



МЕХАНІКО-  
МАТЕМАТИЧНИЙ  
ФАКУЛЬТЕТ  
Львівський університет



Олег Бугрій

# Баєсівський аналіз даних

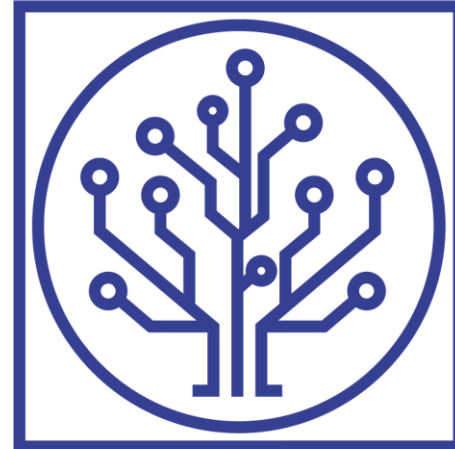
Львівський національний університет  
імені Івана Франка



DES - 2020  
Data Engineering  
and Security

# Статистичний аналіз даних

STATISTICS &  
DATA SCIENCE

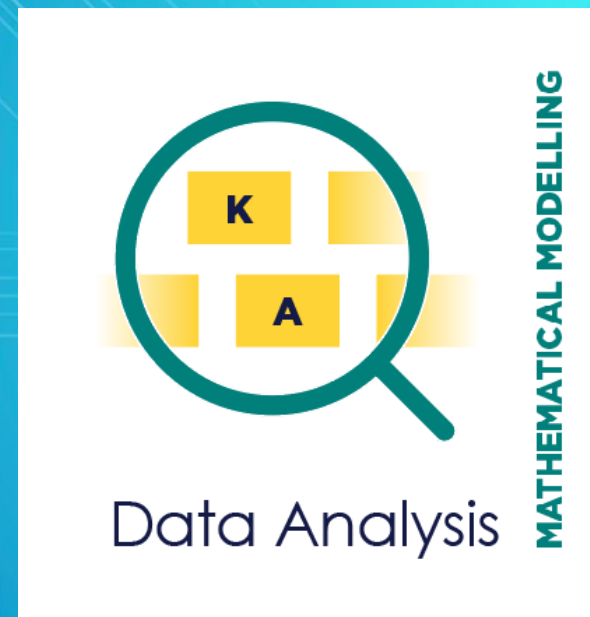


МЕХАНІКО-  
МАТЕМАТИЧНИЙ  
ФАКУЛЬТЕТ  
ЛЬВІВСЬКИЙ УНІВЕРСИТЕТ

Галузь знань 11 – “Математика та статистика”  
Спеціальність 112 – “Статистика”

