Süleyman Demirel Üniversitesi,
Fen Bilimleri Enstitüsü Dergisi,
16-3 (2012), 318-326

Üstel Dağılıma Uygunluk İçin Bazı Uyum İyiliği Testlerinin I.Tip Hata ve Güçleri Bakımından Kıyaslanmaları

Can KÖLE[1], Fikri GÖKPINAR[*1]
[1]*Gazi Üniversitesi, Fen Fakültesi, İstatistik Bölümü / ANKARA*
Alınış Tarihi:19.09.2012, Kabul Tarihi:15.11.2012

Özet: İstatistiksel bir modelin uyum iyiliği gözlenen bir veri setinin istatistiksel modele uyumluluğunu test eder. Bu çalışmada Kolmogorov-Smirnov(1939), Lilliefors(1969), Anderson-Darling(1954), Entropiye dayalı uyum iyiliği testleri (Ebrahimi(1991) , Choi (2004), Correa (1995), Noughabi(2010)) ve Fortiana(2002) 'nın geliştirdikleri uyum iyiliği testleri üstel dağılım için incelenmiştir. Ayrıca bu test istatistikleri deneysel I.tip hata olasılığı ve testin gücü bakımından farklı dağılım şekilleri altında karşılaştırılmıştır.

Anahtar Kelimeler : Kolmogorov- Smirnov- Lilliefors testi, Anderson-Darling testi, Entropi, Fortiana testi.

Comparison of Type-One Errors and Powers of Some Goodness of Fit Tests for Exponentiality

Abstract: The goodness of fit of a statistical model tests how well it fits a set of observations. In this study, some goodness of fit tests called Kolmogorov- Smirnov, Anderson- Darling, based on different entropy estimates tests Ebrahimi(1991) , Choi (2004), Correa (1995), Noughabi(2010)) and Fortiana(2002) and also the tests given by Fortiana(2002) are investigated. Also,these test statistics are compared according to their type 1-error rates and powers under different distribution types.

Keywords : Kolmogorov- Smirnov test , Anderson-Darling test, Entropy, Fortiana test.

Giriş

X rasgele değişkeni λ parametreli üstel dağılıma sahip olsun. Bu durumda olasılık yoğunluk fonksiyonu $f(x|\lambda) = \lambda e^{-\lambda x}, x > 0 \; ve \; \lambda > 0$ şeklinde olur. Bu dağılımda $E(X) = 1/\lambda$'dır. Üstel dağılım kuyruk teorisi, stokastik modeller, güvenilirlik teorisi gibi popüler konularda sıkça kullanılmaktadır. Bu sebepten dolayı herhangi bir örnek verisinin bu dağılımdan gelip gelmediğini bilmek önemlidir. Herhangi bir örnek verisinin belirli bir dağılımdan gelip gelmediği uyum iyiliği testleri ile sınanır. Bu anlamda uyum iyiliği testleri gözlemlenen verilerin hangi istatistiksel modele uyum sağladığını sınamada kullanılır.

1900 'lü yılların başından itibaren uyum iyiliği testleri ile ilgili birçok araştırma yapılmıştır. Pearson (1900) kategorik ve kesikli veriler için beklenen ve gözlenen örnek değerlerinin kıyaslanmasına dayalı olan bir yöntem geliştirmiştir. Kolmogorov-Smirnov(1933,1939), Cramer Von Mises(1928,1930), Anderson Darling(1954), Lilliefors(1969) empirik dağılımlara dayalı ve sürekli dağılımlar için daha güçlü sonuçlar veren uyum iyiliği testleri elde etmişlerdir.

Gail(1978) özel olarak üstel dağılım için ölçekten bağımsız bir test istatistiği geliştirmiş ve Lorenz, Shapiro Wilks ve Pietra istatistikleriyle güç karşılaştırmasını yapmıştır (Gail, 1978).

Son yıllarda özel olarak üstel dağılım için uyum iyiliği testlerinde entropiye dayalı yaklaşımlar ön plana çıkmıştır. Bir bilgi teorisi yöntemi olan ve değişkenler hakkında ön bilgi sağlayıcı bir yöntem olan Entropi, uyum iyiliği testlerinde kullanılan yöntemlerden biridir. Bu yaklaşımda çalışmalarıyla ön plana çıkan araştırmacılar; Ebrahimi(1991) , Choi (2004), Correa

(1995) ve Noughabi(2010) olarak literatürde yer almışlardır. Ayrıca Baratpour ve Rad (2012) üstel dağılıma uyum iyiliği testi için birikimli artık entropiye dayalı bir test istatistiği önermiş ve bunu diğer test istatistikleri ile karşılaştırmıştır. Noughabi ve Arghami (2012) Kullback-Leibler bilgisine dayalı bir uyum iyiliği testi üretmiş ve diğer test istatistikleri ile simülasyon yoluyla karşılaştırmıştır.

Fortiana ve Grane(2002) Hoeffding'in maksimum korelasyonuna dayalı yeni bir uyum iyiliği istatistiği geliştirmişler ve bunu üstel dağılım için Shapiro-Wilk ve Gini istatistikleriyle kıyaslamışlardır. Geliştirdikleri test istatistiğinin gücünün, parametrelerin yokluk(test) hipotezindeki değerlerden uzaklaştıkça Shapiro-Wilk (Wilk, 1972) ve Gini (Gail, 1978) istatistiklerinin gücünden daha düşük olduğunu ancak parametrelerin küçük değerleri için bu iki istatistik kadar güçlü olduğunu göstermiştir.

