Package 'EmpiricalCalibration'

December 19, 2014

Type Package

Title Routines for performing empirical calibration of observational study estimates

Version 1.0.0

Date 2014-07-18

Author Martijn Schuemie

Maintainer Martijn Schuemie <schuemie@ohdsi.org>

Description Routines for performing empirical calibration of observational study estimates. By using a set of negative control hypotheses we can estimate the empirical null distribution of a particular observational study setup. This empirical null distribution can be used to compute a calibrated p-value, which reflects the probability of observing an estimated effect size when the null hypothesis is true taking both random and systematic error into account.

Imports ggplot2,MASS

License Apache License

R topics documented:

`calibrateP` *Calibrate the p-value*

Description

`calibrateP` computes calibrated p-values using the fitted null distribution

Usage

```
calibrateP(logRr, seLogRr, null, pValueConfidenceInterval = FALSE)
```

Arguments

`logRr`	A numeric vector of one or more effect estimates on the log scale
`seLogRr`	The standard error of the log of the effect estimates. Hint: often the standard error = (log(<lower bound 95 percent confidence interval>) - log(<effect estimate>))/qnorm(0.025)
`null`	An object of class `null` created using the `fitNull` function
`pValueConfidenceInterval`	If true, computes the 95 percent confidence interval of the calibrated p-value

Details

This function computes a calibrated two-sided p-value as described in Schuemie et al (2014).

Value

A two-sided calibrated p-value.

References

Schuemie MJ, Ryan PB, Dumouchel W, Suchard MA, Madigan D. Interpreting observational studies: why empirical calibration is needed to correct p-values. Statistics in Medicine 33(2):209-18,2014

Examples

```
data(sccs)
negatives <- sccs[sccs$groundTruth == 0,]
null <- fitNull(negatives$logRr,negatives$seLogRr)
positive <- sccs[sccs$groundTruth == 1,]
calibrateP(positive$logRr,positive$seLogRr, null)
```

`caseControl` *Odds ratios from a case-control design*

Description

A dataset containing the odds ratios (and standard errors) produced using a case-control design. The outcome is upper GI bleeding, the drug of interest (groundTruth = 1) is sertraline. Also included are 46 negative control drugs, for which we believe there to be no causal relation with upper GI bleeding.

Usage

```
data(caseControl)
```

Format

A data frame with 47 rows and 4 variables:

drugName Name of the drug

groundTruth Whether the drug is a positive (1) or negative (0) control

logRr The log of the incidence rate ratio

seLogRr The standard error of the log of the incidence rate ratio

Details

We used a database of medical records from general practices in the USA, the General Electric (GE) Centricity database, which contains data on 11.2 million subjects. We restricted on study period (start of 1990 through November 2003), age requirements (18 years or older), available time prior to event (180 days), number of controls per case (6), and risk definition window (30 days following the prescription). Controls were matched on age and sex. Cases of upper GI bleeding were identified on the basis of the occurrence of ICD-9 diagnosis codes in the problem list. These codes pertain to esophageal, gastric, duodenal, peptic, and gastrojejunal ulceration, perforation, and hemorrhage, as well as gastritis and non-specific gastrointestinal hemorrhage.

For more information on this set see Schuemie et al (2014).

References

Schuemie MJ, Ryan PB, Dumouchel W, Suchard MA, Madigan D. Interpreting observational studies: why empirical calibration is needed to correct p-values. Statistics in Medicine 33(2):209-18,2014

`cohortMethod` *Relative risks from a new-user cohort design*

Description

A dataset containing the relative risks (and standard errors) produced using a new-user cohort design. The outcome is acute liver injury, the drug of interest (groundTruth = 1) is Isoniazid Also included are 30 negative control drugs, for which we believe there to be no causal relation with acute liver injury.

Usage

```
data(cohortMethod)
```

Format

A data frame with 31 rows and 4 variables:

drugName Name of the drug

groundTruth Whether the drug is a positive (1) or negative (0) control

logRr The log of the incidence rate ratio

seLogRr The standard error of the log of the incidence rate ratio

Details

We used the Thomson MarketScan Medicare Supplemental Beneficiaries database, which contains data on 4.6 million subjects. We selected two groups (cohorts): (1) all subjects exposed to isoniazid and (2) all subjects having the ailment for which isoniazid is indicated, in this case tuberculosis, and having received at least one drug that is not known to cause acute liver injury. We removed all subjects who belonged to both groups and subjects for which less than 180 days of observation time was available prior to their first exposure to the drug in question. Acute liver injury was identified on the basis of the occurrence of ICD-9-based diagnosis codes from inpatient and outpatient medical claims and was defined broadly on the basis of codes associated with hepatic dysfunction, as have been used in prior observational database studies. The time at risk was defined as the length of exposure + 30 days, and we determined whether subjects experienced an acute liver injury during their time at risk. Using propensity score stratification, the cohorts were divided over 20 strata, and an odds ratio over all strata was computed using a Mantel-Haenszel test. The propensity score was estimated using Bayesian logistic regression using all available drug, condition, and procedure covariates occurring in the 180 days prior to first exposure, in addition to age, sex, calendar year of first exposure, Charlson index, number of drugs, number of visit days, and number of procedures.

For more information on this set see Schuemie et al (2014).

References

Schuemie MJ, Ryan PB, Dumouchel W, Suchard MA, Madigan D. Interpreting observational studies: why empirical calibration is needed to correct p-values. Statistics in Medicine 33(2):209-18,2014

fitNull *Fit the null distribution*

Description

fitNull fits the null distribution to a set of negative controls

Usage

```
fitNull(logRr, seLogRr)
```

Arguments

logRr	A numeric vector of effect estimates on the log scale
seLogRr	The standard error of the log of the effect estimates. Hint: often the standard error = (log(<lower bound 95 percent confidence interval>) - log(<effect estimate>))/qnorm(0.025)

Details

This function fits a Gaussian function to the negative control estimates as described in Schuemie et al (2014).

