***Opgave 1***

*Anders Andersen kunne godt tænke sig at identificere de parametre, der afgør byggegrundens pris.*

***Spørgsmål 1.1 (20 %)***

*Du bedes derfor udarbejde en lineær regressionsmodel, der forklarer grundens pris ved hjælp af størrelse målt i m2, afstand til folkeskole målt i m og/eller fjordudsigt ja eller nej*

*Alle parametre i modellen skal være signifikante.*

*Fortolk modellens koefficienter.*

*Kan Anders Andersen anvende modellen til at forudsige parcelhusgrundens pris?*

Startmodel:

Vi kan af F-testet konkludere at mindst en forklarende variabel har effekt.

Vi fortsætter med partielle t-tests.

Vi har en p-værdi på 91.27%, vi kan altså ikke forkaste nulhypotesen, derfor udelades den forklarende variabel af modellen.

Vi fortsætter med reduktion til vi når frem til slutmodellen:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Summary of fit** |  | PI enkelt y | KI gns. y |
| R-kvadreret | 0.870797 |  | x-værdi | x-værdi |
| R-kvadreret justeret | 0.866997 | 2.10092204 |  |  |
| Standardfejl | 57.84 | -136.5582637 | 95% Nedre PI | 95% Nedre CI |
| Middel af respons | 610.4166667 | 106.4907002 | -169.6117 | -110.4842 |
| Observationer | 36 |  | 95% Øvre PI | 95% Øvre CI |
| **Analyse af varians** |  |  | 123.2221 | 64.0947 |
| Source | DF | Sum af kvadrater | Mean square | F-teststørrelse |
| Model | 1 | 766709.532 | 766710 | 229.1518 |
| Error SSE/SSR | 34 | 113759.218 | 3346 | p-værdi F-test fuld model |
| Total SST | 35 | 880468.750 |  | 0.00% |
|  |  |  |  |  |
| **Parameter estimater** | Estimater | Std. Fejl | t-teststørrelse | p-værdi koefficienter |
| Skæring | -23.19 | 42.9522 | -0.5400 | 59.27075% |
| x1 | 0.82 | 0.0539 | 15.1378 | 0.00000% |

Hver gang grundens størrelse vokser med 1 kvm, stiger grundprisen med 820 kr.

En grund på 0 kvm vil ifølge modellen koste -23190 kr. Der er tale om ekstrapolation, så resultatet er ikke understøttet af data.

Vi ser af R-kvadreret at 87% af variationen i responsvariablen grundpris kan forklares ved variationen i den forklarende variabel grundstørrelse.

Det er en relativ høj forklaringsgrad, kun 13% af variationen kan ikke forklares af modellen.

Anders Andersen kan altså godt bruge modellen til at forudsige grundprisen.

***Spørgsmål 1.2 (10 %)***

*Undersøg om forudsætningerne for den model, du har foreslået i spørgsmål 1.1 er opfyldt.*

*Forudsætningen om linearitet er opfyldt observationerne ligger pænt omkring den rette linje.*

****

Vi ser af residualplottet at forudsætningen om varianshomogenitet er opfyldt, der er ingen trompetform i residualplottet.

Vi ser af residualplottet at forudsætningen om uafhængighed er opfyldt der er ingen tydelig systematik i residualplottet.



Vi ser at normalfraktildiagrammet at forudsætningen om normalfordelte residualer er opfyldt, da punkterne ligger pænt om den rette linje.



**Opgave 2**

Anders Andersen har en teori om at den gennemsnitlige grundpris er steget siden finanskrisen.

En undersøgelse fra 2008 - ligeledes af 36 grundpriser - gav en gennemsnitlig grundpis på 550.000 kr. med en standardafvigelse på 130.100 kr.

**Spørgsmål 2.1 (10 %)**

Test på 5 % signifikansniveau om standardafvigelserne i de 2 undersøgelser er ens.

Vi tester om varianserne er ens for begge populationer.

Vi kan ikke forkaste nulhypotesen da p-værdien 24.6% er større end signifikansniveauet. Vi kan altså benytte Pooled t-test i den videre analyse, da der er ens varianser.

|  |  |  |  |
| --- | --- | --- | --- |
| Test af varianshomogenitet | |  |  |
| Teststatisic | 1.4862 |  |  |
| p-værdi | 24.60% |  |  |
| H0 | σ21 = σ22 |  |  |
| HA: | σ21 ≠ σ22 |  |  |
| Da p værdien/signifikanssandsynligheden 0.246 > 0.05 signifikansniveauet, kan vi ikke afvise nulhypotesen, varianserne er ens | | | |
|
|
| Forudsætning for test for varianshomogenitet er at begge populationer er normalfordelte | | | |

**Spørgsmål 2.2 (10 %)**

Test på 5 % signifikansniveau om den gennemsnitlige grundpris er steget.

