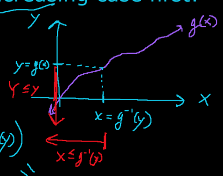
$Y = g(X)$, $f_X(x)$ known, what is $f_Y(y)$?

Let's consider g to be a 1:1 function (i.e. it's strictly increasing or strictly decreasing). Let's do the strictly increasing case first:

Let's try first to derive the CDF and then to get the PDF take its derivative:



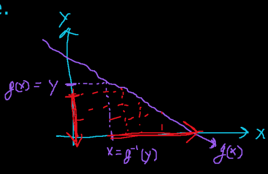
$$F_Y(y) := P(Y \le y) = P(g(X) \le y) \overset{\le}{=} P(X \le g^{-1}(y)) = F_X(g^{-1}(y))$$

$$f_Y(y) = \frac{d}{dy}\left[F_X(g^{-1}(y))\right] = f_X(g^{-1}(y))\, \frac{d}{dy}\left[g^{-1}(y)\right]$$

$$\boxed{\frac{d}{dx}\left[g(x^3)\right] = g'(x^3)\, 3x^2}$$

$$= f_X(g^{-1}(y)) \left|\frac{d}{dy}\left[g^{-1}(y)\right]\right| \qquad \frac{d}{dx}[g(x)] > 0 \Leftrightarrow \frac{d}{dy}[g^{-1}(y)] > 0$$

Let's consider the case where g is strictly decreasing and do the same: find the CDF and take its derivative.



$$F_Y(y) = P(Y \le y) = P(g(X) \le y) \overset{g\ 1:1}{=} P(X \ge g^{-1}(y))$$

$$= 1 - F_X(g^{-1}(y))$$

$$f_Y(y) = \frac{d}{dy}\left[1 - F_X(g^{-1}(y))\right] = -f_X(g^{-1}(y))\, \frac{d}{dy}\left[g^{-1}(y)\right]$$

$$\frac{d}{dx}[g(x)] < 0 \Leftrightarrow \frac{d}{dy}[g^{-1}(y)] < 0$$

$$= f_X(g^{-1}(y)) \left|\frac{d}{dy}\left[g^{-1}(y)\right]\right| \qquad \Rightarrow -\frac{d}{dy}\left[g^{-1}(y)\right] = \left|\frac{d}{dy}\left[g^{-1}(y)\right]\right|$$

We have the same formula for the PDF in both cases for g, thus it is the general formula.

Let's derive some simple consequences of this formula that are very powerful:

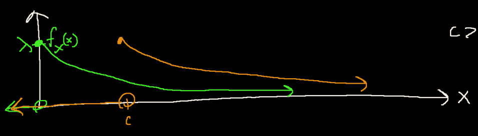
Y = g(X) = aX + c ~ ?   (this is is called a shifted and scaled rv)n

$$y = ax + c \Rightarrow y - c = ax \Rightarrow x = \frac{y-c}{a} = g^{-1}(y) \Rightarrow \left|\frac{d}{dy}\left[g^{-1}y\right]\right| = \left|\frac{1}{a}\right| = \frac{1}{|a|}$$

$$\boxed{f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-c}{a}\right)}$$

$$Y = -X \Rightarrow f_Y(y) = f_X(-y), \quad Y = aX \Rightarrow f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right), \quad Y = X + c \Rightarrow f_Y(y) = f_X(y - c)$$

$$X \sim \text{Exp}(\lambda) = \lambda e^{-\lambda x} \mathbb{1}_{x \in (0,\infty)}, \quad Y = X + c \sim \lambda e^{-\lambda(y-c)} \mathbb{1}_{y - c \in (0,\infty)} = e^{\lambda c} \lambda e^{-\lambda y} \mathbb{1}_{y \in (c,\infty)}$$



$$X \sim \text{Exp}(1) = e^{-x} \mathbb{1}_{x \in (0,\infty)}$$

$$Y = g(X) = -\ln\left(\frac{e^{-x}}{1 - e^{-x}}\right) = \ln\left(\frac{1 - e^{-x}}{e^{-x}}\right) = \ln\left(e^{x} - 1\right) \sim f_Y(y) = ?$$

$$y = \ln(e^x - 1) \Rightarrow e^y = e^x - 1 \Rightarrow e^y + 1 = e^x = \ln(e^y + 1) = x = g^{-1}(y)$$

$$\left|\frac{d}{dy}\left[\ln(e^y + 1)\right]\right| = \left|\frac{e^y}{e^y + 1}\right| = \frac{e^y}{e^y + 1} = \frac{1}{1 + e^{-y}} \overset{\text{FYI}}{\in} (0,1)$$

$$f_Y(y) = f_X\left(\ln(e^y + 1)\right) \frac{1}{1 + e^{-y}} = e^{-\ln(e^y + 1)}\, \mathbb{1}_{\ln(e^y + 1) \in (0,\infty)}\, \frac{e^y}{e^y + 1}$$

$$= \frac{1}{e^y + 1} \frac{e^y}{e^y + 1} = \frac{e^y}{(e^y + 1)^2} = \text{Logistic}(0, 1)$$

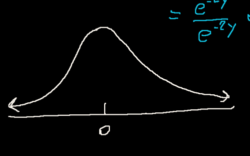
$$= \frac{e^{-2y}}{e^{-2y}} \cdot \frac{e^y}{(e^y+1)^2} = \frac{e^{-y}}{(1 + e^{-y})^2}$$

$e^y + 1 \in (1,\infty)$

$e^y \in (0,\infty)$

$y \in \mathbb{R}$

no need for indicator function!



