

**BERNSTEIN KOPULA İLE HAYAT DIŐI SİGORTASINDA
BAĞIMLILIĞIN MODELLENMESİ**

**DEPENDENCE MODELING OF NON-LIFE INSURANCE
USING THE BERNSTEIN COPULA**

PINAR ÖZER UYAR

Dr. Öğr. Üyesi MURAT BÜYÜKYAZICI
Tez Danışmanı

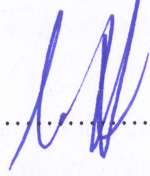
Hacettepe Üniversitesi
Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin
Aktüerya Bilimleri Ana Bilim Dalı için Öngördüğü
YÜKSEK LİSANS TEZİ
olarak hazırlanmıştır.

2018

Pınar Özer Uyar'ın hazırladığı “Bernstein Kopula İle Hayat Dışı Sigortasında Bağımlılığın Modellenmesi” adlı çalışma aşağıdaki jüri tarafından AKTÜERYA BİLİMLERİ ANABİLİM DALI'nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

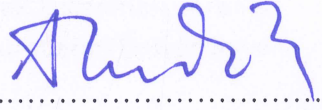
Doç. Dr. Canan HAMURKAROĞLU

Başkan



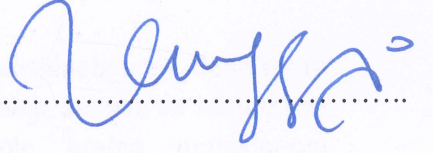
Dr. Öğr. Üyesi Murat BÜYÜKYAZICI

Danışman



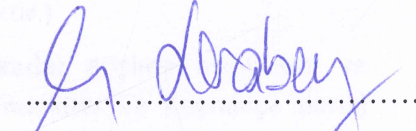
Doç. Dr. Duygu İÇEN

Üye



Dr. Öğr. Üyesi Uğur KARABEY

Üye



Dr. Öğr. Üyesi Başak BULUT KARAGEYİK

Üye



Bu tez Hacettepe Üniversitesi Fen Bilimleri Enstitüsü tarafından YÜKSEK LİSANS TEZİ olarak onaylanmıştır.

Prof.Dr. Menemşe GÜMÜŞDERELİOĞLU

Fen Bilimleri Enstitü Müdürü

YAYINLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI

Enstitü tarafından onaylanan lisansüstü tezimin/raporumun tamamını veya herhangi bir kısmını, basılı (kağıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanması zorunlu metinlerin yazılı izin alarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

- ☐ Tezimin/Raporumun tamamı dünya çapında erişime açılabilir ve bir kısmı veya tamamının fotokopisi alınabilir.

(Bu seçenekle teziniz arama motorlarında indekslenebilecek, daha sonra tezinizin erişim statüsünün değiştirilmesini talep etmeniz ve kütüphane bu talebinizi yerine getirirse bile, tezinin arama motorlarının önbelleklerinde kalmaya devam edebilecektir.)

- ☐ Tezimin/Raporumun tarihine kadar erişime açılmasını ve fotokopi alınmasını (İç Kapak, Özet, İçindekiler ve Kaynakça hariç) istemiyorum.

(Bu sürenin sonunda uzatma için başvuruda bulunmadığım takdirde, tezimin/raporumun tamamı her yerden erişime açılabilir, kaynak gösterilmek şartıyla bir kısmı ve ya tamamının fotokopisi alınabilir)

- ☒ Tezimin/Raporumun 11/06/2018 tarihine kadar erişime açılmasını istemiyorum, ancak kaynak gösterilmek şartıyla bir kısmı veya tamamının fotokopisinin alınmasını onaylıyorum.

- ☐ Serbest Seçenek/Yazarın Seçimi

11 / 06 / 2018

(imza) 

Öğrencinin Adı Soyadı

Pınar ÖZER UYAR

ETİK

Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada,

- tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- ve bu tezin herhangi bir bölümünü bu üniversitede veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı,

beyan ederim.

TARİH

11/06/2018

PL"

PINAR ÖZER UYAR

ÖZET

BERNSTEIN KOPULA İLE HAYAT DIŐI SİGORTASINDA BAĞIMLILIĞIN MODELLENMESİ

Pınar ÖZER UYAR

Yüksek Lisans, Aktüerya Bilimleri Bölümü

Tez Danışmanı: Dr. Öğr. Üyesi Murat BÜYÜKYAZICI

Haziran 2018, 50 Sayfa

Kopula rasgele değişkenler arasındaki bağımlılığın modellenmesinde kullanılan dağılım fonksiyonlarıdır. Bu nedenle bağımlılık çalışmaları kapsamında sıklıkla kullanılmaktadır.

Bu çalışmada hayat dışı sigorta risklerinin Bernstein kopula ile modellenmesinin etkinliği, diğer kopula modelleri ile karşılaştırılarak incelenmiştir. Tezin uygulama kısmında Türkiye’de hayat dışı sigorta alanında hizmet veren bir şirkete ait kasko, trafik, ihtiyari mali mesuliyet, konut/yangın, konut/hırsızlık, konut/cam kırılması, ferdi kaza, kasko koltuk ferdi kaza, bireysel sağlık sigortaları için 72 aylık ödenen hasar verileri alınmıştır. Öncelikle aralarında yüksek korelasyon olan kasko-trafik, trafik-bireysel sağlık, kasko-imm sigortaları seçilmiş ve bu veri çiftlerine ait dağılım ve parametreleri belirlenmiştir. Kopula model seçimi için uyum iyiliği ölçümü olarak öklid uzaklıklarına bakılmıştır. Son olarak hasar veri çiftleri kullanılarak her bir kopula modeli için tahmin edilen toplam hasar verilerinin riske maruz değerleri (VaR) hesaplanmıştır. Çalışmanın sonuçları sigorta şirketleri ve düzenleyicilerinin risk modellemelerinde Bernstein kopula modelini de dikkate almalarının yararlı olabileceğini göstermiştir.

Anahtar Kelimeler: Bağımlılık, hayat dışı sigorta riski, kopula, Bernstein kopula

ABSTRACT

DEPENDENCE MODELING OF NON-LIFE INSURANCE USING THE BERNSTEIN COPULA

Pınar ÖZER UYAR

Master of Science, Department of Actuarial Sciences

Supervisor: Assist. Prof. Dr. Murat BÜYÜKYAZICI

June 2018, 50 pages

Copula is distribution functions which is used in modelling dependency between random variables. For this reason, it is frequently used in dependency studies.

In this study, the effectiveness of the modeling of non-life insurance risks with the Bernstein copula was examined by comparing with other copula models. Study's data got from an insurance company which is operating in the field of non-life insurance in Turkey. This data includes paid incurred for 72 months in the area of traffic insurance, car insurance, liability insurance (imm), residential fire insurance, residential theft insurance, residential glass breakage insurance, personal accident insurance, seat personal accident insurance and personal health insurances. First of all, car insurance-traffic, traffic-personal health and car insurance-imm were selected which have high correlation between them. Then the distribution and parameters of these data pairs were determined. In order to select the model, Euclidean distances were calculated. Finally, for each copula model Risk exposure values (VaR) were calculated.

The results of the study showed that it would be useful for insurance companies and regulators to consider the Bernstein copula model in risk models.

Key Words: Dependency, Non-life insurance risk, Copula, Bernstein Copula

TEŞEKKÜR

Tez çalışmamın her aşamasında bilgi ve deneyimleriyle desteğini esirgemeyen, çalışmam süresince bana yol gösteren danışmanım Sayın Dr. Öğr. Üyesi Murat BÜYÜKYAZICI'ya; değerli görüş ve katkıları için tez jüri üyelerim Sayın Doç. Dr. Canan HAMURKAROĞLU, Doç. Dr. Duygu İÇEN, Dr. Öğr. Üyesi Uğur KARABEY ve Dr. Öğr. Üyesi Başak BULUT KARAGEYİK'e;

bu süreç boyunca desteklerini ve sevgilerini her an hissettiğim Ankara'da ki ailem, Arş. Gör. Müge YELDAN, Melis ERKAN, Tuğçe Seçil TEKTEN ve N. Selvi YILDIRIM'a;

her daim aldığım kararların arkasında duran, sonsuz sevgi ve güven duygularıyla beni büyüten aileme;

hayatımın her aşamasında tüm varlığıyla yanımda olan, sevgisiyle beni güzelleştiren değerli eşim Öğr. Gör. Mevlüt UYAR'a

teşekkürlerimi sunarım.

İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET	i
ABSTRACT	ii
TEŞEKKÜR	iii
İÇİNDEKİLER.....	iv
ŞEKİLLER	vi
ÇİZELGELER.....	vii
1. GİRİŞ.....	1
2. BAĞIMLILIK	4
2.1. Doğrusal Bağımlılık	4
2.2. Tam Bağımlılık.....	5
2.3. Sıra Korelasyonlu Bağımlılık	5
2.3.1. Kendall’ın Tau Katsayısı	5
2.3.2. Spearman’ın Rho Katsayısı	6
2.3.3. Kendall’ın Tau Katsayısı ile Spearman’ın Rho Katsayısı Arasındaki İlişki	7
2.4. Kuyruk Bağımlılığı.....	8
3. KOPULA.....	9
3.1. Kopula Tanımı ve Özellikleri.....	10
3.2. Sklar Teoremi	12
3.3. Frechet-Hoeffding Sınırları	14
3.4. Yaşam Kopulası.....	14
3.5. Parametrik Kopulalar.....	15
3.5.1. Eliptik Kopula.....	15
3.5.1.1. Gaussian Kopula.....	15
3.5.1.2. T-Kopula.....	16

3.5.2. Arşimedyan Kopula.....	17
3.5.2.1. İki Değişkenli Arşimedyan Kopula	17
3.5.2.2. Bazı Önemli Arşimedyan Kopula Aileleri	19
3.5.2.2.1. Gumbel Kopula.....	19
3.5.2.2.2. Clayton Kopula.....	19
3.5.2.2.1. Frank Kopula	19
3.5.2.3. Çok Değişkenli Arşimedyan Kopula.....	20
3.6. Parametrik Olmayan Kopula	20
3.6.1. Deneysel (Empirical) Kopula	21
3.6.2. Kopula Yaklaşımları.....	21
3.6.2.1. Bernstein Polinomları Yaklaşımı.....	21
3.6.2.2. Grid-Type Kopula Yaklaşımı	27
3.7. Bernstein Kopula ile Rasgele Sayı Üretme	28
4. UYGUN KOPULA MODEL SEÇİMİ UYGULAMASI.....	30
4.1. Ödenen hasar verisi ön incelemesi	31
4.2. Uyum İyiliği	37
4.3. Risk Modellemesi.....	39
5. SONUÇ ve ÖNERİLER	42
KAYNAKLAR.....	44
EKLER	47
ÖZGEÇMİŞ.....	49

ŞEKİLLER

Sayfa

Şekil 2. 1: Sürekli rasgele değişken çiftleri için τ ve ρ sınırları	8
Şekil 3. 1: Tam (a) ve tam olmayan (b) üreticiler ve tersleri	18
Şekil 3. 2: $B_{0,1}(t), B_{1,1}(t)$ Bernstein polinomu	22
Şekil 3. 3: $B_{0,2}(t), B_{1,2}(t), B_{2,2}(t)$ Bernstein polinomları	22
Şekil 3. 4: $B_{0,3}(t), B_{1,3}(t), B_{2,3}(t), B_{3,3}(t)$ Bernstein polinomları	23
Şekil 4. 1: Hasar Verilerine Ait Korelasyon Matris Grafiği	31
Şekil 4. 2: Kasko ve Trafik verisine ait saçılım grafiği.....	34
Şekil 4. 3: Trafik ve Bireysel sağlık sigortası verisine ait saçılım grafiği	34
Şekil 4. 4: Kasko ve İmm verisine ait saçılım grafiği.....	34
Şekil 4. 5: Kasko ve Trafik verilerine ait deneysel kopula yoğunlukları yüzey grafiği ile Bernstein kopula yoğunlukları yüzey grafiği	35
Şekil 4. 6: Trafik ve Bireysel sağlık sigortası verilerine ait deneysel kopula yoğunlukları yüzey grafiği ile Bernstein kopula yoğunlukları yüzey grafiği.....	35
Şekil 4. 7: Kasko ve İmm verilerine ait deneysel kopula yoğunlukları yüzey grafiği ile Bernstein kopula yoğunlukları yüzey grafiği	35
Şekil 4. 8: $m=5$ için Kasko Trafik veri çiftine ait Bernstein kopula yoğunlukları yüzey grafiği	36
Şekil 4. 9: $m=10$ için Kasko Trafik veri çiftine ait Bernstein kopula yoğunlukları yüzey grafiği	36
Şekil 4. 10: $m=15$ için Kasko Trafik veri çiftine ait Bernstein kopula yoğunlukları yüzey grafiği.....	36
Şekil 4. 11: Optimal grid grafikleri	39

ÇİZELGELER

Sayfa

Çizelge 3. 1: Kendall'ın sıra korelasyon katsayısı ve kopula aileleri için kuyruk bağımlılık katsayıları	20
Çizelge 4. 1: Hasar Verilerine Ait Korelasyon Matrisi	32
Çizelge 4. 2: Verilere uygun dağılım ve parametreler.....	33
Çizelge 4. 3: Model seçimi için parametre tahminleri.....	37
Çizelge 4. 4: 72'lik örneklem için 1000 simülasyondan elde edilen ortalama öklid uzaklıkları	38
Çizelge 4. 5: %99,5 güven düzeyinde VaR değerleri (Milyon TL)	40

1. GİRİŞ

Rasgele değişkenler arasındaki bağımlılığı ilişkilendirmek olasılık ve istatistikte en çok incelenen konulardan biridir. Bağımlılıkların ölçülmesinde ve derecelendirilmesinde ise çeşitli yaklaşım ve modellemeden yararlanılır. Kopula bağımlılık çalışmalarında seçilen ve rasgele değişkenler arasında ilişkilerin belirlenmesinde kullanılan önemli fonksiyonlardan biridir. Kopula latince sözlük de ‘ilişkilendirmek, birleştirmek, bağlamak’ anlamına gelir. İlk olarak kopula kavramı ise 1959 yılında Abe Sklar tarafından matematiksel ve istatistiksel anlamıyla kullanılmıştır. A. Sklar kopula üzerine yaptığı çalışmasında olasılıksal metrik uzaylar teoremini geliştirmiştir [1]. Kopula fonksiyonu genel anlamda bağımlılık yapısını analiz eder ve bağımlılık yapısının derecesi hakkında bilgi verir [2]. Kopula tek değişkenli marjinal dağılımlarından çok değişkenli dağılım fonksiyonu veren bir yapıdır. Tek değişkenlerin her birinin özelliklerinden ziyade bağımlılık yapısının incelenmesine imkan vermesi onu oldukça kullanışlı kılmıştır. Kopulaya olan ilgi arttıkça literatür genişlemeye devam etmiştir. Doksanlı yılların sonuna doğru bu kavram finans, aktüerya ve sigortacılık gibi bir çok uygulama alanlarındaki sorulara cevaplar aramaya başlamıştır [3].

1997 yılında ki çalışmasında Wang [4] sigorta poliçeleri ile toplam kayıp dağılımı arasındaki ilişkiyi modellemek ve birleştirmeyi sağlamak için kopulayı kullanmıştır. Çalışmasında ayrı fakat birbiriyle ilişkili iş sınıfları için toplam kayıp dağılımlarını ‘nasıl birleştiririm’ sorusuna cevap aramış. 1998 yılında Frees ve Valdez [5] çalışmasında sigorta hasar verileri için iki değişkenli kopula modeli kullanmıştır. 1999 yılında Klugman ve Parsa [6] iki değişkenli sigorta hasar verilerini modellemek için kopulayı kullanmıştır. 1999 yılında Shemyakin ve Noun [7] genelde evli çiftlerin kullandığı ortak yaşam poliçesi üzerine bir çalışma gerçekleştirmiştir. Bu çalışmada çiftler arasındaki yaş farkını değişken olarak belirleyip bu yaş farkına bağlı olarak çiftlerin yaşam fonksiyonunu oluşturmuştur. Oluşturduğu kopula modeli çiftler arasında yaş farkına bağlı olarak ilişki parametresi ile marjinal fonksiyonlarına dayanmıştır. Çalışmanın sonucu son kalan yaşam sigorta fiyatlandırılması (last survivor policy) açısından faydalı olmuştur. 2002 yılında çalışmalarında Embrechts, Mcneil ve Straumann [8] bağımlılık yapılarını incelerken benzetim tekniği ile çalışmalarını desteklemiştirler. Bu çalışmaya dayanarak 2001 yılında Embrechts, Mcneil ve Lindskog [9] kopula, kopula çeşitleri ve kopulanın kullanımını finansal alanda uygulayarak kapsamlı bir bakış açısı sunmuşlardır. Genest ve Rivest [10], [11] ve Genest ve MacKay [12] iki değişkenli kopula ve çok değişkenli kopula ile doğrusal

olmayan bağımlılıkların modellenmesini deneysel verilerle gerçekleştirmiştir. Hayat tablolarının tahmininde kullanılmak üzere Oakes [13],[14] iki değişkenli ve çok değişkenli yaşam modellerini kullanmıştır. Frees ve Valdez [15], Youn ve Shemyakin [16], Youn, Shemyakin ve Herman [17], Anderson [18] bağımlı yaşam sürelerinin doğru tahminlerini yapabilmek, çoklu yaşam sigorta sözleşmelerinde fiyatlandırma ve risk ölçümlerinin değerlendirilebilmesi için çalışmışlardır. Joe [19] ve Nelsen [1] bağımlılık ölçümlerini, kopula ve çeşitlerini, bağımlılıkların modellenmesinde kullanılan kopulaları kapsamlı bir şekilde analiz etmişlerdir.

İki binli yıllarda McNeil ve Demarta [20], Aas [21], Luo, Ryde ve Shevchenko [22] çalışmalarında eliptik kopula ve arşimedyan kopula türlerini kullanarak benzetim çalışmaları yapmışlardır. Eliptik kopula ve Arşimedyan kopulalar ile çalışmalar devam ederken 2004 yılında Sancetta ve Satchell [23],[24] çalışmalarında yeni bir kopula modeli olarak Bernstein kopulayı tanımlamışlardır. Bernstein kopula özelliklerini, kopula tahmin edicilerini ve kopulanın ileri boyutlu analizlerde kullanımı hakkında bilgiler vermişlerdir. 2012 yılında Diers, Elling ve Marek [25] çalışmasında Bernstein kopulayı kullanarak gözleme dayalı verilerle bağımlılığı modellemişlerdir. Pfeifer ve Cottin [26], Bernstein kopulayı bazı deneysel kopula türleri ile karşılaştırıp Bernstein kopulasının da deneysel bir kopula olduğu izlenimine varmıştır. 2015 yılında Yang, Chen ve Wang [27] çalışmasında çok boyutlu kopulaların yeni bir sınıfı olarak bileşik Bernstein kopulayı tanıtmışlardır.

Bu tezin yazılmasındaki amaç hayat dışı sigorta ürünleri arasındaki bağımlılıkların modellenmesinde Bernstein kopula etkinliğini, diğer kopula modelleri ile karşılaştırmaktır. Ortaya konulan modelleme ile sektöre ve literatüre katkı sağlayacağı düşünülmektedir. Çalışmada modelleme oluşturulurken benzetim tekniğinden yararlanılmıştır.

Çalışmanın ikinci bölümünde bağımlılık kavramı, bağımlılık çeşitleri, bağımlılık ölçümleri ve bu ölçümler arasındaki ilişkiler açıklanmıştır.

Çalışmanın üçüncü bölümünde ilk olarak kopulanın tanımı ve özellikleri açıklanmıştır. Daha sonra parametrik ve parametrik olmayan kopula çeşitleri tanımlanarak, bu kopulaların özellikleri üzerinde durulmuştur. Çalışmanın temelini oluşturan Bernstein kopuladan rasgele sayı üretilme aşamaları anlatılmış ve parametrik olmayan kopula yaklaşımları hakkında bilgiler verilmiştir.

