



ISSN 1850-2512 (impreso)
ISSN 1850-2547 (en línea)

UNIVERSIDAD DE BELGRANO

Documentos de Trabajo

Facultad de Tecnología Informática

Estudio semántico de modelos construidos
por elicitadores independientes observando
el mismo problema. Proyecto: Completitud de
Modelos de Requisitos

N° 279

Graciela D. S. Hadad^{1,2},
Claudia Litvak^{1,2}, Jorge H. Doorn²

Departamento de Investigaciones
Agosto 2012

Universidad de Belgrano
Zabala 1837 (C1426DQ6)
Ciudad Autónoma de Buenos Aires - Argentina
Tel.: 011-4788-5400 int. 2533
e-mail: invest@ub.edu.ar
url: <http://www.ub.edu.ar/investigaciones>

1. Facultad de Ingeniería Informática, Universidad de Belgrano
2. DIIT, Universidad Nacional de La Matanza

1. Introducción

Uno de los principales problemas que se presenta en la Ingeniería de Requisitos es la incompletitud [Kotonya 98] [Loucopoulos 95] [Firesmith 05]. Esto se refiere a la dificultad en establecer si se ha elicitado y modelado toda la información requerida para desarrollar un sistema de software que cubra las expectativas y necesidades de los clientes y usuarios.

Entonces el problema de la incompletitud se vincula directamente con un tipo específico de defecto: las omisiones. Las técnicas de verificación y validación ayudan a mitigar este problema [Kotonya 98] [Loucopoulos 95], como así también, el uso de técnicas de elicitación apropiadas permiten lograr una adquisición de conocimiento más acabada sobre el dominio de la aplicación [Goguen 93]. Pero estas técnicas no permiten eliminar el problema ni tampoco estimar el grado de completitud alcanzado, para ello se requiere usar otro tipo de técnicas. El método de captura y recaptura [Otis 78] es un método predictivo sobre el tamaño de poblaciones, que ha sido utilizado en la Ingeniería de Software para establecer la cantidad de defectos que aún faltan detectar [Wohlin 98] y en la Ingeniería de Requisitos para determinar la cantidad de elementos faltantes en un modelo de requisitos [Doorn 03] [Ridao 06].

Según afirma Leite en [Leite 01], el proceso de definir requisitos es inherentemente incompleto, teniendo en vista la gran complejidad del mundo real. No es esperable alcanzar la completitud de un modelo. Se aspira entonces a lograr un modelo lo más completo posible.

Este mismo problema se presenta en otras áreas del conocimiento, tales como Biología, Economía, Medicina y Matemática. En la Ecología, sub-disciplina de la Biología, se usó originalmente el método de captura y recaptura [Otis 78] [White 82] para censar poblaciones cerradas en el mundo animal. Este método se ha aplicado en la Ingeniería de Software para estimar la cantidad de defectos que faltan aún detectar una vez ejecutadas técnicas de verificación y/o validación, principalmente inspecciones, sobre un artefacto de software [Wohlin 98] [Briand 00] [Pettersson 04] o sobre documentos de requisitos [Walia 08].

Una adaptación de este método predictivo, denominado DPM (Detection Profile Method) [Wohlin 98], se aplicó para estimar el tamaño de modelos de requisitos [Doorn 03] [Ridao 06], mediante la estimación de la cantidad de información aún faltante de elicitar. Es decir, propusieron utilizar el método de captura y recaptura para establecer el grado de completitud del modelo de requisitos construido.

Doorn y Ridao en [Doorn 03] estudiaron experimentalmente el uso de DPM para estimar el número total de términos de un glosario, denominado Léxico Extendido del Lenguaje (LEL). Hicieron una prueba utilizando tres grupos de elicidores que habían capturado 96 términos distintos elicitados del dominio del problema y el método predijo la existencia de 122 términos. Al aplicar este método de estimación, utilizando nueve grupos de elicidores sobre el mismo modelo LEL, donde ahora habían capturado 118 términos en total, el método estimó la existencia de 127 términos. Entonces, concluyeron que cuando el número de observadores independientes crece, es razonable suponer que la diferencia entre la cantidad estimada y la cantidad efectivamente obtenida se reduce. Es decir, que al aumentar el número de elicidores (ingenieros de requisitos) la cantidad de elementos estimados será igual a la cantidad realmente existente en el dominio del problema.

En la sección siguiente se plantea una hipótesis para poder dar una explicación a la diferencia de términos detectados por los nueve grupos y a la cantidad de símbolos que aún les faltaba encontrar en conjunto a los grupos según el método DPM aplicado. En la sección 3 se presenta un estudio estadístico para poder establecer el cumplimiento o no de la hipótesis planteada. En la sección 4 se resume una lista de observaciones en base al estudio estadístico realizado. En la sección 5 se presenta el análisis semántico realizado y el posterior estudio estadístico, seguido de la sección 6 donde se realizan observaciones más refinadas teniendo en consideración la parte semántica. Finalmente en la sección 7 se mencionan conclusiones del trabajo realizado y se delinear los pasos a seguir.

2. Hipótesis de Trabajo

Se analizó la aplicación del método de captura y recaptura [Wohlin 98] realizado por Doorn y Ridao [Doorn 03] sobre el modelo Léxico Extendido del Lenguaje (LEL). El LEL es un glosario del dominio de

la aplicación con la definición y connotación de los términos (símbolos) utilizados en él [Leite 93]. Este modelo léxico había sido previamente elaborado por nueve grupos independientes de ingenieros de requisitos, aplicando todos ellos la misma técnica de elicitación sobre el mismo caso Sistema de Planes de Ahorro Previo para la Adquisición de Vehículos 0Km [Mauco 97] [Rivero 98].

Si dos o más grupos construyen el LEL del mismo dominio de la aplicación, se ha encontrado que todos los grupos detectan una cantidad de símbolos diferente y que los conjuntos con menor cantidad de símbolos no son subconjuntos propios de los que tienen mayor cantidad de símbolos. Surgen entonces algunas hipótesis al estudiar qué símbolos de un LEL encuentran diferentes grupos y cómo se relacionan los símbolos hallados por los diferentes grupos.

En otras palabras, dos grupos cualesquiera detectan una cantidad diferente de símbolos, pero con la propiedad que el primero de los grupos detecta algunos símbolos que no detecta el segundo y viceversa. Bajo esta circunstancia, se podrían considerar tres posibilidades:

- 1) Los grupos tienen características que los diferencian.
- 2) Se trata de dos problemas diferentes.
- 3) Las diferencias se producen por el mero azar.

Una de las variantes del método de captura y recaptura se basa en la hipótesis 1) y utiliza un coeficiente de corrección que distingue la probabilidad de detección de los grupos. Lo cual ha inducido a centrarse en dicha hipótesis [Doorn 03]. Frecuentemente, la hipótesis 2) es rechazada por parecer absurda, y la 3) por considerarse de baja incidencia. Analizando estas cuestiones en detalle, planteamos que la diferencia está en la relación entre el grupo y el problema; es decir, que cada grupo tiene cierta tendencia a observar mejor algunos aspectos del problema que otro.

Lo primero que se debe hacer al construir un LEL de un problema es definir los límites del universo de discurso (UdeD). ¿Cuándo hacemos esto? ¡Cuándo menos sabemos del problema! Habitualmente, se están definiendo los límites del UdeD en el momento en que no se conoce nada del mismo. Esto es la confirmación del teorema de Jackson [Jackson 95], en el sentido de no usar el enfoque “top down” en la Ingeniería de Requisitos, dado que la división del sistema en subsistemas no se debe hacer cuando menos sabemos del problema.

Como consecuencia de este análisis, se elaboró una hipótesis de trabajo (ver Figura 1): “Los grupos de ingenieros tratan problemas diferentes”.

