Въведение във вероятностите

Фундаментални понятия, закони и примери

Младен Савов, представено от Демир Тончев

Виртуална Гюлечица, Юли 2021

Вероятностно пространство

Възможните изходи от случаен експеримент означаваме с Ω.

Възможните изходи от случаен експеримент означаваме с Ω.

1. $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$

2. $\Omega = \{\omega_1, \omega_2, \ldots\}$

Вероятностно пространство множество от елементарни събития

Възможните изходи от случаен експеримент означаваме с Ω.

1. $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$
   - $N = 2$, $\omega_1 =$'ези', $\omega_2 =$'тура', $\Omega = \{\omega_1, \omega_2\}$;
   - $N = 2^{100}$, $\Omega = \{\omega_1, \omega_2, \ldots, \omega_{2^{100}}\}$, ω_i са възможните конфигурации 'ези-тура' на 100 различими монети;
   - $N = 13983816$, $\Omega = \{\omega_1, \omega_2, \ldots, \omega_{13983816}\}$.
2. $\Omega = \{\omega_1, \omega_2, \ldots\}$

Възможните изходи от случаен експеримент означаваме с Ω.

1. $\Omega = \{\omega_1, \omega_2, \dots, \omega_N\}$
   - $N = 2$, $\omega_1 =$ 'ези', $\omega_2 =$ 'тура', $\Omega = \{\omega_1, \omega_2\}$;
   - $N = 2^{100}$, $\Omega = \{\omega_1, \omega_2, \dots, \omega_{2^{100}}\}$, ω_i са възможните конфигурации 'ези-тура' на 100 различими монети;
   - $N = 13983816$, $\Omega = \{\omega_1, \omega_2, \dots, \omega_{13983816}\}$.
2. $\Omega = \{\omega_1, \omega_2, \dots\}$

Възможните изходи от случаен експеримент означаваме с Ω.

1. $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$
   - $N = 2$, $\omega_1 =$'ези', $\omega_2 =$'тура', $\Omega = \{\omega_1, \omega_2\}$;
   - $N = 2^{100}$, $\Omega = \{\omega_1, \omega_2, \ldots, \omega_{2^{100}}\}$, ω_i са възможните конфигурации 'ези-тура' на 100 различими монети;
   - $N = 13983816$, $\Omega = \{\omega_1, \omega_2, \ldots, \omega_{13983816}\}$.
2. $\Omega = \{\omega_1, \omega_2, \ldots\}$
   - $\omega_j =$'успех е постигнат на j-тия опит'.

Подмножество $A \subseteq \Omega$ се нарича събитие.

A обединява по признак елементарни събития $\omega \in \Omega$.

Подмножество $A \subseteq \Omega$ се нарича събитие.

A обединява по признак елементарни събития $\omega \in \Omega$.

1. $\Omega = \{1, 2, 3, 4, 5, 6\}$ и събитието 'пада се нечетно число' $A = \{1, 3, 5\}$;
2. $\Omega =$ {гласоподаватели} и събитието A='гласували за партия X';

Подмножество $A \subseteq \Omega$ се нарича събитие.

A обединява по признак елементарни събития $\omega \in \Omega$.

1. $\Omega = \{1, 2, 3, 4, 5, 6\}$ и събитието 'пада се нечетно число' $A = \{1, 3, 5\}$;
2. $\Omega = \{$гласоподаватели$\}$ и събитието A='гласували за партия X';

Ако $A \subseteq \Omega, B \subseteq \Omega$, то

- $A \cap B = \{\omega \in \Omega : \omega \in A \text{ и } \omega \in B\}$;
- $A \cup B = \{\omega \in \Omega : \omega \in A \text{ или } \omega \in B\}$;
- $A^c = \Omega \setminus A = \{\omega \in \Omega : \omega \notin A\}$.

$\Omega = \{$гласоподаватели$\}$, A='гласували за партия X' и B='гласоподаватели с образование от чужбина'

- $A \cap B$ ='гласоподаватели на X с образование от чужбина';
- $A \cup B$ ='гласували за X или гласоподаватели с образование от чужбина';
- A^c ='гласували за партия, различна от X'.

Колекция от събития $\mathcal{F} = \{A : A \subseteq \Omega\}$ наричаме сигма алгебра, ако:

1. $\Omega, \emptyset \in \mathcal{F}$;
2. $A, B \in \mathcal{F} \implies A \cup B \in \mathcal{F}$
   $A_1, A_2, \dots \in \mathcal{F} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$;
3. $A, B \in \mathcal{F} \implies A \cap B \in \mathcal{F}$
   $A_1, A_2, \dots \in \mathcal{F} \implies \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$;
4. $A \in \mathcal{F} \implies A^c \in \mathcal{F}$.

