Gradus Vol 2, No 2(2015) 52-62
ISSN 2064-8014

Eloszláscsaládokhoz való illeszkedés vizsgálata

Goodness of fit to family of distribution

*Oszténayiné Krauczi Éva**

Természet- és Műszaki Alaptudományi Tanszék, Gépipari és Automatizálási Műszaki Főiskolai Kar, Kecskeméti Főiskola, Magyarország

Kulcsszavak:
illeszkedésvizsgálat,
határeloszlás-tétel

Keywords:
goddness-of-fit,
asymptotic distribution theorem

Cikktörténet:
Beérkezett 2015. október 10.
Átdolgozva 2015. október 31.
Elfogadva 2015. november 5.

Összefoglalás

Eloszláscsaládokhoz való illeszkedés vizsgálata a matematikai statisztikának a hipotézisvizsgálathoz tartozó területe. Ennek az összefoglaló cikknek az a célja, hogy áttekintést adjon az ezen a területen elért legelső eredményekről. Felidézzük az első teszteket, amelyekkel rögzített eloszláshoz való illeszkedést lehet ellenőrizni, valamint, hogy hogyan találták meg ezeknek a tesztstatisztikáknak a határeloszlásait. Majd az első összetett illeszkedésvizsgálati módszereket és határeloszlásukat elevenítjük fel. Ezen eljárások két nagy osztályát tárgyaljuk részletesen, az egyik a minta eloszlásának és az eloszláscsalád eloszlásainak távolságán alapuló tesztek, a másik a regresszió-, illetve korrelációtesztek.

Abstract

Goodness of fit to family of distribution belongs to hypothesis tests of mathematical statistics. The goal of this paper is to give a summary of the first results of this area. For the overview we recall the first tests which are suitable for goodness of fit to a fixed distribution paying special attention to the development of the asymptotic theory of goodness of fit tests. The goodness of fit to family of distributions and their asymptotic theories are considered, focusing on two classes of this procedure: tests of fit based on the empirical distribution function (EDF), and the regression and correlation tests of fit.

1. Bevezetés

A hipotézisvizsgálat, és ezen belül az illeszkedésvizsgálat fontos területe a matematikai statisztikának. Arra a kérdésre, hogy mikor merült fel az első ilyen típusú probléma az emberiség történetében, a teljes ismeret hiányában nem tudunk teljes bizonyossággal válaszolni. Annyit tudunk, hogy 1812-ben Laplace csillagászati vizsgálataiban statisztikai módszert használt annak a hipotézisnek az eldöntésére, hogy a naprendszer üstökösei szerves részei a naprendszernek, vagy csak külső behatolók. Ha csak külső behatolók az üstökösök, akkor pályasíkjuk és az ekliptika közötti szög egyenletes eloszlású kell legyen a $(0, 2\pi)$ intervallumon, vagyis egy illeszkedésvizsgálatot kellett elvégeznie.

Az illeszkedésvizsgálat igazi úttörői K. Pearson, E. S. Pearson, A. Fisher és J. Neymann voltak, akik az első eljárásokat dolgozták ki annak a hipotézisnek az eldöntésére, hogy egy véletlen mennyiség eloszlása a minta gyakoriságeloszlása alapján tekinthető-e egy megadott F eloszlással megegyezőnek. Ezt nevezzük egyszerű illeszkedésvizsgálatnak. Később szükség lett olyan eljárásokra is, melyekkel arról a hipotézisről tudtak döntést hozni, hogy a minta egy megadott eloszláscsaládból származik-e. Ezeket az eljárásokat nevezzük összetett illeszkedésvizsgálatnak.

Ennek a cikknek az a célja, hogy bemutassa az összetett illeszkedésvizsgálat első fontos eljárásait. Ehhez del Barrio, Cuesta-Albertos és Matrán [17] cikkét használtuk, amely cikkben egy kitűnő összefoglalás található. Az eljárások bemutatása alatt egyrészt a pontos módszer, a tesztstatisztika, másrészt a tesztstatisztika határeloszlásának megadását értjük. A 2. fejezetben a rögzített eloszláshoz illeszkedés vizsgálatára használt legelső módszereket mutatjuk be. A 3. fejezetben az összetett illeszkedésvizsgálati eljárásokat tárgyaljuk. Ezen eljárások két nagy osztályát mutatjuk be részletesen, az egyik a minta eloszlásának és az eloszláscsalád eloszlásainak távolságán alapuló tesztek (3.1. fejezet), a másik a regresszió-, illetve korrelációtesztek (3.2. fejezet).

*Kapcsolattartó szerző.
E-mail cím: osztenyine.eva@gamf.kefo.hu

A következőkben bevezetjük az általunk használt jelöléseket. A nemnegatív egészek halmazát $\mathbb{N}$, a valós számok halmazát $\mathbb{R}$ és a komplex számok halmazát $\mathbb{C}$ jelöli. Minden véletlen változó ugyanazon $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn van definiálva. Jelölje $\mathbf{I}_A$ az A esemény indikátor változóját. Legyenek $X_1, \dots, X_n$ független azonos eloszlású véletlen változók, vagyis egy statisztikai minta. Jelölje $F(x),\ x \in \mathbb{R}$, a változók közös eloszlásfüggvényét, és

$$Q_F(t) = F^{-1}(t) := \inf\{x \in \mathbb{R} : F(x) \geq t\}, \qquad t \in (0,1), \tag{1}$$

az F eloszlásfüggvény kvantilisfüggvényét. Legyen

$$\bar{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k, \qquad S_n^2 = \frac{1}{n}\sum_{k=1}^{n}(X_k - \bar{X}_n)^2, \text{ illetve } \quad m_i = \frac{1}{n}\sum_{k=1}^{n}(X_k - \bar{X}_n)^i \tag{2}$$

a minta átlaga, szórásnégyzete, illetve i-edik centrális momentuma. Jelölje

$$F_n(x) = \frac{1}{n}\sum_{k=1}^{n} \mathbf{I}_{\{X_k \leq x\}}, \qquad \text{illetve} \qquad \alpha_{F,n}(x) = \sqrt{n}\big(F_n(x) - F(x)\big), \qquad x \in \mathbb{R}, \tag{3}$$

az empirikus eloszlásfüggvényt, illetve az empirikus folyamatot. A rendezett mintára az $X_{1,n}, \dots, X_{n,n}$, a minta kvantilisfüggvényére pedig a $Q_n(t),\ t \in [0,1]$ jelölést használjuk. Vegyük észre, hogy tetszőleges $k = 1, 2, \dots, n$ és $t \in ((k-1)/n, k/n]$ esetén $Q_n(t) = X_{k,n}$.

Jelölje Φ a standard normális eloszlásfüggvényt, φ a hozzá tartozó sűrűségfüggvényt jelöli. Legyen minden $\sigma > 0$ és minden $\mu \in \mathbb{R}$ esetén $N_\sigma^\mu(x) = \Phi((x-\mu)/\sigma),\ x \in \mathbb{R}$, a μ várható értékű és σ szórású normális véletlen változó eloszlásfüggvénye, valamint használjuk az $\mathcal{N} = \{N_\sigma^\mu :\ \sigma > 0,\ \mu \in \mathbb{R}\}$ jelölést a normális eloszláscsaládra, vagyis az összes normális eloszlás osztályára.

