https://doi.org/10.31651/2524-2660-2025-1-83-91

https://orcid.org/0009-0002-1518-5151

КАСЯРУМ Сергій

кандидат педагогічних наук, доцент, доцент кафедри фізико-математичних дисциплін,
Національний університет цивільного захисту України
e-mail: kasarumsergij@gmail.com

УДК: 378.016:519.2(045)

ПИТАННЯ ЗНАХОДЖЕННЯ ОСНОВНИХ СТАТИСТИЧНИХ ФУНКЦІЙ ПРЯМИМ ОБЧИСЛЕННЯМ У ПРОЦЕСІ ВИКЛАДАННЯ ДИСЦИПЛІН «ТЕОРІЯ ЙМОВІРНОСТІ» ТА «МАТЕМАТИЧНА СТАТИСТИКА»

З метою ефективного засвоєння здобувачами знань з математичної статистики і зокрема розділу «функції випадкових величин» показано, як за допомогою відносно простих алгоритмів не складно провести обчислення розподілів χ^2, Стьюдента та Фішера-Снедекора для невеликих ступенів вільності.

Презентовано альтернативний підхід до виведення статистичних функцій без використання спеціальних функцій.

Використовуючи фундаментальні методи диференціювання та інтегрування, автор статті пропонуємо більш доступний підхід для студентів, що дозволяє послідовно будувати статистичні функції крок за кроком. Крім того, дослідження розширює перелік вправ і завдань у розділі «Функції випадкового аргументу» математичної статистики, що сприяє кращому засвоєнню теоретичного матеріалу та формує глибше розуміння виведення статистичних функцій.

Ключові слова: *викладання, математична статистика, функції випадкових величин, інтегральне числення, розподіли χ^2, Стьюдента та Фішера-Снедекора, ступені вільності.*

Постановка проблеми. При викладанні дисциплін «Теорія ймовірності» та «Математична статистика» часто виникають прогалини у чіткому розумінні здобувачами освіти навчальних матеріалів між базовими початковими означеннями та більш складними означеннями та формулами. Нам здається таким місцем у викладанні може бути введення статистичних функцій, пов'язаних з гамма-функцією Ейлера. Розподіли χ^2, Стьюдента та Фішера-Снедекора відіграють надзвичайно велику роль у математичній статистиці. Обґрунтування цих розподілів було у свій час важливою теоретичною роботою, яка виконувалася відомими вченими Ф. Р. Хельмертом (Helmert, 1876), К. Пірсоном (Pearson, 1900), В. Госсетом (Probable error, 1908), Р.А. Фішером (Fisher, 1924). Достатньо детально математичний апарат статистики описаний у ряді інших фундаментальних праць.

Зрозуміло, що, оскільки у теоретичному аспекті усі проблеми, пов'язані із цими розподілами вирішені, обґрунтовані та багатократно перевірені, то у практичному використанні статистичних функцій немає потреби згадувати теоретичні деталі та логіку виведення відповідних математичних формул.

Інша справа – ефективне засвоєння здобувачами вищої освіти знань з математичної статистики. Як правило, навчальна література обмежується лише означеннями статистичних функцій, або формулами χ^2, t та F, що виражені через гамма-функцію Ейлера (Gnedenko, 1998; Барковський та ін., 1997; Кушлик-Дивульська, Горбачук, 2023; Соловко та ін., 2021). Якщо врахувати, що теорія гамма-функції (Г-функції) Ейлера у програмі математичної підготовки здобувачів вищої освіти нематематичних спеціальностей, як правило, відсутня, то поява цих формул у такому форматі дидактично не обґрунтована і часто незрозуміла.

Між тим, за допомогою доступних алгоритмів не складно показати механізм отримання цих формул, якщо не у загальному випадку, що є достатньо складною задачею, то хоч би для декількох частинних випадків. Саме цього бракує у більшості сучасних підручниках.

Мета статті полягає у презентації матеріалу, який уможливлює здобувачам вищої освіти, використовуючи відомий базовий математичний апарат, побудувати алгоритм обчислення складних статистичних функцій у ряді простих випадків для невеликих ступенів вільності, а також сприятиме кращому розумінню ними міжпредметних зв'язків у математичному апараті статистики.

Дидактичною метою презентованої роботи є розширення списку вправ та задач для розділу «Функції випадкового аргументу».

Методологія дослідження. У роботі використаний програмний матеріал з розділів диференціального та інтегрального числення, математичної статистики, теоретичні засади знаходження щільності та розподілу функції суми або частки двох випадкових величин.

Основна ідея обчислень полягає у послідовному знаходженні результатів від малих до більших ступенів вільності за допомогою рекурентних співвідношень. Знаходження необхідних інтегралів не представляє великих ускладнень. Важливо правильно записати вихідні вирази функцій, які ми маємо обчислювати, записати необхідний інтеграл та в разі потреби звести його до табличного.

Далі розглянемо задачі знаходження розподілів χ^2, Стьюдента та F для довільних ступенів вільностей.

Розподілом χ^2 (*хі*-квадрат) з k ступенями вільності називається розподіл суми квадратів незалежних випадкових величин, що розподілені за стандартним нормальним законом, тобто:

$$\chi^2=\sum_{i=1}^{k} Z_i^2,$$

де Z_i (i=1, 2,..., k) має нормальний розподіл з $M(Z_i)=0$ та $\sigma_i=1$.

Густина ймовірності χ^2-розподілу має вигляд($\chi^2=x$):

$$f(x)=\begin{cases}\dfrac{1}{2^{\frac{k}{2}}\cdot\Gamma\left(\dfrac{k}{2}\right)}x^{\frac{k}{2}-1}\cdot e^{-\frac{x}{2}} & \text{при} \quad x\geq 0\\ 0 & \text{при} \quad x<0\end{cases},$$

де $\Gamma(y)=\int\limits_0^{+\infty} e^{-t}t^{y-1}dt$ – гама-функція Ейлера (для цілих додатних значень $y=n$ маємо $\Gamma(n)=(n-1)!$).

