

# Inlämningsuppgift block 4 – hypotestest, konfidensintervall

Statistiska Metoder med R

Grupp 10

Namn på dem som deltagit aktivt:

Jacob Widaeus 950729

Rozh Kader

Karl Wärnberg 940807

Beata Rosenberg

```
# Load dependencies  
library(tidyverse)
```

## 1. En enstaka observation jämförd med en känd population

Du har tidigare beräknat hur stor andel av befolkningen som har blodtrycket 80mmHg eller lägre. Förutsättningen för din beräkning är att variabeln blodtryck är normalfördelad, samt att du känner till medelvärdet och standardavvikelsen för populationen.

Samma typ av beräkning kan användas för ett hypotestest, en beräkning som talar om hur ovanlig en observation är, förutsatt att nollhypotesen är sann. Tänk dig att du mäter blodtryck på en person, resultatet blir 72 mmHg. Frågan som intresserar dig är: Hör den här personen till gruppen vanliga friska, eller hör hon till en grupp sjuka som i genomsnitt har lägre blodtryck än friska?

### Kommenterat exempel

Nollhypotes: Personen hör till populationen med fördelningen blodtryck  $\sim N(90, 10)$

Alternativhypotes: Personen hör till en population med lägre blodtryck än  $N(90, 10)$

**1.1** Är det ett ensidigt eller tvåsidigt test vi har formulerat? Val av signifikansnivå: Vi väljer  $\alpha = 0.05$ . Det innebär att vi förkastar nollhypotesen om p-värdet är lägre än 0.05. Genom att välja  $\alpha = 0.05$  har vi accepterat att om vi gör upprepade tester där nollhypotesen verkligen är sann så kommer vi förkasta nollhypotesen av misstag i 5% av testen på lång sikt.

Svar: Det är ett ensidigt test eftersom alternativhypotesen anger att personen hör till en population med lägre blodtryck än  $N(90, 10)$ .

**1.2** Är det ett typ I eller ett typ II fel att förkasta en sann nollhypotes?

Svar: Att förkasta en sann nollhypotes är ett typ I fel.

Val av statistisk test: z- test

Titta på figuren "ensidigt test" nedan. y är andelen av befolkningen som har ett blodtryck nära x. Areal under hela kurvan är 1. Det betyder att alla i befolkningen har ett blodtryck, alla möjliga utfall ryms under kurvan.

Antag att nollhypotesen stämmer, personen du mätt hör till populationen  $N(90, 10)$ . Hur vanligt är det att man får resultatet 72 eller lägre av en slump? Svaret är arean under kurvan till vänster om  $bt = 72$ . Den arean är det samma som z-testets p-värde.

**1.3** Beräkna blå area i figuren "Ensidigt test" med R

```
# Given information
mean <- 90
sd <- 10
value <- 72

# Beräkna p värdet, för normalfördelad population
p_value <- pnorm(value, mean, sd)
print(p_value)
```

```
[1] 0.03593032
```

**1.4** Återskapa min figur "Ensidigt test" nedan med R

```

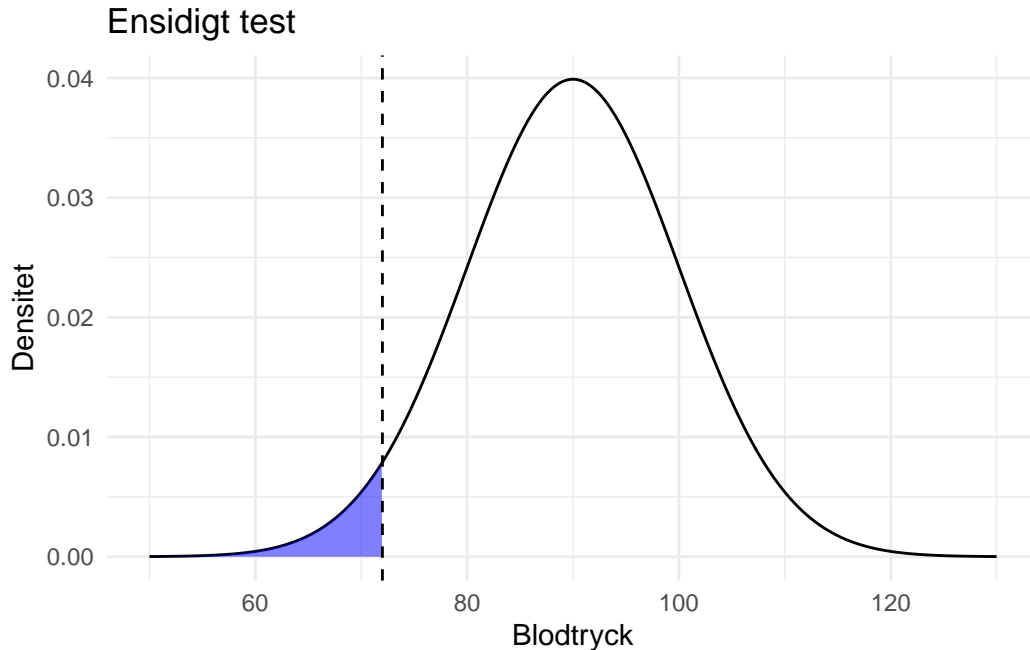
library(ggplot2)

# Givna parametrar
mean <- 90
sd <- 10
value <- 72

# Skapa x värden
x <- seq(mean - 4*sd, mean + 4*sd, length=1000)
# Beräkna fördelning
y <- dnorm(x, mean, sd)
# df
data <- data.frame(x, y)

# Plotta
ggplot(data, aes(x, y)) +
  geom_line() +
  geom_area(data = subset(data, x <= value), aes(x = x, y = y),
    fill = "blue", alpha = 0.5) +
  geom_vline(xintercept = value, linetype = "dashed") +
  labs(title = "Ensidigt test", x = "Blodtryck", y = "Densitet") +
  theme_minimal()

```



**1.5** Blodtrycket  $\leq 72$  eller lägre verkar förekomma i mindre del av befolkningen än 0.05, vår valda signifikansnivå. Ska vi behålla eller förkasta nollhypotesen?

Svar: Eftersom blodtrycket på 72 mmHg är lägre än signifikansnivån på 5%, bör vi förkasta nollhypotesen.

### Tvåsidigt test

Det är vanligt att man måste formulera en tvåsidig alternativhypotes. När jag drar ett värde så ska jag fråga mig "hur vanligt är det att hitta en person som avviker så här mycket från medlet, uppåt eller neråt". Det är lika vanligt att avvika 18 mmHg neråt eller mer som 18 mmHg uppåt eller mer. För att beräkna p-värdet för ett tvåsidigt test måste du summera de två blåa areorna i figuren "tvåsidigt test" nedan. Använd samma mätning igen, 72 mmHg, med en ny alternativhypotes: Nollhypotes: Personen hör till populationen med fördelningen blodtryck  $\sim N(90, 10)$  Alternativhypotes: Personen hör till en population med högre eller lägre blodtryck än  $N(90, 10)$

**1.6** Kan nollhypotesen förkastas med  $\alpha = 0.05$  som signifikansnivå?

```
# Given information
mean <- 90
sd <- 10
value <- 72
alfa <- 0.05

# Beräkna p värde
p_lower <- pnorm(value, mean, sd)
p_upper <- pnorm(mean + (mean - value), mean, sd, lower.tail = FALSE)

# Two-sided
p_value_two_sided <- p_lower + p_upper

if (p_value_two_sided < alfa) {
  print(paste("Nollhypotesen kan förkastas med", round(p_value_two_sided, 4)))
} else {
  print(paste("Nollhypotesen kan inte förkastas med", round(p_value_two_sided, 4)))
}
```

```
[1] "Nollhypotesen kan inte förkastas med 0.0719"
```

Uttrycka skillnader med variation

Att göra ett experiment handlar ofta om att mäta en skillnad. Blev det skillnad i blodtryck mellan kontrollgruppen och gruppen som medicinerats? Ja, alltid blir det någon skillnad mellan två stickprov, även om medicineringen inte påverkar blodtrycket som du hade hoppats. Det blir skillnad av en slump. 3 mmHg blev det kanske. Är det mycket eller lite? Är 10 kg stor skillnad? Kanske för gäddor, kanske inte för människor. För att besvara frågan måste vi byta ut enheten mmHg, kg, cm och istället uttrycka skillnaden med variation som enhet. Övningen som följer ska illustrera detta.

