

مجله علوم آماری، پاییز و زمستان ۱۳۹۵

جلد ۱۰، شماره ۲، ص ۳۲۹-۳۴۴

DOI: 10.18869/acadpub.jss.10.2.329

اثبات رابطه سرگئی وینزکی برای تقریب تابع توزیع نرمال استاندارد

شهرام منصوری

دانشکده علوم ریاضی، دانشگاه شهید بهشتی

تاریخ دریافت: ۱۳۹۴/۵/۲۳ تاریخ آخرین بازنگری: ۱۳۹۵/۲/۲۲

چکیده: در بین تمام توزیع‌های آماری توزیع نرمال استاندارد مهمترین و کاربردی‌ترین توزیع آماری بوده و محاسبه سطح زیر منحنی چگالی و تابع توزیع آن مورد نیاز است. ضابطه این تابع بصورت یک انتگرال معین بیان می‌شود، ولی متاسفانه تابع اولیه آن دارای شکل بسته و تحلیلی نیست، لذا باید آن را تقریب زد. در این مقاله رابطه تقریبی سرگئی وینزکی با یک روش جدید اثبات می‌شود، سپس این تقریب با تغییراتی در رابطه آن بهبود داده و نشان می‌دهیم حداکثر مقدار خطای آن کمتر از $0/0000584$ است. در انتها رابطه‌ای نیز برای محاسبه صدک‌های توزیع نرمال به دست آورده می‌شود.

واژه‌های کلیدی: تابع توزیع نرمال استاندارد تابع خطا، تقریب

آدرس الکترونیک مسئول مقاله: شهرام منصوری، sh_mansouri@sbu.ac.ir

کد موضوع بندی ریاضی (۲۰۱۰): 6۲E۹۹، 6۲E۱۷

تابع خطا و تابع توزیع احتمال نرمال استاندارد به ترتیب به صورت

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt; \quad -\infty < x < \infty \quad (1)$$

$$\Phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt; \quad -\infty < z < \infty$$

تعریف شده‌اند، که در احتمال، آمار و معادلات دیفرانسیل با مشتقات جزئی کاربرد دارند. به عنوان مثال معادله نفوذ $a \frac{\partial^2 U}{\partial x^2} - b \frac{\partial U}{\partial x}$ در قلمرو $t > 0, x > 0$ یا $t > 0, x \in R$ به ازای شرایط مرزی خاص بر حسب تابع خطا بیان می‌شود. به راحتی می‌توان ثابت کرد که بین این دو تابع روابط زیر برقرارند.

$$\Phi(z) = \frac{1}{2} \left(1 + erf\left(\frac{z}{\sqrt{2}}\right) \right) \quad (2)$$

$$erf(x) = 2\Phi(\sqrt{2}x) - 1$$

برخی ویژگی های تابع خطا عبارتند از

الف - $erf(x)$ تابعی فرد است.

ب - $|erf(x)| < 1$

ج - $erf(0) = 0$

د - $erf(\infty) = 1$

البته می‌توان مقادیر این توابع که حاصل یک انتگرال معین هستند را با روش‌های تقریبی انتگرال‌گیری متداول تا هر دقت لازم و مورد نیاز و تا هر چند رقم اعشار با استفاده از نرم افزارهایی مانند Maple، Matlab و S^+ حساب کرد. بدین ترتیب با یک تابع به صورت جدول نمایش داده شده مواجه هستیم. در صورتی که رابطه وینزکی یا وارون آن یک رابطه ساده دستی است که بر حسب تابع معمولی ریاضی معرفی شده و می‌توان آن را در هر فرمولی که $erf(x)$ یا $erf^{-1}(x)$ ظاهر شده جایگزین کرد. علاوه بر این محققین از زمان‌های گذشته و حتی در سال‌های اخیر نیز علاقمند به یافتن توابعی هستند که آنها را با این توابع تقریب بزنند، یا کران‌های دقیق‌تری برای آنها به دست آورند. برخی از این توابع عبارتند از

(۱) کران‌های گوردون (۱۹۴۱):

$$\frac{z \exp\{-\frac{z^2}{2}\}}{z^2 + 1} \leq \int_z^\infty e^{-\frac{1}{2}t^2} dt \leq \frac{\exp\{-\frac{z^2}{2}\}}{z}, \quad z > 0$$

