

Testing a change in Diffusion Constant

Janis

Updated: August 23, 2018

Contents

1	Introduction	2
2	Simulations	3
2.1	Examples	3
2.2	Simulation Results	4

1 Introduction

Here I compare two ways to test the null hypothesis that an individual drifts with the same diffusion constant in two conditions ($H_0 : D_1 = D_2$).

1. **Overlap of 95% Confidence Intervals:** In this case we bootstrap on the drift segments in each condition to estimate the 2.5-97.5 percentiles of each diffusion constant. If these confidence intervals do not overlap, there is strong evidence against the null hypothesis.
2. **Permutation Testing:** In this case we will rephrase the null hypothesis as $H_0 : T = D_1 - D_2 = 0$. Permutation testing uses bootstrapping to estimate the sampling distribution of T under the null hypothesis, to which we can compare our observed test statistic, $\hat{t} = \hat{D}_1 - \hat{D}_2$.
 - Procedure as described in Wasserman 10.5, applied to diffusion constants of drifts.
 - We have drift segments from two different viewing conditions with n_1 and n_2 samples respectively.
 - First pool all drift segments from the two sampling groups so that there are $n = n_1 + n_2$ samples.
 - To estimate the distribution of T under the null hypothesis, for b from 1 to B , the number of bootstrap iterations:
 - Randomly take n samples from the pool with replacement.
 - Estimate $D_1^{(b)}$ from the first n_1 samples and $D_2^{(b)}$ from the remaining n_2 samples.
 - Let $T^{(b)} = D_1^{(b)} - D_2^{(b)}$
 - Then, the estimated p-value for the one-sided¹ test that $T = D_1 - D_2 < 0$ is

$$p = \frac{1}{B} \sum_b \mathbb{1}\{T^{(b)} > \hat{t}\}$$

¹This could of course be done with a two-sided test as well

2 Simulations

Simulations were run in Janis's script `permuationtest_simulation.m` (not yet on gitlab).

Brownian motion traces of 300ms were simulated and compared. The number of bootstrap iterations was 1000.

Here are the number of significant differences found after 100 simulations in several conditions:

D_1	D_2	n_1	n_2	Bootstrap 95% CI # no overlap	Permutation Test # $p < 0.05$
10	20	100	100	100	100
10	11	100	100	11	14
10	10	100	100	10	1

The two tests perform similarly when the diffusion constants are different, but the permutation test seems to produce less false alarms when the diffusion constants are the same.

2.1 Examples

Here are the results of individual simulations in a few conditions:

1. $D_1 = 10, D_2 = 20, n_1 = 100, n_2 = 100$
2. $D_1 = 10, D_2 = 11, n_1 = 100, n_2 = 100$
3. $D_1 = 10, D_2 = 20, n_1 = 300, n_2 = 100$
4. $D_1 = 10, D_2 = 11, n_1 = 300, n_2 = 100$
5. $D_1 = 10, D_2 = 10, n_1 = 100, n_2 = 100$
6. $D_1 = 10, D_2 = 10, n_1 = 300, n_2 = 100$

