Chapter 4. Continuous Random Variables

4.4: Transforming Continuous RVs

(From "Probability & Statistics with Applications to Computing" by Alex Tsun)

Suppose the amount of gold a company can mine is X tons per year, and you have some (continuous) distribution to model this. However, your earning is not simply X - it is actually a function of the amount of product, some $Y = g(X)$. What is the distribution of Y?

Since we know the distribution of X, this will help us model the distribution of Y by *transforming random variables*.

4.4.1 Transforming 1-D (Continuous) RVs via CDF

When we are dealing with discrete random variables, this process wasn't too bad. Let's say X had range $\{-1, 0, 1\}$ and PMF

$$p_X(x) = \begin{cases} 0.3 & x = -1 \\ 0.2 & x = 0 \\ 0.5 & x = 1 \end{cases}$$

and $Y = g(X) = X^2$. Then, $\Omega_Y = \{0, 1\}$, and we could say

$$p_Y(y) = \begin{cases} p_X(-1) + p_X(1) = 0.3 + 0.5 = 0.8 & y = 1 \\ p_X(0) = 0.2 & y = 0 \end{cases}$$

This is because $Y = 1$ if and only if $X \in \{-1, 1\}$, so to find $\mathbb{P}(Y = 1)$, we sum over all values x such that $x^2 = 1$ of its probability. That's all this formula below says (the ":" means "such that"):

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

But for continous random variables, we have density functions instead of mass functions. That means f_X is not actually a probability and so we can't do this same technique. We want to work with the CDF $F_X(x) = \mathbb{P}(X \leq x)$ instead because it actually does represent a probability! It's best to see this idea through an example.

Example(s)

Suppose you know $X \sim \text{Unif}(0, 9)$ (continuous). What is the PDF of $Y = \sqrt{X}$?

Solution We know the range of X,

$$\Omega_X = [0, 9]$$

We also know the PDF of X, which is uniform from 0 to 9, and 0 elsewhere.

$$f_X(x) = \begin{cases} \frac{1}{9} & \text{if } 0 \leq x \leq 9 \\ 0 & \text{otherwise} \end{cases}$$

The CDF of X is derived by taking the integral of the PDF, giving us (can also cite this),

$$F_X(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{x}{9} & \text{if } 0 \le x \le 9 \\ 1 & \text{if } x > 9 \end{cases}$$

Now, we determine the range of Y. The smallest value that Y can take is $\sqrt{0} = 0$, and the largest value that Y can take is $\sqrt{9} = 3$, from the range of X. Since the square root function is monotone increasing, this gives us,

$$\Omega_Y = [0, 3]$$

But can we assume that, because X has a uniform distribution, Y does too?

This is not the case! Notice that values of X in the range $[0, 1]$ will map to Y values in the range $[0, 1]$. But, X values in the range $[1, 4]$ map to Y values in the range $[1, 2]$ and X values in the range $[4, 9]$ map to Y values in the range $[2, 3]$.

So, there is a much larger range of values of X that map to $[2, 3]$ than to $[0, 1]$ (since $[4, 9]$ is a larger range than $[0, 1]$). Therefore, Y's distribution shouldn't be uniform. So, we cannot define the PDF of Y using the assumption that Y is uniform.

Instead, we will first compute the CDF F_Y and then, differentiate that to get the PDF f_Y for $y \in [0, 3]$.

To compute F_Y for any y in $[0, 3]$, we first take the CDF at y:

