金融危機下での日本の株式市場におけるジャンプの存在に関する検定

法政大学大学院工学研究科
茨田佳明

研究目的

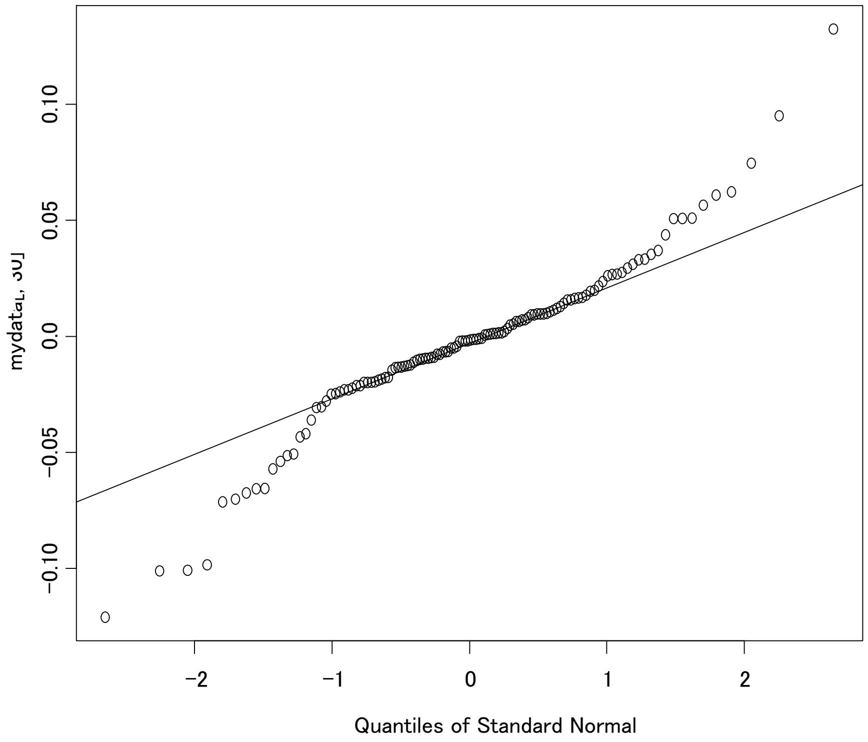
- 本研究の目的は、金融工学における資産価格の「（数理）モデルの選択」である。
- 金融機関における金融派生商品の価格決定には商品を生み出す資産過程に対してあるモデルを設ける。そのモデル次第で、資産価格変動に対してどのようなリスクを考慮に入れることができるかが決まる。
- 本研究では、2008年の金融危機の際に、資産価格がどのように変動したかを実証研究から捉え、そのリスクを考慮に入れる際にはどのようなモデルが良いかを提案することを目的とする。

研究の必要性

- 今まで広く用いられていたモデルでは、リスクを十分に考慮に入れられていない可能性がある。
- そのモデルとは、「株価の対数収益率は正規分布に従う」というものである。対数収益率とは、$S(t)$を時刻 t における株価とする場合 $\log S(t_i) - \log S(t_{i-1})$ を指す。
- 多くの先行研究において、実際の株価の対数収益率は正規分布には“従わない”という結果が得られている。

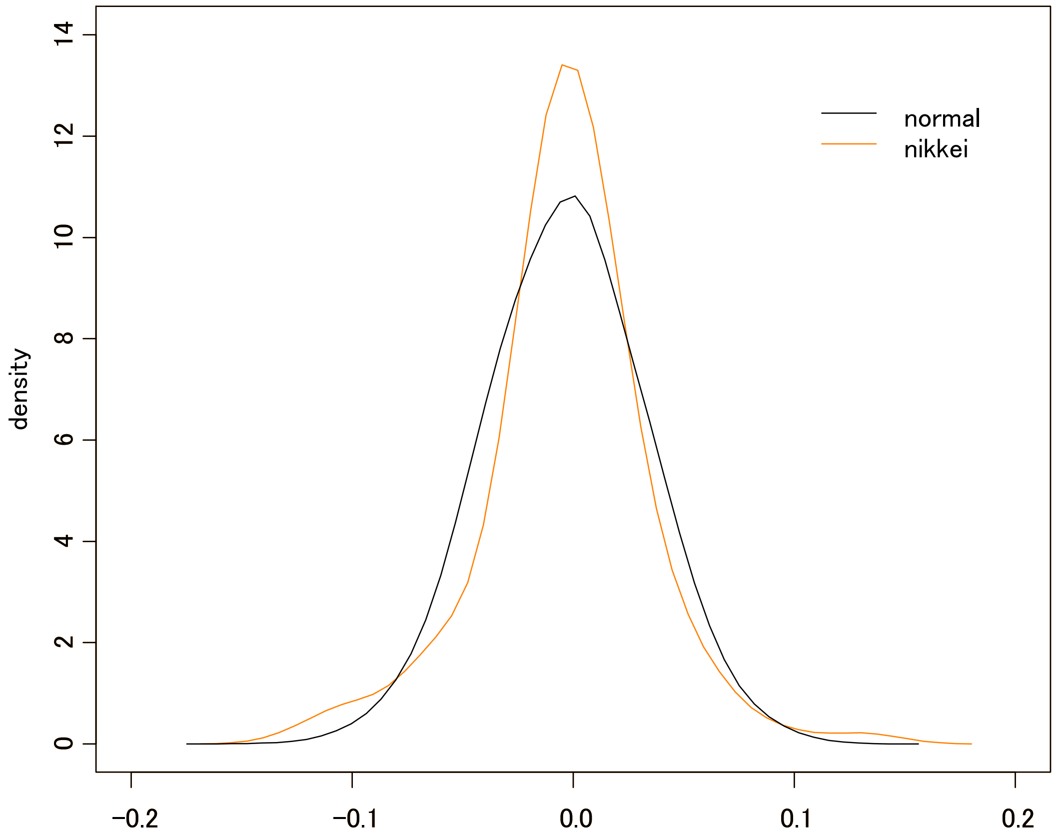
実際の株式市場(QQプロット)

