Теорія ймовірностей:
Лекція 5
Умовна ймовірність

Данило Тавров

15 вересня 2025 р.

План лекції

План лекції

Інтуїція

- Визначення ймовірности (як міри) **формалізує** опис **ступенів упевнености** та **невизначеностей**, пов'язаних із подіями та явищами
- Часто надходження **нової інформації** може **змінити** наші уявлення про ймовірності відповідних подій

Приклад (КЛ3.1.1)

- Нехай одного ранку ми хочемо оцінити ймовірність події $R =$«піде дощ», $\mathbb{P}(R)$
- Якщо $C =$«за вікном хмари», ми вважатимемо, що ймовірність дощу вища: $\mathbb{P}(R) \Rightarrow \mathbb{P}(R \mid C)$
- Такий перехід називають **обумовленням** подією C (conditioning on C)
- Із плином дня можуть надходити нові й нові факти $B_1, \ldots, B_n$, і в результаті матимемо $\mathbb{P}(R \mid C, B_1, \ldots B_n) \equiv \mathbb{P}(R \mid C \cap B_1 \cap \ldots \cap B_n)$ $\square$

Визначення (КЛ3.1.2)

- Нехай маємо вимірний простір $(\Omega, \mathcal{A})$
- Нехай $A, B \in \mathcal{A}$ — деякі події
- Нехай $\mathbb{P}(B) > 0$
- **Умовна ймовірність** A за умови B (conditional probability of A given B):

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \tag{КЛ3.1.1}$$

- $\mathbb{P}(A)$ — **апріорна ймовірність** (prior)
- $\mathbb{P}(A \mid B)$ — **апостеріорна ймовірність** (posterior)

□

Особливості обумовлення

- Нескладно бачити, що

$$\mathbb{P}\left(A \mid A\right) = \frac{\mathbb{P}\left(A \cap A\right)}{\mathbb{P}\left(A\right)} = 1$$

- Якщо ми знаємо, що подія A сталася, то її ймовірність дорівнює 1

Зауваження (КЛ3.1.3)

- **Не існує такої події, як** $A \mid B$
- **Це просто позначення!**
- Насправді, існують **дві різні ймовірнісні міри**:

$$\mathbb{P}\left(\cdot\right) : \mathcal{A} \to \mathbb{R}^{+}\ , \qquad \mathbb{P}\left(\cdot \mid B\right) : \mathcal{A} \to \mathbb{R}^{+}$$

- Іншими словами, з імовірнісного простору $(\Omega, \mathcal{A}, \mathbb{P}\left(\cdot\right))$ можна утворити **новий** імовірнісний простір $(\Omega, \mathcal{A}, \mathbb{P}\left(\cdot \mid B\right))$, $B \in \mathcal{A}$

$\square$

Умовна ймовірність — це ймовірність

Твердження (КЛ3.1.4)

Умовна ймовірність $\mathbb{P}(\cdot \mid B)$ така, що

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \quad A, B \in \mathcal{A}, \ \mathbb{P}(B) > 0$$

є ймовірнісною мірою

Доведення.

- Потрібно перевірити виконання всіх трьох аксіом імовірнісної міри
- $\mathbb{P}(\cdot \mid B)$ невід'ємна, оскільки $\mathbb{P}(\cdot)$ невід'ємна
- Імовірність усього простору дорівнює

$$\mathbb{P}(\Omega \mid B) = \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$$

Доведення.

Продовження...

- Якщо $A_1, A_2, \ldots \in \mathcal{A}$ — послідовність несумісних подій, то

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i \mid B\right) = \frac{1}{\mathbb{P}(B)} \cdot \mathbb{P}\left(\left(\bigcup_{i=1}^{\infty} A_i\right) \cap B\right) = \frac{1}{\mathbb{P}(B)} \cdot \mathbb{P}\left(\bigcup_{i=1}^{\infty}(A_i \cap B)\right)$$
$$= \frac{1}{\mathbb{P}(B)} \cdot \sum_{i=1}^{\infty} \mathbb{P}(A_i \cap B) = \sum_{i=1}^{\infty} \mathbb{P}(A_i \mid B)$$

- Отже σ-адитивність виконується

$\square$

Приклад (КЛ3.1.5)

- Нехай маємо добре перетасовану колоду з 52 гральних карт
- Розгляньмо витяг двох карт **без повторень**:
  - Нехай $A =$«перша карта — чирва»
  - Нехай $B =$«друга карта — червона»
- Імовірність перетину цих подій:

$$\mathbb{P}\left(A \cap B\right) = \frac{13}{52} \cdot \frac{25}{51} = \frac{25}{204}$$

