Package 'forensic'

February 19, 2015

Type Package

Title Statistical Methods in Forensic Genetics

Version 0.2

Date 2007-06-10

Author Miriam Marusiakova (Centre of Biomedical Informatics, Institute
of Computer Science, Academy of Sciences of the Czech Republic)

Acknowledgement The work was supported by the project 1M06014 of the
Ministry of Education, Youth and Sports of the Czech Republic.

Maintainer Miriam Marusiakova <maruskay@gmail.com>

Description The statistical evaluation of DNA mixtures, DNA profile
match probability

LazyLoad yes

Depends R (>= 2.4.0), genetics, combinat

License GPL

Repository CRAN

Date/Publication 2012-10-29 08:58:46

NeedsCompilation no

R topics documented:

1

LR.ind   *Likelihood Ratio for DNA Evidence (Independent Alleles Within and Between Ethnic Groups)*

Description

Calculates likelihood ratio for DNA evidence, comparing two hypotheses about the origin of the mixture at the crime scene. Contributors to the mixture may come from one or more ethnic groups. Independence of alleles within and between ethnic groups is assumed.

Usage

```
LR.ind(alleles, prob, x1, x2, u1 = NULL, u2 = NULL)
```

Arguments

alleles   vector of distinct alleles (from one specific locus) found in the crime sample

prob   matrix of allele proportions in a population. Element (i, j) contains proportion of the i-th allele from alleles in the j-th ethnic group

x1   vector of nonnegative integers. The j-th element contains number of unknown contributors from the j-th ethnic group under the prosecution proposition.

x2   vector of nonnegative integers. The j-th element contains number of unknown contributors from the j-th ethnic group under the defence proposition.

u1   vector of alleles from the mixture, which are not carried by known contributors under the prosecution proposition. If u1=NULL (default), all alleles from the crime sample are carried by known contributors.

u2   vector of alleles from the mixture, which are not carried by known contributors under the defence proposition. Default is u2=NULL.

Details

Likelihood ratio (LR) is computed as a ratio of two conditional probabilities of DNA evidence E

$$LR = \frac{P(E|H_0)}{P(E|H_A)},$$

where H_0 denotes the prosecution proposition and H_A the defence proposition. For calculation of $P(E|H)$ see `Pevid.ind`.

Likelihood ratio is used for assigning the weight of evidence for one locus. The overall LR can be obtained by multiplying over all loci, under the assumption of independent alleles between loci.

Value

likelihood ratio for DNA evidence at the crime scene

Note

We only need to specify the ethnic groups to which the unknown contributors belong to. We do not need to care about the ethnic groups of the known contributors.

Author(s)

Miriam Marusiakova <maruskay@gmail.com>

References

Evett IW, Weir BS (1998), Interpreting DNA evidence; Statistical genetics for forensic scientists. Sinauer, Sunderland, MA.

Fukshansky N, Bar W (1998), Interpreting forensic DNA evidence on the basis of hypotheses testing. International Journal of Legal Medicine 111, 62-66.

Fung WK, Hu YQ (2001), The evaluation of mixed stains from different ethnic origins: general result and common comments. International Journal of Legal Medicine 115, 48-53.

Weir BS, Triggs CM, Starling L, Stowell LI, Walsh KAJ, Buckleton J (1997), Interpreting DNA mixtures. Journal of Forensic Sciences 42, 213-222.

See Also

Pevid.ind, Pevid.gen

Examples

```
## Simpson case (Fung and Hu (2001))
## From the crime scene: mixed profile a,b,c at a locus D2S44
m = c("a", "b", "c")
## profile of the defendant: a,b
## profile of the victim: a,c
## allele frequencies (African - American, Caucasian)
p_aa <- c(0.0316, 0.0842, 0.0926)
p_c <- c(0.0859, 0.0827, 0.1073)
## the number of contributors to the mixed sample is taken 2
## Prosecution proposition: Contributors were the victim and the suspect
## Defence proposition: Contributors were 2 unknown people
##
## Likelihood ratios for DNA evidence for different alternatives:
## two unknown people are African-American
LR.ind(alleles = m, prob = cbind(p_aa, p_c), x1 = c(0, 0),
  x2 = c(2, 0), u2 = m)
LR.ind(alleles = m, prob = p_aa, x1 = 0, x2 = 2, u2 = m)
## one unknown person is African-American and one is Caucasian
LR.ind(alleles = m, prob = cbind(p_aa, p_c), x1 = c(0, 0),
  x2 = c(1, 1), u2 = m)
## two unknown people are Caucasian
LR.ind(alleles = m, prob = cbind(p_aa, p_c), x1 = c(0, 0),
```

```
  x2 = c(0, 2), u2 = m)
LR.ind(alleles = m, prob = p_c, x1 = 0, x2 = 2, u2 = m)
```