Ha a minta a $[0,1]$ intervallumon egyenletes eloszlásból származik, speciálisan jelölje G_n az empirikus eloszlásfüggvényét. Az egyenletes empirikus folyamatot

$$\alpha_n(t) = \sqrt{n}\big(G_n(t) - t\big), \qquad t \in [0,1], \tag{4}$$

a Brown-hidat $B(t),\ t \in [0,1]$ jelöli. Ez utóbbi egy mintafolytonos, $E(B(t)) = 0$ várható értékű és $\operatorname{Cov}(B(s), B(t)) = \min(s,t) - st,\ s,t \in [0,1]$, kovarianciafüggvényű Gauss-folyamat.

Két metrikus térre lesz szükség. Az egyik a $\mathcal{C}[0,1]$ tér, amely definíció szerint az összes $[0,1]$ intervallumon értelmezett, valós értékű, folytonos függvények halmaza. A $\mathcal{C}[0,1]$ tér a

$$\|x\|_\infty := \sup_{0 \leq t \leq 1} |x(t)|, \qquad x \in \mathcal{C}[0,1], \tag{5}$$

úgynevezett suprémum normával van ellátva, mellyel ez a tér teljes, szeparábilis metrikus tér lesz. A másik a $\mathcal{D}[0,1]$ tér, mely azon $[0,1]$ intervallumon értelmezett, valós értékű függvények halmaza, amelyek jobbról folytonosak és van baloldali határértékük. Ez a tér egy olyan távolsággal van ellátva, melyet Szkorohod vezetett be, és amivel ez is teljes, szeparábilis metrikus tér. A Brown-híd a $\mathcal{C}[0,1]$, az egyenletes empirikus folyamat a $\mathcal{D}[0,1]$ tér véletlen eleme.

A cikkben minden konvergencia úgy értendő, amint $n \to \infty$. A $\to_{\mathcal{D}}$ az eloszlásban való, a $\to_{\mathbf{P}}$ pedig a sztochasztikus konvergenciát jelöli. Az eloszlásbeli egyenlőséget $=_{\mathcal{D}}$ jelöli.

2. Illeszkedésvizsgálat rögzített eloszlás esetén

Az egyszerű illeszkedésvizsgálat azt jelenti, hogy a minta egy adott, rögzített $F_0(x),\ x \in \mathbb{R}$, eloszlásfüggvényhez való illeszkedését vizsgáljuk. Adott egy $X_1, \dots, X_n$ véletlen minta egy ismeretlen $F(x),\ x \in \mathbb{R}$, eloszlásfüggvényű véletlen változóból. Döntsük el a minta alapján, igaz-e az az egyszerű nullhipotézis, hogy

$$\mathcal{H}_0 : F = F_0\,.$$

A *Pearson-féle* χ^2*-teszt*et tekinthetjük az első ilyen illeszkedésvizsgálatnak [36]. Az ötlet a következő: osszuk fel a valós egyenest k db páronként diszjunkt cellára, melyek együtt lefedik az egész valós egyenest. Legyenek ezek a $C_1, \dots, C_k$ cellák olyanok, hogy a nullhipotézis mellett annak a valószínűsége, hogy a véletlen változó beleesik ezekbe a cellákba rendre $p_1, \dots, p_k$. Vagyis, ha $F = F_0$, akkor $P(X_1 \in C_i) = p_i,\ i = 1, \dots, k$. Legyen O_i az i-edik cellába eső megfigyelések száma. Ekkor O_i binomiális eloszlású n és p_i paraméterekkel. Így a Moivre–Laplace-tétel szerint

$$\frac{O_i - np_i}{\sqrt{np_i(1-p_i)}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1). \tag{6}$$

A többváltozós centrális határeloszlás tétel azt állítja, hogy ha $l \leq k$, akkor a

$$B_l = \frac{1}{\sqrt{n}} \left(O_1 - np_1, \ldots, O_l - np_l\right)^\top \tag{7}$$

véletlen vektornak van határeloszlása. A határeloszlás a nulla várható értékű és $\Sigma_l = (\sigma_{i,j})_{i,j=1,\ldots,l}$ kovarianciamátrixú normális eloszlás, ahol a kovarianciamátrix elemei $\sigma_{i,j} = -p_i p_j\,, i \neq j$ esetén, és $\sigma_{i,i} = p_i(1-p_i)$. Sőt, ha $p_i > 0$ minden $i = 1, \ldots, k$ esetén, akkor a Σ_{k-1} kovarianciamátrixnak létezik inverze, $\Sigma_{k-1}^{-1} = (\nu_{i,j})_{i,j=1,\ldots,k-1}$, melynek elemei $\nu_{i,j} = p_k^{-1}\,, i \neq j$ esetén, és $\nu_{i,i} = p_i^{-1} + p_k^{-1}$. Ekkor könnyen látható, hogy

$$\chi^2 := \sum_{j=1}^{k} \frac{(O_j - np_j)^2}{np_j} = B_{k-1}^\top \Sigma_{k-1}^{-1} B_{k-1} \xrightarrow{\mathcal{D}} \chi_{k-1}^2, \tag{8}$$

így kapjuk meg a következő jól ismert aszimptotikus eredményt.

2.1. Tétel. *A nullhipotézis teljesülése mellett χ^2 aszimptotikus eloszlása χ_{k-1}^2.*

A teszt hátránya, hogy nagy szabadságot enged a cellák méretének, helyének és számának megválasztásában. Például nem tud különbséget tenni két különböző eloszlás között, melyek a kiválasztott cellákhoz azonos valószínűséget rendelnek.

Az illeszkedésvizsgálati eljárások következő nagy osztálya az *EDF* (Empirical Distribution Function)-*tesztek*. Ezen tesztek alapötlete az, hogy mérjük meg az F_0 hipotetikus eloszlásfüggvény és a mintából számolt F_n empirikus eloszlásfüggvény távolságát, és ezen eltérés nagysága alapján döntsünk a megegyezésről, illetve különbözőségről. Az egyes tesztek abban különböznek egymástól, hogy hogyan mérjük meg a két függvény távolságát.

