



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

La valutazione formale delle conoscenze

Pasquale Anselmi

Dipartimento di Filosofia, Sociologia, Pedagogia e Psicologia Applicata
Università degli Studi di Padova

Stati Generali della Ricerca, 18 Settembre 2018

Knowledge space theory (KST)*

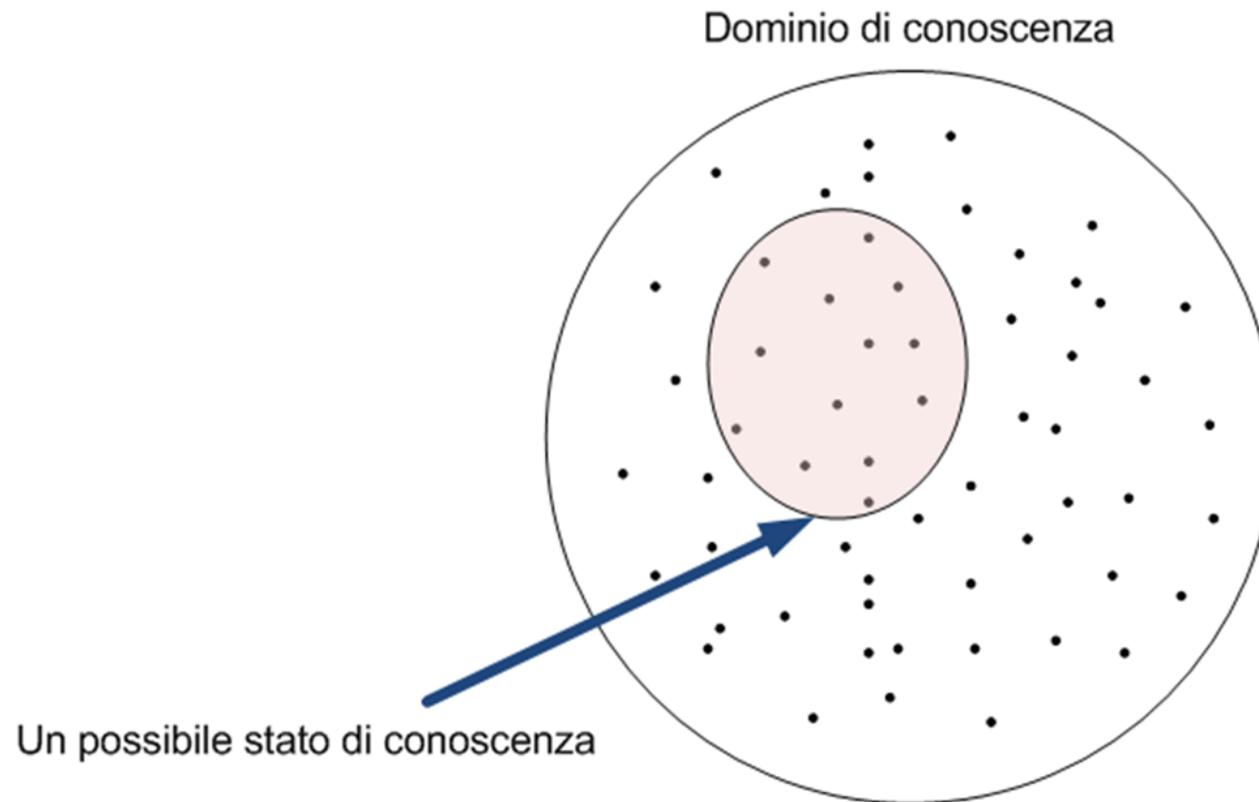
- È un approccio formale alla valutazione delle conoscenze sviluppato da Jean-Claude Falmagne e Jean-Paul Doignon negli anni '80
- L'approccio tradizionale (CTT, IRT) fornisce un punteggio numerico rappresentativo del livello di conoscenza raggiunto dallo studente in un certo dominio disciplinare
- Di converso, la KST fornisce una rappresentazione non-numerica, ma comunque precisa, di ciò che lo studente sa e non sa in quel dominio disciplinare

*Doignon & Falmagne (1985, 1999), Falmagne & Doignon (2011)

Knowledge space theory

Stato di conoscenza

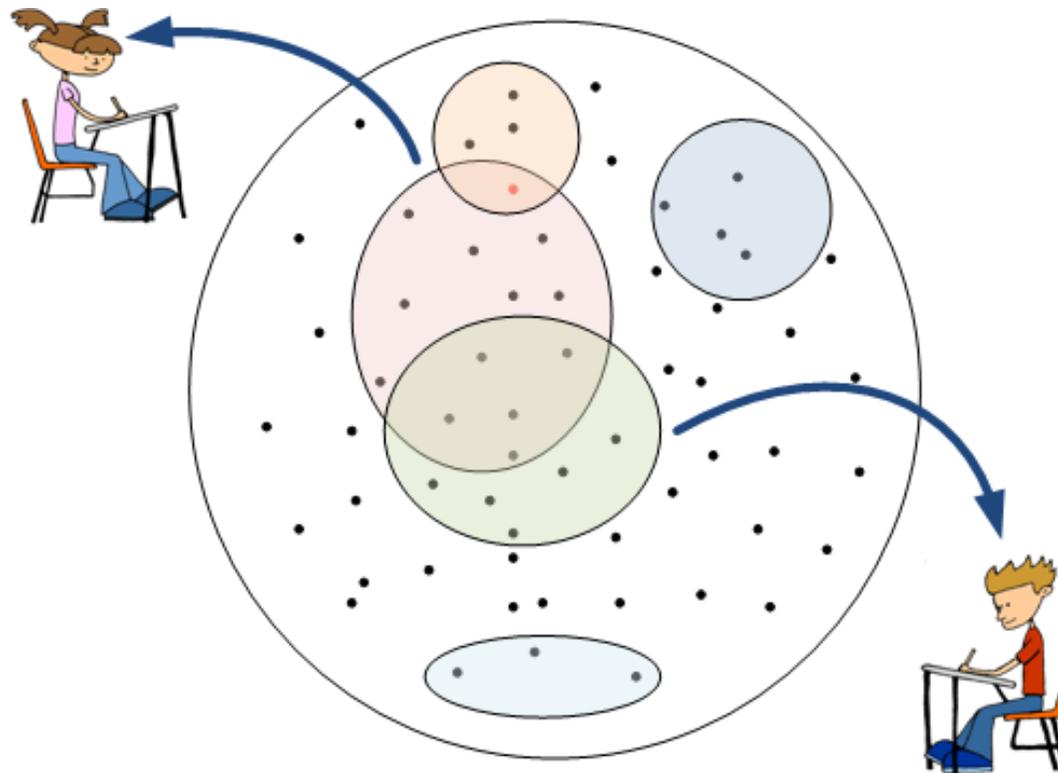
- È l'insieme di tutti gli item che uno studente è in grado di risolvere in un certo dominio disciplinare



