Detecting Parkinson's disease using Vocal Data from Patients

Applied Probability and Statistics for Engineers

Submitted to: Dr. Amar Sabih

Team 1:

Abdulah Alhoothy: 40075668

Muhammad Haris: 40150933

Qamar Bani-Melhem: 40109566

Sourav Patne: 40137125

Yaoxin Li: 40092820

Concordia University
Gina Cody School of Engineering and Computer Science
Montreal, Quebec, Canada

Table of contents

List of Tables

List of Figures

1.Abstract

Through the use of dysphonia (voice), we present a method for distinguishing between individuals who are healthy and those that have Parkinson's disease (PD). The analysis is of practical value because it combines both traditional and non-standard measures. Voice frequencies usually have many variations which are hard to discriminate, which is normal health or environmental acoustic noise. For this problem a new measure of voice has been introduced; Pitch Period Entropy (PPE) is a very robust measurement to any uncontrollable confounding effects including noisy

This dataset is composed of a range of biomedical voice measurements from 31 people, 23 with Parkinson's disease (PD). We then performed cross collation to identify 10 uncorrelated measurements. These 10 uncorrelated measurements were plotted to see if anyone of them could be a good signal determinate for the presence of Parkinson's. Using machine learning and linear regression model we were able to weigh each of the 10 parameters for importance and built a prediction model which was 94.8% accurate. In conclusion, we find that non-standard methods in combination with voice discriminators could be 94.8% capable of separating individuals with PD from healthy ones.

2.Introduction

There are various medical exams that a person has to go through if he/she wants to be diagnosed to identify if they have Parkinson's. The level of severity of Parkinson's disease is identified by the doctor who made the tests. By using dysphonia data taken from both healthy and patients with Parkinson's disease. We hope in our study to be able to discriminate between the two groups. An average of 6 records was taken from each patient [31 patients]. Using calibrated and verified measurement procedures to detect speech signals is the main method by which the data was collected.

Vocal impairment was the main way in which Parkinson's was exhibited in the patients, (nearly 90% of patients) [3, 4]. At the beginning of the illness, vocal impairments are the signs that doctors would look for this makes it particularly convenient because the measurement of vocal impairment is noninvasive and is simple to carry out[5, 6, 7]. So the significance of tracking the progression of symptoms through voice measurement is very high[8].

For this study, Using a Multi-Dimensional Voice Program (MDVP) as the primary of recording the phonations and as a software tool for quantitative acoustic assessment of the quality of a patient's voice, all the recording were carried out in a laboratory and under observation. the MDVP can measure up to 22 parameters on a single voice line from a patient [8].

3.Problem description

Parkinson's disease is a neurodegenerative, progressive disorder of the central nervous system that affects movement and causes tremors and stiffness. This affects dopamine-producing neurons in the brain; every year it affects more than 10 million individuals. Recently, researchers have begun to utilize data science to improve healthcare and services – predicting diseases early will have countless advantages on the prognosis. There are different methods such as Vocal tests, Movement tests etc. to detect whether the person is having Parkinson's or not. In our study, we are using Vocal test results. We use 24 different vocal Phonation such as Jitter, Shimmer, Harmonics to noise ratio etc. which are taken from patients under standard conditions. We minimize these attributes using different statistical techniques to find out which of the attributes are most likely dependent on the final status of the patient.

4.Assumptions and limitations

4.1. Assumptions

1. The measurements detected by the equipment are properly measured.
2. The doctor's diagnosis is assumed to be correct.
3. All the testing equipment are in the best working conditions.
4. Testing machines are properly calibrated as per prescribed guidelines.
5. All the standards and SOP's are followed during the testing of patients.
6. All the designated staff are properly trained for utilizing the testing equipment at its best working condition.

4.2. Limitations

1. The data used for analysis is limited i.e. limited patients under observation.
2. The measuring equipment has limitations on measuring the values.
3. Limited access to the patient's medical history, therefore we can't use any other information.
4. Project team members are limited to work on the data as available in the file.

5.Data analysis

In this report, we are introducing the Parkinson disease (PD) and the type of parameters that had been tested on the patients. We use real data from source [1], and analyze the data with different ways. We studied 195 records for 31 patients, 147 records out of 195 with PD which is 23 patients with PD. Then we select the highly 10 uncorrelated parameters and start working on them.

The period since diagnosis with PD is from 0 to 28 years, and the ages of the subjects are from 46 to 85 [8]. Multiple phonation tests were taken by the subjects (Averages of six phonations were recorded from each subject) ranging from one to 36 seconds in length [8].

The data chosen for the project had several parameters listed in the following table:

Attributes	Definition
MDVP (FO)	Fundamental frequency (Fo) is the vibratory rate of the vocal folds. It can be measured in hertz or cycle per second (CPS). Average fundamental frequency during a conversation for males ranges from 100 to 150 Hz, whereas for females it ranges from 180 to 250 Hz.
MDVP(FHI)	Maximum FO.
MDVP(FLO)	Minimum FO.
MDVP (Shimmer):	Shimmer is a measure of amplitude instability.
(HNR)	Harmonics -to- Noise Ratio.
(D2)	Signal fractal scaling exponent.
Spread 1-2	Two nonlinear measures of fundamental frequency variation.
(RPDE)	Recurrence period density entropy.
(DFA)	Detrended fluctuation analysis.

Table 1 Attributes used in the project.

The following table shows the list of subjects with sex, age, Parkinson's stage and the number of years since diagnosis: Entries labelled "n/a" for healthy subjects, for which Parkinson's stage and years since diagnosis is not applicable. "H&Y" refers to the Hoehn and Yahr PD stage, where higher values indicate a greater level of disability.

