Package ‘jfa’

April 25, 2021

Title Bayesian and Classical Audit Sampling

Version 0.5.3

Date 2021-04-25

Description Implements the audit sampling workflow as discussed in Derks et al. (2019) <doi:10.31234/osf.io/9f6ub>. The package makes it easy for an auditor to plan a statistical sample, select the sample from the population, and evaluate the sample using various methods according to the International Standards on Auditing. Furthermore, the package implements Bayesian equivalents of these methods.

BugReports https://github.com/koenderks/jfa/issues

URL https://github.com/koenderks/jfa, https://koenderks.github.io/jfa/

Suggests testthat, knitr, rmarkdown, kableExtra

Language en-US

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

VignetteBuilder knitr

R topics documented:

auditPrior 2
BuildIt 5
evaluation 6
planning 10
report 13
selection 14

Index **17**

`auditPrior` *Create a prior distribution for audit sampling*

Description

This function creates a prior distribution with audit information to be used in the `planning()` and `evaluation()` functions via their `prior` argument. The function returns an object of class `jfaPrior` which can be used with associated `print()` and `plot()` methods.

For more details on how to use this function see the package vignette: `vignette('jfa',package = 'jfa')`

Usage

```
auditPrior(confidence = 0.95, likelihood = 'binomial', method = 'none',
           expectedError = 0, N = NULL, materiality = NULL, ir = 1, cr = 1,
           pHmin = NULL, pHplus = NULL, factor = 1, sampleN = 0, sampleK = 0)
```

Arguments

`confidence`	a numeric value between 0 and 1 specifying the confidence level desired for the sample planning. Defaults to 0.95 for 95% confidence.
`likelihood`	a character specifying the likelihood assumed when updating the prior distribution. This can be either `binomial` for the binomial likelihood and beta prior distribution, `poisson` for the Poisson likelihood and gamma prior distribution, or `hypergeometric` for the hypergeometric likelihood and beta-binomial prior distribution. See the details section for more information about the available likelihoods.
`method`	a character specifying the method by which the prior distribution is constructed. Defaults to the `none` method, which incorporates no existing information. Other options are `median`, `hypotheses`, `arm`, `sample` or `factor`. See the details section for more information about these methods.
`expectedError`	a numeric value between 0 and 1 specifying the expected errors in the sample relative to the total sample size, or a value (>= 1) that represents the sum of expected errors in the sample. It is advised to set this value conservatively to minimize the probability of the observed errors exceeding the expected errors, which would imply that insufficient work has been done in the end.
`N`	an numeric value larger than 0 specifying the total population size. Only required when `likelihood = 'hypergeometric'`.
`materiality`	a numeric value between 0 and 1 specifying the performance materiality (i.e., maximum upper limit) of the audit as a fraction of the total population size. Can be `NULL` for some methods.
`ir`	if `method = 'arm'`, a numeric value between 0 and 1 specifying the inherent risk in the audit risk model. Defaults to 1 for 100% risk.
`cr`	if `method = 'arm'`, a numeric value between 0 and 1 specifying the internal control risk in the audit risk model. Defaults to 1 for 100% risk.
`pHmin`	if `method = 'hypotheses'`, a numeric value between 0 and 1 specifying the prior probability of the hypothesis $\theta <$ materiality.

`pHplus` if `method = 'hypotheses'`, a numeric value between 0 and 1 specifying the prior probability of the hypothesis $\theta >$ materiality.

`factor` if `method = 'factor'`, a numeric value between 0 and 1 specifying the weighting factor for the results of the earlier sample.

`sampleN` if `method = 'sample'` or `method = 'factor'`, an integer larger than, or equal to, 0 specifying the number of sampling units that were inspected in the earlier sample.

`sampleK` if `sample` or `factor`, a numeric value larger than, or equal to, 0 specifying the sum of errors in the previous sample.

Details

This section elaborates on the available likelihoods and corresponding prior distributions for the `likelihood` argument.

