Forløb om lineær regression afholdt i 2i-MA (mat A, samf A)

|  |  |
| --- | --- |
| Tidslinje (et modul er på 1,5 timer) | Aktivitet |
| 1. modul, mindste kvadraters metode | Lineær regression vises i Nspire på de fire datapunkter i eksemplet fra noterne ”regression\_og\_forklaringsgrad\_version5”. Dernæst vises residualkvadraterne med sum . Tendenslinjen slettes og en flytbar linje med residualkvadrater introduceres.  **Opgave**. Kan du få ned på 1,8?  **Opgave**. Gør den flytbare linje vandret. Vis at minimummet for har . Vis at denne vandrette linje skærer *y*-aksen i gennemsnittet . Beregn forklaringsgraden *r*2 vha. tallene 1,8 og 26.  **Opgave**. Ekstraopgave 2 i noterne  ”regression\_og\_forklaringsgrad\_version5” hvor du beviser at er minimal når den vandrette linje har .  Med Nspire og NetLogo programmerne ”mindste\_kvadrater” vises det til sidst, hvordan *a* og *b* for tendenslinjen kan findes ved at steppe henimod minimummet for . |
| 2. modul | Lineær regression vises på data i solfangereksemplet. Her er *x* opvarmningstiden i minutter og *y* er temperaturen i °C.  **Opgave**. Vis at formlerne for *a*, *b* og *r*2 giver Nspires værdier for lineær regression på data i solfangereksemplet.  Det vises hvordan residualspredningen bestemmes i beregneren og i regnearket som outputtet *s* fra ”statistik -> konfidensintervaller -> lineær reg t-intervaller”. Dernæst vises en normalfordelingstest af residualerne i et ”diagrammer og statistik” vindue. Den bagvedliggende teori om standardnormalfordelingen af repeteres.  Det opklares at residualspredningen er relateret til spredningen i normalfordelingstesten ved .  **Opgaver** fra bogen. |
| 3. modul | Elevgennemgang af bogens opgaver fra sidste gang.  NetLogo programmet ”fordeling\_a\_b\_y” introduceres. Alle *y*’er er normalfordelte omkring den sande linje. Det vises at der opnås en pæn fordeling for en variabel *ta* der er knyttet til hældningen *a*. Den præcise definition af *ta* gemmes til næste modul. 95% konfidensintervallet for hældningen *a* bestemmes for data i solfangereksemplet vha. ”statistik -> konfidensintervaller -> lineær reg t-intervaller”.  **Opgaver** fra bogen. |
| 4. modul | Elevgennemgang af bogens opgaver fra sidste gang.  **Opgave**. Lav arbejdsarket ”t-fordelinger\_opgaver”.  NetLogo programmet ”fordeling\_a\_b\_y” introduceres, men denne gang med formler for *ta* og *tb*.  **Opgave**. Undersøg vha. programmet ”fordeling\_a\_b\_y” om fordelingerne for *ta* og *tb* er t-fordelinger med *n* - 2 frihedsgrader.  Det vises hvordan man selv kan beregne tallene i outputtet fra ”statistik -> konfidensintervaller -> lineær reg t-intervaller”.  Det gennemgås at en ikke-sammenhæng følger fra eller .  **Opgaver** fra bogen. |
| 5. modul, test for | Elevgennemgang af bogens opgaver fra sidste gang.  Der vises en test for på data i ”Regn i Voer” eksemplet. Her er *x* tiden i år efter 2009 og *y* er regnmængden i mm per år. Der anvendes en tosidet ”statistik -> statistiske tests -> lineær regressions t-test”.  Dernæst gives der en forklaring på hvad der ligger bagved testen og forbindelsen til 95% konfidensintervallet for hældningen *a*.  **Opgaver** fra bogen. |