Лекция 5. Непрерывные с.в., часть 2

10 марта 2021 г.

1 Условная плотность вероятности

Напомним трактовку плотность вероятности. Это то, сколько вероятностной массы приходится на маленький интервал значений с.в.:

$$f_X(x)\delta \approx \Pr(X \in [x, x+\delta])$$

Но вероятностная мера может меняться при условии, что произошло событие A. В этом случае определяем

$$f_{X|A}(x)\delta \approx \Pr(X \in [x, x+\delta] \mid A)$$

Точное определение: плотность вероятности с.в. X при условии A есть такая функция $f_{X|A}(x)$, что для любого измеримого множества B верно, что

$$\Pr(X \in B \mid A) = \int_B f_{X|A}(x)dx.$$

NB: у нас опять просто поменялась вероятностная мера. То есть у условной плотности вероятности будут все те же свойства:

- $f_{X|A}(x) \geq 0$
- $\int_{-\infty}^{+\infty} f_{X|A}(x)dx = 1$

Вычисляем аналогично условной функции вероятности. Пусть $x \in A$, тогда

$$f_{X|A}(x)\delta \approx \Pr(X \in [x, x+\delta] \mid A) = \frac{\Pr(X \in [x, x+\delta] \cap X \in A)}{\Pr(A)}$$
$$= \frac{\Pr(X \in [x, x+\delta])}{\Pr(A)} \approx \frac{f_X(x)\delta}{\Pr(A)}$$

Поэтому строго говоря, она вычисляется так:

$$f_{X|A}(x) = \begin{cases} \frac{f_X(x)}{\Pr(A)}, & \text{если } x \in A, \\ 0, & \text{иначе.} \end{cases}$$

То есть масштабируем плотность вероятности по событию A.

Иногда придется иметь дело с условной функцией распределения:

$$F_{X|A}(x) = \Pr(X \leq x \mid A)$$

2 Условное матожидание

Условное матожидание определяем аналогично:

$$E(X \mid A) = \int_{-\infty}^{+\infty} x f_{X|A}(x) dx$$

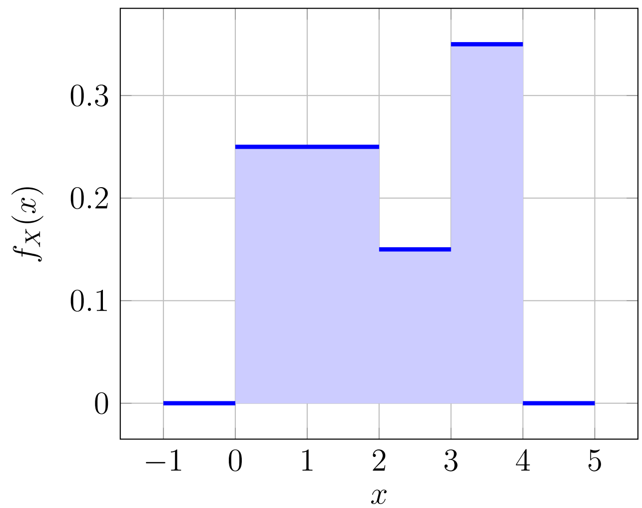
Работает то же самое правило и для функций от с.в. (у нас просто новая плотность вероятности при условии A):

$$E(g(X) \mid A) = \int_{-\infty}^{+\infty} g(x) f_{X|A}(x) dx$$

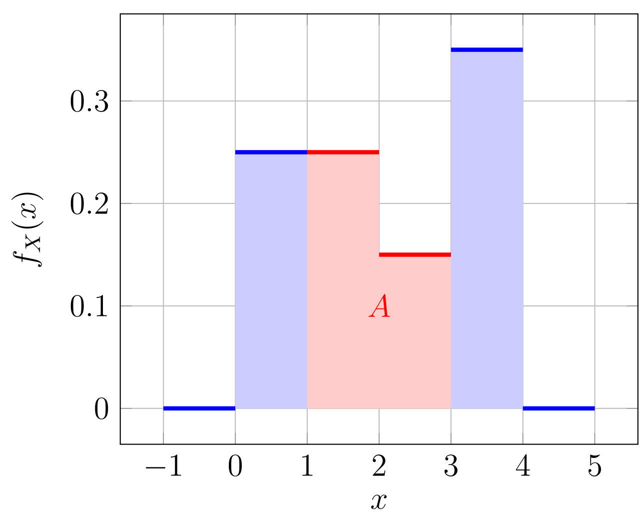
3 Пример условных с.в.

