

1.1 Datapræsentation

1. Opgavenummer 6
2. Tabel 1

n=125	
Colorchange Mean (SD)	23,44 (13,37)
Concentration Mean (SD)	7.9 (4.2)
Light n (%)	72 (57,6)
Erosion n (%)	61 (48,8)
Satisfaction n (%)	
-Vil aldrig bruge produktet igen	37 (29,6)
-Vil måske bruge produktet igen	49 (39,2)
-Vil helt sikkert bruge produktet igen	39 (31,2)

Da Colorchange er en kontinuert variabel kan den enten beskrives gennem middelværdi og spredning eller gennem median og IQR. For at afgøre dette skal observationerne vurderes om de er normalfordelte, og dette kan gøres ved at tegne et histogram. Umiddelbart ser det normalfordelt ud, men hvis man vil lave et ekstra test kan man lave et QQ-plot. Den rette linje i QQ-plottet ser ud til at dække de fleste observationer og jeg kan tillade mig at beskrive Colorchange vha. middelværdi og standard deviation. Det samme gælder Concentration da det også er en kontinuert variabel.

Light og Erosion er binære variabler og skal derfor præsenteres gennem andel og relative frekvenser. Satisfaction er en kategorisk variabel og vil derfor også beskrives gennem andel og relative frekvenser.

1.2 Overordnet effekt af tandblegningsproduktet

1. Gennem t.test kan man undersøge om værdierne for colorchange er lig med 0 (nulhypotese). Hvis de er lig med 0 er der ingen forskel i tandfarven efter behandling sammenlignet med før behandling. Da det er observationer fra den samme population laves der et parret t.test.
T.testen viser at den gennemsnitlige tandfarve har ændret sig med 23.44 i lightness skalaen (konfidensintervaller og p-værdi beskrives i de næste spørgsmål).

2. P-værdien for nulhypotesen er uhyrlig lille ($p=2.2e-16$). Nulhypotesen forkastes, og der kan konstateres, at tandfarven har ændret sig signifikant, og at farveændringen ikke var et tilfældighed.
3. 95% CI= [21.07;25.81]. Vi kan med 95% sikkerhed konkludere at tandfarven vil ændre sig ved blegning med en forskel i lightness skalaen der svarer til en værdi i dette interval i den sande population. Tandblegningsproduktet har bleget tænderne, dette kan ses ud fra p-værdien og konfidensintervallet.
4. 95% referenceintervallet [-3.3;50.18]. Intervallet fortæller os om den øvre og nedre grænse for tandfarveændringen i 95% af populationen. Det fortæller os at nogle har haft en negativ effekt af produktet fordi de har fået mørkere tænder af at bruge det, men hos de fleste har produktet effektivt bleget tænderne.
5. I alle stikprøver vil referenceintervallet være bredere end konfidensintervallet, og det er grundet forskellen i hvordan de to intervaller beregnes. Ved referenceintervallet bruges spredning mens der ved konfidensintervallet bruges standardfejlen af middelværdien (SEM-standard error of the mean) til at udregne intervallet. For at udregne SEM skal spredningen divideres med kvadratroden af antal observationer og derfor får man et mindre tal end ved bare at bruge spredningen alene som man gør ved referenceintervallet.

Her ses hvordan de to intervaller udregnes:

Konfidensintervallet= middelværdien $\pm 2 * SEM$

Referenceintervallet= middelværdien $\pm 2 * SD$

Som før nævnt er SEM altid mindre end standard deviationen (SD) da:

$$SEM = \frac{SD}{\sqrt{n}}$$

1.3 Effekt af koncentration og lys på ændring i tandfarve

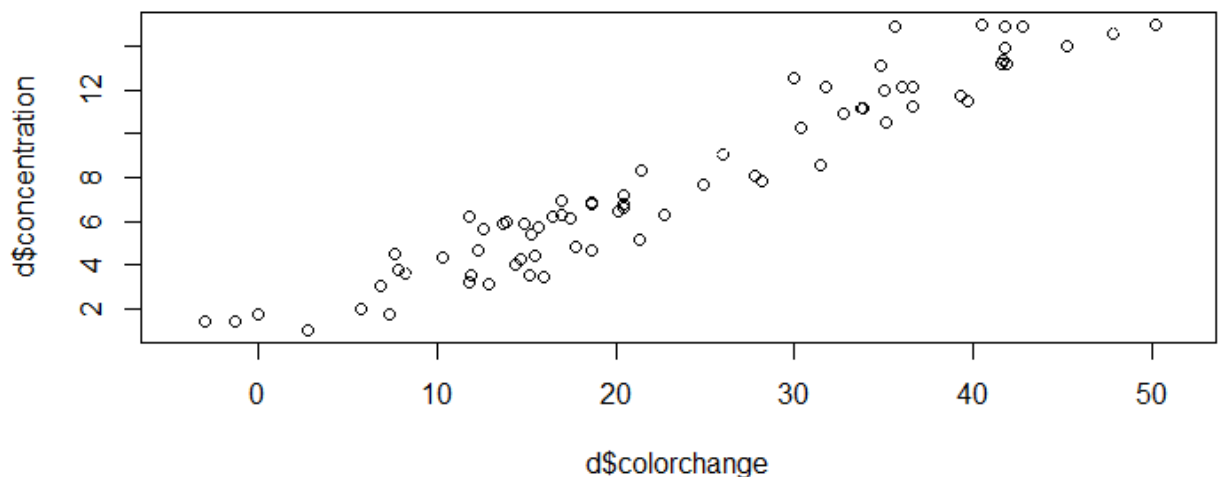
1. Da den afhængige variabel er kontinuert (ændring i tandfarve) og de uafhængige variable er henholdsvis kontinuert (koncentration) og binær (lys) vil der blive brugt multipel lineær regression.
2. Colourchange= $a+b_1*concentration+e$
og med et indikatorvariabel for lys
Colourchange= $a+b_1*concentration+b_2+e$

Colourchange= $-0.74+b_1*3.05+e$
Colourchange= $(-0.74-0.005)+b_1*3.05+e$, med lys
3. Brugen af lys har ingen signifikant betydning for sammenhængen mellem tandfarveændringen og koncentrationen, da $p=0.994$. Nulhypotesen tager udgangspunkt i at brugen af lys ingen effekt har på sammenhængen mellem tandfarveændringen og koncentrationen, og denne hypotese accepteres.
4. Da vi kan konkludere at lys ingen betydelig effekt har, kan vi da bare undlade den i modellen, og således går vi fra at have en multipel lineær regressionsmodel til at have en

simpel lineær regressionsmodel.

Modellen vil nu være således:

5. $\text{Colourchange} = -0.74 + b_1 * 3.05 + e$
 6. Intercept er skæring i y-aksen, og i dette tilfælde er den næsten 0. Ved koncentrationen 0 er der samtidig ingen farveændring.
 7. Nulhypotesen tester for om skæring i y-aksen er lig med 0, med andre ord om der er en farveændring når koncentrationen er lig med 0. P-værdien for intercept er lig med 0.326 og der er derfor statistisk evidens for at den sagtens kan være lig med 0.
 8. 95% CI for skæring i y-aksen = $[-2.23; 0.75]$. Her kan vi også se at 0 er medinddraget i konfidensintervallet og derved bekræfter det vores p-værdi i spørgsmål 7. 95% CI for hældningen = $[2.89; 3.22]$. Dette forklarer at der er en sammenhæng mellem koncentration og farveændringen, da 0 ikke er medinddraget i intervallet.
 9. Nu indsættes koncentrationen 12% i vores model:
 $\text{Colourchange} = -0.74 + 12 * 3.05 + e$
 Farveændringen er på 35.86. Vi ser at det stemmer overens med vores scatterplot nedenfor.
- Fortolkningen er, at ved brug af 12% koncentration hydrogenperoxid vil man gennemsnitligt få en farveændring der svarer til 35.86 i "lightness" skalaen.
10. 95% Referenceintervallet for 10% koncentration $[21.93; 37.6]$. 95% af dem som bruger 10% koncentration vil have en farveændring svarende til en værdi i dette interval.



1.4 Sammenhæng mellem ætsning og brugertilfredshed

1. Da der er tale om en binær (erosion) og en kategorisk (brugertilfredshed) variabler kan χ^2 test anvendes.
2. Der laves en tabel for at få et overblik.

