3 Transformations

3a Linear transformations

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

$$Y = aX + b$$

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

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, $\mathbb{P}\big(X < x \big) \le p \le \mathbb{P}\big(X \le x \big)$ holds 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) $$\big(\, x \text{ is a } p\text{-quantile of } X \, \big) \iff \big(\, y \text{ is a } (1-p)\text{-quantile of } Y \, \big) \, .$$

Proof. $(X < x) \iff (aX + b > ax + b)$, therefore $\mathbb{P}\big(X < x \big) = \mathbb{P}\big(aX + b > ax + b \big) = 1 - \mathbb{P}\big(Y < y \big)$. Similarly, $\mathbb{P}\big(X \le x \big) = 1 - \mathbb{P}\big(Y < y \big)$. So, $\mathbb{P}\big(X < x \big) \le p \le \mathbb{P}\big(X \le x \big)$ holds 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) $$\begin{aligned} &F_Y(y) = F_X(x) && (y = ax + b, \; a > 0) \, , \\ &F_Y(y) = 1 - F_X(x-) && (y = ax + b, \; a < 0) \, . \end{aligned}$$

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

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

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

In terms of quantile functions,

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

it follows easily from (3a1), (3a2) and 2b12.

Let $Y = aX + b$ and $Y_n = aX_n + b$ for $n = 1, 2, \dots$ Then

(3a6) $$\big(X_n \xrightarrow[n\to\infty]{\text{distr}} X \big) \iff \big(Y_n \xrightarrow[n\to\infty]{\text{distr}} Y \big) \, ; \qquad (a \ne 0)$$

it follows easily from (3a4), (2c10) and the fact that $(x_n \to x) \iff (ax_n + b \to ax + b)$.

Smooth case: let X, Y have their densities $f_X(x) = F'_X(x)$, $f_Y(y) = F'_Y(y)$. Differentiating (3a3) we get

$$\begin{aligned} af_Y(y) &= f_X(x) \qquad (y = ax+b,\ a>0)\,, \\ -af_Y(y) &= f_X(x) \qquad (y = ax+b,\ a<0)\,, \end{aligned}$$

that is,

(3a7) $$|a| f_Y(y) = f_X(x) \qquad (y = ax+b,\ a \ne 0)\,.$$

Do not forget the factor $|a|$, responsible for normalization: $\int f_X(x)\,dx = 1$, $\int f_Y(y)\,dy = 1$. Heuristically, $f_X(x) = \left|\frac{dp}{dx}\right|$, $f_Y(y) = \left|\frac{dp}{dy}\right|$, $a = \frac{dy}{dx}$, and so, $\left|\frac{dy}{dx}\right| \cdot \left|\frac{dp}{dy}\right| = \left|\frac{dp}{dx}\right|$.

Atoms, gaps, support (assuming $y = ax + b$, $a \ne 0$):

(3a8) $$\begin{aligned} (\ x \text{ is an atom for } X\ ) &\iff (\ y \text{ is an atom for } Y\ )\,, \\ (\ x \text{ belongs to the support of } X\ ) &\iff (\ y \text{ belongs to the support of } Y\ )\,. \end{aligned}$$

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.

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 "smooth case". The latter requires 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 1a for $a < 0$.

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. An example:

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

Note that $X_n \xrightarrow[n\to\infty]{\text{distr}} X$ does not imply $\operatorname{sgn} X_n \xrightarrow[n\to\infty]{\text{distr}} \operatorname{sgn} X$ (unless $\mathbb{P}\big(\, X = 0 \,\big) = 0$).

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, $X^*\bigl(\frac12\bigr) = \mathrm{Me}(X) = 0$, however, $\varphi(0) = 0$ is less than $Y^*\bigl(\frac12\bigr) = \mathrm{Me}(Y)$ (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.

A more complicated example: $Y = \sin X$, and X takes on integers with some positive probabilities (say, $X \sim \mathrm{Geom}(1/2)$). Here Y has a countable set of atoms, $\{\sin k : k = 0,1,2,\dots\}$, dense on $[-1,1]$. Both F_Y and Y^* are crasy; F_Y increases strongly on $[-1,1]$, having a dense set of jumps (and no density); Y^* has a dense set of flats, and is continuous.

For now we restrict ourselves to transformations that are piecewise monotone with finite number of pieces. For more general φ the question, whether $\varphi(X)$ is a random variable or not, is non-evident. We'll return to the point later.