Лекция 6. Непрерывные с.в., часть 3

17 марта 2021 г.

Закончили на определении плотности вероятности с.в. X при известном значении с.в. Y.

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

И ее свойствах:

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

Еще раз подчеркнем, 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 и определения маргинальной плотности вероятности):

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

Также определено условное матожидание

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

и работает теорема о полном матожидании:

$$
\begin{aligned}
E(X) &= \int_{-\infty}^{+\infty} x f_X(x) dx = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x f_Y(y) f_{X|Y}(x \mid y) dy dx \\
&= \int_{-\infty}^{+\infty} f_Y(y) \left(\int_{-\infty}^{+\infty} x f_{X|Y}(x \mid y) dx \right) dy \\
&= \int_{-\infty}^{+\infty} f_Y(y) E(X \mid Y = y) dy
\end{aligned}
$$

Также можно посчитать условное матожидание функции с.в.

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

1 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 $f_{X|Y}(x \mid y) = f_X(x)$. Свойства 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 независимых с.в.

Пусть есть два трамвая, которые ходят с интервалом 11 и 17 минут, независимо друг от друга. Вы приходите на остановку в случайный момент времени. Сколько ожидаемо вы будете ожидать трамвай, если вам подходит любой из двух?

Определим с.в. Пусть T_1 — время, через которое придет первый трамвай, а T_2 — второй. Заметим, что $T_1 \sim U(0, 11)$, а $T_2 \sim U(0, 17)$, то есть

$$F_{T_1}(t) = \begin{cases} 0, t < 0 \\ \frac{t}{11}, t \in [0, 11] \\ 1, t > 11, \end{cases}$$

$$F_{T_2}(t) = \begin{cases} 0, t < 0 \\ \frac{t}{17}, t \in [0, 17] \\ 1, t > 17, \end{cases}$$

А время, которое придется провести на остановке есть $Y = \min\{T_1, T_2\}$, функция от двух независимых с.в.

Определим функцию распределения Y.

$$\begin{aligned} F_Y(y) &= \Pr(Y \le y) = \Pr(T_1 \le y \cup T_2 \le y) \\ &= 1 - \Pr(T_1 > y \cap T_2 > y) \\ &= 1 - \Pr(T_1 > y)\Pr(T_2 > y) = 1 - (1 - F_{T_1}(y))(1 - F_{T_2}(y)) \end{aligned}$$

Воспользуемся полезной формулой с прошлой практики для неотрицательных с.в. (*которую доказали не все*)

$$\begin{aligned} E(Y) &= \int_0^{+\infty} (1 - F_Y(y))dy = \int_0^{+\infty} (1 - F_{T_1}(y))(1 - F_{T_2}(y))dy \\ &= \int_0^{11} \left(1 - \frac{y}{11}\right)\left(1 - \frac{y}{17}\right) dy \approx 4.31 \end{aligned}$$

Пример про зависимые с.в.

Есть палка длиной ℓ. Ломаем ее в случайном месте $X \in [0, \ell]$, остаток тоже ломаем в случайном месте $Y \in [0, X]$. Какова длина остатка Y? Заметим, что величины зависимы: если $X \ge \ell/2$, то есть шанс, что $Y \ge \ell/2$, а иначе нет.

Совместная функция распределения:

$$f_{X,Y}(x, y) = f_X(x) f_{Y|X}(y \mid x) = \frac{1}{\ell x}$$

Теперь можем найти плотность вероятности Y и его матожидание. Для этого интегрируем совместную ПВ по всем возможным X, то есть от y до ℓ

$$f_Y(y) = \int_y^\ell f_{X,Y}(x, y)dx = \int_0^\ell \frac{1}{\ell x}dx = \frac{1}{\ell}\ln\frac{\ell}{y}.$$

