

Федеральное государственное бюджетное образовательное
учреждение высшего профессионального образования Московский
государственный технический университет имени Н.Э. Баумана

Лабораторная работа №4
«Оценка полученного дохода Пражского метро»
по курсу
«Моделирование»

Студент группы ИУ9-82

Белогуров А.А.

Преподаватель

Домрачева А.Б.

Москва, 2018

Содержание

1	Цель работы	3
2	Постановка задачи	4
2.1	Задача 1	5
2.2	Задача 2	5
2.3	Задача 3	9
3	Теоретические сведения	10
3.1	Задача 1	11
3.2	Задача 2	13
3.3	Задача 3	13
4	Практическая реализация	14
5	Результаты	17
6	Вывод	18

1 Цель работы

Ознакомиться с математической моделью, построенной на основе статистических данных и проанализировать полученную модель.

2 Постановка задачи

Имеются статистические данные пражского метро за 2014 год. Данные представлены в трёх таблицах и отражают информацию о:

- $th\ X$ - кол-во людей, вошедших на станцию X в течение дня, как среднестатистическое за месяц;
- $r\ X$ - кол-во прошедших через турникеты на станции X в течение дня, как среднестатистическое за месяц;
- $II\ X$ кол-во купленных билетов на станции X за месяц;
- Общее количество станций - 61.

Билеты распределены по трем категориям:

- F - взрослый билет, стоимость:
 1. 30 мин - 24 крон
 2. 90 мин - 32 крон
 3. 24 часа - 110 крон
 4. 72 часа - 310 крон
- D - детский билет, стоимость:
 1. 30 мин - 12 крон
 2. 90 мин - 16 крон
 3. 24 часа - 55 крон
 4. 72 часа - 310 крон
- L - льготный билет, стоимость (аналогично десткому).

2.1 Задача 1

Проверить, насколько тесна связь между потоком пассажиров вошедших в метро, и пассажиров прошедших через турникеты.

m/s	thA0	rA0	thA1	rA1	thB0	rB0	thB1	rB1	thC0	rC0	thC1	rC1
1	16,551	14,899	30,746	27,320	32,822	29,553	21,002	18,793	17,084	15,365	4,544	3,118
2	16,810	14,292	22,558	20,155	25,314	22,567	40,022	35,436	29,096	25,876	17,519	16,162
3	14,434	13,046	28,001	24,916	36,918	32,720	35,118	31,145	38,639	34,226	38,841	34,819
4	20,891	18,696	32,958	29,255	46,677	41,259	20,283	18,164	23,690	21,145	37,324	33,492
5	13,773	12,468	28,277	25,159	16,909	15,212	41,746	36,944	29,087	25,868	16,717	15,461
6	14,739	13,313	36,763	32,398	21,889	19,569	40,458	35,817	21,993	20,494	40,099	35,920
7	24,713	22,040	34,650	30,735	34,998	31,040	19,478	17,460	30,082	26,738	42,244	37,797
8	10,127	9,278	33,590	29,808	23,285	20,791	22,974	21,353	18,776	17,263	22,099	20,170
9	14,689	13,269	12,239	11,126	21,561	19,282	25,348	23,430	34,808	31,290	40,895	36,617
10	13,047	11,833	35,848	31,784	37,778	33,472	25,336	22,586	26,192	23,751	17,519	16,162
11	16,487	14,843	38,451	34,061	29,376	26,120	23,743	22,025	18,230	16,784	38,841	34,819
12	14,345	12,968	18,573	16,668	32,822	29,553	29,751	27,282	37,085	33,283	37,324	33,492
E	15,884	14,245	29,388	26,115	30,029	26,762	28,772	25,870	27,064	24,340	29,497	26,502
S	190,606	170,945	352,654	313,385	360,349	321,138	345,259	310,435	324,762	292,083	353,966	318,029

Рис. 1: Входные данные задачи 1

2.2 Задача 2

Оценка общего потока пассажира на основе данных из рисунков 2-5.

Необходимо:

1. Оценить долю турникетов с памятью среди всех турникетов в метрополитене
2. Определить связь между количеством человек, вошедших в метро в день контроля, и количеством человек, оплативших проезд.

