTİK Bilim Dalı TEZLİ YÜKSEK LİSANS öğrencisi Yusuf Arslan'nın GENELLEŞTİRİLMİŞ DOĞRUSAL MODELLER İLE SİGORTA ŞİRKETLERİNDE HASAR REZERVİNİN TAHMİNİ adlı tez çalışması, Enstitümüz Yönetim Kurulunun 10.05.2018 tarih ve 2018-13/19 sayılı kararıyla oluşturulan jüri tarafından oy birliği / oy çokluğu ile Yüksek Lisans Tezi olarak kabul edilmiştir.

Tez Savunma Tarihi 25/05/2018

Öğretim Üyesi Adı Soyadı			**İmzası**
1.	Tez Danışmanı	Prof. Dr. DİLEK ALTAŞ	
2.	Jüri Üyesi	Prof. Dr. ŞAHAMET BÜLBÜL	
3.	Jüri Üyesi	Doç. Dr. ADİL OĞUZHAN	

ÖZET

Hayat-dışı sigortalarda, hasar talepleri ile hasar dosyasının kapanması arasında zaman farkı mevcuttur. Ayrıca, yasal düzenlemeler, hasarın gecikmeli raporlanması ya da yapılan itirazlar, hasar talebinde bulunmanın zaman alması veya kapalı dosyaların yeniden açılmasına neden olabilmektedir.

Sigorta şirketlerinin yükümlülüklerini yerine getirebilmesi için ayrılacak olan hasar rezervlerini doğrun tespit etmesi, şirketin mali yapısının korunması açısından oldukça önemlidir. Rezerv hesabında kullanılan geçmiş veriler genellikle üçgen merdiven metodu ile gösterilmektedir. Bu veriler iki zaman ekseninde birbirinden ayrılır. Yatay eksen hasar gelişim yılını ve dikey eksen hasar oluşum yılını göstermektedir. Genelleştirilmiş Doğrusal Modeller hasar rezervlerinin tahmininde etkin bir yoldur.

Bu çalışmada, hayat-dışı sigorta şirketinin 2007-2016 yılları arasında trafik sigortası ödenen hasar verileri kullanılmıştır. Gelişim üçgenlerindeki verilerin dağılımını belirlemek amacıyla üstel dağılım ailesinden gelen olasılık dağılımları kullanılmıştır. Hasar rezervlerini hesaplanmasında tahmine dayalı tüm dağılımların kullanılması sağlanmıştır. Karlılık için, Pearson artıklarının yeniden örneklendiği bootstrap yöntemi benimsenmiştir. Tahminleyici dağılımlara sahip olmak için, pratikte en çok kullanılan risk ölçümleri VaR ve $CVaR$ hesaplamada dikkate alınmıştır. Önerilen modelleri analiz etmek ve performanslarını karşılaştırmak için gerçek hayattaki veriler üzerinden uygulama yapılmıştır. Uygulama sonuçları tablo ve şekiller ile özetlenmiştir.

Anahtar Kelimeler: Hasar rezervleri, genelleştirilmiş doğrusal modeller, bootstrap metodu, risk ölçümleri.

ABSTRACT

In non-life insurance, there is a time-lag between the occurrence of the claim and the closure of the claim file. In addition to this time-lag, a delay in claims application process or re-opening of the related file may arise due to certain legal regulations, delayed claim reporting or objections arisen during the process.

Determination of reserves accurately and sufficiently in order to meet the liabilities plays an important role in protecting the financial status for an insurance company. Historical data used to calculate reserves are generally presented through claims development triangles. These data diverge on two time axes. The horizontal axis indicates the years of development of a claim, when the vertical one indicates the year of the occurrence of the claim. The Generalized Linear Models may be accepted as one of the most efficient ways to estimate claims reserves.

In this study, a dataset of paid losses in the Motor Third Party Liability Insurance between 2007 and 2016 of the non-life insurance companies are analyzed. Certain relevant probability distributions from exponential distribution family are used to define the distributions of data in development triangles. It is possible to state that almost all predictive distribution methods are used to calculate these claim reserves. The boostrap method, in which the Pearson residuals are resampled, is allocated for profitability estimations. Besides, VaR and $CVaR$ methods, as two of the most commonly preferred in risk measurement are adopted with the purpose of having the predictive distributions. With the intent of analyzing models and comparing the performances, the real-life data serve as the basis of calculations conducted. The results obtained through all these analysis are presented graphically or in the form of tables.

Keywords: Claims reserves, generalized linear models, bootstrap method, risk measures.

TEŞEKKÜR

Tez konusunun seçiminde yol gösteren ve tüm çalışma süresince desteğini esirgemeyen danışmanım Sayın Prof. Dr. Dilek ALTAŞ'a,

Gösterdikleri hoşgörü ve destek için değerli bölüm hocalarıma,

Hayatımın her aşamasında verdikleri destek ve beni yetiştirmek için gösterdikleri çaba için aileme,

Destekleri için tüm yöneticilerim ve çalışma arkadaşlarıma,

Teşekkürlerimi sunuyorum.

İÇİNDEKİLER

ÖZET i
ABSTRACT ii
TEŞEKKÜR iii
TABLO LİSTESİ vi
ŞEKİL LİSTESİ vii
GİRİŞ 1
1 HASAR REZERV TEORİSİNE GİRİŞ 3
1.1 Hasar Süreçleri ve Hasar Rezervleri 4
1.2 Temel Hasar Rezervleri Yöntemi 6
1.2.1 Chain-Ladder Metodu 6
1.2.2 Bornhuetter-Ferguson Metodu 8
2 GENELLEŞTİRİLMİŞ DOĞRUSAL MODELLEME 13
2.1 Temel Varsayımlar ve Terminoloji 13
2.2 Bağ Fonksiyonu 15
2.3 Model Fitleme 17
2.3.1 GDM'de En Çok Olabilirlik Kanunu 17
2.3.2 Yarı (Quasi)-Olabilirlik Yöntemi 19
2.3.3 Uyum İyiliğinin Ölçülmesi 21
2.3.4 Bağımsız Değişkenlerin Test Edilmesi 22
3 HASAR REZERVLERİNDE GENELLEŞTİRİLMİŞ DOĞRUSAL MODELLEME 25
3.1 Hasar Adetlerinin Modellenmesi 25
3.1.1 Aşırı-yayılmış Poisson Modeli 25
3.1.2 Aşırı-yayılmış Negatif Binom Modeli 28
3.2 Hasar Tutarlarının Modellenmesi 31
3.2.1 Gamma Modeli 31
3.2.2 Ters Gauss Modeli 33
3.2.3 Normal Dağılım Modeli 36
Aşırı-yayılmış Negatif Binom Modelinin Yakınlaştırması 38
3.3 Aşırı-yayılmış Poisson Modeli 43
3.4 Aşırı-yayılmış Negatif Binom Modeli 44

3.5 Gamma Modeli 45
3.6 Ters Gauss Modeli 46
3.7 Normal Modeli 46
4 UYGULAMA 49
4.1 Modellerin Analizi 50
4.2 Aşırı-yayılmış Negatif Binom Normal Yaklaşımı 56
SONUÇ 60
KAYNAK KODU 63
KAYNAKÇA 76

TABLO LİSTESİ

1.1 Hayat Dışı Sigortada Hasarın Gelişimi 3

1.2 Hasar Gelişim Üçgeni 4

5.1 2007-2016 Yılları Arasında Trafik Sigortası Aşamalı Hasar Ödemeleri 49

5.2 Hasar Rezervlerinin Nihai Değerleri 50

5.3 Fit Edilen Modellerde Pearson X^2 istatistiği, sapma ve Aaike Ölçütü 54

5.4 Tahmini Rezerv ve risk ölçüleri VaR, CVaR 55

5.5 Aşırı-yayılmış Negatif Binom Normal Yaklaşımı Modeli Fit Edilen Değerler . 57

5.6 Aşırı-yayılmış Negatif Binom Normal Yaklaşımı Modelinin Öngörülen Aşamalı Hasarları 58

5.7 Aşırı-yayılmış Negatif Binom Normal Yaklaşımı Modelinin Rezerv ve Risk Ölçümleri 59

ŞEKİL LİSTESİ

5.1 Birikimli ve Aşamalı Hasar Tutarlarının Gelişimi 50

5.2 Fit Edilen Modellerin Ölçeklendirilmiş Çarpıklık Artıkları 52

5.3 Fit Edilen Modellerin Ölçeklendirilmiş Pearson Artıkları 53

GİRİŞ

Hasar rezervleri sigortacılığın en önemli konusudur. Geçmişte bu problemin çözümü için şirketler birçok yöntem denemiştir. Bu yöntemlerin en popüleri gelecek hasarların yalnızca beklenen değerini tahmin etmeye dayalı deterministik modeldir. Deterministik model tahmin belirsizliklerinin hesaba katılması gereken hasar oluşumları benzeri rastgele olaylar için uygun değildir.

Bu gibi durumları ele alan hasar gelişimi modellemesinin stokastik metodu uygulayıcılar arasında popüler olmaya başlamıştır. Bu modellerin kullanılması bize gelecekteki yükümlülüklerin daha yüksek olasılıkla hesaplanmasını ve dolayısıyla rezerve konu olan hasarlar ile ilgili risk ölçümlemesi yapılabilmesine olanak sağlamaktadır.

Hasar miktarının hesaplanması ve tahmini için stokastik bir model seçimi, bu tezin esas amacını oluşturmaktadır. Bu amaçla doğrusal regresyon uzantısı olan Genelleştirilmiş Doğrusal Modelleme üzerinde çalışılmıştır. Bu modeller normal dağılıma uymayan verilerin doğrusal olmayan ilişki ile bağımlı ve bağımsız değişkenlerin modellenmesi için kullanılmaktadır.

Hasar rezervlerinin hesaplanması ile ilgili olarak, tahminin dağılımı, hasar rezervlerinin rastgele değişkenler toplamı olması sebebiyle karmaşık olabilir. Bu tezde bu problem bootstrap metodu uygulaması ile aşılmaktadır. Ayrıca hayat-dışı sigortalardaki sınırlı sayıda gözlem sorununu ortadan kaldırmaktadır.

Birinci bölümde hasar rezervi problemi ve standart gösterimi tanıtılmıştır. Rastgele değişken kabulü ile üçgen metod ile gösterilen verilerin aşamalı ve kümülatif hasar miktarının ayrımı yapılmıştır. Yaygın olarak kullanılan deterministik hasar tahminleri ile birlikte Zincir Merdiven Metod (Chain-Ladder) ve Bornhuetter-Ferguson metodu açıklanmıştır. Bu metodlar basit ve anlaşılır olması nedeniyle birçok stokastik modelin temelini oluşturmaktadır.

İkinci bölümde genelleştirilmiş doğrusal modelin ilkeleri özetlenmiştir. Bu bölümde modelin yapısı ve muhtemel parametre tahmin prensipleri açıklanmıştır. Model parametrelerinin diğer metodları ayrıca ilave edilmiştir.

Üçüncü bölümde hasar rezevlerinin hesaplanmasında genelleştirilmiş doğrusal modelin uygulanması üzerinde çalışılmıştır. Aktüeryal açıdan ilgili teoriyi tahmin eden dağılımlar birçok modelle belirtilmiştir. Veri bağımlılığını, modelleme olasılıklarını genişletmek amacıyla modelde bağımlı ve bağımsız değişkenler arasındaki ilişkiler kullanılmıştır.

Dördüncü bölümde bootstrap teorisi açıklanmıştır. Algoritma, artıkların yeniden örneklenmesi gereken genelleştirilmiş doğrusal modeller için ayarlanmıştır. Bootstrap yönteminin adımlarının artıkları ve özel karakteristikleri, ayrı modellenmiş dağılımlar için birleştirilmiştir.

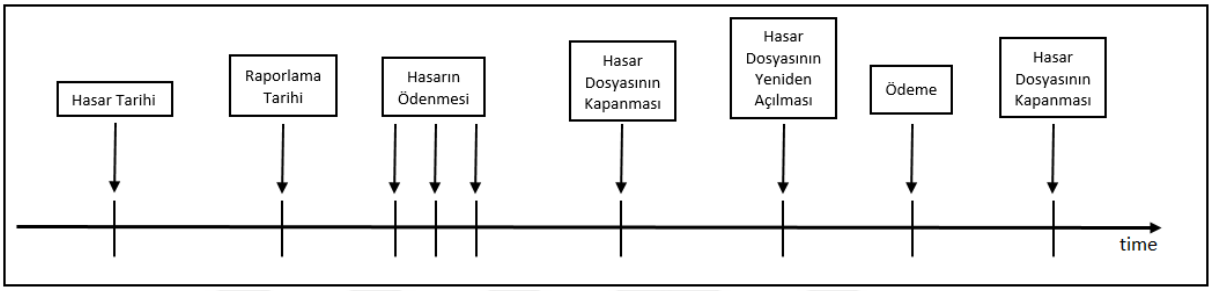
Son olarak beşinci bölümde, güncel veriler üzerinde uygulama gerçekleştirilmiştir. Modelleri verilere uyarlamak ve sonuçta bootstrap tahminli dağılımlar elde etmek için R yazılımı kullanılmıştır. Elde edilen sonuçlar tablo ve şekiller ile açıklanmıştır. Uygun modellerin performans ve yeterliliği tartışılarak değerlendirilmiştir.

1 HASAR REZERV TEORİSİNE GİRİŞ

Bu çalışmada hayat sigortası haricinde her türlü sigorta ürününün hasar rezervlerini incelemek amaçlanmaktadır. Hayat ve hayat-dışı sigorta sözleşmeleri arasında ürünlerin farklı modellenmesine sebep olan farklılıklar mevcuttur.

Hayat-dışı hasar geçmişi Tablo 1.1'de gösterilmektedir. (Wüthrich and Merz (2008))

Tablo 1.1. Hayat Dışı Sigortada Hasarın Gelişimi



Genellikle hasar talebi ile hasar dosyasının kapanması arasında zaman farkı mevcuttur. Bunun temel sebebi hasarın gecikmeli olarak raporlanması, bazı durumlarda hasar talebinde bulunmanın zaman alması veya kapalı dosyaların yeniden açılmasıdır.

Hasar talebi sürecinin bu tip özelliklerinden dolayı, hayat-dışı sigorta şirketleri sigortalıların üzerindeki risklerden doğan yükümlülüklerini karşılamak için rezerv süreçleri oluşturmuşlardır.

Mandl ve Mazurova'nın (1999) belirttiği gibi, hasar rezervlerinde iki ana değerin tahmini söz konusudur. Birincisi ihbar edilmiş ancak kapatılmamış dosya muallaklarını içeren RBNS (Reported But Not Settled) ve ikincisi ise gerçekleşmiş ancak ihbar edilmemiş hasarları içeren IBNR (Incurred But Not Reported) hesabıdır. Rezerv ayrıca hasar dosyalarının çözümlenmesi sırasında ortaya çıkan masrafların tahminini içermektedir.

RBNS hasarlarının karşılanması için hesaplanan rezervin değeri, hayat-dışı sigortacılık uzmanı tarafından tahmin edilebilirken, IBNR hasarlarının karşılanması için

hesaplanan rezervin değeri tarihsel gelişimi kullanarak matematiksel yöntemlerle yaklaşık olarak hesaplanmaktadır.

1.1 Hasar Süreçleri ve Hasar Rezervleri

Hayat-dışı sigorta türlerinde, rezerv hesabında kullanılan geçmiş veriler genellikle üçgen merdiven metodu ile gösterilmektedir. Hasar gelişimini gösteren üçgen merdiven metodunda hasar bilgileri (hasar tutarı, hasar sayısı…) kaza yılı (hasarın oluşum yılı) ve gelişim yılına (hasar ödemesinin devam ettiği yıl) göre sınıflandırılmaktadır.

Son hasarın oluşma yılı I ve maximum hasar gelişim yılı J ile gösterilmektedir. X_{ij} raslantı değişkeni ve i hasar oluşum yılı (kaza yılı) ve j hasar ödemesinin devam ettiği yılı (gelişim yılı) göstermek üzere, birikimli hasar $C_{ij} = \sum_{k=0}^{j} X_{i,k}$ olarak gösterilmektedir.

Tablo 1.2. Hasar Gelişim Üçgeni

Kaza Yılı (i)	Gelişim Yılı (j)						
	0	**1**	**…**	**…**	**…**	**J-1**	**J**
0	$X_{0,0}$	$X_{0,1}$	…	…	…	$X_{0,J-1}$	$X_{0,J}$
1	$X_{1,0}$	$X_{1,1}$	…	…	…	$X_{1,J-1}$	$X_{1,J}$
2	$X_{2,0}$	$X_{2,1}$	…	…	…	$X_{2,J-1}$	$X_{2,J}$
.							
.							
.							
.			Geçmiş hasar bilgisi $X_{i,j}$				
I - J			(i+j ≤ 1)				
.							
.							
.						Gelecek hasar bilgisi $X_{i,j}$	
.						(i+j > 1)	
I - 1	$X_{I-1,0}$	$X_{I-1,1}$	…	…	…	$X_{I-1,J-1}$	$X_{I-1,J}$
I	$X_{I,0}$	$X_{I,1}$	…	…	…	$X_{I,J-1}$	$X_{I,J}$

$C_{i,j}$ değeri, kaza yılı i'nin nihai hasar tutar miktarını temsil etmektedir. Bazen $X_{i,j}$ i'nci yıldaki iddia sayısını ve j raporlama zamanı gecikmesini belirtmektedir. $C_{i,j}$ değeri, i kaza yılında ve $i + j$ zamanına kadar rapor edilen toplam hasar sayılarını ve $C_{i,J}$ i yılda gerçekleşen toplam hasar sayısını ifade etmektedir.

Şekil 1.2'de görüldüğü gibi $X_{i,j}$ değeri sadece $X_{i,j}\ i+j \leq I$ için bilinebilmektedir (üst üçgen), diğer taraftan $X_{i,j}$ değerinin olağan ödemelerinin $i+j > I$ (alt üçgen) için tahmin edilmesi gerekmektedir.

Üst üçgen;

$$D_I^U = \{ X_{i,j}; i+j \leq I, 0 \leq j \leq J \}$$

ve alt üçgen;

$$D_I^L = \{ X_{i,j}; i+j > I, 0 \leq j \leq J \}$$

olarak gösterilmektedir.

Kümülatif hasar tutarı için kümülatif hasar gelişim üçgeninin parçası olarak I zaman diliminde gözlem setini $D_I^{cU} = \{ C_{i,j}; i+j \leq I, 0 \leq j \leq J \}$ ve tahmin edilecek değer $D_I^{cL} = \{ C_{i,j}; i+j > I, 0 \leq j \leq J \}$ ile gösterilmektedir.

Hasar rezervi probleminde asıl önemli husus bekleyen hasar yükümlülüklerinin değeridir. Kaza yılı i için i+j zamanında verilen değer;

$$R_{i,j} = \sum_{k=j+1}^{J} X_{i,k} = C_{i,J} - C_{i,j},$$

burada $X_{i,j}$ artan ödemelerdir. Hasar rezervleri, geçmiş hasar tutarları $C_{i,j}$ ve i kaza yılı nihai hasar tutarı $C_{i,J}$'nin tahminini içeren $R_{i,j}$ tahminine eşittir. Kaza yılı i için olağanüstü zarar yükümlülükleri ve nihai hasar miktarları $R_{i,j} = C_{i,J} - C_{i,I-i}$ ve C_i ile gösterilmektedir.

Bu tezde biz I ve J eşitliği ile tüm j>J değerleri için $X_{i,j} = 0$ eşitliğini temel alan Wüthrich ve Merz (2008) varsayımını temel almaktayız. Bu varsayım kaza yılları $i = 1,2,\dots,I$ için $R_{i,j}$ değerinin tahmin edilebileceğini öngörmektedir.

