

# Statistiska test

## Datorlaboration 3

Skriv ditt namn här

2025-03-13

Syftet med dagens laboration är att du ska

- träna på de grundläggande begreppen inom hypotesprövning (t.ex. signifikansnivå och styrka) samt vilka slutsatser man kan dra från analysen
- bekanta dig med lite av de funktioner som finns i R vad det gäller olika grundläggande statistiska test
- träna på att med hjälp av ett insamlat material konstruera en rimlig statistisk modell samt göra en kritisk granskning av modellen och dess förmåga att beskriva verkligheten
- tillämpa dina kunskaper och med hjälp av R analysera ett biostatistiskt datamaterial
- träna på att skriftligt redovisa antaganden, modeller och slutsatser från en statistisk analys.

## Innehållsförteckning

Förberedelseuppgifter . . . . .	1
<b>1. Grundläggande begrepp vid hypotestestning</b>	<b>2</b>
<b>2. Testets styrka och styrkefunktion</b>	<b>4</b>
Svar till exemplet med muntorrhet . . . . .	7
<b>3. Test av väntevärde i en population (t-test)</b>	<b>7</b>
<b>4. Jämförelse av väntevärden i två populationer (t-test vid två oberoende stickprov)</b>	<b>9</b>
<b>5. Test vid matchade data (stickprov i par)</b>	<b>12</b>
Svar till några av denna sektionens frågor: . . . . .	14
<b>6. Hade vår aktiveringskampanj effekt?</b>	<b>14</b>
Problemställningar i studien . . . . .	14
Frågor . . . . .	14

## Förberedelseuppgifter

Repetera begreppen *hypoteser*, *signifikansnivå*, *styrkefunktion* samt *modell med matchade data* och *modell med två oberoende stickprov*.

Du skall ha gjort följande uppgifter *innan* du kommer till laborationen.

### Hemuppgift 1

För att träna på de grundläggande begreppen i hypotest gör uppgifterna **Dig:4.4.1(felrisker)\_3, \_4 och \_5** på övning 7.

### Hemuppgift 2

För att träna på olika modeller gör **Dig:5.2\_1 och \_2**.

### Innan du sätter igång

Ladda in R-funktioner som behövs för laborationen. Dessa funktioner kommer från R-paketet Räkna med variation, men för att undvika problem med installation har vi plockat ut de ni använder och sparat i en fil. Du "berättar" för R att dessa funktioner finns genom att köra följande kod:

```
source("kod/funktioner_raknamedvariation_light.R")
```

## 1. Grundläggande begrepp vid hypotestestning

Läkemedel kan ge en nedsatt salivkörtelfunktion, vilket är en riskfaktor för karies och andra sjukdomar i munhålan. På 7 slumpmässigt utvalda patienter som alla fick samma medicin mätte man under 5 minuter den så kallade tuggstimulerade saliven. Normal mängd saliv under dessa förhållanden är 1 ml/min och muntorrhet anses föreligga när mängden saliv understiger 0.7 ml/min. Som modell antog man att salivmängden är normalfördelad med väntevärde  $\mu$  och standardavvikelse  $\sigma$ , där  $\sigma$  anses vara 0.5 ml/min. Intressanta frågeställningar är t.ex.:

- Stöder data vår misstanke att medicinen sänker salivproduktionen?
- Om medicinen ger upphov till en genomsnittlig salivproduktion på 0.8 ml/min, hur troligt är det att vi kommer att missa den nedsatta salivproduktionen med vårt test?

- Hur många patienter ska vi mäta på om vi vill att testet ska upptäcka en nedsatt salivproduktion på 0.7,ml/min med sannolikheten 0.95?

På kursens hemsida hittar ni data i filen `saliv.RData`. Kortfattade svar till frågorna som ställs i uppgifterna finns i slutet på denna del av handledningen.

#### Uppgift 1.1

Först vill man undersöka om data från de 7 patienterna stöder vår misstanke att medicinen sänker salivproduktionen. Ställ upp modell och lämpliga hypoteser.

**Modell:**

**Hypoteser:**

#### Uppgift 1.2

Beräkna medelvärde av mätningarna.