Undersök formeln för z-testet.

$$z = (X - \mu) / \sigma$$

X är din uppmätta variabel

är medelvärdet för populationen som X hör till enligt nollhypotesen

$\sigma$  är populationens standardavvikelse

Många variabler X kan vara normalfördelade med olika medel och standardavvikelser. Siffrorna blir olika beroende på vilken enhet man använder.

Titta på variablerna X1 och X2 och deras fördelningar, beskriver de samma fenomen i naturen?

$$X1 \sim N(\mu = 25, \sigma = 2.5)$$

$$X2 \sim N(\mu = 10, \sigma = 1)$$

Javisst, det är längden på en sorts fiskar, X1 är längden angivet i cm, X2 är längden angivet i tum.

z är en speciell variabel med fördelningen  $z \sim N(0, 1)$ , med andra ord, z är normalfördelad med medelvärdet 0 och standardavvikelsen 1. För att rita upp  $z \sim N(0,1)$  i R kan du använda

```
curve(dnorm(x, 0, 1), from= -4, to=4)
```

För att transformera en variabel till z behöver du genomföra stegen A och B nedan:

A. subtrahera  $\mu$  från varje X - då blir medelvärdet noll

exempel:  $X = 80 \text{ mmHg}$ ;  $X \sim N(90, 10)$

$80 \text{ mmHg} - 90 \text{ mmHg} = -10 \text{ mmHg}$

```
bt<- seq(-50, 120, 1)
```

```
y = dnorm(bt, 90, 10)
```

```
plot(bt, y, type="l", xlim=c(-50, 120), main="transformera så  $\mu$  blir noll")
```

```
lines(bt -90, y, xlim=c(-50, 120), col="blue")
```

B. Dividera varje X med standardavvikelsen – då blir skalan på x-axeln standardavvikelse.

Fortsättning på samma exempel:  $-10 \text{ mmHg} / 10 \text{ mmHg} = -1$

1.7 Rita en graf som illustrerar transformeringen av x-axelns skala från mmHg till standardavvikelse

```
# Skapa en sekvens för blodtrycksvärden
bt <- seq(-50, 120, 1)

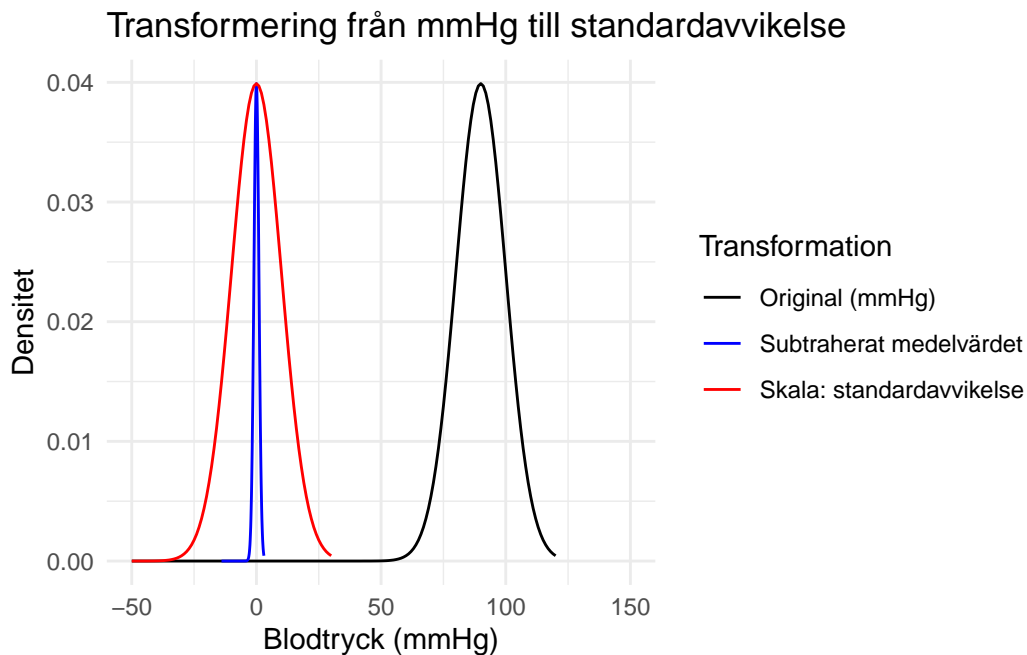
# Beräkna densitetsfunktion för normalfördelning (N(90, 10))
y <- dnorm(bt, mean = 90, sd = 10)

# data.frame
data <- data.frame(
  bt = bt,
  y = y,
  original = bt,
  subtracted_mean = bt - 90,
  standardized = (bt - 90) / 10
)

# Data manipulation för grafen
data_long <- data %>%
  pivot_longer(cols = c(original, subtracted_mean, standardized),
               names_to = "transformation",
               values_to = "value")

# Plotta
ggplot(data_long, aes(x = value, y = y, color = transformation)) +
  geom_line() +
  scale_color_manual(values = c("black", "blue", "red"),
                    labels = c("Original (mmHg)",
                              "Subtraherat medelvärde",
                              "Skala: standardavvikelse")) +
  labs(title = "Transformerings från mmHg till standardavvikelse",
       x = "Blodtryck (mmHg)",
       y = "Densitet",
       color = "Transformation") +
  xlim(-50, 150) +
  theme_minimal()
```

Warning: Removed 90 rows containing missing values or values outside the scale range (`geom\_line()`).



## 2. Ett stickprov jämförs med en känd population

Du har i inlämningsuppgift 3 dragit upprepade stickprov ur populationer med olika underliggande fördelningar. Du har även läst om central limit theorem att samplingsfördelningen går mot normalfördelning om stickprovet är stort nog.

Du vill testa om ett stickprov är draget ur en population med känt medelvärde och standardavvikelse. Rätt metod är z-test, men formeln ser något annorlunda ut än i 1.

$$z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$\bar{X}$  är stickprovets medelvärde

$\mu$  är populationens medelvärde

$n$  är stickprovets storlek

$\sqrt{n}$  är kvadratroten ur stickprovets storlek

**2.1** Varför dividerar man populationens standardavvikelse med kvadratroten ur stickprovets storlek. Du behöver inte härleda matematiken exakt, men anknyt gärna till resultaten i inlämningsuppgift 3 i ditt svar.

Kvalitetskontrollavdelningen på en fabrik testar regelbundet en produkt. Medelvikten är 31.2 g med standardavvikelsen 0.4 g. En morgon vägs ett slumpmässigt stickprov av produkten, 16 exemplar, och stickprovets medelvikt uppmäts till 31.38 g.

Är produkten tyngre än normalt denna morgon?

Svar: Anledningen till att man dividerar populationens standardavvikelse ( $\sigma$ ) med kvadraten av stickprovets storlek ( $\sqrt{n}$ ) är för att hantera osäkerheten i uppskattningen av medelvärdet från stickprovet. Enligt central limit theorem blir samplingsfördelningen av stickprovets medelvärde normalfördelad när stickprovet är stort nog, och dess spridning minskar när stickprovet blir större.

```
# Parametrar
mu <- 31.2          # Populationens medelvärde
sigma <- 0.4        # Populationens standardavvikelse
x_bar <- 31.38      # Stickprovets medelvärde
n <- 16            # Stickprovets storlek

# Beräkna z-värdet
z_value <- (x_bar - mu) / (sigma / sqrt(n))

# Beräkna p-värdet (ensidigt test, höger-svans)
p_value <- 1 - pnorm(z_value)

print(paste("Z-värde:", z_value))
```

```
[1] "Z-värde: 1.8"
```

```
print(paste("P-värde:", p_value))
```

```
[1] "P-värde: 0.035930319112926"
```

```
if (p_value < 0.05) {
  print("Produkten är signifikant tyngre än normalt.")
}
```

```
[1] "Produkten är signifikant tyngre än normalt."
```

## 2.2 Ange nollhypotes och mothypotes

Svar: Nollhypotes ( $H_0$ ): Det finns ingen skillnad i medelvikt, dvs. produkternas medelvikt är fortfarande 31.2 g. Stickprovet kommer från samma population. Mothypotes ( $H_1$ ):



Produkten är tyngre än normalt denna morgon.