Subject code	Sex	Age	Stage (H&Y)	Years since diagnosis
S01	M	78	3.0	0
S34	F	79	2.5	¼
S44	M	67	1.5	1
S20	M	70	3.0	1
S24	M	73	2.5	1
S26	F	53	2.0	1½
S08	F	48	2.0	2
S39	M	64	2.0	2
S33	M	68	2.0	3
S32	M	50	1.0	4
S02	M	60	2.0	4
S22	M	60	1.5	4½
S37	M	76	1.0	5
S21	F	81	1.5	5
S04	M	70	2.5	5½
S19	M	73	1.0	7
S35	F	85	4.0	7
S05	F	72	3.0	8
S18	M	61	2.5	11
S16	M	62	2.5	14
S27	M	72	2.5	15
S25	M	74	3.0	23
S06	F	63	2.5	28
S10 (healthy)	F	46	n/a	n/a
S07 (healthy)	F	48	n/a	n/a
S13 (healthy)	M	61	n/a	n/a
S43 (healthy)	M	62	n/a	n/a
S17 (healthy)	F	64	n/a	n/a
S42 (healthy)	F	66	n/a	n/a
S50 (healthy)	F	66	n/a	n/a
S49 (healthy)	M	69	n/a	n/a

Table 2: List of subjects with sex, age, Parkinson's stage and the number of years since diagnosis [8].

Here we are showing a sample of the dataset we studied :

name	MDVP:Fo(Hz)	MDVP:Fhi(Hz)	MDVP:Flo(Hz)	MDVP:Shimmer	HNR	RPDE	DFA	spread2	D2	spread1
phon_R01	119.992	157.302	74.997	0.04374	21.033	0.414783	0.815285	0.266482	2.301442	-4.81303
phon_R01	122.4	148.65	113.819	0.06134	19.085	0.458359	0.819521	0.33559	2.486855	-4.07519
phon_R01	116.682	131.111	111.555	0.05233	20.651	0.429895	0.825288	0.311173	2.342259	-4.44318
phon_R01	116.676	137.871	111.366	0.05492	20.644	0.434969	0.819235	0.334147	2.405554	-4.1175
phon_R01	116.014	141.781	110.655	0.06425	19.649	0.417356	0.823484	0.234513	2.33218	-3.74779

Table 3 : Sample of dataset

5.1.Methodology

In our study, we have a Dataset of 195 records, 147 of these records shows an indication of having a Parkinson disease. The status attribute shows if the person is an illness with 1 or not with 0. For each patient, after using cross-correlation we decided to study 10 out of 24 attributes. The following steps are our work on the dataset we have:

1. We calculate the cross-correlation for the 24 attributes, then we compare every two attributes. The attributes that they had a high positive correlation - more than (0.65)- we correlated them with the target (0, 1).

   For example, spread 1 and ppe their correlation is equal to (0.96) which is more than (0.65); we know that they are both correlated with each other. First, we are correlating them with the target (status). Then, we will drop whichever has a weaker correlation with the target; in this case, ppe has the least correlation which is (0.53). As shown in the following table :

	status	*spread1*	*PPE*
status	1		
spread1	0.564837997	1	
PPE	0.531039154	0.962435	1

Table 4 : Correlation between PPE and Spread1.

2. We reduced the attributes to 10 according to our results from the correlation. The following table shows the correlation between the 10 attributes

	MDVP:Fo(Hz)	*MDVP:Fhi(Hz)*	*MDVP:Flo(Hz)*	*MDVP:Shimmer*	*HNR*	*RPDE*	*DFA*	*spread2*	*D2*	*spread1*
MDVP:Fo(Hz)	1.00									
MDVP:Fhi(Hz)	0.40	1.00								
MDVP:Flo(Hz)	0.60	0.08	1.00							
MDVP:Shimmer	-0.10	0.00	-0.14	1.00						
HNR	0.06	-0.02	0.21	-0.84	1.00					
RPDE	-0.38	-0.11	-0.40	0.45	-0.60	1.00				
DFA	-0.45	-0.34	-0.05	0.16	-0.01	-0.11	1.00			
spread2	-0.25	0.00	-0.24	0.45	-0.43	0.48	0.17	1.00		
D2	0.18	0.18	-0.10	0.51	-0.60	0.24	-0.17	0.52	1.00	
spread1	-0.41	-0.08	-0.39	0.65	-0.67	0.59	0.20	0.65	0.50	1.00

Table 5 : Correlation for 10 attributes.

3. We plotted a histogram for each attribute (column) and analyzed the distribution for the data. The results will be shown in the statistical analysis.

4. We picked a sample of 150 rows for each attribute. Then we applied the Descriptive Statistical analysis (SD, variance, mean, median, standard error and confidence interval of 95%) for each attribute. The results will be shown in the statistical analysis.

5. We classified the patients into stages according to their records and check whether if they are interfering in their results or not. The results will be shown in the statistical analysis

6. We apply the proportion hypothesis test on the status attribute and analyze the results depending on p-value and alpha. The results will be shown in the statistical analysis.

7. We test the normality for every attribute. The results will be shown in the statistical analysis.

8. We did a multiple linear regression for the attributes and find the equation for the results. The results will be shown in the statistical analysis.

6.Statistical analysis

6.1 histogram

We used histogram for analyzing all attributes of patients. We will provide the visual impression of the shape of each attribute, distribution of the measurements, information about the central tendency and scatter in the data.

6.1.1 MDVP: Fo(Hz)

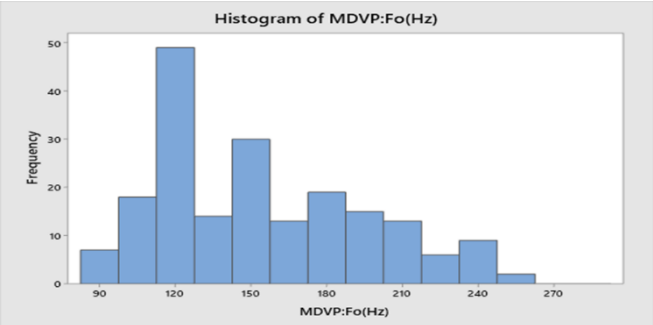


Figure 1: Histogram of MDVP :Fo(Hz)

Shape: The distribution of MDVP: Fo(Hz) is unimodal it has one mode at a value of 120 about which the observations are concentrated. It is right-skewed, larger values at the right tail are greater than the left tail.

Outliers: No outliers exist.

Centre: The centre of distribution approximately 165.

Spread: The data of this attribute range from 85 to 255.

The test of MDVP Fo data is mostly distributed at 120 with a frequency of 49, the 2nd is about 150 with a frequency of 30 and the smallest frequency is at 255. The values at 135 and 165 have almost the same frequency.

