

Resum

En aquest projecte té com a objectiu l'estudi de les interaccions que existeixen entre la variabilitat del ritme cardíac i la duració del cicle respiratori i la seva aplicació com a eina vàlida per a la presa de decisions en el moment de la desconnexió d'un pacient de la ventilació mecànica.

S'han estudiat aquestes dues sèries temporals, la variabilitat del ritme cardíac i la duració del cicle respiratori, provinents de 39 pacients sotmesos a la prova de tub en T, els quals havien rebut suport ventilatori durant un cert període de temps. Aquests pacients estaven classificats a priori en dos grups, els que havien superat la prova, i la desconnexió havia tingut èxit, i els que no.

S'ha buscat una tècnica que permetés trobar diferències entre les interaccions de les senyals d'aquests dos grups anteriors. Degut a l'elevada complexitat de les sèries temporals d'origen biològic, la tècnica d'anàlisi que s'ha utilitzat és la Dinàmica Simbòlica Conjunta. Aquesta tècnica s'ha adequat al tipus de senyals que es disposava i posteriorment s'ha realitzat un estudi estadístic que ha permès trobar un conjunt d'índexs capaços de detectar diferències bastant notables entre els dos grups. Amb aquest índexs s'han generat una sèrie de funcions d'una o dues variables de cara a poder classificar els pacients en un dels dos grups. Els resultats obtinguts han estat bastant satisfactoris, i en la majoria dels casos les funcions generades han classificat correctament els pacients dins el seu grup corresponent. Posteriorment s'han utilitzat aquests índexs per classificar pacients que no s'havien utilitzat en la generació de les funcions i els resultats també han estat bons.

Aquests resultats s'han complementat amb un estudi de la sèrie temporal cicle respiratori dels mateixos pacients també aplicant la tècnica de la Dinàmica Simbòlica a només un senyal. Tenint en compte aquests resultats la conclusió principal que s'extreu és que la Dinàmica Simbòlica Conjunta és una eina per a l'estudi de les interaccions entre les senyals del cicle respiratori i la variabilitat del ritme cardíac i que té un bon potencial a l'hora de cercar l'instant òptim en que un pacient ha d'esser desconnectat del suport ventilatori.





Sumari

| | |
|--|-----------|
| RESUM | 1 |
| SUMARI | 3 |
| GLOSSARI | 8 |
| 1. PREFACI | 9 |
| 1.2. Origen del projecte | 9 |
| 1.3. Motivació | 9 |
| 1.4. Requeriments previs | 9 |
| 2. INTRODUCCIÓ | 11 |
| 2.1. Objectius del projecte..... | 11 |
| 2.2. Abast del projecte..... | 11 |
| 3. VARIABILITAT DEL PATRÓ RESPIRATORI | 13 |
| 3.1. El sistema respiratori..... | 13 |
| 3.2. La ventilació mecànica..... | 13 |
| 3.3. Tècniques de mesura per al patró respiratori..... | 15 |
| 3.3.1. L'espirometria | 15 |
| 3.3.2. Pletismografia respiratòria per inductància..... | 15 |
| 3.3.3. La pneumotacografia..... | 17 |
| 3.4. Preprocessament del senyal pneumotacogràfic..... | 18 |
| 3.5. Base de dades utilitzada | 20 |
| 3.6. Complexitat del sistema respiratori..... | 20 |
| 4. LA VARIABILITAT DEL RITME CARDÍAC | 22 |
| 4.1. El sistema cardiovascular | 22 |
| 4.2. El senyal electrocardiogràfic | 22 |
| 4.3. Tècniques de mesura del patró cardíac | 24 |
| 4.4. Base de dades utilitzada | 26 |
| 4.5. Complexitat del sistema cardíac..... | 26 |
| 4.6. Interaccions entre el sistema respiratori i el cardíac..... | 27 |
| 5. ANÀLISI INICIAL DE LES DADES | 28 |
| 5.1. Introducció..... | 28 |
| 5.2. Característiques de la base de dades | 28 |
| 5.3. Divisió de les dades | 30 |



| | | |
|-----------|---|-----------|
| 5.3.1. | Grup d'aprenentatge i grup de test | 30 |
| 5.4. | Anàlisi preliminar de les sèries temporals | 31 |
| 5.4.1. | Test de significació. Verificació d'hipòtesis. | 31 |
| 5.4.2. | Mètodes paramètrics i mètodes no paramètrics | 32 |
| 5.4.3. | Proves paramètriques | 32 |
| 5.4.4. | Prova de Kolmogorov-Smirnov | 33 |
| 5.4.5. | Test de Levene..... | 36 |
| 5.4.6. | Proves no paramètriques | 37 |
| 6. | TÈCNiques D'ANÀLISI NO LINEAL | 40 |
| 6.1. | Introducció..... | 40 |
| 6.2. | Sistemes Dinàmics | 40 |
| 6.2.1. | Sistemes dinàmics no lineals..... | 43 |
| 6.3. | Teoria del caos..... | 43 |
| 6.4. | Complexitat | 46 |
| 6.5. | Tècniques d'anàlisi no lineal | 47 |
| 6.5.1. | Dimensió..... | 48 |
| 6.5.2. | L'exponent de Hurst | 49 |
| 6.5.3. | Els exponents de Lyapunov..... | 50 |
| 6.5.4. | Entropia | 50 |
| 6.5.5. | Dinàmica Simbòlica..... | 51 |
| 7. | DINÀMICA SIMBÒLICA I ANÀLISI DE LA INTERACCIÓ CARDIORESPIRATÒRIA | 53 |
| 7.1. | Introducció..... | 53 |
| 7.2. | Fonament teòric de la Dinàmica Simbòlica | 53 |
| 7.2.1. | Dinàmica Simbòlica Conjunta | 55 |
| 7.3. | Metodologia de la Dinàmica Simbòlica..... | 56 |
| 7.4. | Paràmetres necessaris per aplicar la Dinàmica Simbòlica | 57 |
| 7.5. | Descripció de l'algorisme | 58 |
| 7.6. | Exemple d'aplicació. | 60 |
| 8. | CÀLCUL DELS PARÀMETRES | 63 |
| 8.1. | Introducció..... | 63 |
| 8.2. | Metodologia per a la determinació dels paràmetres òptims..... | 63 |
| 8.3. | Tècniques estadístiques | 64 |
| 8.3.1. | Realització de les proves d'hipòtesi..... | 64 |
| 8.4. | Determinació del paràmetres òptims mitjançant la prova U de Mann-Whitney | 65 |



| | | |
|------------|---|-----------|
| 8.4.1. | Consideracions prèvies | 65 |
| 8.4.2. | Resultats obtinguts | 66 |
| 9. | ANÀLISI DISCRIMINANT | 69 |
| 9.1. | Introducció | 69 |
| 9.2. | Fonament teòric de l'Anàlisi Discriminant | 70 |
| 9.3. | Aplicació de l'Anàlisi Discriminant | 71 |
| 9.3.1. | Anàlisi de les variables obtingudes utilitzant Dinàmica | 71 |
| | Simbòlica Conjunta | 71 |
| 9.3.2. | Anàlisi de les variables obtingudes amb l'estudi lineal | 76 |
| 10. | ESTUDI DE LA VARIABILITAT DEL PATRÓ RESPIRATORI | 77 |
| 10.1. | Introducció | 77 |
| 10.2. | Mètodes | 77 |
| 10.3. | Alfabet Θ | 78 |
| 10.4. | Alfabet Ω | 80 |
| 11. | CONCLUSIONS I EXTENSIONS FUTURES | 83 |
| | BIBLIOGRAFIA | 86 |
| 11.1. | Referències bibliogràfiques | 86 |
| 11.2. | Bibliografia complementària | 88 |

ANNEXES

| | | |
|-----------|--------------------------------------|----------|
| A. | PRESSUPOST | 1 |
| A.1. | Introducció | 1 |
| A.2. | Costos de Personal | 1 |
| A.3. | Costos dels equips informàtics | 2 |
| A.4. | Costos varis | 4 |
| B. | PROGRAMES | 5 |
| B.1. | DinàmicSimbolicaConjunta.m | 5 |
| B.1.1. | DinàmicSimbolicaConjunta.m | 5 |
| B.1.2. | Funció obtenirdades.m | 9 |



| | | |
|---------|--|----|
| B.1.3. | Funció codificar.m | 11 |
| B.1.4. | Funció mcont.m | 11 |
| B.1.5. | Funció grafiquesfinals.m | 12 |
| B.1.6. | Funció treuredades.m | 13 |
| B.2. | DinamicaSimbolica.m | 16 |
| B.2.1. | DinamicaSimbolica.m | 17 |
| B.2.2. | Funció obtenirdades | 19 |
| B.2.3. | Funció codificar.m (alfabet Θ) | 19 |
| B.2.4. | Funció codificar.m (alfabet Ω) | 20 |
| B.2.5. | Funció mcont.m (alfabet Θ) | 20 |
| B.2.6. | Funció mcont.m (alfabet Ω) | 23 |
| B.2.7. | Funció grafiquesfinals.m (alfabet Θ) | 28 |
| B.2.8. | Funció grafiquesfinals.m (alfabet Ω) | 28 |
| B.2.9. | Funció treuredades.m (alfabet Θ) | 29 |
| B.2.10. | Funció treuredades.m (alfabet Ω) | 30 |
| B.3. | Estadístics.m | 32 |
| B.2.1. | Estadístics.m | 32 |
| B.2.2. | Funció obtenirdades.m | 34 |

C. TAULES RESULTATS ESTUDI VARIABLES LINEALS I DINÀMICA SIMBÒLICA 37

| | | |
|--------|--|----|
| C.1. | Resultats Estudi Variables Lineals | 37 |
| C.1.1. | Resultats càlcul Variables Lineals | 37 |
| C.1.2. | Resultats U de Mann-Whitney i la Prova de Kolmogorov-Smirnov | 39 |
| C.1.3. | Resultats Anàlisi Discriminant | 40 |
| C.1. | Taules de Resultats Dinàmica Simbòlica | 41 |
| C.1.1. | Resultats aplicació Dinàmica Simbòlica | 41 |
| C.1.2. | Resultats U de Mann-Whitney | 71 |
| C.1.3. | Resultats Anàlisi Discriminant | 82 |

D. TAULES RESULTATS DINÀMICA SIMBÒLICA CONJUNTA 89

| | | |
|------|---|-----|
| D.1. | Resultats aplicació Dinàmica Simbòlica Conjunta | 89 |
| D.2. | Resultats U de Mann-Whitney | 126 |
| D.3. | Resultats Anàlisi Discriminant | 139 |





Glossari

| | |
|--------------|---|
| c000 | Paraula pertanyent a la sèrie temporal RR(n) |
| r000 | Paraula pertanyent a la sèrie temporal Ttot(n) |
| P(c000) | Probabilitat d'aparició d'una paraula pertanyent a la sèrie temporal RR(n) |
| P(r000) | Probabilitat d'aparició d'una paraula pertanyent a la sèrie temporal Ttot(n) |
| P(c000,r000) | Probabilitat d'aparició d'una paraula de la sèrie temporal RR(n) en combinació amb una paraula de la sèrie temporal Ttot |
| WcX | Nombre de paraules de la sèrie temporal RR(n) amb una probabilitat d'aparèixer major o igual que X (X=5%, 10%, 15%, 20% i 25%) |
| WrX | Nombre de paraules de la sèrie temporal Ttot(n) amb una probabilitat d'aparèixer major o igual que X (X=5%, 10%, 15%, 20% i 25%, a l'estudi amb Dinàmica Simbòlica Conjunta i X= 1%, 2%, 5%, 10%, 15%, 20% i 25% en l'estudi de la Duració del Cicle Respiratori) |
| Wcr1 | Nombre de paraules de la sèrie temporal RR(n) que apareixen en combinació amb una certa paraula de la sèrie temporal Ttot(n) amb una probabilitat d'aparèixer major o igual que X (1%, 2%, 3%, 4% i 5%) |
| fWc | Nombre de paraules prohibides de la sèrie temporal RR(n), és a dir, aquelles que tenen una probabilitat d'aparèixer menor o igual que 1%. |
| fWr | Paraules prohibides de la sèrie temporal Ttot(n), és a dir, aquelles que tenen una probabilitat d'aparèixer menor o igual que 1% a l'estudi amb Dinàmica Simbòlica Conjunta un 0,01% en l'estudi de la Duració del Cicle Respiratori. |
| fWcr | Nombre de paraules prohibides de la sèrie temporal RR(n) que apareixen en combinació amb una certa paraula de la sèrie temporal Ttot(n), és a dir aquelles que tenen una probabilitat d'aparèixer menor o igual que un 0,1%. |
| SHwc | Entropia de Shannon per als índexs de la sèrie temporal RR(n). |
| SHwr | Entropia de Shannon per als índexs de la sèrie temporal Ttot(n). |
| SHwcr | Entropia de Shannon per als índexs de la sèrie temporal RR(n) combinades amb la sèrie temporal Ttot(n). |



1. Prefaci

1.2. Origen del projecte

L'origen d'aquest projecte és troba en els estudis que es realitzen al Departament de "Enginyeria de Sistemes, Automàtica i Informàtica Industrial" (ESAI), a la Divisió de "Senyals i Sistemes Biomèdics" del centre de recerca en Enginyeria Biomèdica (CREB) de la UPC.

1.3. Motivació

La ventilació mecànica és una de les intervencions més importants que es realitzen en les Unitats de Cures Intensives, essent el moment de la seva retirada uns dels instants crítics de determinar i que pot tenir greus conseqüències. Si és desconnectada prematurament pot ocasionar problemes cardiorespiratoris, problemes deguts a la reintubació i psicològics greus, mentre que el seu ús prolongat pot tenir un risc elevat d'infeccions i debilitament dels músculs pulmonars.

La investigació es centra actualment en la recerca d'un mètode que permeti determinar el millor instant per a interrompre la ventilació mecànica a un pacient. Es pretén trobar un conjunt d'índexs que avaluïn les diferents senyals que es mesuren per a predir el millor instant a partir del qual un pacient pot mantenir la respiració espontània per un temps indefinit.

1.4. Requeriments previs

Per a poder realitzar aquest estudi ha estat necessària una base de dades de diverses sèries temporals de pacients sotmesos a la prova de tub en T. Els senyals han estat enregistrats a l'Hospital de la Santa Creu i Sant Pau de Barcelona i a l'Hospital Universitario de Getafe.



2. Introducció

2.1. Objectius del projecte

L'objectiu d'aquest projecte ha estat l'estudi de les interaccions que existeixen entre la variabilitat del ritme cardíac i la duració del cicle respiratori utilitzant la tècnica de la Dinàmica Simbòlica Conjunta. Per a realitzar-lo s'han utilitzat senyals enregistrats a pacients sotmesos a la prova de tub en T per a la seva desconexió de la ventilació mecànica després d'estar durant un cert temps connectats a ella. La tècnica utilitzada per a l'adquisició del senyal de duració del cicle respiratori ha estat la pneumotacografia i el senyal de variabilitat del ritme cardíac s'ha obtingut a partir de l'electrocardiograma.

Amb l'estudi es pretén observar si a l'aplicar la tècnica de la Dinàmica Simbòlica Conjunta a aquestes dues sèries temporals es pot discriminar correctament entre el grup de pacients que han superat amb èxit la prova de tub en T, i la desconexió del suport ventilatori ha estat possible, i aquells en que ha fracassat. És planteja la hipòtesi, confirmada amb altres senyals biomèdics que una baixa variabilitat a la sèrie temporal indica un nivell de patologia.

2.2. Abast del projecte

En d'aquest projecte es realitza un estudi de senyals provinents de pacients per a trobar un mètode que permeti determinar el moment òptim de desconexió de pacients sotmesos a ventilació mecànica. El seu abast, si el mètode resulta vàlid, és la seva aplicació en el tractament d'aquests pacients per determinar el moment de retirar el suport respiratori.



3. Variabilitat del patró respiratori

3.1. El sistema respiratori

El sistema respiratori és un dels quatre sistemes majors del cos humà conjuntament amb el cardiovascular, el renal i el gastrointestinal [Mompín, 1988]. El sistema respiratori interacciona amb l'entorn i la seva principal funció és subministrar oxigen i eliminar diòxid de carboni per a permetre el manteniment de les cèl·lules.

La musculatura respiratòria actua com a generador de pressió i provoca l'expansió i contracció cíclica de la caixa toràcico-abdominal i dels pulmons, per fer circular els gasos a través de les vies aèries. En el procés respiratori apareixen dues fases diferenciades: inspiració i expiració. La fase on s'agafa aire ric en oxigen del medi és la inspiració. Aquest entra en la cambra pulmonar, que anirà guanyant volum a mesura que es va ingressant aire. Els mecanismes que actuen per dur a terme amb èxit aquest procés consisteixen en un moviment cap avall del diafragma (principal moviment responsable de la respiració) i una elevació de la zona costal. L'expiració és la fase on s'expulsa el CO_2 ja que el pulmó recuperarà el seu volum inicial degut a una elevació del diafragma i a un descens de la posició ventral de les costelles.

Existeixen diversos paràmetres per estudiar el sistema respiratori. Com a característiques de la respiració es solen tractar volums i pressions en funció del temps. Dins el volum hi ha el volum abdominal, el volum toràcic, el volum total i el flux total (derivada del volum total respecte el temps). Com a pressions hi ha la pressió alveolar, la pressió a la cavitat pleural, la pressió a les vies aèries i la pressió transpulmonar. A partir d'aquestes variables es poden calcular diversos paràmetres. Per exemple, a partir dels senyals de la variació de volum en el temps es poden obtenir factors com el volum circulant (volum que intervé en cada cicle) i la freqüència respiratòria (quantitat de cicles per unitat de temps). En funció del cas a investigar s'utilitzen els paràmetres que poden oferir més informació.

3.2. La ventilació mecànica

L'ús de la ventilació mecànica pot ser vital quan un pacient té greus problemes pulmonars o cardíacs, però el seu ús prolongat suposa un risc elevat degut a problemes d'infeccions i debilitament dels músculs pulmonars, així com un gran cost econòmic. Per aquestes motius és important poder detectar l'instant a partir del qual un pacient ja no necessitarà d'aquest suport ventilatori per subsistir i es podrà realitzar l'extubació amb garanties. D'altra banda,



s'ha d'evitar una retirada prematura perquè podria ocasionar una problemes respiratoris cardiorespiratoris i psicològica greus deguts a la reintubació.

Encara que un metge a partir de la seva experiència pot predir amb un gran percentatge d'encert els resultats dels intents d'extubació, actualment no existeix cap mètode objectiu que permeti decidir el moment exacte per interrompre el suport ventilatori. Per aquesta raó, s'està treballant per trobar uns índexs objectius que disminueixin la dependència de l'habilitat o coneixement del metge individual.

Els estudis que s'estan realitzant sobre aquest mètode cerquen poder predir si una extubació tindrà èxit o no, i per tant, identificar el primer moment en que un pacient serà capaç de refer i mantenir la respiració espontàniament sense ajuda.

Per a dur a terme aquest estudi, existeixen una sèrie de variables que tradicionalment s'usen per predir l'èxit o fracàs de la desconexió de la ventilació mecànica [Net i Benito, 1990]. Els paràmetres són els següents:

- Capacitat vital major a 10 - 15 ml/kg de pes corporal.
- Pressió respiratòria negativa menor a -30 cm de columna d'aigua.
- Ventilació per minut menor a 10 l/min.
- Ventilació voluntària màxima superior al doble de la ventilació en repòs.
- Pressió alveolar d'oxigen superior a 60 torr.
- Concentració inspirada d'oxigen inferior a 0,35.
- Gradient alveolar-arterial de la pressió d'oxigen menor a 350 torr.
- Quocient de pressió alveolar i concentració d'oxigen major a 200 torr.

Un dels objectius actuals és intentar poder determinar nous paràmetres com aquests amb tècniques no invasives. En aquest sentit la recerca clínica procura seguir dos camins per a determinar l'èxit de l'extubació: determinar nous paràmetres que puguin aportar major informació i fer estudis més profunds amb les paràmetres més típics com el temps d'inspiració, el temps total d'un cicle respiratori, el volum circulant, entre altres



3.3. Tècniques de mesura per al patró respiratori

Existeixen força diversitat de tècniques per a l'obtenció de variables relacionades amb el patró respiratori. Si l'estudi es centra en senyals respiratoris referits al volum (V) o al flux d'aire ($\phi=dV/dt$) que intervé en la respiració, existeixen tres tècniques: la pneumotacografia, l'espirometria i la pletismografia. Totes elles són bastant utilitzades i no n'hi ha cap que la seva utilització predomini damunt les altres. A continuació es farà una breu introducció a cadascuna d'elles.

3.3.1. L'espirometria

L'espirometria és una tècnica molt antiga que serveix per a realitzar mesures de volum pulmonar. L'any 1848 ja va ser dissenyat un espiròmetre d'aigua per Hutchinson [Mompín, 1988].

L'espiròmetre és un sensor amb un compartiment expansible que té un volum que provoca una pressió interior igual a l'atmosfèrica. Hi ha dos tipus d'espiròmetre: l'espiròmetre d'aigua i l'espiròmetre sec.

L'espiròmetre d'aigua convencional consisteix en una campana cilíndrica amb parets dobles plenes d'aigua. Aquesta campana es desplaça verticalment fins que la pressió al seu interior s'iguali a l'atmosfèrica. L'alçada de la campana és proporcional al volum expirat. Per tant, s'ha passat d'una magnitud de volum a una magnitud d'alçada que es pot representar gràficament en un paper posant una ploma a la campana o que pot passar a senyal elèctric a partir del sensor d'alçada desitjat, per exemple un potenciòmetre. Per tal de que no s'acumuli diòxid de carboni s'incorpora un circuit doble de tubs. En un d'ells, el que pertany a l'expiració, s'hi posa un absorbidor de diòxid i es torna a omplir el volum extret amb oxigen.

Aquest aparell té molta inèrcia i per tant la resposta en freqüències és reduïda. Tot i això, es pot aconseguir que els espiròmetres d'aigua tinguin un error inferior al 2%.

Pel que fa als espiròmetres secs, són molt més petits i manejables que els d'aigua. El fonament és el mateix, mantenir una cavitat a pressió constant per mesurar el volum a partir de la posició d'un pistó. En general els espiròmetres secs són dinàmicament una mica més ràpids, però una mica més inexactes que els d'aigua.

3.3.2. Pletismografia respiratòria per inductància

La pletismografia respiratòria per inductància [Net i Benito, 1990] és una tècnica per a determinar el volum d'aire toràcic i abdominal dels individus. La mesura es fa mitjançant



un transductor inductiu i es col·loquen unes bandes situant a la superfície de la pell de la zona toràcica i abdominal.

Aquesta tècnica s'utilitza des del 1967 i és la més recent de les tres. La principal característica de la pletismografia per inductància és que és la tècnica menys invasiva en comparació a l'espirometria i a la pneumotacografia. Els resultats que s'obtenen de la seva aplicació són força precisos en pacients intubats i a més no necessiten de la col·laboració del pacient, que pot estar adormit durant l'adquisició.

El sensor emprat consta de dues bandes elàstiques amb espines al seu interior que es col·loquen al voltant de l'abdomen i de la caixa toràcica. Un oscil·lador produeix una ona sinusoidal d'amplitud aproximadament 20 mV, a una freqüència de 30 kHz cap a cada banda d'espines. Una modificació del volum toràcic o abdominal varia la inductància del circuit. Aquestes espines actuen com a bobina d'un circuit ressonant, així en canviar el valor de la inductància es modifica la freqüència del circuit oscil·lant. Si s'anomena L a la inductància de les espines i C a la capacitat del circuit, es pot trobar el valor de la freqüència mitjançant l'equació 3.1.

$$f_0 = \frac{1}{\pi} \sqrt{\frac{1}{L \cdot C}} \quad (\text{Eq. 3.1})$$

L'aparell de mesura de la Figura 3.1 agafa el senyal que li arriba i el tracta seguint una sèrie de passos: primer el demodula, amplifica la variació de freqüència i el converteix en voltatge. A continuació es produeix un processament per obtenir tres senyals: volum toràcic, volum abdominal i volum total.

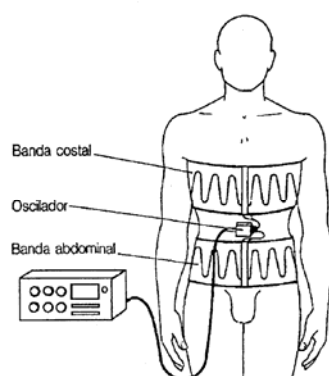


Figura 3.1 Diagrama de l'aplicació del pletismògraf per inductància a un pacient



Es considera acceptable [Net i Benito, 1990] un error de fins un 10% en les mesures que realitza la pletismografia. Aquest nivell d'error és degut a les dificultats en calibrar l'aportació dels moviments toràcics i abdominal sobre el volum total.

