Opgavesættet indeholder fem opgaver. Opgaverne vægter ikke éns, hvorfor den procentvise vægt er angivet ved hver opgave. Der lægges vægt på, at det af besvarelsen tydeligt fremgår hvordan man kommer frem til de forskellige resultater.

Held og lykke!

Opgave 1

Denne opgave vægter med 12%

Lad $X_1, \ldots, X_n$ være i.i.d. stokastiske variable med værdier i $\{1, 2, 3\}$,

$$P(X_i = 1) = \theta, \quad P(X_i = 2) = 2\theta, \quad P(X_i = 3) = 1 - 3\theta,$$

hvor θ er en ukendt parameter, $0 \leq \theta \leq 1/3$.

(a) Brug moment metoden til at udlede en estimator for θ.

Vi har kun brug for $r = 1$ teoretisk moment, nemlig $\mathrm{E}[X^1]$. Da X er diskret og kun tager værdier i $\{1, 2, 3\}$ er

$$\begin{aligned}\mathrm{E}[X] &= \sum_{k=1}^{3} k P(X = k) \\ &= 1 \cdot \theta + 2 \cdot 2\theta + 3 \cdot (1 - 3\theta) \\ &= 3 - 4\theta.\end{aligned}$$

Det teoretiske moment sættes lig med det tilsvarende stikprøvemoment, $\frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$. Det giver

$$\begin{aligned} 3 - 4\hat{\theta} &= \bar{X} \\ \implies \hat{\theta} &= \frac{3 - \bar{X}}{4}.\end{aligned}$$

(b) Er din estimator unbiased? Giv en kort begrundelse for dit svar.

Ja, estimatoren er unbiased, dvs., $\mathrm{E}\,\hat{\theta} = \theta$. Det kan ses ved

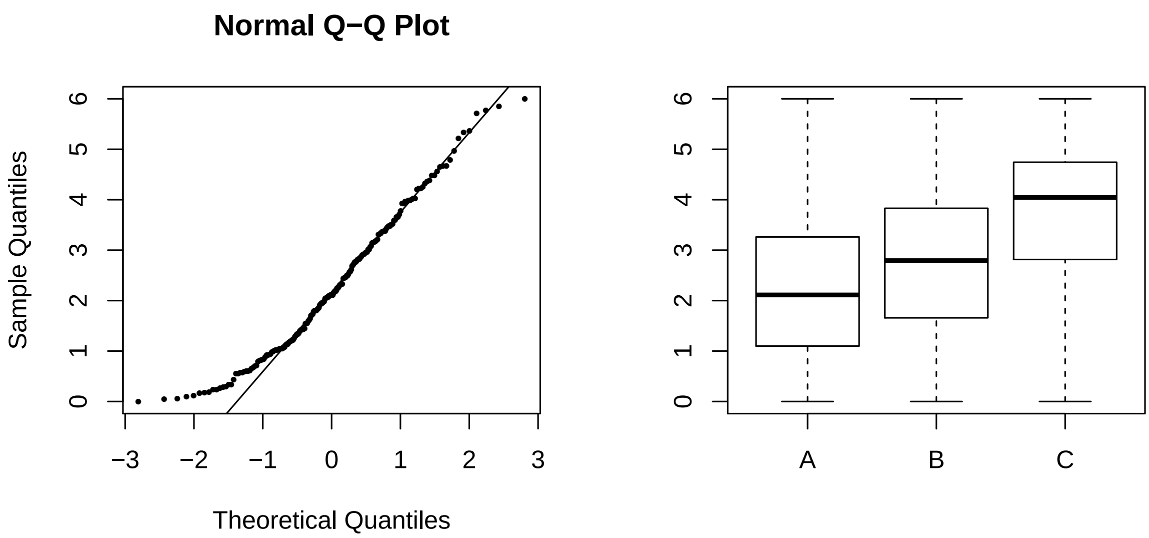
$$\mathrm{E}\,\hat{\theta} = \mathrm{E}[\frac{3 - \bar{X}}{4}] = \frac{3 - \mathrm{E}\,\bar{X}}{4} = \frac{3 - \mathrm{E}\,X}{4} = \frac{3 - (3 - 4\theta)}{4} = \theta.$$