Knowledge space theory

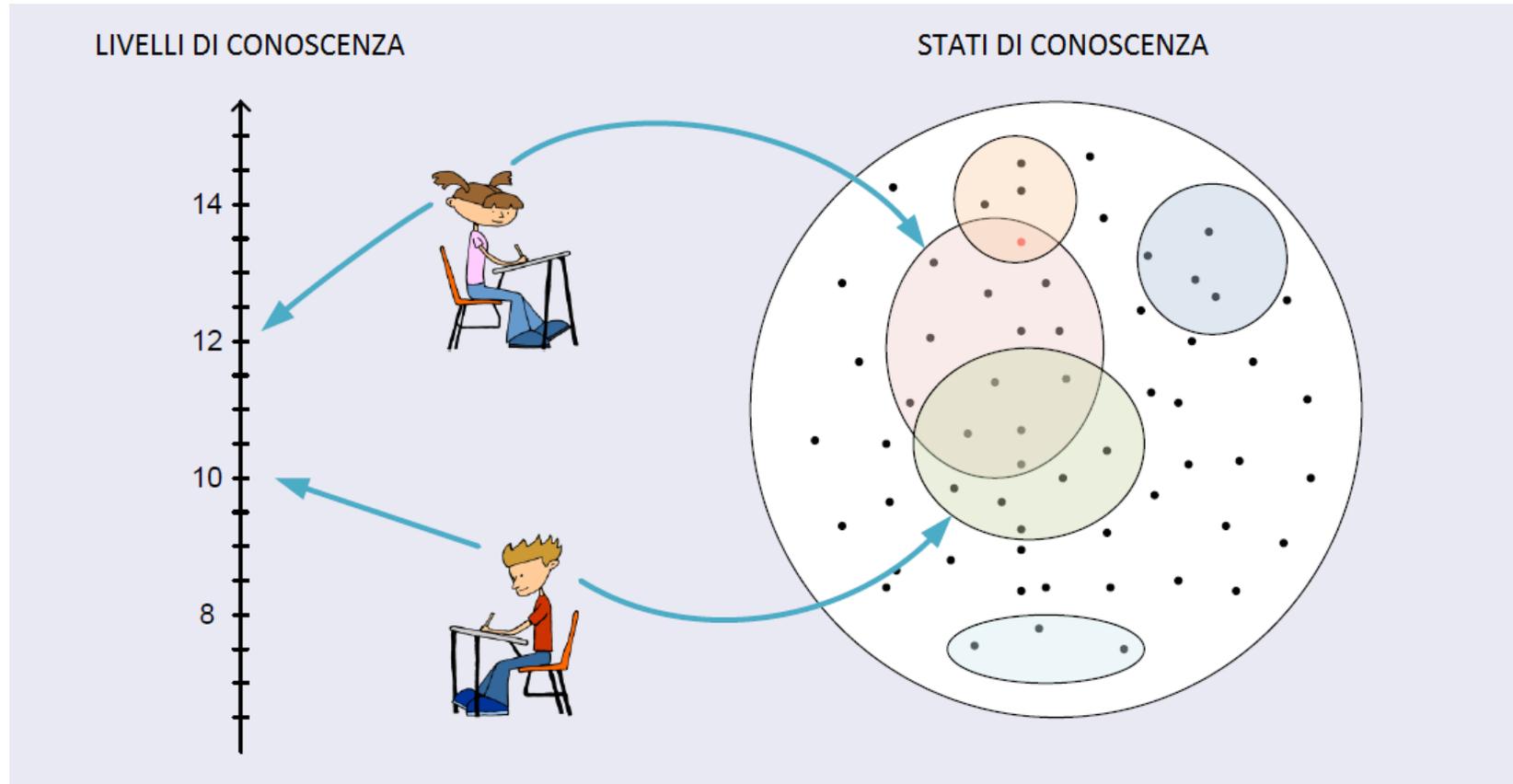
Struttura di conoscenza

- È l'insieme di tutti gli stati di conoscenza osservabili in una popolazione di studenti



Knowledge space theory

Livelli di conoscenza vs. stati di conoscenza



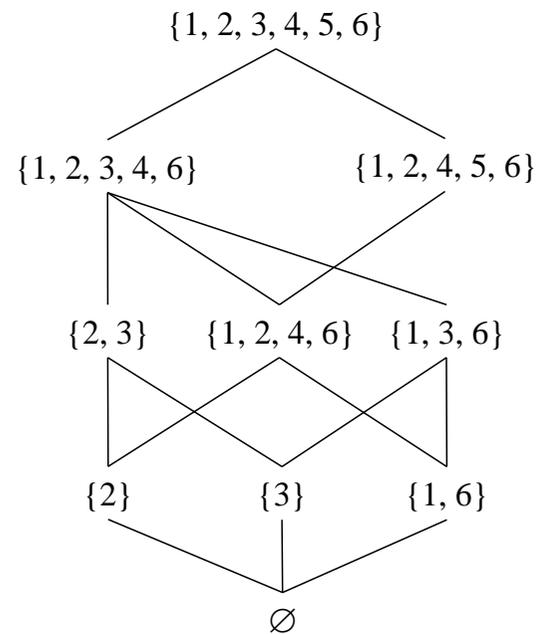
L'assessment adattivo computerizzato

- Mira a individuare lo stato di conoscenza di uno studente in un certo dominio disciplinare attraverso la somministrazione di un numero minimo di item
- La valutazione viene “cucita” sullo studente: gli item presentati non sono nè troppo facili (e quindi noiosi) nè troppo difficili (e quindi demotivanti) per lo studente
- La scelta del nuovo item da presentare si basa sulle risposte che lo studente ha dato agli item già presentati

