

【英語教育評価論】

第5章 多元配置分散分析 (担当箇所 p102~111)

5-5-4 ◆単純・単純主効果の検定 (シンタックス)

対応なし要因 A : 評価者 (Rater: 1. Training なし, 2. Training ありの 2 水準)
 対応あり要因 P : 面接ストラテジー (Strategy: 1. High, 2. Low の 2 水準)
 対応あり要因 Q : スピーキング能力 (Proficiency: 1. High, 2. Mid, 3. Low の 3 水準)

- (1) Strategy と Proficiency の各水準の組み合わせにおける Rater の単純・単純主効果の検定
 - (2) Rater と Proficiency の各水準の組み合わせにおける Strategy の単純・単純主効果の検定
 - (3) Rater と Strategy の各水準の組み合わせにおける Proficiency の単純・単純主効果の検定
- ⇒ 単純・単純主効果の検定では単純交互作用よりも更に分解をして、2 要因のすべての組み合わせにおける Rater の水準間の比較を行う。

単純・単純主効果の検定

- (1) Strategy と Proficiency の各水準における Rater の単純・単純主効果

=1 変量検定

測定変数名: MEASURE 1

Strategy	Proficiency		平方和	自由度	平均平方	F 値	有意確率	偏イータ 2 乗
1	1	対比	8.450	1	8.450	.221	.644	.012
		誤差	688.500	18	38.250			
	2	対比	9.800	1	9.800	.247	.625	.014
		誤差	713.400	18	39.633			
	3	対比	6301.250	1	6301.250	303.512	.000	.944
		誤差	373.700	18	20.761			
2	1	対比	4470.050	1	4470.050	143.245	.000	.888
		誤差	561.700	18	31.206			
	2	対比	42.050	1	42.050	1.434	.247	.074
		誤差	527.700	18	29.317			
	3	対比	.450	1	.450	.040	.843	.002
		誤差	200.500	18	11.139			

F 値は RaterSkill の多変量効果を検定します。この検定は推定周辺平均間で線型に独立したペアごとの比較に基づいています。

(2) Rater と Proficiency の各水準における Strategy の単純・単純主効果
多変量検定

RaterSkill	Proficiency	値	F 値	仮説自由度	誤差自由度	有意確率	偏イータ 2 乗	
Trainingなし	1	Pillai のトレース	.921	211.267 ^a	1.000	18.000	.000	.921
		Wilks のラムダ	.079	211.267 ^a	1.000	18.000	.000	.921
		Hotelling のトレース	11.737	211.267 ^a	1.000	18.000	.000	.921
		Roy の最大根	11.737	211.267 ^a	1.000	18.000	.000	.921
	2	Pillai のトレース	.947	320.184 ^a	1.000	18.000	.000	.947
		Wilks のラムダ	.053	320.184 ^a	1.000	18.000	.000	.947
		Hotelling のトレース	17.788	320.184 ^a	1.000	18.000	.000	.947
		Roy の最大根	17.788	320.184 ^a	1.000	18.000	.000	.947
	3	Pillai のトレース	.970	584.254 ^a	1.000	18.000	.000	.970
		Wilks のラムダ	.030	584.254 ^a	1.000	18.000	.000	.970
		Hotelling のトレース	32.459	584.254 ^a	1.000	18.000	.000	.970
		Roy の最大根	32.459	584.254 ^a	1.000	18.000	.000	.970
Trainingあり	1	Pillai のトレース	.090	1.791 ^a	1.000	18.000	.198	.090
		Wilks のラムダ	.910	1.791 ^a	1.000	18.000	.198	.090
		Hotelling のトレース	.099	1.791 ^a	1.000	18.000	.198	.090
		Roy の最大根	.099	1.791 ^a	1.000	18.000	.198	.090
	2	Pillai のトレース	.934	255.350 ^a	1.000	18.000	.000	.934
		Wilks のラムダ	.066	255.350 ^a	1.000	18.000	.000	.934
		Hotelling のトレース	14.186	255.350 ^a	1.000	18.000	.000	.934
		Roy の最大根	14.186	255.350 ^a	1.000	18.000	.000	.934
	3	Pillai のトレース	.380	11.022 ^a	1.000	18.000	.004	.380
		Wilks のラムダ	.620	11.022 ^a	1.000	18.000	.004	.380
		Hotelling のトレース	.612	11.022 ^a	1.000	18.000	.004	.380
		Roy の最大根	.612	11.022 ^a	1.000	18.000	.004	.380

