

DOI: 10.36910/6775-2524-0560-2020-40-12

УДК: 004.415.3

Пех Петро Антонович, к.т.н., доцент

<https://orcid.org/0000-0002-6327-3319>

Грабинський Богдан Петрович, студент

Луцький національний технічний університет

ПЕРЕВІРКА ГІПОТЕЗИ ПРО ХАРАКТЕР ТА ПАРАМЕТРИ РОЗПОДІЛУ ВИПАДКОВИХ ВЕЛИЧИН**Пех П. А., Грабинський Б. П. Перевірка гіпотези про характер та параметри розподілу випадкових величин.**

В статті наведена методика побудови математичної моделі та перевірки гіпотези про тип та параметри розподілу вхідного потоку та часу обслуговування вимог систем масового обслуговування.

Ключові слова: вхідний потік вимог, час обслуговування вимог, система масового обслуговування, найпростіший потік, розподіл Ерланга**Пех П. А., Грабинский Б. П. Проверка гипотезы о характере и параметрах распределения случайных величин.**

В статье приведена методика построения математической модели и проверки гипотезы о характере и параметрах распределения входящего потока и времени обслуживания требований систем массового обслуживания.

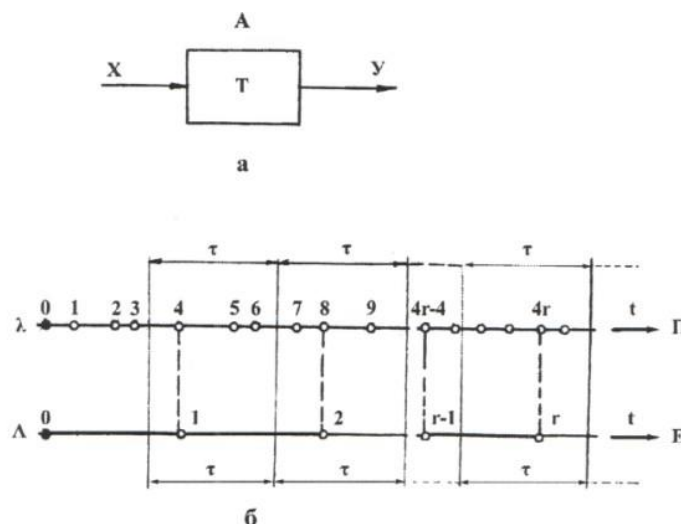
Ключевые слова: входящий поток требований, время обслуживания требований, система массового обслуживания, простейший поток, распределение Эрланга**Pekh Petro, Hrabynskiy Bohdan. Testing the hypothesis about the nature and parameters of the distribution of random variables.** The article presents a method of building a mathematical model and testing the hypothesis about the type and parameters of distribution of input flow and service time requirements of queuing systems.**Keywords:** incoming requirements flow, requirements service time, queuing system, simplest flow, Erlang distribution.**Постановка задачі.** Дискретні випадкові величини часто використовуються у дослідженнях процесів функціонування систем, що відносяться до області комп'ютерної інженерії. Типовим прикладом таких випадкових величин є кількість вимог, що надходять до систем масового обслуговування (СМО) за певний проміжок часу, наприклад, кількість запитів до сервера комп'ютерної мережі [1].Кожну СМО можна розглядати як обслуговуючий апарат A , що функціонує у відповідності з вимогами технології (рис. 1, а). Вхідний потік вимог X , які поступають на обслуговування до апарату A , і випадковий процес їх обслуговування T породжують вихідний потік обслужених вимог Y , а це дає можливість прогнозувати показники якості роботи СМО.

Рисунок 1. а - Структура системи масового обслуговування;

б – Схема утворення потоку з післядією E шляхом k -кратного розрідження найпростішого потоку Π Розглянемо процес надходження вимог до СМО. Відмітимо на осі Ot точками моменти надходження вимог (рис. 1, б). Тоді віддаль між сусідніми точками являє собою тривалість проміжку часу між моментами надходження вимог одна за одною.Кількість вимог, що надходять протягом певного інтервалу часу τ будемо вважати випадковою величиною R . Імовірність того, що за час τ у систему надійде рівно $R = r$ вимог, позначимо $P_r(\tau)$.

