Oppgave 1

a) `pt(-1, df=10, lower.tail=FALSE)`

b) `dt(2.5, df=10)`

c) `qt(.25, df=10)`

Oppgave 2

a) Vi har tilpasset modellen

$$\ln y = \beta_0 + \beta_1 \ln x + \alpha_i + e, \tag{1}$$

hvor $e \sim N(0, \sigma^2)$. Residualleddene antas i tillegg uavhengige. β_0, β_1, $\alpha_1, \alpha_2, \ldots, \alpha_7$ og σ_2 er ukjente parametere.

b) I `drop1`-testen testes nullhypotesene at henholdvis `art` og `log(lengde)` ikke har noen effekt, altså nullhypotesene

$$\ln y = \beta_0 + \beta_1 \ln x + e, \tag{2}$$

og

$$\ln y = \beta_0 + \alpha_i + e, \tag{3}$$

mot modell (1) som alternativ hypotese. Siden begge testene er signifikante kan vi konkludere med at både `log(lengde)` og `art` har en effekt på responsvariabelen `log(vekt)`. Siden `art` er antatt å ha en additiv effekt på `log(vekt)` og blir estimert forskjell i forventet `vekt` for Pike i forhold til Bream (referansekategorien) dersom alle andre forklaringsvariable holdes konstant en faktor på

$$\exp(\hat{\alpha}_6)/\exp(\hat{\alpha}_1) = \exp(-0.72)/\exp(0) = 0.49 \tag{4}$$

d.v.s. 51% mindre vekt.

c) Ved å omskrive ligning (1) får vi at sammenhengen mellom vekt y og lengde x for en gitt art i er gitt ved

$$\begin{aligned} y &= e^{\beta_0 + \beta_1 \ln x + \alpha_i} && (5) \\ &= e^{\beta_0 + \alpha_i} x^{\beta_1}, && (6) \end{aligned}$$

m.a.o. en kurvelineær sammenheng mellom x og y. Hvis $\beta_1 = 1$ har vi direkte proposjonalitet. Dersom fisk av ulike størrelse er formlike vil volum og dermed vekt y om tetthet er konstant være proposjonalt med lengde opphøyd i tredje, x^3. Dette svarer altså til nullhypotesen $\beta_1 = 3$. Under denne nullhyotesen er

$$T = \frac{\hat{\beta}_1 - 3}{\widehat{\mathrm{SE}\,\hat{\beta}_1}} \tag{7}$$

Observert verdi blir $(3.15 - 3)/0.035 = 4.28$. Med $|T| > t_{\alpha/2,149} = t_{0.025,149} = 1.96$ som forkastningsregel forkaster vi dermed H_0 og konkluderer med at β_1 er forskjellig fra 3.

t-fordelt med $n - p = 157 - 8 = 149$ frihetsgrader [1]

d) Nullhypotesen (1) innebærer at stigningstallene i de lineære sammenhengene `log(vekt)` og `log(lengde)` er de samme for alle 7 arter. Inkluderer vi en interaksjon mellom `log(lengde)` og `art` kan stigningstallene være forskjellige (alternativ hypotese). I matematisk notasjon kan modellen med interaskjon skrives

$$\ln y = \beta_0 + \beta_1 \ln x + \alpha_i + \gamma_i \ln x + e, \tag{8}$$
$$= (\beta_0 + \alpha_i) + (\beta_1 + \gamma_i) \ln x + e. \tag{9}$$

Tolkningen av parameterene $\gamma_2, \gamma_3, \ldots, \gamma_7$ (estimatene med navn `art2:log(lengde)` o.s.v. i utskriften) er altså forskjellen i stigningstall for art 2 i forhold til art 1 o.s.v.

F-testen viser at den utvidede modellen med interaksjonsledd i er signifikant bedre (p-verdien er lik 0.93) og vi konkluderer dermed med at dataene ikke gir grunnlag for å påsta at stigningstallene er forskjellig mellom artene og beholder nullhypotesen at stigningstallene er like (parallelle regresjonslinjer på log-log skala).

Oppgave 3

a) Sannsynlighetsmaksimeringsestimatet av θ blir omkring $\hat{\theta} = 1.45$.

b) Under H_0 er $D = 2(\ln L_1 - \ln L_0)$ kji-kvadratfordelt med $p_1 - p_0 = 1 - 0 = 1$ frihetsgrader. Vi forkaster for store verdier av D. Kritisk verdi er dermed $\chi^2_{0.05,1} = 3.84$. Avlesning av grafen gir at maksimalt log likelihood under H_0 og H_1 blir henholdsvis $\ln L_0 = -13.75$ og $\ln L_1 = -12.48$. Observert verdi av D blir dermed 2.54. Vi beholder dermed H_0.

c) Et tilnærmet estimat av variansen til $\hat{\theta}$ er

$$\widehat{\mathrm{Var}\,\hat{\theta}} = \frac{1}{-\frac{d^2}{d\theta^2} \ln L|_{\theta=\hat{\theta}}} = \frac{1}{9.44} = 0.106. \tag{10}$$

Et estimat av standardfeilen er dermed

$$\widehat{\mathrm{SE}\,\hat{\theta}} = \sqrt{0.106} = 0.325. \tag{11}$$

[1] $n = 157$ og ikke 159 som oppgitt i oppgavetekst da to lengder er missing i dataene.