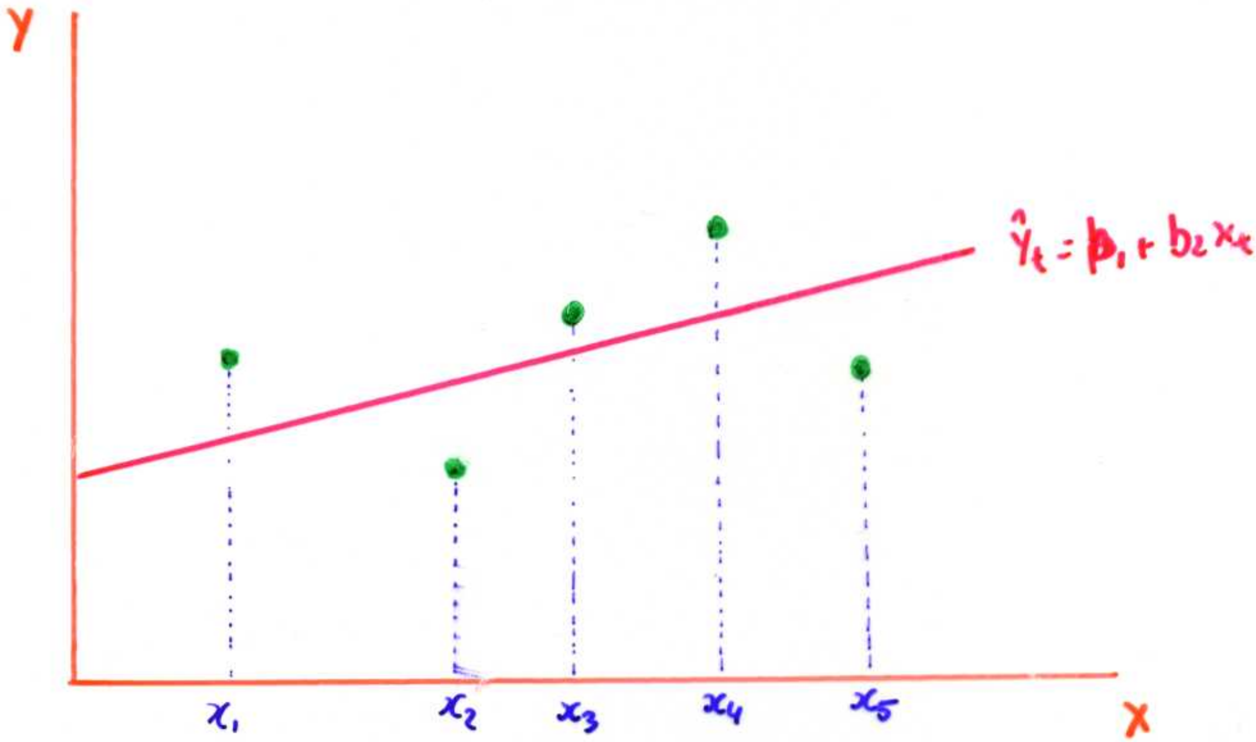