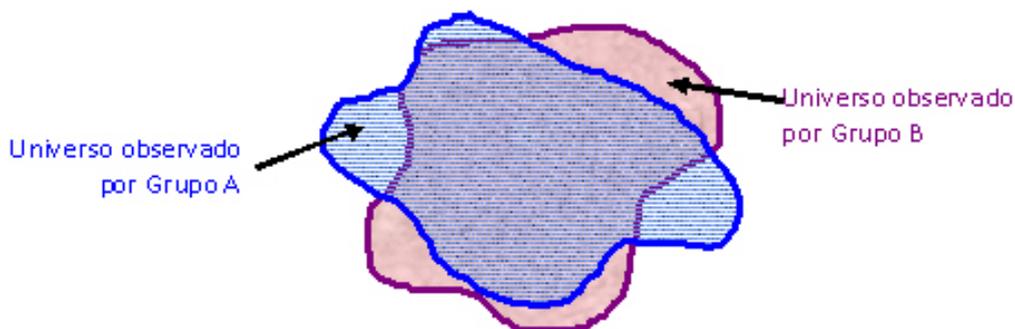


Figura 1. Dos problemas diferentes

Bajo esta visión de problemas diferentes, se realizó un nuevo estudio comparativo de esos nueve modelos léxicos considerando solo componentes de alto nivel del modelo (nombres de los términos léxicos y sus sinónimos). Para ello, se identificaron cinco categorías de sub-problemas o tópicos que abarcaban el problema en estudio y se analizaron estadísticamente cómo cada grupo profundizó o no la elicitación de información correspondiente a cada categoría. Las conclusiones preliminares confirmarían promisoriamente la hipótesis planteada.

3. Estudio Estadístico bajo la Hipótesis de Trabajo

Se elaboró un cuadro estadístico, calculando la cantidad de términos del LEL identificados por cada grupo en cada categoría de sub-problema, el que se presenta en la Tabla 1. Se observa que hay grupos que ahondan más en unos temas que otros, mientras que hay grupos que no observan parte importante de la información. Las conclusiones preliminares de este estudio estadístico se presentaron en [Hadad 12].

Tabla 1. Estadísticas por grupo y por categoría

Categoría	1-Adhesión	2-Adjudicación	3-Gestión de Cuotas	4-Administración de Bien Tipo	5-Administración de Plan	Total del Grupo
Grupo 1	7	13	13	17	4	54
Grupo 2	4	6	9	6	4	29
Grupo 3	8	6	8	4	5	31
Grupo 4	6	7	7	9	6	35
Grupo 5	8	10	21	15	13	67
Grupo 6	8	7	3	4	4	26
Grupo 7	8	10	10	9	9	46
Grupo 8	4	7	5	6	2	24
Grupo 9	7	12	11	12	13	55
Total de la Categoría	60	78	87	82	60	
Total sin repetición	20	26	35	27	21	129
Media	6,67	8,67	9,67	9,11	6,67	
Varianza	2,75	7,00	27,25	22,11	16,50	
Desviación Estándar	1,66	2,65	5,22	4,70	4,06	
t de Student	3,00	3,00	3,00	3,00	3,00	
Límite de Confianza Superior	8,32	11,31	14,89	13,81	10,73	
Límite de Confianza Inferior	5,01	6,02	4,45	4,41	2,60	

Se considera que las curvas obtenidas para todas las categorías son aproximadamente gaussianas, con una leve asimetría a la derecha. Esto nos permite utilizar las siguientes fórmulas aplicables a una cantidad pequeña de casos [Cavanos 88].

Media de la categoría j:
$$\bar{Q}_j = \frac{\sum_{i=1}^N Q_{ij}}{N} \quad (1)$$

Varianza de la categoría j:
$$\sigma_j^2 = \frac{\sum_{i=1}^N (Q_{ij} - \bar{Q}_j)^2}{N-1} \quad (2)$$

Desviación Estándar de la categoría j:
$$\sigma_j = \sqrt{\sigma_j^2} \quad (3)$$

Límite de Confianza Superior para la categoría j:
$$LCS_j = \bar{Q}_j + \frac{t \cdot \sigma_j}{\sqrt{N}} \quad (4)$$

Límite de Confianza Inferior para la categoría j:
$$LCI_j = \bar{Q}_j - \frac{t \cdot \sigma_j}{\sqrt{N}} \quad (5)$$

Siendo,

- i : grupo
- j : categoría
- Q_{ij} : cantidad de símbolos hallados por el grupo i en la categoría j
- N : total de grupos = 9
- t : distribución t-Student

En la Tabla 1 se aplicaron las fórmulas (1) a (5), donde se utilizó la distribución t de Student dado que se tienen muestras de tamaño pequeño N = 9 [Cavanos 88]. Para el cálculo del intervalo de confianza se consideró un nivel de confianza del 99,1%, es decir, se aplicó la distribución t = 3 para 8 grados de libertad (se utiliza N - 1).

En la Tabla 1 podemos observar que hay grupos, marcados en gris, que consideraron escasamente una categoría determinada del problema. Estos grupos se encuentran por debajo del límite de confianza

inferior, lo que implicaría que no consideraron detalladamente esta categoría en particular. También podemos observar los grupos que profundizaron en una categoría determinada, marcados en negro. Estos grupos se encuentran por arriba del límite de confianza superior. En este caso implicaría que se dedicaron a estudiar en mayor medida esa categoría del UdeD.

En la Figura 2 se representan los grupos y el total de símbolos del LEL hallados por cada grupo, con información extraída de la columna Total del Grupo de la Tabla 1. Se puede aquí observar cómo los grupos han capturado una muy diversa cantidad de símbolos, variando desde un grupo con 67 elementos hasta otro grupo con solo 24 símbolos del LEL.

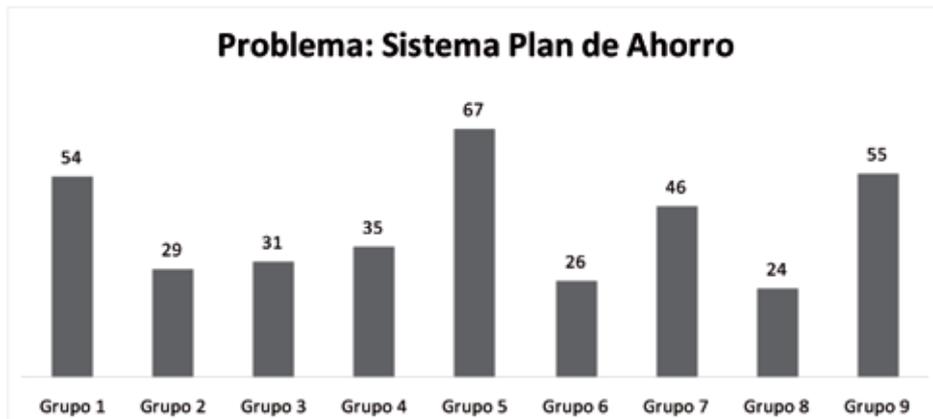


Figura 2. Total de Símbolos del LEL identificados por grupo

Se realizó un estudio estadístico analizando individualmente cada categoría de sub-problemas, con el fin de evaluar la visión de los distintos grupos de ingenieros respecto a cada sub-problema.

3.1. Análisis estadístico para la categoría Gestión de Cuotas

Analizando aisladamente la categoría Gestión de Cuotas, tomando los datos de la columna correspondiente de la Tabla 1, se puede observar que el Grupo 5 supera ampliamente la cantidad de símbolos hallados frente a los restantes grupos (ver Figura 3).