Az első ilyen teszt Cramér (1928), [7], ennek általánosított változata pedig von Mises (1931), [48] nevéhez fűződik. A von Mises-féle tesztstatisztika

$$\omega_n^2 := n \int_{-\infty}^{\infty} \big(F_n(x) - F_0(x)\big)^2 w(x) \mathrm{d}x \tag{9}$$

alakban van definiálva, tehát súlyozott L^2-normában méri a két függvény távolságát, ahol w a különbözőséget alkalmasan mérő súlyfüggvény. Speciálisan a Cramér-teszt a $w \equiv 1$ választással adódik. Kolmogorov (1933), [29] a suprémum normát használja, a kétoldali tesztstatisztikája

$$D_n := \sqrt{n} \sup_{x \in \mathbb{R}} |F_n(x) - F_0(x)|, \tag{10}$$

Szmirnov (1939, [43], 1941, [44]) egyoldali tesztstatisztikái

$$D_n^+ := \sqrt{n} \sup_{x \in \mathbb{R}} \big(F_n(x) - F_0(x)\big), \qquad D_n^- := \sqrt{n} \sup_{x \in \mathbb{R}} \big(F_0(x) - F_n(x)\big), \tag{11}$$

melyekre $D_n = \max(D_n^+, D_n^-)$. A három statisztikát együtt *Kolmogorov–Szmirnov-statisztikák*nak nevezik. Ezen statisztikák előnye, hogy eloszlásmentes statisztikák, vagyis minden folytonos F_0 eloszlásfüggvény esetén, a nullhipotézis mellett

$$D_n \stackrel{\mathcal{D}}{=} \sup_{0 \leq t \leq 1} |\alpha_n(t)|, \qquad D_n^+ \stackrel{\mathcal{D}}{=} \sup_{0 \leq t \leq 1} \alpha_n(t), \qquad \text{és} \qquad D_n^- \stackrel{\mathcal{D}}{=} \sup_{0 \leq t \leq 1} (-1)\alpha_n(t). \tag{12}$$

Így minden folytonos eloszlás esetén, adott szignifikanciaszinthez és mintamérethez ugyanaz a kritikus érték tartozik. Ez a tulajdonság nem teljesül az ω_n^2 statisztikára, de a Szmirnov (1936), [41], (1937), [42] által javasolt

$$W_n^2(\Psi) := n \int_{-\infty}^{\infty} \Psi\big(F_0(x)\big)\big(F_n(x) - F_0(x)\big)^2 \mathrm{d}F_0(x) \tag{13}$$

változatára már igen. Az összes ilyen statisztikát, amit Ψ változtatásával kapunk, *Cramér–von Mises-típusú statisztiká*nak nevezünk, ahol Ψ tetszőleges valós értékű, a valós számok halmazán értelmezett függvény. A különböző súlyfüggvények használata lehetőséget ad különböző alternatívák felismerésére, éppen ezért a Kolmogorov-statisztikának is megadták a súlyozott változatát:

$$K_n(\Psi) := \sqrt{n} \sup_{x \in \mathbb{R}} \frac{|F_n(x) - F_0(x)|}{\Psi\big(F_0(x)\big)}. \tag{14}$$

Bár ez se bírta kompenzálni azt a hiányát a suprémum normának, hogy csak a legnagyobb elterést érzékeli F_n és F_0 között, amíg az L^2-norma ezen két függvény súlyozott átlagos távolságát méri. Ezen heurisztikus megfigyelést a szimuláció is alátámasztja. Például Krauczi (2009), [30] a normális eloszláscsaládhoz való illeszkedésvizsgálat esetében azt találta, hogy a Kolmogorov-tesztnek a legtöbb alternatívával szembeni ereje jóval kisebb, mint más próbák ereje, ahol erő alatt, annak az eseménynek a valószínűségét értjük, hogy a teszt visszautasítja a normális mintát.

Két statisztika különös figyelmet kapott az irodalomban. A $\Psi \equiv 1$ esetén,

$$W_n^2 := n \int_{-\infty}^{\infty} \big(F_n(x) - F_0(x)\big)^2 \mathrm{d}F_0(x) \tag{15}$$

a *Cramér–von Mises-statisztika*; valamint a $\Psi(t) = (t(1-t))^{-1},\ t \in (0,1)$, mellett

$$A_n^2 := n \int_{-\infty}^{\infty} \frac{(F_n(x) - F_0(x))^2}{F_0(t)(1 - F_0(t))} \mathrm{d}F_0(x) \tag{16}$$

az *Anderson–Darling-statisztika* [2], mely utóbbi a szimulációs vizsgálatok alapján a legerősebb ilyen típusú tesztnek tűnik (lásd például Stephens [45], Krauczi [30]).

Ahhoz, hogy használni tudjuk a gyakorlatban ezeket a teszteket, ismernünk kell az eloszlásfüggvényüket tetszőleges $n \in \mathbb{N}$ esetén, vagy legalább az aszimptotikus eloszlásukat. Szmirnov (1941), [44] explicit formában meg tudta adni D_n^+ eloszlásfüggvényét tetszőleges n esetén, Kolmogorov (1933), [29] pedig megadott egy rekurzív kifejezést, amivel kiszámítható $P(D_n < x)$ valószínűség tetszőleges $n \in \mathbb{N}$ és $x \in \mathbb{R}$ esetén. A Cramér–von Mises-típusú statisztikák eloszlásfüggvényének a megtalálása már nehézséget okozott. Akkoriban Monte-Carlo szimuláció hiányában fontos kérdés volt, hogy ki tudják-e számolni a kritikus értékeket rögzített $n \in \mathbb{N}$
esetén. Emellett a határeloszlás kérdése elméleti, de gyakorlati szempontból is érdekes volt. Az első aszimptotikus eredményt is a Kolmogorov–Szmirnov-típusú statisztikákra sikerült megkapni:

2.2. Tétel. *Minden $x > 0$ esetén*
(Kolmogorov 1933, [29])

$$\lim_{n\to\infty} P(D_n \le x) = \sum_{j=-\infty}^{\infty} (-1)^j e^{-2j^2x^2}, \tag{17}$$

(Szmirnov 1941, [44])

$$\lim_{n\to\infty} P(D_n^+ > x) = \lim_{n\to\infty} P(D_n^- < x) = e^{-2x^2}. \tag{18}$$

Feller (1948), [22] megjegyezte, hogy Kolmogorov és Szmirnov teljesen különböző módszerrel bizonyították állításaikat, és megpróbálta egységesíteni a bizonyításukat. Mivel a D_n, D_n^+ és W_n^2 statisztikák az F_n empirikus és az F_0 elméleti eloszlásfüggvények eltérését mérik, vagyis az $\alpha_{F,n}$ empirikus folyamat funkcionáljai, ezért ezen statisztikák $\mathcal{H}_0$ melletti határeloszlásait valami közös technikával lehetne származtatni. Így Feller cikke fontos lépés az empirikus folyamatra épített illeszkedésvizsgálat aszimptotikus elméletének egységesítésében. Bár ekkor még magát az empirikus folyamatot és az ő határeloszlását nem vizsgálták.

Doob (1949), [19] a véges dimenziós eloszlásokat vizsgálva sejtette meg az egyenletes empirikus folyamatnak a Brown-hídhoz való konvergenciáját, de bizonyítani nem tudta. Viszont bizonyította, hogy minden $x > 0$ esetén

$$P\left(\sup_{0\le t\le 1} |B(t)| \le x\right) = \sum_{j=-\infty}^{\infty} (-1)^j e^{-2j^2x^2} \tag{19}$$

és

$$P\left(\sup_{0\le t\le 1} B(t) > x\right) = e^{-2x^2}, \tag{20}$$

vagyis az egyenletes empirikus folyamat abszolút szuprémum és szuprémum funkcionáljainak határeloszlása megegyezik a Brown-híd ugyanezen funkcionáljainak eloszlásával. Ez azt jelenti, hogy ha Doob sejtése igaz, akkor Kolmogorov és Szmirnov eredményeire talán egyszerűbb bizonyítás adható. Donsker (1951), [18] invariancia elve által nyert bizonyítást a sejtés. Az invariancia elv a következőt jelenti. A részletösszeg folyamat minden folytonos funkcionáljának eloszlása konvergál a Brown-mozgás megfelelő funkcionáljának eloszlásához, illetve az egyenletes empirikus folyamat minden folytonos funkcionáljának eloszlása konvergál a Brown-híd megfelelő funkcionáljának eloszlásához.

