



Latent Growth Model in Psicologia: Cosa si rischia se la misura non è (davvero) su scala a intervalli?

ENRICO TOFFALINI
TOMMASO FERACO
MASSIMILIANO PASTORE

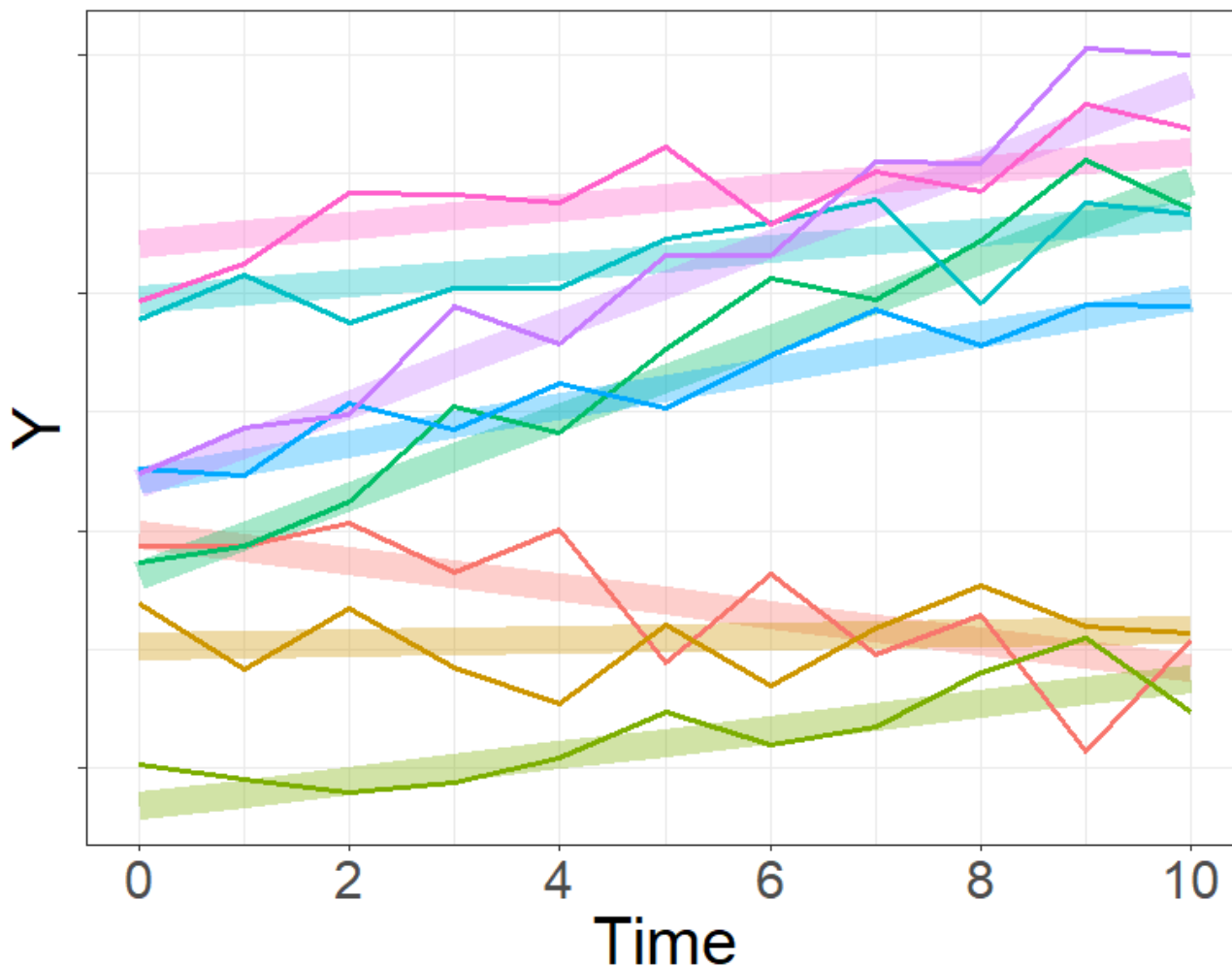
DPG, Università di Padova

DPG, Università di Padova

DPSS, Università di Padova

Utilità dei latent growth model

Ogni percorso ha un punto di partenza e una direzione. I LGM ci aiutano a descriverne i parametri