Bu çalışmada Kolmogorov-Smirnov(1939), Lilliefors (1969), Anderson-Darling (1954), Entropiye dayalı uyum iyiliği testleri (Ebrahimi (1991) , Choi (2004), Correa (1995), Noughabi (2010)) ve Fortiana (2002) 'nın geliştirdikleri uyum iyiliği testleri üstel dağılım için incelenmiştir. Ayrıca bu test istatistikleri deneysel I.tip hata olasılığı ve testin gücü bakımından farklı dağılım şekilleri altında simülasyon yoluyla karşılaştırılmış ve elde edilen sonuçlar yorumlanmıştır.

*fikri@gazi.edu.tr

Uyum İyiliği Testleri

Kolmogorov-Smirnov ve Lilliefors Uyum İyiliği Testi

Kolmogorov-Smirnov test istatistiği deneysel dağılım fonksiyonuna dayalı uyum iyiliği testleri arasında en sık kullanılan yöntemlerden birisidir. Bu test istatistiği ilk olarak Kolmogorov tarafından geliştirilmiş ancak uyum iyiliği testlerinde kullanılması Smirnov tarafından uyarlanmıştır. Bu sebeple literatürde Kolmogorov-Smirnov(K-S) uyum iyiliği testleri olarak geçmektedir.

Önceden belirli bir dağılım olduğu varsayılan kümülatif dağılım fonksiyonu $F_0(x)$ olsun. Yani belirli bir x değerinden küçük ya da eşit değerli verilerin olasılığını $F_0(x)$ fonksiyonu versin. n adet gözlenen veri içerisinde belirli bir x değerinden küçük ya da eşit olanların oranını veren kümülatif deneysel dağılım fonksiyonu ise $S_n(x)$ olsun. K-S uyum iyiliği testi mantığına göre eğer ki deneysel dağılım fonksiyonu sonuçları hipotetik $F_0(x)$'e yeterince yakın değilse, gözlemlenen verilerin dağılımı hipotetik dağılımdan gelmemektedir sonucuna varılır. Yani gözlemlenen veriler hipotezde geldiği öne sürülen dağılıma uymamaktadır. Bu durumun sınanması için öne sürülen test istatistiği de;
$D = \mathrm{MAX}_x |F_0{}^*(x) - S_n(x)|$ olarak tanımlanmıştır (Kolmogorov, 1933), (Smirnov, 1939) D istatistiği $F_0(x)$ sürekli ve tamamen bilinirken test edilen hipotetik dağılımdan bağımsızdır.

Bu istatistiğin dağılımı, dağılımın tüm parametreleri bilinirken elde edilmekte aksi halde D istatistiğinin dağılımı bulunamamaktadır. Lilliefors(1969), yığına ait bilinmeyen parametrelerin gözlemlerden tahmin edilmesi soncunda üstel dağılım için D istatistiğine ait yeni kritik değerlerin bir düzenlemesini sunmuştur. Hazırladığı prosedürde verilen n hacimli örnek için D test istatistiğini;
$$D = \underset{x}{\mathrm{MAX}} |F^*{}_0(x) - S_n(x)|$$

şeklinde tanımlamıştır (Lilliefors, 1969). Burada $S_n(x)$ kümülatif örnek dağılımı ve $F^*{}_0$, $\frac{1}{\lambda} = \bar{x}$ örnek ortalamalı, kümülatif üstel dağılım fonksiyonudur. Eğer hesaplanan D değeri tablodaki kritik değeri aşıyorsa dağılımın üstel dağılımdan geldiği hipotezi reddedilir.

Anderson-Darling Uyum İyiliği Testi

Kolmogorov-Smirnov uyum iyiliği testlerinde varsayılan dağılım ile deneysel dağılım arasındaki fark araştırılırken, üst ve alt kuyruk kısımlarında yeterli duyarlılığa erişilememiştir. Ancak Anderson-Darling uyum iyiliği testinde kuyruk kısımlarında yeterince hassalığa erişilmiş ve daha güçlü test sonuçları elde edilmiştir.

Anderson-Darling(A-D) uyum iyiliği testi Cramer-von-Mises test istatistiğinden türetilmiştir.
W^2'nin ampirik dağılım ve spesifik dağılım arasındaki farkın 0'a eğilim gösterdiği $x \to -\infty$ ve $x \to +\infty$ iken duyarsızlığı önlemek ve uç noktalardaki önyargıları ve belirsizlikleri gidermek için Anderson-Darling yeni bir istatistik geliştirmiştir. Cramer-Von Mises test istatistiğinden farklı bir ağırlıklandırma yöntemi uygulayan A-D testi genel formu ise şu şekilde sunulmuştur (Darling, 1954);
$$A^2 = n \int_{-\infty}^{+\infty} \frac{[S_n(x) - F_0(x)]^2}{[F_0(x)(1 - F_0(x))]} d\,F_0(x)$$

Burada kullanılan ağırlık fonksiyonu olan $[F_0(x)(1-F_0(x))]^{-1}$ bu test istatistiğini Cramer-Von Mises dan daha güçlü ve uçlarda daha hassas hale getirmiştir(Anderson ve Darling,1954).