Розподіл χ^2 асиметричний, має додатну (правосторонню) асиметрію. При k>30 розподіл випадкової величини $Z=2\chi^2-\sqrt{2k}$ стає близьким до нормованого нормального закону, тобто $N(0;1)$.

Причина з якої була введена подібна функція пов'язана із статистичними дослідженнями, для яких дисперсія обчислюється не через математичне сподівання випадкової величини (яке, як правило, невідоме), а через його оцінку, яка на відміну від математичного сподівання є випадковою величиною. От і виходить, що при цьому сума квадратів відхилень представляє собою суму квадратів випадкових величин, як і *хі*-квадрат.

1. Виконуємо перше обчислення. Випадкова величина X розподілена нормально, $M(X)=0$, $\sigma^2=1$. У цьому випадку щільність розподілу випадкової величини описується локальною функцією Лапласа

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}.$$

Поставимо задачу знаходження розподілу немонотонної функції $Y=X^2$, який можна вважати розподілом χ^2 для ступеню вільності k=1.

Ця задача достатньо проста, часто зустрічається у підручниках, а тому наведемо результат обчислень. Щільність розподілу χ^2 для k=1:

$$g(\chi^2)=\frac{1}{\sqrt{2\pi x}}e^{-x/2};\ x\geq 0. \quad (1)$$

2. Поставимо задачу знаходження розподілу функції $\chi^2=z_1^2+z_2^2$, що є розподілом χ^2 для ступеню вільності k=2. Використаємо попередній результат обчислень та скористаємося позначеннями

$$z_1^2=p;\ z_2^2=q;\ \chi^2=y;\ y=p+q.$$

При цьому, шуканий розподіл $g(y)$ знаходимо як згортку розподілів

$$g_1(p)=\frac{1}{\sqrt{2\pi p}}e^{-p/2} \text{ та } g_2(q)=\frac{1}{\sqrt{2\pi q}}e^{-q/2}.$$

Теорія пропонує наступну процедуру:

$$g(y)=\int_0^\infty g_1(p)\cdot g_2(y-p)dp \quad (2)$$

Далі проводимо обчислення, які дають наступний результат:

$$g(y)=\int_0^y\frac{1}{p^{1/2}\cdot\sqrt{2\pi}}\cdot e^{-\frac{p}{2}}\cdot\frac{1}{(y-p)^{1/2}\cdot\sqrt{2\pi}}\cdot e^{-\frac{y-p}{2}}dp=$$

$$=\frac{1}{2\pi}\int_0^y\frac{e^{-\frac{p}{2}-\frac{y-p}{2}}}{p^{1/2}\cdot(y-p)^{1/2}}dp=\frac{e^{-\frac{y}{2}}}{2\pi}\int_0^y\frac{1}{p^{1/2}\cdot(y-p)^{1/2}}dp=$$

$$=\frac{e^{-\frac{y}{2}}}{2\pi}\left[2arctg\sqrt{\frac{p}{y-p}}\right]\Bigg|_0^y=$$

$$=\frac{e^{-\frac{y}{2}}}{2\pi}2arctg(\infty)=\frac{e^{-\frac{y}{2}}}{2\pi}2\frac{\pi}{2}=\frac{e^{-\frac{y}{2}}}{2}.$$

Отримана таким чином щільність розподілу χ^2 для k=2:

$$g(\chi^2)=\frac{e^{-\frac{\chi^2}{2}}}{2} \quad (3)$$

3. Продовжуємо обчислення, використовуючи формулу згортки. Знайдемо розподіл функції $\chi^2_{k=3}=\sum_{i=1}^{3}z_i^2$, який являє собою розподіл χ^2 для ступеню вільності k=3.

Використаємо вже обчислені попередньо розподіли та введемо позначення:

$$z_1^2+z_2^2=p;\ z_2^2=q;\ \chi^2=y;\ y=p+q.$$

Шуканий розподіл g(y) знаходимо як згортку розподілів

$$g_1(p)=\frac{1}{2}e^{-p/2} \text{ та } g_2(q)=\frac{1}{\sqrt{2\pi q}}e^{-q/2}$$

за формулою (2).

Використовуючи результат обчислення інтегралу

$$\int\frac{1}{(y-p)^{1/2}}dp=\begin{vmatrix}x=y-p\\dx=-dp\end{vmatrix}=$$

$$=-\int\frac{dx}{x^{1/2}}=-2x^{1/2}=2(y-p)^{1/2}+c,$$

отримаємо значення щільності розподілу g(y):

$$g(y)=\int_0^y\frac{1}{2}\cdot e^{-\frac{p}{2}}\cdot\frac{1}{(y-p)^{1/2}\cdot\sqrt{2\pi}}\cdot e^{-\frac{y-p}{2}}dp=$$

$$=\frac{1}{2\sqrt{2\pi}}\int_0^y\frac{e^{-\frac{p}{2}-\frac{y-p}{2}}}{(y-p)^{1/2}}dp=$$

$$=\frac{e^{-\frac{y}{2}}}{2\pi}\int_0^y\frac{1}{(y-p)^{1/2}}dp=\frac{e^{-\frac{y}{2}}}{2\sqrt{2\pi}}\left[-2(y-p)^{1/2}\right]\Big|_0^y=$$

$$=\frac{e^{-\frac{y}{2}}}{2\sqrt{2\pi}}2\sqrt{y}=\frac{e^{-\frac{y}{2}}}{\sqrt{2\pi}}\sqrt{y}.$$

Щільність розподілу χ^2 для k=3:

$$g(\chi^2)=\frac{e^{-\frac{\chi^2}{2}}}{\sqrt{2\pi}}\sqrt{\chi^2} \quad (4)$$

4. Очевидно, у такий спосіб, комбінуючи у формулі згортки вже знайдені розподіли, неважко знати розподіли для більших значень ступеню вільності k.

