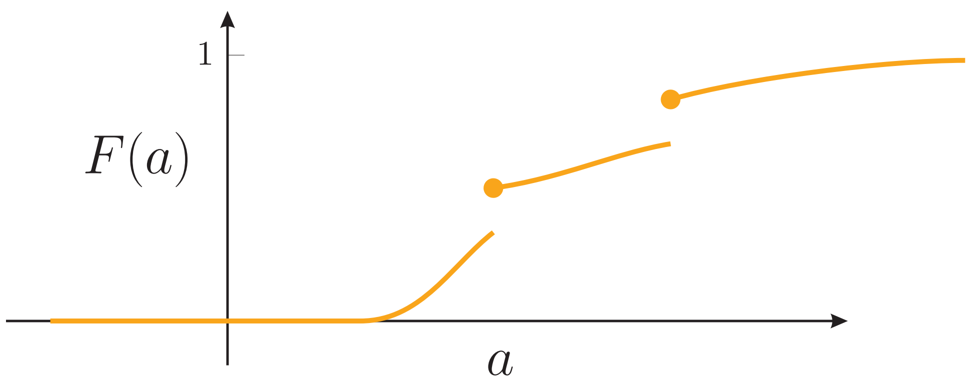
7 - Continuous random variables

- Continuous random variables
- The cumulative distribution function
- The uniform random variable
- Gaussian random variables
- The Gaussian cdf
- Collecting data
- Induced probability density functions
- Example: linear transformations
- Example: non-invertible transformations
- Simulation of random variables

Continuous random variables

A *continuous* random variable $x : \Omega \to \mathbb{R}$ is specified by its (induced) *cumulative distribution function* (cdf)

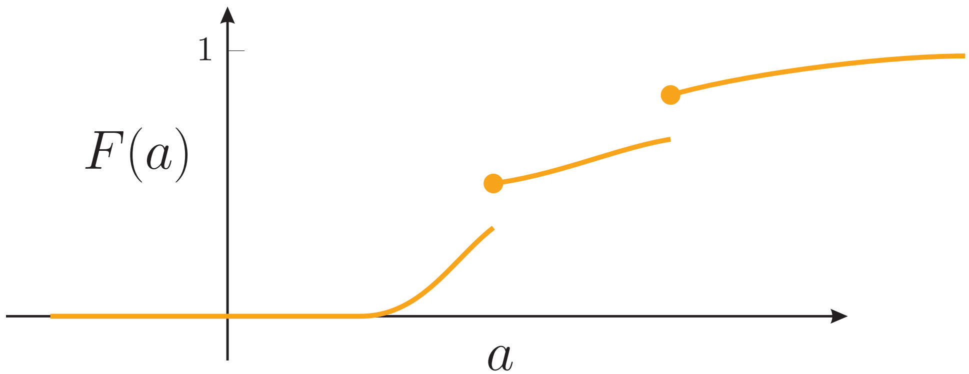
$$F^x(z) = \mathbf{Prob}(x \le z)$$



Then we have

$$\mathbf{Prob}(x \in [a, b]) = F^x(b) - F^x(a^-)$$

Properties of the cumulative distribution function



- $F^x(a) \geq 0$ for all a
- F^x is a non-decreasing function $F^x(a) \leq F^x(b)$ if $a \leq b$
- F^x is *right continuous*, i.e,.

$$\lim_{z \to a^+} F^x(z) = F^x(a)$$

Properties of the cumulative distribution function

If F^x is differentiable, then the *induced probability density function* (pdf) is