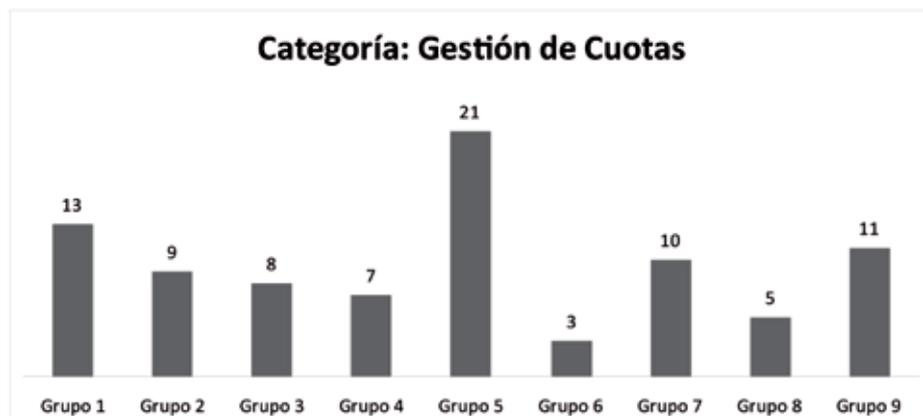


Figura 3. Total de símbolos por grupo para Gestión de Cuotas

Se analizó la distribución por rangos de símbolos identificados por los grupos para esta categoría. Se hicieron varias representaciones y la que sugirió la hipótesis planteada fue el histograma con agrupamientos en pares (ver Figura 4), donde se observa la posible existencia de un único UdeD y un valor atípico (denominado en estadística “outlier”) [Barnett 94], que es el Grupo 5 que halló 21 símbolos del LEL.

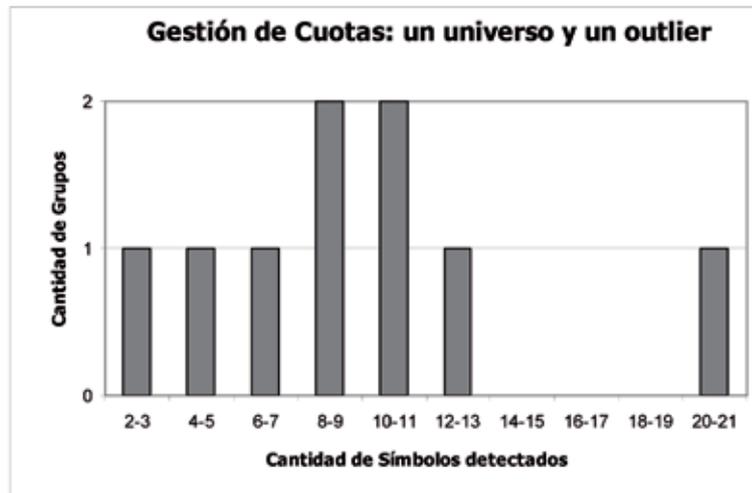


Figura 4. Distribución de grupos para la categoría Gestión de Cuotas

Se realizaron nuevos estudios estadísticos para determinar si el Grupo 5 había observado un universo de discurso distinto. Se utilizaron las siguientes fórmulas para la categoría j y sin el valor atípico α , siendo N el total de grupos y usando la distribución t-Student para una muestra pequeña:

Media de la categoría j sin α :

$$t'_j = \frac{|\alpha - \bar{Q}'_j|}{\sigma'_j / \sqrt{N}} \quad (6)$$

Varianza de la categoría j sin α :

$$\sigma'^2_{j'} = \frac{\sum_{i=1}^N (Q_i - \bar{Q})^2 - (N-1)(\alpha - \bar{Q})^2}{N-2} \quad (7)$$

Desviación Estándar de la categoría j sin α :

$$\sigma'_j = \sqrt{\sigma'^2_{j'}} \quad (8)$$

t-Student de la categoría j sin α :

$$\bar{Q}'_j = \frac{\sum_{i=1}^N Q_i - \alpha}{N-1} \quad (9)$$

Estas fórmulas se aplicaron para la categoría $j = 3$ (Gestión de Cuotas) y el outlier $\alpha = 21$ correspondiente al Grupo 5, dando los siguientes valores:

$$\begin{aligned} \bar{Q}'_3 &= 8,250 \\ \sigma'^2_{3'} &= 10,500 \\ \sigma'_{3'} &= 3,240 \\ t'_{3'} &= 11,129 \end{aligned}$$

En base a la Tabla t de Student (ver anexo A) con el valor $t'_{3'}$ calculado y 7 grados de libertad, se obtuvo p-valor $<0,0005$ indicando un nivel de significación de 0,05, es decir, se puede establecer que hay más de un 99,95% de probabilidad de que el Grupo 5 haya observado al sub-problema Gestión de Cuotas como un UdeD distinto al que observaron los restantes grupos. Del análisis estadístico realizado, se observa que el Grupo 5 ha profundizado el estudio de la categoría Gestión de Cuotas, excediéndose en la identificación de símbolos por encima de la media.

3.2. Análisis estadístico para la categoría Administración de Bien Tipo

Se estudió estadísticamente la cantidad de símbolos del LEL hallados en Administración de Bien Tipo, donde se observa visualmente a través de la Figura 5 una disparidad en la cantidad de símbolos detectados por cada grupo.

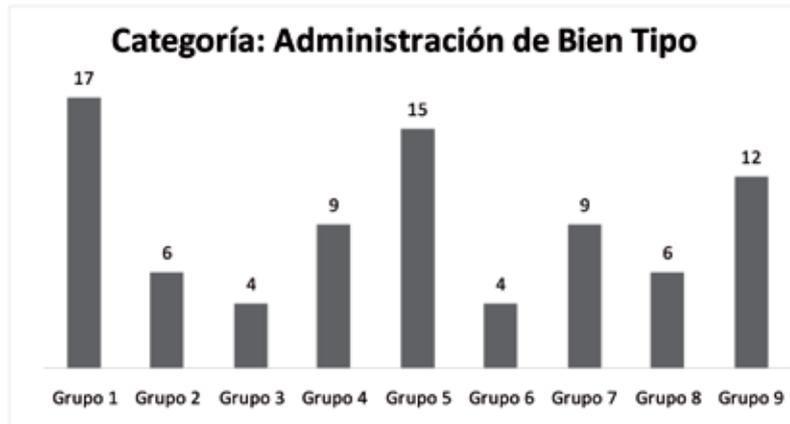


Figura 5. Total de símbolos por grupo para Administración de Bien Tipo

Se analizó la distribución por rangos de símbolos para esta categoría, la que se presenta en la Figura 6. En ella se observan gráficamente tres UdeD distintos: quienes hallaron de 3 a 6 símbolos del LEL, denominado UdeD₁; quienes hallaron de 9 a 12 símbolos, denominado UdeD₂; y quienes hallaron de 15 a 18 símbolos, denominado UdeD₃.

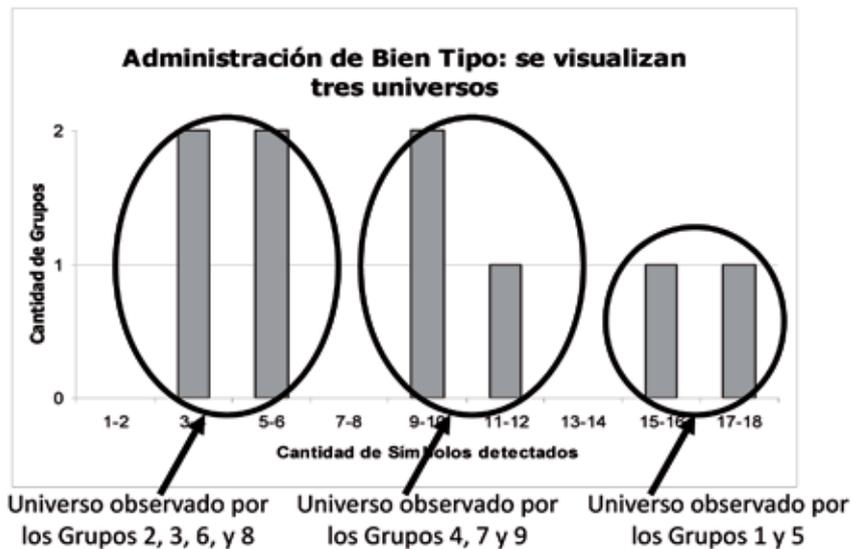


Figura 6. Distribución de grupos para Administración de Bien Tipo

Se realizó un estudio estadístico, basado en t de Student, para comparar grupos de datos o universos y obtener el nivel de significación de que sean universos diferentes [Pértega 01]. Se aplicaron las siguientes fórmulas para cada UdeD en la categoría j:

Media del UdeD_k:

$$\bar{Q}_j, U_k = \frac{\sum_{i=1}^k Q_i}{n_k} \tag{10}$$

Varianza del UdeD_k:

$$\sigma^2, U_k = \frac{\sum_{i=1}^k (Q_i - \bar{Q}_j, U_k)^2}{n_k - 1} \tag{11}$$

Desviación Estándar del UdeD_k:

$$\sigma_j, U_k = \sqrt{\sigma^2, U_k} \tag{12}$$

t-Student del UdeD_k respecto al UdeD_{k+1}:

$$t_j, U_k = \frac{\bar{Q}_j, U_{k+1} - \bar{Q}_j, U_k}{\sqrt{\frac{(n_k - 1)\sigma^2, U_k + (n_{k+1} - 1)\sigma^2, U_{k+1}}{n_k + n_{k+1} - 2} \left(\frac{1}{n_k} + \frac{1}{n_{k+1}} \right)}} \tag{13}$$

Siendo,

k : UdeD

j : categoría

Q_{ij} : cantidad de símbolos del grupo i que pertenece al $UdeD_k$ en la categoría j
 n_k : cantidad de elementos del $UdeD_k$

Los valores obtenidos para la categoría Administración de Bien Tipo ($j = 4$) se muestran en la Tabla 2, donde se ha comparado el $UdeD_2$ contra el $UdeD_1$, y el $UdeD_3$ contra el $UdeD_2$, aplicando las fórmulas (10) a (13). Con los grados de libertad ($n_k+n_{k+1}-2$) y el valor calculado de t-Student ($t_j U_k$), se obtuvo el dato p-valor (ver Anexo A – Tabla t-Student).

Tabla 2. Administración de Bien Tipo: Probabilidad de diferencia estructural

k	Grupos	UdeD _k	n _k	Media	Varianza	Desviación Estándar	t-Student	p-valor
1	2, 3, 6, 8	{4, 4, 6, 6}	4	5	1,333	1,155	4,629	0,004
2	4,7,9	{9, 9, 12}	3	10	3,000	1,732	4,025	0,004
3	1,5	{15, 17}	2	16	2,000	1,414		

En base a los resultados obtenidos en la Tabla 2, se podría concluir que se rechaza la igualdad entre el $UdeD_1$ y el $UdeD_2$, es decir, ambos universos presentan una diferencia estructural. Esto fue observado con un nivel de significación de 0,05 ($p\text{-valor} = 0,004 < 0,05$; probabilidad de diferencia estructural 99,60%). Del mismo modo, se podría concluir que el $UdeD_2$ y el $UdeD_3$ presentan también una diferencia estructural, lo cual fue observado con similar nivel de significación ($p\text{-valor} = 0,004 < 0,05$; probabilidad 99,63%).

3.3. Análisis estadístico para la categoría Administración de Plan

Se analizó también en la categoría Administración de Plan (ver datos en la columna correspondiente en Tabla 1) la distribución por rangos de símbolos identificados por cada grupo. Se hicieron varias representaciones y la que sugirió la hipótesis fue el histograma con agrupamientos en pares (ver Figura 7).

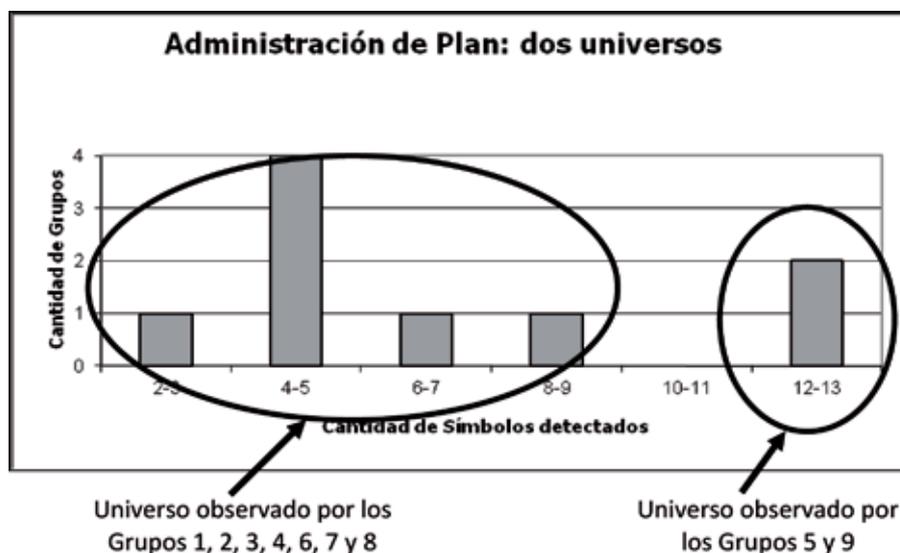


Figura 7. Distribución de grupos para la categoría Administración de Plan

De la distribución por rangos se visualizan dos universos distintos: quienes hallaron de 3 a 9 símbolos del LEL, denominado $UdeD_1$, y los dos grupos 5 y 9 que hallaron 13 símbolos, denominado $UdeD_2$. Se aplicaron las fórmulas de comparación de universos para comprobar que son realmente universos distintos (ver Tabla 3).

Tabla 3. Administración de Plan: Probabilidad de diferencia estructural

k	Grupos	UdeD _k	n _k	Media	Varianza	Desviación Estándar	t-Student	p-valor
1	1, 2, 3, 4, 6, 7, 8	{4, 4, 5, 6, 4, 9, 2}	7	4,857	1,333	1,155	5,960	<0,0005
2	5, 9	{13, 13}	2	13,000	0	0		

En base a los resultados estadísticos presentados en la Tabla 3, se podría concluir que el UdeD₁ y el UdeD₂ presentan una diferencia estructural con un nivel de significación de 0,05 ($p\text{-valor} < 0,0005 < 0,05$; probabilidad de diferencia estructural mayor a 99,95%). La conclusión que obtenemos es que tanto el Grupo 5 como el Grupo 9 pareciera que profundizaron el estudio de la categoría Administración de Plan frente a los restantes grupos.

3.4. Análisis estadístico para la categoría Adjudicación

Respecto a la categoría Adjudicación, se analizó la distribución por rangos de símbolos, la que se presenta en la Figura 8. En ella se observan gráficamente dos UdeD distintos: quienes hallaron de 6 a 7 símbolos del LEL, denominado UdeD₁, y quienes hallaron de 10 a 13 símbolos, denominado UdeD₂.

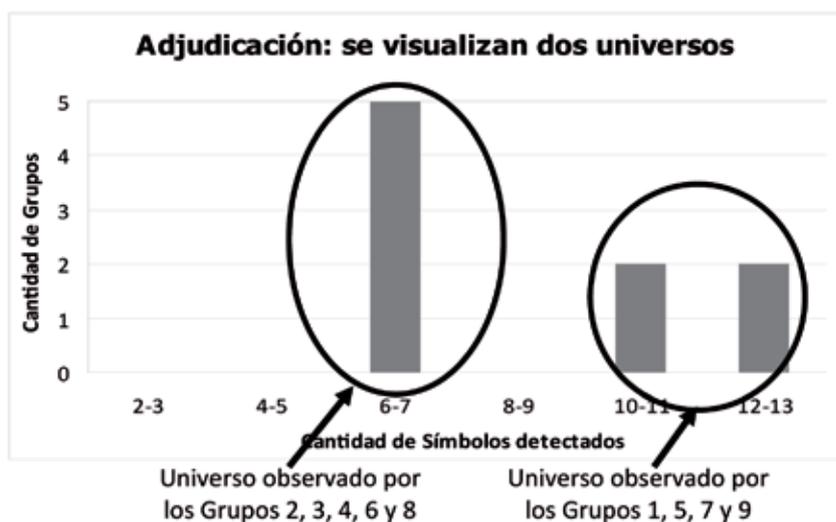


Figura 8. Distribución de grupos para la categoría Adjudicación

Entonces, se realizó el estudio estadístico para comparar universos usando t-Student [Pértega 01]. Los valores obtenidos para la categoría Adjudicación ($j = 2$) se muestran en la Tabla 4, donde se ha comparado el UdeD₂ contra el UdeD₁.