Ezen eredmények hatására fejlődött ki a metrikus terekben való gyenge konvergencia elmélete többek között Kolmogorovnak, Prohorovnak és Szkorohodnak köszönhetően, amely elmélet segített jobban megérteni az invariancia elvet. Erről szól Billingsley 1968-as könyve [3]. Fontos lépés volt, hogy kidolgozták az elméletet a $\mathcal{C}[0,1]$ és a $\mathcal{D}[0,1]$ tereken. Először a részletösszeg és az empirikus folyamatokat lineáris interpolációval kapott folytonos folyamatokkal közelítették, hogy ne kelljen $\mathcal{C}[0,1]$ térből kilépniük. Ezen új folyamat sorozatokra bizonyították a véges dimenziós eloszlások konvergenciáját és a sorozat feszességét, amely kettő tulajdonság együtt a folymatok eloszlásbeli konvergenciáját adja. A folytonos folyamatokkal való közelítés valahogy mesterkélt. Ahhoz, hogy ezt el tudjuk kerülni, egy gazdagabb téren kell dolgoznunk. Ez a gazdagabb tér a $\mathcal{D}[0,1]$ tér, amelynek már maga az empirikus folyamat is eleme.

2.3. Tétel. *Az $\alpha_n \xrightarrow{\mathcal{D}} B$ konvergencia teljesül a $\mathcal{D}[0,1]$ téren.*

A 2.3. Tétel lehetővé teszi a 2.2. Tétel természetesebb bizonyítását. Vegyük észre, hogy az $x \mapsto \|x\|_\infty$ leképezés folytonos a Szkorohod-topológiára nézve egy B mértéke szerint nulla mértékű halmazt kivéve, és mivel $D_n = \|\alpha_n\|_\infty$, ekkor $D_n \xrightarrow{\mathcal{D}} \|B\|_\infty$. Hasonló konvergencia teljesül a D_n^+ és a D_n^- statisztikák esetében.

A 2.3. Tétel teszi lehetővé a Cramér–von Mises-statisztika határeloszlásásának meghatározását is. Az $x \mapsto \int_0^1 x^2(t)\mathrm{d}t$ funkcionál szintén folytonos a Szkorohod-topológiára nézve egy B mértéke szerint nulla mértékű halmazt kivéve. Így a fenti érvelés ismételt alkalmazásával kapjuk, hogy

$$W_n^2 \xrightarrow{\mathcal{D}} \int_0^1 B(t)^2 \mathrm{d}t\,. \tag{21}$$

Innen pedig egy lépés a Cramér–von Mises-típusú statisztikák határeloszlása. Mint a Brown-hidakra vonatkozó iterált logaritmus tétel következményeként Anderson és Darling 1952-ben megmutatta [2], hogy feltéve az

$$\int_0^\delta \Psi(t) t \log\log\frac{1}{t}\mathrm{d}t \quad \text{és} \quad \int_\delta^1 \Psi(t)(1-t)t\log\log\frac{1}{1-t}\mathrm{d}t \tag{22}$$

integrálok végességét valamilyen $\delta \in (0,1)$ esetén, az $x \mapsto \int_0^1 \Psi(t)x^2(t)\mathrm{d}t$ funkcionál folytonos a Szkorohod-topológiára nézve egy B mértéke szerint nulla mértékű halmazt kivéve, és ennek következményeként

$$W_n^2(\Psi) \xrightarrow{\mathcal{D}} \int_0^1 \Psi(t)B(t)^2\mathrm{d}t\,. \tag{23}$$

Ez a konvergencia az Anderson–Darling-féle súlyfüggvény esetén is teljesül.

3. Illeszkedésvizsgálat eloszláscsalád esetén

Ebben a fejezetben azokat a teszteket tekintjük, ahol a kérdés az, hogy a minta egy adott eloszláscsaládból származik-e. Itt legyen $\mathcal{F}$ eloszlásfüggvények egy parametrikus eloszláscsaládja, azaz

$$\mathcal{F} = \{F(\cdot,\theta) : \theta \in \Theta\}\,, \tag{24}$$

ahol Θ valamilyen nyitott paraméterhalmaz $\mathbb{R}^d$-ben.

Az első vizsgálatok normális eloszláscsalád esetében történtek. Fisher (1930), [24], Pearson (1930), [36] és Williams (1935), [52] voltak az elsők, akik a $\sqrt{\beta_1} = m_3/m_2^{2/3}$ és $\beta_2 = m_4/m_2^2$ standardizált harmadik és negyedik momentumok segítségével mérték meg a normalitástól való eltérést. Pearson, D'Agostino és Bowman (1977), [35] a $\sqrt{\beta_1}$ és β_2 két alkalmas függvényét használta erre. Ezekkel a tesztekkel az a probléma, hogy az előbbi lapultsági és a ferdeségi mutató kevés, hogy karakterizálja a normális eloszlást. Ennek az a következménye, hogy ezek a tesztek olyan nemnormális eloszlásból származó minta esetén, amely ugyan szimmetrikus és a lapultsági mutatója ugyanúgy 3, mint normális eloszlásé, de az alakja nagyon különbözik a normálistól, mégis elfogadják a nullhipotézist. Másrészt a gyakorlati alkalmazások szempontjából az is fontos, hogy ha egy eloszlás kicsit különbözik a normális eloszlástól a teszt azt ne vesse el. Például Ali (1977), [1] adott egy olyan sorozatát eloszlásoknak, amely ugyan eloszlásban tart a standard normális eloszláshoz, de a lapultsági mutatója felrobban. Vagyis, ha a sorozat elég nagy indexű tagjából származik a mintánk, akkor nagy eséllyel ezek a tesztek elutasítják, pedig valójában közel normális eloszlásról van szó.

Más normalitás tesztek, például az

$$u := \frac{X_{n,n} - X_{1,n}}{\left(\frac{n}{n-1}\right)^{\frac{1}{2}} m_2^{\frac{1}{2}}}$$

statisztika (David, Hartley és Pearson 1954, [15]) a terjedelem és a szórás, valamint az

$$a := \frac{\sum_{j=1}^n |X_j - \bar{X}_n|}{n \cdot m_2^{\frac{1}{2}}}$$

statisztika (Geary 1947, [25]) a mintaátlagtól való átlagos abszolút eltérés és a szórás hányadosából származtatott tesztek. Ezek a tesztek csak egyes alternatívákkal szemben viselkednek jól, de kicsi erővel bírnak alternatívák széles skálájával szemben.

A következő alfejezetben azokat a teszteket mutatjuk be, amelyeket rögzített eloszláshoz való illeszkedéstesztek átdolgozásaként kapunk.

