МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ РАДІОЕЛЕКТРОНІКИ

Факультет ІТМ

Кафедра прикладної математики

КУРСОВА РОБОТА

ПОЯСНЮВАЛЬНА ЗАПИСКА

Граничні теореми теорії ймовірностей (посилені закони великих чисел).
Перевірка статистичних гіпотез (критерій незалежності χ^2)
(тема роботи)

Теорія ймовірностей і математична статистика
(дисципліна)

Керівник доц. каф. ПМ Гибкіна Н.В.
(підпис, дата, посада, прізвище, ініціали)

Студент Пригорко М.Ю. ППМ-16-1
(група, підпис, дата, прізвище, ініціали)

Робота захищена
з оцінкою________________
«___»______________2018 р.

Харків 2018

Харківський національний університет радіоелектроніки

Факультет Інформаційно-аналітичних технологій та менеджменту
Кафедра Прикладної математики
Дисципліна Теорія ймовірностей і математична статистика
Спеціальність 113 Прикладна математика
Курс 2 Група ППМ-16-1 Семестр 4

ЗАВДАННЯ
НА КУРСОВУ РОБОТУ

студентові Пригорко Максиму Юрійовичу
(прізвище, ім’я, по батькові)

1. Тема роботи Граничні теореми теорії ймовірностей (посилені закони великих чисел).
Перевірка статистичних гіпотез (критерій незалежності χ^2)

2. Термін здачі студентом закінченої роботи 28 травня 2018 р.

3. Вихідні дані до роботи: теоретичні питання та практичні завдання відповідно до варіанту

4. Зміст пояснювальної записки (перелік питань, які потрібно розробити)
Вступ
1 Теоретична частина
1.1 Граничні теореми теорії ймовірностей
1.2 Перевірка статистичних гіпотез
2 Практична частина
2.1 Розв’язання задач на тему «Основні ймовірнісні розподіли»
2.2 Розв’язання задач на тему «Граничні теореми теорії ймовірностей»
2.3 Перевірка гіпотез за допомогою критерію згоди Колмогорова: проста гіпотеза
Висновки

5. Дата видачі завдання 09 березня 2018 р.

КАЛЕНДАРНИЙ ПЛАН

Номер	Назва етапів курсової роботи	Термін виконання етапів роботи	Примітка
1	Вибір теми та отримання завдання на курсову роботу	4 тиждень	виконано
2	Підбір та вивчення літератури за темою роботи	5 – 6 тиждень	виконано
3	Виконання теоретичної частини курсової роботи	7 – 9 тижні	виконано
4	Виконання практичної частини курсової роботи	10 – 12 тижні	виконано
5	Оформлення пояснювальної записки	13 тиждень	виконано
6	Захист курсової роботи	14 тиждень	виконано

Студент ________________________________
(підпис)

Керівник роботи _________________________
(підпис)

РЕФЕРАТ

Пояснювальна записка: 49 с., 8 рис., 3 табл., 1 додаток, 12 джерел.

Під час виконання курсової роботи були отримані практичні навички з наступних тем: функукції випадкових величин, граничні теореми теорії ймовірностей. Також більш детально було розглянуто критерій незалежності χ^2.

ФУНКЦІЯ РОЗПОДІЛУ, ЩІЛЬНІСТЬ,ДИСПЕРСІЯ, КРИТЕРІЙ, ТЕОРЕМА, ЗБІЖНІСТЬ, ГІПОТЕЗА, ВИПАДКОВА ВЕЛИЧИНА

Зміст

ВСТУП

Виникнення теорії ймовірностей як науки відносять до середніх століть і першим спробам математичного аналізу азартних ігор (Орлянка, кістки, рулетка). Спочатку її основні поняття не мали строго математичного вигляду, до них можна було ставитися як до деяких емпіричних фактів, як до властивостей реальних подій, і вони формулювалися в наочних уявленнях. Найбільш ранні роботи вчених в галузі теорії ймовірностей відносяться до XVII століття. Досліджуючи прогнозування виграшу в азартних іграх, Блез Паскаль і П'єр Ферма відкрили перші імовірнісні закономірності, що виникають при киданні кісток. Під впливом піднятих і розглянутих ними питань вирішенням тих же завдань займався і Християн Гюйгенс. При цьому з листуванням Паскаля і Ферма він знайомий не був, тому методику рішення винайшов самостійно. Його робота, в якій вводяться основні поняття теорії ймовірностей, а також використовуються теореми додавання і множення ймовірностей, вийшла в друкованому вигляді на двадцять років раніше (1657 рік) видання листів Паскаля і Ферма (1679 рік).

Важливий внесок в теорію ймовірностей вніс Якоб Бернуллі: він дав доказ закону великих чисел в найпростішому випадку незалежних випробувань. У першій половині XIX століття теорія ймовірностей починає застосовуватися до аналізу помилок спостережень; Лаплас і Пуассон довели перші граничні теореми. У другій половині XIX століття основний внесок внесли російські вчені П. Л. Чебишев, А. А. Марков і А. М. Ляпунов. В цей час були доведені закон великих чисел, центральна гранична теорема, а також розроблена теорія ланцюгів Маркова. Сучасного вигляду теорія ймовірностей отримала завдяки аксіоматизації, запропонованої Андрієм Миколайовичем Колмогоровим. В результаті теорія ймовірностей придбала строгий математичний вигляд і остаточно стала сприйматися як один з розділів математики.

1 ТЕОРЕТИЧНА ЧАСТИНА

1.1 Граничні теореми теорії ймовірностей

1.1.1 Збіжність послідовностей випадкових величин та ймовірнісних розподілів

Розглянемо ймовірнісний простір $\langle \Omega, \mathsf{F}, \mathsf{P} \rangle$ і задані на ньому системи випадкових величин ξ_1, ξ_2, ..., ξ_n, ... та випадкову величину ξ. У теорії ймовірностей розглядають такі види збіжності послідовностей випадкових величин.

Означення 1.1. Послідовність випадкових величин ξ_n збігається за ймовірністю до випадкової величини ξ, якщо для будь-якого $\varepsilon > 0$

$$\lim_{n\to\infty} \mathsf{P}\left(|\xi_n - \xi| \geq \varepsilon\right) = 0 . \tag{1.1}$$

Цей вид збіжності позначають $\xi_n \xrightarrow{\mathsf{P}} \xi$ або $\mathsf{P}\lim_{n\to\infty} \xi_n = \xi$.

Означення 1.2. Послідовність випадкових величин ξ_n називається збіжною з ймовірністю одиниця (майже напевно, майже всюди) до випадкової величини ξ, якщо

$$\mathsf{P}\left\{\varpi : \xi_n \not\to \xi\right\} = 0 , n \to \infty, \tag{1.2}$$

тобто якщо безліч випадків ϖ, для яких $\xi_n(\varpi)$ не сходяться до $\xi(\varpi)$, має нульову ймовірність. Цей вид збіжності позначають у такий спосіб:

$$\xi_n \to \xi, \text{ або } \xi_n \xrightarrow{п.н.} \xi, \text{ або } \xi_n \xrightarrow{п.в.} \xi.$$

Означення 1.3. Послідовність випадкових величин ξ_n називається збіжною в середньому порядку p, $0 < p < \infty$, якщо

$$\mathsf{M}|\xi_n - \xi|^p \to 0,\ n \to \infty. \tag{1.3}$$

Означення 1.4. Послідовність випадкових величин ξ_n називається збіжною за розподілом до випадкової величини ξ (позначення: $\xi_n \xrightarrow{d} \xi$), якщо для будь-якої обмеженою неперервної функції $f = f(x)$

$$\mathsf{M}f(\xi_n) \to \mathsf{M}f(\xi),\ n \to \infty. \tag{1.4}$$

Збіжність за розподілом випадкових величин визначається тільки в термінах збіжності їх функцій розподілу. Тому про цей вид збіжності має сенс говорити і тоді, коли випадкові величини задані на різних імовірнісних просторах.

Теорема. а) Для того щоб $\xi_n \to \xi$, необхідно і достатньо, щоб для будь-якого $\varepsilon > 0$

$$\mathsf{P}\left\{\sup_{k \geq n} |\xi_k - \xi| \geq \varepsilon\right\} \to 0,\ n \to \infty. \tag{1.5}$$

б) Послідовність $\{\xi_n\}_{n \geq 1}$ фундаментальна з ймовірністю одиниця тоді і тільки тоді, коли

$$\mathsf{P}\left\{\sup_{k \geq 0} |\xi_{n+k} - \xi_n| \geq \varepsilon\right\} \to 0,\ n \to \infty, \tag{1.6}$$

для будь-кого $\varepsilon > 0$.

Теорема. Мають місце такі імплікації:

$$\xi_n \xrightarrow{п.н.} \xi \Rightarrow \xi_n \xrightarrow{\mathrm{P}} \xi, \tag{1.7}$$

$$\xi_n \xrightarrow{L^p} \xi \Rightarrow \xi_n \xrightarrow{\mathrm{P}} \xi, \ p > 0, \tag{1.8}$$

$$\xi_n \xrightarrow{\mathrm{P}} \xi \Rightarrow \xi_n \xrightarrow{d} \xi. \tag{1.9}$$

1.1.2 Метод характеристичних функцій

Метод характеристичних функцій був застосований математиком А.М. Ляпуновим для доведення центральної теореми. Надалі цей метод став застосовуватися і для вирішення інших імовірнісних задач, зокрема для відшукання моментів випадкової величини. Поряд з випадковими величинами теорія характеристичних функцій вимагає залучення комплекснозначних випадкових величин.

Означення. Характеристичною функцією випадково величини називається математичне сподівання випадкової величини e^{itX}, де i – уявна одиниця, t – дійсний параметр:

$$\varphi_X(t) = \mathrm{M}[e^{itX}]. \tag{1.11}$$

Дане визначення призводить нас до наступних формул для обчислення характеристичної функції:

– для дискретної випадково величини

$$\varphi_X(t) = \sum_k e^{itx_k} p_k, \tag{1.12}$$

де $p_k = \mathrm{P}(X = x_k)$, k=1,2,...;

– для неперервної випадкової величини

$$\varphi_X(t) = \int_{-\infty}^{+\infty} e^{itx} f(x) dx, \tag{1.13}$$

де *f(x)* – щільність ймовірності випадкової величини X .

Таким чином, характеристична функція неперервної випадкової величини є перетворенням Фур'є щільності ймовірності. З теорії перетворення Фур'є відомо, що щільність ймовірності можна виразити через характеристичну функцію зворотним перетворенням Фур'є:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \varphi_X(t) dt. \tag{1.14}$$

Розглянемо основні властивості характеристичної функції.

1. З формули (1.11) випливає, що

$$\varphi_X(0) = 1. \tag{1.15}$$

2. Характеристична функція випадкової величини $Z = aX + b$ виражається через характеристичну функцію випадкової величини X за формулою

$$\mathrm{M}[e^{it(aX+b)}] = \mathrm{M}[e^{itb} e^{i(at)X}] = e^{itb} \varphi_X(\mathrm{a}t). \tag{1.16}$$

3. Якщо у випадкової величини існує початковий момент k-го порядку $\alpha_k[X]$, то існує k-а похідна характеристичної функції:

$$\varphi_X^{(k)}(t) = \frac{d^k}{dt^k}\text{M}[e^{-itX}] = i^k\text{M}[X^k e^{-itX}]. \tag{1.17}$$

При $t = 0$ отримаємо:

$$\varphi_X^{(k)}(0) = i^k\text{M}[X^k]{=}i^k\alpha_k[X^k]. \tag{1.18}$$

Звідси

$$\alpha_k[X]{=}i^{-k}\varphi_X^{(k)}(0). \tag{1.19}$$

4. Характеристична функція суми взаємно незалежних випадкових величин $X_1,...,X_n$ дорівнює добутку характеристичних функцій доданків

$$\varphi_Z(t) = \prod_{k=1}^{n}\varphi_k(t). \tag{1.20}$$

1.1.3 Закони великих чисел

Розглянемо $\xi_1, \xi_2, ..., \xi_n, ...$ і $a_n = \text{M}\xi_n,\ n = 1,2....$

Означення. Цикли теорем, в яких сформульовані умови, при яких величина $\frac{1}{n}\sum_{i=1}^{n}\xi_i - \frac{1}{n}\sum_{i=1}^{n}\text{M}\xi_i$ збігається до нуля по ймовірності $(\forall\varepsilon > 0 : \lim_{n\to\infty} P(|\xi_n - \xi| \geq 0) = 0)$, називаються законами великих чисел (ЗВЧ).