L'assessment adattivo computerizzato

Un esempio

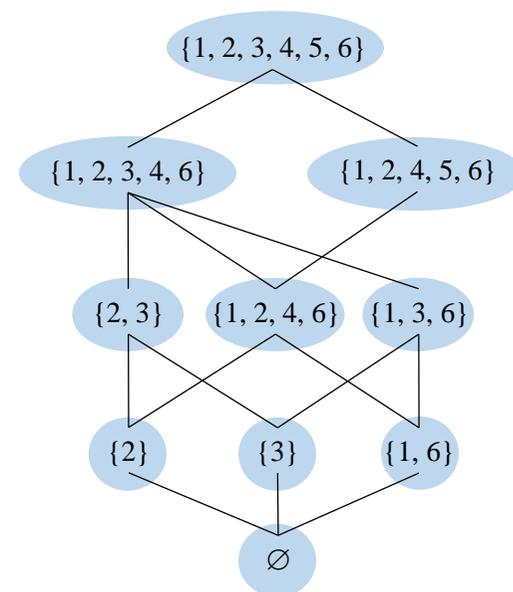
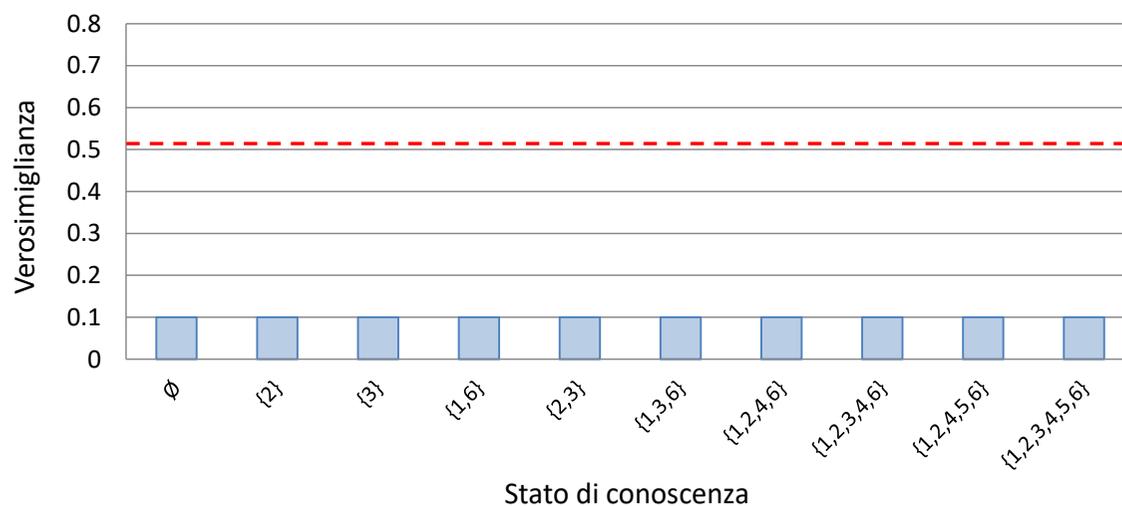
- Consideriamo la seguente struttura di conoscenza composta da 10 stati, su un dominio disciplinare con 6 item



L'assessment adattivo computerizzato

Step 0

- In assenza di informazione sullo studente all'inizio dell'assessment, si assume una distribuzione di verosimiglianza iniziale degli stati di tipo uniforme

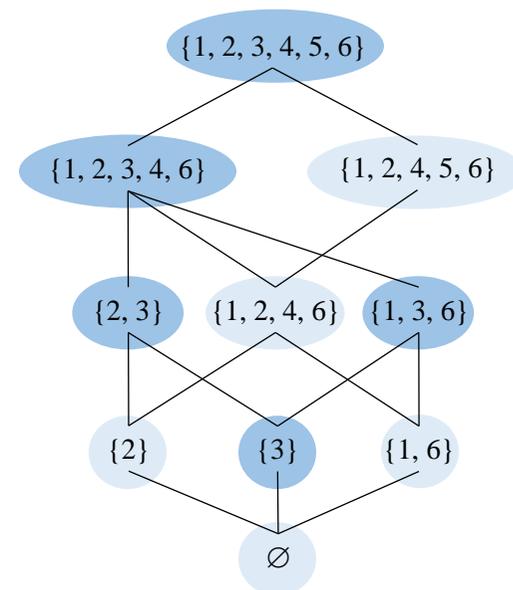
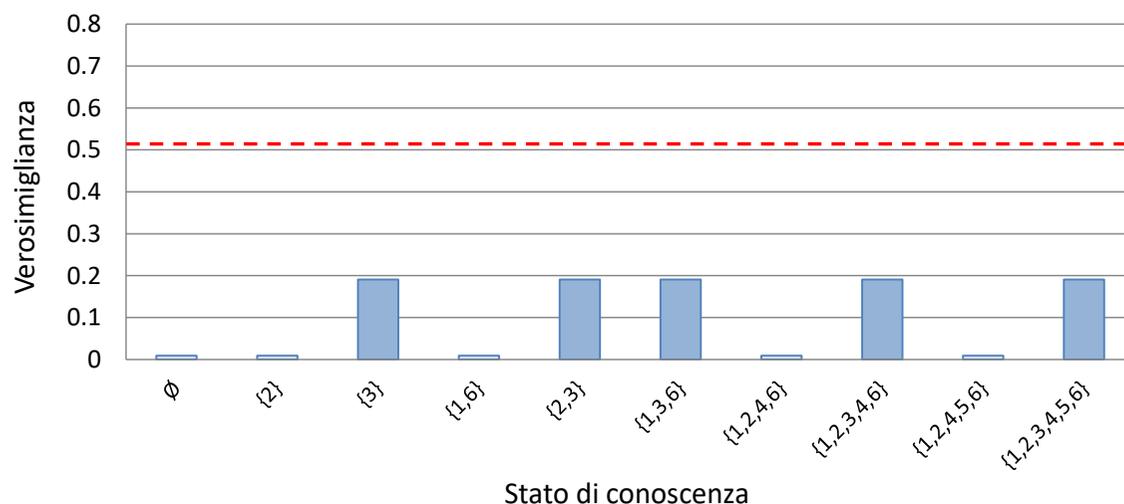


- Ad ogni step, viene selezionato l'item più informativo circa lo stato di conoscenza dello studente
- La procedura termina quando la verosimiglianza di un qualunque stato raggiunge il criterio di terminazione (.51 nell'esempio)

