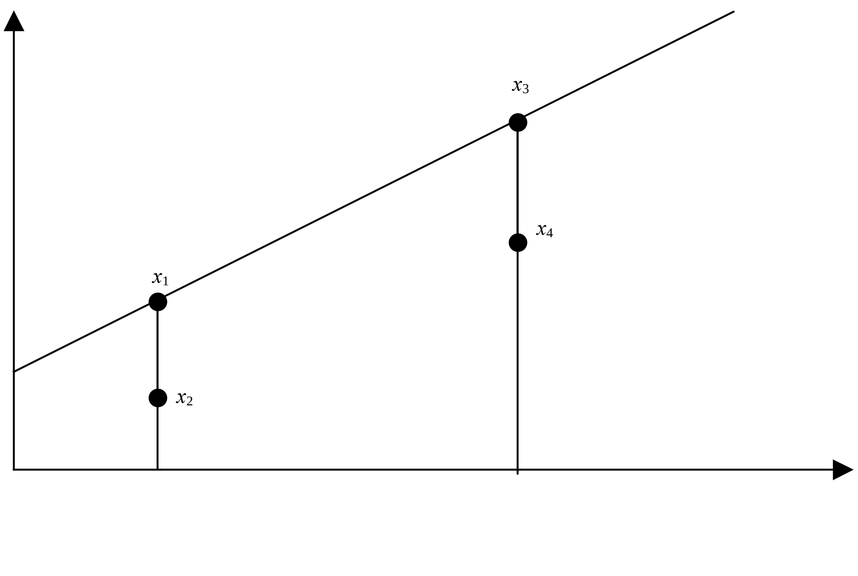
Kontrolskema for regressionsmodel med 1 uafhængig variabel

Hypotese	Grafiske metoder	Numeriske metoder
Lineær hypotese: $E[y] = \beta_0 + \beta_1 \cdot x_i$	- Direktørplot: $(y_i, \hat{y}_i)$ skal være ret linie med hældning 1. - $(x_i, \hat{e}_i^*)$-plottet må ikke have "systematiske krumninger".	- R^2 (forklaringsgraden). Jo tættere på 1, jo bedre. $R^2 \in [0;1]$ $R^2 = \rho^2(x_i, y_i)$
Restled normalfordelte $e_i \sim \text{Nf}(0, \cdot)$	- e_i'erne plottes i normalfraktildiagram, og skal danne en ret linie.	
Homoskedasticitet $\text{var}(e_i) = \sigma^2$ (konstant og dermed uafhængig af x_i)	- $(x_i, \hat{e}_i^*)$-plottet må ikke udvise tendens til hverken stigende eller aftagende varians med stigende x_i.	- Ved flere y'er end x'er kan Bartlett's test anvendes (kap. 10).
Restled uafhængige $\text{cov}(e_i, e_j) = 0$ Vi tester kun for den specielle afhængighed: AR(1)-proces. Derfor kun relevant ved **tidsserier**.	- (tid, $\hat{e}_t^*$) må ikke udvise tegn på autokorrelation, som altså er den eneste type afhængighed vi tester for.	- Durbin-Watson (stadig udelukkende ved tidsserier)
Signifikans Bidrager x_i signifikant til forklaring af y		- t-test for hypotesen $H_0 : \beta_0 = 0$ mod passende alternativ.
Afvigende observationer Er der observationer med stor indflydelse på estimationen af β_0 og β_1?		- Cook's afstandsmål samt Mahalanobis afstandsmål, (og man kan desuden se på de standardiserede residualer, hvor en tommelfingerregel er at $\lvert\hat{e}_i^*\rvert < 2$)

Ad afvigende observationer:

Hvis man i nedenstående figur forestiller sig, at de punkter som regressionen bygger på ligger i nærheden af x_1, så gælder der følgende om de to afstandsmål for de fire punkter x_1, x_2, x_3 og x_4:



	Cook's afstandsmål	**Mahalanobis' afstandsmål**
x_1	Lille	Lille
x_2	Middelstor	Lille
x_3	Lille	Stor
x_4	Stor	Stor