

スタナインを試験科目間の得点調整に用いる方法*

石岡 恒憲¹, 水田 正弘²

¹ 大学入試センター, ² 統計数理研究所

1. はじめに

令和3年度大学入学共通テストより, 成績提供希望者に受験科目の素点に加えて段階表示が提供されるようになった。例年の「大学入学共通テスト受験案内」には, その段階表示の説明があるが, 以下のように説明されている。

(3) 段階表示の方法

① 「科目別得点」を, 「スタナイン (Stanine)」*という方式を用いて, 9 段階に換算します。

* 「スタナイン (Stanine)」とは, 分位点による区分法の一つであり, 受験者を得点順におおよそ 4, 7, 12, 17, 20, 17, 12, 7, 4%の群に分割し, 科目別得点を得点の低い方から順に 1 から 9 の 9 段階に換算する方式です。

スタナインは Flanagan(1948)で紹介されており, これは 316 ページからなる大部な報告書であるが 67 ページ付近に上記の説明がある。結果として得られる分布の平均は 5, 標準偏差は約 2 で, 両端が開いている 1 と 9 の区間を除いて, 区間幅は標準偏差の半分である。Angoff (1984)によれば, スタナイン尺度は 1 桁, すなわち IBM カードの 1 列以上を必要としないため, データ処理機器がまだ開発の初期段階にあり, 1 枚のカードが 80 列でしか個人のデータを表現できない第二次世界大戦中に特に役立ったとしている。ただスタナインは当時も今も, データの効果的な縮約方法として有効である。

また令和7年度共通テストにおける得点調整については以下のように説明がある。

(2) 得点調整の実施条件

得点調整の対象となった各科目間で, 次のいずれかが生じ, これが試験問題の難易差に基づくものと認められる場合には, 得点調整を行います。

- 20 点以上の平均点差が生じた場合

- 15 点以上の平均点差が生じ, かつ, 段階表示の区分点差 (注1) が 20 点以上生じた場合

(注1) ここでいう区分点差とは, 各科目の成績の段階表示 (スタナイン) の各段階の境目となる, 上から 4%, 11%, 23%, 40%, 60%, 77%, 89%, 96%の分位点 (得点) の差を指します。

得点調整の方法としては, 従前の分位点差縮小法を継承しつつも, 以下の通りにしている。

(3) 得点調整の方法

得点調整は, 「分位点差縮小法」 (注2) という方式を用いて, 次のア, イを満たすように行います。

ア. 各科目の元の点数を下げないこと。 (注3)

イ. 各科目間の平均点の順序を保つこと。 (注4)

また, 得点が加算される科目の受験者と加算されない科目の受験者間での公平性の観点から, 区分点差の全てを調整するのではなく, 調整後も区分点差の最も大きいところが 15 点となるようにします。

(注2) 「分位点差縮小法」とは, 分位点差を一定の比率で縮小する方式です。「分位点差」とは, 得点調整の対象となる科目の受験者数の累積割合 (%) が等しいところの分位点 (得点) の差のことです。なお, (2)で得点調整の実施条件として示されている「区分点差」は, 「分位点差」のうち, 各

* 本稿は独立行政法人大学入試センターの公式文書ではない

科目の成績の段階表示（スタナイン）の各段階の境目となる、上から 4 %、11%、23%、40%、60%、77%、89%、96%の分位点（得点）の差を指しています。

（注 3）各科目の元の点数を下げないように、得点調整に当たっては、区分点差が最大となる 2 科目のうち、区分点が低い方の科目の累積分布を高い方の科目に寄せ、区分点が低い方の科目の得点を加算します（図 1 参照）。