$$p^x(z) = \frac{dF^x(z)}{dz}$$

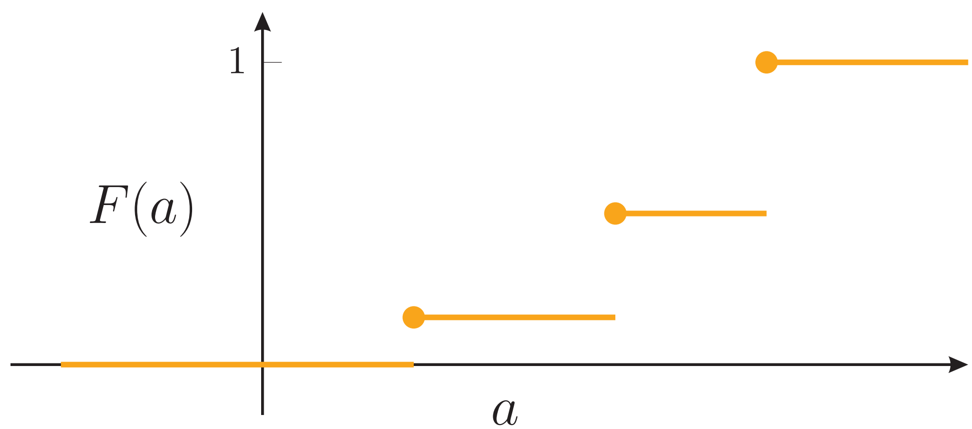
then

$$\mathbf{Prob}(x \in [a, b]) = \int_a^b p^x(z)\, dz$$

- Notice that $p^x(z)$ is not a probability; it may be greater than 1.
- We use notation $x \sim p^x$ to mean x is a random variable with pdf p^x

Properties of the cumulative distribution function

If x is a *discrete random variable*, then F is just a staircase function



- The corresponding probability density function is a sum of δ functions.

The uniform random variable

The *uniform random variable* $x \sim U[a,b]$ has pdf

$$p^x(z) = \begin{cases} \dfrac{1}{b-a} & \text{if } a \le z \le b \\ 0 & \text{otherwise} \end{cases}$$

Gaussian random variables

The *random variable* x is *Gaussian* if it has pdf

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

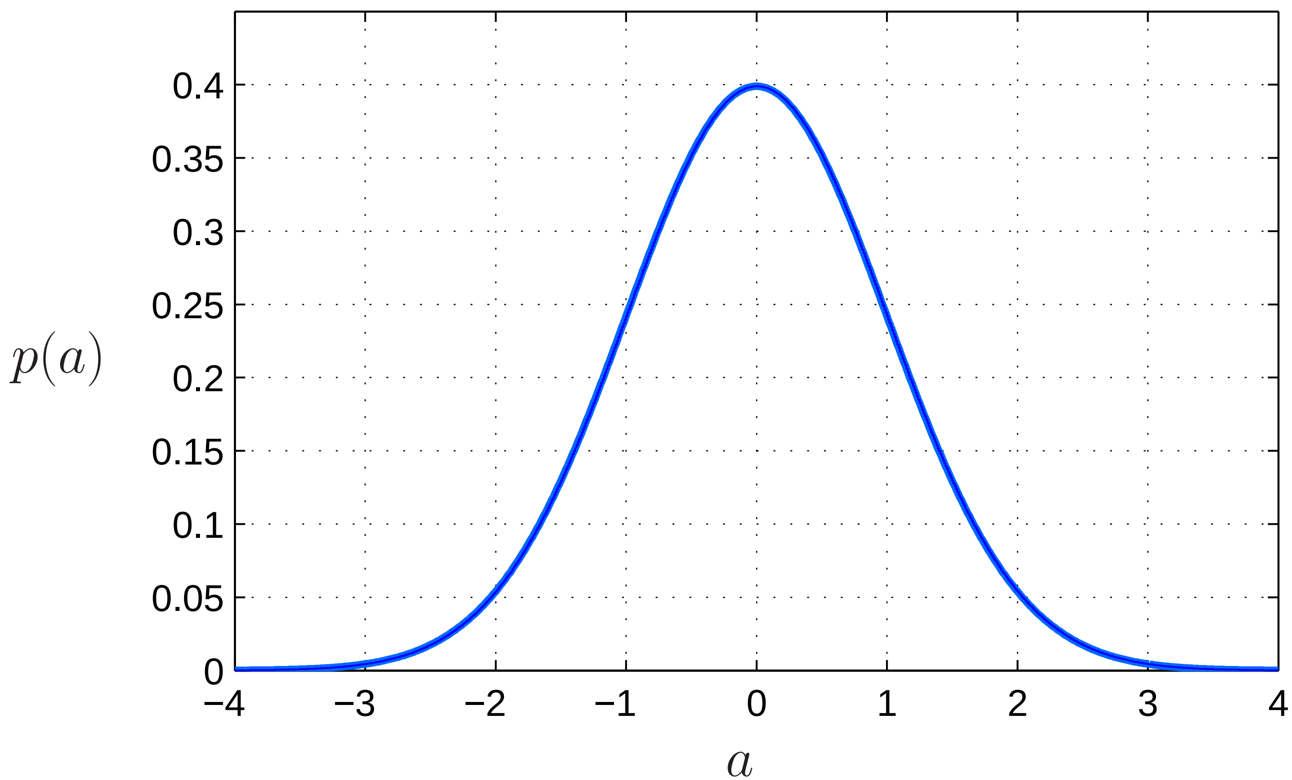
write this as $x \sim \mathcal{N}(\mu, \sigma^2)$

