

**Universidad de las Ciencias Informáticas**

**Facultad 6**



**“Módulo para análisis estadístico de datos del  
RServer 2.0”**

Trabajo de Diploma para optar por el título de  
Ingeniero en Ciencias Informáticas

**Autores:** Yanel de la Caridad Martínez Coipe.  
Ramón Pérez Moshkovskaya.

**Tutor:** Ing. Rachid Alí Grave de Peralta.

**Consultante:** Msc. Yunier Emilio Tejeda Rodríguez.

La Habana, Cuba, junio de 2012

“Año 54 de la Revolución”

DECLARACION DE AUTORIA

Declaramos ser autores de la presente tesis y reconocemos a la Universidad de las Ciencias Informáticas los derechos patrimoniales de la misma, con carácter exclusivo.

Para que así conste firmo la presente a los \_\_\_\_ días del mes de \_\_\_\_\_ del año \_\_\_\_\_.

**Yanel de la Caridad Martínez Coipe**

\_\_\_\_\_  
Firma del Autor

**Rachid Alí Grave de Peralta**

\_\_\_\_\_  
Firma del Tutor

**Ramón Pérez Moshkovskaya**

\_\_\_\_\_  
Firma del Autor

**Rachid Alí Grave de Peralta**

\_\_\_\_\_  
Firma del Tutor

## DATOS DE CONTACTO

**Ing. Rachid Alí Grave de Peralta:** Profesor graduado en Ingeniería de las Ciencias Informáticas en el año 2009, en la Universidad de las Ciencias Informáticas. Contactar al correo [rali@uci.cu](mailto:rali@uci.cu). Se desempeña como profesor en el centro DATEC.

**Msc. Yunier Emilio Tejeda Rodríguez:** Máster en Ciencias Matemáticas, actualmente se desempeña como profesor de Matemáticas en la Universidad de las Ciencias Informáticas. Contactar al correo [yuniere@uci.cu](mailto:yuniere@uci.cu).

## AGRADECIMIENTOS

*Ramón*

*Quiero agradecer en especial a mi papá, pues si no fuera por él, hoy no fuera ingeniero en ciencias informáticas, sino veterinario. A mi mamá que siempre me ha ayudado en todo. Mis padres son especiales para mí, gracias por ayudarme a iniciar mi vida. Quiero que sepan que los quiero mucho y estoy muy orgulloso de ser su hijo.*

*A mis hermanas que tanto las amo, aunque no se lo demuestro con frecuencia y aunque siempre me fajé con ellas, quiero que lo sepan.*

*A mi tática linda, la reina de mi corazón. Gracias por amarme de esa manera tan especial y única. Gracias por haberme dado ese niño tan lindo, que llegó a nuestras vidas para llenarla de alegrías y enfocar nuestro futuro. Te amo con todas mis fuerzas.*

*Quiero agradecer a mi familia, que tanto me quiere. A mis tías Esther, María, Aleida y Meiber. A mis primos y primas. A mi tío Ángel que siempre lo llevo en mi corazón. Una familia numerosa, que siempre ha estado al pendiente.*

*A mis amigos de la infancia Hirán y Hendrik, siempre han estado conmigo en todo momento, que nos consideramos hermanos. Quiero agradecer a todos mis compañeros y compañeras de clase, a los que se han mantenido firme y terminan junto conmigo Lorenzo, Maikel, Arcides, Reinaldo y otros. A mis compañeros de apartamento Noel, Magdiel, Raúl Rondón y Raúl Domínguez. En especial a Marcos y a Enrique.*

*Quiero agradecer a mi compañera de tesis, una muchacha muy ocupada y preocupada, que no me volvió loco de milagro. A Juan José (Pepé) que nos ayudó a traducir el idioma alemán en el que nuestro tutor nos hablaba al principio. Agradecer a nuestro tutor Rachid Ali Grave de Peralta, que si por él fuera replanteara la ingeniería.*

*Yanel*

*Agradezco a mis padres, por siempre confiar en mí. Por dejarme seguir mis sueños y apoyarlos sin ver el precio. Por como son, vivo orgullosa de ser su hija.*

*Agradezco a mi hermano querido, por hacerme sentir que soy su ejemplo a seguir.*

*A mis abuelos y abuelas, por ser ellos las personas a las que uno cuida y nunca quieres que se vayan de tu lado.*

*A mi novio Juan José (Pepe), por estar conmigo incondicionalmente, apoyarme en todas mis decisiones y ser un maestro para mí.*

*A mis tíos, Adelaida, Segunda, Guello, Eduardo y Jesús.*

*A mis otros familiares y amigos de la Universidad. En especial a Lianet Salazar y a Juan Carlos, ambos fueron un sostén durante todo este tiempo. También a Ariadna y a Randiel. Pocos, pero amigos de verdad.*

*De todos ellos yo vivo agradecida, solamente por existir.*

*A mi compañero de tesis, "Kovsky", por ser tan paciente conmigo. A mi tutor, una persona muy profesional y amiga a la vez. Por ayudarme a que mi sueño de graduarme como ingeniera en ciencias informáticas se hiciese realidad.*

*Agradezco a la vida, haber pasado estos cinco años en la UCI, creo que me han hecho mejor persona. Aquí aprendí muchísimo, tanto en lo profesional como en lo personal. Este tiempo transcurrido me va a servir de experiencia para siempre.*

DEDICATORIA

*Ramón*

*A mi mamá y a mi papá por ser los mejores padres del mundo.*

*A mis hermanas Anna y Aleida que tanto las quiero.*

*A mi esposa Ivón y a ese niño tan lindo que me dio.*

*A mis amigos de la infancia, a mis hermanos Hirán y Hendrik, por siempre apoyarme.*

*A mis nuevos amigos de la universidad Enrique y Marcos.*

*Yanel*

*A mi mamá, por vivir para mí y para mi hermano.*

*A mi papá, por su paciencia y comprensión.*

*A mi hermano, ese bichito que mis padres me dieron para tenerlo de compañero y amigo, siempre.*

*A mis abuelos Ismael y Luis, por malcriarme y darme la mejor educación posible, siempre confiando en mí.*

*A mis abuelas Coralía y María Josefa, que aunque no están hoy presentes, se que estuvieran muy orgullosas de mí. Y en especial a mi mima, Coralía, por ser mi ángel.*

*A mi novio Pepe, él ha sido en la Universidad un ejemplo a seguir y un apoyo incondicional.*

## RESUMEN

En muchas empresas el análisis estadístico es fundamental para una correcta toma de decisiones e interpretación de los datos. En la práctica debido al creciente desarrollo tecnológico estas tareas son cada vez más complejas. Sin embargo gracias a la fusión de la informática con la estadística se han creado herramientas computacionales dedicadas a este fin. Una de ellas es el lenguaje de programación R, desarrollado inicialmente por Robert Gentleman y Ross Ihaka en 1993 y considerado como uno de los más potentes en este campo. Por otro lado, las tecnologías web son cada vez más usadas en el ámbito empresarial y económico pero los lenguajes que la soportan no son recomendables para hacer análisis matemáticos precisos. Este trabajo tiene como objetivo el desarrollo de un módulo capaz de brindar un conjunto de tipos de análisis estadísticos propios de R fácilmente integrables en proyectos web, específicamente escritos en PHP<sup>1</sup>. Se lograron implementar 28 tipos de análisis, los cuales pueden ser consumidos como servicios web o usados como API<sup>2</sup>. Finalmente se validó el sistema mediante un conjunto de pruebas exploratorias y se redactó la documentación necesaria para el uso del mismo.

**PALABRAS CLAVES:** análisis estadístico, R, estadística.

---

<sup>1</sup> **PHP:** viene del acrónimo Hypertext Pre-processor. Es un lenguaje de programación interpretado (Lenguaje de alto rendimiento), diseñado originalmente para la creación de páginas web dinámicas.

<sup>2</sup> **Interfaz de programación de aplicaciones o API** (del inglés Application Programming Interface): es el conjunto de funciones y procedimientos (o métodos, en la programación orientada a objetos) que ofrece cierta biblioteca para ser utilizado por otro software como una capa de abstracción. Son usadas generalmente en las bibliotecas.

Índice:

AGRADECIMIENTOS .....	I
DEDICATORIA .....	III
RESUMEN .....	IV
INTRODUCCIÓN .....	1
<b>CAPÍTULO 1: FUNDAMENTO TEÓRICO .....</b>	<b>4</b>
1.1- SOFTWARE DE ANÁLISIS ESTADÍSTICO .....	4
1.2- LENGUAJE DE PROGRAMACIÓN .....	7
1.3- MARCO DE TRABAJO PARA PHP .....	8
1.4- METODOLOGÍA DE DESARROLLO DE SOFTWARE .....	10
1.5- HERRAMIENTA CASE .....	10
1.6- AMBIENTE DE DESARROLLO .....	11
1.7- SERVIDOR WEB .....	11
1.9- BIBLIOTECA UTILIZADA .....	12
1.10- MÉTODOS DE ANÁLISIS ESTADÍSTICO .....	13
<b>CAPÍTULO 2: ANÁLISIS Y DISEÑO DEL SISTEMA .....</b>	<b>21</b>
2.1 - MODELO DE DOMINIO .....	21
2.2- REQUISITOS DE SOFTWARE.....	22
2.2.1- <i>Requisitos Funcionales</i> .....	22
2.2.2- <i>Requisitos no Funcionales</i> .....	23
2.3- DIAGRAMA DE CASO DE USO DEL SISTEMA .....	24
2.4- PATRONES DE DISEÑO Y DE ARQUITECTURA DE SOFTWARE .....	39
2.5- DIAGRAMAS DE CLASES DEL DISEÑO .....	40
2.6- DIAGRAMAS DE SECUENCIA .....	42
2.7- PROPUESTA DE ARQUITECTURA .....	44
<b>CAPÍTULO 3: IMPLEMENTACIÓN Y PRUEBA .....</b>	<b>45</b>
3.1- DIAGRAMA DE COMPONENTES .....	45
3.2- DIAGRAMA DE DESPLIEGUE .....	46
3.3- IMPLEMENTACIONES SIGNIFICATIVAS .....	46
3.4- PRUEBAS AL SISTEMA .....	50
CONCLUSIONES .....	57
RECOMENDACIONES.....	58
REFERENCIAS BIBLIOGRAFICAS .....	59
BIBLIOGRAFIA.....	60
ANEXOS.....	62
GLOSARIO DE TÉRMINOS .....	72

## INTRODUCCIÓN

A medida que se ha ido desarrollando la humanidad, se ha hecho imprescindible la confección de herramientas para la investigación científica y empírica, así como técnicas para contabilizar y registrar de una forma u otra la población y las riquezas existentes. Uno de los problemas universales ha sido la inteligente y la acertada toma de decisiones en condiciones de incertidumbre. El surgimiento y desarrollo de la estadística como ciencia se ha enfocado en solucionar estas necesidades.

“La estadística estudia los métodos científicos para recoger, organizar, resumir y analizar datos, así como para sacar conclusiones válidas y tomar decisiones razonables basadas en tal análisis”. Según expresa Gini, (1953), Murria R. Spiegel, (1991).

“La estadística es la ciencia que trata la recolección, clasificación y presentación de los hechos sujetos a una apreciación numérica como base a la explicación, descripción y comparación de los fenómenos”. Según expresa (Yale y Kendal, 1954).

Cualquiera sea la definición, lo esencial es la importancia científica que aporta esta ciencia, debido al enorme campo de aplicación que posee. Se usa en las diferentes esferas del conocimiento y de las ciencias. Constituyendo uno de los aspectos más relevantes entre los estudiosos de esta última. Es un lenguaje que permite comunicar información basada en datos cuantitativos. La evolución de la estadística ha llegado al punto en que su proyección se percibe en casi todas las áreas de trabajo. También abarca la recolección, presentación y caracterización de información para ayudar tanto en el análisis e interpretación de datos como en el proceso de tomar decisiones.

Con el paso de los años se fueron incrementando las poblaciones y por ende los recursos, por tanto se fueron desarrollando métodos estadísticos más rápidos y fiables para todos. Es entonces que aparece en 1642 la primera máquina de cálculos estadísticos, creada por Blaise Pascal: la cual se fue perfeccionando con el transcurso del tiempo, hasta aparecer los ordenadores digitales actuales, que son sin dudas, el mayor adelanto de la estadística en la actualidad. A partir de este momento se han venido desarrollando grandes equipos y software que proveen una serie de análisis estadísticos, necesarios en variadas ramas de la ciencia, destacándose en la medicina. También existen herramientas que son capaces de reportar análisis, como por ejemplo: MATLAB, Minitab, Statgraphics y STATISTICA, sólo por nombrar algunas. De igual forma existen lenguajes de programación para el análisis estadístico, que ayudan a realizar este tipo de cálculos. Tal es el caso de R, que es un software altamente recomendable por lo robusto y potente de las rutinas que tiene implementadas. (1) Por otro lado, las tecnologías web son cada vez más usadas en el ámbito empresarial y económico pero los lenguajes que la soportan no son recomendables para hacer análisis matemáticos precisos, por ende se hacen necesarias herramientas informáticas que comuniquen las tecnologías web con

lenguajes como el R, como resultado se han hecho algunos aportes en nuestro país y específicamente en la Universidad de las Ciencias Informáticas (UCI).

Esta institución representa la vanguardia en el desarrollo del software en nuestro país, divide su producción de software en departamentos, distribuidos en 7 facultades y diferentes centros de desarrollo. Uno de ellos es el Departamento de Soluciones Integrales del Centro de Tecnologías de Gestión de Datos (DATEC), de la facultad 6. Entre las soluciones confeccionadas en dicho departamento, existe un sistema para el análisis estadístico de datos llamado RServer, que tiene como objetivo realizar análisis estadísticos propios de R en arquitecturas web. En su versión 1.0 presenta un conjunto de deficiencias y limitantes importantes, entre las cuáles se pueden mencionar que no provee un módulo completo que agrupe la mayoría de los tipos de análisis estadísticos brindados por el lenguaje R, teniendo implementado solamente cuatro. No usa buenas prácticas en su implementación, pues se evidencia la carencia de patrones de diseño o arquitectura aplicados. Todo esto trae como consecuencia que la adición de nuevos tipos de análisis, la integración y la reutilización futura sea compleja y difícil de garantizar.

En atención a lo anteriormente expuesto, se determinó el siguiente **Problema de la Investigación**: ¿Cómo realizar el análisis estadístico de datos en el RServer 2.0?

Partiendo del problema planteado, el **Objeto de Estudio**: se define en las aplicaciones y métodos de análisis estadístico.

El **Campo de Acción**: estaría enmarcado en las aplicaciones web para análisis estadístico de datos.

El **Objetivo General**: es desarrollar el módulo para análisis estadístico de datos para el RServer 2.0.

Del objetivo general se desglosan los siguientes **Objetivos Específicos**:

- Realizar el estado del arte de las herramientas que proveen análisis estadísticos y seleccionar las tecnologías a utilizar.
- Definir la arquitectura de la aplicación y diseñar el módulo que brinde solución al problema planteado.
- Implementar las funcionalidades definidas y realizar pruebas que demuestren el correcto funcionamiento del módulo.

A partir de los objetivos trazados se derivan las siguientes **tareas de la investigación** para darle cumplimiento a los mismos:

1. Estudio de los sistemas de análisis estadísticos existentes, para comprender el funcionamiento de los mismos.
2. Selección de las herramientas para el diseño e implementación del software.
3. Definición de los requisitos funcionales y no funcionales del módulo.
4. Selección de patrones a emplear para el desarrollo del sistema.

5. Implementación del diseño realizado para dar cumplimiento a las funcionalidades definidas.
6. Realización de pruebas de caja negra y caja blanca para validar el correcto funcionamiento de la solución propuesta.

**La tesis queda estructurada** en 3 capítulos:

### **Capítulo 1 Fundamento Teórico**

En este capítulo se realiza un estudio de software que realizan análisis estadístico de datos a nivel mundial. También se desarrolla un análisis valorativo y crítico de las herramientas, marco de trabajo para PHP, lenguajes de programación y metodologías a utilizar. Además se determina el ambiente de trabajo en el que se va a desarrollar. Y se brinda una breve descripción matemática de los métodos de análisis estadístico que se van a implementar.

### **Capítulo 2 Análisis y Diseño del Sistema**

En este capítulo se especificarán los requisitos funcionales y no funcionales de la aplicación. Y se exponen el modelo del dominio, el diagrama de caso de uso del sistema, el diagrama de clases del diseño y los diagramas de secuencias más significativos; siendo todos estos los que comprenden el diseño del sistema. También se determinan los patrones de diseño y de arquitectura que se utilizaron en la aplicación. Además se plantea una propuesta de arquitectura a seguir.

### **Capítulo 3 Implementación y Prueba**

En este capítulo se exponen las pruebas realizadas al sistema así como algunos ejemplos de código más significativos, implementados en la solución. También se muestran el diagrama de componentes y el diagrama de despliegue de la aplicación.

## CAPÍTULO 1: FUNDAMENTO TEÓRICO

Este capítulo tiene como objetivos sentar las bases teóricas para la investigación. Realizar el estudio del estado del arte; que recoge las características de diferentes software que realizan análisis estadísticos de datos y una breve valoración y descripción de las ventajas y desventajas de las herramientas, marco de trabajo para PHP, metodologías de desarrollo de software y lenguajes de programación a utilizar en la realización de la aplicación. Además de exponer las fórmulas y aplicaciones matemáticas de los métodos de análisis estadístico de datos a implementar en el módulo de análisis estadístico del RServer 2.0.

