

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**SIRALI İSTATİSTİKLERİN BAZI SİSTEMATİK
MOMENTLERİ**

Gökhan GÖKDERE

YÜKSEK LİSANS TEZİ
MATEMATİK ANABİLİM DALI

DANIŞMAN
Doç. Dr. Mehmet GÜNGÖR

ELAZIĞ, 2005

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

SIRALI İSTATİSTİKLERİN BAZI SİSTEMATİK MOMENTLERİ

Gökhan GÖKDERE

YÜKSEK LİSANS TEZİ
MATEMATİK ANABİLİM DALI

Bu tez, 12.08.2005 tarihinde, aşağıda belirtilen jüri tarafından
oybirliği/oyçokluğu ile başarılı/başarısız olarak değerlendirilmiştir.

Danışman: Doç. Dr. Mehmet GÜNGÖR



Üye: Doç. Dr. Mikail ET



Üye: : Yrd. Doç. Dr. Sinan ÇALIK



Bu tezin kabulü, Fen Bilimleri Enstitüsü Yönetim Kurulu'nun 19/08/2005 tarih ve
2005 - 27 / 8 sayılı kararıyla onaylanmıştır.

TEŞEKKÜR

Bu tezin hazırlanmasında gerekli bütün imkanları sağlayarak bana yardımcı olan, her zaman ilgi ve yardımlarını esirgemeyen çok kıymetli hocam sayın Doç. Dr. Mehmet GÜNGÖR 'e şükranlarımı sunmayı bir borç bilirim. Saygılarımı sunarım.

Ayrıca tezin hazırlanmasında yardımcı olan değerli arkadaşım Yunus BULUT'a teşekkür ederim.

Gökhan GÖKDERE



İÇİNDEKİLER

	<u>Sayfa No</u>
İçindekiler.....	I
Simgeler Listesi.....	II
Özet.....	III
Abstract.....	IV
1. Temel Tanım ve Teoremler.....	1
2. Sıralı İstatistiklerin Bazı Sistematiik Momentleri İçin Temel Formüller.....	10
2.1. Sıralı İstatistiklerin Bazı Sistematiik Momentleri İçin Temel Formüllerin İfade Edilmesi	10
2.2. Sıralı İstatistikler İçin Sistematiik Momentler Arasındaki Bazı Temel Bağıntılar.....	11
3. Sıralı İstatistiklerin Bazı Sistematiik Momentlerinin Yaklaşık Hesaplamaları.....	13
3.1. Sıralı İstatistiklerin Bazı Sistematiik Momentlerinin Yaklaşık Hesaplamalarının İncelenmesi.....	13
3.2. i-inci Sıralı İstatistiğın Ortalaması ve Varyansının Yaklaşık Hesabı.....	15
3.3. Uygulamalar.....	16
3.4. Sonuç.....	18
Kaynaklar.....	19
Özgeçmiş	

SİMGELER LİSTESİ

pdf	: İhtimal yoğunluk fonksiyonu
cdf	: Kümülatif dağılım fonksiyonu
i.i.d.	: Bağımsız aynı dağılımlı
i.n.n.i.d.	: Bağımsız aynı dağılımlı olmayan
R_x	: X tesadüfi değişkeninin tanım aralığı
$\underline{d}$	: Dağılım olarak eşit
μ	: Anakütle ortalaması
σ	: Anakütle standart hatası(standart sapma)



ÖZET

Yüksek Lisans Tezi

SIRALI İSTATİSTİKLERİN BAZI SİSTEMATİK MOMENTLERİ

Gökhan GÖKDERE

Fırat Üniversitesi
Fen Bilimleri Enstitüsü
Matematik Anabilim Dalı

2005, Sayfa: 19

Bu tez üç bölümden oluşmaktadır.

Birinci bölümde, temel tanım ve teoremler verilmiştir.

İkinci bölümde, sıralı istatistiklerin bazı sistematik momentleri için temel formüller incelenmiştir.

Üçüncü bölümde, sıralı istatistiklerin bazı sistematik momentlerinin yaklaşık hesaplamaları incelenmiş ve bazı özel dağılımlar için sayısal sonuçlar verilmiştir.

Anahtar Kelimeler: Sıralı istatistikler, momentler, i.i.d. tesadüfi değişkenler, i.n.n.i.d. tesadüfi değişkenler.

ABSTRACT

Master Thesis

SOME SYSTEMATIC MOMENTS OF ORDER STATISTICS

Gökhan GÖKDERE

Firat University
Graduate School of Natural and Applied Sciences
Department of Mathematics

2005, Page : 19

This thesis consists of three chapters.

In the first chapter, some fundamental definitions and theorems are given.

In the second chapter, fundamental formulas for some systematic moments of order statistics are investigated.

In the third chapter, approximate calculations of some systematic moments of order statistics are examined and numerical results for some special distributions are given.

Keywords: Order statistics, moments, i.i.d. random variables, i.n.n.i.d. random variables.

1. BÖLÜM

1. TEMEL TANIM VE TEOREMLER

Tanım 1.1. (Sürekli tesadüfi değişkenler için olasılık yoğunluk fonksiyonu)

$X, (-\infty, \infty)$ aralığında tanımlanan sürekli tesadüfi değişken olsun. Aşağıdaki koşulları sağlayan $f(x)$ fonksiyonuna X tesadüfi değişkeninin olasılık yoğunluk fonksiyonu denir.

1. $f(x) \geq 0$, $-\infty < x < \infty$

2. $\int_{-\infty}^{\infty} f(x) dx = 1$

Ayrıca $f(x)$, (1) ve (2) şartlarını sağlarsa sürekli X tesadüfi değişkeninin a ve b arasında bulunması olasılığı;

$$P\{a \leq X \leq b\} = \int_a^b f(x) dx \quad a, b \in R \quad \text{ve} \quad a \leq b \quad (1.1)$$

eşitliği ile ifade edilir. Sürekli X tesadüfi değişkeninin belli bir x değerini alması olasılığı sıfırdır. Yani, $P(X = x) = 0$ dir [1].

