Multiple Logistic Regression (Extra)

Dr. Wan Nor Arifin

Unit of Biostatistics and Research Methodology,
Universiti Sains Malaysia.

wnarifin@usm.my / wnarifin.pancakeapps.com

Checking linearity in the logit

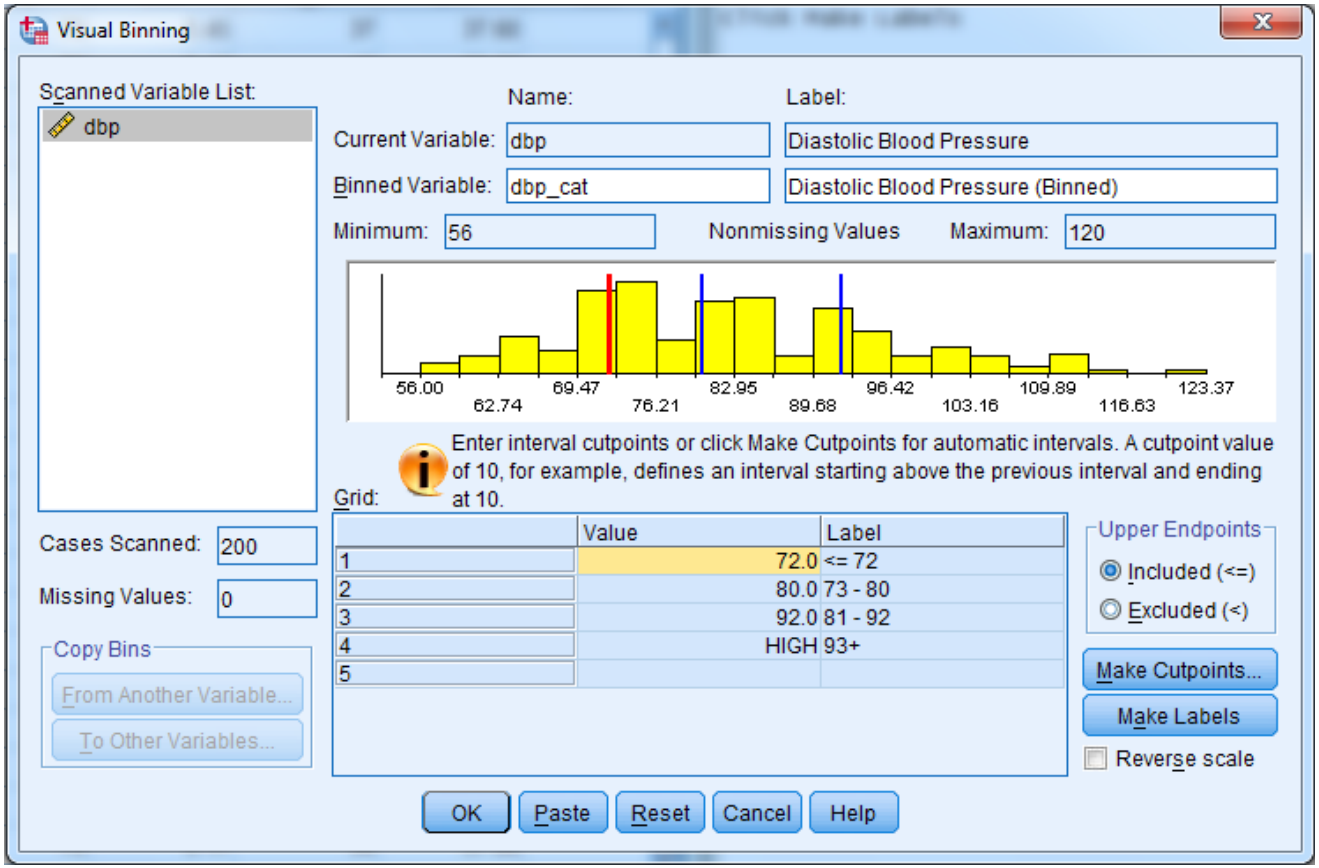
- **Checking linearity in the logit (by right in Step 2c)**
  1. Design variable approach
  2. Fractional polynomials (in STATA)

Checking linearity in the logit

- **Design variable approach:**
  1. Convert original continuous variable into 4-category categorical variable based on quartiles.
  2. Fit multivariable logistic regression model, replacing the continuous variable with the new categorical variable.
  3. Plot estimated coefficients vs group medians.