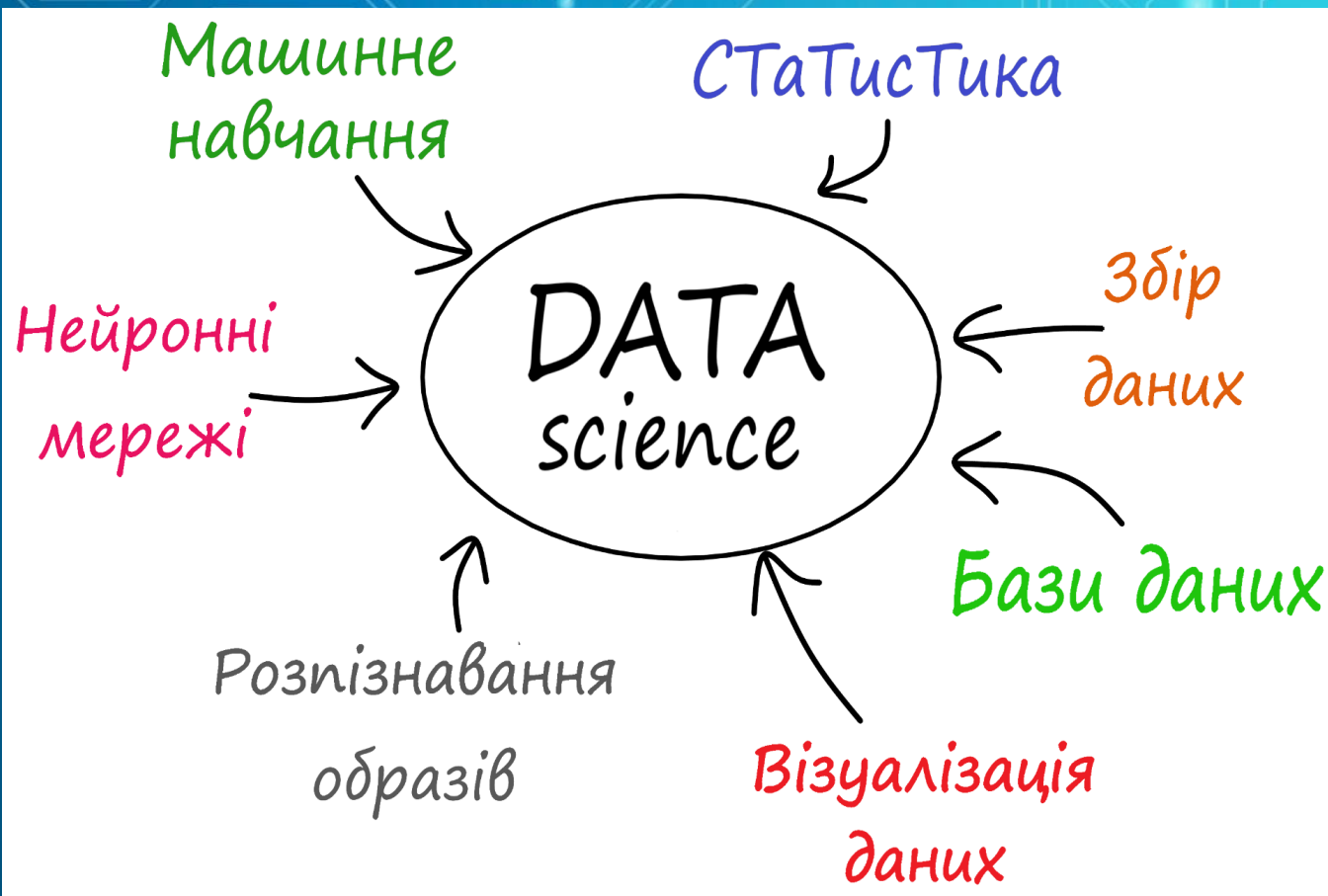


# Комп'ютерний аналіз математичних моделей





# Data Science - це сучасна наука про методи аналізу даних і видобування з них цінної інформації







**Статистика:** Аналізом даних раніше займалася наука “Статистика”, яка зародилася на початку XIX ст. в працях Ф. Гальтона і К. Пірсона. Вона зосереджувалася на статистичному висновку, тобто досить складному наборі процедур для отримання інформації про генеральні сукупності на основі невеликих вибірок. В 1962р. Дж. Тьюкі запропонував перетворити статистику в нову наукову дисципліну – “Аналіз даних”, яку сьогодні прийнято називати “Data Analysis” або “Data Science”.



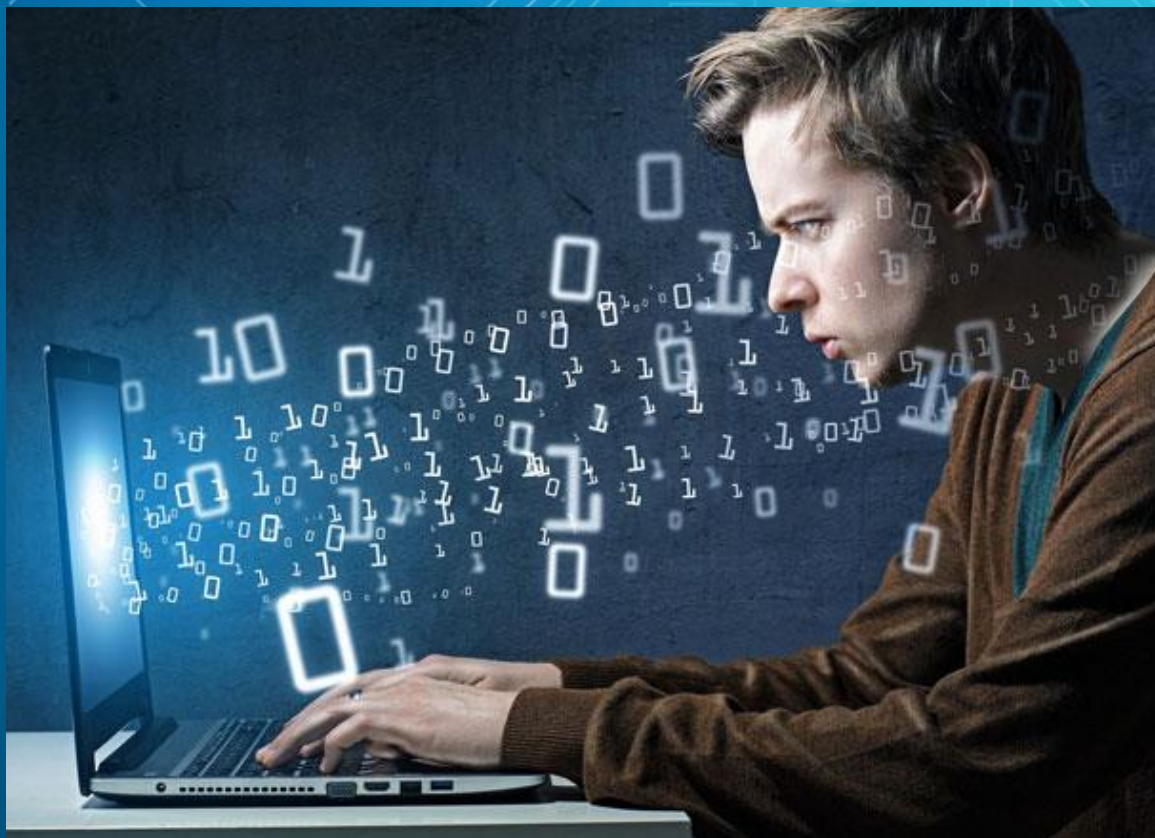


# Програмування:

За час масового поширення технологій людство згенерувало величезну кількість даних. Таку, що ми самі не здатні їх опрацювати і візуалізувати.

Це дані про наші дзвінки і переміщення, поведінку в інтернеті, смаки в магазинах, політичні вподобання, інформація про спортивні події, кліматичні процеси тощо.

З цих даних при правильному підході можна дістати велику користь. І в цьому нам допомагають комп'ютери.







