**Computational Social Intelligence COMPSCI4080  
Assessed Exercise Part-I Report**

Sarthak Ahuja

2707209A

University of Glasgow

MSc Robotics and Artificial Intelligence

**Problem Statement:**

We are provided with a small collection of laughter events (in file “laughter-corpus.csv”) observed during 60 phone conversations between 120 unacquainted speakers. Along with this, also provided is certain metadata for each laughter event recorded as follows:

* Gender of the person laughing: “Male” / “Female”
* Role of the person laughing: “Caller” / “Receiver”
* Duration of each laughter event (in seconds)
* Total speakers: 120 (wherein Male: 57 and Female: 63)
* Callers and Receivers are 60 speakers each

We are expected to use appropriate statistical tests to answer the four major questions regarding the count of laughter events being higher in which gender/role and the duration of laughter events being longer for women/callers separately.

For each of these questions, we need to initially formulate a clear research hypothesis and, correspondingly, a suitable null hypothesis. Then after selecting an appropriate statistical test and applying it, we need to explain whether we reject the null hypothesis and, if yes, what was the confidence level. We need to conclude explaining whether we use a two-tailed or a one-tailed test and the reasons for the same.

**Q1. Is the number of laughter events higher for women than for men?**

**Research hypothesis:** The count of laughter events for females is higher than the count of laughter events for males.

**Null hypothesis:** The count of laughter events for is independent of the gender (female/male).

Description of the data:

e.g., number of laughter occurrences for female and male subjects, average and variance of laughter length for female and male subjects, etc.

Description of the approach: e.g., which test? What are the parameters (degrees of freedom, expectations, etc.) and what is the motivation behind every choice (why a certain test? Why a certain value of the parameters?).

Results of the hypothesis testing: (can the null hypothesis be rejected? What is the confidence level? Etc.)

Explanation of the results: (e.g., in case the test shows that the number of laughter events is higher for female subjects to a statistically significant extent, you should mention that your conclusion is that female subjects tend to laugh more than male ones)

Analysis software you have written (the code must be added in appendix)

You have to show that you know how to calculate the statistic you use for the test (Chi Square or Student’s t).

**Q2. Is the number of laughter events higher for callers than for receivers?**

**Research hypothesis:** The count of laughter events for callers is higher than the count of laughter events for receivers.

**Null hypothesis:** The count of laughter events for is independent of the role (callers/receivers)

Description of the data:

e.g., number of laughter occurrences for female and male subjects, average and variance of laughter length for female and male subjects, etc.

Description of the approach: e.g., which test? What are the parameters (degrees of freedom, expectations, etc.) and what is the motivation behind every choice (why a certain test? Why a certain value of the parameters?).

Results of the hypothesis testing: (can the null hypothesis be rejected? What is the confidence level? Etc.)

Explanation of the results: (e.g., in case the test shows that the number of laughter events is higher for female subjects to a statistically significant extent, you should mention that your conclusion is that female subjects tend to laugh more than male ones)

**Q3. Are laughter events longer for women?**

**Research hypothesis:** The duration of laughter events for females is longer than those of males.

**Null hypothesis:** The duration of laughter events is independent of the gender.

Description of the data:

e.g., number of laughter occurrences for female and male subjects, average and variance of laughter length for female and male subjects, etc.

Description of the approach: e.g., which test? What are the parameters (degrees of freedom, expectations, etc.) and what is the motivation behind every choice (why a certain test? Why a certain value of the parameters?).

Results of the hypothesis testing: (can the null hypothesis be rejected? What is the confidence level? Etc.)

Explanation of the results: (e.g., in case the test shows that the number of laughter events is higher for female subjects to a statistically significant extent, you should mention that your conclusion is that female subjects tend to laugh more than male ones)

**Q4. Are laughter events longer for callers?**

**Research hypothesis:** The duration of laughter events for callers is longer than those of receivers.

**Null hypothesis:** The duration of laughter events is independent of the role (callers/receivers)

Description of the data:

e.g., number of laughter occurrences for female and male subjects, average and variance of laughter length for female and male subjects, etc.

Description of the approach: e.g., which test? What are the parameters (degrees of freedom, expectations, etc.) and what is the motivation behind every choice (why a certain test? Why a certain value of the parameters?).

Results of the hypothesis testing: (can the null hypothesis be rejected? What is the confidence level? Etc.)

Explanation of the results: (e.g., in case the test shows that the number of laughter events is higher for female subjects to a statistically significant extent, you should mention that your conclusion is that female subjects tend to laugh more than male ones)

Appendix

References

1. D.C.Howell, “Statistical Methods for Psychology”, Cengage Learning, 2009
2. F.Camastra and A.Vinciarelli, “Machine Learning for Audio, Image and Video Processing”, Springer Verlag, F.Camastra and A.Vinciarelli, “Machine Learning for Audio, Image and Video Processing”, Springer Verlag, 2008.
3. [A.Vinciarelli, P.Chatziioannou & A.Esposito, “When the Words are not Everything: The Use of Laughter, Fillers, Back-Channel, Silence and Overlapping Speech in Phone Calls”, Frontiers in ICT, 2:4, 2015](https://moodle.gla.ac.uk/mod/resource/view.php?id=2576040)