Теория вероятностей (Interview Preparation Kit)

1 Теория вероятностей

1.1 Основные понятия теории вероятностей

Множество элементарных исходов – множество $\Omega = \{\omega_1, \omega_2, ...\}$.

σ-алгебра событий – система $\mathcal{F}$ подмножеств множества Ω, обладающая следующими свойствами:

- $\Omega \in \mathcal{F}, \varnothing \in \mathcal{F}$
- $A_1, ..., A_n, ... \in \mathcal{F} \Longrightarrow \bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{F}, \bigcap\limits_{i=1}^{\infty} A_i \in \mathcal{F}$
- $A, B \in \mathcal{F} \Longrightarrow A \setminus B \in \mathcal{F}$

Вероятностная мера – отображение $P : \mathcal{F} \to \mathbb{R}$, обладающее следующими свойствами:

- $P(A) \geq 0, \forall A \in \mathcal{F}$
- $P(\Omega) = 1$
- $A_1, ..., A_n, ... \in \mathcal{F}$ и $A_i \cap A_j = \varnothing \Longrightarrow P(A_1 + ... + A_n + ...) = P(A_1) + ... + P(A_n) + ...$

Вероятностное пространство – тройка $(\Omega, \mathcal{F}, P)$.

Классическая вероятностная модель подразумевает выбор такого конечного вероятностного пространства, в котором все элементарные исходы равновероятны. Вероятность события A в этом случае может быть посчитана по формуле $P(A) = \frac{|A|}{|\Omega|}$.

Случайная величина $\xi = \xi(\omega)$ – отображение $\xi : \Omega \to \mathbb{R}$, при котором $\forall x \in \mathbb{R}$ $\{\omega : \xi(\omega) \leq x\} \in \mathcal{F}$.

Функция распределения случайной величины ξ – функция $F_\xi : \mathbb{R} \to \mathbb{R}$, $F_\xi(x) = P\{\omega : \xi(\omega) \le x\}$.

Совместная функция распределения случайных величин $\xi_1, ..., \xi_n$ – функция

$$F_{\xi_1 \dots \xi_n}(x_1, ..., x_n) = P\{\xi_1 \le x_1, ..., \xi_n \le x_n\}.$$

1.2 Условные вероятности

Условная вероятность события A относительно события B – величина $P(A|B) = \frac{P(AB)}{P(B)}$.

Разбиение – такой набор событий $B_1, ..., B_n$, что $B_i \cap B_j = \varnothing$ при $i \ne j$ и $B_1 + ... + B_n = \Omega$.

Формула полной вероятности: $P(A) = P(A|B_1)P(B_1) + ... + P(A|B_n)P(B_n)$, где $B_1, ..., B_n$ – разбиение.

Формула Байеса: $P(B_k|A) = \frac{P(A|B_k)P(B_k)}{\sum\limits_{i=1}^{n} P(A|B_i)P(B_i)}$, где $B_1, ..., B_n$ – разбиение.

1.3 Математическое ожидание, дисперсия, корреляция

Математическое ожидание [дискретный случай] случайной величины ξ – число $E\xi = \sum\limits_{\omega \in \Omega} \xi(\omega)P(\omega)$. Математическое ожидание существует в том и только в том случае, когда этот ряд сходится абсолютно.

Математическое ожидание может быть вычислено по формуле $E\xi = \sum\limits_k x_k P\{\xi = x_k\}$.

$Ef(\xi) = \sum\limits_{\omega \in \Omega} f(\xi(\omega))P(\omega) = \sum\limits_k f(x_k)P\{\xi = x_k\}$, где $f : \mathbb{R} \to \mathbb{R}$.

Непрерывная случайная величина – случайная величина, функция распределения которой непрерывна.

Абсолютно непрерывная случайная величина – случайная величина, для которой существует функция $p_\xi(x)$ такая, что $p_\xi(x) \ge 0$, $\int\limits_{-\infty}^{\infty} p_\xi(x)dx = 1$, $\int\limits_{-\infty}^{t} p_\xi(x)dx = F_\xi(t)$. В этом случае функция $p_\xi(x)$ – **плотность распределения** случайной величины ξ.