- the *mean* or *expected value* of x is $\quad \mathbf{E}(x) = \int_{-\infty}^{\infty} x p(x)\, dx = \mu$

- the *variance* of x is $\quad \mathbf{E}\big((x-\mu)^2\big) = \int_{-\infty}^{\infty} (x-\mu)^2 p(x)\, dx = \sigma^2$

Gaussian random variables

pdf for $x \sim \mathcal{N}(0,1)$ is



- p is symmetric about the mean
- decays very fast; but $p(x) > 0$ for all x

Computing probabilities for Gaussian random variables

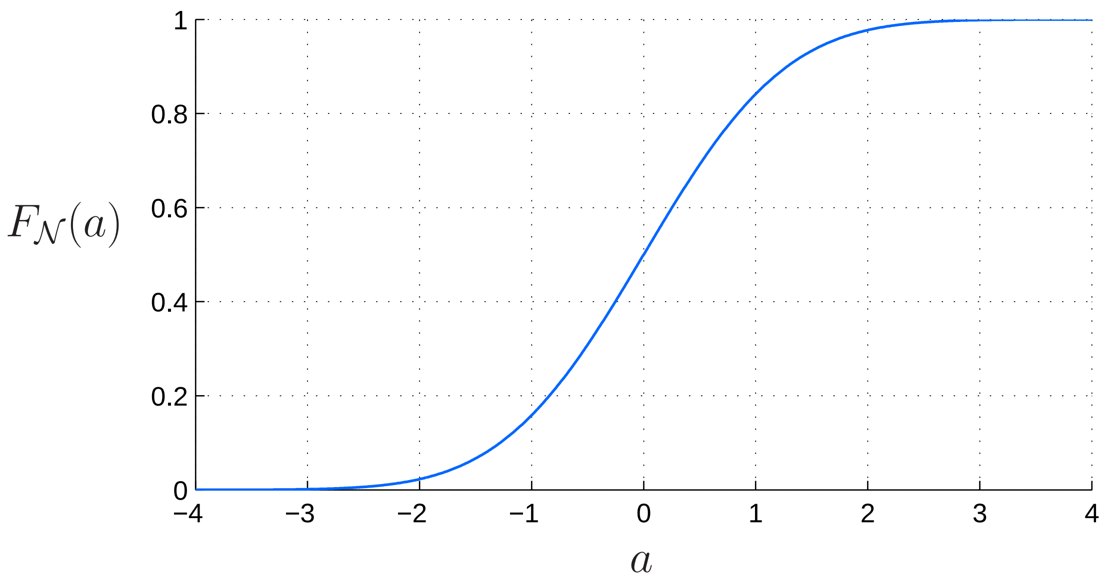
The *error function* is

$$\mathbf{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt$$

The Gaussian CDF is

$$F_{\mathcal{N}}(a) = \frac{1}{2} + \frac{1}{2}\mathbf{erf}\left(\frac{a-\mu}{\sigma\sqrt{2}}\right)$$

When $\mu = 0$ and $\sigma = 1$,



Computing probabilities for Gaussian random variables

so for $x \sim \mathcal{N}(0, \sigma^2)$ we have for $a \geq 0$

$$\mathbf{Prob}(x \in [-a, a]) = \mathbf{erf}\left(\frac{a}{\sigma\sqrt{2}}\right)$$

Some particular values:

$$\mathbf{Prob}(x \in [-\sigma, \sigma]) \approx 0.68$$

$$\mathbf{Prob}(x \in [-2\sigma, 2\sigma]) \approx 0.9545$$

$$\mathbf{Prob}(x \in [-3\sigma, 3\sigma]) \approx 0.9973$$

Collecting data

- For discrete random variables, we can collect data and count the frequencies of outcomes

  This converges to the true pmf.