Tanım 1.2. (Sürekli tesadüfi değişkenler için kümülatif dağılım fonksiyonu)

X , olasılık yoğunluk fonksiyonu $f(x)$ olan sürekli tesadüfi bir değişken olsun. Herhangi bir x reel değeri için X 'in kümülatif dağılım fonksiyonu,

$$F(x) = P\{X \leq x\} = \int_{-\infty}^x f(t) dt \quad (1.2)$$

eşitliği ile ifade edilir [2].

Teorem 1.1.

Eğer X , olasılık yoğunluk fonksiyonu $f(x)$ ve kümülatif dağılım fonksiyonu $F(x)$ 'e sahip bir anakütleden seçilmiş sürekli tesadüfi bir değişken ise $F(x)$ 'in tanımlı olduğu bütün noktalarda

$$f(x) = F'(x) \quad (1.3)$$

eşitliği vardır [3].

Tanım 1.3. (Kesikli tesadüfi değişkenler için olasılık fonksiyonu)

X , sonlu sayıdaki $x_1, x_2, \dots, x_n$ değerlerini $f(x_i) = P(X = x_i)$, $i = 1, 2, \dots, n$ olasılıkları ile alabilen kesikli tesadüfi değişken olsun. Bu durumda,

1. $f(x) \geq 0$, tüm x 'ler için

2. $\sum_i f(x_i) = 1$

koşullarını sağlayan $f(x)$ fonksiyonuna X 'in olasılık fonksiyonu denir [1].

Tanım 1.4. (Kesikli tesadüfi değişkenler için kümülatif dağılım fonksiyonu)

Bir X tesadüfi değişkeninin dağılım fonksiyonu $F(x)$ ile gösterilir ve X 'in x e eşit ya da daha küçük olması olasılığıdır. O halde,

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i) \quad (1.4)$$

dir [1].

Tanım 1.5. (Beklenen değer)

X tesadüfi bir değişken olsun. O zaman X 'in beklenen değeri

$$E(X) = \begin{cases} \sum_i x_i f(x_i), & X \text{ kesikli ise} \\ \int_{R_X} x f(x) dx, & X \text{ sürekli ise} \end{cases} \quad (1.5)$$

şeklinde ifade edilir [3].

Tanım 1.6. (k -ıncı moment)

X tesadüfi bir değişken ve k herhangi bir tamsayı olsun.

i) X tesadüfi değişkeninin sıfıra göre k -ıncı momenti μ_k ,

$$\mu_k = E(X^k) = \begin{cases} \sum_i x_i^k f(x_i), & X \text{ kesikli ise} \\ \int_{R_X} x^k f(x) dx, & X \text{ sürekli ise} \end{cases} \quad (1.6)$$

ii) X tesadüfi değişkeninin μ ortalamaya göre k -ıncı momenti μ'_k ,

$$\mu'_k = E[(X - \mu)^k] = \begin{cases} \sum_i (x_i - \mu)^k f(x_i), & X \text{ kesikli ise} \\ \int_{R_X} (x - \mu)^k f(x) dx, & X \text{ sürekli ise} \end{cases} \quad (1.7)$$

şeklinde ifade edilir [4].

Tanım 1.7. (Varyans)

X tesadüfi değişkeninin ortalamadan sapmasının karesinin beklenen değerine varyans denir ve

$$Var(x) = E[(X - \mu)^2] = E(X^2) - [E(X)]^2 \quad (1.8)$$

olarak ifade edilir [5].

Tanım 1.8. (Kovaryans)

X ve Y tesadüfi değişkenlerinin beklenen değerleri sırasıyla $E(X)$ ve $E(Y)$ olsun. Bu durumda X ve Y 'nin kovaryansı,

$$\begin{aligned} Cov(X, Y) &= E[(X - E(X))(Y - E(Y))] \\ &= E(XY) - E(X)E(Y) \end{aligned} \quad (1.9)$$

şeklinde ifade edilir [4].

Tanım 1.9 (Moment çıkaran fonksiyon)

X tesadüfi değişkeninin moment çıkaran fonksiyonu $m(t)$,

$$m(t) = E(e^{tX}) = \begin{cases} \sum_i e^{tx_i} f(x_i), & X \text{ kesikli ise} \\ \int_{R_X} e^{tx} f(x) dx, & X \text{ sürekli ise} \end{cases} \quad (1.10)$$

şeklinde tanımlanır [4].

Tanım 1.10. (Düzgün dağılım)

X tesadüfi değişkeninin olasılık yoğunluk fonksiyonu,

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{diğer yerlerde} \end{cases} \quad (1.11)$$

şeklinde tanımlansın. Bu durumda X tesadüfi değişkenine düzgün dağılmış tesadüfi bir değişken denir. Düzgün dağılımın kümülatif dağılım fonksiyonu

$$F(x) = \begin{cases} 0, & x < a \quad \text{ise} \\ \frac{x-a}{b-a}, & a \leq x \leq b \quad \text{ise} \\ 1, & x > b \quad \text{ise} \end{cases} \quad (1.12)$$

şeklinde ifade edilir [4].

Tanım 1.11. (Normal dağılım)

Sürekli bir X tesadüfi değişkeni için olasılık yoğunluk fonksiyonu

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2 \right], \quad -\infty < x < \infty, -\infty < \mu < \infty, \sigma^2 > 0 \quad (1.13)$$

ise X normal dağılıma sahiptir denir [1].

Tanım 1.12. (Üstel dağılım)

Negatif olmayan değerler alan sürekli bir X tesadüfi değişkeninin olasılık yoğunluk fonksiyonu aşağıdaki gibi olsun.

$$f(x) = \begin{cases} \alpha e^{-\alpha x}, & x > 0 \\ 0, & \text{diğer yerlerde.} \end{cases} \quad (1.14)$$

Bu taktirde X , $\alpha > 0$ parametresi ile üstel dağılıma sahiptir denir [1].

Tanım 1.13. (Lojistik dağılım)

$-\infty < x < \infty$ için pdf'si $f(x) = \frac{e^{-x}}{(1+e^{-x})^2}$ ve cdf'si $F(x) = \frac{1}{1+e^{-x}}$ olan dağılıma lojistik dağılım denir [6].

