

# HW11

Christian Sallinger

15 6 2021

## 1. Test power in the $z$ -test

Let  $X_1, \dots, X_n$  be i.i.d. random variables with  $X_1 \sim N(\mu, \sigma^2)$ , and  $H_0 : \mu = \mu_0$ .

- (a) Compute the test power of the left-sided  $z$ -test. Express it through cdf of the  $N(0, 1)$  distribution, depending on  $\mu_0, \mu, \sigma, n$  and the significance level  $\alpha$ .
- (b) Comment on the impact of  $\mu_0, \mu, \sigma, n$  and  $\alpha$  on the test power.

**Solution:** (a) In the left-sided  $z$ -test we test

$$H_0 : \mu = \mu_0 \quad \text{vs.} \quad H_1 : \mu < \mu_0$$

with the test statistic

$$\bar{X}_n \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

We then know that

$$Z = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$$

We reject the null at level  $\alpha$  if

$$\alpha = \mathbb{P}_{\mu_0}(\bar{X}_n \leq C) = \mathbb{P}_{\mu_0}\left(\frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}} \leq \frac{C - \mu_0}{\sigma/\sqrt{n}}\right)$$

So our critical value is

$$C = \mu_0 + z_{1-\alpha} \frac{\sigma}{\sqrt{n}}$$

where  $z_{1-\alpha}$  is the  $\alpha$ -quantile of the standard normal. The power of the test (for  $\mu \leq \mu_0$ ) is then

$$\pi(\mu) = \mathbb{P}_\mu(\bar{X}_n \leq C) = \mathbb{P}_\mu\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \leq \frac{C - \mu}{\sigma/\sqrt{n}}\right) = F_Z\left(\frac{\mu_0 - \mu + z_{1-\alpha} \cdot \sigma/\sqrt{n}}{\sigma/\sqrt{n}}\right) = F_Z\left(\frac{\mu_0 - \mu}{\sigma/\sqrt{n}} + z_{1-\alpha}\right)$$

(b) We can observe that the power increases in  $n$  and  $\alpha$ . The power also increases from  $\alpha$  to 1 as the standardized effect  $\frac{\mu_0 - \mu}{\sigma}$  increases.

## 2. Shock absorbers

A manufacturer of automobile shock absorbers was interested in comparing the durability of its shocks with that of the shocks produced by its biggest competitor. To make the comparison, one of the manufacturer's and one of the competitor's shocks were randomly selected and installed on the rear wheel of each of six cars. After the cars had been driven 20000 miles, the strength of each test shock was measured, coded, and recorded. Results are shown in the table.

Car number	Manufacturer's shock	Competitor's shock
1	8.8	8.4
2	10.5	10.1
3	12.5	12.0
4	9.7	9.3
5	9.6	9.0
6	13.2	13.0

Do these data present sufficient evidence to conclude there is a difference in the mean strength of the two types of shocks after 20000 miles of use?

**Solution:** Since the absorbers are presumably put into the same car we are dealing with a paired sample, the underlying random variables are not independent. We will use the test statistic for small sample sizes from page 31, lecture 11 to test

$$H_0 : \mu_d = 0 \quad \text{vs.} \quad H_1 : \mu_d \neq 0$$

where the target parameter is

$$\mu_d = \mu_1 - \mu_2.$$

We will first calculate the difference of the shocks:

Car number	Difference of the shocks
1	0.4
2	0.4
3	0.5
4	0.4
5	0.6
6	0.2

We will do our calculations with the help of R.

```
arr_shock = c(0.4,0.4,0.5,0.4,0.6,0.2)
n = length(arr_shock)

#Calculate our sample mean and variance
d_bar = mean(arr_shock)
S_d = sqrt(var(arr_shock))

#Calculate the test statistic t
t = d_bar/(S_d/sqrt(n))

#p-value for the 2-sided test
paste("p-val =",2*pt(-abs(t),n-1))

## [1] "p-val = 0.000597063879344651"
```

This is a small p-value, so it is evidence against the null. If we want to make life easier we can use the R function `t.test()` and set the logical value `paired` to be `true`.

```
arr_1 = c(8.8, 10.5, 12.5, 9.7, 9.6, 13.2)
arr_2 = c(8.4, 10.1, 12, 9.3, 9, 13)

t.test(arr_1, arr_2, paired = TRUE)

##
## Paired t-test
##
## data: arr_1 and arr_2
## t = 7.6787, df = 5, p-value = 0.0005971
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## 0.2771799 0.5561535
## sample estimates:
## mean of the differences
## 0.4166667
```