1.2 Temel Hasar Rezervleri Yöntemi

Bu bölümde hasar rezervlerinin hesaplanmasında kullanılan Zincir Merdiven Metod (Chain-Ladder (CL)) ve Bornhuetter-Ferguson (BF) tahmin yöntemlerini inceleyeceğiz. Bu yöntemler gelecekteki yükümlülüklerimizi tahmin etmek ve IBNR'ın hesaplanmasında yardımcı olmaktadır. Bununla birlikte rezerv hesaplamasındaki belirsizlikler bu yöntemler ile ölçülemez bu nedenle hesaplamaları uygun bir stokastik modele dayandırmak gerekmektedir.

1.2.1 Zincir Merdiven (Chain-Ladder) Metodu

Zincir Merdiven metodu, nihai hasar maliyetini tahmin etmek amacıyla kullanılan en eski ve en popüler yöntemlerden biridir. Geçmiş dönem tecrübelerine bağlı basit bir algoritmaya dayanmaktadır. Zincir Merdiven metodu sonucu aralık tahmini değil bir nokta tahmini olması sebebiyle deterministiktir.

Zincir Merdiven Yöntemi'nin amacı;

- Rezerv miktarının yeterliliğinin kontrolü
- Rezerv miktarının gelecekteki yükümlülükleri karşılamadaki yeterliliği
- Şirketin poliçe sahiplerine karşı yükümlülüklerini yeterli düzeyde karşılayıp karşılayamadığının kontrolünün sağlanmasıdır.

Zincir Merdiven metodu rezerv tahminlerinin standart hatası için 'dağılımdan-bağımsız (distribution-free)' formül türetmiştir (Mack (1993)). Temeli Schnieper (1991) tarafından geliştirilen 'standart hata' formülüne ek olarak Mack zincir merdiven yönteminde rezervin varyans tahminini elde edilmiştir.

Dağılımdan-bağımsız Zincir Merdiven modelinde birikimli (toplam) hasar tutarı $C_{i,j}$, $0 \leq i \leq I, 0 \leq j \leq J$, daha önce listelendiği gibi $i + j \leq I$ için bilinen değerlere sahip rastgele değişkenler olarak kabul edilir.

Model varsayımları 1.1.

Birikimli hasar tutarı $C_{i,j}$ farklı kaza yılları i'nin bağımsız rastgele değişkenidir. Bu,

$\{C_{i,0}, \dots, C_{i,I}\}$, $\{C_{k,0}, \dots, C_{k,I}\}$ i ≠ k ve birbirinden bağımsız (1.1)

Gelişim faktörleri ve bağlantı oranları $\lambda_0, \dots, \lambda_{j-1} > 0$ ve

$$E(C_{i,j} \mid C_{i,0}, \dots, C_{i,j-1}) = E(C_{i,j} \mid C_{i,j-1}) = \lambda_{i,j-1}\, C_{i,j-1} \quad (1.2)$$

Tüm $0 \le i \le I$, $1 \le j \le J$ değerleri için geçerlidir.

(1.2) varsayımı yalnızca ilk anları dikkate alır ancak gelecekteki hasar miktarlarının koşullu beklenen değerlerinin tahmini için gerekli bilgiyi sağlamaktadır. Eğer tahminlerin çeşitliliği modellemek isteniyorsa, daha uzun anlar için modelleme yapılabilir.

Model (1.1) varsayımı bilinmeyen kalkınma faktörlerini λ_j tahmin etmeye yarayan Zincir Merdiven metodu algoritmasını tanımlamak için yeterlidir.

$$\widehat{\lambda_l} = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}}, \text{ j = 0, 1, …., J-1.}$$

Önceki varsayımlar ve koşullu tahmin edici (beklenti) özellikleri altında aşağıdaki denklemi türetebiliriz.

$$\begin{aligned} E(C_i \mid D_I^{cU}) &= E\,(C_i \mid C_{i,0}, \dots, C_{i,I-i}) \\ &= E\,[E\,(C_i \mid C_{i,I-1}) \mid C_{i,0}, \dots, C_{i,I-i}] \\ &= \lambda_{I-1} E(C_{i,I-1} \mid D_I^{cU}) \\ &= \dots = \lambda_{I-1} \dots \lambda_{I-i}\, C_{i,I-i} \end{aligned}$$

Kaza yılı i için nihai hasar tutarı C_i'nin Zincir Merdiven metodu tahmini,

$$\hat{C}_i^{CL} = \hat{E}(C_i \mid D_I^{cU}) = \hat{\lambda}_{I-1} \dots \hat{\lambda}_{I-i}\, C_{i,I-i} \quad (1.3)$$

ve kaza yılı i için öngörülen muallak hasar yükümlülüklerinin R_i tarafından tahmini,

$$\hat{R}_i^{CL} = \hat{E}(C_i|D_I^{cU}) - C_{i,I-i} = (\hat{\lambda}_{I-1} \dots \hat{\lambda}_{I-i} - 1)\, C_{i,I-i} \quad (1.4)$$

I = J varsayımı altındadır.

$\hat{\lambda}_j{}'nin\ (j = 0, \dots, J-1)$ gelişim faktörlerinin tarafsız ve birbirleriyle ilişkili olmayan tahmincileri olduğunu kanıtlamıştır (Mack (1993)). Sunulan karakteristik nihai hasar tutarı tahmincileri $\hat{C}_i^{CL}$ ve muallak hasar tutarı R_i'nin tahmincisi $\hat{R}_i^{CL}$ tarafsızlığını ispat için kullanılmaktadır.

1.2.2 Bornhuetter-Ferguson Metodu

Hasar rezerv yöntemlerinin bir diğer yaklaşımı Bornhuetter ve Ferguson tarafından 1972 yılında ortaya atılan Bornhuetter-Ferguson metodudur. Hasar-Prim oranını ve buna ait trendi kullanan bir yöntemdir. Tahmini hasar-prim oranı ile Projekte modelin kombinasyonudur. Gözlemlenen verilerdeki sapmaları göz önünde bulundurmaksızın, bu yöntem stokastik yöntemlere dayanan sağlam bir metottur.

BF Metodu varsayımları aşağıdaki gibidir.

Model varsayımları 1.2.

Birikimli hasar tutarı $C_{i,j}$ farklı kaza yılları i'nin bağımsız rastgele değişkenidir.

Burada $\mu_0, \dots, \mu_I > 0$ parametresi ve hasar gelişim deseni $\beta_0, \dots, \beta_J > 0$ ve $\beta_J = 1$ olmak üzere,

$$E(C_{i,0}) = \beta_0 \mu_i\ , \quad (1.5)$$

$$E\ (C_{i,j+k} \mid C_{i,0}, \dots, C_{i,j}) = C_{i,j} + (\beta_{j+k} - \beta_j)\ \mu_i$$

$0 \leq\ i\ \leq I\ , 0 \leq\ j \leq J-1$ ve $1 \leq\ k \leq J-j$ değerleri için geçerlidir.

Yani varsayım (1.5) ile,

$$E(C_{i,j}) = \beta_j \mu_i \text{ ve } E(C_i) = \mu_i$$

çıkarımlarında bulunmaktayız.

Birikimli hasar tutarı $C_{i,j}$ için hasar gelişim deseni aynı zamanda ödeme şekli ve birikimli nakit akış modeli olarak da adlandırılabilir ve zamana göre farklı değerlerle ayrılmış iskonto edilmiş rezervi bulmak için kullanılır.

Varsayım (1.5), birikimli hasar tutarı $C_{i,j}$ farklı kaza yılları "i" ile birlikte BF metodu altında modellenebilen daha zayıf bir varsayımı tanımlamaktadır.

Model varsayımları 1.3.

Birikimli hasar tutarı $C_{i,j}$ farklı kaza yılları i'nin bağımsız rastgele değişkenidir.

Burada $\mu_0, \dots, \mu_I > 0$ parametresi ve desen $\beta_0, \dots, \beta_J > 0$ ve $\beta_J = 1$ olmak üzere,

$$E(C_{i,j}) = \beta_j \mu_i \tag{1.6}$$

tüm $0 \leq\ i \leq I$ ve $0 \leq\ j \leq J$ değerleri için geçerlidir.

BF yönteminin algoritması $E(C_i \mid D_I^{cU})$ faizin tahmin edilmesidir.

(1.2) varsayımı altında,

$$\begin{aligned} E(C_i \mid D_I^{cU}) &= E\,(C_i \mid C_{i,0}, \dots, C_{i,I-i}) \\ &= C_{i,I-i} + E(C_i - C_{i,I-i}) \\ &= C_{i,I-i} + (1 - \beta_{I-i})\mu_i \end{aligned} \tag{1.7}$$

Bunun yerine varsayım (1.6), (1.5) yerine göz önüne alındığında formül,

$$E(C_i \mid D_I^{cU}) = E\ (C_i \mid C_{i,0}, \dots, C_{i,I-i})$$

$$= C_{i,I-i} + E(C_i - C_{i,I-i} \mid C_{i,0}, \dots, C_{i,I-i}) \quad (1.8)$$

son dönemi meşrulaştırmakta zorlandığımızda bu formülü kullanmaktayız. Artan hasarların miktarı $C_i - C_{i,I-i}$ ve hasarlar $C_{i,0}, \dots, C_{i,I-i}$'nin birbirinden bağımsız oldukları varsayılarak kaçınabiliriz. Bu şekilde (1.7) ile aynı sonuca ulaşmış oluruz.

(1.7) ve (1.8) her iki varsayımda da son dönem bilinmemektedir. IBNR hasar tutarları $1 \leq i \leq I$ koşullarında BF metodu ile aşağıdaki formülle tahmin edilmektedir.

$$\hat{C}_{i,j}^{BF} = \hat{E}(C_i|D_I^{cU}) = C_{i,I-i} + (1 - \hat{\beta}_{I-i})\hat{\mu}_i \quad (1.9)$$

$\hat{\beta}_{I-i}$, β_{I-i}'nin tahmin edicisi ve nihai beklenen hasar tutarı $E(C_i)$, $\hat{\mu}_i$, tahmini vermektedir. Schmidt'e göre (2006), $\{\hat{\beta}_j\}_j^J = 0$ iç (hasar gelişim üçgeni) ve dış (market istatistikleri vb.) bilgileri yada her ikisinin birleşimini kullanarak elde edilen bir hasar geliştirme modelinin önceden tahmin edicisi olabilir. Dış bilgilerin kullanılması Bayesian modelin karakteristiğidir.

Gelişim faktörleri λ_j, j= 0, … , J-1 ve hasar rezervlerinin tahmini Zincir Merdiven (Chain-Ladder) metodunda açıkça belirtilirken, Bornhuetter Ferguson yönteminde hasar rezervlerinin tahmini bu kadar basit değildir. Örneğin, Bornhetter Ferguson tahminleyicileri, sözü edilen iki metodun tahmin edicilerinin karşılaştırılması ile elde edilir.

Zincir Merdiven (Chain-Ladder) metodunun altında, Model varsayımı 1.1, aşağıdaki durumlarda,

$$E(C_{i,j}) = E\ [\mathrm{E}(C_{i,j} \mid C_{i,j-1})] = \lambda_{j-1}\mathrm{E}(C_{i,j-1}) = \mathrm{E}(C_{i,0}) \prod_{k=0}^{j-1} \lambda_k\ ,$$

$$E(C_i) = \mathrm{E}\big(C_{i,0}\big) \prod_{k=0}^{J-1} \lambda_k\ ,$$

Sonuç olarak,

$$E(C_{i,j}) = \mathrm{E}(C_i)\ \prod_{k=j}^{J-1} \lambda_k^{-1},$$

Model varsayımı 1.3 ile Bornhuetter Ferguson yöntemini kullanarak, aşağıdaki denklemi

$$\prod_{k=j}^{J-1} \lambda_k^{-1} = \beta_j\ , \qquad (1.10)$$

yazabiliriz. Bununla birlikte, bu model varsayım 1.5 ile çalışmaz, çünkü Zincir Merdiven (Chain-Ladder) metodunun varsayımları tarafından denklem oluşturulmuştur.

Dolayısıyla, Zincir Merdiven (Chain-Ladder) metodunun λ_0, …, λ_{J-1} gelişim faktörleri bilinirken, Bornhuetter Ferguson yöntemini $\{\beta_j\}_j^J = 0$ gelişim deseni oluşturabilir. Bu sebeple 1.10 model varsayımı ile Bornhuetter Ferguson tahmincisini aşağıdaki gibi yeniden yazabiliriz

$$\hat{C}_i^{BF} = C_{i,I-1} + (1 + \frac{1}{\prod_{j=I-i}^{J-1} \widehat{\lambda_J}})\hat{\mu}_i \qquad (1.11)$$

ve Zincir Merdiven Metodu (Chain-Ladder) tahmincisi (1.3) formülüyle yorumlanırsa

$$\hat{C}_i^{CL} = \prod_{j=I-i}^{J-1} \widehat{\lambda_j}$$

$$= C_{i,I-i} + C_{i,I-i}\ (\prod_{j=I-i}^{J-1} \widehat{\lambda_j} - 1)$$

$$= C_{i,I-i} + \frac{\hat{C}_i^{CL}}{\prod_{j=I-i}^{J-1} \widehat{\lambda_J}} (\prod_{j=I-i}^{J-1} \widehat{\lambda_j} - 1) \qquad (1.12)$$

$$= C_{i,I-i} + \left(1 - \frac{1}{\prod_{j=I-i}^{J-1} \widehat{\lambda_J}}\right) \hat{C}_i^{CL}$$

(1.11) ve (1.12) model varsayımları karşılaştırılarak Zincir Merdiven Metodu (Chain-Ladder) ve Bornhuetter Ferguson yöntemleri ile $\{\beta_j\}_j^J = 0$ gelişim deseninin tanımlanması arasındaki temel farklılık tahmin edici $\hat{C}_i^{CL}$ kullanılarak oluşur ve sırasıyla gözlemlere ve

önceden tahmin $\hat{\mu}_i$'ye dayandırılmıştır. İkincisi genellikle bir iş planı tarafından verilen değer veya prim hesaplamaları için kullanılan değerdir ve herhangi bir gözlem kaydedilmeden önce uzman tarafından tanımlanmalıdır.

Zincir Merdiven (Chain-Ladder) metodunun aksine, Bornhuetter Ferguson yönteminde ilk gelişim yılı ödenen nihai hasar oranlarında istikrarsızlık gözlemlenirse verimli olduğu kanıtlanmıştır. Böyle bir durumda uygulanan Zincir Merdiven (Chain-Ladder) metodu tatmin edici sonuçlar vermez. (England ve Verrall (2002))

2 GENELLEŞTİRİLMİŞ DOĞRUSAL MODELLEME

Hasar rezervlerinin stokastik modellemesinde Genelleştirilmiş Doğrusal Modelleme (GDM) çerçevesini sunmadan önce, istatistiksel modellerden olan GDM'nin problemine giriş yapmaktayız. Genelleştirilmiş Doğrusal Modelleme klasik doğrusal modellemelerin genelleştirilmiş biçimi olarak tanımlanabilir. Klasik doğrusal modellerin genelleştirilmiş hali olan bu sınıf, örneğin, lojit ve probit modelleri, log-lineer modeller, multinomial response models ve yaşam modeli (models for survival) datası içerir. Doğrusal regresyon ve varyans analiz modelleri GDM'nin özel durumları olarak değerlendirilmektedir.

2.1 Temel Varsayımlar ve Terminoloji

Klasik doğrusal modellerin gösterimini kullanarak, bir sütun vektörü gözlem,

$$\mathrm{y} = (\mathrm{y}_1, \dots, \mathrm{y}_n)^T$$

rastgele vektör **Y**'nin gerçekleşmesidir. p ≤ n için, **X** bağımsız değişkenleri bir n*p matrisi ve bilinmeyen parametrelerin p-boyutlu vektörü,

$$\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$$

şeklindedir. Y vektörünün ortalama μ vektörü bileşenlerinin bağımsız olarak dağıtıldığı varsayılmaktadır. Klasik doğrusal modellerin sistematik kısmında, ortalama vektör μ matris notasyonundaki Y vektör parametrelerinin yardımı ile,

$$\mu = \boldsymbol{X}\beta$$

şeklinde belirlenmektedir. Modelin rastgele kısmı için varsayımlar, σ^2 sabit varyanslı ve μ ortalamalı bağımsız Normal dağılımdır. Dolayısıyla GDM'ye geçişte McCullagh ve Nelder (1989) tarafından sunulan klasik doğrusal model özelliklerini benimsemekteyiz.

Bu şekilde üç parçalı bir tanımlama yapılabilir:

1. Rastgele Bileşen: Y rastgele değişkenini, σ^2 sabit varyanslı ve $E(Y) = \mu$ ortalamalı bağımsız Normal dağılım olarak tanımlayabiliriz.

2. Sistematik Bileşen: $x_j,\ j = 1,2,\dots,p$ ile verilen bağımsız değişkenler olmak üzere

$$\eta = \sum_{j=1}^{p} x_j \beta_j \text{ doğrusal önkestirimdir.}$$

3. Bağ: Doğrusal önkestirim (η) ile Y rastgele değişkeninin beklenen değerini (μ) ilişkilendiren fonksiyondur.

Doğrusal önkestirim bileşenleri için bir bağlantı fonksiyonu,

$$\eta_i = g(\mu_i) \qquad (2.1)$$

$i = 1,2,\dots,n$ değerleri için yazabiliriz.

Bağ fonksiyonu tüm gerçek sayı aralığında ve normal dağılım varsayımı altında herhangi bir artan türevlenebilir fonksiyon olabilir ve Y rastgele değişkeninin üstel bir dağılım ailesinden geldiği varsayımı altında aşağıdaki şekilde tanımlanabilir.

Tanım 2.1. Y rastgele değişkeninin üstel türden bir dağılıma sahip olduğu varsayımı altında a(.), b(.) ve c(. , .) bazı gerçek değerli özel fonksiyonlar olmak üzere, üstel ailenin olasılık yoğunluk fonksiyonu

$$dF(y) = \exp\left\{\frac{y\theta - b(\theta)}{\alpha(\varphi)} + c(y,\varphi)\right\} dv(y), y \in A \subseteq R, \qquad (2.2)$$

şeklinde ifade edilmektedir. θ parametresi kanonik (standart) bir parametre olarak, $\varphi > 0$ ise dağılım parametresi olarak adlandırılır.

Verilen Y rastgele değişkeninin θ ve φ'nin log-olabilirlik fonksiyonu,

$$l(\theta, y) = \left\{\frac{y\theta - b(\theta)}{\alpha(\varphi)} + c(y, \varphi)\right\}, \quad (2.3)$$

biçiminde yazılabilmektedir.

Y rastgele değişkeninin ortalaması ve varyansı sırasıyla,

$$E(Y) = \mu = b'(\theta)$$

ve

$$Var(Y) = b''(\theta)\alpha(\varphi)$$

şeklinde ifade edilmektedir. $b'(\theta)$ ve $b''(\theta)$, sırasıyla θ'ya göre birinci dereceden ve ikinci dereceden türevleri ifade etmektedir.

(2.2)'deki b(.) fonksiyonu, dağılımın geldiği aileyi daha belirgin bir biçimde tanımlar ve y'nin alanını belirtir.$\alpha(\varphi)$ fonksiyonu, φ gözlemler arası sabit yayılım parametresi ve w gözlemden gözleme farklılık gösteren önsel ağırlık olmak üzere, $\alpha(\varphi) = \varphi / w$ biçiminde ifade edilmektedir. w ağırlık değeri olarak adlandırılır ve sabit bir değerdir. Gruplandırma işlemi gerçekleşmemiş ya da frekans biçiminde ifade edilemeyen veriler için w=1 kabul edilir (McCullagh ve Nelder, 1989).

