Example (lifetime T). Suppose a component has a 5% chance of failing immediately but, given it does not, its lifetime has an exponential(λ) distribution, $\lambda = 10$, i.e. $P(0 < T \le t | T \ne 0) = 1 - e^{-t/\lambda}, \lambda = 10$.
The distribution function $F_T(t)$ is clearly 0 if $t < 0$.
There is a jump at $t = 0$: $F_T(0) = P(T \le 0) = P(T = 0) = 0.05$. For $t > 0$ we have

$$\begin{aligned} F_T(t) &= P(T \le t) = P(T = 0) + P(0 < T \le t \mid T \ne 0)P(T \ne 0) \\ &= 0.05 + 0.95(1 - e^{-t/10}). \end{aligned}$$

Additional assumption: observe until failure or for 30 hours (whichever comes first; "**censored** data") and record either the observed time of failure or 30. Then $\mathcal{S} = [0, 30]$, $1 = P(\mathcal{S}) = P(T \in [0, 30]) = P(T \le 30) = P(T = 30) + P(T < 30)$. Use this to obtain $P(T = 30) = 1 - P(T < 30) = 1 - \{0.05 + 0.95(1 - e^{-3})\} \approx 0.047$. In conclusion, $F_T(t)$ has two jumps, namely from 0 to 0.05 at $t = 0$ and from 0.953 to 1 at $t = 30$. Between $t = 0$ and $t = 30$ it is continuous and strictly increasing.

2 Working with random variables

2.1 Transformations

Aim: Find the distribution of $Y = h(X)$.

Example. Consider the following game: you roll two dice until a total of 7 is observed. The payoff is $Y = 2^X$ where X denotes the number of rolls, i.e. Y takes values $2, 2^2, 2^3, 2^4, \ldots$ (discrete r.v.). The pmf of Y is computed from the distribution of X: for $y = 2, 4, 8, 16, \ldots$ (and thus $\log_2 y \in \mathbb{N}$) we have

$$f_Y(y) = P(Y = y) = P(2^X = y) = P(X = \log_2 y) = f_X(\log_2 y) = \frac{1}{6}\left(\frac{5}{6}\right)^{\log_2 y - 1},$$

with $f_Y(y) = 0$ otherwise.

Example. Let $U \sim U(0, 1)$ and $V = |U - 0.5|$ (distance from the middle). Then V takes values v in $[0, 0.5]$ and $F_V(v) = P(V \le v)$ computes to

$$\begin{aligned} F_V(v) &= P(0.5 - v \le U \le 0.5 + v) = P(U \le 0.5 + v) - P(U \le 0.5 - v) \\ &= 0.5 + v - (0.5 - v) = 2v \end{aligned}$$

for $v \in [0, 0.5]$. The pdf is $f_V(v) = \frac{d}{dv} F_V(v) = 2 \cdot 1_{[0,0.5]}(v)$, i.e. the distribution of V is again uniform, $V \sim U(0, 0.5)$.

Preliminary note: two dice example / inverse image X^{-1}

$$\begin{array}{cccccc} 11 & 12 & 13 & 14 & 15 & \mathbf{16} \\ 21 & 22 & 23 & 24 & \mathbf{25} & 26 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{61} & 62 & 63 & 64 & 65 & 66 \end{array} \quad \begin{array}{c} \text{X = "total"} \\ \longrightarrow \\ X^{-1} \\ \longleftarrow \end{array} \quad 2\ 3\ 4\ 5\ 6\ \mathbf{7}\ 8\ 9\ 10\ 11\ 12$$

$P(X = 7) = P(\{s : X(s) = 7\}) = P(X^{-1}(\{\{7\})) = P(\{(1, 6), (2, 5), \ldots, (6, 1)\}) = 6/36$.

"Compute the distribution of $Y = h(X)$ from the distribution of X".

