



BAGGRUND OG DATABESKRIVELSE (EKSEMPEL 1-2)

Denne opgave beskæftiger sig med den samlede månedlige el-produktion i Danmark i perioden 2000-2015, herunder hvor stor en del af produktionen der stammer vindmøller. Data er venligst stillet til rådighed af Energistyrelsen.

EKSEMPEL 1 (SAMLET EL-PRODUKTION)

- a). **Angiv** de estimerede parametre i en lineær regressionsmodel med den samlede el-produktion i GigaWatt-timer (*El nettoproduktion GWh*) som responsvariabel og gennemsnitstemperaturen (målt over hele døgnet) i grader Celcius (*Middeltemperatur DK C*) som forklarende variabel.
Giv en fortolkning af de estimerede parametre.
Angiv modellens forklaringsgrad.
- b). **Beregn** på baggrund af den estimerede regressionsmodel den forventede samlede månedlige el-produktion i GigaWatt-timer i en måned med en gennemsnitstemperatur på 15 grader Celsius.

EKSEMPEL 2 (EL-PRODUKTION FRA VINDMØLLER)

- a). **Angiv** de estimerede parametre i en lineær regressionsmodel med den samlede el-produktion fra vindkraft i GigaWatt-timer (*El Vindkraft GWh*) som responsvariabel og gennemsnitsvindhastigheden (målt over hele døgnet) i meter per sekund (m/s) (*Mid-delvind m/s*) som forklarende variabel.
- b). **Beregn** et 95%-konfidensinterval for β_0 og **giv** en fortolkning af intervallet.
- c). **Beregn** et 95%-konfidensinterval for betydningen for den forventede samlede el-produktion fra vindkraft af en stigning i gennemsnitsvindhastigheden på 1 m/s og **giv** en fortolkning af intervallet.
Beregn et 95%-konfidensinterval for betydningen for den forventede samlede el-produktion fra vindkraft af en stigning i gennemsnitsvindhastigheden på 0,5 m/s og **giv** en fortolkning af intervallet.
- d). **Gør rede for** om der er statistisk belæg for en nulhypotese om, at gennemsnitsvindhastigheden ikke har nogen betydning for den forventede samlede el-produktion fra



vindkraft (brug $\alpha = 5\%$ som signifikansniveau). **Angiv** i den forbindelse hypoteser samt teknisk og let forståelig konklusion.

- e). **Gør rede for** om der er statistisk belæg for en nulhypotese om, at for hver 1 m/s gennemsnitsvindhastigheden stiger, så stiger den forventede samlede el-produktion fra vindkraft med 100 GigaWatt-timer (brug $\alpha = 5\%$ som signifikansniveau). **Angiv** i den forbindelse hypoteser samt teknisk og let forståelig konklusion.

Gør rede for om der er statistisk belæg for en nulhypotese om, at for hver 1 m/s gennemsnitsvindhastigheden stiger, så stiger den forventede samlede el-produktion fra vindkraft med 150 GigaWatt-timer (brug $\alpha = 5\%$ som signifikansniveau). **Angiv** i den forbindelse hypoteser samt teknisk og let forståelig konklusion.



VEJLEDENDE LØSNINGER

Eksempel 1 (► Løsning i JMP)

1A):

De estimerede parametre hørende til modellen

El nettoproduktion GWh = $\beta_0 + \beta_1 \cdot \text{Middeltemperatur DK C} + \varepsilon$

hvor ε er normalfordelt $N(0, \sigma)$ er givet som

$$\hat{\beta}_0 = 3.947,69 \quad \hat{\beta}_1 = -110,08 \quad \hat{\sigma} = 435,23$$

Det betyder, at den forventede samlede månedlige el-produktion...

- er på 3.947,69 GigaWatt-timer såfremt gennemsnitstemperaturen er 0 grader Celsius
- falder med 110,08 GigaWatt-timer for hver 1 grad Celsius den månedlige gennemsnitstemperatur stiger

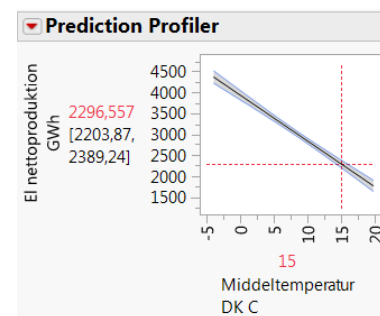
Endvidere betyder det, at den samlede månedlige el-produktion med ca. 95% sandsynlighed vil variere med $\pm 2 \cdot 435,23 = \pm 870,46$ GigaWatt-timer efter at vi har korrigeret for gennemsnitstemperaturens indvirkning på/sammenhæng med el-produktionen.

Den estimerede regressionsmodel har en forklaringsgrad på $R^2 = 0,6951$.

1B):

På baggrund af den i delspg. a) estimerede regressionsmodel er den forventede samlede månedlige el-produktion på 2.296,56 GigaWatt-timer, såfremt gennemsnitstemperaturen i en måned er 15 grader Celsius.