3.3.3. La pneumotacografia

La pneumotacografia [Mompín, 1988] és una tècnica orientada a mesurar el flux d'aire que intervé en la respiració. A partir del flux es pot obtenir el volum mitjançant una integració matemàtica discreta ja que $\phi = dV/dt$.

El seu principi de funcionament es basa en col·locar una resistència pneumàtica al lloc on es vol mesurar el flux, de manera que provoqui una caiguda de pressió proporcional al flux que passi per ella. A l'equació 3.2 es mostra la relació proporcional entre flux i pressió.

$$\Delta P = K \cdot \phi \quad (\text{Eq. 3.2})$$

ΔP = variació de la pressió

K = constant que depèn de la geometria del tub i de la viscositat

ϕ = flux

A la realitat, aquesta proporcionalitat és vàlida per a fluxos baixos, però perd exactitud a mesura que augmenta el flux tendint més a una relació de segon grau. No obstant, mitjançant un disseny optimitzat i un circuit electrònic de linealització, es pot aconseguir que els pneumotacògrafs siguin molt precisos i cometin un error de menys del 4% en mesures de fins a 15 l/s. Els pneumotacògrafs poden efectuar mesures tant en la inspiració com en l'expiració.

Existeixen dos tipus de pneumotacògrafs segons com estigui formada la resistència. Si està formada per un conjunt de malles superposades s'anomenen pneumotacògrafs tipus Lilly, i si està formada per un feix de tubs capil·lars d' 1 o 2 mm de diàmetre cadascun i uns 4cm de longitud, s'anomenen pneumotacògrafs tipus Fleisch.

Una de les característiques principals de la pneumotacografia és que al posar la resistència pneumàtica es pot alterar el valor que tindria el flux sense aquesta. Per aquest motiu, cal que la resistència sigui el més baixa possible i s'exigeix que sigui menor que 0,1 Kpa·s/l. Els pneumotacògrafs porten incorporada una petita resistència elèctrica calefactora perquè el vapor d'aigua present en l'expiració no es condensi en els capil·lars o en les malles. Aquest fet falsejaria el resultat i l'aparell marcaria un valor més alt que l'autèntic.



La pneumotacografia ha estat la tècnica emprada per adquirir els senyals de flux respiratori que seran objecte d'estudi en aquest projecte. L'espirometria no s'ha utilitzar ja que interessa un sistema de mesura amb bona resposta freqüencial i la pletismografia no s'ha utilitzat degut a les dificultats per realitzar una bona calibració.

L'aparell de mesura es mostra a la Figura 3.3 i la Figura 3.2 mostra com s'aplica a un pacient.



Figura 3.3 Pneumotacògraf

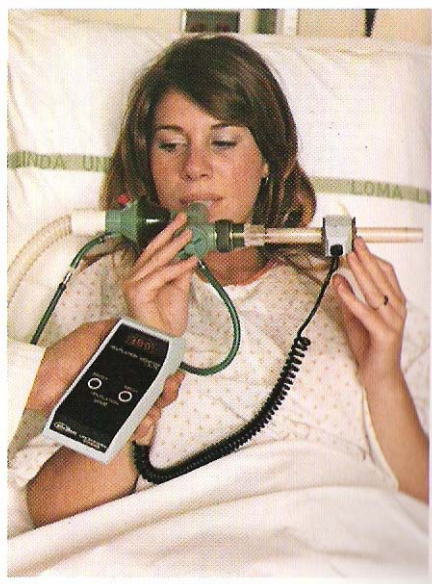


Figura 3.2 Aplicació del pneumotacògraf a un pacient

3.4. Preprocessament del senyal pneumotacogràfic

El preprocessament del senyal de flux consistirà en aplicar un algorisme [Casals, 1999] que detecti el conjunt de punts corresponents als inicis de la inspiració i de l'expiració del pacient. D'aquesta manera, un cop han estat detectades les marques dels cicles respiratoris, es determinaran els valors dels paràmetres a calcular. Aquests paràmetres són els següents:

- T_{ti} : temps d'inspiració
- T_{te} : temps d'expiració
- $T_{tot} = T_{ti} + T_{te}$: duració del cicle respiratori



- f : freqüència respiratòria en cicles/minut
- V_{es} : volum circulant en cada cicle respiratori
- $\frac{T_{ti}}{T_{tot}}$: fracció inspiratòria del temps total del cicle respiratori
- $\frac{V_{es}}{T_{ti}}$: fracció del volum específic del temps d'inspiració

A la Figura 3.4 queden esquematitzades aquest components de manera gràfica, T_{tot} , T_{ti} , T_{te} i V_{es} .

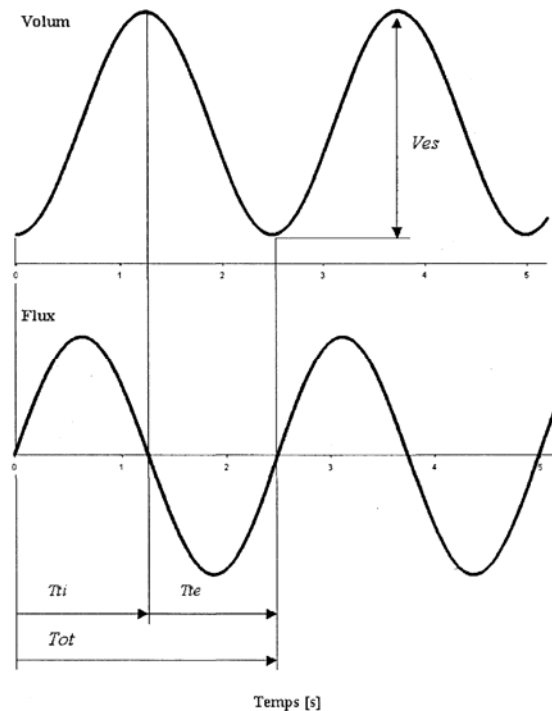


Figura 3.4 Representació gràfica de les magnituds T_{tot} , T_{ti} , T_{te} , V_{es}

Els resultats d'aquesta etapa de preprocessament del senyal de flux, per a cada pacient, és el conjunt de sèries temporals següent: $T_{tot}(n)$, $T_{ti}(n)$, $T_{te}(n)$, $V_{es}(n)$, $T_{ti}/T_{tot}(n)$ i $V_{es}/T_{ti}(n)$. Degut a estudis anteriors [Caminal et al., 2004] s'han obtingut els millors resultats a partir de la sèrie temporal $T_{tot}(n)$, en aquest projecte s'avaluaran les sèries temporals $T_{tot}(n)$ dels pacients estudiats.



3.5. Base de dades utilitzada

La base de dades respiratòries utilitzada en aquest projecte està formada per senyals de flux pulmonar mesurats amb pneumotacografia i adquirits a l'Hospital de la Santa Creu i Sant Pau de Barcelona (HSP) i a l'Hospital Universitario de Getafe (HUG). La freqüència de mostreig dels senyals adquirits és de 250 Hz. La base de dades ha comptat amb 39 pacients en procés de weaning o procés de desconexió del respirador que han estat un temps connectats a la ventilació mecànica. En aquest treball s'estudiaran els senyals respiratoris mesurats en aquests pacients sotmesos a la prova de tub en T per a la desconexió del suport ventilatori durant 30 minuts. Després els pacients han estat classificats en dos grups:

- Els pacients que han superat la prova de tub en T (C1)
- Els pacients que no han superat la prova de tub en T (C2)

S'ha considerat que un pacients no ha superat la prova de tub en T quan el pacient no podia mantenir una respiració espontània durant 30 minuts i era necessari que continués connectat el pacient a la respiració mecànica.

La prova de tub en T és una prova senzilla que consisteix bàsicament en desconnectar al pacient del respirador y deixar-lo respirar espontàniament a través del tub endotraqueal.

3.6. Complexitat del sistema respiratori

Les possibles causes de la variabilitat del patró respiratori han estat discutides molt recentment. Aquesta variabilitat pot ser deguda o bé a processos estocàstics (pertorbacions aleatòries) o a comportaments dinàmics dels reflexos químics i mecànics, i fins i tot, a la interacció entre ambdues causes [Bruce i Danbenspeck, 1995].

La qüestió fonamental és poder determinar fins a quin punt un canvi en l'estat o en les condicions del pacient poden alterar la variabilitat del patró respiratori.

Un estudi recent amb ratolins ha demostrat que la respiració espontània és més variable quan els ratolins han estat exposats a pressions transrespiratòries negatives que causen una reducció en el volum final d'expiració. D'altra banda, la variabilitat del model respiratori es redueix ràpidament quan s'exposa el ratolí a pressions transitàries positives.



Una hipòtesi que es formula és que el comportament no lineal i els retards de temps, propis dels mecanismes respiratoris del sistema de control neuronal, conjuntament amb l'activitat muscular i el mecanisme motriu del pulmó, podrien introduir una variabilitat no estocàstica en el model respiratori. Conseqüentment, el model respiratori seria molt més variable amb pressió negativa que quan està sotmès a pressions positives. Aquesta és una mostra de com poden arribar a influir les condicions i l'estat del pacient en la variabilitat del seu patró respiratori.

Quan s'estudia la variabilitat del patró respiratori, la primera consideració és determinar els índexs que es poden mesurar per extreure la informació oculta. S'han efectuat diversos anàlisi per classificar els patrons de variabilitat respiratòria i aquests, sovint han evidenciat la repetició de correlacions locals en les dades.

Tècniques com els estimadors de dimensió fractals permetrien descriure millor l'estructura dinàmica del comportament respiratori, i també comprendre millor el control de la respiració. Desafortunadament, aquests primers estudis tenen moltes limitacions. Molts d'ells han emprat l'algorisme de Grassberger i Procaccia per estimar dimensions fractals, el qual és molt simple i senzill d'implementar, però recentment s'ha conegut que la seva utilització pot fer perdre informació. El principal problema apareix quan es tenen poques dades o quan el sistema incorpora molt soroll [Carvajal, 1999]. Altres mètodes per analitzar els senyals respiratoris, com l'exponent de Lyapunov, han permès arribar a la conclusió que el sistema respiratori té un comportament caòtic. No obstant, aquesta simple aproximació no distingiria un sistema dinàmic no lineal d'un soroll filtrat linealment [Thelie et al., 1992].



4. La variabilitat del ritme cardíac

4.1. El sistema cardiovascular

El sistema cardiovascular és un dels quatre sistemes majors del cos humà conjuntament amb el respiratori, el renal i el gastrointestinal [Mompín, 1988]. La missió del sistema cardiovascular o circulatori és la circulació de la sang per portar a cada cèl·lula de l'organisme les substàncies que necessita per al seu metabolisme, transportar els productes metabòlics de desguàs, calor, entre altres.

El cor es pot descriure com una bomba dotada de vàlvules, que impulsa la sang a través de l'aparell circulatori [Bartels i Bartels, 1983]. El cor es un òrgan buit, està format per fibres musculars especialment unides entre si formant una espècie de xarxa. La contracció, sístole, del múscul serveix per impulsar sang a través de l'aparell circulatori. Es possible determinar la direcció de la corrent sanguínia gracies a les vàlvules del cor. Les vàlvules formen unes rebaves entre les aurícules i els ventricles per impedir el retrocés de la sang durant la contracció ventricular. Les vàlvules sigmoidees impedeixen el retrocés de la sang des de les grans artèries en la relaxació, diàstole, dels ventricles.

El cor té el seu propi sistema de conducció de l'excitació. Les excitacions parteixen d'un marcapassos, el nòdul sinusal, i s'estenen sobre la musculatura auricular cap el node auriculoventricular. Des d'allí, les excitacions són conduïdes a través de les dos branques del feix de His cap al vèrtex del cor.

4.2. El senyal electrocardiogràfic

En el senyal electrocardiogràfic, al ser l'activitat de les càmeres cardíques rítmica i totalment coordinada, es genera una ona de forma totalment regular. En aquesta ona es poden reconèixer diverses ones, les amplituds, duracions i morfologia estan correctament definides, tal com es mostra a la Figura 4.1, [Pallás, 1988]



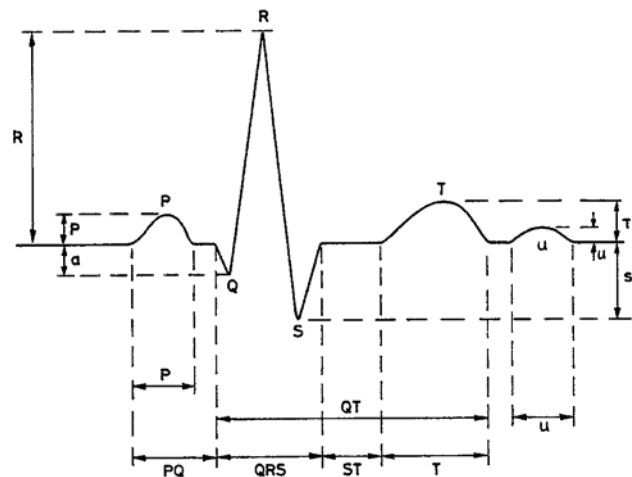


Figura 4.1 Ones característiques a l'ECG

L'electrocardiograma mostra primer una ona positiva, P, que expressa l'excitació de l'aurícula, la segueix una ona negativa, Q, que apareix principalment per l'activació del múscul papil·lar. L'ona R expressa l'excitació que es propaga per el múscul cardíac en direcció cap a la punta del cor. L'ona S indica la penetració de l'excitació a través de les parets ventriculars, i l'ona T, la recuperació de l'excitació.

La Variabilitat del Ritme Cardíac (HRV, Heart Rate Variability) està definida com la variació del temps que transcorre entre intervals RR consecutius del electrocardiograma, **¡Error! No se encuentra el origen de la referencia.** [Task Force, 1996]. Tal com s'il·lustra a la Figura 4.2. En aquest projecte s'estudiarà la variabilitat del ritme cardíac, o senyal RR(n) format pels valors RR al llarg de l'electrocardiograma.

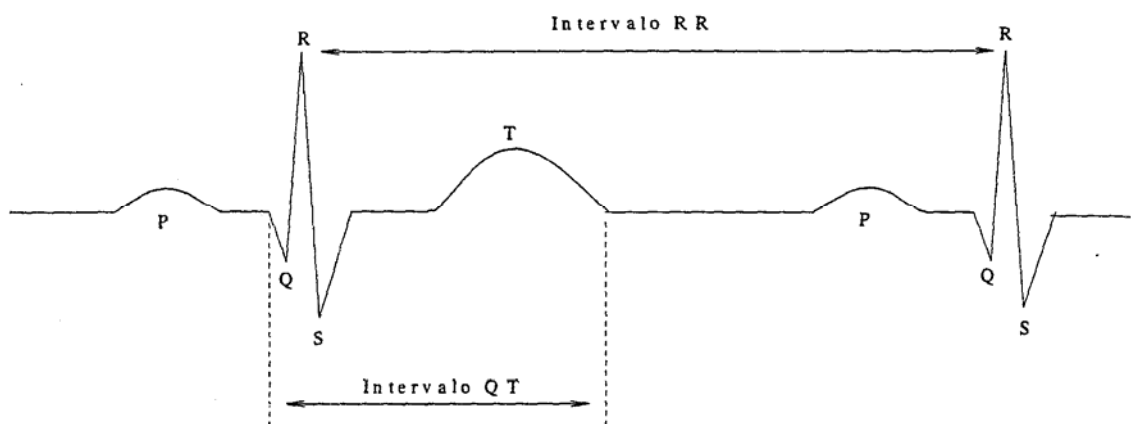


Figura 4.2 Interval RR



4.3. Tècniques de mesura del patró cardíac

Per mesurar els senyals del sistema cardiovascular es realitza un electrocardiograma (ECG). Un electrocardiograma és l'enregistrament de l'activitat elèctrica del cor mesurada normalment entre dos punts de la superfície del cos.

Les corrents iòniques, degudes als potencials d'acció associats a l'activitat cardíaca, circulen pel tòrax produint una distribució de potencials superficials que, en primera aproximació, és similar a la deguda a un dipol de corrent situat a l'hipotètic "centre elèctric" del cor. La direcció i magnitud d'aquest dipol va canviant al llarg del cicle cardíac. El seu moment dipolar constitueix l'anomenat vector cardíac.

El senyal obtingut en la mesura de l'ECG depèn de la localització dels elèctrodes, per aquest motiu s'ha normalitzat. Cada parell d'elèctrodes o combinacions entre ells es denomina "derivació", i el potencial obtingut entre ells és la projecció del vector cardíac en la direcció que defineixen. Encara que per poder determinar un vector és suficient en conèixer la seva projecció en tres direccions diferents, en l'electrocardiograma es sol recórrer a la mesura de 12 derivacions en els plans frontal i transversal, i així millorar el diagnòstic.

Les tres primeres derivacions són les anomenades estàndard, o bipolars, I, II, III. Van ser introduïdes per Einthoven el 1912, basant-se en un model que considera el cor situat al centre d'un triangle equilàter en un pla frontal, i una conducció elèctrica uniforme en totes les direccions. Els elèctrodes es situen un a cada canell i l'altre a la cama esquerra (que està aproximament en la vertical del cor), o en punts equivalents. La polaritat és la indicada a la Figura 4.3. Aquestes derivacions, en un cas normal, la ona R es positiva i té la màxima amplitud a la derivació II. En qualsevol cas, a partir de la relació geomètrica entre derivacions es dedueix que els potencials mesurats compleixen l'equació 4.1:

$$V_I + V_{III} = V_{II} \quad (\text{Eq. 4.1})$$

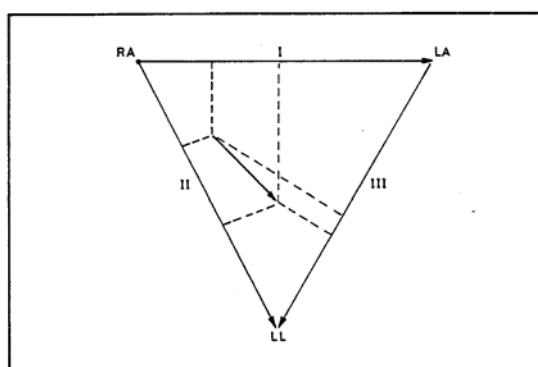


Figura 4.3 Direccions convencionals de les derivacions bipolar en ECG



Les connexions dels tres elèctrodes anteriors mitjançant resistències respectives a un punt comú, permeten obtenir en aquest un promig dels tres potencials, i constitueixen l'anomenat terminal de Wilson. La mesura de cadascun dels elèctrodes i aquest terminal formen les anomenades derivacions unipolars a les extremitats, VL (braç esquerre), VR (braç dret) i VF (cama esquerra). Si la mesura es realitza entre un elèctrode i el promig dels altres dos, s'obté aleshores derivacions unipolars augmentades, aVL, aVR i aVF que són tres de les més utilitzades habitualment. Les direccions que defineixen estan intercalades entre les de les derivacions estàndard. Per això, obtenint aquestes tres derivacions simultàniament es pot determinar fàcilment la direcció del vector cardíac en el pla frontal, a partir de veure quina d'elles dóna la major amplitud.

Les sis derivacions restants són també unipolars i es mesuren entre cada un de sis punts del pit ben definits, i el terminal central de Wilson. Es designen com a derivacions precordials, V₁, V₂, ..., V₆. Aquestes derivacions donen informació en el pla transversal. Algunes vegades, en exploracions molt detallades s'utilitzen altres derivacions.

L'ECG s'utilitza com a senyal per al diagnòstic, com a referència temporal en l'enregistrament d'altres senyals i per a al monitorització.

Els paràmetres més considerats per a analitzar el senyal, normalment de forma empírica, són el ritme cardíac i la seva uniformitat, i també la presència, amplitud, forma i separació de les altres ones. Les amplituds estan associades directament a l'estat de les cèl·lules del múscul cardíac, mentre que els temps venen condicionats per les cèl·lules nervioses. És per tant, molt important no alterar els paràmetres durant la mesura, ja que es poden produir errors d'interpretació.

Per l'enregistrament de les dotze derivacions habituals, la connexió al pacient es realitza mitjançant un 5 o 10 fils de colors diferents, connectant cadascun a un elèctrode: 3 per a les derivacions de les extremitats, 1 o 6 per a les derivacions precordials, i un altre elèctrode de referència, connectat a la cama dreta o a qualsevol altra part del cos.

L'incorporació d'un ordinador als electrocardiogrames permet obtenir informació addicional. Hi ha models que aconseguen detectar el feix de His només utilitzant elèctrodes superficials, a base de processar digitalment el senyal ECG. Tot i que el percentatge de deteccions és molt variable d'uns equips a uns altres, tot i no superat el 80 % en molts casos, la seguretat i la comoditat i la rapidesa de la mesura davant els registres intracardíacs obtinguts mitjançant cateterisme, donen gran interès a aquestes tècniques.

Una altra limitació de l'electrocardiografia clàssic es que sol obté part de la informació sobre la veritable distribució de potencials en la superfície, segons s'ha demostrat [Pallás, 1988].



4.4. Base de dades utilitzada

La base de dades cardíaca utilitzada en aquest projecte està formada per els senyals electrocardiogràfics, senyal RR, adquirits a l'Hospital de la Santa Creu i Sant Pau de Barcelona (HSP) i a l'Hospital Universitario de Getafe (HUG). La base de dades ha comptat amb 39 pacients en procés de weaning o procés de desconnexió del respirador que han estat un temps connectats a la ventilació mecànica assistida. En aquest treball s'estudiaran els senyals RR mesurats en pacients sotmesos a la prova de tub en T per a la desconnexió del suport ventilatori durant 30 minuts.

Aquests senyals pertanyen als mateixos pacients que els senyals descrits a l'apartat 3.5 i el senyal RR ha estat enregistrada simultàniament amb el senyal pneumotacogràfic. Per tant, els senyals RR també estaran classificats en els grups C1, pacients que han superat la prova de tub T, i el grup C2, pacients que no han superat la prova de tub T.

4.5. Complexitat del sistema cardíac

Tradicionalment l'estudi de la variabilitat del senyal RR s'ha realitzat mitjançant índexs obtinguts del domini temporal i freqüencial [Task Force, 1996], per estratificar pacients amb un risc cardíac. Aquests anàlisis han contribuït a la comprensió del funcionament de SNA (Sistema Nerviós Autònom) durant les fluctuacions de l'interval RR. Tot i que, els mètodes espectrals han estat els que més interès han despertat, aquests descriuen l'estructura lineal i fortament periòdica de la variabilitat del senyal RR, no essent capaços de caracteritzar la dinàmica no lineal amagada en la generació dels batecs cardíacs. En els últims anys s'ha trobat que la variabilitat del senyal RR és una senyal complexa i altament no lineal [Kurths et al., 1995] [Carvajal et al., 2002], resultant de la interacció dels diferents sistemes, pel que es busca noves formes d'enfocar l'anàlisi d'aquesta senyal com la dinàmica no lineal de sistemes, teoria del caos o anàlisi de complexitat, tals que permetin obtenir la informació oculta en aquestes senyals.



4.6. Interaccions entre el sistema respiratori i el cardíac

El coneixement de la relació entre la variabilitat del ritme cardíac, amb diversos mecanismes de la complexa funció anatòmica, ha portat a considerar que la variabilitat del senyal RR es atribuïda a la funció dinàmica no lineal i complexa del sistema respiratori i del sistema cardiovascular i a altres sistemes de control de l'organisme [HOYER et al., 1998]. Amb aquest concepte, si s'estudien les complexes interaccions dels mecanismes de control presents en la respiració i en la variabilitat del senyal RR pot contribuir a un millor diagnòstic i caracterització de l'activitat del Sistema Nerviós Autònom [Porta et al., 2000]



5. Anàlisi inicial de les dades

5.1. Introducció

Aquest capítol s'ha estructurat en dues parts. En la primera es descriuran els senyals que s'han utilitzat per fer l'estudi, un total de 39 pacients pertanyents a dos hospitals diferents. Els senyals han estat classificats en el grup C1 o C2 segons la resposta que han tingut a la prova de tub en T a la que han estat sotmesos. També es divideixen els pacients en el grup d'aprenentatge i de test. Aquesta separació es realitza per utilitzar uns pacients per trobar el mètode per a diferenciar entre els grups C1 i C2, i els que s'utilitzaran per verificar el mètode.

En la segona part d'aquest capítol es fa un estudi preliminar de les sèries temporals per veure si únicament amb l'ús de paràmetres estadístics clàssics, la mitja, la desviació tipus i el coeficient de variació de Pearson, són suficients per establir diferències significatives entre els que han respost bé a la prova i els que no. D'aquesta manera, si els resultats no són prou bons, es justificarà l'ús d'altres mètodes més complicats.

