

HLM – Erklärungsversuche zu Folie 19 – 29

Zur Erklärung, wie die HLM aufgebaut wird, wird ein Beispiel von Hox (2002, Kap. 2) verwendet. In dem Beispiel gibt es eine unterschiedliche Anzahl von SchülerInnen N_i in verschiedenen Klassen. Untersucht wird der Effekt von zwei Prädiktoren auf das Kriterium Popularität der SchülerInnen (Y). Auf Level 1, der Schülerebene, befindet sich das Kriterium und der Prädiktor Geschlecht (X). Auf der Klassenebene wird den Einfluss der Erfahrung von LehrerInnen (Z) untersucht.

Level2:

Erfahrung (Z)

Level 1:



Warum die Erfahrung als Moderatorvariable wirkt, wird später deutlich.

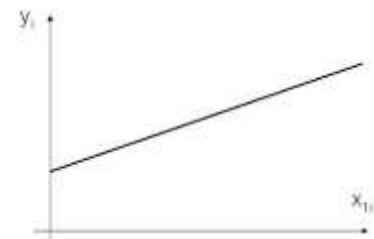
Variablen:

- Popularität der SchülerInnen als AV (Y)
- Geschlecht auf Ebene der SchülerInnen (X: 0=Junge, 1 = Mädchen)
- Erfahrung der LehrerInnen auf der Klassenebene (Z: in Jahren)
- $N=2000$ SchülerInnen, $j=100$ Klassen

Es ergibt sich eine Gleichung für den Zusammenhang zwischen X und Y, die erstmal einfach eine Regression darstellt und die verschiedenen Ebenen nicht berücksichtigt:

$$Y_i = \beta_0 + \beta_1 * X_i + e_i$$

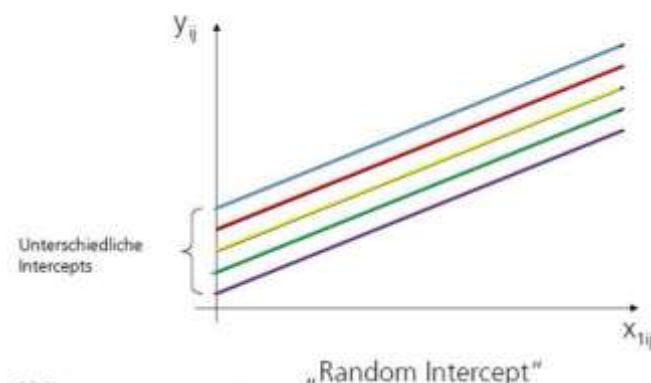
Graphisch lässt sich das dann wie in der Vorlesung darstellen (Folie 19). (Nicht wundern $X_1 = X$, das liegt nur an der Benennung, die ich hier konsistent verwende, in der Vorlesung wird da gewechselt.)



Das ist nun der Level-1 Zusammenhang zwischen Geschlecht und Popularität. Das würde nun aussagen, dass dieser Zusammenhang für alle Schüler gilt – egal in welcher Klasse sie sind. Mittels HLM lässt sich nun betrachten, ob das wirklich für alle Klassen gilt. Dazu wird zugelassen, dass das Kriterium, die Regressionsgewichte, die Prädiktorwerte und das Residuum in den verschiedenen Klassen unterschiedlich sein darf. Dafür wird der Indize j eingeführt, der die Klasse, aus der die SchülerInnen kommen, angibt:

$$y_{ij} = \beta_{0j} + \beta_{1j} x_{1ij} + e_{ij}$$

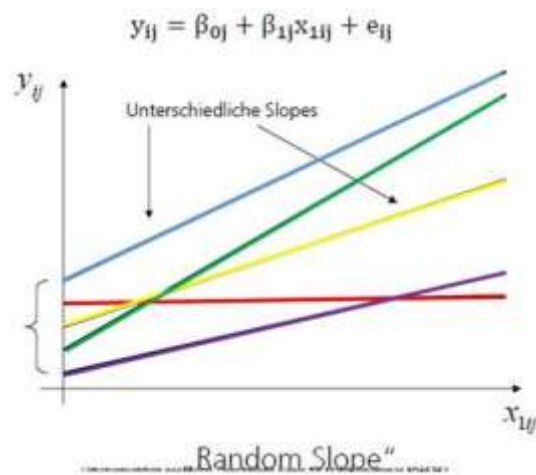
$$Y_{ij} = \beta_{0j} + \beta_{1j} * X_{ij} + e_{ij}$$



Auf Folie 21 ist dargestellt, was passiert, wenn wir zunächst einmal nur dem Intercept (β_{0j}) erlauben zwischen den Klassen unterschiedlich zu sein, der Slope bleibt gleich (β_{1j}). Ich habe die Regressionsgeraden der unterschiedlichen Klassen einfach mal eingefärbt, um das besser zu veranschaulichen. Es verändert sich hier nur der Intercept – er variiert zwischen den Klassen und ist demnach zufällig (random).