6.1.2 MDVP: Fhi(Hz)

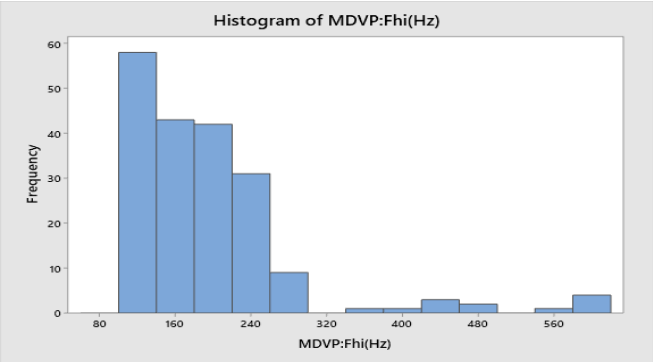


Figure 2: Histogram of MDVP: Fhi(Hz)

Shape: The distribution of MDVP: Fhi(Hz) is unimodal it has one mode at a value of 120 about which the observations are concentrated. It is right-skewed, the larger values at the right tail are greater than the left tail.

Outliers: Outliers exist.

Centre: The centre of distribution approximately at 200.

Spread: The data of this attribute range from 120 to 300.

6.1.3 MDVP: Flo(Hz)

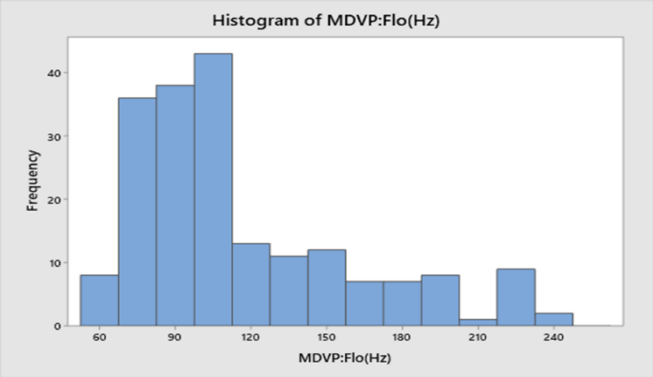


Figure 3: Histogram of MDVP: Flo(Hz)

Shape: The distribution of MDVP: Flo(Hz) is unimodal it has one mode at a value of 105 about which the observations are concentrated. It is right-skewed, the larger values at the right tail are greater than the left tail.

Outliers: No Outliers exist.

Centre: The centre of distribution approximately 135.

Spread: The data of this attribute range from 60 to 240.

6.1.4 MDVP (Shimmer)

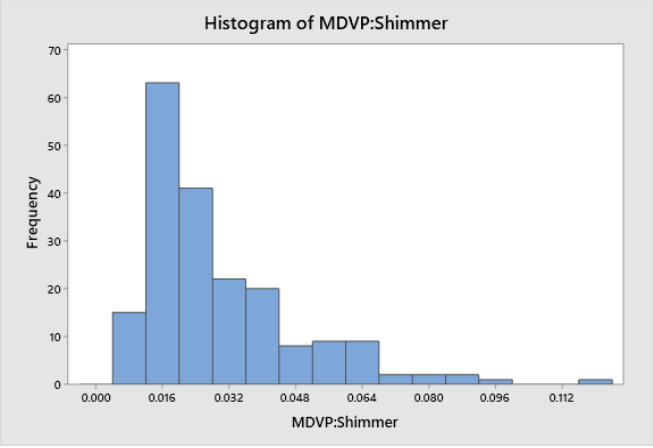


Figure 4: Histogram of MDVP: Shimmer

Shape: The distribution of MDVP Shimmer is unimodal it has one mode at a value of 0.016 about which the observations are concentrated. It is right-skewed the larger values at the right tail is greater than the left tail.

Outliers: Potential Outliers exist.

Centre: The centre of distribution approximately 0.048.

Spread: The data of this attribute range from 0.08 to 0.096.

6.1.5 HNR

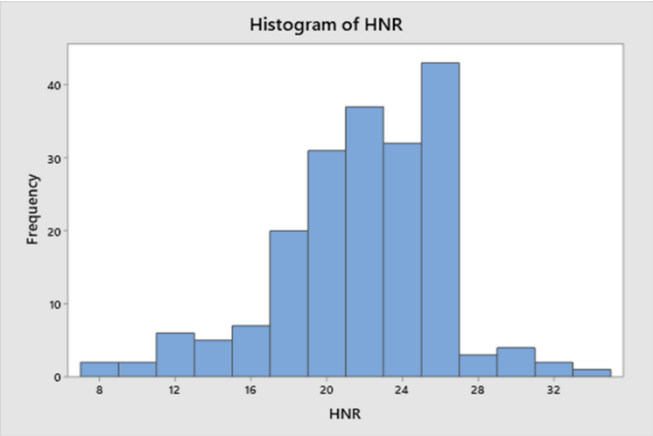


Figure 5: Histogram of HNR

Shape: The distribution of HNR is unimodal it has one mode at a value of 26 about which the observations are concentrated. It is left-skewed the larger values at the left tail is greater than the right tail.

Outliers: No Outliers exist.

Centre: The centre of distribution approximately at 22.

Spread: The data of this attribute range from 8 to 34.

6.1.6 RPDE

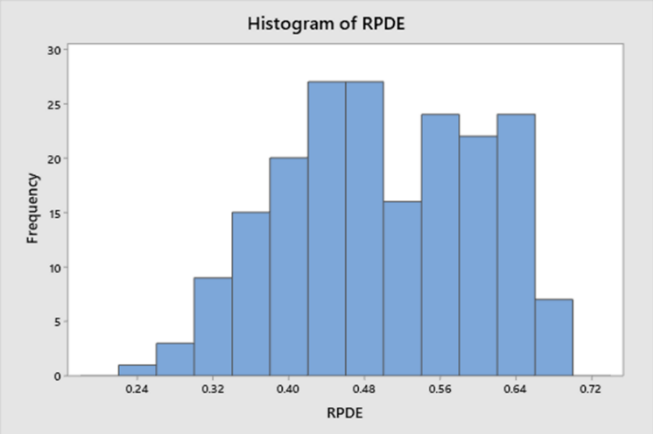


Figure 6: Histogram of RPDE

Shape: The distribution of RPDE is bimodal, it has one mode at a value of 0.44 and another mode at a value of 0.48 about which the observations are concentrated.

