Л2 12.09.24

Условные вероятности и независимость событий

Если о системе известна некоторая доп. информация, то вероятность необходимо подправить. Например, может быть известно, что у кубика чёрные грани с чётными числами.

Пусть $(\Omega, \mathcal{F}, \mathbb{P})$ — вероят. пр-во; пусть произошло событие A, тогда, зная эту информацию, необходимо пересчитать вероят. других событий, $\mathbb{P}(A) > 0$.

def | Условная вероятность события B при условии события A:

$$\mathbb{P}(B|A) \doteq \frac{\mathbb{P}(A\cdot B)}{\mathbb{P}(A)}$$

$\Rightarrow \mathbb{P}(AB) = \mathbb{P}(A)\cdot \mathbb{P}(B|A)$ — иногда удобнее знать усл. вероят., а далее искать вероят. совместного наступления AB.

Вводится $\mathbb{P}_A(B) \doteq \mathbb{P}(B|A) : \mathcal{F} \longrightarrow \mathbb{R}_+$ — условная вероят., обладает всеми св-вами обычной вероятности $\mathbb{P}$, но на $(\Omega, \mathcal{F}, \mathbb{P}_A)$.

Th1 [Теорема умножения]

$A_1, \dots, A_n$ — события. Тогда:

$$\mathbb{P}(A_1 \cdot\ldots\cdot A_n) = \mathbb{P}(A_1)\cdot \mathbb{P}(A_2|A_1)\cdot \mathbb{P}(A_3|A_1 A_2)\cdot\ldots\cdot \mathbb{P}(A_n|A_1\cdot\ldots\cdot A_{n-1}),$$ если $\mathbb{P}(A_1\cdot\ldots\cdot A_n) > 0$.

Замечание | Из условия следует, что $\mathbb{P}(A_1\cdot\ldots\cdot A_k) > 0$, $k=\overline{1,n}$ в силу свойств вероят. $\mathbb{P}$, поэтому все усл. вероят. опр. корректно.

Можно заменить на $\bigcup\limits_{i=1}^{n} H_i = \Omega$, а $\sum$ даёт ещё соб. $H_0 : \mathbb{P}(H_0) = 0$.

Th2 [Формула полной вероятности]

$H_1, \dots, H_n$ — события: $H_i H_j = \varnothing,\ i\neq j,\ \mathbb{P}(H_i) \neq 0,\ \sum\limits_{i=1}^{n} \mathbb{P}(H_i) = 1$. Тогда:

$$\forall A \in \mathcal{F} \mapsto \mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(H_i)\cdot \mathbb{P}(A|H_i) \quad // \quad \mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(AH_i)$$

$\mathbb{P}(H_i)$ — априорная вероят. гипотез; $\mathbb{P}(A|H_i)$ — апостериорная вероят. события A.

$A = A\Omega = A\cdot \bigcup\limits_{i=0}^{n} H_i$, $\mathbb{P}(AH_0) = 0$ т.к. $\mathbb{P}(H_0) = 0$

События H_i называются гипотезами. Смысл формулы такой: высказываем разные гипотезы и смотрим на апост. вероят. $\mathbb{P}(A|H_i)$.

Совокупность $H_1, \ldots, H_n$ из условия называется полной группой событий.

$\boxed{\text{Th3}}$ [Формула Байеса]

$H_1, \ldots, H_n : H_i H_j = \varnothing,\ i \neq j,\ \mathbb{P}(H_i) > 0,\ \sum_{i=1}^{n} \mathbb{P}(H_i) = 1$ и пусть событие $A \in \mathcal{F}: \mathbb{P}(A) > 0$.

Тогда:
$$\mathbb{P}(H_k | A) = \frac{\mathbb{P}(H_k) \cdot \mathbb{P}(A|H_k)}{\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(H_i) \cdot \mathbb{P}(A|H_i)}$$

Из ф-лы Байеса растёт большое количество разных теорий: байесовское семплир., байесовские статистики и т.д. Пока что можем при помощи ф. Байеса определить какая из гипотез H_k наиболее вероятная.