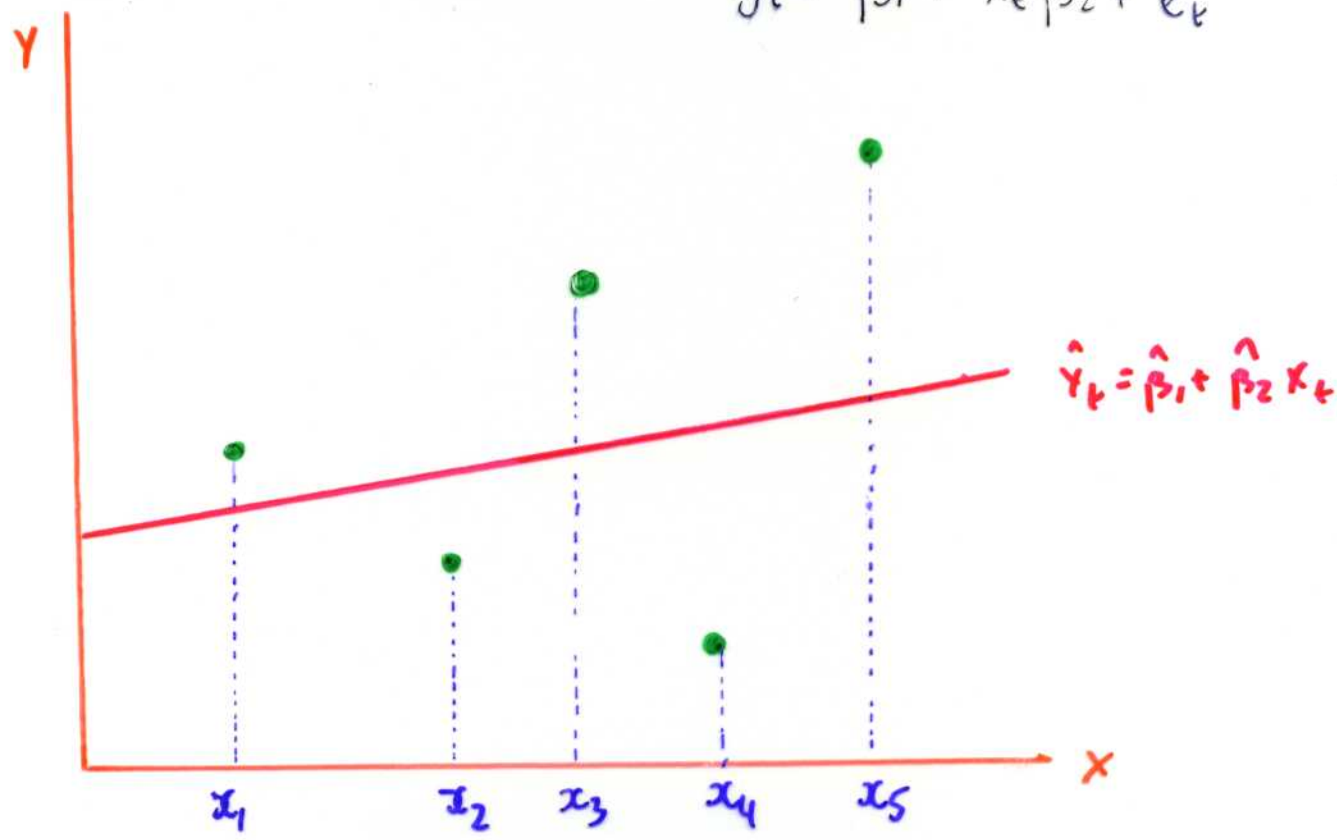