Kommentar(er): Det er også meget fint at argumentere med, at $\bar{X}$ som er en unbiased estimator for $\mathrm{E}\,X$, men så bør nævnes at $\hat{\theta}$ fremkommer ved linær transformation af $\bar{X}$, ellers er argumentet ikke tilstrækkeligt. Tænk på tilfælde hvor $\hat{\theta} = 1/\bar{X}$. Iflg. Jensens ulighed er det generelt ikke en unbiased estimator.

Opgave 2

Denne opgave vægter med 8%

Der er blevet simuleret en stikprøve af størrelse 200 fra hver af tre forskellige fordelinger. Nedenstående figur viser et normalfraktilplot for en af stikprøverne, samt boxplots for alle tre.



(a) Hvilket af de tre boxplots A, B, eller C hører til det viste normalfraktilplot? Giv en kort begrundelse for dit svar.

Boxplot A hører til normalfraktilplottet.

Ved normalfraktilplottet aflæses, at medianen ligger lidt over 2 - det er værdien af sample quantile for theoretical quantile = 0, da standardnormalfordelingens median er 0. Medianerne i boxplots er indtegnet som de tykke streger, og ligger ved hhv lidt over 2 (A), lidt under 3 (B) og ca. 4 (C). Derfor kommer kun boxplot A i betragtning.

Kommentar(er): Andre begrundelser er muligt. Bl.a. kan man se fra fraktildiagrammet, at stikprøven kommer fra en højreskæv fordeling, hvor 50% af værdierne ligger mellem 0 og 2, og 50% ligger mellem 2 og 6. Det afspejles også i asymmetrien af boxplot A.
De empiriske fraktiler Q_1 og Q_3 findes i fraktildiagrammet ved at aflæse grafen over de teoretiske fraktiler $\Phi^{-1}(0.25) = -0.674$ og $\Phi^{-1}(0.75) = 0.674$.

Opgave 3

Denne opgave vægter med 20%

Wafers er siliciumskiver, der danner råmateriale til elektroniske chips. Wafers produceres i renrum, da hver mikropartikel, der lander på overfladen, mindsker udbyttet af chippene. Det kan dog ikke helt undgås, at wafers kontamineres af partikler. Et firma, der producerer renluftanlæg, lover, at i deres anlæg sætter sig gennemsnitligt kun 10 partikler på overfladen af en wafer i løbet af en time. For at teste firmaets påstand tælles antallet X af partikler på 5 wafers der lå i renrummet i 24 timer. Hvis antallet overstiger 1260 partikler, forkaster vi påstanden om, at der kun sætter sig 10 partikler per time på overfladen af en wafer.

Der antages, at partiklerne sætter sig på wafers uafhængigt af hinanden, og at antallet af partikler, der sætter sig på en wafer i løbet af en time, følger en Poisson-fordeling med parameter μ.

(a) Gør rede for, at antallet X er Poisson-fordelt med parameter $\lambda = 120 \cdot \mu$.

Antallet X kan skrives som en sum af variabler Y_{ij}, der betegner antal af partikler der sætter sig på den i-te wafer i løbet af den j-te time, $i = 1, \dots, 5$, $j = 1, \dots, 24$. Da partiklerne agerer uafhængige, er de Y_{ij} også uafhængige, og ifølge antagelsen er de identisk Poisson fordelt med parameter μ. Fra Theorem B.6 følger så, at

$$X = \sum_{i=1}^{5} \sum_{j=1}^{24} Y_{ij} \sim \text{Poisson}(\sum_{i=1}^{5} \sum_{j=1}^{24} \mu) = \text{Poisson}(120\mu).$$

(b) Formuler nulhypotesen og den alternative hypotese i testet.

