Afyon Kocatepe Üniversitesi
FEN BİLİMLERİ DERGİSİ

5 (1-2)
71-76

Afyon Kocatepe University
JOURNAL OF SCIENCE

SÜREKLİ DAĞILIMLAR AİLESİ İÇİN İKİ YANLI İNVARYANT GÜVEN ARALIKLARININ SEVİYELERİNİN TAHMİNİ

Mehmet Fedai KAYA, Buğra SARAÇOĞLU

Selçuk Üniversitesi, Fen-Edebiyat Fakültesi, İstatistik Bölümü, KONYA

ÖZET

$\Im$, dağılım fonksiyonlarının bir sınıfı olmak üzere $X_1, X_2, \ldots, X_n$ dağılımı $F \in \Im$ 'den olan bir örneklem olsun.
f_1 ve f_2, $\forall \underline{u} = (u_1, u_2, \ldots, u_n) \in R^n$ için
$f_1(u_1, u_2, \ldots, u_n) \leq f_2(u_1, u_2, \ldots, u_n)$
özelliğine sahip iki Borel ölçülebilir fonksiyon olmak üzere X_{n+1} yukarıdaki örneklemden bağımsız ve aynı F dağılım fonksiyonuna sahip yeni bir rasgele değişken olsun.
$\forall F \in \Im$ için $P\{X_{n+1} \in (f_1(X_1, X_2, \ldots, X_n), f_2(X_1, X_2, \ldots, X_n))\} = \alpha$ ise $(f_1(X_1, X_2, \ldots, X_n), f_2(X_1, X_2, \ldots, X_n))$ rasgele aralığına $\Im$ sınıfı için ana kitleyi kapsayan α seviyeli invaryant güven aralığı denir.
Bu çalışmada invaryant güven aralıklarının seviyelerinin tahmini örneklem dağılım fonksiyonu yardımıyla elde edilmiştir.

Anahtar Sözcükler: İnvaryant Güven Aralıklar, Örneklem Dağılım Fonksiyonu.

FOR CONTINIOUS DISTRIBUTIONS FAMILY TWO SIDED ESTIMATION OF LEVELS OF INVARIANT CONFIDENCE INTERVALS

ABSTRACT

Let $X_1, \ldots, X_n$ be a sample from a distribution function with $F \in \Im$, where $\Im$ is some class of distribution functions. Suppose f_1 and f_2 are two Borel functions with the following property

$$f_1(u_1, \ldots, u_n) \leq f_2(u_1, \ldots, u_n) \qquad \forall \underline{u} = (u_1, \ldots, u_n) \in R^n$$

Let X_{n+1} be a new random variable from F and be independent of $X_1,\ldots,X_n$. If

$P\{X_{n+1}\in(f_1(X_1,X_2,\ldots,X_n),f_2(X_1,X_2,\ldots,X_n))\}=\alpha$, $\forall F\in\Im$

then $(f_1(X_1,X_2,\ldots,X_n),f_2(X_1,X_2,\ldots,X_n))$ is called an invariant

confidence interval containing the main mass for class of distributions $\Im$ with confidence level α.

In this study estimation of levels of invariant confidence intervals containing the main mass are obtained using empirical distribution function.

Key Words: Invariant Confidence İntervals, Emprical Distribution Function.

1. Giriş

$\Im$, dağılım fonksiyonlarının bir sınıfı ve $X_1,X_2,\ldots,X_n$ de dağılımı $F\in\Im$ olan bir örneklem olsun. f_1 ve f_2,

$$\forall \underline{u}=(u_1,u_2,\ldots,u_n)\in R^n \text{ için } f_1(u_1,u_2,\ldots,u_n)\le f_2(u_1,u_2,\ldots,u_n) \qquad (1)$$

özelliğine sahip iki Borel ölçülebilir fonksiyon olmak üzere $X_1,X_2,\ldots,X_n$ lerden bağımsız ve aynı F dağılımına sahip yeni bir X_{n+1} rasgele değişkeninin çekildiğini düşünelim.

$\forall F\in\Im$ için $P\{X_{n+1}\in(f_1(X_1,X_2,\ldots,X_n),f_2(X_1,X_2,\ldots,X_n))\}=\alpha$ oluyorsa $(f_1(X_1,X_2,\ldots,X_n),f_2(X_1,X_2,\ldots,X_n))$ rasgele aralığına $\Im$ sınıfı için ana kitleyi kapsayan α seviyeli invaryant güven aralığı denir. $\Im_c$, bütün sürekli dağılım fonksiyonlarının bir sınıfı ve $X_1,X_2,\ldots,X_n$ dağılımı $F\in\Im_c$'den olan bir örneklem, $X_{(1)}<X_{(2)}<\ldots<X_{(n)}$ ise bu örneklemden elde edilen sıra istatistikleri olsun. $f_1(x_1,x_2,\ldots,x_n)$ ve $f_2(x_1,x_2,\ldots,x_n)$, (1) özelliğine sahip sürekli, simetrik ve Lebesque anlamında sıfır ölçülü kümeler dışında farklı değerler alan iki fonksiyon olmak üzere,

$$(f_1(X_1,X_2,\ldots,X_n),f_2(X_1,X_2,\ldots,X_n))$$

aralığının ana kitleyi kapsayan invaryant güven aralık olması için gerek ve yeter koşul;