### **1.1- Software de análisis estadístico**

Un paquete estadístico es un cúmulo de programas y subprogramas conectados que funcionan de manera conjunta; en otras palabras, para cambiar de uno al otro, no se hace necesario salir del programa y seguidamente volver a entrar en él. Por lo que permite aplicar a un mismo fichero de datos un conjunto ilimitado de procedimientos estadísticos de manera sincronizada, sin salir del programa. Así la utilidad del conjunto integrado es mayor que la suma de las partes. En cierto modo, un paquete estadístico es similar a un paquete ofimático (por ejemplo, Office 97 de Microsoft). (2)

SPSS (Statistical Package for the Social Sciences) es uno de los principales paquetes estadísticos. Otros ejemplos importantes de paquetes estadísticos son: STATISTICA, MATLAB (abreviatura de Matrix Laboratory), R, Statgraphics y Minitab. En el pasado, había otros paquetes estadísticos, como BMDP (Biomedical Computers Programs) hoy absorbido por la firma SPSS. (2)

### **SPSS**

Es la herramienta estadística más utilizada a nivel mundial en el entorno académico. Puede trabajar con bases de datos de gran tamaño. Además, de permitir la recodificación de las variables y registros según las necesidades del usuario. El programa consiste en un módulo base y módulos anexos que se han ido actualizando constantemente con nuevos procedimientos estadísticos. Esta tecnología automatiza el proceso de descubrimiento del conocimiento, ayuda a centrarse en un área de interés, permite predecir resultados. También encuentra patrones dentro de un fichero de datos y aumenta las capacidades ofrecidas por otras herramientas. Es un software con licencia propietaria. (2)

### **Minitab**

Es otro de los programas más usados en el mundo para el análisis estadístico. Permite calcular la mayoría de las metodologías estadísticas habituales, entre las que se cuentan: análisis exploratorio de datos, gráficos estadísticos, control de calidad, estadística no paramétrica, regresión y sus variantes, entre otras. Ofrece herramientas precisas y fáciles de usar para aplicaciones estadísticas generales y muy especialmente para control de calidad. La confiabilidad de sus algoritmos estadísticos y la sólida base de la combinación de potencia y simplicidad de manejo le han hecho merecer la confianza de los usuarios que lo siguen. Es el único paquete que ofrece todos los métodos estadísticos adecuados: estadística básica y avanzada, regresión y ANOVA (Analysis of Variance), DOE - Diseño de experimentos, análisis de fiabilidad, tamaño de muestra y capacidad, series temporales y predicción, potente importación, exportación y manipulación de datos y lenguaje de macros. Incluye regresión múltiple y no es adecuado para conjuntos de datos pequeños. (2)

### **Statgraphics**

Es un programa de estadística de fácil manejo, es una buena herramienta para la enseñanza de la estadística en secundaria y para la investigación en la educación. Este programa está construido bajo una estructura modular formada por tres módulos diferentes, acumulando en total más de 150 procedimientos distintos de distribución. Prepara todo tipo de informes, gráficos, tablas y analíticas que se van actualizando en tiempo real a medida que el propio usuario varía los datos. Modifica colores, fuentes de letra y tamaños por lo que es un completo editor estadístico. Además es un asistente estadístico, pues brinda una ayuda al escoger, busca siempre el método estadístico que sea más adecuado en cada caso. También los experimentos y procedimientos estadísticos que se pueden utilizar en StatGraphics Centurion son muy variados y ofrecen capacidades de control casi ilimitadas. (3)

### **STATISTICA**

Es otro de los programas más usados a nivel mundial para el análisis estadístico. Entre todas las facilidades y pruebas estadísticas que incluye, destaca la variedad de gráficos y la facilidad en el manejo de bases. El paquete puede ser extendido a través de una interfaz con el lenguaje R. Además, se pueden modificar y añadir nuevas librerías usando el marco de trabajo .NET<sup>3</sup>. Es un software con licencia propietaria. (2)

### **MATLAB**

---

<sup>3</sup> .NET: es un marco de trabajo o framework de Microsoft que hace un énfasis en la transparencia de redes.

Es un software matemático que ofrece un entorno de desarrollo integrado (IDE) con un lenguaje de programación propio (lenguaje M). Está disponible para las plataformas Unix, Windows y Apple Mac OS X. Entre sus prestaciones básicas se hallan: la manipulación de matrices, la representación de datos y funciones, la implementación de algoritmos, la creación de interfaces de usuario (GUI) y la comunicación con programas en otros lenguajes y con otros dispositivos de hardware. Permite una codificación rápida y fácil en un lenguaje de muy alto nivel. Además genera gráficas de alta calidad e instalaciones de visualización. Es un producto propietario de The Mathworks, por lo que los usuarios están sujetos y bloqueados al vendedor. Es un software con licencia propietaria. (2)

### **Lenguaje de programación R**

Es un conjunto de programas integrados para el manejo de datos, simulaciones, cálculos y realización de gráficos. Es también un lenguaje de programación orientado a objetos y además una implementación libre, independiente, código abierto del lenguaje de programación S<sup>4</sup>. Se distribuye bajo la licencia GNU GPL y está disponible para los sistemas operativos Windows, Macintosh, Unix y GNU/Linux. Proporciona un amplio abanico de herramientas estadísticas (modelos lineales y no lineales, test estadísticos, análisis de series temporales, algoritmos de clasificación y agrupamiento) y gráficas. Permite generar gráficos con alta calidad. (2)

Usado en combinación con otro tipo de programas que extienden su capacidad gráfica, hojas de cálculo, procesadores de texto, programas específicos para la representación de datos, Java Script, entre otros, puede realizar con un coste nulo, análisis sumamente sofisticado que, incluyen, tal vez, el 99% de los que precisan el 95% de los usuarios. R proporciona un lenguaje de programación propio. Basado en el lenguaje S, que a su vez tiene muchos elementos del lenguaje C<sup>5</sup>. Sin embargo, la semántica es muy distinta a la de este último. Esto es porque R permite ejecuciones de comandos en línea (compilación y ejecución unidas en un mismo paso), lo cual hace que su semántica esté más próxima a la de un lenguaje de programación funcional. (4)

Consta de características tales como: es un lenguaje de programación bien desarrollado, simple y efectivo, que incluye condiciones, ciclos, funciones recursivas y posibilidad de entradas y salidas. Dispone de un amplio almacenamiento y manipulación efectivo de datos. También tiene una amplia, coherente e integrada colección de herramientas y gráficas para análisis de datos. Además, cuenta con operadores para cálculos sobre las variables indexadas (Array), en particular matrices. Y como característica más notable, se tiene que R es un software libre y que compite con SPSS, por lo que es muy eficiente en cuanto a análisis estadístico se trata. (2)

---

<sup>4</sup> S: es un lenguaje de programación estadístico.

<sup>5</sup> C: es un lenguaje de programación orientado a la implementación de Sistemas Operativos.

Se hizo una breve comparación de un grupo de paquetes que realizan análisis estadístico de datos y R en su versión 2.13, por sus características resalta entre los mencionados. Puesto que es un software de código abierto y libre. Además trae implementado un gran cúmulo de herramientas para el análisis estadístico y genera eficientes gráficos de alta calidad.

### 1.2- Lenguaje de programación

Los lenguajes de programación son creados con el objetivo de que las computadoras o las máquinas entiendan un determinado proceso. Existen un gran número de lenguajes que se pudieran utilizar para la solución final, entre ellos están JAVA<sup>6</sup>, PHP o .ASP (Active Server Pages).

#### PHP (Hypertext Pre-processor)

Es un lenguaje de programación interpretado, diseñado originalmente para la creación de páginas web dinámicas. Se usa principalmente para la interpretación del lado del servidor, pero actualmente puede ser utilizado desde una interfaz de línea de comandos o en la creación de otros tipos de programas incluyendo aplicaciones con interfaz gráfica.

Ventajas que provee este lenguaje:

- Es un lenguaje multiplataforma.
- El código fuente escrito en PHP es invisible al navegador web y al cliente porque es el servidor el que se encarga de ejecutar el código y enviar su resultado HTML (HyperText Markup Language) al navegador. Esto hace que la programación en PHP sea segura y confiable, muy distinto de Javascript<sup>7</sup> que su código está visible para cualquier usuario.
- Posee una amplia documentación en su sitio web oficial, entre la cual se destaca que todas las funciones del sistema están explicadas y ejemplificadas en un único archivo de ayuda y en la UCI existe una gran comunidad que lo sigue.
- Es libre, por lo que se presenta como una alternativa de fácil acceso para todos. A diferencia de .ASP y otros lenguajes que su tecnología es propietaria.
- Permite aplicar técnicas de programación orientada a objetos.
- Biblioteca nativa de funciones sumamente amplia e incluida.
- No requiere definición de tipos de variables aunque sus variables se pueden evaluar también por el tipo que estén manejando en tiempo de ejecución.
- Tiene manejo de excepciones (desde PHP5).

---

<sup>6</sup> **Java:** es un lenguaje de programación de alto nivel, orientado a objetos.

<sup>7</sup> **JavaScript:** es un lenguaje de programación interpretado. Se utiliza principalmente en su forma del lado del cliente.

- Si bien PHP no obliga a quien lo usa a seguir una determinada metodología a la hora de programar (muchos otros lenguajes tampoco lo hacen), aun haciéndolo, el programador puede aplicar en su trabajo cualquier técnica de programación o de desarrollo que le permita escribir código ordenado, estructurado y manejable. Un ejemplo de ello son, los desarrollos que en PHP se han hecho del patrón de diseño Modelo Vista Controlador (MVC), que permiten separar el tratamiento y acceso a los datos, la lógica de control y la interfaz de usuario en tres componentes independientes. (5)

PHP 5.3 es una versión que cuenta con innumerables mejoras que consolidan su éxito. Se ofrece la posibilidad de hacer programas orientados a objetos y la implementación de servicios Web. Este se diferencia en varias partes: la primera sección se centra en los fundamentos para el desarrollo de programas, la segunda parte muestra los avances que han surgido en PHP5, como la programación orientada a objetos, la lectura de archivos XML (Extensible Markup Language) y la última división, orientada a usuarios avanzados, cuenta con la descripción de técnicas actuales de desarrollo. (6)

Se selecciona este lenguaje de programación web para la realización del módulo de análisis estadístico, fundamentalmente porque se ajusta a las necesidades de implementación y por las valiosas características y ventajas que brinda. Además, los productos desarrollados en el centro de DATEC están implementados en dicho lenguaje, lo que permite una futura integración entre ellos y el módulo. Es importante señalar que se usa PHP en su versión 5.3, pues como marco de trabajo para PHP se usará Symfony 2.0, el mismo tiene como requisito el uso de esta versión.

### **1.3- Marco de trabajo para PHP**

Los marcos de trabajo o comúnmente llamados Framework por su traducción en inglés, son desarrollados con el objetivo de brindarle a los programadores y diseñadores una mejor organización y estructura a sus proyectos. Facilita la realización de software, mucho más rápido, que creando toda una estructura desde cero. Entre los marcos de trabajo que se pueden mencionar se encuentran CodeIgniter, Zend Framework, Symfony, entre otros. Siendo este último, sin duda, uno de los mejores actualmente. Cumple con una serie de requisitos que permite la obtención de un glorioso producto final.

- Es fácil de instalar y configurar en la mayoría de plataformas y con la garantía de que funciona correctamente en los sistemas Windows, Unix y Linux.
- Utiliza programación orientada a objetos (POO), de ahí que sea imprescindible PHP 5.

- Es sencillo de usar en la mayoría de los casos, aunque es preferible para el desarrollo de grandes aplicaciones Web que para pequeños proyectos, no así para CodeIgniter<sup>8</sup>.
- Aunque utiliza MVC, tiene su propia forma de trabajo en este punto, con variantes del MVC clásico, el controlador frontal y las acciones.
- Basado en la premisa de “convenir en vez de configurar”, en la que el desarrollador sólo debe configurar aquello que no es convencional.
- Sigue la mayoría de mejores prácticas y patrones de diseño para la web.
- Preparado para aplicaciones empresariales y adaptables a las políticas y arquitecturas propias de cada empresa, además de ser lo suficientemente estable como para desarrollar aplicaciones a largo plazo.
- Es fácil de extender, lo que permite su integración con las bibliotecas de otros fabricantes.
- Una potente línea de comandos que facilitan generación de código, lo cual contribuye a ahorrar tiempo de trabajo, que por su parte, CodeIgniter, no contiene.

### **Symfony 2.0**

Por otra parte, esta versión de Symfony se basa en un conjunto de librerías PHP independientes unidas por una serie de bundles (módulos) modificables. Está construido sobre buenas prácticas como: la inyección de dependencias (término tomado del marco de trabajo Spring), además tiene un sistema de cacheado correcto y permite un desarrollo orientado a Tests (pruebas). Consta de buena seguridad, validación y es muy versátil. Utiliza buenas prácticas, tanto que toma ideas de otros marco de trabajo para PHP (de Spring) para aplicarlas. Plantea que se debe usar el PHP 5.3.2, pues esta versión es más rápida que las anteriores. Esta nueva versión de Symfony ha buscado la manera de agrandar su rendimiento, pone un Reverse Proxy entre la aplicación y el usuario, para cuando el mismo haga una primera petición a la base de datos, este busque la información en dicha base de datos, pero al volver a hacer otra petición, guarda esa información en el Reverse Proxy y así no gastar tiempo yendo a la aplicación y ahorrarse también el tiempo de ir a buscar a la base de datos. (5)

Se utiliza este marco de trabajo con el objetivo de utilizar la nueva característica que brinda (la incorporación de inyección de dependencias). Symfony 2.0 es tres veces más rápido que su versión anterior y que Zend Framework 1.10, por lo que cuenta con un alto rendimiento. Una de las noticias más reconfortantes es que tiene un gran soporte (se piensa que hasta el 2016), una buena documentación y una gran comunidad que lo sigue. Por estas razones y por las anteriormente expuestas se escoge este marco de trabajo para la implementación.

---

<sup>8</sup> **CodeIgniter:** es un entorno de desarrollo abierto que permite crear web dinámicas con PHP.

#### **1.4- Metodología de desarrollo de Software**

El proceso de desarrollo de software es un conjunto de actividades que guían los esfuerzos de las personas implicadas en el proyecto, a modo de plantilla que explica los pasos necesarios para terminar el proyecto (Jacobson, et al., 2000). Las metodologías de desarrollo se clasifican en estructuradas y no estructuradas, ejemplos de ellas son: Programación Extrema (XP), Proceso Unificado Ágil (AUP), Proceso Unificado de Rational (RUP), Método de Desarrollo de Sistemas Dinámicos (DSDM) y Proceso Unificado Abierto (Open Up). (7)

##### **Open Up**

Es un marco de trabajo para procesos de desarrollo de software. Preserva la esencia del Unified Process pues es iterativo e incremental, dirigido por casos de usos, centrado en la arquitectura. La mayoría de los elementos de Open UP están declarados para fomentar el intercambio de información entre los equipos de desarrollo y mantener un entendimiento compartido del proyecto, sus objetivos, alcance y avances. Sólo lo fundamental está incluido, sin dejar de ser completo y extensible (cuenta con menos de 20 artefactos). Está pensado para proyectos pequeños. Se decide emplear esta metodología pues Open Up es modificable y ampliable, está definida por centro para trabajar y por todas estas razones se estuvo de acuerdo en utilizarla. (7)

#### **1.5- Herramienta CASE**

Las Herramientas CASE, es la aplicación de métodos y técnicas a través de las cuales hacen útiles a las personas entender las capacidades de las computadoras, por medio de programas, de procedimientos y sin documentación. Dichas herramientas permiten modelar los procesos de negocios de las empresas y desarrollar los sistemas de información gerenciales. Proveen organizar y manejar la información de un proyecto informático, permitiéndoles a los participantes la comunicación entre ellos y así, tornándose más comprensibles los sistemas. (4)

##### **Visual Paradigm**

Es una herramienta UML (Lenguaje Unificado de Modelado)<sup>9</sup> profesional que soporta el ciclo de vida completo del desarrollo de software: análisis y diseño orientados a objetos, construcción, pruebas y despliegue. El software de modelado UML ayuda a una rápida construcción de aplicaciones de mejor calidad y a un menor coste. Permite dibujar todos los tipos de diagramas de clases, código inverso,

---

<sup>9</sup> **UML:** lenguaje gráfico para visualizar, especificar, construir y documentar un sistema.

generar código desde diagramas y generar documentación. La herramienta UML CASE también proporciona abundantes tutoriales, demostraciones interactivas y proyectos UML.

Además brinda como nueva característica, una interoperabilidad con modelos UML 2 (metamodelos UML 2.x para plataforma Eclipse) a través de XMI. (8)

### **1.6- Ambiente de Desarrollo**

Los Entornos de Desarrollo Integrado (IDE) o ambiente de desarrollo, es un programa informático compuesto por un conjunto de herramientas de programación. Un IDE es un medio de programación que ha sido empaquetado como un programa de aplicación, como un compilador, editor de código y constructor de interfaces gráficas. También suministran un marco de trabajo amigable para la mayoría de los lenguajes de programación. (4)

#### **NetBeans IDE**

Es una herramienta para que los programadores puedan escribir, compilar, depurar y ejecutar programas. Cuenta con un excelente completamiento de código. Está escrito en JAVA, pero puede servir para cualquier otro lenguaje de programación. Es multiplataforma. Existe además un número importante de módulos para extender el NetBeans IDE. Es un producto libre y gratuito sin restricciones de uso. (4)

### **1.7- Servidor Web**

Es la tecnología que tiene implícito programas informáticos que procesan aplicaciones realizando conexiones bidireccionales y/o unidireccionales y síncronas o asíncronas con el cliente, generando o cediendo una respuesta en cualquier lenguaje o aplicación del lado del cliente. (9)

#### **Apache 2.0**

Se utiliza como servidor web pues brinda una serie de características como:

- Es un servidor web conforme al protocolo HTTP (Hypertext Transfer Protocol) /1.1.
- Es modular: puede ser adaptado a diferentes entornos y necesidades, con los diferentes módulos de apoyo que proporciona y con la API (Interfaz de programación de aplicaciones) de programación de módulos, para el desarrollo de módulos específicos.
- Es continuamente actualizado.
- Es multiplataforma.
- Basado en hebras en esta versión 2.0.

- Incentiva la realimentación de los usuarios, obteniendo nuevas ideas, informes de fallos y parches para la solución de los mismos.
- Se desarrolla de forma abierta.
- Es extensible: gracias a ser modular se han desarrollado diversas extensiones entre las que destaca PHP, que es el lenguaje de programación que se utilizará. (9)

### 1.8- Servicio Web

Un servicio es una entidad de software que encapsula funcionalidad de negocios y proporciona dicha funcionalidad a otras entidades a través de interfaces públicas bien definidas. (10)

**Servicio Web** (Service Web) es un sistema de software diseñado para soportar interacción máquina-a-máquina sobre una red. Posee una interfaz descrita en un formato procesable por máquina (específicamente WSDL<sup>10</sup>). Otros sistemas interactúan con el Web Service de una manera prescrita por su descripción utilizando mensajes SOAP (Protocolo de Simple Acceso a Objetos), típicamente transportados usando HTTP con una serialización en XML en conjunción con otros estándares relacionados a la Web. (10)

### SOAP

Es un protocolo que permite la comunicación entre aplicaciones a través de mensajes por medio de Internet. Está basado en XML y es la base de los Web Services. Es independiente de la plataforma y del lenguaje. (10)

En la realización del API de comunicación de PHP con R se brindará un servicio web, por el protocolo SOAP. Existirá una aplicación cliente que enviará peticiones (parámetros para calcular o graficar análisis estadístico de datos) al API (RServer 2.0) y este le proveerá la integración del lenguaje de programación PHP con el lenguaje de análisis estadístico R, mostrándole posteriormente un resultado a la aplicación cliente.

