The University of Sydney
School of Mathematics and Statistics

STUVAC Statistical Snacks 4

MATH1062/MATH1005: Mathematics 1B/Statistical Thinking With Data Semester 1, 2024

Lecturers: J. Baine, T. Cui, J. Spreer, M. Stewart

The author of White (2005) suspected that ravens (intelligent birds whose diet includes carrion) are attracted by the sound of gunshots, presumably in the hope of finding a dead animal to feed on. They conducted an experiment whereby they went to 12 locations in an established "hunting zone", counted how many ravens were visible in the vicinity 10 minutes before and 10 minutes after firing a rifle. The raw data is made available in the *ecostats* R package (Warton, 2022).

```
require(ecostats)
data(ravens)
ravens

##    Before After delta    site treatment trees
## 1       0     2     2 pilgrim         1     1
## 2       0     1     1 pacific         1     1
## 3       0     4     4 uhl hil         1     1
## 4       0     1     1 wolff r         1     1
## 5       0     0     0 teton p         1     1
## 6       2     5     3 glacier         1     1
## 7       1     0    -1 ant fla         1     0
## 8       0     1     1 btb n           1     1
## 9       0     0     0 btb s           1     1
## 10      3     3     0 hay e           1     0
## 11      5     5     0 hay s           1     0
## 12      0     2     2 grove           1     1
## 13      0     0     0 pilgrim         2     1
## 14      1     0    -1 pacific         2     1
## 15      1     1     0 uhl hil         2     1
## 16      0     0     0 wolff r         2     1
## 17      0     1     1 teton p         2     1
## 18      1     1     0 glacier         2     1
## 19      6     3    -3 ant fla         2     0
## 20      3     2    -1 btb n           2     1
## 21      0     0     0 btb s           2     1
## 22      0     0     0 hay e           2     0
## 23      1     0    -1 hay s           2     0
## 24      0     0     0 grove           2     1
## 25      2     1    -1 pilgrim         3     1
## 26      1     0    -1 pacific         3     1
## 27      0     1     1 uhl hil         3     1
## 28      0     0     0 wolff r         3     1
## 29      3     6     3 teton p         3     1
## 30      1     1     0 glacier         3     1
## 31      1     1     0 ant fla         3     0
## 32      0     1     1 btb n           3     1
```



```
## 33       0      0     0 btb s             3     1
## 34       3      1    -2 hay e             3     0
## 35       3      4     1 hay s             3     0
## 36       1      0    -1 grove             3     1
## 37       0      0     0 pilgrim           4     1
## 38       1      0    -1 pacific           4     1
## 39       1      2     1 uhl hil           4     1
## 40       2      3     1 wolff r           4     1
## 41       4      5     1 teton p           4     1
## 42       1      0    -1 glacier           4     1
## 43       0      0     0 ant fla           4     0
## 44       0      0     0 btb n             4     1
## 45       0      1     1 btb s             4     1
## 46       2      0    -2 hay e             4     0
## 47       2      1    -1 hay s             4     0
## 48       0      0     0 grove             4     1
```

The author of the study also tried using other sounds (use `?ravens` to see all the details). The factor `treatment` indicates which sound: the value `1` corresponds to gunshot. We thus extract the first 12 rows.

```
gunshot = ravens[1:12,]
gunshot

##     Before After delta    site treatment trees
## 1        0     2     2 pilgrim         1     1
## 2        0     1     1 pacific         1     1
## 3        0     4     4 uhl hil         1     1
## 4        0     1     1 wolff r         1     1
## 5        0     0     0 teton p         1     1
## 6        2     5     3 glacier         1     1
## 7        1     0    -1 ant fla         1     0
## 8        0     1     1 btb n           1     1
## 9        0     0     0 btb s           1     1
## 10       3     3     0 hay e           1     0
## 11       5     5     0 hay s           1     0
## 12       0     2     2 grove           1     1
```

The column headed `delta` gives `After-Before` differences.

1. If we were to use a T-test to analyse these observations, what kind would that be, and why?

2. Describe an appropriate statistical model for the data, including assumptions appropriate for the kind of T-test given in the previous part. Do the assumptions seem reasonable? Explain.

3. The T-statistic is the ratio of the mean difference and an estimate of the standard error of the mean difference. Determine the estimated standard error. Round to 3 decimal places if necessary. The R output below may be helpful.

```
apply(gunshot[,1:3], 2, sd)

##   Before    After    delta
## 1.621354 1.858641 1.443376
```

4. Formally state the appropriate null and alternative hypotheses for this analysis.

5. Describe the test statistic and state its distribution when the null hypothesis is true (and the assumptions underlying the procedure are reasonable).

6. Using the R output below, state what kind of value the test statistic should take so that we reject the null hypothesis at

- the 5% level of significance;
- the 1% level of significance.

Note: the R function `outer(u, v, fun)` returns a matrix whose `(i,j)`-th element is `fun(u[i],v[j])`).

```
qt.values=outer(c(.95, .975, .99, .995), c(11, 12, 22, 24), qt)
rownames(qt.values)=c("95%", "97.5%", "99%", "99.5%")
colnames(qt.values)=c("11 df", "12 df", "22 df", "24 df")
qt.values

##          11 df    12 df    22 df    24 df
## 95%   1.795885 1.782288 1.717144 1.710882
## 97.5% 2.200985 2.178813 2.073873 2.063899
## 99%   2.718079 2.680998 2.508325 2.492159
## 99.5% 3.105807 3.054540 2.818756 2.796940
```

7. Determine the value taken by the test statistic. Is the result significant at the 5% level? What about at the 1% level?

8. Consider the simulation code and output given below.

```
dif = gunshot$delta
dif

##  [1]  2  1  4  1  0  3 -1  1  0  0  0  2

stat=t.test(dif)$stat
sim.stat=0
n=length(dif)
for(i in 1:10000){
```

```
    samp = sample(dif, size=n, replace=T)
    sim.stat[i] = t.test(samp, mu=mean(dif))$stat
}
sum(abs(sim.stat)>=abs(stat))

## [1] 316

sum(sim.stat>=stat)

## [1] 52
```

Using the simulation output instead of appealing to Student's theory, is the result significant at the 5% level? What about the 1% level?

Note: the link for the Crow White paper below is a Library "Permalink". For it to work you need to be signed into the Library website.

References

David Warton. *ecostats: Code and Data Accompanying the Eco-Stats Text (Warton 2022)*, 2022. URL https://CRAN.R-project.org/package=ecostats. R package version 1.1.11.

Crow White. Hunters ring dinner bell for ravens: Experimental evidence of a unique foraging strategy. *Ecology (Durham)*, 86(4):1057–1060, 2005. ISSN 0012-9658. URL https://sydney.primo.exlibrisgroup.com/permalink/61USYD_INST/2rsddf/cdi_proquest_miscellaneous_17832044.