Вероятностное пространство.

Определение 1. Пусть Ω — множество, тогда $\mathfrak{A} \in 2^{\Omega}$ называется **алгеброй**, если

1. $\Omega \in A$.
2. $\forall A \in \mathfrak{A}\ \overline{A} \in \mathfrak{A}$. Здесь и далее $\overline{A} = \Omega \setminus A$.
3. $\forall A, B \in \mathfrak{A}\ A \cup B \in \mathfrak{A}$.

При этом Ω называется **множеством элементов событий**, $\mathfrak{A}$ — **набор событий**, $A \in \mathfrak{A}$ — **событие**, $A \cup B = A + B$ — **сумма событий**, $\overline{A}$ — **противоположное событие**, $A \cap B = AB$ — **произведение событий**.

Определение 2. Алгебра является **сигма-алгеброй**, если она замкнута относительно объединения счётного количества своих элементов.

Определение 3. Пусть $\mathfrak{A}$ — сигма-алгебра на Ω. Пусть $P\colon \mathfrak{A} \to [0; +\infty)$ и

1. $P(\Omega) = 1$.
2. Если $\{A_i\}_{i=1}^{\infty} \subset \mathfrak{A}$ и $\forall A_i A_j = \varnothing$ то

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Тогда $(\Omega; \mathfrak{A}; P)$ называется **вероятностным пространством**.

Определение 4. Пара событий называется **несовместной**, если их пересечение пусто. Набор событий **несовместен**, если они попарно несовместны.

Определение 5. Пусть $A \subset 2^{\Omega}$ — алгебра. Тогда минимальная по включению сигма-алгебра $\sigma(A) \supset A$ называется **минимальной сигма-алгеброй**.

Утверждение. *Таковая существует.*

Доказательство. Хотя бы одна такая существует (2^{Ω}), причём если пересечь сколько угодно сигма-алгебр, то получится искомая сигма-алгебра. □

Определение 6. Пусть $\mathfrak{A}$ — алгебра на Ω, $P\colon \mathfrak{A} \to [0; +\infty)$ и

- $P(\Omega) = 1$.
- Если $\{A_i\}_{i=1}^{\infty} \subset \mathfrak{A}$ и $\bigsqcup\limits_{i=1}^{\infty} A_i \in \mathfrak{A}$, то

$$P\left(\bigsqcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Тогда $(\Omega; \mathfrak{A}; P)$ — **вероятностное пространство в широком смысле**.

Теорема 1 (О продолжении меры). *Пусть $(\Omega; \mathfrak{A}; P)$ — вероятностное пространство в широком смысле. Тогда существует единственная функция вероятности $Q\colon \sigma(\mathfrak{A}) \to [0 : +\infty)$, такое что $Q\Big|_{\mathfrak{A}} \equiv P$.*
Без доказательства.

Замечание. Эта теорема позволяет нам сказать, например, что мы хотим задать вероятность на отрезках.

Определение 7. Борелевская сигма-алгебра — минимальная σ-алгебра, которая содержит все открытые множества.

Пример. Дискретное вероятностное пространство: $\Omega = \{\omega_i\}_{i=1}^{N}$, $A = 2^{\Omega}$, $P(\{\omega_i\}) = p_i$, $\sum p_i = 1$. Тогда $P(A)$ — сумма вероятностей элементов A.

Пример. Геометрическая вероятность: $\Omega \subset \mathbb{R}^n$, измеримо по Лебегу, $\mu A < +\infty$, $\mathfrak{A}$ состоит из измеримых по Лебегу множеств, $P(A) = \frac{\mu A}{\mu \Omega}$. Обычно при этом $\mathbb{R}^n$ не более чем трёхмерно.

Свойства вероятности.

