

Benford Kanunu ve Genelleştirilmiş Benford Kanunu ile ekosistem doğallığının hesaplanması

Kürşad Özkan^{a,*} 

Özet : Ekosistemlerin doğallık tespitine yönelik yeni bir yöntem önerilmiştir. Bitki türleri ve örnek alanlardan oluşan (S X A) üç veri seti kullanılmıştır. Bunlar Sultan Dağları Alt Bölgesi (BS) veri seti (60 X 96), Dedegül Dağları Alt Bölgesi (BD) veri seti (89 X 119) ve iki alt bölgenin birleşmesinden oluşan Beyşehir Gölü Havzası (B) veri setidir (98 X 215). İlk olarak BS ve BD veri setlerinden belirlenen ilk rakam olasılık değerleri (d_1p_o) ile Benford Kanunu teorik olasılık değerleri (d_1p_e) arasındaki genel uyumu belirlemek için ki kare (χ^2) testleri yapılmıştır. Sonuçta $\chi^2(e_{BS}) = 16,579$ ve $\chi^2(e_{BD}) = 2,406$ olarak bulunmuştur. İkinci olarak BS ve BD'nin gözlenen olasılık değerlerine en uyumlu teorik olasılık değerlerini belirlemek için genelleştirilmiş Benford Kanunu ($GB(d; \gamma)$) kullanılmıştır. BS ve BD için en küçük χ^2 değerleri sırasıyla $\gamma = 0,65$ 'de ve $\gamma = 0,07$ 'de elde edilmiştir ($\chi^2(e_{BD}^{\gamma}) = 4,992$ ve $\chi^2(e_{BD}^{\gamma}) = 2,209$). Beklendiği gibi genelleştirilmiş Benford Kanunu ile her iki alt bölgenin χ^2 değerleri düşmüştür. χ^2 değer düşüşü BS'de çok daha yüksek olmuştur. Alt bölgelerin örnek alan sayıları birbirlerinden farklıdır. Bu yüzden üçüncü aşamada B veri setinden her iki alt bölgenin örnek alan sayıları dikkate alınarak rastlantısal yinelemeli işlemler uygulanmış ve yineleme sayısı (K) kadar ($K = 10000$) χ^2 değerleri elde edilmiştir. Daha sonra kalibrasyon katsayı değerlerini (kd) belirlemek için bu χ^2 değerlerinin ortalamaları (${}^k\bar{\chi}^2(E_{BS}^{\gamma}) = 6,747$ ve ${}^k\bar{\chi}^2(E_{BD}^{\gamma}) = 6,176$) alınmıştır. Sonuçta, BS için $kd = 1$ olduğu için BS doğallık değeri 4,992 ve BD için $kd = 1,093$ olduğu için BD doğallık değeri 2,414 olarak bulunmuştur. Teorik olarak en doğal ekosistemler için tam doğallık değeri=0 kabul edildiğinden, elde edilen doğallık hesaplama sonuçları BD ekosistemlerinin BS ekosistemlerinden daha doğal olduğu göstermiştir.

Anahtar kelimeler: Hemerobi, Doğallık, Orman ekosistemleri, İlk sayı kanunu, Genelleştirme, Rastgele seçim

Estimating ecosystem naturalness using Benford's Law and Generalized Benford's Law

Abstract: A new method was proposed to estimate ecosystem naturalness. Three species-plot (S X A) datasets were used. Those data sets belong to Sultan mountain sub-district (BS) (60 X 96) Dedegül mountain sub-district (BD) (89 X 119) and, Beyşehir Watershed (B) (98 X 215) consisting of both of the sub-districts. Firstly, chi square test (χ^2) was applied to define the statistical goodness of fit between the first digit observed probabilities (d_1p_o) and the theoretical probabilities of Benford's Law (d_1p_e). It was found that $\chi^2(e_{BS}) = 16.579$ and $\chi^2(e_{BD}) = 2.406$.

Secondly, to find the fittest theoretical probabilities for BS and BD, generalized Benford's Law ($GB(d; \gamma)$) was applied. Minimal χ^2 values were obtained at $\gamma = 0.65$ and $\gamma = 0.07$ for BS and BD respectively ($\chi^2(e_{BD}^{\gamma}) = 4.992$, $\chi^2(e_{BD}^{\gamma}) = 2.209$). As expected, χ^2 values of the sub-districts decreased by generalized Benford's Law. The most dramatic χ^2 decrease occurred in BS. The number of sample plots of the sub-districts are different. Two random iterative processes happened 10000 times were therefore performed considering the number of sample plots of the sub-districts in B dataset. As a result 10000 χ^2 values were obtained for each sub-district. Average values of those χ^2 values were then used (${}^k\bar{\chi}^2(E_{BS}^{\gamma}) = 6.747$ and ${}^k\bar{\chi}^2(E_{BD}^{\gamma}) = 6.176$) to calculate calibration coefficients of each sub-district. Naturalness values of BS and BD were found to be 4.992 and 2.414 respectively due to calibration coefficients of $BS = {}^k\bar{\chi}^2(E_{max}^{\gamma}) / {}^k\bar{\chi}^2(E_{BS}^{\gamma}) = 1$ and $BD = {}^k\bar{\chi}^2(E_{max}^{\gamma}) / {}^k\bar{\chi}^2(E_{BD}^{\gamma}) = 1.093$. Since the perfect naturalness value is theoretically equal to 0, the obtained results indicate that BD ecosystems are more natural than BS ecosystems.

Keywords: Hemeroby, Naturalness, Forest ecosystems, First digit rule, Generalization, Randomization

1. Giriş

Doğallık insan etkisi haricindeki her türlü süreç, durum veya sistem olarak tanımlanmaktadır (Anderson, 1991). Haliyle bir ekosistemin doğallık derecesini belirlemek onun temelde insan etkisine ne kadar maruz kaldığını ifade etmek (hemerobi) anlamına gelmektedir. İnsanlar ekosistemleri kesme, otlatma, alan açma, yakma ve hatta silvikültürel

müdahaleler (özellikle ağaçlandırma faaliyetleri) ile etkilemektedir. Bu da ekosistemlerin gittikçe doğallığından uzaklaşmasına sebep olmaktadır. Genel bir kabul olarak söylenebilir ki, doğal sürecin kendi çıktısı haricindeki tüm çıktıları doğallık denkleminin negatif tarafında yer almaktadır.

Doğallık, koruma potansiyeli veya ekolojik değer tespitinde kullanılması gereken girdilerden biridir (Tans,

✉ ^a Isparta Uygulamalı Bilimler Üniversitesi, Orman Fakültesi, Orman Mühendisliği Bölümü, Isparta

@ ^{*} **Corresponding author** (İletişim yazarı): kursadozkan@isparta.edu.tr

✓ **Received** (Geliş tarihi): 31.03.2021, **Accepted** (Kabul tarihi): 17.05.2021



Citation (Atf): Özkan, K., 2021. Benford Kanunu ve Genelleştirilmiş Benford Kanunu ile ekosistem doğallığının hesaplanması. Turkish Journal of Forestry, 22(2): 73-82. DOI: [10.18182/tjf.907217](https://doi.org/10.18182/tjf.907217)

1974; Margules ve Usher, 1981; Smith ve Theberge, 1986). Ekosistemlerin koruma potansiyeli veya ekolojik değer tespitinde doğallık dışında dikkate alınan diğer girdiler ise, biyoçeşitlilik, nadirlik, alan, temsil kabiliyeti, eğitim değeri, tarihsel değer, bilimsel değer, yaban hayatı potansiyeli, ekolojik hassasiyet, ekolojik birim içindeki pozisyon, potansiyel değer, kullanılabilirlik, değiştirilebilirlik ve yönetim anlayışı olarak ifade edilmektedir (Margules ve Usher, 1981).

Doğallık hesabı genelde kavramsal modellere (Reif ve Walentowski, 2008; Côté vd., 2019), ölçeklendirmeye (Machado, 2004; Fanfarillo vd., 2018), gösterge değer tespitine (Erdős vd., 2017) veya süksesyon-klimaks oranına (Dunlop vd., 2014) dayandırılmaktadır. Bu yaklaşımlarla doğallığın tespiti subjektif, zaman alıcı ve masraflı olmaktadır.

Ekolojik değer tespitinde yer alan girdilerden doğal ekosistemin güncel durumunu en iyi şekilde açıklayan parametreler biyoçeşitlilik ve nadirliktir. Biyoçeşitlilik ve nadirlik hesaplarını gerçekleştirmek için türlerin frekans veya bolluk değerlerine dayalı olarak geliştirilmiş birçok indis bulunmaktadır (Özkan, 2016; Leroy, 2016; Fattorini, 2003).

Biyoçeşitlilik ve nadirlik parametrelerinin hesaplanmasında aynı veri havuzu kullanılabilir. Doğallığın da bu ortak veri havuzundan hesaplanmasına yönelik bir yaklaşım önermek hem zaman, kaynak ve emek tasarrufunun sağlanması hem de bir sisteminin büyük oranda tespit edilebilecek koruma değerinin veya nitelik derecesinin biyoçeşitlilik ve nadirlik hesabının yapıldığı aynı veri kaynağından bir bütün olarak değerlendirilmesi açısından önemlidir.

