

BAGGRUND OG DATABESKRIVELSE (EKSEMPEL 1-4)

Denne opgave beskæftiger sig med den samlede månedlige el-produktion i Danmark. Opgaven tager udgangspunkt i datamaterialet i filen El_produktion2. jmp, hvis indhold er beskrevet på eksempelark 9.

EKSEMPEL 1 (SAMLET EL-PRODUKTION)

- a). **Tegn** en figur der viser, hvordan den samlede el-produktion i GigaWatt-timer (*El nettoproduktion GWh*) afhænger af måneden på året (*Måned*).
 - Gør rede for på baggrund af figuren, om der ser ud til at være en lineær eller en kvadratisk sammenhæng mellem den samlede el-produktion og måneden på året.
- b). **Opskriv** en lineær regressionsmodel med den samlede el-produktion i GigaWatt-timer som responsvariabel og en kvadratisk effekt af måneden på året, dvs. en regressionsmodel med måned *og* måned opløftet i 2. potens (dvs. *Måned*²) som forklarende variable.
 - Angiv de estimerede parametre i modellen.
- c). **Beregn** på baggrund af den i delspg. b) estimerede model den forventede samlede el-produktion i GigaWatt-timer for henholdsvis april og oktober måned.

EKSEMPEL 2 (SAMLET EL-PRODUKTION)

- a). **Opskriv** en lineær regressionsmodel med den samlede el-produktion i GigaWatt-timer som responsvariabel og gennemsnitstemperatur (målt over hele døgnet) i grader Celcius (*Middeltemperatur DK C*) og årstid (*Årstid*) som forklarende variable.
 - Angiv de estimerede parametre i modellen.
- b). Fortolk parametrene i den estimerede regressionsmodel.
 - Angiv den forventede forskel i den samlede el-produktion sommer og vinter.
 - Angiv den forventede forskel i den samlede el-produktion forår og efterår.
- c). **Opskriv** nulhypotesen om at årstiden ikke har nogen betydning for den samlede elproduktion (når vi tager højde for forskelle i gennemsnitstemperaturen).
 - **Gør rede for** om der er statistisk belæg for en påstand om, at årstiden ikke har nogen betydning for den samlede el-produktion (når vi tager højde for forskelle i gennemsnitstemperaturen) (brug 5% som signifikansniveau).



d). **Gør rede for** om der er belæg for en påstand om, at der ikke er forskel på den forventede el-produktion sommer og vinter (når vi tager højde for forskelle i gennemsnitstemperaturen) (brug 5% som signifikansniveau).

Gør rede for om der er belæg for en påstand om, at der ikke er forskel på den forventede el-produktion forår og efterår (når vi tager højde for forskelle i gennemsnitstemperaturen) (brug 5% som signifikansniveau).

EKSEMPEL 3 (SAMLET EL-PRODUKTION)

- a). **Tegn** en figur der viser, om der ser ud til at være en forskel på gennemsnitstemperaturens betydning for den samlede el-produktion i GigaWatt-timer i de forskellige årstider.
- b). **Opskriv** en regressionsmodel med den samlede el-produktion som responsvariabel og årstid, gennemsnitstemperatur og en interaktionseffekt mellem gennemsnitstemperatur og årstid som forklarende variable.
 - Angiv de estimerede parametre i modellen.
- c). Fortolk parametrene i den estimerede regressionsmodel.
- d). **Opskriv** nulhypotesen om at gennemsnitstemperaturens betydning for den samlede el-produktion er den samme for alle fire årstider.
 - **Gør rede for** om der er statistisk belæg for en påstand om, at gennemsnitstemperaturens betydning for den samlede el-produktion er den samme for alle fire årstider (brug 5% som signifikansniveau).

EKSEMPEL 4 (SAMLET EL-PRODUKTION)

- a). **Reducér** en lineær regressionsmodel med den samlede el-produktion i GigaWatt-timer som responsvariabel og variablene Å*r*, Måned, Middeltemperatur DK C, Middelvind m/s, Årstid som forklarende variable til kun at indeholde signifikante variable (brug 5% som signifikansniveau).
 - **Angiv** forklaringsgrad og justeret forklaringsgrad hørende til den fundne regressionsmodel.
- b). **Beregn** på baggrund af den i delspg. a) fundne regressionsmodel et 95%-konfidensinterval for den forventede betydning for den samlede el-produktion af en stigning i gennemsnitstemperaturen på 2 grader Celcius og **giv** en fortolkning af intervallet.



