

## Social Robotics

B. Nebel, F. Lindner, T. Engesser,  
B. Kuhnert, L. Wächter  
Winter Semester 2017/18

University of Freiburg  
Department of Computer Science

### Exercise Sheet 5

**Due: December 12, 2017, 10:00**

#### Exercise 5.1 (Statistical distance between t- and normal distribution)

The *statistical distance* is a similarity measure for probability distributions. Intuitively, it can be understood as “the largest possible difference between the probabilities that two probability distributions can assign to the same event”. See the definition on:

[https://en.wikipedia.org/wiki/Total\\_variation\\_distance\\_of\\_probability\\_measures](https://en.wikipedia.org/wiki/Total_variation_distance_of_probability_measures)

- (a) Plot the t-distribution for degrees of freedom of 1, 2, 4, and 8 against the normal distribution in four separate plots. In each plot, shade the area between both distribution functions<sup>1</sup> and calculate the statistical distance between them. *Note that the statistical distance can be calculated very easily from the area between the curves. In R, you can compute integrals using expressions like `integrate(function (x) f(x), -Inf, Inf)`.*
- (b) Plot the statistical distance to the normal distribution for the t-distributions with degrees of freedom from 1 to 50. Calculate the statistical distance for degrees of freedom  $df = 30$ . Discuss the meaning of the value.

#### Exercise 5.2 (T-tests in R)

Run the t-tests for the experiments from exercises 3 to 5 of classwork sheet 5 in R and report the results. In which of the experiments does it make a difference to use the Welch-correction? Discuss. *Note: For the experiment from classwork 5 exercise 5, you are encouraged to try and come up with own sample data that matches the described mean and standard deviation.*

---

<sup>1</sup>For shading between curves, see <http://www.alisonsinclair.ca/2011/03/shading-between-curves-in-r/>