**Hafta 1.**

**Genel Bilgi**

Sigortalanan riskin beklenen maliyetini karşılayacak primi yani risk primini hesaplayabilmek için beklenen hasar sayısı ile beklenen hasar tutarının bilinmesi gerekir.

Sigortacılıkta düşük tutarlı hasarların yanı sıra yüksek tutarlı hasarların da gözlenmesi söz konusu olduğundan, hasar tutarları modellenirken yüksek tutarlı hasarların ortaya çıkma olasılığı da dikkate alınmalıdır. Başka bir deyişle dağılımın kuyruk bölgesine önem verilmelidir.

Uygulamada hasar verilerine ilişkin bilgiler (örneklem bilgisi) kullanılarak elde edilen deneysel dağılım fonksiyonu ile işlem yapılmasına rağmen veriyi en iyi temsil eden( uyum sağlayan) matematiksel özellikleri bilinen bir olasılık dağılımının belirlenmesi önemlidir.

Hasar tutarı ya da hasar sayısına ilişkin bir veri setinin ilk olarak betimsel istatistikleri değerlendirilerek ön inceleme yapılır. Daha sonra bilinen bir dağılıma uygunluğu araştırılır. Bunun için dağılımın parametreleri tahmin edilir.

Veri setine ilişkin betimsel istatistikler;

Konum ölçüleri olarak örneklem ortalaması, mod ve medyan, yayılım ölçüleri olarak varyans, standart sapma, çarpıklık ve basıklık katsayısı yaygın kullanılan istatistiklerdir.

Dağılım parametrelerinin tahmininde en çok olabilirlik tahmin yöntemi, momentler tahmin yöntemi ya da en küçük kareler yönteminden yararlanılabilir.

Parametre tahmininden sonra seçilen dağılıma uygunluğunun testi için Kolmogorov Simirnov, Ki-kare uyum iyiliği testi, Anderson Darling testi, Akaike Bilgi Kriteri (AIC) gibi testlerden yararlanılabilir.

Öte yandan sigortacılıkta düşük tutarlı hasarların yanı sıra yüksek tutarlı hasarların da gözlenmesi söz konusu olduğundan, hasar tutarları modellenirken yüksek tutarlı hasarların ortaya çıkma olasılığı da dikkate alınmalıdır. Başka bir deyişle dağılımın kuyruk bölgesine önem verilmesi gerektiğini belirtmiştik.

Bu bağlamda hasar tutarları modellenirken sıklıkla Pareto, Lognormal gibi uzun kuyruklu dağılımlar kullanılmaktadır. Benzer şekilde hasar sayı verileri genellikle aşırı yayılım gösterdiğinden hasar sayısı verisine Poisson dağılımının iyi uyum sağlamadığı, beklenen hasar sayısı sabit olmayıp bir risk biriminden diğerine değişkenlik gösterdiğinden beklenen hasar sayısının da bir rastlantı değişkeni olarak düşünülmesi ile elde edilen Negatif Binom gibi karışık dağılımlardan birinin daha iyi uyum sağladığı görülmüştür.

Mevcut veriye iyi uyum sağlayan yani veriyi en iyi şekilde temsil eden dağılımın belirlenmesi geleceğe ilişkin sağlıklı tahminlerin yapılmasını mümkün kılacağından veriye klasik dağılımların uyum sağlamaması durumunda dönüşüm yöntemi, karışık dağılım yöntemi, bileşik dağılım yöntemi ve spliced yöntem kullanılarak yeni dağılımlar elde edilir.

Bundan sonra hasar sayısı ve hasar tutarları modellenirken yaygın kullanılan dağılımları inceleyecek ve bu dağılımlardan bazılarının yukarıda bahsedilen yöntemlerle elde edilişini gösterecek ve parametre tahminlerini elde edeceğiz.

**Veri Kavramı**

İstatistik bilimi, ölçüm ya da gözlemlerin toplanması, işlenmesi, sunulması ve yorumlanması işlerini tüm yönleri ile ele alır. **Veri** ise istatistik biliminin ilgilendiği temel materyaldir. Veri, gerçek dünyada var olan ancak araştırmacı tarafından bilinmeyen ve ortaya çıkarılması amaçlanan bilgiyi taşır. Bu nedenle, ilk olarak veri kavramı doğru bir biçimde anlaşılmalıdır.

## Veri Türleri

Temel olarak iki tür veri vardır.

VERİ

Nicel veri, üzerinde toplama, çıkarma, çarpma, bölme gibi aritmetik işlemlerin yapılabildiği veri türüdür ve sayıları içerir. Örneğin,

kişilerin yaşları,

hizmet süreleri,

kandaki bir bileşenin miktarı,

hava sıcaklığı,

**hasar tutarı gibi.**

Burada değinilen nicel verilere aritmetik işlemler uygulanarak anlamlı başka veriler elde edilebilir.

