Örnek: Bir Boeing 720 uçağının havalandırma sisteminin ardışık bozulmalar arası geçen zamanları saat olarak aşağıdaki gibi gözlemlenmiştir.

413(ilk) 14 53 37 100 65 9 169 447 184 36

201 118 34 31 18 67 57 62 7 22 34(son)

Bu veriler birbirinden bağımsız ve $1/\lambda$ parametreli üstel aynı dağılımlı ise bu verilerin modellenmesinde λ oranlı Poisson süreci önerilebilir ya da birbirinden bağımsız ve aynı dağılımlı değil ise homojen olmayan Poisson süreci önerilebilir. Sizce hangi model bu veri seti için uygundur?

H_0: Homojen Poisson Süreci

H_1: Homojen olmayan Poisson Süreci

olarak verilen hipotezleri göz önüne alalım. $S_n = \sum_{i=1}^{n} X_i$, n.varş zamanı olmak üzere H_0 hipotezi altında ilk $(n-1)$ varış zamanı olan $S_1, S_2, \dots, S_{n-1}$ rasgele değişkenlerine $(0, S_n]$ aralığından alınmış düzgün dağılıma sahip $(n-1)$ tane rasgele değişkenin sıra istatistikleri olarak bakılabilir. Bu durumda merkezi limit teoremi yardımıyla

$$U = \frac{\frac{\sum_{i=1}^{n-1} S_i}{n-1} - \frac{S_n}{2}}{S_n\sqrt{1/12(n-1)}}$$

ile verilen test istatistiği büyük n için yaklaşık olarak standart normal dağılıma sahiptir.

S_i'ler aşağıdaki gibi toplayarak elde edilir.

413 427 522 622 687 696 365 1312 1416 1502 1733

1851 1916 1934 1952 2019 2076 2138 2145 2167 2201

Not: Veri setinin modellenmesinden önce o veride bir trendin var (gözlemlerin aynı dağılımdan gelmediği) olup olmadığını anlamak için basit bir grafiksel yönteme başvurulur. (S_n, n) noktaları koordinat sisteminde belirtilir. Elde edilen noktalar lineer bir doğru etrafında yayılım gösteriyorsa o veri seti için bir trendin var olmadığı düşünülebilir. Trendin var olduğu durumda uygun sayma süreci modelinin tespiti için Laplace testi olarak bilinen yukarıdaki test uygulanabilir.

Ödev: Yukarıdaki veri kümesi için grafiği çizerek bir trendin var olduğunu gözleyiniz.

Test istatistiğinde ihtiyaç duyulan değerler $\sum_{i=1}^{n-1} S_i = 30843$ ve $S_n = 2201$ olmak üzere

$$U = \frac{\frac{30843}{22} - \frac{2201}{2}}{2201\sqrt{\frac{1}{12(12)}}} = 2.2254$$

bulunur. $\alpha = 0.05$ anlamlılık düzeyinde $Z_\alpha = 1.96$ ve $U = 2.2254 > 1.96$ olduğundan H_0 hipotezi red edilir. Bu veri seti için uygun modelin homojen olmayan Poisson süreci olduğu söylenir. Şimdi bu sürece ait şiddet fonksiyonunun belirlenmesi gerekmektedir.

b)Uygulamalarda en sık karşılaşılan şiddet fonksiyonu yapıları

1) $\lambda(t) = e^{\alpha+\beta t} \quad t \geq 0 \quad , -\infty < \alpha, \beta < \infty$

2) $\lambda(t) = \alpha\beta t^{\beta-1} \quad t \geq 0\,, \quad \alpha, \beta > 0$

dır. Varsayalım ki veri seti (1) şiddet fonksiyonu yapısına sahip olsun. Bu durumda şiddet fonksiyonunun en çok olabilirlik tahmin edicisini bulalım.

Genel anlamda homojen olmayan bir Poisson süreci için olabilirlik fonksiyonu

$$L = \prod_{i=1}^{n} \lambda(S_i)\, e^{-m(t)}$$

ile verilir.

Ödev. Yukarıda verilen olabilirlik fonksiyonunu elde ediniz.

$\lambda(t) = e^{\alpha+\beta t}$ alınmasıyla

$$M(t) = \int_0^t e^{\alpha+\beta s}\, ds = \frac{e^\alpha}{\beta}\left(e^{\beta t} - 1\right)$$

olarak bulunur. Buradan olabilirlik fonksiyonunun logaritması

$$\ln L = \frac{e^\alpha}{\beta}\left(1 - e^{\beta t}\right) + \sum_{i=1}^{n}(\alpha + \beta S_i) = \frac{e^\alpha}{\beta}\left(1 - e^{\beta t}\right) + n\alpha + \beta \sum_{i=1}^{n} S_i$$

dır. α ve β 'nın en çok olabilirlik tahmin edicilerini bulmak için parametrelere göre türev alalım.

$\frac{d \ln L}{d\alpha} = \frac{e^\alpha}{\beta}\left(1 - e^{\beta t}\right) + n$

olup

$n = -\frac{e^\alpha}{\beta}\left(1 - e^{\beta t}\right) \Rightarrow \frac{n\beta}{e^{\beta t - 1}} = e^\alpha$

$\Rightarrow \hat{\alpha} = \ln\left(\frac{n\beta}{e^{\beta t - 1}}\right)$

olarak bulunur. Benzer olarak

$\frac{d \ln L}{d\beta} = -\frac{e^\alpha}{\beta^2}\left(1 - e^{\beta t}\right) + \frac{e^\alpha}{\beta}\left(-e^{\beta t}\right)t + \sum_{i=1}^{n} S_i$

olup

$\frac{n}{\beta} - \frac{ne^{\beta t}}{(e^{\beta t}-1)} + \sum_{i=1}^{n} S_i = 0$

denkleminin çözülmesiyle β ‘nın en çok olabilirlik tahmin edicisi elde edilir.

Ödev

1. α ve β ‘nın en çok olabilirlik tahmin edicileri olabilirlik fonksiyonunu maksimum yapan değerlerdir. Buna göre herhangi bir paket programı kullanarak bu tahminleri bulunuz.

2. İlk 1 saat içerisinde hiç bozulmama olasılığını bulunuz.

3. Varsayalım ki $\lambda(t) = \alpha\beta t^{\beta-1}$ olsun. Bu durumda parametrelerin tahmin edicilerini teorik olarak bulunuz. Daha sonra herhangi bir paket programı kullanarak bu tahmin edicileri elde ediniz. 2 nolu ödevde verilen olasılığı tekrar hesaplayınız.