Çalışmanın dördüncü bölümünde Türkiye’de hayat dışı sigorta alanında faaliyet göstermekte olan bir sigorta şirketinin farklı branşlara ait hasar verilerinin kopula

modelleri için parametre tahminleri yapılmıştır. Gerçek veri ile farklı kopula tahminleri arasında uyum iyiliği testi yapılmış ve bu koşullar altında en uygun model belirlenmiştir. Bernstein kopula modelinin veri setlerine bağlı olarak optimal grid sayısının bulunması için MATLAB programı kullanılarak tahminler yapılmıştır. Risk değerlendirmesinin yapılabilmesi için tüm kopula modelleri ile rasgele bağımlı veriler üretip toplam hasarın VaR sonuçları karşılaştırılmıştır.

Çalışmanın son bölümünde ise uygulama sonuçları tartışılarak öneriler sunulmuştur.

2. BAĞIMLILIK

Rasgele değişkenler arasındaki bağımlılığın çalışılmasında bir çok ölçüm kullanılmaktadır. Bağımlılık özellikleri ile bu ölçümler birbiriyle ilişkilidirler. Kopulalar ise rasgele değişkenler arasındaki bağımlılığı incelemek ve ölçmek için doğal bir yol sağlar. Kopulanın temel özelliği rasgele değişkenlerin oluşturduğu kesin sürekli, artan dönüşümler altında değişmezliktir [9]. Bu bölümde bağımlılık ölçümleri; doğrusal (lineer) bağımlılık, tam bağımlılık, sıra korelasyonu olarak bağımlılık ve kuyruk bağımlılığı olarak ele alınmaktadır.

2.1. Doğrusal Bağımlılık

Doğrusal korelasyon yaygın olarak kullanılan bağımlılık ölçümlerinden biridir. Doğrusal korelasyon katsayısı Pearson bağımlılık katsayısı olarak da bilinmektedir. (X, Y) varyansları sıfır olmayan sonlu varyanslar ile rasgele değişkenlerin bir vektörü olmak üzere X ve Y arasındaki doğrusal korelasyon katsayısı;

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}, \quad Cov(X, Y) = E(XY) - E(X)E(Y)$$

$Cov(X, Y)$, (X, Y) 'nin kovaryansı, $Var(X)$ ve $Var(Y)$ sırasıyla X ve Y 'nin varyanslarını göstermektedir.

Doğrusal korelasyon bir bağımlılık ölçümüdür. ρ , doğrusal korelasyon katsayısı olarak adlandırılır çünkü bu ölçüm rasgele değişkenler arasındaki doğrusal bağımlılığı verir. Aslında $\rho(X, Y)$ hakkındaki ilişkinin yönüne ilişkin bilgi, doğrusal regresyondaki

$Y = \beta X + \varepsilon$ eşitliğindeki β bilgisine eşittir. ε , doğrusal olarak X 'in ilişkisiz olduğu 'artık' anlamına gelir.

Mükemmel doğrusal bağımlılık için $|\rho(X, Y)| = 1$ 'dir. Doğrusal korelasyon katsayısı ρ ,

$-1 < \rho(X, Y) < 1$ arasındadır. İkinci bir özellik ise $\rho(\alpha X + \delta_1, \gamma Y + \delta_2) = \text{sign}(\alpha\gamma) \rho(X, Y)$ $\alpha, \gamma \in \mathbb{R} \setminus \{0\}$, $\delta_1, \delta_2 \in \mathbb{R}$ eşitliğini sağlar. Diğer bir özellik, doğrusal korelasyon katsayısı kesin artan lineer dönüşümler altında sabittir. Doğrusal korelasyon doğrusal operatörler ile kolayca düzenlenebilir.

Doğrusal korelasyonun kolay hesaplanabilir olmasıyla daha çok kullanılan bir bağımlılık ölçümüdür. Çok değişkenli normal, çok değişkenli t dağılımı gibi eliptik dağılımlarda kullanılan doğal bir bağımlılık ölçümü skaleridir. Ortak eliptik dağılıma uymayan rasgele değişkenler için doğrusal korelasyon yapısı kullanıldığında yanıltıcı sonuçlara varılabilir.

2.2. Tam Bağımlılık

Bütün n-boyutlu kopulalar için Fréchet-Hoeffding eşitsizliği sağlanır.

$$W^n(u_1, u_2, \dots, u_n) \leq C(u_1, u_2, \dots, u_n) \leq M^n(u_1, u_2, \dots, u_n).$$

Bu durumda W, Fréchet-Hoeffding alt sınırı ve negatif tam bağımlılığı; M, Fréchet-Hoeffding üst sınırı ve pozitif tam bağımlılığı ifade eder. Eğer (X, Y), M kopulasına sahip ise X ve Y comonotoniktir. Eğer (X, Y), W kopulasına sahip ise countermonotonik denir.

Teorem 2.1: (X, Y), W ya da M kopulasından herhangi birine sahip olsun. Bu durumda α, β olmak üzere iki monoton fonksiyon vardır. $\alpha, \beta: \mathbb{R} \rightarrow \mathbb{R}$ ve Z bir rasgele değişkeni için

$$(X, Y) = (\alpha(Z), \beta(Z))_d$$

sağlanır.

W kopulasına sahip olması durumunda α artan fonksiyon, β azalan fonksiyondur. M kopulasına sahip ise her iki fonksiyon α, β artan fonksiyonlardır [9].

2.3. Sıra Korelasyonlu Bağımlılık

Sıra korelasyonu, kopulayı verilere uyumlu kılmanın olası bir yolunu sunar [28]. Bir çift rasgele değişkenin birinin büyük değeri diğerinin büyük değeri ile, küçük değerlerinden biri diğerinin küçük değeri ile ilişkilendirilebiliyorsa bu rasgele çift uyumludur. Daha açıklayıcı bir anlatımla, (x_i, y_i) ve (x_j, y_j) , (X, Y) sürekli rasgele değişkenlerinin iki gözlemi olsun. Eğer $x_i < x_j$ iken $y_i < y_j$ oluyorsa (x_i, y_i) ile (x_j, y_j) uyumludur. Benzer şekilde eğer $x_i > x_j$ iken $y_i < y_j$ oluyorsa (x_i, y_i) ile (x_j, y_j) uyumlu değildir denir. Farklı bir şekilde ifade edersek; eğer $(x_i - x_j)(y_i - y_j) > 0$ oluyorsa (x_i, y_i) ile (x_j, y_j) bağımlıdır, $(x_i - x_j)(y_i - y_j) < 0$ oluyorsa uyumlu değildir denir. Sıra korelasyonu ölçümlerinden en önemlileri Kendall'ın tau katsayısı ve Spearman'ın rho katsayısıdır.

2.3.1. Kendall'ın Tau Katsayısı

$\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, (X, Y) sürekli rasgele değişken vektöründen oluşan n tane rasgele gözlemin örnekleme olsun. Örnekleme uyumlu, uyumlu olmayan $\binom{n}{2}$ tane birbirinden farklı (x_i, y_i) ve (x_j, y_j) gözlem çiftleri vardır. c uyumlu çiftlerin sayısını, d uyumlu olmayan çiftlerin sayısını belirtsin. Böylelikle örneklem için Kendall'ın tau katsayısı

$$\tau = \frac{c-d}{c+d} = \frac{c-d}{\binom{n}{2}}$$

eşitliğiyle bulunur. τ , örneklemde rasgele seçilen (x_i, y_i) ve (x_j, y_j) gözlem çiftleri için uyumlu olanların olasılıklarından uyumlu olmayanların olasılıklarının çıkarılmasıyla bulunur. Ortak dağılım fonksiyonu H olan, (X, Y) sürekli rasgele değişkenlerinin oluşturduğu vektör için de Kendall'ın tau katsayısı benzer şekilde bulunur. (X_1, Y_1) ve (X_2, Y_2) bağımsız, aynı dağılımlı rasgele vektörler, ortak dağılım fonksiyonları H olsun. Örneklem için Kendall'ın tau katsayısı uyumlu olan olasılıklardan uyumlu olmayan olasılıkların çıkarılmasıyla bulunur.

$$\tau = \tau_{X,Y} = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]$$

2.3.2. Spearman'ın Rho Katsayısı

$(X_1, Y_1), (X_2, Y_2)$, ve (X_3, Y_3) bağımsız rasgele değişken vektörlerin ortak dağılım fonksiyonu H (marjinalleri F ve G olan) ve C kopula olsun. Spearman'ın rho katsayısı uyumluluk olasılığından, (X_1, Y_1) ile (X_2, Y_3) iki vektörün uyumlu olmayan olasılıklarının çıkarılmasıyla orantılıdır.

$$\rho_{X,Y} = 3(P[(X_1 - X_2)(Y_1 - Y_3) > 0] - P[(X_1 - X_2)(Y_1 - Y_3) < 0])$$

Teorem 2.2: X ve Y, C kopulasına sahip sürekli rasgele değişkenler olsun. X ve Y için Spearman'ın rho katsayısının büyüklüğü şu şekilde bulunur.

$$\rho_{X,Y} = \rho_C = \iint_0^1 uv \, dC(u, v) - 3 = \iint_0^1 C(u, v) \, dudv - 3$$

Spearman'ın rho katsayısı 'sınıf' korelasyon katsayısı olarak da bilinir. Sınıf sıralı çoklukların büyüklüğü demektir. Yani F ve G dağılım fonksiyonların rasgele değişkenleri X ve Y'den x ve y rasgele gözlemleri alınsın. Sırasıyla x ve y'nin sınıfları $u=F(x)$ ve $v=G(y)$ olsun. u ve v ortak dağılım fonksiyonu C olan, (0,1) aralığında $U=F(X)$, $V= G(Y)$ uniform dağılımından elde edilen gözlemlerdir. U ve V'nin ortalamaları $\frac{1}{2}$, varyansları $\frac{1}{12}$, olduğundan Spearman'ın rho katsayısı aşağıdaki eşitliğe dönüşür.

$$\begin{aligned} \rho_{X,Y} = \rho_C &= \iint_0^1 uv \, dC(u, v) - 3 = 12 E(UV) - 3 \\ &= \frac{E(UV) - 1/4}{1/12} = \frac{E(UV) - E(U)E(V)}{\sqrt{Var(U)}\sqrt{Var(V)}} \end{aligned} \quad (2.1)$$

Eşitlik 2.1'de görüldüğü gibi X ve Y için Spearman'ın rho katsayısı, X ve Y sınıflarının Pearson'ın korelasyon katsayısıyla benzerdir.

2.3.3. Kendall'in Tau Katsayısı ile Spearman'ın Rho Katsayısı Arasındaki İlişki

Kendall'in tau katsayısı ile Spearman'ın rho katsayısının bir çok ortak özelliği vardır. ξ sayısal değerini, C kopulasına sahip X ve Y sürekli rasgele değişkenlerin ortak bağımlılık ölçüsü olarak tanımlayalım.

- Her bir çift X, Y sürekli rasgele değişkenleri için κ tanımlıdır.
- $-1 \leq \xi_{X,Y} \leq 1$ arasında değer alır.
- Simetrik bir bağımlılık ölçüsüdür. $\xi_{X,Y} = \xi_{Y,X}$
- X ve Y bağımsız ise $\xi_{X,Y} = \xi_{\Pi} = 0$ 'dır.
- X ve Y kesin artan fonksiyonlar ise $\xi_{X,Y} = \xi_M = 1$. X ve Y kesin azalan fonksiyonlar ise $\xi_{X,Y} = \xi_W = -1$.
- $\{(X_n, Y_n)\}$ C_n kopulasıyla sürekli rasgele değişkenlerinin dizisi ve eğer C_n C'ye noktasal olarak yakınsar ise $\lim_{n \rightarrow \infty} \xi_{C_n} \rightarrow \xi_C$.

Kendall'in tau katsayısı ile Spearman'ın rho katsayısı bir kopula ile rasgele değişkenler arasındaki uyum olasılığının ölçümüdür. Ortak özellikleri oldukça çok olmasına rağmen τ , ρ değerleri sıklıkla farklı değerler alır. Matematiksel olarak hesaplanışları bu farkı yaratır.

Teorem 2.3: X ve Y sürekli rasgele değişkenler ve τ , ρ sırasıyla Kendall'in tau'su ve Spearman'ın rho'su olsun.

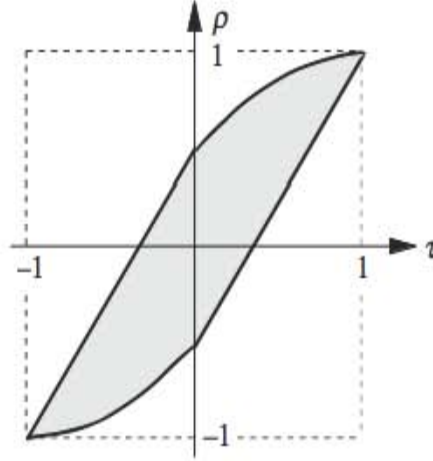
$$-1 \leq 3\tau - 2\rho \leq 1$$

Teorem 2.4: X ve Y sürekli rasgele değişkenler ile C kopula fonksiyonu, τ , ρ sırasıyla Kendall'in tau'su ve Spearman'ın rho'su olsun.

$$\frac{3\tau-1}{2} \leq \rho \leq \frac{1-2\tau-\tau^2}{2}, \tau \geq 0$$

$$\frac{\tau^2+2\tau-1}{2} \leq \rho \leq \frac{3\tau+1}{2}, \tau \leq 0$$

Teorem 2.3 ve teorem 2.4'de ki τ ve ρ için sınırlar şekil 2.1'de gösterilmiştir.



Şekil 2. 1: Sürekli rasgele değişken çiftleri için τ ve ρ sınırları [1]

2.4. Kuyruk Bağımlılığı

Kuyruk bağımlılığı katsayısı sıra korelasyonu katsayıları gibi çift bağımlılık ölçümüdür. Sadece sürekli marjinal dağılım fonksiyonundan rasgele seçilen X ve Y çiftinin kopulasına bağlıdır. Başka bir deyişle iki değişkenli bir dağılımın kuyruklarındaki bağımlılık ölçütüdür [29]. X ve Y sürekli rasgele değişkenler, marjinal fonksiyonları sırasıyla F ve G olsun. Üst kuyruk bağımlılığı parametresi λ_U bir koşullu limit (istenilen limit varsa) olasılığıdır.

$$\lambda_U = \lim_{t \rightarrow 1^-} P[Y > G^{-1}(t) | X > F^{-1}(t)]$$

t değeri için limit 1'e yaklaşıyorsa, X 'in büyük olduğu bilindiğinde Y 'nin de büyük olması olasılığı üst kuyruk bağımlılığı katsayısını verir.

Benzer şekilde alt kuyruk bağımlılığı katsayısı λ_L de bir koşullu limit (istenilen limit varsa) olasılığıdır. t değerinin limiti 0'a yaklaşıyorsa, X 'in olasılığı eşit ya da küçük olduğu bilindiğinde Y 'nin de eşit ya da küçük olması olasılığıdır.

$$\lambda_L = \lim_{t \rightarrow 0^+} P[Y \leq G^{-1}(t) | X \leq F^{-1}(t)]$$

Üst ve alt kuyruk bağımlılığı parametreleri doğrusal değildir ve sadece X ve Y 'nin kopulasına bağlıdır [1].

3. KOPULA

Kopula rasgele bağımlı değişkenlerin modellenmesinde kullanılan önemli bir araçtır. Terim olarak bağlama, birleştirme, ilişkilendirme anlamına gelen latince ‘copulare’ kelimesinden türetilmiştir. Kopula’nın ana amacı bir çok rasgele değişkeni karşılıklı olarak ilişkilendirmektir [28].

M. Fréchet çok değişkenli dağılım fonksiyonları ile bunların daha düşük boyutlu marjinaleri hakkında soruları cevaplamaya çalışırken, kopula kavramı ilk olarak A. Sklar tarafından 1959 yılında tanıtılmıştır [1]. Kopulalar ile ilgili sonuçların çoğu olasılıksal metrik uzay çalışmalarının geliştirilmesi sırasında ve olasılık dağılım fonksiyonları üzerinde ikili işlemler çalışılırken elde edilmiştir. 1942 yılında Karl Menger metrik uzaylar teorisine olasılıksal bir genelleme getirmiştir [30]. Menger, $F_{pq}(x)$ dağılım fonksiyonunu herhangi bir gerçek x değeri için, p ile q arasındaki uzaklığın ($d(p, q)$) x den az olma olasılığı olarak tanımlamıştır [31]. Metrik uzayların olasılıksal yapısında karşılaşılan ilk zorluk üçgen eşitsizliğinin olasılıksal olarak sağlanıp sağlanamamasıdır. Menger bu durumda T bir üçgen norm (t -norm) olduğunda $F_{pr}(x+y) \geq T(F_{pq}(x), F_{qr}(y))$ eşitsizliğini önermiştir. Bazı t -normlar kopuladır buna karşılık bazı kopulalar ise t -normdur [31].

Kopula ile ilgili çalışmalar arttıkça her bir değişken kendi özelliklerinden ziyade bağımlılık yapısı altındaki özellikleri ile ilgi çekmeye başlamıştır. Bu durum literatürün bu alanda genişlemesini sağlamıştır. Rasgele değişkenler arasındaki bağımlılığın parametrik olmadığı durumlar için de kopulanın yararlı olabileceği keşfedilmiştir.

Kopula fonksiyonunu ilk olarak literatüre kazandıran A. Sklar tanımlamasında kopulayı şu şekilde anlatmıştır. Marjinaleri $[0,1]$ aralığında tek düze (uniform) dağılan iki değişkenli dağılım fonksiyonunun $[0,1]^2$ üzerinde kısıtlanmış halidir. Yani H , marjinaleri $F(x)$ ve $G(y)$ olan iki değişkenli dağılım fonksiyonu olsun. O zaman öyle bir C kopula vardır ve

$H(x, y) = C(F(x), G(y))$ sağlanır [31].

Kısaca kopula marjinaleri tek boyutlu olan ve tek düze dağılan dağılım fonksiyonlarıdır [1]. Kopula;

- Matematiksel olarak kolay inşa edilebilir olması,
- Rasgele değişkenlerin çok boyutlarında da kullanımı rahat olması,
- Bağımlılık yapılarının analizinde ve bağımlılık yapısının derecesi hakkında tutarlı bilgiler vermesi,

- Parametrik olan rasgele değişkenlerin yanı sıra yarı parametrik, parametrik olmayan veriler için asimetrik bağımlılık yapısı hakkında bilgi vermesi,
- Sürekli dönüşümler altında değişmez olması,

gibi özelliklerden dolayı tercih edilirler.

3.1. Kopula Tanımı ve Özellikleri

Kopula fonksiyonu bağımlılık fonksiyonları veya tek düze temsiller olarak da bilinir. Genel bir tanımlama olarak, X ve Y sürekli rasgele değişkenler olmak üzere $F(x)=P(X\leq x)$ ve $G(y)=P(Y\leq y)$ tek değişkenli dağılım fonksiyonları, $H(x,y)=P(X\leq x, Y\leq y)$ ortak dağılım fonksiyonu olsun. $[-\infty, \infty]^2$ 'de her (x,y) ikilisi için $[0,1]^3$ alanında $(F(x), G(y), H(x,y))$ noktalarını eşleştiren fonksiyona 'kopula fonksiyonu' denir. Temelde $[0,1]\times[0,1]$ tanım kümesinden aldığı her bir (x, y) rasgele değerlerini $[0,1]$ değer aralığında resmeder [32].

İki değişkenli durum için kopula fonksiyonu aşağıdaki şekilde tanımlanır.

$C: [0,1]^2 \rightarrow [0,1]$ kopula fonksiyonu olsun.

- $[0,1]$ aralığındaki her bir u, v için, $C(u,0) = 0 = C(0,v)$ ve $C(u,1) = u$ $C(1,v) = v$;
- $[0,1]$ aralığındaki her bir u_1, u_2, v_1, v_2 için $u_1 \leq u_2$ ve $v_1 \leq v_2$ olmak üzere

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0$$

özelliklerini sağlar [31].

$C, [0,1]\times[0,1]$ kümesinde tanımlı iki boyutlu kopulaya ikinci dereceden kopula denir. İki boyutlu kopula "subkopula" olarak adlandırılır. Yukarıda tanımlanan özelliklerin yanı sıra Subkopula Lipschitz koşulunu da sağlar.

$$|C(u_2, v_2) - C(u_1, v_1)| \leq |u_2 - u_1| + |v_2 - v_1| \quad (u_1, u_2), (v_1, v_2) \in [0,1]\times[0,1] \quad [33].$$

Lipschitz koşulu: $g(a, b)$ fonksiyonu (a_0, b_0) noktasını içine alan bir R bölgesinde tanımlı ve sürekli olsun. Bölge kapalı ve bir dikdörtgenle sınırlı olsun. R bölgesindeki bütün a, b_1 ve b_2 'ler için $|g(a, b_1) - g(a, b_2)| < L|b_1 - b_2|$ eşitsizliğini sağlayacak şekilde bir L varsa $g(a,b)$ fonksiyonu Lipschitz koşulunu sağlar.