$H_0 : \mu = 10$ vs. $H_A : \mu > 10$.
Kommentar(er): En anden mulighed er $H_0 : \lambda = 1200$ vs. $H_A : \lambda > 1200$.
Det er også rigtigt at skrive $H_0 : \mu \leq 10$, men det er forkert at skrive $H_A : \mu \neq 10$, fordi testet er ensidig.

(c) Angiv sandsynligheden for type 1 fejl.

Sandsynligheden for en type 1 fejl beregnes som

$$\begin{aligned} P(H_0 \text{ forkastes} \mid H_0 \text{ er sand}) &= P(X > 1260 \mid X \sim \text{Poisson}(1200)) \\ &= 1 - P(X \leq 1260 \mid X \sim \text{Poisson}(1200)) \\ &= 1 - F_{\text{Poisson}(1200)}(1260) \\ &= 0.0412. \end{aligned}$$

(d) Angiv sandsynligheden for type 2 fejl, i tilfældet hvor den sande rate er $\mu = 11$ partikler per time og wafer.

Vi ser her på den specielle alternative $H_A : \mu = 11$ som betyder at $X \sim \text{Poisson}(1320)$. Sandsynligheden for en type 2 fejl beregnes som

$$\begin{aligned} P(H_0 \text{ forkastes ikke} \mid H_A \text{ er sand}) &= P(X \leq 1260 \mid X \sim \text{Poisson}(1320)) \\ &= P(X \leq 1260 \mid X \sim \text{Poisson}(1320)) \\ &= F_{\text{Poisson}(1320)}(1260) \\ &= 0.0499. \end{aligned}$$

Vink: til besvarelse af opgave (c) og (d) kan R med fordel bruges.

Opgave 4

Denne opgave vægter med 25%

Selvom revnede æg sorteres fra, inden æg lægges i æggebakkerne til privatkunder, får man alligevel ind imellem knækæg (æg med revner). Bakker med to eller flere knækæg mindsker kundetilfredsheden betydeligt. For at skønne risikoen fik et stort supermarked gennemset en levering på 1000 æggebakker, hver med 10 æg. Under antagelsen om, at æg i en bakke får revner uafhængigt af hinanden, burde antal knækæg følge en binomialfordeling. Under denne antagelse estimeres sandsynligheden for at et æg er et knækæg, og det forventede antal bakker med 0, 1, ..., 10 knækæg beregnes. Tabellen viser en opstilling af de observerede og forventede tal.

	antal knækæg i bakken						
	0	1	2	3	4	5	> 5
observeret	766	194	25	7	5	3	0
forventet	737.42	228.07	31.74	2.62	0.14	0.01	0.00

(a) Gør rede for, at den forventede antal bakker uden knækæg må være 737.42.

Først estimeres sandsynligheden for, at et æg har revner, under antagelsen, at antal X_i af knækæg i den i-te bakke er binomialfordelt, $X_i \sim \text{Binom}(10, p)$, $i = 1, \ldots, 1000$. Så har antal af alle knækæg $X_. = \sum_{i=1}^{1000} X_i$ ligeledes en binomialfordeling, $X_. \sim \text{Binom}(10000, p)$. Maksimum-likelihood estimatet for p bliver så $\hat{p} = \frac{X_.}{10\,000}$, hvor

$$X_. = 0 \cdot 766 + 1 \cdot 194 + 2 \cdot 25 + 3 \cdot 7 + 4 \cdot 5 + 5 \cdot 3 + 0 = 300 \quad \Longrightarrow \quad \hat{p} = 0.03.$$

Så får vi sandsynligheden for, at den i-te bakke er fri for knakæg, som

$$\widehat{P(X_i = 0)} = \binom{10}{0} \hat{p}^0 (1 - \hat{p})^{10} = 0.97^{10} = 0.73742,$$

hvorved den forventede antal bakker uden knakæg estimeres som

$$\widehat{\text{E}[\#i : X_i = 0]} = 1000 \cdot \widehat{P(X_i = 0)} = 737.42.$$

(b) Lav et goodness of fit test for hypotesen, at antal knækæg er binomialfordelt.