(۲) کران‌های پولاک (۱۹۵۶):

$$\frac{2 \exp\{-\frac{z^2}{2}\}}{z + (z^2 + 4)^{\frac{1}{2}}} \leq \int_z^\infty e^{-\frac{1}{2}t^2} dt \leq \frac{2 \exp\{-\frac{z^2}{2}\}}{z + (z^2 + \frac{4}{\pi})^{\frac{1}{2}}}, \quad z > 0$$

(۳) کران‌های فلر (۱۹۶۸):

$$\left(\frac{1}{z} - \frac{1}{z^2}\right) \sqrt{\frac{\pi}{2}} \exp\{-\frac{z^2}{2}\} < 1 - \Phi(z) < \frac{1}{z} \sqrt{\frac{\pi}{2}} \exp\{-\frac{z^2}{2}\}$$

(۴) کران‌های سزارگ (۱۹۹۹):

$$\frac{2 \exp\{-\frac{z^2}{2}\}}{z + (z^2 + 4)^{\frac{1}{2}}} \leq \int_z^\infty e^{-\frac{1}{2}t^2} dt \leq \frac{2 \exp\{-\frac{z^2}{2}\}}{2z + (z^2 + 8)^{\frac{1}{2}}}, \quad z > -1$$

(۵) تقریب تاچر (۱۹۶۳):

$$\Phi(z) \approx \frac{\exp\{2kz\}}{1 + \exp\{2kz\}}; \quad k = \sqrt{\frac{2}{\pi}}$$

(۶) تقریب زلان و سرورو (۱۹۶۴):

$$\Phi(z) \approx 1 - \left(0.4371t - 0.1201t^2 + 0.9373t^3\right) \frac{1}{\sqrt{2\pi}} \exp\{-\frac{z^2}{2}\}$$

که در آن $t = \frac{1}{1 + 0.7374z}$ است.

۳۳۲ اثبات رابطه سرگئی وینزکی برای تقریب تابع توزیع نرمال استاندارد

(۷) تقریب پیچ (۱۹۷۷):

$$\Phi(z) \approx \frac{1}{\sqrt{\pi}}(1 + \tanh y), \quad y = \sqrt{\frac{2}{\pi}}z(1 + 0.044715z^2)$$

(۸) تقریب هاماکار (۱۹۷۸):

$$\Phi(z) \approx 1 - \frac{1}{\sqrt{\pi}}\{1 - \sqrt{1 - e^{-y^2}}\}, \quad y = 0.0806z(1 - 0.018z^2)$$

(۹) تقریب ابرناتی (۱۹۸۸):

$$\Phi(z) \approx \frac{1}{\sqrt{\pi}} + \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n+1}}{2^n n! (2n+1)}, \quad z > 0$$

(۱۰) تقریب لین (۱۹۸۹):

$$\Phi(z) \approx 1 - \frac{1}{\sqrt{\pi}}\{\exp\{-0.717z - 0.416z^2\}\}, \quad z > 0$$

(۱۱) تقریب لین (۱۹۹۰):

$$\Phi(z) \approx 1 - \frac{1}{1 + e^y}, \quad y = 4/2\pi\left(\frac{z}{9-z}\right), \quad z > 0$$

(۱۲) تقریب بگی (۱۹۹۵):

$$\Phi(z) \approx \frac{1}{\sqrt{\pi}} + \frac{1}{\sqrt{\pi}}\left(1 - \frac{1}{3.0}\left(7e^{-\frac{1}{2}z^2} + 16e^{-(2-\sqrt{2})z^2} + \left(7 + \frac{\pi z^2}{4}\right)e^{-z^2}\right)\right)^{\frac{1}{2}}, \quad z > 0$$

(۱۳) تقریب بیرک (۲۰۰۱):

$$\Phi(z) \approx 1 - \frac{z^2 + 5/57519z + 12/77437}{z^2\sqrt{2\pi} + 14/38718z^2 + 31/53531z + 25/548} \exp\left\{-\frac{z^2}{2}\right\}$$

(۱۴) تابع توزیع تجمعی لوژستیک استاندارد:

$$\Phi(z) \approx \frac{1}{1 + \exp\left\{\frac{-\pi z}{\sqrt{3}}\right\}}, \quad z > 0$$

(۱۵) تقریب آمیت چادھاری (۲۰۱۴):

$$\Phi(z) \approx 1 - \frac{1}{\sqrt{2\pi} \cdot 0/226 + 0/64z + 0/33\sqrt{z^2 + 3}} \exp\left\{-\frac{z^2}{2}\right\}, \quad z > 0$$

