Дискретные случайные величины

Антипов Денис Сергеевич

Университет ИТМО, Санкт-Петербург, Россия

УНИВЕРСИТЕТ ИТМО

18 февраля 2021 г.

Случайная величина

Случайная величина (с.в.) — функционал, заданный на Ω.

$$X : \Omega \to \mathbb{R}$$

- NB: этот функционал должен быть измерим по используемой мере $\Pr$

Договоримся об обозначениях

- Большая буква (X, Y, Z, etc) — случайная величина (как отображение)
- Маленькая буква (x, y, z, etc) — значение случайной величины (число)

Примеры случайных величин

На конечной Ω

- Бросок кости: X — число на верхней грани
  - $\{$⚀ $\to 1$, ⚁ $\to 2$, ⚂ $\to 3$, ⚃ $\to 4$, ⚄ $\to 5$, ⚅ $\to 6\}$
- Хоккейный матч: X — очки домашней команды
  - $\{\text{В} \to 2,\ \text{ВО} \to 2,\ \text{ВБ} \to 2,\ \text{ПБ} \to 1,\ \text{ПО} \to 1,\ \text{П} \to 0\}$

На счетной Ω

- Бросаем монету до первого орла: X — число бросков

На несчетной Ω

- Бросаем дротик в мишень для дартса: X — очки согласно правилам игры
- Крутим волчок в ЧГК: X — угол, на котором он остановится
  - Y — население города, из которого пришел вопрос
  - Z — вероятность того, что на выбранный вопрос знатоки ответят

Дискретные случайные величины

Дискретными называются те с.в., которые определены на не более, чем счетной Ω

NB: Множеством значений дискретной с.в. может быть любое подмножество $\mathbb{R}$ (не более, чем счетное, разумеется)

NB-2: Никаких проблем с измеримостью дискретной с.в. как функции обычно нет, поэтому этот вопрос мы будем опускать

Функция вероятности

Можно рассматривать $(X = x)$ как событие, так как оно задает какое-то множество исходов из Ω. Тогда у этого события должна быть вероятность.

$$p_X(x) = \Pr(X = x) = \Pr(\{\omega \in \Omega \mid X(\omega) = x\})$$

p_X называется функцией вероятности с.в. X.
Ее свойства:

- $p_X(x) \geq 0$
- $\sum_x p_X(x) = 1$

Примеры случайных величин

Если нам дано какое-то распределение $\mathcal{D}$ и с.в. X ему следует, мы пишем $X \sim \mathcal{D}$

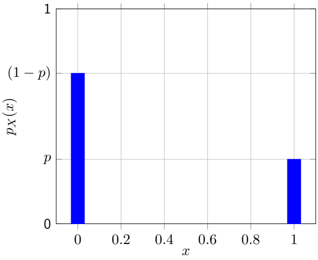
Примеры распределений (законов):

- Бернулли
- Равномерное
- Биномиальное
- Геометрическое
- Гипер-геометрическое
- Степенное

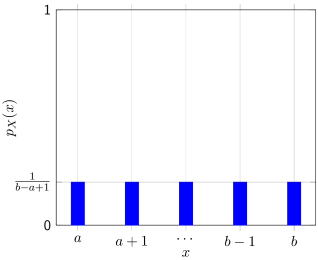
Распределение Бернулли

$$X \sim \mathrm{Bern}(p) \Leftrightarrow \begin{cases} p_X(0) = (1-p), \\ p_X(1) = p, \\ p_X(x) = 0 \text{ для всех остальных } x. \end{cases}$$

- p — параметр распределения, вероятность успеха
- С.в. X может быть индикатором события A:
  - $\omega \in A \Rightarrow X(\omega) = 1$
  - $\omega \notin A \Rightarrow X(\omega) = 0$



Равномерное дискретное распределение



$$X \sim U(a,b) = \begin{cases} p_X(x) = \frac{1}{b-a+1}, & \text{если } x \in [a..b], \\ p_X(x) = 0, & \text{иначе.} \end{cases}$$

- a, b — целочисленные параметры ($a \leq b$)
- $\Omega = [a..b]$
- $X(\omega) = \omega$

Биномиальное распределение

Проведем n независимых опытов по схеме Бернулли с вероятностью успеха p. X — число успехов.

$$X \sim \mathrm{Bin}(n, p) \Leftrightarrow X = \sum_{i=1}^{n} X_i,$$

где все $X_i \sim \mathrm{Bern}(p)$.

