

F10 R Kod

2024-02-29

```
#### Problem 1 ####
## vingfrekvens
calolaemus <- subset(Problem1, underart == "calolaemus")
## Om du inte kan subset()
calolaemus <- c(71.1, 69.3, 72.1, 69.3, 66.7, 69.3, 67.6, 69.2, 69.6, 70.9)

## Delproblem 1.
## Konfidensintervall fC6r medelvC$rdet fC6r calolaemus
t.test(calolaemus$frekvens)
## eller
t.test(calolaemus)
## Jag lC$ser av konfidensintervallet frC%n outputen.

## Delproblem 2.
## LC6sningsvariant 1.

## NormalfC6rdelning? Kolla qqplot.
qqnorm(calolaemus, main = "Problem 1")
qqline(calolaemus)

## Tankar:
## Det C$r vC$ldigt svC%rt att bedC6ma om det C$r normalfC6rdelat
## nC$r man bara har 10 mC$tpunkter. Men QQ-plotten visar i alla fall
## inte pC% nC%gra stora avvikelser frC%n normalfC6rdelning. SC% vi kan
## anta att data C$r nC%gorlunda normalfC6rdelade.

## GC6r ett t-test fC6r tvC% stickprov!
homogenes <- c(70.6, 69.8, 71.7, 73.4, 71.2, 70.4, 70.6, 72.6, 72.3, 72.2)
qqnorm(homogenes, main = "Problem 1")
qqline(homogenes)
t.test(calolaemus, homogenes)

## wilcox.test
wilcox.test(calolaemus, homogenes)

#### Problem 2 ####
## chi2-test:
chisq.test(Problem2)

#### Problem 3 ####
## Del 1: regression
## LinjC$rt samband
plot(Vikt ~ Storlekpixlar, data = Problem3)
```

```

## LinjC$r regression utan nC%gra transformeringsr verkar rimligt!
LR <- lm(Vikt ~ Storlekpixlar, data = Problem3)
summary(LR)
coef(LR)
## Skriv koefficienterna med decimaler istC$llet fC6r "scientific notation":
options(scipen = 1000)
coef(LR)

## Del 2. Prediktion!
predict(LR, newdata = data.frame(Storlekpixlar = 5525945))
-0.289413340803 + 0.000002586817 * 5525945

## Del 3. FC6rutsC$ttningar
## FC6rutsC$ttning 1: linjC$rt samband
plot(Vikt ~ Storlekpixlar, data = Problem3)
## FC6rutsC$ttning 2: lika varians
plot(LR, which = 1)
plot(LR, which = 3)
## FC6rutsC$ttning 3: normalfC6rdelning
plot(LR, which = 2)
## Modellantaganden ser OK ut.

## Om inte normal
boot_summary(LR, method = "residual") # Passa bra med icke-normal
boot_summary(LR, method = "case") # ickenormal och olika varians

#### Problem 4 ####
# ANOVA:
m <- aov(frekvens ~ underart, data = Problem1)
summary(m)

# Parvisa jC$mfc6relser med Tukeys HSD:
TukeyHSD(m)
plot(TukeyHSD(m))

# Modelldiagnostik:
# NormalfC6rdelning? Ser ganska bra ut!
plot(m, which = 2)
# Lika varians? Ser ganska bra ut!
plot(m, which = 1)
plot(m, which = 3)

#### Problem 5 ####
# Modell 1: 3D-scanner
m1 <- lm(Vikt ~ Storlek_pixlar_3D, data = Problem5)
summary(m1) # Noterar R2-vC$rdet: 0.95
plot(Vikt ~ Storlek_pixlar_3D, data = Problem5)
abline(m1)
plot(m1, which = 1:3) # Modellantaganden ser OK ut

# Modell 2: 2D-scanner
m2 <- lm(Vikt ~ Storlek_pixlar_2D, data = Problem5)

```

```

summary(m2) # Noterar R2-vc$rdet: 0.84
plot(Vikt ~ Storlek_pixlar_2D, data = Problem5)
abline(m2)
plot(m2, which = 1:3) # Modellantaganden ser rC$tt OK ut

# Modell 1 ser bC$ttre ut: hC6gre R2 och punkterna ligger nC$rmare den anpassade linjen!

# Tar nu fram ett prediktionsintervall med modell 1:
predict(m1, newdata = data.frame(Storlek_pixlar_3D = 5525945),
        interval = "prediction")

```