

11 - Hvor god er modellen?

Når vi har lavet lineærregression kan vi blive bedt om at vurdere om en model er god.

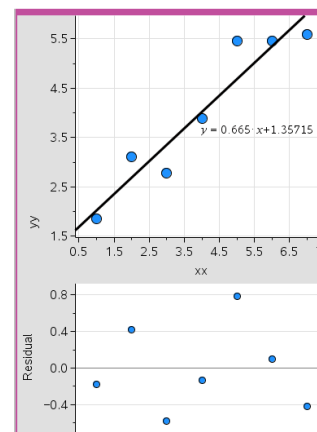
Residualplottet alene siger sjældent noget præcist om en model er god, men ser vi på både dataplot og residualplot sammen kan man godt vurdere modellen.

For at modellen skal være god, så skal punkterne ligge tæt omkring linjen i dataplottet/punktplottet

- Det er det samme som at *residualerne* er ”små”.
- Du kan se på *variationsbredden* for *y*-værdierne (forskellen på den største og den mindste *y*-værdi) og sammenligne med *variationsbredden* for residualerne (forskellen på den største og den mindste residual).
Variationsbredden for *y*-værdierne skal være meget større end for residualerne.

For at modellen skal være god, så skal residualerne (de tilfældige fejl i målingerne) ligge tilfældigt omkring linjen i residualplottet

- Hvis de ikke ligger tilfældigt siger vi, at der er en *systematisk afvigelse*

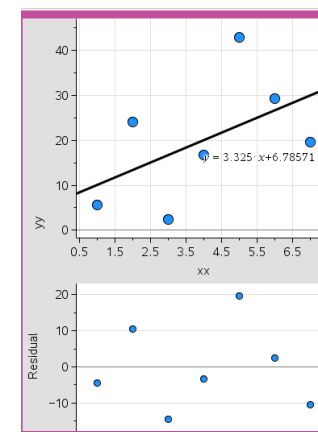


Variationsbredden i *y*'erne (ca):

Variationsbredden i *residualerne* (ca):

Er residualerne fordelt tilfældigt?

Er modellen god?

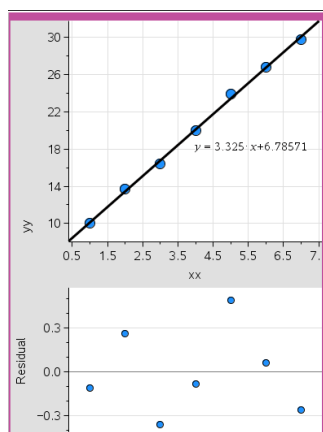


Variationsbredden i *y*'erne (ca):

Variationsbredden i *residualerne* (ca):

Er residualerne fordelt tilfældigt?

Er modellen god?

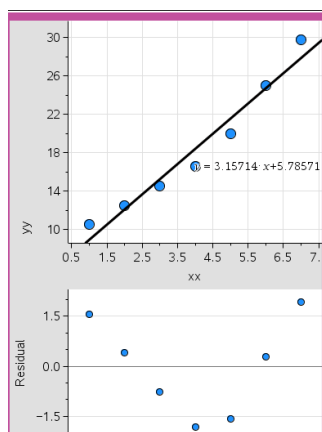


Variationsbredden i *y*'erne (ca):

Variationsbredden i *residualerne* (ca):

Er residualerne fordelt tilfældigt?

Er modellen god?

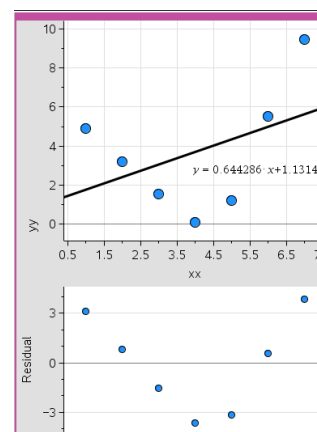


Variationsbredden i *y*'erne (ca):

Variationsbredden i *residualerne* (ca):

Er residualerne fordelt tilfældigt?

Er modellen god?

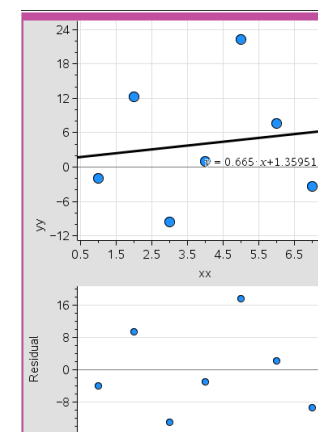


Variationsbredden i *y*'erne (ca):

Variationsbredden i *residualerne* (ca):

Er residualerne fordelt tilfældigt?

Er modellen god?



Variationsbredden i *y*'erne (ca):

Variationsbredden i *residualerne* (ca):

Er residualerne fordelt tilfældigt?

Er modellen god?