Checking linearity in the logit

- Create new variable:
  - **Transform → Visual Binning... → Variables to Bin:** *dbp* **→ Continue**
  - **Binning Variable:** Enter *dbp_cat*
  - Click **Make Cutpoints... → Select Equal Percentiles Based on Scanned Cases → Number of Cutpoints:** *3* **→ Apply**
  - **Click Make Labels → OK**



Checking linearity in the logit

- Perform Enter method with *dbp_cat* & *gender*
- Make sure to properly assign *dbp_cat* as categorical variable properly.
- Copy the results into an Excel sheet.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	gender(1)	.866	.394	4.832	1	.028	2.377	1.098	5.143
	dbp_cat			9.746	3	.021			
	dbp_cat(1)	1.427	.689	4.285	1	.038	4.167	1.079	16.093
	dbp_cat(2)	1.354	.697	3.778	1	.052	3.872	.989	15.165
	dbp_cat(3)	2.130	.693	9.444	1	.002	8.416	2.163	32.746
	Constant	-3.311	.660	25.181	1	.000	.036		

a. Variable(s) entered on step 1: gender, dbp_cat.

Checking linearity in the logit

- Obtain median of dbp for each dbp_cat group:
  - **Data → Split File →** Select **Compare groups**
  - Set **Groups Based on: dbp_cat → OK**
  - **Analyze → Descriptive Statistics → Frequencies**
  - **Variable(s): dbp →** Click **Statistics... → Median under Central Tendency**
  - Copy the results into the Excel sheet.

Statistics

Diastolic Blood Pressure

1	N	Valid	51
		Missing	0
	Median		68.00
2	N	Valid	53
		Missing	0
	Median		76.00
3	N	Valid	55
		Missing	0
	Median		86.00
4	N	Valid	41
		Missing	0
	Median		100.00

Checking linearity in the logit

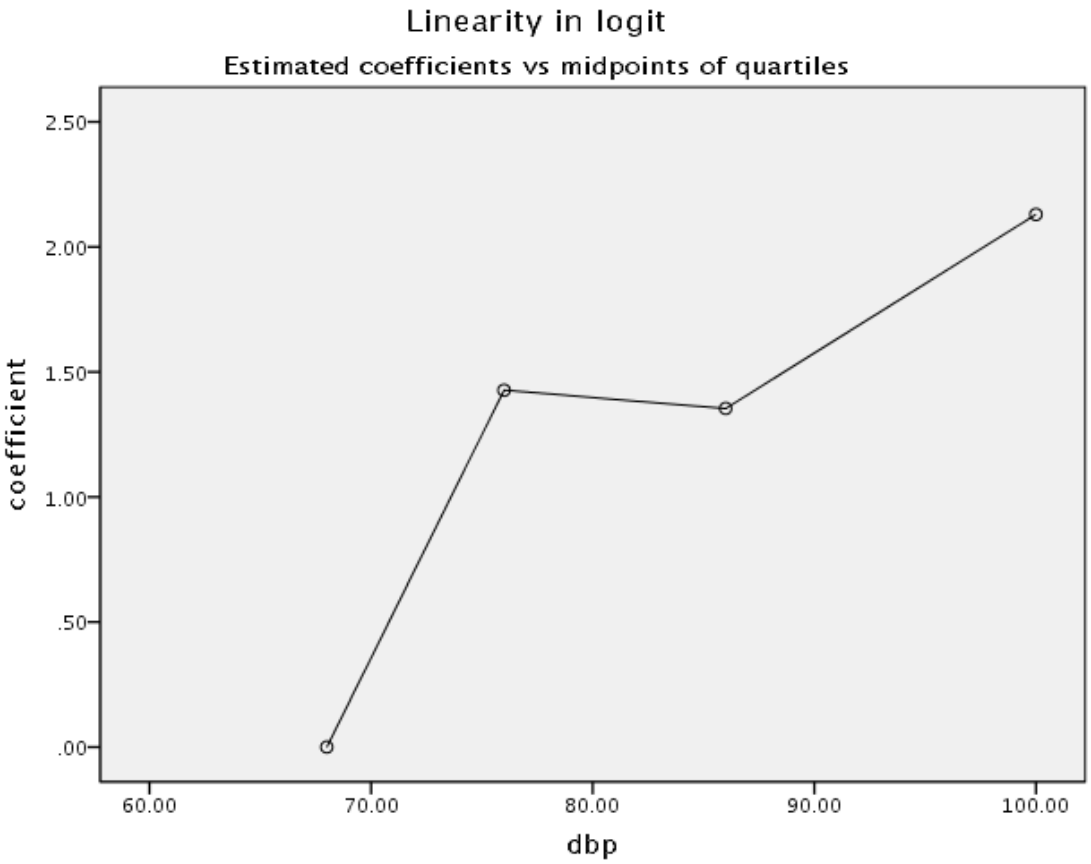
- **Cont…**
  - Copy relevant values as follows in Excel (*design_var.xls*). *Set "0" for the first group.
  - Then create a new SPSS dataset (**File → New → Data**)
  - Copy the values into SPSS Data View.
  - Rename the *VAR00001* & *VAR00002* as *coefficient* and *dbp*.