### 1.9- Biblioteca utilizada

A continuación se muestra la biblioteca utilizada para la realización del módulo de análisis estadístico de datos.

---

<sup>10</sup> **WSDL** (Web Services Description Language): es un formato XML que se utiliza para describir servicios Web.

## Zelig

Una de las bibliotecas utilizadas en la aplicación es Zelig 2.8-5, la cual se estima que es muy fácil de usar y ayuda a interpretar los rangos de resultados existentes en grandes modelos estadísticos. Es literalmente "El software estadístico de todos" porque Zelig es un marco de trabajo simple y unificado que está incorporado en el motor de R. Cuenta con una infraestructura, que facilita el uso de diferentes métodos, permitiendo multiplicar la información de cualquier modelo, el toma la salida de procedimientos estadísticos y traduce esto en resultados de interés directo. Brinda la posibilidad de realizar análisis estadísticos como: arima (Modelo para Series Temporales), blogit (Regresión Logística Bivariada para dos variables dependientes dicotómicas), bprobit (Regresión Probit para variables dependientes dicotómicas), exp (Regresión Exponencial para variables dependientes de duración), factor.bayes (Análisis de Factor de Bayes), gamma (Regresión de Gamma para variables dependientes positivas), logit (Regresión Logística para variables dependientes dicotómicas), logit.bayes (Regresión Logística Bayesiana para variables dependientes dicotómicas), normal (Regresión Normal para variables dependientes continuas), normal.bayes (Regresión Normal Lineal Bayesiana), poisson(Regresión Poisson para eventos de recuento de variables dependientes), poisson.bayes (Regresión de Poisson Bayesiana), probit (Regresión Probit para variables dependientes dicotómicas), probit.bayes (Regresión Probit Bayesiana para variables dependientes dicotómicas), entre otros. Una gran parte de los análisis estadísticos anteriormente mencionados se implementarán en el módulo. (11)

### 1.10- Métodos de análisis estadístico

A continuación se muestran los métodos de análisis estadísticos implementados en la aplicación, así como una breve descripción de sus aplicaciones y fórmulas matemáticas.

#### Mediana

Es el valor que ocupa el lugar central del conjunto de datos, cuando éstos están ordenados de menor a mayor. Se representa por  $Me$ ,  $Md$  ó  $\tilde{x}$  y se puede hallar sólo para variables cuantitativas.

#### Pasos para calcular la Mediana:

- 1-Se ordenan los datos de menor a mayor.
- 2-Si la serie tiene un número impar de medidas la mediana es la puntuación central de la misma.
- 3-Si la serie tiene un número par de puntuaciones la mediana es la media entre las dos puntuaciones centrales.

Cálculo de la mediana para datos agrupados:

La mediana se encuentra en el intervalo donde la frecuencia acumulada llega hasta la mitad de la suma de las frecuencias absolutas.

$$Md = Li + \frac{n/2 - f_{acum(i-1)}}{f_{mediana}} * A$$

Donde:

$Md$  - Mediana.

$Li$  - límite inferior o frontera inferior de donde se encuentra la mediana, la forma de calcularlo es a través de encontrar la posición  $n/2$ . En ocasiones en el intervalo donde se encuentra la mediana se conoce como intervalo mediano.

$n$  - número de observaciones o frecuencia total.

$f_{acum(i-1)}$  - frecuencia acumulada anterior al intervalo mediano.

$f_{mediana}$  - frecuencia del intervalo mediano.

$A$  - amplitud del intervalo en el que se encuentra la mediana. (12)

Su aplicación se ve limitada pues solo considera el orden jerárquico de los datos y no alguna característica propia de estos, como es el caso de la media. Como propiedad se tiene que: es única, simple y los valores extremos no tienen efectos importantes sobre ella, lo que si ocurre con la media. (12)

**Distribuciones Binomiales**

En estadística, la distribución binomial es una distribución de probabilidad discreta que mide el número de éxitos en una secuencia de  $n$  ensayos de Bernoulli independientes entre sí, con una probabilidad fija  $p$  de ocurrencia de éxito entre los ensayos.

Un experimento de Bernoulli se caracteriza por ser dicotómico, quiere decir que, sólo son posibles dos resultados. A uno de ellos se le denomina éxito y tiene una probabilidad de ocurrencia  $p$  y al otro, fracaso, con una probabilidad  $q = 1 - p$ . (13)

Para representar que una variable aleatoria  $X$  sigue una distribución binomial de parámetros  $n$  y  $p$ , se escribe:

$$X \sim B(n, p)$$

$$p(X = k) = \binom{n}{k} p^k * q^{n-k}$$

Donde:

$n$  - es el número de pruebas.

$k$  - es el número de éxitos.

$p$  - es la probabilidad de éxito.

$q$  - es la probabilidad de fracaso. (13)

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

### Distribución de Poisson

Este tipo de leyes se aplican a sucesos con probabilidad muy baja de ocurrir. Es útil para determinar la probabilidad de que ocurran  $x$  éxitos por unidad de tiempo, área, o producto, la fórmula a utilizar sería:

$$p(x, \lambda) = \frac{\lambda^x e^{-\lambda}}{x!}$$

Donde:

$p(x, \lambda)$  - probabilidad de que ocurran  $x$  éxitos, cuando el número promedio de ocurrencia de ellos es  $\lambda$ .

$\lambda$  - media o promedio de éxitos por unidad de tiempo, área o producto.

$e$  - 2.718.

$x$  - variable que denota el número de éxitos que se desea que ocurra.

Hay que hacer notar que en esta distribución el número de éxitos que ocurren por unidad de tiempo, área o producto es totalmente al azar y que cada intervalo de tiempo es independiente de otro intervalo dado, así como cada área es independiente de otra dada y cada producto es independiente de otro producto. (14)

### Distribución de T de Student

En probabilidad y estadística, la distribución  $t$  (de Student) es una distribución de probabilidad que surge del problema de estimar la media de una población normalmente distribuida cuando el tamaño de la muestra es pequeño.

La distribución  $t$  de Student es la distribución de probabilidad del cociente:  $\frac{Z}{\sqrt{V/v}}$

Donde:

$Z$  - tiene una distribución normal de media nula y varianza 1.

$v$  - tiene una distribución chi-cuadrado con grados de libertad.

$Z$  y  $V$  - son independientes.

Si  $\mu$  es una constante no nula, el cociente  $(Z + \mu) / \sqrt{(V / v)}$  es una variable aleatoria que sigue la distribución t de Student no central con parámetro de no-centralidad  $\mu$ . (15)

### Distribución de Chi-Cuadrado

En estadística, la distribución  $X^2$  (de Pearson), llamada Chi cuadrado o Ji cuadrado, es una distribución de probabilidad continua con un parámetro  $k$  que representa los grados de libertad de la variable aleatoria:  $X = Z_1^2 + \dots + Z_k^2$

Donde  $Z_i$  son variables aleatorias normales independientes de media cero y varianza uno. El que la variable aleatoria  $X$  tenga esta distribución se representa habitualmente así:  $X \sim X_k^2$

#### Función de densidad

$$f(x: k) = \begin{cases} \frac{1}{2^{k/2}} * r\left(\frac{k}{2}\right) * x^{(k/2)-1} * e^{-x/2}, & \text{para } x \geq 0 \\ 0, & \text{para } x < 0 \end{cases}$$

Donde  $r$  es la función gamma.

#### Función de distribución acumulada

$$F_k(x) = \gamma \frac{\left(k/2, x/2\right)}{r\left(k/2\right)}$$

Donde  $\gamma(k, z)$  es la función gamma incompleta.

El valor esperado y la varianza de una variable aleatoria  $X$  con distribución  $X^2$  son, respectivamente,  $k$  y  $2k$ .

La distribución  $X^2$  tiene muchas aplicaciones en inferencia estadística. La más conocida es la de la denominada prueba  $X^2$  utilizada como prueba de independencia y como prueba de bondad de ajuste y en la estimación de varianzas. Pero también está involucrada en el problema de estimar la media de una población normalmente distribuida y en el problema de estimar la pendiente de una recta de regresión lineal, a través de su papel en la distribución T de Student. (16)

### Análisis de varianza para variables dependientes continuas

La varianza es la media aritmética del cuadrado de las desviaciones respecto a la media de una distribución estadística. La varianza se representa por  $\sigma^2$ .

$$\sigma^2 = \frac{\sum_{i=1}^n (X_i - \tilde{x})^2}{N}$$

#### Varianza para datos agrupados

$$\sigma^2 = \frac{\sum_{i=1}^n (X_i - \tilde{x})^2}{N} * f_i$$

- La varianza, al igual que la media, es un índice muy sensible a las puntuaciones extremas.
- En los casos que no se pueda hallar la media tampoco será posible hallar la varianza.
- La varianza no viene expresada en las mismas unidades que los datos, puesto que las desviaciones están elevadas al cuadrado. (17)

### Series Temporales (ARIMA)

Una serie temporal es una secuencia de datos, observaciones o valores, medidos en determinados momentos del tiempo, ordenados cronológicamente y normalmente espaciados entre sí de manera uniforme. El análisis de series temporales comprende métodos que ayudan a interpretar este tipo de datos, extrayendo información representativa, tanto referente a los orígenes o relaciones subyacentes como a la posibilidad de extrapolar y predecir su comportamiento futuro. (18)

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + \dots + a_p X_{t-p} + Z_t + b_1 Z_{t-1} + \dots + b_q Z_{t-q}$$

La fórmula general de los modelos denominados ARMA. Está constituido por una combinación de p términos AR (proceso autorregresivo), y q términos MA (proceso de medias móviles). La parte AR modela la influencia de los valores anteriores de la serie ( $X_{t-1}$  hacia atrás), y la parte MA modela la influencia del ruido en valores anteriores de la serie ( $Z_{t-1}$  hacia atrás), junto con el término  $Z_t$  que corresponde al ruido esperado en el mismo momento t en el que se estima el nuevo valor de la variable X. (18)

Una de las ventajas de estos modelos es su gran simplicidad (sumas de términos), frente a los modelos propuestos en la formulación clásica.

La letra I que aparece en el nombre del modelo completo -ARIMA-, corresponde al proceso último a realizar, una vez definido el tipo de modelo y estimados los coeficientes de éste, ya que entonces hay que restablecer las características originales de la serie de datos, que fue transformada para inducir estacionariedad. A ese proceso inverso se denomina en general Integración y aporta esa letra que completa el nombre. (18)

### Regresión Exponencial

Es aquella en la que la función de ajuste será una función exponencial del tipo:

$$y = a \cdot b^x$$

La regresión exponencial aunque no es lineal se puede convertir en una función lineal tomando logaritmos, porque haciendo el cambio de variable  $v = \log$  se tendría que la función anterior generaría: (19)

$$v = \log_y = \log_{(a \cdot b^x)} = \log_a + x \log_b$$

### Regresión Normal

La importancia de esta distribución radica en que permite modelar numerosos fenómenos naturales, sociales y psicológicos. Mientras que los mecanismos que subyacen a gran parte de este tipo de fenómenos son desconocidos, por la enorme cantidad de variables incontrolables que en ellos intervienen, el uso del modelo normal puede justificarse asumiendo que cada observación se obtiene como la suma de unas pocas causas independientes.

El modelo de regresión lineal normal clásico considera que la relación entre la variable dependiente ( $Y$ ) y las independientes ( $X_1, X_2, \dots, X_k$ ) se puede formular matricialmente a partir de la siguiente expresión lineal:  $Y = X\beta + u$  que desarrollándose quedaría así: (20)

$$Y_i = \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_K X_{iK} + u_i \quad , \quad \text{Donde } i = 1, 2, \dots, n$$

### Regresión de Cox

En el análisis de la supervivencia se le conoce como Regresión de Riesgos Proporcionales de Cox a una clase de modelos usados para modelar los riesgos que afectan a la supervivencia de una población de sujetos y en este caso para variables dependientes de duración.

En el modelo de regresión de Cox la función de riesgo se construye como:

$$h_{i(t)} = h_0(t) \exp(\beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip})$$

La función de riesgo base o de referencia  $h_0(t)$  no queda especificada y puede tomar cualquier forma. Al término  $\beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip}$  se le denomina puntuación de riesgo, ya que un valor negativo grande corresponde a un perfil de riesgo menor que la media, mientras que un valor positivo grande de esa puntuación corresponde a un perfil de riesgo mayor que la media. (21)

### Factor de Análisis Bayesiano

Algunos de los campos de aplicación de este factor está en la teoría de la decisión, visión artificial (simulación de la percepción en general) y reconocimiento de patrones por ordenador. (22)

$$P(H_0|E) = \frac{P(E|H_0)P(H_0)}{P(E)}$$

Donde:

$H_0$  - representa la hipótesis nula, que ha sido inferida antes de que la nueva evidencia,  $E$  resultará disponible.

$P(H_0)$  - se llama la probabilidad a priori de  $H_0$ .

$P(H_0|E)$  - se llama la probabilidad condicional de que se cumpla la evidencia  $E$  dado que la hipótesis  $H_0$  es verdadera. Se llama también la función de verosimilitud cuando se expresa como una función de  $E$  dado  $H_0$ .

$P(E)$  - se llama la probabilidad marginal de  $E$ . Se puede calcular como:  $\sum P(E|H_i)P(H_i)$

### Regresión de Poisson para variables de recuento

Esta Regresión se puede hacer derivar de un proceso experimental de observación en el que se tenga las siguientes características:

- Se observa la realización de hechos de cierto tipo durante un cierto período de tiempo o a lo largo de un espacio de observación.
- Los hechos a observar tienen naturaleza aleatoria; pueden producirse o no de una manera no determinada.
- La probabilidad de que se produzcan un número  $x$  de éxitos en un intervalo de amplitud  $t$  no depende del origen del intervalo (Aunque, sí de su amplitud).
- La probabilidad de que ocurra un hecho en un intervalo infinitésimo es prácticamente proporcional a la amplitud del intervalo.
- La probabilidad de que se produzcan 2 ó más hechos en un intervalo infinitésimo es un infinitésimo de orden superior a dos. En consecuencia, en un intervalo infinitésimo podrán producirse 0 ó 1 hecho pero nunca más de uno.
- Si en estas circunstancias se hacen aleatorias de forma que la variable aleatoria  $X$  signifique o designe el "número de hechos que se producen en un intervalo de tiempo o de espacio", la variable  $X$  se distribuye con una distribución de parámetro  $\lambda$  (media o promedio). (11)

Entre los métodos implementados también se encuentran: la Regresión Logística Bivariada de dos variables dependientes y la Regresión Gamma para variables dependientes, continuas y positivas.

**Conclusiones del Capítulo:**

Se hizo una variada y profunda investigación de software y lenguajes que realizan análisis estadístico a nivel mundial, para escoger el lenguaje más apropiado para el módulo de análisis estadístico de datos del RServer 2.0. Llegando a la conclusión de que el lenguaje más indicado es R.

También se seleccionaron las herramientas, tecnologías, lenguajes de programación, servidores web y la metodología apoyándose de sus ventajas y desventajas y se determinó de que: como lenguaje de programación web a PHP en su versión 5.3, como marco de trabajo para PHP Symfony 2.0, como servidor web se escogió Apache 2.0, como metodología de desarrollo de software Open Up y la herramienta CASE Visual Paradigm, para el modelado de diagramas. Por último seleccionó como IDE NetBeans en su versión 7.0. En este capítulo también se expuso una breve descripción matemática de los métodos de análisis estadístico que forman parte de la propuesta de solución. Todos ellos escogidos de la biblioteca Zelig. También se determinó el uso del protocolo SOAP para brindar los servicios web. Todos estos aspectos dan paso al capítulo 2.

## CAPÍTULO 2: ANÁLISIS Y DISEÑO DEL SISTEMA

Este capítulo tiene como objetivos, especificar los requisitos funcionales y no funcionales del sistema. Así como modelar los diagramas correspondientes al diseño. Además de determinar cuáles serán los patrones de diseño y de arquitectura que se van a usar en la implementación del módulo, quedando estructurada la arquitectura para el RServer 2.0.

### 2.1 - Modelo de Dominio

Puede utilizarse para capturar y expresar el entendimiento ganado en un área bajo análisis como paso previo al diseño de un sistema. Por lo que el mapa de conceptos del modelo de dominio constituye una primera versión del sistema. Este modelo permite mostrar de manera visual los principales conceptos que se manejan, ayudando a los usuarios, desarrolladores e interesados; a utilizar un vocabulario común para poder entender el contexto en que se desarrolla el sistema. Además contribuye a identificar personas, eventos, transacciones y objetos involucrados en el sistema. (23)

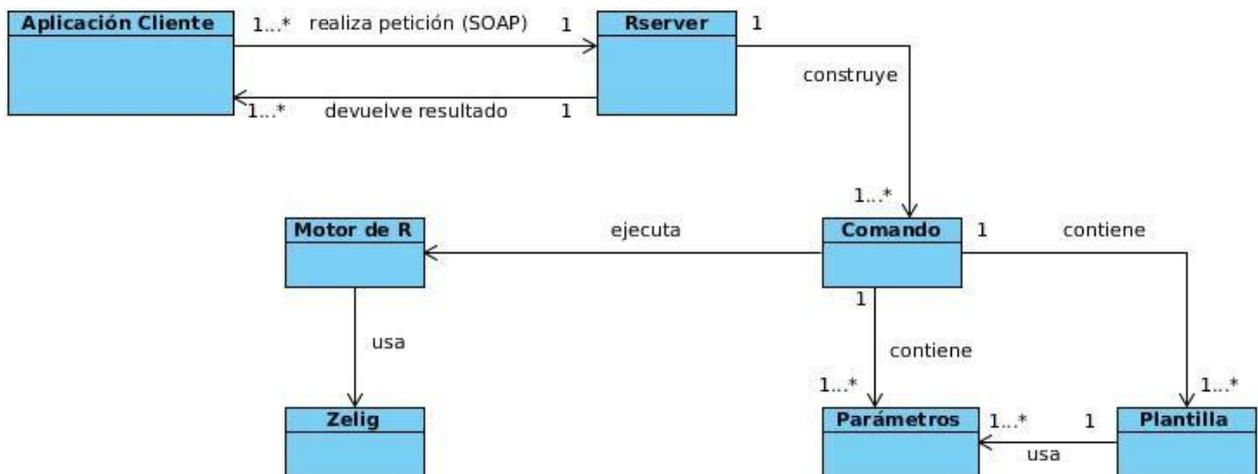


Figura 1. Modelo de Dominio

### Descripción del Modelo del Dominio

La Aplicación Cliente realiza una petición por el protocolo SOAP al RServer 2.0 enviando un conjunto de parámetros, este a su vez construye un Comando. El Comando está constituido por las Plantillas y por los Parámetros enviados por la Aplicación Cliente. La Plantilla usa estos Parámetros para su posterior cálculo o para graficar. Ya conformado el comando, el mismo se ejecuta en el Motor de R. R usa una biblioteca llamada Zelig, que tiene importantes métodos de análisis estadístico que son

necesarios para la implementación del Módulo para Análisis Estadístico. Finalmente el RServer 2.0 devuelve un resultado a dicha Aplicación Cliente en un formato JSON.