Независимость событий

Пусть есть два события A и B. Если знаем, что $AB = \varnothing$ и что произошло A, то точно знаем, что B не произошло. Но нельзя путать независимость и несовместность.

Пусть определили независимость A и B как $\mathbb{P}(B|A) = \mathbb{P}(B)$ $\Rightarrow$ из того что произошло A не получаем информации о событии B. Однако, такая формулировка независимости ущербна т.к. требует $\mathbb{P}(A) > 0$ и несимметрична относ. A и B.

Если $\mathbb{P}(B) > 0$, тогда $\mathbb{P}(A|B) = \dfrac{\mathbb{P}(AB)}{\mathbb{P}(B)} = \dfrac{\mathbb{P}(A) \cdot \cancel{\mathbb{P}(B|A)}}{\cancel{\mathbb{P}(B)}} = \mathbb{P}(A)$ – получили что хотели

Желаемые р-ва $\mathbb{P}(B|A) = \mathbb{P}(B)$ и $\mathbb{P}(A|B) = \mathbb{P}(A)$ вып., если потребовать
$$\mathbb{P}(A \cdot B) = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

$\boxed{\text{def}}$ События A и B назыв. <u>независимыми</u> (попарно), если $\mathbb{P}(A \cdot B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$.

При таком определении м.б. $\mathbb{P}(A) = 0$ или $\mathbb{P}(B) = 0$ и ничего не ломается.

$\boxed{\text{def}}$ События $A_1, \ldots, A_n$ назыв. <u>независимыми в совокупности</u>, если
$$\mathbb{P}\Big(\bigcap_{i \in I} A_i\Big) = \prod_{i \in I} \mathbb{P}(A_i) \quad \forall I \subseteq \{1, 2, \ldots, n\} \quad (*)$$

$\boxed{\text{Упр1}}$ Следует ли из попарной независ. соб. $A_1, \ldots, A_n$ их независ. в совокупности?

$\boxed{\text{Упр2}}$ В р-ве $(*)$ содержится $\sum_{k=2}^{n} C_n^k = 2^n - n - 1$ равенств.

не учитываем одиночные р-ва и $k=0$.

Следует ли из выполнения самой длинной цепочки в $(*)$ выполнение остальных цепочек?

Рассмотрим теперь конкретные вероятностные схемы.

I) Биномиальная схема Бернулли

Пусть некоторый эксп. проводится n раз (испытаний). В каждом испытании следим за выполнением соб. A („успех"), испытания независимы. Пусть $\mathbb{P}(A)=p$, $0<p<1$. $q \doteq 1-p$ – вероят. „неуспеха".

Задача: Пусть $B_n(k)=\{$В n исп. произошло k успехов$\}$, $k=\overline{0,n}$., $\mathbb{P}(B_n(k))=?$

$B_n(k)$ несовместные по построению. Определим элем. соб. как цепочку

$\omega=(\delta_1,\dots,\delta_n)$, $\delta_i=\begin{cases}1, & \text{если } A\\ 0, & \text{если } \overline{A}\end{cases}$ – ω однозначно определяет результат эксп. из n испытаний, $\Omega=\{\omega\}$.

$\Rightarrow |\Omega|=2^n$, но в отличие от класс. опр. вероят. ω не равновероятны. Т.к. исп. независ., то $\mathbb{P}(\{\omega\})=p^{\sum_{i=1}^{n}\delta_i}\cdot q^{n-\sum_{i=1}^{n}\delta_i}$ – Вопрос: задаёт ли $\mathbb{P}(\{\omega\})$ распр. вероят.?

