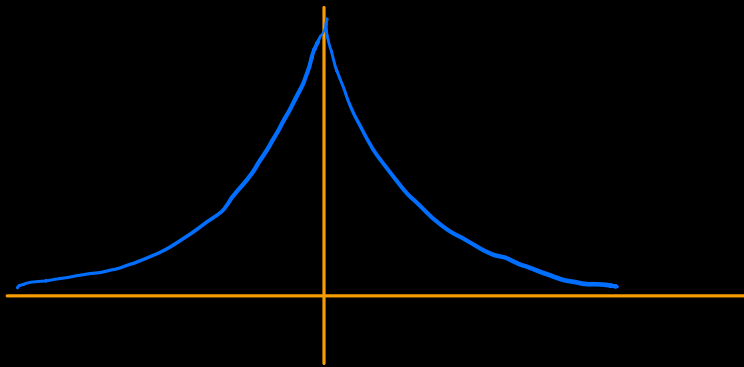


## LAPLACE DIST<sup>n</sup>

- also known as double exp dist<sup>n</sup>
- looks like 2 exp glued together back to back



can be centred around  
zero or o.w.

↳ has 2 params

$\mu$ - centre
$b$ - flatness (bigger is flatter)

more below

Chan: say  $X \sim \text{Laplace}(\mu, b)$

w/ pdf  $f(x | \mu, b) = \frac{1}{2b} \exp\left\{-\frac{|x-\mu|}{b}\right\}$  ← can be written

where  $\mu$  is locat<sup>n</sup> param

$b$  is diversity param

\* The support is  $\mathbb{R}$

\* fatter tail than normal dist<sup>n</sup>.

\* special case:  $\mu=0, b=1 \Rightarrow X \sim \exp(\frac{1}{2})$

\* has closed form cdf that has closed form inverse

## Sampling

1. Generate  $u \sim \text{Unif}\left(-\frac{1}{2}, \frac{1}{2}\right]$

Apply transform  $X = \mu - b \operatorname{sgn}(u) \log(1 - 2|u|)$

$\Rightarrow X \sim \text{Laplace}(\mu, b)$

has cdf 1<sup>o</sup>  
deriv!

2. Generate  $X, Y \stackrel{iid}{\sim} \exp(\frac{1}{b})$   
 $\Rightarrow X - Y \sim \text{Laplace}(0, b)$

note: this is centred at zero always.

note: inversion of param relative to  $\exp(\cdot)$

LASSO - some additional notes.

- does some var. select<sup>n</sup> by penalizing for use of many features when predicting  $\rightarrow$  effectively zero out some  $\beta_i$ .
  - $\hookrightarrow$  forces a simpler model.
- classical approach - Least Sqr.

$$\min_{\beta \in \mathbb{R}^p} \left\{ \frac{1}{N} \|y - X\beta\|_2^2 \right\} \text{ subject to } \|\beta\|_1 \leq t$$

And via the Lagrangian mult.

$$\min_{\beta \in \mathbb{R}^p} \left\{ \frac{1}{N} \|y - X\beta\|_2^2 + \lambda \|\beta\|_1 \right\} \text{ where rel. b/w } \lambda \text{ \& } t \text{ is data dependent.}$$

$\nwarrow$   $\sum_{j=1}^p |\beta_j|$