Рассмотрим частично равномерное распределение:

$$f_X(x) = \begin{cases} 0.25, x \in [0, 2) \\ 0.15, x \in [2, 3) \\ 0.35, x \in [3, 4] \\ 0, \text{ иначе.} \end{cases}$$

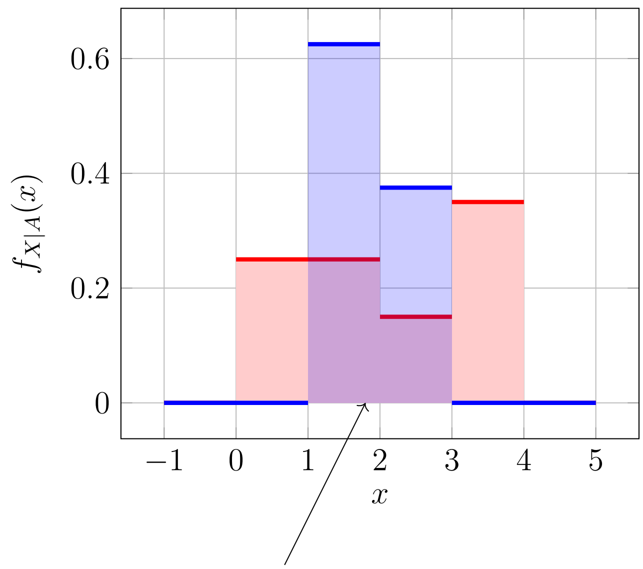


Пусть $A = [1, 3]$. Тогда $\Pr(A) = 0.25 \cdot 1 + 0.15 \cdot 1 = 0.4$



Значит, новая плотность вероятности выглядит так:

$$f_{X|A}(x) = \begin{cases} \frac{0.25}{0.4} = 0.625, x \in [1, 2) \\ \frac{0.15}{0.4} = 0.375, x \in [2, 3) \\ 0, \text{ иначе.} \end{cases}$$



$$E(X \mid A) = \int_1^2 t \cdot 0.625 dt + \int_2^3 t \cdot 0.375 dt = 1.875$$

4 Беспамятство экспоненциального распределения

Как уже говорилось, экспоненциальное распределение очень похоже на геометрическое. В том числе вот почему. Пусть продолжительность жизни лампочки T следует $\mathrm{Exp}(\lambda)$. Следует ли поменять лампочку после того, как она проработала время t? Посмотрим, сколько она еще проживет, то есть распределение $T - t$ при условии $T \geq t$.

$$F_{(T-t)|T \geq t}(x) = \Pr(T - t \geq x \mid T > t) = \frac{\Pr(T - t \geq x \cap T > t)}{\Pr(T > t)}$$
$$= \frac{\Pr(T \geq x + t)}{\Pr(T > t)} = \frac{e^{-\lambda(x+t)}}{e^{-\lambda t}} = e^{-\lambda x},$$

то есть распределение точно то же, как если мы заменим лампочку на новую.

5 Полные вероятность и матожидание

Напомним: пусть есть разбиение Ω на $\{A_i\}$ (не более, чем счетное), тогда

$$\Pr(B) = \Pr(A_1)\Pr(B \mid A_1) + \dots + \Pr(A_n)\Pr(B \mid A_n) + \dots$$
$$p_X(x) = \Pr(A_1) p_{X|A_1}(x) + \dots + \Pr(A_n) p_{X|A_n}(x) + \dots$$

Ничего не меняется и в непрерывном случае. Сначала функция распределения:

$$\begin{aligned} F_X(x) = \Pr(X \le x) &= \Pr(A_1)\Pr(X \le x \mid A_1) + \dots + \Pr(A_n)\Pr(X \le x \mid A_n) + \dots \\ &= \Pr(A_1)F_{X|A_1}(x) + \dots + \Pr(A_n)F_{X|A_n}(x) + \dots \end{aligned}$$

Дифференцируем, получаем:

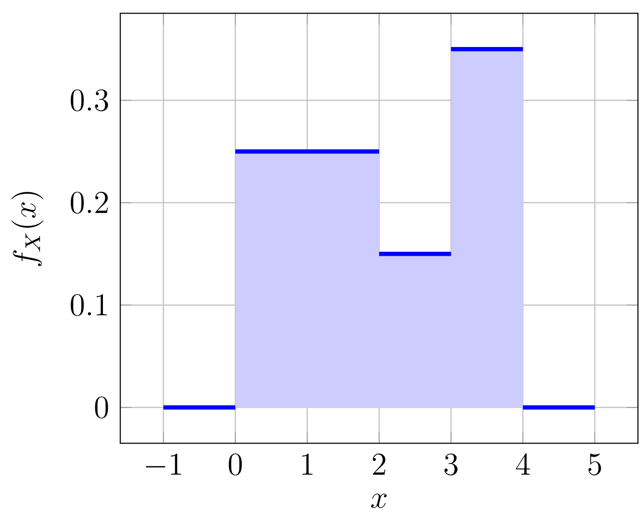
$$f_X(x) = \Pr(A_1)f_{X|A_1}(x) + \dots + \Pr(A_n)f_{X|A_n}(x) + \dots$$

Умножаем на x и интегрируем по всему $\mathbb{R}$:

$$E(X) = \Pr(A_1)E(X \mid A_1) + \dots + \Pr(A_n)E(X \mid A_n) + \dots$$

Пример (который уже был):

$$f_X(x) = \begin{cases} 0.25, x \in [0,2) \\ 0.15, x \in [2,3) \\ 0.35, x \in [3,4] \\ 0, \text{ иначе.} \end{cases}$$



Посчитаем матожидание

$$\begin{aligned} E(X) = &\Pr(X \in [0,2))E(X \mid X \in [0,2)) \\ &+ \Pr(X \in [2,3))E(X \mid X \in [2,3)) \\ &+ \Pr(X \in [3,4))E(X \mid X \in [3,4)) \end{aligned}$$

Заметим, что на каждом отрезке матожидание – это положение центра масс, то есть середина отрезка. Поэтому

$$E(X) = 0.5 \cdot 1 + 0.15 \cdot 2.5 + 0.35 \cdot 3.5 = 2.1$$

6 Смешанные распределения

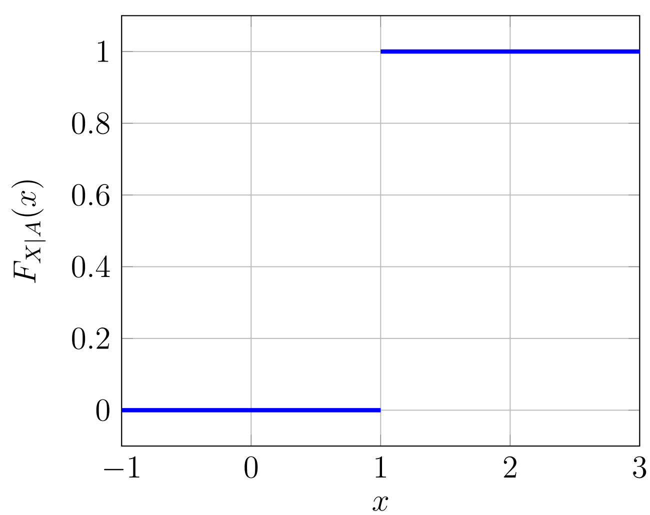
Иногда с.в. могут быть ни дискретными, ни непрерывными, например. Пусть у нас есть слеующий эксперимент. Сначала бросам честную монетку, потом, если выпал орел, то выбираем случайное число из отрезка $[0, 2]$ (равномерно). Случайная величина X при этом равна 1 в случае решки и равна выбранному числу в случае орла.

У данной с.в. нет функции вероятностей, как нет и плотности вероятности. Фукнция распределения все-такие есть. Как ее посчитать? По формуле полной вероятности, которая работает и для функции распределения. Пусть A — событие "выпала решка"

$$F_X(x) = F_{X|A}(x)\Pr(A) + F_{X|\bar{A}}(x)\Pr(\bar{A})$$

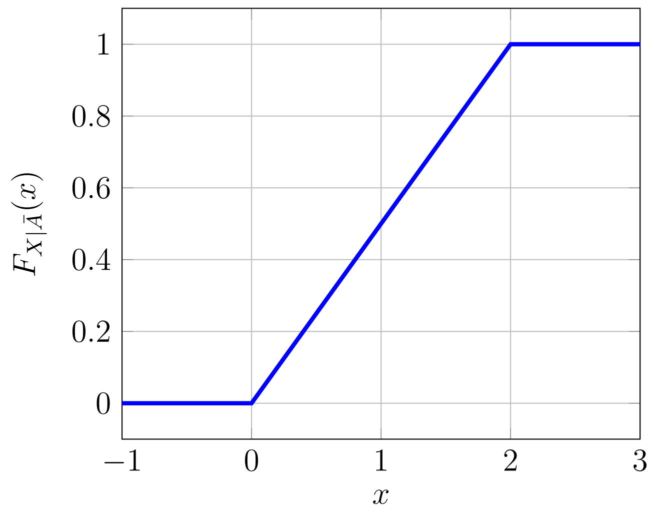
Заметим, что если A, то $X = 1$ с вероятностью 1. То есть,

