УДК 656.2, 519.233

А. Ю. Андрейцев, к.ф.-м.н, доцент
(доцент кафедри «Вища математика» Державного економіко-технологічного університету транспорту, м. Київ)
Т. С. Клецька, к.і.н.
(доцент кафедри «Вища математика» Державного економіко-технологічного університету транспорту, м. Київ)
О. О. Кільчинський, к.ф.-м.н, доцент
(доцент кафедри «Вища математика» Державного економіко-технологічного університету транспорту, м. Київ)

ПРО ВДОСКОНАЛЕННЯ МЕТОДИКИ ПЕРЕВІРКИ СКЛАДНИХ СТАТИСТИЧНИХ ГІПОТЕЗ ПРИ АНАЛІЗІ СВОЄЧАСНОСТІ ДОСТАВКИ ВАНТАЖІВ

У статті запропонована вдосконалена методика перевірки складних статистичних гіпотез відносно законів розподілу досліджуваної ознаки генеральної сукупності для аналізу відхилення фактичного часу доставки вантажів від заданого. Її застосування продемонстроване на прикладі перевірки гіпотези про розподіл Вейбулла відхилення фактичного часу доставки від мінімального.

Ключові слова: *термін доставки вантажу, фактичний час доставки, мінімальний (технологічний) час доставки, стохастична модель, статистична гіпотеза, розподіл Вейбулла.*

В статье предложена усовершенствованная методика проверки сложных статистических гипотез относительно законов распределения исследуемого признака генеральной совокупности для анализа отклонения фактического времени доставки от заданого. Её применение продемонстрировано на примере проверки гипотезы о распределении Вейбулла отклонения фактического времени доставки от минимального.

Ключевые слова: *срок доставки грузов, фактическое время доставки, минимальное (технологическое) время доставки, стохастическая модель, статистическая гипотеза, распределение Вейбулла.*

В умовах економічної нестабільності останніх років актуальною стає задача підвищення ефективності технологічних процесів шляхом їхньої оптимізації. В зв'язку з цим значно зростає роль математичного моделювання вказаних процесів з метою подальшого прогнозування та оптимізації.

Однією з таких задач, наприклад, є задача своєчасної доставки вантажів. При порушенні термінів доставки залізниця несе збитки, пов'язані із штрафами за прострочену доставку, зберіганням вантажів (передчасна доставка) тощо.

Математичні моделі транспортних потоків є стохастичними і базуються на статистичному матеріалі. Суттєвим критерієм для застосування цих моделей є їхня адекватність: відповідність реальному процесу. Це дозволяє досить точно оцінювати витрати та ризики, пов'язані з відхиленнями від заданих умов процесу (наприклад, відхиленнями від графіка доставки вантажів).

З цього приводу безумовно важливим є процес перевірки адекватності, що в умовах стохастичності вимагає перевірки статистичних гіпотез як непараметричних (гіпотез про закон розподілу генеральної сукупності), так і параметричних (про рівність статистичних оцінок значенням параметрів відомого закону розподілу).

В [1] було проведено аналіз та структурну класифікацію витрат, пов'язаних з несвоєчасною доставкою вантажів та запропонована математична модель, що базується на нормальному законі розподілу відхилень реального часу доставки від графікового.

У загальному випадку ці витрати можуть бути обчислені за формулою

$$C = N\int_{t_{\min}}^{+\infty} K(\Delta t^*) f(\Delta t^*) d\Delta t^* , \tag{1}$$

де N – загальний обсяг вантажів, що перевозяться в певний проміжок часу, $t_{\min}$ –мінімальний час доставки, $K(\Delta t^*)$ – ставки плати за одиницю вантажу, пов'язані з обслуговуванням передчасно доставлених вантажів та простроченням термінів доставки, що залежать від відхилення Δt^* фактичного часу доставки від заданого (графікового або мінімального), $f(\Delta t^*)$ – щільність розподілу відхилень.

Аналогічну формулу можна записати для відносних відхилень: $\tau^* = \dfrac{\Delta t^*}{t_{\min}}$.

У [2] була побудована стохастична модель для визначення найімовірніших втрат перевізної плати залізницею, пов'язаних з простроченням термінів доставки вантажів, у припущенні, що час прострочення розподілений за експоненціальним законом.

У [3, 4] запропонована математична модель у припущенні, що відхилення фактичного часу доставки від мінімального підлягає закону розподілу Вейбулла та проведено аналіз зміни відхилень залежно від конкретних значень параметрів даного розподілу. На жаль, у наведених роботах не приділено увагу перевірці адекватності побудованих моделей.