- 日本の株式市場はどのようになっているのか。2008年度の日経平均株価を分析してみた。右の図は、対数収益率をQQプロットしたものである。
- 形状から正規分布よりも裾が厚い分布に従っていることが分かる。



実際の株式市場(分布)

- 右の図は、対数収益率の分布と正規分布を比較したものである。
- 日経平均株価(オレンジ)のほうが、裾が厚く、中央が尖っているのが分かる。



実際の株式市場(正規分布との適合性)

- コルモゴロフ・スミルノフ検定を用いて、対数収益率の標本分布が正規分布からのものかを検討した。(サンプル数124)右の表は、実際にS-PLUSで表示された結果である。
- 正規分布であるという帰無仮説を棄却した。(有意水準1%)

検定結果
ks = 0.1231, p-value = 0.0001
alternative hypothesis:
True cdf is not the normal distn. with estimated parameters
sample estimates:
mean of x standard deviation of x
-0.003385584
0.03630194

実際の株式市場（まとめ）

- 日経平均株価だけでなく、個別銘柄に対しても同様の分析を行い、同様の結果を得た。
- 対数収益率の分布は裾が厚く、中央が尖っていることから、金融機関で対数収益率が正規分布に従うモデルを選択すると、モデル選択の誤りによって損失を被る可能性があると考えられる。
- では、金融機関は具体的にどのようなモデルを設ければよいのか。
- その1つとして、ジャンプ（大きな変化）の導入である。

株式市場におけるジャンプとは

- 株価は「価格変化の強さ」を表すボラティリティと呼ばれるものを持っている。このボラティリティの値が大きければ、激しく株価が乱高下し、小さければ落ち着いた価格変動となる。
- 株式市場におけるジャンプとは、それまでのボラティリティの値からは想定できない「急激な価格変化」を言う。
- 今まで価格変動の大きさの度合いは、前者のボラティリティが決めているとされ、対数収益率が従う正規分布の分散を決定する主因とされてきた。しかし、対数収益率が正規分布に従わないことが指摘されるようになり、ボラティリティ以外に大きな変動を引き起こす要因があると注目されてきた。それが、ジャンプである。
- では、ジャンプは存在するのか。

ジャンプの検定方法

- これまでいくつかのジャンプの存在を確かめる検定方法が提案されてきた。
- 本研究では、Lee-Hannig[2010]*の方法を採り上げる。
- 他の検定方法と比較して、Lee-Hannigの方法の最大の利点は、特定の時点において、そのとき起きた価格変動がボラティリティによるものなのか、ジャンプによるものかを区別できるところにある。
- これは、存在だけではなくジャンプの大きさや頻度などより詳しい分析を可能にする。

- *S.S.Lee, Jan Hannig,"Detecting jumps from Levy jump diffusion processes", Journal of Financial Economics 96 [2010] 271-290

Lee-Hannigの検定方法（前提①）

- 株価S(t)が以下の確率過程に従うと仮定する。
- ジャンプが無い場合

$$d\log S(t) = \mu(t)dt + \sigma(t)dW(t) \qquad \cdots(1)$$

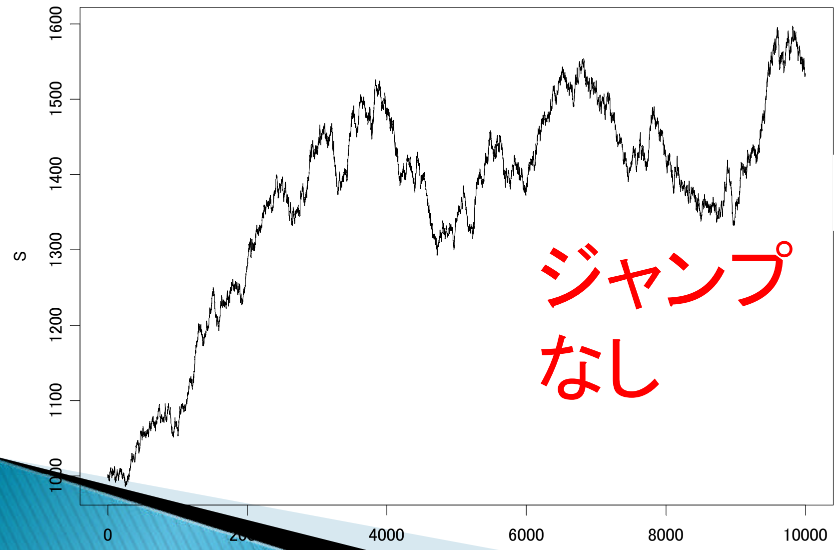
- μ はドリフト（株価の方向性（上昇・下降）を決める）。
- σ はボラティィリティ。
- W はブラウン運動（初期値が0で増分が独立同正規分布に従う確率過程）。
- t は時刻を表し、時間とともにそれぞれの変数が刻々と変化することを表している。

