

3/30

Final Model to interpret is forecast

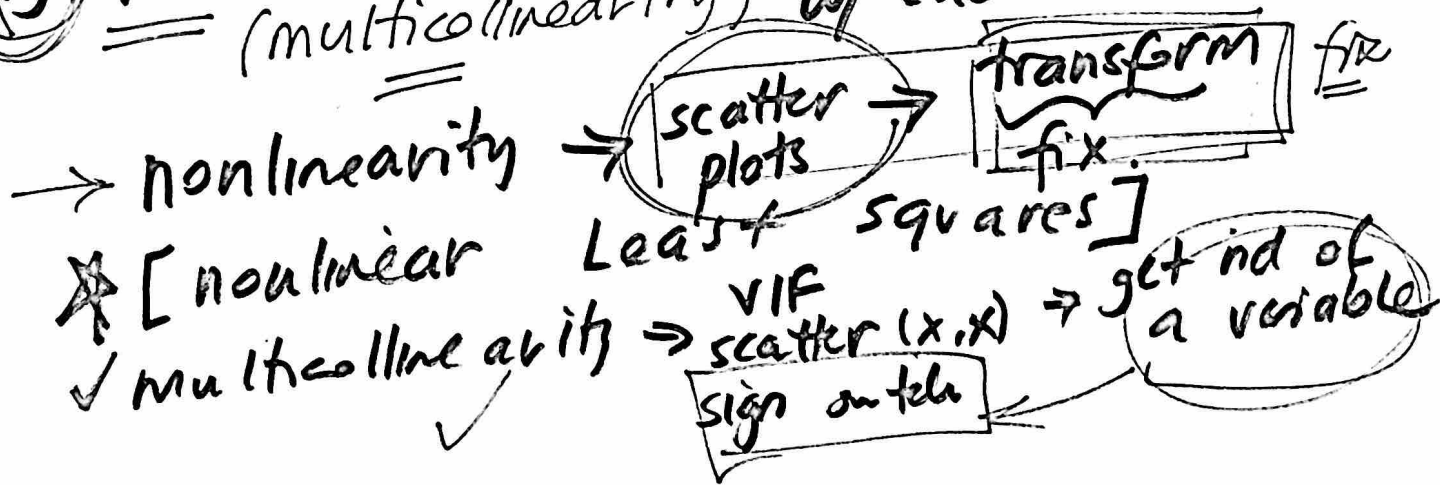
- No sign switch ✓
 - Everything significant ✓
 - Assumptions satisfied ✓
[or violations acknowledged]
- } ★

Today: ~~Go~~ ① Nonlinearity / transformations
→ ② DW test / corrections] Autocorr =
→ 3. Assumptions

