**TP3 – Statistics**

Reminder of theoretical hypotheses:

* People who stutter (PWS) may have greater difficulties than typical persons (PNS) at producing a regular pattern
* PWS may have troubles at anticipating a regular event
* PWS may have troubles at initiating movements

Reminder of operational hypotheses :

* The average Inter Response Interval (IRI) should be close to 0.5 s in PNS. In PWS, it could be faster or slower than 0.5 s, or show a global acceleration or deceleration (i.e. increased or decreased IRI over time).
* The IRI variance (over each train or cycle of taps) should be reduced in PNS, compared to PWS.
* The distribution of reaction times (RT) in the Aperiodic condition should have a Gaussian shape with a small standard deviation in SNPs. This distribution could be almost flat in PWS, or remain Gaussian but with a large standard deviation in PWS.
* The mean RT is expected to be positive in the Aperiodic condition for both groups of subjects, and longer in PWS than in PNS.
* The mean RT should be significantly reduced in PNS in the PeriodicAlong condition, compared with the Aperiodic condition, and become close to 0ms
* For PWS, the mean RT should remain positive in the PeriodicAlong condition, comparable to the mean RT observed in the Aperiodic condition. Or be reduced in the "PeriodicAlong" condition, but significantly less than in PNS.

**1. Descriptive statistics (with Python)**

Used packages:

Matplotlib (import matplotlib.pyplot as plt)

Seaborn (﻿import seaborn as sns)

Pandas (﻿import pandas as pd)

1- Read the data table as a dataframe

Data = pd.read\_table(“Res-AllTaps.txt”, header='infer', delimiter = "\t", decimal='.', na\_values='NaN')

﻿print(Data.columns.values) # to see the names of the different columns

2- Filter the dataframe to consider only a part of the data﻿

Data\_PeriodicAlong = Data[Data[Pattern]==1]

Data\_Aperiodic = Data[Data[Pattern]==2]

Data\_PNS = Data[(Data[Pattern]==1)&(Data['Group']==0)]

Data\_PWS = Data[(Data[Pattern]==1)&(Data['Group']==1)]

3- Calculate the mean and standard deviation of a data set

﻿Data\_PNS['IRI'].mean(skipna = True)

Data\_PNS['IRI'].std(skipna = True)

4- Plot the mean value and standard deviation of the IRI for both groups (PWS and PNS) in the PeriodicAlong condition only (example below)

﻿g = sns.catplot(data= Data\_PeriodicAlong, kind="bar", x="Group", y="IRI", ci="sd", alpha=.6, height=4) # alpha=transparency height = size of the graph

g.set\_axis\_labels ("Group", "IRI(s)")

g.set (ylim=(0.4, 0.6)) # It's up to you to find the right time range to represent

plt.show()

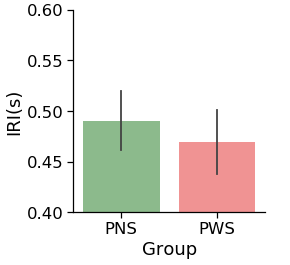
Some elements to configure the graphic aspect ...﻿

sns.set\_context("notebook", font\_scale=1.5, rc={"lines.linewidth": 1}) # font size and error bar thickness

#sns.set\_style("ticks") #sns.set\_style("white") #sns.set\_style("whitegrid")

MyPalette = ["#339933", "#ff3333", "#ff0000", "#e59b8b", "#ff2319"]

sns.set\_palette(MyPalette)



5- Graphically represent, for each group, the variation of the IRI over time (to detect a possible acceleration or deceleration)

Create a new variable TapNb from TrainNumber and BeatNumber (ranging from 1 to 24):

Data['TapNb'] = Data['BeatNumber']+ 8\*(Data['TrainNumber']-1)

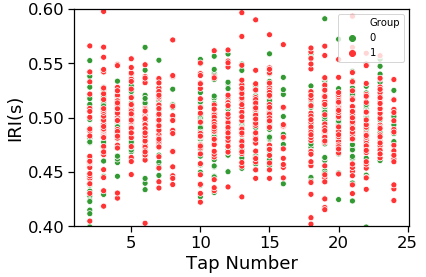
print(Data.columns.values)

g=sns.scatterplot(x="TapNb", y="IRI", hue="Group", data=Data) # legend=False if you don't want a legend on the graphg.set\_ylabel('IRI(s)')

g.set\_xlabel('Tap Number')

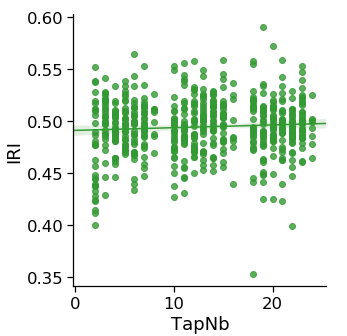
g.legend(loc = 'upper right', fontsize=10)

g.set(ylim=(0.4, 0.6))



Use the function « lmplot » ou « regplot » to superimpose a trend line (linear regression) on each data set:

g=sns.lmplot(x="TapNb", y="IRI", data = Data\_PNS)



6- Plot, for each group, the mean value and standard deviation of the IRI for both groups (PWS and PNS) in the PeriodicAlong condition only

Read the datatable « Res-stdIRI.txt » as a new Dataframe (for example DataTrain)

Visualize, as before, the mean and standard deviation of the variable "StdIRI" for the two groups PWS and PNS.

7- Plot, for each group, the distribution of RT values in the Aperiodic condition for both groups (PWS and PNS)

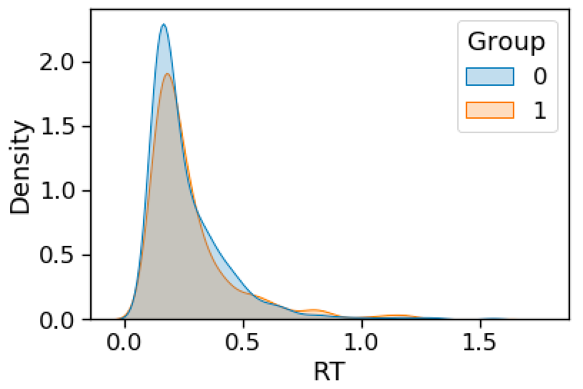
Data\_Ap = Data[Data['Pattern']==2]

sns.kdeplot(data=Data\_Ap, x="RT", hue="Group", common\_norm=True)

﻿ Some elements to configure the graphic aspect …

﻿sns.set\_context("notebook", font\_scale=1.5, rc={"lines.linewidth": 1}) # font size and error bar thickness﻿

sns.kdeplot(data=Data\_Ap, x="RT", hue="Group", common\_norm=True, fill=True)



8- Plot the mean value and standard deviation of RT for the two groups (PWS and PNS) under the two conditions PeriodicAlong and Aperiodic (example below)

g = sns.catplot(data=Data, kind="bar",x="Condition", y="RT", hue="Group",ci="sd", alpha=.6, height=4)

g.set\_axis\_labels("Group", "RT(s)")

g.set(ylim=(-0.25, 0.75))

