Desarrollo de una aplicación web para el estudio de los datos longitudinales para el seguimiento de la infección por el virus de inmunodeficiencia humana, caso de estudio: IAHULA

Stephanie Correa

Indice

- El problema
 - Objetivo general
 - Objetivos específicos
- Fundamentos teóricos
- Metodología
 - Población y muestra
 - Metodología de desarrollo de software en espiral
 - Primera interacción
 - Segunda interacción
- Desarrollo Versión final HIVmlm
- Conclusiones y recomendaciones



Planteamiento del problema

Puntos clave

- Seguimiento de la infección por VIH
- Análisis temporal
- Datos longitudinales

$$y_1 = \{y_{11}, y_{12}, \cdots, y_{1n_i}\}$$

 \vdots
 $y_{i=N} = \{y_{N1}, y_{N2}, \cdots, y_{Nn_i}\}$

Con n_i el número de observaciones para i-ésimo individuo (N).



Planteamiento del problema

Características de los datos longitudinales

- Alta heterogeneidad
- Correlación entre medidas
- Datos faltantes



Planteamiento del problema

Características de los datos longitudinales en VIH



Whitworth L. Reflection on Cohorts and Longitudinal Studies. UNICEF office research. Wellcome Trust. Florence. 2014.



Objetivo general

Objetivo

Desarrollar una aplicación web para el estudio de datos longitudinales para el seguimiento de la infección por el Virus de Inmunodeficiencia Humana.

Caso de estudio: Instituto Autonómo Hospital Universitario de los Andes.



Objetivos especificos

Objetivos

- Diagnosticar la información obtenida en la base de datos suministrada por el Instituto Autonómo Hospital Universitario de los Andes.
- Analizar de manera exploratoria y descriptiva los datos para la selección de las variables a utilizar en el módelo a implementar.
- Implementar un módelo para el manejo de datos longitudinales en una aplicación web.
- Aplicar pruebas de funcionalidad en la aplicación.



Antecedentes

Manejo de datos longitudinales multivariados incompletos de alta dimensión con tipos de datos mixtos por imputación múltiple utilizando un módelo de análisis factorial longitudinal

- Modelo escrito en R
- Análisis de factores y modelo lineales de efectos mixtos
- PAN Imputación Multiple para Panel Multivariado
- MICE Imputación Multiple usando Ecuaciones Encadenadas





Antecedentes

Tablas de modelo de vida calibradas por VIH para países con epidemias de VIH generalizadas

- Paquete de R HIV.LifeTables
- Tasa de mortalidad
- Prevalencia del VIH
- Esperanza de vida al nacer

```
0 2.5 0.075225512 0.084793459
            5 0.010129540 0.039911727
            5 0.002914778 0.014468457
              0.001642622 0.008179518
       15 2.5 0.002155948 0.010721950
            5 0.003335297 0.016538583
       25 2.5 0.004452731 0.022018548
       30 2.5 0.005252820 0.025923669
       35 2.5 0.005791861 0.028545969
       40 2.5 0.006256334 0.030799931
       45 2.5 0.007132402 0.035037262
       50 2.5 0.009135300 0.044656620
       55 2.5 0.013015347 0.063025972
[14,]
       60 2.5 0.020095615 0.095671627
[15.]
       65 2.5 0.032423947 0.149963696
[16.]
       70 2.5 0.053064313 0.234246270
```



Antecedentes

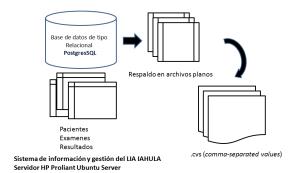
Unión del modelado de datos longitudinales y de supervivencia

- Paquete de R JM
- Datos longitudinales
- Datos de supervivencia
- Eficacia de antirretrovirales

```
> fitLME <- lme(sqrt(CD4) ~ obstime * drug - drug,
                random = \sim 1 | patient, data = aids)
> fitIME
Linear mixed-effects model fit by REML
 nata: aids
 Log-restricted-likelihood: -1360.364
 Fixed: sqrt(CD4) ~ obstime * drug - drug
    (Intercept)
                        obstime obstime:drugddI
    2 508715346
                   -0.033878275
                                    0.005581298
Random effects:
Formula: ~1 | patient
        (Intercept) Residual
          0.8719493 0.4106749
StdDev:
Number of Observations: 1405
Number of Groups: 467
```



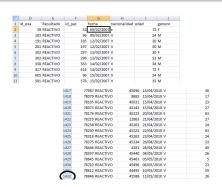
Población y muestra - Adquisición y filtrado



Población de pacientes registrada entre enero del año 2007 hasta diciembre del año 2013: 45649 pacientes.



