AZ $L(z)$ VALÓSZÍNŰSÉG-ELOSZLÁSFÜGGVÉNYRŐL

RÉNYI ALFRÉD

A valószínűségszámításban és a matematikai statisztikában, különösen a rendezett minták elméletében szerepet játszik az

$$L(z) = \frac{4}{\pi} \sum_{k=0}^{\infty} (-1)^k \frac{e^{-\frac{(2k+1)^2\pi^2}{8z^2}}}{2k+1} \qquad (z > 0) \tag{1}$$

eloszlásfüggvény. Három tételt említünk meg, amelyekben ez a függvény előfordul.

1° ERDŐS PÁL és M. KAC 1946-ban bebizonyították ([1]), hogy ha $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ egyforma eloszlású független valószínűségi változók, amelyeknek várható értékük 0 és szórásuk 1, továbbá

$$\eta_n = \xi_1 + \xi_2 + \ldots + \xi_n \quad \text{és} \quad \zeta_n = \max_{1 \leq k \leq n} |\eta_k|,$$

akkor $\zeta_n/\sqrt{n}$ eloszlásfüggvénye $L(z)$-hez konvergál, ha $n \to \infty$.

2° M. KAC 1949-ben kimutatta ([2]), hogy az $L(z)$ függvény adja meg

$$\sqrt{\mu} \sup_{-\infty < x < \infty} |F_\nu(x) - F(x)|$$

határeloszlását, ha $F_\nu(x)$ egy $F(x)$ folytonos eloszlásfüggvényű valószínűségi változó értékeiből vett ν-elemű $(\xi_1, \xi_2, \ldots, \xi_\nu)$ minta empirikus eloszlásfüggvénye, ahol ν maga is valószínűségi változó, amely független a $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ változóktól, és μ várható értékű Poisson-eloszlással rendelkezik: a határeloszlás a $\mu \to \infty$ esetére értendő.

3° A rendezett minták elméletére vonatkozó dolgozatomban ([3]) bebizonyítottam, hogy ha $F_n(x)$ egy $F(x)$ folytonos eloszlásfüggvényű valószínűségi változó értékeire vonatkozó n elemű minta empirikus eloszlásfüggvénye, akkor

$$\sqrt{n} \sup_{0 < a \leq F(x)} \left| \frac{F_n(x) - F(x)}{F(x)} \right|$$

eloszlásfüggvénye, ha $n \to \infty$, az $L\left(z\sqrt{a/(1-a)}\right)$ függvényhez konvergál. $L\left(z\sqrt{a/(1-a)}\right)$ értékeinek táblázata megtalálható J. JANKO [11] táblázatgyüjteményében is (246—247. oldalak). $L(z)$ numerikus meghatározása z kicsiny értékeire (pl. ha $|z| < 3$) az (1) sor első néhány tagjának kiszámításával igen jó pontossággal elvégezhető; ha azonban z értéke nagy, az (1) sor konvergenciája lassú. Ezért bír érdekességgel, hogy $L(z)$-re olyan előállítást találjunk, amely viszont éppen z nagy értékeire teszi lehetővé $L(z)$ értékének viszonylag kevés fáradsággal, nagy pontossággal való meghatározását.[1] Az alábbiakban $L(z)$ egy ilyen előállításával foglalkozunk.

A szóban forgó előállítás a következő:

$$L(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} \operatorname{sgn}\left(\cos \frac{\pi y}{2z}\right) dy \,, \tag{2}$$

ahol $z > 0$ tetszőleges, és $\operatorname{sgn} x$ az x szám előjele, vagyis

$$\operatorname{sgn} x = \begin{cases} 1, & \text{ha} \quad x > 0 \\ 0, & \text{ha} \quad x = 0 \\ -1, & \text{ha} \quad x < 0\,. \end{cases}$$

A (2) jobboldalán álló integrál értéke a

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du \tag{3}$$

normális eloszlásfüggvény értéktáblázata segítségével igen könnyen kiszámítható; e célból célszerű a (2) azonosságot a következő alakra hozni:[2]

$$L(z) = \sum_{k=-\infty}^{+\infty} (-1)^k \{\Phi((2k+1)z) - \Phi((2k-1)z)\} \,. \tag{4}$$

A (2), illetve (4) azonosságra alábbiakban három egyszerű bizonyítást adunk, amelyek tulajdonképpen ugyanannak a bizonyításnak különböző változatai.