Hesaplama formülü olarak;

$$A^2 = -n - \frac{1}{n}\sum_{j=1}^{n}(2j-1)[\log\left(F\left(x_{(j)}\right)\right) + \log(1 - F\left(x_{(n+1-j)}\right))]$$

kullanılır. Burada logaritmalar doğal tabanlı ve $x_{(j)}$ sıralı istatistiklerdir.

Önceleri bu test istatistiğinin de uygulanabilmesi için dağılımın parametrelerinin tahmin edilmemiş olması gerekmekteydi. Parametrelerin tahmin edildiği durumda A-D test istatistiğinin limit dağılımı tahmin edilen parametrelere göre değişiklik gösterir (Stephans, 1976, 1986). Bu yeni dağılıma göre yüzdelik alanlar parametrelerin bilindiği duruma göre çok daha küçülür. Parametrelerin tahmin edilmesi durumunda oluşan bu yeni kritik değerler ve test istatistiğinin dağılımı Stephans(1976, 1986) tarafından oluşturulmuştur.

Entropiye Dayalı Uyum İyiliği Testleri

H(f) entropi fonksiyonu için parametrik olmayan tahminler Vasieck(1976), Theil(1980), Dudewicz ve Van der Meulen(1987) , Van Es(1992) , Ebrahimi ve diğerleri (1994) , Correa(1995) ve Wieczorkowski-Grzegorewski(1999) gibi birçok araştırmacı tarafından tartışılmıştır.
$$H(f) = -\int_{-\infty}^{+\infty} f(x)\log\big(f(x)\big)\,dx$$

Ebrahimi Uyum İyiliği Testi;

Çeşitli tahmin ediciler arasında en geniş çaplı kullanılan Vasieck' in örnek entropisidir.
Vasieck tahmin edicisi ((N.Ebrahimi, 1992));

$$HV_{mn} = \frac{1}{n}\sum_{i=1}^{n}\log\left\{\frac{n}{2m}\left(X_{(i+m)} - X_{(i-m)}\right)\right\} ;$$
burada m çerçeve genişliği, pozitif ve n/2 den küçük bir tamsayıdır. i-m< X1 ise $_{(i-m)}=X_{(1)}$,i>m+n ise $X_{(i)}=X_{(n)}$ olur ve $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$ rasgele örnek büyüklüğü n' e bağlı sıra istatistikleri olur. Vasieck in entropisi yığın entropisi H(f) için tutarlıdır.

Test istatiği elde etmek için Ebrahimi vd(1991) Kullback-Lieber bilgi fonksiyonundan faydalanmışlardır.

$$I(F,F_0)=\int_0^{\infty} f(x)\log\left(\frac{f(x)}{f_0(x)}\right)dx$$
$$=-H(f)-\int_{-\infty}^{\infty} f(x)\log f_0(x)dx$$

K-L bilgi fonksiyonu f(x) ve $f_0(x)$ arasındaki farklılığın bir ölçüsüdür ve bu fonksiyonun tahmin edicisi, kimi araştırmacılar tarafından uyum iyiliği testlerinde kullanılmıştır. Ebrahimi vd(1991) üstel dağılımın testi için şu istatistiği önermişlerdir;

$$TV_{mn}=\frac{\exp(HV_{mn})}{\exp(\log(\overline{X})+1)}$$

Choi Uyum İyiliği Testi;

Choi ve ark. üstelliğin testi için Vasieck entropi tahmin edicisinin geliştirilmiş versiyonunu sunmuşlardır (B. Choi, 2004)

$$HVE_{mn}=\frac{1}{n-m}\sum_{i=1}^{n-m}\left(\frac{n+1}{m}\left(X_{(i+m)}-X_{(i)}\right)\right)+\sum_{k=m}^{n}\frac{1}{k}+\log(m)-\log(n+1)$$

Verilen entropi tahmin edicisine dayalı olarak test istatistiği;

$$TVE_{mn}=\frac{\exp(HVE_{mn})}{\exp(\log(\overline{X})+1)}$$

şeklindedir.

Correa Uyum İyiliği Testi;

Correa entropi tahmin edicisi ((Correa, 1995), (Es, 1992));

$$HC_{mn}=-\frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{\sum_{j=i-m}^{i+m}\left(X_{(j)}-\overline{X}_{(i)}\right)(j-1)}{n\sum_{j=i-m}^{i+m}\left(X_{(j)}-\overline{X}_{(i)}\right)^2}\right);$$

Burada ; $\overline{X}_{(i)}=\frac{1}{2m+1}\sum_{j=i-m}^{i+m}X_{(j)}$ olmak üzere entropi tahmin edicisine dayalı olarak test istatistiği;

$$TC_{mn}=\frac{\exp(HC_{mn})}{\exp(\log(\overline{X})+1)}$$

şeklindedir. Bu istatistik ölçekteki dönüşümlere göre değişmezdir.

