AMS-512 Capital Markets and Portfolio Theory

Extreme Events

Robert J. Frey, Research Professor
Stony Brook University, Applied Mathematics and Statistics

frey@ams.sunysb.edu
http://www.ams.sunysb.edu/~frey

A Look at Real Data

Downloading and processing the data

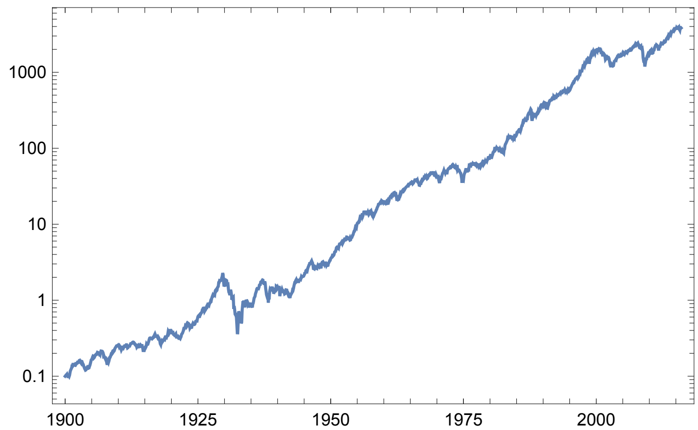
The Import[ ] assumes that the file is in the same directory as this notebook. It contains monthly index data for the S & P 500 Total Return Index from {1899, 12} to {2015, 12}. The term “total return” in this instance includes the reinvestment of dividends back into the index so that it properly represents the total wealth from investing in the index. Although this index did not exist over this period, it is a reconstruction of it by Global Financial Data.

In[1]:=
```
mxSP500Index = Import[FileNameJoin[{NotebookDirectory[], "mxSP500Index.m"}]];
Short[mxSP500Index, 7]
```

Out[•]//Short=
```
{{{1899, 12, 31}, 0.0997052}, {{1900, 1, 31}, 0.101342},
 {{1900, 2, 28}, 0.103484}, {{1900, 3, 31}, 0.104631}, {{1900, 4, 30}, 0.106294},
 <<1383>>, {{2015, 8, 31}, 3660.75}, {{2015, 9, 30}, 3570.17},
 {{2015, 10, 30}, 3871.33}, {{2015, 11, 30}, 3882.84}, {{2015, 12, 31}, 3821.6}}
```

In[3]:=
```
DateListLogPlot[mxSP500Index]
```

Out[∘]=

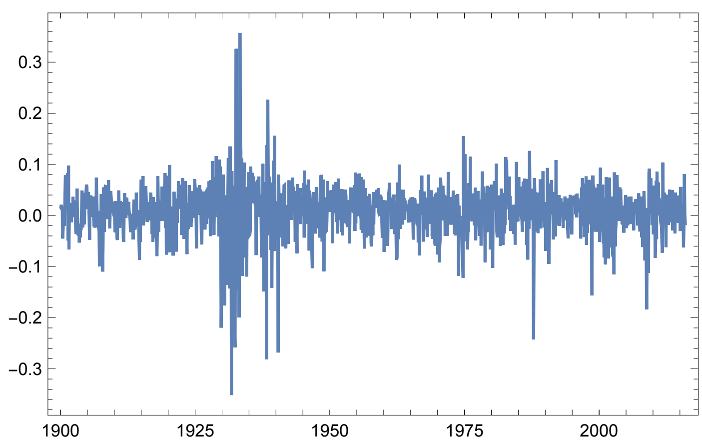


We can compute the log returns of the index by taking the first order differences of the log of the index. Note that we also have to drop the first date.

In[4]:=
```
mxSP500LogReturn =
  Transpose[{Rest[mxSP500Index⟦All, 1⟧], Differences@Log[mxSP500Index⟦All, 2⟧]}];
```

In[5]:=
```
DateListPlot[mxSP500LogReturn, Joined → True, PlotRange → All]
```

Out[∘]=



In[6]:=
```
Dimensions[mxSP500LogReturn]
```

Out[∘]= {1392, 2}

In[7]:=
```
Export[FileNameJoin[{ParentDirectory[NotebookDirectory[]],
   "Workshop1", "mxSP500LogReturn.m"}], mxSP500LogReturn]
```

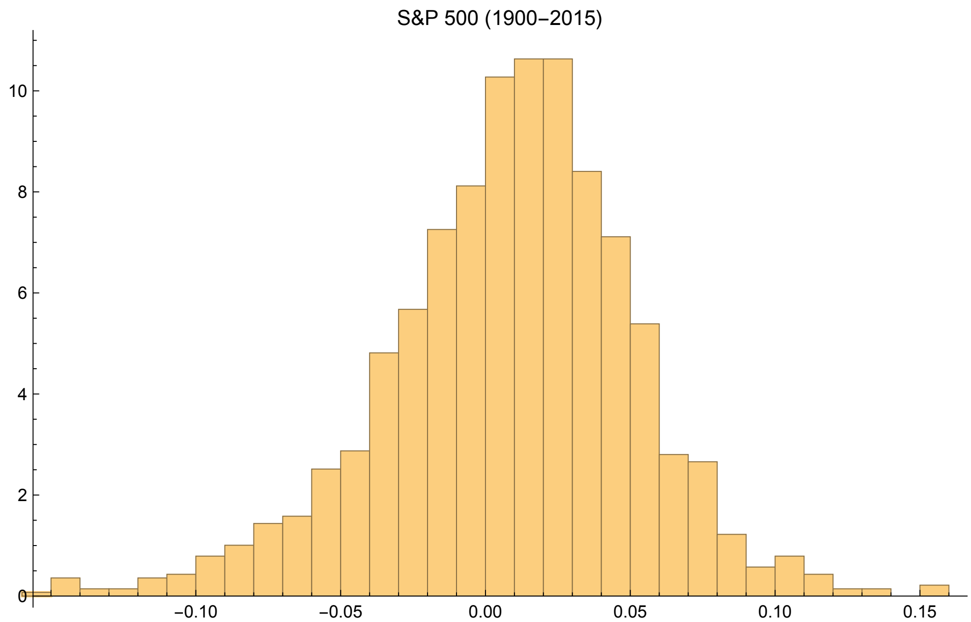
Out[∘]= /Users/robertjfrey/Documents/Work/Stony Brook University/AMS/QF/CourseWork/AMS-512/Modules/Workshop1/mxSP500LogReturn.m

Characterizing the Return Distribution

Fitting Candidate Distributions

In[8]:=
```
Histogram[mxSP500LogReturn[[All, 2]], Automatic,
  "PDF", ImageSize → 500, PlotLabel → "S&P 500 (1900-2015)"]
```

Out[•]=



In[17]:=
```
Through[{Mean, StandardDeviation, Skewness, Kurtosis}[mxSP500LogReturn[[All, 2]]]]
```

Out[•]=
```
{0.00758187, 0.0508859, -0.478152, 11.2052}
```

In[18]:=
```
distSP500N =
  EstimatedDistribution[mxSP500LogReturn[[All, 2]], NormalDistribution[m, s]]
```

Out[•]=
```
NormalDistribution[0.00758187, 0.0508676]
```

In[19]:=
```
distSP500T =
  EstimatedDistribution[mxSP500LogReturn[[All, 2]], StudentTDistribution[m, s, d]]
```

Out[•]=
```
StudentTDistribution[0.0103985, 0.0347629, 3.83284]
```

The computation below takes considerable time and has been commented out for classroom presentation. The result was saved in a prior run and is Import[]'ed below.