Tanım 1.14. (Sıralı istatistikler)

$X_1, X_2, \dots, X_n$, n tane tesadüfi değişken olsun. Eğer $X_1, X_2, \dots, X_n$ 'lerin meydana gelme sırası ile değil, büyüklüklerinin sırası ile ilgilenilirse n birimli bir örneğin sıralı istatistiği,

$$X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n} \quad (1.15)$$

olarak ifade edilir.

$X_{i:n}$, i -inci sıralı istatistik ve $(X_{1:n}, X_{2:n}, \dots, X_{n:n})$ tesadüfi vektörüne de sıralı istatistikler denilir. $X_{1:n}$ örnek minimum ve $X_{n:n}$ örnek maksimumdur. Bundan dolayı

$$X_{1:n} = \min(X_1, X_2, \dots, X_n) \quad (1.16)$$

ve

$$X_{n:n} = \max(X_1, X_2, \dots, X_n) \quad (1.17)$$

yazılabilir. Bazen i -inci sıralı istatistik,

$$Z_{i:n}(x_1, x_2, \dots, x_n) = z_i \quad (1.18)$$

şeklinde de verilebilir. Burada $z_1 \leq z_2 \leq \dots \leq z_n$, $x_1, x_2, \dots, x_n$ 'lerin reel değerlerinin artan sırada düzenlenmiş değerleridir. Buradan

$$X_{i:n} = Z_{i:n}(X_1, X_2, \dots, X_n) \quad (1.19)$$

olur [7].

Tanım 1.15. (Bileşik olasılık yoğunluk fonksiyonu)

X ve Y sürekli tesadüfi değişkenleri xy -düzleminin herhangi bir A bölgesinde tanımlı olsun. X ve Y tesadüfi değişkenlerinin A bölgesinde olma olasılığını veren

$$P\{(X, Y) \in A\} = \iint_A f(x, y) dx dy \quad (1.20)$$

eşitliğindeki $f(x, y)$ fonksiyonuna (X, Y) tesadüfi değişkeninin bileşik olasılık yoğunluk fonksiyonu denir [8].

Tanım 1.16. (Bileşik dağılım fonksiyonu)

X ve Y sürekli tesadüfi değişkenlerinin bileşik dağılım fonksiyonu, $-\infty < x < \infty$ ve $-\infty < y < \infty$ için

$$F(x, y) = P\{X \leq x, Y \leq y\} = \int_{-\infty}^x \int_{-\infty}^y f(s, t) ds dt \quad (1.21)$$

eşitliği ile ifade edilir [8].

Tanım 1.17. (i.i.d. sıralı istatistiklerin bileşik yoğunluk fonksiyonu)

$-\infty < x_{1:n} < x_{2:n} < \dots < x_{n:n} < \infty$ için $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$ sıralı istatistiklerinin bileşik yoğunluk fonksiyonu

$$f_{1,2,\dots,n:n}(x_{1:n}, x_{2:n}, \dots, x_{n:n}) = n! \prod_{i=1}^n f(x_{i:n}) \quad (1.22)$$

olarak ifade edilir [8].

Tanım 1.18. (i.i.d. sıralı istatistiklerinin marjinal yoğunluk fonksiyonu)

$X_{i:n}$ 'nin marjinal yoğunluk fonksiyonu, $1 \leq i \leq n$ ve $-\infty < x_{i:n} < \infty$ için

$$f_{i:n}(x_{i:n}) = \frac{n!}{(i-1)!(n-i)!} [F(x_{i:n})]^{i-1} [1 - F(x_{i:n})]^{n-i} f(x_{i:n}) \quad (1.23)$$

olarak ifade edilir [8].

Teorem 1.2. (i.i.d. sıralı istatistiklerin kümülatif dağılım fonksiyonu)

F ortak dağılım fonksiyonlu n tane $X_1, X_2, \dots, X_n$ i.i.d. tesadüfi değişkenlerin i -inci sıralı istatistiği $X_{i:n}$ olsun. O zaman her x için

$$P\{X_{i:n} \leq x\} = \sum_{m=i}^n \binom{n}{m} F^m(x) (1 - F(x))^{n-m} \quad (1.24)$$

dir.

Özel olarak $X_{1:n}$ minimumun ve $X_{n:n}$ maksimumun kümülatif dağılım fonksiyonları sırasıyla

$$P\{X_{1:n} \leq x\} = 1 - [1 - F(x)]^n \quad (1.25)$$

ve

$$P\{X_{n:n} \leq x\} = F^n(x) \quad (1.26)$$

olarak yazılabilir [7].

Tanım 1.19. (İki sıralı istatistiğin bileşik yoğunluk fonksiyonu)

$X_{i:n}$ ve $X_{j:n}$ nin bileşik yoğunluk fonksiyonu $1 \leq i \leq j \leq n$ ve $-\infty < x_{i:n} < x_{j:n} < \infty$ için

$$f_{i,j:n}(x_{i:n}, x_{j:n}) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} [F(x_{i:n})]^{i-1} [F(x_{j:n}) - F(x_{i:n})]^{j-i-1} [1 - F(x_{j:n})]^{n-j} f(x_{i:n}) f(x_{j:n}) \quad (1.27)$$

olarak ifade edilir [8].

Teorem 1.3. (Bağımlı örneklerde sıralı istatistikler)

$A_1, A_2, \dots, A_n$ olasılık uzayında alınan keyfi olaylar olsun. O zaman;

$$P\left\{\bigcup_{i=1}^n A_i\right\} = \sum_{i=1}^n P\{A_i\} - \sum_{1 \leq i_1 < i_2 \leq n} P\{A_{i_1} \cap A_{i_2}\} + \sum_{1 \leq i_1 < i_2 < i_3 \leq n} P\{A_{i_1} \cap A_{i_2} \cap A_{i_3}\} - \dots + (-1)^{n+1} P\{A_1 \cap A_2 \cap \dots \cap A_n\} \quad (1.28)$$

yazılır ve

$$S_m = \sum_{1 \leq i_1 < i_2 < \dots < i_m \leq n} P\{A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_m}\} \quad (1.29)$$

olarak alınırsa

$$P\left\{\bigcup_{i=1}^n A_i\right\} = 1 - P\left\{\bigcap_{i=1}^n \overline{A_i}\right\} \quad (1.30)$$

olur. Burada $\overline{A_i}$, A_i 'nin tümleyenidir. (1.28)'den

$$P\left\{\bigcap_{i=1}^n \overline{A_i}\right\} = 1 - P\left\{\bigcup_{i=1}^n A_i\right\} = \sum_{i=0}^n (-1)^i S_i \quad (1.31)$$

yazılabilir. Burada $S_0 = 1$ dir [9].