Абсолютно непрерывное распределение случайных величин $\xi_1, ..., \xi_n$ – распределение, для которого существует функция $p_{\xi_1 \dots \xi_n}(x_1, ..., x_n)$ такая, что

- $p_{\xi_1 \dots \xi_n}(x_1, ..., x_n) \ge 0$,

- $\int\limits_{-\infty}^{\infty} ... \int\limits_{-\infty}^{\infty} p_{\xi_1 ... \xi_n}(x_1, ..., x_n) dx_1 ... dx_n = 1$,
- $\int\limits_{-\infty}^{x_n} ... \int\limits_{-\infty}^{x_1} p_{\xi_1 ... \xi_n}(y_1, ..., y_n) dy_1 ... dy_n = F_{\xi_1 ... \xi_n}(x_1, ..., x_n)$ для любых $x_1, ..., x_n \in \mathbb{R}$.

В этом случае функция $p_{\xi_1 ... \xi_n}(x_1, ..., x_n)$ – **совместная плотность распределения** набора случайных величин $\xi_1, ..., \xi_n$.

Математическое ожидание [абсолютно непрерывный случай] случайной величины ξ с плотностью $p_\xi(x)$ – число $E\xi = \int\limits_{-\infty}^{\infty} x p_\xi(x) dx$. Математическое ожидание существует тогда и только тогда, когда интеграл сходится абсолютно.

$E\xi^2 = \int\limits_{-\infty}^{\infty} x^2 p_\xi(x) dx$.

Свойства математического ожидания:

- Если для некоторой константы $C \in \mathbb{R}$ имеет место $P\{\omega : \xi(\omega) = C\} = 1$, то $E\xi = C$.
- $E(c\xi) = cE\xi$ для любого $c \in \mathbb{R}$.
- Если $E\xi_1, ..., E\xi_n$ существуют, то $E(\xi_1 + ... + \xi_n) = E\xi_1 + ... + E\xi_n$.
- Если $E\xi$ и $E\eta$ существуют и $\xi(\omega) \leq \eta(\omega)$ для всех $\omega \in \Omega$, то $E\xi \leq E\eta$.
- $E\xi = E[E(\xi|\eta)]$.
- $E\eta = \sum\limits_{n \geq 1} P\{\eta \geq n\}$, где η принимает неотрицательные целочисленные значения.

Дисперсия случайной величины ξ – число $D\xi = E(\xi - E\xi)^2$.

$D\xi = E\xi^2 - (E\xi)^2$.

Свойства дисперсии:

- $D(\xi + c) = D\xi$ для любого $c \in \mathbb{R}$.
- $D(c\xi) = c^2 D\xi$ для любого $c \in \mathbb{R}$.
- $D(\xi_1 + ... + \xi_n) = D\xi_1 + ... + D\xi_n + 2 \sum\limits_{i<j} Cov(\xi_i, \xi_j)$.

Ковариация двух случайных величин – число $Cov(\xi, \eta) = E(\xi - E\xi)(\eta - E\eta)$.

Свойства ковариации:

- $Cov(\xi, \xi) = D\xi$.
- $Cov(c_1\xi_1 + c_2\xi_2, \eta) = c_1 Cov(\xi_1, \eta) + c_2 Cov(\xi_2, \eta)$, $\forall c_1, c_2 \in \mathbb{R}$.
- $Cov(\xi, \eta) = E\xi\eta - E\xi E\eta$.

Независимые случайные величины $\xi_1, ..., \xi_n$: для всех $x_1, ..., x_n \in \mathbb{R}$ имеет место $P\{\xi_1 = x_1, ..., \xi_n = x_n\} = P\{\xi_1 = x_1\} \cdot ... \cdot P\{\xi_n = x_n\}$.

Если ξ и η независимы, то $Cov(\xi, \eta) = 0$.

Независимые случайные величины $\xi_1, ..., \xi_n$: для всех $x_1, ..., x_n \in \mathbb{R}$ имеет место $F_{\xi_1 ... \xi_n}(x_1, ..., x_n) = F_{\xi_1}(x_1) ... F_{\xi_n}(x_n)$.

Некоррелированные случайные величины ξ и η: $Cov(\xi, \eta) = 0$.

Из независимости случайных величин следует их некоррелированность. Обратное неверно.