**В основі і статистики і  
програмування лежить**

**МАТЕМАТИКА**

**- це основа Data Science !**



Народився в 1702 році в Лондоні. Навчався вдома, у 1719 році вступив до Единбурзького університету. Потім Баєс допомагав батькові проводити службу, а незабаром, у 30-их роках, сам став священником в пресвітеріанській церкві. У 1752 році він вийшов у відставку; помер в 1761 році.

Математичні інтереси Баєса відносяться до теорії ймовірностей. Він сформулював і вирішив одну з основних задач цього розділу математики (теорема Баєса). Робота, присвячена цій формулі, була опублікована в 1763 році, вже після його смерті. Формула Баєса, що дає можливість оцінити ймовірність подій емпіричними шляхом, відіграє важливу роль у сучасній математичній статистиці та теорії ймовірностей.





1. Зісканувати код
2. Перейти за посиланням
3. Ввести email
4. В листі, що прийде натиснути «Confirm email»
5. Ввести ім'я та прізвище
6. Натиснути «Next»
7. Придумати і ввести пароль
8. Натиснути на «Channels»
9. Вибрати канал з лекцією
10. Натиснути «Join channel»





# Money or die?





# Теорема Баєса (Bayes): Означення & Інтуїтивне пояснення

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$



Тут А та В – події;

$P(A)$  і  $P(B)$  – ймовірність подій А та В відповідно (без врахування один одного);

$P(A|B)$  – умовна ймовірність, тобто ймовірність події А за умови виконання події В (probability of A given that B is true);

$P(B|A)$  – умовна ймовірність, тобто ймовірність події В за умови виконання події А.

Вживатимемо таку термінологію:

$$P(A|B) = \frac{\overset{\text{Weighted Average}}{\boxed{P(A)}} * \overset{\text{Assumption}}{\boxed{P(B|A)}}}{\underset{\text{Normalized}}{\boxed{P(B)}}}$$

Тобто,

$P(A)$  – зважене середнє;

$P(B|A)$  – припущення;

$P(B)$  – нормалізуючий множник.



## Приклад 1 - Простий приклад з дайсом (dice)

Припустимо, що наш друг має 3 гральні “кубики” (або гральні кісточки, гральні кості, дайси). Перший гральний “кубик” (далі казатимемо **дайс**) має 4 сторони, другий – 6 сторін, а третій – 8 сторін. Друг вибирає один дайс випадково, котить його без показу нам, і повідомляє результат – випало число 2. Нам треба вгадати який кубик він вибрав і котив. Для цього нам треба обчислити ймовірність того, що дайс був з 4-ма гранями (далі писатимемо 4d дайс), з 6-ма (6d дайс) чи з 8-ма гранями (8d дайс).

### Крок 1 - Визначимо яку ймовірність ми шукаємо, і що спостерігаємо

Ми хочемо знати ймовірність трьох різних можливостей (коротко – 4d, 6d та 8d). Ми спостерігаємо номер, який нам сказали, що він випав.

### Крок 2 - Оцінка початкових ймовірностей

Оскільки маємо 3 дайси, то початкова ймовірність вибору кожного з них однакова і дорівнює  $1/3$  (якщо нема додаткової інформації).



### **Крок 3 - Для кожної початкового випадку, обчислюємо можливість отримання нашого спостереження**

Наше спостереження – випало число 2. Якщо ми припускаємо, що наш друг кидав 4d дайса, то ймовірність випадання числа 2 дорівнює  $1/4$ . Якщо 6d – то  $1/6$ . Якщо 8d – то  $1/8$ .

### **Крок 4 – Перемножаємо отримані раніше ймовірності**

Для 4d дайса: мали ймовірність вибору дайса  $1/3$  та ймовірність  $1/4$  випадання числа 2. Добуток дорівнює  $1/3 * 1/4 = 1/12$ .

Для 6d дайса: добуток дорівнює  $1/3 * 1/6 = 1/18$ .

Для 8d дайса: добуток дорівнює  $1/3 * 1/8 = 1/24$ .





## Крок 5 – Нормалізуємо результати

Сумарна ймовірність вибору одного 3-х дайсів і випадання числа 2 складає

$$1/12 + 1/18 + 1/24 = 13 / 72.$$

Це число менше ніж 1, бо ймовірність випадання числа 2 менша ніж 1. Проте ми знаємо, що число 2 вже випало. Тому нам треба поділити ймовірність для кожного дайса на 13/72 для того, щоб сума усіх ймовірностей для усіх дайсів складала 1.

Цей процес відомий як нормалізація.

Після нормалізації знаходимо ймовірність того, що дайс є одним з вибраних:

$$4d \text{ дайс: } (1/12) / (13/72) = (1*72) / (12*13) = 6/13$$

$$6d \text{ дайс: } (1/18) / (13/72) = (1*72) / (18*13) = 4/13$$

$$8d \text{ дайс: } (1/24) / (13/72) = (1*72) / (24*13) = 3/13$$

Це і є відповідь.

Коли ми почали розв'язувати задачу, ми припускали, що ймовірність вибору кожного дайса дорівнює 33.3%. Після кочення одного дайса і отримання числа 2 ми отримали, що вибір 4d дайса має ймовірність 46.1%, вибір 6d дайса – 30.8%, а 8d дайса – 23.1%.



## Теорема Баєса (Bayes): Термінологія

Досі ми не вживали частину технічної термінології притаманної для теореми Баєса. Зараз, настав час навести її.

Початкова ймовірність, тобто ймовірність кожної можливості перед спогляданням даних, називається **Prior (Пріор)**.

