What does it mean to fatten the tails?

David Salazar (edited by J. R. Minter)

Started: 2020-05-09, Last modified: 2020-05-09

Table of Contents

# What does it mean to fatten the tails?

First, let’s define what we mean by fatter tails.

## What are fatter tails?

Intuitively, fat tails distribution are distributions for which their PDFs decay to zero very slowly. So slowly, that extreme values start gaining traction in the determination of the whole distribution. Thus, a distribution is fatter than another one if its PDF takes longer to decay to zero.

## Fattening the tails

Thus, if we wanted to fatten the tails, the intuitive response is to add more mass at the tails such that the PDF takes more time to decay. However, the PDF must still sum up to 1. **From where do we take the mass that we are going to put in the tails?**

Nassim Taleb has a great chapter in his [latest technical book](https://www.academia.edu/37221402/STATISTICAL_CONSEQUENCES_OF_FAT_TAILS_TECHNICAL_INCERTO_COLLECTION_) wherein he explains a simple heuristic to fatten the tails by randomly switching from two different gaussians. Along the way, we learn to identify where the tails begin and what exactly is so difficult about Black Swan events.

## The heuristic

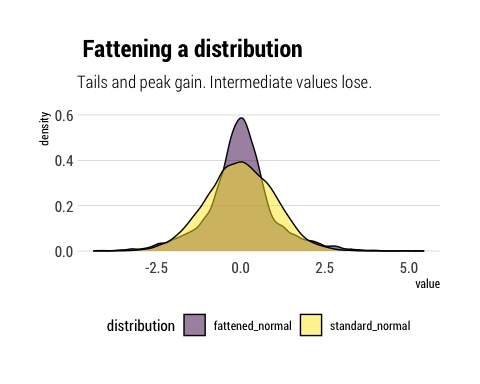
Create a random variable such that:

* with probability ,
* with probability ,

With .

That is, we create a new random variable with an stochastic standard deviation. Let’s simulate samples from this distribution using and . We will compare this fattened samples from the samples of a .

# number of simulated samples  
number\_samples <- 10000  
  
# simulate number\_samples from an uniform  
p\_s <- runif(number\_samples)  
  
# simulate fattened  
fatten\_simulations <- function(p\_s, a) {  
   
 # create vector to store sims  
 sims <- vector(length = length(p\_s))  
 # sample with probability p  
 p <- p\_s < 1/2  
   
 sims[p] <- rnorm(sum(p), sd = sqrt(1-a))  
 sims[!p] <- rnorm(sum(!p), sd = sqrt(1+a))  
 return(sims)  
}  
  
# simulated fattened  
fattened\_normal <- fatten\_simulations(p\_s, 0.8)  
# simulate normal  
standard\_normal <- rnorm(number\_samples)  
  
# plot  
data.frame(n = seq(1, length.out = number\_samples),  
 fattened\_normal,  
 standard\_normal) %>%   
 pivot\_longer(-n, names\_to = "distribution") %>%   
 ggplot(aes(value)) +  
 geom\_density(aes(fill = distribution), alpha = 0.5) +  
 scale\_fill\_viridis\_d() +  
 scale\_color\_viridis\_d() +  
 hrbrthemes::theme\_ipsum\_rc(grid = "Y") +  
 labs(title = " Fattening a distribution",  
 subtitle = "Tails and peak gain. Intermediate values lose.") +  
 theme(legend.position = "bottom")



Therefore, as we stochastize the variance of the distribution, 3 interesting facts arise:

1. The tails of the distribution grow fatter. The PDF at the tails decays to zero more slowly.
2. The peak of the PDF is higher: as Taleb writes, “fatter-tails manifest themselves with higher peaks”.
3. Intermediate events are less likely. Note the yellow are above the fattened tail for intermediate values.

Thus, *we can know answer our question*: **the probability that the tails gain is taken from the probability of the intermediate events**. And then the intermediate events lose even more probability as moderate deviations also gain probability.

## Why does this happen?

We have stochastized the variance of the distribution. Thus, the variance of the distribution has a variance. In particular, this variance of the variance grows with : the larger the , the bigger the difference between the two possible variances of the two different distributions from which we will sample. Taleb writes: “a is roughly the mean deviation of the stochastic volatility parameter”.

