

1 Теория вероятностей.

Монетка:

$\Omega = \{O; P\}$. Ω — множество элементарных исходов; \leq счетное.

Случайное событие: $A \subset \Omega$.

$P: \Omega \rightarrow \mathbb{R}_+$:

1. $P(x) \geq 0$
2. $\sum_{x \in \Omega} P(x) = 1$

(Ω, P) — вероятностное пространство.

Вероятность события A — $P(A) = \sum_{x \in A} P(x)$.

Случайное событие — измеримое подмножество вероятностного пространства.

Сумма событий — объединение множеств.

Произведение событий — пересечение множеств.

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

$P(A + B + C) = P(A) + P(B) + P(C) - P(AB) - P(BC) - P(AC) + P(ABC)$.

Опр. A и B называются независимыми, если $P(AB) = P(A) \cdot P(B)$.

Пример:

Ω_1 и Ω_2 — вероятностные пространства.

$\Omega_1 \times \Omega_2$ — вероятностное пространство.

$A \subset \Omega_1$ — случайное событие.

$P(x, y) = P_1(x) \cdot P_2(y)$.

$P_1(A) = P(A \times \Omega_2)$.

$AB \leftrightarrow (A \times \Omega_2) \cap (\Omega_1 \times B) = A \times B$.

A, B, C — независимы, если $P(ABC) = P(A)P(B)P(C)$.

События A_1, A_2, \dots, A_n независимы, если $\forall A_{i_1}, \dots, A_{i_k}$ верно $P(A_{i_1}A_{i_2} \dots A_{i_k}) = \prod_{j=1}^k P(A_{i_j})$.

Опр. Условная вероятность. $P(A|B)$ — вероятность A при условии, что B произошло.

Теорема. Формула полной вероятности. Пусть A_1, A_2, \dots, A_n — система попарно несовместимых событий, в сумме дающих достоверное. Тогда для любого события B выполнена формула $P(A) = \sum_{i=1}^k P(A|B_i) \cdot P(B_i)$. Доказательство:

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

$$P(A) = \sum_{i=1}^k P(AB_i).$$

1.1 Случайные величины.

Опр. Ω — вероятностное пространство, $f: \Omega \rightarrow \mathbb{R}$. f называется случайной величиной.

Пример:

Бросаем кубик. $\Omega = 1, 2, 3, 4, 5, 6$, $p(x) = \frac{1}{6}$. $f(1) = 1, f(2) = 2$, и т. д.

На \mathbb{R} введем структуру вероятностного пространства: $1 - \frac{1}{6}, 2 - \frac{1}{6}, 3 - \frac{1}{6}, 4 - \frac{1}{6}, 5 - \frac{1}{6}, 6 - \frac{1}{6}$, остальные 0.

Опр. $f: \Omega \rightarrow \mathbb{R} \Rightarrow$ на \mathbb{R} вводится структура вероятностного пространства. $\forall x \in \mathbb{R} P_{\mathbb{R}}(x) = P_{\Omega}(f^{-1}(x))$.

Опр. Распределение случайной величины f — структура вероятностного пространства на \mathbb{R} .

Стандартное распределение:

1. Дискретное равномерное. $x_1 - \frac{1}{n}, x_2 - \frac{1}{n}, \dots, x_n - \frac{1}{n}$.

2. Распределение Бернулли.

$$f(x) = \begin{cases} 1 & p \\ 0 & q = 1 - p \end{cases}$$

$$f = \sum_{i=1}^n f_i.$$

$$0 - q^n, 1 - C_n^1 p q^{n-1}, \dots, n.$$

$$P(f(x) = k) = C_n^k \cdot p^k \cdot q^{n-k}.$$

3. Геометрическое распределение. Бросаем неравновесную монетку до выпадения орла. Случайная величина — количество испытаний.

$$P(f(x) = k) = q^{k-1} \cdot p.$$

4. Распределение Пуассона.

$$P(f = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}. e \approx 2.718281828459045, \lambda > 0.$$

5. Гипергеометрическое распределение. В сосуде N белых шаров и M черных. Мы не глядя вытаскиваем n шаров без возвращения. Случайная величина — количество белых шаров.

$$P(f = k) = \frac{C_N^k \cdot C_M^{n-k}}{C_{N+M}^n}.$$

Опр. Сложение. $f : \Omega \rightarrow \mathbb{R}, g : \Omega \rightarrow \mathbb{R}. (f + g)(x) = f(x) + g(x).$

Опр. $f(x) = a$ — случайное событие.

