Теория вероятностей. Лекция вторая
Вероятность как предел

Дмитрий Валерьевич Хлопин
glukanat@mail.ru

Институт математики и механики им. Н.Н.Красовского

11.09.2018

Снова про схему Бернулли. Заблуждения

Пусть вероятность успеха p, неуспеха — $q = 1 - p$, и для n испытаний:

$$\Omega_n = \{\text{конечные строки из нулей и единиц длины } n\}, \qquad \mathcal{F}_n = 2^{\Omega_n};$$
$$\mathbb{P}_n\{\omega\} = p^{\text{число успехов в строке}} q^{\text{число неуспехов в строке}} \qquad \forall \omega \in \Omega_n;$$
$$\mathbb{P}_n(\text{число успехов ровно } k) = C_n^k p^k q^{n-k}.$$

Факт. Наиболее вероятно число успехов m в случае, если $m \in [np - q, np + p]$, то есть $m \approx np$.
Пусть n четно, $p = 1/2$ (орел/решка); тогда наиболее вероятно число успехов, равное $n/2$.
Заблуждение. Если выпал орел, то следующим выпадает решка, потому что вероятность $50/50$.

Столь же наивное заблуждение. Если достаточно долго подождать, то орлов и решек всегда поровну.

Снова про схему Бернулли. Факты

Пусть вероятность успеха p, неуспеха — $q = 1 - p$, и для n испытаний:

$$\Omega_n = \{\text{конечные строки из нулей и единиц длины } n\}, \qquad \mathcal{F}_n = 2^{\Omega_n};$$

$$\mathbb{P}_n\{\omega\} = p^{\text{число успехов в строке}} q^{\text{число неуспехов в строке}} \qquad \forall \omega \in \Omega_n;$$

$$\mathbb{P}_n(\text{число успехов ровно } k) = C_n^k p^k q^{n-k}.$$

Факт. Пусть n четно, $p = 1/2$, с помощью формулы Валлиса можно доказать

$$\mathbb{P}_n(\text{число успехов} = n/2) = C_n^{n/2} 2^{-n} \approx \sqrt{\frac{2}{\pi n}} \to 0.$$

Вывод: в точности $m/n = p$ бывает чем дальше, тем реже. Однако, при конечном числе n повторений заданных условий доля числа m случаев, когда случится успех, то есть так называемая *частота* m/n, как правило, близка к p.

Почему так? Насколько близка?

Схема Бернулли. Локальная теорема Муавра–Лапласа

Теорема 1. [без д-ва] Пусть k зависит от n так, что $\left|\frac{k(n)-np}{\sqrt{npq}}\right| \leq C$ для некоторого $C > 0$. Тогда имеет место формула Муавра-Лапласа:

$$\mathbb{P}_n(\text{число успехов} = k(n)) \approx \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(k(n)-np)^2}{2npq}} \left(1 + O\Big(\frac{1}{\sqrt{n}}\Big)\right).$$

Основная идея доказательства — формула Стирлинга:
$n! = \sqrt{2\pi n} \cdot n^n e^{-n}(1 + \frac{1}{12n} + \frac{1}{288n^2} + \ldots)$ Подробнее см., например wiki.
Пока доказывать ее не будем. Нам все равно требуется много более общий случай.
Считается, что теорему имеет смысл использовать в качестве точной оценки при $n > 100, npq > 20$. Имеются хорошие оценки погрешности: неравенство Бернштейна, неравенство Берри-Эссеена. В принципе, всю зиму мы будем крутиться вокруг этого и схожих с ним результатов.
Пока ограничимся...

Схема Бернулли. Локальная теорема Муавра–Лапласа

Теорема 1. [без д-ва] Пусть k зависит от n так, что $\left|\frac{k(n)-np}{\sqrt{npq}}\right| \le C$ для некоторого $C > 0$. Тогда имеет место формула Муавра-Лапласа:

$$\mathbb{P}_n(\text{число успехов} = k(n)) \approx \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(k(n)-np)^2}{2npq}} \left(1 + O\left(\frac{1}{\sqrt{n}}\right)\right).$$

Следствие. $\mathbb{P}_n\left(\left|\frac{\text{число успехов}}{n} - p\right| \le \frac{3}{\sqrt{npq}}\right) \to 0{,}9973\ldots.$

Следствие. $\mathbb{P}_n\left(\left|\frac{\text{число успехов}}{n} - p\right| \le \frac{4}{\sqrt{npq}}\right) \to 0{,}999937\ldots.$

Следствие. В рамках схемы независимых испытаний Бернулли [а на самом деле, в гораздо более общем случае]

$$\frac{\text{число успехов}}{n} \approx p \pm O\left(\frac{1}{\sqrt{n}}\right).$$

Вероятностное пространство. Что знаем...