In[12]:=
```
(* distSP500ST=EstimatedDistribution[mxSP500LogReturn[[All,2]],
   TransformedDistribution[s x+m,x≈NoncentralStudentTDistribution[n,b]],
   ParameterEstimator→{"MaximumLikelihood",Method→"NMaximize"}]
Export[FileNameJoin[{NotebookDirectory[],"distSP500SwT.m"}],distSP500ST] *)
```

In[20]:=
```
distSP500ST = Import[FileNameJoin[{NotebookDirectory[], "distSP500ST.m"}]]
```

Out[ ]=
```
TransformedDistribution[0.0321032 - 0.0341353 x,
 x ≈ NoncentralStudentTDistribution[3.90016, 0.577027]]
```

In[22]:=
```
LogLikelihood[#, mxSP500LogReturn[[All, 2]]] & /@ {distSP500N, distSP500T, distSP500ST}
```

Out[ ]=
```
{2170.95, 2318.41, 2328.55}
```

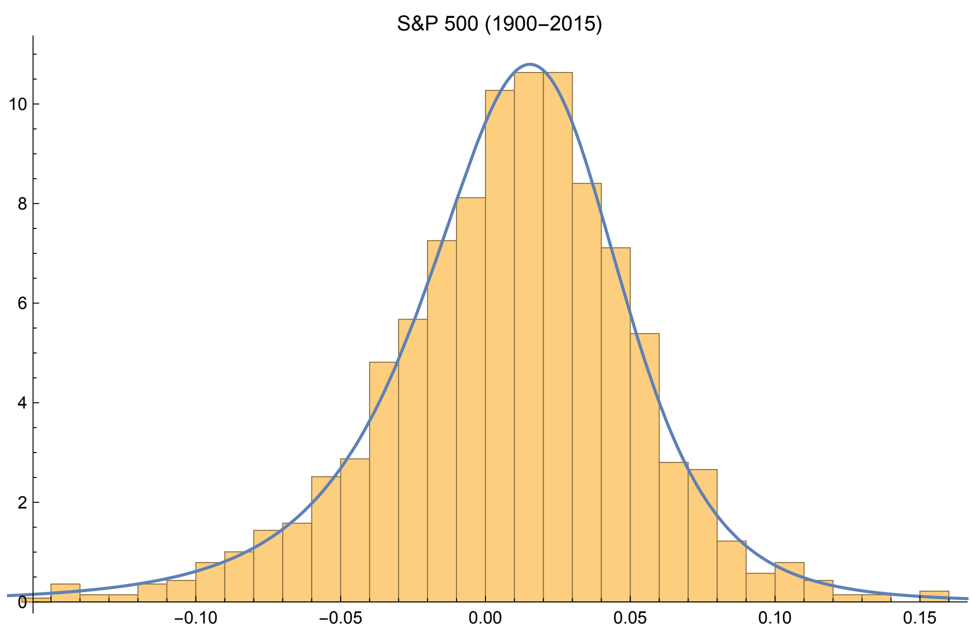
In[23]:=
```
TableForm[
 Join[{
   Through[
    {Mean, StandardDeviation, Skewness, Kurtosis}[mxSP500LogReturn[[All, 2]]]],
   Through[{Mean, StandardDeviation, Skewness, Kurtosis}[distSP500N]],
   Through[{Mean, StandardDeviation, Skewness, Kurtosis}[distSP500T]],
   Through[{Mean, StandardDeviation, Skewness, Indeterminate &}[distSP500ST]]
  }, {Prepend[Log@Likelihood[#, mxSP500LogReturn[[All, 2]]] & /@
      {distSP500N, distSP500T, distSP500ST}, "NA"]}ᵀ, 2],
 TableHeadings → {{"Observed", "Normal Fit", "Student t Fit", "Skew t Fit"},
   {"Mean", "Sdev", "Skew", "Kurt", "Logℒ"}}
]
```

Out[ ]//TableForm=

	Mean	Sdev	Skew	Kurt	Logℒ
Observed	0.00758187	0.0508859	-0.478152	11.2052	NA
Normal Fit	0.00758187	0.0508676	0	3	2170.949198
Student t Fit	0.0103985	0.0502706	0	Indeterminate	2318.414935
Skew t Fit	0.0072429	0.0506947	-1.62173	Indeterminate	2328.547120

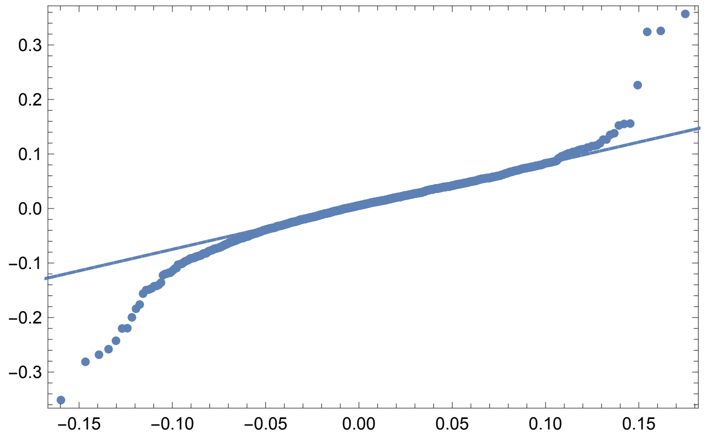
In[24]:=
```
Show[Histogram[mxSP500LogReturn[[All, 2]], Automatic, "PDF", ImageSize → 500,
  PlotLabel → "S&P 500 (1900-2015)"], Plot[PDF[distSP500ST, r], {r, -0.17, 0.17}]]
```

Out[ ]=

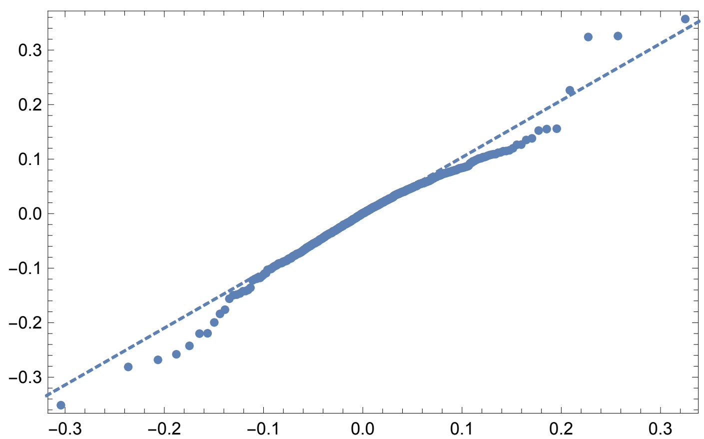


Quantile Plots of the Fits

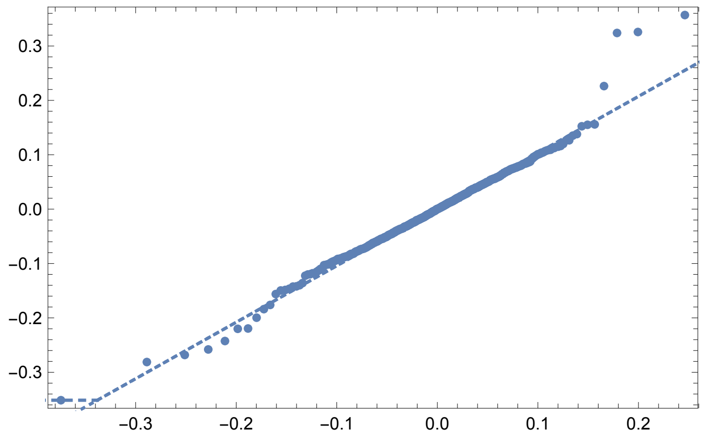
In[25]:=
```
QuantilePlot[mxSP500LogReturn[[All, 2]], distSP500N,
  PlotRange → All, PlotStyle → {PointSize[Medium]}]
```

Out[•]=


In[26]:=
```
QuantilePlot[mxSP500LogReturn[[All, 2]], distSP500T,
  PlotRange → All, PlotStyle → {PointSize[Medium]}]
```

Out[•]=


In[27]:=
```
QuantilePlot[mxSP500LogReturn[[All, 2]], distSP500ST,
  PlotRange → All, PlotStyle → {PointSize[Medium]}]
```

Out[ ]=


Hypothesis Tests

Normal Distribution

In[28]:=
```
hypN =
   DistributionFitTest[mxSP500LogReturn[[All, 2]], distSP500N, "HypothesisTestData"];
```

In[29]:=
```
hypN["TestDataTable", {"KolmogorovSmirnov", "CramerVonMises"}]
```

Out[ ]=

	Statistic	P-Value
Kolmogorov-Smirnov	0.0690767	3.17145×10^{-6}
Cramér-von Mises	2.57698	6.56659×10^{-7}

