Package `InfoTrad`

August 21, 2017

**Type** Package

**Title** Calculates the Probability of Informed Trading (PIN)

**Version** 1.2

**Date** 2017-08-21

**Author** Duygu Celik and Murat Tinic

**Maintainer** Murat Tinic <tinic@bilkent.edu.tr>


**Imports** nloptr

**License** GPL-3

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2017-08-21 05:59:33 UTC

---

**R topics documented:**

- InfoTrad-package
- EA
- EHO
- GAN
- LK
- YZ

Index 13
InfoTrad-package  

Calculates the Probability of Informed Trading (PIN)

Description


Author(s)

Duygu Celik and Murat Tinic

Maintainer: Murat Tinic <tinic@bilkent.edu.tr>

References


Description

It estimates PIN using Ersan and Alici (2016) modified clustering algorithm.

Usage

EA(data, likelihood = c("LK", "EH0"))
## S3 method for class 'EA_class'
print(obj)

Arguments

data Data frame with 2 variables
likelihood Character strings for likelihood algorithm. Default is "LK".
obj object variable

Details

Argument for data must be a data frame with 2 columns that only contain numbers. Not any other type. You do not have to give names to the columns. We will assign first one as "Buy" and second as "Sell", therefore you should put order numbers with respect to this order.

Value

Returns a list of parameter estimates (output)

alpha A Number
delta A Number
mu A Number
eb A Number
es A Number
LikVal A Number
PIN A Number

Warning

This function does not handle NA values. Therefore the datasets should not contain any missing value. This function do not conduct the iterative estimation procedure proposed in the same paper.

Author(s)

Duygu Celik and Murat Tinic
References


Examples

```r
# Sample Data
# Buy Sell
#1 350 382
#2 250 500
#3 500 463
#4 552 550
#5 163 200
#6 345 323
#7 847 456
#8 923 342
#9 123 578
#10 349 455

Buy=c(350,250,500,552,163,345,847,923,123,349)
Sell=c(382,500,463,550,200,323,456,342,578,455)
data=cbind(Buy,Sell)

# Parameter estimates using the LK factorization of Lin and Ke (2011)
# with the modified clustering algorithm of Ersan and Alici (2016).
# Default factorization is set to be "LK"

result=EA(data)
print(result)

# Alpha: 0.9511418
# Delta: 0.2694005
# Mu: 76.7224
# Epsilon_b: 493.7045
# Epsilon_s: 377.4877
# Likelihood Value: 43973.71
# PIN: 0.07728924

# Parameter estimates using the EHO factorization of Easley et. al. (2010)
# with the modified clustering algorithm of Ersan and Alici (2016).

result=EA(data,likelihood="EHO")
print(result)

# Alpha: 0.9511418
# Delta: 0.2694005
# Mu: 76.7224
# Epsilon_b: 493.7045
# Epsilon_s: 377.4877
```
Description
The function calculates the likelihood factorization of Easley et. al. (2010) and computes parameters for estimation of PIN value.

Usage
EHO(data, fixed = c(FALSE, FALSE, FALSE, FALSE, FALSE))

Arguments
data
Data frame with 2 variables
fixed
Initial values for parameters in the following order: alpha, delta, mu, epsilon_b, epsilon_s

Details
In order to use EHO’s return in optimization functions, please omit second argument. With this way, EHO will return a function instead of a value. Moreover, argument for data must be a data frame with 2 columns that contain numbers. Not any other type.

Value
LK_out
Returns an optim() object including parameter estimates for the likelihood factorization of Easley et. al. (2010)

Warning
This function does not handle NA values. Therefore the datasets should not contain any missing values.