2.2 Bağ Fonksiyonu

Önceki bölümde de belirtildiği gibi, bağ fonksiyounu, doğrusal önkestirim (η) ile Y rastgele değişkenin beklenen değerini (μ) ilişkilendiren bir fonksiyondur. Klasik Doğrusal Modelleme'de, ortalama ve doğrusal önkestirim özdeş ve η ve μ'nün her ikisi de reel sayılardan herhangi bir değer alabildiği için özdeş bağ kullanılmaktadır. Dağılımın Poisson olduğu ve sayımlarla ilgilenildiğinde, $\mu > 0$ olması gerektiğinden, özdeş bağ yetersiz kalır ve doğrusal önkestirim ve beklenen değer arasında farklı fonksiyonel ilişkilere ihtiyaç duyulmaktadır. $\eta = g(.)$ bağ fonksiyonu olup, bağımlı değişkenlerin ortalamasıyla bağımsız değişken arasındaki ilişkiyi belirleyen, monoton ve diferansiyellenebilir bir fonksiyondur. Genelleştirilmiş Doğrusal Modelleme'de bağ

fonksiyonu seçimi, oluşturulacak model sonuçlarını direkt etkilemektedir. Bu nedenle, $b(\theta)$'ya bağlı olarak belirlenen g(.) bağ fonksiyonunun seçimi, Genelleştirilmiş Doğrusal Modelleme'nin en önemli adımıdır. (De Jong ve Heller, 2008).

Önemli bağ fonksiyonları,

1. Lojit model;

$$\eta_i = g(\mu_i) = \log(\frac{\mu_i}{1-\mu_i}); \tag{2.4}$$

2. Probit model;

$$\eta_i = g(\mu_i) = \varphi^{-1}(\mu_i), \tag{2.5}$$

$\varphi(.)$ standat normal dağılımın kümülatif dağılım fonksiyonudur.

3. Tamamlayıcı log-log model;

$$\eta_i = g(\mu_i) = \log[-\log(1-\mu_i)]; \tag{2.6}$$

biçimindedir.

$g(.) = \theta$ olduğunda $g(.)$, $b(\theta)$'ya bağlı olarak belirlenen kanonik bağ fonksiyonu olarak adlandırılır. $g(.)$ Bağ fonksiyonunun kanonik belirlenmesi, hesaplamalarda kolaylık sağlamaktadır.

Bazı dağılımlar için kanonik bağlar;

Normal	$\eta = \mu$
Poisson	$\eta = \log \mu$
Binom	$\eta = \log\{\frac{\pi}{(1-\pi)}\}$
Gamma	$\eta = \mu^{-1}$
Ters-Gauss	$\eta = \mu^{-2}$

biçimindedir (McCullagh and Nelder, 1989).

2.3 Model Fitleme

Önceki bölümlerde, belirli bir veri kümesine bir modelin yerleştirilmesinde kullanılabilecek birkaç olası bağlantı fonksiyonunu göstermiştik. Bir bağ fonksiyonunun seçimi verilerin karakterine bağlıdır ve bir şekilde kişinin seçimine bağlıdır. Bağ fonksiyonu, $g(\mu_i)$ için genel gösterim,

$$g(\mu_i) = \boldsymbol{x}_i^T \boldsymbol{\beta}, \text{ i = 1, …, n,}$$

ve μ_i Y_i'nin beklenen değeri, $\boldsymbol{x}_i^T$ tasarım matrisi X'in i'inci dizisi – sırası ve $\boldsymbol{\beta}$ parametresi vektörü p ≤ n boyutundadır. Maksimum parametre sayısına sahip modele, doymuş veya tam model adı verilir ve yanıtlar arasındaki tüm varyasyonların rastgele bir bileşen ile modellendiği bir parametreye sahip modele, boş model adı verilir. Doymuş model pratikte bilgilendirici değildir, ancak fit edilmiş bir modelin yeterliliğini ölçerken faydalıdır.

2.3.1 GDM'de En Çok Olabilirlik Kanunu

Belli bir model seçildikten sonra, onun parametrelerini tahmin etmesi gerekir. GDM'de, parametrelerin tahmincileri en çok olabilirlik yöntemi kullanılarak elde edilir.

Rastgele değişkenin Y bileşenlerinin her biri için log-olabilirlik fonksiyonu;

$$l(\theta_i, \varphi, y_i) = \frac{y_i\theta_i - b(\theta_i)}{a(\varphi)} + c(y_i, \theta_i)$$

ve Y'nin tüm olasılıkları için log-olabilirlik fonksiyonu;

$$l(\boldsymbol{\theta}, \varphi, \boldsymbol{y}) = \sum_{i=1}^{n} l(\theta_i, \varphi, y_i)$$

şeklindedir.

Model parametreleri β'nin maksimum olabilirlik tahmincileri için skor istatisitiği aşağıdaki gibidir.

$$\mathbf{U} = \frac{\partial l}{\partial \beta} = \sum_{i=1}^{n} \left(\frac{\partial l_i}{\partial \beta}\right) = \sum_{i=1}^{n} \left(\frac{\partial l_i}{\partial \theta_i} \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta}\right)$$

$$= \frac{1}{\alpha(\varphi)} \sum_{i=1}^{n} \left[\frac{(y_i - \mu_i)}{var(Y_i)} \left(\frac{\partial \mu_i}{\partial \eta_i}\right) \boldsymbol{x_i}\right]$$

Parametrelerin vektörünü çözen maksimum olabilirlik tahmincisi β ve denklemi,

$$0 = \mathbf{U} = \frac{1}{\alpha(\varphi)} = \sum_{i=1}^{n} \left(\mathrm{w}(\hat{\mu}_i)(y_i - \hat{\mu}_i) \left(\frac{\partial \hat{\eta}_i}{\partial \hat{\mu}_i}\right) \boldsymbol{x_i}\right) \qquad (2.7)$$

burada fonksiyon w(.)'nin işlevi aşağıdaki şekildedir.

$$\mathrm{w}(\mu_i) = \frac{1}{var(Y_i)} \left(\frac{\partial \hat{\eta}_i}{\partial \hat{\mu}_i}\right)^2, i = 1, \dots, n.$$

Terimi ekleyerek

$$\left(\mathbf{X}^T \widehat{\mathbf{W}} \mathbf{X}\right) \hat{\beta} = \left[\sum_{i=1}^{n} (\mathrm{w}(\hat{\mu}_i) \boldsymbol{x_i} \boldsymbol{x_i}^T)\right] \hat{\beta}, \qquad (2.8)$$

ve,

$$\hat{\beta} = \left(\mathbf{X}^T \widehat{\mathbf{W}} \mathbf{X}\right)^{-1} \mathbf{X}^T \widehat{\mathbf{W}} \hat{\boldsymbol{z}}. \qquad (2.9)$$

Vektör $\hat{z} = (\hat{z}_1, \dots, \hat{z}_n)^T$ ayarlanmış bir yanıtı gösterir;

$$\hat{z}_i = \boldsymbol{x_i}^T \hat{\beta} + \left(\frac{\partial \hat{\eta}_i}{\partial \hat{\mu}_i}\right)(y_i - \hat{\mu}_i), i = 1, \dots, n.$$

(2.6) denklemi, ağırlıklı en küçük karelerin (AEKK) tekniğinin kullanıldığı doğrusal bir model için normal denklemler ile aynıdır. Bununla birlikte, pratikte GDM'de

yer alan maksimum olabilirlik tahmincisi β yaygın olarak tekrar edilen ağırlıklı en küçük kareler yöntemi ile elde edilir ve denklem;

$$\hat{\beta}^{(m+1)} = \left(\mathbf{X}^T \widehat{\mathbf{W}}^{(m)} \mathbf{X}\right)^{-1} \mathbf{X}^T \widehat{\mathbf{W}}^{(m)} \hat{\boldsymbol{z}}^{(m)}.$$

şeklindedir.

Prosedürün algoritması dört adıma sahiptir:

1. İlk tahmin ile başlamak $\widehat{\boldsymbol{\mu}}^{(0)}$;
2. Düzeltilmiş bir yanıt vektörü $\hat{\boldsymbol{z}}^{(0)}$ ve ağırlıkların matrisi $\widehat{\mathbf{W}}^{(0)}$ hesaplanması;
3. $\hat{\beta}^{(1)}$'in ağırlıklandırılmış en küçük karelere göre hesaplanması;
4. 2 ve 3. adımlar yakınsanana kadar tekrarlanır.

Fisher'in skorlama yönteminden türetilen ağırlıklıklı en küçük kareler yönetiminin kullanım gerekçesini Dobson (2002)'de bulabiliriz.

2.3.2 Yarı (Quasi)-Olabilirlik Yöntemi

Genelleştirilmiş Doğrusal model parametrelerini tahmin ederken maksimum olabilirlik yöntemini kullanmaktayız. Ancak uygulamada, gözlenen rastgele değişkenlerin dağılımının, uygulamadan sağlanan yetersiz bilgi nedeniyle fark edilemediği ve maksimum olabilirlik fonksiyonunun oluşturulamayacağı görülmektedir. Böyle bir durumda, kovaryantların bağımlı değişkenleri üzerindeki etkisini tahmin etmek için yarı (quasi) – olabilirlik yöntemi kullanılabilir. Yarı (quasi) – olabilirlik fonksiyonunun özelleştirilmesi için sadece gözlemlerin ortalama ve varyansı arasındaki ilişki yeterlidir.

$E(Y) = \mu$ ve varyansın $var(Y) = \varphi \boldsymbol{V}(\mu)$ tarafından verilen ortalamanın bir fonksiyonu olduğunu varsayalım. φ'nın β'den bağımsız bilinmeyen bir sabit olacağı ve $\boldsymbol{V}(\mu)$ bilinen fonksiyonların bir matrisi olacak şekilde $\boldsymbol{V}(\mu) = \boldsymbol{diag}\{V_1(\mu_1), \dots, V_n(\mu_n)\}$, i=1, …., n ve $V_i(\mu_i)$ fonksiyonu yalnızca μ vektörünün i'inci bileşenine bağlıdır.

Y fonksiyonu;

$$\mathrm{U} = u(\mu; \mathrm{Y}) = \frac{Y - \mu}{\varphi V(\mu)}$$

özelliklerle,

$$\mathrm{E}(\mathrm{U}) = 0,$$

$$\mathrm{var}(U) = \frac{1}{\varphi V(\mu)}\,, \tag{2.10}$$

$$-\mathrm{E}\left(\frac{\partial U}{\partial \mu}\right) = \frac{1}{\varphi V(\mu)}\,.$$

Ve yarı (quasi)-olabilirlik fonksiyonu integral ile;

$$\mathrm{Q}(\mu; y) = \int_{y}^{\mu} \frac{y - t}{\varphi V(\mathrm{t})}\, dt,$$

olarak verilmektedir.

U değişkeninin (2.7)'de belirtilen özelliklere sahip olmasıyla $\mathrm{Q}(\mu; y)$ fonksiyonu μ için bir log-olabilirlik fonksiyonu gibi davranmalıdır.

Y'nin tüm bileşenleri için yarı (quasi)-olabilirlik ilgili değişkenleri özetleyerek;

$$\mathrm{Q}(\mu; \boldsymbol{y}) = \sum_{i=1}^{n} \mathrm{Q}_i(\mu_i; y_i)$$

şeklinde elde edilir.

l'nin bir log-olabilirlik fonksiyonu olduğunu kabul edersek;

$$-\mathrm{E}\left(\frac{\partial U}{\partial \mu}\right) \leq -\mathrm{E}\left(\frac{\partial^2 l}{\partial \mu^2}\right) \qquad (2.11)$$

(2.7)'ye göre U'nun özelliklerinden ve Cramer-Rao eşitsizliğinden yukarıdaki sonuca ulaşırız. Wedderburn (1974) bir rastgele değişkeninin varyansının Y ve bir parametreli üstel aile tipine ait olasılık dağılımının varsayımının sadece ortalama-varyans ilişkisinin özelleştirilmesine eşdeğer olduğunu kanıtlamıştır. Dolayısıyla, böyle bir durumda log-olabilirlik ve yarı (quasi)-olabilirlik fonksiyonu aynıdır.

2.3.3 Uyum İyiliğinin Ölçülmesi

Gerçek model parametreleri β'nın maksimum olabilirlik tahmincileri elde edildiğinde, modelin tutarsızlığı odak noktası haline gelebilir. Maksimum olasılık parametreleri ile oluşan tam modelin maksimum log-olabilirlik fonksiyonu $l(\widehat{\boldsymbol{\theta}}^*, \varphi; y)$, burada dağılım parametresi $\varphi' nin$ sabitleneceğini varsayıyoruz. Araştırılan modelin sapması

$$\mathrm{D} = 2\left[l(\widehat{\boldsymbol{\theta}}^*, \varphi; y) - l(\widehat{\boldsymbol{\theta}}, \varphi; y)\right], \qquad (2.12)$$

ve erişilen maksimum log-olabilirlik ile fit edilen model tarafından elde edilen maksimum log-olabilirlik arasındaki fark iki kattır. Dobson'dan (2002) türetildiği üzere, fit edilen modelin sapmasının örnekleme dağılımı $x^2(m-p, v)$, burada $m-p$ tam ve fit edilen model parametrelerinin sayısı arasındaki farktır ve v merkezi olmayan parametredir.

Uyumun tutarsızlığını ölçmek için kullanılan diğer önemli istatistik, Pearson X^2 istatistiği,

$$X^2 = \sum_{i=1}^{n} \frac{(\mathrm{y}_i - \hat{\mu}_i)}{V(\hat{\mu}_i)},$$

burada $V(\hat{\mu}_i)$, ilgilenilen dağılım için tahmini varyans fonksiyonudur. Yine Pearson X^2 istatistiği, normal doğrusal modeller için kesin bir X^2 dağılımına sahiptir. Ancak bu diğer dağılımların bağımlı değişkenleri için genel olarak geçerli değildir.

Bir modelin uygunluğunun yeterliliği, artıklarını kontrol ederek de değerlendirilebilir. Genelleştirilmiş Doğrusal Modellerde kullanılan artıkların tanımları, normal olanın yerine geçen dağılımlara uygulanacak şekilde genişletilir. McCullagh ve Nelder (1989)'da listelenen artıklardan örnek olarak Pearson artıkları şu şekilde tanımlanmaktadır;

$$r_i^{(P)} = \frac{\mathrm{y}_i - \hat{\mu}_i}{\sqrt{V(\hat{\mu}_i)}} \tag{2.13}$$

ve sapma artıkları;

$$r_i^{(D)} = sgn(\mathrm{y}_i - \hat{\mu}_i)\sqrt{d_i}\,,$$

$\mathrm{d}_i = 2\{y_i(\hat{\theta}_i^* - \hat{\theta}_i) - [b(\hat{\theta}_i^*) - b(\hat{\theta}_i)]\}$ ve $\mathrm{D} = \frac{1}{a(\hat{\varphi})}\sum_{i=1}^{n}\left(r_i^{(D)}\right)^2$ için geçerlidir.

Genelleştirilmiş Doğrusal Modellerde, bağlantı fonksiyonunu, varyans fonksiyonunu ve doğrusal tahminleyici terimleri bu araçlar ile inceleyebilirsiniz.

2.3.4 Bağımsız Değişkenlerin Test Edilmesi

Model uyumluluğunda, bağımsız değişkenlerin test edilmesi oldukça önemlidir çünkü modeller fazla parametreleştirilmediğinde daha doğru tahmin vermektedirler. Bu özellikle hayat-dışı sigortalarda yaygın olan küçük data örnekleri için geçerlidir.

Genelleştirilmiş Doğrusal Modelleme için üç test kullanılmaktadır. $\widehat{\beta}$ $p + k \leq n \; ve \; k \geq 1$ parametreli modeldeki parametre tahmincisi ve $\tilde{\beta}$ tahhmincisi p parametreleri olan bir alt model altında $C\beta = r$ kısıtlamaları uygulanmaktadır. Bu durumda C matrisi k satıra sahiptir. Gösterimi basitleştirmek için β fonksiyonunu sırasıyla $L(\beta)$ ve $l(\beta)$ yani olabilirlik ve log-olabilirlik fonksiyonları olarak yazarız.

Olabilirlik Oranı Testi

Olabilirlik Oranı testi (LR), model ve ilgili alt modeli altında maksimum log-olabilirlik fonksiyonlarının karşılaştırılmasına dayanmaktadır. Olabilirlik oranı testi yaparken $\hat{\beta}$ ve $\tilde{\beta}$ gereklidir. Test istatistiği aşağıdaki gibidir.

$$\lambda_i = 2\log\left[\frac{L(\hat{\beta})}{L(\tilde{\beta})}\right] = \left[l(\hat{\beta}) - l(\tilde{\beta})\right]$$

Olabilirlik oranı testi (LR) daima negatif değildir ve sınırlı modelin geçerli olduğu sıfır hipotezi altında X_k^2 dağılımına sahiptir. λ_n'in büyük değerleri gözlemlendiğinde hipotez reddedilir.

Olabilirlik oranı (LR) istatistiğinin dağılımında parametre φ dahil edilir ve bu nedenle tahmin edilmelidir. Bununla birlikte φ tahmincisi $\hat{\varphi}$ tutarlı ise X^2 dağılımının kullanımı halen geçerlidir.

Wald Testi

Wald testi için sadece $\hat{\beta}$ gereklidir. Test istatisiği;

$$W_n = \frac{1}{a(\varphi)}\left(C\hat{\beta} - r\right)^T [\boldsymbol{C}(\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{C}^T]^{-1}\left(C\hat{\beta} - r\right)$$

ve sıfır hipotezi altında, ağırlık matrisi $\boldsymbol{W}$ tahmincisi ile değiştirildiğinde X_k^2 dağılımına sahip olacaktır. Büyük $\boldsymbol{W}$ değerleri $C\beta = r$'nin sıfır hipotezinin reddedildiğini gösterir.

Skor Testi

Wald testinin aksine, skor testinde $\tilde{\beta}$ tahmincisi kullanılmaktadır. Skor test istatistiği, log-olabilirlik fonksiyonunun türevini, $\boldsymbol{U}(\tilde{\beta})$ skor istatistiğini ve varyansını kullanarak

$$R_n = a(\varphi)\boldsymbol{U}^T\left(\tilde{\beta}\right)(\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{U}(\tilde{\beta})$$

şeklinde elde edilir. Test istatistiği sıfır hipotezi altında X_k^2 dağılımına sahiptir ve $\boldsymbol{W}$ ile φ tahmincileri kullanıldığında daha az doğrudur. R_n, X_k^2 dağılımının üst kuyruk değerlerini aldığı zaman, sıfır hipotezi reddedilir.

Yukarıdaki test istatistiklerinin X^2 dağılımının sadece asimptotik olduğunu ve bu nedenle testlerin büyük örneklerde geçerli olduğunu unutmamalıyız.

3 HASAR REZERVLERİNDE GENELLEŞTİRİLMİŞ DOĞRUSAL MODELLEME

Bu bölümün amacı, hasar rezervlerinin hesaplanmasında Genelleştirilmiş Doğrusal Modellemenin uygulanmasıdır. Genelleştirilmiş Doğrusal Modellemeler bize üstel dağılım ailesinden gelen dağılımları olası bağlantı işlevleriyle birleştiren geniş bir model yelpazesi sunmaktadır. Bu birleştirmeleri, hasar gelişim üçgenlerinin kullanımı dahilinde bir sigorta şirketinin gelecekteki yükümlülüklerini tahmin etmek için kullanmaktayız.

Belirli bir dağılım seçilirken modelleme iki kısma bölünür ve bağımlı değişkenlerin hasar sayısını ve hasar tutarlarını belirtip ayırt etmesi önemlidir.

3.1 Hasar Adetlerinin Modellenmesi

Öncelikle hasar adetleri verilerinin olması durumunda modellemeyi düşünmekteyiz. Bu bölümde kaza yılı i ve gelişim yılı j için artan ve kümülatif hasar sayılarını belirtmek için $X_{i,j}$ ve $C_{i,j}$, $0 \leq i,j \leq I = J$, sembollerini kullanmaktayız.

Modelleme yapılırken Poisson ve Negatif binom dağılımları yaygın olarak kullanılır, ancak England ve Verrall (2002) tarafından uygulanmış olan hasar sayılarının modellenmesi yaklaşımı da kullanılmaktadır.