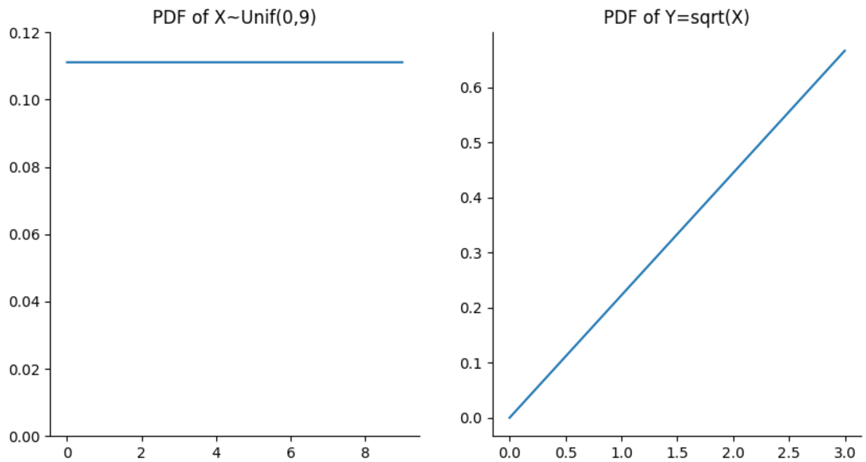
$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \le y) && \text{[def of CDF]} \\ &= \mathbb{P}\left(\sqrt{X} \le y\right) && \text{[def of } Y\text{]} \\ &= \mathbb{P}\left(X \le y^2\right) && \text{[squaring both sides]} \\ &= F_X(y^2) && \text{[def of CDF of } X \text{ evaluated at } y^2\text{]} \\ &= \frac{y^2}{9} && \text{[plug in CDF of } X\text{, since } y^2 \in [0, 9]\text{]} \end{aligned}$$

Be very careful when squaring both sides of an equation - it may not keep the inequality true. In this case we didn't have to worry since X and Y were both guaranteed positive.

Differentiating the CDF to get the PDF f_Y, for $y \in [0, 3]$,

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{2y}{9}$$

Here is an image of the original and transformed PDFs! Remember that $X \sim \text{Unif}(0, 9)$ and $Y = \sqrt{X}$.



□

This is the general strategy for transforming continuous RVs! We'll summarize the steps below.

Definition 4.4.1: Steps to get PDF of $Y = g(X)$ from X (via CDF)

1. Write down the range Ω_X, PDF f_X, and CDF F_X.
2. Compute the range $\Omega_Y = \{g(x) : x \in \Omega_X\}$.
3. Start computing the CDF of Y on Ω_Y, $F_Y(y) = \mathbb{P}(g(X) \leq y)$, in terms of F_X.
4. Differentiate the CDF $F_Y(y)$ to get the PDF $f_Y(y)$ on Ω_Y. f_Y is 0 outside Ω_Y.

Example(s)

Let X be continuous with range $\Omega_X = [-1, +1]$ have density function

$$f_X(x) = \begin{cases} \frac{3}{4}(1-x^2) & -1 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Suppose $Y = X^4$. Find the density function $f_Y(y)$.

Solution We'll follow the 4-step procedure as outlined above.

1. First, we list out the range, PDF, and CDF of the original variable X. We were given the range and PDF, but not the CDF, so let's compute it. For $x \in [-1, +1]$ (note the use of the dummy variable t since x is already taken),

$$F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f_X(t)dt = \int_{-1}^{x} \frac{3}{4}(1-t^2)dt = \frac{1}{4}(2+3x-x^3)$$

So the complete CDF is:

$$F_X(x) = \begin{cases} 0 & x \leq -1 \\ \frac{1}{4}(2+3x-x^3) & -1 \leq x \leq 1 \\ 1 & x \geq 1 \end{cases}$$

2. The range of $Y = X^4$ is $\Omega_Y = \{x^4 : x \in [-1, +1]\} = [0, 1]$, since x^4 is always positive and between 0 and 1 for $x \in [-1, +1]$.

3. Be careful in the third equation below to include *both* lower and upper bounds (draw the function $y = x^4$ to see why). For $y \in \Omega_Y = [0, 1]$, we will compute the CDF:

$$\begin{aligned}
F_Y(y) &= \mathbb{P}(Y \le y) && \text{[def of CDF]} \\
&= \mathbb{P}\left(X^4 \le y\right) && \text{[def of } Y\text{]} \\
&= \mathbb{P}\left(-\sqrt[4]{y} \le X \le \sqrt[4]{y}\right) && \text{[don't forget the negative side]} \\
&= \mathbb{P}\left(X \le \sqrt[4]{y}\right) - \mathbb{P}\left(X \le -\sqrt[4]{y}\right) \\
&= F_X(\sqrt[4]{y}) - F_X(-\sqrt[4]{y}) && \text{[def of CDF of } X\text{]} \\
&= \frac{1}{4}(2 + 3\sqrt[4]{y} - \sqrt[4]{y}^3) - \frac{1}{4}(2 + 3(-\sqrt[4]{y}) - (-\sqrt[4]{y})^3) && \text{[plug in CDF]}
\end{aligned}$$

4. The last step is to differentiate the CDF to get the PDF, which is just computational, so I'll skip it!

□

4.4.2 Transforming 1-D RVs via Explicit Formula

Now, it turns out actually that in some special cases, there is an explicit formula for the density function of $Y = g(X)$, and we don't have to go through all the same steps above. It's important to note that the CDF method *can always be applied*, but this next method has restrictions.