Der er flere klasser i tabellen med forventet antal mindre end 5. De lægges først sammen:

	antal knækæg i bakken		
	0	1	≥ 2
observeret	766	194	$25 + 7 + 5 + 3 + 0 = 40$
forventet	737.42	228.07	$31.74 + 2.62 + 0.14 + 0.01 + 0.00 = 34.51$

Så laves en Pearson's χ^2 test, eller en G-test. Teststatistik:

$$C = \frac{(766 - 737.42)^2}{737.42} + \frac{(194 - 228.07)^2}{228.07} + \frac{(40 - 34.51)^2}{34.51} = 7.070552$$

eller

$$G = 2 \cdot \left(766 \cdot \ln\left(\frac{766}{737.42}\right) + 194 \cdot \ln\left(\frac{194}{228.07}\right) + 40 \cdot \ln\left(\frac{40}{34.51}\right) \right) = 7.2878$$

sammenlignes med en χ^2 fordeling med $df = 3-1-1 = 1$ frihedsgrader (1 parameter estimeret). Det giver

$$p_{\text{obs}} = 1 - F_{\chi^2(1)}(7.070552) = 0.0078$$

hhv.

$$p_{\text{obs}} = 1 - F_{\chi^2(1)}(7.2878) = 0.0069.$$

Begge tests ville forkaste nulhypotesen, at antal knækæg er binomialfordelt, ved signifikansniveauet $\alpha = 0.05$.

(c) Fortolk resultatet af hypotesetestet.

p-værdien er meget lavere end $\alpha = 5\%$. Data strider i mod hypotesen (ved signifikansniveau 5%), at antal knækæg i en bakke er binomialfordelt. Vi tvivler derfor på, at æg knækkes uafhængigt fra hinanden, med samme sandsynlighed hver.

Opgave 5

Denne opgave vægter med 35%

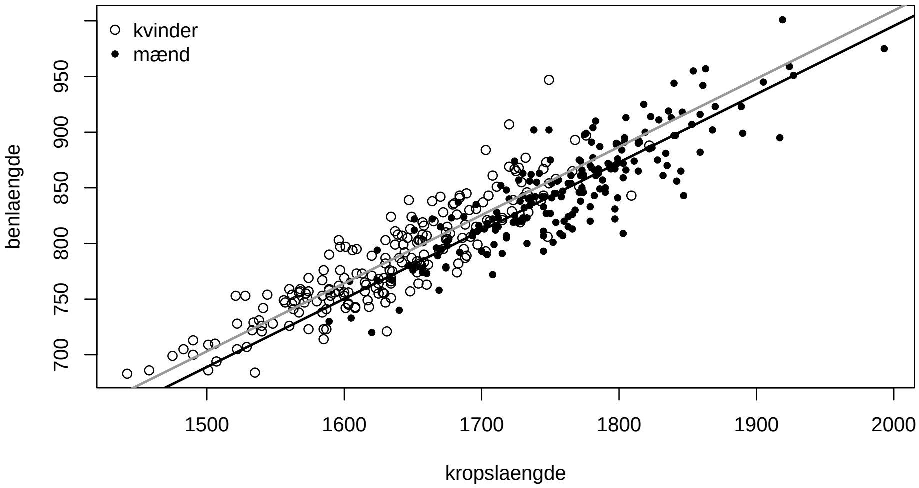
Skal cykler til mænd have en anden geometri end kvindecykler? Eller: kører man som kort mand ganske bekvemt på en stor kvindecykel? Hvis ja, så burde forholdet mellem kropslængde og benlængde være ens hos begge køn. Dette undersøges i nærværende opgave. For 1986 kvindelige og 4082 mandlige medlemmer af US army blev der i årene 2010/2011 taget mål. Datasættet `anthro` til denne opgave består af en tilfældig delstikprøve på 200 mænd og 200 kvinder. Det indeholder faktorvariablen `koen` samt de numeriske variabler `benlaengde` og `kropslaengde`. Længder er angivet i mm.