IIA0		IIA1		IIB0		IIB1		IICO		IIC1		
4,252	2,215	4,902	2,540	6,965	3,511	4,070	2,063	4,344	2,200	1,164	0,671	F
	1,197		1,359		1,635		0,912		0,980		0,425	D
	0,841		1,003		1,819		1,095		1,164		0,069	L
5,158	2,668	6,006	3,092	7,195	3,626	3,840	1,948	5,574	2,815	1,394	0,725	F
	1,423		1,635		1,693		0,854		1,288		0,243	D
	1,067		1,279		1,877		1,038		1,471		0,426	L
4,907	2,543	5,252	2,715	5,088	2,572	8,825	4,441	8,348	4,202	5,489	2,773	F
	1,360		1,447		1,166		2,100		1,981		1,266	D
	1,004		1,091		1,350		2,284		2,165		1,450	L
4,349	2,264	3,275	1,727	5,318	2,687	8,595	4,326	9,578	4,817	5,719	2,888	F
	1,221		0,952		1,224		2,043		2,289		1,324	D
	0,865		0,596		1,407		2,227		2,472		1,508	L
5,242	2,710	4,278	2,228	7,989	4,023	7,599	3,828	11,529	5,793	12,594	6,325	F
	1,444		1,203		1,891		1,794		2,776		3,043	D
	1,088		0,847		2,075		1,978		2,960		3,226	L
4,995	2,587	3,570	1,874	8,219	4,138	7,369	3,713	12,759	6,408	12,827	6,442	F
	1,382		1,026		1,949		1,736		3,084		3,101	D
	1,026		0,670		2,133		1,920		3,268		3,285	L
3,550	1,864	4,369	2,274	10,429	5,243	3,891	1,973	6,546	3,301	12,091	6,074	F
	1,021		1,226		2,501		0,867		1,531		2,917	D
	0,665		0,870		2,685		1,051		1,714		3,101	L
4,458	2,318	5,365	2,772	10,659	5,358	3,661	1,858	7,776	3,916	12,321	6,189	F
	1,248		1,475		2,559		0,809		1,838		2,974	D
	0,892		1,119		2,743		0,993		2,022		3,158	L
4,209	2,194	4,663	2,421	2,987	1,522	9,256	4,656	8,345	4,201	5,222	2,639	F
	1,186		1,299		0,641		2,208		1,980		1,200	D
	0,830		0,943		0,825		2,392		2,164		1,383	L
5,704	2,941	5,350	2,764	3,217	1,637	9,026	4,541	9,575	4,816	5,452	2,754	F
	1,560		1,471		0,698		2,151		2,288		1,257	D

Рис. 2: Входные данные задачи 2(1)

	1,204		1,115		0,882		2,334		2,472		1,441	L
6,605	3,392	6,659	3,419	4,232	2,144	8,934	4,495	5,985	3,021	13,016	6,536	F
	1,785		1,798		0,952		2,128		1,390		3,148	D
	1,429		1,442		1,136		2,311		1,574		3,332	L
6,357	3,268	5,804	2,991	4,462	2,259	8,704	4,380	7,211	3,634	13,242	6,649	F
	1,723		1,585		1,010		2,070		1,697		3,205	D
	1,367		1,229		1,193		2,254		1,881		3,388	L
3,348	1,696	4,419	2,232	7,509	3,783	3,689	1,873	8,672	4,364	10,211	5,134	F
	0,716		0,984		1,771		0,816		2,062		2,447	D
	0,936		1,204		1,955		1,000		2,246		2,631	L
4,239	2,142	5,418	2,731	7,739	3,898	3,459	1,758	9,907	4,982	10,441	5,249	F
	0,939		1,234		1,829		0,759		2,371		2,504	D
	1,159		1,454		2,013		0,943		2,555		2,688	L
3,984	2,014	4,713	2,379	4,581	2,319	4,563	2,310	4,908	2,482	10,526	5,291	F
	0,875		1,057		1,039		1,035		1,121		2,526	D
	1,095		1,277		1,223		1,219		1,305		2,709	L
3,654	1,849	6,119	3,082	4,811	2,434	4,333	2,195	6,138	3,097	5,178	2,617	F
	0,793		1,409		1,097		0,977		1,429		1,189	D
	1,013		1,629		1,281		1,161		1,612		1,372	L
4,556	2,300	7,117	3,581	4,150	2,103	5,156	2,606	10,252	5,154	5,404	2,730	F
	1,018		1,658		0,932		1,183		2,457		1,245	D
	1,238		1,878		1,115		1,367		2,641		1,429	L
4,309	2,177	6,416	3,230	4,380	2,218	4,927	2,492	11,482	5,769	5,489	2,773	F
	0,956		1,483		0,989		1,126		2,765		1,266	D
	1,176		1,703		1,173		1,310		2,948		1,450	L
6,973	3,509	5,695	2,870	8,204	4,130	5,154	2,605	7,382	3,719	9,873	4,965	F
	1,622		1,303		1,945		1,183		1,740		2,362	D
	1,842		1,523		2,129		1,366		1,923		2,546	L
7,872	3,958	6,692	3,368	8,434	4,245	4,923	2,490	8,610	4,333	10,103	5,080	F
	1,847		1,552		2,003		1,125		2,047		2,420	D

