

EpiStats

Jean Pierre Decorps / Esther Kissling

2018-10-08

Epiconcept is made up of a team of doctors, epidemiologists, data scientists and digital specialists. For more than 20 years, Epiconcept has contributed to the improvement of public health by writing software, carrying out epidemiological studies, research, evaluation and training to better detect, monitor and prevent disease and to improve treatment.

Epiconcept provides software, services and studies in the following areas:

- Software for managing public health programs,
- Secure cloud solutions for health data collection, reporting and processing,
- Research projects on vaccine preventable diseases, including measuring the effectiveness and impact of vaccines,
- Services in the field of epidemiology (protocols, analysis, training, etc.),
- Expertise in data analysis,
- Coaching and assistance to professionals in public health,
- Training (in software use and epidemiology: short and longer introductory modules, advanced courses, training through long-term practice).

To achieve such goals Epiconcept :

- Recognized research organization,
- Certified datacenter for hosting personal health data,
- Training organisation.

Epiconcept relies on :

- Its expertise in epidemiology
- Its IT expertise,
- Ethical values rooted in practice (responsibility and quality of services, data security and confidentiality, scientific independence, etc.),
- Capabilities to answer and anticipate tomorrow's challenges (Research - evaluation, e-health, Big Data, IoT, etc.),
- A desire to build long-term relationships with its clients and partners.

Its current customers and partners include some of the greatest names in the world such as: Santé Publique France (and many public health organizations around the world), WHO, ECDC, AFD, MSF, World Bank, etc.

Package EpiStats

Description

The EpiStats package is a set of functions aimed at epidemiologists. They include commands for measures of association and impact for case control studies and cohort studies. They may be particularly useful for outbreak investigations and include univariate and stratified analyses.

The generic function *crossTable* provides a contingency table with optional parameters *percent* and *statistic*

The functions for cohort studies include the *CS*, *CSTable* and *CSInter* commands.

The functions for case control studies include the *CC*, *CCTable* and *CCInter* commands.

All variables used need to be numeric binary variables and coded as 0 and 1 or as factors.

Cohort study functions:

The cohort study functions relate to cohort studies that measure risks, rather than rates in person- time.

The **CS** function provides a 2 by 2 table and measures the association between the outcome and one exposure. It includes the risk ratio and its 95% confidence intervals, the attributable fraction among the exposed and unexposed, and a chi square test and its p-value.

The **CSTable** function displays the measures of association between the outcome and a set of exposures in a table (risk ratios, confidence intervals and p-values). This helps the researcher to compare between exposures and provides a nice table for reports.

The **CSInter** function investigates the effect of a third variable on the association between an exposure and the outcome. It presents two by two tables stratified by the levels of a third value. It provides the Woolf test for homogeneity between stratum-specific risk ratios. It provides the crude risk ratio between an exposure and an outcome and the risk ratio adjusted by the third variable. *CSInter* helps the researcher understand whether a third variable may have an effect modifying or confounding effect on the association between an exposure and the outcome.

Case control study functions:

The **CC** function provides a 2 by 2 table and measures the association between the outcome and one exposure. It includes the odds ratio and its 95% confidence intervals, the attributable fraction among the exposed, and a chi square test and its p-value.

The **CCTable** function displays the measures of association between the outcome and a set of exposures in a table (odds ratios, confidence intervals and p-values). This helps the researcher to compare between exposures and provides a nice table for reports.

The **CCInter** function investigates the effect of a third variable on the association between an exposure and the outcome. It presents two by two tables stratified by the levels of a third value. It provides the Woolf test for homogeneity between stratum-specific odds ratios. It provides the crude odds ratio between an exposure and an outcome and the odds ratio adjusted by the third variable. *CCInter* helps the researcher understand whether a third variable may have an effect modifying or confounding effect on the association between an exposure and the outcome.

The “Tiramisu” dataset

The dataset used in this vignette is from an outbreak investigation carried out in Germany in 1998 by Anja Hauri, Robert Koch Institute. It is used in case studies by organisations including EPIET, ECDC and EpiConcept.

The **CSTable**, **CSInter**, **CCTable** and **CCInter** functions are based on commands written in Stata by *Gilles Desve*, who we gratefully acknowledge.

Working with EpiStats and “Tiramisu” dataset

Loading and recoding the dataset

```
library(EpiStats)
library(dplyr)
library(knitr)

options(knitr.kable.NA = '')
options(width=200)

data(Tiramisu)
DF <- Tiramisu

DF <- DF %>%
  # filter(age != "NA") %>%
  mutate(agegroup = case_when(age < 30 ~ 0, age >= 30 ~ 1)) %>%
  mutate(tportion = case_when(tportion == 0 ~ 0, tpotion == 1 ~ 1, tpotion >= 2 ~ 2)) %>%
  mutate(tportion = as.factor(tportion)) %>%
  as.data.frame()

Colnames <- DF %>%
  select(-ill, -age, -dateonset, -uniquekey, -tportion, -mportion) %>%
  colnames()
```

crossTable

Creates a contingency table of variable of interest and exposure. Percentage are optionals by *row* or by *column*. It can provides an optional statistic (*fisher* or *chisquare*).

