

1 群の比率の推測 — 二項検定

二項検定

```
binom.test(テーブル名, p=設定値)
```

近似検定

```
prop.test(テーブル名, p=設定値)
```

```
>
> setwd("i:¥¥Rdocuments¥¥scripts¥¥")
> d1 <- read.table("1群の比率データ.csv", header=TRUE, sep=",")
> head(d1)
  番号  合否
1     1     1
2     2     1
3     3     0
4     4     1
5     5     1
6     6     1
>

> # 度数分布表
> d1$gouhi <- factor(d1$合否, levels=c(1,0), labels=c("success","failure"))
> (t1 <- table(d1$gouhi))

success failure
      19       6
>
>
> # 設定値
> p0 <- 0.6
```

1	番号	合否
2	1	1
3	2	1
4	3	0
5	4	1
6	5	1
7	6	1
8	7	1
9	8	1
10	9	1
11	10	1
12	11	0
13	12	1
14	13	1
15	14	0
16	15	0
17	16	1
18	17	1
19	18	1
20	19	0
21	20	0

```
> # 二項検定
> binom.test(t1, p=p0)

Exact binomial test

data:  t1
number of successes = 19, number of trials = 25, p-value = 0.1514
alternative hypothesis: true probability of success is not equal to 0.6
95 percent confidence interval:
 0.5487120 0.9064356
sample estimates:
probability of success
                0.76
```

```
> # 近似検定
> prop.test(t1, p=p0)

1-sample proportions test with continuity correction
```

```
data:  t1, null probability p0
X-squared = 2.0417, df = 1, p-value = 0.153
alternative hypothesis: true p is not equal to 0.6
95 percent confidence interval:
 0.5447916 0.8984194
sample estimates:
      p
0.76
```

```
>
>
>
```

対応のある2群の比率の比較 — マクネマーの検定

```
library(descr)
CrossTable(変数名 1, 変数名 2, mcnemar=T)
```

または

```
mcnemar.test(変数名 1, 変数名 2)
mcnemar.test(テーブル名)
```

対応のある比率の差の信頼区間を推定する関数
dPP2(x=変数名 1, y=変数名 2)

conf=で信頼係数を設定できる．デフォルトは0.95.

```
#-----
dPP2 <- function(x, y, conf=0.95) {
  t1 <- table(x, y)
  a <- t1[1,1]
  b <- t1[1,2]
  c <- t1[2,1]
  d <- t1[2,2]
  n <- sum(t1)
  se <- sqrt(b+c-(b-c)^2/n)/n
  p1 <- (a+b)/n
  p2 <- (a+c)/n
  z0 <- qnorm((1-(1-conf)/2))
  ci.l <- p1-p2-z0*se
  ci.u <- p1-p2+z0*se
  title <- "Difference of paired proportions"
  stnames <- c("p1", "p2", "difference", paste("lower(", conf, ")", sep=""), paste("upper(", conf, ")", sep=""))
  statistic <- round(data.frame(p1, p2, (p1-p2), ci.l, ci.u), 4)
  colnames(statistic) <- stnames
  out <- list(title, statistic=statistic)
  return(out)
}
#-----
```

```
> setwd("i:¥¥Rdocuments¥¥scripts¥¥")
> d1 <- read.table("対応のある2群の比率データ.csv", header=TRUE, sep=",")
> head(d1)
  番号 子ども 親
1    1      1  1
2    2      1  1
3    3      1  1
4    4      1  1
5    5      1  1
6    6      1  1
>
>
> # 各群の比率
> colMeans(d1[c("子ども", "親")])
  子ども      親
0.6133333 0.4400000
>
>
```

```
> #クロス表と検定
> library(descr)
> (t1 <- CrossTable(d1$子ども, d1$親, mcnemar=T))
Cell Contents
```

	N
Chi-square contribution	
N / Row Total	
N / Col Total	
N / Table Total	

```
=====
```

	d1\$親		
d1\$子ども	0	1	Total
0	21	8	29
	1.395	1.776	
	0.724	0.276	0.387
	0.500	0.242	
	0.280	0.107	
1	21	25	46
	0.880	1.119	
	0.457	0.543	0.613
	0.500	0.758	
	0.280	0.333	
Total	42	33	75
	0.560	0.440	

```
=====
```