### 3. Simulation of test-power

Simulate the test-power in the two-sample  $t$ -test: Let  $X_1, \dots, X_n, Y_1, \dots, Y_n$  be independent random variables with  $X_i \sim N(0, \sigma^2)$  and  $Y_i \sim N(d, \sigma^2)$  for all  $i = 1, 2, \dots, n$ . Let the null hypothesis be  $H_0: d = 0$  and the significance level  $\alpha = 5\%$ . Simulate the test-power (by computing the relative frequency of rejections) for  $d \in \{-5, -4.5, -4, \dots, 5\}$  in 1000 simulations each. Use the parameters

- (a)  $n = 10$  and  $\sigma = 3$
- (b)  $n = 20$  and  $\sigma = 3$
- (c)  $n = 20$  and  $\sigma = 1$

for each of which you plot the testpower against  $d$ . Comment on your graphic. Hint: You can access the  $p$ -value with `t.test()$p.value`.

**Solution:**

```
#Simulates the test-power by computing the relative frequency of rejections over 1000
#simulations for the different values of d and then plots the simulated testpower against d
test_power = function(n,sigma){

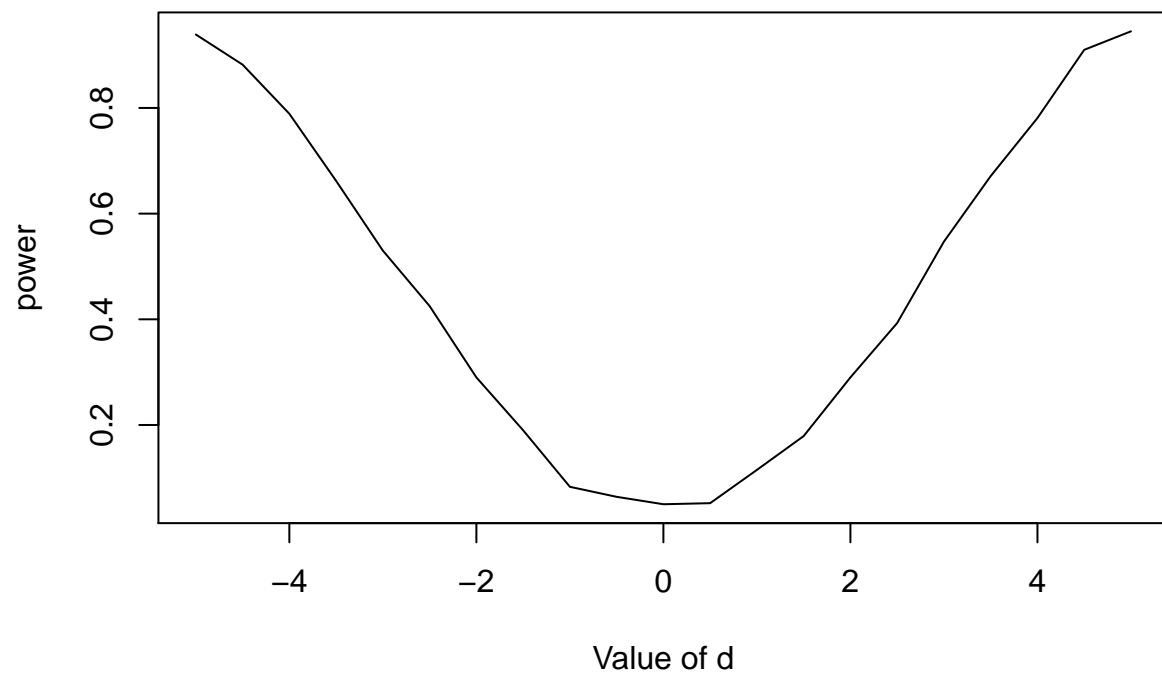
  d_arr = seq(-5,5,0.5)
  nr_rejections_arr = rep(0,length(d_arr))
  i = 0

  for(d in d_arr){
    i = i+1
    nr_rejections_arr[i] = nr_rejections_arr[i] + sum(replicate(1000,test(n,d,sigma))/1000)
  }
  plot(d_arr,nr_rejections_arr, xlab = "Value of d", ylab = "power", type = "l")
  title(paste("Simulation of the t test-power with n =",n," sigma =", sigma))
}

#Does a two sample t-test for n i.i.d. random variables X ~ N(0,sigma) and Y ~ N(d,sigma)
#returns 1 if the p-value is smaller than the alpha = 5% significance level and 0 otherwise
test = function(n,d,sigma){
  X = rnorm(n,0,sigma)
  Y = rnorm(n,d,sigma)
```