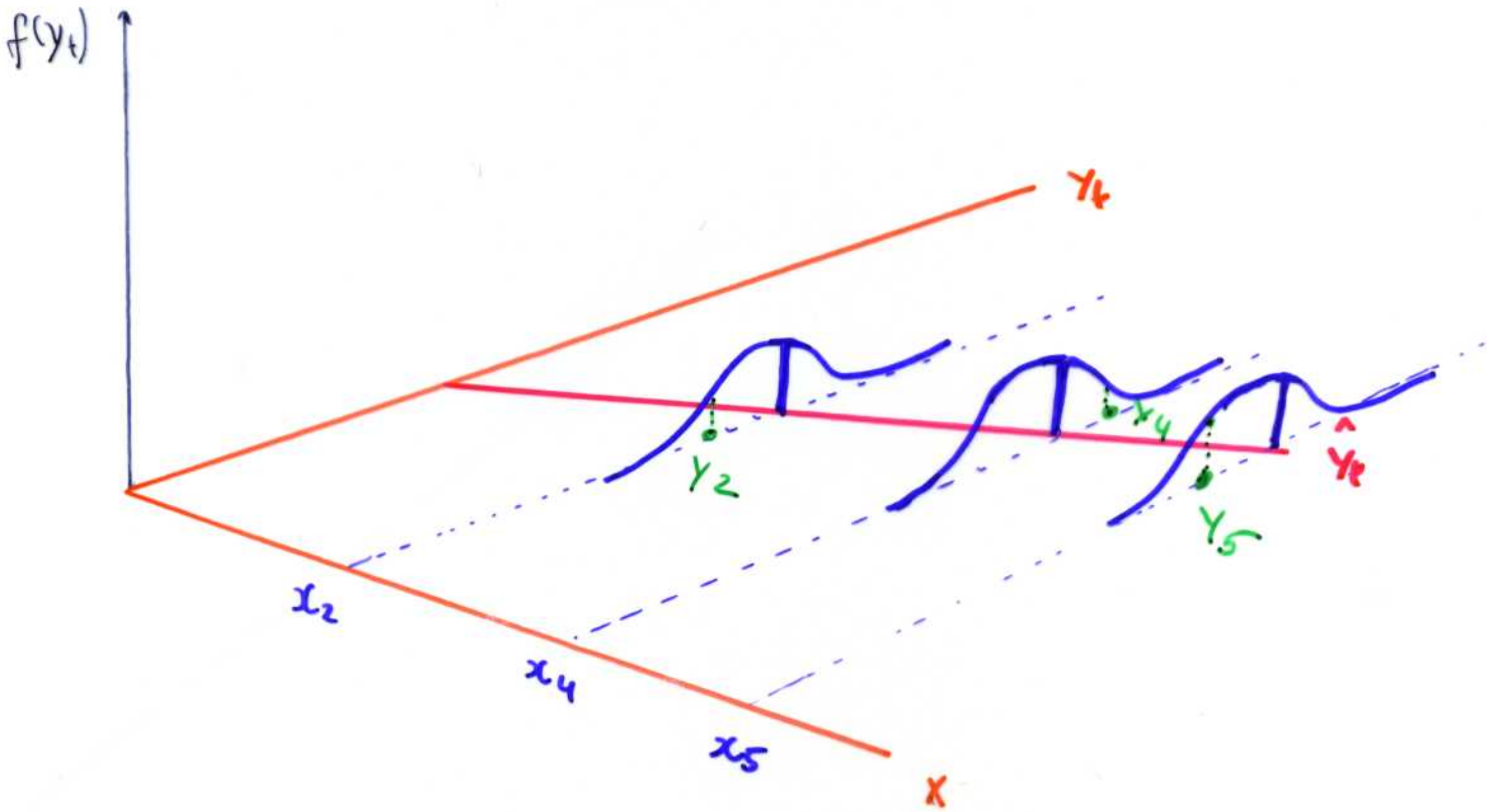
HETEROSKEDASTIC ERROR MODEL