L'assessment adattivo computerizzato

Step 1

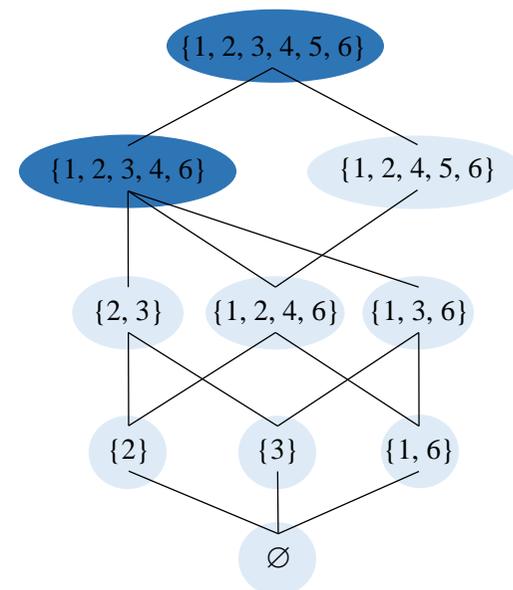
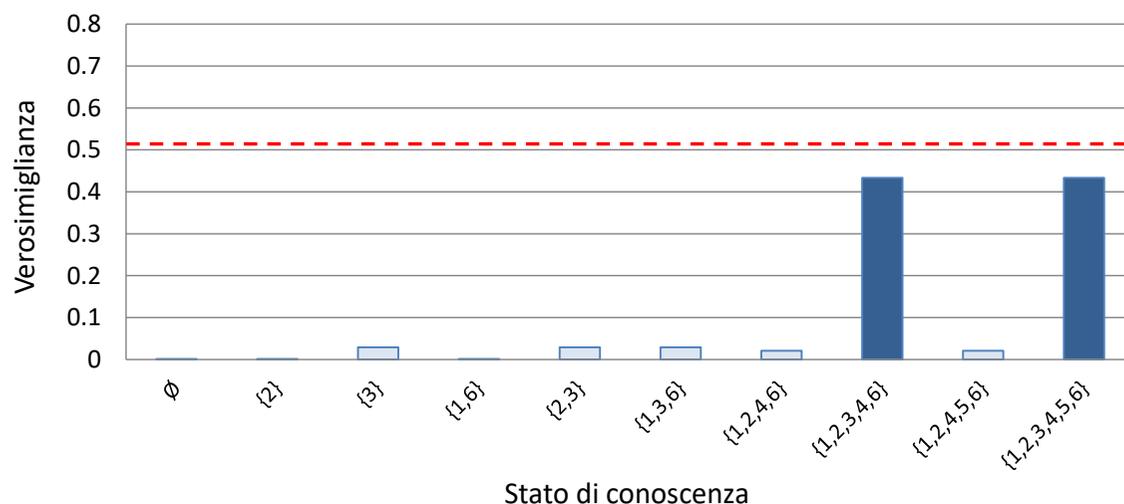
- Viene somministrato l'item 3 e la risposta è corretta
- Viene aumentata la verosimiglianza di tutti gli stati che contengono l'item 3 e diminuita quella di tutti gli stati che non lo contengono



L'assessment adattivo computerizzato

Step 2

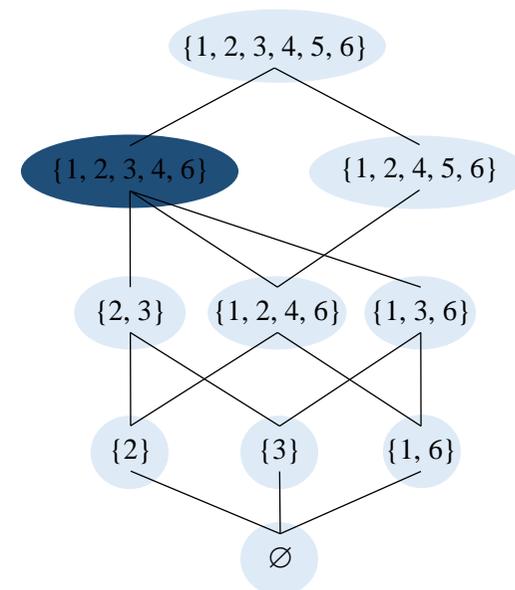
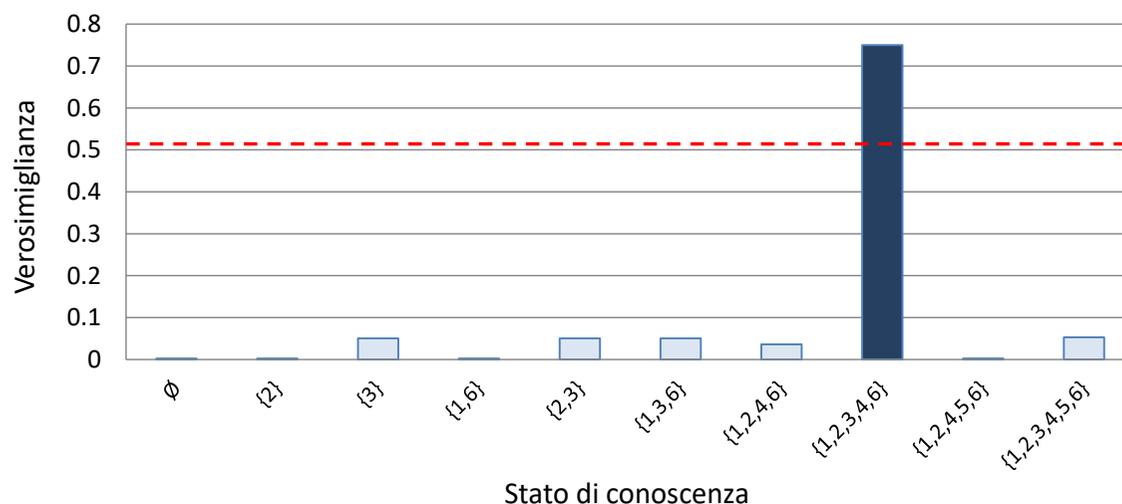
- Viene somministrato l'item 4 e la risposta è corretta
- Viene aumentata la verosimiglianza di tutti gli stati che contengono l'item 4 e diminuita quella di tutti gli stati che non lo contengono



