EGE ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ

(DOKTORA TEZİ)

GENELLEŞTİRİLMİŞ SIRA İSTATİSTİKLERİ VE UYGULAMALARI

Halil TANIL

İstatistik Anabilim Dalı

406.02.01

Sunuş Tarihi: 3.12.2004

Tez Danışmanı: Prof.Dr. İsmihan BAYRAMOĞLU

2004-İZMİR

Sayın Halil Tanıl' ın DOKTORA TEZİ olarak hazırladığı "Genelleştirilmiş Sıra İstatistikleri ve Uygulamaları" adlı bu çalışma, yapılan tez savunması sınavı sonunda jüri tarafından Lisansüstü Eğitim ve Öğretim Yönetmeliği uyarınca değerlendirilerek oybirliği/~~oyçokluğu~~ ile kabul edilmiştir.

3 / 12 / 2004

	Adı soyadı	İmza
Başkan	; Doç. Dr. Şaush Şenol	
Üye	; Prof. Dr. İsmihan Bayramoğlu	
Üye	; Prof. Dr. Ülkü Gürler	
Üye	; Prof. Dr. Onur Başkan	
Üye	; Prof. Dr. Turgut Öziş	

ÖZET

GENELLEŞTİRİLMİŞ SIRA İSTATİSTİKLERİ VE UYGULAMALARI

TANIL, Halil

Doktora Tezi, İstatistik Bölümü
Tez Yöneticisi: Prof.Dr. İsmihan Bayramoğlu
Aralık 2004, 81sayfa

Bu çalışmada, sıra istatistikleri, II. tür ilerleyen sansürlenmiş sıra istatistikleri ve benzeri sıralanmış rasgele değişken modellerini içeren genelleştirilmiş sıra istatistiklerine dayalı olarak; dağılımdan-bağımsız (veya invaryant) güven aralıkları ve güven düzeyleri ortaya koyulmuş, aşan istatistiklerin olasılık dağılımları elde edilmiş ve minimal boşluklara dayalı istatistikler incelenmiştir. Ayrıca, Bairamov ve Petunin (1991b)' in, eğitici örneklemlere dayalı test prosedürü, genelleştirilmiş sıra istatistiklerine uyarlanmıştır. İleri sürülen bulgu ve yöntemler, bazı istatistiksel problemlerin çözümünde kullanılmıştır.

Anahtar kelimeler: *Genelleştirilmiş sıra istatistikleri, Dağılımdan-bağımsız istatistikler, Aşan istatistikler, Sıra istatistikleri, Minimal boşluklar, Eğitici örneklemler.*

ABSTRACT

GENERALIZED ORDER STATISTICS AND THEIR APPLICATIONS

TANIL, Halil

Ph.D., Department of Statistics
Supervisor: Prof.Dr. İsmihan Bayramoğlu
December 2004, 81pages

In this study, based on generalized order statistics which include the models of ordered random variables such as ordinary order statistics and progressively Type-II censored order statistics, we present the distribution-free (or invariant) confidence intervals with their confidence levels, the probability mass functions of exceedance statistics, and the statistics based on minimal spacings. In addition, we adapt the statistical test procedure (based on training samples) proposed by Bairamov and Petunin (1991b) to generalized order statistics. The proposed statistitical results and methods are applied to some statistical problems.

Key words: *Generalized order statistics, Distribution-free (or invariant) confidence intervals, Exceedance statistics, Minimal spacings, Training samples.*

TEŞEKKÜR

Doktora tez çalışmamda bana yön veren, beni yetiştiren ve kıymetli vaktini benden esirgemeyen danışmanım Prof. Dr. İsmihan Bayramoğlu' na; tez çalışmam sırasında bana her türlü imkanı sağlayan, bana güvenen ve manevi destek veren Prof. Dr. Onur Baskan' a; kıymetli hocam ve İstatistik bölüm başkanım Doç. Dr. Şanslı Şenol' a, bana matematiği öğreten sevgili babama, bana sabrı öğreten sevgili anneme, her zaman yanımda olan ve bana her konuda destek veren biricik eşim Zeynep'e ve henüz 5 aylık olan tatlı kızım İdil'e çok teşekkür ederim.

İÇİNDEKİLER

İÇİNDEKİLER (devam)

TABLOLAR DİZİNİ

1 GİRİŞ

Sıra istatistikleri, istatistik teorisinin en önemli kavramlarından biri olup, temel istatistik yöntemlerde ve istatistiksel sonuç çıkarımında kullanılmaktadır. Aynı zamanda, teste tabi tutulan n tane ürünün yaşam zamanlarını gösterdiği için, yaşam analizi (Life-time analysis) ve güvenilirlik teorisinde (reliability theory) de önemli yer tutmaktadır. Sıra istatistikleri, yeterli istatistikler olduklarından örneklem hakkındaki tüm bilgiyi içerirler. Sıra istatistiklerine son yıllarda artan ilgi, bu alanın hızla gelişimini sağlamıştır. Özellikle, sıra istatistiklerine dayalı birçok istatistik, dağılımdan-bağımsızlık (distribution-free) özelliği taşıdığı için, parametrik olmayan istatistiksel yöntemlerde geniş şekilde kullanılmaktadır. Günümüzde, birçok tıbbi uygulamada ortaya çıkan II. tür ilerleyen sansürlenmiş sıra istatistikleri teorisi hızla gelişmektedir.

Sıralanmış veriler sadece sonlu örneklem kitlelerinde değil, sonsuz örneklemlerde de karşımıza çıkmaktadır. Bu duruma örnek olarak rekor değerler (record values) gösterilebilir.

İstatistiksel verilerin sıralanmasının genel bir kuralını ortaya koymak amacıyla, Kamps (1995), "genelleştirilmiş sıra istatistikleri (generalized order statistics)" adını verdiği soyut kavramı ileri sürmüştür. Genelleştirilmiş sıra istatistikleri, birçok parametreye bağlı olarak, birleşik olasılık yoğunluk fonksiyonu ile verilen bir rasgele vektörün elemanlarıdır. Parametrelerin özel durumlarında sıra istatistikleri (ordinary order statistics), II. Tür ilerleyen sansürlenmiş sıra istatistikleri (progressive Type-II censored order statistics), ardışık sıra istatistikleri (sequential order statistics), rekor değerler (record values) ve başka sıralama modelleri elde edilmektedir. Bu yüzden, genelleştirilmiş sıra istatistikleri için elde edilmiş herhangi bir sonuç veya bu istatistiklere dayalı herhangi bir istatistiksel çıkarsama, otomatik olarak sıra istatistiklerine, II. tür ilerleyen sansürlenmiş sıra istatistiklerine, ardışık sıra istatistiklerine ve benzeri modellere uygulanabilir.

Bu tez çalışmasının ikinci bölümünde, sıra istatistikleri, II. tür ilerleyen sansürlenmiş sıra istatistikleri ve genelleştirilmiş sıra istatistikleri hakkında literatürde var olan birtakım genel bilgiler verilmiştir. Üçüncü bölümde, genelleştirilmiş sıra istatistiklerine dayalı aşan (exceedance) istatistiklerin dağılımları ortaya koyulmuş, minimal boşluklara dayalı istatistiklerin dağılımları elde edilmiştir. Dördüncü bölümde ise, genelleştirilmiş sıra istatistiklerine dayalı test prosedürü geliştirilmiş, özel halde sıra istatistiklerine ve II. tür ilerleyen sansürlenmiş sıra istatistiklerine dayalı prosedürler ortaya konmuştur. Beşinci bölüm olan Uygulamalar bölümündeyse, edilen istatistiksel bulgu ve prosedürlerin kullanılabileceği örnek uygulamalara yer verilmiştir.

2 SIRALANMIŞ RASGELE DEĞİŞKENLERE İLİŞKİN MODELLER

Bu bölümde, sıralanmış rasgele değişkenlere ilişkin modellerden sıra istatistikleri ve II. tür ilerleyen sansürlenmiş sıra istatistikleri üzerinde durulacak, ayrıca, bu ve diğer sıralanmış rasgele değişken modellerini içeren genelleştirilmiş sıra istatistikleri hakkında bilgi verilecektir.

2.1 Sıra İstatistikleri

$(X_1, X_2, \ldots, X_n)$ vektörünün elemanları, birbirlerinden bağımsız ve her biri F birikimli dağılım fonksiyonuna sahip rasgele değişkenler olsun. Bu vektörün elemanları büyüklüklerine göre artacak şekilde sıralansın ve bu sıralı vektör $(X_{1:n}, X_{2:n}, \ldots, X_{n:n})$ şeklinde gösterilsin. Elde edilen bu sıralı vektöre "sıra istatistikleri vektörü" adı verilir. Sıra istatistikleri vektörünün elemanları arasında $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$ ilişkisinin olduğu açıktır. $r = 1,2,\ldots,n$ olmak üzere, sıra istatistikleri vektörünün herhangi bir elemanı olan ve $X_{r:n}$ ile gösterilen rasgele değişken, $(X_1, X_2, \ldots, X_n)$ vektörü üzerinde tanımlı "r. sıra istatistiği" olarak adlandırılır (David, 1970, Gibbons, 1971, Arnold vd., 1992).

2.1.1 Marjinal ve Birleşik Olasılık yoğunluk ve dağılım fonksiyonları

"Sıra istatistiği" terimi yukarıdaki gibi tanımlandıktan sonra, şimdi, F dağılım fonksiyonunun sürekli olduğu varsayımı altında, sıra istatistiklerine ilişkin olasılık yoğunluk ve dağılım fonksiyonlarını elde edelim. Örneğin, $X_{n:n} = \max\{X_1, X_2, ..., X_n\}$ olarak tanımlanan "n. sıra istatistiği" nin dağılım fonksiyonu $F_n(x)$ ile gösterilsin. Bu durumda,

$$F_n(x) = P\{X_{n:n} \leq x\} = P\{\max\{X_1, X_2, ..., X_n\} \leq x\} \tag{2.1}$$

yazılabilir. Bu ifadede yer alan "$(X_1, X_2, ..., X_n)$ vektörünün elemanlarından en büyüğünün x' ten küçük veya eşit olması" olayı, "bu vektörün tüm elemanlarının x' ten küçük veya eşit olması" anlamına geldiği için, (2.1) eşitliği,

$$F_n(x) = P\{X_1 \leq x, X_2 \leq x, ..., X_n \leq x\}$$

şekline gelecektir. $X_1, X_2, ..., X_n$ rasgele değişkenlerinin birbirlerinden bağımsız oldukları varsayımı daha önce yapıldığından ve $P\{X_i \leq x\} = F(x)$ olarak tanımlandığından dolayı, sonuç olarak,

$$F_n(x) = (F(x))^n \tag{2.2}$$

eşitliğine ulaşılır ve $X_{n:n}$ sıra istatistiğinin dağılım fonksiyonu elde edilmiş olur. Benzer şekilde, $X_{1:n} = \min\{X_1, X_2, ..., X_n\}$ şeklinde tanımlanan sıra istatistiğinin dağılım fonksiyonunun,

$$F_1(x) = 1 - (1 - F(x))^n \tag{2.3}$$

olduğu kolayca gösterilebilir.

$X_{1:n}$ ve $X_{n:n}$ sıra istatistiklerinin marjinal dağılım fonksiyonları, sırasıyla (2.3) ve (2.2)' de verildikten sonra, şimdi, $1 \leq r \leq n$ olmak üzere, $X_{r:n}$ sıra istatistiğinin dağılım fonksiyonu $F_r(x)$' i elde edelim. $\{X_{r:n} \leq x\}$ olayı, $\{\text{en az } r \text{ adet } X_i \leq x\}$ olayına denk olduğundan,

$$\begin{aligned} F_r(x) &= P\{\text{en az } r \text{ adet } X_i \leq x\} \\ &= P\{r \text{ adet } X_i \leq x \ \textbf{veya} \ (r+1) \text{adet } X_i \leq x \\ &\qquad \textbf{veya} \ldots \textbf{veya} \ n \text{ adet } X_i \leq x\} \end{aligned} \tag{2.4}$$

yazılabilir. Yukarıdaki olasılık ifadesi içinde yer alan "veya" bağlaçları arasındaki her olay birbirlerinden ayrık olduklarından, (2.4) eşitliği, toplamsal olarak,

$$\begin{aligned} F_r(x) &= \sum_{j=r}^{n} P\{\text{tam olarak } j \text{ adet } X_i \leq x\} \\ &= \sum_{j=r}^{n} \binom{n}{j} (F(x))^j (1-F(x))^{n-j} \end{aligned} \tag{2.5}$$

şeklinde yazılacak ve $X_{r:n}$ sıra istatistiğinin dağılım fonksiyonuna ulaşılmış olacaktır. Bu noktada, binom toplamları ve tamamlanmamış Beta fonksiyonu arasındaki iyi tanımlı ilişkiden dolayı, (2.5) eşitliğiyle verilen dağılım fonksiyonu,

$$F_r(x) = I_{F(x)}(r, n-r+1) \tag{2.6}$$

şeklinde de yazılabilir. (2.6) eşitliğinin sağ tarafında yer alan tamamlanmamış beta fonksiyonu, $I_p(a,b)$ olarak gösterilecek olursa,

$$I_p(a,b) = \frac{\int_0^p t^{a-1}(1-t)^{b-1}\,dt}{B(a,b)} \tag{2.7}$$

şeklinde tanımlanır ve paydasındaki beta fonksiyonu olan $B(a,b)$' nin,

$$B(a,b)=\int_0^1 t^{a-1}(1-t)^{b-1}\,dt=\frac{(a-1)!\,(b-1)!}{(a+b-1)!} \tag{2.8}$$

olduğu bilinmektedir (David, 1970).

$X_{r:n}$ sıra istatistiğinin (2.6) eşitliğiyle verilen dağılım fonksiyonundan hareketle, bu rasgele değişkenin olasılık yoğunluk fonksiyonu için,

$$f_r(x)=\frac{1}{B(r,n-r+1)}\frac{d}{dx}\int_0^{F(x)} t^{r-1}(1-t)^{n-r}\,dt \tag{2.9}$$

yazılabilir. (2.9) eşitliğinin sağ tarafındaki integral ve türev birbirlerini götürdükten sonra bu eşitlik,

$$f_r(x)=\frac{1}{B(r,n-r+1)}F^{r-1}(x)(1-F(x))^{n-r}f(x) \tag{2.10}$$

haline gelecek ve r. sıra istatistiğinin olasılık yoğunluk fonksiyonu elde edilmiş olacaktır.

İki veya daha fazla sıra istatistiğini içeren olayların olasılıklarını hesaplayabilmek için, bu istatistiklere ilişkin birleşik olasılık yoğunluk fonksiyonuna ihtiyaç vardır. $1\le r<s\le n$ ve $x\le y$ olmak üzere, $X_{r:n}$ ve $X_{s:n}$ sıra istatistiklerine ait birleşik olasılık yoğunluk fonksiyonu, δx ve δy, çok küçük pozitif reel sayılar olmak üzere,

$$f_{r,s}(x,y)=\frac{d^2F_{r,s}(x,y)}{dxdy}$$
$$=\lim_{\substack{\delta x\to 0\\ \delta y\to 0}}\frac{P\{x<X_{r:n}\le x+\delta x, y<X_{s:n}\le y+\delta y\}}{\delta x\delta y} \tag{2.11}$$

şeklinde yazılabilir. (2.11) eşitliğinin sağ tarafındaki limit içinde yer alan olasılık için,

$$\begin{aligned}&P\{x<X_{r:n}\le x+\delta x, y<X_{s:n}\le y+\delta y\}\\ &\quad=P\{(r-1)\text{ adet } X_i\in(-\infty,x], \text{bir adet } X_i\in(x,x+\delta x],\\ &\qquad(s\text{-}r\text{-}1)\text{ adet } X_i\in(x+\delta x,y], \text{bir adet } X_i\in(y,y+\delta y],\\ &\qquad\qquad(n-s)\text{ adet } X_i\in(y+\delta y,\infty)\}\end{aligned}$$

eşitliği yazılabilir. Bu son eşitliğin sağ tarafındaki olasılığın,

$$\begin{aligned}&C_n^{r-1}(F(x))^{r-1}\times C_{n-r+1}^{1}(F(x+\delta x)-F(x))\\ &\quad\times C_{n-r}^{s-r-1}(F(y)-F(x+\delta x))^{s-r-1}\\ &\qquad\times C_{n-s+1}^{1}(F(y+\delta y)-F(y))\\ &\qquad\quad\times C_{n-s}^{n-s}(1-F(y+\delta y))^{n-s}\end{aligned} \tag{2.12}$$

ifadesine eşit olduğu açıktır. (2.12) ifadesindeki kombinasyonlar açılır ve sadeleştirmeler yapılırsa,

$$\begin{aligned}&\frac{n!}{(r-1)!(s-r-1)!(n-s)!}(F(x))^{r-1}(F(x+\delta x)-F(x))\\ &\quad\times(F(y)-F(x+\delta x))^{s-r-1}(F(y+\delta y)-F(y))\\ &\qquad\times(1-F(y+\delta y))^{n-s}\end{aligned} \tag{2.13}$$

elde edilir. δx ve δy, sıfıra çok yakın pozitif reel sayılar olduğu için $F(x+\delta x)-F(x)$ yerine yaklaşık olarak $f(x)\delta x$; $F(y+\delta y)-F(y)$

yerine de $f(y)\delta y$ yazılabilir. Bu durumda, (2.11) da yer alan eşitlik, $x \leq y$ olmak üzere,

$$f_{r,s}(x,y) = \frac{n!}{(r-1)!(s-r-1)!(n-s)!}(F(x))^{r-1} f(x) \\ \times (F(y)-F(x))^{s-r-1} f(y) \\ \times (1-F(y))^{n-s} \quad (2.14)$$

haline gelecek ve birleşik olasılık yoğunluk fonksiyonu elde edilmiş olacaktır.

$X_{r:n}$ ve $X_{s:n}$ sıra istatistiklerinin yukarıda verilen marjinal olasılık yoğunluk fonksiyonlarının çarpımı, (2.14) eşitliğinde verilen birleşik olasılık yoğunluk fonksiyonunu vermediği için, bu rasgele değişkenlerin bağımlı oldukları sonucuna ulaşılır. Bir başka deyişle, $(X_1, X_2, ..., X_n)$ vektörünün, herbiri sürekli bir F dağılım fonksiyonuna sahip elemanlarının birbirlerinden bağımsız oldukları bilindiğinde, bu vektöre dayalı olarak oluşturulan $(X_{1:n}, X_{2:n}, ..., X_{n:n})$ sıra istatistikleri vektörünün elemanları birbirlerine bağımlıdır.

r. ve s. sıra istatistiklerinin birleşik dağılım fonksiyonu ise, $1 \leq r < s \leq n$ olmak üzere,

$$F_{r,s}(x,y)=\begin{cases}\displaystyle\sum_{i=r}^{n}\sum_{j=\max(0,s-i)}^{n-i}\frac{n!}{i!\,j!(n-i-j)!} & , \quad x<y \\ \quad\times(F(x))^{i}(F(y)-F(x))^{j}(1-F(y))^{n-i-j} & \\ & \\ F_s(y) & , \quad x\geq y\end{cases} \tag{2.15}$$

şeklinde verilmektedir (David, 1970).

n adet sıra istatistiğinin birleşik olasılık yoğunluk fonksiyonu ise, $x_1 \leq x_2 \leq \ldots \leq x_n$ olmak üzere,

$$f_{1,2,\ldots,n}(x_1,x_2,\ldots,x_n)=n!\prod_{j=1}^{n}f(x_j) \tag{2.16}$$

şeklindedir (David, 1970; Gibbons, 1971).

2.1.2 Yeni gözlemleri içeren dağılımdan-bağımsız güven aralıkları ve aşan istatistikler

Dağılımdan-bağımsız (distribution-free) veya invaryant (invariant) güven aralıkları ilk kez Bairamov ve Petunin (1991a) tarafından ortaya atılmış bir konudur (Bairamov ve Eryılmaz, 2000). Yeni bir gözlemin, bir dağılımdan-bağımsız güven aralığına düşmesi olasılığı, aşan (exceedance) istatistiklerin dağılım fonksiyonlarının elde edilmesinde önemli bir rol oynamaktadır. Yeni gözlemleri içeren dağılımdan-bağımsız güven aralıklarının tanımı aşağıdaki gibidir: $\Im$, herhangi bir dağılım fonksiyonları sınıfını göstermek üzere;

$X_1, X_2, ..., X_n$, $F \in \Im$ dağılım fonksiyonuna sahip n büyüklüğünde rasgele bir örneklem olsun. $f_1(.)$ ve $f_2(.)$ ile gösterilen fonksiyonlar, $\forall(x_1, x_2, ..., x_n) \in R^n$ olmak üzere $f_1(x_1, x_2, ..., x_n) \leq f_2(x_1, x_2, ..., x_n)$ koşulunu sağlayan Borel fonksiyonları olarak kabul edilsin. Eğer, $\forall F \in \Im$ için,

$$P\{X_{n+1} \in (f_1(X_1, X_2, ..., X_n), f_2(X_1, X_2, ..., X_n))\} = \beta \qquad (2.17)$$

olacak şekilde $\exists \beta \in (0,1)$ varsa, o zaman $(f_1(X_1, X_2, ..., X_n), f_2(X_1, X_2, ..., X_n))$ rasgele aralığına, $\Im$ sınıfı için "yeni gözlemleri içeren dağılımdan-bağımsız güven aralığı" denir. (2.17) eşitliğinin sağ tarafında yer alan β ise her $F \in \Im$ için sabittir ve "dağılımdan-bağımsız aralığın güven düzeyi" olarak adlandırılır (Bairamov ve Özkaya, 2000).

