# Súlyozott kvantilis korreláció teszt logisztikus eloszláscsaládra

Osztényiné Krauczi Éva
TMAT, Gamf Kar, Kecskeméti Főiskola

**Összefoglalás**: Az aszimptotikus viselkedését vizsgáljuk a súlyozott kvantilis korreláció tesztnek a skála-eltolásos logisztikus eloszláscsalád esetében. Valamint bemutatjuk a tesztek számos alternatívával szembeni erejét.

**Kulcsszavak:** logisztikus eloszlás, határeloszlástételek, illeszkedésvizsgálat, erővizsgálat.

## 1. Bevezetés

A logisztikus eloszlást, mint növekedési görbét Verhulst vezette be populáció dinamikai vizsgálati kapcsán a 19. században [5]-ben. A 20. század végén Balakrishnan pedig egy teljes könyvet szentelt ennek az eloszlásnak [1]. A célja ennek a cikknek, hogy bemutassa a súlyozott kvantilis korreláció tesztet erre az eloszláscsaládra.

Adott $X_1, X_2 \ldots, X_n$ véletlen minta egy ismeretlen $F(x)$ eloszlásfüggvényű véletlen változóból. Döntsük el a minta alapján, igaz-e az az összetett null-hipotézis

$$H_0 : F \text{ logisztikus eloszlású,}$$

vagyis a logisztikus eloszláscsaládból származik-e a minta.

Del Barrio, Cuesta-Albertos, Matrán és Rodríguez-Rodríguez [3] által bevezetett Wasserstein távolságot használó, Csörgő Sándor [2] által javított súlyozott általános tesztet fogjuk használni, hogy döntsünk az összetett nullhipotézis felől.

## 2. Elméleti eredmények

Legyen $G(x) = \dfrac{1}{1+e^{-x}}$ minden valós $x$-re a logisztikus eloszlásfüggvény, a hozzá tartozó kvantilis függvényt $Q_G(t) = G^{-1}(t), 0 < t < 1$ jelöli. A súlyfüggvényt $w(t)$ jelöli, ami ezen eloszlás esetében $w(t) = 6t(1-t), 0 < t < 1$; $G(x)$-nek a hozzá tartozó súlyozott várható értékét $\mu_1(G, w)$ és a súlyozott szórásnégyzetét $\nu(G, w)$ jelöli.

Az adott $X_1, X_2 \ldots, X_n$ véletlen mintához tartozó empirikus eloszlásfüggvényt $F_n(x)$ minden valós $x$-re és empirikus kvantilis függvényt $Q_n(t), 0 < t < 1$ jelöli.

Ekkor az eltolás és skálamentes teszt statisztika a fenti null-hipotézisre

$$V_n = 1 - \frac{\left[\int_0^1 Q_n(t) Q_G(t) w(t) dt - \mu_1(G, w) \int_0^1 Q_n(t) w(t) dt\right]^2}{\nu(G, w)\left[\int_0^1 Q_n^2(t) w(t) dt - \left(\int_0^1 Q_n(t) w(t) dt\right)^2\right]}.$$

A következő tétellel sikerült a teszt statisztika aszimptotikus viselkedését jellemezni.

1. **Tétel.**
*Ha teljesül a null-hipotézis, akkor*

$$nV_n \xrightarrow{D} V,$$

*ahol*

$$V = \frac{1}{\frac{\pi^2}{3} - 2}\left\{\int_0^1 \frac{6B^2(t)}{t(1-t)}dt - \left[\int_0^1 6B(t)dt\right]^2\right\} - \left[\frac{1}{\frac{\pi^2}{3} - 2}\int_0^1 6B(t)\ln\left(\frac{t}{1-t}\right)dt\right]^2$$

*a határeloszlás és* $B(.)$ *a Brown hidat jelöli. Ráadásul a határeloszlást független standard normális véletlen változók négyzetének végtelen lineáris kombinációjakét elő tudjuk állítani*

$$V = \frac{1}{\frac{\pi^2}{3} - 2}\sum_{m=1}^{\infty}\frac{6}{(m+1)(m+2)}Z_m^2 - \left[\frac{1}{\frac{\pi^2}{3} - 2}\sum_{k=0}^{\infty}\frac{3\sqrt{4k+5}}{(k+1)(k+2)(2k+1)(2k+3)}Z_{2k+1}\right]^2.$$