Doğallığı canlı toplum verisine dayanarak değerlendirmek ve bu sayede ekosistemlerin birbirlerine göre doğallık kıyaslamasını gerçekleştirecek bir yaklaşım önermek mümkün olabilir. Bu bağlamda bir kanun bize yardımcı olabilir; doğadaki sayıların eşit dağıldığını gösteren bir kanun...

Şöyle ki;

Bir gökbilimci ve matematikçi olan Simon Newcomb dikkate değer bir gözlem gerçekleştirir. Araştırmacı bir logaritma kitabında ilk sayfaların diğer sayfalarla kıyaslandığında daha çok aşımış olduğunu fark eder. Bu onun fiziksel veya doğal veri setlerindeki sayıların ilk rakamlarının eşit olarak dağıldığını keşfini yapmasını sağlar. Newcomb'un keşfi asıl olarak doğadaki sayıların üstel formda açıklanabileceğini ortaya koymaktadır. Başka bir deyişle sayılar rastgele dağılmamaktadır. Rastgele dağılım - ya da yeknesak dağılım- doğal ya da sosyal sistemlere ait sayıların logaritmik değerlerinden elde edilen mantis değerleri için söz konudur (Newcomb, 1881). Newcomb'un bu keşfi o dönemde bilim camiası tarafından ilgi görmemiştir. Aradan 50 yılı aşkın bir süre geçtikten sonra Newcomb'un keşfi tekrar ele alınır. Frank Benford, Newcomb gibi logaritma kitaplarında aynı gözlemi gerçekleştirir ve oldukça geniş veri seti kullanarak detaylı bir araştırmanın içine girer. Benford'un araştırması kimyasal ve fiziksel sabiteler, başlangıç sayıları, nehir uzunlukları, göl alanları, fibonacci

dizileri, nüfus verileri, hatta kitap ve magazinlerden elde edilen veriler gibi 20 farklı kaynak alandan toplam 20229 sayısal değer kullanılarak gerçekleştirilir. Benford araştırmasını "anormal sayılar kanunu" başlığı altında yayınladı (Benford, 1938). Bu kanun günümüzde "Benford Kanunu" veya "İlk Sayı Kanunu" olarak bilinmektedir.

Benford Kanunu'nda olasılık değerleri sabittir. Bu olasılık değerlerine göre yapılacak değerlendirme bir şablona bağlı kalmak anlamına gelir. Ekosistem verilerinin değerlendirmesinde şablonlar veya eşik değerler pek tercih edilmez. Diğer bir deyişle ekosistem analizlerinde genelde esnek yapıya sahip yöntemler ve tolerans alanına sahip eşik değerler tercih edilir. Ekolojik değerlendirmelere uygun olacak şekilde Benford Kanunu'na esneklik kazandıran form, genelleştirilmiş Benford Kanunu'dur (Hürlimann, 2015).

Bu makalede ekosistemlerin doğallık analizinin ve karşılaştırılmasının Benford Kanunu'na ve genelleştirilmiş Benford Kanunu'na dayandırılarak gerçekleştirilmesi amaçlanmıştır. Bu amaçla, sadece doğallık hesabı için değil, aynı zamanda biyoçeşitlilik ve nadirlik hesaplarının da yapılabileceği gerçek bir veri seti kullanılmıştır. Bu veri seti Beyşehir Gölü Havzası'nda 215 örnek alanda kaydı gerçekleştirilmiş bitki türlerinden oluşan öz nitelik tablosudur (Özkan, 2003).

2. Materyal ve yöntem

2.1. Çalışma materyali

Çalışmada Beyşehir Gölü Havzası'nın (B) iki alt bölgesinden (Sultan Dağları Alt Bölgesi (BS) ve Dedegül Dağları Alt Bölgesi (BD)) elde edilmiş toplam 215 örnek alan (A_B) ve 98 türden (S_B) oluşan veri matrisi kullanılmıştır (Özkan, 2003). BS'de yer alan örnek alan sayısı, $A_{BS} = 96$ olup bu örnek alanlarda kayda geçirilen tür sayısı, $S_{BS} = 60$ 'dir. BD dahilinde yer alan örnek alan sayısı, $A_{BD} = 119$ olup, toplam tür sayısı ise, $S_{BD} = 89$ 'dur (Çizelge 1). Her iki alt bölgedeki ortak tür sayısı 51'dir. Sadece BS'de bulunan tür sayısı 9 iken BD'e özgü tür sayısı 38'dir.

2.2. Tür frekans verileri

Türlerin örnek alanlarda rastlanma değerleri toplamı onların frekans değerlerine (f_i) karşılık gelmektedir. Beyşehir Gölü Havzası (B)'nden alınan örnek alanda 16 türe tek bir örnek alanda rastlanmıştır ($f_i = 16$). Sadece 4 türün rastlanma değeri 100'ü geçmektedir ($f_{103} = 1$, $f_{119} = 1$, $f_{121} = 1$, $f_{134} = 1$). Sultan Dağları Alt Bölgesi (BS)'inde tek bir örnek alanda rastlanan tür sayısı 19 iken ($f_i = 19$), bu değer Dedegül Dağları Alt Bölgesi (BD)'inde 15'dir ($f_i = 15$). BS'de $f_{max} = f_{61} = 1$, BD'de ise $f_{max} = f_{90} = 1$ 'dir (Çizelge 1).

B, BS ve BD için gözlenen veri olasılık değerleri kendilerine ait tür frekans değerlerinin d_i değerlerinden belirlenmiştir.

Çizelge 1. Beyşehir Gölü Havzası (B), Sultan Dağları Alt Bölgesi (BS) ve Dedegül Dağları Alt Bölgesi (BD) içinde toplam ve frekans değerlerine düşen tür sayıları

B		98																		
f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{12}	f_{13}	f_{14}	f_{15}	f_{16}	f_{17}	f_{18}	f_{21}	f_{22}	f_{23}	
16	12	4	8	6	5	3	3	1	3	1	5	2	1	2	1	1	1	2	2	
f_{24}	f_{25}	f_{26}	f_{31}	f_{41}	f_{42}	f_{43}	f_{44}	f_{48}	f_{54}	f_{61}	f_{73}	f_{77}	f_{91}	f_{103}	f_{119}	f_{121}	f_{134}			
1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1			
BS		60																		
f_1	f_2	f_3	f_4	f_5	f_6	f_8	f_9	f_{10}	f_{12}	f_{14}	f_{15}	f_{16}	f_{17}	f_{18}	f_{29}	f_{37}	f_{39}	f_{41}	f_{42}	
19	7	5	2	2	1	2	2	2	4	2	1	1	1	1	1	1	1	1	1	
f_{44}	f_{45}	f_{61}																		
1	1	1																		
BD		89																		
f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}	f_{16}	f_{18}	f_{20}	f_{21}	f_{27}	
15	9	5	6	6	5	3	4	4	2	6	3	1	1	1	1	1	1	1	1	
f_{29}	f_{30}	f_{31}	f_{32}	f_{35}	f_{44}	f_{45}	f_{49}	f_{58}	f_{66}	f_{82}	f_{90}									
1	2	1	1	1	1	1	1	1	1	1	1									

Çalışma da istatistiksel değerlendirmeler bitki türleri (S) ve örnek alanlardan (A) oluşan S X A boyutlarındaki üç veri setinde gerçekleştirilmiştir. Bunlar B (98 X 215), BS (60 X 96) ve BD (89 X 119) veri setleridir. BS ve BD farklı örnek alan sayılarına sahip olduğu için, onların birbiri ile karşılaştırmasında B'nin 96 ve 119'luk yinelemeli rastgele örnek alan seçim ve analiz çıktılarının ortalaması kullanılmıştır.

2.3. Benford kanunu

Newcomb (1881) ve Benford (1938)'e göre ilk sayı kanunun eşitliği aşağıdaki verilmiş olan formül ile belirlenir.

$$Pr(d_1 = d) = \log_{10} \left(1 + \frac{1}{d} \right) \quad (1)$$

Eşitlikte d_1 ilk basamağı (soldaki ilk rakam değerini) ifade etmektedir ($d = 1, 2, 3 \dots 9$).

İkinci rakam dağılım olasılık (d_2) değerlerinin belirlenmesinde eşitlik 2'de bulunan k iki basamaklı sayıların ilk basamak değerlerini ifade etmektedir (Nigrini ve Miller, 2007).

$$Pr(d_2 = d) = \sum_{k=1}^9 \log_{10} \left(1 + \frac{1}{10k+d} \right) \quad (2)$$

Burada $d = 0, 1, 2, 3 \dots, 9$ ve $10k + d = 10, 11, 12 \dots, 99$ değer aralığındaki rakamları ifade etmektedir.

