Test Item Analysis and Statistics
What they are and how to use them

Presented by
Lawrence Preston, *Committee on Accreditation*

Test Item Analysis and Statistics

- Current practices

- Exam results – what they tell us

- Item analysis – what this tells us

- Exercises

- Wrap up

Current Practices

- What does your entity do to validate item and exam results?
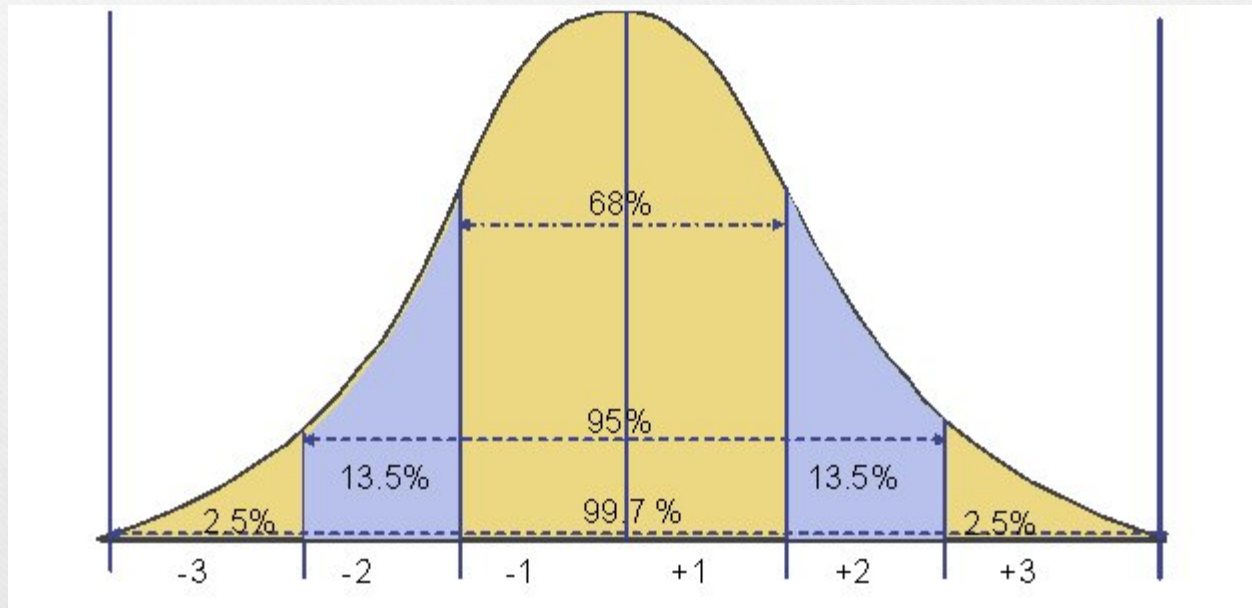
- What is the predominant practice?

Exam Results

- Norm referenced exam results

  - Comparing individuals

  - Expectation of bell-shaped curve

  - Probability of results (0 to 1.0)