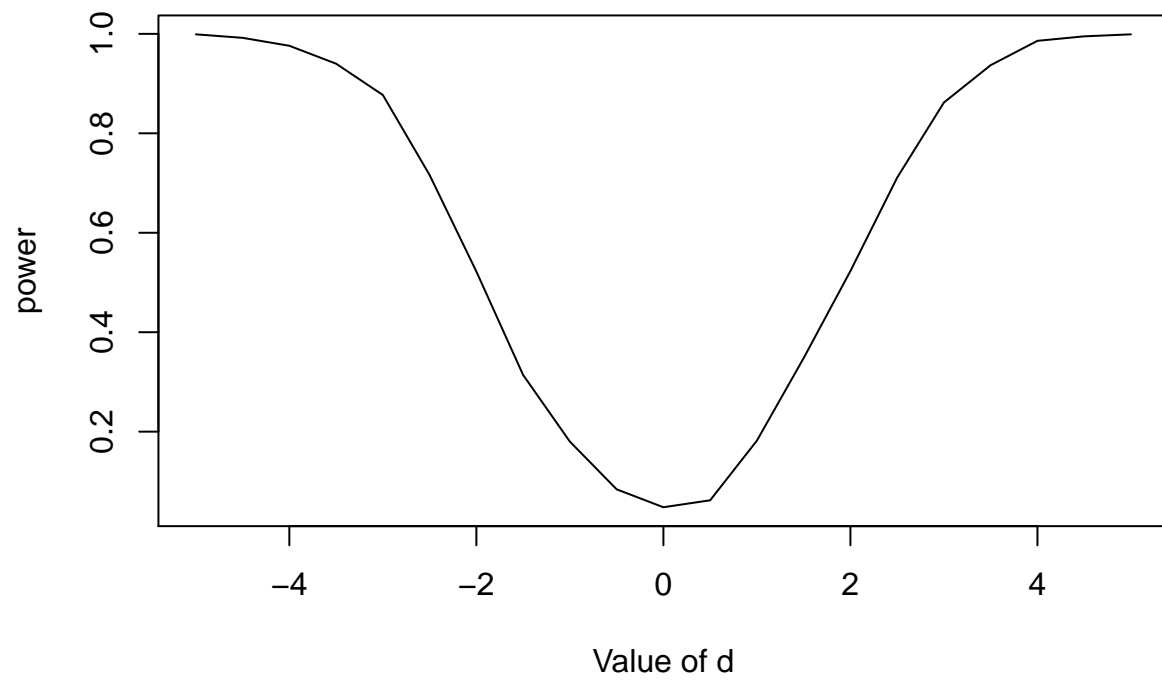
```
if(t.test(X, Y)$p.value < 0.05){  
  return(1)  
}  
else{  
  return(0)  
}  
}  
  
test_power(10,3)
```