*# skriv din R-kod här*

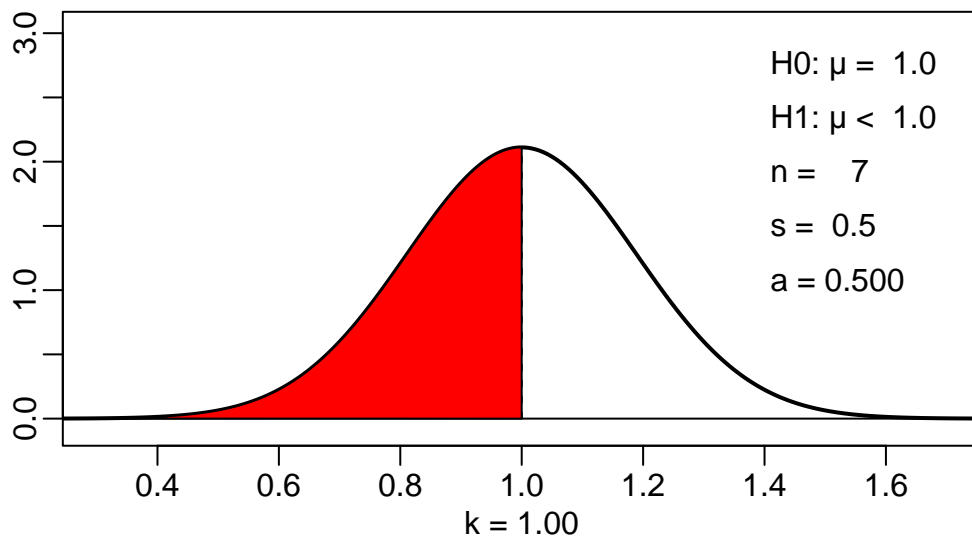
**Svar:**

#### Uppgift 1.3

Använd rutinen `hypotes` för att illustrera testets kritiska område då testet utförs på signifikansnivå  $\alpha = 0.05$ . Det aktuella kommandot är `hypotes( $\sigma$ ,  $n$ ,  $\mu_0$ ,  $\alpha$ , H1-riktn)`. Med hypoteserna  $H_0 : \mu = 1$  och  $H_1 : \mu < 1$  blir kommandot `hypotes(0.5, 7, 1, 0.05, '<')`. Använd rutinen fast med de hypoteser du har formulerat. (Bortse från de felmeddelanden som eventuellt kommer och titta på figuren.)

`hypotes(0.5, 7, 1, 0.5, '<')`

### Kritiskt område, $H_0: \mu = 1.0$ , $H_1: \mu < 1.0$



#### Uppgift 1.4

Rutinen markerar det kritiska området och anger ett värde  $k$  som är gränsen till området. Hur har  $k$  beräknats? Ställ upp den matematiska formeln för  $k$  och använd den för att själv beräkna  $k$ .

**Ekvation för  $k$ :**

*# skriv din R-kod här*

**Egen beräkning av  $k$ :**

#### Uppgift 1.5

Använd ditt beräknade stickprovsmedelvärde för att utföra testet. Vad är din slutsats om nollhypotesen,  $H_0$ ?

**Svar:**

#### Uppgift 1.6

Vad är din konkreta tolkning av signifikansnivån  $\alpha = 0.05$  i detta exempel?

**Svar:**

### Uppgift 1.7

Undersök hur det kritiska området ändras då du ändrar signifikansnivån till  $\alpha = 0.01$ .  
Vad är din slutsats nu?