5.2. Característiques de la base de dades

En aquest treball s'ha disposat d'una base de dades d'un total de 39 pacients pertanyents a dos centres hospitalaris diferents:

- Hospital de la Santa Creu i Sant Pau (HSP), Barcelona
- Hospital Universitario de Getafe (HUG), Getafe.

Per a cada pacient es disposa de les sèries temporals pel període cardíac $RR(n)$ i del període $T_{tot}(n)$.

Per identificar els pacients s'utilitzarà un codi alfanumèric de 5 caràcters, amb la següent estructura:

- Primer caràcter: p
- Segon caràcter: 0 (HSP) o 1 (HUG)
- Resta caràcters: codi numèric identificador dels pacients

A més a més també es disposa de com han reaccionat aquests pacients a la desconexió del ventilador mecànic o weaning, Els pacients han estat classificats en dos grups:



- C1, pacients que han superat la prova de tub en T, moment de desconnexió adequat
- C2, pacients que no han superat la prova, la desconnexió no ha tingut èxit.

| Identificador pacient | Resultat prova T | Hospital | Identificador pacient | Resultat prova T | Hospital |
|-----------------------|------------------|----------|-----------------------|------------------|----------|
| p0004 | C1 | HSP | p1001 | C2 | HUG |
| p0006 | C1 | HSP | p1002 | C2 | HUG |
| p0008 | C1 | HSP | p1003 | C2 | HUG |
| p0012 | C1 | HSP | p1004 | C2 | HUG |
| p0013 | C1 | HSP | p1005 | C2 | HUG |
| p0014 | C2 | HSP | p1007 | C2 | HUG |
| p0015 | C1 | HSP | p1008 | C2 | HUG |
| p0018 | C1 | HSP | p1010 | C2 | HUG |
| p0019 | C1 | HSP | p1012 | C2 | HUG |
| p0020 | C1 | HSP | p1014 | C1 | HUG |
| p0021 | C2 | HSP | p1017 | C1 | HUG |
| p0022 | C2 | HSP | p1018 | C1 | HUG |
| p0024 | C1 | HSP | p1019 | C1 | HUG |
| p0025 | C1 | HSP | p1020 | C1 | HUG |
| p0026 | C1 | HSP | p1023 | C1 | HUG |
| p0027 | C1 | HSP | p1024 | C2 | HUG |
| p0028 | C1 | HSP | p1025 | C2 | HUG |
| p0029 | C1 | HSP | | | |
| p0030 | C1 | HSP | | | |
| p0034 | C1 | HSP | | | |
| p0035 | C1 | HSP | | | |
| p0036 | C1 | HSP | | | |

Taula 5.1 Classificació dels pacients segons la seva resposta a la prova de tub T i el seu centre hospitalari d'origen

A partir d'estudis que s'han realitzat amb anterioritat, [Caminal et al., 2004], s'ha observat que un pacient que presenta una patologia presenta menys variabilitat en les seves senyals que un individu sa, per tant, es realitza la hipòtesi que un pacient candidat a poder superar la desconnexió de la respiració assistida presentarà uns senyals amb més complexitat en el seu patró. Aquesta serà la hipòtesi de partida per poder separar els dos grups.



5.3. Divisió de les dades

5.3.1. Grup d'aprenentatge i grup de test

L'objectiu d'aquest projecte és aconseguir determinar el moment en que ha de produir-se la desconexió dels pacients de la respiració assistida, a partir de l'estudi de les sèries temporals dels senyals respiratoris, concretament el període respiratori, i cardíacs, el senyal RR. El procediment consistirà en utilitzar primer totes les sèries temporals de tots els pacients per decidir quins seran els índexs més adients per poder diferenciar entre els dos grups, C1 i C2. Posteriorment s'utilitzarà una part dels pacients, els pertanyents al grup d'aprenentatge, per trobar un mètode adient per diferenciar-los i una altra part, els del grup de test, per validar si aquest mètode funciona amb pacients que no han intervingut en la determinació del mètode.

Els pacients queden distribuïts tal com es mostra a la Taula 5.2 :

| Identificador pacient | Resultat prova T | Grup | Identificador pacient | Resultat prova T | Grup |
|-----------------------|------------------|--------------|-----------------------|------------------|------|
| p0004 | C1 | Aprenentatge | p0020 | C1 | Test |
| p0006 | C1 | Aprenentatge | p0024 | C1 | Test |
| p0008 | C1 | Aprenentatge | p0025 | C1 | Test |
| p0012 | C1 | Aprenentatge | p0026 | C1 | Test |
| p0013 | C1 | Aprenentatge | p0027 | C1 | Test |
| p0014 | C2 | Aprenentatge | p0028 | C1 | Test |
| p0015 | C1 | Aprenentatge | p0029 | C1 | Test |
| p0018 | C1 | Aprenentatge | p0030 | C1 | Test |
| p0019 | C1 | Aprenentatge | p0034 | C1 | Test |
| p0021 | C2 | Aprenentatge | p0035 | C1 | Test |
| p0022 | C2 | Aprenentatge | p0036 | C1 | Test |
| p1001 | C2 | Aprenentatge | p1007 | C2 | Test |
| p1002 | C2 | Aprenentatge | p1008 | C2 | Test |
| p1003 | C2 | Aprenentatge | p1010 | C2 | Test |
| p1004 | C2 | Aprenentatge | p1012 | C2 | Test |
| p1005 | C2 | Aprenentatge | p1014 | C1 | Test |
| | | | p1017 | C1 | Test |
| | | | p1018 | C1 | Test |
| | | | p1019 | C1 | Test |
| | | | p1020 | C1 | Test |
| | | | p1023 | C1 | Test |
| | | | p1024 | C2 | Test |
| | | | p1025 | C2 | Test |

Taula 5.2 Distribució dels pacients en els grups d'aprenentatge i de test



5.4. Anàlisi preliminar de les sèries temporals

En aquesta part de la memòria es realitzarà un primer estudi de les dades de que es disposa. Aquest estudi consistirà en realitzar el càlcul dels paràmetres mitja, desviació tipus i el coeficient de variació de Pearson de cada sèrie temporal i, per comprovar si són suficients per poder diferenciar entre els dos grups, el C1 i el C2.

La mitja i la desviació tipus són els paràmetres més intuïtius per estudiar la variabilitat de les sèries temporals, així com el coeficient de variació. Abans d'estudiar tècniques d'anàlisi més complicades és convenient comprovar quins són els resultats amb aquests paràmetres més senzills, tan conceptualment com de càlcul.

5.4.1. Test de significació. Verificació d'hipòtesis.

Un test de significació és una regla per decidir si una mostra en particular pertany a un grup o un altre. Consisteix en decidir amb quina probabilitat es certa una hipòtesi que s'ha fet sobre la població. Es realitza una verificació d'aquestes hipòtesis.

La verificació d'hipòtesis té com a objectiu comprovar si les afirmacions que s'han realitzat en referència a certs paràmetres són certes en funció de les dades que es maneja. Hi ha una certa terminologia bàsica [Alvarez, 1999] i la que s'utilitzarà serà:

- Hipòtesi nul·la. Es simbolitza per H_0 i recull la afirmació respecte al paràmetre que es sotmet a la prova d'estimació estadística corresponent.
- Hipòtesi alternativa. Es simbolitza per H_1 i recull el succés complementari de l'expressat a H_0 .
- Estadístic de prova. Expressió matemàtica calculada en funció dels estadístics mostrals que permet determinar si H_0 serà acceptada o no.
- Nivell de confiança. Es denota per $1 - \alpha$ i representa la probabilitat amb que es desitja provar H_0 .
- Nivell de significació. Es simbolitza per α i es la probabilitat complementària del nivell de confiança.

El plantejament de la hipòtesi nul·la dependrà de la tècnica estadística que s'utilitzi, en el cas d'aquest projecte, i de forma general s'utilitza per rebutjar que no hi hagin diferències significatives entre els dos grups de complexitat. Per tant, la hipòtesi alternativa consistirà en que existeixi alguna diferència significativa entre els dos grups.



5.4.2. Mètodes paramètrics i mètodes no paramètrics

A l'hora de triar una mètode per poder realitzar les proves d'hipòtesi sobre les sèries temporals s'ha de tenir en compte si es poden utilitzar mètodes paramètrics o és necessari utilitzar mètodes no paramètrics.

Les distribucions teòriques estan descrites per números anomenats paràmetres, generalment la mitja i la desviació tipus, per això els mètodes que utilitzen suposicions sobre la distribució de les dades, es diuen mètodes paramètrics. La classe de mètodes estadístics que no impliquen suposicions sobre la distribució de les dades són anomenats mètodes no paramètrics o de lliure distribució [Armitage i Berry, 1997].

En els mètodes paramètrics l'interès es centra en estimar o provar una hipòtesi sobre un o més paràmetres de la població, a més s'ha de en suposar que es coneix la distribució de la població. Exemples d'aquestes proves paramètriques són la t de Student o l'ANOVA. S'utilitzen per realitzar proves d'hipòtesi sobre la diferència entre les mitges de les poblacions. Una de les hipòtesis que les fonamenta és que les poblacions tinguin una distribució aproximadament normal. Molts cops aquestes suposicions són clarament falses, i es justifica amb que aquests mètodes són robustos per superar aquests problemes. Altres cops s'utilitzen transformacions.

Per evitar fer suposicions innecessàries sobre les distribucions de les poblacions s'utilitzen els mètodes no paramètrics. Les avantatges d'aquests mètodes són principalment la no necessitat de realitzar afirmacions dels valors dels paràmetres de les poblacions i per tant, poder utilitzar-se quan es desconeix la distribució de la població. Són més fàcils i ràpids de calcular, generalment, tot i que alguna aplicació d'aquestes probes pot arribar a ser molt laboriosa per grans mostres. Una altre inconvenient d'aquestes proves és que si es tracten dades que podrien haver estat tractades amb mètodes paramètrics suposa una pèrdua d'informació. Alguns exemples de proves d'aquest tipus són la U de Mann-Whitney i la prova de Wilcoxon.

5.4.3. Proves paramètriques

Es començarà l'estudi preliminar de les dades mirant si és possible l'aplicació dels mètodes de comparació d'hipòtesis paramètrics, per tal d'evitar una pèrdua d'informació en el cas que es poguessin utilitzar, com s'ha comentat en l'apartat anterior.

Per poder aplicar aquest mètode s'han de complir les següents hipòtesis 0 [Ardanuy i Martín, 1993]:

- Independència de les poblacions. L'origen de les dades posa de manifest si les poblacions poden considerar-se independents o no.



- Aleatorietat de les mostres. És una hipòtesi totalment crítica. L'aleatorietat assegura la representabilitat.
- Normalitat de les dues poblacions d'origen. Tot i que aquesta hipòtesis és poc crítica, ja que el que realment es suposa es que les seves mitjanes es distribueixen segons una distribució normal, el qual en general es podrà considerar cert pel teorema central del límit.
- Igualtat de variàncies poblacionals. Aquesta propietat es coneix com homogeneïtat de les variàncies de les poblacions d'origen.

En les dades d'estudi d'aquest projecte les dues primeres hipòtesis es compleixen, les sèries temporals provenen de pacients independents i són representatives del grup al que pertanyen, s'han escollit pacients de manera aleatòria, no s'han seleccionat de forma prèvia.

Per poder aplicar el mètode farà falta comprovar la normalitat de les poblacions d'origen, que es realitzarà amb la prova de Kolmogorov-Smirnov, i l'homogeneïtat de les variàncies que es realitzarà amb el càlcul de l'Estadístic de Levene.

5.4.4. Prova de Kolmogorov-Smirnov

Per estudiar la distribució de les dades d'una mostra per analitzar si s'ajusta a alguna distribució teòrica es plantegen dues possibilitats, la prova de la ji-quadrada i la prova de Kolmogorov-Smirnov. Com el nombre de dades de que es disposa no és gaire gran el procediment més adient és la prova de Kolmogorov-Smirnov.

Fonament teòric

Per poder aplicar el mètode de Kolmogorov-Smirnov és necessari que es compleixin dues hipòtesis:

- La mostra és aleatòria.
- La distribució teòrica $F_T(x)$ establerta és contínua.

Quan es calcula la bondat de l'ajust de Kolmogorov-Smirnov, es fa una comparació entre alguna funció de distribució acumulada i teòrica, $F_T(x)$, i alguna funció acumulada d'una mostra, $F_S(x)$. Les mostres que s'utilitzen en aquest projecte són aleatòries d'una població amb distribució acumulada $F(x)$ desconeguda.

En aquest mètode les hipòtesis per a la prova de significació seran:

- $H_0: F(x) = F_T(x)$, per tot $x \in (-\infty, \infty)$



- $H_A: F(x) \neq F_T(x)$ per almenys un x .

Aquestes hipòtesis impliquen que si existeix una concordança entre les dues distribucions acumulades, la teòrica i la de la mostra, la hipòtesi H_0 és certa, i per tant, la mostra s'ha extret de la població amb la funció de distribució acumulada $F_T(x)$. Si existeix una discrepància entre les funcions de distribució teòrica i observada massa gran per atribuir-la a l'atzar, es refusa la hipòtesi nul·la i s'accepta la hipòtesi alternativa H_A , la mostra no s'ha extret d'una població amb una funció de distribució acumulada $F_T(x)$.

Per calcular la diferència entre la funció de distribució acumulada teòrica $F_T(x)$ i la funció acumulada de la mostra $F_S(x)$, es mesura l'estadístic D , equació 5.1, que és la distància vertical màxima entre $F_S(x)$ i $F_T(x)$. En el cas de la prova bilateral l'estadístic és [Ardanuy i Martín, 1993]:

La distribució mostral de l'estadístic D , sota H_0 , es troba tabulada i depèn només de la mida de la mostra. [Wayne, 1977]

$$D = \sup_x |F_S(x) - F_T(x)| \quad (\text{Eq. 5.1})$$

Aplicació de la prova de Kolmogorov-Smirnov a les sèries temporals

S'han calculat la mitja, la desviació tipus i el coeficient de variació de totes les sèries temporals, amb les Equacions 5.2, 5.3 i 5.4, i el programa 'Estadistics.m' de Matlab, Annex B. Obtenim tres variables noves a les quals els aplicarem la prova.

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij} \quad (\text{Eq. 5.2})$$

$$s_i^2 = \left(\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \right) \quad (\text{Eq. 5.3})$$

$$cv_i = \frac{s_i}{\bar{x}_i} \quad (\text{Eq. 5.4})$$

En la Taula 5.3 apareixen els resultats obtinguts per cadascuna d'aquestes variables segons grups i per cada sèrie temporal:



| | GRUP | $\mu \pm \sigma$ |
|-----------|------|-------------------------|
| MITJARR | C1 | 708,6350 \pm 136,9095 |
| | C2 | 706,4748 \pm 127,1745 |
| STDRR | C1 | 83,5232 \pm 65,3651 |
| | C2 | 57,4047 \pm 54,8622 |
| CVRR | C1 | 0,1166 \pm 0,0831 |
| | C2 | 0,0788 \pm 0,0612 |
| MITJATTOT | C1 | 2,8931 \pm 0,8932 |
| | C2 | 2,0480 \pm 0,4649 |
| STDTTOT | C1 | 0,5189 \pm 0,2460 |
| | C2 | 0,6181 \pm 0,4137 |
| CVTTOT | C1 | 0,1875 \pm 0,0896 |
| | C2 | 0,3193 \pm 0,2263 |

Taula 5.3 Estadístics descriptius mitja, desviació tipus i coeficient de variabilitat de les diferents sèries temporals

S'ha aplicat el mètode de Kolmogorov-Smirnov a les sèries temporals formades per aquestes noves variables amb el paquet estadístic SPSS i el resultat obtingut apareix a la Taula 5.4:

| | MITJARR | STDRR | CVRR | MITJATTOT | STDTTOT | CVTTOT |
|------------------------|---------|--------|--------|-----------|---------|--------|
| NIVELL DE SIGNIFICACIÓ | 0,9978 | 0,3712 | 0,4065 | 0,0305 | 0,7365 | 0,1086 |

Taula 5.4 Valors dels nivell de significació de la mitja, la desviació tipus i el coeficient de variació per a les diferents sèries temporals en la prova de Kolmogorov-Smirnov.

El nivell de significació, α , és 0,05, per tant, en el cas de la variable mitjaTtot podem rebutjar la hipòtesis nul·la, i per tan la distribució d'aquesta variable no segueix una distribució normal. En el cas de les altres variables no es pot rebutjar la hipòtesis nul·la, i algunes, com ara la mitjaRR, tenen un nivell de significació elevat, per tan la seva probabilitat d'estar distribuïdes amb una llei normal és elevada.



5.4.5. Test de Levene

Fonament teòric

El test de Levene és utilitzat per analitzar si un conjunt de mostres tenen la mateixa variança. Si aquesta suposició es compleix, es diu que existeix homogeneïtat de variàncies. Aquest test és poc sensible a la hipòtesi de normalitat de les poblacions d'origen.

Les hipòtesis nul·la i alternativa que utilitza aquest mètode són:

- $H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$ per a tot k
- $H_A: \sigma_i^2 \neq \sigma_j^2$ per almenys una parella (i, j)

Donada una variable Y amb una mostra de mida N dividida en k subgrups, on N_i és la mida del subgrup i , es defineix l'estadístic del test de Levene com:

$$W = \frac{(N - k) \cdot \sum_{i=1}^k N_i \cdot (\bar{Z}_i - \bar{Z})^2}{(k - 1) \sum_{i=1}^k \sum_{j=1}^{N_i} (Z_{ij} - \bar{Z}_i)^2}$$

Equació 5.5 Estadístic del test de Levene

On Z_{ij} pot estar definit de les següents maneres:

- a) $Z_{ij} = |Y_{ij} - \bar{Y}_i|$ on \bar{Y}_i és la mitja del subgrup i .
- b) $Z_{ij} = |Y_{ij} - \tilde{Y}_i|$ on \tilde{Y}_i és la mediana del subgrup i .

Els dos tipus de definició per Z_{ij} determinen la potència del test i fan que el test de Levene sigui robust. Per robust, s'entén la capacitat del test de no detectar grups amb no homogeneïtat quan les dades estan normalment distribuïdes i en realitat si que presenten homogeneïtat de variàncies. Per potència del test, s'entén la capacitat del test de determinar no homogeneïtat correctament.

Es recomana la definició basada en la mediana, la b), ja que fa que el test sigui robust davant dades que no provenen d'una distribució normal, però també manté la seva potència.



En el test es rebutja la hipòtesi que les variàncies siguin homogènies en el cas que $W > F_{(\alpha, k-1, N-k)}$ on $F_{(\alpha, k-1, N-k)}$ és el valor superior de la distribució F amb $k-1$ i $N-1$ graus de llibertat amb un nivell de significació α . El valor de F es troba tabulat.[Wayne, 1977]

Aplicació del test de Levene d'homogeneïtat de les variàncies a les sèries temporals

S'ha calculat l'estadístic de Levene i la significació que s'ha obtingut queda reflectida en la **Taula 5.5**. Els valors s'ha calculat amb l'ajuda del paquet SPSS:

| | Estadístic de Levene | Significació |
|-----------|----------------------|--------------|
| MITJARR | 0,0007 | 0,9793 |
| STDRR | 1,5425 | 0,2221 |
| CVRR | 2,7862 | 0,1035 |
| MITJATTOT | 9,7474 | 0,0035 |
| STDTTOT | 2,8010 | 0,1026 |
| CVTTOT | 15,6887 | 0,0003 |

Taula 5.5 Resultats del test de Levene per a la prova d'homogeneïtat de les variàncies.

En aquestes dades s'observa que per a les variables mitjaTtot i coefvarTtot és pot rebutjar la hipòtesi nul·la d'igualtat de les variàncies. Per a la resta, excepte per a mitjaRR, tot i que no es pot rebutjar la hipòtesi d'igualtat, el nivell de significació no és gaire elevat.

5.4.6. Proves no paramètriques

Degut a que no totes les variables compleixen les hipòtesis que demanen les proves paramètriques, s'ha pres la decisió d'utilitzar per a totes les variables la mateixa prova d'hipòtesi, i per tant, s'utilitzarà un test de significació no paramètric.

Entre les proves no paramètriques més utilitzades destaquen la prova de Mann-Whitney i la prova de Wilcoxon. Les dues realitzen una comparació entre les medianes dels diferents grups de les poblacions. En aquest projecte s'utilitzarà la prova de Mann-Whitney ja que és la que s'utilitza per a mostres independents.



Fonament teòric

La prova U de Mann-Whitney és un mètode no paramètric que realitza proves d'hipòtesi sobre dades independents. La seva característica principal és que es basa en els rangs de les observacions.

Les suposicions d'aquest mètode són:

- les dues mostres de mides n i m , respectivament, que s'utilitzen per a l'anàlisi s'han extret independentment i a l'atzar de les seves respectives poblacions.
- L'escala de mesura es almenys ordinal. Per escala ordinal s'entén que les observacions no només difereixen de categoria a categoria, sinó que poden classificar-se per rangs amb algun criteri, en aquest cas pacients que han superat la prova amb èxit i els que no.
- Si les poblacions són diferents, es diferencien només en el que respecta a les seves medianes.

Quan es satisfan aquestes suposicions, es pot provar la hipòtesi nul·la de que les medianes de les dues poblacions són iguals contra l'alternativa que les poblacions no tenen medianes iguals:

- $H_0: M_x = M_y$
- $H_A: M_x \neq M_y$

En el cas que les dues poblacions fossin simètriques, de manera que la mitja i la mediana dins de les poblacions coincidissin, les conclusions a les que s'arribaria respecte a les medianes de les dues poblacions serien aplicables també a les seves mitges.

Per calcular l'estadístic de prova es combinen les dues mostres i s'ordenen per rangs totes les observacions, de la més petita a la més gran, tenint present la mostra a la qual pertanyen. L'estadístic de prova T es calcula mitjançant :

$$T = S - \frac{n \cdot (n + 1)}{2} \quad (\text{Eq. 5.5})$$

On n es el nombre d'observacions de la mostra X , i S la suma dels rangs assignats a les observacions de la població de valors de X . L'elecció dels valors de la mostra que es marquen amb X és arbitrària.



En el cas de la prova bilateral els valors de T calculats que siguin suficientment grans o suficientment petits causaran el rebuig de l'hipòtesi nul·la. La regla de decisió que s'utilitzarà serà que es rebutja H_0 si el valor de T calculat és menor que $W_{\alpha/2}$ o major que $W_{1-\alpha/2}$, on $W_{\alpha/2}$ és el valor crític de T per n, m i $\alpha/2$, tabulat [Wayne, 1977]. α és el nivell de significació desitjat.

Aplicació del test de Mann-Whitney a les series temporals

S'ha realitzat la prova U de Mann-Whitney i els resultats queden resumits en la **Taula 5.6**. Els valors s'ha calculat amb l'ajuda del paquet SPSS:

| | MITJARR | STDRR | CVRR | MITJATTOT | STDTTOT | CVTTOT |
|--------------------------------------|----------|----------|----------|-----------|----------|----------|
| U de Mann-Whitney | 170,0000 | 131,0000 | 129,0000 | 74,0000 | 155,0000 | 117,0000 |
| Sig. exacta [2*(Sig. unilateral)] | 0,8965 | 0,2056 | 0,1852 | 0,0025 | 0,5726 | 0,0925 |

Taula 5.6 Resultats del test de Mann-Whitney d'homogeneïtat de les variàncies.

A partir d'aquestes dades s'observa que només en el cas de la variable mitjaTtot podem diferenciar amb una seguretat suficient, segons aquesta prova estadística entre els dos grups, ja que la probabilitat que les medianes dels dos grups siguin iguals es inferior al nivell de significació, $\alpha=0.05$,

Com tant sols s'ha trobat un índex que ens podria servir per diferenciar entre els dos grups, i tot i que el p-valor que apareix per aquest es bastant bo, pareix important continuar l'estudi per trobar algun índex millor.



6. Tècniques d'anàlisi no lineal

6.1. Introducció

A la natura hi ha molts fenòmens que tenen un comportament complex, que no pot ser explicat com la suma de les parts que intervenen en ell. Per a poder estudiar aquests fenòmens s'han de buscar interaccions entre les parts que hi intervenen que sovint produeixen comportaments complicats, alguns d'aparença erràtica i desordenada. Al llarg de la història s'ha considerat que aquests comportaments eren processos de caràcter aleatori i que per tant, no podien ser descrits per equacions, com en el cas de processos deterministes.

Al voltant de la dècada dels 70, van aparèixer noves tècniques matemàtiques que han permès enfocar l'estudi d'aquests fenòmens des d'un altre punt de vista. La seva base està en considerar el procés a analitzar com un sistema determinista amb dinàmica no lineal. A partir d'aquí han nascut noves teories com la Teoria del Caos i l'Anàlisi de Complexitat.

