# Package ‘pbo’

August 29, 2016

- **Type**: Package
- **Title**: Probability of Backtest Overfitting
- **Version**: 1.3.4
- **Date**: 2014-05-18
- **Author**: Matt Barry &lt;mrb@softisms.com&gt;
- **Maintainer**: Matt Barry &lt;mrb@softisms.com&gt;
- **Description**: Following the method of Bailey et al., computes for a collection of candidate models the probability of backtest overfitting, the performance degradation and probability of loss, and the stochastic dominance.
- **License**: MIT + file LICENSE
- **URL**: https://github.com/mrbcuda/pbo
- **BugReports**: https://github.com/mrbcuda/pbo/issues
- **Depends**: utils, lattice
- **Suggests**: PerformanceAnalytics, foreach, grid, latticeExtra, testthat, doParallel, knitr
- **VignetteBuilder**: knitr
- **NeedsCompilation**: no
- **Repository**: CRAN
- **Date/Publication**: 2014-05-31 01:49:51

## R topics documented:

- pbo-package .................................................. 2
- dotplot.pbo .................................................. 2
- histogram.pbo .............................................. 3
- pbo ........................................................... 4
- pbo_show_config ........................................... 5
- xyplot.pbo .................................................... 5

Index 7
pbo-package

*Probability of backtest overfitting.*

**Description**

Computes the probability of backtest overfitting

**Details**

Implements algorithms for computing the probability of backtest overfitting, performance degradation and probability of loss, and first- and second-order stochastic dominance, based on the approach specified in Bailey et al., September 2013. Provides a collection of pre-configured plots based on lattice graphics.

**Author(s)**

Matt Barry <mrb@softisms.com>

**References**


dotplot.pbo

*PBO in-sample selection dot plot.*

**Description**

Draws an annotated dot plot of study selection sorted by in-sample selection frequency.

**Usage**

```r
## S3 method for class 'pbo'
dotplot(x, data = NULL, main, xlab = "Sorted Study Number (N)", ylab = "IS Selection Frequency", show_config = TRUE, show_grid = TRUE, sel_threshold = 0, ...)
```

**Arguments**

- `x` a pbo object as returned by `pbo`.
- `data` should not be used
- `main` plot title, default computed internally, passed to `dotplot`.
- `xlab` x-axis label with default, passed to `dotplot`.
- `ylab` y-axis label with default, passed to `dotplot`.
**histogram.pbo**

```
show_config  whether to show the study dimension annotations, default TRUE
show_grid    whether to show the grid panel, default TRUE
sel_threshold the minimum in-sample frequency subsetting threshold, default 0; selection frequencies at or below this value will be omitted
...
other parameters as passed to dotplot.
```

**See Also**

pbo, histogram.pbo, xyplot.pbo

---

**histogram.pbo**  
*PBO rank logits histogram.*

---

**Description**

Draws an annotated histogram of PBO rank logits.

**Usage**

```r
## S3 method for class 'pbo'
histogram(x, data = NULL, show_pbo = TRUE,
           show_regions = TRUE, show_config = TRUE, col_bar = "#cc99cc",
           col_line = "#3366cc", ...)
```

**Arguments**

- `x`  
an object of class pbo as returned by pbo.
- `data`  
should not be used
- `show_pbo`  
whether to show the PBO value annotation, default TRUE
- `show_regions`  
whether to show the overfit region annotations, default TRUE
- `show_config`  
whether to show the study dimension annotations, default TRUE
- `col_bar`  
histogram bar fill color passed to histogram panel
- `col_line`  
density plot line color passed to density plot panel
- `...`  
other parameters passed to histogram, densityplot, or panel.abline.

**Details**

Uses lattice function histogram, densityplot, and panel.abline panels together with class-specific annotations.

**See Also**

pbo, dotplot.pbo, xyplot.pbo
Description

Performs the probability of backtest overfitting computations.

Usage

```r
pbo(m, s = 4, f = NA, threshold = 0, inf_sub = 6,
    allow_parallel = FALSE)
```

Arguments

- `m` a T x N data frame of returns, where T is the samples per study and N is the number of studies.
- `s` the number of subsets of m for CSCV combinations; must evenly divide m
- `f` the function to evaluate a study’s performance; required
- `threshold` the performance metric threshold (e.g. 0 for Sharpe, 1 for Omega)
- `inf_sub` infinity substitution value for reasonable plotting
- `allow_parallel` whether to enable parallel processing, default FALSE

Details

This function performs the probability of backtest overfitting calculation using a combinatorially-symmetric cross validation (CSCV) approach.

Value

object of class `pbo` containing list of PBO calculation results and settings

References


Examples

```r
## Not run:
require(pbo)
require(PerformanceAnalytics)
N <- 100
T <- 1000
S <- 8
m <- data.frame(matrix(rnorm(N*T,mean=0, sd=1), 
                        nrow=T, ncol=N, byrow=TRUE), 
                dimnames=list(1:T,1:N)),
```
\documentclass{article}

\begin{document}

\section*{pbo\_show\_config}

\begin{verbatim}
check.names=FALSE)
p <- pbo(m, s, fOmegathreshold=1)
## End(Not run)
\end{verbatim}

\textit{pbo\_show\_config} \hspace{1cm} \textit{W}rites grid text to a default predetermined location.