Taleb shows this effect on the tails by analyzing the kurtosis of the transformed variable in terms of . The characteristic function is thus:

The fourth derivative evaluated at zero:

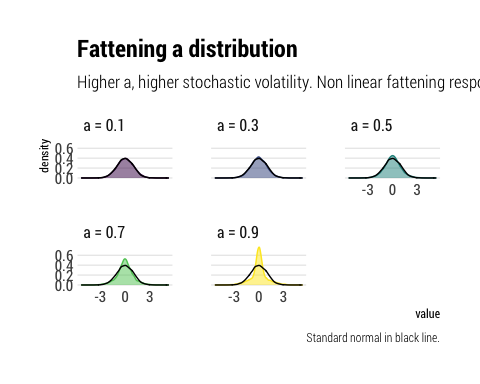
Thus, the higher the , the larger the kurtosis. Thus, we can tinker with the following idea: if the kurtosis is convex to the scale of the distribution, *are tail probabilities also convex to the scale of the distribution?*

## Convexity of the tail

Thus, to find the size of the response to the scale of the distribution, let’s analyze tail probabilities as we grow the parameter which controls the scale of the distribution. Thus, we find the sensitivies of the tail probabilities to perturbations in the variance and we can find out if tail probabilities are indeed convex to the scale of the distribution.

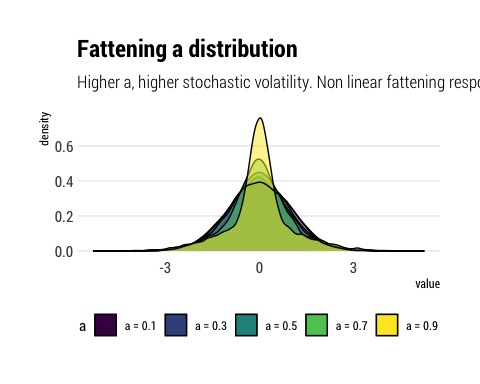
Let’s run the experiments:

standard\_normal <- data.frame(standard\_normal)  
# create the different values for a  
a\_s <- seq(0.1, 0.9, by = 0.2)  
names(a\_s) <- unlist(map(a\_s, ~ glue::glue("a = {.x}")))  
  
a\_s %>%  
 # evaluate the function for different values of a  
 map\_df(~ fatten\_simulations(p\_s, a = .x)) %>%   
 mutate(sim = seq(1, number\_samples)) %>%   
 pivot\_longer(-sim, names\_to = "a") -> different\_as  
different\_as %>%   
 # plot the distributions  
 ggplot(aes(value)) +  
 geom\_density(aes(fill = a,  
 color = a), alpha = 0.5) +  
 geom\_density(data = standard\_normal,  
 aes(standard\_normal)) +  
 scale\_fill\_viridis\_d() +  
 scale\_color\_viridis\_d() +  
 facet\_wrap(~a, ncol = 3) +  
 hrbrthemes::theme\_ipsum\_rc(grid = "Y") +  
 theme(legend.position = "none") +  
 labs(title = "Fattening a distribution",  
 subtitle = "Higher a, higher stochastic volatility. Non linear fattening response.",  
 caption = "Standard normal in black line.")



And now with all curves in one plot:

different\_as %>%   
 ggplot(aes(value)) +  
 geom\_density(aes(fill = a), alpha = 0.5) +  
 geom\_density(data = standard\_normal,  
 aes(standard\_normal)) +  
 scale\_fill\_viridis\_d() +  
 scale\_color\_viridis\_d() +  
 hrbrthemes::theme\_ipsum\_rc(grid = "Y") +  
 theme(legend.position = "bottom") +  
 labs(title = "Fattening a distribution",  
 subtitle = "Higher a, higher stochastic volatility. Non linear fattening response.")



Thus, given the non-linear response of the tail probabilities to changes in , we can conclude, as Taleb does with derivatives, that tail probabilities are indeed convex to the scale of the distribution.

## Where do tails begin?

With this insight, we can respond the important question: **where do the tails begin?**. The tails are the points at the extremes where the probabilities are convex to the stochastic volatility. That is, where changes in the stochastic volatility generate non-linear changes in the probability. Thus, we arrive at a problem: **probability estimation at the tails is unreliable**. If we cannot reliably estimate the standard deviation, then the errors will propagate non-linearly into the estimation of tail probabilities.

Therefore, Taleb concludes that the problem of Black Swans arises not only from the large size of the possible deviations, but also from the unreliability with which we can estimate the probability of these events.

# Conclusions

By investigating how we can fatten a variably by randomly switching from two gaussians, Taleb has shown us a great deal about fat tails. How the distribution changes. And, critically, that the tail probabilities are convex to the scale of the distribution. Thus, small errors in the estimation in the scale will create large problems in the estimation of tail probabilities.