

# Теория вероятностей и математическая статистика

## Полная вероятность. Формула Байеса. Случайные величины.

Глеб Карпов

ВШБ Бизнес-информатика

## Полная вероятность

- Концепция условной вероятности неразрывно связана со следующей, концепцией полной вероятности.

$\Omega$

$B_1$

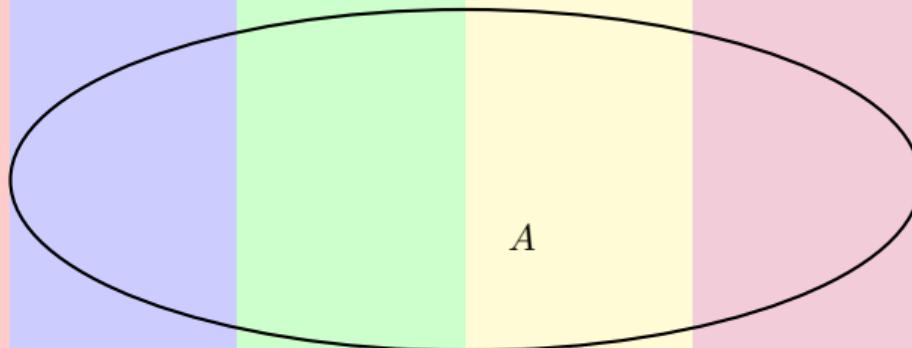
$B_2$

...