### 2.3 välj signifikansnivå

Svar: En vanlig signifikansnivå är 5%, vilket innebär att vi förkastar nollhypotesen om sannolikheten att observera ett stickprov med den uppmätta medelvikten, eller mer extremt, är mindre än 5%.

### 2.4 Beräkna z-test i R.

Du kan transformera X till z med formeln  $z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$  och sedan slå upp p värdet med `pnorm()`. Så måste du göra om du bara har en z-tabell att tillgå för att komma åt p-värdet.

Det går även utmärkt att låta R göra transformeringen genom att fylla i värdet i g i `pnorm()` tillsammans med medel och standardavvikelse i gram.

Ange om nollhypotesen förkastas eller behålls.

```
# Parametrar
mu <- 31.2          # Populationens medelvärde
sigma <- 0.4        # Populationens standardavvikelse
x_bar <- 31.38      # Stickprovets medelvärde
n <- 16             # Stickprovets storlek

# Beräkna z-värdet
z_value <- (x_bar - mu) / (sigma / sqrt(n))

# Beräkna p-värdet (ensidigt test, höger-svans)
p_value <- 1 - pnorm(z_value)

# Skriv ut resultaten
print(paste("Z-värde:", z_value))
```

```
[1] "Z-värde: 1.8"
```

```
print(paste("P-värde:", p_value))
```

```
[1] "P-värde: 0.035930319112926"
```

```
# Kontrollera om vi kan förkasta nollhypotesen
if (p_value < 0.05) {
  print("Nollhypotesen förkastas. Produkten är signifikant tyngre än normalt.")
} else {
  print("Nollhypotesen behålls. Det finns ingen signifikant skillnad i produktens vikt.")
}
```

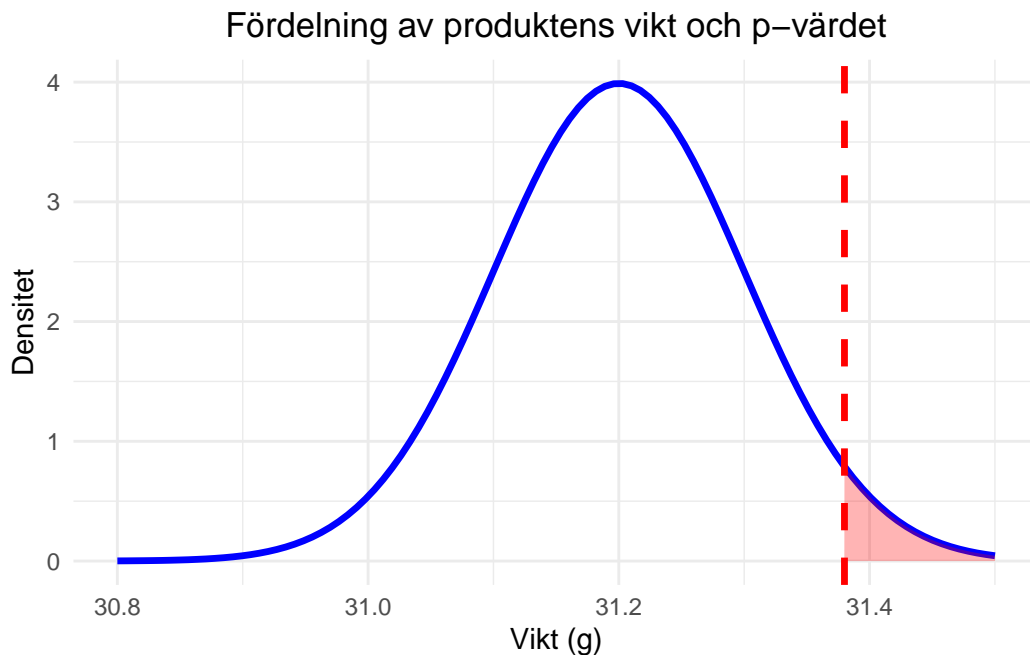
```
[1] "Nollhypotesen förkastas. Produkten är signifikant tyngre än normalt."
```

**2.5** Rita en graf som visar stickprovets förväntade fördelning och arean som motsvarar p-värdet. Välj själv om du vill rita grafen med  $X$  (otransformerade värden i gram) eller med  $z$  (transformerade värden, standardavvikelsen som enhet) som x-axel.

```
# Skapa en data frame för normalfördelningen
bt_values <- seq(30.8, 31.5, length.out = 100)
densities <- dnorm(bt_values, mean = mu, sd = sigma / sqrt(n))
data <- data.frame(bt_values = bt_values, densities = densities)

# Skapa en data frame för p-värdesområdet
fill_values <- seq(x_bar, 31.5, length.out = 100)
fill_densities <- dnorm(fill_values, mean = mu, sd = sigma / sqrt(n))
fill_data <- data.frame(fill_values = fill_values, fill_densities = fill_densities)

# Rita grafen
ggplot(data, aes(x = bt_values, y = densities)) +
  geom_line(color = "blue", linewidth = 1.2) + # Fördelningskurvan
  geom_vline(xintercept = x_bar, color = "red", linetype = "dashed",
    linewidth = 1.2) + # Stickprovets medelvikt
  geom_area(data = fill_data, aes(x = fill_values, y = fill_densities),
    fill = "red", alpha = 0.3) + # Fyll p-värdesområdet
  labs(title = "Fördelning av produktens vikt och p-värdet",
    x = "Vikt (g)", y = "Densitet") +
  theme_minimal() +
  theme(plot.title = element_text(hjust = 0.5)) # Centrerar titeln
```



**2.6** Formulera resultatet från 2.4 utan att använda statistisk terminologi.

Svar: Den produkt som vägdes denna morgon verkar vara lite tyngre än vad som är normalt. Vi har beräknat att sannolikheten för att få ett resultat som detta, om produkten i genomsnitt faktiskt vägde lika mycket som den brukar, är låg. Därför kan vi misstänka att vikten på produkterna denna morgon är högre än vanligt.

### 3. Ett stickprov jämförs med en population, standardavvikelsen är okänd

Ofta har man inte tillgång till standardavvikelsen för populationen. Då blir man tvungen att göra en uppskattning av populationens standardavvikelse  $\sigma$ , utifrån stickprovets standardavvikelse som betecknas SD eller s. Istället för z-test ska man använda t-test

$$t = \frac{\bar{X} - \mu}{\frac{SD}{\sqrt{n}}}$$

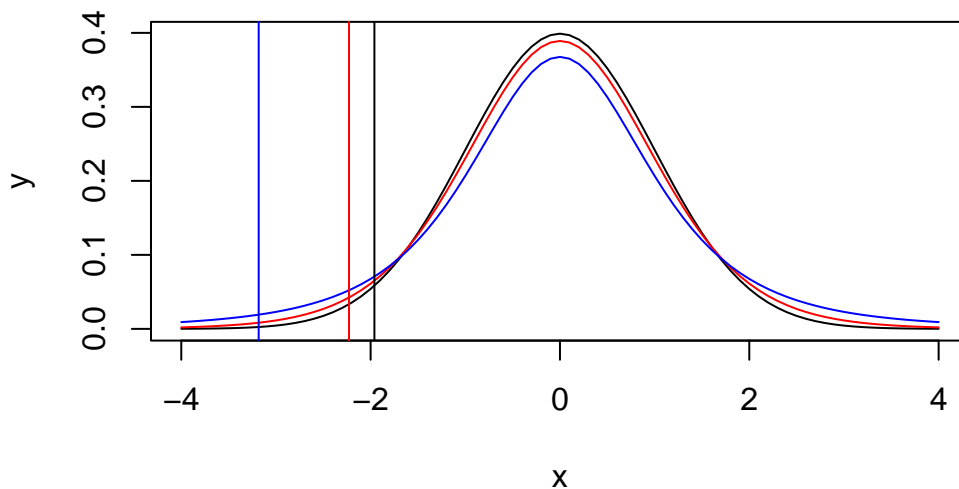
Det finns en enda z-fördelning, men t-fördelningar finns det många, och de ser olika ut beroende på antal frihetsgrader. Frihetsgrader är antalet mätningar, n, minus antalet parametrar som skattas, i vårt fall är det en parameter, medelvärdet.