Опр. f и g независимые, если $\forall A, B \in \mathbb{R} : P(f(x) \in A; g(x) \in B) = P(f(x) \in A) \cdot P(g(x) \in B)$. Или: $P(f(x) = a; g(x) = b) = P(f(x) = a) \cdot P(g(x) = b).$

1.2 Математическое ожидание.

Опр. f — случайная величина. Ее математическим ожиданием называется $Mf = Ef = \sum_{x \in \Omega} f(x)P(x).$

Свойства:

1. c - константа это случайная величина.
2. $f_1 \leq f_2 (\Leftrightarrow f_1(x) \leq f_2(x) \forall x) \Rightarrow Mf_1 \leq Mf_2.$
3. $M(cf) = c \cdot Mf.$
4. $M(f_1 + f_2) = Mf_1 + Mf_2. \sum_{x \in \Omega} (f_1(x) + f_2(x))P(x) = \sum f_1(x) \cdot P(x) + \sum f_2(x) \cdot P(x)$
5. Если f_1, f_2 - независимы $\Rightarrow M(f_1 \cdot f_2) = Mf_1 \cdot Mf_2. Mc = c.$

$$M(f_1 \cdot f_2) = \sum_{x \in \Omega} f_1(x) \cdot f_2(x) \cdot P(x) = \sum a_i \cdot b_j \cdot P(f_1 = a_i; f_2 = b_j) = \sum_{i,j} a_i \cdot b_j \cdot P(f_1 = a_i) \cdot P(f_2 = b_j) = \sum_i a_i \cdot P(f_1 = a_i) \cdot \sum_j b_j \cdot P(f_2 = b_j) = Mf_1 \cdot Mf_2.$$

$$P(f = k) = \frac{C_M^k \cdot C_N^{n-k}}{C_{M+N}^n}.$$

$$Mf = \frac{n \cdot M}{M+N}.$$

$$Mf = \sum_{k=0}^n \frac{k \cdot C_M^k \cdot C_N^{n-k}}{C_{M+N}^n} = \frac{n \cdot M}{M+N} \cdot \frac{(M-1)!N!(n-1)!(M+N-n)!}{(M+N-1)!} \sum_{k=0}^n \frac{1}{(k-1)!(M-k)!(n-k)!(N-n+k)!} \sum_{k=0}^n \frac{(M-1)!N!(n-1)!(M+N-n)!}{(k-1)!(M-k)!(n-k)!(N-n+k)!(M+N-1)!} \\ = \sum_{k=0}^N \frac{C_{n-1}^{k-1} \cdot C_{M+N-n}^{M-k}}{C_{M+N-1}^{M-1}} = 1.$$

1.3 Дисперсия.

Опр. Мера отклонения случайной величины от своего математического ожидания. f — случайная величина; $Df = M(f - Mf)^2.$

Опр. Центрирование случайной величины — $M(f - Mf) = Mf - M(Mf) = 0.$ Условно сдвигаем систему координат в 0.

$$Df = M(f^2 - 2f \cdot Mf + (Mf)^2) = Mf^2 - 2M^2f + M(M^2f) = Mf^2 - M^2f.$$

Свойства:

1. $D(c) = 0 \quad D(f) = 0 \Rightarrow f = const$
2. $D(cf) = c^2 D(f)$
3. $D(f + c) = Df$
4. Если случайные величины f и g независимы, то $D(f + g) = Df + Dg$

Формула Стирлинга. $n! \approx \sqrt{2\pi n} \cdot \left(\frac{n}{e}\right)^n \cdot e^{\frac{\theta_n}{12n}}.$ При больших $n, e^{\frac{\theta_n}{12n}} \approx 1.$

1.4 Неравенство Маркова.

f — случайная величина; $|Mf| < \infty; c > 0; f \geq 0.$

$$P(f \geq c) \leq \frac{Mf}{c}.$$

$f :$

$a_1 < c$	$a_2 < c$	\dots	$a_n \geq c$	$a_{n+1} \geq c$	$a_{n+2} \geq c$	\dots
p_1	p_2	\dots	p_n	p_{n+1}	p_{n+2}	\dots

$$g := \frac{f}{c}.$$

$\frac{a_1}{c} < 1$	$\frac{a_2}{c} < 1$	\dots	$\frac{a_n}{c} \geq 1$	$\frac{a_{n+1}}{c} \geq 1$	$\frac{a_{n+2}}{c} \geq 1$	\dots
p_1	p_2	\dots	p_n	p_{n+1}	p_{n+2}	\dots

Надо: $P(g \geq 1) \leq Mg$.

