МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Центральноукраїнський державний університет
імені Володимира Винниченка
Факультет математики, природничих наук та технологій
Кафедра математики та цифрових технологій

Кваліфікаційна робота
на правах рукопису

Сотник Катерина Олегівна

КВАЛІФІКАЦІЙНА РОБОТА
другий (магістерський) рівень вищої освіти

на тему: «**Дослідження та моделювання ризику банкрутства страхової компанії**»

Виконала: студентка II курсу, групи СТ23М
факультету математики, природничих наук
та технологій
спеціальності 112 Статистика
освітня програма Статистика (Фінансова,
страхова та комп'ютерна статистика)
форма навчання денна
керівник: Халецька Зоя Петрівна, кандидат
фізико-математичних наук; доцент, доцент
кафедри математики та цифрових технологій
рецензент: Гаєвський Микола Вікторович,
кандидат фізико-математичних наук,
консультант відділу розширеної аналітики
(АА) компанії RBC GROUP

Кваліфікаційна робота містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

(підпис, ініціали та прізвище здобувача вищої освіти)

Кропивницький – 2024

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Центральноукраїнський державний університет
імені Володимира Винниченка

Кафедра математики та цифрових технологій

До захисту допустити
Зав. кафедри_/Трифонова О.М./
«_»_2024р.

Сотник Катерина Олегівна

КВАЛІФІКАЦІЙНА РОБОТА
другий (магістерський) рівень вищої освіти
на тему: «**Дослідження та моделювання ризику банкрутства страхової компанії»**

Виконала: студентка II курсу, групи СТ23М
факультету математики, природничих наук та технологій
спеціальності 112 Статистика
освітня програма Статистика (Фінансова, страхова та комп'ютерна статистика)
форма навчання денна
науковий керівник:
Халецька Зоя Петрівна, кандидат фізико-математичних наук; доцент, доцент кафедри математики та цифрових технологій

Кваліфікаційна робота захищена з оцінкою «_» балів,
за шкалою ЄКТС_,
за національною шкалою .

Секретар ЕК / /
«_»_20_р.

Кропивницький – 2024

ЗМІСТ

АНОТАЦІЯ

Сотник К.О. Дослідження та моделювання ризику банкрутства страхової компанії. – Кваліфікаційна робота на правах рукопису. Кваліфікаційна робота на здобуття освітнього ступеня магістра за спеціальністю 112 Статистика. – Центральноукраїнський державний університет імені Володимира Винниченка, Кропивницький, 2024. Дипломна робота присвячена методам дослідження та моделювання ризику банкрутства у страхуванні. У роботі розглянуто основні теоретичні аспекти моделей страхових ризиків, зокрема індивідуальних і колективних моделей, а також методів їх аналізу. Особливу увагу приділено їх практичному застосуванню та створенню методичних рекомендацій щодо вивчення тематики дослідження. На основі використання статистичного програмного забезпечення SPSS, здійснено аналіз статистичних даних щодо страховий випадків смертельних ДТП у штатах США.

Отримані результати можуть бути використані при викладанні актуарної математики, зокрема в темах, пов'язаних із ризиками та фінансовою стійкістю страхових компаній.

Ключові слова: ризик банкрутства, страхова компанія, індивідуальний позов, модель індивідуального ризику, регресія, кластерний аналіз. .

SUMMARY

Sotnyk K.O. Research and modeling of the risk of bankruptcy of an insurance company. – Qualification work in the form of a manuscript.

Qualification work for the degree of master in the specialty 112 Statistics. – Volodymyr Vynnychenko Central Ukrainian State University, Kropyvnytskyi, 2024.

The diploma work is devoted to methods of research and modeling of the risk of bankruptcy in insurance. The work considers the main theoretical aspects of insurance risk models, in particular individual and collective models, as well as methods of their analysis. Particular attention is paid to their practical application and the creation of methodological recommendations for studying the research topic. Based on the use of statistical software SPSS, an analysis of statistical data on insured fatal accidents in the US states was carried out.

The results obtained can be used in teaching actuarial mathematics, in particular in topics related to risks and financial stability of insurance companies.

Keywords: bankruptcy risk, insurance company, individual claim, individual risk model, regression, cluster analysis.

ВСТУП

Актуальність роботи. Актуарна математика пройшла великий шлях розвитку і зараз є одним з найбільш перспективних та актуальних напрямків застосування теорії ймовірностей, теорії випадкових процесів та математичної статистики.

Актуарна математика безпосередньо пов’язана з діяльністю страхової компанії. Взаємодія страхових компаній та їх клієнтів здійснюється шляхом продажу полісів – договорів, в яких вказані умови страхування. При укладанні страхового договору виникає питання про ціну поліса. Ця ціна повинна бути прийнятною і для клієнта, і для страхової компанії, і повинна враховувати такі фактори, як ймовірність настання страхового випадку, величину збитків, що виникають при цьому, тощо. Таким чином, при продажу поліса, страхова компанія повинна провести розрахунки по фінансовим зобов’язанням клієнта, беручи до уваги, що вона сама виконає свої зобов’язання перед ним в майбутньому, до того ж у будь-який випадковий момент.

Договори страхування укладаються для того, щоб позбутися від фінансових збитків, пов’язаних з невизначеністю настання тих чи інших випадкових подій. Клієнт йде на певні витрати для того, щоб позбутися випадкових витрат, які хоча є малоймовірними, проте можуть бути дуже великими для нього. Однак сам ризик не зник - його прийняла на себе страхова компанія. Тому фінансовий ризик і пов’язана з ним небезпека розорення об’єктивно існують в діяльності кожної страхової компанії. Оцінка цього ризику представляє фундаментальний інтерес для компанії і служать основою для прийняття найважливіших рішень.

Проблема забезпечення фінансової стійкості страхової компанії є комплексною; її вивчення і вирішення передбачає зусилля спеціалістів в різних галузях. Однак більшість важливих задач носить чисто математичний характер. В рамках спеціальної математичної теорії – теорії ризиків,

розроблена система понять, моделей та методів, що дозволяють кількісно оцінювати фінансові ризики в діяльності страхової компанії. Беручи до уваги присутність факторів випадковості, загальною математичною базою для теорії ризиків служать теорія ймовірностей і математична статистика.

Об'єкт – моделі страхової математики.

Предмет – моделювання страхових процесів.

Мета роботи полягає у вивченні та застосуванні методів страхової математики та статистичного аналізу у моделюванні страхових процесів.
Відповідно до мети визначено такі ***завдання***:

1. Систематизувати та узагальнити теоретичний матеріал з даної теми.
2. З'ясувати та систематизувати основні властивості дискретної та неперервної моделі індивідуальних позовів.
3. Розглянути систему понять та методів моделювання індивідуального та колективного ризиків.
4. Висвітлити аспекти застосувань логістичної регресії та основних методів статистичного аналізу до тематики дослідження.
5. Застосувати основні ймовірнісні методи оцінки ризику та методи статистичного аналізу до розв'язання практичних задач.
6. Дослідити фактори, які спричиняють ризики смертельних зіткнень у штатах США та відповідні страхові збитки, а також здійснити кластерний аналіз групування штатів США за характерними показниками аварійності.
7. .Створити методичну розробку до практичного модуля з курсу «Актуарна математика».

Методи дослідження: математико-статистичні методи обробки інформації, ймовірнісні методи страхової математики, кореляційний аналіз, кластерний аналіз, системний підхід.

Наукова новизна. Розробка методичних рекомендацій до вивчення методів і моделей страхового ризику.

Апрбація результатів. Основні результати дослідження висвітлені у статті «Математичне моделювання страхових ризиків: оцінка ймовірності банкрутства» (Наукові записки молодих учених. №13 (2024): – Кропивницький: РВВ ЦДУ ім В. Винниченка ISSN 2617-2666; [27]) та тезах «Застосування логістичної регресії до оцінювання ризику банкрутства страхової компанії» (Збірник матеіалів XVI Всеукраїнської студентської наукової конференції «Перспективи розвитку точних наук, економіки та методики їх викладання» (2024); [28])

Практична значущість роботи. Результати і методичні рекомендації роботи будуть корисними для студентів та викладачів при вивченні методів страхової математики.

Структура дипломної роботи. Дипломна робота складається зі вступу, трьох розділів, висновків, списку використаних джерел та додатків. У вступі визначено актуальність, мету, об'єкт, предмет та завдання дослідження.

В першому розділі висвітлені основні теоретичні відомості про модель індивідуальних позовів, моделі індивідуальних та колективних ризиків, логістичну регресію та основні методи статистичного аналізу.

Другий розділ присвячений практичному застосуванню як ймовірнісних методів страхової математики, так і методів статистичного аналізу.

Третій розділ являє собою методичну розробку до практичного модуля курсу «Актуарна математика», яка складається з наступних тем: 1. Застосування методів прикладної статистики в страховій математиці.

2. Моделі індивідуальних позовів.

3. Модель індивідуального ризику.

4. Модель колективного ризику.

РОЗДІЛ 1

ОСНОВНІ ТЕОРЕТИЧНІ ВІДОМОСТІ

1.1. Моделі індивідуальних позовів

Елементарною складовою частиною фінансового ризику страхової компанії є індивідуальний позов клієнта до страхової компанії, тобто сума, необхідна для відшкодування втрат, спричинених страховим випадком. *Дискретні моделі індивідуальних позовів* [7].

Як математичну модель індивідуального позову за даним договором розглядатимемо випадкову величину *X*. Індивідуальний позов *X* з додатною ймовірністю набуває значення 0:

$$P\{X = 0\} = p_0 > 0 \quad (1.1.1)$$

Якщо випадкова величина *X* набуває скінченного або зліченного числа значень, то *X* називатимемо дискретною моделлю індивідуального позову. Розподілом дискретної випадкової величини *X* називатимемо функцію p_X:

$$x_k \to p_X(x_k),\ x_k \in B \subset R^1 \quad (1.1.2)$$

означену на множині *B* різних можливих значень випадкової величини *X*, що ставить у відповідність кожному можливому значенню $x_k \in B$ ймовірність

$$p_X(x_k) = P\{X = x_k\} \quad (1.1.3)$$

Розподіл дискретної випадкової величини *X* часто записують у вигляді таблиці, у верхньому рядку якої вказують різні можливі значення випадкової величини, а в нижньому – імовірності, з якими ці значення відбуваються:

x_k	x_0	x_1	$\cdots$	x_n
p_k	p_0	p_1	$\cdots$	p_n

За розподілом випадкової величини X завжди можна визначити її числові характеристики: математичне сподівання, дисперсію, середнє квадратичне відхилення, моменти різних порядків.

$$MX = \sum_{k=0}^{\infty} x_k p_k(x_k)$$

$$DX = MX^2 - (MX)^2 = \sum_{k=0}^{\infty} x_k^2 p_k - (x_k p_k)^2$$

У теорії ризику зручно подавати випадкову величину X у спеціальному вигляді. Нехай X – величина індивідуального позову за договором страхування:

$$I = \begin{cases} 1, & X > 0 \\ 0, & X = 0 \end{cases}$$

де I – індикатор події – «відбувся страховий випадок».

Визначимо випадкову величину Y як величину реально поданого позову, коли страховий випадок відбувся ($I=1$). Випадкові величини I та Y – незалежні.

$$P\{Y = y_k\} = P\{X = x_k | X > 0\}$$

Тоді X завжди можна подати у вигляді: $X = IY$.

Теорема 1. [6, с.12]

Якщо

$$P\{X = x_k\} = p_k, k = 0, 1, 2, ...$$

розподіл X, то розподілом Y є

$$P\{Y = y_k\} = \frac{p_k}{1-p_0}, k = 0, 1, 2, ... \quad (1.1.4)$$

Зауваження 1. З рівності

$$\sum_{k=1}^{\infty} P\{Y = y_k\} = 1$$

випливає, що *Y* з додатними ймовірностями набуває тільки значень

��$_1$, ��$_2$, …

Зауваження 2. Розподіл *Y* збігається з умовним розподілом *X* за умови {�� > 0}, тобто

��{�� = ��$_{��}$} = ��{�� = ��$_{��}$| �� > 0}, �� = 0, 1, 2, …

*Наслідок***.** Функція розподілу ��$_{��}$(��) дійсно поданого позову *Y* дорівнює умовній функції розподілу позову *X* відносно події {�� > 0}, тобто

��$_{��}$(��) = ��{�� < ��} = ��{�� < ��| �� > 0}, �� > 0,

��$_{��}$(��) = 0, �� ≤ 0.

Теорема 2. [6, с.14] Нехай

(1 0
1 − ��$_0'$ ��$_{0'}$)

- розподіл випадкової величини *I*,

��{�� = ��$_{��}$} = ��$_{��}'$, �� = 0, 1, 2, …

- розподіл випадкової величини *Y*. Тоді

��{�� = 0} = ��$_0'$,

��{�� = ��$_{��}$} = (1 − ��$_0'$)��$_{��}'$, �� = 0, 1, 2, …

У деяких видах страхування один договір за час своєї дії може породити декілька позовів (наприклад, страхування автомобіля). У цьому випадку величину позову *X* за даним договором зручно подати у вигляді: �� = ��$_1$ + ��$_2$ + ⋯ + ��$_{��}$

де *v* – число позовів, породжених даним договором, ��$_{��}$ – величина дійсно поданого позову в *i*-ому страховому випадку за даним договором (*v*, ��$_{��}$ – незалежні). Тоді:

���� = ����$_1$���� (1.1.5)

���� = ����$_1$���� + (����$_1$)^2���� (1.1.6)

Приклади розподілів величини позову [7].

Біноміальний розподіл

Якщо випадкова величина ξ набуває значення *m* з імовірністю

$$P(\xi = m) = C_n^m p^m (1 - p)^{n-m},\ m=0,1,\dots,n,$$

то кажуть, що вона розподілена за біноміальним законом.

У страхуванні цей розподіл зустрічається, якщо:

а) є n незалежних договорів страхування;

б) імовірність позову однакова для всіх договорів і рівна p.

Тоді число позовів *N* має біноміальний розподіл з параметрами *n* і *p* (*N~B(n,p)*).

Розподіл Пуассона

Випадкова величина ξ має розподіл Пуассона з параметром λ>0, якщо множина значень цієї випадкової величини – множина невід'ємних цілих чисел і

$$P(\xi = k) = e^{-\lambda} \frac{\lambda^k}{k!},\ k=0,1,2,\dots \quad (1.1.7)$$

У страхуванні цей розподіл зустрічається, якщо:

а) під час коротких часових інтервалів Δ*t* може бути пред'явлене не більш одного вимоги про виплату;

б) імовірність того, що буде пред'явлена тільки одна вимога про виплату рівна *λ*Δ*t*;

в) непересічні часові інтервали незалежні.

Тоді число позовів за період часу (*t-n*) має розподіл Пуассона з параметром *λt* (пишемо *N~P(λt)*). Випадкова величина N приймає значення r з імовірністю

$$P(N = r) = e^{-\lambda t} \frac{(\lambda t)^r}{r!}. \quad (1.1.8)$$

Розподіл Пуассона може використовуватися в якості гарного наближення біноміального розподілу, якщо n велике, а p мале. Тоді в якості λ потрібно брати np.

Неперервні моделі індивідуальних позовів

Якщо величина *Y* дійсно поданого позову є абсолютно неперервною, то модель �� = ���� індивідуального позову називається неперервною. Оскільки *Y>0*, то її розподіл *F* зосереджений на (0; +∞), при цьому: +∞

���� = ∫ ����(��)����

0

Зокрема, якщо випадкова величина *Y* абсолютно неперервна з щільністю *f*, то:

+∞

���� = ∫ ����(��)����

0

Приклади розподілів дійсно поданого позову [14].

Рівномірний розподіл

Хоча рівномірний розподіл не є типовим в теорії ризику, однак виділяється своєю простотою.

Щільність рівномірного розподілу на проміжку [*a;b*] дорівнює:

1

��(��) = { �� ∈ [��; ��] 0, �� ∉ [��; ��]

�� − ��,

Функція рівномірного розподілу на проміжку [*a;b*] дорівнює:

0, �� < ��

��(��) = { �� ∈ [��; ��] 1, �� ≥ ��

�� − ��

�� − ��,

Моменти рівномірного розподілу на проміжку [*a;b*]:

$$M\xi = \frac{a + b}{2}$$

$$D\xi = \frac{(b - a)^2}{12}$$

Експоненціальний розподіл

Випадкова величина ξ має експоненціальний розподіл з параметром λ>0, якщо вона абсолютно неперервна і її щільність рівна:

$$f_{\xi}(x) = \begin{cases} 0, & x < 0 \\ \lambda e^{-\lambda x}, & x \geq 0 \end{cases}$$

Якщо процес пред'явлення позовів є пуассонівським з показником λ, то випадкова величина T (час до першого позову) розподілена за експоненціальним розподілом (T~E(λ)).

