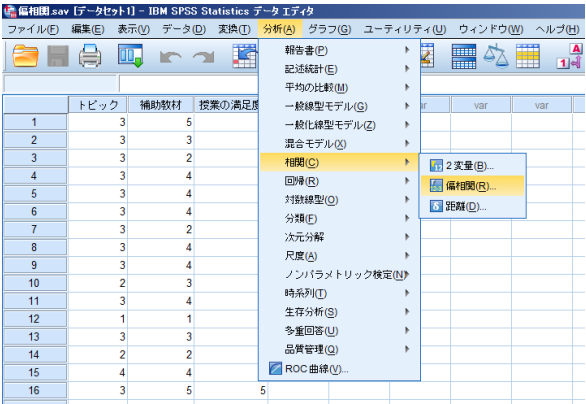


7-2-4 偏相関係数の算出



・ [分析(A)] → [相関(C)] → [偏相関(R)]



- ・ 変数：影響を考慮する
- ・ 制御変数：影響を排除する（＝考慮しない）

相関係数				
制御変数			授業の満足度	補助教材
トピック	授業の満足度	相関	1.000	.645
		有意確率 (両側)	.	.000
		df	0	27
補助教材	授業の満足度	相関	.645	1.000
		有意確率 (両側)	.000	.
		df	27	0

表 7.1 相関係数の解釈	
r の値	相関の強さの判定
.00 ～ ± .20	ほとんど相関がない
± .20 ～ ± .40	弱い相関がある
± .40 ～ ± .70	中程度の相関がある
± .70 ～ ± 1.00	強い相関がある

(田中・山際 [1992] p.188 にもとづく)

・ トピックの影響を除いた、変数（授業の満足度・補助教材）の偏相関係数 $r = .645$

Cf. 第 8 章：回帰分析では、ピアソンの相関係数、偏相関係数、部分相関係数の全てを算出・利用

係数 ^a										
モデル	標準化されていない係数		標準化係数	t値	有意確率	相関			共線性の統計量	
	B	標準誤差	ベータ			ゼロ次	偏	部分	許容度	VIF
1 (定数)	-2.115	4.424		-4.78	.634					
ブレイクメントテスト	-.031	.042	-.035	-.747	.457	-.302	-.069	-.032	.850	1.176
前期試験	.183	.047	.175	3.857	.000	.370	.338	.167	.917	1.090
中間模試	.699	.053	.702	13.101	.000	.851	.774	.568	.654	1.529
後期試験	.215	.050	.210	4.341	.000	.531	.375	.188	.805	1.242

a. 従属変数: 年度末模試

7-3. 相関係数と信頼性

■相関と内的一貫性・評価者信頼性

表 1.4 代表的な信頼性の推定方法

信頼性の推定方法	具体的な方法
再テスト法 (test-retest reliability)	同じテストを同一の受験者に2回実施し、その点数の一致度を相関係数で示す。2回のテスト実施の間に被験者の能力が変わらないこと、テストの内容が被験者が忘れていることが条件。2～3週間の間隔を空けて行う。
同等フォーム法 (equivalent-forms reliability)	ほぼ同一の難易度、標準偏差をもつ2つのテスト（平行テスト）を同一の被験者に実施し、その点数の一致度を相関係数で示す。
内的一貫性 (internal consistency) ・折半法 (split-half method) ・アルファ係数 (α coefficient; Cronbach's α) ・キューダー・リチャードソン 20 (KR-20) ・キューダー・リチャードソン 21 (KR-21)	- たとえば、項目を奇数番号と偶数番号に分けて、この両者の点数の一致度（相関係数）を計算し、スピアマン・ブラウン公式で修正して算出する（第7章）。 - 項目の標準偏差とテストの総得点の標準偏差により算出する（第7章）。 - 項目数、項目分散の合計、テスト全体の分散により算出する。2値データのみ適用。 - テスト項目数、平均値、標準偏差により算出する。2値データのみ適用。
評価者信頼性 (rater reliability) ・評価者間信頼性 (inter-rater reliability) ・評価者内信頼性 (intra-rater reliability)	- 評価者が2名の場合は、相関係数やカッパ係数（第7章）を用いる。評価者が3名以上の場合は、アルファ係数を使用する。また一般化可能性理論や項目応答理論を用いて算出することも可能。 1人の評価者が、同一のパフォーマンスをある一定の期間において複数回採点した結果が一致する度合い。算出方法は、評価者間信頼性に準ずる。

