Pension returns analysis

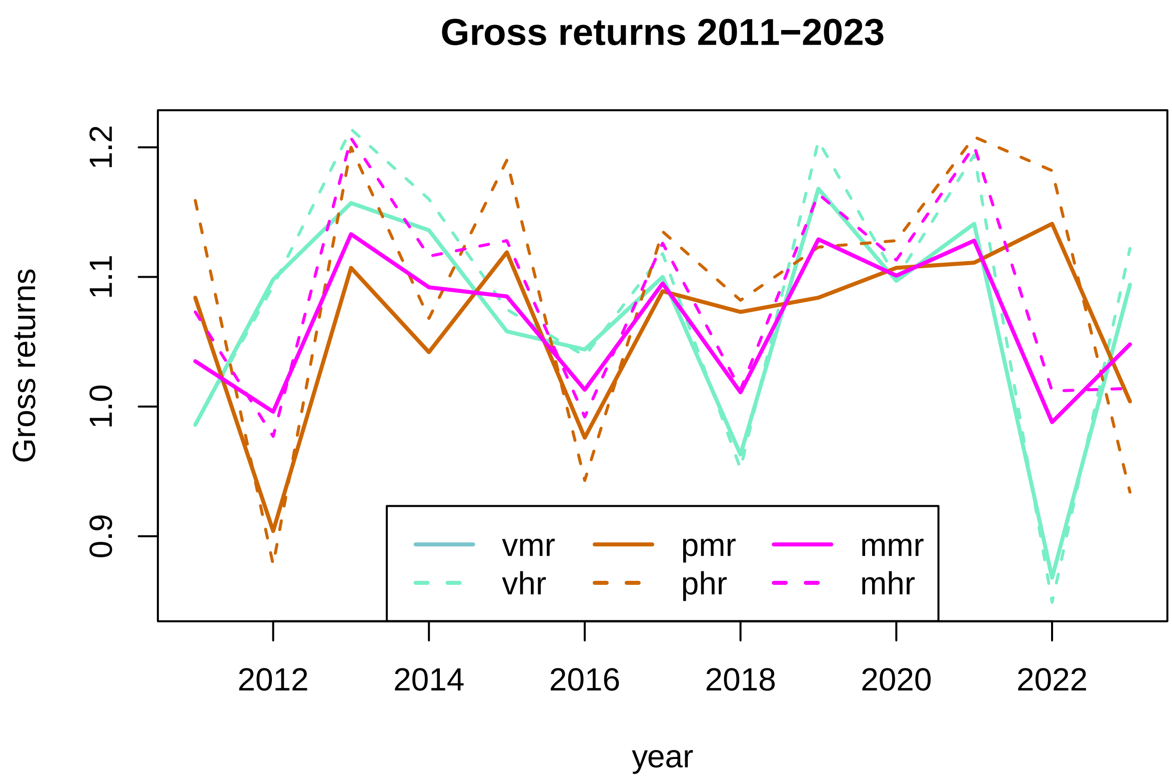
15:11 06 May 2024

Contents

Fit log returns to F-S skew standardized Student-t distribution.
`m` is the location parameter.
`s` is the scale parameter.
`nu` is the estimated shape parameter (degrees of freedom).
`xi` is the estimated skewness parameter.

Returns data 2011-2023.

For 2011, medium risk data is used in the high risk data set, as no high risk fund data is available prior to 2012.
`vmrl` is a long version of Velliv medium risk data, from 2007 to 2023. For 2007 to 2011 (both included) no high risk data is available.



Summary of log-returns

The summary statistics are transformed back to the scale of gross returns by taking $exp()$ of each summary statistic. (Note: Taking arithmetic mean of gross returns directly is no good. Must be geometric mean.)

	vmr	vhr	vmrl	pmr	phr	mmr	mhr	vmr_phr	vhr_pmr
Min. :	0.868	0.849	0.801	0.904	0.878	0.988	0.977	0.979	0.967
1st Qu.:	1.044	1.039	1.013	1.042	1.068	1.013	1.013	1.021	1.011
Median :	1.097	1.099	1.085	1.084	1.128	1.085	1.113	1.102	1.094
Mean :	1.067	1.080	1.057	1.063	1.089	1.064	1.085	1.079	1.072
3rd Qu.:	1.136	1.160	1.128	1.107	1.182	1.101	1.128	1.121	1.107
Max. :	1.168	1.214	1.193	1.141	1.208	1.133	1.207	1.178	1.163

Ranking

Min. :	ranking	1st Qu.:	ranking	Median :	ranking	Mean :	ranking	3rd Qu.:	ranking	Max. :	ranking
0.988	mmr	1.068	phr	1.128	phr	1.089	phr	1.182	phr	1.214	vhr
0.979	vmr_phr	1.044	vmr	1.113	mhr	1.085	mhr	1.160	vhr	1.208	phr
0.977	mhr	1.042	pmr	1.102	vmr_phr	1.080	vhr	1.136	vmr	1.207	mhr
0.967	vhr_pmr	1.039	vhr	1.099	vhr	1.079	vmr_phr	1.128	vmrl	1.193	vmrl
0.904	pmr	1.021	vmr_phr	1.097	vmr	1.072	vhr_pmr	1.128	mhr	1.178	vmr_phr
0.878	phr	1.013	vmrl	1.094	vhr_pmr	1.067	vmr	1.121	vmr_phr	1.168	vmr
0.868	vmr	1.013	mmr	1.085	vmrl	1.064	mmr	1.107	pmr	1.163	vhr_pmr
0.849	vhr	1.013	mhr	1.085	mmr	1.063	pmr	1.107	vhr_pmr	1.141	pmr

Min. :	ranking	1st Qu.:	ranking	Median :	ranking	Mean :	ranking	3rd Qu.:	ranking	Max. :	ranking
0.801	vmrl	1.011	vhr_pmr	1.084	pmr	1.057	vmrl	1.101	mmr	1.133	mmr

Correlations and covariance

Correlations

	vmr	vhr	pmr	phr
vmr	1.000	0.993	-0.197	-0.095
vhr	0.993	1.000	-0.119	-0.016
pmr	-0.197	-0.119	1.000	0.957
phr	-0.095	-0.016	0.957	1.000

Covariances

	vmr	vhr	pmr	phr
vmr	0.007	0.009	-0.001	-0.001
vhr	0.009	0.011	-0.001	0.000
pmr	-0.001	-0.001	0.004	0.007
phr	-0.001	0.000	0.007	0.011

Compare pension plans

Risk of loss

Risk of loss at least as big as x percent for a single period (year).
x values are row names.

	vmr	vhr	pmr	phr	mmr	mhr	vmr_phr	vhr_pmr
0	21.167	21.333	11.833	14.000	12.333	12.667	16.667	16.000
5	12.167	13.167	5.667	8.333	5.833	3.833	8.667	8.167
10	7.000	8.000	3.000	5.000	2.833	0.500	4.333	4.167
25	1.333	1.500	0.500	1.000	0.333	0.000	0.333	0.333
50	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
90	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
99	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Worst ranking for loss percentiles

0	ranking	5	ranking	10	ranking	25	ranking	50	ranking	90	ranking	99	ranking
21.333	vhr	13.167	vhr	8.000	vhr	1.500	vhr	0	vmr	0	vmr	0	vmr
21.167	vmr	12.167	vmr	7.000	vmr	1.333	vmr	0	vhr	0	vhr	0	vhr
16.667	vmr_phr	8.667	vmr_phr	5.000	phr	1.000	phr	0	pmr	0	pmr	0	pmr
16.000	vhr_pmr	8.333	phr	4.333	vmr_phr	0.500	pmr	0	phr	0	phr	0	phr
14.000	phr	8.167	vhr_pmr	4.167	vhr_pmr	0.333	mmr	0	mmr	0	mmr	0	mmr
12.667	mhr	5.833	mmr	3.000	pmr	0.333	vmr_phr	0	mhr	0	mhr	0	mhr
12.333	mmr	5.667	pmr	2.833	mmr	0.333	vhr_pmr	0	vmr_phr	0	vmr_phr	0	vmr_phr
11.833	pmr	3.833	mhr	0.500	mhr	0.000	mhr	0	vhr_pmr	0	vhr_pmr	0	vhr_pmr

Chance of min gains

Chance of gains of at least x percent for a single period (year).
x values are row names.

	vmr	vhr	pmr	phr	mmr	mhr	vmr_phr	vhr_pmr
0	78.833	78.667	88.167	86.000	87.667	87.333	83.333	84.000
5	63.833	66.667	71.667	76.000	71.667	70.167	69.333	69.000
10	40.833	50.167	32.500	59.667	35.500	46.000	47.167	43.833
25	0.000	0.000	0.000	0.000	0.000	0.833	0.000	0.000
50	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
100	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Best ranking for gains percentiles

0	ranking	5	ranking	10	ranking	25	ranking	50	ranking	100	ranking
88.167	pmr	76.000	phr	59.667	phr	0.833	mhr	0	vmr	0	vmr
87.667	mmr	71.667	pmr	50.167	vhr	0.000	vmr	0	vhr	0	vhr
87.333	mhr	71.667	mmr	47.167	vmr_phr	0.000	vhr	0	pmr	0	pmr
86.000	phr	70.167	mhr	46.000	mhr	0.000	pmr	0	phr	0	phr
84.000	vhr_pmr	69.333	vmr_phr	43.833	vhr_pmr	0.000	phr	0	mmr	0	mmr
83.333	vmr_phr	69.000	vhr_pmr	40.833	vmr	0.000	mmr	0	mhr	0	mhr
78.833	vmr	66.667	vhr	35.500	mmr	0.000	vmr_phr	0	vmr_phr	0	vmr_phr
78.667	vhr	63.833	vmr	32.500	pmr	0.000	vhr_pmr	0	vhr_pmr	0	vhr_pmr

MC risk percentiles

Risk of loss at least as big as row name in percent from first to last period.

	vmr	vhr	pmr	phr	mmr	mhr	vmr_phr	vhr_pmr
0	4.81	2.84	1.74	1.07	0.41	0.10	0.16	0.18
5	4.16	2.39	1.59	1.02	0.33	0.08	0.15	0.16
10	3.49	1.93	1.36	0.93	0.28	0.06	0.12	0.11
25	2.12	1.06	1.01	0.60	0.10	0.03	0.04	0.08
50	0.87	0.33	0.51	0.31	0.02	0.01	0.01	0.01
90	0.07	0.03	0.09	0.03	0.00	0.00	0.00	0.00
99	0.02	0.01	0.05	0.01	0.00	0.00	0.00	0.00

1e6 simulation paths of `mhr`:

	0	5	10	25	50	90	99
prob_pct	0.118	0.095	0.076	0.036	0.008	0	0

Worst ranking for MC loss percentiles

0	ranking	5	ranking	10	ranking	25	ranking	50	ranking	90	ranking	99	ranking
4.81	vmr	4.16	vmr	3.49	vmr	2.12	vmr	0.87	vmr	0.09	pmr	0.05	pmr
2.84	vhr	2.39	vhr	1.93	vhr	1.06	vhr	0.51	pmr	0.07	vmr	0.02	vmr
1.74	pmr	1.59	pmr	1.36	pmr	1.01	pmr	0.33	vhr	0.03	vhr	0.01	vhr
1.07	phr	1.02	phr	0.93	phr	0.60	phr	0.31	phr	0.03	phr	0.01	phr
0.41	mmr	0.33	mmr	0.28	mmr	0.10	mmr	0.02	mmr	0.00	mmr	0.00	mmr
0.18	vhr_pmr	0.16	vhr_pmr	0.12	vmr_phr	0.08	vhr_pmr	0.01	mhr	0.00	mhr	0.00	mhr
0.16	vmr_phr	0.15	vmr_phr	0.11	vhr_pmr	0.04	vmr_phr	0.01	vmr_phr	0.00	vmr_phr	0.00	vmr_phr
0.10	mhr	0.08	mhr	0.06	mhr	0.03	mhr	0.01	vhr_pmr	0.00	vhr_pmr	0.00	vhr_pmr

MC gains percentiles

	vmr	vhr	pmr	phr	mmr	mhr	vmr_phr	vhr_pmr
0	95.19	97.16	98.26	98.93	99.59	99.90	99.84	99.82

	vmr	vhr	pmr	phr	mmr	mhr	vmr_phr	vhr_pmr
5	94.52	96.84	98.03	98.84	99.49	99.89	99.77	99.78
10	93.73	96.43	97.84	98.68	99.37	99.88	99.75	99.74
25	91.25	94.88	97.12	98.20	98.85	99.79	99.50	99.53
50	85.80	91.67	95.25	97.31	97.56	99.47	98.99	98.90
100	72.20	83.01	88.55	94.67	89.93	97.65	96.24	94.23
200	39.42	60.89	59.51	85.27	48.53	86.42	80.34	66.41
300	16.21	39.24	22.32	70.63	11.13	62.93	52.20	30.58
400	5.17	23.88	4.42	54.36	1.26	37.79	25.70	9.69
500	1.51	12.81	0.54	38.37	0.09	18.77	9.84	2.54
1000	0.00	0.28	0.01	2.15	0.02	0.06	0.00	0.00

1e6 simulation paths of `mhr`:

	0	5	10	25	50	100	200	300	400	500	1000
prob	99.882	99.854	99.824	99.686	99.301	97.513	86.912	65.992	41.486	21.693	0.086

Best ranking for MC gains percentiles

0	ranking	5	ranking	10	ranking	25	ranking	50	ranking	100	ranking
99.90	mhr	99.89	mhr	99.88	mhr	99.79	mhr	99.47	mhr	97.65	mhr
99.84	vmr_phr	99.78	vhr_pmr	99.75	vmr_phr	99.53	vhr_pmr	98.99	vmr_phr	96.24	vmr_phr
99.82	vhr_pmr	99.77	vmr_phr	99.74	vhr_pmr	99.50	vmr_phr	98.90	vhr_pmr	94.67	phr
99.59	mmr	99.49	mmr	99.37	mmr	98.85	mmr	97.56	mmr	94.23	vhr_pmr
98.93	phr	98.84	phr	98.68	phr	98.20	phr	97.31	phr	89.93	mmr
98.26	pmr	98.03	pmr	97.84	pmr	97.12	pmr	95.25	pmr	88.55	pmr
97.16	vhr	96.84	vhr	96.43	vhr	94.88	vhr	91.67	vhr	83.01	vhr
95.19	vmr	94.52	vmr	93.73	vmr	91.25	vmr	85.80	vmr	72.20	vmr

200	ranking	300	ranking	400	ranking	500	ranking	1000	ranking
86.42	mhr	70.63	phr	54.36	phr	38.37	phr	2.15	phr
85.27	phr	62.93	mhr	37.79	mhr	18.77	mhr	0.28	vhr
80.34	vmr_phr	52.20	vmr_phr	25.70	vmr_phr	12.81	vhr	0.06	mhr
66.41	vhr_pmr	39.24	vhr	23.88	vhr	9.84	vmr_phr	0.02	mmr
60.89	vhr	30.58	vhr_pmr	9.69	vhr_pmr	2.54	vhr_pmr	0.01	pmr
59.51	pmr	22.32	pmr	5.17	vmr	1.51	vmr	0.00	vmr
48.53	mmr	16.21	vmr	4.42	pmr	0.54	pmr	0.00	vmr_phr
39.42	vmr	11.13	mmr	1.26	mmr	0.09	mmr	0.00	vhr_pmr

Summary statistics

Fit summary

Summary for fit of log returns to an F-S skew standardized Student-t distribution.
`m` is the location parameter.
`s` is the scale parameter.
`nu` is the estimated degrees of freedom, or shape parameter.
`xi` is the estimated skewness parameter.

	vmr	vhr	pmr	phr	mmr	mhr	vmr_phr	vhr_pmr
m	0.048	0.063	0.058	0.084	0.059	0.082	0.067	0.062
s	0.120	0.126	0.123	0.121	0.088	0.071	0.091	0.090
nu	3.304	4.390	2.265	3.185	2.773	89.863	4.660	3.892
xi	0.034	0.019	0.477	0.018	0.029	0.770	0.048	0.019
R^2	0.993	0.995	0.991	0.964	0.890	0.961	0.927	0.933

Fit statistics ranking

m	ranking	s	ranking	R^2	ranking
0.084	phr	0.071	mhr	0.995	vhr
0.082	mhr	0.088	mmr	0.993	vmr
0.067	vmr_phr	0.090	vhr_pmr	0.991	pmr
0.063	vhr	0.091	vmr_phr	0.964	phr
0.062	vhr_pmr	0.120	vmr	0.961	mhr
0.059	mmr	0.121	phr	0.933	vhr_pmr
0.058	pmr	0.123	pmr	0.927	vmr_phr
0.048	vmr	0.126	vhr	0.890	mmr

Monte Carlo simulations summary

Monte Carlo simulations of portfolio index values (currency values).
Statistics are given for the final state of all paths.
Probability of down-and_out is calculated as the share of paths that reach 0 at some point. All subsequent values for a path are set to 0, if the path reaches at any point.
0 is defined as any value below a threshold.
`dai_pct` (for down-and-in) is the probability of losing money. This is calculated as the share of paths finishing below index 100.

```
## Number of paths: 10000
```

	vmr	vhr	pmr	phr	mmr	mhr	vmr_phr	vhr_pmr
mc_m	295.71	406.85	344.57	601.85	319.20	504.75	451.78	378.19
mc_s	135.11	211.26	114.07	273.01	102.36	173.81	152.58	121.44
mc_min	0.12	0.55	0.00	0.01	40.95	26.81	47.30	40.29
mc_max	1036.78	1504.07	1308.32	1930.64	4106.75	1414.95	1125.82	1097.82
dao_pct	0.00	0.00	0.03	0.01	0.00	0.00	0.00	0.00
dai_pct	4.47	2.46	1.66	0.99	0.42	0.10	0.14	0.15

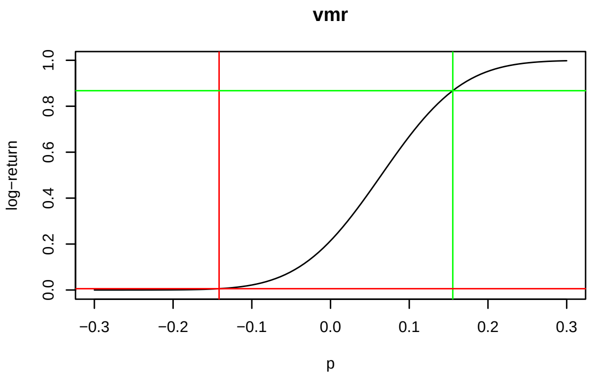
Ranking

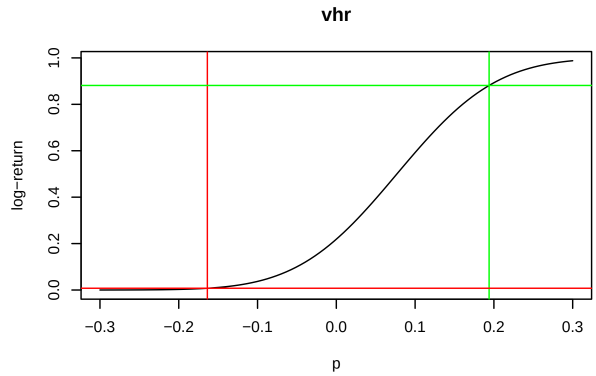
mc_m	ranking	mc_s	ranking	mc_min	ranking	mc_max	ranking	dao_pct	ranking	dai_pct	ranking
601.85	phr	102.36	mmr	47.30	vmr_phr	4106.75	mmr	0.00	vmr	0.10	mhr
504.75	mhr	114.07	pmr	40.95	mmr	1930.64	phr	0.00	vhr	0.14	vmr_phr
451.78	vmr_phr	121.44	vhr_pmr	40.29	vhr_pmr	1504.07	vhr	0.00	mmr	0.15	vhr_pmr
406.85	vhr	135.11	vmr	26.81	mhr	1414.95	mhr	0.00	mhr	0.42	mmr
378.19	vhr_pmr	152.58	vmr_phr	0.55	vhr	1308.32	pmr	0.00	vmr_phr	0.99	phr
344.57	pmr	173.81	mhr	0.12	vmr	1125.82	vmr_phr	0.00	vhr_pmr	1.66	pmr
319.20	mmr	211.26	vhr	0.01	phr	1097.82	vhr_pmr	0.01	phr	2.46	vhr
295.71	vmr	273.01	phr	0.00	pmr	1036.78	vmr	0.03	pmr	4.47	vmr

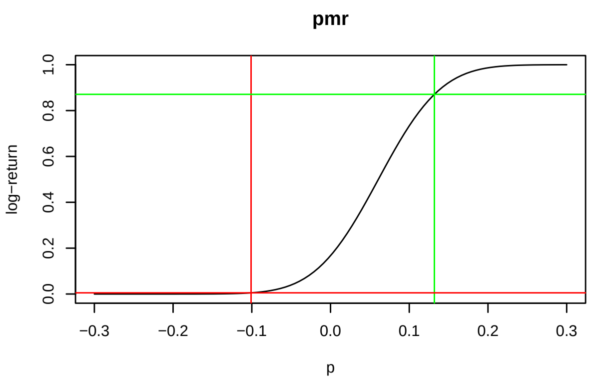
Compare Gaussian and skewed t-distribution fits

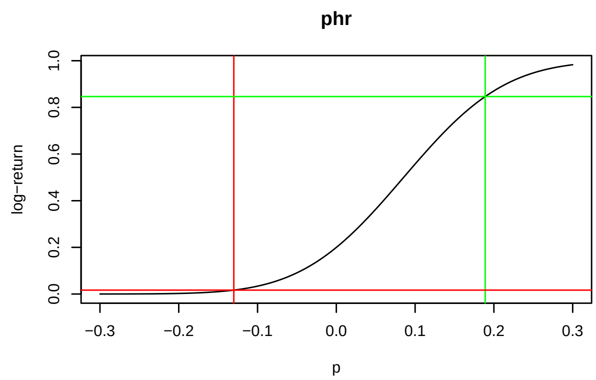
Gaussian fits

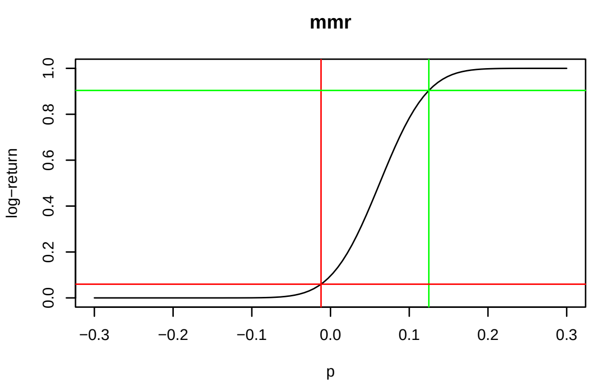
	vmr	vhr	pmr	phr	mmr	mhr	vmr_phr	vhr_pmr
m	0.064	0.077	0.061	0.085	0.062	0.081	0.076	0.069
s	0.081	0.099	0.063	0.101	0.048	0.070	0.062	0.060

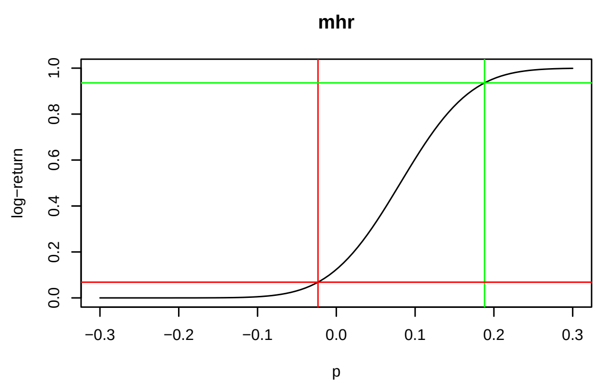

vmr
log−return
0.0 0.2 0.4 0.6 0.8 1.0
−0.3 −0.2 −0.1 0.0 0.1 0.2 0.3
p

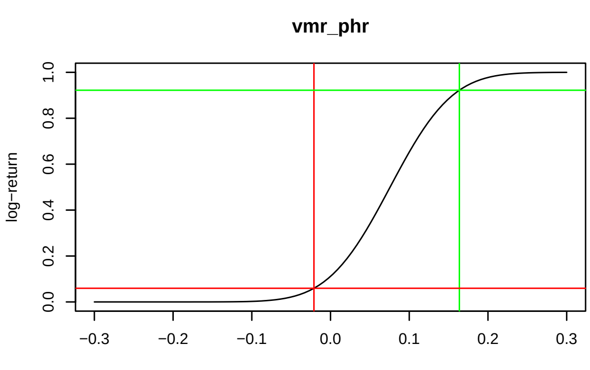

vhr
log−return
0.0 0.2 0.4 0.6 0.8 1.0
−0.3 −0.2 −0.1 0.0 0.1 0.2 0.3
p

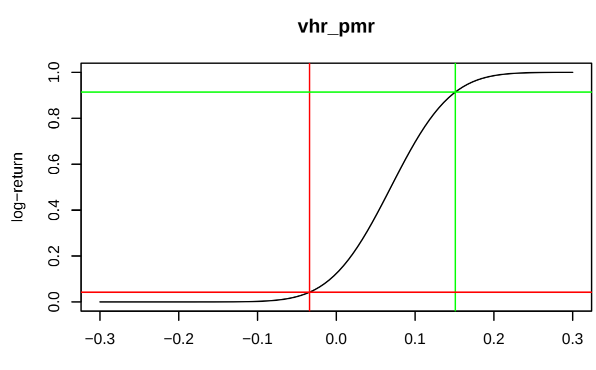

pmr
log−return
0.0 0.2 0.4 0.6 0.8 1.0
−0.3 −0.2 −0.1 0.0 0.1 0.2 0.3
p


phr
log−return
0.0 0.2 0.4 0.6 0.8 1.0
−0.3 −0.2 −0.1 0.0 0.1 0.2 0.3
p

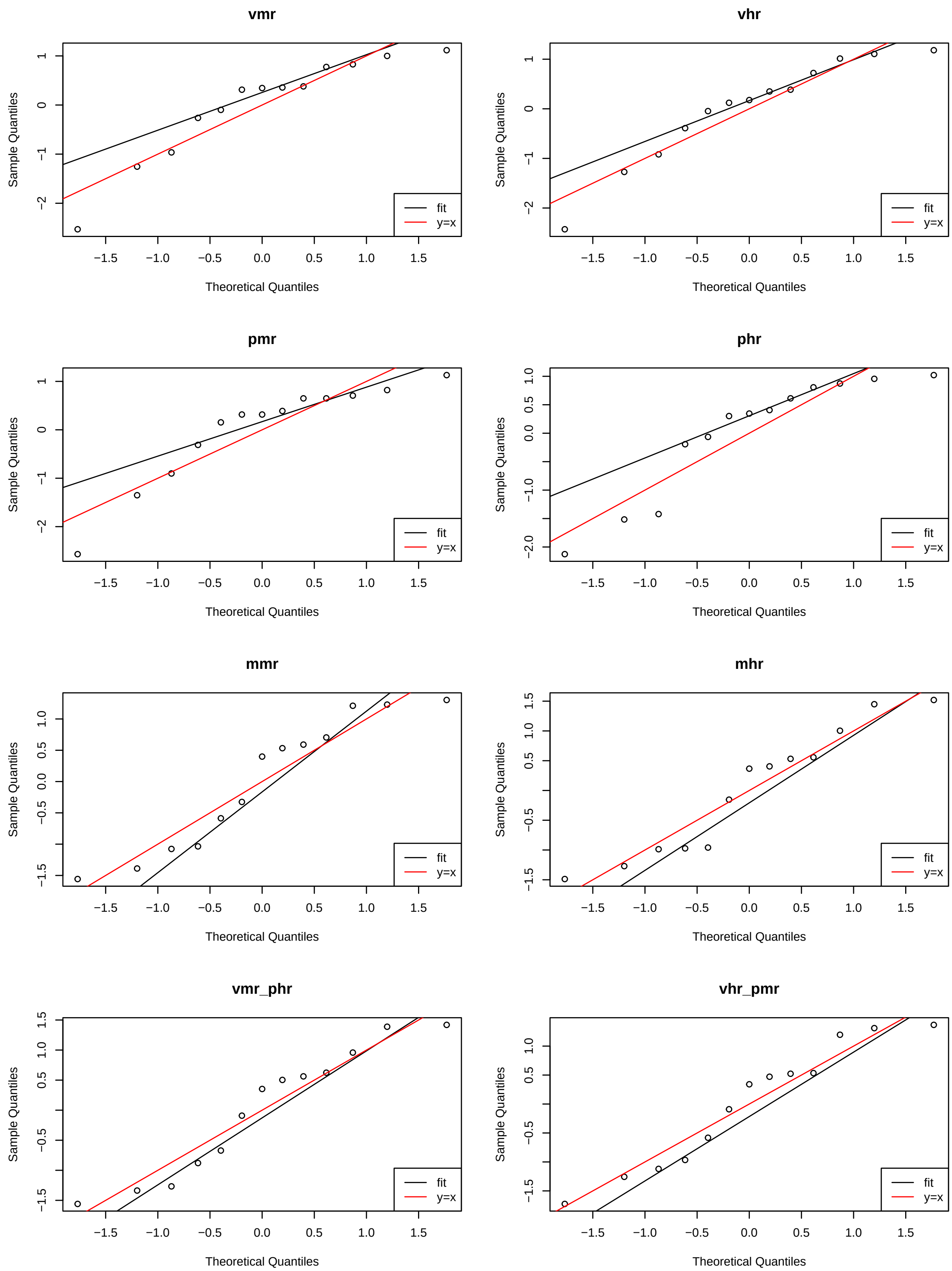

mmr
log−return
0.0 0.2 0.4 0.6 0.8 1.0
−0.3 −0.2 −0.1 0.0 0.1 0.2 0.3
p


mhr
log−return
0.0 0.2 0.4 0.6 0.8 1.0
−0.3 −0.2 −0.1 0.0 0.1 0.2 0.3
p


vmr_phr
log−return
0.0 0.2 0.4 0.6 0.8 1.0
−0.3 −0.2 −0.1 0.0 0.1 0.2 0.3
p


vhr_pmr
log−return
0.0 0.2 0.4 0.6 0.8 1.0
−0.3 −0.2 −0.1 0.0 0.1 0.2 0.3
p

Gaussian QQ plots



Gaussian vs skewed t

Probability in percent that the smallest and largest (respectively) observed return for each fund was generated by a normal distribution:

	vmr	vhr	pmr	phr	mmr	mhr	vmr_phr	vhr_pmr
P_norm(X_min)	0.571	0.758	0.511	1.676	5.971	6.842	5.945	4.228
P_norm(X_max)	13.230	11.876	12.922	15.359	9.628	6.429	7.796	8.592
P_t(X_min)	5.377	5.080	3.489	4.315	10.570	8.015	13.008	10.520
P_t(X_max)	0.118	0.156	2.825	0.188	0.488	5.141	0.229	0.175

Average number of years between min or max events (respectively):

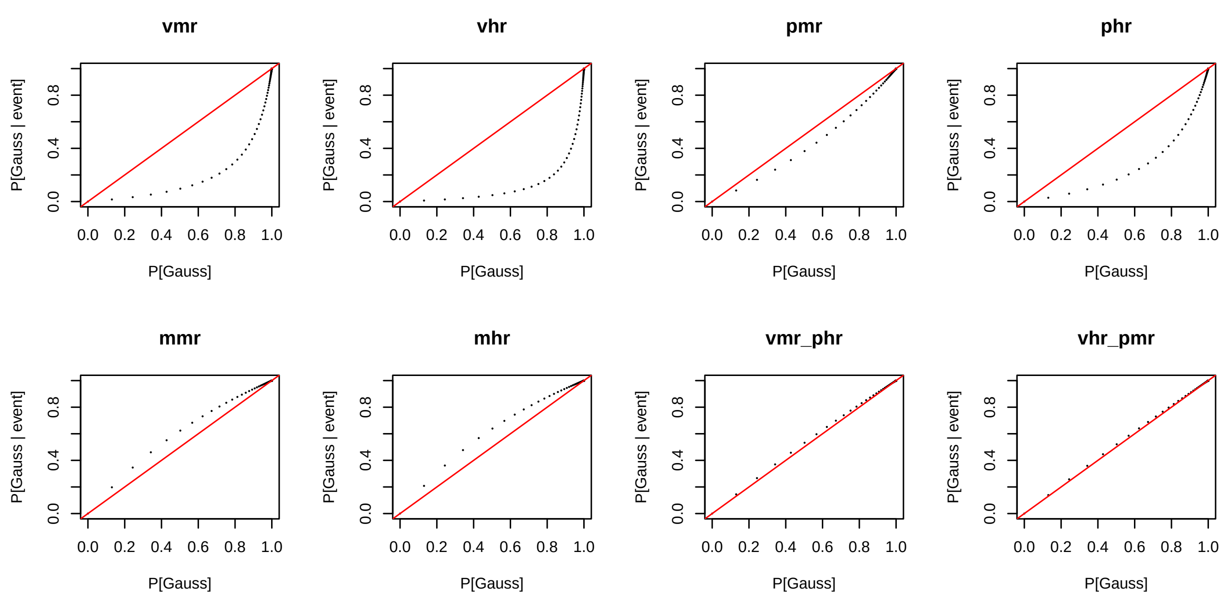
	vmr	vhr	pmr	phr	mmr	mhr	vmr_phr	vhr_pmr
norm: avg yrs btw min	175.248	131.911	195.568	59.669	16.748	14.616	16.820	23.650
norm: avg yrs btw max	7.559	8.420	7.739	6.511	10.386	15.556	12.827	11.639
t: avg yrs btw min	18.596	19.687	28.663	23.173	9.461	12.476	7.688	9.506
t: avg yrs btw max	848.548	640.410	35.400	531.552	205.104	19.450	437.280	572.483

Lilliefors test p-values for Lilliefors test.
Testing H_0, that log-returns are Gaussian.

	vmr	vhr	pmr	phr	mmr	mhr	vmr_phr	vhr_pmr
p value	0.052	0.343	0.024	0.06	0.24	0.137	0.375	0.415

Wittgenstein's Ruler For different given probabilities that returns are Gaussian, what is the probability that the distribution is Gaussian rather than skewed t-distributed, given the smallest/largest observed log-returns?

Conditional probabilities for smallest observed log-returns:

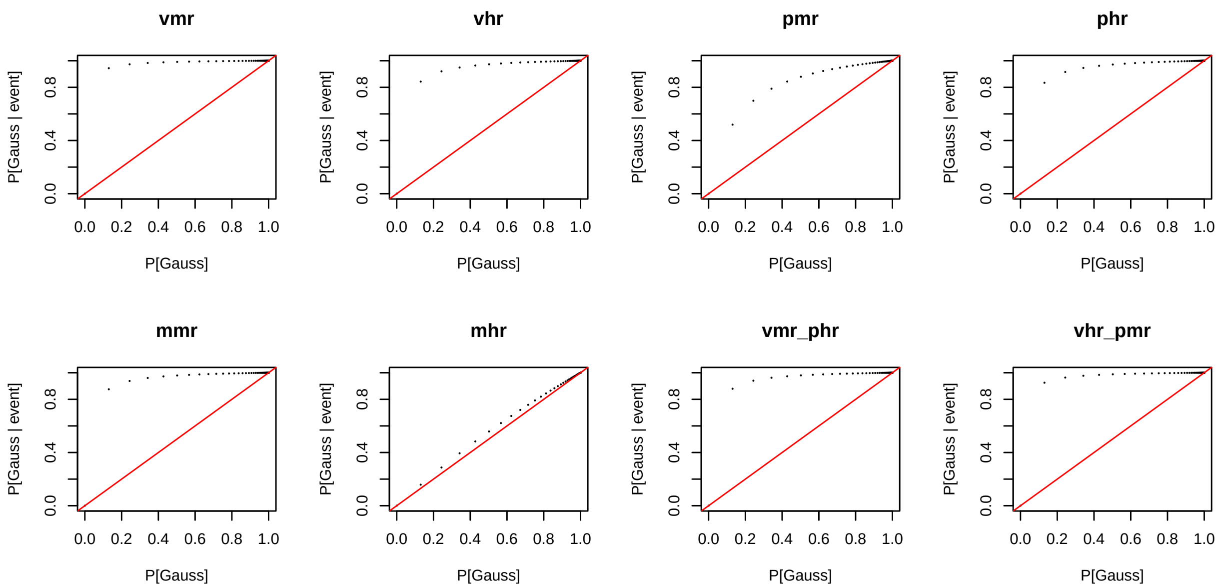


Use $1-$ p-value from Lilliefors test as prior probability that the distribution is Gaussian.
$x_{\text{obs}} = \min(x)$ and $P[\text{Event} \mid \text{Gaussian}] = P_{\text{Gauss}}[X \leq x_{\text{min}}]$:

	vmr	vhr	pmr	phr	mmr	mhr	vmr_phr	vhr_pmr
Lillie p-val	0.052	0.343	0.024	0.060	0.240	0.137	0.375	0.415

	vmr	vhr	pmr	phr	mmr	mhr	vmr_phr	vhr_pmr
Prior prob	0.948	0.657	0.976	0.940	0.760	0.863	0.625	0.585
P[Gauss | Event]	0.661	0.088	0.960	0.754	0.839	0.917	0.653	0.603

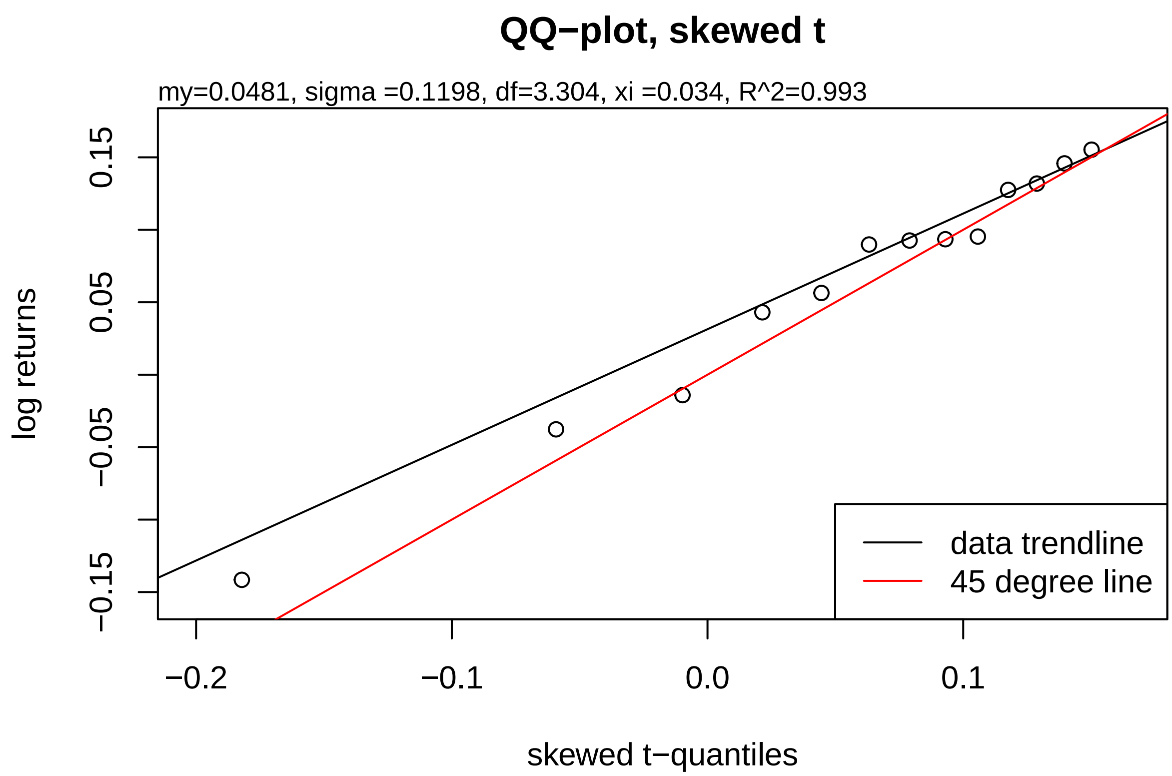
Use $1 -$ p-value from Lilliefors test as prior probability that the distribution is Gaussian.
$x_{\text{obs}} = \max(x)$ and $P[\text{Event} \mid \text{Gaussian}] = P_{\text{Gauss}}[X \geq x_{\text{max}}]$:



	vmr	vhr	pmr	phr	mmr	mhr	vmr_phr	vhr_pmr
Lillie p-val	0.052	0.343	0.024	0.060	0.240	0.137	0.375	0.415
Prior prob	0.948	0.657	0.976	0.940	0.760	0.863	0.625	0.585
P[Gauss | Event]	1.000	0.986	0.997	0.998	0.993	0.888	0.988	0.991

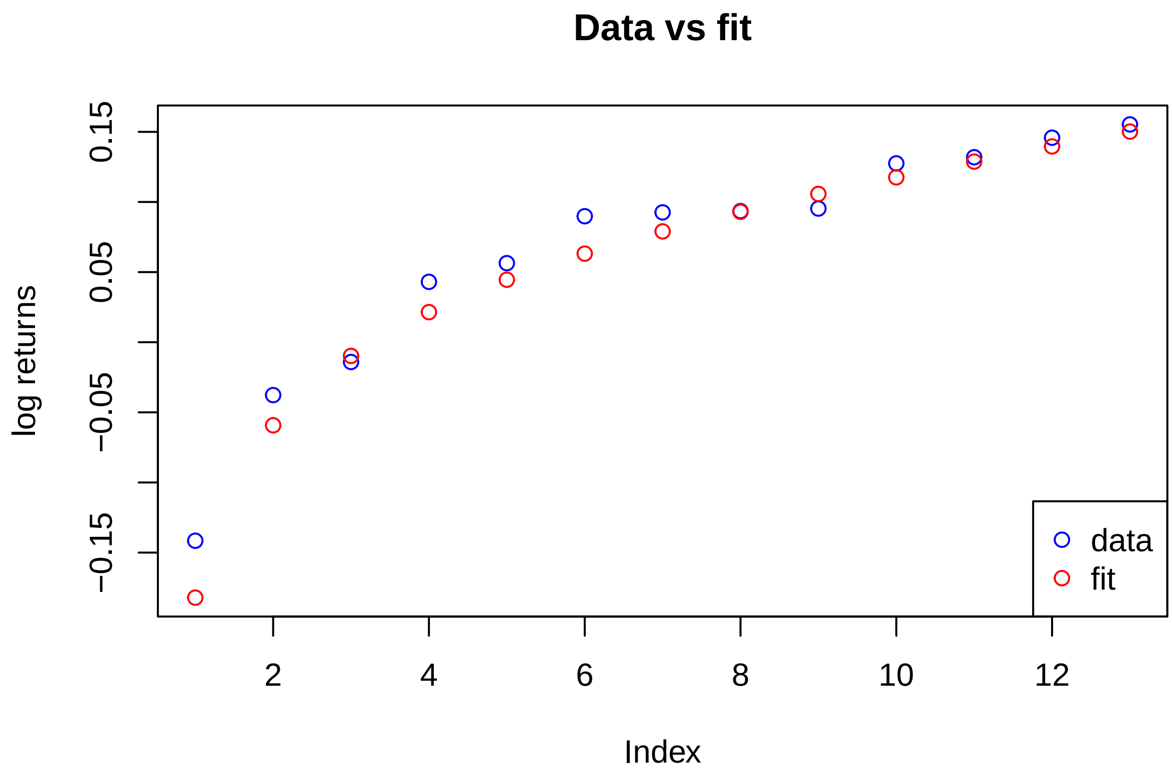
Velliv medium risk (vmr), 2011 - 2023

QQ Plot



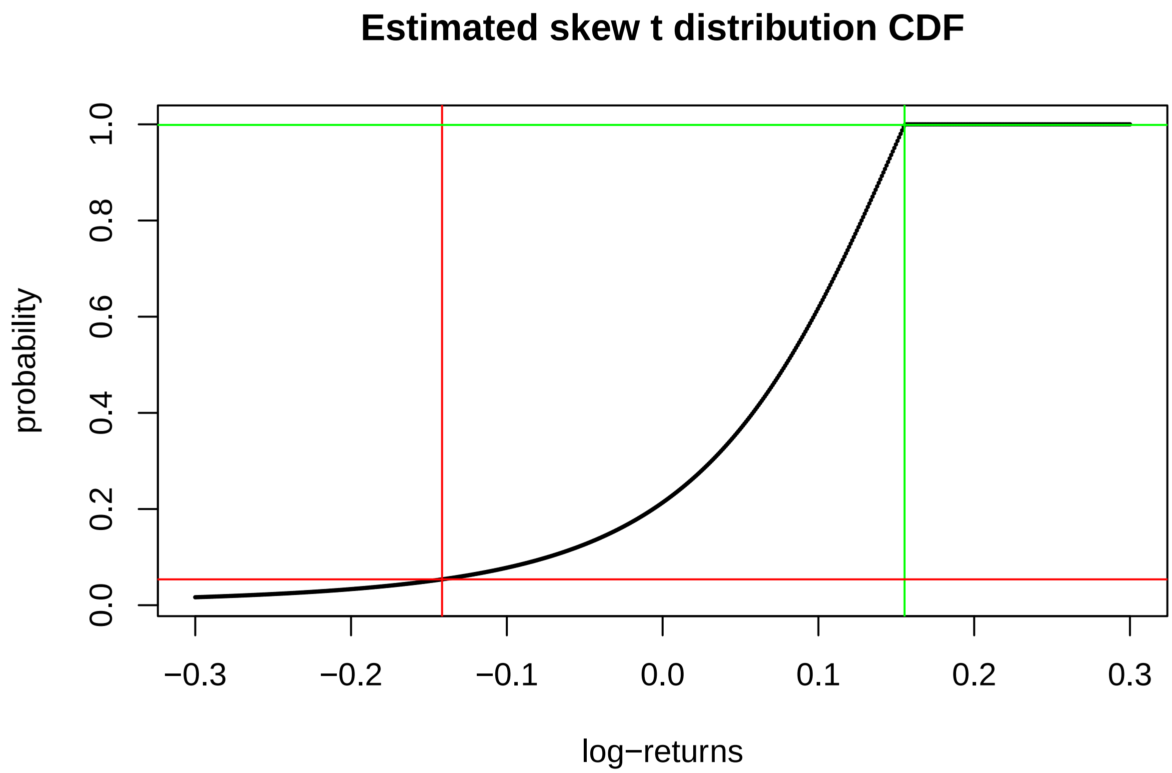
Data vs fit

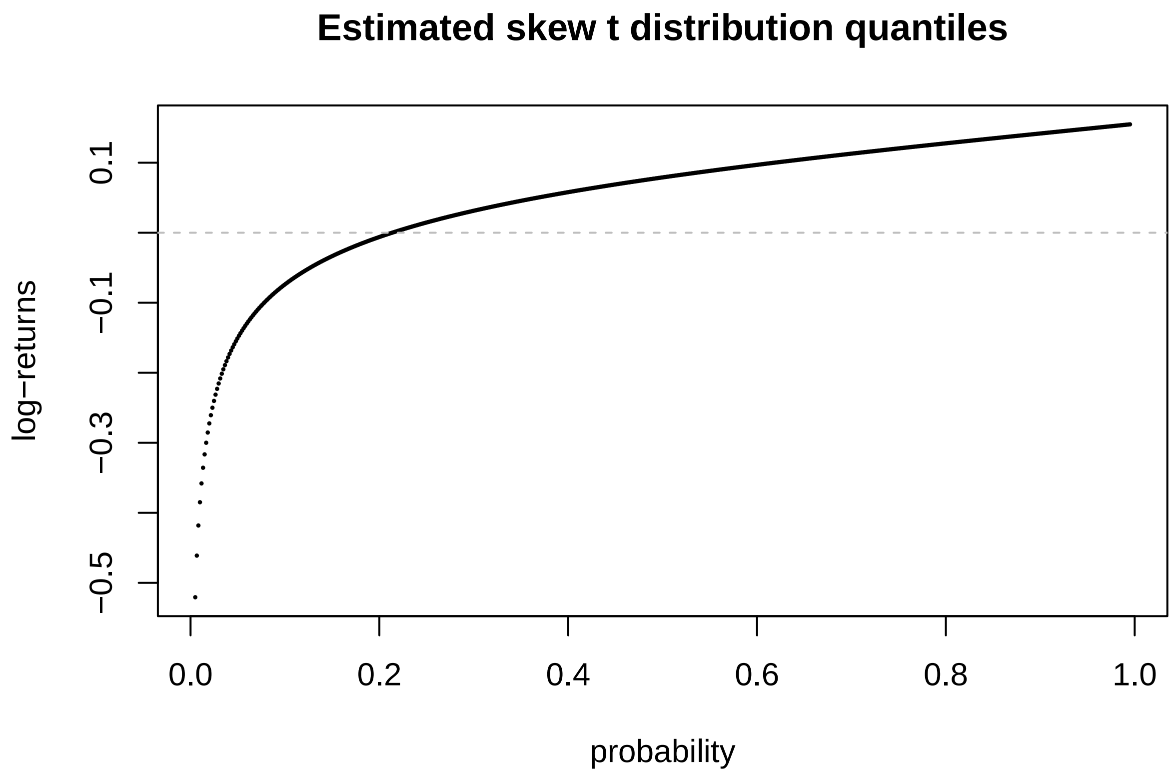
Let’s plot the fit and the observed returns together.