Пока вероятностное пространство — это тройка $(\Omega, \mathcal{F}, \mathbb{P})$, где

Ω — множество элементарных событий (исходов), некоторое непустое множество;

$\mathcal{F}$ — множество событий, равное 2^Ω (наивные...);

$\mathbb{P}$ — вероятность, монотонная функция $\mathbb{P} : \mathcal{F} \to [0,1]$, т.е.

$$0 = \mathbb{P}(\varnothing) \leq \mathbb{P}(A) \leq \mathbb{P}(B) \leq \mathbb{P}(\Omega) = 1 \qquad \forall A, B \in \mathcal{F}, A \subset B.$$

Математическая основа статистики / эмпирическое определение вероятности: при конечном числе n повторений заданных условий доля m/n случаев, когда случится событие A, стремится к $\mathbb{P}(A)$ при $n \to \infty$.

События как подмножества

Если дано непустое множество Ω, то под событиями можно понимать (все или какие-то, пока непринципиально) его подмножества, элементы $\mathcal{F} \subset 2^{\Omega}$. На самом деле (смотрите задачу за 1 балл) можно начинать не с Ω, а с булевой алгебры $\mathcal{F}$, и кто там более перворoден, Ω или $\mathcal{F}$, еще вопрос.
Но можно начать иначе.

События как информация

Пусть событие — суждение с ответом да/нет о некотором эксперименте.

При этом точки ω из Ω можно трактовать как возможные исходы некоторого случайного эксперимента. Каждое подмножество множества Ω теперь связано с некоторым событием, при этом утверждение $\omega \in A$ можно трактовать как "при исходе эксперимента ω произошло событие A".

Если всегда, когда происходит событие A, то происходит и событие B — про такую пару событий можно писать "$A \subset B$".

Как и прежде, $\varnothing, \Omega \in \mathcal{F}$ — невозможное и достоверное события.

Если одновременно события A, B не наступают, то они называются взаимоисключающими (или несовместными).

Операции над событиями

$\overline{A} = \Omega \setminus A$ — "A не произошло";

$A \setminus B$ — "A произошло, а B — нет";

$A \cap B$ — "наступило и событие A, и событие B";

$A \cup B$ — "произошло хотя бы одно из событий A, B";

$\cap_{i=1}^{k} A_i$ — "наступило каждое из событий $A_1, A_2, \dots, A_k$";

$\cup_{i=1}^{k} A_i$ — "наступило хотя бы одно из событий $A_1, A_2, \dots, A_k$";

$\cap_{i \in \mathbb{N}} A_i$ — "наступило каждое из событий $A_1, A_2, \dots$";

$\cup_{i \in \mathbb{N}} A_i$ — "хотя бы для одного $k \in \mathbb{N}$ наступило A_k";

$\cap_{\alpha \in \mathcal{A}} A_\alpha$ — "события A_α наступили для каждого $\alpha \in \mathcal{A}$";

$\cup_{\alpha \in \mathcal{A}} A_\alpha$ — "хотя бы для одного $\alpha \in \mathcal{A}$ событие A_α наступило".

На вырост: пределы событий

Подумать: Для произвольной последовательности событий $(A_i)_{i\in\mathbb{N}}$ рассмотрим

$$\bigcap_{k\in\mathbb{N}}\bigcup_{i\geq k} A_i, \qquad \bigcup_{k\in\mathbb{N}}\bigcap_{i\geq k} A_i;$$

эти события называются соответственно верхним и нижним пределом последовательности $(A_i)_{i\in\mathbb{N}}$. Первое из них — "произошло бесконечное число этих событий (они никогда не кончатся)", второе — "начиная с некоторого произошли все эти события (не произошло только конечное число событий)".

