Package 'jfa'

December 10, 2019

Title A multi-functional R package for auditing

Version 1.0.0

Description The jfa package provides functions for daily use in auditing, such as calculating sample sizes, sampling, and calculating confidence bounds.

License GPL-3 + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.0.1

VignetteBuilder knitr

Suggests testthat, knitr, rmarkdown

R topics documented:

Index																						14
	evaluation planning . sampling .												•									5
	auditPrior BuildIt																					

Create a Prior Distribution

Description

auditPrior

This function creates a prior distribution according to the audit risk model. The returned object is of class jfaPrior and can be used with associated print() and plot() methods. jfaPrior objects can be used as input argument for the prior argument in other functions.

Usage

2 auditPrior

Arguments

materiality a value between 0 and 1 representing the materiality of the audit as a fraction of

the total size or value.

confidence the confidence level desired from the confidence bound (on a scale from 0 to 1).

Defaults to 0.95, or 95% confidence.

method the method by which the prior distribution is constructed. Currently only sup-

ports the arm method, which uses the audit risk model (Derks et al., 2019).

ir the inherent risk probability from the audit risk model. Defaults to 1 for 100%

risk.

cr the inherent risk probability from the audit risk model. Defaults to 1 for 100%

risk.

expectedError a fraction representing the percentage of expected mistakes in the sample relative

to the total size, or a number (>= 1) that represents the number of expected

mistakes.

likelihood can be one of binomial, poisson, or hypergeometric.

N the population size (required for hypergeometric calculations).

Value

An object of class jfaPrior containing:

method the method by which the prior distribution is constructed.

likelihood the likelihood by which the prior distribution is updated.

priorD the name of the probability density function of the prior distribution.

nPrior the prior assumed sample size.
kPrior the prior assumed sample errors

aPrior the prior parameter alpha.
bPrior the prior parameter beta.

materiality that was used to construct the prior distribution.

N if specified as input, the population size.

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.

See Also

planning sampling evaluation

auditPrior 3

```
library(jfa)
set.seed(1)
# Generate some audit data (N = 1000).
population <- data.frame(ID = sample(1000:100000, size = 1000, replace = FALSE),</pre>
                         bookValue = runif(n = 1000, min = 700, max = 1000))
\ensuremath{\mathtt{\#}} Specify the materiality, confidence, and expected errors.
materiality <- 0.05 # 5%
confidence <- 0.95 # 95%
expectedError <- 0.025 # 2.5%
# Specify the inherent risk (ir) and control risk (cr).
ir <- 1
           # 100%
cr <- 0.6 # 60%
# Create a beta prior distribution according to the Audit Risk Model (arm)
# and a binomial likelihood.
prior <- auditPrior(materiality = materiality, confidence = confidence,</pre>
                    method = "arm", ir = ir, cr = cr,
                    expectedError = expectedError, likelihood = "binomial")
print(prior)
# jfa prior distribution for arm method:
# Prior sample size:
# Prior errors:
                         1.27
# Prior:
                         beta(2.275, 50.725)
# Calculate the sample size according to the binomial distribution with the specified prior.
sampleSize <- planning(materiality = materiality, confidence = confidence,</pre>
                   expectedError = expectedError, prior = prior, likelihood = "binomial")
print(sampleSize)
# jfa planning results for beta prior with binomial likelihood:
# Materiality:
                           5%
# Confidence:
                           95%
# Sample size:
                           169
# Allowed sample errors:
                           4.23
# Prior parameter alpha:
                           2.275
# Prior parameter beta:
                           50.725
# Draw sample using random monetary unit sampling.
sampleResult <- sampling(population = population, sampleSize = sampleSize,</pre>
                   algorithm = "random", units = "mus", seed = 1, bookValues = "bookValue")
print(sampleResult)
# jfa sampling results for random monetary unit sampling:
# Population size:
                           1000
# Sample size:
                           169
# Proportion n/N:
                           0.169
# Precentage of value:
                           16.84%
```