Utilità dei latent growth model

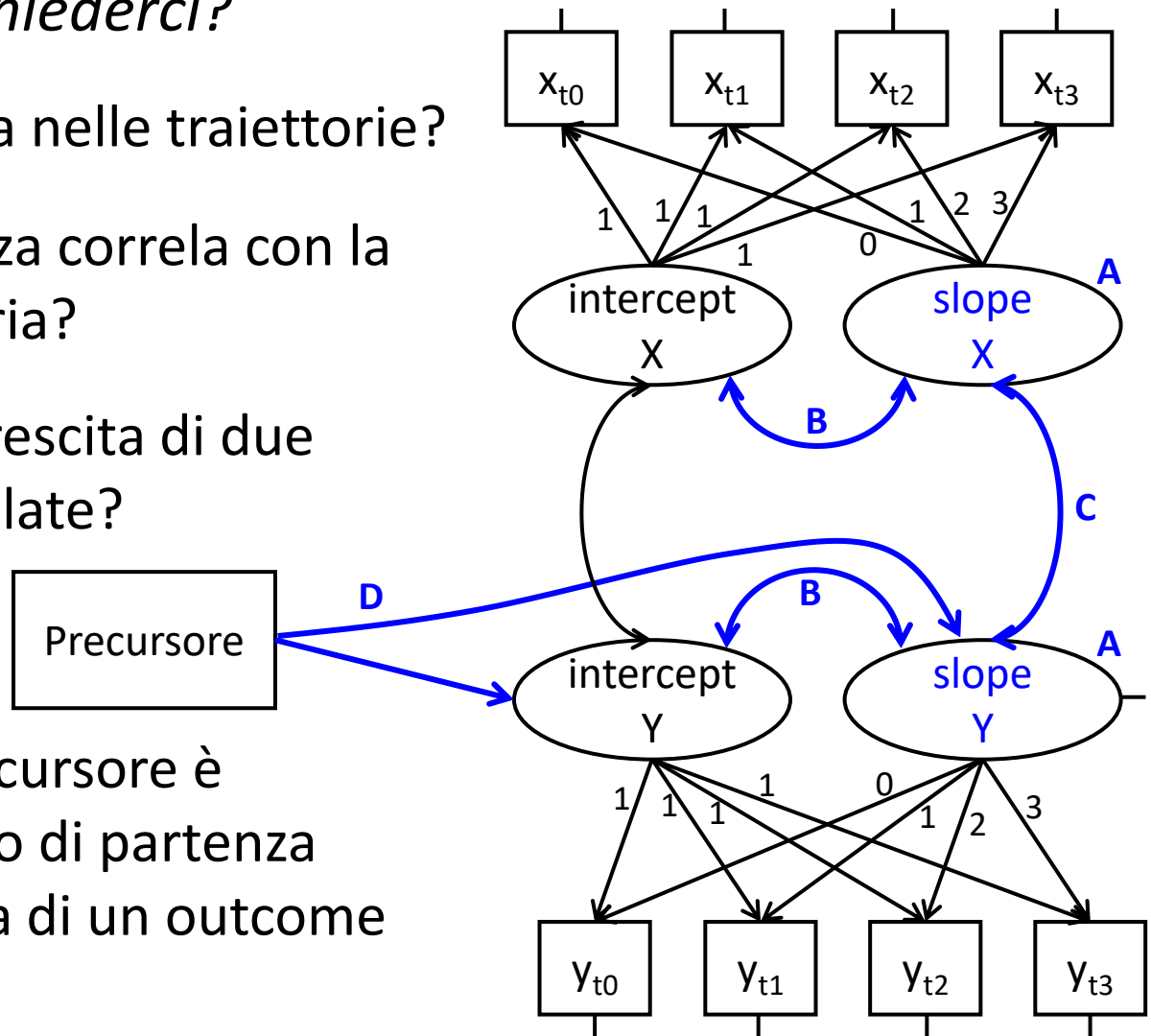
Cosa potremmo chiederci?

A C'è variabilità vera nelle traiettorie?

B Il punto di partenza correla con la successiva traiettoria?

C Le traiettorie di crescita di due variabili sono correlate?

D Un eventuale precursore è predittivo del punto di partenza e/o della traiettoria di un outcome di interesse?



Utilità dei latent growth model

Questi fenomeni sono d'interesse scientifico... e talvolta gli si dà curiosi nomi evangelici...

(...) Growth and stability can also be distinguished from spread (see Figure 1), which relates to whether the distribution of scores over time stays the same, decreases, indicating a **compensatory pattern**, or increases, such that the **“rich get richer and the poor get poorer”** (so called **Matthew effects**; Pfost, Hattie, Dörfler, & Artelt, 2014; Stanovich, 1986). **Matthew effects** have been reported for reading comprehension between 7 and 10 years (Quinn et al., 2015), and for vocabulary knowledge between childhood and 16 years (Duff et al., 2015). However, a compensatory pattern has also been reported for a reading composite between 6 and 12 years (Shaywitz et al., 1995). A recent metaanalysis shows inconsistency across studies but suggests that Matthew effects were more likely for sensitive and reliable reading measures (Pfost et al., 2014)”

Ricketts et al. (2020)

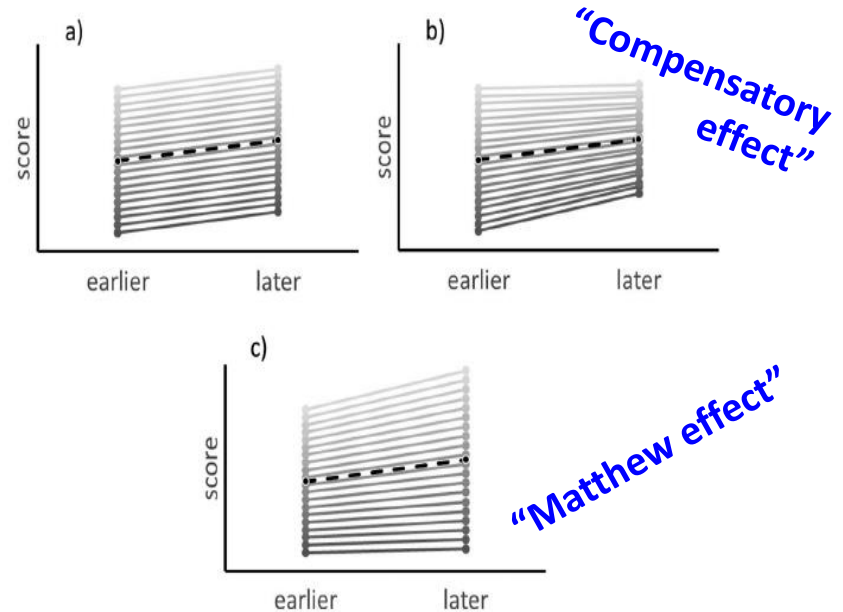


Figure 1. Hypothetical illustration of how spread can be consistent (a), decrease, indicating a compensatory pattern (b), or increase, indicating a Matthew effect (c).


Note. Growth, stability and spread are statistically independent from each other. In this example, there is consistent mean growth (black dashed line) and perfect stability (gray solid lines), with earlier rank in distribution identical to later rank.