Умовна ймовірність, тобто ймовірність спостереження даних в кожному конкретному можливому випадку, називається **Likelihood (Вірогідність)**.

Сумарна ймовірність, тобто сума добутків проміжних обчислень, називається **Normalizing Constant (Нормалізуюча Константа)**.

Нормалізована кінцева відповідь після обчислення ймовірностей кожного можливого випадку називається **Posterior (Постеріор)**.

$$\overset{\text{Posterior}}{P(A|B)} = \frac{\overset{\text{Prior}}{P(A)} * \overset{\text{Likelihood}}{P(B|A)}}{\underset{\text{Normalizing Constant}}{P(B)}} \quad (1)$$



Далі користуватимемося саме такими термінами. Повернемося до прикладу з попереднього підрозділу. Отож, знову припустимо, що наш друг має 3 дайси (4d, 6d та 8d). Він вибирає один дайс випадково, котить його без показу нам, і повідомляє результат – випало число 2. Треба знайти ймовірність того, що дайс був з 4-ма гранями, з 6-ма чи з 8-ма гранями. Перепишемо таблицю, що описує розв’язок цієї задачі в новій формі:

<b>Die</b> [dai]	<b>Prior</b> [‘praɪə]	<b>Likelihood</b> [‘laɪklihud]	<b>Product</b> [‘prɒdʌkt]	<b>Posterior</b> [po’stirɪə]
4d				
6d				
8d				
<b>Normalizing Constant</b>				





Die	Prior	Likelihood	Product	Posterior
4d	$1/3$	$1/4$	$1/12$	$6/13 \approx 0.46$ (46%)
6d	$1/3$	$1/6$	$1/18$	$4/13 \approx 0.31$ (31%)
8d	$1/3$	$1/8$	$1/24$	$3/13 \approx 0.23$ (23%)
Normalizing Constant			$13/72$	

Таблиця 2: Випало 2.

Висновок – найімовірніше наш друг кидав 4d дайс.



**Приклад.** Складіть таблицю, що відповідає випадку випадання на дайсі числа 5 (замість 2). Цей результат принципово відрізняється від першого – від дає можливість нам зразу зробити висновок – вибраний дайс точно не 4-сторонній (бо на 4d дайсі просто нема числа 5).

**Відповідь:**



Die	Prior	Likelihood	Product	Posterior
4d	1/3	0	0	0
6d	1/3	1/6	1/18	$4/7 \approx 0.57$ (57%)
8d	1/3	1/8	1/24	$3/7 \approx 0.43$ (43%)
Normalizing Constant			7/12	

Таблиця 3: Випало 5.

Висновок – цього разу найімовірніше, що наш друг кидав 6d дайс (57%).

Давайте знову повернемося до випадку, коли випало число 2. Висновок з отриманої нами раніше таблиці – найімовірніше наш друг кидав 4d дайс. Проте отримана ймовірність цієї події (46%) мало відрізняється від ймовірності вибору 6d дайса (31%). Подивимося чи допоможе нам отримати більш ймовірний результат наявність додаткової інформації.



Отож, ми просимо друга кинути вибраний ним дайс ще раз. Він погоджується і, наприклад, повідомляє, що отриманий результат – випало число 5. Складемо відповідну таблицю. Оскільки спершу було число 2, то в колонці **Prior** нової таблиці для числа 5 стоятимуть числа з колонки **Posterior** таблиці для числа 2

Die	Prior	Likelihood	Product	Posterior
4d	6/13	0	0	0
6d	4/13	1/6	4/78	$16/25 \approx 0.64$ (64%)
8d	3/13	1/8	3/104	$9/25 \approx 0.36$ (36%)
Normalizing Constant			25/312	

Таблиця 4: Випало 2, потім 5.

Висновок – цього разу найімовірніше, що наш друг кидав 6d дайс. Ймовірність цієї події 64% і ця ймовірність більша за ймовірність кидання 6d дайса при випаданні числа 5 без інформації, що перед тим випало число 2 (57%) і більша за ймовірність вибору довільного дайсу (33%).





**Гіпотеза:** зі збільшенням інформації про випадання чисел на вибраному дайсі росте ймовірність, що характеризує вибраний дайс.

Перевіримо цю гіпотезу ще раз. Отож, ми просимо друга кинути вибраний ним дайс ще раз (вже було спершу 2 а потім 5). Він каже, що випало число 4. результати наших обчислень наведено в такій таблиці

Die	Prior	Likelihood	Product	Posterior
4d	0	1/4	0	0
6d	16/25	1/6	16/150	$64/91 \approx 0.70$ (70%)
8d	9/25	1/8	9/200	$27/91 \approx 0.30$ (30%)
Normalizing Constant			91/600	

Таблиця 5: Випало 2, потім 5, потім 4.

Висновок – це, мабуть, 6d дайс (ймовірність дорівнює вже аж 70%).

Д/З: Прорахувати випадок випадання 5-2-4.



## Приклад 2 - Більше дайсів (die). Більше кочень (rolls)

Давайте розглянемо випадок 6-ти можливих дайсів: 4d, 6d, 8d, 10d, 12d та 20d. Ми збираємося котити вибраний гральний дайс 15 разів і з огляду на отримані результати знайти ймовірність того, що це якийсь конкретний дайс.