- `poisson`: The Poisson likelihood is often used as a likelihood for monetary unit sampling (MUS). The likelihood function is defined as:

  $$p(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

  The conjugate *gamma*(α, β) prior has probability density function:

  $$f(x; \alpha, \beta) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}$$

- `binomial`: The binomial likelihood is often used as a likelihood for attributes sampling *with* replacement. The likelihood function is defined as:

  $$p(x) = \binom{n}{k} p^k (1-p)^{n-k}$$

  The conjugate *beta*(α, β) prior has probability density function:

  $$f(x; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}$$

- `hypergeometric`: The hypergeometric likelihood is used as a likelihood for sampling *without* replacement. The likelihood function is defined as:

  $$p(x = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

  The conjugate *beta-binomial*(α, β) prior (Dyer and Pierce, 1993) has probability density function:

  $$f(k|n, \alpha, \beta) = \binom{n}{k} \frac{B(k+\alpha, n-k+\beta)}{B(\alpha, \beta)}$$

This section elaborates on the available methods for constructing a prior distribution.

- `none`: This method constructs a prior distribution that incorporates negligible information about the values of the misstatement.
- `median`: This method constructs a prior distribution so that the prior probability of tolerable misstatement (H-) is equal to the prior probability of intolerable misstatement (H+).

- `hypotheses`: This method constructs a prior distribution with custom prior probabilities for the hypotheses of tolerable misstatement (H-) and intolerable misstatement (H+). This method requires specification of the `pHmin` and `pHplus` arguments.
- `arm`: This method constructs a prior distribution by translating the risks of material misstatement (inherent risk and internal control risk) from the audit risk model to an implicit sample. The method requires specification of the `ir` (inherent risk) and `cr` (internal control risk) arguments.
- `sample`: This method constructs a prior distribution on the basis of an earlier sample. This method requires specification of the `sampleN` and `sampleK` arguments.
- `factor`: This method constructs a prior distribution on the basis of an earlier sample in combination with a weighting factor. This method requires specification of the `factor`, `sampleN` and `sampleK` arguments.

Value

An object of class `jfaPrior` containing:

`confidence`	a numeric value between 0 and 1 indicating the confidence level.
`likelihood`	a character indicating the specified likelihood.
`method`	a character indicating the method by which the prior distribution is constructed.
`expectedError`	a numeric value larger than 0 indicating the input for the number of expected errors.
`N`	if `N` is specified, an integer larger than 0 indicating the population size.
`materiality`	if `materiality` is specified, a numeric value between 0 and 1 indicating the materiality used to construct the prior distribution.
`description`	a list containing a description of the prior distribution, including the parameters of the prior distribution and the implicit sample on which the prior distribution is based.
`statistics`	a list containing statistics of the prior distribution, including the mean, mode, median, and upper bound of the prior distribution.
`specifics`	a list containing specifics of the prior distribution that vary depending on the `method`.
`hypotheses`	if `materiality` is specified, a list containing information about the hypotheses, including prior probabilities and odds for the hypothesis of tolerable misstatement (H-) and the hypothesis of intolerable misstatement (H+).

Author(s)

Koen Derks, <k.derks@nyenrode.nl>

References

Derks, K., de Swart, J., Wagenmakers, E.-J., Wille, J., & Wetzels, R. (2019). JASP for audit: Bayesian tools for the auditing practice.

Derks, K., de Swart, J., van Batenburg, P. Wagenmakers, E.-J., & Wetzels, R. (2020). Priors in a Bayesian audit: How integrating information into the prior distribution can improve audit transparency and efficiency.

See Also

planning selection evaluation report

Examples

```
# Specify inherent risk (ir) and control risk (cr)
ir <- 1     # 100%
cr <- 0.6   # 60%

# Create the prior distribution
auditPrior(confidence = 0.95, likelihood = 'binomial', method = 'arm',
           expectedError = 0.025, materiality = 0.05, ir = ir, cr = cr)
```

BuildIt *BuildIt Construction financial statements*

Description

Fictional data from a construction company in the United States, containing 3500 observations identification numbers, book values, and audit values. The audit values are added for illustrative purposes, as these would need to be assessed by the auditor in the execution stage of the audit.