Nicel veriler;

Sürekli nicel veri

Kesikli nicel veri

olmak üzere iki sınıfa ayrılabilir.

Sürekli veriler ölçüm ile elde edilen verilerdir. Sürekli verilere örnek olarak **hasar tutarı (miktarı, şiddeti),** boy uzunluğu, hava sıcaklığı, bir sınavdan alınan not, yaş verilebilir. Kesikli veriler ise sayımla elde edilen veridir. **Hasar sayısı,** personel sayısı, birim zamanda emekli olan kişi sayısı, bir kavşakta meydana gelen trafik kazası sayısı kesikli nicel veriye örnektir.

**Nitel veri** ise verinin sayısal olmayan gruplara bölünmesinden elde edilir. Örneğin,

Cinsiyet;

bir kişinin erkek ya da kadın olmasına ilişkin veri erkek için 1, kadın için 2 kodlaması yapılarak kategorik veriye dönüştürülebilir. Burada, 1 ve 2 cinsiyetin düzeyleri olarak adlandırılır.

Akademik unvanlar;

Akademik unvanlara ilişkin veri de kategorik veridir. Düzeyleri; Profesör, Doçent; Dr. Öğretim Üyesi.

Bir kişinin profesör, doçent ya da dr. öğretim üyesi olması kişi profesör ise 1, doçent ise 2 ve yardımcı doçent ise 3 kodlaması yapılarak kategorik veriye dönüştürülebilir. Burada akademik unvanlar 1, 2 ve 3 düzeylerine sahiptir. Bu düzeylere aritmetik işlemlerin uygulanması anlamlı değildir; 1 + 2 gibi bir matematiksel işlemin sonucu anlamlı değildir.

Memnuniyet düzeyi,

**Sigorta türleri gibi.**

Veri türleri daha ayrıntılandırılarak;

Nominal(sınıflanmış),

Ordinal (sıralanmış),

Aralıklı

Orantısal veri

olarak ifade edilebilir.

**Verinin Toplanması**

Araştırma, herhangi bir konuda sorunların incelenmesi için yapılan çalışmadır. Araştırma sonucunun güvenilir olması için araştırmacının bazı kuralları dikkate alması gerekir. Bir araştırmanın aşamaları üç grupta toplanabilir:

1. Araştırma konusunun saptanması,
2. Araştırmanın planlanması,
3. Araştırmanın uygulanması.

Veri toplama yöntemine göre araştırma türleri anket, gözlem, görüşme, deney ve taramaya dayalı araştırmalar olarak sıralanabilir. Tarama yöntemi belge tarama ve bilgi tarama olarak ikiye ayrılabilir. Zaman ekseni dikkate alındığında araştırmalarda veriler iki şekilde toplanır. Veriler gelecekteki olaylar gözlenerek toplanıyorsa ileriye dönük (prospective), geçmişte olmuş olaylar gözlenerek toplanıyorsa geriye dönük (retrospective) çalışma yapılmış olur. Örneğin, geriye dönük bir çalışmada emekli olmuş kişiler ve bu kişilerin emekli oldukları tarihler araştırılabilir. Emeklilik ile ilgili bir kanunun yürürlüğe girmesinden sonra emekli olan kişilerin ele alındığı ve kanunun emekli olma üzerinde yaptığı etkinin incelendiği bir çalışma ileriye dönük çalışmadır.

**Araştırma, Kitle, Örneklem**

Araştırma kapsamına giren, aynı özellikleri taşıyan birimlerin tümüne **kitle** denir.Kitleninbüyüklüğü araştırmanın özelliğine göre değişir. Örneğin, nüfus sayımında kitle, tüm Türkiye’deki bireylerdir. Ankara’da yaşayan yüksek okul mezunu kişilerin televizyon programları hakkındaki görüşlerini yansıtan araştırmada kitle, Ankara’da yaşayan ve yüksek okul mezunu olan kişilerin tümüdür.

Bir kitlenin, örnekleme yöntemleri kullanılarak seçilen, aynı özellikleri taşıyan, kitleyi simgeleyebilme özelliğine sahip bir grup biriminin oluşturduğu topluluğa **örneklem** denir. Kitlenin büyüklüğüne bağlı olarak her zaman bireylerin tümünü ayrı ayrı incelemek zaman ve maliyet gibi nedenlerle olanaksız olduğu için araştırma örneklem üzerinde yürütülür ve örneklemden elde edilen bilgilere bağlı kalarak kitlenin bilinmeyen özellikleri için çıkarsama yapılır.