Syntax

```
crossTable(data, var1, var2, percent="none", statistic="none")
```

Examples

Recoding some data to have ordered factors

```
DF$ill <- factor(DF$ill, levels=c(1,0), ordered = TRUE)
DF$sex <- factor(DF$sex, levels = c("males", "females"), ordered = TRUE)
```

Example 1: crossTable ill - sex

```
ret <- crossTable(DF, var1="ill", var2="tira")
ret
```

```
##  tira / ill   1   0 Total
## 1         0   7 158  165
## 2         1  94  27  121
## 3      Total 101 185  286
```

```
kable(ret, align="r")
```

tira / ill	1	0	Total
0	7	158	165
1	94	27	121
Total	101	185	286

Example 2: crossTable ill - sex with column percentage and chi2 stat

```
ret <- crossTable(DF, "ill", "sex", "col", "chi2")
kable(ret, align="r", caption = "with columns %")
```

Table 2: with columns %

sex / ill	1	0	Total
males	50	102	152
%	48.54	54.26	52.23
females	53	86	139
%	51.46	45.74	47.77
Total	103	188	291
%	100.00	100.00	100.00
-	-	-	-
Pearson CHI2	0.8701	Pr	0.351

Example 3: CrossTable ill - sex with row percentage and Fisher stat

NB: All parameters are unquoted

```
ret <- crossTable(DF, ill, sex, row, fisher)
ret
```

```
##      sex / ill      1      %      0      % Total      %
## 1      males      50 32.89 102 67.11   152 100
## 2      females      53 38.13   86 61.87   139 100
## 3      Total      103 35.40 188 64.60   291 100
## 4      -          -      -      -      -      -      -
## 5 Fisher's exact 0.391
```

```
kable(ret, align="r", caption = "with rows %")
```

Table 3: with rows %

sex / ill	1	%	0	%	Total
males	50	32.89	102	67.11	152
females	53	38.13	86	61.87	139
Total	103	35.40	188	64.60	291
-	-	-	-	-	-
Fisher's exact	0.391				

CrossTable beer - sex with column and row percentages and Chi2 stat

NB: All parameters are unquoted

```
ret <- crossTable(DF, beer, sex, both, chi2)
ret
```

```
##      sex / beer      0      %      1      % Total      %
## 1      males      58 40.85      84 59.15   142 100.00
## 2      %      35.15      79.25
## 3      females      107 82.95      22 17.05   129 100.00
## 4      %      64.85      20.75
## 5      Total      165 60.89      106 39.11   271 100.00
## 6      %      100.00      100.00      100.00
## 7      -          -          -          -          -
## 8 Pearson CHI2 50.3078      Pr      0
```

```
kable(ret, align="r", caption = "% rows and columns")
```

Table 4: % rows and columns

sex / beer	0	%	1	%	Total
males	58	40.85	84	59.15	142
%	35.15		79.25		
females	107	82.95	22	17.05	129
%	64.85		20.75		
Total	165	60.89	106	39.11	271
%	100.00		100.00		100.00
-	-	-	-	-	-
Pearson CHI2	50.3078	Pr	0		

CS

CS analyses cohort studies with equal follow-up time per subject. The risk (the proportion of individuals who become cases) is calculated overall and among the exposed and unexposed. Note that all variables need to be numeric and binary and coded as "0" and "1".

Point estimates and confidence intervals for the risk ratio and risk difference are calculated, along with attributable or preventive fractions for the exposed and the total population. Additionally you can select if you want to display the Fisher's exact test, by specifying `exact = TRUE`. If you specify `full = TRUE` you can easily access useful statistics from the output tables.

Syntax

```
CS(x, cases, exposure, exact, full=FALSE)
```

Example 1: CS ill - mousse (unformatted)

```
CS(DF, "ill", "mousse", exact = FALSE)
```

```
## $df1
##           Cases Non Cases Total Risk
## Exposed      42         81   123 0.34
## Unexposed    144         22   166 0.87
## Total        186        103   289 0.64
##
## $df2
##           Point estimate 95%CI ll 95%CI ul
## Risk difference      -0.53   -0.62   -0.43
## Risk ratio           0.39    0.31    0.51
## Prev. frac. ex.      0.61    0.49    0.69
## Prev. frac. pop      0.26     NA     NA
## chi2(1)             85.22     NA     NA
## Pr>chi2              0.000     NA     NA
```

