ECE 581: Functions of Random Variables

Henry D. Pfister
Duke University

November 15, 2025

1 Discrete Case: PMF and CDF of a Function

Let X be a discrete random variable with PMF $p_X(\cdot)$ and let $Y = g(X)$.

- PMF of Y:
$$p_Y(y) = \sum_{x:g(x)=y} p_X(x).$$
- CDF of Y:
$$F_Y(y) = \Pr(g(X) \leq y) = \sum_{x:g(x)\leq y} p_X(x).$$

Example 1 (Mapping with collapsing values)**.** *Let X take values $\{-2, -1, 0, 1, 2\}$ with probabilities $p_X(-2) = 0.1$, $p_X(-1) = 0.2$, $p_X(0) = 0.4$, $p_X(1) = 0.2$, $p_X(2) = 0.1$ and $g(x) = x^2$. Then*

- $p_Y(0) = p_X(0) = 0.4$
- $p_Y(1) = p_X(-1) + p_X(1) = 0.2 + 0.2 = 0.4$
- $p_Y(4) = p_X(-2) + p_X(2) = 0.1 + 0.1 = 0.2$
- $p_Y(y) = 0$ *otherwise.*

Note how values ± 1 and ± 2 collapse to the same output values under $g(x) = x^2$.

2 Continuous Case: CDF and PDF of a Function

Let X be a continuous random variable with PDF $f_X(\cdot)$, CDF $F_X(\cdot)$, and $Y = g(X)$.

2.1 General CDF Formula for Transformations

For any real y,

$$F_Y(y) = \Pr(g(X) \le y) = \int_{\{u:\, g(u) \le y\}} f_X(u)\, du = \int_{g^{-1}((-\infty, y])} f_X(u)\, du,$$

where $g^{-1}((-\infty, y])$ denotes the inverse image of the real interval $(-\infty, y]$ for the function g. Then, the PDF can be computed by taking the derivative.

Example 2 (Continuous transformation giving point mass)**.** *Let* $X \sim \mathrm{Unif}[0, 1]$ *and define*

$$g(x) = \begin{cases} 0 & \text{if } x \in [0, 1/2] \\ x - 1/2 & \text{if } x \in (1/2, 1] \end{cases}$$

For the CDF of $Y = g(X)$*:*

- *For* $y < 0$*:* $F_Y(y) = 0$
- *For* $y = 0$*:* $F_Y(0) = \Pr(X \in [0, 1/2]) = 1/2$ *(jump discontinuity)*
- *For* $0 < y \le 1/2$*:* $F_Y(y) = 1/2 + \Pr(X - 1/2 \le y, X > 1/2) = 1/2 + y$
- *For* $y > 1/2$*:* $F_Y(y) = 1$

To find the PDF, we differentiate the CDF. At a jump discontinuity, the generalized derivative is a shifted Dirac delta function scaled by value of the jump. Thus, the PDF is $f_Y(y) = \frac{1}{2}\delta(y) + \mathbf{1}_{(0,1/2]}(y)$*, where* δ *is the Dirac delta function representing the point mass at* $y = 0$*. Notice that a function of a random variable is defined by the* inverse images *of the intervals* $(-\infty, y]$ *for* $y \in \mathbb{R}$*.*

2.2 Strictly Monotone Differentiable Transformation

For strictly increasing g, we have

$$F_Y(y) = \Pr(g(X) \le y) = \Pr\left(X \le g^{-1}(y)\right) = F_X\big(g^{-1}(y)\big).$$

Differentiating with respect to y gives

$$f_Y(y) = f_X\big(g^{-1}(y)\big)\, \frac{d}{dy} g^{-1}(y).$$

For strictly decreasing g, we instead get $F_Y(y) = \Pr\left(X \ge g^{-1}(y)\right) = 1 - F_X\big(g^{-1}(y)\big)$, so

$$f_Y(y) = f_X\big(g^{-1}(y)\big) \left(-\frac{d}{dy} g^{-1}(y)\right).$$

Combining both cases yields the following absolute-value form.