در جداول ۱ و ۲ به ترتیب ماکسیمم قدر مطلق خطا و میانگین قدر مطلق خطاها برای برخی روش‌های تقریب تابع توزیع نرمال استاندارد ارائه شده است (چادھاری، ۲۰۰۷). به وضوح دیده می‌شود با توجه به اینکه در بازه‌های مختلف دقت روش‌های تقریب متفاوت می‌باشد لذا بهترین تقریب به‌طور یکنواخت وجود ندارد. در این مقاله رابطه تقریبی سرگئی وینزکی با یک روش جدید در بخش ۲ اثبات می‌شود سپس در بخش ۳ مقدار خطا را در هر نقطه تعیین و نمودار آن رسم شده است در بخش ۴ به بحث و نتیجه‌گیری پرداخته خواهد شد.

۲ اثبات رابطه تقریبی سرگئی وینزکی

در این بخش ابتدا تعریف، لم و نتایج لازم برای اثبات رابطه تقریبی سرگئی وینزکی ارائه می‌شوند.

تعریف ۱: نگارش $f(x) = O(g(x)), x \in D$ بدان معنا است که ثابتی مانند A وجود دارد، به قسمی که $|f(x)| \leq A|g(x)|, x \in D$

جدول ۱: ماکسیمم قدر مطلق خطا برای روش‌های مختلف تقریب

روش تقریبی	$0 < z < 1$	$1 < z < 3$	$3 < z < 4$
تاچر	$9/919 \times 10^{-3}$	$1/767 \times 10^{-2}$	$6/912 \times 10^{-2}$
زلان و سرورو	$1/120 \times 10^{-5}$	$1/095 \times 10^{-5}$	$4/990 \times 10^{-6}$
پاج	$1/530 \times 10^{-5}$	$1/791 \times 10^{-4}$	$1/373 \times 10^{-4}$
هاماکار	$6/229 \times 10^{-4}$	$3/852 \times 10^{-4}$	$2/800 \times 10^{-6}$
لین (۸۹)	$6/585 \times 10^{-3}$	$2/374 \times 10^{-3}$	$2/690 \times 10^{-5}$
لین (۹۰)	$6/688 \times 10^{-3}$	$2/538 \times 10^{-3}$	$1/220 \times 10^{-5}$
بگی	$3/030 \times 10^{-5}$	$2/960 \times 10^{-5}$	$2/710 \times 10^{-6}$
بیرک	$1/185 \times 10^{-5}$	$1/873 \times 10^{-5}$	$2/051 \times 10^{-6}$
لوژستیک	$2/266 \times 10^{-2}$	$1/846 \times 10^{-2}$	$2/063 \times 10^{-3}$

جدول ۲: میانگین قدر مطلق خطا برای روش‌های مختلف تقریب در بازه ۴ - ۰

روش تقریبی	میانگین قدر مطلق خطا
تاچر	$8/592 \times 10^{-3}$
زلان و سرورو	$5/980 \times 10^{-6}$
پاج	$9/470 \times 10^{-5}$
هاماکار	$1/682 \times 10^{-4}$
لین (۸۹)	$1/342 \times 10^{-3}$
لین (۹۰)	$1/365 \times 10^{-3}$
بگی	$1/160 \times 10^{-5}$
بیرک	$6/921 \times 10^{-6}$
لوژستیک	$7/311 \times 10^{-3}$

به سادگی ثابت می‌شود، اگر $\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)}$ متنهایی باشد، آنگاه

$$f(x) = O(g(x))$$

همچنین هرگاه $x \rightarrow \infty$ ، آنگاه

$$CO(x^n) = O(x^n) \quad (۳)$$

$$O(x^n) + O(x^m) = O(x^n) \quad n > m \quad (۴)$$

$$O(x^n)O(x^m) = O(x^{n+m}) \quad (۵)$$

لم ۱: برای مقادیر مثبت و بزرگ x داریم

$$erf(x) = 1 - \frac{e^{-x^2}}{\sqrt{\pi}x} + O\left(\frac{e^{-x^2}}{x^3}\right)$$

برهان: از رابطه (۱) داریم

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^\infty e^{-t^2} dt - \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt \quad (۶)$$

