

尺度得点について

新しい「日本語能力試験」（新試験）で導入された「尺度得点」とは、受験者の日本語能力と試験結果（得点）を、より公平に対応づけるため、共通の尺度（ものさし）に基づいて表示した得点です。尺度得点を利用することで、いつの試験を受験したかにかかわらず、常に同一の基準の下で日本語能力を測定することが可能になります。以下に尺度得点の利点と算出過程について説明します。

【尺度得点の利点】

試験問題の難易度を異なる試験の間で完全に一定に保つことは、試験問題作成の過程において専門家による綿密な分析・検討を経たとしても極めて困難です。このため、受験者が正答した問題の配点を単純に加算する「素点」をテスト得点とすると、テスト得点は試験の難易度に依存することになります。受験の時期が異なり、試験問題の難易度が異なっている場合には、同じ日本語能力の受験者であってもそれぞれの試験のテスト得点は同一にはなりません。そのため、素点で受験者の日本語能力を表した場合には、試験問題の難易度に影響を受け、素点が日本語能力の実態を適切に反映しなくなる可能性について考慮しなければなりません。

これに対して、新試験で導入された「尺度得点」には、「試験の難易度と独立して日本語能力を評価し、統一の尺度に基づいて数値化できる」という、能力測定の方法論上、大変有益な特長があります。この特長により、受験者の日本語能力が同じならば、いつの試験を受験しても、同一のテスト得点を返すことが可能になっています。尺度得点の導入により、同じレベルで同じ得点区分ならば、2つの異なる試験で算出された「尺度得点の差」を「日本語能力の差」として、試験の難易度から独立して解釈することが可能になります。

【尺度得点の算出過程】

試験の難易度に依存しないテスト得点（尺度得点）を算出する具体的な手続きは、項目応答理論（Item Response Theory; IRT）という統計的テスト理論に基づいています。この手続きは、正答数に基づいたテスト得点（素点）の算出法とは全く異なります。

尺度得点は、どの問題にどのように解答しているか（正答か誤答か）、すなわち「解答のパターン」によって数理的に決定されます。例えば10問（項目）の試験問題で構成される試験では、最大で1024通り（ 2^{10} 通り）の解答パターンが存在します。新試験の尺度得点算出の過程では、これらの解答パターンに基づいて受験者の日本語能力を、ひとつの得点区分につき0～60点（N4、N5の「言語知識（文字・語彙・文法）・読解」では0～120点）の尺度上に位置付けていきます。10問の場合で最大で1024通りある解答パターン（実際の試験ではもっと多くなります）を61のグループに分類する

のですから、互いに異なる解答パターンに対して、同一の尺度得点になる場合も生じてきます。したがって、ある2名の受験者について、互いに正答数や解答パターンは違っていても、尺度得点在同一になる場合もあります。逆に、正答数は同一であっても解答パターンが異なるため、尺度得点異なる場合もあります。

尺度得点の算出過程に関する詳細については下記の文献を参照してください。

【文献】

池田央（1994）『現代テスト理論』 朝倉書店

豊田秀樹（2002）『項目反応理論[入門編] -テストと測定の科学-』 朝倉書店

渡辺直登・野口裕之（編）（1999）『組織心理測定論-項目反応理論のフロンティア-』 白桃書房

合格点の設定について

新試験の合否判定水準は、統計分析の結果を踏まえ、旧試験の合否判定水準と整合性をもつように設定されています。これにより、旧試験の1級、2級、3級、4級に合格できる日本語能力をもった受験者は、それぞれ新試験のN1、N2、N4、N5に合格できる日本語能力をもっていると解釈することが可能です。旧試験での対応級が存在しないN3については、旧試験の2級と3級の合否判定水準における日本語能力レベルを統計学的に分析し、旧試験の2級と3級の合否判定水準の間にN3の合格点が収まるように設定されています。