$f_1(X_1, X_2,..., X_n) = X_{(i)}$ ve $f_2(X_1, X_2,..., X_n) = X_{(j)}$ $\quad ,1 \le i < j \le n$
olmasıdır ve seviyesi aşağıdaki gibidir. (I. Bairamov ve Y.I. Petunin, 1990).

$$P\{X_{n+1} \in (X_{(i)} X_{(j)})\} = \frac{j-i}{n+1}$$

2. İNVARYANT GÜVEN ARALIKLARIN SEVİYESİNİN TAHMİNİ

$X_1, X_2,..., X_n$ ve $Y_1, Y_2,..., Y_m$ birbirinden bağımsız, aynı $F \in \Im$ dağılım fonksiyonuna sahip iki örneklem

$$(f_1(X_1, X_2,...X_n), f_2(X_1, X_2,...X_n))$$

rasgele aralığı $\Im$ sınıfı için ana kitleyi kapsayan $\alpha = P\{Y_1 \in (f_1(X_1, X_2,...X_n), f_2(X_1, X_2,...X_n))\}$ seviyeli invaryant güven aralığı olsun.

$$P\{Y_1 \in (f_1(X_1, X_2,..., X_n), f_2(X_1, X_2,..., X_n))\}$$
$$= F(f_2(x_1, x_2,..., x_n)) - F(f_1(x_1, x_2,..., x_n))$$

şeklindedir. Burada F yerine, $Y_1, Y_2,..., Y_m$ ' in örneklem dağılım fonksiyonu yazılırsa

$$\hat{\alpha} = F_m^*(f_2(X_1, X_2,..., X_n)) - F_m^*(f_1(X_1, X_2,..., X_n))$$

tahmin edicisi elde edilir. F_m^*, F için yansız bir tahmin edici olduğundan $\hat{\alpha}$ da α için yansız bir tahmin edicidir.

$$\hat{\alpha} = \mathrm{F}_{m}^{*}(X_{(s)}) - \mathrm{F}_{m}^{*}(X_{(r)}) =$$

$$\begin{cases} 0 \quad , & Y_{(i)} < f_1(X_1,\dots,X_n) < Y_{(i+1)} \,,\, Y_{(i)} < f_2(X_1,\dots,X_n) < Y_{(i+1)} \\ \dfrac{k}{m} \quad , & Y_{(i)} \le f_1(X_1,\dots,X_n) < Y_{(i+1)} \,,\, Y_{(i+k)} < f_2(X_1,\dots,X_n) < Y_{(i+k+1)} \\ 1 \quad , & f_1(X_1,\dots,X_n) \ge Y_{(1)} \quad , \quad f_2(X_1,\dots,X_n) > Y_{(m)} \end{cases}$$

şeklindedir.

Burada $Y_{(m+1)} = \infty$, $Y_{(0)} = -\infty$, $k = 1,\dots,m$, $i = 0,\dots,m-k$ ' dır.

$\mathrm{F}_{m}^{*}(f_2(X_1, X_2,\dots, X_n)) - \mathrm{F}_{m}^{*}(f_1(X_1, X_2,\dots, X_n))$ rasgele değişkeni $0, \frac{1}{m}, \frac{2}{m},\dots,\frac{m-1}{m}, 1$ değerlerini;

$$P\left\{\mathrm{F}_{m}^{*}(f_2(X_1, X_2,\dots, X_n)) - \mathrm{F}_{m}^{*}(f_1(X_1, X_2,\dots, X_n)) = \frac{k}{m}\right\}$$

$$= \sum_{i=0}^{m-k} P\left\{Y_{(i)} \le f_1(X_1,\dots,X_n) < Y_{(i+1)} \,,\, Y_{(i+k)} < f_2(X_1,\dots,X_n) < Y_{(i+k+1)}\right\}, \quad k = 0,\dots,m$$

olasılıklarıyla alan bir rasgele değişkendir. Özel olarak $f_1(X_1, X_2,\dots, X_n) = X_{(r)}$, $f_2(X_1, X_2,\dots, X_n) = X_{(s)}$ (r<s) alınırsa

$$P\left\{\mathrm{F}_{m}^{*}(X_{(s)}) - \mathrm{F}_{m}^{*}(X_{(r)}) = \frac{k}{m}\right\} =$$

$$= \sum_{i=0}^{m-k} P\left\{Y_{(i)} \le X_{(r)} < Y_{(i+1)} \,,\, Y_{(i+k)} < X_{(s)} < Y_{(i+k+1)}, \; k = 0,\dots,m\right\}$$