مقدار انتگرال اول رابطه (۶) برابر $erf(\infty) = 1$ است، با اعمال تغییر متغیر $t = v^2$ در انتگرال دوم، خواهیم داشت

$$erf(x) = 1 - \frac{1}{\sqrt{\pi}} \int_{x^2}^\infty \frac{e^{-t}}{\sqrt{t}} dt$$

حال با روش جزء به جزء به دست می‌آوریم

$$erf(x) = 1 - \frac{1}{\sqrt{\pi}} \left(\frac{e^{-x^2}}{x} - \frac{1}{2} \int_{x^2}^\infty \frac{e^{-t}}{t^{3/2}} dt \right) \quad (۷)$$

ولی به ازای $t > x^2$ ، $\frac{1}{t^{3/2}} < \frac{1}{x^3}$ بنابراین

$$0 < \int_{x^2}^\infty \frac{e^{-t}}{t^{3/2}} dt < \frac{1}{x^3} \int_{x^2}^\infty e^{-t} dt = \frac{e^{-x^2}}{x^3}$$

در نتیجه

$$\frac{1}{2\sqrt{\pi}} \int_{x^2}^\infty \frac{e^{-t}}{t^{3/2}} dt = O\left(\frac{e^{-x^2}}{x^3}\right) \quad (۸)$$

باتوجه به (۷) و (۸) لم ۱ ثابت می‌شود.

۳۳۶ اثبات رابطه سرگئی وینزکی برای تقریب تابع توزیع نرمال استاندارد

تذکر ۱ : از لم ۱ نتیجه می‌گردد که برای مقادیر بزرگ x داریم:

$$\operatorname{erf}(x) \cong 1 + O\left(\frac{e^{-x^2}}{x}\right) \quad (9)$$

طرفین رابطه (۹) را به توان دو می‌رسانیم با استفاده از روابط (۳)، (۴) و (۵) خواهیم داشت:

$$(\operatorname{erf}(x))^2 \cong 1 + O\left(\frac{e^{-x^2}}{x}\right) \quad (10)$$

لم ۲:

$$\operatorname{erf}(x)^2 = \frac{4}{\pi} \left(x^2 - \frac{2x^4}{3} + O(x^6) \right) \quad (11)$$

برهان : با استفاده از بسط $e^{-t} = 1 - \frac{t}{1!} + \frac{t^2}{2!} - \dots$ می‌توان نوشت

$$\begin{aligned} \operatorname{erf}(x) &= \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \\ &= \frac{2}{\sqrt{\pi}} \int_0^x \left(1 - \frac{t^2}{1} + \frac{t^4}{2!} - \dots \right) dt \\ &= \frac{2}{\sqrt{\pi}} \left(x - \frac{x^3}{3} + \frac{x^5}{10} - \dots \right) \end{aligned}$$

از توان دوم طرفین این عبارت نتیجه حاصل می‌شود.

لم ۳:

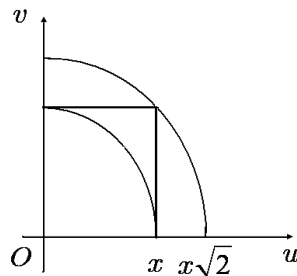
$$1 - e^{-x^2} < (\operatorname{erf}(x))^2 < 1 - e^{-2x^2} \quad (12)$$

برهان : ابتدا فرض می‌کنیم $x \geq 0$ ، با در نظر گرفتن

$$(\operatorname{erf}(x))^2 = \frac{4}{\pi} \left(\int_0^x e^{-u^2} du \right) \left(\int_0^x e^{-v^2} dv \right) = \frac{4}{\pi} \int_0^x \int_0^x e^{-(u^2+v^2)} dudv$$

و تعریف $D = \{(u, v) | 0 < u < x, 0 < v < x\}$ ، می‌توان نوشت:

$$(\operatorname{erf}(x))^2 = \frac{4}{\pi} \int \int_D e^{-(u^2+v^2)} dudv$$



شکل ۱: نواحی انتگرال گیری

نواحی D_1 و D_2 را به صورت

$$D_1 = \{(u, v) \mid 0 < u^2 + v^2 < x^2, 0 < u < x, 0 < v < x\}$$

$$D_2 = \{(u, v) \mid 0 < u^2 + v^2 < 2x^2, 0 < u < x\sqrt{2}, 0 < v < x\sqrt{2}\}$$