Подумать: А что тогда означает запись

$$A = \lim_{n\to\infty} A_n?$$

Что хочется в случае $\mathcal{F} \neq 2^{\Omega}$:

- x0 $\varnothing \in \mathcal{F}$;
- x1 $\Omega \in \mathcal{F}$;
- x2 $\forall A \in \mathcal{F}\ \Omega \setminus A = \bar{A} \in \mathcal{F}$;
- x2п $\forall A \in \mathcal{F}\ \exists k \in \mathbb{N},\ A_1, A_2, \ldots, A_k \in \mathcal{F}\ \Omega \setminus A = \cup_{i=1}^{k} A_i$;
- x3 $\forall A, B \in \mathcal{F}\ B \setminus A \in \mathcal{F}$;
- x3п $\forall A, B \in \mathcal{F}\ \exists k \in \mathbb{N},\ A_1, A_2, \ldots, A_k \in \mathcal{F}\ B \setminus A = \cup_{i=1}^{k} A_i$;
- x4 $\forall A, B \in \mathcal{F}\ A \cap B \in \mathcal{F}$;
- x4k $\forall k \in \mathbb{N}\ A_1, A_2, \ldots, A_k \in \mathcal{F}\ \cap_{i=1}^{k} A_i \in \mathcal{F}$;
- x4σ $\forall A_1, A_2, \ldots, A_k, \ldots \in \mathcal{F}\ \cap_{i \in \mathbb{N}} A_i \in \mathcal{F}$;
- x4α $\forall (A_\alpha)_{\alpha \in \mathcal{A}} \in \mathcal{F}^{\mathcal{A}}\ \cap_{\alpha \in \mathcal{A}} A_\alpha \in \mathcal{F}$;
- x5 $\forall A, B \in \mathcal{F}\ A \cup B \in \mathcal{F}$;
- x5k $\forall k \in \mathbb{N}\ A_1, A_2, \ldots, A_k \in \mathcal{F}\ \cup_{i=1}^{k} A_i \in \mathcal{F}$;
- x5σ $\forall A_1, A_2, \ldots, A_k, \ldots \in \mathcal{F}\ \cup_{i \in \mathbb{N}} A_i \in \mathcal{F}$;
- x5α $\forall (A_\alpha)_{\alpha \in \mathcal{A}} \in \mathcal{F}^{\mathcal{A}}\ \cup_{\alpha \in \mathcal{A}} A_\alpha \in \mathcal{F}$.

В случае тривиальной алгебры ($\mathcal{F} = \{\varnothing, \Omega\}$) всё выполнено.

Совокупность подмножеств множества Ω называется

σ-алгеброй (борелевским полем, σ-полем), если его элементы замкнуты относительно счетного числа теоретико-множественных операций и $\Omega \in \mathcal{F}$;
[достаточно потребовать x1,x2,x4σ; или x1,x2,x5σ]
алгеброй (булевой алгеброй, полем событий), если его элементы замкнуты относительно конечного числа теоретико-множественных операций и $\Omega \in \mathcal{F}$;
[достаточно потребовать x1,x2,x4; или x1,x2,x5]
кольцом (булевым кольцом), если его элементы замкнуты относительно конечного числа теоретико-множественных операций;
[достаточно потребовать x3,x4; или x3,x5]
полукольцом (булевым полукольцом), если выполнены x3п,x4.

Самостоятельно дайте определения σ-кольца или полуалгебры.

Определение измеримого пространства

Множество Ω, снабженное своей σ-алгеброй $\mathcal{F}$, называется *измеримым пространством* и обозначается $(\Omega, \mathcal{F})$.

Подумать. Почему счетное число операций, а не произвольные объединения/пересечения?

Название "измеримое" не вполне удачно, полезнее было бы назвать информационное, но об этом в следующей лекции. Сегодняшняя цель — вероятность.

О вероятности. Хотелки

Пусть задано некоторое измеримое пространство $(\Omega, \mathcal{F})$. Что хочется:

- $x0$ $\mathbb{P}(\varnothing) = 0$;
- $x1$ $\mathbb{P}(\Omega) = 1$;
- $x2$ $\mathbb{P}(A) \leq \mathbb{P}(B)$ для любых $A, B \in \mathcal{F}$ с $A \subset B$;
- $x3$ $\mathbb{P}(\Omega \setminus A) = 1 - \mathbb{P}(A)$ для любых $A \in \mathcal{F}$;
- $x4$ $\mathbb{P}(\cup_{i=1}^{k} A_i) = \sum_{i=1}^{k} \mathbb{P}(A_i)$ для всех $A_i \in \mathcal{F}$ со свойством $A_i \cap A_j = \varnothing$ при $i \neq j$ (для попарно несовместных событий);
- $x4^2$ $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ для всех $A, B \in \mathcal{F}$.