Outliers: No Outliers exist.

Centre: The centre of distribution approximately 0.44.

Spread: The data of this attribute range from 0.24 to 0.68.

6.1.7 DFA

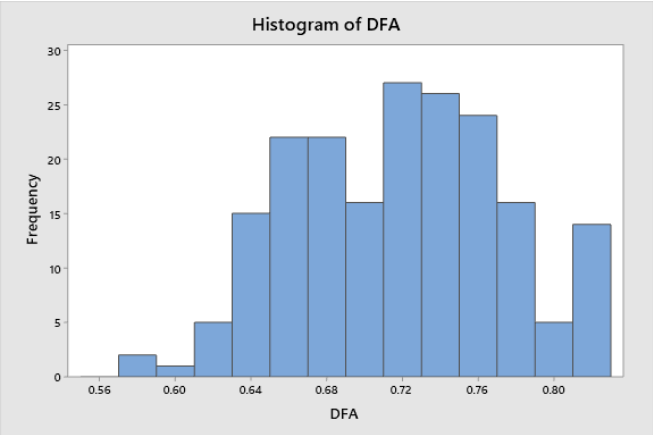


Figure 7: Histogram of DFA

Shape: The distribution of DFA is bimodal or double peak it has one mode at a 0.66 and 0.68and the other at 0.72 about which the observations are concentrated.

Outliers: No potential outliers exist.

Centre: The centre of distribution approximately 0.44.

Spread: The data of this attribute range from 0.24 to 0.68.

6.1.8 Spread 1

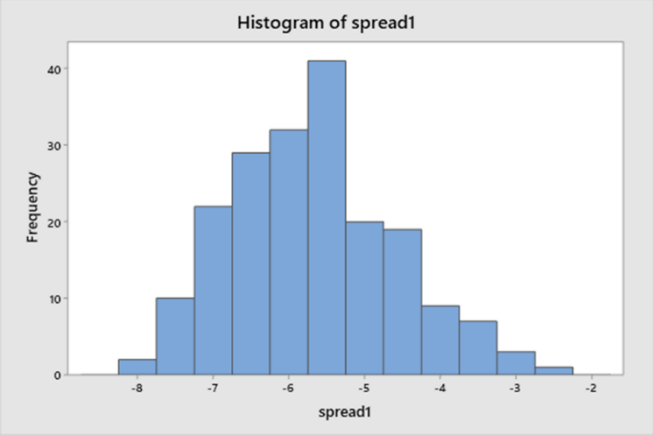


Figure 8: Histogram of Spread 1

Shape: The distribution of spread 1 is unimodal it has one mode at a value of -5.5 about which the observations are concentrated.

Outliers: No potential outliers exist.

Centre: The centre of distribution approximately at -5.5.

Spread: The data of this attribute range from -8 to -2.5.

6.1.9 Spread 2

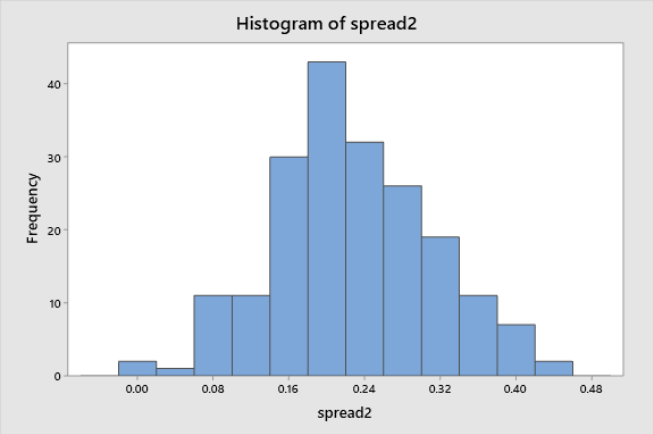


Figure 9: Histogram of Spread 2

Shape: The distribution of spread 2 is unimodal it has one mode at a value of 0.20 about which the observations are concentrated, but compared to other attributes it's approximately symmetrical

Outliers: No potential outliers exist.

Centre: The centre of distribution approximately 0.20.

Spread: The data of this attribute range from 0.0 to 0.44.

6.1.10 D2

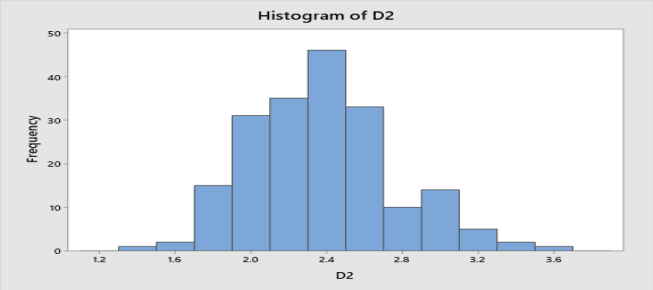


Figure 10: Histogram of D2

Shape: The distribution of D2 is unimodal it has one mode at a value of 2.4 about which the observations are concentrated but comparatively to other attributes its approximately symmetrical

Outliers: No potential outliers exist.

Centre: The centre of distribution approximately 2.4.

Spread: The data of this attribute range from 1.4 to 3.6.

6.2 Descriptive Statistical analysis

We've made our first look on the data and now we are ready to perform some basic data exploration and come up with some inference. Hence, the goal for this section is to take a glimpse on the data as well as any irregularities so that we can correct on the next section.

Attributes	Mean	95%Ci	Standard Error	Median	Standard Deviation	Variance	Mode
MDVP (FO)	154.45	6.47	3.27	151.88	39.97	1597.76	n/a
MDVP (FHI)	193.81	13.39	6.78	179.14	82.73	6845.02	n/a
MDVP (FLO)	115.65	7.06	3.57	104.77	43.59	1899.99	n/a
MDVP (Shimmer)	0.03	0.00	0.00	0.02	0.02	0.00	0.02
HNR	22.05	0.74	0.37	22.24	4.54	20.65	n/a
RPDE	0.50	0.02	0.01	0.50	0.10	0.01	n/a
Spread 1	-5.69	0.17	0.09	-5.66	1.07	1.14	n/a
Spread 2	0.23	0.01	0.01	0.23	0.09	0.01	0.21
DFA	0.72	0.01	0.00	0.73	0.05	0.00	n/a
D2	2.39	0.06	0.03	2.38	0.39	0.15	n/a

Table 6: Statistical analysis

(1) **Standard Deviation and Variance** represent the degree of dispersion between sample values. The standard deviation and variance of MDVP: Shimmer, RPDE, Spread 2, DFA and D2 are very small which means the data are concentrated and close to the mean.