注. Brown [2005]; Hatch & Lazaraton [1991]; 靜 ほか [2002] をもとに作成。

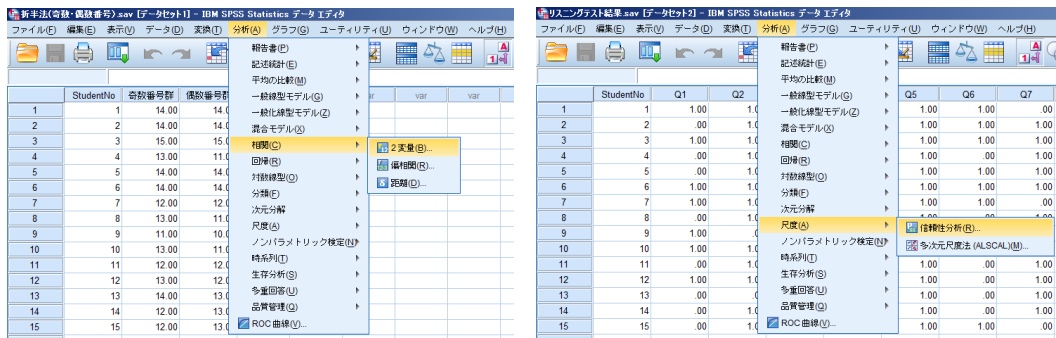
- ・内的一貫性：各受験者のテスト項目の得点は、どの程度一貫しているか？

⇒折半法、アルファ係数

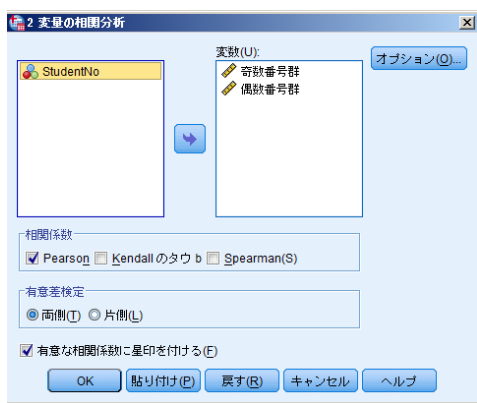
- ・評価者信頼性：複数の評価者による採点は、どの程度一致しているか？

⇒カッパ係数

7-3-1 折半法による信頼性係数の算出



[分析(A)] → [相関(C)] → [2 変量(B)] OR [分析(A)] → [尺度(A)] → [信頼性分析(R)]



奇数番号群と偶数番号群

OR

テストの前半と後半

相関係数			
		奇数番号群	偶数番号群
奇数番号群	Pearson の相関係数	1	.778**
	有意確率 (両側)		.000
	N	36	36
偶数番号群	Pearson の相関係数	.778**	1
	有意確率 (両側)	.000	
	N	36	36

信頼性統計量		
Cronbach のアルファ	部分 1 値 (半割分析)	.685
	項目の数	15 ^a
部分 2 値 (半割分析)	項目の数	.759
	項目の合計数	15 ^b
項目の合計数		30
フォーム間の相関		.690
Spearman-Brown の係数	等しい長さ	.816
	等しくない長さ	.816
Guttman の折半法信頼係数		.815

a. 項目: Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q8, Q9, Q10, Q11, Q12, Q13, Q14, Q15.
b. 項目: Q16, Q17, Q18, Q19, Q20, Q21, Q22, Q23, Q24, Q25, Q26, Q27, Q28, Q29, Q30.