Розглянемо $\xi_n = \frac{1}{n}\sum_{i=1}^{n}\xi_i$ і $\overline{a}_n = \frac{1}{n}\sum_{i=1}^{n}\mathrm{M}\xi_i$.

Теорема (закон великих чисел у формі Чебишова). Нехай $\xi_1, \xi_2, ..., \xi_n, ...$ – послідовність незалежних випадкових величин з скінченними математичними сподіваннями $a_i = \mathrm{M}\xi_i$ і обмеженою в сукупності дисперсією $(\mathrm{D}\xi \le \mathrm{A} < \infty)$.

Тоді $\forall \varepsilon > 0$

$$\lim_{n\to\infty} \mathrm{P}(|\xi_n - \overline{a}_n| \ge \varepsilon) = 0. \tag{1.21}$$

Теорема (закон великих чисел у формі Хинчина). Нехай $\xi_1, \xi_2, ..., \xi_n, ...$ – послідовність незалежних однаково розподілених випадкових величин з скінченними математичними сподіваннями. $a_i = \mathrm{M}\xi_i$, $i = 1, 2....$ Тоді $\forall \varepsilon > 0$:

$$\lim_{n\to\infty} P(|\frac{\xi_1 + \xi_2 + ... + \xi_n}{n} - a| \ge \varepsilon) = 0. \tag{1.24}$$

Теорема (закон великих чисел для схеми Бернуллі). Для будь-якого $\varepsilon > 0$

$$\mathrm{P}(|\upsilon_n - p| \ge \varepsilon) \underset{n\to\infty}{\to} 0, \tag{1.31}$$

де $\upsilon_n = \frac{1}{n}\sum_{i=1}^{n}\xi_i$ – частота появи успіху в випробуваннях.

Теорема (закон великих чисел у формі Маркова). Нехай випадкові величини $\xi_1, \xi_2, ..., \xi_n, ...$ як завгодно залежні і мають моменти другого порядку. Для виконання закону великих чисел достатньо, щоб $\frac{1}{n^2}\mathrm{D}(\sum_{i=1}^{n}\xi_i) \underset{n\to\infty}{\to} 0$.

Якщо $\xi_1, \xi_2, ..., \xi_n$ – незалежні, то співвідношення $\frac{1}{n^2} \mathrm{D}(\sum_{i=1}^{n} \xi_i) \underset{n\to\infty}{\to} 0$ приймає вигляд:

$$\frac{1}{n^2} \mathrm{D}(\sum_{i=1}^{n} \xi_i) = \frac{1}{n^2} \sum_{i=1}^{n} \mathrm{D}\xi_i \underset{n\to\infty}{\to} 0. \tag{1.32}$$

1.1.4 Посилені закони великих чисел

Посиленим законом великих чисел називається твердження, в якому збіжність за ймовірністю в (1.21) замінюються збіжністю з ймовірністю одиниця.

Теорема (Кантеллі). Нехай $\xi_1, \xi_2, \ldots$– незалежні випадкові величини з кінцевим четвертим моментом і такі, що для деякої константи

$$\mathrm{M} \,|\, \xi_n - \mathrm{M}\xi_n \,|^4 \le C,\ n \ge 1. \tag{1.33}$$

тоді при $n \to \infty$

$$\frac{S_n - \mathrm{M}S_n}{n} \to 0\,. \tag{1.34}$$

Доведення. Без обмеження спільності, будемо вважати $\mathrm{M}\xi_n = 0, \quad n \ge 1.$

Для збіжності $\frac{S_n}{n} \to 0$ досить, щоб для будь-якого $\varepsilon > 0$

$$\sum \mathrm{P}\left\{ \left| \frac{S_n}{n} \right| \ge \varepsilon \right\} < \infty$$

У свою чергу, в силу нерівності Чебишева, для цього достатньо виконання умови

$$\sum \mathrm{M}\left|\frac{S_n}{n}\right|^4 < \infty$$

Покажемо, що при зроблених припущеннях це умова дійсно виконано.

Маємо

$$S_n^4 = (\xi_1 + \ldots + \xi_n)^4 = \sum_{i=1}^{n} \xi_i^4 - \sum_{\substack{i,j \\ i<j}} \frac{4!}{2!\,2!}\xi_i^2\xi_j^2 + \sum_{\substack{i \neq j \\ l \neq k \\ i<k}} \frac{4!}{2!1!1!}\xi_i^2\xi_j\xi_k +$$

$$+ \sum_{i<j<k<l} 4!\xi_i\xi_j\xi_k\xi_l + \sum_{i \neq j} \frac{4!}{3!1!}\xi_i^3\xi_j.$$

Тоді, враховуючи, що $\mathrm{M}\xi_k = 0, \quad k \leq n$, звідси знаходимо

$$\mathrm{M}S_n^4 = \sum_{i=1}^{n} \mathrm{M}\xi_i^4 + 6 \sum_{i,j=1}^{n} \mathrm{M}\xi_i^2\mathrm{M}\xi_j^2 \leq nC + 6 \sum_{\substack{i,j=1 \\ i<i}} \sqrt{\mathrm{M}\xi_i^4 \cdot \mathrm{M}\xi_j^4} \leq nC +$$

$$+ \frac{6n(n-1)}{2}C = (3n^2 - 2n)C < 3n^2C.$$

отже,

$$\sum \mathrm{M}\left(\frac{S_n}{n}\right)^4 \le 3C\sum\frac{1}{n^2} < \infty.$$

Теорема доведена.

Теорема (Колмогорова). Нехай $\xi_1, \xi_2, \ldots$ – послідовність незалежних випадкових величин з кінцевими другими моментами, додатні числа b_n такі, що $b_n \uparrow \infty$ і

$$\sum\frac{\mathrm{D}_{\xi_n}}{b_n^2} < \infty\,. \tag{1.35}$$

Тоді

$$\frac{S_n - \mathrm{M}S_n}{b_n} \to 0\,. \tag{1.36}$$

Зокрема, якщо

$$\sum\frac{\mathrm{D}_{\xi_n}}{n^2} < \infty\,, \tag{1.37}$$

то

$$\frac{S_n - \mathrm{M}S_n}{n} \to 0\,. \tag{1.38}$$

Доведення. Для доведення цієї теореми, а також нижченаведеної теореми 3 нам знадобиться наступні два допоміжних твердження.

Лемма 1 (Тепліц). нехай $\{a_n\}$ – послідовність невід'ємних чисел,

$b_n = \sum_{i=1}^{n} a_i$, $b_n > 0$ для усіх $n \ge 1$ і $b_n \uparrow \infty, n \to \infty.$ нехай також $\{x_n\}$ – послідовність чисел, що сходиться до деякого числа x. Тоді

$$\frac{1}{b_n}\sum_{j=1}^{n} a_j x_j \to x.$$

Зокрема, якщо $a_n = 1$, то

$$\frac{x_1 + \ldots + x_n}{n} \to x.$$

Доведення. Нехай $\varepsilon > 0$ *и* $n_0 = n_0(\varepsilon)$ таке, що для всіх $n \ge n_0 \left|x_n - x\right| \le \varepsilon / 2.$ Виберемо $n_1 > n_0$ так, що

$$\frac{1}{b_n}\sum_{j=1}^{n_0}\left|x_j - x\right| < \varepsilon / 2.$$

Тоді для $n > n_1$

$$\left|\frac{1}{b_n}\sum_{j=1}^{n} a_j x_j - x\right| \le \frac{1}{b_n}\sum_{j=1}^{n} a_j \left|x_j - x\right| =$$

$$= \frac{1}{b_n}\sum_{j=1}^{n} a_j \left|x_j - x\right| + \frac{1}{b_n}\sum_{j=n_0+1}^{n} a_j \left|x_j - x\right| \le$$

$$\le \frac{1}{b_{n_1}}\sum_{j=1}^{n_0} a_j \left|x_j - x\right| + \frac{1}{b_n}\sum_{j=n_0+1}^{n} a_j \left|x_j - x\right| \le \frac{\varepsilon}{2} + \frac{n - n_0}{n}\frac{\varepsilon}{2} \le \varepsilon.$$

Лемма доведена.

Лемма (Кронекер). Нехай $\{b_n\}$ – послідовність додатних зростаючих чисел, $b_n \uparrow \infty, n \to \infty$, *и* $\{x_n\}$ – послідовність чисел таких, що ряд $\sum x_n$ збігається. Тоді $\frac{1}{b_n}\sum_{j=1}^{n} b_j x_j \to 0, \quad n \to \infty.$

Зокрема, якщо $b_n = n,\ x_n = \frac{y_n}{n}$ і ряд $\sum \frac{y_n}{n}$ збігається, то

$$\frac{y_1 + \ldots + y_n}{n} \to 0, \quad n \to \infty.$$

Доведення. Нехай $b_0 = 0, S_0 = 0,\ S_n = \sum_{i=1}^{n} x_j$. Тоді ("підсумовування по частинах")

$\sum_{i=1}^{n} b_j x_j = \sum_{i=1}^{n} b_j (S_j - S_{j-1}) = b_n S_n - b_0 S_0 - \sum_{i=1}^{n} S_{j-1}(b_j - b_{i-1})$ і, отже,

$$\frac{1}{b_n}\sum_{j=1}^{n} b_j x_j = S_n - \frac{1}{b_n}\sum_{j=1}^{n} S_{j-1} a_j \to 0,$$

оскільки, якщо $S_n \to x$, то по лемме Тепліца

$$\frac{1}{b_n}\sum_{l=1}^{n} S_{j-1} a_j \to x.$$

Лемма доведена.

Доведення теореми . Оскільки

$$\frac{S_n - \mathrm{M}S_n}{b_n} = -\frac{1}{b_n}\sum_{k=1}^{n} b_k \left(\frac{\xi_k - \mathrm{M}\xi_k}{b_k} \right),$$

то в силу леми Кронекера для виконання (4) досить, щоб (Р-м. н.) збігався ряд $\sum \frac{\xi_k - \mathrm{M}\xi_k}{b_k}$. Але цей ряд дійсно збігається в силу умови (3) і теореми.

Теорема доведена.

Теорема (Колмогорова) Нехай $\xi_1, \xi_2, \ldots$ – послідовність незалежних однаково розподілених випадкових величин з $\mathrm{M}|\xi_1| < \infty$. Тоді

$$\frac{S_n}{n} \to m \qquad \text{(Р-м.н.)}, \tag{1.39}$$

де $m = \mathrm{M}\xi_1$.

Зауваження. Затвердження теореми допускає звернення в наступному сенсі. нехай $\xi_1, \xi_2, \ldots$ – послідовність незалежних однаково розподілених випадкових величин, для яких з ймовірністю одиниця $\frac{\xi_1 + \ldots + \xi_n}{n} \to C$, де деяка (кінцева) константа. Тоді $\mathrm{M}|\xi_1| < \infty$ і $C = \mathrm{M}\xi_1$.