3.1.1 Aşırı-yayılmış Poisson Modeli

Pozitif tamsayıların modellenmesinde (Hasar sayıları) kullanılan Poisson dağılımının yerine, yarı-olabilirlik tahmin modeli kullanılarak oluşturulan aşırı-yayılmış Poisson dağılımı, sürekli verilerin (hasar tutarları) modellenmesinde de kullanılabilmektedir (England and Verrall, 2002).

Sayım verilerinin modellemesi için Poisson dağılımının kullanımı iyi bilinen bir yöntemdir. Bununla birlikte, uygulamanın ortalama ve varyansın eşitliği konusunda sınırlaması mevcuttur. Varyansın eşit olmadığı, ancak ortalamayla orantılı olduğu aşırı-yayılmış Poisson (ODP) modelini tanımlayarak bu problemin üstesinden gelinebilir. Bu

nedenle Mack'in (1991) kanıtladığı Zincir Merdiven (Chain-Ladder) metoduyla aynı tahminleri sağlar. Poisson dağılımının olasılık fonksiyonu

$$f(y) = \frac{\mu^y}{y!} e^{-\mu}, y = 0,1,2, \dots, \quad \mu > 0.$$

burada Genelleştirilmiş Doğrusal Modelleme'nin aşırı dağılmış uzantısını incelemekteyiz. Artan gelişme üçgeninin modeli aşağıdaki varsayımlar ile tanımlanmaktadır.

Model varsayımları 3.1.

Artan hasarların sayısı ($X_{i,j}$) bağımsız aşırı-yayılmış Poisson dağılımına sahip rastlantı değişkenleri olduğu varsayımı altında $0 \leq i, j \leq I$ olmak üzere ortalama ve varyans sırasıyla,

$$E[X_{i,j}] = \mu_{i,j} \quad ve \quad var[X_{i,j}] = \varphi \mu_{i,j} \tag{3.1}$$

şeklindedir.

Doğrusal tahminleyici $\eta_{i,j}$, ortalama $\mu_{i,j}$ ile logaritmik bağlantı fonksiyonu tarafından ilişkilidir.

$$\eta_{i,j} = log(\mu_{i,j}) = c + \alpha_i + \gamma_j, \tag{3.2}$$

$\alpha_0 = \gamma_0 = 0$ durumlarında geçerlidir.

Parametreler $\alpha_i \; ve \; \gamma_j$ i inci olay yılının ve j inci gelişim yılının artan hasar sayılarının beklenen değeri (ve varyansı) üzerindeki etkilerini gösterir. $\alpha_0 \; ve \; \gamma_0$ için köşe kısıtlamaları dikkate alınmalıdır, aksi takdirde model aşırı parametrelendirilir ve benzersiz parametre tahmincileri hesaplanamaz.

Logaritmik bağlantı fonksiyonu, Poisson dağılımı için kanonik bir bağdır ve ortalama kullanımı çarpımsal bir şekle sahip olduğu için pratikte kullanılmaktadır. Tahmin amacı için doğrusal bir yapı ile çalışmak daha elverişlidir.

Denklem (3.2) göz önünde bulundurulduğunda, doğrusal öngörücü $\eta_{i,j}, \forall_{i,j}$, formüle bir tasarım matrisi ve parametre vektörü terimi ile yazılabilir

$$\eta_{i,j} = \boldsymbol{\Gamma}_{i,j}^{T}\boldsymbol{\beta} ,$$

ve

$$\boldsymbol{\Gamma}_{i,j} = \left(1,\ \delta_{1,i}, \dots,\ \delta_{I,i},\ \delta_{1,j}, \dots,\ \delta_{I,j}\right)^{T}, \qquad (3.3)$$

$\delta_{i,j}$ Kronecker'in deltasını ifade etmektedir. Böylece parametre vektörü β aşağıdaki şekilde ifade edilir.

$$\beta = (c, \alpha_1, \dots., \alpha_I,\ \beta_1, \dots.,\ \beta_I)^{T} . \qquad (3.4)$$

Standart matematiksel yazılım paketlerinin kullanımı dahilinde, β'nın tahmini kolayca elde edilir ve hasar sayılarının $X_{i,j}, \forall_{i,j}$, ortalamalarının $\hat{\mu}_{i,j}$ tahminlerini belirlemek için kullanılır.

Tahmini değerler daha sonra dağılım parametresinin φ tahmincisinin hesaplanması için kullanılır.

Pearson X^2 istatistiği ya da bir model sapması, gözlem sayıları ile modelin bir dizi parametresi arasındaki farkla bölünerek hesaplanabilir. Burada sadece β vektörünün elemanlarının sayısını göz önüne almaktayız. Pearson X^2 istatistiğinin kullanılması ile,

$$\hat{\varphi} = \frac{1}{n-p}\sum_{i=0}^{I}\sum_{j=0}^{I-i}\frac{\left(X_{i,j} - \hat{\mu}_{i,j}\right)^2}{V(\hat{\mu}_{i,j})} , \qquad (3.5)$$

şeklinde elde edilir. Aşırı-yayılmış Poisson modelinde $V(\hat{\mu}_{i,j}) = \hat{\mu}_{i,j}$'dir. $(I+1)(I+2)/2$ gözlemlerin sayısı ve $2I+1$ parametrelerin sayısı olmak üzere payda $n-p$, $I(I-1)/2$'ye eşittir.

Model kurulduktan sonra, uyum iyiliğinin ölçülmesi için model sapması kullanılabilir. Aşırı-yayılmış Poisson dağılımında sapma,

$$D = \frac{2}{\hat{\varphi}} \sum_{i=0}^{I} \sum_{j=0}^{I-i} \left(X_{i,j} log \left(\frac{X_{i,j}}{\hat{\mu}_{i,j}} \right) - \left(X_{i,j} - \hat{\mu}_{i,j} \right) \right)$$

şeklinde tanımlanır.

Sapmanın asimptotik olarak $X^2_{I(I-1)/2}$ dağılımına uyar ve doymuş model için 0'a eşittir. Verilen dağılımın üst kuyruk sapması değerleri yetersiz bir uyum göstermektedir.

Alternatif olarak, yapıların ortalama ve varyansları (3.1) ile modellenen hasar sayıları, yarı-olabilirlik açısından tanımlanabilir.

$$Q(\mu; \boldsymbol{X}) = \frac{1}{\hat{\varphi}} \sum_{i=0}^{I} \sum_{j=0}^{I-i} \left(X_{i,j} log\left(\hat{\mu}_{i,j} \right) - \hat{\mu}_{i,j} - \kappa \right),$$

Burada κ, β'dan bağımsız bir sabittir. B'nın tahmini dağılım parametresinden φ etkilenmez, bu durum yarı-olabilirlik yöntemi tahminlerinin kullanıldığı ilişki (3.5) aracılığıyla elde edilir.

Eğer dağılım parametresi $\varphi = 1$ ise, De Jong ve Heller (2008)'de açıkladığı gibi, yarı-olablirlik tahmincileri Poisson regresyonu için maksimum olabilirlik yöntemi ile elde edilen tahmincilere eşittir.

3.1.2 Aşırı-yayılmış Negatif Binom Modeli

Negatif binom dağılımı, μ parametresinin gama dağıtılacağı şekilde Poisson dağılımından elde edilebilir.

$NB(\mu, r)$'nın olasılık fonksiyonu,

$$f(y) = \frac{\Gamma(y+r)}{y!\,\Gamma(r)} \left(\frac{r}{r+\mu} \right)^r \left(\frac{\mu}{r+\mu} \right)^y \quad , y = 0, 1, 2, \ldots, r, \quad \mu > 0.$$

Bu dağılım, sadece r hatalarının sayısı sabitlendiğinde üstel aile tipindedir.

Modelleme için negatif binom dağılımının kanonik bağlantısı olmasa da logaritmik bağlantı fonksiyonunu kullanılmaktadır.

Modelleme için negatif binom dağılımının kanonik bağlantısı olmasa da buradaki logaritmik bağlantı işlevini kullanırız. Bununla birlikte, kullanımı sırasında tahminlerin yorumlanması daha uygundur. $NB(\mu, r)$ için kanonik bağlantı $g(\mu) = log[\mu/(\mu + r)]$ şeklindedir.

Model varsayımları 3.2.

Kümülatif hasar sayıları $C_{i,j}$ $(\forall_i, \forall_j \geq 1)$, ortalama ve varyansı aşağıda verilen aşırı-yayılmış negatif binom dağılımlı rastgele değişkenlerdir ve koşullu olarak bağımsızdır.

$$E(C_{i,j}|C_{i,j-1}) = \mu_{i,j} = \lambda_j C_{i,j-1} \quad ve \quad var(C_{i,j}|C_{i,j-1}) = \varphi\left(\mu_{i,j} + \frac{\mu_{i,j}^2}{r}\right). \tag{3.6}$$

Doğrusal tahminleyici $\eta_{i,j}$ $(\forall_i, \forall_j \geq 1)$, logaritmik bağlantı fonksiyonu vasıtasıyla ortalama $\mu_{i,j}$ ile ilişkildir.

$$\eta_{i,j} = \log(\mu_{i,j}) = c + \gamma_{j-1} + \log(C_{i,j-1}), \tag{3.7}$$

$\gamma_0 = 0$ olduğu durumlar için geçerlidir.

Aşırı-yayılmış Negatif Binom modeli parametreleri yanlızca gelişme yıllarının sayısı ile ilişkili iken, aşırı-yayılmış Poisson modeli parametreleri hem oluşum hem gelişim yıllarının etkileri ile ilişkilidir.

Burada doğrusal tahminleyici $\eta_{i,j}$ hesaplamak için (3.7)

$$\eta_{i,j} = \boldsymbol{\Gamma}_j^T \beta + \log(C_{i,j-1}), \qquad j \geq 1,$$

ve

$$\boldsymbol{\Gamma}_j = \left(1, \delta_{1,j-1}, \ldots, \delta_{I-1,j-1}\right)^T \tag{3.8}$$

$$\beta = (c, \gamma_1, \dots, \gamma_{I-1})^T \tag{3.9}$$

şeklinde ifade edilebilir.

Gelişim üçgenindeki gözlemlerin logaritmik değerleri, tekrarlanan modelde bir dengeleme olarak kullanılır.

Bazen model tanımının $f_{i,j} = C_{i,j}/C_{i,j-1}$ gelişim faktörleri ve ilgili ağırlıkları $w_{i,j} = C_{i,j-1}$ kullanılması daha uygundur. Bu şekilde ortalama ve varyans,

$$E(f_{i,j}|C_{i,j-1}) = \lambda_j \quad ve \quad var(f_{i,j}|C_{i,j-1}) = \varphi\left(\frac{\lambda_j}{w_{i,j}} + \frac{\lambda_j^2}{r}\right),$$

$\forall_i, \forall_j \geq 1$ için geçerlidir.

Parametre vektörü β'nin tahmini elde edildikten sonra tekrar, dağılım parametresinin $\hat{\varphi}$ tahmincisi aşırı-yayılmış Poisson modeli durumunda olduğu gibi Pearson X^2 istatistiği ile hesaplanır. Bu hesaplamada denklemdeki (3.5) payda $n - p$, $I(I-1)/2$ ve $V(\hat{\mu}_{i,j}) = \hat{\mu}_{i,j} + \hat{\mu}_{i,j}^2/r$ şeklindedir.

Model uygunluğunu incelemek için, sapma

$$D = \frac{2}{\hat{\varphi}} \sum_{i=0}^{I} \sum_{j=0}^{I-i} \left(X_{i,j} log\left(\frac{X_{i,j}}{\hat{\mu}_{i,j}}\right) - (X_{i,j} - r) log\left(\frac{X_{i,j} + r}{\hat{\mu}_{i,j} + r}\right)\right)$$

şeklindedir. Aşırı-yayılmış Poisson modelinde olduğu gibi, sapma asimptotik $X^2_{I(I-1)/2}$ olarak dağıtılmış ve uygun olmayan modellerde büyük değerler elde edilmiştir.

(3.6) denklemleri ile tanımlanan ortalama ve varyansa sahip bağımlı değişkenleri için yarı-olabilirlik fonksiyonu,

$$Q(\mu; \boldsymbol{X}) = \frac{1}{\hat{\varphi}} \sum_{i=0}^{I} \sum_{j=0}^{I-i} \left(X_{i,j} log\left(\frac{\hat{\mu}_{i,j}}{r + \mu_{i,j}}\right) r log\left(\frac{1}{r + \mu_{i,j}}\right) + \kappa \right)$$

şeklindedir.

Yani ϰ sabiti β'den bağımsız ve φ dağılım parametresiyle birlikte, $\hat{\beta}$'ye ait yarı-olabilirlik tahmincisini etkilememektedir. Tahmin edici $\hat{\varphi}$, aşırı-yayılmış negatif binom dağıtılan bağımlı değişkenler için mümkün olan en çok olabilirlikte olduğu gibi benzer şekilde elde edilir.

Data üzerinde yarı olabilirlik yöntemi ile çalışıldığında, dağılımsal varsayım olmadığı için kesikli ya da sürekli olmasının bir önem yoktur.

3.2 Hasar Tutarlarının Modellenmesi

Bu bölümde, hasar tutarlarından oluşan gelişim üçgenlerini ele almaktayız. $X_{i,j}$ artan hasar datasını ve $C_{i,j}$ kümülatif hasar datasını ifade etmektedir.

Daha önce belirtildiği gibi, aşırı-yayılmış Poisson ve aşırı-yayılmış Negatif Binom modelleri devam eden datada da kullanılabilir. Bununla birlikte, hasar tutarlarını modellerken negatif değerler gözlemlemek mümkündür ve bu bakımdan aşırı-yayılmış Poisson ve aşırı-yayılmış Negatif binom modelleri sınırlı kalmaktadır. Bu sebeple daha önce belirtildiği gibi yarı-olabilirlik yöntemi kullanılabilir.

Hayat-dışı sigortalarda sürekli veri modellemesi için genellikle kullanılan dağılımlar arasında gamma ve Ters Gauss dağılımları kullanılır. Normal dağılımın simetrik olma kısıtlaması vardır, ancak aşırı-yayılmış Negatif Binom modeline yaklaştırmak ve modeli Mack'in (1993) genelleştirilmiş doğrusal modeli olarak tanımlamak için kullanılabilir.

3.2.1 Gamma Modeli

Gamma dağılımının $\Gamma(a,p)$ olasılık yoğunluk fonksiyonu

$$f(y) = \frac{a^p}{\Gamma(p)} y^{p-1} e^{-ay} \text{ , y} > 0, a, \text{p} > 0,$$

$\text{p} = 1$ durumunda üstel dağılıma eşittir.

Gamma dağılımlı hasar tutarları ve logaritmik bağlantı fonksiyonu olan Genelleştirilmiş Doğrusal Modeli aşağıdaki varsayımlar ile verilmektedir.

Model varsayımları 3.3.

$0 \leq i, j \leq I$ için artan hasar tutarları $X_{i,j}$ bağımsız gamma dağılımının rastgele değişkenleridir.

$$E[X_{i,j}] = \mu_{i,j} \quad ve \quad var[X_{i,j}] = \varphi \mu_{i,j}^2. \tag{3.10}$$

Doğrusal tahminleyici $\eta_{i,j}$ ve ortalama $\mu_{i,j}$ nin ilişkili olduğu logaritmik bağlantı fonksiyonu

$$\eta_{i,j} = \log(\mu_{i,j}) = c + a_i + \gamma_j, \tag{3.11}$$

$a_0 = \gamma_0 = 0$ olduğu durumlar için geçerlidir.

Burada yayılma parametresi dağılımın bir parametresidir ve $\varphi = 1/p$ şeklinde tanımlanmaktadır.

Logaritmik bağlantı fonksiyonu, gamma dağılımı için kanonik bir bağ değildir, ancak ortalamanın çarpımsal yapısı nedeniyle kullanılabilir.

Ortalama değerin karşılıklı işlevi olan gamma dağılımlı hasar tutarları için kanonik bağ yaygın olarak kullanılmaz. Burada, her iki modelin verdiği uygunluk ve tahminleri, logaritmik ve karşılıklı bağlantılar ile karşılaştırmak için göstermekteyiz.

Model varsayımları 3.4.

$0 \leq i, j \leq I$ için artan hasar tutarları $X_{i,j}$ bağımsız gamma dağılımının rastgele değişkenleri ve ortalama ile varyansı

$$E[X_{i,j}] = \mu_{i,j} \quad ve \quad var[X_{i,j}] = \varphi \mu_{i,j}^2$$

şeklindedir.

Doğrusal tahminleyici $\eta_{i,j}$'nin ortalama $\mu_{i,j}$ ile ilişkili olduğu karşılıklı bağlantı fonksiyonu

$$\eta_{i,j} = \frac{1}{\mu_{i,j}} = c + a_i + \gamma_j, \tag{3.12}$$

$a_0 = \gamma_0 = 0$ olduğu durumlar için geçerlidir.

Doğrusal tahminleyici $\eta_{i,j}$ ifadesi, tasarım matrisi ve parametre vektörü, $\eta_{i,j} = \boldsymbol{\Gamma}_{i,j}^T \boldsymbol{\beta}$ için model varsayımları 3.3 ve 3.4 ile aynıdır. $\boldsymbol{\Gamma}_{i,j}^T \; ve \; \beta$ değişkenleri sırasıyla model (3.3) ve model (3.4) ile aynıdır.

Model uygunluğunu incelemek için, sapma

$$D = \frac{2}{\hat{\varphi}} \sum_{i=0}^{I} \sum_{j=0}^{I-i} \left(\frac{X_{i,j} - \hat{\mu}_{i,j}}{\hat{\mu}_{i,j}} - log\left(\frac{X_{i,j}}{\hat{\mu}_{i,j}} \right) \right)$$

ve yakınsak dağılımı $X^2_{I(I-1)/2}$ şeklindedir. Dolayısıyla, dağılımın üst kuyruğundaki değerler, has olmayan modeli işaret etmektedir.

Gelişim üçgeninde negatif değerlerin bulunduğu durumlarda, (3.10)'da verilen ortalama ve varyans arasındaki ilişkiyi, yarı-olabilirlik fonksiyonunu tanımlamak için kullanırız. Bu durumda,

$$Q(\mu; \boldsymbol{X}) = \frac{1}{\varphi} \sum_{i=0}^{I} \sum_{j=0}^{I-i} \left(-\frac{X_{i,j}}{\mu_{i,j}} - \log(\mu_{i,j}) + \kappa \right)$$

burada κ sabiti β'dan bağımsızdır. Elde edilen β'nın yarı-olabilirlik tahmincisi, φ dağılımının parametresinin Pearson X^2 istatistiğine ve $I(I-1)/2$ sayısına oranlanması ile bulunmaktadır.

3.2.2 Ters Gauss Modeli

Ters Gauss ve Gamma dağılımına uyan bağımlı değişkenlere sahip modellerde varyans, ortalamanın güç fonksiyonu ile orantılıdır.

Ters Gauss dağılımı $IG(\mu,\lambda)$, data da büyük çarpıklıklar gözlendiğinde yaygın olarak kullanılmaktadır. Ters Gauss dağılımının olasılık yoğunluk fonksiyonu,

$$f(y)=\sqrt{\frac{\lambda}{2\pi y^3}}exp\left\{-\frac{\lambda(y-\mu)^2}{2\mu^2 y}\right\},y>0,\lambda,\mu>0,$$

şeklindedir.

Model varsayımları 3.5.

$0\le i,j\le I$ için artan hasar tutarları $X_{i,j}$ bağımsız ters gauss dağılımının rastgele değişkenleridir.

$$E\left[X_{i,j}\right]=\mu_{i,j}\quad ve\quad var\left[X_{i,j}\right]=\varphi\mu_{i,j}^3. \tag{3.13}$$

Doğrusal tahminleyici $\eta_{i,j}$ ve ortalama $\mu_{i,j}$ nin ilişkili olduğu logaritmik bağlantı fonksiyonu

$$\eta_{i,j}=\log\left(\mu_{i,j}\right)=c+a_i+\gamma_j, \tag{3.14}$$

$a_0=\gamma_0=0$ olduğu durumlar için geçerlidir.

