

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
“КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ імені ІГОРЯ СІКОРСЬКОГО”
ФАКУЛЬТЕТ ІНФОРМАТИКИ ТА ОБЧИСЛЮВАЛЬНОЇ ТЕХНІКИ



Кафедра інформатики та програмної інженерії

Звіт до РГР

з курсу

**«Ймовірнісні моделі та статистичне оцінювання
в інформаційно-управляючих системах»**

*студента 2 курсу
групи ІТ-02*
Макарова Іллі Сергійовича

Викладач:
к.ф.-м.н., доц. Гарко І.І.

Київ – 2022

Максров Ліне РТР

Завдання 1

Згенеруємо вибірку з нор. розп. собак, з $\mu = 300$, $\sigma = 100$

(22, 88, 107, 126, 138, 138, 145, 147, 163, 171, 177, 179, 182, 195, 195, 197, 198, 201, 203, 204, 204, 207, 209, 223, 225, 237, 240, 252, 259, 259, 260, 262, 262, 265, 265, 266, 266, 271, 275, 275, 279, 291, 293, 293, 311, 312, 320, 321, 323, 328, 331, 335, 339, 339, 345, 348, 348, 348, 348, 350, 353, 354, 355, 359, 362, 364, 368, 368, 372, 374, 383, 384, 384, 394, 395, 395, 395, 396, 396, 398, 401, 405, 406, 410, 414, 423, 424, 424, 428, 429, 430, 435, 450, 450, 469, 475, 478, 483, 530, 624)

~~Вибіркову вибірку з нор. розп. собак, з $\mu = 300$, $\sigma = 100$~~
Вибіркову вибірку з нор. розп. собак, з $\mu = 300$, $\sigma = 100$

Вибіркову інтервальний розподіл: $K = 1 + [3.322 \cdot \lg 100] = 7$

$$h = \frac{X_{\max} - X_{\min}}{K} = \frac{624 - 22}{7} = 86$$

$[X_i; X_{i+1})$ [22; 108) [108; 194) [194; 280) [280; 366) [366; 452)

n_i : 3 10 28 25 28

[452; 538) [538; 624]

5

1

В останньому інтервалі лише одне спостереження, і його можна було б діяти з передбачен. Але робити цього не буду, оскільки

Также, возьмем сред. интервалы,

x_i	65	151	237	323	409	495	581
n_i	3	10	28	25	28	5	1

Теперь найдем + выборочное среднее та ~~гр~~ выб. дисперсию

$$\bar{x} = \frac{65 \cdot 3 + 151 \cdot 10 + 237 \cdot 28 + 323 \cdot 25 + 409 \cdot 28 + 495 \cdot 5 + 581 \cdot 1}{100} \approx 309$$

$$\overline{D}(x) = M(x^2) - M(x)^2 \approx 106681 - 309^2 = 11200$$

$\sigma(x) = \sqrt{D} \approx 105,8$, тут гипотеза не выпр. Д. До выборка достат. велика.

Сколько подл. $n'_i = \frac{h \cdot n}{\sigma} \varphi(z_i)$, где $z_i = \frac{x_i - \bar{x}}{\sigma}$

x_i	n_i	z_i	$\varphi(z_i)$	n'_i
65	3	-2,3	0,03	2,4
151	10	-1,5	0,13	10,5
237	28	-0,68	0,32	26
323	25	0,13	0,4	32,5
409	28	0,95	0,25	20,3
495	5	1,75	0,09	7,3
581	1	2,6	0,014	1,1

Нехай рівень значущості рівний 0.05 , кількість степенів свободи
це $7-2-1=4$, тоді $\chi^2_{кр} = 15.1$

$$\chi^2_{таб} = \sum \frac{(n_i - n'_i)^2}{n'_i} = 0.15 + 0.02 + 0.15 + 1.73 + 2.9 + 0.72 + 0.01 = 5.68$$

Таким чином, оскільки $\chi^2_{таб} < \chi^2_{кр}$, то ми приймаємо
гіпотезу H_0 , що ген. сукуп. розп. норм.

Частина II

Також, тепер нам треба побудувати довірчий інтервал для μ при
заданому δ , для μ при не зад. σ та для дисперсії.

Нехай

$$\delta = 0.99$$

σ генеральної сукуп. як і вже казав $= 100$

Тоді довірчий інтервал для μ :

$$\bar{X}_n - \frac{\sigma \sqrt{t}}{\sqrt{n}} < \mu < \bar{X}_n + \frac{\sigma \sqrt{t}}{\sqrt{n}}, \text{ де } t: \Phi\left(\frac{t}{2}\right) = \frac{\delta}{2}$$

$$\text{Отже } \mu \in \left(309 - \frac{100 \cdot 2.6}{10}; 309 + \frac{100 \cdot 2.6}{10} \right)$$

$$\boxed{\mu \in (283; 335)}$$

Теперь убедились, что \bar{x} нам не подходит. Поскольку $n > 30$,
то для поиска интервала теперь нам треба skorzystać
формулы

$$\Delta \in \left(\bar{x}_B - \frac{st}{\sqrt{n}} ; \bar{x}_B + \frac{st}{\sqrt{n}} \right)$$

Треба найти s , учитывая, что выборка D нам нужна
(в минимуме значений \bar{x}).