In[30]:=
```
hypN["TestDataTable", {"TestConclusion", "CramerVonMises"}]
```

Out[ ]= {

	Statistic	P-Value
Cramér-von Mises	2.57698	6.56659×10^{-7}

,

The null hypothesis that the data is distributed according to the NormalDistribution[0.00758187, 0.0508676] is rejected at the 5 percent level based on the Cramér-von Mises test.}

Student t Distribution

In[31]:=
```
hypT =
   DistributionFitTest[mxSP500LogReturn[[All, 2]], distSP500T, "HypothesisTestData"];
```

In[32]:=
```
hypT["TestDataTable", {"KolmogorovSmirnov", "CramerVonMises"}]
```

Out[ ]=

	Statistic	P-Value
Kolmogorov-Smirnov	0.0240718	0.389226
Cramér-von Mises	0.16364	0.350563

In[33]:=
```
hypT["TestDataTable", {"TestConclusion", "CramerVonMises"}]
```

Out[•]= {

	Statistic	P-Value
Cramér-von Mises	0.16364	0.350563

,
```
The null hypothesis that the data is distributed according to the
  StudentTDistribution[0.0103985, 0.0347629, 3.83284]
 is not rejected at the 5 percent level based on the Cramér-von Mises test.}
```

Noncentral (Skew) Student t Distribution

In[34]:=
```
hypST =
  DistributionFitTest[mxSP500LogReturn[[All, 2]], distSP500ST, "HypothesisTestData"];
```

In[35]:=
```
hypST["TestDataTable", {"KolmogorovSmirnov", "CramerVonMises"}]
```

Out[•]=

	Statistic	P-Value
Kolmogorov-Smirnov	0.0140675	0.942125
Cramér-von Mises	0.0221728	0.994478

In[36]:=
```
hypST["TestDataTable", {"TestConclusion", "CramerVonMises"}]
```

Out[•]= {

	Statistic	P-Value
Cramér-von Mises	0.0221728	0.994478

,
```
The null hypothesis that
 the data is distributed according to the TransformedDistribution[
   0.0321032 - 0.0341353 x, x ≈ NoncentralStudentTDistribution[3.90016, 0.577027]]
 is not rejected at the 5 percent level based on the Cramér-von Mises test.}
```

Value at Risk & Expected Shortfall

General Definition

The Value-at-Risk or VaR is the minimum loss expected over a given time period at a given confidence level. Note there are three components to VaR: the loss tolerance, the time period τ, and the confidence level χ. If F_τ is the cumulative distribution function of returns over a period of length τ, then the $\mathrm{VaR}_{\tau,\chi}$ satisfies the following relationship:

$$\mathrm{VaR}_{\tau,\chi} = \underset{r}{\arg}\{F_\tau(r) == 1 - \chi\}$$

which is the computation

$$\mathrm{VaR}_{\tau,\chi} = F_\tau^{-1}(1-\chi)$$

The expected shortfall, usually called the conditional VaR or CVaR, is the expected value of a loss that exceeds the VaR. Note that we have to include the normalization $1/(1-\chi)$ to account for the fact that we computing a conditional expectation.

$$\mathrm{CVaR}_{\tau,\chi} = \mathrm{E}[\,r \mid r \le \mathrm{VaR}_{\tau,\chi}] = \frac{1}{F_\tau(\mathrm{VaR}_{\tau,\chi})}\int_{-\infty}^{\mathrm{VaR}_{\tau,\chi}} r\, f_\tau(r)\, dr$$

$$= \frac{1}{1-\chi}\int_{-\infty}^{\mathrm{VaR}_{\tau,\chi}} r\, f_{\tau,}(r)\, dr$$

See https://en.wikipedia.org/wiki/Value_at_risk and https://en.wikipedia.org/wiki/Expected_shortfall.

Example - S&P 500

Compute the VaR for the monthly returns of the S&P 500 at a 99.5% confidence level: $\text{VaR}_{\text{month},0.995}$.

In[70]:=
```
nConfLimit = 0.999;
```

Normal Distribution

In[71]:=
```
nVaRN = InverseCDF[distSP500N, 1 - nConfLimit]
```

Out[•]= -0.149611

The $\text{CVaR}_{\text{month},0.995}$ is computed by

In[72]:=
```
nCVaRN = 1/(1 - nConfLimit) NIntegrate[r Evaluate[PDF[distSP500N, r]], {r, -∞, nVaRN}]
```

Out[•]= -0.163694

Student t Distribution

In[73]:=
```
nVaRT = InverseCDF[distSP500T, 1 - nConfLimit]
```

Out[•]= -0.250307

In[74]:=
```
nCVaRT = 1/(1 - nConfLimit) NIntegrate[r Evaluate[PDF[distSP500T, r]], {r, -∞, nVaRT}]
```

Out[•]= -0.3464

Noncentral (Skew) Student t Distribution

In[75]:=
```
nVaRST = InverseCDF[distSP500ST, 1 - nConfLimit]
```

Out[•]= -0.306465

In[76]:=
```
nCVaRST = 1/(1 - nConfLimit) NIntegrate[r Evaluate[PDF[distSP500ST, r]], {r, -∞, nVaRST}]
```

Out[•]= -0.4271

As one can see, as we make the distribution more realistic we see that not capturing the heavy tails and negative skew has serious consequences.

In[81]:=
```
Grid[{
  {"Confidence = 99.9%"},
  {TableForm[{{nVaRN, nCVaRN}, {nVaRT, nCVaRT}, {nVaRST, nCVaRST}},
    TableHeadings → {{"Normal", "Student t", "Skew Student t"}, {"VaR", "CVaR"}}]}
 },
 Frame → All
]
```

Out[ ]=

Confidence = 99.9%		
	VaR	CVaR
Normal	-0.149611	-0.163694
Student t	-0.250307	-0.3464
Skew Student t	-0.306465	-0.4271

Power Law Models of Extreme Events

Extreme Events

For a wide class of "fat tailed" (in a sense that will be made clear shortly) distributions, their extreme values can be characterized by a *power law*. We will consider only the upper tail, which we will define as the values of a variable ξ that are *above* some threshold τ.

First, define the cumulative tail probability or *survival function* as the probability that a random variable Ξ exceeds a given value

$$Q(\xi) \equiv \text{Prob}\{\Xi \geq \xi\}$$

If $F(\xi)$ is the CDF of Ξ, then

$$Q(\xi) = 1 - F(\xi)$$

For ξ *sufficiently large* we assume that the cumulative tail probability can be approximated by a power law:

$$\text{Prob}\{\Xi \geq \xi\} \equiv Q_\alpha(\xi) \tilde{\propto} \xi^{-\alpha}$$

Limiting ourselves to extreme values of ξ (*i.e.*, events that exceed a threshold τ) we approximate the distribution of these extreme values by

$$\text{Prob}\{\Xi \geq \xi \mid \Xi \geq \tau\} \equiv Q_{\alpha,\tau}(\xi) = \frac{1}{Q_\alpha(\tau)} Q_\alpha(\xi) = \zeta\, \xi^{-\alpha}, \quad \tau \leq \xi \leq \infty$$

Power Law (Pareto-Like) Distributions

In general, power law distributions (https://en.wikipedia.org/wiki/Power_law#Power-law_probability_distributions) where

$$P[\Xi > \xi] \sim L(\xi)\, \xi^{-\alpha}, \qquad \text{for } \alpha > 1$$

where L is a slowly varying function such that

$$\underset{\xi \to \infty}{\text{limit}}\left[\frac{L(\lambda\, \xi)}{L(\xi)}\right] = 1, \qquad \forall\, \lambda > 0$$

Thus, in the limit as $\xi \to \infty$, power law distributions approach the Pareto.

Moments and the Power Law Exponent

Far enough out on the tail L is approximately constant and the PDF is approximately proportional to

$$f(\xi) \propto \xi^{-(\alpha+1)}$$

which results from simply differentiating $F_{\alpha,\tau}(\xi)$ with respect to ξ. Consider the contribution of the tail to the n^{th} moment of $f(\xi)$.

$$\int_{\tau}^{\infty} \xi^n f(\xi)\, d\xi \;\propto\; \int_{\tau}^{\infty} \xi^n \xi^{-(1+\alpha)}\, d\xi \;=\; \int_{\tau}^{\infty} \xi^{n-(1+\alpha)}\, d\xi$$

The integral above converges only if $n < \alpha$; alternately, we can say that moments greater than or equal to α do not exist (in the sense that their limits are ∞, or more technically are indeterminate). Thus, the exponent of the power law gives us an indication of the point at which the moments cease to exist and, hence, provides us with a means for characterizing how “fat” the tail is.