**3.1** Rita följande graf:

```

x<-seq(-4, 4, 0.1)
y<- dnorm(x)
plot(x,y, type="l" )
lines(x, dt(x, 10), col="red" )
lines(x, dt(x, 3), col="blue" )
abline(v=qnorm(0.025))
abline(v=qt(0.025, 10), col="red")
abline(v=qt(0.025, 3), col="blue")

```



Grafen visar en z-fördelning och två t-fördelningar. En med 3 frihetsgrader och en med 10.

Jag har ritat en linje som utesluter 0.025 av arean åt vänster för varje fördelning.

**3.2** Viket/vilka av följande påståenden stämmer?

Arean till vänster om abline motsvarar p-värdet för ensidiga test.

**Det krävs en större avvikelse från  $\mu$  för att t-testen ska falla ut som signifikanta än för z-testet.**

Svar: Detta påstående är rätt. T-fördelningarna (speciellt med färre frihetsgrader, t.ex. 3) har bredare svansar än z-fördelningen, vilket innebär att kritiska värden för t-fördelningen ligger längre från medelvärdet än för z-fördelningen. Det innebär att för att ett t-test ska bli signifikant, krävs det en större avvikelse från medelvärdet än för ett z-test, särskilt när

antalet frihetsgrader är lågt.

### Hur ditt experiment påverkas av stickprovets storlek

**3.3** Kommentera skillnaden mellan z och t. När kan det vara rimligt att approximera t-testet med ett z-test?

Svar: Skillnaden mellan z och t: Z-test används när populationens standardavvikelse är känd och stickprovet är stort. T-test används när populationens standardavvikelse är okänd och uppskattas från stickprovet. T-fördelningen liknar normalfördelningen men har bredare svansar, särskilt vid små stickprov. Approximera t med z: Det kan vara rimligt att approximera t-testet med ett z-test när stickprovet är stort, eftersom t-fördelningen närmar sig normalfördelningen ju fler frihetsgrader man har (dvs. när stickprovsstorleken ökar).

En grupp av 32 kvinnor under speciell bevakning av mödravården föder barn. Medelvikten var 3200 g och standardavvikelsen 800 g. Avviker medelvikten från den normala medelvikten 3600 g?

**3.4** Ange nollhypotes och mothypotes

Svar: Nollhypotes ( $H_0$ ): Medelvikten för nyfödda i denna grupp avviker inte från den normala medelvikten på 3600 g.

Mothypotes ( $H_1$ ): Medelvikten för nyfödda i denna grupp avviker från den normala medelvikten.

**3.5** Välj signifikansnivå

$\alpha = 0.05$

**3.6** Beräkna t-test i R och ange p-värde samt om nollhypotesen förkastas eller behålls

Det finns inbyggda funktioner i R för att räkna ttest. Här skulle jag vilja att du räknar ut t från formeln  $t = \frac{\bar{X} - \mu}{\frac{SD}{\sqrt{n}}}$  och slår upp i t-tabellen `pt()` hur ovanligt det beräknade t-värdet är. Fundera på om du ska göra ett ensidigt eller tväsidigt test. Om du vill kan du beräkna ett z-test också och jämföra resultatet.

```
# Parametrar
x_bar <- 3200      # Stickprovets medelvärde
mu <- 3600        # Populationens medelvärde (enligt H0)
sd <- 800         # Stickprovets standardavvikelse
```

```

n <- 32                # Stickprovets storlek

# Beräkna t-värdet
t_value <- (x_bar - mu) / (sd / sqrt(n))

# Tvåsidigt p-värde baserat på t-värdet
p_value <- 2 * pt(abs(t_value), df = n - 1, lower.tail = FALSE)

# Skriv ut resultaten
print(paste("t-värde:", t_value))

```

```
[1] "t-värde: -2.82842712474619"
```

```
print(paste("p-värde:", p_value))
```

```
[1] "p-värde: 0.00812753887944206"
```

```

# Beslut om nollhypotesen förkastas
if (p_value < 0.05) {
  print("Nollhypotesen förkastas. Medelvikten avviker signifikant från den normala.")
} else {
  print("Nollhypotesen behålls. Det finns ingen signifikant skillnad i medelvikt.")
}

```

```
[1] "Nollhypotesen förkastas. Medelvikten avviker signifikant från den normala."
```

### 3.7 Formulera resultatet utan att använda statistisk terminologi.

Svar: Barnen i denna grupp verkar i genomsnitt väga mindre än vad som anses normalt. Skillnaden i vikt är tillräckligt stor för att vi ska misstänka att det inte är en slump att vikten avviker från den förväntade genomsnittsvikten.

## 4. Jämföra två stickprov gruppvis, populationens standardavvikelse är okänd

### 4.1 Skapa en data.frameGruppvis.Vikt med följande data.

```
data <- data.frame(
  vikt = c(90, 91, 93, 106, 97, 108, 97, 105,
           106, 103, 105, 96, 105, 95, 90, 101,
           90, 91, 85, 99, 93, 104, 89, 103,
           102, 95, 103, 95, 105, 87, 86, 101),
  grupp = c('A', 'A', 'A', 'A', 'A', 'A', 'A', 'A',
            'A', 'A', 'A', 'A', 'A', 'A', 'A', 'A',
            'B', 'B', 'B', 'B', 'B', 'B', 'B', 'B',
            'B', 'B', 'B', 'B', 'B', 'B', 'B', 'B')
)
```

I den första räkneuppgiften tänker vi oss att grupperna A och B är slumpmässigt utvalda ur två definierade populationer, kanske utifrån livsstil eller någon diagnos, som vi vill undersöka med avseende på vikt. Sammanlagt deltar 32 personer i studien. Har grupperna olika medelvikt?

#### 4.2 Ange nollhypotes och mothypotes

Svar: Nollhypotes ( $H_0$ ): Det finns ingen skillnad i medelviker mellan grupp A och grupp B.  
 Mothypotes ( $H_1$ ): Det finns en skillnad i medelviker mellan grupp A och grupp B.

#### 4.3 Välj signifikansnivå

Svar:  $\alpha = 0.05$

#### 4.4 Beräkna t-test, tex så här

```
attach(Gruppvis.Vikt)
```

```
t.test(vikt ~ grupp, var.equal = TRUE) #vikt ~grupp ska läsas: vikt förklarar av faktorn grupp
```

Tolka resultattabellen. Ange p-värde, samt om nollhypotesen förkastas eller behålls.

```
result <- t.test(vikt ~ grupp, data = data, var.equal = TRUE)

if (result$p.value < 0.05) {
  cat("Nollhypotesen förkastas.
  Medelvikten avviker signifikant från den normala. p-värde:",
  round(result$p.value, 4))
} else {
  cat("Nollhypotesen behålls.
  Det finns ingen signifikant skillnad i medelvikt. p-värde:",
```

```
round(result$p.value, 4))
}
```

Nollhypotesen behålls.