Таким чином, для незалежних однаково розподілених випадкових величин умова $\mathrm{M}|\xi_1| < \infty$ є необхідною і достатньою для збіжності (з ймовірністю одиниця) від $\frac{S_n}{n}$ до кінцевої межі.

Доведення. Для доказу нам знадобиться наступна Лема 3. Нехай ξ – невід’ємна випадкова величина. Тоді

$$\sum_{n=1}^{\infty} \mathrm{P}(\xi \geq n) \leq \mathrm{M}\xi \leq 1 + \sum_{n=1}^{\infty} \mathrm{P}(\xi \geq n).$$

Доведення випливає з наступного ланцюжка:

$$\sum_{n=1}^{\infty} \mathrm{P}(\xi \geq n) = \sum_{n=1}^{\infty} \sum_{k \geq n} \mathrm{P}(k \leq \xi < k+1) =$$

$$= \sum_{k=1}^{\infty} k\mathrm{P}(k \leq \xi < k+1) = \sum_{k=0}^{\infty} \mathrm{M}\left[kI(k \leq \xi < k+1)\right] \leq \sum_{k=0}^{\infty} \mathrm{M}\left[\xi I(k \leq \xi < k+1)\right] =$$

$$= \mathrm{M}\xi \leq \sum_{k=0}^{\infty} \mathrm{M}\left[(k+1)I(k \leq \xi < k+1)\right] = \sum_{k=0}^{\infty} (k+1)\mathrm{P}(k \leq \xi < k+1) =$$

$$= \sum_{n=1}^{\infty} \mathrm{P}(\xi \geq n) + \sum_{k=0}^{\infty} \mathrm{P}(k \leq \xi < k+1) = \sum_{n=1}^{\infty} \mathrm{P}(\xi \geq n) + 1.$$

Доведення теореми. В силу леми Кронекера і леми Бореля-Кантеллі

$$\mathrm{M}\left|\xi_1\right| < \infty \Leftrightarrow \sum \mathrm{P}\left\{\left|\xi_1\right| \geq n\right\} < \infty \Leftrightarrow$$

$$\Leftrightarrow \sum \mathrm{P}\left\{\left|\xi_n\right| \geq n\right\} < \infty \Leftrightarrow \mathrm{P}\left\{\left|\xi_n\right| \geq n \; б.ч.\right\} = 0.$$

Тому з імовірністю одиниця для всіх n, за винятком лише скінченного числа $\left|\xi_n\right| < n.$

Позначимо $\tilde{\xi}_n = \begin{cases} \xi_n, & \left|\xi_n\right| < n, \\ 0, & \left|\xi_n\right| \geq n, \end{cases}$ і будемо вважати, що $\mathrm{M}\xi_n = 0, n \geq 1.$ Тоді

$\dfrac{\xi_1 + \ldots + \xi_n}{n} \to 0$, якщо і тільки якщо $\dfrac{\tilde{\xi}_1 + \ldots + \tilde{\xi}_n}{n} \to 0$. Зауважимо, що, взагалі кажучи, $\mathrm{M}\tilde{\xi}_n \neq 0,$ але

$$\mathrm{M}\tilde{\xi}_n = \mathrm{M}\xi_n I(\left|\xi_n\right| < n) = \mathrm{M}\xi_1 I(\left|\xi_1\right| < n) \to \mathrm{M}\xi_1 = 0.$$

Тому за лемою Тепліца

$$\frac{1}{i}\sum_{k=1}^{n} \mathrm{M}\tilde{\xi}_k \to 0, \quad n \to \infty,$$

і, отже, $\frac{\xi_1 + \ldots + \xi_n}{n} \to 0$ в тому і тільки тому випадку, коли

$$\frac{(\tilde{\xi}_1 - \mathrm{M}\tilde{\xi}_1) + \ldots + (\tilde{\xi}_n - \mathrm{M}\tilde{\xi}_n)}{n} \to 0, \quad n \to \infty.$$

Позначимо $\tilde{\xi}_n = \tilde{\xi}_n - \mathrm{M}\tilde{\xi}_n$. В силу леми Кронекера для виконання (14) досить лише встановити, що ряд $\sum \frac{\xi_n}{n}$ збігається. У свою чергу, досить показати, що припущення $\mathrm{M}|\xi_1| < \infty$ забезпечує збіжність ряду $\sum \frac{\mathsf{D}\xi_n}{n^2}$.

Маємо

$$\sum \frac{\mathsf{D}\xi_n}{n^2} \le \sum_{n=1}^{\infty\infty} \frac{\mathrm{M}\xi_n^2}{n^2} = \sum_{n=1}^{\infty} \frac{1}{n^2}\mathrm{M}\left[\xi_n I(|\xi_n| < n)\right]^2 =$$

$$= \sum_{n=1}^{\infty} \frac{1}{n^2}\mathrm{M}\left[\xi_1^2 I(|\xi_1| < n)\right] = \sum_{n=1}^{\infty} \frac{1}{n^2} \sum_{k=1}^{n} \mathrm{M}\left[\xi_1^2 I(k-1 \le |\xi_1| < k)\right] =$$

$$= \sum_{k=1}^{\infty} \mathrm{M}\left[\xi_1^2 I(k-1 \le |\xi_1| < k)\right]. \sum_{n=k}^{\infty} \frac{1}{n^2} \le$$

$$\le 2\sum_{k=1}^{\infty} \frac{1}{k}\mathrm{M}\left[\xi_1^2 I(k-1 \le |\xi_1| < k)\right] \le$$

$$\le 2\sum_{k=1}^{\infty} \mathrm{M}\left[|\xi_1| I(k-1 \le |\xi_1| < k)\right] = 2\mathrm{M}|\xi_1| < \infty.$$

Теорема доведена.

1.1.5 Центральна гранична теорема для незалежних однаково розподілених випадкових величин

Нехай $\xi_1,\xi_2,\ldots,\xi_n,\ldots$ – послідовність незалежних, однаково розподілених випадкових величин, яка має скінчені математичне сподівання $M\xi_n = a$ та дисперсію $D\xi_n = \sigma^2$, $S_n = \xi_1 + \xi_2 + \ldots + \xi_n$, а $\Phi(x)$ – функція розподілення нормального закону з параметрами (0,1). Введемо ще послідовність випадкових величин

$$\xi_n = \frac{S_n - an}{\sigma\sqrt{n}}. \tag{1.23}$$

Теорема 1.11. *Якщо* $0 < \sigma^2 < \infty$, то при $n \to \infty$ $P(\xi_n < x) \to \Phi(x)$ рівномірно відносно x $-\infty < x < \infty$.

В цьому випадку послідовність $\{\xi_n\}$ називається асимптотично нормальною.

З того, що $M\xi_n^2 = 1$та з теорем неперервності випливає, що разом із слабкою збіжністю $\xi_n \Rightarrow \xi, \xi \in \Phi_{0,1}(Mf(\xi_n) \to Mf(\xi)$ для будь-якої неперервної обмеженої f має місце також збіжність $Mf(\xi_n) \to Mf(\xi)$для будь-якої неперервної f, такої, що $\left|f(x)\right| < c\left(1 + |x|^{2-\varepsilon}\right)$ при якому-небудь $\varepsilon > 0$ $\varepsilon > 0$.

1.1.6 Теореми Ляпунова і Ліндеберга

Розглянемо умови, в яких ЦГТ має місце в тому випадку коли випадкові величини ξ_1, ξ_2, ..., ξ_n незалежні, і мають різні розподіли. Визначимо $\mathrm{M}\xi_k = a_k$

$, \mathrm{D}\xi_k = \sigma_k^2,\ B_n^2 = \sum_{k=1}^{n} \sigma_k^2 = \sum_{k=1}^{n} \mathrm{D}\xi_k = \mathrm{D}\sum_{k=1}^{n} \xi_k .$

Теорема (Лінденберга). Якщо послідовність незалежних випадкових величин $\xi_1, \xi_2, ..., \xi_n ...$ при будь-яку постійну $\tau > 0$ задовольняє умові Лінденберга

$$\lim_{n\to\infty} \frac{1}{B_n} \sum_{k=1}^{n} \int\limits_{|x-a_k|>\tau B_n} \left(x - a_k\right)^2 dF_k(x) = 0, \tag{1.44}$$

то рівномірно щодо x

$$\mathrm{P}\left\{ \frac{1}{B_n} \sum_{k=1}^{n} \left(\xi_k - a_k\right) < x \right\} \to \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{\frac{-z^2}{2}} dz . \tag{1.45}$$

Теорема (Ляпунова). Якщо послідовність взаємно незалежних випадкових величин $\xi_1, \xi_2, ..., \xi_n ...$ можна підібрати таке додатне $\delta > 0$, що при $n \to \infty$

$$\frac{1}{B_n^{2+\delta}} \sum_{k=1}^{n} \mathrm{M}\left|\xi_k - a_k\right|^{2+\delta} \to 0, \tag{1.46}$$

то при $n \to \infty$ рівномірно по x

$$\mathrm{P}\left\{ \frac{1}{B_n} \sum_{k=1}^{n} \left(\xi_k - a_k\right) < x \right\} \to \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{\frac{-z^2}{2}} dz. \tag{1.47}$$

1.2 Перевірка статистичних гіпотез

1.2.1 Основні задачі математичної статистики та їх стисла характеристика

Математична статистика - розділ математики, який займається розробкою методів отримання науково обгрунтованих висновків про масові процеси та явища за даними спостережень або експериментів.

Основними завданнями математичної статистики є наступні.

1) Завдання визначення закону розподілу випадкової величини за статистичними даними.

При обробці великих за своїм обсягом статистичних даних часто виникає питання про визначення законів розподілу тих чи інших випадкових величин. Виникає характерна для математичної статистики задача згладжування або вирівнювання статистичних даних, подання їх в найбільш компактному вигляді за допомогою простих аналітичних залежностей.

2) Перевірка статистичних гіпотез.

При вирішенні такого роду завдань ми зазвичай не маємо настільки великих статистичних матеріалов, щоб виявлені в ньому статистичні закономірності були в достатній мірі вільні від елементів випадковості. Статистичний матеріал може з більшою чи меншою правдоподібністю підтверджувати або не підтверджує справедливість тієї чи іншої гіпотези.

3) Оцінка невідомих параметрів.

Виникає вузька задача обробки спостережень – визначити тільки деякі параметри випадкової величини або системи випадкових величин. При невеликому числі дослідів, завдання точного визначення цих параметрів не може бути вирішеним; в цих випадках експериментальний матеріал містить в собі неминуче значний елемент випадковості; тому випадковими виявляються і всі параметри, обчислені на основі цих даних.

1.2.2 Перевірка статистичних гіпотез: основні поняття

Означення. Статистичної гіпотезою (H) називається будь-яке припущення щодо розподілу випадкової вибірки яка спостерігається в стохастичному експерименті.

Означення. Параметрична гіпотеза – статистична гіпотеза щодо невідомого істинного значення параметра θ.

Означення. Гіпотеза, що перевіряється називається основною або нульовою (H_0).

Означення . Гіпотеза, що конкурує з H_0, називається альтернативною $\left(H_1\right)$

Означення. Статистична гіпотеза $H_0 : \theta \in \theta_0$ називається простою гіпотезою, якщо множина θ_0 є одноелементною, причому розподіл генеральної сукупності повністю відомий. В іншому випадку гіпотеза називається складною.

Означення. Статистичний критерій перевірки гіпотези H_0 – це правило, відповідно до якого на підставі спостережень $z_{спост.} = \varphi(\vec{X}_n)$ в статистиці ξ гіпотеза H_0 приймається або відкидається.