BuildIt

```
# Isolate the sample.
sample <- sampleResult$sample</pre>
# For this example, we use the book values as audit values. Implies zero errors at this point.
sample$trueValue <- sample$bookValue</pre>
# One overstatement is found.
sample$trueValue[2] <- sample$trueValue[2] - 500</pre>
# Evaluate the sample with one partial error using the posterior distribution.
conclusion <- evaluation(sample = sample, bookValues = "bookValue", auditValues = "trueValue",</pre>
                          prior = prior, materiality = 0.05)
print(conclusion)
# jfa evaluation results for binomial likelihood with prior:
# Materiality:
                          5%
                          95%
# Confidence:
# Upper bound:
                          2.729%
# Sample size:
                          169
# Sample errors:
# Conclusion:
                         Approve population
```

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.

```
data(BuildIt)
```

evaluation	Evaluation of Audit Samples using Confidence / Credible Bounds

Description

This function takes a sample data frame or summary statistics about an evaluated audit sample and calculates a confidence bound accordint to a specified method. The returned object is of class jfaEvaluation and can be used with associated print() and plot() methods.

Usage

Arguments

sample	a data frame containing at least a column of book values and a column of audit
--------	--

(true) values.

bookValues the column name for the book values in the sample.

auditValues the column name for the audit (true) values in the sample.

confidence the required confidence level for the bound.

nSumstats the number of observations in the sample. If specified, overrides the sample,

bookValues and auditValues arguments and assumes that the data comes from

summary statistics specified by nSumstats and kSumstats.

kSumstats the sum of the errors found in the sample. If specified, overrides the sample,

bookValues and auditValues arguments and assumes that the data comes from

summary statistics specified by kSumstats and nSumstats.

method can be either one of poisson, binomial, hypergeometric, stringer, stringer-meikle,

stringer-lta, stringer-pvz, rohrbach, moment, direct, difference, quotient,

or regression.

materiality if specified, the function also returns the conclusion of the analysis with respect

to the materiality. This value must be specified as a fraction of the total value of the population (a value between 0 and 1). The value is discarded when direct,

difference, quotient, or regression method is chosen.

N the total population size.

prior whether to use a prior distribution when evaluating. Defaults to FALSE for fre-

quentist evaluation. If TRUE, the prior distribution is updated by the specified likelihood. 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 the prior parameter α (number of errors in the assumed prior sample).

kPrior the prior parameter β (total number of observations in the assumed prior sam-

ple).

rohrbachDelta the value of Δ in Rohrbach's augmented variance bound.

momentPoptype can be either one of accounts or inventory. Options result in different meth-

ods for calculating the central moments, for more information see Dworin and

Grimlund (1986).

populationBookValue

the total value of the audit population. Required when method is one of direct,

difference, quotient, or regression.

csA if method = "coxsnell", the α parameter of the prior distribution on the mean

taint. Default is set to 1, as recommended by Cox and Snell (1979).

if method = "coxsnell", the β parameter of the prior distribution on the mean

taint. Default is set to 3, as recommended by Cox and Snell (1979).

csMu if method = "coxsnell", the mean of the prior distribution on the mean taint.

Default is set to 0.5, as recommended by Cox and Snell (1979).

Details

This section lists the available options for the methods argument.

- poisson: The confidence bound taken from the Poisson distribution. If combined with prior = TRUE, performs Bayesian evaluation using a *gamma* prior and posterior.
- binomial: The confidence bound taken from the binomial distribution. If combined with prior = TRUE, performs Bayesian evaluation using a *beta* prior and posterior.
- hypergeometric: The confidence bound taken from the hypergeometric distribution. If combined with prior = TRUE, performs Bayesian evaluation using a *beta-binomial* prior and posterior.
- stringer: The Stringer bound (Stringer, 1963).
- stringer-meikle: Stringer bound with Meikle's correction for understatements (Meikle, 1972).
- stringer-1ta: Stringer bound with LTA correction for understatements (Leslie, Teitlebaum, and Anderson, 1979).
- stringer-pvz: Stringer bound with Pap and van Zuijlen's correction for understatements (Pap and van Zuijlen, 1996).
- rohrbach: Rohrbach's augmented variance bound (Rohrbach, 1993).
- moment: Modified moment bound (Dworin and Grimlund, 1986).
- coxsnell: Cox and Snell bound (Cox and Snell, 1979).
- direct: Confidence interval using the direct method (Touw and Hoogduin, 2011).
- difference: Confidence interval using the difference method (Touw and Hoogduin, 2011).
- quotient: Confidence interval using the quotient method (Touw and Hoogduin, 2011).
- regression: Confidence interval using the regression method (Touw and Hoogduin, 2011).

