

Taleb's kappa

11:17 11 May 2024

Contents

Minimum number of terms (summands) needed for convergence	1
Problem statement	1
Calculating κ	1
Calculating n_{\min}	2
Notes	2
Note 1	2
Note 2	2
Note 3	3
Note 4	3
Note 5	3
Note 6	3

Minimum number of terms (summands) needed for convergence

See Taleb: The Statistical Consequences Of Fat Tails, Ch. 8

Problem statement

Let $\{X_{g,i}\}$ be Gaussian distributed with mean μ and scale σ .

Let $\{X_{\nu,i}\}$ be t -distributed, scaled such that $\mathbb{M}^{\nu}(1) = \mathbb{M}^g(1) = \sqrt{\frac{2}{\pi}}\sigma$.

Given n_g , we want to determine and n_{ν}^* such that

$$\text{Var} \left[\sum_i^{n_g} X_{g,i} \right] = \text{Var} \left[\sum_i^{n_{\nu}^*} X_{\nu,i} \right]$$

Calculating κ

For iid. r.v $\{X_i\}$:

$$S_n = X_1 + X_2 + \cdots + X_n$$

$$\mathbb{M}(n) = \mathbb{E}(|S_n - \mathbb{E}(S_n)|)$$

Taleb's convergence metric (κ):

The “rate” of convergence for n summands vs n_0 , i.e. the improved convergence achieved by $n - n_0$ additional terms, is given by $\kappa(n_0, n)$:

$$\kappa(n_0, n) = 2 - \frac{\log(n) - \log(n_0)}{\log\left(\frac{\mathbb{M}(n)}{\mathbb{M}(n_0)}\right)}$$

Calculating n_{\min}

The minimum number of summands needed to achieve same variance as the sum of n_g Gaussian summands:

$$n_{\min} := n_{\nu}^*$$

We don't need to calculate $\mathbb{M}(n)$ and $\mathbb{M}(n_0)$. Instead we use the properties

$$\mathbb{M}^{\nu}(1) = \mathbb{M}^g(1) = \sqrt{\frac{2}{\pi}}\sigma$$

where $\sqrt{\frac{2}{\pi}}$ is the ratio between the sd of a Gaussian r.v. and its MAD,

and

$$\mathbb{M}(n) = n^{1/\alpha}\mathbb{M}(1)$$

where α is the degrees of freedom for the Student t . This is true for Stable distributions only.

This seems to not work. Why?

Notes

Note 1

The relation $\mathbb{M}^{\nu}(1) = \sqrt{\frac{2}{\pi}}\sigma$ only holds for the standard normal distribution, and only asymptotically. The ratio $\frac{\text{MAD}(X)}{\text{sd}(X)}$ varies between 0 and 1.

https://en.wikipedia.org/wiki/Average_absolute_deviation#Mean_absolute_deviation_around_the_mean

For a Student t with 3 df, the ratio is $\frac{\pi}{2}$. (See Taleb ch. 4 for further discussion.)

Note 1a The tail exponent α of a Student t distribution with ν degrees of freedom is $\alpha = \nu + 1$, which can be seen from the density function, which is proportional to

$$x^{-(\nu+1)}$$

Confusingly, on p. 147 Taleb uses the notation “Student $T(\alpha)$ ”, where $\alpha = 3$. Does this mean 2 or 3 degrees of freedom??

Taleb et al seem confused by this themselves On p. 150 they write:

“a Student T with 3 degrees of freedom ($\alpha = 3$) requires 120 observations to get the same drop in variance from averaging (hence confidence level) as the Gaussian with 30, that is 4 times as much. The one-tailed Pareto with the same tail exponent $\alpha = 3$ requires 543 observations to match a Gaussian sample of 30”.

-So in one case α is the degrees of freedom of a Student t , and in the next sentence α is the tail exponent.

Note 2

Notice that if $\mathbb{M}(n) = n^{1/\alpha}\mathbb{M}(1)$, then $\kappa(1, n)$ becomes

$$2 - \frac{\log(n) - \log(n_0)}{\log\left(\frac{\mathbb{M}(n)}{\mathbb{M}(n_0)}\right)} = 2 - \frac{\log(n) - \log(1)}{\log\left(\frac{n^{1/\alpha}\mathbb{M}(1)}{\mathbb{M}(1)}\right)} = 2 - \frac{\log(n)}{\log(n^{1/\alpha})}$$

For $\alpha = 2$, this is

$$2 - \frac{\log(n)}{\log(\sqrt{n})} = 2 - \log_{\sqrt{n}}(n) = 0$$

as $\log_{\sqrt{n}}(n) = 2$ for all n .

So if we use $\tilde{\alpha} = \alpha I\{\alpha < 2\} + 2I\{\alpha \geq 2\}$, κ is always 0 for $\alpha \geq 2$. Is the F-S Skewed t Stable? According to the table in Taleb, p. 148, the Student t has a κ far from 0, even when $\alpha > 2$.