$B_{n-1}$

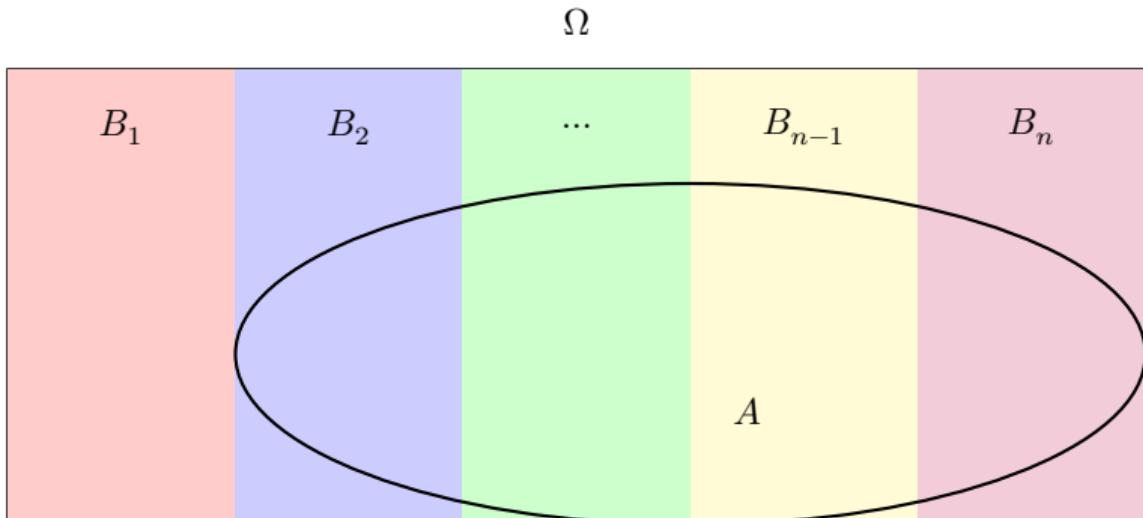
$B_n$

$A$



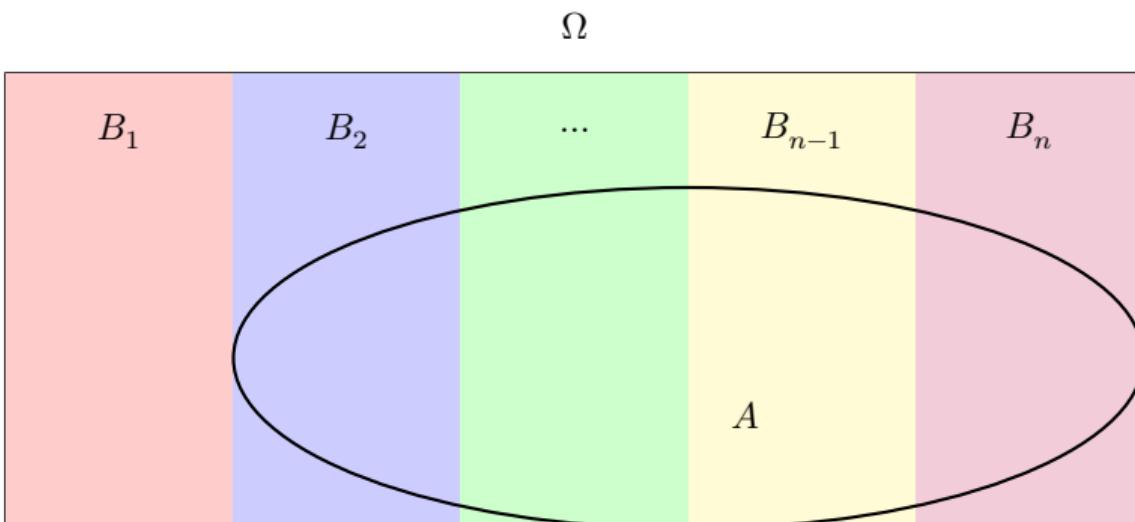
## Полная вероятность

- Концепция условной вероятности неразрывно связана со следующей, концепцией полной вероятности.
- Рассмотрим зафиксированное пространство  $(\Omega, \mathcal{F}, P)$ . Назовем **разбиением**  $\Omega$  коллекцию событий  $\{B_k, k \in I\}$ , таких что  $B_i \cap B_j = \emptyset$  при  $i \neq j$  и  $\bigcup B_i = \Omega$ .



## Полная вероятность

- Концепция условной вероятности неразрывно связана со следующей, концепцией полной вероятности.
- Рассмотрим зафиксированное пространство  $(\Omega, \mathcal{F}, P)$ . Назовем **разбиением**  $\Omega$  коллекцию событий  $\{B_k, k \in I\}$ , таких что  $B_i \cap B_j = \emptyset$  при  $i \neq j$  и  $\bigcup B_i = \Omega$ .
- Вдобавок, рассмотрим какое-то другое событие  $A$ , которое пересекается с какими-то событиями из разбиения, но не обязано пересекаться со всеми.



## Полная вероятность

Если  $\{B_1, B_2, \dots\}$  - разбиение  $\Omega$ , с  $P(B_i) > 0 \forall i$ , то:

$$P(A) = \sum_i P(A|B_i)P(B_i), \forall A \in \mathcal{F}$$

**Доказательство.** Заметим, что мы можем реконструировать событие  $A$  из его частичек-пересечений со всеми  $B_i$ :  $A = \bigcup_i (A \cap B_i)$ . Эти кусочки  $\{A \cap B_i\}$  попарно не пересекаются, как и оригинальные элементы разбиения.

Поэтому далее можем применить свойство аддитивности вероятности:

$$\begin{aligned} P(A) &= P\left(\bigcup_i (A \cap B_i)\right) \\ &= \sum_i P(A \cap B_i) = \sum_i P(A|B_i)P(B_i) \end{aligned}$$

## Теорема Байеса

Пусть  $\{B_1, B_2, \dots\}$  - разбиение  $\Omega$ , с  $P(B_i) > 0 \forall i$ . Тогда для любого события  $A$ ,  $P(A) > 0$ , справедливо:

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_i P(A|B_i)P(B_i)}$$

### Доказательство.

Опираемся на то, что  $P(X \cap Y)$  может быть выражена через разные условные вероятности:

$$P(X \cap Y) = P(X|Y)P(Y) = P(Y|X)P(X).$$

Применим это к условию

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{P(A)},$$

где  $P(A) = \sum_i P(A|B_i)P(B_i)$  как полная вероятность.

## Пример 1

Вероятность победы лошади в скачках равна 0.3, если погода сухая, и 0.5 если идет дождь. Прогноз погоды дает вероятность дождя 40%.

1. Найдите вероятность победы лошади.
2. Если вы знаете, что лошадь проиграла гонку, какова вероятность того, что погода была сухой в день скачек? (Предположим, что вы не помните!)

## Пример 1

Вероятность победы лошади в скачках равна 0.3, если погода сухая, и 0.5 если идет дождь. Прогноз погоды дает вероятность дождя 40%.