Розподіл Парето

Випадкова величина T має розподіл Парето з параметрами α і λ, якщо

$$F(x) = 1 - \left(\frac{\lambda}{\lambda + x}\right)^{\alpha}$$

Розподіл Парето має функцію щільності:

$$f(x) = \frac{\alpha \lambda^{\alpha}}{(\lambda + x)^{\alpha+1}}$$

Пишемо T~P(α,λ).

Гамма - розподіл

Випадкова величина T має гамма-розподіл з параметрами α і δ, якщо щільність рівна:

$$f(x) = \frac{\delta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\delta x}$$

Функція Γ(α) - це відома гамма-функція, обрахована в такий спосіб:

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$$

Пишемо T~G(α,δ).

Нормальний розподіл

Випадкова величина ξ має нормальний розподіл з параметрами (m,σ^2), якщо її щільність рівна:

$$f_{\xi}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}, \quad -\infty < x < \infty$$

Пишемо $T \sim N(m,\sigma^2)$.

Відповідна функція розподілу має вигляд:

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

Імовірність потрапляння випадкової величини T в інтервал (α, β) обчислюється за формулою:

$$P(\alpha < T < \beta) = \int_{\alpha}^{\beta} f(x) dx = \Phi\left(\frac{\beta - m}{\sigma}\right) - \Phi\left(\frac{\alpha - m}{\sigma}\right)$$

де Φ(z) - функція Лапласа.

Нормальний розподіл обмежено використовується для опису розподілу розміру позовів, оскільки останній дуже несиметричний.

Логнормальний розподіл

Щільність логнормального розподілу з параметрами (μ,σ):

$$f(x)=\begin{cases}\frac{1}{\sigma x\sqrt{2\pi}}e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}, & x>0\\ 0, & x\le 0\end{cases}$$

Функція логнормального розподілу:

$$F(x)=\begin{cases}\Phi_{0;1}\left(\frac{\ln x-\mu}{\sigma}\right), & x>0\\ 0, & x\le 0\end{cases}$$

Моменти:

$$E\xi^{k}=e^{k\mu+\frac{k^2\sigma^2}{2}}$$

Розподіл Вейбулла

Щільність розподілу Вейбулла з параметрами (α,b):

$$f(x)=\begin{cases}\frac{\alpha}{b}\left(\frac{x}{b}\right)^{\alpha-1}e^{-\left(\frac{x}{b}\right)^{\alpha}}, & x>0\\ 0, & x\le 0\end{cases}$$

Функція розподілу Вейбулла:

$$F(x)=\begin{cases}1-e^{-\left(\frac{x}{b}\right)^{\alpha}}, & x>0\\ 0, & x\le 0\end{cases}$$

Моменти розподілу Вейбулла:

$$E\xi^{k}=b^{k}\Gamma\left(1+\frac{k}{\alpha}\right)$$

1.2. Модель індивідуального ризику

Використано джерела [6; 7; 10; 14; 25; 26].

Модель індивідуального ризику – найпростіша з моделей функціонування страхової компанії, за допомогою якої можна обчислити ймовірність банкрутства компанії. Вона базується на таких припущеннях:

1. Аналізується короткий проміжок часу (можна знехтувати інфляцією та не враховувати прибуток від інвестування);

2. Число договорів страхування N фіксоване;

3. Плата за страховку вноситься на початку періоду страхування (ніяких надходжень протягом цього періоду немає), розрахунок проводиться в кінці періоду;

4. Кожен договір страхування розглядається незалежно від інших, розподіл індивідуального позову X_i за i-им договором відомий [7]. За наявності цих припущень ймовірність банкрутства визначається сумарним позовом

$$S = \sum_{i=1}^{N} X_i$$

до страхової компанії та величиною резервного капіталу компанії u. Якщо сумарний позов S більший, ніж капітал компанії u, то компанія не зможе виконати свої зобов'язання і збанкрутує. Тому ймовірність банкрутства компанії дорівнює:

$$R = P\left\{\sum_{i=1}^{N} X_i > u\right\} \quad (1.2.1)$$

У припущенні, що розподіли величин позовів $X_1, X_2, \dots, X_N$ відомі й випадкові величини $X_1, X_2, \dots, X_N$ незалежні, розподіл суми $\sum_{i=1}^{N} X_i$ можна обчислити як згортку розподілів цих випадкових величин.

За розподілами F та G незалежних випадкових величин X_1 та X_2 розподіл Q суми $X_1 + X_2$ одержуємо як згортку розподілів F та G [6, с.75]:

$$Q(x) = \int F(x - y)G(dy) = \int G(x - y)F(dy)$$

Якщо ��({��}) = ����, ��({��}) = ����, то:

�� ��

$$��(\{��\}) = ���� = \sum_{��=0} ������ -\, ������ = \sum_{��=0} ������ -\, ������$$

Число договорів *N*, як правило, велике. Тому конкретні розрахунки за допомогою згорток, виявляються досить складними, і виникає необхідність знаходження апроксимаційних формул для функції розподілу сумарних виплат. Теоретичним підгрунтям для розв'язку такої задачі можуть служити граничні теореми для розподілу сум випадкових величин і, у першу чергу, центральна гранична теорема (ЦГТ).

Нормальна апроксимація. [6]. Із ЦГТ для сум незалежних однаково розподілених випадкових величин випливає, що

$$\lim_{��\to\infty} ��\,(���� - ����_1\ ��\sqrt{����}) = \Phi(��),$$

$$\lim_{��\to\infty}(���� \le ����_1 + ��\sqrt{��} \le ��) = \lim$$

��

$$\Phi(��) = \Phi_{0,1}(��) = \frac{1}{\sqrt{2��}} \int_{-\infty} e^{-��^2/2}\, ���� \quad (1.2.2)$$

- функція розподілу стандартного нормального закону.

Отже, як апроксимацію для ����(��) розглядають

$$����(��) = ��(���� \le ��) \approx \Phi_{����,����(��)} \quad (1.2.3)$$

де $\Phi_{��,��(��)}$- функція розподілу нормального закону з параметрами *a* і дисперсією *b* , ���� = ����$_1$, ���� = ����^{2}, або еквівалентну

формулу: $����(��) \approx \Phi_{0,1}(��)\left(\frac{�� - ����}{\sqrt{����}}\right)$. (1.2.4)

Узагальнення ЦГТ для різнорозподілених і /або залежних випадкових величин дає можливість застосувати (1.2.3) і (1.2.4) для широкого класу задач. У випадку різнорозподілених незалежних ��$_{��}$, �� = 1, … , �� ��

$$���� = ����_{��} = \sum_{��=1} ����_{��},$$

$$��(��) = ��(��_{��}) = \sum_{��=1}^{��} ��_{��}{}^{2},$$

Пуассонівська апроксимація. Припустимо, що *n* (кількість договорів) велике, і в разі настання страхової події за кожним договором виплачується однакова сума *b* (яку можна прийняти за одиницю), а ймовірність настання страхової події *p* мала. Тоді для апроксимації розподілу ��$_{��}$ має сенс розглядати розподіл Пуассона з параметром *pn* , тобто

$$��_{��}(��) \approx ��(��; ����) \quad (1.2.5)$$

Теоретичним підгрунтям для такої апроксимації є граничні теореми стосовно збіжності до розподілу Пуассона.

Г – апроксимація. Дуже часто розподіл сумарних виплат має багато спільних рис з Г – розподілом: він виявляється одномодальним, носій – невід'ємна піввісь, коефіцієнт асиметрії �� > 0. Тому слушним виявляється підхід, при якому��$_{��}$(��) наближають за допомогою Г – розподілу із зсувом, тобто

$$��_{��}(��) \approx ��(�� - ��_0; ��, ��), \quad (1.2.6)$$

де

$$��(��; ��, ��) = \frac{1}{��} \; Γ(��) \int ��^{��-1} � \; �^{��}_{0} \; e^{-����} � \; ���$$

а параметри ��, ��, ��$_0$ підбирають так, щоб перші три моменти суми ��$_{��}$ (точніше, середнє ����, дисперсія ��$_{��}{}^{2}$, і коефіцієнт асиметрії ��$_{��}$) дорівнювали відповідним моментам апроксимуючого Г-розподілу. Нагадаємо, що для Г розподілу середнє ��$_1$ = ��/��, дисперсія ��^{2} = �� ��^{2} / , коефіцієнт асиметрії �� = 2/√��.

Отже ��, ��, ��$_0$ знаходимо з рівнянь:

���� = ��$_0$ + ��/��,

��$_{��}{}^2$ = �� ��^{2} / ,

��$_{��}$ = 2/√��.

Звідси для знаходження ��, ��, ��$_0$ отримаємо співвідношення

�� = ^{4} ��$_{��2}$, �� =^{2}

��$_{��}$��$_{��}$, ��$_0$ = ��$_1$ –2��$_{��}$

��$_{��}$(1.2.7)

1.3. Модель колективного ризику

Використано джерела [6; 7; 13;14; 36].

Модель колективного ризику базується на таких припущеннях:

1. Аналізується фіксований короткий інтервал часу;

2. Плата за страховку вноситься на початку періоду страхування, розрахунок проводиться в кінці, надходжень протягом цього періоду немає; 3. Позови ��$_1$, ��$_2$, ..., що надходять до компанії, не пов'язані з конкретними договорами, а розглядаються як результат сумарного ризику компанії;

4. Випадкова величина *ν* – загальне число позовів за період страхування [6, 92].

Щоб обчислити ймовірність банкрутства �� = ��{��$_{��}$ > ��} застосуємо формулу повної ймовірності:

∞ ∞

�� = ��{��$_{��}$ > ��} = �� } = ∑��{��$_{��}$ > ��, �� =

{��$_{��}$ > ��,∪{�� = ��} ��} ��=0

��=0 ∞

∞

= ∑��{��$_{��}$ > ��, �� = ��} ��=1

= ∑��{��$_{��}$ >

��}��{�� = ��}   ��=1

Позначивши ��{�� = ��} = ����, ймовірність банкрутства запишемо у вигляді:

∞

�� = ∑��{���� > ��} ��=1 ���� (1.3.1)

Для обчислення ймовірності ��{���� > ��} знайдемо розподіл випадкової величини

���� = ��1 + ��2 + ⋯ + ����

як розподіл суми незалежних випадкових величин ��1, ��2, … , ���� (він дорівнює згортці розподілів доданків), а потім обчислимо ��{���� > ��}.

Якщо випадкові величини Y_i – дискретні, то:

$$��\{���� > ��\} = \sum_{��=��+1}^{\infty} ��\{���� = ��\}$$

Ймовірність банкрутства дорівнює:

∞ ∞ ∞ ∞

�� = ∑(∑ ���� ��{���� = ��}) = ∑���� ∑ ��{���� = ��}

��=1 +1 ��=�� ��=1 +1 ��=��

∞ ∞

= ∑ ∑��{���� = ��}���� (1.3.2)

��=��+1 ��=1

Якщо випадкові величини Y_i – абсолютно неперервні, то:

∞

��{���� > ��} = ∫ ��������(��) ���� . Тоді ймовірність банкрутства:

��

де ��������(��) – щільність розподілу ����

∞

∞

∞

∞

∞

�� = > = ∑��₍��₎ ∫ �) (∑��₍��₎��) ∑��{��₍��₎ ��}��₍��₎ ��₍����₎(� ���� = ∫ ₍����₎(��) ���� (1.3.3)

��=1 ��=1 �� �� ��=1

1.4. Логістична регресія

Логістична функція, також відома як сигмоїдна функція, використовується у логістичній регресії для того, щоб відображати будь-яку дійсну числову величину $z \in R^1$в діапазон між 0 і 1, а значення з цього діапазону можна було розглядати як ймовірність події:

$$\sigma: z \rightarrow (0,1)$$

Аргумент сигмоїдної функції визначається як лінійна згортка ваг ��^�� = (��₀, ��₁, … . . ��₍��₎) та ознак ��₍��₎, досліджуваного об’єкта, представлених у числовому вигляді �� = (1, ��₁, … . . ��₍��₎)^��, m – кількість ознак об’єкта ^��₍��=0₎ ∈ ��^1.

z = ��^���� = ∑ ��₍��₎��₍��₎

Формула логістичної функції має вигляд:

σ(z) = p =^1

1+�� ^−�� ∈ (0,1)

і визначає ймовірність p = P(Y = 1) належності об’єкта до певного класу1 [18]. Графіки сигмоїдної функції та її похідної зображено на рис. 1.

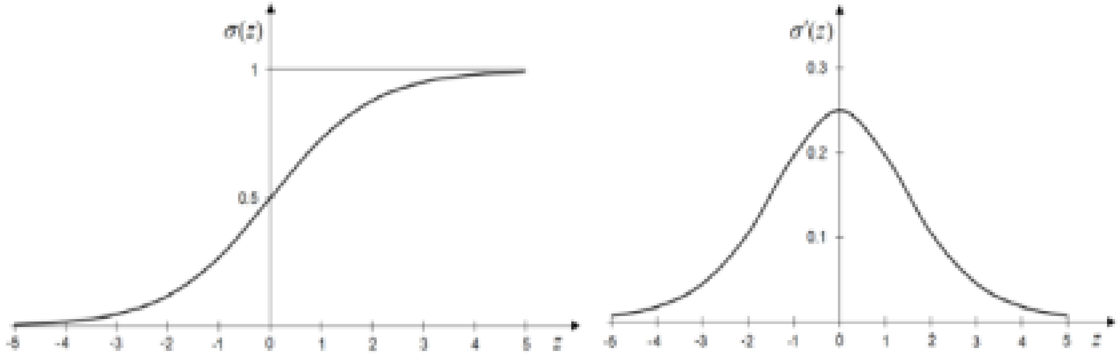

Рис.1

а) Стандартна логістична функція, б) Похідна логістичної функції На початковому етапі побудови моделі для оцінки ймовірності банкрутства необхідно визначити показники, які найбільше впливають на фінансову стійкість страхової компанії. До таких, наприклад, можна віднести: *платоспроможність, ліквідність активів*, *рівень фінансової незалежності*, *рентабельність активів* (ROA), *коефіцієнт боргового навантаження,* тощо.

Модель логістичної регресії оцінки ймовірності банкрутства може бути представлена у вигляді:

$$P(Y=1)=\frac{1}{1+e^{-(\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_n X_n)}}$$

де Y — залежна змінна, що набуває значення 1 у разі банкрутства і 0 в іншому випадку; $X_1,X_2,\dots,X_n$ — незалежні змінні (значення фінансових показників); $\beta_0,\beta_1,\dots,\beta n$ — параметри моделі (ваги), які визначаються за допомогою методу максимальної правдоподібності (maximum likelihood method).

1.5. Основні моделі первинного аналізу

Використано джерела [4; 20; 21; 22; 37].

1.5.1. Описова статистика

Описова статистика забезпечує первинний аналіз даних шляхом їх узагальнення та візуалізації. Основні показники включають: *Міри центральної тенденції:*

· Середнє арифметичне характеризує середнє значення розподілу. · Медіана Me відображає центральне значення у впорядкованому ряді. · Мода *Mo* — значення, яке зустрічається найчастіше.

Міри варіації: дисперсія (σ^2), стандартне відхилення (σ), коефіцієнт варіації V, що вимірює рівень розсіювання даних.

Графічні методи: гістограми, діаграми розкиду, коробкові діаграми. Реалізація таких показників та графічних методів часто автоматизується в статистичних пакетах, таких як SPSS, R, або Python. **1.5.2. Кореляція**

Кореляційний аналіз використовується для оцінки сили та напряму взаємозв’язків між змінними.

Коефіцієнт кореляції Пірсона: вимірює лінійну кореляцію між двома наборами даних *X* та *Y*

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n * s_x * s_y}.$$

Він варіюється в межах [−1,1], де $r > 0$ означає додатний зв'язок, а $r < 0$ — від’ємний. Чим ближче до 1 або -1 значення кореляції, тим сильніший зв'язок. Якщо значення знаходиться ближче до 0, то це означає слабкий зв'язок.

Коефіцієнт Спірмена використовується для оцінки монотонних зв’язків між порядковими змінними (досліджуються не самі дані, а їх ранги).

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

де $d_i = x_i - y_i$ – різниця рангів для *i*-го об’єкту серед n спостережень.

Як і для звичайного коефіцієнта кореляції, значення коефіцієнта *R* змінюється від –1 до 1. Чим ближче абсолютне значення коефіцієнта рангової кореляції до одиниці, тим більш щільним є зв’язок між факторами.