Метою даної роботи є вдосконалення методики перевірки складних гіпотез відносно законів розподілу досліджуваної ознаки генеральної сукупності на прикладі, відхилення фактичного часу доставки вантажів від заданого.

Нехай ми маємо сукупність вибіркових даних відносно часу доставки вантажів на деякій лінії залізниці. Виходячи з цих даних ми можемо побудувати статистичний ряд та визначити такі характеристики: мода $t_ì$ – час, за який доставляється максимальна кількість вантажів; медіана t_m – час, протягом якого буде доставлено половину усіх вантажів, t_c – середній час доставки. Розглянемо абсолютні відхилення цих величин від мінімального часу доставки на даній ділянці: $\Delta t_ì^* = t_ì - t_{\min}$,

$\Delta t_m^* = t_m - t_{\min}$, $\Delta t_c^* = t_c - t_{\min}$, а також відносні відхилення: $\tau_M^* = \dfrac{\Delta t_M^*}{t_{\min}}$, $\tau_m^* = \dfrac{\Delta t_m^*}{t_{\min}}$, $\tau_c^* = \dfrac{\Delta t_c^*}{t_{\min}}$. Закон розподілу фактичного відхилення Δt^* генеральної сукупності є дискретним.

Але, враховуючи, що об'єм генеральної сукупності (кількість транспортних одиниць) є досить велике число, доцільно вважати, що Δt^* асимптотично прямує до деякої неперервної випадкової величини. Дійсно, оскільки об'єм вантажоперевезень не є сталою величиною, то для дискретних процесів кожного разу доведеться проводити аналіз із самого початку.

Розглянемо припущення (гіпотезу): функція розподілу Δt^* асимптотично прямує до функції деякого відомого, наперед заданого, розподілу. Такі гіпотези називають асимптотично непараметричними. Але, оскільки функції розподілу залежать від певної кількості параметрів, то необхідно одночасно перевірити гіпотезу про вид закону розподілу та значущість його параметрів. Таким чином, виникає задача перевірки складної гіпотези:

$$H_0: \quad F(\Delta t^*) = F_0(\Delta t^*, \tilde{\theta}(\Delta t^*)), \quad \theta \in \Theta \subset R^d.$$

де $F(\Delta t^*)$ – функція розподілу ознаки генеральної сукупності, $F_0(\Delta t^*, \tilde{\theta}(\Delta t^*))$ теоретична функція розподілу, що залежить від d -вимірного вектора $\tilde{\theta}(\Delta t^*)$ – оцінок теоретичного вектора-параметра θ.

Задача перевірки цієї гіпотези розпадається на два етапи. На першому нам необхідно отримати оцінки параметрів відомого розподілу. Оскільки у випадку вантажних перевезень ми можемо оперувати досить великим обсягом вибіркових даних відносно t_m, t_m, t_c, то пропонуємо для цього спрощену методику, що може бути застосована до дво- або трипараметричних моделей.

Знаходимо моду, медіану та математичне сподівання для гіпотетичної випадкової величини і прирівнюємо до відповідної вибіркової. Залежно від того, скільки параметрів входить до щільності розподілу, розглядаємо систему двох або трьох рівнянь для визначення оцінок невідомих параметрів. Дана методика є аналогом методу моментів, але не вимагає обчислення вибіркових моментів вищих порядків, а базується лише на даних статистичної звітності.

На другому етапі перевіряємо гіпотезу про вид розподілу. Через те, що обсяг вибірки може бути досить великим, ми можемо скористатись критерієм Пірсона. Для цього розбиваємо числову піввісь $\Delta t^* \geq 0$ на r проміжків. Одразу зауважимо, що довжини проміжків можуть бути різними, причому останній з них – $[\Delta t_{r-1}^*; +\infty)$. Кількість і довжина кожного інтервалу визначається, виходячи з нормативних документів та правил перевезень, а також з умови, що частота потрапляння в певний інтервал $[\Delta t_{k-1}^*; \Delta t_k^*]$ повинна бути не меншою п'яти ($n_k \geq 5$). Тоді випадкова величина:

$$\chi^2_{cn} = \sum_{k=1}^{r} \frac{\left(n_k - np_k\left(\tilde{\theta}\right)\right)^2}{np_k\left(\tilde{\theta}\right)},$$

де

$$p_k\left(\tilde{\theta}\right) = F_0(\Delta t_k^*;\tilde{\theta}) - F_0(\Delta t_{k-1}^*;\tilde{\theta}) = \int_{\Delta t_{k-1}}^{\Delta t_k} f_0(\Delta t^*;\tilde{\theta}) d\Delta t^*,$$

розподілена за законом *хі*-квадрат з $r-d-1$ степенями свободи. Порівнюємо її з табличним значенням $\chi^2(1-\alpha, r-d-1)$, де α – рівень значущості.

