Лекция 2. Дискретные с.в. (продолжение)

17 февраля 2022 г.

1 Дисперсии различных распределений

Распределение Бернулли. $X \sim \mathrm{Bern}(p)$

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

Максимальное отклонение от матожидания при $p = \frac{1}{2}$. Заметим, что при этом с.в. всегда отклоняется от матожидания, равного $\frac{1}{2}$, на $\frac{1}{2}$. Поэтому

$$\sigma = \sqrt{\mathrm{Var}(X)} = \sqrt{\frac{1}{2} \cdot \left(1 - \frac{1}{2}\right)} = \frac{1}{2}$$

Равномерное распределение. $X \sim U(a, b)$. Можем посчитать вариацию $Y = X - a \sim U(0, n)$, где $n = b - a$.

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

2 Условные с.в.

Когда мы знаем, что произошло какое-то событие, у нас меняется вероятностная мера, а поэтому меняется и распределение с.в., заданной на нашем множестве исходов Ω.

Поэтому мы вводим понятие условной функции вероятностей:

$$p_X(x \mid A) = \Pr(X = x \mid A)$$

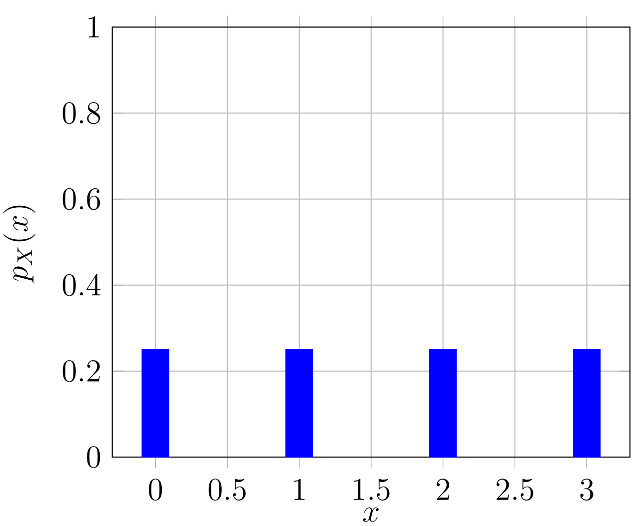
Но в целом, ничего нового не происходит, просто меняется функция веротяностей. Легко проверить свойства функции вероятностей:

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

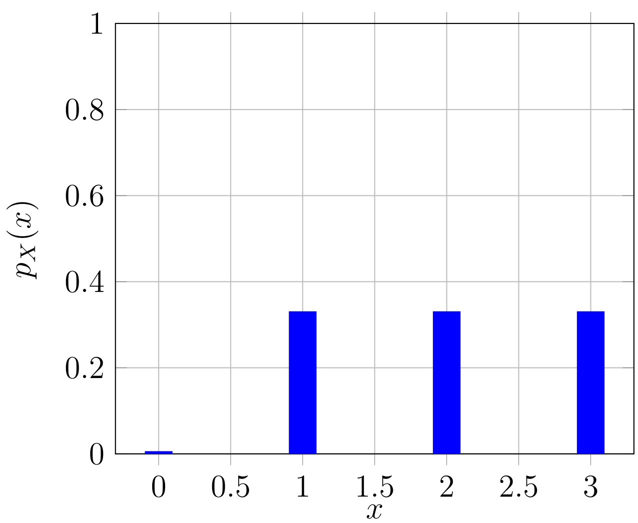
А также можно определить условное матожидание и условное матожидание функции от с.в.:

- $E(X \mid A) = \sum_x x \cdot p_X(x \mid A)$
- $E(g(X) \mid A) = \sum_x g(x) \cdot p_X(x \mid A)$

Пример условной с.в. Пусть есть $X \sim U(0, 3)$. Его функция распределения выглядит так:



Условное распределение при $X \mid X \geq 1$ будет выглятеть так:



Когда мы ввелди условные вероятности, мы также доказали теорему о полной вероятности, которая звучала так:

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

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

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

Мы пришли к так называемой формуле полного матожидания:

$$\boxed{E(X) = \sum_i \Pr(A_i) E(X \mid A_i)}$$

С помощью этой формулы посчитаем матожидание геометрического распределения, но сначала заметим одно свойство этого распределения, называемое *беспамятством*. Пусть $X \sim \mathrm{Geom}(p)$, то есть мы проводим опыт Бернулли, пока не получим первый успешный опыт. И, допустим, мы провели один опыт Бернулли, который не был успешным. Тогда какое распределение у числа оставшихся опытов до ближайшего успеха?