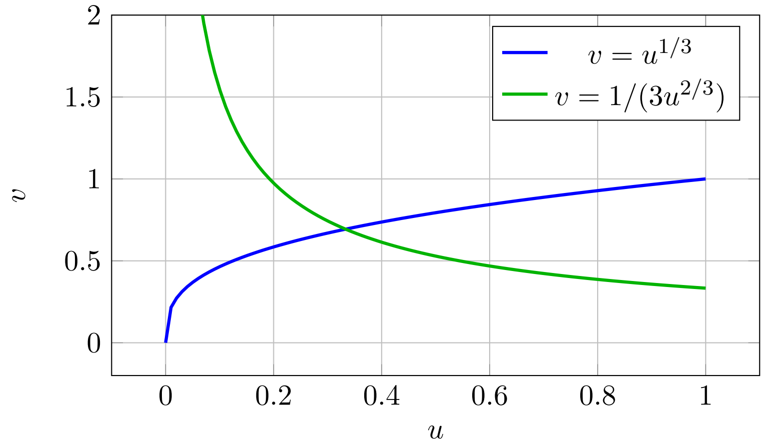
If g is differentiable and strictly monotone so that $x = g^{-1}(y)$ is unique, then

$$f_Y(y) = f_X\big(g^{-1}(y)\big)\left|\frac{d}{dy}g^{-1}(y)\right| = \frac{f_X(x)}{|g'(x)|} \quad \text{with } x = g^{-1}(y).$$

Example 3 (Cubic transformation with visualization)**.** *Let* $X \sim \text{Unif}[0,1]$ *and* $Y = X^3$. *Since* $g(x) = x^3$ *is strictly increasing, we have* $g^{-1}(y) = y^{1/3}$ *and* $\frac{d}{dy}g^{-1}(y) = \frac{1}{3}y^{-2/3}$. *Thus,*

$$f_Y(y) = f_X(y^{1/3}) \cdot \frac{1}{3}y^{-2/3} = 1 \cdot \frac{1}{3}y^{-2/3} = \frac{1}{3y^{2/3}}, \quad y \in (0,1].$$

Curves for transformation $Y = X^3$ *for* $X \sim \text{Unif}[0,1]$



Proposition 1 (Derivative of inverse function)**.** *If* g *is differentiable and strictly monotone with* $g'(x) \neq 0$, *then*

$$\frac{d}{dy}g^{-1}(y) = \frac{1}{g'(g^{-1}(y))} = \frac{1}{g'(x)} \quad \textit{where } x = g^{-1}(y).$$

Proof. By the chain rule applied to $g(g^{-1}(y)) = y$:

$$g'(g^{-1}(y)) \cdot \frac{d}{dy}g^{-1}(y) = 1,$$

which gives the result. □

2.3 Monotone Transformations: CDF Simplifications

When g is monotone but may have flat regions or jump discontinuities, the preimage

$$A_y := \{u : g(u) \leq y\}$$

is an interval (possibly empty or unbounded). Using the right-continuity of F_X, we obtain the correct CDF by taking the boundary of this set: for non-decreasing g, the boundary is $\sup A_y$; for non-increasing g, it is $\inf A_y$. These choices handle the possibilities that many x map to the same

y (flat regions) or that $g^{-1}(y)$ is empty at a discontinuity. If, in addition, g is Lipschitz (hence differentiable almost everywhere), then at points where derivatives exist we may differentiate these CDFs to obtain the density formulas; in the strictly monotone case this reduces to the change-of-variables rule above.

If g is non-decreasing,

$$F_Y(y) = F_X(\sup\{u : g(u) \leq y\}) .$$

If g is non-increasing,

$$F_Y(y) = 1 - F_X(\inf\{u : g(u) \leq y\}) .$$

For example, try applying the first definition to the function defined in Example 2.

2.4 Examples of Continuous Transformations

The following examples include brief derivations illustrating the CDF method, change of variables, and handling multiple preimages.