    - .6826 + or – 1 Standard Deviation

    - .9554 + or – 2 Standard Deviations

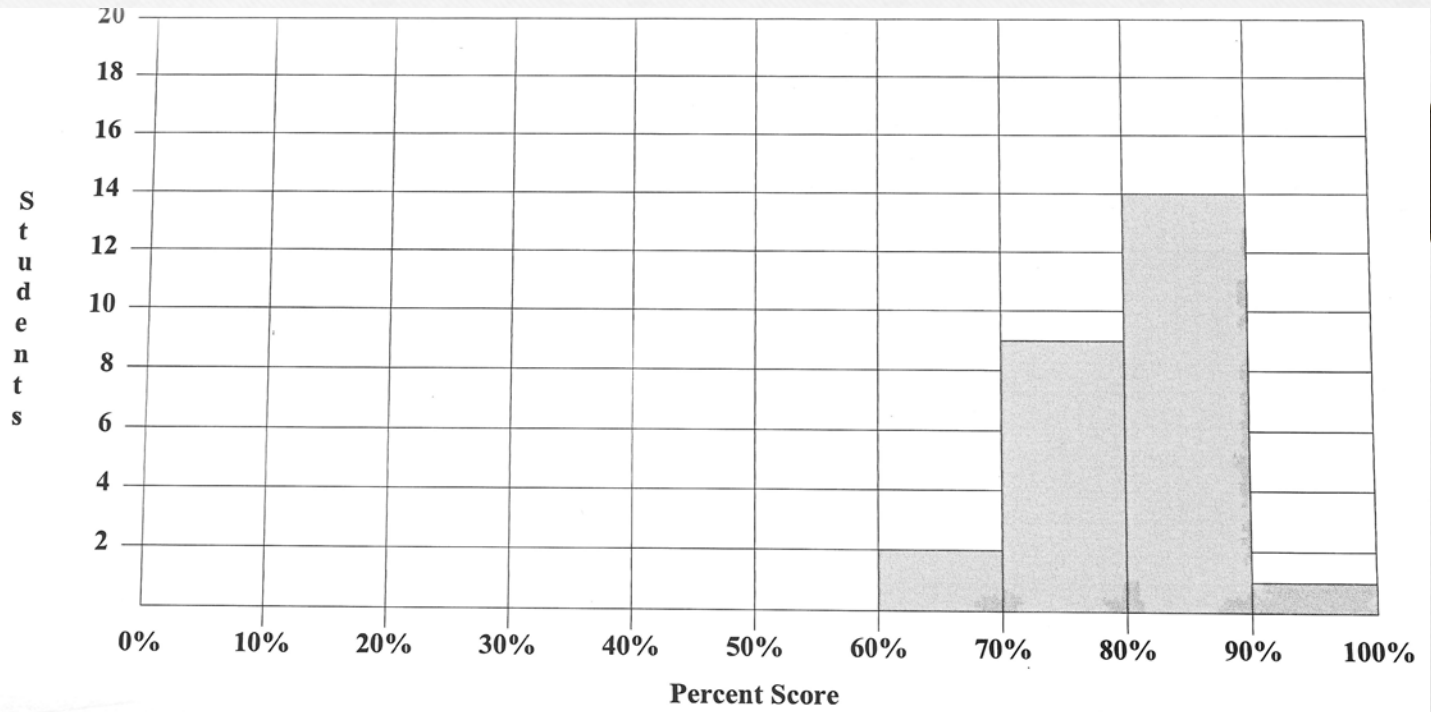   - .9974 + or – 3 Standard Deviations

Exam Results

Exam Results

- Criterion referenced exam results

  - Compares individuals against a standard or set of criteria

  - May or may not produce a bell-shaped curve

  - The graph of test results is provided by several commercial packages

Graph of Test Results

Exam Results

- Don't expect bell-shaped curve for your certification exams

- There are a lot of variables

Test Item Statistics

- Difficulty – P Factor

- Discrimination – Point Biserial Correlation Coefficient, rpb

- These 2 should be considered together in making decisions

Difficulty

- Often called the P Factor

- Values from .0 to 1.0

- .75 means that 75% of candidates got the answer correct

- Difficulty levels range from low (1.0) to high (0.)

- What are the ramifications of too hard or easy?

Difficulty

- Question #1 P = .12

- Question #2 P = .56

- Question #3 P = .91

- Does the criticality of the JPR affect this decision?

Discrimination

- Known as Point Biserial Correlation Coefficient or rpb

- It correlates the average score of those who answered correctly versus the average score of those who answered incorrectly for each question. It uses the scores of the highest and lowest scoring 25%* of students for each

- Correlations range from -1.0 to +1.0

Discrimination – Correct Answers

- Positive 1.0 correlation means there is a perfect correlation between the higher scoring group with correct answers on the test (All the higher scoring students picked the correct answer and none of the lower scoring students did)

- Negative 1.0 means there is a perfect negative correlation between the lower scoring group with incorrect answers on the test (All picked wrong) and none of the higher scoring students did

- WHAT?

Discrimination

- Let's do a deep dive into this by looking at the correct answer statistics

    - A question's correct response with a rpb of .95 indicates that a high percentage of high performing students selected the correct response compared to the low scoring group

    - A question's correct response with a rpb of -.75 indicates that a high percentage of high performing students did not select the correct response compared to the low scoring group

Discrimination

- Examples – **correct response rpb** =
- 1.0,
- -.26,
- .34,
- -.76
- 0.00 what does this mean?

Discrimination Continued

- Let's look at the low performing group

- In theory the lower scoring group should select more wrong answers on any particular question than the higher group

- This statistic identifies the differences between the 2 groups or it discriminates between the 2 groups

- This would result in a negative rpb

Discrimination Continued

- An example on a test with 40 administrations:
  - **Response A** is the correct answer on a test
  - Response B is the incorrect answer on a test

Question 1	40 Responses	**A**	B
	P Factor .50	**20**	20
		1.0	-1.0

Discrimination Continued

- Overview of Item Statistic Report for FO II

<mark>5.2.1 HR POL & PROC 19</mark>

6	25	0	2	6	**15**	2
MCS	0.60	0.00	0.08	0.24	**0.60**	0.08
1.00	83.17	0.00	71.46	78.24	**83.17**	73.76
	+0.156	-	-0.281	+0.071	**+0.156**	-0.112

Discrimination Continued

- Overview of Item Statistic Report

5.2.1 HR POL & PROC 19

6	25	0	2	6	15	2
MCS	0.60	0.00	0.08	0.24	**0.60**	0.08
1.00	83.17	0.00	71.46	78.24	**83.17**	73.76
	+0.156	-	-0.281	+0.071	**+0.156**	-0.112

Discrimination Continued

- Overview of Item Statistic Report

5.2.1 HR POL & PROC 19

6	25	0	2	6	**15**	2
MCS	0.60	0.00	0.08	0.24	**0.60**	0.08
1.00	83.17	0.00	71.46	78.24	**83.17**	73.76
	+0.156	-	-0.281	+0.071	**+0.156**	-0.112

Discrimination Continued

- Overview of Item Statistic Report

5.2.1 HR POL & PROC 19

6	25	0	2	6	15	2
MCS	0.60	0.00	0.08	0.24	**0.60**	0.08
1.00	83.17	0.00	71.46	78.24	**83.17**	73.76
	+0.156	-	-0.281	+ 0.071	**+0.156**	-0.112

Discrimination Continued

- Overview of Item Statistic Report

5.2.1 HR POL & PROC 19

6	25	0	2	6	**15**	2
MCS	0.60	0.00	0.08	0.24	**0.60**	0.08
1.00	83.17	0.00	71.46	78.24	**83.17**	73.76
	+0.156	-	-0.281	+ 0.071	**+0.156**	-0.112

The Next Step

- Having identified questions with potential problems what could be the possible source of the problems?

1.

2.

3.

4.

What would you do for each of the above?

Test Item Analysis and Statistics

- Questions?

- Contact lpreston@theproboard.org if needed or Brian Brauer at proboardbrian@gmail.com