Variansstabiliserande transformasjonar.

Eit alternativ til generaliserte lineære modellar er å bruke varians-stabiliserande transformasjonar på responsvariabelen. Tanken er at ved å gjere variansen nær konstant vil ein normalt vere nærare å ha normalfordelte variable.

La $Y = g(X)$. Då har vi $Y = g(\mu) + g'(X)\big|_{X=\mu}(X-\mu) + R$

og $\text{Var}(Y) \approx (g'(X)\big|_{X=\mu})^2\,\text{Var}(X)$.

For at $\text{Var}(Y) \approx k$ må vi ha at $(g'(X)\big|_{X=\mu})^2 \approx \dfrac{k}{\text{Var}(X)}$

Nokre situasjonar .

$\text{VAR}(X) \propto \mu \Rightarrow g'(\mu) \approx \dfrac{k'}{\sqrt{\mu}}$. Ein eigna transformasjon bør vere $g(X) = \sqrt{X}$.

$\text{VAR}(X) \propto \mu^2 \Rightarrow g'(\mu) \approx \dfrac{k'}{\mu}$. Ein eigna transformasjon bør vere $g(X) = \mathit{ln}(X)$.

$\text{VAR}(X) \propto \mu^4 \Rightarrow g'(\mu) \approx \dfrac{k'}{\mu^2}$. Ein eigna transformasjon bør vere $g(X) = 1/X$.

Med konvensjonen at $X^\lambda = \mathit{ln}(X)$ for $\lambda = 0$ ser ein alle transformasjonane ovanfor er på forma X^λ. Målet blir altså å finne ein λ slik at variansen til X^λ blir konstant.

Sidan ein ofte ynskjer å bruke transformasjonar som ikkje gjer fortolkninga altfor komplisert er ein ofte brukt praksis å nytte den av dei 3 transformasjonane ovanfor der variansen til residuala ser ut til å vere mest mogeleg konstant.

<u>Eks.1</u> $X_i \sim \text{Pois}(\mu_i)$. Då er $P(X_i = x) = \dfrac{\mu_i^x e^{-\mu_i}}{x!}$. $E(X_i) = \text{VAR}(X_i) = \mu_i$

For store nok μ_i bør ein eigna transformasjon vere $g(X_i) = \sqrt{X_i}$.

<u>Eks. 2</u> $X_i \sim \exp(\mu_i)$. Då er $f_{X_i}(x) = \dfrac{1}{\mu_i} e^{-\frac{x}{\mu_i}}$, $x > 0$. $E(X_i)^2 = \text{VAR}(X_i) = \mu_i^2$.

Ein mogeleg transformasjon er difor $g(X_i) = \mathit{ln}(X_i)$.