3 One-dimensional transformations of distributions

3a Linear transformations

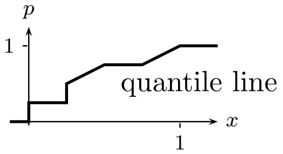
Let two random variables $X, Y : \Omega \to \mathbb{R}$ be related by the equality

$$Y = aX + b \qquad \text{(that is, } \forall \omega \;\; Y(\omega) = aX(\omega) + b\text{)}$$

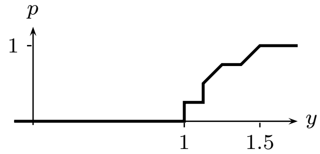
with some (nonrandom) parameters $a, b \in \mathbb{R}$.

Let $p \in (0,1)$, $x \in \mathbb{R}$, and $y = ax + b$. If $a > 0$ then

(3a1) $$\big(\, x \text{ is a } p\text{-quantile of } X \, \big) \iff \big(\, y \text{ is a } p\text{-quantile of } Y \, \big) \, .$$



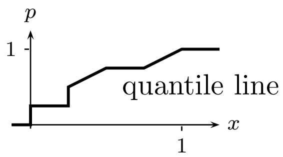
$$\Big(y = 1 + \frac{1}{2} x \Big)$$



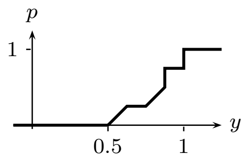
Proof. $(X < x) \iff (aX + b < ax + b)$, therefore $\mathbb{P}\big(X < x \big) = \mathbb{P}\big(aX + b < ax + b \big) = \mathbb{P}\big(Y < y \big)$. Similarly, $\mathbb{P}\big(X \le x \big) = \mathbb{P}\big(Y \le y \big)$. So, $\mathbb{P}\big(X < x \big) \le p \le \mathbb{P}\big(X \le x \big)$ if and only if $\mathbb{P}\big(Y < y \big) \le p \le \mathbb{P}\big(Y \le y \big)$. □

If $a < 0$ then

(3a2) $$\big(\, x \text{ is a } p\text{-quantile of } X \, \big) \iff \big(\, y \text{ is a } (1-p)\text{-quantile of } Y \, \big) \, .$$



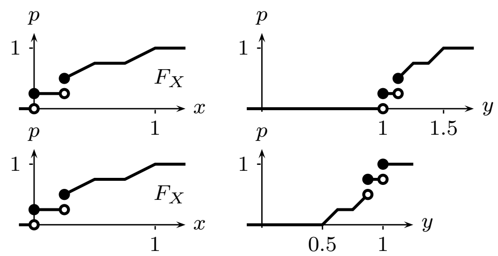
$$\Big(y = 1 - \frac{1}{2} x \Big)$$



Proof. $(X < x) \iff (aX + b > ax + b)$, therefore $\mathbb{P}\big(X < x \big) = \mathbb{P}\big(aX + b > ax + b \big) = \mathbb{P}\big(Y > y \big) = 1 - \mathbb{P}\big(Y \le y \big)$. Similarly, $\mathbb{P}\big(X \le x \big) = 1 - \mathbb{P}\big(Y < y \big)$. So, $\mathbb{P}\big(X < x \big) \le p \le \mathbb{P}\big(X \le x \big)$ if and only if $1 - \mathbb{P}\big(Y \le y \big) \le p \le 1 - \mathbb{P}\big(Y < y \big)$, which means $\mathbb{P}\big(Y < y \big) \le 1 - p \le \mathbb{P}\big(Y \le y \big)$. □

In terms of cumulative distribution functions,

(3a3)
$$\begin{aligned} F_Y(y) &= F_X(x) \qquad && (y = ax + b, \; a > 0) \, , \\ F_Y(y) &= 1 - F_X(x-) \qquad && (y = ax + b, \; a < 0) \, . \end{aligned}$$