Наприклад, знайдемо розподіл функції

$$\chi^2_{k=4}=\sum_{i=1}^{4}z_i^2.$$

Введемо позначення: $\chi^2=\sum_{i=1}^{4}z_i^2;\ z_1^2+z_2^2=p;\ z_{,3}^2+z_4^2=q;\ \chi^2=y$ та отримаємо згортку g(y) розподілів: $g(p)=\frac{1}{2}\cdot e^{-\frac{p}{2}};\ g(q)=\frac{1}{2}\cdot e^{-\frac{q}{2}}$:

y=p+q;

$$g(y)=\int_0^\infty f_1(p)\cdot f_2(q)dp=\int_0^\infty f_1(p)\cdot f_2(s-p)dp.$$

$$g(y)=\int_0^y\frac{1}{2}\cdot e^{-\frac{p}{2}}\cdot\frac{1}{2}\cdot e^{-\frac{y-p}{2}}dp=\frac{1}{4}\int_0^y e^{-\frac{p}{2}-\frac{y-p}{2}}dp=$$

$$=\frac{e^{-\frac{y}{2}}}{4}\int_0^y dp=\frac{e^{-\frac{y}{2}}}{4}\left[p\right]_0^y=\frac{e^{-\frac{y}{2}}}{4}y.$$

Щільність розподілу χ^2 для k=4:

$$g(\chi^2)=\frac{e^{-\frac{\chi^2}{2}}}{4}\chi^2 \quad (5)$$

Зрозуміло, що така ідея обчислень дозволяє на підставі вже зроблених розрахунків (1); (3)–(5) відносно просто отримати результати для більших значень ступенів вільності.

Отже, для k=n потрібно обчислити згортку з відповідними попередньо обчисленими розподілами із ступенями вільності k<n.

5. Виконаємо перевірку отриманих значень розподілу χ^2 із результатами підстановки у загальний вираз цього розподілу через гамма-функцію:

$$f(x)=\begin{cases}\dfrac{1}{2^{\frac{k}{2}}\cdot\Gamma\left(\dfrac{k}{2}\right)}x^{\frac{k}{2}-1}\cdot e^{-\frac{x}{2}} & \text{при} \quad x\geq 0\\ 0 & \text{при} \quad x<0,\end{cases}$$

де $\Gamma(y)=\int\limits_0^{+\infty}e^{-t}t^{y-1}dt$ – Г-функція Ейлера.

Відомо, що для цілих додатних значень y=n маємо: $\Gamma(n)=(n-1)!$, а для напівцілих

$$\Gamma(1/2+n)=\frac{(2n)!}{4^n n!}\sqrt{\pi}\ .$$

Отже, для y=1 отримаємо Г(1)=1, для y=2 також Г(2)=1, для y=1/2 отримаємо $\Gamma(1/2)=\sqrt{\pi}$, а для y=3/2 – $\Gamma(3/2)=\frac{1}{2}\sqrt{\pi}$.

Зробимо підстановку

$$f_{k=1}(x)=\frac{1}{2^{\frac{1}{2}}\cdot\Gamma\left(\frac{1}{2}\right)}x^{\frac{1}{2}-1}\cdot e^{-\frac{x}{2}}=\frac{e^{-\frac{x}{2}}}{\sqrt{2\pi x}}\ ,$$

$$f_{k=2}(x)=\frac{1}{2^{\frac{2}{2}}\cdot\Gamma\left(\frac{2}{2}\right)}x^{\frac{2}{2}-1}\cdot e^{-\frac{x}{2}}=\frac{e^{-\frac{x}{2}}}{2}\ ,$$

$$f_{k=3}(x)=\frac{1}{2^{\frac{3}{2}}\cdot\Gamma\left(\frac{3}{2}\right)}x^{\frac{3}{2}-1}\cdot e^{-\frac{x}{2}}=\frac{e^{-\frac{x}{2}}}{\sqrt{2\pi}}\sqrt{x}\ ,$$

$$f_{k=4}(x)=\frac{1}{2^{\frac{4}{2}}\cdot\Gamma\left(\frac{4}{2}\right)}x^{\frac{4}{2}-1}\cdot e^{-\frac{x}{2}}=\frac{e^{-\frac{x}{2}}}{4}x\ .$$

що збігається з формулами (1), (3)–(5).

Отже, маємо повну відповідність результатів підстановки у загальну формулу χ^2 з результатами вище проведених обчислень.

Розподілом Стьюдента (або t-розподілом) називається розподіл випадкової величини:

$$t=\frac{Z}{\sqrt{\frac{1}{k}\chi^2}}$$

де Z – випадкова величина, розподілена за нормованим нормальним законом, тобто $M(Z)=0$ та σ=1;

χ^2 – незалежна від Z випадкова величина, що має χ^2 – розподіл з k степенями вільності

Густина ймовірності розподілу Стьюдента має вигляд:

$$f(t)=\frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{\pi k}}\left(1+\frac{t^2}{n}\right)^{-\frac{k+1}{2}}\ ,$$

де Г(y) – гама-функція в точці у.

При $k\to\infty$ t-розподіл наближається до нормального. Практично вже для k>30 t-розподіл мало відрізняється від нормального.

Математичне сподівання випадкової величини, що має t-розподіл, в силу симетрії її кривої розподілу дорівнює нулю, а її дисперсія (це можна довести) дорівнює $\frac{k}{k-2}$, тобто

$$M(t)=0;\ D(t)=\frac{k}{k-2}.$$

Розподіл Стьюдента в статистиці використовують замість нормального розподілу, аргументом якого є величина $z=\frac{x-a}{\sqrt{D(x)}}$ в тих випадках, коли дисперсія розподілу невідома. При цьому в знаменнику аргументу z з'являється обчислене значення дисперсії, яка є випадковою величиною, розподіленою за розподілом χ^2. Часто розподіл Стьюдента використовують для висновків про середні значення випадкових величин.

Для знаходження розподілу Стьюдента використаємо дещо складнішу процедуру обчислень порівняно із попередньою задачею. Для того, щоб скористатися способом знаходження розподілу результату ділення однієї випадкової величини на іншу $F\left(\frac{Z}{y}\right)=\iint\limits_{z,y}g_1(Z)g_3(y)dzdy$, потрібно спочатку знайти щільність розподілу величини $y=f(\chi^2)=\sqrt{\frac{1}{k}\chi^2}$ (щільність розподілу величини y – знаменника t).