また、区分点差が最大となる 2 科目以外の科目についても、区分点差が同一の比率で縮小されるよう調整します。

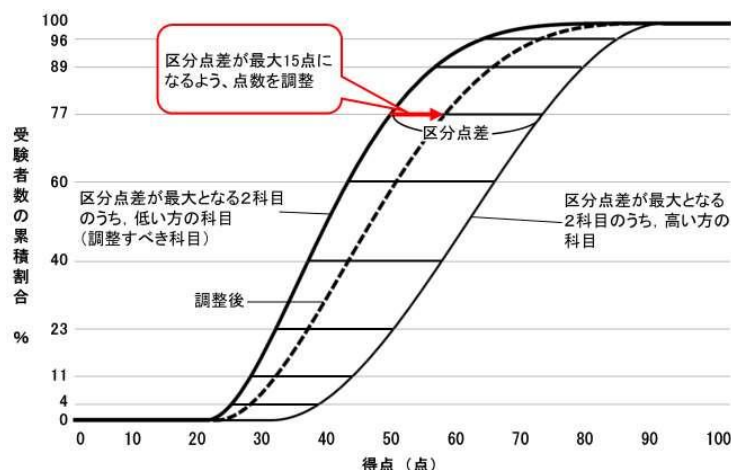


図 1．分位点差縮小法

この方式により、区分点差が最大となる 2 科目のうち、区分点が低い方の科目の得点の累積分布は、図中の点線で描かれた分布に移動することとなります。この点線の分布が調整後の得点の累積分布となり、横軸上の素点から矢印に沿って進み、再び横軸上に戻った点が調整後の得点となります。

（注 4）区分点差の最も大きいところが 15 点となるようにした際に、各科目間の平均点の順序が調整前から入れ替わる場合は、順序が入れ替わらない範囲で調整することとします。その結果、区分点差が 15 点以上に止まる場合もあります。

（転載：大学入試センター, 2024）

ただ上記の説明は、新しい得点調整法の分かりやすさを優先させるために枠組みについてのみ述べたものである。このため、たとえば累積 96%の分位点はスタナイン 8 の上限なのか、スタナイン 9 の下限なのか、はたまた両方があり得るのか（場合によるのか）、スタナインの上下限はいつでも存在するのかについては直ちにはわからない。また得点は整数しかとらない離散分布であるから、分位点も必ずしも一意に定まらない。図 1 に示すスタナイン区分点差の最大になる上位科目（得点の高い方）も複数存在した場合にどの科目を上位科目として選ぶのか、さらに最大区分点差の生じる段階の境目が複数存在した場合（たとえば累積 77%と 89%でともに 23 点差が生じた場合）にどちらの段階の境目を基準にするのか、については省略されている。「各科目間の平均点の順序を保つ」ためのアルゴリズムについても同様に省略されてある。そこで本稿では、スタナイン区分点を得点調整に用いる場合の詳細について説明し、省略された部分について実装の提案を行なう。令和 7 年度共通テストにおける得点調整は、所定の発動条件を満たした場合に本提案が参考にされる予定である。本稿は大学入試センターによる公式文書ではなく、記述に誤りがあるとしたらそれは著者らの責によるものである。

2. スタナインの計算手順

2. 1 経験累積分布関数

経験累積分布 F_n は観測値 i/n （ここで i は観測値の番号）でジャンプする階段関数である。

観測値を $x' = (x_1, x_2, \dots, x_n)$ とするとき、 F_n は t 以下（等号に注意）の観測値を n で割ったもの、すなわち次式で与えられる。ここで Indicator はカッコ内を満たすとき 1、満たさないときに 0 を与える指示関数である。

$$F_n(t) = \# \{x_i \leq t\} / n = 1/n \sum_{i=1}^n \text{Indicator}(x_i \leq t) \quad (1)$$

統計言語の R を使うなら組み込みの標準関数 `ecdf()` を用いることで容易に得ることができる。

2. 2 スタナイン

スタナインの定義に従い最下位の 4% のスタナインスコアを 1、次の 7% のスタナインスコアを 2 のように、以下の表にしたがって定める。等号に注意されたい。

表 1. スタナインスコアを与え方

スタナインスコア	スコアのパーセント割合	経験累積分布 $F_n(t)$
1	最下位の 4 %	$0\% \leq F_n(t) \leq 4\%$
2	次の下位 7 %	$4\% < F_n(t) \leq 11\%$
3	次の下位 12 %	$11\% < F_n(t) \leq 23\%$
4	次の下位 17 %	$23\% < F_n(t) \leq 40\%$
5	中位 20 %	$40\% < F_n(t) \leq 60\%$
6	次の上位 17 %	$60\% < F_n(t) \leq 77\%$
7	次の上位 12 %	$77\% < F_n(t) \leq 89\%$
8	次の上位 7 %	$89\% < F_n(t) \leq 96\%$
9	最上位の 4 %	$96\% < F_n(t) \leq 100\%$

ここで注目して欲しいのは、 t の値を定めれば、仮に t の値を取得する人がいなくても (1) 式により $F_n(t)$ を計算することができ、表 1 により スタナインスコアを計算することができる、ことである。