Gamma modelinde olduğu gibi, logaritmik bağlantı, ters gauss dağılan bağımlı değişkenler için kanonik (standart) bir bağ fonksiyonu değildir. Ancak özelliklerinden dolayı pratikte kullanılmaktadır.

Bu durumda kanonik (standart) bağ, ortalamanın karşılıklı kare fonksiyonudur.

Model varsayımları 3.6.

$0\le i,j\le I$ için artan hasar tutarları $X_{i,j}$ bağımsız ters gauss dağılımının rastgele değişkenleridir.

$$E\left[X_{i,j}\right]=\mu_{i,j}\quad ve\quad var\left[X_{i,j}\right]=\varphi\mu_{i,j}^3. \tag{3.15}$$

Doğrusal tahminleyici $\eta_{i,j}$ ve ortalama $\mu_{i,j}$ nin ilişkili olduğu logaritmik bağlantı fonksiyonu

$$\eta_{i,j} = \frac{1}{\mu_{i,j}^2} = c + a_i + \gamma_j, \tag{3.16}$$

$a_0 = \gamma_0 = 0$ olduğu durumlar için geçerlidir.

Her iki model varsayımında da doğrusal tahminleyici $\eta_{i,j} = \boldsymbol{\Gamma}_{i,j}^T \boldsymbol{\beta}$ biçiminde yeniden yazılabilir, burada $\boldsymbol{\Gamma}_{i,j}$ ve $\boldsymbol{\beta}$ sırasıyla (3.3) ve (3.4)'te verilmiştir.

Verilen modellerin parametre tahmincileri, daha önce belirtildiği gibi maksimum olabilirlik yöntemi ile elde edilmektedir.

Model uygunluğunu incelemek için, sapma

$$D = \frac{1}{\hat{\varphi}} \sum_{i=0}^{I} \sum_{j=0}^{I-i} \left(\frac{\left(X_{i,j} - \hat{\mu}_{i,j}\right)^2}{\hat{\mu}_{i,j}^2 X_{i,j}} \right),$$

ve yakınsak dağılımı $X^2_{I(I-1)/2}$ şeklindedir. Modelin belirli bir veri seti için uygun olmadığı durumlarda, bu dağılımın üst kuyruğunda D değerleri ortaya çıkmaktadır.

Gelişim üçgeninde negatif değerler varsa, yarı-olabilirlik yöntemi, aralarındaki bağı (3.13) tarafından verilen ortalama ve varyans yapısı ile, modeli dataya uydurmak için kullanılabilir. Burada yarı-olabilirlik fonksiyonu

$$Q(\beta; \boldsymbol{X}) = \frac{1}{\varphi} \sum_{i=0}^{I} \sum_{j=0}^{I-i} \left(-\frac{X_{i,j}}{2\mu_{i,j}^2} + \frac{1}{\mu_{i,j}} + \varkappa \right)$$

şeklindedir.

Burada $\varkappa$ sabiti β'dan bağımsızdır. Elde edilen β'nın yarı-olabilirlik tahmincisi, φ dağılımının parametresinin Pearson X^2 istatistiğine ve $I(I-1)/2$ sayısına oranlanması ile bulunmaktadır. Burada $V\left(\hat{\mu}_{i,j}\right) = \hat{\mu}_{i,j}^3$'dir.

3.2.3 Normal Dağılım Modeli

Normal dağılımın simetrisi konusunda bir sınırlaması mevcuttur. Buna rağmen, Aşırı-yayılmış Negatif Binom modelinin bir yaklaşımı olarak, Mack'in Mack Modelinde (1993) kullanılabilir.

İlk olarak ortalaması μ ve varyansı σ^2 olan normal dağılmış rastgele bir değişkenin olasılık yoğunluk fonksiyonu,

$$f(y) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{(y-\mu)^2}{2\sigma^2}\right\}, y \in \mathbb{R},$$

şeklindedir.

Normal dağılıma sahip bağımlı değişkenler için düşünülen modeller aşağıdaki varsayımlar ile gösterilmektedir.

Model varsayımları 3.7.

$0 \leq i, j \leq I$ için artan hasar tutarları $X_{i,j}$ bağımsız normal dağılımın rastgele değişkenleridir.

$$E[X_{i,j}] = \mu_{i,j} \quad ve \quad var[X_{i,j}] = \varphi. \tag{3.16}$$

Doğrusal tahminleyici $\eta_{i,j}$ ve ortalama $\mu_{i,j}$ nin ilişkili olduğu logaritmik bağlantı fonksiyonu

$$\eta_{i,j} = \mu_{i,j} = c + a_i + \gamma_j, \tag{3.17}$$

$a_0 = \gamma_0 = 0$ olduğu durumlar için geçerlidir.

Görüldüğü gibi, Model varsayımları 3.7 ile tanımlanan model, normal dağılım için kanonik (standart) bir bağ fonksiyonudur ve normal dağılmış hata terimlerinin klasik bir doğrusal modelidir.

Logaritmik bağlantı fonksiyonlu model ise aşağıdaki gibidir.

Model varsayımları 3.8.

$0 \leq i, j \leq I$ için artan hasar tutarları $X_{i,j}$ bağımsız normal dağılımın rastgele değişkenleridir.

$$E[X_{i,j}] = \mu_{i,j} \quad ve \quad var[X_{i,j}] = \varphi.$$

Doğrusal tahminleyici $\eta_{i,j}$ ve ortalama $\mu_{i,j}$ nin ilişkili olduğu logaritmik bağlantı fonksiyonu

$$\eta_{i,j} = \log(\mu_{i,j}) = c + a_i + \gamma_j, \tag{3.18}$$

$a_0 = \gamma_0 = 0$ olduğu durumlar için geçerlidir.

Burada hata terimleri, normal dağılmış bağımsız rastgele değişkenlerdir ve tepki değişkenlerinin logaritmik dönüşümü ile modellerdeki hatanın aksine log-normal dağılımına uymaktadır.

Yukarıdaki modellerde, dağılım parametresi φ, σ^2'ye eşittir. Varyans fonksiyonu, ortalamadan bağımsızdır ve 1'e eşittir.

Doğrusal tahminleyici $\eta_{i,j} = \boldsymbol{\Gamma}_{i,j}^T \boldsymbol{\beta}$ biçiminde yeniden yazılabilir, burada $\boldsymbol{\Gamma}_{i,j}$ ve $\boldsymbol{\beta}$ sırasıyla (3.3) ve (3.4)'te verilmiştir.

Model uygunluğunu incelemek için, sapma

$$D = \frac{1}{\hat{\varphi}} \sum_{i=0}^{I} \sum_{j=0}^{I-i} \left(X_{i,j} - \hat{\mu}_{i,j}\right)^2,$$

ve yakınsak dağılımı $X^2_{I(I-1)/2}$ şeklindedir. Modelin belirli bir veri seti için uygun olduğu durumlarda, D'nin küçük değerleri ortaya çıkmaktadır.

Gözlemlenen veriler, ortalama ve varyansı (3.16)'da gösterildiği şekildeyse, normal dağılım varsayımını rahatlatabilir ve model parametrelerinin tahminlerini elde etmek için yarı-olabilirlik fonksiyonunu kullanabiliriz. Burada yarı-olabilirlik fonksiyonu

$$Q(\beta; \boldsymbol{X}) = -\frac{1}{2\varphi}\sum_{i=0}^{I}\sum_{j=0}^{I-i}\left(X_{i,j} - \mu_{i,j}\right)^2$$

şeklindedir.

Dağılım parametresi φ normal doğrusal modeller için olağan varyans tahminini veren Pearson X^2 istatistiğine ve $I(I-1)/2$ sayısına oranlanması ile bulunmaktadır.

Aşırı-yayılmış Negatif Binom Modelinin Yakınlaştırması

Normal dağılan bağımlı değişkenlere sahip modellerden biri, aşırı-yayılmış negatif binom (ONB) modelinin normal yaklaşımıdır. Mack'te (1993) açıklanan model ile benzer şekilde hasar tutarlarının tahminlerini veren bu modelde faydalıdır.

Model varsayımı 3.2'de gösterildiği gibi, aşırı-yayılmış Negatif Binom (ONB) modeli özyinelemelidir ve normal yaklaşım ile benzer biçimdedir. Burada England and Verrall (2002)'de sunulan model çözümlemesini benimsemekteyiz. Kümülatif hasar tutarları $C_{i,j},\ \forall_i, \forall_j \geq 1$ için koşullu bağımsız normal dağılımın rastgele değişkenleri olarak kabul edilir ve

$$E\left(C_{i,j}|C_{i,j-1}\right) = \lambda_j C_{i,j-1} \quad ve \quad var\left(C_{i,j}|C_{i,j-1}\right) = \varphi_j C_{i,j-1}.$$

şeklinde ifade edilir.

$C_{i,j-1}$'yi $w_{i,j}$ sembolü ile ve oran $C_{i,j}/w_{i,j}$'yi $f_{i,j}$ ile gösterirsek, model varsayımlarını aşağıdaki biçimde gösterebiliriz.

Model varsayımları 3.9.

Gelişim faktörü $f_{i,j}$, $\forall_i, \forall_j \geq 1$ için, normal dağılıma uyan rastgele değişkenlerdir ve

$$E\left(f_{i,j}|C_{i,j-1}\right) = \mu_{i,j} = \lambda_j \quad ve \quad var\left(f_{i,j}|C_{i,j-1}\right) = \frac{\varphi_j}{w_{i,j}}. \tag{3.19}$$

doğrusal tahminleyici $\eta_{i,j}$ ve ortalama $\mu_{i,j}$ nin ilişkili olduğu özdeşlik bağlantı fonksiyonu

$$\eta_{i,j} = \mu_{i,j} = c + a_i + \gamma_{j-1}, \;\; j \geq 1 \qquad (3.20)$$

$\gamma_0 = 0$ olduğu durumlar için geçerlidir.

Model varsayımı 3.9 ile verilen model ağırlıkları, $W_{i,j} = w_{i,j}/\varphi_j$ olan ağırlıklandırılmış genelleştirilmiş doğrusal bir modeldir. Yine aşırı-yayılmış negatif binom (ONB) modelinde olduğu gibi doğrusal tahminleyici sadece "j" gelişim yılına bağlıdır. Doğrusal tahminleyici $\eta_{i,j} = \boldsymbol{\Gamma}_{i,j}^T \boldsymbol{\beta}$ biçiminde yeniden yazılabilir, burada $\boldsymbol{\Gamma}_{i,j}$ ve $\boldsymbol{\beta}$ sırasıyla (3.8) ve (3.9)'te verilmiştir.

Dağılım parametresi φ_j "j" gelişim yılına bağlıdır. Bu bağımlılık, "Ortak Modelleme" adı verilen model parametrelerini tahmin etmenin özel yöntemini gerektirmektedir. Bu yöntemin uygun tanımlanması için McCullagh and Nelder (1989)'a bakarız. Ayrıca böyle bir durum için yarı-olabilirlik fonksiyonunun genişletilmiş formunu sunmaktadır.

Model uygunluğunu incelemek için, sapma

$$D = \sum_{i=0}^{I-1} \sum_{j=0}^{I-i} \frac{w_{i,j}}{\varphi_j} \left(f_{i,j} - \hat{\lambda}_j\right)^2,$$

ve yakınsak dağılımı $X^2_{I(I-1)/2}$ şeklindedir.

GELİŞİM ÜÇGENİNDE BOOTSTRAP YÖNTEMİ

Bir önceki bölümde, Genelleştirilmiş Doğrusal Modelleme ile hasar rezervleri hesabını incelemiştik. Model parametrelerinin nasıl tahmin edileceğini ve uyumluluğun yeterliliğini test etmeyi açıkladık. Uygun bir model olması durumunda, sigortacının gelecekteki yükümlülüklerini karşılamak amacıyla gerekli rezerv tahminlerini hesaplaması gerekmektedir. Ayrıca ortalama hataları karesinin tahmini (MSEP) yapılabilir ve böylece belirsizlikler ortadan kaldırılabilir.

Ortalama hata karesi tahmini

$$E\left[\left(X_{i,j}-\hat{X}_{i,j}\right)^2\right]=E\left\{\left[\left(X_{i,j}-EX_{i,j}\right)-\left(\hat{X}_{i,j}-EX_{i,j}\right)\right]^2\right\}$$

$$=E\left[\left(X_{i,j}-EX_{i,j}\right)^2\right]+E\left[\left(\hat{X}_{i,j}-E\hat{X}_{i,j}\right)^2\right], \qquad (4.1)$$

şeklinde tanımlanabilir. Görüleceği üzere ortalama hata karesi tahmini (MSEP), sürecin varyansı ve tahminin varyansı olmak üzere iki bileşenden oluşmaktadır.

Bunula birlikte, bur yaklaşımın bir dezavantajı vardır, çünkü (4.1)'in analitik çözümünü elde etmek güç olabilir. Model varsayımları 3.1, 3.2, 3.3 ve 3.9 ile verilen modellerde, ortalama hataların karesi (MSEP) formülleri England and Verrall (2002)'de bulunabilir.

Bununla birlikte, rezerv tahmini probleminde, tahminin tüm dağılımının türetilmesi yararlıdır. Bu türetme, gelişim üçgenlerinde rastgele değişkenlerin toplamının dağılımını belirlememiz gereken verilere sahip olduğu durumlarda karmaşık olabilir.

Bu gibi durumlarda bootstrap metodunu kullanabiliriz. Bu yöntem, sigorta şirketlerinin gelecekteki yükümlülüklerini yerine getirmeme riskini ölçme imkanı tanıyan rezerv tahminlerinin tam tahmini dağılımını sağlar. McNeil, Frey and Embrechts (2005)'te tanımlanan, Risk Değeri (VaR) veya Şartlı Risk Değeri (CVaR) gibi belirli risk ölçümlerini veya kantilleri kullanarak riski ölçebiliriz.

Efron and Tibshirani (1993) tarafından ayrıntıları ile gösterilen bootstrap yöntemi, verilerin yeniden örneklenmesine dayanmaktadır. Rastgele örnekleri $(y_1, \dots, y_n)$ olan bilinmeyen olasılık dağılımı F ve F dağılımının parametre tahmincisi θ, verilen rastgele örnek $\hat{\theta}_n = s(y_1, \dots, y_n)$'nin bir fonksiynu olarak yazılabilir.

Bu yöntemin kullanılması gözlemlerin bağımsız ve benzer şekilde dağıtılmasını gerektirir. Dolayısıyla, Genelleştirilmiş Doğrusal Modellerde verilen gözlemleri tekrar örnekleyemiyoruz, çünkü kovaryantlara bağımlı ortalamalar (ve muhtemelen varyanslar) mevcuttur. Böyle bir durumda, bootstarp yönteminin varsayımlarının tahminen tuttuğu model artıklarını yeniden örnekleme yaygın olarak uygulanır.

England and Verrall (1999)'da gösterildiği gibi, kullanılan artıklar, belirli bir modelin uygunluğu için uygulanan artıklarla tutarlı olmalıdır. Bu sebeple, (2.10)'da tanımlanan Pearson artıklarını ve buradaki ölçeklendirilmiş versiyonunu göz önüne alacağız. Gelişim üçgeni için ölçeklendirilmiş Pearson artıkları

$$r_{i,j}^{(PS)} = \frac{X_{i,j} - \hat{\mu}_{i,j}}{\sqrt{a(\hat{\varphi})V(\hat{\mu}_{i,j})}}, \tag{4.2}$$

$0 \leq i + j \leq I$ için geçerlidir. φ parametresinin tüm gözlemler için sabit veya sadece "j" gelişim dönemine bağlı olduğunu unutmamalıyız. Bu sebeple $\hat{\varphi}$ değişkenini (4.2)'de $\hat{\varphi}_j$ ile değiştirmeliyiz.

Gelişim üçgenindeki bootstrap süreci birkaç adımda gerçekleştirilir:

1. Verilen gözlemlere uyan modeli ve parametre vektörünün tahmincisi ile dağılım parametresi $\beta \; ve \; \varphi$ değişenlerini elde edelim.
2. Uygun modeli elde edelim $\hat{X}_{i,j} = \hat{\mu}_{i,j}, 0 \leq i + j \leq I$ için geçerlidir.
3. $0 \leq i + j \leq I$ için, uygun model için ölçeklendirilmiş Pearson artıklarını (4.2) denklemi ile belirleyebiliriz.

Yukarıdaki üç adımdan sonra bootstrap döngüsü başlar ve bu B kez tekrarlanır. $1 \leq b \leq B$ için o halde,

4. Bootstrap artıklarının üçgenini oluşturma $\left\{ {}_{(b)}r_{i,j}^{(PS)*} \right\}_{0 \le i+j \le I}$;
5. Bootstrap üçgeni hesaplama

$$ {}_{(b)}\hat{X}_{i,j}^{*} = {}_{(b)}r_{i,j}^{(PS)*}\sqrt{a(\hat{\varphi})V(\hat{\mu}_{i,j})} + \hat{\mu}_{i,j}, $$

$0 \le i + j \le I$ için geçerlidir.

6. Kullanılan modeli bootstrap üçgenine uyduralım ve bootstraplı ${}_{(b)}\hat{\beta}^{*}$'yı tahmin edelim.
7. Gelecek üçgeninin bootstrap değerlerini hesaplayalım ${}_{(b)}\tilde{X}_{i,j}^{*} = {}_{(b)}\tilde{\mu}_{i,j}^{*}, I < i + j \le 2I$ için geçerlidir.
8. Tüm (i,j) değerleri, $I < i + j \le 2I$ için, gelecekteki hasar tutarlarını simüle edelim.
9. Bootstrap rezerv modelini hesaplayalım.

$$ {}_{(b)}R_{i}^{*} = \sum_{j=I-i+1}^{I} {}_{(b)}\hat{X}_{i,j}^{*} \quad ve \quad {}_{(b)}\hat{R}^{*} = \sum_{i=1}^{I}\sum_{j=I-i+1}^{I} {}_{(b)}\hat{X}_{i,j}^{*} $$

B bootstrap rezervi ya her "i" olay yılı için ayrı ayrı veya toplam için tüm "i" değerlerine sahip olduğunda, son adım gerçekleştirilir.

10. B boyutu olan bootstrap rezervleri için ampirik dağılım fonksiyonu $\hat{F}_{B}^{R_i*}$ ($\hat{F}_{B}^{R*}$) belirlenir.

Verilen dağılımlar, bir parametre hatasının bootstrap tahmininin yanı sıra süreç varyansını içermektedir. Gelişim üçgenlerinde bootstrap yönteminin bu yaklaşımı England (2002) tarafından tanımlanmştır; burada ortalama hataların karesi (MSEP) için analitik olarak hesaplanan sonuçlarla artıkların $n/(n-p)$ faktörü ile çarpımı tutarlılık amacıyla uygulanmaktadır. Bunun nedeni, kabul edilen modelin serbestlik derecesine göre ayarlanmasıdır.

England (2002) tarafından ortaya atılan yöntem, England and Verrall (1999) tarafından incelenen yöntemin bir uzantısıdır. Böyle bir durumda, rezervlerin ortalama

hataları karesi (MSEP) değerini değerlendirirken, bootstrap sürecindeki adım 8'i atlayabilir ve aşağıda yer alan formülü kullanabiliriz.

$$MSEP^{*}(R) = \hat{\varphi}\hat{R} + \frac{n}{n-p} var^{*}\big(\hat{R}\big). \tag{4.3}$$

(4.3)'ün sağındaki ilk terimi, $\hat{\varphi}\ ve\ \hat{R}$'nin dağılım parametresinin sabit değerleri ve başlangıç modelinin toplam rezervleri için süreç varyansının tahminini ifade etmektedir. İkinci terim ise bootstrap tahmininin varyanısıdır.

$$var^{*}\big(\hat{R}\big) = \frac{1}{B}\sum_{b=1}^{B}\big({}_{(b)}\hat{R}^{*} - \hat{R}\big)^{2}.$$

p, modelin parametre sayısı için ayarlanamaz. Bu nedenle, analitik olarak değerlendirilen ortalama hataların karesi (MSEP) sonuçları ile karşılaştırılabilir olması için $n/(n-p)$ faktörü ile çarpılması gerekmektedir.