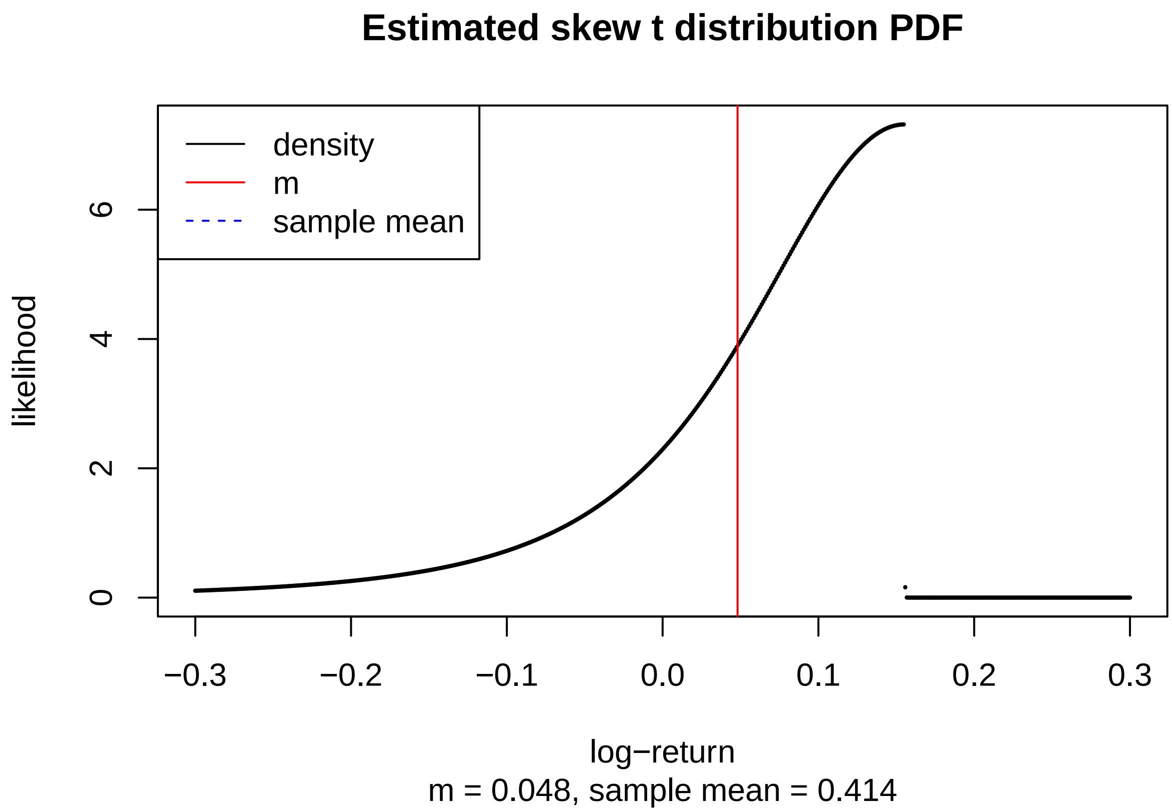


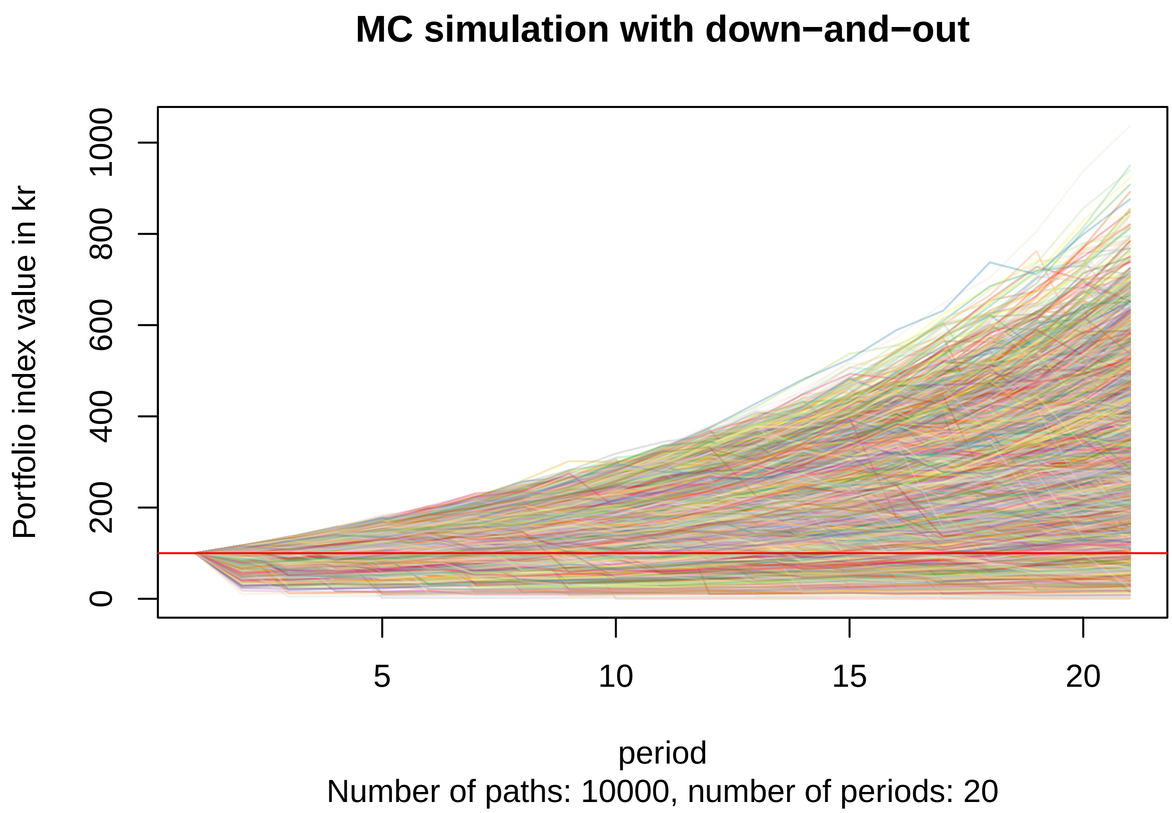
Estimated distribution

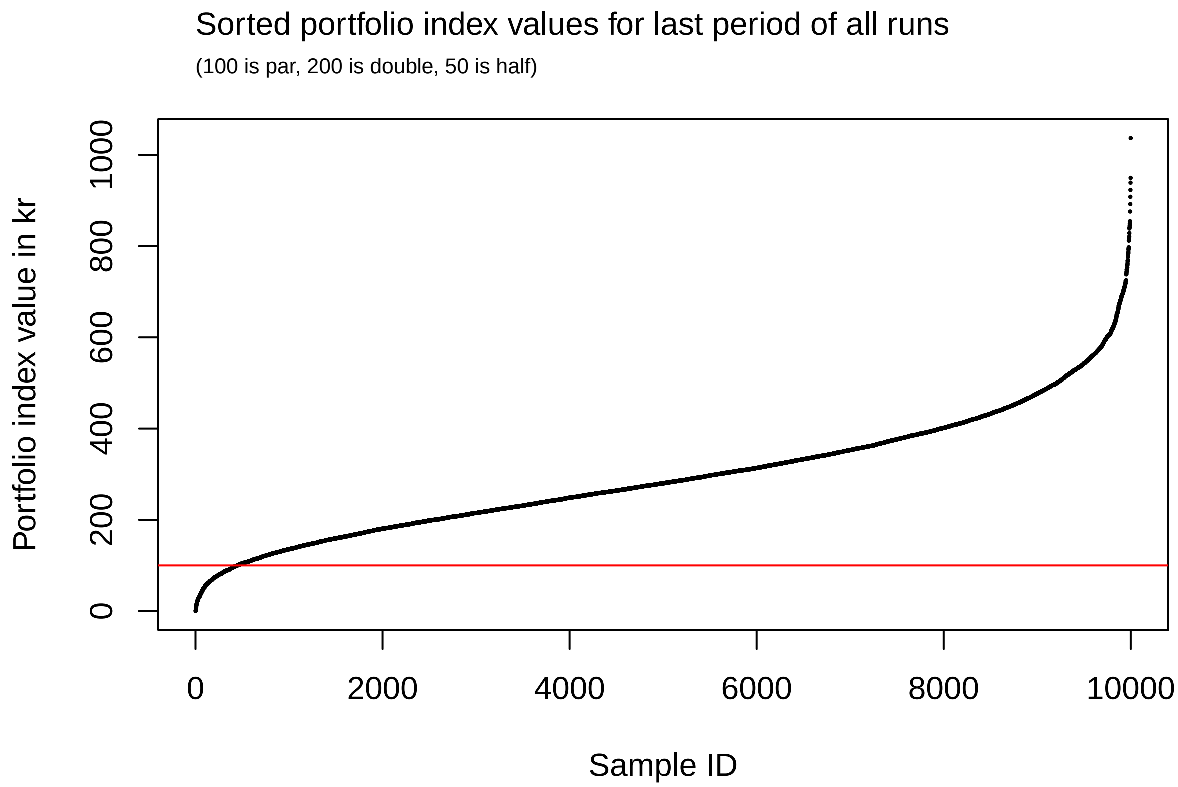
Now lets look at the CDF of the estimated distribution for each 0.1% increment between 0.5% and 99.5% for the estimated distribution:


Estimated skew t distribution CDF
probability
0.0
0.2
0.4
0.6
0.8
1.0
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
log−returns


Estimated skew t distribution quantiles
log−returns
0.1
−0.1
−0.3
−0.5
0.0
0.2
0.4
0.6
0.8
1.0
probability


Estimated skew t distribution PDF
density
m
sample mean
likelihood
0
2
4
6
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
log−return
m = 0.048, sample mean = 0.414

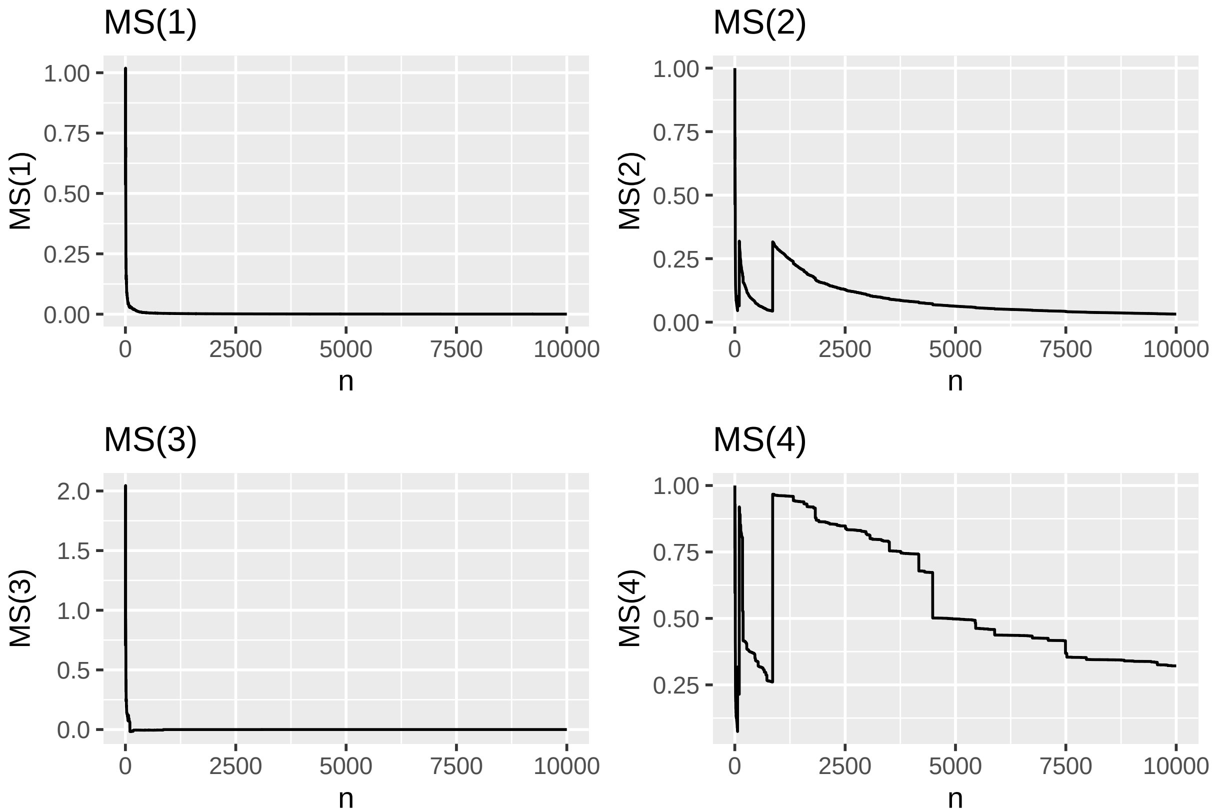

MC simulation with down−and−out
Portfolio index value in kr
0
200
400
600
800
1000
5
10
15
20
period
Number of paths: 10000, number of periods: 20



Convergence

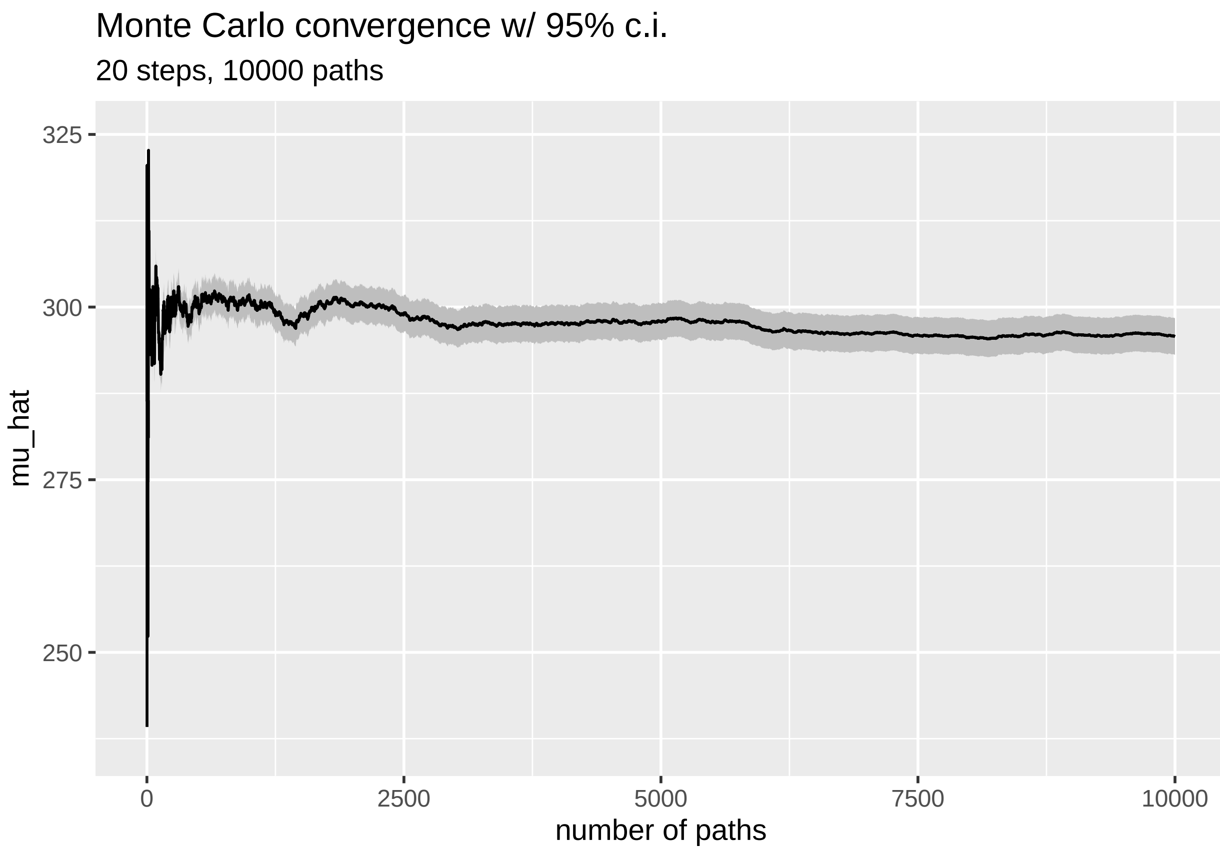
Max vs sum

Max vs sum plots for the first four moments:



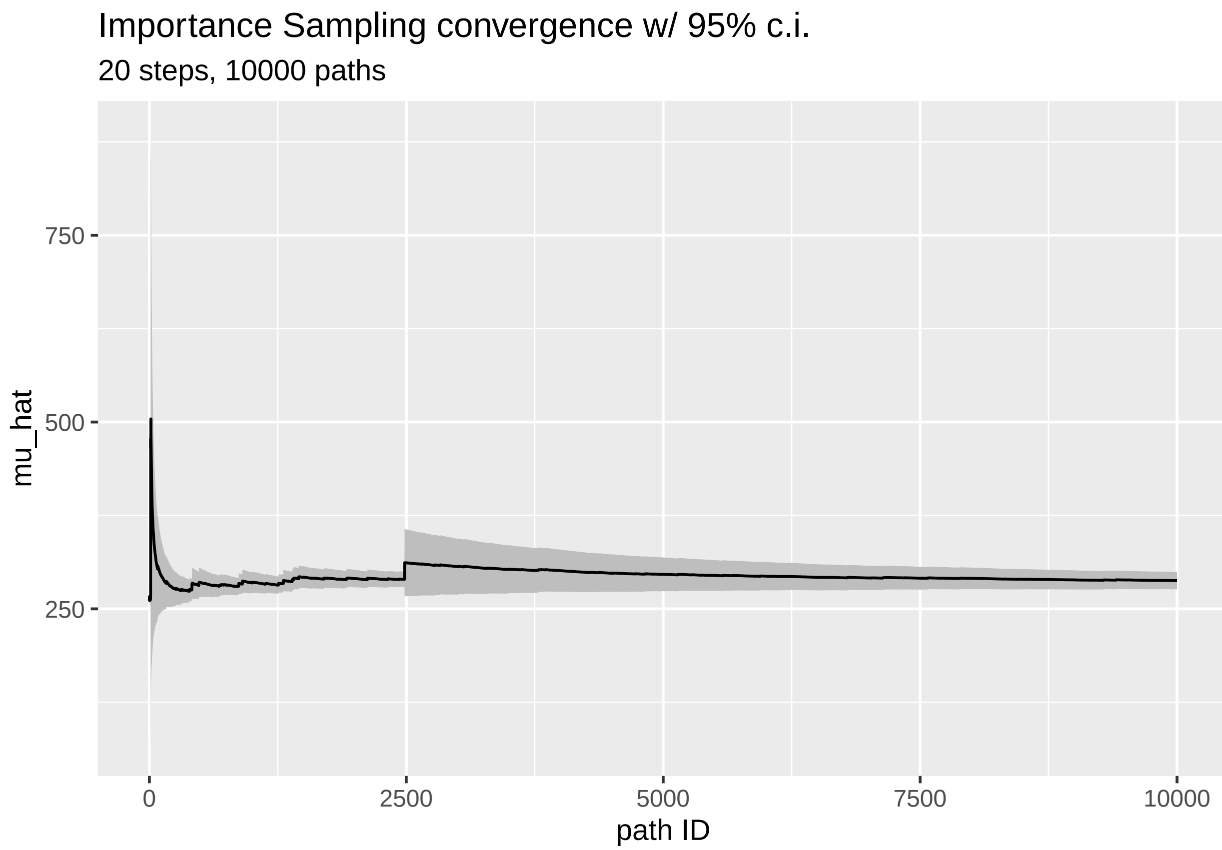
MS(1)
MS(2)
MS(3)
MS(4)
n

MC



Monte Carlo convergence w/ 95% c.i.
20 steps, 10000 paths
325
300
275
250
mu_hat
0
2500
5000
7500
10000
number of paths

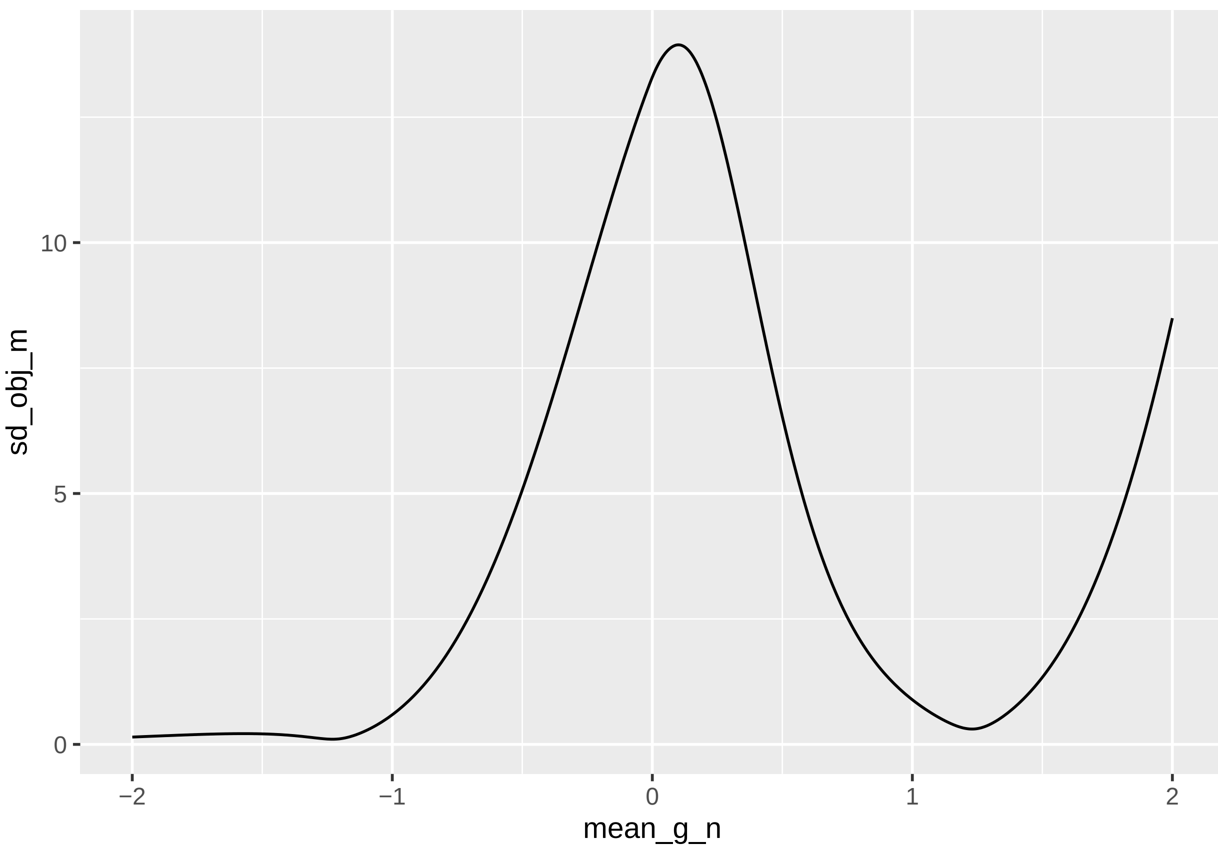
IS

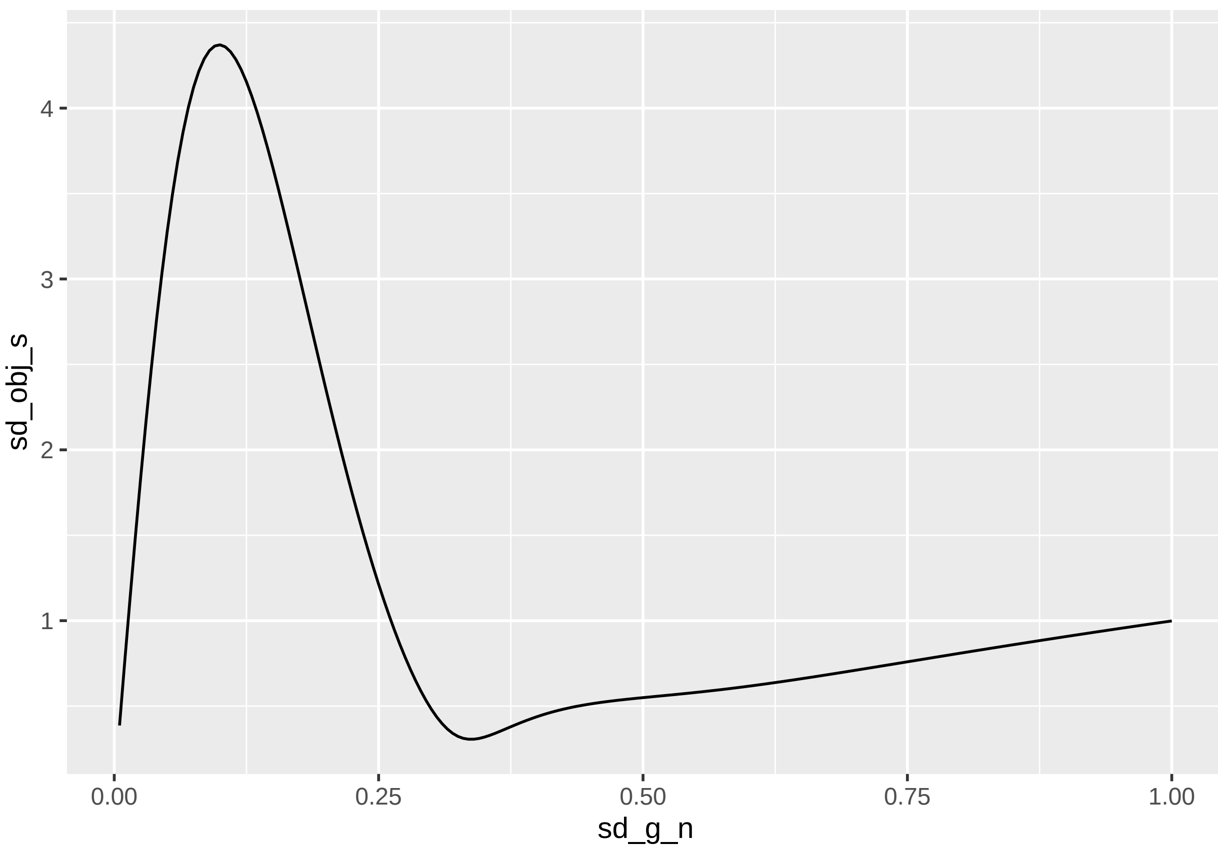


Parameters

```
## [1] 1.2294983 0.3373312
```

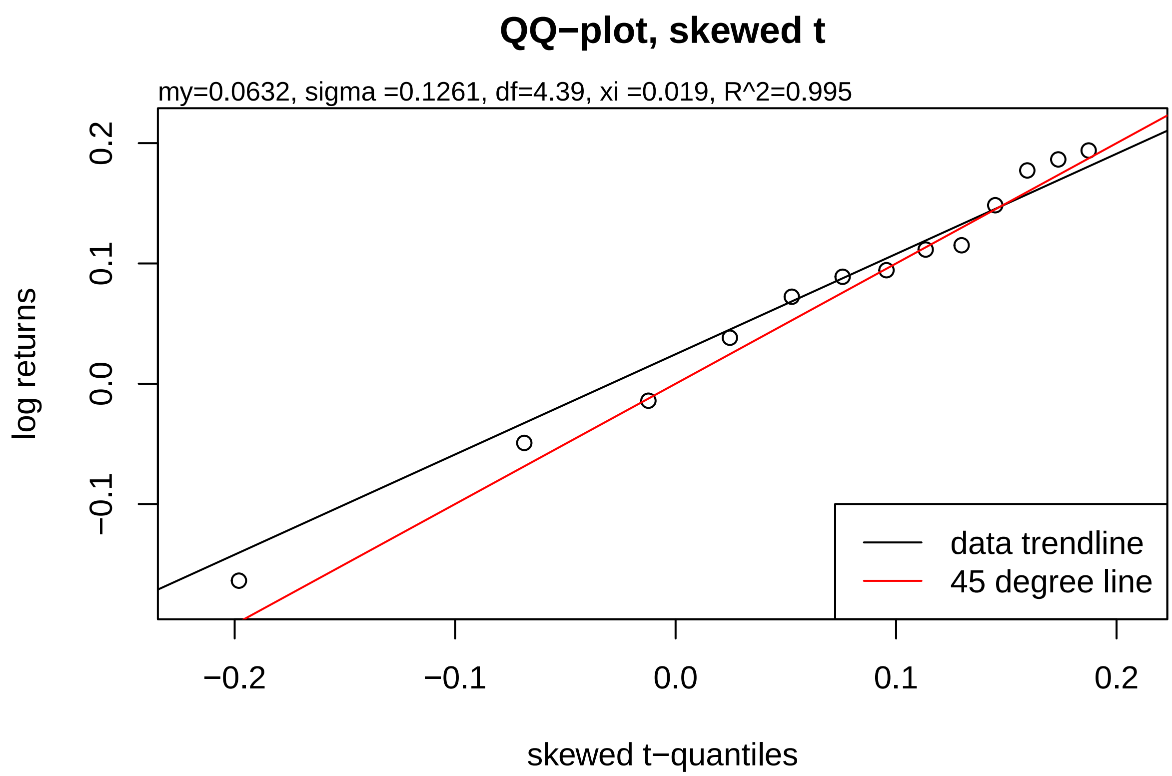
Objective function plots


10
5
0
sd_obj_m
−2
−1
0
1
2
mean_g_n


4
3
2
1
sd_obj_s
0.00
0.25
0.50
0.75
1.00
sd_g_n

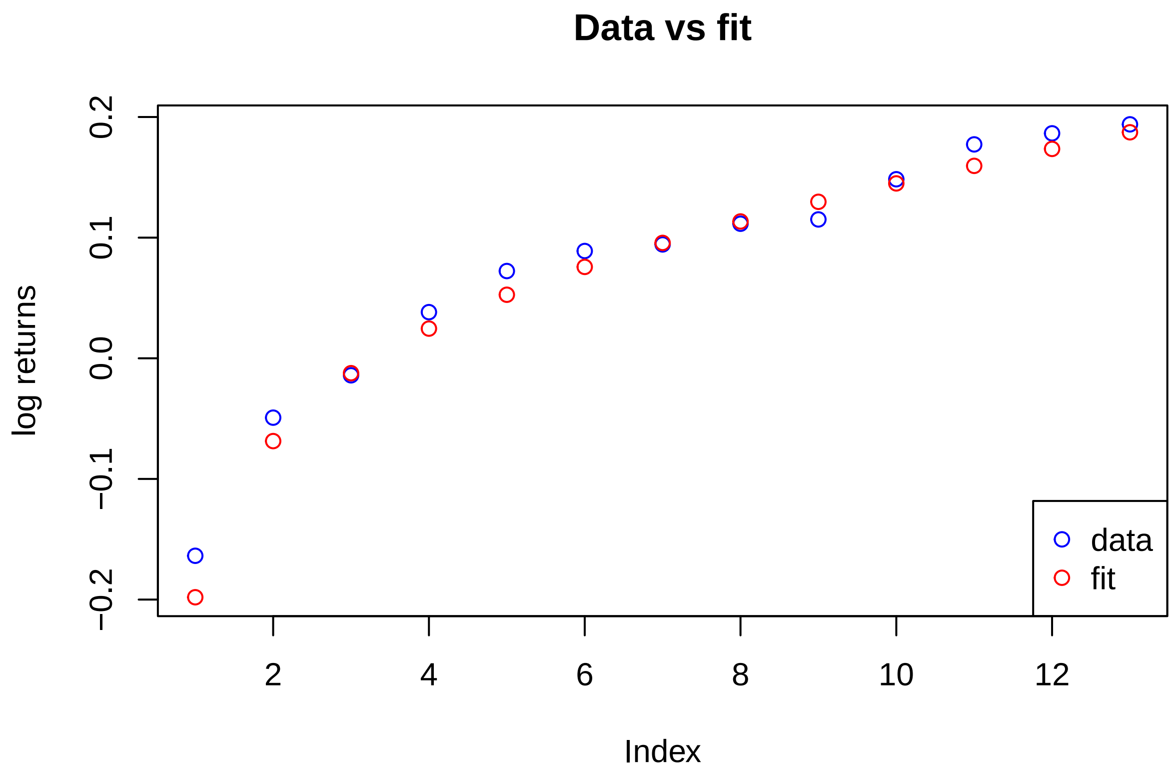
Velliv high risk (vhr), 2011 - 2023

QQ Plot



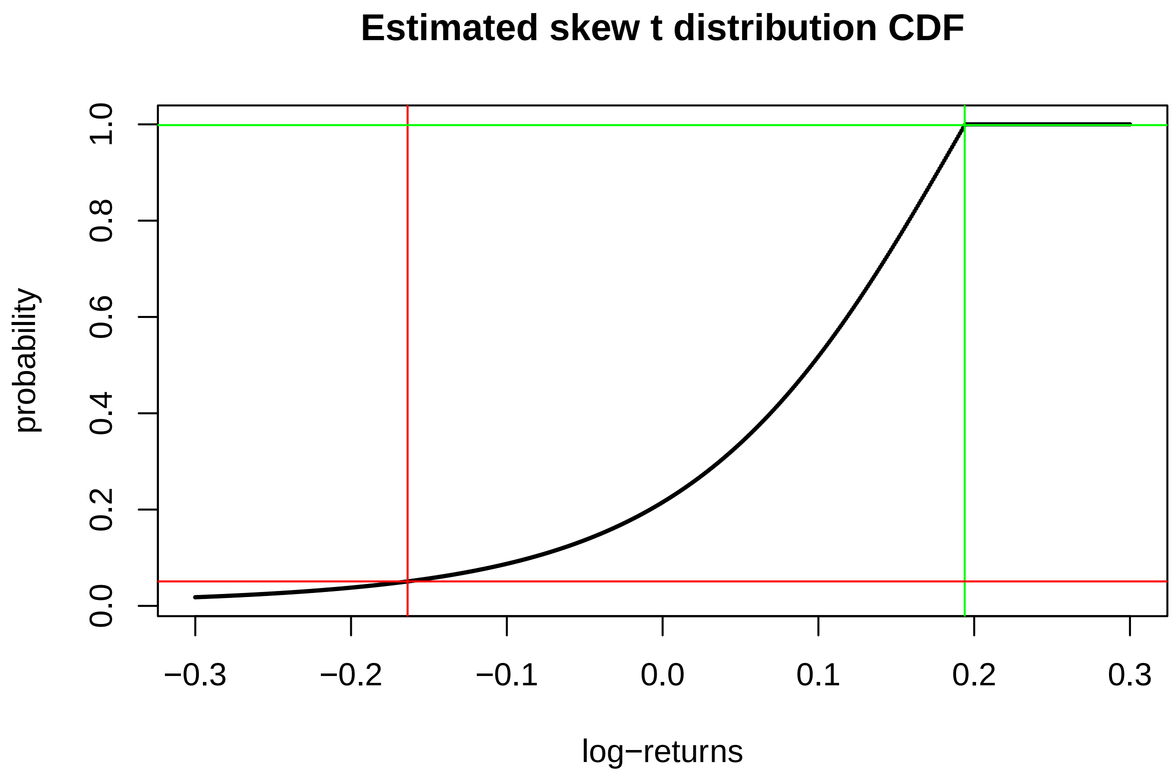
Data vs fit

Let's plot the fit and the observed returns together.

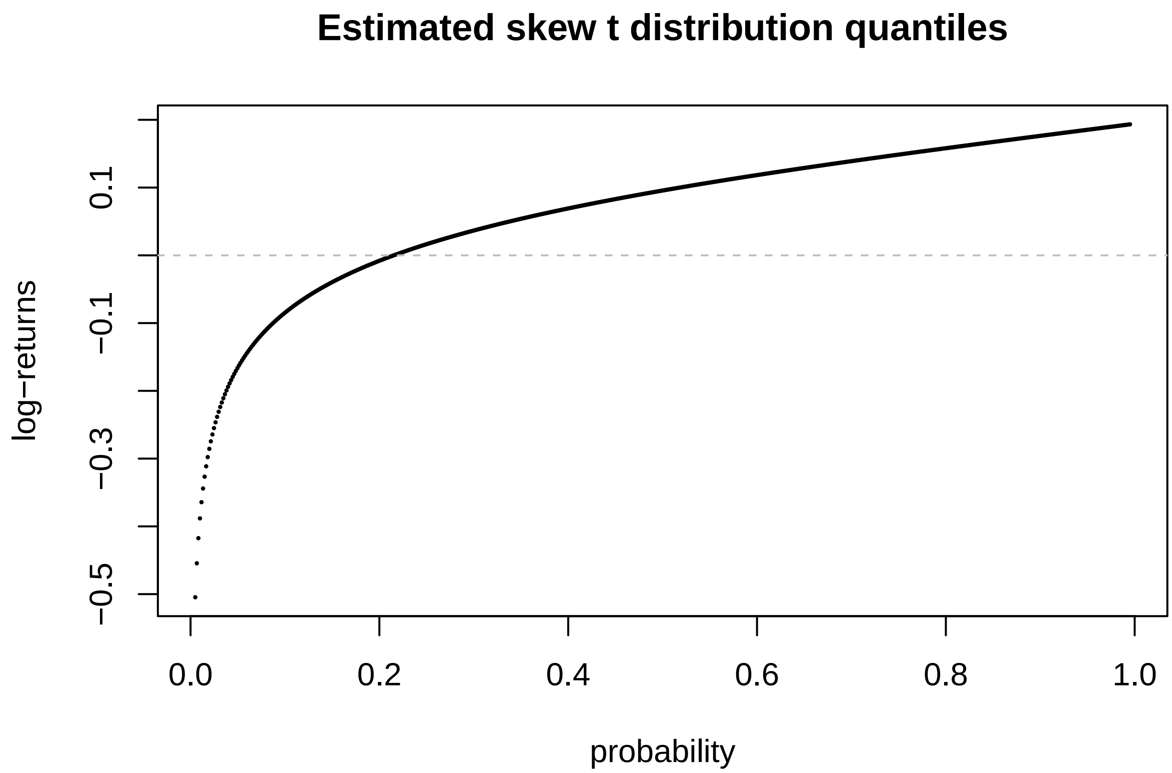


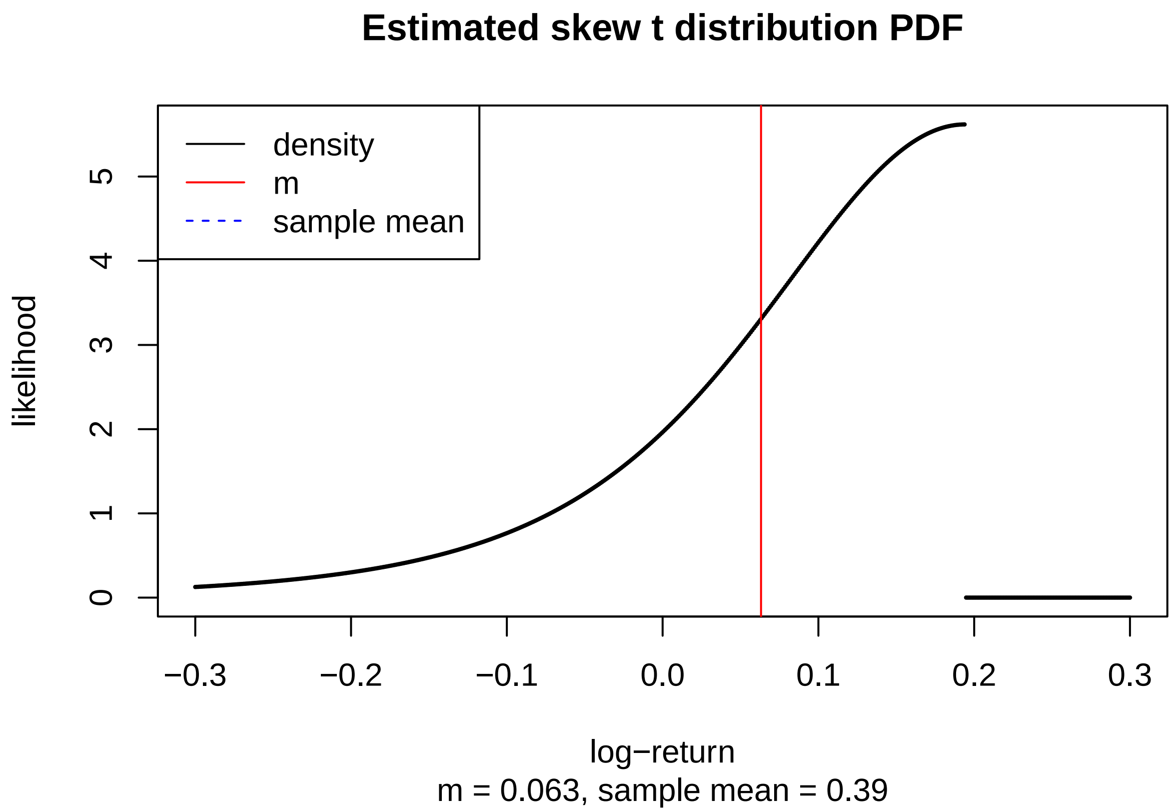
Estimated distribution

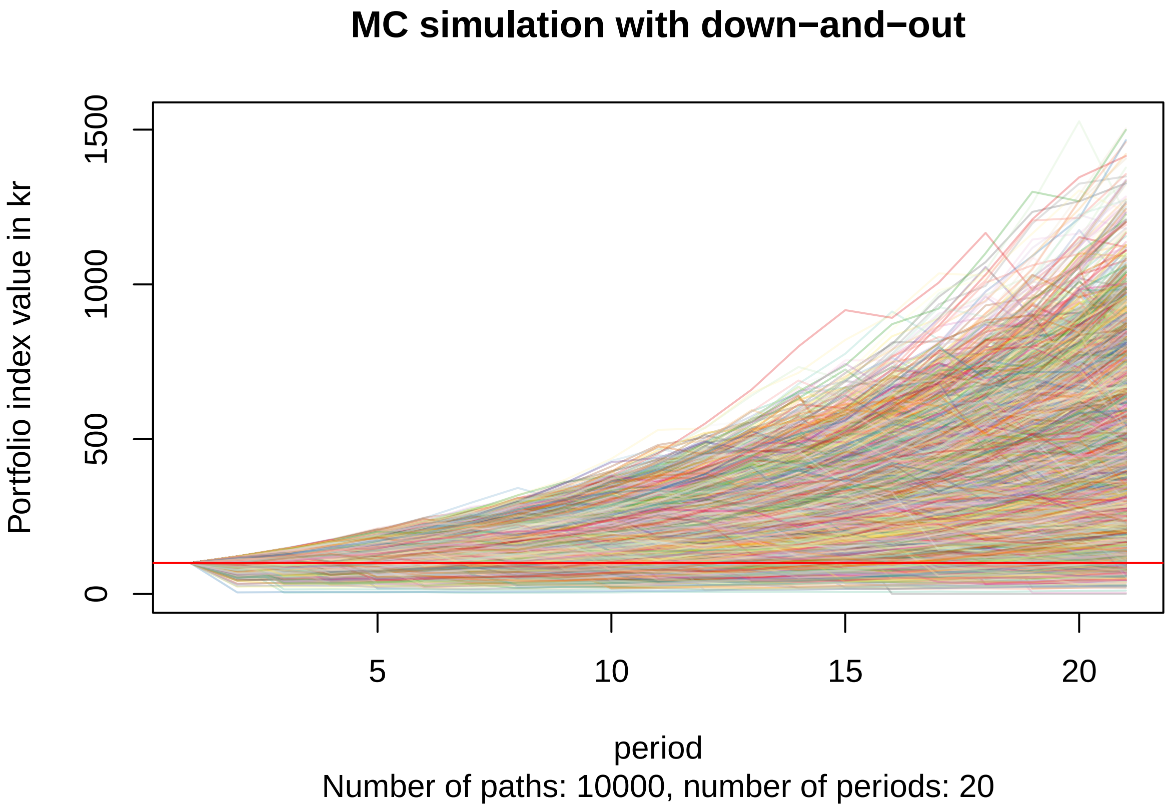
Now lets look at the CDF of the estimated distribution for each 0.1% increment between 0.5% and 99.5% for the estimated distribution:

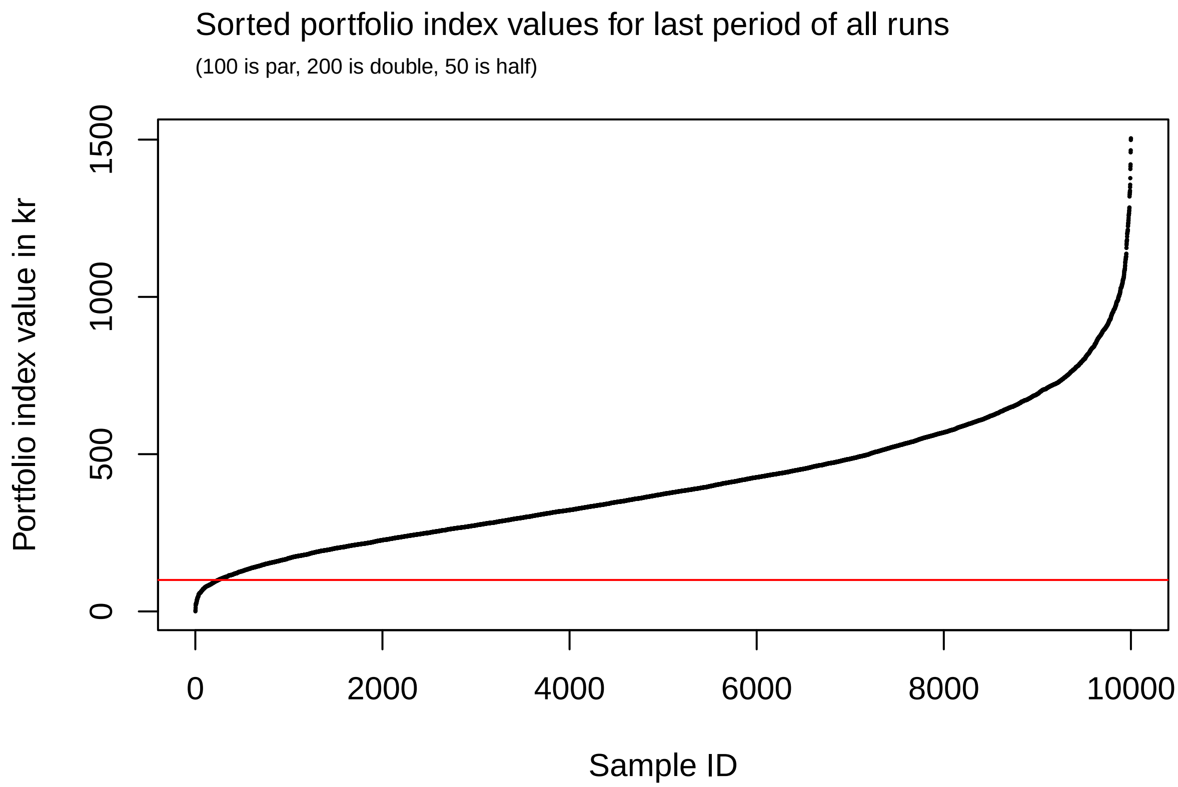


Estimated skew t distribution CDF
probability
1.0
0.8
0.6
0.4
0.2
0.0
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
log−returns


Estimated skew t distribution quantiles
log−returns
probability
0.1
−0.1
−0.3
−0.5
0.0
0.2
0.4
0.6
0.8
1.0


Estimated skew t distribution PDF
density
m
sample mean
likelihood
0
1
2
3
4
5
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
log−return
m = 0.063, sample mean = 0.39

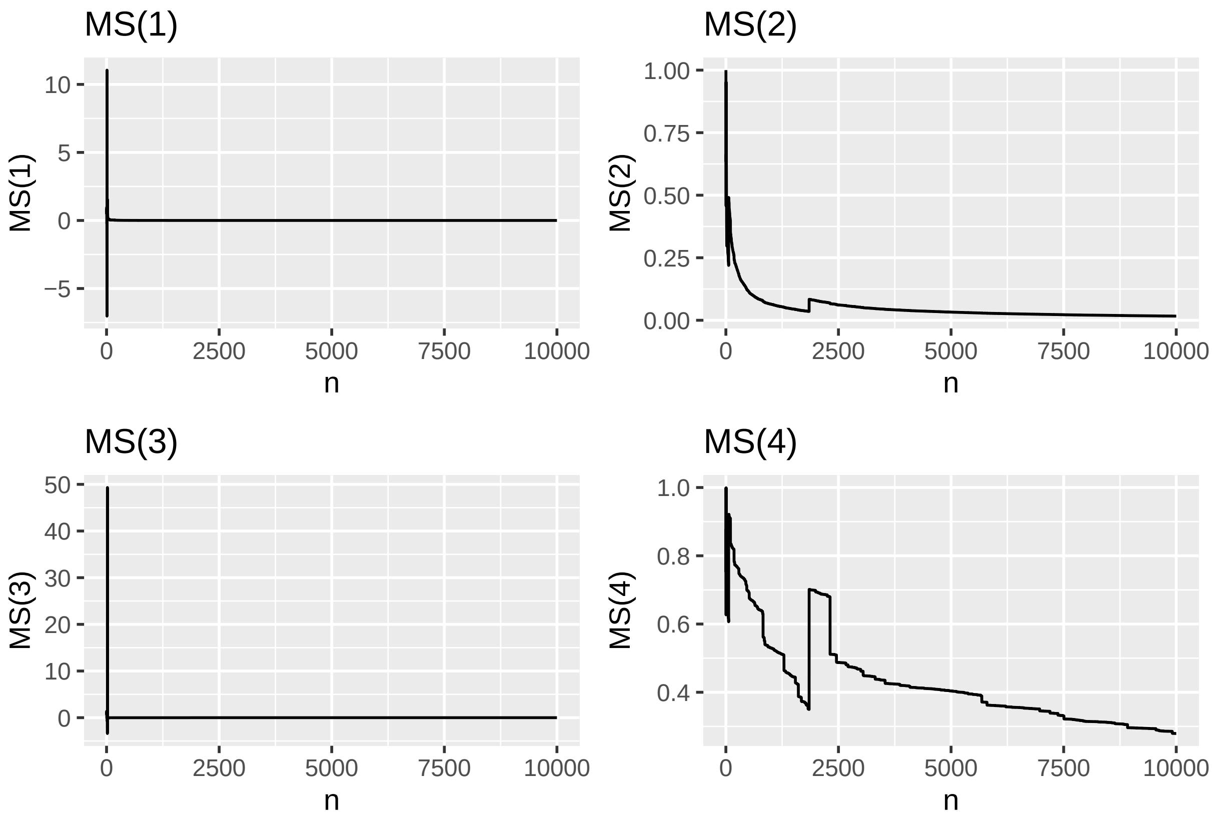

MC simulation with down−and−out
Portfolio index value in kr
0
500
1000
1500
5
10
15
20
period
Number of paths: 10000, number of periods: 20



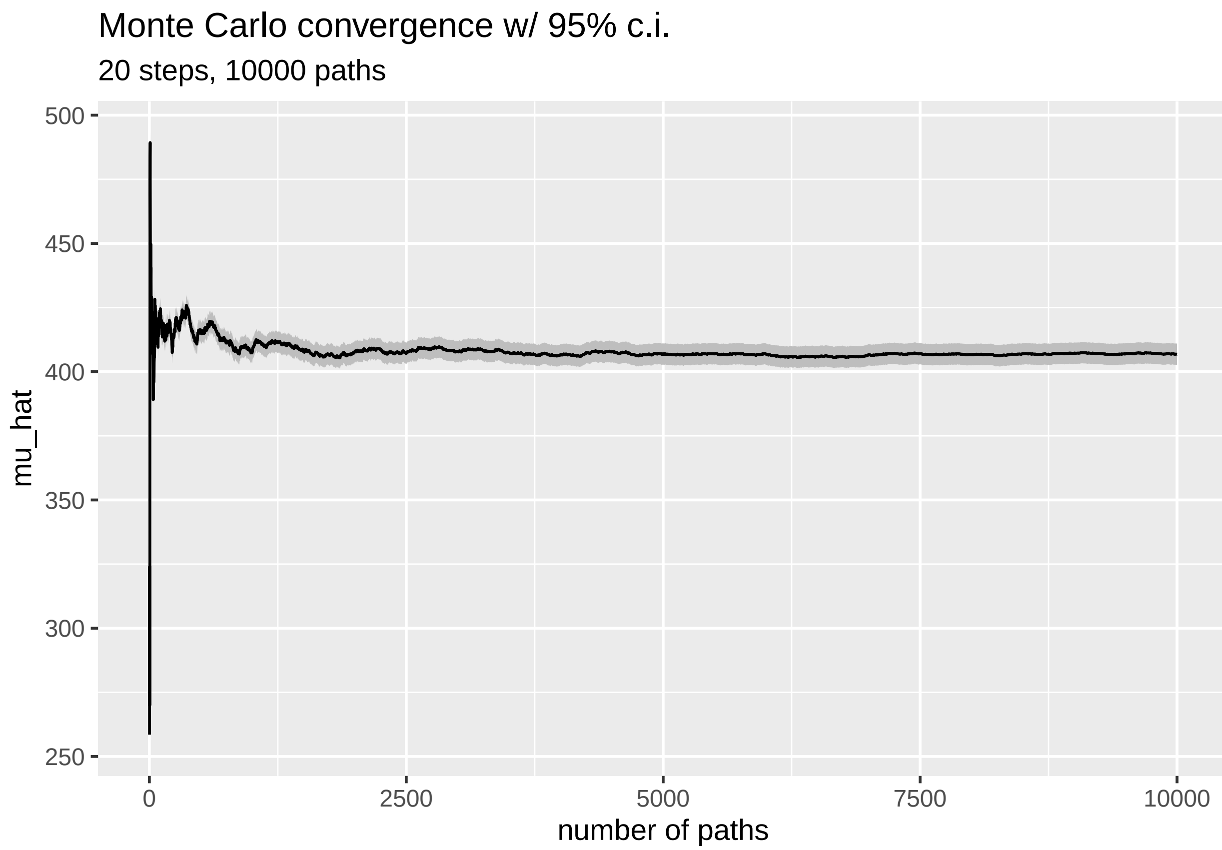
Convergence

Max vs sum

Max vs sum plots for the first four moments:

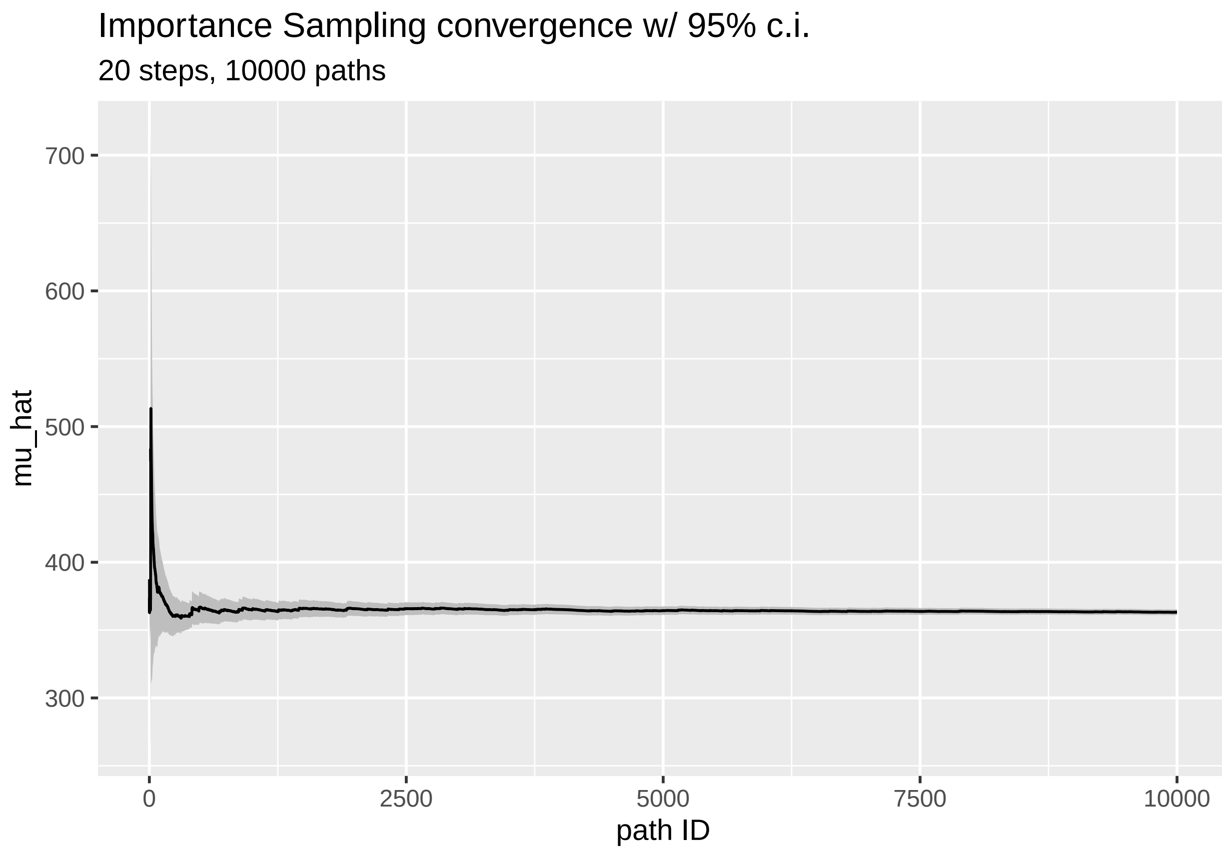

MS(1)
MS(2)
MS(3)
MS(4)
n
0
2500
5000
7500
10000
10
5
0
−5
1.00
0.75
0.50
0.25
0.00
50
40
30
20
10
1.0
0.8
0.6
0.4

MC



Monte Carlo convergence w/ 95% c.i.
20 steps, 10000 paths
mu_hat
500
450
400
350
300
250
0
2500
5000
7500
10000
number of paths

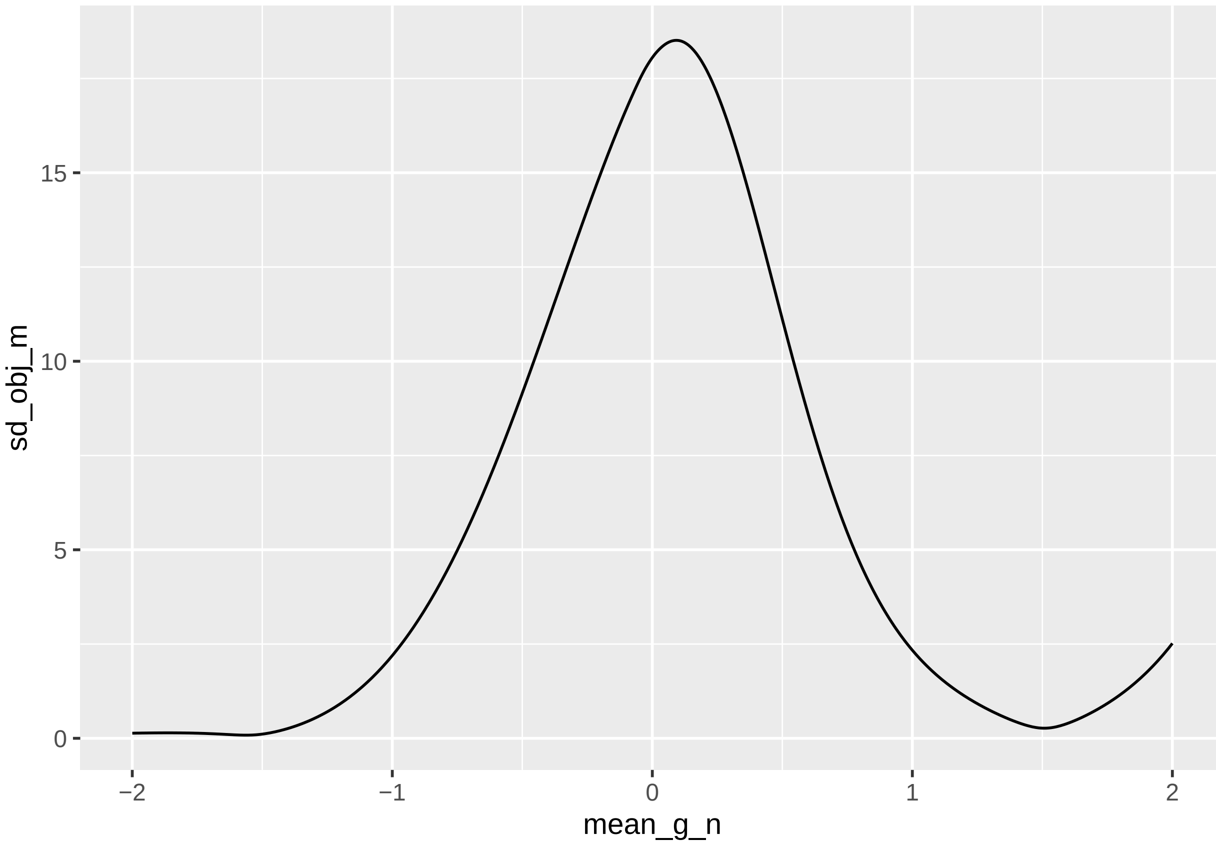
IS

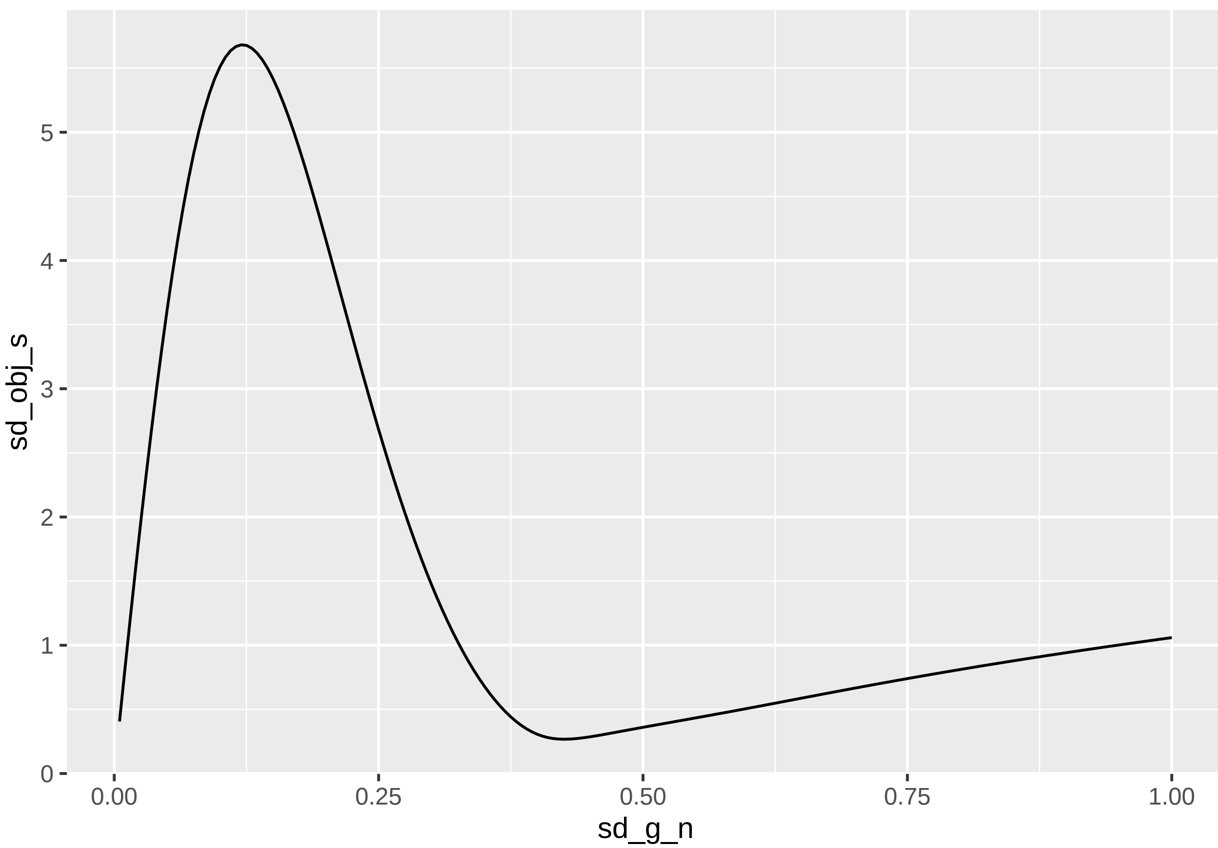


Parameters

```
## [1] 1.5074609 0.4255322
```

Objective function plots

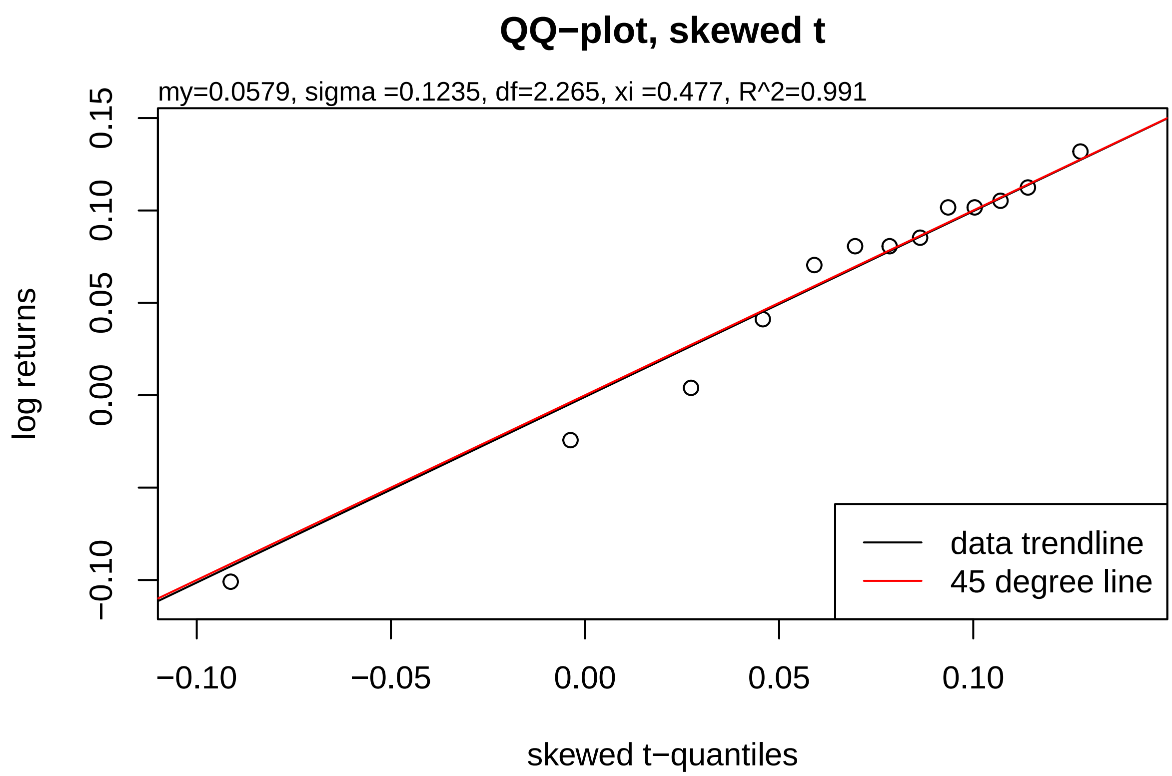

sd_obj_m
15
10
5
0
−2
−1
0
1
2
mean_g_n



5
4
3
2
1
0
sd_obj_s
0.00
0.25
0.50
0.75
1.00
sd_g_n

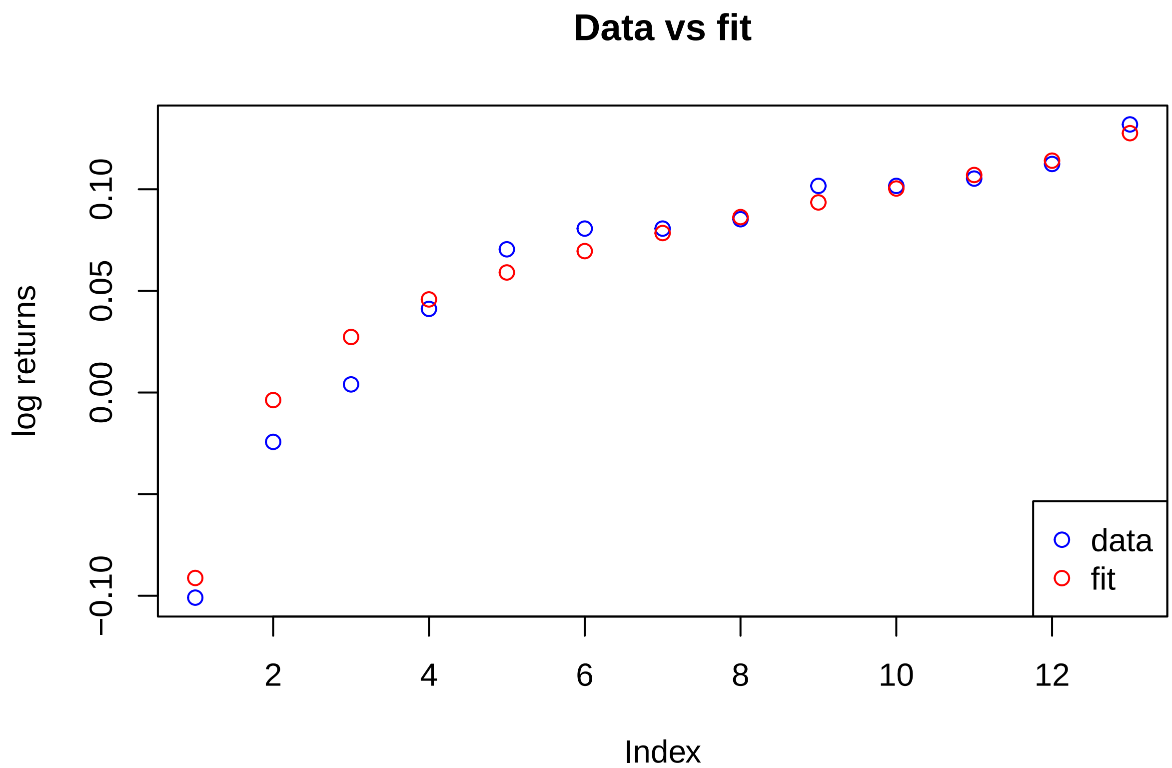
PFA medium risk (pmr), 2011 - 2023

QQ Plot



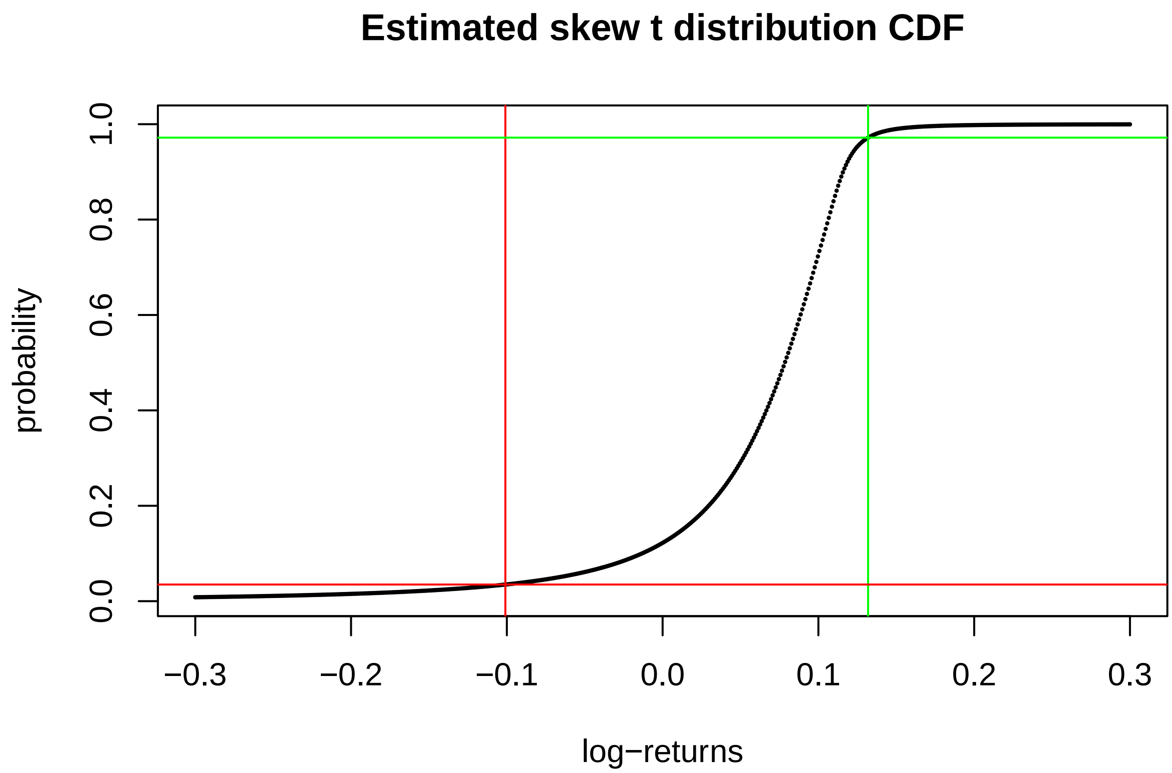
Data vs fit

Let's plot the fit and the observed returns together.

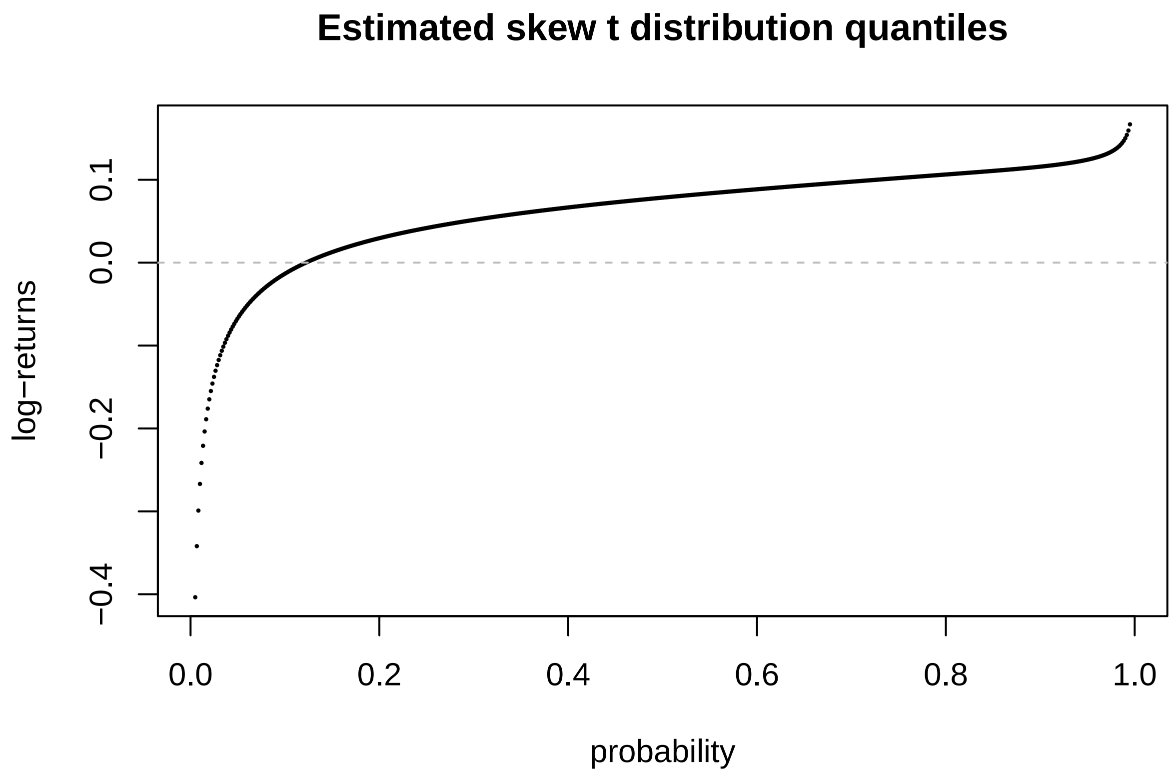


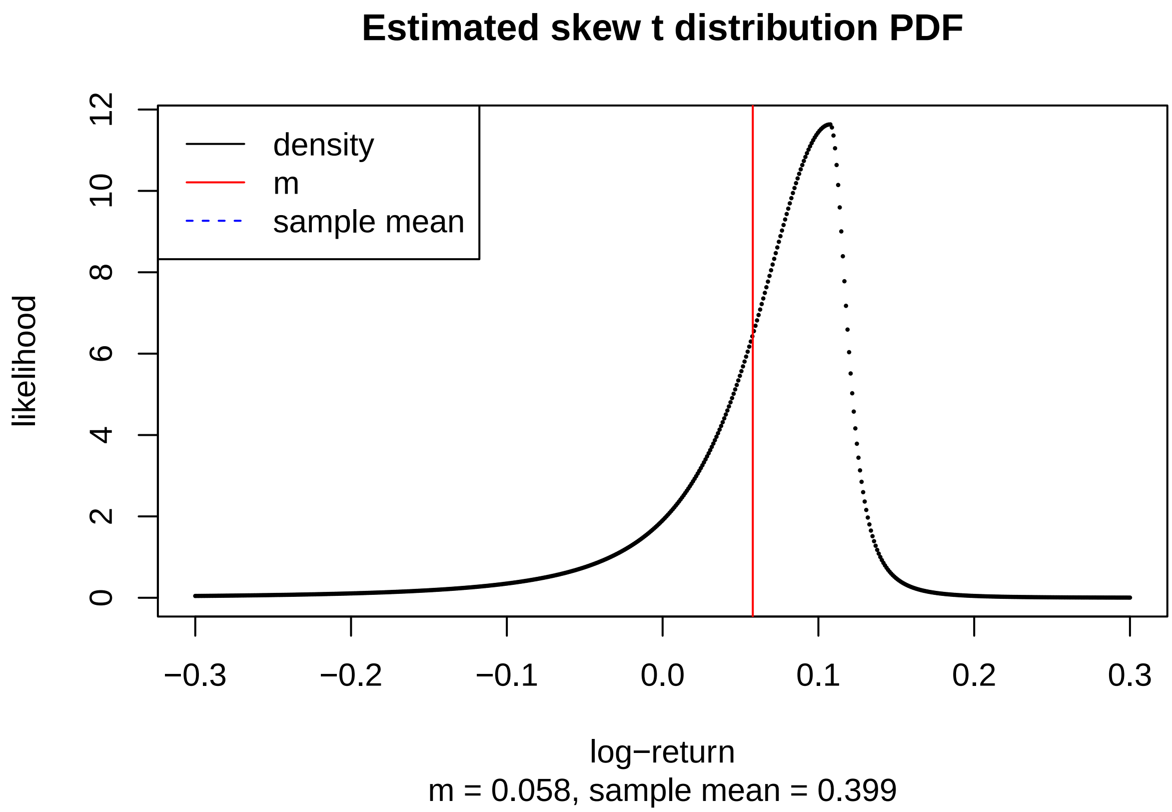
Estimated distribution

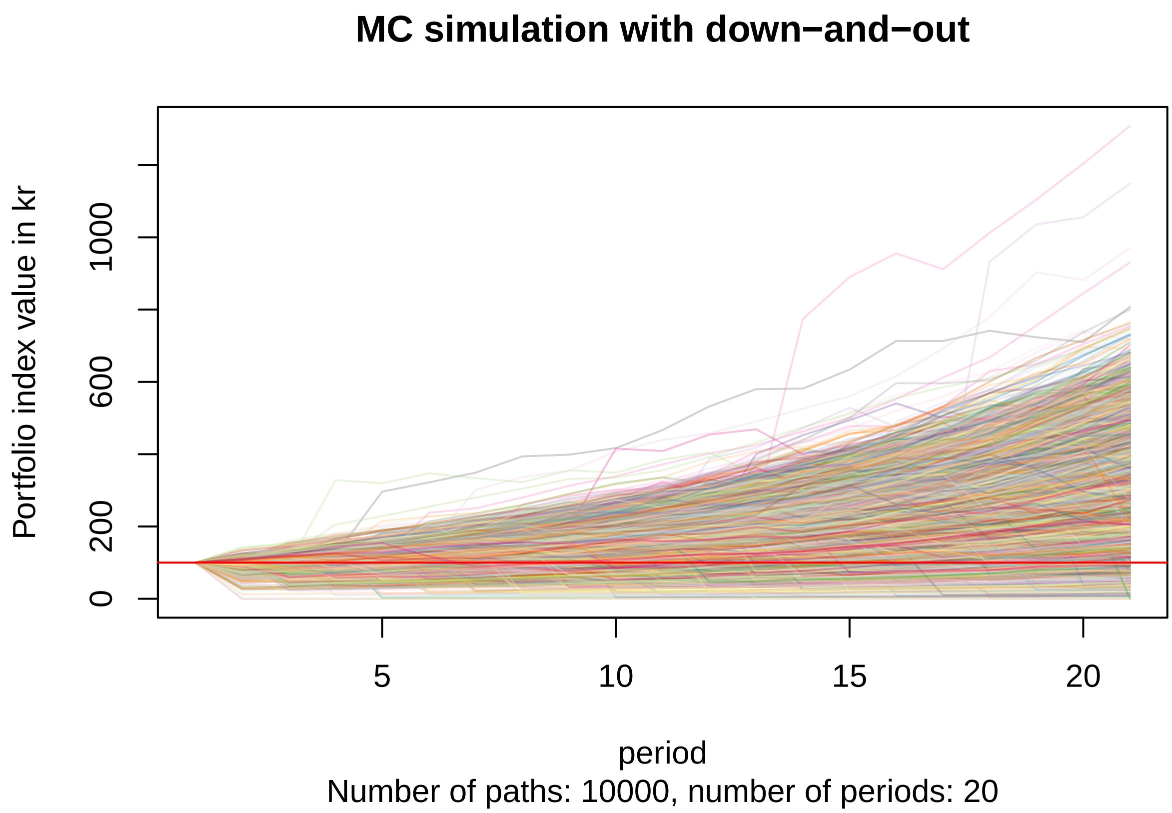
Now lets look at the CDF of the estimated distribution for each 0.1% increment between 0.5% and 99.5% for the estimated distribution:

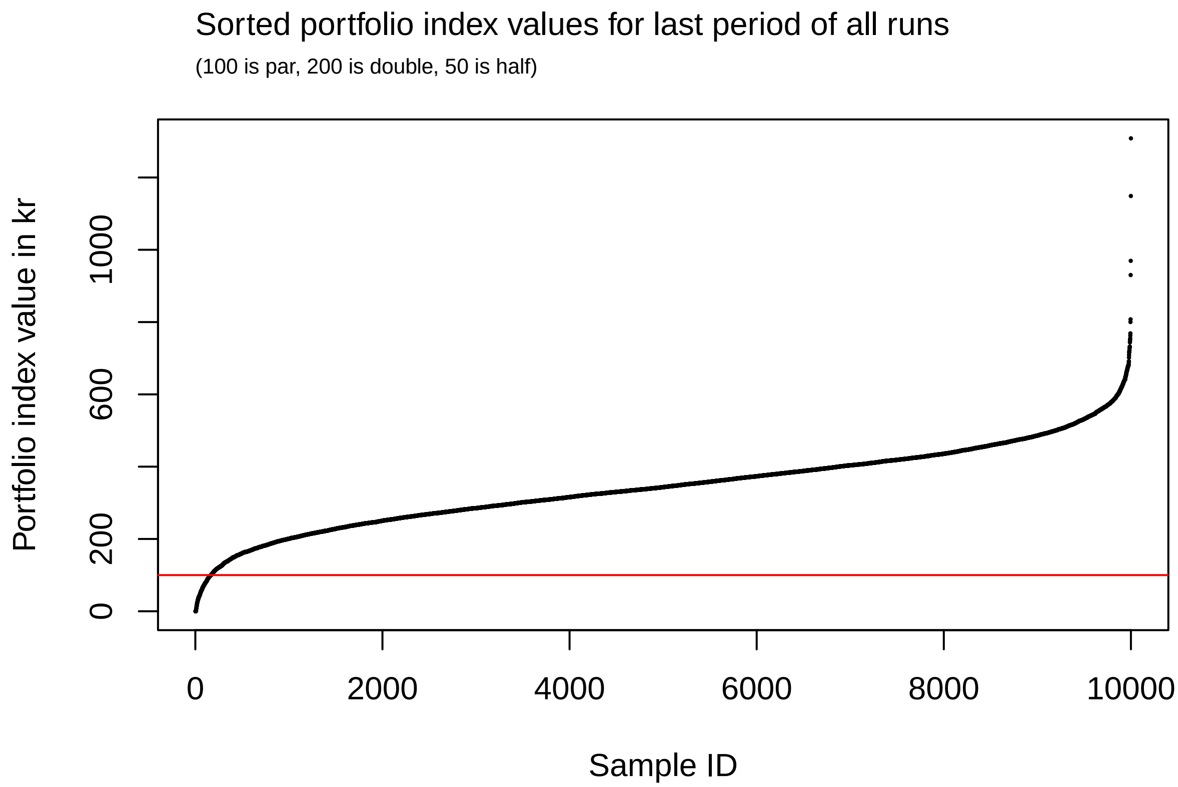


Estimated skew t distribution CDF
probability
log−returns
0.0
0.2
0.4
0.6
0.8
1.0
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3


Estimated skew t distribution quantiles
log−returns
0.1
0.0
−0.2
−0.4
0.0
0.2
0.4
0.6
0.8
1.0
probability


Estimated skew t distribution PDF
density
m
sample mean
likelihood
0
2
4
6
8
10
12
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
log−return
m = 0.058, sample mean = 0.399

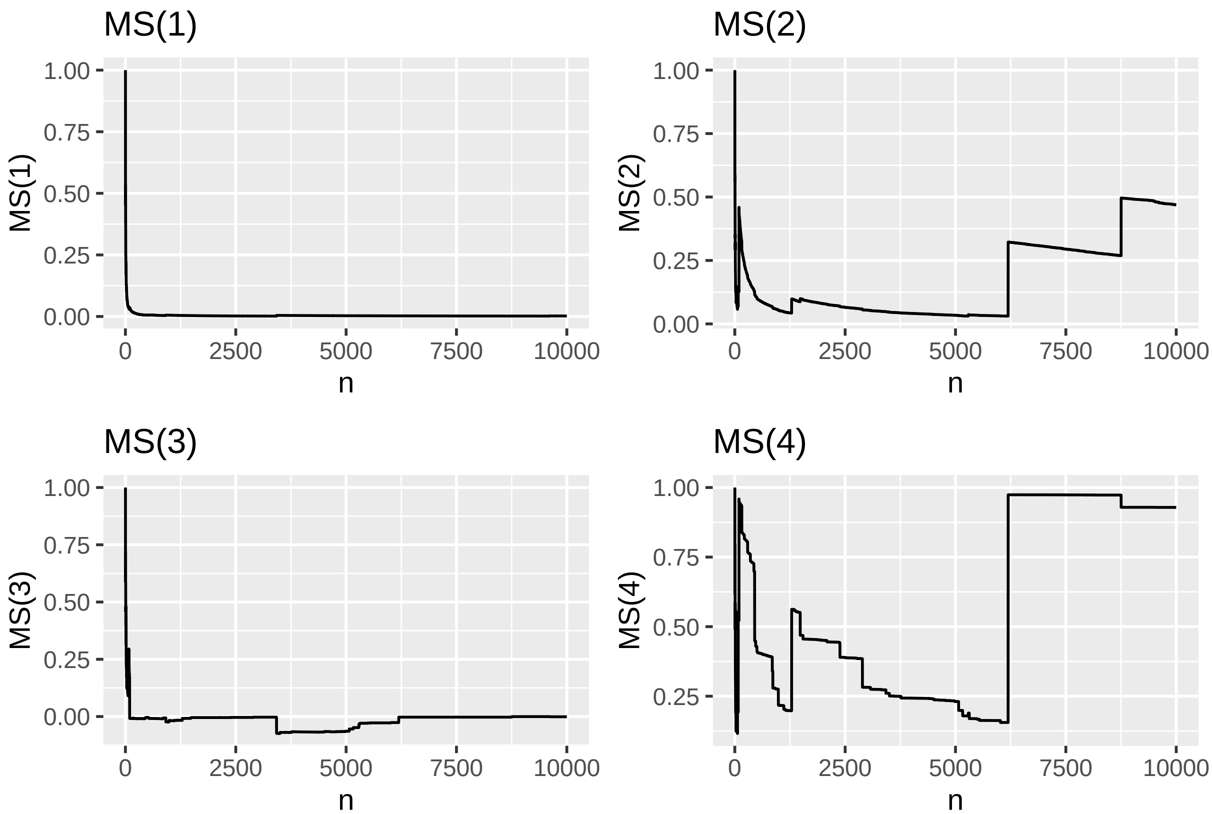

MC simulation with down−and−out
Portfolio index value in kr
0
200
600
1000
5
10
15
20
period
Number of paths: 10000, number of periods: 20