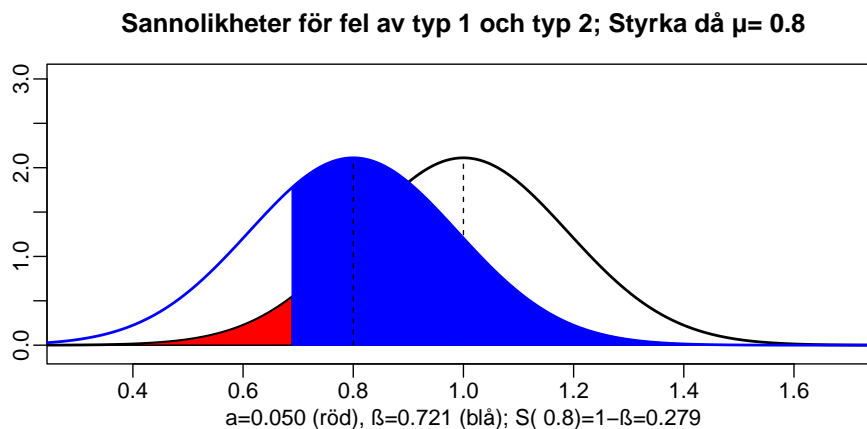
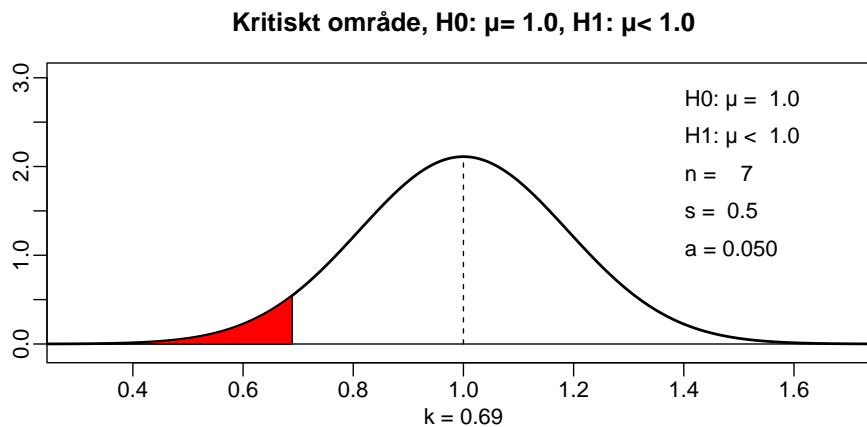
*# skriv din R-kod här*

Svar:

## 2. Testets styrka och styrkefunktion

Antag nu att genomsnittlig salivutsöndring i riskgruppen är 0.8. Då är förstås  $H_0 : \mu = 1$  falsk och vi vill att vårt test ska upptäcka detta och förkasta denna hypotes till förmån för hypotesen  $H_1 : \mu < 1$ . Sannolikheten att testet verkligen klarar av detta kallas för testets styrka i punkten 0.8. Använd rutinen `hypotes` för att illustrera testets styrka i punkten 0.8. Kommandot är nu `hypotes( $\sigma$ ,  $n$ ,  $\mu_0$ ,  $\alpha$ , H1-riktn, sant  $\mu$ )`, så i detta fall skriver du `hypotes(0.5, 7, 1, 0.05, '<', 0.8)`.

```
hypotes(0.5, 7, 1, 0.05, '<', 0.8)
```



### Uppgift 2.1

Rutinen ger dig ytterligare en figur som, förutom signifikansnivån  $\alpha$  (felrisk av typ I), även visar  $\beta$  (felrisk av typ II). Vad är den konkreta tolkningen av  $\beta$  i detta exempel? Hur hänger  $\beta$  ihop med testets styrka?

Generellt är testets styrka i punkten  $\mu$  sannolikheten att nollhypotesen förkastas om  $\mu$  är verklig genomsnittlig salivutsöndring i riskgruppen, d.v.s.

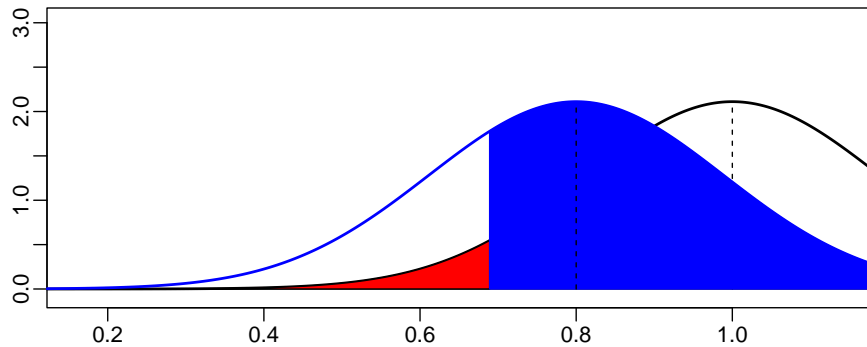
$$S(\mu) = P(H_0 \text{ förkastas} | \mu)$$

Observera att styrkan beror på värdet  $\mu$ . I detta exempel gäller att ju mindre  $\mu$  är i förhållande till  $\mu_0 = 1$  desto större är chansen att testet ska upptäcka att  $H_0$  inte gäller. Därför är det intressant att studera styrkan som en funktion av  $\mu$ , denna funktion betecknas ofta  $S(\mu)$ .