Det finns ingen signifikant skillnad i medelvikt. p-värde: 0.1221

#### 4.5 Formulera resultatet utan att använda statistisk terminologi

Svar: Efter att ha analyserat vikten hos personer i två olika grupper visar vår undersökning att även om det finns en observerad skillnad i genomsnittsvikt mellan grupperna, är denna skillnad inte statistiskt signifikant. Med ett p-värde på 0.1221 kan vi inte säkert säga att en grupp väger mer än den andra utifrån de data som samlats in. Detta innebär att eventuella skillnader i vikt som observerats potentiellt kan bero på slumpmässiga variationer.

#### Jämföra två stickprov där mätningarna hör ihop parvis, populationens standardavvikelse är okänd

En kraftfullare experimentell design när man arbetar med människor, där den individuella variationen kan vara mycket stor, är att göra upprepade mätningar på samma individ. Tänk dig att mätningen A är gjord före en tre veckors kontrollerad diet, och mätningen B är gjord efter. Nu tänker vi oss alltså att data från 4.1 avser mätningar på bara 16 personer.

Nollhypotes:

Det är krångligt att få till en stram formulering av nollhypotesen utan att skriva en formel.

Nollhypotesen för ett gruppvis test lyder: "Stickproven  $\bar{X}_A$  och  $\bar{X}_B$  är dragna ur samma population" Det kan illustreras med formeln  $A = B$ . Lägg märke till att det inte är samma sak som  $\bar{X}_A = \bar{X}_B$ . Även om stickproven är dragna ur samma population så tror vi inte att de ska ge exakt samma medelvärde. Det är den underliggande populationen som vi tror har samma medelvärde hela tiden.

I ett parvis test beräknas inte gruppernas medelvärde. Istället mäter man förändringen för varje individ,  $d$ , därefter beräknar man medelvärdet för alla förändringar,  $\bar{d}$ . Med hjälp av t-testet kan vi ställa  $\bar{d}$  i relation till variationen i mätningarna.  $\delta$  är den underliggande populationens sanna skillnad.

$$t = \frac{\bar{d} - \delta}{\frac{SD}{\sqrt{n}}}$$

Oftast är nollhypotesen att  $\delta = 0$

Så här brukar formeln se ut:



$$t = \frac{\bar{d}}{\frac{SD}{\sqrt{n}}}$$

n är i vårt exempel 16. Vid parvis test är n inte antalet mätningar, utan antalet par. Antalet frihetsgrader är n-1, antalet par minus ett.

Den goda strategin att räkna parvis t-test är att skapa en data.frame som visar att mätningarna hör ihop, enligt mottot en rad ett case, en kolumn en variabel.

#### 4.6 Läs in data från 4.1 formatterat på ett nytt sätt

```
# Parvis.Vikt<- read.table("clipboard", header=T)

Parvis.Vikt <- data.frame(
  A = c(90, 91, 93, 106, 97, 108, 97, 105,
        106, 103, 105, 96, 105, 95, 90, 101),
  B = c(90, 91, 85, 99, 93, 104, 89, 103,
        102, 95, 103, 95, 105, 87, 86, 101)
)
```

Det är även möjligt att återanvända din gamla data.frame genom att lägga till paried=T. Normalt känner man ju till i förväg om data hör ihop parvis eller ej och jag rekommenderar att ni använder strukturen i data.frame.

```
t.test(vikt ~ grupp, paired = T)
```

#### 4.7 Beräkna parvis ttest, kommentera varför resultatet blir annorlunda än vid gruppvis test

```
# Förutsatt att data frame 'Parvis.Vikt' redan är skapad
t_result <- t.test(Parvis.Vikt$A, Parvis.Vikt$B, paired = TRUE)

# Visa resultaten från det parvisa t-testet
print(t_result)
```

##### Paired t-test

```
data: Parvis.Vikt$A and Parvis.Vikt$B
t = 4.6967, df = 15, p-value = 0.0002866
alternative hypothesis: true mean difference is not equal to 0
95 percent confidence interval:
 2.048174 5.451826
```

```
sample estimates:
mean difference
      3.75
```

```
if (t_result$p.value < 0.05) {
  cat("Nollhypotesen förkastas.
  Medelvikten avviker signifikant från den normala. p-värde:", round(t_result$p.value, 4))
} else {
  cat("Nollhypotesen behålls.
  Det finns ingen signifikant skillnad i medelvikt. p-värde:", round(t_result$p.value, 4))
}
```

Nollhypotesen förkastas.

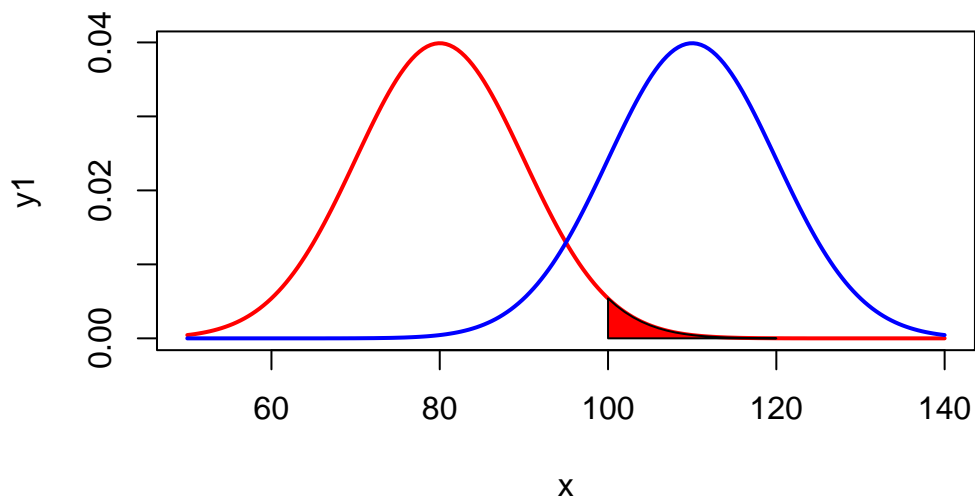
Medelvikten avviker signifikant från den normala. p-värde: 3e-04

Svar: Resultaten från ett parvis t-test och ett oberoende t-test kan skilja sig åt eftersom de hanterar datastrukturer olika. Ett parvis t-test tar hänsyn till beroenden mellan par av observationer, vilket minskar effekterna av individuell variation och kan öka den statistiska kraften. Däremot behandlar ett oberoende t-test observationerna som om de kommer från helt separata grupper, vilket kan leda till större inverkan av individuella skillnader på resultaten. Således kan ett parvis t-test avslöja signifikanta skillnader där ett oberoende t-test inte gör det, vilket speglar de olika antagandena och metodernas känslighet.

## 5. Illustrera typ I och typ II fel

Koden nedan ritar upp två normalfördelningar. Tänk er att de är stickprovsfördelningar och att man kan använda z-fördelningen därför att stickprovets storlek är stort. Den röda fördelningen,  $\mu_1$ , gäller om nollhypotesen är sann. Den blåa fördelningen,  $\mu_2$ , gäller om alternativhypotesen är sann. I den röda fördelningen har jag ritat in arean 0.025.