Value

An object of type null containing the mean and standard deviation (both on the log scale) of the null distribution.

References

Schuemie MJ, Ryan PB, Dumouchel W, Suchard MA, Madigan D. Interpreting observational studies: why empirical calibration is needed to correct p-values. Statistics in Medicine 33(2):209-18,2014

Examples

```
data(sccs)
negatives <- sccs[sccs$groundTruth == 0,]
null <- fitNull(negatives$logRr,negatives$seLogRr)
```

plotCalibration *Create a calibration plot*

Description

plotCalibration creates a plot showing the calibration of our calibration procedure

Usage

```
plotCalibration(logRr, seLogRr)
```

Arguments

logRr	A numeric vector of effect estimates on the log scale
seLogRr	The standard error of the log of the effect estimates. Hint: often the standard error = (log(<lower bound 95 percent confidence interval>) - log(<effect estimate>))/qnorm(0.025)

Details

Creates a calibration plot showing the number of effects with p < alpha for every level of alpha. The empirical calibration is performed using a leave-one-out design: The p-value of an effect is computed by fitting a null using all other negative controls. Ideally, the calibration line should approximate the diagonal. The plot shows both theoretical (traditional) and empirically calibrated p-values.

Value

A Ggplot object. Use the ggsave function to save to file.

Examples

```
data(sccs)
negatives <- sccs[sccs$groundTruth == 0,]
plotCalibration(negatives$logRr,negatives$seLogRr)
```

---

plotCalibrationEffect *Plot the effect of the calibration*

Description

plotCalibrationEffect creates a plot showing the effect of the calibration.

Usage

```
plotCalibrationEffect(logRrNegatives, seLogRrNegatives, logRrPositives,
  seLogRrPositives, null = NULL, xLabel = "Relative risk")
```

Arguments

logRrNegatives	A numeric vector of effect estimates of the negative controls on the log scale
seLogRrNegatives	The standard error of the log of the effect estimates of the negative controls
logRrPositives	A numeric vector of effect estimates of the positive controls on the log scale
seLogRrPositives	The standard error of the log of the effect estimates of the positive controls
null	An object representing the fitted null distribution as created by the fitNull function
xLabel	The label on the x-axis: the name of the effect estimate

Details

Creates a plot with the effect estimate on the x-axis and the standard error on the y-axis. Negative controls are shown as blue dots, positive controls as yellow diamonds. The area below the dashed line indicated estimates with p < 0.05. The orange area indicates estimates with calibrated p < 0.05.

Value

A Ggplot object. Use the ggsave function to save to file.

Examples

```
data(sccs)
negatives <- sccs[sccs$groundTruth == 0,]
positive <- sccs[sccs$groundTruth == 1,]
plotCalibrationEffect(negatives$logRr,negatives$seLogRr,positive$logRr,positive$seLogRr)
```

---

plotForest *Create a forest plot*

Description

plotForest creates a forest plot of effect size estimates.

Usage

```
plotForest(logRr, seLogRr, names, xLabel = "Relative risk")
```

Arguments

logRr	A numeric vector of effect estimates on the log scale
seLogRr	The standard error of the log of the effect estimates. Hint: often the standard error = (log(<lower bound 95 percent confidence interval>) - log(<effect estimate>))/qnorm(0.025)
names	A vector containing the names of the drugs or outcomes
xLabel	The label on the x-axis: the name of the effect estimate

Details

Creates a forest plot of effect size estimates (ratios). Estimates that are significantly different from 1 (alpha = 0.05) are marked in orange, others are marked in blue.

Value

A Ggplot object. Use the ggsave function to save to file.

Examples

```
data(sccs)
negatives <- sccs[sccs$groundTruth == 0,]
plotForest(negatives$logRr,negatives$seLogRr, negatives$drugName)
```

sccs *Incidence rate ratios from Self-Controlled Case Series*

Description

A dataset containing the incidence rate ratios (and standard errors) produced using a Self-Controlled Case Series (SCCS) design. The outcome is upper GI bleeding, the drug of interest (groundTruth = 1) is sertraline. Also included are 45 negative control drugs, for which we believe there to be no causal relation with upper GI bleeding.

Usage

```
data(sccs)
```

Format

A data frame with 46 rows and 4 variables:

drugName Name of the drug

groundTruth Whether the drug is a positive (1) or negative (0) control

logRr The log of the incidence rate ratio

seLogRr The standard error of the log of the incidence rate ratio

Details

We used a database of medical records from general practices in the USA, the General Electric (GE) Centricity database, which contains data on 11.2 million subjects. We restricted on study period (start of 1990 through November 2003), age requirements (18 years or older), available time prior to event (180 days), and risk definition window (30 days following the prescription). Time 30 days prior to the first prescription was removed to account for possible contra-indications. Cases of upper GI bleeding were identified on the basis of the occurrence of ICD-9 diagnosis codes in the problem list. These codes pertain to esophageal, gastric, duodenal, peptic, and gastrojejunal ulceration, perforation, and hemorrhage, as well as gastritis and non-specific gastrointestinal hemorrhage.

For more information on this set see Schuemie et al (2014).

References

Schuemie MJ, Ryan PB, Dumouchel W, Suchard MA, Madigan D. Interpreting observational studies: why empirical calibration is needed to correct p-values. Statistics in Medicine 33(2):209-18,2014

Index