  - Перший витяг — 13 чирвових карт із 52 можливих
  - Другий витяг — 25 червоних із 51, що лишилися
- Імовірність кожної події окремо: $\mathbb{P}\left(A\right) = \frac{1}{4}$, $\mathbb{P}\left(B\right) = \frac{1}{2}$:
  - Перша карта може бути будь-якої з 4 мастей
  - Друга карта може бути будь-якого з 2 кольорів
- Тоді умовні ймовірності дорівнюють

$$\mathbb{P}\left(A \mid B\right) = \frac{\mathbb{P}\left(A \cap B\right)}{\mathbb{P}\left(B\right)} = \frac{25}{102}\,, \qquad \mathbb{P}\left(B \mid A\right) = \frac{\mathbb{P}\left(B \cap A\right)}{\mathbb{P}\left(A\right)} = \frac{25}{51}$$

Зауваження (КЛ3.1.6)

- Обумовлення однієї події іншою **не є комутативним**: $\mathbb{P}(A \mid B) \neq \mathbb{P}(B \mid A)$
- Обидві ймовірності мають сенс, **причиново-наслідкові зв'язки** між подіями **не важливі**
- Ми просто враховуємо інформацію, яку одна подія може надати про іншу
- Замість «першої» та «другої» карти можна розглядати «карту в лівій руці» і «карту в правій руці», які витягнуто **одночасно**

$\square$

- Обумовлення події A подією B означає, що ми **перенормалізовуємо** ймовірнісну міру

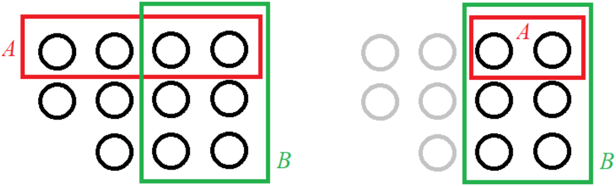


Рис. КЛ3.1.1: $\mathbb{P}(A) = \frac{4}{11}$, проте $\mathbb{P}(A \mid B) = \frac{2}{6} = \frac{1}{3}$

- Пояснення для двох підходів до інтерпретації ймовірностей:
  - **Частотна інтерпретація**: $\mathbb{P}(A \mid B)$ як відносна частота події A **серед тих експериментів**, які закінчилися подією B
  - **Беєсівська інтерпретація**: оновлення суб'єктивної ймовірности після надходження додаткової інформації

Ланцюгове правило для умовних імовірностей

- Із визначення випливає:

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \quad \Rightarrow \quad \mathbb{P}(A \cap B) = \mathbb{P}(B)\,\mathbb{P}(A \mid B)\ ,\ \mathbb{P}(B) > 0$$
$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \quad \Rightarrow \quad \mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B \mid A)\ ,\ \mathbb{P}(A) > 0 \tag{КЛ3.1.2}$$

- Відповідну властивість можна узагальнити

Твердження (КЛ3.1.7)

- Нехай $\mathbb{P}(A_1 \cap \ldots \cap A_n) > 0$
- Тоді

$$\begin{aligned}\mathbb{P}(A_1 \cap \ldots \cap A_n) &= \mathbb{P}(A_1)\,\mathbb{P}(A_2 \mid A_1)\,\mathbb{P}(A_3 \mid A_1, A_2) \ldots \mathbb{P}(A_n \mid A_1, \ldots, A_{n-1}) \\ &= \mathbb{P}(A_2)\,\mathbb{P}(A_1 \mid A_2)\,\mathbb{P}(A_3 \mid A_1, A_2) \ldots \mathbb{P}(A_n \mid A_1, \ldots, A_{n-1}) \\ &= \ldots \end{aligned} \tag{КЛ3.1.3}$$

і т.д. для всіх можливих комбінацій подій

Доведення.

