**Kommandoer i RStudio**

Data frame ved navn myData som indeholder både x og y variabler: myData$x for at se x-variablen. myData$y for at se y-variablen.

Indhold

[Standarddeviation (SD): 2](#_Toc25258642)

[Gennemsnit: 2](#_Toc25258643)

[Median: 2](#_Toc25258644)

[Histogram: 2](#_Toc25258645)

[Boxplot 2](#_Toc25258646)

[Navnene for variablerne I data frame’en 2](#_Toc25258647)

[IQR (interquartile range) 2](#_Toc25258648)

[Tabel - opsummeringstabel af frekvenserne (fx **hvor mange** der er gift, ikke gift, skilt): 2](#_Toc25258649)

[Kvantil/Kvartil 2](#_Toc25258650)

[Udtrække en delmængde af et datasæt 2](#_Toc25258651)

[Simpel lineær regression 2](#_Toc25258652)

[Se resultatet (sd, r^2 osv) af vores lineære regressionsmodel som hvordan den plottes 3](#_Toc25258653)

[Plot vores lineære regressionsmodel 3](#_Toc25258654)

[Referenceinterval 3](#_Toc25258655)

[Konfidensinterval 3](#_Toc25258656)

[Uparret t-test 4](#_Toc25258657)

[Parret t-test 4](#_Toc25258658)

# Standarddeviation (SD):

sd(myData$y)

# Gennemsnit:

mean(myData$y)

Median:
median(myData$x)

# Histogram:

hist(myData$x)

# Boxplot

boxplot(myData$x)

# Navnene for variablerne I data frame’en

Names(myData)

# IQR (interquartile range)

IQR(myData). Bruges sammen med median hvis obs. Ikke er normalfordelte. Hvis normalfordelte: brug spredning (sd) og middelværdi. SE OGSÅ KVARTIL/KVANTIL og Summary. Undersøg ved at tegne et histogram om det er normalfordelt/symmetrisk.

# Tabel - opsummeringstabel af frekvenserne (fx **hvor mange** der er gift, ikke gift, skilt):

table(myData$marital.stat)

# Kvantil/Kvartil

Quantile. Summary kan måske også bruges.

# Udtrække en delmængde af et datasæt

Subset.

Fx: dvisitYes <- subset(dvisit, dental.visit == "Yes")

# Simpel lineær regression

model=**lm**(myData$y~myData$x). Husk at skrive y først. Dette betyder at y afhænger af x. lm står for lineær model. Kommandoen er lm. ~ =afhænger af. Dernæst bruges summary(model) for at aflæse b (skæring med y-aksen (intercept)) samt hældningskoefficienten. Se næste punkt nedenfor.

# Se resultatet (sd, r^2 osv) af vores lineære regressionsmodel som hvordan den plottes

summary(model). Model er defineret ovenfor: **lm**(myData$y~myData$x). “Model” er et navn vi selv har valgt.
Residual standardfejl: Hvor meget linjen passer til vores observationer (gennemsnitlige afstand mellem obs. og linjen).
**Intercept: skæring med y aksen, tallet nedenunder tallet for skæring med y-aksen er hældningskoefficienten. Disse to tal bruges til at plotte vores lineære regressionsmodel i formen y=b+ax.**

# Plot vores lineære regressionsmodel

Først skal denne kommando være kørt: lm, se under **Simpel lineær regression**Dernæst plottes punkterne: plot(myData$x,myData$y)for at se scatterplottet før vores regression.
Derefter plottes den lige linje som resultat af vores lineære regression: **abline**(tallet for skæring med y-aksen , tallet for hældningskoefficient). Se **Se resultatet (sd, r^2 osv) af vores lineære model.**
fx abline(0.06319,1.88868).

Fx:

model=lm(myData$y~myData$x) Først laves lineære regression. (lm for linear model)

summary(model) Så findes b ved “Intercept” og dernæst tallet under som er hældningen

plot(myData$x,myData$y) Scatterplottet skal plottes først

abline(0.06319,1.88868) Den lige linje plottes. Husk rækkefølgen: abline(skæring med y-aksen,hældning)

# Referenceinterval

95% referenceinterval=middelværdi ± 2·spredning

# Konfidensinterval

95% konfidensinterval= middelværdien ± 2· SEM (standardfejlen af middelværdien).

Brug kommandoen t.test eller udregn med hånden. $SEM=\frac{SD}{√n}$

95% konf. interval beskriver usikkerheden på vores gennemsnitlige alder, mens 95% referenceintervallet beskriver grænserne som 95% af pigerne vil ligge indenfor.

Ved regressionsmodeller (efter brug af lm-kommandoen) kan kommandoen ”confint” bruges til at finde konf. Intervallerne for skæring i y-aksen og hældningen. Fx ”confint(model)”.

Ved 2x2 tabel (binær x binær) bruges prop.test. ved den relative risiko, fx:

**prop.test**(148, 148 **+** 423)

Ved odds-ration bruges fisher.test(model). Model er definition af din 2x2 tabel, se hjemmeopgave 6 for hvordan det definieres.

# Uparret t-test

Bruges for sammenligning af 2 forskellige grupper.

t.test(periodontitis$IgG~periodontitis$group)

Her bruges ~ imellem de 2 grupper.

# Parret t-test

Bruges for samme forsøgspersoner, fx ved sammenligning før og efter medicinbehandling.

t.test(rotte$tape-rotte$sutur)

Her bruges – imellem de 2 grupper.

Forskellen bliver udregnet og derfor bruges et minustegn.