(2) **Standard Error** is a measure to describe the dispersion of sampling distribution of the corresponding sample mean and the sampling error of the corresponding sample mean. The smaller the standard error is, the more accurate the estimation of the population mean is, and the more representative the sample data is. Such as MDVP: Shimmer, RPDE, Spread 1-2, DFA and D2.

(3) **95% Confidence interval** is an estimate given in interval form for an unknown parameter value in the parameter distribution of the population generating this sample. The MDVP: Shimmer, RPDE, Spread 2, DFA and D2 demonstrate their allowable error of average value are small.

(4) **The mode** of a set of data values is the value that appears most often. **The mean and median** are the statistic describing the degree of data concentration to determine the equilibrium point of a set of data.

6.3 Stages comparing Stages of the disease

Before attempting to build a linear regression model that would be able to inform us which parameter is most likely to predict the existence of a Parkinson's disease. We are going to do a cross-comparison with each of the 10 parameters. Each of the 10 parameters will be plotted in a box and whisker plot, each plot will have 7 box plots representing a stage of the disease. Essentially what we are trying to do here is evaluating the data to see if there is a clear difference between the data gathered from people in stage 1 of the disease vs people who are in stage 3. For example, if there was a clear distinction between the measurement of any two different groups of any of the 10 parameters then that could potentially indicate that said parameters would be a good reference to be more thoroughly investigate as a potential Parkinson's measure. Unfortunately since in each parameter, the box plots tend to overlap, this gives us no way of decerning a patient's health based on looking at his/her data compared to healthy patients.

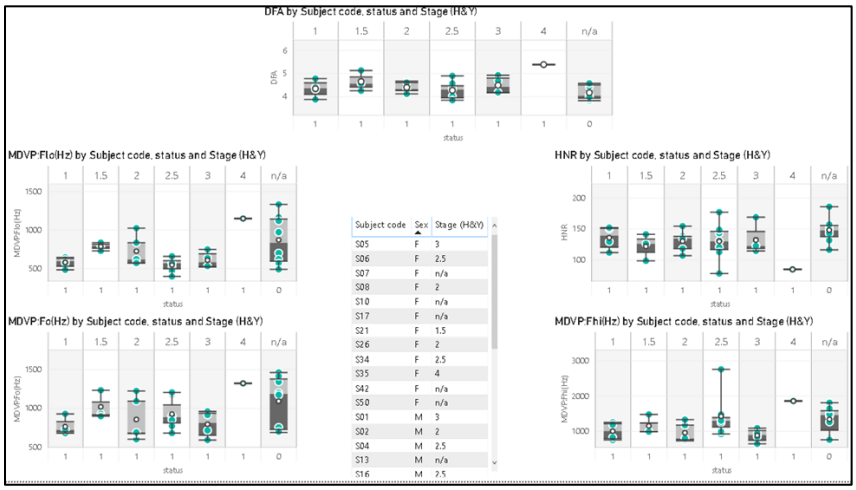


Figure 11: Stages of illness

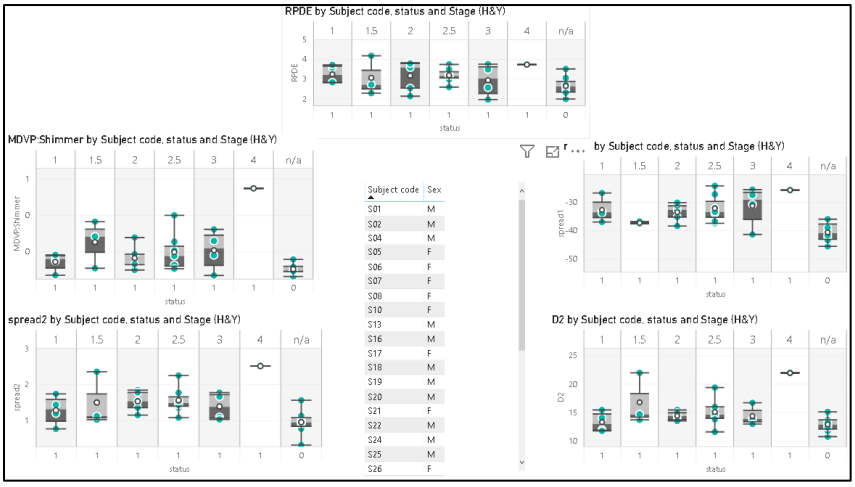


Figure 12: Stages of illness

6.4 Hypothesis test

According to our data, we applied the proportion hypothesis test using Minitab on the status attribute which has a (0,1) values.

Using the seven-step hypothesis-testing procedure as follows:

1. Parameter of Interest: The parameter of interest is the process fraction defective p.

2. Null hypothesis: H_0:p = 0.7.

3. Alternative hypothesis: H1: p ≠ 0.7.

4. The test statistic is:

$$z_0 = \frac{x - np_0}{\sqrt{np_0(1 - p_0)}}$$

5. Reject H_0 if: H_0:p = 0.7 if the p-value is less than 0.05.

6. Computations: The test statistic is equal to $z_0 = 1.64$.

7. Conclusions: Since $z_0 = 1.64$, the P-value is 0.101. So we fail to reject H_0 and conclude that 70% of the patients are illness.

Descriptive Statistics

N	Event	Sample p	95% CI for p
195	147	0.753846	(0.693385, 0.814307)

Test

		Z-Value	P-Value
Null hypothesis	H_0: p = 0.7		
Alternative hypothesis	H_1: p ≠ 0.7	1.64	0.101

6.5 Normality Test

We test the normality using Minitab. To decide which attribute follow the normality we check the P-value. If the P-value was less than (0.005), then the attribute fails to be normally distributed. On the other hand, if the P-value was more than (0.005), then the attribute follow the normality.