Pevid.gen   *Probability of DNA Evidence (Allowing for Population Substructure and Dependence)*

Description

Calculates conditional probability of DNA evidence, given proposition about who the known and unknown contributors to the mixture were. All the individuals involved in a case are assumed to come from the same subpopulation with a given coancestry coefficient. Independence of alleles in the subpopulation (i.e., relatives are excluded) and dependence in the whole population is assumed.

Usage

```
Pevid.gen(alleles, prob, x, T = NULL, V = NULL, theta = 0)
```

Arguments

alleles	vector of distinct alleles (from one specific locus) found in the crime sample.
prob	vector of corresponding allele proportions in a population. The allele proportions are estimated from the whole population.
x	nonnegative integer. The number of unknown contributors to the mixture.
T	object of class genotype (package **genetics**), or a vector of strings where each string contains two alleles seperated by '/', corresponding to one known contributor. The length of the vector equals the number of known contributors. Default is NULL.
V	object of class genotype (package **genetics**), or a vector of strings where each string contains two alleles seperated by '/', corresponding to one known noncontributor. The length of the vector equals the number of known noncontributors. Default is NULL.
theta	number from the interval [0,1]. Coancestry coefficient theta describes variation in allele proportions among subpopulations. Default is 0 (no variation, whole population in Hardy-Weinberg equilibrium). The recommended values of theta are 0.01 for large subpopulations such as USA, and 0.03 for small isolated subpopulations (National Research Council, 1996).

Details

The general formula for the evaluation of the probability of DNA evidence (and its derivation) can be found in Zoubkova and Zvarova (2004).

If theta = 0, Pevid.gen returns the same numerical result as Pevid.ind.

Value

probability of the DNA evidence

Author(s)

Miriam Marusiakova <maruskay@gmail.com>

The work was supported by the project 1M06014 of the Ministry of Education, Youth and Sports of the Czech Republic.

References

Curran JM, Triggs CM, Buckleton J, Weir BS (1999), Interpreting DNA mixtures in structured populations, Journal of Forensic Sciences 44, 987-995.

Evett IW, Weir BS (1998), Interpreting DNA evidence; Statistical genetics for forensic scientists. Sinauer, Sunderland, MA

Fung WK, Hu YQ (2000), Interpreting forensic DNA mixtures: allowing for uncertainty in population substructure and dependence, Journal of the Royal Statistical Society A 163, 241-254.

National Research Council (1996), The evaluation of forensic DNA evidence National Academy Press, Washington, DC.

Zoubkova K, Zvarova J (2004), Statisticke metody ve forenzni genetice, Master's thesis, Charles University, Prague.

See Also

Pevid.ind, LR.ind, Pmatch

Examples

```
## Rape case
## The evidence profile:
m <- c(13, 14, 15)
##  the victim's genotype:
victim <- "13/14"
## the suspect's genotype
suspect <- "15/15"
## frequencies of alleles {13, 14, 15}:
p <- c(0.042, 0.166, 0.106)
## consensual partner of the victim
partner <- "15/16"
## Prosecution proposition:
## Contributors were the victim and the suspect.
Pevid.gen( alleles = m, prob = p, T = c(victim, suspect),
  V = partner, x = 0, theta = 0.03)
## Defence proposition:
## Contributors were the victim and one unknown person.
##
## Likelihood ratio for DNA evidence:
## structured population
1/Pevid.gen( alleles = m, prob = p, T = victim,
```

```
   V = c(suspect, partner), x = 1, theta = 0.03)
## Note: a person carrying both alleles different from the alleles
## in the crime sample (e.g., with genotype "16/16") has no effect
## on the value of LR:
1/Pevid.gen( alleles = m, prob = p, T = victim,
  V = c(suspect, partner, "16/16"), x = 1, theta = 0.03)
## But the consensual partner of the victim having genotype "15/16"
## influences the value of LR, compare:
1/Pevid.gen( alleles = m, prob = p, T = victim, V = suspect,
  x = 1, theta = 0.03)
##
## population in Hardy - Weinberg equilibrium
1/Pevid.gen( alleles = m, prob = p, T = victim,
  V = c(suspect, partner), x = 1)
1/Pevid.gen( alleles = m, prob = p, T = victim, x = 1)
## compare
1/Pevid.ind( alleles = m, prob = p, u = 15, x = 1)
```