Tanım 1.20. (Bağımlı Sıralı İstatistikler)

$X_1, X_2, \dots, X_n$ bağımlı tesadüfi değişkenler olsunlar. $X_1, X_2, \dots, X_n$ tesadüfi değişkenlerin bileşik dağılım fonksiyonunu $P_n(x_1, x_2, \dots, x_n)$ ile gösterelim. Özel olarak

$$P\{X_{n:n} \leq x\} = P_n(x, x, \dots, x) \quad (1.32)$$

eşitliği yazılabilir. (1.32) eşitliği i -inci sıralı istatistik için

$$F_{i:n}(x) = P\{X_{i:n} \leq x\} = \sum_{m=i}^n (-1)^{m-i} \binom{m-1}{i-1} S_m \quad (1.33)$$

şeklinde genelleştirilebilir. Burada $S_m, i_1 < i_2 < \dots < i_m$ olmak üzere $\binom{n}{m}$ tane $P\{X_{i_1} \leq x, X_{i_2} \leq x, \dots, X_{i_m} \leq x\}$ olasılıklarının toplamıdır. Özel olarak X_i tesadüfi değişkenleri değiştirilebilir tesadüfi değişkenler (yani; P_n olasılığı $x_1, x_2, \dots, x_n$ 'lere göre simetrik) ise (1.33) ifadesi

$$F_{i:n}(x) = P\{X_{i:n} \leq x\} = \sum_{m=i}^n (-1)^{m-i} \binom{m-1}{i-1} \binom{n}{m} P\{X_{m:m} \leq x\} \quad (1.34)$$

şeklinde ifade edilebilir [8].

Teorem 1.4.

$f_1(x_1)$ ve $f_2(x_2)$ sırasıyla X_1 ve X_2 'nin marjinal yoğunluk fonksiyonları ve $f(x_1, x_2)$ de X_1 ve X_2 'nin bileşik yoğunluk fonksiyonu olmak üzere X_1 ve X_2 tesadüfi değişkenlerinin stokastik bağımsız olması için gerek ve yeter şart,

$$f(x_1, x_2) = f_1(x_1) f_2(x_2) \quad (1.35)$$

olmasıdır [3].

Teorem 1.5. (i.n.n.i.d. tesadüfi değişkenler)

$X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}, X_1, X_2, \dots, X_n$ i.n.n.i.d. tesadüfi değişkenlerinin sıralı istatistikleri ve F_i, X_i 'nin dağılım fonksiyonu olsun. O zaman her B Borel cümlesi için

$$P\{(X_{1:n}, X_{2:n}, \dots, X_{n:n}) \in B\} = \sum_{m=1}^n (-1)^{n-m} \frac{m^n}{n!} \sum_{|S|=m} P\{(X_{1:n}^S, X_{2:n}^S, \dots, X_{n:n}^S) \in B\} \quad (1.36)$$

ifadesi yazılır. Burada toplam, $\{1, 2, \dots, n\}$ 'nin m elemanlı tüm S alt kümeleri üzerinden alınır. Bununla birlikte $X_{1:n}^S \leq X_{2:n}^S \leq \dots \leq X_{n:n}^S$,

$$F^S = |S|^{-1} \sum_{i \in S} F_i \quad (1.37)$$

ortak dağılım fonksiyonlu n tane i.i.d. tesadüfi değişkenin sıralı istatistiklerini göstermektedir [7].

II. BÖLÜM

2. SIRALI İSTATİSTİKLERİN BAZI SİSTEMATİK MOMENTLERİ İÇİN TEMEL FORMÜLLER

2.1. Sıralı istatistiklerin bazı sistematik momentleri için temel formüllerin ifade edilmesi

$X_1, X_2, \dots, X_n$ olasılık yoğunluk fonsiyonu $f(x)$ ve kümülatif dağılım fonsiyonu $F(x)$ olan sürekli bir anakütleden seçilen bir tesadüfi örnek olsun ve bu örneğin sıralı istatistikleri $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$ olacak şekilde elde edilmiş olsun. Tekli moment olan $E(X_{i:n}^k)$ ifadesi $\alpha_{i:n}^{(k)}$ ile gösterilirse Tanım 1.18 ve Tanım 1.6'dan $i = 1, 2, \dots, n$ ve $k \geq 1$ için

$$\begin{aligned} \alpha_{i:n}^{(k)} &= E(X_{i:n}^k) = \int_{-\infty}^{\infty} x^k f_{i:n}(x) dx \\ &= \frac{n!}{(i-1)!(n-i)!} \int_{-\infty}^{\infty} x^k \{F(x)\}^{i-1} \{1-F(x)\}^{n-i} f(x) dx \end{aligned} \quad (2.1.1)$$

elde edilir. Kolaylık için $\alpha_{i:n}^{(1)}$ basit olarak $\alpha_{i:n}$ ile gösterilirse $X_{i:n}$ 'nin ilk iki tekli momentlerinden faydalanarak varyans, Tanım 1.7'den

$$\beta_{i,i:n} = Var(X_{i:n}) = \alpha_{i:n}^{(2)} - \alpha_{i:n}^2 \quad 1 \leq i \leq n \quad (2.1.2)$$

şeklinde elde edilir.

Şimdi $E(X_{i:n} X_{j:n})$ çarpım momentini $\alpha_{i,j:n}$ olarak gösterelim. O zaman Tanım 1.19'dan çarpım momenti, $1 \leq i < j \leq n$ için

$$\begin{aligned} \alpha_{i,j:n} &= E(X_{i:n} X_{j:n}) = \iint_{-\infty < x < y < \infty} xy f_{i,j:n}(x, y) dy dx \\ &= \frac{n!}{(i-1)!(j-i-1)!(n-j)!} \iint_{-\infty < x < y < \infty} xy \{F(x)\}^{i-1} \\ &\quad \{F(y) - F(x)\}^{j-i-1} \{1-F(y)\}^{n-j} f(x) f(y) dy dx \end{aligned} \quad (2.1.3)$$

elde edilir. Uyum için bazen $\alpha_{i:n}^{(2)}$ ifadesi $\alpha_{i,i:n}$ olarak gösterilebilir. $X_{i:n}$ 'nin ilk iki tekli momentlerinden ve $\alpha_{i,j:n}$ çarpım momentlerinden $X_{i:n}$ ve $X_{j:n}$ 'nin kovaryansı Tanım 1.8'den

$$\beta_{i,j:n} = Cov(X_{i:n}, X_{j:n}) = \alpha_{i,j:n} - \alpha_{i:n} \alpha_{j:n} \quad 1 \leq i < j \leq n \quad (2.1.4)$$

şeklinde elde edilir.

