



# (Elementary) Regression Methods & Computational Statistics (405.952)

## Part IV: Hypothesis Testing and Confidence Intervals (cont.)

**Assoz. Prof. Dr. Wolfgang Trutschnig**

Arbeitsgruppe Stochastik/Statistik

Fachbereich Mathematik

Universität Salzburg

[www.trutschnig.net](http://www.trutschnig.net)

Salzburg, January 2019





- ▶ We again return to the two-sided  $t$ -test.
- ▶ Suppose that  $X \sim \mathcal{N}(\mu_x, \sigma^2)$ , we do not know  $\mu_x$  and  $\sigma^2$ .
- ▶ Suppose that  $Y \sim \mathcal{N}(\mu_y, \sigma^2)$ , we do not know  $\mu_y$  and  $\sigma^2$ .
- ▶ Notice that the variance of  $X$  and  $Y$  is the same (and unknown).
- ▶ Given a sample  $X_1, \dots, X_n$  from  $X$  and a sample  $Y_1, \dots, Y_m$  from  $Y$  with  $m, n \geq 2$  we now want to calculate a confidence interval for the parameter  $\mu_D := \mu_x - \mu_y$ .
- ▶ Remember that when testing for  $H_0 : \mu_D = 0$  R also returned a 95%-confidence interval





```

1 mux <- muy <- 0
2 sigmax <- 1; sigmay <- 2
3 n <- 1000
4 x <- rnorm(n, mean=muy, sd=sigmax)
5 y <- rnorm(n, mean=muy, sd=sigmay)
6
7 test <- t.test(x,y, paired=FALSE, alternative="two.sided")
8 test

```

► yields

```

1 Welch Two Sample t-test
2
3 data:  x and y
4 t = 0.32697, df = 1515.4, p-value = 0.7437
5 alternative hypothesis: true difference in means is not equal to
6 0
7 95 percent confidence interval:
8 -0.1125703  0.1576068
9 sample estimates:
10 mean of x    mean of y
0.009047213 -0.013471049

```





- ▶ How is this 95%-confidence interval calculated?
- ▶ We know that  $S_{n,m}$  given by

$$S_{n,m} = \frac{\bar{X}_n - \bar{Y}_m - (\mu_x - \mu_y)}{\sqrt{S_p^2 \sqrt{\frac{1}{n} + \frac{1}{m}}}},$$

follows a  $t_{n+m-2}$ -distribution.

- ▶ A consequence

$$\mathbb{P}\left(S_{n,m} \in \left[t_{n+m-2; \frac{\alpha}{2}}, t_{n+m-2; 1-\frac{\alpha}{2}}\right]\right) = 1 - \alpha.$$

- ▶ Based on this we can easily derive the following confidence interval  $C_{n,m}^{1-\alpha}$  with coverage probability  $1 - \alpha$ :

$$C_{n,m}^{1-\alpha}(X_1, \dots, X_n, Y_1, \dots, Y_m) = C_{n,m}^{1-\alpha} = \left[\bar{X}_n - \bar{Y}_m - \Delta, \bar{X}_n - \bar{Y}_m + \Delta\right]$$

with  $\Delta = t_{n+m-2; 1-\frac{\alpha}{2}} \sqrt{S_p^2 \sqrt{\frac{1}{n} + \frac{1}{m}}}$ .





```

1 # t-test for H0: muD=mux-muy=0
2 mux <- 0
3 muy <- 0.2
4 sigmax <- sigmay <- 1
5 n <- m <- 50
6 x <- rnorm(n, mean=mux, sd=sigmax)
7 y <- rnorm(m, mean=muy, sd=sigmay)
8 test <- t.test(x, y, paired=FALSE, alternative="two.sided", var.
   equal=TRUE)
9 test
10
11 #confidence interval for muD manually
12 alpha <- 0.05
13 sp <- ((n-1)*var(x)+(m-1)*var(y))/(n+m-2)
14 Delta <- qt(p=1-alpha/2, df=n+m-2)*sqrt(sp *(1/n+1/m))
15 conf.int <- c(mean(x)-mean(y) - Delta, mean(x)-mean(y) + Delta)
16 test$conf.int[1:2]
17 conf.int