Оскільки ми не використовуємо загальну формулу розподілу Стьюдента, то змушені знаходити розподіл знаменника для кожного значення ступеню вільності k.

1. Розв'яжемо задачу для k=1. У цьому випадку $t=\frac{Z}{\sqrt{\chi^2}}$, а розподіли Z та χ^2 вважаємо такими, що

$$g_1(Z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}};\ \ g_2(\chi^2)=\frac{1}{\sqrt{2\pi\chi^2}}e^{-\frac{\chi^2}{2}}$$

(для k=1).

Спочатку знаходимо розподіл $y=\sqrt{\chi^2}$ за процедурою знаходження розподілу функції випадкової величини з відомим розподілом $y=f(\chi^2)=\sqrt{\chi^2}$, звідки $\chi^2=f^{-1}(y)=y^2$; $(f^{-1}(y))'=2y$;

$$g_3(y)=\frac{1}{\sqrt{2\pi y^2}}e^{-\frac{y_2}{2}}\cdot 2y=\frac{2}{\sqrt{2\pi}}e^{-\frac{y_2}{2}}.$$

Отже,

$$g_3(y)=\frac{2}{\sqrt{2\pi}}e^{-\frac{y^2}{2}} \text{ або } g_3(\sqrt{\chi^2})=\frac{2}{\sqrt{2\pi}}e^{-\frac{\chi^2}{2}}.$$

Далі знаходимо розподіл частки подібно знаходженню розподілу суми:

$$F\left(\frac{Z}{y}\right)=\iint\limits_{z,y} g(Z)g(y)dzdy=$$

$$=\iint\limits_{z,y}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dzdy.$$

Якщо $\frac{Z}{y}=x$, то $Z=xy$; $dz=ydx$ і тоді

$$F\left(\frac{Z}{y}\right)=\int_{-\infty}^{t}dx\int_0^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2y^2}{2}}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}ydy=$$

$$=\frac{1}{\pi}\int_{-\infty}^{t}dx\int_0^{\infty}e^{-\frac{(x^2+1)y^2}{2}}ydy=\left|\begin{array}{l}-\frac{(x^2+1)y^2}{2}=p\\ dp=(x^2+1)ydy\end{array}\right|=$$

$$=\int_{-\infty}^{t}dx\left[\frac{1}{\pi(x^2+1)}\right]\int_0^{\infty}e^{-p}dp=\int_{-\infty}^{t}dx\left[\frac{1}{\pi(x^2+1)}\right].$$

Таким чином, функція розподілу Стьюдента для k=1 інтеграл $F(x)$, а щільність розподілу – підінтегральний вираз:

$$g(t)=\frac{1}{\pi(t^2+1)};\tag{6}$$

від

$$F(t)=\frac{1}{\pi}\int_{-\infty}^{t}\frac{1}{(x^2+1)}dx=\frac{1}{\pi}\cdot arctgx\Big|_{-\infty}^{t}=\frac{1}{\pi}\left(arctgt+\frac{\pi}{2}\right).$$

У подальших обчисленнях ми не будемо обчислювати $F(x)$, а лише щільність $g(x)$. У статистичних задачах для обчислень використовують саме щільність розподілу, а не функцію розподілу.

2. Задача знаходження розподілу Стьюдента для k=2 розв'язується аналогічно попередній.

Спочатку знаходимо щільність розподілу $g_3\left(\sqrt{\frac{1}{2}\chi^2}\right)$:

$$g_1(Z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}};\quad g_2(\chi^2)=\frac{1}{2}e^{-\frac{\chi^2}{2}};$$

$$t=\frac{Z}{\sqrt{\frac{1}{2}\chi^2}};\quad y=f(\chi^2)=\sqrt{\frac{1}{2}\chi^2},$$

звідки $\chi^2=f^{-1}(y)=2y^2$; $(f^{-1}(y))'=4y$;

$$g_3(y)=\frac{1}{2}e^{-y^2}\cdot 4y=2y\cdot e^{-y^2}$$

$$\text{або } g_3\left(\sqrt{\frac{1}{2}\chi^2}\right)=\sqrt{2\chi^2}\cdot e^{-\frac{1}{2}\chi^2}.$$

Далі шукаємо щільність та функцію розподілу $t=\frac{Z}{\sqrt{\frac{1}{2}\chi^2}}=\frac{Z}{y}$; використавши позначення $\frac{Z}{y}=x$, звідки $Z=xy$; $dz=ydx$;

$$F\left(\frac{Z}{y}\right)=\iint\limits_{z,y} g(Z)g(y)dzdy=$$

$$=\iint\limits_{z,y}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\cdot 2y\cdot e^{-y^2}dzdy=$$

$$=\int_{-\infty}^{t}dx\int_0^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2y^2}{2}}\cdot 2ye^{-y^2}ydy=$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t}dx\int_0^{\infty}e^{-\left(\frac{x^2}{2}+1\right)y^2}2y^2dy=$$

$$=\left|\begin{array}{l}\left(\frac{x^2}{2}+1\right)y^2=p^2;\\ \left(\frac{x^2}{2}+1\right)^{\frac{1}{2}}y=p;\\ dp=\left(\frac{x^2}{2}+1\right)^{\frac{1}{2}}ydy\end{array}\right|=$$

$$=\int_{-\infty}^{t}dx\left[\frac{1}{\sqrt{2\pi}\left(\frac{x^2}{2}+2\right)^{\frac{3}{2}}}\int_0^{\infty}e^{-p^2}2p^2dp\right]\Rightarrow$$

$$\Rightarrow\int_{-\infty}^{t}dx\left[\frac{1}{\sqrt{2\pi}\left(\frac{x^2}{2}+2\right)^{\frac{3}{2}}}\cdot\frac{\sqrt{\pi}}{2}\right]=$$

$$=\int_{-\infty}^{t} dx\left[\frac{1}{2\sqrt{2}\left(\frac{x^2}{2}+1\right)^{\frac{3}{2}}}\right].$$