Для обчислень використаємо Excel, щоб створити в цій програмі ймовірністну таблицю. Дані в ній подібні до даних з попереднього прикладу проте перетворимо її до вигляду, в якому зручніше враховувати багато випробувань. Отож, раніше ми мали Табл. 2 для ситуації “випало число 2”:

Die	Prior	Likelihood	Product	Posterior
4d	1/3	1/4	1/12	$6/13 \approx 0.46$ (46%)
6d	1/3	1/6	1/18	$4/13 \approx 0.31$ (31%)
8d	1/3	1/8	1/24	$3/13 \approx 0.23$ (23%)
Normalizing Constant			13/72	



Тепер запишемо цю інформацію у такому вигляді:

		<b>Die</b> →	4d	6d	8d
Номер кидання ↓	Число, що спостері гаємо ↓	<b>Prior</b> →	1/3	1/3	1/3
1	2	<b>Posterior</b> № 1 →	6/1 3	4/13	3/13

Таблиця 6: Інформація про випадання числа 2 (в новому вигляді).

Далі кожне нове кидання дайсу дасть нам нову стрічку з фінальними ймовірностями. Всі обчислення зробимо в Excel.



Автозбереження Problem\_2 - Excel

Файл Основне Вставлення Макет сторінки Формули Да

Вставити Шрифт Вирівнювання Число

В14 =RANDBETWEEN(1;\$B\$9)

	A	B	C	D	E
8					
9		8			
10					
11					
12					
13	Номер	Випадковий результат	Результат		
14	1	4	4		
15	2	8	5		
16	3	6	7		
17	4	8	5		
18	5	3	4		
19	6	8	5		
20	7	5	3		
21	8	4	6		
22	9	8	7		
23	10	8	4		
24	11	1	2		
25	12	5	5		
26	13	1	4		
27	14	2	4		
28	15	6	7		
29					

Рис. 1: Генерування випадкових чисел в Excel





## Генерація випадкових результатів кочень

В задачі 1 ми припустили, що на дайсі випало 2, як ілюстративний приклад. Проте коли мова йде про багато експериментів, то люди, взагалі кажучи, погано вибирають справді випадкові номери. Тому для генерації 15 випадкових результатів кочень дайсу використаємо функцію = RANDBETWEEN().

Отож, створимо Excel-файл і в комірку B9 запишемо число, що відповідає за вибраний нами дайс, наприклад, число 8 (вважаємо, що вибрано 8d дайс). Починаючи з комірки A14 і вниз генеруємо номери експерименту (від 1 до 15). В клітинку B14 записуємо формулу для генерації випадкових чисел = RANDBETWEEN(1;\$B\$9) і заповнюємо нею клітинки B14:B28. Затим виділяємо їх і копіюємо в буфер обміну. Стаємо в клітинку C14 і виконуємо команди **Основне => Буфер обміну => Вставити => Вставити значення**. Після цього в клітинках C14:C28 матимемо незмінні результати першого виконання команди = RANDBETWEEN() (див. Рис. 1).



## Початок формування таблиці відповідей

Для зручності, за допомогою кнопки “+” внизу вікна Excel, створимо новий аркуш Excel і, клацнувши мишкою двічі по його назві, назовемо його “Таблиця”. Аркуш1 перейменуємо на “Допоміжні обчислення” (див. Рис. 2).

6	5	5		
7	4	6		
8	7	7		
9	6	2		
10	4	7		
11	7	7		
12	3	1		
13	8	5		
14	5	7		
15	6	6		

Допоміжні обчислення | Таблиця | +

Рис. 2. Новий аркуш

Перейдемо на аркуш Таблиця і створимо там таку таблицю, зображену на Рис. 3 (коміркам, в яких стоять НЕ числа зробить текстовий формат).



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1			Допоміжна Таблиця для проведення обчислень							*		Таблиця нормалізованих ймовірностей					
2		Die →	4	6	8	10	12	20		*	Die →	4	6	8	10	12	20
3	№ Roll ↓	Results ↓	-	-	-	-	-	-	Sum ↓	*	№ Roll ↓	-	-	-	-	-	-
4	0	-								*	0						
5	1									*	1						
6	2									*	2						
7	3									*	3						
8	4									*	4						
9	5									*	5						
10	6									*	6						
11	7									*	7						
12	8									*	8						
13	9									*	9						
14	10									*	10						
15	11									*	11						
16	12									*	12						
17	13									*	13						
18	14									*	14						
19	15									*	15						
20																	

Рис. 3. Структура таблиці

## Початкова Ймовірність

Наступний крок – ввести початкову ймовірність. Оскільки є шість дайсів і ми припускаємо, що вони випадково відтягуються від сумки, то початкова ймовірність вибору кожного дайсу дорівнює числу  $1/6$ , тобто наближено – числу 0.1667.

Заповнюємо цими числами діапазони комірок C4:H4 та L4:Q4, збільшуючи при потребі відображення на екрані кількості розрядів числа. Також заповніть колонку Results результатами, випадковим чином згенерованими на першому аркуші, а в комірку I4 поміщаємо формулу =SUM(C4:H4) (в ідеалі результат мав би дорівнювати одиниці, але за рахунок наближених обчислень тут матимемо 1,0002) (див. рис. 4).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1			Допоміжна Таблиця для проведення обчислень							*		Таблиця нормалізованих ймовірностей					
2		Die →	4	6	8	10	12	20		*	Die →	4	6	8	10	12	20
3	№ Roll ↓	Results ↓	-	-	-	-	-	-	Sum ↓	*	№ Roll ↓	-	-	-	-	-	-
4	0	-	0,1667	0,1667	0,1667	0,1667	0,1667	0,1667	1,0002	*	0	0,1667	0,1667	0,1667	0,1667	0,1667	0,1667
5	1	6								*	1						
6	2	3								*	2						

Рис. 4. Заповнення таблиці початковими даними.