C , bir kopula olsun. $(0,1)$ aralığındaki herhangi bir v için $\frac{\partial C(u,v)}{\partial u}$ kısmi türevi neredeyse tüm u 'lar için vardır ve $0 \leq \frac{\partial}{\partial u} C(u, v) \leq 1$ sağlanır. Benzer şekilde $(0,1)$ aralığındaki herhangi bir u için $\frac{\partial C(u,v)}{\partial v}$ kısmi türevi neredeyse tüm v 'ler için vardır ve $0 \leq \frac{\partial}{\partial v} C(u,v) \leq 1$ sağlanır.

$u \rightarrow \frac{\partial C(u,v)}{\partial u}$ ve $v \rightarrow \frac{\partial C(u,v)}{\partial v}$ fonksiyonları tanımlıdır ve (0,1) üzerinde neredeyse her yerde azalmayandır [1].

İkiden fazla değişkenli durum için kopula fonksiyonu ise şu şekilde tanımlanır.

$X = (X_1, X_2, \dots, X_n)$ dağılım fonksiyonu F ve marjinal dağılım fonksiyonları F_i , $(X_i \sim F_i)$, $1 \leq i \leq n$ olan bir rasgele vektör olsun. $[0,1]$ üzerinde tek düze marjinalleri olan C dağılım fonksiyonu eğer, $F = C(F_1, \dots, F_n)$ şeklinde yazılabiliyorsa C 'ye X 'in kopula fonksiyonu denir. Kopula kümesi; $\mathcal{U}(t) = t$, $t \in [0,1]$ tek düze marjinal dağılım fonksiyonları olan bütün C dağılım fonksiyonlarının $\mathcal{F}(\mathcal{U}, \dots, \mathcal{U})$ Fréchet sınıfına özdeştir. F 'in marjinal fonksiyonlarının sürekli olması durumunda kopulanın tanımlanması da kolaydır.

C 'yi $(F_1(x_1), \dots, F_n(x_n))$ 'nin dağılım fonksiyonu olarak tanımlayalım. $F_i \sim U(0,1)$ tek düze dağılımlı olduğu için C bir kopuladır. Ayrıca aşağıdaki şekilde de yazılabilir.

$$\begin{aligned} C(u_1, \dots, u_n) &= P(F_1(X_1) \leq u_1, \dots, F_n(X_n) \leq u_n) \\ &= P(X_1 \leq F_1^{-1}(u_1), \dots, X_n \leq F_n^{-1}(u_n)) \\ &= F_X(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)) \end{aligned}$$

F_i^{-1} fonksiyonu F_i 'nin ters fonksiyonudur [34]. Bu fonksiyona yüzdelik dönüşüm denir. $F_i^{-1}(t) = \inf \{x \in R; F_i(x) \geq t\}$ şeklinde tanımlanır. F , C tarafından tanımlandığı için C F 'in bir kopulasıdır.

$$\begin{aligned} F(x_1, x_2, \dots, x_n) &= P(X_1 \leq x_1, \dots, X_n \leq x_n) \\ &= P(F_1(X_1) \leq F_1(x_1), \dots, F_n(X_n) \leq F_n(x_n)) \\ &= C(F_1(x_1), \dots, F_n(x_n)) \end{aligned} \quad (3.1)$$

Kopula yapısının kurulması için argüman, F_i sürekli dağılım fonksiyonlarının $F_i(X_i)$, $(0,1)$ aralığında tek düze dağılan fonksiyon $(F_i(X_i) \sim U(0,1))$ olma özelliğine dayanmalıdır. Eşitlik 3.1'de tanımlanan dönüşüme dağılımsal dönüşüm (distributional transform) denir.

Dağılımsal dönüşüm: F dağılım fonksiyonun bir gerçek rasgele değişkeni Y ; V , Y 'den bağımsız rasgele bir değişken ve V , $(0,1)$ aralığında tek düze dağılım. ($V \sim U(0,1)$). Modifiye edilmiş dağılım fonksiyonu $F(x, \lambda)$; $F(x, \lambda) := P(X < x) + \lambda P(Y = x)$ şeklinde tanımlanır. $U := F(Y, V)$ fonksiyonuna Y 'nin genelleştirilmiş dağılımsal dönüşümü denir. Eğer F fonksiyonu sürekli dağılım fonksiyonu ise tüm λ 'lar için $F(x, \lambda) = F(x)$ ve

$U = F(Y) =^d U(0,1)$ dağılımdaki eşitliği gösterir. Dağılımsal dönüşüm kopulanın anlaşılması için çok önemli olan Sklar teoremi'nin kanıtında kullanılan ve kanıtı kolaylaştıran dönüşümdür [34].

Bağımsız kopula: $U = (U_1, U_2, \dots, U_d)$ her bir bileşenleri birbirinden bağımsız $(0,1)$ üzerinde tek düze dağılan rasgele bir vektör olmak üzere d boyutlu kopula

$$\Pi_d(u) = u_1 u_2 \dots u_d$$

şeklinde yazılır.

Comonotonik kopula: $U = (U_1, U_2, \dots, U_d)$ vektörü $(0,1)$ üzerinde tek düze dağılan rasgele bir vektör ve $U_1 = U_2 = \dots = U_d$ olmak üzere d boyutlu comonotonik kopula

$$M_d(u) = \min \{u_1, u_2, \dots, u_d\}$$

şeklindedir.

The countermonotonik kopula: $U = (U_1, U_2)$ vektörü $(0,1)$ üzerinde tek düze dağılan rasgele bir vektör $U_1 = 1 - U_2$ olmak üzere iki boyutlu countermonotonik kopula

$$W_2(u) = \max \{u_1 + u_2 - 1, 0\}$$

şeklindedir [35].

Comonotonik kopula aynı zamanda bir Fréchet-Hoeffding üst sınırıdır. Ayrıca tam pozitif bağımlı olma durumunu ifade eder. Countermonotonik kopula sadece iki boyutlu durumda bulunabilir. Bu kopula bir Fréchet-Hoeffding alt sınırıdır ve tam negatif bağımlı olma durumunu ifade eder. Bu iki kopula bir kopula dağılım fonksiyonu değildir. İkisinin de bükülme noktaları vardır ve bu nedenle türevlenemezler. Örneğin, $U_1 = U_2$ comonotonik kopula olacak şekilde U_1, U_2 alalım. $u_1 = u_2$ olması durumu kopulanın tek bir noktada yoğunlaşması demektir. Aynı durumda countermonotoniklik için $u_1 = 1 - u_2$ tek bir değerdedir. Bu iki kopula bu yüzden bir dağılım fonksiyonu değildir [28].

3.2. Sklar Teoremi

H , marjinalleri F ve G olan bir ortak dağılım fonksiyonu olsun. $\mathbb{R}$ 'den seçilen bütün x, y 'ler için bir C kopulası vardır ve $H(x, y) = C(F(x), G(y))$ 'dir.

Eğer $[0,1]$ 'de tanımlı F ve $[0,1]$ de tanımlı G sürekli fonksiyonlar ise C kopulası tektir. Diğer taraftan C kopulasının tek olması onun $[0,1] \times [0,1]$ 'de tanımlı olmasıdır. Ters durumda, C bir kopula ve F ve G dağılım fonksiyonları olsun. O zaman H marjinalleri F ve

G olan bir ortak dağılım fonksiyonudur. F bir dağılım fonksiyonu olsun. F 'in yarı tersi $F^{(-1)}$ fonksiyonunun tanım kümesi $(0, 1)$ dir. Aşağıdaki özellikleri sağlar.

1. Eğer t değeri F 'in tanım bölgesi $[0,1]$ 'de ise, $F^{(-1)}(t) \in \mathbb{R}$ 'de bir x için $F(x) = t$ 'dir. Ayrıca F 'nin tanım bölgesi $[0,1]$ 'deki bütün t değerleri için $F(F^{(-1)}(t)) = t$ 'dir.
2. Eğer t değeri $[0,1]$ 'de değil ise, $F^{(-1)}(t) = \inf \{x : F(x) \geq t\} = \sup \{x : F(x) \leq t\}$

Eğer F kesin artan fonksiyon ise tek bir yarı terse sahiptir ve bu F 'in normal tersidir. Notasyon olarak F^{-1} kullanılır.

H , marjinalleri F ve G olan ortak dağılım fonksiyonu C' bir subkopula olsun. $F^{(-1)}$ ve $G^{(-1)}$ sırasıyla F ve G 'nin yarı tersi olmak üzere, C' 'nin tanım kümesindeki her (u,v) ikilisi için $C'(u, v) = H(F^{(-1)}(u), G^{(-1)}(v))$ 'dir. F ve G sürekli olduğu zaman bu özellik kopuların ortak dağılım fonksiyonlarının oluşturulmasında kullanılır.

Teorem 3.1: X ve Z sırasıyla F ve G dağılım fonksiyonların rasgele değişkenleri, H ortak dağılım fonksiyonları olsun. Bu durumda bir C kopulası vardır ve subkopuladır. Eğer F ve G $[0,1]$ 'de tanımlı sürekli fonksiyonlar ise, C tektir. Diğer taraftan C kopulası $[0,1] \times [0,1]$ 'de tanımlıdır.

Teorem 2.1'deki C kopulası X ve Z 'nin kopulası olarak adlandırılır. C_{XZ} olarak ifade edilir.

Teorem 3.2: X ve Z sürekli rasgele değişkenler olsun. X ve Z bağımsızdır ancak ve ancak

$$C_{XZ} = \Pi.$$

Sklar teoremi çok boyutlu olarak da yazılabilir. H marjinalleri $F_1, F_2, \dots, F_n$ olan n boyutlu dağılım fonksiyonu olsun. $\mathbb{R}^n$ 'de ki her x elemanı için bir n -boyutlu C kopulası vardır ve $H(x_1, x_2, \dots, x_n) = C(F_1(x_1), F_2(x_2), \dots, F_n(x_n))$ 'dir.

Her biri $[0,1]$ 'de tanımlı tüm $F_1, F_2, \dots, F_n$ 'lerin sürekli olması durumunda C kopulası tektir. Diğer taraftan C tek ve $[0,1]^n$ 'de tanımlıdır. Tersine C bir n -kopula ve $F_1, F_2, \dots, F_n$ 'ler dağılım fonksiyonları olsun. Bu durumda H marjinalleri $F_1, F_2, \dots, F_n$ olan n -boyutlu bir dağılım fonksiyonudur.

$F_1^{(-1)}, F_2^{(-1)}, \dots, F_n^{(-1)}$ ler sırasıyla $F_1, F_2, \dots, F_n$ 'lerin yarı-ters fonksiyonları olsun. $(0,1)^n$ 'de ki her u için,

$$C(u_1, u_2, \dots, u_n) = H(F_1^{(-1)}(u_1), F_2^{(-1)}(u_2), \dots, F_n^{(-1)}(u_n))$$

sağlanır.

3.3. Fréchet-Hoeffding Sınırları

Her bir kopula C için ve (0,1) tanım aralığında her u, v'ler için

$$W(u, v) = \max(u + v - 1, 0) \leq C(u, v) \leq \min(u, v) = M(u, v)$$

sınırlarına Fréchet-Hoeffding sınırları denir. Sklar teoreminin bir sonucu olarak, X ve Y rasgele değişkenlerin ortak marjinal dağılım fonksiyonu H, marjinalleri sırasıyla F ve G olsun. R'de ki bütün x, y'ler için $\max(F(x) + G(y) - 1, 0) \leq H(x, y) \leq \min(F(x), G(y))$ eşitsizliğinde sınırlar marjinalleri F ve G olan H ortak dağılım fonksiyonu için Fréchet-Hoeffding sınırları olarak adlandırılır [1].

C_{XZ} kopula ve X, Z sürekli rasgele değişkenler olsun. Sırasıyla f, g fonksiyonlarının X'in ve Z'nin tanım bölgesinde kesin monoton fonksiyonlar olduğu aşağıda gösterilmiştir.

1. Eğer f ve g kesin artan fonksiyonlarsa, $C_{f(X),g(Z)}(u, v) = C_{XZ}(u, v)$.
2. Eğer f kesin artan ve g kesin azalan fonksiyonlarsa, $C_{f(X),g(Z)}(u, v) = u - C_{XZ}(u, 1-v)$.
3. Eğer f kesin azalan ve g kesin artan fonksiyonlarsa, $C_{f(X),g(Z)}(u, v) = v - C_{XZ}(1-u, v)$.
4. Eğer f ve g kesin azalan fonksiyonlarsa, $C_{f(X),g(Z)}(u, v) = u + v - 1 + C_{XZ}(1-u, 1-v)$ [31].

3.4. Yaşam Kopulası

Bir çok uygulamada rasgele değişkenler bir popülasyondaki bireylerin veya nesnelerin yaşam süresini temsil eder. Bir bireyin x zamanın ötesinde yaşama ya da hayatta kalma olasılığını veren fonksiyon yaşam fonksiyonlarıdır (Güvenilirlik fonksiyonu). F, X'in dağılım fonksiyonu olmak üzere $\bar{F}(x) = P[X > x] = 1 - F(x)$ yaşam fonksiyonudur. Yaşam süresi ile ilgilenildiği zaman rasgele değişkenlerin tanımlı olduğu aralık $(0, \infty)$ 'dir. Ancak tanım aralığı $\mathbb{R}$ olduğunda bile $P[X > x]$ yaşam fonksiyonu olarak kullanılabilir.

(X, Y) rasgele değişkenlerin herhangi bir çifti ve S ortak dağılım fonksiyonu olsun. Ortak yaşam fonksiyonu $\bar{S}(x, y) = P[X > x, Y > y]$ 'dir. $\bar{S}$ 'nin marjinalleri $\bar{S}(x, -\infty)$, $\bar{S}(-\infty, y)$ ve tek değişkenli yaşam fonksiyonları sırasıyla $\bar{F}$, $\bar{G}$ olsun. X ve Y'nin kopulası C,

$$\begin{aligned}\bar{S}(x, y) &= 1 - F(x) - G(y) + S(x, y) \\ &= \bar{F}(x) + \bar{G}(y) - 1 + S(F(x), G(y)) \\ &= \bar{F}(x) + \bar{G}(y) - 1 + S(1 - \bar{F}(x), 1 - \bar{G}(y))\end{aligned}$$

biçiminde ifade edilir.

$$\hat{S} : (0,1)^2 \rightarrow (0,1) \text{ olarak alındığında } \hat{S}(u, v) = u + v - 1 + S(1-u, 1-v)$$

$\bar{S}(x,y) = \hat{S}(\bar{F}(x), \bar{G}(y))$ eşitliği elde edilir. Eşitlikte olduğu gibi $\hat{S}$ bir kopuladır aynı zamanda X ve Y 'nin yaşam kopulasıdır. $\hat{S}$ çiftlerin ortak yaşam fonksiyonuyla tek değişkenli marjinallerini bağlar. Aynı zamanda ortak dağılım fonksiyonlarını da marjinallerine bağlar. $\hat{S}$ yaşam kopulası ile $\bar{S}$ ortak yaşam fonksiyonu karıştırmamalıdır. $\bar{S}$ rasgele iki değişkeni (0,1) aralığında tek düze dağılan ve ortak dağılım fonksiyonu S olan yaşam fonksiyonudur. Aralarındaki ilişki,

$$\bar{S}(u, v) = P[U > u, V > v] = 1 - u - v + S(u, v) = \hat{S}(1 - u, 1 - v)$$

biçimindedir [1].

3.5. Parametrik Kopulalar

Kopulalar bu bölümde eliptik kopula ve arşimedyan kopula aileleri ile sınıflandırılmıştır. Eliptik kopulalardan Gaussian kopula ve T-kopula, Arşimedyan kopula ailesinden Clayton kopula, Gumbel kopula ve Frank kopula aileleri ele alınmıştır.

3.5.1. Eliptik Kopula

Eliptik dağılımlar çok değişkenli aşırılıkların modellenme, çok değişkenli normal dağılımların özelliklerini paylaşabilme ve de normal olmayan diğer bağımlılık formlarını içeren çok değişkenli dağılımlar açısından zengin bir kaynaktır. Eliptik kopulalar eliptik dağılımlardan elde edilir. Sklar teoremi'nin de sonucu olduğu gibi eliptik kopulalar da benzetim çalışması yapılması kolaydır [9].

Bu bölümde önemli eliptik kopula ailelerinden olan Gaussian kopula ve t-kopuladan bahsedilmiştir.

3.5.1.1. Gaussian Kopula

Gaussian kopula standart çok değişkenli normal dağılımdan elde edilmiştir.

ϕ_m m değişkenli standart normal dağılım fonksiyonu ve ϕ^{-1} ters fonksiyonu olmak üzere ϕ standart normal dağılım fonksiyonu olsun. Gaussian kopula aşağıdaki formdadır.

$$C_{\rho}^{Ga}(u_1, u_2, \dots, u_m) = \phi_m(\phi_{(u_1)}^{-1}, \phi_{(u_2)}^{-1}, \dots, \phi_{(u_m)}^{-1})$$

Rasgele değişkenlerin her bir çifti için, tahmin edilen τ (tau) ($-1 \leq \tau \leq 1$) katsayısı hesaplanır ve $r = \sin(\pi\tau/2)$ formülüyle hesaplanır. Tüm r 'ler için Gaussian kopulanın yayılım matrisindeki bileşenleri kullanılır. Bu yayılım matrisi bağımlılık altındaki parametrelerden oluşur. Yayılım matrisi pozitif tanımlı ve simetrik olmalıdır. İlişki katsayısı ρ (rho) her bir rasgele değişken çifti için doğrudan hesaplanamaz. Nedeni ise

sadece lineer bağımlılığı ölçebilir olması ve kullanılan marjinallerin eliptik olmamasındandır. Dolayısıyla r 'yi bulmak için ρ kullanılır.

Yayılım matrisini oluşturduktan sonra, m değişkenli Gaussian kopulaya ait veriler aşağıdaki adımlar kullanarak benzetim yapılır. Yayılım matrisi pozitif tanımlı ve simetriktir.

1. $Z_1, Z_2, \dots, Z_m$ m değişkenli standart normal dağılımdan elde edilen örneklem (doğrusal korelasyon matrisi gibi yayılım matrisi)
2. Her $i=1,2, \dots, m$ değeri için $V_i = \phi(Z_i)$ hesaplanır.
3. Her $i=1,2, \dots, m$ değeri için $X_i = F_{X_i}^{-1}(V_i)$ hesaplanır.

3.5.1.2. T-Kopula

T-kopula çok değişkenli t dağılımından elde edilir. Bu kopula Student t dışındaki marjinal dağılımlarla birleştirilebilir. Serbestlik derecesi ν , t_v^{-1} ters fonksiyonu olmak üzere t_v tek değişkenli Student t dağılım fonksiyonu ve $t_{v,m}$ m -değişkenli t dağılım fonksiyonu olsun. T-kopula aşağıdaki formdadır.

$$C(u_1, u_2, \dots, u_m) = t_{v,m}(t_v^{-1}(u_1), t_v^{-1}(u_2), \dots, t_v^{-1}(u_m))$$

T-kopula'nın Gaussian kopula ile uyumluluğu benzerdir. Nedeni ise hem çok değişkenli t dağılımının hem de çok değişkenli normal dağılımın eliptik dağılım olmasındandır. Serbestlik derecesi arttıkça da t -kopula Gaussian kopulaya yakınsar.

Rasgele değişkenlerin her bir çifti için tahmin edilen τ (tau) ($-1 \leq \tau \leq 1$) ve $r = \sin(\pi\tau/2)$ formülüyle hesaplanır. τ , t -kopulanın yayılım matrisini oluşturmak için kullanılır. ν parametresi kuyruk bağımlılığındaki genel tahminle tanımlanır veya ν 'nün farklı deneysel değerleri tahmin edilir.

Yayılım matrisini oluşturduktan ve ν 'nün değerine karar verdikten sonra; m değişkenli t -kopula ya ait veriler aşağıdaki adımlara göre simülasyon yapılır.