$$b_i = \frac{a_i}{c}.$$

$$Mg = b_1 p_1 + b_2 p_2 + \dots + b_n p_n + b_{n+1} p_{n+1} + \dots \geq p_{n+1} + p_{n+2} + \dots$$

1.5 Неравенство Чебышева.

f — случайная величина, Mf и $Df \exists, c > 0$.

$$\Rightarrow P(|f - Mf| \geq c) \leq \frac{Df}{c^2}.$$

Нормировка случайной величины. $g := \frac{f - Mf}{\sqrt{Df}}$. $Mg = 0$, $Dg = 1$.

$$|f - Mf| \geq c \Leftrightarrow \sqrt{Df} \cdot g \geq c.$$

$$P(|g| \geq \frac{c}{\sqrt{Df}}) \leq \frac{Df}{c^2}.$$

$$P(|g| \geq \frac{c}{\sqrt{Df}}) = P(g^2 \geq \frac{c^2}{Df}) \leq \frac{Mg^2}{(\frac{c^2}{Df})} = \frac{Df}{c^2}.$$

1.6 Закон больших чисел.

f_1, f_2, f_3, \dots — последовательность. $M := Mf_i$, $D = Df_i$.

Одинаковое распределение случайных величин.

$$S_n := \frac{f_1 + f_2 + \dots + f_n}{n} \Rightarrow \forall \varepsilon > 0 \lim_{n \rightarrow \infty} P(|S_n - M| \geq \varepsilon) = 0.$$

$$MS_n = M; DS_n = \frac{D}{n}.$$

$$P(|S_n - M| \geq \varepsilon) \leq \frac{D(S_n)}{\varepsilon^2} = \frac{D}{n\varepsilon^2} \rightarrow 0. n \rightarrow \infty, \varepsilon — \text{фикс.}$$

1.7 Ковариация и корреляция.

Опр. $cov(f; g) = M(f - Mf)(g - Mg)$. $cov(f; f) = Df = Mf^2 - M^2f$. $cov(f, g) = M(fg) - Mf \cdot Mg$.

Главное свойство ковариации: f, g — независимы $\Rightarrow cov(f, g) = 0$.

Опр. $r(f, g) = \frac{cov(f, g)}{\sqrt{Df \cdot Dg}}$ — коэффициент корреляции.

Теорема.

$$-1 \leq r(f, g) \leq 1.$$

$$r(f, g) = 1 \Leftrightarrow \frac{f}{g} = const > 0.$$

$$r(f, g) = -1 \Leftrightarrow \frac{f}{g} = const < 0.$$

1.8 Применение к комбинаторике.

Опр. Перестановки. $f: \{1 \dots n\} \rightarrow \{1 \dots n\}$; биекция. Таких $n!$. Неподвижная точка — $f(x) = x$.

1.9 Теорема Рамсея.

Задача. Есть полный граф K_N , ребра которого покрашены в 2 цвета. Всегда ли можно выбрать либо K_m с ребрами 1 цвета, либо K_n с ребрами 2 цвета?

Теорема. $\forall m; n \exists N \gg 0$: утверждение верно. Наименьшее такое N — число Рамсея $R(m; n)$.

Например $R(3, 3) = 6$. Но вот $R(5, 5)$ уже не известно.

Два аспекта:

1. Верхняя оценка = теорема Рамсея.

2. Нижняя оценка.

Теорема (все та же). $R(m; n) \leq R(m-1; n) + R(m; n-1)$. $R(m; n) \leq 2^{m+n}$.

$R(m; n) \leq C_{m+n}^m$ — более точная оценка.

Теорема Эрдеша. $R(n; n) \geq 2^{\frac{n}{2}}$.

Вероятностный метод. Суть: рассмотрим случайный (со случайной раскраской ребер) граф на $2^{\frac{n}{2}}$ вершин. f —

количество одноцветных клик K_n . Если $M(f) < 1 \Rightarrow \text{Ок.}$

Задача. Пьяница стоит на краю обрыва. С вероятностью p — шаг к обрыву, $q = 1 - p$ — от обрыва.

$P(\text{упадет ровно после } 2k + 1 \text{ шага}) = p^{k+1} \cdot q^k \cdot c_k.$

c_k имеет рекуррентную формулу: $c_{n+1} = c_0 \cdot c_n + c_1 \cdot c_{n-1} + \dots + c_{n-1} \cdot c_1 + c_n \cdot c_0$; $c_0 = 1$.

$P(\text{упадет}) = \frac{1 - \sqrt{1 - 4pq}}{2q}.$