SCIENTIFIC STUDIES OF READING
2020, VOL. 24, NO. 5, 380–396
<https://doi.org/10.1080/10888438.2019.1689244>

 **Routledge**
Taylor & Francis Group

 OPEN ACCESS  Check for updates

Reading and Oral Vocabulary Development in Early Adolescence

Jessie Ricketts ^a, Arne Lervåg^b, Nicola Dawson^a, Lucy A. Taylor^a, and Charles Hulme^c

^aDepartment of Psychology, Royal Holloway, University of London, London, UK; ^bDepartment of Education, University of Oslo, Oslo, Norway; ^cDepartment of Education, University of Oxford, Oxford, UK

Fenomeni e spiegazioni

Spiegazioni teoriche? Es. possibile *Matthew effect* in abilità di lettura (Pfos et al. 2014) come circolo virtuoso: leggo bene → mi diverto → leggo di più → leggo sempre meglio. *Oppure*: SES che incide sia su punto di partenza che su traiettoria. *Compensatory effect*, invece, talvolta interpretato come temporaneo ritardo nello sviluppo che poi si recupera (Pfos et al., 2014)

Ma prima di lanciarsi in sofisticate spiegazioni psicologiche
bisogna escludere semplici spiegazioni psicometriche

Banali artefatti da cui guardarsi? Se ho solo 2 time point potrei correlare X_0 con $(X_1 - X_0)$? NO. la correlazione sarà negativa («*compensatory effect*»?!) per banale regressione verso la media. Per modellare le variabili latenti dei LGM servono almeno 3 time point

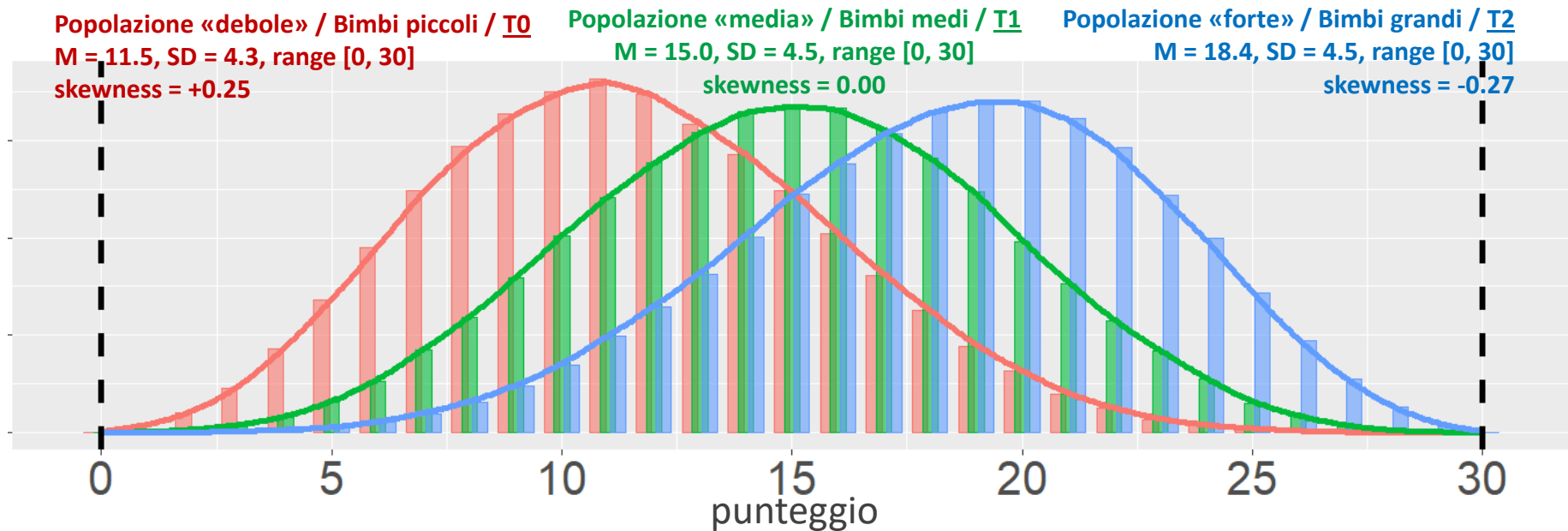
Un problema meno scontato, anche scegliendo il modello giusto, riguarda le distribuzioni dei dati...

Distribuzioni dei dati

In psicologia abbiamo spesso punteggi *bounded* (min-max) non-normali, ma apparentemente neanche troppo skewed, es.

- Somma risposte item scala Likert a questionario
- Somma risposte corrette a test (es. comprensione lettura, matematica, ragionamento)