Якщо $\chi^2_{cn} < \chi^2$, то не має підстав відхиляти гіпотезу H_0. У випадку $\chi^2_{cn} > \chi^2$ гіпотеза H_0 може бути відхилена на рівні значущості α.

Одним із розподілів, що досить часто використовується при моделюванні випадкових величин, пов'язаних з часовими потоками, є розподіл Вейбулла. Розглянемо застосування описаної методики для перевірки гіпотез про те, що відхилення часу доставки вантажів від мінімального розподілене за цим законом. Щільність розподілу Вейбулла:

$$f(\Delta t^*) = \begin{cases} 0, & \Delta t^* < 0 \\ \alpha\beta(\Delta t^*)^{\beta-1} e^{-\alpha(\Delta t^*)^{\beta}}, & \Delta t^* \geq 0, \quad \alpha > 0, \beta > 1 \end{cases}.$$

Даний розподіл є двопараметричним. Обчислимо для нього моду та медіану. Для знаходження моди продиференціюємо $f(\Delta t^*)$, прирівнявши отриману похідну до нуля. Звідки:

$$MoX = \alpha^{-\frac{1}{\beta}} \left(\frac{\beta-1}{\beta}\right)^{\frac{1}{\beta}}.$$

Розв'язавши рівняння:

$$\int_0^{\Delta t_m^*} \alpha\beta(\Delta t)^{\beta-1} e^{-\alpha(\Delta t^*)^{\beta}} d\Delta t^* = \frac{1}{2},$$

маємо

$$MeX = \alpha^{-\frac{1}{\beta}} (\ln 2)^{\frac{1}{\beta}}.$$

Прирівняємо одержані значення до вибіркових

$$\overline{\Delta t_м^*} = \alpha^{-\frac{1}{\beta}} \left(\frac{\beta-1}{\beta}\right)^{\frac{1}{\beta}}, \qquad \overline{\Delta t_m^*} = \alpha^{-\frac{1}{\beta}} (\ln 2)^{\frac{1}{\beta}}$$

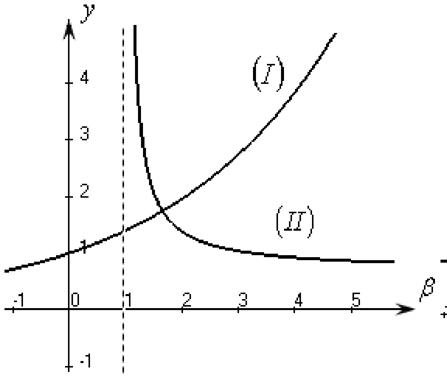
і розглянемо відношення:

$$P = \frac{\overline{\Delta t_m^*}}{\overline{\Delta t_м^*}} = \left(\frac{\beta \ln 2}{\beta - 1}\right)^{\frac{1}{\beta}}$$