1. Uniform scaling: Let $X \sim \text{Unif}[0,1]$ and $Y = 2X$. Then for $y \in [0,2]$,

$$F_Y(y) = \Pr\left(X \leq \frac{y}{2}\right) = \frac{y}{2}, \qquad f_Y(y) = \frac{1}{2}, \quad y \in [0,2].$$

2. 2D Gaussian to Rayleigh: Let (X_1, X_2) be independent zero-mean Gaussian random variables with variance σ^2, so $f_{X_i}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/(2\sigma^2)}$. Let $R = \sqrt{X_1^2 + X_2^2}$.

   To find the CDF of R, we compute:

$$\begin{aligned} F_R(r) &= \Pr(R \leq r) = \Pr(X_1^2 + X_2^2 \leq r^2) && (1) \\ &= \int_{x_1^2+x_2^2 \leq r^2} f_{X_1}(x_1) f_{X_2}(x_2)\, dx_1 dx_2 && (2) \\ &= \int_{x_1^2+x_2^2 \leq r^2} \frac{1}{2\pi\sigma^2} e^{-(x_1^2+x_2^2)/(2\sigma^2)}\, dx_1 dx_2 && (3) \end{aligned}$$

   Converting to polar coordinates with $x_1 = \rho\cos\theta$, $x_2 = \rho\sin\theta$, and $dx_1 dx_2 = \rho\, d\rho d\theta$:

$$\begin{aligned} F_R(r) &= \int_0^{2\pi} \int_0^r \frac{1}{2\pi\sigma^2} e^{-\rho^2/(2\sigma^2)} \rho\, d\rho d\theta && (4) \\ &= \frac{1}{\sigma^2} \int_0^r \rho e^{-\rho^2/(2\sigma^2)}\, d\rho && (5) \end{aligned}$$

   Using substitution $u = \rho^2/(2\sigma^2)$, so $du = \rho d\rho/\sigma^2$:

$$F_R(r) = \int_0^{r^2/(2\sigma^2)} e^{-u}\, du = 1 - e^{-r^2/(2\sigma^2)}, \quad r > 0.$$

   Differentiating gives $f_R(r) = \frac{r}{\sigma^2} e^{-r^2/(2\sigma^2)}$ for $r > 0$, which is the Rayleigh distribution with scale parameter σ.

3. Rayleigh to exponential: Let X be Rayleigh with scale parameter $\sigma = 1$, i.e., $f_X(x) = xe^{-x^2/2}$ for $x \geq 0$, and let $Y = X^2$. Here $g(x) = x^2$, $g'(x) = 2x$, and for $y \geq 0$, $x = \sqrt{y}$. Then

$$f_Y(y) = \frac{f_X(\sqrt{y})}{|g'(\sqrt{y})|} = \frac{\sqrt{y}\, e^{-y/2}}{2\sqrt{y}} = \frac{1}{2}e^{-y/2}, \quad y \geq 0,$$

i.e., Y is exponential with mean 2.

4. Affine transform of Gaussian: If $X \sim \mathcal{N}(0,1)$ and $Y = aX + b$ with $a \neq 0$, then

$$f_Y(y) = \frac{1}{|a|\sqrt{2\pi}} \exp\left(-\frac{(y-b)^2}{2a^2}\right),$$

so Y is Gaussian with mean b and variance a^2.

5. Cosine of a uniform phase (arcsine law): Let $X \sim \text{Unif}[0, 2\pi)$ and $Y = \cos X$.

For $y \in (-1, 1)$, we need to find all $x \in [0, 2\pi)$ such that $\cos x = y$. These are:

- $x_1 = \arccos y \in [0, \pi]$
- $x_2 = 2\pi - \arccos y \in [\pi, 2\pi)$

First, let's find the CDF using the direct method. For $y \in (-1, 1)$, we need:

$$F_Y(y) = \Pr(\cos X \leq y) = \Pr(X \in \{x \in [0, 2\pi) : \cos x \leq y\})$$

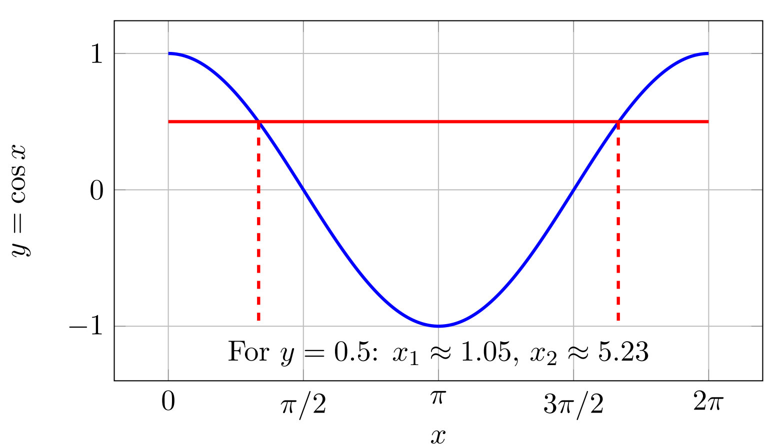
The set $\{x \in [0, 2\pi) : \cos x \leq y\}$ consists of the interval $[\arccos y, 2\pi - \arccos y]$. Since $X \sim \text{Unif}[0, 2\pi)$, we have:

$$F_Y(y) = \frac{(2\pi - \arccos y) - \arccos y}{2\pi} = \frac{2\pi - 2\arccos y}{2\pi} = 1 - \frac{\arccos y}{\pi}$$

Differentiating with respect to y:

$$f_Y(y) = \frac{d}{dy}F_Y(y) = -\frac{1}{\pi} \cdot \frac{d}{dy}\arccos y = -\frac{1}{\pi} \cdot \frac{-1}{\sqrt{1-y^2}} = \frac{1}{\pi\sqrt{1-y^2}}, \quad -1 < y < 1.$$

$Y = \cos X$ for $X \sim \text{Unif}[0, 2\pi)$



3 Generating Random Variables via Inverse Transform

3.1 Continuous Case: Inverse CDF Method

Let X be a continuous random variable where F_X is continuous and strictly increasing. Then, we can define $U = F_X(X)$ and observe that, for $u \in [0, 1]$, we have

$$\Pr(U \le u) = \Pr\big(F_X(X) \le u\big) = \Pr\big(X \le F_X^{-1}(u)\big) = F_X\big(F_X^{-1}(u)\big) = u.$$

This implies that $U \sim \mathrm{Unif}(0, 1)$. Conversely, if $U \sim \mathrm{Unif}(0, 1)$ and $V = F_X^{-1}(U)$, then for any x,

$$\Pr(V \le x) = \Pr(F_X^{-1}(U) \le x) = \Pr\big(U \le F_X(x)\big) = F_X(x),$$

so V has CDF F_X (and PDF f_X when it exists).

In summary, if F_X is a continuous, strictly increasing CDF, then:

- If $U = F_X(X)$, with X having CDF F_X, then $U \sim \mathrm{Unif}(0, 1)$.
- Conversely, if $U \sim \mathrm{Unif}(0, 1)$ and $V = F_X^{-1}(U)$, then V has PDF f_X.

Example 4 (Exponential with mean $\lambda > 0$)**.** *$F_X(x) = 1 - e^{-\lambda x}$ for $x \ge 0$. Then $F_X^{-1}(u) = -\frac{1}{\lambda}\log(1 - u)$. Thus, if $U \sim \mathrm{Unif}(0, 1)$,*

$$X = -\frac{1}{\lambda}\log(1 - U) \sim \mathrm{Exp}(\lambda).$$

Example 5 (Standard Gaussian via Box-Muller)**.** *For a standard Gaussian X, $F_X(x)$ does not have closed-form inverse CDF. Thus, to generate a 2D standard Gaussian (X_1, X_2), a key insight is to use polar coordinates (R, θ). As noted in Section 2.4, the radius $R = \sqrt{X_1^2 + X_2^2}$ is Rayleigh and the angle θ is uniform on $[0, 2\pi)$. Moreover, the variables ar independent because their joint density, in polar coordinates, separates into a product form since θ is uniform.*

If the Rayleigh variable is generated from a uniform, then this method of generating two standard Gaussians is known as the Box-Muller transform.