1	番号	子ども	親
2	1	1	1
3	2	1	1
4	3	1	1
5	4	1	1
6	5	1	1
7	6	1	1
8	7	1	1
9	8	1	1
10	9	1	1
11	10	1	1
12	11	1	1
13	12	1	1
14	13	1	1
15	14	1	1
16	15	1	1
17	16	1	1
18	17	1	1
19	18	1	1
20	19	1	1
21	20	1	1

McNemar's Chi-squared test

```
Chi^2 = 5.827586      d.f. = 1      p = 0.01577676
```

McNemar's Chi-squared test with continuity correction

```
Chi^2 = 4.965517      d.f. = 1      p = 0.02585758
```

```
>
>
```

```
> #マクネマーの検定
```

```
> mcnemar.test(d1$子ども, d1$親)
```

McNemar's Chi-squared test with continuity correction

data: d1\$子ども and d1\$親

McNemar's chi-squared = 4.9655, df = 1, p-value = 0.02586

```
>
```

```
> #対応のある比率の差の信頼区間
```

```
> dPP2(x=d1$子ども, y=d1$親)
```

```
[[1]]
```

```
[1] "Difference of paired proportions"
```

```
$statistic
```

```
      p1      p2 difference lower(0.95) upper(0.95)
1 0.1193 0.2927   -0.1733   -0.3085   -0.0382
```

対応のある多群の比率の比較 — コクランのQ検定

```
cochranq.test(データ名)
```

あらかじめCVSTパッケージをインストールしておく必要がある。
データは、分析したい変数だけ残して、不要な変数は削除しておく。

同時信頼区間

```
dPPs(data=データ名)
```

```
#-----
dPPs <- function(data, conf=0.95) {
  nc <- ncol(data)
  npair <- nc*(nc-1)
  pairs <- matrix(c(0), npair, 5)
  k <- 0
  for(i in 1:nc) {
    for(j in c(1:nc)[-i]) {
      d2 <- data[,c(i,j)]
      t1 <- table(d2)
      a <- t1[1,1]
      b <- t1[1,2]
      c <- t1[2,1]
      d <- t1[2,2]
      n <- sum(t1)
      se <- sqrt(b+c-(b-c)^2/n)/n
      p1 <- (a+b)/n
      p2 <- (a+c)/n
      z0 <- qnorm((1-((1-conf)/(npair/2))/2))
      ci.l <- p1-p2-z0*se
      ci.u <- p1-p2+z0*se
      k <- k+1
      pairs[k,] <- round(c(p1, p2, (p1-p2), ci.l, ci.u), 4)
    }
  }
  rownames(pairs) <- c(1:npair)
  jconf <- round(100*(1-((1-conf)/(npair/2))), 1)
  colnames(pairs) <- c("p1", "p2", "diff", paste("lower(", jconf, ")", sep=""), paste("upper(", jconf, ")", sep=""))
  title <- paste("Joint ", 100*conf, "% Confidence Intervals", sep="")
  out <- list(title, CI= pairs)
  return(out)
}
#-----
```

```
> setwd("i:¥¥Rdocuments¥¥scripts¥¥")
> d1 <- read.table("対応のある多群の比率データ.csv", header=TRUE, sep=",")
> head(d1)
  番号 福祉系 教育系 幼保系
1     1      0      0      0
2     2      1      0      0
3     3      1      1      1
4     4      1      1      0
5     5      0      0      1
6     6      1      0      0
>
>
> # 各群の比率
> colMeans(d1[,c("福祉系", "教育系", "幼保系")])
  福祉系 教育系 幼保系
0.7232143 0.4196429 0.6071429
>
>
> # cochran.Qを使う方法
> library(CVST)
> cochranq.test(d1[,c("福祉系", "教育系", "幼保系")])
```

Cochran's Q Test

```
data: d1[, c("福祉系", "教育系", "幼保系")]
Cochran's Q = 22.075, df = 2, p-value = 1.609e-05
```

>

>

>

> # 対応のある比率の差の信頼区間

> dPPs(data=d1[, c("福祉系", "教育系", "幼保系")])

[[1]]

[1] "Joint 95% Confidence Intervals"

\$CI

	p1	p2	diff	lower(98.3)	upper(98.3)
1	0.2768	0.5804	-0.3036	-0.4572	-0.1499
2	0.2768	0.3929	-0.1161	-0.2469	0.0148
3	0.5804	0.2768	0.3036	0.1499	0.4572
4	0.5804	0.3929	0.1875	0.0289	0.3461
5	0.3929	0.2768	0.1161	-0.0148	0.2469
6	0.3929	0.5804	-0.1875	-0.3461	-0.0289