Example 2: CS ill - beer (formatted)

The following results tables are outputs in “markdown” using the *kable* function.

```
result <- CS(DF, "ill", "beer", exact = TRUE, full = TRUE)
kable(result$df1)
```

	Cases	Non Cases	Total	Risk
Exposed	76	30	106	0.72
Unexposed	96	69	165	0.58
Total	172	99	271	0.63

```
kable(result$df2, align = result$df2.align )
```

	Point estimate	95%CI ll	95%CI ul
Risk difference	0.14	0.02	0.25
Risk ratio	1.23	1.03	1.47
Attr. frac. ex.	0.19	0.03	0.32
Attr. frac. pop	0.08	NA	NA
chi2(1)	5.09	NA	NA
Pr>chi2	0.024	NA	NA
Fisher p.value	0.028	NA	NA

By storing the results in the object “result”, you are able to use the result tables in Markdown as shown above. By specifying “full = TRUE” you can also easily use individual elements of the results. For example if you would like to view just the risk ratio, you can view it by typing:

```
result$st$risk_ratio$point_estimate
```

```
## [1] 1.232311
```

CSTable - Summary table for cohort studies

CSTable is used for univariate analysis of cohort studies with several exposures. The results are summarised in one table with one row per exposure making comparisons between exposures easier and providing a useful table for integrating into reports. Note that all variables need to be numeric and binary and coded as “0” and “1”.

The results of this function contain: The name of exposure variables, the total number of exposed, the number of exposed cases, the attack rate among the exposed, the total number of unexposed, the number of unexposed cases, the attack rate among the unexposed, risk ratios, 95% confidence intervals, 95% p-values.

You can optionally choose to display the Fisher’s exact p-value instead of the Chi squared p-value, with the option `exact = TRUE`.

You can specify the sort order, with the option `sort=“rr”` to order by risk ratios. The default sort order is by p-values.

The option `“full = TRUE”` provides you with useful formatting information, which can be handy if you’re using “markdown”.

Syntax

```
CSTable(x, cases, exposure=c(), exact=FALSE, sort = “pvalue”, full=FALSE)
```

Example 1: CSTable results ordered by p-value (unformatted)

```
CSTable(DF,
  "ill",
  exposure = c("sex", "agegroup", "tira", "beer", "mousse", "wmousse", "dmousse",
    "redjelly", "fruitsalad", "tomato", "mince", "salmon", "horseradish",
    "chickenwin", "roastbeef", "pork"))
```



```
## $df
##          Tot.Exp. Exp.Cases  AR% Tot.Unex. Unex.Cases  AR%  RR CI ll CI ul p(Chi2)
## tira          121      27 22.31      165      158 95.76 0.23 0.17 0.33 0.000
## mousse        123      42 34.15      166      144 86.75 0.39 0.31 0.51 0.000
## wmousse        72      23 31.94      205      156 76.10 0.42 0.30 0.59 0.000
## dmousse       113      37 32.74      174      148 85.06 0.38 0.29 0.51 0.000
## redjelly       79      34 43.04      212      154 72.64 0.59 0.45 0.77 0.000
## fruitsalad    71      25 35.21      220      163 74.09 0.48 0.34 0.66 0.000
## beer          106      76 71.70      165       96 58.18 1.23 1.03 1.47 0.024
## tomato         83      48 57.83      208      140 67.31 0.86 0.70 1.06 0.127
## pork          120      72 60.00      169      115 68.05 0.88 0.74 1.05 0.158
## horseradish   72      42 58.33      217      145 66.82 0.87 0.70 1.08 0.192
## sex           139      86 61.87      152      102 67.11 0.92 0.78 1.09 0.351
## roastbeef     29      21 72.41      262      167 63.74 1.14 0.89 1.45 0.354
## chickenwin    84      51 60.71      207      137 66.18 0.92 0.75 1.12 0.377
## mince         87      55 63.22      204      133 65.20 0.97 0.80 1.17 0.747
## agegroup      68      43 63.24      215      140 65.12 0.97 0.79 1.19 0.777
## salmon       104      67 64.42      183      120 65.57 0.98 0.82 1.17 0.844
```


Example 2: CStable results ordered by risk ratio (formatted)

The following results tables are outputs in “markdown” using the kable function.

```
res = CStable(DF, "ill", sort = "rr", exposure = Colnames, full = TRUE)
kable(res$df, digits=res$digits, align=res$align)
```