Пусть $\omega\in B_n(k)$, тогда есть k успехов $\Rightarrow \mathbb{P}(\{\omega\})=p^k\cdot q^{n-k}$, т.е. внутри $B_n(k)$ все элем. события равновероятны и эти элем. события отличаются расположением в них успехов $\Rightarrow \boxed{\mathbb{P}(B_n(k))=\underbrace{C_n^k}_{=|B_n(k)|}\cdot p^k\cdot q^{n-k}}$ $\left(\mathbb{P}(A)=\sum_{\omega\in A}\mathbb{P}(\omega)\right)$

Упр 3 Проверить, что $\mathbb{P}(B_n(k))$ задаёт распр. вероят.

II) Полиномиальная схема

n независ. испытаний, исходы: $A_1,\dots,A_r$ ($r=2$ – бином. схема), $p_k=\mathbb{P}(A_k)$ в отдельном испытании, $p_k>0$, $\sum_{k=1}^{r}p_k=1$.

Обозначим $B_n(k_1,\dots,k_r)$ – события, в которых k_i – число исп., в которых появилось A_i, $k_1+\dots+k_r=n$. Задача: $\mathbb{P}(B_n(k_1,\dots,k_r))=?$

Ω: $\omega=(\delta_1,\dots,\delta_n)$, $\delta_i\in\{\underset{\uparrow A_1}{1},\dots,\underset{\uparrow A_r}{r}\}$

В силу независ. испытаний: $\mathbb{P}(\{\omega\})=p_1^{k_1}\cdot\ldots\cdot p_r^{k_r}$, но такие цепочки не равновероятны для разных $B_n(\dots)$.

(Если $\omega \in B_n(k_1,\dots,k_r)$, то внутри $B_n(k_1,\dots,k_r)$ цепочки равновероятны и тогда определим $|B_n(k_1,\dots,k_r)|$.

Зафиксируем в послед. n исп. те, в которых $A_1 \to C_n^{k_1}$ вариантов и т.д.

$$\Rightarrow |B_n(k_1,\dots,k_r)| = \underset{\uparrow\ \text{расст. } A_1}{C_n^{k_1}} \cdot \underset{\uparrow\ \text{расст. } A_2}{C_{n-k_1}^{k_2}} \cdot \ldots \cdot C_{k_r}^{k_r} =$$

def| Полиномиальный коэффициент.

$$= \frac{n!}{k_1!\cdot\cancel{(n-k_1)!}} \cdot \frac{\cancel{(n-k_1)!}}{k_2!\cdot(n-k_1-k_2)!}\cdot\ldots\cdot 1 = \overbrace{\frac{n!}{k_1!\cdot\ldots\cdot k_r!}}$$

$$\Rightarrow \boxed{\mathbb{P}(B_n(k_1,\dots,k_r)) = \frac{n!}{k_1!\cdot\ldots\cdot k_r!}\cdot p_1^{k_1}\cdot\ldots\cdot p_r^{k_r}.}$$

Рассмотрим два предельных варианта схемы Бернулли.

Th4 [Теорема Пуассона]

Вопрос: имеем ли право в схеме Бернулли $n\to\infty$? Есть ли такое вероят. пр-во?

Пусть в биномиальной схеме: $n\to\infty$, $p\to 0$, $n\cdot p\to\lambda>0$. Тогда:

$$\mathbb{P}(B_n(k)) = C_n^k\cdot p^k\cdot q^{n-k} \xrightarrow[n\to\infty]{} e^{-\lambda}\cdot\frac{\lambda^k}{k!},\ k=0,1,2,\ldots \text{ — „Закон редких событий".}$$

$$\triangleleft\ \mathbb{P}(B_n(k)) = \frac{n!}{k!\cdot(n-k)!}\cdot p^k\cdot \underset{1-p}{\underset{\parallel}{q}}{}^{n-k}\ \Big|\cdot\frac{n^k}{n^k} = \frac{(n\cdot p)^k}{k!}\cdot\frac{n\cdot(n-1)\cdot\ldots\cdot(n-k+1)}{n^k}\cdot(1-p)^{n-k} =$$