F 値はそれぞれ表示された他の効果の各水準の組み合わせ内の Strategy の多変量単純効果を検定します。このような検定は推定周辺平均間で線型に独立したペアごとの比較に基づいています。

a. 正確統計量

(3) Rater と Strategy の各水準における Proficiency の単純・単純主効果の検定[t1]

多変量検定

RaterSkill	Strategy	値	F 値	仮説自由度	誤差自由度	有意確率	偏イータ 2 乗	
Trainingなし	1	Pillai のトレース	.015	.130 ^a	2.000	17.000	.879	.015
		Wilks のラムダ	.985	.130 ^a	2.000	17.000	.879	.015
		Hotelling のトレース	.015	.130 ^a	2.000	17.000	.879	.015
		Roy の最大根	.015	.130 ^a	2.000	17.000	.879	.015
	2	Pillai のトレース	.643	15.298 ^a	2.000	17.000	.000	.643
		Wilks のラムダ	.357	15.298 ^a	2.000	17.000	.000	.643
		Hotelling のトレース	1.800	15.298 ^a	2.000	17.000	.000	.643
		Roy の最大根	1.800	15.298 ^a	2.000	17.000	.000	.643
Trainingあり	1	Pillai のトレース	.942	139.236 ^a	2.000	17.000	.000	.942
		Wilks のラムダ	.058	139.236 ^a	2.000	17.000	.000	.942
		Hotelling のトレース	16.381	139.236 ^a	2.000	17.000	.000	.942
		Roy の最大根	16.381	139.236 ^a	2.000	17.000	.000	.942
	2	Pillai のトレース	.962	214.867 ^a	2.000	17.000	.000	.962
		Wilks のラムダ	.038	214.867 ^a	2.000	17.000	.000	.962
		Hotelling のトレース	25.278	214.867 ^a	2.000	17.000	.000	.962
		Roy の最大根	25.278	214.867 ^a	2.000	17.000	.000	.962

F 値はそれぞれ表示された他の効果の各水準の組み合わせ内の Proficiency の多変量単純効果を検定します。このような検定は推定周辺平均間で線型に独立したペアごとの比較に基づいています。

a. 正確統計量

⇒ 表 5.7 での Proficiency のように 3 水準で結果が有意になっている場合、差がどこにあるのかを明らかにするために多重比較を行う必要がある。

多重比較の結果

(1) Strategy と Proficiency における多重比較

ペアごとの比較

測定変数名:MEASURE 1

Strategy	Proficiency	(I) RaterSkill	(J) RaterSkill	平均値の差 (I-J)	標準誤差	有意確率 ^a	95% 平均差信頼区間 ^a	
							下限	上限
1	1	Trainingなし	Trainingあり	-1.300	2.766	.644	-7.111	4.511
		Trainingあり	Trainingなし	1.300	2.766	.644	-4.511	7.111
	2	Trainingなし	Trainingあり	1.400	2.815	.625	-4.515	7.315
		Trainingあり	Trainingなし	-1.400	2.815	.625	-7.315	4.515
	3	Trainingなし	Trainingあり	35.500 [*]	2.038	.000	31.219	39.781
		Trainingあり	Trainingなし	-35.500 [*]	2.038	.000	-39.781	-31.219
2	1	Trainingなし	Trainingあり	-29.900 [*]	2.498	.000	-35.149	-24.651
		Trainingあり	Trainingなし	29.900 [*]	2.498	.000	24.651	35.149
	2	Trainingなし	Trainingあり	-2.900	2.421	.247	-7.987	2.187
		Trainingあり	Trainingなし	2.900	2.421	.247	-2.187	7.987
	3	Trainingなし	Trainingあり	-.300	1.493	.843	-3.436	2.836
		Trainingあり	Trainingなし	.300	1.493	.843	-2.836	3.436