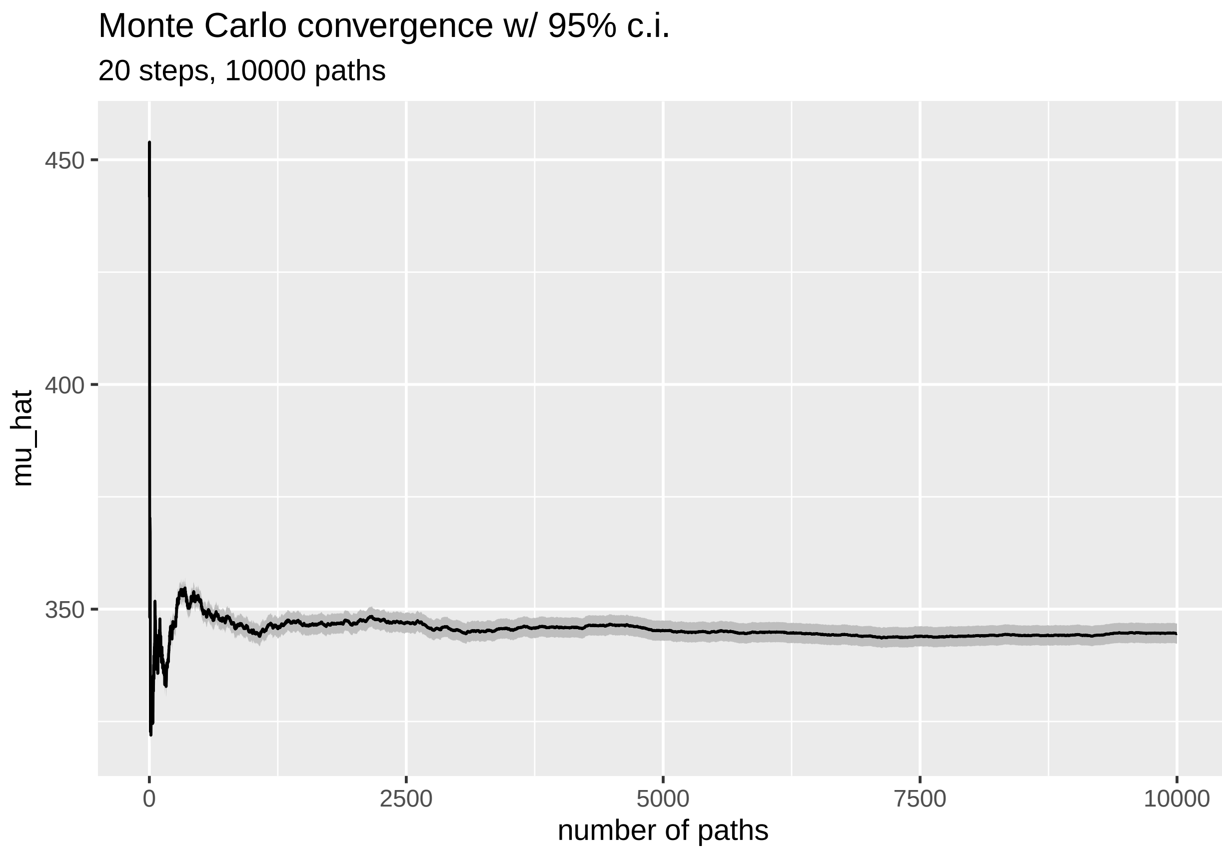


Convergence

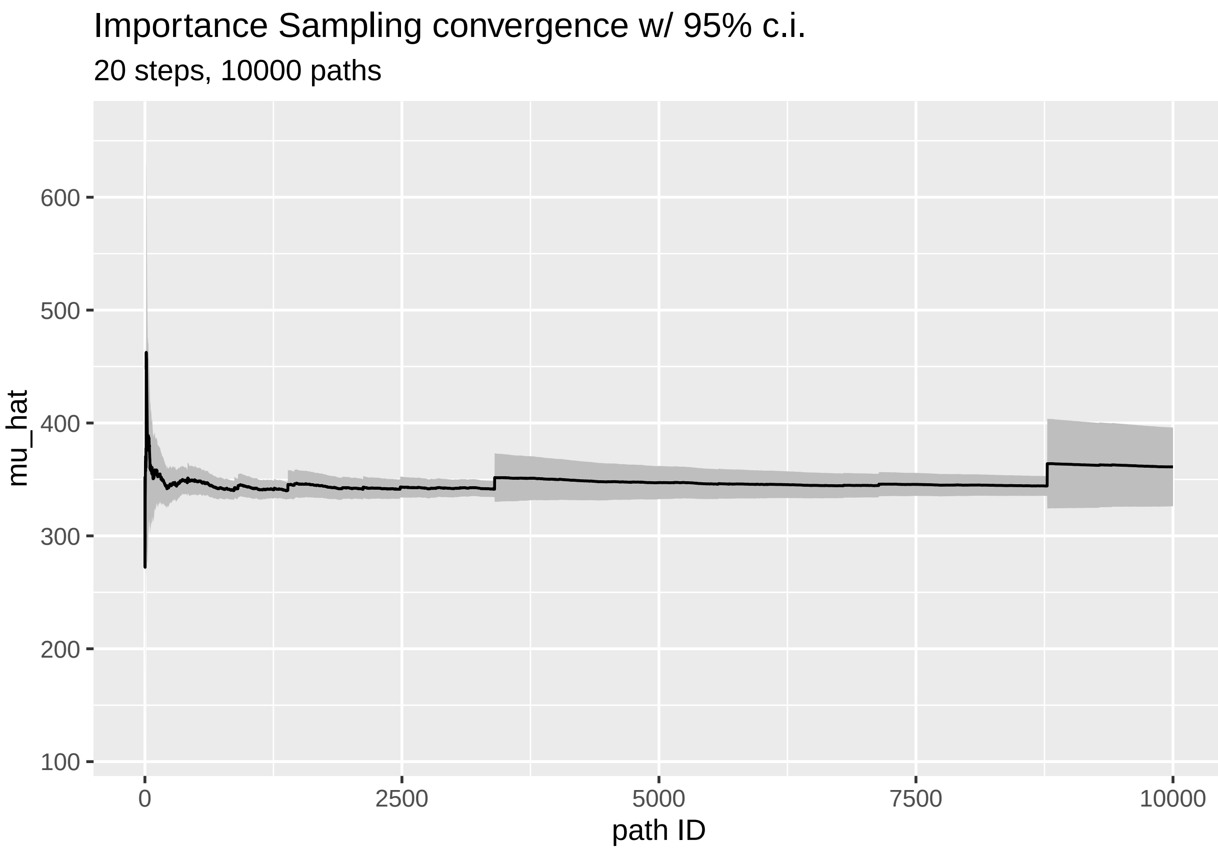
Max vs sum

Max vs sum plots for the first four moments:


MS(1)
MS(2)
MS(3)
MS(4)
1.00
0.75
0.50
0.25
0.00
0
2500
5000
7500
10000
n


Monte Carlo convergence w/ 95% c.i.
20 steps, 10000 paths
450
400
350
mu_hat
0
2500
5000
7500
10000
number of paths

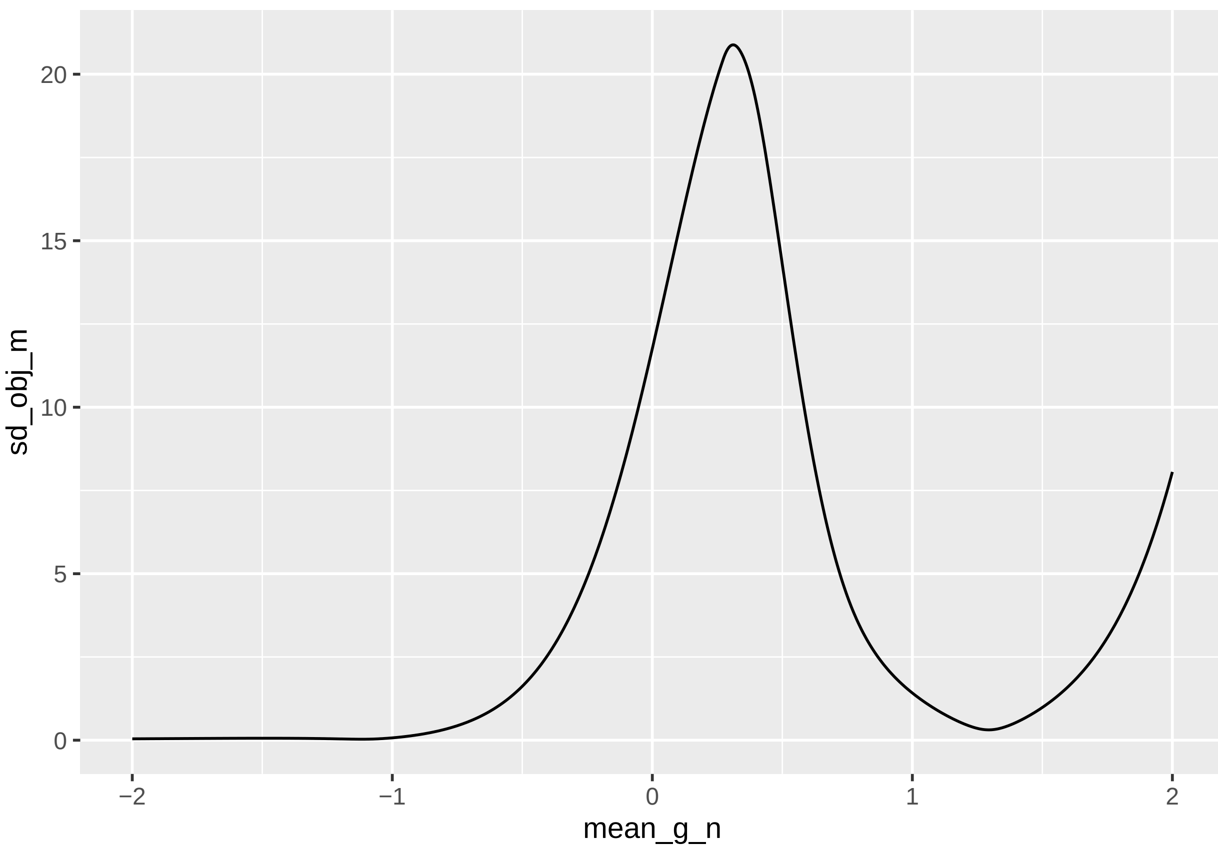
IS



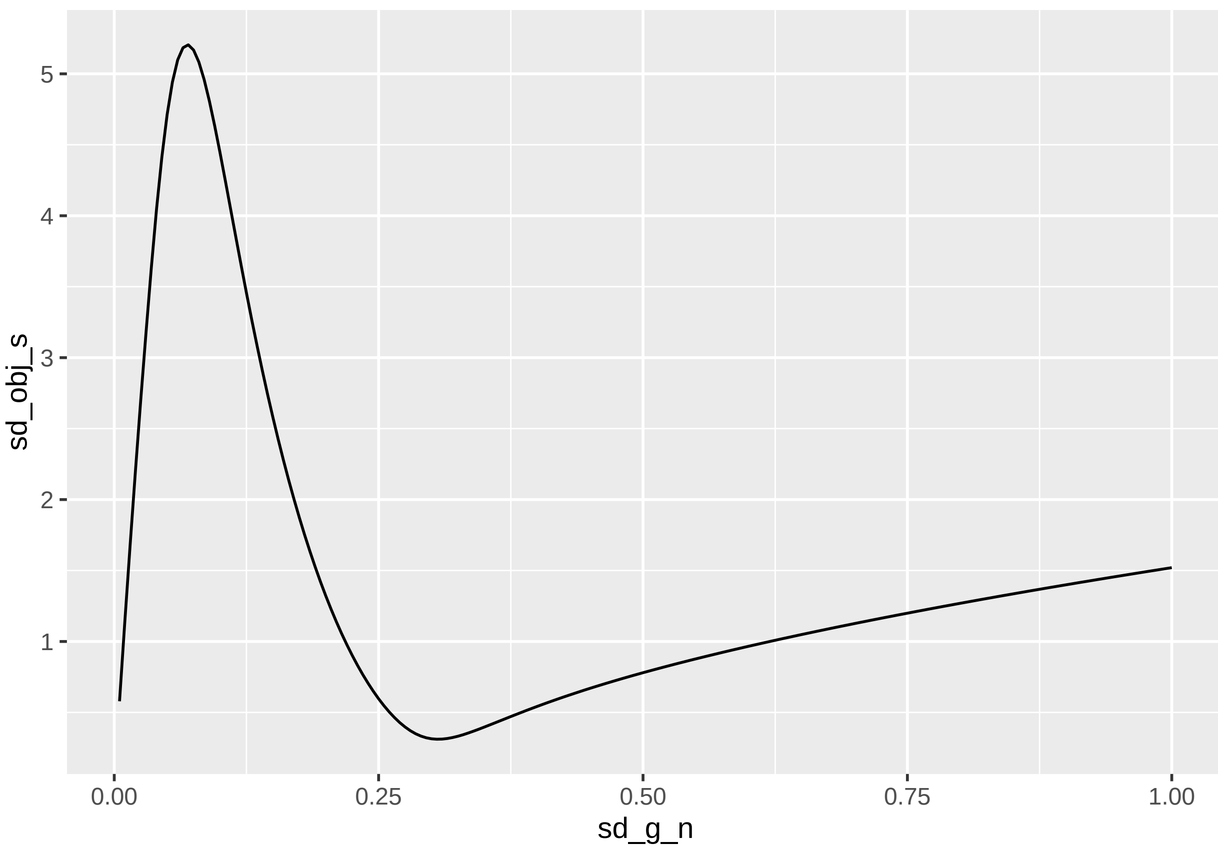
Parameters

```
## [1] 1.2936284 0.3062685
```

Objective function plots

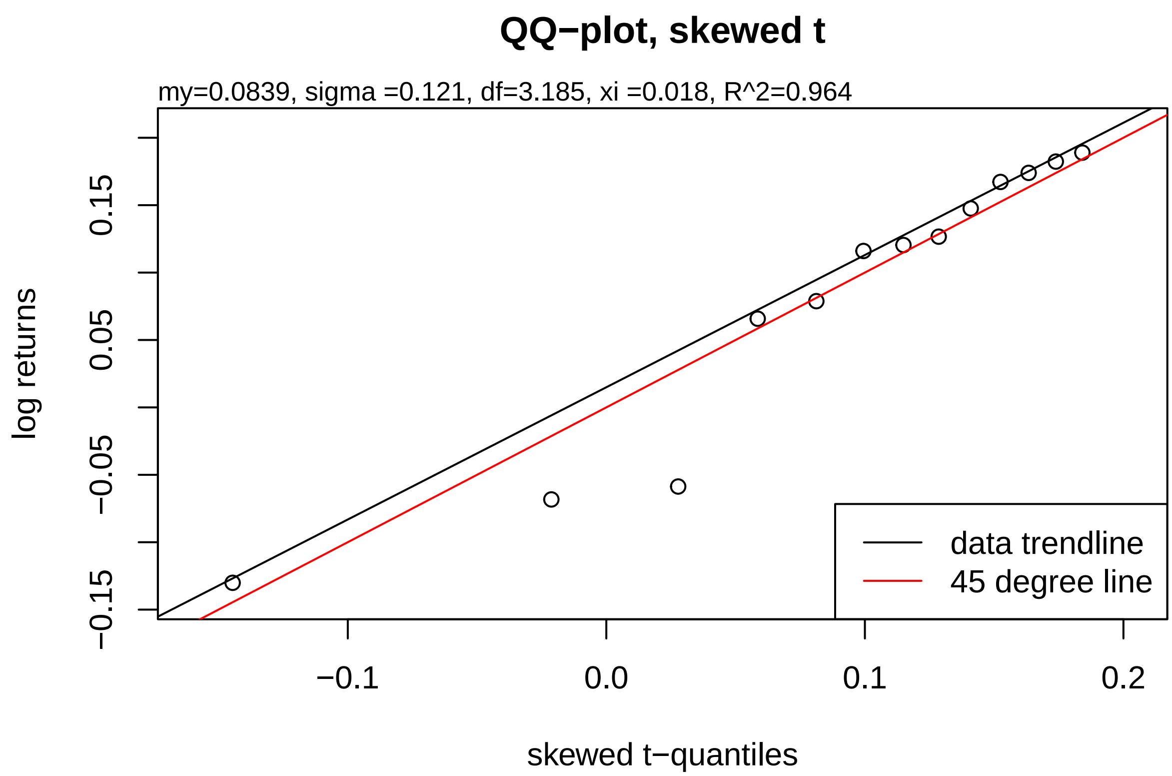


sd_obj_m
20
15
10
5
0
−2
−1
0
1
2
mean_g_n


5
4
3
2
1
sd_obj_s
0.00
0.25
0.50
0.75
1.00
sd_g_n

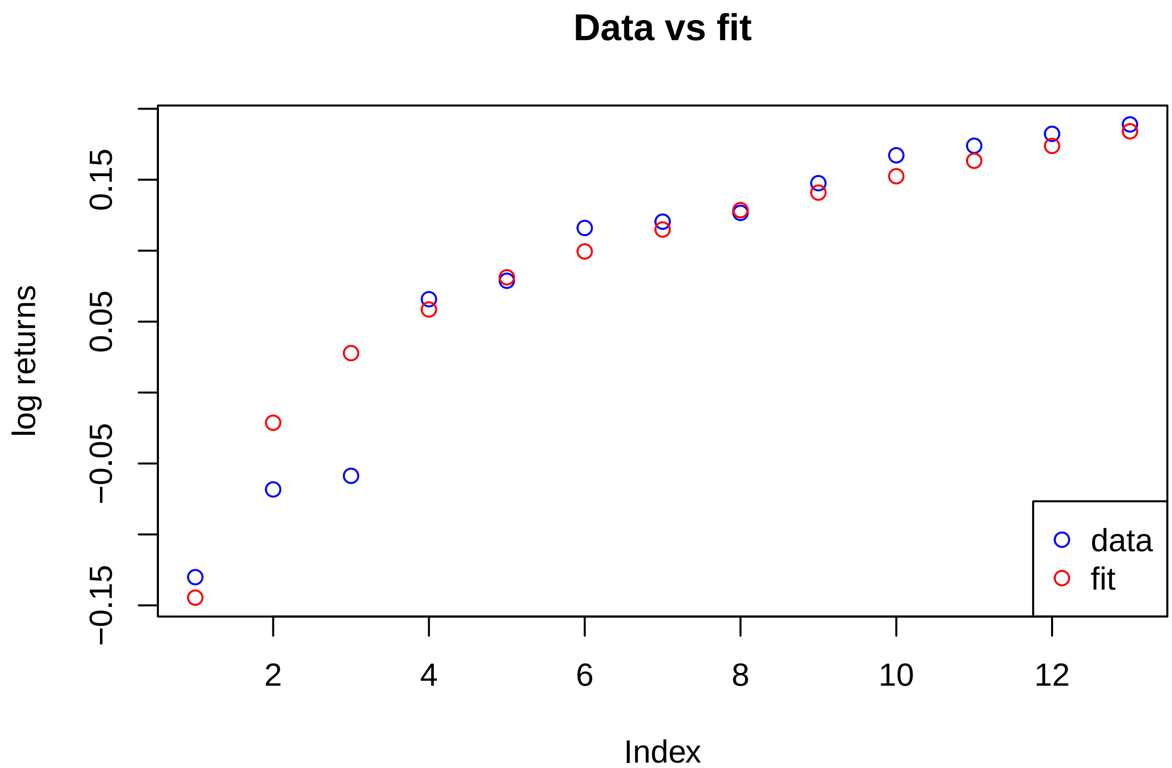
PFA high risk (phr), 2011 - 2023

QQ Plot



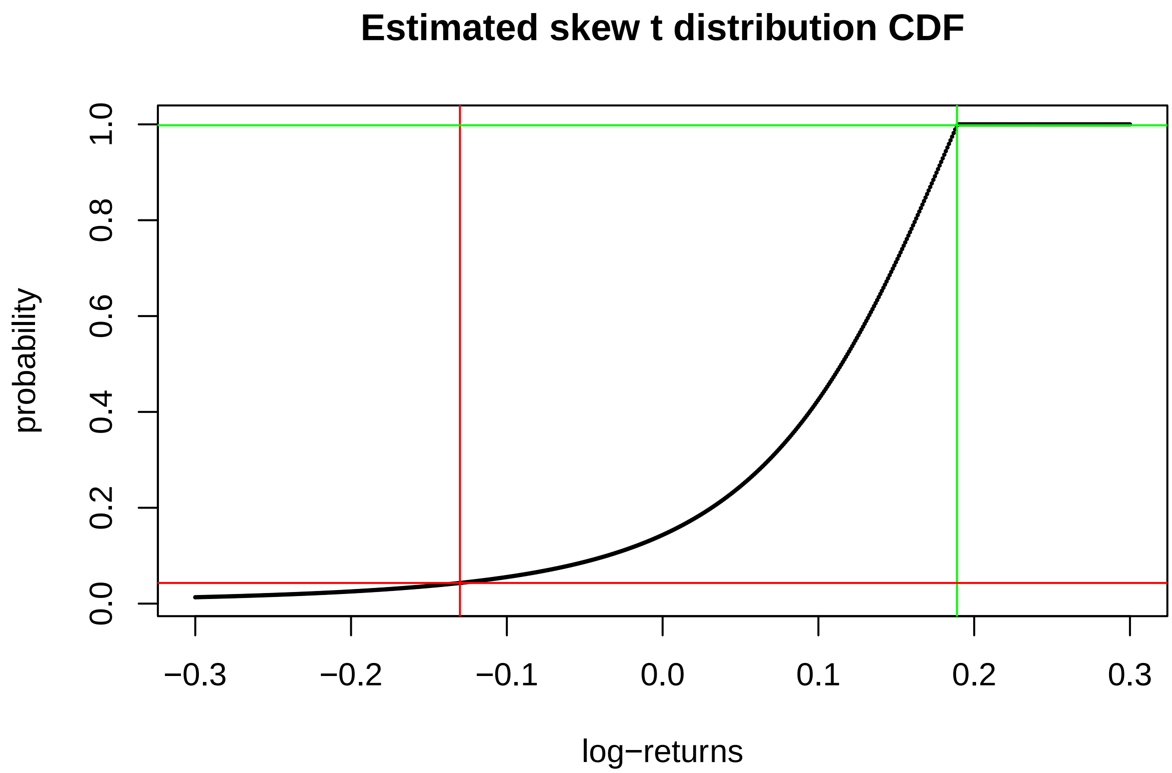
Data vs fit

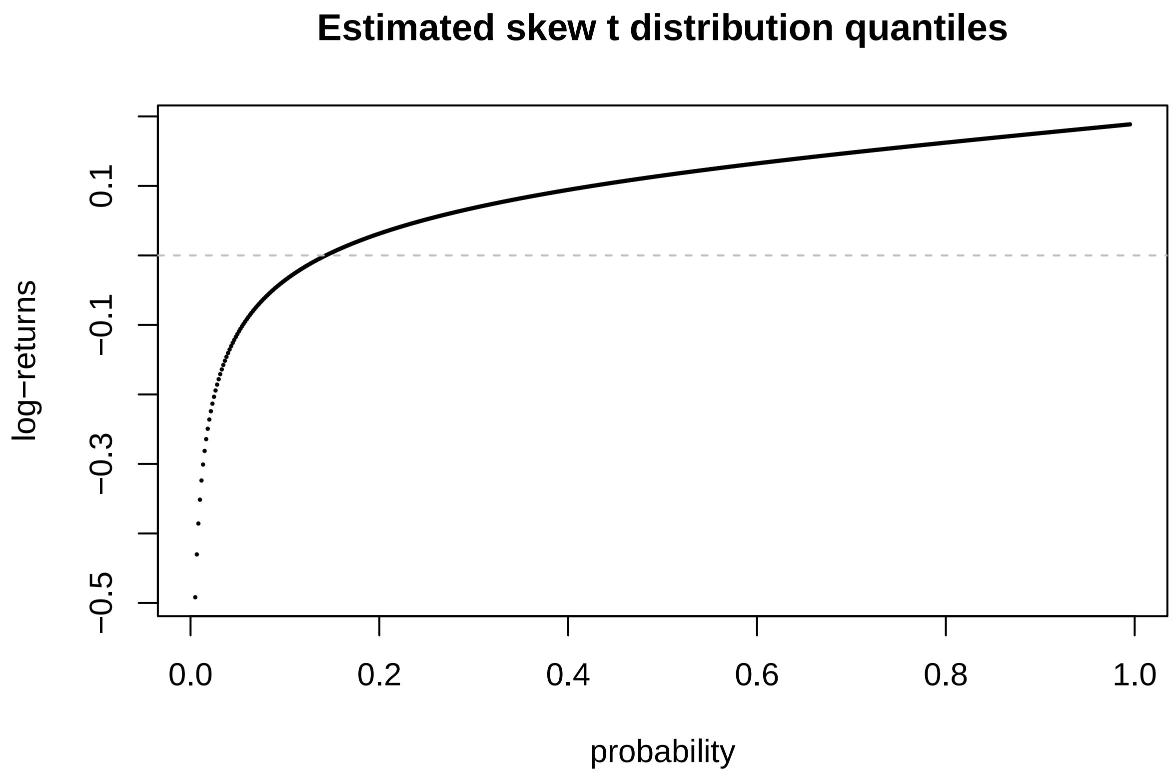
Let’s plot the fit and the observed returns together.



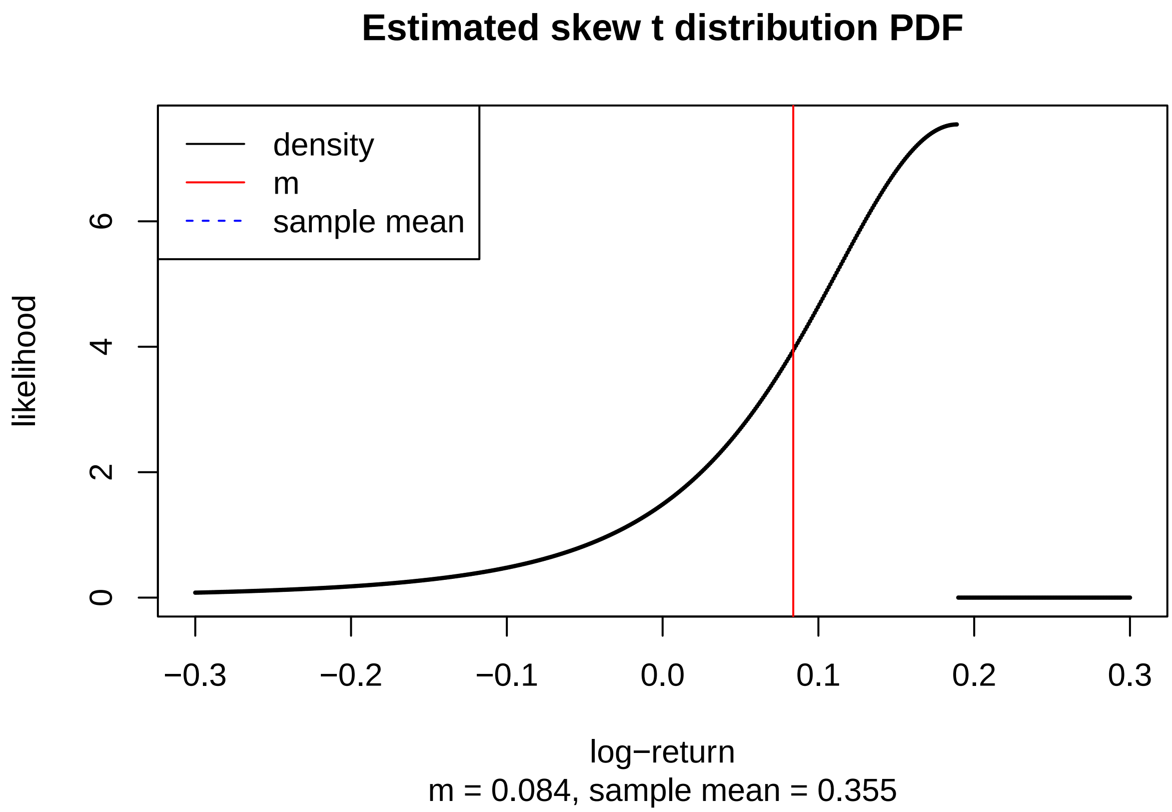
Estimated distribution

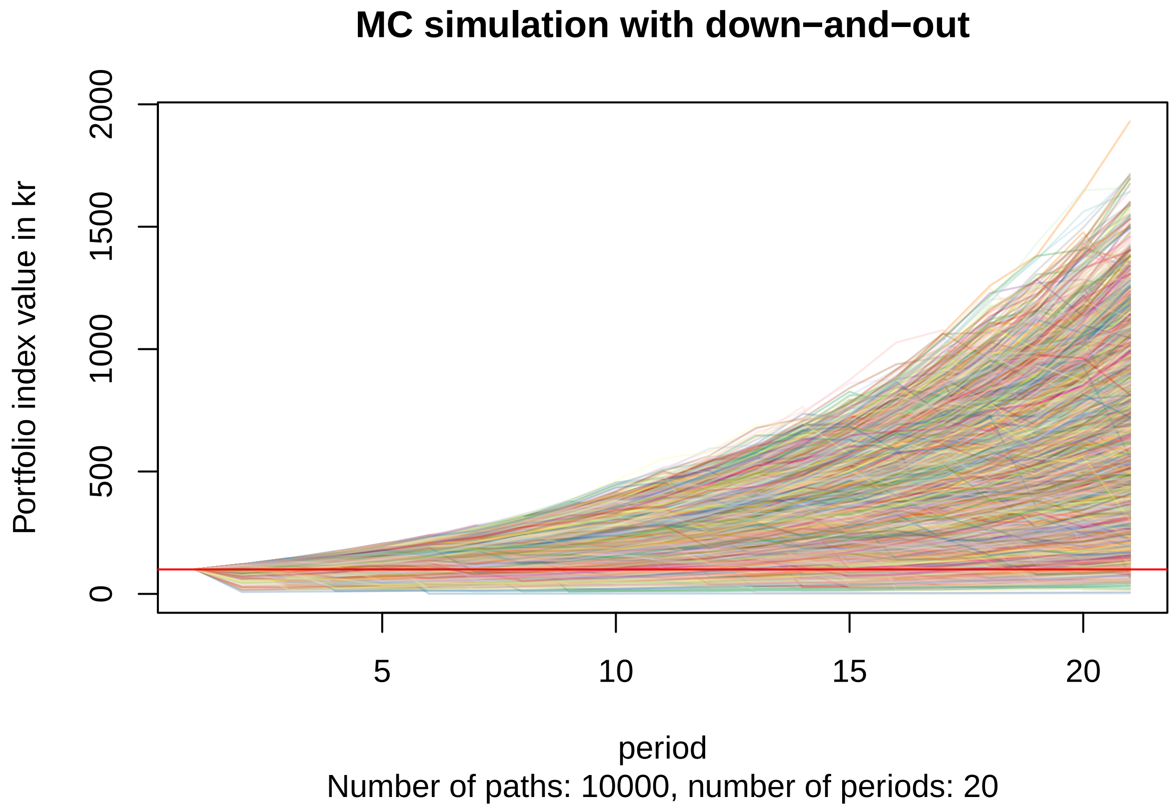
Now lets look at the CDF of the estimated distribution for each 0.1% increment between 0.5% and 99.5% for the estimated distribution:

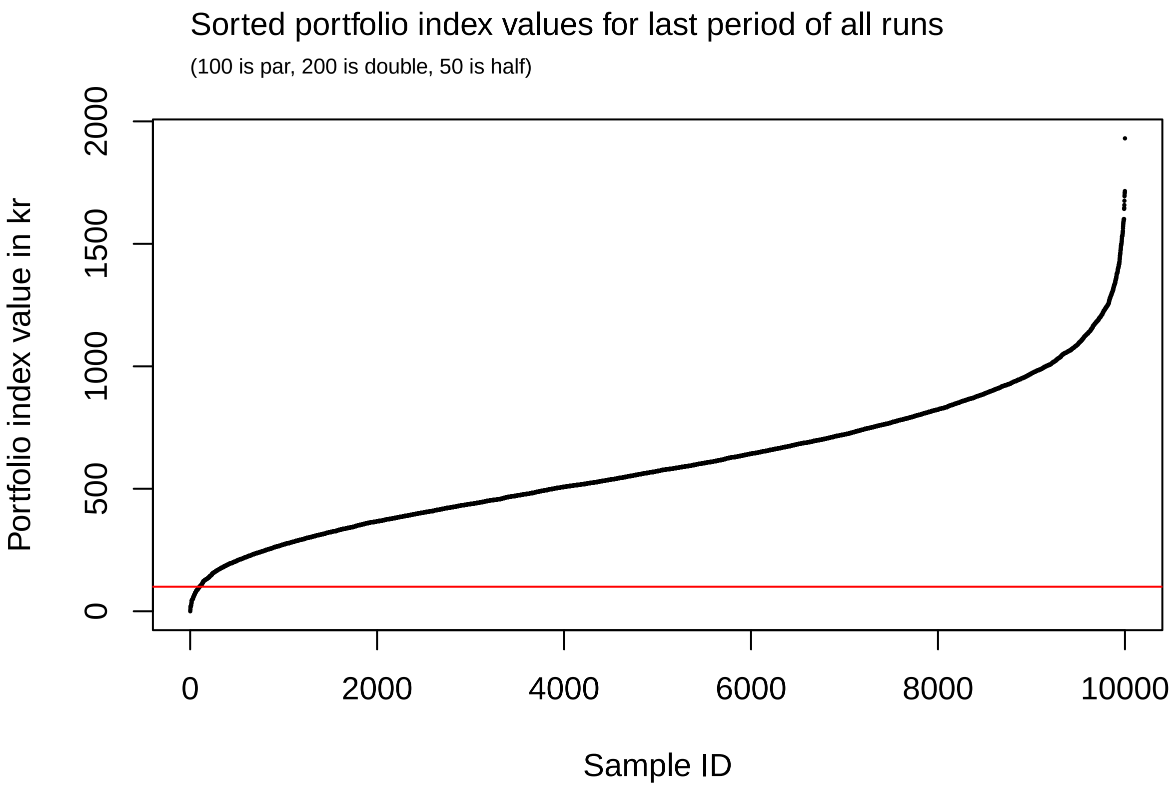

Estimated skew t distribution CDF
probability
0.0
0.2
0.4
0.6
0.8
1.0
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
log−returns



Estimated skew t distribution quantiles
0.1
−0.1
−0.3
−0.5
log−returns
0.0
0.2
0.4
0.6
0.8
1.0
probability


Estimated skew t distribution PDF
density
m
sample mean
likelihood
0
2
4
6
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
log−return
m = 0.084, sample mean = 0.355

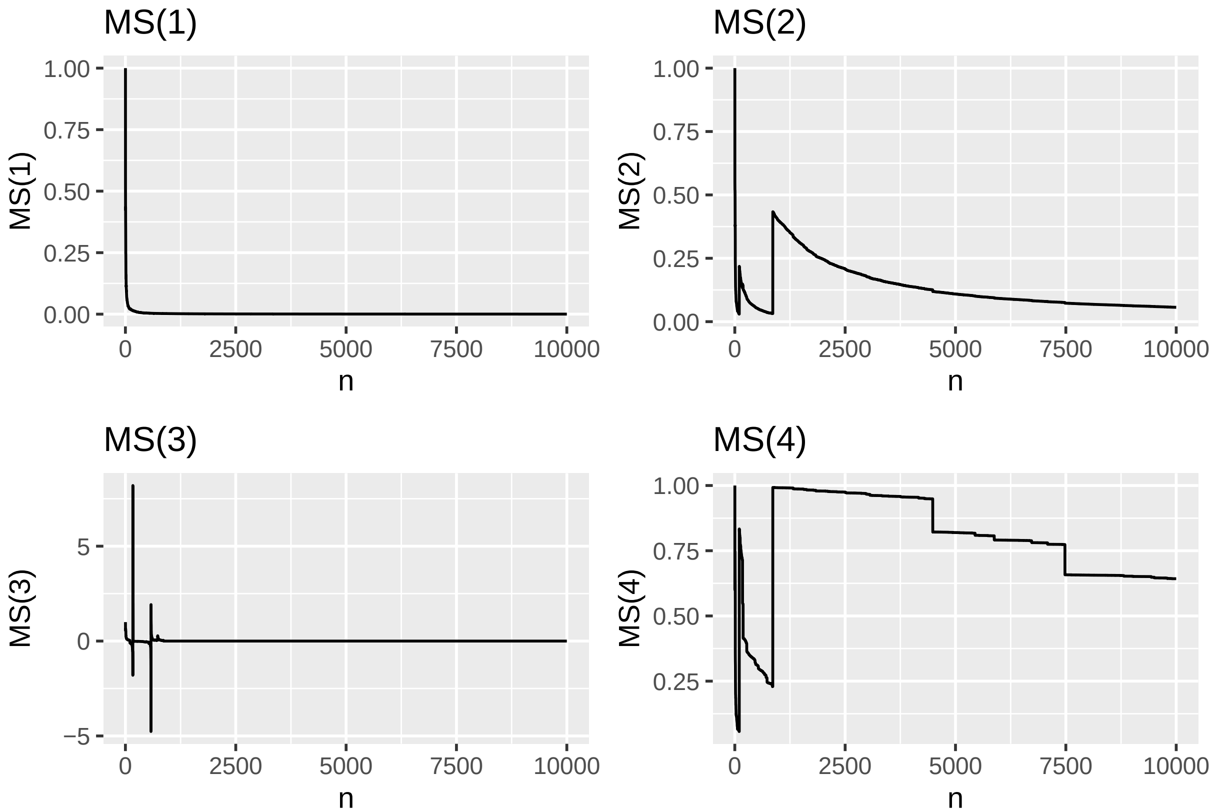

MC simulation with down−and−out
Portfolio index value in kr
0
500
1000
1500
2000
5
10
15
20
period
Number of paths: 10000, number of periods: 20

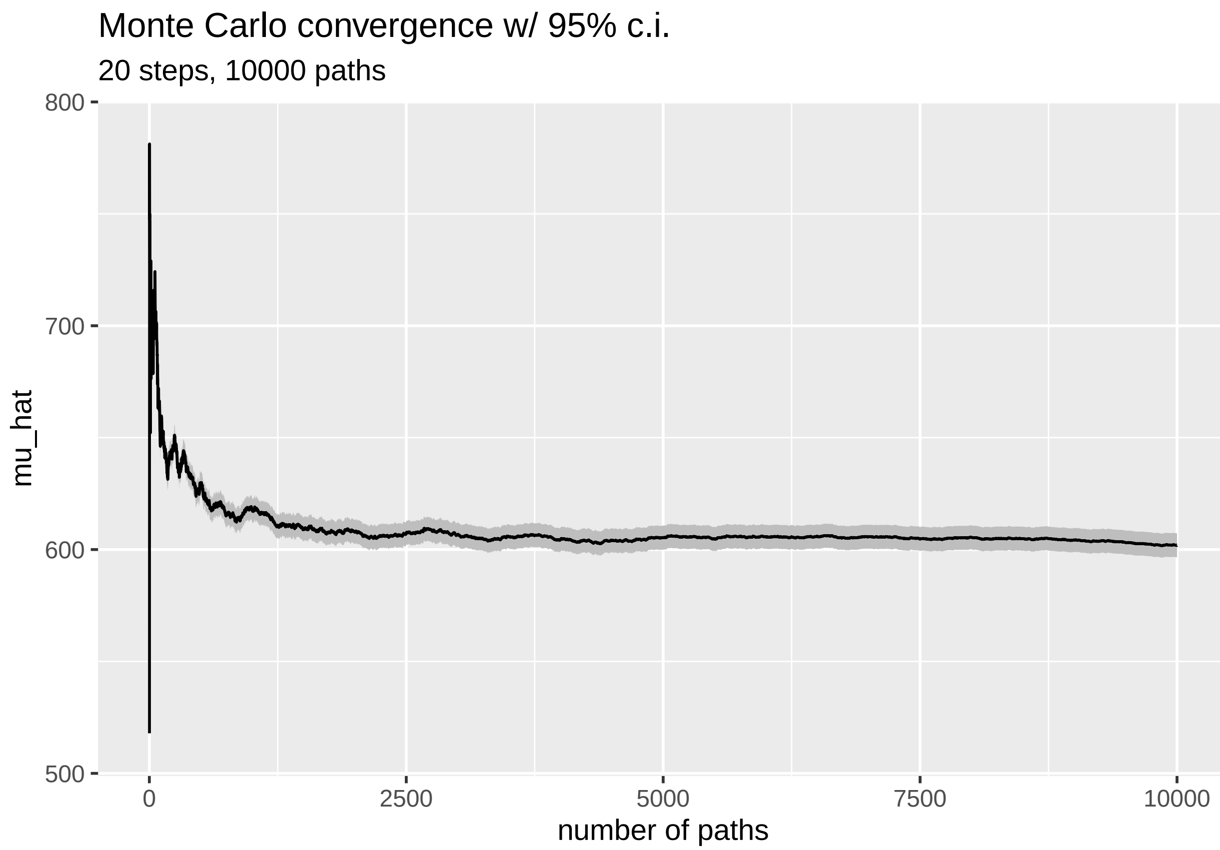


Convergence

Max vs sum

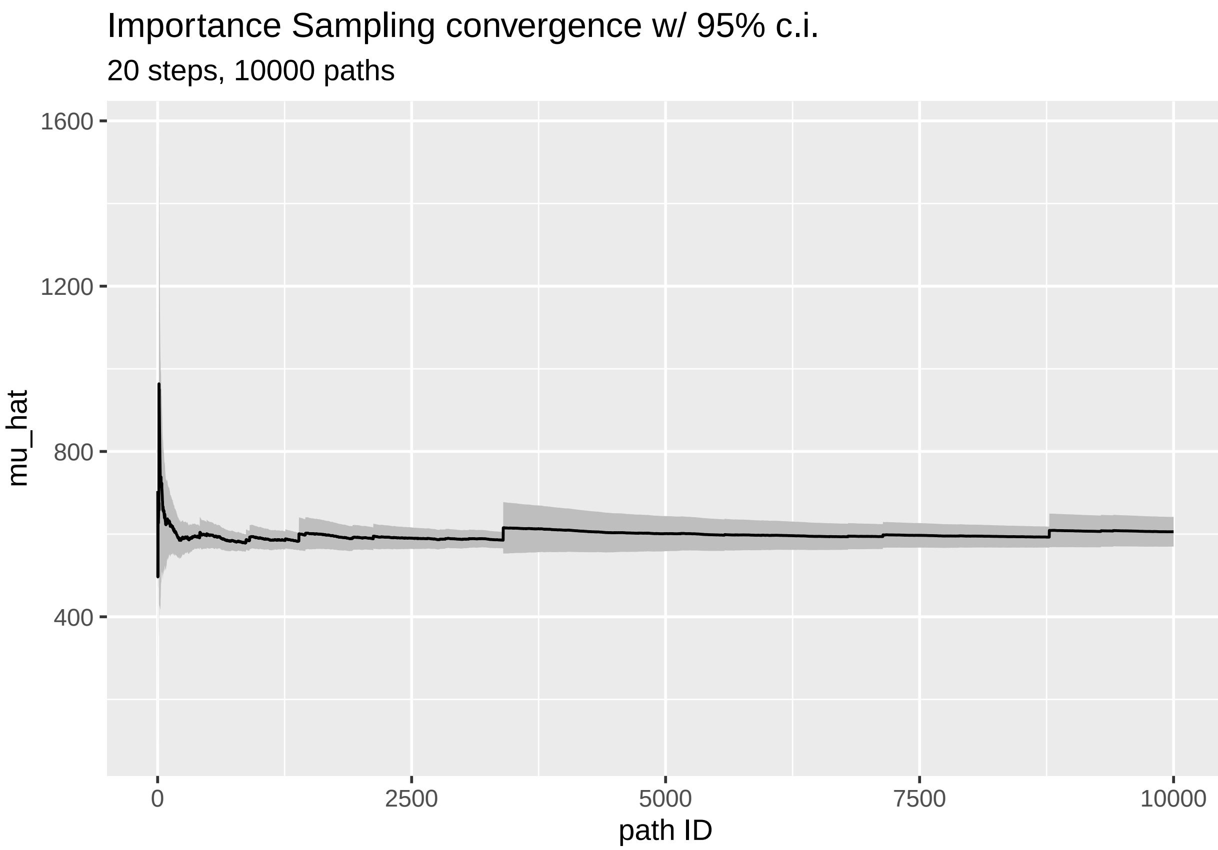
Max vs sum plots for the first four moments:


MS(1)
1.00
0.75
0.50
0.25
0.00
0
2500
5000
7500
10000
n
MS(2)
1.00
0.75
0.50
0.25
0.00
0
2500
5000
7500
10000
n
MS(3)
5
0
−5
0
2500
5000
7500
10000
n
MS(4)
1.00
0.75
0.50
0.25
0
2500
5000
7500
10000
n



Monte Carlo convergence w/ 95% c.i.
20 steps, 10000 paths
mu_hat
number of paths
800
700
600
500
0
2500
5000
7500
10000