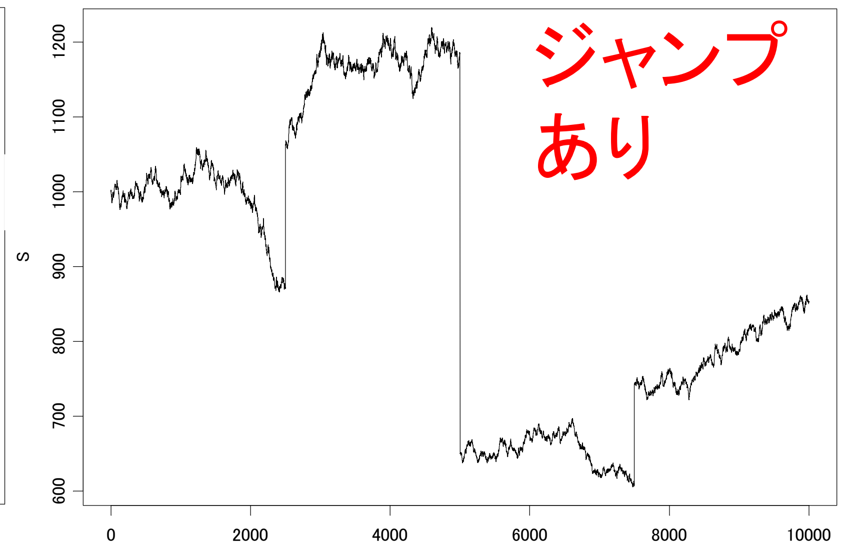
Lee-Hannigの検定方法（前提②）

- ジャンプがある場合

$$d\log S(t)=\mu(t)dt+\sigma(t)dW(t)+dL(t)$$

- L はレヴィジャンプ過程（ジャンプを発生させる確率過程）。
- 下のグラフは、それぞれの場合の疑似株価を表している。縦軸が価格、横軸が時刻である。





Lee-Hannigの検定方法(検定統計量)

- $t \in (t_{i-1}, t_i]$ にジャンプが存在するかどうかを検討する検定統計量を以下のように定める。$(\Delta t = t_i - t_{i-1})$

$$J(t) \equiv \frac{\log(S(t_i)/S(t_{i-1}))}{\hat{\sigma}(t_i)\Delta t^{1/2}}, t \in (t_{i-1}, t_i]$$

- ただし、$t \in (t_{i-1}, t_i]$において

$$\hat{\sigma}(t)^2 \equiv \frac{\Delta t^{-1}}{K} \sum_{j=i-K}^{i-1} \left(\log \frac{S(t_{j-m+1})}{S(t_{j-m})} \right)^2 1_{\left\{ \left| \log \frac{S(t_{j-m+1})}{S(t_{j-m})} \right| \leq g \Delta t^{\omega} \right\}}$$

- $g > 0, 0 < \omega < 1/2$ の定数である。
- K はウインドウサイズと呼ばれ、どのくらい過去のデータを使うかを表すものである。

検定統計量の意味

- 検定統計量は、「対数収益率」と「それまでの価格変動のボラティリティ（標準偏差）」の比を表している。

$$J(t) \equiv \frac{\boxed{\log(S(t_i)/S(t_{i-1}))}}{\boxed{\hat{\sigma}(t_i)\Delta t^{1/2}}}$$

- この比の大きさで、今までの価格変動とは違う変動（ジャンプ）を検討することになる。
- ウインドウサイズ K は「それまでの価格変動の標準偏差」にジャンプの要素が混入しない値を選択する必要がある。

Lee−Hannigの検定方法(分布定理)

- Lee−Hannig[2010]では次のことが述べられている。株価(1)式に従い $\Delta t \to 0$のとき、すなわち観測時間間隔を極限まで小さくしたとき、次が成り立つ:

$$\frac{\max\limits_{t\in(t_{i-1},t_i]\ for\ 0\le i\le n} |J(t)| - C_n}{B_n} \xrightarrow{D} \xi, \cdots (2)$$