## 2.2- Requisitos de Software

Los requisitos de software se dividen en dos categorías, en requisitos funcionales que son los que están relacionados con las condiciones que el sistema debe cumplir y los no funcionales que por su parte, son los que imponen restricciones en el diseño o la implementación. Son las propiedades o cualidades que el producto debe tener.

### 2.2.1- Requisitos Funcionales

- RF\_1 Realizar análisis estadístico de datos por el método de la distribución acumulada de Poisson.
- RF\_2 Realizar análisis estadístico de datos por el método de la distribución de cuantiles de Poisson.
- RF\_3 Realizar análisis estadístico de datos por el método de la densidad de Poisson.
- RF\_4 Realizar análisis estadístico de datos generando números aleatorios de Poisson.
- RF\_5 Realizar análisis estadístico de datos por el método de la distribución acumulada de T-Student.
- RF\_6 Realizar análisis estadístico de datos por el método de la distribución de cuantiles de T-Student.
- RF\_7 Realizar análisis estadístico de datos por el método de la densidad de T-Student.
- RF\_8 Realizar análisis estadístico de datos generando números aleatorios de T-Student.
- RF\_9 Realizar análisis estadístico de datos por el método de la distribución acumulada Binomial.
- RF\_10 Realizar análisis estadístico de datos por el método de la distribución simple Binomial.
- RF\_11 Realizar análisis estadístico de datos por el método de la Mediana.
- RF\_12 Realizar análisis estadístico de datos por el método de la distribución acumulada de Chi-Cuadrado.
- RF\_13 Realizar análisis estadístico de datos por el método de la densidad de Chi-Cuadrado.
- RF\_14 Realizar análisis estadístico de datos por el método de la distribución de cuantiles de Chi-Cuadrado.
- RF\_15 Realizar análisis estadístico de datos generando números aleatorios de Chi-Cuadrado.
- RF\_16 Realizar análisis estadístico de datos de varianza para variables dependientes continuas.
- RF\_17 Realizar análisis estadístico para datos de series temporales, estimando un modelo ARIMA con regresores exógenos, además de los errores y retardados valores de la variable dependiente.
- RF\_18 Realizar análisis estadístico para datos de series temporales, estimando el modelo ARIMA, y resumiendo los resultados.
- RF\_19 Realizar análisis estadístico para datos de regresión logística bivariada de dos variables dependientes dicotómicas, estimando el modelo (blogit).

- RF\_20 Realizar análisis estadístico para datos de regresión logística bivariada de dos variables dependientes dicotómicas, con la estimación conjunta de un modelo con diferentes conjuntos de variables explicativas (blogit).
- RF\_21 Realizar análisis estadístico para datos de regresión logística bivariada de dos variables dependientes dicotómicas, estimando el modelo (probit).
- RF\_22 Realizar análisis estadístico para datos de regresión logística bivariada de dos variables dependientes dicotómicas, con la estimación conjunta de un modelo con diferentes conjuntos de variables explicativas (probit).
- RF\_23 Realizar análisis estadístico para datos de regresión de riesgos proporcionales de Cox para variables dependientes de duración, estimando los valores de los parámetros de la regresión coxph.
- RF\_24 Realizar análisis estadístico para datos de regresión exponencial de las variables dependientes de duración.
- RF\_25 Realizar análisis estadístico para datos de regresión del factor de análisis bayesiano.
- RF\_26 Realizar análisis estadístico para datos de regresión Gamma para variables dependientes continuas y positivas.
- RF\_27 Realizar análisis estadístico para datos de regresión normal para variables dependientes continuas.
- RF\_28 Realizar análisis estadístico para datos de regresión Poisson para variables de recuento de eventos dependientes.
- RF\_29 Graficar análisis estadístico.

### 2.2.2- Requisitos no Funcionales

Entre los requisitos no funcionales se encuentran los de hardware, que son los elementos que se debe disponer, para que el sistema implementado cumpla con sus funcionalidades. Y otro de los requerimientos existentes son los de software. Estos últimos plantean las tecnologías que debe tener instalada en el servidor donde se despliegue.

#### Requisitos de Software

Se debe tener instalado en el servidor donde se despliegue el RServer 2.0, el motor de R en su versión 2.13, Symfony 2.0, Apache 2.0 y PHP 5.3.

#### Requisitos de Hardware

Se investigó sobre los requerimientos mínimos que requiere el servidor web Apache 2.0 y el lenguaje de análisis estadístico R, llegando a la conclusión de que el sistema requiere de una memoria RAM de

256MB con un ordenador Pentium III ó equivalente. Dejando un espacio de disco duro de 20GB para la instalación de las tecnologías necesarias para el despliegue de la aplicación.

### 2.3- Diagrama de Caso de Uso del Sistema

Los diagramas de caso de uso modelan el comportamiento del sistema. Muestra un los actores, los casos de uso y las relaciones que pueden existir entre ellos. Una vez especificado los requisitos funcionales, se agruparon en dos casos de uso, Realizar Análisis Estadístico y Graficar Análisis Estadístico.

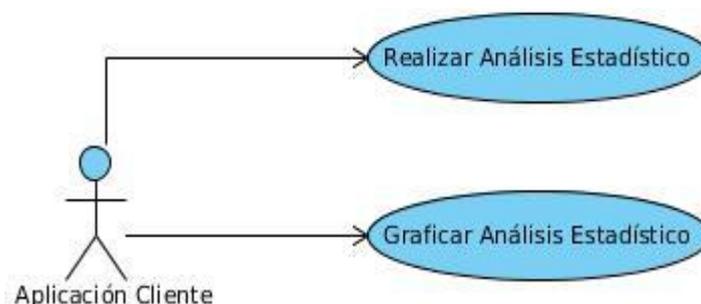


Figura 2. Diagrama de Caso de Uso del Sistema

#### Descripción del Diagrama de Casos de Uso del Sistema

El actor de este diagrama es cualquier aplicación cliente que necesite realizar un análisis estadístico o generar una gráfica.

<b>Objetivo</b>	Realizar un análisis estadístico con los parámetros que la aplicación cliente envía por el protocolo SOAP.
<b>Actores</b>	Aplicación Cliente
<b>Resumen</b>	Realiza análisis estadístico de los parámetros enviados por la aplicación cliente. Estos parámetros tienen un formato JSON. A partir de ellos, se ejecuta la funcionalidad correspondiente con el tipo de análisis y a esta funcionalidad se le pasa el resto de los parámetros que serían los datos a calcular. Una vez realizado el cálculo, el RServer devuelve una respuesta a la aplicación cliente en formato JSON.
<b>Complejidad</b>	Alta
<b>Prioridad</b>	Crítico
<b>Precondiciones</b>	El cliente ha enviado los parámetros correctamente.
<b>Postcondiciones</b>	Se realizó correctamente el análisis y se envió un resumen del análisis en formato JSON.
<b>Flujo de eventos</b>	

Flujo básico Realizar análisis estadístico de datos		
	Actor	Sistema
1.	Envía los parámetros a calcular en formato JSON con la funcionalidad a ejecutar.	
2.		<p>Según los parámetros recibidos por la aplicación cliente que serían los datos a calcular, se ejecuta la funcionalidad correspondiente al tipo de análisis y se envía a una de las secciones siguientes.</p> <ol style="list-style-type: none"> <li>1- Realizar análisis estadístico de acumulada de Poisson.</li> <li>2- Realizar análisis estadístico de cuantiles de Poisson.</li> <li>3- Realizar análisis estadístico de densidad de Poisson.</li> <li>4- Generar números aleatorios de Poisson.</li> <li>5- Realizar análisis estadístico de acumulada de T-Student.</li> <li>6- Realizar análisis estadístico de cuantiles de T-Student.</li> <li>7- Realizar análisis estadístico de densidad de T-Student.</li> <li>8- Generar números aleatorios de T-Student.</li> <li>9- Realizar análisis estadístico de acumulada Binomial.</li> <li>10- Realizar análisis estadístico de simple Binomial.</li> <li>11- Realizar análisis estadístico de Mediana.</li> <li>12- Realizar análisis estadístico de acumulada Chi-Cuadrado.</li> <li>13- Realizar análisis estadístico de densidad Chi-Cuadrado.</li> <li>14- Realizar análisis estadístico de cuantiles Chi-Cuadrado.</li> <li>15- Generar números aleatorios de Chi-Cuadrado.</li> <li>16- Realizar análisis estadístico de Varianza.</li> <li>17- Realizar análisis estadístico de Series Temporales estimando un modelo.</li> <li>18- Realizar análisis estadístico de Series Temporales estimando un modelo y resumiendo los resultados.</li> <li>19- Realizar análisis estadístico de Blogit (regresión logística estimando un modelo)</li> <li>20- Realizar análisis estadístico de Blogit (regresión logística con estimación conjunta)</li> <li>21- Realizar análisis estadístico de Probit (regresión logística estimando un modelo)</li> <li>22- Realizar análisis estadístico de Probit (regresión logística con estimación</li> </ol>

		conjunta) 23- Realizar análisis estadístico de regresión de Riesgos Proporcionales. 24- Realizar análisis estadístico de regresión Exponencial. 25- Realizar análisis estadístico de Factor de Bayes. 26- Realizar análisis estadístico de regresión Gamma. 27- Realizar análisis estadístico de regresión Normal. 28- Realizar análisis estadístico de Variables de Recuento de Poisson.
3.		Termina el caso de uso.
<b>Flujos alternos</b>		
<b>Nº 1. Datos enviados mal estructurados</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Enviará un mensaje de error. Por lo que es recomendable visitar la ayuda para ver la estructura que deben llevar los parámetros.
<b>Sección 1: “Realizar análisis estadístico de acumulada de Poisson”</b>		
<b>Flujo básico Realizar análisis estadístico de acumulada de Poisson</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un vector de cuantiles, un vector de media y dos variables lógico <sup>11</sup> . Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 2: “Realizar análisis estadístico de cuantiles de Poisson”</b>		
<b>Flujo básico Realizar análisis estadístico de cuantiles de Poisson</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un vector de probabilidades, un vector de media y dos variables lógico. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 3: “Realizar análisis estadístico de densidad de Poisson”</b>		
<b>Flujo básico Realizar análisis estadístico de densidad de Poisson</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un vector de número entero no negativo, un vector de media no negativo y dos variables lógico. Todos, en

<sup>11</sup> **Lógico:** estructura de datos, que puede tomar valor verdadero o falso.

		ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 4: “Generar números aleatorios de Poisson”</b>		
<b>Flujo básico Generar números aleatorios de Poisson</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un número de valores aleatorios para generar, un vector de media no negativo y dos variables lógico. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 5: “Realizar análisis estadístico de acumulada de T-Student”</b>		
<b>Flujo básico Realizar análisis estadístico de acumulada de T-Student</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un vector de cuantiles, los grados de libertad, un parámetro que determina la centralidad de una gráfica t-Student. Si se omite, el estudio se realiza con la gráfica centralizada en 0, en este caso se puede omitir, pues no se va a graficar y dos variables lógico. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 6: “Realizar análisis estadístico de cuantiles de T-Student”</b>		
<b>Flujo básico Realizar análisis estadístico de cuantiles de T-Student</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un vector de probabilidades, los grados de libertad, un parámetro que determina la centralidad de una gráfica t-Student. Si se omite, el estudio se realiza con la gráfica centralizada en 0, en este caso se puede omitir, pues no se va a graficar y dos variables lógico. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 7: “Realizar análisis estadístico de densidad de T-Student”</b>		
<b>Flujo básico Realizar análisis estadístico de densidad de T-Student</b>		

	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un vector de cuantiles, los grados de libertad y un parámetro que determina la centralidad de una gráfica t-Student. Si se omite, el estudio se realiza con la gráfica centralizada en 0, en este caso se puede omitir, pues no se va a graficar. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 8: “Generar números aleatorios de T-Student”</b>		
<b>Flujo básico Generar números aleatorios de T-Student</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser números de observaciones, los grados de libertad y un parámetro que determina la centralidad de una gráfica t-Student. Si se omite, el estudio se realiza con la gráfica centralizada en 0, en este caso se puede omitir, pues no se va a graficar. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 9: “Realizar análisis estadístico de acumulada Binomial”</b>		
<b>Flujo básico Realizar análisis estadístico de acumulada Binomial</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un número de éxitos, número de repeticiones y la probabilidad de éxitos. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 10: “Realizar análisis estadístico de simple Binomial”</b>		
<b>Flujo básico Realizar análisis estadístico de simple Binomial</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un vector ó número de éxitos, número de repeticiones y la probabilidad de éxitos. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 11: “Realizar análisis estadístico de Mediana”</b>		

<b>Flujo básico Realizar análisis estadístico de Mediana</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser cualquier vector de una o dos dimensiones ó un solo número. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 12: “Realizar análisis estadístico de acumulada Chi-Cuadrado”</b>		
<b>Flujo básico Realizar análisis estadístico de acumulada Chi-Cuadrado</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un vector de cuantiles, los grados de libertad, un parámetro que determina la centralidad de una gráfica t-Student. Si se omite, el estudio se realiza con la gráfica centralizada en 0, en este caso se puede omitir, pues no se va a graficar y dos variables lógico. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 13: “Realizar análisis estadístico de densidad Chi-Cuadrado”</b>		
<b>Flujo básico Realizar análisis estadístico de densidad Chi-Cuadrado</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un vector de cuantiles, los grados de libertad, un parámetro que determina la centralidad de una gráfica t-Student. Si se omite, el estudio se realiza con la gráfica centralizada en 0, en este caso se puede omitir, pues no se va a graficar y una variable lógico. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 14: “Realizar análisis estadístico de cuantiles Chi-Cuadrado”</b>		
<b>Flujo básico Realizar análisis estadístico de cuantiles Chi-Cuadrado</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un vector de probabilidades, los grados de libertad, un parámetro que determina la centralidad de una gráfica t-Student. Si se omite, el estudio se realiza con la gráfica centralizada

		en 0, en este caso se puede omitir, pues no se va a graficar y dos variables lógico. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 15: “Generar números aleatorios de Chi-Cuadrado”</b>		
<b>Flujo básico Generar números aleatorios de Chi-Cuadrado</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un número de observaciones, los grados de libertad y un parámetro que determina la centralidad de una gráfica t-Student. Si se omite, el estudio se realiza con la gráfica centralizada en 0, en este caso se puede omitir, pues no se va a graficar. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 16: “Realizar análisis estadístico de Varianza”</b>		
<b>Flujo básico Realizar análisis estadístico de Varianza</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un arreglo de datos de igual dimensión, el número de filas, el número de columnas, los nombres de las columnas, la variable independiente, un arreglo de datos de igual dimensión de variables dependientes y un valor lógico que ordena por filas si es TRUE, en caso de ser FALSE ordena por columnas. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 17: “Realizar análisis estadístico de Series Temporales estimando un modelo”</b>		
<b>Flujo básico Realizar análisis estadístico de Series Temporales estimando un modelo</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un arreglo de datos de igual dimensión, el número de filas, el número de columnas, los nombres de las columnas, la variable independiente, un arreglo de datos de igual dimensión de variables dependientes y un valor lógico que ordena por filas si es TRUE, en caso de ser FALSE ordena por columnas. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.

2.		Devuelve el resultado en formato JSON.
<b>Sección 18: “Realizar análisis estadístico de Series Temporales estimando un modelo y resumiendo los resultados”</b>		
<b>Flujo básico Realizar análisis estadístico de Series Temporales estimando un modelo y resumiendo los resultados</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un arreglo de datos de igual dimensión, el número de filas, el número de columnas, los nombres de las columnas, la variable independiente y un valor lógico que ordena por filas si es TRUE, en caso de ser FALSE ordena por columnas. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 19: “Realizar análisis estadístico de Blogit (regresión logística estimando un modelo)”</b>		
<b>Flujo básico Realizar análisis estadístico de Blogit (regresión logística estimando un modelo)</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un arreglo de datos de igual dimensión, el número de filas, el número de columnas, los nombres de las columnas, la primera variable dependiente, la segunda variable dependiente, un arreglo de datos de igual dimensión de variables independientes y un valor lógico que ordena por filas si es TRUE, en caso de ser FALSE ordena por columnas. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 20: “Realizar análisis estadístico de Blogit (regresión logística con estimación conjunta)”</b>		
<b>Flujo básico Realizar análisis estadístico de Blogit (regresión logística con estimación conjunta)</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un arreglo de datos de igual dimensión, el número de filas, el número de columnas, los nombres de las columnas, la primera variable dependiente, la segunda variable dependiente, la variable independiente, el primer arreglo de datos de igual dimensión de variables independientes, el segundo arreglo de datos de igual dimensión de variables independientes y un valor lógico que ordena por filas si es TRUE, en caso de ser FALSE ordena por columnas. Todos, en ese orden.

		Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 21: “Realizar análisis estadístico de Probit (regresión logística estimando un modelo)”</b>		
<b>Flujo básico Realizar análisis estadístico de Probit (regresión logística estimando un modelo)</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un arreglo de datos de igual dimensión, el número de filas, el número de columnas, los nombres de las columnas, la primera variable dependiente, la segunda variable dependiente, un arreglo de datos de igual dimensión de variables independientes y un valor lógico que ordena por filas si es TRUE, en caso de ser FALSE ordena por columnas. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 22: “Realizar análisis estadístico de Probit (regresión logística con estimación conjunta)”</b>		
<b>Flujo básico Realizar análisis estadístico de Probit (regresión logística con estimación conjunta)</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un arreglo de datos de igual dimensión, el número de filas, el número de columnas, los nombres de las columnas, la primera variable dependiente, la segunda variable dependiente, un primer arreglo de datos de igual dimensión de variables independientes, un segundo arreglo de datos de igual dimensión de variables independientes y un valor lógico que ordena por filas si es TRUE, en caso de ser FALSE ordena por columnas. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 23: “Realizar análisis estadístico de regresión de Riesgos Proporcionales”</b>		
<b>Flujo básico Realizar análisis estadístico de regresión de Riesgos Proporcionales</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un arreglo de datos de igual dimensión, el número de filas, el número de columnas, los nombres de las columnas, la primera variable dependiente, la segunda variable dependiente, un arreglo de datos de igual dimensión de variables independientes y dos valores lógico uno que ordena por filas si es TRUE, en caso de ser FALSE

		ordena por columnas y el otro es TRUE por defecto. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 24: “Realizar análisis estadístico de regresión Exponencial”</b>		
<b>Flujo básico Realizar análisis estadístico de regresión Exponencial</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un arreglo de datos de igual dimensión, el número de filas, el número de columnas, los nombres de las columnas, una variable dependiente, un arreglo de datos de igual dimensión de variables independientes y un valor lógico que ordena por filas si es TRUE, en caso de ser FALSE ordena por columnas. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 25: “Realizar análisis estadístico de Factor de Bayes”</b>		
<b>Flujo básico Realizar análisis estadístico de Factor de Bayes</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un arreglo de datos de igual dimensión, el número de filas, el número de columnas, los nombres de las columnas, un arreglo de datos de igual dimensión de variables independientes y un valor lógico que ordena por filas si es TRUE, en caso de ser FALSE ordena por columnas. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 26: “Realizar análisis estadístico de regresión Gamma”</b>		
<b>Flujo básico Realizar análisis estadístico de regresión Gamma</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un arreglo de datos de igual dimensión, el número de filas, el número de columnas, los nombres de las columnas, una variable independiente, un arreglo de datos de igual dimensión de variables dependientes y un valor lógico que ordena por filas si es TRUE, en caso de ser FALSE ordena por columnas. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.

<b>Sección 27: “Realizar análisis estadístico de regresión Normal”</b>		
<b>Flujo básico Realizar análisis estadístico de regresión Normal</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un arreglo de datos de igual dimensión, el número de filas, el número de columnas, los nombres de las columnas, una variable independiente, un arreglo de datos de igual dimensión de variables dependientes y un valor lógico que ordena por filas si es TRUE, en caso de ser FALSE ordena por columnas. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 28: “Realizar análisis estadístico de Variables de Recuento de Poisson”</b>		
<b>Flujo básico Realizar análisis estadístico de Variables de Recuento de Poisson</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un arreglo de datos de igual dimensión, el número de filas, el número de columnas, los nombres de las columnas, una variable independiente, un arreglo de datos de igual dimensión de variables dependientes y un valor lógico que ordena por filas si es TRUE, en caso de ser FALSE ordena por columnas. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Relaciones</b>	<b>CU Incluidos</b>	-
	<b>CU Extendidos</b>	-
<b>Requisitos no funcionales</b>	Se investigó sobre los requerimientos mínimos que requiere el servidor web Apache 2.0 y el lenguaje de análisis estadístico R, llegando a la conclusión de que el sistema requiere de una memoria RAM de 256MB con un ordenador Pentium III. Dejando un espacio de disco duro de 20GB para la instalación de las tecnologías necesarias para el despliegue de la aplicación.	
<b>Asuntos pendientes</b>	-	

Tabla 1. Descripción del Caso de Uso Realizar Análisis Estadístico

<b>Objetivo</b>	Graficar análisis estadísticos de los parámetros enviados por la aplicación cliente, por el protocolo SOAP.	
<b>Actores</b>	Aplicación Cliente	
<b>Resumen</b>	La aplicación cliente envía parámetros en formato JSON, especificando la funcionalidad del tipo de análisis que quiere graficar. Una vez realizada la gráfica, el sistema envía un JSON a la aplicación cliente conformado por el nombre y la dirección donde se encuentra la imagen.	
<b>Complejidad</b>	Alta	
<b>Prioridad</b>	Media	
<b>Precondiciones</b>	El cliente ha enviado los parámetros correctamente.	
<b>Postcondiciones</b>	Se realizó correctamente el análisis para obtener la gráfica. Enviando a la aplicación cliente una dirección de donde está ubicada en el servidor la imagen, previamente conformada con su nombre, en un formato JSON.	
<b>Flujo de eventos</b>		
<b>Flujo básico Graficar Análisis Estadístico de Datos</b>		
	<b>Actor</b>	<b>Sistema</b>
1.	Envía los parámetros a calcular en formato JSON.	
2.		<p>Según los parámetros recibidos por la aplicación cliente que serían los datos a graficar, se ejecuta la funcionalidad correspondiente al tipo de análisis y se envía a una de las secciones siguientes.</p> <ul style="list-style-type: none"> <li>1- Graficar análisis estadístico de Mediana.</li> <li>2- Graficar análisis estadístico de Varianza.</li> <li>3- Graficar análisis estadístico de Series Temporales estimando un modelo y resumiendo los resultados.</li> <li>4- Graficar análisis estadístico de Probit (regresión logística estimando un modelo).</li> <li>5- Graficar análisis estadístico de Probit (regresión logística con estimación conjunta).</li> <li>6- Graficar análisis estadístico de regresión de Riesgos Proporcionales.</li> <li>7- Graficar análisis estadístico de regresión Exponencial.</li> <li>8- Graficar análisis estadístico de regresión Normal.</li> <li>9- Graficar análisis estadístico de Variables de Recuento de Poisson.</li> </ul>
3.		Termina el caso de uso.
<b>Flujos alternos</b>		
<b>Nº 1 Datos enviados mal estructurados.</b>		

	<b>Actor</b>	<b>Sistema</b>
1.		Enviará un mensaje de error. Por lo que es recomendable visitar la ayuda para ver la estructura que deben llevar los parámetros.
<b>Sección 1: “Graficar análisis estadístico de Variables de Recuento de Mediana”</b>		
<b>Flujo básico Graficar análisis estadístico de Variables de Recuento de Mediana</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser cualquier vector de una o dos dimensiones ó un solo número. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 2: “Graficar análisis estadístico de Varianza”</b>		
<b>Flujo básico Graficar análisis estadístico de Varianza</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un arreglo de datos de igual dimensión, el número de filas, el número de columnas, los nombres de las columnas, la variable independiente, un arreglo de datos de igual dimensión de variables dependientes y un valor lógico que ordena por filas si es TRUE, en caso de ser FALSE ordena por columnas. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 3: “Graficar análisis estadístico de Series Temporales estimando un modelo y resumiendo los resultados”</b>		
<b>Flujo básico Graficar análisis estadístico de Series Temporales estimando un modelo y resumiendo los resultados</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un arreglo de datos de igual dimensión, el número de filas, el número de columnas, los nombres de las columnas, la variable independiente y un valor lógico que ordena por filas si es TRUE, en caso de ser FALSE ordena por columnas. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.

<b>Sección 4: “Graficar análisis estadístico de Probit (regresión logística estimando un modelo)”</b>		
<b>Flujo básico Graficar análisis estadístico de Probit (regresión logística estimando un modelo)</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un arreglo de datos de igual dimensión, el número de filas, el número de columnas, los nombres de las columnas, la primera variable dependiente, la segunda variable dependiente, un arreglo de datos de igual dimensión de variables independientes y un valor lógico que ordena por filas si es TRUE, en caso de ser FALSE ordena por columnas. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 5: “Graficar análisis estadístico de Probit (regresión logística con estimación conjunta)”</b>		
<b>Flujo básico Graficar análisis estadístico de Probit (regresión logística con estimación conjunta)</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un arreglo de datos de igual dimensión, el número de filas, el número de columnas, los nombres de las columnas, la primera variable dependiente, la segunda variable dependiente, un primer arreglo de datos de igual dimensión de variables independientes, un segundo arreglo de datos de igual dimensión de variables independientes y un valor lógico que ordena por filas si es TRUE, en caso de ser FALSE ordena por columnas. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Sección 6: “Graficar análisis estadístico de regresión de Riesgos Proporcionales”</b>		
<b>Flujo básico Graficar análisis estadístico de regresión de Riesgos Proporcionales</b>		
	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un arreglo de datos de igual dimensión, el número de filas, el número de columnas, los nombres de las columnas, la primera variable dependiente, la segunda variable dependiente, un arreglo de datos de igual dimensión de variables independientes y dos valores lógico uno

		que ordena por filas si es TRUE, en caso de ser FALSE ordena por columnas y el otro es TRUE por defecto. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.

**Sección 7: “Graficar análisis estadístico de regresión Exponencial”**

**Flujo básico Graficar análisis estadístico de regresión Exponencial**

	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un arreglo de datos de igual dimensión, el número de filas, el número de columnas, los nombres de las columnas, una variable dependiente, un arreglo de datos de igual dimensión de variables independientes y un valor lógico que ordena por filas si es TRUE, en caso de ser FALSE ordena por columnas. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.

**Sección 8: “Graficar análisis estadístico de regresión Normal”**

**Flujo básico Graficar análisis estadístico de regresión Normal**

	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un arreglo de datos de igual dimensión, el número de filas, el número de columnas, los nombres de las columnas, una variable independiente, un arreglo de datos de igual dimensión de variables dependientes y un valor lógico que ordena por filas si es TRUE, en caso de ser FALSE ordena por columnas. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.

**Sección 9: “Graficar análisis estadístico de Variables de Recuento de Poisson”**

**Flujo básico Graficar análisis estadístico de Variables de Recuento de Poisson**

	<b>Actor</b>	<b>Sistema</b>
1.		Ejecuta una plantilla predeterminada en R, pasándole dichos parámetros. Cuyos parámetros de esta sección deberían ser un arreglo de datos de igual dimensión, el número de filas, el número de columnas, los nombres de las columnas, una variable independiente, un arreglo de datos de igual dimensión de variables dependientes y un

		valor lógico que ordena por filas si es TRUE, en caso de ser FALSE ordena por columnas. Todos, en ese orden. Esta a la vez va al motor de R y realiza el cálculo.
2.		Devuelve el resultado en formato JSON.
<b>Relaciones</b>	<b>CU Incluidos</b>	-
	<b>CU Extendidos</b>	-
<b>Requisitos no funcionales</b>	Se investigó sobre los requerimientos mínimos que requiere el servidor web Apache 2.0 y el lenguaje de análisis estadístico R, llegando a la conclusión de que el sistema requiere de una memoria RAM de 256MB con un ordenador Pentium III ó equivalente. Dejando un espacio de disco duro de 20GB para la instalación de las tecnologías necesarias para el despliegue de la aplicación.	
<b>Asuntos pendientes</b>	-	

Tabla 2. Descripción del Caso de Uso Graficar Análisis Estadístico

## 2.4- Patrones de Diseño y de Arquitectura de Software

Los patrones arquitectónicos son patrones de diseño de software que ofrecen soluciones a problemas en la arquitectura de un sistema. Dan una descripción de los elementos y el tipo de relación que tienen unidos con un conjunto de restricciones sobre cómo pueden ser usados y expresan un esquema de organización estructural esencial. En comparación con los patrones de diseño, los patrones arquitectónicos tienen un nivel de abstracción mayor. (24)

Durante la elaboración del sistema se hizo empleo de diferentes patrones de diseño y/o arquitectónicos. Dentro de los cuales se destacan la inyección de dependencias y el MVC.

- **Modelo-Vista-Controlador-** Este patrón de arquitectura de software separa los datos de la aplicación, la interfaz de usuario, y la lógica de negocio en tres componentes distintos. (25).

Dentro de las tecnologías utilizadas se encuentra Symfony 2.0, la misma proporciona un avanzado MVC, el cual se utiliza para crear la estructura de la solución final. En el diagrama de clases del diseño se modeló solamente el controlador, puesto que la vista estaría dada por el envío JSON que se produce al ejecutar una funcionalidad por la aplicación cliente, siendo esta, quien proponga la interfaz. No es objetivo de este trabajo usar bases de datos, por lo que no se presenta un modelo; pero queda abierta la opción para posteriores integraciones de otros módulos al RServer 2.0.

- **Inyección de Dependencias-** El patrón de inyección de dependencias consiste en hacer que piezas de software sean independientes comunicándose únicamente a través de una interfaz. Esto implica muchas modificaciones en el código fuente como el uso de implementaciones, la eliminación de la instanciación de objetos mediante la instrucción new o la necesidad de un modo de configuración que indique que clases se instanciarán en el caso de solicitarlo. (26)  
La inyección de dependencias es un patrón orientado a objetos, que suministra a una clase un objeto, evitando que esta clase sea quien lo cree. En la solución propuesta se hace uso de este patrón, puesto que garantiza un diseño modular, independiente, flexible y que se pueda reutilizar con facilidad.

### **2.5- Diagramas de Clases del Diseño**

Un diagrama de clases describe la estructura del sistema. Exponiendo sus clases, atributos y las relaciones que pueden existir entre ellas. Así como los subsistemas que intervienen en este proceso. A continuación se muestra el diagrama de clases del diseño que se generó como solución.



### Descripción del Diagrama de Clases del Diseño

Inicialmente se encuentra el controlador frontal `app_dev.php`, el mismo inicia el flujo de Symfony. A la clase `RserverApiController.php` llegan los parámetros enviados por la aplicación cliente, esta, a su vez le pasa los parámetros a `Rservices.php` mediante la inyección de dependencias, previamente configurada. `Rservices.php` es la clase que se encarga de vincular el motor de R con el lenguaje PHP, la misma ejecuta la plantilla correspondiente al tipo de análisis en cuestión que se encuentra en el paquete `Methods`. Estas plantillas utilizan una biblioteca implementada en R llamada `Zelig`.

### 2.6- Diagramas de Secuencia

El diagrama de secuencia se encuentra entre los diagramas de interacción, ellos modelan el comportamiento dinámico de un sistema, describen la interacción entre objetos. El diagrama de secuencia en específico muestra la interacción de un conjunto de objetos en el tiempo. A continuación se muestran dos diagramas de secuencia, uno por cada caso de uso. Los casos de uso por independientes tienen varias secciones, el de Realizar Análisis Estadístico cuenta de 28 secciones y el de Graficar Análisis Estadístico cuenta con 9 secciones, todos siguen el mismo flujo, donde únicamente cambia es en la funcionalidad que se va a ejecutar, por lo que es distinta la plantilla a renderizar. Por estas razones no se ve necesario plasmar todos estos diagramas, ni siquiera en los anexos.

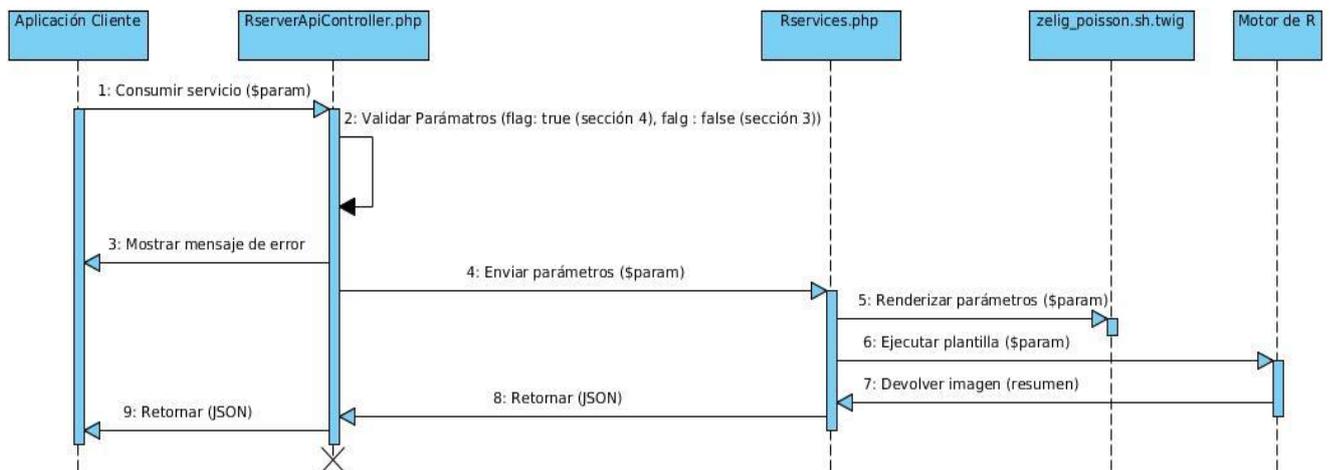


Figura 4. Diagrama de Secuencia del Caso de Uso: Realizar Análisis Estadístico, Sección Realizar Análisis Estadístico de Datos por el método de Poisson

### Descripción del diagrama de Secuencia del Caso de Uso: Realizar Análisis Estadístico de Datos

La Aplicación Cliente envía una petición al RServer 2.0 pasándole un conjunto de parámetros, la clase RserverApiController.php valida que esos datos entrados estén correctamente estructurados, de no ser así, muestra un mensaje de error, de lo contrario continúa ejecutándose correctamente. Dicha clase envía esos parámetros ya evaluados y codificados a la clase Rservices.php, este, los decodifica y le pasa los parámetros a la plantilla predeterminada para el tipo de análisis en cuestión, en el caso de este diagrama es la Plantilla de Zelig\_Poisson. El Rservices.php ejecuta la plantilla. R realiza el cálculo predeterminado y devuelve un resumen al Rservices, esta clase, vuelve a codificarlo y lo retorna al RserverApiController.php, esta, vuelve a decodificarlo y finalmente se lo retorna al cliente en un formato de JSON.