$X_1, X_2, ..., X_n$ birbirlerinden bağımsız ve aynı F dağılım fonksiyonuna sahip rasgele değişkenler olmak üzere, F dağılım fonksiyonuna sahip $X_{n+1}, X_{n+2}, ..., X_{n+m}$' ler yeni gözlemler olsun. Varsayalım ki, $X_{n+1}, X_{n+2}, ..., X_{n+m}$ gözlemleri, $X_1, X_2, ..., X_n$' lerden bağımsız olarak elde edilsin. $i = 1,2,...,m$ için ξ_i rasgele değişkeni,

$$\xi_i = \begin{cases} 1 & , \; X_{n+i} \in (f_1(X_1, X_2, ..., X_n), f_2(X_1, X_2, ..., X_n)) \\ 0 & , \; X_{n+i} \notin (f_1(X_1, X_2, ..., X_n), f_2(X_1, X_2, ..., X_n)) \end{cases}$$

olmak üzere, $S_m = \sum_{i=1}^{m} \xi_i$ şeklinde bir rasgele değişken tanımlansın. Bu rasgele değişken, birbirlerinden bağımsız ve ortak bir $F \in \Im$ dağılım fonksiyonuna sahip m adet yeni gözlemden kaç tanesinin, dağılımdan-bağımsız güven aralığı $(f_1(X_1,X_2,...,X_n), f_2(X_1,X_2,...,X_n))$' e düştüğünü göstermektedir. S_m istatistiklerine, literatürde "aşan (exceedance) istatistikler" adı verilir.

Şimdi, yukarıdaki bilgileri sıra istatistiklerine uygulayalım. Eğer $\Im = \Im_c$, tüm sürekli dağılım fonksiyonlarını gösteren sınıf olarak tanımlanırsa, o zaman, $1 \le r < s \le n$ ve $X_{r:n} \le X_{s:n}$ olmak üzere, $f_1(X_1,X_2,...,X_n) = X_{r:n}$ ve $f_2(X_1,X_2,...,X_n) = X_{s:n}$ olarak seçilebilir. Bu durumda, $(X_{r:n}, X_{s:n})$ aralığı, yeni gözlemler için bir dağılımdan-bağımsız güven aralığı olur. Bairamov ve Petunin (1991a), $F \in \Im_c$ dağılım fonksiyonuna sahip kitleden gelen yeni bir gözlemin bu dağılımdan-bağımsız aralığa düşme olasılığının,

$$P\{X_{n+1} \in (X_{r:n}, X_{s:n})\} = \frac{s-r}{n+1} \tag{2.18}$$

olduğunu göstermişlerdir. $(X_{r:n}, X_{s:n})$ dağılımdan-bağımsız güven aralığına, F dağılım fonksiyonuna sahip ve birbirlerinden bağımsız l adet yeni gözlemin düşmesi olasılığı,

$$P\{X_{n+1}, X_{n+2}, ..., X_{n+l} \in (X_{r:n}, X_{s:n})\} = \frac{n!(s-r-1+l)!}{(s-r-1)!(n+l)!} \tag{2.19}$$

dir. Ayrıca, m adet yeni gözlemden ilk l tanesinin bu dağılımdan-bağımsız güven aralığına düşmesi ve geriye kalan $(m-l)$ tanesinin ise düşmemesi olasılığı ise,

$$P\{X_{n+1},X_{n+2},...,X_{n+l}\in(X_{r:n},X_{s:n}),$$
$$X_{n+l+1},X_{n+l+2},...,X_{n+m}\notin(X_{r:n},X_{s:n})\}$$
$$=\frac{(s-r)(s-r+1)...(s-r+l-1)}{(n+1)(n+2)...(n+m)}$$
$$\times(n+1-s+r)...(n+m-l-s+r) \qquad (2.20)$$

şeklinde verilmektedir. $i=1,2,...,m$ olmak üzere, ξ_i rasgele değişkeni,

$$\xi_i=\begin{cases}1 & , \ X_{n+i}\in(X_{r:n},X_{s:n})\\ 0 & , \ X_{n+i}\notin(X_{r:n},X_{s:n})\end{cases}$$

olmak üzere, $S_m=\sum_{i=1}^{m}\xi_i$ şeklinde tanımlanan aşan istatistiğin değerinin l ' ye eşit olması olasılığı ise, $l\in\{0,1,2,...,m\}$ için,

$$P\{S_m=l\}=\binom{m}{l}\frac{(s-r)(s-r+1)...(s-r+l-1)}{(n+1)(n+2)...(n+m)}$$
$$\times(n+1-s+r)...(n+m-l-s+r) \qquad (2.21)$$

şeklindedir (Bairamov ve Kotz, 2001).

2.1.3 Minimal boşluğa (minimal spacing) dayalı istatistikler

$\Im$, herhangi bir dağılım fonksiyonları sınıfı olmak üzere, $X_1,X_2,...,X_n$, değer kümesi negatif olmayan sürekli $F\in\Im$ dağılım fonksiyonuna sahip bir örneklem olsun. $f_1(.),f_2(.),...,f_n(.)$ ile gösterilen fonksiyonlar, $\forall(x_1,x_2,...,x_n)\in R^n$ olmak üzere, $f_1(x_1,x_2,...,x_n)\le f_2(x_1,x_2,...,x_n)\le...\le f_n(x_1,x_2,...,x_n)$ koşulunu sağlayan Borel fonksiyonları olsun. $f_0(x_1,x_2,...,x_n)=0$ için,

$$f_1(x_1,x_2,...,x_n)-f_0(x_1,x_2,...,x_n),$$

$$f_2(x_1,x_2,...,x_n)-f_1(x_1,x_2,...,x_n),$$

$$\dots$$

$$f_n(x_1,x_2,...,x_n)-f_{n-1}(x_1,x_2,...,x_n)$$

boşlukları dikkate alınsın. Minimal uzunluğa sahip boşluğun indeksini gösteren v rasgele değişkeni, $i=1,2,...,n$ olmak üzere,

$$v=r \Leftrightarrow f_r(x_1,x_2,...,x_n)-f_{r-1}(x_1,x_2,...,x_n) \\ \leq f_i(x_1,x_2,...,x_n)-f_{i-1}(x_1,x_2,...,x_n)$$

şeklinde tanımlanır. Bu durumda $f_v(x_1,x_2,...,x_n)-f_{v-1}(x_1,x_2,...,x_n)$ boşluğuna "minimal boşluk" denir.

Sıra istatistiklerinde minimal boşluğu tanımlamak için,

$$f_1(x_1,x_2,...,x_n)=X_{1:n},$$

$$f_2(x_1,x_2,...,x_n)=X_{2:n},$$

$$\dots$$

$$f_n(x_1,x_2,...,x_n)=X_{n:n}$$

olarak alınmalıdır. Bu durumda v rasgele değişkeni, $i=1,2,...,n$ olmak üzere,

$$v=r \Leftrightarrow (X_{r:n}-X_{r-1:n})\leq(X_{i:n}-X_{i-1:n})$$

şeklinde tanımlanacaktır (Bairamov ve Eryılmaz, 2000). F dağılım fonksiyonu, $\lambda>0$ parametreli üstel dağılım olduğunda, $i=1,2,...,n$

olmak üzere, ($X_{i:n} - X_{i-1:n}$) olarak tanımlanan rasgele değişkenler birbirlerinden bağımsızdırlar. Ayrıca,

$$\left(X_{1:n}, X_{2:n} - X_{1:n}, X_{3:n} - X_{2:n}, \ldots, X_{n:n} - X_{n-1:n}\right)$$

boşluklar vektörünün, sıra istatistikleri $X_{1:n}, X_{2:n}, \ldots, X_{n-1:n}$' lerden de bağımsız olduğu bilinmektedir (Arnold ve diğ., 1998). Bairamov ve Eryılmaz (2000), çalışmalarında, $F(x) = 1 - e^{-\lambda x}, x \geq 0, \lambda > 0$ olmak üzere, v rasgele değişkeninin r'ye eşit olması olasılığının,

$$P\{v = r\} = \frac{2(n - r + 1)}{n(n+1)} \tag{2.22}$$

olduğunu göstermişlerdir. Yine aynı çalışmada, minimal boşluğa dayalı olarak, $1 \leq l \leq m$ olmak üzere,

$$P\{X_{n+1} \in (X_{v-1:n}, X_{v:n})\} = \frac{4}{3} \frac{(n+2)}{(n+1)(n^2 + n + 2)}, \tag{2.23}$$

$$P\{X_{n+1}, X_{n+2}, \ldots, X_{n+l} \in (X_{v-1:n}, X_{v:n})\}$$

$$= \frac{2(n + l + 1)}{(n+1)(l+2)} \binom{n(n+1)/2 + l}{l}^{-1} \tag{2.24}$$

ve

$$P\{X_{n+1}, X_{n+2}, \ldots, X_{n+l} \in (X_{v-1:n}, X_{v:n}),$$
$$X_{n+l+1}, X_{n+l+2}, \ldots, X_{n+m} \notin (X_{v-1:n}, X_{v:n})\}$$

$$= \frac{2}{n+1} \sum_{j=0}^{m-l} (-1)^j \binom{m-l}{j} \frac{n + l + j + 1}{l + j + 2} \left(\binom{n(n+1)/2 + l + j}{l + j}\right)^{-1} \tag{2.25}$$

olasılıkları verilmiş, ayrıca, $i = 1,2,\ldots,m$ olmak üzere, ξ_i rasgele değişkeni,

$$\xi_i = \begin{cases} 1 & , \; X_{n+i} \in (X_{v-1:n}, X_{v:n}) \\ 0 & , \; X_{n+i} \notin (X_{v-1:n}, X_{v:n}) \end{cases}$$

olmak üzere, $S_m = \sum_{i=1}^{m} \xi_i$ şeklinde tanımlanan rasgele değişkenin değerinin l' ye eşit olması olasılığının,

$$P\{S_m = l\} = \binom{m}{l} \frac{2}{n+1} \sum_{j=0}^{m-l} (-1)^j \binom{m-l}{j} \frac{n+l+j+1}{l+j+2} \times \left(\binom{n(n+1)/2 + l + j}{l+j} \right)^{-1} \qquad (2.26)$$

olduğu ispatlanmıştır.

2.2 II. Tür İlerleyen Sansürlenmiş Sıra İstatistikleri

II-tür ilerleyen sansürlenmiş sıra istatistikleri, genellikle, yaşam-testi deneylerinde kullanılmaktadır. Örneğin, N adet birbirinden bağımsız elemanın belirli koşullar altındaki bozulma süreleri üzerine yapılan bir deneyde, ilk bozulan elemanla birlikte geriye kalan $(N-1)$ elemandan rasgele olarak R_1 tanesinin deneyden çıkarıldığı, daha sonra, ikinci bozulan elemanla birlikte geriye kalan $(N - R_1 - 2)$ elemandan rasgele olarak R_2 tanesinin deneyden çıkarıldığı düşünülsün. i. bozulan elemanla birlikte geriye kalan $(N - R_1 - R_2 - \ldots - R_{i-1} - i)$ elemandan R_i tanesinin bu deneyden çıkarılacağı açıktır. Bu deneyin ilk n adet bozulma sonunda bitirilmesi durumunda, deney süresince n adet bozulma süresi gözlemlenmiş ve $(R_1 + R_2 + ... + R_n)$ adet eleman sansürlenmiş olmaktadır. Burada $R_n = N - R_1 - R_2 - ... - R_{n-1} - n \geq 0$ şeklinde hesaplanır ki bu, n bozulma sonunda deneyden, geriye kalan tüm elemanların çıkarılması anlamına gelir. Böyle bir deneyde $\widetilde{\mathbf{R}} = (R_1, R_2, ..., R_n)$ vektörüne, "sansürleme planı" denmektedir (Balakrishnan vd., 2001).

Yukarıdaki deneyde, $i = 1,2,...,n$ olmak üzere, i. bozulma zamanı $X_{i:n:N}^{\widetilde{\mathbf{R}}}$ rasgele değişkeni ile gösterilecek olursa, $X_{1:n:N}^{\widetilde{\mathbf{R}}} \leq X_{2:n:N}^{\widetilde{\mathbf{R}}} \leq ... \leq X_{n:n:N}^{\widetilde{\mathbf{R}}}$ rasgele değişkenlerine, "II. tür ilerleyen sansürlenmiş sıra istatistikleri" adı verilir.

II. tür ilerleyen sansürlenmiş sıra istatistiklerine ilişkin birleşik olasılık yoğunluk fonksiyonu Balakrishnan ve Aggarwala (2000) tarafından, F sürekli bir olasılık dağılım fonksiyonu, $c = N(N - R_1 - 1)...(N - R_1 - R_2 - ... - R_{n-1} - n + 1)$ ve $x_1 \leq x_2 \leq ... \leq x_n$ olmak üzere,

$$f^{X_{1:n:N}^{\widetilde{\mathbf{R}}}, ..., X_{n:n:N}^{\widetilde{\mathbf{R}}}}(x_1, x_2, ..., x_n) = c\prod_{j=1}^{n} f(x_j)(1 - F(x_j))^{R_j} \tag{2.27}$$

şeklinde verilmiştir. Eğer sansürleme planı $\widetilde{\mathbf{R}} = (0,0,...,0)$ olarak seçilirse, II. Tür ilerleyen sansürlenmiş sıra istatistikleri, sıra istatistiklerine indirgenmiş olacaktır. Bu durumda, $j = 1,2,...,n$ ve $N = n$ olmak üzere, (2.27)' de yeralan $c = n!$ olacaktır. Sonuç olarak, (2.27)' de verilen olasılık yoğunluk fonksiyonunun, (2.16)' de verilmiş olan n adet sıra istatistiğinin birleşik olasılık yoğunluk fonksiyonu haline geleceği açıkça görülebilir.

$X_{r:n:N}^{\tilde{R}}$ ile gösterilen, r. II. tür ilerleyen sansürlenmiş sıra istatistiğinin marjinal olasılık yoğunluk ve dağılım fonksiyonları, $c_{r-1} = \prod_{j=1}^{r} \gamma_j$, $1 \le i \le r \le n$ için $a_i(r) = \prod_{\substack{j=1 \\ j \ne i}}^{r} \left(1/(\gamma_j - \gamma_i)\right)$, $\prod_{\varnothing} = 1$, $\forall i \ne j$ için $\gamma_i \ne \gamma_j$, $N = n + \sum_{i=1}^{n} R_i$ ve $\gamma_j = n - j + \sum_{i=j}^{n} R_i + 1 > 0$ olmak üzere, sırasıyla,

$$f^{X_{r:n:N}^{\tilde{R}}}(x) = c_{r-1} f(x) \sum_{i=1}^{r} a_i(r) (1 - F(x))^{\gamma_i - 1} \tag{2.28}$$

ve

$$F^{X_{r:n:N}^{\tilde{R}}}(x) = 1 - c_{r-1} \sum_{i=1}^{r} \frac{a_i(r)}{\gamma_i} (1 - F(x))^{\gamma_i} \tag{2.29}$$

olarak verilmektedir (Kamps ve Cramer, 2001).

2.3 Genelleştirilmiş Sıra İstatistikleri

Sıralanmış rasgele değişkenlere ilişkin çeşitli modeller vardır. Örneğin; sıra istatistikleri, II. tür ilerleyen sansürlenmiş sıra istatistikleri, rekor değerler, k-rekorlar, Pfeifer' in rekor modeli, ardışık (sequential) sıra istatistikleri gibi. Genelleştirilmiş sıra istatistikleri, tüm bu modelleri veya bunların sınırlandırılmış uygun versiyonlarını içermektedir.

Genelleştirilmiş sıra istatistikleri, kendisine ait parametrelerle oynanarak, herbir alt modele indirgenebilir.

Genelleştirilmiş sıra istatistiklerini tanımlamadan önce, "Uniform genelleştirilmiş sıra istatistikleri" nin tanımını vermemiz gerekir. "r. uniform genelleştirilmiş sıra istatistiği", $\tilde{\mathbf{m}} = (m_1, m_2, \dots, m_{n-1}) \in R^{n-1}$, $k \geq 1$ ve $1 \leq r \leq n$ olmak üzere, $U(r, n, \tilde{\mathbf{m}}, k)$ ile gösterilir. Ayrıca, n adet uniform genelleştirilmiş sıra istatistiği arasında $0 \leq U(1, n, \tilde{\mathbf{m}}, k) \leq U(2, n, \tilde{\mathbf{m}}, k) \leq \dots \leq U(n, n, \tilde{\mathbf{m}}, k) \leq 1$ ilişkisi vardır ve bu istatistiklere ait birleşik olasılık yoğunluk fonksiyonu, $0 \leq u_1 \leq u_2 \leq \dots \leq u_n \leq 1$, $n \in \{2,3,\dots\}$, $m_1, \dots, m_{n-1} \in R$, $\tilde{\mathbf{m}} = (m_1, m_2, \dots, m_{n-1})$, $k \geq 1$, $j = 1,2,\dots,n-1$ için $\gamma_j = k + n - j + M_j > 0$, $\gamma_n = k$ ve $M_j = \sum_{i=j}^{n-1} m_i$ olmak üzere,

$$f^{U(1,n,\tilde{\mathbf{m}},k), U(2,n,\tilde{\mathbf{m}},k), \dots, U(n,n,\tilde{\mathbf{m}},k)}(u_1, u_2, \dots, u_n)$$
$$= k\left(\prod_{j=1}^{n-1} \gamma_j\right)\left(\prod_{i=1}^{n-1} (1-u_i)^{m_i}\right)(1-u_n)^{k-1} \tag{2.30}$$

şeklinde verilmektedir (Kamps, 1995).

Yukarıda verdiğimiz tanımdan sonra, şimdi, "genelleştirilmiş sıra istatistikleri" nin tanımını verelim: Sürekli F dağılım fonksiyonuna bağlı olarak, $X(r, n, \tilde{\mathbf{m}}, k) = F^{-1}(U(r, n, \tilde{\mathbf{m}}, k))$ şeklinde tanımlanan rasgele değişkene "r. genelleştirilmiş sıra istatistiği" adı verilir. Genelleştirilmiş sıra istatistiklerinin birleşik olasılık yoğunluk fonksiyonu ise, $F^{-1}(0) < x_1 \leq x_2 \leq \dots \leq x_n < F^{-1}(1)$ olmak üzere,

$$f^{X(1,n,\tilde{\mathbf{m}},k),X(2,n,\tilde{\mathbf{m}},k),\ldots,X(n,n,\tilde{\mathbf{m}},k)}(x_1,x_2,\ldots,x_n)$$

$$=k\left(\prod_{j=1}^{n-1}\gamma_j\right)\left(\prod_{i=1}^{n-1}(1-F(x_i))^{m_i}f(x_i)\right)(1-F(x_n))^{k-1}f(x_n) \qquad (2.31)$$

olarak tanımlanmaktadır (Kamps, 1995).

Kamps ve Cramer (2001), yaptıkları çalışmada genelleştirilmiş sıra istatistiklerine ilişkin marjinal olasılık yoğunluk ve marjinal dağılım fonksiyonlarını, $c_{r-1}=\prod_{j=1}^{r}\gamma_j$, $1\le i\le r\le n$ için $a_i(r)=\prod_{\substack{j=1\\ j\ne i}}^{r}(1/(\gamma_j-\gamma_i))$,

$a_i(r)=\prod_{\varnothing}=1$ ve $\forall i\ne j$ için $\gamma_i\ne\gamma_j$ olmak üzere,

$$f^{X(r,n,\tilde{\mathbf{m}},k)}(x)=c_{r-1}f(x)\sum_{i=1}^{r}a_i(r)(1-F(x))^{\gamma_i-1} \qquad (2.32)$$

ve

$$F^{X(r,n,\tilde{\mathbf{m}},k)}(x)=1-c_{r-1}\sum_{i=1}^{r}\frac{a_i(r)}{\gamma_i}(1-F(x))^{\gamma_i} \qquad (2.33)$$

şeklinde vermişlerdir. Yine aynı çalışmada, $1\le r<s\le n$ için $X(r,n,\tilde{\mathbf{m}},k)$ ve $X(s,n,\tilde{\mathbf{m}},k)$ genelleştirilmiş sıra istatistiklerinin birleşik olasılık yoğunluk fonksiyonu, $r+1\le j\le s$ için $a_j^{(r)}(s)=\prod_{\substack{i=r+1\\ i\ne j}}^{s}1/(\gamma_i-\gamma_j)$

olmak üzere,

$$f^{X(r,n,\tilde{\mathbf{m}},k),X(s,n,\tilde{\mathbf{m}},k)}(x_r,x_s)$$
$$=c_{s-1}\left(\sum_{j=r+1}^{s}a_j^{(r)}(s)\left(\frac{1-F(x_s)}{1-F(x_r)}\right)^{\gamma_j}\right)\left(\sum_{i=1}^{r}a_i(r)(1-F(x_r))^{\gamma_i}\right)$$
$$\times\frac{f(x_r)}{1-F(x_r)}\frac{f(x_s)}{1-F(x_s)} \tag{2.34}$$

biçiminde verilmektedir.