- c). **Beregn** på baggrund af den i delspg. a) fundne regressionsmodel et 95%-konfidensinterval for den forventede forskel i den samlede el-produktion mellem forår og vinter og **giv** en fortolkning af intervallet.
 - **Beregn** på baggrund af den i delspg. a) fundne regressionsmodel et 95%-konfidensinterval for den forventede forskel i den samlede el-produktion mellem forår og efterår og **giv** en fortolkning af intervallet.
- d). **Beregn** på baggrund af den i delspg. a) fundne regressionsmodel den forventede samlede el-produktion i GigaWatt-timer for november 2018, forudsat at gennemsnitstemperaturen er 6 grader Celsius og middelvindhastigheden er 5 m/s.

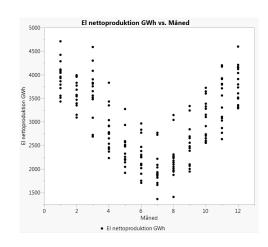


VEJLEDENDE LØSNINGER

EKSEMPEL 1 (► Løsning i JMP)

1a):

Figuren af den samlede el-produktion henover årets 12 måneder viser ikke overraskende en kvadratisk sammenhæng, hvor el-produktionen er højest i årets første og sidste måneder, mens den er lavest midt på året.



0.674587

0,67093

450,8937

2984,436

118,2683

42,07051

3,160792

Estimate Std Error t Ratio Prob>|t|

42,27

-19,20

18,85

<,0001*

<.0001*

■ Summary of Fit RSquare

Mean of Response

Observations (or Sum Wgts)

Parameter Estimates

4999,4785

-807,5659

59,577174

RSquare Adj Root Mean Square Error

Term

Intercept

1_B):

De estimerede parametre hørende til modellen

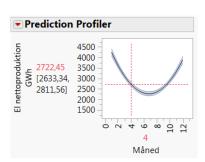
El nettoproduktion GWh
$$=\beta_0+\beta_{Md}\cdot \text{Måned}+\beta_{Md2}\cdot \text{Måned}^2+\varepsilon$$

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

$$\hat{eta}_0 = 4.999,48$$
 $\hat{eta}_{Md} = -807,57$ $\hat{eta}_{Md2} = 59,58$ $\hat{\sigma} = 450,89$

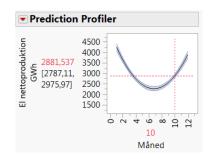
1c):

På baggrund af den i delspg. b) estimerede regressionsmodel er den forventede samlede månedlige elproduktion på 2.722,45 GigaWatt-timer i april måned.





På baggrund af den i delspg. b) estimerede regressionsmodel er den forventede samlede månedlige elproduktion på 2.881,54 GigaWatt-timer i oktober måned.





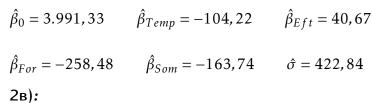
EKSEMPEL 2 (► Løsning i JMP)

2a):

De estimerede parametre hørende til modellen

El nettoproduktion GWh $= \beta_0 + \beta_{Temp} \cdot \text{Middeltemperatur DK C} \\ + \beta_{Eft} \cdot \text{Årstid=efterår} \\ + \beta_{For} \cdot \text{Årstid=forår} \\ + \beta_{Som} \cdot \text{Årstid=sommer} + \varepsilon$

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



Summary of Fit 0,717039 RSquare RSquare Adj 0,710608 Root Mean Square Error 422,8375 Mean of Response 2984,436 Observations (or Sum Wgts) 181 **Effect Tests** Sum of Source DF Squares F Ratio Prob > F Middeltemperatur DK C 16799000 93,9586 4,5494 2440189 Indicator Function Parameterization Estimate Std Error t Ratio Prob>|t| Intercept 3991.3304 65,1419 61,27 Middeltemperatur DK C -104,2227 10,75212 -9,69 <.0001 Arstid[Efterår] 40,67373 123,9695 0,33 0,7432 Årstid[Forår] -258,4756 107,731 0,0175* -2,40 -163,7403 Arstid[Sommer] 178.2416 0.3595

Den forventede samlede månedlige el-produktion...

