



FACULTAD DE INGENIERÍA
BIOINGENIERÍA
Bioseñales y Sistemas

1. OBJETIVOS

General

Aplicar técnicas para el análisis de señales biomédicas en el dominio del tiempo mediante técnicas estadísticas.

Específicos

- Comprender las características que diferencian las señales determinísticas de las estocásticas.
- Implementar diferentes técnicas para el análisis estadístico de señales, e interpretar los resultados obtenidos.

2. MATERIALES Y EQUIPOS

Para la realización de esta práctica se requiere del paquete computacional de Python.

3. MARCO TEÓRICO

Naturaleza aleatoria o determinista de las señales

De forma amplia, las señales pueden clasificarse en uno de dos tipos: señales deterministas o estocásticas. Una señal determinista es aquella que puede reproducirse exactamente en diferentes momentos de registro. Por ejemplo, el voltaje resultante en un circuito RC cuando se activa una fuente conocida es una señal determinista. Ahora, tomemos como ejemplo la señal de EEG ante la aplicación de un estímulo visual: la presentación de una imagen durante un determinado período de tiempo. En este caso, hay muchas razones que nos llevan a imaginarnos que si se aplica repetidas veces el mismo estímulo, es probable que la respuesta sea diferente cada vez.

Una señal aleatoria o estocástica, es una señal que no se puede repetir de una manera predecible. Algunas señales pueden considerarse como deterministas o como aleatorias, dependiendo de la aplicación y las técnicas de análisis que se vayan a implementar, como se verá más adelante en el curso.

Por lo general, las señales biomédicas son de naturaleza estocástica. Para el estudio de las señales aleatorias, es necesario describir dichas señales en términos de probabilidades, o de su comportamiento promedio. Es por esto que los conceptos de probabilidad y estadística son de gran utilidad en el tratamiento de señales aleatorias.

Por otra parte, otra clasificación frecuente de señales se da respecto de la estacionariedad de estas. Lo cual se refiere a si las señales presentan características estadísticas constantes en el tiempo, es decir, si se evalúan ciertas características en un tiempo t , y luego pasado un tiempo τ se vuelve a observar la señal, las características no presentan diferencias estadísticas significativas. Estrictamente hablando, las bioseñales no son estacionarias, por ejemplo, un ECG en un momento dado puede presentar unos valores, y por cambios en el ambiente, el sujeto de quien se adquiere el ECG puede alterar su fisiología y esto cambiar la señal ECG. Sin embargo, la estacionariedad es un supuesto en muchos de los análisis, y se evalúa en términos probabilísticos y estadísticos, a través de la observación de características estadísticas como la media y la varianza, así como de la función autocorrelación, o por medio de estadísticos de prueba como el Dickey-Fuller.

4. PARA TENER EN CUENTA

Consulte el uso de las funciones en Python: mean, var, std.

5. PROCEDIMIENTO

Implementación de funciones (0.5)

- Implemente una función que permita calcular el valor cuadrático medio (RMS) de una señal. El valor RMS se define en forma discreta con N número de muestras, como:

$$x_{RMS} = \left[\frac{1}{N} \sum_{i=1}^N x(i)^2 \right]^{\frac{1}{2}}$$

Para este laboratorio, se le suministra el archivo de datos denominado *signals.mat*, el cual corresponde al registro de señales ECG y EMG.

Las señales se adquirieron a una tasa de muestreo de 1024 Hz. El registro correspondiente a las señales tal cual se registraron corresponde al campo *asRecording* de los archivos entregados. Las señales se sometieron a un proceso de filtrado el cual corresponde al campo *filtered*.