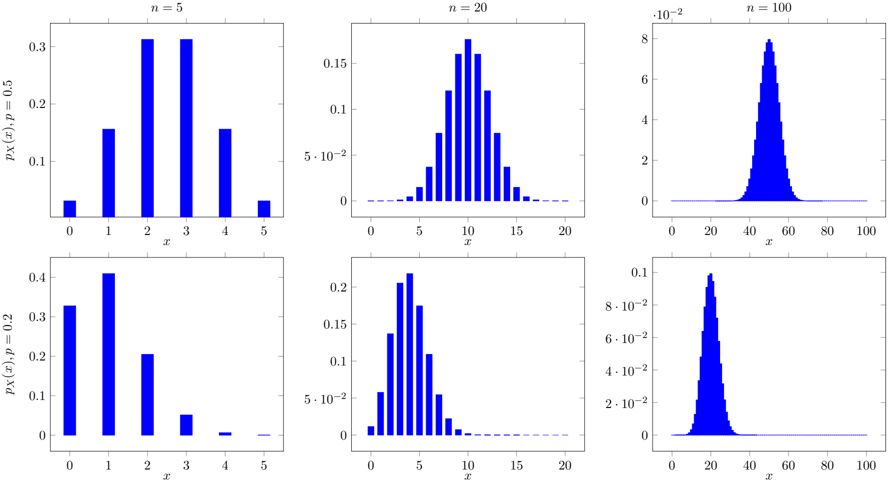
$$\boxed{p_X(x) = \binom{n}{x} p^x (1-p)^{n-x}}$$

- Вероятность конкретного исхода, в котором ровно x успехов:

$$\Pr(\omega) = p^x (1-p)^{n-x}$$

- Всего таких исходов $\binom{n}{x}$

Функции вероятностей биномиального распределения



NB: везде разные масштабы оси OY

Геометрическое распределение

Проводим независимые эксперименты по схеме Бернулли с вероятностью успеха p до первого успешного исхода. X — число экспериментов, которые мы проводим.

$$X \sim Geom(p) \Leftrightarrow X = \min\{i \in \mathbb{N} \mid X_i = 1\},$$

где $X_i \sim \mathrm{Bern}(p)$.

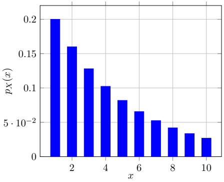
$$p_X(x) = \quad p(1-p)^{x-1}$$

- Вероятность, что первые $x-1$ неуспешны:

$$\Pr(X_1 = X_2 = \dots X_{x-1} = 0) = (1-p)^{x-1}$$

- Вероятность, что x-ый исход успешен:

$$\Pr(X_x = 1) = p$$



Уточнение

Есть исход, при котором $X = +\infty$, но его вероятность равна нулю.

$$\Pr(X = \infty) \leq \Pr(X_1 = \cdots = X_n = 0) = (1 - p)^n$$

для любого n. Если допустить, что $\Pr(X = \infty) = q > 0$, то возьмем $n \geq \log_{1-p} q$, получим, что вероятность подсобытия больше, чем вероятность события (Противоречие со свойствами вероятностной меры).

Математическое ожидание с.в.

Если много раз повторить один и тот же эксперимент, то каков будет средний результат?

- Бросаем честную кость d6. Что в среднем?
- $\frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 + \frac{1}{6} \cdot 5 + \frac{1}{6} \cdot 6 = 3.5$

$$E(X) = \sum_x x \cdot p_X(x)$$

NB: Матожидание определено только когда этот ряд абсолютно сходится

NB-2: Матожидание — функционал на множестве с.в.

Матожидание некоторых с.в.