Tabla 4. Adjudicación: Probabilidad de diferencia estructural

k	Grupos	UdeD _k	n _k	Media	Varianza	Desviación Estándar	t-Student	p-valor
1	2, 3, 4, 6, 8	{6, 6, 7, 7, 7}	5	6,6	0,3	0,548	5,014	0,0002
2	1, 5, 7, 9	{13, 10, 10, 12}	4	11,0	2,0	1,414		

Se podría concluir de estos cálculos que el UdeD₁ y el UdeD₂ presentan una diferencia estructural, lo cual fue observado con un nivel de significación de 0,05 ($p\text{-valor} = 0,0002 < 0,05$; probabilidad de diferencia estructural 99,98%).

3.5. Análisis estadístico para la categoría Adhesión

Al analizar la categoría Adhesión (ver datos en la columna correspondiente en Tabla 1) se observó una distribución muy pareja de los datos obtenidos, como se muestra en la Figura 9.

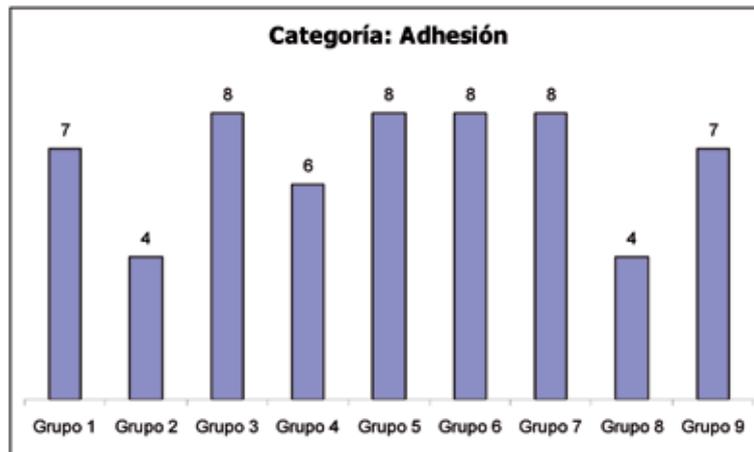


Figura 9. Total de símbolos por grupo para Adhesión

Como los datos presentan una distribución aproximadamente normal con pequeñas diferencias entre ellos (ver Figura 9), se aplicó la distribución chi cuadrado (χ^2) [Cavanos 88]. Un valor bajo de χ^2 es un indicador de rechazo de independencia. Entonces, se ha calculado una estadística χ^2 utilizando las siguientes fórmulas:

Chi-cuadrado:

$$\chi^2_j = \frac{\sum_{i=1}^N (Q_{ij} - E_{ij})^2}{E_{ij}} \quad (14)$$

Porcentaje Global de Observados:

$$PG_j = \frac{\sum_{i=1}^N Q_{ij}}{\sum_{i=1}^N \sum_{j=1}^N Q_{ij}} \quad (15)$$

Siendo,

- Q_{ij} : cantidad de símbolos observados por el grupo i en la categoría j
- E_{ij} : frecuencia esperada para el grupo i en la categoría j
- N : Total de grupos = 9
- i : grupo
- j : categoría

En la Tabla 5 se muestra el resultado de la prueba chi-cuadrado para la categoría Adhesión ($j=1$). Se utilizó la fórmula (15), dando el valor $PG_1 = 11,11 \%$, para el cálculo de la frecuencia esperada (E_{ij}).

Tabla 5. Categoría Adhesión: Aplicación de χ^2 cuadrado

Adhesión	Cantidad de símbolos observados	Complemento de cantidad de símbolos	Total de símbolos	Porcentaje de Observados / Total	Frecuencia esperada	Complemento de Frecuencia esperada	χ^2
Grupo 1	7	53	60	11,67%	6,67	53,33	0,02
Grupo 2	4	56	60	6,67%	6,67	53,33	1,07
Grupo 3	8	52	60	13,33%	6,67	53,33	0,27
Grupo 4	6	54	60	10,00%	6,67	53,33	0,07
Grupo 5	8	52	60	13,33%	6,67	53,33	0,27
Grupo 6	8	52	60	13,33%	6,67	53,33	0,27
Grupo 7	8	52	60	13,33%	6,67	53,33	0,27

Grupo 8	4	56	60	6,67%	6,67	53,33	1,07
Grupo 9	7	53	60	11,67%	6,67	53,33	0,02
Total	60	480	540				3,30

Como todas las frecuencias esperadas obtenidas son mayores a 5, entonces se puede aplicar la estadística chi-cuadrado. Cabe aclarar que no importa que dos frecuencias reales (cantidad de símbolos observados) sean menores (Grupos 2 y 8). Se obtuvo el valor $\chi^2_1 = 3,30$ para la categoría Adhesión. De la Tabla chi-cuadrado (ver Anexo B) para el nivel de significación 5% y 8 grados de libertad ($N - 1$), se obtiene el valor de $\chi^2 = 15,51$. Dado que $\chi^2_1 = 3,3 < 15,51$, no se rechaza la igualdad. Como conclusión, rechazamos que son independientes, por lo tanto todos los datos obtenidos de la categoría Adhesión están asociados.

Por ende, podemos concluir que los grupos han observado para la categoría Adhesión el mismo UdeD. Es la única categoría donde todos los grupos observaron el mismo universo de discurso. En las demás categorías observaron varios universos, como en Adjudicación donde observaron dos UdeD, en Administración de Plan donde observaron dos UdeD, y en Administración de Bien Tipo donde observaron tres UdeD. En el caso de Gestión de Cuotas se encontró un único UdeD y un outlier.

4. Observaciones sobre los resultados estadísticos

Podemos observar en base al análisis estadístico realizado que:

- El Grupo 5 “presuntamente” exageró en el estudio de las categorías Gestión de Cuotas, Administración de Bien Tipo, Administración de Plan y Adjudicación. Es el grupo que aparentemente profundizó en 4 de las 5 categorías, siendo la quinta categoría aquella que todos la observaron en un mismo nivel de detalle: Adhesión.
- El Grupo 1 “presuntamente” exageró en el estudio de las categorías Administración de Bien Tipo y Adjudicación.
- El Grupo 9 “presuntamente” exageró en el estudio de las categorías Administración de Plan y Adjudicación.
- El Grupo 7 “presuntamente” exageró en el estudio de la categoría Adjudicación.
- Los Grupos 2, 3, 6 y 8 “presuntamente” estudiaron débilmente las categorías Administración de Bien Tipo y Adjudicación.
- El Grupo 4 “presuntamente” observó débilmente la categoría Adjudicación.
- Los Grupos 4 y 7 elicitron información para todas las categorías con similar nivel de detalle, con excepción de la categoría Adjudicación, donde el primero la estudió débilmente y el segundo la profundizó. Esto indicaría que sólo dos grupos observaron el mismo UdeD (Grupos 4 y 7) en 4 de las 5 categorías.
- Ningún grupo estudió una categoría con mayor profundidad y otra con menor profundidad. Es decir, hubo grupos que “presuntamente” profundizaron en alguna/s categorías mientras que fueron otros los grupos que “presuntamente” estudiaron débilmente alguna/s categorías.
- La categoría Adhesión fue considerada por todos los grupos con un nivel de detalle homogéneo. “Presuntamente” todos los grupos observaron el mismo universo para esta categoría. Con lo cual, esta podría ser considerada una categoría “central” para el dominio de la aplicación.
- La categoría Adjudicación polarizó el estudio sobre ella: varios grupos la estudiaron con “aparente” profundidad y varios grupos la estudiaron con “aparente” poco detalle.

Estas conclusiones basadas en el análisis estadístico realizado, permitieron confirmar la hipótesis de trabajo, pues se ha podido constatar que varios grupos observaron distintos UdeD para una misma categoría.