- ただし、ξ は分布関数 $P(\xi \le x) = \exp(-e^{-x})$ を持つガンベル分布である。また、

$$C_n = (2\log n)^{1/2} - \frac{\log \pi + \log(\log n)}{2(2\log n)^{1/2}}, B_n = \frac{1}{(2\log n)^{1/2}}$$

n:観測したデータ数

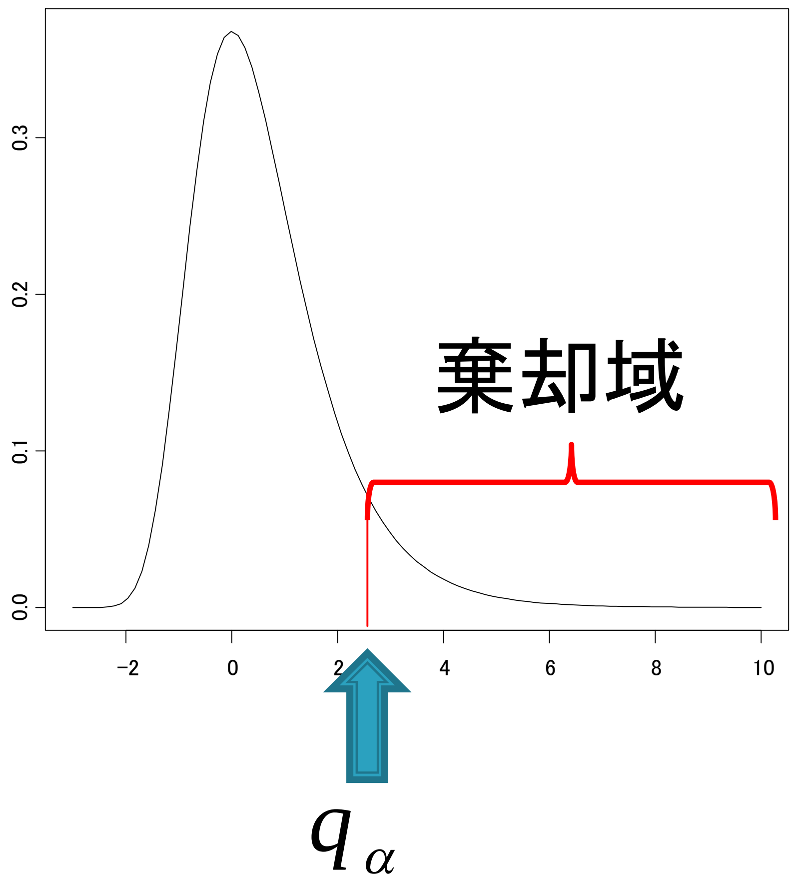
Lee-Hannigの検定方法（棄却域）

- q_αをξのαパーセント点とする。
- $$\frac{\max_{t \in (t_{i-1}, t_i] \text{ for } 0 \le i \le n} |J(t)| - C_n}{B_n} > q_\alpha$$

のとき、帰無仮説（ジャンプがない）を棄却する。

- 実際の検定では、$\max_{t \in (t_{i-1}, t_i] \text{ for } 0 \le i \le n} |J(t)|$を$|J(t_i)|$とする。

ガンベル分布の密度関数

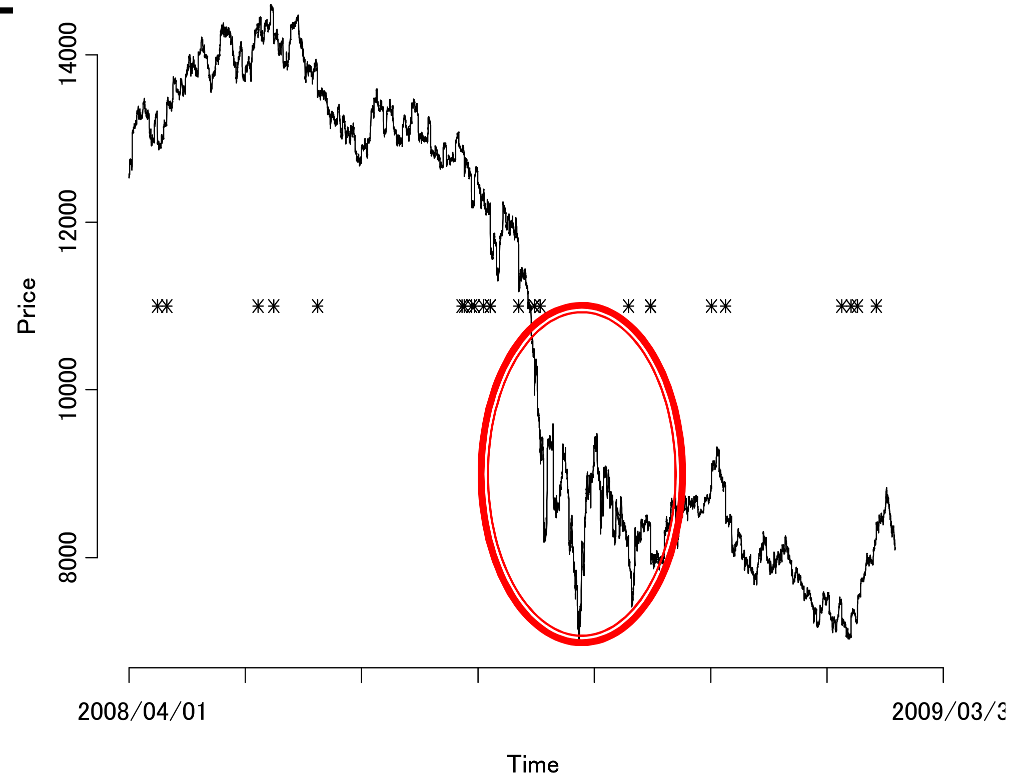


実データに対する検定(日経平均株価)