Вираз результату обчислень у квадратних дужках є щільність розподілу Стьюдента для k=2:

$$g(t)=\frac{1}{2\sqrt{2}\left(\frac{t^2}{2}+1\right)^{\frac{3}{2}}}, \qquad (7)$$

3. Щільність розподілу Стьюдента для k=3. Задача знаходження розподілу розв'язується аналогічно попереднім:

$$g(Z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}};\ g(\chi^2)=\frac{\sqrt{\chi^2}}{\sqrt{2\pi}}e^{-\frac{\chi^2}{2}}.$$

Спочатку знаходимо щільність розподілу

$$g_3\left(\sqrt{\frac{1}{3}\chi^2}\right),\ y=f(\chi^2)=\sqrt{\frac{1}{3}\chi^2},$$

звідки $\chi^2=f^{-1}(y)=3y^2$; $(f^{-1}(y))'=6y$;

$$g(y)=\frac{\sqrt{3y^2}}{\sqrt{2\pi}}e^{-\frac{3y^2}{2}}\cdot 6y=\frac{1}{\sqrt{2\pi}}\cdot 6\sqrt{3}y^2\cdot e^{-\frac{3y^2}{2}}.$$

Далі шукаємо щільність та функцію розподілу $t=\frac{Z}{\sqrt{\frac{1}{2}\chi^2}}=\frac{Z}{y}$ при $\frac{Z}{y}=x$, звідки $Z=xy$; $dz=ydx$;

$$F\left(\frac{Z}{y}\right)=\iint_{z,y} g(Z)g(y)dzdy=$$

$$=\iint_{z,y}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}\cdot\frac{1}{\sqrt{2\pi}}\cdot 6\sqrt{3}y^2\cdot e^{-\frac{3y^2}{2}}dzdy=$$

$$=\int_{-\infty}^{t}dx\int_0^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2y^2}{2}}\cdot\frac{1}{\sqrt{2\pi}}\cdot 6\sqrt{3}y^2\cdot e^{-\frac{3y^2}{2}}ydy=$$

$$=\frac{2\sqrt{3}}{2\pi}\int_{-\infty}^{t}dx\int_0^{\infty}e^{-\left(\frac{x^2}{3}+1\right)\frac{3y^2}{2}}y^2 3ydy=$$

$$=\left|\begin{matrix}\left(\frac{x^2}{3}+1\right)\frac{3y^2}{2}=p;\\ \left(\frac{x^2}{3}+1\right)3ydy=dp\end{matrix}\right|=$$

$$=\frac{\sqrt{3}}{\pi}\cdot\frac{2}{3}\int_{-\infty}^{t}dx\left[\frac{1}{\left(\frac{x^2}{3}+1\right)^2}\int_0^{\infty}e^{-p}pdp\right]=$$

$$=\left|\int_0^{\infty}e^{-p}dp=1\right|=\int_{-\infty}^{t}dx\left[\frac{2}{\pi\sqrt{3}}\cdot\frac{1}{\left(\frac{x^2}{3}+1\right)^2}\right].$$

Вираз у квадратних дужках результату обчислень є щільність розподілу Стьюдента для k=3:

$$g(t)=\frac{2}{\pi\sqrt{3}}\frac{1}{\left(\frac{t^2}{3}+1\right)^2}, \qquad (8)$$

Продовжуючи обчислення подібним чином можемо знаходити розподіл Стьюдента для інших ступенів вільності, попередньо знайшовши розподіл відповідної χ^2.

4. Порівняємо отримані результати із загальною формулою для розподілу Стьюдента:

$$f(t)=\frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{\pi k}}\left(1+\frac{t^2}{k}\right)^{-\frac{k+1}{2}}$$

Зробимо підстановку наведених вище відповідних значень Г-функції:

$$f_{k=1}(t)=\frac{\Gamma\left(\frac{1+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\sqrt{\pi\cdot 1}}\left(1+\frac{t^2}{1}\right)^{-\frac{1+1}{2}}=\frac{1}{\pi}\left(1+t^2\right)^{-1};$$

$$f_{k=2}(t)=\frac{\Gamma\left(\frac{2+1}{2}\right)}{\Gamma\left(\frac{2}{2}\right)\sqrt{\pi\cdot 2}}\left(1+\frac{t^2}{2}\right)^{-\frac{2+1}{2}}=\frac{\sqrt{\pi}}{2\sqrt{\pi\cdot 2}}\left(1+\frac{t^2}{2}\right)^{-\frac{3}{2}}=$$

$$=\frac{1}{2\sqrt{2}}\left(1+\frac{t^2}{2}\right)^{-\frac{3}{2}};$$

$$f_{k=3}(t)=\frac{\Gamma\left(\frac{3+1}{2}\right)}{\Gamma\left(\frac{3}{2}\right)\sqrt{\pi\cdot 3}}\left(1+\frac{t^2}{3}\right)^{-\frac{3+1}{2}}=\frac{1}{\frac{1}{2}\sqrt{\pi}\sqrt{\pi\cdot 3}}\left(1+\frac{t^2}{3}\right)^{-\frac{4}{2}}=$$

$$=\frac{2}{\pi\sqrt{3}}\left(1+\frac{t^2}{3}\right)^{-2}.$$

Маємо повну відповідність результатів підстановки у загальну формулу з результатами вище проведених обчислень (6)–(8).

Розподілом Фішера-Снедекора (або F-розподілом) називається розподіл випадкової величини

$$F=\frac{\frac{1}{k_1}\chi^2(k_1)}{\frac{1}{k_2}\chi^2(k_2)}$$

де $\chi^2(k_1)$ і $\chi^2(k_2)$ – випадкові величини, що мають χ^2-розподіл відповідно з k_1 і k_2 степенями вільності.