Колекция от събития $\mathcal{F} = \{A : A \subseteq \Omega\}$ наричаме сигма алгебра, ако:

1. $\Omega, \emptyset \in \mathcal{F}$;
2. $A, B \in \mathcal{F} \implies A \cup B \in \mathcal{F}$
   $A_1, A_2, \dots \in \mathcal{F} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$;
3. $A, B \in \mathcal{F} \implies A \cap B \in \mathcal{F}$
   $A_1, A_2, \dots \in \mathcal{F} \implies \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$;
4. $A \in \mathcal{F} \implies A^c \in \mathcal{F}$.

Колекция от събития $\mathcal{F} = \{A : A \subseteq \Omega\}$ наричаме сигма алгебра, ако:

1. $\Omega, \emptyset \in \mathcal{F}$;
2. $A, B \in \mathcal{F} \implies A \cup B \in \mathcal{F}$
   $A_1, A_2, \dots \in \mathcal{F} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$;
3. $A, B \in \mathcal{F} \implies A \cap B \in \mathcal{F}$
   $A_1, A_2, \dots \in \mathcal{F} \implies \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$;
4. $A \in \mathcal{F} \implies A^c \in \mathcal{F}$.

Колекция от събития $\mathcal{F} = \{A : A \subseteq \Omega\}$ наричаме сигма алгебра, ако:

1. $\Omega, \emptyset \in \mathcal{F}$;
2. $A, B \in \mathcal{F} \implies A \cup B \in \mathcal{F}$
   $A_1, A_2, \dots \in \mathcal{F} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$;
3. $A, B \in \mathcal{F} \implies A \cap B \in \mathcal{F}$
   $A_1, A_2, \dots \in \mathcal{F} \implies \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$;
4. $A \in \mathcal{F} \implies A^c \in \mathcal{F}$.

Колекция от събития $\mathcal{F} = \{\mathrm{A} : \mathrm{A} \subseteq \Omega\}$ наричаме сигма алгебра, ако:

1. $\Omega, \emptyset \in \mathcal{F}$;
2. $\mathrm{A}, \mathrm{B} \in \mathcal{F} \implies \mathrm{A} \cup \mathrm{B} \in \mathcal{F}$
   $\mathrm{A}_1, \mathrm{A}_2, \dots \in \mathcal{F} \implies \bigcup_{\mathrm{i}=1}^{\infty} \mathrm{A}_\mathrm{i} \in \mathcal{F}$;
3. $\mathrm{A}, \mathrm{B} \in \mathcal{F} \implies \mathrm{A} \cap \mathrm{B} \in \mathcal{F}$
   $\mathrm{A}_1, \mathrm{A}_2, \dots \in \mathcal{F} \implies \bigcap_{\mathrm{i}=1}^{\infty} \mathrm{A}_\mathrm{i} \in \mathcal{F}$;
4. $\mathrm{A} \in \mathcal{F} \implies \mathrm{A}^\mathrm{c} \in \mathcal{F}$.

$$\Omega = \{\omega_1, \omega_2\}, \mathcal{F} = \{\emptyset, \Omega, \{\omega_1\}, \{\omega_2\}\}.$$

$\Omega = \{\omega_1, \omega_2\}, \mathcal{F} = \{\emptyset, \Omega, \{\omega_1\}, \{\omega_2\}\}.$

$\Omega = \{1, 2, 3, 4, 5, 6\}, \mathcal{F} = 2^{\Omega}$ множеството на всички подмножества на Ω.

При зададени $\Omega, \mathcal{F}$, вероятността е функция $\mathbb{P} : \mathcal{F} \mapsto [0, 1]$ със свойствата:

- $\mathbb{P}(\Omega) = 1$;
- $\mathbb{P}(A^c) = 1 - \mathbb{P}(A), A \in \mathcal{F}$;
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B), A \cap B = \emptyset$
  $A_1, A_2, \cdots$ непресичащи се $\implies \mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$.

Вероятностно пространство

Тройката

$$(\Omega, \mathcal{F}, \mathbb{P})$$

задава вероятностно пространство.