Value

An object of class jfaEvaluation containing:

n the sample size.

k an integer specifyng the number of observed errors.

t a number specifying the sum of observed taints.

confidence the confidence level of the result.

popBookvalue if specified as input, the total book value of the population.

pointEstimate if method is one of direct, difference, quotient, or regression, the value

of the point estimate.

lowerBound if method is one of direct, difference, quotient, or regression, the value

of the lower bound of the interval.

upperBound if method is one of direct, difference, quotient, or regression, the value

of the upper bound of the interval.

confBound the upper confidence bound on the error percentage.

method the evaluation method that was used.

materiality the materiality.

conclusion if materiality is specified, the conclusion about whether to approve or not

approve the population.

N if specified as input, the population size.

populationK the assumed total errors in the population. Used for inferences with hypergeometric

method.

prior a logical, indicating whether a prior was used in the analysis.

nPrior if a prior is specified, the prior assumed sample size. kPrior if a prior is specified, the prior assumed sample errors.

multiplicationFactor

if method = "coxsnell", the multiplication factor for the F-distribution.

df1 if method = "coxsnell", the df1 for the *F*-distribution. df2 if method = "coxsnell", the df2 for the *F*-distribution.

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., and Grimlund, R. A. (1986). Dollar-unit sampling: A comparison of the quasi-Bayesian and moment bounds. *Accounting Review*, 36-57.

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 behaviour of the Stringer bound 1. *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 sampling
```