İkiden fazla basamaklı değerler için olasılık değerlerini eşitlik 3 verir (Hindls ve Hronová, 2015):

$$Pr(d_m = d) = \sum_{d_1=1}^9 \sum_{d_2=0}^9 \dots \sum_{d_{m-1}=0}^9 \log_{10} \left(1 + \frac{1}{\sum_{i=1}^{10} d_i 10^{m-i}} \right) \quad (3)$$

Burada $d_m = 0, 1, 2, 3 \dots 9$ değerlerini ifade etmektedir.

Bu makalede d_1 olasılık dağılımları kullanılmıştır. Benford Kanunu'na göre d_1 olasılık dağılımları 1'den 9'a kadar sırası ile 0,3010; 0,1761; 0,1249; 0,0969; 0,0792; 0,0669; 0,0580; 0,0512; 0,0458 şeklindedir.

Her bir d_1 sayısı için gözlenen ve teorik değerler arasında fark olup olmadığı Z istatistiği ile araştırılmıştır. Z istatistiği uygulanmasında her bir d_1 olasılık dağılımının uyumu için % 5 önem seviyesinde $u_{0,975} = 1,96$ kritik değeri esas alınmıştır (Durtschi vd., 2004; Cleary ve Thibodeau, 2005).

$$Z(d_1) = \frac{|p_o - p_e| - \frac{1}{2s}}{s(d_1)} \quad (4)$$

Burada $s(d_1)$ standart sapmayı ifade etmektedir.

$$s(d_1) = \sqrt{\frac{p_e(1-p_e)}{s}} \quad (5)$$

Her bir beklenen değer için % 95 güven düzeyinde güven aralığı (alt sınır ve üst sınır) aşağıda verilen formüllerle belirlenmiştir (Nigrini ve Mittermaier, 1997).

$$\text{alt sınır} = p_e - \left[1,96 * s(d_1) - \left(\frac{1}{2s} \right) \right] \quad (6.1)$$

$$\text{üst sınır} = p_e + \left[1,96 * s(d_1) + \left(\frac{1}{2s} \right) \right] \quad (6.2)$$

Gözlenen dağılımla teorik dağılım arasında genel anlamda uyum olup olmadığının kontrolü için ki-kare (χ^2) testi uygulanmıştır (Hindls ve Hronová, 2015).

$$\chi^2 = S \sum_{d_1=1}^9 \frac{(p_o - p_e)^2}{p_e} \quad (7)$$

Eşitliklerde yer alan p_o gözlenen (ampirik) değer, p_e teorik veya beklenen değer, S tür sayısı ya da örnek boyutudur. %95 güven düzeyinde $n - 1 = \sum d_1 - 1 = 8$ değerinde χ^2 kritik değeri, $\chi_{0,95}^2[8] = 15,51$ 'dir.

2.4. Genelleştirilmiş Benford Kanunu

Benford Kanunu'nun d_1 olasılık değerleri ile sınırlı kalmamak için ideal olan seçeneği genelleştirilmiş Benford Kanunu'dur. Genelleştirilmiş Benford Kanunu ile bir ekosistem için (e_i) gözlenen olasılık değerlerine (p_o) maksimum uyumunu sağlayacak teorik olasılık değerleri (p_γ) tespit edilebilir. Böylece her bir ekosistem kendisi için en uygun olasılık dağılımlarına göre ifade edilebilir.

Hürlimann (2015) Bayesian olasılık modeline dayanarak Benford Kanunu'nu ($GB(d; \alpha)$) genelleştirmiştir ($\gamma \in (-\infty, \infty)$).

$$GB(d; \gamma) = \begin{cases} \frac{d^{-\gamma} - (1+d)^{-\gamma}}{1 - 10^{-\gamma}}, \gamma \neq 0 \\ \log(1 + d^{-1}), \gamma = 0 \end{cases} \quad (8)$$

Burada $GB(d; \gamma) = \log(1 + d^{-1}) = Pr(d_m = d)$ olup, Benford Kanunu'na denk gelmektedir. Ayrıca $GB(d; \gamma) = \frac{d^{-\gamma} - (1+d)^{-\gamma}}{1 - 10^{-\gamma}}$ formunda $\gamma \rightarrow 0$ iken teorik olasılık değerleri Benford Kanunu'nun teorik olasılık değerlerine yaklaşır.

Bir ekosistem (e_i) için en uygun teorik olasılık değerlerinin belirlenmesi sonrası güven aralıkları ile birlikte Z istatistiği ve χ^2 testi p_e yerine p_γ kullanılarak belirlenmiştir.

$$Z(d_j) = \frac{|p_0 - p_\gamma| - \frac{1}{2S}}{\sqrt{\frac{p_\gamma(1-p_\gamma)}{S}}} \quad (9)$$

$$\text{alt sınır} = p_\gamma - \left[1,96 * \sqrt{\frac{p_\gamma(1-p_\gamma)}{S}} - \left(\frac{1}{2S}\right) \right] \quad (10.1)$$

$$\text{üst sınır} = p_\gamma + \left[1,96 * \sqrt{\frac{p_\gamma(1-p_\gamma)}{S}} + \left(\frac{1}{2S}\right) \right] \quad (10.2)$$

$$\chi^2 = S \sum_{d_i=1}^9 \frac{(p_0 - p_\gamma)^2}{p_\gamma} \quad (11)$$

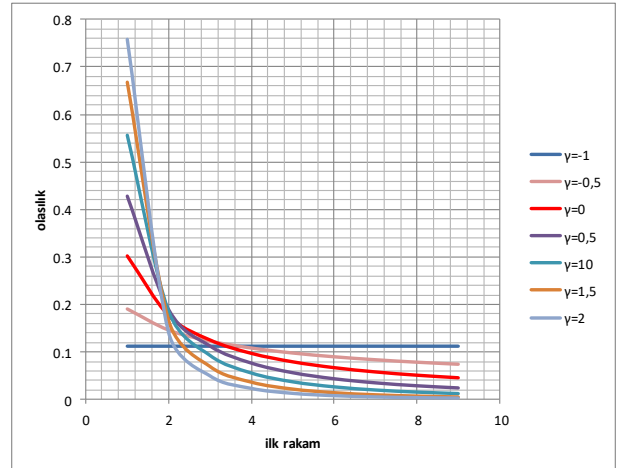
Bu makalede, bir ekosistem (e_i) için $GB(d; \gamma)$ uygulaması ile elde edilecek beklenen olasılık değerlerine (p_γ) yönelik tek kabul $d_i p_{\gamma(i)} > d_i p_{\gamma(i+1)}$ ($i = 1, 2, \dots, 8$) şeklindedir. Örnek olması amacıyla, Eşitlik 5 kullanılarak $\gamma = -1$ 'den $\gamma = 2$ 'e kadar oluşturulan teorik olasılık eğrileri Şekil 1'de verilmiştir.

2.5. İterasyonlar

İterasyon işlemleri BS ve BD 'nin örnek alan boyutlarından kaynaklanacak sapmanın belirlenmesi ve hesaplamalara aktarılması amacıyla gerçekleştirilmiştir. Bu bağlamda B veri matrisine iki kere iterasyon işlemi uygulanmıştır. Bunlardan birincisi BS 'ye atfen B veri matrisinden toplamda 10000 kere ($K = 10000$) olmak üzere ($k = 1, 2, \dots, 10000$) her defasında rastgele seçilen 96 örnek alan verisinden ${}^k p_0$ değerlerinin belirlenmesi, ikincisi BD 'ye atfen B veri matrisinde aynı işlemlerin 119 örnek alan için gerçekleştirilmesi şeklindedir. Genelleştirilmiş Benford Kanunu ile B 'nin gözlenen olasılık dağılımına en uygun olan teorik olasılık dağılımları (p_γ) tespit edildikten sonra bu teorik dağılım değerleri ile iterasyonlara ait olasılık dağılımlarının (${}^k p_0$) her biri arasında ${}^k \chi^2$ hesaplamaları gerçekleştirilmiştir.

$${}^k \chi^2 = {}^k S \sum_{d_i=1}^9 \frac{({}^k p_0 - p_\gamma)^2}{p_\gamma} \quad (12)$$

Eşitliklerde yer alan ${}^k S$ her bir iterasyonda yer alan tür sayısını ifade etmektedir. Zira rastgele seçilen örnek alanlar sebebiyle iterasyonların hepsinde tür sayısının eşit olması beklenemez.



Şekil 1. Genelleştirilmiş Benford Kanunu'nun farklı γ değerlerinde olasılık değer eğrileri

Tüm istatistiksel çıktılarının ortalaması alınarak (K değerine bölünerek) her iki alt bölge için ${}^k \chi^2$ değerleri standart sapma $\left(\sigma = \frac{\sqrt{\sum_{k=1}^{10000} ({}^k \chi^2 - {}^k \chi^2)^2}}{K-1} \right)$ değerleri ile birlikte elde edilmiştir.

$${}^k \chi^2 = \frac{\sum_{k=1}^K {}^k \chi^2}{K} \quad (13)$$

2.6. Doğallık hesabı

Bu makalede ekosistemlerin doğallık kestirimine yönelik temel kabul "Bir ekosistemin (e_i) doğallığı onun kendisine en uygun olasılık dağılım eğrisi ve parçası olduğu bütünlü açıklanabilir" şeklindedir.