Definition. Let $h : \mathbb{R} \to \mathbb{R}$. For any set $A \subset \mathbb{R}$ the **inverse image** of A by h is

$$h^{-1}(A) = \{x : h(x) \in A\}.$$

In particular, for $y \in \mathbb{R}$, $h^{-1}((-\infty, y]) = \{x : h(x) \leq y\}$.

Theorem. Let X be a r.v. and h a real-valued function. The cdf for $Y = h(X)$ is then

$$F_Y(y) = P(X \in h^{-1}(-\infty, y]).$$

Proof. $F_Y(y) = P(Y \leq y) = P\{h(X) \in (-\infty, y]\} = P(X \in h^{-1}(-\infty, y])$. □

Example. Let $Y = \#$ successes in n independent Bernoulli(p) trials, $W = n - Y$ ($\#$ failures).
What is the distribution of W?

Short answer: $W = \#$ negative 'successes' in n independent Bernoulli($1-p$) trials
$\Rightarrow W \sim binomial(n, 1-p)$.

Slowly: Let $w \in \{0, 1, \ldots, n\}$. Then $f_W(w) = P(n - Y = w) = P(Y = n - w) = \binom{n}{n-w} p^{n-w}(1-p)^w = \frac{n!}{(n-w)!(n-n+w)!} p^{n-w}(1-p)^w = \binom{n}{w}(1-p)^w(1-(1-p))^{n-w}$; now introduce $\tilde{p} = 1 - p$.

Corollary. Assume, additionally, that Y is discrete.
(i) The pmf for Y is $f_Y(y) = P(X \in h^{-1}\{y\})$.
(ii) If X is also discrete with pmf f_X then $f_Y(y) = \sum_{x \in h^{-1}\{y\}} f_X(x)$.
(iii) If h is 1-1 with inverse h^{-1} then X must also be discrete and $f_Y(y) = f_X(h^{-1}\{y\})$.

Proof. (i) holds since $f_Y(y) = P(h(X) \in \{y\}) = P(X \in h^{-1}\{y\})$, (ii) since P is additive, and (iii) since h is 1-1 and therefore $h^{-1}\{y\}$ is *one* value. □

Corollary. Consider X and $Y = h(X)$ as before; let Y be (absolutely) continuous.
(i) The pdf for Y is $f_Y(y) = \frac{d}{dy} P(X \in h^{-1}(-\infty, y])$.
(ii) If h is 1-1 with inverse h^{-1} and differentiable then X is (absolutely) continuous and

$$f_Y(y) = \frac{1}{|h'\{h^{-1}(y)\}|} f_X\{h^{-1}(y)\}.$$

Proof. (i) Y is absolutely continuous, i.e. for those y at which F' exists, $f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} P(X \in h^{-1}(-\infty, y])$ by the previous theorem.
(ii) Assume h *strictly increasing.* Then, applying the chain rule,

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} P\{h(X) \leq y\} = \frac{d}{dy} P\{X \leq h^{-1}(y)\} = \frac{d}{dy} F_X\{h^{-1}(y)\} \\ &= \frac{d}{dy} h^{-1}(y) f_X\{h^{-1}(y)\} = \frac{1}{h'\{h^{-1}(y)\}} f_X\{h^{-1}(y)\} \end{aligned}$$

with $h' > 0$. We can therefore write $\frac{1}{|h'\{h^{-1}(y)\}|} f_X\{h^{-1}(y)\}$.

Assume h *decreasing.* Then, similarly, $f_Y(y) = \frac{d}{dy} P\{h(X) \leq y\} = \frac{d}{dy} P\{X \geq h^{-1}(y)\} = \frac{d}{dy}[1 - F_X\{h^{-1}(y)\}] = -\frac{1}{h'\{h^{-1}(y)\}} f_X\{h^{-1}(y)\}$ with $h' < 0$. Again we can write $\frac{1}{|h'\{h^{-1}(y)\}|} f_X\{h^{-1}(y)\}$. □

It is easier to recall the previous density transformation formula if one writes

$$f_Y(y) = \left|\frac{dx}{dy}\right| f_X(x).$$