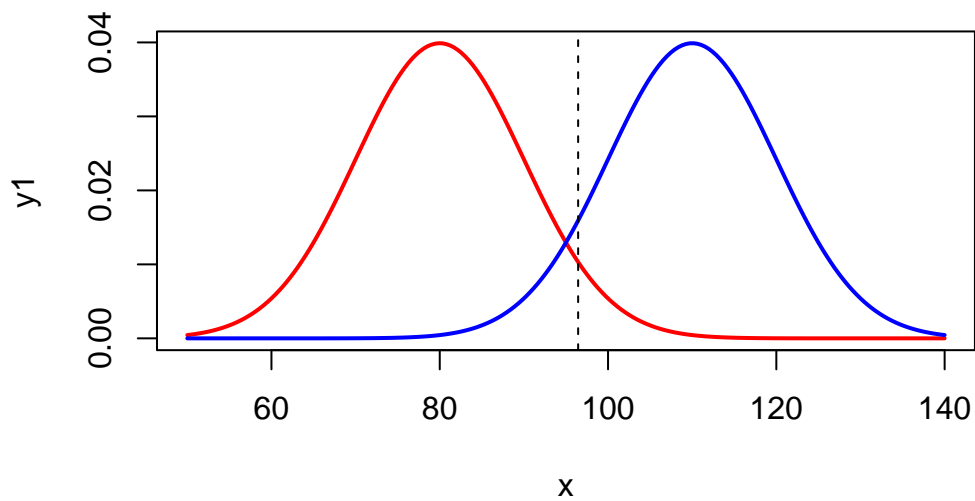
```
x=seq(50,140,length=200)
y1=dnorm(x,80, 10)
plot(x,y1,type='l',lwd=2,col='red')
y2=dnorm(x,110, 10)
lines(x,y2,type='l',lwd=2,col='blue')
cord.x1 <- c((round(qnorm(0.975, 80, 10, lower.tail=T))),
seq((round(qnorm(0.975, 80, 10, lower.tail=T))), 120,1),120)
cord.y1 <- c(0,dnorm(
  seq((round(qnorm(0.975, 80, 10, lower.tail=T))), 120, 1), 80, 10),0)
polygon(cord.x1,cord.y1,col='red')
```



Gör ett antal nya grafer:

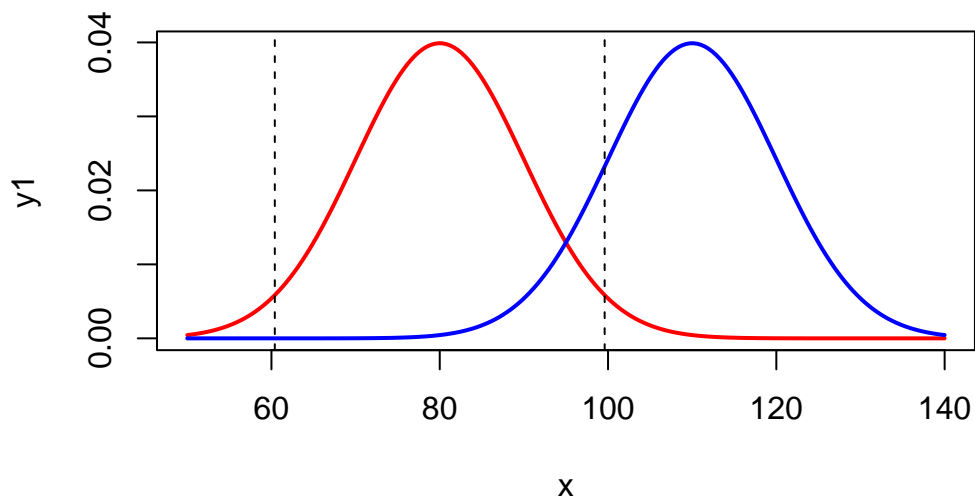
**5.1** Rita in en abline som visar gränsen utanför vilken du skulle förkasta nollhypotesen  $\mu_1 = \mu_2$  till förmån för alternativhypotesen  $\mu_1 < \mu_2$ . Signifikansnivån ska vara 0.05

```
x = seq(50, 140, length=200)
y1 = dnorm(x, 80, 10)
y2 = dnorm(x, 110, 10)
plot(x, y1, type='l', lwd=2, col='red')
lines(x, y2, type='l', lwd=2, col='blue')
abline(v=qnorm(0.95, 80, 10), col='black', lty=2)
```



**5.2** Rita in två abline som visar gränserna utanför vilka du skulle förkasta nollhypotesen  $\mu_1 = \mu_2$  till förmån för alternativhypotesen  $\mu_1 \neq \mu_2$ . Signifikansnivån ska vara 0.05

```
plot(x, y1, type='l', lwd=2, col='red')
lines(x, y2, type='l', lwd=2, col='blue')
abline(v=qnorm(0.025, 80, 10, lower.tail=TRUE), col='black', lty=2)
abline(v=qnorm(0.975, 80, 10, lower.tail=TRUE), col='black', lty=2)
```

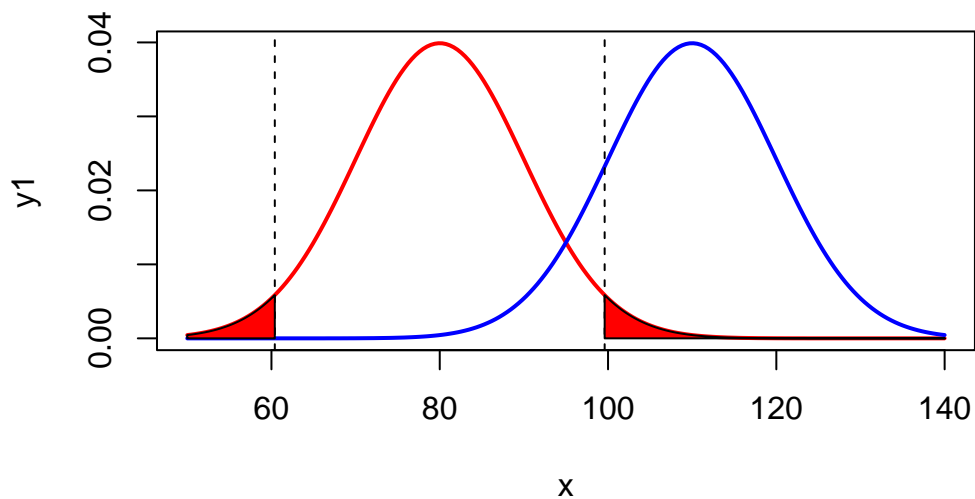


**5.3** Gör en ny 5.2 där du färgar arean som visar risken att förkasta en sann nollhypotes. Typ I felet.

```
plot(x, y1, type='l', lwd=2, col='red')
lines(x, y2, type='l', lwd=2, col='blue')
abline(v=qnorm(0.025, 80, 10, lower.tail=TRUE), col='black', lty=2)
abline(v=qnorm(0.975, 80, 10, lower.tail=TRUE), col='black', lty=2)

# Färga områden för Typ I fel
x_fill_left <- seq(min(x), qnorm(0.025, 80, 10), length.out=100)
x_fill_right <- seq(qnorm(0.975, 80, 10), max(x), length.out=100)
y_fill_left <- dnorm(x_fill_left, 80, 10)
y_fill_right <- dnorm(x_fill_right, 80, 10)

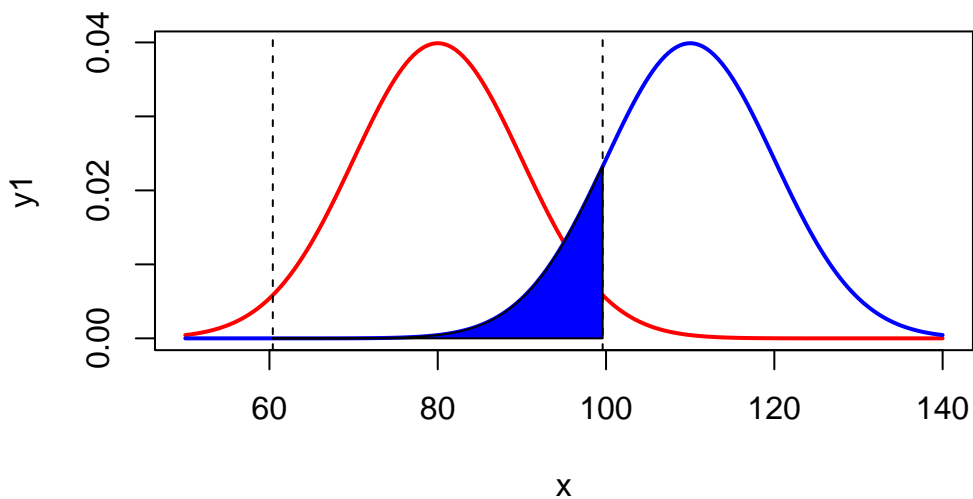
polygon(c(x_fill_left, rev(x_fill_left)),
c(rep(0, length(x_fill_left)), rev(y_fill_left)), col='red')
polygon(c(x_fill_right, rev(x_fill_right)),
c(rep(0, length(x_fill_right)), rev(y_fill_right)), col='red')
```