İstatistiksel olarak uygun örneklemleri seçmek için uygulanan yöntemler topluluğuna **örnekleme yöntemleri** denir.

Kitlede ve örneklemde yer alan her bireye / birime **denek** denir. Genellikle kitledeki denek sayısı “N” ile, örneklemdeki denek sayısı ise “n” ile gösterilir.

Kitlenin özelliklerinin sayı ile belirtilen değerine **parametre** denir. Parametreler genellikle bilinmeyen sabitlerdir. Kitledeki değerlerin ortalaması (), kitledeki değerlerin varyansı () kitlenin parametrelerine örnektir. Araştırma kitle üzerinde uygulanmıyorsa ya da kitle parametreleri bilinmiyorsa kitle parametreleri tahmin edilir. Bunun için araştırma kitleden örnekleme yöntemleri kullanılarak alınan bir örneklem üzerinden yürütülür. Örneklemden elde edilen değerlere örneklem değeri (istatistik) denir. Örneklem ortalaması  ve örneklem varyansı  örneklem değerlere örnek olarak gösterilebilir. Örneklem değerin parametreden farkı, parametre gibi sabit olmaması, örneklemden örnekleme farklı değerleri farklı olasılıklarla alan bir raslantı değişkeni olmasıdır. Uygulamada örneklem değerin gözlenen değeri ile ilgilenilir.

Nicel ya da nitel anlamda bir özellik ya da karakterde belirgin olarak görülen farklılık **değişken** ile gösterilir. Boy uzunluğu, ağırlık, yaş, cinsiyet birer değişkendir. Bir değişken, denekten deneğe değişen değerler alır. Değişkenlere karşılık gelen denek değerine **veri** denir.

Veriler değişken sayısına göre tek değişkenli ve çok değişkenli olarak gruplandırılır. Örneğin 50 üniversite öğrencisi bitirdikleri ortaöğretim kurumu türüne göre sınıflanırsa tek değişkenli veri, bu öğrenciler bitirdikleri ortaöğretim kurumu türüne ve cinsiyetlerine göre sınıflanırsa iki değişkenli veri elde edilir. Öğrencilerin bitirdiği okul türü birinci değişken, cinsiyet ise ikinci değişkendir. Bir bölgede yaşayan 150 kişi yaş gruplarına, medeni durumlarına ve cinsiyetlerine göre sınıflandırılırsa üç değişkenli veri elde edilir.

**Örnekleme Yöntemleri**

Kitleden örneklem seçmek için yapılan uygulamaya **örnekleme** denir. Kitle çok büyük ve kitlenin tüm deneklerini inceleme olanağı yoksa örnekleme uygulanır. Elde edilen bilgilerden yararlanılarak *kitle parametreleri* için tahmin yapılır. Kitledeki tüm birimlere ulaşılabiliniyorsa kitlenin parametreleri *tahmin yapılmaksızın* doğrudan hesaplanır.

Örnekleme daha ekonomiktir. Bilgiler daha kısa sürede elde edilir. Bununla birlikte, kullanılacak örnekleme yönteminin verilerin özelliğine göre en doğru şekilde belirlenmesi gerekir.

Çizge 1. Kitle, örneklem ve örnekleme

Olasılığa bağlı örnekleme

**KİTLE**

İlgilenilen verinin tümü

Kitle parametreleri

**ÖRNEKLEM**

İstatistikler

Örnekleme dahil olan veri

Olasılığa bağlı olmayan örnekleme

Örnekleme

Yöntemi

**Olasılığa bağlı olmayan örnekleme** yöntemlerinde N denekli kitleden kişisel görüşe göre rasgele olarak n denek seçilip örneklem oluşturulur. **Olasılığa bağlı olan örnekleme** yöntemlerinde ise kitleden belli olasılık kuralları uyarınca çekilen deneklerden bir örneklem oluşturulur. Olasılığa bağlı olan örnekleme yöntemleri ile olasılığa bağlı olmayan örnekleme yöntemlerine göre kitleyi temsil edebilme yeteneği daha yüksek olan bir örneklem oluşturulabilir.

Olasılığa bağlı örnekleme yöntemleri genel olarak Çizge 2’de açıklanmıştır.

Olasılığa Bağlı Örnekleme Yöntemleri

Basit Rasgele Örnekleme

Tabakalı Örnekleme

Sistematik Örnekleme

Yerine koyarak

(Eşit olasılık)

Yerine koymadan

(Eşit olmayan olasılık)

Çizge 2. Olasılığa bağlı örnekleme yöntemleri

**Kaynak:** Prof. Dr. Canan Hamurkaroğlu