Dağılım parametresi tüm gözlemler için sabit olduğunda mümkün olan başka bir basitleştirme yöntemi daha mevcuttur. Daha sonra (2.10)'da gösterilen ölçeklenmemiş Pearson artıkları kullanılabilir, çünkü 5. adımda ölçeklendirme kaybolmaktadır. Ölçek parametresi, modelin ortak değişkenlerine bağlı olduğu zaman bu işlem yapılamamaktadır.

3.3 Aşırı-yayılmış Poisson Modeli

İlk düşünülen model varsayımı 3.1 tarafından belirtilen aşırı-yayılmış Poisson modelidir. Logaritmik fonksiyonun kullanılması dahilinde

$$\hat{\mu}_{i,j} = exp\big\{c + \ \hat{\alpha}_i + \hat{\gamma}_j\big\}, \tag{4.4}$$

modelin fit edilmiş değerleri için geçerlidir. Artan hasar sayıları (tutarları) $X_{i,j}$'nin ortalama ve varyansının (3.1)'de tanımlanması göz önüne alındığında ölçeklendirilmiş Pearson artıklarını hesaplarken

$$r_{i,j}^{(PS)} = \frac{X_{i,j} - \hat{\mu}_{i,j}}{\sqrt{\hat{\varphi}\hat{\mu}_{i,j}}}$$

$0 \leq i + j \leq I$ değerleri için geçerlidir. Görüldüğü gibi dağılım parametresi kovaryantlara bağımlı değildir ve bu nedenle hesaplamadan çıkarılabilir.

3.4 Aşırı-yayılmış Negatif Binom Modeli

Hasar rezervleri için özyenilemeli model, model varsayımları 3.2'de verilen aşırı-yayılmış Negatif Binom modelidir. Bu durumda,

$$\hat{\lambda}_j = exp\{c + \gamma_{j-1}\},$$

λ_j parametre tahmincisi ve $j \geq 1$ için geçerlidir.

$\hat{\mu}_{i,j} = \hat{\lambda}_j C_{i,j-1}$ değerine sahip olan ölçekli Pearson artıkları

$$r_{i,j}^{(PS)} = \frac{(C_{i,j} - \hat{\mu}_{i,j})}{\sqrt{\hat{\varphi}\left(\hat{\mu}_{i,j} + \frac{\hat{\mu}_{i,j}^2}{r}\right)}},$$

$1 \leq i + j \leq I \; ve \; j \geq 1$ değerleri için geçerlidir. Bootstrap kümülatif hasarları,

$$_{(b)}C_{i,j}^{*} = {}_{(b)}r_{i,j}^{(PS)*}\sqrt{\hat{\varphi}\left(\hat{\mu}_{i,j} + \frac{\hat{\mu}_{i,j}^2}{r}\right)} + \hat{\mu}_{i,j},$$

burada $1 \leq b \leq B$, bootstrap döngüsünün sayısıdır.

Gelişim üçgeninde sabit değerler olarak $C_{i,j}, (1 \leq i + j \leq I \; ve \; j \geq 1$ değerleri için) gözlemlerini düzeltmek önemlidir ve bu bootstrap prosedüründe yeniden örneklenemez. Modelin uyumunu sağlamak için aynı varsayımın kullanılması gerekmektedir. Aksi takdirde, tutarsızlıklardan dolayı analitik ve bootstrap sonuçları karşılaştırılabilir olamaz.

Süreç varyansını hesaplamak için, dağılımda yer alan değerler simüle edilir

$$_{(b)}C^*_{i,j}|C'_{i,j-1} \sim ONB\left[\hat{\mu}'_{i,j}, \hat{\varphi}\left(\hat{\mu}'_{i,j} + \frac{\hat{\mu}'^{\,2}_{i,j}}{r}\right)\right],$$

$\hat{\mu}'_{i,j} = \hat{\lambda}_j C'_{i,j-1}$ için geçerlidir ve

$$C'_{i,j} = \begin{cases} C_{i,j}, & j = I - i + 1; \\ {}_{(b)}\hat{C}^*_{i,j}, & I - i + 2 \le j \le I, \end{cases} \tag{4.5}$$

tüm i değerleri için, b bootstrap döngüsünün sayısını belirtmektedir. England and Verrall (2006)'da belirtildiği gibi, ortalama hataların karesinin (MSEP) öngörülen dağılımının ve değerlendirmesinin ilk iki momenti sadece (4.3) ile elde edilebilir ve nihai hasarlar aşağıdaki gibi hesaplanır

$$\hat{C}^*_i = C_{i,I-i}\hat{\lambda}_{I-i+1} \ldots . \hat{\lambda}_I$$

tüm i değerleri için geçerlidir.

3.5 Gamma Modeli

Önceki bölümde, gamma dağılımına uyan artan hasarlar $X_{i,j}$'yi iki model ile belirledik. Modellerden biri, logaritmik bağlantı fonksiyonu (Model varsayımları 3.3) ve diğeri gamma dağılımı için doğal bir bağ olan karşılıklı bağlantı fonksiyonu (Model varsayımları 3.4) olarak tanımlanmıştır.

İlk model için fit edilen değerler, aşırı-yayılmış Poisson modeli için tanımlanan (4.4) maddesi ile hesaplanır. Karşılıklı bağlantı, fit edilen değerleri vermektedir.

$$\hat{\mu}_{i,j} = \frac{1}{c + \hat{\alpha}_i + \hat{\gamma}_j}$$

tüm $i\ ve\ j$ değerleri için geçerlidir.

Gamma modelleri için ölçeklendirilmiş Pearson artıkları

$$r^{(PS)}_{i,j} = \frac{X_{i,j} - \hat{\mu}_{i,j}}{\sqrt{\hat{\varphi}\hat{\mu}^2_{i,j}}}$$

$0 \leq i + j \leq I$ ve her iki modelde (3.10) formülleri ile verilen ortalama ve varyans yapısına sahip olduğunda için geçerlidir. Aynı bağlantılar, bootstrap yöntemi ile gelecekteki üçgenin bootstrap değerlerini simüle etmek için kullanılır.

3.6 Ters Gauss Modeli

Gamma dağılımına benzer şekilde, ters Gauss dağılımına uyan hasar tutarları $X_{i,j}$'yi iki model ile belirleyebiliriz. İlki model varsayımları 3.5'te verilen logaritmik bağlantı fonksiyonu ve diğeri ortalamanın karşılıklı kare fonksiyonu olan ters Gauss dağılımı için kanonik (standart) bağ fonksiyonudur (Model varsayımı 3.6).

İlk modelin fit edilen değerleri, özdeşlik (4.4) ile verilmiştir. İkinci model, fit edilen değerleri

$$\hat{\mu}_{i,j} = \frac{1}{\sqrt{c + \hat{\alpha}_i + \hat{\gamma}_j}}$$

tüm $i \; ve \; j$ değerleri için geçerlidir. Bu tahminler daha sonra ölçekli Pearson artıklarını hesaplamak için kullanılır. Ters Gauss dağılımı ve (3.13) ile belirtilen ortalama ve varyans tarifine göre

$$r_{i,j}^{(PS)} = \frac{X_{i,j} - \hat{\mu}_{i,j}}{\sqrt{\hat{\varphi}\hat{\mu}_{i,j}^3}}$$

$0 \leq i + j \leq I$ değerleri için geçerlidir.

Bootstrap ortalama şartıyla temel alınan modelin dağılımı simüle edebilmek için standart bir yaklaşım olarak kullanılmaktadır.

3.7 Normal Modeli

Model varsayımları 3.7 ve 3.8 ile verilen normal modellerde $\mu_{i,j}$'nin tahmini değerleri hesaplanmaktadır. Logaritmik link fonksiyonlu model için denklem (4.4)'te verilmektedir.

Daha sonra her iki model için artıklar aşağıdaki gibi hesaplanmaktadır.

$$r_{i,j}^{(PS)} = \frac{X_{i,j} - \hat{\mu}_{i,j}}{\sqrt{\hat{\varphi}}}$$

$0 \leq i + j \leq I$ değerleri için geçerlidir ve bunlar sadece dağılım parametresi ile ölçeklenen gözlenen ve tahmini değerler arasındaki farklardır.

Aşırı-yayılmış negatif binom (ONB) modelinin normal yaklaşımı (NONB), yinelemeli olarak tanımlanmıştır ve bu nedenle aşırı-yayılmış negatif binom modeline benzemektedir. (3.19)'da verilen ortalama ve varyans yapısı göz önüne alındığında,

$$r_{i,j}^{(PS)} = \frac{\sqrt{w_{i,j}}\left(f_{i,j} - \hat{\lambda}_j\right)}{\sqrt{\hat{\varphi}_j}}$$

$1 \leq i + j \leq I, j \geq 1$ değerleri için geçerlidir.

Dağılım parametreleri φ_j, başlangıç fit modelinde "ortak-modelleme" yöntemi ile hesaplanmaktadır.

Aldatıcı veri, geliştirme faktörlerinin bootstrap değerlerini içermektedir.

$$_{(b)}\hat{f}_{i,j}^{*} = {}_{(b)}r_{i,j}^{(PS)*}\sqrt{\frac{\hat{\varphi}_j}{w_{i,j}}} + \hat{\lambda}_j$$

$1 \leq b \leq B$ bootstrap döngüsü için geçerlidir.

$w_{i,j}, 1 \leq i + j \leq I, j \geq 1$ değerleri için, modelin ağırlığı olmak üzere aşırı-yayılmış Negatif Binom modelinde olduğu gibi, geliştirme üçgeninde gözlemlenen sabit değerlerdir.

Hasar rezervlerinin tam tahminli bir dağılımını elde etmek için, bootstrap adımlarından adım 8'in simülasyonu

$$_{(b)}C_{i,j}^{*}|C'_{i,j-1} \sim N\left(\hat{\lambda}_j C'_{i,j-1}, \hat{\varphi}_j C'_{i,j-1}\right), \qquad (4.6)$$

$C'_{i,j-1}$ formülü (4.5)'te verilmektedir.

Kümülatif hasarlar normal dağılıma uyduğunda, simülasyon sırasında negatif değerler elde edilebilir. Bu sorunun üstesinden gelebilmek için, England and Verrall (2006) tarafından (4.6) denklemini gamma dağılımı kullanarak çözülmesi önerilmektedir.

4 UYGULAMA

Çalışmanın bu bölümünde, Genelleştirilmiş Doğrusal modellemelerin gerçek dönem verileri kullanılarak, sonuç ve performanslarını karşılaştırılması yapılmıştır.

Stokastik bir veri modellemesinden doğacak rezerv riskini değerlendirmek amacıyla, 10.000 döngü tekrarının yürütüldüğü bootstrap yöntemi yardımıyla nihai rezervlerin tahmini dağılımlarını oluşturmaktayız. Dağılım tahminleri oluşturulduktan sonra, uygulamada kullanılan en yaygın risk ölçütü olarak rezervlerin riske maruz değerleri (VaR) ve koşullu riske maruz değerleri (CVaR) hesaplanmıştır.

Hesaplama, R Çekirdek Ekibi (2014) tarafından geliştirilen R yazılımı ile gerçekleştirilmiştir.

Çalışmada hayat-dışı sigorta şirketinin 10 yıllık (2007-2016) trafik sigortası ödenen hasar verisi yer almaktadır. Veriler, hasar tazminatı talebinde yaygın bir uygulama olan yılsonunda tahsis edilen giderler için düzenlenmiştir, çünkü tazminat bedellerinin hesaplanması, rezerv tutarlarının değerlendirilmesi için gereklidir.

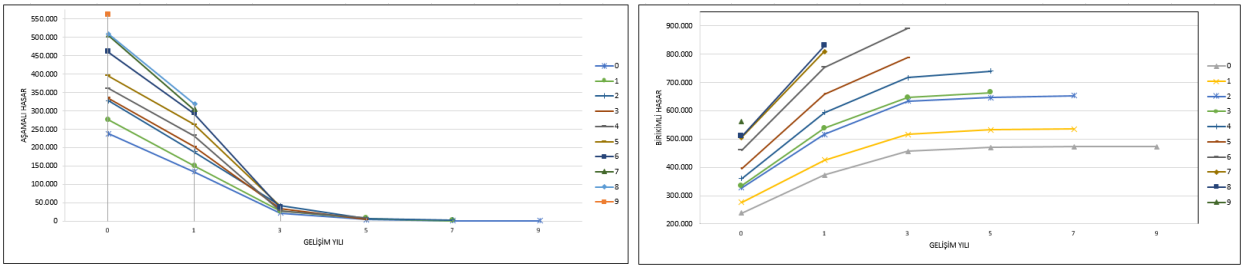
Genelleştirilmiş Doğrusal Modelleme ile rezerv tahminlerinde kurulan modeller, aşamalı hasar ödemelerini baz alan gelişim üçgeni üzerinden hesaplanmaktadır. Buna bağlı olarak, 2007-2016 yılları arasında aşamalı hasar ödemeleri verisi Tablo 5.1'de yer almaktadır.

Tablo 5.1. 2007-2016 Yılları Arasında Trafik Sigortası Aşamalı Hasar Ödemeleri

Hasar Yılı i	Gelişim Yılı j									
	0	1	2	3	4	5	6	7	8	9
0	238.206,30	134.816,40	61.487,90	21.863,50	9.935,30	4.299,70	1.737,40	569,40	43,80	649,70
1	276.100,60	149.942,00	63.772,80	26.732,60	8.935,20	7.190,50	2.525,80	912,50	613,20	**73,00**
2	327.426,90	188.953,20	76.810,60	41.427,50	6.862,00	5.810,80	2.774,00	3.277,70	**3.949,30**	**1.087,70**
3	335.150,30	203.034,90	75.686,40	33.331,80	13.081,60	4.504,10	2.233,80	**2.270,30**	**569,40**	**43,80**
4	360.795,20	232.490,40	96.046,10	28.302,10	14.629,20	8.066,50	**3.080,60**	**4.109,90**	**1.686,30**	**839,50**
5	395.287,70	263.099,30	88.957,80	39.872,60	20.702,80	**8.504,50**	**3.496,70**	**708,10**	**109,50**	**14,60**
6	460.374,50	293.649,80	99.265,40	36.777,40	**20.746,60**	**9.212,60**	**2.438,20**	**1.445,40**	**620,50**	**605,90**
7	506.014,10	303.329,60	105.784,30	**36.565,70**	**17.877,70**	**5.832,70**	**4.307,00**	**2.087,80**	**- 423,40**	**1.284,80**
8	509.547,30	319.499,10	**102.667,20**	**39.193,70**	**12.957,50**	**9.446,20**	**2.095,10**	**2.686,40**	**4.255,90**	**854,10**
9	561.669,30	**329.470,90**	**117.274,50**	**45.106,70**	**22.213,90**	**9.227,20**	**2.168,10**	**1.087,70**	**649,70**	**1.036,60**

Verilerin hem birikimli hem de aşamalı formlardaki grafiksel gösterimi, Şekil 5.1'de yer almaktadır.

Şekil 5.1. Birikimli ve Aşamalı Hasar Tutarlarının Gelişimi



(a) Aşamalı Hasarlar (b) Birikimli Hasarlar

Her iki formdaki gelişimde görebileceğimiz gibi, birikimli hasar ödemeleri değerlerinin kaza yılları ile birlikte arttığı ve ileriki gelişim yıllarında gerçekleşen yükselişlerin başlangıçtaki yükselişten daha az olduğu görülmektedir.

Aşamalı hasar taleplerinin ortaya çıkması, gelişim yıllarının ilerlemesiyle yavaşlar ve hasar tutarları arasındaki farklar ortadan kalkma eğilimindedir.

Üçgen metodu ile alt üçgenden hesapladığımız değerleri dikkate alarak Tablo 5.2'de nihai rezerv sonuçlarını elde etmekteyiz.

Tablo 5.2. Hasar Rezervlerinin Nihai Değerleri

Hasar Yılı i	1	2	3	4	5	6	7	8	9	TOPLAM
Rezerv	73,00	5.037,00	2.883,50	9.716,30	12.833,40	35.069,20	67.532,30	174.156,10	528.235,30	835.536,10

4.1 Modellerin Analizi

Çalışmanın bu bölümünde, aşırı-yayılmış Poisson, aşırı-yayılmış Negatif Binom, aşırı-yayılmış Negatif Binom'un normal yaklaşımı, Gamma ve Normal dağılımlara uyan değişkenlerin logaritmik link fonksiyonları olan modellerini analiz etmekteyiz. Model varsayımları sırasıyla 3.1, 3.2, 3.9, 3.3 ve 3.8 ile verilmektedir.

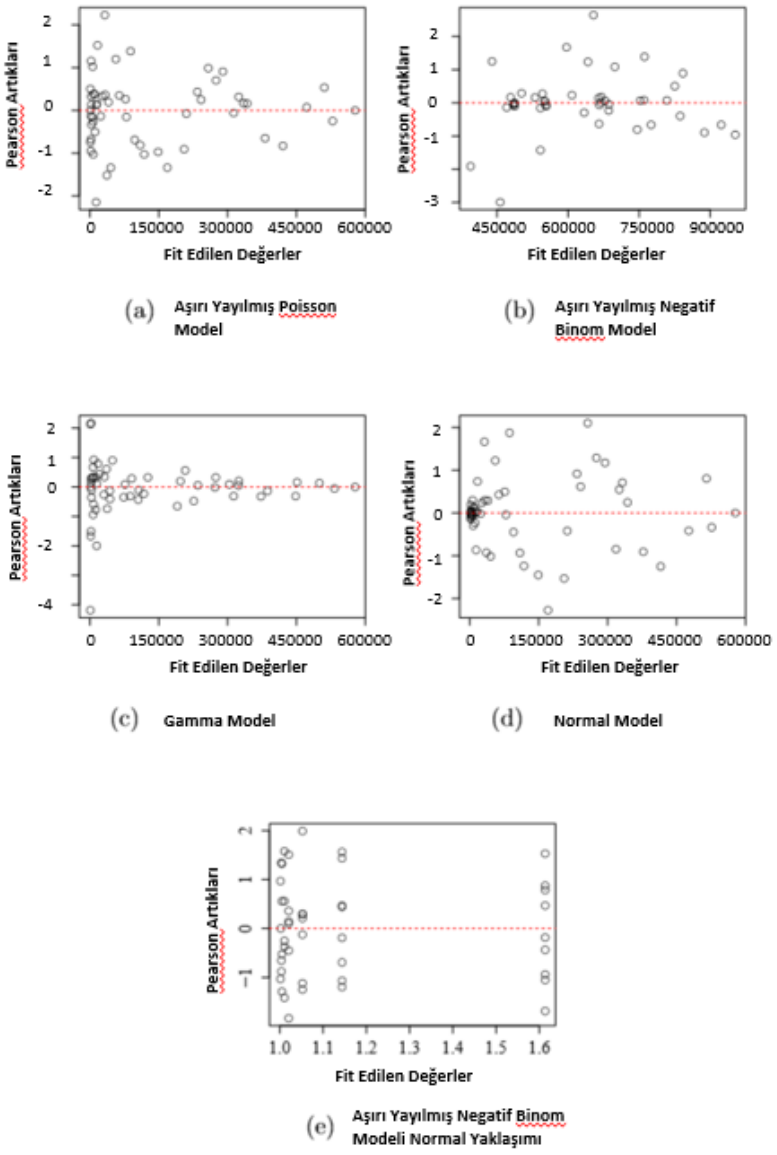
Ayrıca Ters Gauss ile sabit (kanonik) bağları olan gamma ve normal dağılmış bağımlı değişkenler için model içerisinde kısa bir inceleme sunmaktayız. Bu modeller, öncekilere göre daha iyi performans göstermiştir.

