項目反応理論

1 項目分析

1.1 信頼性

テストで測られる学力や性格、技能は真の得点呼ばれ、無視できない誤差を伴ってしか測定できないものである。そこで、テスト得点 x を、真の得点 t と誤差 e の和

$$x = t + e \tag{1}$$

として表現する。

ここで誤差 e と t に関する期待値を

仮定１： $E[e] = 0$

$+$ の誤差の人、$-$ の誤差の人がそれぞれ存在するが平均すれば 0 になるだろうという仮定。

仮定 2： $E[et] = 0$

真の点数が高いほど誤差が大きくなるなどのような関連性はなく、誤差はに無相関に散らばるという仮定。

これら 2 の仮定の下で、テスト得点 x の母平均は

$$\mu_x = E[x] = E[e + t] = E[t] = \mu_t \tag{2}$$

となり真の得点の母平均に一致する。また、テスト得点 x 分散は、

$${\sigma_x}^2 = {\sigma_e}^2 + {\sigma_t}^2 \tag{3}$$

のように真の得点の分散と、誤差の分散の単純な和に分解される。このとき、真の得点の分散とテスト得点の分散の比を信頼性係数という。

テスト得点 x の信頼性係数

$$\rho = \frac{{\sigma_t}^2}{{\sigma_e}^2 + {\sigma_t}^2} = 1 - \frac{{\sigma_e}^2}{{\sigma_x}^2} = \frac{{\sigma_t}^2}{{\sigma_x}^2} \tag{4}$$

1.1.1 タウ等価測定

テスト得点 x だけでなく、項目得点 x_i もまた、

$$x_i = t_i + e_i \tag{5}$$

というように古典的テストモデルに従っているものとし、項目 i と項目 j に関して、

1. $E[e_i] = 0$
2. $E[t_i e_i] = 0$ (i=j のときを含む)
3. $E[e_i e_j] = 0$ ($i \neq j$ の場合のみ)

を仮定する。この仮定を満たしているとき、項目を互いに同族測定という。

また、

$$t_1 - \mu_{t_1} = \cdots = t_i - \mu_{t_i} = \cdots = t_n - \mu_{t_n} = t^* \tag{6}$$

が、成り立つとき項目は互いにタウ等価測定であるという。言い換えると、タウ等価測定とは、

$$t - \mu_t = nt^* \tag{7}$$

であり、すべての項目は同じ真の得点を測っているという仮定である。したがって、タウ等価測定のもとでは

$$\sigma_{t_1}{}^2 = \cdots = \sigma_{t_i}{}^2 = \cdots = \sigma_{t_n}{}^2 = \sigma_{t^*}{}^2 \tag{8}$$

のように、真の得点の分散は添え字によらず一定値 $\sigma_{t^*}{}^2$ となる。さらに任意の 2 つの項目得点の共分散も

$$C[x_i, x_j] = \sigma_{t^*}{}^2 \tag{9}$$

となる。

1.1.2 クロンバックの α 係数

タウ等価測定が成り立てば、信頼係数は

$$\rho = \frac{V[t]}{\sigma_x{}^2} = \frac{V[nt^*]}{\sigma_x{}^2} \tag{10}$$

項目 i と項目 j の標本共分散 s_{ij} はすべて $\sigma_{t^*}{}^2$ の推定量に使用できる。今回は精度を上げるために、$n(n-1)$ 個の標本共分散の平均

$$\hat{\sigma}_{t^*}^2 = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{i \neq j, j-1}^{n} s_{ij} \tag{11}$$

を推定量として利用する。式 (11) を式 (10) に代入して $s_x{}^2$ を $\sigma_x{}^2$ の推定量に利用すると

$$\alpha = \frac{n}{n-1} \frac{\sum_{i=1}^{n} \sum_{i \neq j, j-1}^{n} s_{ij}}{s_x{}^2} = \frac{n}{n-1} \left(\frac{1 - \sum_{i=1}^{n} s_i{}^2}{s_x{}^2} \right) \tag{12}$$

のようにテスト得点 x の信頼性係数の推定量を構成することができる。またこれをクロンバックの α 係数という。

1.1.3 マクドナルドの Ω 係数

クロンバックの α 係数はタウ等価測定を仮定して導出した。しかし、実際のテストにおいては t に対して等価であるとは言い難い場面が多い。そこで、同族測定の仮定だけで信頼性を推定してみる。最初の仮定を確認すると

1. $E[e_i] = 0$
2. $E[t_i e_i] = 0$ (i=j のときを含む)
3. $E[e_i e_j] = 0$ ($i \neq j$ の場合のみ)

これより

$$\sigma_e{}^2 = \sum_{i=1}^{n} \sigma_{ei}{}^2 \tag{13}$$

を得る。式 (13) を式 (4) に代入した式

$$\Omega = 1 - \frac{\sum_{i=1}^{n} {\sigma_{ei}}^2}{{\sigma_x}^2} \tag{14}$$