This "standard logistic" rv looks very similar to a standard normal rv but it has slightly thicker tails. Used in chess ratings (Elo), deep learning, implicitly used in "logistic regression".

X ~ Logistic(0, 1), σ > 0, Y = μ + σX ~ Logistic(μ, σ) = ?

$$f_Y(y) = \frac{1}{\sigma} \frac{e^{\frac{y-\mu}{\sigma}}}{\left(e^{\frac{y-\mu}{\sigma}} + 1\right)^2}, \quad E[Y] = \mu, \quad SD[Y] = \sigma \frac{\pi}{\sqrt{3}} \approx 1.8\sigma$$

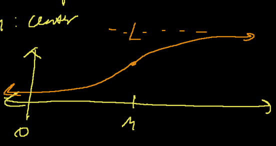
$$X \sim U(0,1), \quad Y = -\ln\left(\frac{1}{X} - 1\right) \sim \text{Logistic}(0,1) \quad (\text{HW})$$

Why is it called "logistic"?

$$\ell(x) = \frac{L}{1 + e^{-k(x-\mu)}} \in (0, L)$$

L: max value,
k: steepness,
μ: center

$$\overset{L=1,\ k=1,\ \mu=0}{=} \frac{1}{1 + e^{-x}} = \frac{e^x}{1 + e^x}$$



This function ell is very useful and it's used to model population changes.

Y ~ Logistic(0, 1)

this is the std. logistic funct.

$$F_Y(y) = \int_{t=-\infty}^{t=y} \frac{e^t}{(1 + e^t)^2} dt = \int_{u=1}^{u=1+e^y} \frac{\cancel{u-1}}{u^2} \frac{1}{\cancel{u-1}} du = \left[-\frac{1}{u}\right]_1^{1+e^y} = 1 - \frac{1}{1+e^y} = \frac{e^y}{1+e^y} = \frac{1}{1+e^{-y}}$$

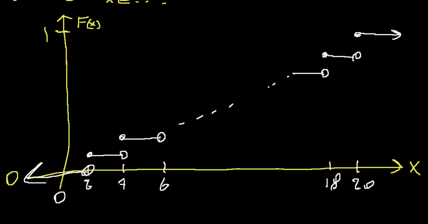
let $u = 1 + e^t \Rightarrow \frac{du}{dt} = e^t \Rightarrow dt = e^{-t} du = \frac{1}{u-1} du$, $t = y \Rightarrow u = 1 + e^y$

$\Rightarrow u - 1 = e^t$ $t = -\infty \Rightarrow u = 1 + e^{-\infty} = 1$

New concept. Solve for min x s.t. q ≤ P(X ≤ x) i.e. q ≤ F(x). This x is called the "qth quantile" or "100qth percentile". If q = 0.5 that x is called the "median" i.e. Med[X]. Otherwise the notation is Q[X, q], the quantile operator.

$$X \sim U(\{2, 4, 6, \ldots, 20\}) = \frac{1}{10} \mathbb{1}_{x \in \ldots}$$

x	p(x)	F(x)
2	0.1	0.1
4	0.1	0.2
6	0.1	0.3
⋮	⋮	⋮
18	0.1	0.9
20	0.1	1.0



$$Q[X, 0.3] = 6$$

6 is the 30%ile of the X distribution

$$Q[X, 0.9] = 18$$

$Q[X, 0.85] = 18$ is the min x s.t. $F(x) \ge 0.85$

$\cancel{\exists}\, x$ s.t. $F(x) = 0.85$

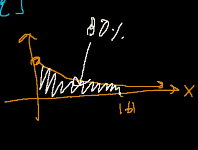
If X was a continuous rv with "contiguous support" meaning basically no gaps e.g. [0,10], (0, ∞), ℝ but not [0,1] ∪ [8,9] then the min x s.t. F(x) ≥ q is where F(x) = q i.e x = $F^{-1}(q)$ which is then called the "quantile function".

Let $X \sim \text{Exp}(\lambda) = \lambda e^{-\lambda x} \Rightarrow F(x) = 1 - e^{-\lambda x} = q$

$$\Rightarrow 1 - q = e^{-\lambda x} \Rightarrow \ln(1-q) = -\lambda x \Rightarrow \ln\left(\frac{1}{1-q}\right) = \lambda x$$

$$\Rightarrow x = \frac{1}{\lambda} \ln\left(\frac{1}{1-q}\right) = F_X^{-1}(q) = Q[X, q]$$

If $\lambda = 1$, $F_X^{-1}(0.8) = \ln(5) \approx 1.61$



$$\text{Med}[X] = F_X^{-1}\left(\tfrac{1}{2}\right) = \ln(2)$$

The quantile function is rarely in closed form e.g.

T ~ Erlang(k, λ) => F(t) = P(k, λt) whose inverse is not available in closed form. So... you use a computer to solve for t where q = P(k, λt).