2.2. Sıralı istatistikler için sistematik momentler arasındaki bazı temel bağıntılar

Çalışmamızın bu kısmında sıralı istatistiklerin tekli veya çarpım momenti ve kovaryansı tarafından sağlanan bazı temel bağıntılar verilmiştir. İlk olarak

$$\sum_{i=1}^n X_{i:n}^k = \sum_{i=1}^n X_i^k \quad (2.2.1)$$

ve

$$\sum_{i=1}^n \sum_{j=1}^n X_{i:n} X_{j:n} = \sum_{i=1}^n \sum_{j=1}^n X_i X_j \quad (2.2.2)$$

ifadelerini yazalım. Burada (2.2.1) ve (2.2.2) ifadelerinin her iki tarafının beklenen değeri alınır,

$$\sum_{i=1}^n \alpha_{i:n}^{(k)} = n \alpha_{1:1}^{(k)} = n E(X^k) \quad (2.2.3)$$

ve

$$\sum_{i=1}^n \sum_{j=1}^n \alpha_{i,j:n} = n \alpha_{1:1}^{(2)} + n(n-1) \alpha_{1:1}^2 \quad (2.2.4)$$

eşitlikleri elde edilir. Özel olarak $k = 1$ ve $k = 2$ alınır (2.2.3) ifadesi sırasıyla,

$$\sum_{i=1}^n \alpha_{i:n} = n \alpha_{1:1} = n E(X) \quad (2.2.5)$$

ve

$$\sum_{i=1}^n \alpha_{i:n}^{(2)} = n \alpha_{1:1}^{(2)} = n E(X^2) \quad (2.2.6)$$

eşitliklerine döndürür. (2.2.6) eşitliği (2.2.4) eşitliğinde kullanılırsa

$$\sum_{i=1}^n \sum_{j=i+1}^n \alpha_{i,j:n} = \binom{n}{2} \alpha_{1:1}^2 \quad (2.2.7)$$

eşitliği elde edilir. Burada,

$$\alpha_{1,2:2} = E(X_{1:2} X_{2:2}) = E(X_1 X_2) = \alpha_{1:1}^2 \quad (2.2.8)$$

ifadesinin kullanılmasıyla (2.2.7) ifadesi

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^n \alpha_{i,j:n} = \binom{n}{2} \alpha_{1,2:n} \quad (2.2.9)$$

olarak da gösterilebilir. (2.2.4) ve (2.2.5) ifadelerinden

$$\begin{aligned} \sum_{i=1}^n \sum_{j=1}^n \beta_{i,j:n} &= \sum_{i=1}^n \sum_{j=1}^n \alpha_{i,j:n} - \left(\sum_{i=1}^n \alpha_{i:n} \right) \left(\sum_{j=1}^n \alpha_{j:n} \right) \\ &= n \left(\alpha_{1,1}^{(2)} - \alpha_{1,1}^2 \right) \\ &= n \beta_{1,1:n} = n \text{Var}(X) \end{aligned} \quad (2.2.10)$$

eşitliği elde edilir. Sıfır civarında simetrik olan dağılımlar için $0 < U_{i:n} < 1$ olmak üzere,

$$F^{-1}(U_{i:n}) \stackrel{d}{=} X_{i:n} \quad (2.2.11)$$

ifadesinden $1 \leq i \leq n$ için

$$X_{i:n} \stackrel{d}{=} (-X)_{n-i+1:n} \quad (2.2.12)$$

ve $1 \leq i < j \leq n$ için

$$(X_{i:n}, X_{j:n}) \stackrel{d}{=} ((-X)_{n-j+1:n}, (-X)_{n-i+1:n}) \quad (2.2.13)$$

olduğu açıktır. Bundan dolayı sıfır civarında simetrik olan dağılımlar için $1 \leq i \leq n$ ve $k \geq 1$ iken

$$\alpha_{n-i+1:n}^{(k)} = (-1)^k \alpha_{i:n}^{(k)} \quad (2.2.14)$$

$1 \leq i < j \leq n$ için

$$\alpha_{n-j+1,n-i+1:n} = \alpha_{i,j:n} \quad (2.2.15)$$

ve $1 \leq i \leq j \leq n$ için

$$\begin{aligned} \beta_{n-j+1,n-i+1:n} &= \alpha_{n-j+1,n-i+1:n} - \alpha_{n-j+1:n} \alpha_{n-i+1:n} \\ &= \alpha_{i,j:n} - \alpha_{j:n} \alpha_{i:n} \\ &= \beta_{i,j:n} \end{aligned} \quad (2.2.16)$$

ifadeleri yazılabilir.

III. BÖLÜM

3. SIRALI İSTATİSTİKLERİN BAZI SİSTEMATİK MOMENTLERİNİN YAKLAŞIK HESAPLAMALARI

3.1. Sıralı istatistiklerin bazı sistematik momentlerinin yaklaşık hesaplamalarının incelenmesi

$X_1, X_2, \dots, X_n$ pdf'si f ve cdf'si F olan aynı dağılımlı bağımsız sürekli tesadüfi değişkenler olsun. Verilen bu sürekli tesadüfi değişkenler küçükten büyüğe doğru $X_{1:n} < X_{2:n} < \dots < X_{n:n}$ şeklinde sıralanabilir. Burada $X_{i:n}$ i -inci sıralı istatistik olarak ifade edilir ayrıca standart düzgün dağılımlı sıralı istatistikler de $U_{1:n} < U_{2:n} < \dots < U_{n:n}$ şeklinde belirtilir.