Рис. 3: Входные данные задачи 2(2)

	2,067		1,772		2,186		1,309		2,230		2,604	L
7,626	3,785	5,991	2,968	6,104	3,080	4,755	2,406	4,726	2,391	10,188	5,122	F
	2,023		1,614		1,420		1,083		1,076		2,441	D
	1,819		1,410		1,604		1,267		1,259		2,625	L
2,113	1,029	5,434	2,689	6,334	3,195	4,525	2,291	5,956	3,006	4,029	2,043	F
	0,644		1,475		1,478		1,025		1,383		0,901	D
	0,440		1,271		1,661		1,209		1,567		1,085	L
3,018	1,481	6,475	3,210	6,965	3,511	6,257	3,157	11,011	5,534	4,252	2,043	F
	0,871		1,735		1,635		1,458		2,647		0,901	D
	0,667		1,531		1,819		1,642		2,831		1,085	L
2,762	1,353	5,778	2,861	7,195	3,626	6,021	3,039	12,241	6,149	4,344	2,110	F
	0,807		1,561		1,693		1,399		2,954		1,112	D
	0,603		1,357		1,877		1,583		3,138		1,030	L
3,636	1,790	1,209	0,577							9,360	2,156	F
	1,025		0,418								1,135	D
	0,821		0,214								1,053	L
4,532	2,238	2,212	1,078							9,592	4,664	F
	1,249		0,669								2,389	D
	1,045		0,465								2,307	L
4,286	2,115	1,504	0,724							9,675	4,780	F
	1,188		0,492								2,447	D
	0,984		0,288								2,365	L
3,089	1,517	5,929	2,937							8,981	4,822	F
	0,888		1,598								2,468	D
	0,684		1,394								2,386	L
3,989	1,967	6,929	3,437							9,211	4,475	F
	1,113		1,848								2,294	D
	0,909		1,644								2,212	L
3,739	1,842	6,229	3,087							9,296	4,590	F
	1,051		1,673								2,352	D

Рис. 4: Входные данные задачи 2(3)

	0,847		1,469								2,270	L
4,235	2,090	6,450	3,197									
	1,175		1,729									
	0,971		1,525									
5,135	2,540	7,454	3,699									
	1,400		1,980									
	1,196		1,776									
4,885	2,415	6,750	3,347									
	1,337		1,804									
	1,133		1,600									
3,521	1,733	2,474	1,209									
	0,996		0,735									
	0,792		0,531									
4,421	2,183	3,474	1,709									
	1,221		0,985									
	1,017		0,781									
4,171	2,058	2,772	1,358									
	1,159		0,809									
	0,955		0,605									

Рис. 5: Входные данные задачи 2(4)

2.3 Задача 3

Найти среднюю стоимость билета пражского метро. Входные данные взять из таблиц, изображенных на рисунках 2-5.