2. 3 スタナイン換算表の計算例

素点をスタナインスコアに変換する方法を説明する。手順はきわめてシンプルである。

【ステップ 1】0 点から満点に対して該当する累積経験分布を計算する (2.1 節)

【ステップ 2】累積経験分布に対応するスタナインスコアを出力する (2.2 節)

【ステップ 3】これより素点に対する換算表を出力する。表 2 はその出力例である。

表 2: スタナイン換算表の一例

素点	0-30	31-39	40-51	52-63	64-75	76-83	84-90	91-96	97-100
スタナイン	1	2	3	4	5	6	7	8	9

2. 4 補遺

1:100 (1 点から 100 点まで 1 個ずつの 100 個のデータ) に対してスタナインスコアの頻度は

スタナイン	1	2	3	4	5	6	7	8	9
頻度	4	7	12	17	20	17	12	7	4

であり、表 1 にある「スコアのパーセント割合」とびたりと一致する。また 0:100 (0 点から 100 点まで 1 個ずつの 101 個のデータ) に対しては

スタナイン	1	2	3	4	5	6	7	8	9
頻度	4	7	12	17	20	17	12	7	<u>5</u>

となり、スタナイン 9 の頻度のみが 1 つ増える。また、以下の事項に注意する。

- (1) スタナインの算出において正規化スコアを求めることは必要ではない。スタナインの定義がそうだからである。
- (2) スタナインスコアの 9 は必ず出るが、1 は出ない場合がある。また人数が少ない、あるいは特定の得点の頻度が極端に多いと抜け番となるスタナインスコアはある。
- (3) たとえば満点取得者が 20% を占める場合、満点では $F_n(t) = 1$ となりスタナインスコアの 9 の条件を満たすので、スタナインスコアは 9 となる。この場合、スタナインスコア 8 は抜け番となる。満点取得者の 20% が 11% (=4%+7%) を超えているためである。
- (4) 一人しか受けないときは、その人のスタナインスコアは 9 になる。二人しか受けないときは、点数の高い方のスタナインスコアは 9 になり、低い方のスタナインスコアは 5 となる。
- (5) 上記のことから分かるように 4%, 11%, 23%, 40%, 60%, 77%, 89%, 96% の分位点は（存在しても）スタナイン区分点の上下限になるとは限らない。スタナイン区分点の上下限は 2.3 節の述べた手順にしたがって計算された表 2 のような換算表によって求まる。またスタナインスコアは抜け番が生じることがあり、そのときは当然ながら、その上下限のいずれも存在しない。

3. 得点調整

3. 1 分位点差縮小法

令和 7 年度大学入学共通テストにおける得点調整法では、スタナイン分位 (4, 11, 23, 40, 60, 77, 89, 96%) における最大分位点差が 20 点以上の場合に 15 点に縮めるわけだが、そのためには経験累積分布における分位点を計算する必要がある。累積分布関数 (Cdf: cumulative distribution function) の逆関数は、表現関数 (representing function), あるいは分位関数 (quantile function) といわれるが、その値は必ずしも一意に定まらない。図 2 は 96% 点の求め方を示した一例である。経験累積分布は右連続の階段関数であるため、図中の黒丸を含み、白丸を含まない。左側の図は累積分布 (図中 Cdf と表記) のジャンプが 96% をまたぐ場合である。この場合は黒丸の位置の x 軸座標値が 96% 分位点となる。一方、右側の図はジャンプの頂上が 96% にピッタリと一致する場合である。サンプルサイズが十分に大きい場合には 25 分の 1 の確率で累積 96% になるデータは存在し、かつそのデータが偶々ジャンプの頂上になった場合には分位点は一意には定まらない。統計言語 R の標準ライブラリには分位関数 `quantile()` が用意されているが、離散分布に対しては Type 1 から Type 3 までのオプションが用意されている。Type 1 は順序統計量 $x[j]$ ($j=1, \dots, n$) が 96% に一致したとき (あるいは初めて 96% を超えたときに) にその $x[j]$ の値を返すもので、Type 2 は $x[j]$ が 96% に一致したとき $x[j]$ と $x[j+1]$ の平均を取るもの、Type 3 は $x[j]$ が 96% に一致したとき j の偶奇によって $x[j]$ か $x[j+1]$ のいずれかを選ぶ (偶数が選ばれる) としている。特に指定しなければ Type 1 が選ばれる。R の `help(quantile)` によると Type 2 は SAS のデフォルトで Type 3 は 2010 年までの SAS のデフォルトだとしている。我々は R のデフォルトである Type 1 を採用する。もっとも実際には分位点差が最大になるのは累積 77% であることが多いが、累積 77% にピッタリと一致するデータが存在する確率はサンプルサイズが十分に大きい場合 100 分の 1 であり、かつそのデータが偶々ジャンプの頂上に一致する確率はさらに低くなり、このタイプの違いが結果に及ぼす影響は微小といっていいたいだろう。