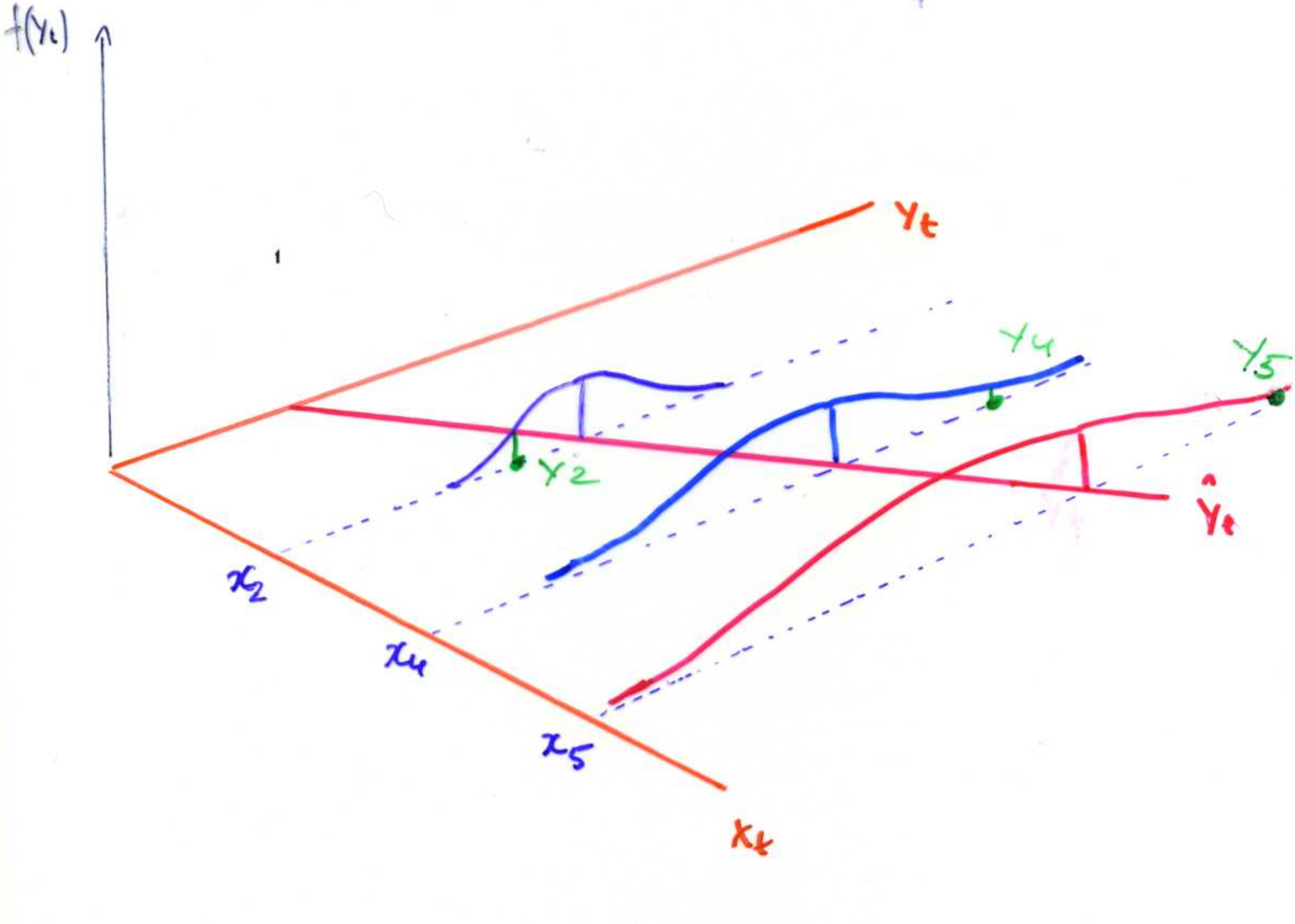
1. Homoskedastic Linear Model : $y_t = \beta_1 + x_t\beta_2 + e_t$