Pevid.ind   *Probability of DNA Evidence (Independent Alleles Within and Between Ethnic Groups)*

Description

Calculates conditional probability of DNA evidence, given proposition about who the known and unknown contributors to the mixture were. Contributors to the mixture may come from one or more ethnic groups. Independence of alleles within and between ethnic groups is assumed.

Usage

```
Pevid.ind(alleles, prob, x, u = NULL)
```

Arguments

alleles	vector of distinct alleles (from one specific locus) found in the crime sample
prob	matrix of allele proportions in a population. Element (i, j) contains proportion of the i-th allele from alleles in the j-th ethnic group
x	vector of nonnegative integers. The j-th element contains number of unknown contributors from the j-th ethnic group.
u	vector of alleles from the mixture, which are not carried by known contributors (and have to be carried by some unknown contributors). If u=NULL (default), all alleles from the crime sample are carried by known contributors.

Details

The formula for the evaluation of the probability of DNA evidence (and its derivation) can be found in Fukshansky and Bar (1998) and in Fung and Hu (2001) (different derivation of the formula).

Value

probability of DNA evidence

Author(s)

Miriam Marusiakova <maruskay@gmail.com>

References

Evett IW, Weir BS (1998), Interpreting DNA evidence; Statistical genetics for forensic scientists. Sinauer, Sunderland, MA.

Fukshansky N, Bar W (1998), Interpreting forensic DNA evidence on the basis of hypotheses testing. International Journal of Legal Medicine 111, 62-66.

Fung WK, Hu YQ (2000), Interpreting DNA mixtures based on the NRC-II recommendation 4.1. Forensic Sci Commun. Available at http://www.fbi.gov/hq/lab/fsc/backissu/oct2000/fung.htm

Fung WK, Hu YQ (2001), The evaluation of mixed stains from different ethnic origins: general result and common comments. International Journal of Legal Medicine 115, 48-53.

Fung WK, Hu YQ (2002), The statistical evaluation of DNA mixtures with contributors from different ethnic groups. International Journal of Legal Medicine 116, 79-86.

Weir BS, Triggs CM, Starling L, Stowell LI, Walsh KAJ, Buckleton J (1997), Interpreting DNA mixtures. Journal of Forensic Sciences 42, 213-222.