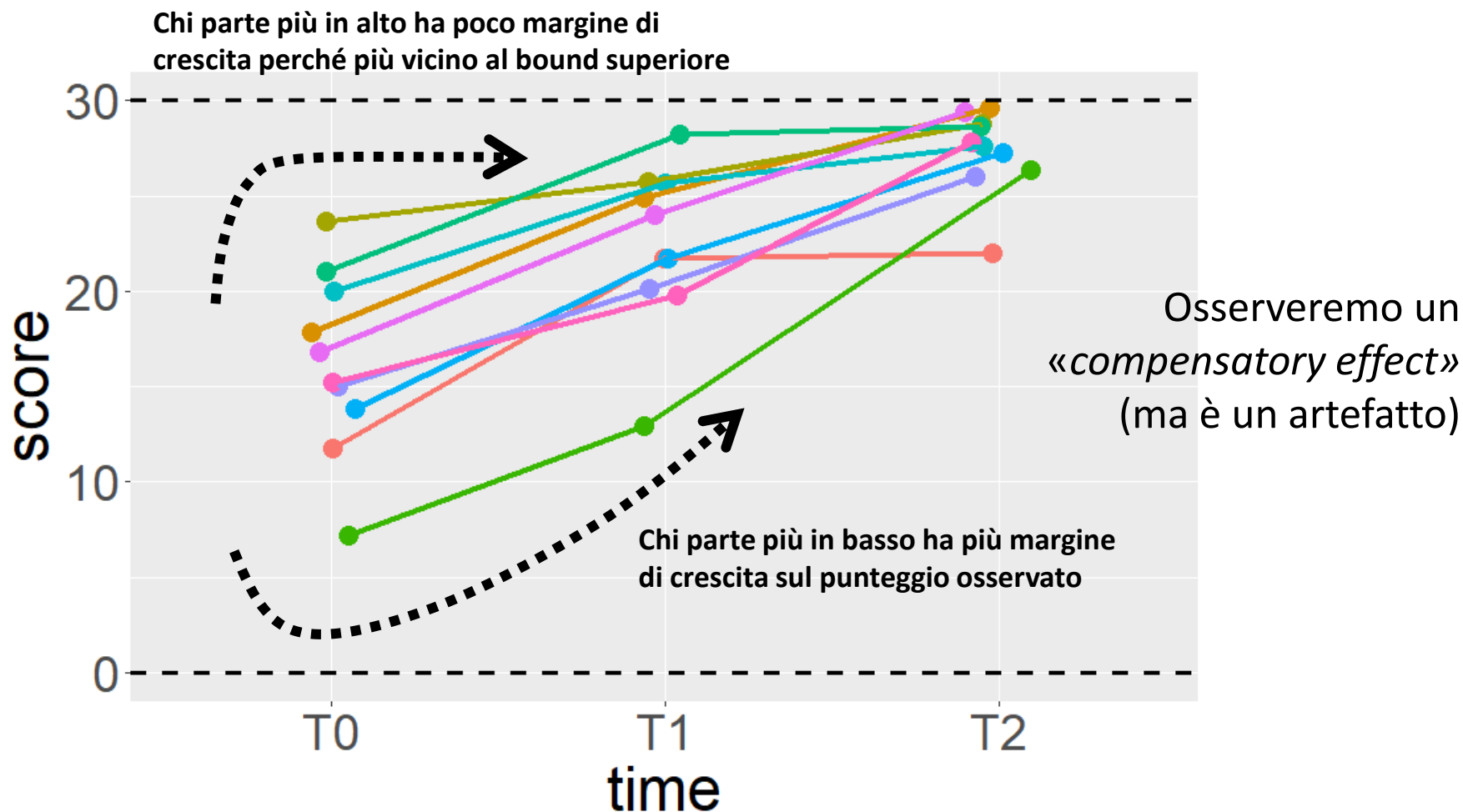
ESEMPIO – Dati da test di comprensione con 30 item di risposta...



Procediamo con un caso analogo, simulando dati...