Теорема 2 [С-но]. Для всякой неотрицательной аддитивной функции на алгебре событий со свойством $x1$ выполнено $x0$–$x4, x4^2$.

Аддитивность

Функцию из $\mathcal{F}$ в $\mathbb{R} \cup \{+\infty\}$ называют *аддитивной*, если выполнено свойство $x4$: для всех попарно несовместных событий $A_i \in \mathcal{F}$

$$\mathbb{P}(\cup_{i=1}^{k} A_i) = \sum_{i=1}^{k} \mathbb{P}(A_i).$$

Теорема 2 [С-но] Для всякой неотрицательной аддитивной функции на алгебре событий со свойством $x1$ выполнено $x0$–$x4, x4^2$.

Схема Бернулли. Ответы и снова вопросы

Как и прежде,

$$\Omega_n = \{\text{конечные строки из нулей и единиц длины } n\}, \qquad \mathcal{F}_n = 2^{\Omega_n};$$
$$\mathbb{P}_n\{\omega\} = p^{\text{число успехов в строке}} q^{\text{число неуспехов в строке}} \qquad \forall \omega \in \Omega_n.$$

Теперь, по теореме 2, вероятность $\mathbb{P}_n$ может быть продолжена на всё $\mathcal{F}$.

[С-но] Докажите, что $\mathbb{P}_n\{\text{число успехов ровно } k\} = C_n^k p^k q^{n-k}$,

что $\mathbb{P}_n\{\text{первый успех был на } k\text{-м испытании}\} = pq^{k-1}$ при $k = n$, $k < n$.

Подумать. Найдите вероятность того, что успех наступит хоть когда-то, хоть в каком-то испытании. А точно хватило $(\Omega_n, \mathcal{F}, \mathbb{P}_n)$? А точно хватило аддитивности вероятности?

О вероятности. Хотелки побольше

Что хотелось бы дополнительно:

$x5$ $\mathbb{P}(\cup_{i\in\mathbb{N}}A_i) = \sum_{i\in\mathbb{N}} \mathbb{P}(A_i)$ для всех таких $A_i \in \mathcal{F}, i \in \mathbb{N}$, что $A_i \cap A_j = \varnothing$ при $i \neq j$ (для попарно несовместных событий);

$x5_\uparrow$ $\mathbb{P}(\cup_{i\in\mathbb{N}}B_i) = \lim_{i\to\infty} \mathbb{P}(B_i)$ для всех таких $B_i \in \mathcal{F}, i \in \mathbb{N}$, что $B_1 \subset B_2 \subset \ldots \subset B_k \subset \ldots$ (непрерывность снизу);

$x5_\downarrow$ $\mathbb{P}(\cap_{i\in\mathbb{N}}C_i) = \lim_{i\to\infty} \mathbb{P}(C_i)$ для всех таких $C_i \in \mathcal{F}, i \in \mathbb{N}$, что $C_1 \supset C_2 \supset \ldots \supset C_k \supset \ldots$ (непрерывность сверху);

$x5_\varnothing$ $\lim_{i\to\infty} \mathbb{P}(D_i) = 0$ для всех таких $D_i \in \mathcal{F}, i \in \mathbb{N}$, что $D_1 \supset D_2 \supset \ldots \supset D_k \supset \ldots$ и $\cap_{i\in\mathbb{N}}D_i = \varnothing$ (непрерывность в $\varnothing$).

Теорема 3 [С-но или [Колмогоров, Т. II.1.1],[Синай, Т. 1.36]].
Для всякой неотрицательной аддитивной функции на алгебре событий свойства $x5$, $x5_\uparrow$, $x5_\downarrow$, $x5_\varnothing$ эквивалентны.

Д-во $x5_\downarrow \Leftrightarrow x5_\uparrow$. Достаточно задать $C_k = \Omega \setminus D_k, D_k = \Omega \setminus C_k$, теперь

$$\cup_{i\in\mathbb{N}} D_i = \Omega \setminus (\cap_{i\in\mathbb{N}} C_i)$$

осталось воспользоваться посылкой и свойством $x2$.

Д-во $x5_\downarrow \Leftarrow x5_\varnothing$. Достаточно задать $C = \cap_{i\in\mathbb{N}} C_i$, $D_k = C_k \setminus C$, теперь

$$\cap_{i\in\mathbb{N}} C_i = C \cup (\cap_{i\in\mathbb{N}} D_i)$$

осталось воспользоваться посылкой и свойством $x4$.