Note, the α used here as the exponent of a power law is *not* the same as the α parameter of a Stable Distribution.

Comparing the Normal and Cauchy Distributions

It is useful at this point to compare the thinly tailed distribution, the Normal, with a heavily tailed one, the Cauchy.

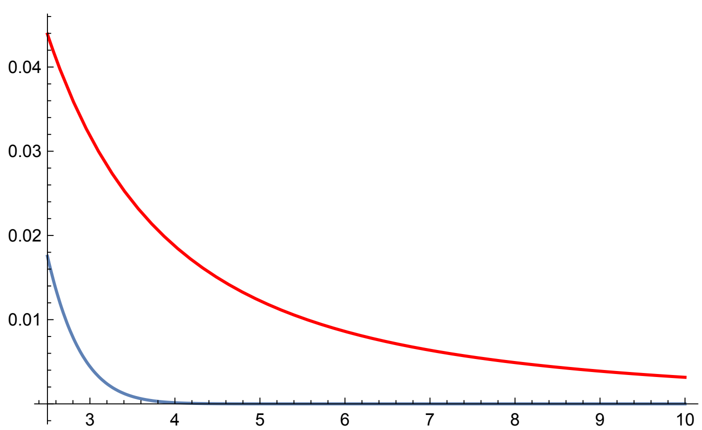
```
PDF[CauchyDistribution[], x]
```

$$\frac{1}{\left(1 + x^2\right) \pi}$$

```
Mean[CauchyDistribution[]]
```

```
Indeterminate
```

```
Show[
 Plot[Evaluate[PDF[NormalDistribution[], x]], {x, 2.5, 10}, PlotRange → All],
 Plot[Evaluate[PDF[CauchyDistribution[], x]],
  {x, 2.5, 10}, PlotRange → All, PlotStyle → Red]
]
```

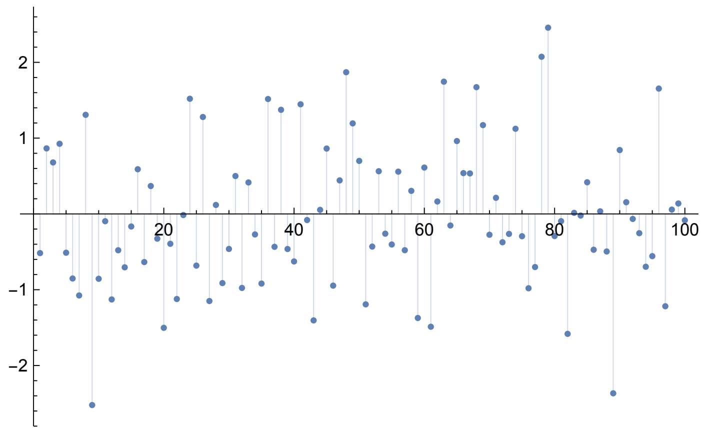


(Fractal-like) Scaling of Cauchy Samples

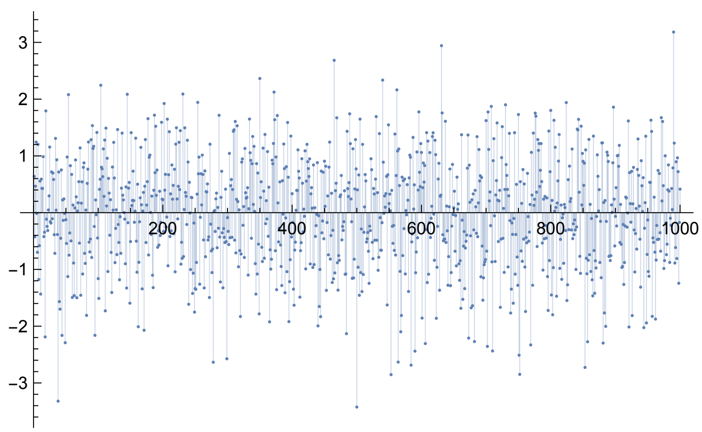
Normal Samples

Below we compare samples of increasing size from a standard Normal distribution. Note that as the sample size increases, the sample quickly settles into a fairly stable picture; *viz*., the data are almost all ±3 with rare excursions out as far as ±4.

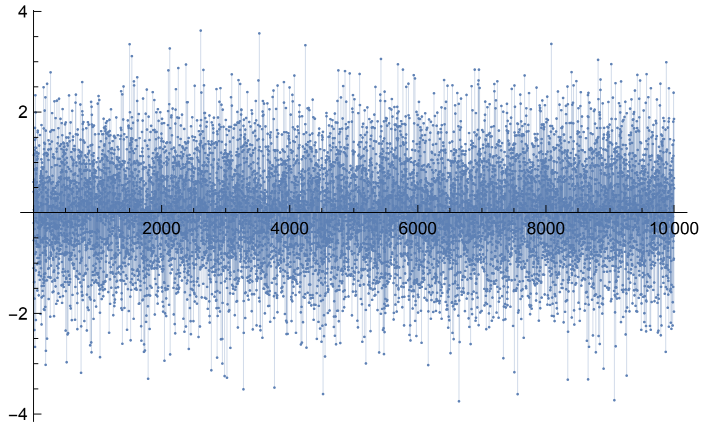
```
ListPlot[RandomVariate[NormalDistribution[], 100], Filling → 0, PlotRange → All]
```



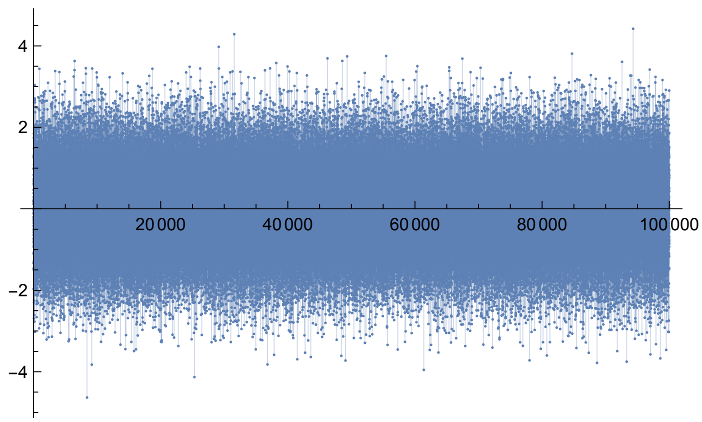

```
ListPlot[RandomVariate[NormalDistribution[], 1000], Filling → 0, PlotRange → All]
```



```
ListPlot[RandomVariate[NormalDistribution[], 10 000], Filling → 0, PlotRange → All]
```



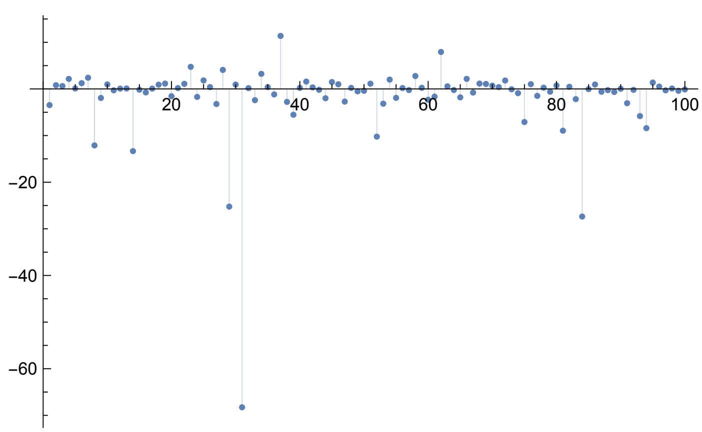
```
ListPlot[RandomVariate[NormalDistribution[], 100 000], Filling → 0, PlotRange → All]
```