See Also

LR.ind, Pevid.gen

Examples

```
## Simpson case (Fung and Hu (2001))
## From the crime scene: mixed profile a,b,c at the locus D2S44
m = c("a", "b", "c")
## profile of the defendant: a,b
## profile of the victim: a,c
## the number of contributors to the mixed sample is taken 2
## Prosecution proposition: Contributors were the victim and the suspect.
## (The defendant was an African-American and the victim was a Caucasian,
## but we do not need this information for the calculation of LR)
## Defence proposition: Contributors were 2 unknown people.
## allele frequencies (African - American, Caucasian)
p_aa <- c(0.0316, 0.0842, 0.0926)
p_c <- c(0.0859, 0.0827, 0.1073)
##
## Likelihood ratios for DNA evidence for different alternatives:
## two unknown contributors were African-American
1/Pevid.ind(alleles = m, prob = cbind(p_aa, p_c), x = c(2, 0), u = m)
1/Pevid.ind(alleles = m, prob = p_aa, x = 2, u = m)
```

```
## one unknown contributor was African-American and one was Caucasian
1/Pevid.ind(alleles = m, prob = cbind(p_aa, p_c), x = c(1, 1), u = m)
## two unknown contributors were Caucasian
1/Pevid.ind(alleles = m, prob = cbind(p_aa, p_c), x = c(0, 2), u = m)
1/Pevid.ind(alleles = m, prob = p_c, x = 2, u = m)

## Rape case in Hong Kong (Fung and Hu (2000), Fung and Hu (2002))
## the mixed stain at D3S1358
m = c(14, 15, 17, 18)
## suspect's genotype (14, 17)
## If the mixed stain did not originate from the victim,
## we may consider the propositions:
## Prosecution: The contributors were the suspect and an unknown
## Defence: The contributors were two unknowns
## Arrested suspect was a Caucasian
## Possible ethnic groups of the unknowns:
## Caucasian, Chinese
p_ca = c(0.187, 0.213, 0.223, 0.127)
p_ch = c(0.033, 0.331, 0.239, 0.056)
## the both unknowns are Caucasian
Pevid.ind(alleles = m, prob = p_ca, x = 1, u = c(15, 18))/
  Pevid.ind(alleles = m, prob = p_ca, x = 2, u = m)
## first unknown is Caucasian, the second one Chinese
Pevid.ind(alleles = m, prob = p_ca, x = 1, u = c(15, 18))/
  Pevid.ind(alleles = m, prob = cbind(p_ca, p_ch), x = c(1, 1), u = m)
## first unknown is Chinese, the second one Caucasian
Pevid.ind(alleles = m, prob = p_ch, x = 1, u = c(15, 18))/
  Pevid.ind(alleles = m, prob = cbind(p_ca, p_ch), x = c(1, 1), u = m)
## the both unknowns are Chinese
Pevid.ind(alleles = m, prob = p_ch, x = 1, u = c(15, 18))/
  Pevid.ind(alleles = m, prob = p_ch, x = 2, u = m)
##
## Suppose the second suspect (Caucasian, (15, 18)) was arrested
## Prosecution: The contributors were the two suspects
## Defence: The contributors were two unknowns
## the both unknowns are Caucasian
1/Pevid.ind(alleles = m, prob = p_ca, x = 2, u = m)
## first unknown is Caucasian, the second one Chinese
1/Pevid.ind(alleles = m, prob = cbind(p_ca, p_ch), x = c(1, 1), u = m)
## first unknown is Chinese, the second one Caucasian
1/Pevid.ind(alleles = m, prob = cbind(p_ca, p_ch), x = c(1, 1), u = m)
## the both unknowns are Chinese
1/Pevid.ind(alleles = m, prob = p_ch, x = 2, u = m)
```

Pevid.rel   *Probability of DNA Evidence With the Presence of Relatives*

Description

Computes probability of DNA evidence, given the proposition about who the contributors to the mixture were. Two main cases are considered: tested suspect (noncontributor) with an unknown relative (contributor), or unknown suspect (contributor) with a tested relative (noncontributor) and two unknown related people as contributors.

Usage

```
Pevid.rel(alleles, prob, x, u = NULL, k = c(1, 0, 0), S = NULL)
```

Arguments

alleles	vector of distinct alleles (from one specific locus) found in the crime sample.
prob	vector of corresponding allele proportions in a population
x	nonnegative integer. The number of unknown contributors to the mixture.
u	vector of alleles from the mixture, which are not carried by known contributors (and have to be carried by some unknown contributors). If u=NULL (default), all alleles from the crime sample are carried by known contributors.
k	vector of kinship coefficients (k_0, k_1, k_2), where k_i is the probability that two people will share i alleles identical by descent, $i = 0, 1, 2$.
S	object of class genotype (package **genetics**), or a string of length 1 with two alleles seperated by '/'. Genotype of the typed person (declared noncontributor) whose untyped relative is assumed to be a contributor to the sample. If S = NULL (default), we consider two related unknown contributors.

Details

Table of kinship coefficients for commonly encountered relationships:

Relationship	k_0	k_1	k_2
Parent - child	0	1	0
Grandparent - grandchild	1/2	1/2	0
Full sibs	1/4	1/2	1/4
Halfsibs	1/2	1/2	0
Uncle - nephew	1/2	1/2	0
First cousins	3/4	1/4	0
Second cousins	15/16	1/16	0
Unrelated	1	0	0

The formulas (and their derivations) for the evaluation of the probabilities of DNA evidence with the presence of at most 2 relatives can be found in Hu and Fung (2003).