- The analogous procedure for continuous random variables uses the *cumulative distribution function*.

  Suppose $S = \{z_1, \ldots, z_n\}$ are n samples of a real-valued random variable.

  Let $F(a)$ be the fraction of samples less than or equal to a, given by

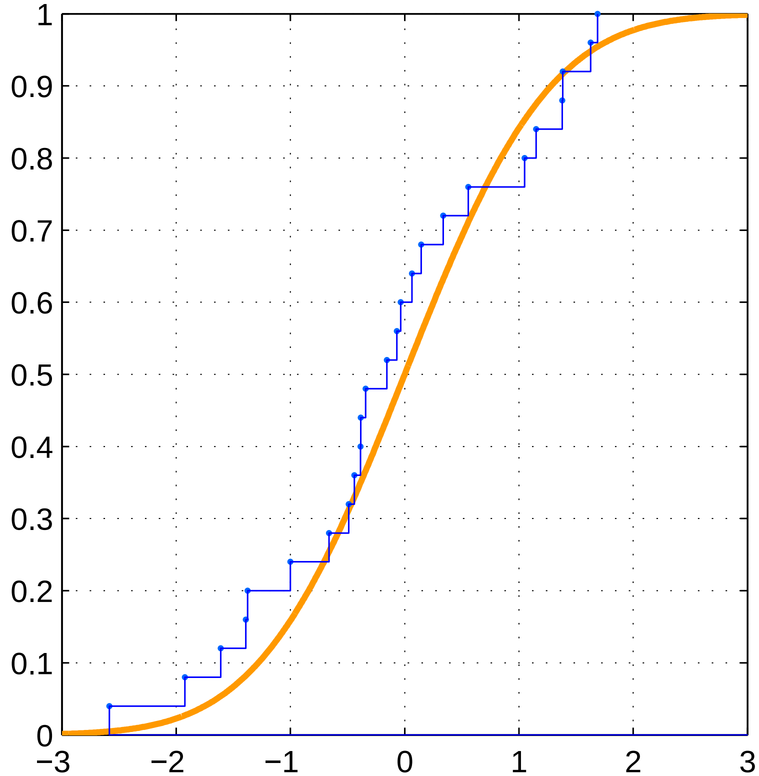
$$F(a) = \frac{|\{\, z \in S \mid z \leq a \,\}|}{n}$$

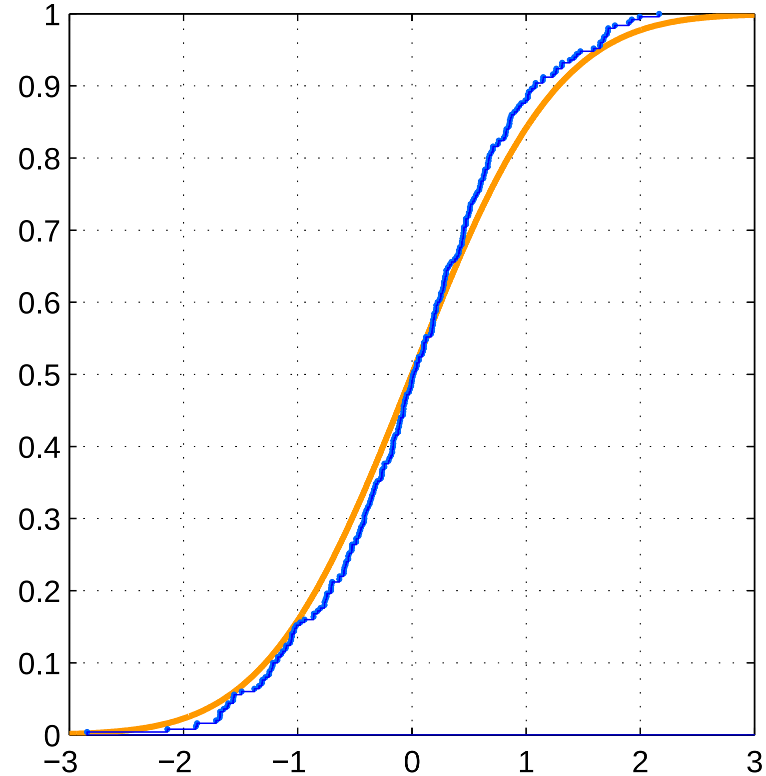
- F is a piecewise constant function, called the *empirical cdf*

Example: collecting data

Suppose $x \sim \mathcal{N}(0,1)$.