A (3) képlet módot nyújt az $L(z)$ eloszlás momentumainak meghatározására is, ezért (2) bebizonyítása után röviden kitérünk $L(z)$ momentumainak (2) alapján való kiszámítására is.

Első bizonyítás. Az első bizonyítás, amit bemutatunk, a ϑ-függvényekre vonatkozó ismert sorfejtések és transzformációs képletek felhasználásán alapszik. Könnyen belátható ugyanis, hogy

$$L(z) = \int_{-\frac{1}{2}}^{+\frac{1}{2}} \vartheta_2\left(v \,\middle|\, \frac{i\pi}{2z^2}\right) dv \,, \tag{5}$$

[1] A (4) előállítás felhasználásra került az $L(z\sqrt{a/(1-a)})$ függvény [11]-ben szereplő táblázatának elkészítésénél. A numerikus számításokat Palásti Ilona végezte.

[2] $L(z)$ (4) alatti alakja nem tekinthető újnak, megtalálható pl. [10]-ben. Azonban az irodalomban nem találtam meg $L(z)$ (1), illetve (4) alakjai azonosságának közvetlen és egyszerű bizonyítását.

ahol (lásd: [4], 195. oldal) $\mathfrak{Im}\ \tau > 0$ esetében

$$\vartheta_2(v|\tau) = 2\sum_{k=1}^{\infty} e^{\pi\tau\left(\frac{2k+1}{2}\right)^2} \cos(2k+1)\pi v \,. \tag{6}$$

Mármost felhasználva a

$$\vartheta_2(v|\tau) = \sqrt{\frac{i}{\tau}}\, e^{-\frac{i\pi v^2}{\tau}}\, \vartheta_0\left(\frac{v}{\tau}\middle| -\frac{1}{\tau}\right) \tag{7}$$

ismert transzformációs képletet (lásd [4], 248. oldal), ahol

$$\vartheta_0(v|\tau) = 1 + 2\sum_{k=1}^{\infty} (-1)^k e^{i\pi\tau k^2} \cos 2k\pi v \,, \tag{8}$$

ha $\mathfrak{Im}\ \tau > 0$, és tagonként integrálva nyerjük a (4) és abból a (2) azonosságot.

Második bizonyítás. A (2) képlet egy másik bizonyítását is bemutatjuk, ami a hővezetési egyenlet elméletének egy jólismert tételén alapszik.

Keressük meg azt az $u(x, t)$ függvényt, amely a $-\infty < x < +\infty$, $t > 0$ félsíkban x szerint kétszer, t szerint pedig egyszer folytonosan deriválható, korlátos, eleget tesz a

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} \tag{9}$$

hővezetési egyenletnek és a

$$\lim_{t\to+0} u(x, t) = \operatorname{sgn}(\cos \pi x) \qquad -\infty < x < +\infty \tag{10}$$

peremfeltételnek. Ismeretes (lásd: [5], III. tétel), hogy az előírt feltételek az $u(x, t)$ függvényt egyértelműen meghatározzák. Mármost nem nehéz verifikálni, hogy az

$$U(x, t) = \frac{4}{\pi}\sum_{k=0}^{\infty} (-1)^k \frac{\cos(2k+1)\pi x}{2k+1}\, e^{-\frac{\pi^2 t(2k+1)^2}{2}} \tag{11}$$

függvény eleget tesz a (9) és (10) feltételeknek. Másrészt ismeretes (lásd: [5]), hogy a (9) egyenletnek eleget tevő $u(x, t)$ függvény az x-tengely menti peremértékei segítségével általában a következőképpen fejezhető ki:

$$u(x, t) = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{2t}}\, u(y, 0)\, dy \,. \tag{12}$$