Figur 1 nedenfor viser benlængde plottet mod kropslængde, med forskellige symboler afhængigt af køn. Tilsyneladende kan data godt beskrives ved en simpel lineær regressionsmodel. Regressionslinjer er indtegnet separat for mænd og kvinder.

Vi starter med modellen

$$M_0 : Y_{ij} \sim N(\alpha_i + \gamma_i x_{ij}, \sigma_i^2), \quad i = 1, 2, \quad j = 1, \dots, 200,$$

hvor Y_{ij} er benlængde og x_{ij} er kropslængde, og $i = 1, 2$ er køn ($i = 1$ kvinde, $i = 2$ mand).



Figur 1: Benlængde tegnet op mod kropslængde. Udhulede prikker: data fra kvinder, sorte prikker: mænd. Desuden er regressionslinjer indtegnet (grå: kvinder, sort: mænd).

Til besvarelse af de følgende spørgsmål kan man bruge relevante resultater fra R-udskrifterne nedenfor.

(a) Gør rede for, at modellen M_0 passer til data.

Bemærk, at modellen M_0 indeholder to antagelser: for hvert køn stammer data fra en *normalfordeling* hvis middelværdi afhænger *lineært* af kropslængden. Variansen kan være forskellig for mænd og kvinder.
Vi ser på plots fra R-protokol 1, hvor M_0 undersøges separat for begge køn, $i = 1$ og $i = 2$. Ved plottet "residuals vs. prediction" kan ses, at residualerne fordeler sig symmetrisk omkring 0, og der er ingen tendens eller struktur at se. Det antyder, at den lineære sammenhæng, som antages for middelværdien, passer fint. Residualernes fraktilplots danner en linje, som antyder, at residualerne er normalfordelt. Der er kun

nogle få (ca. 2 – 3) outliers ved kvinderne. Disse to plots understøtter antagelsen, at model M_0 beskriver data vel.

(b) Vis ved et test, at der kan antages, at de to varianser i M_0 er ens.

Vi vil gerne teste $H_0 : \sigma_1^2 = \sigma_2^2$ mod $H_A : \sigma_1^2 \neq \sigma_2^2$. Da vi kun har to stikprøver, bruger vi et F-test. Estimater til σ_1^2 og σ_2^2 fås fra R-protokol 2:

$$s_1^2 = 22.0628^2 = 486.7671, \quad s_2^2 = 23.6309^2 = 558.4194,$$

begge to er estimeret med 198 frihedsgrader hver.
Teststørrelse:

$$F = \frac{s_1^2}{s_2^2} = 0.8716 \quad \text{(eller, hvis man vender } s_1^2 \text{ og } s_2^2 \text{ om, } F = 1.1472\text{)}$$

p-værdi:

$$\begin{aligned} p_{\text{obs}} &= 2\min(F_{F(198,198)}(0.8716), 1 - F_{F(198,198)}(0.8716)) \\ &= 2\min(0.16738, 0.83262) = 0.33476. \end{aligned}$$

Da $p_{\text{obs}} > 0.05$, accepterer vi H_0.

Kommentar(er): Det er også fint at bruge Bartlett's test, men der skulle tages højde for, at de $n_i - 1$ som bruges ved Bartlett's test i ANOVA skal erstattes med frihedsgraderne

(c) Opskriv modellerne M_1, M_2, M_{3a} og M_{3b} fra R-protokol 3 i samme form som M_0 er givet.