$1 \leq i \leq n$ için $X_{i:n}$ 'nin marjinal olasılık yoğunluk fonksiyonu $-\infty < x_{i:n} < \infty$ olmak üzere

$$f_{i:n}(x_{i:n}) = \frac{n!}{(i-1)!(n-i)!} [F(x_{i:n})]^{i-1} [1 - F(x_{i:n})]^{n-i} f(x_{i:n}) \quad (3.1.1)$$

şeklinde, $1 \leq i < j \leq n$ için $X_{i:n}$ ve $X_{j:n}$ 'nin bileşik olasılık yoğunluk fonksiyonu da $-\infty < x_{i:n} < x_{j:n} < \infty$ olmak üzere

$$f_{i,j:n}(x_{i:n}, x_{j:n}) = \frac{n!}{(i-1)!(j-i-1)!(n-i)!} [F(x_{i:n})]^{i-1} [F(x_{j:n}) - F(x_{i:n})]^{j-i-1} [1 - F(x_{j:n})]^{n-j} f(x_{i:n}) f(x_{j:n}) \quad (3.1.2)$$

şeklinde ifade edilir. Burada f ve F arasında

$$F(x) = \int_{-\infty}^x f(t) dt \quad (3.1.3)$$

bağıntısı mevcuttur.

Sürekli X tesadüfi değişkeninin orjine göre k -ıncı momenti,

$$E(X^k) = \int_{-\infty}^{\infty} x^k f(x) dx \quad (3.1.4)$$

şeklinde yazılabilir. Burada (3.1.1) ve (3.1.4) eşitlikleri kullanılarak

$$E(U_{i:n}^k) = \frac{(i+k-1)(i+k-2)\dots(i+1)i}{(n+k)(n+k-1)\dots(n+2)(n+1)} \quad (3.1.5)$$

elde edilir [10]. Benzer düşünceyle X ve Y sürekli tesadüfi değişkenlerinin orjine göre r -inci ve s -inci çarpım momentleri

$$E(X^r Y^s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^r y^s f(x, y) dx dy \quad (3.1.6)$$

şeklinde yazılabilir. Burada (3.1.2) ve (3.1.6) kullanılarak

$$E(U_{i:n}^r U_{j:n}^s) = \frac{i(i+1) \dots (i+r-1)(j+r)(j+r+1) \dots (j+r+s-1)}{(n+1)(n+2) \dots (n+r)(n+r+1)(n+r+2) \dots (n+r+s)} \quad (3.1.7)$$

elde edilir. Burada,

$$\mu = E(U_{i:n}) = \frac{i}{n+1} \quad (3.1.8)$$

ifadesi (3.1.5) ve (3.1.7) eşitliklerinden ayrı ayrı elde edilebilir. Ek olarak sürekli X tesadüfi değişkeninin varyansı

$$Var(X) = E(X^2) - [E(X)]^2 \quad (3.1.9)$$

ve X, Y sürekli tesadüfi değişkenlerinin kovaryansı

$$Cov(X, Y) = E(XY) - E(X)E(Y) \quad (3.1.10)$$

şeklinde ifade edilebilir. Burada (3.1.5) ve (3.1.9) kullanılarak

$$Var(U_{i:n}) = \frac{i(n-i+1)}{(n+2)(n+1)^2} \quad (3.1.11)$$

ifadesini (3.1.7) ve (3.1.10) kullanılarak

$$Cov(U_{i:n}, U_{j:n}) = \frac{i(n-j+1)}{(n+2)(n+1)^2} \quad (3.1.12)$$

ifadesini elde ederiz[11]. (3.1.8), (3.1.11) ve (3.1.12) denklemleri $U_{i:n}$ 'nin sadece kendisine değil $U_{i:n}$ 'nin anakütle büyüklüğüne ve sırasına bağlıdır. Bu sıralı istatistiklerin koşulsuz dağılıma ve parametrik olmayan istatistiklere bağlı olduğunu gösterir.

Düzgün standart dağılımlı i -inci sıralı istatistiğin tam momentleri (3.1.5) denkleminde i -inci ve j -inci sıralı istatistiğin tam momentleri (3.1.7) denkleminde kolayca hesaplanabilir. Fakat herhangi sürekli bir dağılımdan sıralı istatistiklerin tam momentlerinin hesaplanması birçok sıkıcı sayısal integrasyonu içerir.

Olasılık-İntegral dönüşümü:

$$F(X_{i:n}) = U_{i:n} \quad (3.1.13)$$

şeklinde $U_{i:n}$ sıralı istatistiği $X_{i:n}$ 'ye dönüştür. Bu sebepten (3.1.13) ifadesi herhangi sürekli bir dağılımdan sıralı istatistiklerin momentlerini hesaplamayı kolaylaştırır.

3.2. i-inci sıralı istatistiğin ortalaması ve varyansının yaklaşık hesabı

Genellikle n 'nin küçük değerleri için momentlerin karakteristik yapıları tamamen incelenmiştir. n 'nin büyük değerleri için herhangi sürekli bir dağılımdan sıralı istatistiklerin tam momentlerinin hesabı çok zor ve oldukça sıkıcı yöntemler içerir. Bazı durumlarda imkansız bile olabilir. Bundan dolayı sıralı istatistiklerin momentlerini elde etmek için yaklaşımlar kaçınılmazdır. Bu sebepten (3.1.13) denkleminin tersini alarak

$$X_{i:n} = F^{-1}(U_{i:n}) \quad (3.2.1)$$

elde ederiz. Eğer herhangi bir $g(z)$ (3.1.8) noktası etrafında Taylor serisine açılırsa

$$g(z) = g(\mu) + \sum_{i=1}^{\infty} \frac{(z-\mu)^i}{i!} g^{(i)}(\mu) \quad (3.2.2)$$

elde ederiz. $\mu < z_1 < z$ için $\lim_{n \rightarrow \infty} \frac{(z-\mu)^n}{n!} g^{(n)}(z_1) = 0$ olursa (3.2.2) ifadesi yakınsak olur. (3.2.2) ifadesinin beklenen değerini alırsak aşağıdaki denklem elde edilir.