	M	N
	coefficient	median
	0	68
	1.427	76
	1.354	86
	2.13	100

	coefficient	dbp	va
1	.00	68.00	
2	1.43	76.00	
3	1.35	86.00	
4	2.13	100.00	
5			
6			

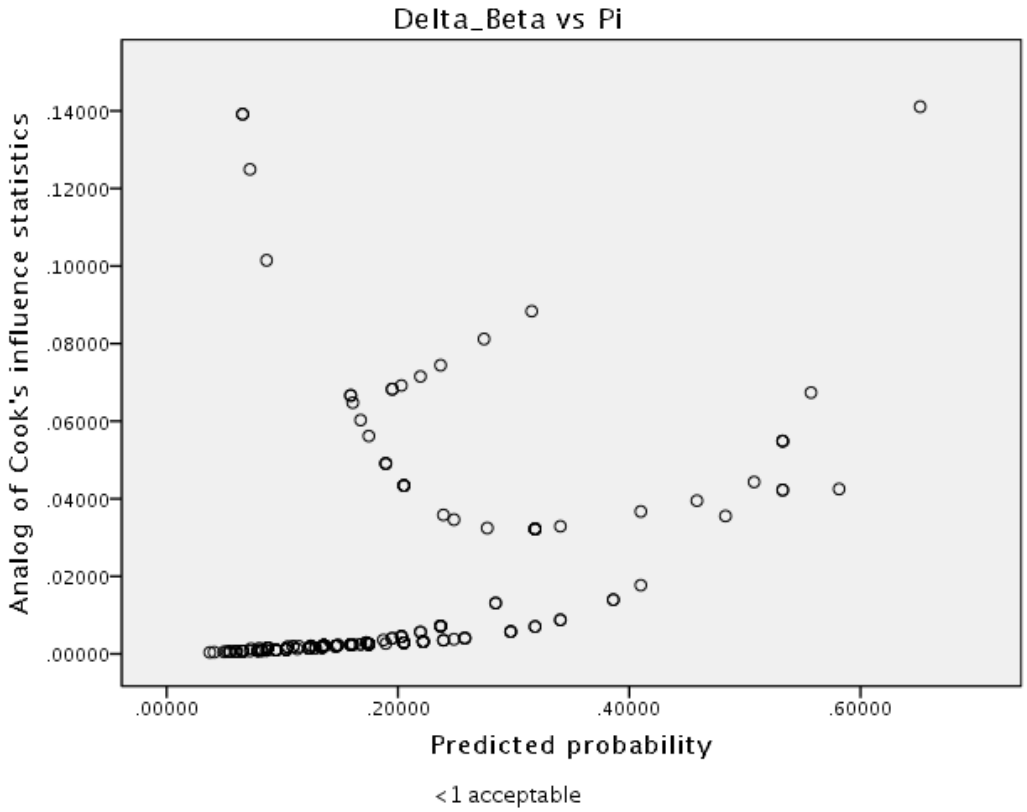
Checking linearity in the logit

- Cont…
  - Plot estimated coefficients vs group medians
  - **Graphs -> Legacy Dialogs -> Scatter/Dot -> Simple Scatter -> Define**
  - **Y Axis:** *coefficient*, **X Axis:** *dbp* **-> OK**
  - Double click on the plot → **Elements → Interpolation Line**
  - Should have an approximately straight line → Linearity in logit assumption fullfilled.



Logistic regression diagnostics

- **In STATA, based on covariate patterns.**
- **In SPSS, limited and not based on covariate patterns:**
  - Change in estimated coefficients (after deleting a case) vs predicted probabilities
  - Click **Save...** → Tick **Cook's** under **Influence**
  - A new variable *COO_1* will be created.
  - Plot *COO_1* vs *PRE_1*
  - Values should be < 1 (Hosmer & Lemeshow, 2000).



Q&A