Bu kabul çerçevesinde doğallık tespiti iki aşamalı olarak gerçekleştirilmiştir.

İlk aşama Bölüm 2.4'de açıklandığı üzere ekosistem $i(e_i)$ 'nin genelleştirilmiş Benford Kanunu ile en uygun olduğu teorik olasılık dağılımları (p_γ) belirlenmiştir. Bu durumda ekosistem $i(e_i)$ artık e_i^γ olarak tanımlanmıştır.

İkinci aşamada ekosistemlerin bütününden oluşan veri matrisinden (E_i) p_γ değerleri tespit edilmiş ve bu p_γ değerleri kullanılarak her bir ekosistem için iterasyon ile elde edilen olasılık değerleri kullanarak istatistiksel analizler gerçekleştirilmiştir. Bu durumda e_i^γ 'in karşılığı E_i^γ olmuştur. E_i^γ 'lerin Bölüm 2.5'de açıklandığı üzere hesaplanan ${}^k \chi^2$ değerlerinden maksimum değere sahip olanı E_{max}^γ olmak üzere, ekosistem $i(e_i)$ için doğallığın (nat_i) χ^2 testi ile kestirimine yönelik form eşitlik 14'de verilmiştir.

$$nat_i(\chi^2) = \chi^2(e_i^\gamma) \left(\frac{{}^k \chi^2(E_{max}^\gamma)}{{}^k \chi^2(E_i^\gamma)} \right) \quad (14)$$

Ekosistem $i(e_i)$ 'nin $nat_i(\chi^2)$ değeri ne kadar küçük ise o kadar doğaldır. Haliyle teorik olarak bir ekosistemin (e_i) tamamen doğal olması $nat_i(\chi^2) = 0$ olması anlamına gelmektedir.

3. Bulgular

3.1. Benford kanunu'na dayalı bulgular

Beşşehir Gölü Havzası (B), Sultan Dağları Alt Bölgesi (BS) ve Dedegül Dağları Alt Bölgesi (BD) veri setlerinden gözlenen olasılık değerleri ile Benford Kanunu olasılık değerleri arasında gerçekleştirilen istatistiksel analiz sonuçları Şekil 2 ve Çizelge 2'de verilmiştir.

B'nin ilk rakam değerleri (B_{d_1}) ile Benford Kanunu'nun ilk rakam değerleri (K_{d_1}) arasında genel olarak bir uyumun olduğu görülmektedir. B_{d_1} eğrisi K_{d_1} eğrisi ile en fazla $d_1 = 3$, $d_1 = 1$ ve $d_1 = 4$ rakamlarında uyumsuzluk göstermektedir. Fakat bu değerlerin Z_{d_1} sonuçları 1,96 değerinden daha düşüktür. Bunun sonucu olarak B_{d_1} eğrisi K_{d_1} eğrisinin güven aralıkları içinde kalmıştır. B_{d_1} ve K_{d_1} arasında farklılık söz konusu değildir ($\chi^2(e_B) = 8,733 < \chi^2_{0,95}[8] = 15,51$) (Şekil 2a, Çizelge 2).

BS'nin ilk rakam değerlerinden (BS_{d_1}) $d_1 = 1$ için $Z_1 = 3,501$ olup bu değer $u_{0,975} = 1,96$ kritik değerini aşmıştır. Ayrıca $d_1 = 7$ için $Z_7 = 1,646$ ile kritik değere oldukça yaklaşmıştır. Bunların sonucu olarak BS_{d_1} eğrisi K_{d_1} eğrisinin güven aralıkları içinde kalmamıştır ve $\chi^2(e_{BS}) = 16,579 > \chi^2_{0,95}[8] = 15,51$ olduğundan BS_{d_1} dağılımı K_{d_1} dağılımına uyum göstermemiştir (Şekil 2b).

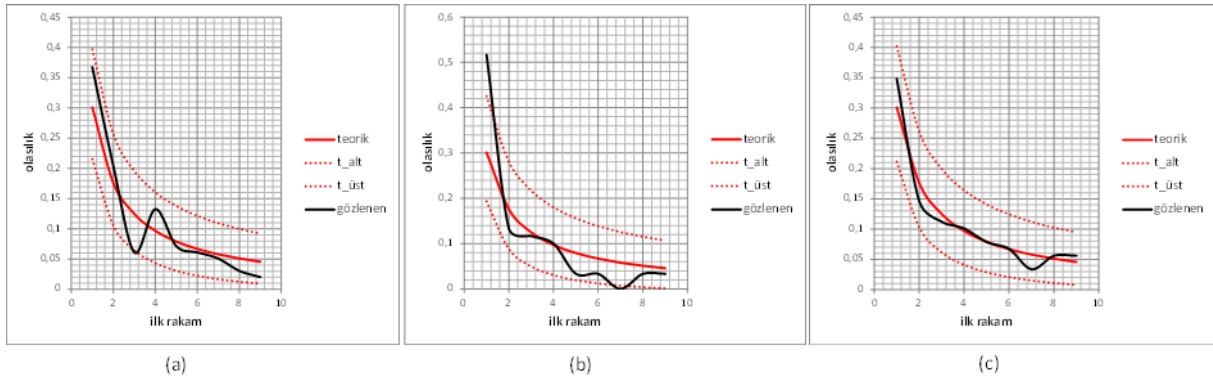
BD'nin ilk rakam değer (BD_{d_1}) eğrisi ile K_{d_1} eğrisi birbirleri ile oldukça paralel seyir izlemektedir ve BD_{d_1} eğrisinin bu seyri K_{d_1} eğrisinin güven aralıkları içinde

gerçekleşmektedir (Şekil 2c). Bütün d_1 değerleri için Z istatistiği sonuçları $u_{0,975} = 1,96$ kritik değerinin oldukça altındadır. En uyumsuz BD_{d_1} değeri $d_1 = 7$ olup bunun için $Z_7 = 0,753$ 'tür. $\chi^2(e_{BD}) = 2,406 < \chi^2_{0,95}[8] = 15,51$ olup bu sonuç, BD_{d_1} ile K_{d_1} arasında oldukça yüksek bir uyumun varlığına işaret etmektedir.

3.2. Genelleştirilmiş Benford Kanunu'na dayalı bulgular

Genelleştirilmiş Benford Kanunu uygulanırken B, BS ve BD'nin her biri için $\gamma = -1$ 'den $\gamma = 2$ 'ye kadar 0,01 birim aralıklarla χ^2 hesaplamaları gerçekleştirilmiştir. B, BS ve BD'nin maksimum uyumlarına (minimum χ^2 değerlerine) sırası ile $\gamma = 0,25$ 'de, $\gamma = 0,65$ 'de ve $\gamma = 0,07$ 'de ulaşılmıştır (Şekil 2).

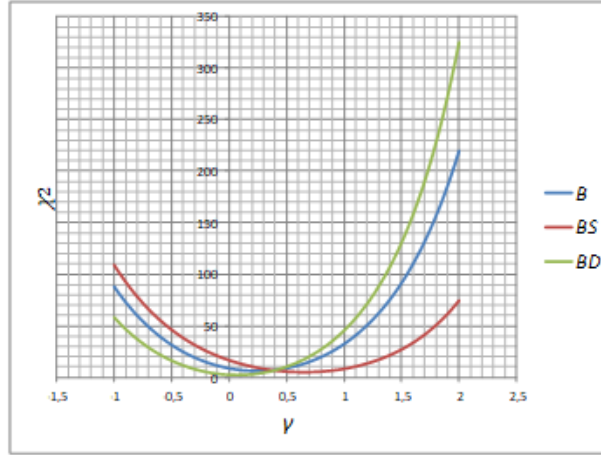
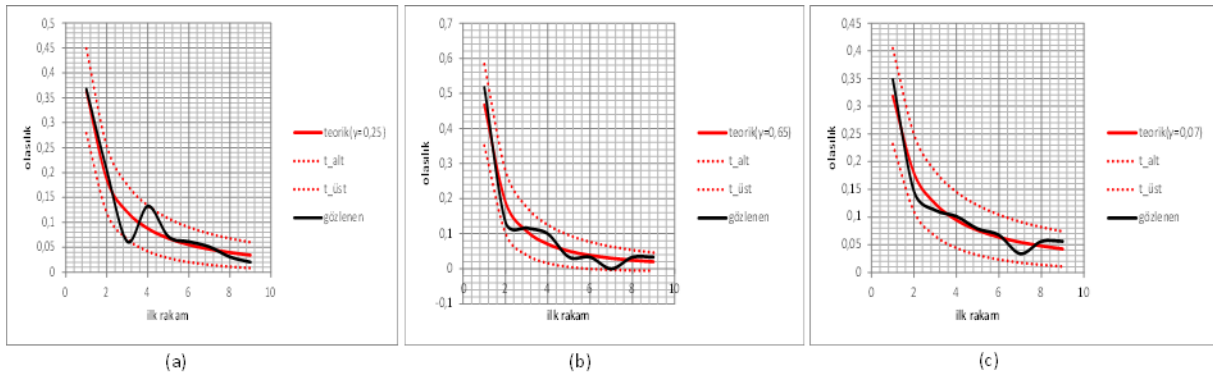
B, BS ve BD'nin gözlenen olasılık değerleri ile teorik olasılık değerleri arasında dikkate değer bir şekilde uyum artışı olmuştur (Şekil 2 ve Şekil 4). Çizelge 2 ve Çizelge 3 karşılaştırıldığında genel olarak Z_i değerlerinde uyuma katkı sağlayacak sonuçlar elde edilmiştir. Beşşehir Gölü Havzası ve onun alt bölgeleri için uyum artışındaki en büyük paya $d_1 = 1$ sahip olmuştur. Havza ve alt bölgeler kıyaslandığında en fazla uyum artışı BS'de gerçekleşmiştir. Zira Çizelge 2'de görüleceği üzere BS'nin χ^2 değeri genelleştirme öncesi ($\gamma = 0$) 16,589 (Çizelge 2) iken, bu değer genelleştirme sonrası ($\gamma = 0,25$), 4,992'ye kadar düşmüştür (Çizelge 3). Bu düşüşü B takip ederken en az uyum kazancı BD'de olmuştur (Şekil 3).