Два способа посчитать матожидание. Первый:

$$E(Y) = \int_0^\ell y f_Y(y)dy = \int_0^\ell \frac{y}{\ell}\ln\frac{\ell}{y}dy = \left(\frac{\ln(\ell)}{\ell}\cdot\frac{y^2}{2} - \frac{y^2\ln(y)}{2\ell} + \frac{y^2}{4\ell}\right)\Bigg|_0^\ell = \frac{\ell}{4}$$

Второй:

$$E(Y) = \int_0^\ell f_X(x)E(Y \mid X = x)dx = \int_0^\ell \frac{1}{\ell}\cdot\frac{x}{2}dx = \frac{x^2}{4\ell}\Bigg|_0^\ell = \frac{\ell}{4}$$

Интуитивная логика: первый раз палка ломается в среднем посередине, потом снова в среднем посередине, то есть средняя длина должна быть $\ell/4$. Она работает не всегда, а только для Y, матожидание которых линейно относительно X. Пусть $E(Y \mid X = x) = g(x)$ тогда:

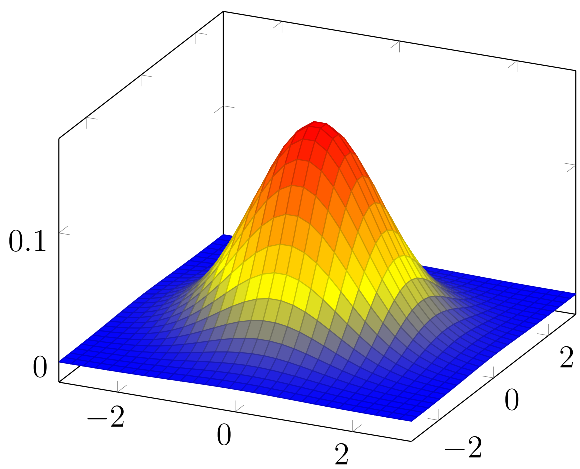
$$E(Y) = \int_{-\infty}^{+\infty} f_X(x) E(Y \mid X = x) dx = \int_{-\infty}^{+\infty} f_X(x) g(x) dx = E(g(X)) \neq g(E(X)).$$

Независимые нормальные распределения

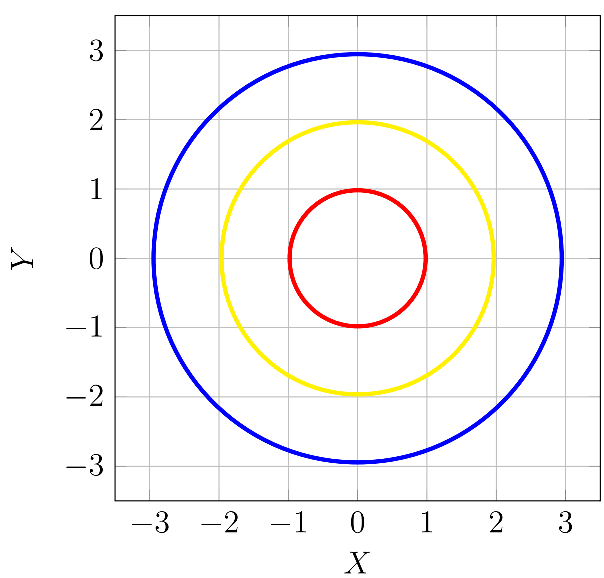
Пусть есть $X \sim N(0,1)$ и $Y \sim N(0,1)$, и они независимы. Тогда

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2} = \frac{1}{2\pi} \exp\left(-\frac{x^2+y^2}{2}\right).$$