Свойство 7.1.

$$\forall A, B \in \mathfrak{A}\ A \subset B \Rightarrow P(A) \leqslant P(B)$$

Доказательство. Понятно, что $B \setminus A \in \mathfrak{A}$. Тогда

$$B = A \sqcup (B \setminus A) \Rightarrow P(B) = P(A) + P(B \setminus A) \geqslant P(A)$$

□

Следствие 0.1.

$$\forall A \in \mathfrak{A}\ P(A) \leqslant 1$$

Свойство 7.2.

$$P(A) = 1 - P(\overline{A})$$

Свойство 7.3.

$$P(A + B) = P(A) + P(B) - P(AB)$$

Доказательство.

$$B = (B \setminus AB) \sqcup AB \Rightarrow P(B) = P(B \setminus AB) + P(AB)$$

Тогда

$$P(A + B) = P(A) + P(B \setminus AB) = P(A) + P(B) - P(AB)$$

□

Утверждение (Формула включений-исключений)**.**

$$P(A_1 + \dots + A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{\substack{i,j=1 \\ i<j}}^{n} P(A_i A_j) + \sum_{\substack{i,j,k=1 \\ i<j<k}}^{n} P(A_i A_j A_k) - \dots + (-1)^n P(A_1 \cdots A_n)$$

Доказательство. Мне лень это писать, докажите сами по индукции. □

Утверждение.

$$P\left(\bigcup_i A_i\right) \leqslant \sum_i P(A_i)$$

Доказательство. Пусть $B_1 = A_1$, $B_2 = A_2\overline{A_1}$, $B_3 = A_3\overline{A_1 \cup A_2}$ и так далее. Тогда

$$\bigcup_i A_i = \bigsqcup_i B_i$$

При этом $B_i \subset A_i$, а значит

$$\sum_i P(A_i) \geqslant \sum_i P(B_i)$$

□

Теорема 2. *Пусть* $(\Omega; \mathfrak{A}; P)$ *— вероятностное пространство. Тогда следующие утверждения равносильны:*

1. P *счётно-аддитивна.*
2. P *конечно-аддитивна и* $\forall \{B_i\}_{i=1}^{\infty} : B_{i+1} \subset B_i,\ B = \bigcap\limits_{i=1}^{\infty} B_i\ \lim\limits_{n\to\infty} P(B_i) = P(B)$ *(непрерывность сверху).*

3. *P конечно-аддитивна и $\forall\{C_i\}_{i=1}^{\infty} : C_{i+1} \supset B_i$, $C = \bigcup\limits_{i=1}^{\infty} C_i$ $\lim\limits_{n\to\infty} P(C_i) = P(C)$ (непрерывность сверху).*

Доказательство. Равносильность двух непрерывностей тривиально из формул де Моргана.
Докажем, что из 1 следует 2. Конечная аддитивность есть, докажем непрерывность сверху. Пусть $A_1 = B_1\overline{B_2}$, $A_2 = B_2\overline{B_3}$ и так далее. Очевидно, A_i несовместны. Также очевидно, что A_i несовместны с B. Также заметим, что

$$B_n = B \sqcup \bigsqcup_{i=n+1}^{\infty} A_i$$

Отсюда $P(B_n) = P(B) + \sum\limits_{i=n+1}^{\infty} P(A_i)$, а справа остаток (очевидно, сходящегося) ряда, который стремится к нулю при $n \to \infty$.
Теперь из 2 докажем 1. Рассмотрим $\{A_i\}_{i=1}^{\infty}$ несовместные. Очевидно,

$$\sum_{i=1}^{\infty} P(A_i) = \lim_{n\to\infty} \sum_{i=1}^{n} P(A_i)$$

А ещё мы знаем, что

$$\bigsqcup_{i=1}^{\infty} A_i = \bigsqcup_{i=1}^{n} A_i \sqcup \bigsqcup_{i=n+1}^{\infty} A_i$$

То есть

$$\lim_{n\to\infty} P\left(\bigsqcup_{i=1}^{n} A_i\right) = \lim_{n\to\infty}\left(P\left(\bigsqcup_{i=1}^{\infty} A_i\right) - P\left(\bigsqcup_{i=n+1}^{\infty} A_i\right)\right) = P\left(\bigsqcup_{i=1}^{\infty} A_i\right) - \lim_{n\to\infty} P\left(\bigsqcup_{i=n+1}^{\infty} A_i\right)$$

Второе слагаемое — ноль но непрерывности меры, а отсюда счётная аддитивность. □

Условная вероятность.