L'assessment adattivo computerizzato

Step 3

- Viene somministrato l'item 5 e la risposta è errata
- Viene aumentata la verosimiglianza di tutti gli stati che non contengono l'item 5 e diminuita quella di tutti gli stati che lo contengono



- L'assessment termina: $\{1, 2, 3, 4, 6\}$ è lo stato dello studente

L'assessment adattivo computerizzato

- L'accuratezza (individuazione dello stato vero) e l'efficienza (numero di item somministrati) dell'assessment dipendono fortemente:
 - dalle probabilità di careless error e lucky guess degli item
 - dalla distribuzione di verosimiglianza iniziale degli stati
- Tuttavia:
 - queste quantità potrebbero essere sconosciute o inaccurate
 - se la struttura contiene molti stati e/o molti item, non è possibile raccogliere dati completi su un numero adeguato di studenti al fine di calcolare stime appropriate

Una procedura che “apprende”*

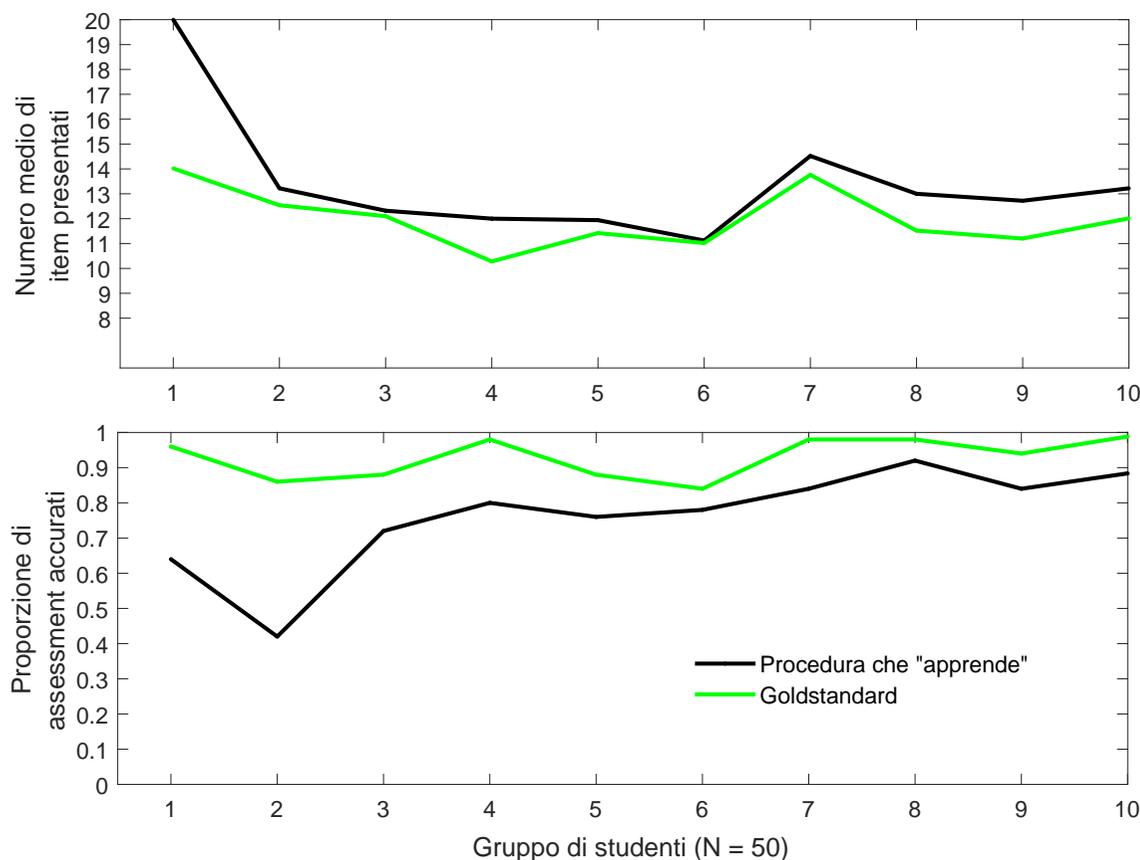
- La procedura accumula informazione dagli studenti già valutati e la utilizza per migliorare la valutazione degli studenti che seguono
- Le prime valutazioni utilizzano valori dei parametri che favoriscono l'accuratezza sull'efficienza
- Le valutazioni che seguono utilizzano stime dei parametri ottenute sui dati raccolti fino a quel momento
- Con l'aumentare dei dati raccolti, vengono calcolate stime sempre più accurate. Ciò contribuisce ad accrescere l'efficienza dell'assessment pur garantendone l'accuratezza

*Anselmi, Robusto, Stefanutti, & de Chiusole (2016)

Una procedura che “apprende”

Alcuni risultati su dati empirici

- La procedura è stata applicata alle risposte di più di 500 studenti a 20 item sulle operazioni con frazioni
- Gli studenti sono stati divisi in 10 gruppi da 50 studenti ciascuno



La valutazione delle abilità

- Nell'ambito della competence-based knowledge space theory (Cb-KST)*, si assume l'esistenza di abilità astratte sottostanti le risposte agli item di un test
- Viene definita una *mappa di abilità* che associa ciascun item alle abilità necessarie per risolverlo
- Dalle risposte (corrette o errate) agli item è possibile risalire alle abilità possedute dallo studente (cioè al suo *stato di competenza*)
- In alcuni casi, la valutazione delle abilità non è univoca: diversi stati di competenza possono delineare lo stesso stato di conoscenza

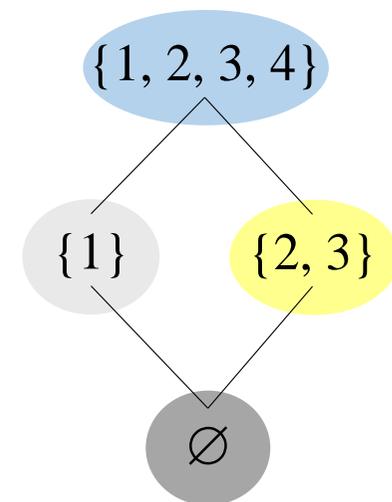
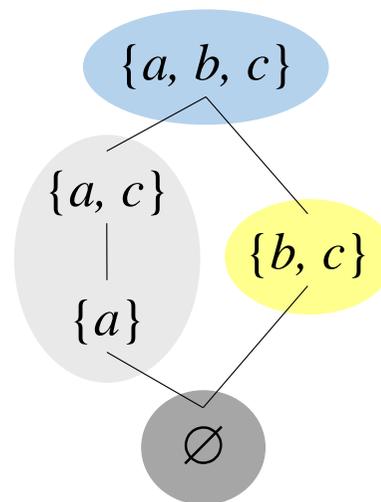
*Doignon (1994); Heller, Ünlü, & Albert (2013)