تعریف می کنیم. با توجه به شکل ۱ به روشنی $D_1 \subset D \subset D_2$ است. چون $e^{-(u^2+v^2)} > 0$ می توان نوشت

$$\frac{4}{\pi} \int \int_{D_1} e^{-(u^2+v^2)} dudv < \frac{4}{\pi} \int \int_D e^{-(u^2+v^2)} dudv < \frac{4}{\pi} \int \int_{D_2} e^{-(u^2+v^2)} dudv$$

با استفاده از تبدیل مختصات قطبی، انتگرال های فوق به صورت زیر ساده می شوند.

$$\frac{4}{\pi} \int \int_{D_1} e^{-(u^2+v^2)} dudv = \frac{4}{\pi} \int_0^{\frac{\pi}{4}} \int_0^x re^{-r^2} dr d\theta = 1 - e^{-x^2}$$

و به طور مشابه $\frac{4}{\pi} \int \int_{D_2} e^{-(u^2+v^2)} dudv = 1 - e^{-2x^2}$ است. چون سه جمله نابرابری (۱۲) توابعی زوج هستند به ازای $x < 0$ هم برقرار هستند به این ترتیب نامساوی (۱۲) ثابت شده است.

قضیه ۱: رابطه تقریبی (رابطه سرگئی وینزکی)

$$erf(x) \cong [1 - \exp(-x^2 \frac{\frac{4}{\pi} + ax^2}{1 + ax^2})]^{\frac{1}{2}}; \quad x \geq 0 \quad (13)$$

را برای محاسبه تابع خطا داریم، که در آن $0.14001 \cong \frac{14001}{100000} \cong \frac{14001}{100000} \cong \frac{14001}{100000}$

۳۳۸ اثبات رابطه سرگئی وینزکی برای تقریب تابع توزیع نرمال استاندارد

برهان : از نابرابری‌های (۱۲) نتیجه می‌شود که تابع $k = k(x)$ وجود دارد به طوری که $1 < k(x) < 2$ و $(erf(x))^2 = 1 - e^{-k(x)x^2}$. چون $(erf(x))^2$ تابعی زوج است، بنابراین $k(x)$ هم تابعی زوج است. از بین توابع تقریبی برای $k(x)$ تابع گویای

$$k(x) \cong \frac{Ax^2 + B}{Cx^2 + D} = \frac{\frac{A}{D}x^2 + \frac{B}{D}}{\frac{C}{D}x^2 + 1}$$

یعنی $k(x) \cong \frac{ax^2 + b}{cx^2 + 1}$ را انتخاب می‌کنیم (با این انتخاب این فرم استخراج x بر حسب $erf(x)$ نیز امکان پذیر است). پس

$$(erf(x))^2 = 1 - e^{-\frac{ax^2 + b}{1 + cx^2}x^2} \quad (14)$$

حال اعداد a, b, c را چنان انتخاب می‌کنیم که الف) وقتی $x \rightarrow \infty$ طرفین (۱۴) به طور مجانبی به هم نزدیک باشند. ب) این نزدیکی برای مقادیر کوچک x هم برقرار باشد به عبارت دیگر در بسط طرفین (۱۴) بر حسب توان‌های اولیه x با هم برابر باشند. اعمال شرط الف: می‌توان نوشت

$$\frac{(ax^2 + b)x^2}{cx^2 + 1} = \frac{a}{c}x^2 + \frac{(b - \frac{a}{c})x^2}{cx^2 + 1}$$

در نتیجه

$$\lim_{x \rightarrow \infty} \frac{(b - \frac{a}{c})x^2}{cx^2 + 1} = \frac{1}{c}(b - \frac{a}{c})$$

پس وقتی $x \rightarrow \infty$ از (۱۴) خواهیم داشت

$$1 - e^{-\frac{(ax^2 + b)x^2}{cx^2 + 1}} = 1 - e^{-\frac{a}{c}x^2} e^{-\frac{(b - \frac{a}{c})x^2}{cx^2 + 1}} = 1 + O(e^{-\frac{a}{c}x^2}) \quad (15)$$