Adquisición - Ejemplo de una vista minable



Vista minable de la ventana de tiempo entre enero del año 2007 y marzo del año 2010 reportaba 1431 pacientes reactivo VIH.

Según el Informe nacional de seguimiento de la declaración política sobre VIH y el SIDA de 2011. MPPS. ente los años 2 2010 se reportaron 47.771 nuevos casos (Tabla 4. Pag. 21)

Filtrado de la vistas minables

- La base de datos del Laboratorio de Investigaciones Hormonales IAHULA, no sólo almacena información del Programa de Pacientes con VIH, las vistas minables estan conformadas por pacientes confirmados por las pruebas Elisa y Western Blot para VIH
- En las vistas minables para las pruebas sólo se categorizarán pacientes con residencia en el Estado Mérida
- Las vistas minables sólo las integraran pacientes con más de tres años en seguimiento. Dado que los modelos son para datos longitudinales.
- Se categorizaron los pacientes por semestres y en cada lapso en medicados y no medicados.



Filtrado de la vistas minables

'nnnnnnnnn

Archivo para las pruebas funcionales.

- 115 registros de pacientes (Todos medicados con antiguedad < 6 meses) mayores a 18 años.
- Sección transversal: edad el género y el municipio de residencia
- Sección longitudinal (Enero del año 2007 Diciembre del año 2013) Log₁₀ de la carga viral plasmática y las sub-poblaciones linfocitaria de células T⁺ CD4 y T⁺ CD8.

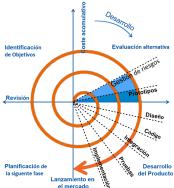
Según estadísticas del Programa para pacientes con VIH del IAHULA, en el año 2007 existían 615 pacientes adscritos al programa.

Metodología de desarrollo de software en espiral

Características del desarrollo de software

Ciclo de vida de software

- Determinar objetivos
- 2 Análisis de riesgos
- Desarrollar, verificar y validar
- Planificar





Prototipo 1

Primera interacción

- Base de datos del LIA IAHULA
- Implementación en lenguaje R
- Desarrollo WEB en Java
- Conexión R- Java con rJava y JDK eclipse.
- Servidor WEB (Ubuntu server)





Prototipo 2

Segunda interacción

- RStudio
- R
- Shiny





Etapa [.]

HIVmlm Creación del paquete en R formatR

Módulos:

- Carga de datos
- Exploración de los datos
- Visualización en el mapa
- Ajuste del modelo
- Validación del modelo
- Generar reporte



Modelo mixto implementado

$$y$$
 = $X\beta$ + Zb + ϵ
Variable respuesta Efectos fijos Efectos aleatorios

Donde:

La variable respuesta: Log₁₀CVP - Conteo de células T CD4 Efectos fijos: edad, género, lugar de residencia, grupo de riesgo, co-infecciones, (Conteo de células T CD4 - Log₁₀CVP), Conteo de células T CD8, entre otros.

Efectos aleatorios: Se considera que el comportamiento de cada paciente es igual o diferente.

← - error.

Etapa 2

- Creación del paquete
- GitHub
- global.R
- ui.R
- server.R
- reporte.Rmd

Estructura del paquete y proyecto





Etapa 3

- IAHULA
- Período 2007 2013
- 1610 datos registrados
- 115 pacientes
- Gráfico de torta
- Histograma
- Marcadores fisiológicos

Plan de integración y pruebas





Carga de Archivos







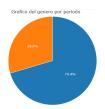
Visualización en el mapa

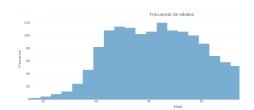
HIVmlm





Exploración de los datos







Ajuste del modelo

HIVmlm





Conclusiones

- Diagnosticar la información obtenida en la base de datos
- Analizar de manera exploratoria y descriptiva los datos
- Implementar un módelo para el manejo de los datos longitudinales





Recomendaciones

Implementar el módelo a los diferentes estados del país



Unificar esfuerzos con organismos de salud públicos y privados