>

>

1	番号	福祉系	教育系	幼保系
2	1	0	0	0
3	2	1	0	0
4	3	1	1	1
5	4	1	1	0
6	5	0	0	1
7	6	1	0	0
8	7	1	0	1
9	8	1	1	1
10	9	0	1	1
11	10	0	1	0
12	11	1	1	1
13	12	0	1	1
14	13	1	1	1
15	14	1	1	1
16	15	1	1	1
17	16	0	0	0
18	17	1	0	1
19	18	1	0	1
20	19	0	0	0
21	20	1	1	1

リスク差・リスク比・オッズ比の推測

リスク差

prop.test(テーブル名)

リスク比・オッズ比

```
library(epitools)
epitab(テーブル名, method=c("riskratio"))
epitab(テーブル名, method=c("oddsratio"))
```

データ値を、「0=なし, 1=あり」(「なし」のほうが小さい値)としておかなければならない
 CrossTable関数による見やすい表と, table関数で作成して, リスク比・オッズ比を計算する元になる表で,
 行および列の1, 0の並び順が逆になることに注意する.

あらかじめepitoolsパッケージをインストールしておく必要がある.
 method で "oddsratio" か "riskratio" を指定する. "rateratio" というオプションもある.
 vcdパッケージにもoddsratioという関数がある. オッズ比の計算をしてくれる.

```
> setwd("i:¥¥Rdocuments¥¥scripts¥¥")
>
```

> # 前向き研究

```
> d1 <- read.table("対応のない2群の比率データ前.csv", header=TRUE, sep=",")
> head(d1)
  番号 身内患者 医療関係
1    1         0         0
2    2         0         0
3    3         1         1
4    4         1         0
5    5         1         1
6    6         0         1
>
```

> # クロス表

```
> d1$身内患者rev <- factor(d1$身内患者, levels=c(1,0))
> d1$医療関係rev <- factor(d1$医療関係, levels=c(1,0))
> library(descr)
> CrossTable(d1$身内患者rev, d1$医療関係rev, prop.chisq=F, prop.c=F, prop.t=F)
Cell Contents
```

	N	
	N / Row Total	

```
=====
d1$身内患者rev  d1$医療関係rev
                1         0      Total
1                30        45        75
                0.400    0.600    0.500
-----
0                 15         60        75
                0.200    0.800    0.500
-----
Total              45       105       150
=====
>
```

1	番号	身内患者	医療関係
2	1	0	0
3	2	0	0
4	3	1	1
5	4	1	0
6	5	1	1
7	6	0	1
8	7	0	0
9	8	1	0
10	9	1	0
11	10	1	0
12	11	0	0
13	12	0	0
14	13	0	0
15	14	0	0
16	15	0	0
17	16	1	0
18	17	0	0
19	18	1	0
20	19	0	0
21	20	1	1

```

> # リスク差
> (t1 <- table(d1$身内患者, d1$医療関係))

      0  1
0 60 15
1 45 30

> prop.test(t1)

      2-sample test for equality of proportions with continuity
      correction

data:  t1
X-squared = 6.2222, df = 1, p-value = 0.01262
alternative hypothesis: two.sided
95 percent confidence interval:
 0.04353114 0.35646886
sample estimates:
prop 1 prop 2
 0.8    0.6

>

> # リスク比
> library(epitools)
> epitab(t1, method=c("riskratio"))
$tab

      0 p0  1  p1 riskratio      lower      upper    p.value
0 60 0.8 15 0.2          1         NA         NA         NA
1 45 0.6 30 0.4          2 1.176315 3.400451 0.01215822

$measure
[1] "wald"

$conf.level
[1] 0.95

$pvalue
[1] "fisher.exact"

>

> # オッズ比
> library(epitools)
> epitab(t1, method=c("oddsratio"))
$tab

      0      p0  1      p1 oddsratio      lower      upper    p.value
0 60 0.5714286 15 0.3333333 1.000000         NA         NA         NA
1 45 0.4285714 30 0.6666667 2.666667 1.284531 5.535958 0.01215822

$measure
[1] "wald"

$conf.level
[1] 0.95

$pvalue
[1] "fisher.exact"