حال نزدیکترین وضعیت عبارت‌های (۱۰) و (۱۵) به هم وقتی رخ می‌دهد که $\frac{a}{c} = 1$ یعنی $a = c$ ، و با توجه به (۱۴) داریم

$$(erf(x))^2 = 1 - e^{-\frac{b + ax^2}{1 + ax^2}x^2} \quad (16)$$

اعمال شرط ب: از سری $e^{-t} = 1 - \frac{t}{1!} + \frac{t^2}{2!} - \dots$ و رابطه (۱۶) می توان نتیجه گرفت

$$\begin{aligned} (erf(x))^2 &= 1 - [1 - \frac{1}{1!}(\frac{b+ax^2}{1+ax^2}x^2) + \frac{1}{2!}(\frac{b+ax^2}{1+ax^2}x^2)^2 - \dots] \\ &= \frac{1}{1+ax^2}(bx^2+ax^4) - \frac{1}{2} \frac{1}{(1+ax^2)^2}(bx^2+ax^4)^2 \\ &+ \dots \end{aligned} \quad (17)$$

می دانیم که با بسط طرفین رابطه (۱۷) بر حسب توان های x لازم است ضرائب مربوط به توان های اولیه x با هم برابر باشند. حال طرفین این رابطه را بسط می دهیم، با توجه به زوج بودن توان های x و اینکه به دو معادله روی a و b نیاز داریم و ثابت های دو طرف (۱۷) هم حذف می شوند کفایت که ضرائب x^2 و x^4 را مساوی هم قرار دهیم. پس عملاً به بسط های طرفین (۱۷) تا توان چهار نیاز داریم. حال با استفاده از سری هندسی $\frac{1}{1+ax^2} = 1 - ax^2 + a^2x^4 - \dots$ و رابطه (۱۷) می توان نتیجه گرفت

$$\begin{aligned} (erf(x))^2 &= (1 - ax^2 + \dots)(bx^2 + ax^4) \\ &- \frac{1}{2}(1 + \dots)(b^2x^4 + \dots) + \dots \\ &= bx^2 + (-\frac{b^2}{2} - ab + a)x^4 + \dots \end{aligned} \quad (18)$$

از مساوی قرار دادن ضرائب x^2 و x^4 در روابط (۱۱) و (۱۸) دستگاه معادلات

$$\begin{cases} b = \frac{2}{\pi} \\ -\frac{b^2}{2} - ab + a = -\frac{a}{3\pi} \end{cases}$$

حاصل می شود. به این ترتیب $a = \frac{\Lambda(\pi-3)}{3\pi(\frac{4}{3}-\pi)}$ و $b = \frac{2}{\pi}$ به دست می آید.

۳ خطای تقریب تابع توزیع نرمال استاندارد با رابطه تقریبی سرگئی وینزکی

چون تابع خطا فرد است، بنابراین $erf(-x) = -erf(x)$ ، از این رو رابطه تقریبی سرگئی وینزکی را می توان به صورت

$$erf(x) \cong sgn(x)[1 - exp(-x^2 \frac{\frac{2}{\pi} + ax^2}{1+ax^2})]^{1/2} \quad (19)$$

۳۴۰ اثبات رابطه سرگئی وینزکی برای تقریب تابع توزیع نرمال استاندارد

نوشت. خطای تقریب تابع خطا با رابطه (۱۹) در نزدیکی صفر با مقادیر بزرگ $|x|$ ناچیز بوده و خطای کمتر از $0/00035$ است و به روش سعی و خطا می توان نشان داد که در صورت استفاده از مقدار $a = 0/147$ به جای $a = 0/14$ خطای آن به کمتر از $0/00012$ کاهش یابد. حال با فرض $x > 0$ ، رابطه (۱۹) را بر حسب x^2 حل می کنیم

$$\exp(-x^2 \frac{\frac{2}{\pi} + ax^2}{1 + ax^2}) = 1 - (\operatorname{erf}(x))^2$$

$$x^2 \frac{\frac{2}{\pi} + ax^2}{1 + ax^2} = -\ln[1 - (\operatorname{erf}(x))^2]$$

با قرار دادن $\ln[1 - (\operatorname{erf}(x))^2] = t$ داریم $\ln(1 + ax^2) = -t(1 + ax^2)$ یا $ax^4 + \frac{2}{\pi}x^2 = -t(1 + ax^2)$ داریم $ax^4 + (\frac{2}{\pi} + at)x^2 + t = 0$

$$x^2 = \frac{-(\frac{2}{\pi} + at) \pm \sqrt{(\frac{2}{\pi} + at)^2 - 4at}}{2a}$$

که جواب قابل قبول عبارت است از

$$x^2 = -(\frac{2}{\pi a} + \frac{t}{2}) + \sqrt{(\frac{2}{\pi a} + \frac{t}{2}) - \frac{t}{a}}$$