推定周辺平均に基づいた

a. 多重比較の調整: Bonferroni。

*. 平均の差は .05 水準で有意です。

(2) Rater と Proficiency における多重比較

ペアごとの比較

測定変数名:MEASURE 1

RaterSkill	Proficiency	(I) Strategy	(J) Strategy	平均値の差 (I-J)	標準誤差	有意確率 ^a	95% 平均差信頼区間 ^a	
							下限	上限
Trainingなし	1	1	2	31.500 [*]	2.167	.000	26.947	36.053
		2	1	-31.500 [*]	2.167	.000	-36.053	-26.947
	2	1	2	40.200 [*]	2.247	.000	35.480	44.920
		2	1	-40.200 [*]	2.247	.000	-44.920	-35.480
	3	1	2	41.500 [*]	1.717	.000	37.893	45.107
		2	1	-41.500 [*]	1.717	.000	-45.107	-37.893
Trainingあり	1	1	2	2.900	2.167	.198	-1.653	7.453
		2	1	-2.900	2.167	.198	-7.453	1.653
	2	1	2	35.900 [*]	2.247	.000	31.180	40.620
		2	1	-35.900 [*]	2.247	.000	-40.620	-31.180
	3	1	2	5.700 [*]	1.717	.004	2.093	9.307
		2	1	-5.700 [*]	1.717	.004	-9.307	-2.093

推定周辺平均に基づいた

*. 平均の差は .05 水準で有意です。

a. 多重比較の調整: Bonferroni。

(3) Rater と Strategy における多重比較

ペアごとの比較

測定変数名: MEASURE_1

RaterSkill	Strategy	(I) Proficiency	(J) Proficiency	平均値の差 (I-J)	標準誤差	有意確率 ^a	95% 平均差信頼区間 ^a	
							下限	上限
Trainingなし	1	1	2	-.200	2.929	1.000	-7.930	7.530
			3	1.000	2.631	1.000	-5.943	7.943
		2	1	.200	2.929	1.000	-7.530	7.930
			3	1.200	2.536	1.000	-5.493	7.893
		3	1	-1.000	2.631	1.000	-7.943	5.943
			2	-1.200	2.536	1.000	-7.893	5.493
	2	1	2	8.500 [*]	2.377	.006	2.227	14.773
			3	11.000 [*]	1.944	.000	5.869	16.131
		2	1	-8.500 [*]	2.377	.006	-14.773	-2.227
			3	2.500	2.107	.753	-3.062	8.062
		3	1	-11.000 [*]	1.944	.000	-16.131	-5.869
			2	-2.500	2.107	.753	-8.062	3.062
Trainingあり	1	1	2	2.500	2.929	1.000	-5.230	10.230
			3	37.800 [*]	2.631	.000	30.857	44.743
		2	1	-2.500	2.929	1.000	-10.230	5.230
			3	35.300 [*]	2.536	.000	28.607	41.993
		3	1	-37.800 [*]	2.631	.000	-44.743	-30.857
			2	-35.300 [*]	2.536	.000	-41.993	-28.607
	2	1	2	35.500 [*]	2.377	.000	29.227	41.773
			3	40.600 [*]	1.944	.000	35.469	45.731
		2	1	-35.500 [*]	2.377	.000	-41.773	-29.227
			3	5.100	2.107	.079	-.462	10.662
		3	1	-40.600 [*]	1.944	.000	-45.731	-35.469
			2	-5.100	2.107	.079	-10.662	.462

推定周辺平均に基づいた

a. 多重比較の調整: Bonferroni.

*. 平均の差は .05 水準で有意です。

◆ 2 次の交互作用が有意ではなく、1 次の交互作用が有意な場合
単純主効果の検定を行う。

・ A×P の交互作用が有意な場合

① 要因 A の各水準における要因 P の単純主効果の検定

② 要因 P の各水準における要因 A の単純主効果の検定

→単純主効果が有意であれば必要に応じて多重比較。

* そのほかの組み合わせも同様。

◆ 2 次および 1 次の交互作用が有意でない場合

・ 交互作用が見られなかった場合：要因それぞれの主効果の検定結果を見る。

・ いずれかの主効果が有意である場合：必要に応じて多重比較。

・ 主効果がいずれも有意でない場合：差はないという結論になる。