```

```
>
> # 後ろ向き研究

> d1 <- read.table("対応のない2群の比率データ後.csv", header=TRUE, sep=",")
> head(d1)
  番号 医療関係 身内患者
1    1         0         0
2    2         0         0
3    3         1         0
4    4         1         0
5    5         0         0
6    6         0         0
>
```

```
> # クロス表
> d1$身内患者rev <- factor(d1$身内患者, levels=c(1,0))
> d1$医療関係rev <- factor(d1$医療関係, levels=c(1,0))
> library(descr)
> CrossTable(d1$身内患者rev, d1$医療関係rev, prop.chisq=F, prop.r=F, prop.t=F)
```

```
Cell Contents
-----
N
N / Col Total
-----

=====
d1$身内患者rev    d1$医療関係rev
                   1         0    Total
-----
1                   25         12       37
                   0.333    0.158
-----
0                   50         64      114
                   0.667    0.842
-----
Total              75         76      151
                   0.497    0.503
=====
>
```

1	番号	医療関係	身内患者
2	1	0	0
3	2	0	0
4	3	1	0
5	4	1	0
6	5	0	0
7	6	0	0
8	7	1	0
9	8	0	0
10	9	0	0
11	10	1	1
12	11	1	0
13	12	1	0
14	13	0	0
15	14	0	0
16	15	0	1
17	16	0	1
18	17	0	0
19	18	1	0
20	19	0	0
21	20	0	0

```
> # オッズ比
> library(epitools)
> epitab(t1, method=c("oddsratio"))
$tab

      0      p0  1      p1 oddsratio   lower   upper   p.value
0 60 0.5714286 15 0.3333333  1.000000      NA      NA      NA
1 45 0.4285714 30 0.6666667  2.666667 1.284531 5.535958 0.01215822

$measure
[1] "wald"

$conf.level
[1] 0.95

$pvalue
[1] "fisher.exact"
```

```
>
>
```

```
>
>
```

```
> # 横断研究
> d1 <- read.table("対応のない2群の比率データ横.csv", header=TRUE, sep=",")
> head(d1)
  番号 身内患者 医療関係
1     1         0         0
2     2         1         0
3     3         0         1
4     4         0         0
5     5         0         0
6     6         1         0
>
```

1	番号	身内患者	医療関係
2	1	0	0
3	2	1	0
4	3	0	1
5	4	0	0
6	5	0	0
7	6	1	0
8	7	0	0
9	8	0	0
10	9	0	0
11	10	0	0
12	11	0	0
13	12	0	1
14	13	0	0
15	14	0	0
16	15	0	1
17	16	0	0
18	17	1	0
19	18	0	1
20	19	0	0
21	20	0	0

```
> # クロス表
> d1$身内患者rev <- factor(d1$身内患者, levels=c(1,0))
> d1$医療関係rev <- factor(d1$医療関係, levels=c(1,0))
> library(descr)
> CrossTable(d1$身内患者rev, d1$医療関係rev, prop.chisq=F)
```

Cell Contents

	N
N / Row Total	
N / Col Total	
N / Table Total	

```
=====
d1$身内患者rev      d1$医療関係rev
                    1         0      Total
-----
1                   18        12        30
    0.600    0.400    0.200
    0.375    0.118
    0.120    0.080
-----
0                   30        90       120
    0.250    0.750    0.800
    0.625    0.882
    0.200    0.600
-----
Total                48       102       150
    0.320    0.680
=====
>
```

```
> # リスク差
> (t1 <- table(d1$身内患者, d1$医療関係))
```

```
      0  1
0 90 30
1 12 18
> prop.test(t1)
```

2-sample test for equality of proportions with continuity correction

```
data:  t1
X-squared = 11.9504, df = 1, p-value = 0.0005463
alternative hypothesis: two.sided
95 percent confidence interval:
 0.1375056 0.5624944
sample estimates:
prop 1 prop 2
 0.75   0.40
```

```
>
```

```
> # リスク比
> library(epitools)
> epitab(t1, method=c("riskratio"))
$tab
```

	0	p0	1	p1	riskratio	lower	upper	p.value
0	90	0.75	30	0.25	1.0	NA	NA	NA
1	12	0.40	18	0.60	2.4	1.567615	3.674372	0.0004215309

```
$measure
[1] "wald"
```

```
$conf.level
[1] 0.95
```

```
$pvalue
[1] "fisher.exact"
```

```
>
> # オッズ比
> library(epitools)
> epitab(t1, method=c("oddsratio"))
$tab
```