\section*{Description}

\begin{quote}
W\v{r}ites grid text to a default predetermined location.
\end{quote}

\section*{Usage}

\begin{verbatim}
pbo\_show\_config(p)
\end{verbatim}

\section*{Arguments}

\begin{verbatim}
p \hspace{1cm} \text{an object of class pbo as returned by pbo.}
\end{verbatim}

\section*{Note}

\begin{quote}
Meant for internal use only.
\end{quote}

\section*{xyplot.pbo}

\textit{PBO xy-plots}

\section*{Description}

\begin{quote}
Draws an annotated plot of performance degradation and probability of loss.
\end{quote}

\section*{Usage}

\begin{verbatim}
## S3 method for class 'pbo'
yxplot(x, data = NULL, plotType = "cscv", show_eqn = TRUE,
\text{show_threshold = TRUE, show_config = TRUE, show_rug = TRUE,}
\text{show_prob = TRUE, show_grid = TRUE, increment = 0.01,}
\text{osr_threshold = 0, sel_threshold = 0, xlab, ylab, main, lwd = 1,}
\text{ylab_left, ylab_right, col_bar, col_line, col_sd1 = "#3366cc",}
\text{col_sd2 = "#339999", lty_sd = c(1, 2, 4), ...}
\end{verbatim}

\end{document}
Arguments

- **x**  
a pbo object as returned by `pbo`.
- **data**  
should not be used
- **plotType**  
one of cscv, degradation, dominance, pairs, ranks or selection.
- **col_bar**  
histogram bar fill color
- **col_line**  
density plot line color
- **col_sd1**  
color of two first-order stochastic dominance lines
- **col_sd2**  
color of the single second-order stochastic dominance line
- **xlab**  
x-axis label, default computed if not provided
- **ylab**  
y-axis label, default computed if not provided
- **main**  
plot title, default computed if not provided
- **lwd**  
line width, default 1, passed to panels and legends
- **lty_sd**  
line type array for stochastic dominance plot, e.g. c(2,3,5)
- **ylab_left**  
dominance plot left-hand axis label
- **ylab_right**  
dominance plot right-hand axis label
- **increment**  
stochastic dominance distribution generator increment, e.g. 0.1 steps
- **osr_threshold**  
out-of-sample rank threshold for filtering, default 0
- **sel_threshold**  
selection frequency threshold for filtering, default 0
- **show_eqn**  
whether to show the line equation annotation, default TRUE
- **show_threshold**  
whether to show the probability of loss annotation, default TRUE
- **show_config**  
whether to show the study dimension annotations, default TRUE
- **show_grid**  
whether to show the panel grid, default TRUE
- **show_prob**  
whether to show the probability value in dominance plot, default TRUE
- **show_rug**  
whether to show scatter rugs near the axes, default TRUE
- **...**  
other parameters passed to `xyplot` or its panels

Details

Provides several variations of xy-plots suitable for presentation of PBO analysis results. Use the `plotType` argument to indicate which variation or result to plot:

- The cscv type shows in-sample and out-of-sample results by CSCV iteration case (default).
- The degradation type shows the performance degradation regression fit results and the probability of loss.
- The dominance type shows the results of the first-order and second-order stochastic dominance analysis using two axes.
- The pairs type shows the in-sample and out-of-sample case selections.
- The ranks type shows the sorted performance ranks results.
- The selection type shows the case selection frequencies.

See Also

- `pbo`, `histogram.pbo`, `xyplot.pbo`
## Index

<table>
<thead>
<tr>
<th>Topic</th>
<th>Functions</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CSCV</strong></td>
<td>pbo, pbo-package</td>
<td>4, 2</td>
</tr>
<tr>
<td></td>
<td>xyplot.pbo</td>
<td>5</td>
</tr>
<tr>
<td><strong>PBO</strong></td>
<td>xyplot.pbo</td>
<td>5</td>
</tr>
<tr>
<td><strong>PBO</strong></td>
<td>pbo, pbo-package</td>
<td>4, 2</td>
</tr>
<tr>
<td></td>
<td>backtest</td>
<td>pbo, 2</td>
</tr>
<tr>
<td></td>
<td>dotplot.pbo</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>pbo-package</td>
<td>2</td>
</tr>
<tr>
<td><strong>overfitting</strong></td>
<td>dotplot.pbo</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>pbo, pbo-package</td>
<td>2</td>
</tr>
<tr>
<td><strong>pbo</strong></td>
<td>dotplot.pbo</td>
<td>2</td>
</tr>
<tr>
<td><strong>probability</strong></td>
<td>pbo, pbo-package</td>
<td>4, 2</td>
</tr>
<tr>
<td><strong>densityplot</strong></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td><strong>dotplot</strong></td>
<td>2, 3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>dotplot.pbo</td>
<td>2</td>
</tr>
<tr>
<td><strong>histogram</strong></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>histogram.pbo</td>
<td>3</td>
</tr>
<tr>
<td><strong>panel.abline</strong></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td><strong>pbo</strong></td>
<td>2, 3, 4, 5, 6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>pbo-package</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>pbo_show_config</td>
<td>5</td>
</tr>
<tr>
<td><strong>xyplot</strong></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td><strong>xyplot.pbo</strong></td>
<td></td>
<td>5</td>
</tr>
</tbody>
</table>