Modifiye-Ebrahimi Uyum İyiliği Testi;

Ebrahimi ve ark. tekrar düzenledikleri örnek entropisi tahmin edicisini şu şekilde sunmuşlardır (N. Ebrahimi,K.Ppflughoeft ve E. Soofi,1994);

$$HE_{mn}=\frac{1}{n}\sum_{i=1}^{n}log\left(\frac{n}{c_i m}\left(X_{(i+m)}-X_{(i-m)}\right)\right);$$

$$c_i=\begin{cases}1+\frac{i-1}{m} & ,\quad 1\le i\le m\\ 2 & ,m+1\le i\le n-m\\ 1+\frac{n-i}{m} & ,n-m+1\le i\le n\end{cases}$$

Verilen entropi tahmin edicisine dayalı olarak test istatistiği;

$$TE_{mn}=\frac{exp(HE_{mn})}{exp(log(\overline{X})+1)}$$

şeklindedir.

Alizadeh Uyum İyiliği Testi;

Alizadeh de yaptığı araştırmada istatistikte uyum iyiliği testlerinde sıkça kullanılan bir yöntem olan entropiden faydalanmıştır. Yazar araştırmasında kendi entropi tahmin edicisini geliştirmiş ve entropiye dayanan test istatistiği yardımı ile, ilgili dağılımın üstel dağılımdan gelip gelmediğini test etmiştir. Ayrıca Alizadeh bu araştırmada daha önce geliştirilen entropi tahmin edicilerini de kullanarak her bir testin gücünü farklı dağılım alternatifleri bakımından araştırmış ve kendi yönteminin daha etkin ve güçlü olduğunu göstermiştir (H.A Noughabi, 2011).

Alizadeh Noughabi entropi tahmin edicisi olarak aşağıdaki istatistiği geliştirmiştir (Noughabi, 1992);

$$HA_{mn}=-\frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{\hat{f}\left(X_{(i+m)}\right)+\hat{f}\left(X_{(i-m)}\right)}{2}\right),$$

$$\hat{f}(X_i)=\frac{1}{nh}\sum_{j=1}^{n}k\left(\frac{X_i-X_j}{h}\right),$$

k, Kernel fonksiyonu, standart normal yoğunluk fonksiyonu olarak seçildi ve genişlik, h normal optimum düzgün formül olarak seçildi; s, standart hata olmak üzere $h=1.06\,sn^{-1/5}$ dır.
Aşağıda HA_{mn}ye dayalı üstelliğin testiyle ilgili istatistik verilmiştir (Noughabi ve Arghami, 2011).

$$TA_{mn}=\frac{\exp(HA_{mn})}{\exp(\log(\overline{X})+1)}.$$

Tüm bu Entropiye dayalı uyum iyiliği test istatistikleri parametreden bağımsızdır.

Fortiana Uyum İyiliği Testi

Fortiana yaptığı araştırmada üstel dağılım için Hoeffding' in en büyük korelasyonuna dayalı bir uyum iyiliği istatistiği vermiştir. Hoeffding maksimum korelasyonu $\rho^{+}(F_1, F_2)$ olarak tanımlanır. Frechet sınıfının üst sınırı, Hoeffding maksimum korelasyonu olarak tanımlanmıştır. Frechet Sınıfı stokastik süreçlerde kullanılan, bağımlılık etkisine ilişkin alt ve üst kuyruk değerlerinin tahmininde kullanılan bir yöntemdir.

Cuadras ve Fortiana(1993); verilen dağılım F_0 ve deneysel dağılım S_n için $\rho^{+}(S_n, F_0)$ istatistiğini bağımsız rasgele değişkenler için uyum iyiliği testlerinde nitel ölçüm olarak düzenlemişlerdir. Yapılan bu araştırmada yazarlar bu test istatistiğinin özellikleri üzerinde çalışmıştır.

Fortiana ve Grane (2002) üstel dağılıma uyum iyiliği için $\rho^{+}(S_n, F_0)$ fonksiyonundan türetilen ve aşağıdaki gibi tanımlanan test istatistiğini önermiştir .

$$Q_n = \sum_{j=1}^{n} l_{nj} y_{(j)} \Big/ \sum_{j=1}^{n} y_{(j)} \quad \text{ve}$$

$l_{nj} = (n-j)\log(n-j) - (n-j+1)\log(n-j+1) + \log(n)$ ve burada 0log0≡0 olarak tanımlanmıştır.

Fortiana ve Grane (2002) bu Q_n istatistiğinin olasılık yoğunluk fonksiyonu ise aşağıdaki şekilde tanımlanmıştır

$$f(t) = (n-1)\sum_{j=1}^{n-1} a_{nj}^{-3} \prod_{k=1,k\neq j}^{n-1} \left(a_{nj} - a_{nk}\right)^{-1} \left(1 - \frac{t}{a_{nj}}\right)^{n-2} \mathbf{1}_{(0,a_{nj})}(t)$$

Burada $a_{nj} = \log\left(\frac{n}{n-j}\right)$ ve j=1,…,n-1 şeklindedir ayrıca **1** gösterge fonksiyonunu belirtir. Bu olasılık yoğunluk fonksiyonundan faydalanarak istenilen değerlere ilişkin kritik değerler hesaplanabilir. Örnek çapının küçük değerleri için bu fonksiyon kullanılarak kritik değerler hesaplanabilmekte iken örnek çapı arttıkça bu fonksiyonu kullanmak oldukça zor ve vakit alıcıdır bunun yerine bu çalışmada Monte Carlo yaklaşımıyla kritik değerler elde edilmiştir.