3 Теоретические сведения

Методика построения стохастической модели:

1. Выполняется описание предметной области на естественном языке в общей форме с входными и выходными данными и параметрами системы.
2. Рисуются схема сложной системы с входными и выходными данными.
3. Формирование статистической гипотезы относительно распределений потоков входных и выходных данных. Проверка выдвинутых гипотез.
4. Оценивается возможность реализации модели. При избыточности данных модель упрощается. При недостатке данных изыскивается возможность их получения.
5. Формируется общее уравнение, связывающее входные и выходные данные.
6. По результатам эксперимента оцениваются параметры уравнения, полученного на этапе 5.
7. Оценивается применимость полученной модели на основе контрольных выборок натурального эксперимента для возможности повышения адекватности построенной модели.
8. Построение концептуальной модели на основе формального описания для проведения натурального эксперимента.

Критерий Колмогорова-Смирнова:

Критерий используется для проверки гипотезы H_0 : "случайная величина X имеет распределение $F(x)$ ".

Пусть X_n - выборка независимых одинаково распределённых случайных величин, $F_n(x)$ - эмпирическая функция распределения, $F(x)$ - некоторая "истинная" функция распределения с известными параметрами. Статистика критерия определяется выражением:

$$D_n = \sup_x |F_n(x) - F(x)|. \quad (1)$$

Схема пражского метро:

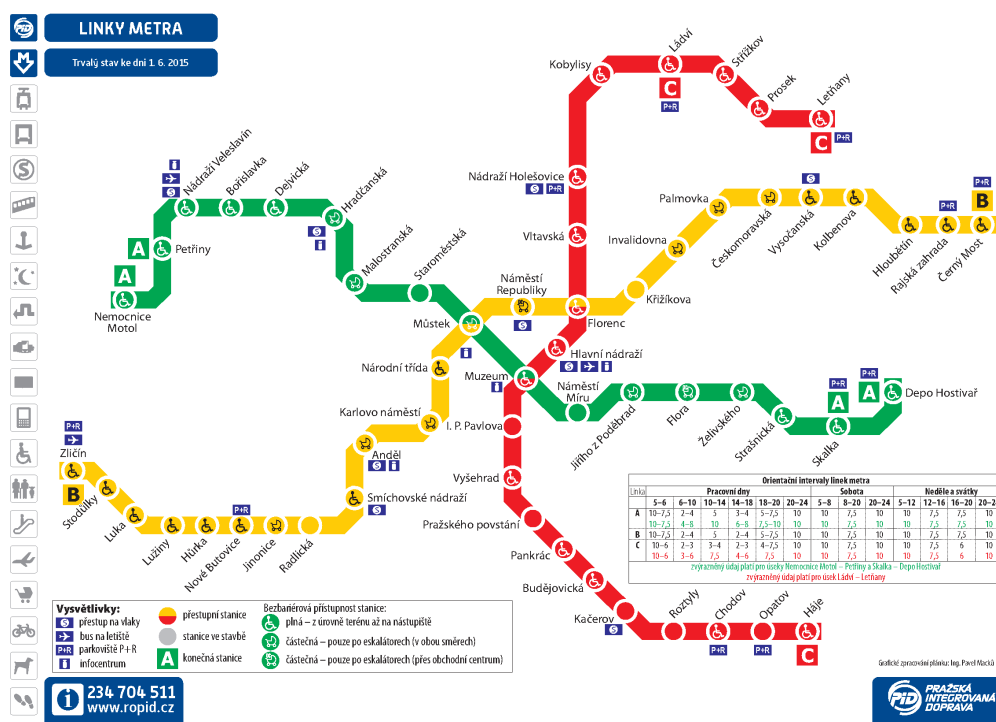


Рис. 6: Схема метро.

3.1 Задача 1

Пусть

- ε_1 - количество людей, вошедших на станцию в течение дня;

- ε_2 - количество людей, прошедших через турникеты на станции в течение дня.