(2.31) eşitliğinde birleşik olasılık yoğunluk fonksiyonu verilen genelleştirilmiş sıra istatistikleri modeli, Tablo1'de verilen özel durumlardaki gibi çeşitli sıralanmış rasgele değişken modellerini içermektedir (Bieniek ve Szynal, 2003; Kamps, 1995).

$1 \le r < s \le n$ olmak üzere, r. ve s. genelleştirilmiş sıra istatistikleri arasındaki fark $W_{r,s} = X(s,n,\tilde{\mathbf{m}},k) - X(r,n,\tilde{\mathbf{m}},k)$ şeklinde tanımlanırsa, Kamps ve Cramer (2001), bu rasgele değişkene ilişkin olasılık yoğunluk ve dağılım fonksiyonlarını aşağıdaki gibi vermektedir:

$$f^{W_{r,s}}(w)=c_{s-1}\int_0^{\infty}\left(\sum_{i=r+1}^{s}a_i^{(r)}(s)\left(\frac{1-F(v+w)}{1-F(v)}\right)^{\gamma_i}\right)$$
$$\times\left(\sum_{i=1}^{r}a_i(r)(1-F(v))^{\gamma_i}\right)\frac{f(v)f(v+w)}{(1-F(v))(1-F(v+w))}dv \tag{2.35}$$

ve $H(z;r,s)=(c_{s-1}/c_{r-1})\sum_{i=r+1}^{s}a_i^{(r)}(s)(1/\gamma_i)z^{\gamma_i}$, $z\in[0,1]$ olmak üzere,

$$F^{W_{r,s}}(w)=1-\int_0^{\infty}H\left(\frac{1-F(v+w)}{1-F(v)};r,s\right)dF^{X(r,n,\tilde{\mathbf{m}},k)}(v)\,. \tag{2.36}$$

Eğer dağılım fonksiyonu F, $\lambda > 0$ parametreli üstel dağılım olarak seçilirse, o zaman $X(r-1, n, \widetilde{\mathbf{m}}, k)$ ile $W_{r-1,r}$ rasgele değişkenlerinin birbirlerinden bağımsız olduğu gösterilmiştir (Ahsanullah, 2000).

Tablo 1 Genelleştirilmiş sıra istatistiklerinin alt modellere indirgenmesi

Sıralanmış rasgele değişken modeli adı	$\widetilde{\mathbf{m}}$ **vektörü**	k **parametresi**
1. *Sıra istatistikleri*	$i = 1,2,\ldots,n-1$ için $m_i = 0$	$k = 1$
2. *II. tip ilerleyen sansürlenmiş sıra istatistikleri*	$i = 1,2,\ldots,n-1$ için $m_i = R_i$	$k = R_n + 1$
3. Ardışık (sequential) sıra istatistikleri	$i = 1,2,\ldots,n-1$ ve $\alpha_1,\ldots,\alpha_n > 0$ için $m_i = (n-i+1)\alpha_i - (n-i)\alpha_{i+1} - 1$	$k = \alpha_n$
4. *Rekor değerler*	$i = 1,2,\ldots,n-1$ için $m_i = -1$	$k = 1$
5. k. *rekor değerler*	$i = 1,2,\ldots,n-1$ için $m_i = -1$	$k \in \{1,2,\ldots\}$
6. *Pfeifer' in rekor değerleri*	$i = 1,2,\ldots,n-1$, $j \geq 1$ ve $\beta_j > 0$ için $m_i = \beta_i - \beta_{i+1} - 1$	$k = \beta_n$

3 GENELLEŞTİRİLMİŞ SIRA İSTATİSTİKLERİNE DAYALI İSTATİSTİKLER

Bu bölümde, genelleştirilmiş sıra istatistiklerine dayalı dağılımdan-bağımsız güven aralıkları oluşturulmuş, aşan (exceedance) istatistiklerin olasılık fonksiyonları ve üstel dağılıma sahip kitleden alınan yeni gözlemlerin minimal boşluğa düşme olasılıkları elde edilmiştir.

3.1 Dağılımdan-Bağımsız (veya İnvaryant) Güven Aralıkları ve Aşan İstatistikler

Genelleştirilmiş sıra istatistiklerine dayalı aşan istatistiklerin olasılık fonksiyonlarının elde edilebilmesi için, öncelikle, genelleştirilmiş sıra istatistiklerine dayalı dağılımdan-bağımsız güven aralıkları oluşturulmalı ve güven düzeylerinin ne olduğu bulunmalıdır.

Eğer, Bölüm 1.2.1' de verilen $f_1(.)$ ve $f_2(.)$ Borel fonksiyonlarını, $X_1, X_2, ..., X_n$ birbirlerinden bağımsız ve sürekli $F \in \Im_c$ dağılım fonksiyonuna sahip bir örneklem ve $1 \leq r < s \leq n$ olmak üzere, $f_1(X_1, X_2, ..., X_n) = X(r, n, \tilde{\mathbf{m}}, k)$ ve $f_2(X_1, X_2, ..., X_n) = X(s, n, \tilde{\mathbf{m}}, k)$ şeklinde alacak olursak, $(X(r, n, \tilde{\mathbf{m}}, k), X(s, n, \tilde{\mathbf{m}}, k))$ rasgele aralığı $\Im_c$ sınıfı için bir dağılımdan-bağımsız güven aralığı olacaktır.

Şimdi *ilk adım olarak*, $X(1, n, \tilde{\mathbf{m}}, k), ..., X(n, n, \tilde{\mathbf{m}}, k)$, sürekli F dağılım fonksiyonu üzerine kurulu genelleştirilmiş sıra istatistikleri ve $1 \leq r < s \leq n$ olmak üzere, F dağılım fonksiyonuna sahip kitleden yeni bir gözlem olarak alınan X_{n+1}' in , $(X(r, n, \tilde{\mathbf{m}}, k), X(s, n, \tilde{\mathbf{m}}, k))$

dağılımdan-bağımsız güven aralığına düşmesi olasılığını bulalım: Eğer $X_{n+1} = X$, $X(r,n,\tilde{\mathbf{m}},k) = Y_r$ ve $X(s,n,\tilde{\mathbf{m}},k) = Y_s$ ile gösterilecek olursa,

$$P\{X \in (Y_r, Y_s)\} = \int_{-\infty}^{\infty}\int_{y_r}^{\infty}\int_{y_r}^{y_s} f(x) f^{Y_r,Y_s}(y_r, y_s) dx dy_s dy_r$$

$$= \int_{-\infty}^{\infty}\int_{y_r}^{\infty}\left(\int_{y_r}^{y_s} f(x)dx\right) f^{Y_r,Y_s}(y_r, y_s) dy_s dy_r \qquad (3.1)$$

eşitliği yazılabilir. Bu eşitliğin sağ tarafında parantez içinde yer alan integralin değeri yerine $(F(y_s) - F(y_r))$ ve $f^{Y_r,Y_s}(y_r, y_s)$ yerine (2.34) eşitliğindeki karşılığı yazılacak ve gerekli düzenlemeler yapılacak olursa, (3.1) eşitliğinin sağ tarafı,

$$c_{s-1}\sum_{j=1}^{r}\sum_{i=r+1}^{s} a_i^{(r)}(s)\, a_j(r) \int_{-\infty}^{\infty}(1 - F(y_r))^{\gamma_j - \gamma_i - 1}$$

$$\times \int_{y_r}^{\infty}(F(y_s) - F(y_r))(1 - F(y_s))^{\gamma_i - 1} dF(y_s) dF(y_r) \qquad (3.2)$$

olur. (3.2) ifadesinde, içteki integralde önce $F(y_s) = 1 - u$, sonra $u = (1 - F(y_r))v$ dönüşümleri yapılırsa, içteki integral için,

$$\int_{y_r}^{\infty}(F(y_s) - F(y_r))(1 - F(y_s))^{\gamma_i - 1} dF(y_s)$$

$$= (1 - F(y_r))^{\gamma_i + 1}\int_0^1 v^{\gamma_i - 1}(1 - v) dv \qquad (3.3)$$

eşitliği yazılabilir. Bu eşitliğin sağ tarafında yer alan integralin $B(\gamma_i, 2)$' ye eşit olduğu bilinmektedir. Bu noktada, (3.3) ifadesi (3.2) de yerine konacak olursa, (3.2) ifadesi

$$c_{s-1}\sum_{j=1}^{r}\sum_{i=r+1}^{s}a_i^{(r)}(s)\,a_j(r)B(\gamma_i,2)\int_{-\infty}^{\infty}(1-F(y_r))^{\gamma_j}dF(y_r) \qquad (3.4)$$

şekline gelecektir. Daha sonra, (3.4) ifadesindeki integralde $F(y_r)=u$ dönüşümü yapılıp integral alınır, beta fonksiyonu yerine değeri yazılır ve (3.4) ifadesi düzenlenirse,

$$c_{s-1}\sum_{j=1}^{r}\frac{a_j(r)}{\gamma_j+1}\sum_{i=r+1}^{s}\frac{a_i^{(r)}(s)}{\gamma_i(\gamma_i+1)} \qquad (3.5)$$

elde edilecektir. Sonuç olarak, $\sum_{j=1}^{r}\frac{a_j(r)}{\gamma_j+1}=\left(\prod_{j=1}^{r}(\gamma_j+1)\right)^{-1}$ olduğundan,

$$P\{X\in(Y_r,Y_s)\}=\frac{c_{s-1}}{\prod_{j=1}^{r}(\gamma_j+1)}\times\sum_{i=r+1}^{s}\frac{a_i^{(r)}(s)}{\gamma_i(\gamma_i+1)} \qquad (3.6)$$

eşitliğine ulaşılır ve yeni bir gözlemin genelleştirilmiş sıra istatistiklerine dayalı dağılımdan-bağımsız güven aralığına düşmesi olasılığı elde edilmiş olur.

İkinci adımda, $X(1,n,\tilde{\mathbf{m}},k),\ldots,X(n,n,\tilde{\mathbf{m}},k)$, sürekli F dağılım fonksiyonu üzerine kurulu genelleştirilmiş sıra istatistikleri, $1\le r<s\le n$ ve $1\le l$ olmak üzere, F dağılım fonksiyonuna sahip kitleden gelen l adet birbirlerinden bağımsız yeni gözlemin $(X_{n+1},X_{n+2},\ldots,X_{n+l})$, $(X(r,n,\tilde{\mathbf{m}},k),X(s,n,\tilde{\mathbf{m}},k))$ dağılımdan-bağımsız güven aralığına düşmesi olasılığının ne olduğunu gösterelim: Bu sefer, $X_{n+i}=X_i$, $X(r,n,\tilde{\mathbf{m}},k)=Y_r$, $X(s,n,\tilde{\mathbf{m}},k)=Y_s$ olmak üzere,

$$P\{X_1,...,X_l \in (Y_r,Y_s)\} = \int_{-\infty}^{\infty}\int_{y_r}^{\infty}\int_{y_r}^{y_s}...\int_{y_r}^{y_s} f(x_1)f(x_2)...f(x_l)$$

$$\times f^{Y_r,Y_s}(y_r,y_s)dx_1dx_2...dx_l dy_s dy_r \qquad (3.7)$$

yazılabilir. Bu eşitlik aynı zamanda,

$$P\{X_1,...,X_l \in (Y_r,Y_s)\}$$

$$= \int_{-\infty}^{\infty}\int_{y_r}^{\infty}\left(\int_{y_r}^{y_s} f(x_1)dx_1\right)...\left(\int_{y_r}^{y_s} f(x_l)dx_l\right) f^{Y_r,Y_s}(y_r,y_s)dy_s dy_r \qquad (3.8)$$

şeklinde de ifade edilebilir. Parantez içindeki l adet integralin çarpımı $(F(y_s)-F(y_r))^l$ ' ye eşit olduğu için (3.8) eşitliğinin sağ tarafı yerine,

$$\int_{-\infty}^{\infty}\int_{y_r}^{\infty}(F(y_s)-F(y_r))^l f^{Y_r,Y_s}(y_r,y_s)dy_s dy_r \qquad (3.9)$$

ifadesi yazılabilir. Bu ifadede $f^{Y_r,Y_s}(y_r,y_s)$ yerine (2.34)' teki karşılığı yazılıp gerekli düzenlemeler yapılırsa,

$$c_{s-1}\sum_{j=1}^{r}\sum_{i=r+1}^{s} a_i^{(r)}(s)\, a_j(r) \int_{-\infty}^{\infty}(1-F(y_r))^{\gamma_j-\gamma_i-1}$$

$$\times \int_{y_r}^{\infty}(F(y_s)-F(y_r))^l (1-F(y_s))^{\gamma_i-1} dF(y_s)dF(y_r) \qquad (3.10)$$

elde edilir. Daha sonra, sırasıyla, $F(y_s)=1-u$ ve $u=(1-F(y_r))v$ dönüşümleri yapılır ve integral alınırsa, (3.10) ifadesi,

$$c_{s-1}\sum_{j=1}^{r}\sum_{i=r+1}^{s} a_i^{(r)}(s)\, a_j(r)\, B(\gamma_i,l+1)\int_{-\infty}^{\infty}(1-F(y_r))^{\gamma_j+l-1}dF(y_r) \qquad (3.11)$$

şekline gelir. (3.11) ifadesinde yer alan integral $(\gamma_j + l)^{-1}$ 'e eşittir. Sonuç olarak, $\{X_1,\ldots,X_l \in (Y_r, Y_s)\}$ olayının olasılığı,

$$P\{X_1,\ldots,X_l \in (Y_r,Y_s)\} = c_{s-1} \times \sum_{j=1}^{r} \frac{a_j(r)}{\gamma_j + l} \times \sum_{i=r+1}^{s} a_i^{(r)}(s) B(\gamma_i, l+1). \qquad (3.12)$$

olarak elde edilir.

Üçüncü adımda ise, $X(1,n,\tilde{\mathbf{m}},k),\ldots,X(n,n,\tilde{\mathbf{m}},k)$, sürekli F dağılım fonksiyonu üzerine kurulu genelleştirilmiş sıra istatistikleri, $1 \le r < s \le n$ ve $1 \le l \le m$ olmak üzere, F dağılım fonksiyonuna sahip kitleden gelen m adet birbirlerinden bağımsız yeni gözlemden $(X_{n+1}, X_{n+2}, \ldots, X_{n+m})$, ilk l tanesinin $(X(r,n,\tilde{\mathbf{m}},k), X(s,n,\tilde{\mathbf{m}},k))$ dağılımdan-bağımsız güven aralığına düşmesi ve geriye kalan $(m-l)$ tanesinin ise düşmemesi olasılığını bulalım: $X_{n+i} = X_i$, $X(r,n,\tilde{\mathbf{m}},k) = Y_r$, $X(s,n,\tilde{\mathbf{m}},k) = Y_s$, $(-\infty, Y_r] = I_r$, $(Y_s, \infty] = I_s$ ve $(Y_r, Y_s) = I_{r,s}$ ile gösterilsin. m adet yeni gözlemden ilk l tanesinin $I_{r,s}$ aralığına düşmesi ve geriye kalan $(m-l)$ tanesinin ise bu aralığa düşmemesi olayı

$$\begin{aligned}\{X_1, X_2, \ldots, X_l \in I_{r,s} \text{ ve } & (X_{l+1} \in I_r \text{ veya } X_{l+1} \in I_s) \\ & \text{ve } \ldots \text{ ve } (X_m \in I_r \text{ veya } X_m \in I_s)\}\end{aligned}$$

şeklinde ifade edilebilir. Öyleyse,

$$P\{X_1, X_2, ..., X_l \in I_{r,s} \text{ ve } X_{l+1}, X_{l+2}, ..., X_m \notin I_{r,s}\}$$
$$= P\{X_1, X_2, ..., X_l \in I_{r,s} \text{ ve } (X_{l+1} \in I_r \text{ veya } X_{l+1} \in I_s)$$
$$\text{ve ... ve } (X_m \in I_r \text{ veya } X_m \in I_s)\} \tag{3.13}$$

yazılabilir. Bu eşitliğin sağ tarafının,

$$\sum_{t=0}^{m-l} \binom{m-l}{t} \times P\{X_1, X_2, ..., X_l \in I_{r,s}$$
$$\text{ve } t \text{ adet } X_i \in I_r \text{ ve } (m-l-t) \text{adet } X_i \in I_s\} \tag{3.14}$$

ifadesine eşit olduğu açıktır. (3.14)' de yer alan olasılık ifadesinin açılımı,

$$\int_{-\infty}^{\infty}\int_{y_r}^{\infty} \times \left(\underbrace{\int_{y_r}^{y_s} ... \int_{y_r}^{y_s}}_{(l \text{ adet})} \times \underbrace{\int_{-\infty}^{y_r} ... \int_{-\infty}^{y_r}}_{(t \text{ adet})} \times \underbrace{\int_{y_s}^{\infty} ... \int_{y_s}^{\infty}}_{(m-l-t \text{ adet})} \prod_{i=1}^{m} f(x_i) dx_j ... dx_1 \right)$$
$$\times f^{Y_r, Y_s}(y_r, y_s) dy_s dy_r \tag{3.15}$$

şeklindedir. Parantez içinde yer alan toplam m adet integralin $(F(y_s) - F(y_r))^l (F(y_r))^t (1 - F(y_s))^{m-l-t}$' ye eşit olduğu açıktır. Bu noktada, $f^{Y_r, Y_s}(y_r, y_s)$' nin (2.34) teki karşılığı da (3.15) ifadesinde yerine yazılır ve (3.15) ifadesi (3.14)' deki olasılık ifadesinin yerine yazılacak ve düzenlenecek olursa, (3.14) ifadesi,

$$c_{s-1} \sum_{t=0}^{m-l} \binom{m-l}{t} \times \sum_{j=1}^{r} \sum_{i=r+1}^{s} a_i^{(r)}(s)\, a_j(r) \times \int_{-\infty}^{\infty} (F(y_r))^t (1 - F(y_r))^{\gamma_j - \gamma_i - 1}$$
$$\times \int_{y_r}^{\infty} (F(y_s) - F(y_r))^l (1 - F(y_s))^{\gamma_i + m - l - t - 1} dF(y_s) dF(y_r) \tag{3.16}$$

şekline gelir. Bu ifadede, önce $F(y_s)=1-u$ ve $u=(1-F(y_r))v$ dönüşümleri sırayla yapılıp içteki integral alınır, sonra $F(y_r)=z$ dönüşümü yapılıp dıştaki integral alınırsa,

$$\begin{aligned} &P\{X_1,\dots,X_l \in (Y_r,Y_s)\} \\ &\quad = c_{s-1}\times\sum_{t=0}^{m-l}\binom{m-l}{t}\times\left(\sum_{j=1}^{r}a_j(r)B(t+1,\gamma_j+m-t)\right) \\ &\quad\quad \times\left(\sum_{i=r+1}^{s}a_i^{(r)}(s)B(\gamma_i+m-l-t,l+1)\right) \end{aligned} \tag{3.17}$$

olasılığı elde edilmiş olur.

Yukarıdaki dağılımdan bağımsız güven aralıkları ve güven düzeylerine dayalı olarak, aşan istatistiğin olasılık fonksiyonu aşağıdaki teoremde verilmektedir.

Teorem 1. $X(1,n,\tilde{\mathbf{m}},k),\dots,X(n,n,\tilde{\mathbf{m}},k)$, sürekli F dağılım fonksiyonu üzerine kurulu genelleştirilmiş sıra istatistikleri, $(X_{n+1},X_{n+2},\dots,X_{n+m})$ F dağılım fonksiyonuna sahip kitleden gelen m adet birbirlerinden bağımsız yeni gözlem ve $i=1,2,\dots,m$ olmak üzere, ξ_i rasgele değişkeni,

$$\xi_i=\begin{cases}1 & , \; X_{n+i}\in(X(r,n,\tilde{\mathbf{m}},k),X(s,n,\tilde{\mathbf{m}},k)) \\ 0 & , \; X_{n+i}\notin(X(r,n,\tilde{\mathbf{m}},k),X(s,n,\tilde{\mathbf{m}},k))\end{cases}$$

şeklinde tanımlansın. $S_m=\sum_{i=1}^{m}\xi_i$ olarak tanımlanan ve $X_{n+1},X_{n+2},\dots,X_{n+m}$ gözlemlerinden $(X(r,n,\tilde{\mathbf{m}},k),X(s,n,\tilde{\mathbf{m}},k))$ dağılımdan-bağımsız güven aralığına düşenlerin sayısını gösteren aşan istatistiğin değerinin l' ye eşit olması olasılığı, yani S_m rasgele değişkeninin olasılık fonksiyonu,

$$P\{S_m = l\} = \binom{m}{l} \times c_{s-1} \times \sum_{t=0}^{m-l} \binom{m-l}{t} \times \left(\sum_{j=1}^{r} a_j(r) B(t+1, \gamma_j + m - t) \right)$$
$$\times \left(\sum_{i=r+1}^{s} a_i^{(r)}(s) B(\gamma_i + m - l - t, l + 1) \right)$$

dir.