$$= \frac{(n\cdot p)^k}{k!}\cdot\frac{n\cdot(n-1)\cdot(n-k+1)}{n^k}\cdot(1-p)^{-k}\cdot(1-p)^n =$$

$$= \frac{(\overset{\to\lambda}{np})^k}{k!}\cdot\underbrace{1\cdot\left(1-\frac{1}{n}\right)\cdot\left(1-\frac{2}{n}\right)\cdot\ldots\cdot\left(1-\frac{k+1}{n}\right)}_{\to 1}\cdot\underbrace{(1-p)^{-k}}_{\to 1}\cdot\Big[\underbrace{(1-p)^{\frac{1}{p}}}_{\to e^{-1}}\Big]^{\overset{\to\lambda}{np}} \longrightarrow \frac{\lambda^k}{k!}\cdot e^{-\lambda}.\ \triangleright$$

Величина $\lambda \leftarrow np$ имеет смысл интенсивности появления успехов.

def| Распр. Пуассона: $p_k = e^{-\lambda}\cdot\dfrac{\lambda^k}{k!}$

А что если $n\to\infty$, но $p\not\to 0$?

Th5* [Локальная теорема Муавра-Лапласа]

Пусть в схеме Бернулли $0<p<1$, $\sigma \doteq \sqrt{n\cdot pq}$, $x(k) = \dfrac{k-np}{\sigma}$. Тогда:

$\forall M>0$ равн. по всем k: $|x(k)|\leq M \mapsto$ при $n\to\infty$:

$$C_n^k\cdot p^k\cdot q^{n-k} = \frac{1}{\sigma\cdot\sqrt{2\pi}}\cdot e^{\frac{-x(k)^2}{2}}\cdot(1+\bar{o}(1))$$

Это частный случай ЦПТ, $\varphi(x) = \dfrac{1}{\sqrt{2\pi}}\cdot e^{-x^2/2}$ — плот. стандартного равн. распр.

Случайные величины

До этого работали с абстрактными пр-вами $(\Omega, \mathcal{F}, \mathbb{P})$ и просто приписывали событиям их вероятности вручную. Необходима более гибкая метрика описания $(\Omega, \mathcal{F}, \mathbb{P})$, а значит и самого случ. эксперимента.

def Если Ω не более чем счётно и $\mathcal{F} = 2^{\Omega}$, то <u>случайной величиной</u> (с.в.) назовём функцию $\xi : \Omega \to \mathbb{R}$ (с этим опр. до этого и работали)

Пример Схема Бернулли:

$$\omega = (i_1, \dots, i_n),\ i_j \in \{0,1\} \Rightarrow \xi(\omega) \doteq \sum_{j=1}^{N} i_j,\ N \in \mathbb{N}$$

$\boxed{!}$ А что если Ω несчётное? Знать значение $\xi(\omega)$ не можем до реализации конкретной Ω $\Rightarrow$ смотрим на какие-то вероятности, касающиеся с.в.

Пример $\omega = (i_1, \dots, i_{10})$ – оценки за экзамен, $\xi(\omega) = \dfrac{i_1 + \dots + i_{10}}{10}$

Пусть хотим посмотреть на $\mathbb{P}(\xi(\omega) \geqslant 8.6) = \mathbb{P}(\{\omega : \xi(\omega) \geqslant 8.6\})$

Если $\mathcal{F} \neq 2^{\Omega}$, то м.б., что $\{\omega : \xi(\omega) \geqslant x\} \notin \mathcal{F}$, а это крайне важно т.к. $\mathbb{P}$ определена именно на $\mathcal{F}$. $\Rightarrow$ необходима модификация ~~определения~~ ξ.

def Пусть $\mathcal{K}$ – некоторый класс подмножеств Ω. Через $\sigma(\mathcal{K})$ будем обозначать σ-алг. подмножеств Ω, удовлетворяющую условиям:

(1): $\mathcal{K} \subset \sigma(\mathcal{K})$

(2): Если $\mathcal{A}$ – некоторая σ-алг. подмнож. Ω и $\mathcal{K} \subset \mathcal{A}$, то $\sigma(\mathcal{K}) \subseteq \mathcal{A}$.