Der Intercept bedeutet, dass bei einer Ausprägung von Null im Prädiktor (X; Geschlecht), das

Kriterium als Mittelwert den Wert des Intercepts annimmt. $X = 0$ ist in diesem Beispiel definiert als „Junge“. Jungen haben also den Wert des Intercepts als vorhergesagten, mittleren Y-Wert innerhalb einer Klasse. Das lässt sich gut an der Abbildung erkennen: Jungen in der Klasse „lila“ haben bspw. die geringsten Popularitätswerte und Jungen in der Klasse „blau“ die höchsten. Zwar unterscheiden sich jetzt die Intercepts, aber es ist immer noch so, dass in allen Klassen die Popularitätswerte der Mädchen höher sind als die der Jungen und auch im selben Maß ansteigen (Steigung ist überall gleich = fixed Slope).



Als nächstes lassen wir zu, dass sich auch der Slope zwischen den Klassen verändern darf (random slope). Das entspricht dann der Abbildung auf Folie 22. Der Intercept in den unterschiedlichen Klassen ist unterschiedlich und ebenso der Slope. Es ist jetzt also möglich, dass der Zusammenhang zwischen Geschlecht und Popularität sich unterscheidet. In Klasse „rot“ sind Jungen beliebter als Mädchen, in den anderen Klassen ist es umgekehrt. Beim Vergleich von Klasse gelb und grün fällt auf, dass die Unterschiede der Popularität zwischen Jungen und Mädchen in der gelben Klasse viel geringer sind als in der grünen Klasse.

Das lässt sich auch mathematisch darstellen (Folie 23). Die Regressionsgewichte auf Level 1, setzen sich auf Level 2 aus einem Regressionsgewicht γ und einer Abweichung μ zusammen. So dass gilt:

$$\text{Level 1: } Y_{ij} = \beta_{0j} + \beta_{1j} * X_{ij} + e_{ij}$$

$$\text{Level 2: } \beta_{0j} = \gamma_{00} + \mu_{0j}$$

$$\beta_{1j} = \gamma_{10} + \mu_{1j}$$

(Folie 25 und 26 und 27 sind ganz am Ende)

Bis hierhin können wir nur feststellen, dass es Unterschiede gibt, aber nicht warum. Dafür müssen weitere Variablen diskutiert werden. Es könnten zum Beispiel Variablen auf Level 1 hinzugenommen werden (Im Buch von Hox zum Beispiel Extraversion). Es gibt aber auch Variablen auf Level 2. Der Vorteil von HLM ist, dass eben auch Level 2 (oder noch höhere) Variablen berücksichtigt werden können. In diesem Beispiel wird die Erfahrung der LehrerInnen (Z) als eine Level 2 Variable verwendet, um die Variabilität der Regressionskoeffizienten zwischen den Klassen zu erklären. Die Regressionsgleichung auf Level 1 wird dafür durch weitere Gleichungen auf Level 2 ergänzt, die einen linearen Zusammenhang zwischen den Regressionsgewichten auf Level 1 und der Level 2 Variable ausdrücken:

$$\text{Level 1: } Y_{ij} = \beta_{0j} + \beta_{1j} * X_{ij} + e_{ij}$$

$$\text{Level 2: } \beta_{0j} = \gamma_{00} + \gamma_{01} * Z_j + \mu_{0j} \text{ und}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} * Z_j + \mu_{1j}$$

Die erste Gleichung auf Level 2 sagt den Mittelwert des Kriteriums (Popularität; Y) in einer Klasse vorher. γ_{00} ist der Wert, den β_{0j} annimmt, wenn die Variable Z die Ausprägung Null (keine Berufserfahrung) hat. γ_{01} gibt an, um wie viele Einheiten sich β_{0j} verändert, wenn Z_j sich um eine Einheit ändert (die Interpretation ist also wie bei der Regression. Wenn γ_{01} positiv ist, steigt der

Mittelwert der Popularität mit zunehmender Erfahrung an. Wenn γ_{01} negativ ist, dann sinkt die Popularität mit zunehmender Erfahrung.

Die zweite Gleichung auf Level 2 bezieht sich auf den Zusammenhang (dargestellt durch β_{1j}) zwischen Popularität (Y) und Geschlecht (X) in Abhängigkeit von der Erfahrung des Lehrers. Wenn γ_{11} positiv ist, dann ist der Effekt von Geschlecht auf Popularität größer bei größerer Erfahrung → die Unterschiede zwischen den Geschlechtern werden größer. Wenn γ_{11} negativ ist, dann ist der Effekt von Geschlecht auf Popularität geringer bei größerer Erfahrung → die Unterschiede zwischen den Geschlechtern werden kleiner. Das entspricht inhaltlich einer Moderation.