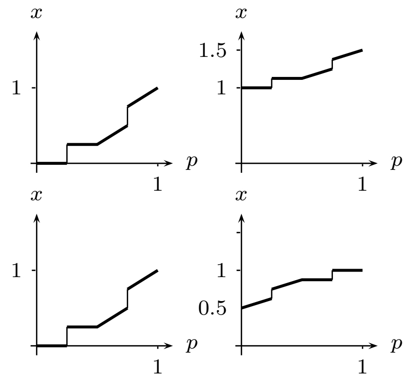
Or, more symmetrically (assuming $y = ax + b$)

(3a4)
$$\begin{aligned} F_Y(y\pm) &= F_X(x\pm) \qquad && \text{when } a > 0 \, , \\ F_Y(y\pm) &= 1 - F_X(x\mp) \qquad && \text{when } a < 0 \, ; \end{aligned}$$

do not forget that $F(u+) = F(u)$.

In terms of quantile functions,

(3a5)
$$
\begin{aligned}
Y^*(p\pm) &= aX^*(p\pm) + b \qquad &&\text{when } a > 0\,, \\
Y^*(p\pm) &= aX^*((1-p)\mp) + b \qquad &&\text{when } a < 0\,;
\end{aligned}
$$



it follows easily from (3a1), (3a2) and the fact that $F_X(x-) \le p \le F_X(x+) \iff X(p-) \le x \le X(p+)$ (recall (2e8)).

If X has a density f_X, then Y also has a density f_Y, and

(3a6)
$$|a| f_Y(y) = f_X(x) \qquad (y = ax + b,\ a \ne 0)\,.$$

Proof. [35] The case $a > 0$: $\mathbb{P}\big(u < Y < v\big) = \mathbb{P}\big(u < aX + b < v\big) = \mathbb{P}\big(\frac{u-b}{a} < X < \frac{v-b}{a}\big) = \int_{(u-b)/a}^{(v-b)/a} f_X(x)\,dx$; a change of variable $y = ax+b$, $x = (y-b)/a$, $dx = \frac1a dy$ gives $\mathbb{P}\big(u < Y < v\big) = \int_u^v f_X(\frac{y-b}{a})\frac1a\,dy$; so, the function f_Y defined by $f_Y(y) = \frac1a f_X(\frac{y-b}{a})$ is a density of Y. □

3a7 Exercise. Prove it for the other case, $a < 0$.

Hint: the sign is changed twice; first, $dx = \frac1a\,dy = -\frac{1}{|a|}\,dy$; second, $\int_{(u-b)/a}^{(v-b)/a} = -\int_{(v-b)/a}^{(u-b)/a}$.

3a8 Exercise. Assuming smoothness, derive (3a6) by differentiating equalities

(3a9)
$$
\begin{aligned}
F_Y(y) &= F_X(x) \qquad &&\text{when } a > 0\,, \\
F_Y(y) &= 1 - F_X(x) \qquad &&\text{when } a < 0\,;
\end{aligned}
\qquad (y = ax + b)
$$

these are (3a3) for the case of continuous F_X, F_Y.

Is density a substitute for probabilities of points? True, $f_X(\cdot)$ is a function of a point, not a set function (in contrast to $P_X(\cdot)$). However,

$$
\begin{aligned}
\mathbb{P}\big(Y = y\big) &= \mathbb{P}\big(X = x\big), \\
|a| f_Y(y) &= f_X(x)\,;
\end{aligned}
\qquad (y = ax + b,\ a \ne 0)
$$

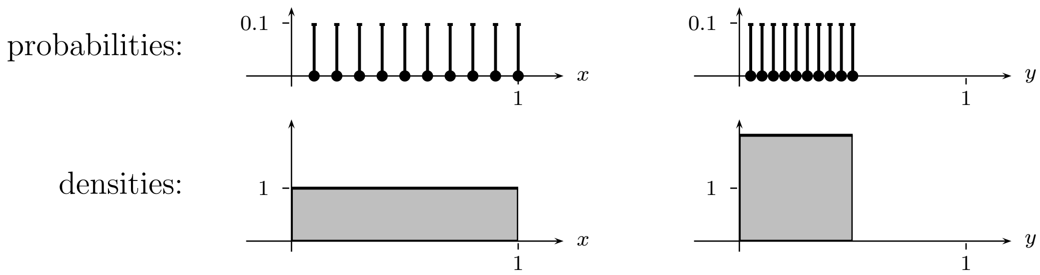
the coefficient $|a|$ is an essential distinction between $\mathbb{P}\big(X = x\big)$ and $f_X(x)$. Intuitively,

$$f_Y(y)\,|dy| = dp = f_X(x)\,|dx|\,; \qquad dy = a\,dx\,.$$

[35] We'll prove only that there is *some* density of Y satisfying $|a| f_Y(y) = f_X(x)$. Both densities may be changed arbitrarily on a set of zero measure.