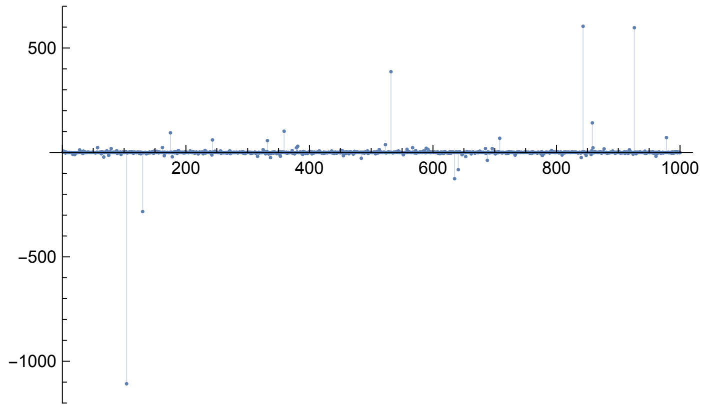
Cauchy Samples

In contrast, the behavior of the Cauchy samples of increasing size qualitatively look similar: many small values punctuated with extreme excursions. However, note how the range of extreme values increases with the sample size. This tendency to look the same at increasing scales is fractal-like. This means that a Cauchy driven process never settles into any long-term steady state behavior.

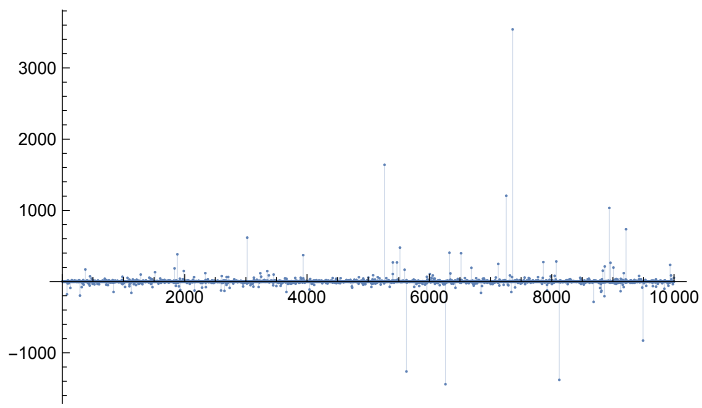
```
ListPlot[RandomVariate[CauchyDistribution[], 100], Filling → 0, PlotRange → All]
```



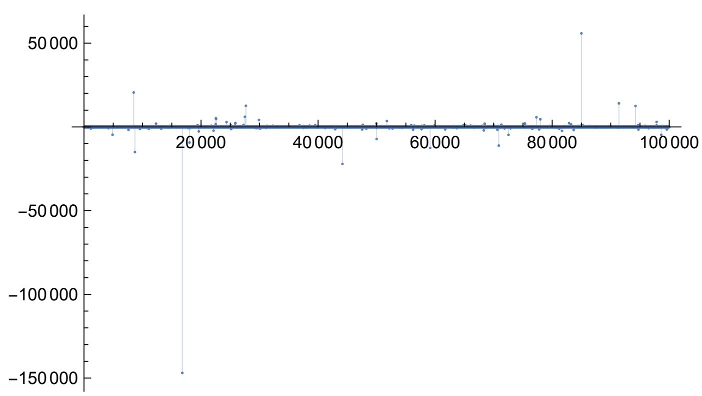
```
ListPlot[RandomVariate[CauchyDistribution[], 1000], Filling → 0, PlotRange → All]
```




```
ListPlot[RandomVariate[CauchyDistribution[], 10 000], Filling → 0, PlotRange → All]
```




```
ListPlot[RandomVariate[CauchyDistribution[], 100 000], Filling → 0, PlotRange → All]
```



Taleb’s Mediocristan and Extremistan

Nassim N. Taleb in his book *The Black Swan* described two different conditions: thin tailed (Mediocristan) and heavy-tailed (Extremistan). Consider a group of 1,000 individuals and two measurements: their mean height and their mean wealth.

- *Mediocristan*: What is likely to happen to the mean if you replace an “average” person in the sample with the tallest person in the world: ~9 feet? If we replace one average person with the tallest person in the world, then the mean of the sample increases by 0.003.

$$\frac{(6 \times 999) + 9}{1000.}$$

6.003

- *Extremistan*: What is likely to happen to the mean if you replace an “average” person in the sample with the richest person in the world: ~ $100 billion? If we replace one average person with the richest person in the world, then the mean of the sample increases from $100 thousand to about $100 million.

In[82]:= $$\frac{(100\,000 \times 999) + 100 \times 10^9}{1000.}$$

Out[]= 1.001×10^8

- What is likely to happen to the median values of these two measures in each case above? Naturally, the median would not be affected.

Adapting to Lower Tail Models

Generally, though not always, we are concerned with one tail of the distribution from a risk point of view; *e.g.*, the upper tail when we are trying to hedge out a price increase of a commodity or the lower tail when assessing the risk of loss for a long stock portfolio. It is straightforward to apply the necessary conversions so we are always looking at an upper tail.

Caveats

Don’t Apply a Power Law without Clear Evidence

A power law characterization of a distribution's tail is meant to be an approach for characterizing distributions that show clear evidence of power-law behavior. Consider, for example, the Normal distribution. There is no such power law that bounds its tails from below because that would contradict the fact that all of its moments exist.

However, real systems are finite and don't have the entire real line as their domains; nevertheless, the theory above is a useful framework for characterizing the extreme values of distributions that are poorly modeled by the Normal and for estimating the likelihood of large events based on empirical observations.

The Power Law Characterization is an Improper Probability Distribution

While the power law characterization of the tail resembles closely the form of a Pareto distribution, it is important to remember that *it is an empirical characterization of only a relatively small portion of the distribution in terms of its survival function*. Thus, it is not in itself a complete distribution. In performing computations we can often treat it as a distribution, but we must be careful to restrict those computations to the appropriate range of values.

Fitting a Power Law

A power law approximation of the tail of a distribution for x is defined on some range $0 < \tau \le x \le \infty$. We are looking for rare events multiple deviations from the mean so τ is positive. Thus, if we are interested in the upper tail of a variable x, then $\xi = x$. If we are interested in the lower tail, then $\xi = -x$. This we think of extreme events, whether gains or losses, in positive units.

Starting with a sample $\boldsymbol{\xi}$ of N observations, our tail sample $\boldsymbol{\xi}_\tau$ is the ordered list sorted from smallest to largest with I tail observations that exceed a threshold τ

$$\boldsymbol{\xi}_\tau = \text{Sort}[\text{Select}[\boldsymbol{\xi}, \# \ge \tau \,\&]] = \{\xi_{N-I+1}, \ \ldots, \xi_{N-I+i}, \, , \ \ldots, \xi_N\}$$

A reasonable estimate for $\hat{Q}_\tau(\xi_i)$, which "centers" the discrete observations, is

$$\hat{Q}_\tau(\xi_i) = 1 - \frac{i - \frac{1}{2}}{N}, \; i = N - I + 1, \; \ldots, \; N$$

We can linearize the data by taking the logarithm of both sides. We'll use ζ for the normalizing constant.

$$\hat{Q}_\tau(\xi_i) \cong \zeta \, \xi_i^{-\alpha} \quad \Longrightarrow \quad \mathrm{Log}\left[\hat{Q}_\tau(\xi_i)\right] \cong \mathrm{Log}[\zeta] - \alpha \, \mathrm{Log}[\xi_i]$$

Plotting these data on a log-log plot gives us a simple visual check, the slope of a straight line being easier to interpret than the power of a curved one.

Example - Stock Portfolio for 1900 to 2015

It is easier in selecting reasonable values for τ to work in terms of standard units of dispersal above the central tendency. Here we use the median for the central tendency and median deviation for the measure of dispersal. They are robust measures that will produce consistent results even in the presence of fat tails.

Picking τ

In[83]:=
```
{nMednSP500, nMdevSP500} =
  Through[{Median, MedianDeviation}[mxSP500LogReturn⟦All, 2⟧]]
```

Out[•]=
```
{0.0113887, 0.0267734}
```

Often looking at 3 units of dispersal out from the center gives reasonable results, although this needs to be carefully checked.

