ME 5824: Homework-5

Shaunak Mehta, Ananth Jonavittula

Q 1

We completed the paired t-test and got the following values:

- **Degrees of Freedom**: 19
- **t-value**: -5.504
- **Significance**: $< .001$

It is generally accepted that results are statistically significant if the p value (significance) < 0.05. Therefore our results are significant. The detailed results from SPSS are attached at the end.

Q 2

Two contrasting questions were asked to the participants to see if they trusted the robot. On performing reliability test on the responses for these two questions, it was found that the scale was reliable with a Cronbach's Alpha value of 0.801. Thus, we can group the results of these questions into a combined score. The mean, standard deviation and standard error are given as:

- **Mean:** 9.10
- **Std. Deviation:** 2.644
- **Standard Error:** 0.8361

Q 3

Effect of task

The statistics from the repeated measures ANOVA for tasks are given below

- **Degrees of Freedom**: 2
- **F**: 128.099
- **Significance**: < 0.001

The task has a significant effect on the results obtained. We can say this because the significance reported for tasks is < 0.001.

Effect of method

The statistics from the repeated measures ANOVA for tasks are given below

- **Degrees of Freedom**: 1
- **F**: 44.927
- **Significance**: < 0.001

The method also has a significant effect on the results obtained. We can say this because the significance reported for tasks is < 0.001.

Effect of Task 2

The statistics from the two tailed t-test for task 2 are given below

- **Degrees of Freedom**: 9

- **t**: -0.607

- **Significance**: 0.559

The methods for task 2 do not have a significant effect on the results obtained. We can say this because the significance reported for tasks is 0.559, which is not < 0.05.

Q 4

To avoid the confounding factors and make sure that the results are unbiased:

- Since we have only 10 participants, we perform a within subjects study to ensure that the the results are not biased by some users (say working with Ours) having more skill over others (working with Baseline).

- To overcome the bias arising due to participants gaining skills from the first method, we flip the order of methods for half of the participants.

- We do not reveal which which method the user is working with (Ours or the Baseline) to the user before or during the user study.

T-Test

Notes

Output Created		26-MAR-2022 16:16:32
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	20
Missing Value Handling	Definition of Missing	User defined missing values are treated as missing.
	Cases Used	Statistics for each analysis are based on the cases with no missing or out-of-range data for any variable in the analysis.
Syntax		T-TEST PAIRS=tasktimewithmeth odA WITH tasktimewithmethodB (PAIRED) /ES DISPLAY(TRUE) STANDARDIZER(SD) /CRITERIA=CI(.9500) /MISSING=ANALYSIS.
Resources	Processor Time	00:00:00.01
	Elapsed Time	00:00:00.00

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	task time with method A	8.12537840	20	1.04489667	.233645999
	task time with method B	9.85870717	20	.760166873	.169978480

Paired Samples Correlations

		N	Correlation	Significance One-Sided p	Two-Sided p
Pair 1	task time with method A & task time with method B	20	-.198	.202	.403

Paired Samples Test

		Paired Differences			95% Confidence ...
		Mean	Std. Deviation	Std. Error Mean	Lower
Pair 1	task time with method A - task time with method B	-1.73332877	1.40849364	.314948752	-2.39252409

Paired Samples Test

		Paired ...			Significance
		95% Confidence Interval of the ...			
		Upper	t	df	One-Sided p
Pair 1	task time with method A - task time with method B	-1.07413346	-5.504	19	<.001

Paired Samples Test

		Significance
		Two-Sided p
Pair 1	task time with method A - task time with method B	<.001

Paired Samples Effect Sizes

			Standardizer[a]	Point Estimate	95% ...
					Lower
Pair 1	task time with method A - task time with method B	Cohen's d	1.40849364	-1.231	-1.807
		Hedges' correction	1.46731789	-1.181	-1.735

Paired Samples Effect Sizes

			95% ...
			Upper
Pair 1	task time with method A - task time with method B	Cohen's d	-.636
		Hedges' correction	-.610

a. The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation of the mean difference.
Hedges' correction uses the sample standard deviation of the mean difference, plus a correction factor.

Reliability

Notes

Output Created		26-MAR-2022 16:27:30
Comments		
Input	Active Dataset	DataSet3
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	10
	Matrix Input	
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics are based on all cases with valid data for all variables in the procedure.