Sovint, els termes dinàmica no lineal, caos i complexitat són utilitzats com a sinònims per a referir-se a un mateix tipus de tècniques d'estudi de sistemes complexos. Aquestes tècniques han demostrat ser útils per a una àmplia varietat d'aplicacions científiques en camps tant diversos com la física, la biologia, l'economia, la sociologia ... i l'estudi de sistemes biològics. En aquest últim camp, s'han efectuat estudis sobre el diagnòstic, tractament i classificació de diverses patologies. Exemples d'això podrien ser l'estudi del senyal electrocardiogràfic per a la detecció de patologies cardíaques [Voss et al., 1990], o l'estudi del sistema respiratori [Bruce i Danbenspeck, 1995].

L'objectiu d'aquest capítol es fer una breu introducció a l'estudi de sistemes no lineals. Es comentaran temes referents a la teoria del caos i a la complexitat dels sistemes dinàmics. A més, es comentaran les tècniques d'estudi de sistemes no lineals i es justificarà el fet d'escollir la tècnica de la Dinàmica Simbòlica en aquest projecte. Finalment, es farà una introducció a la variabilitat del patró respiratori i al seu caràcter no lineal.

6.2. Sistemes Dinàmics

Un sistema dinàmic és aquell que presenta una variació de les magnituds que el defineixen en el temps. Des del punt de vista matemàtic pot estar definit com una



formulació determinista per a descriure l'estat d'un sistema al llarg del temps [Ott, et al., 1994].

Una manera d'establir una primera classificació dels sistemes dinàmics és segons si tenen representada la variable temps com una variable contínua o discreta, fet que repercutirà en el seu tractament.

Un exemple de sistema dinàmic en que el temps (denotat per t) és una variable contínua, és un sistema d'equacions diferencials ordinàries, autònomes, de primer ordre, es troba expressat a l'equació 6.1:

$$\begin{aligned} \frac{dx^{(1)}}{dt} &= F_1(x^{(1)}, x^{(2)}, \dots, x^{(N)}) \\ \frac{dx^{(2)}}{dt} &= F_2(x^{(1)}, x^{(2)}, \dots, x^{(N)}) \\ &\dots \\ \frac{dx^{(N)}}{dt} &= F_N(x^{(1)}, x^{(2)}, \dots, x^{(N)}) \end{aligned} \quad (\text{Eq. 6.1})$$

Que es pot representar de manera vectorial com a l'equació 6.2 on x és un vector de dimensió N , $x = (x^{(1)}, x^{(2)}, \dots, x^{(N)})$

$$\frac{dx(t)}{dt} = F(x(t)) \quad (\text{Eq. 6.2})$$

En els sistemes dinàmics per qualsevol valor inicial del sistema $x(0)$, en principi es poden resoldre les equacions per a obtenir l'estat futur del sistema $x(t)$, per a $t > 0$. A la Figura 3.1 [Ott et al., 1994] es pot veure el camí seguit per l'estat d'un sistema d'aquest tipus al llarg del temps, per el cas $N=3$. Aquest espai és conegut com l'espai de fase tridimensional i la ruta seguida per el sistema en l'espai de fases al llarg del temps, s'anomena òrbita o trajectòria.



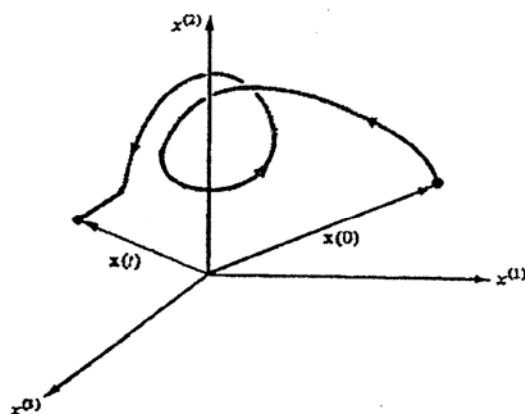


Figura 6.1 Òrbita en l'espai de fase tridimensional

Si s'observa l'espai de fases que és de dimensió N , es veu que la variable temps no apareix de manera explícita en cap eix. Així la geometria de les solucions en l'espai de fase proporciona una gran quantitat d'informació sobre la dinàmica del sistema [Solé i Manubria, 1996].

En el cas que el sistema dinàmic tingui representat el temps com a variable discreta, la seva representació es farà mitjançant un mapa de retorn, que vectorialment quedarà descrit a l'equació 6.3. En aquest projecte les sèries considerades estan expressades d'aquesta manera.

$$x_{n+1} = M(x_n) \quad (\text{Eq. 6.3})$$

n és la variable temps $n = 0, 1, 2, \dots, i$

x el vector de dimensió N , $x_n = (x_n^1, x_n^2, \dots, x_n^N)$

En aquests sistemes a partir d'un estat inicial x_0 , es pot obtenir l'estat en l'instant $n=1$ per aplicació directa de l'equació 6.3, quan $n=0$. Mitjançant la iteració d'aquesta equació es poden anar aconseguint la resta d'estats. Així donada una condició inicial x_0 , es pot generar l'òrbita o trajectòria del sistema en temps discret.

Els sistemes dinàmics es caracteritzen per la presència de conjunts atractors en l'espai de fases o en els mapes de retorn. La majoria dels sistemes dinàmics coneguts evolucionen en el temps cap a un equilibri o romanen repetint-se en cicles. Si evolucionen cap a l'equilibri ho fan formant una trajectòria en forma d'espiral cap a un punt límit, el qual és un atractor ja que les trajectòries properes a aquest punt són atretes cap a ell. En canvi, si romanen repetint-se



en cicles, formen trajectòries circulars o cicles límit que també són atractors. A part dels punts límit i els cicles límit, existeixen més punts característics com els nodes, els punts de sella o comportaments encara més complexos.

6.2.1. Sistemes dinàmics no lineals

Els sistemes dinàmics que estan descrits per equacions diferencials de primer grau, són els anomenats sistemes dinàmics lineals, mentre que els sistemes on apareix alguna de les seves variables elevada a alguna potència superior a un, o si hi ha productes entre variables, són anomenats sistemes dinàmics no lineals. D'aquesta manera s'estableix una segona classificació dels sistemes dinàmics.

Els sistemes no lineals són sistemes que no responen proporcionalment a les entrades. La major part dels sistemes no lineals generalment no tenen solució analítica, la qual cosa complica el seu anàlisi. No va ser fins l'aparició de les computadores, que es va començar a plantejar el seu estudi de manera més exhaustiva. Fins llavors s'acostumaven a utilitzar aproximacions lineals o simplement es desestimava la seva anàlisi.

S'ha anat acceptant lentament l'existència de propietats emergents que apareixen en aquests sistemes com a resultat de la interacció entre els subsistemes que els integren i que no poden explicar-se a partir de les propietats dels seus elements corresponents. Es tracta doncs, de sistemes complexos, on les propietats apareixen de les interaccions entre les seves parts constituents.

6.3. Teoria del caos

En certes circumstàncies, els sistemes deterministes no lineals es comporten erràticament, entrant en un estat anomenat caos. Aquest caos de naturalesa no lineal es refereix a una aleatorietat de classe restringida que es pot associar a la geometria fractal, tema que es tractarà posteriorment.

Una de les primeres característiques del comportament caòtic dels sistemes dinàmics no lineals, va ser descoberta per E. Lorenz durant l'estudi de la convecció de fluids per a poder explicar alguns canvis climàtics. Lorenz va descobrir que el sistema exhibia sensibilitat a les condicions inicials, és a dir, que diferències molt petites en els valors inicials provocaven comportaments totalment diferents en el temps, que feien que divergissin exponencialment. Això implica que com que les condicions inicials mai es coneixen amb suficient precisió, no tenen sentit les prediccions a llarg termini. Per aquesta raó, els sistemes caòtics estan considerats com a sistemes deterministes amb comportament impredecible [Lorenz, 1963].



En 1971, D. Ruelle i F. Takens, que estudiaven el fenomen de la turbulència utilitzant sistemes no lineals, van visualitzar noves estructures a l'espai de fases multidimensional que van anomenar atractors estranys. El seu nom es deu a que eren atractors que tenien una estructura molt diferent a la dels atractors que es coneixien.

En 1976 R. May, fent models ecològics amb la funció logística l'equació 6.4 que és una equació no lineal en temps discret molt simple, va trobar comportaments molt complicats.

$$x_{n+1} = k \cdot x_n \cdot (1 - x_n)$$

(Eq. 6.4)

Va descobrir que al variar el valor del coeficient k es produïen bifurcacions, que portaven a comportaments caòtics com es pot observar a la Figura 6.2 [Gleick, 1994]

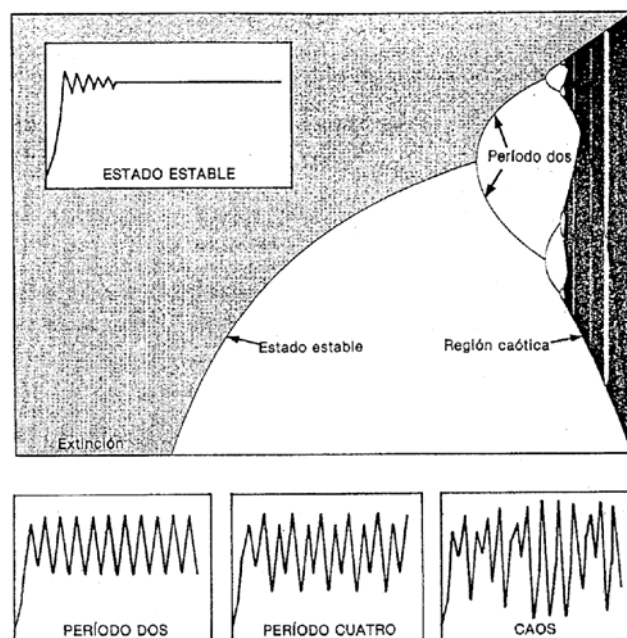


Figura 6.2 Diagrama de bifurcacions corresponent a l'equació logística al variar els valors del paràmetre k .

De forma paral·lela Mandelbrot va introduir els fractals. Un fractal consta de fragments geomètrics d'orientació i mida variable, però d'aspecte similar. Els detalls d'un fractal a certa escala, són semblants (encara que no necessàriament idèntics) als d'una estructura visible a



una escala major o menor. Els fractals posseeixen la propietat d'autosemblança que fa que s'assemblin a sí mateixos. A la Figura 6.3 [Gleick, 1994] es mostra una imatge del conegut Home de Mandelbrot que s'obté de la iteració en el pla complex de l'**Error! No se encuentra el origen de la referencia.**

$$z \rightarrow z^2 + c \quad (\text{Eq. 6.5})$$

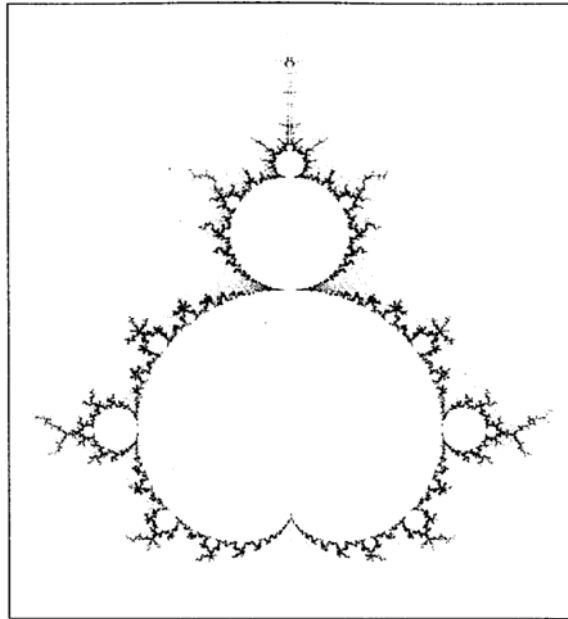


Figura 6.3 Imatge de l'Home de Mandelbrot de naturalesa fractal

Donat que un fractal està compost per estructures similars de "gra" cada cop més fi, la seva longitud no està estrictament definida. La noció de longitud perd el seu significat en el cas dels fractals, per això s'utilitza la dimensió per a quantificar de quina manera omple l'espai un fractal.

Els fractals tenen dimensió no entera (fraccionària). Tant els diagrames de bifurcacions com alguns dels atractors estranys tenen naturalesa fractal, així com moltes formes trobades en la natura.

En 1978 M. Feigenbaum va fer un descobriment sorprenent. Va descobrir que hi ha certes lleis universals que governen la transició de conducta regular a caòtica. Aixà vol dir que hi ha sistemes totalment diferents que es tornen caòtics de la mateixa manera. El seu treball va establir una relació entre el caos i les transicions de fase.



Les conseqüències de tots aquests descobriments han estat enormes. Gran part del desordre que ens envolta ha resultat ser només aparent. Darrera seu s'amaga un ordre que es pot arribar a traduir en models matemàtics simples, els quals han modificat la visió clàssica on l'ordre i el desordre eren conceptes oposats.

6.4. Complexitat

El caos determinista es només una peça d'un enorme conjunt de nous conceptes, que genèricament s'agrupen sota la denominació de sistemes complexos. La cerca de les lleis del complex, s'ha convertit a finals del segle XX en l'objectiu d'estudiosos procedents de camps molt diversos. Dins d'aquesta difícil cerca, s'han generat noves teories i models. Les xarxes neuronals, els autòmats cel·lulars, els objectes fractals o la criticalitat autoorganitzada han permès formular, de manera simple, les primeres hipòtesis generals [Solé i Manubria, 1996]. La teoria de complexitat és una teoria encara en fase de creixement.

Durant la dècada dels 70, un nou marc conceptual vindria a introduir nous elements en el cos teòric de la ciència. A partir dels treballs de I. Prigogine es van començar a aplicar certes eines de la física a l'estudi de sistemes químics o biològics lluny de l'equilibri.

S'havia observat que en determinades reaccions químiques sota condicions adequades, un sistema que havia de tendir cap a un estat d'equilibri homogeni, mostrava oscil·lacions periòdiques i adquiria una sorprenent estructura espacial en forma d'ones espirals. Es tractava d'un exemple de l'emergència espontània del complex.

No es disposa d'una definició simple de complexitat. La majoria de les que s'han fet involucren conceptes com l'aleatorietat, l'entropia i la informació, termes que per sí sols ja són difícils de definir. Una manera senzilla d'explicar la complexitat podria ser: "un objecte és complex si conté informació difícil d'obtenir" [Thelie et al., 1992].

El complex es troba a mig camí entre l'ordenat (un cristall) i el desordenat (un gas). Es manifesta en unes condicions molt especials, en els punts crítics en els que tenen lloc les transicions de fase. En aquests punts, ordre i desordre coexisteixen. Aquesta coexistència porta a l'aparició d'una propietat en els sistemes complexos naturals, les estructures fractals, caracteritzades per presentar el mateix aspecte bàsic a diferents escales.

Altres fenòmens complexos com els moviments sísmics, la dinàmica de muntanyes de sorra o la macroevolució biològica, tendeixen de forma espontània cap al punt crític. Tot i que la seva naturalesa és diferent, aquests sistemes exhibeixen propietats comunes. En tots ells es troben elements d'interacció que posseeixen la capacitat d'amplificar perturbacions fortuïtes. Aquestes poden ser, segons s'escaigui, petites tensions en les plaques tectòniques, la caiguda d'un gra de sorra sobre la muntanya o l'aparició d'una nova espècie. L'amplificació



de la pertorbació inicial pot provocar grans canvis com terratrèmols immensos, allaus de sorra o extinció d'espècies.

Una de les propietats dels sistemes complexos és la seva capacitat d'interaccionar amb el seu ambient i processar la informació rebuda. Per aconseguir-ho es requereix un cert grau d'ordre intern que permeti emmagatzemar informació, però a la vegada, aquesta informació ha de ser manipulable, el que obliga al sistema a tenir certa flexibilitat interna, amb cert grau de desordre. Quan es dona aquesta situació s'està a la frontera del caos.

Problemes tant diferents com les fluctuacions de l'economia, la resposta del sistema immunològic, la propagació dels impulsos nerviosos o la comunicació en una xarxa informàtica, tenen entre sí característiques similars. Tots aquests sistemes poden ser classificats com a sistemes complexos adaptatius. Això significa que hi ha principis generals que governen la conducta d'aquest tipus de sistemes.

6.5. Tècniques d'anàlisi no lineal

L'objectiu de l'anàlisi de dades experimentals, provinents de sistemes no lineals, és trobar un patró o estructura que ajudi a comprendre el comportament del sistema. Aquestes dades s'analitzen com una sèrie temporal $x(t)$ d'una variable x .

La presència del que és anomenat com atractor estrany pot donar lloc a sèries temporals molt complicades. Un atractor estrany té la mateixa propietat que els cicles límit o els punts límit, Apartat 6.2, on el sistema és atret cap a un estat final. En aquest cas però, l'estat no serà periòdic i a sobre serà extremadament complex.

La presència d'aquest atractor es revela millor en les representacions en l'espai de fases quan es parla de temps continu, on es representa $x(t)$ front les seves derivades $dx(t)/dt$ i derivades superiors. En el cas d'un senyal en temps discret x_1, x_2, \dots, x_N , es representaria en el mapa de retorn de m dimensions que seria $x_i = [x_i, x_{i+\tau}, \dots, x_{i+(m-1)\tau}]$ essent τ el retard escollit.

Així doncs, el conjunt de dades pot ser representat com una trajectòria en l'espai de fases o en el mapa de retorn. La idea és incrementar la dimensió d'aquest espai fins que la topologia de l'atractor no variï i quedi identificat. Un cop un atractor estrany ha estat identificat, la seva naturalesa pot ser quantificada calculant mesures de dimensió, com la de correlació o la de capacitat, l'exponent de Hurst o l'exponent de Lyapunov.

A continuació es farà una breu introducció a les mesures més utilitzades a l'actualitat a l'hora de realitzar l'anàlisi no lineal sobre sèries temporals. En aquest projecte s'ha decidit utilitzar



la denominada Dinàmica Simbòlica, tècnica que s'explicarà més extensament en el Capítol 7. En el present capítol es limitarà a introduir aquesta tècnica i explicar les raons per les que s'ha adoptat davant les altres possibles.

6.5.1. Dimensió

La dimensió és una mesura de la complexitat de la trajectòria en l'espai de fases. Teòricament, per tal d'obtenir una mesura precisa de la dimensió de l'atractor és necessari immers-lo en un espai de dimensió almenys $2d+1$, on d és la dimensió sencera més petita que conté l'atractor. Això és degut a que s'aprofita la propietat de que si per una dimensió de l'espai de fases queda identificada la trajectòria d'un atractor, si s'augmenta la dimensió de l'espai la trajectòria continuarà existint.

La dimensió permet una estimació dels graus de llibertat del sistema, és a dir, del nombre de variables independents que serien necessàries per a descriure el sistema dinàmic. Existeixen diverses definicions, però la més usada és la dimensió de correlació per ser considerada el millor estimador.

a) Dimensió de capacitat

La dimensió de capacitat és la forma més simple de dimensió i es coneix també com a dimensió fractal. Primer es cobreix l'atractor amb elements de volum (esferes, cubs, ...) cadascun de longitud característica ε i s'anomena $N(\varepsilon)$ al nombre d'elements de volum que són necessaris per a cobrir l'atractor. Quant més petit sigui ε , la suma dels elements de volum s'aproximarà més a l'atractor. Si l'atractor és D -dimensional (amb D sencer) llavors per a ε petit el nombre d'elements de volum que seran necessaris per cobrir l'atractor serà inversament proporcional a ε^D . És a dir, la relació que s'obté és $N(\varepsilon)=k \cdot \varepsilon^D$, per a alguna k . Així la definició de dimensió de capacitat s'obté de l'equació 6.6.

$$D_{cap} = \lim_{\varepsilon \rightarrow 0} \frac{\ln(N(\varepsilon))}{\ln\left(\frac{1}{\varepsilon}\right)} \quad (\text{Eq. 6.6})$$

Si el límit no existís, la dimensió de capacitat es diu que és indefinida. Quan el límit existeix, la qüestió és si amb un altre element de volum (esferes enlloc de cubs, o variant les mides) la D_{cap} varia. En cas que s'obtinguin valors diferents fent servir diferents elements de volum, la dimensió de capacitat es considerarà el valor més petit.

b) Dimensió de correlació



Un altre tipus de dimensió probabilística és la dimensió de correlació D_c que conté informació sobre l'estructura del sistema. Aquest paràmetre permet conèixer el nombre de dimensions mínimes d per a caracteritzar el sistema caòtic.

Si es suposen coneguts N punts d'una trajectòria, la correlació es defineix segons l'algorisme de Grassberg i Procaccia [Grassberg i Procaccia, 1991] com:

$$C(\varepsilon) = \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \Theta(r - |x_i - x_j|) \quad (\text{Eq. 6.7})$$

on Θ és la funció de Heaviside, o el que és el mateix:

$$C(\varepsilon) = \lim_{N \rightarrow \infty} \frac{M}{N^2} \quad (\text{Eq. 6.8})$$

on M és el nombre de parells de punts tals que la seva distància euclídea d_{ij} , compleixi que $d_{ij} = \|x_i - x_j\| < \varepsilon$. Llavors:

$$D_c = \lim_{\varepsilon \rightarrow 0} \frac{\ln C(\varepsilon)}{\ln \varepsilon} \quad (\text{Eq. 6.9})$$

La dimensió de correlació pot calcular-se directament a partir de la definició utilitzant una xarxa d'hipercubs de costat ε , però és més eficient i robust usar la definició de correlació.

6.5.2. L'exponent de Hurst

L'exponent de Hurst H quantifica la dinàmica caòtica del sistema a analitzar. Un exponent de Hurst elevat indicarà la presència d'una complexitat baixa (correlació positiva), mentre que un exponent de Hurst baix indica la presència d'una complexitat alta (correlació negativa).

Aquest exponent permet ponderar fins a quin punt les dades poden ser representades per un moviment de tipus brownià, que es donarà quan $H=0.5$. Exponents majors a 0,5 indiquen que les tendències del passat persisteixen en el futur, mentre que els valors menors de 0,5



indiquen que les tendències del passat tendeixen a ser inverses en el futur. A l'equació 6.10 apareix la relació d'on prové H:

$$\frac{R}{s} = \left(\frac{\tau}{2}\right)^H \quad (\text{Eq. 6.10})$$

R = rang, diferència entre valor màxim i mínim

s = desviació estàndard

τ = retard

Així H s'estimarà com el pendent de la representació logarítmica de R/s respecte τ .

6.5.3. Els exponents de Lyapunov

Els exponents de Lyapunov poden quantificar, en certa forma, la impredictibilitat del senyal estudiat i la seva dependència de les condicions inicials. En concret, és el major exponent de Lyapunov el que indicarà, quan sigui positiu, que el senyal estudiat és impredictible, de naturalesa caòtica. Si el seu valor és zero o negatiu voldrà dir que les dades són periòdiques, i si és infinit, significarà que les dades són aleatòries i no estan correlades.

Una estimació del màxim exponent de Lyapunov es basa en un algorisme proposat per [Moon, 1992]. Sigui x_k un punt en l'espai de fase de dimensió N. Sigui $L(k)$ la distància euclídea entre aquest punt i el punt més pròxim x_i . Sigui $L(k+\Delta k)$ la distància entre x_k i el punt $x_{k+\Delta k}$. Llavors es pot estimar l'exponent màxim de Lyapunov segons l'equació 6.11:

$$LE = \frac{1}{M \cdot \Delta k} \sum_{k=1}^M \log_2 \frac{L(k + \Delta k)}{L(k)} \quad (\text{Eq. 6.11})$$

on M és el nombre de punts en l'espai de fase i Δk és l'interval de predicció per un senyal periòdic, és a dir, el seu període.

6.5.4. Entropia

Les entropies són un mètode per quantificar el caos d'un sistema dinàmic. Hi han dos tipus d'entropies, l'entropia mètrica i l'entropia topològica. Aquestes entropies són positives per a sistemes caòtics i zero per a sistemes no caòtics. Les dues juguen un paper fonamental en la



teoria matemàtica del caos. No obstant, són usades menys que els exponents de Liapunov per examinar situacions que passen a la pràctica, ja que l'entropia topològica, sobretot, és molt més complicada de determinar numèricament.

En aquest projecte s'ha calculat l'entropia mètrica amb la finalitat de d'utilitzar-la com un índex més a l'hora de discriminar entre els dos grups de pacients.