In[84]:=
```
nTau = nMdevSP500 3 + nMednSP500
```

Out[•]=
```
0.0917088
```

Extracting and Transforming Data from the Lower Tail

Next we compute the tail sample and produce the survival function. Note that because we are interested in the loss tail we take the *negative* of the S & P 500 returns.

In[85]:=
```
vnTailSample = Sort[Select[-mxSP500LogReturn⟦All, 2⟧, # ≥ nTau &]]
iSample = Length[mxSP500LogReturn⟦All, 2⟧]
iTail = Length[vnTailSample]
```

Out[•]=
```
{0.0933283, 0.0947389, 0.0956101, 0.0964682, 0.0976655, 0.0993924, 0.101422,
 0.101707, 0.102635, 0.103043, 0.109211, 0.10985, 0.112584, 0.115051, 0.117898,
 0.117952, 0.119558, 0.120206, 0.12238, 0.136009, 0.139935, 0.14198, 0.142439,
 0.146393, 0.14858, 0.149598, 0.156159, 0.176159, 0.183863, 0.199612,
 0.219511, 0.220065, 0.242516, 0.258023, 0.268177, 0.281166, 0.351359}
```

Out[•]=
```
1392
```

Out[•]=
```
37
```

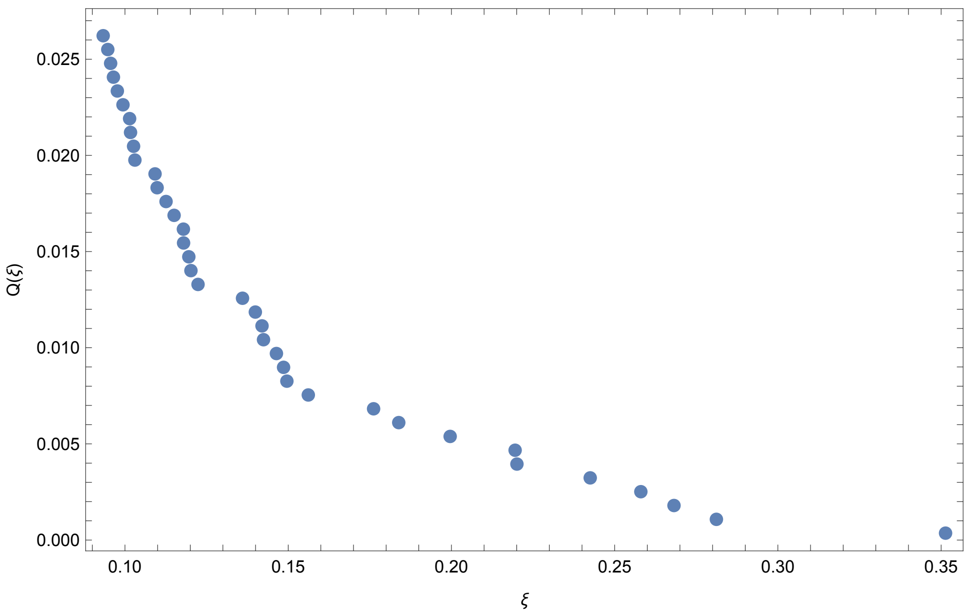
Computing the Survival Function Data

In[88]:=
```
mnTailSF =
  Transpose[{Sort[vnTailSample], 1 - (Range[iSample - iTail + 1, iSample] - 0.5)/iSample}];
mnLogTailSF = Log[mnTailSF];
```

Plotting the survival function and its log transformation. The latter indicates the linear relation we are looking for.

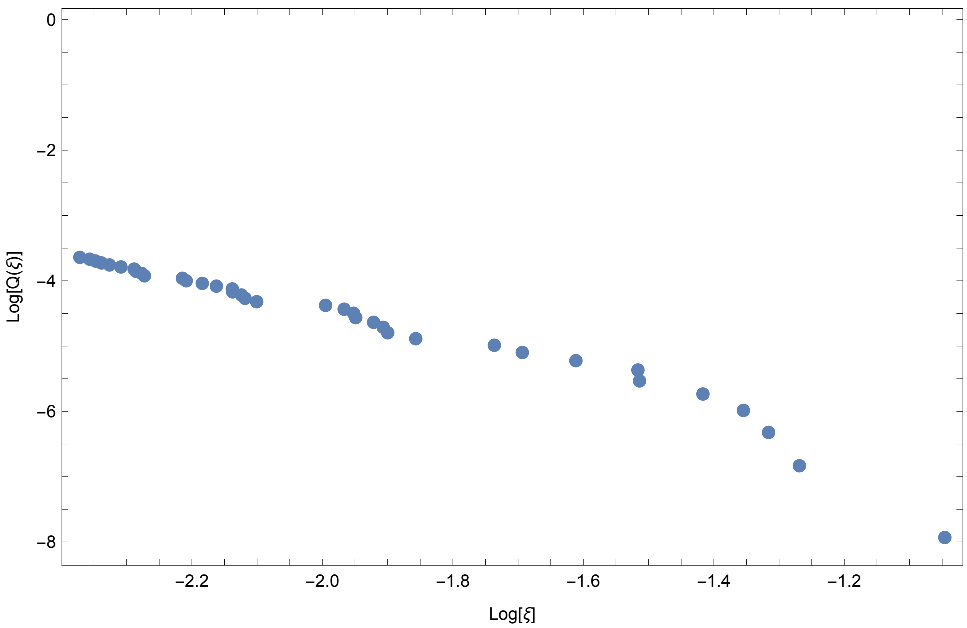
In[90]:=
```
ListPlot[mnTailSF, PlotStyle → {PointSize[Large]}, Axes → False, Frame → True,
 FrameLabel → {"ξ", "Q(ξ)", Style["Tail Data", FontSize → 16], ""}, ImageSize → 500]
```

Out[ ]=



In[91]:=
```
ListPlot[mnLogTailSF, PlotStyle → {PointSize[Large]}, Axes → False, Frame → True,
 FrameLabel → {"Log[ξ]", "Log[Q(ξ)]", Style["Tail Data", FontSize → 16], ""},
 ImageSize → 500]
```

Out[•]=



Fitting the Power Law

We fit a regression to the log transformed data to estimate ζ and α.

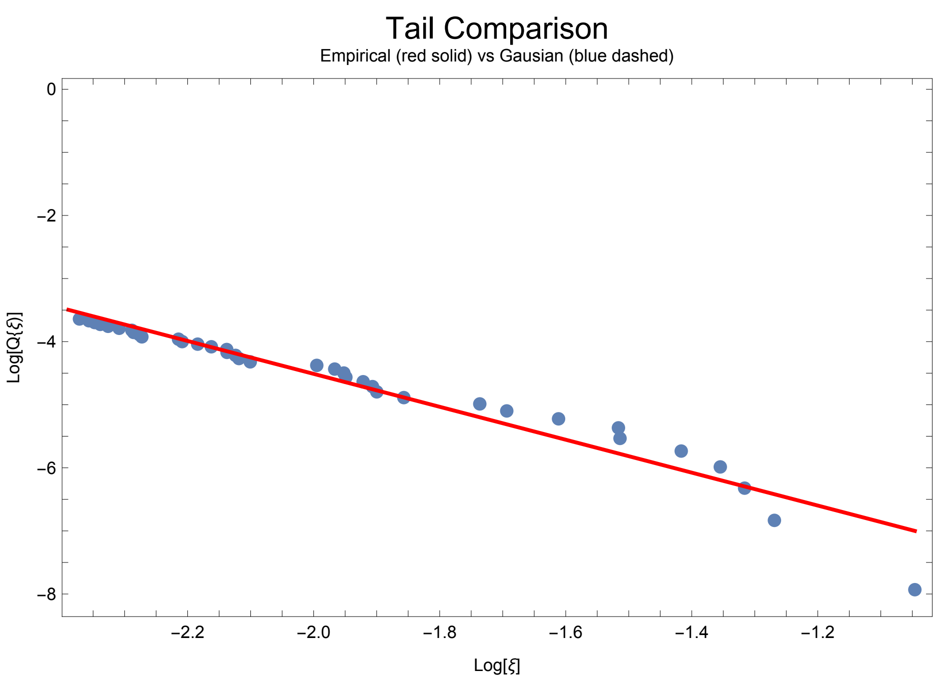
In[92]:=
```
{nConstant, nPowerLaw} = {Exp[First[#]], -Last[#]} &[
  LinearModelFit[mnLogTailSF, logξ, logξ]["BestFitParameters"]]
```

Out[•]=
```
{0.0000595497, 2.60876}
```

We plot the fit against the data in both log and untransformed space.