	Tot.Exp.	Exp.Cases	AR%	Tot.Unex.	Unex.Cases	AR%	RR	CI ll	C
beer	106	76	71.70	165	96	58.18	1.23	1.03	
roastbeef	29	21	72.41	262	167	63.74	1.14	0.89	
salmon	104	67	64.42	183	120	65.57	0.98	0.82	
mince	87	55	63.22	204	133	65.20	0.97	0.80	
agegroup	68	43	63.24	215	140	65.12	0.97	0.79	
sex	139	86	61.87	152	102	67.11	0.92	0.78	
chickenwin	84	51	60.71	207	137	66.18	0.92	0.75	
pork	120	72	60.00	169	115	68.05	0.88	0.74	
horseradish	72	42	58.33	217	145	66.82	0.87	0.70	
tomato	83	48	57.83	208	140	67.31	0.86	0.70	
redjelly	79	34	43.04	212	154	72.64	0.59	0.45	
fruitsalad	71	25	35.21	220	163	74.09	0.48	0.34	
wmousse	72	23	31.94	205	156	76.10	0.42	0.30	
mousse	123	42	34.15	166	144	86.75	0.39	0.31	
dmousse	113	37	32.74	174	148	85.06	0.38	0.29	
tira	121	27	22.31	165	158	95.76	0.23	0.17	

Example 3: CStable results ordered by p-value from the Fisher's exact test (formatted)

The following results tables are outputs in "markdown" using the kable function.

```
res = CStable(DF, "ill", exact = TRUE, exposure = Colnames, full = TRUE)
kable(res$df, digits=res$digits, align=res$align)
```

	Tot.Exp.	Exp.Cases	AR%	Tot.Unex.	Unex.Cases	AR%	RR	CI ll	C
tira	121	27	22.31	165	158	95.76	0.23	0.17	
wmousse	72	23	31.94	205	156	76.10	0.42	0.30	
dmousse	113	37	32.74	174	148	85.06	0.38	0.29	
mousse	123	42	34.15	166	144	86.75	0.39	0.31	
redjelly	79	34	43.04	212	154	72.64	0.59	0.45	
fruitsalad	71	25	35.21	220	163	74.09	0.48	0.34	
beer	106	76	71.70	165	96	58.18	1.23	1.03	
tomato	83	48	57.83	208	140	67.31	0.86	0.70	
pork	120	72	60.00	169	115	68.05	0.88	0.74	
horseradish	72	42	58.33	217	145	66.82	0.87	0.70	
sex	139	86	61.87	152	102	67.11	0.92	0.78	
roastbeef	29	21	72.41	262	167	63.74	1.14	0.89	
chickenwin	84	51	60.71	207	137	66.18	0.92	0.75	
agegroup	68	43	63.24	215	140	65.12	0.97	0.79	
mince	87	55	63.22	204	133	65.20	0.97	0.80	
salmon	104	67	64.42	183	120	65.57	0.98	0.82	

By storing the results in the object "res", you are able to use the result table in Markdown as shown above. You can also use individual elements of the results. For example if you would like to view just the risk ratio, you can view it by typing (for example):

```
res$df$'Risk Ratio'[2]
```

```
## NULL
```

CSInter - Stratified analysis for cohort studies

CSInter is useful to determine the effects of a third variable on the association between an exposure and an outcome. CSInter produces 2 by 2 tables with stratum specific risk ratios, attributable risk among exposed and population attributable risk. Note that the outcome and exposure variable need to be numeric and binary and coded as "0" and 1". The third variable needs to be numeric, but may have more categories, such as "0", "1" and "2".

CSInter displays a summary with the crude RR, the Mantel Haenszel adjusted RR and the result of a "Woolf" test for homogeneity of stratum-specific RR.

The option "full = TRUE" provides you with useful formatting information, which can be handy if you're using "markdown".

Syntax

```
CSInter(x, cases, exposure, by, full=FALSE)
```

Example 1 : CSInter ill - wmousse by tira (unformatted)

```
CSInter(DF, cases="ill", exposure = "wmousse", by = "tira")
```

```
## , , tira = 0
##
##      ill
## wmousse  0  1
##      1  13  4
##      2 141  3
##
## , , tira = 1
##
##      ill
## wmousse  0  1
##      1   9 43
##      2  14 46

## $df1
##   CSInter ill - wmousse by(tira) Total Cases Risk %      P.est. Stats 95%CI-l1 95%CI-ul
## 1          tira = 1    112 <NA>      NA Risk difference -0.06    -0.21    0.09
## 2          Exposed     52   9 17.31    Risk ratio  0.74     0.35    1.57
## 3          Unexposed    60  14 23.33 Prev. frac. ex. 0.26    -0.57    0.65
## 4          NA <NA>      NA Prev. frac. pop 0.12     NA      NA
## 5          tira = 0   161 <NA>      NA Risk difference -0.21    -0.42   -0.01
## 6          Exposed     17  13 76.47    Risk ratio  0.78     0.60    1.02
## 7          Unexposed   144 141 97.92 Prev. frac. ex. 0.22    -0.02    0.40
## 8          NA <NA>      NA Prev. frac. pop 0.02     NA      NA
## 9          Missing / Missing % 18 6.2%    NA          <NA>    NA      NA
##
## $df2
##          Point Estimate Chi2 p.value Stats 95%CI-l1 95%CI-ul
## 1      Woolf test of homogeneity 0.02  0.899    NA      NA      NA
## 2          Crude RR for wmousse    NA     NA 0.42    0.29    0.60
## 3 MH RR wmousse adjusted for tira    NA     NA 0.77    0.57    1.03
## 4 Adjusted/crude relative change    NA     NA 83.27    NA      NA
```