$$E[g(Z)] = g(\mu) + \frac{Var(Z)}{2!} g^{(2)}(\mu) + \sum_{i=3}^{\infty} \frac{E[(Z-\mu)^i]}{i} g^{(i)}(\mu) \quad (3.2.3)$$

İlave olarak (3.2.2) ifadesinin varyansı (3.1.5) ve (3.1.9) denklemlerinden

$$Var[g(Z)] = Var(Z) \left[g^{(1)}(\mu) \right]^2 - \frac{1}{4} \left[g^{(2)}(\mu) \right]^2 Var^2(Z) + E[h(Z)] \quad (3.2.4)$$

şeklinde ifade edilir. Burada $E[h(Z)]$ ifadesi Z 'nin üçüncü ve daha yüksek momentlerini içerir. (3.2.1) denklemi,

$$g(u_{i:n}) = x_{i:n} = F^{-1}(u_{i:n}) \quad (3.2.5)$$

şeklinde yazılabilir. Ayrıca (3.2.3) ve (3.2.5) denklemlerini kullanarak aşağıdaki ifadeleri elde edebiliriz.

$$E(X_{i:n}) \approx g(\mu) + \frac{Var(U_{i:n})}{2} g^{(2)}(\mu) \quad (3.2.6)$$

ve

$$Var(X_{i:n}) \approx Var(U_{i:n}) \left[g^{(1)}(\mu) \right]^2 - \frac{1}{4} \left[g^{(2)}(\mu) \right]^2 Var^2(U_{i:n}) \quad (3.2.7)$$

Burada (3.2.6) ve (3.2.7) denklemleri g 'nin türevlerine bağlıdır.

Elde edilen bu denklemlerde g 'nin birinci ve ikinci türevleri (3.2.5) kullanılarak aşağıdaki gibi hesaplanabilir.

$$g^{(1)}(\mu) = \left. \frac{dx_{i:n}}{du_{i:n}} \right|_{u_{i:n}=\mu} = \left. \frac{1}{f(x_{i:n})} \right|_{x_{i:n}=F^{-1}(\mu)} \quad (3.2.8)$$

ve

$$g^{(2)}(\mu) = \left. \frac{d}{du_{i:n}} \left[\frac{1}{f(x_{i:n})} \right] \right|_{x_{i:n}=F^{-1}(\mu)} = \left. \frac{-f^{(1)}(x_{i:n})}{[f(x_{i:n})]^3} \right|_{x_{i:n}=F^{-1}(\mu)} \quad (3.2.9)$$

(3.2.8) ve (3.2.9) değerleri (3.2.6) ve (3.2.7) de yerine yazılırsa $X_{i:n}$ 'nin ortalaması ve varyansı için yaklaşım,

$$E(X_{i:n}) \approx x_{i:n} - \frac{1}{2} Var(U_{i:n}) \left. \frac{f^{(1)}(x_{i:n})}{[f(x_{i:n})]^3} \right|_{x_{i:n}=F^{-1}(\mu)} \quad (3.2.9)$$

ve

$$Var(X_{i:n}) \approx \frac{Var(U_{i:n})}{[f(x_{i:n})]^2} - \frac{1}{4} Var^2(U_{i:n}) \left. \frac{[f^{(1)}(x_{i:n})]^2}{[f(x_{i:n})]^6} \right|_{x_{i:n}=F^{-1}(\mu)} \quad (3.2.10)$$

şeklinde elde edilir[10]. Burada (3.2.9) ve (3.2.10) denklemleri f 'nin türevlerine bağlıdır.

3.3. Uygulamalar

Standart düzgün, üstel ve lojistik dağılımlar ile $X_{i:n}$ 'nin ortalama ve varyansı aşağıdaki ifadeler sırasıyla kullanılarak (3.2.9) ve (3.2.10)'dan elde edildi.

$$f(x) = 1 \quad 0 < x < 1 \quad \text{ve} \quad F^{-1}(\mu) = \mu, \quad (3.3.1)$$

$$f(x) = e^{-x} \quad 0 \leq x < \infty \quad \text{ve} \quad F^{-1}(\mu) = -\ln(1 - \mu), \quad (3.3.2)$$

ve

$$f(x) = \frac{e^{-x}}{(1 + e^{-x})^2} \quad -\infty < x < \infty \quad \text{ve} \quad F^{-1}(\mu) = \ln(\mu) - \ln(1 - \mu), \quad (3.3.3)$$

(3.3.1)-(3.3.3) denklemlerindeki $F^{-1}(\mu)$, $0 < \mu < 1$ için F 'nin ters fonksiyonudur. Elde edilen yaklaşık hesaplamalar kullanılarak (3.3.1)-(3.3.3) denklemlerinin i -inci sıralı istatistiklerinin beklenen değer ve varyanslarının sonuçları QBASIC 4.5 programından elde edilen Tablo 1'de verilmiştir.