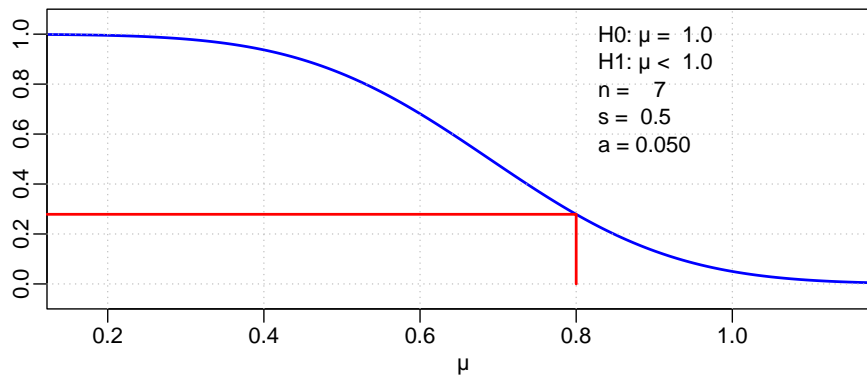
Kommandot `styrka(0.5, 7, 1, 0.05, '<', 0.8)` ger dig den tidigare figuren plus en graf av testets styrka som en funktion av  $\mu$ .

```
styrka(0.5, 7, 1, 0.05, '<', 0.8)
```

### Sannolikheter för fel av typ 1 och typ 2; Styrka då $\mu = 0.8$



$$S(\mu) = P(\text{förkasta } H_0); S(0.8) = 0.28$$



#### Uppgift 2.2

Uppskatta utifrån styrkefunktionen hur stor sannolikheten är att vi med vårt test kommer upptäcka att en grupp som bör klassas som muntorra ( $\mu = 0.7$ ) har en sänkt salivproduktion.

*# skriv din R-kod här*

**Svar:**

#### Uppgift 2.3

Hur många patienter bör vi mäta på om vi med sannolikheten 0.95 verkligen ska upptäcka att muntorra har en sänkt salivproduktion? Tips: Testa med olika värden på  $n$  i styrka och läs av ur grafen.

*# skriv din R-kod här*

**Svar:**

## Svar till exemplet med muntorrhet

1. Del 1
  - 1.1  $H_0 : \mu = 1; H_1 : \mu < 1$
  - 1.2  $k = \mu_0 - z_{1-\alpha} \frac{\sigma}{\sqrt{n}} = 1 - 1.6445 \cdot \frac{0.05}{\sqrt{7}} = 0.689$
  - 1.5 Eftersom medelvärde  $< 0.689$  förkastas  $H_0$  på nivå 0.05
  - 1.6 Det finns en chans på 5% att vi påstår att en person i riskgruppen har en sänkt salivproduktion när den i själva verket är normal
  - 1.7  $H_0$  kan ej förkastas på nivå 0.01
2. Del 2
  - 2.1  $\beta = P(\text{ej förkasta } H_0 | \mu)$ , d.v.s.  $\beta$  när det sanna väntevärdet är 0.8 är 1-strykan i punkten 0.8 = 1- S(0.8)
  - 2.2 Styrkan i punkten 0.7, S(0.7), kan enligt figuren uppskattas till 0.48
  - 2.3 Det krävs  $n = 30$  patienter för att styrkan ska vara 0.95 i punkten 0.7

## 3. Test av väntevärde i en population (t-test)

En blandning av blodserum innehåller exakt 42 g albumin per liter. Två laboratorier (A och B) får göra sex bestämningar var av koncentrationen. Vi vill undersöka om det finns någon systematisk avvikelse från det sanna värdet (42 g/l) i var och en av dessa två grupper.

```
load("data/lab3_filer/albumin.RData")
```

### Uppgift 3.1

Ange modell för A och B och specificera hypoteser för att testa om det finns en systematisk avvikelse från det sanna värdet i respektive grupp. Ledtråd: det ska vara en tvåsidig mothypotes.