Example 2 : CSInter ill - beer by tira (formatted)

The following results tables are outputs in “markdown” using the kable function.

```
res <- CSInter(DF, "ill", "beer", "tira", full = TRUE)
```

```
## , , tira = 0
##
##   ill
## beer 0 1
##   1 60 3
##   2 83 4
##
## , , tira = 1
##
##   ill
## beer 0 1
##   1 14 27
##   2 12 63
```

CSInter ill - beer by(tira)	Total	Cases	Risk %	P.est.	Stats	95%CI-ll
tira = 1	116		NA	Risk difference	0.18	0.01
Exposed	41	14	34.15	Risk Ratio	2.13	1.09
Unexposed	75	12	16.00	Attrib.risk.exp	0.53	0.08
			NA	Attrib.risk.pop	0.29	NA
tira = 0	150		NA	Risk difference	-0.00	-0.07
Exposed	63	60	95.24	Risk ratio	1.00	0.93
Unexposed	87	83	95.40	Prev. frac. ex.	0.00	-0.07
			NA	Prev. frac. pop	0.00	NA
Missing / Missing %	25	8.6%	NA		NA	NA

Point Estimate	Chi2	p.value	Stats	95%CI-ll	95
Woolf test of homogeneity	4.88	0.027	NA	NA	
Crude RR for beer	NA		1.21	1.02	
MH RR beer	NA		1.12	1.00	
adjusted for tira					
Adjusted/crude relative change	NA		-7.57	NA	

Example 3: CSInter ill - beer by tportion (formatted)

The following results tables are outputs in “markdown” using the kable function.

```
res <- CSInter(DF, "ill", "beer", "tportion", full = TRUE)
```

```
## , , tportion = 0
##
##   ill
## beer 0 1
##   1 60 3
##   2 83 4
##
## , , tportion = 1
##
##   ill
## beer 0 1
##   1 12 10
##   2  8 33
##
## , , tportion = 2
##
##   ill
## beer 0 1
##   1  2 17
##   2  4 30
```

```
kable(res$df1, align="r")
```

CSInter ill - beer by(tportion)	Total	Cases	Risk %	P.est.	Stats	95%CI
tportion = 2	53		NA	Risk difference	-0.01	-0.01
Exposed	19	2	10.53	Risk ratio	0.89	0.89
Unexposed	34	4	11.76	Prev. frac. ex.	0.11	-0.01
			NA	Prev. frac. pop	0.04	0.04
tportion = 1	63		NA	Risk difference	0.35	0.35
Exposed	22	12	54.55	Risk Ratio	2.80	1.02
Unexposed	41	8	19.51	Attrib.risk.exp	0.64	0.64
			NA	Attrib.risk.pop	0.39	0.39
tportion = 0	150		NA	Risk difference	-0.00	-0.00
Exposed	63	60	95.24	Risk ratio	1.00	1.00
Unexposed	87	83	95.40	Prev. frac. ex.	0.00	-0.00
			NA	Prev. frac. pop	0.00	0.00
Missing / Missing %	25	8.6%	NA		NA	NA

```
kable(res$df2, align="r")
```

Point Estimate	Chi2	p.value	Stats	95%CI-ll	95%CI-ul
Woolf test of homogeneity	7.60	0.022	NA	NA	NA
Crude RR for beer	NA		1.21	1.02	1.40
MH RR beer adjusted for tportion	NA		1.12	1.00	1.24
Adjusted/crude relative change	NA		-7.45	NA	NA

By storing the results in the object “res”, you are able to use the result table in Markdown as shown above. You can also use individual elements of the results. For example if you would like to view just the Mantel-Haenszel risk ratio for beer adjusted for tportion, you can view it by typing:

```
res$df2$Stats[3]
```

```
## [1] 1.12
```

```
## Levels: NA -7.45 1.12 1.21
```

CC

CC is used for case control studies to determine the association between an exposure and an outcome. Variables need to be binary and coded as “0” and “1”. Point estimates and confidence intervals for the odds ratio are calculated along with attributable or preventive fractions for the exposed and total population. Additionally you can select if you want to display the Fisher’s exact test, by specifying `exact = TRUE`. If you specify `full = TRUE` you can easily access useful statistics from the output tables.