### Simulation of the t test–power with $n = 10$ , $\sigma = 3$



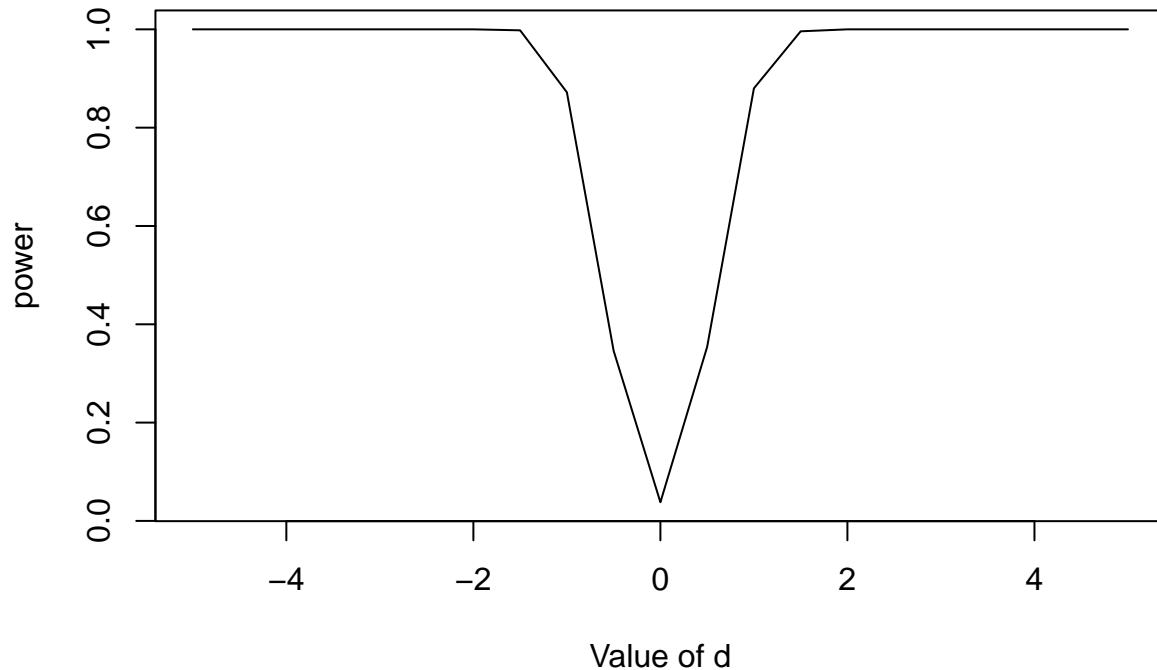
```
test_power(20,3)
```

### Simulation of the t test–power with $n = 20$ , $\sigma = 3$



```
test_power(20,1)
```

### Simulation of the t test–power with n = 20 , sigma = 1



We can see that the power increases both with the absolute value of  $d$  and the sample size. The power is also higher for lower values of  $\sigma$ . This coincides with our theoretical results for the one-sample t-test.

#### 4. Mechanics

In order to compare the means of two populations, independent random samples of 400 observations are selected from each population, with the following results:

Sample 1	Sample 2
$\bar{x}_1 = 5,275$	$\bar{x}_2 = 5,240$
$s_1 = 150$	$s_2 = 200$

- Use a 95% confidence interval to estimate the difference between the population means  $(\mu_1 - \mu_2)$ . Interpret the confidence interval.
- Test the null hypothesis  $H_0 : (\mu_1 - \mu_2) = 0$  versus the alternative hypothesis  $H_1 : (\mu_1 - \mu_2) \neq 0$ . Give the  $p$ -value of the test, and interpret the result.
- Suppose the test in the previous part were conducted with the alternative hypothesis  $H_1 : (\mu_1 - \mu_2) > 0$ . How would your answer change?
- Test the null hypothesis  $H_0 : (\mu_1 - \mu_2) = 25$  versus the alternative  $H_1 : (\mu_1 - \mu_2) \neq 25$ . Give the  $p$ -value, and interpret the result. Compare your answer with that obtained from the test conducted in part (b).
- What assumptions are necessary to ensure the validity of the inferential procedures applied in parts (a)-(d)?

#### Solution:

- Since our sample is large we can use the results from page 32 (lecture 11), which states that a  $100(1-\alpha)\%$  CI for  $\mu_1 - \mu_2$  is given by

$$\bar{x}_1 - \bar{x}_2 \pm z_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

so we can easily calculate our CI with R:

```
x_1 = 5.275
x_2 = 5.240
s_1 = 150
s_2 = 200
n = 400

z_alphah = qnorm(1-0.05/2)

paste("100(1-0.05)% CI = [", x_1 - x_2 - z_alphah*sqrt((s_1+s_2)/n), ", "
      , x_1 - x_2 + z_alphah*sqrt((s_1+s_2)/n) , "]" )

## [1] "100(1-0.05)% CI = [ -1.79837843014129 , 1.86837843014129 ]"
```

This random interval contains the true difference between the means with the specified confidence level.

- (b) On the same page we can also find a test statistic we can use for this test, to calculate the p-value we will again use R

```
z = (x_1-x_2)/sqrt((s_1+s_2)/n)

paste("p-value = ", 2*pnorm(-abs(z)))

## [1] "p-value = 0.970152857888392"
```

Since the p-value is high we will favor the null instead of the alternative hypothesis.

- (c) The test does not change but our critical value does, the p-value for this test is  $\mathbb{P}(Z \geq z)$  instead of  $2\mathbb{P}(Z \leq -|z|)$  as it was in the two-sided test in (b). (Where  $Z \sim \mathcal{N}(0,1)$ ).

```
paste("p-value = ", 1-pnorm(z))

## [1] "p-value = 0.485076428944196"
```

This p-value is still rather large and again supports the null, which is still the same as in (b).

- (d)

```
z_new = (x_1-x_2 - 25)/sqrt((s_1+s_2)/n)

paste("p-value = ", 2*pnorm(-abs(z_new)))

## [1] "p-value = 6.3653647108698e-157"
```

Here we see that the p-value is, unsurprisingly, very small so we will reject the null in favor of the alternative.