$\Omega = \{\omega_1, \omega_2, \cdots, \omega_{\mathrm{N}}\}$, то $\left(\Omega, 2^{\Omega}, \mathbb{P}\right)$ е снабдено с равномерна вероятност, ако:

$$\mathbb{P}(\omega) = \frac{1}{\mathrm{N}}, \text{ за всяко } \omega \in \Omega.$$

$\mathrm{N} = 13983816$, то Ω са всички шесторки в играта 6 от 49 и

$$\mathbb{P}(\omega) = \frac{1}{13983816}.$$

Две кози и една кола са скрити зад три затворени врати.

Избирате една врата, без да я отваряте, обаче.

След избора се отваря една от останалите врати и виждаме една от козите.

Две кози и една кола са скрити зад три затворени врати.

Избирате една врата, без да я отваряте, обаче.

След избора се отваря една от останалите врати и виждаме една от козите.

Бихте ли сменили избора си, ако можете?

Условна вероятност и независимост

Ако се е сбъднало $A \subseteq \Omega$, то това променя вероятностното пространство:

- $\Omega \to A$;
- $\mathcal{F} \to \mathcal{F} \cap A = \{B \in \mathcal{F}, A \cap B\}$;
- $\mathbb{P} \to \mathbb{P}_A$, така че:

$$\mathbb{P}_A(B) = \mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}.$$

Ако $\Omega = H_1 \cup H_2 \cdots H_n$ и $H_j, j \leq n$ са непресичащи се, то

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}(A \cap \Omega) \\ &= \mathbb{P}(A \cap H_1 \cup H_2 \cdots \cup H_n) = \sum_{i=1}^{n} \mathbb{P}(A \cap H_i) \\ &= \sum_{i=1}^{n} \mathbb{P}(H_i)\, \mathbb{P}(A|H_i)\,.\end{aligned}$$

Две събития A, B са независими, тогава и само тогава когато

$$\mathbb{P}(\mathrm{A} \cap \mathrm{B}) = \mathbb{P}(\mathrm{A})\,\mathbb{P}(\mathrm{B}) \implies \mathbb{P}_{\mathrm{B}}(\mathrm{A}) = \mathbb{P}(\mathrm{A}|\mathrm{B}) = \mathbb{P}(\mathrm{A}).$$

В6 от 49 са изтеглени 10001 тиража.

Каква е вероятността поне веднъж в два последователни тиража да са изтеглени едни и същи числа?

В6 от 49 са изтеглени 10001 тиража.

Каква е вероятността поне веднъж в два последователни тиража да са изтеглени едни и същи числа?

$$\mathrm{A} = \bigcup_{\mathrm{i}=1}^{10000} \{\text{числа от i-то теглене съвпадат с тези от } \mathrm{i}+1\}$$

В6 от 49 са изтеглени 10001 тиража.

Каква е вероятността поне веднъж в два последователни тиража да са изтеглени едни и същи числа?

$$\mathrm{A} = \bigcup_{\mathrm{i}=1}^{10000} \{\text{числа от i-то теглене съвпадат с тези от } \mathrm{i}+1\}$$

$$\begin{aligned}\mathbb{P}(\mathrm{A}) &= 1 - \mathbb{P}\left(\bigcap_{\mathrm{i}=1}^{10000} \{\text{числа на i-то теглене не съвпадат с тези от } \mathrm{i}+1\}\right) \\ &= 1 - \mathbb{P}\left(\text{числа на 1-то теглене не съвпадат с тези от } 2\right)^{10000} \\ &= 1 - \left(\frac{13983815}{13983816}\right)^{10000} \sim \frac{7}{1000}.\end{aligned}$$

В стая има n човека, нека:

$$\mathrm{A} = \{\text{поне двама човека имат една и съща дата за рожден ден}\}$$

колко човека($\mathrm{n} =?$) трябва да има в една стая така, че $\mathbb{P}(\mathrm{A}) = 0.5$
приемете, че в годината има 365 дни

hint:

$$\mathrm{A}^c = \{\text{няма хора с еднакви дати}\}$$

След като сме наблюдавали събитието $\mathbb{P}(A)$ какво може да кажем за събитието $\mathbb{P}(B)$

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

Пример:
Нека от някаква болест са болни 20% от населението. Имаме тест, който хваща 95% от болните правилно(дава положителен резултат при положение, че пациента е болен) и греши (дава положителен резултат) при 10% при положение, че пациента е здрав. Идва нов човек, каква е вероятността да е болен ако теста е положителен?