L'entropia mètrica va ser introduïda per Kolmogorov i es coneguda de vegades con entropia K-S, degut a Sinai que també col·laborar en la seva definició final. L'entropia mètrica pot ser entesa com un número que mesura l'índex de temps de creació d'informació de com les òrbites caòtiques es desenvolupen. És a dir, degut a la sensibilitat a les condicions inicials dels sistemes caòtics, a prop d'òrbites divergents. Si en un principi es calculen les condicions inicials sense molta exactitud, pot semblar que dues òrbites siguin iguals a l'espai de fases. Si les òrbites evolucionen en el temps, pot ser que aquestes acabin sent diferents. Alternativament, si una òrbita és iterada, i s'observa la seva localització amb suficient precisió, els díigits que s'havien despreciat en un principi les poden fer diferents, i poden ser calculats. Assumint que es poden calcular les òrbites amb suficient precisió i que es saben les equacions que donen l'òrbita, es poden usar les observacions per obtenir més informació sobre els díigits imprecisos de les condicions inicials. En aquest sentit es diu que les òrbites creen informació.

La definició de l'entropia mètrica està basada en la formulació de Shannon del grau d'incertesa que són capaces de predir les sortides en un esdeveniment probabilístic. Si l'esdeveniment té r possibles sortides, i p_1, p_2, \dots, p_r , són les probabilitats de cada sortida. L'entropia de Shannon dona un nombre que caracteritza la quantitat d'incertesa que hi ha referent a quina sortida es donarà i ve donada per l'equació 6.12:

$$H_S = \sum_{i=1}^r p_i \ln(1/p_i) \quad [\text{Ott, 1999}] \quad (\text{Eq. 6.12})$$

6.5.5. Dinàmica Simbòlica

Aquesta ha estat la tècnica que s'ha utilitzat per al tractament de les sèries temporals objecte d'estudi en aquest projecte. Val a dir que es farà una explicació detallada d'aquesta tècnica i dels paràmetres que hi intervenen en el Capítol 7. En aquest apartat simplement es vol fer una petita introducció i la justificació del perquè s'utilitza.

La Dinàmica Simbòlica és una tècnica que mesura la variabilitat d'una sèrie temporal mitjançant la transformació de cada element de la sèrie amb els símbols d'un alfabet donat.



Tot i que es perd informació en aquesta transformació, la informació robusta i invariant del sistema continua apareixent i serà la que permetrà l'anàlisi de la sèrie i conseqüentment, del sistema que s'estudia.

L'avantatge principal d'aquesta tècnica, front a les altres tècniques que s'han exposat, és que treballa correctament tant amb senyals amb elevat nombre de mostres, com amb senyals amb poques mostres. Les sèries temporals considerades contenen entre 400 i 3000 mostres, depenent del pacient. Com algunes, en especial, contenen pocs nombres de mostres s'ha fet fa desaconsellable utilitzar les altres tècniques, com la dimensió de correlació, ja que aquestes necessiten d'un elevat nombre de mostres per a obtenir uns resultats fiables.



7. Dinàmica simbòlica i Anàlisi de la Interacció Cardiorespiratòria

7.1. Introducció

L'anàlisi de la interacció cardiorespiratòria es realitzarà mitjançant l'estudi de la Dinàmica Simbòlica Conjunta entre sèries temporals.

La Dinàmica Simbòlica és una tècnica que ens proporciona una de les formes més rigoroses de definir la complexitat en els sistemes dinàmics. El concepte es basa en un 'granulat gruixut' per explicar l'evolució d'un procés dinàmic [Hao, 1989]. La Dinàmica Simbòlica és un mètode no lineal. La sèrie temporal es transforma en una seqüència de símbols d'un alfabet donat, en la transformació es perd algun detall del senyal inicial, però això no impedeix l'anàlisi del comportament general del sistema ja que la informació invariant i robusta del sistema no es perd.

Un cop obtinguts els resultats amb el mètode no lineal de transformació, s'analitzaran les dades utilitzant mètodes estadístics. La utilització d'un mètode estadístic o un altre està molt lligada a les hipòtesis que s'han de realitzar per poder-los aplicar. Una part destacada en aquest estudi és la presa de decisions per escollir-los, s'hauran d'analitzar les hipòtesis dels mètodes abans d'aplicar-los.

7.2. Fonament teòric de la Dinàmica Simbòlica

Con ja s'ha comentat anteriorment, la Dinàmica Simbòlica és un mètode per estudiar els sistemes dinàmics. En el cas d'aquest projecte s'analitzaran variables discretes extretes d'un sistema dinàmic. La descripció més completa de l'evolució d'una variable discreta extreta d'un sistema dinàmic s'expressa mitjançant l'equació 7.1 [Hao, 1989]:

$$x_{n+1} = f(\mu, x_n), x \in V \quad (\text{Eq. 7.1})$$

On μ és el valor mitjà de la sèrie temporal. Aquesta equació requereix el coneixement de tota la sèrie temporal $\{x_i \quad i = 0, 1, \dots\}$ Per a tots els possibles valors inicials. No obstant, es pot desenvolupar una descripció del sistema amb un "granulat gruixut", ignorant els nombres actuals de la sèrie, però retenint el caràcter essencial de la seva evolució.



Hi ha dues possibilitats a l'hora d'escollir la manera de codificar amb Dinàmica Simbòlica les sèries temporals:

- Estudiar com varia el valor de la sèrie temporal, és a dir, si el valor de la variable que s'està estudiant creix, respecte al seu valor anterior, o disminueix, la qual cosa ens porta a l'equació 7.2 [Baumert et al., 2002], partint d'aquesta transformació es realitzarà l'estudi de les dues sèries temporals conjuntes:

$$s = \begin{cases} 0 : (x_n - x_{n+1}) \leq I \\ 1 : (x_n - x_{n+1}) > I \end{cases} \quad (\text{Eq. 7.2})$$

I = valor llindar que es fixa a un cert valor, normalment 0.

- Es divideix el rang de valors del senyal, per exemple, l'interval V de la seqüència d'una sèrie, en varies regions i s'anomena cada regió amb un símbol. Es reemplaça cada número x_i en amb el símbol de l'interval on està inclòs, cara sèrie $\{x_i\}$ es transformarà en una seqüència de símbols. Amb aquesta transformació es crea la possibilitat d'introduir un esquema de classificació per les sèries numèriques. Totes les seqüències que corresponguin a una mateixa seqüència de símbols poden ser equivalents.

Si es divideix l'interval de manera acurada, tenint en compte la física essencial del sistema, es poden definir algunes regles per a la construcció i ordre per a les sèries simbòliques, aconseguint així un millors resultats.

Un exemple d'aquesta transformació és l'equació 7.3 [Caminal et al, 2004]

$$S_n = \begin{cases} 0 & \text{si} & b_n > (1 + \alpha / 2)\mu \\ 1 & \text{si} & (1 - \alpha / 2)\mu < b_n \leq (1 + \alpha / 2)\mu \\ 2 & \text{si} & b_n \leq (1 - \alpha / 2)\mu \end{cases} \quad (\text{Eq. 7.3})$$

μ = Mitja de la sèrie temporal

La partició en regions dels valors de la sèrie temporal, la seva etiquetació o simplificar l'evolució són les idees de les que parteix la Dinàmica Simbòlica, per tal de poder fer un estudi més intuïtiu sense perdre la part essencial de la que es parteix. L'elecció de les diferents equacions per codificar el senyal és un dels punts més important d'aquest mètode i sempre s'ha de tenir en compte la naturalesa del senyal a estudiar per que aquesta transformació sigui el més efectiva possible.



7.2.1. Dinàmica Simbòlica Conjunta

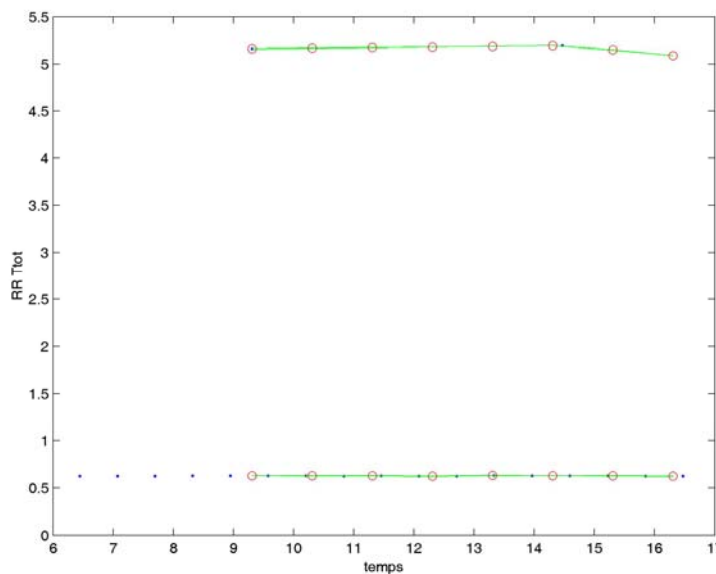
En els sistemes dinàmics moltes variables diferents poden estar interactuant, per estudiar la conseqüència d'aquestes interaccions es pot aplicar Dinàmica Simbòlica Conjunta.

La Dinàmica Simbòlica Conjunta consisteix en la transformació de les diferents sèries temporals que poden estar relacionades en símbols. La transformació s'aplica a les dues sèries temporals a la vegada i s'estudia l'aparició d'un símbol en una sèrie amb combinació amb els símbols de l'altra o altres sèries.

Un punt molt important per a la correcta aplicació d'aquest mètode es tenir en el mateix instant de temps els valors de cada sèrie temporal, per la qual cosa és necessari normalment interpolat les dades disponibles i remostrejar sincronitzadament.

En aquest projecte s'han analitzat amb Dinàmica Simbòlica Conjunta les sèries temporals $RR(n)$ i $Ttot(n)$, la Figura 7.1. S'han considerat dues freqüències de mostreig, 1 Hz ($T=1s$), i 2 Hz ($T=0,5s$), per tal d'avaluar la seva incidència en els resultats finals de l'estudi. El tipus d'interpolació que s'ha utilitzat ha estat la interpolació lineal.

A la Figura 7.1 es veu com s'aconsegueixen dues senyals en el mateix instant de temps. També és important destacar el fet que les dues senyals han de començar en el mateix instant de temps, i finalitzar al mateix instant, cosa que porta de vegades a no estudiar alguns valors inicials o finals d'algun senyal.



Punts blaus: Mostres inicials del senyal

Rodones vermelles: Punts resultants després del mostreig.

Figura 7.1 Interpolació i mostreig de les senyals $RR(n)$ i $Ttot(n)$



7.3. Metodologia de la Dinàmica Simbòlica

En aquest projecte s'estudien dos senyals temporals, la seva interacció i com es comporten independentment.

Es disposa del senyal de variació del ritme cardíac, RR, i del senyal de duració del cicle respiratori, Ttot. Aquests senyals han estat mesurats conjuntament, per tant, abans de la transformació es disposa de la sèrie de vectors x , equació 7.4:

$$x = \left\{ \left\{ x_n^{RR}, x_n^{Ttot} \right\} \right\}_{n=0,1,\dots} x \in \mathfrak{R} \quad (\text{Eq. 7.4})$$

Aquest vector després de la transformació es convertirà amb el vector s , equació 7.5 :

$$s = \left\{ \left\{ s_n^{RR}, s_n^{Ttot} \right\} \right\}_{n=0,1,\dots} s \in 0,1 \quad (\text{Eq. 7.5})$$

La transformació que s'utilitzarà per transformar un vector amb l'altre, per realitzar la codificació queda expressada en l'equació 7.6, [Baumert et al., 2002]:

$$s_n^{RR} = \begin{cases} 0 : (x_n^{RR} - x_{n+1}^{RR}) \leq I^{RR} \\ 1 : (x_n^{RR} - x_{n+1}^{RR}) > I^{RR} \end{cases} \quad s_n^{Ttot} = \begin{cases} 0 : (x_n^{Ttot} - x_{n+1}^{Ttot}) \leq I^{Ttot} \\ 1 : (x_n^{Ttot} - x_{n+1}^{Ttot}) > I^{Ttot} \end{cases} \quad (\text{Eq. 7.6})$$

$n=1, 2, 3, \dots, N$ N =nombre de mostres

x_n = valor en un instant de temps concret de cada sèrie temporal

La I^{xx} representa un valor llindar.

Aquest alfabet utilitza únicament dos símbols, 1 i 0. Amb aquesta transformació obtenim informació més simplificada i intuïtiva que amb els senyals reals de l'evolució del sistema. El símbol que es determinarà serà 1 si el senyal disminueix de valor entre dues mostres, i 0 si passa el contrari, que augmenta el seu valor. D'aquesta manera és codifica la sèrie temporal, tenint en compte la física essencial del sistema.

Un cop transformada la sèrie temporal, s'ha d'emmagatzemar la informació obtinguda del sistema no lineal de manera que es pugui utilitzar tota la informació obtinguda.



7.4. Paràmetres necessaris per aplicar la Dinàmica Simbòlica

Per aplicar la tècnica de la Dinàmica Simbòlica i per al tractament de les dades que s'extreuen de la seva aplicació fan falta una sèrie de paràmetres. En aquest apartat es farà una descripció de cadascun d'ells.

Els valors d'aquests paràmetres no es poden predir a priori, tot i que es coneixen valors aproximats d'estudis anteriors realitzats per altres investigadors. Es tractarà de trobar els millors valors per als quals s'obtinguin uns resultats més bons a l'hora de complir l'objectiu final, predir el moment de la desconexió del pacient de la màquina respiratòria.

Els paràmetres que s'utilitzaran són:

- **Nombre de símbols.** En l'apartat anterior, apartat 7.3, s'han presentat les equacions que fixaran la transformació que s'utilitzarà en aquest projecte. L'alfabet utilitzat està format per 2 símbols {0,1}. El nombre de símbols que s'utilitzaran depenen de les equacions que es determinin com les més indicades per analitzar la complexitat del sistema.
- **Paraules. Longitud de les paraules.** La unitat que s'utilitzarà per emmagatzemar la informació s'anomena paraula. Una paraula és un tros de seqüència de longitud donada, és a dir, per cada longitud de paraula existiran un nombre diferent de paraules que podran aparèixer. Aquest nombre be donat per la relació, $n^{\circ} \text{ paraules} = \text{simbols alfabet}^{\ell}$, on ℓ és la longitud de paraula. Partint de dades d'estudis anteriors, es decideix fixar la longitud de paraula a 3. Per tant, el nombre de paraules que s'utilitzaran serà 2^3 , hi hauran 8 paraules diferents en cada sèrie temporal.
- **Nombre de dades.** És el nombre de dades de la sèrie temporal sobre les que s'aplicarà la Dinàmica Simbòlica, el nombre de dades que seran tractades. S'intenten obtenir resultats que permetin discriminar entre els dos grups de pacients de la manera més eficient possible, i a poder ser amb el mínim temps de càlcul. Això s'aconsegueix tractant el mínim de dades per arribar als resultats òptims o millors.
- **Tau o solapament.** Al anar llegint els símbols de les sèries temporals s'ha de decidir com s'anirà avançant, és a dir, el nombre de símbols que compartiran dues paraules consecutives. Les paraules tenen una longitud de 3 símbols, per tan tau=0 o no



solapament implica que les paraules no compartiran cap símbol, $\tau=1$ en compartiran 1, i $\tau=2$ en compartiran 2. Com interessa que el temps de còmput sigui el més petit possible, s'intentarà que el solapament sigui mínim.

- **Llindar de probabilitat.** Un cop obtingudes les probabilitats d'aparèixer de cada paraula pot resultar interessant saber el nombre de paraules que tenen una probabilitat d'aparèixer superior a un determinat valor o més petit que un altre, Podria ser que el nombre de paraules variés depenent del nivell de complexitat de la sèrie i per tant es podria utilitzar fer diferenciar entre els grups. Els valors d'aquests llindars que s'utilitzaran en l'estudi de les sèries temporals relacionades seran el nombre de paraules que tenen una probabilitat d'aparèixer superior a 1%, 2%, 3%, 4% i 5%. Així com el nombre de paraules prohibides, aquelles que tenen una probabilitat d'aparèixer menor que 0,1%. I per a les sèries independents 5%, 10%, 15%, 20% i 25% i les prohibides seran les que la seva probabilitat d'aparèixer sigui menor que un 1%. Aquests valors són diferents quan parlem de l'estudi de la interacció entre les variables de quan ho fem de manera independent, ja que la probabilitat que apareixi una paraula interaccionant amb una altra es molt inferior de que ho faci de manera independent.

7.5. Descripció de l'algorisme

Per aplicar aquest mètode als senyals temporals i que el seu posterior processat permeti treure'n tota la informació possible de manera clara i ràpida s'ha desenvolupat un programa que s'executarà en l'entorn Matlab, que és un entorn de computació i desenvolupament per realitzar projectes amb grans nombres de càlcul i que permet la seva visualització gràfica.

El programa que s'utilitzarà es s'anomena 'DinamicaSimbolicaConjunta' i es troba a l'Annex D.

El programa segueix el següent esquema:

- Es busquen les dues sèries temporals per al pacient i. Les senyal que pertanyen a un mateix pacient tenen una part de l'identificador inicial igual, concretament una sèrie de 4 nombres. Per tal de realitzar-ho es creen unes variables on hi ha emmagatzemats els identificadors dels pacients RR i Ttot a tractar. Aquestes variables contindran els identificadors dels pacients les sèries temporals dels quals es trobin en les carpetes: file_RR i file_Ttot, respectivament. Es llegeix l'identificador inicial d'un pacient RR i es fa una cerca per els noms dels arxius Ttot fins trobar el coincident. Si d'un pacient només es disposa del senyal RR, no s'estudiarà, es



passarà al següent pacient. Per seleccionar les series temporal a processar s'utilitza la funció **obtenirdades.m**.

- Es realitzen les gràfiques de les series temporals inicials en funció del temps i del número de mostra i es guardaran a les carpetes: graf_ini_RR_t, graf_ini_Ttot_t, graf_ini_RR_n i graf_ini_Ttot_n, respectivament.
- S'interpolen linealment els senyals i es mostregen per disposar del seu valor en el mateix instant de temps per a les dues mostres. El temps de mostreig t , es una variable que es podrà fixar al valor desitjat, en aquest projecte s'utilitzarà $t=1$ (freqüència de mostreig 1 Hz) i $t=0,5$ (freqüència de mostreig 2 Hz). En aquest punt és important determinar correctament l'instant d'inici i final de les mostres, ja que es necessari que els valors de les sèries temporals existeixin per a les dues senyals i es produeixin en el mateix instant. Tot seguit es realitzen les gràfiques dels valors interpolats i es guarden a les carpetes graf_RR_i i graf_Ttot_i.
- Per aplicar la transformació de les dades s'utilitzarà la funció **codificar.m**, d'aquesta manera si es decidís canviar les equacions que s'utilitzen per a l'estudi de les dades només faria falta canviar aquesta funció per poder fer un nou estudi. Per tal de codificar la sèrie temporal es realitza un recorregut per tota ella i s'apliquen les equacions esmentades a l'apartat 7.3.
- El següent punt és calcular la probabilitat d'aparèixer de cada paraula. Es fixa la longitud de paraula ℓ , com ja s'ha comentat serà igual a 3, tot i que al tractar-se d'una variable pot ésser modificada si pogués resultar d'utilitat. Es crea una matriu on s'emmagatzemaran les probabilitats d'aparèixer de cada paraula. També en aquest punt es fixa el solapament o tau entre dues paraules consecutives, $s=3$, no hi ha solapament, $s=2$, un símbol de solapament, o $s=1$, les paraules tenen en comú dos símbols. Totes aquestes dades, més les sèries temporals i la seva longitud es passen a la funció **mcont.m**
- La funció mcont.m realitza un recorregut pels senyals a estudiar codificades. Compta els cops que apareix cada paraula de la sèrie RR conjuntament amb de la sèrie Ttot i ho guarda en una matriu, que té per files les paraules RR i per columnes les paraules Ttot. A la última fila d'aquesta matriu s'hi emmagatzemarà la suma de senyals RR que surten en combinació amb una determinada paraula de la sèrie Ttot, és a dir la probabilitat d'aparició de cada paraula Ttot. El mateix passarà a la última columna de la matriu, però a l'inrevés, apareixerà la suma de senyals Ttot per una paraula RR concreta, la probabilitat d'aparèixer de cada paraula RR.



- Per tal d'estudiar aquest resultat es grafiquen les probabilitats obtingudes cridant a l'acció **grafiquesfinal.m** que realitza els histogrames de la matriu de probabilitats per als senyals conjunts, més un histograma per als senyals RR i un altre per als Ttot, aquestes gràfiques es guarden dins les carpetes graf_mat_RR, graf_mat_Ttot i graf_mat, respectivament.
- L'últim pas consistirà en generar uns fitxers per a facilitar les posteriors proves estadístiques que es realitzaran. Es crida a la funció **treuredades.m** que guarda els identificadors de pacients, transforma la matriu en una taula i calcula les paraules majors que una certa probabilitat d'aparèixer així com les prohibides. També es aquí on es calcula l'entropia de Shannon.
- Tots aquestes passos es fan per cada pacient del qual es disposa de les dues senyals, quan ja s'ha fet per tots els pacients es guarden en arxius de dades els resultats, els identificadors dels pacients es guarden a l'arxiu pacients.dat, les probabilitats d'aparèixer de cada paraula RR relacionades amb les paraules de la sèrie Ttot es guarden a dades1.dat, les de les paraules RR i Ttot de manera separada es guarden a dades2.dat i el nombre de paraules amb una certa probabilitat major d'aparèixer, les paraules prohibides i la entropia de Shannon es guarden a l'arxiu dades3.dat.

7.6. Exemple d'aplicació.

Com ja s'ha comentat en els apartats anteriors s'estudiarà la interacció entre els dos senyals i cadascun dels senyals per separat.

S'aplicarà la tècnica de la Dinàmica Simbòlica al següent conjunt de dades:

- El senyal de variació del ritme cardíac, RR, i el senyal del cicle respiratori, Ttot, un cop interpolats i remostrejats tenen els següents valors.

| | | | | | | | | |
|----------|--------|---------|---------|---------|---------|---------|---------|---------|
| RR (ms) | 628,00 | 627,337 | 627,051 | 624,000 | 631,692 | 628,000 | 627,439 | 624,000 |
| Ttot (s) | 5,1600 | 5,1670 | 5,1740 | 5,1809 | 5,1879 | 5,1949 | 5,1481 | 5,0912 |

- Les dades s'aniran guardant en una matriu per fer-ne una representació clara i molt intuïtiva. Els senyals codificats són els següents, així com la seva representació a la matriu de probabilitats. Es faix $s=2$, i $t=1$.



| | | | | | | | |
|-------|---|---|---|---|---|---|---|
| RRc | 1 | 1 | 1 | 0 | 1 | 1 | 1 |
| Ttotc | 0 | 0 | 0 | 0 | 0 | 1 | 1 |

| RR/Ttot | 000 | 001 | 010 | 011 | 100 | 101 | 110 | 111 | Σ Ttot | RR |
|-------------|-----|-----|-----|-----|-----|-----|-----|-----|---------------|----|
| 000 | | | | | | | | | | |
| 001 | | | | | | | | | | |
| 010 | | | | | | | | | | |
| 011 | | | | | | | | | 1 | |
| 100 | | | | | | | | | | |
| 101 | 1 | | | | | | | | 1 | |
| 110 | 1 | | | | | | | | 1 | |
| 111 | 1 | | | 1 | | | | | 2 | |
| Σ RR | 3 | 1 | | 1 | | | | | | |
| Ttot | | | | | | | | | | |

En aquesta matriu s'emmagatzemaran el nombre de vegades que surt cada paraula RR amb combinació amb una Ttot al centre de la matriu i de les paraules independents a les columnes externes, Σ RR o Ttot, i Σ Ttot o RR. Dividint la matriu pel nombre de paraules tractades s'obté la probabilitat d'aparèixer cada paraula.

També s'obtenen histogrames, Figura 7.2, Figura 7.3 i Figura 7.4 que poden servir per un realitzar un anàlisi visual de les dades, que poden ajudar a una millor comprensió dels resultats.