1. $Z_1, Z_2, \dots, Z_m$ m değişkenli standart normal dağılımdan elde edilen örneklem (doğrusal korelasyon matrisi gibi yayılım matrisi)
2. ν serbestlik derecesi için örneklemin Ki-Kare test istatistik değeri bulunur. ($S \sim \chi_\nu^2$)
3. Bütün i değerleri için $V_i = t_v(Z_i \sqrt{\nu/S})$ hesaplanır.
4. Her $i=1,2, \dots, m$ için $X_i = F_{X_i}^{-1}(V_i)$ hesaplanır.

T-kopula her bir çift rasgele değişkenler için aynı yüksek ve düşük kuyruk bağımlılığına sahiptir.

$$\lambda_{yüksek} = \lambda_{düşük} = 2 - 2 t_{v+1}(\sqrt{(v+1)(1-r)(1+r)})$$

eşitiyle bulunur. Serbestlik derecesi (v) sonsuza gittikçe bu değer sıfıra eşit olur.

τ (tau)'nın sıfır ya da negatif olması serbestlik derecesi v 'nün küçük değerleri için kuyruk bağımlılığı yüksektir.

Genel sigorta yükümlülüklerinde T kopula modellenmesinin iki dezavantajı vardır. Birincisi Gaussian kopula gibi T kopula simetriktir. Yani çok değişkenli bir örneklemede her bir çift rasgele değişken için kuyruk bağımlılığının en yüksek değeri ile en düşük değeri birbirine eşittir. ($\lambda_{yüksek} = \lambda_{düşük}$)

İkincisi T kopula rasgele değişkenlerin her biri için yalnızca bir parametre (v) vardır. Parametre (v) tüm rasgele değişken çiftleri için kuyruk bağımlılığının belirlenmesinde yer alır bu nedenle modelleme esnekliği daha azdır [36].

3.5.2. Arşimedian Kopula

Arşimedian kopulalar çeşitli nedenlerden dolayı istatistik, matematik, sigorta ve daha birçok alanda kullanılmaktadır. Bu nedenlerden bazıları bu kopulanın kolay inşa edilebilir olması, bu sınıfa ait oldukça geniş ailelerin olması ve de bu kopulanın yararlı özelliklere sahip olmasıdır. Bu bölümde arşimedian kopulalar iki değişkenli ve çok değişkenli olma durumlarına göre incelenmiştir.

3.5.2.1. İki Değişkenli Arşimedian Kopula

Arşimedian kopula üretici fonksiyonu ile tanımlanır.

φ üretici fonksiyonu, φ^{-1} tersi olmak üzere iki değişkenli Arşimedian kopula,

$$C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v))$$

formundadır.

$\varphi : [0,1] \rightarrow [0,\infty]$, $\varphi(1) = 0$ ve $\varphi(0) = \infty$ üretici fonksiyonudur.

Üretici fonksiyonun $\varphi(0) = \infty$ olmaması durumunda φ 'nin tersi yerine **sözde tersi** fonksiyonu kullanılır.

$\varphi : [0,1] \rightarrow [0,\infty]$ sürekli, kesin azalan fonksiyon $\varphi(1) = 0$, $\varphi(0) \leq \infty$ olmak üzere φ 'nin **sözde tersi** fonksiyonu $\varphi^{[-1]}$ aşağıdaki formdadır.

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t) & 0 \leq t \leq \varphi(0), \\ 0 & \varphi(0) \leq t \leq \infty, \end{cases}$$

$\varphi^{[-1]}$, $[0, \infty]$ aralığında sürekli, artmayan, $[0, \varphi(0)]$ aralığında kesin azalan,

$\text{Dom } \varphi^{[-1]} = [0, \infty]$, $\text{Ran } \varphi^{[-1]} = [0, 1]$ özelliklerini sağlayan fonksiyondur.

Ayrıca $[0, 1]$ aralığı üzerinde $\varphi^{[-1]}(\varphi(u)) = u$ ve

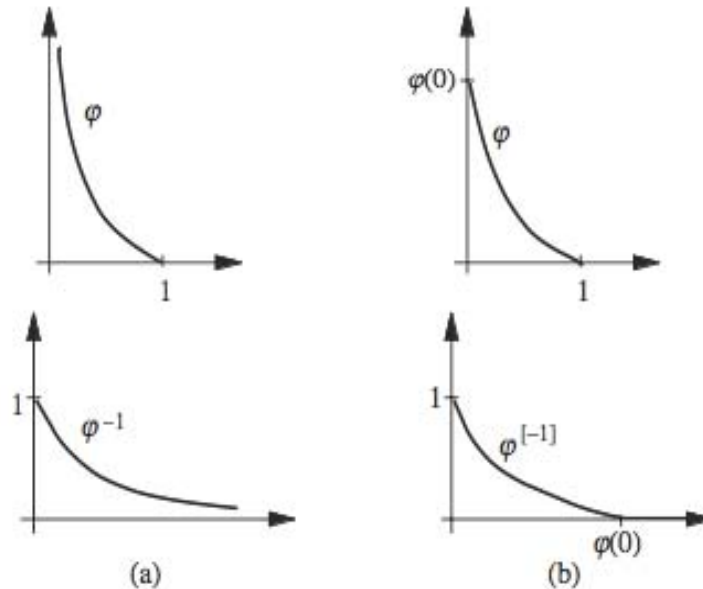
$$\begin{aligned} \varphi(\varphi^{[-1]}(t)) &= \begin{cases} t & 0 \leq t \leq \varphi(0), \\ \varphi(0) & \varphi(0) \leq t \leq \infty, \end{cases} \\ &= \min(t, \varphi(0)) \end{aligned}$$

Sonuç olarak $\varphi(0) = \infty$ ise $\varphi^{[-1]} = \varphi^{-1}$ 'dir.

$C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v))$ formu Arşimedyen kopula olarak adlandırılır. φ üretici fonksiyonu olmak üzere eğer $\varphi(0) = \infty$ eşitliği sağlanıyorsa bu üretici mutlak, tam üreticidir (strict generator).

C , φ üretici fonksiyonu olan Arşimedyen kopula olsun.

1. $[0, 1]$ aralığındaki tüm u, v 'ler için C simetriktir. $C(u, v) = C(v, u)$
2. $[0, 1]$ aralığındaki tüm u, v, w 'lar için birleştirilebilirdir.
 $C(C(u, v), w) = C(u, C(v, w))$
3. Her $c > 0$ sabit sayısı için $c\varphi$ 'de kopula üreticisidir [1].



Şekil 3. 1: Tam (a) ve tam olmayan (b) üreticiler ve tersleri [1]

3.5.2.2. Bazı Önemli Arşimedyen Kopula Aileleri

Literatür incelendiği zaman bir çok arşimedyen kopula ailesi olduğu görülmektedir. Ancak bu çalışmada sadece çalışmamıza konu olan arşimedyen kopula aileleri ele alınmıştır.

3.5.2.2.1. Gumbel Kopula

Üretici fonksiyonu; $\varphi_{Gu}(u) = (-\ln u)^\theta$, $\theta \in [1, \infty)$ olmak üzere Gumbel kopula,

$$C_{\theta}^{Gu}(u_1, u_2) = \exp [- ((-\ln u_1)^\theta + (-\ln u_2)^\theta)]^{\frac{1}{\theta}}$$

formundadır.

$\theta = 1$ olduğunda bağımsız kopula elde edilirken,

$\theta \rightarrow \infty$ (sonsuzaya yaklaşıırken) comonotonik kopulaya eğilimlidir.

Gumbel kopula bağımsız ve pozitif bağımlılık arasında anlam verir.

3.5.2.2.2. Clayton Kopula

Üretici fonksiyonu; $\varphi_{Cl}(u) = \frac{1}{\theta} (u^{-\theta} - 1)$, $\theta \in [1, \infty) \setminus \{0\}$ olmak üzere Clayton kopula,

$$C_{\theta}^{Cl}(u_1, u_2) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1}{\theta}}$$

formundadır.

$\theta \rightarrow 0$ (limiti sıfıra yaklaştıkça) bağımsız kopula elde edilirken,

$\theta \rightarrow \infty$ (sonsuzaya yaklaşıırken) comonotonik kopulaya ulaşır.

$\theta = -1$ olduğunda Fréchet Hoeffding alt sınırına eşit olur.

Gumbel kopula gibi Clayton kopula da tam bağımlı yapılar (countermonotonik), bağımsız ve comonotonik arasında anlamlandırılır.

3.5.2.2.1. Frank Kopula

Üretici fonksiyonu; $\varphi_{Fr}(u) = \ln \frac{e^{-\theta} - 1}{e^{-\theta u} - 1}$, $\theta \in \mathbb{R} \setminus \{0\}$ olmak üzere Frank kopula,

$$C_{\theta}^{Fr}(u_1, u_2) = -\frac{1}{\theta} \ln \left(1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1} \right)$$

formundadır [28].

Çizelge 3. 1: Kendall’ın sıra korelasyon katsayısı ve kopula aileleri için kuyruk bağımlılık katsayıları [29]

Kopula	τ	λ_U	λ_L
C_θ^{Cl}	$\theta/(\theta + 2)$	0	$\begin{cases} 2^{-1/\theta}, & \theta \geq 0, \\ 0, & \theta < 0, \end{cases}$
C_θ^{Fr}	$1 - 4\theta^{-1}(1 - \mathcal{D}_1(\theta))$	0	0
C_θ^{Gu}	$1 - 1/\theta$	$2 - 2^{1/\theta}$	0

$\mathcal{D}_1(\theta)$ Debye fonksiyonudur. $\mathcal{D}_1(\theta) = \theta^{-1} \int_0^\theta t/(\exp(t) - 1) dt$ [29].

3.5.2.3. Çok Değişkenli Arşimedyen Kopula

$\varphi : [0,1] \rightarrow [0,\infty]$ kesin azalan, sürekli, Arşimedyen kopulanın üretici fonksiyonu, $\varphi(0)=\infty$ ve $\varphi(1)=0$,

φ^{-1} , φ fonksiyonunun sözde tersi olsun. Çok değişkenli arşimedyen kopula aşağıdaki formdadır.

$$C^n(u) = \varphi^{-1}(\varphi(u_1) + \varphi(u_2) + \dots + \varphi(u_n))$$

Bu kopula n-kopula (n boyutlu) olarak adlandırılır.

C^n , n boyutlu kopuladır ancak ve ancak φ^{-1} , $[0,\infty)$ aralığında kesin monotondur.

$(-1)^k \frac{d^k}{dt^k} \varphi^{-1}(t) \geq 0$ her $t \in [0,\infty)$ için integrallenebilirdir, $k=0,1,2,\dots$

Çok değişkenli Arşimedyen kopula ailelerinden bazıları:

n-boyutlu Clayton kopula,

$$C_\theta^{Cl}(u) = (u_1^{-\theta} + \dots + u_n^{-\theta} - n + 1)^{-1/\theta}, \theta > 0$$

n-boyutlu Frank kopula,

$$C_\theta^{Fr}(u) = -\frac{1}{\theta} \ln \left(1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1) \dots (e^{-\theta u_n} - 1)}{(e^{-\theta} - 1)^{n-1}} \right), \theta > 0$$

Yukarıdaki formlarda ifade edilebilir [37].

3.6. Parametrik Olmayan Kopula

Parametrik olmayan kopula modelleri deneysel (empirical) kopula modelleri anlamına gelir. Bu kısımda deneysel kopula ve kopula yaklaşımlarından Bernstein polinomları yaklaşımı ve Grid-type kopula yaklaşımı anlatılmıştır.

3.6.1. Deneyisel (Empirical) Kopula

$\{(x_k, y_k)_{k=1}^n$ sürekli çift değişkenli bir dağılımdan elde edilen n boyutlu bir örnekleme gösterebilir. C_n deneyisel kopula fonksiyonu;

$$C_n\left(\frac{i}{n}, \frac{j}{n}\right) = \frac{x \leq x_i, y \leq y_j \text{ ile örneklemdaki } (x, y) \text{ çiftlerinin sayısı}}{n}$$

eşitliğinden elde edilir. x_i ve y_j , $1 \leq i, j \leq n$, örneklemdaki sıralı istatistikleri belirtir. Deneyisel kopula sıklığı c_n ;

$$c_n\left(\frac{i}{n}, \frac{j}{n}\right) = \begin{cases} \frac{1}{n}, & (x_i, y_j) \text{ örneklemden ise,} \\ 0, & \text{diğer durumlar için.} \end{cases}$$

biçiminde tanımlanır. C_n ile c_n arasındaki ilişki $C_n\left(\frac{i}{n}, \frac{j}{n}\right) = \sum_{p=1}^i \sum_{q=1}^j c_n\left(\frac{p}{n}, \frac{q}{n}\right)$ dir.

$$c_n\left(\frac{i}{n}, \frac{j}{n}\right) = C_n\left(\frac{i}{n}, \frac{j}{n}\right) - C_n\left(\frac{i-1}{n}, \frac{j}{n}\right) - C_n\left(\frac{i}{n}, \frac{j-1}{n}\right) + C_n\left(\frac{i-1}{n}, \frac{j-1}{n}\right).$$

Deneyisel kopula ilk olarak 1979 yılında Deheuvels tarafından çalışılmış ve tanıtılmıştır. Deneyisel kopulayı bir bağımlılık fonksiyonu olarak literatüre kazandırmıştır [1].

3.6.2. Kopula Yaklaşımları

Kopula yaklaşımları altta yatan bağımlılık yapısını parametrik olmayan bir şekilde tahmin etmek için kullanılır. Bu yaklaşımlar gerçek bağımlılık yapısına daha güçlü yaklaşmayı sağlar. Kopula yaklaşımları Bernstein polinomları yaklaşımı ve Grid-type kopula yaklaşımı olarak bu bölümde anlatılmıştır.

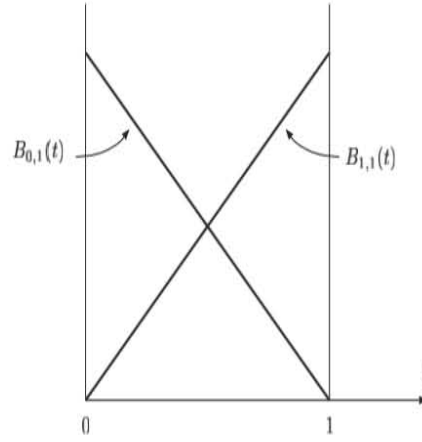
3.6.2.1. Bernstein Polinomları Yaklaşımı

Bernstein kopula, Bernstein polinomları yaklaşımıyla üretilmiştir. Polinomlar tanımlaması basit, bilgisayar sisteminde inşası oldukça hızlı ve de çeşitli fonksiyonları temsil etmede oldukça kullanışlı matematiksel araçlardır. Türevlenebilmesi ve integralinin alınması kolaydır. En genel haliyle $p(t) = a_n t^n + a_{n-1} t^{n-1} + \dots + a_1 t + a_0$ şeklindedir. Bir polinom $\{(1, t, t^2, \dots, t^n)\}$ vektörünün doğrusal kombinasyonu olarak gösterilmiş halidir. Derecesi n 'e eşit ya da n 'den küçük olan polinom kümesi bir vektör uzayı oluşturur. $\{(1, t, t^2, \dots, t^n)\}$ vektörü vektör uzayının tabanı, bazı olarak adlandırılır. Yani derecesi n 'e eşit ya da n 'den küçük olan bir polinom bu tabanın (bazın) doğrusal kombinasyonu olarak tek bir şekilde yazılabilir. Bir n 'inci dereceden Bernstein polinomu,

$$\binom{n}{i} = \frac{n!}{i!(n-i)!}, i = 0, 1, \dots, n \text{ olmak üzere;}$$

$B_{(i,n)}(t) = \binom{n}{i} t^i (1-t)^{n-i}$ biçiminde tanımlanır. Tanımlanan ifadeye $n+1$ tane n 'inci dereceden Bernstein polinomu denir. Matematiksel kolaylık olarak $i < 0$ ya da $i > n$ olması durumunda $B_{(i,n)} = 0$ kabul edilir. $\binom{n}{i}$ katsayısı Pascal üçgeni'nden elde edilir. Öncelikle geometrik anlamda Bernstein kopula incelenecektir.

Bernstein polinomunun derecesi 1 iken $B_{0,1}(t) = 1-t$, $B_{1,1}(t) = t$ 'dir. $0 \leq t \leq 1$ için;

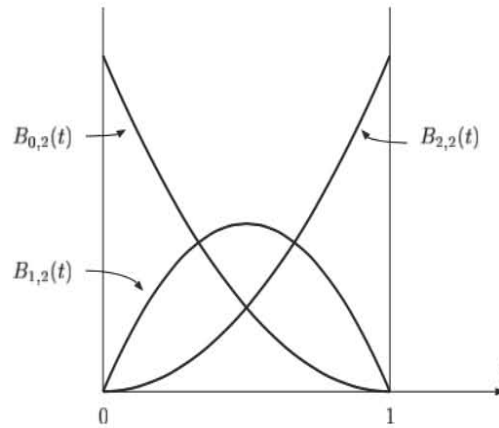


Şekil 3. 2: $B_{0,1}(t)$, $B_{1,1}(t)$ Bernstein polinomu [38]

Bernstein polinomunun derecesi 2 iken $B_{0,2}(t) = (1-t)^2$, $B_{1,2}(t) = 2t(1-t)$,

$B_{2,2}(t) = t^2$ 'dir.

$0 \leq t \leq 1$ için;

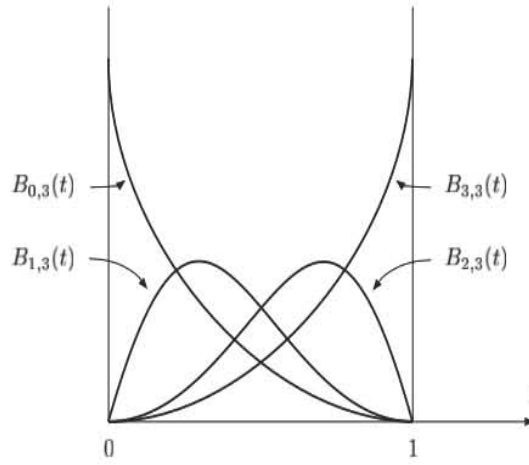


Şekil 3. 3: $B_{0,2}(t)$, $B_{1,2}(t)$, $B_{2,2}(t)$ Bernstein polinomları [38]

Bernstein polinomunun derecesi 3 iken $B_{0,3}(t) = (1-t)^3$, $B_{1,3}(t) = 3t(1-t)^2$,

$B_{2,3}(t) = 3t^2(1-t)$, $B_{3,3}(t) = t^3$ 'dür.

$0 \leq t \leq 1$ için;



Şekil 3. 4: $B_{0,3}(t)$, $B_{1,3}(t)$, $B_{2,3}(t)$, $B_{3,3}(t)$ Bernstein polinomları [38].

Derecesi n olan Bernstein polinomları derecesi $n-1$ olan iki Bernstein polinomu ile yazılabilir. Bu özellik Bernstein polinomlarının yinelemeli özelliği olarak bilinir.

$$B_{k,n}(t) = (1-t) B_{k,n-1}(t) + t B_{k-1,n-1}(t)$$

$t \in [a,b]$ için $f(t) \geq 0$ ise bir $f(t)$ fonksiyonu $[a,b]$ aralığı üzerinde negatif fonksiyon değildir. Bu durum n 'inci dereceden bir Bernstein polinomu için de geçerlidir. Yani $[0,1]$ aralığı üzerinde negatif değildirler. $0 < t < 1$ aralığında bütün Bernstein polinomları pozitifdir.