Author(s)
Duygu Celik and Murat Tinic

References
Examples

# Sample Data
# Buy Sell
#1  350  382
#2  250  500
#3  500  463
#4  552  550
#5  163  200
#6  345  323
#7  847  456
#8  923  342
#9  123  578
#10 349  455

Buy<-c(350, 250, 552, 163, 345, 847, 923, 123, 349)
Sell<-c(382, 500, 463, 550, 200, 323, 456, 342, 578, 455)
data=cbind(Buy,Sell)

# Initial parameter values
# par0 = (alpha, delta, mu, epsilon_b, epsilon_s)
par0 = c(0.5, 0.5, 300, 400, 500)

# Call EHO function
EHO_out = EHO(data)
model = optim(par0, EHO_out, gr = NULL, method = c("Nelder-Mead"), hessian = FALSE)

## Parameter Estimates
model$par[1] # Estimate for alpha
# [1] 0.9111102
model$par[2] # Estimate for delta
# [1] 0.0001231429
model$par[3] # Estimate for mu
# [1] 417.1497
model$par[4] # Estimate for eb
# [1] 336.075
model$par[5] # Estimate for es
# [1] 466.2539

## Estimate for PIN
(model$par[1]*model$par[3])/((model$par[1]*model$par[3])+model$par[4]+model$par[5])
# [1] 0.3214394

###

GAN

GAN et al.(2015) Clustering based PIN Estimates

Description

It estimates PIN using hierarchical agglomerative clustering.
Usage

```r
GAN(data, likelihood = c("LK", "EHO"))
## S3 method for class 'GAN_class'
print(obj)
```

Arguments

- `data`: Data frame with 2 variables
- `likelihood`: Character strings for likelihood algorithm. Default is "LK".
- `obj`: object variable

Details

Argument for data must be a data frame with 2 columns that only contain numbers. Not any other type. You do not have to give names to the columns. We will assign first one as "Buy" and second as "Sell", therefore you should put order numbers with respect to this order. This package uses the hclust() function of Mullner (2013) to cluster the data at default settings.

Value

Returns a list of parameter estimates (output)

- `alpha`: A Number
- `delta`: A Number
- `mu`: A Number
- `eb`: A Number
- `es`: A Number
- `LikVal`: A Number
- `PIN`: A Number

Warning

This function does not handle NA values. Therefore, the dataset should not contain any missing values.

Author(s)

Duygu Celik and Murat Tinic

References


Examples

```r
# Sample Data
# Buy Sell
#1  350  382
#2  250  500
#3  500  463
#4  552  550
#5  163  200
#6  345  323
#7  847  456
#8  923  342
#9  123  578
#10 349  455

Buy<-c(350,250,500,552,163,345,847,923,123,349)
Sell<-c(382,500,463,550,200,323,456,342,578,455)
data<-cbind(Buy,Sell)

# Parameter estimates using the LK factorization of Lin and Ke (2011)
# with the algorithm of Gan et. al. (2015).
# Default factorization is set to be "LK"

result=GANE(data)
print(result)

# Alpha: 0.3999998
# Delta: 0
# Mu: 442.1667
# Epsilon_b: 263.3333
# Epsilon_s: 424.9
# Likelihood Value: 44371.84
# PIN: 0.2044464

# Parameter estimates using the EHO factorization of Easley et. al. (2010)
# with the algorithm of Gan et. al. (2015)

result=GANE(data, likelihood="EHO")
print(result)

# Alpha: 0.3230001
# Delta: 0.4780001
# Mu: 481.3526
# Epsilon_b: 356.6359
# Epsilon_s: 313.136
# Likelihood Value: Inf
# PIN: 0.1884001
```

**Likelihood factorization of Lin and Ke (2011) - LK Factorization**
Description

The function calculates the likelihood factorization of Lin and Ke (2011) and computes parameters for estimation of PIN value.

Usage

LK(data, fixed = c(FALSE, FALSE, FALSE, FALSE, FALSE))

Arguments

data Data frame with 2 variables
fixed Initial values for parameters in the following order: alpha, delta, mu, epsilon_b, epsilon_s

Details

In order to use LK's return in optimization functions, please omit second argument. With this way, LK will return a function instead of a value. Moreover, argument for data must be a data frame with 2 columns that contain numbers. Not any other type.