$X \sim \mathrm{Bern}(p)$

$$E(X) = 0 \cdot (1 - p) + 1 \cdot p = p$$

$X \sim U(a, b)$

$$E(X) = \sum_{i=a}^{b} \frac{i}{b - a + 1} = \frac{a + b}{2}$$

Элементарные свойства матожидания

- Неотрицательность: $X \geq 0 \Rightarrow E(X) \geq 0$
- Ограниченность: $X \in [a, b] \Rightarrow E(X) \in [a, b]$
- Матожидание константы: $E(c) = c$

Функции от с.в. и их матожидание

Пусть есть с.в. X и функция $g : \mathbb{R} \to \mathbb{R}$. Тогда $Y = g(X)$ — тоже с.в.

Как посчитать $E(Y)$?

- по определению: $E(Y) = \sum_y y p_Y(y)$
- $\boxed{E(Y) = \sum_x g(x) p_X(x)}$

$$\sum_x g(x) p_X(x) = \sum_y \sum_{x:g(x)=y} g(x) p_X(x)$$

$$= \sum_y y \sum_{x:g(x)=y} p_X(x) = \sum_y y p_Y(y) = E(Y)$$

Линейность матожидания

$$E(aX + b) = aE(X) + b$$

Позволяет легко пересчитывать матожидание с.в. не по определению.

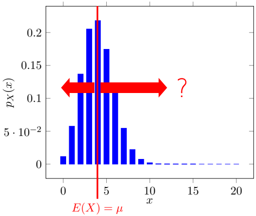
- Пусть X принимает четные значения от 0 до 100 с равной вероятностью.
- $X = 2Y$, где $Y \sim U(0, 50)$
- $E[X] = 2E[Y] = 50$

NB: Для всех линейных функций g выполнено $E(g(X)) = g(E(X))$, но в общем случае это неверно

Дисперсия

Иногда хотим оценить то, насколько с.в. отклоняется от своего среднего значения μ.

- Первая мысль: посчитать $E(X - \mu)$
- $E(X - \mu) = E(X) - \mu = 0$



$$\boxed{\mathrm{Var}(X) = E((X - \mu)^2)}$$

Несколько замечаний

$$\boxed{\mathrm{Var}(X) = E((X-\mu)^2)}$$

- Дисперсия считается как мотожидание функции от с.в.
- Всегда неотрицательна, равна нулю только у детерминированной с.в.
- Почему не посчитать $|X-\mu|$? Просто так удобнее. Чем — узнаем позже
- Часто полезно знать среднеквадратичное отклонение

$$\boxed{\sigma(X) = \sqrt{\mathrm{Var}(X)}}$$

- В дальнейшем часто будем обозначать $\mu := E(X)$, если понятно, про какую с.в. речь

Свойства дисперсии

$$\boxed{\operatorname{Var}(aX+b) = a^2\operatorname{Var}(X)}$$

- Умножая с.в. на a — увеличиваем ее вариацию в a^2 раз
- Прибавляя к с.в. b — не меняем вариацию

$$\begin{aligned}
\operatorname{Var}(aX+b) &= E\left((aX+b-E(aX+b))^2\right)\\
&= E\left((aX+b-aE(X)-b)^2\right)\\
&= E\left((aX-aE(X))^2\right)\\
&= E\left(a^2(X-E(X))^2\right)\\
&= a^2E\left((X-E(X))^2\right)\\
&= a^2\operatorname{Var}(X)
\end{aligned}$$

Свойства дисперсии

$$\boxed{\mathrm{Var}(X) = E(X^2) - (E(X))^2}$$

$$\begin{aligned} \mathrm{Var}(X) &= E\left((X-\mu)^2\right) \\ &= E\left(X^2 - 2X\mu + \mu^2\right) \\ &= E(X^2) - 2\mu^2 + \mu^2 \\ &= E(X^2) - (E(X))^2 \end{aligned}$$

Дисперсия распределения Бернулли

- $X \sim \mathrm{Bern}(p)$

$$\begin{aligned}\mathrm{Var}(X) &= E(X^2) - (E(X))^2 = E(X) - (E(X))^2 \\ &= p - p^2 = p(1-p)\end{aligned}$$

- Если кидаем честную монету, мы всегда отклоняемся от μ на $\frac{1}{2}$

$$\sigma(x) = \sqrt{p(1-p)} = \left[p = \frac{1}{2}\right] = \frac{1}{2}$$

Дисперсия равномерного распределения

- $X \sim U(a, b)$ — можем посчитать вариацию $Y = X - a \sim U(0, n)$, где $n = b - a$.