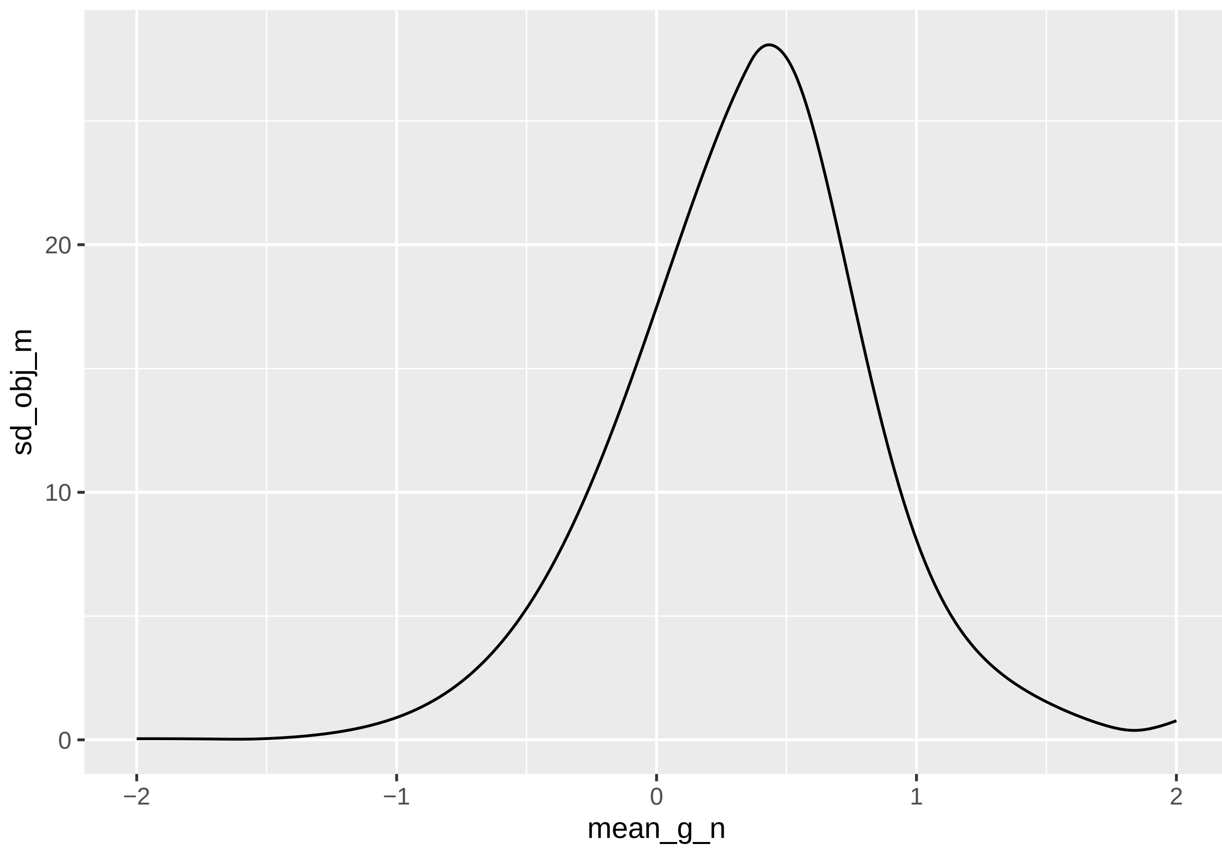
IS



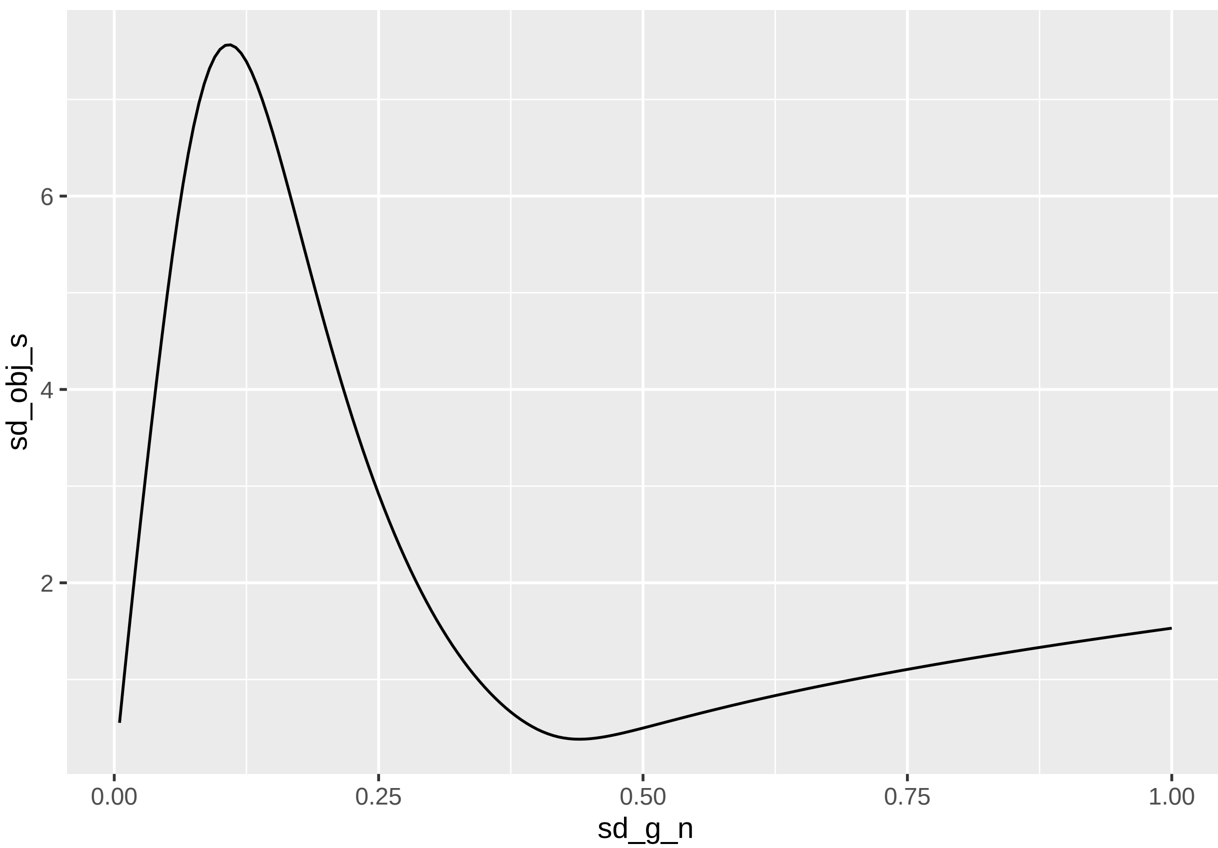
Parameters

```
## [1] 1.8379614 0.4397688
```

Objective function plots

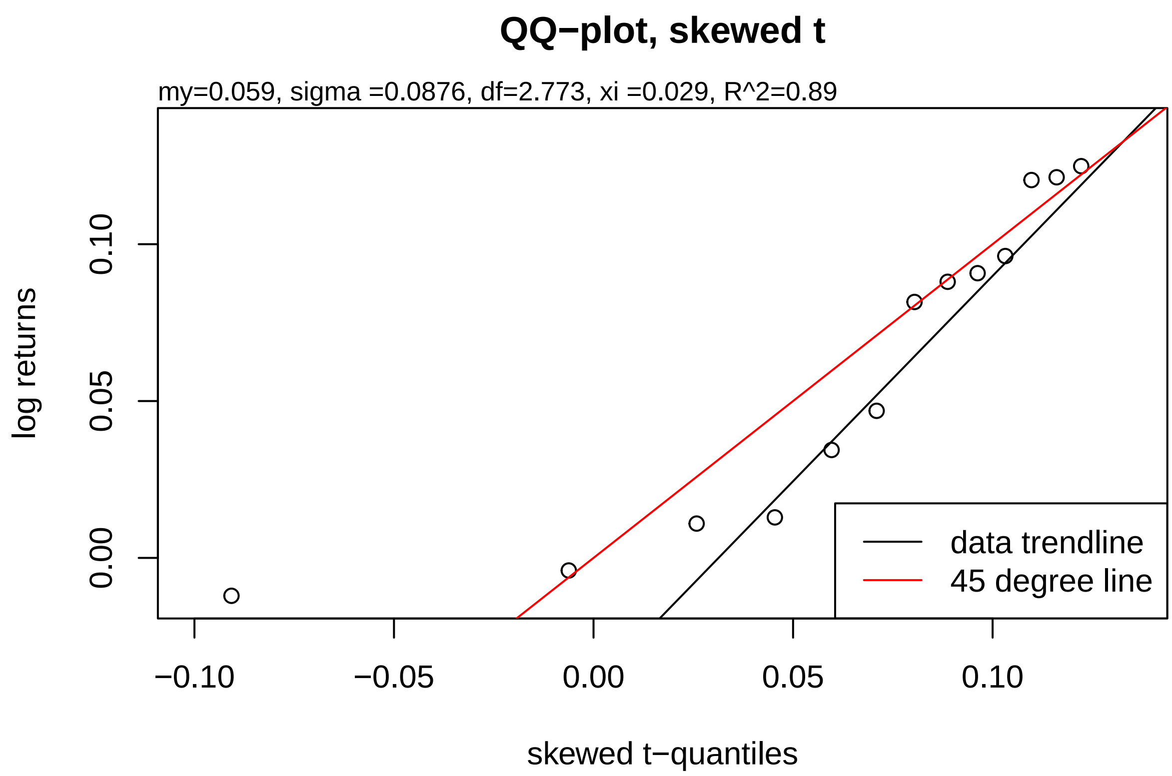


20
10
0
sd_obj_m
−2
−1
0
1
2
mean_g_n


sd_obj_s
6
4
2
0.00
0.25
0.50
0.75
1.00
sd_g_n

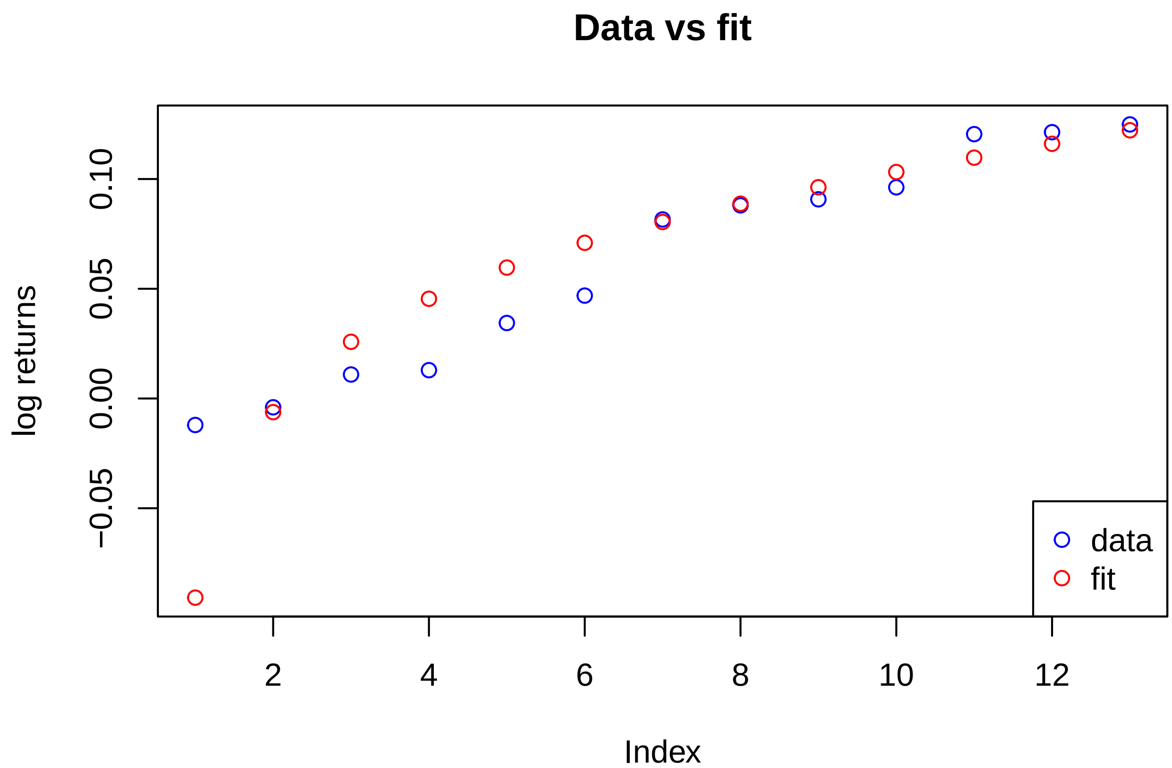
Mix medium risk (mmr), 2011 - 2023

QQ Plot



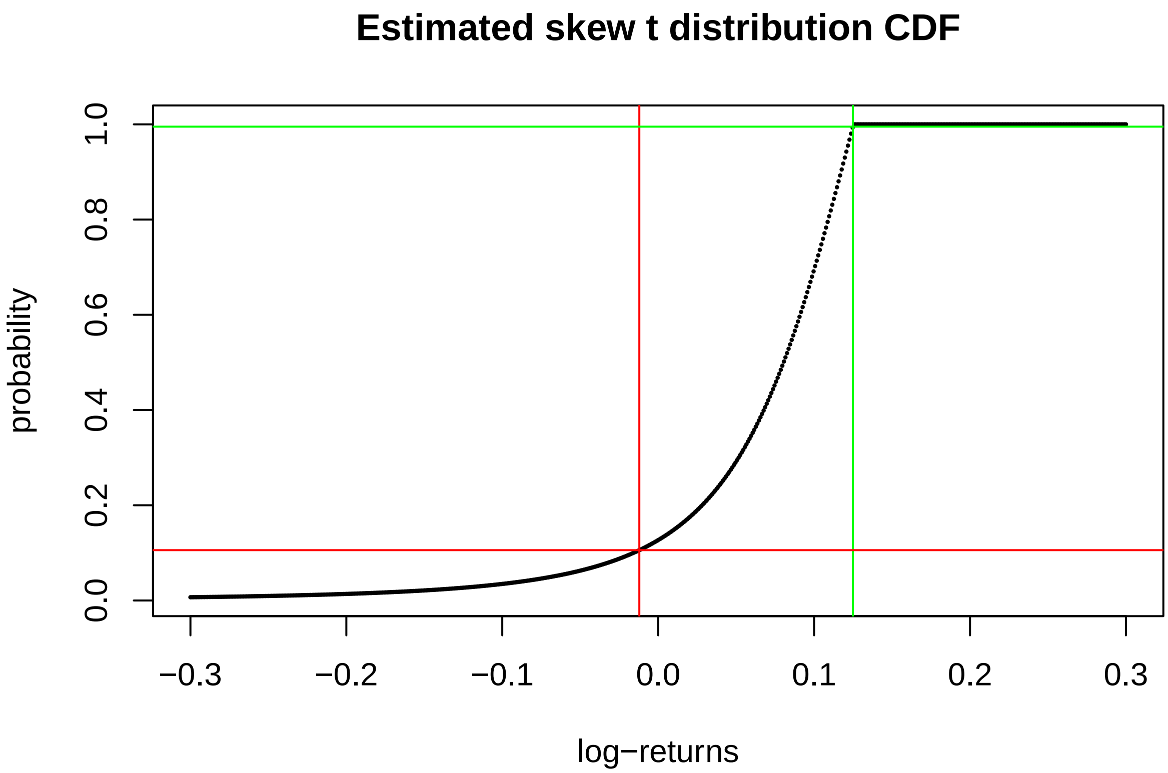
Data vs fit

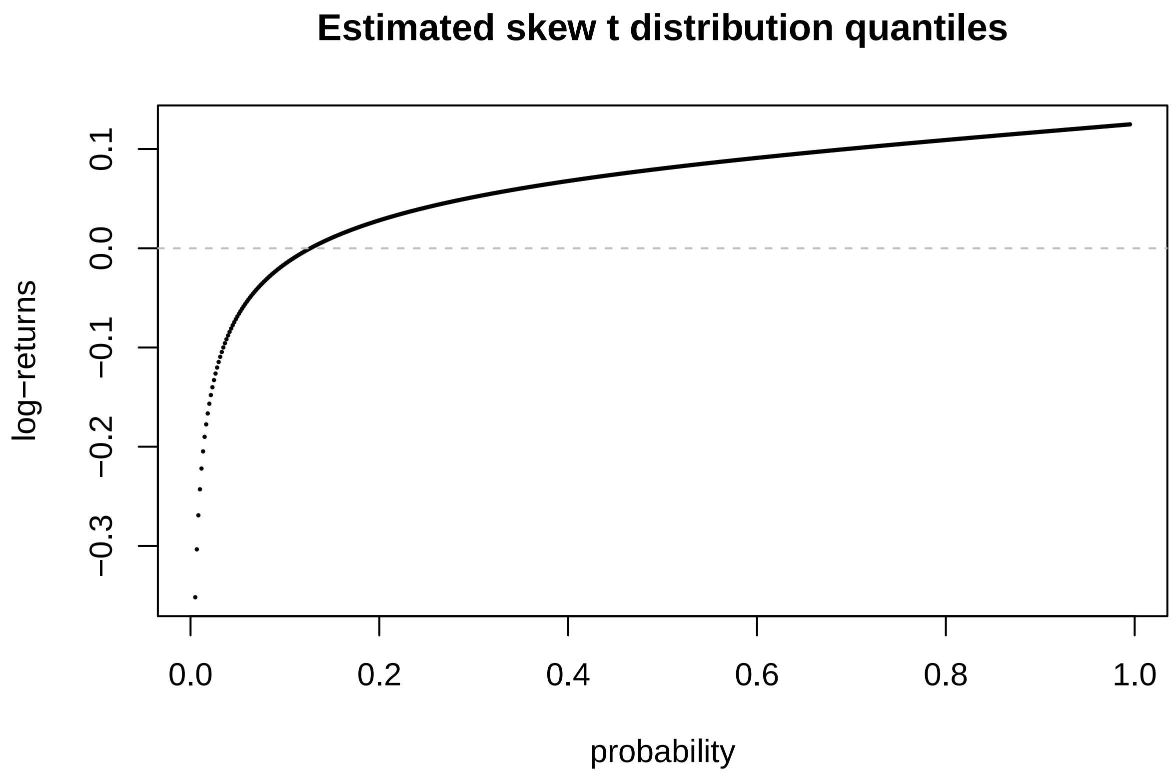
Let's plot the fit and the observed returns together.

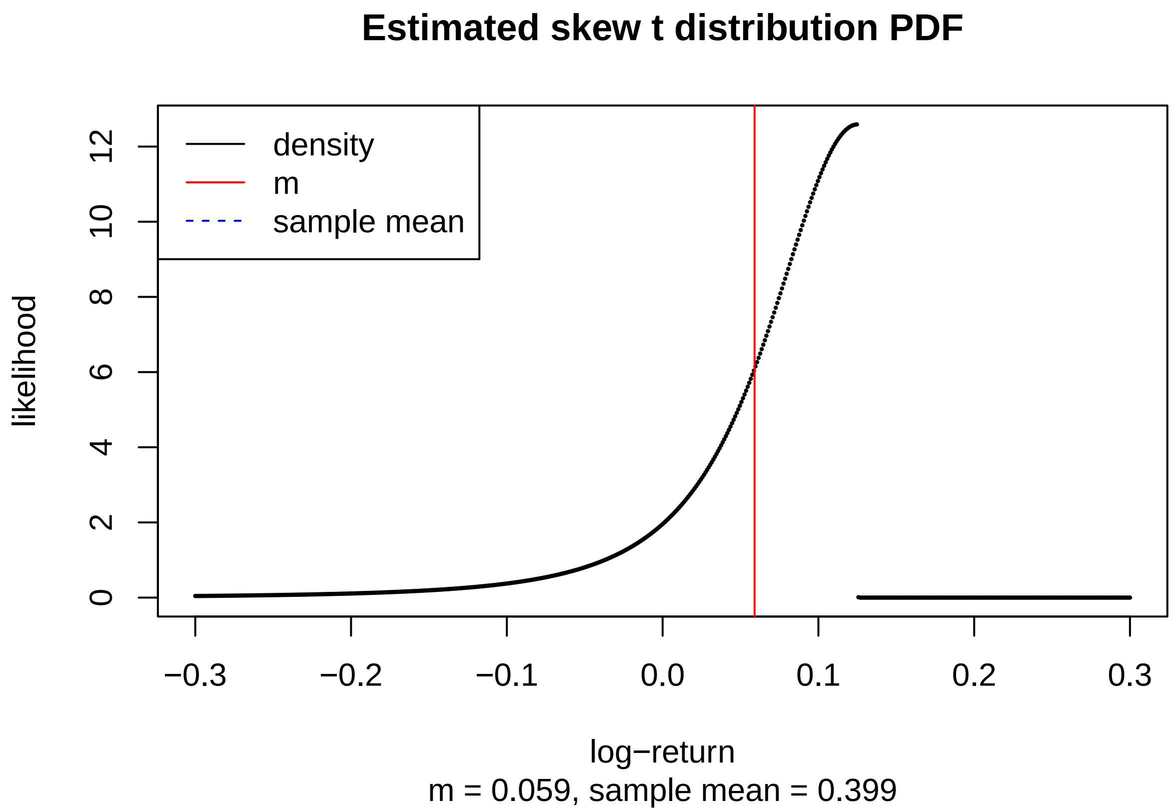


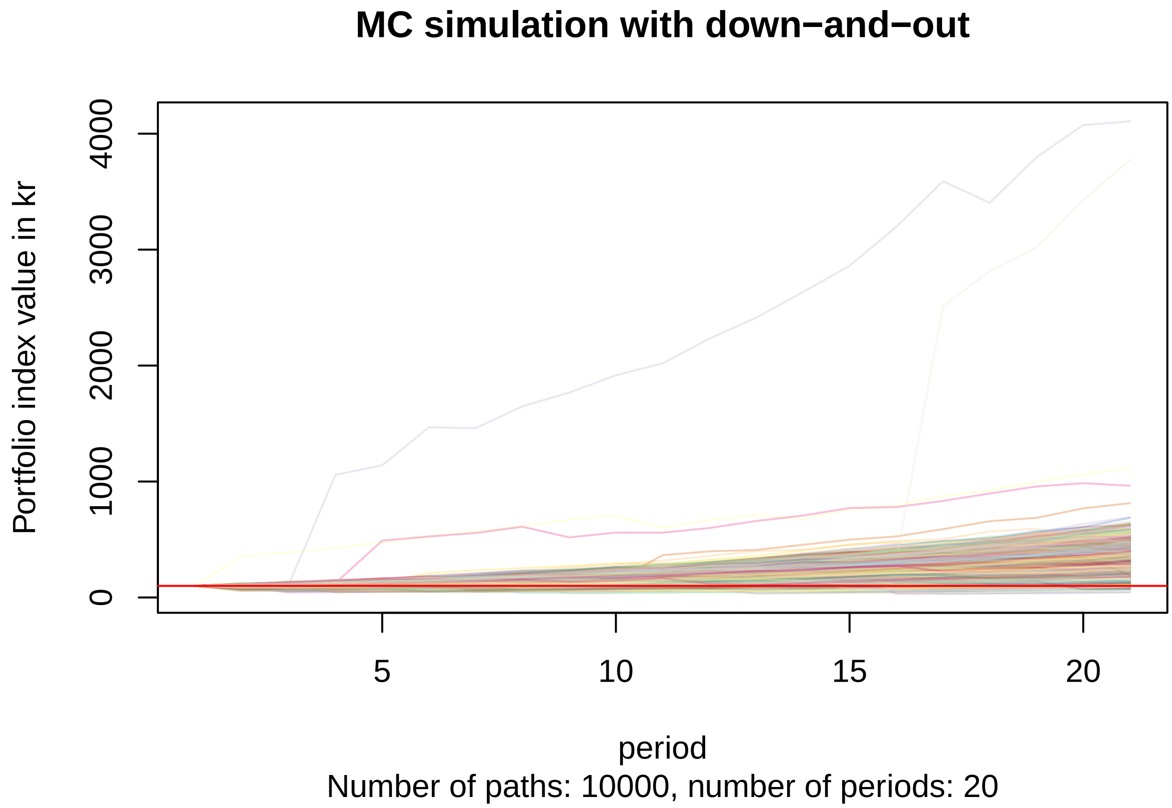
Estimated distribution

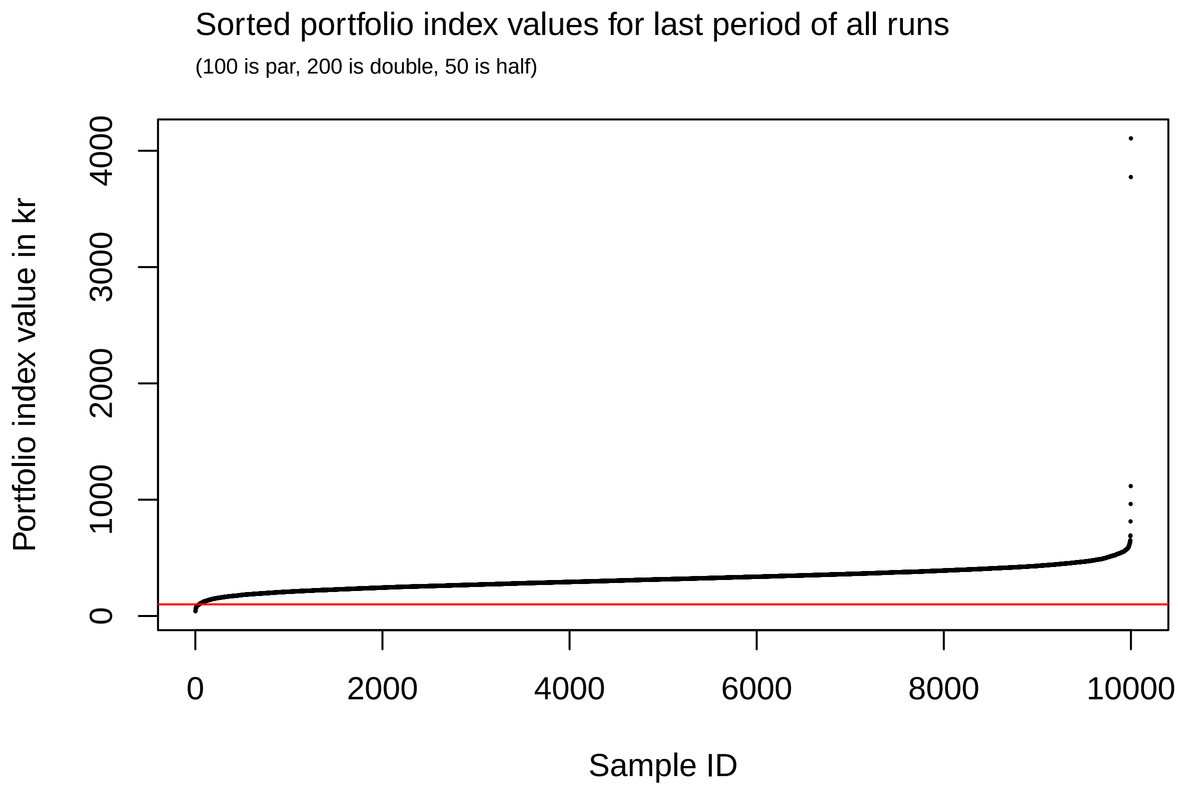
Now lets look at the CDF of the estimated distribution for each 0.1% increment between 0.5% and 99.5% for the estimated distribution:


Estimated skew t distribution CDF
probability
0.0
0.2
0.4
0.6
0.8
1.0
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
log−returns


Estimated skew t distribution quantiles
log−returns
0.1
0.0
−0.1
−0.2
−0.3
0.0
0.2
0.4
0.6
0.8
1.0
probability


Estimated skew t distribution PDF
density
m
sample mean
likelihood
0
2
4
6
8
10
12
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
log−return
m = 0.059, sample mean = 0.399

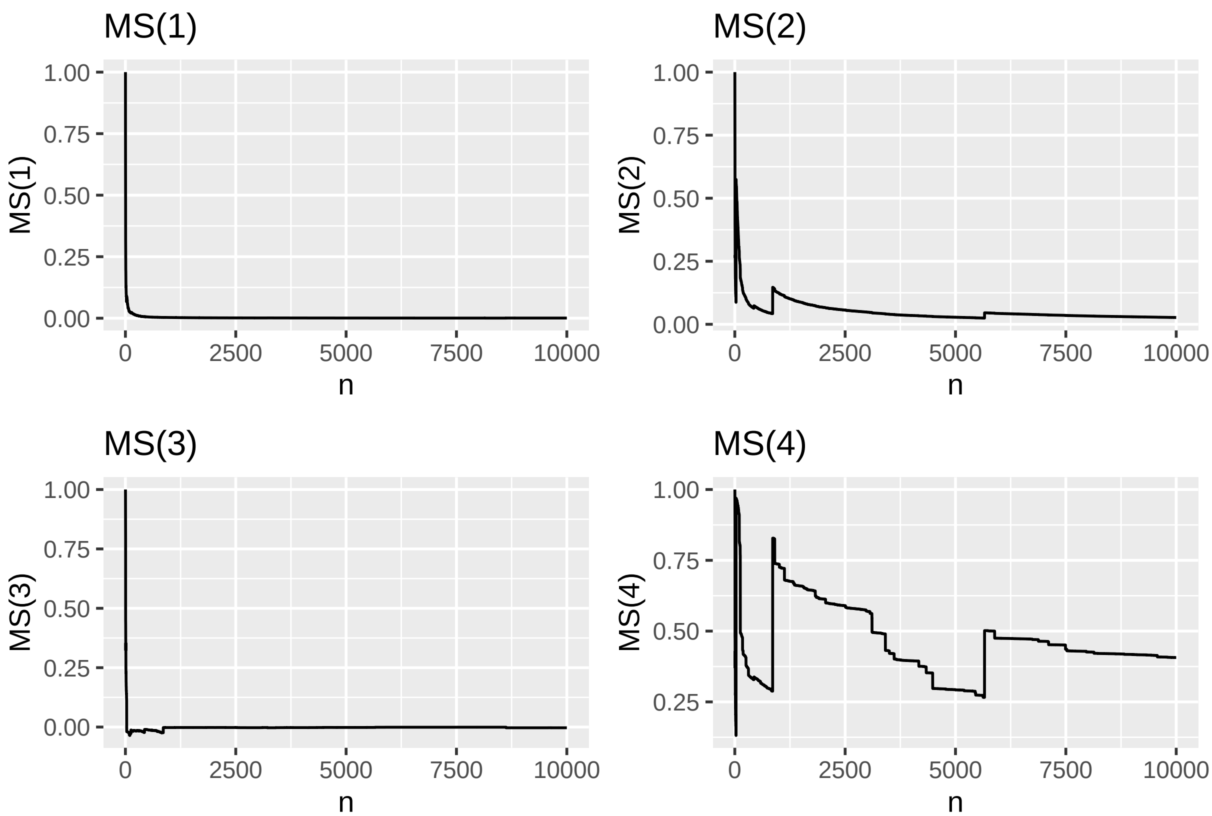

MC simulation with down−and−out
Portfolio index value in kr
0
1000
2000
3000
4000
5
10
15
20
period
Number of paths: 10000, number of periods: 20



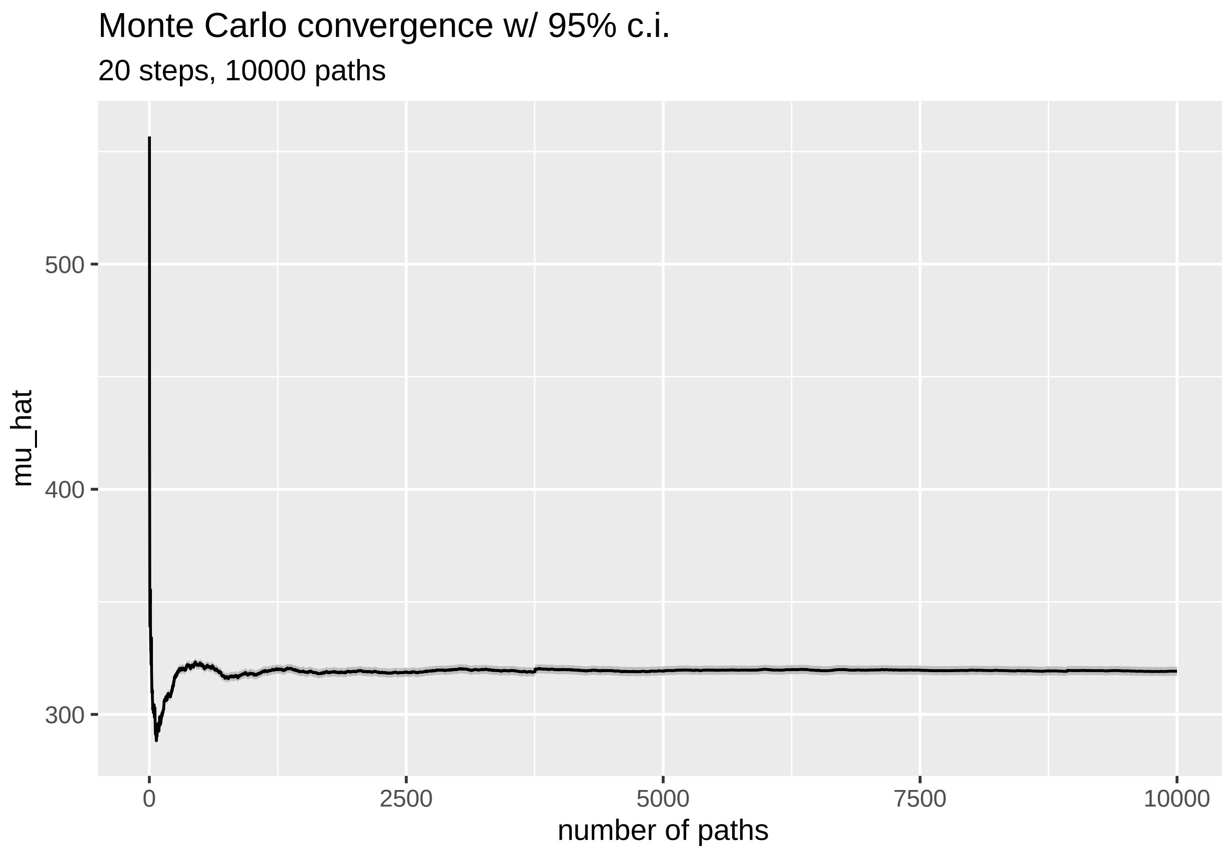
Convergence

Max vs sum

Max vs sum plots for the first four moments:

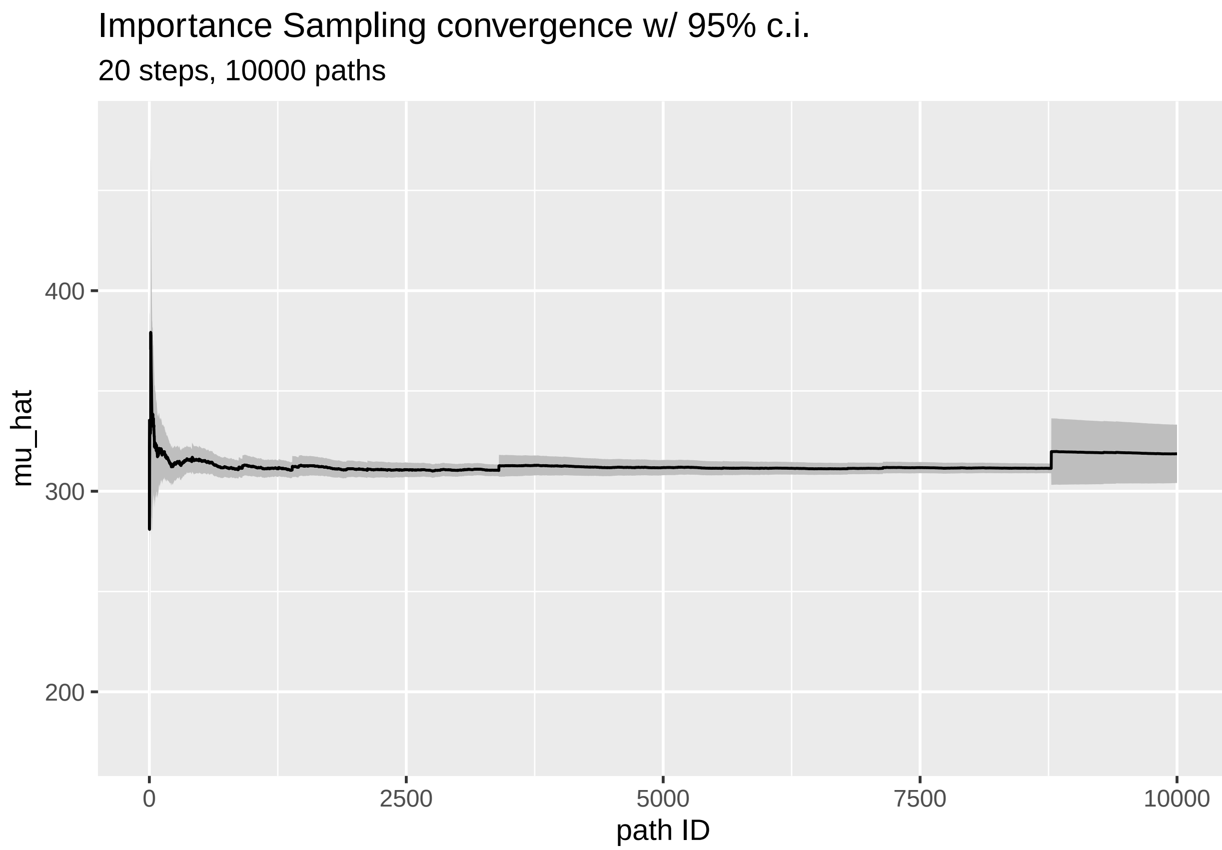

MS(1)
MS(2)
MS(3)
MS(4)
1.00
0.75
0.50
0.25
0.00
0
2500
5000
7500
10000
n

MC



Monte Carlo convergence w/ 95% c.i.
20 steps, 10000 paths
mu_hat
500
400
300
0
2500
5000
7500
10000
number of paths

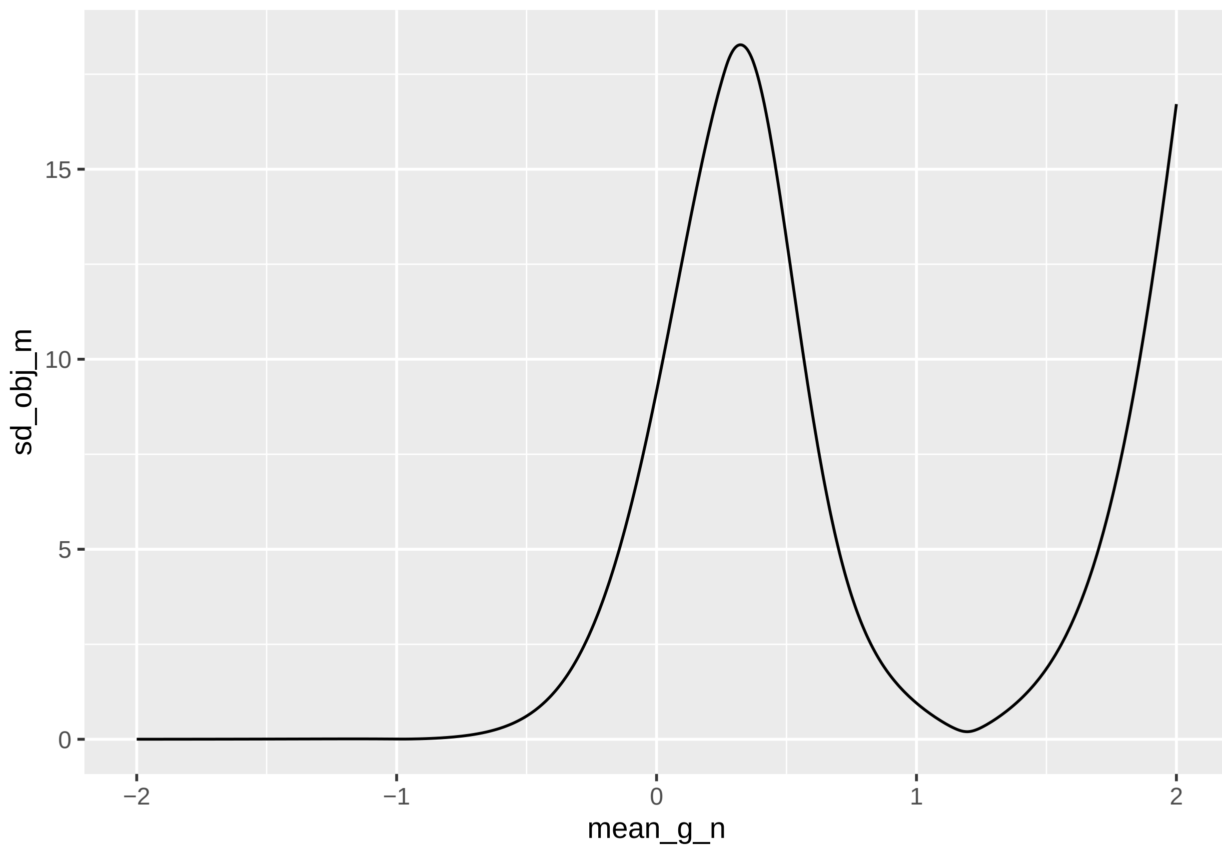
IS



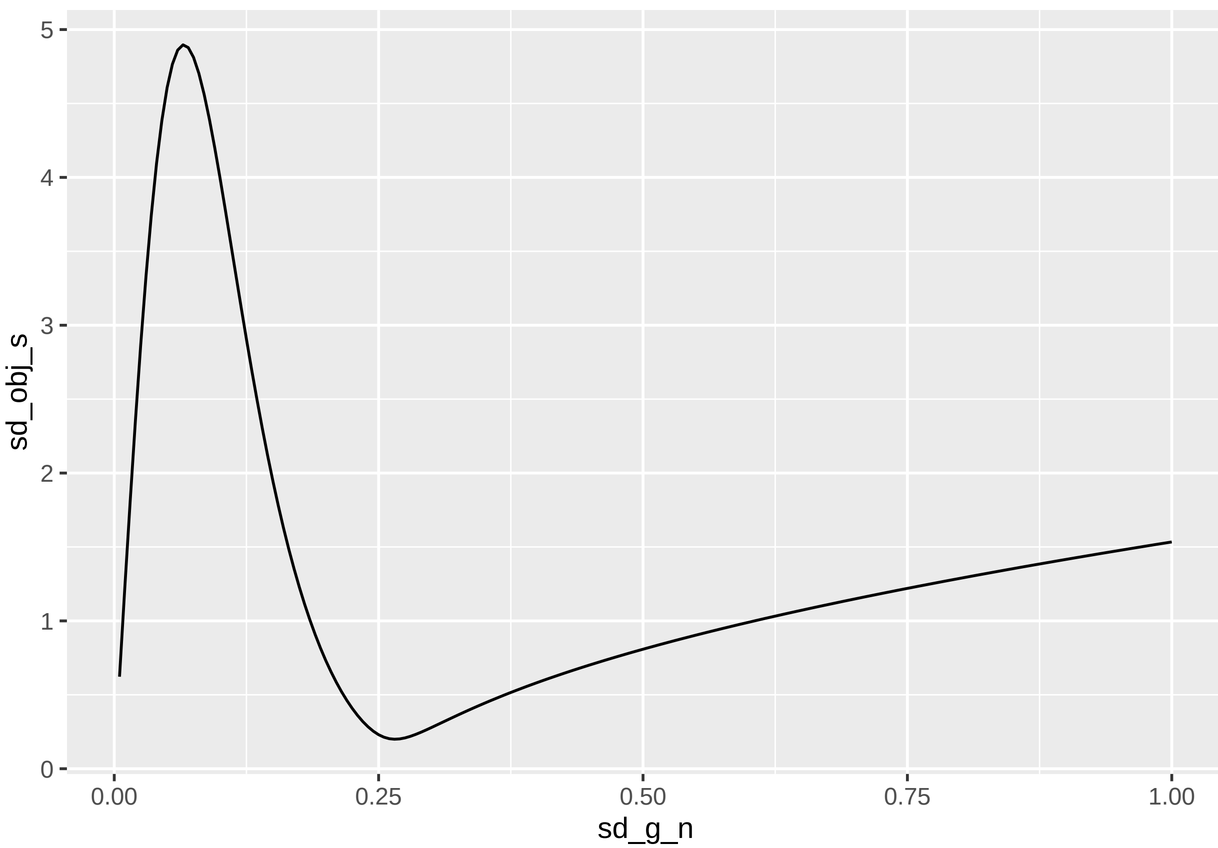
Parameters

```
## [1] 1.1948623 0.2654885
```