Формула на Бейс - продължение

$$\mathbb{P}(B) = 0.2$$
$$\mathbb{P}(A|B) = 0.95$$
$$\mathbb{P}(A|B^c) = 0.10$$
$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B)\,\mathbb{P}(B)}{\mathbb{P}(A)}$$

Формула на Бейс - продължение

$$
\begin{aligned}
\mathbb{P}(B) &= 0.2\\
\mathbb{P}(A|B) &= 0.95\\
\mathbb{P}(A|B^c) &= 0.10\\
\mathbb{P}(B|A) &= \frac{\mathbb{P}(A|B)\,\mathbb{P}(B)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B)\,\mathbb{P}(B)}{\mathbb{P}(A|B)\,\mathbb{P}(B) + \mathbb{P}(A|B^c)\,\mathbb{P}(B^c)}
\end{aligned}
$$

Случайни величини

Нека $(\Omega, \mathcal{F}, \mathbb{P})$ е вероятностно пространство.

Тогава всяко 'хубаво' изображение

$$\mathrm{X} : \Omega \mapsto (-\infty, \infty)$$

се нарича случайна величина.

Предимства

- числово описват характеристики на елементарни събития;
- позволяват математически операции;
- по-лесна компютърна обработка;
- позволяват моделиране на различни случайни събития с еднаква вероятностна структура;
- позволяват числови характеристики като средно и дисперсия.

$X : \Omega \mapsto (-\infty, \infty)$ приема краен или изброим брой стойности.

- $X \in \{0, 1\}$, като $\mathbb{P}(X = 0) = 1 - p$, $\mathbb{P}(X = 1) = p$;
- $X \in \{0, 1, \cdots, n\}$, като
$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k};$$
- $X \in \{0, 1, 2, \cdots\}$, като
$$\mathbb{P}(X = k) = (1 - p)^k p.$$

$X : \Omega \mapsto (-\infty, \infty)$ приема <u>краен или изброим</u> брой стойности.

- $X \in \{0, 1\}$, като $\mathbb{P}(X = 0) = 1 - p$, $\mathbb{P}(X = 1) = p$;
- $X \in \{0, 1, \cdots, n\}$, като
$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k};$$
- $X \in \{0, 1, 2, \cdots\}$, като
$$\mathbb{P}(X = k) = (1 - p)^k p.$$

$X : \Omega \mapsto (-\infty, \infty)$ приема краен или изброим брой стойности.

- $X \in \{0, 1\}$, като $\mathbb{P}(X = 0) = 1 - p$, $\mathbb{P}(X = 1) = p$;
- $X \in \{0, 1, \cdots, n\}$, като
$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k};$$
- $X \in \{0, 1, 2, \cdots\}$, като
$$\mathbb{P}(X = k) = (1 - p)^k p.$$

$X : \Omega \mapsto (-\infty, \infty)$ приема краен или изброим брой стойности.

- $X \in \{0, 1\}$, като $\mathbb{P}(X = 0) = 1 - p$, $\mathbb{P}(X = 1) = p$;
- $X \in \{0, 1, \cdots, n\}$, като

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k};$$

- $X \in \{0, 1, 2, \cdots\}$, като

$$\mathbb{P}(X = k) = (1 - p)^k p.$$

$X : \Omega \mapsto (-\infty, \infty)$ приема <u>неизброимо</u> много стойности.

Вероятностите се изчисляват с помощта на вероятностната плътност или разпределението се задава

$$\mathbb{P}\left(a < X < b\right) = \int_a^b f_X(y)dy.$$

Примерите включват нормална, експоненциална, Бета случайни величини.

Две случайни величини X, Y са независими ($X \perp Y$), ако за всеки $a < b, c < d$

$$\mathbb{P}\left(a < X < b \cap c < Y < d\right) = \mathbb{P}\left(a < X < b\right)\mathbb{P}\left(c < Y < d\right).$$

Две случайни величини X, Y са еднакви по разпределение ($X \stackrel{d}{=} Y$), ако за всеки $a < b$

$$\mathbb{P}\left(a < X < b\right) = \mathbb{P}\left(a < Y < b\right).$$

Математическо очакване и дисперсия

Ако X е случайна величина, то математическото очакване се задава чрез:

- $\mathbb{E}[X] = \sum x\mathbb{P}(X = x)$, ако X е дискретна,
- $\mathbb{E}[X] = \int_{-\infty}^{\infty} xf_X(x)dx$, ако X е непрекъсната.