Матриця кореляцій: у багатовимірних даних дозволяє оцінити взаємозв’язки між кількома змінними одночасно. Аналіз матриці коефіцієнтів парної кореляції використовується при побудові моделей множинної регресії. Програмне забезпечення надає можливості обчислення кореляцій та побудови графіків для аналізу зв’язків.

1.5.3. Кластеризація

Кластерний аналіз — це метод групування об'єктів на основі схожості їх характеристик. Основні методи:

Ієрархічна кластеризація:

- Результат представлений у вигляді дендрограми,
- Відстань між об'єктами часто обчислюється за допомогою евклідової

метрики ��2���� = √∑ (������ − ������)^2 ��=1, де t – розмірність простору.

Метод K-середніх: будує задане число кластерів, розташованих якнайдалі один від одного.

DBSCAN: алгоритм кластеризації на основі щільності, подібний до методу зсуву середнього значення.

Ці методи широко реалізовані в сучасних статистичних пакетах і використовуються для класифікації, сегментації, тощо.

1.5.4. Регресія

Регресійний аналіз дозволяє моделювати залежність однієї змінної (залежної) від однієї або кількох незалежних змінних.

*Лінійна регресія:*Модель: $Y = b_1 X + c$, де Y – значення ознаки по лінії регресії, тобто теоретичні значення, b_1 – кутовий коефіцієнт регресії, X –

значення ознаки-фактору (предиктора), с – вільний член, константа. Якщо незалежна змінна одна, то регресія називається парною.

Множинна регресія: Розширює модель для декількох незалежних змінних:

$Y = b_1 \cdot X_1 + b_2 \cdot X_2 + b_3 \cdot X_3 + \ldots + b_k \cdot X_k + c$

РОЗДІЛ 2

ЗАСТОСУВАННЯ ЙМОВІРНІСНИХ МЕТОДІВ СТРАХОВОЇ МАТЕМАТИКИ ТА МЕТОДІВ СТАТИСТИЧНОГО АНАЛІЗУ ДО РОЗВ’ЯЗАННЯ ПРАКТИЧНИХ ЗАВДАНЬ

2.1. Моделі позовів

Використано джерела [6; 10; 25;30;36].

Приклад 1. Розподіл величини дійсно поданого позову.

Розподіл величини дійсно поданого позову Y для договорів страхування автомобілів дорівнює:

Величина дійсно поданого позову	20	30	40	50	60	70	80
Ймовірність	0,15	0,10	0,05	0,20	0,10	0,10	0,30

Обчислити ймовірність того, що величина дійсно поданого позову Y відрізняється від свого середнього менше, ніж на одне стандартне відхилення.

Розв’язання:

$$P\{|Y - EY| < \sqrt{DY}\} = P\{|Y - 55| < \sqrt{475}\} =$$

$$P\{|Y - 55| < 21{,}8\} = P\{33{,}2 < Y < 76{,}8\}$$

$$= P\{Y = 40\} + P\{Y = 50\} + P\{Y = 60\} + P\{Y = 70\}$$

$$= 0{,}05 + 0{,}2 + 0{,}1 + 0{,}1 = 0{,}45$$

Відповідь: $P\{|Y - EY| < \sqrt{DY}\}$ =0,45.

Приклад 2. Перевірка якості підгонки.

У портфелі автотранспортного страхування зібрана інформація про страхові позови по 100000 річним договорам страхування. У таблиці показане зареєстроване число страхувальників, що пред'явили 0,1,2,3,4,5 позовів протягом року.

Число вимог	Число договорів
0	81056
1	16174
2	2435
3	295
4	36

5	4

Перевірити, чи дає пуассонівська модель гарний опис цих даних.

Розв'язання:

Виберемо для підгонки пуассонівську модель із математичним сподіванням, що співпадає із середнім значенням наведених даних. $\lambda = \bar{x} = 0,22093$

По формулі

$$P(X = k) = e^{-\lambda} \frac{\lambda^{k}}{k!}$$

обчислимо ймовірність того, що за договором страхування буде пред'явлено 0,1,2,... позовів. Результати розрахунків наведемо в таблиці 2.1. Таблиця 2.1

Число вимог	Імовірність
0	0,801773
1	0,177136
2	0,019567
3	0,001440
4	0,000080
5	0,000004

Зареєстровані (емпіричні) й очікувані (теоретичні) частоти позовів представимо в таблиці 2.2.:

Таблиця 2.2.

Число позовів	Емпіричні частоти	Теоретичні частоти
0	81056	80177,3
1	16174	17713,6
2	2435	1956,7

3	295	144
4	36	8
5	4	0,4
Разом	100000	100000

Обчислимо

$$\chi^2 = \sum_{i=1}^{5} \frac{(n_i - n_i')^2}{n_i'} = 549$$

Визначимо $\chi^2_{кр}$ з рівнем значущості 0,05 і числом ступенів свободи 4.

$\chi_{кр}^2 = \chi^2(0{,}05; 4) = 9{,}5.$

Оскільки $\chi^2_{кр} < \chi^2$, то гіпотеза про опис даних розподілом Пуассона відкидається.

2.2. Знаходження числових характеристик випадкової величини та оцінка параметрів розподілу

Використано джерела [19; 20; 21; 22; 36].

Приклад 3. Кількісні характеристики випадкової величини. Нехай випадкова величина X розподілена за законом E(λ). Знайти $\varphi_X(t)$, MX,DX.

Розв'язання:

Маємо

$$\varphi_X(t) = \int_0^\infty e^{itx} \lambda e^{-\lambda x} dx = \lambda \int_0^\infty e^{-(\lambda - it)x} dx = \frac{\lambda}{\lambda - it}$$

При обчисленні інтеграла була використана наступна властивість: якщо f(t) - це довільна функція щільності, то

$$\int_{-\infty}^{\infty} f(t)\,dt = 1$$

Цей факт ми використали, помітивши, що підінтегральний вираз в другому інтегралі пропорційний функції щільності експоненціального розподілу E(λ-s).

$$M_X(s) = \frac{\lambda}{\lambda - s} = \left(1 - \frac{s}{\lambda}\right)^{-1}$$

$$M_X(s) = 1 + \frac{s}{\lambda} + \left(\frac{s}{\lambda}\right)^2 + \left(\frac{s}{\lambda}\right)^3 + \cdots$$

$$EX = M_X'(0) = 1/\lambda$$

$$E(X^2) = M_X''(0) = 2/\lambda^2 \Rightarrow D(X) = E(X^2) - E^2(X) = 1/\lambda^2$$

Відповідь: $M_X(s) = \frac{\lambda}{\lambda - s}$, $EX = 1/\lambda$, $D(X) = 1/\lambda^2$.

Приклад 4. Математичне сподівання розподілу Парето.

Нехай $X \sim P(\alpha, \beta)$. Довести, що $EX = \frac{\beta}{\alpha - 1}$.

Розв'язання:

Розподіл Парето має функцію щільності:

$$f(x) = \frac{\alpha \beta^{\alpha}}{(x + \beta)^{\alpha + 1}}$$

Математичне сподівання абсолютно неперервної випадкової величини *X* обчислюється за формулою:

$$EX = \int_{-\infty}^{\infty} x f(x)\,dx$$

��)��+1���� ��

∞

= ������(�� +

���� = ∫ ��   = [��)−��

26

.������

(�� +

��)��+1����   �� = �� ���� = ����   ]

0   ���� =������   −��

(�� +

∞

− ∫

= �� ·   ����   ��   =������

���� ��   ((−��|0 ∞

�� +   �� +

��)−��   ��)−��   −������

∞   0

−��.(�� + ��)−��+1

=��������−��+1

(−��)(−�� + 1)=��   −�� + 1|0

�� − 1

Доведено.

Приклад 5. Оцінка параметру експоненціального розподілу методом максимальної правдоподібності.

Знайти оцінку параметра α для експоненціального розподілу методом максимальної правдоподібності.

Розв’язання:

��(����) = ��e−������

��=1

��   ��

∏��(����)   ��

��=1   = ����.

exp(−��∑���

��   ��

=   ����∏��(�

∏��e−������   )   ���) ��=1

��=1

= ������ ^�� �� ��=1) =
+ ���� ��(���� ∏ ��������� −
(exp(−��∑�� ��(����)) ��∑�� �� ��=1

$$���� = \frac{1}{��} ��−∑��_{��} = 0$$

$$��−∑��_{��} = 0$$

$$\hat{��} = (\sum \frac{��}{��})^{-1} = (\bar{x})^{-1}$$

Відповідь: $\hat{��} = (\bar{x})^{-1}$

2.3. Модель індивідуального ризику

Використано джерела [6; 10; 25; 26;30].

Приклад 6. Знаходження мінімальної величини капіталу, необхідної для мінімізації ймовірності банкрутства в моделі індивідуального ризику. Портфель страхової компанії складається з 1800 договорів страхування життя строком на 1 рік. Страхові виплати поділяються на 2 групи розміром 100 і 200 у.о. для застрахованих з імовірністю настання страхового випадку 0,02 і 0,1 відповідно. Далі наведена кількість застрахованих n_k у кожній із чотирьох груп, що відповідає різним значенням страхової виплати b_k і ймовірності позову q_k.

k	q_k	b_k	n_k
1	0,02	100	500
2	0,02	200	500
3	0,1	100	300
4	0,1	200	500

Визначити мінімальну величину капіталу, за наявності якого ймовірність банкрутства компанії не перевищує 0,05.

Розв’язання:

Візьмемо величину 100 у.о. за одиницю виміру грошових сум. Тоді індивідуальний позов X_i до компанії застрахованого першої групи має розподіл:

$$\begin{pmatrix} 0 & 1 \\ 0,98 & 0,02 \end{pmatrix}$$

другої групи:

$$\begin{pmatrix} 0 & 2 \\ 0,98 & 0,02 \end{pmatrix}$$

третьої групи:

$$\begin{pmatrix} 0 & 1 \\ 0,9 & 0,1 \end{pmatrix}$$

четвертої групи:

$$\begin{pmatrix} 0 & 2 \\ 0,9 & 0,1 \end{pmatrix}$$

При цьому $MX_1 = 0,02$, $MX_2 = 0,04$, $MX_3 = 0,1$, $MX_4 = 0,2$, а $DX_1 = 0,0196$, $DX_2 = 0,0784$, $DX_3 = 0,09$, $DX_4 = 0,36$

$X_1, X_2, \ldots, X_{1800}$ – незалежні випадкові величини.

Тоді $S = X_1 + X_2 + \ldots + X_{1800}$.

$$MS = M(X_1 + X_2 + \cdots + X_{1800}) = 500 \cdot 0,02 + 500 \cdot 0,04 + 300 \cdot 0,1 + 500 \cdot 0,2 = 160$$

$$DS = D(X_1 + X_2 + \cdots + X_{1800}) = 500 \cdot 0,0196 + 500 \cdot 0,0784 + 300 \cdot 0,09 + 500 \cdot 0,36 = 256$$

Знайдемо величину капіталу компанії, за якого ймовірність банкрутства рівна 0,05.

$$0,05 = P\{S > u\} = 1 - \Phi_{0;1}\left(\frac{u - MS}{\sqrt{DS}}\right)$$

$��_{0;1}$ ($�� - 160$

$\sqrt{256}) = 0,95$

$�� - 160$

$16 = 1,645$

Отже, мінімальний капітал компанії має бути не меншим *u=186,32* або 18632 у.о.

Приклад 7. Модель індивідуального ризику та величина дійсно поданого позову, що має зрізаний експоненціальний розподіл.

Клієнти компанії, що займається страхуванням автомобілів, поділяються на 2 групи:

Група	Кількість у групі	Імовірність позову	Параметри зрізаного експоненціального розподілу	
k	n_k	q_k	λ	L
1	500	0,10	1	2,5
2	2000	0,05	2	5,0

Величина дійсно поданого позову має зрізаний експоненціальний розподіл, який задається функцією розподілу

0, якщо $x < 0$

$��(��) = \{$ $1 - e^{-����}$, якщо $0 \leq �� \leq$ $��$ 1, якщо $�� \geq ��$

Визначити мінімальну величину капіталу, за наявності якого ймовірність банкрутства компанії не перевищує 0,05.

Розв’язання:

Обчислимо середнє та дисперсію величини дійсно поданого позову Y:

$��$ $���� = \int$

����e⁻���� �� ��²��e⁻���� ��

�� + �� 0 �� + ��² 0

e⁻���� =1 − e⁻���� ��, e⁻���� =² ��² (1 − e⁻����) − 2�� ��e⁻����,

��

����² = ∫

���� =1 − 2����e⁻���� − e⁻²����

��²

Для першої групи клієнтів параметри λ=1, L=2,5, тоді MY=0,9179, DY=0,5828. Для другої групи клієнтів параметри λ=2, L=5, тоді MY=0,5, DY=0,2498.

Отже, середнє та дисперсія індивідуального позову X до компанії дорівнюють

���� = ���� · �� {�� = 1},

���� = ���� · ��{�� = 1} + (����)²��{�� = 1}��{�� = 0}.

Для клієнтів першої групи маємо MX=0,09179, DX=0,13411. Для клієнтів другої групи маємо MX=0,025, DX=0,02436.

Математичне сподівання MS та дисперсія DS сумарного позову S до компанії дорівнюють

���� = 500 · 0,09179 + 2000 · 0,025 = 95,89,

���� = 500 · 0,13411 + 2000 · 0,02436 = 115,78.

Знайдемо величину капіталу компанії, за якого ймовірність банкрутства компанії дорівнює 0,05. Маємо

0,05 = ��(�� > ��) = 1 − ��$_{0;1}$ (�� − ���� / √����)

��$_{0;1}$ (�� − 95,89

$\sqrt{115{,}78}$) = 0,95

Звідси

�� − 95,89

$\sqrt{115{,}78}$ = 1,645.

Мінімальний капітал компанії має бути не менший ніж *u*=113,59.

2.4. Модель колективного ризику

Приклад 8. Знаходження математичного сподівання сумарного позову. Сумарний позов до страхової компанії дорівнює ��₍��₎ = ��₁ + ��₂ + ⋯ + ��₍��₎, де *v* – загальне число позовів набуває значення 0, 1, 2 з однаковими ймовірностями. Кожна з випадкових величин ��₍��₎ експоненціально розподілена з середнім 1/2. Випадкові величини ��, ��₍��₎ незалежні. Знайти ��������{��₍��₎}. *Розв'язання:*

Події {�� = ��}, �� = 0,1,2 утворюють повну групу подій.

Згідно з формулою повного математичного сподівання маємо:

2

��������{��₍��₎} = ∑��(������{��₁ + ��₂ + ⋯ + ��₍��₎} · {�� = ��}) ��=0 · ��{�� = ��}

� · ��{�� = ��} ��=0

= ∑��(������{��₁})� 2 = 1/3(1 + ��������{��₁} + (��������{��₁})��)

Оскільки ��������{��₁} = 2, то ��������{��₍��₎} = 7/3.

2.5. Аналіз даних в комп'ютерному статистичному пакеті SPSS

Приклад 9. Виявити фактори, які сприяють підвищенню ризиків смертельних зіткнень у штатах США, та згрупувати регіони за характерними показниками аварійності.

Завдання аналізу містять:

1. Оцінку статистичних характеристик основних показників. 2.

Визначення взаємозв'язків між змінними за допомогою кореляційного аналізу.

3. Класифікацію штатів за рівнем аварійності методом кластерного аналізу.

4. Побудову регресійної моделі для оцінки впливу окремих факторів на ризики смертельних зіткнень.

Дослідження базується на статистичних даних автомобільної аварійності та страхових збитків в США *(Додаток Г)*, а саме: 1) кількість водіїв, які потрапили у смертельні зіткнення на мільярд миль, 2) відсоток водіїв, залучених у смертельні зіткнення, які перевищували швидкість, 3) відсоток водіїв, залучених у смертельні зіткнення, які перебували в стані алкогольного сп'яніння, 4) відсоток водіїв, залучених у смертельні зіткнення, які не відволікалися, 5) відсоток водіїв, залучених у зіткнення зі смертельними наслідками, які раніше не брали участі в аваріях, 6) премії за страхування автомобіля ($), 7) збитки, понесені страховими компаніями через зіткнення, на одного застрахованого водія ($). Дані охоплюють 51 штат США, що забезпечує репрезентативність вибірки [39; 40].

Розв'язання:

1. Описова статистика

Для первинної оцінки даних використано методи описової статистики, які застосуємо до наступних факторів (змінних) - відсоткових даних водіїв, залучених до смертельних ДТП:

- ✔ що перевищили швидкість;
- ✔ які перебували у стані алкогольного сп'яніння;
- ✔ які не відволікалися під час аварій;
- ✔ що раніше не були причетні до аварій.

Це дозволило ідентифікувати основні характеристики кожної змінної, включаючи середнє значення, стандартне відхилення, мінімальні та максимальні показники (подано в табл.2.3).