Şekil 2. Beşşehir Gölü Havzası (a), Sultan Dağları Alt Bölgesi (b) ve Dedegül Dağları Alt Bölgesi (c) $d_1 p_e$ eğrileri ve % 95 önem seviyesinde alt ve üst güven aralıkları ile $d_1 p_e$ teorik eğrisi

Çizelge 2 Beşşehir Gölü Havzası (B), Sultan Dağları Alt Bölgesi (BS) ve Dedegül Dağları Alt Bölgesi (BD) Benford olasılık değerlerine dayalı Z_{d_1} ve χ^2 sonuçları

e_i	S	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6	Z_7	Z_8	Z_9	χ^2
e_B	98	1,321	0,595	1,755	1,025	0,097	0,025	0,079	0,694	0,959	8,733
e_{BS}	60	3,501	0,700	-0,001	-0,137	1,076	0,783	1,646	0,334	0,152	16,589
e_{BD}	89	0,857	0,605	0,199	-0,045	-0,178	-0,194	0,753	-0,025	0,217	2,406

Şekil 3. B, BS ve BD'nin $\gamma = -1$ 'den $\gamma = 2$ 'ye χ^2 değer dağılımlarıŞekil 4. Beyşehir Gölü Havzası (a), Sultan Dağları Alt Bölgesi (b) ve Dedegül Dağları Alt Bölgesi (c) d_1p_0 eğrileri ve maksimum uyumu sağlayan d_1p_γ teorik eğrileri ile onların % 95 önem seviyesinde alt ve üst güven aralıklarıÇizelge 3. Beyşehir Gölü Havzası (B), Sultan Dağları Alt Bölgesi (BS) ve Dedegül Dağları Alt Bölgesi (BD) maksimum uyum eğrilerine dayalı Z_{d_1} ve χ^2 sonuçları

e_i^γ	S	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6	Z_7	Z_8	Z_9	χ^2
e_B^γ	98	-0,03	0,351	1,647	1,396	-0,07	0,04	-0,01	0,189	0,478	6,187
e_{BS}^γ	60	0,636	0,959	0,018	0,635	0,315	-0,13	0,991	0,021	0,249	4,992
e_{BD}^γ	89	0,497	0,674	0,173	0,033	-0,11	-0,07	0,632	0,127	0,386	2,209

3.3. İterasyon bulguları

215 örnek alandan oluşan B veri setine BS'ye atfen 96 örnek alan boyut sınırlandırması ile rastlantısal 10000 yinelemeli işlemlerden ve BD'ye atfen 119 örnek alan boyut sınırlandırması ile rastlantısal 10000 yinelemeli işlemlerden elde edilen olasılık değer eğrileri Şekil 5a ve Şekil 5b'de gösterilmektedir.

96 ve 119 örnek alan boyut sınırlandırmasının her birinde her bir yinelemeye ait gözlenen olasılık değerleri ile B'nin genelleştirme ile elde edilen teorik olasılık değerleri ($\gamma = 0,25$) arasında χ^2 testleri yapılmıştır. Her bir yinelemede örnek alan boyut sınırlandırması söz konusu olduğundan B veri setinde olan bütün türlerin (toplam 98 tür) işlemde yer alması beklenemez. Bu yüzden 96 örnek alan sınırlandırmasına ve 119 örnek alan sınırlandırmasına dayanan yineleme işlemlerinden tam sayıya yuvarlanmış ortalama tür sayıları ($[S_{ort}]$) sırası ile 84 ve 88 olarak bulunmuştur. Elde edilen sonuçlara göre 96 örnek alan sınırlandırmasının yineleme sayısına denk ($K = 10000$) χ^2 değerlerinin ortalaması, $k\chi^2(E_{BS}^\gamma) = 6,747$ ($\sigma_{BS} = 3,302$) olarak bulunmuştur. 119

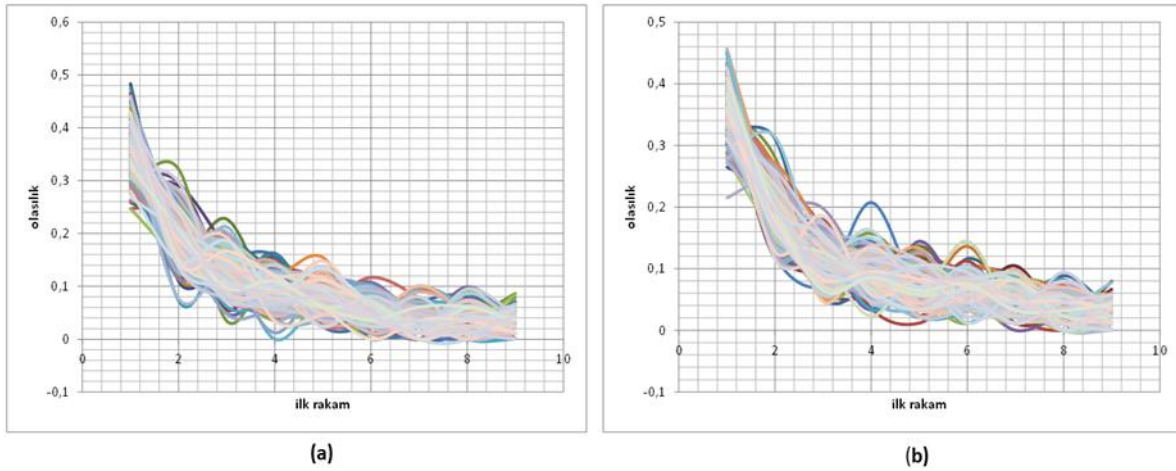
örnek alan sınırlandırmasında ise yinelemelerin ortalama χ^2 değeri ise $k\chi^2(E_{BS}^\gamma) = 6,176$ ($\sigma_{BS} = 3,147$) çıkmıştır (Çizelge 4).

Çizelge 5'deki değerler Bölüm 2.6'da verilen formülde (eşitlik 14) yerine konulduğunda, BS'nin doğallık değeri $nat_{BS}(\chi^2) = \chi^2(e_{BS}^\gamma) \left(\frac{k\chi^2(E_{max}^\gamma)}{k\chi^2(E_{BS}^\gamma)} \right) = 4,992(6,747/6,747) = 4,992$ ve BS'nin doğallık değeri $nat_{BD}(\chi^2) = \chi^2(e_{BD}^\gamma) \left(\frac{k\chi^2(E_{max}^\gamma)}{k\chi^2(E_{BD}^\gamma)} \right) = 2,209(6,747/6,176) = 2,414$ olarak ortaya çıkmaktadır.

Bölüm 2.6'da ifade edildiği üzere bir ekosistemin (e_i) doğallık değerinin yüksekliği $nat_i(\chi^2)$ değerinin düşüklüğü ile orantılıdır. Elde edilen sonuçlar göstermektedir ki, Dedegül Dağları Yöresi(BD)'nin doğallık değeri Sultan Dağları Yöresi(BS)'den daha yüksektir.

3.4. Doğallık hesaplama bulguları

BS ve BD'nin doğallık hesaplamaları için ihtiyaç duyulan tüm girdiler Çizelge 3 ve Çizelge 4'den özetlenerek Çizelge 5'de verilmiştir.