Value

LK_out Returns an optim() object including parameter estimates for the likelihood factorization of Lin and Ke (2011)

Warning

This function does not handle NA values. Therefore the datasets should not contain any missing value

Author(s)

Duygu Celik and Murat Tinic

References


Examples

# Sample Data
# Buy Sell
#1 350 382
#2 250 500
#3 500 463
#4 552 550
#5 163 200
#6 345 323
#7 847 456
# Initial parameter values
# par0 = (alpha, delta, mu, epsilon_b, epsilon_s)
par0 = c(0.5, 0.5, 300, 400, 500)

# Call LK function
LK_out = LK(data)
model = optim(par0, LK_out, gr = NULL, method = c("Nelder-Mead"), hessian = FALSE)

## Parameter Estimates
model$par[1] # Estimate for alpha
# [1] 0.480277
model$par[2] # Estimate for delta
# [1] 0.830850
model$par[3] # Estimate for mu
# [1] 315.259805
model$par[4] # Estimate for eb
# [1] 296.862318
model$par[5] # Estimate for es
# [1] 434.3046

## Estimate for PIN
# [1] 0.178391

---

**Yan and Zhang (2012) Grid-Search based PIN Estimates**

**Description**


**Usage**

YZ(data, likelihood = c("LK", "EMO"))

# S3 method for class 'YZ_class'
print(obj)

**Arguments**

- **data**: Data frame with 2 variables
- **likelihood**: Character strings for likelihood algorithm. Default is "LK".
- **obj**: object variable
Details
Argument for data must be a data frame with 2 columns that only contain numbers. Not any other type. You do not have to give names to the columns. We will assign first one as "Buy" and second as "Sell", therefore you should put order numbers with respect to this order.

Value
Returns a list of parameter estimates (output)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>alpha</td>
<td>A Number</td>
</tr>
<tr>
<td>delta</td>
<td>A Number</td>
</tr>
<tr>
<td>mu</td>
<td>A Number</td>
</tr>
<tr>
<td>eb</td>
<td>A Number</td>
</tr>
<tr>
<td>es</td>
<td>A Number</td>
</tr>
<tr>
<td>LikVal</td>
<td>A Number</td>
</tr>
<tr>
<td>PIN</td>
<td>A Number</td>
</tr>
</tbody>
</table>

Warning
This function does not handle NA values. Therefore the datasets should not contain any missing value

Author(s)
Duygu Celik and Murat Tinic

References

Examples
```r
# Sample Data
# Buy Sell
#1 350 382
#2 250 500
#3 500 463
#4 552 550
#5 163 200
#6 345 323
#7 847 456
#8 923 342
#9 123 578
#10 349 455

Buy<-c(350,250,500,552,163,345,847,923,123,349)
Sell<-c(382,500,463,550,200,323,456,342,578,455)
```
data<-cbind(Buy,Sell)

# Parameter estimates using the LK factorization of Lin and Ke (2011)
# with the algorithm of Yan and Zhang (2012).
# Default factorization is set to be "LK"

result=YZ(data)
print(result)

# Alpha: 0.3999999
# Delta: 0
# Mu: 442.1667
# Epsilon_b: 263.3333
# Epsilon_s: 424.9
# Likelihood Value: 44371.84
# PIN: 0.2004457

# Parameter estimates using the EHO factorization of Easley et. al. (2010)
# with the algorithm of Yan and Zhang (2012).

result=YZ(data,likelihood="EHO")
print(result)

# Alpha: 0.9000001
# Delta: 0.9000001
# Mu: 489.1111
# Epsilon_b: 396.1803
# Epsilon_s: 28.72002
# Likelihood Value: Inf
# PIN: 0.3321033
Index

*Topic Probability of Informed Trading (PIN), LK factorization, EHO factorization, YZ Algorithm, CPIN

InfoTrad-package, 2

EA, 3
EA_class (EA), 3
EH0, 5

GAN, 6
GAN_class (GAN), 6

InfoTrad (InfoTrad-package), 2
InfoTrad-package, 2

LK, 8

print (EA), 3
print.GAN_class (GAN), 6
print.YZ_class (YZ), 10

YZ, 10
YZ_class (YZ), 10