Por otro lado, cabe aclarar que la palabra “presuntamente”, utilizada al puntualizar las observaciones, indica que estas han sido realizadas considerando exclusivamente el estudio estadístico realizado. Con lo cual, las conclusiones provenientes de este estudio son preliminares. Las mismas están supeditadas a que la visión semántica las confirme o modifique.

5. Análisis semántico aplicado a los resultados estadísticos

Finalizado el estudio estadístico, se realizó un estudio de la semántica de todos los términos del LEL definidos por los nueve grupos. Se analizó la denotación y connotación de los términos que se consideraron dudosos. Los términos dudosos fueron examinados para determinar su relevancia en el UdeD (si aportan o no conocimiento) y su pertenencia al UdeD (existen realmente en él o fueron creados artificialmente por el grupo de elicidores). Se encontraron términos cuyo nombre no existía en el dominio del problema y, en algunos casos, sus denotaciones y connotaciones contenían información relevante. Asimismo, se analizó la denotación y connotación de todos aquellos términos cuyos nombres no eran exactamente iguales o no eran sinónimos evidentes de símbolos definidos por otros grupos, con el fin de determinar aquellos símbolos que habían sido identificados por un único grupo, es decir, ningún otro grupo los había detectado. En unos pocos casos, se encontraron sinónimos dentro del mismo grupo de elicidores: correspondían a dos términos definidos separadamente pero que tenían el mismo significado (denotación más connotación).

En resumen, los símbolos calificados de no relevantes involucran distintos aspectos: i) información que estaba fuera del límite del UdeD, ii) información con un exceso de nivel de detalle innecesario para un glosario, que podía ser mencionada dentro de otro símbolo, y iii) información que estaba contenida en otros símbolos.

La Tabla 6 presenta los resultados obtenidos por grupo y por categoría considerando los símbolos relevantes.

Tabla 6. Resultados post-análisis semántico

Categoría	1-Adhesión	2-Adjudicación	3-Gestión de Cuotas	4-Administración de Bien Tipo	5-Administración de Plan	Total del Grupo
Grupo 1	6	11	13	12	4	46
Grupo 2	4	6	9	6	3	28
Grupo 3	6	5	7	4	2	24
Grupo 4	6	7	7	9	6	35
Grupo 5	7	9	20	15	11	62
Grupo 6	4	4	1	3	2	14
Grupo 7	7	11	11	9	6	44
Grupo 8	4	5	5	6	2	22
Grupo 9	7	10	11	11	12	51
Total de la Categoría	51	68	84	75	48	
Total sin repetición	13	20	31	19	14	97

A continuación se detalla el resultado del análisis semántico realizado sobre cada categoría de subproblema, y se compara con el análisis previo netamente estadístico.

5.1. Análisis semántico para la categoría Gestión de Cuotas

Según el análisis estadístico de la categoría Gestión de Cuotas (ver Figura 4), se encontró un universo observado por ocho grupos y un valor atípico sobredimensionado respecto a lo hallado por los otros grupos. En la Tabla 7 se detallan los resultados del análisis semántico para esta categoría. Analizando la distribución por rangos de símbolos, considerando ahora los símbolos relevantes por grupo (ver Figura 10), se visualizó en el histograma por ternas la posible existencia de un universo y un valor atípico ($\alpha = 20$) correspondiente al mismo Grupo 5.

Tabla 7. Gestión de Cuotas: resultados del análisis semántico

Gestión de Cuotas	Total de Símbolos (Estadísticos)	Total de Símbolos Relevantes (Semánticos)	Símbolos No Relevantes	Nombres de Símbolos Inexistentes	Símbolos Únicos	Símbolos Únicos Relevantes
Outlier						
Grupo 5	21	20	1	0	7	6
UdeD						
Grupo 1	13	13	0	0	2	2
Grupo 2	9	9	0	0	1	1
Grupo 3	8	7	1	0	1	0
Grupo 4	7	7	0	0	0	0
Grupo 6	3	1	2	1	2	0
Grupo 7	10	11	0	0	3	3
Grupo 8	5	5	0	0	0	0
Grupo 9	11	11	0	0	2	2
Total	87	84	4	1	18	14

Total de Símbolos Sin Repetición	35
Total de Símbolos Sin Repetición No Relevantes	4
Total de Símbolos Sin Repetición Relevantes	31

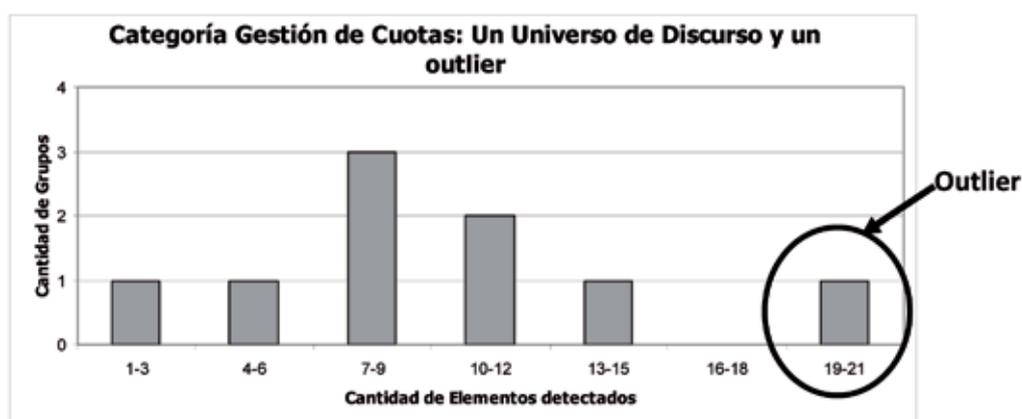


Figura 10. Distribución de grupos para Gestión de Cuotas post-análisis semántico

Se rehicieron los estudios estadísticos para determinar si el Grupo 5 había observado un universo distinto al de los restantes grupos. Se aplicaron entonces las fórmulas (6) a (9) anteriormente planteadas descartando el outlier ($\alpha = 20$) perteneciente al Grupo 5 y usando la distribución t-Student para una muestra pequeña, dando los siguientes valores:

$$\begin{aligned}\bar{Q}'_3 &= 8,000 \\ \sigma^2'_3 &= 14,865 \\ \sigma'_3 &= 3,854 \\ t'_3 &= 8,806\end{aligned}$$

Con el nuevo valor de t-Student t'_3 (usando la Tabla del Anexo A) con 7 grados de libertad, se obtuvo el p-valor $< 0,0005$ indicando un nivel de significación de 0,05, es decir, se puede establecer que hay más de un 99,95% de probabilidad de que el Grupo 5 haya observado al sub-problema Gestión de Cuotas como un universo distinto al que observaron los restantes grupos.

El Grupo 5 agregó 6 símbolos relevantes que no habían sido considerados por los restantes grupos. De los 21 símbolos detectados por dicho grupo, solo uno fue considerado no relevante, mostrando que el Grupo 5 estudió en profundidad la categoría Gestión de Cuotas, con mucha diferencia en la cantidad

de símbolos detectados respecto a los encontrados por los otros grupos, quienes observaron un universo con menos información (en promedio 7 símbolos), aunque algunos grupos aportaron entre 1 a 3 símbolos únicos relevantes. En resumen, el Grupo 5 detectó el 65% del total de símbolos no repetidos hallados por todos los grupos en esta categoría. En oposición, el Grupo 6 solo vio un símbolo relevante en esta categoría, símbolo encontrado junto a otros 5 más.

Esta ha sido la categoría con más símbolos relevantes (31), habiéndose observado solo 4 símbolos no relevantes en conjunto. No hubo ningún símbolo visto por los 9 grupos, aunque hubo 6 símbolos vistos entre 5 y 7 grupos, siendo que más de la mitad de los símbolos (58%) fueron vistos por 1 o 2 grupos (muchos de ellos encontrados por el Grupo 5).