A density is not a probability; rather, it is the quotient of (infinitesimal) measures, namely, (probability)/(Lebesgue measure). Here are two examples, discrete and continuous, for $Y = \frac{1}{2}X$:

(3a10)



Note that a density can exceed 1; moreover, it can be unbounded (recall 2c3).

3a11 Exercise. Prove that

$$\begin{aligned} (\ x \text{ is an atom for } X\ ) &\iff (\ y \text{ is an atom for } Y\ )\,, \\ (\ x \text{ belongs to the support of } X\ ) &\iff (\ y \text{ belongs to the support of } Y\ ) \end{aligned}$$

assuming $y = ax + b$, $a \ne 0$.

3b Monotone transformations

As you know, a function is called monotone, if it either increases or decreases (that is, either increases everywhere, or decreases everywhere). Note that a monotone function need not be continuous, and a continuous function need not be monotone. Note also the distinction between 'monotone' and 'strictly monotone'.

Let two random variables $X, Y : \Omega \to \mathbb{R}$ be related by the equality

$$Y = \varphi(X)\,, \qquad \text{that is,} \quad \forall \omega \in \Omega \;\; Y(\omega) = \varphi(X(\omega))\,,$$

which may be written also as $Y = \varphi \circ X$.

If $\varphi : \mathbb{R} \to \mathbb{R}$ is increasing and invertible (that is, strictly increasing, continuous, and $\varphi(-\infty) = -\infty$, $\varphi(+\infty) = +\infty$), then all said in 3a for the case $a > 0$ remains true (of course, after replacing $ax + b$ by $\varphi(x)$), except for densities. These require smoothness of φ, giving

(3b1) $$|\varphi'(x)| f_Y(y) = f_X(x) \qquad \text{when } y = \varphi(x)\,.$$

In the same sense the case of a decreasing invertible φ generalizes 3a for $a < 0$.

3b2 Exercise. Prove (3b1) in two ways, by variable change in an integral, and by differentiating a distribution function.

Similar statements hold for a monotone invertible $\varphi : (a, b) \to (c, d)$, provided that $\mathbb{P}\big(\, a < X < b \,\big) = 1$.

In general, an increasing φ may have discontinuities (jumps) and constant intervals (flats). Still,

$$Y = \varphi(X) \quad \Longrightarrow \quad Y^* = \varphi(X^*) \qquad \text{for all increasing } \varphi$$

in the sense that $\varphi(X^*)$ is a quantile function of Y (since it is an increasing function distributed like Y). Other relations may be violated.

3b3 Exercise. Consider the monotone transformation

$$Y = \operatorname{sgn} X = \begin{cases} -1 & \text{when } X < 0, \\ 0 & \text{when } X = 0, \\ +1 & \text{when } X > 0 \end{cases}$$

of an arbitrary X. Show that $Y^* = \varphi(X^*)$. Find F_Y in terms of F_X. What can you say about atoms, supports, and densities?

3b4 Exercise. Draw a picture similar to (3a10) for $Y = X^2$. What happens near the origin?

3c Non-monotone transformations

A simple example: $X \sim \mathrm{U}(-1,+1)$ and $Y = X^2$. The function $\varphi(x) = x^2$ is not monotone. Because of that, Y^* has nothing in common with $\varphi(X^*)$. Say, $X^*\big(\frac12\big) = \mathrm{Me}(X) = 0$, however, $\varphi(0) = 0$ is less than $Y^*\big(\frac12\big) = \mathrm{Me}(Y)$ (and any other $Y^*(p)$).