・ピアソンの相関係数 r は 1/2 のテスト項目間の数値なので、スピアマン・ブラウンの公式によってテスト全体の信頼性係数 ρ を算出する必要あり

⇒テスト項目を 2 群に折半した場合の式は、
$$\rho = \frac{2 \times r}{1 + r(2 - 1)} = \frac{2 \times .778}{1 + .778} = \frac{1.556}{1.778} = .875$$

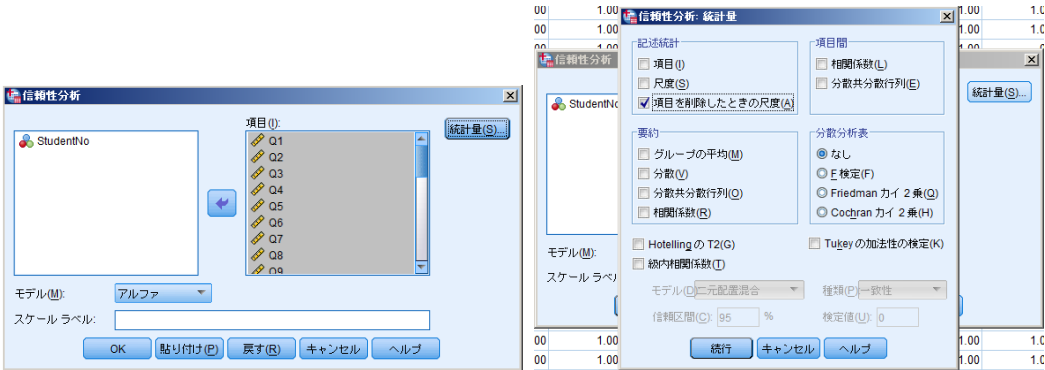
Cf. テスト項目を 3 群に折半した場合は、2 の箇所を 3 に置き換える

・共に高い信頼性計数（.875 と.816）が得られたが、折半の方法によって異なる数値（問題点）

コメント [G1]: 教科書では公式の分母部分に誤植？

7-3-2 アルファ係数の算出

- ・折半法の問題点を解決するため、全折半方法による信頼性の平均値を算出



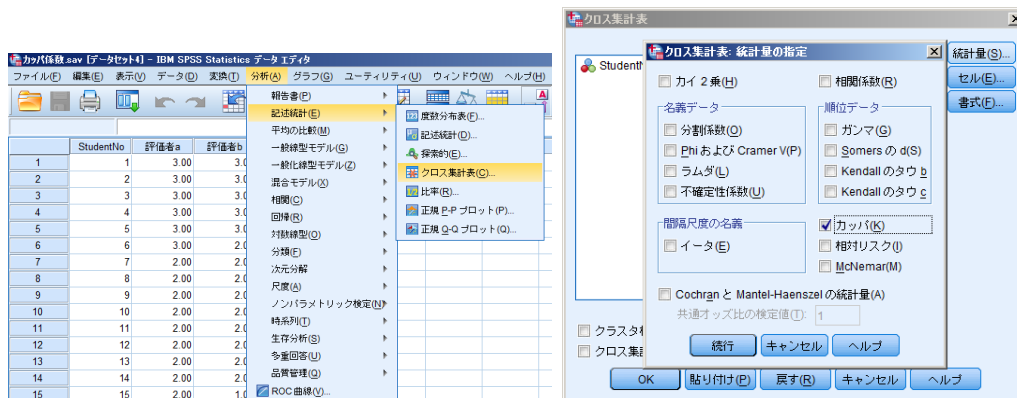
- ・[分析(A)] → [尺度(A)] → [信頼性分析(R)]
- ・[統計量(S)] → [記述統計] → [項目を削除したときの尺度(A)]

信頼性統計量		項目合計統計量				
		項目が削除された場合の尺度の平均値	項目が削除された場合の尺度の分散	修正済み項目合計相関	項目が削除された場合の Cronbach のアルファ	
Cronbach のアルファ	項目の数	Q1	21.8056	27.990	-.365	.857
		Q2	21.6667	24.514	.427	.829
		Q3	21.9167	23.279	.584	.822
		Q4	21.6111	25.502	.211	.835
		Q5	21.7222	23.521	.624	.822
		Q6	22.0278	25.456	.130	.840

