

УДК 624.21.059.1

Давиденко О.О.

**ПЕРЕВІРКА ВІДПОВІДНОСТІ ОБРАНОЇ МОДЕЛІ РОЗПОДІЛУ
СПОСТЕРЕЖУВАНИМ ДАНИМ ТЕРМІНІВ СЛУЖБИ
ЗАЛІЗОБЕТОННИХ АВТОДОРОЖНИХ МОСТІВ УКРАЇНИ**

Анотація. Стаття присвячена статистичному аналізу термінів служби мостів. Досліджується відповідність закону розподілу спостережуваних даних до теоретичних моделей розподілу за непараметричним критерієм згоди Колмогорова-Смірнова.

Ключові слова: залізобетонні мости, критерій згоди Колмогорова-Смірнова, термін служби.

УДК 624.21.059.1

Давыденко А.А.

**ПРОВЕРКА СООТВЕТСТВИЯ ВЫБРАННОЙ МОДЕЛИ
РАСПРЕДЕЛЕНИЯ НАБЛЮДАЕМЫМ ДАННЫМ СРОКОВ СЛУЖБЫ
ЖЕЛЕЗОБЕТОННЫХ АВТОДОРОЖНЫХ МОСТОВ УКРАИНЫ**

Аннотация. Статья посвящена статистическому анализу сроков службы мостов. Исследуется соответствие закону распределения наблюдаемых данных и теоретических моделей распределения используя непараметрический критерий согласия Колмогорова-Смирнова.

Ключевые слова: железобетонные мосты, критерий согласия Колмогорова-Смирнова, срок службы.

UDC 624.21.059.1

Davydenko O.O.

**VERIFICATION OF COMPLIANCE OF THE SELECTED MODEL
DISTRIBUTION TO THE OBSERVED DATA OF SERVICE LIFE OF
REINFORCED CONCRETE HIGHWAY BRIDGES OF UKRAINE**

Abstract. This article is devoted to statistical analysis of the service life of bridges. Conformance with the distribution of the observed data to theoretical models for the distribution-free test of fit Kolmogorov-Smirnov.

Keywords: concrete bridges, distribution-free test of fit Kolmogorov-Smirnov, service life.

Вступ

Дослідження базується на даних отриманих з галузевої Аналітичної експертної системи управління мостами (АЕСУМ) Укравтодору. Науковий супровід програмного комплексу АЕСУМ веде ДерждорНДІ Укравтодору. Починаючи з 2006 р. АЕСУМ функціонує, постійно розвивається і вдосконалюється, накопичується банк даних [1].

База даних, за станом на 01.01.2016 р., містить 16191 од. автодорожніх мостів підпорядкованих Укравтодору. З цієї загальної кількості паспортизовано і описано в базі даних 7031 мостів.

Предмет нашого дослідження – вибірка залізобетонних мостів кількістю в 6919 од., що складає 48,4 % від загальної кількості залізобетонних мостів (14305 од.) підпорядкованих Укравтодору. В теорії математичної статистики 48 – відсоткова вибірка вважається достатньо представницькою. Таким чином, результати аналізу вибірки можуть бути достовірно розповсюджені на загальну кількість автодорожніх залізобетонних мостів [3, 5].

Опис кількісних даних залежить від типу їх розподілу. Тому, для підтвердження кількісних даних, необхідно перевірити тип їх розподілу. Також перевірка розподілу на основі отриманих даних необхідна, коли досліднику невідомо розподіл кількісних ознак в генеральній сукупності, як в нашому випадку. В нашому дослідженні для великої вибірки будемо використовувати критерій згоди Колмогорова-Смірнова.

Мета перевірки гіпотези про згоду дослідного розподілу з теоретичним – це прагнення впевнитися в тому, що модель теоретичного закону не суперечить спостережуваним даним, і використання її не призведе до суттєвих помилок при імовірнісних розрахунках в подальшому.