Objective function plots

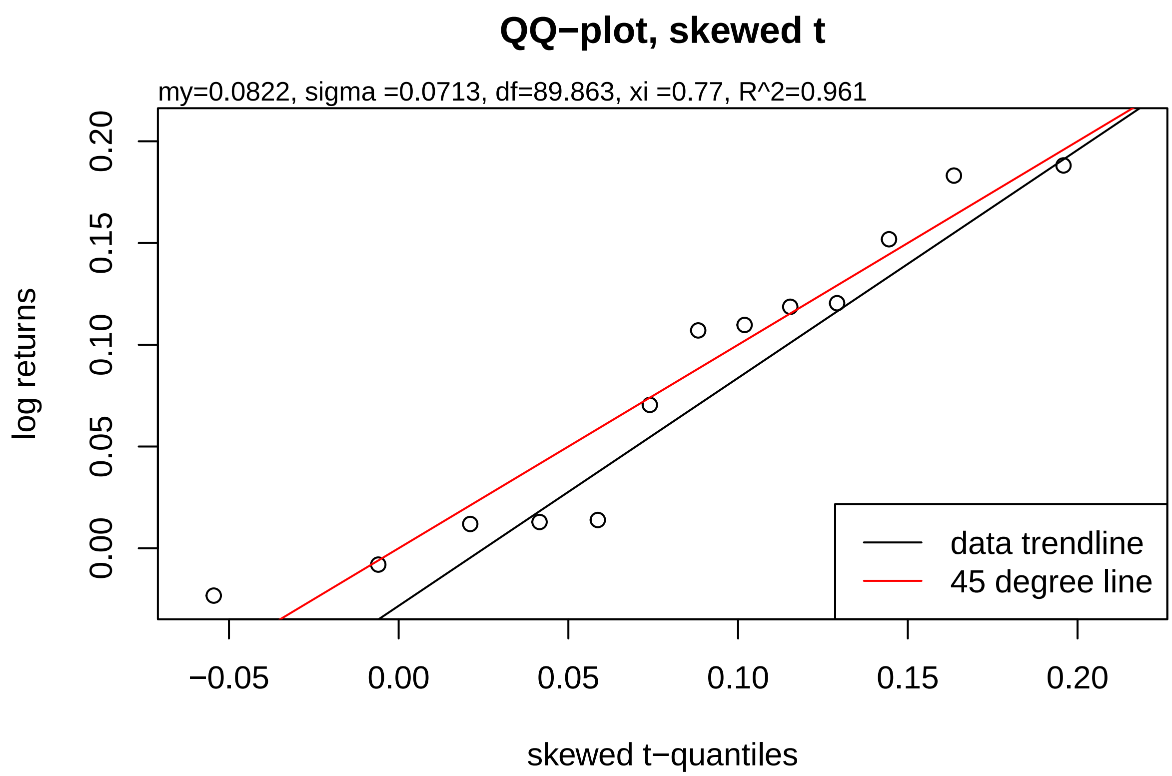


sd_obj_m
15
10
5
0
−2
−1
0
1
2
mean_g_n


sd_obj_s
0
1
2
3
4
5
0.00
0.25
0.50
0.75
1.00
sd_g_n

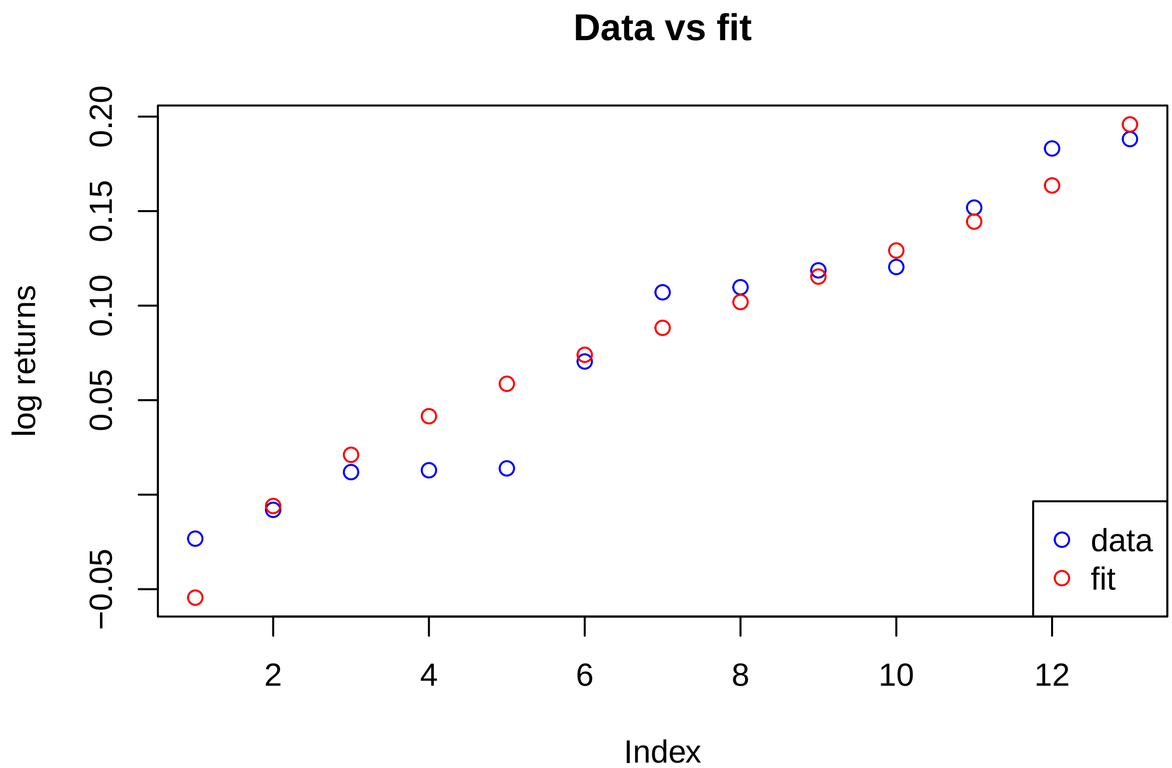
Mix high risk (mhr), 2011 - 2023

QQ Plot



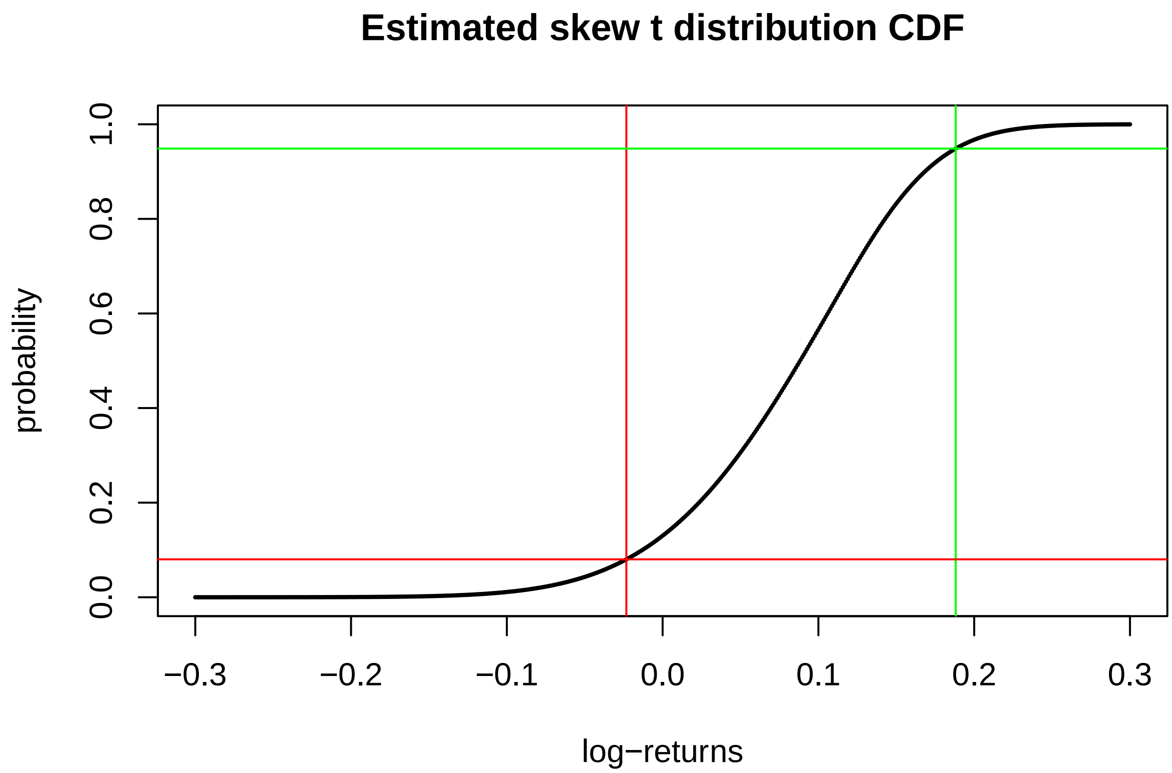
Data vs fit

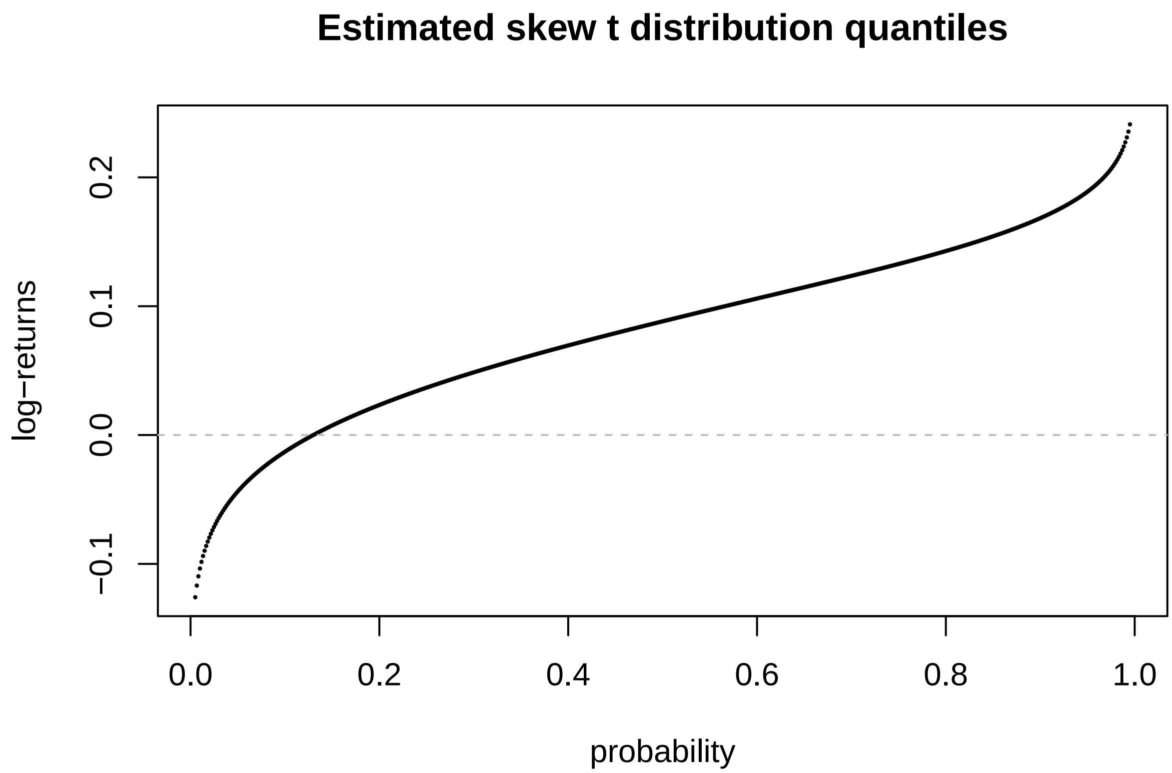
Let's plot the fit and the observed returns together.

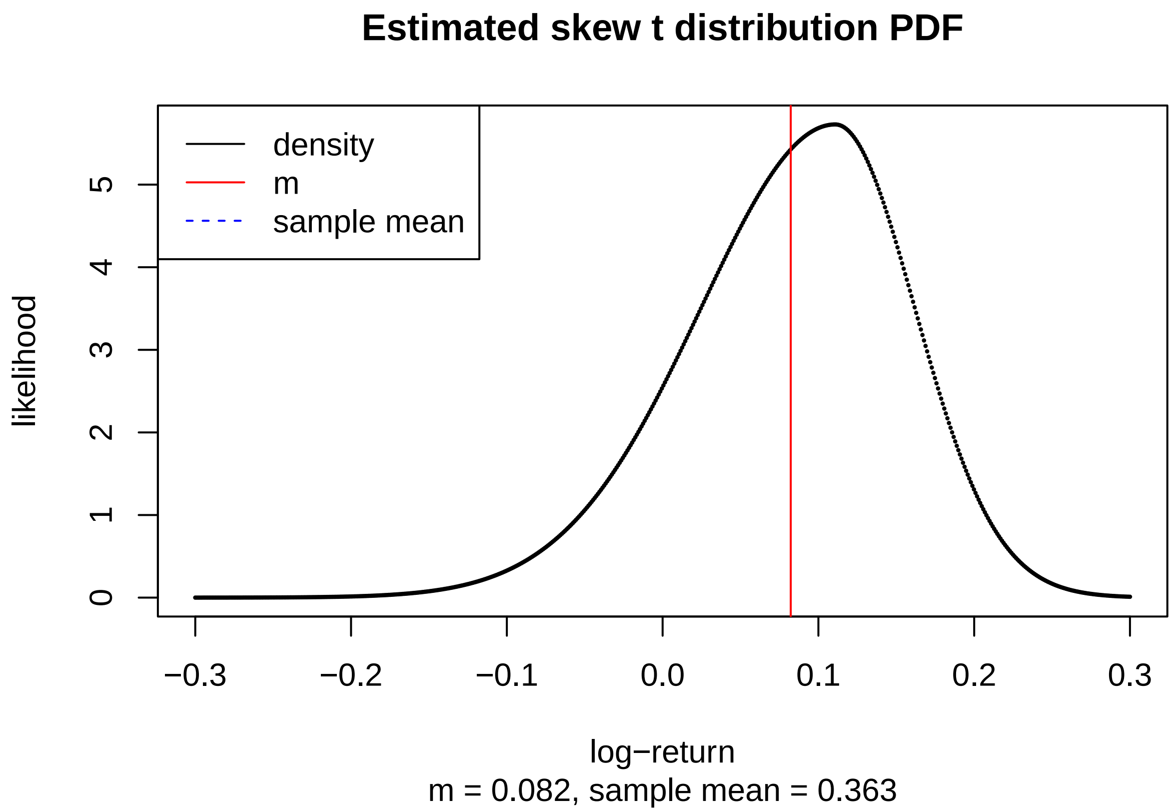


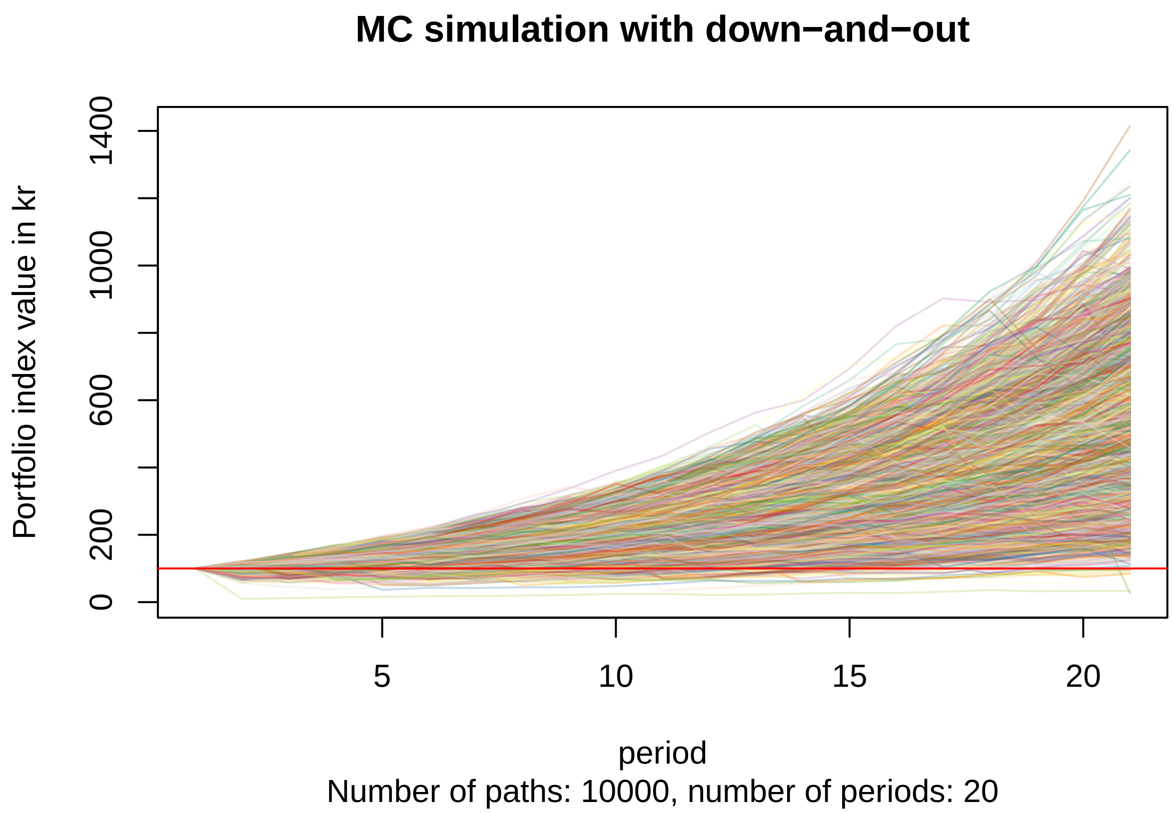
Estimated distribution

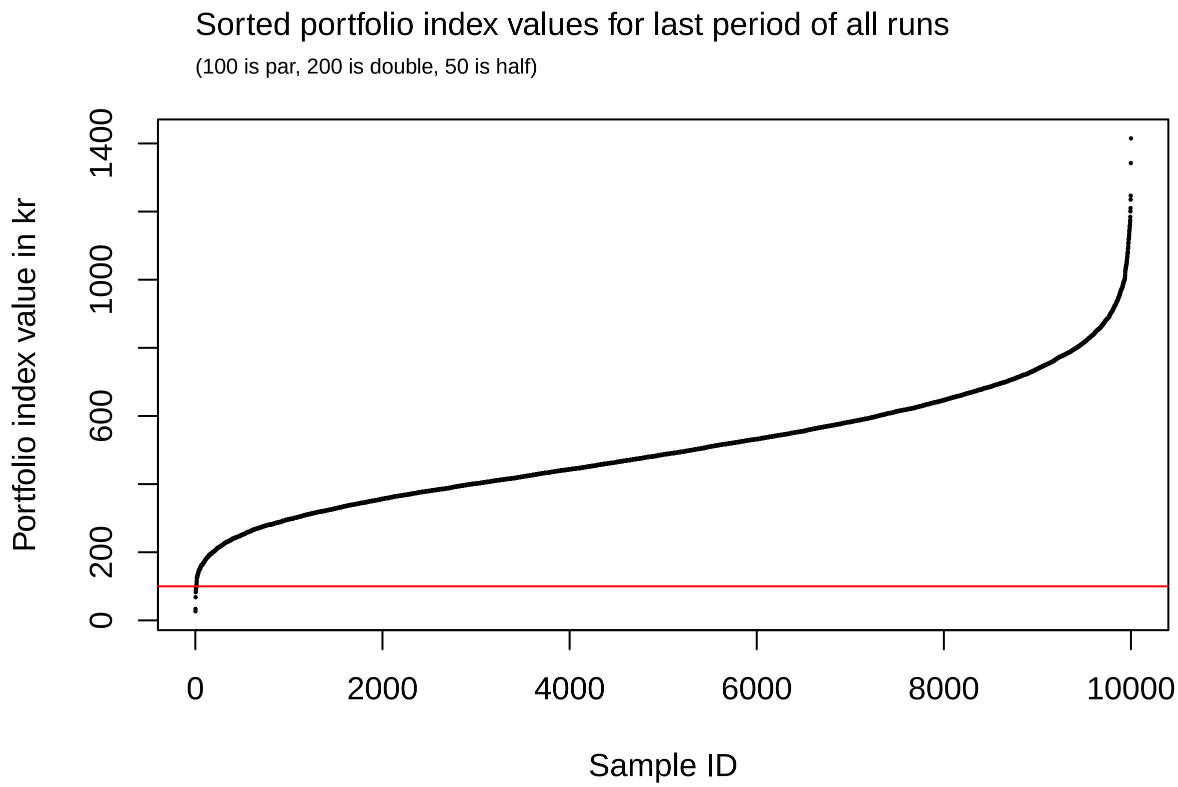
Now lets look at the CDF of the estimated distribution for each 0.1% increment between 0.5% and 99.5% for the estimated distribution:


Estimated skew t distribution CDF
probability
log−returns
0.0
0.2
0.4
0.6
0.8
1.0
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3


Estimated skew t distribution quantiles
log−returns
−0.1
0.0
0.1
0.2
0.0
0.2
0.4
0.6
0.8
1.0
probability


Estimated skew t distribution PDF
density
m
sample mean
likelihood
0
1
2
3
4
5
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
log−return
m = 0.082, sample mean = 0.363

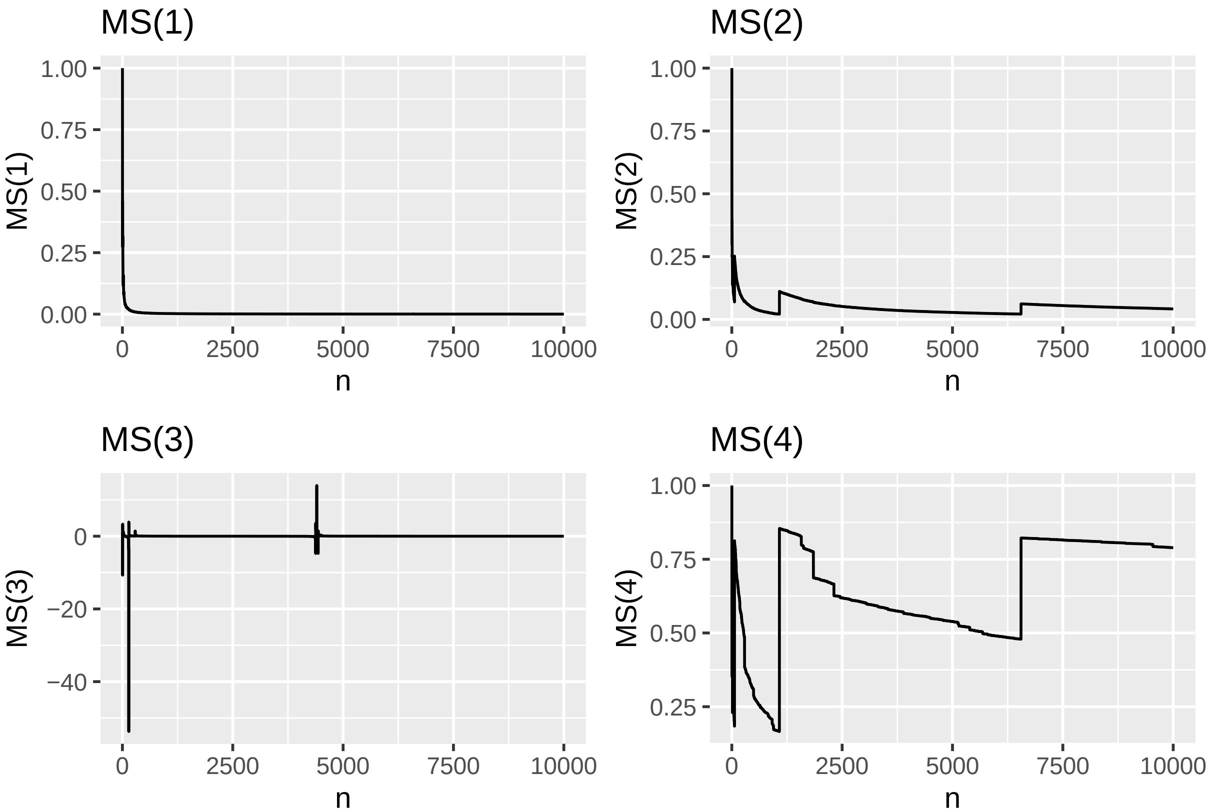

MC simulation with down−and−out
Portfolio index value in kr
0
200
600
1000
1400
5
10
15
20
period
Number of paths: 10000, number of periods: 20



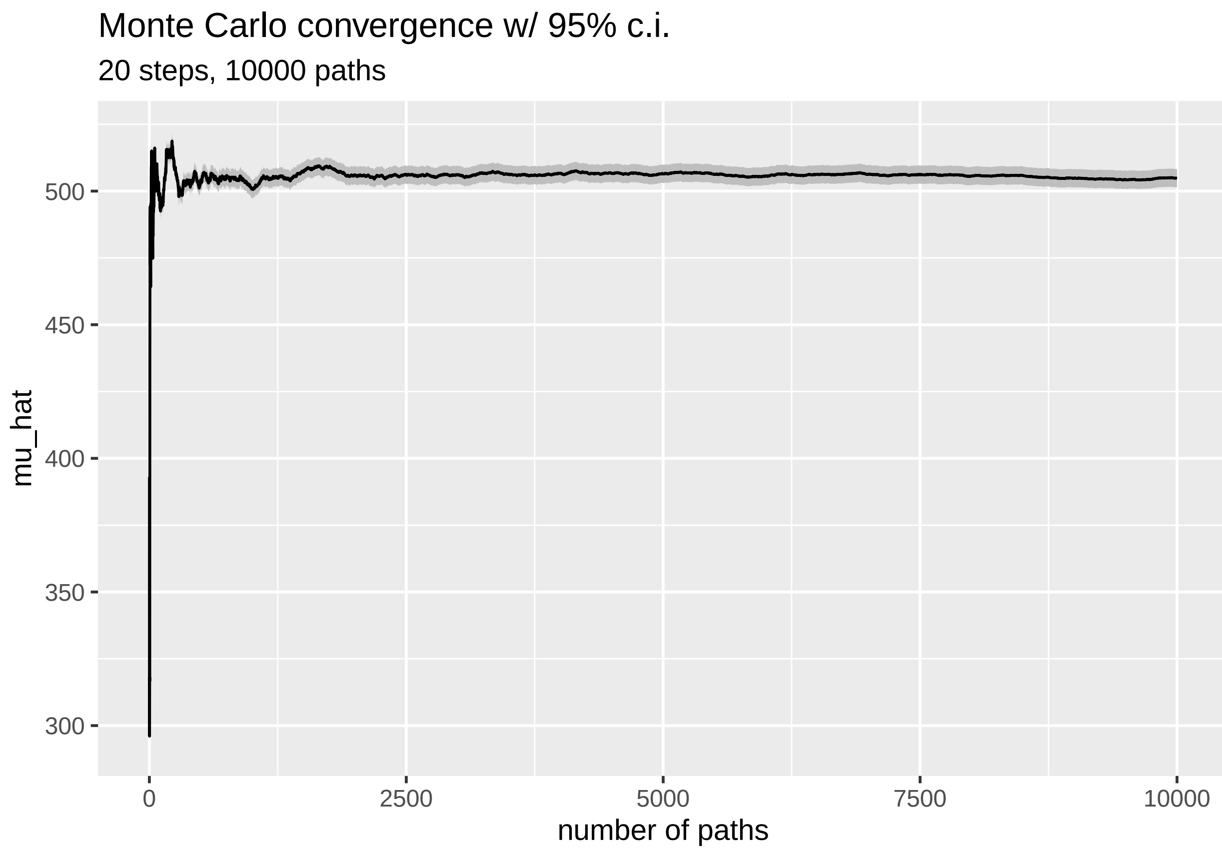
Convergence

Max vs sum

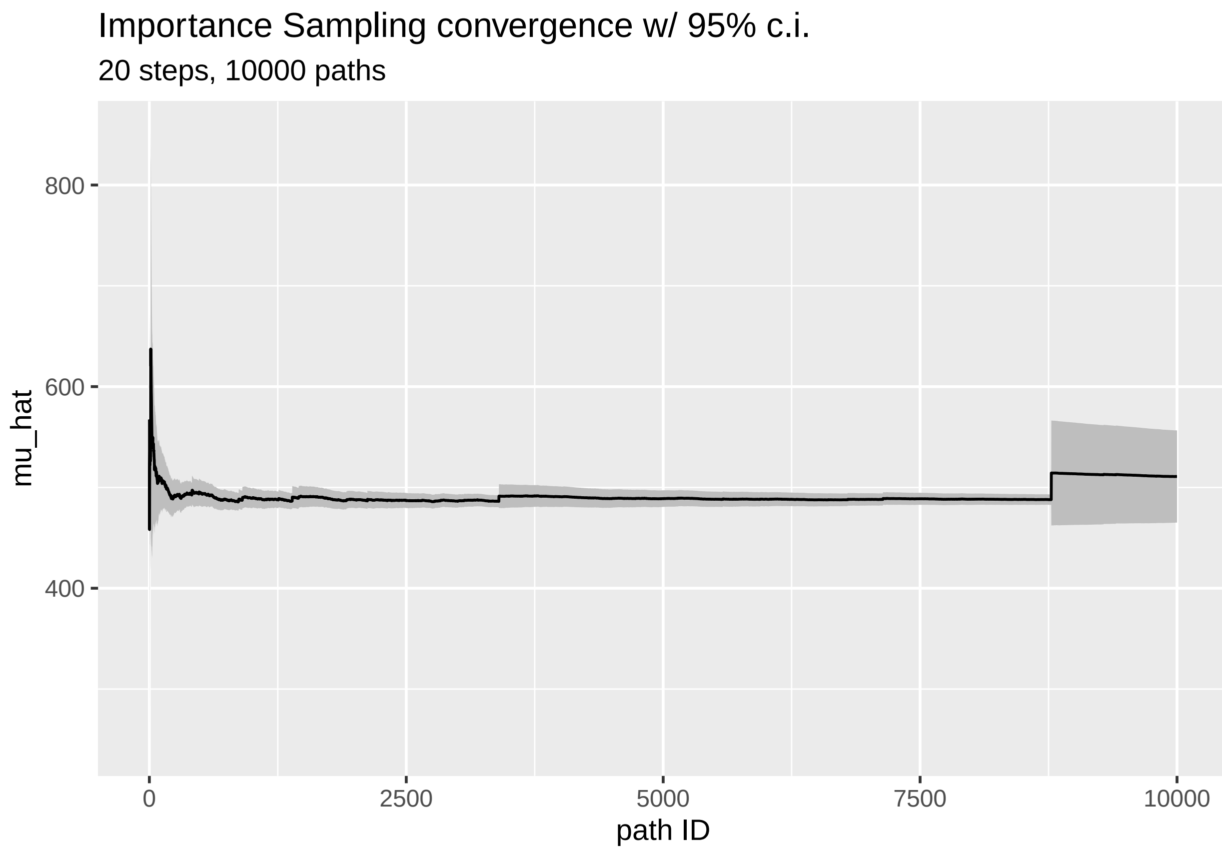
Max vs sum plots for the first four moments:



MS(1)
1.00
0.75
0.50
0.25
0.00
0
2500
5000
7500
10000
n
MS(2)
1.00
0.75
0.50
0.25
0.00
0
2500
5000
7500
10000
n
MS(3)
0
−20
−40
0
2500
5000
7500
10000
n
MS(4)
1.00
0.75
0.50
0.25
0
2500
5000
7500
10000
n


Monte Carlo convergence w/ 95% c.i.
20 steps, 10000 paths
mu_hat
500
450
400
350
300
0
2500
5000
7500
10000
number of paths

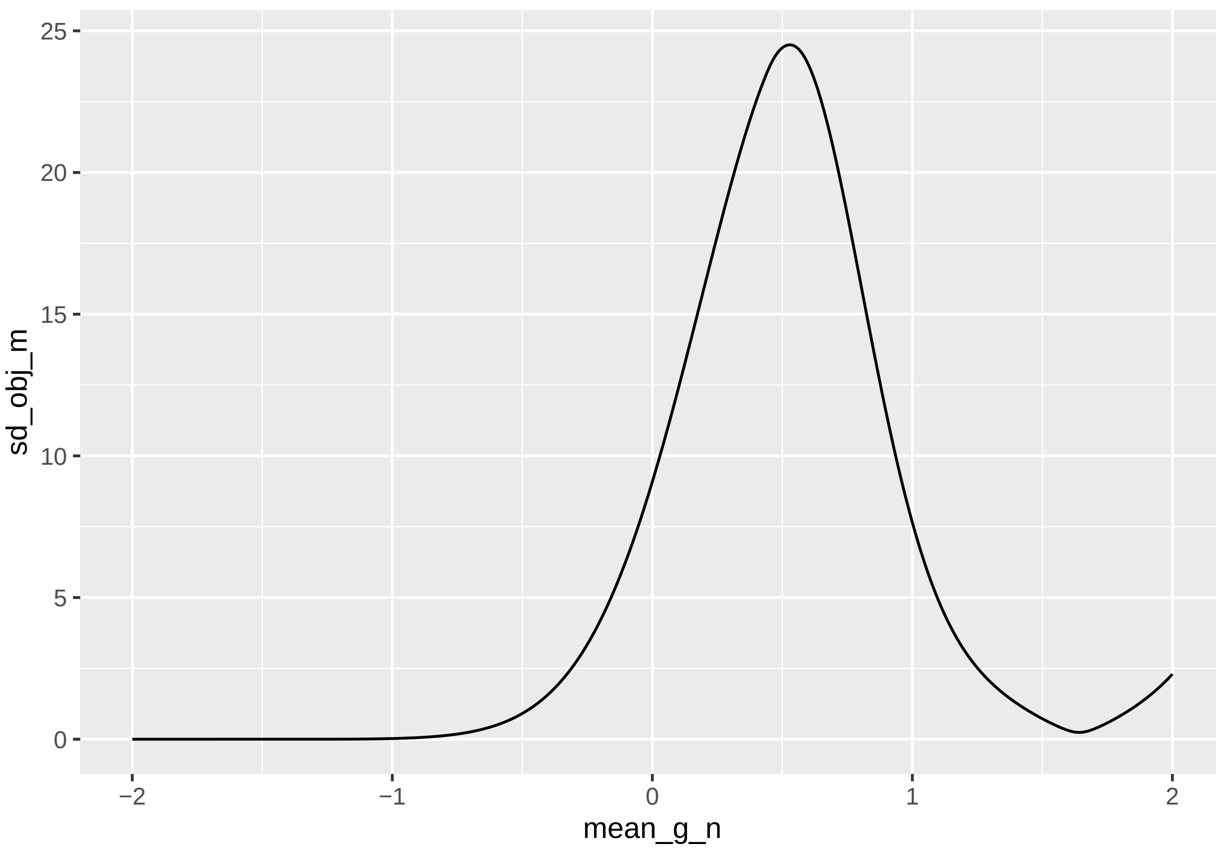
IS

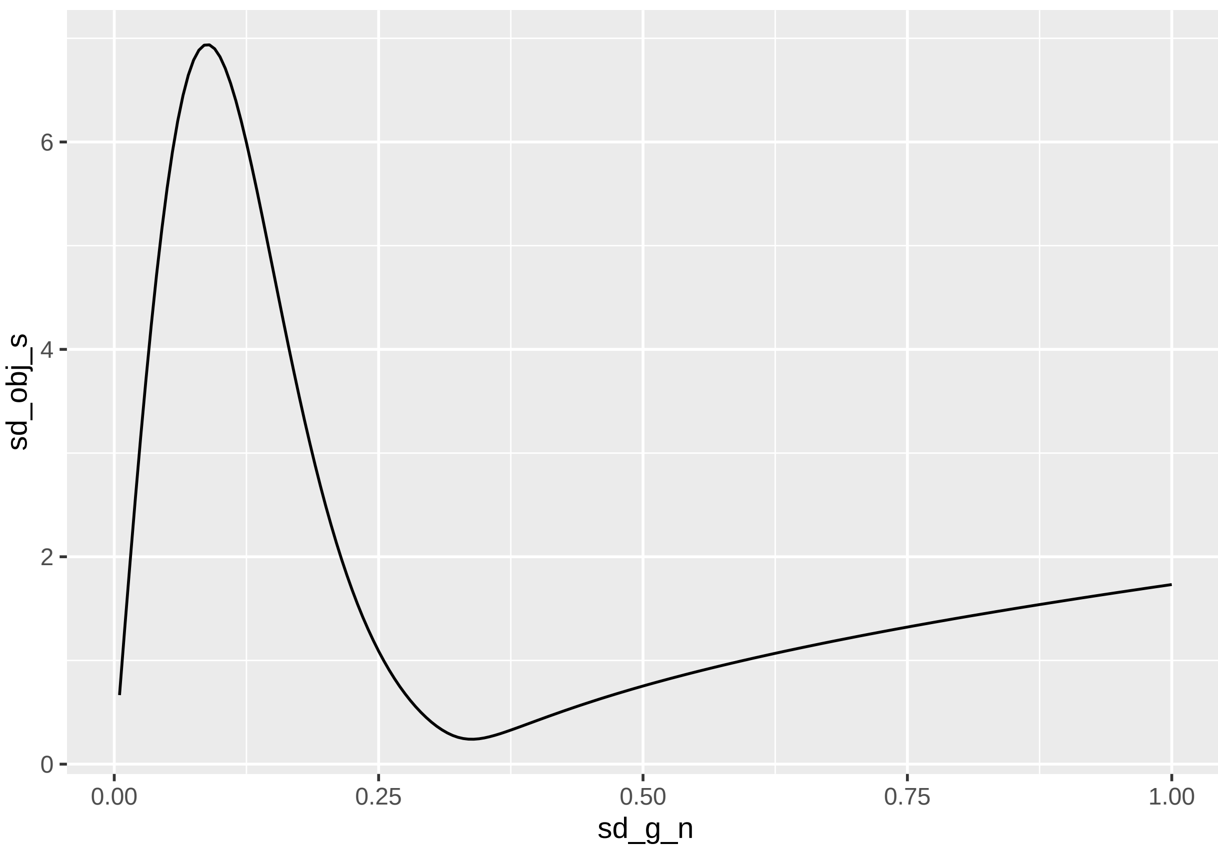


Parameters

```
## [1] 1.6413478 0.3380133
```

Objective function plots

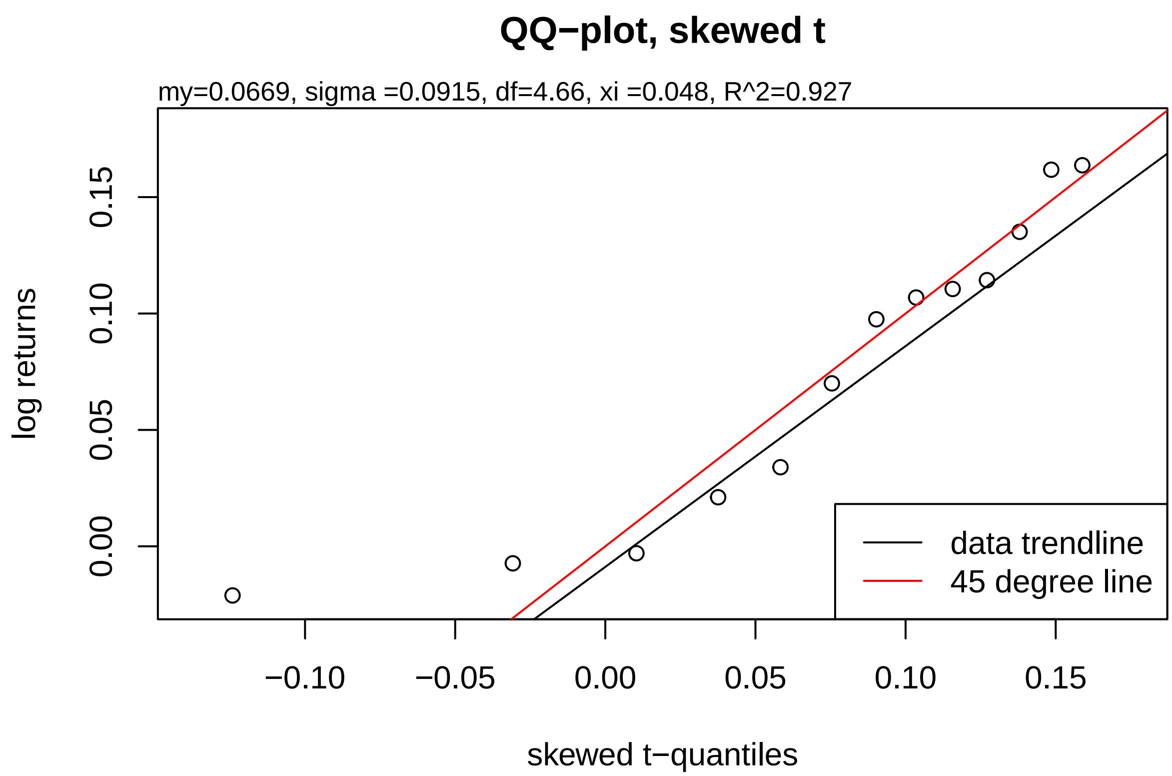

sd_obj_m
25
20
15
10
5
0
−2
−1
0
1
2
mean_g_n



sd_obj_s
sd_g_n
0
2
4
6
0.00
0.25
0.50
0.75
1.00

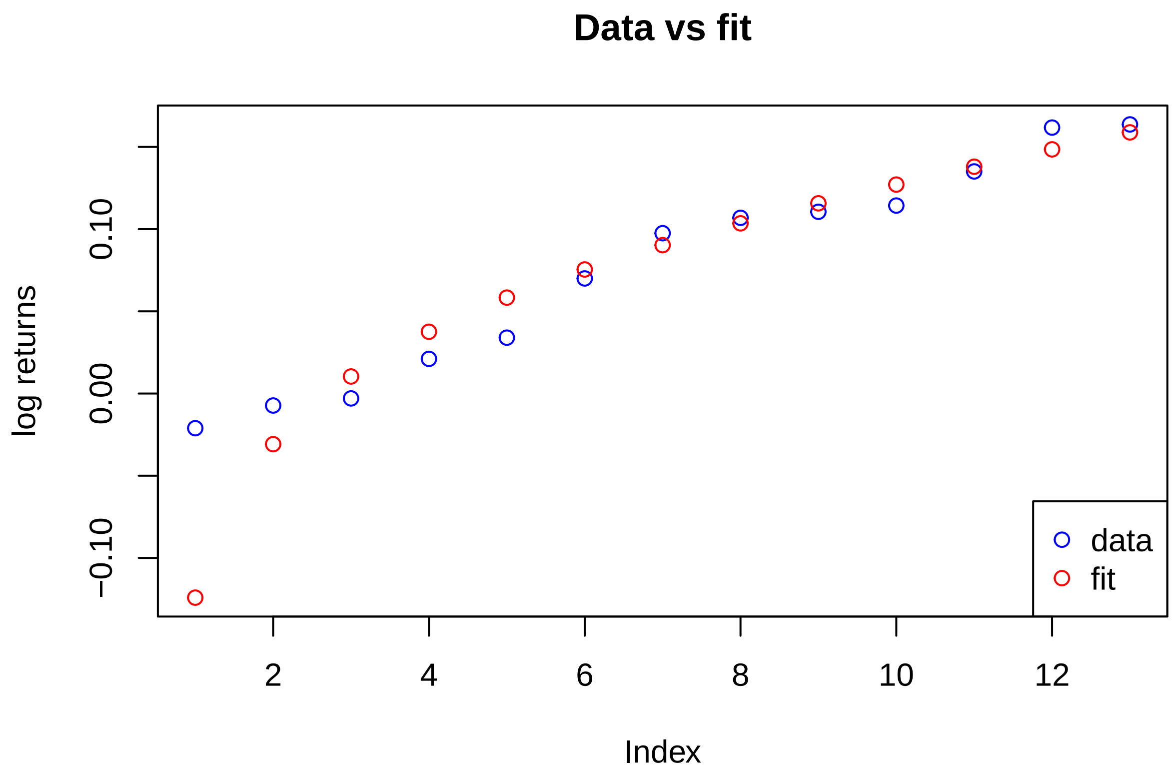
Mix vmr+phr (vm_ph), 2011 - 2023

QQ Plot



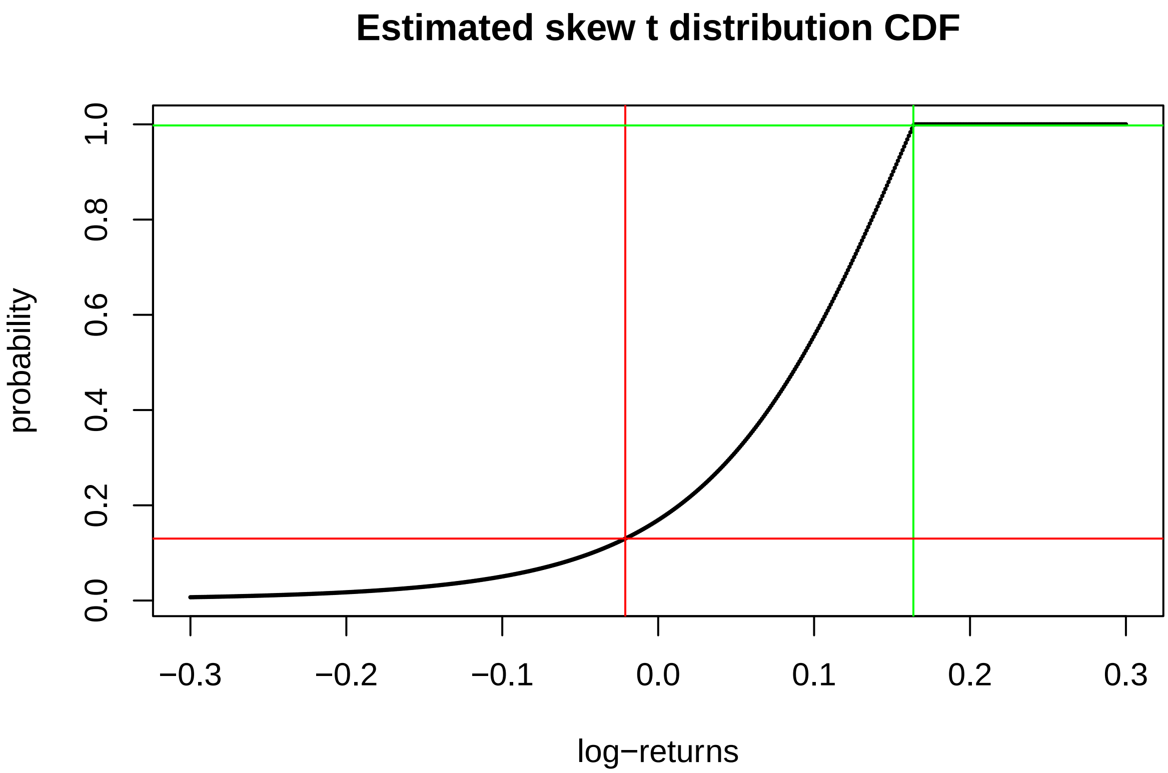
Data vs fit

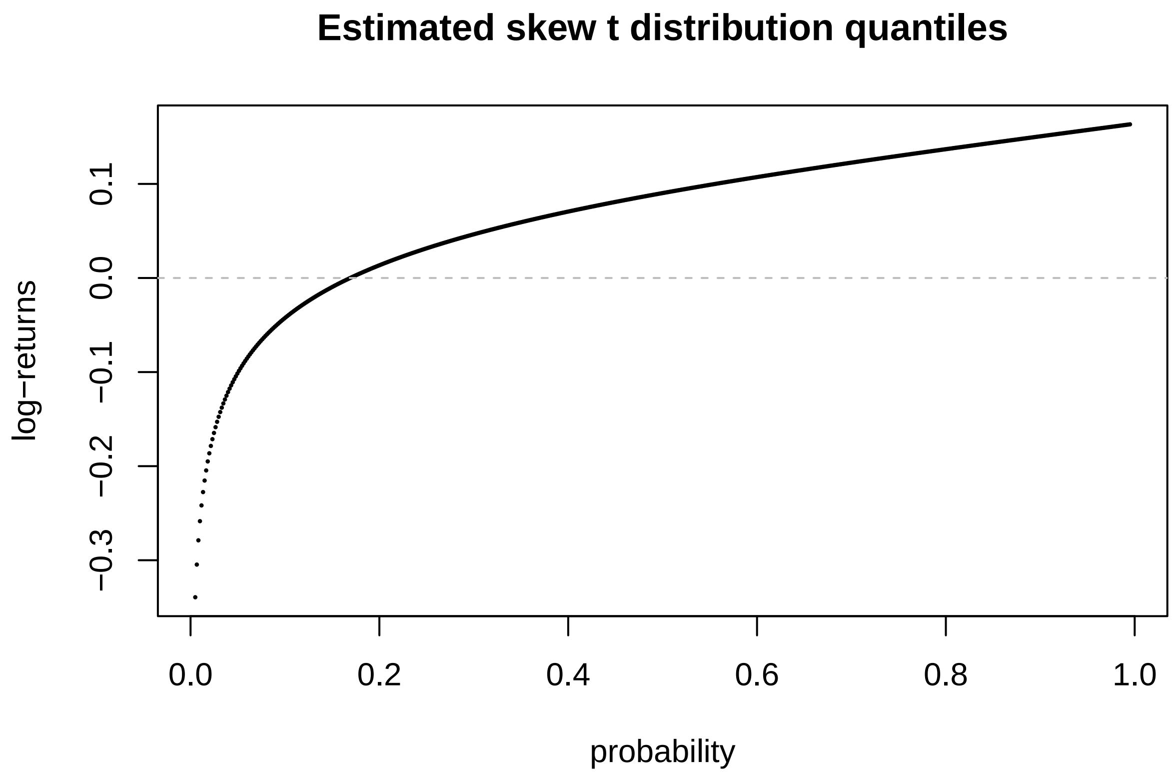
Let's plot the fit and the observed returns together.