- ・ α (Cronbach のアルファ) の値次第では、項目の削除を検討

⇒但し、仮に項目の削除によって α を上げられるとしても、削除は必ずしも良い手段とは限らないため要検討（内的一貫性 vs. 内容的妥当性・構成概念妥当性）

7-3-3 カッパ係数の算出



- [分析(A)] → [記述統計(E)] → [クロス集計表(C)]
- [統計量(S)] → [カップ(K)]

表 7.4 カップ係数(k)の判定基準

値	判定 (一致度)
.00- .20	slight (低い)
.21- .40	fair (やや低い)
.41- .60	moderate (中程度)
.61- .80	substantial (かなり高い)
.81-1.00	almost perfect (ほぼ一致)

注. Landis & Koch [1977] p.165 をもとに作成

対称性による類似度				
	値	漸近標準誤差 ^a	近似 T 値 ^b	近似有意確率
一致の測定方法 カップ	.684	.139	4.396	.000
有効なケースの数	20			

a. 帰無仮説を仮定しません。
b. 帰無仮説を仮定して漸近標準誤差を使用します。

- 2 人の評価者の**評価基準が同一**である必要
- 全データを 2 人の評価者で分析できない場合、データ全体の何%なら容認されるのか？
⇒一般的には **25%** とされているが、**以下の記述**も併せて参照

コメント [G2]: 強調と下線は筆者 (石井) による

*Loewen, S., & Philp, J. (2006). Recasts in the adult L2 classroom: Characteristics, explicitness, and effectiveness. *Modern Language Journal*, 90, 536-556. では以下のような記述がありました。

“Lombard, Snyder-Duch, and Campanella Bracken (2005) recommended **10% as a minimal, acceptable subsample** to use for interrater reliability coding.” (Mizumoto, n.d.)

This sample must also be selected using a random or other justifiable procedure. **The appropriate size of the sample depends on many factors but it should not be less than 50 units or 10% of the full sample, and it rarely will need to be greater than 300 units;** larger reliability samples are required when the full sample is large and/or when the expected reliability level is low (Neuendorf, 2002; see Lacy and Riffe, 1996 for a discussion). (Lombard, Snyder-Duch, & Bracken, 2010)

コメント [G3]: many のタイプミス？

- ・ 3人以上の評価者の場合はアルファ係数を用いる
- ・ 信頼性は集団に依存する（＝集団毎に異なる可能性がある）という点は、常に留意すべき

7-3-4 相関係数の希薄化

- ・ 相関係数の希薄化：観測変数の誤差による相関係数の低下

- ・ 真の相関係数を算出する希薄化の修正公式は、

$$\frac{\text{修正前の相関係数}}{\sqrt{\text{変数xの信頼性係数} \times \text{変数yの信頼性係数}}}$$

⇒修正前の相関係数は下限値、修正後の相関係数は上限値であり、真の相関係数が含まれる範囲を示している

- ・ 希薄化の修正を行った場合、修正前の相関係数・修正後の相関係数だけでなく、計算に使用した信頼性係数の種類・係数の値も併せて報告
- ・ 但し、変数の信頼性係数を算出する必要性・信頼性係数の種類による影響も考慮すると希薄化の修正は稀であり、大前提として信頼性・妥当性が高いデータを集める事が重要

コメント [G4]: 公式を使用した教科書の例は、p. 138 の図 7.9 を参照しながら読むと理解しやすい。尚、「このそれぞれの信頼性係数を算出した場合、... だったとします。」は、この後の考慮すべき点に繋がる

7-3-5 論文への記載（アルファ係数とカッパ係数）

- ・ pp. 153-154 参照

参考文献

Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2010). *Intercoder reliability: Practical resources for assessing and reporting intercoder reliability in content analysis research projects*. Retrieved from <http://matthewlombard.com/reliability/>

Mizumoto, A. (n.d.). *Kappa keisu no keisan* [Calculating Kappa coefficients]. Retrieved from <http://www.mizumoto.com/stats/kappa.htm>