Note 3

Also note, that for Stable distributions with $1 \leq \tilde{\alpha}$,

$$\kappa_{(n_0, n)} = 2 - \tilde{\alpha}$$

Note 4

The problem calculating the MAD directly is that we need the mean, which is what we are estimating in the first place! How much data do we need to estimate κ ???

A method then is to estimate the tail index α (which will be Gaussian) using MLE and from there get the theoretical mean. See Taleb Ch. 13.

For the analytical mean of the F-S Skewed t , see Li and Nadarajah: A review of Student's t distribution and its generalizations, p. 10. (Typo: Should be x instead of s . See Fernandez-Steel: On Bayesian Modelling Of Fat Tails And Skewness, eq. 2.5+2.6)

Note 5

In this implementation we calculate the MAD of S_n as

$$\mathbb{M}(n) = \mathbb{E}(|S_n - \mathbb{E}(S_n)|) = \mathbb{E}(|S_n - n\hat{\mu}_{X_i}|)$$

where $\hat{\mu}_{X_i}$ is the estimated mean of X_i . We estimate the mean of X_i by simulation.

Note 6

Test if the sample means of a sum of F-S Skewed t r.v.'s are equal to n times the mean of a single r.v. from that distribution.

Even with only 50 samples to estimate the sample mean, the fit is spot on.

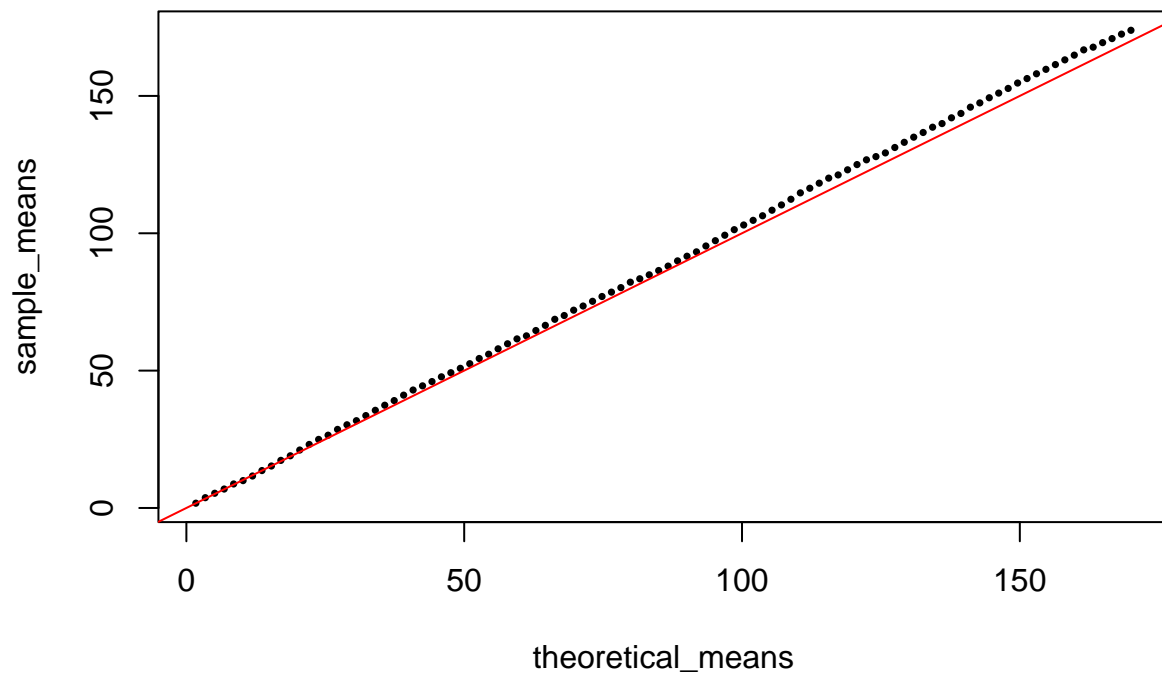
```
num_sim <- 50
n <- 100
mu <- 1.7
sigma <- 2
nu <- 3
xi <- 0.25

x_df <- replicate(num_sim, cumsum(rsstd(n, mu, sigma, nu, xi)))

theoretical_means <- (1:n) * mu
sample_means <- lapply(
  1:n,
  function(i) {
    mean(x_df[i,])
  }
)
```

```
)
```

```
plot(theoretical_means, sample_means, pch = 16, cex = 0.5)  
abline(0, 1, col = "red")
```



```
f_mad <- function(x) {  
  sum(abs(x - mean(x))) / length(x)  
}  
  
## MAD of S_n, given the mean of X  
f_mad_n <- function(Sn, mean_X) {  
  diff <- numeric(length(Sn))  
  for(i in seq_along(Sn)) {  
    diff[i] <- abs(Sn[i] - mean_X)  
  }  
  sum(diff) / length(Sn)  
}  
  
## mean is the mean of a single Gaussian r.v. (same for t-distribution)  
## sd is the sd of a single Gaussian r.v. (same for t-distribution)  
f_kappa <- function(n0, n, mean, sd = 1, nu = 3, xi = 1, num_sim = 1e4) {  
  Sn_sim <- replicate(  
    num_sim,  
    rsstd(n = n, mean = mean, sd = nu/(nu - 2) * sd, nu = nu, xi = xi)  
  )  
  #x_n0 <- rfun(n0, ...)
```

```

#x_n <- rfun(n, ...)
#mad_1 <- sqrt(2 / pi) * sd_g
#mad_n0 <- mad_1 * n0^(1 / nu)
#mad_n <- mad_1 * n^(1 / nu)
mad_n0 <- f_mad_n(
  unlist(lapply(1:num_sim, function(i) sum(Sn_sim[1:n0, i]))),
  n0 * mean
)
mad_n <- f_mad_n(
  unlist(lapply(1:num_sim, function(i) sum(Sn_sim[1:n, i]))),
  n * mean
)
nominator <- log(n) - log(n0)
denominator <- log(mad_n / mad_n0)
2 - (nominator / denominator)
}

## Use approximation if approx = TRUE
f_n_min <- function(n_g, mean, sd_g = 1, nu = 3, xi = 1, num_sim = 1e4, approx = FALSE) {
  ifelse(
    approx,
    exponent <- - 1 / (f_kappa(
      1, 2, mean, sd = sd_g * (nu / (nu - 2)), nu = nu, xi = xi, num_sim = num_sim
    ) - 1),
    exponent <- - 1 / (f_kappa(
      1, n_g, mean, sd = sd_g * (nu / (nu - 2)), nu = nu, xi = xi, num_sim = num_sim
    ) - 1)
  )
  n_g^exponent
}