The plots below show 25 and 250 data points, respectively.





Induced probability density

Suppose we have

- $x : \Omega \to \mathbb{R}$ is a random variable with induced pdf $p^x : \mathbb{R} \to \mathbb{R}$.
- y is a function of x, given by $y = g(x)$

What is the induced pdf of y?

The key idea is that we need to *change variables* for integration of probabilities. Recall the following.

If f and h' are continuous, then

$$\boxed{\int_{h(a)}^{h(b)} f(x)\,dx = \int_a^b f\big(h(y)\big)h'(y)\,dy}$$

Induced probability density

Assume g' is *continuous*, and g is *strictly increasing*; i.e.,

$$\text{if } a < b \text{ then } g(a) < g(b)$$

This implies that g is *invertible*, i.e., for every y there is a unique x such that $y = g(x)$.

We would like to find the pdf of y is p^y, which satisfies for $a \leq b$,

$$\mathbf{Prob}(y \in [a,b]) = \int_a^b p^y(y)\,dy$$

We also know that this probability is

$$\begin{aligned} \mathbf{Prob}(y \in [a,b]) &= \mathbf{Prob}(g(x) \in [a,b]) \\ &= \mathbf{Prob}\big(x \in [g^{-1}(a), g^{-1}(b)]\big) \quad \text{since } g \text{ is increasing} \\ &= \int_{g^{-1}(a)}^{g^{-1}(b)} p^x(x)\,dx \end{aligned}$$

Induced probability density

We have

$$\int_a^b p^y(y)\,dy = \int_{g^{-1}(a)}^{g^{-1}(b)} p^x(x)\,dx$$

Now we can apply the change of variables $x = h(y)$ to the integral on the right hand side, where $h = g^{-1}$. We have

$$h'(y) = \frac{1}{g'\big(g^{-1}(y)\big)}$$

because $g(h(y)) = y$, so $\dfrac{d}{dy}g(h(y)) = 1$, i.e., $g'(h(y))h'(y) = 1$

Therefore, by the change of variables formula

$$\int_a^b p^y(y)\,dy = \int_a^b \frac{p^x\big(g^{-1}(y)\big)}{g'\big(g^{-1}(y)\big)}\,dy$$

Induced probability density

Since this holds for all a and b, we have the following.

If $y = g(x)$, and g is strictly increasing with g' continuous, then the pdf of y is

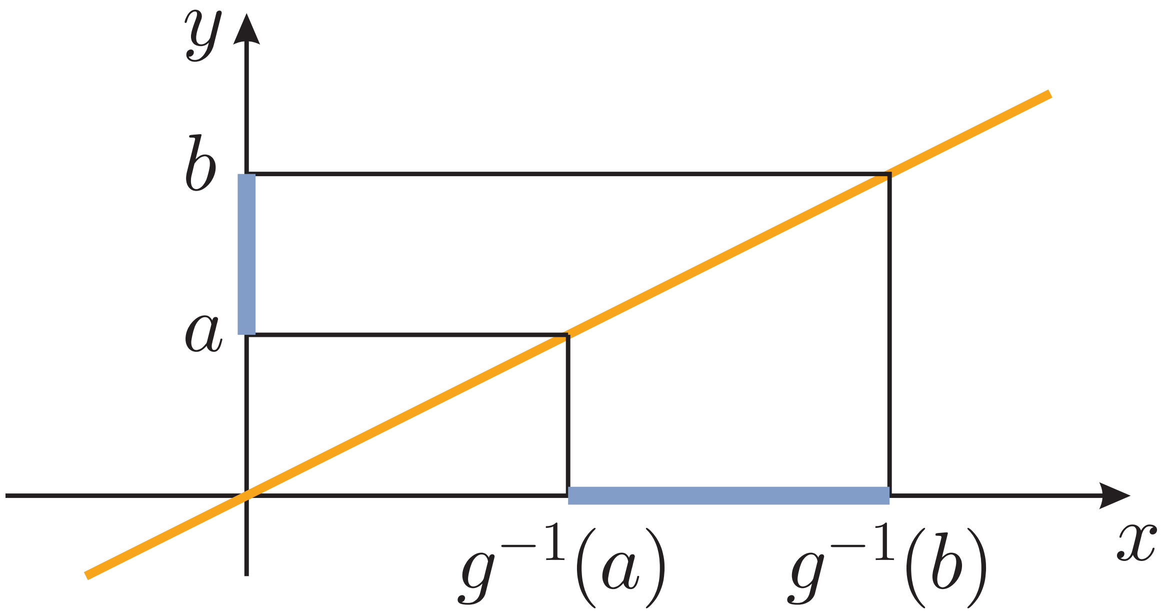
$$p^y(y) = \frac{p^x\big(g^{-1}(y)\big)}{g'\big(g^{-1}(y)\big)}$$