Model testi için yapılacak olabilirlik oranı testinde anlamlılık düzeyini α =% 5 kabul etmekteyiz. Bu test, küçük veriler için en uygun olan test türüdür.

İlk model grubu iyi performans gösterir ve benzer sonuçlar veriri. Aşırı-yayılmış Negatif Binom'a yaklaşan model haricinde, söz konusu diğer modellerin logaritmik bağlantı fonksiyonları benzerlik göstermektedir. Logaritmik bağlantı fonksiyonu, basit yorumlanabilmesi nedeniyle hayat-dışı sigortalar rezerv tahmininde yaygın olarak kullanılmaktadır.

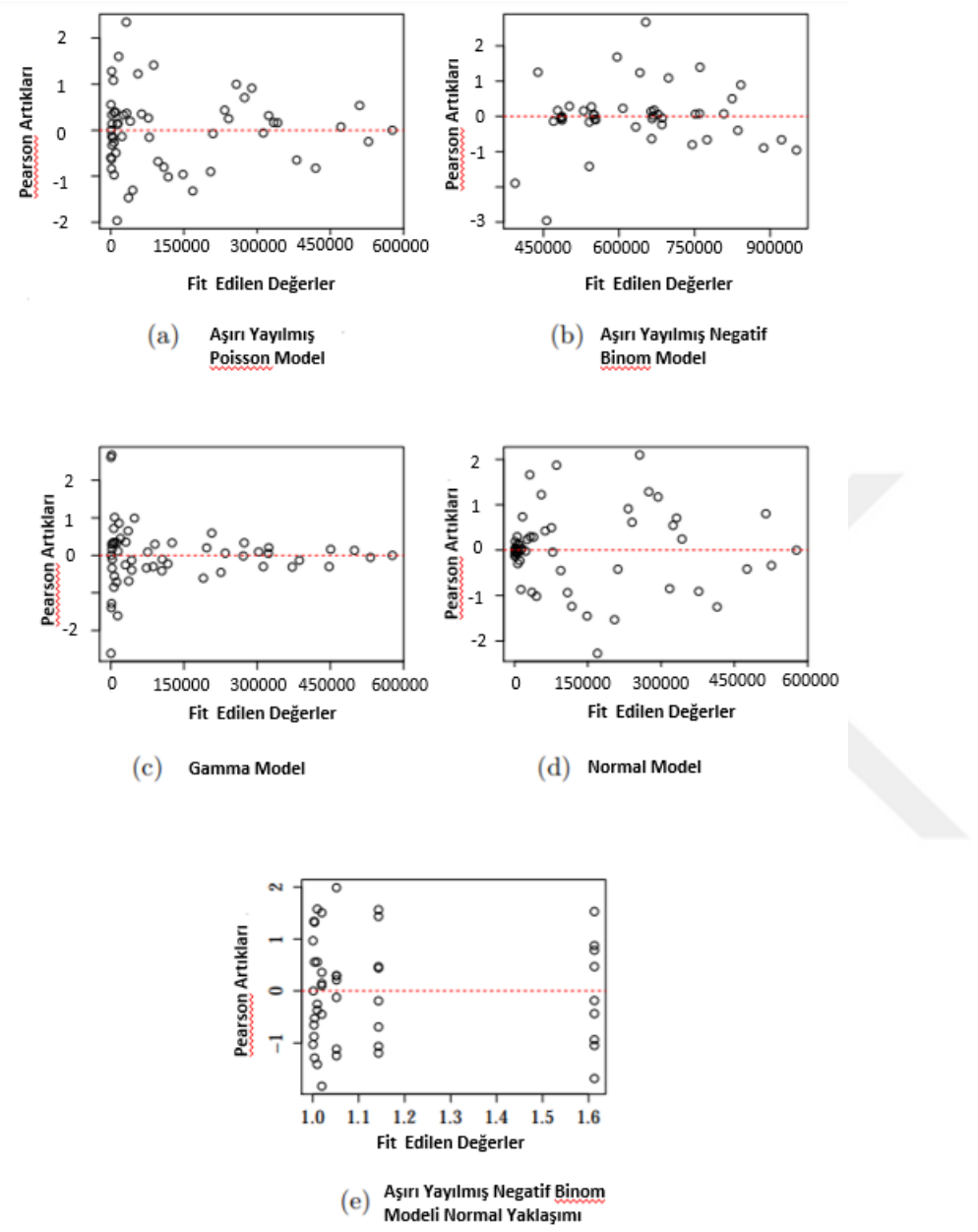
Hesaplamaya konu olan modellerin yeterliliğini incelemek amacıyla, Şekil 5.2 ve Şekil 5.3'te ölçeklendirilmiş sapma ve Pearson artıkları gösterilmiştir.

Şekil 5.2. Fit Edilen Modellerin Ölçeklendirilmiş Çarpıklık Artıkları



Aşırı-yayılmış Negatif Binom modelinde, bağımlı değişkenlerinin kümülatif hasar talepleri olduğu ve Aşırı-yayılmış Negatif Binom modelinin Normal yaklaşımı modeli için ise bunların geliştirme faktörleri olduğu unutulmamalıdır. Diğer durumlarda, aşamalı hasarlar modellenmektedir.

Şekil 5.3. Fit Edilen Modellerin Ölçeklendirilmiş Pearson Artıkları



Şekillerden de anlaşılacağı gibi, artıklar eşvaryanslı değildir ve varyansları artan fit edilmiş değerler ile azalmaktadır. Bu durum, Aşırı-yayılmış Poisson modeli, Gamma modeli ve logaritmik bağlantı fonksiyonuna sahip normal model için daha yüksek sayıda aşamalı hasarlar için daha düşük değerdeki gözlemlerin bir sonucu olabilir. (Tablo 5.1'deki gözlemlenen değerlerin sadece üçte biri 150.000'den büyüktür.) Aşırı-yayılmış Negatif Binom modelinde, kümülatif hasarlar olası değer aralıkları boyunca homojen bir şekilde yayılır, dolayısıyla varyansın sabit olduğu görülmektedir.

Aşırı-yayılmış Negatif Binom modeli Normal yaklaşımı modeli özel bir durumdur, çünkü gelişim faktörleri, gelişim yıllarına bağımlı varyans ile ağırlıklı bir Genelleştirilmiş Doğrusal Modelleme ile modellenmiştir. Bununla birlikte, ölçeklendirilmiş artıklar, tahminlenen gelişim faktörlerinin tüm değerleri için eşvaryanslı görünmektedir.

Modellerin karşılaştırılması ile ilgili olarak, modellerin serbestlik derecelerinin sayısına eşit olan asimptotik X^2 dağılımı elde etmek için ölçeklendirilen sapma D ve Pearson X^2 istatistiklerini değerlendirmektedir.

Bu istatistiklerle birlikte, Tablo 5.3'te modellerin Akaike Bilgi Ölçütü sonuçlarını sunmaktayız.

Tablo 5.3. Fit Edilen Modellerde Pearson X^2 istatistiği, Sapma ve Akaike Ölçütü

Model	Serbestlik Derecesi	X^2	Sapma (D)	Akaike Ölçütü
Aşırı-Yayılmış Poisson	36,00	36,00	36,30	(3.814,30)
Aşırı-Yayılmış Negatif Binom	36,00	36,00	36,06	(769,08)
Gamma	36,00	36,00	44,59	978,71
Normal	36,00	37,95	37,95	921,56
Aşırı-Yayılmış Negatif Binom Normal Yaklaşımı	36,00	44,00	44,00	- 317,72

Ölçeklendirilmiş Pearson X^2 istatistiği ve sapmasının, tüm modellerde benzer yada serbestlik derecesine (Df) yakın bir sonuç olduğunu gözlemlemekteyiz. Bu nedenle modellerin geçerliliğini reddedemeyiz. Bununla birlikte, Aşırı-yayılmış Poisson ve aşırı-yayılmış Negatif Binom dağılımları için, Pearson X^2 istatistiğinin değerinin φ dağılım parametresinin tahminine göre hesaplandığı unutulmamalıdır. Normal dağılım için her iki istatistik de karelerin toplamına eşittir.

Akaike Bilgi Ölçütü'nün değerlerine bakıldığında, Aşırı-yayılmış Negatif Binom normal yaklaşımı modeli en iyi performansa sahiptir. Aşırı-yayılmış Poisson ve Aşırı-yayılmış Negatif Binom modelleri için Akaike Ölçütü değerlerinin başkalarıyla karşılaştırılması uygun değildir. Bu değerler, aşırı-yayılım için ayarlama olmaksızın başlangıçta hesaplanmıştır.

İncelenen modellerin karşılaştırılmasının ardından, tahmini toplam rezervlerin ve bunların bootstrap eşdeğerinin sonuçlarını Tablo 5.4'te sunulmaktadır. Ayrıca veri

setinin tahminsel dağılımlarından, mevcut risk ölçütleri olan riske maruz değerler (VaR) ve koşullu riske maruz değerler (CVaR) hesabı anlamlılık düzeyini α = 0,95 için hesaplanmıştır.

Tablo 5.4. Tahmini Rezerv ve Risk Ölçüleri VaR, CVaR

MODEL	Tahmini Rezerv	Bootstrap Rezerv	$VaR_{0.95}$	$CVaR_{0.95}$
ODP	922.680,58	943.168,03	1.013.948,83	1.035.219,57
ONB	923.458,03	923.426,64	1.196.084,56	1.279.151,99
Gamma	919.218,19	945.538,34	1.356.458,99	1.594.282,77
Normal	932.962,63	933.401,36	1.006.872,21	1.026.823,11
NONB	922.680,58	923.270,42	986.978,98	1.003.190,09

Tablo 5.4'ten tahmin edilen rezervler arasında farklılıklar olduğu ve bunların her birinin Tablo 5.2'de verilen asıl yükümlülüklerden daha yüksek olduğu görülmektedir. Bootstrap rezervleri ile ilgili olarak, modeller için değerleri benzerdir. En yüksek risk dağılımı gamma modelindedir, burada VaR ve CVaR risk ölçümleri en yüksek değerleri almaktadır. Bu durum verilen dağılımın çarpıklığının yüksek oluşu sebebiyle oluşmaktadır.

Analiz edilen modellerin performansı yeterli ve tatmin edicidir. Bu nedenle aktüerlerin, pratik yönden uygulanabilir modeli seçme olasılığı vardır. Seçimi riskten kaçma şeklinde de belirleyebilmekteyiz. Çalışmada model olarak aşırı-yayılmış Negatif Binomun normal yaklaşımı modeli seçilmiştir.

Analizde fit edilen değerler bazı durumlarda gözlenen değerlerin katları olduğundan, modelin anlamsız parametreleri silinmiştir. Bu nedenle Ters Gauss dağılımı ve model varsayımı 3.5 tarafından verilen logaritmik bağlantı fonksiyonu ile model ayrıntılı olarak ele alınmamıştır. Büyük değerlerin sonucu olarak, tüm olay yıllarının en yüksek hasar talebi, Tablo 5.2'de yer alan nihai değerlerin neredeyse 1.5 katı olan 1.227.086,93 TL değerine sahip olduğu tahmin edilmektedir.

Ters Gauss dağılımı için sabit bağlantıya sahip modelde (Model varsayımları 3.6) iyi performans göstermemiştir. Tahmini nihai rezerv sonucu orantısızdır.

Gamma dağılımı durumunda, model varsayımı 3.4 tarafından verilen, doğrusal öngörücü ve ortalama arasındaki ilişki için karşılıklı bağa sahip aşamalı hasar söz konusu olduğunda, parametreler gelişim yılları ile ilişkilenmektedir. Tablo 5.1 veya Şekil 5.1'de, aşamalı hasar değerlerinin ilerleyen yıllar ile birlikte arttığı ve bu parametrelerin verilen veri kümesi için uygun olmadığını görmekteyiz.

Normal dağılıma sahip ve model varsayımı 3.7 tarafından tanımlanan aynı bağlantı fonksiyonuna sahip klasik doğrusal model üzerinden yapılan nihai rezerv tahmini çok yüksek bir değer çıkmaktadır. Bunun nedeni, uygun olmayan benzerlik bağının kullanılmasıdır.

4.2 Aşırı-yayılmış Negatif Binom Normal Yaklaşımı

Burada odaklandığımız karşılaştırılabilir modeller arasında en düşük Akaike Ölçütüne sahip olan Aşırı-yayılmış Negatif Binom Normal yaklaşımı modelinin performansının açıklanmasıdır. Ancak daha önce de belirttiğim üzere, açıklanan modellerden birinin seçimi keyfidir ve aktüer, nihai rezerv sonucunu belirlemede daha elverişli özellikleri nedeniyle başka bir modeli seçebilir.

Aşırı-yayılmış Negatif Binom Normal yaklaşımı ile ilgili modelin kaynak kodu Ek-A'da yer almaktadır.

"Ortak Modelleme" algoritması, gelişim yıllarına göre $W_{i,j} = w_{i,j}$ ağırlıklarına bağlı olarak gelişim faktörleri için ağırlıklı Genelleştirilmiş Doğrusal Modellemeyi başlatmaktadır. İkinci adım Pearson artıklarının hesaplanmasıdır.

$$r_{i,j}^{(P)^2} = w_{i,j}(f_{i,j} - \hat{\lambda}_j)^2,$$

$j \geq 1$ gelişim yıllarına bağlı yayılma parametrelerinin φ_j tahmini için ikinci Genelleştirilmiş Doğrusal Modellemede bir cevap olarak kullanılır.

Değerleri $\varphi_j = \hat{r}_{i,j}^{(P)^2}$ olan $W_{i,j}$ ağırlıkları $W_{i,j} = w_{i,j}/\hat{\varphi}_j$ şeklinde güncellenmelidir. φ_9 değerini tahmin etmek için eksik bilgi (sadece bir artık) olduğunu

ve bu nedenle, England and Verrall (2002) örneğini izleyerek, φ_9 değerinin φ_8 değerine göre ayarlanır.

Son adımda, güncellenmiş ağırlıklar $W_{i,j}$ ile ağırlıklandırılmış Genelleştirilmiş Doğrusal Modelleme, $f_{i,j}$ ve gelişim yılları $j \geq 1$ arasındaki ilişkiyi modellemek için kullanılır. Bu adımlar yakınsamaya kadar tekrarlanır, ancak bizim hesaplamalarımızda bir tekrarlama sonrası sonuç elde edilmiştir. Gelişim faktörleri $\hat{\lambda}_j$'nin yayılma parametrelerinin $\hat{\varphi}_j$ tahminleri ile birleştirilmiş değerleri Tablo 5.5'de görülebilir.

Tablo 5.5. Aşırı-yayılmış Negatif Binom Normal Yaklaşımı Modeli Fit Edilen Değerler

Gelişim Yılı j	$\hat{\lambda}_j$	$\hat{\varphi}_j$
1	1,6128	64,5765
2	1,1430	11,8005
3	1,0516	6,7982
4	1,0197	2,0512
5	1,0099	0,4208
6	1,0040	0,0227
7	1,0029	0,2330
8	1,0007	0,0190
9	1,0014	0,0190

Tablo 5.5'te yer alan değerler, Şekil 5.1'de yer alan gelişim yılları itibari ile büyük eğimin görüldüğü verilerin yorumlanmasına yardımcı olmaktadır. Dağılım parametrelerinin bozulması, daha sonraki gelişim yıllarında kümülatif hasarların yükselişine benzer bir eğilim ile açıklanabilir.

Tahmin edilen değerler ile gözlemlenen değerler karşılaştırılırken, tahminlerin nihai değerlerden daha büyük olduğu görülmektedir. Bu, Aşırı-yayılmış Negatif Binom Normal Yaklaşımı modeli için tahmini toplam IBNR rezervi değerinin, gerçek ödenmemiş borçlardan büyük olmasına karşılık gelmektedir.

Hasar taleplerinin tahmin edilmesi ve rezervlerin sonuçlarının elde edilmesi ile ilgili risk tahmini için 10.000 tekrar döngüsü ile bootstrap yöntemini gerçekleştirdik.

Tablo 5.6. Aşırı-yayılmış Negatif Binom Normal Yaklaşımı Modelinin Öngörülen Aşamalı Hasarları

Hasar Yılı i	1	2	3	4	5	6	7	8	9
0									
1									737,30
2								425,37	898,05
3							1.915,23	435,59	919,51
4						2.964,38	2.134,30	485,38	1.024,63
5					7.975,98	3.266,97	2.352,13	534,87	1.129,24
6				17.553,73	8.960,24	3.670,08	2.642,38	600,94	1.268,59
7			47.254,87	18.979,93	9.688,27	3.968,28	2.857,07	649,77	1.371,67
8		118.587,62	48.933,43	19.654,16	10.032,39	4.109,24	2.958,54	672,84	1.420,36
9	344.164,41	129.571,35	53.465,71	21.474,56	10.961,61	4.489,87	3.232,59	735,11	1.551,98

Hasar yılı ilerledikçe simülasyonların medyanları gözlenen değerlerden önemli ölçüde yüksek çıkmaktadır. Tablo 5.6'da sunulan değerler yüksek tahmin edilen değerlere karşılık gelmektedir.

Bootstrap değerlerinin daha büyük olduğunu ve gerçek değerin nihai rezerv tahmini dağılımının alt kuyruğunda olduğunu görebiliriz. Bu daha önceden tahmin edilen ve bootstrap prosedüründe kullanılan yaklaşıl 0 civarında değişen ilgili artıklar nedeniyle beklenen bir durumdur.

Rezerv hesaplaması yapılırken bir aktüerin karşılaştığı riskin değerlendirilmesine odaklanarak, her bir kaza yılı için rezervlerin ayrı ayrı ve birlikte tahmini için dağılımların riske maruz değerleri ($VaR_{0,95}$) ve koşullu riske maruz değerleri ($CVaR_{0,95}$) hesaplanmıştır. Söz konusu risk ölçümleri ile birlikte, nihai hasarlar ve bunların tahmini ile bootstrap değerleri Tablo 5.7'de verilmiştir.

Tablo 5.7. Aşırı-yayılmış Negatif Binom Normal Yaklaşımı Modelinin Rezerv ve Risk Ölçümleri

Hasar Yılı i	Gerçekte Hesaplanan	Tahmini Rezerv	Bootstrap Rezerv	$VaR_{0.95}$	$CVaR_{0.95}$
1	73,00	737,30	732,26	1.495,62	1.650,97
2	5.037,00	1.323,49	1.315,46	2.445,28	2.718,08
3	2.883,50	3.270,33	3.264,56	5.743,57	6.375,97
4	9.716,30	6.608,69	6.601,61	9.404,08	10.161,53
5	12.833,40	15.259,19	15.267,51	19.573,93	20.661,26
6	35.069,20	34.695,95	34.775,81	43.535,81	45.680,70
7	67.532,30	84.769,86	84.788,77	100.880,31	105.105,55
8	174.156,10	206.368,52	206.615,62	230.709,35	236.970,41
9	528.235,30	569.647,03	569.908,81	616.191,69	627.747,88
TOPLAM	835.536,10	922.680,58	923.270,42	986.978,98	1.003.190,09

Tahmini risk ölçümlerine göre riske maruz değerleri ($VaR_{0,95}$) ve koşullu riske maruz değerleri ($CVaR_{0,95}$), sigorta şirketinin zararının %5 gibi düşük bir olasılıkla 63.708,56 TL'yi aşacaktır. Tahmine dayalı dağılımın üst kuyruğundan alınan değerler dikkate alındığında, zararın ortalaması 79.919,67 TL olacaktır. Bu durum, hesaplanan rezerv miktarı ile bootstrap ile elde edilen tahmin dağılımının ortalamaya eşit olduğu durumlarda geçerlidir.