La valutazione delle abilità

Un esempio

- Consideriamo un test composto da 4 item che richiedono 3 abilità (a , b , c) per essere risolti

| Item | a | b | c |
|------|-----|-----|-----|
| 1 | ✓ | | |
| 2 | | ✓ | |
| 3 | | ✓ | ✓ |
| 4 | ✓ | ✓ | ✓ |



- I due stati di competenza $\{a\}$ e $\{a, c\}$ delineano lo stato di conoscenza $\{1\}$

La valutazione *univoca* delle abilità*

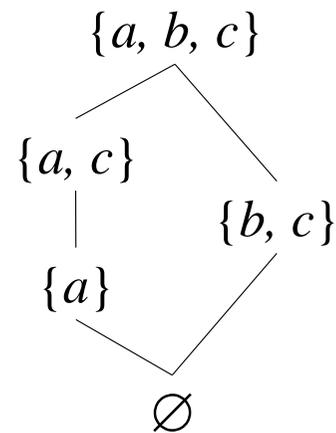
- Sono state definite le proprietà di una mappa di abilità che consentono di valutare in maniera univoca le abilità sottostanti alle risposte agli item
- È stata sviluppata una procedura iterativa che consente di:
 - sviluppare un test da zero
 - migliorare un test esistente
 - sviluppare la versione breve di un test

*Anselmi, Heller, Stefanutti, & Robusto (2018); Anselmi, Stefanutti, Heller, & Robusto (2017); Heller, Anselmi, Stefanutti, & Robusto (2017); Heller, Stefanutti, Anselmi, & Robusto (2015)

La valutazione *univoca* delle abilità

Esempio di sviluppo di un test da zero

- Consideriamo la seguente struttura di competenza definita sulle 3 abilità a , b e c



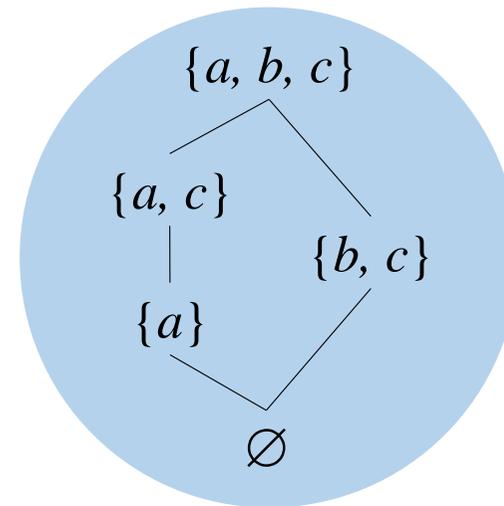
- Vogliamo costruire un test che consenta la valutazione univoca delle 3 abilità

La valutazione *univoca* delle abilità

Esempio di sviluppo di un test da zero

- Step 0: Nessun item presente

| Item | <i>a</i> | <i>b</i> | <i>c</i> |
|------|----------|----------|----------|
| ... | ... | ... | ... |



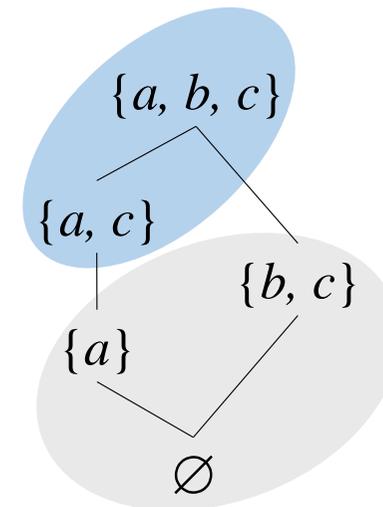
- Ad ogni step, viene introdotto l'item che divide la struttura di competenza in due insiemi (stati che risolvono l'item e stati che non lo risolvono) di ampiezza il più possibile simile

La valutazione *univoca* delle abilità

Esempio di sviluppo di un test da zero

- Step 1: Viene introdotto il primo item

| Item | <i>a</i> | <i>b</i> | <i>c</i> |
|------|----------|----------|----------|
| 1 | ✓ | | ✓ |

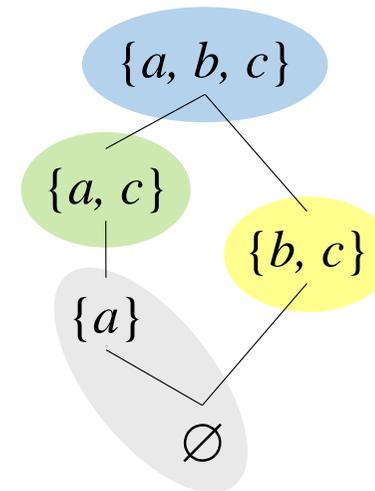


La valutazione *univoca* delle abilità

Esempio di sviluppo di un test da zero

- Step 2: Viene introdotto il secondo item

| Item | <i>a</i> | <i>b</i> | <i>c</i> |
|------|----------|----------|----------|
| 1 | ✓ | | ✓ |
| 2 | | ✓ | ✓ |

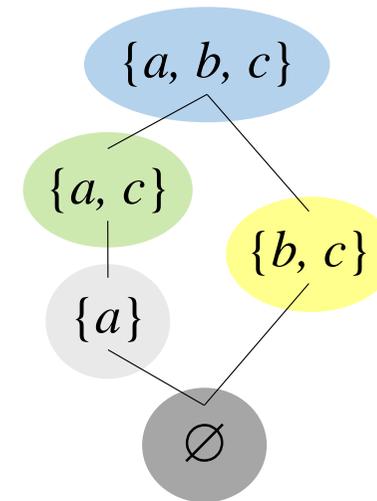


La valutazione *univoca* delle abilità

Esempio di sviluppo di un test da zero

- Step 3: Viene introdotto il terzo item

| Item | <i>a</i> | <i>b</i> | <i>c</i> |
|------|----------|----------|----------|
| 1 | ✓ | | ✓ |
| 2 | | ✓ | ✓ |
| 3 | ✓ | | |