Usage

```
data(BuildIt)
```

Format

A data frame with 3500 rows and 3 variables.

ID unique record identification number.

bookValue book value in US dollars ($14.47–$2,224.40).

auditValue true value in US dollars ($14.47–$2,224.40).

References

Derks, K., de Swart, J., Wagenmakers, E.-J., Wille, J., & Wetzels, R. (2019). JASP for audit: Bayesian tools for the auditing practice.

Examples

```
data(BuildIt)
```

`evaluation` *Evaluate a statistical audit sample*

Description

This function takes a data frame (using `sample`, `bookValues`, and `auditValues`) or summary statistics (using `nSumstats` and `kSumstats`) and performs inference on the misstatement in the sample. The function returns an object of class `jfaEvaluation` which can be used with associated `print()` and `plot()` methods.

For more details on how to use this function see the package vignette: `vignette('jfa',package = 'jfa')`

Usage

```
evaluation(confidence = 0.95, method = 'binomial', N = NULL,
           sample = NULL, bookValues = NULL, auditValues = NULL, counts = NULL,
            nSumstats = NULL, kSumstats = NULL,
            materiality = NULL, minPrecision = NULL,
            prior = FALSE, nPrior = 0, kPrior = 0,
            rohrbachDelta = 2.7, momentPoptype = 'accounts',
            populationBookValue = NULL,
            csA = 1, csB = 3, csMu = 0.5)
```

Arguments

`confidence`	a numeric value between 0 and 1 specifying the confidence level used in the evaluation. Defaults to 0.95 for 95% confidence.
`method`	a character specifying the method to be used in the evaluation. Possible options are `poisson`, `binomial` (default), `hypergeometric`, `mpu`, `stringer`, `stringer-meikle`, `stringer-lta`, `stringer-pvz`, `rohrbach`, `moment`, `direct`, `difference`, `quotient`, or `regression`. See the details section for more information.
`N`	an integer larger than 0 specifying the total number of items in the population.
`sample`	a data frame containing the sample to be evaluated. The sample must at least contain a column of book values and a column of audit (true) values.
`bookValues`	a character specifying the column name for the book values in the `sample`.
`auditValues`	a character specifying the column name for the audit values in the `sample`.
`counts`	a integer vector specifying the number of times each transaction in the sample should be counted in the evaluation (due to it being selected multiple times for the sample).
`nSumstats`	an integer larger than 0 specifying the number of items in the sample. If specified, overrides the `sample`, `bookValues` and `auditValues` arguments and assumes that the data come from summary statistics specified by both `nSumstats` and `kSumstats`.
`kSumstats`	a numeric value larger than 0 specifying the sum of errors found in the sample. If specified, overrides the `sample`, `bookValues` and `auditValues` arguments and assumes that the data come from summary statistics specified by both `kSumstats` and `nSumstats`.