3.1. Eloszláscsalád tesztelése rögzített eloszláshoz való illeszkedésvizsgálat segítségével

A 2. fejezetben rögzített eloszláshoz való illeszkedés teszteket tekintettünk. Egy lehetőség, hogy eloszláscsaládhoz való illeszkedést teszteljünk ezekkel a tesztekkel, ha a θ paraméternek a $\mathcal{H}_0$ mellett egy $\hat{\theta}_n$ becslését véve azt ellenőrizzük, hogy a minta $F(x, \hat{\theta}_n)$, $x \in \mathbb{R}$, eloszlásfüggvényű-e. Ezt javasolta Pearson a χ^2-tesztje esetében. Legyen

$$\hat{\chi}^2 := \sum_{j=1}^{k} \frac{(O_j - np_j(\hat{\theta}_n))^2}{np_j(\hat{\theta}_n)}, \tag{25}$$

ahol $p_j(\theta)$ annak a valószínűsége, hogy X_1 a j-edik cellába esik $F(x, \theta)$, $x \in \mathbb{R}$, mellett. Pearson nem tudta megadni $\hat{\chi}^2$ aszimptotikus eloszlását. Fisher volt az, aki rámutatott arra, hogy a határeloszlás függ a paraméter becslésének módszerétől, és megmutatta, hogy a szokásos feltételek mellett, ha a θ maximum-likelihood becslését vesszük a csoportosított $(O_1, \dots, O_k)$ adatokon, akkor a $\hat{\chi}^2$ statisztikának χ^2_{k-d-1} a határeloszlása (lásd Cochran 1952, [6]).

Fisher azt is megfigyelte, hogy a csoportosított $(O_1, \dots, O_k)$ mintából származó $\hat{\theta}_n$ becslésből adódó információvesztés erőcsökkenést eredményez. Ezért Fisher abban az esetben is megvizsgálta $\hat{\chi}^2$ határeloszlását, amikor a θ paraméter egydimenziós, és a teljes mintából vesszük a θ paraméter maximum-likelihood becslését. Az eredményét Chernoff és Lehmann (1954), [5] d-dimenziós paraméterre általánosította, nevezetesen, hogy megfelelő feltételek mellett

$$\hat{\chi}^2 \xrightarrow{\mathcal{D}} \sum_{j=1}^{k-d-1} Z_j^2 + \sum_{j=k-d}^{k-1} \lambda_j Z_j^2, \tag{26}$$

ahol Z_j független standard normális változók, és $\lambda_j \in [0,1]$, $j = k-d, \dots, k-1$, olyan konstans, amely függhet a θ paraméter igazi értékétől. Ez a függés mutatja az egyik nagy hátrányát a $\hat{\chi}^2$-teszt használatának eloszláscsalád esetében.

A másik nehézség a $\hat{\chi}^2$-teszt használatában a cellák választása. Az O_i cellagyakoriságok aszimptotikus normalitásának a következménye a Pearson-féle statisztika aszimptotikus χ^2_{k-1}-eloszlása. Viszont egy kicsi várható gyakorisággal rendelkező cella esetében az O_i változó nagyon lassan konvergál a normális eloszláshoz, ami azt eredményezi, hogy a (26) konvergencia lassú. Vagyis az asszimptotikus kritikus értékek használatának létjogosultsága sérülne ebben az esetben. A gyakorlatban ezt úgy próbálják meg elkerülni, hogy „olyan cellákat használnak, amelyekbe legalább 10 megfigyelés esik” (lásd Cochran 1952, [6]).

A cellák jó választására nézve Mann és Wald (1942), [34] valamint Gumbel (1943), [26] azt javasolták rögzített eloszlás esetén, hogy a nullhipotézis mellett azonos valószínűségű cellákat használjunk, ezáltal csökkentve a cellák választásának esetlegességét. Ez a gondolat paraméteres eloszláscsalád esetére úgy vihető át, hogy először vegyük valamilyen alkalmas becslését θ-nak, majd $F(x, \hat{\theta}_n)$, $x \in \mathbb{R}$, mellett azonos valószínűségű cellákat használjunk. Vagyis megint véletlenül fogunk cellákat választani! Ugyanúgy a minta határozza meg, hogy melyik cellákat használjuk, mint amikor olyan cellákat választunk, amelyekbe legalább 10 megfigyelés esik. Watson(1957, [49], 1958 [50]) megmutatta, ha $\hat{\theta}_n$ a teljes mintából származó maximum-likelihood becslése θ-nak, valamint a j-edik cella végpontjai $F^{-1}((j-1)/k, \hat{\theta}_n)$ és $F^{-1}(j/k, \hat{\theta}_n)$, akkor (26) teljesül. Továbbá, ha $\mathcal{F}$ eltolás-skála család, akkor a λ_j együtthatók nem függnek a θ paramétertől, csak az eloszláscsaládtól.

Az EDF-tesztek adaptációja eloszláscsaládok esetére könnyen kivitelezhető, és hasonlóan a rögzített eloszlás esetére, ezek a tesztek jobb erővel bírnak, mint a $\hat{\chi}^2$-tesztek. Legyen $\hat{\theta}_n$ valamilyen becslése θ-nak. Ekkor a megfelelő becsléses statisztikák

$$\widehat{W}_n^2(\Psi) := n \int_{-\infty}^{\infty} \Psi\left(F(x, \hat{\theta}_n)\right) \left(F_n(x) - F(x, \hat{\theta}_n)\right)^2 \mathrm{d}F(x, \hat{\theta}_n) \tag{27}$$

és

$$\hat{K}_n(\Psi) := \sqrt{n} \sup_{x \in \mathbb{R}} \frac{|F_n(x) - F(x, \hat{\theta}_n)|}{\Psi\left(F(x, \hat{\theta}_n)\right)}. \tag{28}$$

A $\Psi \equiv 1$ esetben a két statisztikát a $\widehat{W}_n^2$ és $\hat{K}_n$ jelöli. A kívánatos eloszlásmentesség, ami a rögzített esetben teljesült, itt sajnos nem igaz. Legyen $Z_i = F(X_i, \hat{\theta}_n)$, és $\hat{G}_n(t)$, $t \in [0,1]$ jelölje a $Z_1, \dots, Z_n$ változókhoz tartozó empirikus eloszlásfüggvényt. Ekkor

$$\widehat{W}_n^2(\Psi) = n \int_0^1 \Psi(t)(\hat{G}_n(t) - t)^2 \mathrm{d}t \tag{29}$$

és

$$\hat{K}_n(\Psi) = \sqrt{n} \sup_{0<t<1} \frac{|\hat{G}_n(t) - t|}{\Psi(t)}. \tag{30}$$

Tehát a két statisztika értéke csak a $\hat{G}_n$ függvénytől függ. Viszont $Z_1, \ldots, Z_n$ nem független, azonosan egyenletes eloszlású véletlen változók, ami azt eredményezi, hogy $\hat{G}_n$ függvény funkcionáljainak eloszlására nem alkalmazhatók az eddigiek. Éppen ezért $\hat{G}_n$ nem olyan, amivel klasszikus értelemben tudunk dolgozni. Számos fontos esetben $Z_1, \ldots, Z_n$ eloszlása nem függ a θ paramétertől, csak az eloszláscsaládtól, vagyis ekkor $\widehat{W}_n^2(\Psi)$ és $\hat{K}_n^2(\Psi)$ paramétermentes. Ez történik az eltolás-skála családok esetében, amikor olyan $\hat{\theta}$ becslést használunk, amiben a becslés felcserélhető a skálázással, illetve az eltolással (lásd David és Johnson 1948, [14]). Az eltolás-skála eloszláscsalád olyan család, ahol adott egy H_0 standardizált (0 várható értékű és 1 szórású) eloszlásfüggvény, és az eloszláscsalád többi tagja lineáris transzformációval kapható belőle. Lilliefors (1967), [33] ezt használta fel és készítette el a népszerű táblázatát a normális eloszláscsalád esetére a Kolmogorov–Szmirnov-statisztikához.