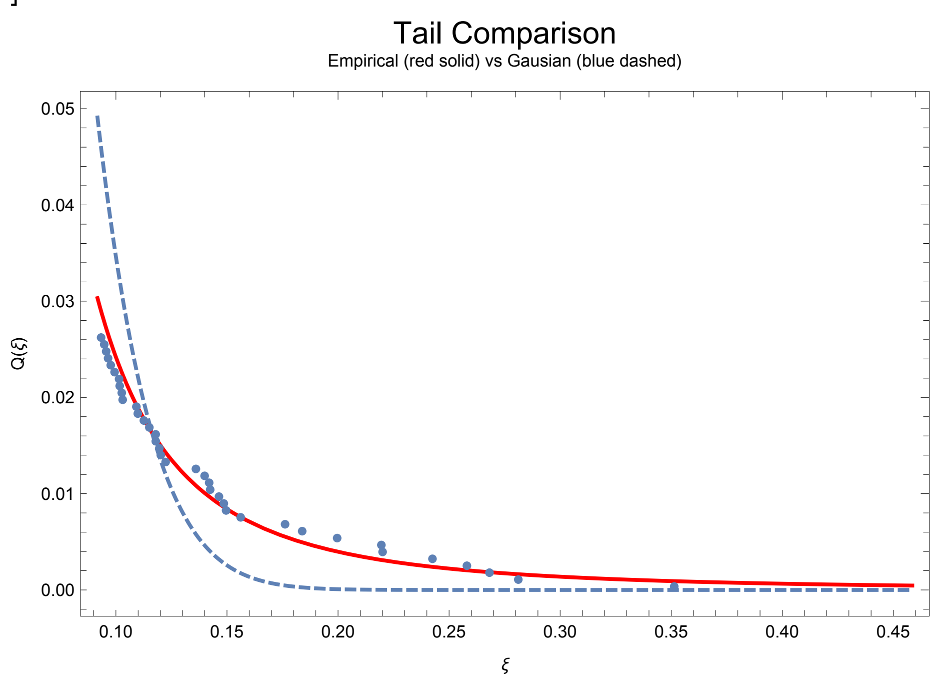
In[93]:=
```
Show[
 ListPlot[
  mnLogTailSF,
  PlotStyle → {PointSize[Large]},
  Axes → False,
  Frame → True,
  FrameLabel →
   {{"Log[Q{ξ}]", ""}, {"Log[ξ]", Column[{Style["Tail Comparison", FontSize → 16],
       "Empirical (red solid) vs Gausian (blue dashed)"}, Center]}},
  ImageSize → 500
 ],
 Plot[
  Log[nConstant] - nPowerLaw logXi ,
  {logXi, Log[nTau], Max[First /@ mnLogTailSF]},
  PlotStyle → {Red, Thick}
 ]
]
```

Out[•]=



In[94]:=
```
Show[
 Plot[
  nConstant z^-nPowerLaw,
  {z, nTau, 5. nTau},
  PlotRange → All,
  PlotStyle → {Red, Thick},
  Frame → True,
  FrameLabel → {"ξ", "Q(ξ)", Column[{Style["Tail Comparison", FontSize → 16],
      "Empirical (red solid) vs Gausian (blue dashed)"}, Center], ""},
  ImageSize → 500
 ],
 Plot[
  1 - Evaluate[CDF[distSP500N, z]],
  {z, nTau, 5. nTau},
  PlotStyle → {Thick, Dashed},
  PlotRange → All
 ],
 ListPlot[
  mnTailSF,
  PlotStyle → {PointSize[Medium]}
 ]
]
```

Out[ ]=



As the plot above clearly indicates, we have successfully modeled the tail above τ. Also note, that the data clearly indicates that the Normal approximation models the tail behavior poorly.

Again, it is important to note that the power law model is not a proper distribution. It is a *local* approximation of the tail in terms of the survival function for values above τ.

VaR and CVaR for a Power Law Tail

VaR

If we apply the power law model for the size of the loss, then the VaR (*where the loss in this case is expressed as a positive number*) satisfies the following relationship:

$$\mathrm{VaR}_\chi = \{\xi \mid Q(\xi) = 1 - \chi\} = \{\xi \mid \zeta \xi^{-\alpha} = 1 - \chi\} = \left(\frac{\zeta}{1-\chi}\right)^{1/\alpha}$$

We can use Solve[] which produces the mathematically equivalent result

In[95]:=
```
Solve[ζ VaR^-α == 1 - χ, VaR, InverseFunctions → True]
```

Out[•]= $\left\{\left\{\mathrm{VaR} \to \left(-\frac{-1+\chi}{\zeta}\right)^{-1/\alpha}\right\}\right\}$

The VaR for the power law tail is

In[96]:=
```
nVaRPL = (ζ/(1 - χ))^(1/α) /. {χ → nConfLimit, ζ → nConstant, α → nPowerLaw}
```

Out[•]= 0.339142

CVaR

The CVaR is then

$$\mathrm{CVaR} = \mathrm{E}[\xi \mid \xi \geq \mathrm{VaR}] = \frac{1}{1-\chi}\int_{\mathrm{VaR}}^{\infty} \alpha\,\zeta\,\xi^{-\alpha}\,d\xi = \frac{\alpha}{\alpha-1}\left(\frac{\zeta}{1-\chi}\right)^{1/\alpha} = \frac{\alpha}{\alpha-1}\,\mathrm{VaR}$$

An example of how to use *Mathematica* to solve the above is immediately below. Note the Assumptions options in Integrate and FullSimplify; try running the expressions without them.

First we will check that $Q(\mathrm{VaR}) = 1 - \chi$

In[97]:=
```
(ζ ξ^-α /. {ζ → nConstant, α → nPowerLaw, ξ → nVaRPL})
```

Out[•]= 0.001

Next, we derive the PDF of the power law tail by differentiating the CDF.

In[98]:=
```
D[1 - ζ ξ^-α, ξ]
```

Out[•]= $\alpha\,\zeta\,\xi^{-1-\alpha}$

As a sanity check, we can verify that integrating the PDF over the limits $\mathrm{VaR} \leq \xi \leq \infty$ yields $1 - \chi$.

In[99]:=
```
Integrate[Evaluate[α ζ ξ^(-1-α) /. {ζ → nConstant, α → nPowerLaw}], {ξ, nVaRPL, ∞}]
```

Out[•]= 0.001

We now have the confidence that our conditional expectation will yield the proper CVaR.

In[101]:=
```
nCVaRPL = NIntegrate[
  Evaluate[(α ξ ξ^-α)/(1 - χ) /. {χ → nConfLimit, ξ → nConstant, α → nPowerLaw}], {ξ, nVaRPL, ∞}]
```

Out[•]=
```
0.549952
```

As with the VaR for a power law, we can also derive a closed-form expression for the CVaR.

```
Integrate[(α ξ ξ^-α)/(1 - χ), {ξ, VaR, ∞}, Assumptions → α > 1 && 0 < χ < 1 && ξ > 0 && VaR > 0]
```

$$-\frac{\mathrm{VaR}^{1-\alpha}\,\alpha\,\xi}{(-1+\alpha)\ (-1+\chi)}$$

```
FullSimplify[% /. VaR -> (-(-1 + χ)/ξ)^(-1/α), Assumptions → α > 1 && 0 < χ < 1 && ξ > 0]
```

$$\frac{\alpha\left(\frac{\xi}{1-\chi}\right)^{\frac{1}{\alpha}}}{-1+\alpha}$$

But $\left(\frac{\xi}{1-\chi}\right)^{\frac{1}{\alpha}}$ is just the VaR, so this simplifies further to $\frac{\alpha}{\alpha-1}$ VaR. We can check this against the numerical computation for the CVaR above. Note, however, that $\alpha > 1$.