Аналіз останніх досліджень і публікацій

В багатьох задачах теорії надійності, в яких оперують даними типу часу або інтенсивністю відмов, є необхідність визначення відповідності вибірки статистичних даних до певного закону розподілу. Для інженерних задач, в більшості випадків, перевіряють непараметричні гіпотези про зв'язки функції розподілу декількох вибірок в разі, коли не виконується класичні умови незалежності і однакового розподілу елементів вибірок [9].

Більшість наукових публікацій [2, 6, 9, 11] українською та російською мовами на тему дослідження статистичної вибірки та її розподілу розглядаються в рамках, лише математичної статистики.

Математична статистика є інструментом дослідника в багатьох різноманітних задачах, але наукових публікацій, які описують як використовувати теорію в рамках певної спеціалізації значно менше. В статті [10] наводять аналіз статистичної вибірки фізичних показників кваліфікованих фехтувальників-шаблістів, метою якого є покращення результатів змагань. Статистичний аналіз також використовують для визначення ступеня однорідності великої вибірки пацієнтів з гострим панкреатитом [8]. Цілю роботи є визначення причин виникнення та розвитку захворювань, що дозволить в майбутньому проводити диференційовану терапію даного захворювання і запобігати частим рецидивам хвороби. Наукова публікація [7] підіймає питання визначення стандартів в'язкості мастики для різних партій та двох різних методів заміру. Для рішення такої задачі науковець розглядає статистичні критерії однорідності двох незалежних статистичних вибірок.

Нами не знайдено україномовних та російськомовних опублікованих наукових праць, які досліджують статистичні вибірки термінів служби залізобетонних мостів та їх відповідність закону розподілу за критерієм згоди Колмогорова-Смірнова.

Алгоритм дослідження варіаційного ряду термінів служби мостів

Досліджується варіаційний ряд термінів служби автодорожніх мостів підпорядкованих Державній адміністрації автомобільних доріг України (Укравтодор). Варіаційний ряд термінів служби T , тобто випадкової величини T , досліджується за класичною схемою оцінки статистичних даних [2].

Варіаційний ряд характеризується обсягом вибірки n та розмахом варіювання:

$$R = x_{\max} - x_{\min}, \quad (1)$$

де x_{\max} , x_{\min} – максимальна і мінімальна варіанти ряду;

n – кількість варіант ряду.

Для варіаційного ряду будемо визначати емпіричну функцію щільності розподілу, яка визначає для кожного значення x відносну частоту подій:

$$F_n(x) = \frac{n_x}{n}, \quad (2)$$

де n_x – кількість варіант, які менші від x ;

n – обсяг варіаційного ряду.

Емпірична функція щільності розподілу відображається графічно гістограмою відносних частот – ступінчатим графіком, який складається з прямокутників з основами з частинних інтервалів варіант довжиною $h = x_k - x_{k-1}$ та висот n_k/h . Тут k – кількість інтервалів варіаційного ряду. Для рядів близьких до нормального розподілу орієнтовне значення кількості інтервалів (ціле число) визначається за формулою:

$$k = 1 + 3,322 \lg(n), \quad (3)$$

Для варіаційного ряду будемо також обчислювати статистичні оцінки:

- медіану – значення середнього елемента варіаційного ряду

$$\text{Me} = \begin{cases} x_m; n = 2m - 1 \\ \frac{x_m + x_{m+1}}{2}; n = 2m \end{cases} \quad (4)$$

- середнє значення варіант

$$\mu = \frac{1}{n} \sum_{i=1}^k x_i n_i; \quad (5)$$

- середнє квадратичне відхилення – квадратний корінь із вибіркової дисперсії

$$\sigma_B = \sqrt{D_B}; \quad (6)$$

- вибіркoву дисперсію – середня квадратів відхилення варіант від вибіркової середньої з врахуванням відповідних відносних частот

$$D_B = \frac{1}{n} \sum_{k=1}^m n_k (x_k - x_B)^2. \quad (7)$$

Досліджуються варіаційні ряди термінів служби (віку) прогонових будов в третьому експлуатаційному стані (тут і далі стан визначається за нормативним документом [3]), який представлено обсягом $n = 4288$ од. (табл.1).