Означення. Критична область $\bar{G}$ статистичного критерію – це область тих значень $z_{спост.}$ статистики ξ, при яких гіпотеза H_0 відкидається.

Означення. Довірча область G статистики критерію – це область тих значень $z_{спост.}$ статистики ξ, при яких гіпотеза H_0 приймається.

При цьому $D = G \cup \bar{G}$.

Загальна схема перевірки статистичних гіпотез:

1) сформулювати основну гіпотезу H_0 та альтернативну H_1;

2) вибрати рівень значимості α;

3) вибрати статистику $\varsigma = \varphi\left(\overrightarrow{X}_n\right)$ для перевірки гіпотези H_0;

4) знайти розподіл цієї статистики $F_{\xi}\left(z,H_0\right)$ за умови, що вірна H_0;

5) в залежності від сформульованої гіпотези H_1 побудувати критичну область $\overline{G}$;

6) отримати вибірку спостережень і розрахувати на підставі цієї вибірки спостережуване значення статистики;

7) прийняти статистичний висновок на рівні значущості $1-\alpha$: якщо спостережуване значення статистики належить $z_{спост.}\in\overline{G}$, то слід відхилити гіпотезу H_0, як таку, яка не узгоджується з результатами спостережень; якщо $z_{спост.}\in G$, то слід прийняти гіпотезу H_0, як таку, що не суперечить результатам спостережень.

1.2.3 Критерій незалежності χ^2

Простий критерій згоди для гіпотези незалежності H_0 можна побудувати, виходячи з методики хі-квадрат. Цю методику застосовують для дискретних моделей з кінцевим числом результатів, тому домовимося вважати, це випадкова величина ξ_1 приймає кінцеве число деяких значень, які будемо позначати $a_1,...,a_s$, а друга компонента ξ_2 – k деяких значень $b_1,...,b_k$. Якщо дані мають іншу структуру, то їх попередньо групують окремо по першій і другій компоненті. У цьому випадку безліч можливих значень ξ_1 розбивається на s непересічних інтервалів $\varepsilon_1^{(1)},\ldots,\varepsilon_s^{(1)}$, множина значень ξ_2 - на k інтервалів $\varepsilon_1^{(2)},\ldots,\varepsilon_s^{(2)}$ і тим самим множина значень двовимірної величини $\xi=(\xi_1,\xi_2)$ розбивається на $N=sk$ прямокутників $\varepsilon_i^{(1)}\times\varepsilon_j^{(2)}$.

Нехай v_{ij} є число спостережень пари (a_i,b_j) (число елементів вибірки, що належать прямокутнику $\varepsilon_i^{(1)} \times \varepsilon_j^{(2)}$, якщо дані групуються), так що

$$\sum_{i=1}^{s}\sum_{j=1}^{k} v_{ij} = n.$$

Результати спостережень зручно для наочності розташувати у вигляді таблиці спряженості двох ознак (табл 1.1)

Таблиця 1.1

ξ_1	ξ_2				$\sum$
	b_1	b_2	$\dots$	b_k	
a_1	v_{11}	v_{12}	$\dots$	v_{1k}	$v_{1\bullet}$
a_2	v_{21}	v_{22}	$\dots$	v_{2k}	$v_{2\bullet}$
a_s	v_{s1}	v_{s2}	$\dots$	v_{sk}	$v_{s\bullet}$
$\sum$	$v_{\bullet 1}$	$v_{\bullet 2}$	$\dots$	$v_{\bullet k}$	n

Нехай $p_{ij} = \mathrm{P}\left\{\xi_1 = a_i, \xi_2 = b_j\right\}, i = 1,...,s, j = 1,...,k.$ Тоді гіпотеза незалежності означає, що існує $s + k$ сталих $p_{i\bullet}, p_{\bullet j}$ таких, що

$$\sum_{i=1}^{s} p_{i\bullet} = \sum_{j=1}^{k} p_{\bullet j} = 1 \text{ и } p_{ij} = p_{i\bullet} p_{\bullet j}.$$

Оскільки спостереження незалежні за умовою, то частоти $(v_{ij}, i = 1,\dots,s, j = 1,\dots,k)$ розподілені по поліномному закону і при гіпотезі H_0

ймовірності наслідків p_{ij} мають зазначену специфічну структуру (вони визначаються значеннями $r=s+k-2$ невідомих параметрів $p_{i\bullet}, i=1,\ldots,s, j=1,\ldots,k-1$). Отже, гіпотеза незалежності в даному випадку є специфічною формою складної гіпотези для поліномної моделі. Щоб застосувати критерій хі-квадрат, треба спочатку знайти оцінки максимальної правдоподібності для визначальних розглянуту модель невідомих параметрів при справедливості гіпотези H_0, тобто максимізуючи в даному випадку функцію правдоподібності

$$\prod_{i,j}(p_{i\bullet}p_{\bullet j})^{\nu_{ij}} = \prod_i p_{i\bullet}^{\nu_{i\bullet}} \prod_j p_{\bullet j}^{\nu_{\bullet j}} \cdot$$

за параметрами $p_{i\bullet}, p_{\bullet j}$. Звідси маємо

$$\hat{p}_{i\bullet} = \frac{\nu_{i\bullet}}{n}, \quad \hat{p}_{\bullet j} = \frac{\nu_{\bullet j}}{n},$$

тобто це звичайні оцінки параметрів поліномних розподілів

$$£(\nu_{1\bullet},\ldots,\nu_{s\bullet}) = M(n; p_{1\bullet},\ldots,p_{s\bullet}).$$

і

$$£(\nu_{\bullet 1},\ldots,\nu_{\bullet k}) = M(n; p_{\bullet 1},\ldots,p_{\bullet k}).$$

З огляду на це тестова статистика набирає вигляду

$$\hat{\chi}_n^2 = \sum_{i,j}\frac{(\nu_{ij}-n\hat{p}_{i\bullet}\hat{p}_{\bullet j})^2}{n\hat{p}_{i\bullet}\hat{p}_{\bullet j}} = n\sum_{i=1}^{s}\sum_{j=1}^{k}\frac{1}{\nu_{i\bullet}\nu_{\bullet j}}\left(\nu_{ij}-\frac{\nu_{i\bullet}\nu_{\bullet j}}{n}\right)^2 .$$

з числом ступенів свободи

$$N-1-r=sk-1-(s+k-2)=(s-1)(k-1).$$

Отже, асимптотичний (при великих n) варіант критерію незалежності χ^2 при заданому рівні значущості α має вигляд

$$H_0 \text{ відкидається, якщо } \hat{\chi}_n^2 > \chi^2_{1-\alpha,(s-1)(k-1)},$$

де $\chi^2_{1-\alpha,(s-1)(k-1).}$

2 ПРАКТИЧНА ЧАСТИНА

2.1 Задача на тему «Основні розподіли теорії ймовірностей»

Задача 1. При $\beta = 5$дослідити усічений вправо нормальний розподіл з щільністю:

$$p_\xi(x) = \frac{1}{\gamma\sigma\sqrt{2\pi}} e^{-\frac{(x-\alpha)^2}{2\sigma^2}}, \ x \le \beta, \ \sigma > 0,$$

де $\gamma = \frac{1}{2} + \Phi_0(\frac{\beta - a}{\sigma})$.

зробити наступне.

1. Дати визначення випадкової величини, що має заданий розподіл.

2. Побудувати графіки щільності розподілу $p_\xi(x)$:

а) при фіксованому значенні параметра $a = 2$ для наступних значень параметра b (5 графіків на одному рисунку): $\sigma = \frac{1}{2}, \sigma = 1, \sigma = 2, \sigma = 3, \sigma = 4;$

б) при фіксованому значенні параметра $\sigma = 3$ для наступних значень параметра a (5 графіків на одному рисунку): $a = -2, a = -1, a = 0, a = 1, a = 2;$

3. Визначити функцію розподілу $F_\xi(x)$ випадкової величини ξ.

4. Побудувати графік функції розподілу $F_\xi(x)$.

а) при фіксованому значенні параметра $a = 2$ для наступних значень параметра b (5 графіків на одному рисунку): $\sigma = \frac{1}{2}, \sigma = 1, \sigma = 2, \sigma = 3, \sigma = 4;$

б) при фіксованому значенні параметра $\sigma = 3$ для наступних значень параметра a (5 графіків на одному рисунку): $a = -2, a = -1, a = 0, a = 1, a = 2;$

5. Визначити:

а) математичне сподівання, дисперсію та середньоквадратичне відхилення;

б) коефіцієнт варіації;

в) асиметрію та ексцес;

г) моду;

д) p-квантиль, верхню і нижню квартилі та медіану;

е) характеристичну функцію.

6. Навести значення числових характеристик, розрахованих у п.5, для випадку параметрів $a=1, \sigma=2$. Побудувати графіки щільності розподілу та функції розподілу, позначити на графіках математичне сподівання, моду та квартилі.

7. Побудувати щільність розподілу випадкової величини $\eta=f(\xi)$, де випадкова величина ξ має рівномірний розподіл на проміжку $[a,b]$, за умови, що:

а) $\eta=3\xi+1;$ б) $\eta=\sqrt{|x|}$.

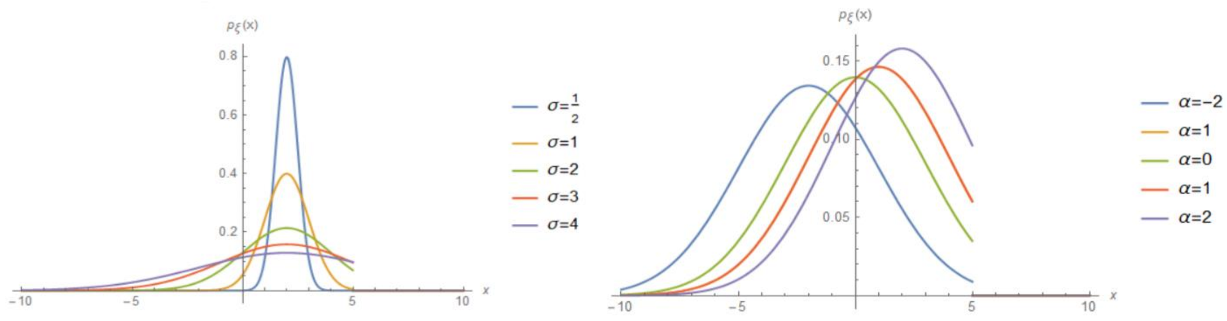
8. Навести галузі застосування та типові задачі, у яких використовується заданий закон розподілу.

Розв’язання.

1.Усіченим нормальним розподілом випадкової величини називається розподіл, що отримується з нормального при обмеженні інтервалу можливих значень цієї величини. Цей розподіл має щільність:

$$p_{\xi}(x)=\frac{1}{\gamma\sigma\sqrt{2\pi}}e^{-\frac{(x-\alpha)^2}{2\sigma^2}}, x\le\beta, \sigma>0.$$

2. Графіки щільності розподілу $p_{\xi}(x)$ при фіксованому $a=2$ та різних значеннях параметру σ наведено на рис. 2.1(а). Графіки щільності розподілу $p_{\xi}(x)$ при фіксованому $\sigma=3$ та різних значеннях параметру a наведено на рис.2.1(б).