Syntax		RELIABILITY
		
		/VARIABLES=Itrustedther obotduringinteraction InverseofC
		  /SCALE('ALL VARIABLES') ALL
		  /MODEL=ALPHA
		
		/STATISTICS=DESCRIPTIV E SCALE CORR
		  /SUMMARY=TOTAL.

Resources	Processor Time	00:00:00.01
	Elapsed Time	00:00:00.00

[DataSet3]

Scale: ALL VARIABLES

Case Processing Summary

		N	%
Cases	Valid	10	100.0
	Excluded[a]	0	.0
	Total	10	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.801	.821	2

Item Statistics

	Mean	Std. Deviation	N
"I trusted the robot during interaction."	5.30	1.636	10
Inverse of C	3.80	1.229	10

Inter-Item Correlation Matrix

	"I trusted the robot during interaction."	Inverse of C
"I trusted the robot during interaction."	1.000	.696
Inverse of C	.696	1.000

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation
"I trusted the robot during interaction."	3.80	1.511	.696	.484
Inverse of C	5.30	2.678	.696	.484

Item-Total Statistics

	Cronbach's Alpha if Item Deleted
"I trusted the robot during interaction."	.
Inverse of C	.

Scale Statistics

Mean	Variance	Std. Deviation	N of Items
9.10	6.989	2.644	2

General Linear Model

Notes

Output Created		26-MAR-2022 16:40:47
Comments		
Input	Active Dataset	DataSet4
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	10
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics are based on all cases with valid data for all variables in the model.
Syntax		GLM t1ma t1mb t2ma t2mb t3ma t3mb /WSFACTOR=Task 3 Polynomial Methods 2 Polynomial /METHOD=SSTYPE(3) /PLOT=PROFILE (Task*Methods) TYPE=BAR ERRORBAR=CI MEANREFERENCE=NO /EMMEANS=TABLES (Task) /EMMEANS=TABLES (Methods) /EMMEANS=TABLES (Task*Methods) /PRINT=DESCRIPTIVE /CRITERIA=ALPHA(.05) /WSDESIGN=Task Methods Task*Methods.
Resources	Processor Time	00:00:00.56
	Elapsed Time	00:00:01.00

Within-Subjects Factors

Measure: MEASURE_1

Task	Methods	Dependent Variable
1	1	t1ma
	2	t1mb
2	1	t2ma
	2	t2mb
3	1	t3ma
	2	t3mb

Descriptive Statistics

	Mean	Std. Deviation	N
t1ma	3.86228674	.981744379	10
t1mb	5.39376155	.841951468	10
t2ma	7.99561986	.928154831	10
t2mb	8.30803692	1.13825017	10
t3ma	2.08388428	1.01167692	10
t3mb	5.43984765	.836397790	10

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.
Task	Pillai's Trace	.971	131.798 [b]	2.000	8.000	<.001
	Wilks' Lambda	.029	131.798 [b]	2.000	8.000	<.001
	Hotelling's Trace	32.950	131.798 [b]	2.000	8.000	<.001
	Roy's Largest Root	32.950	131.798 [b]	2.000	8.000	<.001
Methods	Pillai's Trace	.833	44.927 [b]	1.000	9.000	<.001
	Wilks' Lambda	.167	44.927 [b]	1.000	9.000	<.001
	Hotelling's Trace	4.992	44.927 [b]	1.000	9.000	<.001
	Roy's Largest Root	4.992	44.927 [b]	1.000	9.000	<.001
Task * Methods	Pillai's Trace	.682	8.579 [b]	2.000	8.000	.010
	Wilks' Lambda	.318	8.579 [b]	2.000	8.000	.010
	Hotelling's Trace	2.145	8.579 [b]	2.000	8.000	.010
	Roy's Largest Root	2.145	8.579 [b]	2.000	8.000	.010

a. Design: Intercept
   Within Subjects Design: Task + Methods + Task * Methods

b. Exact statistic

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser
Task	.859	1.217	2	.544	.876
Methods	1.000	.000	0	.	1.000
Task * Methods	.973	.220	2	.896	.974

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Epsilon[b] Huynh-Feldt	Lower-bound
Task	1.000	.500
Methods	1.000	1.000
Task * Methods	1.000	.500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept
   Within Subjects Design: Task + Methods + Task * Methods