Şekil 5. Beyşehir gölü havzası, (215 örnek alan), rastgele seçilen 96 örnek alan (a) ve rastgele seçilen 119 örnek alan (b) veri setlerinin $d_1 p_0$ dağılımları (örnek alan sayısı boyutlandırma işlemlerinin her biri için yineleme sayısı=10000)

Çizelge 4. $\gamma = 0,25$ 'de B veri setinden 96 örnek alandan oluşan alt toplumlar ve 119 örnek alandan oluşan alt toplumlara ait $k\overline{\chi^2}$ sonuçları (uygulamaların her biri için alt toplum (alt veri seti) sayısı (yineleme sayısı) =10000)

	E_{BS}^Y	σ_{BS}	E_{BD}^Y	σ_{BD}
$k\overline{\chi^2}$	6,747	3,302	6,176	3,147
$[S_{ort}]$	84		88	

$[S_{ort}]$ yinelemelerin ortalama tür sayılarının tam sayıya yuvarlanmış değerlerini ifade etmektedir.

Çizelge 5. Doğallık hesaplama girdileri

	$\chi^2(e_i^Y)$	$k\overline{\chi^2}(E_i^Y)$	$k\overline{\chi^2}(E_{max}^Y)$	$\frac{k\overline{\chi^2}(E_{max}^Y)}{k\overline{\chi^2}(E_i^Y)}$
BS	4,992	6,747	6,747	1,000
BD	2,209	6,176		1,093

4. Tartışma ve sonuç

Ekosistemlerin doğallık hesabında Benford Kanunu'na odaklanılmasının sebebi, onun verdiği mesajın çok net olmasından kaynaklanmaktadır.

Matematiksel işlemlerin ürünü olmayan, haliyle özgeçmişe sahip olmayan (rastgele seçilmiş) on binlerce rakamın $d_1 = 1$ 'den $d_1 = 9$ 'a kadar elde edilen olasılık değer dağılımlarının birbirlerinden farklı olması beklenmez. Beklenti homojen dağılımdır. Oysaki yeterince işlem görmüş (çarpılmış çırpılmış ve/veya bölünmüş parçalanmış; eğer elde edilen sonuç bir (1) değerinden büyük ise kendisi aksi durumda tersi alınmış) ve bu sayede bir özgeçmişe sahip olan on binlerce rakamdan $d_1 = 1$ 'den $d_1 = 9$ 'a kadar elde edilen olasılık değerleri ($d_1 p_{o(1)} = 0,301, d_1 p_{o(9)} = 0,046$) azalan bir trende sahiptir. Bu trendi oluşturan eğri Benford Kanunu'nun olasılık değerleri tarafından oluşturulmuştur. Buna merkezi limit teorisinin vücut bulmuş hali de diyebiliriz (Hill, 1995; Fewster, 2009).

Benford olasılık eğrisi gerçekte sayı ordusundan oluşan bir bütünün "işlem bilgisi" içerdiği anlamına gelmektedir. Bu "işlem bilgisi" olasılık hesabına giren sayıların ne olduğunu ve hangi formüller kullanılarak elde edildiğini açıklamaz. Buradaki "işlem bilgisi" sadece özgeçmişe sahip sayılar

anlamına gelir. Benford Kanunu'nun bu mesajı onun teorik olasılık değerleri kadar nettir.

Özgeçmişe sahip sayılar ne anlama gelmektedir?

Sistem anlamına gelmektedir. Şöyle ki; bir bütünü oluşturan sayısal değerler eğer Benford olasılık eğrisine uyumlu bir davranış trendi gösteriyor ise, o bütünün var olması ve bunun devamını sağlayan bir "doğal bilinç" vardır. Diğer bir deyişle bir bütün eğer "doğal bilince" sahip ise Benford eğrisine benzer bir görüntü sunması beklenir. Bu durumda artık "bütün" ifadesi yerine "sistem" ifadesi kullanılır.

Doğal bilince uyumsuzluk ne anlama gelmektedir?

Doğal bilince dışardan harici faktörlerin etki ettiği veya ediyor olduğu anlamına gelir. Diğer bir deyişle, doğal bilinçte rol oynayan değişkenlerin oynadığı rollere müdahale ya da manipülasyon söz konusudur. Manipülasyon Benford Kanunu'nun kullanıldığı yere göre farklı anlamlara gelebilir. Örneğin Benford Kanunu şirket ya da devlet kuruluşlarının mali denetiminde kullanıyor ise ve gözlenen olasılık değerlerinin Benford Kanunu teorik olasılık değerlerine olan bariz uyumsuzluğu için makul bir açıklama söz konusu değil ise, bu durumda manipülasyon sahtekarlık veya yolsuzluk anlamına gelmektedir (Akkaş, 2007; Aybars ve Ataunal, 2016).

İnsanlar tarafından kurulan finansal ya da ekonomik sistemlerde Benford Kanunu'nu teorik olasılık değerlerini kontrol amacıyla kullanmak mantıklıdır. Zira bu sistemlerin dikkate alınan sayısal değerlerinin gerisinde bilinen matematik kuralları işlemektedir. Fakat doğal sistemlerde Benford Kanunu'nun teorik olasılık değerlerini eşik değerler olarak almak ve bu eşik değerler temelinde elde edilen sonuçlara göre kararlar vermek yapılacak değerlendirmelerde önemli hatalara sebep olabilir. Bu yüzden doğal sistemler için Benford Kanunu yerine geliştirilmiş Benford Kanunu'nun kullanılması çok daha iyi bir seçenektir.

Bu çalışmanın işlem süreçleri üç aşamadan oluşmaktadır.

İlk aşamada BS ve BD 'nin gözlenen olasılık değerleri (BD_{d_i} , BS_{d_i}) ile standart Benford Kanunu'nun teorik olasılık değerlerini (K_{d_i}) karşılaştırmak amacıyla χ^2 analizleri gerçekleştirilmiştir. Sonuçta $\chi^2(e_{BS}) = 16,579$ ve $\chi^2(e_{BD}) = 2,406$ olarak bulunmuştur. Standart Benford Kanunu'na göre BS , BD 'den 6,89 kat daha büyük bir χ^2 değerine sahiptir.

İkinci aşamada, geliştirilmiş Benford Kanunu kullanarak BS ve BD 'nin gözlenen olasılık değerlerine en uyumlu teorik olasılık değerleri belirlenmiştir. BS ve BD için en küçük χ^2 değerleri sırasıyla $\gamma = 0,65$ 'de ve $\gamma = 0,07$ 'de elde edilmiştir ($\chi^2(e_{BS}^Y) = 4,992$ ve $\chi^2(e_{BD}^Y) = 2,209$). Geliştirilmiş Benford Kanunu devreye sokulduğunda χ^2 değerleri itibarıyla BS 'de yaklaşık 3,5 kat azalma, BD 'de ise sadece 1,08 kat azalma tespit edilmiştir. Bunun sonucu olarak BS ve BD 'nin χ^2 değerleri itibarıyla oransal farkı 6,89'dan 2,26'ya düşmüştür.

BS ve BD için geliştirilmiş Benford Kanunu ile elde edilen χ^2 değerlerinin ekosistem doğallığına atfen kullanılması her iki alt bölgeye ait örnek alan sayılarının eşit olmaması sebebiyle güvenilir değildir. Diğer bir değişle doğallık hesabının sonlanması için kalibrasyon işlemi ile alt bölgelerin içerdiği örnek alan sayı farklılıklarından kaynaklanan sapmanın ortadan kaldırılması gerekmektedir.

Bu sebeple son aşama olan üçüncü aşamada BS ve BD 'nin birleşmesinden oluşan B veri setine 96 ve 119 örnek alan sınırlandırılması ile yineleme işlemleri gerçekleştirilmiştir. Bu yineleme işlemleri sonucunda $k\chi^2(E_{BS}^Y) = 6,747$ ve $k\chi^2(E_{BD}^Y) = 6,176$ bulunmuştur. Bu sonuçlar itibarıyla BS 'nin kalibrasyon katsayı değeri $k\chi^2(E_{max}^Y)/k\chi^2(E_{BS}^Y) = 1$ olduğu için doğallık değeri 4,992 ve BS 'nin kalibrasyon katsayı değeri $k\chi^2(E_{max}^Y)/k\chi^2(E_{BD}^Y) = 1,093$ olduğu için doğallık değeri 2,414 olarak bulunmuştur. Kalibrasyon sonucu BS ve BD 'nin χ^2 değerlerine dayalı doğallık değerlerinin oransal farkı 2,26 değerinden bir miktar daha düşerek 2,06 ile gerçek değerine ulaşmıştır. Elde edilen sonuçlar BD 'nin BS 'den yaklaşık olarak 2 kat daha doğal olduğunu göstermektedir. "Beşşehir Gölü Havzası'nın Yetiştirme Ortamı Özellikleri ve Sınıflandırılması" isimli çalışmada alt bölgeler ile ilgili yapılan açıklamalar da bu sonuçları teyit eder niteliktedir (Özkan, 2003).

Bu bölümü sonlandırmadan önce ekosistem doğallık konusunda gelecekteki çalışmalar için faydalı olabilecek ek bilgileri vermek uygun olacaktır.