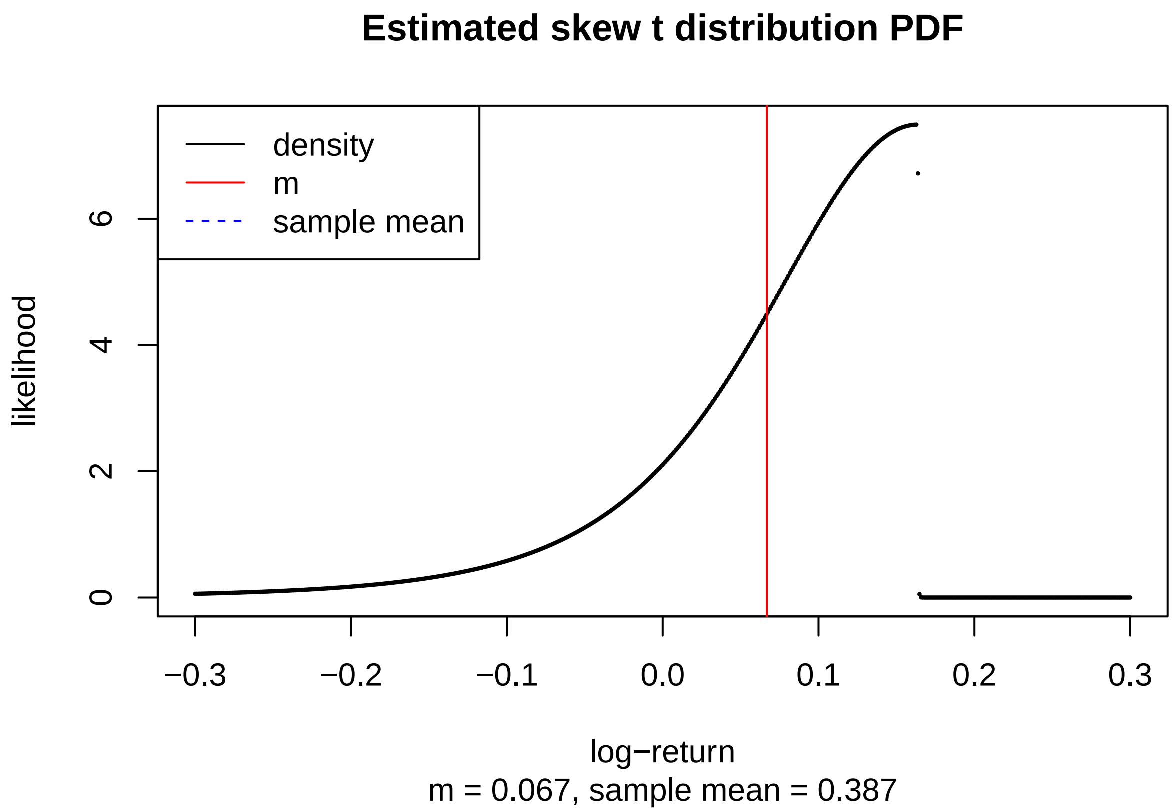
Estimated distribution

Now lets look at the CDF of the estimated distribution for each 0.1% increment between 0.5% and 99.5% for the estimated distribution:

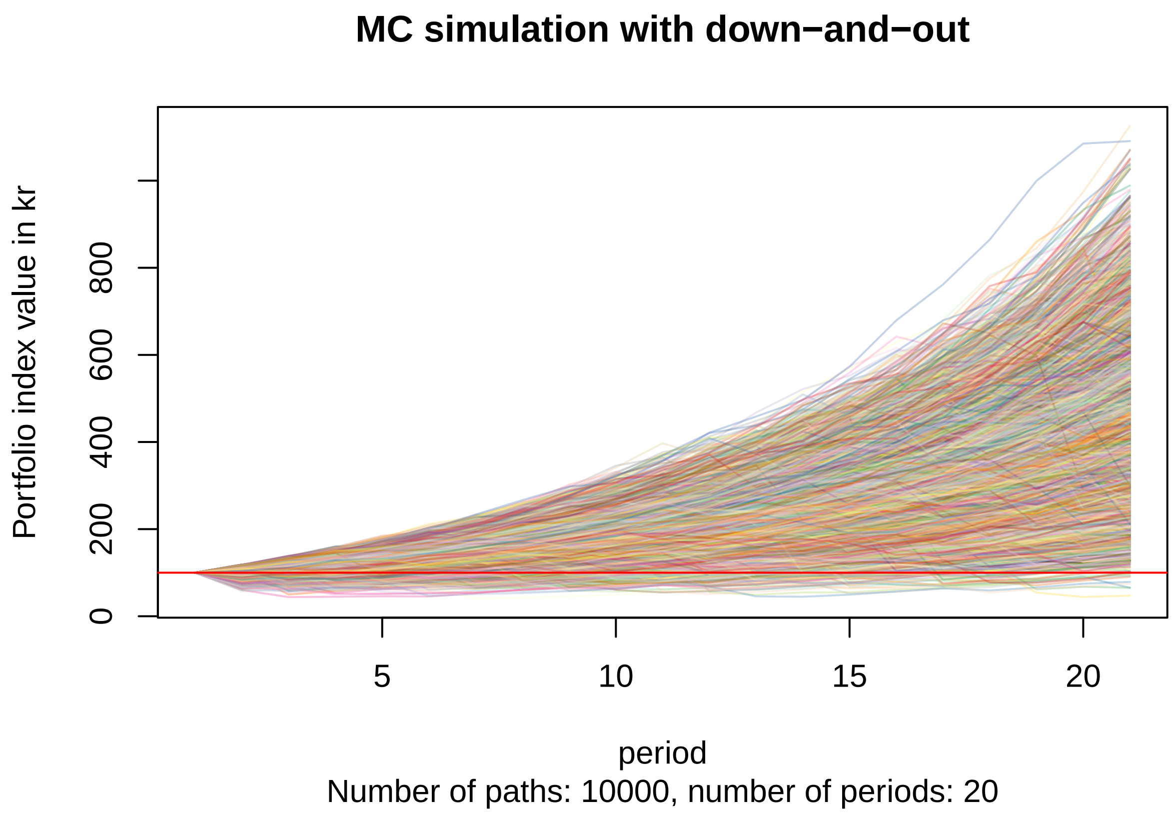

Estimated skew t distribution CDF
probability
0.0
0.2
0.4
0.6
0.8
1.0
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
log−returns

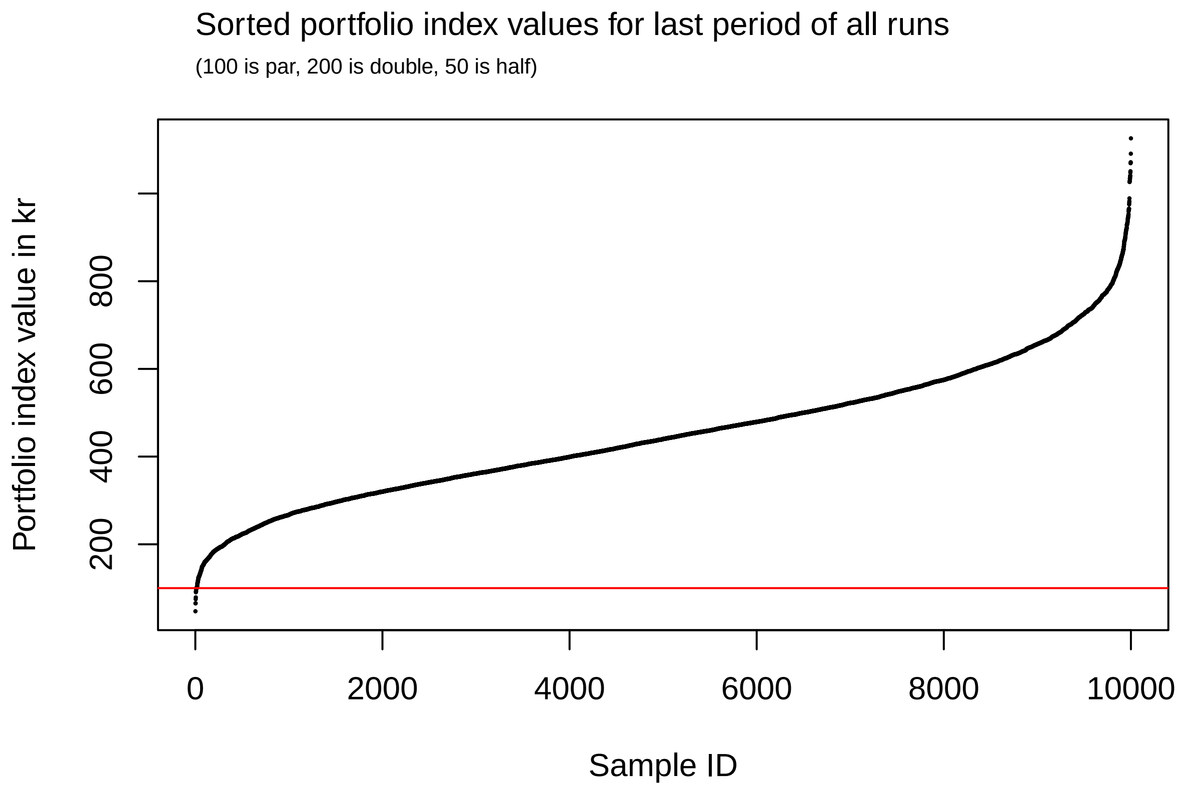


Estimated skew t distribution quantiles
log−returns
0.1
0.0
−0.1
−0.2
−0.3
0.0
0.2
0.4
0.6
0.8
1.0
probability



Estimated skew t distribution PDF
density
m
sample mean
likelihood
0
2
4
6
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
log−return
m = 0.067, sample mean = 0.387

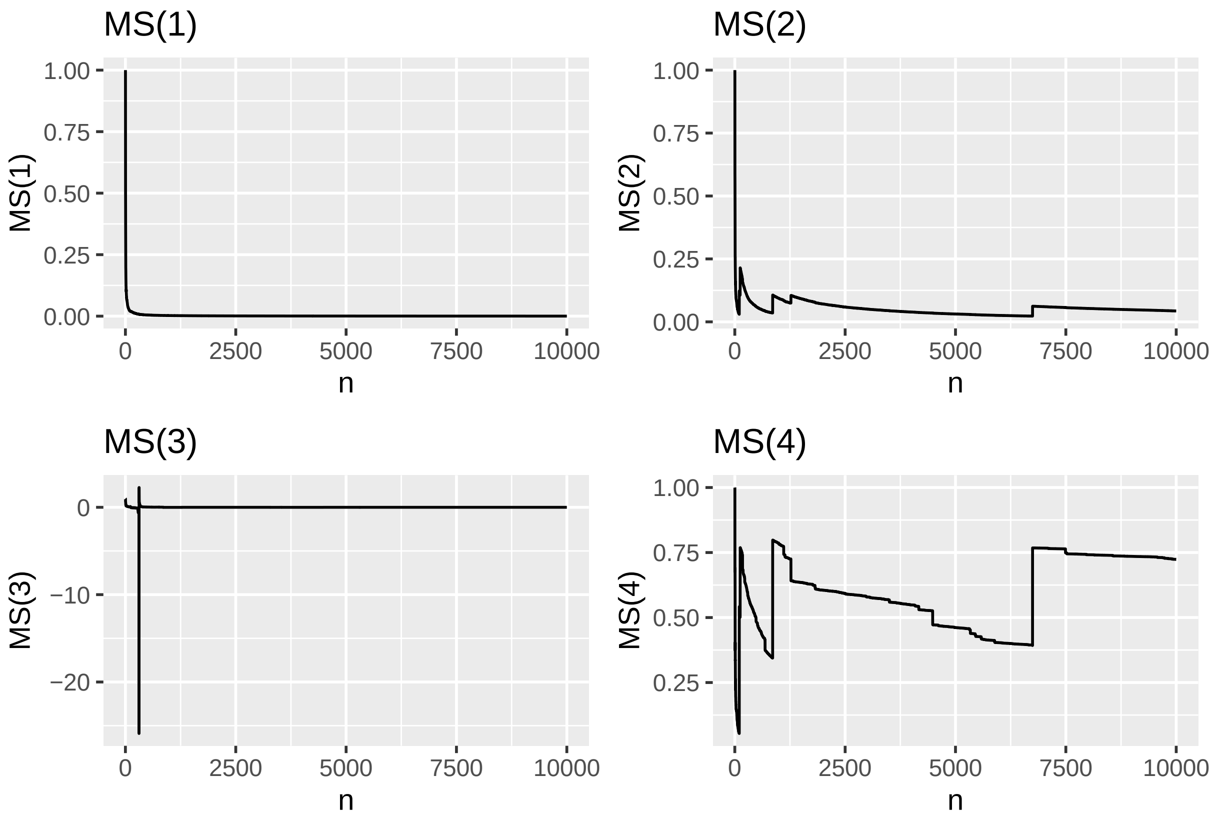

MC simulation with down−and−out
Portfolio index value in kr
0
200
400
600
800
5
10
15
20
period
Number of paths: 10000, number of periods: 20

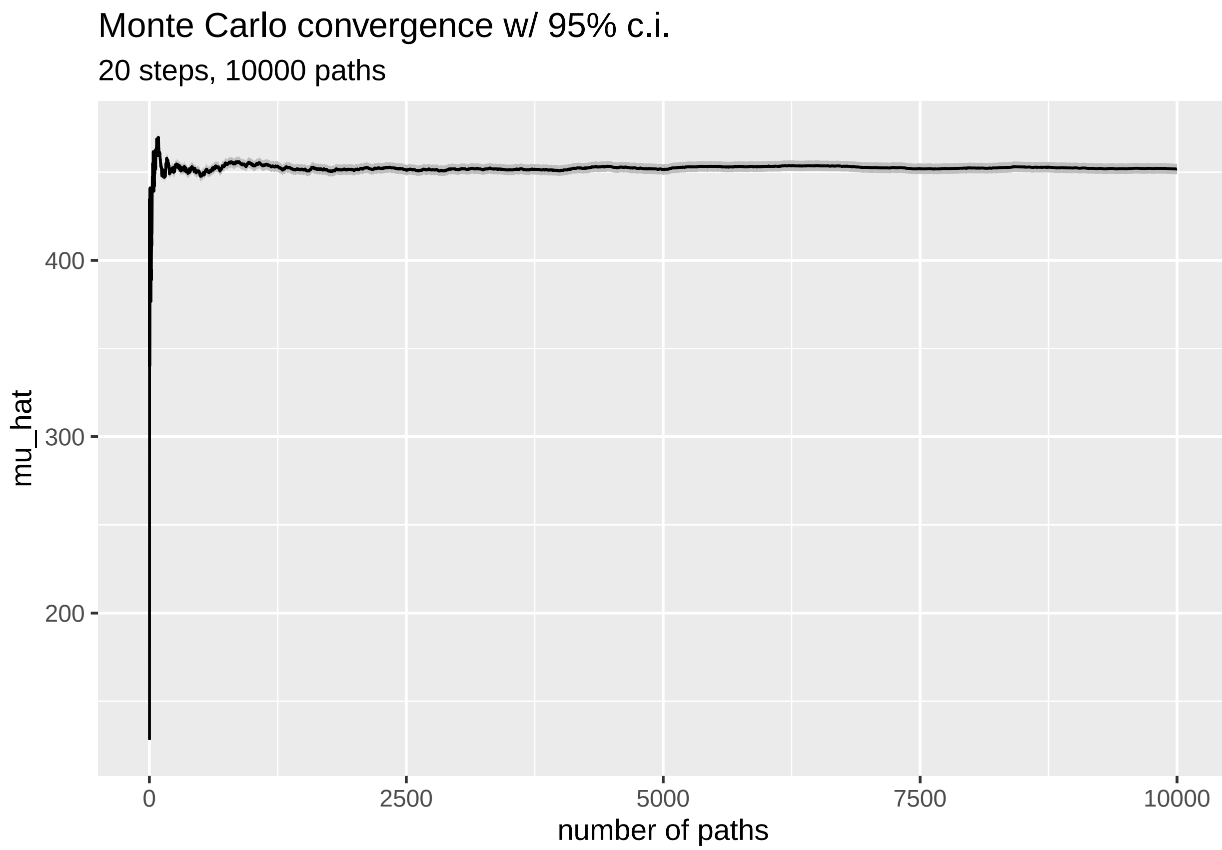


Convergence

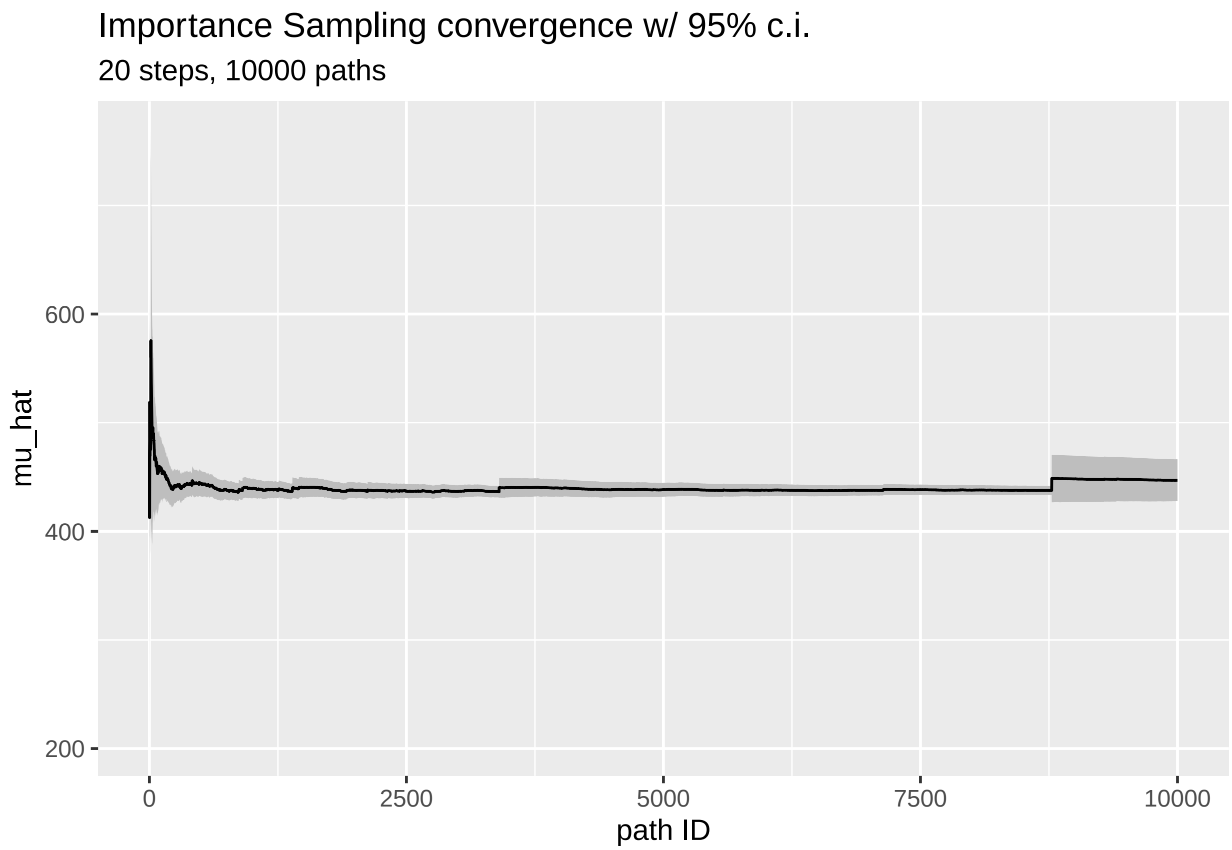
Max vs sum

Max vs sum plots for the first four moments:


MS(1)
MS(2)
MS(3)
MS(4)
1.00
0.75
0.50
0.25
0.00
0
−10
−20
0
2500
5000
7500
10000
n


Monte Carlo convergence w/ 95% c.i.
20 steps, 10000 paths
mu_hat
400
300
200
0
2500
5000
7500
10000
number of paths

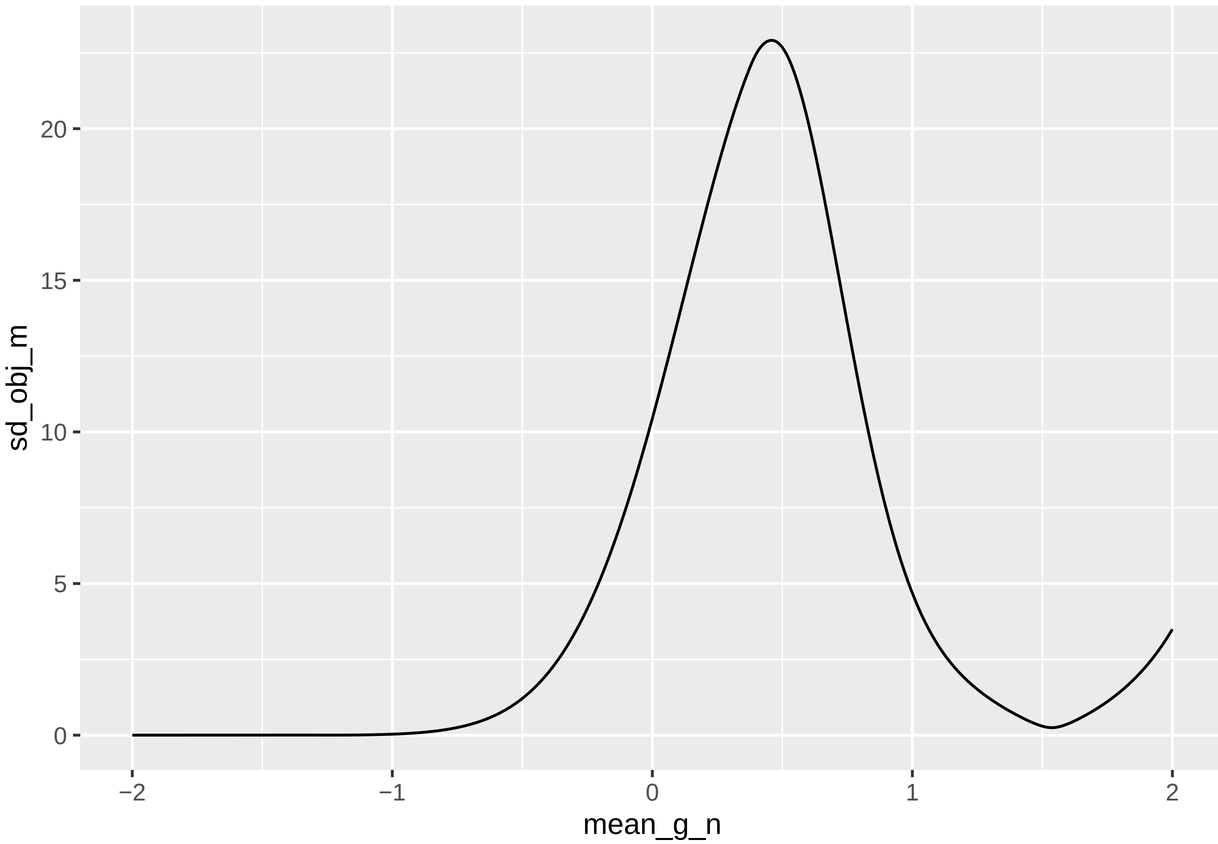
IS



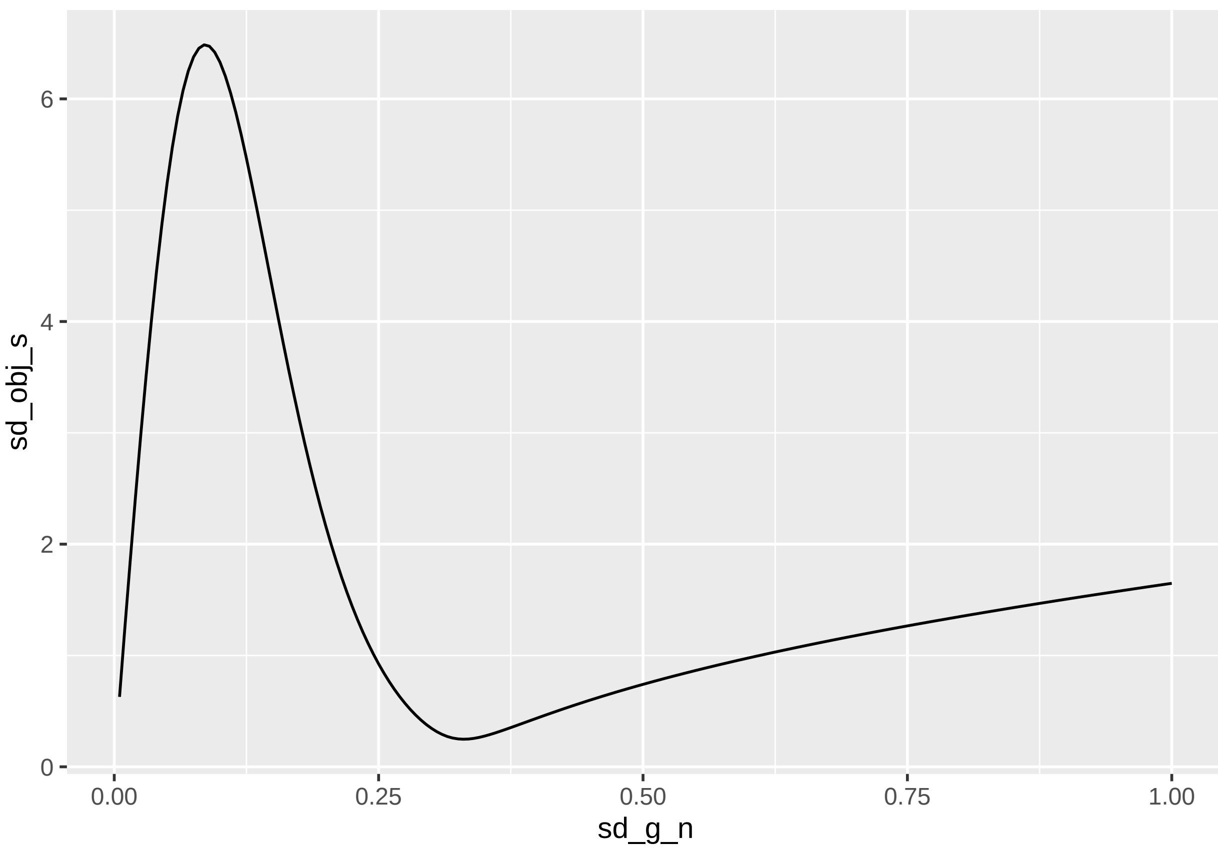
Parameters

```
## [1] 1.5363616 0.3304634
```

Objective function plots



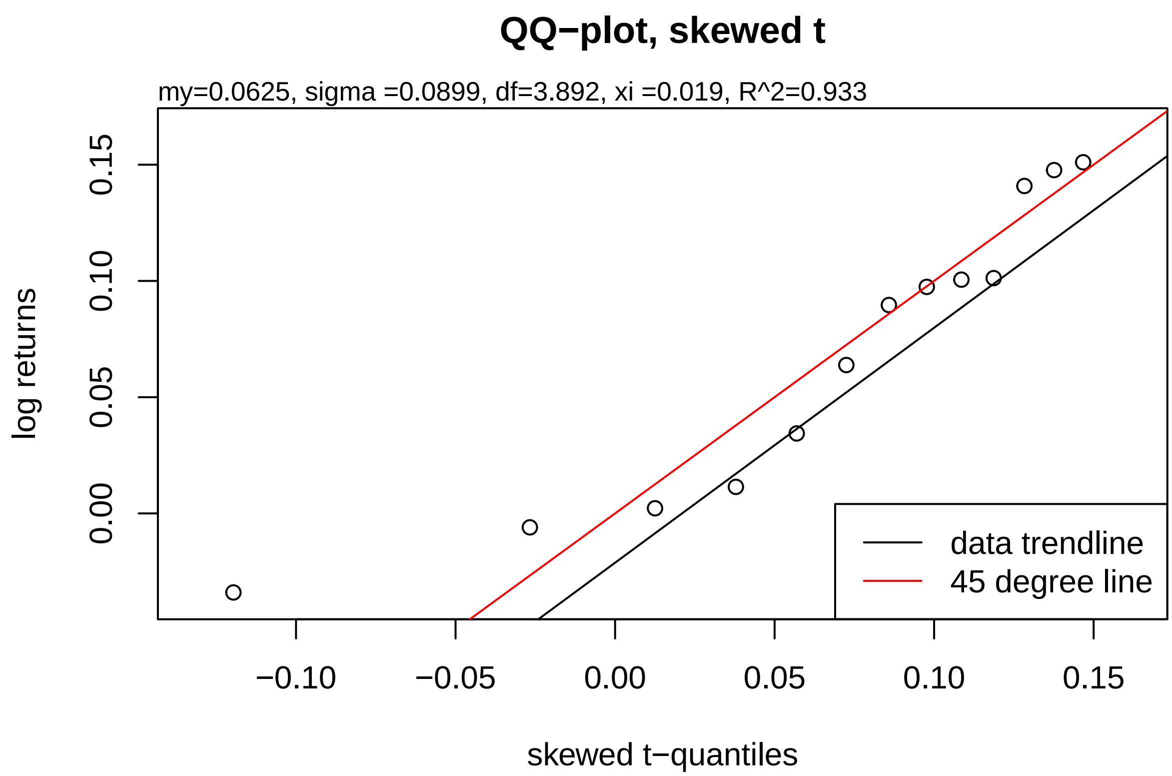
sd_obj_m
0
5
10
15
20
−2
−1
0
1
2
mean_g_n



6
4
2
0
0.00
0.25
0.50
0.75
1.00
sd_obj_s
sd_g_n

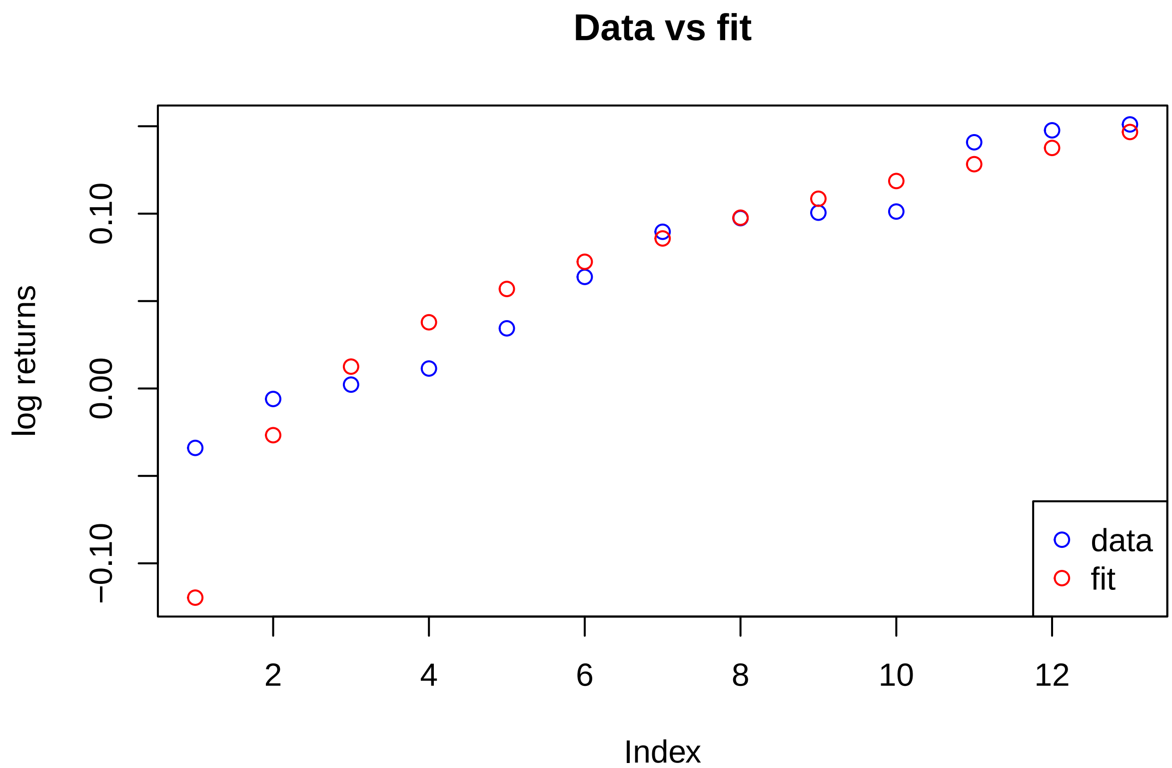
Mix vhr+pmr (mh_pm), 2011 - 2023

QQ Plot



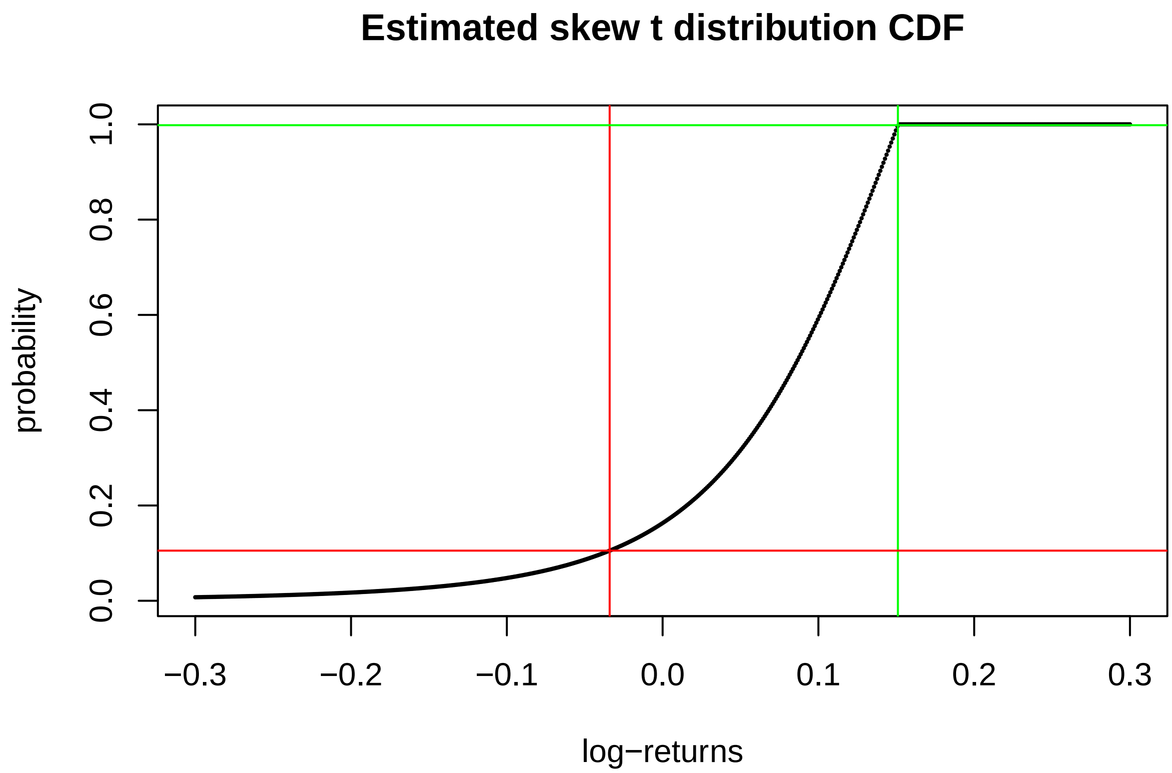
Data vs fit

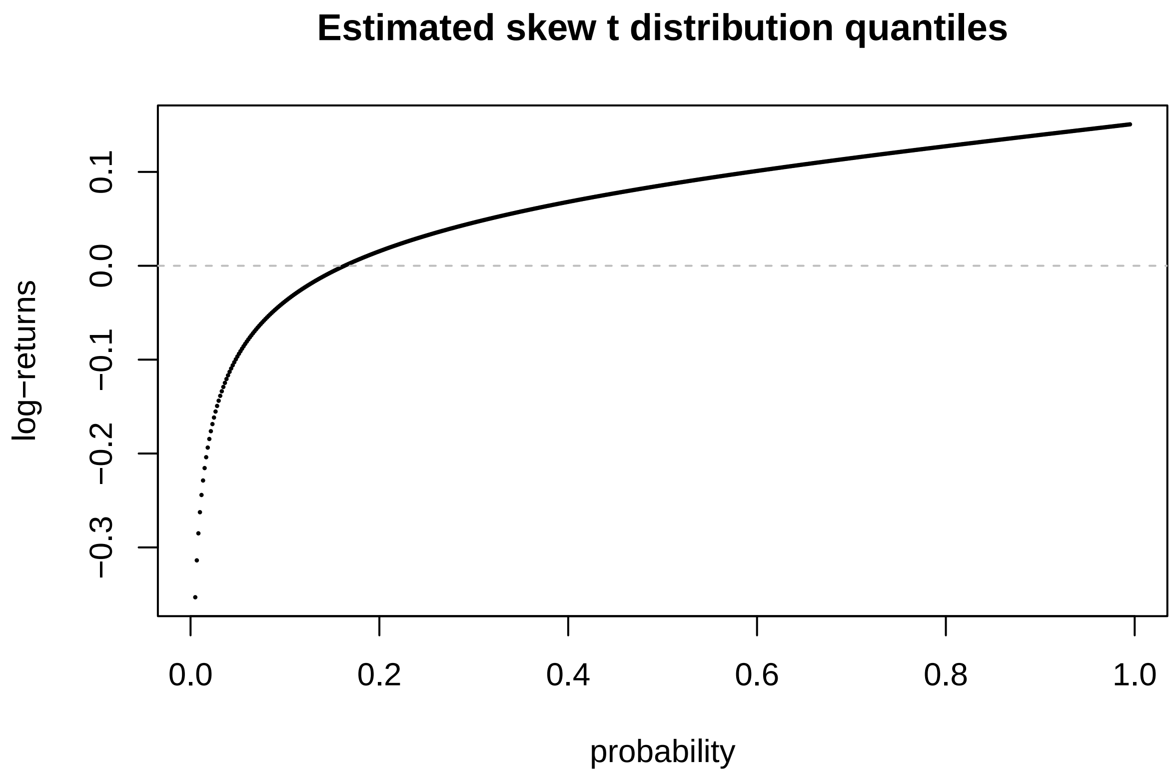
Let's plot the fit and the observed returns together.



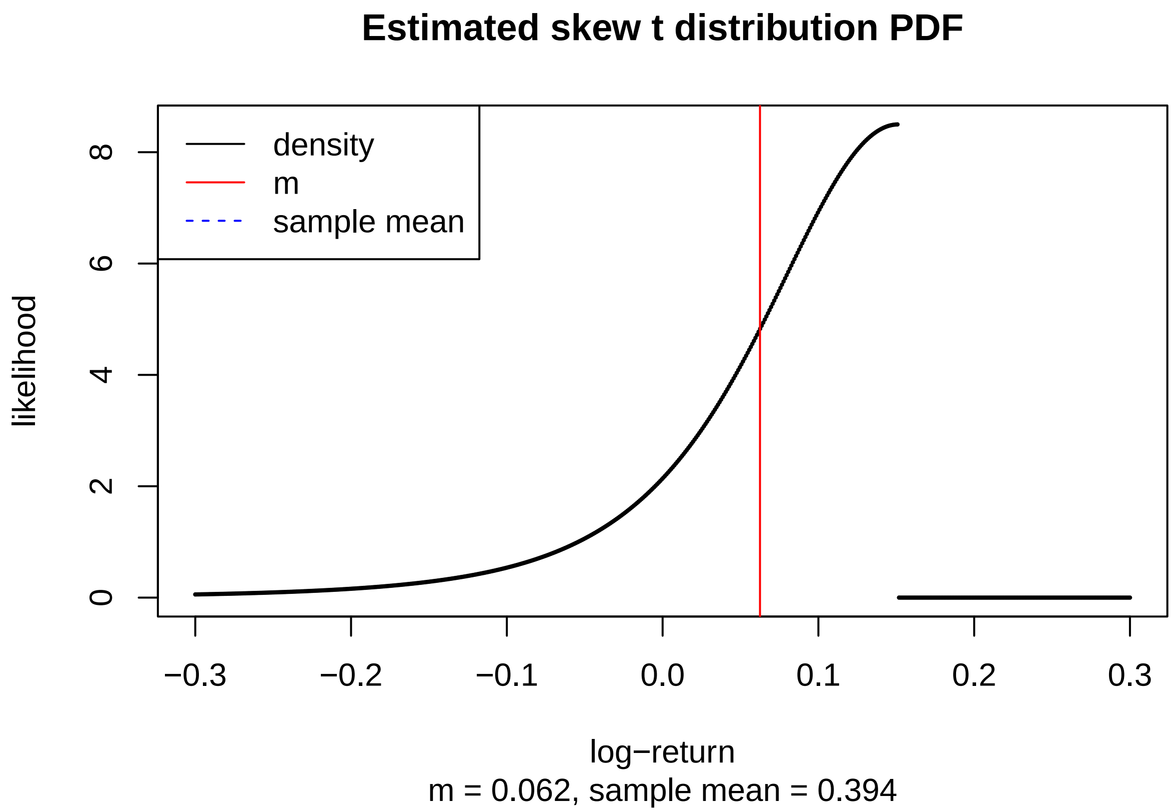
Estimated distribution

Now lets look at the CDF of the estimated distribution for each 0.1% increment between 0.5% and 99.5% for the estimated distribution:

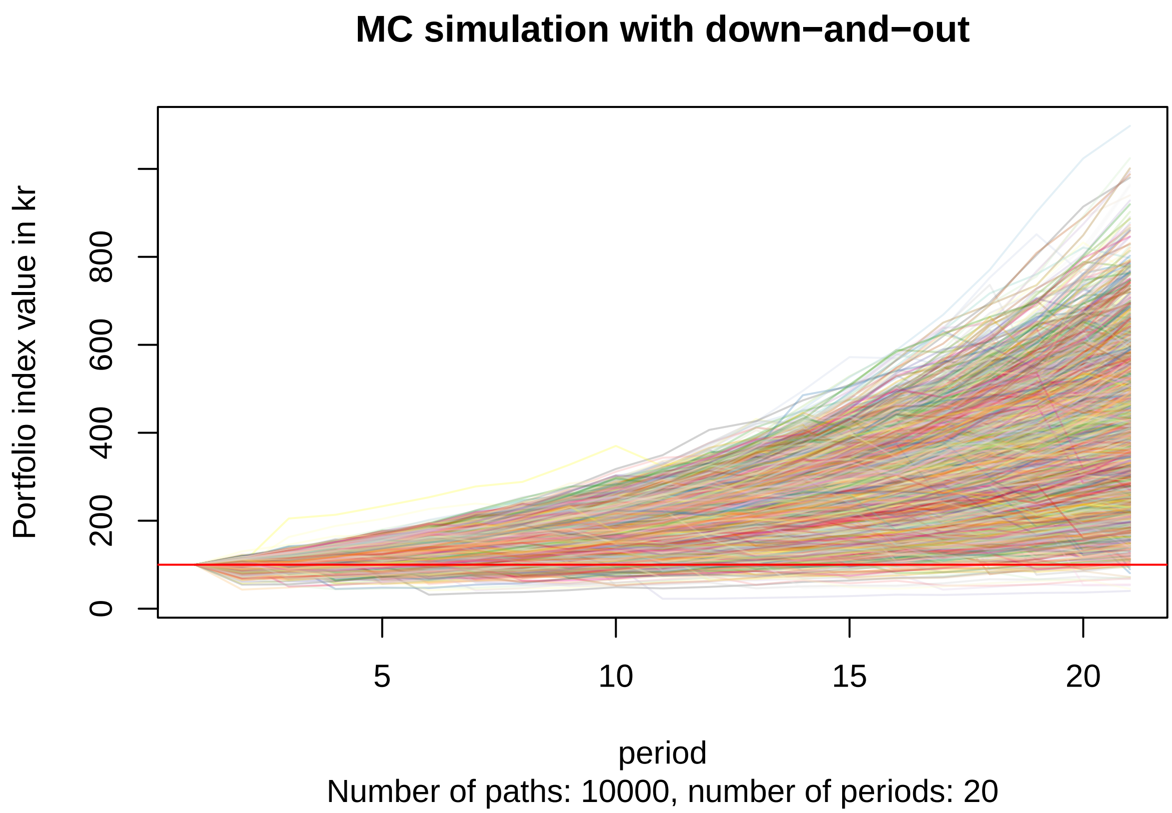

Estimated skew t distribution CDF
probability
1.0
0.8
0.6
0.4
0.2
0.0
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
log−returns