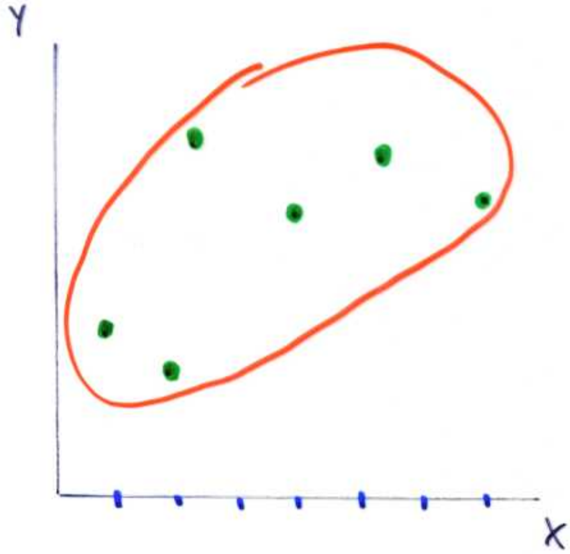
2. Heteroskedastic Linear Model: $y_t = \beta_1 + x_t\beta_2 + \tilde{e}_t$

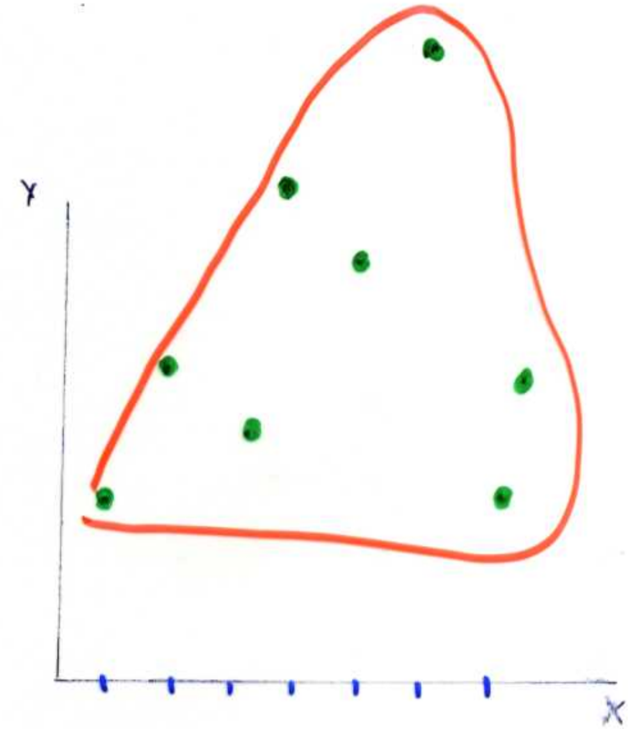



f(yt)
Yt
Y2
Y4
Y5
Ŷt
x2
x4
x5
X


f(yt)
Yt
Y4
Y5
Y2
Ŷt
x2
x4
x5
Xt

To visualize:





assumptions:

$$var(e_t) = \sigma_t^2$$

say, that in our case variance increases proportionally to the explanatory x_t's. We write it

$$\sigma_t^2 = \sigma^2 \cdot x_t$$

OUR APPROACH WILL BE TO TRANSFORM THE HETEROSKEDASTIC MODEL INTO A HOMOSKEDASTIC MODEL.

CONSIDER: $y_t = \beta_1 + \beta_2 x_t + e_t$

WHERE $e_t \sim (0, \sigma^2 x_t)$

ERRORS ARE ASSUMED ALSO UNCORRELATED,

$E(e_t e_s) = 0$ for $t \neq s$

TO MAKE THINGS EASY WE ASSUME

X is NONSTOCHASTIC !

transformation:

$$\mathrm{var}(e_t) = \sigma^2 x_t \text{ , hence } \mathrm{var}\left(\frac{e_t}{\sqrt{x_t}}\right) = \frac{1}{x_t} \cdot \mathrm{var}(e_t) =$$

$$= \frac{1}{x_t} \cdot \sigma^2 x_t = \sigma^2$$

now:

$$y_t = \beta_1 \cdot 1 + \beta_2 x_t + e_t$$

$$\Downarrow$$

$$\frac{y_t}{\sqrt{x_t}} = \beta_1 \frac{1}{\sqrt{x_t}} + \beta_2 \frac{x_t}{\sqrt{x_t}} + \frac{e_t}{\sqrt{x_t}}$$

$$\boxed{y_t^* = \beta_1 x_{1t}^* + \beta_2 x_{2t}^* + e_t^*}$$

is a transformed model with

$$e_t^* \sim (0, \sigma^2)$$

We use simply OLS to estimate this model, i.e. we build new regressand $y_t/\sqrt{x_t}$ and as regressors use vectors of $\frac{1}{\sqrt{x_t}}$ and $\sqrt{x_t}$.

- IF WE DON'T ACCOUNT FOR HETEROSKEDASTICITY THE OLS produces wrong variance estimator σ^2:

⇒ THE ESTIMATOR'S variances, i.e. $var(b_1)$ and $var(b_2)$ ARE WRONG

⇒ CONFIDENCE INTERVALS WRONG (NOT VALID)

⇒ HYPOTHESES TEST NOT VALID

- THE TRANSFORMATION ABOVE YIELDS AN ESTIMATOR

$$\hat{\beta} = \left(X^{*\prime}X^{*}\right)^{-1} X^{*\prime}y^{*}$$

which theoretically is obtained by minimizing

$$\sum e_t^{*2} = \sum \left(\frac{e_t^2}{x_t}\right)$$

and can be interpreted as an WEIGHTED LEAST SQUARES

$\text{cov}(\hat{\beta}) = \sigma^2 (X^{*\prime}X^*)^{-1}$

where $\hat{e}^* = y^* - X^*\hat{\beta}$ and $\hat{\sigma}^2 = \dfrac{\hat{e}^{*\prime}\hat{e}^*}{T-K}$

$\hat{\beta}$ is the minimum variance linear unbiased. (NOTE: AGAIN WE HAVE EXACT PROPERTIES, AS X'S NONSTOCHASTIC)

obviously, it yields valid tests and confidence intervals. and IS MORE EFFICIENT THAN OLS.