■ $0 \leq p \leq 1$ に対し $x \leftarrow \text{Cdf}(p)$

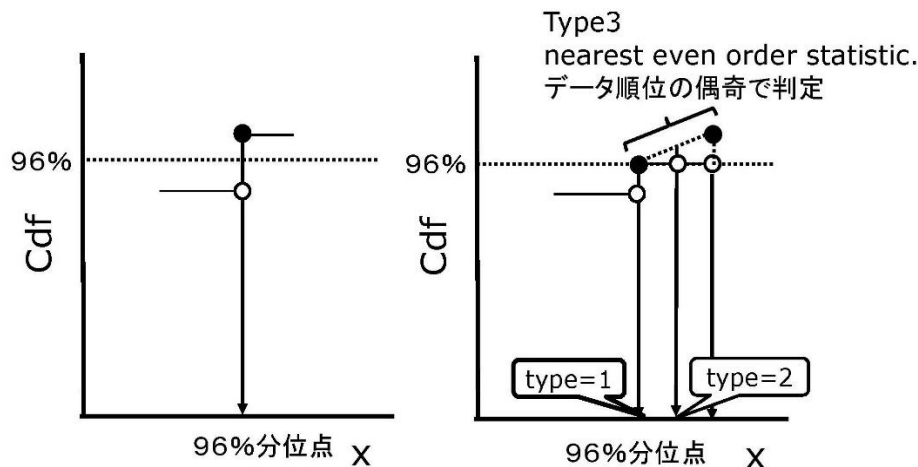


図 2. 96%分位点の求め方

3. 2 平均点の順序を保ちつつ最大分位点差を縮める

今回の得点調整は、スタナイン分位点差だけでなく平均点差も同時に重要視している。

$F_A \Rightarrow F_B$ すなわち「A: 対象分布」を「B: 目標分布」へ寄せるに際し、過去 20 年の共通試験データを用いた数値シミュレーションの結果、採用したポイントは以下の 3 つである。

ポイント 1. 「B: 目標分布」は分位点差最大の上位科目とする（平均値最大分布としない）

ポイント 2. 「A: 対象分布の平均値」 $\geq$ 「B: 目標分布の平均値」なら A はそのままとする

ポイント 3. A は平均値の大きい順に実施（適用）する。調整平均点が前の調整平均点を越えないように加点を調整する

「B: 目標分布」については、複数の分布の組み合わせである最右分布（同じ Cdf に対して得点の大きい科目の組みあわせ分布）も検討したが、特に階段関数の重なり部分が離散値となるため、最右分布をどのように決めるかの判定ロジックがあまりに煩雑になり、プログラムの信頼性を損ねる危険があるため採用を見送った。

上記 3 つのポイントを踏まえ、疑似コードで示したアルゴリズムを図 3 に示す。

```

B を決める（スタナイン区分点最大の分布が複数なら平均値がより大きい方）
w ← 15/スタナイン区分点における最大分位点差
for 科目 #A を平均値の大きい順に実施
  s ← 0:100 # 得点
  If  $\mu_A \geq \mu_B$  なら何もしない； 調整点 ← s とし終了。
  else
     $F_A(s)$  を計算 # ecdf()
     $X_B \leftarrow F_B^{-1}(F_A(s))$  # 対応する B の分位点 quantile(type=1)
    diff ←  $\max(X_B - s, 0)$  # 分位点差
     $\alpha \leftarrow 1$  # 平均点の順番保持のためのパラメータ
    Repeat
      ww ← Ceiling( $w * \text{diff} * \alpha$ ) # つけるべき差
      調整点 ←  $\max(X_B - ww, s)$  # 素点を減らさない
      If 調整点による A の平均がこれまでの科目の調整平均点を越えたら
         $\alpha \leftarrow \alpha + 0.01$  # 加点を少なくする
      else 終了
    endfor 科目
  endfor 科目