1. Найдите вероятность победы лошади.
  2. Если вы знаете, что лошадь проиграла гонку, какова вероятность того, что погода была сухой в день скачек? (Предположим, что вы не помните!)
- $P(V) = P(V|D)P(D) + P(V|W)P(W) = 0.3 \cdot 0.6 + 0.5 \cdot 0.4 = 0.38$

## Пример 1

Вероятность победы лошади в скачках равна 0.3, если погода сухая, и 0.5 если идет дождь. Прогноз погоды дает вероятность дождя 40%.

1. Найдите вероятность победы лошади.
  2. Если вы знаете, что лошадь проиграла гонку, какова вероятность того, что погода была сухой в день скачек? (Предположим, что вы не помните!)
- $P(V) = P(V|D)P(D) + P(V|W)P(W) = 0.3 \cdot 0.6 + 0.5 \cdot 0.4 = 0.38$
  - $$P(D|\bar{V}) = \frac{P(\bar{V}|D)P(D)}{P(\bar{V})} = \frac{0.7 \cdot 0.6}{1 - 0.38} \approx 0.677$$

## Пример 2 (False positives)

- Редкое, но опасное заболевание имеет распространенность  $1 \text{ на } 10^5$  в общей популяции. Существует диагностический тест, но он несовершенен. Если у человека есть заболевание, тест положительный с вероятностью 0.9; если нет - тест положительный с вероятностью  $\frac{1}{20}$ . Тест дал положительный результат. Какова вероятность того, что у человека **действительно** есть заболевание?

## Пример 2 (False positives)

- Редкое, но опасное заболевание имеет распространенность  $1 \text{ на } 10^5$  в общей популяции. Существует диагностический тест, но он несовершенен. Если у человека есть заболевание, тест положительный с вероятностью 0.9; если нет - тест положительный с вероятностью  $\frac{1}{20}$ . Тест дал положительный результат. Какова вероятность того, что у человека **действительно** есть заболевание?
- Обозначим вероятности: ( $H = \text{healthy}$ ,  $I = \text{ill}$ ,  $+$  for detected positive,  $-$  for detected negative)  $P(I) = 10^{-5}$ ,  $P(+|I) = 0.9$ ,  $P(+|H) = \frac{1}{20}$ . Мы заинтересованы в  $P(I|+)$ .

## Пример 2 (False positives)

- Редкое, но опасное заболевание имеет распространенность  $1 \text{ на } 10^5$  в общей популяции. Существует диагностический тест, но он несовершенен. Если у человека есть заболевание, тест положительный с вероятностью 0.9; если нет - тест положительный с вероятностью  $\frac{1}{20}$ . Тест дал положительный результат. Какова вероятность того, что у человека **действительно** есть заболевание?
- Обозначим вероятности: ( $H = \text{healthy}$ ,  $I = \text{ill}$ ,  $+$  for detected positive,  $-$  for detected negative)  $P(I) = 10^{-5}$ ,  $P(+|I) = 0.9$ ,  $P(+|H) = \frac{1}{20}$ . Мы заинтересованы в  $P(I|+)$ .
- $P(+) = P(+|I)P(I) + P(+|H)P(H) = 0.9 \cdot 10^{-5} + 0.05 \cdot (1 - 10^{-5}) = 0.0500085$

## Пример 2 (False positives)

- Редкое, но опасное заболевание имеет распространенность  $1 \text{ на } 10^5$  в общей популяции. Существует диагностический тест, но он несовершенен. Если у человека есть заболевание, тест положительный с вероятностью 0.9; если нет - тест положительный с вероятностью  $\frac{1}{20}$ . Тест дал положительный результат. Какова вероятность того, что у человека **действительно** есть заболевание?
- Обозначим вероятности: ( $H = \text{healthy}$ ,  $I = \text{ill}$ ,  $+$  for detected positive,  $-$  for detected negative)  $P(I) = 10^{-5}$ ,  $P(+|I) = 0.9$ ,  $P(+|H) = \frac{1}{20}$ . Мы заинтересованы в  $P(I|+)$ .
- $P(+) = P(+|I)P(I) + P(+|H)P(H) = 0.9 \cdot 10^{-5} + 0.05 \cdot (1 - 10^{-5}) = 0.0500085$
- 

$$P(I|+) = \frac{P(+|I)P(I)}{P(+)} = \frac{0.9 \cdot 10^{-5}}{0.0500085} \approx 0.0002$$

## Случайные величины: Мотивация

- Элементарный исход в общем может иметь различный характер: цвет, буква, звук и т.д. Но для более углубленного и полного анализа случайного эксперимента лучше работать с числами.