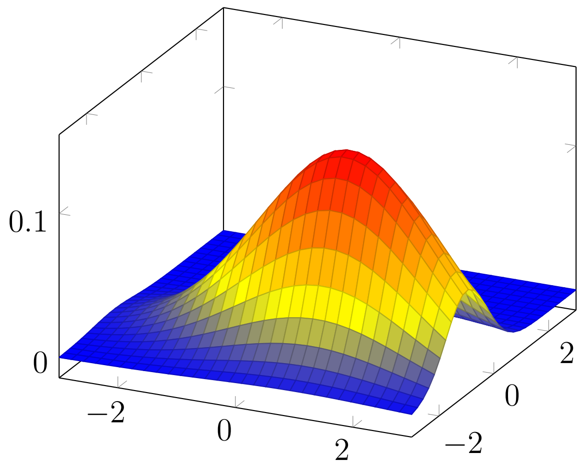
То есть их совместная плотность вероятности пропорциональна $e^{-r^2/2}$, где r — расстояние до точки (0, 0).



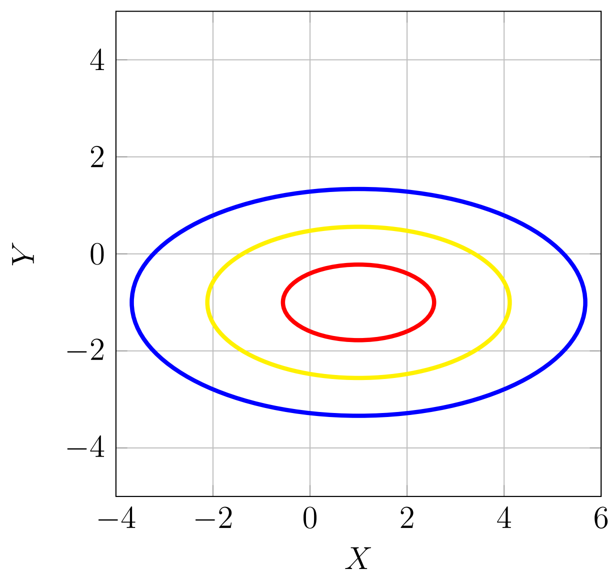
Эквиплотные линии:



Если у нас нестандартные нормальные распределения, то бугорок может сплющиваться или растягиваться вдоль оси, и его центр будет свдигаться



Эквиплотные линии превратятся в эллипсы.



Заметим, что оси эллипса должны быть направлены вдоль осей, иначе с.в. будут зависимые

2 Формула Байеса для случайных величин

Напомним смысл формулы Байеса. С помощью нее мы выражаем вероятность события, которое не можем пронаблюдать (A_i в формуле) через априорные оценки вероятностей другого события(ий) (B в формуле), которое мы можем наблюдать, после его наблюдения.

Априорная оценка вероятности A_i

Априорная оценка условной вероятности наблюдаемого события B

$$\boxed{\Pr(A_i \mid B) = \frac{\Pr(A_i)\Pr(B \mid A_i)}{\sum_j \Pr(A_j)\Pr(B \mid A_j)}}$$

Апостериорная оценка вероятности A_i

Полная вероятность B

Легко вывести из правила умножения и формулы полной вероятности для двух дискретных и двух непрерывных случанйых величин. Начнем с дискретных

$$\begin{aligned}
p_{X,Y}(x,y) &= p_Y(y)p_{X|Y}(x \mid y) \\
&= p_X(x)p_{Y|X}(y \mid x) \\
\Rightarrow p_{X|Y}(x \mid y) &= \frac{p_X(x)p_{Y|X}(y \mid x)}{p_Y(y)} \\
&= \frac{p_X(x)p_{Y|X}(y \mid x)}{\sum_x p_X(x)p_{Y|X}(y \mid x)}
\end{aligned}$$

$$\boxed{p_{X|Y}(x \mid y) = \frac{p_X(x)p_{Y|X}(y \mid x)}{\sum_x p_X(x)p_{Y|X}(y \mid x)}}$$