```

According to table on p. 148, for Student t with 3 df,

- $\kappa_{1,2} = 0.29$

```

nu <- 3
f_kappa(n0 = 1, n = 2, mean = 0, sd = 1, nu = nu, xi = 1, num_sim = 1e5)

```

```
## [1] 0.291821
```

- $\kappa_{1,30} = 0.191$

```
f_kappa(n0 = 1, n = 30, mean = 0, sd = 1, nu = nu, xi = 1, num_sim = 1e5)
```

```
## [1] 0.1929398
```

- $\kappa_{1,100} = 0.159$

```
f_kappa(n0 = 1, n = 100, mean = 0, sd = 1, nu = nu, xi = 1, num_sim = 1e5)
```

```
## [1] 0.1578102
```

```
f_kappa(n0 = 1, n = 100, mean = 0, sd = 1, nu = nu, xi = 1, num_sim = 1e5)
```

```
## [1] 0.1556043
```

```
f_kappa(n0 = 1, n = 100, mean = 0, sd = 1, nu = nu, xi = 1, num_sim = 1e5)
```

```
## [1] 0.1626357
```

```
f_kappa(n0 = 1, n = 100, mean = 0, sd = 1, nu = nu, xi = 1, num_sim = 1e5)
```

```
## [1] 0.1665746
```

```
f_kappa(n0 = 1, n = 100, mean = 0, sd = 1, nu = nu, xi = 1, num_sim = 1e5)
```

```
## [1] 0.1547186
```

```
f_kappa(n0 = 1, n = 100, mean = 0, sd = 1, nu = nu, xi = 1, num_sim = 1e5)
```

```
## [1] 0.1610288
```

```
f_kappa(n0 = 1, n = 100, mean = 0, sd = 1, nu = nu, xi = 1, num_sim = 1e5)
```

```
## [1] 0.1546279
```

```
f_kappa(n0 = 1, n = 100, mean = 0, sd = 1, nu = nu, xi = 1, num_sim = 1e5)
```

```
## [1] 0.1566946
```

```
f_kappa(n0 = 1, n = 100, mean = 0, sd = 1, nu = nu, xi = 1, num_sim = 1e5)
```

```
## [1] 0.1556008
```

“a Student T with 3 degrees of freedom ($\alpha = 3$) requires 120 observations to get the same drop in variance from averaging (hence confidence level) as the Gaussian with 30”. This seems only true for the approximation - which seems to be not very good:

Without approximation:

$$n_{g\nu} = n_g^{\frac{1}{\kappa_{1,n_g}-1}}$$

```
n_g <- 30 ## Number of Gaussian terms
sd_g = 1
df = 3
```

```
n_min <- f_n_min(n_g, mean = 0, sd_g = sd_g, nu = df, xi = 1, num_sim = 1e5, approx = FALSE)
n_min
```

```
## [1] 67.93561
```

With approximation:

$$n_{g\nu} = n_g^{\frac{1}{\kappa_{1,2}-1}}$$

```
n_g <- 30 ## Number of Gaussian terms
sd_g = 1
df = 3
```

```
n_min <- f_n_min(n_g, mean = 0, sd_g = sd_g, nu = df, xi = 1, num_sim = 1e5, approx = TRUE)
n_min
```

```
## [1] 116.1906
```