Introduction to Bayesian Statistics with R

2: Exercises

Jack Kuipers

Exercise 2.1 - confidence intervals

Take a sample from a normal distribution extract its 95% confidence interval. For example with mean of -0.25, sd of 1 and sample size of 50, we can use the `t.test` function:

```r
set.seed(42) # set a seed
test_sample <- rnorm(50, mean = -0.25, sd = 1)
t.test(test_sample)$conf.int
```

```
## [1] -0.61291775  0.04157418
## attr(,"conf.level")
## [1] 0.95
```

In the lecture notes we looked at the following questions, which you might like to check and code for yourself: Is the true mean inside your confidence interval? If you repeat this procedure a large number of times, how often is the true mean in the confidence interval? And is your result in line with questions 4 and 5 of the confidence interval quiz?

Here instead we focus on question 6 of the confidence interval quiz and see how often new sample means are in the original confidence interval above.

For this, sample from the same process as above a large number of times and see how often the sample means lie within the first confidence interval.

Is it 95%? How does this align with statement 6 if we replace "true mean" with "sample mean"?

Bonus Exercise 2.2 - a testing example

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

This example comes from **Eddy (1982)**, and asked of medical doctors to see if they can get the right ballpark probability in the end:

- 1% of women at age forty who participate in routine screening have breast cancer.
- 80% of women with breast cancer will get positive mammographies.
- 9.6% of women without breast cancer will also get positive mammographies.

A woman in this age group had a positive mammography in a routine screening. What is the probability that she actually has breast cancer?