İspat: $\{S_m = l\}$ olayı, her $j \neq t$ için $i_j \neq i_t$ olmak üzere,

$$\{X_{n+i_1}, \ldots, X_{n+i_l} \in (X(r,n,\tilde{\mathbf{m}},k), X(s,n,\tilde{\mathbf{m}},k)),$$
$$X_{n+i_{l+1}}, \ldots, X_{n+i_m} \notin (X(r,n,\tilde{\mathbf{m}},k), X(s,n,\tilde{\mathbf{m}},k))\}$$

şeklinde de ifade edilebilir. Bu durumda,

$$\begin{aligned} &P\{S_m = l\} \\ &=_{i_1,\ldots,i_m} P\{X_{n+i_1}, \ldots, X_{n+i_l} \in (X(r,n,\tilde{\mathbf{m}},k), X(s,n,\tilde{\mathbf{m}},k)), \\ &\qquad X_{n+i_{l+1}}, \ldots, X_{n+i_m} \notin (X(r,n,\tilde{\mathbf{m}},k), X(s,n,\tilde{\mathbf{m}},k))\} \end{aligned} \tag{3.18}$$

yazılabilir. $(X(r,n,\tilde{\mathbf{m}},k), X(s,n,\tilde{\mathbf{m}},k))$ aralığına, m tane rasgele değişkenden l tanesinin, $C_m^l = \binom{m}{l}$ farklı şekilde düşeceği açıktır. Ayrıca, $\binom{m}{l}$ farklı durumun her birinin olasılığı (3.17)' de verilen olasılıkla aynıdır. Bu durumda, (3.18) yerine,

$$P\{S_m = l\} = \binom{m}{l} \times P\{X_{n+1},\dots,X_{n+l} \in (X(r,n,\tilde{\mathbf{m}},k), X(s,n,\tilde{\mathbf{m}},k)),$$
$$X_{n+l+1},\dots,X_{n+m} \notin (X(r,n,\tilde{\mathbf{m}},k), X(s,n,\tilde{\mathbf{m}},k))\}$$
$$= \binom{m}{l} \times c_{s-1} \times \sum_{t=0}^{m-l} \binom{m-l}{t} \times \left(\sum_{j=1}^{r} a_j(r) B(t+1, \gamma_j + m - t) \right)$$
$$\times \left(\sum_{i=r+1}^{s} a_i^{(r)}(s) B(\gamma_i + m - l - t, l+1) \right)$$

yazılabilir ve teoremin ispatı tamamlanmış olur. □

3.2 Üstel Dağılım İçin Bağımsızlık Özelliği

Ahsanullah (2000), genelleştirilmiş sıra istatistiklerinin üstel dağılım üzerine kurulması durumunda, $X(0,n,\tilde{\mathbf{m}},k)=0$ ve $1 \le r \le n$ olmak üzere, $X(r-1,n,\tilde{\mathbf{m}},k)$ ile $X(r,n,\tilde{\mathbf{m}},k)-X(r-1,n,\tilde{\mathbf{m}},k)$ rasgele değişkenlerinin birbirlerinden bağımsız olduğunu bir teorem olarak vermiştir. Şimdi, Ahsanullah' ın bu teoremini, $1 \le r < s \le n$ olmak üzere, $X(r,n,\tilde{\mathbf{m}},k)$ ile $X(s,n,\tilde{\mathbf{m}},k)-X(r,n,\tilde{\mathbf{m}},k)$ rasgele değişkenlerinin bağımsız olduğu şeklinde genellemek için aşağıdaki teoremi verelim:

Teorem 2. $X(1,n,\tilde{\mathbf{m}},k),\dots,X(n,n,\tilde{\mathbf{m}},k)$, $F(x)=1-e^{-\lambda x}$ $(x \ge 0, \lambda > 0)$ dağılım fonksiyonu üzerine kurulu genelleştirilmiş sıra istatistikleri, $X(0,n,\tilde{\mathbf{m}},k)=0$ ve $1 \le r < s \le n$ olmak üzere, $X(r,n,\tilde{\mathbf{m}},k)$ ile $X(s,n,\tilde{\mathbf{m}},k)-X(r,n,\tilde{\mathbf{m}},k)$ rasgele değişkenleri bağımsızdır.

İspat: $X(r,n,\tilde{\mathbf{m}},k)=Y_r$ ve $X(s,n,\tilde{\mathbf{m}},k)=Y_s$ olmak üzere,

$$F^{Y_r,Y_s-Y_r}(x,w)=P\{Y_r\le x,Y_s-Y_r\le w\}$$

$$=\int_0^x\int_{y_r}^{y_r+w}f^{Y_r,Y_s}(y_r,y_s)dy_s dy_r \tag{3.20}$$

yazılabilir. r ' inci ve s ' inci genelleştirilmiş sıra istatistiklerinin birleşik olasılık yoğunluk fonksiyonu (2.34)' teki gibi yazılır ve bu birleşik olasılık yoğunluk fonksiyonunda yer alan f ve F yerine sırayla λ parametreli üstel dağılımın olasılık yoğunluk fonksiyonu ve dağılım fonksiyonu yazılacak olursa, (3.20) eşitliğinin sağ tarafı,

$$\lambda^2 c_{s-1}\sum_{j=r+1}^{s}\sum_{i=1}^{r}a_j^{(r)}(s)a_i(r)\int_0^x e^{-\lambda(\gamma_i-\gamma_j)y_r}\left(\int_{y_r}^{y_r+w}e^{-\lambda\gamma_j y_s}dy_s\right)dy_r \tag{3.21}$$

olacaktır. Parantez içindeki integral, $e^{-\lambda\gamma_j y_r}\left(1-e^{-\lambda\gamma_j w}\right)/\lambda\gamma_j$ ' ye eşittir. Bu durumda (3.21) ifadesi,

$$\lambda c_{s-1}\sum_{j=r+1}^{s}\sum_{i=1}^{r}\frac{a_j^{(r)}(s)}{\gamma_j}\left(1-e^{-\lambda\gamma_j w}\right)a_i(r)\int_0^x e^{-\lambda\gamma_i y_r}dy_r \tag{3.22}$$

haline gelecektir. Yukarıdaki ifadede yer alan integral alınır ve bu ifade düzenlenlenirse,

$$c_{s-1}\sum_{j=r+1}^{s}\sum_{i=1}^{r}\frac{a_j^{(r)}(s)}{\gamma_j}\left(1-e^{-\lambda\gamma_j w}\right)\frac{a_i(r)}{\gamma_i}\left(1-e^{-\lambda\gamma_i x}\right) \tag{3.23}$$

ifadesine ulaşılır. (3.23) ifadesi,

$$c_{s-1}\left(\sum_{j=r+1}^{s}\frac{a_j^{(r)}(s)}{\gamma_j}\left(1-e^{-\lambda\gamma_j w}\right)\right)\left(\sum_{i=1}^{r}\frac{a_i(r)}{\gamma_i}\left(1-e^{-\lambda\gamma_i x}\right)\right) \tag{3.24}$$

şeklinde de yazılabilir. Bu noktada, toplam sembolleri içindeki çarpma işlemleri yapılır ve toplam sembolleri terimlere dağıtılacak olursa, (3.24) ifadesi,

$$c_{s-1}\left(\sum_{j=r+1}^{s}\frac{a_j^{(r)}(s)}{\gamma_j}-\sum_{j=r+1}^{s}\frac{a_j^{(r)}(s)}{\gamma_j}e^{-\lambda\gamma_j w}\right)\left(\sum_{i=1}^{r}\frac{a_i(r)}{\gamma_i}-\sum_{i=1}^{r}\frac{a_i(r)}{\gamma_i}e^{-\lambda\gamma_i x}\right) \tag{3.25}$$

olarak yazılabilecektir. Bu son ifadede yer alan $\sum_{j=r+1}^{s}a_j^{(r)}(s)/\gamma_j=c_{r-1}/c_{s-1}$ ve $\sum_{i=1}^{r}a_i(r)/\gamma_i=1/c_{r-1}$ olduğundan, (3.20) eşitliğinin sağ tarafı olan (3.25) ifadesi, ilgili düzenlemeler yapıldıktan sonra (3.20)' de yerine koyulursa,

$$\begin{aligned}F^{Y_r,Y_s-Y_r}(x,w)=1-c_{r-1}&\sum_{i=1}^{r}\frac{a_i(r)}{\gamma_i}e^{-\lambda\gamma_i x}\\&-\frac{c_{s-1}}{c_{r-1}}\sum_{j=r+1}^{s}\frac{a_j^{(r)}(s)}{\gamma_j}e^{-\lambda\gamma_j w}\\&+c_{s-1}\sum_{j=r+1}^{s}\frac{a_j^{(r)}(s)}{\gamma_j}e^{-\lambda\gamma_j w}\sum_{i=1}^{r}\frac{a_i(r)}{\gamma_i}e^{-\lambda\gamma_i x}\end{aligned} \tag{3.26}$$

elde edilecektir. Böylelikle, Y_r ve Y_s-Y_r rasgele değişkenlerinin birleşik dağılım fonksiyonu elde edilmiş olmaktadır. Y_r ve Y_s-Y_r rasgele değişkenlerinin marjinal dağılım fonksiyonları ise, sırayla,

$$F^{Y_r}(x)=1-c_{r-1}\sum_{i=1}^{r}\frac{a_i(r)}{\gamma_i}e^{-\lambda\gamma_i x} \tag{3.27}$$

ve

$$F^{Y_s-Y_r}(w)=1-\frac{c_{s-1}}{c_{r-1}}\sum_{j=r+1}^{s}\frac{a_j^{(r)}(s)}{\gamma_j}e^{-\lambda\gamma_j w} \tag{3.28}$$

şeklinde olup, $F^{Y_r,Y_s-Y_r}(x,w)=F^{Y_r}(x)F^{Y_s-Y_r}(w)$ olduğu görülmektedir. Sonuç olarak, Y_r ve Y_s-Y_r rasgele değişkenlerinin bağımsız olduğu ispatlanmış olmaktadır. □

3.3 Üstel Dağılım İçin Minimal Boşluk'a Dayalı İstatistikler

$X_1,X_2,...,X_n$, değer kümesi negatif olmayan sürekli F dağılım fonksiyonundan gelen bir örneklem, $X(1,n,\tilde{\mathbf{m}},k)<X(2,n,\tilde{\mathbf{m}},k)<...<X(n,n,\tilde{\mathbf{m}},k)$ ise dağılım fonksiyonu F ' e dayalı olan genelleştirilmiş sıra istatistikleri olsun. $X(0,n,\tilde{\mathbf{m}},k)=0$ olmak üzere, $(X(1,n,\tilde{\mathbf{m}},k)-X(0,n,\tilde{\mathbf{m}},k))$, $(X(2,n,\tilde{\mathbf{m}},k)-X(1,n,\tilde{\mathbf{m}},k))$,..., $(X(n,n,\tilde{\mathbf{m}},k)-X(n-1,n,\tilde{\mathbf{m}},k))$ boşlukları dikkate alındığında, minimal genişliğe sahip boşluğun indeksini gösteren v rasgele değişkeni, $i=1,2,...,n$ olmak üzere,

$$v=r \Leftrightarrow X(r,n,\tilde{\mathbf{m}},k)-X(r-1,n,\tilde{\mathbf{m}},k)\leq X(i,n,\tilde{\mathbf{m}},k)-X(i-1,n,\tilde{\mathbf{m}},k)$$

şeklinde tanımlanır (Bairamov ve Eryılmaz, 2000). F dağılım fonksiyonu, $\lambda>0$ parametreli üstel dağılım olduğunda, $i=1,2,...,n$ olmak üzere, $X(i,n,\tilde{\mathbf{m}},k)-X(i-1,n,\tilde{\mathbf{m}},k)$ boşluklarının birbirlerinden bağımsız oldukları bilinmektedir. Ayrıca, bu boşlukların genelleştirilmiş sıra istatistiklerinden de bağımsız olduğu bilinmektedir. Buna ilave olarak, $F(x)=1-e^{-\lambda x}, x\geq 0, \lambda>0$, için

$$f^{X(i,n,\tilde{\mathbf{m}},k)}(x) = c_{i-1}\lambda\sum_{j=1}^{i} a_j(i)e^{-\lambda\gamma_j x} \quad , \tag{3.29}$$

$$F^{X(i,n,\tilde{\mathbf{m}},k)}(x) = 1 - c_{i-1}\sum_{j=1}^{i}\frac{a_j(i)}{\gamma_j}e^{-\lambda\gamma_j x} \, , \tag{3.30}$$

$$f^{X(i,n,\tilde{\mathbf{m}},k)-X(i-1,n,\tilde{\mathbf{m}},k)}(w) = \gamma_i\lambda\, e^{-\lambda\gamma_i w} \tag{3.31}$$

ve

$$F^{X(i,n,\tilde{\mathbf{m}},k)-X(i-1,n,\tilde{\mathbf{m}},k)}(w) = 1 - e^{-\lambda\gamma_i w} \tag{3.32}$$

olmaktadır. Bu bilgilerin ışığı altında, aşağıdaki teorem verilebilir.

Teorem 3. $X(1,n,\tilde{\mathbf{m}},k),\ldots,X(n,n,\tilde{\mathbf{m}},k)$, $F(x) = 1 - e^{-\lambda x}$ $(x \geq 0, \lambda > 0)$ dağılım fonksiyonu üzerine kurulu genelleştirilmiş sıra istatistikleri, $X(0,n,\tilde{\mathbf{m}},k) = 0$ ve v minimal boşluğun indeksini gösteren rasgele değişken olmak üzere, $r = 1,2,\ldots,n$ için,

$$P\{v = r\} = \frac{\gamma_r}{\sum_{i=1}^{n}\gamma_i}$$

dir.

İspat: $j = 1,2,\ldots,n$ ve $X(0,n,\tilde{\mathbf{m}},k) = 0$ için $X(j,n,\tilde{\mathbf{m}},k) - X(j-1,n,\tilde{\mathbf{m}},k) = W_j$ olmak üzere, $\{v = r\} \equiv \{W_r \leq W_1, W_r \leq W_2, \ldots, W_r \leq W_n\}$ ve $j \neq i$ için W_j ile W_i bağımsız olduğundan,

$$P\{v=r\}=\int_0^{\infty}\int_{w_r}^{\infty}\ldots\int_{w_r}^{\infty}f^{W_1}(w_1)f^{W_2}(w_2)\ldots f^{W_n}(w_n)dw_n$$

$$\times dw_{n-1}\ldots dw_{r+1}dw_{r-1}dw_{r-2}\ldots dw_1dw_r \tag{3.33}$$

yazılabilir. Yukarıdaki eşitliğin sağ tarafı,

$$\int_0^{\infty}\left(\prod_{j=1}^{r-1}\left(\int_{w_r}^{\infty}f^{W_j}(w_j)dw_j\right)\Bigg/\int_{w_r}^{\infty}f^{W_r}(w_r)dw_r\right)\times f^{W_r}(w_r)dw_r \tag{3.34}$$

şeklinde de yazılabilir. $\int_{w_r}^{\infty}f^{W_j}(w_j)dw_j=1-F^{W_j}(w_r)$ ve $\int_{w_r}^{\infty}f^{W_r}(w_r)dw_r=1-F^{W_r}(w_r)$ olduğundan, (3.34) ifadesi,

$$\int_0^{\infty}\left(\prod_{j=1}^{r-1}\left(1-F^{W_j}(w_r)\right)\Big/\left(1-F^{W_r}(w_r)\right)\right)\times f^{W_r}(w_r)dw_r \tag{3.35}$$

haline gelecektir. Bir sonraki aşamada, $1-F^{W_j}(w_r)=e^{-\lambda\gamma_j w_r}$ ve $f^{W_r}(w_r)=\gamma_r\lambda\, e^{-\lambda\gamma_r w_r}$ olduğundan, (3.35) ifadesi,

$$\lambda\, c_{r-1}\sum_{i=1}^{r-1}\frac{a_i(r-1)}{\gamma_i}\int_0^{\infty}e^{-\lambda\sum_{j=1}^{n}\gamma_j w_r}dw_r \tag{3.36}$$

şeklinde yazılabilir. Bu ifadedeki integral alınacak olursa,

$$\frac{1}{\sum_{j=1}^{n}\gamma_j}c_{r-1}\sum_{i=1}^{r-1}\frac{a_i(r-1)}{\gamma_i} \tag{3.37}$$

ifadesine ulaşılır. Son olarak, $c_{r-1}\sum_{i=1}^{r-1}\frac{a_i(r-1)}{\gamma_i}=\gamma_r$ olduğu için, $\{v=r\}$ olayının olasılığı,

$$P\{v=r\}=\frac{\gamma_r}{\sum_{i=1}^{n}\gamma_i} \tag{3.38}$$

şeklinde yazılır ve böylece teorem ispatlanmış olur.□

Şimdi, minimal boşluğa yeni gözlemlerin düşme olasılıklarını elde edelim. İlk olarak, $X(1,n,\widetilde{\mathbf{m}},k),\ldots,X(n,n,\widetilde{\mathbf{m}},k)$, $F(x)=1-e^{-\lambda x}$ ($x\geq 0,\lambda>0$) dağılım fonksiyonu üzerine kurulu genelleştirilmiş sıra istatistikleri, $X(0,n,\widetilde{\mathbf{m}},k)=0$ ve v minimal boşluğun indeksini gösteren rasgele değişken olmak üzere, $F(x)$ dağılım fonksiyonuna sahip yeni bir gözlem olan X_{n+1}' in minimal boşluğa düşmesinin olasılığını bulmaya çalışalım: $i=1,2,\ldots,n$ için $Y_i=X(i,n,\widetilde{\mathbf{m}},k)$, $W_i=Y_i-Y_{i-1}$, $Y_0=X(0,n,\widetilde{\mathbf{m}},k)=0$ ve $X=X_{n+1}$ olmak üzere, $\{X\in(Y_{v-1},Y_v)\}$ olayının olasılığı,

$$P\{X\in(Y_{v-1},Y_v)\}=P\{X\in(Y_0,Y_1),v=1\}$$
$$+\sum_{r=2}^{n}P\{X\in(Y_{r-1},Y_r),v=r\} \tag{3.39}$$

şeklinde hesaplanabilir. (3.39) eşitliğinin sağ tarafındaki ilk olasılık dikkate alınacak olursa, bu olasılık,

$$P\{X\in(Y_0,Y_1),v=1\}$$
$$=P\{X\in(Y_0,Y_1),W_1\leq W_2,W_1\leq W_3,\ldots,W_1\leq W_n\} \tag{3.40}$$

olarak yazılabilir. $W_1 = Y_1$ tanımlandığı için, eşitliğin sağ tarafındaki olasılık ifadesi,

$$\int_0^\infty P\{X \in (Y_0, Y_1), Y_1 \le W_2, Y_1 \le W_3, ..., Y_1 \le W_n | Y_1 = t_1\} f^{Y_1}(t_1) dt_1 \qquad (3.41)$$

biçiminde yazılabilir. Bu ifade, aynı zamanda,

$$\int_0^\infty P\{X \in (0, t_1), t_1 \le W_2, t_1 \le W_3, ..., t_1 \le W_n\} f^{Y_1}(t_1) dt_1 \qquad (3.42)$$

ifadesi şeklinde de yazılabilir. $j \ne i$ olmak üzere, W_j, W_i' den ve X, W_j' den bağımsız olduğu için,

$$\begin{aligned} &P\{X \in (0, t_1), t_1 \le W_2, t_1 \le W_3, ..., t_1 \le W_n\} \\ &\quad = P\{X \in (0, t_1)\} P\{t_1 \le W_2\} P\{t_1 \le W_3\} ... P\{t_1 \le W_n\} \end{aligned} \qquad (3.43)$$

olacaktır. (3.43) eşitliğinin sağ tarafı,

$$(1 - e^{-\lambda t_1}) e^{-\lambda \gamma_2 t_1} e^{-\lambda \gamma_3 t_1} ... e^{-\lambda \gamma_n t_1}$$

ifadesine eşittir. Bu son ifade, $(1 - e^{-\lambda t_1}) e^{-\lambda t_1 \sum_{j=2}^{n} \gamma_j}$ şeklinde de yazılabilir. O zaman (3.42) ifadesi,

$$\int_0^\infty (1 - e^{-\lambda t_1}) e^{-\lambda t_1 \sum_{j=2}^{n} \gamma_j} f^{Y_1}(t_1) dt_1 \qquad (3.44)$$

olacaktır. $f^{Y_1}(t_1)$ yerine $\lambda \gamma_1 e^{-\lambda \gamma_1 t_1}$ yazılıp (3.44) ifadesi düzenlenirse,

$$\lambda \gamma_1 \int_0^{\infty} \left(1 - e^{-\lambda t_1}\right) e^{-\lambda t_1 \sum_{j=1}^{n} \gamma_j} dt_1 \tag{3.45}$$

elde edilir. (3.45) ifadesinde yer alan integral alınıp, bu ifade düzenlenirse ve (3.40)' ta yerine koyulursa,

$$P\{X \in (Y_0, Y_1), v = 1\} = \frac{\gamma_1}{\left(\sum_{j=1}^{n} \gamma_j\right)\left(1 + \sum_{j=1}^{n} \gamma_j\right)} \tag{3.46}$$

eşitliğine ulaşılır. Diğer yandan, (3.39) eşitliğinin sağ tarafındaki toplam içinde yer alan ikinci olasılık dikkate alınacak olursa, bu olasılık, $r = 2,3,\ldots,n$ olmak üzere,

$$\begin{aligned} &P\{X \in (Y_{r-1}, Y_r), v = r\} \\ &\quad = P\{X \in (Y_{r-1}, Y_r), W_r \leq Y_1, W_r \leq W_2, \ldots, W_r \leq W_n\} \end{aligned} \tag{3.47}$$