Theorem 4.4.1: Formula to get PDF of $Y = g(X)$ from X

If $Y = g(X)$ and $g : \Omega_X \to \Omega_Y$ is **strictly monotone** and **invertible** with inverse $X = g^{-1}(Y) = h(Y)$, then

$$f_Y(y) = \begin{cases} f_X(h(y)) \cdot |h'(y)| & \text{if } y \in \Omega_Y \\ 0 & \text{otherwise} \end{cases}$$

That is, the PDF of Y at y is the PDF of X evaluated at $h(y)$ (the value of x that maps to y) multiplied by the absolute value of the derivative of $h(y)$.

Note that the formula method is not as general as the previous method (using CDF), since g must satisfy monotonicity and invertibility. So transforming via CDF always works, but transforming may not work with this explicit formula all the time.

Proof of Formula to get PDF of $Y = g(X)$ from X.

Suppose $Y = g(X)$ and g is strictly monotone and invertible with inverse $X = g^{-1}(Y) = h(Y)$. We'll assume g is strictly monotone *increasing* and leave it to you to prove it for the case when g is strictly monotone *decreasing* (it's very similar).

$$\begin{aligned}
F_Y(y) &= \mathbb{P}(Y \leq y) && \text{[def of CDF]} \\
&= \mathbb{P}(g(X) \leq y) && \text{[def of } Y\text{]} \\
&= \mathbb{P}\left(X \leq g^{-1}(y)\right) && \text{[invertibility, AND monotone increasing keeps the sign]} \\
&= F_X(g^{-1}(y)) && \text{[def of CDF of } X \text{ evaluated at } g^{-1}(y)\text{]} \\
&= F_X(h(y)) && [h(y) = g^{-1}(y)]
\end{aligned}$$

Hence, by the chain rule (of calculus),

$$f_Y(y) = \frac{d}{dy}F_Y(y) = f_X(h(y)) \cdot h'(y)$$

A similar proof would hold if g were monotone decreasing, except in the third line we would flip the sign of the inequality and make the $h'(y)$ become an absolute value: $|h'(y)|$.

□ Now let's try the same example as we did earlier, but using this new method instead.

Example(s)

Suppose you know $X \sim \text{Unif}(0, 9)$ (continuous). What is the PDF of $Y = \sqrt{X}$?

Solution Recall, we know the range of X,

$$\Omega_X = [0, 9]$$

We also know the PDF of X, which is uniform from 0 to 9 and 0 elsewhere.

$$f_X(x) = \begin{cases} \frac{1}{9} & \text{if } 0 \leq x \leq 9 \\ 0 & \text{otherwise} \end{cases}$$

Our goal is to use the formula given $f_Y(y) = f_X(h(y)) \cdot |h'(y)|$, after verifying some conditions on g.

Let $g(t) = \sqrt{t}$. This is strictly monotone increasing on $\Omega_X = [0, 9]$. This means that as t increases, $\sqrt{t}$ also increases - therefore, $g(t)$ is an increasing function.

What is the inverse of this function g? The inverse of the square root function is just the squaring function:

$h(y) = g^{-1}(y) = y^2$

Then, we find it's derivative:

$h'(y) = 2y$

Now, we can use the explicit formula to find the PDF of Y.

For $y \in [0, 3]$,

$$f_Y(y) = f_X(h(y)) \cdot |h'(y)| = \frac{1}{9}|2y| = \frac{2}{9}y$$

Note that we dropped the absolute value because we already assume $y \in [0, 3]$ and hence $2y$ is always positive. This gives the same formula as earlier, as it should! □

4.4.3 Transforming Multidimensional RVs via Formula

For completion, we've cited a formula to transform n random variables to n other random variables. For example, this might be useful if you have a system of two equations. For example, (R, Θ) (polar) coordinates which are random variables, and wanting to convert to Cartesian coordinates to the two random variables (X, Y) where $X = R\cos(\Theta)$ and $Y = R\sin(\Theta)$. This extends the formula we just learned to multi-dimensional random variables!

