IRT

Natural Language Processing
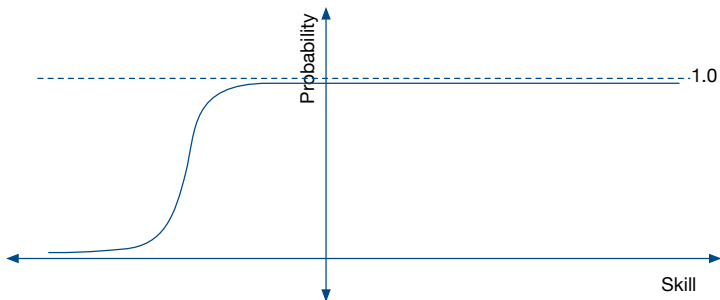
University of Maryland

Exercises

Drawing IRT Curves

- Low difficulty
- High difficulty
- High discrimination
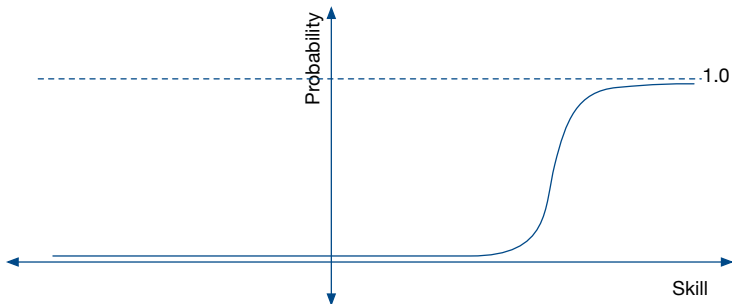- Low discrimination
- Low feasibility

Drawing IRT Curves

- Low difficulty

- High difficulty
- High discrimination
- Low discrimination
- Low feasibility

Drawing IRT Curves

- Low difficulty
- High difficulty

- High discrimination
- Low discrimination
- Low feasibility

Drawing IRT Curves

- Low difficulty
- High difficulty
- High discrimination

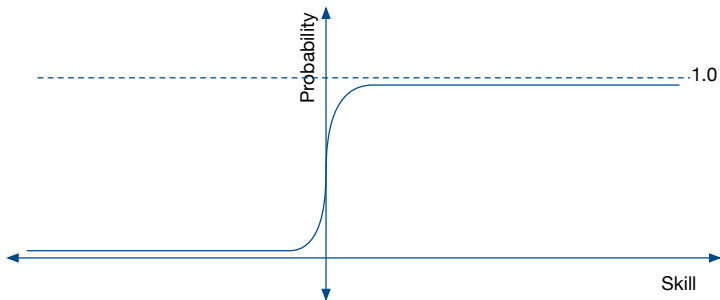- Low discrimination
- Low feasibility

Drawing IRT Curves

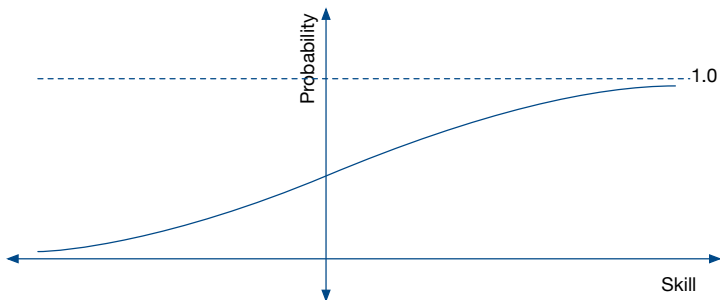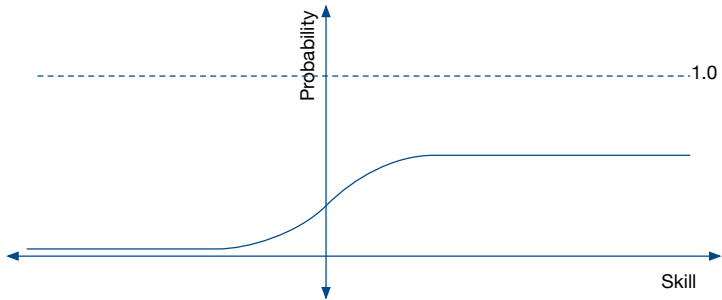- Low difficulty
- High difficulty
- High discrimination
- Low discrimination

- Low feasibility

Drawing IRT Curves

- Low difficulty
- High difficulty
- High discrimination
- Low discrimination
- Low feasibility

Examples of Questions

What's an example of:

- A low difficulty question
- A high difficulty question
- An infeasible question
- A high discrimination question
- A negative discrimination question

- **What is the analog to difficulty, skill, and correctness in the Bradley Terry model?**

- **What is the analog to difficulty, skill, and correctness in the Bradley Terry model?**
  - ▶ Difficulty = Rating / value of response 1
  - ▶ Skill = Rating / value of response 2
  - ▶ Corectness = Preference

Bradley Terry vs. IRT

- **What is the analog to difficulty, skill, and correctness in the Bradley Terry model?**
- **A Bradley Terry model doesn't have something like feasibility. What would that mean? Why does it not make sense.**

- **What is the analog to difficulty, skill, and correctness in the Bradley Terry model?**

- **A Bradley Terry model doesn't have something like feasibility. What would that mean? Why does it not make sense.**
  In original BT models, this is called "home field advantage". A cap on probability of one of the responses being preferred. It doesn't make sense because Bradley Terry is symmetric.

- **What is the analog to difficulty, skill, and correctness in the Bradley Terry model?**
- **A Bradley Terry model doesn't have something like feasibility. What would that mean? Why does it not make sense.**
- **A Bradley Terry model doesn't have something like discrimination. What would that mean?**

- **What is the analog to difficulty, skill, and correctness in the Bradley Terry model?**

- **A Bradley Terry model doesn't have something like feasibility. What would that mean? Why does it not make sense.**

- **A Bradley Terry model doesn't have something like discrimination. What would that mean?**
  That for some prompts, there is a clear unambiguous answer.

Bradley Terry vs. IRT

- **What is the analog to difficulty, skill, and correctness in the Bradley Terry model?**

- **A Bradley Terry model doesn't have something like feasibility. What would that mean? Why does it not make sense.**

- **A Bradley Terry model doesn't have something like discrimination. What would that mean?**

- **This isn't so crazy. . . what's a circumstance where that might apply?**

Bradley Terry vs. IRT

- **What is the analog to difficulty, skill, and correctness in the Bradley Terry model?**

- **A Bradley Terry model doesn't have something like feasibility. What would that mean? Why does it not make sense.**

- **A Bradley Terry model doesn't have something like discrimination. What would that mean?**

- **This isn't so crazy. . . what's a circumstance where that might apply?**
  Factoid QA without explanation, quotes.