Here, a single $y \in (0,1)$ corresponds to two values of x, namely, $x_1 = -\sqrt{y}$, $x_2 = +\sqrt{y}$. Both contribute to the density of Y:

$$f_Y(y) = \frac{f_X(x_1)}{|\varphi'(x_1)|} + \frac{f_X(x_2)}{|\varphi'(x_2)|} \,.$$

Similarly, if a smooth φ has n intervals of monotonicity, then f_Y is a sum of n terms.

The bizarre distribution of Example 2b8 results from a simple discrete distribution by a non-monotone transformation $Y = \sin X$. Atoms of Y correspond to atoms of X, but F_Y strongly increases on $[-1,1]$ in contrast to the step function F_X, and Y^* is continuous, in contrast to the step function X^*.

3d Borel functions

If φ is an *arbitrary* function and X is a random variable, then $Y = \varphi(X)$ is a function $\Omega \to \mathbb{R}$ but, in general, not a random variable, since the set $\{\omega \in \Omega : Y(\omega) \le y\}$ need not belong to the σ-field $\mathcal{F}$ of events (recall 2a3).

3d1 Definition. A function $\varphi : \mathbb{R} \to \mathbb{R}$ is *Borel measurable,* or a *Borel function,* if[36]

$$\forall y \in \mathbb{R} \quad \{x \in \mathbb{R} : \varphi(x) \le y\} \in \mathcal{B} \,.$$

[36]Recall that $\mathcal{B}$ stands for the σ-field of all Borel subsets of $\mathbb{R}$.

Sometimes we use the probability space $(\Omega, \mathcal{F}, P) = (\mathbb{R}, \mathcal{B}, P)$ (choosing a probability measure P on $(\mathbb{R}, \mathcal{B})$); in such a case random variables $\Omega \to \mathbb{R}$ are just Borel functions $\mathbb{R} \to \mathbb{R}$ (irrespective of P). Indeed, 3d1 is a special case of 2a3 for $(\Omega, \mathcal{F}) = (\mathbb{R}, \mathcal{B})$.

3d2 Exercise. If $\varphi : \mathbb{R} \to \mathbb{R}$ is a Borel function and $B \subset \mathbb{R}$ a Borel set, then the set $\varphi^{-1}(B) = \{x \in \mathbb{R} : \varphi(x) \in B\}$ is also a Borel set. Prove it. (Hint: use 2d1.)

It means that the next definition is equivalent to 3d1.

3d3 Definition. A function $\varphi : \mathbb{R} \to \mathbb{R}$ is a Borel function, if for every Borel set $B \subset \mathbb{R}$ its inverse image $\varphi^{-1}(B) = \{x \in \mathbb{R} : \varphi(x) \in B\}$ is also a Borel set.

3d4 Exercise. If $X : \Omega \to \mathbb{R}$ is a random variable and $\varphi : \mathbb{R} \to \mathbb{R}$ a Borel function, then $Y : \Omega \to \mathbb{R}$ defined by $\forall \omega \;\; Y(\omega) = \varphi\big(X(\omega)\big)$ is also a random variable. Prove it. (Hint: combine 2a3, 3d1 and 2d1.)

3d5 Exercise. Every continuous function $\varphi : \mathbb{R} \to \mathbb{R}$ is a Borel function. Prove it. (Hint: the set $(-\infty, y]$ is closed, therefore its inverse image is also closed. Use 1f10.)

3d6 Exercise. Every monotone function $\varphi : \mathbb{R} \to \mathbb{R}$ is a Borel function. Prove it. (Hint: the set $(-\infty, y]$ is a ray, therefore its inverse image is also a ray.) Generalize it for piecewise monotone functions. Does 3d6 follow from 3d5? Does 3d5 follow from 3d6?

3d7 Exercise. If $\varphi, \psi : \mathbb{R} \to \mathbb{R}$ are Borel functions, then the function $\xi : \mathbb{R} \to \mathbb{R}$ defined by $\forall x \;\; \xi(x) = \psi\big(\varphi(x)\big)$ is also a Borel function. Prove it. (Hint: use 3d3.)