2.2 Simulation Results

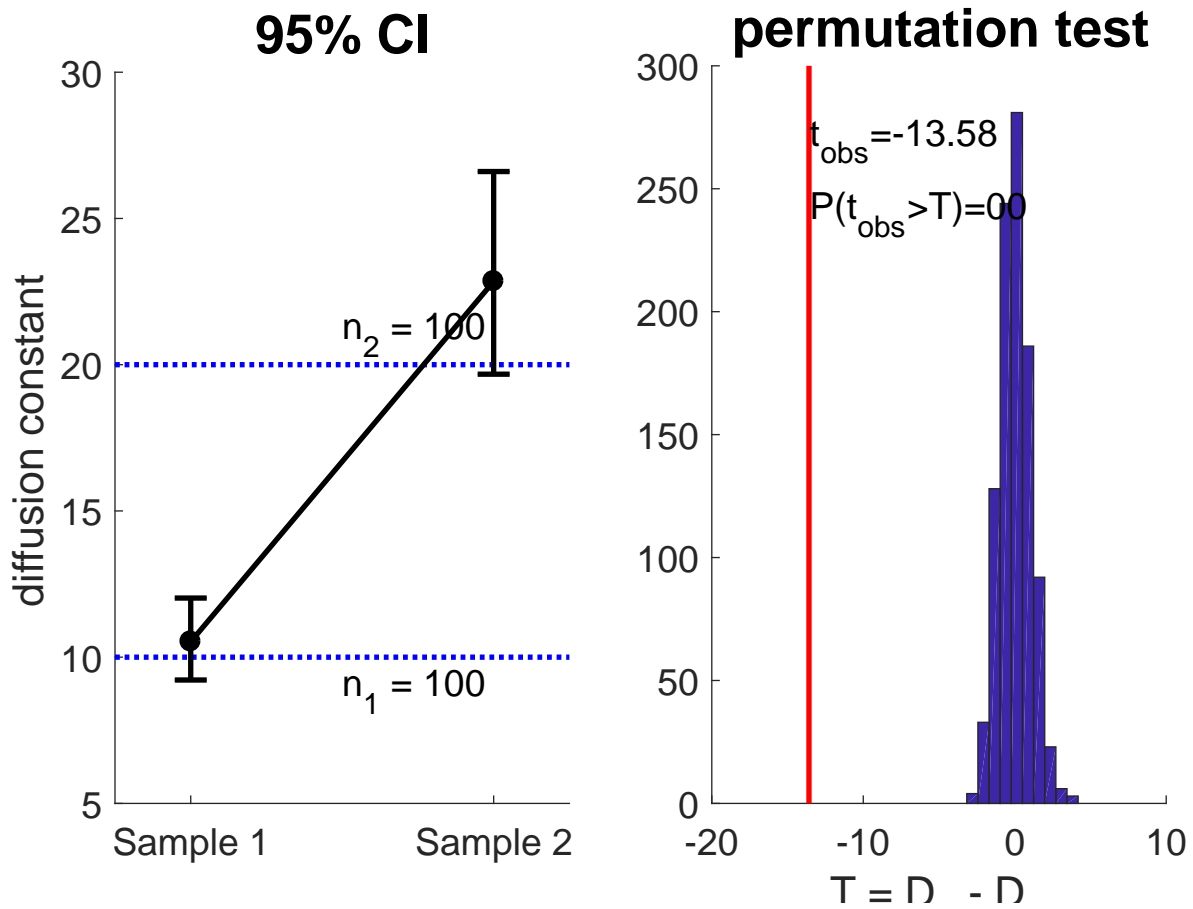


Figure 1: Simulation 1: $D_1 = 10$, $D_2 = 20$, $n_1 = 100$, $n_2 = 100$. The 95% confidence intervals do *not* overlap (left) and the estimated p-value of the hypothesis test is $p = 0$ (right).