$$M_1 : Y_{ij} \sim N(\alpha_i + \gamma_i x_{ij}, \sigma^2), \quad i = 1, 2, \; j = 1, \ldots, 200$$
(to regressionslinier, ens varians)

$$M_2 : Y_{ij} \sim N(\alpha_i + \gamma x_{ij}, \sigma^2), \quad i = 1, 2, \; j = 1, \ldots, 200$$
(to regressionslinier, ens varians, fælles hældning)

$$M_{3a} : Y_{ij} \sim N(\alpha + \gamma x_{ij}, \sigma^2), \quad i = 1, 2, \; j = 1, \ldots, 200$$
(én regressionslinie, ens varians)

$$M_{3b} : Y_{ij} \sim N(\alpha_i, \sigma^2), \quad i = 1, 2, \; j = 1, \ldots, 200$$
(forskellig middelværdi, ens varians)

(d) Vis ved et test, at der kan antages at hældningerne er ens i modellen med ens varianser.

Hypotesen $H_0 : \gamma_1 = \gamma_2$, at hældningerne er éns i model M_1, svarer til modelreduktion fra M_1 til M_2.
Der kan enten laves et t-test eller alternativt et ensided F-test.
Resultatet af t-testet aflæses simpelthen i summary af M_1, under linjen `kropslaengde:koenm`, med en p-værdi på $p_{\text{obs}} = 0.984$.
Til F-test beregnes teststørrelsen fra $s_{01}^2 = 22.8603^2$ med $df_1 = 396$ og $s_{02}^2 = 22.8315^2$ med $df_2 = 397$,

$$F = \frac{(s_{02}^2 \cdot df_2 - s_{01}^2 \cdot df_1)(df_2 - df_1)}{s_{01}^2} = 0.000328,$$

som giver

$$p_{\text{obs}} = 1 - F_{1,396}(0.000328) = 0.9856.$$

(Der kan forekomme afvigelser pga. runding.)
I begge tilfælde er p-værdien meget større end gængse signifikansniveauer, og derfor forkastes H_0 ikke.

(e) Vis at modellen ikke kan reduceres yderligere.

Igen kan der aflæses t-tests:

- Reduktion fra M_2 til M_{3a} svarende til $H_{0a} : \alpha_1 = \alpha_2$ testes med en p-værdi på $p_{\text{obs}} = 7.94 \cdot 10^{-6}$, som er meget mindre end $\alpha = 0.05$ og som forkastes derfor. Det aflæses i `summary(M2)`, linje `koenm`.
- Reduktion fra M_2 til M_{3b} svarende til $H_{0b} : \gamma = 0$ testes med en p-værdi på $p_{\text{obs}} < 2 \cdot 10^{-16}$, som er meget mindre end $\alpha = 0.05$ og som forkastes derfor. Det aflæses i `summary(M2)`, linje `kropslaengde`.

Alternativt fås de tilsvarende F-tests som

- Reduktion fra M_2 til M_{3a}:

$$F = \frac{(s_{03a}^2 \cdot df_{3a} - s_{02}^2 \cdot df_2)/(df_{3a} - df_2)}{s_{02}^2} = \frac{23.3838^2 \cdot 398 - 22.8315^2 \cdot 397}{22.8315^2} = 20.488$$

p-værdi: $p_{\text{obs}} = 1 - F_{F(1,397)}(F) = 7.94 \cdot 10^{-6}$.

- Reduktion fra M_2 til M_{3b}:

$$F = \frac{(s_{03b}^2 \cdot df_{3b} - s_{02}^2 \cdot df_2)/(df_{3b} - df_2)}{s_{02}^2} = \frac{48.0642^2 \cdot 398 - 22.8315^2 \cdot 397}{22.8315^2} = 1366.8$$

p-værdi: $p_{\text{obs}} = 1 - F_{F(1,397)}(F) = 0$.

I opgaverne (f) og (g) betragtes modellen med to regressionslinier med fælles hældning.

(f) Angiv et 95% konfidensinterval for forskellen i skæringen med andenaksen (intercept) for regressionslinjerne for mænd og kvinder.