In the following figures, we will see which attribute is normally distributed and which one fail to be normal:

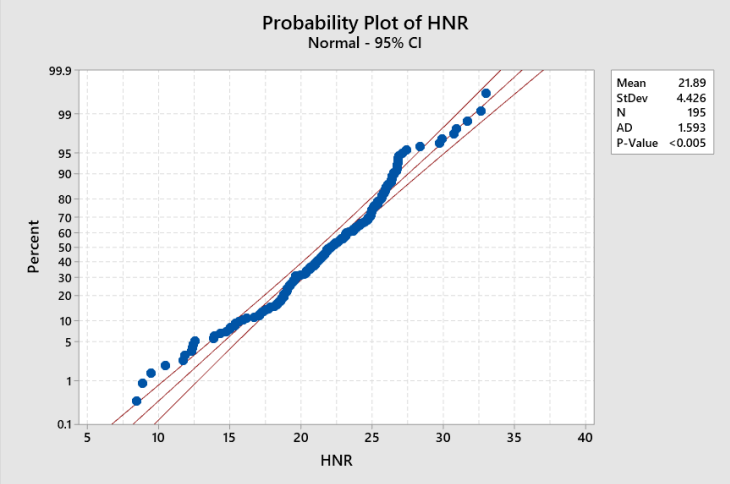


Figure 13: Probability plot of HNR- not normally distributed

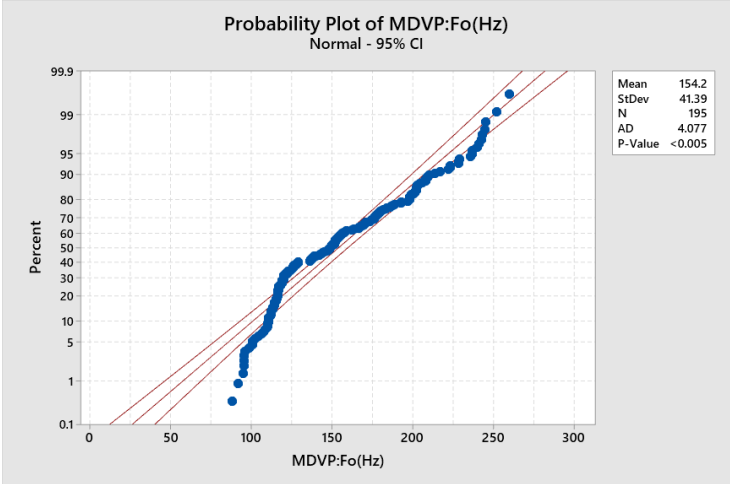


Figure 14: Probability plot of MDVP:Fo(Hz)- not normally distributed

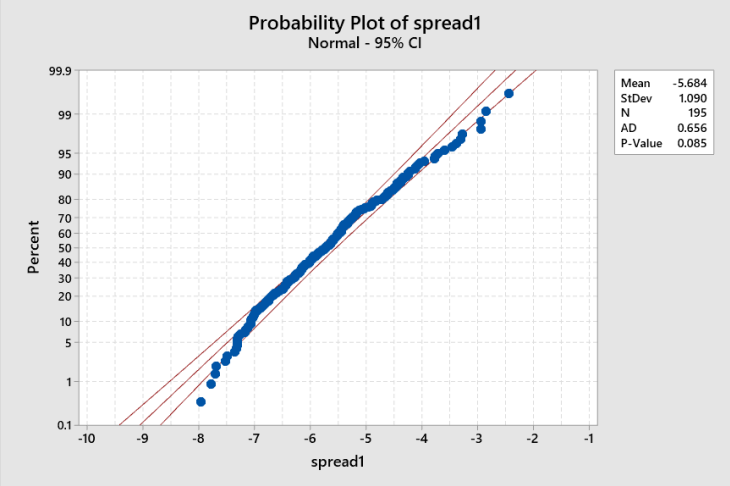


Figure 15: Probability plot of Spread1- normally distributed

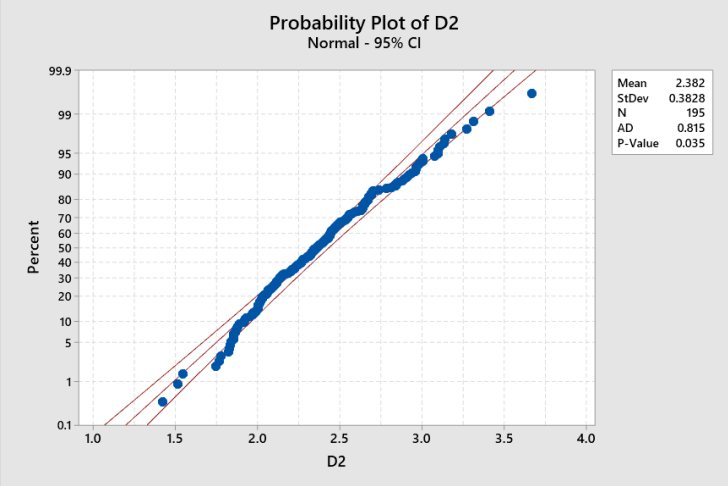


Figure 16: Probability plot of D2- normally distributed

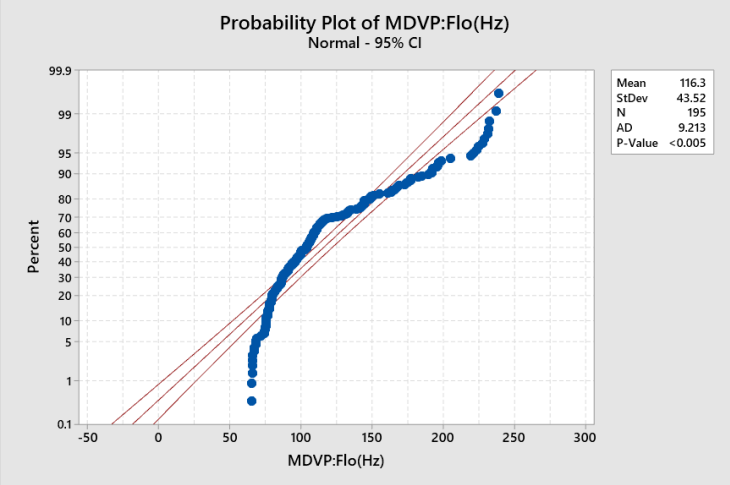


Figure 17: Probability plot of MDVP:Flo(Hz)- not normally distributed

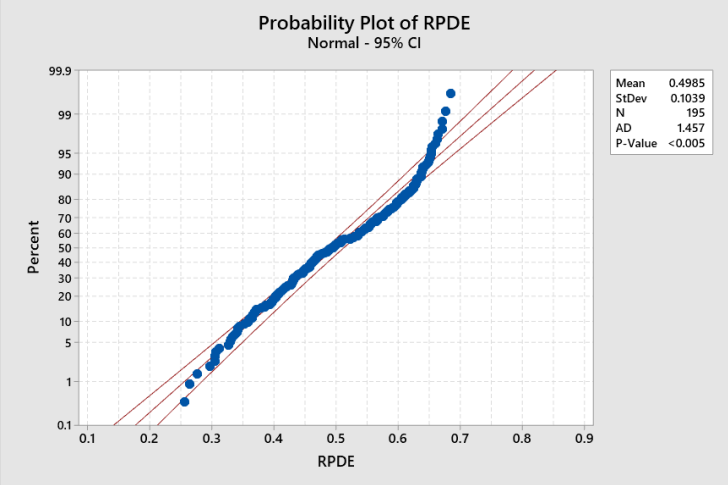


Figure 18: Probability plot of RPDE- not normally distributed

