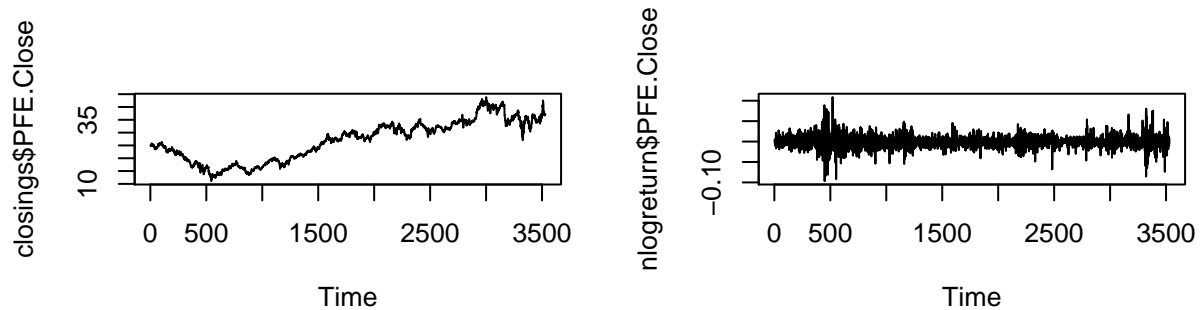# Extreme Value Theory: Application to Pfizer Stock Prices.
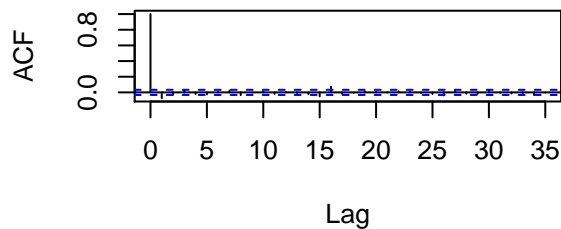
Sherman ALINE, Sungyun HWANG

We examine the stock prices of Pfizer.

```
closings <- data.frame(Date = index(PFE),Close = PFE$PFE.Close)
nlogreturn <- data.frame(Date = index(PFE),nlogreturn = -diff(log(PFE$PFE.Close)))
par(mfrow = c(2,2))
plot.ts(closings$PFE.Close); plot.ts(nlogreturn$PFE.Close)
acf(na.omit(nlogreturn$PFE.Close))
```


Series na.omit(nlogreturn$PFE.Close

As is common in finance, we take the negative difference of the log. This removes non-stationarity and mitigates volatility, as stocks often have exponentials growths in valuation. We can see from the ACF that this was effective and we have no serial correlation.

```
years <- format(nlogreturn$Date, "%Y")
bydate <- tapply(nlogreturn$PFE.Close, years, max)
```

1. Block Maxima Approach

i) fit a GEV

```
model <- fgev(bydate); par(mfrow = c(2,2)); plot(model)
```

Probability Plot

Quantile Plot

Density Plot


Return Level Plot
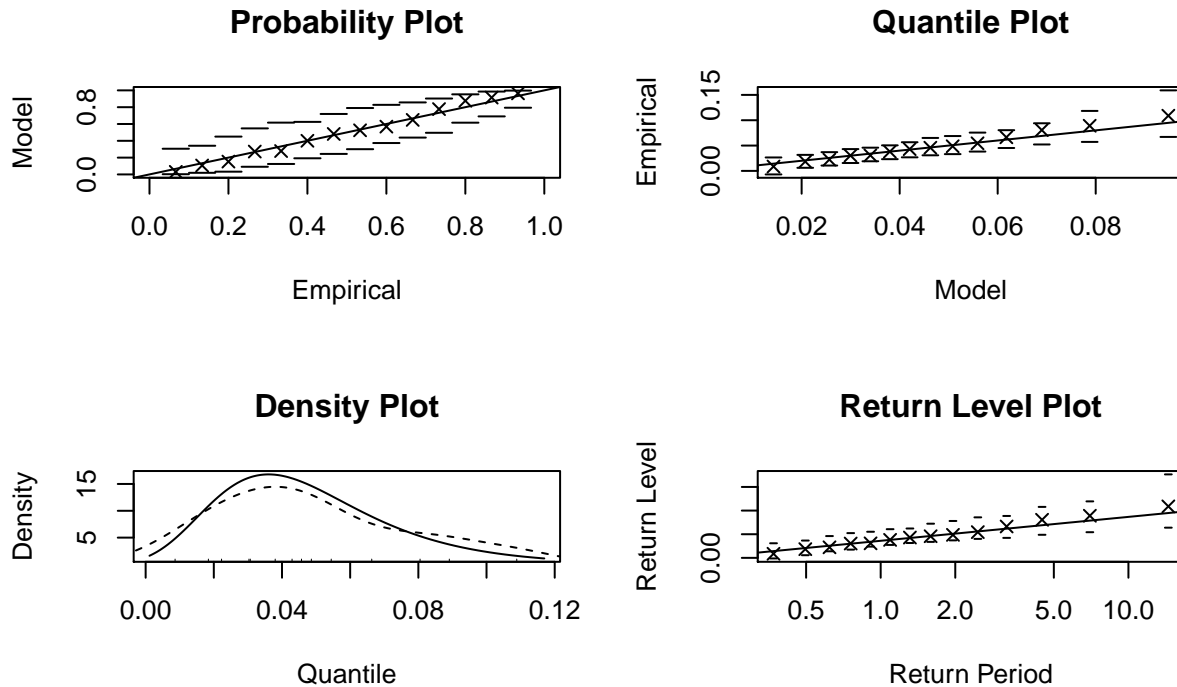
Diagnostic plots for assessing the accuracy of the EV model for these data are shown above. Probability plot (top left), quantiles from a sample drawn from the fitted GEV df against the empirical data quantiles with 95% confidence bands (top right), density plots of empirical data and fitted GEV df (bottom left), and return level plot with 95% pointwise normal approximation confidence intervals (bottom right).