Гипотеза 1: Между величинами ε_1 и ε_2 существует стохастическая связь, которая может быть выражена как:

$$\varepsilon_1 = \varphi(\varepsilon_2) = \alpha \varepsilon_2^\beta. \quad (2)$$

Функция распределения будет иметь вид:

$$F(\varepsilon_1, \alpha_1, \beta_1) = 1 - e^{-\lambda_1 \varepsilon_1^{\beta_1}}, \quad (3)$$

$$F(\varepsilon_2, \alpha_2, \beta_2) = 1 - e^{-\lambda_2 \varepsilon_2^{\beta_2}}, \quad (4)$$

Из метода моментов, логарифмируя левую и правую часть равенства, и используя выборочное среднее и среднеквадратичное отклонение для выборок значений случайных величин ε_1 и ε_2 , получаем решение для параметров статистик:

$$\hat{\beta} = \sqrt{\frac{\hat{S}^2(\ln(\varepsilon_1))}{\hat{S}^2(\ln(\varepsilon_2))}} \quad (5)$$

$$\hat{\alpha} = \exp(\overline{\ln(\varepsilon_1)} - \hat{\beta} \overline{\ln(\varepsilon_2)}), \quad (6)$$

где

- $\hat{\alpha}$ и $\hat{\beta}$ - выборочные значения α и β ,
- $\ln(\varepsilon_i)$ - средневывборочное значение случайной величины,
- \hat{S}^2 - среднеквадратичное отклонение выборки.

Далее с помощью критерия Колмогорова-Смирнова можно проверить гипотезу о принадлежности выборки закону распределения

3.2 Задача 2

Пусть

- ε_1 - количество людей, вошедших в метро в день контроля;
- ε_3 - количество людей, оплативших проезд.

Гипотеза 2: Существует стохастическая зависимость между случайными величинами ε_1 и ε_3 , которую в пределе можно описать функцией вида:

$$\varepsilon_3 = \varphi_2(\varepsilon_1) = \alpha_2 \varepsilon_1^{\beta_2} \quad (7)$$

3.3 Задача 3

Находится количество билетов для каждой категории из трех данных:

- F - взрослый билет ($count_tickets_F$);
- D - детский билет ($count_tickets_D$);
- L - льготный билет ($count_tickets_L$).

Далее находится среднее арифметическое стоимости билета и общая стоимость всех билетов с помощью формулы:

$$cost_all_tickets = \sum_{i=1}^3 (avg_i * count_tickets_i), \quad (8)$$

где $i = 1$ для F , $i = 2$ для D , $i = 3$ для L .

Тогда средняя стоимость всех билетов будет вычисляться по формуле:

$$average = \frac{cost_all_tickets}{3 \times \sum_{i=1}^3 (count_tickets_i)}. \quad (9)$$

4 Практическая реализация

Листинг 1. Задача 1

```
1  # work
2  psi_1w = thA0 + thB0 + thC0
3  psi_2w = rA0 + rB0 + rC0
4  alpha, betta = alpha_betta(psi_1w, psi_2w)
5
6  print("psi_1: a =", alpha, "b =", betta)
7
8  # control
9  psi_1c = thA1 + thB1 + thC1
10 psi_2c = rA1 + rB1 + rC1
11
12 # _ = *
13 psi_2_ = psi_2c
14 psi_1_ = [alpha * (psi_2_[i] ** betta) for i in range(len(psi_2_))]
15
16 psi_1c.sort()
17 psi_1_.sort()
18
19 print(get_sections(psi_1c))
20 print(get_sections(psi_1_))
21
22 psi_1c_len = [len(data) / 36 for data in get_sections(psi_1c)]
23 psi_1__len = [len(data) / 36 for data in get_sections(psi_1_)]
24
25 result = [abs(psi_1c_len[i] - psi_1__len[i]) for i in range(len(psi_1c_len))]
26
27 if max(result) < 0.05:
28     print("Ok")
29     print(max(result))
30 print()
```

