Example: Diabetes Data.

In the SPSS parameterization:

β_{30}, β_{31}	Group 3 Intercept and Slope
$\beta_{10} = \beta_{30} + \delta_{10}$ $\beta_{20} = \beta_{30} + \delta_{20}$	Changes in the Intercepts in Groups 1 and 2 are δ_{10} and δ_{20}
$\beta_{11} = \beta_{31} + \delta_{11}$ $\beta_{21} = \beta_{31} + \delta_{21}$	Changes in the Slopes in Groups 1 and 2 are δ_{11} and δ_{21}

Thus the six new parameters are

$$\beta_{30}, \beta_{31}, \delta_{10}, \delta_{20}, \delta_{11}, \delta_{21}$$

MODEL 0 $\beta_{31} = 0$
$\delta_{10} = \delta_{20} = \delta_{11} = \delta_{21} = 0$

MODEL 1 $\beta_{31} = \delta_{11} = \delta_{21} = 0$

MODEL 2 $\delta_{10} = \delta_{20} = \delta_{11} = \delta_{21} = 0$

MODEL 3 $\delta_{11} = \delta_{21} = 0$

Note: $\beta_{31} = 0 \Longrightarrow \delta_{11} = \delta_{21} = 0$, as X_1 is not included in the model.

Counting Parameters

- Whenever we remove a **continuous covariate**, from a model, we set **one** parameter to zero.
- Whenever we remove a **factor predictor** with L levels from a model, we set $L-1$ parameters to zero.
- Whenever we remove a two-way interaction between these variables from a model, we set $1.(L-1) = L-1$ parameters to zero.

Models 0,1,2,3 are nested inside Model 4.

Two approaches to finding the best model are used:

1. Start with Model 0 and try to **add** terms that improve the model fit (**Forward Selection**)
2. Start with Model 4 and try to remove terms that improve the model fit (**Backward Selection**)

Note: Models 0,1 and 2 are nested inside Model 3. Model 0 is nested inside Models 1 and 2. Therefore we can begin with Model 4, or Model 3 or Model 1 or 2, and simplify to a nested model.

Example: Diabetes Data.

Here $n = 144$. From SPSS output handouts:

Model	Description	SSE	p
0	1	28.504	1
1	X_2	4.160	3
2	X_1	3.738	2
3	$X_1 + X_2$	1.472	4
4	$X_1 + X_2 + X_1.X_2$	1.318	6

p is the number of non-zero parameters; k or g is always $p - 1$ in the following analysis.

Backward Selection:

Complete Model : Model 4
Reduced Model : Model 3

Here $k = 5, g = 3$ so $k - g = 2$, and
$n - k - 1 = 144 - 5 - 1 = 138$.

$$F = \frac{(SSE_R - SSE_C)/(k-g)}{SSE_C/(n-k-1)} = \frac{(1.472 - 1.318)/2}{1.318/138} = 8.062$$

We compare this with the

$$\text{Fisher-F}(k-g, n-k-1) = \text{Fisher-F}(2, 138)$$

distribution; we have $F_\alpha(2, 138) = 3.061$, so we

$$\text{Reject } H_0 \text{ at } \alpha = 0.05$$

i.e. Model 4 ($X_1 + X_2 + X_1.X_2$) fits **significantly better** than Model 3 ($X_1 + X_2$).

- we cannot simplify the complete model to the reduced model without the loss of significant explanatory power.

The Interaction is Necessary in the Model

Backward selection stops here; we cannot simplify further from the complete model.