Tablo 1

i	n		Düzgün	Üstel	Lojistik
1	4	$E(X_{i:n})$ $Var(X_{i:n})$	.2 $2.666667E-02$	.2439769 $4.123264E-02$	-1.698794 .9440105
2	9	$E(X_{i:n})$ $Var(X_{i:n})$	.2 $1.454545E-02$	.2345072 $2.259814E-02$	-1.556749 .5391271
3	14	$E(X_{i:n})$ $Var(X_{i:n})$	.2 .01	.230956 $1.556396E-02$	-1.556749 .5391271
4	19	$E(X_{i:n})$ $Var(X_{i:n})$	.2 $7.619048E-03$	.2290959 $1.186933E-02$	-1.47558 .2896471
5	24	$E(X_{i:n})$ $Var(X_{i:n})$	.2 $6.153846E-03$	.2279512 $9.59227E-03$	-1.45841 .235184
6	29	$E(X_{i:n})$ $Var(X_{i:n})$	.2 $5.161291E-03$	.2271758 $8.048257E-03$	-1.44778 .1979546
7	34	$E(X_{i:n})$ $Var(X_{i:n})$	.2 $4.44444E-03$	.2266158 $6.932388E-03$	-1.438378 .1708985
8	39	$E(X_{i:n})$ $Var(X_{i:n})$	.2 $3.902439E-03$	.2261923 $6.088266E-03$	-1.432026 .1503476
9	44	$E(X_{i:n})$ $Var(X_{i:n})$	.2 $3.478261E-03$	.2258609 $5.427398E-03$	-1.427055 .1342081
10	49	$E(X_{i:n})$ $Var(X_{i:n})$	.2 $3.137255E-03$	.2255945 $4.895953E-03$	-1.423059 .1211974
14	69	$E(X_{i:n})$ $Var(X_{i:n})$	.2 $2.253521E-03$	.2249041 $3.518027E-03$	-1.412703 $8.733077E-02$
18	89	$E(X_{i:n})$ $Var(X_{i:n})$	.2 $1.758242E-03$	.2245172 $2.745366E-03$	-1.406899 $6.825678E-02$
22	109	$E(X_{i:n})$ $Var(X_{i:n})$	.2 $1.441441E-03$	.2242697 $2.250984E-03$	-1.403186 $5.602098E-02$
30	149	$E(X_{i:n})$ $Var(X_{i:n})$	.2 $1.059603E-03$	.2239714 $1.654944E-03$	-1.398712 $4.123654E-02$
35	174	$E(X_{i:n})$ $Var(X_{i:n})$	.2 $9.090909E-04$	.2238538 $1.41995E-03$	-1.396948 $3.539787E-02$
40	199	$E(X_{i:n})$ $Var(X_{i:n})$	.2 $7.960199E-04$	.2237654 $1.243394E-03$	-1.395623 $3.100751E-02$
80	399	$E(X_{i:n})$ $Var(X_{i:n})$	.2 $3.990025E-04$	.2234553 $6.233442E-04$	-1.39097 $1.556417E-02$
100	499	$E(X_{i:n})$ $Var(X_{i:n})$	.2 $3.193613E-04$	.2233931 $4.989397E-04$	-1.390037 $1.246104E-02$
140	699	$E(X_{i:n})$ $Var(X_{i:n})$	.2 $2.282454E-04$	.2233219 $3.566016E-04$	-1.388969 $8.908681E-03$
180	899	$E(X_{i:n})$ $Var(X_{i:n})$	.2 $1.775805E-04$	.2232823 $2.774502E-04$	-1.388375 $6.932407E-03$
200	999	$E(X_{i:n})$ $Var(X_{i:n})$	.2 $1.508402E-04$	.2232684 $2.497346E-04$	-1.388168 $6.240248E-03$

3.4. Sonuç

Bu yaklaşım metodu n 'nin çok büyük değerleri için, ekstrem değerlerde yavaş işleyebilir. Bu sıkıntıyı aşmak için ekstrem değer dağılımları olarak bilinen Fréchet, Weibull ve Gumbell dağılımlarını kullanabiliriz. Bu çalışmada standart üstel dağılım ve düzgün dağılım, yaklaşım metodunun yakınsaklığını kontrol etmek için verildi. Bundan başka standart düzgün dağılımın i -inci sıralı istatistiğinin ortalama ve varyansı tam olarak belirlendi ve sırasıyla (3.1.8) ve (3.1.11) ile verildi.

Aynı zamanda, standart üstel dağılımlı i -inci sıralı istatistiğin ortalama ve varyansı da tam olarak sırasıyla

$$E(X_{i:n}) = \sum_{l=1}^i \frac{1}{n-l+1} \quad (3.4.1)$$

ve

$$Var(X_{i:n}) = \sum_{l=1}^i \frac{1}{(n-l+1)^2} \quad (3.4.2)$$

olarak bilinir.

Standart düzgün ve üstel dağılımların i -inci sıralı istatistiğinin ortalama ve varyansının yaklaşık sonuçları sırasıyla (3.1.8), (3.1.11) ve (3.4.1), (3.4.2) ifadelerinden elde edilen sonuçlara çok yakındır.

Standart normal dağılımlı i -inci sıralı istatistiğin tam momentleri de n 'nin küçük değerleri için elde edilebilir.

KAYNAKLAR

- [1] Akdeniz, Fikri, 2002, **Olasılık ve İstatistik**, Baki Kitabevi, Adana.
- [2] Rees, D.G., 1987, **Foundations of Statistics**, J.W. Arrowsmith Ltd., Bristol.
- [3] Hogg, R.V. and Craig, A.T., 1978, **Introduction to Mathematical Statistics**, Macmillan Publishing Co., Inc., New York.
- [4] Aytaç, M., 1994, **Matematiksel İstatistik**, Uludağ Üniversitesi Basımevi, Bursa.
- [5] Balakrishnan, N., and Basu, A.P., 1995, **The Exponential Distribution**, Gordon and Breach Science Publishers S.A. Netherlands.
- [6] Balakrishnan, N., and Cohen, A.C., 1991, **Order Statistics and Inference**, Academic Press, Inc., London.
- [7] Reiss, R.D., 1989, **Approximate Distributions of Order Statistics**, Springer, Verlag, New York Inc., USA.
- [8] David, H.A., 1981, **Order Statistics**, John Wiley and Sons Inc., USA.
- [9] Galambos, J., 1987, **The Asymtotic Theory of Extreme Order Statistics**, 2nd ed., Malabar, Florida, Krieger.
- [10] Güngör, Mehmet., **Approximate Calculation of Some Moments of Order Statistics**, Jour. of Inst. of Math. & Comp. Sci. (Math. Ser.), Vol. 13, No. 2 (2000) 135-140.
- [11] Güngör, Mehmet., **David-Johnson Approximation for Calculation of Some Moments of Order Statistics**, Jour. of Inst. of Math. & Comp. Sci. (Math. Ser.), Vol. 13, No. 2 (2000) 239-245.

ÖZGEÇMİŞ

1979 yılında Bursa’da doğmuşum. İlk ve orta öğrenimi Malatya’da tamamladıktan sonra 1997 yılında Fırat Üniversitesi Fen-Edebiyat Fakültesi Matematik Bölümünü kazandım. 2001 yılında bu bölümden mezun oldum. 11 Eylül 2001 tarihinde M.E.B’ de matematik öğretmeni olarak göreve başladım. Halen aynı görevi sürdürmekteyim. 2002 yılında Fen Bilimleri Enstitüsünün açmış olduğu Matematik Anabilim Dalının Uygulamalı Matematik Programında tezli yüksek lisansa başladım.

Gökhan GÖKDERE