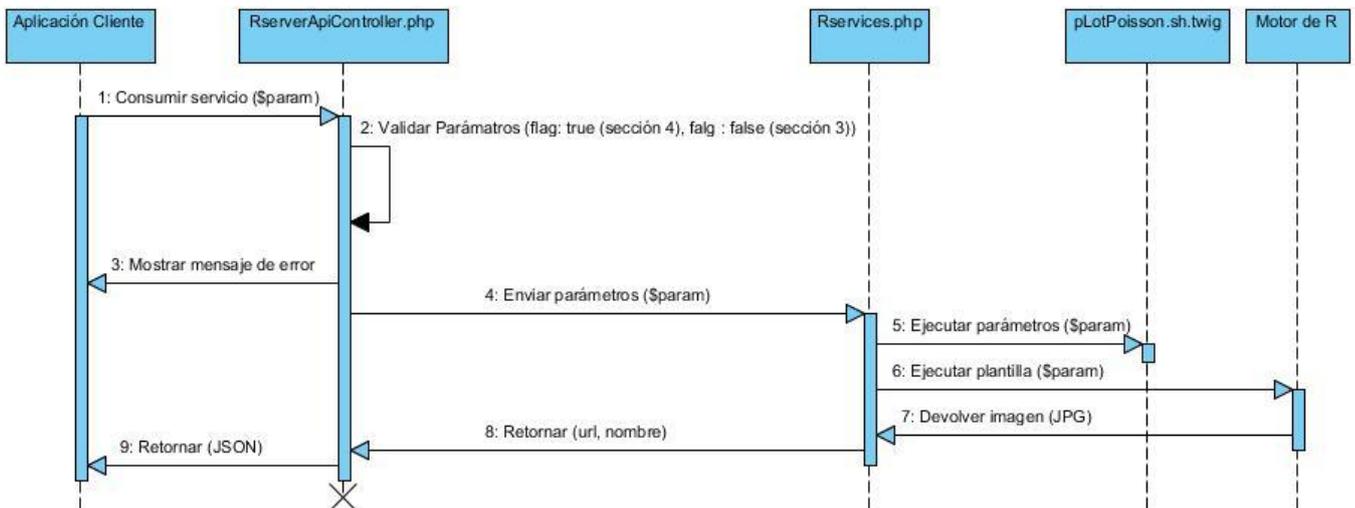


Figura 5. Diagrama de Secuencia del Caso de Uso: Graficar Análisis Estadístico, Sección Graficar Análisis Estadístico por el método de Poisson

### Descripción del diagrama de Secuencia del Caso de Uso: Graficar Análisis Estadístico

La Aplicación Cliente envía una petición al RServer 2.0 pasándole un conjunto de parámetros, la clase RserverApiController.php valida que esos datos entrados estén correctamente estructurados, de no ser así muestra un mensaje de error, de lo contrario continúa ejecutándose correctamente. Dicha clase envía esos parámetros ya evaluados y codificados a la clase Rservices.php, este a su vez los decodifica y envía los parámetros a la plantilla predeterminada para el tipo de análisis en cuestión, en este caso la Plantilla de Zelig\_Poisson. El Rservices.php ejecuta la plantilla. R realiza el cálculo predeterminado y devuelve una imagen al Rservices, este, vuelve a codificarlo y lo retorna al RserverApiController.php, esta clase finalmente retorna el nombre de la imagen y la dirección de donde está ubicada en el servidor del RServer 2.0, todo este envío lo hace en formato JSON.

## 2.7- Propuesta de Arquitectura

Para el RServer 2.0 se propone una arquitectura aparte de la que Symfony 2.0 trae constituida. La arquitectura propuesta cumple con buenas prácticas de programación; puesto que utiliza:

- La **inyección de dependencias** que brinda el marco de trabajo Symfony 2.0, esta hace más independiente al módulo y ahorra espacio en memoria virtual, pues ayuda a no tener que volver a instanciar una clase creada, gracias a una configuración previa.
- También brinda una excelente manera de agregar nuevos métodos de análisis estadísticos, mediante el **empleo de plantillas .twig**<sup>12</sup>. Se escribe el código correspondiente a un nuevo análisis estadístico del lenguaje R en una nueva plantilla.

Se puede describir de la siguiente forma: una aplicación cliente enviará parámetros para ser ejecutados por el RServer 2.0 mediante el protocolo SOAP. Este a su vez renderiza esos parámetros enviados en una plantilla previamente confeccionada. Esta plantilla se ejecuta en el motor de R, devolviendo un resultado. Todo queda estructurado de modo tal que para agregar nuevas funcionalidades y métodos de análisis estadísticos sólo se tendría que insertar una nueva plantilla .twig con el código a ejecutar en R. En la clase Rservices.php se escribe una nueva funcionalidad en la cual se ejecuta dicha plantilla, se debe corresponder el nombre del método con el análisis. Dentro de la clase RserverApiController.php se escribe el código correspondiente para publicar la nueva funcionalidad como un servicio web.

En el diagrama de componentes queda reflejada esta estructura.

## Conclusiones del Capítulo

Se dio cumplimiento a los objetivos trazados en este capítulo. Pues se confeccionó el modelo de dominio. También se hizo un estudio de los requisitos funcionales y no funcionales del sistema, quedando, los requisitos funcionales, en dos casos de uso; que dio lugar a la elaboración del diagrama de caso de uso del sistema. Además se confeccionaron los diagramas de clases del diseño y los diagramas de secuencia. Se dio una breve explicación de cada uno de los modelos para obtener un mejor entendimiento. Todos estos elementos dan paso al capítulo de implementación y pruebas. Se definieron un conjunto de pautas arquitectónicas a seguir para la integración de nuevas funcionalidades y módulos del RServer 2.0.

---

<sup>12</sup> **.twig**: es un lenguaje de plantillas muy potente llamado Twig. Permite escribir plantillas concisas y fáciles de leer que son más amigables para los diseñadores web y de varias maneras.

## CAPÍTULO 3: IMPLEMENTACIÓN Y PRUEBA

En este capítulo se tienen como objetivos, implementar lo diseñado en el capítulo anterior y probar que los requisitos especificados fueron implementados correctamente. Partiendo del resultado del diseño del sistema, en el presente capítulo se generan los artefactos correspondientes a esta fase. Se modela un diagrama de componentes, quedando reflejada la arquitectura del sistema y se modela el diagrama donde se desplegará el RServer 2.0. También se especifican las pruebas realizadas a la aplicación, para validar que el software. Se muestran algunos ejemplos de código, mostrando cómo se programaron los métodos de análisis estadísticos implementados en la aplicación.

### 3.1- Diagrama de Componentes

Un diagrama de componentes representa cómo un sistema de software es dividido en componentes, mostrando las dependencias que existen entre ellos. Los componentes físicos incluyen: archivos, cabeceras, bibliotecas compartidas, módulos, ejecutables y paquetes.

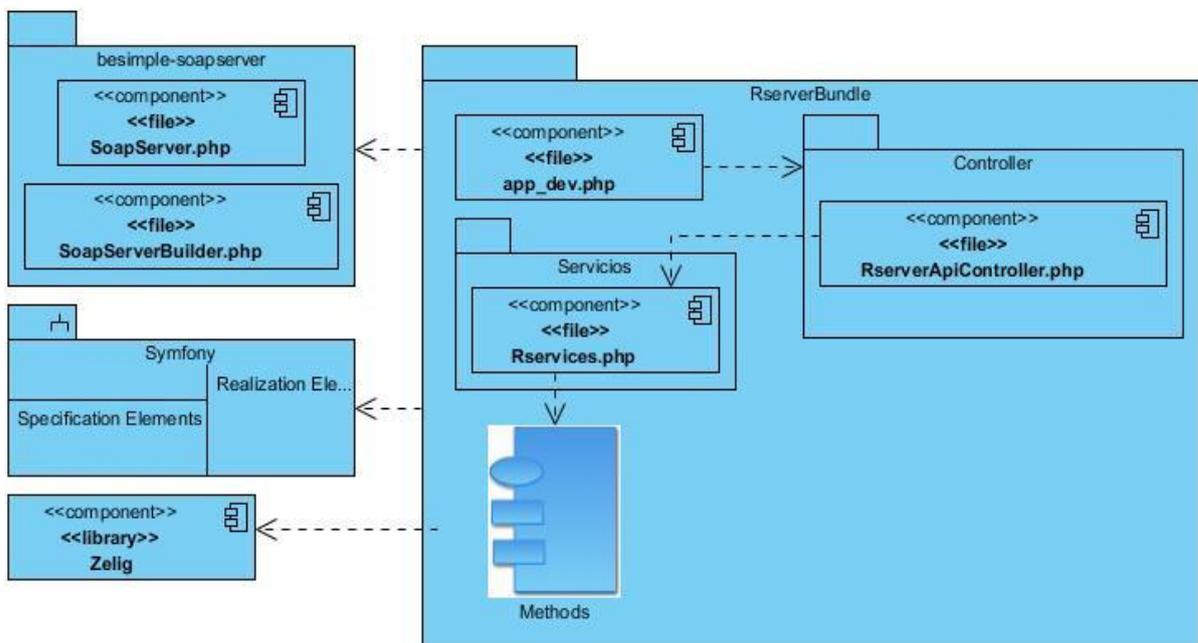


Figura 6. Diagrama de Componentes de la Aplicación

### Descripción del diagrama de Componentes

En este diagrama se representa como queda estructurada la aplicación arquitectónicamente. Dentro de RserverBundle (módulo de análisis estadístico) se encuentran el controlador frontal, que da inicio al

paquete de controladores. Dentro de este paquete se encuentra la clases MethodController.php (clase que se utiliza para probar las funcionalidades de la aplicación) y RserverApiController.php, que es la clase que realiza el análisis estadístico. Ambas van a la clase Rservices.php. Pues esta es la que vincula al lenguaje R con PHP. Rservices.php ejecuta las plantillas que se encuentran en el subsistema de componentes Methods. Estas plantillas, usan la biblioteca Zelig del motor de R. RserverBundle utiliza el paquete besimple-soapserver para el envío de JSON a la aplicación cliente.

### 3.2- Diagrama de Despliegue

Este diagrama presenta una vista actualizada del diagrama de despliegue, muestra los principales artefactos y componentes de la aplicación, distribuidos en sus correspondientes nodos, así como las dependencias que existen entre cada uno de ellos.



Figura 7. Diagrama de Despliegue de la Aplicación

#### Descripción del diagrama de Despliegue

El nodo PC Cliente representa donde va a estar la aplicación cliente que va a consumir el servicio que brindará el RServer 2.0, mediante el protocolo SOAP. Puede ser cualquier aplicación que desee realizar un análisis o graficar un análisis estadístico de los datos. En el nodo RServer 2.0 es donde estará desplegada la aplicación quedando allí instalados todos los motores de cálculo y las aplicaciones y servidores necesarios para el correcto funcionamiento del mismo.

### 3.3- Implementaciones Significativas

Siguiendo las pautas arquitectónicas propuestas en el capítulo anterior se muestran implementaciones hechas al módulo de análisis estadístico de datos. A continuación se muestra primeramente el flujo del cálculo de la densidad de Poisson hasta llegar a un resultado y después se muestra el flujo del cálculo de la Regresión de Poisson para variables de recuento de eventos dependientes. Este último, es un análisis que utiliza la biblioteca Zelig que trae implementada R.

```

/**
 * @Soap\Method("poisson_Densidad")
 * @Soap\Param("params", phpType = "string")
 * @Soap\Result(phpType = "string")
 */
public function poisson_DensidadAction($params) {
    $rservices = $this->get('rserver.stats');
    $result = $rservices->poisson_DensidadAction($params);
    return new SoapResponse($result);
}

```

Figura 8. Método que devuelve la densidad de Poisson en formato JSON

En esta figura se muestra un segmento de código de cómo se calcula el análisis estadístico de la densidad de Poisson. En la parte superior se encuentra un comentario por el cual el WSDL puede acceder al método. El método recibe un parámetro de tipo JSON. Una vez dentro, se realiza una inyección de dependencias (figura 11), guardándose en la variable \$rservices. Usando esta variable se puede acceder a todas las funcionalidades de la clase services.php. Posteriormente la variable \$result guardará el resultado de lo que devuelva la función de poisson\_DensidadAction(\$params) (Figura9). Por último se devuelve un objeto JSON, en una respuesta SOAP.

```

public function poisson_DensidadAction($json_params) {
    $params = json_decode($json_params, true);
    $result = array();

    try {
        $comando = $this->plantilla->render('RserverBundle:Methods:poisson_Densidad.sh.twig', $params);
        $data = exec($comando, $out);
    } catch (\Exception $exc) {
        $data = array('success' => false, 'error_code' => $exc->getCode(), 'mensaje' => $exc->getMessage());
    }
    $result['success'] = true;
    $result['data'] = $data;
    return json_encode($result);
}

```

Figura 9. Método que calcula la densidad de Poisson

Esta funcionalidad recibe un objeto JSON, que es decodificado y guardado en la variable \$params. Se inicializa la variable \$result como un arreglo, donde se guardará la respuesta. Luego en un bloque try catch para capturar posibles errores, son encerradas las líneas, dentro de éstas se guardan en una variable la ejecución de los parámetros en una plantilla definida para ese método (Figura 10). Y en la

otra se ejecuta dicho comando. De haber algún error es capturado y luego devuelto como respuesta. De ocurrir una ejecución correcta, se devuelve el resultado codificado como un JSON.

```
#!/bin/sh
R --slave << EOF
dpois(c({{x}}), {{lambda}}, log = {{log}})
EOF
```

Figura 10. Cálculo de la densidad de Poisson en el lenguaje R

Este es el cálculo en el lenguaje R de la densidad de Poisson, que está implementado predeterminadamente en una plantilla llamada poisson\_Densidad.sh.twig. Las dos primeras líneas de código son las que levantan el motor de R. Y la última es el cierre. La tercera línea es la función de la densidad de poisson (dpois), donde x es un vector de número entero no negativo, lambda es un vector de media no negativo y la variable log es una variable lógica.

```
services:
  rserver.stats:
    class: Datec\RserverBundle\Servicios\Rservices
    arguments: ['@templating']
```

Figura 11. Configuración de la Inyección de Dependencias

En este fragmento de código es donde se configura la inyección de dependencias.

A continuación se muestra otros ejemplos del flujo que se va generando, pero con otro tipo de análisis más complejo que el anterior.

```

/**
 * @Soap\Method("Regresion_Poisson")
 * @Soap\Param("params", phpType = "string")
 * @Soap\Result(phpType = "string")
 * @Route("/ww")
 */
public function Regresion_PoissonAction($params) {
    $rservices = $this->get('rserver.stats');
    $result = $rservices->Regresion_PoissonAction($params);

    return new SoapResponse($result);
}

```

Figura 12. Método que devuelve la Regresión para variables de Recuento Poisson en formato JSON

En esta figura se muestra un segmento de código de cómo se calcula el análisis estadístico de la regresión de Poisson para variables de recuento. En la parte superior se encuentra un comentario por el cual el WSDL puede acceder al método. El método recibe un parámetro de tipo JSON. Una vez dentro, se realiza una inyección de dependencias (figura 11), guardándose en la variable \$rservices. Usando esta variable se puede acceder a todas las funcionalidades de la clase services.php. Posteriormente la variable \$result guardará el resultado de lo que devuelva la función de Regresion\_PoissonAction (\$params) (Figura13). Por último se devuelve un objeto JSON, en una respuesta SOAP.

```

public function Regresion_PoissonAction($json_params) {
    $params = json_decode($json_params, true);
    $result = array();

    try {
        $comando = $this->plantilla->render('RserverBundle:Methods:zelig_poisson.sh.twig', $params);
        $data = exec($comando, $out);
    } catch (\Exception $exc) {
        $data = array('success' => false, 'error_code' => $exc->getCode(), 'mensaje' => $exc->getMessage());
    }
    $salida = array_slice($out, 25);
    $result['success'] = true;
    $result['data'] = $salida;
    return json_encode($result);
}

```

Figura 13. Método que calcula la Regresión de Poisson para variables de recuento

Esta funcionalidad recibe un objeto JSON, que es decodificado y guardado en la variable \$params. Se inicializa la variable \$result como un arreglo, donde se guardará la respuesta. Luego en un bloque try

catch para capturar posibles errores, son encerradas las líneas, dentro de éstas se guardan en una variable la ejecución de los parámetros en una plantilla definida para ese método (Figura 14). Y en la otra se ejecuta dicho comando. De haber algún error es capturado y luego devuelto como respuesta. De ocurrir una ejecución correcta se devuelve el resultado codificado como un JSON.

```
#!/bin/sh
R --slave << EOF

library(Zelig)
#arregloDatos
#{%for dato in arregloDatos%}{% if(loop.last)>{{dato}}{%else%}{{dato}}{%endif%}{%endfor%} #}

dataset <- data.frame(matrix(c({{arregloDatos}}), {{nf}}, {{nc}}, byrow = {{byrow}}))
colnames(dataset)<-c({%for col in arregloNombresCol%}{%if(loop.last)%} "{{col}}"{%else%} "{{col}}",{%endif%}{%endfor%}

result <- zelig( {{varIndep}} ~ {{arregloVarDep}} , model = "poisson" , dataset)
summary(result)

EOF
```

Figura 14. Cálculo de la Regresión de Poisson para variables de Recuento en el lenguaje R

Este es el cálculo en el lenguaje R de la Regresión de Poisson para variables de recuento, que está predeterminado en una plantilla llamada Zelig\_poisson.sh.twig. Las dos primeras líneas de código son las que arrancan el motor de R. Y la última es el cierre. En la tercera línea se inicia los análisis de la biblioteca Zelig. Seguido de esto, la variable dataset guarda un dataFrame, que está compuesto por una matriz que contiene los datos a los cuales se les realizará el análisis en cuestión, por la variable nf (número de filas), la variable nc (número de columnas) y por la variable byrow, que determina como se organizará la matriz, por filas o por columnas. La variable result guardará el resultado, está compuesta por las variables independientes, por un arreglo de variables dependientes, el tipo de modelo (en este caso poisson) y por el dataset anteriormente confeccionado. Seguido de este paso se encuentra el código summary (result), lo que enviará será un resumen del cálculo de este análisis.

### 3.4- Pruebas al Sistema

Probar es el proceso ejecución de un programa con el fin de descubrir errores, de validar, encontrar defecto o fallo. Una prueba tiene éxito si descubre un error no detectado hasta entonces. Existen las pruebas de caja blanca y las de caja negra. Las pruebas de caja blanca garantizan que se ejercite por lo menos una vez todos los caminos independientes de cada módulo y todas las decisiones lógicas en sus vertientes verdadera y falsa. Que se ejecuten todos los bucles en sus límites y con sus límites

operacionales. Y que se prueben las estructuras internas de datos para asegurar su validez. Por otra parte las pruebas de caja negra intentan encontrar errores en funciones incorrectas o ausentes, errores de interfaz, errores en estructuras de datos o acceso a bases de datos externas, errores de rendimiento y errores de inicialización y terminación. (27)

### Pruebas de Caja Blanca

El método utilizado para hacerle pruebas de caja blanca es el método *Regresion\_PoissonAction*, anteriormente explicado. Primeramente se hacen separaciones del código en bloques, para posteriormente realizarle las pruebas que se requieren.