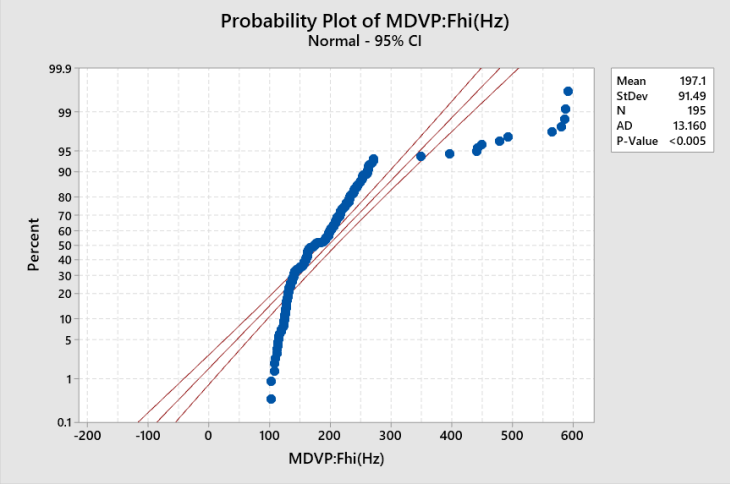


Figure 19: Probability plot of MDVP:Fhi(Hz)- not normally distributed

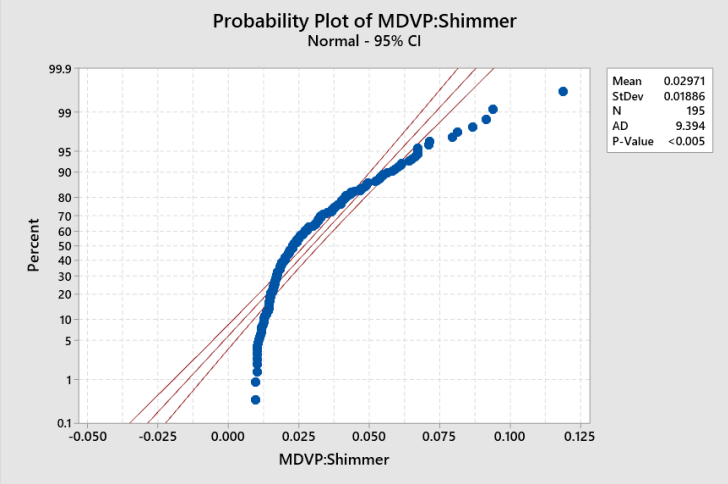


Figure 20: Probability plot of MDVP:Shimmer- not normally distributed

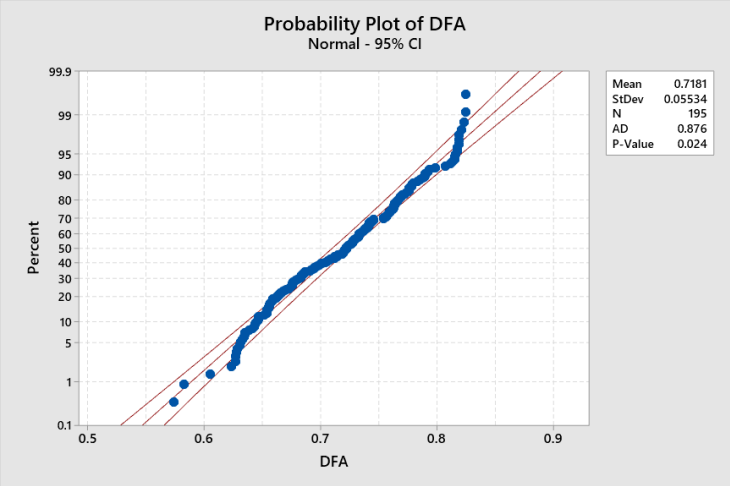


Figure 21: Probability plot of DFA- normally distributed

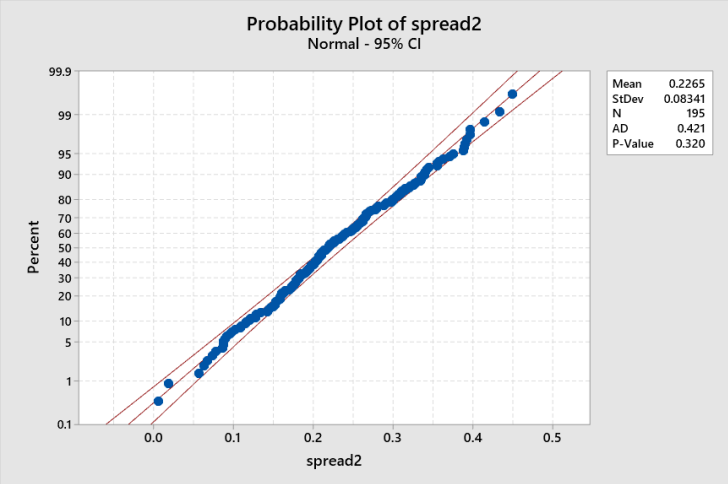


Figure 22: Probability plot of Spread2- normally distributed

6.6 Simple linear regression

Regression analysis is a way of mathematically sorting out which variables do have an impact. It also investigates the relationship between two or more variables and estimates one variable based on the others. It answers the questions: Which factors matter most?