	0	p0	1	p1	oddsratio	lower	upper	p.value
0	90	0.8823529	30	0.625	1.0	NA	NA	NA
1	12	0.1176471	18	0.375	4.5	1.94424	10.41538	0.0004215309

```
$measure
[1] "wald"
```

```
$conf.level
[1] 0.95
```

```
$pvalue
[1] "fisher.exact"
```

対応のない多群の比率の等質性の推測

prop.test(テーブル名)

同時信頼区間

dNPs(x=群分け変数名, y=0/1変数名)

```
#-----
dNPs <- function(x, y, conf=0.95) {
  t1 <- table(x, y)
  nc <- length(levels(x))
  npair <- nc*(nc-1)
  pairs <- matrix(c(0), npair, 5)
  k <- 0
  for(i in 1:nc) {
    for(j in c(1:nc)[-i]) {
      t12 <- t1[c(i, j), ]
      tt <- prop.test(t12, conf.level=(1-((1-conf)/(npair/2))))
      p1 <- tt$estimate[1]
      p2 <- tt$estimate[2]
      ci.l <- tt$conf.int[1]
      ci.u <- tt$conf.int[2]
      k <- k+1
      pairs[k,] <- round(c(p1, p2, (p1-p2), ci.l, ci.u), 4)
    }
  }
  rownames(pairs) <- c(1:npair)
  jconf <- round(100*(1-((1-conf)/(npair/2))), 1)
  colnames(pairs) <- c("p1", "p2", "diff", paste("lower(", jconf, ")", sep=""), paste("upper(", jconf, ")", sep=""))
  title <- paste("Joint ", 100*conf, "% Confidence Intervals", sep="")
  out <- list(title, CI= pairs)
  return(out)
}
#-----
```

```
> setwd("i:¥¥Rdocuments¥¥scripts¥¥")
> d1 <- read.table("対応のない多群の比率データ.csv", header=TRUE, sep=",")
```

```
> head(d1)
  番号 地域 罹患
```

```
1    1    A    1
2    2    C    0
3    3    B    0
4    4    B    0
5    5    C    0
6    6    A    1
```

>

>

> # クロス表

> d1\$地域 <- factor(d1\$地域, levels=c("A", "B", "C"))

> d1\$罹患 <- factor(d1\$罹患, levels=c(1, 0))

> library(descr)

> CrossTable(d1\$地域, d1\$罹患, prop.chisq=F, prop.c=F, prop.t=F)

Cell Contents

			N
		N / Row Total	

```
=====
```

	d1\$罹患		
d1\$地域	1	0	Total
A	12	22	34
	0.353	0.647	0.239
B	8	38	46
	0.174	0.826	0.324

番号	地域	罹患
1	A	1
2	C	0
3	B	0
4	B	0
5	C	0
6	A	1
7	A	1
8	C	1
9	B	0
10	B	0
11	C	0
12	B	0
13	C	0
14	C	0
15	A	0
16	A	0
17	C	1
18	A	0
19	C	1
20	B	0
21	C	0

C	26	36	62
	0.419	0.581	0.437
Total	46	96	142

```
>
>
```

```
> # 対応のない多群の比率
```

```
> (t1 <- table(d1$地域, d1$罹患))
```

```

  1  0
A 12 22
B  8 38
C 26 36
```

```
> prop.test(t1)
```

```

3-sample test for equality of proportions without continuity
correction
```

```

data:  t1
X-squared = 7.4355, df = 2, p-value = 0.02429
alternative hypothesis: two.sided
sample estimates:
  prop 1    prop 2    prop 3
0.3529412 0.1739130 0.4193548
```

```
>
>
```

```
>
> # 対応のない比率の差の信頼区間
> dNPs(x=d1$地域, y=d1$罹患)
```

```
[[1]]
[1] "Joint 95% Confidence Intervals"
```

```
$CI
      p1      p2      diff lower(98.3) upper(98.3)
1 0.3529 0.1739  0.1790    -0.0840     0.4421
2 0.3529 0.4194 -0.0664    -0.3362     0.2033
3 0.1739 0.3529 -0.1790    -0.4421     0.0840
4 0.1739 0.4194 -0.2454    -0.4654    -0.0255
5 0.4194 0.3529  0.0664    -0.2033     0.3362
6 0.4194 0.1739  0.2454     0.0255     0.4654
```

```
>
>
```