`materiality`	a numeric value between 0 and 1 specifying the performance materiality (maximum tolerable error) as a fraction of the total size of the population. If specified, the function also returns the conclusion of the analysis with respect to the performance materiality. The value is discarded when `direct`, `difference`, `quotient`, or `regression` method is chosen.
`minPrecision`	a numeric value between 0 and 1 specifying the required minimum precision (upper bound minus most likely error) as a fraction of the total size of the population. If specified, the function also returns the conclusion of the analysis with respect to the required minimum precision.
`prior`	a logical specifying if a prior distribution must be used, or an object of class 'jfaPrior' containing the prior distribution. Defaults to `FALSE` for frequentist planning. If `TRUE`, a negligible prior distribution is chosen by default, but can be adjusted using the 'kPrior' and 'nPrior' arguments. Chooses a conjugate gamma distribution for the Poisson likelihood, a conjugate beta distribution for the binomial likelihood, and a conjugate beta-binomial distribution for the hypergeometric likelihood.
`nPrior`	if `prior = TRUE`, a numeric value specifying the number of sampling units in the implicit sample on which the prior distribution is based.
`kPrior`	if `prior = TRUE`, a numeric value specifying the assumed sum of errors in the implicit sample on which the prior distribution is based.
`rohrbachDelta`	if `method = 'rohrbach'`, a numeric value specifying Δ in Rohrbach's augmented variance bound (Rohrbach, 1993).
`momentPoptype`	if `method = 'moment'`, a character specifying the type of population (Dworin and Grimlund, 1984). Possible options are `accounts` and `inventory`. This argument affects the calculation of the central moments in the bound.
`populationBookValue`	if `method` is one of `direct`, `difference`, `quotient`, or `regression`, a numeric value specifying the total value of the transactions in the population. This argument is optional otherwise.
`csA`	if `method = "coxsnell"`, a numeric value specifying the α parameter of the prior distribution on the mean taint. Defaults to 1 as recommended by Cox and Snell (1979).
`csB`	if `method = "coxsnell"`, a numeric value specifying the β parameter of the prior distribution on the mean taint. Defaults to 3 as recommended by Cox and Snell (1979).
`csMu`	if `method = "coxsnell"`, a numeric value between 0 and 1 specifying the mean of the prior distribution on the mean taint. Defaults to 0.5 as recommended by Cox and Snell (1979).

Details

This section lists the available options for the `methods` argument.

- `poisson`: Evaluates the sample with the Poisson distribution. If combined with `prior = TRUE`, performs Bayesian evaluation using a *gamma* prior and posterior.
- `binomial`: Evaluates the sample with the binomial distribution. If combined with `prior = TRUE`, performs Bayesian evaluation using a *beta* prior and posterior.
- `hypergeometric`: Evaluates the sample with the hypergeometric distribution. If combined with `prior = TRUE`, performs Bayesian evaluation using a *beta-binomial* prior and posterior.

- `mpu`: Evaluates the sample with the mean-per-unit estimator.
- `stringer`: Evaluates the sample with the Stringer bound (Stringer, 1963).
- `stringer-meikle`: Evaluates the sample with the Stringer bound with Meikle's correction for understatements (Meikle, 1972).
- `stringer-lta`: Evaluates the sample with the Stringer bound with LTA correction for understatements (Leslie, Teitlebaum, and Anderson, 1979).
- `stringer-pvz`: Evaluates the sample with the Stringer bound with Pap and van Zuijlen's correction for understatements (Pap and van Zuijlen, 1996).
- `rohrbach`: Evaluates the sample with Rohrbach's augmented variance bound (Rohrbach, 1993).
- `moment`: Evaluates the sample with the modified moment bound (Dworin and Grimlund, 1984).
- `coxsnell`: Evaluates the sample with the Cox and Snell bound (Cox and Snell, 1979).
- `direct`: Evaluates the sample with the direct estimator (Touw and Hoogduin, 2011).
- `difference`: Evaluates the sample with the difference estimator (Touw and Hoogduin, 2011).
- `quotient`: Evaluates the sample with the quotient estimator (Touw and Hoogduin, 2011).
- `regression`: Evaluates the sample with the regression estimator (Touw and Hoogduin, 2011).