olarak yazılabilir. Eşitliğin sağ tarafındaki olasılık ifadesi,

$$\begin{aligned} &\int_0^{\infty}\int_0^{\infty} P\{X \in (Y_{r-1}, Y_r), W_r \leq Y_1, W_r \leq W_2, \ldots, W_r \leq W_n | \\ &\qquad W_r = t_1, Y_{r-1} = t_2\} f^{W_r}(t_1) f^{Y_{r-1}}(t_2) dt_1 dt_2 \end{aligned} \tag{3.48}$$

biçiminde açılabilir. Bu ifade, aynı zamanda,

$$\begin{aligned} &\int_0^{\infty}\int_0^{\infty} P\{X \in (t_2, t_1 + t_2), t_1 \leq Y_1, t_1 \leq W_2, \ldots, t_1 \leq W_n\} \\ &\qquad \times f^{W_r}(t_1) f^{Y_{r-1}}(t_2) dt_1 dt_2 \end{aligned} \tag{3.49}$$

ifadesi şeklinde de yazılabilir. $j \neq i$ olmak üzere W_j ve W_i, W_j ve X, W_j ve Y_1, X ve Y_1 birbirlerinden bağımsız oldukları için,

$$P\{X \in (t_2, t_1 + t_2), t_1 \le Y_1, t_1 \le W_2, \ldots, t_1 \le W_n\}$$
$$= P\{X \in (t_2, t_1 + t_2)\} P\{t_1 \le Y_1\} P\{t_1 \le W_2\} \ldots P\{t_1 \le W_n\} \qquad (3.50)$$

olacaktır. (3.50) eşitliğinin sağ tarafı,

$$e^{-\lambda t_2}\left(1 - e^{-\lambda t_1}\right) e^{-\lambda \gamma_1 t_1} e^{-\lambda \gamma_2 t_1} \ldots e^{-\lambda \gamma_{r-1} t_1} e^{-\lambda \gamma_{r+1} t_1} e^{-\lambda \gamma_{r+2} t_1} \ldots e^{-\lambda \gamma_n t_1}$$

ifadesine eşittir. Bu son ifade, $e^{-\lambda t_2}\left(1 - e^{-\lambda t_1}\right) e^{-\lambda \sum_{j=1}^{r-1} \gamma_j t_1} e^{-\lambda \sum_{j=r+1}^{n} \gamma_j t_1}$ şeklinde de yazılabilir. O zaman (3.47) ifadesinin sağ tarafı,

$$\int_0^\infty \int_0^\infty e^{-\lambda t_2}\left(1 - e^{-\lambda t_1}\right) e^{-\lambda \sum_{j=1}^{r-1} \gamma_j t_1} e^{-\lambda \sum_{j=r+1}^{n} \gamma_j t_1} f^{W_r}(t_1) f^{Y_{r-1}}(t_2) dt_1 dt_2 \qquad (3.51)$$

olacaktır. $f^{W_r}(t_1)$ ve $f^{Y_{r-1}}(t_2)$ yerine sırayla $\lambda \gamma_r e^{-\lambda \gamma_r t_1}$ ve $c_{r-2} \lambda \sum_{i=1}^{r-1} a_i (r-1) e^{-\lambda \gamma_i t_2}$ yazılıp (3.51) ifadesi düzenlenirse,

$$\lambda^2 c_{r-1} \sum_{i=1}^{r-1} a_i (r-1) \times \left(\int_0^\infty e^{-\lambda(\gamma_i + 1) t_2} dt_2 \right)$$
$$\times \int_0^\infty \left(e^{-\lambda \sum_{j=1}^{n} \gamma_j t_1} - e^{-\lambda \left(1 + \sum_{j=1}^{n} \gamma_j\right) t_1} \right) dt_1 \qquad (3.52)$$

elde edilir. (3.52) ifadesinde yer alan integraller alınıp, bu ifade düzenlenirse,

$$\frac{c_{r-1} \sum_{i=1}^{r-1} \dfrac{a_i (r-1)}{(\gamma_i + 1)}}{\left(\sum_{j=1}^{n} \gamma_j \right)\left(1 + \sum_{j=1}^{n} \gamma_j \right)} \qquad (3.53)$$

ifadesine ulaşılır. Burada yer alan $c_{r-1}\sum_{i=1}^{r-1} a_i(r-1)/(\gamma_i+1)$ ifadesinin $(\gamma_r+1)\prod_{i=1}^{r}\gamma_i/(\gamma_i+1)$' e eşit olduğu kolayca gösterilebilir. O zaman, (3.53) ifadesi yeniden düzenlenir ve (3.47)' de yerine koyulursa,

$$P\{X\in(Y_{r-1},Y_r),v=r\}=\frac{(\gamma_r+1)\prod_{i=1}^{r}\frac{\gamma_i}{(\gamma_i+1)}}{\left(\sum_{j=1}^{n}\gamma_j\right)\left(1+\sum_{j=1}^{n}\gamma_j\right)}$$

eşitliği elde edilir. Bu son ifade ve (3.46), (3.39)' da yerine koyulursa, $\{X\in(Y_{v-1},Y_v)\}$ olayının olasılığı,

$$P\{X\in(Y_{v-1},Y_v)\}=\frac{\sum_{r=1}^{n}(\gamma_r+1)\prod_{i=1}^{r}\frac{\gamma_i}{(\gamma_i+1)}}{\left(\sum_{j=1}^{n}\gamma_j\right)\left(1+\sum_{j=1}^{n}\gamma_j\right)} \tag{3.54}$$

şeklinde elde edilmiş olur.

Şimdi, "Tek bir yeni gözlem yerine, l adet yeni gözlem olsaydı, bunların minimal aralığa düşme olasılığı ne olurdu?" sorusunun cevabını arayalım. $X(1,n,\tilde{\mathbf{m}},k),\ldots,X(n,n,\tilde{\mathbf{m}},k)$, $F(x)=1-e^{-\lambda x}$ $(x\geq 0,\lambda>0)$ dağılım fonksiyonu üzerine kurulu genelleştirilmiş sıra istatistikleri, $X(0,n,\tilde{\mathbf{m}},k)=0$ ve v minimal boşluğun indeksini gösteren rasgele değişken olmak üzere, $F(x)$ dağılım fonksiyonuna sahip l adet bağımsız yeni gözlemin minimal boşluğa düşmesi olasılığı, $i=1,2,\ldots,n$ için

$X(i,n,\tilde{\mathbf{m}},k)=Y_i$, $Y_i-Y_{i-1}=W_i$, $Y_0=X(0,n,\tilde{\mathbf{m}},k)=0$ ve $j=1,2,...$ için $X_{n+j}=X_j$ olmak üzere,

$$P\{X_1,X_2,...,X_l\in(Y_{v-1},Y_v)\}=P\{X_1,X_2,...,X_l\in(Y_0,Y_1),v=1\}$$
$$+\sum_{r=2}^{n}P\{X_1,X_2,...,X_l\in(Y_{r-1},Y_r),v=r\} \quad (3.55)$$

olarak yazılabilir. (3.55)' de yer alan $P\{X_1,X_2,...,X_l\in(Y_0,Y_1),v=1\}$ olasılığı,

$$P\{X_1,X_2,...,X_l\in(0,Y_1)\text{ve } W_2,...,W_n\in[Y_1,\infty)\} \quad (3.56)$$

şeklinde ifade edilebilir. Bu olasılık ifadesini, koşullu olasılık kuralını kullanılarak,

$$\int_0^\infty P\{X_1,X_2,...,X_l\in(0,Y_1),W_2,...,W_n\in[Y_1,\infty)|Y_1=t_1\}$$
$$\times f^{Y_1}(t_1)dt_1 \quad (3.57)$$

yazılabilir. Yukarıdaki ifadenin,

$$\int_0^\infty P\{X_1,X_2,...,X_l\in(0,t_1),W_2,...,W_n\in[t_1,\infty)\}f^{Y_1}(t_1)dt_1 \quad (3.58)$$

ifadesi şeklinde de yazılabileceği açıktır. (3.58) ifadesindeki olasılık, X' lerin ve W' ların kendi içlerinde ve aralarında bağımsız oldukları bilindiğinden,

$$P\{X_1,X_2,...,X_l\in(0,t_1)\text{ve } W_2,...,W_n\in[t_1,\infty)\}$$
$$=\left(\prod_{i=1}^{l}P\{X_i\in(0,t_1)\}\right)\times\left(\prod_{j=2}^{n}P\{W_j\in(t_1,\infty)\}\right) \quad (3.59)$$

şeklinde yazılabilir. F, üstel dağılımın dağılım fonksiyonu olduğunda, $\{X_i \in (0,t_1)\}$ olayının olasılığı $(1-e^{-\lambda t_1})$ ve $\{W_j \in (t_1,\infty)\}$ olayının olasılığı ise $e^{-\lambda\gamma_j t_1}$ olacaktır. Bu bilgilere dayalı olarak, (3.59) eşitliğinin sağ tarafı yerine,

$$\left(1-e^{-\lambda t_1}\right)^l e^{-\lambda\sum_{j=2}^{n}\gamma_j t_1} \tag{3.60}$$

yazılabilir. Y_1 rasgele değişkenlerinin olasılık yoğunluk fonksiyonu ile (3.60) ifadesi (3.58)' de yerine yazılacak ve bu ifade düzenlenecek olursa,

$$\lambda\gamma_1\int_0^{\infty}\left(1-e^{-\lambda t_1}\right)^l e^{-\lambda\sum_{j=1}^{n}\gamma_j t_1}\,dt_1 \tag{3.61}$$

elde edilir. t_1'e göre integral alınırsa ve (3.57)' de yerine koyulursa,

$$P\{X_1,X_2,\ldots,X_l \in (0,Y_1)\,\text{ve}\; W_2,\ldots,W_n \in [Y_1,\infty)\}$$

$$=\gamma_1 B\left(\sum_{j=1}^{n}\gamma_j, l+1\right) \tag{3.62}$$

eşitliğine ulaşılır. Bu noktadan sonra, (3.55)' deki toplam içinde yer alan $P\{X_1,X_2,\ldots,X_l \in (Y_{r-1},Y_r), v=r\}$ olasılığı ($r=2,3,\ldots,n$),

$$P\{X_1,X_2,\ldots,X_l \in (Y_{r-1},Y_r)\,\text{ve}\; W_1,W_2,\ldots,W_n \in [W_r,\infty)\} \tag{3.63}$$

şeklinde ifade edilebilir. (3.63)' deki olasılık yerine, koşullu olasılık kuralı kullanılarak,

$$\int_0^\infty\int_0^\infty P\{X_1,X_2,\dots,X_l \in (Y_{r-1},Y_r),$$
$$W_1,W_2,\dots,W_n \in [W_r,\infty) | W_r = t_1, Y_{r-1} = t_2\}$$
$$\times f^{W_r}(t_1) f^{Y_{r-1}}(t_2) dt_1 dt_2 \tag{3.64}$$

yazılabilir. Yukarıdaki ifadenin,

$$\int_0^\infty\int_0^\infty P\{X_1,X_2,\dots,X_l \in (t_2,t_1+t_2), W_1,W_2,\dots,W_n \in [t_1,\infty)\}$$
$$\times f^{W_r}(t_1) f^{Y_{r-1}}(t_2) dt_1 dt_2 \tag{3.65}$$

ifadesi şeklinde de yazılabileceği açıktır. (3.65) ifadesindeki olasılık, X' lerin ve W' ların kendi içlerinde ve aralarında bağımsız oldukları bilindiğinden,

$$P\{X_1,X_2,\dots,X_l \in (t_2,t_1+t_2) \text{ ve } W_1,W_2,\dots,W_n \in [t_1,\infty)\}$$
$$= \left(\prod_{i=1}^{l} P\{X_i \in (t_2,t_1+t_2)\}\right) \times \left(\prod_{j=1}^{r-1} P\{W_j \in (t_1,\infty)\}\right)$$
$$\times \left(\prod_{j=r+1}^{n} P\{W_j \in (t_1,\infty)\}\right) \tag{3.66}$$

şeklinde yazılabilir. F, üstel dağılımın dağılım fonksiyonu olduğunda, $\{X_i \in (t_2,t_1+t_2)\}$ olayının olasılığı $e^{-\lambda t_2}(1-e^{-\lambda t_1})$ ve $\{W_j \in (t_1,\infty)\}$ olayının olasılığı ise $e^{-\lambda\gamma_j t_1}$ olacaktır. Bu bilgilere dayalı olarak, (3.66) eşitliğinin sağ tarafı yerine,

$$\left(e^{-\lambda t_2}\left(1-e^{-\lambda t_1}\right)\right)^l e^{-\lambda\sum_{j=1}^{r-1}\gamma_j t_1} e^{-\lambda\sum_{j=r+1}^{n}\gamma_j t_1} \tag{3.67}$$

yazılabilir. W_r ve Y_{r-1} rasgele değişkenlerinin olasılık yoğunluk fonksiyonları ile (3.67) ifadesi (3.64)' de yerine yazılacak ve bu ifade düzenlenecek olursa,

$$\lambda^2 \sum_{r=1}^{n} c_{r-1} \sum_{i=1}^{r-1} a_i(r-1) \int_0^\infty e^{-\lambda(\gamma_i+l)t_2} dt_2 \int_0^\infty \left(1-e^{-\lambda t_1}\right)^l e^{-\lambda \sum_{j=1}^{n} \gamma_j t_1} dt_1$$

elde edilir. t_1 ve t_2' ye göre integraller alındıktan sonra,

$$B\left(\sum_{j=1}^{n} \gamma_j, l+1\right) \sum_{i=1}^{r-1} \frac{c_{r-1} a_i(r-1)}{\gamma_i + l} \tag{3.68}$$

ifadesine ulaşılır. $\sum_{i=1}^{r-1} c_{r-1} a_i(r-1)/(\gamma_i + l)$, $(\gamma_r + l)\prod_{i=1}^{r} \gamma_i/(\gamma_i + l)$' e eşit olduğundan, (3.68) yerine,

$$B\left(\sum_{j=1}^{n} \gamma_j, l+1\right)(\gamma_r + l)\prod_{i=1}^{r} \frac{\gamma_i}{\gamma_i + l}$$

yazılabilir. Sonuç olarak, bu son ifade ve (3.62), (3.55)' de yerine konacak olursa, $\{X_1, X_2, \ldots, X_l \in (Y_{v-1}, Y_v)\}$ olayının olasılığı,

$$P\{X_1, X_2, \ldots, X_l \in (Y_{v-1}, Y_v)\}$$
$$= B\left(\sum_{j=1}^{n} \gamma_j, l+1\right)\left(\sum_{r=1}^{n} (\gamma_r + l)\prod_{i=1}^{r} \frac{\gamma_i}{\gamma_i + l}\right) \tag{3.69}$$

olarak bulunacaktır.

Yukarıdaki olasılığı da bulduktan sonra, $X(1, n, \tilde{\mathbf{m}}, k), \ldots, X(n, n, \tilde{\mathbf{m}}, k)$, $F(x) = 1 - e^{-\lambda x}$ $(x \geq 0, \lambda > 0)$ dağılım fonksiyonu üzerine kurulu genelleştirilmiş sıra istatistikleri, v minimal boşluğun indeksini gösteren rasgele değişken, $X(0,n,\tilde{\mathbf{m}},k)=0$ ve $1 \le l \le m$ olmak üzere, $F(x)$ dağılım fonksiyonuna sahip m adet bağımsız yeni gözlemin ilk l tanesinin minimal boşluğa düşmesi ve geriye kalan $m-l$ tanesinin düşmemesi olasılığını da bulalım: $j=1,2,..,n$ için $X(j,n,\tilde{\mathbf{m}},k)=Y_j$, $(Y_{j-1},Y_j)=I_{j-1,j}$, $i=1,2,...,m$ için $X_{n+i}=X_i$, $X(j,n,\tilde{\mathbf{m}},k)-X(j-1,n,\tilde{\mathbf{m}},k)=W_j$ ve $Y_0=X(0,n,\tilde{\mathbf{m}},k)=0$ olsun. Bu durumda, $\{X_1,...,X_l \in I_{v-1,v},X_{l+1},...,X_m \notin I_{v-1,v}\}$ olayının olasılığı toplamsal olarak,

$$
\begin{aligned}
P\{X_1,...,X_l &\in I_{v-1,v},X_{l+1},...,X_m \notin I_{v-1,v}\} \\
&= P\{X_1,...,X_l \in I_{0,1},X_{l+1},...,X_m \notin I_{0,1},v=1\} \\
&+ \sum_{r=2}^{n} P\{X_1,...,X_l \in I_{r-1,r},X_{l+1},...,X_m \notin I_{r-1,r},v=r\} \qquad (3.70)
\end{aligned}
$$

şeklinde yazılabilir. (3.70) eşitliğinin sağ tarafındaki, $P\{X_1,...,X_l \in I_{0,1},X_{l+1},...,X_m \notin I_{0,1},v=1\}$ olasılığı,

$$
\begin{aligned}
P\{X_1,...,X_l &\in I_{0,1},X_{l+1},...,X_m \notin I_{0,1}, \\
&Y_1 \le W_2,Y_1 \le W_3,...,Y_1 \le W_n\} \qquad (3.71)
\end{aligned}
$$

ifadesine eşittir. Bu son ifade yerine,

$$
\begin{aligned}
\int_0^{\infty} P\{X_1,...,X_l &\in I_{0,1},X_{l+1},...,X_m \notin I_{0,1}, \\
&Y_1 \le W_2,Y_1 \le W_3,...,Y_1 \le W_n | Y_1=t_1\} f^{Y_1}(t_1)dt_1 \qquad (3.72)
\end{aligned}
$$

yazılabilir. Bu noktada, yukarıdaki integralden hemen sonra gelen koşullu olasılık,

$$P\{X_1,\dots,X_l \in (0,t_1); X_{l+1},\dots,X_m \notin (0,t_1),\ W_2,W_3,\dots,W_n \in [t_1,\infty)\}$$

şeklinde yazılabildiğinden, (3.72) ifadesi,

$$\int_0^\infty P\{X_1,\dots,X_l \in (0,t_1), X_{l+1},\dots,X_m \in [t_1,\infty),$$

$$W_2,W_3,\dots,W_n \in [t_1,\infty)\} f^{\gamma_1}(t_1)dt_1 \qquad (3.73)$$

şeklinde gösterilebilir. (3.73) ifadesi içerisinde yer alan, m adet yeni gözlemden ilk l tanesinin $(0,t_1)$ aralığına düşmesi, geriye kalan $(m-l)$ tanesinin ise $[t_1,\infty)$ aralığına düşmesi ve her W_j ' nin t_1 ' den büyük veya eşit olması olayı,

$$\{X_1,X_2,\dots,X_l \in (0,t_1), X_{l+1},X_{l+2},\dots,X_m \in [t_1,\infty),$$

$$W_2,W_3,\dots,W_n \in [t_1,\infty)\} \qquad (3.74)$$

şeklinde de ifade edilebilir. Öyleyse, (3.74) olayının olasılığı,

$$P\{X_1,X_2,\dots,X_l \in (0,t_1), X_{l+1},X_{l+2},\dots,X_m \in [t_1,\infty),$$
$$W_2,W_3,\dots,W_n \in [t_1,\infty)\}$$

$$= \left(1-e^{-\lambda t_1}\right)^l \left(e^{-\lambda t_1}\right)^{m-l+\sum_{i=2}^{n}\gamma_i} \qquad (3.75)$$

olacaktır. Bu durumda (3.73) ifadesi yerine,

$$\lambda\gamma_1 \int_0^\infty \left(1-e^{-\lambda t_1}\right)^l \left(e^{-\lambda t_1}\right)^{m-l+\sum_{i=1}^{n}\gamma_i} dt_1 \qquad (3.76)$$

yazılacaktır. Yukarıdaki ifadedeki integral alınacak olursa,

$$P\{X_1,\dots,X_l \in I_{0,1}, X_{l+1},\dots,X_m \notin I_{0,1}, v=1\}$$

$$= \gamma_1 B\left(m-l+\sum_{i=1}^{n}\gamma_i, l+1\right) \qquad (3.77)$$

elde edilir. Diğer yandan, (3.70) eşitliğinin sağ tarafında yer alan toplam olasılık,

$$\sum_{r=2}^{n} P\{X_1,\ldots,X_l \in I_{r-1,r}, X_{l+1},\ldots,X_m \notin I_{r-1,r},$$
$$W_r \le W_1, W_r \le W_2,\ldots,W_r \le W_n\} \tag{3.78}$$

şeklinde yazılabilir. Bu son ifade yerine, $W_r = t_1$, $Y_{r-1} = t_2$ olmak üzere ve daha önce verdiğimiz teorem-4' e göre $f^{W_r,Y_{r-1}}(t_1,t_2) = f^{W_r}(t_1) f^{Y_{r-1}}(t_2)$ olduğundan,

$$\sum_{r=2}^{n} \int_0^\infty \int_0^\infty P\{X_1,\ldots,X_l \in I_{r-1,r}, X_{l+1},\ldots,X_m \notin I_{r-1,r},$$
$$W_r \le W_1, W_r \le W_2,\ldots,W_r \le W_n | W_r = t_1, Y_{r-1} = t_2\}$$
$$\times f^{W_r}(t_1) f^{Y_{r-1}}(t_2) dt_1 dt_2 \tag{3.79}$$

yazılabilir. Bu noktada, yukarıdaki integralden hemen sonra gelen koşullu olasılık,

$$P\{X_1,\ldots,X_l \in (t_2, t_1+t_2); X_{l+1},\ldots,X_m \notin (t_2, t_1+t_2),$$
$$W_1, W_2,\ldots,W_{r-1}, W_{r+1}, W_{r+2},\ldots,W_n \in [t_1,\infty)\}$$