5.2. Análisis semántico para la categoría Administración de Bien Tipo

La categoría Administración de Bien Tipo había sido estudiada con tres visiones diferentes del UdeD según el análisis estadístico realizado (ver Figura 6). Podemos observar en la Tabla 8 los datos obtenidos luego del análisis semántico de las denotaciones y las connotaciones de los símbolos no exactamente iguales y de los símbolos considerados no relevantes. Se encontró que el grupo que más símbolos había detectado (Grupo 1) era también el grupo con más símbolos no relevantes: un tercio de sus símbolos. Al realizar nuevamente la distribución por rangos, surgió que este Grupo 1 pasó ahora a formar parte de la población del UdeD₂ (antes pertenecía al UdeD₃), como se muestra en la Figura 11.

Tabla 8. Administración de Bien Tipo: resultados del análisis semántico

Administración de Bien Tipo	Total de Símbolos (Estadísticos)	Total de Símbolos Relevantes (Semánticos)	Símbolos No Relevantes	Nombres de Símbolos Inexistentes	Símbolos Únicos	Símbolos Únicos Relevantes
UdeD₁						
Grupo 2	6	6	0	0	0	0
Grupo 3	4	4	0	0	0	0
Grupo 6	4	3	2	3	2	0
Grupo 8	6	6	0	0	0	0
UdeD₂						
Grupo 4	9	9	0	0	0	0
Grupo 7	9	9	0	0	0	0
Grupo 9	12	11	0	0	1	1
UdeD₃						
Grupo 1	17	12	6	0	8	2
Grupo 5	15	15	0	0	2	2
Total	82	75	8	3	13	5

Total de Símbolos Sin Repetición	27
Total de Símbolos Sin Repetición No Relevantes	8
Total de Símbolos Sin Repetición Relevantes	19

La Tabla 8 muestra que los grupos que observaron el UdeD₁ encontraron de 3 a 6 símbolos relevantes y no aportaron ningún símbolo único relevante, no considerado por los otros grupos. Los grupos que observaron el UdeD₂ encontraron de 9 a 12 símbolos y agregaron tres símbolos únicos relevantes. Ahora, el UdeD₃ fue observado por un solo grupo (Grupo 5), que detectó 15 símbolos relevantes, aportando 2 símbolos relevantes únicos.

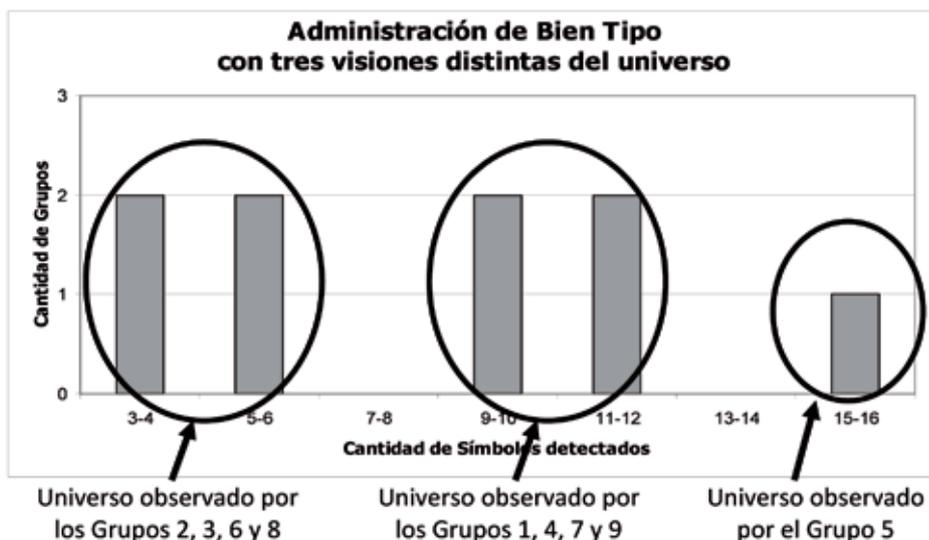


Figura 11. Distribución de grupos para Administración de Bien Tipo según análisis semántico

Realizando nuevamente el estudio comparativo de universos [Pértega 01], considerando los símbolos del LEL relevantes en cada grupo, obtenemos los resultados de la Tabla 9, donde se confirma la existencia de tres universos dado que se observa un nivel de significación $< 0,05$ (probabilidad de diferencia estructural de 99,69% y 99,91% respectivamente).

Tabla 9. Administración de Bien Tipo: Probabilidad de diferencia estructural considerando símbolos relevantes

k	Grupos	UdeD _k	n _k	Media	Varianza	Desviación Estándar	t-Student	p-valor
1	2, 3, 6, 8	{6, 4, 3, 6}	4	4,75	2,250	1,500	5,185	0,003
2	1, 4, 7, 9	{12, 9, 9, 11}	4	10,25	2,250	1,500	2,832	0,001
3	5	{15}	1	15	0	0		

El Grupo 5 detectó el 79% (15 símbolos) del total de símbolos relevantes hallados por todos los grupos, solo le faltó detectar 4 símbolos relevantes. Mientras que el Grupo 6 fue el que encontró menos símbolos relevantes (3 símbolos) aunque detectó 1 símbolo no visto por el Grupo 5.

De este estudio semántico en la categoría Administración de Bien Tipo podemos concluir que el UdeD₁ es relativamente pobre de información respecto a los otros dos universos, dado lo cual se ha confirmado la debilidad en el estudio de esta categoría por parte de los grupos 2, 3, 6 y 8. Mientras que, los grupos 1, 4, 7 y 9 han sido relativamente homogéneos al estudiar esta categoría, aportando un total de 3 símbolos únicos relevantes, con lo cual el UdeD₂ tiene un nivel medio de detalle. Podemos confirmar que el Grupo 5 profundizó el estudio de esta categoría.

Una observación con respecto a la dispersión de los símbolos en esta categoría es que del total de 19 símbolos, hubo solo 2 símbolos vistos por todos los grupos, aunque 4 símbolos más vistos por 6 o 7 grupos, mientras que 8 símbolos fueron vistos por 1 o 2 grupos (gran parte encontrados por el Grupo 5).

5.3. Análisis semántico para la categoría Administración de Plan

En la categoría Administración de Plan, se tenía estadísticamente la visión de un universo observado por ocho grupos y otro universo observado por 2 grupos. Luego del análisis semántico realizado en base a la denotación y connotación de los símbolos del LEL referidos a esta categoría (ver Tabla 10), se determinó que el Grupo 5 observó 2 símbolos no relevantes, a pesar de ello persistió en tener uno de los valores más

altos de símbolos relevantes, junto con el Grupo 9, el cual además observó 2 símbolos relevantes únicos. Luego del análisis semántico, se mantuvo la misma distribución de grupos, observándose dos universos. Se aplicó el análisis estadístico para comparar universos (ver Tabla 11) y poder establecer lo observado.

Tabla 10. Administración de Plan: resultados del análisis semántico

Administración de Plan	Total de Símbolos (Estadísticos)	Total de Símbolos Relevantes (Semánticos)	Símbolos No Relevantes	Nombres de Símbolos Inexistentes	Símbolos Únicos	Símbolos Únicos Relevantes
UdeD₁						
Grupo 1	4	4	0	0	0	0
Grupo 2	4	3	1	1	1	0
Grupo 3	5	2	3	1	1	0
Grupo 4	6	6	0	0	0	0
Grupo 6	4	2	2	2	2	0
Grupo 7	9	6	1	2	1	0
Grupo 8	2	2	0	0	0	0
UdeD₂						
Grupo 5	13	11	2	0	0	0
Grupo 9	13	12	0	0	2	2
Total	60	48	9	6	7	2

Total de Símbolos Sin Repetición	21
Total de Símbolos Sin Repetición No Relevantes	7
Total de Símbolos Sin Repetición Relevantes	14

En base a los resultados mostrados en la Tabla 11, se puede concluir que el UdeD₁ y el UdeD₂ presentan una diferencia estructural con un nivel de significación de 0,05 (probabilidad de diferencia estructural mayor a 99,95%).