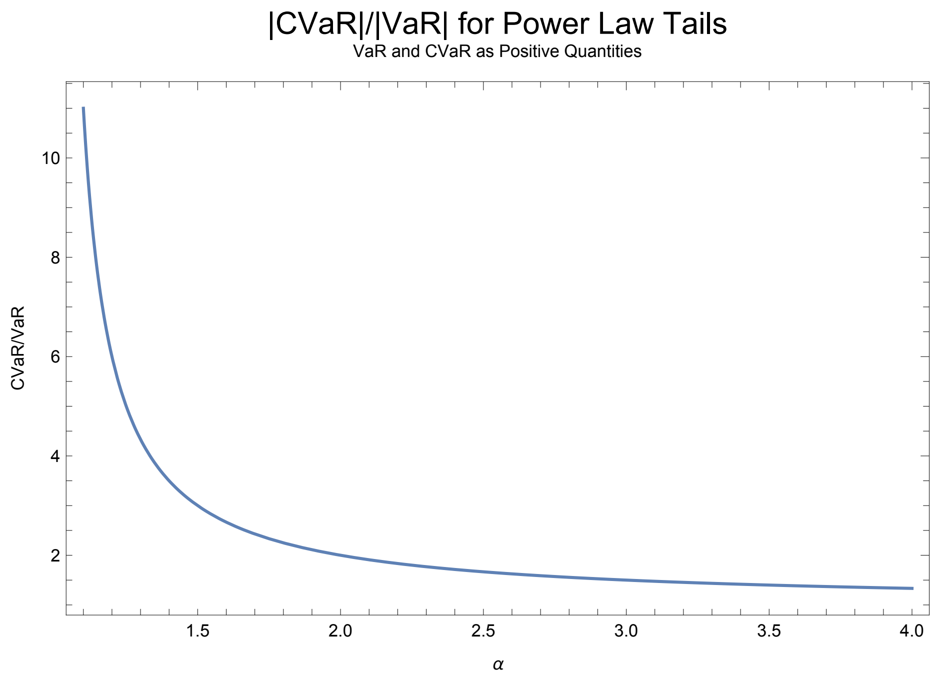
In[102]:=
```
α/(α - 1) nVaRPL /. {α → nPowerLaw}
```

Out[•]=
```
0.549952
```

Because the relationship between a power law VaR and CVaR is a function of α, we can examine this with the plot below.

In[103]:=
```
Plot[
  α
 ─────,
 α - 1
 {α, 1.1, 4},
 PlotRange → All,
 Frame → True,
 FrameLabel → {{"CVaR/VaR", ""},
   {"α", Column[{Style["|CVaR|/|VaR| for Power Law Tails", FontSize → 16],
      "VaR and CVaR as Positive Quantities"}, Center]}},
 ImageSize → 500
]
```

Out[ ]=



VaR and CVaR Using the Empirical Distribution

Working with an Empirical Distribution

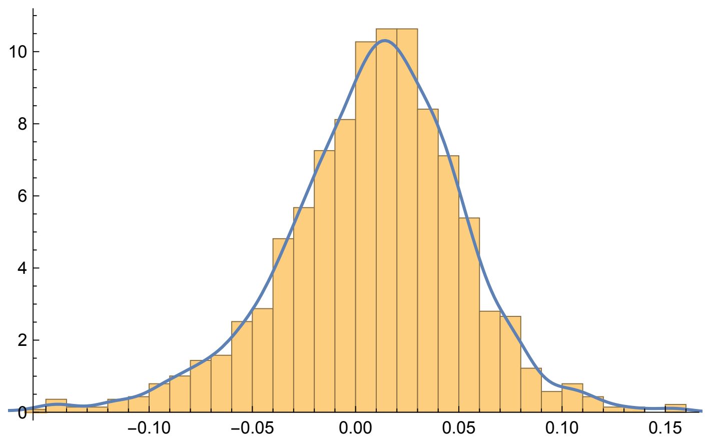
Rather than attempting to fit a distribution to the data, we bootstrap estimates of VaR and CVaR directly from the data. First, we approximate an empirical distribution by using a Kernel smoothed estimate.

In[104]:=
```
distSP500E = SmoothKernelDistribution[mxSP500LogReturn[[All, 2]]]
```

Out[ ]=
```
DataDistribution[ Type: SmoothKernel Data points: 1392 ]
```

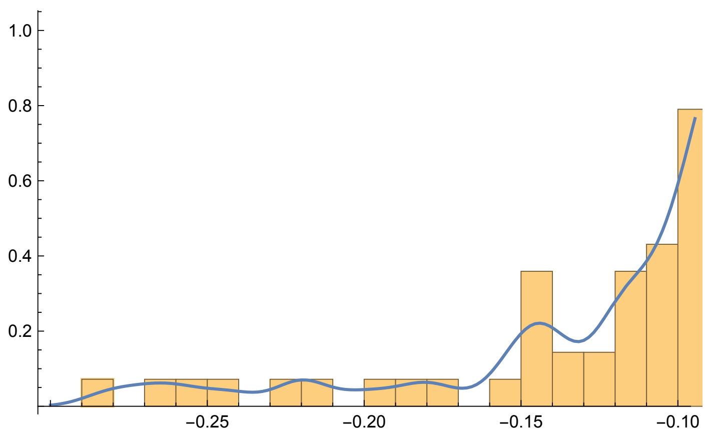
The overall fit appears quite reasonable.

In[105]:=
```
Show[Histogram[mxSP500LogReturn[[All, 2]], Automatic, "PDF"],
 Plot[PDF[distSP500E, x], {x, -0.3, 0.3}]
]
```

Out[•]=


Looking at the lower tail, however, raises some issues.

In[106]:=
```
Show[Histogram[mxSP500LogReturn[[All, 2]], Automatic, "PDF",
  PlotRange → {{-0.1, -0.3}, {0., 1}}], Plot[PDF[distSP500E, x], {x, -0.3, -0.075}]
]
```

Out[•]=


In looking at the lower tail of the PDF we see an important problem with using an empirical distribution: We are basing our computations on a very small sample. We could smooth the kernel estimate more but we are still faced with the problem of over-fitting to the data. For example, we probably do not believe that the local peak in the PDF at about -0.15 is real. Having some theoretical framework to work within, the power law fit is an example, can often give us better estimates of out of sample behavior.

Another important issue when using the empirical distribution directly in the presence of fat tails is that we almost always underestimate the size of extreme events. This is what we saw in simulations of the Cauchy distribution above: As the sample size increased the incidence of large excursions did as well. In contrast, with a Normally distributed random variable we rarely see events outside $\pm 3\sigma$—even in large samples.

Estimating VaR and CVaR

The empirical $\text{VaR}_{\text{monthly},0.999}$ is computed by

In[107]:=
```
nVaRE = InverseCDF[distSP500E, 1 - nConfLimit]
```

Out[]=
```
-0.284225
```

The $\text{CVaR}_{\text{month},0.999}$ is computed by

In[109]:=
$$\text{nCVaRE} = \frac{1}{1 - \text{nConfLimit}} \text{NIntegrate}[r\ \text{Evaluate}[\text{PDF}[\text{distSP500E}, r]], \{r, -\infty, \text{nVaRE}\}]$$

Out[]=
```
-0.334042
```

Summary

Here is a summary of the different estimates generated for the VaR and CVaR of the S & P 500. Note that we have used the negative of the power law VaR and CVaR so that they properly reflect as loses.

In[112]:=
```
Grid[{
  {"Confidence = 99.9%"},
  {TableForm[{{nVaRN, nCVaRN}, {nVaRT, nCVaRT}, {nVaRST, nCVaRST},
     {-nVaRPL, -nCVaRPL}, {nVaRE, nCVaRE}}, TableHeadings → {{"Normal", "Student t",
       "Skewed t", "PowerLaw", "Kernel Smoothed"}, {"VaR", "CVaR"}}]}
 },
 Frame →
  All]
```

Out[]=

Confidence = 99.9%		
	VaR	CVaR
Normal	-0.149611	-0.163694
Student t	-0.250307	-0.3464
Skewed t	-0.306465	-0.4271
PowerLaw	-0.339142	-0.549952
Kernel Smoothed	-0.284225	-0.334042