обслуговуючого апарату системи, а віддаль між сусідніми точками – це не що інше, як тривалість процесу обслуговування вимог.

Введемо у розгляд випадкову величину T – тривалість часу обслуговування вимоги – це проміжок часу між моментами надходження вимог на обслуговування. Ця випадкова величина є неперервною. Для неї можуть бути знайдені диференціальна $f(t)$ та інтегральна $F(t)$ функції розподілу імовірностей.

Диференціальна функція $f(t)$ розподілу тривалості часу обслуговування вимог у випадку найпростішого характеру процесу обслуговування має вигляд:

$$f(t) = \mu e^{-\mu t}. \quad (9)$$

де μ – інтенсивність процесу обслуговування, що має найпростіший характер – часто у цьому випадку кажуть, що час обслуговування є експоненціальним.

Інтегральна функція $F(t)$ розподілу імовірностей тривалості часу обслуговування вимог у випадку найпростішого характеру процесу обслуговування має вигляд:

$$F(t) = 1 - e^{-\mu t}. \quad (10)$$

Зауважимо, що інтегральна функція $F(t)$ – це імовірність того, що випадкова величина T тривалості часу обслуговування вимог прийме значення, яке буде менше або дорівнюватиме деякому значенню t , тобто $F(t) = P(T \leq t)$.

Математичне сподівання M_t та дисперсія D_t тривалості часу обслуговування у найпростішому процесі обслуговування вимог відповідно дорівнюють:

$$M_t = \frac{1}{\mu}; \quad (11)$$

$$D_t = \frac{1}{\mu^2}. \quad (12)$$

Рівність

$$\frac{M_t^2}{D_t} = 1 \quad (13)$$

є ознакою того, що процес обслуговування вимог є найпростішим.

Інтенсивність μ_E Ерлангівського процесу обслуговування можна визначити за формулою:

$$\mu_E = \frac{\mu}{k}. \quad (14)$$

Диференціальна $f^k(t)$ функція розподілу тривалості часу обслуговування вимог у випадку Ерлангівського характеру процесу обслуговування має вигляд:

$$f^k(t) = \frac{k\mu_E(k\mu_E t)^{k-1}}{(k-1)!} e^{-k\mu_E t}. \quad (15)$$

Інтегральна функція розподілу імовірностей тривалості часу обслуговування вимог у випадку Ерлангівського характеру процесу обслуговування має вигляд:

$$F^k(t) = 1 - e^{-k\mu_E t} \sum_{n=0}^{k-1} \frac{(k\mu_E t)^n}{n!}. \quad (16)$$

Математичне сподівання M_t^k та дисперсія D_t^k тривалості часу обслуговування у Ерлангівському процесі обслуговування відповідно дорівнюють:

$$M_t^k = \frac{1}{\mu_E}; \quad (17)$$

$$D_t^k = \frac{1}{k\mu_E^2}. \quad (18)$$

Параметр стабільності k процесу Ерланга можна визначити за формулою:

$$k = \frac{(M_t^k)^2}{D_t^k}. \quad (19)$$

Ознакою того, що процес обслуговування вимог дійсно є Ерлангівським, є виконання нерівності $k > 1$. Чим більше значення цього параметру, тим стабільнішим є процес обслуговування, а характер його стає все ближчим до детермінованого.

Основна частина. Перевірка гіпотези про характер вхідного потоку вимог. Нехай ми маємо можливість вести спостереження за роботою СМО і експериментальним шляхом визначити кількість вимог, що надходять до системи за проміжки часу конкретної величини. Заданою метою перевірити гіпотезу про характер вхідного потоку вимог і обчислити його параметри. Нижче у табл. 1 наведені дані $n=100$ спостережень кількості надходжень вимог до СМО протягом п'ятихвилинних проміжків часу ($\tau = 5$ хв.). Для вирішення поставленої задачі необхідно виконати наступне.