Таблиця 1 – Статистичні характеристики вибірки для віку мостів, що знаходяться в третьому експлуатаційному стані

	Медіана, роки	μ_{\min} , роки	μ_{\max} , роки	Середнє значення елементів μ , роки	Середнє квадратичне відхилення σ , роки	Дисперсія D , роки ²	Вибірка кількість мостів n , од.
	Стан 3						
Монолітні	52	39	69	54	15	223	591
Збірно-монолітні	38	22	52	37	15	222	199
Збірні	40	24	56	40	16	242	3179
Всі типи + немає відомостей про тип	42	25	59	42	17	278	4288

Аналіз термінів служби автодорожніх мостів України

Вибірка статистичних даних термінів служби автодорожніх мостів України, отримана з ПК АЕСУМ, досліджується на відповідність до закону розподілу нормального, логнормального, логістичного та розподілу Коші. Дослідження проводиться за критерієм згоди Колмогорова-Смірнова. Критерій був обраний так, як в своєму класичному вигляді є «більш потужним» ніж «критерій χ^2 » і може використовуватися для перевірки гіпотези про належність емпіричного розподілу до будь-якого теоретичного розподілу $F(x)$ з відомими статистичними параметрами.

Приклад наведено для найчисельнішою вибірки всіх типів мостів в третьому експлуатаційному стані (рис. 1). Далі опишемо алгоритм дослідження.

Маємо вибірку $\mathbf{X} = (X_1, X_2, \dots, X_n)$ з деякого розподілу F . Перевіряємо просту гіпотезу $H_0 = \{ F = F_1 \}$ проти складної альтернативної $H_1 = \{ F \neq F_1 \}$.

Статистика критерію Колмогорова-Смірнова визначається за формулою:

$$T_n = \sqrt{n} \cdot \sup_x \left(F_n^*(x) - F_1(x) \right) \stackrel{\sim}{=} \sqrt{n} \cdot \max_{X_i} \left(F_n^*(X_i) - F_1(X_i) \right), \quad (8)$$

де $F_n^*(x)$ – емпірична функція розподілу; $F_1(x)$ – теоретична функція розподілу; n – кількість елементів вибірки, 4288 од.