Таблиця 2.3

Основні статистичні характеристики змінних

Стандартна похибка

відхилення

Percentage Of Drivers Involved In

N Мінімум Максимум Среднє

13,00%	54,00%	31,7255%
16,00%	44,00%	30,6863%
10,00%	100,00%	85,9216%
76,00%	100,00%	88,7255%

Fatal Collisions Who Were Speeding Percentage Of Drivers Involved In Fatal Collisions Who Were Alcohol Impaired

Percentage Of Drivers Involved In Fatal Collisions Who Were Not Distracted

Percentage Of Drivers Involved In Fatal Collisions Who Had Not Been Involved In Any Previous Accidents

51 9,63344% 51 5,13221%

51 15,15895% 51 6,96011%

N валидных (по списку) 51

Проаналізуємо ці результати:

Варіативність між змінними:

- Найбільша варіативність (стандартне відхилення SD=15.16%) спостерігається для водіїв, які не відволікалися під час смертельних зіткнень. Це свідчить про значну різницю між штатами за цим показником.
- Найменша варіативність (SD=5.13%) спостерігається для водіїв, які перебували у стані алкогольного сп’яніння. Це вказує на те, що частка таких водіїв є більш стабільною серед штатів.

Середні значення:

- Середній відсоток водіїв, які перевищували швидкість (M=31.73%), та водіїв у стані алкогольного сп’яніння (M=30.69%) є майже однаковими, що

підтверджує, що обидва фактори є поширеними причинами смертельних зіткнень.

- Високий середній відсоток водіїв, які не відволікалися (M=85.92%), та тих, хто раніше не був причетний до смертельних зіткнень (M=88.73%), вказує на те, що ці фактори є менш значущими в контексті аварійності. *Мінімальні та максимальні значення:*

· Найнижчий мінімум (10%) спостерігається для водіїв, які не відволікалися, що може свідчити про значні проблеми із концентрацією в окремих штатах. · Максимум у 100% для цієї ж змінної вказує на те, що в деяких штатах водії демонструють максимальну увагу під час керування.

2. Кореляційний аналіз

Для оцінки взаємозв'язків між змінними проведено кореляційний аналіз за методом Пірсона (Д*одаток Д*). Це дозволило визначити як силу, так і напрямок зв'язків між показниками і отримати наступні результати: *Значущі кореляції:*

· Помірний позитивний зв'язок ($r=0.286$, $p=0.042$) між перевищенням швидкості та алкогольним сп'янінням вказує на те, що ці два фактори часто взаємопов'язані. Це свідчить про високий ризик аварійності у штатах, де поширені ці дві поведінки водіїв.

Незначущі кореляції:

· Інші взаємозв'язки, включаючи вплив відсутності відволікань чи попередньої участі в смертельних зіткненнях, не досягли статистичної значущості. Це може вказувати на те, що ці фактори є незалежними від інших змінних у вибірці.

Особливості змінних:

· Низький зв'язок між відсутністю відволікань та іншими показниками може свідчити про те, що цей фактор менш важливий у контексті аварійності.

· Слабка негативна кореляція ($r=-0.245$) між попередньою відсутністю в смертельних зіткненнях і алкогольним сп'янінням потребує додаткового вивчення.

3. Кластерний аналіз

Для групування штатів за характерними показниками використано в комп'ютерному статистичному пакеті SPSS ієрархічний кластерний аналіз із застосуванням методу K-середніх. За метрику відстані обрано евклідову відстань. Насамперед знайдено початкові центри кластерів (поданов таб.2.4).

Таблиця 2.4

Початкові центри кластерів

Кластер

1 2 3

19,00%
27,00%
67,00%
98,00%

Percentage Of Drivers Involved In Fatal Collisions Who Were Speeding 50,00% 15,00%

Percentage Of Drivers Involved In Fatal Collisions Who Were Alcohol Impaired

Percentage Of Drivers Involved In Fatal Collisions Who Had Not Been Involved In Any Previous Accidents

Percentage Of Drivers Involved In Fatal Collisions Who Were Not Distracted 31,00% 31,00% 96,00% 10,00% 88,00% 100,00%

Далі здійснювався ітераційний процес, що включав наступні кроки (Таб. 2.5).

Таблиця 2.5

Хронологія ітерацій

Зміни центрів кластерів

1,350
,000

Ітерація

1 2 3

17,237
3,144

1 14,194 ,000

2 ,439 19,634

3 ,870 ,000

4 ,000 ,000

В результаті одержано розбиття штатів США на 3 кластери *(Додаток Д)* та знайдено кінцеві центри кластерів (Таб.2.6).

Таблиця 2.6

Кінцеві центри кластерів

Кластер

1 2 3

Percentage Of Drivers Involved In Fatal Collisions Who Were Speeding

Percentage Of Drivers Involved In Fatal Collisions Who Were Alcohol-Impaired

Percentage Of Drivers Involved In Fatal Collisions Who Were Not Distracted

Percentage Of Drivers Involved In Fatal Collisions Who Had Not Been Involved In Any Previous Accidents

88,44%

22,56%
28,33%
84,67%

37,45% 25,50% 31,97% 32,00% 90,61% 24,50%

88,68% 92,00%

Також, визначено відстані між кластерами (подано в таб.2.7). Таблиця 2.7

Відстань між кінцевими центрами кластерів

Кластер 1 2 3

16,447
60,455

1 67,267 2 16,447 60,455 3 67,267

Визначено кількість штатів в кожному кластері (подано в таб.2.8). Таблиця 2.8

Число спостережень у кожному кластері

р 1 31,000 2 18,000

3 2,000

Валідні 51,000 Пропущені ,000

Отже, кластерний аналіз, виконаний методом **K-середніх**, розділив дані на три кластери за характеристиками сукупностей водіїв, які брали участь у смертельних аваріях (Додаток Д) (маємо к*інцеві центри кластерів****): Кластер 1*** (31 штат):

- Високий відсоток водіїв, які не відволікались (90,61%).
- Помірний рівень перевищення швидкості (37,45%).
- Рівень алкогольного сп'яніння — 31,97%.

- Високий відсоток водіїв без попередніх аварій (88,68%).

Кластер 2 (18 штатів):

· Нижчий відсоток перевищення швидкості (22,56%).
· Середній рівень алкогольного сп'яніння (28,33%).
· Відносно високий рівень водіїв, які не відволікались (84,67%). · Високий рівень водіїв без попередніх аварій (88,44%).

Кластер 3 (2 штати):

· Найнижчий рівень концентрації водіїв, які не відволікались (24,50%). · Високий рівень перевищення швидкості (25,50%) та алкоголю (32,00%). · Водії, які не потрапляли в аварії раніше — 92,00%.

Відстань між кластерами:

Найбільша відстань між кластерами 1 та 3 (67,267), що вказує на суттєві відмінності. Найменша — між кластерами 1 та 2 (16,447). *Розподіл спостережень:*

Більшість штатів потрапили в кластер 1 (31 штат). Кластер 2 охоплює 18 штатів. Лише 2 штати увійшли до кластеру 3, що свідчить про його унікальність.

Найбільший внесок у розділення кластерів внесла змінна "Відсоток водіїв, які не відволікались" (F = 61,271). Найменший вплив мала змінна "Відсоток водіїв із алкогольним сп'янінням" (F = 3,178) (показано в таб.2.9). Таблиця 2.9

ANOVA

Кластер похибка

Середній квадрат ст.св. | Середній квадрат ст.св. | F

2	67,372	48
2	49,984	48

Percentage Of Drivers Involved In Fatal Collisions Who Were Speeding

Percentage Of Drivers Involved In Fatal Collisions Who Were Alcohol-Impaired

Percentage Of Drivers Involved In Fatal Collisions Who Were Not Distracted

Percentage Of Drivers Involved In Fatal Collisions Who Had Not Been Involved In Any Previous Accidents 1303,767 30,788 77,006 3,178 4127,916 61,271 11,469

2	42,346	48
2	24,228	48

,229

Візуалізацію проробленого кластерного аналізу щодо групувань штатів подано нижче на рис. 1.

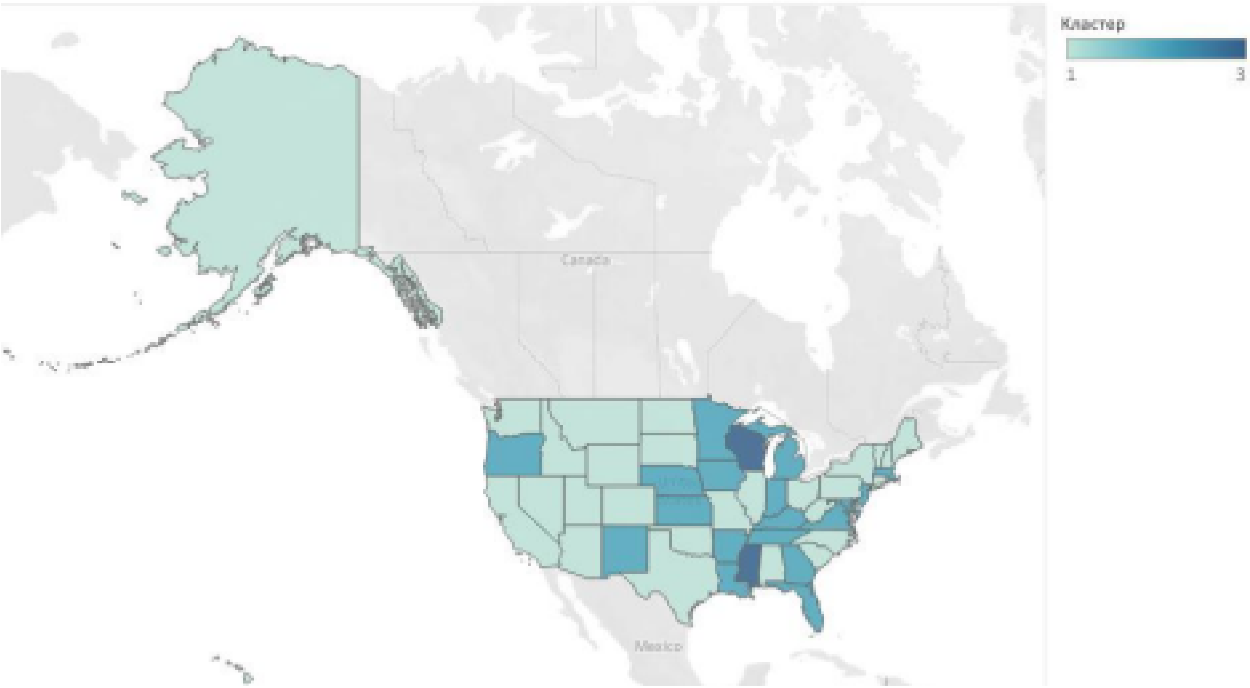

Рис. 1 ***Картограма кластеризації штатів США за факторами смертельних автомобільних аварій***

4. Регресійний аналіз

Результати регресійного аналізу показали, що жодна з обраних незалежних змінних, включаючи відсоток водіїв, які перевищували швидкість, були у стані алкогольного сп'яніння, не відволікались або не мали попередніх аварій, не має статистично значущого впливу на залежну змінну — збитки, понесені страховими компаніями через зіткнення, на одного застрахованого водія ($) *(Додаток Е)*.

Загальна характеристика моделі:

- R: 0.107 — низький рівень кореляції між предикторами та залежною змінною.
- R-квадрат: 0.011 — лише 1,1% варіації у втратах можна пояснити незалежними змінними.
- Скоригований R-квадрат: -0.075 — вказує, що модель навіть гірша за середнє значення у поясненні

варіації залежної змінної.

• Стандартна похибка оцінки: 25.746 — свідчить про суттєвий розкид даних навколо передбачених

значень.

ANOVA:

• Значення F = 0.132, значимість моделі (p) = 0.970.
• Ці результати вказують на те, що модель не є статистично значущою.

Коефіцієнти регресії:

Незалежні змінні (предиктори) не мають статистично значущого впливу на залежну змінну:

• Константа: $B = 148.551$, $p = 0.028$ — значуща на рівні 5%, вказує на середнє значення втрат за умови, що всі предиктори дорівнюють нулю. • Відсоток водіїв, які перевищували швидкість: $B = -0.093$, $p = 0.818$. • Відсоток водіїв, які були у стані алкогольного сп'яніння: $B = -0.325$, $p = 0.674$
• Відсоток водіїв, які не відволікались: $B = -0.078$, $p = 0.755$ • Відсоток водіїв, які не мали попередніх аварій: $B = 0.062$, $p = 0.911$.

РОЗДІЛ 3

МЕТОДИЧНІ РЕКОМЕНДАЦІЇ ДО ПРАКТИЧНИХ ЗАНЯТЬ

Використано матеріали джерел [1; 6; 7; 10; 14; 21; 27; 30].

Практичне заняття №1

Тема: «Застосування методів прикладної статистики в страховій математиці»

Контрольні запитання

1. Назвіть дискретні розподіли, які зустрічаються в страховій справі. 2. Назвіть неперервні розподіли, які зустрічаються в страховій справі. 3.

Який розподіл можна використати, якщо ймовірність позову однакова для всіх договорів?

4. Випадкова величина розподілена за експоненціальним розподілом, якщо…
5. Сформулюйте центральну граничну теорему.
6. Які методи оцінки параметрів розподілу ви знаєте?
7. В чому полягає метод моментів?
8. Сформулюйте алгоритм використання методу максимальної правдоподібності.
9. Дайте означення функції правдоподібності.
10. Що називають оцінками максимальної правдоподібності? **Приклади розв'язання задач**

Приклад 1. Знаходження ймовірності пред'явлення позову. Ймовірність пред'явлення позову по 4 незалежним договорам страхування життя в календарному році рівна 0,25. Знайти ймовірність того, що буде пред'явлено 0,1,2,3,4 страхові позови. Визначити ймовірність того, що буде пред'явлено 2 або більше позовів.

Розв'язання:

Маємо 4 незалежних договорів страхування; імовірність позову однакова для всіх договорів і рівна 0,25. Тоді число позовів N має біноміальний розподіл з параметрами n=4 і p=0,25.

Імовірності знайдемо за формулою:

$P(N = m) = C_n^m p^m (1 - p)^{n-m}$, m=0,1,…,n,

$$P(N = 0) = C_4^0 (0,25)^0 (1 - 0,25)^{4-0} = 0,75^4 = 0,32$$

$$P(N = 1) = C_4^1 (0,25)^1 (1 - 0,25)^{4-1} = 4 \cdot 0,25 \cdot 0,75^3 = 0,42$$

$$P(N = 2) = C_4^2 (0,25)^2 (1 - 0,25)^{4-2} = 6 \cdot 0,25^2 \cdot 0,75^2 = 0,21$$

$$P(N = 3) = C_4^3 (0,25)^3 (1 - 0,25)^{4-3} = 4 \cdot$$

$$0{,}25^3 \cdot 0{,}75 = 0{,}045 \quad P(X = 4) = C_4^{\,4}(0{,}25)^4(1 - 0{,}25)^{4-4} = 0{,}25^4 = 0{,}005$$

$$P(X \geq 2) = 0{,}21 + 0{,}045 + 0{,}005 = 0{,}26$$

Відповідь: $P(X = 0) = 0{,}32$, $P(X = 1) = 0{,}42$, $P(X = 2) = 0{,}21$, $P(X = 3) = 0{,}045$, $P(X = 4) = 0{,}005$, $P(X \geq 2) = 0{,}26$.

Приклад 2. Оцінка параметру розподілу методом максимальної правдоподібності.

Певна частина портфеля автотранспортного страхування складається з n ідентичних і незалежних договорів страхування. Позначимо через x_i число позовів, що надійшли (до даного моменту часу) по і-му договору страхування. Процес пред'явлення позовів - пуассонівський з показником λ. Оцінити параметр λ.

Розв'язання:

Складаємо функцію правдоподібності:

$$L(x_1, \ldots, x_n, \lambda) = \prod_{i=1}^{n} e^{-\lambda} \frac{\lambda^{x_i}}{x_i!}$$

Логарифм функції правдоподібності виглядає в такий спосіб:

$$\ln L(\lambda) = (-n\lambda) + \left(\sum_{i=1}^{n} x_i\right) \ln \lambda$$

З умови

$$\frac{d\ln(L)}{d\lambda} = 0$$

Одержуємо

/�� = ��̅ ��̅. ��
^
�� = ∑��₍��₎ ��=1

Відповідь: �� =

Приклад 3. Використання центральної граничної теореми до знаходження ймовірності пред'явлення позовів.