Behelyettesítve (12) jobboldalán az $u(y, 0) = \operatorname{sgn}(\cos \pi y)$ függvényt, nyerjük, hogy a (11) alatti $U(x, t)$ függvény a következőképpen fejezhető ki:

$$U(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} \operatorname{sgn}\left[\cos \pi(x + y\sqrt{t})\right] dy \,. \tag{13}$$

Összehasonlítva (11)-et és (13)-at, következik, hogy

$$(14)\qquad \frac{4}{\pi}\sum_{k=0}^{\infty}\frac{(-1)^k\cos(2k+1)\pi x}{2k+1}\,e^{-\frac{\pi^2 t(2k+1)^2}{2}} =$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\,\mathrm{sgn}\left[\cos\pi(x+y\sqrt{t})\right]dy\,.$$

Helyettesítve (14)-be az $x = 0$ és $t = 1/4z^2$ értékeket, kapjuk a (2) azonosságot. A (14) azonosság egyben (2) általánosításának is tekinthető.

Harmadik bizonyítás. Végiggondolva a második bizonyítást, azt úgy jellemezhetjük, hogy a (9) hővezetési egyenletet a (10) peremfeltétel mellett kétféleképpen oldottuk meg: először a megoldást az

$$e^{-\frac{k^2\pi^2 t}{2}}\cos k\pi x \quad (k = 0, 1, \ldots)$$

alapmegoldásokból, másodszor az

$$\frac{1}{\sqrt{2\pi t}}\,e^{-\frac{(x-a)^2}{2t}} \quad -\infty < a < +\infty$$

alapmegoldásokból raktuk össze, és a két eredmény azonosságából következtettünk a (2) azonosság fennállására. Ez a módszer általános esetben is alkalmazható. Ha $f(x)$ egy tetszőleges 2π periódusú integrálható függvény, amelynek Fourier-sora

$$f(x) \sim \sum_{n=0}^{\infty}(a_n\cos nx + b_n\sin nx)\,,$$

akkor

$$(15)\qquad \sum_{n=0}^{\infty}(a_n\cos nx + b_n\sin nx)\,e^{-\frac{n^2 t}{2}} = \frac{1}{\sqrt{2\pi t}}\int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{2t}} f(y)\,dy\,.$$

A (15) azonosság Fouriertől származik ([6]). (15) jobboldalán $f(x)$ Fourier-sorának úgynevezett $(A, 2)$ közepei állnak ([7]). Weierstrass a folytonos függvények trigonometrikus polinomokkal való egyenletes approximálhatóságára vonatkozó híres tételét éppen a (15) azonosság segítségével bizonyította be. (Lásd: [8], Vol. 3., p. 20.) (2) tehát felfogható, mint a Fourier-féle (15) azonosság speciális esete is.

A momentumok kiszámítása

(4)-ből könnyen következik, hogy

$$(16)\qquad L'(z) = l(z) = \frac{4}{\sqrt{2\pi}}\sum_{k=0}^{\infty}(-1)^k(2k+1)\,e^{-\frac{(2k+1)^2 z^2}{2}}\,.$$

Másrészt (1)-ből

$$l(z) = \frac{\pi}{z^3} \sum_{k=0}^{\infty} (-1)^k (2k+1)\, e^{-\frac{(2k+1)^2\pi^2}{8z^2}}. \tag{17}$$

Tehát a

$$\lambda(z) = l(z)\, z^{3/2} \tag{18}$$

függvény eleget tesz a

$$\lambda(z) = \lambda\left(\frac{\pi}{2z}\right) \tag{19}$$

függvényegyenletnek. (Ezt itt csak közbevetőleg jegyeztük meg; (19)-re a momentumok meghatározásához nincs szükségünk.)