Summary of Fit				
RSquare	0,695096			
RSquare Adj	0,693393			
Root Mean Square Error	435,2326			
Mean of Response	2984,436			
Observations (or Sum Wgts)	181			
Analysis of Variance				
Parameter Estimates				
Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	3947,6862	57,62214	68,51	<,0001*
Middeltemperatur DK C	-110,0753	5,449068	-20,20	<,0001*



**Eksempel 2** (► Løsning i JMP)**2A):**

De estimerede parametre hørende til modellen

$$El \text{ Vindkraft } GWh = \beta_0 + \beta_1 \cdot \text{Middelvind } m/s + \varepsilon$$

hvor ε er normalfordelt $N(0, \sigma)$ er givet som

$$\hat{\beta}_0 = -307,85 \quad \hat{\beta}_1 = 193,96 \quad \hat{\sigma} = 259,43$$

Det betyder, at den forventede samlede månedlige el-produktion fra vindkraft...

- er på -307,85 GigaWatt-timer såfremt gennemsnitsvindhastigheden er 0 m/s.
(denne fortolkning er naturligvis en anelse aparte, idet en gennemsnitsvindhastighed på 0 betyder, at det overhovedet ikke blæser i en hel måned. Læg desuden mærke til, at der ikke er nogen måneder i datamaterialet, hvor det ikke i gennemsnit blæser mindst 3 m/s. Derfor er tilfældet med 0 m/s udelukkende et teoretisk scenarie, og derfor er det ikke overraskende, at modellens forventede produktion (negativ produktion) i dette tilfælde er ren nonsens.)
- stiger med 193,96 GigaWatt-timer for hver 1 m/s den månedlige gennemsnitsvindhastighed stiger

Endvidere betyder det, at den samlede månedlige el-produktion fra vindkraft med ca. 95% sandsynlighed vil variere med $\pm 2 \cdot 259,43 = \pm 518,86$ GigaWatt-timer efter at vi har korrigeret for gennemsnitsvindhastighedens indvirkning på/sammenhæng med el-produktionen fra vindkraft.

2B):

Et 95%-konfidensinterval for β_0 er givet som $[-535,58; -80,11]$. Det betyder, at i en måned med en gennemsnitsvindhastighed på 0 vil den forventede el-produktion fra vindkraft med 95% sandsynlighed være på mellem -535,58 GigaWatt-timer og -80,11 GigaWatt-timer (som ovenfor er fortolkningen også her ren nonsens, fordi det ikke giver mening at alene se på β_0 , idet det svarer til at se på el-produktionen fra vindkraft i en måned, hvor det overhovedet ikke blæser bare det mindste).

Summary of Fit

RSquare	0,272061
RSquare Adj	0,267994
Root Mean Square Error	259,4333
Mean of Response	622,8343
Observations (or Sum Wgts)	181

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob> t	Lower 95%	Upper 95%
Intercept	-307,8472	115,4084	-2,67	0,0083*	-535,5833	-80,11111
Middelvind m/s	193,95895	23,71361	8,18	<,0001*	147,16476	240,75314



2c):

Et 95%-konfidensinterval for β_1 er givet som $[147,16; 240,75]$. Det betyder, at med 95% sandsynlighed vil den forventede el-produktion fra vindkraft stige med mellem 147,16 og 240,75 GigaWatt-timer for hver gang gennemsnitsvindhastigheden stiger med 1 m/s.

Et 95%-konfidensinterval for $\beta_1 \cdot 0,5$ er givet som $[147,16 \cdot 0,5; 240,75 \cdot 0,5] = [73,58; 120,38]$. Det betyder, at med 95% sandsynlighed vil den forventede el-produktion fra vindkraft stige med mellem 73,58 og 120,38 GigaWatt-timer for hver gang gennemsnitsvindhastigheden stiger med 0,5 m/s.

2d):

Test af nulhypotesen om at gennemsnitsvindhastigheden ikke har nogen betydning for den samlede månedlige el-produktion fra vindkraft

$$H_0: \beta_1 = 0 \quad \text{og} \quad H_a: \beta_1 \neq 0$$

forkastes idet P-værdien er mindre end 0,01% (alternativt: fordi 0 ikke tilhører 95%-konfidensintervallet for β_1 , jf. delspg. 2c)). Der er således ikke på baggrund af datamaterialet belæg for at hævde, at gennemsnitsvindhastigheden ikke har en sammenhæng med den samlede el-produktion fra vindkraft (hvilket i øvrigt heller ikke ville give meget mening).

2e):

Test af nulhypotesen om at for hver 1 m/s gennemsnitsvindhastigheden stiger, så stiger den forventede samlede månedlige el-produktion fra vindkraft med 100 GigaWatt-timer

$$H_0: \beta_1 = 100 \quad \text{og} \quad H_a: \beta_1 \neq 100$$

forkastes, fordi 100 ikke tilhører 95%-konfidensintervallet for β_1 (jf. delspg. 2c)). Der er således ikke på baggrund af datamaterialet belæg for at hævde, at en stigning i gennemsnitsvindhastigheden på 1 m/s medfører en stigning i den forventede samlede månedlige el-produktion fra vindkraft på 100 GigaWatt-timer.

Test af nulhypotesen om at for hver 1 m/s gennemsnitsvindhastigheden stiger, så stiger



den forventede samlede månedlige el-produktion fra vindkraft med 150 GigaWatt-timer

$$H_0: \beta_1 = 150 \quad \text{og} \quad H_a: \beta_1 \neq 150$$

forkastes ikke, fordi 150 tilhører 95%-konfidensintervallet for β_1 (jf. delspg. 2c)). Der er således ikke på baggrund af datamaterialet belæg for at afvise en påstand om, at en stigning i gennemsnitsvindhastigheden på 1 m/s medfører en stigning i den forventede samlede månedlige el-produktion fra vindkraft på 150 GigaWatt-timer.