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

Можно обобщить и показать, что $p_{X-n}(x \mid X > n) = p_X(x)$. То есть при условии, что первые n исходов — неудачные, число оставшихся опытов следует тому же распределению $\mathrm{Geom}(p)$. С помощью этого свойства мы можем посчитать матожидание $X \sim \mathrm{Geom}(p)$

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

Конечно, то же самое вычислить в лоб, посчитав сумму ряда:

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

но это более заморочно.

3 Случайный вектор

Несколько с.в. (случайный вектор)

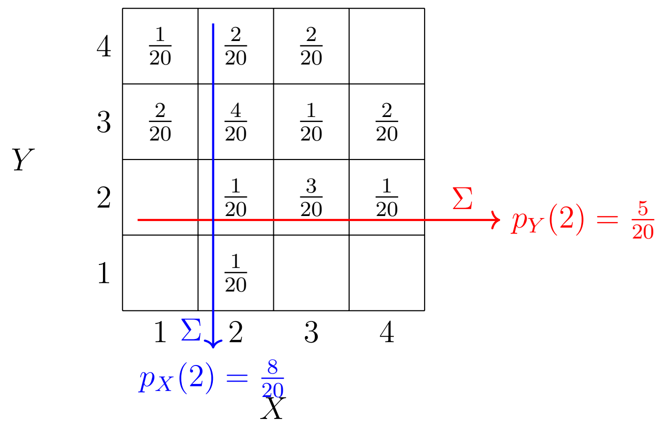
Совместная функция вероятностей

$$p_{X,Y}(x,y) = \Pr(X = x \cap Y = y)$$

Маргинальные функции вероятностей

$$p_X(x) = \sum_y p_{X,Y}(x,y)$$
$$p_Y(y) = \sum_x p_{X,Y}(x,y)$$

Пример случайного вектора В клетках — $p_{X,Y}(x,y)$



NB: все работает точно также и для большего числа с.в.

4 Линейность матожидания

Линейность матожидания (опять) Теперь у нас есть возможность смотреть $Z = g(X, Y)$

$$p_Z(z) = \Pr(g(X,Y) = z) = \sum_{(x,y):g(x,y)=z} p_{X,Y}(x,y)$$

$$E(Z) = \sum_x \sum_y g(x,y) p_{X,Y}(x,y)$$

Пусть $Z = X + Y$ как вычислить ее матожидание? Хорошо, что оно *линейно*

$$\boxed{E(Z) = E(X) + E(Y)}$$

Линейность матожидания: доказательство

$$\begin{aligned}
E(Z) = E(X+Y) &= \sum_x \sum_y (x+y) p_{X,Y}(x,y) \\
&= \sum_x \sum_y x p_{X,Y}(x,y) + \sum_x \sum_y y p_{X,Y}(x,y) \\
&= \sum_x x \sum_y p_{X,Y}(x,y) + \sum_y y \sum_x p_{X,Y}(x,y) \\
&= \sum_x x p_X(x) + \sum_y y p_Y(y) \\
&= E(X) + E(Y)
\end{aligned}$$

Окончательная линейность матожидания

$$E\left(\sum_i a_i X_i\right) = \sum_i a_i E(X_i)$$

NB: случайная величина может быть равна *константе*, поэтому мы не рассматриваем прибавление к сумме константы

Матожидание биномиального распределения *Напоминание:* $X \sim \mathrm{Bin}(n, p)$, если X равен числу успехов в n независимых испытаниях Бернулли с вероятностью успеха p.