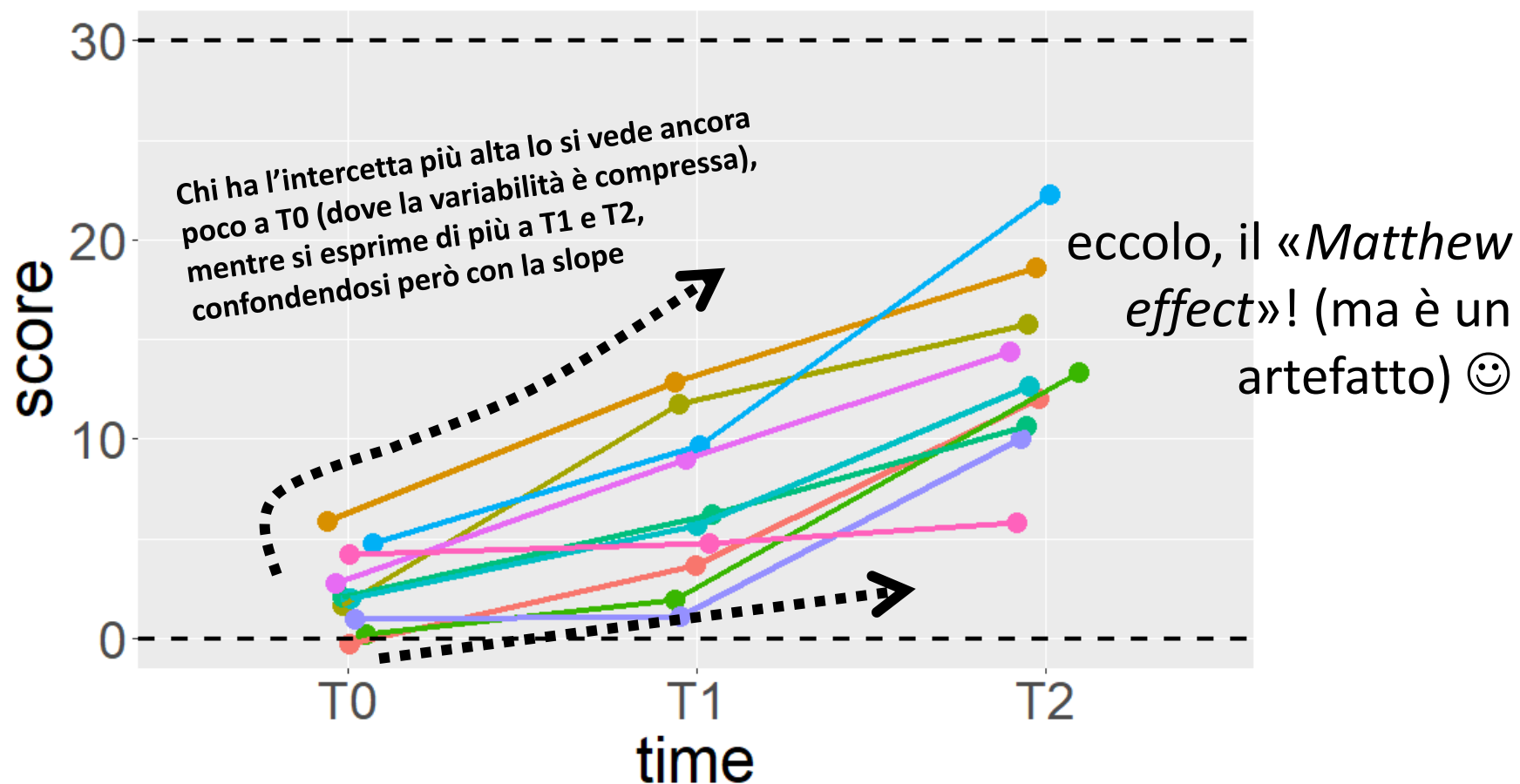
Distribuzioni dei dati

Qui iniziano i guai: prendiamo il caso di un campione preso a un punteggio *intermedio* al T0 e seguito a T1 e T2



Distribuzioni dei dati

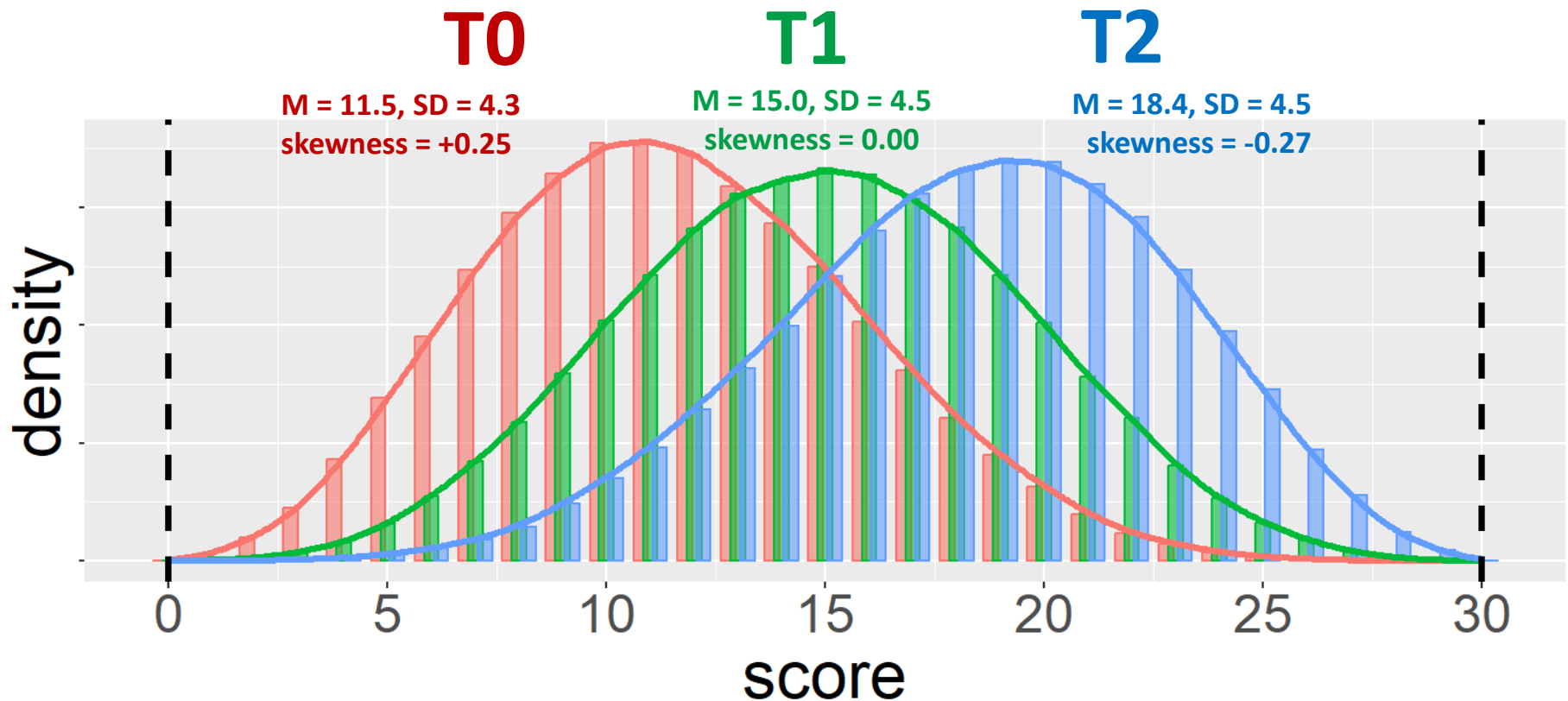
Ma succedere anche l'opposto se prendiamo un campione con punteggio *basso* al T0, poi seguito a T1 e T2