Aşırı-yayılmış Negatif Binom Normal Yaklaşımı modelinin ayrıntılı analizinde görüldüğü üzere, performans, ihtiyaç duyulandan daha fazla rezerv tahmininin yapıldığı hasar tutarlarının tahmin edilmesi için uygundur. Aktüer için, Mack (1993) modeli benzeri aynı tahminleri sağlayan avantajlı bir modeldir.

SONUÇ

Tez çalışmasının amacı, sigortacılıkta önemli bir yeri olan, hasar rezervlerinin hesaplanmasında Genelleştirilmiş Doğrusal Modellerinin uygulanmasını incelemektir. GDM'lerin kullanımı, modellerin klasik doğrusal regresyonun bir uzantısı olarak sunduğu olanaklarla doğrulanmıştır.

Öncelikle, pratikte yaygın olarak kullanılan hasar rezervleri ve ilgili temel tahmin yöntemleri tanıtılmıştır. Ancak Zincir Merdiven Metod (Chain-Ladder) ve BF yöntemleri sadece ödenmemiş hasarların beklenen değerlerinin tahmini ile ilgilenirken, rezerv riskini hesaplamamaktadır. Bu durum, stokastik modellemenin kullanılmasını önemli kılmaktadır.

Kullanılan stokastik modelleme Genelleştirilmiş Doğrusal Modellemedir. Hesaplama çerçevesinde rezerv hesaplaması sürecini geliştirmek amacıyla, ilişkili parametre tahmin yöntemleri ve model analizi ilkeleri gösterilmiştir. Üstel dağılım ailesinin tanımları da dahil edilmiştir.

Gelişim üçgenlerindeki verilerin dağılımını belirlemek amacıyla bu aileden gelen olasılık dağılımları kullanılmıştır. Hayat-dışı sigortada bağımlı ve bağımsız değişkenler arasındaki doğrusal bağ varsayımı çoğu zaman ihlal edildiğinden, GDM'lerin özellikleri nedeniyle avantajı olmuştur ve bu ilişki birkaç tekdüze gerçek değerli fonksiyonların yardımıyla tanımlanmıştır. Dağılımların bağlantı fonksiyonları ile birleştirilmesiyle, kesikli ve sürekli rastgele değişkenlerin modellenmesi arasında ayrıştırılan gelişim üçgeni için dokuz model önerilmiştir.

Tezin ilgi konusunun rezerv risklerinin tahmin edilmesi olması sebebiyle, hasar rezervlerinin hesaplanmasında tahmine dayalı tüm dağılımların kullanılması sağlanmıştır. Karlılık için, Pearson artıklarının yeniden örneklendiği bir bootstrap yöntemi benimsenmiştir. Tahminleyici dağılımlara sahip olmak için, pratikte en çok kullanılan risk ölçümleri olarak VaR ve $CVaR$ hesaplamada dikkate alınmıştır.

Son olarak, önerilen modelleri analiz etmek ve performanslarını karşılaştırmak için gerçek hayattaki veriler üzerinden uygulama yapılmıştır. Elde edilen değerlerde,

birbirine benzeyen ve gerçek değerlere karşılık gelen beş modelden oluşan grup için elde edilen sonuçların detaylarını açıkladık. Modellerin analizini ve karşılaştırmasını takiben, aşırı-yayılmış Negatif Binom Normal Yaklaşımı modeli tam olarak açıklanmak üzere seçilmiştir. Uygulama sonuçları tablo ve şekiller ile özetlenmiştir.

Analizlerimiz sonucu elde edilen sonuçlar tatmin edicidir. Soruları cevaplanan iddiaların altta yatan teorisini geliştirmek için çeşitli önerilerimiz mevcuttur. Örneğin; GDM'lerin varsayımlarından biri, hayat-dışı sigorta gerçekliği ile bağlantılı olan gözlemlerin bağımsızlığıdır. Bu nedenle, bu tez boyunca kullanılan teorinin uzantısı olarak karşılıklı veri korelasyonunu inceleme ve düşünmeye teşvik etmekteyiz. Hasar tutarlarının bu tür gelişim ve tahminlerini incelemek ve tahmini risk ölçümlerindeki değişimi araştırmak yararlı olacaktır.

EKLER

Kaynak Kodu

Burada verilerin okunması için kaynak kodunu ve "ortak-modelleme" yöntemi ile Aşırı-yayılmış Negatif Binom Normal yaklaşımı modelini hesaplamayı göstermekteyiz. Ayrıca bu tip modelde bootstrap prosedürünün gerçekleştirilmesi için algortiması yer almaktadır.

Ek 1: Verinin aktarımı ve gösterimi

```
----------------------------------------------------------------
## Paket ve Kütüphane
install.packages('ChainLadder')
install.packages('Epi')
library(ChainLadder)
library(stats)
library(statmod)
library(Hmisc)
library(Epi) library(MASS)

## Veri

a=read.csv(choose.files(), header=TRUE)
grp.code=unique(a$GRKOD)

# Grupta verilen verileri alma işlevi
ins.line.data=function(g.kod){
        b=subset(a, a$GRKOD==g.kod)
        name=b$GRAD
        grpcode=b$GRKOD
```

```
        ay=b$HasarYılı
        dev=b$GelişimLag

        cum_pdloss=b[,7]
        data.out=data.frame(name,grpcode,ay,dev,cum_pdloss)
        return(data.out)
}

# Veri seçimi ve dönüştürme

data=ins.line.data(grp.code[12])
data$ay=data$ay-2007
data$dev=data$dev-1

# Artan hasar ödemeleri
incr_pdloss=c() for(i in 9:0){
incr_pdloss <- c(incr_pdloss,
                subset(data[,5],data[,4]==0 & data[,3]==9-i),
        diff(subset(data[,5],data[,3]==9-i)))
}
data$incr_pdloss=incr_pdloss
attach(data)

## Üst Üçgen

Insample=subset(data,ay+dev<=9)
```

```
## Alt Üçgen

Outsample=subset(data,ay+dev>9)

## Üçgen için Veri

UUcgen_cum=as.triangle(Insample,origin="ay", dev="dev",
                                    value="cum_pdloss")
AUcgen_cum=as.triangle(Outsample,origin="ay", dev="dev",
                                    value="cum_pdloss")
UUcgen_incr=as.triangle(Insample,origin="ay", dev="dev",
                                    value="incr_pdloss")
AUcgen_incr=as.triangle(Outsample,origin="ay", dev="dev",
                                    value="incr_pdloss")

## Veri Alanları

xticks <- seq(2,11,2)

# Kümülatif Veri Alanları

(UUcgen_cum, xlab="Gelişim yılı",
        ylab="Kümülatif hasarlar",lwd="2", pch=c(48:57), xaxt="n")
axis(1, at = xticks, labels = xticks-1)
```

```
plot(UUcgen_cum, lattice=TRUE)

# Aşamalı Veri Alanları

(UUcgen _incr, xlab="Gelişim yılı",
        ylab="Aşamalı hasarlar",lwd="2", pch=c(48:57), xaxt="n")
axis(1, at = xticks, labels = xticks-1)

plot(UUcgen _incr, lattice=TRUE)

# Mullak Hasarlar

total_res <- (sum(subset(Outsample$cum_pdloss,
                         Outsample[,4]==9))
                         -sum(subset(Insample[,6],Insample[,3]!=0)))
ay_res <- (subset(Outsample[,5],Outsample[,4]==9)
            -subset(Insample,
                    (Insample$ay+Insample$dev)==9
                     & Insample$ay!=0)[,5])
```

Ek 2: Aşırı-yayılmış Negatif Binom Normal Yaklaşım Modeli

```
---------------------------------------------------------------------------------------------------------------------------

## Veri Hazırlığı (gelişim faktörleri & ağırlıklar)


dev_factor <- c()

for (i in 1:10){

        for (j in 1:(11-i)){

                ifelse(j>1,

                        dev_factor <- c(dev_factor,

                                        UTriangle_cum[i,j]/ UUcgen_cum[i,j-1]),

                        dev_factor <- c(dev_factor,0))

        }

}

Insample$dev_factor <- dev_factor


ağırlıklar <- c()

for (i in 1:10){

for (j in 1:(11-i)){

        ifelse(j>1,

                ağırlıklar <- c(ağırlıklar,UUcgen_cum[i,j-1]),

                ağırlıklar <- c(ağırlıklar,0))

        }

}

Insample$ağırlıklar  <- ağırlıklar

NONB_data <- subset(Insample, Insample$dev_factor!=0)
```

```
## GDM ortalama ile ilişkisi

NONB0_model <- glm(dev_factor ~ factor(dev),
                   family = gaussian(link="identity"),
                   data = NONB_data, ağırlıklar = ağırlıklar)
summary(NONB0_model)

## Yayılma cevabının hesaplanması

disp <- residuals(NONB0_model, "pearson")^2
NONB_data$disp <- disp

## GDM için yayılma parametresi

NONB_model_disp <- glm(disp ~ factor(dev),
                   family = Gamma(link="log"),
                   data = NONB_data)
summary(NONB_model_disp)

NONB_phi <- fitted(NONB_model_disp)
NONB_phi[9] <- NONB_phi[8]

## Güncellenen ağırlıklar ile ortalama GDM

yeni_ağırlıklar <- NONB_data$ağırlıklar/NONB_phi
```

```
NONB_data$NONB_ağırlıklar <- yeni_ağırlıklar
NONB_model <- glm(dev_factor ~ factor(dev),
                  family = gaussian(link="identity"),
                  data = NONB_data, ağırlıklar = NONB_ağırlıklar)
summary(NONB_model)

## Üst Üçgenin Fit edilmesi

NONB_fitted <- fitted(NONB_model)
NONB_fitted_cum <- NONB_fitted*NONB_data$ağırlıklar

## Model Parametreleri

NONB_coef <- (coef(NONB_model))
n <- length(NONB_data$dev_factor)
p <- length(NONB_coef)

## Ölçekli sapma ve Pearson artıkları

NONB_dresid <- residuals(NONB_model, "sapma")
plot(NONB_dresid ~ NONB_fitted,
     ylab="Deviance residuals", xlab="Fit edilen değer")
abline(h=0, "lty"=2, col="red")

NONB_presid <- residuals(NONB_model, "pearson")
```

```
plot(NONB_presid ~ NONB_fitted,
        ylab="Pearson residuals", xlab="Fit edilen değer")
abline(h=0, "lty"=2, col="red")

## Modelin sapması

NONB_D <- sum(residuals(NONB_model, "sapma")^2)

## Ölçekli Pearson İstatistiği

NONB_P <- sum(residuals(NONB_model, "pearson")^2)

## Gözlenen vs. fit edilen değerler
# Fit Edilen Değerler için Sapma Faktörü

plot(NONB_data$dev_factor ~ NONB_fitted,
        ylab="Observed values", xlab="Fit edilen değer")
abline(0, 1, "lty"=2, col="red")

# Fit Edilen Değerler için Kümülatif Hasarlar

plot(NONB_data$cum_pdloss ~ NONB_fitted_cum,
        ylab="Observed values", xlab="Fit edilen değer")
bline(0, 1, "lty"=2, col="red")
```

```
## Alt Üçgenin Tahmini


NONB_predicted <- c()
for(i in 1:9){
        for(j in 2:10){
                ifelse(Outsample[,3]==i & Outsample[,4]==j-1,
                        NONB_predicted <- c(NONB_predicted,
                                                NONB_fitted[j-1]
                                                *(ifelse(i+j==11,
                                                        UUcgen_cum[1+i,j-1],
                0)
        }
}

## Tahmini Nihai Rezervler


NONB_res_ay <- (subset(NONB_predicted,Outsample[,4]==9)
                -subset(Insample,
                        (Insample$ay+Insample$dev)==9
                        & Insample$ay!=0)[,5])
NONB_res_total <- (sum(subset(NONB_predicted,
                        Outsample[,4]==9))
                -sum(subset(Insample[,6],Insample[,3]!=0)))
```

```
## Artıkların Serbestlik Derecesine göre ayarlanması

NONB_resid <- (n/(n-p))*(residuals(NONB_model, "pearson"))

## Bootstrap Hazırlığı

set.seed(19951201)

## Bootstrap rezervlerinin matris / vektörünün hazırlanması

NONB_boot_res_ay <- c()
NONB_boot_res_total <- c()

## Bootstrap döngüsü

for (b in 1:10000){

# Residuals resampling & bootstrap üçgeni

resam <- sample(NONB_resid,n,replace=TRUE)
NONB_boot_triangle <- resam*sqrt(1/NONB_data$NONB_weights)
                      +NONB_fitted
```

```
# Model fit

NONB_boot_model <- glm(NONB_boot_triangle ~ factor(dev),
                       family = gaussian(link="identity"),
                       data = NONB_data, weights = NONB_weights)
NONB_boot_fitted <- fitted(NONB_boot_model)

# Gelecek Üçgeni

 NONB_boot_coef <- (coef(NONB_boot_model))
NONB_boot_predicted <- c()
for(i in 1:9){
        for(j in 2:10){
                ifelse(Outsample[,3]==i & Outsample[,4]==j-1,
                        NONB_boot_predicted <- c(NONB_boot_predicted,
                                                 NONB_boot_fitted[j-1]),
                0)
        }
}

# Normal Dağılımın Simülasyonu

value <- function(i,j){
        ifelse(Outsample[,3]==i & Outsample[,4]==j-1,
                x <- ifelse(i+j==11,UUcgen_cum[1+i,j-1],
```

```
                    tail(NONB_boot_sim,1)),0)
        return(x)
}
NONB_boot_sim <- c()
for(i in 1:9){
        for(j in 2:10){
                ifelse(Outsample[,3]==i & Outsample[,4]==j-1,
                        NONB_boot_sim<- c(NONB_boot_sim,
                                rnorm(1,mean=NONB_boot_fitted[j-1]
                                        *value(i,j)
                                ,sd=sqrt(NONB_phi[j-1]
                                        *value(i,j)))),
                0)
        }
}

# Bootstrap Rezerv Hesaplaması

NONB_boot_res_tmp <- c()
NONB_boot_res_tmp <- (subset(NONB_boot_sim,Outsample[,4]==9)
                    -subset(Insample,
                            (Insample$ay+Insample$dev)==9
                            & Insample$ay!=0)[,5])
NONB_boot_res_ay <- cbind(NONB_boot_res_ay,NONB_boot_res_tmp)
NONB_boot_res_total <- c(NONB_boot_res_total,
```

```
                                    sum(subset(NONB_boot_sim,
                                            Outsample[,4]==9))
                                    -sum(subset(Insample[,6],
                                            Insample[,3]!=0))) }

## Rezerv dağılımlarının analizi
# Hasar yılı için ayrı ayrı Rezervler

 i=1 # select accident year of interest (1, ..., 9)
hist(NONB_boot_res_ay[i,])
ozet(NONB_boot_res_ay[i,])
quantile(NONB_boot_res_ay[i,], prob=0.95)
ortalama(subset(NONB_boot_res_ay[i,],
                NONB_boot_res_ay[i,]
                >quantile(NONB_boot_res_ay[i,],
                        olasılık=0.95)))
# Toplam Rezerv
hist(NONB_boot_res_total)
ozet(NONB_boot_res_total)
NONB_VaR <- quantile(NONB_boot_res_total, prob=0.95)
NONB_CVaR <- mean(subset(NONB_boot_res_total,
                                    NONB_boot_res_total
                                    >quantile(NONB_boot_res_total,
                                            olasılık=0.95)))
```

KAYNAKÇA

ALAI, D. H., MERZ, M. and WUTHRICH, M. V. (2009). *Mean square error of prediction in the Bornhuetter-Ferguson claims reserving method*. Annals of Actuarial Science, Vol. 4, No. 1, p. 7–31.

BORNHUETTER, R. L. and FERGUSON, R. E. (1972). *The actuary and IBNR*. Proc. CAS, Vol. 59, p. 181-195.

DE JONG, P. and HELLER, G. Z. (2008). *Generalized linear models for insurance data*. Cambridge University Press, New York.

DOBSON, A. J. (2002). *An introduction to generalized linear models*. 2nd Ed. CRC Press LLC, Boca Raton.

EFRON, B. and TIBSHIRANI, R. J. (1993). *An introduction to the bootstrap*. Chapman & Hall, New York.

ENGLAND, P., VERRALL, R.J., (2002), *Stochastic claims reserving in general insurance*, British Actuarial Journal, Faculty of actuaries and Instute of actuaries, 8, 3 443-518.

HOSSACK, I.B, POLLARD, j.H., ZEHNWIRTH, B., (1983), *Introductory Statistics with Applications in General Insurance*, Cambridge University Press, London,275p

KAAS, R, GOOVAERTS, M., DHAENE, J., DENUIT, M., (2008), *Modern Actuarial Risk Theory: Using R*, Springer Verlag Inc. , Berlin, 381p.

KREMER, E., (1982), *IBNR claims and the two way model of ANOVA*. Scandinavian Actuarial Journal, 47-55

LINDSEY, J.K., (1997), *Applying Generalized Linear Models*, Springer Verlag Inc., New York, 258 p.

MACK, T., (1991), *A simple parametric model for rating automobile insurance or estimating IBNR claims reserves*, ASTIN Bulletin, 22, 1, 93-109.

MACK, T., (1994), *Which stochastic models underlying the chain-ladder model Insurance: Mathematics and Economics*, 15, 133-138.

MCCULLAGH, P. and NELDER, J.A., (1989), *Generalized Linear Models*, Chapman and Hall, London, 511p.

PINHEIRO, P.J.R., ANDRADE E SILVA, J.M., CENTENO, M.L., (2003), *Bootstrap methodology in claim reserving*, The Journal of Risk and Insurance,70, 4, 701-714.

TAYLOR, G., MCGUIRE, H., (2004), *Loss reserving with GLMs: a case study*, Casualty Actuarial Society 2004 Discussion Paper Program, 327-392.

ENGLAND, P. D. (2002). *Addendum to 'Analytic and bootstrap estimates of prediction errors in claims reserving'. Insurance: Mathematics and Economics*, Vol. 31, p. 461-466.

ENGLAND, P. D. and VERRALL, R. J. (1999). *Analytic and bootstrap estimates of prediction errors in claims reserving. Insurance: Mathematics and Economics*, Vol. 25, p. 281-293.

ENGLAND, P. D. and VERRALL, R. J. (2006). *Predictive distributions of outstanding liabilities in general insurance.* Annals of Actuarial Science, Vol. 1, No. 2, p. 221-270.

MACK, T. (1993). *Distribution-free calculation of the standard error of chainladder reserve estimates.* ASTIN Bulletin, Vol. 23, No. 2, p. 213-225.

MCNEIL, A. J., FREY, R. and EMBRECHTS, P. (2005). *Quantitative risk management: concepts, techniques and tools*. Princeton University Press, Princeton.

MEYERS, G. G. and SHI, P. (2011). *Loss reserving data pulled from NAIC Schedule P*. Casualty Actuarial Society. http://www.casact.org/research/index.cfm?fa=loss_reserves_data (Accessed: 9th November 2014).

BROWN, R.L., (1993), *Introduction to Ratemeking and Loss Reserving for Property and Casualty Insurance,* ACTEX Publications, Connecticut, 157 p.

ENGLAND, P., VERRALL, R.J., (1998), *Standard errrors of prediction in claims reserving: a comparison of methods*, Proceedings of the General Insurance Convention & ASTIN Colloquium in Glasgow, 1, 459-478.

R CORE TEAM (2014). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna. http://www.R-project.org/ (Accessed: 31st October 2014).

SCHMIDT, K. D. (2006). *Methods and models of loss reserving based on run-off triangles: a unifying survey*. CAS Forum (Fall), p. 269-317.

VERRALL, R. J. (2000). *An investigation into stochastic claims reserving models and the chain- ladder technique*. Insurance: Mathematics and Economics, Vol. 26, p. 91-99.

WEDDERBURN, R. W. M. (1974). *Quasi-likelihood functions, generalized linear models and the Gauss-Newton method*. Biometrika, Vol. 61, No. 3, p. 439-447.

WUTHRICH, M. V. and MERZ, M. (2008). *Stochastic claims reserving methods in insurance*. Wiley, Chichester.