Theorem 4.4.2: Formula to get PDF of $Y = g(X)$ from X (Multidimensional Case)

Let $\mathbf{X} = (X_1, ..., X_n)$, $\mathbf{Y} = (Y_1, ..., Y_n)$ be continuous random vectors (each component is a continuous rv) with the same dimension n (so $\Omega_{\mathbf{X}}, \Omega_{\mathbf{Y}} \subseteq \mathbb{R}^n$), and $\mathbf{Y} = g(\mathbf{X})$ where $g : \Omega_{\mathbf{X}} \to \Omega_{\mathbf{Y}}$ is invertible and differentiable, with differentiable inverse $\mathbf{X} = g^{-1}(\mathbf{y}) = h(\mathbf{y})$. Then,

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(h(\mathbf{y})) \left| \det \left(\frac{\partial h(\mathbf{y})}{\partial \mathbf{y}} \right) \right|$$

where $\left(\frac{\partial h(\mathbf{y})}{\partial \mathbf{y}} \right) \in \mathbb{R}^{n \times n}$ is the Jacobian matrix of partial derivatives of h, with

$$\left(\frac{\partial h(\mathbf{y})}{\partial \mathbf{y}} \right)_{ij} = \frac{\partial (h(\mathbf{y}))_i}{\partial \mathbf{y}_j}$$

Hopefully this formula looks very similar to the one for the single-dimensional case! This formula is just for your information and you'll never have to use it in this class.

4.4.4 Exercises

1. Suppose X has range $\Omega_X = (1, \infty)$ and density function

$$f_X(x) = \begin{cases} \frac{2}{x^3} & x > 1 \\ 0 & \text{otherwise} \end{cases}$$

For reference, the CDF is also given

$$F_X(x) = \begin{cases} 1 - \frac{1}{x^2} & x > 1 \\ 0 & \text{otherwise} \end{cases}$$

Let $Y = \frac{e^X - 1}{2}$.

(a) Compute the density function of Y via the CDF transformation method.

(b) Compute the density function of Y using the formula, but explicitly verify the monotonicity and invertibility conditions.

Solution:

(a) The range of Y is $\Omega_Y = \left(\frac{e-1}{2}, \infty\right)$. For $y \in \Omega_Y$,

$$
\begin{aligned}
F_Y(y) &= \mathbb{P}(Y \le y) && \text{[def of CDF]} \\
&= \mathbb{P}\left(\frac{e^X - 1}{2} \le y\right) && \text{[def of } Y\text{]} \\
&= \mathbb{P}\left(e^X \le 2y + 1\right) \\
&= \mathbb{P}(X \le \ln(2y+1)) \\
&= F_X(\ln(2y+1)) && \text{[def of CDF]} \\
&= 1 - \frac{1}{[\ln(2y+1)]^2} && \left[F_X(x) = 1 - \frac{1}{x^2}\right]
\end{aligned}
$$

The derivative is (don't forget the chain rule)

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{2}{[\ln(2y+1)]^3} \cdot \frac{1}{2y+1} \cdot 2 = \frac{4}{(2y+1)[\ln(2y+1)]^3}$$

This density is valid for $y \in \Omega_Y$, and 0 everywhere else.

(b) The function $g(t) = \frac{e^t - 1}{2}$ is monotone increasing (since e^t is, and we shift and scale it by a positive constant), and has inverse $h(y) = g^{-1}(y) = \ln(2y+1)$. We have $h'(y) = \frac{2}{2y+1}$. By the formula, we get

$$
\begin{aligned}
f_Y(y) &= f_X(h(y))|h'(y)| && \text{[formula]} \\
&= \frac{2}{[\ln(2y+1)]^3} \cdot \frac{2}{2y+1} && \left[f_X(x) = \frac{2}{x^3}\right] \\
&= \frac{4}{(2y+1)[\ln(2y+1)]^3}
\end{aligned}
$$

This gives the same answer as part (a)!