```
library(jfa)
# Generate some audit data (N = 1000)
set.seed(1)
data <- data.frame(ID = sample(1000:100000, size = 1000, replace = FALSE),</pre>
                  bookValue = runif(n = 1000, min = 700, max = 1000))
# Using the binomial likelihood, calculates the upper 95% confidence bound for a
# materiality of 5\% when 1\% full errors are found in a sample (n = 93).
jfaRes <- planning(materiality = 0.05, confidence = 0.95, expectedError = 0.01,
                   likelihood = "binomial")
# Using monetary unit sampling, draw a random sample from the population.
samp <- sampling(population = data, sampleSize = jfaRes, units = "mus",</pre>
                 bookValues = "bookValue", algorithm = "random")
samp$sample$trueValue <- samp$sample$bookValue</pre>
samp$sample$trueValue[2] <- samp$sample$trueValue[2] - 500 # One overstatement is found
# Evaluate the sample using the stringer bound.
conclusion <- evaluation(sample = samp$sample, bookValues = "bookValue",</pre>
                      auditValues = "trueValue", method = "stringer", materiality = 0.05)
print(conclusion)
# jfa evaluation results for stringer method:
                         5%
# Materiality:
                         95%
# Confidence:
# Upper bound:
                         4.244%
# Sample size:
                        93
# Sample errors:
# Conclusion:
                         Approve population
# Evaluate the sample using summary statistics (n = 93, k = 1).
conclusion <- evaluation(nSumstats = 93, kSumstats = 1, method = "binomial",</pre>
                         materiality = 0.05)
print(conclusion)
# jfa evaluation results for binomial method:
# Materiality:
                         5%
                         95%
# Confidence:
# Upper bound:
                         4.999%
# Sample size:
                         93
# Sample errors:
# Conclusion:
                         Approve population
```

planning 9

planning	Frequentist and Bayesian Planning for Audit Samples	

Description

This function calculates the required sample size for an audit, based on the poisson, binomial, or hypergeometric likelihood. A prior can be specified to perform Bayesian planning. The returned object is of class jfaPlanning and can be used with associated print() and plot() methods.

Usage

Arguments

materiality	a value between 0 and 1 representing the materiality of the audit as a fraction of the total size or value.
confidence	the confidence level desired from the confidence bound (on a scale from 0 to 1). Defaults to 0.95 , or 95% confidence.
expectedError	a fraction representing the percentage of expected mistakes in the sample relative to the total size, or a number (>= 1) that represents the number of expected mistakes.
likelihood	can be one of binomial, poisson, or hypergeometric.
N	the population size (required for hypergeometric calculations).
maxSize	the maximum sample size that is considered for calculations. Defaults to 5000 for efficiency. Increase this value if the samle size cannot be found due to it being too large (e.g., for low materialities).
prior	whether to use a prior distribution when planning. Defaults to FALSE for frequentist planning. If TRUE, the prior distribution is updated by the specified likelihood. 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.
kPrior	the prior parameter α (number of errors in the assumed prior sample).
nPrior	the prior parameter β (total number of observations in the assumed prior sample).

Details

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

• poisson: The Poisson likelihood is used as a likelihood for monetary unit sampling (MUS). Its likelihood function is defined as:

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

The conjugate $gamma(\alpha, \beta)$ prior has probability density function:

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

10 planning

binomial: The binomial likelihood is used as a likelihood for record sampling with replacement. Its likelihood function is defined as:

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

The conjugate $beta(\alpha, \beta)$ 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 record sampling *without* replacement. Its 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:

materiality the value of the specified materiality.

confidence the confidence level for the desired population statement.

sampleSize the resulting sample size.

expectedSampleError

the number of full errors that are allowed to occur in the sample.

expectedError the specified number of errors as a fraction or as a number.

likelihood the specified likelihood.

errorType whether the expected errors where specified as a percentage or as an integer.

N the population size (only returned in case of a hypergeometric likelihood).

populationK the assumed population errors (only returned in case of a hypergeometric likeli-

hood).

prior a list containing information on the prior parameters.

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 sampling evaluation

sampling 11

Examples

sampling

Sampling from Audit Populations

Description

This function takes a data frame and performs sampling according to one of three popular algorithms: random sampling, cell sampling, or fixed interval sampling. Sampling is done in combination with one of two sampling units: records or monetary units The returned object is of class jfaSampling and can be used with associated print() and plot() methods.

Usage

Arguments

population	a data frame containing the population the auditor wishes to sample from.					
sampleSize	the number of observations that need to be selected from the population. Can also be an object of class jfaPlanning.					
bookValues	a character specifying the name of the column containing the book values (as in the population data).					
units	can be either records (default) for record sampling, or mus for monetary unit sampling. $ \\$					
algorithm	can be either one of random (default) for random sampling, cell for cell sampling, or interval for fixed interval sampling.					
intervalStartingPoint						
	the starting point in the interval (used only in fixed interval sampling)					
ordered	if TRUE (default), the population is first ordered according to the value of their book values.					

12 sampling

ascending if TRUE (default), order the population in ascending order. withReplacement

whether sampling should be performed with replacement. Defaults to FALSE.

seed seed to reproduce results. Default is 1.

Details

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

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

This second part of this section elaborates on the 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 (intervalStartingPoint).

Value

An object of class jfaSampling containing:

population a data frame containing the input population.

sample a data frame containing the selected observations.

bookValues if specified, the name of the specified book value column.

algorithm that was used for sampling.

units the sampling units that were used for sampling.

Author(s)

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

References

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

See Also

auditPrior planning evaluation

sampling 13

Index

```
*Topic audit
    auditPrior, 1
    evaluation, 5
    planning, 9
    sampling, 11
*Topic bound
    evaluation, 5
*Topic confidence
    evaluation, 5
*Topic datasets
    BuildIt, 4
*Topic distribution
    auditPrior, 1
*Topic evaluation
    evaluation, 5
*Topic planning
    planning, 9
*Topic prior
    auditPrior, 1
*Topic sample
    planning, 9
    {\tt sampling}, {\tt 11}
*Topic sampling
    sampling, 11
*Topic size
    planning, 9
auditPrior, 1, 8, 10, 12
BuildIt, 4
evaluation, 2, 5, 10, 12
planning, 2, 8, 9, 12
sampling, 2, 8, 10, 11
```