Bu çalışmada gözlenen olasılık değerleri ile beklenen olasılık değerleri arasındaki genel uyumu belirlemek için sadece χ^2 testi kullanılmıştır. χ^2 testi dışında kullanılabilir başka alternatif testler de vardır. Bunlardan ilki χ^2 ile birlikte içinde birçok formülü barındıran $D^2(\lambda)$ eşitliğidir (Cressie ve Read, 1984).

$$D^2(\lambda) = \frac{2}{\lambda(\lambda+1)} \sum_{i=1}^9 S p_o \left\{ \left(\frac{p_o}{p_e} \right)^\lambda - 1 \right\} \quad (15)$$

Burada S örnek boyutu (tür sayısı), p_o gözlenen olasılık değeri, p_e beklenen olasılık değeridir. Eşitlikte $\lambda = 1$ değerinde Pearson χ^2 , $\lambda = 0$ 'da ($\lambda \rightarrow 0$) benzerlik oran testi (log likelihood ratio) $\lambda = -1/2$ değerinde Freman-Tukey istatistiği, $\lambda = -1$ 'de modifiye edilmiş benzerlik testi ve $\lambda = -2$ 'de modifiye edilmiş Neyman istatistiği sonuçları elde edilmektedir. Ayrıca Cressie ve Read (1984) yeni bir uyum istatistik testi önermişlerdir. Araştırmacıların önerdiği formül güç sapma test istatistiği olup bu formülde $\lambda = 2/3$ 'tür. Güç sapma test istatistiğinde önem düzeyi için serbestlik derecesi $n - 2 = \sum d_i - 2$ 'ye göre belirlenmektedir. Daha detaylı bir açıklama ile %95 güven düzeyinde $\sum d_i - 1 = 8$ değerinde χ^2 kritik değeri, $\chi_{0,95}^2[8] = 15,51$ iken $D^2(\lambda = 2/3)$ için %95 güven düzeyinde kritik değer ($\sum d_i - 2 = 7$) $D^2(\frac{2}{3})_{0,95}[7] = 14,07$ 'dir. Belirtmek gerekir ki, ekolojik verinin özellikleri sebebiyle $D^2(\lambda)$ formülünde negatif λ değerlerinin kullanılması uygun olmayabilir. Zira özellikle iterasyon işlemlerinde $\forall d_i p_{o(i)} > 0$ durumu geçerli olmayabilir.

$D^2(\lambda)$ formülü dışında mutlak farkların ortalaması (MAD) (Hinds ve Hronová, 2015), maksimum fark (d_{max}) (Judge ve Schechter, 2009), öklit mesafesi (d_e) (Judge ve Schechter, 2009), normalize edilmiş öklit mesafesi (d^*) (Crocetti ve Randi, 2016; Branets, 2019), FSD dağılımına dayalı fark (d_a) (Branets, 2019) ve Kullback ve Leibler ayrışması (d_{KL}) (Ausloos vd., 2016) kullanılabilir diğer formüllerdir.

$$d_{max} = \max_{i=1..9} |p_o - p_e| \quad (16)$$

$$MAD = \frac{\sum_{i=1}^9 |p_o - p_e|}{9} \quad (17)$$

$$d_e = \sqrt{\sum_{i=1}^9 (p_o - p_e)^2} \quad (18)$$

$$d^* = \frac{\sqrt{\sum_{i=1}^9 (p_o - p_e)^2}}{\sqrt{\sum_{i=1}^9 (p_e)^2 + (0,954243)^2}} \quad (19)$$

$$d_a = \frac{|\mu_o - \mu_e|}{9 - \mu_e} \quad (20)$$

$$d_{KL} = \sum_{i=1}^9 p_o \ln \left(\frac{p_o}{p_e} \right) \quad (21)$$

Bu formlar içinde d_{KL} iterasyon işlemlerinde $\forall d_i p_{o(i)} > 0$ ise kullanılabilir.

Eşitliklerde p_o gözlenen olasılık değeri, p_e beklenen olasılık değeridir. d^* eşitliğinde $0,954243 = 1 - 0,045757$ olup; $0,045757 d_i = 9$ 'un olasılık değeridir. d_a formülünde yer alan μ_o ve μ_e gözlenen ve beklenen olasılık değerlerinin kendilerine ait d_i değerleri ile çarpımlarının toplam değerleridir.

Uyum testlerinde olduğu gibi bu makalede kullanılmış olan Benford Kanunu'nun geliştirilmiş formu da ($GB(d; \gamma)$) (Hürlimann, 2015) alternatif değildir. $GB(d; \gamma)$ dışında kullanılabilir hem tek hem de iki parametrelili geliştirilmiş eşitlikler bulunmaktadır. Tüm

genelleştirilmiş eşitliklerde $d \in \{1, \dots, 9\}$ olup, tek parametrelilikten $GB(d; a)$ (Tseng vd., 2017) ve $GB(d; c)$ (Hürlimann, 2003) ve $GB(d; s)$ (Clippe ve Ausloos, 2012) a, c ve s modellere ait parametrelerdir ($c \geq 1$). $GB(d; a)$ eşitliği $a \rightarrow \infty$ doğru, $GB(d; s)$ eşitliği $s = 0$ 'da ve $GB(d; c)$ eşitliği $c = 1$ veya $c = 2$ 'de Benford Kanunu'na indirgenir.

$$GB(d; a) = \sum_{j=-\infty}^{\infty} \int_{10^j d}^{10^{j+1} d} f(x) dx = \sum_{j=-\infty}^{\infty} \left[\frac{1}{10^{j+1} d^{a+1}} - \frac{1}{10^j d^{a+1}} \right] \quad (22)$$

$$GB(d; c) = \frac{1}{2} \{ [\log_{10}(I + d)]^c - [\log_{10}(d)]^c - [1 - \log_{10}(I + d)]^c + [1 - \log_{10}(d)]^c \} \quad (23)$$

$$GB(d; s) = N \log_{10} \left(\frac{1}{d} + I + sd \right) \quad (24)$$

İki parametrelilikten genelleştirilmiş eşitliklerden $GB(d; s; m)$ 'de s ve m (Fu vd., 2007), $GB(d; a; b)$ 'de a ve b (Tseng vd., 2017), $GB(d; \alpha; \beta)$ 'de α ve β (Hürlimann, 2009) modellere ait parametreleri temsil etmektedir.

$$GB(d; s; m) = N \log_{10} \left(I + \frac{1}{s+am} \right) \quad (25)$$

$$GB(d; a; b) = \frac{b}{a+b} \sum_{j=-\infty}^{-1} \left[10^{\frac{j}{b}} (I + d)^{\frac{1}{b}} - 10^{\frac{j}{b}} d^{\frac{1}{b}} \right] - \frac{a}{a+b} \sum_{j=0}^{\infty} \left[10^{-\frac{j}{a}} d^{-\frac{1}{a}} - 10^{-\frac{j}{a}} (I + d)^{-\frac{1}{a}} \right] \\ = \frac{b}{a+b} \left[\frac{(I+d)^{\frac{1}{b}-\frac{1}{b}}}{10^{\frac{1}{b}-1}} \right] + \frac{a}{a+b} \left[\frac{d^{-\frac{1}{a}-(I+d)^{-\frac{1}{a}}}}{1-10^{-\frac{1}{a}}} \right] \quad (26)$$

$$GB(d; \alpha; \beta) = \frac{\alpha}{\alpha+\beta} \{ [\log_{10}(I + d)]^\beta - [\log_{10}(d)]^\beta \} + \frac{\beta}{\alpha+\beta} \sum_{k=1}^{\infty} \{ [k + \log_{10}(d)]^{-\alpha} - [k + \log_{10}(I + d)]^{-\alpha} \} \quad (27)$$

$GB(d; s; m)$ eşitliği $s = 0$ ve $m = 1$ 'de, $GB(d; a; b)$ eşitliği $a \rightarrow \infty$ ve/veya $b \rightarrow \infty$ doğru ve, $GB(d; \alpha; \beta)$ eşitliği $\alpha \rightarrow \infty$ ve $\beta = 1$ 'de Benford Kanunu'na indirgenir. $GB(d; s)$ ve $GB(d; s; m)$ 'de bulunan N normalleştirme faktörü olup beklenen olasılık değerleri toplamı=1 olmasını sağlar.