	Vil aldrig bruge produktet igen	Vil måske bruge produktet igen	Vil helt sikkert bruge produktet igen
Erosion	31	24	6
Ingen erosion	6	25	33

Der laves en χ^2 test og p-værdi=1.9e-8. Der er en signifikant forskel i tilfredsheden mellem gruppen der fik erosioner og gruppen der ikke fik erosioner. Nulhypotesen som p-værdien tilhører tager udgangspunkt i at begge grupper vil have den samme tilfredshed.

3. Sandsynligheden for helt sikkert at bruge produktet igen er på 9,8%. 95% CI= [0.04;0.21] Vi er 95% sikre på at den sande sandsynlighed i populationen vil ligge et sted mellem 4% og 21%.
4. Den estimerede risiko for aldrig at bruge produktet igen hos dem der fik erosion var på 50,8% mens dem der *ikke* fik erosion var på 9,4%.
5. Odds-ratioen er på 1.012. Det vil sige at der er 1.2% højere odds for måske at bruge produktet igen hos dem der havde erosion sammenlignet med dem der ikke havde erosion. Da dette er et usikkert estimat vil man normalt supplere dette med et 95% konfidensinterval som nok vil have 1 i intervallet, hvilket betyder at der ikke er forskel mellem de to grupper da odds-ratioen viser at grupperne næsten er ens.

1.5 Sammenhæng mellem koncentration og ætsning

1. Da den afhængige variabel er binær (erosion, ingen erosion) og den uafhængige variabel er kontinuert (concentration), vil der her bruges logistisk regression.
2. $\log\text{Odds(erosion)}=a+b*\text{concentration}$
 $\log\text{Odds(erosion)}=-4,27+0,54*\text{concentration}$
3. Odds-ratio = 1,72. Odds for at få erosion er 1,72 gange højere end for ikke at få erosion når man tager koncentrationen af hydrogenperoxid i betragtning. 95% CI=[1.47;2.1]. Vi er 95% sikre på at odds-ratioen i den sande population ligger et sted mellem 1,47 og 2,1.
4. Jeg er uenig i dette udsagn. Der er statistisk evidens for at koncentrationen af hydrogenperoxid har en sammenhæng med forekomsten af erosion (p-værdi=2,32e-9). Nulhypotesen tager udgangspunkt i at der ingen sammenhæng er mellem koncentration og erosion, og i det tilfælde skal hældningen være lig med 0. Nulhypotesen forkastes da p-værdien forklarer at den estimerede hældning ikke er en tilfældighed.

5. Den prædikterede odds er 0.357. Hver gang der er en der ikke får erosion er der 0.357 der får erosion når der bruges 6% hydrogenperoxid.
6. Risikoen for erosion ved brug af 6% hydrogenperoxid er på 26,3%.
- 7.

2. Læsning og forståelse af videnskabelige artikel

1. delta L beskriver hvor hvid tanden er (lysintensiteten), og når tænderne ligger i te i 7 dage vil man forvente et negativt delta L da tænderne bliver misfarvet og mørkere når de ligger i te sammenlignet med før forsøget.
2. 95% CI = [-6.94;-9.76]. Ændringen i lysintensiteten for tænderne i den sande population efter at tænderne har ligget i te i 7 dage vil i lightness skalaen svare til en ændring svarende til en værdi i intervallet -6.94 og -9.76. Denne konfidensinterval er baseret på gruppen for produktet CR før blegning/brug af produktet.
3. Nulhypotesen undersøger om der har været en signifikant forskel i farveændringen mellem de respektive tænder, altså om alle tænder har samme farve efter at de har været i te i 7 dage. Her vil man gerne have at der ikke er en signifikant forskel, da alle tænder i alle grupper skal have samme farve fra udgangspunkt for at kunne sammenligne blegningseffektiviteten af de forskellige produkter.
4. Alle tænder i alle grupper ligner stort set hinanden farvemæssigt, og det skal de gøre som udgangspunkt for senere at kunne undersøge blegningseffektiviteten af produkterne.
5. Efter behandling med produktet CR er tænderne blevet bleget og produktet har virket. Vi ser at farveændringen er på 12.42 med SD=4.47. Samtidig ses at farveforskellen har ændret sig signifikant fra Time 2 til Time 3 i CR gruppen, mens den fra Time 1 og Time 2 har været et mindre forskel i blegningen som ikke var signifikant.
6. Parret t-test da det er den samme gruppe jeg sammenligner med. Her er det 2 middelværdier fra samme gruppe (Time 2 og Time 3) hvor der testes om der er en signifikant forskel i middelværdierne.
7. Det er ikke muligt at konkludere om de fire produkter er bedre til at blege end gruppe AS da delta a kun forklarer forskellen efter behandling i røde og grønne nuancer og ikke i lysintensitet. Men produkterne har i hvert fald haft en effekt på tandfarven. Desuden er gruppe AS en "negativ kontrol" gruppe og blev ikke behandlet med et blegeprodukt.
8. Ja det tror jeg da der ud fra figur 2 ses en stor forskel mellem gruppe AS og de andre grupper til Time 3.
9. Jeg er ikke enig i dette påstand. Hvis man kigger på tabel 3 under produktet OP i Time 3 har de selv angivet at OP er statistisk signifikant anderledes i farveforskellen ved farveparameteren delta E. Ved figur 3 ses produktet OP også at have medført en stor farveforskel sammenlignet med de andre produkter på farveparameteren delta b.