- Доведення здійснюватимемо за індукцією
- **База індукції**: для $n = 2$ теорема є просто повторенням (КЛ3.1.2):

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\,\mathbb{P}(A_2 \mid A_1)$$

- **Індукційний перехід:** нехай теорема виконується для $n-1$:

$$\mathbb{P}(A_1 \cap \ldots \cap A_{n-1}) = \mathbb{P}(A_1)\,\mathbb{P}(A_2 \mid A_1) \ldots \mathbb{P}(A_{n-1} \mid A_1, \ldots, A_{n-2})$$

- Тоді

$$\begin{aligned}
\mathbb{P}(A_1 \cap \ldots \cap A_n) &= \mathbb{P}((A_1 \cap \ldots \cap A_{n-1}) \cap A_n) \\
&= \mathbb{P}(A_1 \cap \ldots \cap A_{n-1}) \cdot \mathbb{P}(A_n \mid A_1, \ldots, A_{n-1}) \\
&= \mathbb{P}(A_1)\,\mathbb{P}(A_2 \mid A_1) \ldots \mathbb{P}(A_{n-1} \mid A_1, \ldots, A_{n-2}) \\
&\times \mathbb{P}(A_n \mid A_1 \cap \ldots \cap A_{n-1})
\end{aligned}$$

□

Розбиття на підзадачі

- Обумовлення корисно використовувати для розбиття складних задач на простіші підзадачі

Приклад (КЛ3.1.8)

- Нехай у вазі містяться 10 однакових кульок, 5 із яких — чорні, 3 — червоні і 2 — білі
- Нехай по черзі і **без повторень** витягають 4 кульки
- Чому дорівнює ймовірність того, що:
  - Перша кулька — чорна (B_1)?
  - Друга кулька — червона (R_2)?
  - Третя кулька — біла (W_3)?
  - Четверта — знову чорна (B_4)?
- Згідно з ланцюговим правилом,

$$\mathbb{P}(B_1 \cap R_2 \cap W_3 \cap B_4) = \mathbb{P}(B_4 \mid B_1, R_2, W_3)\,\mathbb{P}(W_3 \mid B_1, R_2)\,\mathbb{P}(R_2 \mid B_1)\,\mathbb{P}(B_1)$$

□

Приклад

Продовження...

- Кожну окрему ймовірність легко обчислити:

$$\mathbb{P}\left(B_1\right) = \frac{5}{10}, \quad \mathbb{P}\left(R_2 \mid B_1\right) = \frac{3}{9}, \quad \mathbb{P}\left(W_3 \mid B_1, R_2\right) = \frac{2}{8},$$
$$\mathbb{P}\left(B_4 \mid B_1, R_2, W_3\right) = \frac{4}{7}$$

- Отже

$$\mathbb{P}\left(B_1 \cap R_2 \cap W_3 \cap B_4\right) = \frac{4}{7} \cdot \frac{2}{8} \cdot \frac{3}{9} \cdot \frac{5}{10} \approx 0.024$$

□

План лекції

- Задача від Мартина Гарднера (Martin Gardner, 1914–2010) у журналі *Scientific American* у 1956 р.:
  Містер Джонс має двох дітей. ***Старша*** *дитина — дівчинка. Яка ймовірність того, що в нього дві дівчинки?*
  Містер Смит має двох дітей. ***Принаймні один*** *із них — хлопчик. Яка ймовірність того, що в нього два хлопчики?*
- Протягом десятиліть точилися суперечки, як правильно інтерпретувати відповідні твердження
- Розгляньмо обидва випадки в термінах дівчат
- Вважатимемо, що ймовірності народження хлопчика і дівчинки однакові
- Також вважатимемо, що народження дітей незалежні одне від одного

Формалізація задачі

- Щоб дати відповіді на ці питання, спочатку потрібно з'ясувати, який у нас простір елементарних подій та самі події
- Нехай $XY =$«старша дитина — X, молодша дитина — Y», $X, Y \in \{B, G\}$
- Тоді простір елементарних подій є $\Omega = \{GG, GB, BG, BB\}$
- Уведімо символ $\cdot$ на позначення того, що нам непринципово, якої стати дитина
- Наприклад, $G\cdot = \{GG, GB\}$
- Тоді подію $G =$«принаймні одна дитина — дівчинка» можна розписати як $G = (G\cdot) \cup (\cdot G)$

Розв'язання задачі

- Імовірність обох дівчат, якщо старша — дівчинка:

$$\mathbb{P}\left(GG \mid G\cdot\right) = \frac{\mathbb{P}\left(GG \cap (G\cdot)\right)}{\mathbb{P}\left(G\cdot\right)} = \frac{\mathbb{P}\left(\{GG\}\right)}{\mathbb{P}\left(\{GG, GB\}\right)} = \frac{1/4}{1/2} = \frac{1}{2}$$

- Також маємо ймовірність обох дівчат, якщо принаймні одна — дівчинка:

$$\mathbb{P}\left(GG \mid G\right) = \frac{\mathbb{P}\left(GG \cap G\right)}{\mathbb{P}\left(G\right)} = \frac{\mathbb{P}\left(\{GG\}\right)}{\mathbb{P}\left(\{GG, GB, BG\}\right)} = \frac{1/4}{3/4} = \frac{1}{3}$$