Derecesi k olan $k+1$ tane Bernstein polinomunun toplamı derecesi $k-1$ olan k tane Bernstein polinomunun toplamına eşittir.

$$\sum_{i=0}^k B_{i,k}(t) = \sum_{i=0}^{k-1} B_{i,k-1}(t)$$

n 'inci dereceden Bernstein polinomunun türevi $n-1$ 'inci dereceden Bernstein polinomudur.

$$\frac{d}{dt} B_{k,n}(t) = n (B_{k-1,n-1}(t) - B_{k,n-1}(t))$$

Bernstein polinomları polinomlar uzayında geniş yer kaplarlar. Derecesi n 'den küçük ya da n 'e eşit olan herhangi bir polinom Bernstein polinomunun doğrusal kombinasyonu olarak yazılabilir. Bir diğer özellikleri, doğrusal bağımsızdırlar. Yani $c_0, c_1, \dots, c_n$ sabit katsayıları varsa $0 = c_0 B_{0,n}(t) + c_1 B_{1,n}(t) + \dots + c_n B_{n,n}(t)$ özdeşliği bütün t 'ler için yazılır ve her bir c_i 'nin değeri 0 olmalıdır. Bernstein temel polinomlarının doğrusal bileşimi olarak yazılan bir polinom ele alındığında,

$B(t) = c_0 B_{0,n}(t) + c_1 B_{1,n}(t) + \dots + c_n B_{n,n}(t)$ polinomunu matris olarak iki vektör halinde yazılır.

$$B(t) = [B_{0,n}(t) \quad B_{1,n}(t) \quad \dots \quad B_{n,n}(t)] \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{bmatrix}$$

$$B(t) = \begin{bmatrix} 1 & t & t^2 & \dots & t^n \end{bmatrix} \begin{bmatrix} b_{0,0} & 0 & 0 & \dots & 0 \\ b_{1,0} & b_{1,1} & 0 & \dots & 0 \\ b_{2,0} & b_{2,1} & b_{2,2} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{n,0} & b_{n,1} & b_{n,2} & \dots & b_{n,n} \end{bmatrix}$$

$b_{i,j}$ 'ler ilgili bernstein polinom temelinin katsayılarıdır. $n=2$ olması durumunda matris,

$$B(t) = [1 \quad t \quad t^2] \begin{bmatrix} 1 & 0 & 0 \\ -2 & 2 & 0 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} \text{şeklinde tanımlıdır.}$$

$n=3$ iken matris,

$$B(t) = [1 \quad t \quad t^2 \quad t^3] \begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 3 & 0 & 0 \\ 3 & -6 & 3 & 0 \\ 1 & 3 & -3 & 1 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} \text{şeklindedir [38].}$$

Genel bir bakış açısıyla; $f(x)$, $[0,1]$ aralığında tanımlı bir fonksiyon olsun.

$B_n(x) = B_n^f(x) = \sum_{v=0}^n f\left(\frac{v}{n}\right) \binom{n}{v} x^v (1-x)^{n-v}$, $v \in \{0, \dots, n\}$ fonksiyonu $f(x)$ fonksiyonun n sıralısının Bernstein polinomu olarak adlandırılır. $n+1$ tane

$b_{v,n}(x) = \binom{n}{v} x^v (1-x)^{n-v}$, $v \in \{0, \dots, n\}$ ifadesi Bernstein temel polinomudur. Daha geniş bir tanım olarak, Bernstein temel polinomlarının herhangi bir doğrusal kombinasyonu $B(x) = \sum_{v=0}^n \beta(v) b_{v,n}(x)$ ifadesi n 'inci dereceden Bernstein polinomu olarak söylenir. $\beta(v)$ katsayılarına Bernstein katsayıları denir [39]. Bernstein polinomu olasılık teorisiyle ilişkilidir. $b_{v,n}(x)$ polinomu olasılık teorisinde kullanılan binom olasılığı ya da Newton olasılığı olarak bilinir. Bir E olayının olasılığı $0 \leq x \leq 1$ arasındadır. $b_{v,n}(x)$ olasılığı bize E olayının n bağımsız deneme sonrası v kez gerçekleşmesi anlamına gelir. Bu polinomu S. Bernstein [40] Weistrass teoremi yaklaşımının kanıtını açıklayıcı bir biçimde yapabilmek amacıyla geliştirmiştir. Bu yaklaşım Bernstein kopulanın özellikleri hakkında bilgi verir.

Teorem 3.3 (Weistrass teoremi): Bernstein polinomları $f(x)$ fonksiyonu ile doğrusaldır:

$f(x) = a_1 f_1(x) + a_2 f_2(x)$ ise $B_n^f(x) = a_1 B_n^{f_1}(x) + a_2 B_n^{f_2}(x)$ 'dir.

$b_{v,n}(x) = \binom{n}{v} x^v (1-x)^{n-v} \geq 0$, $0 \leq x \leq 1$ ve $\sum_0^n b_{v,n}(x) = 1$ için, $m \leq f(x) \leq M$ aralığında $m \leq B_n^f(x) \leq M$ 'dir.

$[a, b]$ kapalı aralığında sürekli her bir $f(x)$ fonksiyonu, $\varepsilon > 0$ için $f(x)$ 'e tek düze yaklaşan bir $P(x)$ polinomu vardır ve $|f(x) - P(x)| > \varepsilon$ 'dur. Doğrusal olması durumunda $[a, b]$ kapalı aralığı $[0,1]$ olarak dönüştürülebilir.

Teorem 3.4: $f(x)$ $[0,1]$ aralığında sınırlı fonksiyonu, f 'in sürekli olduğu her bir x için

$\lim_{n \rightarrow \infty} B_n(x) = f(x)$ 'dir. Eğer $f(x)$ $[0,1]$ aralığında sürekli ise bu ilişki $[0,1]$ aralığında düzgündür [40].

Sklar'in çok değişkenli dağılımlar için kullandığı teoremini göz önüne alınsın. H marjinalleri tek boyutlu $F_1, \dots, F_k$ olan k boyutlu dağılım fonksiyonu olmak üzere,

$[0,1]^k \rightarrow [0,1]$ 'e götüren bir C fonksiyonu vardır.

$H(x_1, x_2, \dots, x_n) = C(F_1(x_1), F_2(x_2), \dots, F_n(x_n))$ sağlayan C , n -boyutlu bir kopuladır. Eğer F_j 'ler sürekli ise C kopulası tektir.

$\alpha(\frac{v_1}{m_1}, \dots, \frac{v_k}{m_k})$ gerçekteğerli bir sabit $(v_1, \dots, v_k)$, $v_j \in \mathbb{N}$, $0 \leq v_j \leq m_j$.

$B_{v_j, m_j}(u_j) = \binom{m_j}{v_j} u_j^{m_j - v_j} (1 - u_j)^{v_j}$ olsun.

$C_B: [0,1]^k \rightarrow [0,1]$ olmak üzere,

$$C_B(u_1, \dots, u_k) = \sum_{v_1} \dots \sum_{v_k} \alpha(\frac{v_1}{m_1}, \dots, \frac{v_k}{m_k}) B_{v_1, m_1}(u_1) \dots B_{v_k, m_k}(u_k) \quad (3.2)$$

kopula fonksiyonu her bir $m_j \geq 1$ için C_B Bernstein kopuladır. Bernstein kopula polinom kopula ailesinin genelleştirilmiş halidir. Polinom kopulası tek veya daha çok değişkenli polinomlar ile oluşan kopulaların özel bir halidir. İki boyutlu durum için,

$C(u_1, u_2) = u_1^3 a(u_2) + u_1^2 b(u_2) + u_1 c(u_2) + d(u_2)$ kopulası $a(\dots)$, $b(\dots)$, $c(\dots)$, $d(\dots)$ seçilen uygun fonksiyonlarla bir kopuladır. Bu kopula kübik bir parçadır ve seçilen fonksiyonlar polinom olabilir. Bernstein kopula (C_B) ile diğer tüm kopulaların ortak özellikleri aşağıda listelenmiştir.

1. C_B tüm durumlarda artan fonksiyonlardır.
2. C_B Fréchet sınırlarına uyar:

$\min (0, u_1 + \dots + u_k - (k - 1)) \leq C_B (u_1, \dots, u_k) \leq \min (u_1, \dots, u_k)$, en az bir $u_j=0$ olsa bile $C_B (u_1, \dots, u_k) = 0$; ve j ne olursa olsun $C_B (1, \dots, 1, u_j, 1, \dots, 1) = u_j$.

3. $\prod_{j=1}^k u_j$ bağımsız rasgele değişkenler için kopula; üretici kopula
4. C_B Lipschitz koşulunu sağlar:

$$|C_B(x_1, \dots, x_k) - C_B(y_1, \dots, y_k)| \leq \sum_{j=1}^k |x_j - y_j|$$

Teorem 3.5: $C_B (u_1, \dots, u_k)$ Bernstein kopula olması için gerek ve yeter koşul

$$\sum_{l_1=0}^1 \dots \sum_{l_k=0}^1 (-1)^{l_1+\dots+l_k} \alpha\left(\frac{v_1+l_1}{m}, \dots, \frac{v_k+l_k}{m}\right) \geq 0 \text{ ve}$$

$$\min\left(0, \frac{v_1}{m} + \dots + \frac{v_k}{m} - (k - 1)\right) \leq \alpha\left(\frac{v_1}{m}, \dots, \frac{v_k}{m}\right) \leq \min\left(\frac{v_1}{m}, \dots, \frac{v_k}{m}\right), \quad \text{özellikle}$$

$$\lim_{v_j=0} \alpha\left(\frac{v_1}{m}, \dots, \frac{v_k}{m}\right) = 0, \text{ her } j=1, \dots, k \text{ ve } \alpha\left(1, \dots, 1, \frac{v_j}{m}, 1, \dots, 1\right), \text{ her } j=1, \dots, k.$$

Bazı durumlarda Bernstein kopula üretici kopula ve düzensiz terimlerin toplamı şeklinde yazılır.

$$\begin{aligned} C_B (u_1, \dots, u_k) &= u_1 \dots u_k + \sum_{v_1=0}^m \dots \sum_{v_k=0}^m \gamma\left(\frac{v_1}{m}, \dots, \frac{v_k}{m}\right) B_{v_1,m}(u_1) \dots B_{v_k,m}(u_k) \\ &= \sum_{v_1=0}^m \dots \sum_{v_k=0}^m \alpha\left(\frac{v_1}{m}, \dots, \frac{v_k}{m}\right) B_{v_1,m}(u_1) \dots B_{v_k,m}(u_k). \end{aligned}$$

$$\gamma\left(\frac{v_1}{m}, \dots, \frac{v_k}{m}\right) = \alpha\left(\frac{v_1}{m}, \dots, \frac{v_k}{m}\right) \cdot \frac{v_1}{m} \dots \frac{v_k}{m} \text{ eşitliğinden yola çıkarak}$$

$$u_1 \dots u_k = \sum_{v_1=0}^m \dots \sum_{v_k=0}^m \alpha\left(\frac{v_1}{m}, \dots, \frac{v_k}{m}\right) B_{v_1,m}(u_1) \dots B_{v_k,m}(u_k) \text{ elde edilir.}$$

Teorem 3.6: Herhangi bir $C (u_1, \dots, u_k)$ kopulası $u_1 \dots u_k$ üretici(bağımsız) kopula ile ve $G(u_1, \dots, u_k)$, $u_1, \dots, u_k$ 'nin bağımlılık yapısı ile ilgili bütün bilgileri içeren düzensiz terimin toplamı $u_1 \dots u_k + G(u_1, \dots, u_k)$ şeklinde yazılabilir.

Teorem 2.6'da ki $G (u_1, \dots, u_k)$ terimi bağımsız kopulanın kopuladan uzaklığı ile tanımlanır. Bu aşağıdan ve yukarıdan Fréchet sınırı ile sınırlıdır.

Bernstein kopula kesin süreklidir ve herhangi bir Bernstein kopula bir kopula dağılım fonksiyonudur. $\Delta_{1,\dots,k}$, k boyutlu fark operatörü olarak tanımlayalım.

$$\Delta_{1,\dots,k} \alpha\left(\frac{v_1}{m}, \dots, \frac{v_k}{m}\right) \equiv \sum_{l_1=0}^1 \dots \sum_{l_k=0}^1 (-1)^{k+l_1+\dots+l_k} \alpha\left(\frac{v_1+l_1}{m}, \dots, \frac{v_k+l_k}{m}\right)$$

Bernstein kopula dağılımı Bernstein polinomunun özelliklerinden dolayı aşağıdaki forma sahiptir. c_B bernstein dağılım fonksiyonunu göstermek üzere;

$$c_B(u_1, \dots, u_k) = m^k \sum_{v_1=0}^{m-1} \dots \sum_{v_k=0}^{m-1} \Delta_{1,\dots,k} \alpha\left(\frac{v_1}{m}, \dots, \frac{v_k}{m}\right) B_{v_1, m-1}(u_1) \dots B_{v_k, m-1}(u_k)$$

biçimindedir.

$c_B = \frac{\partial^k c_B}{\partial u_1 \dots \partial u_k}$ bernstein kopula tanımının türevinden elde edilir. Kopula dağılımı daima pozitifdir. Bir diğer gösterimle

$$c_B(u_1, \dots, u_k) = \sum_{v_1=0}^m \dots \sum_{v_k=0}^m \beta\left(\frac{v_1}{m}, \dots, \frac{v_k}{m}\right) \prod_{j=1}^k \binom{m}{v_j} u_j^{v_j} (1 - u_j)^{m - v_j},$$

$$\beta\left(\frac{v_1}{m}, \dots, \frac{v_k}{m}\right) \equiv m^{k+1} \Delta_{1,\dots,k} \alpha\left(\frac{v_1}{m+1}, \dots, \frac{v_k}{m+1}\right)$$

biçimindedir.

Bernstein kopula dağılımının kullanılmasında bir çok temel amaç vardır. Bunlardan birincisi; Bernstein kopulayı yüksek boyutlarda kullanmak matematiksel ve istatistiksel açıdan daha kolay olmasıdır. Bu durum ise kalibrasyon ve benzetim çalışmalarında diğer parametrik ya da parametrik olmayan bağımlılık modelleri ile karşılaştırıldığında tercih sebebidir. İkincisi eliptik ve diğer kopula yaklaşımları belli bir simetri derecesine sahip ve sınırlı korelasyon yapıları verirken, Bernstein kopula bu sınırlamalara bağlı değildir. Bağımlılık yapısına daha yakın tahminler verir [25]. Üçüncüsü Bernstein kopula keyfi bir kopula ile kolayca uyum sağlayabilir [41].

3.6.2.2. Grid-Type Kopula Yaklaşımı

$d \in \mathbb{N}$, $U = (U_1, \dots, U_d)$ rasgele vektörünün marjinal bileşenleri U_i 'ler kesikli uniform dağılım. $T_i = \{0, 1, \dots, m_i - 1\}$, $m_i \in \mathbb{N}$, $i=1, \dots, d$ olmak üzere

$p(k_1, \dots, k_d) := P\left(\bigcap_{i=1}^d \{U_i = k_i\}\right)$ olasılıkları tüm $(k_1, \dots, k_d) \in \times_{i=1}^d T_i$ için Bernstein kopula

$$c(u_1, \dots, u_d) := \sum_{k_1=0}^{m_1-1} \dots \sum_{k_d=0}^{m_d-1} p(k_1, \dots, k_d) \prod_{i=1}^d m_i B(m_i - 1, k_i, u_i)$$

$(u_1, \dots, u_d) \in [0, 1]^d$ U'lar tarafından üretilen d-boyutlu kopuladır.

Bütün $(k_1, \dots, k_d) \in \times_{i=1}^d T_i$ olası seçimler için $I_{(k_1, \dots, k_d)} := \times_{j=1}^d \left(\frac{k_j}{m_j}, \frac{k_j+1}{m_j}\right)$ aralıkları tanımlansın.

$c_{(u_1, \dots, u_d)}^* := \sum_{k_1=0}^{m_1-1} \dots \sum_{k_d=0}^{m_d-1} p(k_1, \dots, k_d) \mathbf{1}_{I_{(k_1, \dots, k_d)}}$ U'lar tarafından üretilen d-boyutlu kopula yoğunluk fonksiyonuna grid-type kopula denir. $\mathbf{1}_A$, A olayı için sırasıyla aralıklarda gözlenen olasılık değeri var ise 1, olmaması durumunda 0 değerini almasıdır.

$[0,1]$ aralığında negatif olmayan $\{\phi(m, k, .) | 0 \leq k \leq m-1, m \in \mathbb{N}\}$ fonksiyonlar ailesi;

- $\int \phi(m, k, .) du = \frac{1}{m}; \quad k=0,1,\dots,m-1$
- $\sum_{k=0}^{m-1} \phi(m, k, .) = 1; \quad m \in \mathbb{N}$

koşulları Bernstein kopula ve grid-type kopula için gridlere parçalanmaya (birimsel bölünmeler) uygun olması durumunda dikkate alınır. Bu durumda U'lar tarafından üretilen d-boyutlu kopula yoğunluk fonksiyonu,

$$c_{(u_1, \dots, u_d)}^{\phi} := \sum_{k_1=0}^{m_1-1} \dots \sum_{k_d=0}^{m_d-1} P(\cap_{i=1}^d \{U_i = k_i\}) \prod_{i=1}^d m_i \phi(m_i, k_i, u_i), (u_1, \dots, u_d) \in [0,1]^d$$

kopula fonksiyonu,

$$C_{(u_1, \dots, u_d)}^{\phi} := \sum_{k_1=0}^{m_1-1} \dots \sum_{k_d=0}^{m_d-1} P(\cap_{i=1}^d \{U_i < k_i\}) \prod_{i=1}^d \phi(m_i, k_i, u_i), (u_1, \dots, u_d) \in [0,1]^d$$

şeklinde elde edilirler. m grid sayısı verilerin iki boyutlu olması durumunda $m_1=m_2=m$ olarak kabul edilir. Bernstein kopula olması durumunda polinomlar kullanılırken,

$$\phi(m, k, u) = B(m-1, k, u) = \binom{m-1}{k} u^k (1-u)^{m-1-k}$$

grid-type kopula için,

$$\phi(m, k, u) = \mathbf{1}_{I_{(\frac{k}{m}, \frac{k+1}{m})}}(u), 0 \leq k \leq m-1$$

gösterge fonksiyonu kullanılır [41].

3.7. Bernstein Kopula ile Rasgele Sayı Üretme

Bernstein kopula dağılım fonksiyonu, eşitlik 3.2'de gösterilen deneysel kopula olasılıklarının Bernstein polinomları yaklaşımından elde edilmiştir.

İlk adım olarak d=2 boyutlu hasar verilerinden elde edilen deneysel kopula olasılıkları bulunmuştur. Bu olasılıklar gridler ile uyumlu hale getirilip $m_1=m_2=m=10$ grid sayısı için $m_1 \times m_2 = 100$ hücreden oluşan verinin gerçek sıklık tablosu oluşturulmuştur. Elde edilen sıklık tablosu $[a_{ij}]$ 'ler (i=satır dizini, j=sütun dizini) ile temsil edilmektedir. Bu tabloda marjinal dağılım olasılıkları yani satır ve sütun toplamaları $\frac{1}{m} = \frac{1}{10}$ eşit olarak dağılmamaktadır. Bu durumda en uygun $[x_{ij}]$ olasılıklarını bulabilmek için

$$\min \sum_{i=1}^m \sum_{j=1}^m (x_{ij} - a_{ij})^2$$

$$\sum_{i=1}^m x_{ik} = \sum_{j=1}^m x_{lj} = \frac{1}{m} = \frac{1}{10} \quad \text{ve} \quad x_{l,k} \geq 0 \quad l, k = 1, \dots, m$$

optimizasyon problemi çözülmüştür.

Bernstein kopula dağılımı d değişkenli polinomlar olduğundan $[0,1]^d$ birim küpünde sınırlıdır. Bu açıdan en uygun yol çok değişkenli kabul-red etme yöntemidir. M , tek bir rasgele sayı elde etmek için çok sayıda reddetmenin önüne geçip kabul etme olasılıklarını göreceli olarak arttırmak için kullanılan bir sabittir. Yöntem için uygulanan adımlar sırasıyla;

- Adım 1: $d+1$ tane bağımsız uniform dağılım $u_1, \dots, u_{d+1}$ rasgele sayıları üretilir.
- Adım 2: Eğer $c(u_1, \dots, u_d) > M u_{d+1}$ ise 3. adıma geçilir aksi durumda 1. Adıma geri dönülür.
- Adım 3: eşitsizliği sağlayan $u_1, \dots, u_d$ değerleri Bernstein kopula örnekleme olarak kullanılır.

4. UYGUN KOPULA MODEL SEÇİMİ UYGULAMASI

Bu bölümde Türkiye’de hayat dışı alanda faaliyet göstermekte olan bir sigorta şirketinin; kasko, trafik, ihtiyari mali mesuliyet (imm), konut/yangın, konut/hırsızlık, konut/cam kırılması, ferdi kaza, kasko koltuk ferdi kaza, bireysel sağlık sigortaları olmak üzere 9 farklı branşlarında 2008-2013 yıllarına ait 72 aylık ödenen hasar verileri kullanılmıştır. Veriyi enflasyon etkisinden arındırıp 2018 Nisan değerleri ile ifade edebilmek için TÜİK’den alınan (2003=100 bazlı) tüketici fiyat endeks değerlerinden yararlanılmıştır.