Simülasyon Çalışması

Bu bölümde testler 1.tip hata ve testin gücü bakımından kıyaslanmıştır. Bu kıyaslama için nominal alfa değeri 0.05 olarak alınmıştır. Testin gücü ve deneysel alfa değerlerini elde etmek için 100000 deneme yapılmıştır.Bu simulasyon çalışması MATLAB R2008A programı kullanılarak oluşturulmuştur. Deneysel I.tip hata değerleri üstel(1) dağılımında üretilen verilerden elde edilen test istatistiklerinin red sayılarının toplam deneme sayısına bölünmesiyle elde edilmiştir. Testin gücü için ise Üstel dağılım gibi (0,∞) arasında tanımlı ve aşağıda verilen çeşitli parametreli Gamma, Weibull ve Rayleigh dağılımlarından faydalanılarak, deneysel I.tip hata hesaplanırken kullanılan süreç kullanılmıştır. Bununla birlikte bu durumda farklı olarak veriler Gamma, Weibull ve Rayleigh' ın çeşitli parametreleri altında üretilmiş ve test istatistikleri hesaplanarak red sayıları toplam deneme sayısına oranlanmıştır. Ayrıca Entropiye dayalı testlerde çerçeve genişliğini belirten m değerleri tüm tablolarda $m = [\![n/2]\!]$ tam değer fonksiyonu olarak alınmıştır. Çizelge 2. de ilgili dağılımlara ilişkin deneysel α değerleri sonuçlar tablolar halinde verilmiştir. Çizelge 3. ve Çizelge 4. de Gamma dağılımına ilişkin testin gücü sonuçları verilmiştir. Çizelge 5.- Çizelge 8. arasında ise Weibull dağılımı için testin gücü sonuçları verilmiştir. Çizelge 9.- Çizelge 12. arasında Rayleigh dağılımına ilişkin test güçleri verilmiştir.

Çizelge 1. *Kullanılan dağılımlar, olasılık yoğunluk fonksiyonları ve parametreleri*

Dağılım	Parametreler	Olasılık Yoğunluk Fonksiyonu
Üstel(a)	(1)	$y = \frac{1}{a} e^{-\frac{x}{a}}$
Weibull(a,b)	(1,2) (1,3) (2,2) (2,3)	$y = ba^{-b}x^{b-1}e^{-(\frac{x}{a})^{b}}$
Gamma(a,b)	(2,1) (3,1)	$y = \frac{1}{b^{a}\Gamma(a)} x^{a-1} e^{-\frac{x}{b}}$ $\Gamma(a) = (a-1)!$
Rayleigh(b)	(1) (2) (3) (4)	$y = \frac{x}{b^{2}} e^{(\frac{-x^{2}}{2b^{2}})}$

Çizelge 2. *α=0.05 iken Üstel dağılım altında uyum iyiliği testlerinin deneysel I.tip hata olasılıkları*

n	Fortiana	Alizadeh	Modified-Ebrahimi	Choi	Ebrahimi	Correa	Liliefors	Anderson-Darling
10	0.0491	0.0483	0.0472	0.0475	0.0472	0.0480	0.0489	0.0483
15	0.0500	0.0478	0.0490	0.0491	0.0490	0.0484	0.0488	0.0487
20	0.0495	0.0526	0.0508	0.0523	0.0508	0.0503	0.0508	0.0503
25	0.0515	0.0527	0.0501	0.0519	0.0501	0.0499	0.0513	0.0507
30	0.0504	0.0499	0.0487	0.0521	0.0487	0.0485	0.0501	0.0491
50	0.0495	0.0497	0.0505	0.0494	0.0505	0.0488	0.0515	0.0498

Çizelge 2. de deneysel I.tip hata olasılıkları verilmiştir.Nominal α değeri 0.05 olarak alınmış ve hesaplamalar buna göre yapılmıştır. I.tip hata değerleri kıyaslanacak olursa 0.05 olarak alınan nominal α değerine genel olarak tüm testler yaklaşık sonuçlar vermiştir.

Çizelge 3*. Gamma(2,1) dağılımı için elde edilen testin gücü sonuçları*

n	Fortiana	Alizadeh	Modified-Ebrahimi	Choi	Ebrahimi	Correa	Liliefors	Anderson-Darling
10	0.1884	0.3204	0.3182	0.2772	0.3182	0.3154	0.2240	0.1874
15	0.2890	0.5004	0.4716	0.3720	0.4716	0.4730	0.3074	0.3188
20	0.3776	0.5738	0.5440	0.5072	0.5440	0.5402	0.4104	0.4438
25	0.4452	0.6802	0.6022	0.5946	0.6022	0.5932	0.4976	0.5654
30	0.5386	0.7324	0.6494	0.6720	0.6494	0.6348	0.5972	0.6852
50	0.7758	0.8644	0.7326	0.8760	0.7326	0.7172	0.8370	0.9176

Çizelge 4. *Gamma(3,1) dağılımı için testin gücü sonuçları*

n	Fortiana	Alizadeh	Modified-Ebrahimi	Choi	Ebrahimi	Correa	Liliefors	Anderson-Darling
10	0.4258	0.6730	0.6674	0.5608	0.6674	0.6720	0.4602	0.4560
15	0.6234	0.8584	0.8088	0.7258	0.8088	0.7950	0.6524	0.7144
20	0.7716	0.9310	0.8952	0.8664	0.8952	0.8872	0.8134	0.8858
25	0.8606	0.9832	0.9304	0.9382	0.9304	0.9110	0.9104	0.9640
30	0.9220	0.9866	0.9450	0.9744	0.9450	0.9328	0.9558	0.9864
50	0.9932	0.9982	0.9748	0.9988	0.9748	0.9648	0.9988	0.9998