1. Побудувати емпіричний розподіл частот надходження вимог у кількості r_i штук протягом п'ятихвилинних проміжків часу.

2. Визначити середню кількість вимог r_c , що надходять у СМО протягом п'ятихвилинних проміжків часу:

$$r_c = \frac{\sum r_i n_i}{n}. \quad (20)$$

3. Визначити інтенсивність вхідного потоку вимог, що надходять у СМО протягом п'ятихвилинних проміжків часу:

$$\lambda = \frac{1}{r_c}. \quad (21)$$

Таблиця 1. Дані $n=100$ спостережень кількості надходжень вимог до СМО протягом п'ятихвилинних проміжків часу

i	r_i	i	r_i	i	r_i	i	r_i
1	3	26	2	51	1	76	8
2	4	27	4	52	6	77	6
3	8	28	3	53	4	78	2
4	1	29	7	54	3	79	5
5	4	30	5	55	7	80	3
6	3	31	7	56	5	81	8
7	5	32	3	57	4	82	4
8	2	33	4	58	2	83	8
9	4	34	8	59	6	84	2
10	5	35	6	60	10	85	9
11	5	36	9	61	0	86	0
12	3	37	4	62	6	87	4
13	8	38	8	63	7	88	7
14	4	39	6	64	4	89	3
15	2	40	5	65	9	90	7
16	6	41	7	66	6	91	5
17	5	42	2	67	2	92	1
18	1	43	5	68	14	93	6
19	4	44	4	69	5	94	4
20	3	45	7	70	5	95	5
21	5	46	3	71	6	96	6
22	4	47	7	72	6	97	5
23	3	48	5	73	3	98	3
24	6	49	6	74	9	99	7
25	6	50	5	75	5	100	4

4. Визначити дисперсію кількості вимог, що надходять у СМО протягом п'ятихвилинних проміжків часу:

$$D = \frac{\sum (r_i - r_c)^2 n_i}{n}. \quad (22)$$

5. Визначити середнє квадратичне відхилення кількості вимог, що надходять у СМО протягом п'ятихвилинних проміжків часу:

$$\sigma = \sqrt{D}. \quad (23)$$

6. Визначити відносні частоти (частоті) кількості вимог, що надходять у СМО протягом п'ятихвилинних проміжків часу:

$$\gamma_i = \frac{n_i}{n}. \quad (24)$$

7. Визначити теоретичні імовірності кількості вимог, що надходять у СМО протягом п'ятихвилинних проміжків часу:

$$P_r^k(\tau) = e^{-k\tau/r_c} \sum_{n=kr}^{kr+k-1} \frac{(k\tau/r_c)^n}{n!}, \quad r = 0, 1, 2, \dots \quad (25)$$

Розглянути випадки: $k=1$; $k=2$; $k=3$.

8. Визначити теоретичні частоти кількості вимог, що надходять у СМО протягом п'ятихвилинних проміжків часу:

$$n_{mi} = n * P_r^k(\tau). \quad (26)$$

9. Визначити величину χ^2_{ϕ} , що характеризує відхилення теоретичних частот кількості вимог, що надходять у СМО протягом п'ятихвилинних проміжків часу, від емпіричних значень:

$$\chi^2_{\phi} = \sum \frac{(n_i - n_{mi})^2}{n}. \quad (27)$$

10. Порівняти отримані значення χ^2_{ϕ} з табличним значенням $\chi^2_{табл}$, прийнявши рівень значущості $\alpha = 0.05$. Число степенів свободи визначити за формулою

$$f = l - s - 1 \quad (28)$$

де l – кількість інтервалів після об'єднання деяких з них (об'єднувати можна ті крайні інтервали, для яких є дуже малою величиною);

S - кількість додатково накладених зв'язків.