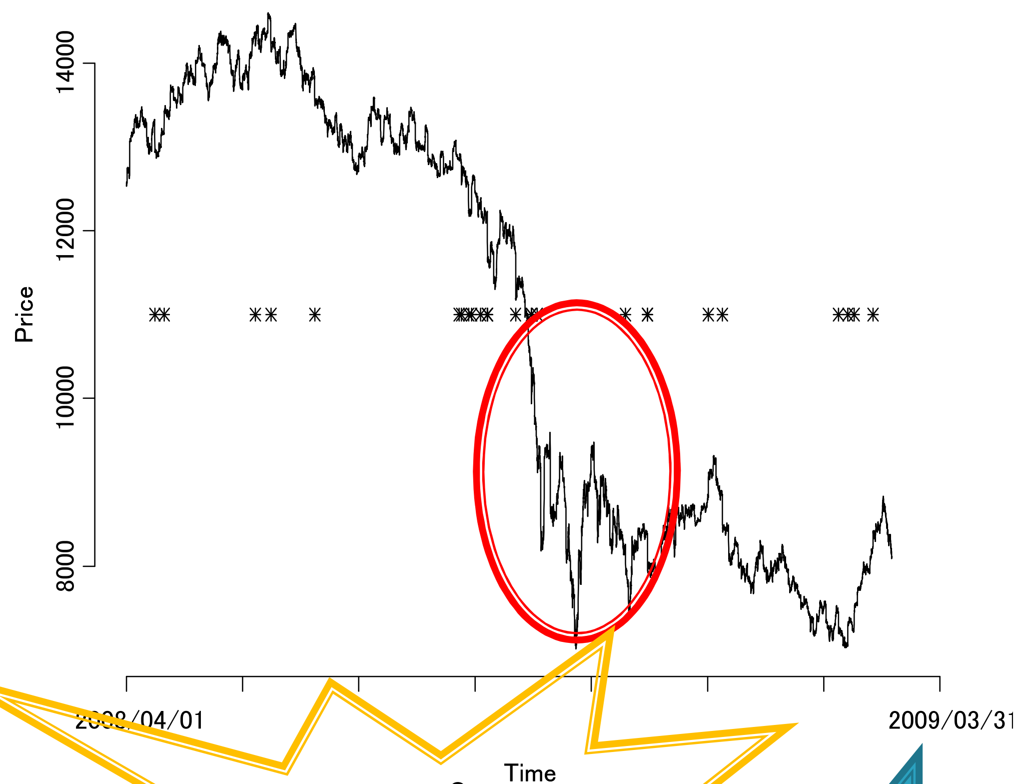
- データは日経NEED-TICKデータを使用した。ここでは、日経平均株価の2008年度(2008年4月1日から2009年3月31日)、2009年度(2009月4月1日から2010年3月31日)のデータに対する結果を示す。観測間隔は5分である。
- 取引がされている時間を検定対象とし、取引がされていない時間帯のデータは除いて検定を行った。
- 日本の株式市場では、取引可能時間は9時から11時、12時30分から15時の4時間30分である。
- 有意水準は1%で、ウインドウサイズ K はLee-Hannigが提案する値を採用した。

日経平均株価(ジャンプ時間)

- 右の図は、日経平均株価(2008年度)の推移とジャンプの位置(ポイント)を重ねたものである。
- 合計で26個のジャンプを検出した。
- 赤い部分でもう少しジャンプを検出してもよいと考えられる。
- 2009年度は合計で30個を検出した。



