

# Assignment (statistics): Fisher's exact test and Pearson's chi square test

## Instructions

Answer the following questions, which all have equal mark value. Show your working by including output from your R session.

### Question 1

Consider the following dataset drawn from AUT student services:

```
M <- matrix(c(10,2,11,7),2,2)
dimnames(M) <- list(OS=c("windows","mac"),major=c("science","arts"))
M
```

```
##           major
## OS      science arts
## windows      10   11
##  mac           2    7
```

we suspect arts students are more likely to use a mac than science students.

- State your null clearly
- State the precise definition of p-value
- state what “more extreme” means here
- use `fisher.test()`, calculate your pvalue and interpret

### Question 2

A medical researcher is investigating whether vitamin C helps to cure the common cold. He takes a sample of 20 subjects, of whom 7 are given vitamin C and 13 a placebo (a “placebo” is a dummy pill that has no medical effect). The researcher interviews the subjects after one week, and 9 subjects report that their cold is cured. Of these 9 subjects who are cured, 5 were given real vitamin C. We may represent the data in R as follows:

```
M <- matrix(c(5,4,2,9),2,2)
dimnames(M) <- list(
  given=c("vitamin C","placebo"),state=c("cured","cold")
)
M
```

```
##           state
## given      cured cold
## vitamin C      5    2
## placebo        4    9
```

- State a sensible null hypothesis
- State the precise definition of p-value and explain what “more extreme” means in this context
- Is a one-sided or two-sided test needed? justify

- Estimate the probability of: (a), a person who was given vitamin C, and (b), a person given the placebo, of being cured.
- Perform a Fisher test using `fisher.test()` and interpret

## Question 3

Consider a situation with 3 white and 5 black balls in a bag. Four balls are drawn from the bag, without replacement. Write down every possible sample and calculate its probability.

## Question 4

Three AUT students and four UoA students are given a problem in statistics. All three of the AUT students answer the problem correctly, and none of the UoA students answer correctly. Discuss.

```
fisher.test(diag(3:4)) # two sided? . . . . .

##
## Fisher's Exact Test for Count Data
##
## data:  diag(3:4)
## p-value = 0.02857
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.9258483      Inf
## sample estimates:
## odds ratio
##      Inf
# strong evidence
```

## Question 5

Give an example of Fisher's exact test in your daily life. Give a 2x2 contingency table, with labelled rows and columns. State your null *clearly*, and your alternative. State and justify your use of a one-sided or two-sided test. Carry out your test, report the p-value, and interpret. Excellence question: find the "most extreme" observation that is consistent with your marginal totals.