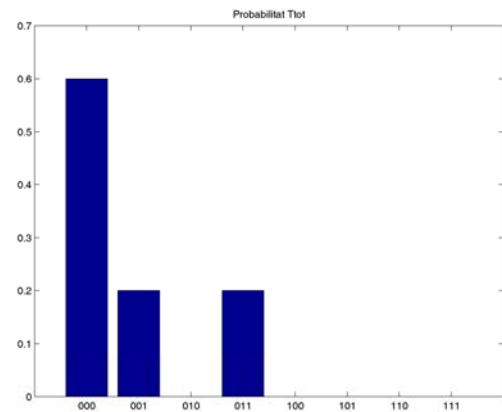
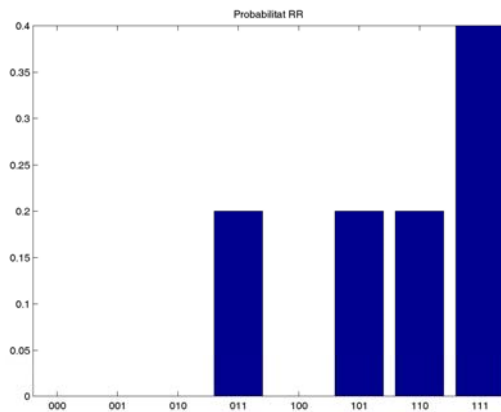


Figura 7.2 Histograma probabilitat aparició paraules de la sèrie temporal RR

Figura 7.3 Histograma probabilitat aparició paraules de la sèrie temporal Ttot



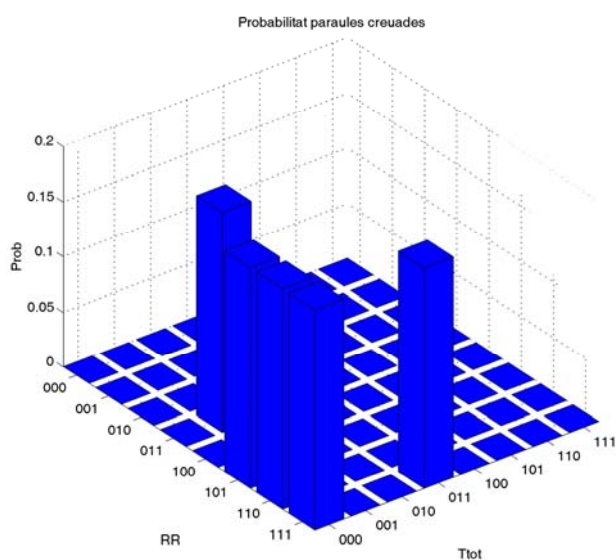


Figura 7.4 Histograma probabilitat aparició paraules de les sèries temporals RR i Ttot interactuant



8. Càlcul dels paràmetres

8.1. Introducció

En aquest apartat es fixarà el procediment per trobar els valors idonis per a obtenir els millors resultats possibles amb el mètode utilitzat, la Dinàmica Simbòlica. Com ja s'ha comentat en l'apartat anterior, alguns valors ja estaran fixats des del principi, ja que són el resultat d'estudis anteriors.

Per a la determinació d'aquests paràmetres, s'utilitzaran el programa que s'ha realitzat especialment i es realitzarà un processat de les dades que surten utilitzant mètodes estadístics, s'utilitzarà la U de Mann-Whitney comentada en l'apartat 5.4.6.

8.2. Metodologia per a la determinació dels paràmetres òptims

L'objectiu d'aquest projecte és veure si la Dinàmica Simbòlica és una eina vàlida a l'hora de prendre decisions per poder diferenciar entre dos grups de pacients els quals han respost de manera satisfactòria o no satisfactòria a la desconexió de la respiració artificial, és a dir trobar diferències en el patró dels senyals que genera cada grup.

Com ja s'ha comentat al capítol 5, s'han dividit els 39 pacients inicials en dos grups per realitzar l'estudi. El grup d'aprenentatge, 18 pacients, el grup de test, la resta, classificats tal com apareixen en la Taula 5.2. El mètode consistirà en determinar tots els paràmetres necessaris amb tots el pacients, és a dir posar a punt el mètode de la Dinàmica Simbòlica, i posteriorment amb els resultats obtinguts es generaran funcions discriminants per a poder distingir entre els dos nivells de complexitat, utilitzant únicament els pacients d'aprenentatge. Després aquestes funcions discriminants seran provades sobre el grup de test, Capítol 9.

La metodologia que es seguirà per determinar els paràmetres òptims d'aplicació de la Dinàmica Simbòlica és el primer que s'ha de determinar per tal de poder diferenciar entre els dos grups de pacients. La metodologia a seguir consistirà en determinar els nivells de significació de cadascuna de les variables que resulten de l'aplicació de la Dinàmica Simbòlica. Apareixeran quatre tipus de variables:

- Variables tipus paraula d'una sèrie temporal que tenen associada la seva probabilitat d'aparèixer en combinació amb les paraules de l'altra sèrie temporal.



- Variables de tipus paraula d'una sèrie temporal que tenen associada la seva probabilitat d'aparèixer de forma independent, és a dir sense tenir en compte amb quina paraula de l'altra sèrie apareixen a la vegada.
- Variables del tipus nombre de paraules per sobre o sota d'un cert llindar de probabilitat d'aparèixer per a paraules d'una sèrie temporal amb combinació amb l'altra sèrie.
- Variables del tipus nombre de paraules per sobre o sota d'un cert llindar de probabilitat d'aparèixer per a paraules d'una sèrie temporal sense tenir en compte l'altra sèrie.

La raó de determinar el nivell de significació de cada variable és que s'està intentant buscar una variable o un conjunt de variables que presenti valors diferents per a cada nivell de complexitat. D'aquesta manera, les variables que presentin un millor nivell de significació seran les candidates a formar les funcions discriminants que es generaran en el següent capítol.

Per tant, la metodologia a seguir consistirà en buscar els paràmetres òptims del mètode de la Dinàmica Simbòlica que permetin obtenir un major nombre de variables significatives i amb un millor nivell de significació a l'hora de diferenciar entre els dos grups de pacients.

8.3. Tècniques estadístiques

S'ha parlat de buscar paraules amb diferents nivells de significació depenent del grup de pacients al que pertanyin, és a dir diferenciar entre els dos nivells de complexitat de les senyals respiratòries i cardíaques. En aquest capítol es definiran les proves d'hipòtesi que s'utilitzaran per trobar els nivells de significació per a les totes les variables i les tècniques associades a aquests mètodes.

8.3.1. Realització de les proves d'hipòtesi

L'estratègia de la cerca de valors òptims per als paràmetres es centrarà en contrastar una hipòtesi nul·la (H_0) contra la hipòtesi alternativa (H_1). Es buscaran aquelles variables capaces de rebutjar la hipòtesi nul·la amb un major nivell de significació per tal d'assegurar que el risc d'error sigui el mínim possible. Per tant, es seleccionaran aquelles variables capaces de rebutjar la hipòtesi nul·la amb el millor nivell de significació.

El plantejament de la hipòtesi nul·la dependrà de la tècnica estadística que s'utilitzi. En general, la hipòtesi nul·la que s'intenta rebutjar consisteix en l'afirmació que existeixin



diferències significatives entre dos grups de dades, en el cas d'aquest projecte que hi hagi diferències en la complexitat de les sèries temporals per als dos grups de pacients.

En el capítol 5 s'han presentat diferents mètodes estadístics. S'ha fet una distinció entre mètodes estadístics paramètrics i no paramètrics, apartat 5.4.2. La principal diferència entre aquests dos grups de mètodes eren les hipòtesis que feien els paramètrics respecte a les condicions de les dades que s'analitzaven. S'ha demostrat que no sempre es complien les hipòtesis que requereix la seva aplicació, per tant, tal com s'ha realitzat al final del capítol i també d'ara en endavant s'utilitzaran els mètodes no paramètrics, i concretament s'utilitzarà la prova U de Mann-Whitney, descrita a l'apartat 5.4.6.

8.4. Determinació del paràmetres òptims mitjançant la prova U de Mann-Whitney

Els resultats de l'aplicació d'aquest mètode es troben a l'Annex D

8.4.1. Consideracions prèvies

En apartats anteriors s'ha fet una presentació del paràmetres que es necessita determinar, alguns vindran ja donats, procedents d'estudis anteriors capítol 7.4, altres seran determinats en aquest capítol. Els paràmetres són els següents:

- **Nombre de símbols.** 2 símbols {0,1}, fixats per les equacions que són utilitzades
- **Paraules. Longitud de les paraules.** La longitud de paraula serà igual a 3 símbols. El nombre de paraules diferents és de 8 en cada sèrie temporal
- **Nombre de dades.** Només es reduirà el nombre de dades a estudiar si s'arriba a un 100% de pacients ben classificats quan s'apliqui la funció discriminat al grup de test. Al reduir el nombre de dades també es redueix la informació que es té del sistema i per tant, no és útil perdre informació quan no s'aconsegueix l'objectiu fixat. Per tant, aquest paràmetre només s'utilitzarà si s'aconsegueix arribar a un 100% d'eficiència al classificar els pacients del grup de test.
- **Periode de mostreig.** S'estudiarà aquest paràmetre amb 2 valors, $t = 1s$ (freqüència de mostreig 1 Hz) i $t = 0,5s$ (freqüència de mostreig 2 Hz).
- **Tau o solapament.** Aquest paràmetre tindrà els següents valors: $s=1$, dos símbols de solapament entre paraules, $s=2$, un símbol de solapament entre les paraules, $s=3$, no hi ha solpament entre paraules.



- **Llindar de probabilitat.** Els valors d'aquests llindars que s'utilitzaran en l'estudi de les sèries temporals relacionades seran el nombre de paraules que tenen una probabilitat d'aparèixer superior a 1%, 2%, 3%, 4% i 5%. Així com el nombre de paraules prohibides, aquelles que tenen una probabilitat d'aparèixer menor que 0,1%. I per a les sèries independents 5%, 10%, 15%, 20% i 25% i les prohibides seran les que la seva probabilitat d'aparèixer sigui menor que un 1.

8.4.2. Resultats obtinguts

Les dades s'han generat amb el programa DinamicaSimbolicaConjunta.m, dades1.dat., dades2.dat, dades3.dat, apartat 7.5. Han estat gestionades i transformades en un fitxer de tipus xls Annex D Per a preparar-les per a ser manipulades amb el paquet estadístic SPSS. Amb aquest paquet s'ha realitzat la prova U de Mann-Whitney per trobar els nivells de significació de cada variable, així com la seva mitja i desviació tipus, que serveixen d'utilitat per fer un reconeixement ràpid i intuïtiu de les dades. Els resultats per a totes les combinacions de paràmetres que s'ha utilitzat, estan detallats a l'Annex D, queden resumits en les següents taules, Taula 8.1 i Taula 8.2, per als períodes $t=1s$ i $t=0,5s$ respectivament. En aquesta taula apareixen els valors de significació més petits que 0,005 de color taronja i de color blau els més petis que 0,01. Les variables amb aquests nivells de significació són les candidates per a generar les funcions discriminants.

t = 1

| s | Probabilitat | grup 1 | grup 2 | p-valor |
|---|--------------|-----------------|-----------------|---------|
| 1 | P(c001,r000) | 0,0305 ± 0,0146 | 0,0204 ± 0,0132 | 0,0057 |
| | P(c001,r111) | 0,0347 ± 0,0152 | 0,0221 ± 0,0105 | 0,0038 |
| | P(c010,r001) | 0,0199 ± 0,0093 | 0,0301 ± 0,0107 | 0,0092 |
| | P(c010,r010) | 0,0009 ± 0,0016 | 0,0038 ± 0,0036 | 0,0076 |
| | P(c010,r100) | 0,0203 ± 0,0108 | 0,0322 ± 0,0119 | 0,0025 |
| | P(c010,r110) | 0,0204 ± 0,0103 | 0,0307 ± 0,0084 | 0,0025 |
| | P(c100,r000) | 0,0306 ± 0,0147 | 0,0189 ± 0,0114 | 0,0057 |
| | P(c100,r111) | 0,0348 ± 0,0154 | 0,0232 ± 0,0070 | 0,0092 |
| | P(c101,r010) | 0,0014 ± 0,0021 | 0,0040 ± 0,0032 | 0,0063 |
| | P(c101,r101) | 0,0015 ± 0,0022 | 0,0045 ± 0,0034 | 0,0025 |
| | P(r010) | 0,0059 ± 0,0106 | 0,0192 ± 0,0178 | 0,0046 |
| | P(r011) | 0,1086 ± 0,0278 | 0,1333 ± 0,0177 | 0,0046 |
| | P(r101) | 0,0071 ± 0,0132 | 0,0223 ± 0,0197 | 0,0063 |
| | P(r110) | 0,1086 ± 0,0279 | 0,1333 ± 0,0175 | 0,0042 |
| | P(r111) | 0,2972 ± 0,0678 | 0,2336 ± 0,0496 | 0,0083 |
| | SHwr | 2,4422 ± 0,2209 | 2,6760 ± 0,1692 | 0,0022 |
| 2 | P(c001,r111) | 0,0346 ± 0,0165 | 0,0205 ± 0,0077 | 0,0092 |
| | P(c010,r001) | 0,0193 ± 0,0094 | 0,0308 ± 0,0138 | 0,0083 |



| | | | | |
|---|--------------|-----------------|-----------------|--------|
| | P(c010,r100) | 0,0194 ± 0,0111 | 0,0330 ± 0,0144 | 0,0069 |
| | P(c010,r110) | 0,0211 ± 0,0106 | 0,0302 ± 0,0077 | 0,0038 |
| | P(c011,r111) | 0,0388 ± 0,0137 | 0,0263 ± 0,0132 | 0,0083 |
| | P(c101,r101) | 0,0014 ± 0,0020 | 0,0043 ± 0,0035 | 0,0057 |
| | P(c111,r000) | 0,0170 ± 0,0088 | 0,0093 ± 0,0062 | 0,0083 |
| | P(r000) | 0,2581 ± 0,0710 | 0,1965 ± 0,0575 | 0,0083 |
| | P(r011) | 0,1059 ± 0,0296 | 0,1337 ± 0,0195 | 0,0031 |
| | P(r100) | 0,1041 ± 0,0281 | 0,1335 ± 0,0224 | 0,0025 |
| | SHwr | 2,4395 ± 0,2205 | 2,6701 ± 0,1723 | 0,0020 |
| 3 | P(c001,r000) | 0,0290 ± 0,0136 | 0,0190 ± 0,0131 | 0,0092 |
| | P(c001,r111) | 0,0358 ± 0,0185 | 0,0208 ± 0,0106 | 0,0069 |
| | P(c010,r001) | 0,0199 ± 0,0113 | 0,0310 ± 0,0116 | 0,0092 |
| | P(c101,r101) | 0,0010 ± 0,0019 | 0,0036 ± 0,0035 | 0,0008 |
| | P(r000) | 0,2570 ± 0,0723 | 0,2002 ± 0,0561 | 0,0092 |
| | P(r101) | 0,0066 ± 0,0116 | 0,0236 ± 0,0233 | 0,0076 |
| | SHwr | 2,4352 ± 0,2201 | 2,6618 ± 0,1727 | 0,0038 |

Taula 8.1 Nivells de significació a partir de la prova U de Mann-Whitney al variar el solapament entre les paraules, per a $t = 1$

t = 0,5

| s | Probabilitat | grup 1 | grup 2 | p-valor |
|---|--------------|------------------|------------------|---------|
| 1 | P(c001,r001) | 0,0077 ± 0,0028 | 0,0126 ± 0,0050 | 0,0008 |
| | P(c001,r110) | 0,0081 ± 0,0029 | 0,0128 ± 0,0043 | 0,0010 |
| | P(c010,r110) | 0,0030 ± 0,0028 | 0,0067 ± 0,0051 | 0,0042 |
| | P(c011,r001) | 0,0095 ± 0,0041 | 0,0136 ± 0,0042 | 0,0069 |
| | P(c011,r011) | 0,0089 ± 0,0040 | 0,0130 ± 0,0049 | 0,0069 |
| | P(c011,r100) | 0,0087 ± 0,0032 | 0,0134 ± 0,0039 | 0,0004 |
| | P(c100,r001) | 0,0081 ± 0,0030 | 0,0121 ± 0,0042 | 0,0013 |
| | P(c100,r100) | 0,0081 ± 0,0028 | 0,0115 ± 0,0028 | 0,0028 |
| | P(c100,r110) | 0,0079 ± 0,0031 | 0,0113 ± 0,0035 | 0,0069 |
| | P(c111,r000) | 0,0488 ± 0,0209 | 0,0267 ± 0,0160 | 0,0011 |
| | P(c111,r111) | 0,0563 ± 0,0272 | 0,0328 ± 0,0252 | 0,0057 |
| | P(c111) | 0,1386 ± 0,0582 | 0,0841 ± 0,0536 | 0,0051 |
| | P(r001) | 0,0576 ± 0,0187 | 0,0770 ± 0,0169 | 0,0022 |
| | P(r011) | 0,0576 ± 0,0188 | 0,0773 ± 0,0173 | 0,0022 |
| | P(r100) | 0,0576 ± 0,0187 | 0,0769 ± 0,0168 | 0,0025 |
| | P(r110) | 0,0576 ± 0,0188 | 0,0772 ± 0,0172 | 0,0025 |
| | Wcr1 | 22,5600 ± 7,7732 | 30,2857 ± 6,7985 | 0,0038 |



| | | | | |
|------|-----------------|------------------|------------------|--------|
| | SHwcr | 4,8055 ± 0,2180 | 5,0047 ± 0,2144 | 0,0046 |
| | SHwr | 1,9833 ± 0,1978 | 2,1837 ± 0,1525 | 0,0022 |
| 2 | P(c001,r001) | 0,0077 ± 0,0029 | 0,0124 ± 0,0057 | 0,0046 |
| | P(c001,r110) | 0,0083 ± 0,0034 | 0,0129 ± 0,0046 | 0,0013 |
| | P(c010,r110) | 0,0030 ± 0,0028 | 0,0063 ± 0,0038 | 0,0011 |
| | P(c011,r001) | 0,0096 ± 0,0039 | 0,0135 ± 0,0040 | 0,0083 |
| | P(c011,r100) | 0,0087 ± 0,0034 | 0,0130 ± 0,0045 | 0,0028 |
| | P(c100,r001) | 0,0083 ± 0,0030 | 0,0130 ± 0,0047 | 0,0018 |
| | P(c100,r100) | 0,0081 ± 0,0028 | 0,0122 ± 0,0036 | 0,0022 |
| | P(c111,r000) | 0,0481 ± 0,0204 | 0,0277 ± 0,0162 | 0,0025 |
| | P(c111,r111) | 0,0556 ± 0,0275 | 0,0331 ± 0,0265 | 0,0076 |
| | P(c111) | 0,1371 ± 0,0582 | 0,0854 ± 0,0548 | 0,0076 |
| | P(r001) | 0,0585 ± 0,0178 | 0,0784 ± 0,0165 | 0,0016 |
| | P(r011) | 0,0566 ± 0,0207 | 0,0758 ± 0,0186 | 0,0046 |
| | P(r100) | 0,0570 ± 0,0179 | 0,0770 ± 0,0195 | 0,0028 |
| | P(r110) | 0,0582 ± 0,0205 | 0,0771 ± 0,0155 | 0,0031 |
| | Wcr1 | 23,0400 ± 7,3455 | 30,5000 ± 6,3821 | 0,0034 |
| | SHwcr | 4,7936 ± 0,2195 | 4,9949 ± 0,2129 | 0,0038 |
| SHwr | 1,9822 ± 0,1982 | 2,1826 ± 0,1525 | 0,0025 | |
| 3 | P(c001,r001) | 0,0073 ± 0,0033 | 0,0117 ± 0,0037 | 0,0007 |
| | P(c001,r110) | 0,0078 ± 0,0027 | 0,0121 ± 0,0050 | 0,0016 |
| | P(c010,r110) | 0,0029 ± 0,0029 | 0,0068 ± 0,0056 | 0,0057 |
| | P(c011,r001) | 0,0089 ± 0,0041 | 0,0144 ± 0,0052 | 0,0057 |
| | P(c101,r110) | 0,0034 ± 0,0033 | 0,0080 ± 0,0049 | 0,0028 |
| | P(c110,r100) | 0,0099 ± 0,0049 | 0,0141 ± 0,0037 | 0,0076 |
| | P(c111,r000) | 0,0495 ± 0,0218 | 0,0276 ± 0,0176 | 0,0016 |
| | P(c111,r111) | 0,0566 ± 0,0290 | 0,0334 ± 0,0253 | 0,0092 |
| | P(c111) | 0,1396 ± 0,0607 | 0,0856 ± 0,0553 | 0,0063 |
| | P(r001) | 0,0586 ± 0,0195 | 0,0770 ± 0,0141 | 0,0011 |
| | P(r011) | 0,0568 ± 0,0192 | 0,0752 ± 0,0174 | 0,0083 |
| | P(r110) | 0,0558 ± 0,0192 | 0,0780 ± 0,0162 | 0,0004 |
| | Wcr1 | 23,0800 ± 7,0055 | 30,1429 ± 5,6002 | 0,0018 |
| | SHwcr | 4,7868 ± 0,2191 | 4,9858 ± 0,2111 | 0,0042 |
| SHwr | 1,9787 ± 0,1998 | 2,1789 ± 0,1430 | 0,0018 | |

Taula 8.2 Nivells de significació a partir de la prova U de Mann-Whitney al variar el solapament entre les paraules, per a t = 1



9. Anàlisi discriminant

9.1. Introducció

En aquest capítol es pretén trobat una eina capaç de diferenciar entre els diferents grups de pacients. Es persegueix que amb les resultats obtinguts amb aquesta eina quan es disposi de noves sèries temporals permeti decidir a quin grup de pacients pertany, que permeti classificar aquest 'pacient nou' dins el grup correcte. La eina que s'utilitzarà serà la tècnica estadística de la Funció Discriminant., que consisteix en fer combinacions lineals de variables independents per a generar funcions que puguin classificar cada cas en un dels nivells d'estudi.

En el capítol anterior, capítol 8, s'han trobat tota una sèrie de variables candidates a generar les funcions discriminants. Les variables candidates seran aquelles que hagin superat el nivell de significació fixat. L'estratègia consistirà en buscar funcions discriminants formades amb una única variable, això és equivalent a buscar un llindar, i funcions discriminants amb dues variables per aquells índexs que hagin donats millors nivells de significació.

El criteri que s'utilitzarà per avaluar les funcions discriminants serà maximitzar el nombre de pacient ben classificats, és a dir, es seleccionaran com a bones aquelles funcions que cometin el menor nombre d'error classificant els pacients.

Com ja s'ha comentat en els apartats anteriors, els pacients pertanyen a dos grups, al grup C1 pertanyen els pacients que han respost de manera satisfactòria a la desconexió, i que generalment les seves senyals presenten un nivell de complexitat major, i al grup C2 pertanyen els pacients que han hagut de tornar a ser connectats a la respiració artificial, ja que no han superat la prova de desconexió, i que en general les seves senyals presenten un nivell de complexitat menor.

S'han separat els pacients en dos grups, el grup d'aprenentatge i el grup de test. En aquest capítol es generaran funcions discriminants utilitzant els pacients del grups d'aprenentatge, que posteriorment serà estudiada la seva eficiència en els pacients del grup de test.

Un cop aplicat el mètode de la Dinàmica Simbòlica a les sèries temporals, es disposa d'un conjunt de variables, amb valors coneguts per una sèrie de casos, i es vol distingir entre dos grups mútuament excloents, el grup C1 i el grup C2. Per realitzar aquesta possible distinció s'utilitzarà l'Anàlisi Discriminant que consisteix en fer combinacions lineals de variables independents que serviran per com a base per a classificar els casos dins de cada grup de



complexitat. Els valors d'aquestes variable seleccionades per a cada pacient del grup d'aprenentatge seran el punt de partida per a l'anàlisi discriminant.

9.2. Fonament teòric de l'Anàlisi Discriminant

Els tests de significació i els estadístics descriptius aporten informació bàsica sobre les distribucions de les variables en cada grup i ajuden a identificar diferències entre ells.

En l'Anàlisi Discriminant es busca una regla per discriminar entre diferents grups, de manera que per qualsevol individu que es sap que procedeix d'un grup, es pogués aplicar la regla i assignar aquest individu al grup més adequat. L'objectiu es centra en definir una regla que permeti assignar un nou pacient a un d'aquests grups mesurant les mateixes variables.

Funció Discriminant

A l'Anàlisi Discriminant es forma una combinació lineal de les variables independents que serveixen com a base per assignar casos a cadascun dels grups. Així la informació continguda en múltiples variables independents queda resumida en un índex D, equació 9.1, que serà utilitzat per fer la separació entre grups. L'equació de la funció discriminant és la següent:

$$D = B_0 + B_1 \cdot X_1 + B_2 \cdot X_2 + \dots + B_p \cdot X_p \quad (\text{Eq. 9.1})$$

On les X_i són les variables independents i els B_i són els coeficients estimats. Com es pot veure l'equació de la funció discriminant és molt semblant a la regressió lineal múltiple.