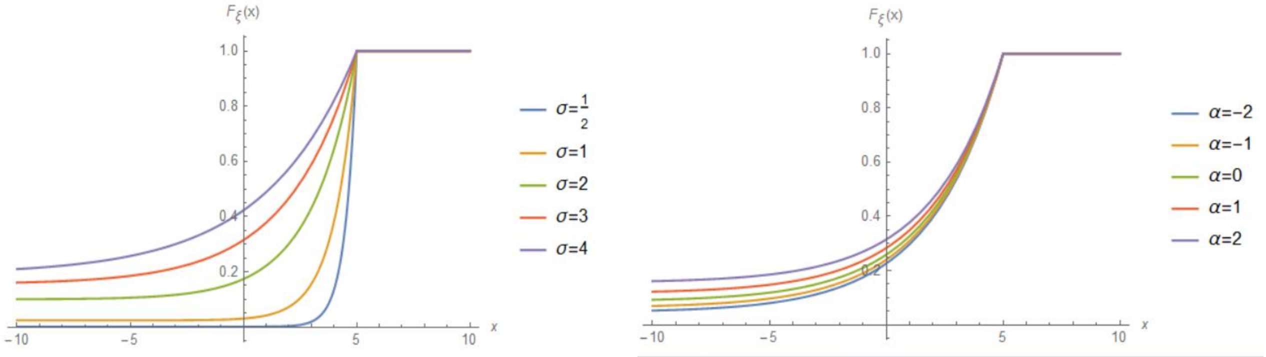
а) $a=2$ б) $\sigma=3$

Рисинок 2.1 Графік щільності $p_\xi(x)$ усіченого вправо нормального розподілу

3. Оскільки щільність випадкової величини ξ, що має усічений вправо нормальний розподіл, визначена при $x\in(-\infty;\beta)$, то її функція розподілу матиме вигляд:

$$F_\xi(x)=\int_{-\infty}^{x}\frac{1}{\gamma\sigma\sqrt{2\pi}}e^{-\frac{(\mathrm{t}-\alpha)^2}{2\sigma^2}}dt=\left|\begin{matrix} z=\dfrac{t-a}{\sigma} \\ \dfrac{x-a}{\sigma}<z<-\infty \end{matrix}\right|=\frac{1}{\gamma\sigma\sqrt{2\pi}}\int_{-\infty}^{x}\sigma e^{-\frac{z^2}{2}}dz=$$
$$=\frac{1}{\gamma}\left(\Phi_0\left(\frac{x-a}{\sigma}\right)+\frac{1}{2}\right).$$

Графіки функції розподілу $F_\xi(x)$ при фіксованому $a=2$ та різних значеннях параметру σ наведено на рис. 2.2(а). Графіки функції розподілу $F_\xi(x)$ при фіксованому $\sigma=3$ та різних значеннях параметру а наведено на рис. 2.2(б).



а) $a=2$ б) $\sigma=3$

Рисунок 2.2 Графік функції розподілу $F_\xi(x)$ усіченого вправо нормального розподілу

5. а) Математичне сподівання випадкової величини ξ дорівнює:

$$\mathrm{M}_\xi=\int_{-\infty}^{\beta}x\frac{1}{\gamma\sigma\sqrt{2\pi}}e^{\frac{-(\mathrm{x}-\alpha)^2}{2\sigma^2}}dx=\left|\begin{matrix}z=\dfrac{x-a}{\sigma} & x=\sigma z+a\\ dx=\sigma dz & \dfrac{\beta-a}{\sigma}<z<-\infty\end{matrix}\right|=$$

$$=\frac{1}{\gamma\sigma\sqrt{2\pi}}\int_{-\infty}^{\frac{\beta-a}{\sigma}}(\sigma z+a)\,e^{-\frac{z^2}{2}}\sigma dz=\frac{\sigma}{\gamma\sqrt{2\pi}}\int_{-\infty}^{\frac{\beta-a}{\sigma}}z\,e^{-\frac{z^2}{2}}dz-\frac{a}{\gamma}\left(\Phi_0\left(\frac{\beta-a}{\sigma}\right)+\frac{1}{2}\right)=$$

$$=-\frac{\sigma}{\gamma\sqrt{2\pi}}e^{\frac{(\beta-a)^2}{2\sigma^2}}+a=a-\frac{\sigma}{\gamma}p\left(\frac{\beta-a}{\sigma}\right).$$

Дисперсія та середньоквадратичне відхилення цієї величини дорівнюють відповідно:

$$\mathrm{D}_{\xi} = \int\limits_{-\infty}^{\beta} x^2 \frac{1}{\gamma\sigma\sqrt{2\pi}} e^{-\frac{(\mathrm{x}-\alpha)^2}{2\sigma^2}} dx - (a - \frac{\sigma}{\gamma} p(\frac{\beta - a}{\sigma}))^2 = \left| \begin{matrix} z = \dfrac{x-a}{\sigma} & x = \sigma z + a \\ dx = \sigma dz & \dfrac{\beta - a}{\sigma} < z < -\infty \end{matrix} \right| =$$

$$= \frac{1}{\gamma\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{\frac{\beta-a}{\sigma}} (\sigma z + a)^2 \cdot e^{-\frac{z}{2}} \sigma dz - (a - \frac{\sigma}{\gamma} p(\frac{\beta - a}{\sigma}))^2 = \frac{\sigma^2}{\gamma\sqrt{2\pi}} \int_{-\infty}^{\frac{\beta-a}{\sigma}} z^2 e^{-\frac{z^2}{2}} dz +$$

$$+ \frac{2\sigma a}{\gamma\sqrt{2\pi}} \int_{-\infty}^{\frac{\beta-a}{\sigma}} z e^{-\frac{z^2}{2}} + \frac{a^2}{\gamma\sqrt{2\pi}} \int_{-\infty}^{\frac{\beta-a}{\sigma}} e^{-\frac{z^2}{2}} = \frac{2\sigma a}{\gamma} p\left(\frac{\beta - a}{\sigma}\right) + a^2 -$$

$$- \frac{\sigma^2}{\gamma\sqrt{2\pi}} \left(\frac{\beta - a}{\sigma}\right) e^{\frac{\left(\frac{\beta-a}{\sigma}\right)^2}{2}} + \left(\Phi_0\left(\frac{\beta - a}{\sigma}\right) + \frac{1}{2}\right) - \left(a - \frac{\sigma}{\gamma} p\left(\frac{\beta - a}{\sigma}\right)\right)^2 =$$

$$= -\frac{\sigma^2}{\gamma} \cdot \left(\frac{\beta - a}{\sigma}\right) p\left(\frac{\beta - a}{\sigma}\right) + \gamma + \frac{2\sigma a}{\gamma} p\left(\frac{\beta - a}{\sigma}\right) + a^2 - \left(a - \frac{\sigma}{\gamma} p\left(\frac{\beta - a}{\sigma}\right)\right)^2 =$$

$$= \sigma^2 - \left(\beta - \left(a - \frac{\sigma}{\gamma} p\left(\frac{\beta - a}{\sigma}\right)\right)\right)\left(a - \left(a - \frac{\sigma}{\gamma} p\left(\frac{\beta - a}{\sigma}\right)\right)\right);$$

$$\sigma_x = \sqrt{\sigma^2 - \left(\beta - \left(a - \frac{\sigma}{\gamma} p\left(\frac{\beta - a}{\sigma}\right)\right)\right)\left(a - \left(a - \frac{\sigma}{\gamma} p\left(\frac{\beta - a}{\sigma}\right)\right)\right)} =$$

$$= \sigma - \sqrt{\left(\beta - \left(a - \frac{\sigma}{\gamma} p\left(\frac{\beta - a}{\sigma}\right)\right)\right)\left(a - \left(a - \frac{\sigma}{\gamma} p\left(\frac{\beta - a}{\sigma}\right)\right)\right)}.$$

б) Коефіцієнт варіації випадкової величини ξ дорівнює:

$$\mathrm{V} = \frac{\sigma_{\xi}}{\mathrm{M}_{\xi}} = \frac{\sigma - \sqrt{\left(\beta - \left(a - \frac{\sigma}{\gamma} p\left(\frac{\beta - a}{\sigma}\right)\right)\right)\left(a - \left(a - \frac{\sigma}{\gamma} p\left(\frac{\beta - a}{\sigma}\right)\right)\right)}}{\left(a - \frac{\sigma}{\gamma} p\left(\frac{\beta - a}{\sigma}\right)\right)}.$$

в) Для обчислення коефіцієнтів асиметрії та ексцесу випадкової величини ξ, розрахуємо центральні моменти μ_3 та μ_4 цієї випадкової величини :

$$\mu_3 = a_3 - 3a_2\left(a - \frac{\sigma}{\gamma}p\left(\left(\frac{\beta-a}{\sigma}\right)\right)\right) + 2\left(a - \frac{\sigma}{\gamma}p\left(\left(\frac{\beta-a}{\sigma}\right)\right)\right)^3 =$$

$$= \left(a - \mathrm{M}_\xi\right)\cdot \mathrm{D}_\xi + \left(\mathrm{M}_\xi - a\right)\left(\beta - \mathrm{M}_\xi\right)^2;$$

$$\mu_4 = a_4 - 4a_3\left(a - \frac{\sigma}{\gamma}p\left(\left(\frac{\beta-a}{\sigma}\right)\right)\right) + 6a_2\left(a - \frac{\sigma}{\gamma}p\left(\left(\frac{\beta-a}{\sigma}\right)\right)\right)^2 -$$

$$-3\left(a - \frac{\sigma}{\gamma}p\left(\left(\frac{\beta-a}{\sigma}\right)\right)\right)^4 = \left(a - \mathrm{M}_\xi\right)\cdot \mu_3 + 3\sigma^2 \mathrm{D}_\xi + \left(\mathrm{M}_\xi - a\right)\left(\beta - \mathrm{M}_\xi\right)^3;$$

Тепер коефіцієнти асиметрії та ексцесу дорівнюватимуть, відповідно:

$$A_\xi = \frac{\mu_3}{\sigma_\xi^3} = \frac{\left(a - \mathrm{M}_\xi\right)\cdot \mathrm{D}_\xi + \left(\mathrm{M}_\xi - a\right)\left(\beta - \mathrm{M}_\xi\right)^2}{\left(\sigma - \sqrt{\left(\beta - \left(a - \frac{\sigma}{\gamma}p\left(\frac{\beta-a}{\sigma}\right)\right)\right)\left(a - \left(a - \frac{\sigma}{\gamma}p\left(\frac{\beta-a}{\sigma}\right)\right)\right)}\right)^3} =$$

$$= \left(\frac{a - \mathrm{M}_\xi}{\sigma_\xi}\right)\left(1 - \frac{\left(\mathrm{M}_\xi - \beta\right)^2}{\mathrm{D}_\xi}\right);$$

$$E_\xi = \frac{\mu_4}{\sigma_\xi^4} - 3 = \frac{\left(a - \mathrm{M}_\xi\right)\cdot \mu_3 + 3\sigma^2 \mathrm{D}_\xi + \left(\mathrm{M}_\xi - a\right)\left(\beta - \mathrm{M}_\xi\right)^3}{\left(\sigma - \sqrt{\left(\beta - \left(a - \frac{\sigma}{\gamma}p\left(\frac{\beta-a}{\sigma}\right)\right)\right)\left(a - \left(a - \frac{\sigma}{\gamma}p\left(\frac{\beta-a}{\sigma}\right)\right)\right)}\right)^4} - 3 =$$

$$= \left(\frac{a - \mathrm{M}_\xi}{\mathrm{D}_\xi}\right)\left(a - 4\mathrm{M}_\xi + 3\beta - \frac{\left(\beta - \mathrm{M}_\xi\right)^2}{\mathrm{D}_\xi}\left(a - 2\mathrm{M}_\xi + \beta\right)\right).$$

г) Розглянемо щільність випадкової величини ξ, що має усічений вправо на проміжку $\left(-\infty;\beta\right)$ розподіл:

$$p_\xi(x) = 0,$$

$$p_{\xi}(x)=\frac{1}{\gamma\sigma\sqrt{2\pi}}e^{-\frac{(x-\alpha)^2}{2\sigma^2}}=0,$$

тоді

$$M_0=\begin{cases}a, & \beta\geq a;\\ \beta, & \beta<a\end{cases}.$$

д) Для визначення квантиля рівня p рівномірного розподілу необхідно розв’язати рівняння $F_{\xi}(x_p)=p$, яке має вигляд:

$$\frac{1}{\gamma}\left(\Phi_0\left(\frac{x_p-a}{\sigma}\right)+\frac{1}{2}\right)=p.$$

При $p=\frac{1}{4}, p=\frac{1}{2}, p=\frac{3}{4}$ дістанемо квартилі рівномірного розподілу:

- нижня квартиль: $x_{\frac{1}{4}}=\Phi_0\left(\frac{x_{\frac{1}{4}}-a}{\sigma}\right)=\frac{1}{4}+\frac{1}{4}\Phi_0\left(\frac{\beta-a}{\sigma}\right)$;

- медіана: $x_{\frac{1}{2}}=\Phi_0\left(\frac{x_{\frac{1}{2}}-a}{\sigma}\right)=\frac{1}{4}+\frac{1}{2}\Phi_0\left(\frac{\beta-a}{\sigma}\right)$;

- верхня квартиль: $x_{\frac{3}{4}}=\Phi_0\left(\frac{x_{\frac{3}{4}}-a}{\sigma}\right)=\frac{1}{4}+\frac{3}{4}\Phi_0\left(\frac{\beta-a}{\sigma}\right)$.