То же самое для непрерывных с.в., но говорим про плотности вероятности, а не про функцию вероятности, и суммы заменяем интегралами

$$\begin{aligned}
f_{X,Y}(x,y) &= f_Y(y)f_{X|Y}(x \mid y) \\
&= f_X(x)f_{Y|X}(y \mid x) \\
\Rightarrow f_{X|Y}(x \mid y) &= \frac{f_X(x)f_{Y|X}(y \mid x)}{f_Y(y)} \\
&= \frac{f_X(x)f_{Y|X}(y \mid x)}{\int_{\mathbb{R}} f_X(x)f_{Y|X}(y \mid x)dx}
\end{aligned}$$

$$\boxed{f_{X|Y}(x \mid y) = \frac{f_X(x)f_{Y|X}(y \mid x)}{\int_{\mathbb{R}} f_X(x)f_{Y|X}(y \mid x)dx}}$$

Но иногда могут быть случаи, когда есть две с.в.: одна дискретная, другая непрерывная. Можем сделать примерно следующее. Пусть X — дискретная, а Y — непрерывная

$$\begin{aligned}
&\Pr(X = x \cap y \le Y \le y + \delta) \\
&\qquad\qquad = \Pr(X = x)\Pr(y \le Y \le Y + \delta \mid X = x) \approx p_X(x) f_{Y|X}(y \mid x)\delta \\
&\qquad\qquad = \Pr(y \le Y \le y + \delta)\Pr(X = x \mid y \le Y \le y + \delta) \approx f_Y(y)\delta p_{X|Y}(x \mid y)
\end{aligned}$$

Откуда следует, что

$$p_X(x) f_{Y|X}(y \mid x) = f_Y(y) p_{X|Y}(x \mid y)$$

Чтобы доказать более строго, надо просто δ устремить к нулю, тогда вместо "≈" будет "=". Получаем две формулы.

$$\boxed{p_{X|Y}(x \mid y) = \frac{p_X(x) f_{Y|X}(y \mid x)}{f_Y(y)}} \qquad \boxed{f_{Y|X}(y \mid x) = \frac{f_Y(y) p_{X|Y}(x \mid y)}{p_X(x)}}$$

И теперь в левую формулу можно подставить формулу полной вероятности, которая работает для $f_Y(y)$:

$$f_Y(y) = \sum_{x'} p_X(x') f_{Y|X}(y \mid x').$$

С правой чуть посложнее, но пока поверьте наслово, что

$$p_X(x) = \int_{\mathbb{R}} f_Y(y') p_{X|Y}(x \mid y') dy.$$

Пример: наблюдаем непрерывную с.в., оцениваем дискретную

$$\boxed{p_{X|Y}(x \mid y) = \frac{p_X(x) f_{Y|X}(y \mid x)}{f_Y(y)}}$$

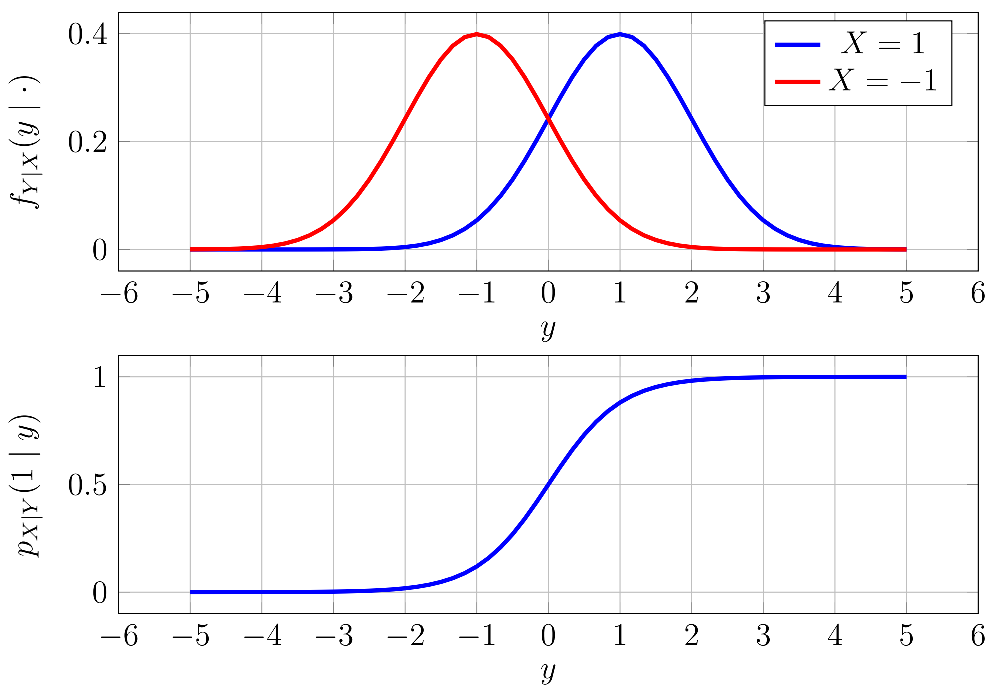
Ситуация: посылаем дискретный сигнал $X \in [-1, 1]$, но к нему добавляется шум $Z \sim N(0, 1)$. В итоге мы можем замерить только $Y = X + Z$. Давайте определим вероятности каждого варианта посланного сигнала, если изначально отправка каждого равновероятна.