şeklinde yazılabildiğinden, (3.79) ifadesi,

$$\sum_{r=2}^{n} \int_0^\infty \int_0^\infty P\{X_1,\ldots,X_l \in (t_2, t_1+t_2), X_{l+1},\ldots,X_m \notin (t_2, t_1+t_2),$$
$$W_1, W_2,\ldots,W_{r-1}, W_{r+1}, W_{r+2},\ldots,W_n \in [t_1,\infty)\}$$
$$\times f^{W_r}(t_1) f^{Y_{r-1}}(t_2) dt_1 dt_2 \tag{3.80}$$

şeklinde gösterilebilir. (3.80) ifadesi içerisinde yer alan, m adet yeni gözlemden ilk l tanesinin (t_2, t_1+t_2) aralığına düşmesi, geriye kalan

$(m-l)$ tanesinin ise bu aralığa düşmemesi ve her W_j ' nin t_1 ' den büyük olması olayı,

$$\begin{aligned}\{&X_1, X_2, \ldots, X_l \in (t_2, t_1+t_2)\, ve \\ &(X_{l+1} \in (0, t_2] \text{ veya } X_{l+1} \in [t_1+t_2, \infty))\, ve \\ &(X_{l+2} \in (0, t_2] \text{ veya } X_{l+2} \in [t_1+t_2, \infty))\, ve \\ &\ldots\, ve\, (X_m \in (0, t_2] \text{ veya } X_m \in [t_1+t_2, \infty))\, ve \\ &W_1, W_2, \ldots, W_{r-1}, W_{r+1}, W_{r+2}, \ldots, W_n \in [t_1, \infty)\} \end{aligned} \tag{3.81}$$

şeklinde de ifade edilebilir. Öyleyse, (3.81) olayının olasılığı,

$$\begin{aligned}\sum_{t=0}^{m-l} \binom{m-l}{t} \times P\{&X_1, X_2, \ldots, X_l \in (t_2, t_1+t_2)\, ve \\ &t \text{ adet } X_i \in (0, t_2]\, ve \\ &(m-l-t) \text{adet } X_i \in [t_1+t_2, \infty)\, ve \\ &W_1, W_2, \ldots, W_{r-1}, W_{r+1}, W_{r+2}, \ldots, W_n \in [t_1, \infty)\} \end{aligned} \tag{3.82}$$

olacaktır. (3.82) ifadesinde yer alan tüm rasgele değişkenler birbirlerinden bağımsız olduklarından, bu ifade yerine,

$$\begin{aligned}&\sum_{t=0}^{m-l} \binom{m-l}{t} \times (P\{X \in (t_2, t_1+t_2)\})^l \times (P\{X \in (0, t_2]\})^t \\ &\times (P\{X \in [t_1+t_2, \infty)\})^{m-l-t} \times \left(\prod_{j=1}^{r-1} P\{W_j \in [t_1, \infty)\}\right) \\ &\times \left(\prod_{j=r+1}^{n} P\{W_j \in [t_1, \infty)\}\right) \end{aligned} \tag{3.83}$$

yazılabilir. Üstel dağılımın dağılım fonksiyonu F ile gösterdiğine göre,

$P\{X \in (t_2, t_1+t_2)\} = e^{-\lambda t_2}(1-e^{-\lambda t_1})$, $P\{X \in (0, t_2]\} = 1-e^{-\lambda t_2}$,

$P\{X \in [t_1 + t_2, \infty)\} = e^{-\lambda(t_1+t_2)}$ ve $P\{W_j \in [t_1, \infty)\} = e^{-\lambda \gamma_j t_1}$ olacaktır. Bu durumda (3.83) ifadesi yerine,

$$\sum_{t=0}^{m-l} \binom{m-l}{t} \times \left(e^{-\lambda t_2}\left(1 - e^{-\lambda t_1}\right)\right)^l \times \left(1 - e^{-\lambda t_2}\right)^t$$

$$\times \left(e^{-\lambda(t_1+t_2)}\right)^{m-l-t} \times \left(\prod_{j=1}^{r-1} e^{-\lambda \gamma_j t_1}\right) \times \left(\prod_{j=r+1}^{n} e^{-\lambda \gamma_j t_1}\right) \tag{3.84}$$

yazılır ve bu ifade düzenlenecek olursa,

$$\sum_{t=0}^{m-l} \binom{m-l}{t} \times \left(1 - e^{-\lambda t_1}\right)^l \left(1 - e^{-\lambda t_2}\right)^t e^{-\lambda\left(m-l-t+\sum_{j=1}^{r-1}\gamma_j + \sum_{j=r+1}^{n}\gamma_j\right)t_1} e^{-\lambda(m-t)t_2} \tag{3.85}$$

elde edilir. Ayrıca, $f^{W_r}(t_1) = \gamma_r \lambda\, e^{-\lambda \gamma_r t_1}$ ve $f^{Y_{r-1}}(t_2) = c_{r-2}\lambda \sum_{i=1}^{r-1} a_i(r-1) e^{-\lambda \gamma_i t_2}$ olduğundan, bu eşitliklerin sağ tarafları ve (3.85) ifadesi, (3.80) ifadesinde yerlerine konacak olursa,

$$\sum_{r=2}^{n} \int_0^\infty \int_0^\infty \sum_{t=0}^{m-l} \binom{m-l}{t} \times \left(1 - e^{-\lambda t_1}\right)^l \left(1 - e^{-\lambda t_2}\right)^t$$

$$\times e^{-\lambda\left(m-l-t+\sum_{j=1}^{r-1}\gamma_j + \sum_{j=r+1}^{n}\gamma_j\right)t_1} e^{-\lambda(m-t)t_2}$$

$$\times \gamma_r \lambda\, e^{-\lambda \gamma_r t_1} c_{r-2}\lambda \sum_{i=1}^{r-1} a_i(r-1) e^{-\lambda \gamma_i t_2}\, dt_1 dt_2 \tag{3.86}$$

elde edilir. Bu ifade düzenlenecek olursa,

$$\sum_{r=2}^{n}\sum_{t=0}^{m-l}\sum_{i=1}^{r-1}\binom{m-l}{t}\lambda^2 c_{r-1}a_i(r-1)$$
$$\times\left(\int_0^{\infty}\left(1-e^{-\lambda t_1}\right)^l e^{-\lambda\left(m-l-t+\sum_{j=1}^{n}\gamma_j\right)t_1}dt_1\right)$$
$$\times\left(\int_0^{\infty}\left(1-e^{-\lambda t_2}\right)^t e^{-\lambda(m-t+\gamma_i)t_2}dt_2\right) \tag{3.87}$$

ifadesine ulaşılır. Bu son ifadede, t_1'e göre integralin sonucunun $B\left(m-l-t+\sum_{j=1}^{n}\gamma_j, l+1\right)\Big/\lambda$ ve t_2'ye göre integralin sonucunun ise $B(m-t+\gamma_i, t+1)/\lambda$ olduğu için, (3.87) ifadesi,

$$\sum_{r=2}^{n}\sum_{t=0}^{m-l}\sum_{i=1}^{r-1}\binom{m-l}{t}\lambda^2 c_{r-1}a_i(r-1)$$
$$\times\frac{B\left(m-l-t+\sum_{j=1}^{n}\gamma_j, l+1\right)}{\lambda}\times\frac{B(m-t+\gamma_i, t+1)}{\lambda} \tag{3.88}$$

halini alacaktır. Bu son ifade düzenlenecek olursa,

$$\sum_{r=2}^{n}P\{X_1,\ldots,X_l\in I_{r-1,r}, X_{l+1},\ldots,X_m\notin I_{r-1,r}, v=r\}$$
$$=\sum_{t=0}^{m-l}\binom{m-l}{t}B\left(m-l-t+\sum_{j=1}^{n}\gamma_j, l+1\right)$$
$$\times\sum_{r=2}^{n}\sum_{i=1}^{r-1}c_{r-1}a_i(r-1)B(m-t+\gamma_i, t+1) \tag{3.89}$$

eşitliğine ulaşılır. Sonuç olarak, (3.77) ve (3.89), (3.70)' de yerine koyulursa,

$$P\{X_1,\ldots,X_l \in I_{v-1,v}, X_{l+1},\ldots,X_m \notin I_{v-1,v}\}$$

$$= \gamma_1 B\left(m-l+\sum_{i=1}^{n}\gamma_i, l+1\right) + \sum_{t=0}^{m-l}\binom{m-l}{t} B\left(m-l-t+\sum_{j=1}^{n}\gamma_j, l+1\right)$$

$$\times \sum_{r=2}^{n}\sum_{i=1}^{r-1} c_{r-1} a_i (r-1) B(m-t+\gamma_i, t+1) \qquad (3.90)$$

elde edilecektir.

Yukarıdaki olasılıkları bulduktan sonra, artık aşağıdaki teoremi verebiliriz:

Teorem 4. $X(1,n,\tilde{\mathbf{m}},k),\ldots,X(n,n,\tilde{\mathbf{m}},k)$, $F(x)=1-e^{-\lambda x}$ $(x \geq 0, \lambda > 0)$ dağılım fonksiyonu üzerine kurulu genelleştirilmiş sıra istatistikleri, v minimal boşluğun indeksini gösteren rasgele değişken, $X(0,n,\tilde{\mathbf{m}},k)=0$ ve $1 \leq l \leq m$ olmak üzere,

$$\xi_i = \begin{cases} 1 & , \ X_{n+i} \in \left(X(v-1,n,\tilde{\mathbf{m}},k), X(v,n,\tilde{\mathbf{m}},k)\right) \\ 0 & , \ X_{n+i} \notin \left(X(v-1,n,\tilde{\mathbf{m}},k), X(v,n,\tilde{\mathbf{m}},k)\right) \end{cases}$$

şeklinde tanımlansın. $S_m = \sum_{i=1}^{m} \xi_i$ olarak tanımlanan ve $F(x)$ dağılım fonksiyonuna sahip birbirlerinden bağımsız $X_{n+1}, X_{n+2},\ldots,X_{n+m}$ gözlemlerinden minimal boşluğa düşenlerin sayısını gösteren bu istatistiğin değerinin l' ye eşit olması olasılığı, yani S_m rasgele değişkeninin olasılık fonksiyonu,

$$P\{S_m = l\} = \binom{m}{l}\gamma_1 B\left(m - l + \sum_{i=1}^{n}\gamma_i, l+1\right)$$
$$+\binom{m}{l}\sum_{t=0}^{m-l}\binom{m-l}{t}B\left(m-l-t+\sum_{j=1}^{n}\gamma_j, l+1\right)$$
$$\times\sum_{r=2}^{n}\sum_{i=1}^{r-1}c_{r-1}a_i(r-1)B(m-t+\gamma_i, t+1)$$

dir.

İspat: $\{S_m = l\}$ olayı, her $j \neq t$ için $i_j \neq i_t$ olmak üzere,

$$\{X_{n+i_1},\ldots,X_{n+i_l} \in (X(v-1,n,\tilde{\mathbf{m}},k), X(v,n,\tilde{\mathbf{m}},k)),$$
$$X_{n+i_{l+1}},\ldots,X_{n+i_m} \notin (X(v-1,n,\tilde{\mathbf{m}},k), X(v,n,\tilde{\mathbf{m}},k))\}$$

şeklinde de ifade edilebilir. Bu durumda,

$$P\{S_m = l\} =_{i_1,\ldots,i_m} P\{X_{n+i_1},\ldots,X_{n+i_l} \in (X(v-1,n,\tilde{\mathbf{m}},k), X(v,n,\tilde{\mathbf{m}},k)),$$
$$X_{n+i_{l+1}},\ldots,X_{n+i_m} \notin (X(v-1,n,\tilde{\mathbf{m}},k), X(v,n,\tilde{\mathbf{m}},k))\} \qquad (3.91)$$

56

yazılabilir. Minimal aralığa, m tane rasgele değişkenden l tanesinin, $C_m^l = \binom{m}{l}$ farklı şekilde düşeceği açıktır. Ayrıca, $\binom{m}{l}$ farklı durumun her birinin olasılığı (3.90)' da verilen olasılıkla aynıdır. Bu durumda, (3.91) yerine,

$$P\{S_m = l\} = \binom{m}{l} \times P\{X_{n+1},\ldots,X_{n+l} \in (X(v-1,n,\tilde{\mathbf{m}},k), X(v,n,\tilde{\mathbf{m}},k)),$$

$$X_{n+l+1},\ldots,X_{n+m} \notin (X(v-1,n,\tilde{\mathbf{m}},k), X(v,n,\tilde{\mathbf{m}},k))\}$$

$$= \binom{m}{l}\gamma_1 B\left(m-l+\sum_{i=1}^{n}\gamma_i, l+1\right)$$

$$+\binom{m}{l}\sum_{t=0}^{m-l}\binom{m-l}{t}B\left(m-l-t+\sum_{j=1}^{n}\gamma_j, l+1\right)$$

$$\times\sum_{r=2}^{n}\sum_{i=1}^{r-1}c_{r-1}a_i(r-1)B(m-t+\gamma_i, t+1)$$

yazılabilir ve teorem ispatlanmış olur.□

3.4 Sonuçlar

Yukarıda verilen teoremlere ilişkin olarak elde edilen bazı sonuçlar aşağıdaki gibidir:

(i) Eğer $k=1$ ve $\tilde{\mathbf{m}} = (m_1, m_2, \ldots, m_{n-1}) = (0,0,\ldots,0)$ seçilirse, $X(r,n,0,1)$ artık r. sıra istatistiği olacaktır. Bu durumda, $m = l = 1$ için Teorem-1, (3.6) haline gelecektir. (3.6)' dan hareketle X_{n+1} ' in r ' inci ve s ' inci sıra istatistikleri arasına düşme olasılığının (2.18)' deki gibi $\frac{s-r}{n+1}$ olduğu aşağıdaki gibi kolayca gösterilebilir: $c_{s-1} = \frac{n!}{(n-s)!}$,

$\gamma_i = n+1-i$, $a_i(r) = \frac{(-1)^{r-i}}{(r-i)!(i-1)!}$, $a_j^{(r)}(s) = \frac{(-1)^{s-j}}{(s-j)!(j-r-1)!}$ ve

$B(n+1-i,2) = \frac{1}{(n-i+2)(n-i+1)}$ olacaktır. Ayrıca,

$$b_1 = \sum_{j=1}^{r} \frac{(-1)^{r-j}}{(r-j)!(j-1)!(n-j+2)}$$

ve

$$b_2 = \sum_{i=r+1}^{s} \frac{(-1)^{s-i}}{(s-i)!(i-r-1)!(n-i+2)(n-i+1)}$$

olarak tanımlanırsa, Teorem-1' e göre,

$$P\{X_{n+1} \in (X(r,n,0,1), X(s,n,0,1))\} = \frac{n!}{(n-s)!} \times b_1 \times b_2 \qquad (3.92)$$

eşitliği yazılabilir. (3.92) eşitliğindeki b_1 ifadesi,

$$\begin{aligned}
b_1 &= \sum_{j=1}^{r} \frac{(-1)^{r-j}}{(r-j)!(j-1)!(n-j+2)} \\
&= \frac{(-1)^{r-1}}{(r-1)!(0)!(n+1)} + \frac{(-1)^{r-2}}{(r-2)!(1)!(n)} + \dots + \frac{(-1)^{0}}{(0)!(r-1)!(n-r+2)} \\
&= \frac{(-1)^{r-1}}{(r-1)!}\left[\binom{r-1}{0}\frac{1}{n+1} - \binom{r-1}{1}\frac{1}{n} + \binom{r-1}{2}\frac{1}{n-1} - \dots + \binom{r-1}{r-1}\frac{(-1)^{1-r}}{n-r+2}\right] \\
&= \frac{(-1)^{r-1}}{(r-1)!}\left[\frac{\dfrac{(1-r)(-1)^{r}(r-2)!}{(n+1)!}}{(n-r+1)!}\right] \\
&= \frac{(-1)^{2r-1}(1-r)(r-2)!(n-r+1)!}{(r-1)!(n+1)!} \\
&= \frac{(n-r+1)!}{(n+1)!}
\end{aligned}$$

şeklinde yazılabilir. Benzer şekilde, $b_2 = \dfrac{(n-s)!(s-r)}{(n-r+1)!}$ olarak bulunacaktır. Sonuç olarak, b_1 ve b_2 (3.92)' de yerine koyulursa,

$$\begin{aligned} P\{X_{n+1} \in (X(r,n,0,1), X(s,n,0,1))\} \\ &= \frac{n!}{(n-s)!} \times b_1 \times b_2 \\ &= \frac{n!}{(n-s)!} \times \frac{(n-r+1)!}{(n+1)!} \times \frac{(n-s)!(s-r)}{(n-r+1)!} \\ &= \frac{s-r}{n+1} \end{aligned}$$

elde edilecektir.

(ii) Teorem-2' de, $s = r+1$ seçilmesi durumunda, Ahsanullah (2000)' ın, $X(r-1,n,\tilde{\mathbf{m}},k)$ ile $X(r,n,\tilde{\mathbf{m}},k) - X(r-1,n,\tilde{\mathbf{m}},k)$ rasgele değişkenlerinin birbirlerinden bağımsız olduğununa ilişkin teoremine ulaşılır. Yani, Teorem-2, Ahsanullah' ın bu teoreminin genelleştirilmiş halidir.

(iii) Teorem-3'den hareketle, $k=1$ ve $\tilde{\mathbf{m}} = (0,0,...,0)$ seçilirse, (2.22) eşitliği,

$$P\{v=r\} = \frac{n-r+1}{\sum_{i=1}^{n} n-i+1}$$

$$= \frac{n-r+1}{n+(n-1)+(n-2)+...+(n-(n-1))}$$

$$=\frac{n-r+1}{n^2-\frac{n(n-1)}{2}}$$
$$=\frac{2(n-r+1)}{n(n+1)}$$

şeklinde elde edilebilmektedir.

(iv) $m=l=1$ için Teorem-4, (3.54) haline gelecektir ve (3.54)' de $k=1$, $\widetilde{\mathbf{m}}=(0,0,...,0)$ seçilirse, (2.23) eşitliği,

$$P\{X_{n+1}\in(X(v-1,n,0,1),X(v,n,0,1))\}$$

$$=\frac{\sum_{r=1}^{n}(n+2-r)\prod_{i=1}^{r}\frac{n+1-i}{n+2-i}}{\left(\sum_{j=1}^{n}(n+1-j)\right)\left(1+\sum_{j=1}^{n}(n+1-j)\right)}$$

$$=\frac{\sum_{r=1}^{n}\frac{n(n-1)...(n-r+1)}{(n+1)n(n-1)...(n-r+3)}}{\left(\frac{n(n+1)}{2}\right)\left(1+\frac{n(n+1)}{2}\right)}$$

$$=\frac{4}{n(n+1)^2(n^2+n+2)}\times\sum_{r=1}^{n}(n-r+2)(n-r+1)$$

$$=\frac{4}{n(n+1)^2(n^2+n+2)}\times[(n+1)n+n(n-1)+(n-1)(n-2)$$
$$+...+(n-(n-2))(n-(n-1))]$$

$$=\frac{4}{n(n+1)^2(n^2+n+2)}\times[2n^2+(n-1)(n-2)$$
$$+...+(n-(n-2))(n-(n-1))]$$

$$= \frac{4}{n(n+1)^2(n^2+n+2)} \times \left(2n^2 + \frac{n^3 - 3n^2 + 2n}{3} \right)$$

$$= \frac{4}{n(n+1)^2(n^2+n+2)} \times \frac{n(n+2)(n+1)}{3}$$

$$= \frac{4}{3} \frac{(n+2)}{(n+1)(n^2+n+2)}$$

şeklinde elde edilmektedir.

4 GENELLEŞTİRİLMİŞ SIRA İSTATİSTİKLERİNE DAYALI HİPOTEZ TESTLERİ

Genelleştirilmiş sıra istatistikleri, sıra istatistiklerine ilişkin literatürde var olan birçok modeli içerdiği için, bu istatistiklere dayalı olarak geliştirilen herhangi bir yöntem hepsi için geçerli olacaktır.

Aşağıdaki problemi gözönüne alalım. Varsayalım ki, $X_1, X_2, ..., X_n$ ve $Y_1, Y_2, ..., Y_n$, sırasıyla, sürekli F_x ve F_y dağılımlarına sahip olan iki örneklem olsun. Genel olarak, F_x ve F_y' nin bilinmediği varsayılmaktadır. Bu örneklemlere eğitici (training) örneklemler de denilmektedir. Örneğin, $(X_1, X_2, ..., X_n)$ vektörü, kesin olarak şekerli diyabet hastalığına yakalanmış n tane hastanın kan şekeri miktarlarını gösteren vektör, $(Y_1, Y_2, ..., Y_n)$ ise şekersiz diyabete kesin olarak yakalanmış n tane hastaya ilişkin kan şekeri miktarlarını gösteren vektör olarak tanımlansın. Yeni bir olası diyabet hastasından kan örneği alınsın ve kan şekeri miktarı ölçülsün ve bu yeni ölçüm "Z" ile, Z' nin dağılım fonksiyonu ise F_z ile gösterilsin. Bu durumda yeni gelen olası bir diyabet hastasının kan şekeri miktarına bakarak şekerli diyabet mi yoksa şekersiz diyabet mi olduğuna karar vermek için $H_0 : F_z = F_x$ hipotezini $H_1 : F_z = F_y$ hipotezine karşı test etmemiz gerekmektedir. Bu karar problemini çözmek için Bairamov ve Petunin (1991b), bir test prosedürü önermişlerdir. Bu test prosedürü, genelleştirilmiş sıra istatistiklerine aşağıdaki şekilde uyarlanabilir.