I R-protokol 3, `summary(M2)`, aflæses estimat for forskellen af skæringen til at være

$$\widehat{\alpha_1 - \alpha_2} = -13.95712$$

og estimat til standardfejlen til at være

$$\widehat{SE}(\widehat{\alpha_1 - \alpha_2}) = 0.308352.$$

Til at konstruere et konfidensinterval bruges fraktilen fra t-fordelingen med 397 frihedsgrader,

$$q = F_{t(397)}^{-1}(1 - \alpha/2) = F_{t(397)}^{-1}(0.975) = 1.965957.$$

(Ved så mange frihedsgrader er det også tilladt at bruge det tilsvarende fraktil fra standardnormalfordelingen.) Grænserne på konfidensintervallet fås som

$$\widehat{\alpha_1 - \alpha_2} \pm q \cdot \widehat{SE}(\widehat{\alpha_1 - \alpha_2}) = -13.95712 \pm 1.965957 \cdot 3.08352 = [-20.019, -7.895].$$

(g) Angiv et estimat for den gennemsnitlige benlængde for en kvinde og for en mand, som begge har en kropslængde på 1700 mm.

Lad $\mu_{is} = \alpha_i + \gamma x_{is}$, $i = 1, 2$ betegne middelværdien af benlængden af en kvinde ($i = 1$) eller mand ($i = 2$) med kropslængde x_{is}. Så får vi et estimat for μ_{is} ved at indsætte estimater til α_i og γ: $\hat{\mu}_{is} = \hat{\alpha}_i + \hat{\gamma} x_{is}$. I opgaven er $x_{is} = 1700$. Fra summary til M_2 aflæses, at

$$\begin{aligned} \hat{\alpha}_1 &= -215.87526 && \texttt{(Intercept)}, \\ \widehat{\alpha_2 - \alpha_1} &= -13.95712 && \texttt{koenm}, \\ \hat{\gamma} &= 0.61261 && \texttt{kropslaengde} \end{aligned}$$

Altså bliver den estimerede gennemsnitlige benlængde:

$$\hat{\mu}_{is} = \begin{cases} -215.87526 + 0.61261 \cdot 1700 = 825.5617, & i = 1 \text{ (kvinde)}, \\ \underbrace{-215.87526 - 13.95712}_{=-229.8324} + 0.61261 \cdot 1700 = 811.6046, & i = 2 \text{ (mand)}. \end{cases}$$

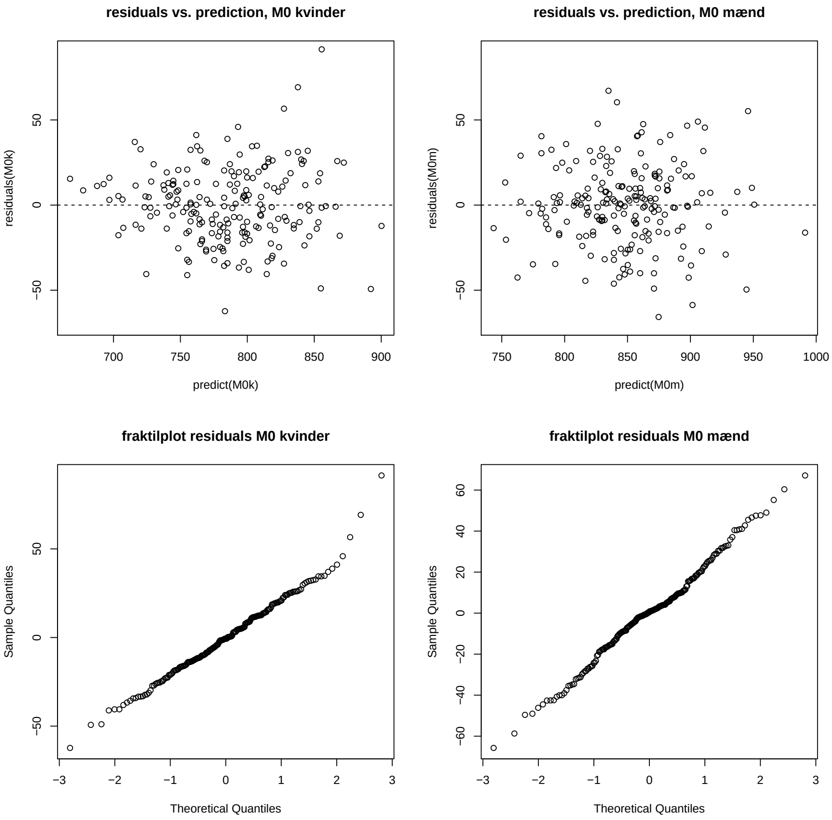
R-protokol 1

```
> kvinder <- subset(anthro, koen == "k")
> maend <- subset(anthro, koen == "m")

> M0k <- lm(benlaengde ~ kropslaengde, data = kvinder))
> M0m <- lm(benlaengde ~ kropslaengde, data = maend))

> plot(residuals(M0k) ~ predict(M0k), main = "residuals vs. prediction, M0 kvinder")
> abline(h = 0, lty = "dashed")
> plot(residuals(M0m) ~ predict(M0m), main = "residuals vs. prediction, M0 maend")
> abline(h = 0, lty = "dashed")

> qqnorm(residuals(M0k), main = "fraktilplot residuals M0 kvinder")
> qqnorm(residuals(M0m), main = "fraktilplot residuals M0 maend")
```