Elimizdeki gözlemler Gamma(2,1) dağılımından geldiği durum için, yanlış olan H_0 hipotezini reddetme olasılığı olan testin gücü sonuçları bulunmaktadır. Sonuçlardan aşikardır ki örnek çapı arttıkça testin gücü artmaktadır. Buna karşın simülasyon çalışmasına göre örnek çapı küçükken en yüksek testin gücü değerini Alizadeh testi vermiştir. Buna karşın n=50 iken en güçlü test, 0.9176 testin gücü değeri ile Anderson-Darling olmuştur. Burada göze çarpan diğer bir sonuç ise Fortiana uyum iyiliği testinin genel olarak Gamma(2,1) için zayıf sonuçlar verdiğidir.Gamma(3,1) için genel olarak elde edilen testin gücü simülasyon çalışmaları Gamma(2,1) için elde edilen testin güçlerinden daha yüksektir. Bunun sebebi de a=3 parametresidir. Bu parametre arttığı için dağılım üstel dağılımdan daha da uzaklaşmış ve testin gücü sonuçlarında genel olarak bir artma gözlenmiştir. Bu sonuçlar içerisinde n=10 iken en güçlü test , Alizadeh uyum iyiliği testi olmuştur.n=50 iken ise Alizadeh, Choi, Lilliefors ve Anderson Darling testleri aynı güç sonucuna ulaşmışlardır.

Çizelge 5. Weibull(1,2) dağılımı için testin gücü sonuçları

n	Fortiana	Alizadeh	Modified-Ebrahimi	Choi	Ebrahimi	Correa	Liliefors	Anderson-Darling
10	0.5596	0.7358	0.7590	0.5968	0.7590	0.7532	0.5190	0.5324
15	0.8046	0.8966	0.9082	0.7800	0.9082	0.9096	0.7110	0.7902
20	0.9188	0.9390	0.9582	0.8984	0.9582	0.9564	0.8524	0.9200
25	0.9672	0.9776	0.9846	0.9436	0.9846	0.9820	0.9258	0.9708
30	0.9908	0.9792	0.9926	0.9806	0.9926	0.9902	0.9648	0.9944
50	0.9998	0.9958	0.9988	0.9996	0.9988	0.9982	0.9988	1.000

Çizelge 6. *Weibull(1,3) için testin gücü sonuçları*

n	Fortiana	Alizadeh	Modified-Ebrahimi	Choi	Ebrahimi	Correa	Liliefors	Anderson-Darling
10	0.9642	0.9856	0.9924	0.9576	0.9924	0.9916	0.9052	0.9506
15	0.9992	0.9992	0.9998	0.9952	0.9998	0.9998	0.9848	0.9978
20	1	1	1	0.9994	1	1	0.9988	0.9998
25	1	1	1	1	1	1	1	1
30	1	1	1	1	1	1	1	1
50	1	1	1	1	1	1	1	1

Çizelge 7. *Weibull(2,2) dağılımı için elde edilen güç sonuçları*

n	Fortiana	Alizadeh	Modified-Ebrahimi	Choi	Ebrahimi	Correa	Liliefors	Anderson-Darling
10	0.5492	0.6794	0.7238	0.5750	0.7238	0.7296	0.5026	0.5236
15	0.7898	0.8808	0.8986	0.7766	0.8986	0.8892	0.7112	0.7822
20	0.9186	0.9314	0.9626	0.9024	0.9626	0.9622	0.8486	0.9208
25	0.9708	0.9718	0.9808	0.9412	0.9808	0.9778	0.9272	0.9726
30	0.9892	0.9800	0.9908	0.9738	0.9908	0.9874	0.9624	0.9916
50	1	0.9944	0.9988	0.9996	0.9988	0.9986	0.9990	1

Çizelge 8. *Weibull(2,3) dağılımı için testin gücü sonuçları*

n	Fortiana	Alizadeh	Modified-Ebrahimi	Choi	Ebrahimi	Correa	Liliefors	Anderson-Darling
10	0.5726	0.7160	0.7562	0.6028	0.7562	0.7618	0.5208	0.5354
15	0.7964	0.8948	0.9124	0.7958	0.9124	0.9068	0.6986	0.7880
20	0.9088	0.9318	0.9498	0.8976	0.9498	0.9500	0.8552	0.9212
25	0.9714	0.9742	0.9822	0.9398	0.9822	0.9804	0.9228	0.9738
30	0.9912	0.9800	0.9914	0.9764	0.9914	0.9902	0.9648	0.9922
50	1	0.9956	0.9986	1	0.9986	0.9980	1	1

Weibull(1,2) parametreli dağılım için elde edilen (Çizelge 5) sonuçları (Çizelge 4) – (Çizelge 3) ve (Çizelge 2)' de ki güç sonuçlarına göre daha yüksek çıkmıştır. Bu da ilk 3 dağılıma göre bu dağılımın üstel dağılımdan daha uzak olduğunu ifade etmektedir. Bunun sebebi Weibull dağılımının özellikle b şekil parametresi 1den büyük iken dağılımın şeklinin simetriğe yakın olmasından kaynaklanmaktadır. Örnek çapı küçük olduğunda Entropiye dayalı uyum iyiliği testleri kıyaslanan diğer test istatistiklerine göre daha yüksek güç değerleri vermiştir. Örnek çapı arttıkça diğer test istatistikleri de yüksek testin gücü değerleri vermiştir.