Syntax

`CC(x, cases, exposure, exact, full=FALSE)`

Example 1: CC ill - mousse (unformatted)

```
cc(DF, "ill", "mousse", exact = TRUE)
```

```
## $df1
##           Cases Controls Total
## Exposed           42      81  123
## Unexposed          144      22  166
## Total              186     103  289
## Proportion exposed 0.23     0.79  0.43
##
## $df2
##           Point estimate 95%CI-l1 95%CI-ul
## Odds ratio           0.08     0.04     0.15
## Prev. frac. ex.       0.92     0.85     0.96
## Prev. frac. pop       0.72        NA        NA
## chi2(1)              85.22        NA        NA
## Pr>chi2               0.000        NA        NA
## Fisher p.value       0.000        NA        NA
```


Example 2: CC ill - beer (formatted)

The following results tables are outputs in “markdown” using the kable function.

```
result <- CC(DF, "ill", "beer", exact = TRUE, full = TRUE)
kable(result$df1, align="r")
```

	Cases	Controls	Total
Exposed	76	30	106
Unexposed	96	69	165
Total	172	99	271
Proportion exposed	0.44	0.30	0.39

```
kable(result$df2, align=result$df2.align)
```

	Point estimate	95%CI-ll	95%CI-ul
Odds ratio	1.82	1.05	3.20
Attr. frac. ex.	0.45	0.05	0.69
Attr. frac. pop	0.20	NA	NA
chi2(1)	5.09	NA	NA
Pr>chi2	0.024	NA	NA
Fisher p.value	0.028	NA	NA

By storing the results in the object “result”, you are able to use the result tables in Markdown as shown above. By specifying “full = TRUE” you can also easily use individual elements of the results. For example if you would like to view just the odds ratio, you can view it by typing:

```
result$st$odds_ratio$point_estimate
```

```
## [1] 1.820833 1.047409 3.196975
```

CCTable - Summary table for case control studies

CCTable is used for univariate analysis of case control studies with several exposures. The results are summarised in one table with one row per exposure making comparisons between exposures easier and providing a useful table for integrating into reports. Note that all variables need to be numeric and binary and coded as “0” and “1”.

The results of this function contain: The name of exposure variables, the total number of cases, the number of exposed cases, the percentage of exposed among cases, the number of controls, the number of exposed controls, the percentage of exposed among controls, odds ratios, 95%CI intervals, p-values.

You can optionally choose to display the Fisher’s exact p-value instead of the Chi squared p-value, with the option `exact = TRUE`.

You can specify the sort order, with the option `sort=“or”` to order by odds ratios. The default sort order is by p-values.

The option `“full = TRUE”` provides you with useful formatting information, which can be handy if you’re using “markdown”.

Syntax

```
CCTable(x, cases, exposure=c(), exact=FALSE, sort = “pvalue”, full=FALSE)
```

Example 1: CCTable results ordered by p-value (unformatted)

```
CCTable(DF, "ill",
  exposure = c("sex", "agegroup", "tira", "beer", "mousse", "wmousse", "dmousse",
    "redjelly", "fruitsalad", "tomato", "mince", "salmon", "horseradish",
    "chickenwin", "roastbeef", "pork"))
```

```
## $df
##          Tot.Cases Exposed    % Tot.Ctrls Exposed    %  OR CI l1 CI ul p(Chi2)
## tira          185      27 14.59         101      94 93.07 0.01  0.00  0.03  0.000
## mousse        186      42 22.58         103      81 78.64 0.08  0.04  0.15  0.000
## wmousse       179      23 12.85          98      49 50.00 0.15  0.08  0.28  0.000
## dmousse       185      37 20.00         102      76 74.51 0.09  0.05  0.16  0.000
## redjelly      188      34 18.09         103      45 43.69 0.28  0.16  0.51  0.000
## fruitsalad    188      25 13.30         103      46 44.66 0.19  0.10  0.35  0.000
## beer          172      76 44.19          99      30 30.30 1.82  1.05  3.20  0.024
## tomato        188      48 25.53         103      35 33.98 0.67  0.38  1.17  0.127
## pork          187      72 38.50         102      48 47.06 0.70  0.42  1.18  0.158
## horseradish   187      42 22.46         102      30 29.41 0.70  0.39  1.25  0.192
## sex           188      86 45.74         103      53 51.46 0.80  0.48  1.32  0.351
## roastbeef     188      21 11.17         103       8  7.77 1.49  0.61  4.05  0.354
## chickenwin    188      51 27.13         103      33 32.04 0.79  0.45  1.39  0.377
## mince         188      55 29.26         103      32 31.07 0.92  0.53  1.61  0.747
## agegroup      183      43 23.50         100      25 25.00 0.92  0.51  1.70  0.777
## salmon        187      67 35.83         100      37 37.00 0.95  0.56  1.63  0.844
```