е) Характеристична функція рівномірного розподілу дорівнює:

$$f_\xi(t)=\int_{-\infty}^{\beta} e^{itx}\frac{1}{\gamma\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}dx=\frac{1}{\gamma\sigma\sqrt{2\pi}}\int_{-\infty}^{\beta} e^{\frac{itx2\sigma^2-x^2+2xa-a^2}{2\sigma^2}}dx=\frac{1}{\gamma\sigma\sqrt{2\pi}}\int_{-\infty}^{\beta} e^{\frac{-x^2-2x(a+it\sigma^2)+a^2}{2\sigma^2}}dx=$$

$$=\frac{1}{\gamma\sigma\sqrt{2\pi}}\int_{-\infty}^{\beta} e^{\frac{-x^2-2x(a+it\sigma^2)+a^2+(\mathrm{a}+\mathrm{it}\sigma^2)^2-(\mathrm{a}+\mathrm{it}\sigma^2)^2}{2\sigma^2}}dx=\frac{1}{\gamma\sigma\sqrt{2\pi}}e^{\frac{-2ait\sigma^2+(\mathrm{it}\sigma^2)^2}{2\sigma^2}}\int_{-\infty}^{\beta} e^{\frac{-(\mathrm{x}-(\mathrm{a}+\mathrm{it}\sigma^2))^2}{2\sigma^2}}dx=$$

$$=\frac{1}{\gamma\sigma\sqrt{2\pi}}e^{-ait+\frac{(\mathrm{it}\sigma^2)^2}{2\sigma^2}}\int_{-\infty}^{\beta} e^{\frac{-(\mathrm{x}-(\mathrm{a}+\mathrm{it}\sigma^2))^2}{2\sigma^2}}dx=\frac{\sigma\sqrt{2}}{\gamma\sigma\sqrt{2\pi}}e^{ait-\frac{(\mathrm{it}\sigma^2)^2}{2\sigma^2}}\int_{-\infty}^{\beta} e^{\frac{-(\mathrm{x}-(\mathrm{a}+\mathrm{it}\sigma^2))^2}{2\sigma^2}}dx=$$

$$=\left|\begin{matrix}\frac{x-a(it\sigma^2)}{\sigma}=z & dx=\sigma dz\\ x=z\sigma+a(it\sigma^2) & \frac{\beta-a}{\sigma}<z<-\infty\end{matrix}\right|=\frac{1}{\gamma}\left(\frac{1}{2}+\Phi_0\left(\frac{\beta-a}{\sigma}-i\sigma t\right)\right)e^{i\sigma t-\frac{\sigma^2t^2}{2}}.$$

6. Якщо $a=1,\sigma=2$, то випадкова величина ξ має рівномірний розподіл. Щільність цієї випадкової величини дорівнює:

$$p_\xi(x)=5{,}1299\cdot e^{\frac{-(x-1)^2}{8}},\ x\le 5,$$

а її функція розподілу має вигляд:

$$F_\xi(x)=1{,}0232\cdot\left(\Phi_0\left(\frac{x-1}{2}\right)+\frac{1}{2}\right).$$

Графіки щільності розподілу $p_\xi(x)$ та функції розподілу наведені на рисунках 3 та 4 відповідно.

Обчислимо значення основних числових характеристик цього розподілу відповідно до загальних співвідношень, отриманих у п.5:

- математичне сподівання $\mathrm{M}_\xi=1-\frac{2}{0{,}97725}p(\frac{5-1}{2})=0{,}8895;$
- дисперсія $\mathrm{D}_\xi=2^2-(5-0{,}8895)(1-0{,}8895)=3{,}5457;$
- середньоквадратичне відхилення $\sigma_\xi=\sqrt{3{,}5457}=1{,}88303;$

- коефіцієнт варіації $V=\dfrac{\mathrm{M}_\xi}{\sigma_\xi}=\dfrac{0{,}8895}{1{,}88303}=0{,}4723;$

-асиметрія $A_\xi=\dfrac{\mu_3}{\sigma^3}=\dfrac{(1-0{,}8895)\cdot 3{,}5457+(0{,}8895-1)(5-0{,}8895)^2}{1{,}88303^3}=-0{,}2209;$

- ексцес $E_\xi=\dfrac{\mu_4}{\sigma^4}-3=\dfrac{(1-0{,}8895)\cdot(-1{,}4753)+3\cdot 4\cdot 3{,}5457+(0{,}8895-1)(5-0{,}8895)^3}{1{,}88303^4}-3=$

$$=-0{,}2391;$$

- мода $M_0=\begin{cases}1, & 5\ge 1;\\ 5, & 5<1;\end{cases}$

- *p*-квантиль $\dfrac{1}{\gamma}\left(\Phi_0\left(\dfrac{x_p-a}{\sigma}\right)+\dfrac{1}{2}\right)=p;$

- нижня квартиль $x_{\frac{1}{4}}=3{,}55;$

- медіана $x_{\frac{1}{2}}=4{,}3;$

- верхня квартиль $x_{\frac{3}{4}}=4{,}7.$

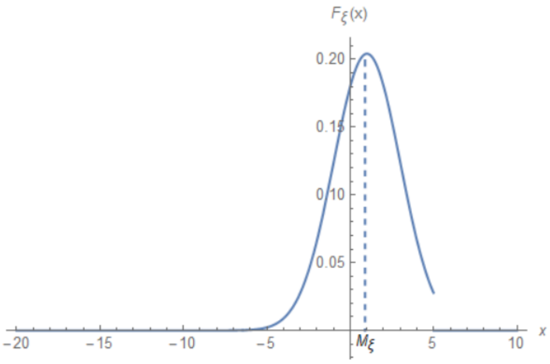


Рисунок2.3 Графік щільності $p_\xi(x)$

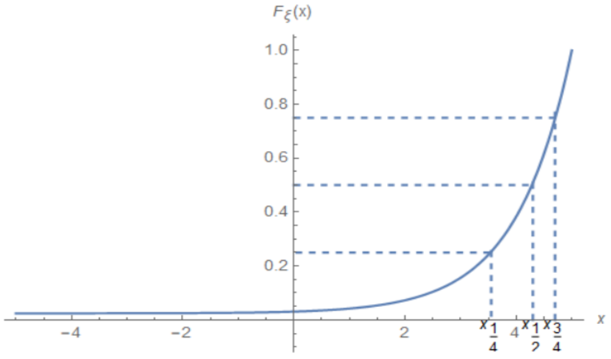


Рисунок2.4 Графік функції розподілу $F_\xi(x)$

7. а) Визначимо щільність розподілу випадкової величини $\eta=3\xi+1$, де ξ має рівномірний розподіл зі щільністю:

$$p_{\xi}(x)=\frac{1}{\gamma\sigma\sqrt{2\pi}}e^{-\frac{(x-\alpha)^2}{2\sigma^2}},\ x\le\beta,\ \sigma>0,$$

Множиною можливих значень величини ξ є проміжок $\left(-\infty;\beta\right)$. Функція $f(\text{x})=3x+1$ визначена, монотонна і неперервно диференційовна на проміжку $x\in\left(-\infty;\beta\right)$, отже, множиною її можливих значень буде область $\left(-\infty;3\beta+1\right)$ при $x\in\left(-\infty;\beta\right)$(рис. 2.5).

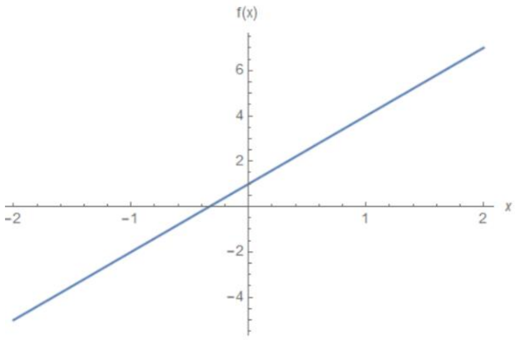


Рисунок 2.5 Графік функції $f(x)$

Обернена до неї функція має вигляд $x=f^{-1}(y)=\frac{y-1}{3}$, визначена на проміжку $y\in\left(-\infty;3\beta+1\right)$, а її похідна дорівнює $[f^{-1}(y)]'=\frac{1}{3}$.

Щільність випадкової величини η дорівнюватиме:

$$p_{\eta}(y)=p_{\xi}[f^{-1}(y)]\left|[f^{-1}(y)]'\right|=\frac{1}{3\gamma\sigma\sqrt{2\pi}}e^{-\frac{\left(\frac{y-1}{3}-a\right)^2}{2\sigma^2}},y\in\left(-\infty;3\beta+1\right).$$

Безпосередньою перевіркою переконуємось, що щільність $p_{\eta}(y)$ задовольняє умові нормування:

$$\int_{-\infty}^{\infty}p_{\eta}(y)dy=\int_{-\infty}^{3\beta+1}\frac{1}{3\gamma\sigma\sqrt{2\pi}}e^{-\frac{\left(\frac{y-1}{3}-a\right)^2}{2\sigma^2}}dy=\begin{vmatrix}\frac{\frac{y-1}{3}-a}{\sigma}=z & y=3a+3z\sigma+1\\ dy=3\sigma dz & \frac{\beta-a}{\sigma}<z<-\infty\end{vmatrix}=$$

$$=\frac{1}{3\gamma\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{z^2}{2}}3\sigma dz=\frac{1}{\gamma}\left(\Phi_0\left(\frac{\beta-a}{\sigma}\right)+\frac{1}{2}\right)=1.$$

б) Визначимо щільність розподілу випадкової величини $\eta = \sqrt{|x|}$. Множиною можливих значень випадкової величини ξ є область $[0;+\infty]$. Як видно з рис. 6, функція $f(x) = \sqrt{|x|}$ не є монотонною функцією. Монотонні гілки цієї функції належать проміжкам $x \in [-\infty; 0]$ та $x \in [0; -\infty]$.