Розглянемо договір страхування життя строком на 10 років. Імовірність пред'явлення позову за цим договором страхування рівна 0,25. Страхова компанія уклала 100 незалежних договорів страхування. Знайти ймовірність того, що до закінчення терміну дії договору страхування (через 10 років) буде пред'явлено точно 30 позовів.

Розв'язання:

Нормальний розподіл - неперервний. Загальне число позовів X, яке буде пред'явлено страховій компанії, належить біноміальному розподілу B(100;0,25), який є дискретним. Таким чином, ми шукаємо наближення дискретного біноміального розподілу за допомогою неперервного нормального розподілу.

Умовно замінимо значення 30 дискретної випадкової величини X інтервалом (29,5;30,5), якому належить випадкова величина Y, що має нормальний розподіл.

Використовуючи математичне сподівання $M(X)=np$ і дисперсію $D(X)=npq$ біноміально розподіленої випадкової величини X, знайдемо, що апроксимуючий нормальний розподіл повинен мати наступні параметри:

$m=np=100\cdot 0,25=25$

$\sigma^2=npq=100\cdot 0,25\cdot 0,75=18,75.$

Тобто, Y належить N(25;18,75).

Згідно центральної граничної теореми:

$P(X=30)=P(29,5<Y<30,5).$

Обчислимо останню ймовірність, переходячи до стандартного нормального розподілу. Використовуючи таблиці для стандартного нормального розподілу (див. додаток Б) знаходимо

$$P(29{,}5 < X < 30{,}5) = P\left(\frac{29{,}5-25}{4{,}33} < Z < \frac{30{,}5-25}{4{,}33}\right) = P(1{,}04 < Z < 1{,}27)$$

$$= \Phi(1{,}27) - \Phi(1{,}04) = 0{,}898 - 0{,}851 = 0{,}047.$$

Точне обчислення P(X=30) дає 0,046.

Відповідь: ймовірність того, що до закінчення терміну дії договору страхування (через 10 років) буде пред'явлено точно 30 позовів рівна 0,046.

Завдання для самостійного розв'язання:

1. Ймовірність пред'явлення страхового позову у зв'язку із смертю по кожному з 4 договорів страхування життя рівна 0,25. Показати, що математичне сподівання числа позовів дорівнює 1.

2. Нехай існує договір страхування життя строком на 10 років. Ймовірність пред'явлення позову по цьому договору страхування дорівнює 0,18. Припустимо, що страхова компанія уклала 75 незалежних договорів страхування. Знайдіть ймовірність того, що до закінчення строку дії договорів страхування буде пред'явлено більше 30 позовів.

3. Число позовів по договорам автотранспортного страхування на протязі календарного року розподілене згідно закону Пуасона з математичним сподіванням рівним 0,2. Існує 100 незалежних договорів страхування. Знайдіть ймовірність того, що на протязі року буде пред'явлено 19, 20, 21 позов. Чи можна в даному випадку застосувати центральну граничну теорему для знаходження апроксимації розподілу?

4. Річна кількість автомобільних аварій без смертельного випадку розподілена згідно розподілу Пуассона з середнім 4,6. Число осіб в

автомобілях, що потрапили в якусь з цих аварій, може бути описана наступним чином:

Кількість осіб в автомобілях	1	2	3	4	5	6	7	8	9	10
Ймовірність	0,1	0,3	0,1	0,1	0,1	0,1	0,05	0,05	0,05	0,05

Кількість робочих днів, втрачених особою, яка потрапила в одну з аварій, може бути описана розподілом Пуассона з середнім 1,5. Знайдіть середнє значення загальної кількості втрачених робочих днів особами, що потрапили на протязі року в аварії без смертельного випадку.

Тестові завдання для самоконтролю:

1. Нормально розподілена випадкова величина ( ) ξзадана щільністю

2

1 + 1 − ( ) x

ξ. Параметри цього розподілу m і σ^2 рівні: $f\ x\ e$

= 2 π 4

А) 1 і

2 Б) -12

і 2 В)

1 і Г) 2(*)

-1 і

2. Якщо імовірність позову однакова для всіх договорів і рівна р, то розподіл позовів є розподілом:

А) біноміальним (*)

Б) Пуассона

В) експоненціальним

Г) нормальним

3. Центральна гранична теорема говорить про те, що випадкова

величина ��– –��

��/√��при досить великому n має розподіл, близький до:

А) розподілу Парето

Б) експоненіального розподілу

В) стандартного нормального розподілу (*)

Г) від’ємно-біноміального розподілу

4. Функцією правдоподібності називається функція вигляду:

А)��(��$_1$, … , ��$_{��}$, ��) = ∑$^{��}_{��=1}$ ��(��$_{��}$, ��)

Б) ��(��$_1$, … , ��$_{��}$, ��) = ∫$^{��}_{��=1}$ ��(��$_{��}$, ��)

В) ��(��$_1$, … , ��$_{��}$, ��) = ∏$^{��}_{��=1}$ ��(��$_{��}$, ��) (*)

Г) ��(��$_1$, … , ��$_{��}$, ��) = �� ∏$^{��}_{��=1}$ ��(��$_{��}$)

5. Метод максимальної правдоподібності полягає в тому, що параметри θ вибираються так, щоб функція правдоподібності набула:

А) найбільшого значення (*)

Б) найменшого значення

В) нульового значення

Г) одиничного значення

6. Випадкова величина Т має гамма-розподіл з параметрами α і δ, якщо щільність рівна:

А) ��(��) = $^{����}$ Γ(α)��$^{��}$e$^{-����}$

Б) ��(��) = $^{����}$ Γ(α)��$^{��}$e$^{����}$

В) ��(��) = $^{����}$ Γ(α)��$^{��-1}$e$^{1-����}$

Г) ��(��) = $^{����}$

Γ(α)�� ��−1 e−���� (*)

7. Рівняння правдоподібності має вигляд:

А) ��������

����= 0

Б) ��������

���� = 0 (*)

В) ��������

���� = 0

Г) ����

���� = 0

8. Оцінками максимальної правдоподібності називають:

А) розв’язки рівняння правдоподібності (*)

Б) параметри функції правдоподібності

В) нульові значення функції правдоподібності

Г) максимальні значення функції правдоподібності

9. В методі моментів статистики ����,�� * =�� 1 �� +⋯+�� �� �� ��є:

А) оцінками параметрів

Б) оцінками моментів (*)

В) оцінками ризиків

Г) оцінками ймовірностей

Практичне заняття №2

Тема: «Моделі індивідуальних позовів»

Контрольні питання

1. Дайте означення дискретної моделі індивідуального позову. 2. Дайте означення функції розподілу дискретної випадкової величини. 3. Які числові характеристики можна знайти, знаючи розподіл дискретної випадкової

величини?

4. Що таке індикатор події? Для чого він використовується? 5. Сформулюйте основні теореми для величини дійсно поданого позову.

6. Як можна перевірити, який розподіл мають дані про величини позовів?

7. Що таке критерій χ^2? Як його застосовують?

Приклади розв’язання задач

Приклад 1. Розглянемо страхування життя на один рік з величиною страхової виплати b=100 000 грн., імовірність смерті застрахованого q=0,0025. Знайти розподіли випадкових величин I та Y.

Розв’язання. Розподілом X є

$$\begin{pmatrix} 0 & 100\,000 \\ 0{,}9975 & 0{,}0025 \end{pmatrix}.$$

Розподілом I є

$$\begin{pmatrix} 0 & 1 \\ 0{,}9975 & 0{,}0025 \end{pmatrix}$$

Тоді

$$P\{Y = 100\,000\} = P\{X = 100\,000 \mid X > 0\} = \frac{P\{X = 100\,000\}}{P\{X > 0\}} = \frac{0{,}0025}{1 - 0{,}9975} = 1$$

Приклад 2. Знаходження розподілів величин структурованої моделі Розглянемо страхування життя на один рік з величиною страхової виплати b_1=500 000 грн. у випадку смерті від нещасного випадку (ймовірність цієї події q_1=0,0004) і виплатою b_2=100 000 грн. у випадку смерті внаслідок природних причин (ймовірність цієї події q_2=0,0020). Знайти розподіли величин I та Y.

Розв’язання. Розподілом X є

$$\begin{pmatrix} 0 & 100\,000 & 500\,000 \\ 0{,}9976 & 0{,}0020 & 0{,}0004 \end{pmatrix}$$

Знайдемо розподіл *I*.

��{�� = 1} = ��{�� > 0} = ��{�� = 500 000} + ��{�� = 100 000} = 0,0024, ��{�� = 0} = ��{�� = 0} = 0,9976,

тобто розподілом *I* є

$$\begin{pmatrix} 0 & 1 \\ 0{,}9976 & 0{,}0024 \end{pmatrix}$$

Тоді

��{�� = 500 000} = ��{�� = 500 000| �� > 0} = ��{�� = 500 000} / ��{ �� > 0} = 1/6

��{�� = 100 000} = ��{�� = 100 000| �� > 0} = ��{�� = 100 000} / ��{ �� > 0} = 5/6

отже, розподілом *Y* є

$$\begin{pmatrix} 100\,000 & 500\,000 \\ 5/6 & 1/6 \end{pmatrix}$$

Приклад 3. Статистичний аналіз величин дійсно поданих позовів за портфелем договорів дає підстави стверджувати, що коли Y – величина дійсно поданого позову, то величина �� = ln �� має нормальний розподіл із середнім 6,012 та дисперсією 1,792. Обчислити ймовірність того, що величина дійсно поданого позову знаходиться в межах від 200 до 500.

Розв'язання:

��{200 < �� < 500} = ��{ln 200 < �� < ln 500}

= �� ((ln 200 − ����) / √���� < (�� − ����) / √���� < (ln 500 − ����) / √����

√����)

= ��$_{0;1}$ (ln 500 − ����

√����) − ��$_{0;1}$ (ln 200 − ����

√����) = 0,263

Відповідь: ��{200 < �� < 500} = 0,263.

Завдання для самостійного розв'язання

1. Імовірність виникнення пожежі в деякому будинку протягом заданого проміжку часу становить 0,02. Якщо пожежа виникає, то величина пошкодження *Y*, завданого будинку, рівномірно розподілена на інтервалі від 0 до її максимально можливого значення *a*. Обчислити математичне сподівання і дисперсію величини пошкодження *X*, завданого пожежею будинку протягом заданого проміжку часу.

2. Розподіл величини індивідуального позову *X* для договорів страхування складу від пожеж дорівнює:

Величина	0	500	1000	10000	50000	100000
Ймовірність	0,9	0,06	0,03	0,008	0,001	0,001

Знайти розподіли величин *I* та *Y*. Обчислити середню величину дійсно поданого позову.

3. У невеликому приморському містечку річні витрати від штормів, пожеж та розкрадання майна є незалежними експоненціально розподіленими

випадковими величинами із середніми 1; 1,5; 2,4 відповідно. Обчислити ймовірність того, що максимальна величина із зазначених збитків перевищує 3.

4. Далі наведені дані про число позовів, поданих за один день протягом року до страхової компанії:

Число позовів за день	Число днів

0	50
1	122
2	101
3	92
≥4	0

Застосовуючи критерій χ^2, перевірити гіпотезу про пуассонів розподіл числа позовів, поданих за один день протягом року.

Тестові завдання для самоконтролю

1. Дискретною моделлю індивідуальних позовів називається випадкова величина, яка набуває:

А) скінченного числа значень

Б) нескінченного числа значень

В) зліченного числа значень

Г) скінченного або зліченного числа значень (*)

2. Розподілом дискретної випадкової величини X називатимемо функцію ��$_{��}$: ��$_{��}$ → ��$_{��}$(��$_{��}$), ��$_{��}$ ∈ �� ⊂ ��^{1}, що ставить у відповідність кожному можливому значенню ��$_{��}$ ∈ �� ймовірність:

А) ��$_{��}$(��$_{��}$) = ��{�� ≥ ��$_{��}$}

Б) ��$_{��}$(��$_{��}$) = ��{�� = ��$_{��}$} (*)

В) ��$_{��}$(��$_{��}$) = ��{�� ≤ ��$_{��}$}

Г) ��$_{��}$(��$_{��}$) = ��{�� < ��$_{��}$}

3. Індикатор події визначається як:

А) �� = { 1, �� > 0
0, �� = 0 (*)

Б) �� = { 1, �� < 0
0, �� = 0

В) �� = { 1, �� = 0
0, �� > 0

Г) �� = {1, �� = 0
0, �� < 0

4. Індикатор вводиться для:

А) характеристики величини збитків

Б) реєстрації позовів

В) оцінки ймовірності настання страхового випадку

Г) оцінки залежності ймовірності банкрутства від капіталу компанії. 5. Величина *Y* реально поданого позову, коли страховий випадок відбувся визначається як:

А) ��{�� = ����} = ��{�� = ����|�� > 0} (*)

Б) ��{�� = ����} = ��{�� = ����|�� < 0}

В) ��{�� = ����} = ��{�� = ����|�� ≥ 0}

Г) ��{�� = ����} = ��{�� = ����|�� ≤ 0}

6. При перевірці гіпотези про тип розподілу сукупності визначаються частоти:

А) загальні та особливі

Б) емпіричні та критичні

В) емпіричні і теоретичні (*)

Г) статичні і динамічні

7. Критерій χ^2обчислюється за формулою:

А) ��² = ∑(�� ��������–�� ������)²
�� �� ��������
=1 �����

Б) ��² = ∑(�� ������−�� ��������)²
�� �� ��������
=1 ���

В) ��² = ∑(�� ������+�� ��������)²
�� �� ��������
=1 �����

Г) ��² = ∑(�� ������−�� ��������)²
�� ��=1(*)
�� ��������

8. Згідно критерію χ^2нульову гіпотезу приймають, якщо:

А) ��$^2 >$ ��$_{кр}2$

Б) ��$^2 <$ ��$_{кр}{}^2$(*)

В) ��$^2 =$ ��$_{кр}2$

Г) ��$^2 \geq$ ��$_{кр}2$

9. В моделях ризику умовний розподіл використовується для:
А) оцінки числа позовів
Б) оцінки величини збитку
В) оцінки величини резерву
Г) оцінки ймовірності банкрутства

10. Модель �� $=$ ���� індивідуального позову називається неперервною, якщо:

А) величина *X* є абсолютно неперервною

Б) величина *Y* є абсолютно неперервною (*)

В) величина *I* є абсолютно неперервною

Г) величина *Z=lnX* є абсолютно неперервною

Практичне заняття №3

Тема: «Модель індивідуального ризику»

Контрольні питання

1. Назвіть припущення на яких базується модель індивідуального ризику.

2. Що таке сумарний позов?

3. Що таке резервний капітал компанії?

4. Як визначається ймовірність розорення компанії в моделі індивідуального ризику?

5. Що таке згортка розподілів? Для чого вона використовується?

6. Коли використовують центральну граничну теорему?

7. Як визначити ймовірність розорення компанії, використовуючи центральну граничну теорему?

8. Дайте означення генератриси розподілу

9. В чому полягає метод генератрис?

Приклади розв'язання задач

Приклад 1

Портфель компанії складається з 4 однакових договорів страхування життя. Якщо смерть застрахованого настала через нещасний випадок, то спадкоємцям виплачують 500000 грн. Якщо смерть застрахованого наступила природно, то страхова виплата дорівнює 250000 грн. у певній віковій категорії. Імовірність природної смерті дорівнює 0,1, а ймовірність смерті від нещасного випадку – 0,001. Знайти залежність імовірності банкрутства R від величини капіталу компанії u.

Розв'язання:

1 спосіб

Для розрахунків зручно прийняти 250000 грн. за одиницю виміру грошових сум. Тоді розподіл індивідуального позову X_i за i-им договором (i=1,2,3,4) відомий і дорівнює

0	1	2
0,899	0,1	0,001

причому X_1, X_2, X_3, X_4 – незалежні випадкові величини. Розподіл суми X_1+X_2 індивідуальних позовів знаходимо як згортку розподілів. Спільний розподіл (X_1, X_2) має такий вигляд:

X1/X2	0	1	2
0	0,808201	0,0899	0,000899
1	0,0899	0,01	0,0001
2	0,000899	0,0001	0,000001

$$P\{X_1 = 0, X_2 = 0\} = P\{X_1 = 0\} \cdot P\{X_2 = 0\} = 0{,}899 \cdot 0{,}899 = 0{,}808201 \quad P\{X_1 = 0, X_2 = 1\} = P\{X_1 = 0\} \cdot$$

$P\{X_2 = 1\} = 0{,}899 \cdot 0{,}1 = 0{,}0899$ Звідси розподіл суми X_1+X_2:

0	1	2	3	4
0,808201	0,1798	0,011798	0,0002	0,000001

$$P\{X_1 + X_2 = 0\} = P\{X_1 = 0, X_2 = 0\} = 0{,}808201$$

$$P\{X_1 + X_2 = 1\} = P\{X_1 = 0, X_2 = 1\} + P\{X_1 = 1, X_2 = 0\} = 0{,}0899 + 0{,}0899 = 0{,}1798$$

Аналогічно знаходимо розподіл суми X_1+X_2+X_3.