- er på 3.991,33 GigaWatt-timer om vinteren, såfremt gennemsnitstemperaturen er 0 grader Celsius
- falder med 104,22 GigaWatt-timer for hver 1 grad Celsius den månedlige gennemsnitstemperatur stiger
- er 40,67 GigaWatt-timer højere om efteråret end om vinteren
- er 258,48 GigaWatt-timer lavere om foråret end om vinteren
- er 163,74 GigaWatt-timer lavere om sommeren end om vinteren

Endvidere vil den samlede månedlige el-produktion med ca. 95% sandsynlighed variere $med \pm 2 \cdot 422, 84 = \pm 845, 68$ GigaWatt-timer efter at vi har korrigeret for gennemsnitstemperaturens og årstidens indvirkning på/sammenhæng med el-produktionen.



Den forventede el-produktion er 163,74 GigaWatt-timer lavere om sommeren end om vinteren.

Den forventede el-produktion er -258,48-(40,67)=-299,15 GigaWatt-timer lavere om foråret end om efteråret.

Indicator Function Parameterization									
Term	Estimate	Std Error	t Ratio	Prob> t					
Intercept	4032,0041	122,9286	32,80	<,0001*					
Middeltemperatur DK C	-104,2227	10,75212	-9,69	<,0001*					
Årstid[Forår]	-299,1494	92,70289	-3,23	0,0015*					
Årstid[Sommer]	-204,414	112,1036	-1,82	0,0699					
Årstid[Vinter]	-40,67373	123,9695	-0,33	0,7432					

2c):

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

$$H_0: \ \beta_{Eft} = \beta_{For} = \beta_{Som} = 0$$
 og $H_a: \ Mindst \ \acute{e}n \ af \ parametrene \ \beta_{Eft}, \beta_{For}, \beta_{Som} \ er \ forskellig \ fra \ 0$

forkastes idet P-værdien er 0,43%.

Effect Tests										
Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F					
Middeltemperatur DK C Årstid	1 3	1	16799000 2440189	93,9586 4,5494	<,0001* 0.0043*					

Der er således ikke på baggrund af datamaterialet belæg for at hævde, at årstiden ikke har en sammenhæng med den samlede el-produktion. Med andre ord, der <u>er</u> forskel på den samlede el-produktion henover de fire årstider, selv når vi korrigerer for forskelle i gennemsnitstemperaturen.

2_D):

Test af nulhypotesen om der ikke er forskel på den samlede månedlige el-produktion sommer og vinter (når vi tager højde for forskelle i gennemsnitstemperaturen)

$$H_0: \beta_{Som} = 0$$
 og $H_a: \beta_{Som} \neq 0$

forkastes ikke, idet P-værdien er 35,95%.



Indicator Function Parameterization									
Term	Estimate	Std Error	t Ratio	Prob> t					
Intercept	3991,3304	65,1419	61,27	<,0001*					
Middeltemperatur DK C	-104,2227	10,75212	-9,69	<,0001*					
Årstid[Efterår]	40,67373	123,9695	0,33	0,7432					
Årstid[Forår]	-258,4756	107,731	-2,40	0,0175*					
Årstid[Sommer]	-163,7403	178,2416	-0,92	0,3595					

Der er således ikke på baggrund af datamaterialet belæg for at afvise en påstand om, at den samlede el-produktion er den samme sommer og vinter, bortset fra forskelle i gennemsnitstemperaturen. Med andre ord, at forskellen i el-produktion sommer og vinter alene skyldes temperaturforskelle.

Test af nulhypotesen om der ikke er forskel på den samlede månedlige el-produktion forår og efterår (når vi tager højde for forskelle i gennemsnitstemperaturen)

$$H_0: \beta_{For} = \beta_{Eft}$$
 og $H_a: \beta_{For} \neq \beta_{Eft}$

forkastes, idet P-værdien er 0,15%.

