

ЛАБОРАТОРНА РОБОТА № 4. НЕЗАЛЕЖНІ ЛІНГВІСТИЧНІ

ВИПРОБУВАННЯ В ТЕКСТІ

ЗАВДАННЯ 1.

- 1 Середня довжина простого речення або синтаксично оформленої частини складного речення в англійських науково-технічних текстах є в межах від 10 до 11 словоформ. Априорна ймовірність появи іменників у вказаній підмові дорівнює $1/3$. Вважаючи появу окремих словоформ у типових синтаксично оформлених сегментах цих текстів незалежними подіями, визначити ймовірність того, що на 10 слововживань, що складають типовий сегмент тексту, рівно x будуть іменниками. Розв'язати задачу для $x=1, 3, 4, 5$.
- 2 Визначити ймовірність того, що в 10-слівному сегменті англійського науково-технічного тексту з'явиться рівно 4 іменники, 2 дієслівні форми і 4 словоформи, що належать іншим класам. Задана нормою априорна ймовірність появи іменників дорівнює 0.33, ймовірність дієслівних форм складає 0.16, ймовірність решти граматичних класів дорівнює 0.51.

ЗАВДАННЯ 2.

- 3 Вважаючи, що ймовірність використання іменників у розмовній мові дорівнює 0.11, в поезії – 0.19, в художній прозі – 0.32, в науково-технічних текстах – 0.39, знайти ймовірності появи серед одночасно взятих чотирьох словосполучень жодної, однієї, двох, трьох, чотирьох іменних груп.
- 4 Нехай із текстів з радіоелектроніки випадковим чином взято 1000 слововживань. Знайти ймовірність того, що словоформа струм зустрінеться хоча б один раз, якщо її відносна частота дорівнює 0.0012.
- 5 Визначити об'єм вибірки текстів з радіоелектроніки (див. зад. 2.6), який необхідний для того, щоб з імовірністю 90% словоформа струм з'явилась у ній хоча б один раз. Розв'язати цю ж задачу для ймовірності 95%.

ЗАВДАННЯ 3.

1 Визначимо величини c, w, g, t, s :

Порядковий номер цифри в номері студентського квитка студента				
1	2	3	4	5
c	w	g	t	s
Сортувати за зростанням ($c < w$)				

▪ Знайти імовірність появи слова *портал* хоча б один раз в сотні випробувань, якщо імовірність появи цього слова складає 0,03. Знайти кількість випробувань, які необхідно здійснити, щоб з точністю $s\%$ можна було очікувати появу слова портал принаймні один раз.

▪ Для визначення стилістико-семантичних особливостей деякого тексту з нього було навмання вибрано 100 відрізків по 10 слів кожен. Частоти появи іменників у цих серіях подано в таблиці.

x	0	1	2	3	4	5	6	7	8	9	10
S_x	0	1	4	15	33	27	11	4	2	1	2

▪ Намалювати графік залежності $P_N(x)$ від x . Пересвідчитись, чи відповідають графічні результати чисельному значенню x_0 .

▪ Знайти імовірність того, що в одній серії іменник з'явиться від c до w разів; не більше g разів; не менше t разів.