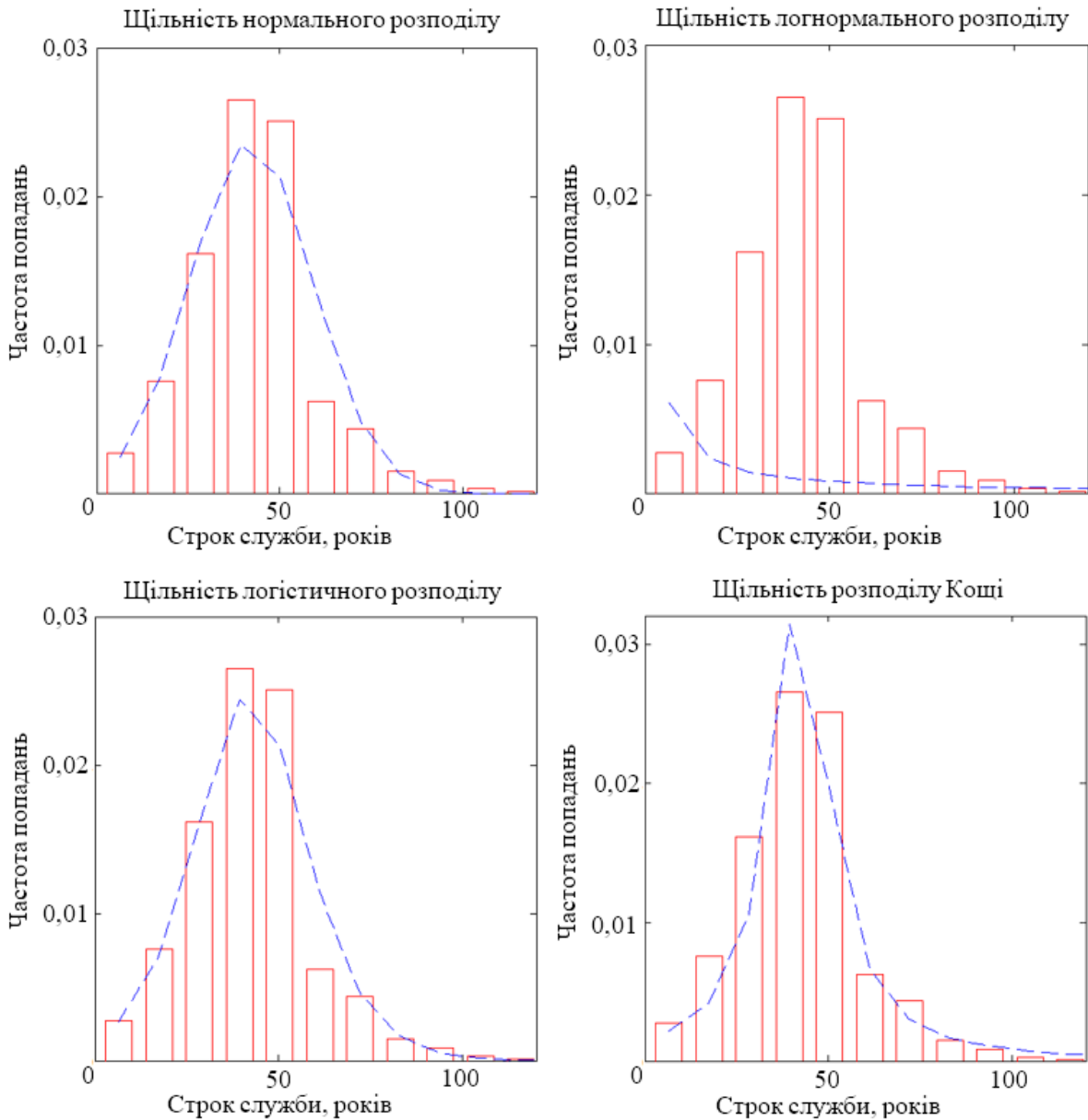


Рисунок 1 – Гістограма вибірки строків служби для всіх типів мостів в 3-у експлуатаційному стані та щільність розподілу за нормальним, логнормальним, логістичним та розподілом Коші

Критичні значення статистичних критеріїв можна визначити з таблиць або за допомогою функції:

$$C(q) = \sqrt{\frac{-\ln\left(\frac{q}{2}\right)}{2}}, \quad (9)$$

де q – рівень значимості, $q = 0,1$.

Основна гіпотеза H_0 приймається у випадку коли $T_n < C(q)$. Всі отримані дані представлено в табл. 2.

Таблиця 2 – Гіпотеза про вид розподілу

Теоретична функція	Статистика критерію, T_n	Критичні значення статистичного критерію, $C(0,1)$	Гіпотеза, що приймається
Нормального розподілу	3,975	1,224	H_1
Логнормального розподілу	25,375	1,224	H_1
Логістичному розподілу	3,664	1,224	H_1
Розподілу Коші	3,485	1,224	H_1

Висновки

1. За непараметричним критерієм згоди Колмогорова-Смірнова для всіх типів розподілу досліджуваної вибірки була прийнята альтернативна складна гіпотеза про вид розподілу $H_1 = \{ F \neq F_1 \}$.

2. За критерієм згоди Колмогорова-Смірнова найбільша відмінність між теоретичною функцією логнормального розподілу та функцією розподілу спостережуваних даних. Найбільшу схожість функція розподілу статистичної вибірки має з теоретичною функцією розподілу Коші, яка відрізняється на 4 відсотки від функції нормального розподілу, в порівнянні з граничним критерієм. Вид функції Коші і логістичної функції розподілу залежать від двох параметрів: μ – математичне очікування; s – параметр масштабу, що підбирається емпірично. В нашому випадку при мінімальному розходженні зручніше використовувати функцію нормального розподілу яка має два параметри: μ – математичне очікування; σ – середнє квадратичне відхилення, що вже відомі з дослідження варіаційного ряду.