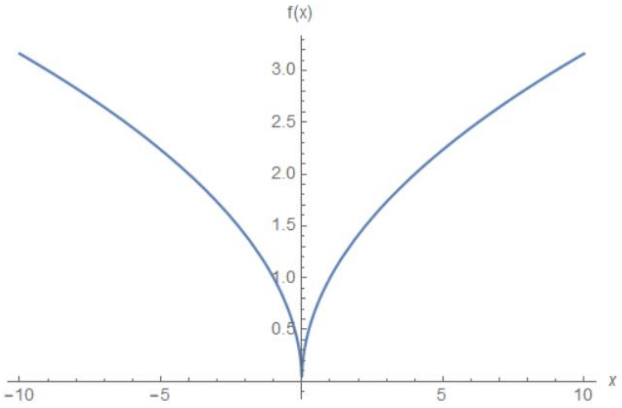


Рисунок 2.6 Графік функції $f(x)$

Обернена функція на цих проміжках дорівнює відповідно $x = f_1^{-1}(y) = -y^2$ Та $x = f_2^{-1}(y) = y^2$, а її похідні: $[f_1^{-1}(y)]' = -2y$ і $[f_2^{-1}(y)]' = 2y$.Тоді щільність випадкової величини η дорівнюватиме:

$$p_\eta(y) = \frac{2y}{\sqrt{2\pi}\sigma}\left(e^{\frac{-(-y^2-a)}{2\sigma^2}} + e^{\frac{-(y^2-a)}{2\sigma^2}} \right).$$

Оскільки $y \in [0; \infty]$, то остаточно отримаємо

$$p_\eta(y) = \begin{cases} \frac{2y}{\sqrt{2\pi}\sigma}\left(e^{\frac{-(-y^2-a)}{2\sigma^2}} + e^{\frac{-(y^2-a)}{2\sigma^2}} \right), & y \in [0; \infty], \\ 0, \quad y \notin [0; \infty]. \end{cases}$$

Безпосередньою перевіркою переконуємось, що щільність $p_\eta(y)$ задовольняє умові нормування:

$$\int_{-\infty}^{\infty} p_\eta(y)dy = \int_0^{\infty} \frac{2y}{\sqrt{2\pi\sigma}}(\mathrm{e}^{\frac{-(-y^2-a)}{2\sigma^2}} + \mathrm{e}^{\frac{-(y^2-a)}{2\sigma^2}})dy = \int_0^{\infty} \frac{2y}{\sqrt{2\pi\sigma}} \cdot \mathrm{e}^{\frac{-(-y^2-a)}{2\sigma^2}} +$$

$$+\int_0^{\infty} \frac{2y}{\sqrt{2\pi\sigma}} \cdot \mathrm{e}^{\frac{-(y^2-a)}{2\sigma^2}} = \frac{1}{\sqrt{2\pi\sigma}} \int_0^{\infty} 2y \cdot \mathrm{e}^{\frac{-(-y^2-a)}{2\sigma^2}} + \frac{1}{\sqrt{2\pi\sigma}} \int_0^{\infty} 2y \cdot \mathrm{e}^{\frac{-(y^2-a)}{2\sigma^2}} =$$

$$= \frac{1}{\sqrt{2\pi\sigma}} \int_{-\frac{a}{\sigma}}^{\infty} \mathrm{e}^{\frac{-t^2}{2}} dt + \frac{1}{\sqrt{2\pi\sigma}} \int_{\frac{a}{\sigma}}^{\infty} \mathrm{e}^{\frac{-t^2}{2}} dt = 1.$$

8. Нормальний розподіл часто зустрічається в природі. Наприклад, такі випадкові величини добре моделюються нормальним розподілом:

- відхилення при стрільбі;
- похибки вимірювань (проте похибки деяких вимірювальних приладів мають інший розподіл);
- деякі властивості живих організмів у популяції.

Багатовимірний нормальний розподіл використовується при дослідженні багатовимірних випадкових величин (випадкових векторів). Одним з численних прикладів таких застосувань є дослідження властивостей особистості людини в психології і психіатрії.

2.2 Задача на тему «Граничні теореми теорії ймовірностей»

Задача 1. Інтеграл $I = \int_1^5 \ln x dx$ обчислений методом Монте-Карло.

а) Знайти ймовірність того, що абсолютна похибка у визначенні величини I не перевищує 0,005, якщо проведено 1300 незалежних дослідів.

б) Скільки дослідів потрібно зробити, щоб з імовірністю більшої 0,9, можна було вважати абсолютну похибку обчислення значення інтеграла що не перевищує 0,25% від I?

Розв’язання.

а) Інтеграл $I = \int_1^5 \ln x dx$ можна записати у вигляді

$$I = \int_1^5 \ln x dx = \frac{1}{4}\int_1^5 4 \ln x dx.$$

Розглянемо випадкову величину $\eta = 4\xi$, де ξ – випадкова величина, рівномірно розподілена на [1; 5]. Щільність ξ має вигляд:

$$p_\xi(x) = \begin{cases} \frac{1}{4}, x \in [1; 5]; \\ 0, x \notin [1; 5]. \end{cases}$$

Відомо, що математичне сподівання випадкової величини $\eta = (b-a) f(\xi)$, де $p_\xi(x)$ – щільність відповідної випадкової величини ξ, дорівнює:

$$\mathrm{M}\eta = \int_{-\infty}^{+\infty} (b-a) f(x) p_\xi x dx.$$

Тому при $f(\xi)=\xi$ отримаємо:

$$\mathrm{M}\eta=\frac{1}{4}\int_1^5 4\ln x dx,$$

тобто $I=\mathrm{M}\eta.$

$$I_n=\frac{1}{n}\sum_{k=1}^{n}4\xi_k,$$

де ξ_k – незалежні випадкові числа з відрізка [1;5].

Обчислимо математичне сподівання і дисперсію випадкової величини I_n. Оскільки

$$M\left[f(\xi_k)\right]=M\xi_k=\int_{-\infty}^{+\infty}f(x)p_{\xi_k}(x)dx=\frac{1}{4}\int_1^5\ln xdx\approx 1{,}0118,$$

$$D\left[f(\xi_k)\right]=D\xi_k=\int_{-\infty}^{+\infty}f^2(x)p_{\xi_k}(x)dx-(M\xi_k)^2=\frac{1}{4}\int_1^5\ln^2 xdx-(1{,}0118)^2\approx 0{,}190534,$$

то

$$\mathrm{M}I_n=\frac{1}{n}\sum_{k=1}^{n}4\mathrm{M}\xi_k=\frac{4*1{,}0118}{1300}\approx 0{,}00311322,$$

$$DI_n=\frac{1}{n^2}\sum_{k=1}^{n}16D\xi_k=\frac{16*0{,}190534}{1300}\approx 0{,}00234504.$$

$$P(|I_n-I|<\varepsilon)=p\left(\frac{|I_n-I|}{\sqrt{DI_n}}<\frac{\varepsilon}{\sqrt{DI_n}}\right)\approx 2\Phi_0\left(\frac{\varepsilon}{\sqrt{DI_n}}\right).$$

При $n=1300$ і $\varepsilon=0,005$ отримуємо

$$P(|I_{1300}-I||<0,005)\approx 2\Phi_0\left(\frac{0,005}{\sqrt{0,00234504}}\right)\approx 0,0822364.$$

Таким чином, ймовірність того, що при проведенні 1300 незалежних дослідів абсолютна похибка у визначенні величини $I=\int\limits_1^5 \ln xdx$ не перевищує 0,005, складе 0,822364.

б) Інтеграл можна розглядати як математичне сподівання випадкової величини $\eta=4\xi$, ξ - випадкова величина, рівномірний розподілена на [1;5]. Тоді:

$$p_\xi(x)=\begin{cases}\frac{1}{4}, x\in[1;5];\\ 0, x\notin[1;5],\end{cases}$$

$$\mathrm{M}\eta=\frac{1}{4}\int\limits_1^5 4\ln xdx=I.$$

Наближене значення інтеграла I, обчислене методом Монте-Карло, дорівнює

$$I_n=\frac{1}{n}\sum_{k=1}^{n}4\xi_k,$$

де ξ_k – незалежні випадкові числа з відрізка [1;5]. При цьому

$$M\left[\ln\xi_k\right] = M\xi_k = \int\limits_{-\infty}^{+\infty} f(x)p_{\xi_k}(x)dx = \frac{1}{4}\int\limits_1^5 \ln x dx = \frac{1}{4}(5\ln(5)-4),$$

$$D\left[\ln\xi_k\right] = D\xi_k = \int\limits_{-\infty}^{+\infty} f^2(x)p_{\xi_k}(x)dx - (M\xi_k)^2 =$$

$$= \frac{1}{4}\int\limits_1^5 \ln^2 x dx - (\frac{1}{4}(5\ln(5)-4))^2 = \frac{1}{16}(16-5\ln^2(5)),$$

тоді

$$\mathrm{M}I_n = \frac{1}{n}\sum_{k=1}^{n} 4\mathrm{M}[\ln\xi_k] = 5\ln(5)-4,$$

$$DI_n = \frac{1}{n^2}\sum_{k=1}^{n} 16D\xi_k = \frac{16-5\ln^2(5)}{n},$$

$$P\left(\left|I_n - I\right| < \varepsilon\right) = P\left(\frac{\left|I_n - I\right|}{\sqrt{DI_n}} < \frac{\varepsilon}{\sqrt{DI_n}}\right) \approx 2\varPhi_0(\frac{\varepsilon}{\sqrt{DI_n}}).$$

Якщо з ймовірністю більше 0,9 абсолютна похибка обчислення інтеграла I не перевищує 0,25% від $I = \mathrm{M}I_n$, то $\varepsilon = 0,0025I = 0,0025(5\ln(5)-4)$ і

$$P\left(\left|I_n - I\right| < \varepsilon\right) \approx 2\varPhi_0\left(\frac{0,0025*(5\ln(5)-4)}{\sqrt{\frac{16-5\ln^2(5)}{n}}}\right) = 2\varPhi_0(0,00579491\sqrt{n}) > 0,9.$$

Звідси

$$0,00579491\sqrt{n} > \varPhi_0^{-1}(0,45) = 1,64, \text{ тобто } n > 0,08*10^6.$$

Це означає, що при кількості дослідів не менше $0,08*10^6$ ймовірність того, що абсолютна похибка відхилення значення I від I_n не перевищить 0,25% від I, буде більше 0,9.

2.3 Задача на перевірку статистичних гіпотез критерієм незалежності χ^2

Задача 1. Досліджується наявність бронхіту у працівників цеху в залежності від звички курити. За результатами досліджень складено таблицю спряженості ознак

Ознака	Наявність бронхіту	Відсутність бронхіту
Не курить	6	22
Кинув курити	9	38
Курить	16	9

Критерієм незалежності χ^2 потрібно на рівнях значущості $\alpha = 0{,}1;\ 0{,}05;\ 0{,}01$ перевірити гіпотезу про те, чи є залежними наявність бронхіту (випадкова величина ξ) і звичка палити (випадкова величина η).

Розв’язання.

В таблиці 2.1 наведено суму опитуваних.

Таблиця 2.1

Ознака	Наявність бронхіту	Відсутність бронхіту	$\sum$
Не курить	6	22	28
Кинув курити	9	38	47
Курить	16	9	75
$\sum$	31	69	100

Обчислимо спострежежуване значення статистики:

$$\hat{\chi}^2(\vec{x}_n, \vec{y}_n) = n\sum_{i=1}^{n}\sum_{j=1}^{s}\frac{(n_{ij} - n_{i\bullet} \cdot n_{\bullet j})^2}{n_{i\bullet} \cdot n_{\bullet j}} = 17,013.$$

Обчислимо число ступенів свободи:

$$(S-1)(k-1) = 2$$

На рівні значущості $\alpha = 0,1$:

$$\hat{\chi}_n^2 > \hat{\chi}_{0,9,2}^2 = 4,6052,$$

тоді гіпотеза H_0 відкидається, тобто наявність бронхіту і звичка палити є залежними.