Durch Einsetzen der Level 2 Gleichungen in die Level 1 Gleichung ergibt sich:

$$Y_{ij} = \gamma_{00} + \gamma_{01} * Z_j + \mu_{0j} + \gamma_{10} * X_{ij} + \gamma_{11} * Z_j * X_{ij} + \mu_{1j} * X_{ij} + e_{ij}$$

Gelb hervorgehoben in dieser Gleichung ist der Interaktionsterm zwischen X und Z (Cross-Level-Interaktion). Hier zeigt sich jetzt mathematisch, dass die Variable auf Level 2 als Moderatorvariable in das Modell aufgenommen wird.

Residuen und Varianzen (Folie 25 und 26)

Das Level 1-Residuum e_{ij} lässt sich als Differenz der beobachteten Werte von den vorhergesagten Werten darstellen: $e_{ij} = y_{ij} - \hat{y}_{ij}$. Ganz äquivalent lassen sich auch die Abweichungen μ als Differenz der β (auf Level 1) und der vorhergesagten β -Werte verstehen (Folie 25):

- Abweichung des klassenspezifischen Intercepts β_{0j} vom vorhergesagten Intercept (= random intercept; u_{0j})
 $u_{0j} = \beta_{0j} - \hat{\beta}_{0j}$
- Abweichung des klassenspezifischen Slopes β_{1j} vom vorhergesagten Slope (= random slope; u_{1j})
 $u_{1j} = \beta_{1j} - \hat{\beta}_{1j}$
- Abweichung des y_{ij} -Wertes einer/s SchülerIn i in Klasse j vom vorhergesagten Wert $\hat{y}_{ij}$ (= Level-1-Residuum; e_{ij})
 $e_{ij} = y_{ij} - \hat{y}_{ij}$

Auch die Varianzen der Abweichungen μ und des Residuums auf Level 1 lassen sich interpretieren (Folie 26)

- Abweichung des klassenspezifischen Intercepts β_{0j} vom vorhergesagten Intercept (= random intercept; u_{0j})
 Die Varianz der random intercepts $\sigma_{u_0}^2$ zeigt an, in welchem Ausmaß die Popularität unsystematisch von Klassen zu Klasse variiert
- Abweichung des klassenspezifischen Slopes β_{1j} vom vorhergesagten Slope (= random slope; u_{1j})
 Die Varianz der random slopes $\sigma_{u_1}^2$ zeigt an, in welchem Ausmaß der Effekt des Geschlechts auf die Popularität unsystematisch zwischen Klassen variiert
- Abweichung des y_{ij} -Wertes einer/s SchülerIn i in Schule j vom vorhergesagten Wert $\hat{y}_{ij}$ (= Level-1-Residuum; e_{ij})
 Die Varianz der Level-1-Residuen σ_e^2 zeigt an, in welchem Ausmaß die Popularität innerhalb einer Klasse unsystematisch zwischen den SchülerInnen variiert

Anzahl der zu schätzenden Parameter.

In der HLM müssen die Regressionskoeffizienten und Varianzkomponenten geschätzt werden. Für die HLM-Gleichung $Y_{ij} = \gamma_{00} + \gamma_{10} * X_{ij} + \mu_{0j} + \mu_{1j} * X_{ij} + e_{ij}$ also 2 Regressionskoeffizienten, 2 Varianzen der Abweichungen μ , 1 Varianz des Level 1-Residuums und die Kovarianz zwischen den Abweichungen. Die Schätzung erfolgt iterativ, z.B. über Maximum Likelihood.

- Es müssen also insgesamt 6 Modellparameter geschätzt werden:
 - Das durchschnittliche Intercept (γ_{00})
 - Der durchschnittliche Slope (γ_{10})
 - Die Varianz der *random intercepts* ($\sigma_{u_0}^2$)
 - Die Varianz der *random slopes* ($\sigma_{u_1}^2$)
 - Die Kovarianz zwischen u_0 und u_1 ($\sigma_{u_{01}}$)
 - Die Varianz der Level-1-Residuen (σ_e^2)

Für $Y_{ij} = \gamma_{00} + \gamma_{01} * Z_j + \mu_{0j} + \gamma_{10} * X_{ij} + \gamma_{11} * Z_j * X_{ij} + \mu_{1j} * X_{ij} + e_{ij}$ müsste entsprechend gelten: 4 Regressionskoeffizienten (grün), 2 Varianzen der Abweichungen μ , die Varianz des Residuums e und die Kovarianz zwischen den Abweichungen. Insgesamt also 8 Modellparameter. (So habe ich das jetzt zumindest verstanden).