4.1 Genelleştirilmiş Sıra İstatistiklerine Dayalı Test Prosedürü

Bairamov ve Petunin' in (1991b) önerdikleri test prosedürüne uygun olarak, genelleştirilmiş sıra istatistiklerine dayalı test prosedürü iki farklı şekilde ele alınabilir: (i) Parametreleri aynı olan iki eğitici örneklem için test prosedürü, (ii) Parametreleri farklı olan iki eğitici örneklem için test prosedürü. Bu iki durum için önerilen test prosedürleri bu bölümde verilecektir.

4.1.1 Parametreleri aynı olan iki eğitici örneklem için test prosedürü

$X(1,n,\tilde{\mathbf{m}},k), X(2,n,\tilde{\mathbf{m}},k),\ldots,X(n,n,\tilde{\mathbf{m}},k)$ ve $Y(1,n,\tilde{\mathbf{m}},k)$, $Y(2,n,\tilde{\mathbf{m}},k),\ldots,Y(n,n,\tilde{\mathbf{m}},k)$ genelleştirilmiş sıra istatistikleri, sırasıyla, sürekli F_x ve F_y dağılımları üzerine kurulu oldukları varsayılsın.Ayrıca,

$$X(1,n,\tilde{\mathbf{m}},k) \leq Y(1,n,\tilde{\mathbf{m}},k) \leq X(n,n,\tilde{\mathbf{m}},k) \leq Y(n,n,\tilde{\mathbf{m}},k)$$

koşulu sağlanıyor olsun. Yeni bir gözlem olan Z, F_z dağılım fonksiyonuna sahip olsun. $H_0 : F_z = F_x$ hipotezini $H_1 : F_z = F_y$ hipotezine karşı test ederken kullanılacak karar kriterleri,

Eğer $Z \in (-\infty, Y(1,n,\tilde{\mathbf{m}},k))$ ***ise,*** H_0 ***reddedilemez.***

Eğer $Z \in (X(n,n,\tilde{\mathbf{m}},k), \infty)$ ***ise,*** H_0 ***reddedilir.***

Eğer $Z \in (Y(1,n,\tilde{\mathbf{m}},k), X(n,n,\tilde{\mathbf{m}},k))$ ***ise, karar verilemez***

şeklindedir. Bu test prosedürü için I. tip hata $\alpha = P\{H_1|H_0\}$ şeklinde tanımlanır ve

$$\begin{aligned}
\alpha &= P\{H_1|H_0\} \\
&= P\{Z \in (X(n,n,m,k),\infty)|H_0\} \\
&= P\{X_{n+1} \in (X(n,n,m,k),\infty)\} \\
&= 1 - P\{X_{n+1} \leq X(n,n,m,k)\} \qquad (4.1)
\end{aligned}$$

yazılabilir. (4.1) eşitliğinin sağ tarafındaki olasılık,

$$\begin{aligned}
P\{X_{n+1} \leq X(n,n,m,k)\} &= \int_{-\infty}^{\infty}\int_{x}^{\infty} f_x(x) f^{X(n,n,m,k)}(y)\,dy\,dx \\
&= \int_{-\infty}^{\infty} f_x(x) \int_{x}^{\infty} f^{X(n,n,m,k)}(y)\,dy\,dx \\
&= \int_{-\infty}^{\infty} f_x(x)\left(1 - F^{X(n,n,m,k)}(x)\right)dx \qquad (4.2)
\end{aligned}$$

şeklinde ifade edilebilir. Bu eşitliğin sağ tarafında yer alan $F^{X(n,n,m,k)}(x)$ yerine, (2.33)' deki karşılığı yazılacak olursa, (4.2) eşitliğinin sağ tarafı,

$$c_{n-1}\sum_{i=1}^{n}\frac{a_i(n)}{\gamma_i}\int_{-\infty}^{\infty}(1 - F_x(x))^{\gamma_i}\,dF_x(x) \qquad (4.3)$$

olacaktır. Bu son ifadede yer alan integral alınır ve (4.3) ifadesi, (4.1)' de yerine koyulursa,

$$\begin{aligned}
\alpha = P\{H_1|H_0\} &= 1 - c_{n-1}\sum_{i=1}^{n}\frac{a_i(n)}{\gamma_i(\gamma_i+1)} \\
&= c_{n-1}\sum_{i=1}^{n}\frac{a_i(n)}{\gamma_i} - c_{n-1}\sum_{i=1}^{n}\frac{a_i(n)}{\gamma_i(\gamma_i+1)}
\end{aligned}$$

$$= c_{n-1} \sum_{i=1}^{n} \frac{a_i(n)}{\gamma_i + 1} \tag{4.4}$$

şeklinde I. tip hata elde edilmiş olur. Benzer şekilde, II. tip hata,

$$\begin{aligned}
\beta &= P\{H_0 | H_1\} \\
&= P\{Z \in (-\infty, Y(1, n, \tilde{\mathbf{m}}, k)) | H_1\} \\
&= P\{Y_{n+1} \in (-\infty, Y(1, n, \tilde{\mathbf{m}}, k))\} \\
&= P\{Y_{n+1} < Y(1, n, \tilde{\mathbf{m}}, k)\} \\
&= \frac{1}{\gamma_1 + 1}
\end{aligned} \tag{4.5}$$

olacaktır. β , n ' in bir fonksiyonu olduğundan, $\beta = \beta\ (n)$ şeklinde ifade edilebilir. Bu durumda, $n \to \infty$ için $\beta\ (n) \to 0$ olacağı açıktır:

$$\begin{aligned}
\lim_{n\to\infty} \beta\ (n) &= \lim_{n\to\infty} \frac{1}{k + n - 1 + \sum_{i=1}^{n} m_i} \\
&= \frac{1}{k + \infty - 1 + \sum_{i=1}^{\infty} m_i} \\
&= \frac{1}{\infty} = 0\ .
\end{aligned}$$

4.1.2 Parametreleri farklı olan iki eğitici örneklem için test prosedürü

$n^{(\cdot)} \in \{2,3,\ldots\}$, $i = 1,2,\ldots,(n^{(\cdot)} - 1)$, $m_i^{(\cdot)} \in R$, $\tilde{\mathbf{m}}^{(\cdot)} = (m_1^{(\cdot)},\ldots,m_{n_x-1}^{(\cdot)})$,

$M_i^{(\cdot)} = \sum_{j=i}^{n^{(\cdot)}-1} m_j^{(\cdot)}$, $\gamma_i^{(\cdot)} = k^{(\cdot)} + n^{(\cdot)} - i + M_i^{(\cdot)} > 0$, $\gamma_{n^{(\cdot)}}^{(\cdot)} = k^{(\cdot)}$, $k^{(\cdot)} > 0$ olmak

üzere,

$$X\left(1,n^{(x)},\widetilde{\mathbf{m}}^{(x)},k^{(x)}\right),X\left(2,n^{(x)},\widetilde{\mathbf{m}}^{(x)},k^{(x)}\right),...,X\left(n^{(x)},n^{(x)},\widetilde{\mathbf{m}}^{(x)},k^{(x)}\right)$$

ve

$$Y\left(1,n^{(y)},\widetilde{\mathbf{m}}^{(y)},k^{(y)}\right),Y\left(2,n^{(y)},\widetilde{\mathbf{m}}^{(y)},k^{(y)}\right),...,Y\left(n^{(y)},n^{(y)},\widetilde{\mathbf{m}}^{(y)},k^{(y)}\right)$$

genelleştirilmiş sıra istatistikleri, sırasıyla, sürekli F_x ve F_y dağılımları üzerine kurulu olsunlar. Ayrıca,

$$X\left(1,n^{(x)},\widetilde{\mathbf{m}}^{(x)},k^{(x)}\right)\leq Y\left(1,n^{(y)},\widetilde{\mathbf{m}}^{(y)},k^{(y)}\right)\leq X\left(n^{(x)},n^{(x)},\widetilde{\mathbf{m}}^{(x)},k^{(x)}\right)$$

ve

$$X\left(n^{(x)},n^{(x)},\widetilde{\mathbf{m}}^{(x)},k^{(x)}\right)\leq Y\left(n^{(y)},n^{(y)},\widetilde{\mathbf{m}}^{(y)},k^{(y)}\right)$$

koşulu sağlanıyor olsun. Yeni bir gözlem olan Z, F_z dağılım fonksiyonuna sahip olsun. $H_0 : F_z = F_x$ hipotezini $H_1 : F_z = F_y$ hipotezine karşı test ederken kullanılacak karar kriterleri,

Eğer $Z \in \left(-\infty, Y\left(1,n^{(y)},\widetilde{\mathbf{m}}^{(y)},k^{(y)}\right)\right)$ ***ise,*** H_0 ***reddedilemez.***

Eğer $Z \in \left(X\left(n^{(x)},n^{(x)},\widetilde{\mathbf{m}}^{(x)},k^{(x)}\right),\infty\right)$ ***ise,*** H_0 ***reddedilir.***

Eğer $Z \in \left(Y\left(1,n^{(y)},\widetilde{\mathbf{m}}^{(y)},k^{(y)}\right),X\left(n^{(x)},n^{(x)},\widetilde{\mathbf{m}}^{(x)},k^{(x)}\right)\right)$ ***ise, karar verilemez***

şeklinde olacaktır. Bu durumda, $a_i^{(.)}\left(n^{(.)}\right)=\prod\limits_{\substack{j=1\\j\neq i}}^{n^{(.)}}\left(1/\left(\gamma_j^{(.)}-\gamma_i^{(.)}\right)\right)$ ve

$c_{n^{(x)}-1}^{(.)}=\gamma_1^{(.)}...\gamma_{n^{(x)}}^{(.)}$ olmak üzere,

$$
\begin{aligned}
\alpha &= P\{H_1 | H_0\} \\
&= P\left\{Z \in \left(X\left(n^{(x)}, n^{(x)}, \tilde{\mathbf{m}}^{(x)}, k^{(x)}\right), \infty\right) \middle| H_0\right\} \\
&= P\left\{X_{n+1} \in \left(X\left(n^{(x)}, n^{(x)}, \tilde{\mathbf{m}}^{(x)}, k^{(x)}\right), \infty\right)\right\} \\
&= 1 - P\left\{X_{n+1} \leq X\left(n^{(x)}, n^{(x)}, \tilde{\mathbf{m}}^{(x)}, k^{(x)}\right)\right\} \\
&= 1 - c^{(x)}_{n^{(x)}-1} \sum_{i=1}^{n_x} \frac{a_i^{(x)}\left(n^{(x)}\right)}{\gamma_i^{(x)}\left(\gamma_i^{(x)}+1\right)} \\
&= c^{(x)}_{n^{(x)}-1} \sum_{i=1}^{n_x} \frac{a_i^{(x)}\left(n^{(x)}\right)}{\left(\gamma_i^{(x)}+1\right)}
\end{aligned}
\tag{4.6}
$$

ve

$$
\begin{aligned}
\beta &= P\{H_0 | H_1\} \\
&= P\left\{Z \in \left(-\infty, Y\left(1, n^{(y)}, \tilde{\mathbf{m}}^{(y)}, k^{(y)}\right)\right) \middle| H_1\right\} \\
&= P\left\{Y_{n+1} \in \left(-\infty, Y\left(1, n^{(y)}, \tilde{\mathbf{m}}^{(y)}, k^{(y)}\right)\right)\right\} \\
&= P\left\{Y_{n+1} < Y\left(1, n^{(y)}, \tilde{\mathbf{m}}^{(y)}, k^{(y)}\right)\right\} \\
&= \frac{1}{\gamma_1^{(y)}+1}
\end{aligned}
\tag{4.7}
$$

şeklinde elde edilir.

4.2 II. Tür İlerleyen Sansürlenmiş Sıra İstatistiklerine Dayalı Test Prosedürü

Son yıllarda istatistik literatüründe II. Tür ilerleyen sansürlenmiş (progressive Type-II Censoring) verilere dayalı istatistiksel sonuç çıkarımına ilişkin araştırmalara sık sık raslanmaktadır (Balakrishnan ve diğ., 2002; Balakrishnan ve diğ., 2001; Aggarwala ve diğ., 1998). Örneğin, birçok uygulamada örneklemin tamamına ulaşılması mümkün olmadığından, sadece sansürlenmiş (censored) verilere ulaşılabilir. İstatistiksel sonuç çıkarmayı bu kısıtlı sayıdaki verilere dayalı olarak yapmamız gerekmektedir.

4.2.1 Örneklem büyüklükleri ve sansürleme planları aynı olan iki eğitici örneklem için test prosedürü

$X^{\tilde{\mathbf{R}}}_{1:n:N}, X^{\tilde{\mathbf{R}}}_{2:n:N}, \ldots, X^{\tilde{\mathbf{R}}}_{n:n:N}$ ve $Y^{\tilde{\mathbf{R}}}_{1:n:N}, Y^{\tilde{\mathbf{R}}}_{2:n:N}, \ldots, Y^{\tilde{\mathbf{R}}}_{n:n:N}$ II. tür ilerleyen sansürlenmiş sıra istatistikleri, sırasıyla, sürekli F_x ve F_y dağılımları üzerine kurulu olsunlar. Ayrıca, $X^{\tilde{\mathbf{R}}}_{1:n:N} \le Y^{\tilde{\mathbf{R}}}_{1:n:N} \le X^{\tilde{\mathbf{R}}}_{n:n:N} \le Y^{\tilde{\mathbf{R}}}_{n:n:N}$ koşulu sağlanıyor olsun. Yeni bir gözlem olan Z, F_z dağılım fonksiyonuna sahip olsun. $H_0 : F_z = F_x$ hipotezini $H_1 : F_z = F_y$ hipotezine karşı test ederken kullanılacak karar kriterleri,

Eğer $Z \in \left(-\infty, Y^{\tilde{\mathbf{R}}}_{1:n:N}\right)$ ***ise,*** H_0 ***reddedilemez.***

Eğer $Z \in \left(X^{\tilde{\mathbf{R}}}_{n:n:N}, \infty\right)$ ***ise,*** H_0 ***reddedilir.***

Eğer $Z \in \left(Y^{\tilde{\mathbf{R}}}_{1:n:N}, X^{\tilde{\mathbf{R}}}_{n:n:N}\right)$ ***ise, karar verilemez***

şeklindedir.

Eğer $m_i = R_i$ ve $k = R_n + 1$ olarak seçilirse, genelleştirilmiş sıra istatistikleri $X\left(r, n, \tilde{\mathbf{R}}, R_n + 1\right) = X^{\tilde{\mathbf{R}}}_{r:n:N}$ ve $Y\left(r, n, \tilde{\mathbf{R}}, R_n + 1\right) = Y^{\tilde{\mathbf{R}}}_{r:n:N}$ şeklinde II. tür ilerleyen sansürlenmiş sıra istatistikleri haline gelecektir(bkz. Tablo-1). Sonuç olarak, $\gamma_i = n - i + 1 + \sum_{j=i}^{n} R_j$,

$$c_{n-1} = \prod_{i=1}^{n}\left(n-i+1+\sum_{j=i}^{n} R_j\right),\; a_i(n) = \prod_{\substack{j=1\\ j\neq i}}^{n}\left(i-j+\sum_{u=j}^{n} R_u - \sum_{j=i}^{n} R_j\right)^{-1}$$ olacak,

I. tip (α) ve II. tip (β) hatalar, sırasıyla (4.4) ve (4.5) kullanılarak,

$$\begin{aligned}\alpha &= c_{n-1}\sum_{i=1}^{n}\frac{a_i(n)}{(\gamma_i+1)}\\ &= \left(\prod_{i=1}^{n}\left(n-i+1+\sum_{j=i}^{n} R_j\right)\right)\\ &\quad\times\sum_{i=1}^{n}\prod_{\substack{j=1\\ j\neq i}}^{n}\frac{1}{\left(n-i+2+\sum_{u=i}^{n} R_u\right)\left(i-j+\sum_{u=j}^{n} R_u-\sum_{j=i}^{n} R_j\right)}\end{aligned} \tag{4.8}$$

ve

$$\beta = \frac{1}{n+1+\sum_{i=1}^{n} R_i} \tag{4.9}$$

şeklinde elde edilecektir.

4.2.2 Örneklem büyüklükleri ve sansürleme planları farklı olan iki eğitici örneklem için test prosedürü

$X_{1:n^{(x)}:N^{(x)}}^{\tilde{\mathbf{R}}^{(x)}}, X_{2:n^{(x)}:N^{(x)}}^{\tilde{\mathbf{R}}^{(x)}}, \ldots, X_{n^{(x)}:n^{(x)}:N^{(x)}}^{\tilde{\mathbf{R}}^{(x)}}$ ve $Y_{1:n^{(y)}:N^{(y)}}^{\tilde{\mathbf{R}}^{(y)}}, Y_{2:n^{(y)}:N^{(y)}}^{\tilde{\mathbf{R}}^{(y)}}, \ldots, Y_{n^{(y)}:n^{(y)}:N^{(y)}}^{\tilde{\mathbf{R}}^{(y)}}$ II. tür ilerleyen sansürlenmiş sıra istatistikleri, sırasıyla, sürekli F_x ve F_y dağılımları üzerine kurulu olsunlar. Ayrıca, $X_{1:n^{(x)}:N^{(x)}}^{\tilde{\mathbf{R}}^{(x)}} \le Y_{1:n^{(y)}:N^{(y)}}^{\tilde{\mathbf{R}}^{(y)}} \le X_{n^{(x)}:n^{(x)}:N^{(x)}}^{\tilde{\mathbf{R}}^{(x)}} \le Y_{n^{(y)}:n^{(y)}:N^{(y)}}^{\tilde{\mathbf{R}}^{(y)}}$ koşulu sağlanıyor olsun. Yeni bir gözlem olan Z, F_z dağılım fonksiyonuna sahip olsun.

$H_0 : F_z = F_x$ hipotezini $H_1 : F_z = F_y$ hipotezine karşı test ederken kullanılacak karar kriterleri,

Eğer $Z \in \left(-\infty, Y^{\tilde{\mathbf{R}}^{(y)}}_{1:n^{(y)}:N^{(y)}}\right)$ ***ise,*** H_0 ***reddedilemez.***

Eğer $Z \in \left(X^{\tilde{\mathbf{R}}^{(x)}}_{n^{(x)}:n^{(x)}:N^{(x)}}, \infty\right)$ ***ise,*** H_0 ***reddedilir.***

Eğer $Z \in \left(Y^{\tilde{\mathbf{R}}^{(y)}}_{1:n^{(y)}:N^{(y)}}, X^{\tilde{\mathbf{R}}^{(x)}}_{n^{(x)}:n^{(x)}:N^{(x)}}\right)$ ***ise, karar verilemez***

şeklindedir.

II. tür ilerleyen sansürlenmiş sıra istatistiklerinde, (4.6) ve (4.7)' de verilen hataların kullanılabilmesi için $m_i^{()} = R_i^{()}$ ve $k^{()} = R_n^{()} + 1$ olarak seçilmelidir (bkz. Tablo-1). Bu durumda, genelleştirilmiş sıra istatistikleri $X\left(r, n^{(x)}, \tilde{\mathbf{R}}^{(x)}, R^{(x)}_{n^{(x)}} + 1\right) = X^{\tilde{\mathbf{R}}^{(x)}}_{r:n^{(x)}:N^{(x)}}$ ve $Y\left(r, n^{(y)}, \tilde{\mathbf{R}}^{(y)}, R^{(y)}_{n^{(y)}} + 1\right) = Y^{\tilde{\mathbf{R}}^{(y)}}_{r:n^{(y)}:N^{(y)}}$ şeklinde II. tür ilerleyen sansürlenmiş sıra istatistikleri haline gelecektir. Sonuç olarak, $\gamma_i^{()} = n^{()} - i + 1 + \sum_{j=i}^{n^{()}} R_j^{()}$, $c^{()}_{n^{()}-1} = \prod_{i=1}^{n^{()}} \left(n^{()} - i + 1 + \sum_{j=i}^{n^{()}} R_j^{()} \right)$,

$a_i^{()}\left(n^{()}\right) = \prod_{\substack{j=1 \\ j \neq i}}^{n^{()}} \left(i - j + \sum_{u=j}^{n^{()}} R_u^{()} - \sum_{j=i}^{n^{()}} R_j^{()} \right)^{-1}$ olacak, I. tip (α) ve II. tip (β)

hatalar,

$$\alpha = c^{(\cdot)}_{n^{(\cdot)}-1}\sum_{i=1}^{n}\frac{a_i^{(\cdot)}(n)}{\left(\gamma_i^{(\cdot)}+1\right)}$$
$$=1-\left(\prod_{i=1}^{n^{(\cdot)}}\left(n^{(\cdot)}-i+1+\sum_{j=i}^{n^{(\cdot)}}R_j^{(\cdot)}\right)\right)$$
$$\times\sum_{i=1}^{n}\prod_{\substack{j=1\\ j\neq i}}^{n^{(x)}}\frac{1}{\left(n^{(x)}-i+2+\sum_{u=i}^{n}R_u^{(x)}\right)\left(i-j+\sum_{u=j}^{n^{(x)}}R_u^{(x)}-\sum_{j=i}^{n^{(x)}}R_j^{(x)}\right)} \qquad (4.10)$$

ve

$$\beta=\frac{1}{n^{(y)}+1+\sum_{i=1}^{n^{(y)}}R_i^{(y)}} \qquad (4.11)$$

şeklinde elde edilecektir.