Distribuzioni dei dati

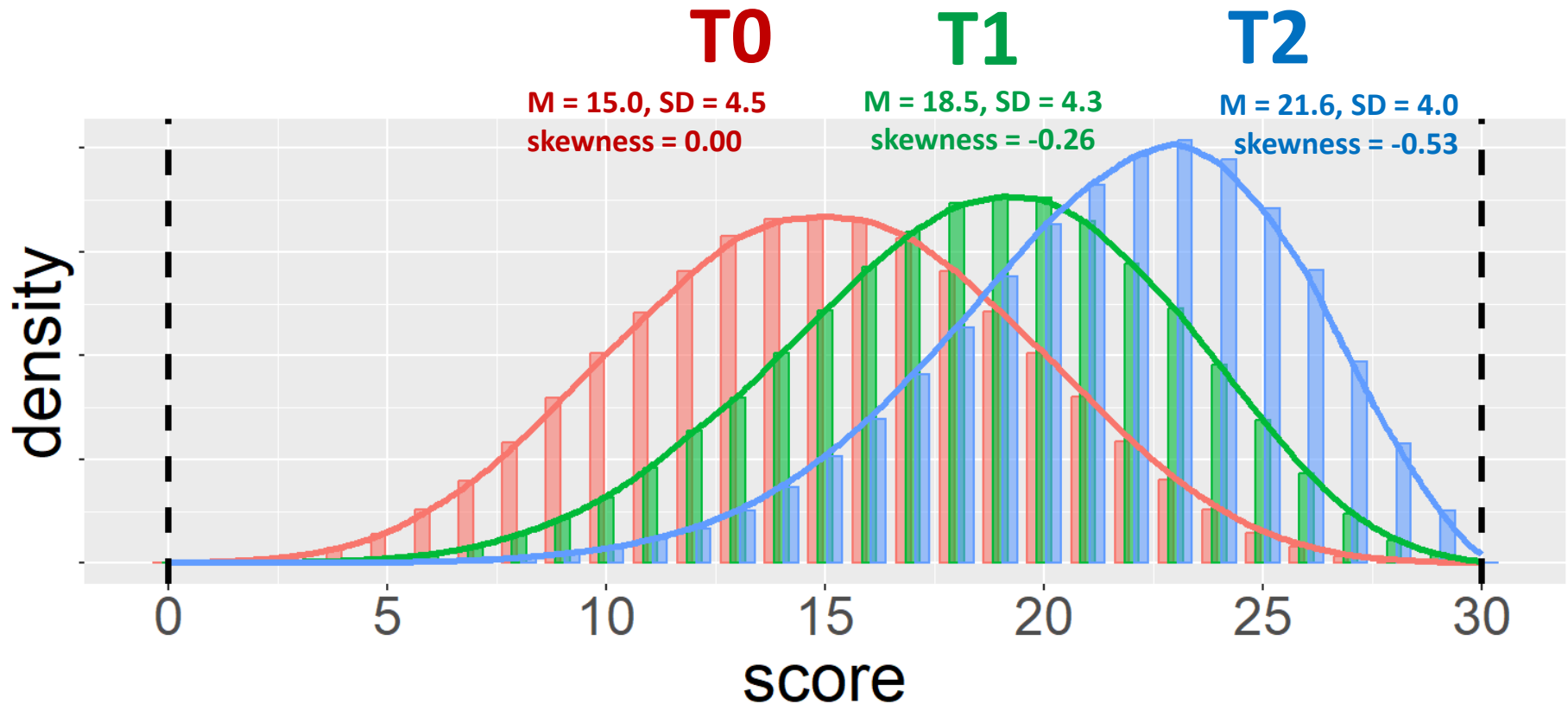
I precedenti esempi sono estremi (*ceiling* and *floor effects* evidenti), ma cosa succederebbe con un caso come quello visto prima?

Già così potremmo avere una **correlazione intercetta-slope di -0.19** (e dunque un «compensatory effect» che NON è nel processo generativo dei dati):



Distribuzioni dei dati

Nel caso qui sotto, un po' più «sospetto» ma neanche troppo per le usuali abitudini, la **correlazione intercetta-slope diventa addirittura di -0.46**