$X = \sum_{i=1}^{n} X_i$, где все $X_i \sim \mathrm{Bern}(p)$

$$E(X) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} p = np$$

Всегда можно посчитать *в лоб*

$$E(X) = \sum_{i=0}^{n} i \binom{n}{i} p^i (1-p)^{n-i}$$

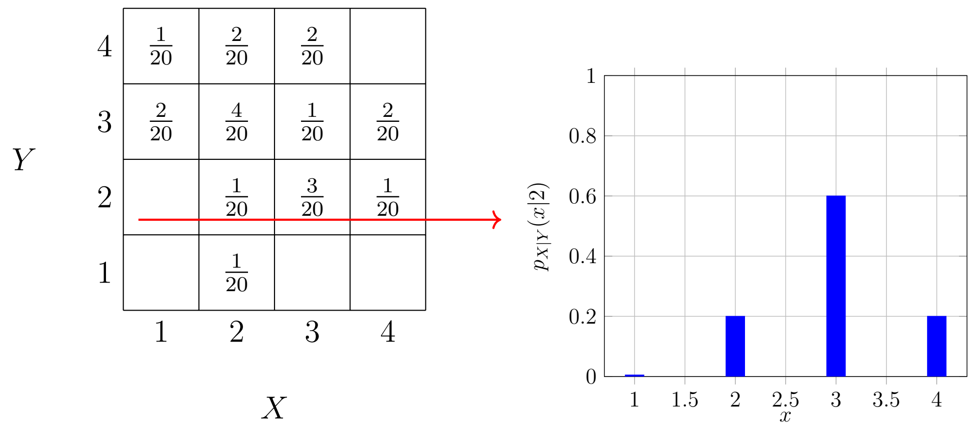
5 С.в., условная на другой с.в.

Условная с.в. (на другой с.в.)

Было:
$p_X(x \mid A) = \Pr(X = x \mid A)$
Стало:
$p_{X|Y}(x \mid y) = \Pr(X = x \mid Y = y)$

$$p_{X|Y}(x \mid y) = \frac{p_{X,Y}(x, y)}{p_Y(y)}$$

NB: Условная вероятность определена только для y, у которых $p_Y(y) > 0$



Условные случайные векторы

$$p_{X_1,\dots X_n|Y_1,\dots,Y_m}(x_1,\dots x_m \mid y_1,\dots,y_m)$$
$$= \Pr(X_1 = x_1 \cap \dots \cap X_n = x_n \mid Y_1 = y_1 \cap \dots Y_m = y_m)$$

Чуть проще:

- $p_{X|Y,Z}(x \mid y, z) = \Pr(X = x \mid Y = y \cap Z = z)$
- $p_{X,Y|Z}(x, y \mid z) = \Pr(X = x \cap Y = y \mid Z = z)$

Работает правило умножения:

- Было: $\Pr(A \cap B \cap C) = \Pr(A)\Pr(B \mid A)\Pr(C \mid A \cap B)$
- Стало: $p_{X,Y,Z}(x, y, z) = p_X(x)p_{Y|X}(y \mid x)p_{Z|X,Y}(z \mid x, y)$

Условное маотжидание

$$E(X \mid Y = y) = \sum_x x p_{X|Y}(x \mid y)$$

Условия просто *меняют вероятностную меру*, поэтому с условными с.в. все работает так же, как и с безусловными

Полные вероятность и матожидание Формула полной вероятности (оригинал): для разбиения Ω на A_i

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

Давайте теперь скажем, что $A_i = (Y = y)$, получим

$$p_X(x) = \sum_y p_Y(y) p_{X|Y}(x \mid y)$$

Полное матожидание — по аналогии

$$E(X) = \sum_y p_Y(y) E(X \mid Y = y)$$

NB: работает только когда ряд сходится *абсолютно*

6 Независимость

Независимость и условная независимость с.в.

- *Напоминание*:A и B независимы $\Leftrightarrow \Pr(A \cap B) = \Pr(A)\Pr(B)$
- Событие A и с.в. X независимы $\Leftrightarrow$

$$\forall x \quad \Pr(A \cap X = x) = \Pr(A) p_X(x)$$

- X и Y — *независимые* с.в. $\Leftrightarrow$

$$\forall x, y \quad p_{X,Y}(x, y) = p_X(x) p_Y(y)$$

NB: независимость по-прежнему означает, что одно событие (или с.в.) *не дает никакой информации* о другом событии (или с.в)

Условная независимость с.в. Условие порождает новую меру $\Rightarrow$ зависимость с.в. может поменяться в зависимости от условия