A becsléses $\widehat{W}_n^2(\Psi)$ és $\hat{K}_n^2(\Psi)$ típusú statisztikák határeloszlásának a meghatározására tett első kísérlet Darling nevéhez fűződik (1955), [13]. A becsléses Cramér–von Mises-statisztika aszimptotikus eloszlását tudta meghatározni abban az esetben, amikor a θ paraméter egydimenziós. Sukhatme (1972), [46] kiterjesztette Darling eredményét többdimenziós paraméterekre. Ezekben a cikkekben egy segédfolyamaton keresztül találták meg $\widehat{W}_n^2$ határeloszlását.

Kac, Kiefer és Wolfowitz (1955), [28] viszont közvetlenül az

$$\hat{\alpha}_n(t) = \sqrt{n}(\hat{G}_n(t) - t), \qquad t \in [0,1],$$

becsléses empirikus folyamatot tanulmányozva kapták meg $\widehat{W}_n^2$ határeloszlását normális eloszláscsalád esetén a maximum-likelihood paraméterbecslésekkel: $\hat{\theta}_n = (\hat{X}_n, S_n^2)$. Ugyan a becsléses empirikus folyamatnak a gyenge konvergenciáját nem bizonyították, de megmutatták, hogy

$$\hat{W}_n^2 \xrightarrow{\mathcal{D}} \int_0^1 (Z(t))^2 \, \mathrm{d}t\,, \tag{31}$$

ahol $Z(t),\, t \in (0,1)$ egy 0 várható értékű és

$$K(s,t) = \min(s,t) - st - \varphi\big(\Phi^{-1}(s)\big)\varphi\big(\Phi^{-1}(t)\big) - \frac{1}{2}\Phi^{-1}(s)\varphi\big(\Phi^{-1}(s)\big)\Phi^{-1}(t)\varphi\big(\Phi^{-1}(t)\big) \tag{32}$$

kovarianciafüggvényű Gauss folyamat.

A becsléses empirikus folyamat gyenge konvergenciájának általános vizsgálata Durbin nevéhez fűződik (1973), [20]. Az eloszláscsaládra és a paraméterre tett megfelelő regularitási feltételek mellett az $\hat{\alpha}_n$ empirikus folymat gyengén konvergál a 0 várható értékű és $K(s,t),\ s,t \in [0,1]$ kovarianciafüggvényű Gauss folyamathoz. Durbin cikkjében explicit formulát adott a $K(s,t)$ kovarianciafüggvényre, és standard számolással megmutatható, hogy ennek speciális esete a Kac, Kiefer és Wolfowitz által megadott kovariancia.

Megjegyezzük, hogy Burke, Csörgő M., Csörgő S. és Révész (1979), [4] cikkéből következik Durbin eredménye. Ebben a cikkben a becsléses empirikus folyamatot Gauss folyamatok sorozatával közelítik. Azon túl, hogy Durbin tételéből következik a $\hat{W}_n^2(\Psi)$ és $\hat{K}_n^2(\Psi)$ típusú statisztikák nullhipotézis melletti eloszlásbeli konvergenciája, a [4] cikk eredménye az aszimptotikus erők tanulmányozásának is eszköze lehet.

Az empirikus folyamatot tanulmányozó elmélet fejlődésének következményeként további illeszkedést vizsgáló technikák jelentek meg. Például Feuerverger és Mureika (1977), [23], valamint Csörgő S. (1981), [8] az empirikus karakterisztikus függvény aszimptotikus eloszlását vizsgálták. A Durbin-tétel analóg változatát empirikus karakterisztikus és kvantilis függvényekre Csörgő S. (1981), [9] és LaRiccia és Mason (1986), [31] dolgozták ki. Ezen eredmények segítségével új normalitástesztek születtek, melyek közül Murota és Takeuchi (1981), Hall és Wels (1983), [27], Epps és Pulley (1983), [21] valamint Csörgő S. (1986a, 1989), [10], [11] eredményeit említjük meg.

Egy másik ötlet, hogy hogyan tudjuk a rögzített eloszlás esetében használt tesztelési eljárást parametrikus eloszláscsalád esetében használni, a *minimum távolság módszere*. Legyen δ egy metrika az eloszlásfüggvények halmazán. Ekkor $\Delta(F_n, \mathcal{F}) = \inf_\theta \delta(F_n, F(\cdot\,, \theta))$ egy lehetséges mértéke az empirikus eloszlásfüggvény $\mathcal{F}$ parametrikus eloszláscsaládtól való távolságának. Pollard használta ezt először (1980), [37] és meghatározta $\Delta(F_n, \mathcal{F})$ határeloszlását, tetszőleges normált lineáris tér értékű véletlen változók esetében.

3.2. Regresszió- és korrelációtesztek

Ebben a fejezetben tegyük fel, hogy $\mathcal{F}$ eltolás-skála család.

Az ötlet a következő. Legyen $X_1, \ldots, X_n$ az $\mathcal{F}$ eloszláscsaládból származó μ várható értékű és σ^2 szórásnégyzetű minta. A korábbi jelöléseknek megfelelően legyen $\mathbf{X}_n = (X_{1,n}, \ldots, X_{n,n})$ a mintához tartozó rendezett minta. Továbbá legyen $\mathbf{Z}_n = (Z_{1,n}, \ldots, Z_{n,n})$ H_0 eloszlásfüggvényű rendezett minta, és jelölje $\mathbf{m} = (m_1, \ldots, m_n)$ a várható érték vektorát és $V = (v_{ij})_{i,j=1\ldots,n}$ a kovarianciamátrixát. Könnyen látszik, hogy

$$X_{i,n} \overset{\mathcal{D}}{=} \mu + \sigma Z_{i,n}, \qquad i = 1, \ldots, n\,. \tag{33}$$

Ha kétdimenziós koordinátarendszerben ábrázoljuk az $(m_i, X_{n,i}),\ i = 1, \ldots, n$ pontokat, akkor ezeknek közelítőleg egy egyenesre kell esniük, és a linearitás hiánya azt sugallja, hogy X_1 eloszlásfüggvénye nem $\mathcal{F}$-beli. Gyakran ezt csak „szemre" ellenőrzik, de vannak analitikus eljárások is ennek az ellenőrzésére. Két nagy osztálya van ezeknek az eljárásoknak: az egyik a *regresszió*-, a másik a *korrelációtesztek*, mely különböző eljárások valójában ekvivalens tesztekre vezetnek.