$$
\begin{aligned}
\mathrm{Var}(X) = \mathrm{Var}(Y + a) = \mathrm{Var}(Y) &= E(Y^2) - (E(Y))^2 \\
&= \sum_{i=0}^{n} \frac{i^2}{n+1} - \left(\frac{n}{2}\right)^2 \\
&= \frac{1}{n+1} \cdot \frac{n(n+1)(2n+1)}{6} - \frac{n^2}{4} \\
&= \frac{4n^2 + 2n - 3n^2}{12} = \frac{n(n+2)}{12} \\
&= \frac{(b-a)(b-a+2)}{12}
\end{aligned}
$$

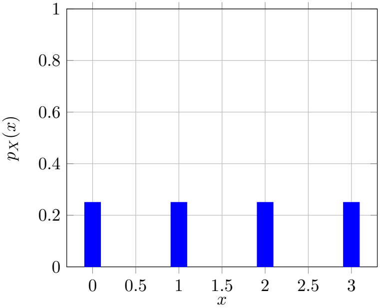
Условная с.в. (на событии)

Ничего особо нового:

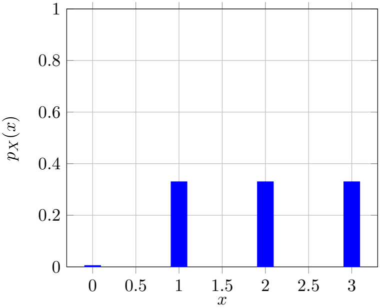
- $p_X(x \mid A) = \Pr(X = x \mid A)$
- Легко проверить свойства функции вероятностей:
  - $p_X(x \mid A) \geq 0$
  - $\sum_x p_X(x \mid A) = 1$
- $E(X \mid A) = \sum_x x \cdot p_X(x \mid A)$
- $E(g(X) \mid A) = \sum_x g(x) \cdot p_X(x \mid A)$

Пример условной с.в.

Безусловное распределение
$X \sim U(0, 3)$:



Условное распределение
$X \mid X \geq 1$:



Теорема о полном матожидании

Слышали о полной вероятности?

$$\Pr(B) = \sum_i \Pr(A_i) \Pr(B \mid (A_i))$$

Но событием B может быть событие $X = x$

$$p_X(x) = \sum_i \Pr(A_i) p_X(x \mid A_i)$$

Домножим левую часть на x и просуммируем по всем x

$$\begin{aligned}\sum_x x \cdot p_X(x) &= \sum_x \sum_i x \Pr(A_i) p_X(x \mid A_i) \\ &= \sum_i \Pr(A_i) \sum_x x \cdot p_X(x \mid A_i) \\ &= \sum_i \Pr(A_i) E(X \mid A_i)\end{aligned}$$

Теорема о полном матожидании

$$E(X) = \sum_i \Pr(A_i) E(X \mid A_i)$$

Беспамятство геометрического распределения

У геометрического распределения нет памяти. Пусть $X \sim \mathrm{Geom}(p)$.

$$p_{(X-1)}(x \mid X > 1) = \frac{\Pr(X = x + 1 \cap X > 1)}{\Pr(X > 1)} = \frac{\Pr(X = x + 1)}{1 - p}$$
$$= \frac{p(1-p)^x}{1-p} = p(1-p)^{x-1} = p_X(x)$$

Можно также показать, что $p_{X-n}(x \mid X > n) = p_X(x)$

При условии, что первые n исходов — неудачные, число оставшихся бросков следует тому же распределению $\mathrm{Geom}(p)$

Матожидание геометрического распределения

$X \sim \mathrm{Geom}(p)$

$$\begin{aligned}
E(X) &= 1 + E(X-1) \\
&= 1 + \Pr(X=1)E(X-1 \mid X=1) \\
&+ \Pr(X>1)E(X-1 \mid X>1) \\
&= 1 + 0 + (1-p)E(X) \\
pE(X) &= 1 \\
E(X) &= \frac{1}{p}
\end{aligned}$$

Но можно то же самое вычислить в лоб, посчитав сумму ряда:

$$E(X) = \sum_{k=1}^{+\infty} kp(1-p)^{k-1}$$