Çizelge 6.' da ki sonuçlara genel olarak bakıldığında Weibull(1,3) dağılımının üstel dağılımdan iyice uzaklaştığı için bu dağılıma ilişkin deneysel test güçleri genel olarak yüksek çıkmıştır. Tabloda ki sonuçlara bakıldığında en zayıf olarak göze çarpan test istatistikleri Lilliefors ve Choi uyum iyiliği testleri olmuştur. Örnek çapı 25 olduktan sonra tüm testlerin gücü 1 olarak hesaplanmıştır.

Weibull(2,2) dağılımı genel olarak Weibull(1,2) dağılımında elde edilen sonuçlara paralel güç sonuçları vermiştir. Örnek çapı 10 iken en güçlü test 0.7296 ile Correa uyum iyiliği testi olmuştur. Fortiana bu dağılımda da küçük örnek çapında zayıf sonuç vermiştir. Buna karşın küçük örnek çapında en zayıf sonuç Lilliefors uyum iyiliği testinde elde edilmiştir. Örnek çapı 50 olduğunda en güçlü test istatistikleri Fortiana - Grane, Anderson-Darling olarak göze çarpmıştır.

Weibull(2,3) dağılımı Weibull(1,3) ile kıyaslandığında genel olarak daha zayıf test güçleri vermiştir. Küçük örnek çapında Entropiye dayalı uyum iyiliği testleri olan Alizadeh, Modified Ebrahimi, Choi, Ebrahimi ve Correa uyum iyiliği testleri daha yüksek güç değerlerine sahiptir. Örnek çapı arttıkça tüm testlerin güç değerleri 1'e oldukça yakın sonuçlar vermektedir**.**

Çizelge 9. *Rayleigh(1) dağılımı için testin gücü sonuçları*

N	Fortiana	Alizadeh	Modified-Ebrahimi	Choi	Ebrahimi	Correa	Liliefors	Anderson-Darling
10	0.5582	0.7448	0.7696	0.5958	0.7696	0.7724	0.5048	0.5266
15	0.7944	0.8886	0.9024	0.7710	0.9024	0.8988	0.7114	0.7898
20	0.9124	0.9294	0.9566	0.9064	0.9566	0.9582	0.8514	0.9244
25	0.9690	0.9724	0.9840	0.9404	0.9840	0.9826	0.9192	0.9696
30	0.9888	0.9796	0.9890	0.9764	0.9890	0.9862	0.9656	0.9908
50	0.9996	0.9928	0.9982	0.9992	0.9982	0.9980	0.9986	1

Çizelge 10. *Rayleigh(2) dağılımı için testin gücü sonuçları*

N	Fortiana	Alizadeh	Modified-Ebrahimi	Choi	Ebrahimi	Correa	Liliefors	Anderson-Darling
10	0.5588	0.7210	0.7264	0.5992	0.7264	0.7298	0.5036	0.5264
15	0.7940	0.8958	0.9062	0.7698	0.9062	0.9058	0.7020	0.7854
20	0.9134	0.9402	0.9572	0.8964	0.9572	0.9576	0.8530	0.9176
25	0.9692	0.9750	0.9806	0.9434	0.9806	0.9790	0.9336	0.9726
30	0.9878	0.9822	0.9884	0.9758	0.9884	0.9862	0.9668	0.9924
50	1	0.9948	0.9984	0.9992	0.9984	0.9974	0.9994	1

Çizelge 11. *Rayleigh(3) dağılımı için testin gücü sonuçları*

n	Fortiana	Alizadeh	Modified-Ebrahimi	Choi	Ebrahimi	Correa	Liliefors	Anderson-Darling
10	0.5532	0.7184	0.7424	0.6122	0.7424	0.7428	0.5042	0.5200
15	0.7918	0.8798	0.8944	0.7510	0.8944	0.8922	0.7126	0.7860
20	0.9112	0.9398	0.9584	0.8984	0.9584	0.9554	0.8498	0.9148
25	0.9664	0.9794	0.9834	0.9414	0.9834	0.9806	0.9190	0.9738
30	0.9906	0.9762	0.9912	0.9694	0.9912	0.9912	0.9648	0.9918
50	1	0.9932	0.9992	0.9994	0.9992	0.9986	0.9994	1

Çizelge 12. *Rayleigh(4) dağılımı için testin gücü sonuçları*

n	Fortiana	Alizadeh	Modified-Ebrahimi	Choi	Ebrahimi	Correa	Liliefors	Anderson-Darling
10	0.5550	0.7144	0.7274	0.6068	0.7274	0.7420	0.5164	0.5334
15	0.7926	0.8954	0.9102	0.7738	0.9102	0.9030	0.7022	0.7826
20	0.9164	0.9372	0.9588	0.8970	0.9588	0.9548	0.8478	0.9174
25	0.9696	0.9734	0.9848	0.9404	0.9848	0.9824	0.9274	0.9722
30	0.9898	0.9804	0.9900	0.9772	0.9900	0.9888	0.9650	0.9926
50	0.9998	0.9952	0.9982	0.9994	0.9982	0.9976	0.9996	1