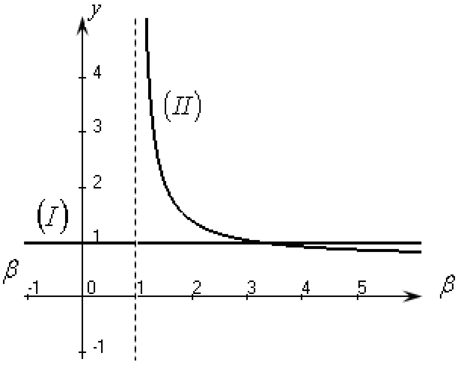
Або, еквівалентне йому

$$P^{\beta} = \ln 2\left(1 + \frac{1}{\beta - 1}\right). \tag{2}$$

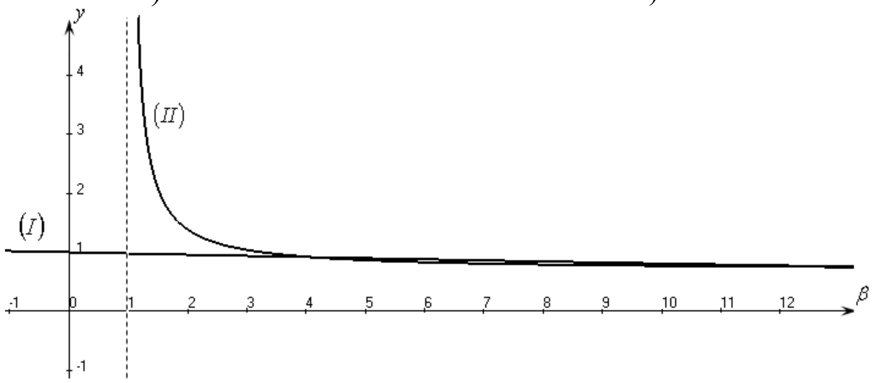
З рисунків 1, а), 1, б) видно, що при $P \geq 1$ ($\Delta t_m \geq \Delta t_м$) це рівняння має один розв'язок: $\beta > 1$. При $0{,}96 < P < 1$ – два розв'язки (Рис. 1в). При $P < 0{,}96$ (рис. 1, г) рівняння розв'язків не має, тому вже на цьому етапі гіпотезу про те, що розподіл відхилення часу доставки від мінімального асимптотично прямує до розподілу



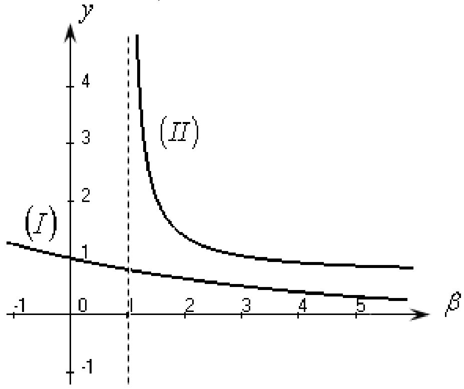
а) $P > 1$



б) $P = 1$



в) $0{,}96 < P < 1$



г) $P < 0{,}96$

Рис. 1. **Графічне дослідження розв'язності рівняння (2):**

(I) – графік функції $y = P^{\beta}$, (II) – графік функції $y = \ln 2\left(1 + \frac{1}{\beta - 1}\right)$.

Вейбулла, відхиляють. У випадку, коли $P \approx 1$, має сенс перевірка про нормальний розподіл відхилень, для якого $MoX = MeX$.

Оцінку параметра β можна знайти наближеними методами з необхідною точністю. Після чого знаходимо:

$$\tilde{\alpha} = \frac{\ln 2}{\left(\overline{\Delta t_m^*}\right)^{\beta}}. \tag{3}$$

Для перевірки гіпотези про розподіл Вейбулла відхилень з параметрами $\tilde{\alpha}$ та $\tilde{\beta}$ необхідно розбити піввісь $\Delta t^* \geq 0$ на r проміжків.

У статистичній звітності відхилення часу доставки фіксуються з точністю до 0,1 доби, штрафи за прострочену доставку, зафіксовані у відповідних документах, збільшуються з інтервалом в одну добу і є незмінними, починаючи з чотиридобового прострочення (див. [5, 6]).

Крім того, кількість інтервалів повинна бути більшою за $\log_2 n - 1$, і для кожного з інтервалів $np_k\left(\tilde{\theta}\right) \geq 5$.

Виходячи з цього, при розбитті часової півосі $\Delta t^* \geq 0$ на проміжки доцільно:

1) при Δt^* близьких до нуля обирати достатньо великий крок від 0,5 до 1 доби, оскільки частка вантажів, що доставляються набагато раніше графікового часу є досить малою;

2) надалі крок розбиття зменшувати, можливо навіть до 0,1 доби в околі моди;

3) при $2 \leq \Delta t^* \leq 4$ крок покласти рівним одній добі, оскільки штрафи стягуються при простроченні доставки на 2 доби, зростають щодоби і є незмінними при простроченні від чотирьох діб;

4) останній інтервал покласти $[4; +\infty)$, враховуючи незмінність штрафів та досить малий обсяг вантажів, що доставляються з відхиленням $\Delta t^* \geq 4$.

Зазначимо, що такий підхід надалі дає можливість замість формули (1) для оцінки витрат користуватись дискретною формулою:

$$C = N \sum_{k=1}^{r} K_k \, p_k\left(\tilde{\theta}\right), \tag{4}$$

оскільки в межах кожного з інтервалів $K_k = K(\Delta t^*) = const$. Крім цього, необхідно враховувати, що при $n = 100,\ r$ повинно бути не менше шести, при $n = 100 - r \geq 9$, при $n = 1000 - r \geq 13$.