A	B	C	D	E	F	G	H	I	J	K	L
	i	n _i	n _i	n _i /n	n _i /r _c	(n _i -r _c) ²	n/n	P ₁ (k=1)	n*P ₁	(n _i -n*P ₁) ² /(n*P ₁)	
104											
105	1	0	2	0	-4,93	24,30490	0,02	0,00722650	0,722650000000	2,257833007	
106	2	1	4	4	-3,93	15,44490	0,04	0,03662666	3,662670000000	0,036883762	
107	3	2	8	16	-2,93	8,58490	0,08	0,08781972	8,781970000000	0,069628692	
108	4	3	13	39	-1,93	3,72490	0,13	0,14431707	14,431710000000	0,142034002	
109	5	4	17	68	-0,93	0,86490	0,17	0,17787079	17,787080000000	0,034828366	
110	6	5	18	90	0,07	0,00490	0,18	0,17538060	17,538060000000	0,012167177	
111	7	6	15	90	1,07	1,14490	0,15	0,14410439	14,410440000000	0,024120062	
112	8	7	10	70	2,07	4,28490	0,10	0,10149067	10,149070000000	0,002189547	
113	9	8	7	56	3,07	9,42490	0,07	0,06254362	6,254360000000	0,088894629	
114	10	9	4	36	4,07	16,56490	0,04	0,03426001	3,426000000000	0,096169294	
115	11	10	1	10	5,07	25,70490	0,01	0,01689018	1,689020000000	0,261079301	
116	12	11	0	0	6,07	36,84490	0,00	0,00756987	0,756990000000	0,75699	
117	13	12	0	0	7,07	49,98490	0,00	0,00310996	0,311000000000	0,311	
118	14	13	0	0	8,07	65,12490	0,00	0,00117939	0,117940000000	0,11794	
119		14	1	14	9,07	82,26490	0,01	0,00041531	0,041530000000		
120			100	493		344,27360	1,00		99,980490000000	4,24856785	
121				t _c = 4,93	D _r =	3,47751					
122					σ _r =	1,86481					
123											
124											
125											
126											
127											

Рисунок 2 - Розрахунок параметрів вхідного потоку вимог до СМО за кількістю вимог, що надходять до системи протягом п'ятихвилинних проміжків часу і апроксимація його найпростішим потоком з параметром стабільності k=1

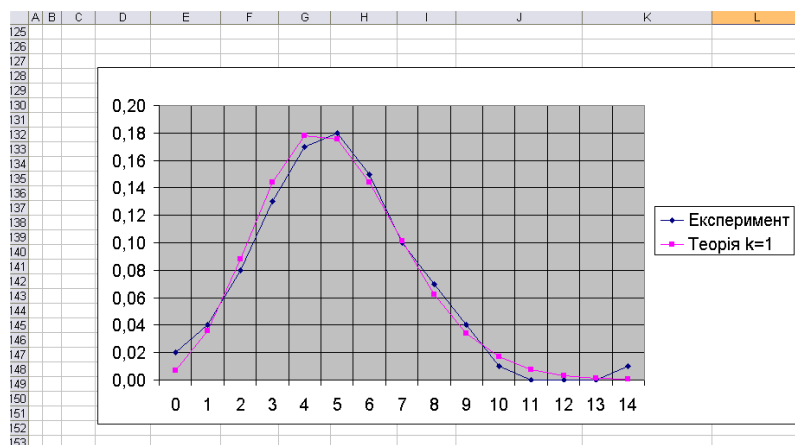


Рисунок 3 - Графіки розподілу частот та імовірностей надходження вимог до системи протягом п'ятихвилинних проміжків часу у випадку k=1

Перевірка гіпотези про характер розподілу часу обслуговування вимог та визначення його параметрів

Нижче у табл. 2 наведені дані $n=100$ спостережень тривалості часу обслуговування вимог у СМО. Необхідно дослідити характер і обчислити параметри процесу обслуговування вимог, для чого необхідно виконати наступне.