A HETEROSKEDASTIC PARTITION

consider example on page 491. We estimate a model

$$q_t = \beta_1 + \beta_2 p_t + \beta_3 t + e_t$$

Where t is trend, ie: a sequence 1, 2, 3, 4 for years (dates) at which variables were observed. It is a "time series" type of model where obs on quantities in time are recorded and not accross individuals.

We believe that in the sample of 26 obs (for 26 years) the variance of q_t in the first 13 years was different as in the last 13 years.

hence:

$E(e_t) = 0$

$var(e_t) = \sigma_1^2$ for $t = 1, \dots, 13$

$var(e_t) = \sigma_2^2$ for $t = 14, \dots, 26$

I split the model

$$\begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_{13} \\ \hline q_{14} \\ \vdots \\ q_{26} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ \hline 1 \\ \vdots \\ 1 \end{bmatrix} \beta_1 + \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_{13} \\ \hline p_{14} \\ \vdots \\ p_{26} \end{bmatrix} \beta_2 + \begin{bmatrix} 1 \\ 2 \\ \vdots \\ 13 \\ \hline 14 \\ \vdots \\ 26 \end{bmatrix} \beta_3 + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_{13} \\ \hline e_{14} \\ \vdots \\ e_{26} \end{bmatrix}$$

$$q_t = \beta_1 + p_t \beta_2 + t \beta_3 + e_t$$

$t = 1, \dots 26$

into 2 regressions

- for $t = 1, \dots 13$:

$$q_t^1 = \beta_1^1 + p_t^1 \beta_2^1 + t^1 \beta_3 + e_t^1$$

- for $t = 14, \dots 26$

$$q_t^2 = \beta_1^2 + p_t^2 \beta_2^2 + t^2 \beta_3 + e_t^2$$

This is not an optimal procedure as both subsamples yield different estimates: (by OLS)

$$b^1 = (X^{1\prime}X^1)^{-1} X^{1\prime} q^1 \quad \text{and} \quad b^2 = (X^{2\prime}X^2)^{-1} X^{2\prime} q^2$$

which are different and don't exploit the entire info in the entire (whole) sample. Besides you can't combine them optimally.

This sample split has been done to highlight the next test and estimation procedures, namely:

1) GOLDFELD - QUANDT TEST to detect heteroskedasticity

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_1: \sigma_2^2 < \sigma_1^2$$

i.e. the variance in both subsamples is same against heteroskedasticity presence.

$$GQ = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} \sim F_{[(T_1 - K_1), (T_2 - K_2)]}$$

reject H_0 when $GQ > F_c$. THIS AGAIN IS AN EXACT TEST AS REGRESSORS ARE ASSUMED NONSTOCHASTIC

WE USE FOR THE ONE SIDED TEST F_c at $\alpha = 0.05$

$\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ are estimated from the 2 subsamples of 13 obs in our example

HOW TO APPLY THE GQ TEST TO DETECT HETEROSKEDASTICI. OF TYPE $\sigma_t^2 = \sigma^2 \cdot x_t$?

(one regressor only: order all observations by value of x_t from smallest to longest, take 2 samples of equal number of obs from the top and the bottom of the sample, ex. the first and the last 20 obs in a sample of 50 obs, compute variances in subsamples by OLS, apply test.)