Корреляция двух случайных величин – коэффициент $\rho(\xi, \eta) = \frac{Cov(\xi, \eta)}{\sqrt{D\xi}\sqrt{D\eta}}$.

Свойства коэффициента корреляции:

- $|\rho(\xi, \eta)| \leq 1$.
- Если $|\rho(\xi, \eta)| = 1$, то $\exists\ c_1, c_2 \in \mathbb{R}, c_1^2 + c_2^2 > 0 : c_1\xi + c_2\eta = 0$.

1.4 Независимость событий

Независимые в совокупности события $A_1, ..., A_n$: для всех k и для любых $1 \leq i_1 < ... < i_k \leq n$ верно $P(A_{i_1} ... A_{i_k}) = P(A_{i_1}) ... P(A_{i_k})$.

Попарно независимые события $A_1, ..., A_n$: $\forall i \neq j$ верно $P(A_i A_j) = P(A_i) P(A_j)$.

Из независимости событий следует их попарная независимость. Обратное неверно.

1.5 Основные теоремы теории вероятностей

Первое неравенство Чебышева. Если $\xi \geq 0$, то $P\{\xi \geq \varepsilon\} \leq \frac{E\xi}{\varepsilon}$.

Второе неравенство Чебышева. $P\{|\eta - E\eta| > \varepsilon\} \leq \frac{D\eta}{\varepsilon^2}$.

Выпуклая вниз функция g, определенная на (a, b): для любых точек $x_1, ..., x_n$ и вероятностей $p_1, ..., p_n$ ($p_i \geq 0$, $p_1 + ... + p_n = 1$) выполняется неравенство $g(\sum\limits_j p_j x_j) \leq \sum\limits_j p_j g(x_j)$.

Неравенство Йенсена. g – выпуклая вниз функция, существуют $E\xi$ и $Eg(\xi)$. Тогда $Eg(\xi) \geq g(E\xi)$.

Закон больших чисел [Чебышев]. $\xi_1, ..., \xi_n, ...$ – последовательность независимых с.в. и $\forall\ i\ D\xi_i \leq C$. Тогда $\forall\ \varepsilon > 0\ P\{|\frac{\xi_1+...+\xi_n}{n} - \frac{E\xi_1+...+E\xi_n}{n}| > \varepsilon\} \to 0$ при $n \to \infty$.

Закон больших чисел [Хинчин]. $\xi_1, ..., \xi_n, ...$ – последовательность независимых одинаково распределенных с.в. Существует $E\xi_i = a$. Тогда $\forall\ \varepsilon > 0\ P\{|\frac{\xi_1+...+\xi_n}{n} - a| > \varepsilon\} \to 0$ при $n \to \infty$.

Центральная предельная теорема. $\xi_1, ..., \xi_n, ...$ – последовательность независимых одинаково распределенных с.в. с конечной дисперсией. Обозначим $E\xi_i = a$, $D\xi_i = \sigma^2 > 0$, $S_n = \xi_1 + ... + \xi_n$. Тогда $\forall x \in \mathbb{R}\ P\{\frac{S_n - ES_n}{\sqrt{DS_n}} \leq x\} \to \Phi(x)$ при $n \to \infty$, где $\Phi(x) = (2\pi)^{-1/2} \int_{-\infty}^{x} e^{-y^2/2} dy$.

1.6 Распределения

Производящая функция распределения – функция $\varphi_\xi(s) = Es^\xi$.

$E\xi = \varphi'_\xi(s)\big|_{s=1}, D\xi = (\varphi''_\xi(s) + \varphi'_\xi(s) - (\varphi'_\xi(s))^2)\big|_{s=1}$.

Если ξ_1 и ξ_2 – независимые с.в., тогда $\varphi_{\xi_1+\xi_2}(s) = \varphi_{\xi_1}(s)\varphi_{\xi_2}(s)$.

Производящая функция моментов – функция $M_\xi(t) = \varphi_\xi(e^t)$.

$E\xi = M'_\xi(t)\big|_{t=0}, D\xi = (M''_\xi(t) - (M'_\xi(t))^2)\big|_{t=0}$.

Характеристическая функция – функция $\psi_\xi(t) = Ee^{it\xi}$.