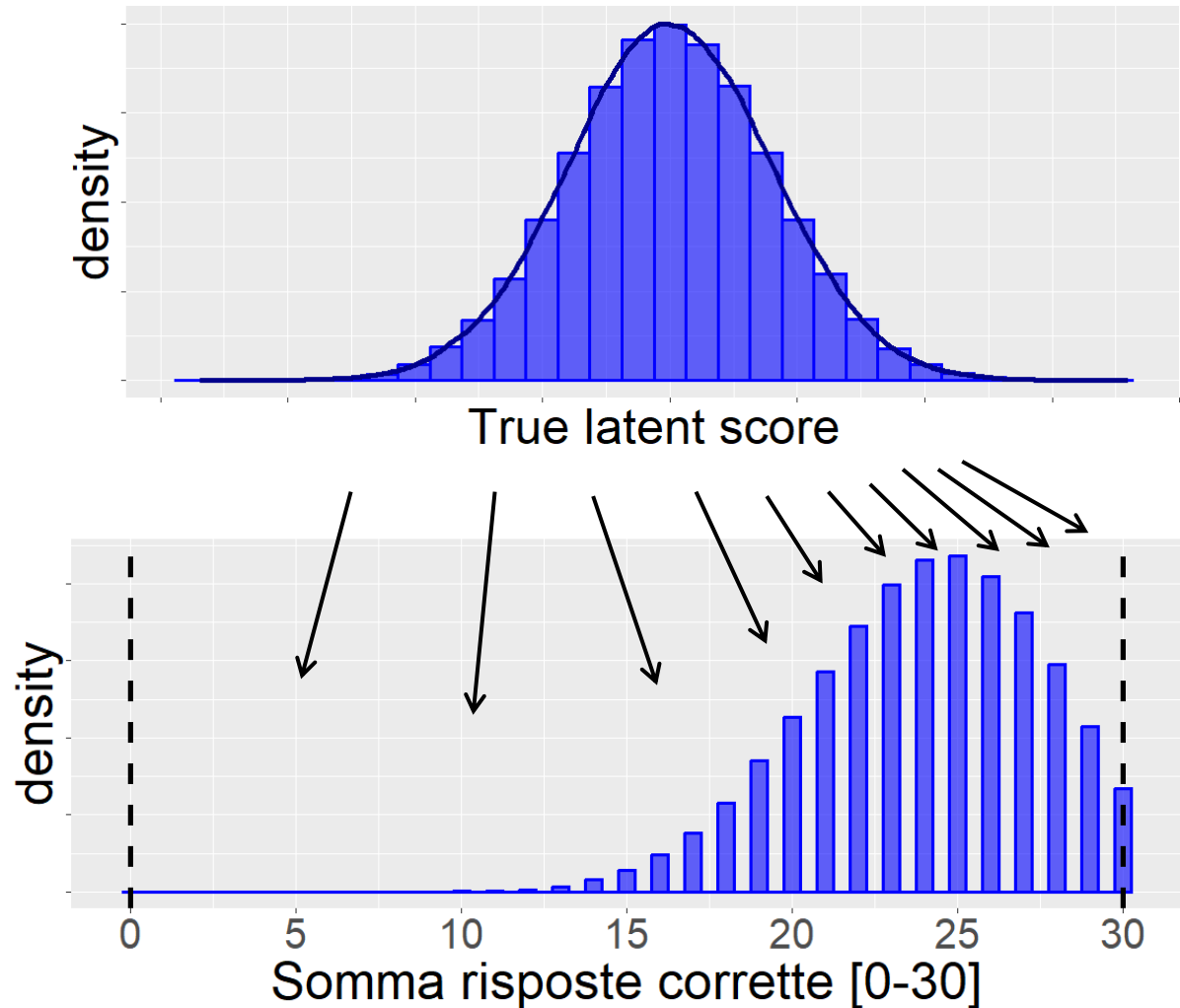
Distribuzioni dei dati

PROCESSO GENERATIVO

‘VERO’: le variabili latenti sono ipotizzate esistere su un continuum lineare e distribuite normalmente

PROCESSO GENERATIVO DELLA MISURA: per motivi pratici siamo costretti a misurarle attraverso risposte su scale binomiali (giusto-sbagliato) o ordinali (scale Likert) che, se non trattate come tali, non possono coprire e rispecchiare l'intero continuum latente

INTERVALLI UGUALI SUL PUNTEGGIO OSSERVATO **NON** CORRISPONDONO A INTERVALLI UGUALI SUL COSTRUTTO LATENTE



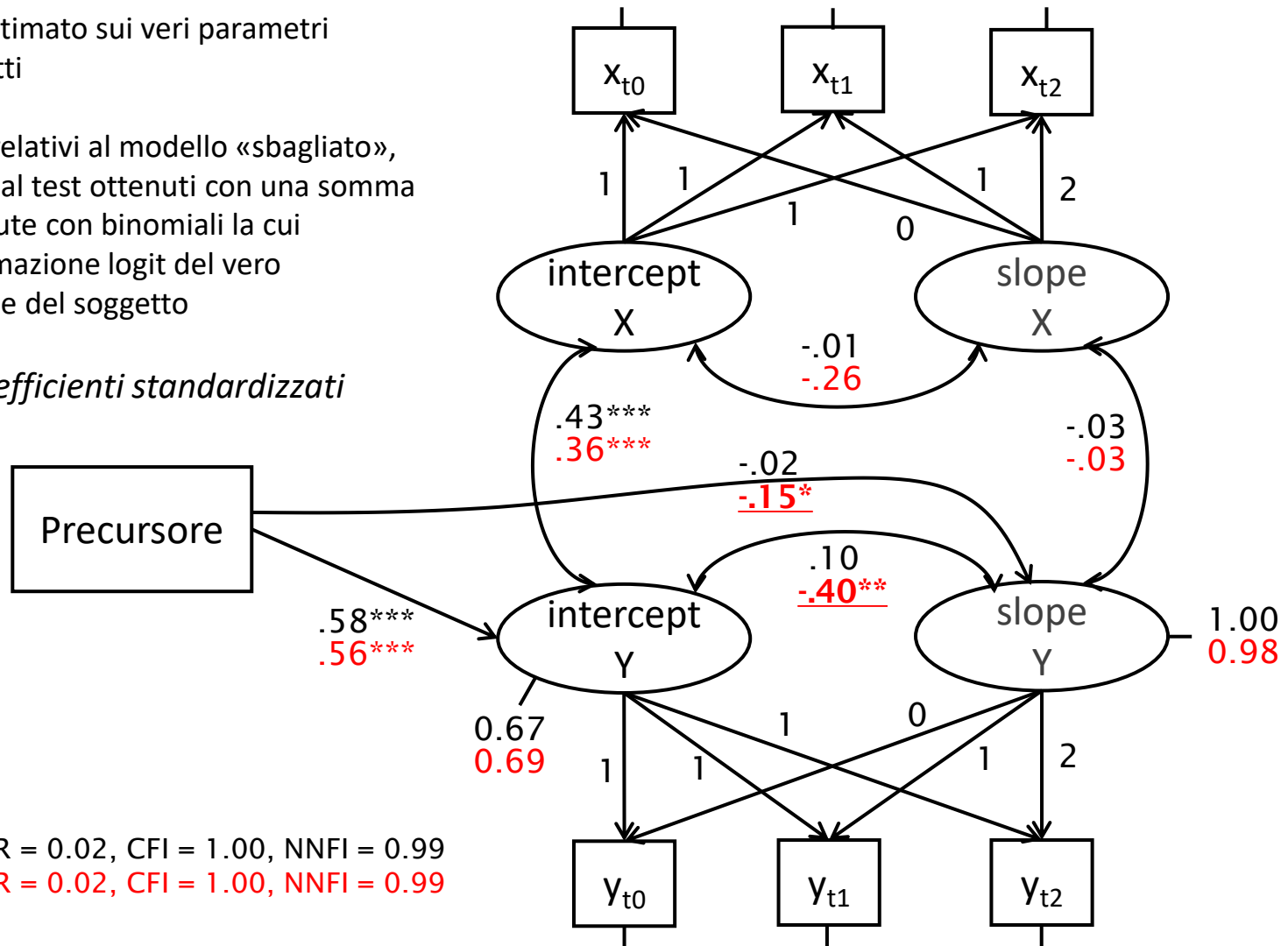
Esempio di stime «sbagliate» sotto condizioni plausibili

in **nero** i parametri relativi al modello «corretto»
generativo dei dati, stimato sui veri parametri
individuali dei soggetti

in **rosso** i parametri relativi al modello «sbagliato»,
stimato sui punteggi al test ottenuti con una somma
di risposte 0/1 ottenute con binomiali la cui
probabilità è trasformazione logit del vero
parametro individuale del soggetto