```

図 3. $F_A \Rightarrow F_B$ # (A: 対象分布) を (B: 目標分布) へ寄せるアルゴリズム

また「B：目標分布」は以下の方針により決定する。

(1) 分位点差最大となるスタナイン分位（4%, ..., 96%）が複数存在する場合は、累積値のより大きな方を採用し、そこにおける上位科目分布を「B：目標分布」とする。たとえば累積 77%と累積 89%とでともに分位点差 23 点の差がつき、これが最大分位点差であったなら、累積 89%における上位科目を「B：目標分布」とする。これは、より高得点者における不都合を解消すべきであるという判断による。

(2) 上の手順において、累積 89%における上位科目が複数あったときは、より平均点の大きな科目の分布を「B：目標分布」とする。これは平均点もまた重要であるという判断による。

4. FAQ

得点調整、およびスタナインに関して予想されうる質問とその答えを示す。理解の助けになれば幸いである。

Q1： 目標分布あるいは対象分布の取得可能な得点が限定されていて、これがギザギザの階段分布であつても良好に機能するか？

A1： 良好に機能する。分位点差は（取得できない得点を含め）素点が違えば加減幅が違う。滑らかな加減幅になる。

Q2： 96%分位点はスタナイン 9 の下限なのか？それともスタナイン 8 の上限なのか？

A2： 通常はスタナイン 9 の下限になる。ただし標本サイズが 25 の倍数ならぴったり 96%に対応するデータが存在し、かつ 96%が階段関数のジャンプの到達点ならスタナイン 8 の上限以下になる。素点が決まればその経験累積分布値（cdf）が定まり、それが定まればスタナインスコアが定まる。スタナインの上下限は自ずとそこから定まる。これによりスタナインスコアの抜けにも対応できる。「スタナインの上下限≠スタナイン区分点」であることに注意されたい。

Q3： 最大分位点差 20 点以上を 15 点に縮小しても、別の箇所で 15 点以上残ることを許容しているが、それはどのようなときか？

A3： 得点調整による得点変換後においても平均点順序を保持するという制約のため、上記のことは起こり得る。この事象が起きるためには以下の 2 つの条件のいずれもが必要である。①上位科目（目標分布となる科目）と平均値最大科目が異なる；このとき上位科目と平均値最大科目（の累積分布）は（必ず）クロスする。② 平均値最大科目と下位科目で、詰めるべきスタナイン分位とは異なった箇所で通常 20 点以上の差がある。得点調整では上下位科目間で詰めるが、平均値最大科目の得点は維持される（素点を下げない）ため、20 点以上の差を 15 点以下に詰め切ることができない場合がある。ただし①②の現象はそれぞれが起こりにくく、更に同時に起きることは極めて稀である。過去の例では生じない。もちろん、上位科目と平均値最大科目とで 15 点を越える差が生じる場合も（これら 2 つの科目は素点を下げないから）15 点以下に詰めることはできない。ただこの事象が現実には起きることは想定しにくいといつてよいだろう。

謝辞

仕様策定の議論に貴重な意見をくださった前川眞一氏、荘島宏二郎氏、橋本貴充氏、大津起夫氏、ならびに投稿にあたり見識あるコメントをくださった山地弘起氏（いずれも大学入試センター）に心より感謝いたします。

引用文献

Angoff, W.H. (1984) Scales, Norms, and Equivalent Scores. Princeton, NJ: Educational Testing Service. (Reprint of chapter in R. L. Thorndike (Ed.) (1971) Educational measurement (2nd ed.). Washington, DC: American Council on Education.

<https://www.ets.org/Media/Research/pdf/Angoff.Scales.Norms.Equiv.Scores.pdf>

大学入試センター(2024).令和 7 年度大学入学者選抜に係る大学入学共通テストの得点調整の実施条件・方法について, <https://www.dnc.ac.jp/albums/abm.php?d=487&f=abm00003584.pdf>

Flanagan, J.C. (Ed.) (1948) The aviation psychology program in the Army Air Forces. (Report No. 1) Washington: Government Printing Office. <https://apps.dtic.mil/sti/tr/pdf/AD0655590.pdf>