

Uno strumento di analisi statistica delle prove d'esame di programmazione

**Corso di Laurea Triennale in Informatica
Tesi di Laurea di Marco Cordoni, matricola 855535**

Relatore: Prof. Massimo Santini

Anno Accademico 2016/2017

In questo progetto ho sviluppato uno strumento che permette di studiare dati e consente al docente di analizzare le valutazioni ottenute dagli studenti negli esami e nei laboratori di Programmazione.

L'obiettivo principale della tesi è la stima delle difficoltà degli esercizi e la formulazione di ipotesi relative al livello di complessità riscontrato, usufruendo dei dati ricevuti e utilizzando la libreria di funzioni definita.

Per ciascuna delle sessioni d'esame e di laboratorio conosco quali studenti vi hanno partecipato, gli esercizi che sono stati affrontati e il livello di correttezza delle soluzioni proposte.

Lo strumento teorico fondamentale che ho acquisito per l'elaborazione della tesi è detto "Item response theory" (IRT) che permette di misurare la conoscenza dei soggetti e la difficoltà degli esercizi inerenti ai dati raccolti tramite test, l'approccio che l'IRT mette a disposizione è legato al modello parametrico adottato.

Per la realizzazione della libreria ho applicato il linguaggio R e ho adottato metodi di codifica che favoriscono la leggibilità e comprensione del codice, come l'uso di nomi esplicativi, la parametrizzazione delle costanti presenti nelle funzioni, la modularizzazione del codice e l'introduzione della documentazione.

Dopo aver definito gli stimatori per i parametri ricercati ho validato la loro correttezza mediante un processo che inizia con la simulazione dei dati di una sessione d'esame, procede applicando ai dati generati gli stimatori definiti e si conclude confrontando i valori ottenuti con quelli utilizzati per simulare la sessione analizzata, evidenziando le differenze riscontrate.

Procedendo nello sviluppo del progetto ho affrontato un problema che ricorre frequentemente quando si trattano queste informazioni, ovvero l'assenza di dati per il mancato invio dell'esercizio da parte dello studente.

Ho implementato quindi dei metodi per gestire i valori mancanti: li ho considerati come esercizi falliti oppure li ho rimossi prima della computazione delle stime, ho confrontato i risultati ottenuti da entrambi i metodi e ne ho tratto le dovute considerazioni.

Ho trasformato i dati forniti dal docente perché gli stimatori potessero analizzarli, quindi ho introdotto una funzione per modificare la loro forma e poter procedere con la fase di analisi dei dati e la formulazione di ipotesi.

Infine ho confrontato gli stimatori impiegati con una tecnica di stima naïf e ho formulato osservazioni sulle differenze che ho rilevato nei risultati.

Un futuro sviluppo per questo progetto potrebbe essere l'implementazione di stimatori più complessi e tecniche di validazione più accurate; la trattazione diretta dei dati evitando di trasformarli per adattarli agli stimatori da me definiti rimane un problema rilevante lasciato aperto.