3. Дослідження виду розподілу показало, що кожна статистична вибірка розподіляється за своїм законом який неможливо описати жодною відомою функцією розподілу. В теоретичних дослідженнях термінів служби автодорожніх залізобетонних мостів України будемо використовувати нормальний закон розподілу, як найзручніший і той, що є одним з найбільш наближених.

Ця робота була виконана під керівництвом д-ра техн. наук, професора Лантуха-Лященко А.І. Висловлюю йому мою щирю подяку.

Література

1. Боднар Л.П. Програмний комплекс АЕСУМ. Сучасний стан та концепція подальшого розвитку / Л.П. Боднар // «Дороги і мости»: зб. наук. п – К.: ДерждорНДІ, 2010. – Вип.12. - С. 31-39.
2. Горбань І.І. Теорія ймовірностей і математична статистика для наукових працівників та інженерів / Горбань І.І. – Київ, 2003. – С. 244.
3. Давиденко О.О. Моніторинг безпечної експлуатації автодорожніх мостів України [Текст] / О.О. Давиденко // Збірник наукових праць «Мости та тунелі: теорія, дослідження, практика» – Д., 2015. – Вип. 7. – С. 4–12.
4. ДСТУ-Н Б В.2.3-23:2012. Настанова з оцінювання і прогнозування технічного стану автодорожніх мостів: – На заміну ДСТУ-Н Б В.2.3-23:2009. – [Чинні від 2013-12-01]. – К. : Мінрегіонбуд України, 2013. – 49 с. – (Національний стандарт України).
5. Лантух-Лященко, А.І. Уточнення оцінки експлуатаційного стану мостів [Текст] / А.І. Лантух-Лященко // Збірник. Дороги і мости. ДерждорНДІ – К, 2008. – Вип. 9. – С. 12–18.
6. Мазур, Н. П. Статистичні критерії як інструмент аналізу отриманих даних моніторингового дослідження [Текст] / Н. П. Мазур // Проблеми сучасної педагогічної освіти. Се: Педагогіка і психологія – К., 2014. – Вип. № 44. – С. 190–197.
7. Орлов, А. И. О проверке однородности связанных выборок [Текст] / А. И. Орлов // Научный журнал КубГАУ. – Кубань, 2016. – Вып. № 123. – С. 1–19.
8. Спиринов, А. Н., Солнцев, В. Н., Назаров, В. Е. Применение методов многомерного статистического анализа для определения степени однородности большой выборки пациентов с острым отечным панкреатитом [Текст] / А. Н. Спиринов, В. Н. Солнцев, В. Е. Назаров // Известия Российского государственного педагогического университета им. А.И. Герцена – М., 2012. – Вып. № 147. – С. 87–97.
9. Тянникова, Н. Д., Тимонин, В. И. Метод вычисления точных распределений статистик типа Колмогорова-Смирнова в случае нарушения однородности и независимости анализируемых выборок [Текст] / Н. Д. Тянникова, В. И. Тимонин // Наука и Образование. МГТУ им. Н.Э. Баумана. Электрон. журн. – М., 2014. – Вып. № 11. – С. 227–237.
10. Хохла, А. І. Структура фізичної підготовленості кваліфікованих фехтувальників-шаблестів [Текст] / А. І. Хохла // Вісник «Проблеми формування й удосконалення спортивно-технічної майстерності» – 2014. – Вип. № 118. – С. 215–218.
11. Чимитова, Е. В., Ведерникова, М. А., Галанова, Н. С. Непараметрические критерии согласия в задачах проверки адекватности моделей надежности по цензурированным данным [Текст] / Е. В. Чимитова, М. А. Ведерникова, Н. С. Галанова // Вестник Томского государственного университета – Томск, 2013. – Вып. № 25. – С. 115–124.

Рецензенти:

Мішутін А.В., д-р техн. наук, Одеська державна академія будівництва і архітектури.
Снітко В.П. канд. нехн. наук. Національний транспортний університет.

Reviewers:

Mishutin A.V., Dr. Tech. Sci., Odessa State Academy of Construction and Architecture.
Snitko V. P. candidate. nehn. Sciences. National Transport University.

Стаття надійшла до редакції: **31.01.2017 р.**