$$s^2 = \frac{n}{n-1} \cdot D_B = \frac{100 \cdot 11200}{99} = ~~11313~~$$

$$s = \sqrt{s^2} = 106,4$$

$$\text{Таким образом } \Delta \in \left(309 - \frac{106,4 \cdot 2,6}{10} ; 309 + \frac{106,4 \cdot 2,6}{10} \right)$$

$$\Delta \in (281,3 ; 336,7)$$

Ну, минимум мне найти \bar{x} , тут все просто: $n=100$, $\bar{x}=0,99 \Rightarrow$

$$\Rightarrow q = 0,198, \text{ где } \bar{x} \in (0 ; s(1+q))$$

Ой, же не треба было
найти D , а не \bar{x} ,

а в книге работи записав же оценку D

$$\bar{x} \in (0 ; 106,4 \cdot 1,198)$$

$$\bar{x} \in (0 ; 127,5)$$

000
000

Задача III Ви скадали, що тут треба за доп. метода найменш. кв. знайти коеф. ~~у~~ для лінійної регресії. Що взагалі дивно, ~~якщо~~ бо норм. розпод. виражається за доп. ~~формули~~ Гауса, що точно не є лінійною, але ладно.

Для погляду тут ~~я~~ треба зробити ~~дискретний~~ стат. розподіл:

X: [22, 88, 107, 126, 138, 145, 147, 163, 171, 177, 179, 182, 195, 197, 198, 201, 203, 204, 207, 209, 223, 225, 237, 240, 252, 259, 260, 262, 265, 266, 271, 275, 279, 291, 293, 311, 312, 320, 321, 323, 328, 331, 335, 339, 345, 348, 350, 353, 354, 355, 359, 362, 364, 368, 372, 374, 383, 384, 394, 395, 396, 398, 401, 405, 406, 410, 414, 423, 424, 428, 429, 430, 435, 450, 469, 475, 478, 483, 530, 624]

N: [1, 1, 1, 1, 2, 1, 1, 1, 1, 1, 1, 1, 2, 1, 1, 1, 1, 2, 1, 1, 1, 1, 1, 1, 2, 1, 2, 2, 2, 1, 2, 1, 1, 2, 1, 1, 1, 1, 1, 1, 1, 2, 1, 4, 1, 1, 1, 1, 1, 1, 2, 1, 1, 1, 2, 1, 3, 2, 1, 1, 1, 1, 1, 1, 1, 2, 1, 1, 1, 2, 1, 1, 1, 1, 1]

Тепер треба знайти коеф. k та b . Зробимо це за доп. методу найменш. кв., розв'яжемо систему

$$\begin{cases} k \sum x_i^2 + b \sum x_i = \sum x_i y_i \\ k \sum x_i + nb = \sum y_i \end{cases} \Leftrightarrow \left\{ \begin{array}{l} \text{На ~~линійній~~ калькуляторі знайдемо} \\ \sum x_i^2 = 8717483 \\ \sum x_i = 24845 \\ \sum y_i = 100 \\ n = 80 \\ \sum x_i y_i = 31197 \end{array} \right. \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} 8717483 \cdot k + 24845 \cdot b = 31197 \\ k \cdot 24845 + 80 \cdot b = 100 \end{cases} \Leftrightarrow \begin{cases} b = \frac{100 - 24845k}{80} \\ 8717483 \cdot k + 24845 \left(\frac{100 - 24845k}{80} \right) = 31197 \end{cases}$$

$$\Leftrightarrow \begin{cases} b = \frac{100 - 24845k}{80} \\ 8717483k - 776857k = 31197 - 31056 \end{cases} \Leftrightarrow \begin{cases} b = \frac{100 - 24845k}{80} \\ 1000626k = 141 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} k = 14 \cdot 10^{-5} \\ b = 1,208 \end{cases} \text{ Спрівимо з тим, що було розраховано:}$$

✓ [20] slope, intercept, r, p, std_err = stats.linregress(x, y)

✓ slope, intercept

(0.000140531096467671, 1.206356311353259)

Ну а ще там трошки округлював, тому b в тисячних відр.
Таким чином рівн. регр. $y = 14 \cdot 10^{-5} x + 1,208$, де x -зв., y -кільк.

~~Так, думав~~ Так, думав очевидно, що таке рівн.
Дуже і дуже погано описує нормальний розподіл. Але
когда віддалено зогось розбедного треба використовувати
багатолирну регресію. Але ви сказали розв'язати лінійну. :D

Так, а тепер я ще напишу оцінку для D , що раніше випадково отримав

Оберу надійність $\gamma = 0,95$

Формула для оцінки D : $\frac{ns^2}{\chi_{1n}^2} \leq D \leq \frac{ns^2}{\chi_{2n}^2}$

$s^2 = 11313$, це я вже рахував

$$\chi_{1n} = \chi^2\left(\frac{1+\gamma}{2}; n\right) = \chi^2\left(0, \overset{0,95}{\cancel{0,975}}; 100\right) = 129,561$$

$$\chi_{2n} = \chi^2\left(\frac{1+\gamma}{2}; n\right) = \chi^2\left(0, \overset{0,05}{\cancel{0,975}}; 100\right) = 74,222$$

Таким чином

$$\frac{100 \cdot 11313}{129,561} \leq D \leq \frac{100 \cdot 11313}{74,222}$$

$$8732 \leq D \leq 15242$$