赤い部分で少ない理由

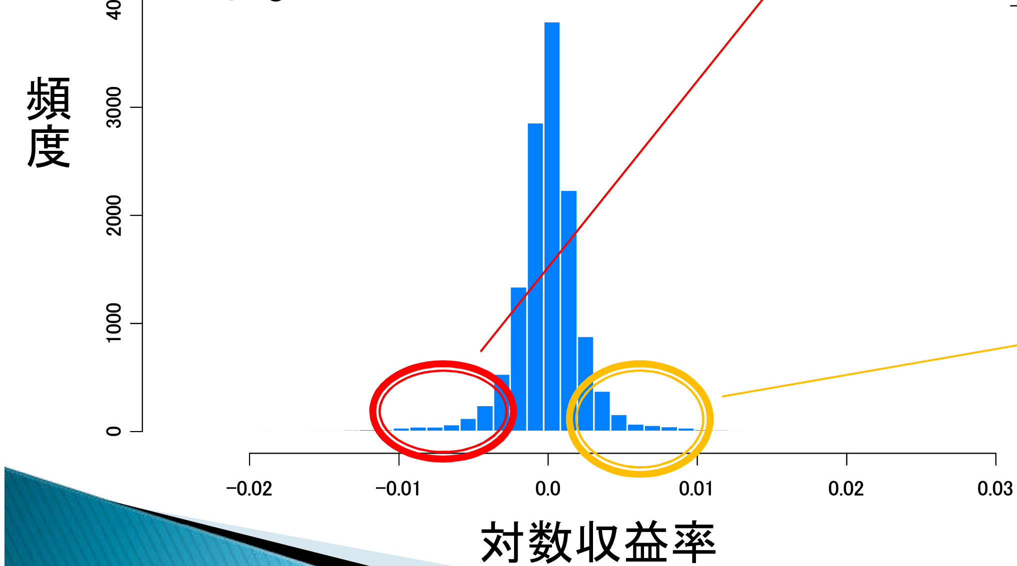


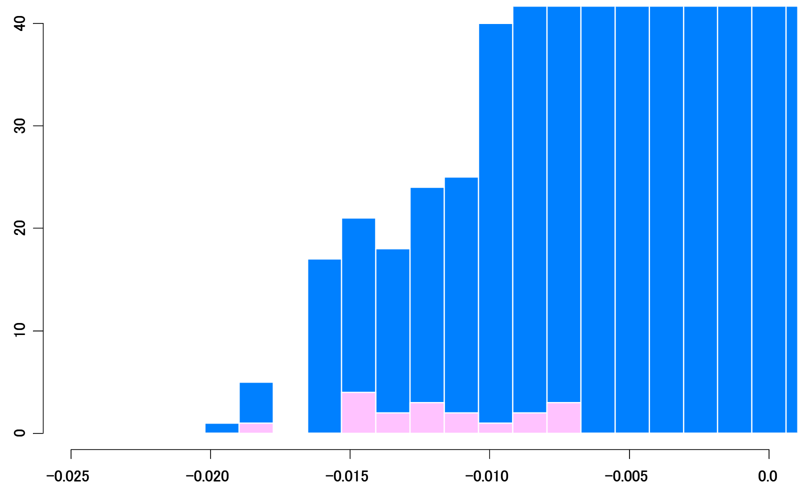
ジャンプと判定しにくくなる。

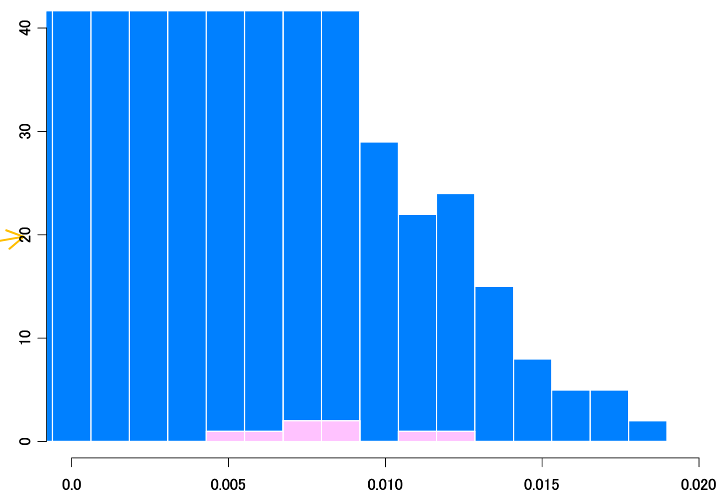
- 大きな変動が起こる。
- その大きな変動を$\hat{\sigma}(t)$の計算に使う。
- $\hat{\sigma}(t)$が大きくなる。
- 検定統計量の絶対値の値が小さくなる。

日経平均株価(ヒストグラム①)

- これらは、対数収益率(2008年度)をヒストグラムにしたものである。
- 青が全体、ピンクがジャンプと判定したものである。

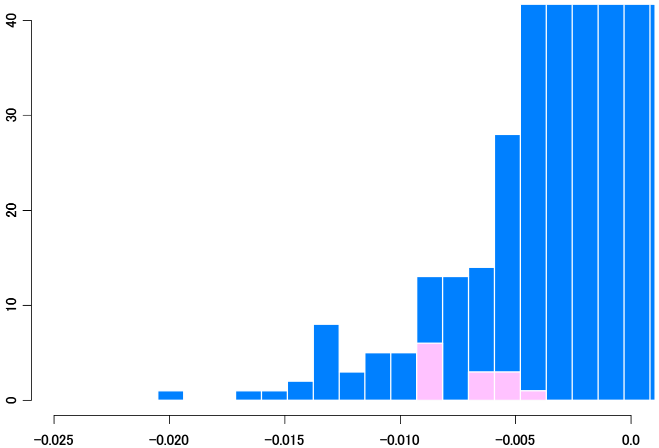


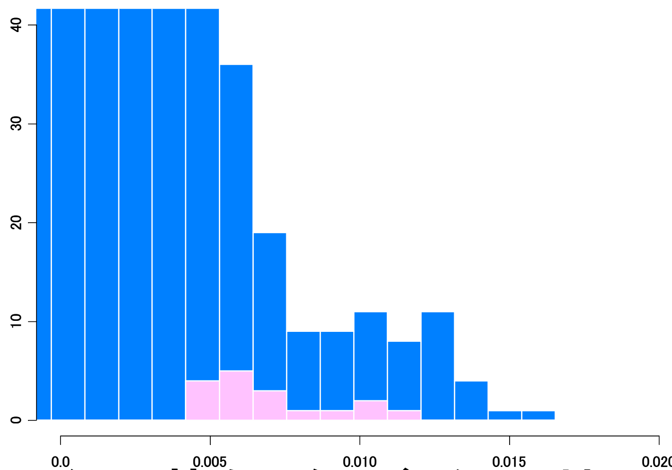




日経平均株価(ヒストグラム②)