Дисперсията като грешка на типичната стойност

Математическото очакване $\mathbb{E}\left[X\right]$ е оценка за "типичната"стойност, но колко е добра? Възможна мярка е дисперсията:

$$\mathrm{Var}(X) = \mathbb{E}\left(X - \mathbb{E}\left[X\right]\right)^2$$

Математическото очакване $\mathbb{E}[X]$ е оценка за "типичната"стойност, но колко е добра? Възможна мярка е дисперсията:

$$\mathrm{Var}(X) = \mathbb{E}\left(X - \mathbb{E}[X]\right)^2$$

Математическото очакване е оценка за типичната стойност на случайната величина погледната през призмата на най-малките квадрати.

Закон за големите числа и Централна гранична теорема

Нека $(X_i)_{i\geq 1}$ са независими, еднакво разпределени случайни величини. Означаваме

$$S_n = \sum_{i=1}^{n} X_i.$$

Нека $(X_i)_{i \geq 1}$ са независими, еднакво разпределени случайни величини. Означаваме

$$S_n = \sum_{i=1}^{n} X_i.$$

Тогава

$$\lim_{n \to \infty} \frac{S_n}{n} = \mathbb{E}\left[X_1\right].$$

Сред огромна популация мерим наличността на свойство например 'гласоподавател за X'. Как ще оценим пропорцията p от популацията с това свойство?

Сред огромна популация мерим наличността на свойство например 'гласоподаветел за X'. Как ще оценим пропорцията p от популацията с това свойство?

Случайно задаваме номер на членовете на популацията и тогава:

- $X_i = 1$ ако $i-$тия член притежава свойството и $\mathbb{P}(X_i = 1) = p$;
- $X_i = 0$ ако $i-$тия член не притежава свойството и $\mathbb{P}(X_i = 0) = 1 - p$.

Сред огромна популация мерим наличността на свойство например 'гласоподаветел за X'. Как ще оценим пропорцията p от популацията с това свойство?

Случайно задаваме номер на членовете на популацията и тогава:

- $X_i = 1$ ако $i-$тия член притежава свойството и $\mathbb{P}(X_i = 1) = p$;
- $X_i = 0$ ако $i-$тия член не притежава свойството и $\mathbb{P}(X_i = 0) = 1 - p$.

Ако разгледаме първите n члена, то

$$\frac{\sum_{i=1}^{n} X_i}{n} = \frac{S_n}{n} \overset{?}{\sim} p = \mathbb{E}[X_1].$$

Разтърсваме и отваряме кутия с N монети. Преброяваме монетите показващи ези. Какво можем да кажем за техния брой?

Разтърсваме и отваряме кутия с N монети. Преброяваме монетите показващи ези. Какво можем да кажем за техния брой?

Нека $X_i = 1$ ако $i-$тата монета е ези и $\mathbb{P}\left(X_i = 1\right) = 1/2$, иначе $X_i = 0$. Тогава

$$\frac{\sum_{i=1}^{N} X_i}{N} = \frac{S_N}{N} \sim \frac{1}{2}.$$

Централна гранична теорема (ЦГТ) като грешка в закона за големите числа

ЦГТ е грешката в закона за големите числа и е основен резултат и инструмент в модерната наука.

ЦГТ е грешката в закона за големите числа и е основен резултат и инструмент в модерната наука.

Отново $S_n = \sum_{i=1}^{n} X_i$, където $(X_i)_{i \geq 1}$ независими и еднакво разпределени. Тогава

$$\frac{\frac{S_n}{n} - \mathbb{E}\left[X_1\right]}{\sqrt{\frac{\mathrm{Var}(X_1)}{n}}} \Rightarrow \xi \sim N(0,1).$$

ЦГТ е грешката в закона за големите числа и е основен резултат и инструмент в модерната наука.

Отново $S_n = \sum_{i=1}^{n} X_i$, където $(X_i)_{i\geq 1}$ независими и еднакво разпределени. Тогава

$$\frac{\frac{S_n}{n} - \mathbb{E}[X_1]}{\sqrt{\frac{Var(X_1)}{n}}} \Rightarrow \xi \sim N(0,1).$$

Следва, че за големи n

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}[X_1]\right| > a\sqrt{\frac{Var(X_1)}{n}}\right) \sim \frac{1}{\sqrt{2\pi}} \int_{|x|>a} e^{-\frac{x^2}{2}} dx.$$

БЛАГОДАРЯ!