1. Побудувати емпіричний розподіл частот тривалості часу обслуговування t_i вимог за частотами їх появи n_i .

2. Визначити середнє значення t_c тривалості часу обслуговування вимог:

$$t_c = \frac{\sum t_i n_i}{n}. \quad (29)$$

3. Визначити інтенсивність μ_E процесу обслуговування вимог у СМО:

$$\mu_E = \frac{1}{t_c}. \quad (30)$$

4. Визначити дисперсію D тривалості часу обслуговування вимог:

$$D = \frac{\sum (r_i - r_c)^2 n_i}{n}. \quad (31)$$

5. Визначити середнє квадратичне відхилення σ тривалості часу обслуговування вимог у СМО:

$$\sigma = \sqrt{D}. \quad (32)$$

Таблиця 2. Дані $n=100$ спостережень тривалості часу t_i обслуговування вимог у СМО

i	t_i	i	t_i	i	t_i	i	t_i
1	5	26	96	51	16	76	147
2	172	27	14	52	176	77	24
3	17	28	160	53	12	78	62
4	56	29	22	54	34	79	38
5	13	30	35	55	72	80	11
6	128	31	108	56	37	81	182
7	26	32	28	57	46	82	30
8	66	33	60	58	6	83	48
9	198	34	18	59	64	84	118
10	44	35	99	60	203	85	10
11	52	36	50	61	19	86	69
12	7	37	73	62	76	87	140
13	82	38	88	63	33	88	78
14	32	39	9	64	90	89	5
15	18	40	29	65	19	90	93
16	42	41	83	66	72	91	16
17	58	42	23	67	13	92	103
18	6	43	36	68	37	93	38
19	54	44	50	69	45	94	58
20	25	45	10	70	32	95	10
21	67	46	30	71	113	96	30
22	40	47	83	72	130	97	49
23	86	48	24	73	9	98	16
24	7	49	188	74	116	99	124
25	17	50	8	75	21	100	146

6. Визначити параметр стабільності процесу обслуговування вимог у СМО:

$$k = \frac{t_c^2}{D}. \quad (33)$$

7. Визначити емпіричні відносні частоти (частоти) γ_i тривалості часу обслуговування вимог у СМО:

$$\gamma_i = \frac{n_i}{n}. \quad (34)$$

8. Визначити емпіричні накопичені частоти γ_i тривалості часу обслуговування вимог у СМО:

$$\gamma_{ki} = \sum_{i=1}^i n_i. \quad (35)$$

9. Визначити емпіричні накопичені частоти θ_{ki} тривалості часу обслуговування вимог у СМО:

$$\theta_{ki} = \frac{\gamma_{ki}}{n}. \quad (36)$$

10. Визначити теоретичні імовірності тривалості часу обслуговування вимог у СМО:

$$F^k(T \leq t_i) = 1 - e^{-\frac{kt_i}{t_c}} \sum_{n=1}^{k-1} \frac{(\frac{kt_i}{t_c})^n}{n!}. \quad (37)$$

Розглянути випадки: $k=1$; $k=2$; $k=3$.

11. Визначити теоретичні накопичені частоти тривалості часу обслуговування вимог у СМО:

$$n_{kmi} = n * F^k(T \leq t_i). \quad (38)$$

12. Визначити теоретичні частоти тривалості часу обслуговування вимог у СМО:

$$n_{mi} = n_{km(i-1)} - n_{kmi}. \quad (28)$$

13. Визначити величину χ_{ϕ}^2 , що характеризує відхилення теоретичних частот тривалості часу обслуговування вимог у СМО, від відповідних емпіричних значень:

$$\chi_{\phi}^2 = \sum \frac{(n_i - n_{mi})^2}{n}. \quad (39)$$

14. Порівняти отримані значення χ_{ϕ}^2 з табличним значенням $\chi_{табл}^2$, прийнявши рівень значущості $\alpha = 0.05$. Число степенів свободи визначити за формулою

$$f = l - s - 1 \quad (40)$$

де l – кількість інтервалів після об'єднання деяких з них (об'єднувати можна ті крайні інтервали, для яких є дуже малою величиною);

s – кількість додатково накладених зв'язків.