Genel olarak Rayleigh dağılımına ilişkin parametreye bağlı durumlara bakıldığında testlerin gücü birbirine paralel sonuçlar vermiştir. Yani Rayleigh(1), Rayleigh(2), Rayleigh(3), Rayleigh(4) dağılımları birbirine yakın güç sonuçları vermiştir. Dolayısıyla bu dört durum genel olarak yorumlanabilir. Örneğin Rayleigh(4) durumu için Çizelge 12.' ye bakılırsa örnek çapı 10 iken en güçlü sonuçları Entropiye dayalı uyum iyiliği testleri vermiştir Örnek çapı arttıkça incelenen tüm uyum iyiliği testleri birbirine yakın sonuçlar vermiştir.

Sonuç

Genel olarak çizelgeler incelendiğinde tüm uyum iyiliği testleri için örnek çapı arttıkça ve dağılım karakteristik olarak üstel dağılımdan uzaklaştıkça daha yüksek test güçleri hesaplanmıştır ki bu beklenen bir durumdur. Ayrıca entropiye dayalı testlerin diğerlerine göre daha üstün olduğu görülmektedir.Uyum iyiliği testlerinde karşılaşılan sorunlardan biri, küçük örnek çaplarında genel olarak düşük test gücü sonuçları elde edilmesidir. Bununla birlikte bu çalışma, entropiye dayalı uyum iyiliği testlerinin küçük örnek çaplarında, diğer uyum iyiliği testlerine göre nispeten daha yüksek güç değerlerine sahip olduğunu göstermiştir. Ampirik dağılım fonksiyonuna dayalı testler içinde, Anderson-Darling uyum iyiliği testinin tüm durumlar için Lilliefors uyum iyiliği testinden daha güçlü olduğu sonuçlardan gözlemlenmiştir. Diğer bir sonuç olarak ise Fortiana uyum iyiliği testinin küçük örnek çaplarında diğer testlerden daha zayıf olduğudur.

Entropiye dayalı testlerle ilgili elde edilen sonuçlara bakıldığında, Noughabi ve Arghami (2011)'de elde edilen sonuçlarla paralellik gösterdiği görülmektedir.

Kaynaklar

Choi, K.K. (2004). Goodness of fit tes for exponentiality based on Kullback-Leibler information. *Commun. Stat. Simul. Comput. 33* , 525-536.

Cochran, W.G. (1952). The chi-square test of goodness of fit. *The Annals of Mathematical Statistics* , 315-345.

Correa, J. (1995). A new estimator of entropy. *Commun. Stat. Theory Methods 24* , 2439-2449.

Cramer, H. (1928). on the composition of elementary errors. *Skand. Aktuar. 11* , 141-180.

Darling, T.W. (1954). A Test of Goodness of Fit. *Journal of the American Statistical Association* , 765-769.

Es, B.V. (1992). Estimating functionals related to a density by class of statistics based on spacings. *Scand. J. Stat* , 61-72.

Noughabi, H. A, Arghami N.R. (2011). Monte Carlo comprassion of five exponentiality test using different entropy estimates. *Journal of Statstical Computataion and Simulation* , 1-14.

Fortiana, A.G. (2002). A scale-free goodness-of-fit statistic for the exponential distrubution based on maximum correlations. *journal of statistical planning and inference* , 85-97.

Kolmogorov, A.N. (1933). Sulla determinazone empirica di une legge di distribuzione. *G. İst. Attuari* , 83-91.

Lilliefors, H. W. (1969). On the Kolmogorov-Smirnov Test for the Exponential Distribution with Mean. *Journal of the American Statistical Association* , 387-389.

Gail, J. L. (1978). A scale-free goodness of fit test for the exponential distribution based on the Gini statistic. *Journal of the Royal Statistical Society* , 350-357.

N.Ebrahimi, M.H. (1992). Testing Exponentiality based on Kullback-Leiber information. *J.R. Statist. Soc. Ser. B54* , 739-748.

Noughabi, H.A (1992). A new estimator of entropy and its application in testing normality. *J.Statis. Soc. Ser. B 54* , 1151-1162.

Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine series 5* , 157-175.

Shannon, C. (1948). A mathematical thory of communications. *Bell Syst. Tech. J. 27* , 379-423.

Smirnov, N. (1939). On the estimation of the discrepancy between emprical curves of distribution for two independent samples. *Bulletin of Moscow University* , 3-16.

Stephens, M.A. (1976). Asymptotic results for goodness of fit statistics with unknown parameters. *The Annals of Statistics* , 357-369.

Wilk, S. (1972). An analysis of variance test for the exponential distrubutions complete samples. *Technometrics* , 355-370.

Noughabi, H.A and Arghami, N. R. General treatment of goodness-of-fit tests based on Kullback–Leibler information. Journal of Statistical Computation and Simulation, DOI: 10.1080/00949655.2012. 667100

Baratpour, S. and Rad, A.H. (2012) Testing Goodness-of-Fit for Exponential Distribution Based on Cumulative Residual Entropy, Communications in Statistics - Theory and Methods, 41(8): 1387-1396