- 前スライドから、プラス方向のジャンプが少ないように感じるが、これは検定方法の精度によるものだろうか。以下は2009年度のデータのヒストグラムである。



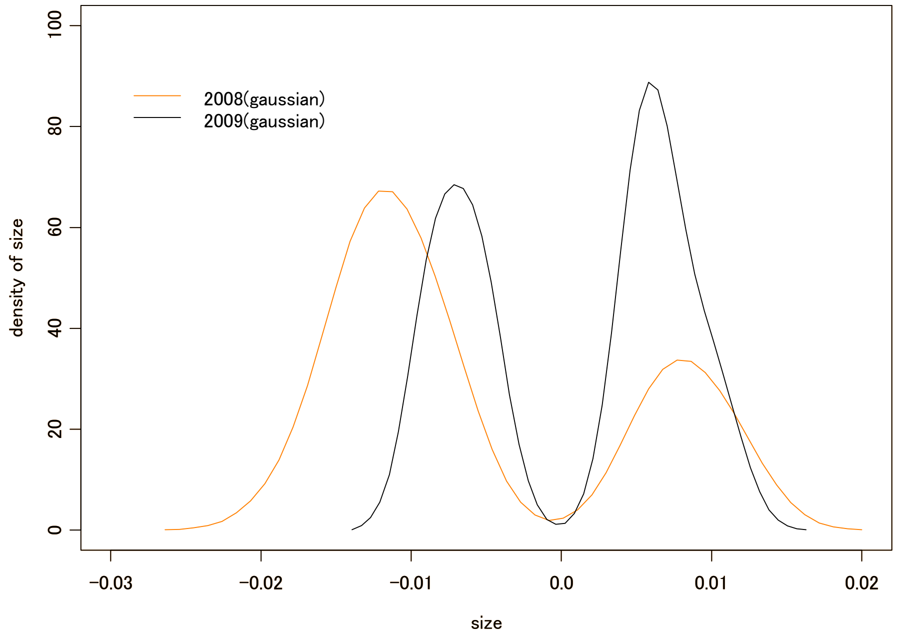


- 2008年度はリーマン・ショックなど下落傾向が強い状況だったことからそれが検出したジャンプの特徴に反映されたと考えられる。

日経平均株価(ジャンプサイズの分布)

- 右の図は、ジャンプサイズの分布を正規カーネル法で推定したものである。
- コルモゴロフ・スミルノフ検定の結果、同分布であるという帰無仮説を棄却した。(有意水準1%)
- 表はそれぞれの年度の平均と分散を