A	B	C	D	E	F	G	H	I	J	K	L
		i	t _i	n _i	n _i *t _i	t _i -t _c	(t _i -t _c) ²	n _i /n	P _{t_i} (k=1)	n*P _{t_i}	(n _i -n*P _{t_i}) ² /(n*P _{t_i})
104											
105		1	0	2	0	-4,93	24,30490	0,02	0,00722650	0,722650000000	2,257833007
106		2	1	4	4	-3,93	15,44490	0,04	0,03562666	3,562670000000	0,053683762
107		3	2	8	16	-2,93	8,58490	0,08	0,08781972	8,781970000000	0,069628692
108		4	3	13	39	-1,93	3,72490	0,13	0,14431707	14,431710000000	0,142034002
109		5	4	17	68	-0,93	0,86490	0,17	0,17787079	17,787080000000	0,034828366
110		6	5	18	90	0,07	0,00490	0,18	0,17538060	17,538060000000	0,01216717
111		7	6	15	90	1,07	1,14490	0,15	0,14410439	14,410440000000	0,024120082
112		8	7	10	70	2,07	4,28490	0,10	0,10149067	10,149070000000	0,002189547
113		9	8	7	56	3,07	9,42490	0,07	0,06254362	6,254360000000	0,088894629
114		10	9	4	36	4,07	16,56490	0,04	0,03426001	3,426000000000	0,096169294
115		11	10	1	10	5,07	25,70490	0,01	0,01689018	1,689020000000	0,281079301
116		12	11	0	0	6,07	36,84490	0,00	0,00756987	0,756990000000	0,75699
117		13	12	0	0	7,07	49,98490	0,00	0,00310996	0,311000000000	0,311
118		14	13	0	0	8,07	65,12490	0,00	0,00117939	0,117940000000	0,11794
119			14	1	14	9,07	82,26490	0,01	0,00041531	0,041530000000	
120				100	493		344,27350	1,00		99,980490000000	4,24855785
121				t _c =	4,93	D _r =	3,47751				
122						σ _r =	1,86481				
123											
124											
125											
126											
127											

Рисунок 4 - Розрахунок параметрів процесу обслуговування вимог у СМО за тривалістю часу обслуговування вимог у СМО і апроксимація його найпростішим процесом з параметром стабільності $k=1$

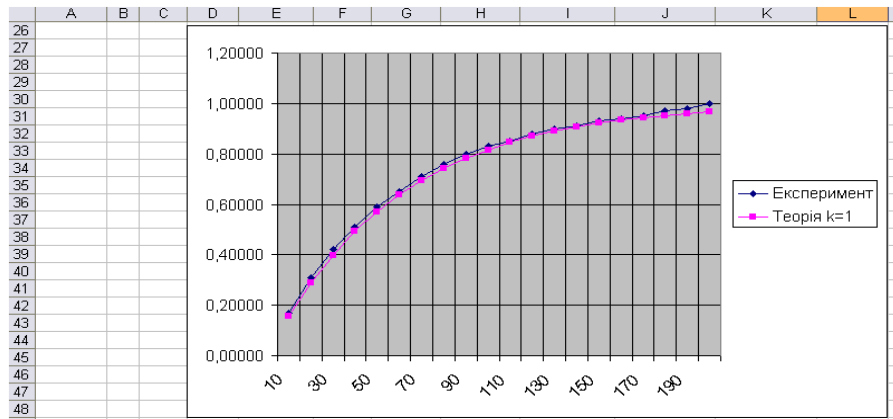


Рисунок 5 - Графіки розподілу частот та імовірностей тривалості часу обслуговування вимог у СМО у випадку $k=1$

Висновки.