Замечание. Пусть $|\Omega| = n$, $|A| = k$, $|B| = m$, $|AB| = l$. Если мы знаем, что B произошло, как узнать вероятность того, что произошло A? Ну, это

$$\frac{l}{m} = \frac{l/n}{m/n} = \frac{P(AB)}{P(B)}$$

Определение 8. Пусть $(\Omega; \mathfrak{A}; P)$ — вероятностное пространство, $B \in \mathfrak{A}$, $P(B) > 0$. Тогда **условной вероятностью** A при условии B называется

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

Также обозначается $P_B(A)$.

Свойство 8.1. *Несложно проверить, что условная вероятность является вероятностью.*

Утверждение (Произведение вероятностей)**.** *Несложно по определению проверить*

$$P(A_1 \cdots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2) \cdots P(A_n|A_1A_2 \cdots A_{n-1})$$

Теорема 3 (Формула полной вероятности)**.** *Пусть $A \in \mathfrak{A}$, $B_i \in \mathfrak{A}$ несовместны, $A \subset \bigsqcup\limits_{i=1}^{n} B_i$ (обычно объединение равно Ω), и $\forall i \in [0 : n]$ $P(B_i) > 0$. Тогда*

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$$

Доказательство.

$$P(A) = P\left(A \cap \bigsqcup_{i=1}^{n} B_i\right) = P\left(\bigsqcup_{i=1}^{n} A \cap B_i\right)$$

Всё. □

Теорема 4 (Формула Байеса)**.** *Пусть $A, B \in \mathfrak{A}$, $P(A) > 0$, $P(B) > 0$. Тогда*

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

Доказательство. Очевидно из определения. □

Определение 9. События $A, B \in \mathfrak{A}$ называются **независимыми**, если

$$P(AB) = P(A)P(B)$$

Определение 10. Говорят, что $A_1, A_2, \ldots, A_n \in \mathfrak{A}$ **независимы в совокупности**, если $P(A_1A_2\cdots A_n) = P(A_1)P(A_2)\cdots P(A_n)$

Свойство 10.1. *Несложно проверить, что независимость событий A, B равносильна $P(A|B) = P(A)$.*

Свойство 10.2. *Независимые в совокупности события попарно независимы. Обратное неверно.*

Определение 11. Пусть у нас есть два вероятностных пространства: $(\Omega_1, \mathfrak{A}_1; P_1)$ и $(\Omega_2, \mathfrak{A}_2; P_2)$. Рассмотрим вот такое вероятностное пространство: $(\Omega, \mathfrak{A}; P)$, где $\Omega = \Omega_1 \times \Omega_2$, $\mathfrak{A}$ — минимальная σ-алгебра, включающая в себя $\mathfrak{A}_1 \times \mathfrak{A}_2$,

$$P((A_1; A_2)) = P_1(A_1)P_2(A_2)$$

Тогда $(\Omega_1, \mathfrak{A}_1; P_1)$ и $(\Omega_2, \mathfrak{A}_2; P_2)$ — независимые испытания.

Пример. Схема Бернулли: $\Omega_1 = \{0; 1\}$, $\mathfrak{A}_1 = 2^{\Omega_1}$, $P_1(1) = p$, $P_1(0) = 1 - p = q$. Хочется рассмотреть эту штуку в степени n (то есть n одинаковых независимых испытаний). Тогда что у нас получается для $\omega \in \Omega = \Omega_1^n$?

$$P(\omega) = \sum_{i=1}^{n} P_i(\omega_i) = p^{\sum \omega_i} q^{n - \sum \omega_i}$$

Посчитаем тут такую вероятность: пусть S_n — количество успехов в n испытаниях? Посчитаем вероятность того, что $S_n = k$? Очевидно, оно равно $\binom{n}{k} p^k q^{n-k}$.

Утверждение. *Пусть k^* — наиболее вероятное число успехов в Бернуллиевских испытаниях. Тогда*

$$k^* = \begin{cases} p(n-1) \text{ или } p(n-1)+1 & p(n-1) \in \mathbb{N} \\ \lceil p(n-1) - 1 \rceil & p(n-1) \notin \mathbb{N} \end{cases}$$

Доказательство. Давайте рассмотрим вот такое частное:

$$\frac{P(S_n = k+1)}{P(S_n = k)}$$

Чему оно равно?

$$\frac{P(S_n = k+1)}{P(S_n = k)} = \frac{\binom{n}{k+1} p^{k+1} q^{n-k-1}}{\binom{n}{k} p^k q^{n-k}} = \frac{p}{q} \cdot \frac{n-k}{k+1}$$