На рівні значущості $\alpha = 0,05$:

$$\hat{\chi}_n^2 > \hat{\chi}_{0,95,2}^2 = 5,9915,$$

тоді гіпотеза H_0 відкидається, тобто наявність бронхіту і звичка палити є залежними.

На рівні значущості $\alpha = 0,01$:

$$\hat{\chi}_n^2 > \hat{\chi}_{0,99,2}^2 = 9,2103,$$

тоді гіпотеза H_0 відкидається, тобто наявність бронхіту і звичка палити є залежними.

ВИСНОВКИ

Теорія ймовірностей і математична статистика – це наука, що займається вивченням закономірностей масових випадкових явищ, тобто статистичних закономірностей.

У ході виконання курсової роботи я зміг не тільки вивчити теорію з розглянутих тем, але й використати свої навички на практиці. Ми розглянули основні теореми та закони граничних теорем та розібрали важливі критерії. Після цього ми виконали практичну частину для закріплення отриманих знань з таких тем,як «Перевірка статистичних гіпотез» та «Граничні теореми теорії ймовірності».

На мою думку ця дисципліна заслужено стоїть на такому високому та важливому місці серед інших дисциплін на спеціальності «Прикладна математик

ПЕРЕЛІК ПОСИЛАНЬ

1. Ивченко Г.И., Медведев Ю.И. Введение в математическую статистику. – М.: Изд-во ЛКИ, 2010. – 600 с.

2. Ивченко Г.И., Медведев Ю.И. Математическая статистика. – М.: Высш. шк., 1984. – 248 с.

3. Кобзарь А.И. Прикладная математическая статистика. Для инженеров и научных работников. – М.: ФИЗМАТЛИТ, 2006. – 816 с.

4. Математическая статистика / В.Б. Горяинов, И.В. Павлов, Г.М. Цветкова и др.; Под ред. В.С. Зарубина, А.П. Крищенко. – Изд. 3-е, испр. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2008. – 424 с.

5. Боровков А. А. Теория вероятностей. – М.: Наука, 1986. – 432 с.

6. Гихман И. И., Скороход А. В., Ядренко М. И. Теория вероятностей и математическая статистика. – К.: Выща школа, 1979. – 440 с.

7. Гнеденко Б. В. Курс теории вероятностей. – М.: Изд–во ЛКИ, 2007. – 448с.

8. Розанов Ю. А. Теория вероятностей, случайные процессы и математическая статистика. – М.: Наука, 1989. – 320 с.

9. Ширяев А. Н. Вероятность. – 2–е изд., перераб. и доп. – М.: Наука, 1989. – 640 с.

10. Жлуктенко В. І., Наконечний С. І. Теорія ймовірностей і математична статистика. У 2 ч. – К.: КНЕУ, 2000. – 304 с.

11. Вадзинский Р. Н. Справочник по вероятностным распределениям. – Спб.: Наука, 2001. – 295 с.

12. Хастингс Н., Пикок Дж. Справочник по статистическим распределениям. – М.: Статистика, 1980. — 95 с.

Додаток А

Наведемо деякі інтеграли, що були використані при розв’язанні задачі 2.1

$$\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\beta} e^{-\frac{t^2}{2}}\,dt=\Phi_0(\beta)-\Phi_0(-\infty)=\Phi_0(\beta)+\frac{1}{2};$$

$$\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\beta} t\cdot e^{-\frac{t^2}{2}}\,dt=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\beta} e^{-\frac{t^2}{2}}\,d\left(\frac{t^2}{2}\right)=-\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\Bigg|_{-\infty}^{\beta}=-\frac{1}{\sqrt{2\pi}}e^{-\frac{\beta^2}{2}};$$

$$\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\beta} t^2\cdot e^{-\frac{t^2}{2}}\,dt=\left|\begin{matrix} u=t & du=dt \\ dv=t\,e^{-\frac{t^2}{2}}dt & v=-e^{-\frac{t^2}{2}}\end{matrix}\right|=-t\,e^{-\frac{t^2}{2}}\Bigg|_{-\infty}^{\beta}+\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\beta} e^{-\frac{t^2}{2}}dt=$$

$$=-\frac{1}{\sqrt{2\pi}}\,\beta\,e^{-\frac{\beta^2}{2}}+\Phi_0(\beta)+\frac{1}{2};$$

$$\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\beta} t^3\cdot e^{-\frac{t^2}{2}}\,dt=\left|\begin{matrix} z=t^2 \\ dz=2tdt\end{matrix}\right|=\frac{1}{\sqrt{2\pi}}\cdot\frac{1}{2}\int\limits_{-\infty}^{\beta} z\,e^{-\frac{t^2}{2}}dz=$$

$$=\left|\begin{matrix} u=z & du=dz \\ dv=e^{-\frac{z}{2}} & v=-2e^{-\frac{z}{2}}\end{matrix}\right|=-z\,e^{-\frac{z}{2}}+\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\beta} e^{-\frac{z}{2}}\,dz=$$

$$=\frac{1}{\sqrt{2\pi}}-z\,e^{-\frac{z}{2}}-2e^{-\frac{z}{2}}=-\frac{1}{\sqrt{2\pi}}e^{-\frac{\beta^2}{2}}(2+\beta^2);$$

$$a_2=\int\limits_{-\infty}^{\beta} x^2\,\frac{1}{\gamma\sigma\sqrt{2\pi}}e^{-\frac{(x-\alpha)^2}{2\sigma^2}}\,dx=\left|\begin{matrix} z=\dfrac{x-a}{\sigma} \\ \dfrac{\beta-a}{\sigma}<z<-\infty\end{matrix}\right|=\frac{1}{\gamma\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{\frac{\beta-a}{\sigma}}(\sigma z+a)^2\,e^{-\frac{z^2}{2}}\,\sigma dz=$$

$$=\frac{\sigma^2}{\gamma\sqrt{2\pi}}\int\limits_{-\infty}^{\frac{\beta-a}{\sigma}} z^2\,e^{-\frac{z^2}{2}}\,dz+\frac{2\sigma a}{\gamma\sqrt{2\pi}}\int\limits_{-\infty}^{\frac{\beta-a}{\sigma}} z\,e^{-\frac{z^2}{2}}\,dz+\frac{a^2}{\gamma\sqrt{2\pi}}\int\limits_{-\infty}^{\frac{\beta-a}{\sigma}} e^{-\frac{z^2}{2}}\,dz=$$

$$=-\frac{\sigma^2}{\gamma}\,p\left(\frac{\beta-a}{\sigma}\right)\left(\frac{\beta-a}{\sigma}\right)+\gamma+\frac{2\sigma a}{\gamma}\,p\left(\frac{\beta-a}{\sigma}\right)+a^2;$$

$$a_3 = \int\limits_{-\infty}^{\beta} x^3 \frac{1}{\gamma\sigma\sqrt{2\pi}} e^{-\frac{(x-\alpha)^2}{2\sigma^2}} dx = \left| \begin{array}{l} z = \dfrac{x-a}{\sigma} \\ \dfrac{\beta-a}{\sigma} < z < -\infty \end{array} \right| = \frac{1}{\gamma\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{\frac{\beta-a}{\sigma}} (\sigma z + a)^3 e^{-\frac{z^2}{2}} \sigma dz =$$

$$= \frac{\sigma^3}{\gamma\sqrt{2\pi}} \int\limits_{-\infty}^{\frac{\beta-a}{\sigma}} z^3 e^{-\frac{z^2}{2}} dz + \frac{3\sigma^2 a}{\gamma\sqrt{2\pi}} \int\limits_{-\infty}^{\frac{\beta-a}{\sigma}} z^2 e^{-\frac{z^2}{2}} dz + \frac{3\sigma a^2}{\gamma\sqrt{2\pi}} \int\limits_{-\infty}^{\frac{\beta-a}{\sigma}} z\, e^{-\frac{z^2}{2}} dz + \frac{a^3}{\gamma\sqrt{2\pi}} \int\limits_{-\infty}^{\frac{\beta-a}{\sigma}} e^{-\frac{z^2}{2}} dz =$$

$$= -\frac{\sigma^3}{\gamma} p\left(\frac{\beta-a}{\sigma}\right)\left(2 + \frac{(\beta-a)^2}{\sigma^2}\right) - \frac{3\sigma a^2}{\gamma} p\left(\frac{\beta-a}{\sigma}\right) + a^3 - \frac{3\sigma a}{\gamma}\left(p\left(\frac{\beta-a}{\sigma}\right) + \gamma\right) =$$

$$= a\left(a\left(a - \frac{\sigma}{\gamma} p\left(\left(\frac{\beta-a}{\sigma}\right)\right)\right) + \sigma^2 + \left(a - \frac{\sigma}{\gamma} p\left(\left(\frac{\beta-a}{\sigma}\right)\right) - a \right)\beta \right) +$$

$$+2\sigma^2\left(a - \frac{\sigma}{\gamma} p\left(\left(\frac{\beta-a}{\sigma}\right)\right)\right) + \left(\left(a - \frac{\sigma}{\gamma} p\left(\left(\frac{\beta-a}{\sigma}\right)\right)\right) - a\right)\beta^2;$$

$$a_4 = \int\limits_{-\infty}^{\beta} x^4 \frac{1}{\gamma\sigma\sqrt{2\pi}} e^{-\frac{(x-\alpha)^2}{2\sigma^2}} dx = \left| \begin{array}{l} z = \dfrac{x-a}{\sigma} \\ \dfrac{\beta-a}{\sigma} < z < -\infty \end{array} \right| =$$

$$= \frac{1}{\gamma\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{\frac{\beta-a}{\sigma}} (\sigma z + a)^4 e^{-\frac{z^2}{2}} \sigma dz = \frac{\sigma^4}{\gamma\sqrt{2\pi}} \int\limits_{-\infty}^{\frac{\beta-a}{\sigma}} z^4 e^{-\frac{z^2}{2}} dz +$$

$$+ \frac{4\sigma^3 a}{\gamma\sqrt{2\pi}} \int\limits_{-\infty}^{\frac{\beta-a}{\sigma}} z^3 e^{-\frac{z^2}{2}} dz + \frac{6\sigma^2 a^2}{\gamma\sqrt{2\pi}} \int\limits_{-\infty}^{\frac{\beta-a}{\sigma}} z^2\, e^{-\frac{z^2}{2}} dz + \frac{4\sigma a^3}{\gamma\sqrt{2\pi}} \int\limits_{-\infty}^{\frac{\beta-a}{\sigma}} z e^{-\frac{z^2}{2}} dz +$$

$$+ \frac{a^4}{\gamma\sqrt{2\pi}} \int\limits_{-\infty}^{\frac{\beta-a}{\sigma}} e^{-\frac{z^2}{2}} dz = \frac{\sigma^4}{\gamma} p\left(\frac{\beta-a}{\sigma}\right)\left(2 + \frac{(\beta-a)^2}{\sigma^2}\right) + \frac{4\sigma^3 a}{\gamma} p\left(\frac{\beta-a}{\sigma}\right) -$$

$$- \frac{6\sigma^2 a^2}{\gamma}\left(p\left(\frac{\beta-a}{\sigma}\right)\left(\frac{\beta-a}{\sigma}\right) + \gamma \right) - \frac{4\sigma a^3}{\gamma}\left(p\left(\frac{\beta-a}{\sigma}\right)\right) + a^4 =$$

$$= aa_3 + 3\sigma^2 a_2 + \left(a - \frac{\sigma}{\gamma} p\left(\left(\frac{\beta-a}{\sigma}\right)\right) - a \right)\beta^3.$$