- $p_X(-1) = p_X(1) = \frac{1}{2}$

- $Y \sim N(0,1) + X$, то есть если $X = 1$, то $Y \mid X = 1 \sim N(1,1)$. Аналогично $Y \mid X = -1 \sim N(-1,1)$
- Из предыдущего понимаем, что
  - $f_{Y|X}(y,1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-1)^2}{2}}$
  - $f_{Y|X}(y,-1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(y+1)^2}{2}}$
- $f_Y(y) = \frac{1}{2} f_{Y|X}(y,1) + \frac{1}{2} f_{Y|X}(y,-1)$

$$\begin{aligned} p_{X|Y}(1 \mid y) &= \frac{\frac{1}{2} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-1)^2}{2}}}{\frac{1}{2} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-1)^2}{2}} + \frac{1}{2} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(y+1)^2}{2}}} \\ &= \frac{1}{1 + e^{-\frac{(y+1)^2 - (y-1)^2}{2}}} = \frac{1}{1 + e^{-2y}} \end{aligned}$$

Иллюстрация: распределение наблюдаемого значения при разных сигналов и вероятность, что посланный сигнал равне единице, в зависимости от наблюдаемого значения



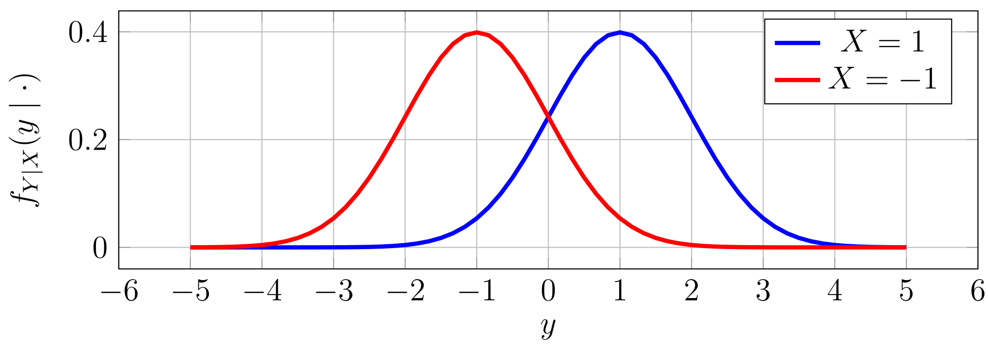
Пример: наблюдаем дискретную с.в., оцениваем непрерывную