NOTE FOR:

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_A \quad \sigma_1^2 \neq \sigma_2^2$$

build GQ as ratio : $\frac{\text{large}}{\text{small}}$ variance and use $F[(T_1 - K_1)(T_2 - K_2)]$ at $\alpha = 0.025$, REJECT H_0 when

$$GQ > F_c$$

if you have heteroskedasticity due to more than one regressor you can't make the ordering, and should USE:

BREUSCH-PAGAN TEST to detect heteroskedasticity

variance of type

$$\sigma_t^2 = \underbrace{\alpha_1 + \alpha_2 Z_{t2} + \alpha_3}_{\text{coefficients}} Z_{t3}$$

where Z_{t2}, Z_{t3} any variable, like a regressor in the model, but not necessary, $\alpha_1, \alpha_2, \alpha_3$ are coefficients

$$H_0: \alpha_2 = \alpha_3 = 0$$

$$H_1: \alpha_1 \text{ or } \alpha_2 \text{ or } (\alpha_1 \text{ and } \alpha_2) \text{ are zero.}$$

if H_0 true $\sigma_t^2 = \alpha_1$ and hence the model is homoskedastic

STEPS TO FOLLOW:

1. estimate the model by OLS $\Rightarrow \hat{e} = y - Xb$, and $\tilde{\sigma}^2 = \Sigma \hat{e}_t^2 / T$.

2. $\sigma_t^2 = E(e_t^2)$, hence $\sigma_t^2 = \hat{e}_t^2 + v_t$. write

$$\hat{e}_t^2 = \alpha_1 + \alpha_2 Z_{t2} + \alpha_3 Z_{t3} + v_t$$

run this regression and keep SSR (sum of squared errors) note : if z_{t2}, z_{t3} are good explanatory variables of the variance, the ESS (explained sum of squares) is LARGE

IF z_{t2}, z_{t3} are poor explanatory variables, it is not true that $\sigma_t^2 = \alpha_1 + \alpha_2 z_{t2} + \alpha_3 z_{t3}$ and ESS (explained s.s) is SMALL.

3. compute

$$BP = \frac{ESS}{2\tilde{\sigma}^4}$$

ASYMPTOTICALLY VALID TEST! (ASYMPTOTIC TEST)

ASYMPTOTICALLY $BP \sim \chi^2_{(s)}$

where s = number of regressors in the variance model other than intercept ($s=2$ here)

REJECT H_0 WHEN

$$BP > \chi^2_c \qquad \text{at } \alpha = 0.05$$

GENERALIZED LEAST SQUARES (GLS)

$$y = X\beta + e$$

$E(e) = 0$

$cov(e) = E(ee') = W = \sigma^2 V$

(holds for any type of heteroskedasticity as V can be matrix) In our first example: $\sigma_t^2 = \sigma^2 x$ can be written as:

$$cov(e) = E(ee') = W = \begin{bmatrix} \sigma^2 x_1 & 0 & 0 \\ 0 & \sigma^2 x_2 & 0 \\ 0 & 0 & \sigma^2 x_3 \end{bmatrix} = \sigma^2 \begin{bmatrix} x_1 & 0 & 0 \\ 0 & x_2 & 0 \\ 0 & 0 & x_3 \end{bmatrix}$$

$$= \sigma^2 \cdot V$$

for a sample of 3 obs. Hence $V = \begin{bmatrix} x_1 & 0 & 0 \\ 0 & x_2 & 0 \\ 0 & 0 & x_3 \end{bmatrix}$.

WE USE THE FOLLOWING RESULT:

THERE ALWAYS EXIST A TRANSFORMATION MATRIX P SUCH THAT $P'P = V^{-1}$

and the error defined as

$$e^* = Pe$$

is homoskedastic.

P is $(T \times T)$. PREMULTIPLY THE MODEL:

$$Py = PX\beta + Pe$$

to get a homoskedastic model:

$$y^* = X^*\beta + e^*$$

where $y^* = Py$, $X^* = PX$ and $e^* = Pe$

GLS:

$$\boxed{\hat{\beta} = (X^{*\prime}X^*)^{-1}X^{*\prime}y^*}$$

can also be written as:

$$\hat{\beta} = (X^{*\prime}X^*)^{-1}X^{*\prime}y^* = (X'P'PX)^{-1}X'P'Py =$$

$$\boxed{(X'V^{-1}X)^{-1}X'V^{-1}y}$$

and since $W^{-1} = \frac{1}{\sigma^2} V^{-1}$ and $V^{-1} = \sigma^2 W^{-1}$

$$\hat{\beta} = (X' \sigma^2 W^{-1} X)^{-1} X' \sigma^2 W^{-1} Y$$

$$= \frac{1}{\sigma^2} (X' W^{-1} X)^{-1} X' \sigma^2 W^{-1} Y$$

$$\boxed{\hat{\beta} = (X' W^{-1} X)^{-1} X' W^{-1} Y}$$

GLS is Min Var. linear unbiased.