Густина ймовірності F-розподілу має вигляд:

$$f(x)=\frac{\Gamma\left(\frac{k_1+k_2}{2}\right)\cdot k_1^{\frac{k_1}{2}}\cdot k_2^{\frac{k_2}{2}}}{\Gamma\left(\frac{k_1}{2}\right)\cdot\Gamma\left(\frac{k_2}{2}\right)}\cdot x^{\frac{k_1}{2}-1}\cdot\left(k_1x+k_2\right)^{-\frac{k_1+k_2}{2}},$$

де $\Gamma(y)$ – гама-функція Ейлера в точці y.

З самого виду розподілу Фішера-Снедекора видно, що це відношення двох розподілів χ^2 ($\chi^2(k_1)$ та $\chi^2(k_2)$), а отже може служити порівнянню між собою дисперсій цих розподілів.

1. Поставимо задачу знайти розподіл функції Фішера-Снедекора – F для частинного випадку, де $\chi^2(k_1=3)$, а $\chi^2(k_2=4)$. При цьому використаємо попередні обчислення щільності розподілів відповідних величин. Скористаємося формулами попередніх обчислень (4) та (5):

$$\chi^2(k_1=3)=p-g_1(p)=\frac{e^{-\frac{p}{2}}}{\sqrt{2\pi}}\sqrt{p};$$

$$\chi^2(k_2=4)=q-g_2(q)=\frac{e^{-\frac{q}{2}}}{4}q.$$

Шукаємо функцію розподілу

$$G\left(\frac{k_2p}{k_1q}\right)=\iint_{p,q}g_1(p)g_2(q)dpdq=\iint_{p,q}\frac{e^{-\frac{p}{2}}}{\sqrt{2\pi}}\sqrt{p}\cdot\frac{e^{-\frac{q}{2}}}{4}q\cdot dpdq;$$

$\frac{k_2p}{k_1q}=x$, звідки $p=\frac{k_1}{k_2}xq;\ dp=\frac{k_1}{k_2}qdx$ і

$$G\frac{4}{3}\frac{p}{q}=\int_{p,q}g_1(p)g_2(q)dpdq=$$

$$=\iint_{x,q}\frac{e^{-\frac{3}{4}\frac{xq}{2}}}{\sqrt{2\pi}}\sqrt{\frac{3}{4}xq}\cdot\frac{e^{-\frac{q}{2}}}{4}q\cdot\frac{3}{4}qdxdq;$$

$$G\left(\frac{4}{3}\frac{p}{q}\right)=\int_0^F dx\int_0^\infty\frac{1}{4\sqrt{2\pi}}e^{-\frac{3}{4}\frac{xq}{2}}\sqrt{\frac{3}{4}x}\cdot\frac{k_1}{k_2}e^{-\frac{q}{2}}q^{\frac{5}{2}}dq=$$

$$=\frac{1}{4\sqrt{2\pi}}\left(\frac{3}{4}\right)^{\frac{3}{2}}\int_0^F\sqrt{x}dx\int_0^\infty e^{-\frac{\left(\frac{3}{4}x+1\right)q}{2}}q^{\frac{5}{2}}dq=$$

$$=\left|\begin{array}{l}\dfrac{\left(\frac{3}{4}x+1\right)q}{2}=y;\\ dy=\dfrac{\left(\frac{3}{4}x+1\right)}{2}dq\\ q=\dfrac{2y}{\left(\frac{3}{4}x+1\right)}\end{array}\right|=$$

$$=\frac{1}{4\sqrt{2\pi}}\left(\frac{3}{4}\right)^{\frac{3}{2}}\int_0^F\frac{2^{\frac{2}{7}}\sqrt{x}dx}{\left(\frac{3}{4}x+1\right)^{\frac{7}{2}}}\int_0^\infty e^{-y}y^{\frac{5}{2}}dy=$$

$$=\left|\int_0^\infty e^{-y}y^{\frac{5}{2}}dy=\frac{15}{8}\sqrt{\pi}\right|=$$

$$=\int_0^F\frac{1}{4\sqrt{2\pi}}\left(\frac{3}{4}\right)^{\frac{3}{2}}\frac{2^{\frac{7}{2}}\sqrt{x}}{\left(\frac{3}{4}x+1\right)^{\frac{7}{2}}}\cdot\frac{15}{8}\sqrt{\pi}dx=$$

$$=\int_0^F\frac{180\sqrt{3x}}{(3x+4)^{\frac{7}{2}}}dx.$$

Отже, маємо щільність F- розподілу, для k_1=3, а k_2=4:

$$g(x)=\frac{180\sqrt{3x}}{(3x+4)^{\frac{7}{2}}}. \tag{9}$$

2. Знайдемо розподіл функції Фішера-Снедекора – F для частинного випадку, де де $\chi^2(k_1=4)$, а знаменник також – $\chi^2(k_2=4)$. Попередні обчислення щільності розподілів відповідних величин та позначення у цьому випадку:

$$\chi^2(k_1=4)=p-g_1(p)=\frac{e^{-\frac{p}{2}}}{4}p;$$

$$\chi^2(k_2=4)=q-g_2(q)=\frac{e^{-\frac{q}{2}}}{4}q.$$

Шукаємо функцію розподілу, врахувавши, що:

$$\frac{k_2p}{k_1q}=x,\ p=\frac{4}{4}xq=xq;\ dp=qdx;$$

$$G\left(\frac{k_2p}{k_1q}\right)=\iint_{p,q}g_1(p)g_2(q)dpdq=$$

$$=\iint_{p,q}\frac{e^{-\frac{p}{2}}}{4}p\cdot\frac{e^{-\frac{q}{2}}}{4}q\cdot dpdq=$$

$$=\iint_{x,q}\frac{e^{-\frac{xq}{2}}}{4}xq\cdot\frac{e^{-\frac{q}{2}}}{4}q\cdot qdxdq=\int_0^F\frac{xdx}{16}\int_0^\infty e^{-\frac{(x+1)q}{2}}q^3dq=$$

$$=\left|\begin{array}{l}\dfrac{(x+1)q}{2}=y;\\ dy=\dfrac{(x+1)}{2}dq\\ q=\dfrac{2y}{x+1}\end{array}\right|=\int_0^F\frac{2^4xdx}{16(x+1)^4}\int_0^\infty e^{-y}y^3dy=$$