In this study, the target variable that we are trying to predict is the "Status" and the predictors that we are going to use are the 10 selected features shown above. In the following image, you see an implementation of a regression model, it has been implemented through python.

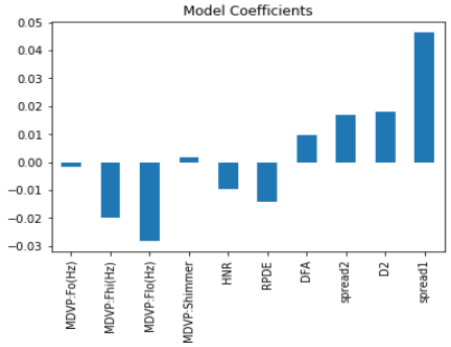
```
from sklearn.linear_model import LinearRegression
LR = LinearRegression(normalize=True)
predictors = ["MDVP:Fo(Hz)","MDVP:Fhi(Hz)","MDVP:Flo(Hz)","MDVP:Shimmer","HNR","RPDE","DFA","spread2","D2","spread1"]
target ="Status"
modelfit(LR, Rtrain, Rtest, predictors, target, 'LR.csv')

coef1 = pd.Series(LR.coef_, predictors)
coef1.plot(kind='bar', title='Model Coefficients')
```

```
[-0.00157734 -0.0197791  -0.02834955  0.00183221 -0.00973751 -0.01420971
  0.00981635  0.01692955  0.01816465  0.04656127]

Model Report
RMSE : 0.3399
CV Score : Mean - 0.3509 | Std - 0.0937 | Min - 0.1754 | Max - 0.4986

<matplotlib.axes._subplots.AxesSubplot at 0x22eb84060c8>
```



As seen from above, it seems that spread1 is the most important for predicting the status followed by MDVP: Flo(Hz) and so on, as shown in the below equation:

X represents the values gotten from the voice records, where (x_1) is the value of MDVP: Fo(Hz) and respectively for the other attributes. Y represents a range between (0,1) which indicates if the person has PD or not.

$$y = -0.00157734x_1 - 0.0197791x_2 - 0.02834955x_3 + 0.00183221x_4 - 0.00973751x_5 - 0.01420971x_6 + 0.00981635x_7 + 0.01692955x_8 + 0.01816465x_9 + 0.04656127x_{10}$$

Using the above model which was trained on 80% of the data, we tested it on the remaining 20%. We asked the model to use the 10 parameters of each voice line and give us a prediction of whether this voice line is of someone who has parkin's or not. The model was able to predict from the voice line the presence of Parkinson's disease correctly 94% of the time(meaning in every hundred 94 predictions are correct). We think that this accuracy is very good however more clinical trials are necessary to determine if this measure could be used under all circumstances.

7. Conclusion

In Conclusion, Through the use of dysphonia(voice) we present a method for distinguishing between individuals who are healthy and those that have Parkinson's disease (PD). With a total of 31 patients and a sample size of 195 an average of six Phonation's were taken from each patient; with different voice tests like Shimmer, Jitter, HNR, Spread 1 etc. ranging from one to 36 seconds in length [8]. By using different statistical tools like histogram, Cross-Correlation, Regression and Normality test the data was analysed. We have succeeded in reducing the unmanageable statistic data of 24 different voice attributes to 10 manageable attributes which are not correlated to each other. Using these 10 attributes we conducted the descriptive statistical analysis and the proportion hypothesis test for p-value with a 95% confidence interval. Relationship between these reduced attributes is found out using the linear regression. Our aim of predicting the status of the patient is achieved by this we can say that just by calculating/finding these 10 attributes we can predict the status of the patient. In conclusion, we find that non-standard methods in combination with voice discriminators could be 94.8% capable of separating individuals with PD from healthy ones.

8. References

[1] Little, M.A., McSharry, P.E., Roberts, S.J. et al. Exploiting Nonlinear Recurrence and Fractal Scaling Properties for Voice Disorder Detection. BioMed Eng OnLine 6, 23 (2007). https://doi.org/10.1186/1475-925X-6-23

[2] A. E. Lang and A. M. Lozano, "Parkinson's disease - First of two parts," New England Journal of Medicine, vol. 339, pp. 1044-1053, 1998.

[3] A. K. Ho, R. Iansek, C. Marigliani, J. L. Bradshaw, and S. Gates, "Speech impairment in a large sample of patients with Parkinson's disease," Behavioural Neurology, vol. 11, pp. 131-137, 1998.

[4] J. A. Logemann, H. B. Fisher, B. Boshes, and E. R. Blonsky, "Frequency and Co-Occurrence of Vocal-Tract Dysfunctions in Speech of a Large Sample of Parkinson Patients," Journal of Speech and Hearing Disorders, vol. 43, pp. 47-57, 1978.

[5] J. R. Duffy, Motor speech disorders: substrates, differential diagnosis, and management, 2nd ed. St. Louis, Mo.: Elsevier Mosby, 2005.

[6] S. Sapir, J. L. Spielman, L. O. Ramig, B. H. Story, and C. Fox, "Effects of Intensive Voice Treatment (the Lee Silverman Voice Treatment [LSVT]) on Vowel Articulation in Dysarthric Individuals With Idiopathic Parkinson Disease: Acoustic and Perceptual Findings," J Speech Lang Hear Res, vol. 50, pp. 899-912, 2007.

[7] D. A. Rahn, M. Chou, J. J. Jiang, and Y. Zhang, "Phonatory impairment in Parkinson's disease: Evidence from nonlinear dynamic analysis and perturbation analysis," Journal of Voice, vol. 21, pp. 64-71, 2007.

[8] M. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman and L. O. Ramig, "Suitability of Dysphonia Measurements for Telemonitoring of Parkinson's Disease," in IEEE Transactions on Biomedical Engineering, vol. 56, no. 4, pp. 1015-1022, April 2009.