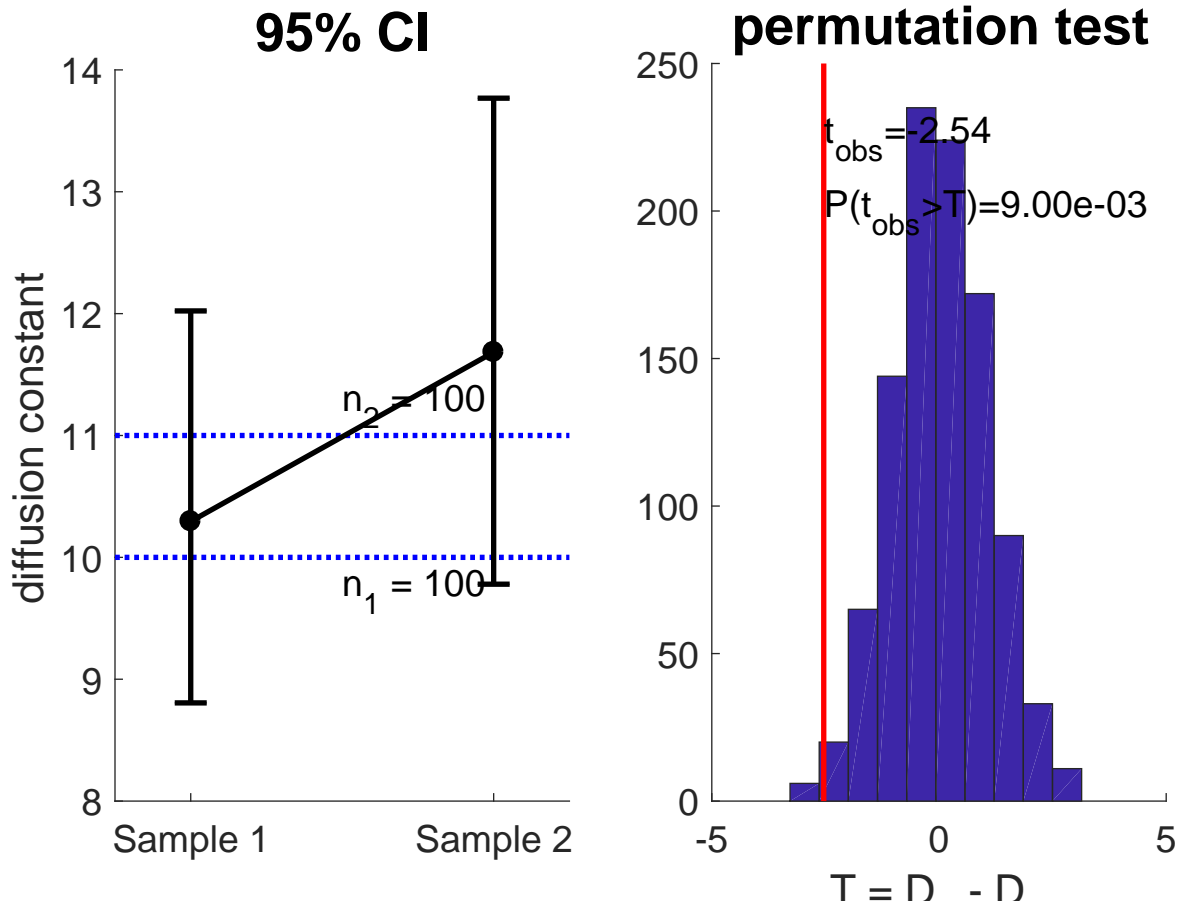


Figure 2: Simulation 2: $D_1 = 10$, $D_2 = 11$, $n_1 = 100$, $n_2 = 100$. The 95% confidence intervals do overlap (left) and the estimated p-value of the hypothesis test is $p = .009$.