```

► yields

```

1 [1] -0.5283843  0.2504880
2 [1] -0.5283843  0.2504880

```





- ▶ Check if the confidence interval does what it should.

```

1 R <- 10000
2 error <- rep(0,R)
3 CI <- data.frame(lower=rep(0,R), upper=rep(0,R))
4 for(i in 1:R){
5   mux <- 0
6   muy <- 0.2
7   sigmax <- sigmay <- 1
8   n <- m <- 50
9   x <- rnorm(n, mean=mux, sd=sigmax)
10  y <- rnorm(m, mean=muy, sd=sigmay)
11  test <- t.test(x,y, paired=FALSE, alternative="two.sided", var.
    equal=TRUE)
12  CI[i,] <- test$conf.int[1:2]
13 }
14
15 CI$contained <- ifelse(CI$lower <= mux- muy & CI$upper >= mux- muy
    ,1,0)
16 coverage <- mean(CI$contained)
17 coverage
18 [1] 0.9504

```





- ▶ What happens if we change the values of  $\mu_x$  and  $\mu_y$ ?
- ▶ What happens if we change  $n$  and  $m$ ?
- ▶ How is the hypothesis test for  $H_0 : \mu_D = 0$  vs. the two-sided alternative related with the confidence interval?
- ▶ Answer: We reject  $H_0$  if and only if  $0 \notin C_{n,m}^{1-\alpha}$ , i.e. if the confidence interval does not contain 0.

**Exercise 39:** Confirm the just-stated answer by simulations and proceed as follows:

- ▶ Choose some values for  $\mu_x$  and  $\mu_y$  and simulate samples of  $X$  and  $Y$ .
- ▶ Apply the two-sided  $t$ -test and save the  $p$ -value as well as the confidence interval.
- ▶ Repeat the two steps  $R = 1.000$  times and verify if in all  $R$  cases we have that the  $p$ -value is less than 0.05 if and only if  $0 \notin C_{n,m}^{1-\alpha}$ .





- ▶ Suppose that  $x_1, \dots, x_n$  is a sample from  $X$  and that  $y_1, \dots, y_m$  is a sample from  $Y$ .
- ▶ We repeat the following steps  $R$  times:
  - ▶ Randomly draw  $n$  values from  $x_1, \dots, x_n$  and  $m$  values from  $y_1, \dots, y_m$  with (!) replacement
  - ▶ The resulting samples  $x_1^*, \dots, x_n^*, y_1^*, \dots, y_m^*$  are called bootstrap samples or bootstrap replications.
  - ▶ Calculate  $\bar{x}_n^* - \bar{y}_m^*$  and save this value.
- ▶ Let  $d_1^*, \dots, d_R^*$  denote the resulting values (i.e. the differences of the means of the bootstrap samples).
- ▶ The bootstrap confidence interval  $C_{n,m}^{*,1-\alpha}$  is then defined as the interval formed by the  $\frac{\alpha}{2}$ -quantile and the  $(1 - \frac{\alpha}{2})$ -quantile of the sample  $d_1^*, \dots, d_R^*$ , i.e.

$$C_{n,m}^{*,1-\alpha} = \left[ (F_d^*)^{-1} \left( \frac{\alpha}{2} \right), (F_d^*)^{-1} \left( 1 - \frac{\alpha}{2} \right) \right]$$

- ▶ Let's check the details in R.







```

1 mux <- 0
2 muy <- 0.2
3 sigmax <- sigmay <- 1
4 n <- m <- 50
5 x <- rnorm(n, mean=mux, sd=sigmax)
6 y <- rnorm(m, mean=muy, sd=sigmay)
7 test <- t.test(x, y, paired=FALSE, alternative="two.sided", var.
  equal=TRUE)           #just
8 test$conf.int[1:2]
9
10 boot.diff <- rep(0, R)
11 for(i in 1:R){
12   x.boot <- sample(x, size = n, replace = TRUE)
13   y.boot <- sample(y, size = m, replace = TRUE)
14   boot.diff[i] <- mean(x.boot) - mean(y.boot)
15 }
16 ci.boot <- as.numeric(quantile(boot.diff, probs = c(alpha/2, 1-
  alpha/2)))
17 #compare the two intervals
18 test$conf.int[1:2]
19 ci.boot

```

▶ yields (lucky coincidence?)

```

1 [1] -0.8296607 -0.0334093
2 [1] -0.82211378 -0.03919495