$E\xi = \frac{1}{i} \cdot \psi'_\xi(t)\big|_{t=0}, D\xi = (-\psi''_\xi(t) + (\psi'_\xi(t))^2)\big|_{t=0}$.

Если ξ_1 и ξ_2 – независимые с.в., тогда $\psi_{\xi_1+\xi_2}(t) = \psi_{\xi_1}(t)\psi_{\xi_2}(t)$.

Коэффициент эксцесса – величина $\gamma_2 = \frac{E(\xi - E\xi)^4}{(D\xi)^2} - 3$.

Геометрическое распределение – число испытаний Бернулли до наблюдения первого успеха.

$$P\{\xi = k\} = (1-p)^k p, k = 0, 1, 2, ...$$

$E\xi = \frac{1-p}{p}, D\xi = \frac{1-p}{p^2}$.

Отсутствие памяти. Количество прошлых "неудач"не влияет на количество будущих "неудач".

Биномиальное распределение – число успехов в последовательности n независимых испытаний Бернулли. $\nu_n(\omega) = \sum\limits_{i=1}^{n} \theta_i$, где $\omega = (\theta_1, ..., \theta n)$.

$$P\{\nu_n = k\} = C_n^k p^k (1-p)^{n-k}, k = 0, 1, ..., n.$$

$E\nu_n = np, D\nu_n = np(1-p)$. $\psi_{\nu_n}(t) = (pe^{it} + 1 - p)^n$.

Пуассоновское распределение с параметром $\lambda > 0$ – число событий, произошедших за фиксированное время (λ – среднее количество событий за фиксированный промежуток времени), счётчик.

$$P\{\Pi = k\} = \frac{\lambda^k}{k!} e^{-\lambda}, k = 0, 1, ...$$

$E\Pi = \lambda, D\Pi = \lambda$.

Равномерное распределение в отрезке $[c; d]$.

$$p_\xi(x) = \begin{cases} \frac{1}{d-c}, & x \in [c; d], \\ 0, & x \notin [c; d]. \end{cases}$$

$E\xi = \frac{c+d}{2}, D\xi = \frac{(d-c)^2}{12}$.

Нормальное распределение – $N(a, \sigma^2), a \in \mathbb{R}, \sigma > 0$.

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right).$$

$E\xi = a, D\xi = \sigma^2$.

Стандартное нормальное распределение – $N(0, 1)$.

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

$E\xi = 0, D\xi = 1$.

Показательное распределение с параметром $\lambda > 0$ – время между двумя последовательными совершениями одного и того же события ($\frac{1}{\lambda}$ – среднее время ожидания нового события).

$$p_\xi(x) = \begin{cases} 0, & x < 0, \\ \lambda e^{-\lambda x}, & x \geq 0. \end{cases}$$

$E\xi = \frac{1}{\lambda}, D\xi = \frac{1}{\lambda^2}$. $M_\xi(t) = (1 - \frac{t}{\lambda})^{-1}, E\xi^n = \frac{n!}{\lambda^n}$.

Отсутствие памяти. Пусть автобусы приходят на остановку случайно, но с некоторой фиксированной средней интенсивностью. Тогда количество времени, уже затраченное пассажиром на ожидание автобуса, не влияет на время, которое ему ещё придётся прождать.

1.7 Разное

Преобразования случайных величин. Пусть $\xi : \Omega \to \mathbb{R}$ – абсолютно непрерывная с.в., $g(x) : \mathbb{R} \to \mathbb{R}$ – непрерывная и монотонная функция. Тогда $\eta = g(\xi)$ – абсолютно непрерывна и $p_\eta(y) = p_\xi(g^{-1}(y))\left|\frac{dg^{-1}(y)}{dy}\right|$.

Формула свертки. Пусть ξ_1 и ξ_2 абсолютно непрерывны с плотностями $p_{\xi_1}(x)$ и $p_{\xi_2}(x)$ и независимы. Тогда $p_{\xi_1+\xi_2}(x) = \int\limits_{-\infty}^{\infty} p_{\xi_1}(y)p_{\xi_2}(x-y)dy$.

Формула Стирлинга. $n! = \sqrt{2\pi n}n^n e^{-n+\theta(n)}$, где $\frac{1}{12n+1} < \theta(n) < \frac{1}{12n}$.