Indicator Function Parameterization										
Term	Estimate	Std Error	t Ratio	Prob> t						
Intercept	4032,0041	122,9286	32,80	<,0001*						
Middeltemperatur DK C	-104,2227	10,75212	-9,69	<,0001*						
Årstid[Forår]	-299,1494	92,70289	-3,23	0,0015*						
Årstid[Sommer]	-204,414	112,1036	-1,82	0,0699						
Årstid[Vinter]	-40,67373	123,9695	-0,33	0,7432						

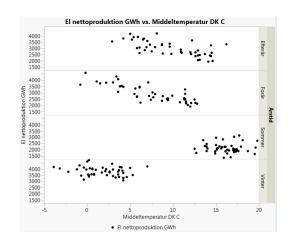
Der er således ikke på baggrund af datamaterialet belæg for at hævde, at den samlede elproduktion er den samme forår og efterår, bortset fra forskelle i gennemsnitstemperaturen. Med andre ord, der er ikke belæg for at hævde, at forskellen i el-produktion forår og efterår alene skyldes temperaturforskelle.



EKSEMPEL 3 (► Løsning i JMP)

3a):

Figuren af sammenhængen mellem den samlede elproduktion og gennemsnitstemperaturen indenfor hver af de fire årstider viser, at det primært ser ud til at være forår og efterår, at gennemsnitstemperaturen indvirker på el-produktionen (der er en nedadgående tendens forår og efterår, mens der ikke er nogen klar retning i observationerne for sommer og vinter).



3в):

De estimerede parametre hørende til modellen (hvor variablen Middeltemperatur DK C for nemheds skyld er forkortet Temp)

El nettoproduktion GWh

$$=\beta_0+\beta_{Temp}\cdot Temp\\ +\beta_{Eft}\cdot \mathring{A}rstid=efter\mathring{a}r\\ +\beta_{For}\cdot \mathring{A}rstid=for\mathring{a}r\\ +\beta_{Som}\cdot \mathring{A}rstid=sommer\\ +\beta_{Temp,Eft}\cdot Temp,\mathring{A}rstid=efter\mathring{a}r\\ +\beta_{Temp,For}\cdot Temp,\mathring{A}rstid=for\mathring{a}r\\ +\beta_{Temp,Som}\cdot Temp,\mathring{A}rstid=sommer+\varepsilon$$

Summary of Fit								
RSquare	0,752857							
RSquare Adj	0,742857							
Root Mean Square Error	398,5816							
Mean of Response	2984,436							
Observations (or Sum Wgts)	181							
Effect Tests								
			Su	ım of				
Source	Nparm	DF	Sq	uares	FI	Ratio	Pro	b > F
Middeltemperatur DK C	1	1	5579	260,7	35	,1190	<,0	0001*
Årstid	3	3	2106	250,9	4	,4193	0,0	0051*
Middeltemperatur DK C*Årst	id 3	3	3983	267,8	8	,3576	<,(0001*
Indicator Function Par	ameteriza	tion						
Term		Esti	mate	Std Er	ror	t Rat	io P	rob>
Intercept		386	2,978	73,958	334	52,2	23	<,0001
Middeltemperatur DK C		-31,1	L5085	25,563	308	-1,2	22	0,2247
Årstid[Efterår]		264,4	19073	193,5	152	1,3	37	0,1735
Årstid[Forår]		184	,8004	151,50	006	1,2	22	0,2242
Årstid[Sommer]			8,282	564,54		-3,:		0,0023
Middeltemperatur DK C*Årstid[Efterår]			79774	30,80		-2,0		0,0079
Middeltemperatur DK C*Årst			,3498	30,08		-3,8		0,0002
Middeltemperatur DK C*Årst		22.0/	59968	42,920	-0-	0,7		0.4421

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

$$\hat{\beta}_{0} = 3.862,98 \qquad \hat{\beta}_{Temp} = -31,15 \qquad \hat{\beta}_{Eft} = 264,49 \qquad \hat{\beta}_{For} = 184,80 \qquad \hat{\beta}_{Som} = -1.748,28$$

$$\hat{\beta}_{Temp,Eft} = -82,80 \qquad \hat{\beta}_{Temp,For} = -115,35 \qquad \hat{\beta}_{Temp,Som} = 33,07 \qquad \hat{\sigma} = 398,58$$



3c):

Den forventede samlede månedlige el-produktion...

