

# Introduction to Bayesian Statistics with R

1: Exercises

Jack Kuipers

## Exercise 1.1 - a statistical report

A small clinical trial on asthma patients has been run measuring the lung function of a control group on a placebo and a treatment group on a new drug.

- Read in the trial data (`lung_data.csv`),
- visualize the data for each group,
- test whether there is a difference in function between the two groups.

## Bonus Exercise 1.2 - normality and outliers

**NOTE:** This exercise is an optional bonus for when you have sufficient free time.

The t-test assumes normality and no outliers. To get a feel for how important those assumptions are, we can break them and check with simulated data.

From the example in the slides with a small difference in means between two groups

```
# Generate some Gaussian samples with mean -0.25 and control with mean 0
test_samples <- rnorm(50, mean = -0.25, sd = 1)
control_samples <- rnorm(50, mean = 0, sd = 1)
```

the power is actually quite low:

```
power.t.test(n = 50, delta = -0.25, sd = 1)$power
```

```
## [1] 0.2350874
```

Therefore if we were to repeat the experiment many times, we would expect around 24% to be significant.

Although we can work out the power exactly in this case, we could brute force it and run many such 'experiments' in the computer

```
set.seed(42) # set a seed
n_reps <- 4e3 # how many repetitions
p_vals <- rep(NA, n_reps) # to store the p-values
for (ii in 1:n_reps) {
  test_samples <- rnorm(50, mean = -0.25, sd = 1) # treatment group
  control_samples <- rnorm(50, mean = 0, sd = 1) # control group
  p_vals[ii] <- t.test(test_samples, control_samples)$p.value # t-test
}
mean(p_vals < 0.05) # the power given by the fraction of significant tests
```

```
## [1] 0.2365
```

Now we can use this simulation-based approach to play around with the assumptions:

- What happens to the power if we use a different distribution (with the same mean and sd) instead of a normal?
- What happens if we add an outlier (for example, shift one of the treatment group by a large negative value)?