## Случайные величины: Мотивация

- Элементарный исход в общем может иметь различный характер: цвет, буква, звук и т.д. Но для более углубленного и полного анализа случайного эксперимента лучше работать с числами.
- Можем ли мы также прикрепить число к каждому элементарному исходу независимо от его природы, а затем оперировать **только** в пространстве чисел? Спойлер: это даст нам новые инструменты, которые недоступны, если мы работаем с исходами разной природы.

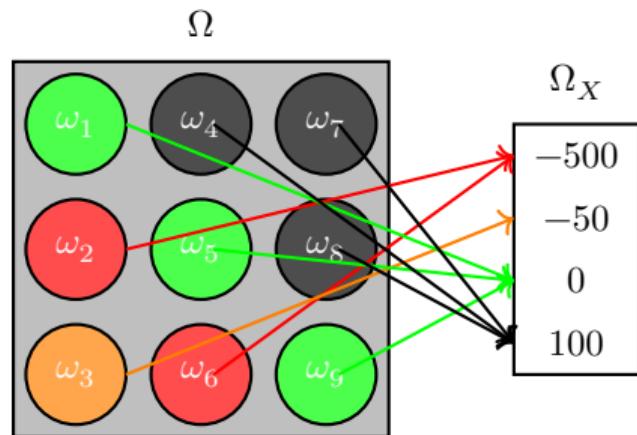
## Случайные величины: Мотивация

- Элементарный исход в общем может иметь различный характер: цвет, буква, звук и т.д. Но для более углубленного и полного анализа случайного эксперимента лучше работать с числами.
- Можем ли мы также прикрепить число к каждому элементарному исходу независимо от его природы, а затем оперировать **только** в пространстве чисел? Спойлер: это даст нам новые инструменты, которые недоступны, если мы работаем с исходами разной природы.
- Пример: игральная кость с 6 гранями, 2 грани желтые, 2 грани зеленые, 2 грани черные, без чисел - мы можем искусственно прикрепить числа 1, 2, 3 к каждому цвету.

## Случайные величины: Формальное определение

Результирующее 'прикрепленное' число мы называем случайной величиной.

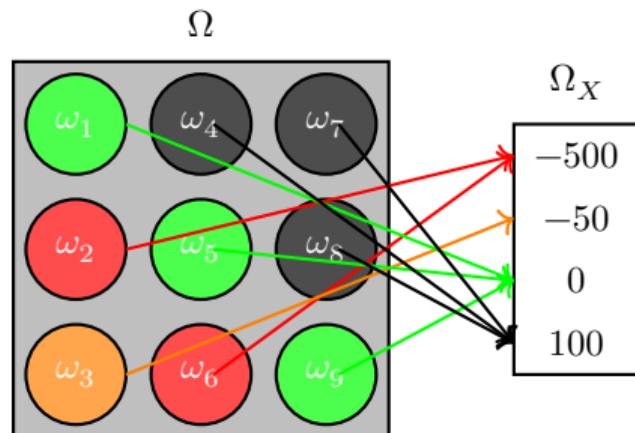
- Дискретная случайная величина  $X$  в пространстве вероятностей  $(\Omega, \mathcal{F}, P)$  является отображением  $X : \Omega \rightarrow \mathbb{R}$ , такое что:



## Случайные величины: Формальное определение

Результирующее 'прикрепленное' число мы называем случайной величиной.