Figur 2: Plots lavet med R-koden ovenfor.

R-protokol 2

```
> summary(M0m)
Coefficients:
               Estimate Std. Error t value Pr(>t)
(Intercept)  -230.42803   43.72827   -5.27 3.56e-07 ***
kropslaengde    0.61295    0.02484   24.68  < 2e-16 ***
---
Residual standard error: 23.6309 on 198 degrees of freedom

> summary(M0k)
Coefficients:
               Estimate Std. Error t value Pr(>t)
(Intercept)  -215.37097   36.21261  -5.947 1.21e-08 ***
kropslaengde    0.61230    0.02214  27.659  < 2e-16 ***
---
Residual standard error: 22.0628 on 198 degrees of freedom
```

R-protokol 3

```
> M1 <- lm(benlaengde ~ kropslaengde * koen, data = anthro)
> M2 <- lm(benlaengde ~ kropslaengde + koen, data = anthro)
> M3a <- lm(benlaengde ~ kropslaengde, data = anthro)
> M3b <- lm(benlaengde ~ koen, data = anthro)

> summary(M1)
```

Opgavesættet fortsættes

```
Coefficients:
                     Estimate Std. Error t value Pr(>t)
(Intercept)        -215.37100   37.52166  -5.740 1.89e-08 ***
kropslaengde          0.61231    0.02294  26.694  < 2e-16 ***
koenm               -15.05706   56.54514  -0.266    0.790
kropslaengde:koenm    0.00065    0.00332   0.019    0.984
---
Residual standard error: 22.8603 on 396 degrees of freedom

> summary(M2)
Coefficients:
               Estimate Std. Error t value Pr(>t)
(Intercept)  -215.87526   27.12834  -7.958 1.86e-14 ***
kropslaengde    0.61261    0.01657  36.971  < 2e-16 ***
koenm         -13.95712    3.08352  -4.526 7.94e-06 ***
---
Residual standard error: 22.8315 on 397 degrees of freedom

> summary(M3a)
Coefficients:
               Estimate Std. Error t value Pr(>t)
(Intercept)  -137.31535   21.35414   -6.43 3.65e-10 ***
kropslaengde    0.56220    0.01257   44.74  < 2e-16 ***
---
Residual standard error: 23.3838 on 398 degrees of freedom

> summary(M3b)
            Estimate Std. Error t value Pr(>t)
Coefficients:
(Intercept)  785.300      3.399  231.06   <2e-16 ***
koenm         62.665      4.806   13.04   <2e-16 ***
---
Residual standard error: 48.0642 on 398 degrees of freedom
```