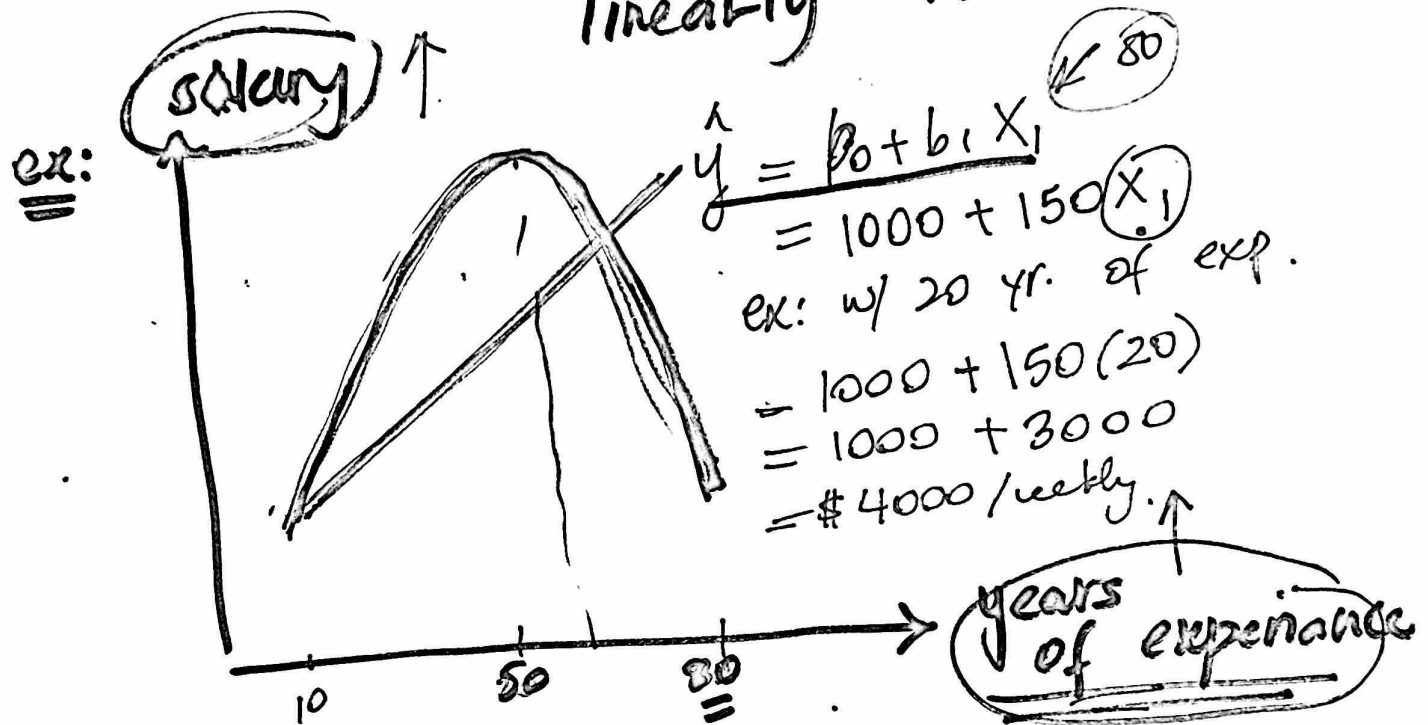
11

Assumptions CLRM (p: 230, 239, 339)

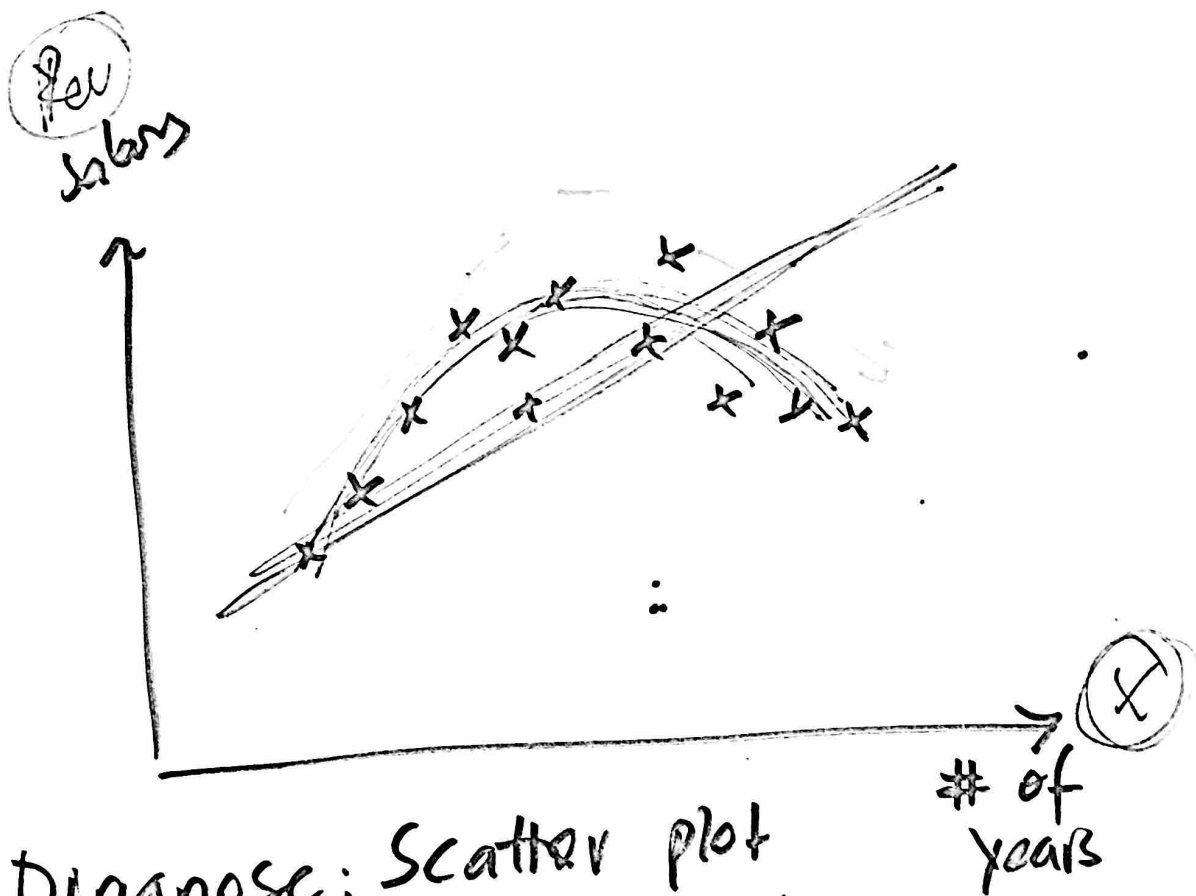
1. $E(e) = 0$ $X \leftarrow [4\text{-in-1 plots}] \searrow$
2. Variance (error) = σ^2 (constant) heteroskedasticity
3. Covariance (e_i, e_{i-1}) = 0 (autocorrelation) independent errors
4. X shld be linearly related to y (nonlinear)
5. X should NOT be highly correlated (multicollinearity) w/ each other.