Sono riportati i *coefficienti standardizzati*

N = 1,000



Indici di fit

RMSEA = 0.03, SRMR = 0.02, CFI = 1.00, NNFI = 0.99

RMSEA = 0.02, SRMR = 0.02, CFI = 1.00, NNFI = 0.99

Conclusioni e suggerimenti

- L'uso di misure bounded [min-max] date da somme di item:
 - genera correlazioni intercetta-slope nei LGMs che sono artefatti;
 - laddove un predittore esterno sia associato all'intercetta, finisce per essere predittivo anche della slope
- Il problema emerge anche con livelli di *skewness* che normalmente verrebbero considerati più che accettabili ([0.20, 0.40]; ma un campanello d'allarme può essere la variazione di skewness tra i tempi)
- Che fare? Bisogna garantire che la misura osservata sia su scala a intervalli, ad esempio ottenendola come stima dei parametri individuali da un modello IRT, oppure lavorare sui SEM a partire dalle risposte osservate ai questionari ...

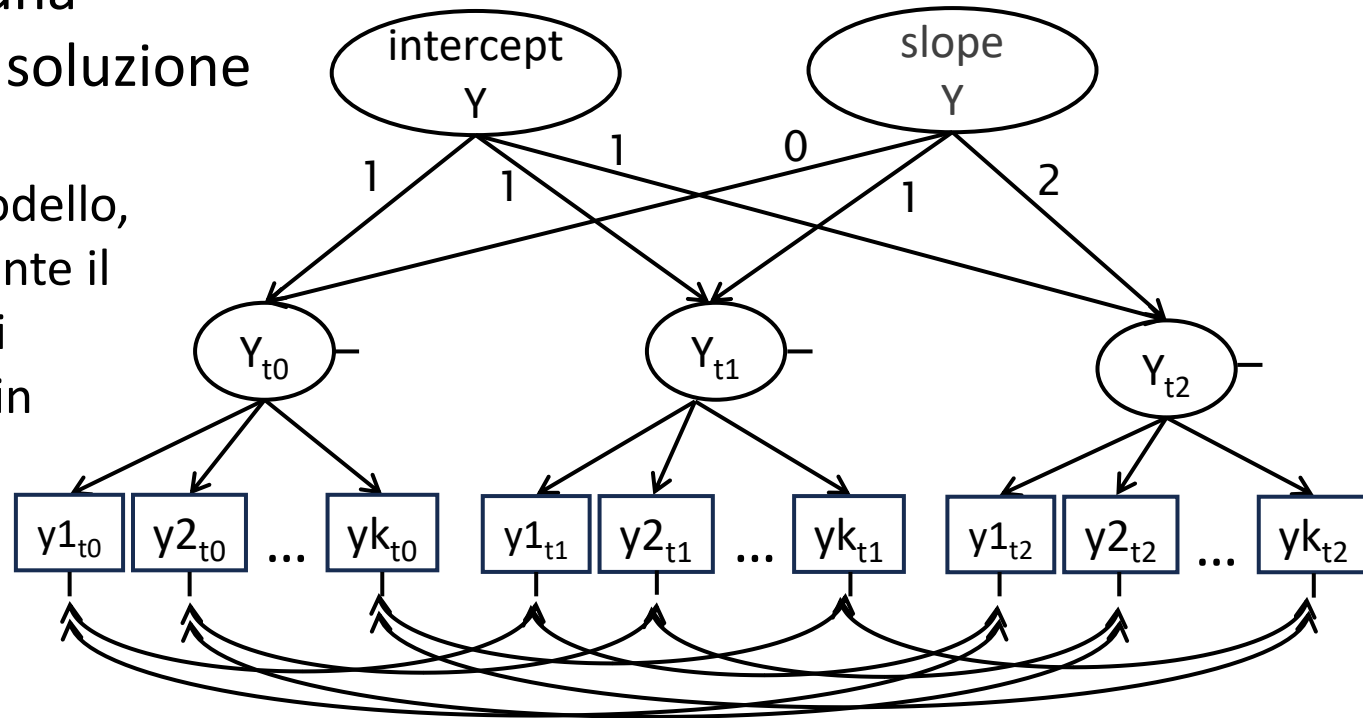
Conclusioni e suggerimenti

...

Questa sarebbe una
ragionevolissima soluzione

MA

complica molto il modello,
aumenta enormemente il
numero di parametri
stimati, imponendo in
genere sample size
maggiori



NB. poiché i punteggi $y1, y2, \dots, yk$ sono su scala Likert/ordinale o perfino dicotomici, ricordarsi di specificare `ordered=TRUE` in lavaan



Associazione
Italiana
di Psicologia

SEZIONE
SPERIMENTALE

XXIX CONGRESSO
NAZIONALE

GRAZIE PER L'ATTENZIONE

ENRICO TOFFALINI
TOMMASO FERACO
MASSIMILIANO PASTORE

DPG, Università di Padova

DPG, Università di Padova

DPSS, Università di Padova

Sessione: Metodologia 2

**LUCCA - Mercoledì,
20 Settembre 2023**