Спочатку знаходимо спільний розподіл (X_1+X_2,X_3):

X1+X2/X3	0	1	2
0	0,726572699	0,0808201	0,000808201
1	0,1616402	0,01798	0,0001798
2	0,010606402	0,0011798	0,000011798
3	0,0001798	0,00002	0,0000002
4	0,000000899	0,0000001	0,000000001

Тоді розподіл суми X_1+X_2+X_3:

0	1	2	3	4	5	6
0,726572699	0,24246	0,029395	0,001539	0,000032697	0,0000003	0,000000001

Спільний розподіл X_1+X_2+X_3+X_4 дорівнює

X1+X2+X3/X4	0	1	2
0	0,653188856	0,0726573	0,000726573
1	0,21797181	0,024246	0,00024246
2	0,026425748	0,0029395	2,93946E-05
3	0,001383921	0,0001539	1,5394E-06
4	2,93946E-05	3,27E-06	3,2697E-08

5	2,697E-07	3E-08	3E-10
6	8,99E-10	1E-10	1E-12

Розподіл суми $X_1+X_2+X_3+X_4$

0	1	2	3	4	5	6	7	8
0,653188	0,290629	0,051398	0,004566	0,000212	5E-06	6E-08	4E-10	1E-12

Отже, залежність ймовірності банкрутства R від капіталу компанії u дорівнює ��(��) = ��{��1 + ��2 + ��3 + ��4 > ��} і задається таблицею

u	R(u)
0	0,346811
1	0,056182
2	0,004783
3	0,000217
4	5,14E-06
5	6,39E-08
6	4,01E-10
7	1E-12
8	0

2 спосіб.

Застосуємо генератриси для знаходження розподілу суми $X_1+X_2+X_3+X_4$, а саме кожен індивідуальний позов за і-им договором (і=1,2,3,4) має одну й ту ж генератрису:

��$_1$(��) = ��$_2$(��) = ��$_3$(��) = ��$_4$(��) = 0,899��^{0} + 0,1��^{1} + 0,001��^{2} Тоді генератриси суми $X_1+X_2+X_3+X_4$ дорівнює

��(��) = ��$_1$(��)��$_2$(��)��$_3$(��)��$_4$(��) = (0,899��^{0} + 0,1��^{1} + 0,001��^{2})^4 = 0,6532 + 0,2906�� + 0,0514��^{2} + 4,57 · 10^{-3}��^{3} + 2,13 · 10^{-4}��^{4} + 5,08 · 10^{-6}��^{5} + 6,36 · 10^{-8}��^{6} + 4 ·

10^{-10}��$^{7} + 10^{-12}$��^{8}

Коефіцієнт при ��$^{��}$ дорівнює ймовірності

��$_{��=1}$}, n=0,1,…,8.

��{∑ ��$_{��}$ = ��

Отже, залежність ймовірності банкрутства R від капіталу компанії u задається таблицею

u	R(u)
0	0,3468
1	0,0562
2	0,0048
3	$2,18\cdot10^{-4}$
4	$5,14\cdot10^{-6}$
5	$6,4\cdot10^{-8}$
6	$4,01\cdot10^{-10}$
7	10^{-12}
8	0

Приклад 2.

У страховій компанії застраховано N=3000 осіб з імовірністю смерті протягом року 0,003. Компанія виплачує нащадкам суму 250000 грн. у випадку смерті застрахованого протягом року і не платить нічого, якщо особа доживе до кінця року. Визначити мінімальну величину капіталу, за якого ймовірність банкрутства компанії не перевищує 0,05.

Розв'язання:

Візьмемо величину 250000 грн. за одиницю виміру грошових сум. Тоді індивідуальний позов X_i до компанії має розподіл

0	1
0,997	0,003

при цьому $MX_i = 0{,}003$, $DX_i = 0{,}002991$ та $X_1, X_2, \ldots, X_{3000}$ – незалежні випадкові величини. Тоді математичне сподівання *MS* та дисперсія *DS* сумарного позову $S=X_1+X_2+\ldots+X_{3000}$ до компанії дорівнюють:

$$MS = M(X_1 + X_2 + \cdots + X_{3000}) = 3000 \cdot 0{,}003 = 9$$

$$DS = D(X_1 + X_2 + \cdots + X_{3000}) = 3000 \cdot 0{,}002991 = 8{,}973$$

Знайдемо величину капіталу компанії, при якому ймовірність банкрутства компанії дорівнює 0,05. Маємо

$$0{,}05 = P\{S > u\} = 1 - \Phi_{0;1}\left(\frac{u - MS}{\sqrt{DS}}\right)$$

$$\Phi_{0;1}\left(\frac{u - 9}{\sqrt{8{,}973}}\right) = 0{,}95$$

$$\frac{u - 9}{\sqrt{8{,}973}} = 1{,}65$$

Отже, мінімальний капітал компанії має бути не меншим *u=13,94* або 3485642,4 грн.

Приклад 3.

Нехай *n*=1000 людей придбали поліси страхування життя строком на 1 рік. Імовірність смерті застрахованого *p*=0,001, виплата в разі його смерті

b=1. Скориставшись різними видами апроксимації, знайдіть імовірність того, що сумарні виплати для такого портфелю полісів будуть більшими за 3,5. Підрахуємо математичне сподівання і дисперсію сумарних виплат:

$$MS = nMX_i = np = 1000 \cdot 0{,}001 = 1$$

$$DS = np(1 - p) = 1000 \cdot 0{,}001 \cdot 0{,}999 = 0{,}999 \approx 1$$

Нормальна апроксимація

$$F_S(x) \approx \Phi\left(\frac{x - MS}{\sqrt{DS}}\right) = \Phi(x - 1)$$

Тоді шукана ймовірність:

$$P(S_N > 3{,}5) = 1 - P(S_N \le 3{,}5) \approx 1 - \Phi(3{,}5 - 1) = 1 - \Phi(2{,}5) = 1 - 0{,}9938 = 0{,}0062$$

Пуассонівська апроксимація

$$\lambda = np = 1000 \cdot 0{,}001 = 1$$

$$P(S_N > 3{,}5) = 1 - P(S_N \le 3{,}5) \approx 1 - p_0 - p_1 - p_2 - p_3$$

$$p_k = \frac{(np)^k}{k!} e^{-np}$$

$$P(S_N > 3{,}5) \approx 1 - e^{-1} - e^{-1} - \frac{1}{2}e^{-1} - \frac{1}{6}e^{-2} \approx 0{,}01899$$

Зауваження. У даному прикладі сумарні виплати мають біноміальний розподіл з параметрами *n*=1000 та *p*=0,001. Шукана ймовірність для біноміального розподілу дорівнює 0,01893. Отже, у даному випадку пуасонівська апроксимація значно точніша за нормальну.

Г-апроксимація. Щоб скористатися формулами (1.2.7) підрахуємо коефіцієнт асиметрії

$$\gamma_S = \frac{E(S_N - ES_N)^3}{\sigma_S^3}$$

$ES = 1$, $DS = 0{,}999$ – знайдені вище. У даному випадку

$$E(S_N - ES_N)^3 = \frac{d^3}{dt^3} M_S(t)\Big|_{t=0} = \frac{d^3}{dt^3} \log(pe^t + 1 - p)^n\Big|_{t=0},$$

де $M_S(t)$ – генератриса моментів для біноміального розподілу.

Виконавши всі необхідні дії, маємо

$$E(S_N - ES_N)^3 = np(1 - 3p + 2p^2) = 1000 \cdot 0{,}01(1 - 0{,}003 + 2 \cdot 10^{-6}) \approx 0{,}997$$ Тоді

$$\gamma_S \approx \frac{0{,}997}{\sqrt{0{,}999}^3} \approx 1$$

Значення параметрів Г-розподілу знаходимо з (1.2.7) при

���� ≈ ��$_{��}$ ≈ ��$_{��}$ ≈ 1

Звідси �� = 4, �� = 2, ��$_0$ = −1.

Отже ��(��$_{��}$ > 3,5) ≈ 1 − ��(3,5 + 1; 4,2) ≈ 1 − ��(4,5; 4,2) ≈ 0,0212 Бачимо, що в даному випадку Г-апроксимація дає трохи гірше наближення до точного значення, ніж пуассонівська, проте значно краще, ніж нормальна.

Завдання для самостійного розв'язання:

1. Портфель компанії складається з 4 однакових договорів страхування життя. Якщо смерть застрахованого настала внаслідок нещасного випадку, то нащадкам виплачують 500000 грн. Якщо смерть застрахованого настала в результаті природних причин, то страхова виплата дорівнює 250000 грн. у певній віковій категорії. Для кожного із застрахованих імовірність смерті через нещасний випадок дорівнює 0,1, ймовірність смерті у зв'язку з природними причинами дорівнює 0,1, і, отже, ймовірність дожиття дорівнює 0,8. Знайти залежність імовірності банкрутства R від величини капіталу компанії u: а) за методом згорток; б) методом генератрис.

2. Портфель компанії складається з двох договорів страхування будівель від пожежі. Вартість першої будівлі c_1=1 млн.грн., а другої – c_2=2 млн.грн. Імовірність пожежі на першому об'єкті за проміжок часу, який розглядається – q_1=0,2, на другому – q_2=0,1, а збитки від пожежі (Y_1, Y_2), якщо вона має місце, рівномірно розподілені від 0 до повної вартості об'єкта. Знайти залежність імовірності банкрутства компанії від капіталу компанії.

3. Портфель страхової компанії складається з 16000 договорів страхування життя строком на 1 рік. Далі наведена кількість застрахованих n_k

у кожній з 5 груп з різними значеннями страхової виплати b_k і ймовірності позову q_k:

k	q_k	b_k	n_k
1	0,01	10000	8000
2	0,02	20000	3500

3	0,03	30000	2500
4	0,05	50000	1500
5	0,1	100000	500

Визначити мінімальну величину капіталу, за наявності якого ймовірність банкрутства компанії не перевищує 0,05.

4. Нехай 5000 людей придбали поліси страхування життя строком на 1 рік. Імовірність смерті застрахованого 0,001, виплата в разі його смерті дорівнює 1. Скориставшись різними видами апроксимації, знайдіть імовірність того, що сумарні виплати для такого портфелю полісів будуть більшими за 5.

Тестові завдання для самоконтролю:

1. В моделі індивідуального ризику плата за страховку здійснюється: А) на початку періоду (*)

Б) в кінці періоду

В) і на початку, і в кінці періоду

Г) на протязі періоду

2. Ймовірність банкрутства компанії дорівнює:

А) $�� = ��\{\sum_{��=1}^{��} ��_{��} > 0\}$

Б) $�� = ��\{\sum_{��=1}^{��} ��_{��} < 0\}$

В) $�� = ��\{\sum_{��=1}^{��} ��_{��} > ��\}$ (*)

Г) $�� = ��\{\sum_{��=1}^{��} ��_{��} < ��\}$

3. В моделі індивідуального ризику число договорів страхування є:

А) випадковим

Б) незалежним

В) фіксованим (*)

Г) змінним

4. В моделі індивідуального ризику позови розглядаються як:

А) позови по окремим договорам (*)

Б) один сумарний позов

В) окремі позови, а потім як згруповані позови

Г) позови, що згруповані в залежності від умов настання страхового випадку

5. Величини позовів по окремим договорам в моделі індивідуального ризику є:

А) нульовими

Б) незалежними (*)

В) однаковими

Г) залежними

6. В моделі індивідуального ризику виключається можливість страхового випадку, що призводить:

А) до позовів, що взаємопов'язані між собою

Б) до позовів, великих за своєю величиною

В) до позовів одразу по кільком договорам

Г) до позовів, спричинених нещасними випадками

7. Якщо випадкові величини *X₁,X₂* незалежні, то генератриса суми випадкових величин дорівнює:

А) $P_{X_1+X_2}(z) = P_{X_1}(z) + P_{X_2}(z)$

Б) $P_{X_1+X_2}(z) = P_{X_1}(z) - P_{X_2}(z)$

В) $P_{X_1+X_2}(z) = P_{X_1}(z) / P_{X_2}(z)$

Г) $P_{X_1+X_2}(z) = P_{X_1}(z) \cdot P_{X_2}(z)$ (*)

8. Розподіл суми $\sum_{i=1}^{n} X_i$ можна обчислити як згортку розподілів цих випадкових величин, якщо величини $X_1, X_2, \dots, X_n$:

А) фіксовані

Б) змінні

В) залежні

Г) незалежні (*)

9. Апроксимаційна формула при нормальній апроксимації має вигляд:

А) $��_{��}(��) \approx \Phi_{0,1}(��)\left(\frac{��-����}{\sqrt{����}}\right)$ (*)

Б) $��_{��}(��) \approx \Phi_{0,1}(��)\left(\frac{��+����}{\sqrt{����}}\right)$

В) $��_{��}(��) \approx \Phi_{0,1}(��)\left(\frac{��-����}{\sqrt{����}}\right)$

Г) $��_{��}(��) \approx \Phi_{0,1}(��)\left(\frac{��+����}{\sqrt{����}}\right)$

10.Застосування пуассонівської апроксимації має сенс, якщо ймовірність настання страхової події:

А) велика

Б) нульова

В) рівна 1

Г) мала (*)

Практичне заняття №4

Тема: «Модель колективного ризику»

Контрольні питання

1. Назвіть припущення на яких базується модель колективного ризику. 2. Як обчислити ймовірність банкрутства в моделі колективного ризику, якщо позови є дискретними випадковими величинами? 3. Як обчислити ймовірність банкрутства в моделі колективного ризику, якщо позови є абсолютно неперервними випадковими величинами? 4. Як знайти математичне сподівання сумарного позову?

5. Чому дорівнює дисперсія сумарного позову?

6. Яка формула теорії ймовірностей застосовується при знаходженні

ймовірності банкрутства в моделі колективного ризику?

Приклади розв’язання задач

Приклад. Знаходження ймовірності банкрутства від капіталу компанії

Нехай за фіксований короткий проміжок часу портфель договорів страхової компанії може породити 0, 1, 2, 3 позови з ймовірностями 0,2; 0,3; 0,4; 0,1 відповідно. У випадку, якщо позов поданий, його величина дорівнює 1, 2, 3 (умовних грошових одиниць) з ймовірностями 0,5; 0,4; 0,1 відповідно. Визначити залежність ймовірності банкрутства від капіталу компанії. *Розв’язання:*

Спочатку знайдемо розподіли випадкових величин $��_1$, $��_1 + ��_2$, $��_1 + ��_2 + ��_3$. Кожна випадкова величина $��_{��} (�� = 1,2,3)$ має розподіл:

1	2	3
0,5	0,4	0,1

Спільний розподіл $��_1$, $��_2$:

$��_1$	$��_2$		
	1	2	3
1	0,25	0,2	0,05
2	0,2	0,16	0,04
3	0,05	0,04	0,01

Тоді розподіл суми $��_1 + ��_2$ має вигляд:

1	2	3	4	5	6
0	0,25	0,4	0,26	0,08	0,01

Спільний розподіл $(��_1 + ��_2, ��_3)$ дорівнює:

$��_3$	$��_1 + ��_2$					
	1	2	3	4	5	6
1	0	0,125	0,2	0,13	0,04	0,005
2	0	0,1	0,16	0,104	0,032	0,004
3	0	0,025	0,04	0,026	0,008	0,001

Тоді розподіл суми $(��_1 + ��_2 + ��_3)$ наступний:

1	2	3	4	5	6	7	8	9
0	0	0,125	0,3	0,315	0,184	0,063	0,012	0,001

Розподіл числа позовів *v* до страхової компанії:

0	1	2	3
0,2	0,3	0,4	0,1

Знайдемо розподіл сумарного позову $��_{��} = ��_1 + ��_2 + \cdots + ��_{��}$ за формулою:

$$\infty_{��=1} = \sum ��\{��_1 + ��_2 + \cdots + ��_{��} = ��\}��\{�� =$$

$$\infty_{��=1} \quad \sum ��\{��_{��} = ��\}��_{��}$$

$$��\}. \quad ��\{��_{��} = ��\} =$$

Звідси одержимо:

$$��\{��_{��} = 0\} = ��\{�� = 0\} = 0,2$$

$$��\{��_{��} = 1\} = ��\{��_1 = 1\}��\{�� = 1\} + ��\{��_1 + ��_2 = 1\}��\{�� = 2\} + ��\{��_1 + ��_2 + ��_3 = 1\}��\{�� = 3\} = 0,5 \cdot 0,3 + 0 \cdot 0,4 + 0 \cdot 0,1 = 0,15$$

$$��\{��_{��} = 2\} = ��\{��_1 = 2\}��\{�� = 1\} + ��\{��_1 + ��_2 = 2\}��\{�� = 2\} + ��\{��_1 + ��_2 + ��_3 = 2\}��\{�� = 3\}$$

$$= 0,4 \cdot 0,3 + 0,25 \cdot 0,4 + 0 \cdot 0,1 = 0,22$$

$$��\{��_{��} = 3\} = 0,1 \cdot 0,3 + 0,4 \cdot 0,4 + 0,125 \cdot 0,1 = 0,2025$$

$$��\{��_{��} = 4\} = 0 \cdot 0,3 + 0,26 \cdot 0,4 + 0,3 \cdot 0,1 = 0,134$$

$$��\{��_{��} = 5\} = 0 \cdot 0,3 + 0,08 \cdot 0,4 + 0,315 \cdot 0,1 = 0,0635$$

$$��\{��_{��} = 6\} = 0 \cdot 0,3 + 0,01 \cdot 0,4 + 0,184 \cdot 0,1 = 0,0224$$

$$��\{��_{��} = 7\} = 0 \cdot 0,3 + 0 \cdot 0,4 + 0,063 \cdot 0,1 = 0,0063$$

$$��\{��_{��} = 8\} = 0 \cdot 0,3 + 0 \cdot 0,4 + 0,012 \cdot 0,1 = 0,0012$$

$$��\{��_{��} = 9\} = 0 \cdot 0,3 + 0 \cdot 0,4 + 0,001 \cdot 0,1 = 0,0001$$

Отже, розподіл сумарного позову $��_{��}$ дорівнює:

k	$��\{��_{��} = ��\}$
0	0,2
1	0,15
2	0,22
3	0,2025
4	0,134
5	0,0635
6	0,0224
7	0,0063
8	0,0012
9	0,0001

Знаючи розподіл $��_{��}$, знайдемо ймовірність банкрутства *R(u)*:

$$��(��) = ��\{��_{��} > ��\} = \sum_{��=��+1}^{\infty} ��\{��_{��} = ��\}$$

та подамо значення *R(u)* у таблиці:

u	*R(u)*
0	0,8

1	0,65
2	0,43
3	0,2275
4	0,0935
5	0,03
6	0,0076
7	0,0013
8	0,0001
9	0

Завдання для самостійного розв'язання

1. За фіксований короткий проміжок часу портфель договорів страхової компанії може породити 0 ,1, 2, 3 позови з ймовірностями 0,1; 0,3; 0,4; 0,2 відповідно. У випадку, якщо позов поданий, його величина дорівнює 1, 2, 3 з ймовірностями 0,5; 0,4; 0,1 відповідно. Визначити залежність імовірності банкрутства від капіталу компанії.

2. Нехай число позовів *v* за фіксований проміжок часу має зміщений геометричний розподіл з параметром *p*:

$$p_n = P\{\nu = n\} = p(1-p)^{n-1}$$

а величини позовів розподілені показниково з параметром λ:

$$F_{Y_i}(x) = P\{Y_i < x\} = \begin{cases} 0, & x < 0 \\ 1 - e^{-\lambda x}, & x \geq 0 \end{cases}$$

Знайти залежність ймовірності банкрутства від капіталу компанії методом згорток.

3. Величина позову в разі його подання до страхової компанії дорівнює 1, 2, 3 з ймовірностями 0,5; 0,4; 0,1 відповідно. Число позовів до страхової компанії має пуассонів розподіл із середнім 1,7. Знайти розподіл сумарного

позову до компанії.

Тестові завдання для самоконтролю

1. В моделі колективного ризику аналізується такий проміжок часу, що:

А) обов’язково враховує інфляцію

Б) дозволяє не враховувати доходи компанії

В) обов’язково враховує рівень доходів клієнтів

Г) дозволяє не враховувати інфляцію та доходи від інвестування (*)

2. В моделі колективного ризику плата за страховку здійснюється:

А) на початку періоду (*)

Б) в кінці періоду

В) і на початку, і в кінці періоду

Г) на протязі періоду

3. В моделі колективного ризику позови розглядаються як:

А) позови по окремим договорам

Б) один сумарний позов (*)

В) окремі позови, а потім як один єдиний позов

Г) позови, що згруповані в залежності від умов настання страхового випадку.

4. В якості основної характеристики портфеля в моделі колективного ризику виступає:

А) кількість укладених договорів

Б) середня кількість позовів, що надійшли за період

В) загальна кількість позовів, що надійшли за період (*)

Г) середня кількість укладених договорів

5. В моделі колективного ризику розорення визначається:

А) сумарним позовом до страхової компанії

Б) кількістю укладених договорів

В) середньою кількістю позовів, що надійшли до страхової компанії Г) співвідношенням між кількістю укладених договорів та кількістю позовів, що надійшли до страхової компанії

6. Якщо випадкові величини Y_i (реально подані позови)– дискретні, то ймовірність розорення страхової компанії визначається формулою:

∞ ��= ∞ ��=
��+1 1

А) �� = ∑ ∑ ��{�� = ��}

∞ ��= ∞ ��=
��+1 1

Б) �� = ∑ ∑ ��{��$_{��}$ = ��}

∞ ��=��+1 ��{�� = ��}
∞ ��=1

В) �� = ∑ ∑ ��{��$_{��}$ = ��}

∞ ��=��+1 ��{�� = ��}
∞ ��=1

Г) �� = ∑ ∑ ��{��$_{��}$ = ��}

7. Якщо випадкові величини Y_i – абсолютно неперервні, то ймовірність розорення страхової компанії визначається формулою:

∞
∞ ��=1)

А) �� = ∫ (∑ ��{��$_{��}$ =

��}��$_{����}$(��)

$_{��}$����

∞
∞ ��=1)

Б) �� = ∫ (∑ ��{�� =

��}��$_{����}$(��)

$_{��}$����

∞
∞ ��=1)

В) �� = ∫ (∑ ��{�� =

��}��$_{��}$(��)

$_{��}$����

∞ ∞ ��=1)

Г) �� = ∫ (∑ ��{�� = ��}������(��)

������

При розробці методичних рекомендацій використані матеріали посібників [1; 6;10; 14; 21; 27; 30].

ВИСНОВКИ

У результаті проведених теоретичних досліджень та розробки методичних рекомендацій до практичного модуля дисципліни «Актуарна математика», можна зробити наступні висновки:

1. Систематизовано та узагальнено теоретичний матеріал з теми ризиків банкрутства страхової компанії.
2. Ґрунтовно розглянуто, досліджено та проаналізовано методи розрахунку ймовірності банкрутства в моделях індивідуального і колективного ризику.
3. Застосовано основні ймовірнісні методи оцінки ризику до розв’язання практичних задач.
4. Засобами комп’ютерного статистичного пакету досліджено фактори, які спричиняють ризики смертельних зіткнень у штатах США та відповідні страхові збитки, також здійснено кластерний аналіз групування штатів США за характерними показниками аварійності.
5. Розроблено методичні рекомендації до практичного модуля дисципліни «Актуарна математика» за темами:
   - Застосування методів прикладної статистики в страховій математиці
   - Моделі індивідуальних позовів
   - Модель індивідуального ризику
   - Модель колективного ризику.

До кожної частини модуля наведено приклади розв’язання типових задач з детальним поясненням, підібрано систему вправ для самостійного розв'язання, розроблені тестові завдання.

Отримані результати можуть бути використані викладачами та студентами ВНЗ математичних та економічних спеціальностей, що вивчають дисципліни страхової математики

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Базилевич В.Д. Страхування: Підручник (за ред.В. Д. Базилевича). Київ:Знання, 208. 1019 с.

2. Базилевич В.Д., Базилевич К.С. Страхова справа. Київ: Т-во «Знання», 2003.- 250с.

3. Баранов А.И. Управління страховим портфелем. Вісник Київського національного університету ім.Т.Шевченка.2007.(№94-95). С.78- 135.

4. Бахрушин В.Є. Методи аналізу даних : навчальний посібник для студентів / В.Є.Бахрушин. – Запоріжжя : КПУ, 2011. – 268 с. ISBN 978-966-414-103-8

5. Бірд, Р. Теорія ризику: Стохастична основа страхування. 3-тє видання. Берлін: Springer, 1984. 408 с.

6. Бондаренко Я. С. Теорія ризику у страхуванні. Основні поняття, приклади, задачі: навч. посібник./ Я.С.Бондаренко, В.М.Турчин, Є.В.Турчин. – Д.: РВВ ДНУ, 2010 – 180 с.

7. Бондаренко Я. С. Теорія ризику в страхуванні: навч. посіб./ Я. С. Бондаренко, В.М. Турчин, Є.В. Турчин. – Д.:РВВ ДНУ, 2008 – 112 с 8. В.О.Кофанов. Основи актуарної математики // ДНУ: Дніпропетровськ – 2005. – 96 с.

9. Волков Ю. І., Войналович Н. М. Початки стохастики: Навчальний посібник. – Кіровоград, 2008. – 168 с.

10. Гербер, Ганс У. Вступ до математичної теорії ризику. Перше видання. Іллінойс: R. D. Irwin, 1979. 164 с.

11. Гербер, Ганс У. Математика страхування життя. Берлін: Springer Berlin Heidelberg, 1997.

12. Джерроу, Роберт А. Економічні основи управління ризиками: Теорія, практика та застосування. Сінгапур: World Scientific, 2016.

13. Дрібноход А.О. Підхід до моделювання ризику банкрутства у страхуванні. Науковий журнал. 2012. (№9).

14. Зінченко Н. М. Математичні методи в теорії ризику. – К.: ВПЦ "Київський університет", 2008. – 213 с.

15. Кінаш О.М., Сороківський В.М., Папка М.В. Основи актуарних розрахунків: навч.-метод.посіб. Львів, 2012. 188с. (вступ про страх) 16. Ковтун І.О., Денисенко М.П., Кабанов В.Г. Основи актуарних розрахунків:навчальний посібник. Київ : «ВД «Професіонал», 2008. 480 с 17. Конет І.М. Теорія ймовірностей та математична статистика в прикладах і задачах.– Кам'янець–Подільський: Абетка, 2011. – 220 с. 18. Кравець П., Пасічник В., Проданюк М. Математична модель логістичної регресії для бінарної класифікації. Ч.1. Регресійні моделі узагальнення даних. - INFORMATION SYSTEMS AND NETWORKS Issue 15, 2024

19. Крамер, Гаральд. Математичні методи статистики. Нью-Йорк: Princeton University Press, 1946. (перше видання: 1922).

20. Кушлик-Дивульська О.І., Поліщук Н.В., Орел Б.П., Штабалюк П.І. Теорія ймовірностей та математична статистика. Навчальний посібник – К.: НТУУ «КПІ», 2014.–212 с.

21. Леоненко М.М., Мішура Ю.С., Пархоменко В.М. Ядренко М.Й. Теоретико-ймовірнісні та статистичні методи в економетриці та фінансовій математиці. Київ, 1995, 379с.

22. Лупан І.В., Авраменко О.В., Акбаш К.С. Комп'ютерні статистичні пакети: навчально-методичний посібник. – 2-е вид. – Кіровоград: «КОД», 2015. – 236 с.

23. Пономаренко О.І. Основи математики фінансів і страхування. Київ : ІВЦДержкомстату України, 2004. 256 с.

24. Пономаренко О.І. Фінансовий аналіз. Вип.1,2. – К.: ЕМЦ, 2001.

25. Пономаренко О. І., Ковальчук Ю. О. Актуарна математика: навчальний посібник. – Ніжин: Видавництво НДУ ім. М. Гоголя, 2008. – 182 с.

26. Роженко Н. О. Модель індивідуальних ризиків: метод. посіб. Одеса.: Одеський національний університет, 2012. 34с.

27. Сотник К.О. Математичне моделювання страхових ризиків: оцінка ймовірності банкрутства. - №13 (2024): Наукові записки молодих учених. – Кропивницький: РВВ ЦДУ ім В. Винниченка ISSN 2617-2666

28. Сотник К.О. «Застосування логістичної регресії до оцінювання ризику банкрутства страхової компанії» - Збірник тез. XVI Всеукраїнська студентська наукова конференція «Перспективи розвитку точних наук, економіки та методики їх викладання» (2024)

29. Терещенко Т. Є., Заволока Л. О., Пономарьова О. Б. Страхування (у схемах, таблицях, коментарях) : навч. посібник. – Дніпро : Університет митної справи та фінансів, 2020. − 221 с. − (Серія «Бізнес. Економіка. Фінанси»)

30. Bowers, Newton, Gerber Hans, Hickey, James, et al. Actuarial Mathematics. Second Edition. Schaumburg: Society of Actuaries, 1997. 753 p. 31. E. T. Jaynes. Probability Theory: The Logic of Science. Cambridge: Cambridge University Press, 2003.

32. E.M. Mirkes (2011) «K-means and K-medoids applet»

33. Erich Schubert, Jörg Sander, Martin Ester, Hans Peter Kriegel, Xiaowei Xu. (2017) «DBSCAN: Why and How You Should (Still) Use DBSCAN» // ACM Trans. Database Syst.

34. Hastie Trevor, Tibshirani Robert, Friedman Jerome (2001) «The EM algorithm» — New York: Springer. — С. 236—24

35. Jain A., Murty M., Flynn P. (1999) Data Clustering: A Review. // ACM Computing Surveys

36. Kaas R., Goovarts M., Dhane J., Denuit M. Morden Actuarial Risk Theory: Kluver.2003.185c.

37. Stuart, Alan, and Keith Ord. Kendall’s Advanced Theory of Statistics. Volume 1: Distribution Theory. 6th ed. London: Edward Arnold, 1994. 38. Yinfei Yang, Daniel Cer, Amin Ahmad (2019) «Multilingual Universal Sentence Encoder for Semantic Retrieval»

39. Five Thirty Eight. URL: https://fivethirtyeight.datasettes.com/ 40. Find Open Datasets and Machine Learning Projects | Kaggle. URL: https://www.kaggle.com/datasets

ДОДАТКИ

Додаток А

Основні теоретичні відомості зі статистики

Центральна гранична теорема

Нехай $X_1, X_2, ..., X_n$ - незалежні й однаково розподілені випадкові величини. Позначимо

��(����) = ��, ��(����) = ��², ‾�� = ∑ ����/��

Центральна гранична теорема говорить про те, що випадкова величина

‾�� − ��

��/√��

при досить великому n має розподіл, близький до стандартного нормального розподілу, тобто

‾��~�� (��,��² / ��)

Чим більше n, тим кращим є це наближення.

Методи оцінки параметрів розподілу

Нехай θ – параметр, який характеризує розподіл. Однією з задач математичної статистики є задача знаходження цього параметра, використовуючи вибірку. Точно цю задачу розв’язати не можна. Тому шукають наближені значення θ, ці значення називають оцінками.

Метод моментів

Нехай $��(��, ��_1, \dots, ��_{��})$ – розподіл випадкової величини ξ, який залежить від параметрів $��_1, \dots, ��_{��}$, і нехай існують моменти $��_1$, $��_2, \dots, ��_{��}$, відповідно, порядків 1,2,…,r.

Візьмемо за оцінки цих моментів статистики

$$��^{*}_{��,��} = ��_1^{��} + \cdots + ��_{��}^{��}$$

k=1,2,…,r

В силу закону великих чисел

$$��^{*}_{��,��} \to ��_{��} \text{ при } �� \to \infty$$

k=1,2,…,n

З іншого боку

$$��_{��} = \int_{-\infty}^{\infty} ��^{��}��(��, ��_1, \dots, ��_{��})���� = ��_{��}(��_1, \dots, ��_{��})$$

Розглянемо систему алгебраїчних рівнянь відносно $��_1, \dots, ��_{��}$

$$\begin{cases} ��_1(��_1, \dots, ��_{��}) = ��^{*}_{1,��} \\ ��_2(��_1, \dots, ��_{��}) = ��^{*}_{2,��} \\ \dots \\ ��_{��}(��_1, \dots, ��_{��}) = ��^{*}_{��,��} \end{cases}$$

Якщо ця система має єдиний розв’язок, то його беруть за оцінки параметрів $\theta_1, \dots, \theta_k$.

Метод максимальної правдоподібності

Якщо вибірка береться з абсолютно неперервного розподілу з щільністю $f(x, \theta)$, то функція

$$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^{n} f(x_i, \theta)$$

називається функцією правдоподібності.