بنابراین معکوس تابع خطا عبارت است از

$$\operatorname{erf}^{-1}(x) \cong \sqrt{\sqrt{(\frac{2}{\pi a} + \frac{\ln(1-x^2)}{2})^2 - \frac{\ln(1-x^2)}{a}} - (\frac{2}{\pi a} + \frac{\ln(1-x^2)}{2})}$$

با توجه به اینکه تابع خطا فرد است، معکوس آن نیز تابعی فرد است، در نتیجه:

$$\operatorname{erf}^{-1}(x) \cong \operatorname{sgn}(x)$$

$$\times \sqrt{\sqrt{(\frac{2}{\pi a} + \frac{\ln(1-x^2)}{2})^2 - \frac{\ln(1-x^2)}{a}} - (\frac{2}{\pi a} + \frac{\ln(1-x^2)}{2})}$$

حال فرض کنیم q_α نمایش صدک α م توزیع نرمال استاندارد باشد، با استفاده از رابطه (۲) داریم

$$\alpha = \frac{1}{2} (1 + \operatorname{erf}(\frac{q_\alpha}{\sqrt{2}}))$$

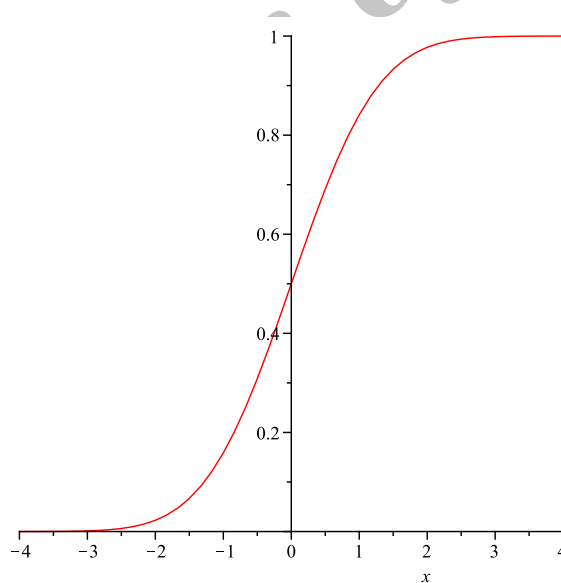
لذا صدک مذکور از رابطه

$$q_\alpha = \sqrt{2} \operatorname{erf}^{-1}(2\alpha - 1)$$

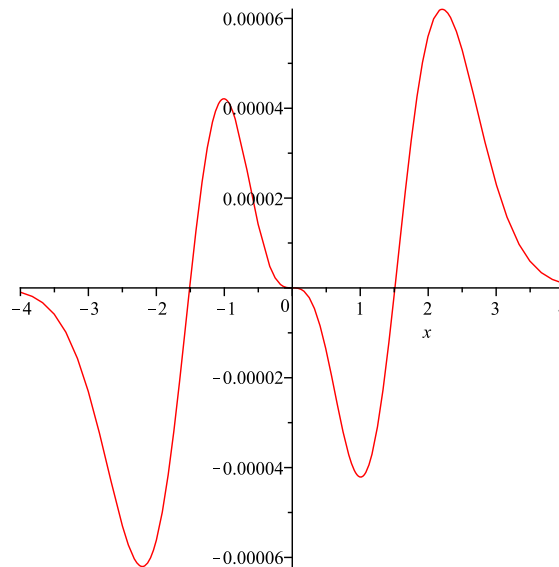
به دست می آید. با استفاده از روابط (۲) و (۱۹) می توان تقریب تابع $\Phi(x)$ را به کمک رابطه

$$\hat{\Phi}(x) = \frac{1}{\sqrt{\pi}} \left(1 + \operatorname{sgn}(x) \left[1 - \exp\left(-\frac{1}{\sqrt{\pi}} x^2 \frac{\frac{a}{\pi} + ax^2}{2 + ax^2}\right) \right]^{\frac{1}{2}} \right) \quad (20)$$

به دست آورد که حداکثر خطای آن $0/00016$ است. اگر در رابطه (۲۰) قرار داده شود $a = 0/1473$ ، خطای تقریب تابع توزیع نرمال استاندارد در نزدیکی صفر یا مقادیر بزرگ $|x|$ ناچیز بوده و حداکثر خطای آن کمتر از $0/0000584$ است. در شکل ۲ نمودار این تابع و در شکل ۳ نمودار خطای آن یعنی $\hat{\Phi}(x) - \Phi(x)$ توسط نرم افزار *Maple* رسم شده است.



