COMP 5517 HCI: Quantitative Evaluation Tutorial Exercise

For each of the scenarios 1 to 3 below, answer the following questions:

a. State the hypothesis clearly.
b. What is/are the independent variable(s) and the dependent variable(s)?
c. What are the two population groups?
d. What are the mean and variance of each population group?
e. What is the pooled variance, if applicable?
f. What is the t value?
g. Are these groups significantly different?

(For Questions b-g, verify your results.)

Scenario 1:

You are writing an application and are now working on the interface part. You have two choices for the menu: the traditional *drop-down menu*, but there's also something that you have heard about recently: the *pie menu*. You want to use the most efficient one possible, so you create two working prototypes of your application, keeping everything else the same, but just changing the interface styles. You then recruit *two groups of people* to do **Task A** on the application, and count the number of mouse clicks.

The results are as follows:

Menu Used	Number of mouse clicks
Drop-down	4, 4, 5, 5, 6, 6, 7, 7
Pie	3, 4, 4, 4, 5, 5, 5, 6

Scenario 2:

Company X has asked you to test a new drug for them. They claim that the new drug is supposed to boost people's performance, without causing any side effects. They have doctors who have verified the second claim, and now they want you to verify the first.

You bring in *two groups of experiment subjects*. One group is given the drug and the other group, a placebo. You then ask both groups to finish a **programming problem** in Java. The results are as follows:

Drug/Placebo	Time needed to finish program
Drug	165.9, 210.3, 166.8, 182.3, 182.1, 218, 170
Placebo	212.1, 203.5, 210.3, 228.4, 206.2, 203.2, 224.9, 202.1

Scenario 3:

Institute Y has asked you to test for a new education method for studying. They claim that taking the new education module could improve a student's test performance. You are asked to verify the claim. Compute also the correlation of the pre/post scores for the subjects.

You ask *a group of experiment subjects* to take a benchmark test and record the scores. After completing the education module, the subjects take another benchmark test at the same level of difficulty and the scores are recorded. The results as aligned with the same subject are as follows:

Pre/Post-education	Scores received
Pre-score	70, 62, 58, 61, 73, 82, 70, 88, 40, 66
Post-score	78, 70, 63, 68, 75, 95, 85, 87, 73, 70

Scenario 4:

Expert computer users like to show off by using the keyboard for *all* their controls: some of them know the keyboard short-cuts so well that the mouse could be missing and they couldn't tell the difference! They claim that this makes their performance faster.

COMP 5517 HCI: Quantitative Evaluation Tutorial Exercise

You want to test this claim experimentally. Describe your experiment. What are some sources of confounding? Discuss.

Worksheet

Recall that sample variance s^2 can be computed as:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \overline{y})^2$$
$$= \frac{1}{n-1}\left(\sum_{i=1}^{n} y_i^2 - n\overline{y}^2\right)$$

and the t-statistics value can be computed as:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1 - \bar{X}_2}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \text{where } s^2_{\bar{X}_1 - \bar{X}_2} = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{df}$$

Some partial computation results are given in the following tables.

Q2:

x	165.9	210.3	166.8	182.3	182.1	218	170		*1295.4*
y	212.1	203.5	210.3	228.4	206.2	203.2	224.9	202.1	*1690.7*
x^2	27522.81	44226.09	27822.24	33233.29	33160.41	47524	28900		*242388.84*
y^2	44986.41	41412.25	44226.09	52166.56	42518.44	41290.24	50580.01	40844.41	*358024.41*

Q3:

x	70	62	58	61	73	82	70	88	40	66	*670*
y	78	70	63	68	75	95	85	87	73	70	*764*
x^2	4900	3844	3364	3721	5329	6724	4900	7744	1600	4356	*46482*
y^2	6084	4900	3969	4624	5625	9025	7225	7569	5329	4900	*59250*
xy	5460	4340	3654	4148	5475	7790	5950	7656	2920	4620	*52013*

y-x	8	8	5	7	2	13	15	-1	33	4	*94*
$(y-x)^2$	64	64	25	49	4	169	225	1	1089	16	*1706*

Homework (recommended)

When you go home, try to perform similar analysis on interfaces A and B, measured by time using between subject design, and manual versus electric typewriters, measured by typing speed using within subject design. Is there any significant difference between interfaces A and B? Is electric typewriter superior to manual typewriter?

COMP 5517 HCI: Quantitative Evaluation Tutorial Exercise

The answers:

Q1:

Same as example walkthrough in slides, except sample set 1 and 2 are interchanged.

Means are 5.5 and 4.5, t value is 1.871

H0 is not rejected.

Q2:

Mean: 185.06, 211.34
Variance: 444.30, 102.30
Pooled variance: 260.15
DF: 13
t-value: 3.15
One-tailed value: 1.77

H0 is rejected.

Q3:

Correlation r = 0.697.

This is within subject design, so should use paired study, using the value x-y:
Mean: 9.4
Variance: 91.38
DF: 9
t-value: 3.11
One-tailed value: 1.83

H0 is rejected.

If you assume between subject design, you will get
Mean: 67.0, 76.4
Variance: 176.89, 97.82
Pooled variance: 137.36
DF: 18
t-value: 1.79
One-tailed value: 1.73
Two-tailed value: 2.10

H0 is marginally rejected, but will not be rejected based on two-tailed test.

Homework: no (t = 0.76) and yes (t = 5.88)