**Modell för grupp A:** Låt  $X$  vara koncentration g albumin per liter i mätningar från laboration A. Väntevärde och varians för  $X$  är  $\mu_x$  och  $\sigma_x^2$ . Antag att mätningar är oberoende och likafördelade. Väntevärdet skattas med stickprovsmedelvärde  $\hat{\mu}_x = \bar{x}$ . Enligt centrala gränsvärdessatsen är  $\bar{x} \overset{A}{\sim} N(\mu_x, \frac{\sigma_x}{\sqrt{n_x}})$ .

**Hypoteser för grupp A:**

**Modell för grupp B:** Låt  $Y$  vara koncentration g albumin per liter i mätningar från laboration B.

**Hypoteser för grupp B:**

### Uppgift 3.2

I R gör man dessa två t-test med nedanstående rutinen **t-test**. Tolka utskrifterna när du kör kommandot. Hur stora är  $p$ -värdena och vad blir slutsatserna för respektive



grupp? Vad blir konfidensintervallen för de förväntade koncentrationerna?

```
t.test(AlbuminA, mu=42)
```

One Sample t-test

```
data: AlbuminA
t = 2.9194, df = 5, p-value = 0.03304
alternative hypothesis: true mean is not equal to 42
95 percent confidence interval:
 42.05974 42.94026
sample estimates:
mean of x
    42.5
```

```
t.test(AlbuminB, mu=42)
```

One Sample t-test

```
data: AlbuminB
t = -2.1843, df = 5, p-value = 0.08067
alternative hypothesis: true mean is not equal to 42
95 percent confidence interval:
 35.68716 42.51284
sample estimates:
mean of x
    39.1
```

Svar:

### Uppgift 3.3

Antag att vi vill undersöka om **laboration B** under- respektive överskattar koncentrationer. I R gör man detta som ett ensidigt test med rutinen **t-test**. Tolka utskriften för ensidiga test för laboration B. Vad händer med *p*-värden och intervall när man går från ensidigt till tvåsidiga alternativ?

```
t.test(AlbuminB, mu=42, alternative="less")
```

One Sample t-test

```
data: AlbuminB
t = -2.1843, df = 5, p-value = 0.04034
```

```

alternative hypothesis: true mean is less than 42
95 percent confidence interval:
    -Inf 41.77529
sample estimates:
mean of x
    39.1

t.test(AlbuminB, mu=42, alternative="greater")

```

#### One Sample t-test

```

data: AlbuminB
t = -2.1843, df = 5, p-value = 0.9597
alternative hypothesis: true mean is greater than 42
95 percent confidence interval:
    36.42471      Inf
sample estimates:
mean of x
    39.1

```

Svar:

## 4. Jämförelse av väntevärden i två populationer (t-test vid två oberoende stickprov)

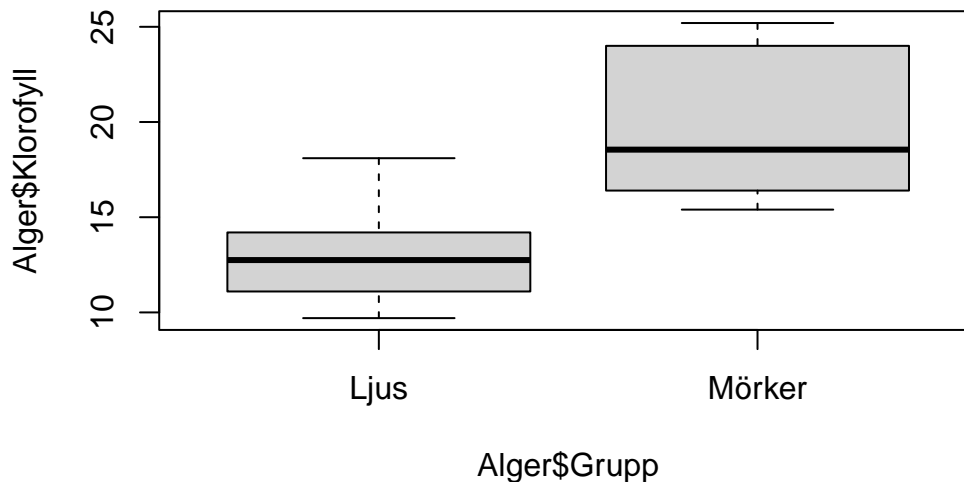
Alger i fick växa under ljusa respektive mörka förhållanden och därefter mättes klorofyllhalt.

```
load("data/lab3_filer/alger.RData")
```