Gerçek hasar verisinde gözlenen bağımlılık yapılarına en uyumlu tahminleri yapabilmek için parametrik ve parametrik olmayan kopula modelleri incelenmiştir. Bu çalışmada incelenen kopula modelleri; Gaussian, t, Clayton, Frank ve Bernstein kopuladır. Ayrıca sonuçları karşılaştırmak için bağımsızlık durumu da incelenmiştir.

$(C_{\theta}^{Cl}, C_{\theta}^{Fr}, C_{\rho}^{Ga}, C_{\rho, \nu}^t, C_B)$ kopulalarının her biri için elde edilen kopula yoğunluk fonksiyonları sırasıyla $(\hat{C}_{\theta}^{Cl}, \hat{C}_{\theta}^{Fr}, \hat{C}_{\rho}^{Ga}, \hat{C}_{\rho, \nu}^t, \hat{C}_B)$ ’dir. Burada C_B ve $\hat{C}_B$ sırasıyla Bernstein kopula dağılım ve yoğunluk fonksiyonlarıdır. Bu modeller içinde parametrik olanların kopula fonksiyonları sırasıyla,

$$C_{\theta}^{Cl}(u_1, u_2) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{\frac{-1}{\theta}}$$

$$C_{\theta}^{Fr}(u_1, u_2) = -\frac{1}{\theta} \ln \left(1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1} \right)$$

$$C_{\rho}^{Ga}(u_1, u_2) = \phi_{\rho}(\phi_{(u_1)}^{-1}, \phi_{(u_2)}^{-1})$$

$$C_{\rho, \nu}^t = t_{\nu, m}(t_{\nu}^{-1}(u_1), t_{\nu}^{-1}(u_2))$$

biçimindedir.

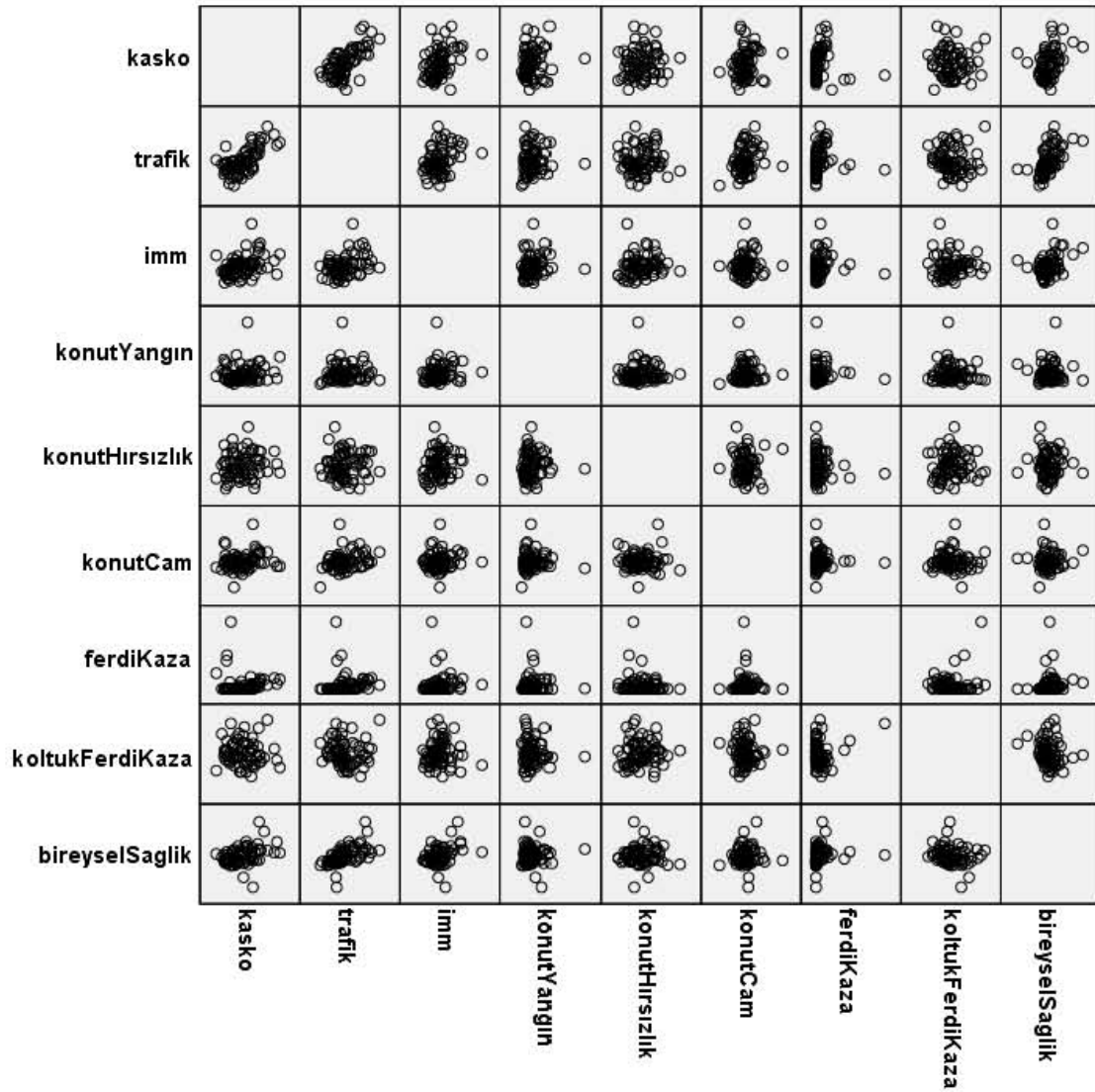
Parametrik modeller arasında uyumluluk açısından seçim yapabilmek için AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion) değerleri ve bir çok uyum iyiliği testlerine bakılabilirken; bu değerler parametrik olmayan Bernstein kopula dağılımı için elde edilememektedir. Bu nedenle bu çalışmada uyum iyiliği ölçümü olarak öklid uzaklıklarından faydalanılmıştır. Öklid uzaklıkları,

$$ED_{c,\hat{c}} = \sqrt{\sum_{j=1}^p (c_{\theta} - \hat{c}_{\theta})^2}$$

formülünden elde edilmiştir. c_{θ} , deneysel (gözlenen) kopula yoğunlukları; $\hat{c}_{\theta}$, tahmin edilen kopula yoğunluklarıdır.

4.1. Ödenen hasar verisi ön incelemesi

Sigorta branşlarına ait ödenen hasarlar arasındaki korelasyon incelenmiş korelasyon matris grafiği Şekil 4.1’de ve korelasyon matrisi Çizelge 4.1’de verilmiştir.



Şekil 4. 1: Hasar Verilerine Ait Korelasyon Matris Grafiği

Çizelge 4. 1: Hasar Verilerine Ait Korelasyon Matrisi

		Kasko	Trafik	İmm	Konut Yangın	Konut Hırsızlık	Konut Cam	Ferdi Kaza	Koltuk Ferd Kaza	Bireys el Saglık
Kasko	Pearson Correlation	1	0,669**	0,388**	0,146	0,120	0,176	-0,014	-0,033	0,357**
	Sig. (2-tailed)		0,000	0,001	0,220	0,314	0,138	0,910	0,781	0,002
	N	72	72	72	72	72	72	72	72	72
Trafik	Pearson Correlation	0,669**	1	0,358**	0,078	0,030	0,274*	0,091	-0,067	0,573**
	Sig. (2-tailed)	0,000		0,002	0,512	0,803	0,020	0,448	0,578	0,000
	N	72	72	72	72	72	72	72	72	72
İmm	Pearson Correlation	0,388**	0,358**	1	0,136	0,093	0,088	0,051	0,006	0,371**
	Sig. (2-tailed)	0,001	0,002		0,253	0,436	0,464	0,670	0,963	0,001
	N	72	72	72	72	72	72	72	72	72
Konut Yangın	Pearson Correlation	0,146	0,078	0,136	1	0,046	-0,002	-0,038	-0,04	0,015
	Sig. (2-tailed)	0,220	0,512	0,253		0,703	0,985	0,750	0,739	0,898
	N	72	72	72	72	72	72	72	72	72
Konut Hırsızlık	Pearson Correlation	0,120	0,030	0,093	0,046	1	-0,007	-0,129	0,055	0,004
	Sig. (2-tailed)	0,314	0,803	0,436	0,703		0,952	0,280	0,646	0,970
	N	72	72	72	72	72	72	72	72	72
Konut Cam	Pearson Correlation	0,176	0,274*	0,088	-0,002	-0,007	1	0,013	-0,043	0,064
	Sig. (2-tailed)	0,138	0,020	0,464	0,985	0,952		0,917	0,720	0,593
	N	72	72	72	72	72	72	72	72	72
Ferd Kaza	Pearson Correlation	-0,014	0,091	0,051	-0,038	-0,129	0,013	1	0,191	0,164
	Sig. (2-tailed)	0,910	0,448	0,670	0,750	0,280	0,917		0,109	0,168
	N	72	72	72	72	72	72	72	72	72
Koltuk Ferd Kaza	Pearson Correlation	-0,033	-0,067	0,006	-0,040	0,055	-0,043	0,191	1	-0,195
	Sig. (2-tailed)	0,781	0,578	0,963	0,739	0,646	0,720	0,109		0,101
	N	72	72	72	72	72	72	72	72	72
Bireysel Sağlık	Pearson Correlation	0,357**	0,573**	0,371**	0,015	0,004	0,064	0,164	-0,195	1
	Sig. (2-tailed)	0,002	0,000	0,001	0,898	0,970	0,593	0,168	0,101	
	N	72	72	72	72	72	72	72	72	72

** . %99,9 güven düzeyi

* . %99,5 güven düzeyi

Korelasyonun %99 güven düzeyinde anlamlı olduğu kasko-trafik, trafik-bireysel sağlık, kasko-imm sigortalı arasındaki bağımlılıkların modellenmesine karar verilmiştir. Bu üç adet ikili bağımlı hasar verisinde yer alan; kasko, trafik, imm, bireysel sağlık, konut

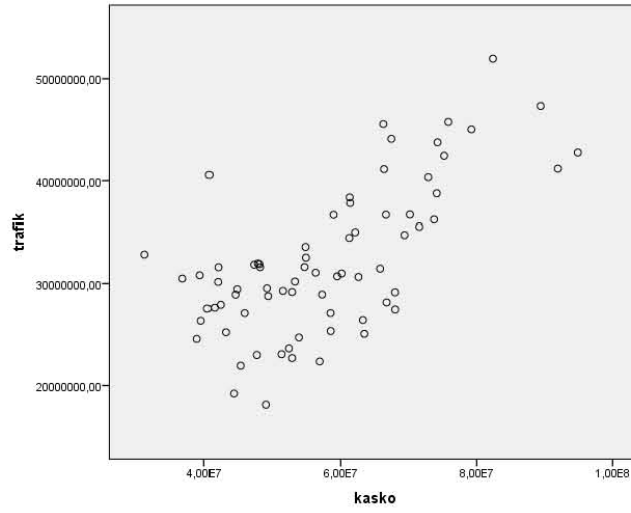
hırsızlık sigortaları ödenen hasar verilerine uygun dağılımlar belirlenmiştir (Bknz. Ek-1). Belirlenen dağılımların adları ve parametreleri aşağıda Çizelge 4.2’de verilmiştir.

Çizelge 4. 2: Verilere uygun dağılım ve parametreler

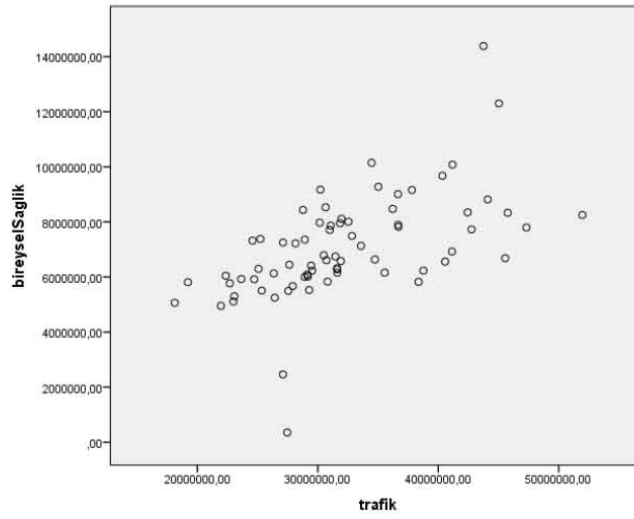
	Kasko	Trafik	İmm	Sağlık
Dağılım	Lognormal	Lognormal	Lognormal	Normal
Kolmogorov-Smirnov değeri	0,057	0,093	0,079	0,109
Parametreler	$\mu=17,843$ $\sigma=0,23423$	$\mu=17,263$ $\sigma=0,21869$	$\mu=13,801$ $\sigma=0,39492$	$\mu=7,0622E+6$ $\sigma=1,9141E+6$

Örneklem büyüklüğü 72 ve güven düzeyi 0,99 için Kolmogorov-Smirnov test istatistiği kritik değeri 0,18903’tür. Bu durumda Çizelge 4.2’deki dört veri seti için de hipotezler red edilememektedir. Bu durumda, test edilen dağılımlara uydukları söylenebilir.

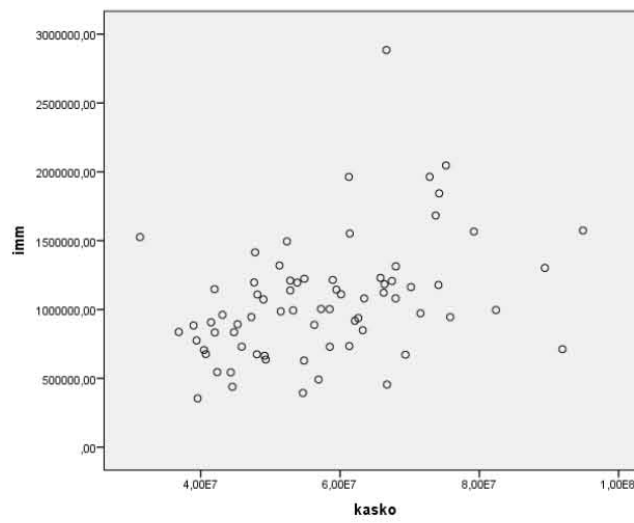
Bağımlılık modellemesinde, sırasıyla kasko-trafik, trafik-bireysel sağlık, kasko-imm sigortaları çiftlerine ait saçılım grafikleri aşağıda Şekil 4.2-4.4’de verilmiştir.



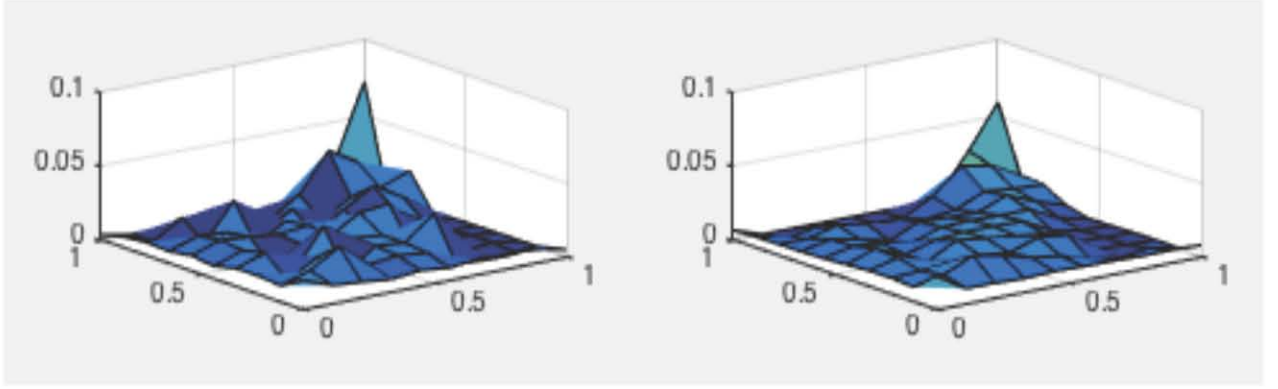
Şekil 4. 2: Kasko ve Trafik verisine ait saçılım grafiği



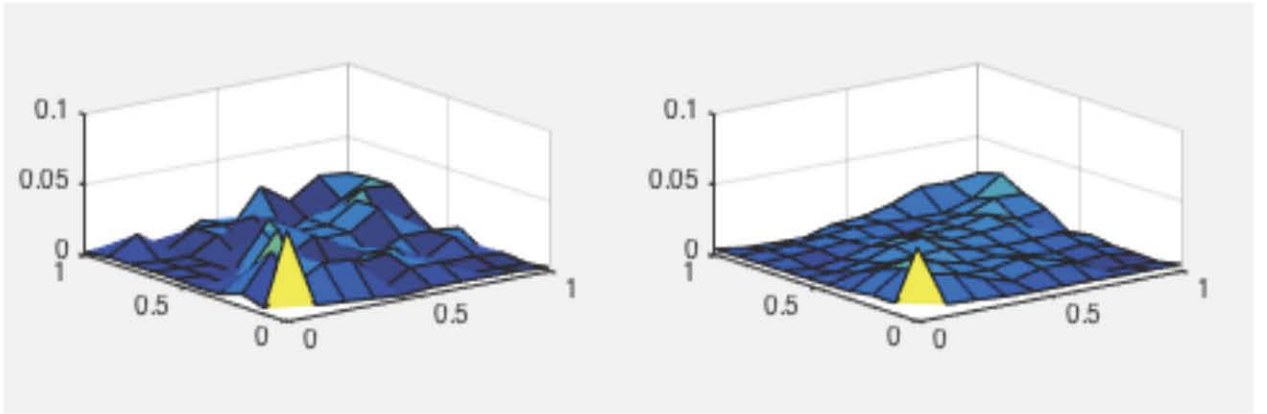
Şekil 4. 3: Trafik ve Bireysel sağlık sigortası verisine ait saçılım grafiği



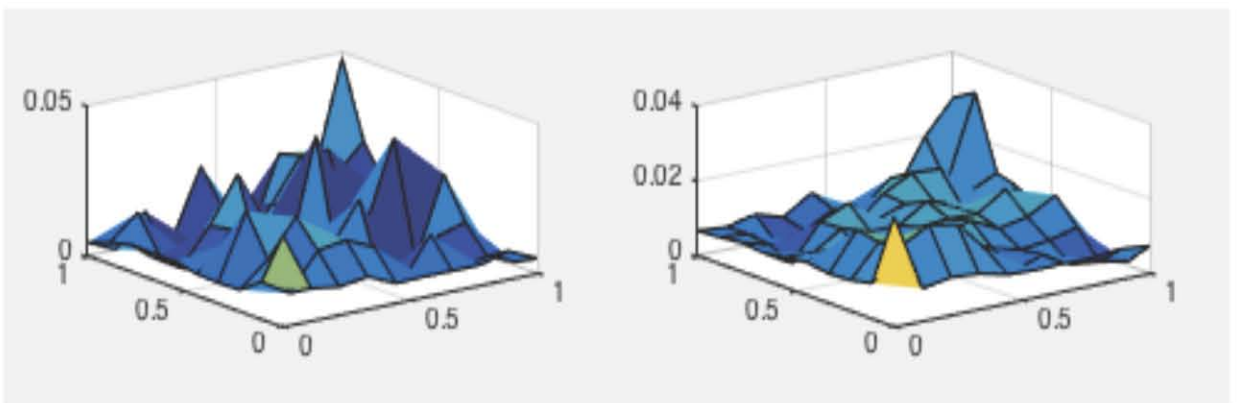
Şekil 4. 4: Kasko ve İmm verisine ait saçılım grafiği



Şekil 4. 5: Kasko ve Trafik verilerine ait deneysel kopula yoğunlukları yüzey grafiği ile Bernstein kopula yoğunlukları yüzey grafiği