olur. O halde $\hat{\alpha} = \mathrm{F}_{m}^{*}(X_{(s)}) - \mathrm{F}_{m}^{*}(X_{(r)})$ rasgele değişkeninin olasılık fonksiyonu;

$$f_{F^*_m(X_{(s)})-F^*_m(X_{(r)})}(x)=$$

$$=\begin{cases} \dfrac{n!(m+n-s)!}{(n-s)!(m+n)!}+A & , \; x=0 \\[2ex] \dfrac{mn!(m+n-s-1)!(s-r)}{(n-s)!(n+m)!}+B & , \; x=\dfrac{1}{m} \\[2ex] \dfrac{n!m!(m+n-s-k)!(s+k-r-1)!}{k!(s-r-1)!(n-s)!(m-k)!(m+n)!}+C+D & , \; x=\dfrac{k}{m} \\[2ex] \dfrac{n!(m+s-r-1)!}{(s-r-1)!(m+n)!} & , \; x=1 \end{cases} \qquad k=2,3,\ldots,m-1$$

şeklindedir. Burada A,B,C,D aşağıdaki gibidir.

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

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

$$C=\sum_{i=1}^{m-k-1}\frac{rn!m!(m+n-s-i-k)!(k+s-r-1)!}{k!(s-r-1)!(n-s)!(m-k-i)!(m+n-i+1)!}$$

$$D=\frac{n!m!(m+r-k-1)!(s+k-r-1)!}{k!(m-k)!(r-1)!(m+n)!}$$

Örnek 1. X_1, X_2, X_3, X_4, $F\in \Im_c$ dağılımına sahip örneklem , Y_1, Y_2, Y_3 ' de aynı F dağılımına sahip X_1, X_2, X_3, X_4 örnekleminden bağımsız başka bir örneklem olsun. $(f_1(X_1,X_2,X_3,X_4), f_2(X_1,X_2,X_3,X_4))$ rasgele aralığı $\alpha = P\{Y_1 \in (f_1(X_1,X_2,\ldots X_n), f_2(X_1,X_2,\ldots X_n))\}$ seviyeli invaryant güven aralığı olsun.

$$f_1(X_1,X_2,\ldots,X_n)=X_{(2)}\;,\; f_2(X_1,X_2,\ldots X_n)=X_{(3)}$$

alınırsa

$$\alpha = P\{Y_1 \in (X_{(2)}, X_{(3)})\} = \frac{3-2}{4+1} = \frac{1}{5}$$

olur.

$$E(\hat{\alpha}) = E\left[F_m^*(X_{(3)}) - F_m^*(X_{(2)})\right]$$

$$= 0\left(\frac{4}{35}+\frac{6}{35}+\frac{6}{35}+\frac{4}{35}\right)+\frac{1}{3}\left(\frac{3}{35}+\frac{4}{35}+\frac{3}{35}\right)+\frac{2}{3}\left(\frac{2}{35}+\frac{2}{35}\right)+\frac{1}{35}=\frac{1}{5}$$

Örnek 2. $X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8$, $F \in \Im_c$ dağılımına sahip örneklem, Y_1, Y_2, Y_3, Y_4 F dağılımına sahip ilk örneklemden bağımsız başka bir örneklem olmak üzere $(f_1(X_1, X_2, \ldots X_8), f_2(X_1, X_2, \ldots X_8))$ rasgele aralığı
$\alpha = P\{Y_1 \in (f_1(X_1, X_2, \ldots X_8), f_2(X_1, X_2, \ldots X_8))\}$ seviyeli invaryant güven aralığı olsun.

$$f_1(X_1, X_2, \ldots X_8) = X_{(2)}\ ,\ f_2(X_1, X_2, \ldots X_8) = X_{(4)}$$

$$\alpha = P\{Y_1 \in (X_{(2)}, X_{(4)})\} = \frac{4-2}{8+1} = \frac{2}{9}$$

$$E(\hat{\alpha}) = E\left[F_m^*(X_{(4)}) - F_m^*(X_{(2)})\right]$$

$$= 0\left(\frac{14}{99}+\frac{3}{11}+\frac{1}{99}\right)+\frac{1}{4}\left(\frac{14}{99}+\frac{2}{11}+\frac{8}{495}\right)$$

$$+\frac{2}{4}\left(\frac{1}{11}+\frac{2}{33}+\frac{1}{55}\right)+\frac{3}{4}\left(\frac{4}{99}+\frac{8}{495}\right)+\frac{1}{99}=\frac{2}{9}$$

Yukarıdaki örneklerden de görüldüğü gibi $E(\alpha^*) = \alpha$ bulunmuştur.

KAYNAKLAR

1. Bairamov, I. G., Petunin Yu. I., Structure of Invariant Confidence Intervals Containing The Main Distributed Mass, Theor. Probab. Appl., 35:1, 15-26, (1990).
2. David H. A., Order Statistics. John Wiley & Sons, Inc., New York, (1970)