- I tre item consentono di distinguere tutti gli stati di competenza. La procedura termina

La valutazione *univoca* delle abilità

Esempio di sviluppo della versione breve di un test

- Consideriamo il seguente test di 15 item sulle sottrazioni con frazioni
- Si assume l'esistenza di 8 abilità sottostanti la soluzione dei 15 item:
 (a) converti un numero intero in frazione, (b) separa un numero intero da una frazione, (c) semplifica prima di sottrarre, (d) trova il comune denominatore, (e) prendi in prestito da un numero intero, (f) prendi in prestito per sottrarre il secondo numeratore dal primo, (g) sottrai i numeratori, (h) semplifica il risultato

| Item | | a | b | c | d | e | f | g | h |
|------|---------------------------------|---|---|---|---|---|---|---|---|
| 1 | $\frac{5}{3} - \frac{3}{4}$ | | | | ✓ | | ✓ | ✓ | |
| 2 | $\frac{4}{8} - \frac{3}{4}$ | | | | ✓ | | | ✓ | |
| 3 | $3\frac{1}{3} - 2\frac{3}{2}$ | | ✓ | ✓ | | ✓ | | ✓ | |
| 4 | $4\frac{3}{5} - 3\frac{4}{10}$ | | ✓ | | ✓ | | | ✓ | ✓ |
| 5 | $\frac{6}{7} - \frac{4}{7}$ | | | | | | | ✓ | |
| 6 | $3 - 2\frac{1}{5}$ | ✓ | ✓ | | | | | ✓ | |
| 7 | $3\frac{7}{8} - 2$ | | ✓ | | | | | | |
| 8 | $4\frac{4}{12} - 2\frac{7}{12}$ | | ✓ | | | ✓ | | ✓ | ✓ |
| 9 | $4\frac{1}{3} - 2\frac{4}{3}$ | | ✓ | | | ✓ | | ✓ | |
| 10 | $1\frac{1}{8} - \frac{1}{8}$ | | | | | | | ✓ | ✓ |
| 11 | $3\frac{3}{8} - 2\frac{5}{8}$ | | ✓ | | ✓ | ✓ | | ✓ | |
| 12 | $3\frac{4}{5} - 3\frac{2}{5}$ | | ✓ | | | | | ✓ | |
| 13 | $2 - 1\frac{1}{3}$ | ✓ | | | | | | ✓ | |
| 14 | $4\frac{1}{10} - 2\frac{8}{10}$ | | ✓ | | | ✓ | ✓ | ✓ | |
| 15 | $4 - 1\frac{4}{3}$ | ✓ | ✓ | ✓ | | ✓ | | ✓ | |

La valutazione *univoca* delle abilità

Esempio di sviluppo della versione breve di un test

- È sufficiente utilizzare i 9 item seguenti per ottenere la stessa differenziazione tra gli stati di competenza consentita dai 15 item

| Item | | <i>a</i> | <i>b</i> | <i>c</i> | <i>d</i> | <i>e</i> | <i>f</i> | <i>g</i> | <i>h</i> |
|------|---------------------------------|----------|----------|----------|----------|----------|----------|----------|----------|
| 1 | $5\frac{5}{3} - 3\frac{3}{4}$ | | | | ✓ | | ✓ | ✓ | |
| 2 | $3\frac{3}{4} - 3\frac{3}{8}$ | | | | ✓ | | | ✓ | |
| 3 | $3\frac{1}{2} - 2\frac{3}{2}$ | | ✓ | ✓ | | ✓ | | ✓ | |
| 4 | $4\frac{3}{5} - 3\frac{4}{10}$ | | ✓ | | ✓ | | | ✓ | ✓ |
| 5 | $6\frac{6}{7} - 4\frac{4}{7}$ | | | | | | | ✓ | |
| 6 | $3 - 2\frac{1}{5}$ | ✓ | ✓ | | | | | ✓ | |
| 7 | $3\frac{7}{8} - 2$ | | ✓ | | | | | | |
| 8 | $4\frac{4}{12} - 2\frac{7}{12}$ | | ✓ | | | ✓ | | ✓ | ✓ |
| 9 | $4\frac{1}{3} - 2\frac{4}{3}$ | | ✓ | | | ✓ | | ✓ | |
| 10 | $1\frac{1}{8} - \frac{1}{8}$ | | | | | | | ✓ | ✓ |
| 11 | $3\frac{3}{8} - 2\frac{5}{6}$ | | ✓ | | ✓ | ✓ | | ✓ | |
| 12 | $3\frac{4}{5} - 3\frac{2}{5}$ | | ✓ | | | | | ✓ | |
| 13 | $2 - \frac{1}{3}$ | ✓ | | | | | | ✓ | |
| 14 | $4\frac{1}{10} - 2\frac{8}{10}$ | | ✓ | | | ✓ | ✓ | ✓ | |
| 15 | $4 - 1\frac{4}{3}$ | ✓ | ✓ | ✓ | | ✓ | | ✓ | |

KnowLab*

- È un intelligent tutoring system per l'assessment adattivo delle conoscenze e l'apprendimento personalizzato a livello universitario
- Utilizza concetti teorici e modelli matematici sviluppati nell'ambito della KST e della Cb-KST
- Si compone di due moduli:
 - Modulo di assessment: valuta in maniera adattiva le conoscenze dello studente; produce un report di valutazione
 - Modulo di apprendimento: fornisce allo studente *soltanto* contenuti didattici relativi alle abilità che è pronto ad apprendere
- Lo studente interagisce con il sistema passando di continuo tra i due moduli

*de Chiusole, Stefanutti, Anselmi, Robusto, & Longo (2013)

Stat-KnowLab*

www.stat-knowlab.unipd.it



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

| CHE COS'È STAT KNOWLAB | CHI SIAMO | AMBITI DISCIPLINARI | HELP |
|--|-----------|---------------------|---|
| | | | <h3>PERCHÈ ISCRIVERSI</h3> <p>Stat Knowlab è una piattaforma gratuita aperta a tutti per migliorare la conoscenza della statistica. Attraverso un sistema interattivo, Stat Knowlab testa il livello raggiunto dall'utente e propone esercizi sempre più impegnativi.</p> <p>ISCRIVITI ></p> <p>E-MAIL <input type="text"/></p> <p>PASSWORD <input type="text"/></p> <p>SUBMIT</p> |
| <p>© 2018 Università di Padova - Tutti i diritti riservati</p> | | | |

*de Chiusole, Robusto, Stefanutti, & Anselmi (2017)

Stat-KnowLab

Esempio di report di valutazione

ARGOMENTO: Test della Media con il t di Student

Le tue conoscenze attuali, relative a questo argomento, coprono il 62% del totale. Quindi la tua preparazione in questo ambito è al momento **sufficiente**.