Assumption #4: Y and X shld be linearly related.

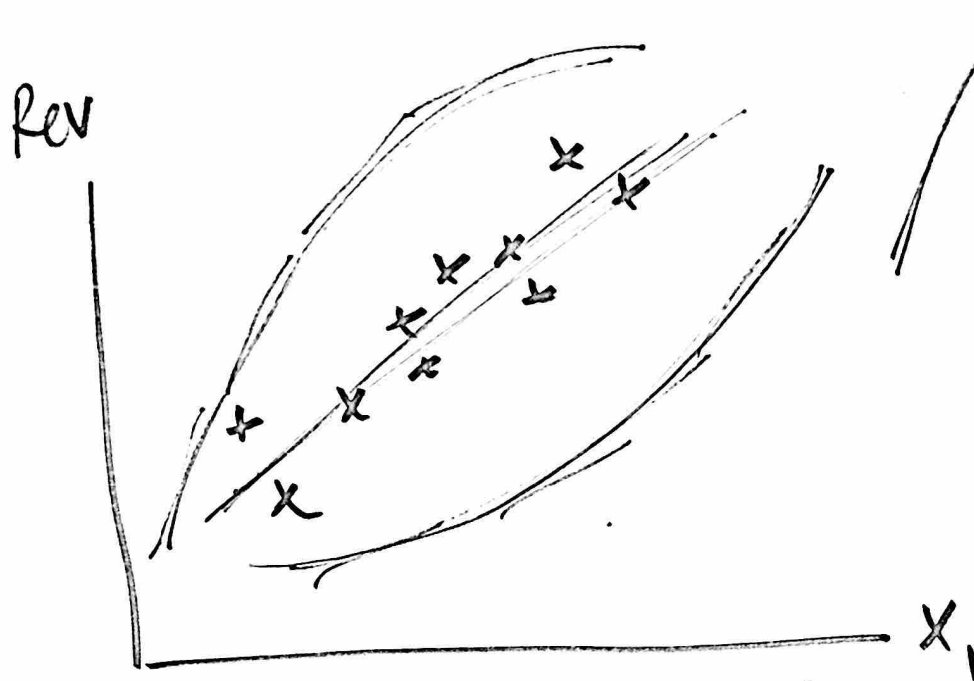


ex: 105 years of age
 $\Rightarrow$ 80 years of exp.



Diagnose: Scatter plot
Fix: Nonlinear Regression

47

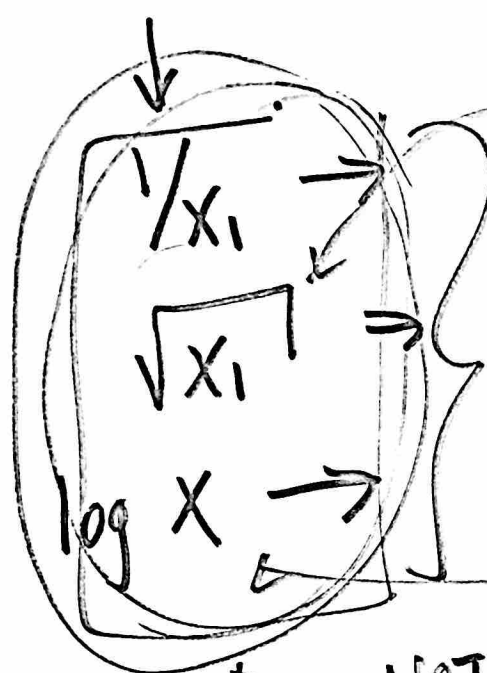


Any curve $\rightarrow$ nonlinearity
 $\hat{y} = b_0 + \underbrace{b_1}_{\text{slope}} X$

5]

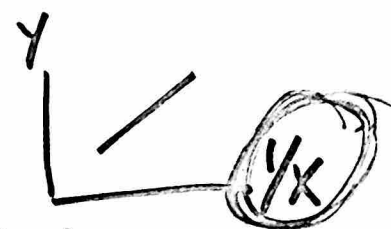
We will transfer X variable

$X_1 \rightarrow$ information



different
functions

of X



* Do NOT do more than 2-3 transformations

When you re-run the regression.

$$\hat{y} = b_0 + b_1 \left[\frac{1}{x_1} \right] + b_2 \underline{x_2} + b_3 \underline{x_3}$$

↑
click
on transformed X

~o~

#3 | Covariance ($\underline{e_t}, \underline{e_{t-1}} = 0$) } 4-in-1
(independent errors) } plot