Value

An object of class `jfaEvaluation` containing:

`confidence`	a numeric value between 0 and 1 indicating the confidence level.
`method`	a character indicating the evaluation method.
`N`	if `N` is specified, in integer larger than 0 indicating the population size.
`n`	an integer larger than 0 indicating the sample size.
`k`	an integer larger than, or equal to, 0 indicating the number of items in the sample that contained an error.
`t`	a value larger than, or equal to, 0, indicating the sum of observed taints.
`materiality`	if `materiality` is specified, a numeric value between 0 and 1 indicating the performance materiality as a fraction of the total population size.
`minPrecision`	if `minPrecision` is specified, a numeric value between 0 and 1 indicating the minimum required precision as a fraction of the total population size.
`mle`	a numeric value between 0 and 1 indicating the most likely error in the population as a fraction of its total size.
`precision`	a numeric value between 0 and 1 indicating the difference between the most likely error and the upper bound in the population as a fraction of the total population size.
`popBookvalue`	if `populationBookValue` is specified, a numeric value larger than 0 indicating the total value of the population.
`pointEstimate`	if `method` is one of `direct`, `difference`, `quotient`, or `regression`, a numeric value indicating the point estimate of the population misstatement as a fraction the total population size.
`lowerBound`	if method is one of `direct`, `difference`, `quotient`, or `regression`, a numeric value indicating the lower bound of the interval around the population misstatement as a fraction the total population size.

`upperBound`	if method is one of `direct`, `difference`, `quotient`, or `regression`, a numeric value indicating the upper bound of the interval around the population misstatement as a fraction the total population size.
`confBound`	a numeric value indicating the upper bound on the population misstatement as a fraction the total population size.
`conclusion`	if `materiality` is specified, a character indicating the conclusion about whether to approve or not approve the population with respect to the performance materiality.
`populationK`	if `method = 'hypergeometric'`, an integer indicating the assumed total errors in the population.
`prior`	an object of class 'jfaPrior' that contains the prior distribution.
`posterior`	an object of class 'jfaPosterior' that contains the posterior distribution.
`data`	a data frame containing the relevant columns from the `sample`.

Author(s)

Koen Derks, <k.derks@nyenrode.nl>

References

Cox, D. and Snell, E. (1979). On sampling and the estimation of rare errors. *Biometrika*, 66(1), 125-132.

Dworin, L. D. and Grimlund, R. A. (1984). Dollar-unit sampling for accounts receivable and inventory. *The Accounting Review*, 59(2), 218–241

Leslie, D. A., Teitlebaum, A. D., & Anderson, R. J. (1979). *Dollar-unit Sampling: A Practical Guide for Auditors*. Copp Clark Pitman; Belmont, Calif.: distributed by Fearon-Pitman.

Meikle, G. R. (1972). *Statistical Sampling in an Audit Context: An Audit Technique*. Canadian Institute of Chartered Accountants.

Pap, G., and van Zuijlen, M. C. (1996). On the asymptotic behavior of the Stringer bound. *Statistica Neerlandica*, 50(3), 367-389.

Rohrbach, K. J. (1993). Variance augmentation to achieve nominal coverage probability in sampling from audit populations. *Auditing*, 12(2), 79.

Stringer, K. W. (1963). Practical aspects of statistical sampling in auditing. *In Proceedings of the Business and Economic Statistics Section* (pp. 405-411). American Statistical Association.

Touw, P., and Hoogduin, L. (2011). *Statistiek voor Audit en Controlling*. Boom uitgevers Amsterdam.

See Also

`auditPrior` `planning` `selection` `report`

Examples

```
data('BuildIt')

# Draw a sample of 100 monetary units from the population using
# fixed interval monetary unit sampling
sample <- selection(population = BuildIt, sampleSize = 100,
          algorithm = 'interval', units = 'mus', bookValues = 'bookValue')$sample
```

```
# Evaluate using the Stringer bound
evaluation(confidence = 0.95, materiality = 0.05,
           method = 'stringer', sample = sample,
           bookValues = 'bookValue', auditValues = 'auditValue')
```

`planning` *Plan a statistical audit sample*

Description

This function calculates the required sample size for a statistical audit sample based on the Poisson, binomial or hypergeometric likelihood. The function returns an object of class `jfaPlanning` which can be used with associated `print()` and `plot()` methods.