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F
Task	Sphericity Assumed	216.261	2	108.131	128.099
	Greenhouse-Geisser	216.261	1.753	123.391	128.099
	Huynh-Feldt	216.261	2.000	108.131	128.099
	Lower-bound	216.261	1.000	216.261	128.099
Error(Task)	Sphericity Assumed	15.194	18	.844	
	Greenhouse-Geisser	15.194	15.774	.963	
	Huynh-Feldt	15.194	18.000	.844	
	Lower-bound	15.194	9.000	1.688	
Methods	Sphericity Assumed	45.064	1	45.064	44.927
	Greenhouse-Geisser	45.064	1.000	45.064	44.927
	Huynh-Feldt	45.064	1.000	45.064	44.927
	Lower-bound	45.064	1.000	45.064	44.927
Error(Methods)	Sphericity Assumed	9.028	9	1.003	
	Greenhouse-Geisser	9.028	9.000	1.003	
	Huynh-Feldt	9.028	9.000	1.003	
	Lower-bound	9.028	9.000	1.003	
Task * Methods	Sphericity Assumed	23.463	2	11.732	11.107
	Greenhouse-Geisser	23.463	1.947	12.050	11.107
	Huynh-Feldt	23.463	2.000	11.732	11.107
	Lower-bound	23.463	1.000	23.463	11.107
Error(Task*Methods)	Sphericity Assumed	19.012	18	1.056	
	Greenhouse-Geisser	19.012	17.525	1.085	
	Huynh-Feldt	19.012	18.000	1.056	
	Lower-bound	19.012	9.000	2.112	

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Sig.
Task	Sphericity Assumed	<.001
	Greenhouse-Geisser	<.001
	Huynh-Feldt	<.001
	Lower-bound	<.001
Error(Task)	Sphericity Assumed	
	Greenhouse-Geisser	
	Huynh-Feldt	
	Lower-bound	
Methods	Sphericity Assumed	<.001
	Greenhouse-Geisser	<.001
	Huynh-Feldt	<.001
	Lower-bound	<.001
Error(Methods)	Sphericity Assumed	
	Greenhouse-Geisser	
	Huynh-Feldt	
	Lower-bound	
Task * Methods	Sphericity Assumed	<.001
	Greenhouse-Geisser	<.001
	Huynh-Feldt	<.001
	Lower-bound	.009
Error(Task*Methods)	Sphericity Assumed	
	Greenhouse-Geisser	
	Huynh-Feldt	
	Lower-bound	

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Task	Methods	Type III Sum of Squares	df	Mean Square	F
Task	Linear		7.502	1	7.502	10.370
	Quadratic		208.759	1	208.759	216.381
Error(Task)	Linear		6.511	9	.723	
	Quadratic		8.683	9	.965	
Methods		Linear	45.064	1	45.064	44.927
Error(Methods)		Linear	9.028	9	1.003	
Task * Methods	Linear	Linear	8.322	1	8.322	8.745
	Quadratic	Linear	15.141	1	15.141	13.044
Error(Task*Methods)	Linear	Linear	8.565	9	.952	
	Quadratic	Linear	10.447	9	1.161	

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Task	Methods	Sig.
Task	Linear		.010
	Quadratic		<.001
Error(Task)	Linear		
	Quadratic		
Methods		Linear	<.001
Error(Methods)		Linear	
Task * Methods	Linear	Linear	.016
	Quadratic	Linear	.006
Error(Task*Methods)	Linear	Linear	
	Quadratic	Linear	

Tests of Between-Subjects Effects

Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	1824.190	1	1824.190	2435.332	<.001
Error	6.741	9	.749		

Estimated Marginal Means

1. Task

Measure: MEASURE_1

Task	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	4.628	.197	4.184	5.073
2	8.152	.204	7.690	8.613
3	3.762	.204	3.300	4.223

2. Methods

Measure: MEASURE_1

Methods	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	4.647	.203	4.187	5.107
2	6.381	.131	6.085	6.676

3. Task * Methods

Measure: MEASURE_1

Task	Methods	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	3.862	.310	3.160	4.565
	2	5.394	.266	4.791	5.996
2	1	7.996	.294	7.332	8.660
	2	8.308	.360	7.494	9.122
3	1	2.084	.320	1.360	2.808
	2	5.440	.264	4.842	6.038