ii) do model selection and model checking

```
qq(model,cex=0.3)
```

Quantile Plot
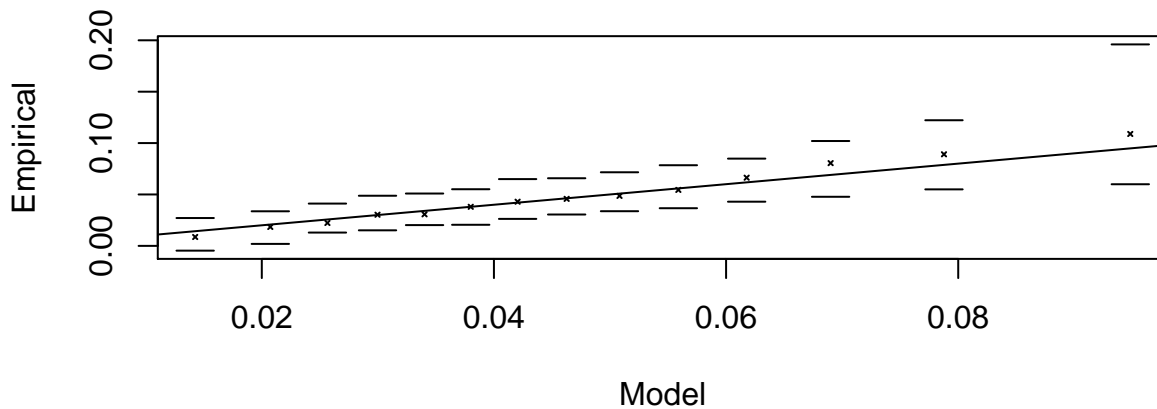

```
#QQ-plots are useful for model checking, identifying possible outliers
```

QQplot- The plot shows 95% pointwise confidence intervals obtained by parametric bootsrap.

We make another model and compare with likelihood ration test to checkif trend is zero.

```
cov <- data.frame(year = scale(1:nrow(bydate), scale=FALSE)/100)
model2 <- fgev(bydate, nsloc=cov)
W <- 2*(logLik(model2) - logLik(model) )
```
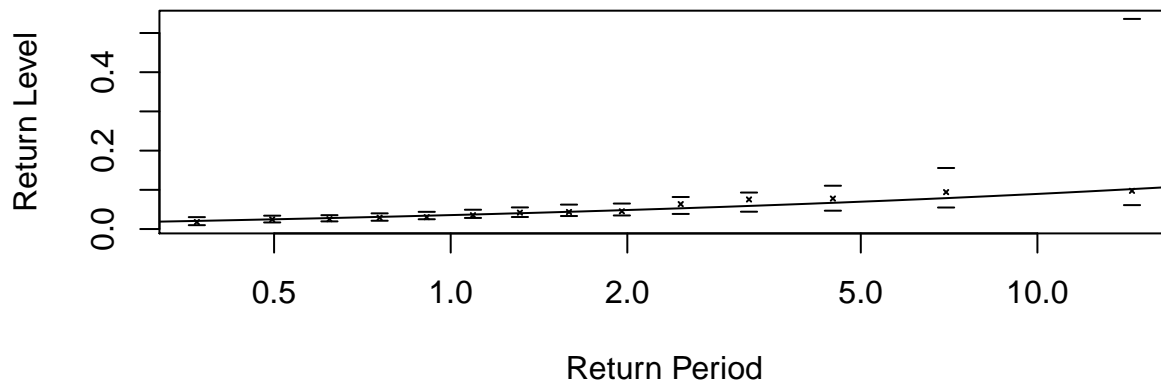
2

```
pchisq(W, df=1, lower.tail=FALSE)
```