For more details on how to use this function see the package vignette: `vignette('jfa',package = 'jfa')`

Usage

```
planning(confidence = 0.95, expectedError = 0, likelihood = 'poisson', N = NULL,
          materiality = NULL, minPrecision = NULL,
          prior = FALSE, nPrior = 0, kPrior = 0,
          increase = 1, maxSize = 5000)
```

Arguments

`confidence`	a numeric value between 0 and 1 specifying the confidence level used in the planning. Defaults to 0.95 for 95% confidence.
`expectedError`	a numeric value between 0 and 1 specifying the expected errors in the sample relative to the total sample size, or a number (>= 1) that represents the number of expected errors in the sample. It is advised to set this value conservatively to minimize the probability of the observed errors exceeding the expected errors, which would imply that insufficient work has been done in the end.
`likelihood`	a character specifying the likelihood assumed in the calculation. This can be either `binomial` for the binomial likelihood, `poisson` for the Poisson likelihood, or `hypergeometric` for the hypergeometric likelihood. See the details section for more information about the available likelihoods.
`N`	an integer larger than 0 specifying the total population size. Only required when `likelihood = 'hypergeometric'`.
`materiality`	a numeric value between 0 and 1 specifying the performance materiality (i.e., maximum upper limit) as a fraction of the total population size. Can be `NULL`, but `minPrecision` should be specified in that case.
`minPrecision`	a numeric value between 0 and 1 specifying the minimum precision (i.e., upper bound minus most likely error) as a fraction of the total population size. Can be `NULL`, but `materiality` should be specified in that case.
`prior`	a logical specifying whether to use a prior distribution when planning, or an object of class `jfaPrior` containing the prior distribution. Defaults to `FALSE` for frequentist planning. If `TRUE`, a negligible prior distribution is chosen by default, but can be adjusted using the `kPrior` and `nPrior` arguments. Chooses a conjugate gamma distribution for the Poisson likelihood, a conjugate beta distribution

for the binomial likelihood, and a conjugate beta-binomial distribution for the hypergeometric likelihood.

`nPrior`	if `prior = TRUE`, a numeric value larger than, or equal to, 0 specifying the number of sampling units in the implicit sample on which the prior distribution is based.
`kPrior`	if `prior = TRUE`, a numeric value larger than, or equal to, 0 specifying the assumed sum of errors in the implicit sample on which the prior distribution is based.
`increase`	an integer larger than 0 specifying the desired increase step for the sample size calculation.
`maxSize`	an integer larger than 0 specifying the maximum sample size that is considered in the calculation. Defaults to 5000 for efficiency. Increase this value if the sample size cannot be found due to it being too large (e.g., for a low materiality).

Details

This section elaborates on the available likelihoods and corresponding prior distributions for the `likelihood` argument.

- `poisson`: The Poisson likelihood is often used as a likelihood for monetary unit sampling (MUS). The likelihood function is defined as:

  $$p(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

  The conjugate *gamma(α, β)* prior has probability density function:

  $$f(x; \alpha, \beta) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}$$

- `binomial`: The binomial likelihood is often used as a likelihood for attributes sampling *with* replacement. The likelihood function is defined as:

  $$p(x) = \binom{n}{k} p^k (1-p)^{n-k}$$

  The conjugate *beta(α, β)* prior has probability density function:

  $$f(x; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}$$

- `hypergeometric`: The hypergeometric likelihood is used as a likelihood for sampling *without* replacement. The likelihood function is defined as:

  $$p(x = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

  The conjugate *beta-binomial(α, β)* prior (Dyer and Pierce, 1993) has probability density function:

  $$f(k|n, \alpha, \beta) = \binom{n}{k} \frac{B(k+\alpha, n-k+\beta)}{B(\alpha, \beta)}$$

Value

An object of class `jfaPlanning` containing:

`confidence`	a numeric value between 0 and 1 indicating the confidence level.
`expectedError`	a numeric value larger than, or equal to, 0 indicating the expected errors input.
`likelihood`	a character indicating the specified likelihood.
`N`	an integer larger than 0 indicating the population size (only returned if `N` is specified).
`materiality`	a numeric value between 0 and 1 indicating the specified materiality. Can be `NULL`.
`minPrecision`	a numeric value between 0 and 1 indicating the minimum precision to be obtained. Can be `NULL`.
`sampleSize`	an integer larger than 0 indicating the required sample size.
`errorType`	a character indicating whether the expected errors where specified as a percentage or as an integer.
`expectedSampleError`	a numeric value larger than, or equal to, 0 indicating the number of errors that are allowed in the sample.
`expectedBound`	a numeric value between 0 and 1 indicating the expected upper bound if the sample goes according to plan.
`expectedPrecision`	a numeric value between 0 and 1 indicating the expected precision if the sample goes according to plan.
`populationK`	if `likelihood = 'hypergeometric'`, an integer larger than 0 indicating the assumed population errors.
`prior`	if a prior distribution is specified, an object of class `jfaPrior` that contains information about the prior distribution.
`expectedPosterior`	if a prior distribution is specified, an object of class `jfaPosterior` that contains information about the expected posterior distribution.

Author(s)

Koen Derks, <k.derks@nyenrode.nl>

References

Dyer, D. and Pierce, R.L. (1993). On the choice of the prior distribution in hypergeometric sampling. *Communications in Statistics - Theory and Methods*, 22(8), 2125 - 2146.

See Also

`auditPrior` `selection` `evaluation` `report`

Examples

```
# Frequentist planning using a binomial likelihood
planning(confidence = 0.95, expectedError = 0.025, likelihood = 'binomial',
         materiality = 0.05)

# Bayesian planning using a negligible beta prior
planning(confidence = 0.95, expectedError = 0.025, likelihood = 'binomial',
         materiality = 0.05, prior = TRUE)
```

---

report — *Create a statistical audit sampling report*

Description

This function takes an object of class `jfaEvaluation` as returned by the `evaluation()` function automatically generates a `html` or `pdf` report containing the analysis results and their interpretation.

For more details on how to use this function see the package vignette: `vignette('jfa',package = 'jfa')`

Usage

```
report(object, file = 'report.html', format = 'html_document')
```

Arguments

`object`	an object of class `jfaEvaluation` as returned by the `evaluation()` function.
`file`	a character specifying the name of the report (e.g. `report.html`). By default, the report is created in your current working directory.
`format`	a character specifying the output format of the report. Possible options are `html_document` (default) and `pdf_document`, but compiling to `pdf` format requires a local version of MikTex.

Value

A `html` or `pdf` file containing a report of the evaluation.

Author(s)

Koen Derks, <k.derks@nyenrode.nl>

See Also

`auditPrior` `planning` `selection` `evaluation`

Examples

```
data('BuildIt')

# Draw a sample of 100 monetary units from the population using
# fixed interval monetary unit sampling
sample <- selection(population = BuildIt, sampleSize = 100,
          algorithm = 'interval', units = 'mus', bookValues = 'bookValue')$sample

# Evaluate using the Stringer bound
result <- evaluation(confidence = 0.95, materiality = 0.05,
                     method = 'stringer', sample = sample,
                     bookValues = 'bookValue', auditValues = 'auditValue')

## Not run:
 report(result)

## End(Not run)
```

`selection` *Select a statistical audit sample*

Description

This function takes a data frame and performs statistical sampling according to one of three algorithms: random sampling, cell sampling, and fixed interval sampling. Sampling is done on the level of two possible sampling units: records or monetary units. The function returns an object of class `jfaSelection` which can be used with associated `print()` and a `plot()` methods.