$$=\left|\int_0^\infty e^{-y}y^3dy=6\right|=\int_0^F\frac{6xdx}{(x+1)^4}.$$

Отже, маємо щільність F-розподілу, для k_1=4, та k_2=4:

$$g(x)=\frac{6x}{(x+1)^4}. \qquad (10)$$

Перевіримо відповідність отриманих результатів (9)–(10) загальній формулі розподілу Фішера-Снедекора:

$$f(x)=\frac{\Gamma\left(\frac{k_1+k_2}{2}\right)\cdot k_1^{\frac{k_1}{2}}\cdot k_2^{\frac{k_2}{2}}}{\Gamma\left(\frac{k_1}{2}\right)\cdot\Gamma\left(\frac{k_2}{2}\right)}\cdot x^{\frac{k_1}{2}-1}\cdot(k_1x+k_2)^{-\frac{k_1+k_2}{2}}$$

Для цього підставимо у цей вираз значення ступенів вільності:

k_1=3, k_2=4:

$$f(x)=\frac{\Gamma\left(\frac{3+4}{2}\right)\cdot 3^{\frac{3}{2}}\cdot 4^{\frac{4}{2}}}{\Gamma\left(\frac{3}{2}\right)\cdot\Gamma\left(\frac{4}{2}\right)}\cdot x^{\frac{3}{2}-1}\cdot(3x+4)^{-\frac{3+4}{2}}=$$

$$=\frac{\Gamma\left(3+\frac{1}{2}\right)\cdot 3^{\frac{3}{2}}\cdot 4^{2}}{\Gamma\left(\frac{3}{2}\right)\cdot\Gamma(2)}\cdot x^{\frac{1}{2}}\cdot(3x+4)^{-\frac{7}{2}}=$$

$$=\frac{15\sqrt{\pi}\cdot 3\sqrt{3}\cdot 16\cdot 2}{8\sqrt{\pi}}\cdot x^{\frac{1}{2}}\cdot(3x+4)^{-\frac{7}{2}}=$$

$$=180\sqrt{3}\cdot x^{\frac{1}{2}}\cdot(3x+4)^{-\frac{7}{2}}.$$

Тут враховано, що $\Gamma=\left(3+\frac{1}{2}\right)=\frac{15\sqrt{\pi}}{8}$; $\Gamma=\left(\frac{3}{2}\right)=\frac{\sqrt{\pi}}{2}$; Г(2)=1.

Отже, $g(x)=\frac{180\sqrt{3x}}{(3x+4)^{\frac{7}{2}}}$, а для k_1=4, k_2=4 маємо:

$$f(x)=\frac{\Gamma\left(\frac{4+4}{2}\right)\cdot 4^{\frac{4}{2}}\cdot 4^{\frac{4}{2}}}{\Gamma\left(\frac{4}{2}\right)\cdot\Gamma\left(\frac{4}{2}\right)}\cdot x^{\frac{4}{2}-1}\cdot(4x+4)^{-\frac{4+4}{2}}=$$

$$=\frac{\Gamma(4)\cdot 4^2\cdot 4^2}{\Gamma(2)\cdot\Gamma(2)}\cdot x\cdot\frac{(x+1)^{-4}}{4^4}=6x(x+1)^{-4}.$$

Отже, $g(x)=6x(x+1)^{-4}$ (тут враховано, що Г(4)=6; Г(2)=1).

Для обох розглянутих прикладів маємо повну відповідність результатів обчислень.

Отже, використавши базові поняття функції випадкового аргументу та згортка двох розподілів таких функцій, можна послідовно обчислювати значення основних статистичних функцій для будь яких значень ступенів вільності.

Висновки. У роботі були досліджені алгоритми розв'язання наступних задач з курсу «математична статистика»: задачі знаходження щільності та функції розподілу функції випадкової величини, задачі знаходження щільності та функції розподілу функції суми або частки двох випадкових величин. Ці задачі основані на відомих для здобувачів вищої освіти правилах диференціювання та інтегрування неперервних функцій. Було встановлено, що відомі алгоритми розв'язання перелічених задач потребують спрощення і вдосконалення, оскільки, використовуючи відомий здобувачем базовий математичний апарат, нескладно показати алгоритми отримання виразів цих статистичних функцій у ряді простих випадків для невеликих ступенів вільності та порівняти результати цих обчислень із результатами, отриманими від загальних виразів на прикладі задачі знаходження розподілів χ^2 та Стьюдента для довільних ступенів вільностей. Основна ідея спрощення цих алгоритмів полягає у послідовному знаходженні результатів від довільних ступенів вільності за допомогою отримання рекурентних залежностей. Тому були отримані прості і зручні алгоритми розрахунку базових статистичних функцій, крім того, використання запропонованих алгоритмів розширює список вправ та задач для розділу математичної статистики «Функції випадкового аргументу», які можна застосувати у освітньому процесі.

Було доведено, що використавши базові поняття функції випадкового аргументу та згортки двох розподілів таких функцій, можна послідовно обчислювати значення основних статистичних функцій для довільних значень ступенів вільності. Це заповнює прогалину у освітньому процесі між базовими початковими означеннями та складними формулами статистичних функцій пов'яза-них с гамма-функцією Ейлера.

Матеріали роботи розширюють список вправ та задач для розділу математичної статистики «Функції випадкового аргументу», які можна застосувати у освітньому процесі.

Список бібліографічних посилань

Барковський та ін., 1997 – Барковський, В.В., Барковська, Н.В., Лопатін, О.К. (1997). Математика для економістів: Теорія ймовірностей та математична статистика. Київ: Національна академія управління. 424 с.

Кушлик-Дивульська, Горбачук, 2023 – Кушлик-Дивульська, О.І., Горбачук, В.М. (2023). Теорія ймовірностей та математична статистика. Київ: КПІ ім. Ігоря Сікорського. 351 с. URL: https://ela.kpi.ua/server/api/core/bitstreams/4b0ef359-532f-44b9-9436-4203204734db/content

Соловко та ін., 2015 – Соловко, Я.Т., Остафійчук, П.Г., Гарпуль, О.З., Войтик, С.А. (2015). Теорія ймовірностей та математична статистика (конспект лекцій + тести): навчальний посібник. 2-ге вид., допов. Івано-Франківськ: Симфонія Форте. 153 с.