Value

probability of DNA evidence

Author(s)

Miriam Marusiakova <maruskay@gmail.com>

References

Balding DJ, Nichols RA (1994), DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands. Forensic Science International 64, 125-140.

Fung WK, Hu YQ (2000), Interpreting DNA mixtures based on the NRC-II recommendation 4.1. Forensic Sci Commun. Available at http://www.fbi.gov/hq/lab/fsc/backissu/oct2000/fung.htm

Hu YQ, Fung WK (2003), Interpreting DNA mixtures with the presence of relatives. International Journal od Legal Medicine 117, 39-45.

Evett IW, Weir BS (1998), Interpreting DNA evidence. Statistical genetics for forensic scientists. Sinauer, Sunderland, MA.

See Also

[Pmatch](#)

Examples

```
## Rape case in Hong Kong (Fung and Hu (2000))
## mixture (loci: D3S1358, vWA, FGA)
m1 <- c(14, 15, 17, 18)
m2 <- c(16, 18)
m3 <- c(20, 24, 25)
## genotype of the victim (loci: D3S1358, vWA, FGA):
victim_1 = "15/18"
victim_2 = "18/18"
victim_3 = "20/24"
## genotype of the suspect (loci: D3S1358, vWA, FGA):
suspect_1 = "14/17"
suspect_2 = "16/16"
suspect_3 = "25/25"
## allele proportions
p1 <- c(0.033, 0.331, 0.239, 0.056)
p2 <- c(0.155, 0.158)
p3 <- c(0.042, 0.166, 0.106)
##
## Likelihood ratio for DNA evidence:
## Prosecution proposition:
## Contributors were the victim and the suspect.
## Defence proposition 1:
## Contributors were the victim and one relative of the suspect (sibling).
print(LR11 <- 1 / Pevid.rel(alleles = m1, prob = p1, x = 1,
  k = c(1/4, 1/2, 1/4), S = suspect_1, u = c(14, 17)))
print(LR12 <- 1 / Pevid.rel(alleles = m2, prob = p2, x = 1,
  k = c(1/4, 1/2, 1/4), S = suspect_2, u = 16))
```

```
print(LR13 <- 1 / Pevid.rel(alleles = m3, prob = p3, x = 1,
  k = c(1/4, 1/2, 1/4), S = suspect_3, u = 25))
##
## Defence proposition 2:
## Contributors were one relative of the suspect (sibling) and one unknown.
print(LR21 <- 1 / Pevid.rel(alleles = m1, prob = p1, x = 2,
  k = c(1/4, 1/2, 1/4), S = suspect_1, u = m1))
print(LR22 <- 1 / Pevid.rel(alleles = m2, prob = p2, x = 2,
k = c(1/4, 1/2, 1/4), S = suspect_2, u = m2))
print(LR23 <- 1 / Pevid.rel(alleles = m3, prob = p3, x = 2,
k = c(1/4, 1/2, 1/4), S = suspect_3, u = m3))
##
## Defence proposition 3:
## Contributors were two related people (siblings).
print(LR31 <- 1 / Pevid.rel(alleles = m1, prob = p1, x = 2,
k = c(1/4, 1/2, 1/4), u = m1))
print(LR32 <- 1 / Pevid.rel(alleles = m2, prob = p2, x = 2,
k = c(1/4, 1/2, 1/4), u = m2))
print(LR33 <- 1 / Pevid.rel(alleles = m3, prob = p3, x = 2,
k = c(1/4, 1/2, 1/4), u = m3))
##
## Likelihood ratios overall:
## for defence proposition 1
LR11*LR12*LR13
## for defence proposition 2
LR21*LR22*LR23
## for defence proposition 3
LR31*LR32*LR33
```

Pmatch   *Match Probabilities of Genotype*

Description

Calculates match probability of the genotype of the suspect and that of the crime stain presumed
to have come from an offender other than the suspect. Possible assumptions: the suspect and an
unknown offender are unrelated, or are members of the same subpopulation with a given coancestry
coefficient, or are close relatives.