For more details on how to use this function see the package vignette: `vignette('jfa',package = 'jfa')`

Usage

```
selection(population, sampleSize, units = 'records', algorithm = 'random',
          bookValues = NULL, intervalStartingPoint = 1, ordered = TRUE,
          ascending = TRUE, withReplacement = FALSE, seed = 1)
```

Arguments

`population`	a data frame containing the population of items the auditor wishes to sample from.
`sampleSize`	an integer larger than 0 specifying the number of sampling units that need to be selected from the population. Can also be an object of class `jfaPlanning`.
`units`	a character specifying the sampling units used. Possible options are `records` (default) for selection on the level of items or `mus` for selection on the level of monetary units.
`algorithm`	a character specifying the sampling algorithm used. Possible options are `random` (default) for random sampling, `cell` for cell sampling, or `interval` for fixed interval sampling.
`bookValues`	a character specifying the name of the column in the `population` that contains the book values of the items.

intervalStartingPoint
: if algorithm = 'interval', an integer larger than 0 specifying the starting point of the algorithm.

ordered
: a logical specifying whether to first order the items in the population according to the value of their bookValues. Defaults to TRUE.

ascending
: if ordered = TRUE, a logical specifying whether to order the population bookValues from smallest to largest. Defaults to TRUE.

withReplacement
: if algorithm = 'random', a logical specifying whether sampling should be performed with replacement. Defaults to FALSE.

seed
: if algorithm = 'random' or algorithm = 'cell', an integer specifying a seed to reproduce results. Defaults to 1.

Details

The first part of this section elaborates on the two possible options for the units argument:

- records: In record sampling each item in the population is seen as a sampling unit. An item of $5000 is therefore equally likely to be selected as an item of $500.
- mus: In monetary unit sampling each monetary unit in the population is seen as a sampling unit. An item of $5000 is therefore ten times more likely to be selected as an item of $500.

The second part of this section elaborates on the three possible options for the algorithm argument:

- random: In random sampling each sampling unit in the population is drawn with equal probability.
- cell: In cell sampling the sampling units in the population are divided into a number (equal to the sample size) of intervals. From each interval one sampling unit is selected with equal probability.
- interval: In fixed interval sampling the sampling units in the population are divided into a number (equal to the sample size) of intervals. From each interval one sampling unit is selected according to a fixed starting point (specified by intervalStartingPoint).

Value

An object of class jfaSelection containing:

population
: a data frame containing the input population.

sample
: a data frame containing the selected sample of items.

units
: a character indicating the sampling units that were used to create the selection.

algorithm
: a character indicating the the algorithm that was used to create the selection.

bookValues
: if bookValues is specified, a character indicating the name of the book value column.

Author(s)

Koen Derks, <k.derks@nyenrode.nl>

References

Leslie, D. A., Teitlebaum, A. D., & Anderson, R. J. (1979). *Dollar-unit Sampling: A Practical Guide for Auditors*. Copp Clark Pitman; Belmont, Calif.: distributed by Fearon-Pitman.

Wampler, B., & McEacharn, M. (2005). Monetary-unit sampling using Microsoft Excel. *The CPA journal*, 75(5), 36.

See Also

auditPrior planning evaluation report

Examples

```
data('BuildIt')

# Draw a sample of 100 monetary units from the population using
# fixed interval monetary unit sampling
selection(population = BuildIt, sampleSize = 100,
          algorithm = 'interval', units = 'mus', bookValues = 'bookValue')
```

Index

∗ **audit**
auditPrior, 2
evaluation, 6
planning, 10
report, 13
selection, 14
∗ **bound**
evaluation, 6
∗ **confidence**
evaluation, 6
∗ **datasets**
BuildIt, 5
∗ **distribution**
auditPrior, 2
∗ **evaluation**
evaluation, 6
report, 13
∗ **planning**
planning, 10
∗ **prior**
auditPrior, 2
∗ **report**
report, 13
∗ **sample**
planning, 10
selection, 14
∗ **selection**
selection, 14
∗ **size**
planning, 10

auditPrior, 2, *9*, *12*, *13*, *16*

BuildIt, 5

evaluation, *5*, 6, *12*, *13*, *16*

planning, *5*, *9*, 10, *13*, *16*

report, *5*, *9*, *12*, 13, *16*

selection, *5*, *9*, *12*, *13*, 14