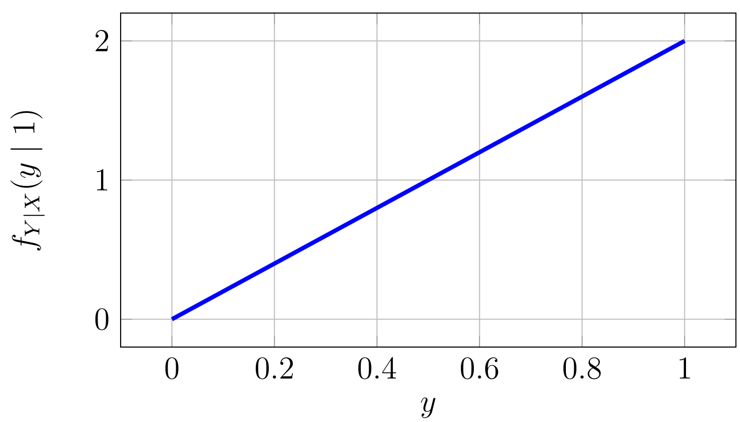
$$f_{Y|X}(y \mid x) = \frac{f_Y(y) p_{X|Y}(x \mid y)}{p_X(x)}$$

Эксперимент: берем нечестную монету, бросаем ее и исходя из результата хотим оценить степень ее нечестности. Наблюдаемая с.в. $X \in \{0, 1\}$ и неизвестная с.в. $Y = \Pr(X = 1) \sim U(0, 1)$.

- $f_Y(y) = 1$ на отрезке $[0, 1]$
- $p_{X|Y}(1 \mid y) = y$
- $p_{X|Y}(0 \mid y) = 1 - y$
- $p_X(1) = \int_0^1 f_Y(y')p_{X|Y}(1 \mid y')dy' = \int_0^1 y'dy = \frac{y^2}{2}\Big|_0^1 = \frac{1}{2}$

$$f_{Y|X}(y \mid 1) = \frac{1 \cdot y}{1/2} = 2y$$





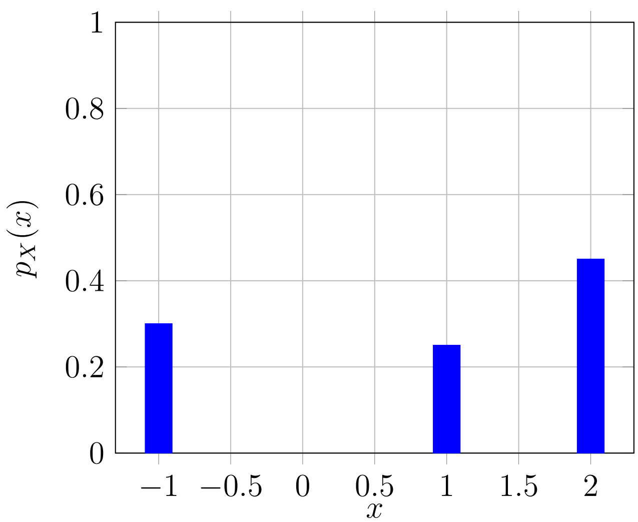
3 Линейные функции от с.в.

Дискретные с.в.

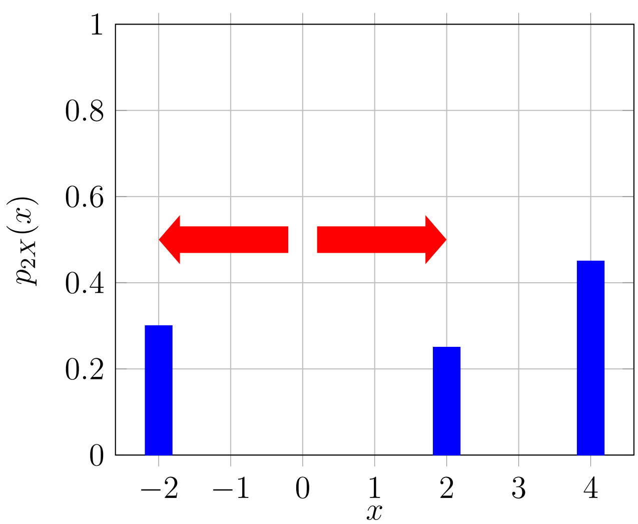
Если $Y = g(X)$ и мы знаем функцию вероятности $p_X(x)$, то не ноствит труда посчитать