'log Lik.' 0.07497936 (df=4)

We reject the null hypothes $\beta = 0$ and so use the model with trend. (model2)

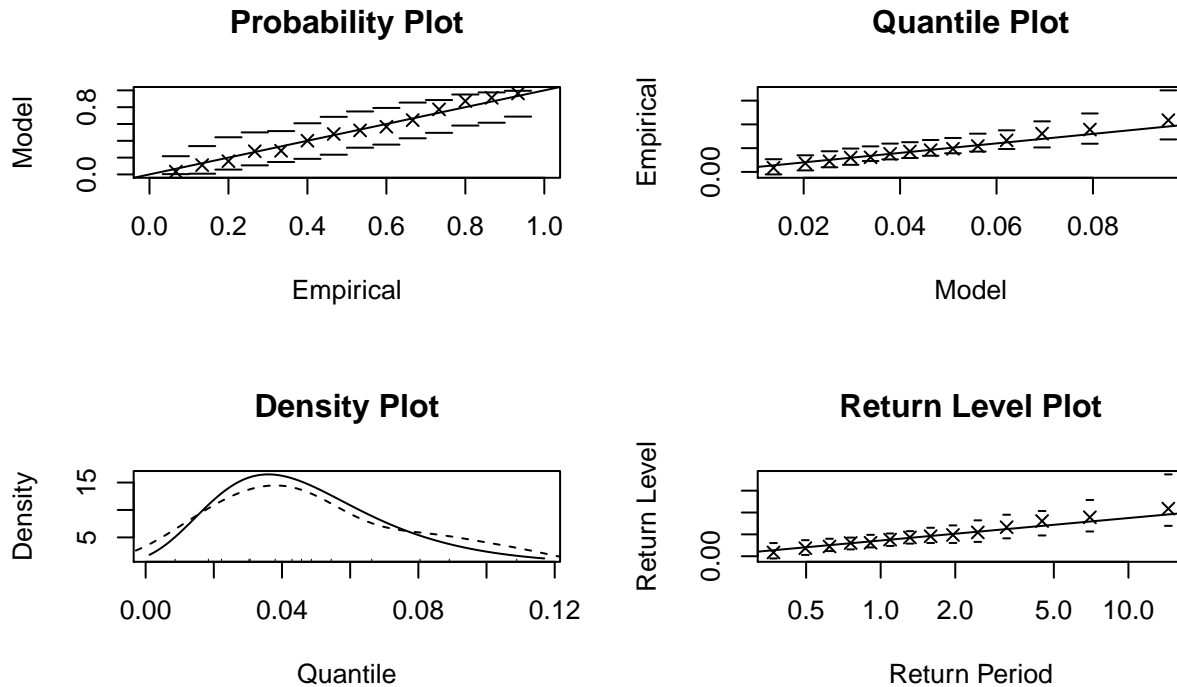**iii) derive a return level plot**

```
rl(model2,cex=0.3)
```

Return Level Plot

Our points lie within the confidence interval and

iv) get an estimate of the 100–year return level (or V aR1/100) and its associated confidence interval.

```
fit10 <- fgev(bydate, prob = 1/100)
par(mfrow = c(2,2))
plot(fit10)
```

3

Probability Plot

Quantile Plot

Density Plot
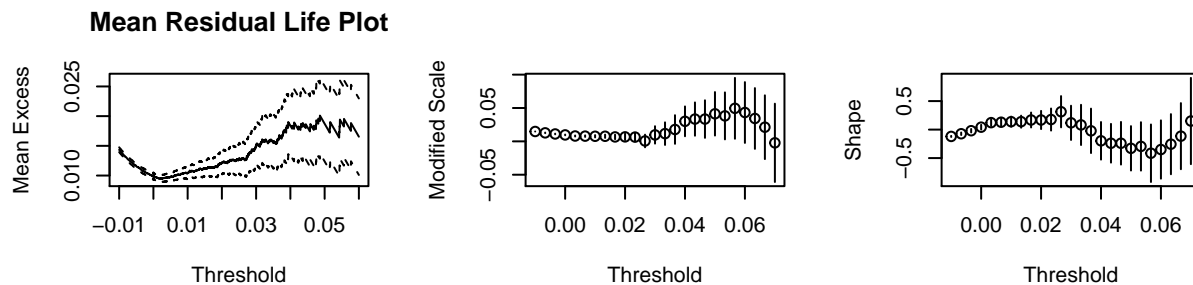

Return Level Plot

2. Peaks over Threshold Approach

i) choose appropriate threshold(s)