Bunların dışında Benford Kanunu'nun Tsallis entropisi ile genelleştirilmiş bir formuda ($GB(d; q; \beta)$) bulunmaktadır (Shao ve Ma, 2010).

$$\widehat{GB}(d; q; \beta) = \sum_{n=-\infty}^{\infty} \left\{ [1 - 10^n \beta d(I - q)]^{\frac{q-2}{q-1}} - [1 - 10^n \beta (d + 1)(I - q)]^{\frac{q-2}{q-1}} \right\} \quad (28)$$

Bu eşitlikle $1 < q \leq 2$ olup, $d_1 = 1$ 'den $d_9 = 9$ 'a kadar olasılık değer toplamları ≤ 1 olma ihtimali sebebiyle, olasılık değerlerinin $GB(d; q; \beta) = \frac{\widehat{GB}(d; q; \beta)}{\sum_{i=1}^9 \widehat{GB}(d; q; \beta)}$ formülü ile düzenlenmesine ihtiyaç vardır. Bu eşitlik ekolojik çalışmalar için $\beta = 1$ ve q 'nun model parametresi alınması durumunda işe yarayabilir. $GB(d; q; \beta)$ ile teorik olasılık değerleri dar aralıkta değişim göstermektedir. Bu yüzden bu formdan diğer genelleştirilmiş formlar kadar verimli sonuçlar alınması beklenmemelidir.

Burada belirtmek gerekir ki, Benford Kanunu'nun genelleştirilmiş herhangi bir j formu (GB_j) seçildiğinde, bu formun model parametresine ya da parametrelerine ait değerler $d_j p_{GB_j(i)} > d_{j+1} p_{GB_j(i+1)}$ ($i = 1, 2, \dots, 8$) kuralını ihlal etmeyecek sınırlarda kalmalıdır.

Sonuç olarak, bu makale ile ekosistem veya meta toplum seviyesindeki karşılaştırmalar için ilk defa önerilmiş olan genelleştirilmiş Benford Kanunu'na dayalı doğallık hesabı alışılmadık model yaklaşımlarından tamamen farklıdır. Bu model yaklaşımı olasılık temelli olup, uyum ve genelleştirmeye yönelik birçok alternatif formülle birlikte kullanılabilir ve ekosistemlerin doğallık hesabına yönelik sonuçların elde edilmesinde etkin rol oynayabilir. Bununla birlikte, önerilen bu model yaklaşımının etkinliğini gerçek anlamda görebilmek için farklı bölge ekosistemlerinde ve farklı canlı gruplarında gerçekleştirilecek yeni çalışmalara da ihtiyaç olduğu bilinmelidir.

Kaynaklar

Akkaş, M., 2007. Denetimde Benford kanunu'nun uygulaması. Gazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 9(1): 191-206.

Anderson, J.E., 1991. A conceptual framework for evaluating and quantifying naturalness. Conservation Biology, 5: 347-352.

Ausloos, M., Castellano, R., Cerqueti, R., 2016. Regularities and discrepancies of credit default swaps: A data science approach through Benford's law. Chaos, Solitons & Fractals, 90: 8-17.

Aybars A., Ataunal L., 2016. An application of Benford's Law to fundamental accounting figures reported by Borsa Istanbul (BIST) companies. Journal of Economics, Finance and Accounting, 3(3): 234-243.

Benford, F., 1938. The law of anomalous numbers. Proceedings of The American Philosophical Society, 78: 551-572.

Branets, S., 2019. Detecting money laundering with Benford's law and machine learning . Masters Thesis, University of Tartu, Faculty of Social Sciences, Estonia.

Cleary, R., Thibodeau, J.C., 2005. Applying digital analysis using Benford's law to detect fraud: The dangers of type I errors. Auditing: A Journal of Practice & Theory, 24(1): 77-81.

Clippe, P., Ausloos, M., 2012. Benford's law and their transform of financial data. Physica A: Statistical Mechanics and Its Applications, 391: 6556- 6567.

Côté, S., Bélanger, L., Beaugard, R., Thiffault, É., Margni, M., 2019. A conceptual model for forest naturalness assessment and application in Quebec's Boreal Forest. Forests, 10(4): 325. <https://doi.org/10.3390/f10040325>

Cressie, N., Read, T.R.C., 1984. Multinomial goodness-of-fit tests. Journal of the Royal Statistical Society. Series B (Methodological), 46: 440-464.

Crocetti, E., Randi, G., 2016. Using the Benford's Law as a first step to assess the quality of the cancer registry data. Frontiers in Public Health, 4:225. DOI:10.3389/fpubh.2016.00225.

Dunlop, S., Lanfranco, S., Schembri, J.S., 2014. The role of 'Naturalness' and seral stage in the assessment and management of coastal sites, <https://www.um.edu.mt/library/oar/bitstream/123456789/21115/1/Dunlop%20C%20Lanfranco%20%26%20Schembri%20%282014%29.%20The%20role%20of%20%27Naturalness%27%20and%20seral%20stage%20in%20the%20assessment%20and%20management%20of%20coastal%20sites.pdf> , Erişim: 20.12.2020.

Durtschi, C., Hillison, W., Pacini, C., 2004. The effective use of Benford's law to assist in detecting fraud in accounting data. Journal of Forensic Accounting, 5(1): 17-34.

Erdős, L., Bátori, Z., Penksza, K., Dénes, A., Kevey, B., Kevey, D., Magnes, M., Sengl, P., Tölgyesi, C., 2017. Can naturalness indicator values reveal habitat degradation? A test of four methodological approaches. Polish Journal Of Ecology, 65: 1-13.

- Fanfarillo, E., Kasperski, A., Giuliani, A., Cicinelli, E., Latini, M., Abbate, G., 2018. Assessing naturalness of arable weed communities: A new index applied to a case study in central Italy. *Biological Agriculture & Horticulture*, 34(4): 232-244, DOI: 10.1080/01448765.2018.1434832.
- Fattorini, L., 2003. Statistical analysis of ecological diversity (Eds., El-Shaarawi, A.H., Jureckova, J.), *Environmetrics*, EOLSS: Paris, France, 1, 18–29.
- Fewster, R.M., 2009. A simple explanation of Benford's Law. *The American Statistician*, 63: 26-32. <http://dx.doi.org/10.1198/tast.2009.0005>.
- Fu, D., Shi, Y.Q., Su, W., 2007. A generalized Benford's Law for JPEG coefficients and its applications in image forensics. *Proceedings of SPIE*, 6506, 1L1- 1L11.
- Hill, T.P., 1995. A statistical derivation of the significant-digit law. *Statistical Science*, 10: 354–363.
- Hindls, R., Hronová, S., 2015. Benford's Law and possibilities for its use in governmental statistics. *Statistika*, 95(2): 54-64.
- Hürlimann, W., 2009. Generalizing Benford's law using power laws: Application to integer sequences. *International Journal of Mathematics and Mathematical Sciences*, Article ID 970284, Doi:10.1155/2009/970284.
- Hürlimann, W., 2003. A generalized Benford law and its application. *Mathematics Preprint Archive*, 3(3): 217-228.
- Hürlimann, W., 2015. On the uniform random upper bound family of first significant digit distributions. *Journal of Informetrics*, 9(2): 349-358.
- Judge, G., Schechter, L., 2009. Detecting problems in survey data using Benford's Law. *Journal of Human Resources*, 44(1): 1-24.
- Leroy, B., 2016. Calculation of rarity indices for species and assemblages of species. R package version 1.3-4. Available at: <https://CRAN.R-project.org/package=Rarity>.
- Machado, A., 2004. An index of naturalness. *Journal for Nature Conservation*, 12: 95-110.
- Margules, C., Usher, M.B., 1981. Criteria used in assessing wildlife conservation potential: A review. *Biological Conservation*, 21:79–109.
- Newcomb, S., 1881. Note on the frequency of use of the different digits in natural numbers. *American Journal of Mathematics*, 4(1): 39–40.
- Nigrini, M.J., Mittermaier, L.J., 1997. The use of Benford's law as an aid in analytical procedures. *Auditing*, 16(2): 52.
- Nigrini, M.J., Miller, S.J., 2007. Benford's Law applied to hydrology data—Results and relevance to other geophysical data. *Mathematical Geology*, 39(5): 469–490. <http://doi.org/10.1007/s11004-007-9109-5>.
- Özkan, K., 2016. *Biyolojik Çeşitlilik Bileşenleri (α , β ve γ) Nasıl Ölçülür (1. Basım)*. Süleyman Demirel Üniversitesi, Orman Fakültesi Yayınları, Isparta.
- Özkan, K., 2003. *Beyşehir Gölü Havzasının Yetiştirme Ortamı Özellikleri ve Sınıflandırılması*. Doktora Tezi, İÜ Araştırma Fonu Proje Numarası T-981/19022001, 189s.
- Reif, A., Walentowski, H., 2008. The assessment of naturalness and its role for nature conservation and forestry in Europe. *Waldökologie Landschaftsforschung Naturschutz*, 6:63–76.
- Shao, L., Ma, B.Q., 2010. First-digit law in nonextensive statistics. *Physical Review E*, 82(4): 1-4, 10.1103/PhysRevE.82.041110.
- Smith, P.G.R., Theberge, J.B., 1986. A review of criteria for evaluating natural areas. *Environmental Management*, 10(6): 715-734.
- Tans, W., 1974. Priority ranking of biotic natural areas. *Michigan Botanist*, 13: 31–39.
- Tseng, H.C., Huang, W.N., Huang, D.W., 2017. Modified Benford's law for two-exponent distributions. *Scientometrics*, 110(3): 1403–1413.