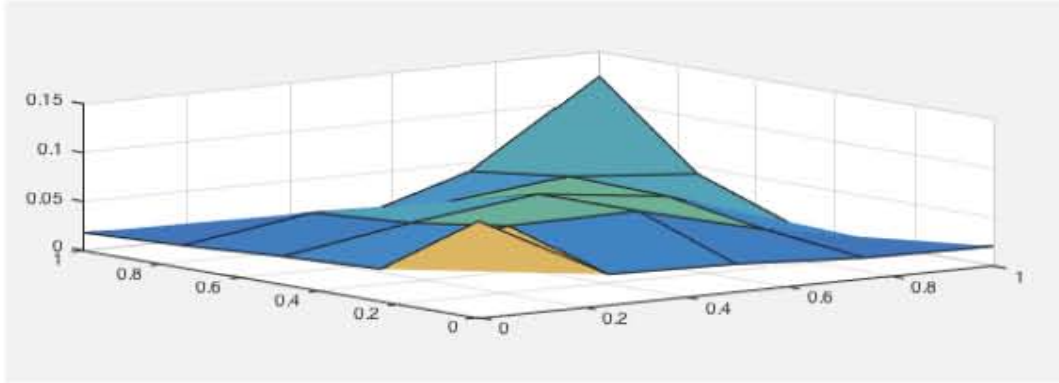


Şekil 4. 6: Trafik ve Bireysel sağlık sigortası verilerine ait deneysel kopula yoğunlukları yüzey grafiği ile Bernstein kopula yoğunlukları yüzey grafiği

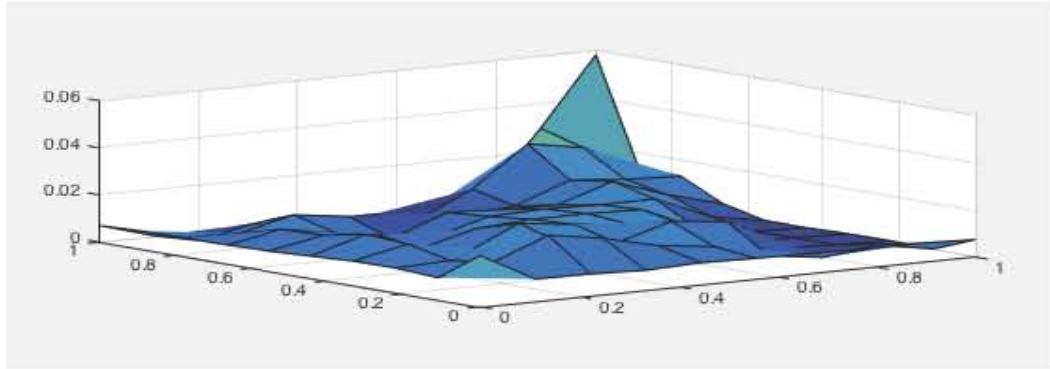


Şekil 4. 7: Kasko ve İmm verilerine ait deneysel kopula yoğunlukları yüzey grafiği ile Bernstein kopula yoğunlukları yüzey grafiği

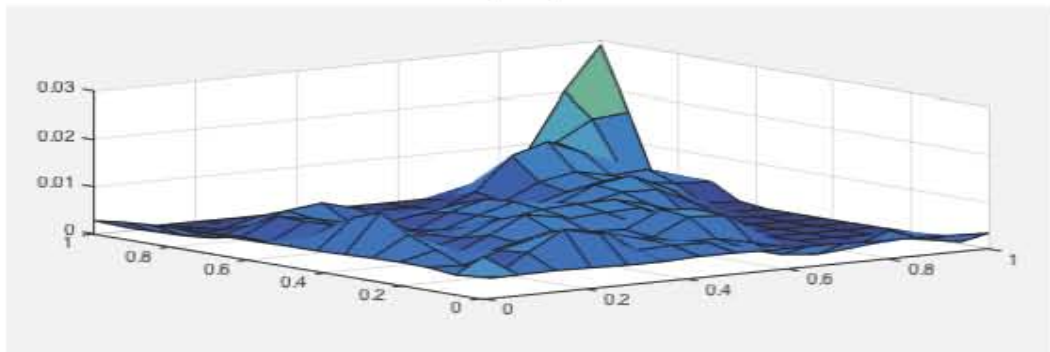
Şekil 4.5-4.7’de kasko-trafik, trafik-bireysel sağlık, kasko-imm sigortaları hasar verileri için grid sayısı $m=10$ alınarak hesaplanan deneysel kopula yoğunlukları ve bu verilerin Bernstein kopula yoğunlukları verilmiştir. Her bir veri seti için bağımlılığın, homojen ve simetrik bir yapıda olmadıkları görülmektedir. Bernstein kopula yoğunlukları deneysel kopula yoğunlukları ile karşılaştırıldığında veri setleri için daha pürüzsüz bir yapı sağladığı gözlemlenmiştir.



Şekil 4. 8: $m=5$ için Kasko Trafik veri çiftine ait Bernstein kopula yoğunlukları yüzey grafiği



Şekil 4. 9: $m=10$ için Kasko Trafik veri çiftine ait Bernstein kopula yoğunlukları yüzey grafiği



Şekil 4. 10: $m=15$ için Kasko Trafik veri çiftine ait Bernstein kopula yoğunlukları yüzey grafiği

Şekil 4.8-4.10’da kasko trafik veri çiftine ait grid sayısının $m=5,10,15$ olması durumundaki Bernstein kopula yoğunlukları verilmiştir. Grid sayısı arttıkça kopula modelinin daha pürüzsüzleştiği görülmektedir. Ancak daha pürüzsüz olmasının daha iyi bir model olduğu anlamına gelmediği söylenebilir.

4.2. Uyum İyiliği

Çalışmada kopula model seçimi için uyum iyiliği testi olarak öklid uzaklıkları kullanılmıştır. Bu amaçla; $\rho = 0.65$, $(\rho=0.65, \nu = 4)$, $\theta = 2, \theta=5$ parametreleri kullanılarak Gaussian, t, Clayton ve Frank parametrik kopula modellerinin her biri ile iki değişkenli 72’lik rasgele örneklemelerden oluşan dört veri seti oluşturulmuştur. Ayrıca, üç adet ikili bağımlı gerçek hasar veri seti kullanılmıştır. Gerçek veri setleri de 72’şer çift değerden oluşmaktadır. Böylece, çeşitli kopula modelleri ile tahmin edilmek üzere 72’şer çift değerden oluşan yedi adet veri seti elde edilmiştir.

Öncelikle parametrik kopula modellerinin verilen parametreleri ile üretilmiş 72’lik rasgele örneklemelerden oluşan dört veri seti için yine bu dört parametrik kopula modeli ile parametre tahminleri yapılmıştır. Çizelge 4.3’de yatay ekseninde gösterilen $\rho = 0.65$, $(\rho=0.65, \nu = 4)$, $\theta = 2, \theta=5$ değerleri sırasıyla Gaussian, t, Clayton ve Frank kopula parametreleridir. Gaussian, t, Clayton ve Frank kopula modelleri için sırasıyla $\hat{c}_\rho^{Ga}$, $\hat{c}_{\rho,\nu}^t$, $\hat{c}_\theta^{Cl}$, $\hat{c}_\theta^{Fr}$ tahmini kopula parametreleri maksimum olabilirlik fonksiyonu ile elde edilmiştir. Elde edilen parametre tahmin değerleri Çizelge 4.3’de verilmiştir.

Çizelge 4. 3: Model seçimi için parametre tahminleri

	c_ρ^{Ga}	$c_{\rho,\nu}^t$	c_θ^{Cl}	c_θ^{Fr}
θ	$\rho = 0.65$	$(\rho=0.65, \nu = 4)$	$\theta = 2$	$\theta=5$
$\hat{c}_\rho^{Ga}$	0.6939	0.5885	0.7439	0.7274
$\hat{c}_{\rho,\nu}^t$	$\rho=0.6135,$ $\nu = 1.5844$	$\rho=0.6251,$ $\nu = 6.4599$	$\rho=0.7481,$ $\nu = 8.4545$	$\rho=0.7456,$ $\nu = 4.6692$
$\hat{c}_\theta^{Cl}$	0.9053	0.9629	2.0610	1.6298
$\hat{c}_\theta^{Fr}$	5.7284	4.8158	7.1350	6.8786

Uyum iyiliği testleri arasında yer alan AIC (Akaike Information Criterion) ve BIC (Bayesian Information Criterion)’in parametrik kopula modelleri arasında sıklıkla tercih edilmesine rağmen bu test istatistiğinin parametrik olmayan kopula modelleri için uygulanabilir olmadığı görülmüştür. Bu nedenle kopula model seçimi için uyum iyiliği ölçümü olarak öklid uzaklıklarına bakılmıştır.

72 çift veriden oluşan bir örneklemin $72 \times 72 = 5184$ deneysel kopula yoğunluk değeri elde edilir. Aynı şekilde bir o kadar da tahmini kopula değeri bulunmaktadır. Böylece sütunda yer alan yedi veri seti için elde edilen 5184'er deneysel kopula değerleri ile; satırda yer alan altı tahmin modeli için elde edilen 5184'er tahmini kopula değerlerinin 42 kere öklid uzaklıkları 1000'er simülasyon denemesi için hesaplanıp ortalamaları alınmış ve aşağıda bu ortalama uzaklıklar Çizelge 4.4'te verilmiştir. Çizelge 4.4'te Bernstein kopula tahminleri grid sayısı $m=5,10,15$ için incelenmiştir.

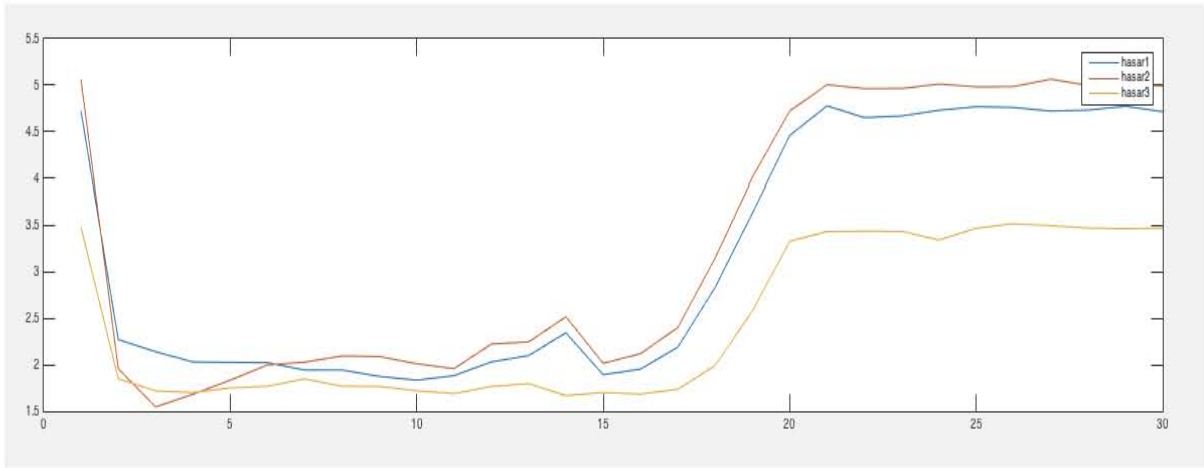
Çizelge 4. 4: 72'lik örneklemler için 1000 simülasyondan elde edilen ortalama öklid uzaklıkları

	Gaussian	t	Clayton	Frank	Kasko Trafik Verisi	Trafik Bireysel Sağlık Verisi	Kasko İmm Verisi
Gaussian	1.1548	1.1562	1.4765	1.3751	2.3787	2.4566	1.7669
t	1.1953	1.1610	1.4908	1.3633	3.0024	3.1873	2.2770
Clayton	1.8972	1.7085	1.1646	2.1941	3.1617	1.5438	1.6921
Frank	1.2049	1.1684	1.4486	1.1560	1.6915	1.1699	1.0012
Bernstein ($m=5$)	1.7742	1.6937	1.6954	1.8086	2.0300	1.7956	1.7471
Bernstein ($m=10$)	1.7920	1.6899	1.7139	1.8723	1.8689	2.0446	1.7262
Bernstein ($m=15$)	2.2163	2.2822	2.9520	2.0445	1.8570	2.0766	1.7012
Bağımsız	4.6635	4.6081	5.1784	4.8495	4.7900	5.0247	3.5293

Çizelge 4.4'te parametrik kopula modelleri (Gaussian, t, Clayton ve Frank) ile üretilmiş veri setlerinin yine parametrik kopula modelleri ile tahmin edildiğinde, parametrik olmayan Bernstein kopula modeline göre daha iyi sonuçlar verdiği görülmektedir. Kasko-Trafik, Trafik-Bireysel sağlık sigortası, Kasko-İmm hasar veri setleri için yapılmış olan Gaussian, t ve Clayton kopula tahmin değerleri için öklid uzaklıklarının en kötü üç sonucu verdiği görülmüştür. Frank kopulanın üç gerçek hasar veri setinin tahmininde de öklid uzaklığını en aza indirdiği görülmektedir. Bernstein kopulanın $m=10$ grid sayısı için ise kasko-trafik ve kasko-imm veri setlerinin tahmininde Frank kopuladan sonra en iyi sonuçları verdiği görülmüştür. Trafik-bireysel sağlık veri setinde ise Frank ve Clayton kopulalarından sonra üçüncü en iyi sonucu vermektedir. Bernstein kopula, parametrik kopula modelleri ile üretilmiş veri setlerinin tahmininde başarısız sonuçlar verirken gerçek veri setlerinde daha başarılı sonuçlar vermesi, gerçek veri setlerindeki bağımlılık yapısının

homojen ve simetrik olmaktan daha uzak olması ile açıklanabilir. Bernstein kopulanın $m=5$ ve $m=15$ grid sayıları için de kasko-trafik veri setinin tahmininde Frank kopuladan sonra en iyi sonuçları verdiği, diğer iki hasar veri setinin tahmininde ise Frank ve Clayton kopulalarından sonra üçüncü en iyi sonucu verdikleri görülmüştür. Bu durumda farklı grid sayılarının seçilmesi, Bernstein kopulanın tahmin sonuçlarını etkilediği görülmektedir. Bu nedenle Bernstein kopula modellemesinde optimal grid sayısının belirlenmesinin önemi ortaya çıkmaktadır.

Bernstein kopulanın Kasko-Trafik, Trafik-Bireysel sağlık sigortası, Kasko-İmm hasar veri setleri için 500 yineleme ve 1-30 değerleri arasındaki grid sayıları için belirlenmiş ortalama öklid uzaklıkları aşağıda Şekil 4.11’de verilmiştir.



Şekil 4. 11: Optimal grid grafikleri

Şekil 4.11’de hasar1: Kasko-Trafik, hasar2: Trafik-Bireysel sağlık sigortası, hasar3: Kasko-İmm hasar veri setlerine ait grafikleri göstermektedir. Kasko-trafik veri seti için optimal grid sayılarının 10 ya da 15 olabileceği, Trafik-Bireysel veri seti için optimal grid sayısının 3 olabileceği, kasko-imm hasar veri seti için ise 3 ile 17 değerleri arasında olabileceği görülmektedir. Optimal grid sayılarındaki bu farklılıklar hasar veri setlerinin simetriklik ve homojenlik açısından karakteristik özelliklerinden kaynaklandığı düşünülmektedir.

4.3. Risk Modellemesi

Kasko-Trafik, Trafik-Bireysel sağlık sigortası, Kasko-İmm bağımlı hasar veri çiftleri; Gaussian, t, Clayton, Frank ve $m=5,10,15$ grid sayıları için Bernstein kopulaları ile modellenmiş, rasgele sayılar üretilmiştir. Bu rasgele sayılardan elde edilen VaR değerleri risk değerlendirilmesi yapılarak kopula modellerinin karşılaştırılmasında kullanılmıştır. VaR değerleri MATLAB programı ile tahmin edilmiştir.

Öncelikle her bir bağımlı hasar veri çifti için Gaussian, t, Clayton ve Frank kopula parametreleri tahmin edilmiştir. Bu parametreler kullanılarak her bir parametrik kopula modeli için 1.000.000'lük bağımlı rasgele örneklemeler oluşturulmuştur. Parametrik olmayan Bernstein kopula için rasgele u, v değerleri m=5,10,15 grid sayıları için Bölüm 3.7'de anlatıldığı gibi üretilmiştir. Üretilen bağımlı rasgele u, v değerleri; veri ön incelemesi aşamasında Çizelge 4.2'de gerçek hasar verileri için belirlenen dağılım ve parametreler dikkate alınarak ters dönüşüm yöntemi ile rasgele hasar değerlerine dönüştürülmüştür. Gaussian, t, Clayton, Frank, Bernstein kopula modellerine ve bağımsızlık durumuna ait tahmini toplam hasar verileri üretilmiştir. Her bir model için tahmin edilen toplam hasar verilerinin %99,5 güven düzeyinde riske maruz değerleri (VaR) hesaplanmıştır. Hesaplanan VaR değerleri aşağıda Çizelge 4.5'de verilmiştir.

Çizelge 4. 5: %99,5 güven düzeyinde VaR değerleri (Milyon TL)

	Kasko&Trafik	Trafik&bireysel sağlık sigortası	Kasko&İmm
Gaussian	147,20	647,59	104,06
t	144,82	642,85	103,76
Clayton	140,91	637,54	103,97
Frank	145,14	646,61	104,07
Bernstein (m=5)	144,50	636,62	103,95
Bernstein (m=10)	147,39	639,60	103,82
Bernstein (m=15)	147,87	637,43	103,75
Bağımsız	138,39	626,43	103,74

Sigorta şirketleri için VaR değerleri önemli bir risk ölçütüdür. Risk değerlendirmesindeki eksiklik ya da fazlalık şirketin mali durumunu olumsuz etkileyecek potansiyele sahiptir. Dolayısıyla bağımlılık riskinin uyumluluk ölçütleri açısından en iyi kopula ile modellenip bu kopuladan VaR değerlerinin tahmin edilmesi sigorta şirketleri açısından önemlidir. Çizelge 4.4'de görüldüğü gibi farklı kopula modelleri ile farklı VaR değerleri elde edilmektedir. Örneğin kasko-trafik toplam hasar ödemesinin Bernstein kopula modeli ile tahmininden elde edilen %99,5 güven düzeyinde VaR değeri 147,87 Milyon TL iken, Clayton kopula ile bu değer 140,91 Milyon TL olarak hesaplanmıştır. Bernstein kopula ile yapılan tahminlerin kasko-trafik gerçek verisi için daha uyumlu sonuç verdiği düşünüldüğünde, toplam hasar verisinin risk değerinin diğer kopulalar ile ya da

bağımsızlık durumunda elde edilmesinin riskin eksik değeriendirilmesine neden olacağı ortaya çıkmaktadır.

5. SONUÇ ve ÖNERİLER

Bu çalışmada hayat dışı sigorta risklerinin Bernstein kopula ile modellenmesinin etkinliği, diğer kopula modelleri ile karşılaştırılarak incelenmiştir.

Bu amaçla öncelikle literatürde yapılan önemli çalışmaların sonuçları incelenmiştir. Martin Eling, Dorothea Diers ve Sebastian D. Marek; 2012 yılında hayat dışı sigorta riskinin bağımlılık çalışmasında Bernstein kopulayı kullanmıştır. Çalışmalarında iki boyutlu uyum iyiliği testi olarak The Cramér-von Mises test istatistiği sonuçları karşılaştırılmıştır. Clayton ($\theta=2$), parametresiyle 100 tane iki değişkenli rasgele sayı üretilmiş, bu değerler Gaussian, Student, Clayton, Gumbel, Bernstein, Grid-type, Kernel kopula modelleriyle tahmin edildiğinde sonuçlar karşılaştırılmıştır. Bağımsızlık durumunu da incelemişlerdir. Clayton modeli ($\theta=2$), ile üretilmiş veri seti için tahmin değerlerine bakıldığında parametrik kopula modelleri arasında en iyi tahmin edicinin Clayton kopula olduğu, parametrik olmayan modeller arasında ise Grid-type kopula modelinin olduğunu belirtmişlerdir. Çalışmalarında, tüm modeller arasında en iyi tahmini Clayton kopula modelinin sağladığı, Bernstein kopula modelinin grid-type ve Clayton tahmin edicilerine yakın değerler verdiği sonucuna varmışlardır. Doro Rose; 2015 yılında doktora çalışmasında kopula modelleri arasından; Gaussian, t, Clayton ve Frank kopula modelleri ile $\rho = 0.65$, ($\rho=-0.3$, $v=4$), $\theta = 1$, $\theta=5$ parametrelerini kullanarak deneysel veri setleri üretmiştir. Kopula modelleri ile üretilen bu verileri Gaussian, t, Clayton, Frank ve Bernstein kopula modelleriyle tahmin etmiş ve sonuçları karşılaştırmıştır. Karşılaştırmada uyum iyiliği istatistiği olarak göreceli uzaklıklara bakılmıştır. Bernstein kopula modelinin Clayton ve Frank kopulaları için iyi bir tahmin edici olduğu, Gaussian ve T kopula modeli için ise ortalamanın üzerinde bir sonuç verdiği belirtilmiştir. P.D. Pfeifer, D. Straburger, P. Jörg, 2009 yılında bağımlılık çalışmasında sel ve fırtına hasar verilerini kullanarak Bernstein, Gaussian kopula modelleri ve bağımsızlık durumu için 1000'er adet iki değişkenli rasgele veri setleri üretmiştir. Modellerin sonuçlar üzerindeki etkisini gözlemlemek için en yüksek olası hasar değerlerini (Probable Maximum Lose) hesaplamışlardır. Ancak burada tahminlerin gerçek veriye uyumları karşılaştırılmadığından modeller arasında bir tercih yapılmamıştır.