$$F_{X|A}(x) = \begin{cases} 0, x < 1 \\ 1, x \geq 1 \end{cases}$$



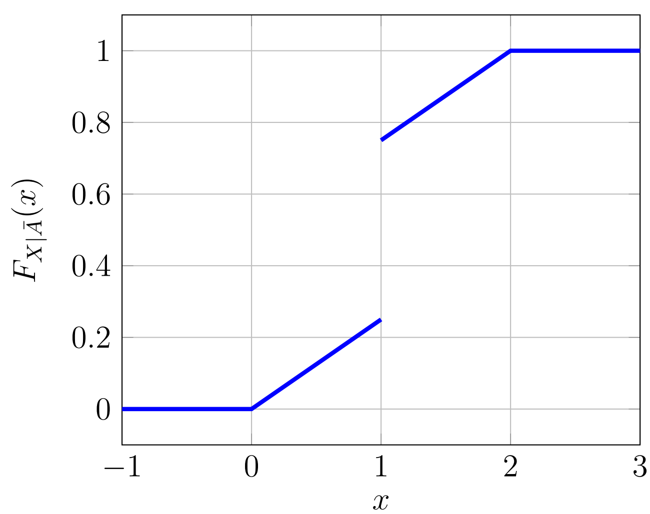
А если $\bar{A}$, то X следует равномерному распределению на отрезке $[0, 2]$.

$$F_{X|\bar{A}}(x) = \begin{cases} 0, x < 0 \\ x/2, x \in [0, 2] \\ 1, x > 2 \end{cases}$$



Итоговая функция распределения выглядит так:

$$F_X(x) = \frac{1}{2}F_{X|A}(x) + \frac{1}{2}F_{X|\bar{A}}(x) = \begin{cases} 0, x < 0 \\ x/4, x \in [0, 1) \\ x/4 + 1/2, x \in [1, 2] \\ 1, x > 2 \end{cases}$$



7 Векторы из непрерывных с.в.

Пусть у нас есть 2 непрерывных с.в. X и Y. Тогда можно говорить о совместной плотности вероятности:

$$f_{X,Y}(x, y)\delta^2 \approx \Pr(X \in [x, x + \delta] \cap Y \in [y, y + \delta])$$

Более строго: если у нас есть такая функция $f_{X,Y}(x,y)$, такая что для любого измеримого множества $A \subset \mathbb{R}^2$ верно

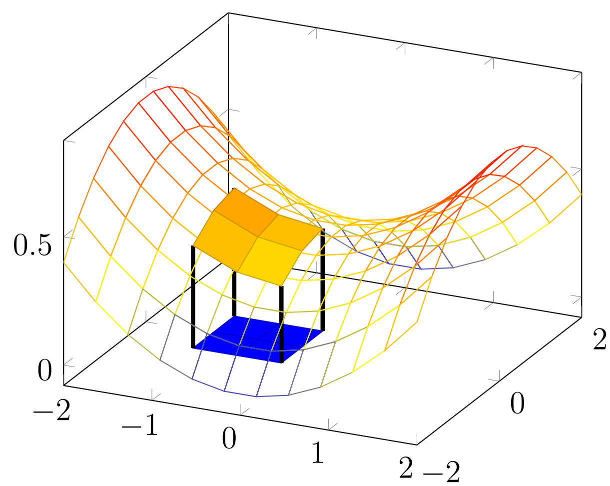
$$\Pr((X,Y) \in A) = \iint\limits_A f_{X,Y}(x,y)dxdy,$$

то с.в. X и Y называются совместно непрерывными, а $f_{X,Y}(x,y)$ называется их совместной плотностью вероятности

Свойства, аналогичные совместной функции вероятности, только суммы заменены интегралами:

- $f_{X,Y}(x,y) \geq 0$
- $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dxdy = 1$

Как визуализировать:



Чтобы посчитать вероятность, что X попадает в какое-то событие, надо просто посчитать объем подграфика на этом событии

NB: Одномерные события имеют вероятность ноль. Например, событие $Y = X$.

