Let $X_1, X_2 \overset{iid}{\sim} \text{Exp}(1) := e^{-x}\mathbb{1}_{x\in(0,\infty)}$. $D = \underset{X}{X_1} - \underset{Y = -X_2}{X_2} = X + Y$

$Y \sim f_Y(y) = e^{-\cdot y}\mathbb{1}_{\cdot y\in(0,\infty)} = e^{y}\mathbb{1}_{y\in(-\infty,0)}$

$$f_D(d) = \int_{x\in \text{Supp}[X]} f_X^{old}(x)\, f_Y^{old}(d-x)\,\mathbb{1}_{d-x\in\text{Supp}[Y]}\,dx$$

$$= \int_{x\in(0,\infty)} e^{-x}\,\underbrace{e^{d-x}}_{e^d e^{-x}}\,\mathbb{1}_{d-x\in(-\infty,0)}\,dx = e^d\int_{x\in(0,\infty)} e^{-2x}\mathbb{1}_{x\in(d,\infty)}\,dx$$

($d - x \in (-\infty, 0)$, $x - d \in (0,\infty)$, $x \in (d, \infty)$)

$$= e^d\begin{cases}\int_d^\infty e^{-2x}dx & \text{if } d\ge 0\\ \int_0^\infty e^{-2x}dx & \text{if } d<0\end{cases} = e^d\begin{cases}\left[-\frac{1}{2}e^{-2x}\right]_d^\infty & \text{if } d\ge 0\\ \left[-\frac{1}{2}e^{-2x}\right]_0^\infty & \text{if } d<0\end{cases}$$

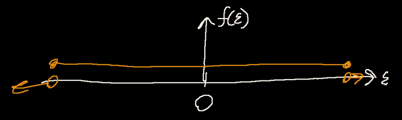
$$= e^d\begin{cases}\frac{1}{2}e^{-2d} & \text{if } d\ge 0\\ \frac{1}{2} & \text{if } d<0\end{cases} = \frac{1}{2}\begin{cases}e^{-d} & \text{if } d\ge 0\\ e^{d} & \text{if } d<0\end{cases} = \frac{1}{2}e^{-|d|}$$

$d \ge 0 \Rightarrow d = |d|$, $d < 0 \Rightarrow d = -|d|$

$=$ Laplace(0, 1) i.e. "standard Laplace" or "double-exponential"

X ~ Laplace(0, 1), Y = μ + σX ~ Laplace(μ, σ) := $f_Y(y) = \frac{1}{2\sigma}e^{-\frac{|y-\mu|}{\sigma}}$, $\sigma > 0$

Laplace published this distribution in 1774 calling it the "first law of errors". Imagine you measure a quantity v but your measuring device has random error ε so M = v + ε where ε is a rv. It makes sense that E[ε] = 0 => E[M] = v, further Med[ε] = 0 (i.e. 50% of the time ε > 0 and 50% of the time ε < 0) and f(ε) = f(-ε). What distributions have these properties?



The problem with the above is small errors should be more likely than large errors. Thus, another consideration is f'(ε) < 0 for ε > 0 and f'(ε) > 0 for ε < 0.

Then he reasoned if f''(ε) = f'(ε) => f(ε) = $ce^{-d|\varepsilon|}$ when you solve this simple differential equation. Solving this for c,d to make it a valid PDF, you get c = 1/2, d = 1. (should be d = 1/2. This is only one such valid solution). c=d is full solution set.

$X \sim \text{Exp}(1) = e^{-x}\mathbb{1}_{x\in(0,\infty)}$. $Y = \frac{1}{\lambda}X^{\frac{1}{k}} = g(X)$ s.t. $\lambda > 0$, $k > 0$.

$y = \frac{1}{\lambda}x^{\frac{1}{k}} \Rightarrow \lambda y = x^{\frac{1}{k}} \Rightarrow x = \lambda^k y^k = g^{-1}(y)$

$\frac{d}{dy}[\;] = k\lambda^k y^{k-1}$

$$f_Y(y) = f_X(g^{-1}(y))\left|\frac{d}{dy}[g^{-1}(y)]\right| = e^{-\lambda^k y^k}\mathbb{1}_{\lambda^k y^k\in(0,\infty)}\left|k\lambda^k y^{k-1}\right|$$

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

$$= k\lambda(\lambda y)^{k-1}e^{-(\lambda y)^k}\mathbb{1}_{y\in(0,\infty)} = \text{Weibull}(k, \lambda)$$

Weibull(1, λ) = $\lambda e^{-\lambda y}\mathbb{1}_{y\in(0,\infty)}$ = Exp(λ)

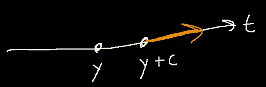
Weibull is a generalization of the exponential making a more flexible "survival distribution". The k parameter is very important. Let's first find the CDF.

$$F(y) = \int_{t=0}^{t=y} k\lambda(\lambda t)^{k-1}e^{-(\lambda t)^k}dt = \int_{u=0}^{u=(\lambda y)^k} \cancel{k\lambda(\lambda t)^{k-1}}\, e^{-u}\,\frac{1}{\cancel{\lambda^k k t^{k-1}}}\,du$$

let $u = (\lambda t)^k = \lambda^k t^k \Rightarrow \frac{du}{dt} = \lambda^k k t^{k-1} \Rightarrow dt = \frac{1}{\lambda^k k t^{k-1}}du$, $t = 0 \Rightarrow u = 0$, $t = y \Rightarrow u = (\lambda y)^k$

$$= \left[-e^{-u}\right]_0^{(\lambda y)^k} = 1 - e^{-(\lambda y)^k} \Rightarrow S(y) = e^{-(\lambda y)^k}$$

Let's consider the conditional probability w = P(Y ≥ y + c | Y ≥ c), c > 0.



$$= \frac{P(Y\ge y+c, Y\ge c)}{P(Y\ge c)} = \frac{P(Y\ge y+c)}{P(Y\ge c)} = \frac{e^{-(\lambda(y+c))^k}}{e^{-(\lambda c)^k}} = \boxed{e^{\lambda^k(c^k - (y+c)^k)}}$$

k is called the "Weibull modulus".

$k = 1 \Rightarrow w = e^{\lambda(c - (y+c))} = e^{-\lambda y} = \text{Exp}(\lambda)$

This is the "memorylessness property" of the exponential rv. The geometric also has this property due to the underlying iid Bernoullis.

$k > 1 \Rightarrow e^{\lambda^k(c^k - (y+c)^k)} < e^{-(\lambda y)^k} = P(Y \ge y)$

$k < 1 \Rightarrow e^{\lambda^k(c^k - (y+c)^k)} > e^{-(\lambda y)^k} = P(Y \ge y)$

Order Statistics (p160). Let $X_1, \ldots, X_n$ be continuous rv's then sort them from smallest to largest and denote them $X_{(1)}, \ldots, X_{(n)}$ which are called the order statistics of the original set of rv's.

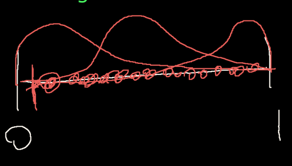
$X_{(1)} = \min\{X_1, \ldots, X_n\}$

$\vdots$

$X_{(k)} = k^{th}$ largest

$\vdots$

$X_{(n)} = \max\{X_1, \ldots, X_n\}$



$R := X_{(n)} - X_{(1)}$ which is called the "range"

The goal is to find the distribution (PDF and CDF) of all order statistics given the distribution of the original collection.