Az egyik esetben a (33) lineáris model segítségével adunk egy $\hat{\sigma}_n^2$ becslést a σ^2 szórásnégyzetre és ezt hasonlítjuk össze az S_n^2 becsléssel. Ekkor a nullhipotézis mellett a $\hat{\sigma}_n^2/S_n^2$ tesztstatisztika értéke közel kell legyen 1-hez, ellenkező esetben elvetjük a nullhipotézist. Ezeket az eljárásokat nevezik *regressziótesztek*nek. A másik osztálya ezen teszteknek a ρ korrelációs együttható segítségével ellenőrzi, van-e lineáris kapcsolat az $\mathbf{X}_n$ véletlen vektor és az $\mathbf{m}$ determinisztikus vektor között. Ekkor a nullhipotézis mellett a $\rho^2(\mathbf{m}, \mathbf{X}_n)$ tesztstatisztika értéke közel kell legyen 1-hez, ellenkező esetben elvetjük a nullhipotézist. Ezeket az eljárásokat nevezik *korrelációtesztek*nek.

A regressziótesztek első változata *Wilk és Shapiro* (1965), [40] W *normalitásteszt*je. A μ és σ paraméterek legjobb lineáris torzítatlan becslése a (33) model alapján

$$\hat{\mu}_n = \bar{X}_n \qquad \text{és} \qquad \hat{\sigma}_n = \frac{\mathbf{m}^\top V^{-1}\mathbf{X}_n}{\mathbf{m}^\top V^{-1}\mathbf{m}}. \tag{34}$$

Wilk és Shapiro a W tesztstatisztikát a $\hat{\sigma}_n^2/S_n^2$ tesztstatisztika normalizált változataként definiálta

$$W := \frac{(\mathbf{m}^\top V^{-1}\mathbf{X}_n)^2}{\mathbf{m}^\top V^{-1}V^{-1}\mathbf{m}\sum_i (X_i - \bar{X})^2} \tag{35}$$

alakban. Ezzel egy regressziótesztet kaptak. Másrészt ez egy korrelációteszt is, ami a normalizációból következik, ugyanis $W = \rho^2(V^{-1}\mathbf{m}, \mathbf{X}_n)$. Shapiro, Wilk és Chen (1968), [38], szimulációs vizsgálatából kiderült, hogy a W-teszt egyike a legerősebb normalitásteszteknek alternatívák széles skálájával szemben, valamint Krauczi (2009), [30] szimulációs vizsgálata is ezt támasztotta alá. Ezért népszerű módszer a mai napig, annak ellenére, hogy rejteget egy-két nehézséget a használata.

Egyik probléma, hogy magát a W tesztstatisztikát bonyolult kiszámítani. Ahhoz, hogy W-t meg tudjuk határozni, előzetesen ki kell számolnunk az $\mathbf{m}$ vektort és a V^{-1} mátrixot. Ez a mintaméret növekedésével egyre nehezebb feladat, és valójában amikor W-t bevezették, legfeljebb 20 elemű minta esetén tudták megadni a V^{-1} mátrix elemeit pontosan. Ezért már Wilk és Shapiro is numerikus közelítéssel számolta W értékeit 50-es mintaméretig. Egy másik probléma, hogy az $n = 3$ esetet kivéve nem ismerjük W eloszlásfüggvényét. Mivel az $n = 3$ esetben a W-teszt megegyezik az u-teszttel, ekkor W pontos eloszlása is ismert. Wilk és Shapiro $n = 50$ mintaméretig szimulációval adták meg a kritikus értékeket. A határeloszlás viszont 1986-ig ismeretlen volt, amikor is Leslie, Stephens és Fotopoulos (1986), [32] megmutatták a W-teszt aszimptotikus ekvivalenciáját egy másik korrelációteszttel, amely teszt határeloszlása akkor már ismert volt.

Ezek a problémák a W-teszt módosításaihoz vezettek. Az első példányai ezeknek a próbálkozásoknak a *D'Agostino* (1971), [12] és a *Shapiro–Francia-korrelációtesztek* (1972), [39], melyek használatát 50-nél nagyobb elemű minták esetén javasolták. A D'Agostino-tesztstatisztika a

$$D := \frac{\sum_{i=1}^n (i - \frac{n+1}{2})X_{i,n}}{n^2 S_n}\,, \tag{36}$$

a Shapiro–Francia-tesztstatisztika pedig a

$$W' := \frac{(\mathbf{m}^\top \mathbf{X}_n)^2}{\mathbf{m}^\top \mathbf{m}\sum_i (X_i - \bar{X})^2} \tag{37}$$

formulával van definiálva. Mindkét cikk szimulációs tanulmánya azt sugallta, hogy ezen tesztek aszimptotikusan ekvivalensek a W-teszttel.

A W' további egyszerűsítését javasolta Weisberg és Bingham (1975), [51]. Az $\mathbf{m}$ vektort helyettesítsük az $\tilde{\mathbf{m}} = (\tilde{m}_1, \ldots, \tilde{m}_n)$ vektorral, ahol

$$\tilde{m}_i = \Phi^{-1}\left(\frac{i - 3/8}{n + 1/4}\right), \quad i = 1, \ldots, n. \tag{38}$$

Ez a statisztika még könnyebben számolható, mint W', valamint Weisberg és Bingham empirikus vizsgálata szerint aszimptotikusan ekvivalens a W statisztikával.

A következő fontos változata a W-tesztnek *de Wet és Venter korrelációtesztje* (1972), [16]. Az ő tesztstatisztikájuk

$$W^* := \sum_{i=1}^{n} \left(\frac{X_{i,n} - \bar{X}_n}{S_n} - \Phi^{-1} \left(\frac{i}{n+1} \right) \right)^2 . \tag{39}$$

Azon túl, hogy ők vezették be a korrelációteszt fogalmát, ez volt az első olyan típusú normalitásteszt, amely határeloszlását is sikerült meghatározni. De Wet és Venter megmutatták, hogy ha $Z_1, Z_2, \ldots$ független, standard normális véletlen változók sorozata, akkor

$$W^* - a_n \xrightarrow{\mathcal{D}} \sum_{i=3}^{\infty} \frac{Z_i^2 - 1}{i} , \tag{40}$$

megfelelő $(a_n)_{n=1,2\ldots}$ konstansok sorozatára. Ezzel a tétellel megnyílt a lehetőség arra, hogy más korreláció normalitástesztek határeloszlását megkaphatjuk a W^*-teszttel való aszimptotikus ekvivalencia által. Fontos lépés volt ebben a programban Verril és Johnson (1987), [47] eredménye, ahol megmutatták a korrelációtesztek bizonyos általános feltételek melletti aszimptotikus ekvivalenciáját. Így vált világossá, hogy a Shapiro–Francia- és a Weisberg–Bingham-tesztek határeloszlása megegyezik a de Wet–Venter-teszt határeloszlásával. Továbbá a Wilk–Shapiro- és Shapiro–Francia-tesztek aszimptotikus ekvivalenciájából következett a kiindulási W-teszt határeloszlásának ismerete.

Hivatkozások

[1] M. M. Ali. Stochastic ordering and kurtosis measure. *Journal of the American Statistical Association*, 69:543–545, 1974.

[2] T. W. Anderson and D. A. Darling. Asymptotic theory of certain „goodness of fit” criteria based on stochastic processes. *Ann. Math. Statistics*, 23:193–212, 1952.

[3] P. Billingsley. Convergence of probability measures. New York-London-Sydney-Toronto: John Wiley and Sons, Inc. XII, 1968.