Нам хочется оценить, больше это чем 1 или меньше (это позволит нам найти K^*). Ну,

$$\frac{p}{q}\cdot\frac{n-k}{k+1} > 1 \Leftrightarrow p(n-k) > q(k+1) \Leftrightarrow pn - pk > k = pk + 1 - p \Leftrightarrow pn > k + 1 - p \Leftrightarrow pn + p - 1 > k$$

То есть возрастание достигается при $k < p(n-1) - 1$, а иначе убывание. Тогда где экстремум? Рассмотрим $k = p(n-1) - 1$. Если это целое число, то там $P(S_n = k+1) = P(S_n = k)$, и это самое k даёт значение больше остальных. То есть $k^* = p(n-1) - 1$ или $k^* = p(n-1)$.
А что если оно не целое? То надо куда-то округлить. А именно вверх, потому что тогда оно больше, чем следующее, а предыдущее меньше его. $\square$

Пример. Пусть $n = 10000$, $p = \frac{1}{10000}$. Давайте посчитаем $P(S_n > 3)$. Ну, это

$$1 - P(S_n \leqslant 3) = 1 - q^{10000} - 10000pq^{10000-1} - \binom{10000}{2}p^2q^{10000-2} - \binom{10000}{3}p^3q^{10000-3}$$

Фиг мы такое посчитаем.

Пример. Или если взять $p = q = 0.5$, то при $n = 5\cdot 10^3$ мы не сможем нормально посчитать $P(S_n = 2349)$.

Замечание. Ну и как такое считать?

Теорема 5 (Теорема Пуассона)**.** *Пусть у нас есть несколько схем Бернулли. В первой одно испытание и вероятность успеха p_1, во второй — 2 и вероятность успеха p_2, в n-ной n испытаний и вероятность p_n. Пусть $np_n \xrightarrow[n\to\infty]{} \lambda > 0$. Тогда*

$$P(S_n = k) \xrightarrow[n\to\infty]{} e^{-\lambda}\frac{\lambda^k}{k!}$$

Доказательство. Известно,

$$P(S_n = k) = \frac{1}{k!}n(n-1)\cdots(n-k+1)p_n^k(1-p_n)^{n-k}$$

Известно, что

$$np_n = \lambda + o(1) \Rightarrow p_n = \frac{\lambda}{n} + o\left(\frac{1}{n}\right)$$

Тогда

$$P(S_n = k) = \frac{1}{k!}\cancelto{1}{n(n-1)\cdots(n-k+1)\frac{1}{n^k}}\;\cancelto{\lambda^k}{(\lambda + o(1))^k}\;\frac{\cancelto{e^{-\lambda}}{\left(1 - \frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right)^n}}{\cancelto{1}{\left(1 - \frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right)^k}} \xrightarrow[n\to\infty]{} e^{-\lambda}\frac{\lambda^k}{k!}$$

$\square$

Лемма 1. *Пусть $p \in (0;1)$, $H(x) = x\ln\frac{x}{p} + (1-x)\ln\frac{1-x}{1-p}$. Пусть $p^* = \frac{k}{n}$. Пусть $k \to +\infty$, $n-k \to +\infty$. Тогда*

$$P(S_n = k) \sim \frac{1}{\sqrt{2\pi np^*(1-p^*)}}\exp(-nH(p^*))$$

Доказательство. Мы знаем формулу Стирлинга

$$n! \sim \sqrt{2\pi n}n^ne^{-n}$$

Тогда

$$P(S_n = k) = \frac{\sqrt{2\pi n}n^ne^{-n}}{\sqrt{2\pi\underbrace{k}_{np^*}}k^ke^{-k}\sqrt{2\pi\underbrace{(n-k)}_{n(1-p^*)}}(n-k)^{n-k}e^{-n+k}}p^k(1-p)^{n-k} =$$

$$= \frac{n^np^k(1-p)^{n-k}}{\sqrt{2\pi np^*(1-p^*)}k^k(n-k)^{n-k}} = \frac{1}{\sqrt{2\pi np^*(1-p^*)}}\exp\underbrace{\ln\frac{n^np^k(1-p)^{n-k}}{k^k(n-k)^{n-k}}}_{L}$$