Análisis estadístico de ECG (2.0)

- A partir de la información suministrada en el archivo .mat, determine el tiempo de duración de las señales, cree un vector de tiempo.
- Grafique la señal sin filtrar y filtrada, compárelas. Comente lo que observa, y describa cómo debe ser el comportamiento del filtro que se utilizó. Recuerde las frecuencias de interés de la señal ECG.
- Tome la señal tal cual se registró. De dicha señal, seleccione un ciclo cardiaco, grafique. Calcule el promedio, el valor RMS, la varianza y la desviación estándar del ciclo. Analice los resultados.
- Tome la señal filtrada. Repita el procedimiento anterior con esta señal, compare con los resultados obtenidos en el ítem anterior.
- Extraiga 15 ciclos de la señal filtrada, y para cada ciclo calcule el valor promedio y la varianza, grafique cada ciclo, y analice (gráfica y estadísticamente). ¿Se puede considerar que la señal es estacionaria?
- Realice dos veces comparaciones entre dos ciclos cualquiera utilizando una prueba t. Determine estadísticamente si existe diferencia estadística entre los ciclos, Compruebe los supuestos necesarios para realizar una prueba t, esto es: Normalidad de la variable, independencia (se asume que los grupos son independientes), y homocedasticidad (use una prueba de Levene), finalmente realice la prueba t para determinar si existen diferencias entre los ciclos. Qué indica este análisis respecto de la estacionariedad, explique. De no cumplirse los requisitos, realice entonces un análisis no paramétrico (prueba U de Mann-Whitney).
- Compruebe los resultados anteriores aplicando una prueba de Dickey-Fuller a la señal (Estacionariedad)

Comparación estadística de Energía de señales EEG (2.5)

Se cuenta con señales de EEG de dos grupos de personas, un grupo control y un grupo de pacientes con enfermedad de Parkinson. Se tiene evidencia que la energía de las señales de EEG puede conducir a diferenciar entre señales de pacientes con enfermedad de Parkinson y sanos. Se pide calcular la Energía promedio por grupo poblacional de cada canal.

1. Implemente una función que reciba una señal de múltiples canales y épocas y calcule la Energía de promedio de cada canal.
2. Calcule la energía de cada canal promediada por épocas para cada sujeto, esto para ambos grupos poblacionales. Guarde esta información en un DataFrame de columnas 'canal' y filas '#sujeto' con los valores de energía calculados, un DataFrame para cada grupo poblacional.
3. Determine si existe diferencia estadística entre canales de cada grupo de sujetos a través de una prueba t. Compruebe los supuestos necesarios para realizar una prueba t, esto es: Normalidad de la variable, independencia (se asume que los grupos son independientes), y homocedasticidad (use una prueba de Levene), finalmente realice la prueba t para determinar si existen diferencias entre los canales entre grupos de sujetos. De no cumplirse los requisitos, realice entonces un análisis no paramétrico (prueba U de Mann-Whitney). Este numeral tiene como objetivo identificar los canales que entregan información diferencial entre pacientes Sanos y con enfermedad de Parkinson.

6. INFORME

Cada parte tiene su respectivo informe, por lo que se generan 2 archivo .ipynb donde presente la realización de las tareas descritas en el procedimiento, anexe análisis, conclusiones y referencias en el mismo archivo. Recuerde usar celdas de texto para los textos y comentarios solo para comentar el código. Separar cada literal o numeral de la guía.

7. ARCHIVOS ADJUNTOS O ANEXOS

Archivo

- signals.mat
- dataset/control
- dataset/parkinson

8. BIBLIOGRAFÍA

- Ejemplo prueba de Tukey <https://www.statology.org/tukey-test-python/>
- Ejemplos Test de normalidad https://github.com/thomas-haslwanter/statsintro_python/blob/master/ISP/Code_Quantlets/07_CheckNormality_CalcSampleSize/checkNormality/ISP_checkNormality.py
- Pruebas para homocedasticidad: <https://www.cienciadedatos.net/documentos/pystats07-test-homocedasticidad-heterocedasticidad-python.html>
- Prueba t: <https://www.cienciadedatos.net/documentos/pystats10-t-test-python.html>
- Prueba U Mann-Whitney: https://jmyao17.github.io/Statistics/Nonparametric_Statistical_Significance_Tests.html

9. LIBRERIAS DE INTERES PARA ESTADISTICA

- <https://docs.scipy.org/doc/scipy/reference/stats.html>
- <https://seaborn.pydata.org/>
- <https://pingouin-stats.org/build/htm>

Ejemplo de Dickey-Fuller

```
from statsmodels.tsa.stattools import adfuller

result = adfuller(np.squeeze(data))

print('ADF Statistic: %f' % result[0])
```

```
print('p-value: %f' % result[1])
```

ADF Statistic: -3.210042
p-value: 0.019411