Якщо цей розподіл дискретний, то функція правдоподібності матиме вигляд

$$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^{n} p_{x_i}(\theta)$$

$$p_{x_i}(\theta) = P(X = x_i)$$

Метод максимальної правдоподібності полягає в тому, що параметри θ вибираються так, щоб функція правдоподібності набула найбільшого значення. Для знаходження цих значень потрібно розв’язати рівняння правдоподібності

$$\frac{\partial \ln L}{\partial \theta} = 0$$

бо функція $\ln L$ при фіксованих $X_1, X_2, \dots, X_n$ досягає максимуму при тих же значеннях θ, що і функція L.

Розв’язки рівняння правдоподібності називаються оцінками максимальної правдоподібності.

Перевірка якості підгонки

Будь-яке твердження про генеральну сукупність, яке перевіряється на вибірці називається статистичною гіпотезою.

Статистичні гіпотези діляться на гіпотези про параметри розподілу і гіпотези про тип розподілу.

Варто відмітити, що статистичними методами гіпотезу можна тільки спростувати або прийняти, але не довести.

Правило, за допомогою якого гіпотеза приймається, або відхиляється, називається статистичним критерієм.

Розглянемо питання перевірки гіпотези про тип розподілу. Нехай гіпотеза *H* полягає в тому, що вибірка об'єкту *n* береться із заданого розподілу $F_0(x)$ і вважатимемо, що цей розподіл повністю визначений, тобто не містить невідомих параметрів.

Критерій згоди Пірсона χ^2 є найпотужнішим непараметричним критерієм. При порівнянні емпіричного розподілу з теоретичним необхідно розрахувати частоти відповідного теоретичного розподілу. Значення критерію дорівнює:

$$\chi^2=\sum_{i=1}^{k}\frac{(n_{емп}-n_{теор})^2}{n_{теор}}$$

$n_{емп}$ – емпірична частота, $n_{теор}$ – теоретична частота.

Критичне значення визначають для (k-r-1)-степенів вільності (k – кількість класових інтервалів, r – кількість оцінюваних параметрів розподілу) та рівня значущості p.

Тепер ми можемо сформулювати правило для перевірки гіпотези про розподіл.

Критерій згоди Пірсона χ^2. Фіксуємо такий малий рівень значущості, щоб можна було практично не сумніватись в тому, що при одному випробуванні подія $\{\chi^2>\chi_{кр}^2\}$ не відбудеться. Якщо реалізація χ^2 виявилась такою, що $\chi^2>\chi_{кр}^2$, то нульову гіпотезу відхиляємо, в протилежному випадку, гіпотеза може бути прийнятою.

Додаток Б

Критичні точки розподілу χ^2

Число ступенів свободи, *k*	Рівень значущості, α					
	0,01	0,025	0,05	0,95	0,975	0,999
1	6,6	5,0	3,8	0,0039	0,00098	0,00016
2	9,2	7,4	6,0	0,103	0,051	0,020
3	11,3	9,4	7,8	0,352	0,216	0,115
4	13,3	11,1	9,5	0,711	0,484	0,297
5	15,1	12,8	11,1	1,15	0,831	0,554
6	16,8	14,4	12,6	1,64	1,24	0,872
7	18,5	16,0	14,1	2,17	1,69	1,24
8	20,1	17,5	15,5	2,73	2,18	1,65
9	21,7	19,0	16,9	3,33	2,70	2,09
10	23,2	20,5	18,3	3,94	3,25	2,56
11	24,7	21,9	19,7	4,57	3,82	3,05
12	26,2	23,3	21,0	5,23	4,40	3,57
13	27,7	24,7	22,4	5,89	5,01	4,11
14	29,1	26,1	23,7	6,57	5,63	4,66
15	30,6	27,5	25,0	7,26	6,26	5,23
16	32,0	28,8	26,3	7,96	6,91	5,81
17	33,4	30,2	27,6	8,67	7,56	6,41
18	34,8	31,5	28,9	9,39	8,23	7,01
19	36,2	32,9	30,1	10,1	8,91	7,63
20	37,6	34,2	31,4	10,9	9,59	8,26
21	38.9	35,5	32,7	11,6	10,3	8,90
22	40,3	36,8	33,9	12,3	11,0	9,54
23	41,6	38,1	35,2	13,1	11,7	10,2
24	43,0	39,4	36,4	13,8	12,4	10,9
25	44,3	40,6	37,7	14,6	13,1	11,5
26	45,6	41,9	38,9	15,4	13,8	12,2
27	47,0	43,2	40,1	16,2	14,6	12,9

28	48,3	44,5	41,3	16,9	15,3	13,6
29	49,6	45,7	42,6	17,7	16,0	14,3
30	60,9	47,0	43,8	18,5	16,8	15,0

x
–
z
ток В

2

1 *Дода*

Таблиця значень функції *Φ x e dz*
Лапласа

( ) = − 2 π ∫ 0 2

x	*Φ*(*x*)	*x*	*Φ*(*x*)	*x*	*Φ*(*x*)	*x*	*Φ*(*x*)
0,00	0,0000	0,26	0,1026	0,52	0,1985	0,78	0,2823
0,01	0,0040	0,27	0,1064	0,53	0,2019	0,79	0,2852
0,02	0,0080	0,28	0,1103	0,54	0,2054	0,80	0,2881
0,03	0,0120	0,29	0,1141	0,55	0,2088	0,81	0,2910
0,04	0,0160	0,30	0,1179	0,56	0,2123	0,82 0	0,2939
0,05	0,0199	0,31	0,1217	0,57	0,2157	0,83	0,2967
0,06	0,0239	0,32	0,1255	0,58	0,2190	0,84	0,2995
0,07	0,0279	0,33	0,1293	0,59	0,2224	0,85	0,3023
0,08	0,0319	0,34	0,1331	0,60	0,2257	0,86	0,3051
0,09	0,0359	0,35	0,1368	0,61	0,2291	0,87	0,3078
0,10	0,0398	0,36	0,1406	0,62	0,2324	0,88	0,3106
0,11	0,0438	0,37	0,1443	0,63	0,2357	0,89	0,3133
0,12	0,0478	0,38	0,1480	0,64	0,2389	0,90	0,3159
0,13	0,0517	0,39	0,1617	0,65	0,2422	0,91	0,3186
0,14	0,8557	0,40	0,1564	0,66	0,2454	0,92	0,3212
0,15	0,0596	0,41	0,1691	0,67	0,2486	0,93	0,3238
0,16	0,0636	0,42	0,1628	0,68	0,2517	0,94	0,3264
0,17	0,0675	0,43	0,1664	0,69	0,2549	0,95	0,3289
0,18	0,0714	0,44	0,1700	0,70	0,2580	0,96	0,3315

0,19	0,0753	0,45	0,1736	0,71	0,2611	0,97	0,3340
0,20	0,0793	0,46	0,1772	0,72	0,2642	0,98	0,3365
0,21	0,0832	0,47	0,1808	0,73	0,2673	0,99	0,3389
0,22	0,0871	0,48	0,1844	0,74	0,2703	1,00	0,3413
0,23	0,0910	0,49	0,1879	0,75	0,2734	1,01	0,3438
0,24	0,0948	0,50	0,1915	0,76	0,2764	1,02	0,3461
0,25	0,0987	0,51	0,1950	0,77	0,2794	1,03	0,3485
x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1,04	0,3508	1,33	0,4082	1,62	0,4474	1,91	0,4719
1,05	0,3531	1,34	0,4099	1,63	0,4484	1,92	0,4726
1,06	0,3554	1,35	0,4115	1,64	0,4495	1,93	0,4732
1,07	0,3577	1,36	0,4131	1,65	0,4505	1,94	0,4738
1,08	0,3599	1,37	0,4147	1,66	0,4515-	1,95	0,4744
1,09	0,3621	1,38	0,4162	1,67	0,4525	1,96	0,4750
1,10	0,3643	1,39	0,4177	1,68	0,4535	1,97	0,4756
1,11	0,3665	1,40	0,4192	1,69	0,4545	1,98	0,4761
1,12	0,3686	1,41	0,4207	1,70	0,4554	1,99	0,4767
1,13	0,3708	1,42	0,4222	1,71	0,4564	2,00	0,4772
1,14	0,3729	1,43	0,4236	1,72	0,4573	2,02	0,4783
1,15	0,3749	1,44	0,4251	1,73	0,4582	2,04	0,4793
1,16	0,3770	1,45	0,4265	1,74	0,4591	2,06	0,4803

1,17	0,3790	1,46	0,4279	1,75	0,4599	2,08	0,4812
1,18	0,3810	1,47	0,4292	1,76	0,4608	2,10	0,4821
1,19	0,3830	1,48	0,4306	1,77	0,4616	2,12	0,4830
1,20	0,3849	1,49	0,4319	1,78	0,4625	2,14	0,4838
1,21	0,3869	1,50	0,4332	1,79	0,4633	2,16	0,4846
1,22	0,3883	1,51	0,4345	1,80	0,4641	2,18	0,4854
1,23	0,3907	1,52	0,4357	1,81	0,4649	2,20	0,4861

1,24	0,3925	1,53	0,4370	1,82	0,4656	2,22	0,4868
1,25	0,3944	1,54	0,4382	1,83	0,4664	2,24	0,4875
1,26	0,3962	1,55	0,4394	1,84	0,4671	2,26	0,4881
1,27	0,3980	1,56	0,4406	1,85	0,4678	2,28	0,4887
1,28	0,3997	1,57	0,4418	1,86	0,4686	2,30	0,4893
1,29	0,4015	1,58	0,4429	1,87	0,4693	2,32	0,4898
1,30	0,4032	1,59	0,4441	1,88	0,4699	2,34	0,4904
1,31	0,4049	1,60	0,4452	1,89	0,4706	2,36	0,4909
1,32	0,4066	1,61	0,4463	1,90	0,4713	2,38	0,4913
2,40	0,4918	2,60	0,4953	2,80	0,4974	3,20	0,49931
2,42	0,4922	2,62	0,4956	2,82	0,4976	3,40	0,49966
2,44	0,4927	2,64	0,4959	2,84	0,4977	3,60	0,49984
2,46	0,4931	2,66	0,4961	2,86	0,4979	3,80	0,499928
2,48	0,4934	2,68	0,4963	2,90	0,4981	4,00	0,499968
2,50	0,4938	2,70	0,4965	2,92	0,4982	5,00	0,499997
2,52	0,4941	2,72	0,4967	2,94	0,4984		
2,54	0,4945	2,74	0,4969	2,96	0,49846		
2,56	0,4948	2,76	0,4971	2,98	0,49856		
2,58	0,4951	2,78	0,4973	3,00	0,49865	x > 5	0,5

Додаток Г

Статистичні дані по штатам США:

1) Кількість водіїв, які потрапили у смертельні зіткнення на мільярд миль, 2) Відсоток водіїв, залучених у смертельні зіткнення, які перевищували швидкість, 3) Відсоток водіїв, залучених у смертельні зіткнення, які перебували в стані алкогольного сп’яніння, 4) Відсоток водіїв, залучених у смертельні зіткнення, які не відволікалися, 5) Відсоток водіїв, залучених у зіткнення зі смертельними наслідками, які раніше не брали участі в аваріях, 6) Премії за страхування автомобіля ($), 7) Збитки, понесені страховими

компаніями через зіткнення, на одного застрахованого водія ($).

State	1)Numb er of drivers involv ed in fatal collisi on s per billion miles	2)Perce nta ge Of Drivers Involved In Fatal Collision s Who Were Speedin g	3)Perce nta ge Of Drivers Involved In Fatal Collisio ns Who Were Alcohol Impaired	4)Perce nta ge Of Drivers Involved In Fatal Collisio ns Who Were Not Distracted	5)Perce nta ge Of Drivers Involved In Fatal Collisio ns Who Had Not Been Involved In Any Previous Accidents	6)Car Insur anc e Premiu ms ($)	7)Loss es incurre d by insuranc e compani es for collision s per insured driver ($)
Alabama	18.8	39%	30%	96%	80%	784.55	145.08
Alaska	18.1	41%	25%	90%	94%	1053.48	133.93
Arizona	18.6	35%	28%	84%	96%	899.47	110.35
Arkansas	22.4	18%	26%	94%	95%	827.34	142.39
California	12	35%	28%	91%	89%	878.41	165.63
Colorado	13.6	37%	28%	79%	95%	835.5	139.91
Connecticut	10.8	46%	36%	87%	82%	1068.73	167.02
Delaware	16.2	38%	30%	87%	99%	1137.87	151.48
District of Columbia	5.9	34%	27%	100%	100%	1273.89	136.05
Florida	17.9	21%	29%	92%	94%	1160.13	144.18
Georgia	15.6	19%	25%	95%	93%	913.15	142.8
Hawaii	17.5	54%	41%	82%	87%	861.18	120.92
Idaho	15.3	36%	29%	85%	98%	641.96	82.75
Illinois	12.8	36%	34%	94%	96%	803.11	139.15
Indiana	14.5	25%	29%	95%	95%	710.46	108.92
Iowa	15.7	17%	25%	97%	87%	649.06	114.47
Kansas	17.8	27%	24%	77%	85%	780.45	133.8
Kentucky	21.4	19%	23%	78%	76%	872.51	137.13
Louisiana	20.5	35%	33%	73%	98%	1281.55	194.78

Maine	15.1	38%	30%	87%	84%	661.88	96.57
Maryland	12.5	34%	32%	71%	99%	1048.78	192.7
Massachusetts	8.2	23%	35%	87%	80%	1011.14	135.63
Michigan	14.1	24%	28%	95%	77%	1110.61	152.26
Minnesota	9.6	23%	29%	88%	88%	777.18	133.35
Mississippi	17.6	15%	31%	10%	100%	896.07	155.77
Missouri	16.1	43%	34%	92%	84%	790.32	144.45
Montana	21.4	39%	44%	84%	85%	816.21	85.15
Nebraska	14.9	13%	35%	93%	90%	732.28	114.82
Nevada	14.7	37%	32%	95%	99%	1029.87	138.71
New Hampshire	11.6	35%	30%	87%	83%	746.54	120.21
New Jersey	11.2	16%	28%	86%	78%	1301.52	159.85
New Mexico	18.4	19%	27%	67%	98%	869.85	120.75
New York	12.3	32%	29%	88%	80%	1234.31	150.01
North Carolina	16.8	39%	31%	94%	81%	708.24	127.82
North Dakota	23.9	23%	42%	99%	86%	688.75	109.72
Ohio	14.1	28%	34%	99%	82%	697.73	133.52
Oklahoma	19.9	32%	29%	92%	94%	881.51	178.86
Oregon	12.8	33%	26%	67%	90%	804.71	104.61
Pennsylvania	18.2	50%	31%	96%	88%	905.99	153.86
Rhode Island	11.1	34%	38%	92%	79%	1148.99	148.58
South Carolina	23.9	38%	41%	96%	81%	858.97	116.29
South Dakota	19.4	31%	33%	98%	86%	669.31	96.87
Tennessee	19.5	21%	29%	82%	81%	767.91	155.57
Texas	19.4	40%	38%	91%	87%	1004.75	156.83
Utah	11.3	43%	16%	88%	96%	809.38	109.48
Vermont	13.6	30%	30%	96%	95%	716.2	109.61
Virginia	12.7	19%	27%	87%	88%	768.95	153.72
Washington	10.6	42%	33%	82%	86%	890.03	111.62

West Virginia	23.8	34%	28%	97%	87%	992.61	152.56
Wisconsin	13.8	36%	33%	39%	84%	670.31	106.62
Wyoming	17.4	42%	32%	81%	90%	791.14	122.04

Додаток Г

Кореляції

	Percentage Of Drivers Involved In Fatal Collisions Who Were Not Distracted	Percentage Of Drivers Involved In Fatal Collisions Who Had Not Been Involved In Any Previous Accidents	Percentage Of Drivers Involved In Fatal Collisions Who Were Speeding	Percentage Of Drivers Involved In Fatal Collisions Who Were Alcohol Impaired

Percentage Of Drivers Involved In Fatal Collisions Who Were Not Distracted

Percentage Of Drivers Involved In Fatal Collisions Who Had Not Been Involved In Any Previous Accidents Percentage Of Drivers Involved In Fatal Collisions Who Were Speeding

Percentage Of Drivers Involved In Fatal Collisions Who Were Alcohol Impaired

51	51
-,245	,286*
,083	,042
51	51

Кореляція Пірсона 1 ,043 Знач. (двустороння) ,762 N 51 51

Кореляція Пірсона -,195 -,245 Знач. (двустороння) ,170 ,083 N 51 51

Кореляція Пірсона ,132 ,286* Знач. (двустороння) ,357 ,042 N 51 51

Кореляція Пірсона ,043 1 Знач. (двустороння) ,762 N 51 51

-,195	,132
,170	,357
51	51
1	,014
	,922
51	51
,014	1
,922	

*. Кореляція значима на рівні 0,05 (двохстороння).