В силу (2): $\sigma(\mathcal{K})$ называется <u>минимальной σ-алг., порожд. классом $\mathcal{K}$</u>.

При этом $\forall \mathcal{K}\ \exists!\ \sigma(\mathcal{K})$, т.е. определение $\sigma(\mathcal{K})$ корректно.

Пример $\Omega = \mathbb{R}$, $\mathcal{K}$ – класс всех откр. мн-в в $\mathbb{R}$.

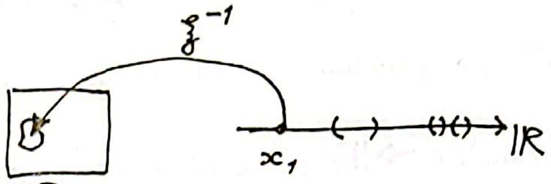
$\sigma(\mathcal{K}) \doteq \mathcal{B}(\mathbb{R}) \equiv \mathcal{B}$ – борелевская σ-алгебра.

Th$^{*}_{6}$ Пусть $\mathcal{T}$ – совокуп. всех полуинтервалов вида $(-\infty, x]$, $x \in \mathbb{R}$, тогда: $\sigma(\mathcal{T}) = \mathcal{B}$. Можно использовать $(-\infty, x)$, $[x, +\infty)$ и т.д.

def | Случайной величиной ξ называется числовая функ. $\Omega \to \mathbb{R}$ такая, что

$\forall B \in \mathcal{B} \mapsto \{\omega : \xi(\omega) \in B\} \equiv \xi^{-1}(B) \in \mathcal{F}$, т.е. какое бы борелевское мн-во B ни взяли, в $\mathcal{F}$ всегда найдётся его ξ-прообраз.

σ-алг. можно ассоциировать с информацией:



пусть мы знаем какое значение приняла $\xi = x_1$ после эксперимента, но не знаем какое именно элем. событие ω произошло.

Какое бы $B \in \mathcal{B}$ не взяли, можем знать $x_1 \in B$ или нет. При этом прообраз B лежит в $\mathcal{F}$. Т.о. о любом $B \in \mathcal{B}$ можно сказать произошло оно или нет в терминах элементов $\mathcal{F}$.

• Однако, если брать $\xi^{-1}(B)$, то можем получить не всё $\mathcal{F}$, например, для дискр. с.в. это будут атомы разбиения Ω.

$\Rightarrow$ Вводим $\sigma_\xi = \{\xi^{-1}(B), B \in \mathcal{B}\}$ – все ξ-прообразы борелевских множеств.

def | σ-алг., порождённая с.в. ξ.

Это опр. также явл. корректным: σ_ξ явл. σ-алг., $\sigma_\xi \subseteq \mathcal{F}$.

σ_ξ нужно ассоциировать с информацией, которую несёт в себе ξ т.к. относительно любого события из σ_ξ мы можем сказать произошло оно или нет, если знаем какое значение приняла ξ.

Можно показать, что если $\pi(\xi) = \{\xi^{-1}((-\infty, x]) : x \in \mathbb{R}\}$, то $\sigma(\pi(\xi)) = \sigma_\xi$

def | Пусть есть $(\Omega, \mathcal{F}, \mathbb{P})$. Классы множеств $\mathcal{K}_i \in \mathcal{F}$, $i = \overline{1,n}$ назыв. независимыми, если $\forall A_1 \in \mathcal{K}_1, \ldots, A_n \in \mathcal{K}_n \mapsto A_1, \ldots, A_n$ независ. в совокупности.

def | Пусть с.в. $\xi_1, \ldots, \xi_n$ опр. на $(\Omega, \mathcal{F}, \mathbb{P})$. $\xi_1, \ldots, \xi_n$ – независимы, если $\sigma_{\xi_1}, \ldots, \sigma_{\xi_n}$ – независимы.