## Ймовірність випадання числа для кожного дайсу

Зауважимо таке: якщо число, що випало, більше за розмірність дайсу, то ймовірність його появи на дайсі складає 0 (тобто, якщо випало 6, то ймовірність появи цього числа на 4d дайсі нульова);

якщо число, що випало, менше за розмірність дайсу, то ймовірність його появи на дайсі складає  $1/(\text{розмірність дайсу})$  (тобто, якщо випало 6, то ймовірність появи цього числа на 8d дайсі дорівнює  $1/8$ );

Для обчислень, що призводять до знаходження Posterior використаємо оператор

=IF(equation to evaluate; value if true; value if false)

Таким чином маємо записати таке:

=IF(результат кочення > розмірність дайсу; 0;  
1/( розмірність дайсу) \* попередню  
нормаліз. ймовірність ) (2)

Отож, в комірку C5 маємо записати чисельник дроби з формули (1), а саме

**Likelihood\*Prior.**



Вираз Likelihood обчислюємо за формулою (2), порівнюючи вміст комірок B5 та C2, а Prior для комірки C5 міститься в комірці L4 – це нормалізована ймовірність для попереднього кидання (тут попереднього ще не було і тому тут – це 0,1667). Отже, в комірку C5 записуємо таку формулу

$$=IF(\$B5>C\$2;0;(1/C\$2)*L4)$$

Копіюємо цю формулу в комірки D5:H5. В комірку I5 копіюємо формулу з I4 – це **Normalizing Constant**. Після цього записуємо **Posterior** в комірку L5

$$=\$C5/\$I5$$

та заповнюємо цією формулою діапазок комірок M5:Q5. В результаті в комірках L5:Q5 містяться **Posterior** для результату експерименту «Кинули дайс і випало число 6» (див. Рис. 5)



L5    :    ✕    ✓     $\sum$     =\$C5/\$I5

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1			Допоміжна Таблиця для проведення обчислень							*			Таблиця нормалізованих ймовірностей				
2		Die →	4	6	8	10	12	20		*	Die →	4	6	8	10	12	20
3	№ Roll ↓	Results ↓	-	-	-	-	-	Sum ↓	*	№ Roll ↓	-	-	-	-	-	-	-
4	0	-	0,1667	0,1667	0,1667	0,1667	0,1667	0,1667	1,0002	*	0	0,1667	0,1667	0,1667	0,1667	0,1667	0,1667
5	1	6	0,0000	0,0278	0,0208	0,0167	0,0139	0,0083	0,0875	*	1	0,0000	0,3175	0,2381	0,1905	0,1587	0,0952

Рис. 5. Повні дані для першого експерименту

Висновки про ймовірності вибору дайсів подамо у таблиці

4d	6d	8d	10d	12d	20d
0 %	32 %	24 %	19 %	16 %	10 %

Таблиця 7: Інформація про випадання числа 6.

Зауважимо, що оскільки ми проводимо наближені обчислення, то сума останнього рядка в Таб. 1 може НЕ дорівнювати 100 %.



**Приклад 2.1.** Заповніть стрічку № 6 у Excel файлі, яка відповідає за експеримент «Другий раз кинути дайс і випало число 3». Продумайте спосіб модифікації (захисту комірок) та копіювання вже заповнених комірок існуючого Excel-файлу.

**Відповідь:** див. Рис. 6 та Таб. 8.

C6 : X ✓ fx =IF(\$B6>C\$2;0;(1/C\$2)\*L5)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1			Допоміжна Таблиця для проведення обчислень							*			Таблиця нормалізованих ймовірностей				
2		Die →	4	6	8	10	12	20		*	Die →	4	6	8	10	12	20
3	№ Roll ↓	Results ↓	-	-	-	-	-	-	Sum ↓	*	№ Roll ↓	-	-	-	-	-	-
4	0	-	0,1667	0,1667	0,1667	0,1667	0,1667	0,1667	1,0002	*	0	0,1667	0,1667	0,1667	0,1667	0,1667	0,1667
5	1	6	0,0000	0,0278	0,0208	0,0167	0,0139	0,0083	0,0875	*	1	0,0000	0,3175	0,2381	0,1905	0,1587	0,0952
6	2	3	0,0000	0,0529	0,0298	0,0190	0,0132	0,0048	0,1197	*	2	0,0000	0,4420	0,2486	0,1591	0,1105	0,0398

Рис. 6. Повні дані для другого експерименту

4d	6d	8d	10d	12d	20d
0 %	44 %	25 %	16 %	11 %	4 %

Таблиця 8: Інформація про випадання числа 3.





**Приклад 2.2.** Заповніть стрічку № 7 у Excel файлі, яка відповідає за експеримент «Третій раз кинули дайс і випало число 2». Продумайте спосіб копіювання вже заповнених комірок існуючого Excel-файлу.