Декілька коментарів

- У силу симетрії $\mathbb{P}(GG \mid G\cdot) = \mathbb{P}(GG \mid \cdot G) = 1/2$
- Інформація про **старшу** дівчинку або **молодшу** дівчинку має однакову силу
- Натомість, імовірність $\mathbb{P}(GG \mid G)$ інша
- Це тому, що кажучи про **принаймні** одну дівчинку, ми **не уточнювали**, яка саме
- Обумовлюючи подією $G\cdot$ або $\cdot G$, ми «викидаємо з простору Ω» по два варіанти в кожному випадку (BB, BG та BB, GB, відповідно)
- Обумовлюючи подією G, ми викидаємо тільки **один** варіант — BB

Альтернативна інтерпретація

Містер Смит ***гуляє в парку*** *з дівчинкою. На питання, чи це його дочка, він відповідає ствердно. Яка ймовірність того, що в нього дві дівчинки, якщо на прогулянку він вибрав одного зі своїх дітей абсолютно випадково?*

- У цьому випадку

$$\Omega = \{GG/g\cdot, GG/\cdot g, GB/g\cdot, GB/\cdot b, BG/b\cdot, BG/\cdot g, BB/b\cdot, BB/\cdot b\}$$

- Наприклад, $GB/\cdot b =$«старша дитина — дівчинка, молодша — хлопчик, і на прогулянці побачили хлопчика»
- Тоді маємо обумовлювальну подію $G' =$«на прогулянці містера Смита побачили з дівчинкою»: $G' = \{GG/g\cdot, GG/\cdot g, GB/g\cdot, BG/\cdot g\}$
- Отже

$$\mathbb{P}\left(GG \mid G'\right) = \frac{\mathbb{P}\left(GG \cap G'\right)}{\mathbb{P}\left(G'\right)} = \frac{\mathbb{P}\left(\{GG/g\cdot, GG/\cdot g\}\right)}{\mathbb{P}\left(\{GG/g\cdot, GG/\cdot g, GB/g \cdot BG/\cdot g\}\right)} = \frac{2/8}{4/8} = \frac{1}{2}$$

- Розгляньмо варіант цієї задачі
- Додаткова інформація: пора року народження дитини
- Простір елементарних подій Ω містить елементи такого виду:

$$\{XS, YT\}\ ,\ X, Y \in \{G, B\}\ ,\ S, T \in \{W, Sp, Su, F\}$$

- Наприклад, $GW, BSp =$«старша дитина — дівчинка, що народилася взимку, молодша — хлопчик, що народився навесні»
- Як і раніше, символ $\cdot$ для позначення **будь-якої** стати чи пори року

BW,BW	BW,BSp	BW,BSu	BW,BF	BW,GW	BW,GSp	BW,GSu	BW,GF
BSp,BW	BSp,BSp	BSp,BSu	BSp,BF	BSp,GW	BSp,GSp	BSp,GSu	BSp,GF
BSu,BW	BSu,BSp	BSu,BSu	BSu,BF	BSu,GW	BSu,GSp	BSu,GSu	BSu,GF
BF,BW	BF,BSp	BF,BSu	BF,BF	BF,GW	BF,GSp	BF,GSu	BF,GF
GW,BW	GW,BSp	GW,BSu	GW,BF	GW,GW	GW,GSp	GW,GSu	GW,GF
GSp,BW	GSp,BSp	GSp,BSu	GSp,BF	GSp,GW	GSp,GSp	GSp,GSu	GSp,GF
GSu,BW	GSu,BSp	GSu,BSu	GSu,BF	GSu,GW	GSu,GSp	GSu,GSu	GSu,GF
GF,BW	GF,BSp	GF,BSu	GF,BF	GF,GW	GF,GSp	GF,GSu	GF,GF

Рис. КЛЗ.2.1а

Додаткова інформація

- Чому дорівнює ймовірність того, що обидві дитини — дівчатка, якщо **принаймні одна** з дітей — дівчинка, **що народилася взимку**?
- Маємо обумовлювальну подію $GW = (\cdot\cdot, GW) \cup (GW, \cdot\cdot)$
- За формулою (КЛ3.1.1) маємо:

$$\mathbb{P}(G\cdot, G\cdot \mid GW) = \frac{\mathbb{P}((G\cdot, G\cdot) \cap GW)}{\mathbb{P}(GW)}$$