Після визначення $p_k\left(\tilde{\theta}\right)$ та n_k обчислюємо χ^2_{cn} і, порівнюючи з $\chi^2(1-\alpha, r-d-1)$, робимо висновок щодо справедливості висунутої гіпотези про те, що розподіл відхилень часу доставки вантажів від мінімального асимптотично прямує до розподілу Вейбулла. Це дає нам підстави оцінювати витрати, пов'язані з несвоєчасною доставкою вантажів за формулою (4).

На завершення зазначимо, що дана методика дає можливість досить просто та досить точно вирішувати питання про адекватність багатопараметричних стохастичних моделей. Одним з її недоліків є те, що на другому етапі вона не є універсальною, оскільки розбиття на інтервали часової півосі проводиться для абсолютних відхилень. Однак на першому етапі для відносних відхилень підхід не змінюється.

Наприклад, якщо для розподілу Вейбулла замість абсолютних взяти відносні відхилення, то усі викладки при заміні Δt^* на τ^* не зміняться. При цьому для функції $f(\tau^*)$ $\tilde{\beta}$ знаходиться із (2), а $\tilde{\alpha} = \frac{\ln 2}{\left(\overline{\tau_m^*}\right)^{\beta}}$.

Тому, маючи універсальну модель для описання розподілу відносних відхилень, за допомогою відповідного перетворення можна перейти до моделі розподілу абсолютних відхилень на конкретній ділянці.

Крім того, даний підхід базується на незалежності відхилень, тобто час доставки для кожної транспортної одиниці не залежить від часу доставки інших одиниць. Однак на практиці несвоєчасна доставка часто фіксується для груп вагонів, кількість вагонів у групах є різною. Це може суттєво вплинути на адекватність відповідної моделі. Даний аспект вимагає подальших досліджень з метою вдосконалення описаної методики.

ЛІТЕРАТУРА

1. *Андрейцев А.Ю.* Стохастична модель оцінки витрат, пов'язаних із відхиленням від графіка доставки вантажів / А.Ю. Андрейцев, Г.С. Висоцька // Проблеми та перспективи розвитку транспортних систем в умовах реформування залізничного транспорту: управління, економіка і технології: Матеріали VI Міжнародної науково-практичної конференції. – Сер. «Техніка, технологія». – Київ: ДЕТУТ, 2013. – с. 210 – 211
2. *Висоцька Г.С.* Визначення імовірнісних характеристик процесу доставки вантажів / Г.С. Висоцька // Зб. наук. праць ДЕТУТ. Серія «Транспортні системи і технології». – К., 2012. – № 20. – С. 240 – 245
3. *Myronenko V.K.* Evaluation of stochastic characteristics of goods delivery schedules by rail / Myronenko V.K., Andreytsev A.Yu., Vysotska G.S.// Зб. наук. праць ДЕТУТ. Серія «Транспортні системи і технології». – Київ, 2014. – № 24. – С. 157 – 162
4. *Андрейцев А.Ю.* Стохастическая модель оценки своевременности доставки грузов / А.Ю. Андрейцев, Г.С. Высоцкая // Сучасні енергетичні установки на транспорті, технології та обладнання для їх обслуговування: V Міжнар. наук.-практ. конф., 1-3 жовтня 2014 р.: тези доп. – Херсон: ХДМА, 2014. – С. 128 – 130
5. Статут залізниць України [Електронний ресурс]. Режим доступу: http://zakon1.rada.gov.ua/laws/show/457-98-%D0%BF.
6. Соглашение о международном железнодорожном грузовом сообщении (СМГС) [Електронний ресурс]. Режим доступу: http://www.uz.gov.ua/files/file/cargo_transportation/smgs/SMGS_01.07.2013_aktual.2014.pdf

***Andrij Ju. Andreytsev, PhD (Physical and Math. Sciences),** Associate Professor*

(Associate Professor of High Mathematics Chair, State University for Transport Economy and Technologies)

Tetiana S. Kletska, PhD (Historical Sciences)

(Associate Professor of High Mathematics Chair, State University for Transport Economy and Technologies)

***Oleksandr O. Kilchynskyi, PhD (Physical and Math. Sciences),** Associate Professor*

(Associate Professor of High Mathematics Chair, State University for Transport Economy and Technologies)

ON IMPROVING THE TECHNIQUE OF VERIFICATION COMPLEX STATISTICAL HYPOTHESES IN THE ANALYSIS OF GOODS DELIVERY TIMELINESS

Verifying the adequacy of mathematical models is a prerequisite for their use in predicting the real processes. Most of the models that describe the technological

processes are stochastic. In this case, the statistical hypotheses verifying is of high topicality.