8 Маргинальные распределения

Покажем, что

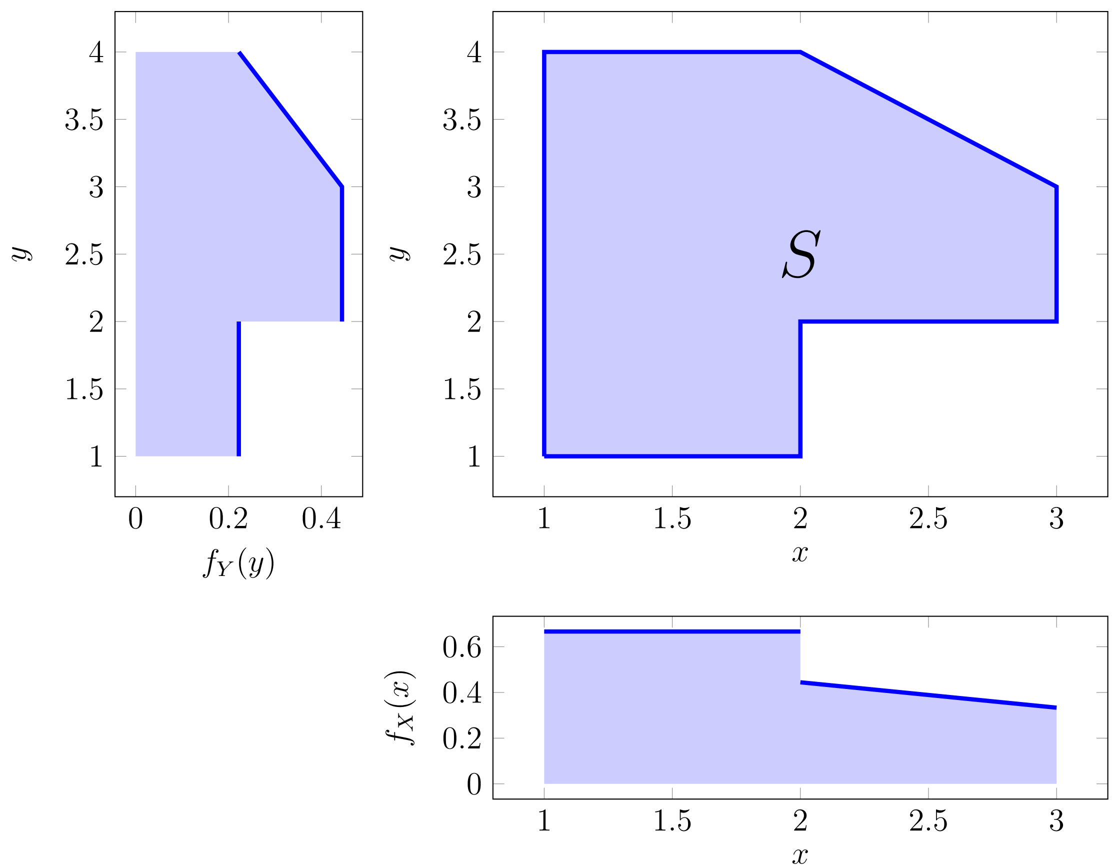
$$\begin{cases} f_X(x) &= \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy \\ f_Y(y) &= \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dx \end{cases}$$

Для этого рассмотрим функцию распределения X:

$$F_X(x) = \Pr(X \leq X) = \int_{-\infty}^{x} \left(\int_{-\infty}^{+\infty} f_{X,Y}(t, y) dy \right) dt f_X(x) \quad = F'_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dy$$

Пример: равномерное распределение на множестве S площадью 4.5

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{|S|}, & \text{если } (x, y) \in S, \\ 0, & \text{иначе.} \end{cases}$$



9 Более двух с.в. и функции от многих с.в.

Совместное распределение может быть задано на более, чем одной с.в., тогда есть плотность вероятности $f_{X,Y,Z}(x, y, z)$. Все то же самое, что было со многими дискретными с.в., только вместо сумм интегралы:

- $f_{X,Y,Z}(x, y, z) \geq 0$
- $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{X,Y,Z}(x, y, z) dxdydz = 1$

И на многих с.в. мы также можем задавать функции, которые будут по сути новыми с.в.: $Z = g(X, Y)$. Их матожидание считается так:

$$E(g(X,Y)) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x,y) f_{X,Y}(x,y) dx dy$$

При этом имеет место быть линейность матожидания:

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

10 Совместная функция распределения

Для нескольких с.в. определим функцию распределения:

$$F_{X,Y}(x,y) = \Pr(X \leq x \cap Y \leq Y) = \int_{-\infty}^{x} \left(\int_{-\infty}^{y} f_{X,Y}(s,t) dt \right) ds$$

Также можно заметить:

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$$

11 С.в., условные на других с.в.