**Відповідь:** див. Рис. 7 та Таб. 9.

C7 : X ✓ fx =IF(\$B7>C\$2;0;(1/C\$2)\*L6)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1			Допоміжна Таблиця для проведення обчислень							*		Таблиця нормалізованих ймовірностей					
2		Die →	4	6	8	10	12	20		*	Die →	4	6	8	10	12	20
3	№ Roll ↓	Results ↓	-	-	-	-	-	-	Sum ↓	*	№ Roll ↓	-	-	-	-	-	-
4	0	-	0,1667	0,1667	0,1667	0,1667	0,1667	0,1667	1,0002	*	0	0,1667	0,1667	0,1667	0,1667	0,1667	0,1667
5	1	6	0,0000	0,0278	0,0208	0,0167	0,0139	0,0083	0,0875	*	1	0,0000	0,3175	0,2381	0,1905	0,1587	0,0952
6	2	3	0,0000	0,0529	0,0298	0,0190	0,0132	0,0048	0,1197	*	2	0,0000	0,4420	0,2486	0,1591	0,1105	0,0398
7	3	2	0,0000	0,0737	0,0311	0,0159	0,0092	0,0020	0,1319	*	3	0,0000	0,5587	0,2357	0,1207	0,0698	0,0151

Рис. 7. Повні дані для третього експерименту

4d	6d	8d	10d	12d	20d
0 %	56 %	24 %	12 %	7 %	2 %

Таблиця 9: Інформація про випадання числа 2.

**Приклад 2.3.** Внесіть дані всіх 15-х експериментів у Excel файл. Продумайте спосіб копіювання вже заповнених комірок існуючого Excel-файлу.

**Відповідь:** див. Рис. 8 та Таб. 10.

C19 : =IF(\$B19>C\$2;0;(1/C\$2)\*L18)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1			Допоміжна Таблиця для проведення обчислень							*		Таблиця нормалізованих ймовірностей					
2		Die →	4	6	8	10	12	20		*	Die →	4	6	8	10	12	20
3	N° Roll ↓	Results ↓	—	—	—	—	—	Sum ↓	*	N° Roll ↓	—	—	—	—	—	—	—
4	0	—	0,1667	0,1667	0,1667	0,1667	0,1667	0,1667	1,0002	*	0	0,1667	0,1667	0,1667	0,1667	0,1667	0,1667
5	1	6	0,0000	0,0278	0,0208	0,0167	0,0139	0,0083	0,0875	*	1	0,0000	0,3175	0,2381	0,1905	0,1587	0,0952
6	2	3	0,0000	0,0529	0,0298	0,0190	0,0132	0,0048	0,1197	*	2	0,0000	0,4420	0,2486	0,1591	0,1105	0,0398
7	3	2	0,0000	0,0737	0,0311	0,0159	0,0092	0,0020	0,1319	*	3	0,0000	0,5587	0,2357	0,1207	0,0698	0,0151
8	4	6	0,0000	0,0931	0,0295	0,0121	0,0058	0,0008	0,1412	*	4	0,0000	0,6594	0,2086	0,0855	0,0412	0,0053
9	5	8	0,0000	0,0000	0,0261	0,0085	0,0034	0,0003	0,0383	*	5	0,0000	0,0000	0,6805	0,2230	0,0896	0,0070
10	6	5	0,0000	0,0000	0,0851	0,0223	0,0075	0,0003	0,1152	*	6	0,0000	0,0000	0,7385	0,1936	0,0648	0,0030
11	7	6	0,0000	0,0000	0,0923	0,0194	0,0054	0,0002	0,1172	*	7	0,0000	0,0000	0,7875	0,1651	0,0461	0,0013
12	8	7	0,0000	0,0000	0,0984	0,0165	0,0038	0,0001	0,1189	*	8	0,0000	0,0000	0,8282	0,1389	0,0323	0,0005
13	9	2	0,0000	0,0000	0,1035	0,0139	0,0027	0,0000	0,1201	*	9	0,0000	0,0000	0,8617	0,1157	0,0224	0,0002
14	10	7	0,0000	0,0000	0,1077	0,0116	0,0019	0,0000	0,1212	*	10	0,0000	0,0000	0,8890	0,0955	0,0154	0,0001
15	11	7	0,0000	0,0000	0,1111	0,0095	0,0013	0,0000	0,1220	*	11	0,0000	0,0000	0,9112	0,0783	0,0105	0,0000
16	12	1	0,0000	0,0000	0,1139	0,0078	0,0009	0,0000	0,1226	*	12	0,0000	0,0000	0,9290	0,0638	0,0072	0,0000
17	13	5	0,0000	0,0000	0,1161	0,0064	0,0006	0,0000	0,1231	*	13	0,0000	0,0000	0,9433	0,0519	0,0048	0,0000
18	14	7	0,0000	0,0000	0,1179	0,0052	0,0004	0,0000	0,1235	*	14	0,0000	0,0000	0,9547	0,0420	0,0033	0,0000
19	15	6	0,0000	0,0000	0,1193	0,0042	0,0003	0,0000	0,1238	*	15	0,0000	0,0000	0,9639	0,0339	0,0022	0,0000

Рис. 8. Повні дані всіх експериментів



4d	6d	8d	10d	12d	20d
0 %	99 %	1 %	0 %	0 %	0 %

Таблиця 10: Інформація про результат всіх експериментів.