<pre>public function Regresion_PoissonAction(\$json_params) {     \$params=json_decode(\$json_params, true);     \$result=array();     try {         \$comando=\$this-&gt;plantilla-&gt;render('RserverBundle:Methods:zelig_poisson.sh.twig',         \$params);         \$data=exec(\$comando, \$out);     }     catch (\Exception \$exc)     {         \$data=array('success'=&gt;false,'error_code'=&gt;\$exc-&gt;getCode(),'mensaje'=&gt;\$exc-         &gt;getMessage());     } }</pre>	1
<pre>\$salida=array_slice(\$out, 25); \$result['success']=true; \$result['data']=\$salida;</pre>	2
<pre>return json_encode(\$result); }</pre>	3

Quedando conformado el grafo de la siguiente manera:



Entre las variantes de las pruebas de caja blanca están la prueba del camino básico, prueba de condición, prueba de flujo de datos y prueba de bucles. En este caso se hará la prueba del camino básico.

#### Prueba del camino básico:

Complejidad Ciclomática (La complejidad de McCabe  $V(G)$ ). Es un indicador del número de caminos independientes que existen en un grafo.

$$V(G) = a - n + 2$$

$$V(G) = r$$

$$V(G) = c + 1$$

Donde:

$a$  - número de arcos o aristas del grafo.

$n$  - número de nodos.

$r$  - número de regiones cerradas del grafo.

$c$  - número de nodos de condición.

$V(G)$  - marca el límite mínimo de casos de prueba para un programa. Cuando  $V(G) > 10$  la probabilidad de defectos en el módulo o programa crece mucho entonces quizás sea interesante dividir el módulo.

Quedando de la siguiente forma:

$$a = 2, \quad n = 3, \quad r = 1, \quad c = 0$$

$$V(G) = 2 - 3 + 2 = \mathbf{1}$$

$$V(G) = \mathbf{1}$$

$$V(G) = 0 + 1 = \mathbf{1}$$

Resultado de esta prueba:  $V(G)$  no es mayor que 10, por lo que la probabilidad de defectos en el módulo no crece mucho, entonces no es necesario dividir dicho módulo.

### Pruebas de Caja Negra

Entre las variantes de las pruebas de caja negra están métodos de prueba basados en grafos, partición equivalente (Casos de Pruebas), análisis de valores límite, prueba de comparación, conjetura de errores.

Los casos de prueba son actividades en las cuales un sistema o componente es ejecutado bajo condiciones o requerimientos especificados, los resultados son observados y registrados, haciendo una evaluación de algún aspecto o componente del sistema. Un buen caso de prueba es aquel que tiene una alta probabilidad de mostrar un error no descubierto hasta entonces. Al generar casos de prueba, se deben incluir tanto datos de entrada válidos y esperados como no válidos e inesperados.

(28)

A continuación se muestran casos de pruebas de la prueba partición equivalente con algunas de la secciones del caso de uso Realizar Análisis Estadístico de Datos. Ver más en Anexos.

**Descripción general**

Realiza análisis estadístico de los datos enviados según el tipo de análisis que se desea calcular. Culmina cuando se tenga el resultado esperado.

**Condición de ejecución**

Que todos los valores entrados para el cálculo, estén bien estructurados.

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 1.1 Calcular Densidad de Poisson de forma correcta.	Se realiza el cálculo correctamente y se devuelve a la aplicación cliente un resumen del cálculo de la Densidad de Poisson.	V (Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	Se envía el resultado de dicho cálculo a la aplicación cliente.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Realizar análisis estadístico. 4- Enviar resultado obtenido por el sistema.
EC 1.2 Calcular Densidad de Poisson de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Mostrar un mensaje de error.

*Tabla 5. Realizar Análisis Estadístico de Datos sección Densidad de Poisson.*

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 2.1 Generar los números aleatorios de Poisson de forma correcta.	Se realiza el cálculo correctamente y se devuelve a la aplicación cliente un resumen con los números aleatorios de Poisson generados.	V (Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	Se envía el resultado de dicho cálculo a la aplicación cliente.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Realizar análisis estadístico. 4- -Enviar resultado obtenido por el sistema.
EC 2.2 Generar los números aleatorios de	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente)	Se envía un mensaje de error.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren

Poisson de forma incorrecta.		correspondiendo con el tipo de análisis en cuestión o faltan datos)		correctamente estructurados. 3- Mostrar un mensaje de error.
------------------------------	--	---	--	---

Tabla 6. Realizar Análisis Estadístico de Datos sección Generar números aleatorios de Poisson.

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 3.1 Calcular la distribución acumulada de Poisson de forma correcta.	Se realiza el cálculo correctamente y se devuelve a la aplicación cliente un resumen con la distribución acumulada de Poisson.	V (Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	Se envía el resultado de dicho cálculo a la aplicación cliente.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Realizar análisis estadístico. 4- Enviar resultado obtenido por el sistema.
EC 3.2 Calcular la distribución acumulada de Poisson de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Mostrar un mensaje de error.

Tabla 7. Realizar Análisis Estadístico de Datos sección distribución acumulada de Poisson.

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 4.1 Calcular la distribución cuantiles de Poisson de forma correcta.	Se realiza el cálculo correctamente y se devuelve a la aplicación cliente un resumen con la distribución cuantiles de Poisson.	V (Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	Se envía el resultado de dicho cálculo a la aplicación cliente.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Realizar análisis estadístico. 4- Enviar resultado obtenido por el sistema.
EC 4.2 Calcular la distribución de cuantiles de Poisson de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Mostrar un mensaje de error.

Tabla 8. Realizar Análisis Estadístico de Datos sección distribución de cuantiles de Poisson.

Al sistema se le hicieron otras dos pruebas de Caja Negra durante la implementación: una se hizo para comprobar el correcto funcionamiento de los métodos de análisis estadísticos; cuando la aplicación se encontraba en un 70% de desarrollo y la otra se realizó para comprobar el WSDL, esta por su parte se comprobó cuando el módulo estaba implementado al 100%.

- 1- Comprobar el correcto funcionamiento de los métodos de análisis estadísticos: en esta prueba se confeccionó una clase llamada MethodController.php, la misma trae implementada todas las funcionalidades correspondientes a los métodos de análisis estadísticos con datos bien estructurados, para simular el proceso del módulo.

```
/**
 * @Route("/zelig_poisson", name="zelig_poisson")
 * @Template()
 */
public function Regresion_PoissonAction() {
    $params = array(
        'arregloDatos' => " 20, 65, 174, 22, 70, 180 ",
        'nf' => 3,
        'nc' => 2,
        'arregloNombresCol' => array("name", "edad"),
        'varIndep' => "name",
        'arregloVarDep' => "edad",
        'byrow' => 'FALSE'
    );
    $json_params = json_encode($params);

    $rserver = $this->get('rserver.stats');
    $result = $rserver->Regresion_PoissonAction($json_params);

    return $result;
}
```

Figura 15. Ejemplo de código, con datos entrados

- 2- Comprobar el WSDL: se desarrolló un cliente en java para probar el consumo de los servicios web.

```
.import java.rmi.RemoteException;

public class Regresion Poisson {
    @SuppressWarnings("unchecked")
    public static void main(String[] args) {
        RserverPortTypeProxy ws = new RserverPortTypeProxy();

        JSONObject obj = new JSONObject();
        obj.put("arregloDatos", "20, 65, 174, 22, 70, 180");
        obj.put("nf", new Integer(3));
        obj.put("nc", new Integer(2));

        JSONArray list = new JSONArray();
        list.add("name");
        list.add("edad");
        obj.put("arregloNombresCol", list);
        obj.put("varIndep", "name");
        obj.put("arregloVarDep", "edad");
        obj.put("byrow", "FALSE");

        try {
            String resp = ws.regresion_Poisson(obj.toString());
            System.out.print(resp);
        } catch (RemoteException e) {
            e.printStackTrace();
        }
    }
}
```

Figura 16. Ejemplo de código, en JAVA

Estas dos pruebas arrojaron resultados satisfactorios.

### Conclusiones del Capítulo

Se dio cumplimiento a los objetivos trazados en el inicio del capítulo. Se confeccionaron los diferentes modelos correspondientes a la fase en cuestión; el diagrama de componentes y el diagrama de despliegue del sistema. También se realizaron pruebas de caja negra y de caja blanca a la aplicación. Además de mostrar algunos ejemplos del código implementado que ayudan a comprender el flujo de cómo se realiza un cálculo de análisis estadístico internamente.

## CONCLUSIONES

La estadística se hace cada vez más imprescindible tanto en las ramas de las ciencias como en la vida cotidiana. Gracias a ella se han conseguido innumerables logros científicos y tecnológicos, que han llevado a pretender que la humanidad viva de una forma más civilizada y organizada. Como resultado de la investigación para realizar el módulo de análisis estadístico de datos para el RServer 2.0 se concluye que:

- Se efectuó un estudio del marco conceptual referente a los sistemas que realizan análisis estadísticos de datos para dar solución al problema planteado.
- Se escogieron las herramientas y tecnologías a utilizar para el desarrollo de la aplicación.
- Se propuso e implementó la arquitectura que quedará como plataforma base para posteriores integraciones de módulos al RServer 2.0.
- Se implementó el módulo de análisis estadísticos de datos con 28 análisis, donde 9 de ellos permiten ser graficados.
- Se probó el módulo de análisis estadísticos de datos.

## RECOMENDACIONES

- Usar la arquitectura propuesta para la incorporación de nuevos módulos y nuevos métodos estadísticos en el Rserver 2.0.

## REFERENCIAS BIBLIOGRAFICAS

1. [Online] [Cited: 10 02, 2011.] [http://www.estadisticaparatodos.es/software/software\\_libre.html](http://www.estadisticaparatodos.es/software/software_libre.html).
2. Otros software estadísticos. *Estadística para todos*. [Online] [Cited: 10 11, 2011.] [http://www.estadisticaparatodos.es/software/software\\_otros.html](http://www.estadisticaparatodos.es/software/software_otros.html).
3. [Online] [Cited: 10 04, 2011.] <http://statgraphics.programas-gratis.net/>.
4. **Hernández, Leyser Hernández**. *Desarrollo de algoritmos de Pruebas no Paramétricas para el servidor de análisis estadístico: Rserver*. Ciudad de La Habana : s.n., 2010.
5. [Online] <http://php.uci.cu>.
6. [Online] [Cited: 11 12, 2011.] <http://www.agapea.com/libros/PHP-5-isbn-8441517851-i.htm>.
7. **Jordana, Garcilaso**. *Introducción Open UP*.
8. Visual Paradigm International Ltd. [Online] 05 2007. [Cited: 11 20, 2011.] [http://www.freedownloadmanager.org/es/downloads/Paradigma\\_Visual\\_para\\_UML\\_\(M%C3%8D\)\\_14720\\_p/...](http://www.freedownloadmanager.org/es/downloads/Paradigma_Visual_para_UML_(M%C3%8D)_14720_p/...)
9. [Online] <http://acsblog.es/articulos/trunk/LinuxActual/Apache/html/x31.html>.
10. [Online] <http://www.desarrolloweb.com/faq/360.php>.
11. Ayuda de R.
12. **Hamza, K**. The smallest uniform upper bound on the distance between the mean and the median of the binomial and Poisson distributions. s.l. : Lett. 23 21-25., 1995.
13. [Online] [http://www.vitutor.com/pro/3/b\\_2.html](http://www.vitutor.com/pro/3/b_2.html).
14. [Online] [http://www.itch.edu.mx/academic/industrial/sabaticorita/\\_private/05Distr%20Poisson.htm](http://www.itch.edu.mx/academic/industrial/sabaticorita/_private/05Distr%20Poisson.htm).
15. **Myers, Roland Walpole and Raymond y Ye. Keying**. Probability and Statistics for Engineers and Scientists Pearson Education. 2002.
16. **Weisstein, Eric W**. Chi-Squared Test.
17. [Online] [Cited: 02 15, 2012.] [http://www.vitutor.com/estadistica/descriptiva/a\\_15.html](http://www.vitutor.com/estadistica/descriptiva/a_15.html).
18. **Molinero, Luis M**. Análisis de Series Temporales. . *Asociación de la Sociedad Española de Hipertensión Liga Española para la lucha contra la Hipertensión Arterial*. [Online] 2004. [Cited: 01 15, 2012.] <http://www.seh-lelha.org/tseries.htm>.
19. [Online] [Cited: 01 16, 2012.] <http://www.uv.es/ceaces/base/regresion/exponenci.htm>.
20. **María de Mar Zamora Sanz y Julio Estavillo Dorado**. Modelo de regresión normal clásico. . [Online] 2001. [Cited: 02 10, 2012.] [http://ciberconta.unizar.es/leccion/ecoreg/..](http://ciberconta.unizar.es/leccion/ecoreg/)
21. [Online] [Cited: 02 2012, 16.] <http://www.seh-lelha.org/residuosCox.htm>.
22. **Bolstad, William M., John Wiley**. Introduction to Bayesian Statistics. 2004.
23. [Online] [Cited: 05 21, 2012.] [http://www.ecured.cu/index.php/Modelo\\_de\\_dominio](http://www.ecured.cu/index.php/Modelo_de_dominio).
24. **Mühlrad, Daniel**. *Patrones de diseño*.
25. Applications Programming in Smalltalk-80(TM): How to use Model-View-Controller (MVC). [Online] <http://st-www.cs.illinois.edu/users/smarch/st-docs/mvc.html>.
26. Genbeta: dev. *Páginas de Programación*. [Online] <http://www.genbetadev.com/paradigmas-de-programacion/que-es-la-inyeccion-de-dependencias>.
27. Entorno Virtual de Aprendizaje. [Online] <http://eva.uci.cu>.
28. [Online] [Cited: 05 12, 2012.] [http://www.ecured.cu/index.php/Prueba\\_de\\_RUP](http://www.ecured.cu/index.php/Prueba_de_RUP).

## BIBLIOGRAFIA

1. [Online] [Cited: 10 02, 2011.] [http://www.estadisticaparatodos.es/software/software\\_libre.html](http://www.estadisticaparatodos.es/software/software_libre.html).
2. Otros software estadísticos. *Estadística para todos*. [Online] [Cited: 10 11, 2011.] [http://www.estadisticaparatodos.es/software/software\\_otros.html](http://www.estadisticaparatodos.es/software/software_otros.html).
3. [Online] [Cited: 10 04, 2011.] <http://statgraphics.programas-gratis.net/>.
4. **Hernández, Leyser Hernández**. *Desarrollo de algoritmos de Pruebas no Paramétricas para el servidor de análisis estadístico: Rserver*. Ciudad de La Habana : s.n., 2010.
5. [Online] <http://php.uci.cu>.
6. [Online] [Cited: 11 12, 2011.] <http://www.agapea.com/libros/PHP-5-isbn-8441517851-i.htm>.
7. **Jordana, Garcilaso**. *Introducción Open UP*.
8. Visual Paradigm International Ltd. [Online] 05 2007. [Cited: 11 20, 2011.] [http://www.freedownloadmanager.org/es/downloads/Paradigma\\_Visual\\_para\\_UML\\_\(M%C3%8D\)\\_14720\\_p/...](http://www.freedownloadmanager.org/es/downloads/Paradigma_Visual_para_UML_(M%C3%8D)_14720_p/...)
9. [Online] <http://acsblog.es/articulos/trunk/LinuxActual/Apache/html/x31.html>.
10. [Online] <http://www.desarrolloweb.com/faq/360.php>.
11. Ayuda de R.
12. **Hamza, K**. The smallest uniform upper bound on the distance between the mean and the median of the binomial and Poisson distributions. s.l. : Lett. 23 21-25., 1995.
13. [Online] [http://www.vitutor.com/pro/3/b\\_2.html](http://www.vitutor.com/pro/3/b_2.html).
14. [Online] [http://www.itch.edu.mx/academic/industrial/sabaticorita/\\_private/05Distr%20Poisson.htm](http://www.itch.edu.mx/academic/industrial/sabaticorita/_private/05Distr%20Poisson.htm).
15. **Myers, Roland Walpole and Raymond y Ye. Keying**. Probability and Statistics for Engineers and Scientists Pearson Education. 2002.
16. **Weisstein, Eric W**. Chi-Squared Test.
17. [Online] [Cited: 02 15, 2012.] [http://www.vitutor.com/estadistica/descriptiva/a\\_15.html](http://www.vitutor.com/estadistica/descriptiva/a_15.html).
18. **Molinero, Luis M**. Análisis de Series Temporales. . *Asociación de la Sociedad Española de Hipertensión Liga Española para la lucha contra la Hipertensión Arterial*. [Online] 2004. [Cited: 01 15, 2012.] <http://www.seh-lelha.org/tseries.htm>.
19. [Online] [Cited: 01 16, 2012.] <http://www.uv.es/ceaces/base/regresion/exponenci.htm>.
20. **María de Mar Zamora Sanz y Julio Estavillo Dorado**. Modelo de regresión normal clásico. . [Online] 2001. [Cited: 02 10, 2012.] [http://ciberconta.unizar.es/leccion/ecoreg/..](http://ciberconta.unizar.es/leccion/ecoreg/)
21. [Online] [Cited: 02 2012, 16.] <http://www.seh-lelha.org/residuosCox.htm>.
22. **Bolstad, William M., John Wiley**. Introduction to Bayesian Statistics. 2004.
23. [Online] [Cited: 05 21, 2012.] [http://www.ecured.cu/index.php/Modelo\\_de\\_dominio](http://www.ecured.cu/index.php/Modelo_de_dominio).
24. **Mühlrad, Daniel**. *Patrones de diseño*.
25. Applications Programming in Smalltalk-80(TM): How to use Model-View-Controller (MVC). [Online] <http://st-www.cs.illinois.edu/users/smarch/st-docs/mvc.html>.
26. Genbeta: dev. *Páginas de Programación*. [Online] <http://www.genbetadev.com/paradigmas-de-programacion/que-es-la-inyeccion-de-dependencias>.
27. Entorno Virtual de Aprendizaje. [Online] <http://eva.uci.cu>.
28. [Online] [Cited: 05 12, 2012.] [http://www.ecured.cu/index.php/Prueba\\_de\\_RUP](http://www.ecured.cu/index.php/Prueba_de_RUP).
29. [Online] <http://www.json.org/json-es.html>.
30. [Online]. <http://netbeans.org/community/releases/70/index.html>.
31. Hernández, Pedro Veloso. Uso de patrones de arquitectura.
32. Pressman, Roger S. Software Engeneering, a practitioner's approach.
33. Introducción al lenguaje R. [bookauth] Dra. Diana M. 2006.