A (16) képlet alkalmas $L(z)$ momentumainak kiszámítására. Ugyanis (16)-ból

$$M_n = \int_0^{\infty} z^n l(z)\, dz = \frac{2^{\frac{n}{2}+1}\,\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi}} \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)^n}. \tag{20}$$

Speciálisan

$$M_1 = \sqrt{\frac{\pi}{2}} \tag{21}$$

és

$$M_2 = 2G \tag{22}$$

ahol

$$G = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)^2} = 0{,}915965594\ldots. \tag{23}$$

az úgynevezett CATALAN-féle állandó (lásd: [9], II. kötet, 2. oldal), és így a szórásnégyzet;

$$D^2 = 2G - \frac{\pi}{2}. \tag{24}$$

(Beérkezett: 1957. II. 15.)

IRODALOM

[1] ERDŐS, P.—KAC, M.: „On certain limit theorems of the theory of probability." *Bulletin of the American Mathematical Society* **52** (1946) 292—302.

[2] KAC, M.: „On deviations between theoretical and empirical distributions." *Proceedings of the National Academy of Sciences of the USA* **35** (1949) 252—257.

[3] RÉNYI, A.: „On the theory of order statistics." *Acta Mathematica Academiae Scientiarum Hungaricae* **4** (1953) 191—231.

[4] Hurwitz, A.—Courant, R.: *Funktionentheorie*. Springer, Berlin, 1939.
[5] Czipszer J.: „Hővezetés a végtelen rúdban, I." *A Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei* **3** (1954) 395—408. és II., *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei* **1** (1956) 185—192.
[6] Fourier, J. B.: „Théorie analytique de la chaleur." *Oeuvres de Fourier,* Vol. 1. Paris, Gauthier-Villars, 1888, pp. 426—427.
[7] Hardy, G. H.: *Divergent series*. Oxford, 1949.
[8] Weierstrass, K.: „Über die analytische Darstellbarkeit sogenannter willkürlichen Funktionen reeller Argumente." *Mathematische Werke von K. Weierstrass,* Mayer—Müller, Berlin, 1903.
[9] Gröbner, W.—Hofreiter, W.: *Integraltafeln*. Springer, Wien, 1950.
[10] Grenander, U.—Rosenblatt, M.: „Statistical spectral analysis of time series arising from stationary stochastic processes." *Annals of Mathematical Statistics* **24** (1953) 537—558; „Comments on statistical spectral analysis." *Skandinavisk Aktuarietidskrift* **36** (1953) 82—202.
[11] Janko, J.: *Statisticke tabulky*. Československé Akademie Věd, Praha, 1958.

О ФУНКЦИИ РАСПРЕДЕЛЕНИЯ $L(z)$

A. RÉNYI

Резюме

В теории вероятностей и математической статистики функция распределения

$$L(z)=\frac{4}{\pi}\sum_{k=0}^{\infty}(-1)^k\,\frac{e^{-\frac{(2k+1)^2\pi^2}{8z^2}}}{2k+1} \tag{1}$$

играет известный роль. (См. напр. [1], [2], [3].)

Ряд (1) быстро сходится для малых z, но быстрота сходимости уменьшается если z возрастает. Поэтому нужно другое выражение для $L(z)$, которое выгодно для больших значений от z. Для этой цели служит формула

$$L(z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{y^2}{2}}\operatorname{sgn}\left(\cos\frac{\pi y}{2z}\right)dy \tag{2}$$

($z>0$) (где sgn a означает знак от a), которое может быть преображено также в

$$L(z)=\sum_{k=-\infty}^{\infty}(-1)^k\{\Phi((2k+1)z)-\Phi((2k-1)z)\}, \tag{3}$$

где

$$\Phi(z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}e^{-\frac{u^2}{2}}du\,. \tag{4}$$

Хотя (3) имеется в [10], автор статы не нашёл в литературе простой вывод формулы (3) соотв. (2) из (1). В работе даны три доказательства формулы (3). Первое доказательство использует преобразование от ϑ-функций, второе формулы для решения уравнения теплопроводности и третье