3d8 Exercise. If $\varphi : \mathbb{R} \to \mathbb{R}$ is a Borel function, then $x \mapsto \varphi^2(x)$ (it means $\big(\varphi(x)\big)^2$, of course) is a Borel function, and $x \mapsto \varphi(x^2)$ is a Borel function. Prove it. What about $1/\varphi(x)$, $\sin \varphi(x)$, $\sqrt{\varphi(x)}$?

3d9 Proposition. If $\varphi, \psi : \mathbb{R} \to \mathbb{R}$ are Borel functions, then $\varphi + \psi$ (that is, $x \mapsto \varphi(x) + \psi(x)$) is also a Borel function. The same for $a\varphi + b\psi$ ($a, b \in \mathbb{R}$), for $\varphi\psi$, and for φ/ψ provided that $\forall x \;\; \psi(x) \neq 0$.

There is a simple and natural proof; it uses Borel maps $\mathbb{R}^2 \to \mathbb{R}$ such as $(x, y) \mapsto x + y$; we'll return to the point later (in 5a).

3d10 Exercise. Let $\varphi_1, \varphi_2, \dots : \mathbb{R} \to \mathbb{R}$ be Borel functions, $\varphi : \mathbb{R} \to \mathbb{R}$, and $\varphi_n(x) \uparrow \varphi(x)$ for every $x \in \mathbb{R}$. Then φ is a Borel function.[37] Prove it. (Hint: $\{x : \varphi(x) \le y\} = \bigcap_n \{x : \varphi_n(x) \le y\}$.) What about a decreasing sequence?

3d11 Exercise. Let $\varphi_1, \varphi_2, \dots : \mathbb{R} \to \mathbb{R}$ be Borel functions, $\varphi : \mathbb{R} \to \mathbb{R}$, and $\varphi_n(x) \to \varphi(x)$ for every $x \in \mathbb{R}$. Then φ is a Borel function.[38] Prove it.
(Hint: $\varphi(x) = \lim_{n\to\infty} \sup\{\varphi_n(x), \varphi_{n+1}(x), \dots\}$; apply 3d10 twice.)

[37] Note that convergence need not be uniform.

[38] Note that convergence need not be uniform.

3d12 Exercise. Calculate the function

$$\varphi(x) = \lim_{n\to\infty} \lim_{k\to\infty} \cos^{2k}(\pi n! x) \,.$$

Is φ a Borel function?

3d13 Exercise. Prove that the random variable Y of Example 2b6 is well-defined, that is, the function $Y : (0,1) \to \mathbb{R}$ defined there is indeed a Borel function. (Hint: use 3d11.) The same for Example 2b7.

It is quite difficult, to construct an example of a non-Borel function.[39] Having a single, explicitly defined function, you may be pretty sure that it is a Borel function (unless its definition is terribly complicated, or uses uncountably many arbitrary choices).

Now we may treat bizarre random variables of Examples 2b6, 2b7 in two ways. One way: Y is a random variable defined on the probability space $(0,1)$ (with Lebesgue measure); $X(\omega) = \omega$, while Y is a bizarre Borel function $(0,1) \to \mathbb{R}$. The other way: X, Y are random variables defined on an arbitrary probability space; $X \sim \mathrm{U}(0,1)$, while $Y = \varphi(X)$ where $\varphi : (0,1) \to \mathbb{R}$ is a bizarre Borel function.

3d14 Exercise. Let X, Y be identically distributed random variables[40] (possibly, on different probability spaces), and $\varphi : \mathbb{R} \to \mathbb{R}$ a Borel function. Then random variables $\varphi(X), \varphi(Y)$ are identically distributed. Prove it. (Hint: find the distribution $P_{\varphi(X)}$ in terms of P_X.)

3d15 Exercise. Let X be a random variable and $\varphi : \mathbb{R} \to \mathbb{R}$ a Borel function. Then the three random variables

$$\varphi(X) \,, \qquad \varphi(X^*) \,, \qquad \big(\varphi(X)\big)^*$$

are identically distributed. Prove it. (Hint: use 2e8 and 3d14.) What about $\big(\varphi(X^*)\big)^*$? What about a monotone φ?

[39]Existence follows from the (non-evident) fact that the set of all Borel functions is of cardinality continuum, while the set of *all* functions is of a higher cardinality.

[40]Recall 2d7.