34. Rumbaugh, Jim, Jacobson, Ivan y Booch, Grady. UML el lenguaje Unificado de modelado. [Online] <http://www.uml.org/>.

## ANEXOS

## Casos de Pruebas

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 5.1 Calcular Mediana de forma correcta.	Se realiza el cálculo correctamente y se devuelve a la aplicación cliente un resumen del cálculo de la Mediana.	V (Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	Se envía el resultado de dicho cálculo a la aplicación cliente.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Realizar análisis estadístico. 4- Enviar resultado obtenido por el sistema.
EC 5.2 Calcular Mediana de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Mostrar un mensaje de error.

*Tabla 9. Realizar Análisis Estadístico de Datos, sección Mediana.*

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 6.1 Calcular la distribución acumulada de T-Student de forma correcta.	Se realiza el cálculo correctamente y se devuelve a la aplicación cliente un resumen con la distribución acumulada de T-Student.	V (Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	Se envía el resultado de dicho cálculo a la aplicación cliente.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Realizar análisis estadístico. 4- Enviar resultado obtenido por el sistema.
EC 6.2 Calcular la distribución acumulada de T-Student de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Mostrar un mensaje de error.

*Tabla 10. Realizar Análisis Estadístico de Datos, sección de distribución de acumulada de T-Student.*

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 7.1 Calcular la distribución de cuantiles de T-Student de forma correcta.	Se realiza el cálculo correctamente y se devuelve a la aplicación cliente un resumen con la distribución de cuantiles de T-Student.	V (Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	Se envía el resultado de dicho cálculo a la aplicación cliente.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Realizar análisis estadístico. 4- Enviar resultado obtenido por el sistema.
EC 7.2 Calcular la distribución de cuantiles de T-Student de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Mostrar un mensaje de error.

Tabla 11. Realizar Análisis Estadístico de Datos, sección de distribución de cuantiles de T-Student.

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 8.1 Calcular la densidad de T-Student de forma correcta.	Se realiza el cálculo correctamente y se devuelve a la aplicación cliente un resumen con la densidad de T-Student.	V (Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	Se envía el resultado de dicho cálculo a la aplicación cliente.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Realizar análisis estadístico. 4- Enviar resultado obtenido por el sistema.
EC 8.2 Calcular la densidad de T-Student de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Mostrar un mensaje de error.

Tabla 12. Realizar Análisis Estadístico de Datos, sección de distribución de densidad de T-Student.

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 9.1 Generar los números aleatorios de	Se realiza el cálculo correctamente y se devuelve a la	V (Se envían los parámetros correctamente)	Se envía el resultado de dicho cálculo a la aplicación cliente.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren

T-Student de forma correcta.	aplicación cliente un resumen con los números aleatorios de T-Student generados.	correspondiendo con el tipo de análisis en cuestión)		correctamente estructurados. 3- Realizar análisis estadístico. 4- Enviar resultado obtenido por el sistema.
EC 9.2 Generar los números aleatorios de T-Student de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Mostrar un mensaje de error.

Tabla 13. Realizar Análisis Estadístico de Datos, sección generar números aleatorios de T-Student.

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 10.1 Calcular la distribución acumulada Binomial de forma correcta.	Se realiza el cálculo correctamente y se devuelve a la aplicación cliente un resumen de la distribución acumulada Binomial.	V (Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	Se envía el resultado de dicho cálculo a la aplicación cliente.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Realizar análisis estadístico. 4- Enviar resultado obtenido por el sistema.
EC 10.2 Calcular la distribución acumulada Binomial de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Mostrar un mensaje de error.

Tabla 14. Realizar Análisis Estadístico de Datos, sección distribución acumulada Binomial.

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 11.1 Calcular la distribución simple Binomial de forma correcta.	Se realiza el cálculo correctamente y se devuelve a la aplicación cliente un resumen de la distribución simple Binomial.	V (Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	Se envía el resultado de dicho cálculo a la aplicación cliente.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Realizar análisis estadístico. 4- Enviar resultado obtenido por el sistema.

EC 11.2 Calcular la distribución simple Binomial de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	<ol style="list-style-type: none"> <li>1- Recibir datos enviados a calcular.</li> <li>2- Validar datos enviados, verificar que se encuentren correctamente estructurados.</li> <li>3- Mostrar un mensaje de error.</li> </ol>
--	--	---	-------------------------------	---

*Tabla 15. Realizar Análisis Estadístico de Datos, sección distribución simple Binomial.*

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 13.1 Calcular la distribución acumulada de Chi-Cuadrado de forma correcta.	Se realiza el cálculo correctamente y se devuelve a la aplicación cliente un resumen de la densidad de Chi-Cuadrado.	V (Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	Se envía el resultado de dicho cálculo a la aplicación cliente.	<ol style="list-style-type: none"> <li>1- Recibir datos enviados a calcular.</li> <li>2- Validar datos enviados, verificar que se encuentren correctamente estructurados.</li> <li>3- Realizar análisis estadístico.</li> <li>4- Enviar resultado obtenido por el sistema.</li> </ol>
EC 13.2 Calcular la distribución acumulada de Chi-Cuadrado de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	<ol style="list-style-type: none"> <li>1- Recibir datos enviados a calcular.</li> <li>2- Validar datos enviados, verificar que se encuentren correctamente estructurados.</li> <li>3- Mostrar un mensaje de error.</li> </ol>

*Tabla 16. Realizar Análisis Estadístico de Datos, sección distribución acumulada Chi\_Cuadrado.*

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 14.1 Calcular la distribución de cuantiles de Chi-Cuadrado de forma correcta.	Se realiza el cálculo correctamente y se devuelve a la aplicación cliente un resumen de la distribución de cuantiles de Chi-Cuadrado.	V (Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	Se envía el resultado de dicho cálculo a la aplicación cliente.	<ol style="list-style-type: none"> <li>1- Recibir datos enviados a calcular.</li> <li>2- Validar datos enviados, verificar que se encuentren correctamente estructurados.</li> <li>3- Realizar análisis estadístico.</li> <li>4- Enviar resultado obtenido por el sistema.</li> </ol>
EC 14.2 Calcular la distribución de cuantiles de	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente)	Se envía un mensaje de error.	<ol style="list-style-type: none"> <li>1- Recibir datos enviados a calcular.</li> <li>2- Validar datos enviados, verificar que se encuentren</li> </ol>

Chi-Cuadrado de forma incorrecta.		correspondiendo con el tipo de análisis en cuestión o faltan datos)		correctamente estructurados. 3- Mostrar un mensaje de error.
-----------------------------------	--	---	--	---

Tabla 17. Realizar Análisis Estadístico de Datos, sección distribución de cuantiles de Chi\_Cuadrado.

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 15.1 Generar los números aleatorios de Chi-Cuadrado de forma correcta.	Se realiza el cálculo correctamente y se devuelve a la aplicación cliente un resumen de los números aleatorios de Chi-Cuadrado generados.	V (Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	Se envía el resultado de dicho cálculo a la aplicación cliente.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Realizar análisis estadístico. 4- Enviar resultado obtenido por el sistema.
EC 15.2 Generar los números aleatorios de Chi-Cuadrado de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Mostrar un mensaje de error.

Tabla 18. Realizar Análisis Estadístico de Datos, sección generar números aleatorios de Chi\_Cuadrado.

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 16.1 Calcular la Varianza de forma correcta.	Se realiza el cálculo correctamente y se devuelve a la aplicación cliente un resumen de la Varianza.	V (Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	Se envía el resultado de dicho cálculo a la aplicación cliente.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Realizar análisis estadístico. 4- Enviar resultado obtenido por el sistema.
EC 16.2 Calcular la Varianza de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Mostrar un mensaje de error.

Tabla 19. Realizar Análisis Estadístico de Datos, sección Varianza.

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 17.1 Calcular series temporales, estimando un modelo de forma correcta.	Se realiza el cálculo correctamente y se devuelve a la aplicación cliente un resumen de series temporales, estimando un modelo.	V (Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	Se envía el resultado de dicho cálculo a la aplicación cliente.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Realizar análisis estadístico. 4- Enviar resultado obtenido por el sistema.
EC 17.2 Calcular series temporales, estimando un modelo de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Mostrar un mensaje de error.

Tabla 20. Realizar Análisis Estadístico de Datos, sección Series Temporales, estimando un modelo.

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 18.1 Calcular series temporales resumiendo los resultados de forma correcta.	Se realiza el cálculo correctamente y se devuelve a la aplicación cliente un resumen series temporales resumiendo los resultados.	V (Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	Se envía el resultado de dicho cálculo a la aplicación cliente.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Realizar análisis estadístico. 4- Enviar resultado obtenido por el sistema.
EC 18.2 Calcular series temporales resumiendo los resultados de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Mostrar un mensaje de error.

Tabla 21. Realizar Análisis Estadístico de Datos, sección Series Temporales, resumiendo los resultados.

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
-----------	-------------	---------------------	-----------------------	---------------

EC 19.1 Calcular la regresión logística bivariada estimando el modelo (blogit) de forma correcta.	Se realiza el cálculo correctamente y se devuelve a la aplicación cliente un resumen de la regresión logística bivariada estimando el modelo (blogit).	V (Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	Se envía el resultado de dicho cálculo a la aplicación cliente.	<ol style="list-style-type: none"> <li>1- Recibir datos enviados a calcular.</li> <li>2- Validar datos enviados, verificar que se encuentren correctamente estructurados.</li> <li>3- Realizar análisis estadístico.</li> <li>4- Enviar resultado obtenido por el sistema.</li> </ol>
EC 19.2 Calcular la regresión logística bivariada estimando el modelo (blogit) de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	<ol style="list-style-type: none"> <li>1- Recibir datos enviados a calcular.</li> <li>2- Validar datos enviados, verificar que se encuentren correctamente estructurados.</li> <li>3- Mostrar un mensaje de error.</li> </ol>

*Tabla 22. Realizar Análisis Estadístico de Datos, sección Regresión Logística Bivariada, estimando modelo (Blogit).*

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 20.1 Calcular la regresión logística bivariada estimación conjunta (blogit) de forma correcta.	Se realiza el cálculo correctamente y se devuelve a la aplicación cliente un resumen de la regresión logística bivariada estimación conjunta (blogit).	V (Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	Se envía el resultado de dicho cálculo a la aplicación cliente.	<ol style="list-style-type: none"> <li>1- Recibir datos enviados a calcular.</li> <li>2- Validar datos enviados, verificar que se encuentren correctamente estructurados.</li> <li>3- Realizar análisis estadístico.</li> <li>4- Enviar resultado obtenido por el sistema.</li> </ol>
EC 20.2 Calcular la regresión logística bivariada estimación conjunta (blogit) de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	<ol style="list-style-type: none"> <li>1- Recibir datos enviados a calcular.</li> <li>2- Validar datos enviados, verificar que se encuentren correctamente estructurados.</li> <li>3- Mostrar un mensaje de error.</li> </ol>

*Tabla 23. Realizar Análisis Estadístico de Datos, sección Regresión Logística Bivariada, estimación conjunta (Blogit).*

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 21.1	Se realiza el	V	Se envía el	1- Recibir datos enviados a

Calcular la regresión logística bivariada estimando el modelo (probit) de forma correcta.	cálculo correctamente y se devuelve a la aplicación cliente un resumen de la regresión logística bivariada estimación conjunta (blogit).	(Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	resultado de dicho cálculo a la aplicación cliente.	<ul style="list-style-type: none"> <li>calcular.</li> <li>2- Validar datos enviados, verificar que se encuentren correctamente estructurados.</li> <li>3- Realizar análisis estadístico.</li> <li>4- Enviar resultado obtenido por el sistema.</li> </ul>
EC 21.2 Calcular la regresión logística bivariada estimando el modelo (probit) de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	<ul style="list-style-type: none"> <li>1- Recibir datos enviados a calcular.</li> <li>2- Validar datos enviados, verificar que se encuentren correctamente estructurados.</li> <li>3- Mostrar un mensaje de error.</li> </ul>

Tabla 24. Realizar Análisis Estadístico de Datos, sección Regresión Logística Bivariada, estimando modelo (Probit).

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 22.1 Calcular la regresión logística bivariada estimación conjunta (probit) de forma correcta.	Se realiza el cálculo correctamente y se devuelve a la aplicación cliente un resumen de la regresión logística bivariada estimación conjunta (probit).	V (Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	Se envía el resultado de dicho cálculo a la aplicación cliente.	<ul style="list-style-type: none"> <li>1- Recibir datos enviados a calcular.</li> <li>2- Validar datos enviados, verificar que se encuentren correctamente estructurados.</li> <li>3- Realizar análisis estadístico.</li> <li>4- Enviar resultado obtenido por el sistema.</li> </ul>
EC 22.2 Calcular la regresión logística bivariada estimación conjunta (probit) de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	<ul style="list-style-type: none"> <li>1- Recibir datos enviados a calcular.</li> <li>2- Validar datos enviados, verificar que se encuentren correctamente estructurados.</li> <li>3- Mostrar un mensaje de error.</li> </ul>

Tabla 25. Realizar Análisis Estadístico de Datos, sección Regresión Logística Bivariada, estimación conjunta (Probit).

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 23.1 Calcular la regresión de	Se realiza el cálculo correctamente y	V (Se envían los parámetros	Se envía el resultado de dicho cálculo a la	1- Recibir datos enviados a calcular.

riesgos de forma correcta.	se devuelve a la aplicación cliente un resumen de la regresión de riesgos.	correctamente correspondiendo con el tipo de análisis en cuestión)	aplicación cliente.	<ol style="list-style-type: none"> <li>2- Validar datos enviados, verificar que se encuentren correctamente estructurados.</li> <li>3- Realizar análisis estadístico.</li> <li>4- Enviar resultado obtenido por el sistema.</li> </ol>
EC 23.2 Calcular la regresión de riesgos de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	<ol style="list-style-type: none"> <li>1- Recibir datos enviados a calcular.</li> <li>2- Validar datos enviados, verificar que se encuentren correctamente estructurados.</li> <li>3- Mostrar un mensaje de error.</li> </ol>

*Tabla 26. Realizar Análisis Estadístico de Datos, sección Regresión de Riesgos.*

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 24.1 Calcular la regresión exponencial de forma correcta.	Se realiza el cálculo correctamente y se devuelve a la aplicación cliente un resumen de la regresión exponencial.	V (Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	Se envía el resultado de dicho cálculo a la aplicación cliente.	<ol style="list-style-type: none"> <li>1- Recibir datos enviados a calcular.</li> <li>2- Validar datos enviados, verificar que se encuentren correctamente estructurados.</li> <li>3- Realizar análisis estadístico.</li> <li>4- resultado obtenido por el sistema.</li> </ol>
EC 24.2 Calcular la regresión exponencial de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	<ol style="list-style-type: none"> <li>1- Recibir datos enviados a calcular.</li> <li>2- Validar datos enviados, verificar que se encuentren correctamente estructurados.</li> <li>3- Mostrar un mensaje de error.</li> </ol>

*Tabla 27. Realizar Análisis Estadístico de Datos, sección Regresión Exponencial.*

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 25.1 Calcular la regresión del factor de	Se realiza el cálculo correctamente y se devuelve a la	V (Se envían los parámetros correctamente	Se envía el resultado de dicho cálculo a la aplicación cliente.	<ol style="list-style-type: none"> <li>1- Recibir datos enviados a calcular.</li> <li>2- Validar datos enviados, verificar que se encuentren</li> </ol>

análisis bayesiano de forma correcta.	aplicación cliente un resumen de la regresión del factor de análisis bayesiano.	correspondiendo con el tipo de análisis en cuestión)		correctamente estructurados. 3- Realizar análisis estadístico. 4- Enviar resultado obtenido por el sistema.
EC 25.2 Calcular la regresión del factor de análisis bayesiano de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Mostrar un mensaje de error.

*Tabla 28. Realizar Análisis Estadístico de Datos, sección Regresión del Factor de Bayes.*

Escenario	Descripción	Variable-Parámetros	Respuesta del sistema	Flujo central
EC 26.1 Calcular la regresión de Gamma de forma correcta.	Se realiza el cálculo correctamente y se devuelve a la aplicación cliente un resumen de la regresión de Gamma.	V (Se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión)	Se envía el resultado de dicho cálculo a la aplicación cliente.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Realizar análisis estadístico. 4- Enviar resultado obtenido por el sistema.
EC 26.2 Calcular la regresión de Gamma de forma incorrecta.	Se le envía a la aplicación cliente un mensaje de error.	F (No se envían los parámetros correctamente correspondiendo con el tipo de análisis en cuestión o faltan datos)	Se envía un mensaje de error.	1- Recibir datos enviados a calcular. 2- Validar datos enviados, verificar que se encuentren correctamente estructurados. 3- Mostrar un mensaje de error.

*Tabla 29. Realizar Análisis Estadístico de Datos, sección Regresión de Gamma.*

## GLOSARIO DE TÉRMINOS

1. **API** - Interfaz de Programación de Aplicación. Es el conjunto de funciones y procedimientos (o métodos, en la programación orientada a objetos) que ofrece cierta biblioteca para ser utilizado por otro software como una capa de abstracción. Son usadas generalmente en las bibliotecas.
2. **AUP** - Proceso Unificado Ágil. Es una metodología de desarrollo de software.
3. **DATEC** - Centro de Tecnologías de Gestión de Datos.
4. **DSDM** - Método de desarrollo de sistemas dinámicos. Es una metodología de desarrollo de software.
5. **CUS** - Caso de Uso del Sistema.
6. **HTTP** - Protocolo de Hipertexto de Transferencia de Texto.
7. **IDE** - Entorno de Desarrollo Integrado.
8. **MVC** - Modelo Vista Controlador. Patrón de diseño arquitectónico.
9. **Open Up** - Proceso Unificado Abierto. Es una metodología de desarrollo de software.
10. **PHP** - Viene del acrónimo Hypertext Pre-processor. Es un lenguaje de programación interpretado (Lenguaje de alto rendimiento), diseñado originalmente para la creación de páginas web dinámicas.
11. **POO** - Programación Orientada a Objetos.
12. **Plantillas** - Es un archivo de texto que puede generar cualquier formato basado en texto.
13. **RF** - Requisito Funcional.
14. **RNF** - Requisito No Funcional.
15. **RUP** - Proceso Unificado de Rational. Es una metodología de desarrollo de software.
16. **SOAP** - Protocolo de Simple Acceso a Objetos.
17. **UCI** - Universidad de las Ciencias Informáticas.
18. **UML** - Lenguaje de Modelado Unificado.
19. **XML** - Lenguaje de Etiquetas Extensibles.
20. **XP** - Programación Extrema. Es una metodología de desarrollo de software.