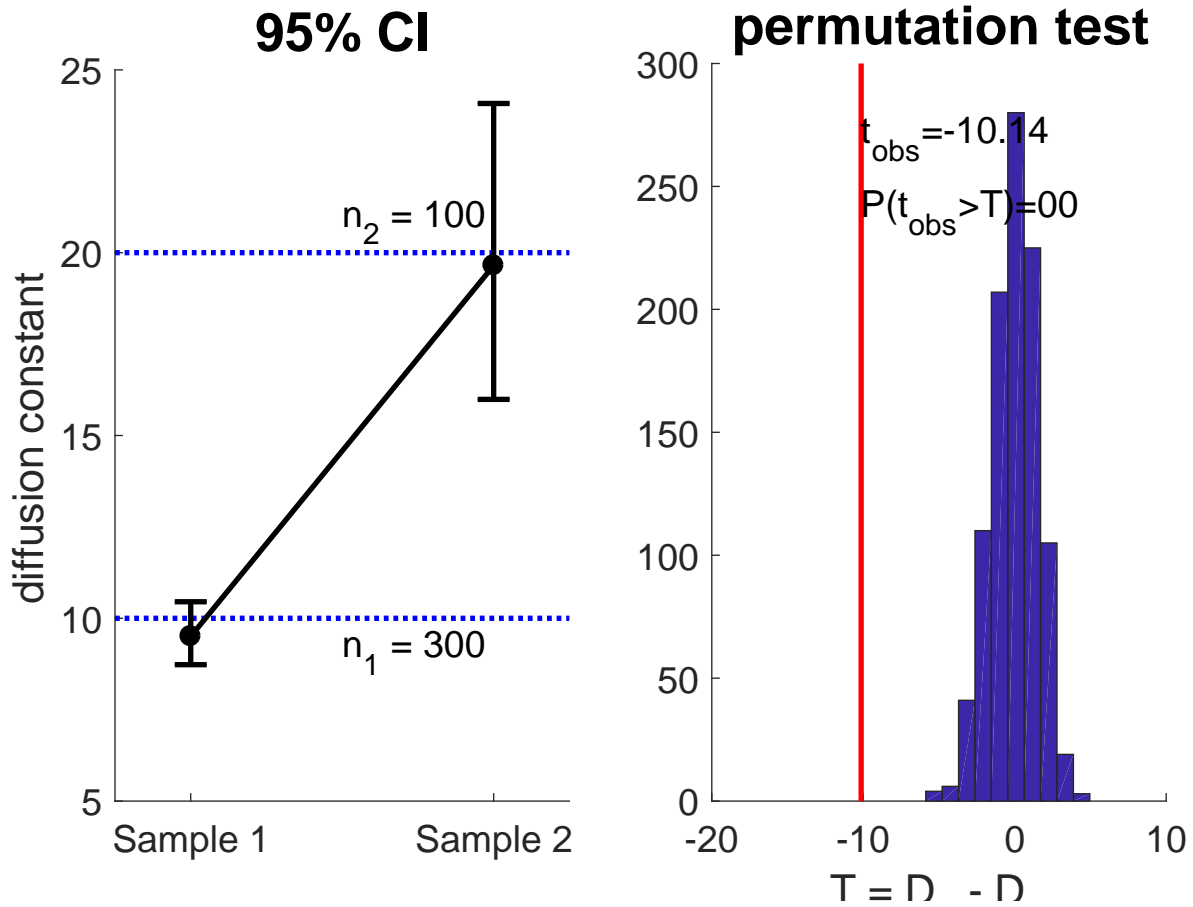


Figure 3: Simulation 2: $D_1 = 10$, $D_2 = 20$, $n_1 = 300$, $n_2 = 100$. The 95% confidence intervals do *not* overlap (left) and the estimated p-value of the hypothesis test is $p = 0$.