Literatürde yapılmış olan bu çalışmaların hiçbirinde gerçek hasar veri seti için uyum iyiliği karşılaştırılması yapılmamıştır. Bu çalışma da ise parametrik kopula modelleri ile üretilmiş veri setlerinin yanısıra üç adet ikili bağımlı gerçek hasar veri seti için de uyum iyiliği karşılaştırılması yapılmıştır.

Tezin uygulama aşamasında Türkiye’de hayat dışı sigorta alanında faaliyet göstermekte olan bir sigorta şirketinden ödenen hasar tutarlarına ait veriler kullanılmıştır. Öncelikle sigorta branşlarına ait ödenen hasarlar arasındaki korelasyon incelenmiş, korelasyonun %99,5 güven düzeyinde anlamlı olduğu kasko-trafik, trafik-bireysel sağlık, kasko-imm sigortaları arasındaki bağımlılıkların modellenmesine karar verilmiş, bu verilere ait uygun dağılım ve parametreleri belirlenmiştir. Kopula model seçimi için uyum iyiliği ölçümü olarak öklid uzaklıklarına bakılmış ve deneysel kopula değerleri ile tahmini kopula olasılık değerleri arasındaki göreceli uzaklıklar elde edilmiştir. Bernstein kopula, parametrik kopula modelleri ile üretilmiş veri setlerinin tahmininde başarısız sonuçlar verirken gerçek veri setlerinde daha başarılı sonuçlar vermesi, gerçek veri setlerindeki bağımlılık yapısının homojen ve simetrik olmaktan daha uzak olması ile açıklanmıştır. Farklı grid sayılarının seçilmesinin, Bernstein kopulanın tahmin sonuçlarını etkilediği görülmektedir. Bu nedenle Bernstein kopula modellemesinde optimal grid sayısının belirlenmesinin önemi ortaya çıkmıştır. Optimal grid sayılarındaki bu farklılıklar hasar veri setlerinin simetriklik ve homojenlik açısından karakteristik özelliklerinden kaynaklandığı düşünülmüştür. Risk değerlendirmesi olarak VaR değerlerine bakılmış, Bernstein kopula ile yapılan tahminlerin kasko-trafik gerçek verisi için daha uyumlu sonuç verdiği düşünüldüğünde, toplam hasar verisinin risk değerinin diğer kopulalar ile ya da bağımsızlık durumunda elde edilmesinin riskin eksik değerlendirilmesine neden olacağı ortaya çıkmaktadır. Risk değerlendirmesindeki eksiklik ya da fazlalık şirketin mali durumunu olumsuz etkileyecek potansiyele sahiptir. Dolayısıyla bağımlılık riskinin uyumluluk ölçütleri açısından en iyi kopula ile modellenip bu kopuladan VaR değerlerinin tahmin edilmesi sigorta şirketleri açısından önemlidir.

Çalışmanın sonuçları sigorta şirketleri ve düzenleyicilerinin risk modellemelerinde Bernstein kopula modelini de dikkate almalarının yararlı olabileceğini göstermektedir. Literatürde Bernstein kopulanın diğer kopulalar ile uyum iyiliği karşılaştırılmasının gerçek veri üzerinden yapıldığı bir çalışmaya rastlanmamıştır. Bu açıdan bilimsel literatüre de bir katkısı bulunmaktadır. Literatüre ve sektöre katkı sağlayacağı düşünülen bu çalışmanın ileride Bernstein kopula modeli ile ilgili yapılacak olan çalışmalar için bir örnek teşkil edeceği düşünülmektedir. Ayrıca Bernstein kopula ile farklı bağımlılık yapılarına sahip veri setleri kullanılarak optimal grid sayısına ilişkin çeşitli değerlendirmeler yapılabilir.

KAYNAKLAR

- [1] Nelsen, R.B., *An Introduction to Copulas*, second edition, Springer, **2006**.
- [2] Karakaş, A.M., and Doğan, M., Archimedean Copula Parameter Estimation with Kendall Distribution Function Kendall. *Iğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 7(3):187–198, **2017**.
- [3] Purvono, Y., Copula Inference for Multiple Lives Analysis Preliminaries. *Submitted to the 13th East Asian Actuarial Conference*, 392–407, **2009**.
- [4] Wang, S., Aggregation of Correlated Risk Portfolios: Models and Algorithms. *Proceedings of the Casualty Actuarial Society*, 85(163): 848-939, **1998**.
- [5] Frees, E.W., and Valdez, E.A., Understanding Relationships Using Copulas. *North American Actuarial Journal*, 2(1):137-141, **1998**.
- [6] Klugman, S.A., and Parsa, R., Fitting bivariate loss distributions with copulas. *Insurance Mathematics Economics*, 24(1–2):139–148, **1999**.
- [7] Youn, H., and Shemyakin, A., Statistical Aspects of Joint Life Insurance Pricing. *American Statistical Association*, pages 34–38, **1999**.
- [8] Embrechts, P., A. McNeil, and D. Straumann, Correlation and Dependence in Risk Management: Properties and Pitfalls, in: M. A. H. Dempster, ed., *Risk Management: Value at Risk and Beyond*, (Cambridge, UK) Cambridge University Press, 176-223, **2002**.
- [9] Embrechts, P., Lindskog, F., and McNeil, A., Modelling dependence with copulas and applications to risk management, pages 329–384, **2001**.
- [10] Genest, C., and Rivest, P., Statistical Inference Procedures for Bivariate Archimedean Copulas. *American Statistical Association*, 88(423):1034–1043, **1993**.
- [11] Genest, C., Rivest, P., and Ghoudi, K., A Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distributions. *Biometrika*, 82(3):543–552, **1995**.
- [12] Genest, C., and Mackay, J., The Joy of Copulas: Bivariate Distributions with Uniform Marginals. *American Statistical Association*, 40(4):280–283, **1986**.
- [13] Oakes, D., Bivariate Survival Models Induced by Frailties. *American Statistical Association*, 84(406):487–493, **1989**.
- [14] Oakes, D., Multivariate survival distributions. *Journal of Nonparametric Statistics*, 3(3–4):343–354, **1994**.
- [15] Frees, E.W., Valdes, E., and Carriere, J., Annuity Valuation with Dependent Mortality. *American Statistical Association*, 63(2):229–261, **1996**.
- [16] Shemyakin, A., and Youn, H., Copula Models of Joint Survival Analysis, **2014**.
- [17] Youn, H., Shemyakin, A., Herman, E., Thomas, S., and Paul, S., A Re-examination of the Joint Mortality Functions. *North American Actuarial Journal*, 6(1):166–170, **2002**.
- [18] Anderson J.E., Time-Dependent Association Measures for Bivariate Survival Distributions. *American Statistical Association*, 87(419), 641–650, **1992**.
- [19] Joe, H., *Dependence Modeling with Copulas*, first edition, A Chapman and Hall

Book, **2015**.

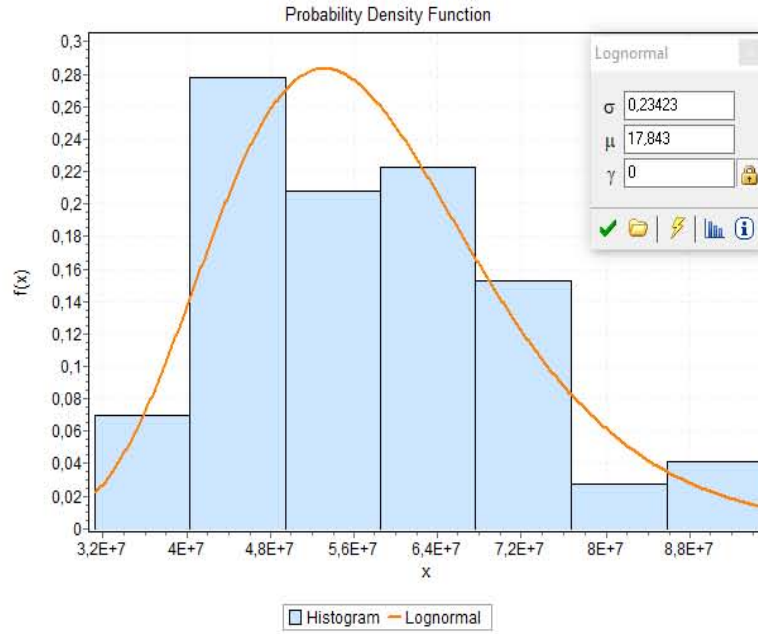
- [20] Demarta, S., and Mcneil, A.J., The t Copula and Related Copulas. *International Statistical Review*, 73(1):111–129, **2005**.
- [21] Aas, K., Modeling The Dependence Structures of Financial Assets: A Survey of Four Copulas. *Samba*, 22(04): 18, **2004**.
- [22] Luo, X., and Shevchenko, P.V., The t copula with Multiple Parameters of Degrees of Freedom: Bivariate Characteristics and Application to Risk Management, *Quantitative Finance*, 10(9):1–31, **2009**.
- [23] Sancetta, A., and Satchell, S., Bernstein Approximations to the Copula Function and Portfolio Optimization. *Cambridge Working Papers in Economics*, **2001**.
- [24] Sancetta, A., and Satchell, S., The Bernstein Copula and Its Applications to Modeling and Approximations of Multivariate Distributions. *Econometric Theory*, 20(3): 535–562, **2004**.
- [25] Diers, D., Eling, M., and Marek, S.D., Dependence modeling in non-life insurance using the Bernstein copula. *Insurance Mathematics Economics*, 50(3):430–436, **2012**.
- [26] Cottin, C., and Pfeifer, D., From Bernstein Polynomials to Bernstein Copulas. **2014**.
- [27] Yang, J., Chen, Z., Wang, F., and Wang, R., Composite Bernstein Copulas. *Astin Bulletin*, 45(02): 445–475, **2015**.
- [28] Schmidt, T., Coping with Copulas.
https://www.researchgate.net/publication/228876267_Coping_with_copulas. (May 15, **2018**).
- [29] Mcneil, A., Frey, R., and Embrechts, P., *Quantitative Risk Management: Concepts, Techniques, and Tools*, Princeton Series in Finance, **2005**.
- [30] Menger, K., Statistical Metrics. *Proceedings of the National Academy of Sciences of the United States of America*, 28(12): 526–529, **1942**.
- [31] Molina, J.J.Q., What are copulas? *Monografías del Semin. Matem. García de Galdeano*, 27: 499–506, **2003**.
- [32] Nelsen, R.B., Properties and applications of copulas: A brief survey. *First Brazilian Conference on Statistical Modelling Insurance Financial*, 3: 1–18, **2003**.
- [33] Kort, J., Modeling tail dependence using copulas — literature review.
<https://pdfs.semanticscholar.org/0253/ff4324bc96e32d715aaf6fe54dbdbecf8d0>. (May 15 **2018**).
- [34] Rüschendorf, L., *Mathematical Risk Analysis*, first edition, Springer, **2013**.
- [35] Bickel, E.P., Diggle, P., Fienberg, S., and Gather, U., *Copula Theory and Its Applications*, first edition, Springer, **2010**.
- [36] Li, J., Modelling dependency between different lines of business with copulas.
https://fbe.unimelb.edu.au/_data/assets/pdf_file/0003/2592021/146. (May 15 **2018**).
- [37] C. C., Theory and Application of Copula Econometrics.
www.math.pku.edu.cn/.../Lecture09_MAOGY_copula. (May 15 **2018**).
- [38] Joy, K.I., Bernstein polynomials, On-Line Geometric Modeling Notes.
http://en.wikipedia.org/wiki/Bernstein_polynomial. (May 15 **2018**).

- [39] Rose, D., *Modeling and estimating multivariate dependence structures with the Bernstein copula*, Doktora, Der Ludwig–Maximilians–Universität, München, **2015**.
- [40] Bernstein, S., Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités, *Communications of the Kharkov Mathematical Society*, 13(1–2), **1912**.
- [41] Pfeifer, P.D., Straburger, D., and Jörg, P., Modelling and simulation of dependence structures in nonlife insurance with Bernstein copulas. Working Paper. Carl von Ossietzky University, Oldenburg, **2009**.

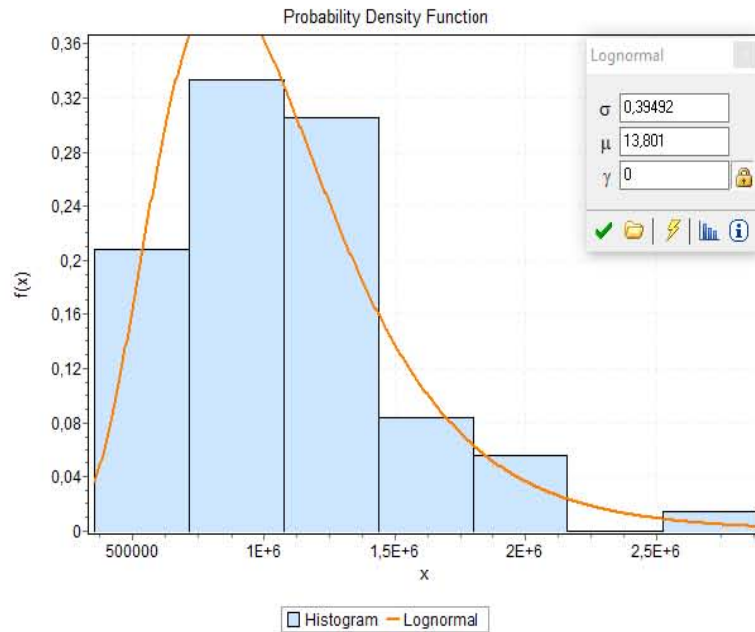
EKLER

EK 1: Hasar Verilerine Ait Dağılım ve Parametre Tahminleri

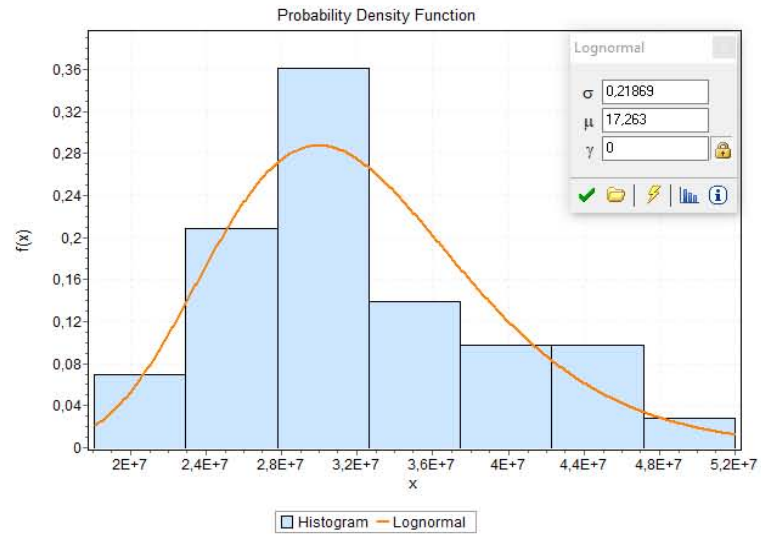
KASKO



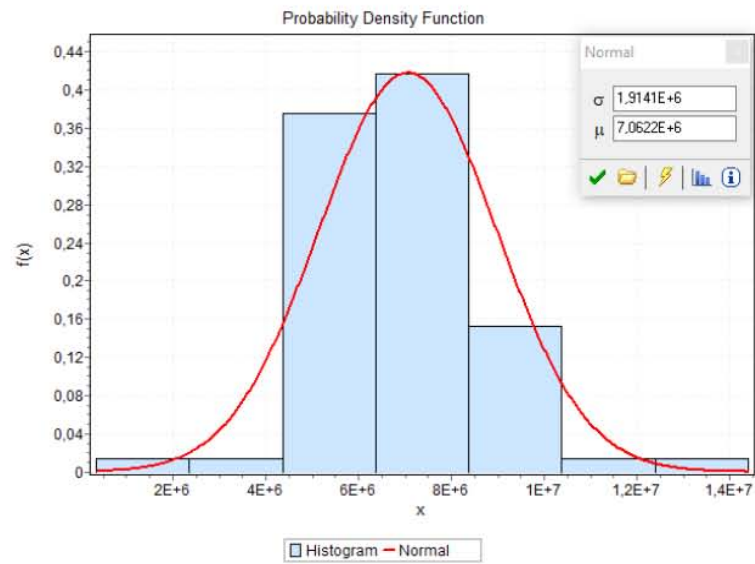
İHTİYARİ MALİ MESULİYET



TRAFİK



SAĞLIK



ÖZGEÇMİŞ

Kimlik Bilgileri

Adı Soyadı : Pınar ÖZER UYAR

Doğum yeri : Alanya

Medeni Hali : Evli

E-posta : pnarozerm@gmail.com

Adresi : Hacettepe Üniversitesi Aktüerya Bilimleri Bölümü

Eğitim

Lise : 2005-2009 Hasan Çolak Anadolu Lisesi

Lisans : 2009-2014 Hacettepe Üniversitesi Matematik Bölümü

Yüksek lisans : 2014-2018 Hacettepe Üniversitesi Aktüerya Bilimleri Bölümü

Yabancı Dil ve Düzeyi

İngilizce, İleri

Almanca, Başlangıç

İş Deneyimi

2016-2017 Bahçeşehir Koleji Matematik Öğretmeni

Deneyim Alanları

Tezden Üretilmiş Projeler ve Bütçesi

Tezden Üretilmiş Yayınlar

Tezden Üretilmiş Tebliğ ve/veya Poster Sunumu ile Katıldığı Toplantılar



HACETTEPE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
YÜKSEK LİSANS/DOKTORA TEZ ÇALIŞMASI ORJİNALLİK RAPORU

HACETTEPE ÜNİVERSİTESİ
FEN BİLİMLER ENSTİTÜSÜ
AKTÜERYA BİLİMLERİ ANABİLİM DALI BAŞKANLIĞI'NA

Tarih: 11.06.2018

Tez Başlığı / Konusu: BERNSTEIN KOPULA İLE HAYAT DIŞI SİGORTASINDA BAĞIMLILIĞIN MODELLENMESİ

Yukarıda başlığı/konusu gösterilen tez çalışmamın a) Kapak sayfası, b) Giriş, c) Ana bölümler d) Sonuç kısımlarından oluşan toplam 43 sayfalık kısmına ilişkin, 11/06/2018 tarihinde tez danışmanım tarafından Turnitin adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı % 14 'tür.

Uygulanan filtrelemeler:

- 1- Kaynakça hariç
- 2- Alıntılar dâhil
- 3- 5 kelimedenden daha az örtüşme içeren metin kısımları hariç

Hacettepe Üniversitesi Fen Bilimleri Enstitüsü Tez Çalışması Orjinallik Raporu Alınması ve Kullanılması Uygulama Esasları'nı inceledim ve bu Uygulama Esasları'nda belirtilen azami benzerlik oranlarına göre tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Gereğini saygılarımla arz ederim.

11.06.2018
P.O.

Tarih ve İmza

Adı Soyadı: PINAR ÖZER UYAR
Öğrenci No: N14229766
Anabilim Dalı: AKTÜERYA BİLİMLERİ BÖLÜMÜ
Programı: AKTÜERYA BİLİMLERİ BÖLÜMÜ
Statüsü: ☒ Y.Lisans ☐ Doktora ☐ Bütünleşik Dr.

DANIŞMAN ONAYI

UYGUNDUR.

DR. ÖĞR. ÜYESİ MURAT
BÜYÜKYAZICI

(Unvan, Ad Soyad, İmza)