**5.4** Gör en ny 5.2. Vi tänker oss att alternativhypotesen,  $\mu_2$ , är sann. Färga en area som visar risken att behålla en falsk nollhypotes. Typ II felet, kallas även  $\beta$ . Beräkna och ange  $\beta$ .

```
plot(x, y1, type='l', lwd=2, col='red')
lines(x, y2, type='l', lwd=2, col='blue')
abline(v=qnorm(0.025, 80, 10, lower.tail=TRUE), col='black', lty=2)
abline(v=qnorm(0.975, 80, 10, lower.tail=TRUE), col='black', lty=2)

# Färga områden för Typ II fel ( $\beta$ )
x_fill_beta <- seq(qnorm(0.025, 80, 10), qnorm(0.975, 80, 10), length.out=100)
y_fill_beta <- dnorm(x_fill_beta, 110, 10)
polygon(c(x_fill_beta, rev(x_fill_beta)),
c(rep(0, length(x_fill_beta)), rev(y_fill_beta)), col='blue')
```

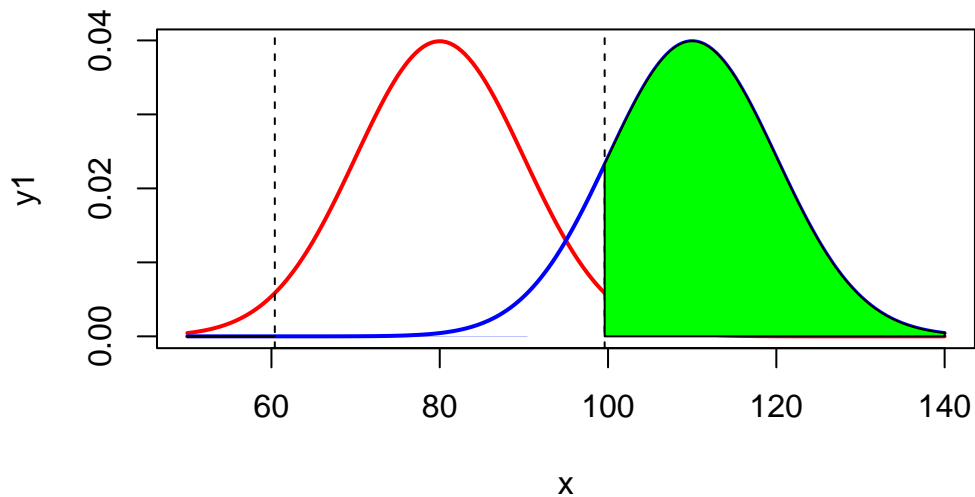


**5.5** Gör en ny 5.4 där du även färgar arean som visar sannolikheten att förkasta en falsk nollhypotes. Tips: Det ska bli  $1 - \alpha$ . Detta kallas även statistisk kraft, statistical power.

```
plot(x, y1, type='l', lwd=2, col='red')
lines(x, y2, type='l', lwd=2, col='blue')
abline(v=qnorm(0.025, 80, 10, lower.tail=TRUE), col='black', lty=2)
abline(v=qnorm(0.975, 80, 10, lower.tail=TRUE), col='black', lty=2)
polygon(c(qnorm(0.025, 80, 10), qnorm(0.025, 110, 10)), c(0, 0), col='blue', border=NA)
polygon(c(qnorm(0.975, 110, 10), max(x)), c(0, 0), col='blue', border=NA)

x_fill_power_left <- seq(min(x), qnorm(0.025, 80, 10), length.out=100)
x_fill_power_right <- seq(qnorm(0.975, 80, 10), max(x), length.out=100)
y_fill_power_left <- dnorm(x_fill_power_left, 110, 10)
y_fill_power_right <- dnorm(x_fill_power_right, 110, 10)

polygon(c(x_fill_power_left, rev(x_fill_power_left)),
c(rep(0, length(x_fill_power_left)), rev(y_fill_power_left)), col='green')
polygon(c(x_fill_power_right, rev(x_fill_power_right)),
c(rep(0, length(x_fill_power_right)), rev(y_fill_power_right)), col='green')
```



## 6. Ickeparametriskt alternativ till t-test

Data inspirerade av: DIFFERENTIATING DENGUE VIRUS INFECTION FROM SCRUB TYPHUS IN THAI ADULTS WITH FEVER, GEORGE WATT et al 2003

“[...] simple criteria to differentiate scrub typhus from dengue infection are needed [...], particularly where rapid confirmatory diagnostic tests are not available.”

Vita blodkroppar har räknats hos två patientgrupper som är svåra att skilja åt: Denguefeber, respektive scrub typhus. Använd en hypotestest-strategi för att avgöra om antal vita blodceller/ mm<sup>3</sup> kan ligga till grund för differentiell diagnos av scrub typhus och denguefeber. Man kan **inte** anta att mätvärdena är normalfördelade. Välj  $\alpha = 0.05$ .

| count | diagnosis |
|-------|-----------|
| 3000  | dengue    |
| 3200  | dengue    |
| 3500  | dengue    |
| 5068  | dengue    |
| 5679  | dengue    |
| 6200  | dengue    |
| 6300  | dengue    |
| 7020  | dengue    |
| 4400  | scrub     |
| 4500  | scrub     |



| count | diagnosis |
|-------|-----------|
| 5900  | scrub     |
| 6839  | scrub     |
| 7561  | scrub     |
| 9047  | scrub     |
| 12300 | scrub     |
| 14000 | scrub     |

## 6.1 Läs in data och beräkna ett two sample Wilcoxon test (ISwR 5.4, sid 103)

```
# Skapa data
vbc_counts <- c(3000, 3200, 3500, 5068, 5679, 6200, 6300, 7020,
               4400, 4500, 5900, 6839, 7561, 9047, 12300, 14000)
diagnosis <- c(rep("dengue", 8), rep("scrub", 8))

# Läs in data som en data frame
data <- data.frame(count = vbc_counts, diagnosis = diagnosis)

# Utför Wilcoxon test
wilcox_test <- wilcox.test(count ~ diagnosis, data = data, exact = FALSE)

# Visa resultatet av testet
print(wilcox_test)
```

Wilcoxon rank sum test with continuity correction

```
data: count by diagnosis
W = 14, p-value = 0.06608
alternative hypothesis: true location shift is not equal to 0
```

## 6.2 Hur formuleras nollhypotes och alternativhypotes i Wilcoxon testet?

Nollhypotes ( $H_0$ ): Medianen av vita blodkroppar är lika mellan patienter med dengue och scrub typhus. Det innebär att det inte finns någon skillnad mellan de två gruppernas medianvärden.

Alternativhypotes ( $H_1$ ): Medianen av vita blodkroppar skiljer sig mellan patienter med dengue och scrub typhus.

### 6.3 Visar testet att en enkel räkning av vita blodkroppar kan stå till grund för differentiell diagnos?

Eftersom p-värdet är större än signifikansnivån kan vi inte förkasta nollhypotesen. Detta innebär att det inte finns tillräckliga bevis i de data som samlats för att hävda att det finns en statistiskt signifikant skillnad i medianantalet vita blodkroppar mellan de två sjukdomsgrupperna.

### 6.4 Ange nollhypotes och mothypotes för t-testet

Nollhypotes ( $H_0$ ): Medelvärdet av vita blodkroppar är lika mellan de två diagnosgrupperna.  
Alternativhypotes ( $H_1$ ): Medelvärdet av vita blodkroppar skiljer sig mellan de två diagnosgrupperna.