3. Generelle spørgsmål

1. a

2. a,b,c,e

3.

- a: 8%.
- b: 2%

Appendiks

1.1 Datapræsentation

```
d <- read.csv("http://causal.sund.ku.dk/e19/data6.csv", header=TRUE)
```

```
hist(d$colorchange)
```

```
qqnorm(d$colorchange)
```

```
qqline(d$colorchange)
```

```
hist(d$concentration)
```

```
qqnorm(d$concentration)
```

```
qqline(d$concentration)
```

```
mean(d$colorchange)
```

```
sd(d$colorchange)
```

```
mean(d$concentration)
```

```
sd(d$concentration)
```

```
table(d$light)
```

```
72/(72+53)
```

```
table(d$erosion)
```

```
61/(61+64)
```

```
table(d$satisfaction)
```

```
37/125
```

49/125

39/125

1.2 Overordnet effekt af tandblegningsproduktet

```
t.test(d$colorchange,mu=0)
```

```
mean(d$colorchange)
```

```
sd(d$colorchange)
```

```
23.44+2*13.37
```

```
23.44-2*13.37
```

1.3 Effekt af koncentration og lys på ændring i tandfarve

```
plot(d$colorchange,d$concentration,col=d$light)
```

```
model=lm(d$colorchange~d$concentration+d$light)
```

```
summary(model)
```

```
simpelmodel=lm(d$colorchange~d$concentration)
```

```
summary(simpelmodel)
```

```
confint(simpelmodel)
```

```
-0.74+(12*3.05)
```

```
-0.74+(10*3.05)
```

```
29.76+2*3.916
```

```
29.76-2*3.916
```

1.4 Sammenhæng mellem ætsning og brugertilfredshed

```
erosion=subset(d,erosion=="1")
```

```
table(erosion$satisfaction)
```

```
noerosion=subset(d,erosion=="0")
```

```
table(noerosion$satisfaction)
```

```
spg.1.4<-matrix(c(31, 6, 24, 25, 6, 33), nrow = 2, ncol = 3)
```

chisq.test(spg.1.4)

6/(31+24+6)

prop.test(6,61)

31/(31+24+6)

6/(6+25+33)

24/(31+6)

25/(33+6)

0.6486/0.641

1.5 Sammenhæng mellem koncentration og ætsning

logr=glm(d\$erosion ~ d\$concentration, family = binomial)

summary(logr)

exp(0.54261)

confint(logr)

exp(0.3827)

exp(0.7429)

-4.27+(0.54*6)

exp(-1.03)

exp(-4.272+(0.54*6))/(1+exp(-4.272+(0.54*6)))

0.8/(exp(-4.272+0.54)/(1+exp(-4.272+0.54)))

exp(-4.272+(0.54*34.21))/(1+exp(-4.272+(0.54*34.21)))

2. Læsning og forståelse af videnskabelige artikel

2.23/sqrt(10)

-8.35+2*0.705

-8.35-2*0.705

3. Generelle spørgsmål

10/1.25

10-8