Example 2: CCTable results ordered by odds ratio (formatted)

The following results tables are outputs in “markdown” using the kable function.

```
res = CCTable(DF, "ill", sort = "or", exposure = Colnames)
kable(res$df)
```

	Tot.Cases	Exposed	%	Tot.Ctrls	Exposed	%	OR	CI ll	CI ul
beer	172	76	44.19	99	30	30.30	1.82	1.05	3.20
roastbeef	188	21	11.17	103	8	7.77	1.49	0.61	4.05
salmon	187	67	35.83	100	37	37.00	0.95	0.56	1.63
mince	188	55	29.26	103	32	31.07	0.92	0.53	1.61
agegroup	183	43	23.50	100	25	25.00	0.92	0.51	1.70
sex	188	86	45.74	103	53	51.46	0.80	0.48	1.32
chickenwin	188	51	27.13	103	33	32.04	0.79	0.45	1.39
horseradish	187	42	22.46	102	30	29.41	0.70	0.39	1.25
pork	187	72	38.50	102	48	47.06	0.70	0.42	1.18
tomato	188	48	25.53	103	35	33.98	0.67	0.38	1.17
redjelly	188	34	18.09	103	45	43.69	0.28	0.16	0.51
fruitsalad	188	25	13.30	103	46	44.66	0.19	0.10	0.35
wmousse	179	23	12.85	98	49	50.00	0.15	0.08	0.28
dmousse	185	37	20.00	102	76	74.51	0.09	0.05	0.16
mousse	186	42	22.58	103	81	78.64	0.08	0.04	0.15
tira	185	27	14.59	101	94	93.07	0.01	0.00	0.03

Example 3: CCTable results ordered by p-value from the Fisher's exact test (formatted)

The following results tables are outputs in “markdown” using the kable function.

```
res = CCTable(DF, "ill", exposure = Colnames, exact=TRUE)
kable(res$df)
```

	Tot.Cases	Exposed	%	Tot.Ctrls	Exposed	%	OR	CI ll	CI ul
tira	185	27	14.59	101	94	93.07	0.01	0.00	0.03
wmousse	179	23	12.85	98	49	50.00	0.15	0.08	0.28
dmousse	185	37	20.00	102	76	74.51	0.09	0.05	0.16
mousse	186	42	22.58	103	81	78.64	0.08	0.04	0.15
redjelly	188	34	18.09	103	45	43.69	0.28	0.16	0.51
fruitsalad	188	25	13.30	103	46	44.66	0.19	0.10	0.35
beer	172	76	44.19	99	30	30.30	1.82	1.05	3.20
tomato	188	48	25.53	103	35	33.98	0.67	0.38	1.17
pork	187	72	38.50	102	48	47.06	0.70	0.42	1.18
horseradish	187	42	22.46	102	30	29.41	0.70	0.39	1.25
sex	188	86	45.74	103	53	51.46	0.80	0.48	1.32
roastbeef	188	21	11.17	103	8	7.77	1.49	0.61	4.05
chickenwin	188	51	27.13	103	33	32.04	0.79	0.45	1.39
agegroup	183	43	23.50	100	25	25.00	0.92	0.51	1.70
mince	188	55	29.26	103	32	31.07	0.92	0.53	1.61
salmon	187	67	35.83	100	37	37.00	0.95	0.56	1.63

By storing the results in the object “res”, you are able to use the result table in Markdown as shown above. You can also use individual elements of the results. For example if you would like to view just the odds ratio, you can view it by typing (for example):

```
res$df$'Odds Ratio'[1]
```

```
## NULL
```

CCInter - Stratified analysis for case control studies

CCInter is useful to determine the effects of a third variable on the association between an exposure and an outcome. CCInter produces 2 by 2 tables with stratum specific odds ratios, attributable risk among exposed and population attributable risk.

Note that the outcome and exposure variable need to be numeric and binary and coded as "0" and 1". The third variable needs to be numeric, but may have more categories, such as "0", "1" and "2".

CCInter displays a summary with the crude OR, the Mantel Haenszel adjusted OR and the result of a Woolf test for homogeneity of stratum-specific OR.

The option "full = TRUE" provides you with useful formatting information, which can be handy if you're using "markdown".