- Нехай імовірності народження в будь-яку пору року однакові
- Нехай народження дітей незалежні одне від одного
- Імовірність кожної конкретної комбінації «стать-пора року» дорівнює $(1/2) \cdot (1/4) = (1/8)$
- Імовірність усіх комбінацій, **окрім конкретної**, дорівнює $1 - (1/8) = 7/8$
- Відповідно,

$$\mathbb{P}(GW) = 1 - \mathbb{P}(GW^c) = 1 - \left(\frac{7}{8}\right)^2 = \frac{15}{64}$$

Додаткова інформація

- Імовірність події $\mathbb{P}\left((G\cdot, G\cdot) \cap GW\right)$ дорівнює 7/64 (див. рисунок далі)
- Отже
$$\mathbb{P}\left(G\cdot, G\cdot \mid GW\right) = \frac{7/64}{15/64} = \frac{7}{15} > \frac{1}{3}$$
- Це **контрінтуїтивно**: чому інформація про пору року народження несе якесь смислове навантаження?
- Насправді це дає змогу **уточнити**, про якого саме з дітей мова
- Якби ми використали подію «принаймні одна з двох дітей — дівчинка, що народилася 31 березня о 8:30», то результат був би ще ближчий до $\frac{1}{2}$

BW,BW	BW,BSp	BW,BSu	BW,BF	BW,GW	BW,GSp	BW,GSu	BW,GF
BSp,BW	BSp,BSp	BSp,BSu	BSp,BF	BSp,GW	BSp,GSp	BSp,GSu	BSp,GF
BSu,BW	BSu,BSp	BSu,BSu	BSu,BF	BSu,GW	BSu,GSp	BSu,GSu	BSu,GF
BF,BW	BF,BSp	BF,BSu	BF,BF	BF,GW	BF,GSp	BF,GSu	BF,GF
GW,BW	GW,BSp	GW,BSu	GW,BF	GW,GW	GW,GSp	GW,GSu	GW,GF
GSp,BW	GSp,BSp	GSp,BSu	GSp,BF	GSp,GW	GSp,GSp	GSp,GSu	GSp,GF
GSu,BW	GSu,BSp	GSu,BSu	GSu,BF	GSu,GW	GSu,GSp	GSu,GSu	GSu,GF
GF,BW	GF,BSp	GF,BSu	GF,BF	GF,GW	GF,GSp	GF,GSu	GF,GF

Рис. КЛ3.2.1б: події $G\cdot, G\cdot$ (помаранчевий) та $G\cdot$ (синій і помаранчевий разом)

BW,BW	BW,BSp	BW,BSu	BW,BF	BW,GW	BW,GSp	BW,GSu	BW,GF
BSp,BW	BSp,BSp	BSp,BSu	BSp,BF	BSp,GW	BSp,GSp	BSp,GSu	BSp,GF
BSu,BW	BSu,BSp	BSu,BSu	BSu,BF	BSu,GW	BSu,GSp	BSu,GSu	BSu,GF
BF,BW	BF,BSp	BF,BSu	BF,BF	BF,GW	BF,GSp	BF,GSu	BF,GF
GW,BW	GW,BSp	GW,BSu	GW,BF	GW,GW	GW,GSp	GW,GSu	GW,GF
GSp,BW	GSp,BSp	GSp,BSu	GSp,BF	GSp,GW	GSp,GSp	GSp,GSu	GSp,GF
GSu,BW	GSu,BSp	GSu,BSu	GSu,BF	GSu,GW	GSu,GSp	GSu,GSu	GSu,GF
GF,BW	GF,BSp	GF,BSu	GF,BF	GF,GW	GF,GSp	GF,GSu	GF,GF

Рис. КЛ3.2.1в: події $(G\cdot, G\cdot) \cap GW$ (помаранчевий) та GW (синій і помаранчевий разом)

- Шукана умовна ймовірність — відношення помаранчевих комірок до зафарбованих

План лекції

Теорема Беєса

Теорема (Теорема Беєса (Bayes' Theorem), КЛ3.3.1)

- Нехай маємо вимірний простір $(\Omega, \mathcal{A})$
- Для подій $A, B \in \mathcal{A}$, $\mathbb{P}(B) > 0$, справедливо:

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A)\,\mathbb{P}(A)}{\mathbb{P}(B)} \tag{КЛ3.3.1}$$

- Це безпосередньо випливає з визначення умовної ймовірности

Закон повної ймовірности

Теорема (Закон повної ймовірности (Law of total probability), КЛ3.3.2)

