

# 教育評価

岸 学

(東京学芸大学名誉教授・教育インキュベーションセンター共同研究員)

manabu@u-gakugei.ac.jp 042-329-7922

<https://kishilab.sakura.ne.jp/wb/>



枠組みと進め方

なぜ教育評価の知識と理解が必要なのか？

教育評価のしくみを5W1Hでとらえる

ルーブリック活用評価とポートフォリオ評価

評価データ集計と分析のための基礎知識

測定の信頼性と妥当性の考え方を知る

評価項目・テスト項目を作るときの注意点

ルーブリックを作ってみよう

# 教育評価をとらえる枠組み(5W1H)とは？

Manabu  
KISHI



# ループリック活用による方法の例 (reportの評価)

Manabu  
KISHI

	1:許容で きる	2:無難な 出来	3:期待以 上	4:すばら しい
読みやす さ	専門用語 多い	文章は大 体良い	文章が明 瞭	最適な文 章
適切さ	論点が妥 当でない	論点の半 分は同意	論点が明 確	論点すべて 興味深い
有用性	利用できる 情報なし	いくつかの 情報は有 益	情報の殆ど が有益	すべての内 容が有益

# ルーブリックを用いた 評価のためのヒント

Manabu  
KISHI

山口陽弘 (2013). 教育評価におけるルーブリック作成のためのいくつかのヒントの提案  
群馬大学教育学部紀要, 人文・社会科学編, 62, 157-168. より

## ルーブリックの定義

- ・ルーブリック(rubric)とは、いくつかの段階に分けて教育の達成度の目安を記述して、学習者の達成度を判断する基準を示したものである。

## 作成の手順

- Step1** 事前に予想される学習者(児童・生徒など)のさまざまな振る舞い方(問題、解決の方法、行動)を可能な限りリストアップする。 ※正しいものだけではない点に注意
- Step2** 可能であれば複数の教師(できれば3人くらい)でリストアップされた解答・行動パターンや振る舞い方を段階に分けて採点する。 ※合議が望ましい
- Step3** この作業をしながら、すべての段階を埋めていく。

## 注意点

- ・無理に段階数をそろえようとするのではなく、妥当な段階数を検討する
- ・尺度の段階数を多くすることがルーブリック作成の目的ではなく、教師が確実にみとれる行動指標を見出すこと

# 分析法の選び方

推測統計

関係

差

名義尺度  
同士

- ・ $\chi^2$ 検定
- ・オッズ比
- ・連関係数 ( $\phi$  係数)
- ・ユールの関連係数
- など

順序尺度  
同士

- ・名義尺度で使えるもの
- ・順位相関係数 (スピアマン・ケンドール)
- など

間隔・比尺  
度 同士

- ・順序尺度で使えるもの
- ・ピアソンの積率相関係数
- ・相関比
- ・回帰分析
- など

比率の差

- ・ $\chi^2$ 検定
- ・CR検定
- など

平均値の差

- ・ $t$ 検定 (2つの平均値)
- ・分散分析 (2つ以上の平均値)
- など

中央値の差

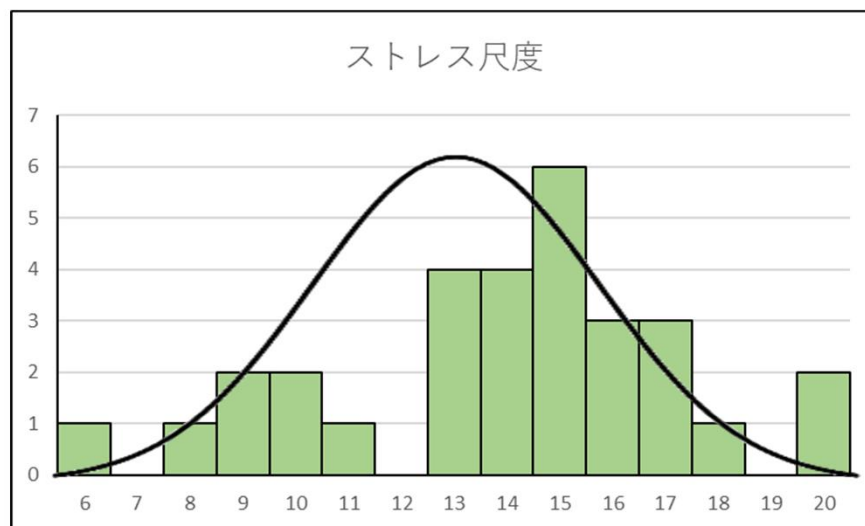
- ・マン・ホイットニーの $U$ 検定
- ・ウイルコクソンの符号付き順位検定
- ・クラスカル・ウォリスの $H$ 検定
- ・フリードマンの検定

標準偏差の  
差

- ・ $F$ 検定
- ・Leveneの検定

# Excelの基本統計量の出力内容

ストレス尺度		正規分布
平均	13.93333333	13
標準誤差	0.61388781	0.500574
中央値 (メジアン)	14.5	13
最頻値 (モード)	15	13
標準偏差	3.36240199	2.741759
分散	11.3057471	7.517241
尖度	0.06480793	1.528509
歪度	-0.4513622	0
範囲	14	14
最小	6	6
最大	20	20
合計	418	390
データの個数	30	30