Y \ X	1	2	3	4
4	$\frac{1}{20}$	$\frac{2}{20}$	$\frac{2}{20}$	
3	$\frac{2}{20}$	$\frac{4}{20}$	$\frac{1}{20}$	$\frac{2}{20}$
2		$\frac{1}{20}$	$\frac{3}{20}$	$\frac{1}{20}$
1		$\frac{1}{20}$		

- X и Y, очевидно, зависимы:
  - $p_X(1) = \frac{3}{20}$
  - $p_{X|Y}(1, 1) = 0$
- Но при условии $C = (X \leq 2 \cap Y \geq 3)$ становятся независимы:
  - $p_X(1 \mid C) = \frac{1}{3}$, $p_X(2 \mid C) = \frac{2}{3}$
  - $p_Y(3 \mid C) = \frac{2}{3}$, $p_Y(4 \mid C) = \frac{1}{3}$
- Поэтому,
  - $p_{X,Y}(1, 4 \mid C) = \frac{1}{9} = p_X(1 \mid C)p_Y(4 \mid C)$
  - $p_{X,Y}(1, 3 \mid C) = \frac{2}{9} = p_X(1 \mid C)p_Y(3 \mid C)$
  - $p_{X,Y}(2, 4 \mid C) = \frac{2}{9} = p_X(2 \mid C)p_Y(4 \mid C)$
  - $p_{X,Y}(2, 3 \mid C) = \frac{4}{9} = p_X(2 \mid C)p_Y(3 \mid C)$

Матожидание независимых с.в.
Пусть X и Y — *независимые* с.в., тогда

$$\boxed{E(XY) = E(X)E(Y)}$$

$$\begin{aligned} E(XY) &= \sum_x \sum_y xyp_{X,Y}(x, y) = \sum_x \sum_y xyp_X(x)p_Y(y) \\ &= \sum_x xp_X(x) \sum_y yp_Y(y) = E(X)E(Y) \end{aligned}$$

Дисперсия независимых с.в.
Пусть X и Y — *независимые* с.в., тогда

$$\boxed{\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)}$$

Предположим, что $E(X) = E(Y) = 0$.
Тогда $E(XY) = E(X)E(Y) = 0$

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

Дисперсия биномиалки *Напоминание:* $X \sim \mathrm{Bin}(n,p) \Leftrightarrow X = \sum_{i=1}^{n} X_i$, где все $X_i \sim \mathrm{Bern}(p)$ — *независимы*

$$\mathrm{Var}(X) = \mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$$
$$= \sum_{i=1}^{n} p(1-p) = np(1-p)$$

7 Более сложные распределения

Гипер-геометрическое распределение

Эксперимент: у нас в корзине есть N шаров, D из которых белые, $N-D$ — черные. Мы достаем случайные n шаров (не возвращая их в корзину). X — число белых шаров, которые мы достали. Тогда $X \sim HG(D,N,n)$

Матожидание: представляем X_i как сумму индикаторных величин для всех белых шаров, что он был вытащен. Вероятность успеха $p = \frac{n}{N}$

$$E[X] = \sum_{i\text{- белый шар}} E[X_i] = D \cdot \frac{n}{N} = \frac{nD}{N}$$

Дисперсия:

$$E[X^2] = \sum_{i\text{- белый шар}} E[X_i^2] + \sum_{i,j\text{- белые шары}} E[X_i X_j] = E[X] + D(D-1) \cdot \frac{n}{N} \cdot \frac{n-1}{N-1}$$
$$= \frac{nD(N-n)(N-D)}{N^2(N-1)}$$

При больших N очень близко к биномиальному распределению $\mathrm{Bin}(n, \frac{D}{N})$.

Степенное распределение

$X \sim \mathrm{pow}(u,\beta)$ значит, что мы выбираем число из $[1..u]$, причем $\Pr(X=i) \sim i^{-\beta}$. Более точно:

$$\Pr[X=i] = \begin{cases} C_{u,\beta} i^{-\beta}, \text{ если } i \in [1..u], 0, \text{иначе.} \end{cases}$$

$C_{u,\beta}$ — коэффициент нормализации

$$C_{u,\beta} = \left(\sum_{i=1}^{u} i^{-\beta}\right)^{-1} = H_{u,\beta}^{-1}$$

$H_{u,\beta}^{-1}$ — обобщенное гармоническое число

Замечание: если $\beta > 1$, то u может быть $+\infty$

Рассказать про приближение интегралами