В роботі наведені методики визначення:

- типу та параметрів розподілу імовірностей надходження кількості вимог (дискретної випадкової величини) до системи масового обслуговування протягом проміжків часу певної величини;
- типу та параметрів розподілу імовірностей розподілу тривалості часу обслуговування вимог (неперервної випадкової величини) у системах масового обслуговування.

Список бібліографічного опису.

1. Пех П.А., Баландюх В.В. Розробка інтернет-магазину з продажу косметики та засобів гігієни на основі CVS WordPress. Науковий журнал "Комп'ютерно-інтегровані технології: освіта, наука, виробництво" - Луцьк: Видавництво ЛНТУ. – Вип. 26. – 2017. – С. 154-158.
2. Пех П.А., Бортник К.Я., Яручик О.В. Програмний C++ комплекс та результати дослідження швидкодії алгоритмів сортування методом імітаційного моделювання. Науковий журнал "Комп'ютерно-інтегровані технології: освіта, наука, виробництво" - Луцьк: Видавництво ЛНТУ. – Вип. 27. – 2017. – С. 54-59.
3. Пех П.А., Войтович А.О. Програмний C++ комплекс для дослідження процесу функціонування автоматизованих ліній та його верифікація. Науковий журнал "Комп'ютерно-інтегровані технології: освіта, наука, виробництво" - Луцьк: Видавництво ЛНТУ. – Вип. 27. – 2017. – С. 60-66.
4. Пех П.А., Корець Р.С. Реалізація основної функції калькулятора (обчислення виразів) за допомогою Java-класів. Науковий журнал "Комп'ютерно-інтегровані технології: освіта, наука, виробництво" - Луцьк: Видавництво ЛНТУ. – Вип. 35. – 2019. – С. 166-171.
5. Пех П.А., Яковлюк С.М. Створення колоди карток з програмування мовою C++ засобами програми Anki. Науковий журнал "Комп'ютерно-інтегровані технології: освіта, наука, виробництво" - Луцьк: Видавництво ЛНТУ. – Вип. 35. – 2019. – С. 172-177..
6. Пех П.А., Костюк Ю.Ю., Кравченко М.Б. До питання конструювання класів з конструкторами різного типу. Науковий журнал "Комп'ютерно-інтегровані технології: освіта, наука, виробництво" - Луцьк: Видавництво ЛНТУ. – Вип. 35. – 2019. – С. 197-202.

References.

1. Pekh Petro, Balandyukh Vladimir. Development of an online store for the sale of cosmetics and hygiene products based on CMS WordPress. Scientific journal "Computer-integrated technologies: education, science, production" - Lutsk: LNTU Publishing House. - №26. - 2017. - P. 154-158.
2. Pekh Petro, Bortnik Katerina, Yaruchik Olexander. The C ++ program complex and results of sorting algorithms speed study by the simulsion method. Scientific journal "Computer-integrated technologies: education, science, production" - Lutsk: LNTU Publishing House. - №27. - 2017. - P. 54-59.
3. Pekh Petro, Voitovich Andriy. Software C++ complex for research of an automated lines functioning and it's verification. Scientific journal "Computer-integrated technologies: education, science, production" - Lutsk: LNTU Publishing House. - №27. - 2017. - P. 60-66.
4. Pekh Petro, Korets Roman. Implementation of the main calculator function (expression calculation) be means of Java-classes. Scientific journal "Computer-integrated technologies: education, science, production" - Lutsk: LNTU Publishing House. - №35. - 2019. - P. 166-171.
5. Pekh Petro, Kostyuk Yuriy, Kravchenko Maxim. On the issue of designing classes with cjnstructors of different types. Scientific journal "Computer-integrated technologies: education, science, production" - Lutsk: LNTU Publishing House. - № 35. - 2019. - P. 197-202.