тождество от Fourier: если $f(x)$ периодическая интегрируемая функция с периодом 2π, и ряд Fourier от $f(x)$ есть

$$f(x) \sim \sum_{n=0}^{\infty} (a_n \cos nx + b_n \sin nx) \tag{5}$$

то имеет место для $t > 0$

$$\sum_{n=0}^{\infty} (a_n \cos nx + b_n \sin nx)\, e^{-\frac{n^2 t}{2}} = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{-\frac{(y-x)^2}{2t}} f(y)\, dy\ . \tag{6}$$

Формула (2) получается из (6) для $f(y) = \operatorname{sgn}(\cos y)$.

В конце работы вычислены моменты функции распределения. В частности получается

$$\int_0^{\infty} z\, dL(z) = \sqrt{\frac{\pi}{2}} \quad \text{и} \quad \int_0^{\infty} z^2\, dL(z) = 2G \tag{7}$$

где G постоянное Catalan.

ON THE DISTRIBUTION FUNCTION $L(z)$

A. RÉNYI

Summary

Many problems of probability theory and statistics lead to the cumulative distribution function

$$L(z) = \frac{4}{\pi} \sum_{k=0}^{\infty} (-1)^k \frac{e^{-\frac{(2k+1)^2\pi^2}{8z^2}}}{2k+1} \qquad z > 0 \tag{1}$$

(see e. g. [1], [2], [3]). The series on the right of (1) converges rapidly for small values of z, but for large values of z the convergence is slow. Therefore an other formula for $L(z)$ is needed, which enables to calculate easily the values of $L(z)$ for large values of z. This need is met by the formula

$$L(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} \operatorname{sgn}\left(\cos \frac{\pi y}{2z}\right) dy \qquad z > 0 \tag{2}$$

where $\operatorname{sgn} x$ is the sign of x ($\operatorname{sgn} x = +1$ for $x > 0$, $\operatorname{sgn} x = 0$ for $x = 0$ and $\operatorname{sgn} x = -1$ for $x < 0$). Putting

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} du \tag{3}$$

the formula (2) may be brought also to the form

$$(4) \qquad L(z) = \sum_{k=-\infty}^{\infty} (-1)^k \{\Phi((2k+1)z) - \Phi((2k-1)z)\} .$$

By means of (4) we can calculate the value of $L(z)$ for large values of z, using a table for $\Phi(x)$.

For the identity (2) three variants of the same proof are given. (The formula (4) can be found e. g. in [10], but the author did not find in the literature a simple deduction of the formulae (2) resp. (4) from (1).)

The first proof makes use of a transformation formula for thetafunctions, the second of the representation in two different forms of the solution of the heat equation, the third is based on a classical formula, due to FOURIER [7], according to which if $f(x)$ is an integrable periodic function with period 2π, the Fourier-series of which is

$$(5) \qquad f(x) \sim \sum_{n=0}^{\infty} (a_n \cos nx + b_n \sin nx)$$

then for $t > 0$ we have

$$(6) \qquad \sum_{n=0}^{\infty} (a_n \cos nx + b_n \sin nx)\, e^{-\frac{n^2 t}{2}} = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{2t}} f(y)\, dy .$$

Applying this formula to

$$f(x) = \operatorname{sgn}(\cos x)$$

whose Fourier-series is

$$f(x) = \sum_{k=0}^{\infty} \frac{(-1)^k \cos(2k+1)x}{2k+1}$$

and putting

$$t = \left(\frac{\pi}{2z}\right)^2$$

we obtain (2) from (6) for $x = 0$.

Finally it is pointed out that (2) enables us to calculate easily the moments of the distribution function $L(z)$; we obtain

$$M_n = \int_0^{\infty} z^n\, dL(z) = \frac{2^{\frac{n}{2}+1}\, \Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi}} \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)^n} .$$

Especially $M_1 = \sqrt{\pi/2}$ and the variance of $L(z)$ is equal to $2G - \pi/2$ where G is CATALAN's constant.