Donat que la funció discriminant és una funció utilitzada per diferenciar el grup al que pertany en cada cas, en aquest cas el nivell de complexitat dels senyals, els grups diferiran en els valors de D. Així, els coeficients B_i seran escollits de manera que els valors de D difereixin tant com sigui possible entre els grups. El ratio que s'utilitzarà serà el que maximitzi la relació de l'equació 9.2, qualsevol altra combinació de variables tindrà un valor d'aquest ratio menor, [Armitage i Perry, 1997]:

$$\frac{\text{suma de quadrats de les diferències entre els grups}}{\text{suma de quadrats de les diferències dins els grups}} \quad (\text{Eq. 9.2})$$



Un cop obtinguts els coeficients es calcula el valor de D per cada cas, substituint els valors de les variables independents corresponents dins d'equació discriminant, equació 9.1. Es pot comprovar que la mitja dels valors de D per a tots el casos combinats serà 0 i la desviació tipus serà 1.

9.3. Aplicació de l'Anàlisi Discriminant

9.3.1. Anàlisi de les variables obtingudes utilitzant Dinàmica

Simbòlica Conjunta

Com s'ha comentat en apartats anteriors aplicarem el mètode de l'Anàlisi Discriminant sobre el grup de pacients d'aprenentatge, per després comprovar la seva eficiència sobre el grup de test. Per aplicar el mètode de la funció discriminant s'utilitzarà el paquet estadístic SPSS.

Els resultats de l'aplicació d'aquest mètode són una funció lineal calculada amb les sèries temporals que pertanyen als pacients del grup d'aprenentatge, i que posteriorment es aplicada sobre els pacients del grup de test. Les funcions discriminants considerant una variable obtingudes per a les variables seleccionades en l'apartat anterior i el seu resultat al aplicar-les al grup de test són les indicades a la Taula 9.1 i Taula 9.2. Es detalla el percentatge de pacients ben classificats pel grup C1, el percentatge de pacients ben classificats pel grup C2 i el percentatge de pacients ben classificats al grup de test.

t = 1

| s | Probabilitat | Funció Discriminant | Grup C1 | Grup C2 | Total |
|---|--------------|--------------------------------|--------------------|--------------------|--------------------|
| | | | % Ben classificats | % Ben classificats | % Ben classificats |
| 1 | P(c001,r000) | 68,2125 *P(c001,r000) -1,7572 | 52,9412 | 83,3333 | 60,8696 |
| | P(c001,r111) | 81,9037 *P(c001,r111) -2,3033 | 58,8235 | 100,0000 | 69,5652 |
| | P(c010,r001) | 115,0480 *P(c010,r001) -3,2980 | 76,4706 | 16,6667 | 60,8696 |
| | P(c010,r010) | 363,5320 *P(c010,r010) -0,8680 | 82,3529 | 50,0000 | 73,9130 |
| | P(c010,r100) | 93,7076 *P(c010,r100) -2,6726 | 82,3529 | 50,0000 | 73,9130 |
| | P(c010,r110) | 124,2789 *P(c010,r110) -3,4137 | 70,5882 | 33,3333 | 60,8696 |
| | P(c100,r000) | 71,6429 *P(c100,r000) -1,8331 | 47,0588 | 83,3333 | 56,5217 |
| | P(c100,r111) | 91,0824 *P(c100,r111) -2,5423 | 52,9412 | 66,6667 | 56,5217 |
| | P(c101,r010) | 317,6345 *P(c101,r010) -0,9480 | 82,3529 | 50,0000 | 73,9130 |
| | P(c101,r101) | 346,0069 *P(c101,r101) -1,0811 | 82,3529 | 50,0000 | 73,9130 |
| | P(r010) | 96,5345 *P(r010) -1,0065 | 82,3529 | 66,6667 | 78,2609 |
| | P(r011) | 42,3251 *P(r011) -5,1725 | 64,7059 | 83,3333 | 69,5652 |
| | P(r101) | 86,8022 *P(r101) -1,0147 | 82,3529 | 66,6667 | 78,2609 |
| | P(r110) | 42,4690 *P(r110) -5,1931 | 64,7059 | 83,3333 | 69,5652 |
| | P(r111) | 16,3263 *P(r111) -4,2128 | 64,7059 | 66,6667 | 65,2174 |



| | | | | | | | |
|---|--------------|----------|---------------|----------|---------|----------|---------|
| | SHwr | 5,5576 | *SHwr | -14,2292 | 58,8235 | 66,6667 | 60,8696 |
| 2 | P(c001,r111) | 90,7680 | *P(c001,r111) | -2,5458 | 64,7059 | 100,0000 | 73,9130 |
| | P(c010,r001) | 90,1580 | *P(c010,r001) | -2,5586 | 88,2353 | 33,3333 | 73,9130 |
| | P(c010,r100) | 97,1037 | *P(c010,r100) | -2,7277 | 76,4706 | 33,3333 | 65,2174 |
| | P(c010,r110) | 121,8208 | *P(c010,r110) | -3,3540 | 82,3529 | 33,3333 | 69,5652 |
| | P(c011,r111) | 75,8957 | *P(c011,r111) | -2,2671 | 64,7059 | 50,0000 | 60,8696 |
| | P(c101,r101) | 332,3796 | *P(c101,r101) | -1,0263 | 70,5882 | 50,0000 | 65,2174 |
| | P(c111,r000) | 129,8599 | *P(c111,r000) | -1,6073 | 70,5882 | 83,3333 | 73,9130 |
| | P(r000) | 17,5690 | *P(r000) | -4,0950 | 52,9412 | 83,3333 | 60,8696 |
| | P(r011) | 36,7572 | *P(r011) | -4,4348 | 64,7059 | 83,3333 | 69,5652 |
| | P(r100) | 40,5832 | *P(r100) | -4,8825 | 70,5882 | 83,3333 | 73,9130 |
| | SHwr | 5,5267 | *SHwr | -14,1326 | 58,8235 | 66,6667 | 60,8696 |
| 3 | P(c001,r000) | 72,7246 | *P(c001,r000) | -1,8624 | 52,9412 | 100,0000 | 65,2174 |
| | P(c001,r111) | 63,5874 | *P(c001,r111) | -1,8635 | 58,8235 | 100,0000 | 69,5652 |
| | P(c010,r001) | 87,0749 | *P(c010,r001) | -2,6225 | 82,3529 | 16,6667 | 65,2174 |
| | P(c101,r101) | 366,2773 | *P(c101,r101) | -0,8484 | 82,3529 | 50,0000 | 73,9130 |
| | P(r000) | 17,1169 | *P(r000) | -3,9946 | 52,9412 | 83,3333 | 60,8696 |
| | P(r101) | 94,6606 | *P(r101) | -1,0361 | 82,3529 | 66,6667 | 78,2609 |
| | SHwr | 5,5576 | *SHwr | -14,2292 | 64,7059 | 66,6667 | 65,2174 |

Taula 9.1 Funció Discriminant d'una variable, Resultats de la seva aplicació sobre el grup de pacients de test, per a $t=1s$.

t = 0,5

| s | Probabilitat | Funció Discriminant | Grup C1 | Grup C2 | Total | | |
|---|--------------|---------------------|--------------------|--------------------|--------------------|---------|---------|
| | | | % Ben classificats | % Ben classificats | % Ben classificats | | |
| 1 | P(c001,r001) | 263,5373 | *P(c001,r001) | -2,5579 | 76,4706 | 83,3333 | 78,2609 |
| | P(c001,r110) | 265,4475 | *P(c001,r110) | -2,7305 | 82,3529 | 66,6667 | 78,2609 |
| | P(c010,r110) | 348,6943 | *P(c010,r110) | -1,4858 | 88,2353 | 66,6667 | 82,6087 |
| | P(c011,r001) | 241,8701 | *P(c011,r001) | -2,9549 | 64,7059 | 50,0000 | 60,8696 |
| | P(c011,r011) | 223,8276 | *P(c011,r011) | -2,4042 | 76,4706 | 66,6667 | 73,9130 |
| | P(c011,r100) | 262,6803 | *P(c011,r100) | -2,9231 | 76,4706 | 83,3333 | 78,2609 |
| | P(c100,r001) | 333,4927 | *P(c100,r001) | -3,3483 | 76,4706 | 66,6667 | 73,9130 |
| | P(c100,r100) | 384,5520 | *P(c100,r100) | -3,7396 | 58,8235 | 50,0000 | 56,5217 |
| | P(c100,r110) | 316,9740 | *P(c100,r110) | -3,1270 | 76,4706 | 50,0000 | 69,5652 |
| | P(c111,r000) | 67,2856 | *P(c111,r000) | -2,3096 | 76,4706 | 66,6667 | 73,9130 |
| | P(c111,r111) | 45,8517 | *P(c111,r111) | -1,7709 | 70,5882 | 66,6667 | 69,5652 |
| | P(c111) | 22,8329 | *P(c111) | -2,2510 | 82,3529 | 66,6667 | 78,2609 |
| | P(r001) | 63,2505 | *P(r001) | -4,2042 | 70,5882 | 66,6667 | 69,5652 |
| | P(r011) | 63,2505 | *P(r011) | -4,2042 | 70,5882 | 66,6667 | 69,5652 |
| | P(r100) | 63,5061 | *P(r100) | -4,2200 | 70,5882 | 66,6667 | 69,5652 |
| | P(r110) | 63,5061 | *P(r110) | -4,2200 | 70,5882 | 66,6667 | 69,5652 |
| | Wcr1 | 0,1554 | *Wcr1 | -4,1096 | 76,4706 | 66,6667 | 73,9130 |
| | SHwcr | 4,7068 | *SHwcr | -23,0791 | 70,5882 | 83,3333 | 73,9130 |
| | SHwr | 6,2751 | *SHwr | -13,0591 | 64,7059 | 83,3333 | 69,5652 |
| | 2 | P(c001,r001) | 254,5090 | *P(c001,r001) | -2,5550 | 76,4706 | 50,0000 |



| | | | | | | | |
|---|--------------|----------|---------------|----------|---------|---------|---------|
| | P(c001,r110) | 226,6970 | *P(c001,r110) | -2,4270 | 82,3529 | 66,6667 | 78,2609 |
| | P(c010,r110) | 386,9990 | *P(c010,r110) | -1,6290 | 82,3529 | 83,3333 | 82,6087 |
| | P(c011,r001) | 300,1030 | *P(c011,r001) | -3,6660 | 76,4706 | 50,0000 | 69,5652 |
| | P(c011,r100) | 265,2140 | *P(c011,r100) | -2,8030 | 64,7059 | 83,3333 | 69,5652 |
| | P(c100,r001) | 293,7060 | *P(c100,r001) | -3,0830 | 82,3529 | 66,6667 | 78,2609 |
| | P(c100,r100) | 306,6340 | *P(c100,r100) | -3,0460 | 64,7059 | 66,6667 | 65,2174 |
| | P(c111,r000) | 68,4390 | *P(c111,r000) | -2,3440 | 70,5882 | 66,6667 | 69,5652 |
| | P(c111,r111) | 46,5710 | *P(c111,r111) | -1,7770 | 76,4706 | 66,6667 | 73,9130 |
| | P(c111) | 23,0800 | *P(c111) | -2,2720 | 82,3529 | 66,6667 | 78,2609 |
| | P(r001) | 64,3740 | *P(r001) | -4,3970 | 64,7059 | 83,3333 | 69,5652 |
| | P(r011) | 59,1050 | *P(r011) | -3,8190 | 70,5882 | 66,6667 | 69,5652 |
| | P(r100) | 58,6690 | *P(r100) | -3,8860 | 70,5882 | 83,3333 | 73,9130 |
| | P(r110) | 64,2000 | *P(r110) | -4,2790 | 76,4706 | 66,6667 | 73,9130 |
| | Wcr1 | 0,1630 | *Wcr1 | -4,4140 | 82,3529 | 66,6667 | 78,2609 |
| | SHwcr | 4,6910 | *SHwcr | -22,9620 | 70,5882 | 83,3333 | 73,9130 |
| | SHwr | 6,2730 | *SHwr | -13,0500 | 64,7059 | 66,6667 | 65,2174 |
| 3 | P(c001,r001) | 312,6000 | *P(c001,r001) | -2,9450 | 70,5882 | 83,3333 | 73,9130 |
| | P(c001,r110) | 241,9120 | *P(c001,r110) | -2,2920 | 70,5882 | 66,6667 | 69,5652 |
| | P(c010,r110) | 279,6001 | *P(c010,r110) | -1,3086 | 82,3529 | 66,6667 | 78,2609 |
| | P(c011,r001) | 250,2180 | *P(c011,r001) | -3,0300 | 64,7059 | 33,3333 | 56,5217 |
| | P(c101,r110) | 254,8910 | *P(c101,r110) | -1,4210 | 70,5882 | 50,0000 | 65,2174 |
| | P(c110,r100) | 246,4770 | *P(c110,r100) | -3,0490 | 64,7059 | 66,6667 | 65,2174 |
| | P(c111,r000) | 61,6290 | *P(c111,r000) | -2,1590 | 76,4706 | 66,6667 | 73,9130 |
| | P(c111,r111) | 42,8630 | *P(c111,r111) | -1,6920 | 76,4706 | 66,6667 | 73,9130 |
| | P(c111) | 21,2740 | *P(c111) | -2,1250 | 76,4706 | 66,6667 | 73,9130 |
| | P(r001) | 74,3130 | *P(r001) | -4,9950 | 76,4706 | 66,6667 | 73,9130 |
| | P(r011) | 68,8780 | *P(r011) | -4,4050 | 64,7059 | 83,3333 | 69,5652 |
| | P(r110) | 71,3020 | *P(r110) | -4,7100 | 70,5882 | 66,6667 | 69,5652 |
| | Wcr1 | 0,1930 | *Wcr1 | -5,0600 | 70,5882 | 66,6667 | 69,5652 |
| | SHwcr | 4,8580 | *SHwcr | -23,7240 | 76,4706 | 83,3333 | 78,2609 |
| | SHwr | 6,8710 | *SHwr | -14,2620 | 64,7059 | 83,3333 | 69,5652 |

Taula 9.2 Funció Discriminant d'una variable, Resultats de la seva aplicació sobre el grup de pacients de test, per a $t=0,5$ s.

Analitzant les taules 9.1, i 9,2 es treuen les següent conclusions:

- Per al període de mostreig més petit ($t=0,5$) hi ha un nombre més elevat de paràmetres amb significació estadística per a diferenciar els dos pacients.
- A menor tau (s) més elevat és el nombre de paraules amb significació estadística.



- Els paràmetres vinculats a una única sèrie temporal (cardíaca o respiratòria) són més significatius quan corresponen a les sèries temporals respiratòries.

La Taula 9.3 i la Taula 9.4 presenten els resultats considerat funcions discriminants de dues variables. Per generar aquestes funcions s'han considerat aquells índexs que tenien els millors nivells de significació.

t = 1

| s | Variabes a la Funció Discriminant | % grup 1 ben classificats | % grup 2 ben classificats | % grup test ben classificats |
|-------------|-----------------------------------|---------------------------|---------------------------|------------------------------|
| 1 | P(c001,r111)P(c101,r101) | 82,3529 | 50,0000 | 73,9130 |
| | P(c001,r111)P(r011) | 70,5882 | 83,3333 | 73,9130 |
| | P(c001,r111)P(r110) | 70,5882 | 83,3333 | 73,9130 |
| | P(c001,r111)SHwr | 70,5882 | 66,6667 | 69,5652 |
| | P(c010,r100)P(c101,r101) | 76,4706 | 66,6667 | 73,9130 |
| | P(c010,r100)P(r011) | 82,3529 | 50,0000 | 73,9130 |
| | P(c010,r100)P(r110) | 82,3529 | 50,0000 | 73,9130 |
| | P(c010,r100)PSHwr | 100,0000 | 0,0000 | 73,9130 |
| | P(c010,r110)SHwr | 76,4706 | 33,3333 | 65,2174 |
| | P(c101,r101)P(r011) | 76,4706 | 66,6667 | 73,9130 |
| | P(c101,r101)P(r110) | 76,4706 | 66,6667 | 73,9130 |
| | P(c101,r101)SHwr | 82,3529 | 50,0000 | 73,9130 |
| | P(r011)P(r110) | 64,7059 | 83,3333 | 69,5652 |
| | P(r011)SHwr | 76,4706 | 66,6667 | 73,9130 |
| P(r110)SHwr | 76,4706 | 66,6667 | 73,9130 | |
| 2 | P(c010,r110)P(r011) | 76,4706 | 50,0000 | 69,5652 |
| 3 | P(c101,r101)SHwr | 76,4706 | 66,6667 | 73,9130 |

Taula 9.3 Funció Discriminant d'una variable, Resultats de la seva aplicació sobre el grup de pacients de test, per a t=1s.

t = 0,5

| s | Variabes a la Funció Discriminant | % grup 1 ben classificats | % grup 2 ben classificats | % grup test ben classificats |
|---|-----------------------------------|---------------------------|---------------------------|------------------------------|
| 1 | P(c001,r001)P(c011,r100) | 76,4706 | 83,3333 | 78,2609 |
| | P(c001,r001)P(c111,r000) | 82,3529 | 66,6667 | 78,2609 |
| | P(c001,r110)P(c011,r100) | 76,4706 | 100,0000 | 82,6087 |
| | P(c001,r110)P(c111,r000) | 88,2353 | 66,6667 | 82,6087 |
| | P(c011,r100)P(c111,r000) | 94,1176 | 66,6667 | 86,9565 |
| | P(c100,r001)P(c111,r000) | 88,2353 | 50,0000 | 78,2609 |
| 2 | P(c001,r110)P(c010,r110) | 82,3529 | 83,3333 | 82,6087 |



| | | | | |
|---|--------------------------|---------|----------|---------|
| | P(c001,r110)P(c100,r001) | 82,3529 | 66,6667 | 78,2609 |
| | P(c100,r001)P(r001) | 82,3529 | 66,6667 | 78,2609 |
| 3 | P(c001,r001)P(c001,r110) | 76,4706 | 100,0000 | 82,6087 |
| | P(c001,r001)P(c111,r000) | 88,2353 | 66,6667 | 82,6087 |
| | P(c001,r001)P(r001) | 70,5882 | 83,3333 | 73,9130 |
| | P(c001,r001)P(r011) | 70,5882 | 83,3333 | 73,9130 |
| | P(c001,r001)P(r110) | 76,4706 | 66,6667 | 73,9130 |
| | P(c001,r001)PWcr1 | 70,5882 | 83,3333 | 73,9130 |
| | P(c001,r001)PSHwr | 76,4706 | 66,6667 | 73,9130 |
| | P(c001,r110)P(c111,r000) | 76,4706 | 100,0000 | 82,6087 |
| | P(c001,r110)P(r001) | 76,4706 | 83,3333 | 78,2609 |
| | P(c001,r110)P(r011) | 82,3529 | 100,0000 | 86,9565 |
| | P(c001,r110)P(r110) | 76,4706 | 100,0000 | 82,6087 |
| | P(c001,r110)PWcr1 | 76,4706 | 83,3333 | 78,2609 |
| | P(c001,r110)PSHwr | 76,4706 | 83,3333 | 78,2609 |
| | P(c111,r000)P(r001) | 82,3529 | 66,6667 | 78,2609 |
| | P(c111,r000)P(r011) | 88,2353 | 66,6667 | 82,6087 |
| | P(c111,r000)P(r110) | 82,3529 | 66,6667 | 78,2609 |
| | P(c111,r000)PWcr1 | 76,4706 | 66,6667 | 73,9130 |
| | P(c111,r000)PSHwr | 82,3529 | 66,6667 | 78,2609 |
| | P(r001)P(r011) | 76,4706 | 66,6667 | 73,9130 |
| | P(r001)PSHwr | 76,4706 | 66,6667 | 73,9130 |
| | P(r011)P(r110) | 82,3529 | 83,3333 | 82,6087 |
| | P(r011)PWcr1 | 76,4706 | 66,6667 | 73,9130 |
| | P(r011)PSHwr | 70,5882 | 66,6667 | 69,5652 |
| | P(r110)PWcr1 | 70,5882 | 66,6667 | 69,5652 |
| | P(r110)PSHwr | 82,3529 | 83,3333 | 82,6087 |
| | PWcr1PSHwr | 70,5882 | 66,6667 | 69,5652 |

Taula 9.4 Funció Discriminant de dos variables, Resultats de la seva aplicació sobre el grup de pacients de test, per a $t=1s$.

Analitzant les taules 9.3, i 9,4 es treuen les següent conclusions:

- En el període de mostreig $t=1s$, no s'aprecien millores considerables al realitzar l'anàlisi discriminant amb dues variables. Els percentatges de pacients ben classificats ja s'obtenien amb les funcions discriminants d'una variable.
- En el període de mostreig $t=0,5$, al realitzar l'anàlisi discriminant amb dues variables s'obté el nombre màxim de pacients ben classificat amb un 86,96% de pacients correctament classificats.



9.3.2. Anàlisi de les variables obtingudes amb l'estudi lineal

Amb la finalitat de comparar els resultats obtinguts amb el mètode de la Dinàmica Simbòlica amb les tècniques lineals en aquest apartat es realitzarà l'anàlisi discriminant a les variable lineals de les sèries temporals.

A l'apartat 5.4.6 s'ha realitzat la prova U de Mann-Whitney a les variables lineals de les sèries temporals, els resultats apareixien a la Taula 5.6. A partir d'aquesta taula s'arriba a la conclusió que la única variable que pot arribar a diferenciar correctament entre els grups es la MitjaTtot, que és la única que té un bon nivell de significació, menor que un 0,05%

Els resultats es mostren a la Taula 9.5:

| | Grup C1 | Grup C2 | Total |
|---------------------------|--------------------|--------------------|--------------------|
| Funció Discriminant | % Ben classificats | % Ben classificats | % Ben classificats |
| 1,6176 *mitjaTtot -3,9768 | 52,9412 | 83,3333 | 60,8696 |

Taula 9.5 Funció Discriminant de la variable lineal mitjaTtot

Observant aquest resultat s'arriba la conclusió que la Dinàmica Simbòlica és un eina millor per a discriminar entre els dos grups que les variables lineals, degut a que el nombre d'índexs amb significació estadística que apareixen és més elevat i l'anàlisi discriminat dona ba a millors resultats



10. Estudi de la variabilitat del patró respiratori

10.1. Introducció

Fins aquests moments en aquest projecte s'ha estudiat la interacció entre dues variables, per predir el moment de la desconexió dels pacients de la respiració assistida. Per tal de comprovar la necessitat de l'estudi de les dues variables o si seria suficient l'estudi de només la variable que correspon al cicle respiratori, s'ha realitzat aquí l'estudi més detallat d'aquesta variable amb dos alfabetos diferents, és a dir utilitzant dues transformacions diferents en l'aplicació del mètode de la Dinàmica Simbòlica.

10.2. Mètodes

El mètode que s'utilitzarà en aquest apartat serà molt similar al de l'apartat anterior. S'ha creat un programa, `DinàmicaSimblioca.m`, que seguirà el següent esquema:

- Es llegeix l'identificador del primer pacient T_{tot} de la carpeta `file_Ttot`. Per seleccionar les sèries temporals a processar s'utilitza la funció **obtenirdades.m**.
- Es realitzen les gràfiques de la sèrie temporal inicial en funció del temps i del número de mostra i es guardaran a les carpetes `graf_ini_Ttot_t` i `graf_ini_Ttot_n`, respectivament.
- Per aplicar la transformació de les dades s'utilitzarà la funció **codificar.m**, d'aquesta manera per aplicar una transformació o una altra només serà necessari substituir aquesta funció en el programa principal.
- El següent punt és calcular la probabilitat d'aparèixer de cada paraula. Es fixa la longitud de paraula l , com ja s'ha comentat serà igual a 3, tot i que al tractar-se d'una variable pot ésser modificada si pogués resultar d'utilitat. Es crea una variable on s'emmagatzemaran les probabilitats d'aparèixer de cada paraula. També en aquest punt es fixa el solapament o tau entre dues paraules consecutives, $s=2, 1$ o 0 . Totes aquestes dades, més les sèries temporals i la seva longitud es passen a la funció **mcont.m**
- La funció `mcont.m` realitza un recorregut per el senyal a estudiar codificades. Conta els cops que apareix cada paraula de la sèrie T_{tot} i ho guarda a la variable.



- Per tal d'estudiar aquest resultat es grafiquen les probabilitats obtingudes cridant a l'acció **grafiquesfinal.m** que realitza l'histograma i els guarda a l'arxiu graf_mat_Ttot.
- L'últim pas consistirà en generar uns fitxers per a facilitar les posteriors proves estadístiques que es realitzaran. Es crida a la funció **treuredades.m** que guarda els identificadors de pacients, guarda els valors de la variable on s'emmagatzemen les probabilitats d'aparèixer de cada paraula i calcula les paraules majors que una certa probabilitat d'aparèixer així com les prohibides. També es aquí on es calcula l'entropia de Shannon.
- Tots aquestes passos es fan per cada pacient, es guarden en arxius de dades els resultats, els identificadors del pacients es guarden a l'arxiu pacients.dat, les probabilitats d'aparèixer de cada paraula Ttot a l'arxiu dades1.dat, les el nombre de paraules amb una certa probabilitat major d'aparèixer, les paraules prohibides i la entropia de Shannon es guarden a l'arxiu dades2.dat.