Vi tester

|  |  |  |
| --- | --- | --- |
| Et-sidet alternativ hypotese | |  |
| p-værdi | 4.08% |  |
| H0 | μ1 -μ2 ≤ D |  |
| HA: | μ1 -μ2 > D |  |
| Da p værdien/signifikanssandsynligheden 0.0408 < 0.05 signifikansniveauet, afviser vi nulhypotesen | | |
|

Vi forkaster nulhypotesen da p-værdien er 4.08% hvilket er mindre end signifikansniveauet. Den gennemsnitlige grundpris er altså steget fra 2008 til 2015.

**Spørgsmål 2.3 (10 %)**

Redegør for signifikansniveauets betydning for din konklusion.

Bemærk konklusionen er niveaufølsom, havde vi testet på 1% signifikansniveau, kunne vi ikke forkaste nulhypotesen, og havde nået den modsatte konklusion, da p-værdien så ville være større end 1%.

***Opgave 3***

*Anders Andersen er meget interesseret i grunde med kort afstand til skole.*

***Spørgsmål 3.1 (10 %)***

*Udarbejd et 95 % konfidensinterval for andelen af grunde, der har under 650 m til skole?*

Vi benytter test af andele, vi kan af nedenstående plot se at forudsætningen for approximation til normalfordelingen ikke helt er opfyldt.

|  |  |  |  |
| --- | --- | --- | --- |
| Succeser | Stikprøve str. | Stikprøve andel | n·p·(1-p)>9 Forudsætningen er ikke opfyldt |
| 17 | 36 | 0.4722 | 8.9722 |
| Nedre grænse 95 KI andelen p | |  | Fejlmargin |
| 0.3091 |  |  | 16.31% |
| Øvre grænse 95 KI andelen p | |  |  |
| 0.6353 |  |  |  |

Vi kan med 95% sikkerhed sige at andelen af huse med mindre end 650 meter til skole i populationen ligger mellem 30,91% og 63,53%

***Spørgsmål 3.2 (10 %)***

*Test ved et 5 % signifikansniveau om den gennemsnitlige skoleafstand er større end 650 m.*

Da p-værdien 2.68% er mindre end signifikansniveauet på 5% forkaster vi nulhypotesen.

Den gennemsnitlige afstand er altså større end 650 meter.

|  |  |
| --- | --- |
| Et-sidet alternativ hypotese |  |
| p-værdi | 2.68% |
| H0 | μ ≤ 650 |
| HA: | μ > 650 |
| Da p værdien/signifikanssandsynligheden 0.0268 < 0.05 signifikansniveauet, afviser vi nulhypotesen | |
|

***Spørgsmål 3.3 (10 %)***

*Udarbejd et 90 % konfidensinterval for den gennemsnitlige skoleafstand.*

Vi kan med 90% sikkerhed sige at den gennemsnitlige afstand til skole i populationen ligger mellem 667 meter og 852 meter.

|  |  |
| --- | --- |
| **Nedre grænse 90% KI middelværdien μ** | **666.9471** |
| **Øvre grænse 90% KI middelværdien μ** | **852.4974** |

**Opgave 4**

Anders Andersen har en fornemmelse af, at der bliver færre og færre grunde til salg med fjordudsigt.

I 2008 lavede han en undersøgelse, der ligeledes omfattede 36 grunde. Heraf havde 17 fjordudsigt.

**Spørgsmål 4.1 (10 %)**

Test på 5 % signifikansniveau om andelen af grunde med fjordudsigt er faldende.



Vi benytter test af 2 andele, bemærk forudsætningerne om approximation til normalfordelingen er ikke opfyldt for de 2 stikprøver.

|  |  |  |  |
| --- | --- | --- | --- |
| Et-sidet alternativ hypotese | | |  |
| Prob>z = | 16.948992% |  |  |
| H0 | p1 - p2 ≤ D |  |  |
| HA: | p1 - p2 > D |  |  |
| Da p værdien/signifikanssandsynligheden 0.1695 > 0.05 signifikansniveauet, kan vi ikke afvise nulhypotesen. | | | |
|

Vi kan ikke forkaste nulhypotesen da p-værdien 16.95% er større end 5% signifikansniveauet.

Andelen af grunde med fjordudsigt er ikke faldende, Anders Andersen har altså ikke ret i sin fornemmelse.på fornel