Formally, if (X_1, X_2) are independent standard normals, then, in polar coordinates:

- *$R = \sqrt{X_1^2 + X_2^2}$ follows a Rayleigh distribution with $f_R(r) = re^{-r^2/2}$ for $r > 0$*
- *$\Theta = \arctan(X_2/X_1)$ is uniform on $[0, 2\pi)$ and independent of R*

From the inverse transform method:

- *For the Rayleigh: $F_R(r) = 1 - e^{-r^2/2}$, so $R = \sqrt{-2\log(1 - U_1)} = \sqrt{-2\log U_1}$ (since $1 - U_1$ has the same distribution as U_1)*
- *For the uniform angle: $\Theta = 2\pi U_2$*

Converting back to Cartesian coordinates:

$$X_1 = R\cos\Theta = \sqrt{-2\log U_1}\cos(2\pi U_2) \tag{6}$$

$$X_2 = R\sin\Theta = \sqrt{-2\log U_1}\sin(2\pi U_2) \tag{7}$$

If $U_1, U_2 \sim \text{Unif}(0,1)$ *are independent, then* (X_1, X_2) *are independent standard normal random variables.*

3.2 Not Strictly Increasing

When F_X is not strictly increasing, the generalized (maximal and right-continuous) inverse can be defined in two equivalent ways,

$$F^{-1}(u) := \inf\{x : F(x) \geq u\} = \sup\{x : F(x) < u\}, \quad u \in (0,1). \tag{8}$$

The first form using the infimum is the most common because defines the quantile function of the distribution. For generating random variables, the same arguments hold with these definitions.

Proposition 2 (Generalized inverse)**.** *For any CDF* F*, the generalized (right-continuous) inverse defined by* (8) *satisfies:*

- $F^{-1}(F(x)) \leq x$ *for all* x
- $F(F^{-1}(u)) \geq u$ *for all* $u \in (0,1)$
- *If* F *is continuous and strictly increasing, then* $F^{-1}(F(x)) = x$ *and* $F(F^{-1}(u)) = u$

Proof. The key property is that $F^{-1}(u) \leq x$ if and only if $u \leq F(x)$, which follows directly from the definition and the monotonicity of F. The two definitions are equivalent because CDFs are non-decreasing and right-continuous. □

3.3 Discrete Case: Inverse Transform Sampling

Let X take ordered values $x_1 < x_2 < \cdots$ with probabilities $p_X(x_i)$, and define the cumulative $F_X(x_i) = \sum_{j \leq i} p_X(x_j)$, with $F_X(x_0) \equiv 0$. Given $U \sim \text{Unif}(0,1)$, set

$$X = x_i \quad \text{if } F_X(x_{i-1}) < U \leq F_X(x_i).$$

Then $\Pr(X = x_i) = F_X(x_i) - F_X(x_{i-1}) = p_X(x_i)$.

The algorithm works because the intervals $(F_X(x_{i-1}), F_X(x_i)]$ partition $(0,1]$ and have lengths equal to $p_X(x_i)$. Since U is uniform, it falls in each interval with probability equal to the interval's length.

Example 6 (Discrete inverse transform)**.** *Let* X *take values* $\{1,3,5\}$ *with probabilities* $\{0.2, 0.5, 0.3\}$*. The cumulative probabilities are:*

- $F_X(1) = 0.2$
- $F_X(3) = 0.7$
- $F_X(5) = 1.0$

Given $U \sim \text{Unif}(0, 1)$*:*

- *If* $0 < U \leq 0.2$*, then* $X = 1$
- *If* $0.2 < U \leq 0.7$*, then* $X = 3$
- *If* $0.7 < U \leq 1.0$*, then* $X = 5$

In practice, precompute cumulative probabilities and use a binary search to locate i.