Было в дискретном случае:

- $p_{X,Y}(x,y)$ — совместная функция вероятности
- $p_{X|A}(x) = \frac{p_X(x) \cdot [x \in A]}{\Pr(A)}$ — условная функция вероятности (где $[\cdot]$ — скобка Айверсона)
- $p_{X|Y}(x \mid y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$ — функция вероятности X, условная на Y (определена только для таких y, что $p_Y(y) > 0$)

То же самое есть для непрерывного случая:

- $f_{X,Y}(x,y)$ — совместная плотность вероятности
- $f_{X|A}(x) = \frac{f_X(x) \cdot [x \in A]}{\Pr(A)}$ — условная плотность вероятности
- $f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$ — плотность вероятности X, условная на Y (определена только для таких y, что $p_Y(y) > 0$)

Чуть ближе к формальному определению:

$$
\begin{aligned}
\Pr\left(X \in [x, x+\delta] \mid Y \in [y, y+\varepsilon]\right) &= \frac{\Pr\left(X \in [x, x+\delta] \cap Y \in [y, y+\varepsilon]\right)}{\Pr(Y \in [y, y+\varepsilon])} \\
&\approx \frac{f_{X,Y}(x,y)\delta\varepsilon}{f_Y(y)\varepsilon} = \frac{f_{X,Y}(x,y)}{f_Y(y)}\delta
\end{aligned}
$$

И совсем формальное. Если существует такая функция $f_{X|Y}(x \mid y)$, что для всех y и для всех A верно

$$
\Pr(X \in A \mid Y = y) = \int_A f_{X|Y}(x \mid y) dx,
$$

то эта функция называется условной плотностью вероятности X при условии Y.

Опять при условиях у нас просто появляется новая вероятностная мера. Для всех $y : f_Y(y) > 0$ мы имеем

- $f_{X|Y}(x \mid y) \geq 0$
- $\int_{-\infty}^{+\infty} f_{X|Y}(x \mid y) dx = \frac{\int_{-\infty}^{+\infty} f_{X,Y}(x|y) dx}{f_Y(y)} = 1$

Из последнего видно, что стоит воспринимать условную плотность как срез совместной плотности по какому-то значению с.в.. Заметьте, что при этом вероятность этого среза равна нулю (одномерное множество), поэтому только терминами условности на события тут не обойтись.

Работает правило умножения:

$$
\begin{aligned}
f_{X,Y}(x,y) &= f_X(x) f_{Y|X}(y \mid x) \\
&= f_Y(y) f_{X|Y}(x \mid y)
\end{aligned}
$$

А значит, работают полные вероятность и матожидание. Полная вероятность (следует из правила умножения):

$$
f_X(x) = \int_{-\infty}^{+\infty} f_Y(y) f_{X|Y}(x, y) dy
$$

Также определено условное матожидание

$$
E(X \mid Y = y) = \int_{-\infty}^{+\infty} x f_{X|Y}(x, y) dx
$$

и работает теорема о полном матожидании (TODO: доказать)

$$
E(X) = \int_{-\infty}^{+\infty} f_Y(y) E(X \mid Y = y) dy
$$

Также можно посчитать условное матожидание функции с.в.

$$
E(g(X) \mid Y = y) = \int_{-\infty}^{+\infty} g(x) f_{X|Y}(x, y) dx
$$

12 Независимость с.в.

Было у дискретных:

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

Логично определить непрерывные с.в. X и Y независимыми, если

$$f_{X,Y}(x,y) = f_X(x)f_Y(y)$$

Это автоматически подразумевает, что для всех y верно, что $f_{X|Y}(x \mid y) = f_X(x)$.

Свойств независимых с.в. те же, что и у дискретных

- $E(XY) = E(X)E(Y)$
- $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$
- $g(X)$ и $h(Y)$ тоже будут независимы

Пример про зависимые с.в.

Есть палка длиной ℓ. Ломаем ее в случайном месте $X \in [0, \ell]$, остаток тоже ломаем в случайном месте $Y \in [0, X]$. Найти:

- $f_{X,Y}$
- $f_Y(y)$
- $E[Y]$