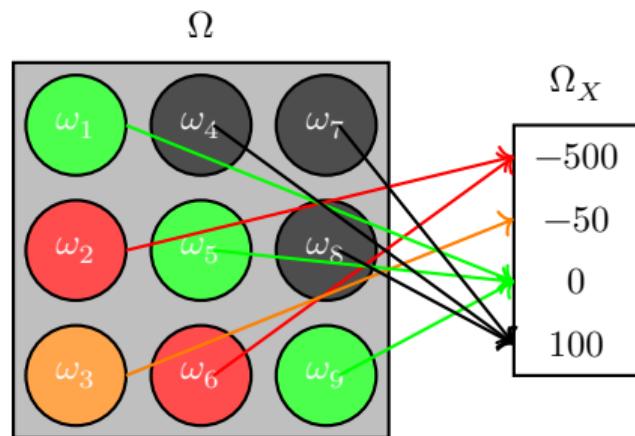
- Дискретная случайная величина  $X$  в пространстве вероятностей  $(\Omega, \mathcal{F}, P)$  является отображением  $X : \Omega \rightarrow \mathbb{R}$ , такое что:
1. Образ  $X(\Omega) = \Omega_X = \text{Im } X$  (все возможные результаты) является счетным подмножеством  $\mathbb{R}$ ,



## Случайные величины: Формальное определение

Результирующее 'прикрепленное' число мы называем случайной величиной.

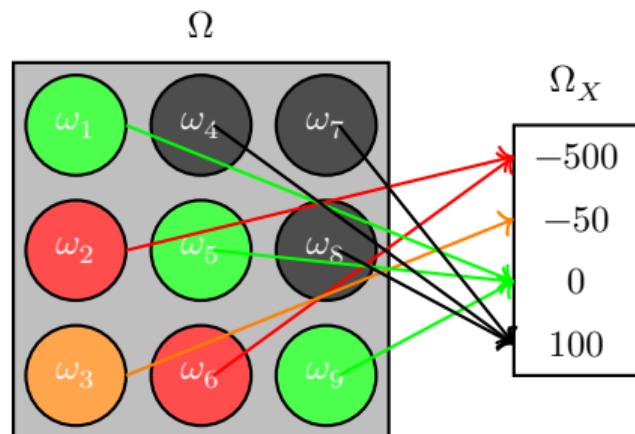
- Дискретная случайная величина  $X$  в пространстве вероятностей  $(\Omega, \mathcal{F}, P)$  является отображением  $X : \Omega \rightarrow \mathbb{R}$ , такое что:
  1. Образ  $X(\Omega) = \Omega_X = Im X$  (все возможные результаты) является счетным подмножеством  $\mathbb{R}$ ,
  2.  $\{\omega \in \Omega : X(\omega) = x\} \in \mathcal{F}$ , для  $x \in \mathbb{R}$ .



## Случайные величины: Формальное определение

Результирующее 'прикрепленное' число мы называем случайной величиной.

- Дискретная случайная величина  $X$  в пространстве вероятностей  $(\Omega, \mathcal{F}, P)$  является отображением  $X : \Omega \rightarrow \mathbb{R}$ , такое что:
  1. Образ  $X(\Omega) = \Omega_X = \text{Im } X$  (все возможные результаты) является счетным подмножеством  $\mathbb{R}$ ,
  2.  $\{\omega \in \Omega : X(\omega) = x\} \in \mathcal{F}$ , для  $x \in \mathbb{R}$ .
- Для упрощения мы сокращаем события вида  $\{\omega \in \Omega : X(\omega) = x\}$  до более простого вида  $\{X = x\}$ .



## Вероятность случайной величины

- Основной оставшийся вопрос - как находить вероятности различных значений С.В.

## Вероятность случайной величины

- Основной оставшийся вопрос - как находить вероятности различных значений С.В.
- Мы знаем, что каждое уникальное значение С.В. привязано к какому-то событию из  $\mathcal{F}$ . Тогда вероятность этого события и будет вероятностью получить данное значение С.В.

## Вероятность случайной величины

- Основной оставшийся вопрос - как находить вероятности различных значений С.В.
- Мы знаем, что каждое уникальное значение С.В. привязано к какому-то событию из  $\mathcal{F}$ . Тогда вероятность этого события и будет вероятностью получить данное значение С.В.
- Определение. Функция вероятности дискретной случайной переменной  $X$  - это функция  $p_X(x) : \mathbb{R} \rightarrow [0, 1]$ , заданная как:

$$p_X(x) = P(X = x) = P(A), \text{ где } A = \{\omega \in \Omega : X(\omega) = x\}$$