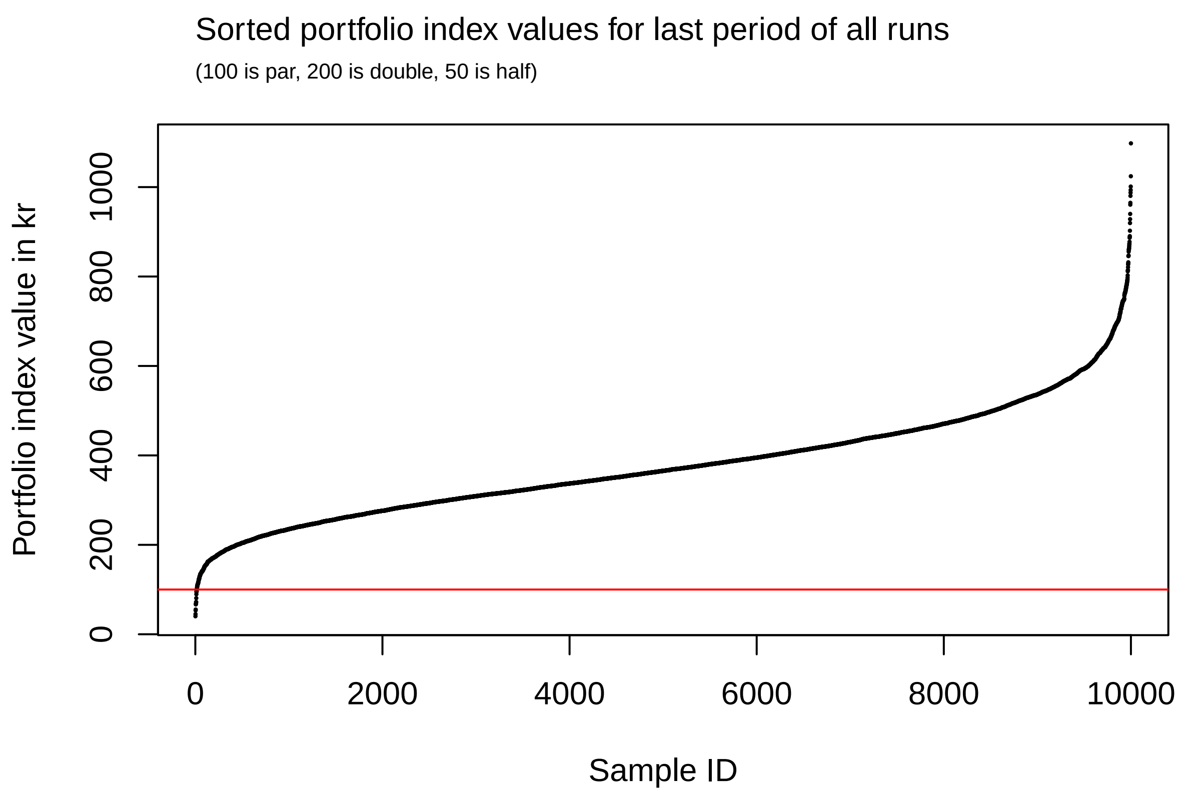


Estimated skew t distribution quantiles
log−returns
0.1
0.0
−0.1
−0.2
−0.3
0.0
0.2
0.4
0.6
0.8
1.0
probability



Estimated skew t distribution PDF
density
m
sample mean
likelihood
0
2
4
6
8
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
log−return
m = 0.062, sample mean = 0.394

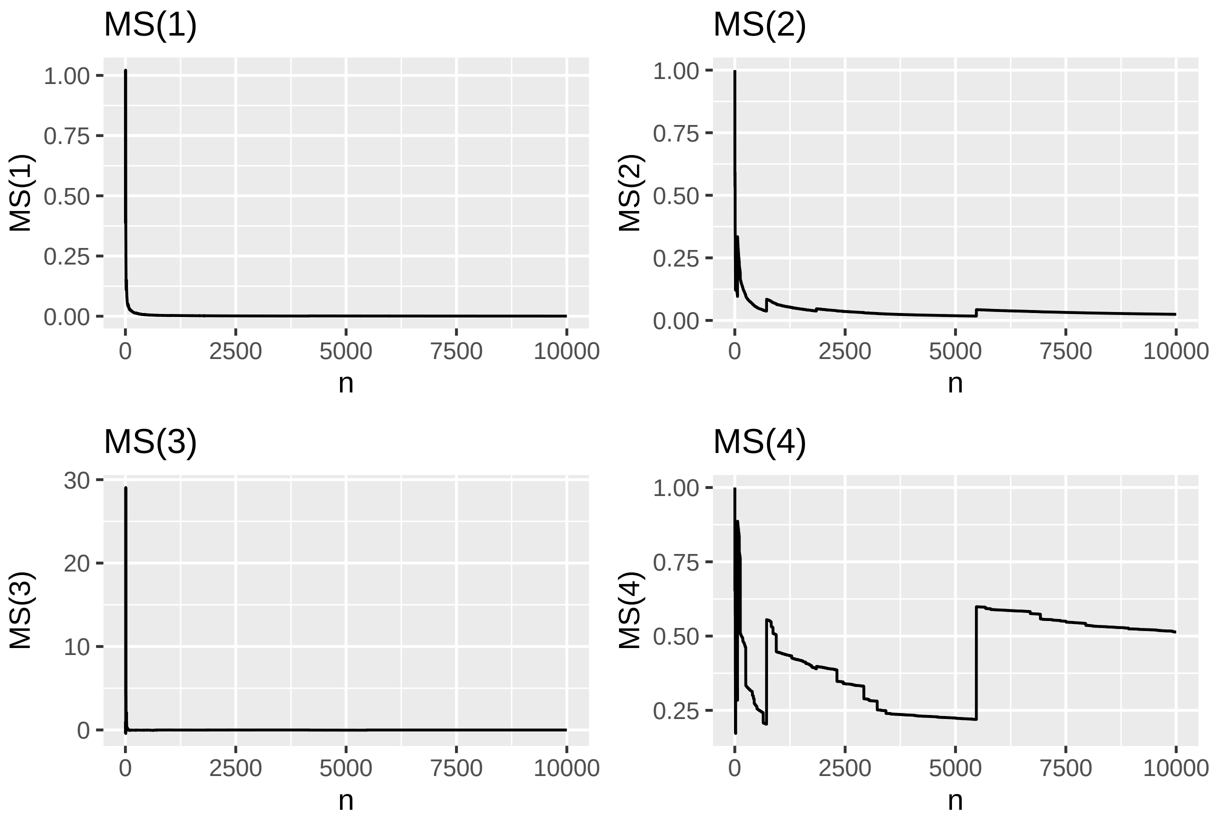

MC simulation with down−and−out
Portfolio index value in kr
0
200
400
600
800
5
10
15
20
period
Number of paths: 10000, number of periods: 20

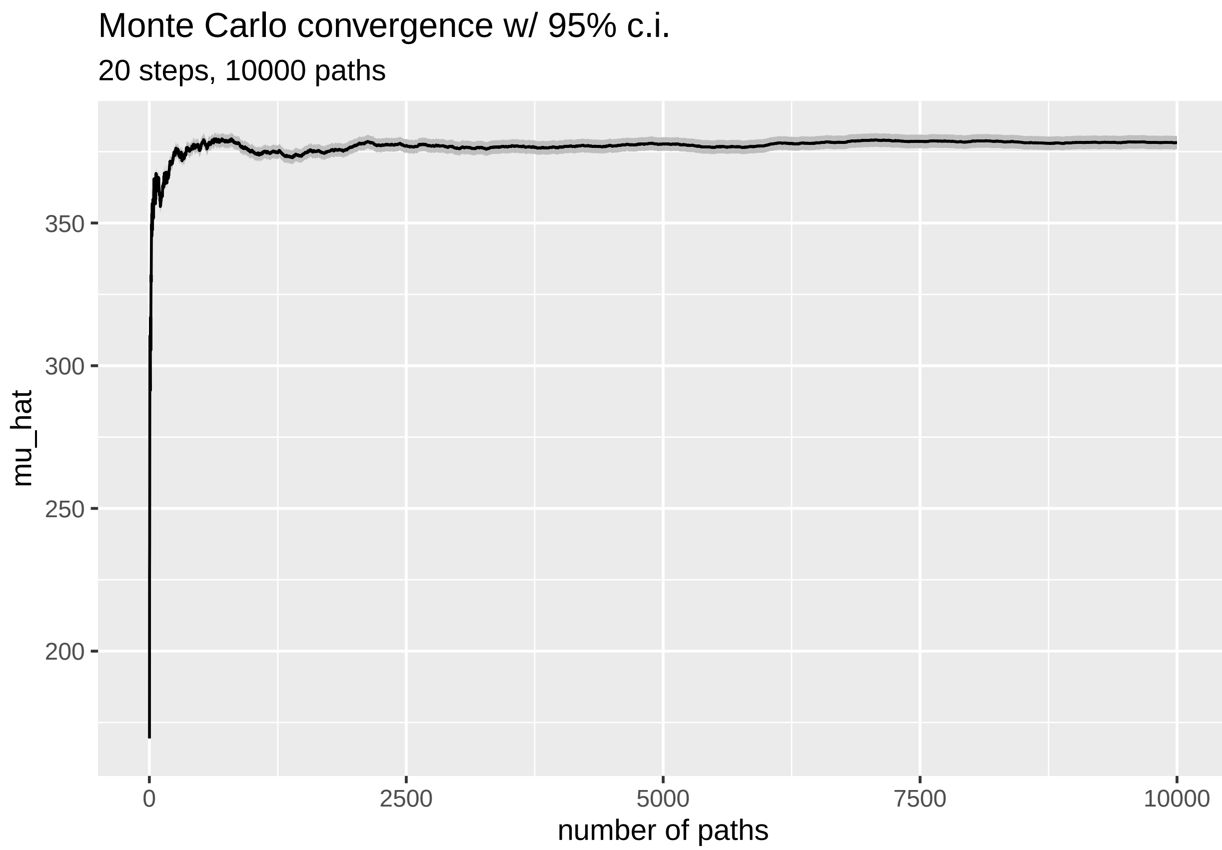


Convergence

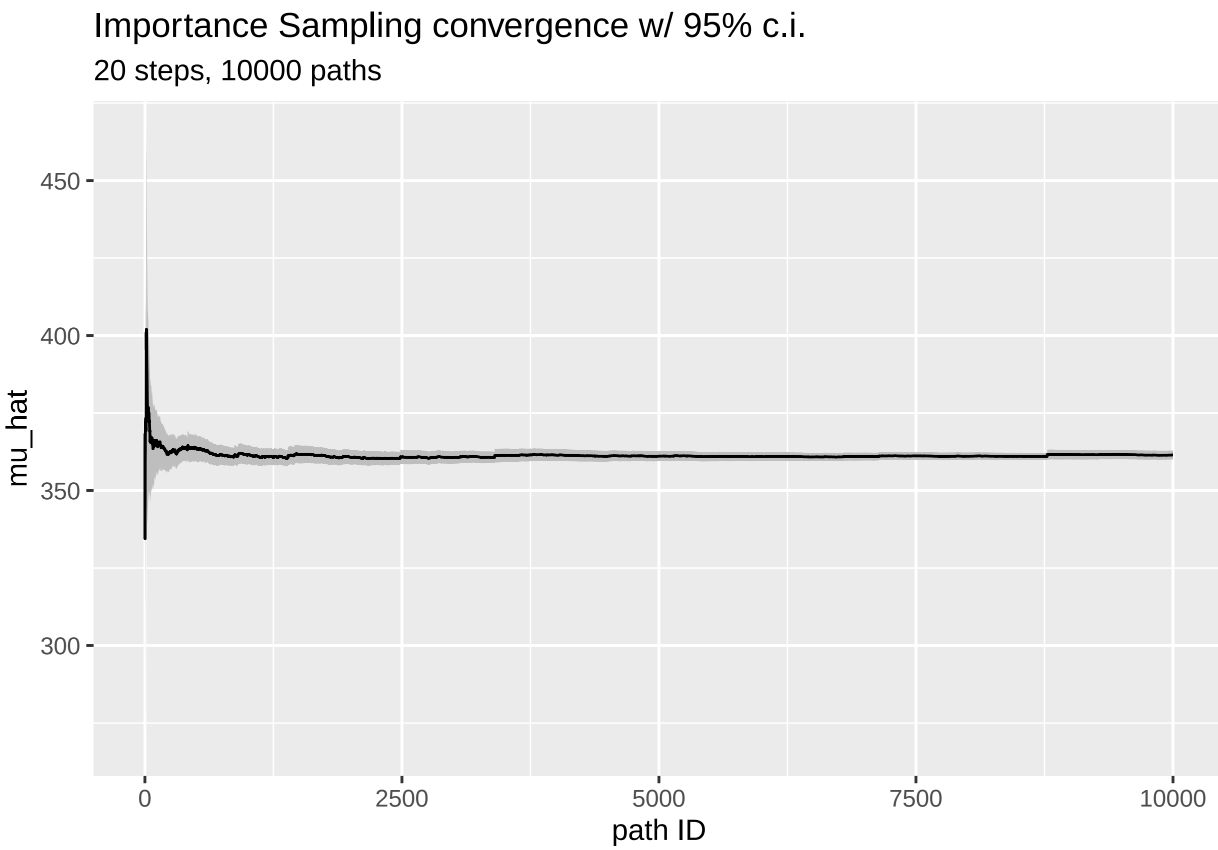
Max vs sum

Max vs sum plots for the first four moments:


MS(1)
1.00
0.75
0.50
0.25
0.00
0
2500
5000
7500
10000
n
MS(2)
1.00
0.75
0.50
0.25
0.00
0
2500
5000
7500
10000
n
MS(3)
30
20
10
0
0
2500
5000
7500
10000
n
MS(4)
1.00
0.75
0.50
0.25
0
2500
5000
7500
10000
n


Monte Carlo convergence w/ 95% c.i.
20 steps, 10000 paths
mu_hat
350
300
250
200
0
2500
5000
7500
10000
number of paths

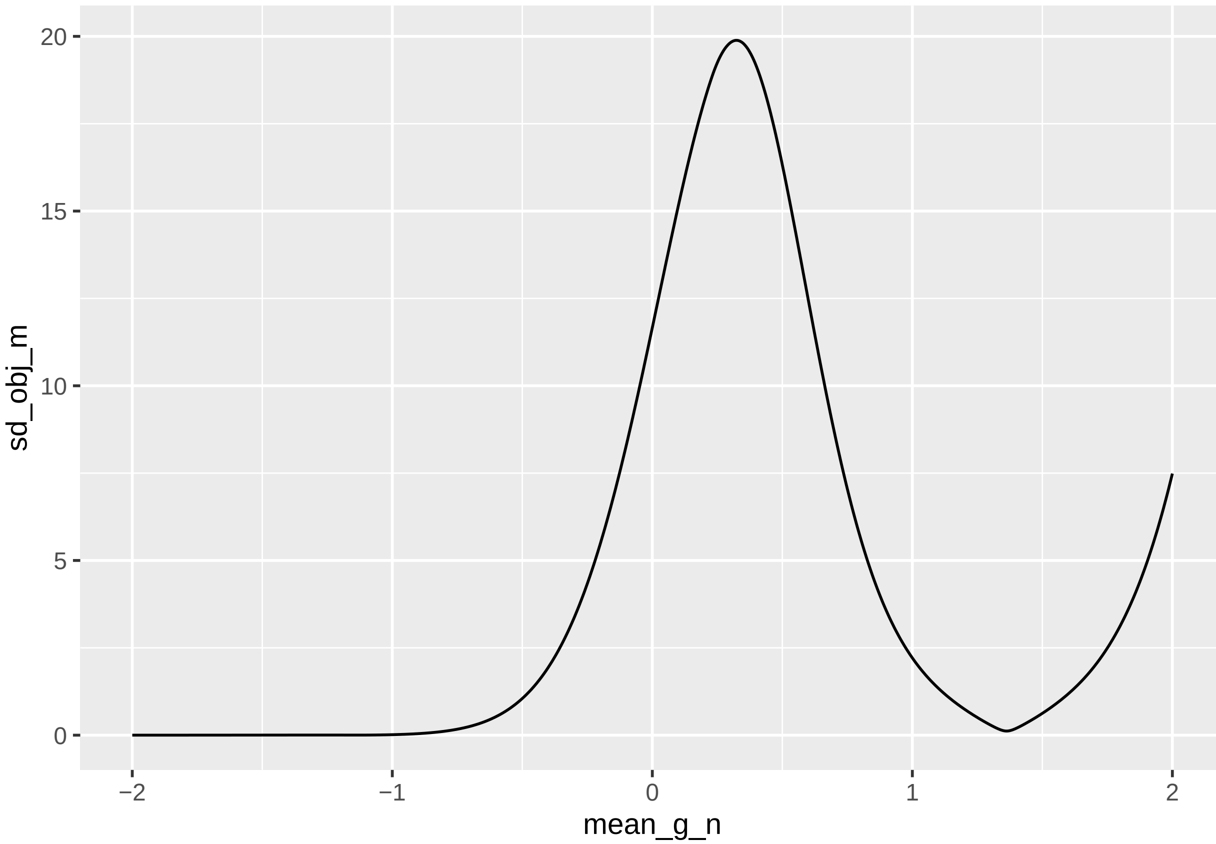
IS



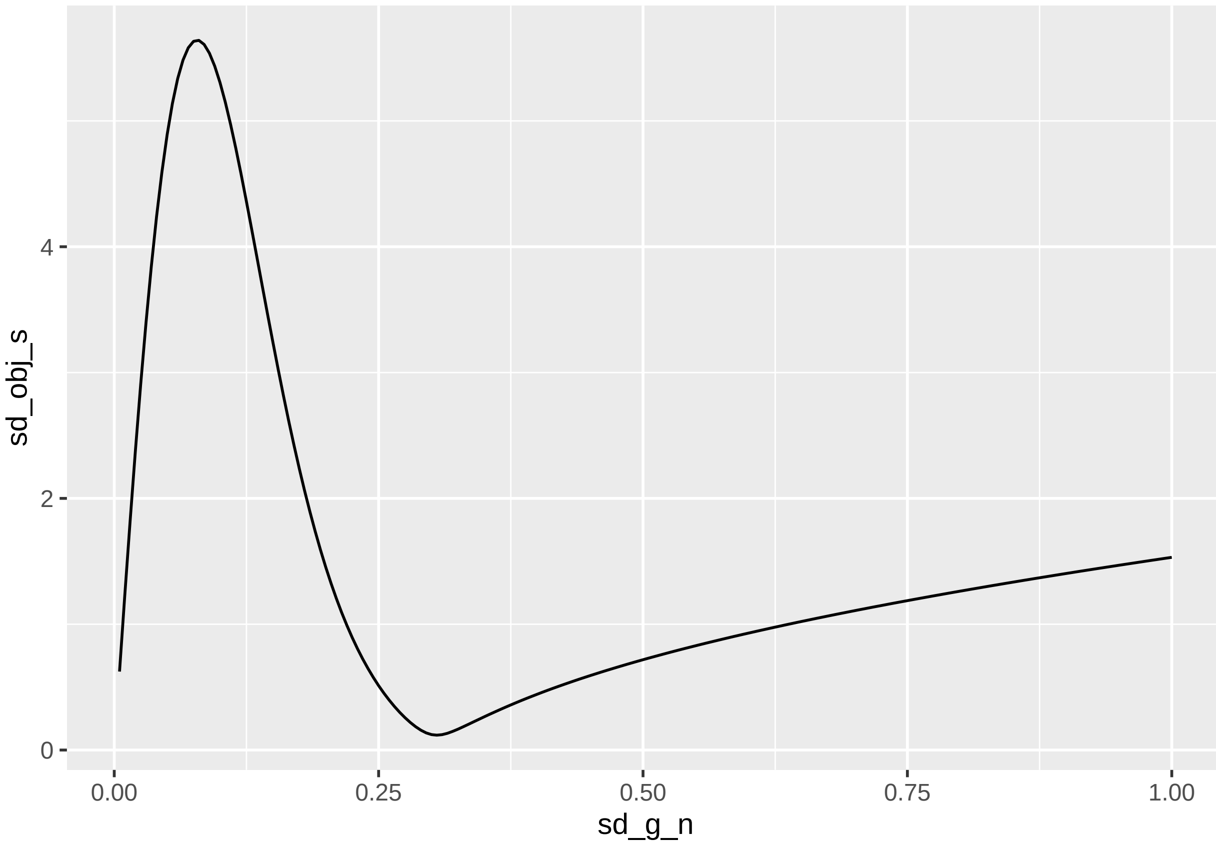
Parameters

```
## [1] 1.3625460 0.3050122
```

Objective function plots



20
15
10
5
0
sd_obj_m
−2
−1
0
1
2
mean_g_n



Comments

(Ignoring `mhr_a`…)

`mhr` has some nice properties:

- It has a relatively high `nu` value of 90, which means it is tending more towards exponential tails than polynomial tails. All other funds have `nu` values close to 3, except `phr` which is even worse at close to 2. (Note that for a Gaussian, `nu` is infinite.)
- It has the lowest losing percentage of all simulations, which is better than 1/6 that of `phr`.
- It has a DAO percentage of 0, which is the same as `mmr`, and less than `phr`.
- Only `phr` has a higher `mc_m`.
- It has a smaller `mc_s` than the individual components, `vhr` and `phr`.
- It has the highest `xi` of all fits, suggesting less left skewness. Density plots for `vmr`, `phr` and `mmr` have an extremely sharp drop, as if an upward limiter has been applied, which corresponds to extremely low `xi` values. The density plot for `mhr` is by far the most symmetrical of all the fits. As seen in the section "Compare Gaussian and skewed t-distribution fits", the other skewed t-distribution fits don't capture the max observed returns at all.
- Only `mmr` has as higher `mc_min`. However, that of `mmr` is 18 times higher with 62, so `mmr` is a clear winner here.
- Naturally, it has a `mc_max` smaller than the individual components, `vhr` and `phr`, but ca. 1.5 times higher then `mmr`.
- All the first 4 moments converge nicely. For all other fits, the 4th moment doesn't seem to converge.

Taleb, Statistical Consequences Of Fat Tails, p. 97:
"the variance of a finite variance random variable with tail exponent < 4 will be infinite".

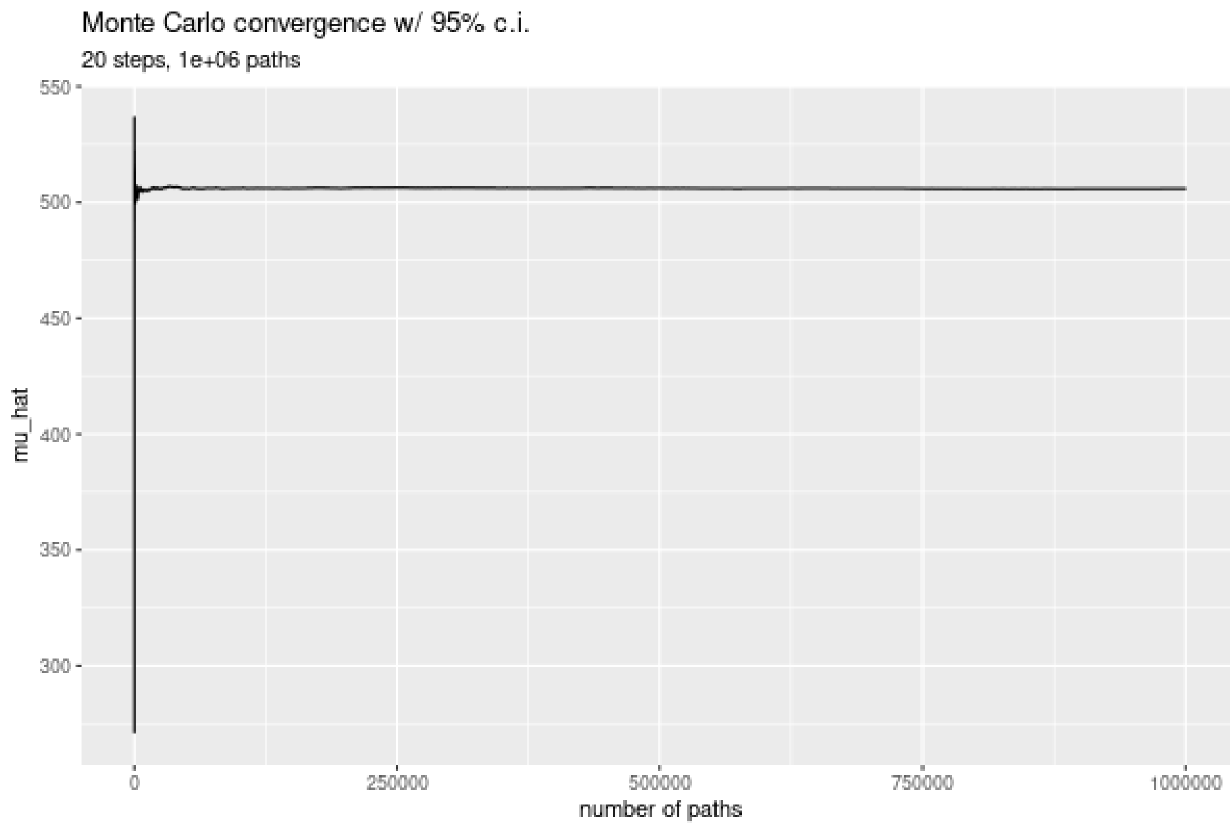
And p. 363:
"The hedging errors for an option portfolio (under a daily revision regime) over 3000 days, un- der a constant volatility Student T with tail exponent $\alpha = 3$. Technically the errors should not converge in finite time as their distribution has infinite variance."

- Note: QQ lines by design pass through 1st and 3rd quantiles. They are not trendlines in the sense of linear regression.

Appendix

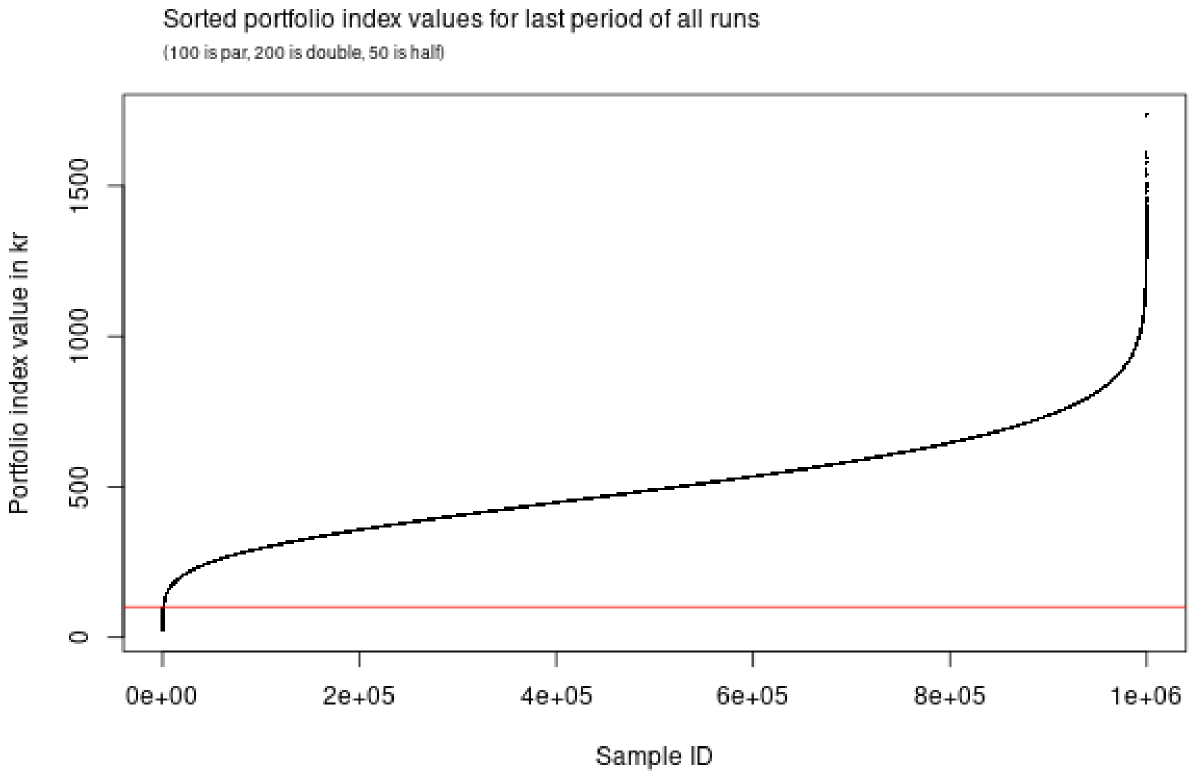
Many simulations of mc_mhr: `num_paths = 1e6`

1e6 paths:



Compare 10^6 and 10^4 paths for mhr:

	mc_m	mc_s	mc_min	mc_max	dao_pct	dai_pct
mc_mhr_1e6	505.90695	173.22176	21.09569	1734.83520	0.00000	0.07330
mc_mhr_1e4	504.75125	173.80504	26.81367	1414.94530	0.00000	0.10000
is_mhr_1e4	510.836	2331.167	205.398	232384.846	ibid.	ibid.



Arithmetic vs geometric mean

Let m be the number of steps in each path and n be the number of paths. a is the initial capital. Use arithmetic mean for mean of all paths at time t:

$$\frac{a(e^{z_1} + e^{z_2} + \cdots + e^{z_n})}{n}$$

where

$$z_i := x_{i,1} + x_{i,2} + \cdots + x_{i,m}$$

Use geometric mean for mean of all steps in a single path i:

$$ae^{\frac{x_{i,1}+x_{i,2}+\cdots+x_{i,m}}{m}} = a\sqrt[m]{e^{x_{i,1}+x_{i,2}+\cdots+x_{i,m}}}$$

So for **Monte Carlo** of returns after m periods, we

- fit a skewed t-distribution to log-returns and use that distribution to simulate $\{x_{i,j}\}_j^m$,
- for each path i, calculate $100 \cdot e^{z_i}$,
- calculate the mean of $\{z_i\}_i^n$:
  - $$\bar{z} = 100\frac{e^{z_1} + e^{z_2} + \cdots + e^{z_n}}{n}$$

For **Importance Sampling**, we

- model log-returns on a skewed t-distribution,
- for each path i, calculate $100 \cdot e^{z_i}$,
- fit a skewed t-distribution to $\{z_i\}_i^n$ and use it as our f density function from which we simulate $\{h_i\}_i^n$,
  - In our case h and z are identical, because we have an idea for a distribution to simulate z, but in general for IS h could be a function of z.
- calculate $w* = \frac{f}{g^*}$, where $g*$ is our proposal distribution, which minimizes the variance of $h \cdot w$.
- calculate the arithmetic mean of $\{h_i w_i^*\}_i^n$:
  - $$100\frac{e^{h_1 w_1^*} + e^{h_2 w_2^*} + \cdots + e^{h_n w_n^*}}{n}$$

Average of returns vs returns of average

Math

$$\text{Avg. of returns} := \frac{\left(\frac{x_t}{x_{t-1}} + \frac{y_t}{y_{t-1}}\right)}{2}$$

$$\text{Returns of avg.} := \left(\frac{x_t + y_t}{2}\right) \Big/ \left(\frac{x_{t-1} + y_{t-1}}{2}\right) \equiv \frac{x_t + y_t}{x_{t-1} + y_{t-1}}$$

For which x_1 and y_1 are Avg. of returns $=$ Returns of avg.?

$$\frac{\left(\frac{x_t}{x_{t-1}} + \frac{y_t}{y_{t-1}}\right)}{2} = \frac{x_t + y_t}{x_{t-1} + y_{t-1}}$$

$$\frac{x_t}{x_{t-1}} + \frac{y_t}{y_{t-1}} = 2\frac{x_t + y_t}{x_{t-1} + y_{t-1}}$$

$$(x_{t-1} + y_{t-1})x_t y_{t-1} + (x_{t-1} + y_{t-1})x_{t-1} y_t = 2(x_{t-1} y_{t-1} x_t + x_{t-1} y_{t-1} y_t)$$

$$(x_{t-1} x_1 y_{t-1} + y_{t-1} x_t y_{t-1}) + (x_{t-1} x_{t-1} y_t + x_{t-1} y_{t-1} y_t) = 2(x_{t-1} y_{t-1} x_t + x_{t-1} y_{t-1} y_t)$$

This is not generally true, but true if for instance $x_{t-1} = y_{t-1}$.

Example

Definition: `R = 1+r`

```
## Let x_0 be 100.

## Let y_0 be 200.

## So the initial value of the pf is 300 .

## Let R_x be 0.5.

## Let R_y be 1.5.
```

Then,

```
## x_1 is R_x * x_0 = 50.

## y_1 is R_y * y_0 = 300.
```

Average of returns:

```
## 0.5 * (R_x + R_y) = 1
```

So here the value of the pf at t=1 should be unchanged from t=0:

```
## (x_0 + y_0) * 0.5 * (R_x + R_y) = 300
```

But this is clearly not the case:

```
## 0.5 * (x_1 + y_1) = 0.5 * (R_x * x_0 + R_y * y_0) = 175
```

Therefore we should take returns of average, not average of returns!

Let’s take the average of log returns instead:

```
## 0.5 * (log(R_x) + log(R_y)) = -0.143841
```

We now get:

```
## (x_0 + y_0) * exp(0.5 * (log(Rx) + log(Ry))) = 259.8076
```

So taking the average of log returns doesn't work either.

Simulation of mix vs mix of simulations

Test if a simulation of a mix (average) of two returns series has the same distribution as a mix of two simulated returns series.

```
## m(data_x): 0.005386465
## s(data_x): 0.3568481
## m(data_y): 9.876111
## s(data_y): 2.57217
##
## m(data_x + data_y): 4.940749
## s(data_x + data_y): 1.335916
```

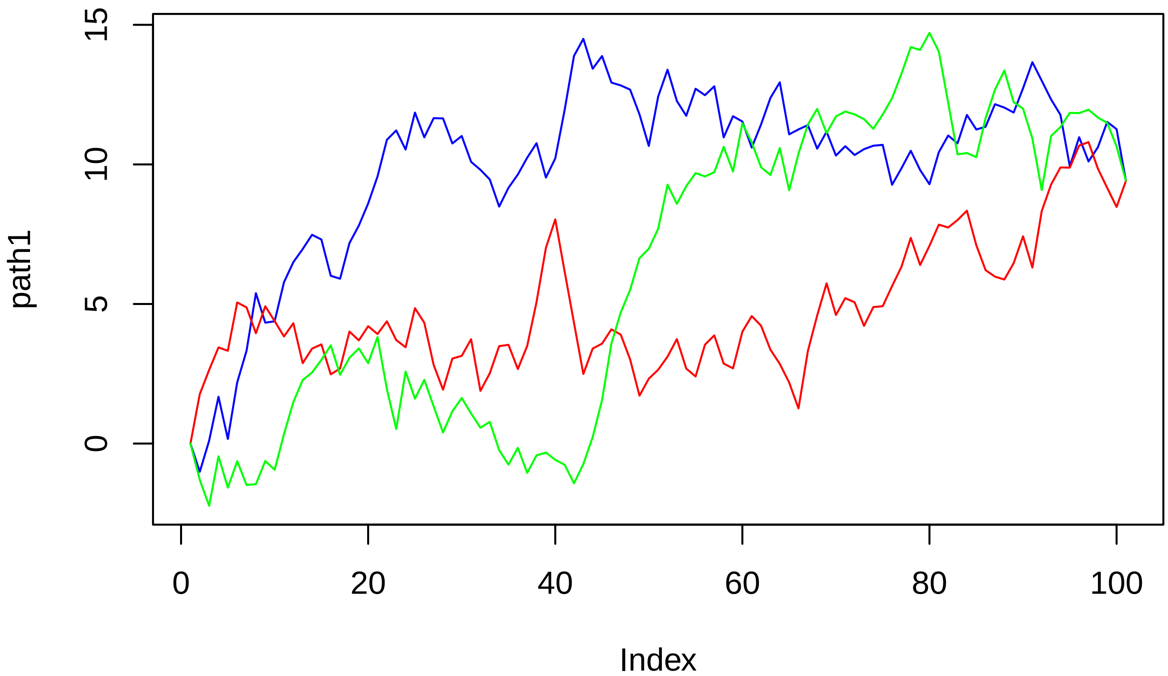
m and s of final state of all paths.
`_a` is mix of simulated returns.
`_b` is simulated mixed returns.

m_a	m_b	s_a	s_b
98.886	98.693	5.580	6.091
98.720	98.830	5.627	5.949
98.988	98.486	5.790	6.115
98.778	98.980	5.694	5.719
98.599	99.013	5.630	5.849
98.718	98.529	5.797	5.813
98.612	98.812	5.833	5.909
98.885	98.839	5.951	5.741
98.909	98.743	5.924	6.083
98.947	98.877	5.998	6.057

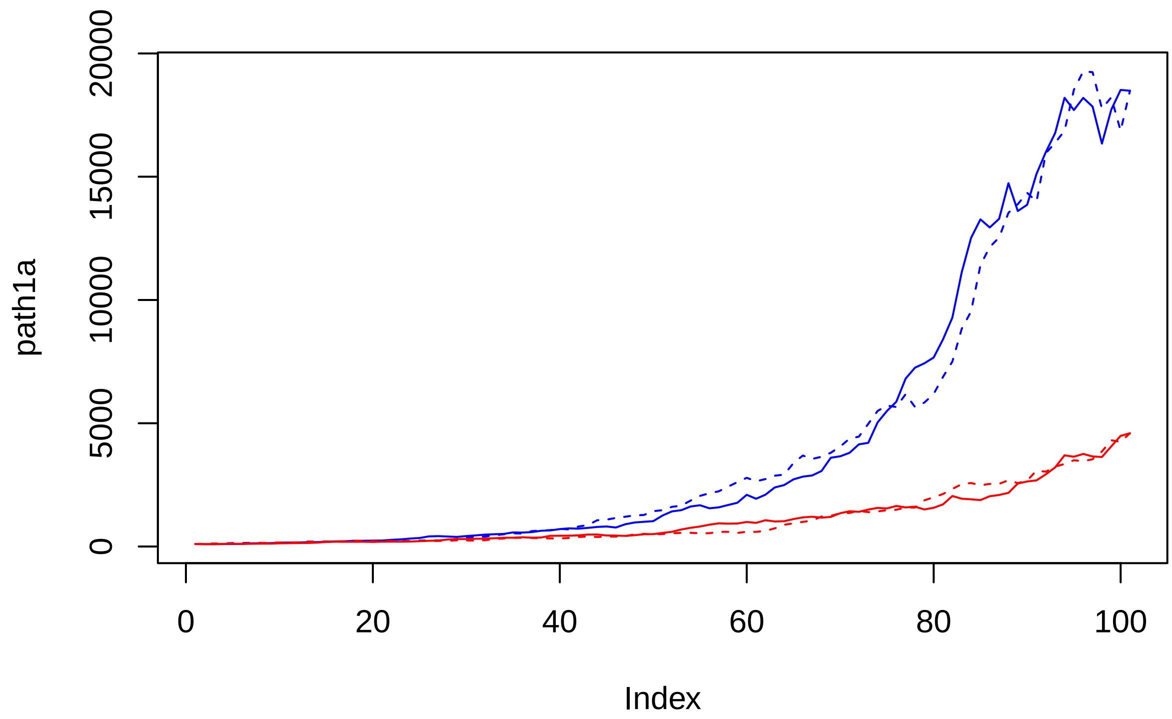
```
##       m_a            m_b             s_a             s_b
##  Min.   :98.60   Min.   :98.49   Min.   :5.580   Min.   :5.719
##  1st Qu.:98.72   1st Qu.:98.71   1st Qu.:5.646   1st Qu.:5.822
##  Median :98.83   Median :98.82   Median :5.793   Median :5.929
##  Mean   :98.80   Mean   :98.78   Mean   :5.782   Mean   :5.933
##  3rd Qu.:98.90   3rd Qu.:98.87   3rd Qu.:5.901   3rd Qu.:6.076
##  Max.   :98.99   Max.   :99.01   Max.   :5.998   Max.   :6.115
```

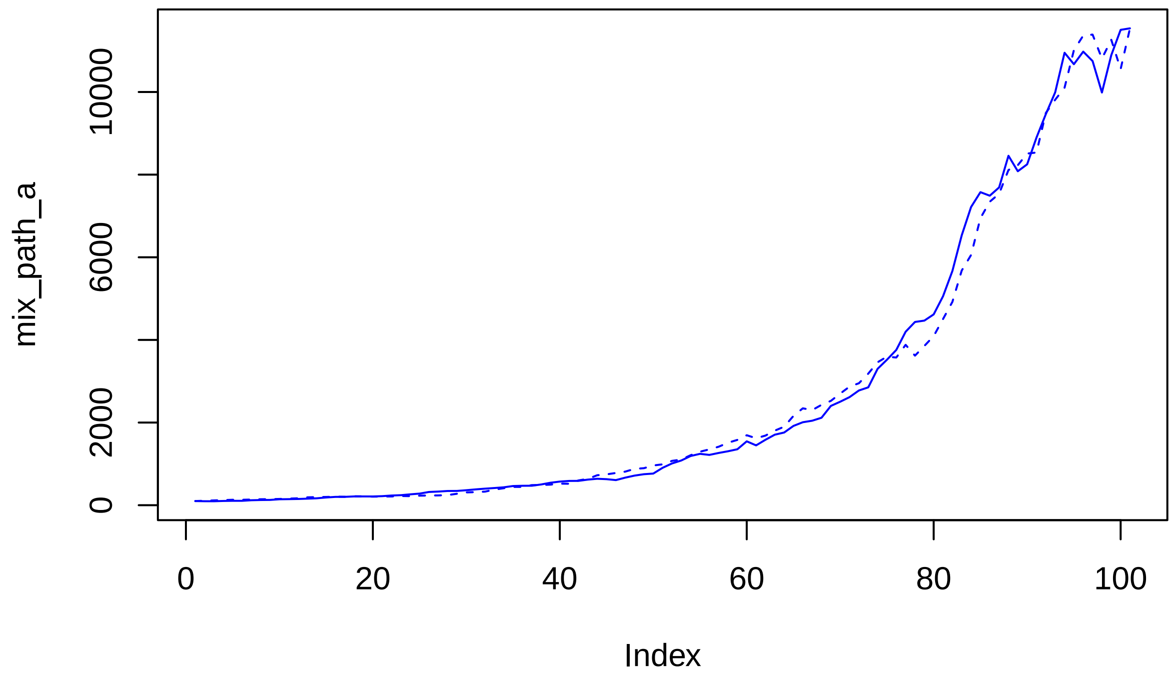
`_a` and `_b` are very close to equal.
We attribute the differences to differences in estimating the distributions in version a and b.

The final state is independent of the order of the preceding steps:



So does the order of the steps in the two processes matter, when mixing simulated returns?


path1a
0
5000
10000
15000
20000
0
20
40
60
80
100
Index



The order of steps in the individual paths do not matter, because the mix of simulated paths is a sum of a sum, so the order of terms doesn't affect the sum. If there is variation it is because the sets preceding steps are not the same. For instance, the steps between step 1 and 60 in the plot above are not the same for the two lines.

Recall,

$$\text{Var}(aX + bY) = a^2\text{Var}(X) + b^2\text{Var}(Y) + 2ab\text{Cov}(a, b)$$

```
var(0.5 * vhr + 0.5 * phr)
```

```
## [1] 0.005355618
```

```
0.5^2 * var(vhr) + 0.5^2 * var(phr) + 2 * 0.5 * 0.5 * cov(vhr, phr)
```

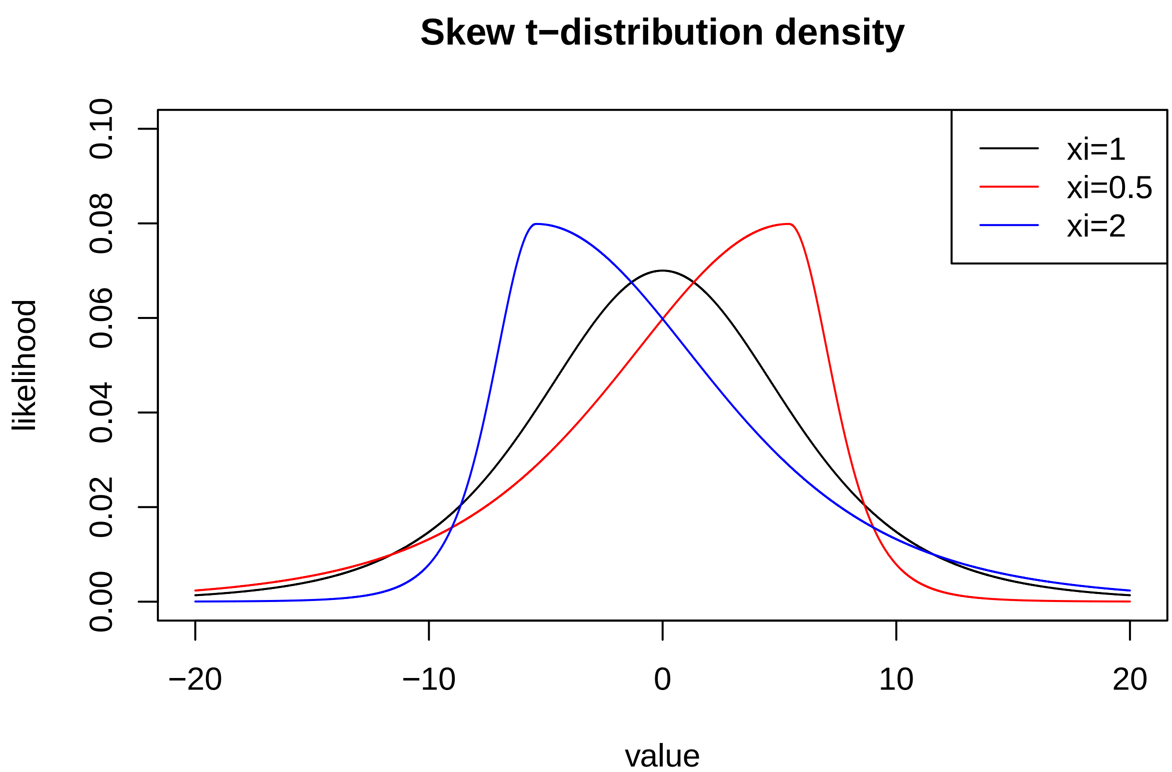
```
## [1] 0.005355618
```

Our distribution estimate is based on 13 observations. Is that enough for a robust estimate? What if we suddenly hit a year like 2008? How would that affect our estimate?
Let's try to include the Velliv data from 2007-2010.
We do this by sampling 13 observations from `vmrl`.

```
##        m                  s
##  Min.   :0.06075   Min.   :0.04867
##  1st Qu.:0.06611   1st Qu.:0.05963
##  Median :0.06883   Median :0.06639
##  Mean   :0.07020   Mean   :0.06694
##  3rd Qu.:0.07330   3rd Qu.:0.07329
##  Max.   :0.08361   Max.   :0.08975
```

The meaning of `xi`

The fit for `mhr` has the highest `xi` value of all. This suggests right-skew:



Max vs sum plot

If the Law Of Large Numbers holds true,

$$\frac{\max(X_1^p, ..., X^p)}{\sum_{i=1}^{n} X_i^p} \to 0$$

for $n \to \infty$.

If not, X doesn't have a p'th moment.

See Taleb: The Statistical Consequences Of Fat Tails, p. 192