

第8章 テスト得点の比較

8.1 テストの等化

■ 等化 (equating): 同一能力を測る複数のテスト形式間で不公平がないように、尺度を調整する。

- ・ equal construct requirement: 測定対象の構成概念が同一
- ・ equal reliability requirement: 信頼性が等しい
- ・ symmetry requirement: 対称性が保たれている
- ・ equality requirement: どちらのテストを受けても同質
- ・ population invariance requirement: 母集団が不変

※ 実際には全ての概念を満たす尺度調整は非常に難しい。簡単な等化方法では、いくつかは既に満たされているという前提で行っている。

■ 主得点尺度 (primary score scale): 複数のテスト形式の得点を調整する際、その中心となる形式の得点尺度 (Peterson et al., 1989)。

8.2 平均等化法

■ 平均等化法 (mean equating): 線形等化法の一つ。得点分布 (平均以外) などが等しいという仮定の下、平均を揃える。

$$\rightarrow Y = X + a \text{ (点)}$$

※ 2 フォームの得点分布が異なる場合は、正確な等化が不可能。この場合、加点することで満点を超過してしまうケースが生じる。また得点が高いほどさらに高得点を取るのが難しいので、全受験者に一定の得点を加算するのが公平かという問題もある。

8.3 線形等化法

■ 線形等化法 (linear equating method): 得点の分布型 (平均と標準偏差以外) が同一という前提の下、同一の集団、または同一と見られる 2 つの集団に 2 つのフォームのテストを受験してもらう。平均と標準偏差から算出される z 得点を等しいとし、変換式を導く。

$$X_z = Y_z \Rightarrow (X - \bar{X}) / S_x = (Y - \bar{Y}) / S_y \Rightarrow Y = S_y / S_x \times X + \bar{Y} - S_y / S_x \times \bar{X} \Rightarrow Y = aX + b$$

※ $a = 1$ のとき平均等化。X, Y の逆も成り立つ。

- ・ 一度等化式を確立すれば、異なるテストフォームで受験した集団がいても、比較が可能になる。
- ・ 係留項目デザイン: 一部共通問題 (anchor item) を含むテストデザイン。同一グループに 2 つのテストを受けてもらえない場合、2 つ以上のフォームを 1 つの主得点に等化させたい場合に等化が可能になる (9章参照)。

8.4 等パーセンタイル等化法

■ 等パーセンタイル等化法 (equipercentile equating method): 非線形等化法の一つ。2 テストのパーセンタイル順位が等しくなるように得点を等化する。

- ・ パーセンタイル: 得点の小さい方から順位を数えたときの得点の相対的な順位。

$$\text{パーセンタイル} = / \text{低いその順位までの度数} \times 100$$

- ・ excel で「データポイントを円滑図でつないだ散布図」を描いたり、換算表を作成しておくとう便利。

(ex.) テスト Y が 30 点のときパーセンタイル順位が 57.2%

→ テスト X でこれに最も近いパーセンタイル順位は 54.6% で、その時の得点は 40 点

※ 2 テスト間で必ずしも同じパーセンタイル順位が得られるわけではない。特にテスト項目数が大きく異なっている場合は、あまり厳密な変換表が作成できない。

- ・ 等パーセンタイル等化法の利点
- ✓ 2つのテストの得点分布に関係なく等化できる。
- ✓ 線形等化と異なり、上限と下限の下の素点を超える心配がない。最高点・最低点がテスト間で対応している。

8.5 偏差値比較法

- 一方のテスト得点を主得点にする理由がない場合、両方とも同じ基準にして比較する。
 - 偏差値比較 (Z-score comparison): 各テスト得点の平均・標準偏差を同等にして比較する。
(ex.) z 得点: 平均: 0 点、標準偏差: 1 点 / 偏差値 (Z, T 得点): 平均: 50 点、標準偏差: 10 点
- ・ 異なるテストフォームの比較が素点より正確で、全体における自分の位置づけもできる。
- ・ 複数科目の成績を合計する場合も、標準得点を足し合わせる方法が用いられることがある。素点だと標準偏差の大きな科目の影響が大きい (Petersen et al., 1989)。

8.6 正規化された偏差値比較

- 線形比較による標準得点比較は、2つの得点分布が異なる歪み方をしているとき、同じ標準得点でも順位が同じになるとは限らない。
- 正規化された偏差値 (T-score, normalized T-score): どちらのテスト得点も平均 50、標準偏差 10 の正規分布になる尺度に換算する。偏差値の値で、偏差値比較より正確にパーセンタイルを特定することができる。
- ※ この方法は、受験集団における自分の位置しか示していない。1, 2 回目の受験集団が異なると、集団における自分の相対的位置が高くなっただけなのか、実際に実力が付いたのか特定できない。

8.7 回帰分析

- 回帰分析による得点の予測: X の得点から Y の得点を最も予測できる一時式をたてる。
- ※ 線形等化と異なり、テスト Y から X の回帰式で得られる得点と、X から Y の回帰式で得られる得点は一致しない (等化条件の対称性が保たれていない)。
 - ・ 活用法 1: 外部試験を目安にしたいが、学生全員に高い料金を払って受けてもらうことが難しい場合に利用できる。
- (ex.) 同一グループに TOEFL と同じ能力を測定する自作のテストを受験してもらう。そして自作テストの得点が TOEFL の何点に相当するかを変換式や換算表を作成しておく。
 - 自作テストを受けるだけで TOEFL 得点を予測できる。
- ・ 活用法 2: 実際に受験できなかった場合に、複数のテストから回帰分析で得られる変換式により得点を予測できる。