S'observa com a diferència en aquesta codificació que aquí les dades no són interpolades, ja que com tal sols s'analitza una senyal, no fa falta tenir-ne en un instant de temps determinat el seu valor.

Els resultats de l'aplicació de la Dinàmica Simbòlica complets es troben a l'Annex C.

10.3. Alfabet Θ

La primera transformació que s'utilitzarà és la següent:

$$S_n = \begin{cases} 0 & \text{si} & b_n > (1 + \alpha/2)\mu \\ 1 & \text{si} & (1 - \alpha/2)\mu < b_n \leq (1 + \alpha/2)\mu \\ 2 & \text{si} & b_n \leq (1 - \alpha/2)\mu \end{cases} \quad (\text{Eq. 10.1})$$

μ = valor mig de la sèrie temporal

En aquesta transformació és té en compte el valor mig de cada mostra, i com es desvien els diferents valors d'aquest valor mig. En aquesta codificació un dels paràmetres de més importància i més crític a l'hora de ser escollit és el paràmetre α . En aquest estudi s'utilitzarà $\alpha = 0,5$, valor que es va donar com a òptim en estudis anteriors [Caminal et al., 2004]

Després d'aplicar la transformació a les sèries temporals, els arxius generats han estat transformats en un .xls, per al seu millor processament. Un cop realitzat aquest pas, s'ha



realitzat la prova d'hipòtesis U de Mann-Whitney per decidir quines variables eren les indicades per realitzar l'Anàlisi Discriminant. Els resultats apareixen a la Taula 10.1

a=0,5

| s | Probabilitat | grup 1 | grup 2 | p-valor |
|---|--------------|-----------------|-----------------|---------|
| 1 | P(r201) | 0,0028 ± 0,0053 | 0,0009 ± 0,0028 | 0,0379 |
| 2 | P(r112) | 0,0227 ± 0,0194 | 0,0114 ± 0,0130 | 0,0583 |
| 3 | P(r112) | 0,0238 ± 0,0208 | 0,0102 ± 0,0112 | 0,0351 |
| | P(r211) | 0,0214 ± 0,0157 | 0,0105 ± 0,0094 | 0,0326 |

Taula 10.1 Nivells de significació a partir de la prova U de Mann-Whitney al variar el solapament entre les paraules, per a l'alfabet Θ

S'observa que per aquest alfabet Ω els resultats obtinguts són molt pitjors que per a l'estudi anterior en que s'estudiava la interacció entre les variables. Com no s'ha trobat cap índex amb un nivell de significació inferior al 0,05%, s'ha decidit aplicar l'Anàlisi Discriminant amb les variables que tenen un índex de significació inferior al 5%. Els resultats obtinguda després de calcular la Funció Discriminant amb el grup d'aprenentatge, i utilitzar les funcions generades amb el grup de pacients de test apareixen a la Taula 10.2:

a=0,5

| s | Probabilitat | Funció Discriminant | Grup C1 | Grup C2 | Total |
|---|--------------|---------------------------|--------------------|--------------------|--------------------|
| | | | % Ben classificats | % Ben classificats | % Ben classificats |
| 1 | P(r201) | 191,2224 *P(r201) -0,3205 | 41,1765 | 66,6667 | 47,8261 |
| 2 | P(r112) | 66,0601 *P(r112) -1,2476 | 35,2941 | 50,0000 | 39,1304 |
| 3 | P(r112) | 74,6272 *P(r112) -1,2177 | 52,9412 | 50,0000 | 52,1739 |
| | P(r211) | 65,0435 *P(r211) -1,1443 | 52,9412 | 83,3333 | 60,8696 |

Taula 10.2 Funció Discriminant, Resultats de la seva aplicació sobre el grup de pacients de test, per a l'alfabet Θ .



10.4. Alfabet Ω

La segona transformació és:

$$S_n \begin{cases} 0 & \text{si } b_n > (1 + \alpha)\mu \\ 1 & \text{si } \mu < b_n \leq (1 + \alpha)\mu \\ 2 & \text{si } (1 - \alpha)\mu < b_n \leq \mu \\ 3 & \text{si } b_n \leq (1 - \alpha)\mu \end{cases} \quad (\text{Eq. 10.2})$$

μ = valor mig de la sèrie temporal

En aquesta transformació és té en compte el valor mig de cada mostra, i com es desvien els diferents valors d'aquest valor mig, al igual que en el cas anterior. En aquesta codificació el paràmetre α torna a ser decisiu per a l'aplicació del mètode. En aquest estudi s'utilitzarà $\alpha = 0,2$ i $\alpha = 0,25$ valor que es va donar com a òptim en estudis anteriors, [Caminal et al., 2004]

Després d'aplicar la transformació a les sèries temporals, els arxius generats han estat transformats en un .xls, per al seu millor processament. Un cop realitzat aquest pas, s'ha realitzat la prova d'hipòtesis U de Mann-Whitney per decidir quines variables eren les indicades per realitzar l'Anàlisi Discriminant. Els resultats es troben a la Taula 10.3 i la Taula 10.4:

a=0,2

| s | Probabilitat | grup 1 | grup 2 | p-valor |
|---|--------------|-----------------|------------------|---------|
| 1 | P(r113) | 0,0045 ± 0,0048 | 0,0007 ± 0,0011 | 0,0051 |
| | P(r222) | 0,1855 ± 0,0864 | 0,3645 ± 0,1587 | 0,0002 |
| | Wr25 | 0,2800 ± 0,4583 | 0,7857 ± 0,4258 | 0,0083 |
| 2 | P(r123) | 0,0051 ± 0,0050 | 0,0016 ± 0,0034 | 0,0042 |
| | P(r222) | 0,1861 ± 0,0859 | 0,3625 ± 0,1610 | 0,0002 |
| | Wr25 | 0,2800 ± 0,4583 | 0,7857 ± 0,4258 | 0,0083 |
| 3 | P(r210) | 0,0080 ± 0,0072 | 14,7857 ± 6,3267 | 0,0057 |
| | P(r213) | 0,0042 ± 0,0038 | 0,0010 ± 0,0017 | 0,0057 |
| | P(r222) | 0,1840 ± 0,0883 | 0,3650 ± 0,1577 | 0,0002 |
| | Wr15 | 0,9600 ± 0,7348 | 1,6429 ± 0,4972 | 0,0092 |
| | Wr25 | 0,2800 ± 0,4583 | 0,7857 ± 0,4258 | 0,0083 |

Taula 10.3 Nivells de significació a partir de la prova U de Mann-Whitney al variar el solapament entre les paraules, per a l'alfabet Ω i $\alpha = 0,2$



a=0,25

| s | Probabilitat | grup 1 | grup 2 | p-valor |
|---|--------------|-----------------|-----------------|---------|
| 1 | P(r113) | 0,0034 ± 0,0039 | 0,0006 ± 0,0009 | 0,0063 |
| | P(r213) | 0,0027 ± 0,0031 | 0,0006 ± 0,0010 | 0,0083 |
| | P(r222) | 0,2206 ± 0,0987 | 0,4045 ± 0,1553 | 0,0003 |
| | SHwr | 3,7472 ± 0,6425 | 3,0521 ± 0,6902 | 0,0051 |
| 2 | P(r222) | 0,2190 ± 0,0987 | 0,4018 ± 0,1575 | 0,0003 |
| | SHwr | 3,6986 ± 0,6358 | 3,0256 ± 0,6740 | 0,0076 |
| 3 | P(r213) | 0,0037 ± 0,0037 | 0,0005 ± 0,0012 | 0,0025 |
| | P(r222) | 0,2192 ± 0,1003 | 0,4043 ± 0,1554 | 0,0002 |
| | SHwr | 3,6520 ± 0,6051 | 3,0017 ± 0,6659 | 0,0069 |

Taula 10.4 Nivells de significació a partir de la prova U de Mann-Whitney al variar el solapament entre les paraules, per a l'alfabet Ω i $\alpha = 0,25$

S'observa que amb l'alfabet Ω s'obté millor diferència estadística entre els dos grups C1 i C2 que amb l'alfabet Θ . El nombre de variables amb bona significació estadística són semblants per als valors de $a=0,2$ i $a=0,25$. Els resultats obtinguts al realitzar l'Anàlisi Discriminant es presenten a la Taula 10.5 i la Taula 10.6.

a=0,2

| s | Probabilitat | Funció Discriminant | Grup C1 | Grup C2 | Total |
|---|--------------|---------------------------|--------------------|--------------------|--------------------|
| | | | % Ben classificats | % Ben classificats | % Ben classificats |
| 1 | P(r113) | 255,6750 *P(r113) -0,9180 | 35,2941 | 100,0000 | 52,1739 |
| | P(r222) | 7,9250 *P(r222) -1,9860 | 76,4706 | 83,3333 | 78,2609 |
| | wr25 | 2,1602 *wr25 -1,0801 | 70,5882 | 83,3333 | 73,9130 |
| 2 | P(r123) | 197,4484 *P(r123) -0,8293 | 41,1765 | 100,0000 | 56,5217 |
| | P(r222) | 7,8478 *P(r222) -1,9568 | 82,3529 | 83,3333 | 82,6087 |
| | wr25 | 2,1602 *wr25 -1,0801 | 70,5882 | 83,3333 | 73,9130 |
| 3 | P(r210) | 157,2220 *P(r210) -0,9110 | 52,9412 | 83,3333 | 60,8696 |
| | P(r213) | 459,5040 *P(r213) -1,2660 | 58,8235 | 66,6667 | 82,6087 |
| | P(r222) | 7,7750 *P(r222) -1,9340 | 82,3529 | 83,3333 | 82,6087 |
| | wr15 | 1,9010 *wr15 -2,0200 | 64,7059 | 66,6667 | 65,2174 |
| | wr25 | 2,1600 *wr25 -1,0800 | 70,5882 | 83,3333 | 73,9130 |

Taula 10.5 Funció Discriminant, Resultats de la seva aplicació sobre el grup de pacients de test, per a l'alfabet Ω , per a $\alpha=0,2$



a=0,25

| s | Probabilitat | Funció Discriminant | Grup C1 % Ben classificats | Grup C2 % Ben classificats | Total % Ben classificats |
|---|--------------|---------------------------|-------------------------------|-------------------------------|-----------------------------|
| 1 | P(r113) | 329,1880 *P(r113) -0,7960 | 35,2941 | 83,3333 | 47,8261 |
| | P(r213) | 312,2070 *P(r213) -0,6470 | 41,1765 | 66,6667 | 47,8261 |
| | P(r222) | 7,9080 *P(r222) -2,3150 | 82,3529 | 83,3333 | 82,6087 |
| | SHwr | 1,4948 *SHwr -5,1483 | 47,0588 | 66,6667 | 52,1739 |
| 2 | P(r222) | 7,9009 *P(r222) -2,2986 | 82,3529 | 83,3333 | 82,6087 |
| | SHwr | 1,5007 *SHwr -5,1162 | 52,9412 | 66,6667 | 56,5217 |
| 3 | P(r213) | 418,4470 *P(r213) -0,9550 | 58,8235 | 83,3333 | 65,2174 |
| | P(r222) | 7,7390 *P(r222) -2,2480 | 82,3529 | 83,3333 | 82,6087 |
| | SHwr | 1,5336 *SHwr -5,1862 | 41,1765 | 66,6667 | 47,8261 |

Taula 10.6 Funció Discriminant, Resultats de la seva aplicació sobre el grup de pacients de test, per a l'alfabet Ω , per a $\alpha=0,2$

La variable que aconseguix una millor discriminació és P(r222), tant per als valors de $a=0,2$ com per a $a=0,25$, resultant un 82,6087 % de pacients ben classificats en el grups de tests per quasi totes les combinacions de paràmetres.

Utilitzant els alfabetos Ω_i i Θ no s'obtenen funcions discriminants de dues variables amb resultats sensiblement millors als de la Taula 10.5 i la Taula 10.6.



11. Conclusions i extensions futures

11.1. Introducció

En aquest capítol es farà una explicació de les principals conclusions a les que s'ha arribat mitjançant la realització d'aquest projecte, així com d'algunes de les possibles extensions que es proposen realitzar en el futur.

11.2. Conclusions

En aquest projecte s'han analitzat les interaccions que existeixen entre la variabilitat del ritme cardíac i la duració del cycle respiratori i la seva aplicació en la determinació del instant òptim de desconnexió d'un pacient de la ventilació mecànica. La sèrie temporal de duració del cycle respiratori s'han obtingut mitjançant la tècnica de la pneumotacografia, un mètode que mesura el flux d'aire que intervé en la respiració. El senyal de variabilitat del ritme cardíac s'ha mesurat a partir del senyal electrocardiogràfic.

Per realitzar aquest projecte s'ha disposat dels dos senyals anteriors de 39 pacients que han estat sotmesos a la prova de tub en T per a la seva possible desconnexió de la respiració mecànica. Els senyals han estat classificats en dos grups: el grup C1, pacients que han superat la prova de tub en T i la desconnexió ha estat possible, i el grup C2, pacients que no han superat la prova i ha fet falta reintubar-los amb les complicacions que això comporta. El moment de la desconnexió de la màquina ha estat determinat a partir de l'experiència dels responsables clínics d'aquests pacients. En aquest projecte es busca una eina que ajudi a determinar aquest instant amb tècniques més objectives.

Un cop realitzada aquesta classificació, C1 C2, s'ha dividit el grup de 39 pacients inicials dels que s'ha efectuat l'adquisició del senyal en dos subgrups: un grup d'aprenentatge format per 16 pacients i un grup de test format per 23 pacients. Aquesta divisió s'ha realitzat perquè s'utilitzarà el grup d'aprenentatge per a la posta a punt dels mètodes per determinar les diferències entre les interaccions de la variabilitat ritme cardíac i el cycle respiratori i el grup de test per fer una validació d'aquests mètodes amb pacients que no han estat involucrats en la posta a punt del mètode. D'aquesta manera es podran valorar amb més objectivitat els resultats obtinguts.



El primer estudi que s'ha realitzat ha estat mitjançant la mitjana, la desviació tipus, el coeficient de variació de Pearson de les dades. Així, es pretenia veure fins a quin punt la utilització d'aquests estadístics tant intuïtius permetria realitzar una diferenciació en el nivell de complexitat de les sèries temporals. Els resultats obtinguts han reflexat que només es podia efectuar aquesta diferenciació per a les sèries que contenen la mitja de la sèrie temporal $T_{tot}(t)$.

Donada la complexitat de la interacció entre aquests dos senyals, i com que l'objectiu és trobar una metodologia que permeti obtenir diferents índexs que detectin la diferent variabilitat de les sèries temporals, s'ha decidit fer ús de tècniques d'anàlisi no lineal. La tècnica d'anàlisi no lineal que s'ha decidit utilitzar és la de la Dinàmica Simbòlica Conjunta. La seva aplicació requereix que els senyals de què s'utilitzen siguin simultanis en el temps, i per tant que els senyals tinguin valors en el mateix moment de mostreig, per la qual cosa es fa necessària una interpolació prèvia de les dades.

De l'aplicació d'aquesta metodologia sobre els pacients del grup d'aprenentatge, s'ha arribat a la important conclusió que existeixen una sèrie de variables amb un nivell de significació estadística molt bo com per a poder diferenciar entre grups de C1 i C2. Amb el període d'interpolació $t=0,5$ s s'obtenen més variables significatives que amb $t=1$ s, però qualitativament hi ha més variables amb un nivell molt bo de significació per a $t=1$ s. El valor del solapament, utilitzat en la metodologia de la Dinàmica Simbòlica, canvia poc aquests resultats.

S'han decidit generar funcions discriminants formades per combinacions d'aquestes variables que han resultat significatives estadísticament, per a diferenciar entre sèries que dels grups C1 i C2. L'objectiu que s'ha perseguit ha estat generar funcions amb el mínim nombre de variables independents possibles. Així, les funcions que s'han generat han estat funcions d'una variable (posant límits a les variables) i de dues variables (rectes en un espai de dues dimensions).

Un cop generades aquestes funcions discriminants s'han aplicat sobre el grup de sèries temporals que formen el grup de test per validar si les funcions generades poden ser útils per discriminar entre els dos grups. Els resultats han estat bastant bons, sobretot per al període de mostreig $t=0,5$ s, on s'ha arribat a més d'un 80% de pacients ben classificats del grup de test, en el cas de la variable $P(c010,r110)$, tant per solapament $s=1$ com $s=2$, en el cas de funcions discriminants d'una variable. Al aplicar les funcions discriminants de dues variables sobre el grup de test s'ha aconseguit el màxim nombre de pacients ben classificats en aquest estudi per a la combinació de variables $P(c011,r100)$ i $P(c111, r000)$ amb un 86,9565 % de pacients ben classificats.



Aquests resultats s'han comparat amb els que ens dóna l'estudi de variables lineals, amb el càlcul de la funció discriminant utilitzant la variable mitja de la sèrie temporal $T_{tot}(t)$, que era la única variable lineal que tenia un bon nivell de significació i s'ha arribat a un 60,8696 % de pacients ben classificats, un valor bastant inferior a l'obtingut amb la Dinàmica Simbòlica Conjunta.

L'última part d'aquest projecte s'ha dedicat a l'estudi del senyal de duració del cicle respiratori, aplicant Dinàmica Simbòlica i utilitzant dos alfabetos diferents, Θ i Ω . Amb l'alfabet Ω s'ha arribat a bons resultats utilitzant com a variable de la funció discriminant $P(r222)$ on s'ha arribat al 82,6087 % de pacients ben classificats i qualsevol grup de paràmetres. $P(r213)$. En aquest cas, els resultats obtinguts no han estat tan bons com els obtinguts amb la Dinàmica Simbòlica Conjunta, ni un nombre d'índexs tan elevat.

11.3. Extensions futures

En finalitzar aquest Projecte de Fi de Carrera i tenint en compte els resultats obtinguts, es plantegen dues extensions futures.

La primera extensió que s'hauria de plantejar és intentar corroborar els resultats obtinguts mitjançant uns grups d'aprenentatge i de test formats per un major nombre de pacients.

La segona extensió seria considerar altres alfabetos en la metodologia de la Dinàmica Simbòlica Conjunta, diferents del considerat en aquest projecte.



Bibliografia

11.4. Referències bibliogràfiques

ÀLVAREZ SAINZ, M. 'Estadística'. Universidad de Deusto, Bilbao, 1999.

ARDANUY ALBAJAR, R. / MARTÍN MARTÍN, Q. 'Estadística para Ingenieros' Editorial Hespérides. Salamanca, 1993.

ARMITAGE, P. / BERRY, G. 'Estadística para la Investigación Biomedica'. Harcourt Brace, Barcelona 1997.

BARTELS, H / BARTELS, R. 'Fisiología, Manual y Atlas' Ediciones Toray, S. A. Barcelona, 1983.

BAUMERT, M.¹ / WALTHER, T.² / HOPFE, J.¹ / STEPAN, H.³ / FABER, R.³ / VOSS, A.¹ 'Joint symbolic dynamic analysis of beat-to-beat interactions of heart rate and systolic blood pressure in normal pregnancy' . ¹Faculty of Medical Engineering, University of Applied Sciences Jena, Jena, German. ²Department of Cardiology, Benjamin Franklin Hospital, Free University Berlin, Berlin, Germany. ³Department of Obstetrics & Gynaecology, University of Leipzig, Leipzig Germany, 2002

BRUCE, N. / DANBENSPECK, J.A. 'Mechanisms and analysis of respiratory variability' Marvel Dekker, pp 285-317, 1995.

CAMINAL, P. / DOMINGO, L. / GIRALDO, B. / VALLVERDÚ, M. / VAZQUEZ G./ KAPLAN, D, 'Variability analysis of the respiratory volume based on nonlinear prediction methods'. Medical and Biological Engineering and Computing, vol. 42, n.1, pp. 86-91, 2004.

CARVAJAL, R. 'Análisis de la complejidad de la variabilidad de intervalos característicos del electrocardiograma' Tesi Doctoral ESAIL, CREB, 1999.

CARVAJAL, R. / ZEBROWSKI, J.J. / VALLVERDÚ, M. / BARANOWSKI, R. / CHOJNOWSKA, L. / POPLAWSKA, W / CAMINAL, P. 'Dimensional analysis of HRV in hypertrophic cardiomyopathy patients'. IEEE Engineering in Medicine and Biology Magazine, vol. 6, pp. 71-78. 2002

CASALS, A. 'Sistema de processament i anàlisi de senyals respiratoris' Projecte Fi de Carrera ESAIL, CREB, 1999. VOSS, A. / WESSEL, N. / KURTHS, J. '24 Hour Heart Rate



Variability Based on New Methods of Non-Linear Dynamics' Computers in Cardiology, pp. 693-696. 1995.

DOUGLAS G. ALTMAN. 'Practical Statistics for Medical Research'. Chapman & Hall. London 1993.

GLEICK, J. 'Caos. La creación de una ciencia' Seix Barral, Barcelona, 1994.

GRASSBERG, P. / PROCACCIA, I. 'Characterization of strange attractors' Physica D.51, 1991

HAO, B. L., 'Elementary Symbolic Dynamics and Chaos in Dissipative Systems'. London: World Scientific Publishing Co, Pte, Ltd, 1989.

HOYER, D. / POMPE, B. / HERZEL, H. / ZWIENER, U. 'Nonlinear coordination of the cardiovascular and respiratory autonomic control'. IEEE Eng, And Med. And Biol., vol. 17, pp. 17-21., 1998.

KURTHS J. / VOSS, A. / WITT, A. / SAPARIN, P. / KLEINER, H. J, / WESSEL, N. 'Quantitative analysis of heart rate variability'. *Chaos*, vol. 5, pp. 88-94. 1995

LORENZ, E. N. 'Deterministic nonperiodic flow' Journal of Atmospheric Science, vol. 20 pp. 130-141, 1963.

MOMPÍN POBLET, J. [et al.]. 'Introducción a la bioingeniería' Mancombo Boixareu, Barcelona, 1988.

MOON, F. C. 'Chaotic and Fractal Dynamics: an introduction for applied scientist an engineers' Wiley Interscience, New York, 1992.

NET, A. / BENITO, S. 'Función pulmonar en el paciente ventilado' Edicions Doyma, Barcelona 1990.

OTT, E. 'Chaos in dynamical systems'. Cambridge University Press. Maryland University, 1997

OTT, E. / SAUER, T. / YORKE, J. A. 'Coping with chaos' Wiley, New York, 1994.

PALLÁS ARENY, R. 'Introducción a la bioingeniería' Serie mundo electrónico, pp. 79-81. Boixareu Editores, Barcelona, 1088

PORTA, A. / GUZZETTI, S. / MONTANO, N. / PAGANI, M. / SOMERS, V. / MALLIANI, A. / BASELLI, G. / CERUTTI, S. ' Information domain analysis of cardiovascular variability of regularity, synchronisation and co-ordination' Med. & Biol. Eng. & Comp., vol 38, pp. 180-188, 2000.



PRAT BARTÉS, A. / TORT-MARTORELL LLABRÉS, X. / GRIMA CINTAS, P. / POZUETA FERNÁNDEZ, L. 'Métodos Estadísticos'. Edicions UPC, Barcelona 1997.

SOLÉ, R. V. / MANUBRIA, S. C. 'Orden y caos en sistemas complejos. Fundamentos' Edicions UPC, Barcelona, 1996.

TASK FORCE of the European society of cardiology and the North American society of pacing and electrophysiology, heart rate variability – standards of measurements, Physiological interpretation and clinical use. *Circulation*, vol. 93, pp. 1043-1065. 1996.

THELIER J. / EUBANK, A. / LONGTIN, B. / GALDRIKIAN, D. / FARMER, D. 'Testing for nonlinearity in time series: The method of surrogate data" *Physica D*, vol. 58, pp. 77-94, 1992.

WAYNE, D. 'Bioestadística. Base para el análisis de las ciencias de la salud' Limusa Noriega Editores, Georgia State University, 1977.

11.5. Bibliografia complementària

FERRÁN ARANAZ, M. 'SPSS para Windows. Programación y Anàlisis Estadístico'. McGraw-Hill/Interamericana de Espanya, S. A. Madrid, 1996.

GARCÍA DE JALÓN, J. / RODRIGUEZ, J. I. / BRAZÁLEZ, A. 'Aprenda Matlab 6.1 como si estuviera en primero' Universidad Politecnica de Madrid. Madrid, 2001.

NORUŠIS, M. J. 'SPSS Professional Statistics 6.1'. Chicago, 1994