При этом

$$\begin{aligned}
L &= \ln\frac{n^n p^k (1-p)^{n-k}}{k^k (n-k)^{n-k}} = \ln\frac{n^n p^k (1-p)^n (n-k)^k}{k^k (n-k)^n (1-p)^k} = \\
&= \ln\left(\underbrace{\frac{n^n}{(n-k)^n}}_{(1-p^*)^{-n}}(1-p)^n\right) + \ln\frac{p^k (n-k)^k}{(np^*)^k (1-p)^k} = \\
&= n\ln\frac{1-p}{1-p^*} + k\ln\frac{p}{p^*} + k\ln\frac{(n-k)}{n(1-p)} = n\ln\frac{1-p}{1-p^*} + k\ln\frac{p}{p^*} + k\ln\frac{1-p^*}{1-p} = \\
&= -(n-k)\ln\frac{1-p^*}{1-p} - k\ln\frac{p^*}{p} = -n\underbrace{\left(p^*\ln\frac{p^*}{p} + (1-p^*)\ln\frac{1-p^*}{1-p}\right)}_{H(p^*)}
\end{aligned}$$

Это ли не то, что нам надо? □

Лемма 2.

$$H(x) = \frac{(x-p)^2}{2p(1-p)} + O((x-p)^3)$$

Доказательство.

$$H'(x) = \ln\frac{x}{p} + x\cdot\frac{p}{x}\cdot\frac{1}{p} - \ln\frac{1-x}{1-p} - 1 = \ln\frac{x}{p} - \ln\frac{1-x}{1-p}$$

$$H''(x) = \frac{1}{x} + \frac{1}{1-x}$$

Тогда $H'(p) = 0$, $H''(p) = \frac{1}{p(1-p)}$. По Тейлору получаем искомое. □

Теорема 6 (Локальная теорема Муавра — Лапласа)**.** *Пусть* $p \in (0;1)$, $H(x) = x\ln\frac{x}{p} + (1-x)\ln\frac{1-x}{1-p}$*.*
Пусть $p^* = \frac{k}{n}$*. Пусть* $k \to +\infty$, $n-k \to +\infty$*.*
Пусть $k = np = p(n^{2/3})$*. Тогда*

$$P(S_n = k) \sim \frac{1}{\sqrt{2\pi np(1-p)}}\exp\left(-\frac{(k-np)^2}{2np(1-p)}\right)$$

Доказательство. Известно

$$P(S_n = k) \sim \frac{1}{\sqrt{2\pi np(1-p)}}\exp(-nH(p^*))$$

Отсюда

$$P(S_n = k) \sim \frac{1}{\sqrt{2\pi np(1-p)}}\exp(-n\frac{(p^*-p)^2}{2p(1-p)} + n\cdot O((p^*-p)^3))$$

Заметим, что $\frac{k}{n} - p = o(n^{-1/3})$. Тогда

$$P(S_n = k) \sim \frac{1}{\sqrt{2\pi np(1-p)}}\exp\left(-n\frac{(p-k/n)^2}{2p(1-p)} + O(n(k/n-p)^3)\right) \sim \frac{1}{\sqrt{2\pi np(1-p)}}\exp\left(-n\frac{(np-k)^2}{2p(1-p)n^2} + o(1)\right)$$

Что и требовалось доказать. □

Теорема 7 (Интегральная теорема Муавра — Лапласа)**.** *Пусть*

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}}\,\mathrm{d}t$$

Далее мы будем называть эту функцию функцией стандартного нормального распределения. Тогда

$$\sup_{-\infty<x_1<x_2<+\infty} \left| P\left(x_1 \leqslant \frac{S_n - np}{\sqrt{npq}} \leqslant x_2\right) (\Phi(x_2) - \Phi(x_1)) \right| \underset{n\to\infty}{\longrightarrow} 0$$

Иными словами

$$P\left(x_1 \leqslant \frac{S_n - np}{\sqrt{npq}} \leqslant x_2\right) \approx \frac{1}{\sqrt{2\pi}} \int\limits_{x_1}^{x_2} e^{-\frac{t^2}{2}} \, \mathrm{d}t$$

Пока без доказательства.