Fisher, 1924 – Fisher, R.A. (1924). The distribution of the partial correlation coefficient. *Metron,* 3: 329–332.

Gnedenko, 1998 – Gnedenko, B.V. (1998). Theory of probability (6th ed.). London: Routledge. 520 p. Doi: https://doi.org/10.1201/9780203718964

Helmert, 1876 – Helmert, F.R. (1876). Über die Wahrscheinlichkeit von Potenzsummen der Beobachtungsfehler etc. *Zeitschrift für Mathematik und Physik,* 21: 102–219.

Pearson, 1900 – Pearson, R. (1900). On the criterion that a given system of deviations from the probable in the case a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine,* 50: 157–175.

Probable error, 1908 – The probable error of a means (1908). *Biometrika,* 6(1), 1–25.

References

Barkovskyi, V.V., Barkovska, N.V., Lopatin, O.K. (1997). Mathematics for Economists: Probability Theory and Mathematical Statistics. Kyiv: National Academy of Management. 424 p.

Kushlyk-Dyvulska, O.I., Horbachuk, V.M. (2023). Probability Theory and Mathematical Statistics. Kyiv: Igor Sikorsky Kyiv Polytechnic Institute. 351 p. Retrieved from https://ela.kpi.ua/server/api/core/bitstreams/4b0ef359-532f-44b9-9436-4203204734db/content

Solovko, Ya.T., Ostafiichuk, P.H., Harpul, O.Z., & Voityk, S.A. (2021). Probability theory and mathematical statistics (lecture notes + tests): textbook. 2nd ed., supplemented. Ivano-Frankivsk: Symphonia Forte. 153 p.

Fisher, 1924 – Fisher, R.A. (1924). The distribution of the partial correlation coefficient. *Metron,* 3: 329–332.

Gnedenko, B.V. (1998). Theory of probability (6th ed.). London: Routledge. 520 p. Doi: https://doi.org/10.1201/9780203718964

Helmert, F.R. (1876). Über die Wahrscheinlichkeit von Potenzsummen der Beobachtungsfehler etc. *Zeitschrift für Mathematik und Physik,* 21: 102–219 [in Germ.].

Pearson, R. (1900). On the criterion that a given system of deviations from the probable in the case a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine,* 50: 157–175.

Probable error, 1908 – The probable error of a means (1908). *Biometrika,* 6(1), 1–25.

KASIARUM Serhii

Ph.D in Pedagogy, Associate Professor,
Associate Professor of the Department of Physical and Mathematical Disciplines,
National University of Civil Protection of Ukraine

THE ISSUE OF FINDING THE MAIN STATISTICAL FUNCTIONS BY DIRECT CALCULATION IN THE TEACHING OF THE DISCIPLINES "PROBABILITY THEORY" AND "MATHEMATICAL STATISTICS"

Summary. *Introduction. For students' qualitative mathematical training, it is crucial to carefully and consistently present theoretical material. This ensures that students develop logical and continuous understanding of theoretical problems within a discipline, as well as the principles of their solutions from the beginning to the end of the course.*

In the teaching of «The Theory of Probability and Mathematical Statistics», we have identified a gap in this logical chain – specifically in the determination of the basic statistical functions of Pearson, Student, and Fisher through the Euler gamma function. Typically, neither the gamma function itself, nor its properties, nor its applications are included in students' mathematical training. This omission creates difficulties in understanding statistical functions at a deeper level.

Purpose. The purpose of this study is to present an alternative approach to deriving statistical functions without relying on special functions such as the gamma function. By using fundamental differentiation and integration methods, we aim to provide a more accessible approach for students. This approach facilitates understanding statistical function derivations by constructing them sequentially, step by step.

Methods. To construct the formulas for the statistical functions, we employ a stepwise convolution method. For finding the formula of the Pearson distribution, we use the formula for the convolution of two densities χ^2*. First, the convolution was applied to the degrees of freedom* $k_1=1$ *and* $k_2=1$*, which allowed to be determined* χ^2 *for* $k = 2$*. Then for the pair* $k_1=1$ *and* $k_2=2$ *we obtain* χ^2 *with degrees* $k = 3$ *and* $k = 4$*, and then, consistently, we match* χ^2 *with* k*, equal to 5, 6, 7, 8, and so on.*

To find the density of the Student's distribution, we use a somewhat more complicated computation procedure, since it is a function of the distribution of the particle density, and the density of the divisor must be calculated each time separately. Similarly, you can find the Fisher function.

To verify the accuracy of our results, we compare our derived formulas with those obtained using general expressions based on the gamma function.

Results. The comparison between our derived formulas and the standard gamma function-based formulas shows complete correspondence. This confirms the validity of our proposed approach. Our method provides an intuitive, step-by-step construction of statistical function formulas without requiring prior knowledge of the gamma function.

Originality. This study presents a new pedagogical approach to teaching statistical functions in probability theory and mathematical statistics. By avoiding reliance on special functions, we make the learning process more accessible to students who may lack prior exposure to the Euler gamma function. Our stepwise approach enhances comprehension and supports deeper engagement with the mathematical structures underlying statistical distributions.

Conclusion. The proposed approach does not seek to replace the existing optimal computation methods using the gamma function. Rather, it serves as a pedagogical tool to bridge the gap between fundamental definitions of statistical functions and their more complex formulations.

Furthermore, the study expands the range of exercises and problems available in the section «Functions of a Random Argument» in mathematical statistics. This contributes to better assimilation of theoretical concepts and reinforces students' understanding of statistical function derivations.

Keywords: *teaching; mathematical statistics; functions of random variables; integral calculus;* χ^2*, Student's and Fisher-Snedecor distributions; degrees of freedom.*

Одержано редакцією 01.03.2025
Прийнято до публікації 13.03.2025