Сейчас снова попали в ту же самую ловушку, что были перед введением опр. с.в. – сейчас не знаем как численно определить является ли набор с.в. независимым.

Для этого вводится функция распределения с.в.

Именно понятие с.в. позволяет перенести все рассуждения и вычисления в пр-во, где всё удобно считать.

$(\Omega, \mathcal{F}, \mathbb{P})$, $\xi : \Omega \to \mathbb{R}$ – с.в.

Рассмотрим $\mathbb{P} \circ \xi^{-1}$. Пусть $B \in \mathcal{B}(\mathbb{R}) \Rightarrow \xi^{-1}(B) \in \mathcal{F}$ по опр. с.в.

На $\mathcal{F}$ определена $\mathbb{P} \Rightarrow$ можно рассмотреть $\mathbb{P}(\xi^{-1}(B)) = \mathbb{P} \circ \xi^{-1}(B)$, при этом $\mathbb{P} \circ \xi^{-1}(B)$ яв. вероят. мерой (вероятностью) т.к. сохраняет при взятии прообраза все теоретико-множ. операции.

$\Rightarrow$ Возникает $\mathbb{P}_\xi(B) = \mathbb{P}(\{\omega : \xi(\omega) \in B\})$, $B \in \mathcal{B}$ – <u>def</u>| Распр. с.в. ξ.

Весь эффект переноса вероятности состоит в том, что все вычисления с абстрактного $(\Omega, \mathcal{F}, \mathbb{P})$ на $(\mathbb{R}, \mathcal{B}, \mathbb{P}_\xi)$.

(Th Каратеодори) + (Th. Однознач. $F_\xi \leftrightarrow \mathbb{P}_\xi$)

Оказывается, что $\mathbb{P}_\xi$ полностью определяется действием на $(-\infty, x]$ и функцией распр.

$F_\xi(x) \doteq \mathbb{P}_\xi((-\infty, x]) \equiv \mathbb{P}(\{\omega : \xi(\omega) \leq x\})$ – ф.р. с.в. ξ.

<u>Св-ва $F_\xi(x)$:</u>

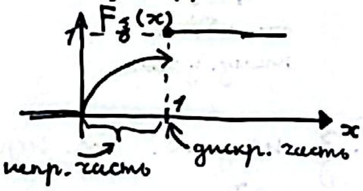
1) $F_\xi(x)$ не убывает
2) $F_\xi(x)$ непр. справа $\forall x \in \mathbb{R}$ и сущ. левосторонние пределы $F_\xi(x-0)$
3) $F_\xi(-\infty) = 0$, $F_\xi(+\infty) = 1$

Выразим вероятности попадания ξ в отрезки/интервалы через F_ξ:

1) $\mathbb{P}(\xi \leq a) = F_\xi(a)$, 2) $\mathbb{P}(\xi > a) = 1 - F_\xi(a)$

3) $\mathbb{P}(a < \xi \leq b) = F(b) - F(a)$, 4) $\mathbb{P}(a < \xi < b) = F(b-0) - F(a)$

Почему важно это: пусть с.в. ξ имеет ф.р. $F_\xi(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-x}, & x \in [0,1) \\ 1, & x \geq 1 \end{cases}$



$$\mathbb{E}\xi = \int_0^1 x \cdot F'(x)\,dx + 1 \cdot \mathbb{P}(\xi = 1) =$$

u dv $\int u\,dv = uv - \int v\,du$

$$= \int_0^1 x \cdot e^{-x}\,dx + (1 - (1 - e^{-1})) \cdot 1 =$$

$$= (-x \cdot e^{-x} - e^{-x})\Big|_0^1 + \frac{1}{e} = -\frac{2}{e} + 1 + \frac{1}{e} = 1 - \frac{1}{e}.$$