We don't observe serial dependence (as expected, since we took the difference of the log).

```
par(mfrow = c(1, 3))
mrlplot(nlogreturn$PFE.Close, c(-0.01, 0.06))
tcplot(nlogreturn$PFE.Close, c(-0.01, 0.07))
```

Mean Residual Life Plot

For the MRL plot, the choice of threshhold is based on drawing a straight line of any slope, starting at the right side, so it stays in the confidence region. With this approach we choose a threshold of approx 0.005. This is to satisfy the GPD assumption that the data is linear in u.

For the TCplot we choose a threshold by drawing a horizontal line from the right side of the graph to as far as we can of the left side, while remaining in the confidence region. With this approach we choose a threshold of 0.03. This is to satisfy the assumption that the shape parameter is constant.

We arrive at a threshold of 0.03 to maintain stability and linearity.

```
thr <- 0.03
plot( nlogreturn$Date[which(nlogreturn$PFE.Close > thr)], nlogreturn$PFE.Close[which(nlogreturn$PFE.Clos
points( nlogreturn$Date[which(nlogreturn$PFE.Close <= thr)], nlogreturn$PFE.Close[which(nlogreturn$PFE.C
abline(h=thr, col="red",cex=0.3)
```

4
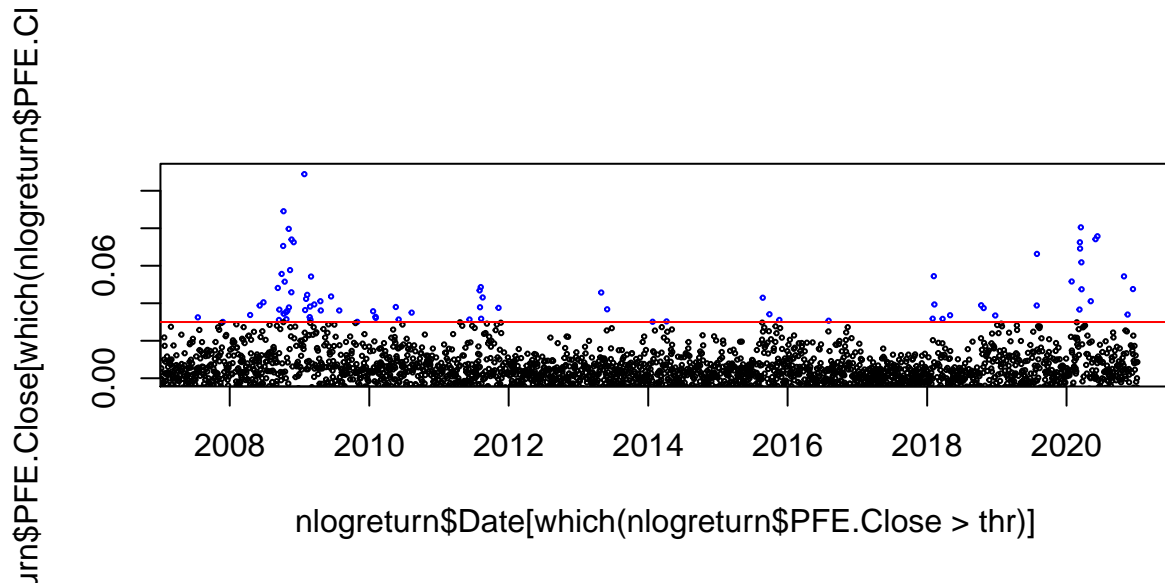
```r
pat <- mean(nlogreturn$PFE.Close > thr, na.rm = TRUE); pat
```
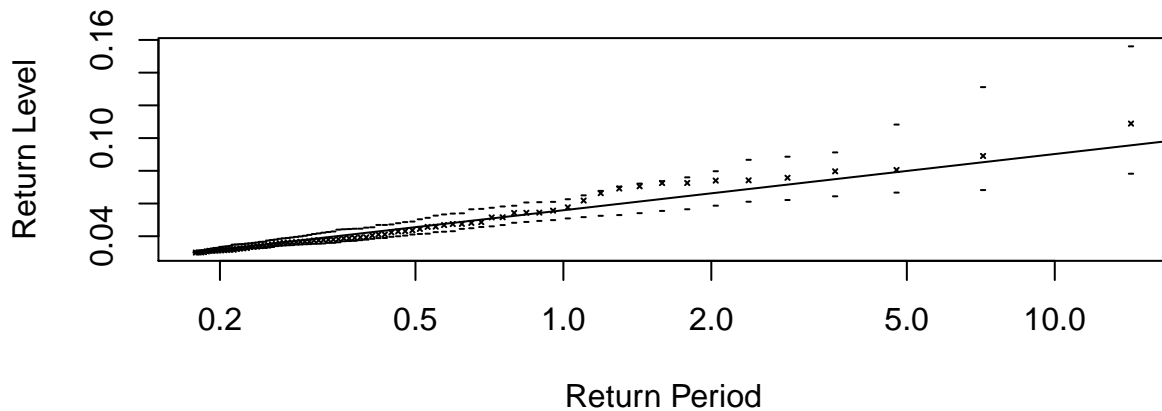
```
## [1] 0.02267574
```