### Uppgift 4.1

Gör en grafisk beskrivning av skillnaderna. Vad visar låd-diagrammet (boxplotten)?

```
boxplot(Alger$Klorofyll ~ Alger$Grupp)
```



Svar:

#### Uppgift 4.2

Nedan visar vi hur man kan sätta upp modell och hypoteser, och undersök med ett t-test om det finns skillnader i förväntad klorofyllhalt mellan de två grupperna. Läs igenom och se till att du förstår modellen.

Bekanta dig med hur rutinen `t.test` kan användas genom att lägga till gruppvariabeln. Se nedan. Lär dig att tolka utskriften från rutinen.

**Modell:** Låt  $X$  vara tillväxt hos alger i mörka förhållanden med väntevärde  $\mu_x$  och varians  $\sigma_x^2$ , och  $Y$  vara tillväxt hos alger i ljusa förhållanden med väntevärde  $\mu_y$  och varians  $\sigma_y^2$ . Väntevärden skattas med stickprovsmedelvärden  $\hat{\mu}_x = \bar{x}$  och  $\hat{\mu}_y = \bar{y}$ . Enligt centrala gränsvärdessatsen kommer skattningen av respektive väntevärde vara approximativt normalfördelad, d.v.s.  $\bar{x} \stackrel{A}{\sim} N(\mu_x, \frac{\sigma_x}{\sqrt{n_x}})$  och  $\bar{y} \stackrel{A}{\sim} N(\mu_y, \frac{\sigma_y}{\sqrt{n_y}})$ .

**Hypoteser:**  $H_0 : \mu_x = \mu_y$  mot  $H_1 : \mu_x \neq \mu_y$

```
t.test(Alger$Klorofyll ~ Alger$Grupp)
```

Welch Two Sample t-test

```
data: Alger$Klorofyll by Alger$Grupp
```

```
t = -3.2089, df = 9.0652, p-value = 0.01058
```

```
alternative hypothesis: true difference in means between group Ljus and group Mörker is not  
95 percent confidence interval:
```

```
-11.219229 -1.947438
sample estimates:
 mean in group Ljus mean in group Mörker
      13.10000      19.68333
```

Svar:

#### Uppgift 4.3

Om vi inte säger något annat förutsätter rutinen `t.test` att varianserna i de två grupperna, d.v.s.  $\sigma_x^2$  och  $\sigma_y^2$  är **olika** och kompenserar för det. Om vi vet (eller antar) att varianserna är lika, kan vi utnyttja det för att få högre frihetsgrader eftersom det är en parameter färre att skatta.

Undersök om varianserna är lika med ett så kallat  $\chi^2$ -test. Nedan visar vi hur man kan formulera hypoteser för att testa lika varians. Testet kan utföras i R genom rutinen `var.test`. Tolka utskriften!

**Hypoteser:**  $H_0 : \sigma_x^2 = \sigma_y^2$  mot  $H_1 : \sigma_x^2 \neq \sigma_y^2$

```
var.test(Alger$Klorofyll ~ Alger$Grupp)
```

F test to compare two variances

```
data: Alger$Klorofyll by Alger$Grupp
F = 0.51386, num df = 5, denom df = 5, p-value = 0.4825
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.07190457 3.67222210
sample estimates:
ratio of variances
 0.5138575
```

Svar:

#### Uppgift 4.4

Gör om t-testet i uppgift 4.2 där du antar att varianser är lika genom att lägga till `var.equal=TRUE` i R-rutinen. Vad blir dina slutsatser nu?