Excelのメニュー→データ→データ分析→基本統計量。下のヒストグラムは、データ分析→ヒストグラム

平均⇒算術平均。ヒストグラムでは重心。

標準誤差⇒標本分布の標準偏差 標準偏差/ $\sqrt{N}$ 。平均値の標本分布の変動の程度。

中央値(メジアン)⇒データの真ん中。ヒストグラムでは面積を2等分する値。

最頻値(モード)⇒最も柱が高いところの値

標準偏差(SD)⇒データの散らばり具合。平均 $\pm 1$ SDの中に、60~70%のデータが入る。

分散⇒標準偏差の2乗  $3.36^2=11.31$

尖度⇒ピークが尖った分布かどうかの程度。正規分布は0、それより尖っていれば正、平らなら負の値

歪度⇒分布全体の歪みの様子。左右対称なら0、左に偏れば正、右に偏れば負の値。

範囲⇒最大と最小の差

最小⇒最小値

最大⇒最大値

合計⇒raw dataの総和

# 良い評価を行うための検討方法

## 学生指導

### 評価項目

栗山先生による評価結果

		評価項目				
		項目1	項目2	項目3	項目4	合計
学 生	Aさん	3	2	3	2	10
	Bさん	2	2	1	1	6
	Cさん	3	2	2	3	10
	Dさん	1	1	2	1	5
	Eさん	3	3	3	2	11

$\alpha$  係数 = 0.840

① 評価項目が適切かどうか⇒信頼性と妥当性の高さ

- 1) 評価者（先生）が代わったとしても安定した結果になること
- 2) 構成する項目が、似たような（等質の）内容をチェックしていること ⇒ IT 相関 (item-total 相関) による検討。合計と個々の項目との相関の程度を見してみる
- 3) このような条件を満たすような項目群を選んでいくこと⇒項目分析

### 指導者

指導者の評価結果

		評価項目 No.1			
		栗山	源田	森	山川
学 生	Aさん	3	2	1	2
	Bさん	2	1	1	2
	Cさん	3	2	2	2
	Dさん	1	2	3	2
	Eさん	3	1	1	2

① 指導者の間で観察や到達度の視点が共通すること

② どの先生の評定値が他の先生と異なるかを検討すること

- 1) 共通しない場合は研修や指導者間カンファレンスなどで共通化を図ること
- 2) 学生の具体的な行動や様子を取り上げて視点を揃える練習をすること
- 3) 学生の学習や習熟や発達による変化の要素についてのとらえ方も共通化すること



## ○再テスト法

- 同一テストの繰り返し実施。再現性・安定性
- 信頼性係数(相関係数)

## × 平行テスト法

## × 折半法

- 信頼性係数(スピアマン・ブラウンの公式)

## ◎内部一貫法(内的整合法)

- テスト項目の間関係の強さ。等質性・識別性
- 信頼性係数( $\alpha$ 係数、KR-20の式 など)

## ◎一般化可能性理論による方法

- 分散成分の推定による

# 内的整合法(内部一貫法)の $\alpha$ 係数の 計算法と考え方

$$\alpha \text{ 係数} = \frac{\text{問題数}}{\text{問題数} - 1} \times \left( 1 - \frac{\text{問題項目の分散の合計}}{\text{個人得点の分散}} \right)$$

番号	問1	問2	問3	問4	問5	個人得点
1	2	2	2	2	1	9
2	2	2	1	2	2	9
3	2	2	1	2	2	9
4	2	2	2	2	1	9
5	2	1	0	2	2	7
6	1	0	0	1	0	2
7	1	0	1	1	0	3
8	2	0	2	2	1	7
9	2	1	2	2	1	8
10	2	1	2	2	2	9
分散	0.178	0.767	0.678	0.178	0.622	6.844

①分散の求め方は、VARという関数を用いる。標準偏差(SD)を求めてそれを2乗しても同じである。

② $\alpha$ 係数を求める手順は次の通り。

問題数 = 5

問題項目の分散の合計 =  $0.178+0.767+0.678+0.178+0.622=2.423$

個人得点の分散 = 6.844

$\alpha$ 係数 =  $\frac{\text{問題数}}{\text{問題数}-1} \times (1 - \frac{\text{問題項目の分散の合計}}{\text{個人得点の分散}})$

=  $5/(5-1) \times (1 - 2.423/6.844)$

= 0.807

	Q1	Q2	Q3	Q4	合計
1	2	2	1	1	6
2	2	3	2	3	10
3	1	3	1	1	6
4	3	3	2	3	11
5	1	1	3	3	8
6	2	3	2	3	10
7	2	1	3	2	8
8	3	2	3	3	11
分散	0.571	0.786	0.696	0.839	4.214
				分散の合計	2.893
				$\alpha =$	<u>0.418</u>

	Q1	Q2	Q3	Q4	合計
1	1	1	2	2	6
2	2	2	3	3	10
3	1	1	2	2	6
4	3	2	3	3	11
5	2	1	3	2	8
6	2	3	2	3	10
7	2	1	3	2	8
8	3	2	3	3	11
分散	0.571	0.554	0.268	0.286	4.214
				分散の合計	1.679
				$\alpha =$	<u>0.802</u>

教えた形式と設問の形式とを一致させる

何を学習していたら正答できるかのイメージ

- 誤答だった人が、何を勉強すればよいかの方針が立てられるか？

正答の基準を明確に。設問の表現に注意

- 雪が溶けたら何になるか？⇒
- 次の方程式を解け⇒
- イデオロギーとは何か？(須賀,1980)⇒
- 神奈川県のある県庁所在地は？⇒
- 高性能気圧計を使って高層ビルの高さを知る方法を示せ(池田,1992)⇒

選択枝の数をなるべく揃える

- チャンスレベルの均一化

# テストで出題される問題のタイプ分類