- (e) The assumption that the samples are independent and that the sample size is large ( $n > 30$ ) for both populations.

## 5. Comparing groups

Health professionals warn that transmission of infectious diseases may occur during the traditional handshake greeting. Two alternative methods of greeting (popularized in sports) are the high five and the first bump. Researchers compared the hygiene of these alternative greetings in a designed study and reported the results in the American Journal of Infection Control (Aug. 2014). A sterile-gloved hand was dipped into a culture of bacteria, then made contact for three seconds with another sterile-gloved hand via either a handshake, high

five, or fist bump. The researchers then counted the number of bacteria present on the second, recipient, gloved hand. This experiment was replicated five times for each contact method. Simulated data (recorded as a percentage relative to the mean of the handshake), based on information provided by the journal article, are provided in the table.

Handshake:	131	74	129	96	92
High five:	44	70	69	43	53
Fist bump:	15	14	21	29	21

- The researchers reported that more bacteria were transferred during a handshake compared with a high five. Use a 95% confidence interval to support this statement statistically.
- The researchers also reported that the first bump gave a lower transmission of bacteria than the high five. Use a 95% confidence interval to support this statement statistically.
- Based on the results, parts (a) and (b), which greeting method would you recommend as being the most hygienic?

### Solution:

We will first calculate the sample mean and variance for the greeting methods

```
#Put the values in arrays
handshake = c(131, 74, 129, 96, 92)
high_five = c(44, 70, 69, 43, 53)
fist_bump = c(15, 14, 21, 29, 21)

#Calculate the sample means
x_1 = mean(handshake)
x_2 = mean(high_five)
x_3 = mean(fist_bump)

#Calculate the sample variances
S_1 = sqrt(var(handshake))
S_2 = sqrt(var(high_five))
S_3 = sqrt(var(fist_bump))
print(paste(S_1, S_2, S_3))
```

```
## [1] "24.8052413816113 13.1034346642398 6"
```

- Under the assumption of independence we can guess that the variances are unequal and since we only have a small sample size we will use the CI as given in the lecture notes (page 34, lecture 11):

$$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2}(\nu) \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

where  $\nu$  is specified in the lecture notes.

```
#Function that returns the value of nu as given in the lecture notes
nu = function(var_1, var_2, n_1, n_2){
  num = (var_1^2/n_1 + var_2^2/n_2)^2
  denom = (var_1^2/n_1)^2/(n_1-1) + (var_2^2/n_2)^2/(n_2-1)
  return(num/denom)
}

t_alphah_1 = qt(1 - 0.025, nu(S_1, S_2, 5, 5))

pm = t_alphah_1 * sqrt((S_1^2 + S_2^2)/5)
```



```
paste("95% CI = [", x_1 - x_2 - pm, ",", x_1 - x_2 + pm, "]")
```

```
## [1] "95% CI = [ 17.9882239868181 , 79.2117760131819 ]"
```

```
t.test(handshake,high_five)
```

```
##
```

```
## Welch Two Sample t-test
```

```
##
```

```
## data: handshake and high_five
```

```
## t = 3.8738, df = 6.0711, p-value = 0.008044
```

```
## alternative hypothesis: true difference in means is not equal to 0
```

```
## 95 percent confidence interval:
```

```
## 17.98822 79.21178
```

```
## sample estimates:
```

```
## mean of x mean of y
```

```
## 104.4 55.8
```

As we can see the p-value is very small and the 95% confidence interval has its lower bound in the positive numbers, which supports the statement statistically.

(b) We can do the same as in (a)

```
t_alphah_2 = qt(1 - 0.025, nu(S_2, S_3, 5, 5))
```

```
pm = t_alphah_2 * sqrt((S_2^2 + S_3^2) / 5)
```

```
paste("95% CI = [", x_2 - x_3 - pm, ",", x_2 - x_3 + pm, "]")
```

```
## [1] "95% CI = [ 19.7571981448642 , 51.8428018551358 ]"
```

```
t.test(high_five, fist_bump)
```

```
##
```

```
## Welch Two Sample t-test
```

```
##
```

```
## data: high_five and fist_bump
```

```
## t = 5.5546, df = 5.6067, p-value = 0.0018
```

```
## alternative hypothesis: true difference in means is not equal to 0
```

```
## 95 percent confidence interval:
```

```
## 19.7572 51.8428
```

```
## sample estimates:
```

```
## mean of x mean of y
```

```
## 55.8 20.0
```

The same things we observed in (a) apply here.

(c) The most hygienic greeting method seems to be the fist bump.