## Вероятность случайной величины

- Основной оставшийся вопрос - как находить вероятности различных значений С.В.
- Мы знаем, что каждое уникальное значение С.В. привязано к какому-то событию из  $\mathcal{F}$ . Тогда вероятность этого события и будет вероятностью получить данное значение С.В.
- Определение. Функция вероятности дискретной случайной переменной  $X$  - это функция  $p_X(x) : \mathbb{R} \rightarrow [0, 1]$ , заданная как:

$$p_X(x) = P(X = x) = P(A), \text{ где } A = \{\omega \in \Omega : X(\omega) = x\}$$

- Свойства:

## Вероятность случайной величины

- Основной оставшийся вопрос - как находить вероятности различных значений С.В.
- Мы знаем, что каждое уникальное значение С.В. привязано к какому-то событию из  $\mathcal{F}$ . Тогда вероятность этого события и будет вероятностью получить данное значение С.В.
- Определение. Функция вероятности дискретной случайной переменной  $X$  - это функция  $p_X(x) : \mathbb{R} \rightarrow [0, 1]$ , заданная как:

$$p_X(x) = P(X = x) = P(A), \text{ где } A = \{\omega \in \Omega : X(\omega) = x\}$$

- Свойства:

1.  $p_X(x) \geq 0$  для  $\forall x \in \Omega_X$ ,  $p_X(x) = 0$  для  $x \notin \Omega_X$

## Вероятность случайной величины

- Основной оставшийся вопрос - как находить вероятности различных значений С.В.
- Мы знаем, что каждое уникальное значение С.В. привязано к какому-то событию из  $\mathcal{F}$ . Тогда вероятность этого события и будет вероятностью получить данное значение С.В.
- Определение. Функция вероятности дискретной случайной переменной  $X$  - это функция  $p_X(x) : \mathbb{R} \rightarrow [0, 1]$ , заданная как:

$$p_X(x) = P(X = x) = P(A), \text{ где } A = \{\omega \in \Omega : X(\omega) = x\}$$

- Свойства:

1.  $p_X(x) \geq 0$  для  $\forall x \in \Omega_X$ ,  $p_X(x) = 0$  для  $x \notin \Omega_X$
2.  $\sum_{x_i \in \Omega_X} p_X(x_i) = 1$ .

$$\sum_{x_i \in \Omega_X} p_X(x_i) = \sum_{x_i \in \Omega_X} P(\{\omega \in \Omega : X(\omega) = x_i\}) = P\left(\bigcup_{x_i \in \Omega_X} \{\omega \in \Omega : X(\omega) = x_i\}\right) = P(\Omega) = 1.$$

## Случайные величины: Пример

Подбрасываются две монеты. Первая монета выпадает орлом с вероятностью 0.6, вторая с вероятностью 0.7. Предположим, что результаты подбрасываний независимы, и пусть  $X$  равно общему количеству выпавших орлов. Постройте дискретную случайную величину и её функцию вероятности.

## Новое пространство вероятности

- Как и прежде, мы можем комбинировать случайные переменные в события.

## Новое пространство вероятности

- Как и прежде, мы можем комбинировать случайные переменные в события.
- Например,  $A = \{x_1, \dots, x_k\}$ ,  $x_1, \dots, x_k \in \Omega_X$ .

## Новое пространство вероятности

- Как и прежде, мы можем комбинировать случайные переменные в события.
- Например,  $A = \{x_1, \dots, x_k\}$ ,  $x_1, \dots, x_k \in \Omega_X$ .
- Для вычисления  $P(A)$  мы используем принцип аддитивности:

$$P(A) = P(\{X = x_1\} \cup \dots \cup \{X = x_k\}) = \sum_{i=1}^k P(\{X = x_i\}).$$

## Новое пространство вероятности

- Как и прежде, мы можем комбинировать случайные переменные в события.
- Например,  $A = \{x_1, \dots, x_k\}$ ,  $x_1, \dots, x_k \in \Omega_X$ .
- Для вычисления  $P(A)$  мы используем принцип аддитивности:

$$P(A) = P(\{X = x_1\} \cup \dots \cup \{X = x_k\}) = \sum_{i=1}^k P(\{X = x_i\}).$$

- В конечном счете, мы можем забыть о нашем оригинальном  $(\Omega, \mathcal{F}, P)$ , скрыть их ‘под капотом’ нашего случайного эксперимента. Вместо этого мы используем новое пространство  $\Omega_X$  всех возможных значений С.В. и  $p_X(x)$  как новую функцию вероятности.