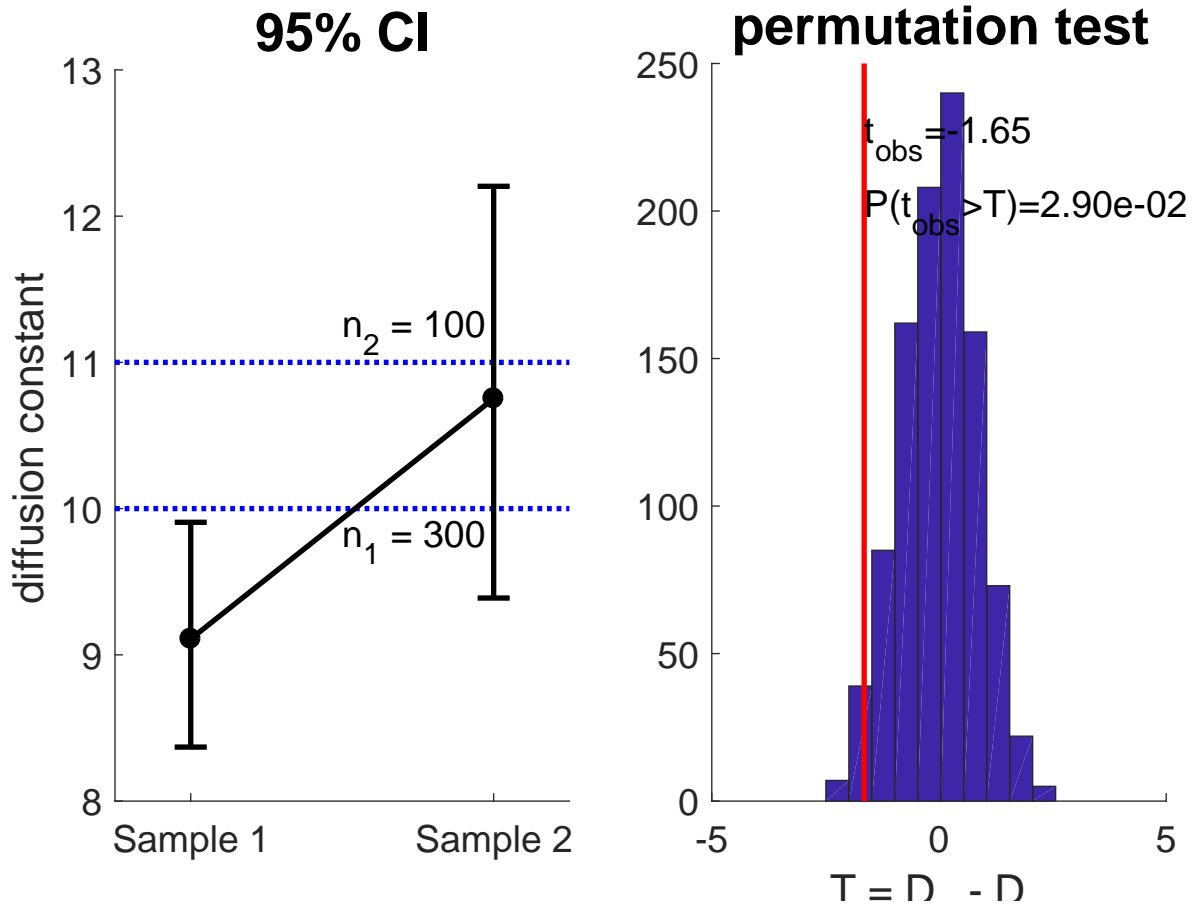


Figure 4: Simulation 2: $D_1 = 10$, $D_2 = 11$, $n_1 = 300$, $n_2 = 100$. The 95% confidence intervals do overlap (left) and the estimated p-value of the hypothesis test is $p = 0.029$.

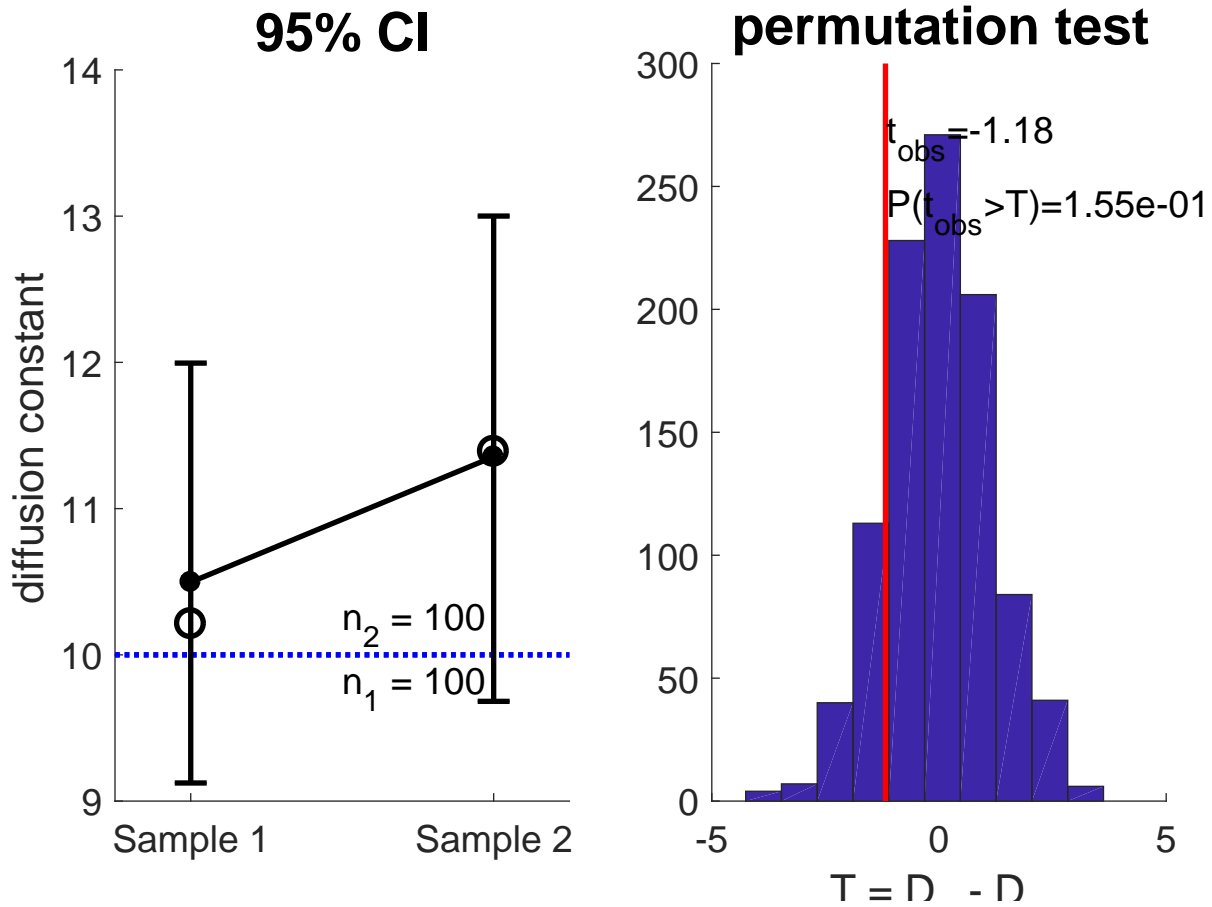


Figure 5: Simulation 2: $D_1 = 10$, $D_2 = 10$, $n_1 = 100$, $n_2 = 100$. The 95% confidence intervals do overlap (left) and the estimated p-value of the hypothesis test is $p = 0.155$.