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

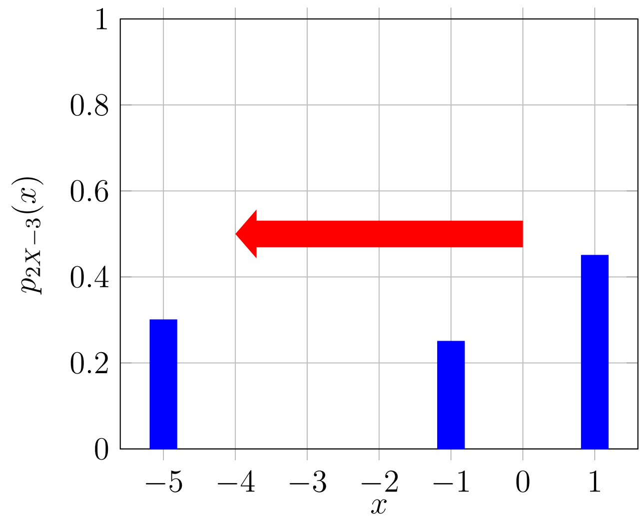
рассмотрим простой случай: $g(x) = ax + b$ — линейная функция. Что происходит с функцией распределения?



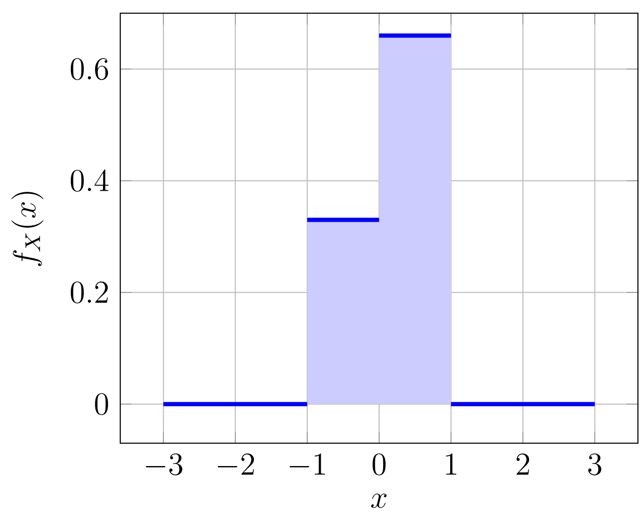
Расмотрим сначала, например, $g(x) = 2x$. Легко понять, что функция вероятностей сохранила свою форму, просто столбики отъехали от оси OY в два раза.



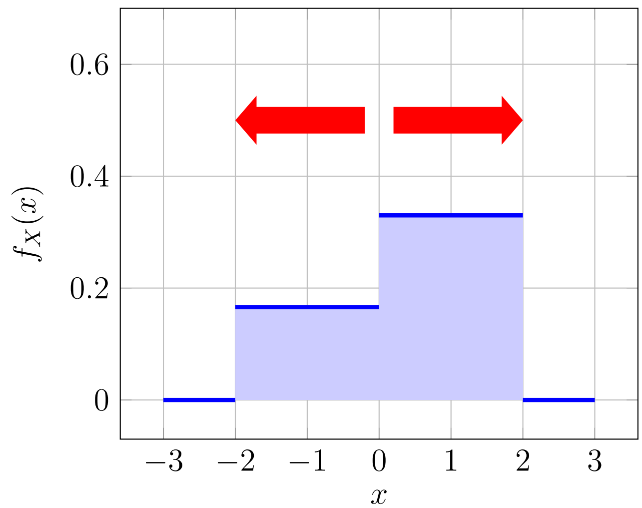
Если мы еще прибавим константу $b = -3$, то просто сдвинем все столбики влево на 3.