- Нехай маємо вимірний простір $(\Omega, \mathcal{A})$
- Розгляньмо розбиття: $\Omega = \bigcup_{i=1}^{n} A_i$, $A_i \cap A_j = \emptyset$, $i \neq j$
- Нехай $\mathbb{P}(A_i) > 0$ для всіх i
- Тоді

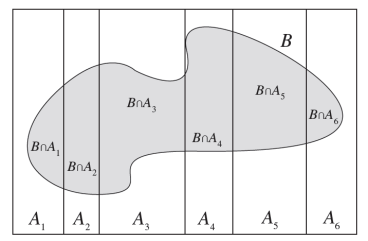
$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(B \mid A_i)\, \mathbb{P}(A_i) \tag{КЛ3.3.2}$$

Доведення.

- Оскільки $A_1, \ldots, A_n$ — розбиття, маємо

$$B = (B \cap A_1) \cup \ldots \cup (B \cap A_n)$$

- До того ж $(B \cap A_i) \cap (B \cap A_j) = \emptyset$, $i \neq j$



- Для неперетинних множин можемо застосувати **адитивність** імовірности:

$$\mathbb{P}(B) = \mathbb{P}(B \cap A_1) + \ldots + \mathbb{P}(B \cap A_n)$$

Закон повної ймовірности

Доведення.

Продовження...

- Кожний перетин розпишімо через умовну ймовірність:

$$\mathbb{P}(B) = \mathbb{P}(B \mid A_1)\,\mathbb{P}(A_1) + \ldots + \mathbb{P}(B \mid A_n)\,\mathbb{P}(A_n)$$

□

Зауваження (КЛ3.3.3)

- Безумовна ймовірність $\mathbb{P}(B)$ — це фактично **зважене середнє** умовних імовірностей
- Вагами виступають імовірності обумовлюючих подій

□

- Правило Беєса зручно застосовувати у випадках, коли обчислення умовної ймовірности $\mathbb{P}(A \mid B)$ простіше звести до обчислення $\mathbb{P}(B \mid A)$ і навпаки

Помилка обвинувачення

Приклад (КЛ3.3.4)

- В 1998 р. Селлі Кларк (Sally Clark) обвинуватили в умисному вбивстві двох новонароджених синів
- Свідок-експерт заявив, що ймовірність смерти немовляти від **синдрому раптової дитячої смерти** складає 1/8500
- Смерть обох дітей одночасно взагалі складає $(1/8500)^2$, або 1 на 73 мільйони випадків
- Відтак, **стверджував експерт**, Селлі невинувата з імовірністю 1 на 73 мільйони
- **Що тут не так?**

□

Приклад (КЛ3.3.4)

Продовження...

- Експерт **переплутав** дві умовні ймовірності
- Нехай I =«обвинувачений невинуватий»
- Нехай E =«дитина загинула внаслідок синдрому»
- Експерт зазначив, що ймовірність $\mathbb{P}(E \mid I)$ дуже мала
- Але нас цікавить **інша ймовірність**:

$$\mathbb{P}(I \mid E) = \frac{\mathbb{P}(E \mid I)\,\mathbb{P}(I)}{\mathbb{P}(E)} = \frac{\mathbb{P}(E \mid I)\,\mathbb{P}(I)}{\mathbb{P}(E \mid I)\,\mathbb{P}(I) + \mathbb{P}(E \mid I^c)\,\mathbb{P}(I^c)}$$

- Тобто ймовірність невинуватости **після того, як смерть уже сталася**
- Апріорна ймовірність $\mathbb{P}(I^c)$ надзвичайно мала
  - Осіб, які вчинили подвійне вбивство, у суспільстві не так уже й багато!
- А тому весь дріб дуже близький до 1
- Селлі Кларк випустили через 3 роки перебування у в'язниці

□

Приклад (КЛ3.3.5)

- Нехай пацієнта тестують на наявність дуже рідкого захворювання
- На нього страждає $100p_1$ відсотків населення
- Якщо результат тесту позитивний, то вважають, що пацієнт хворий
- Нехай $D =$«пацієнт хворий», $T =$«результат тесту позитивний»
- Тест має такі параметри:
  - *Чутливість*: $\mathbb{P}(T \mid D) = 1 - p_2$, тобто ймовірність **хибнонегативного результату** є p_2
  - *Специфічність*: $\mathbb{P}(T^c \mid D^c) = 1 - p_3$, тобто ймовірність **хибнопозитивного результату** є p_3
- Чому дорівнює ймовірність, що пацієнт **насправді** хворий?