$$\underline{var(\hat{\beta})} = \sigma^2 (X^{*\prime} X^*)^{-1} = \sigma^2 (X' V^{-1} X)^{-1} = \underline{(X' W^{-1} X)^{-1}}$$

it accounts for the heteroskedasticity.

homoscedastic model

X are "CONSTANT"

$$y_t = \beta_1 + \beta_2 x_t + e_t$$

$$\text{var } e_t = \sigma^2 \quad \Rightarrow \underline{\text{var}(e) = \sigma^2 I_T}$$

ex $T=4$

$$\begin{bmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{bmatrix} = \text{var}(\underline{\underline{e}})$$

heteroscedastic model 1

$$y_t = \beta_1 + \beta_2 x_t + e_t$$

$$\text{var } e_t = \sigma^2 x_t$$

ex $T=4$

$$\text{var}(e) = \sigma^2 \begin{bmatrix} x_1 & 0 & 0 & 0 \\ 0 & x_2 & 0 & 0 \\ 0 & 0 & x_3 & 0 \\ 0 & 0 & 0 & x_4 \end{bmatrix} = \underline{W = \sigma^2 \cdot V}$$

heteroscedastic model 2

$$y_t = \beta_1 + \beta_2 x_t + e_t$$

$$\text{var } e_t = \sigma_1^2 \quad \text{for } t = 1, 2$$

$$\text{var } e_t = \sigma_2^2 \quad \text{for } t = 3, 4$$

ex $T=4$

$$\text{var}(e) = \begin{bmatrix} \boxed{\sigma_1^2 I_2} & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} & \boxed{\sigma_2^2 I_2} \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_1^2 & 0 & 0 \\ 0 & 0 & \sigma_2^2 & 0 \\ 0 & 0 & 0 & \sigma_2^2 \end{bmatrix} = W$$

In general: $Y = X\beta + e$

1. When errors e are homoscedastic

$$e \sim (0, \sigma^2 I_T)$$

use OLS estimator

$$b = (X'X)^{-1}X'Y$$

2. When errors are heteroscedastic

$$e \sim (0, W)$$

$$\hat{\beta}_{GLS} = \left(X'W^{-1}X\right)^{-1} X'W^{-1}Y$$

theoretical GLS

- unbiased
- efficient in the class of linear estimators of β

$$Y = X\beta + e \quad , \quad e \sim (0, W)$$

$$\hat{\beta}_G = (X'W^{-1}X)^{-1} X'W^{-1}Y = \underline{(X'W^{-1}X)^{-1} X'W^{-1}(X\beta + e)}$$

$$= (X'W^{-1}X)^{-1}(X'W^{-1}X)\beta + (X'W^{-1}X)^{-1}X'W^{-1}e$$

$$= \beta + (X'W^{-1}X)^{-1}\underline{X'W^{-1}e}$$

$$E\hat{\beta}_G = \beta$$

$$\text{var } \hat{\beta}_G = E\left[(\hat{\beta}_G - \beta)(\hat{\beta}_G - \beta)'\right]$$

$$= (X'W^{-1}X)^{-1} X'W^{-1} \underbrace{Eee'}_{W} W^{-1}X(X'W^{-1}X)^{-1}$$

$$= (X'W^{-1}X)^{-1} X'W^{-1} \cancel{W} \cancel{W^{-1}} X (X'W^{-1}X)^{-1}$$

$$= (X'W^{-1}X)^{-1}$$

In practice W is almost always unknown.

Therefore, the theoretical GLS NEEDS TO BE REPLACED BY

FEASIBLE GLS

$$\text{FGLS: } \hat{\hat{\beta}} = \left(X'\hat{W}^{-1}X\right)^{-1} X'\hat{W}^{-1}Y$$

where the true, unknown W is replaced by $\hat{W}$

we can hope that FGLS is consistent

Its asymptotic efficiency depends on the estimator $\hat{W}$