### 6.5 Beräkna testet, tolka resultatet, formulera slutsats.

```
# Utför t-test
t_test <- t.test(count ~ diagnosis, data = data)

# Visa resultatet av t-testet
print(t_test)
```

Welch Two Sample t-test

```
data: count by diagnosis
t = -2.256, df = 9.6671, p-value = 0.04856
alternative hypothesis: true difference in means between group dengue and group scrub is not
95 percent confidence interval:
 -6121.34819 -23.65181
sample estimates:
mean in group dengue mean in group scrub
      4995.875      8068.375
```

P-värdet är mindre än signifikansnivån och således så förkastas nollhypotesen.

### 6.6 Varför blir det olika resultat i 6.3 och 6.5?

Svar: Skillnaderna i resultaten kan bero på att de två testerna hanterar data på olika sätt. Wilcoxontestet är mindre känsligt för extremvärden och icke-normal datafördelning, vilket kan

leda till olika slutsatser jämfört med t-testet som antar normalfördelning och kan påverkas mer av outlier-värden och ojämn fördelning av data. Om data inte är normalfördelade, kan t-testets antaganden vara brutna, vilket ger missvisande resultat.

## 7. Konfidensintervall som metod för hypotestestning

I resultattabellerna som t.test skapat finner du även konfidensintervall. Att beräkna konfidensintervall är mycket användbart. Både för att illustrera hur bra en skattning av en parameter är och för resonemang som liknar hypotestest.

I ett parvis t-test undersöker vi om  $\delta \neq 0$ .

Vi kan skatta  $\delta$  med  $\bar{d}$  och titta efter om intervallet  $\bar{d} \pm 95\% \text{ CI}$ .

Ni har mött tekniken hur man räknar konfidensintervall för medelvärden i inlämningsuppgift 3, i beräkning av ett blodtryckintervall som gäller 95% av befolkningen. Här vill jag poängtera att beräkningen blir lite olika beroende på om populationens standardavvikelse  $\sigma$ , är känd, eller om man måste skatta den utifrån ett stickprovs standardavvikelse, SD. Var vänlig slå upp formlerna i en bok. När  $\sigma$  är känd går det att använda z.

**7.1** En däckfabrikör mäter 10 däck som körts 50000 miles. I medel är 0.32 tum av mönstret kvar, med standardavvikelsen 0.09 tum. Beräkna ett 95% konfidensintervall kring medlet.

```
# Data för däckmönster
n <- 10
mean_x <- 0.32
sd_x <- 0.09

# Beräkna konfidensintervall
t_critical <- qt(0.975, df = n-1) # 95% CI, tvåsidigt, n-1 frihetsgrader
error_margin <- t_critical * (sd_x / sqrt(n))

# Konfidensintervall
ci_lower <- mean_x - error_margin
ci_upper <- mean_x + error_margin

# Skriv ut resultaten
cat("95% konfidensintervall för däckmönster är mellan", ci_lower, "och", ci_upper, "\n")
```

95% konfidensintervall för däckmönster är mellan 0.2556179 och 0.3843821

**7.2** Kan fabrikören påstå att 0.30 tum brukar vara kvar efter 50000 miles utifrån mätningen?

```

claimed_value <- 0.30

# Kontrollera om påstått värde ligger inom konfidensintervallet
if (claimed_value >= ci_lower && claimed_value <= ci_upper) {
  cat("Fabrikören kan påstå att 0.30 tum brukar vara kvar efter 50000 miles,
    eftersom detta värde ligger inom 95% konfidensintervallet [", ci_lower, ",", ci_upper, "].\n")
} else {
  cat("Fabrikören kan inte påstå att 0.30 tum brukar vara kvar efter 50000 miles,
    eftersom detta värde inte ligger inom 95% konfidensintervallet [", ci_lower, ",", ci_upper, "].\n")
}

```

Fabrikören kan påstå att 0.30 tum brukar vara kvar efter 50000 miles,  
eftersom detta värde ligger inom 95% konfidensintervallet [ 0.2556179 , 0.3843821 ].

**7.3** Tio slumpmässigt utvalda individer har ett medelkolesterolvärde på 5,4 mmol/l och standardavvikelsen 0.5 mmol/l. Beräkna ett 95% konfidensintervall för medlet.

```

# Data för kolesterol
n <- 10
mean_k <- 5.4
sd_k <- 0.5

# Beräkna konfidensintervall
t_critical <- qt(0.975, df = n-1) # 95% CI, tvåsidigt, n-1 frihetsgrader
error_margin <- t_critical * (sd_k / sqrt(n))

# Konfidensintervall
ci_lower_k <- mean_k - error_margin
ci_upper_k <- mean_k + error_margin

# Skriv ut resultaten
cat("95% konfidensintervall för medelkolesterolvärde är mellan",
  ci_lower_k, "och", ci_upper_k, "\n")

```

95% konfidensintervall för medelkolesterolvärde är mellan 5.042322 och 5.757678

## 8 ANOVA – ANalysisOfVariance

I denna övning nöjer vi oss med en enklare användning av ANOVA. Vi betraktar ANOVA som ett sätt att hantera flera simultana t-test.

Nollhypotes:  $\mu_1 = \mu_2 = \mu_3$

Alternativhypotes:  $\mu_1 = \mu_2 = \mu_3$  gäller ej

Vi ställer upp ett antal stickprov och kontrollerar om någon grupp avviker från de andra.

**8.1Läs in datasetet med blodtryck hos tre olika grupper av försökspersoner och beräkna ANOVA** med funktionen `anova(lm(blodtryck ~behandling))`

| blodtryck | behandling   |
|-----------|--------------|
| 77        | kontroll     |
| 77        | kontroll     |
| 78        | kontroll     |
| 80        | kontroll     |
| 81        | kontroll     |
| 89        | kontroll     |
| 90        | kontroll     |
| 96        | kontroll     |
| 99        | kontroll     |
| 107       | kontroll     |
| 59        | medicinering |
| 66        | medicinering |
| 70        | medicinering |
| 73        | medicinering |
| 76        | medicinering |
| 77        | medicinering |
| 78        | medicinering |
| 81        | medicinering |
| 81        | medicinering |
| 91        | medicinering |
| 64        | träning      |
| 66        | träning      |
| 69        | träning      |
| 72        | träning      |
| 73        | träning      |
| 74        | träning      |
| 74        | träning      |
| 80        | träning      |
| 84        | träning      |
| 99        | träning      |

```
# Skapa datasetet
data <- data.frame(
  blodtryck = c(77, 77, 78, 80, 81, 89, 90, 96, 99, 107,
```

```

59, 66, 70, 73, 76, 77, 78, 81, 81, 91,
64, 66, 69, 72, 73, 74, 74, 80, 84, 99),
  behandling = c(rep("kontroll", 10), rep("medicinering", 10), rep("träning", 10))
)

# Utför ANOVA
anova_resultat <- aov(blodtryck ~ behandling, data = data)
summary(anova_resultat)

```

```

              Df Sum Sq Mean Sq F value Pr(>F)
behandling    2  968.5    484.2    4.948 0.0148 *
Residuals   27 2642.5     97.9
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

**8.2** ANOVA räknar ut ett F värde utifrån hur variationen är fördelad över matrisen. Sök efter värdet  $\text{Pr}(>F)$  i tabellen, sannolikheten att få värdet F eller större under förutsättning att nollhypotesen stämmer. Om denna sannolikhet är mindre än ditt  $\alpha$ -värde kan du förkasta nollhypotesen.

Svar:  $\text{Pr}(>F) = 0.0148$

Således kan nollhypotesen förkastas.

### 8.3 Välj ett sätt att visualisera data.

```

ggplot(data, aes(x = behandling, y = blodtryck, fill = behandling)) +
  geom_boxplot() +
  labs(title = "Blodtryck efter behandlingstyp",
       x = "Behandlingskategori",
       y = "Blodtryck") +
  theme_minimal()

```