Usage

```
Pmatch(prob, k = c(1, 0, 0), theta = 0)
```

Arguments

prob   matrix with 2 rows and L columns (L is the number of loci, each locus has 2
                alleles). Contains frequencies of alleles in a population found in the crime stain.
                For homozygous locus just one entry is nonzero. prob can also be a vector with
                odd number of elements (it is then easily transformed to a matrix with two rows,
                the matrix is filled by columns)

k	vector of kinship coefficients (k_0, k_1, k_2), where k_i is the probability that two people (the suspect and an unknown offender) will share i alleles identical by descent, $i = 0, 1, 2$.
theta	number from the interval [0,1]. Coancestry coefficient theta describes variation in allele proportions among subpopulations. Default is 0 (no variation, whole population in Hardy-Weinberg equilibrium). The recommended values of theta are 0.01 for large subpopulations such as USA, and 0.03 for small isolated sub-populations (National Research Council, 1996). If theta is nonzero, the allele proportions are taken from the whole population (not from the subpopulation).

Details

The match probability is calculated as

$$k_2 + k_1 Z_1 + k_0 Z_2,$$

where k_0, k_1, k_2 are the kinship coefficients (for more information see Details of `Pevid.rel`),

$$Z_2 = \frac{[2\theta + (1-\theta)p_i]\,[3\theta + (1-\theta)p_i]}{(1+\theta)(1+2\theta)},$$

$$Z_2 = \frac{2\,[\theta + (1-\theta)p_i]\,[\theta + (1-\theta)p_j]}{(1+\theta)(1+2\theta)}$$

are the match probabilities in the unrelated case for homozygotes and heterozygotes, respectively, and

$$Z_1 = \frac{2\theta + (1-\theta)p_i}{1+\theta}$$

for the homozygote case and

$$Z_1 = \frac{2\theta + (1-\theta)(p_i + p_j)}{2(1+\theta)}$$

for the heterozygote case. The quantity θ is the coancestry population theta. The formula is derived in Balding and Nichols (1994).

The match probability at all loci is calculated as a product of all single locus probabilities. We assume independence across loci.

Value

Pmatch returns a list with the following components:

prob	matrix of allele proportions at each locus (input value in `Pmatch`)
match	single locus match probabilities
total_match	match probability of genotype = multiplication of single locus match probabilities

Author(s)

Miriam Marusiakova <maruskay@gmail.com>

References

Balding DJ, Nichols RA (1994), DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands. Forensic Science International 64, 125-140.

Evett IW, Weir BS (1998), Interpreting DNA evidence; Statistical genetics for forensic scientists. Sinauer, Sunderland, MA.

National Research Council (1996), The evaluation of forensic DNA evidence National Academy Press, Washington, DC.

See Also

Pevid.rel, Pevid.gen

Examples

```
## match probability of thirteen-locus genotype
    ## (11 heterozygous and 2 homozygous loci)
    p<-c(0.057,0.160,0.024,0.122,0.078,0.055,0.035,0.150,
      0.195,0.027,0.084,0.061,0.122,0.083,0.164,0.065,0.143,
      0.151,0.167,0.180,0.099,0.182,0.120,0,0.182,0)
    ## suspect and offender are unrelated
    Pmatch(p)
    ## suspect and offender are unrelated but members of the same
    ## subpopulation with the coancestry coefficient theta
    Pmatch(p, theta = 0.03)
    ## suspect and offender are close relatives (cousins)
    Pmatch(p, k = c(3/4, 1/4, 0))
    ## suspect and offender are close relatives (cousins) and
    ## members of the same subpopulation with the coancestry
    ## coefficient theta
    Pmatch(p, k = c(3/4, 1/4, 0), theta = 0.03)
    ##
    ## one locus
    Pmatch(p[1:2], theta = 0.03)
    Pmatch(p[25:26], theta = 0.03)
    ## compare
    Pevid.gen(alleles = c(1, 2), prob = p[1:2], V = "1/2", x = 1,
      theta = 0.03)
    Pevid.gen(alleles = "a", prob = p[25], V = "a/a", x = 1,
      theta = 0.03)
```

Index