More generally, if $g'(x) \neq 0$ for all x, then

$$p^y(y) = \frac{p^x\big(g^{-1}(y)\big)}{\big|g'\big(g^{-1}(y)\big)\big|}$$

Example: linear transformations

Suppose $x : \Omega \to \mathbb{R}$, and $y = \alpha x + \beta$

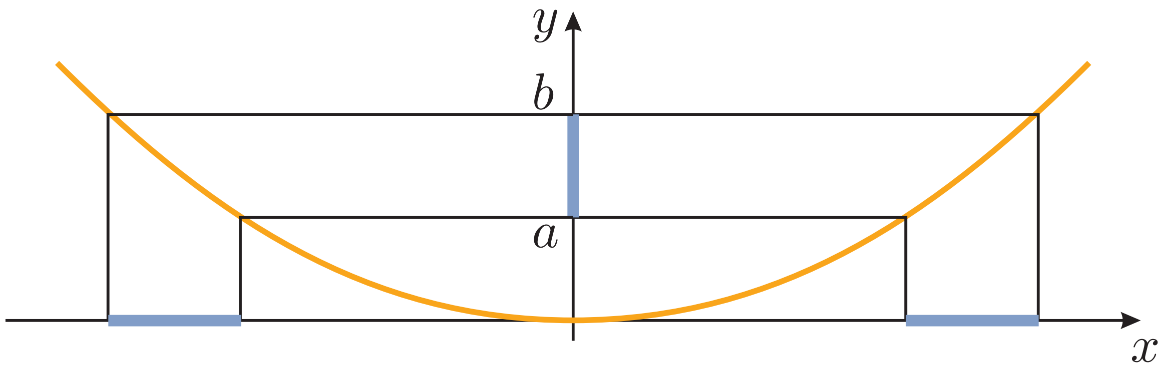


We have

$$p^y(y) = \frac{1}{|\alpha|} p^x \left(\frac{y - \beta}{\alpha} \right)$$

Non-invertible transformations

What happens when g is not invertible? e.g., when $y = x^2$



$$\begin{aligned}
\textbf{Prob}(y \in [a,b]) &= \textbf{Prob}(x^2 \in [a,b]) \\
&= \textbf{Prob}(x \in [-\sqrt{b}, -\sqrt{a}]) + \textbf{Prob}(x \in [\sqrt{a}, \sqrt{b}]) \\
&= \int_{-\sqrt{b}}^{-\sqrt{a}} p^x(x)\,dx + \int_{\sqrt{a}}^{\sqrt{b}} p^x(x)\,dx \\
&= \int_a^b \frac{p^x(-\sqrt{y})}{2\sqrt{y}}\,dy + \int_a^b \frac{p^x(\sqrt{y})}{2\sqrt{y}}\,dy
\end{aligned}$$

$$y = x^2 \qquad \Longrightarrow \qquad p^y(y) = \frac{1}{2\sqrt{y}}\Big(p^x(-\sqrt{y}) + p^x(\sqrt{y})\Big)$$

Simulation of random variables

We are given $F : \mathbb{R} \rightarrow [0, 1]$

- We would like to simulate a random variable y so that it has *cumulative distribution function* F
- We have a source of *uniform* random variables $x \sim U[0, 1]$

To construct y, set

$$\boxed{y = F^{-1}(x)}$$

Because

$$\begin{aligned} \mathbf{Prob}(y \leq a) &= \mathbf{Prob}(F^{-1}(x) \leq a) \\ &= \mathbf{Prob}(x \leq F(a)) \\ &= F(a) \end{aligned}$$

- This works when F is invertible and continuous