$\square$

Приклад (КЛ3.3.5)

Продовження...

- Згідно з правилом Беєса та законом повної ймовірности:

$$\begin{aligned}\mathbb{P}(D \mid T) &= \frac{\mathbb{P}(T \mid D)\,\mathbb{P}(D)}{\mathbb{P}(T)} \\ &= \frac{\mathbb{P}(T \mid D)\,\mathbb{P}(D)}{\mathbb{P}(T \mid D)\,\mathbb{P}(D) + \mathbb{P}(T \mid D^c)\,\mathbb{P}(D^c)} \\ &= \frac{(1-p_2)\cdot p_1}{(1-p_2)\cdot p_1 + p_3\cdot(1-p_1)}\end{aligned}$$

- Наприклад, якщо $p_1 = 0.01$, $p_2 = p_3 = 0.05$, матимемо $\mathbb{P}(D \mid T) \approx 0.16$
- **Тобто тест не помиляється з імовірністю 95%, але пацієнт хворий з імовірністю 16%!**

□

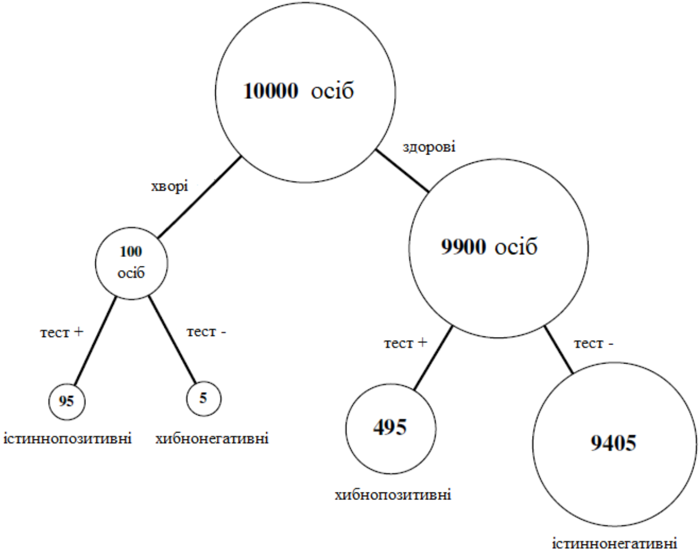

10000 осіб
хворі
здорові
100 осіб
9900 осіб
тест +
тест -
тест +
тест -
95
5
495
9405
істиннопозитивні
хибнонегативні
хибнопозитивні
істиннонегативні

Рис. КЛ3.3.1

Приклад (КЛ3.3.5)

Продовження...

- Нічого дивного тут немає, адже треба враховувати **апріорну** ймовірність (яка дуже низька)
- Хворі 100 осіб із 10 000
- Серед позитивних результатів:
  - 95 істиннопозитивних
  - 495 хибнопозитивних
- Відповідно, **більшість** осіб із позитивним результатом тесту насправді **цілком здорові**

□

Обумовлення декількома подіями

Твердження (КЛ3.3.6)

- Нехай маємо вимірний простір $(\Omega, \mathcal{A})$
- Нехай $A, B, C \in \mathcal{A}$, $\mathbb{P}(B \cap C) > 0$
- Тоді

$$\mathbb{P}(A \mid B, C) = \frac{\mathbb{P}(B \mid A, C)\,\mathbb{P}(A \mid C)}{\mathbb{P}(B \mid C)} \qquad \text{(КЛ3.3.3)}$$

Твердження (КЛ3.3.7)

- Нехай маємо вимірний простір $(\Omega, \mathcal{A})$
- Нехай $A_1, \ldots, A_n$ — деяке його розбиття
- Нехай $\mathbb{P}(A_i \cap B) > 0$ для всіх i
- Тоді

$$\mathbb{P}(B \mid C) = \sum_{i=1}^{n} \mathbb{P}(B \mid A_i, C)\,\mathbb{P}(A_i \mid C) \qquad \text{(КЛ3.3.4)}$$

Приклад (КЛ3.3.8)