Листинг 2. Задача 2

```
1  I0 = IIA0 + IIB0 + IIC0
2
3  # work
4  ksi_w1 = []
5  for value_list in I0:
6      ksi_w1.append(sum(map(int, value_list)))
```

```

7
8 ksi_w2 = []
9 for value_list in I0:
10     if len(value_list) == 3:
11         ksi_w2.append(value_list[0] + value_list[1])
12     else:
13         ksi_w2.append(value_list[0])
14
15 ksi_w3 = []
16 for value_list in I0:
17     ksi_w3.append(value_list[len(value_list) - 1])
18
19 # control
20 ksi_c1 = []
21 for value_list in I0:
22     ksi_c1.append(sum(map(int, value_list)))
23
24 ksi_c2 = []
25 for value_list in I0:
26     if len(value_list) == 3:
27         ksi_c2.append(value_list[0] + value_list[1])
28     else:
29         ksi_c2.append(value_list[0])
30
31 I1 = IIA1 + IIB1 + IIC1
32
33 ksi_c3 = []
34 for value_list in IIA1:
35     ksi_c3.append(value_list[len(value_list) - 1])
36
37 a_12, b_12 = alpha_betta(ksi_w1, ksi_w2)
38 a_23, b_23 = alpha_betta(ksi_w2, ksi_w3)
39
40 a, b = alpha_betta(ksi_w1, ksi_w3)
41 print("a_12 = {}, b_12 = {}".format(a_12, b_12))
42 print("a_23 = {}, b_23 = {}".format(a_23, b_23))
43
44 ksi_calc3 = [(x ** (1 / b_23)) / a_23 for x in ksi_c2]
45 ksi_temp = [a_23 * (x ** b_23) for x in ksi_calc3]
46 ksi_calc1 = [a_12 * (x ** b_12) for x in ksi_temp]
47
48 ksi_c1.sort()
49 ksi_calc1.sort()
50 dif = []
51 n = len(ksi_w1)
52
53 for i in range(0, 30, 10):
54     func_control = len([x for x in ksi_c1 if x < i]) / n
55     func_calc = len([x for x in ksi_calc1 if x < i]) / n

```

```

56         dif.append(abs(func_control - func_calc))
57     print("dif ", max(dif))
58     print()

```

Листинг 3. Задача 3

```

1     # work
2     FDL0 = np.asarray(FDLA0 + FDLB0 + FDLC0)
3     # FDL1 = FDLA1 + FDLB1 + FDLC1
4
5     # F - взрослый
6     # D - детский
7     # L - льготный
8
9     count_tickets_F = sum(FDL0[:, 0])
10    count_tickets_D = sum(FDL0[:, 1])
11    count_tickets_L = sum(FDL0[:, 2])
12
13    avg_F = average(F_30, F_90, F_24, F_72)
14    avg_D = avg_L = average(D_30, D_90, D_24, D_72)
15
16    cost_all_tickets = avg_F * count_tickets_F + avg_D * count_tickets_D + avg_L
17    ↪ * count_tickets_L
18    average_cost = cost_all_tickets / (count_tickets_F + count_tickets_D +
19    ↪ count_tickets_L)
20    print("Average - {}".format(average_cost / 3))

```


5 Результаты

Ниже приведен результат работы программы из **Листинга 1, 2, 3.**

Листинг 4. Результат работы программы для 3 задач.

```
1      --1--
2  psi_1: a = 1.0644531215676833 b = 1.015231400769257
3  Hypothesis 1: Ok
4  0.02777777777777779
5
6      --2--
7  a_12 = 1.7310412290615682, b_12 = 0.9976777068524383
8  a_23 = 1.3961675695768274, b_23 = 0.8984112750936865
9  Hypothesis 2: Ne Ok
10 0.055555555555555547
11
12      --3--
13 Average - 19.965618399165084
```

Задача 1. Гипотеза о наличии связи между между потоком пассажиров, вошедших в метро, и пассажиров, прошедших через турникет подтвердилась, так как критерий Колмогорова-Смирнова дал результат $D = 0.02778$, что соответствует $D < 0.05$ (если $D > 0.05$ - то гипотеза бы не подтвердилась).

Задача 2. Гипотеза о наличии связи между между кол-вом людей, вошедших в метро в день контроля, и кол-вом людей, оплативших проезд не подтвердилась, так как критерий Колмогорова-Смирнова дал результат $D = 0.0556$ ($D > 0.05$).

Задача 3. Была получена средняя стоимость билета пражского метро - 19.97 крон.

6 Вывод

В ходе лабораторной работы была изучена стохастическая модель на примере пражского метро, применен критерий Колмогорова-Смирнова, получена средняя стоимость билета метро на основе входных данных, полученных с помощью натурного эксперимента.