Syntax

`CCInter(x, cases, exposure, by, full=FALSE)`

Example 1: CCInter ill - wmousse by tira (unformatted)

```
CCInter(DF, cases="ill", exposure = "wmousse", by = "tira")

## $df1
##   CCInter ill - wmousse by(tira) Cases Controls      P.est. Stats 95%CI-l1 95%CI-ul
## 1          tira = 1 <NA>      <NA>      Odds ratio 0.69      0.24      1.92
## 2          Exposed   9         43 Prev. frac. ex. 0.31      -0.92     0.76
## 3          Unexposed 14         46 Prev. frac. pop 0.15         NA        NA
## 4          Total    23         89          <NA>      NA        NA        NA
## 5          Exposed % 39.1%    48.3%    <NA>      NA        NA        NA
## 6          ----- <NA>      <NA>      <NA>      NA        NA        NA
## 7          tira = 0 <NA>      <NA>      Odds ratio 0.07      0.01      0.47
## 8          Exposed   13         4 Prev. frac. ex. 0.93      0.53     0.99
## 9          Unexposed 141         3 Prev. frac. pop 0.53         NA        NA
## 10         Total    154         7          <NA>      NA        NA        NA
## 11         Exposed % 8.4%     57.1%    <NA>      NA        NA        NA
## 12         ----- <NA>      <NA>      <NA>      NA        NA        NA
## 13         Number of obs 273      <NA>      <NA>      NA        NA        NA
## 14         Missing    18         <NA>      <NA>      NA        NA        NA
##
## $df2
##           P.estimate  Stats 95%CI-l1 95%CI-ul
## 1      MH test of Homogeneity  0.01      NA        NA
## 2      Crude OR for wmousse    0.15     0.08     0.28
## 3 MH OR wmousse adjusted for tira 0.44     0.20     0.99
## 4 Adjusted/crude relative change 199.81     NA        NA
```

Example 2: CCInter ill - beer by tira (formatted)

The following results tables are outputs in "markdown" using the kable function.

```
res <- CCInter(DF, cases="ill", exposure = "beer", by = "tira", full = TRUE)
kable(res$df1, align=res$df1.align)
```

CCInter ill - beer by(tira)	Cases	Controls	P.est.	Stats	95%CI-l1	95%CI-ul
tira = 1			Odds ratio	2.72		1.01
Exposed	14	27	Attrib.risk.exp	0.63		0.01
Unexposed	12	63	Attrib.risk.pop	0.34		NA
Total	26	90		NA		NA
Exposed %	53.8%	30.0%		NA		NA
				NA		NA

CCInter ill - beer by(tira)	Cases	Controls	P.est.	Stats	95%CI-ll
tira = 0			Odds ratio	0.96	0.16
Exposed	60	3	Prev. frac. ex.	0.04	-5.82
Unexposed	83	4	Prev. frac. pop	0.02	NA
Total	143	7		NA	NA
Exposed %	42.0%	42.9%		NA	NA
				NA	NA
Number of obs	266			NA	NA
Missing	25			NA	NA

`kable(res$df2)`

P.estimate	Stats	95%CI-ll	95%CI-ul
MH test of Homogeneity	0.22	NA	NA
Crude OR for beer	1.74	1.00	3.07
MH OR beer adjusted for tira	2.07	0.96	4.47
Adjusted/crude relative change	18.80	NA	NA

Example 3: CCIInter ill - beer by tportion (formatted)

The following results tables are outputs in “markdown” using the kable function.

```
res <- CCIInter(DF, cases="ill", exposure = "beer", by = "tportion", full = TRUE)
kable(res$df1, align=res$df1.align)
```

CCIInter ill - beer by(tportion)	Cases	Controls	P.est.	Stats	95%CI-ll
tportion = 2					
Exposed	2	17	Odds ratio	0.88	0.07
Unexposed	4	30	Prev. frac. ex.	0.12	-5.94
Total	6	47	Prev. frac. pop	0.04	NA
Exposed %	33.3%	36.2%		NA	NA

tportion = 1					
Exposed	12	10	Odds ratio	4.95	1.38
Unexposed	8	33	Attrib.risk.exp	0.79	0.27
Total	20	43	Attrib.risk.pop	0.48	NA
Exposed %	60.0%	23.3%		NA	NA

tportion = 0					
Exposed	60	3	Odds ratio	0.96	0.16
Unexposed	83	4	Prev. frac. ex.	0.04	-5.82
Total	143	7	Prev. frac. pop	0.02	NA
Exposed %	42.0%	42.9%		NA	NA

Number of obs	266			NA	NA
Missing	25			NA	NA

```
kable(res$df2, align=res$df2.align)
```

P.estimate	Stats	95%CI-ll	95%CI-ul
MH test of Homogeneity	0.13	NA	NA
Crude OR for beer	1.74	1.00	3.07
MH OR beer adjusted for tportion	2.14	0.98	4.69
Adjusted/crude relative change	23.04	NA	NA

By storing the results in the object “res”, you are able to use the result table in Markdown as shown above. You can also use individual elements of the results. For example if you would like to view just the Mantel-Haenszel odds ratio for beer adjusted for tportion, you can view it by typing:

```
res$df2$Stats[3]
```

```
## [1] 2.14
## Levels: 0.13 1.74 2.14 23.04
```

```
)
```