QUELLO CHE GIA' SAI FARE

- Calcolare la statistica t, nel caso in cui, in una ricerca empirica, sia coinvolto un solo campione
- Prendere la decisione statistica opportuna (Accettare o non l'ipotesi nulla) nel caso della verifica di ipotesi con il t test e un ipotesi monodirezionale sinistra
- Prendere la decisione statistica opportuna (Accettare o non l'ipotesi nulla) nel caso della verifica di ipotesi con il t test e un ipotesi bidirezionale
- Prendere la decisione statistica opportuna (Accettare o non l'ipotesi nulla) nel caso della verifica di ipotesi con il t test e un ipotesi monodirezionale destra
- Individuare correttamente il valore del t critico, nel caso della verifica di un ipotesi bidirezionale
- Individuare correttamente il valore del t critico, nel caso della verifica di un ipotesi monodirezionale destra
- Individuare correttamente il valore del t critico, nel caso della verifica di un ipotesi monodirezionale sinistra
- Individuare correttamente quando applicare il test t, nel caso di un campione, per procedere alla verifica di ipotesi



QUELLO CHE ANCORA NON SAI FARE

- Calcolare la statistica t, nel caso in cui, in una ricerca empirica, siano coinvolti due campioni indipendenti
- Calcolare la statistica t, nel caso in cui, in una ricerca empirica, siano coinvolti due campioni non indipendenti
- Individuare correttamente quando applicare il test t, nel caso siano coinvolti due campioni indipendenti, per procedere alla verifica di ipotesi
- Individuare correttamente quando applicare il test t, nel caso siano coinvolti due campioni non indipendenti (o appaiati), per procedere alla verifica di ipotesi
- Condurre la verifica di ipotesi sulla media con il t test



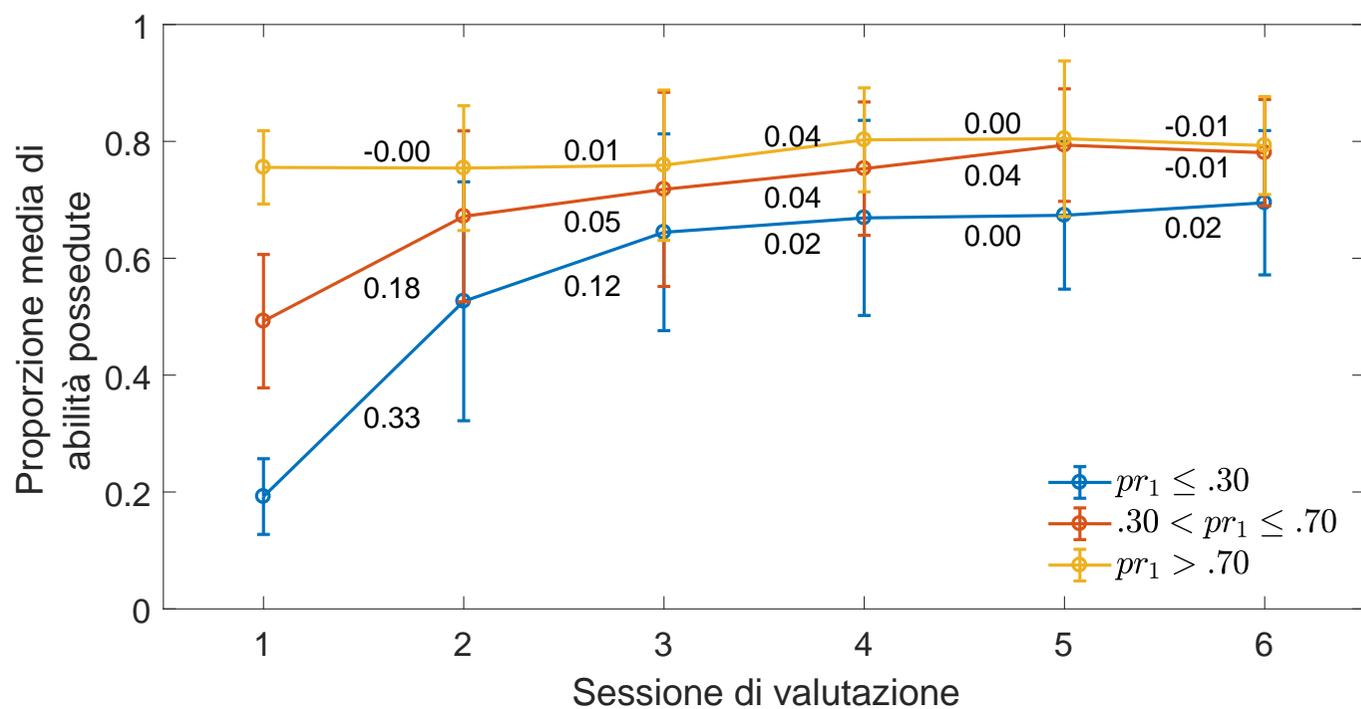
QUELLO CHE SEI PRONTO AD APPRENDERE

- Condurre la verifica di ipotesi sulla media con il t test

Stat-KnowLab

Alcuni risultati della sperimentazione empirica

- 308 studenti del corso di Psicometria (Università di Padova) hanno utilizzato Stat-KnowLab per un mese
- Dominio “statistiche descrittive”: 24 item, 20 abilità, 4302 stati



Sviluppi futuri

- Sviluppo della piattaforma KnowLab (ad es., implementazione di strumenti che consentano l'interazione tra gli studenti)
- Utilizzo di KnowLab in altri contesti di formazione (ad es., scolastici e lavorativi) e domini disciplinari
- Sviluppo e applicazione di modelli formali per l'analisi dei processi di apprendimento

Grazie per l'attenzione!

pasquale.anselmi@unipd.it