ii) fit a stationary GPD

```r
fitted <- fpot(nlogreturn$PFE.Close, thr, npp = 250, model="gpd")
```

iii) derive a return level plot

```r
rl(fitted,cex=0.3)
```

Return Level Plot

Based on the return level plot our fit looks good, althought at 1.5 it gets very close to leaving confidence intervals.

iv) get an estimate of the 100–year return level (or V aR1/100) and its associated confidence interval.

```r
qgpd(1 - 1 / (pat * 250 * 100), thr, fitted$par["scale"], fitted$par["shape"], )
```

```
##      scale
## 0.1246241
```

```r
plot(profile(fpot(nlogreturn$PFE.Close, thr, mper = 10, npp = 250), "rlevel", xmin = 0.1, xmax = 0.5))
```

```
## [1] "profiling rlevel"
```


Profile Log–likelihood of Rlevel
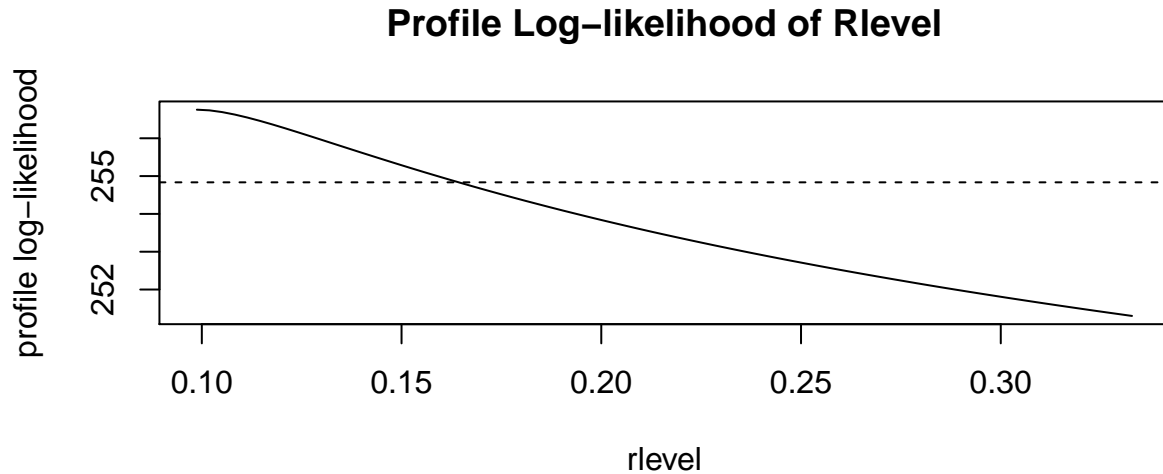
3. Compare the two previous modelling strategies and discuss your results.

```r
model2$estimate #block maxima approach
```

```
##          loc      locyear        scale        shape
##   0.03555007  -0.22758180   0.01696734   0.26783676
```

```r
fitted$estimate #peaks over threshhold
```

```
##        scale        shape
## 1.492452e-02 8.624304e-12
```

With the peaks over thresshold apporach our scale and shape are much smaller, however it is difficult to the first model relies on aggregated data. Without aggregated data, the fit was very poor.

```r
model2$std.err #block maxima approach
```

```
##          loc      locyear        scale        shape
## 0.005851011 0.112120786 0.004769325 0.358739563
```

```r
fitted$std.err #peaks over threshhold
```

```
##        scale        shape
## 0.002660731 0.122005688
```

The block maxima approach has higher variance.

In comparing the return level plots, we can see that the block maxima approach and the peaks over threshhold approach have similar estimates, however the peaks over threshold approach leads to much smaller confidence intervals.