The paper deals with the questions of statistical hypotheses verifying, arising in problems of timely goods delivery. It is assumed that the distribution of the real-time delivery deviations from the minimum one asymptotically converges to a preset continuous distribution. The problem is divided into two stages. The first step is to find the parameters estimates of population distribution on the basis of statistical data. The second step is to verify a complex statistical hypothesis on the form of the distribution using the non-parametrical Pearson's test.

The approach to verify the hypothesis on the distribution of deviations asymptotically converges to the Weibull distribution has been demonstrated. The parameters estimation of the two-parameter distribution has been carried out using the sample values of mode and median of the real-time delivery deviations from the minimum one. The system of equations for determining the distribution of the parameter estimates has been written. Its solvability has been analyzed on the basis of the graphical representation. The recommendations for partitioning a half-axis into intervals have been given. The number of intervals of the partition has been determined from the statistical formulas. The length of the interval has been determined from the group data according to regulatory documents on rail transportation.

Conclusions about the possibility of using the described approach for absolute and relative deviations have been made.

Keywords: *time of goods delivery, actual delivery time, minimum (technological) delivery time, stochastic model, statistical hypothesis, the Weibull distribution*

REFERENCES

1. *Andreytsev A. Yu*. Stokhastychna model otsinky vytrat, pov'iazanykh iz vidkhylenniam vid hrafika dostavky vantazhiv [Stochastic model of costs evaluation caused by deviations of goods delivery from the schedules] / A. Yu. Andreytsev, G. S. Vysotska // Problemy ta perspektyvy rozvytku transportnykh system v umovakh reformuvannia zaliznychnoho transportu: upravlinnia, ekonomika i tekhnolohii: Materialy VI Mizhnarodnoi naukovo-praktychnoi konferentsii. – Seriia «Tekhnika, tekhnolohiia». – K.: DETUT [State Economy and Technology University of Transport], 2013. – P. 210 – 211.

2. *Vysotska G. S.* Vyznachennia imovirnisnykh kharakterystyk protsesu dostavky vantazhiv [Definition of probabilistic characteristics of the process of goods delivery] / G. S. Vysotska // Zbirnyk naukovyh prats DETUT [Proceedings of the State Economy and Technology University of Transport]. Seriia «Transportni systemy i tekhnolohii». – Kyiv, 2012. – issue 20. – P. 240 – 245.

3. *Myronenko V. K.* Evaluation of stochastic characteristics of goods delivery schedules by rail / Myronenko V. K., Andreytsev A. Yu., Vysotska G. S.// Zbirnyk naukovyh prats DETUT [Proceedings of the State Economy and Technology University of Transport]. Seriia «Transportni systemy i tekhnolohii». – Kyiv, 2014. – issue 24. – P. 157 – 162.

4. *Andreytsev A. Yu*. Stokhastycheskaia model otsenky svoevremennosty dostavky hruzov [Stochastic model for evaluation of goods delivery timeliness] / A. Yu. Andreytsev, G. S. Vysotska // Suchasni enerhetychni ustanovky na transporti, tekhnolohii ta obladnannia dlia yikh obsluhovuvannia: 5-ta mizhnar. nauk.-prakt. konf., 1-3 zhovtnia 2014 r.: tezy dop. [Proc. of the 5th Int. Scientific and Practical Conf. «Modern power plants in transport, technologies and maintenance equipment» Oktober 01-03. 2014]. – Kherson: KhDMA [Kherson State Maritime Academy], 2014. – P. 128 – 130.

5. Statut zaliznyts Ukrainy [Charter for the Ukrainian railways]. Rezhym dostupu: http://zakon1.rada.gov.ua/laws/show/457-98-%D0%BF.

6. Sohlashenye o mezhdunarodnom zheleznodorozhnom hruzovom soobshchenyy (SMHS) [Agreement on international rail freight traffic]. Rezhym dostupu:

http://www.uz.gov.ua/files/file/cargo_transportation/smgs/SMGS_01.07.2013_aktual.2014.pdf.