Д-во $x5_\uparrow \Leftrightarrow x5$. Достаточно задать $B_k = \cup_{i\leq k} A_k$ или $A_k = B_k \setminus B_{k-1}$ и подставить в уже известное; после $\cup_{i\in\mathbb{N}} A_i = \cup_{i\in\mathbb{N}} B_i$,
$\mathbb{P}(B_k) = \sum_{i\leq k} \mathbb{P}(A_k)$ останется лишь заметить, что сумма ряда это предел его частичных сумм.

Определения. От аддитивной функции до вероятности

Функцию из $\mathcal{F}$ в $\mathbb{R} \cup \{+\infty\}$ называют *аддитивной*, если выполнено свойство $x4$, σ-*аддитивной* — если выполнено свойство $x5$.

Вероятностью *(распределением вероятности, вероятностным распределением)* называют неотрицательную σ-аддитивную функцию из σ-алгебры событий $\mathcal{F}$ (над Ω) со свойством $\mathbb{P}(\Omega) = 1$.
Вероятностным пространством называют совокупность непустого множества, некоторой его σ-алгебры и вероятности, определенной на этой σ-алгебре.

Зарядом называют произвольную σ-аддитивную функцию из σ-алгебры событий $\mathcal{F}$ (над Ω) в $\mathbb{R} \cup \{+\infty\}$.
Мерой — произвольную неотрицательную σ-аддитивную функцию из σ-алгебры событий $\mathcal{F}$ (над Ω).

Геометрия *vs* аддитивность: $\mathcal{F} \neq 2^\Omega$

Плохой пример 1. Геометрическая вероятность на окружности $\Omega = S^1$ должна выдерживать поворот. Если геометрическую вероятность можно продолжить на 2^{S^1}, то окружность нельзя разделить на счетное число множеств, переводящихся друг в друга движением. Но так сделать можно [пример Витали, с-но].
Подумать. Но может дело в том, что мы потребовали σ-аддитивность?

Плохой пример 2. Возьмем параллелепипед $[0,1/2] \times [0,1] \times [0,1]$. По парадоксу Банаха–Тарского можно его разделить на конечное число частей, поперемещать их и составить из них куб $[0,1] \times [0,1] \times [0,1]$. Если бы объем у таких частей был всегда, то $1/2 = 1$. Значит хорошо продолжить объем, на все подмножества, например куба $[0,1] \times [0,1] \times [0,1]$, нельзя.
В общем случае — любые два (в хотя бы трехмерном пр-ве) объекта с непустой внутренностью равносоставлены.

Вероятностное пространство. Итог

Вероятностное пространство — это тройка $(\Omega, \mathcal{F}, \mathbb{P})$, где

Ω — множество элементарных событий (исходов), некоторое непустое множество;

$\mathcal{F}$ — σ-алгебра событий, $\mathcal{F} \subset 2^{\Omega}$;

$\mathbb{P}$ — вероятность, счетно-аддитивная функция $\mathbb{P} : \mathcal{F} \to [0,1]$, т.е.$\mathbb{P}(\Omega) = 1$ и

$$\mathbb{P}(A \sqcup B) = \mathbb{P}(A) + \mathbb{P}(B) \qquad \forall A, B \in \mathcal{F}, A \cap B = \varnothing.$$

Смысл вероятности: при конечном числе n повторений заданных условий доля m/n случаев, когда случится событие A, стремится к $\mathbb{P}(A)$ при $n \to \infty$.

Задача теории вероятностей — рассчитать вероятность сложных событий.

Одна из задачек на эту лекцию, и вообще говоря, краеугольный камень и статистики тоже

Задача [О тривиальности горизонта; 1 балл] Пусть на одном и том же вероятностном пространстве дана последовательность событий A_n. Назовем событие A *далеким от народа*, если для всех $k \in \mathbb{N}$, по событиям A_n начиная с k-го, можно определить выполнено ли событие A. Докажите, что вероятность любого далекого от народа события равна или нулю, или единице.

В качестве тривиального следствия...

Задача [предел должен быть; 1 балл] Докажите, опираясь на предыдущую задачу, что в схеме Бернулли с бесконечным числом независимых испытаний

$$\mathbb{P}\Big(\frac{\text{число успехов за первые } n \text{ испытаний}}{n} \text{ имеет предел при } n \uparrow \infty\Big) = 1.$$