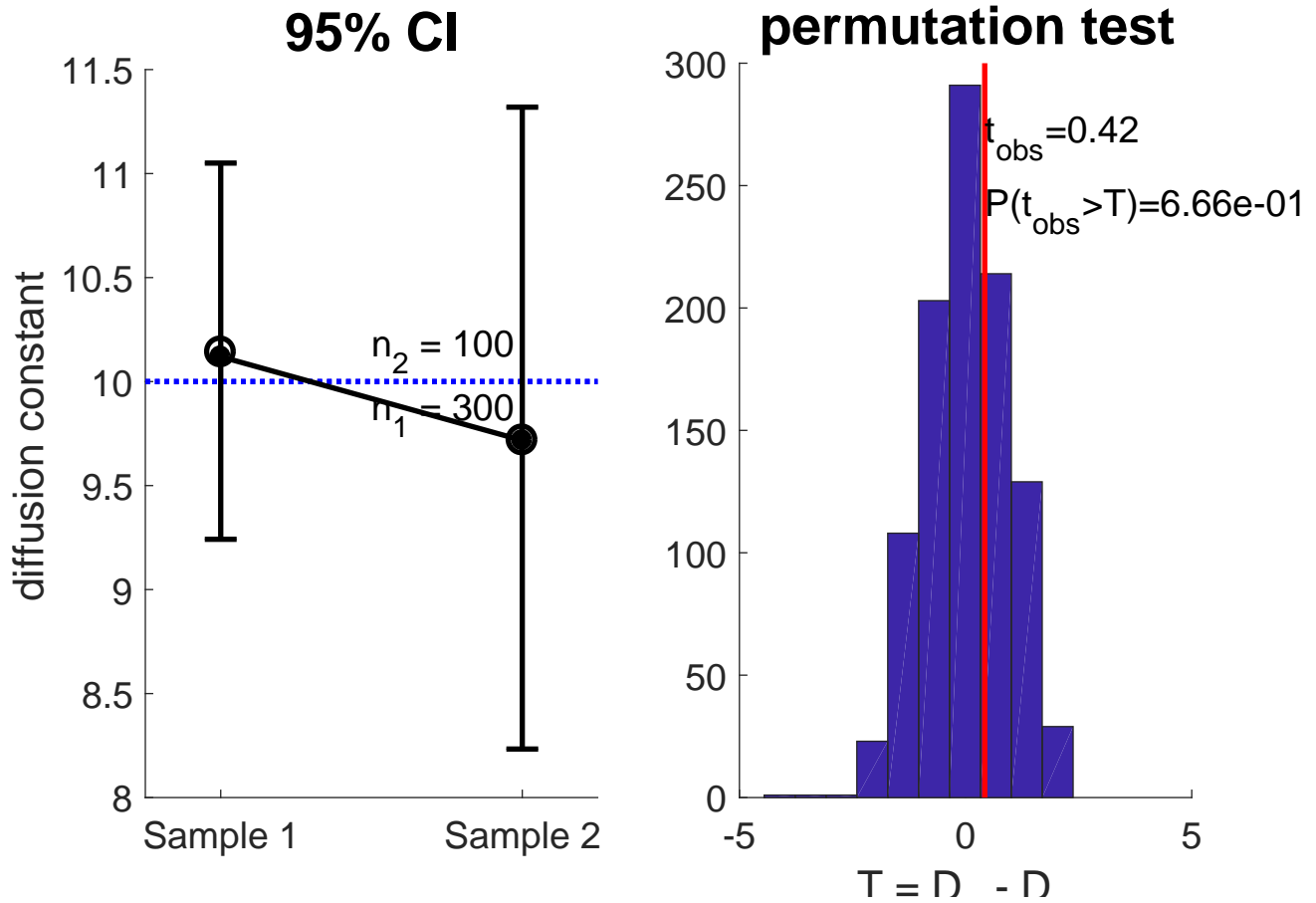


Figure 6: Simulation 2: $D_1 = 10$, $D_2 = 10$, $n_1 = 300$, $n_2 = 100$. The 95% confidence intervals do overlap (left) and the estimated p-value of the hypothesis test is $p = 0.66$.