- er på 3.862,98 GigaWatt-timer om vinteren, såfremt gennemsnitstemperaturen er 0 grader Celsius
- falder med 31,15 GigaWatt-timer om vinteren for hver 1 grad Celsius den månedlige gennemsnitstemperatur stiger
- er alt andet lige 246,49 GigaWatt-timer højere om efteråret end om vinteren¹
- er alt andet lige 184,80 GigaWatt-timer højere om foråret end om vinteren
- er alt andet lige 1.748,28 GigaWatt-timer lavere om sommeren end om vinteren
- falder med 82,80 GigaWatt-timer mere om efteråret end <u>om vinteren</u> (dvs. falder med i alt 31,15+82,80=113,95) for hver 1 grad Celsius den månedlige gennemsnitstemperatur stiger
- falder med 115,35 GigaWatt-timer mere om foråret end <u>om vinteren</u> (dvs. falder med i alt 31,15+115,35=146,50) for hver 1 grad Celsius den månedlige gennemsnitstemperatur stiger
- falder med 33,07 GigaWatt-timer mindre om sommeren end <u>om vinteren</u> (dvs. falder med i alt 31,15-33,07=-1,92, dvs. stiger med i alt 1,92) for hver 1 grad Celsius den månedlige gennemsnitstemperatur stiger

Endvidere vil den samlede månedlige el-produktion med ca. 95% sandsynlighed variere $med \pm 2 \cdot 398, 58 = \pm 797, 16$ GigaWatt-timer efter at vi har korrigeret for gennemsnitstemperaturens og årstidens indvirkning på/sammenhæng med el-produktionen.

¹Med "alt andet lige" menes: "efter vi har korrigeret for en eventuel forskel i el-produktionen, der skyldes forskel i gennemsnitstemperaturen mellem efterår og vinter".



3_D):

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

$$H_0: \beta_{Temp,Eft} = \beta_{Temp,For} = \beta_{Temp,Som} = 0$$

 H_a : Mindst én af parametrene $\beta_{Temp,Eft}$, $\beta_{Temp,For}$, $\beta_{Temp,Som}$ er forskellig fra 0

forkastes idet P-værdien er mindre end 0,01%.

Effect Tests											
Source	Nparm	DF	Sum of Squares	F Ratio	Prob						
Middeltemperatur DK C	1	1	5579260,7	35,1190	<,00						
Årstid	3	3	2106250,9	4,4193	0,00						
Middeltemperatur DK C*Årstid	3	3	3983267,8	8,3576	<,00						

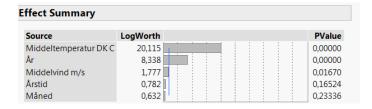
Der er således ikke på baggrund af datamaterialet belæg for at hævde, at gennemsnitstemperaturen påvirker den samlede el-produktion på samme måde i de fire årstider. Med andre ord, der <u>er</u> forskel på hvor meget ændringer i gennemsnitstemperaturen påvirker den samlede el-produktion henover de fire årstider.



EKSEMPEL 4 (► Løsning i JMP)

4_A):

Variablen Måned viser sig at være insignifikant med en P-værdi på 0,2333, og kan derfor udelades af modellen.