5 UYGULAMALAR

Bu bölümde, üçüncü bölümde aşan istatistikler ve minimal boşluklarla ilgili verdiğimiz Teorem-1 ve Teorem-4' ü ve dördüncü bölümde verdiğimiz test prosedürlerini kullanabileceğimiz dört uygulamaya yer verilecektir.

5.1 Uygulama 1: Aşan İstatistikler

Nelson (1982)' un kitabında yer alan, izolasyonu yapılmış sıvılar üzerindeki deneyden elde edilen, 34 Kilovolt' taki bozulma sürelerine ilişkin ilerleyen sansürlenmiş veriler Tablo-2' de yer almaktadır.

Tablo 2 Bozulma sürelerine ilişkin ilerleyen sansürlenmiş veriler ***(Nelson, 1982)***

i	1	2	3	4	5	6	7	8
$X^{\tilde{\mathbf{R}}}_{i:8:19}$	0.19	0.78	0.96	1.31	2.78	4.85	6.5	7.35
R_i	0	0	3	0	3	0	0	5

Bu deneye toplam $N=19$ sıvı ile başlanmış ve sansürleme planı $\tilde{\mathbf{R}}=(0,0,3,0,3,0,0,5)$ olarak belirlenmiştir. Yani, üçüncü ve beşinci bozulma anlarında rasgele olarak seçilen bozulmamış üçer sıvı materyal rasgele olarak seçilerek deneyden çıkarılmış, sekizinci bozulma anında ise geriye kalan beş bozulmamış sıvı materyal deneyden çıkarılmıştır. Bu durumda, deney süresince toplam gözlenen bozulma süresi $n=8$ ve toplam sansürlenen veri sayısı ise $\sum_{i=1}^{8} R_i = 11$ olmuştur.

Bu deneye bağlı olarak, $m=3$ yeni bozulma süresinden l tanesini $\left(X_{2:8:19}^{\tilde{R}}, X_{6:8:19}^{\tilde{R}}\right)=(0.78,4.85)$ dağılımdan-bağımsız güven aralığına düşmesi olasılıkları, yani aşan istatistiğin olasılık fonksiyonu $P\{S_3=l\}$, Teorem-1 kullanılarak elde edilebilir: Teorem-1' i, bu örnekte kullanabilmek için $m_i=R_i$, $k=R_n+1$ olarak alınmalıdır. Elde edilen olasılıklar Tablo-3' te gösterilmiştir.

Tablo 3 ***Aşan istatistiğin olasılık fonksiyonu***

N	n	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	l	$P\{S_3=l\}$
19	8	0	0	3	0	3	0	0	5	0	0.4434
										1	0.3818
										2	0.1492
										3	0.0256

Tablo-3' te, $\left(X_{2:8:19}^{\tilde{R}}, X_{6:8:19}^{\tilde{R}}\right)$ dağılımdan bağımsız güven aralığına üç yeni gözlemden hiçbirinin düşmemesi olasılığının 0.4434 olduğu görülmektedir. Ayrıca l arttıkça olasılık düşmektedir.

Aynı sansürlenmiş verilere ilişkin farklı sansürleme planları için ilgili olasılıklar ise Tablo-4' te yeralmaktadır.

Tablo 4 ***Farklı sansürleme planları için, aşan istatistiğin olasılık fonksiyonları***

N	n	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	l	$P\{S_3 = l\}$
19	8	3	3	0	0	0	0	0	5	0	0.3769
										1	0.3952
										2	0.1884
										3	0.0395
19	8	3	0	3	0	0	0	0	5	0	0.3925
										1	0.3927
										2	0.1789
										3	0.0359
19	8	2	0	2	1	0	2	0	4	0	0.4293
										1	0.3861
										2	0.1568
										3	0.0278
19	8	2	0	2	1	2	1	1	2	0	0.4040
										1	0.3899
										2	0.1723
										3	0.0338

Yukarıdaki tabloda, sansürleme planı değiştikçe aşan istatistiğin olasılık fonksiyonunu değiştiği görülmektedir.

5.2 Uygulama 2: Minimal Boşluklar

Bir önceki uygulamada verilmiş olan Tablo-2' deki II. Tür ilerleyen sansürlenmiş örneklemi, $F(x) = 1 - \exp(-\lambda x)$, $\lambda > 0$, $x > 0$ olduğu varsayımı altında ele alacak olursak, aynı sansürleme planı için, $m = 3$ yeni gözlemin minimal boşluğa düşmesi olasılıkları, Teorem-4 kullanılarak, Tablo-5' de verilmiştir.

Tablo 5 $m = 3$ **yeni gözlemin minimal boşluğa düşmesi olasılıkları**

N	n	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	l	$P\{S_3 = l\}$
19	8	0	0	3	0	3	0	0	5	0	0.9750
										1	0.0246
										2	0.0002
										3	0.0000

Yukarıdaki tabloda, $l = 0$ için, yani 3 yeni gözlemden hepsinin minimal boşluğun dışına düşmesi olayının olasılığının %97.5 olduğu görülmektedir.

5.3 Uygulama 3: Sıra İstatistikleri için Hipotez Testi

Büyüklükleri $n = 12$ olan iki eğitici örneklem Tablo-6' da verilmektedir. Bu örneklemlerden ilki, kesin olarak şekersiz diyabet hastalığına yakalanmış 12 hastadan alınan mmol/L birimine sahip kan şekeri miktarlarından; ikincisi ise, kesin olarak şekerli diyabet hastalığına yakalanmış 12 hastanın kan şekeri miktarlarından oluşmaktadır. Bu durumda, şekersiz diyabet hastalığına sahip hastaların kan şekeri miktarlarına ilişkin dağılım fonksiyonunun F_x, şekerli diyabet hastalarınınki ise F_y olduğu varsayılsın. $F_x \neq F_y$ olmak üzere, F_x ve F_y' nin sürekli oldukları bilinmekte, ancak ne oldukları bilinmemektedir.

Tablo 6 Kesin olarak şekersiz ve şekerli diyabet hastalarının kan şekeri miktarlarına (mmol/L) ait eğitici örneklemler

i	1	2	3	4	5	6	7	8	9	10	11	12
$X_{i:12}$	6.8	6.9	7.2	7.3	7.3	7.5	7.5	7.6	7.8	7.8	7.9	8.1
$Y_{i:12}$	7.8	7.8	7.9	7.9	7.9	8.1	8.3	8.4	8.4	8.6	8.9	9.1

Olası bir diyabet hastasının kan şekeri ölçüldükten sonra, şekersiz diyabet mi yoksa şekerli diyabet mi olduğuna kara verilmek istendiğinde, yeni hastanın kan şekeri miktarı Z, kan şekeri miktarının dağılım fonksiyonu F_z olmak üzere, $H_0 : F_z = F_x$ hipotezini $H_1 : F_z = F_y$ hipotezine karşı test edebiliriz. Bu örnekte $X_{1:12} \leq Y_{1:12} \leq X_{12:12} \leq Y_{12:12}$ koşulu sağlanmaktadır ve karar kriterleri,

Eğer $Z \in (0,7.8)$ ***ise,*** H_0 ***reddedilemez.***

Eğer $Z \in (8.1,\infty)$ ***ise,*** H_0 ***reddedilir.***

Eğer $Z \in (7.8,8.1)$ ***ise, karar verilemez***

şeklinde belirlenebilir. (4.4) ve (4.5) teki hataların bu örnekte kullanılabilmesi için $\tilde{\mathbf{m}} = (0,0,...,0)$ ve $k = 1$ alınmalıdır.

Üç yeni gelen olası diyabet hastasının kan şekeri miktarlarına bakılarak verilecek kararlara ait hatalar (4.4) ve (4.5) kullanılarak hesaplanmış ve Tablo-7'de verilmiştir.

Tablo 7 Kan şekeri miktarına ilişkin olarak verilen kararlar ve hata olasılıkları

Kişi No	**Kan şekeri (mmol/L) (z)**	**KARAR**	**I. tip Hata (α)**	**II. tip Hata (β)**
1	6.9	*Kişi şekersiz diyabet hastasıdır.*	0.077	yok
2	8.2	*Kişi şekerli diyabet hastasıdır.*	yok	0.077
3	7.9	*Karar verilemez*	yok	yok

Yukarıdaki tablo incelendiğinde, birinci ve ikinci tür hataların birbirlerine eşit ve 0.077 olduğu, kan şekeri miktarı 6.9 mmol/L olan hastanın % 92.3 güven düzeyinde "şekersiz diyabet hastası" olduğu (yani H_0' ın kabulü) , 8.2 mmol/L olan hastanın yine %92.3 güven düzeyinde "şekerli diyabet hastası" olduğu (yani H_0' ın reddi) anlaşılmaktadır. H_0' ın reddedilmesi durumunda ortaya çıkan ikinci tip hatanın 0.077 olduğu görülmektedir. Ayrıca, kan şekeri miktarı 7.9 mmol/L olan hastanın hangi tür diyabet hastası olduğuna karar verilememektedir.

5.4 Uygulama 4: II. Tür İlerleyen Sansürlenmiş Sıra İstatistikleri için Hipotez Testi

Varsayalım ki, bir laboratuarda, deneylerde kullanılan farelerde dişi ebeveynlerinden %100 geçen bir ölümcül bir virüs saptanmış (A virüsü) ve bu virüsün iki versiyonu (A_1 ve A_2) belirlenmiştir. A virüsünü taşıyan fareler en fazla 730 gün yaşamaktadır. Deneylerde kullanılan farelerin, genetik yapıları itibariyle, A virüsü dışında herhangi bir sebepten dolayı 730 günden önce ölmelerinin mümkün olmadığı varsayılsın. Ayrıca, A_1'li farelerin yaşam sürelerinin, A_2'li farelerinkinden daha küçük olduğu bilinmektedir. Dişi ebeveyni hakkında A virüsünü taşıyıp taşımadığına ilişkin bilgiye sahip olunmadığı durumda, yeni doğan bir farenin A virüsünü taşıyıp-taşımadığı ve eğer taşıyorsa A virüsünün hangi versiyonunu taşıdığı, çok pahalı bir yöntemle (D yöntemiyle) belirlenebilmektedir. Yani, D yönteminin maliyeti, bir deney faresinin maliyetine göre çok fazladır.

Araştırmacılar, bir yandan fareler üzerinde farklı amaçlı pek çok deney yaparken, diğer yandan 730 günden önce ölen farelerin yaşam sürelerine bakarak, A_1 ve A_2 virüslerinden hangisini taşıdıklarına karar verebilmek ve bu bilgileri kaydetmek istiyorlar. Kaydedilen bu bilgilerin, A virüsünün versiyonlarının görünme oranlarının tahmin edilmesinde kullanılacağı düşünülsün.

Araştırmacılar, yukarıdaki karar problemini çözebilmek için, A_1 virüsünü taşıdığı D yöntemiyle belirlenen 100 günlük 20 tane fareyi ve A_2 virüsünü taşıdığı D yöntemiyle belirlenen 100 günlük 20 tane fareyi iki grup (sırasıyla, Grup1 ve Grup2) olarak almışlardır. Her iki grupta da ayrı ayrı yapacakları bir deneyde, ilk gözlemlenen ölüm anında yaşam süresi kaydedilecek ve rasgele olarak seçilen bir canlı denek başka deneylerde kullanılmak üzere deney grubundan çıkarılacaktır. Geriye kalan 18 canlı denekle deney devam edecek ve ikinci ölüm anında yaşam süresi kaydedilecek, geriye kalan canlı deneklerden rasgele olarak seçilen bir tanesi diğer deneylerde kullanılmak üzere deney grubundan yine çıkarılacaktır. Deney böylece devam edecek ve her iki grupta da 10. ölüm gerçekleştiğinde sona erecektir.

Deneyin sonucunda, her iki grup için, yaşam sürelerinden oluşan II. tür ilerleyen sansürlenmiş eğitici örneklemler elde edilmiş olur. Bu deneyde her iki grupta da kullanılan sansürleme planı $\tilde{\mathbf{R}} = (1,1,1,1,1,1,1,1,1,1)$ şeklindedir.

Grup1' deki yaşam süreleri (saat), $i = 1,2,\ldots,10$ olmak üzere $X_{i:10:20}^{\tilde{\mathbf{R}}}$ ile ve Grup2' dekiler ise $Y_{i:10:20}^{\tilde{\mathbf{R}}}$ ile gösterilsin. Ayrıca, A_1 ve A_2 virüslerini taşıyan farelerin yaşam sürelerinin olasılık dağılım fonksiyonları, sırasıyla, F_x ve F_y ile gösterilsin. Varsayalım ki, bu dağılım fonksiyonlarının sürekli ve $F_x \neq F_y$ olduğu bilinmekte, ancak ne oldukları bilinmemektedir.

730 günden önce ölen bir farenin yaşam süresi Z ile, Z' nin dağılım fonksiyonu ise F_z ile gösterilsin. İki gruba ait deney sonucunda elde edilen II. tür ilerleyen sansürlenmiş eğitici örneklemler Tablo-8'de yer almaktadır.

Tablo 8 A_1 ve A_2 virüsü taşıyan farelerin yaşam sürelerine ilişkin II. Tür ilerleyen sansürlenmiş eğitici örneklemler

i	1	2	3	4	5	6	7	8	9	10
$X_{i:10:20}^{\tilde{R}}$	100.2	190.3	215.8	320.1	329.6	440.6	450.1	451.8	559.8	564.5
$Y_{i:10:20}^{\tilde{R}}$	326.2	430.6	433.7	548.4	566.3	580.0	591.5	597.0	683.6	703.1
$\tilde{R}$	1	1	1	1	1	1	1	1	1	1

Sonuç olarak; $X_{1:10:20}^{\tilde{R}} \le Y_{1:10:20}^{\tilde{R}} \le X_{10:10:20}^{\tilde{R}} \le Y_{10:10:20}^{\tilde{R}}$ varsayımı altında, bölüm 4.2.1' deki test prosedürü ve (4.8) ve (4.9)' daki hatalar, 730 günden önce ölen bir farenin hangi virüsü taşıdığını, istatistiksel olarak tahmin etmek için kullanılabilir. Bu amaçla, $H_0 : F_z = F_x$ hipotezini $H_1 : F_z = F_y$ hipotezine karşı test etmek için kullanılacak karar kriterleri,

Eğer $Z \in (0,326.2)$ ***ise,*** H_0 ***reddedilemez.***

Eğer $Z \in (564.5,\infty)$ ***ise,*** H_0 ***reddedilir.***

Eğer $Z \in (326.2,564.5)$ ***ise, karar verilemez***

şeklinde olacaktır. Ömrü 730 günden az olan üç yeni fareye ilişkin ölüm yaşları, test sonuçları ve ilgili hatalar aşağıdaki Tablo-9'da yer almaktadır.

Tablo 9 Ömrü 730 günden az olan üç yeni fareye ilişkin yaşam süreleri ve test sonuçları

Fare No	z (gün)	KARAR	α	β
1	220.2	*Fare A_1 virüsünü taşıyordu.*	0.270	Yok
2	593.0	*Fare A_2 virüsünü taşıyordu.*	yok	0.047
3	428.5	*Farenin hangi virüsü taşıdığına karar verilemez.*	yok	Yok

Tablo-9 incelendiğinde, I.Tip hatanın 0.27, II.Tip hatanın ise 0.047 olduğu, ömrü 220.2 gün olan fare için %73 güven düzeyinde H_0' ın reddedilemediği (yani farenin A_1 virüsünü taşıdığına karar verilmiş olduğu); ömrü 593 gün olan fare içinse %73 güven düzeyinde H_0 reddedilmiş olduğu (yani A_2 virüsünü taşıdığına karar verildiği), bu son karara ilişkin ikinci tip hatanın ise %4.7 olduğu görülmektedir. Ayrıca, ömrü 428.5 gün olan fareye ilişkin karar verilememiştir.

6 GENEL SONUÇLAR

Bu çalışmada verilen 4 teoremle; genelleştirilmiş sıra istatistiklerine dayalı dağılımdan-bağımsız (veya invaryant) güven aralıkları kurulmuş ve güven düzeyleri verilmiş, yine genelleştirilmiş sıra istatistiklerine dayalı olarak aşan istatistiklerin olasılık dağılımları incelenmiş ve elde edilmiş, genelleştirilmiş sıra istatistiklerine dayalı istatistiklerin üstel dağılım için karakterizasyonu yapılmış, m adet yeni gözlemden l tanesinin minimal boşluğa düşme olasılıkları, parametrelere bağlı olarak verilmiştir.

Diğer yandan, Bairamov ve Petunin' in (1991b) önerdikleri test prosedürüne uygun olarak, genelleştirilmiş sıra istatistiklerine dayalı test prosedürü ortaya koyulmuştur.

Sonuç olarak elde edilen istatistiksel bulgu ve yöntemler, genelleştirilmiş sıra istatistiklerinin alt modellerinden uygun olanlarında, örneğin, sıra istatistiklerinde, II. Tür ilerleyen sansürlenmiş sıra istatistiklerinde, ardışık sıra istatistiklerinde vb. kullanılabilecektir.

KAYNAKLAR DİZİNİ

Aggarwala, R., Balakrishnan, N., 1998, *Some properties of progressive censored order statistics from arbitrary and uniform distributions with applications to inference and simulation*, Journal of Statistical Planning and Inference 70, 35-49.

Ahsanullah M., 2000, *Generalized order statistics from exponential distribution*, Journal of Statistical Planning and Inference 85, 85-91.

Arnold, B.C., Balakrishnan, N., (1989), Relations, Bounds & Approximations for Order Statistics (Lecture Notes in statistics, Vol.53), Springer-Verlag.

Arnold, B.C., Balakrishnan, N., Nagaraja, H.N., (1992), A First Course in Order Statistics, John Wiley & Sons, New York.

Arnold, B.C., Balakrishnan, N., Nagaraja, H.N., (1998), Records, Wiley, New York.

Bairamov, I.G., Eryılmaz, S.E., 2000, *Distributional properties of statistics based on minimal spacing and record exceedance statistics*, Journal of Statistical Planning and Inference 90, 21-33.

Bairamov, I.G., Kotz, S., 2001, On distributions of exceedances associated with order statistics and record values for arbitrary distributons, Statistical Papers, Vol.42, No.2, 171-185.

Bairamov, I.G., Özkaya, N., 2000, *On the nonparametric test for two sample problem based on spacings*, Journal of Applied Statistical Science, Vol.10, No.1, 57-68.

Bairamov, I.G., Petunin, Yu.I., 1991a, *Structure of invariant confidence intervals containing the main distributed mass*, Theor. Probab. Appl., Vol.35, Num.1, 15-26.

Bairamov, I.G., Petunin, Yu.I., 1991b, *Statistical test based of training samples*, Cybernetics, Vol.27, Num.3, 408-413.

Balakrishnan, N., Aggarwala, R., (2000), Progressive censoring: Theory, methods and applications, Birkhauser, Boston.

Balakrishnan, N., Childs, A., and Chandrasckar, B., 2002, *An efficient computational method for moments of order statistics under progressive censoring*, Statistics & Probability Letters 60, 359-365.

Balakrishnan, N., Cramer, E., and Kamps, U., 2001, *Bounds for means and variances of progressive type II censored order statistics*, Statistics & Probability Letters 54, 301-315.

Bieniek, M., Szynal, D., 2003, *Characterizations of distributions via linearity of regression of generalized order statistics*, Metrica 58, 259-271.

David, H.A., (1970), Order Statistics, Wiley, New York.

Gibbons, J.D., (1971), Nonparametric Statistical Inference, McGraw-Hill, Tokyo.

Kamps, U., (1995), A Concept of Generalized Order Statistics / von Udo Kamps, Teubner, Stuttgart.

Kamps U., Cramer E., 2001, *On distribution of generalized order statistics*, Statistics 35, 269-280.

Nelson, W., 1982, Applied Life Data Analysis, Wiley, New York.

ÖZGEÇMİŞ

Dört erkek kardeşin en büyüğü olarak 1974 yılında İstanbul' da dünyaya gelen Halil Tanıl, ilk, orta ve lise öğrenimini İzmir' de tamamlamıştır. 1992 yılında Ege Üniversitesi Fen Fakültesi İstatistik Bölümü' ne girmiş ve 1996 yılında bu bölümden birincilikle mezun olmuştur. Yüksek Lisansını, Ege Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı'nda, "Benzetimde Otomatik Olarak Girdi Analizi Yapan Bir Sistem Geliştirme" konulu tezi vererek 2000 yılında tamamlamıştır. Yine aynı yıl, Ege Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı'nda Doktoraya başlamıştır.

Yüksek Lisans sırasında 2.5 yıl bilgisayar eğitim uzmanı olarak özel sektörde çalıştıktan sonra, 1999' da lisansını tamamladığı bölümde Araştırma Görevlisi olarak çalışmaya başlamıştır. Halen bu görevini sürdürmekte olan Halil Tanıl, evli ve bir çocuk babasıdır.