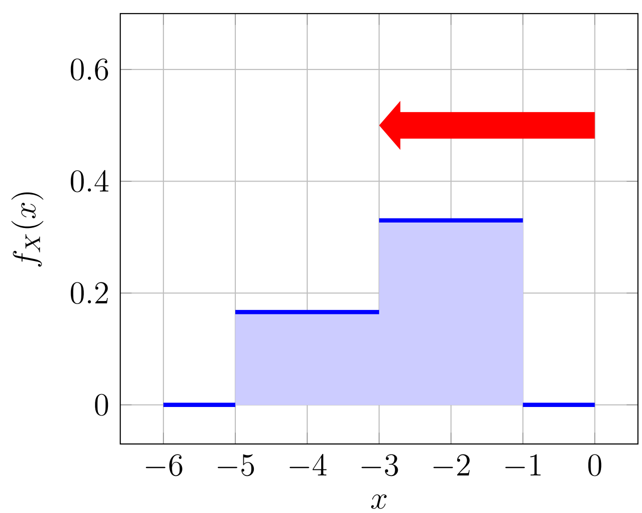
Очень похожая история с непрерывными. Рассмотрим пример.



Для того, чтобы получить функцию плотности вероятности для с.в. $Y = 2x$, надо растянуть ее от оси OY в два раза. Заметьте, что при растягивании сама плотность упадет также в два раза.



Ну и сдвиг при добавлении константы аналогичный. Рассмотрим, например, $g(x) = 2x - 3$



Как это в общем случае делать с.в.? Рассмотрим $Y = aX + b$ (где $a \neq 0$, иначе это скучный случай, когда $Y = b$ с вероятностью 1). Пусть сначала X дискретная, и нам известна ее функция вероятностей. Тогда

$$p_Y(y) = \Pr(Y = y) = \Pr(aX + b = y) = \Pr\left(X = \frac{y - b}{a}\right) = p_X\left(\frac{y - b}{a}\right)$$

С непрерывными такое не работает, так как вероятность, что непрерывная с.в. равна конкретному числу, есть ноль. Но мы можем работать с функциями распределения! Допустим мы знаем $f_X(x)$ и $F_X(x)$. Рассмотрим случай $a > 0$

$$
\begin{aligned}
F_Y(y) &= \Pr(Y \le y) = \Pr(aX + b \le y) = \Pr\left(X \le \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right), \\
f_Y(y) &= F'_Y(y) = f_X\left(\frac{y-b}{a}\right) \cdot \frac{1}{a}.
\end{aligned}
$$

И отдельно $a < 0$

$$
\begin{aligned}
F_Y(y) &= \Pr(Y \le y) = \Pr(aX + b \le y) = \Pr\left(X \ge \frac{y-b}{a}\right) = 1 - F_X\left(\frac{y-b}{a}\right), \\
f_Y(y) &= F'_Y(y) = -f_X\left(\frac{y-b}{a}\right) \cdot \frac{1}{a}.
\end{aligned}
$$

Объединяя два случая, получаем:

$$
f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right).
$$

Докажем теперь, что линейное преобразование нормального распределения оставляет его нормальным. Пусть $X \sim N(\mu, \sigma^2)$ и $Y = aX + b$. Значит,

$$
\begin{aligned}
f_Y(y) &= \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right) \\
&= \frac{1}{|a|\sigma\sqrt{2\pi}} \exp\left(\frac{\left(\frac{y-b}{a} - \mu\right)^2}{2\sigma^2}\right) \\
&= \frac{1}{(|a|\sigma)\sqrt{2\pi}} \exp\left(\frac{(y - (b + a\mu))^2}{2(a\sigma)^2}\right),
\end{aligned}
$$

что есть функция плотности вероятности для $N(a\mu + b, (a\sigma)^2)$.