## 3. A új teszt és ereje

A teszt kritikus értékeit és erejét Monte-Carlo módszerrel határoztuk meg. Az alábbi táblázat tartalmazza a kritikus értékeket.

| **n** | **0,85** | **0,90** | **0,95** | **0,99** |
|---|---|---|---|---|
| **20** | 2,07 | 2,34 | 2,83 | 4,02 |
| **50** | 2,21 | 2,49 | 2,99 | 4,17 |
| **100** | 2,24 | 2,52 | 2,99 | 4,13 |
| **200** | 2,24 | 2,52 | 2,99 | 4,14 |
| **500** | 2,23 | 2,51 | 2,97 | 4,06 |
| ∞ | 2,22 | 2,49 | 2,95 | 4,02 |

**1. táblázat.** Empirikus kritikus értékek a $nV_n$ teszt statisztikára és a $V$ határeloszlásra.

Számos alternatívával szembeni erejét pedig az alábbi táblázat mutatja.

| **Alternatívák** | **20** | **50** | **100** | **20** | **50** | **100** |
|---|---|---|---|---|---|---|
| **N(0,1)** | 5 | 6 | 8 | 2 | 2 | 4 |
| **Egyenletes** | 13 | 47 | 93 | 5 | 29 | 82 |
| **Cauchy** | 88 | 99 | * | 84 | 99 | * |
| **Laplace** | 26 | 39 | 55 | 17 | 29 | 43 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Exp(1)** | 70 | 99 | * | 56 | 97 | * |
| **Triangle(1)** | 4 | 7 | 13 | 2 | 3 | 6 |
| **Triangle(2)** | 21 | 61 | 97 | 11 | 43 | 91 |
| **Beta(2,2)** | 6 | 15 | 40 | 2 | 7 | 24 |
| **Weibull(2)** | 12 | 25 | 54 | 5 | 15 | 38 |
| **Gamma(2,1)** | 40 | 81 | 99 | 27 | 69 | 98 |
| **Lognormal** | 86 | * | * | 79 | * | * |
| **Student(5)** | 16 | 19 | 21 | 10 | 12 | 13 |
| $\chi^2(1)$ | 94 | * | * | 88 | * | * |
| **Negatív Exp** | 69 | 99 | * | 56 | 97 | * |

**2. táblázat.** Becsült ereje a 10%-os (2., 3. és 4. oszlop) és az 5%-os (5., 6. és 7. oszlop) $nV_n$ teszt statisztikának, az első oszlopban jelölt alternatívákkal szemben, n=20, 50, 100 mintaméret mellett ( * a 100%-os erőt jelöli).

Általában az mondható el a tesztről, hogy könnyen számolható, lényegében bármelyik véges mintaméretnél használhatók az aszimptotikus kritikus értékek és az ereje viszonylag jó más logisztikus eloszláscsalád tesztekhez képest.

## Irodalomjegyzék

[1] Balakrishnan, N., Handbook of the logistic distribution, Marcel Dekker, New York. (1992)

[2] Csőrgö, S., Weighted correlation tests for location-scale families, Mathematical and Computer Modelling, 38, 753--762. (2003)

[3] del Barrio, E., J.A. Cuesta-Albertos, C. Matrán and J.M. Rodríguez-Rodríguez, Tests of goodness of fit based on the $L_2$-Wasserstein distance, The Annals of Statistics, 27, 1230--1239. (1999)

[4] Balogh, F., É. Oszténо yiné Krauczi.: Weighted quantile correlation test for the logistic family, ACTA, (submitting)

[5] Verhulst, P-F., Notice sur la loi que la population poursuit dans son accroissement, Correspondance mathématique et physique 10, 113--121. (1838)

## Szerző

Oszténуiné Krauczi Éva: TMAT, GAMF Kar, Kecskeméti Főiskola. 6000 Kecskemét Izsáki út 10, Magyarország. E-mail: osztenyine.eva@gamf.kefo.hu