Alle øvrige forklarende variable forbliver signifikante i modellen, efter at Måned er fjernet.

ffect Summary						
Source	LogWorth					PValue
Middeltemperatur DK C	19,914					0,00000
År	8,426					0,00000
Årstid	2,338					0,00459
Middelvind m/s	1,576					0,02654

De estimerede parametre hørende til den reducerede model

El nettoproduktion GWh

$$= \beta_0 + \beta_{\mathring{A}r} \cdot \mathring{A}r$$

 $+\beta_{Temp} \cdot Middeltemperatur DK C$

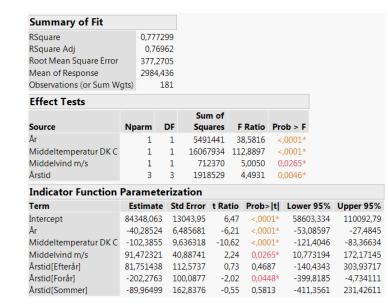
 $+\beta_{Vind} \cdot Middelvind m/s$

$$+\beta_{Eft} \cdot \text{Årstid} = \text{efterår}$$

$$+\beta_{For} \cdot \text{Årstid=forår}$$

$$+\beta_{Som} \cdot \text{Årstid} = \text{sommer} + \varepsilon$$

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



$$\hat{\beta}_0 = 84.348,06$$
 $\hat{\beta}_{\mathring{A}r} = -40,29$ $\hat{\beta}_{Temp} = -102,39$ $\hat{\beta}_{Vind} = 91,47$

$$\hat{\beta}_{Ar} = -40,29$$

$$\hat{\beta}_{Temp} = -102,39$$

$$\hat{\beta}_{Vind} = 91,47$$

$$\hat{\beta}_{Eft} = 81,75$$

$$\hat{\beta}_{Eft} = 81,75$$
 $\hat{\beta}_{For} = -202,28$ $\hat{\beta}_{Som} = -89,96$ $\hat{\sigma} = 377,27$

$$\hat{\beta}_{Som} = -89,96$$

$$\hat{\sigma} = 377, 27$$



4_B):

Et 95%-konfidensinterval for β_{Temp} er givet som [-121,40; -83,37]. Det betyder, at med 95% sandsynlighed vil den forventede el-produktion falde med mellem 83,37 og 121,40 GigaWatt-timer for hver gang gennemsnitstemperaturen stiger med 1 grad Celsius.

Et 95%-konfidensinterval for β_{Temp} ·2 er givet som $[-121,40\cdot2; -83,37\cdot2] = [-242,80; -166,74]$. Det betyder, at med 95% sandsynlighed vil den forventede el-produktion falde med mellem 166,74 og 242,80 GigaWatt-timer for hver gang gennemsnitstemperaturen stiger med 2 grader Celsius.

4c):

Et 95%-konfidensinterval for β_{For} er givet som [-399,82; -4,73]. Det betyder, at med 95% sandsynlighed vil den forventede el-produktion om foråret alt andet lige (dvs. når vi korrigerer for ændrer i modellens øvrige forklarende variable) være mellem 4,73 og 399,82 GigaWatt-timer lavere end om vinteren.

Et 95%-konfidensinterval for $\beta_{For} - \beta_{Eft}$ er givet som [-447,85; -120,20]. Det betyder, at med 95% sandsynlighed vil den forventede el-produktion om foråret alt andet lige (dvs. når vi korrigerer for ændringer i modellens øvrige forklarende variable) være mellem 120,20 og 447,85 GigaWatt-timer lavere end om efteråret.

Indicator Function Parameterization										
Term	Estimate	Std Error	t Ratio	Prob> t	Lower 95%	Upper 95%				
Intercept	84429,815	13040,84	6,47	<,0001*	58691,226	110168,4				
År	-40,28524	6,485681	-6,21	<,0001*	-53,08597	-27,4845				
Middeltemperatur DK C	-102,3855	9,636318	-10,62	<,0001*	-121,4046	-83,36634				
Middelvind m/s	91,472321	40,88741	2,24	0,0265*	10,773194	172,17145				
Årstid[Forår]	-284,0277	83,00545	-3,42	0,0008*	-447,8549	-120,2006				
Årstid[Sommer]	-171,7164	101,0075	-1,70	0,0909	-371,0741	27,641213				
Årstid[Vinter]	-81,75144	112,5737	-0,73	0,4687	-303,9372	140,43429				



4_D):

På baggrund af den i delspg. a) estimerede regressionsmodel er den forventede samlede månedlige elproduktion på 2.977,25 GigaWatttimer i november måned 2018, såfremt gennemsnitstemperaturen er 6 grader Celsius og middelvindhastigheden er 5 m/s.