**Antagande:**  $\sigma^2 = \sigma_x^2 = \sigma_y^2$

```
t.test(Alger$Klorofyll ~ Alger$Grupp, var.equal=TRUE)
```

Two Sample t-test

```

data: Alger$Klorofyll by Alger$Grupp
t = -3.2089, df = 10, p-value = 0.009349
alternative hypothesis: true difference in means between group Ljus and group Mörker is not
95 percent confidence interval:
 -11.154519 -2.012148
sample estimates:
 mean in group Ljus mean in group Mörker
      13.10000      19.68333

```

Svar:

## 5. Test vid matchade data (stickprov i par)

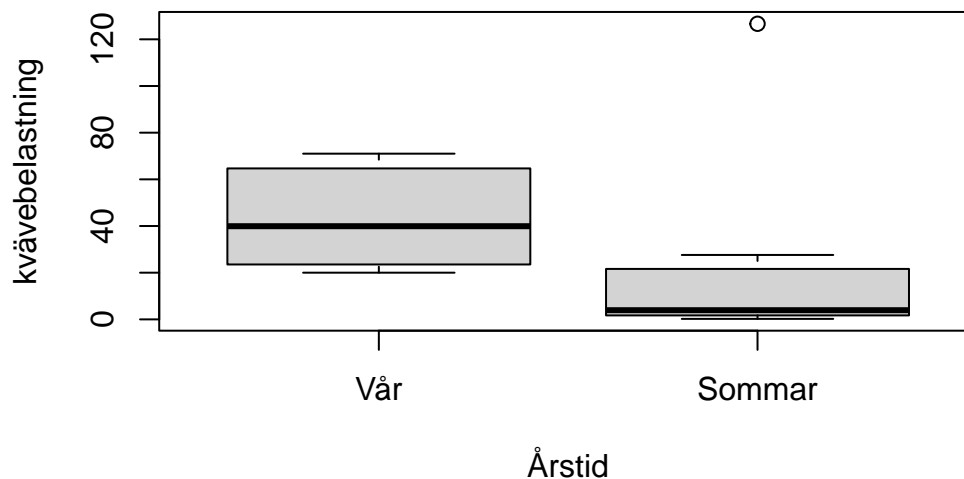
Det har gjorts mätningar av kvävebelastning under vår respektive sommar på ett antal dammar. Om man vill jämföra kvävebelastningen mellan de två årstiderna är en rimlig modell "stickprov i par".

```

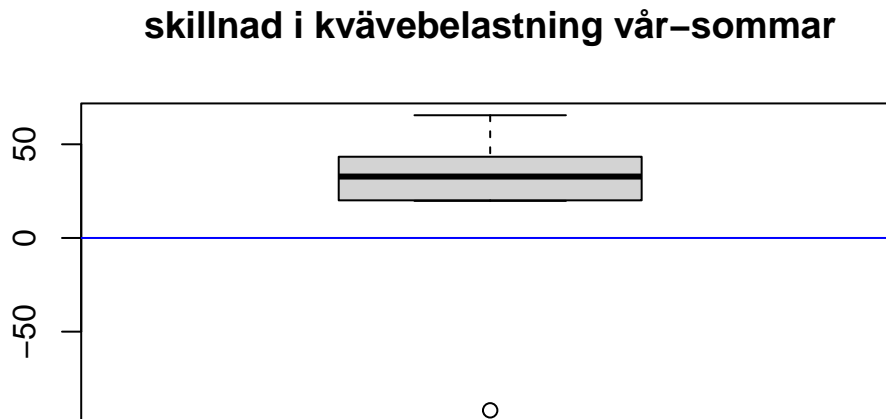
load("data/lab3_filer/dammar.RData")

boxplot(Dammar$N_belast_V, Dammar$N_belast_S,
        ylab = "kvävebelastning",
        xlab = "Årstid",
        names = c("Vår", "Sommar"))

```



```
boxplot(Dammar$N_belast_V-Dammar$N_belast_S,
        main = "skillnad i kvävebelastning vår-sommar")
abline(a=0,b=0,col='blue')
```



**Modell:** Låt  $D$  vara skillnaden i mätningar under vår och sommar för en damm. Slumpvariabeln  $D$  har väntevärde  $\mu$  och varians  $\sigma^2$ . Vi skapar ett stickprov  $(d_1, \dots, d_n)$  för skillnader i mätvärden. Väntevärdet skattas med stickprovsmedelvärdet,  $\hat{\mu} = \bar{d}$ . Enligt centrala gränsvärdessatsen är  $\bar{d} \overset{A}{\sim} N(\mu, \frac{\sigma}{\sqrt{n}})$ .