- Розгляньмо **повторний** тест після позитивного результату
- Нехай T_1 і T_2 — позитивні результати за 1-им і 2-им разом
- Маємо:

$$\mathbb{P}(D \mid T_1 \cap T_2) = \frac{\mathbb{P}(T_1 \cap T_2 \mid D)\,\mathbb{P}(D)}{\mathbb{P}(T_1 \cap T_2 \mid D)\,\mathbb{P}(D) + \mathbb{P}(T_1 \cap T_2 \mid D^c)\,\mathbb{P}(D^c)}$$
$$= \frac{0.95^2 \cdot 0.01}{0.95^2 \cdot 0.01 + 0.05^2 \cdot 0.99} \approx 0.78$$

- Цей же результат можна дістати **послідовно**:
  - Ми вже знаємо $\mathbb{P}(D \mid T_1) = 0.16$
  - Тому

$$\mathbb{P}(D \mid T_1, T_2) = \frac{\mathbb{P}(T_2 \mid D, T_1)\,\mathbb{P}(D \mid T_1)}{\mathbb{P}(T_2 \mid D, T_1)\,\mathbb{P}(D \mid T_1) + \mathbb{P}(T_2 \mid D^c, T_1)\,\mathbb{P}(D^c \mid T_1)}$$
$$= \frac{0.95 \cdot 0.16}{0.95 \cdot 0.16 + 0.05 \cdot 0.84} \approx 0.78$$

- Повторний тест збільшує ймовірність того, що пацієнт справді хворий, із 0.16 до 0.78!

Зауваження (КЛ3.3.9)

Непринципово, які саме події вважати обумовлюючими:

$$\mathbb{P}(A \mid B, C) = \frac{\mathbb{P}(C \mid A, B)\,\mathbb{P}(A \mid B)}{\mathbb{P}(C \mid B)} = \frac{\mathbb{P}(A \cap B \cap C)}{\mathbb{P}(B \cap C)}$$

□

Приклад (КЛ3.3.10)

Табл. КЛ3.3.1

	Метод A			Метод B		
	Операцій	Успішних	Відсоток	Операцій	Успішних	Відсоток
Малі	87	81	93%	270	234	87%
Великі	263	192	73%	80	55	69%
Разом	350	273	78%	350	289	83%

- Розгляньмо результати лікування малих та великих каменів у нирках двома різними методами
- Метод A для кожної категорії **ліпший**
- Проте **загалом** метод A **гірший**
- Як таке може бути?

□

Приклад (КЛ3.3.10)

Продовження...

- Нехай S =«лікування успішне», A =«застосовано метод А», B =«застосовано метод В» (очевидно, що $A = B^c$), C =«лікували малі камені»
- Тоді

$$\mathbb{P}(S \mid A) = \mathbb{P}(S \mid C, A)\,\mathbb{P}(C \mid A) + \mathbb{P}(S \mid C^c, A)\,\mathbb{P}(C^c \mid A)$$
$$\mathbb{P}(S \mid B) = \mathbb{P}(S \mid C, B)\,\mathbb{P}(C \mid B) + \mathbb{P}(S \mid C^c, B)\,\mathbb{P}(C^c \mid B)$$

- Імовірності успішного лікування $\mathbb{P}(S \mid A)$ і $\mathbb{P}(S \mid B)$ є **зваженими середніми** ймовірностей успішного лікування **каменів певного типу**
- Ваговими коефіцієнтами є $\mathbb{P}(C \mid A)$, $\mathbb{P}(C^c \mid A)$, $\mathbb{P}(C \mid B)$ і $\mathbb{P}(C^c \mid B)$ — **частоти** відповідних операцій

$\square$

Парадокс Симпсона

Приклад (КЛ3.3.10)

Продовження...

- Якби вагові коефіцієнти були однакові, парадоксу не було б
- Але

$$\mathbb{P}(C \mid B) > \mathbb{P}(C \mid A)\,, \qquad \mathbb{P}(C^c \mid B) < \mathbb{P}(C^c \mid A)$$

- На наших даних

$$\mathbb{P}(S \mid B) = 0.87 \cdot \frac{270}{350} + 0.69 \cdot \frac{80}{350} \approx 0.83$$
$$\mathbb{P}(S \mid B^c) = 0.93 \cdot \frac{87}{350} + 0.73 \cdot \frac{263}{350} \approx 0.78$$

- У загальному випадку, проблема є, коли присутній завадний чинник (confounder), як тип каменю у нашому прикладі
- Інший відомий приклад:
  - Вступ в аспірантуру Університету Каліфорнії в Берклі (1973 р.)
  - Набрали більше чоловіків, ніж жінок ⇒ звинувачення в дискримінації
  - Проте насправді в **більшості** факультетах набрали більше **жінок**
  - Просто жінки подавались на **факультети з вищим конкурсом**

□