DW test > Minitab has it

7] Violation: Autocorrelation

ϵ_t vs. ϵ_{t-1}

$\epsilon_t = \rho \epsilon_{t-1} + v_t$

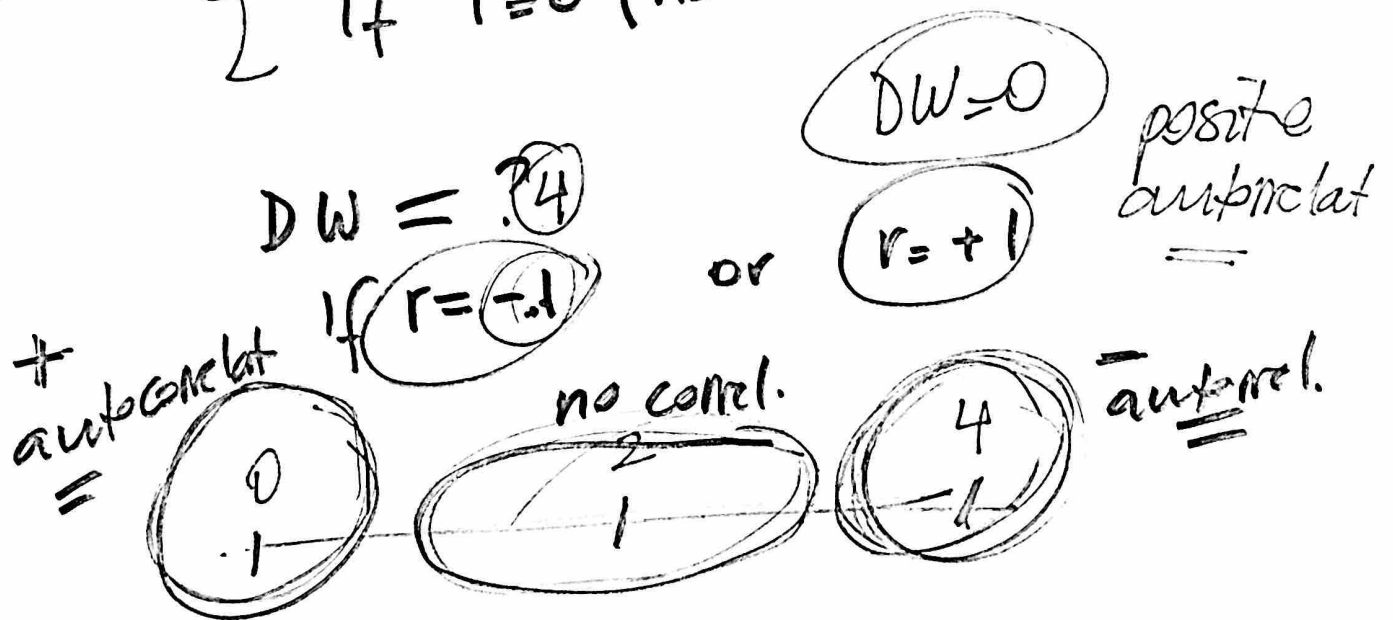
(DW test) ϵ_t dependent ρ slope ϵ_{t-1} indep. v_t error independent + random

$H_0: \rho = 0$ ✓ No autocorrelation
 $H_1: \rho \neq 0$ No relation btw ϵ_t & ϵ_{t-1}
 If I fail to reject $H_0 \Rightarrow$ we have autocorrelation

Min itab output $\Rightarrow \underline{DW} = 0 - 4$ 344

$$\underline{DW} = \underline{2} (1 - \underline{r_1}) \text{ correlation coef.}$$

no autocorrel $\left\{ \begin{array}{l} DW = ? 2 \\ \text{if } r=0 \text{ (no relation} \equiv \text{no autocorrel)} \end{array} \right.$



Why autocorrelation is ^{how to} fix it?

1. TS data $\rightarrow$ time series data
2. Omitted variables (X variable)
- $\{X \quad Y = b_0 + b_1 X_i$

Fix: Generalized least squares (GLS)
vs. OLS (ordinary least squares)

Q is D: $\rightarrow$ Introduce an autoregressive
lag
fix: If we have autocorr.
problem.

Auto. $\rightarrow$ regressive $\rightarrow$ lag.

Diag $\rightarrow$ DW
 Fix $\rightarrow$ w/ autoregressive
 recalc DW again
 fixed DW ≈ 2 ✓
 not $\rightarrow$ acknowledge it as
 on

$$y_t = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \text{error}$$

~~$\beta_3 y_{t-1}$~~

DW = + autocorrelation
 $DW \approx 0$

Minitab $>$ Lag $y_t \rightarrow y_{t-1}$

y_t	y_{t-1}
100	100
200	200
500	500
1000	1000

||