[4] M. Burke, M. Csörgő, S. Csörgő, and P. Révész. Approximations of the empirical process when parameters are estimated. *Ann. Probab.*, 7(5):790–810, 1979.

[5] H. Chernoff and E. Lehmann. The use of maximum likelihood estimates in χ^2 tests for goodness of fit. *Ann. Math. Stat.*, 25:579–586, 1954.

[6] W. Cochran. The χ^2 test of goodness of fit. *Annals of Mathematical Statistics*, 23:315–345, 1952.

[7] H. Cramér. On the composition of elementary errors. I. Mathematical deductions. II. Statistical applications. *Skand. Aktuarietidskr.*, 11:13–74, 141–180, 1928.

[8] S. Csörgő. Limit behaviour of the empirical characteristic function. *Ann. Probab.*, 9:130–144, 1981.

[9] S. Csörgő. The empirical characteristic process when parameters are estimated. Contributions to probability, Collect. pap. dedic. E. Lukacs, 215-230 , 1981.

[10] S. Csörgő. Testing for normality in arbitrary dimension. *Annals of Statistics*, 14:708–723, 1986.

[11] S. Csörgő. Consistency of some tests for multivariate normality. *Metrika*, 36:107–116, 1989.

[12] R. B. D'Agostino. An omnibus test of normality for moderate and large sample sizes. *Biometrika*, 58:341–348, 1971.

[13] D. Darling. The Cramér-Smirnov test in the parametric case. *Ann. Math. Stat.*, 26:1–20, 1955.

[14] F. David and N. Johnson. The probability integral transformation when parameters are estimated from the sample. *Biometrika*, 35:182–190, 1948.

[15] H. David, H. Hartley, and E. Pearson. The distribution of the ratio, in a single normal sample, of range to standard deviation. *Biometrika*, 41:482–493, 1954.

[16] T. de Wet and J. Venter. Asymptotic distributions of certain test criteria of normality. *S. Afr. Stat. J.*, 6:135–149, 1972.

[17] E. del Barrio, J. A. Cuesta-Albertos, and C. Matrán. Contributions of empirical and quantile processes to the asymptotic theory of goodness-of-fit tests. *Test*, 9(1):1–96, 2000. With discussion.

[18] M. D. Donsker. An invariance principle for certain probability limit theorems. *Mem. Am. Math. Soc.*, 6:12, 1951.

[19] J. L. Doob. Heuristic approach to the Kolmogorov-Smirnov theorems. *Ann. Math. Stat.*, 20:393–403, 1949.

[20] J. Durbin. Weak convergence of the sample distribution function when parameters are estimated. *Ann. Stat.*, 1:279–290, 1973.

[21] T. Epps and L. B. Pulley. A test for normality based on the empirical characteristic function. *Biometrika*, 70:723–726, 1983.

[22] W. Feller. On the Kolmogorov-Smirnov limit theorems for empirical distributions. *Ann. Math. Stat.*, 19:177–189, 1948.

[23] A. Feuerverger and R. A. Mureika. The empirical characteristic function and its applications. *Ann. Stat.*, 5:88–97, 1977.

[24] R. A. Fisher. The moments of the distribution for normal samples of measures of departure from normality. *Proc. R. Soc. Lond., Ser. A*, 130:16–28, 1930.

[25] R. Geary. Testing for normality. *Biometrika*, 34:209–242, 1947.

[26] E. Gumbel. On the reliability of the classical chi-square test. *Ann. Math. Stat.*, 14:253–263, 1943.

[27] P. Hall and A. H. Welsh. A test for normality based on the empirical characteristic function. *Biometrika*, 70:485–489, 1983.

[28] M. Kac, J. Kiefer, and J. Wolfowitz. On tests of normality and other tests of goodness of fit based on distance methods. *Ann. Math. Stat.*, 26:189–211, 1955.

[29] A. Kolmogorov. Sulla determinazione empirica di una legge di distribuzione. *Giornale del Istituto Italiano degli Attuari*, 4:83–91, 1933.

[30] É. Krauczi. A study of the quantile correlation test of normality. *Test*, 18(1):156–165, 2009.

[31] V. LaRiccia and D. M. Mason. Cramér-von Mises statistics based on the sample quantile function and estimated parameters. *J. Multivariate Anal.*, 18:93–106, 1986.

[32] J. Leslie, M. Stephens, and S. Fotopoulos. Asymptotic distribution of the Shapiro-Wilk W for testing for normality. *Ann. Stat.*, 14:1497–1506, 1986.

[33] H. W. Lilliefors. On the kolmogorov-smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, 62:399–402, 1967.

[34] H. Mann and A. Wald. On the choice of the number of class intervals in the application of the chi square test. *Ann. Math. Stat.*, 13:306–317, 1942.

[35] E. Pearson, R. D'Agostino, and K. Bowman. Tests for departure from normality: Comparison of powers. *Biometrika*, 64:231–246, 1977.

[36] E. S. Pearson. A further development of tests for normality. *Biometrika*, 22:239–249, 1930.

[37] D. Pollard. The minimum distance method of testing. *Metrika*, 27:43–70, 1980.

[38] M. W. Shapiro, S.S. and H. Chen. An approximate analysis of variance test for normality. *Journal of the American Statistical Association*, 63:1343–72, 1968.

[39] S. Shapiro and R. Francia. An approximate analysis of variance test for normality. *Journal of the American Statistical Association*, 67:215–216, 1972.

[40] S. Shapiro and M. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52:591–611, 1965.

[41] N. Smirnov. Sur la distribution de ω^2 (Critérium de M.R. von Mises). *Comptes Rendus de l'Académie des Sciences Paris*, 202:449–452, 1936.

[42] N. Smirnov. Sur la distribution de ω^2 (Critérium de M.R. von Mises). *Matematicheskij Sbornik*, 2:973–993, 1937.

[43] N. Smirnov. Sur les écarts de la courbe de distribution empirique. *Matematicheskij Sbornik*, 6:3–26, 1939.

[44] N. Smirnov. Approximate laws of distribution of random variables from empirical data. *Uspekhi Matematicheskikh Nauk*, 10:179–206, 1941.

[45] M. A. Stephens. Edf statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, 69:730–737, 1974.

[46] S. Sukhatme. Fredholm determinant of a positive definite kernel of a special type and its application. *Ann. Math. Stat.*, 43:1914–1926, 1972.

[47] S. Verrill and R. Johnson. The asymptotic equivalence of some modified shapiro-wilk statistics - complete and censored sample cases. *Annals of Statstics*, 15:413–419, 1987.

[48] R. von Mises. *Wahrscheinlichkeitsrechnung*. Wein, Leipzig, 1931.

[49] G. Watson. The χ^2 goodness-of-fit test for normal distributions. *Biometrika*, 44:336–348, 1957.

[50] G. Watson. On chi-square goodness-of-fit tests for continuous distributions. *J. R. Stat. Soc., Ser. B*, 20:44–72, 1958.

[51] S. Weisberg and B. C. An approximate analysis of variance test for non-normality suitable for machine calculation. *Technometrics*, 17:133–134, 1975.

[52] P. Williams. Note on the sampling distribution of $\sqrt{\beta_1}$ where the population is normal. *Biometrika*, 27:269–271, 1935.