#### Uppgift 5.1

Finns det en skillnad i väntevärden mellan vår och sommar?

Vi testar det med följande **hypoteser**:  $H_0 : \mu = 0$  mot  $H_1 : \mu \neq 0$

Detta är ett vanligt test och man behöver inte räkna ut ett nytt stickprov för att köra rutinen i R. För att tala om för `t.test` att data är matchade i par lägger man till `paired=TRUE` som argument.

```
t.test(Dammar$N_belast_V, Dammar$N_belast_S, paired=TRUE)
```

Paired t-test

data: Dammar\$N\_belast\_V and Dammar\$N\_belast\_S

t = 1.2189, df = 7, p-value = 0.2624

alternative hypothesis: true mean difference is not equal to 0

95 percent confidence interval:

```
-19.49477 60.97227
sample estimates:
mean difference
20.73875
```

Svar:

### Svar till några av denna sektionens frågor:

1. Svar
  - 3.2 A: p-värde= 0.033; B: p-värde= 0.081; Vi kan påvisa en skillnad för A men inte för B. 95% konfidensintervall: A: (42,06, 42.94); B: (35.69, 42.51)
  - 4.3 Varianserna är inte olika (p-värde= 0.48)
  - 4.4 t-test: p-värde= 0.009; Vi kan påvisa en skillnad i klorofyll.
  - 5.1 p-värde= 0.262 Vi kan inte påvisa någon skillnad

## 6. Hade vår aktiveringskampanj effekt?

### Problemställningar i studien

De flesta forskare anser att hög kolesterolhalt i blodet är en riskfaktor för hjärt- och kärlsjukdomar. I en studie ville vi undersöka om man genom ett aktiveringsprogram bestående av flera faktorer (rökstopp, mental och fysisk träning) kan minska halten av kolesterol. Vi utgick från en grupp med 40 rökande män som samtliga hade något förhöjda kolesterolhalter i blodet. Av dessa 40 valde vi slumpmässigt ut 20 (A-grupp) som fick genomgå vårt aktiveringsprogram.

De övriga 20 (B-grupp) levde som vanligt, åtminstone såsom vi uppfattade det. Efter ett halvår mätte vi kolesterolhalten (mmol/l) på samtliga 40 män igen.

I filen `blodprov.Rdata` finns samtliga data och variablerna heter 'Afore', 'Aefter', 'Bfore' samt 'Befter'.

### Frågor

#### Fråga 6.1

Eftersom vi slumpmässigt valt ut de 20 som ska genomgå aktiveringsprogrammet bör det inte finnas några skillnader mellan A- och B-gruppen beträffande genomsnittlig kolesterolhalt **innan** studien börjar. Vi vill verkligen försäkra oss om detta så vi inte från början introducerar en systematisk skillnad mellan grupperna. Finns en skillnad mellan grupperna? Undersök med lämpligt statistiskt test.

### Fråga 6.2

Har den förväntade kolesterolhalten minskat hos grupp A? Undersök med lämpligt statistiskt test.

### Fråga 6.3

Det är inte otroligt att patienterna i grupp B, även om de inte genomgår aktiveringsprogrammet, ändå påverkas i sin kolesterolhalt eftersom uppmärksamhet kring frågorna kan ge effekt (en så kallad placebo-effekt).

Verkar det vara så i vår undersökning? Undersök om det finns en skillnad i kolesterol hos patienterna i grupp B före och efter programmet, med lämpligt statistiskt test.

### Fråga 6.4

Den övergripande frågan är om aktiveringsprogrammet sänkte den förväntade kolesterolhalten hos grupp A. Undersök med lämpligt statistiskt test om aktiveringsprogrammet har påverkat de två grupperna på olika sätt?

## Instruktioner för rapportering i del 6

Ange (om så är lämpligt):

- vilka antaganden du gör om data,
- vilka modeller du ställer upp,
- vilka hypoteser du testar,
- vilka satser du använder
- Redovisa resultat av analysen och vilka tolkningar och slutsatser du gör.
- För att bli godkänd på del 6, ska du undvika hänvisa till R-kod för att beskriva vad du har gjort.