Profile Plots

Error bars: 95% CI
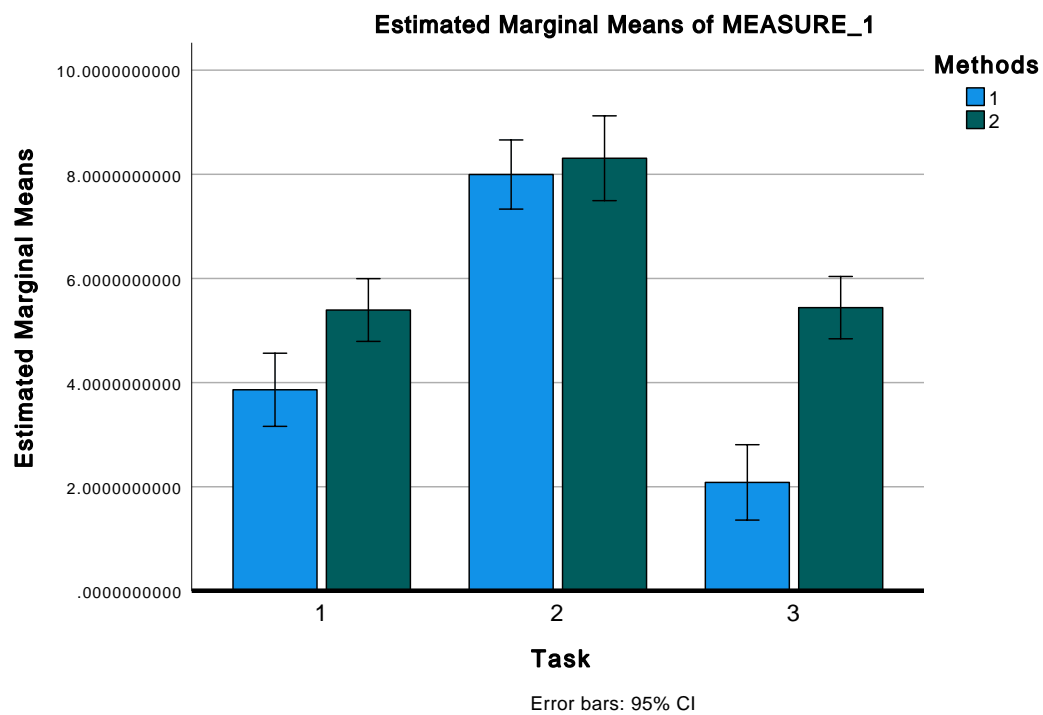
T-Test

Notes

Output Created		26-MAR-2022 16:41:34
Comments		
Input	Active Dataset	DataSet4
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	10
Missing Value Handling	Definition of Missing	User defined missing values are treated as missing.
	Cases Used	Statistics for each analysis are based on the cases with no missing or out-of-range data for any variable in the analysis.
Syntax		T-TEST PAIRS=t2ma WITH t2mb (PAIRED) /ES DISPLAY(TRUE) STANDARDIZER(SD) /CRITERIA=CI(.9500) /MISSING=ANALYSIS.
Resources	Processor Time	00:00:00.01
	Elapsed Time	00:00:00.00

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	t2ma	7.99561986	10	.928154831	.293508329
	t2mb	8.30803692	10	1.13825017	.359946308

Paired Samples Correlations

					Significance	
		N	Correlation	One-Sided p	Two-Sided p
Pair 1	t2ma & t2mb	10	-.233	.259	.518

Paired Samples Test

		Paired Differences			
					95% Confidence ...
		Mean	Std. Deviation	Std. Error Mean	Lower
Pair 1	t2ma - t2mb	-.312417057	1.62744897	.514644553	-1.47662392

Paired Samples Test

		Paired …			Significance	
		95% Confidence Interval of the …				
		Upper	t	df	One-Sided p	Two-Sided p
Pair 1	t2ma - t2mb	.851789805	-.607	9	.279	.559

Paired Samples Effect Sizes

			Standardizer[a]	Point Estimate	95% Confidence Interval	
					Lower	Upper
Pair 1	t2ma - t2mb	Cohen's d	1.62744897	-.192	-.813	.439
		Hedges' correction	1.78082251	-.175	-.743	.401

a. The denominator used in estimating the effect sizes.
   Cohen's d uses the sample standard deviation of the mean difference.
   Hedges' correction uses the sample standard deviation of the mean difference, plus a correction factor.