شکل ۲: نمودار تقریب تابع توزیع نرمال استاندارد با رابطه (۲۰)



شکل ۳: نمودار خطا تقریب تابع توزیع نرمال استاندارد با رابطه (۲۰)

۴ بحث و نتیجه گیری

در این مقاله رابطه تقریبی سرگئی وینزکی برای تقریب تابع توزیع نرمال استاندارد برای تخمین مساحت زیر منحنی نرمال استاندارد به روشی جدید ثابت و با روش سعی و خطا و تغییر یکی از پارامترهای آن رابطه‌ای با دقت بالاتر نسبت به رابطه تقریبی سرگئی وینزکی به دست آورده شده است. همانطور که از جداول ۱ و ۲ شکل ۳ نتیجه می‌شود دقت این رابطه در بیشتر موارد از سایر روابط بیشتر است. البته ادعا نمی‌شود که جدول توزیع احتمال نرمال باید کنار گذاشته شود اما تقریب ارائه شده می‌تواند رقیبی برای جدول توزیع احتمال نرمال باشد.

تقدیر و تشکر

نویسنده لازم می‌داند که از جناب آقای دکتر عبدالحسین هورفر استاد برجسته دانشگاه تهران، داوران و مسئولین محترم مجله علوم آماری به دلیل دقت در بازبینی مقاله و پیشنهادات ارزنده‌شان که منجر به بهبود این مقاله شده است، تشکر و

مراجع

- Abernathy, R. W. (1988), Finding Normal Probabilities with Hand held Calculators, *Mathematics Teacher*, **81**, 651-652.
- Bagby, R. J. (1995), Calculating Normal Probabilities, *American Mathematical Monthly*, **102**, 46-49.
- Byrc, W. (2001), A Uniform Approximation to the Right Normal Tail Integral, *Applied Mathematics and Computation*, **127**, 365-374.
- Choudhury, A., Ray, S. and Sarkar, P. (2007), Approximating the Cumulative Distribution of the Normal Distribution, *Mathematics and Journal of Statistical Research*, **41**, 59-67.
- Choudhury, A. (2014), A Simple Approximation to the Area Under Standard Normal Curve, *Mathematics and Statistics*, **2**, 147-149.
- Feller, W. (1968), *An Introduction to Probability Theory and its Applications*, John Wiley, New York.
- Gordon, R. D. (1941), Values of Mills' Ratio of Area Bounding Ordinate and of the Normal Probability Integral for Large Values of the Argument, *Annales of Mathematical Statistics*, **12**, 364-366.
- Hammakar, H. C. (1978), Approximating the Cumulative Normal Distribution and its Inverse, *Applied Statistics*, **27**, 76-77.

۳۴۴ اثبات رابطه سرگئی وینزکی برای تقریب تابع توزیع نرمال استاندارد

Lin, J. T. (1989), Approximating the Normal Tail Probability and its Inverse for Use on a Pocket Calculator, *Applied Statistics*, **38**, 69-70.

Lin, J. T. (1990), A Simpler Logistic Approximation to the Normal Tail Probability and its Inverse, *Applied Statistics*, **39**, 255-257.

Page, E. (1977), Approximations to the Cumulative Normal Function and its Inverse for Use on a Pocket Calculator, *Applied Statistics*, **26**, 75-76.

Pollak, H. O. (1956), A Remark on "Elementary Inequalities for Mills' Ratio" by Y. Komatu, *Reports of Statistical Application Research, Union of Japanese Scientists and Engineers*, **4**, 110.

Szarek, S. J. (1999), A Nonsymmetric Correlation Inequality for Gaussian Measure, *Journal of Multivariate Analysis*, **68**, 193-211.

Tocher, K. D. (1963), *The Art of Simulation*, English University Press, London.

Winitzki, S., [http://sites.google.com/site/winitzki/sergei winitzki's files](http://sites.google.com/site/winitzki/sergei_winitzki's_files).

Zelen, M. and Severo, N. C. (1964), Probability Function (pp 925-995) in Handbook of Mathematical Functions, Abramowitz, M. and Stegun, I. A., *Applied Mathematics Series*, 55.