

note13

Matts966

2018/01/30

```
# matrix は縦に数字を振っていく関数 (byrow=Tなどで変更)
```

```
x <- matrix(c(25, 10, 15, 30), ncol=2)
```

```
x
```

```
##      [,1] [,2]
```

```
## [1,]   25   15
```

```
## [2,]   10   30
```

```
chisq.test(x, correct=F)
```

```
##
```

```
## Pearson's Chi-squared test
```

```
##
```

```
## data:  x
```

```
## X-squared = 11.429, df = 1, p-value = 0.0007232
```

```
d1 <- matrix(c(12, 7, 2, 9), nrow=2, byrow=T)
```

```
fisher.test(d1, alternative = "g")
```

```
##
```

```
## Fisher's Exact Test for Count Data
```

```
##
```

```
## data:  d1
```

```
## p-value = 0.02119
```

```
## alternative hypothesis: true odds ratio is greater than 1
```

```
## 95 percent confidence interval:
```

```
##  1.331695      Inf
```

```
## sample estimates:
```

```
## odds ratio
```

```
##   7.166131
```

```
hair <- matrix(c(32, 43, 16, 9, 55, 65, 64, 16), nrow=2, byrow=T)
```

```
fisher.test(hair)
```

```
##
```

```
## Fisher's Exact Test for Count Data
```

```
##
```

```
## data:  hair
```

```
## p-value = 0.0241
```

```
## alternative hypothesis: two.sided
```

```
#2000 回試行のモンテカルロシミュレーションも実行可能
```

```
hair <- matrix(c(32, 43, 16, 9, 55, 65, 64, 16), nrow=2, byrow=T)
fisher.test(hair, simulate.p.value = T)
```

```
##
## Fisher's Exact Test for Count Data with simulated p-value (based
## on 2000 replicates)
##
## data: hair
## p-value = 0.02299
## alternative hypothesis: two.sided
```

【考察】モンテカルロシミュレーションのような、膨大な量の計算を実行して、表を埋めていき、正確な値を推定するような手法は、確かにコンピュータの速度、メモリ容量、仮想記憶の性能向上などがなければ実行できていないと考えられる。

```
# 演習問題 12.1
```

```
x2 = (49 / 31) + (49 / 31) + (49 / 29) + (49 / 29)
x2
```

```
## [1] 6.540601
```

```
pchisq(x2, df=1)
```

```
## [1] 0.989456
```

```
x <- matrix(c(38, 22, 24, 36), ncol = 2, nrow = 2, byrow = T)
chisq.test(x, correct = F)
```

```
##
## Pearson's Chi-squared test
##
## data: x
## X-squared = 6.5406, df = 1, p-value = 0.01054
```

```
# 合っている
```

```
print('-----
```

```
## [1] "-----
"
```

```
# モンテカルロ実行
```

```
e <- matrix(c(31, 31, 29, 29), ncol = 2, nrow = 2, byrow = T)
chisq.test(x, p=e, rescale = T, simulate.p.value = T, B = 10000)
```

```
##
## Pearson's Chi-squared test with simulated p-value (based on 10000
## replicates)
```

```
##  
## data:  x  
## X-squared = 6.5406, df = NA, p-value = 0.0182
```

【考察】さすがに、毎回実際に一万回計算して結果を算出しているので、毎回異なる数値が出力される。概して、カイ 2 乗検定よりも p-value は大きくなってしまっている。