```





```

1 #systematical comparison of the two CIs
2 outer.R <- 100
3 Results <- data.frame(lower.t=rep(0, outer.R), lower.boot=rep(0,
4   outer.R), upper.t=rep(0, outer.R), upper.boot=rep(0, outer.R))
5 for(k in 1:outer.R){
6   mux <- 0; muy <- 0.2
7   sigmax <- sigmay <- 1
8   n <- m <- 50
9   x <- rnorm(n, mean=mux, sd=sigmax)
10  y <- rnorm(m, mean=muy, sd=sigmay)
11  test <- t.test(x,y, paired=FALSE, alternative="two.sided", var.
12    equal=TRUE) #just
13  Results[k,c(1,3)] <- test$conf.int[1:2]
14
15 R <- 1000; boot.diff <- rep(0,R)
16 for(i in 1:R){
17   x.boot <- sample(x, size = n, replace = TRUE)
18   y.boot <- sample(y, size = m, replace = TRUE)
19   boot.diff[i] <- mean(x.boot)-mean(y.boot)
20 }
21 Results[k,c(2,4)] <- as.numeric(quantile(boot.diff, probs = c(
22   alpha/2,1-alpha/2)))
23 }

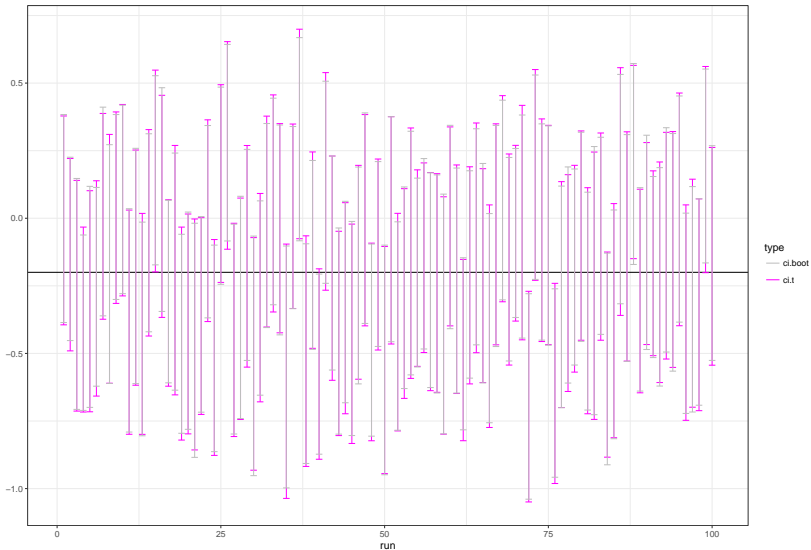
```



## Confidence intervals



The bootstrap confidence interval for  $\mu_D = \mu_x - \mu_y$



**Exercise 40:**

- ▶ Fix  $\mu \in \mathbb{R}$  and  $\sigma^2 > 0$ .
- ▶ Generate a sample  $X_1, \dots, X_n$  from  $X \sim \mathcal{N}(\mu, \sigma^2)$ .
- ▶ Calculate a bootstrap confidence-interval  $C_n^{1-\alpha}$  for the parameter  $\mu$  based on the sample.
- ▶ Use the  $t$ -test to get an exact confidence interval and compare the interval with the bootstrap interval.
- ▶ Repeat the previous steps to get a more systematic picture of the performance of the bootstrap confidence interval.



**Exercise 41:**

- ▶ Fix  $\mu \in \mathbb{R}$  and  $\sigma^2 > 0$ .
- ▶ Generate a sample  $X_1, \dots, X_n$  from  $X \sim \mathcal{N}(\mu, \sigma^2)$ .
- ▶ Calculate a bootstrap confidence-interval  $C_n^{1-\alpha}$  for the parameter  $\sigma^2$  based on the sample.
- ▶ Compare the exact confidence interval

$$I = \left[ \frac{(n-1)S_n^2}{\chi_{n-1; 1-\frac{\alpha}{2}}^2}, \frac{(n-1)S_n^2}{\chi_{n-1; \frac{\alpha}{2}}^2} \right]$$

and compare the interval with the bootstrap interval.

- ▶ Repeat the previous steps to get a more systematic picture of the performance of the bootstrap confidence interval.





## Exercise 42:

- ▶ We have already mentioned the correspondence between two-sided hypothesis tests and confidence intervals.
- ▶ Return to the situation discussed in the slides (confidence interval for  $\mu_D = \mu_x - \mu_y$ ) and use the bootstrap confidence interval to derive a bootstrap hypothesis test.
- ▶ Evaluate the performance of the test via simulations.



**Exercise 43 @mathematicians:**

- ▶ Fix  $\lambda > 0$ .
- ▶ Generate a sample  $X_1, \dots, X_n$  from  $X \sim \mathcal{E}(\lambda)$  (exponential distribution).
- ▶ Calculate a bootstrap confidence-interval  $C_n^{1-\alpha}$  for the parameter  $\lambda$  based on the sample.
- ▶ Evaluate the performance of the bootstrap confidence interval via simulations and compare the interval with an exact confidence interval (as derived in the UV 'Angewandte Statistik').