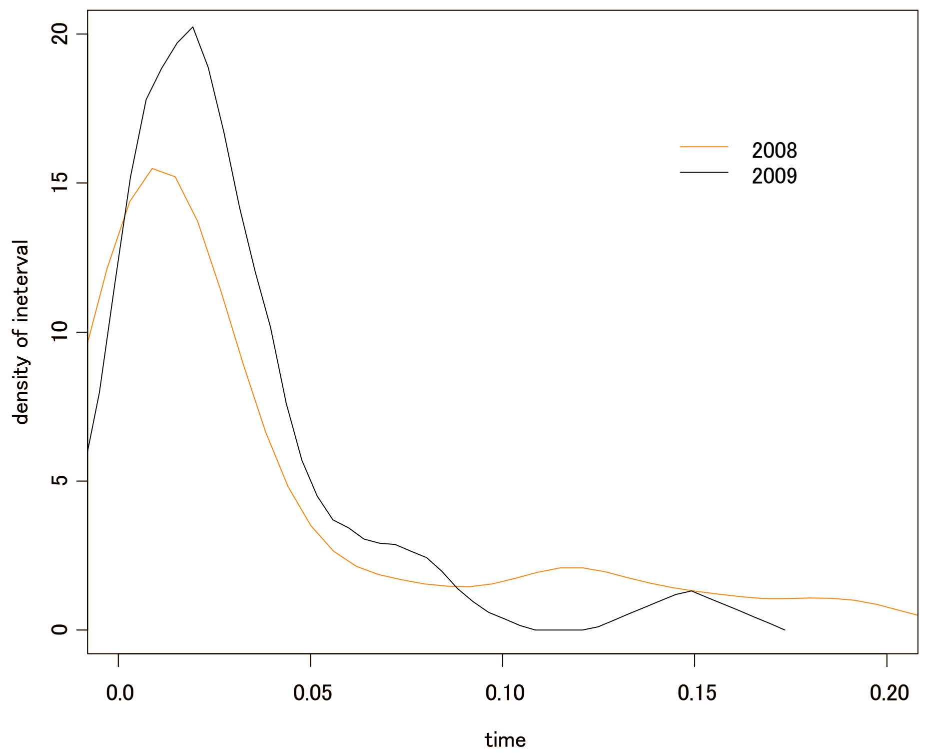
まとめたものである。



	2008年度	2009年度
平均	-5.48×10^{-3}	9.00×10^{-4}
分散	9.62×10^{-5}	5.20×10^{-5}

日経平均株価（頻度①）

- 右の図は、ジャンプ時間の間隔の分布を推定したものである。
- 1年を1と考えている。
- 多くのジャンプが約2週間の間隔で起きていることが分かる。
- コルモゴロフ・スミルノフ検定の結果、同分布に従うという帰無仮説を採択した。P値は0.51であった。



日経平均株価（頻度②）

- 前スライドの分布が理論分布に従うかどうかをコルモゴロフ・スミルノフ検定を用いて調べた。下記の表はそれぞれの理論分布に対するp値である。

	指数分布 (mean)	対数正規分布 (logmean,logsd)	ワイブル分布 （shape,scale）
2008年度	0.0153 (26.6)	0.135 (-4.69,2.35)	0.656 (0.593,0.0260)
2009年度	0.500 (34.9)	0.184 (-4.14,1.38)	0.426 (0.993,0.0286)

- 総合的には、頻度の分布にはワイブル分布が適当のように考えられる。

まとめ

- これまでの多くの金融機関は、ジャンプがないモデルを選択していたため価格変動のリスクを十分に考慮できていなかった可能性があった。
- Lee-Hannigの検定方法においては、日本の株式市場にジャンプが存在することとなり、ジャンプがあるモデルの選択を勧める根拠となる。
- ジャンプがあるモデルの選択に際しては、そのジャンプの特徴も考慮に入れなければならない。本研究より、市場の状況次第でジャンプの大きさの分布が変化することが分かり、金融危機下のような特殊な場合においては、マイナス方向のジャンプを多く含むモデルを選択する必要がある。
- また、ジャンプ頻度の分布において、多くの場合、指数分布を仮定してきたが、本研究より、対数正規分布やワイブル分布など他の分布の採用を検討する必要があることを指摘できる。

今後の課題

- 今回は比較的大きめのジャンプの検出を行った。Lee-hannig[2010]や他の先行研究では、比較的小さめのジャンプの存在にも考察を加えている。今後はそのようなジャンプに対しても研究をする必要がある。

S-PLUS code ①

- ここでは、Lee-Hannigの検定方法のコードを載せる。

```
#データ読み込み
s<-importData(file="5min(hosei).csv",type="ASCII")
hv<-importData(file="270minHVave.csv",type="ASCII")
hv<-hv[1,29:30]*3
omega<-0.47
count<-nrow(s)#列数
delta<-1/count

#株価のみのデータ行列
ss<-s[,31:32]
sss<-ss[2:nrow(ss),]
x<-rep(0,length(ss))
sss<-rbind(sss,x)

#対数収益率の計算
logs<-log(sss)-log(ss)

#σ^の計算
logs<-abs(logs)
logs<-logs[1:nrow(logs)-1,]
t<-s[,2]
t<-t[2:length(t)]
k=116
sigma2<-rep(0,nrow(ss))
sigma2<-t(t(sigma2))
for(j in 1:length(ss)){
alpha<-hv[,j]
alpha<-alpha*delta^omega
logsj<-rbind(t,logs[,j])
logsj<-t(logsj)
for(i in 1:nrow(ss)){
if(i==1){sigma<-NA}
else{
if(i<=k){x<-NA}
else{x<-logsj[(i-k):((i-k)+(k-1-1)),]
tt<-x[,1]
x<-x[(tt!=900)&(tt!=1230),2]
x<-x[x<=alpha]
x<-abs(x)*abs(x)
x<-sum(x)/(k*delta)}
sigma<-rbind(sigma,x)
}
}
sigma2<-cbind(sigma2,sigma)
}
sigma2<-sigma2[,2:3]
t<-s[,2]
sigma2<-cbind(t(t(t)),sigma2)
tt<-sigma2[,1]
sigma2<-sigma2[(tt!=900)&(tt!=1230),]
sigma2<-sigma2[,2:3]
```

S-PLUS code ②

```
#検定統計量の計算
R<-log(sss)-log(ss)
x<-rep(0,length(ss))
R<-rbind(x,R)
R<-R[1:nrow(ss),]
t<-s[,2]
R<-cbind(t,R)
tt<-R[,1]
R<-R[(tt!=900)&(tt!=1230),2:3]
L<-(R)/(sqrt(sigma2)*sqrt(delta))
#検定
beta<-4.6001
Cn<-sqrt(2*log(nrow(s)))-(log(3.14159265358979)+log(log(nrow(s))))/(2*(sqrt(2*log(nrow(s)))))
Sn<-1/(sqrt(2*log(nrow(s))))
L<-(abs(L)-Cn)/Sn
L<-ifelse((L<=beta)|(L==Inf),0,1)
```