## Візуалізація результатів

Для відображення отриманих результатів скопіюємо діапазон нормалізованих ймовірностей K2:Q19 і вставимо їх на перший аркуш, починаючи з клітинки A20. Видаливши відповідні текстові поля, отримаємо такі дані експериментів

	A	B	C	D	E	F	G
20							
21		4	6	8	10	12	20
22	0	0,1667	0,1667	0,1667	0,1667	0,1667	0,1667
23	1	0,0000	0,3175	0,2381	0,1905	0,1587	0,0952
24	2	0,0000	0,4420	0,2486	0,1591	0,1105	0,0398
25	3	0,0000	0,5587	0,2357	0,1207	0,0698	0,0151
26	4	0,0000	0,6594	0,2086	0,0855	0,0412	0,0053
27	5	0,0000	0,0000	0,6805	0,2230	0,0896	0,0070
28	6	0,0000	0,0000	0,7385	0,1936	0,0648	0,0030
29	7	0,0000	0,0000	0,7875	0,1651	0,0461	0,0013
30	8	0,0000	0,0000	0,8282	0,1389	0,0323	0,0005
31	9	0,0000	0,0000	0,8617	0,1157	0,0224	0,0002
32	10	0,0000	0,0000	0,8890	0,0955	0,0154	0,0001
33	11	0,0000	0,0000	0,9112	0,0783	0,0105	0,0000
34	12	0,0000	0,0000	0,9290	0,0638	0,0072	0,0000
35	13	0,0000	0,0000	0,9433	0,0519	0,0048	0,0000
36	14	0,0000	0,0000	0,9547	0,0420	0,0033	0,0000
37	15	0,0000	0,0000	0,9639	0,0339	0,0022	0,0000

Рис. 9. Нормалізовані ймовірності (таблиця)





Підсумовуючи отримані результати бачимо, що після 15-и кочень ми з ймовірністю 96.4% можемо стверджувати, що спершу було вибрано 8d дайсі. Крім того, з ймовірністю 3.4% -- вибрали 10d дайс, ймовірність вибору 12d дайсу дуже мала, а 4d, 6d та 20d дайсів – нульова.

Тепер поглянемо про тенденцію зміни Posterior. Виділивши діапазон A21:G37 натискаємо **Вставлення => Діаграми => Рекомендовані діаграми => Лінійчаста діаграма** та після певного оформлення результату, отримуємо діаграму з Рис. 10



Результат	Номер	
6	1	
3	2	
2	3	
6	4	
8	5	
5	6	
6	7	
7	8	
2	9	
7	10	
7	11	
1	12	
5	13	
7	14	
6	15	

## Ймовірність вибору дайсів vs Кочення

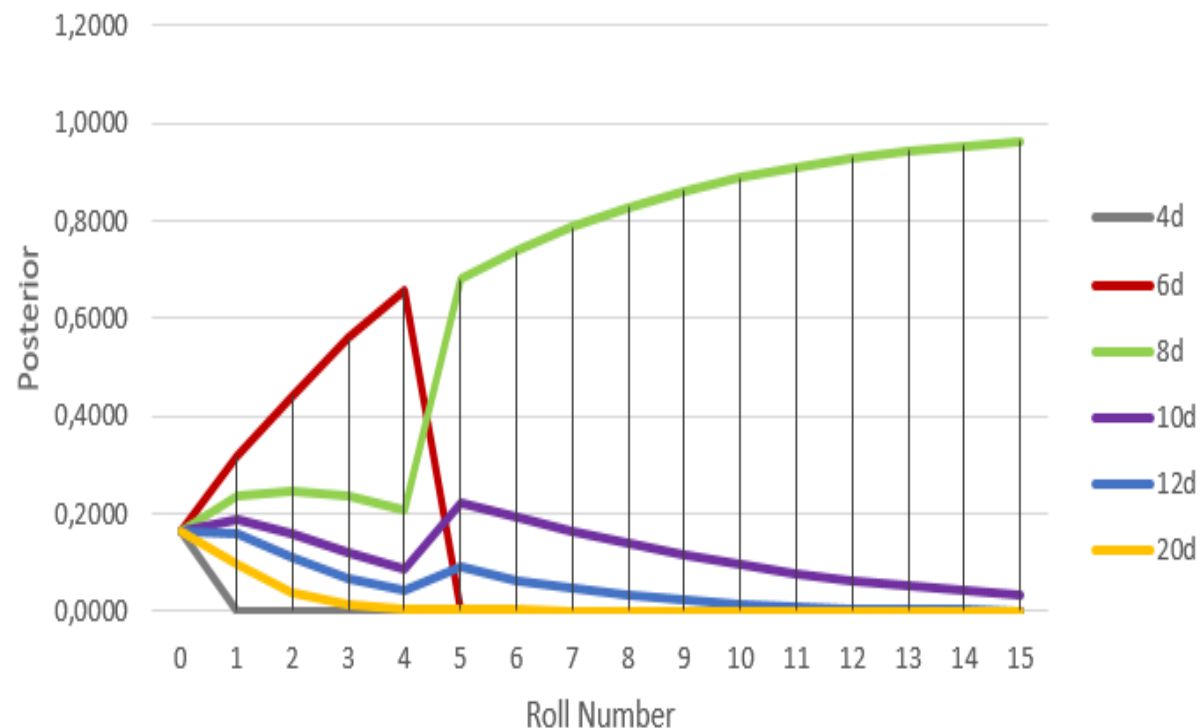


Рис. 10. Нормалізовані ймовірності (графік)

Аналіз Рис. 10 більш наочний:

Після першого кочення ймовірність вибору 4d дайса стає нульовою. Це природно, бо випало число 6, якого нема на 4d дайсі.

До четвертого кочення фаворитом вибору є 6d дайс. Але на 5-му коченні випадає 8, яке виключає з розгляду 6d дайс і лідером вибору стає 8d дайс. Ймовірність того, що вибрано 10d, 12d чи 20d дуже мала, хоча і відмінна від нуля.



**Дякую за увагу!**