Tabla 11. Administración de Plan: Probabilidad de diferencia estructural considerando símbolos relevantes

k	Grupos	UdeD _k	n _k	Media	Varianza	Desviación Estándar	t-Student	p-valor
1	1,2, 3, 4, 6, 7, 8	{4, 3, 2, 6, 2, 6, 2}	7	3,75	3,28	1,813	5,819	<0,0005
2	5, 9	{11, 12}	2	11,50	0,50	0,707		

A través del análisis semántico no se ha modificado la situación evaluada estadísticamente. El universo observado por los Grupos 5 y 9 contiene la mayor cantidad de símbolos relevantes, mientras que el universo observado por los restantes grupos presenta escasos símbolos relevantes: un promedio de 3,75 símbolos por grupo, no aportaron en conjunto ningún símbolo único relevante, y el grupo que más detectó encontró menos de la mitad del total símbolos encontrados en conjunto. Es decir, podemos afirmar que los Grupos 5 y 9 profundizaron el estudio de la categoría Administración de Plan, quienes detectaron en promedio el 82% de símbolos del total de 14 símbolos relevantes (sin repetición) hallados por todos los grupos. Esto comprueba lo “presuntamente” concluido estadísticamente para los Grupos 5 y 9.

Una observación interesante del análisis semántico de esta categoría es que en total definieron 6 símbolos cuyos nombres no existían en el UdeD, siendo 5 de ellos no relevantes. Un tercio del total de símbolos detectados (sin repetición) fueron calificados como no relevantes.

Respecto a la dispersión de símbolos identificados por los grupos, hubo solo 2 símbolos en común detectados por todos los grupos, 4 símbolos fueron detectados por 5, 4 o 3 grupos y la mayoría de los símbolos (57%) fueron detectados por solo 1 o 2 grupos (principalmente los Grupos 5 y 9).

5.4. Análisis semántico para la categoría Adjudicación

En el estudio estadístico de la categoría Adjudicación se habían observado dos universos diferentes (ver Figura 8). El análisis semántico realizado se resume en la Tabla 12, donde en base al total de símbolos relevantes se construyó la Figura 12 (histograma por temas) donde se visualiza una distribución aproximadamente normal con pequeñas diferencias entre ellos, lo que pareciera indicar que se ha observado un único universo.

Tabla 12. Adjudicación: resultados del análisis semántico

Adjudicación	Total de Símbolos (Estadísticos)	Total de Símbolos Relevantes (Semánticos)	Símbolos No Relevantes	Nombres de Símbolos Inexistentes	Símbolos Únicos	Símbolos Únicos Relevantes
Grupo 1	13	11	1	1	1	1
Grupo 2	6	6	0	0	0	0
Grupo 3	6	5	1	0	0	0
Grupo 4	7	7	0	0	0	0
Grupo 5	10	9	1	0	1	1
Grupo 6	7	4	3	0	2	0
Grupo 7	10	11	0	0	5	5
Grupo 8	7	5	2	0	1	0
Grupo 9	12	10	2	0	2	1
Total	78	68	10	1	12	8

Total de Símbolos Sin Repetición	26
Total de Símbolos Sin Repetición No Relevantes	6
Total de Símbolos Sin Repetición Relevantes	20

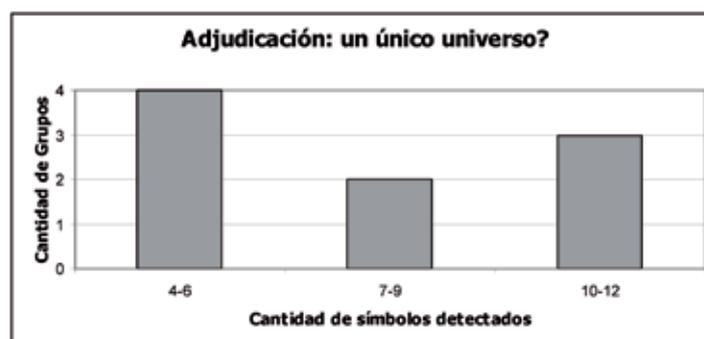


Figura 12. Distribución de grupos para Adjudicación post-análisis semántico

Aplicando la estadística chi-cuadrado observamos la existencia de un único universo (ver Tabla 13). Se aplicó la fórmula (15), dando el valor $PG_2 = 12,50\%$ que se utiliza en el cálculo de las frecuencias esperada (E_{ij}). Estas dieron todas mayores a 5, dado lo cual se puede aplicar la prueba chi-cuadrado, fórmula (14).

Tabla 13. Adjudicación: Aplicación de χ^2 cuadrado post-análisis semántico

Adjudicación	Cantidad de símbolos observados	Complemento de cantidad de símbolos	Total de símbolos	Porcentaje de Observados / Total	Frecuencia esperada	Complemento de Frecuencia esperada	χ^2
Grupo 1	11	40	68	16,18%	8,50	42,50	0,74
Grupo 2	6	45	68	8,82%	8,50	42,50	0,74
Grupo 3	5	46	68	7,35%	8,50	42,50	1,44
Grupo 4	7	44	68	10,29%	8,50	42,50	0,26
Grupo 5	9	42	68	13,24%	8,50	42,50	0,03

Grupo 6	4	47	68	5,88%	8,50	42,50	2,38
Grupo 7	11	40	68	16,18%	8,50	42,50	0,74
Grupo 8	5	46	68	7,35%	8,50	42,50	1,44
Grupo 9	10	41	68	14,71%	8,50	42,50	0,26
Total	68	391	544				8,029

Como se muestra en la Tabla 13, se obtuvo el valor $\chi^2_2 = 8,029$ para la categoría Adjudicación considerando sólo símbolos relevantes. Dado que $\chi^2_2 = 8,029 < 15,51$ (de la Tabla chi-cuadrado en Anexo B para el nivel de significación 5% y 8 grados de libertad), podemos concluir que todos los datos obtenidos de esta categoría están asociados y, entonces, Por ende, los grupos han observado para la categoría Adjudicación el mismo UdeD.

En esta categoría del total de símbolos relevantes (sin repetición), el máximo que encontró un grupo fueron 11 símbolos (correspondió tanto al Grupo 1 y como al Grupo 7), lo que representa el 55% del total de símbolos. Es decir, hubo una gran dispersión en el total de símbolos encontrados por los grupos: 1 símbolo en común, 4 símbolos detectados por 8 o 7 grupos, y la gran mayoría detectado por 1 o 2 grupos (el 60% de los símbolos) siendo detectados por distintos grupos. Cabe destacar que el Grupo 7 aportó 5 símbolos únicos relevantes.

5.5. Análisis semántico para la categoría Adhesión

Para la categoría Adhesión se tenía, en base al análisis estadístico, que todos los grupos habían observado el mismo universo. El análisis semántico para esta categoría se resume en la Tabla 14, donde se observa una distribución aproximadamente normal de símbolos relevantes entre grupos con pequeñas diferencias entre ellos, lo que presume un único universo observado. Dado lo cual se aplicó la estadística chi-cuadrado.

Tabla 14. Adhesión: resultados del análisis semántico

Adhesión	Total de Símbolos (Estadísticos)	Total de Símbolos Relevantes (Semánticos)	Símbolos No Relevantes	Nombres de Símbolos Inexistentes	Símbolos Únicos	Símbolos Únicos Relevantes
Grupo 1	7	6	1	0	2	1
Grupo 2	4	4	0	0	0	0
Grupo 3	8	6	2	0	2	1
Grupo 4	6	6	0	1	0	0
Grupo 5	8	7	1	0	2	2
Grupo 6	8	4	4	1	4	0
Grupo 7	8	7	0	2	1	1
Grupo 8	4	4	0	0	0	0
Grupo 9	7	7	0	0	1	1
Total	60	51	8	4	12	6

Total de Símbolos Sin Repetición	20
Total de Símbolos Sin Repetición No Relevantes	7
Total de Símbolos Sin Repetición Relevantes	13

En la Tabla 15 se muestra el resultado de la prueba chi-cuadrado observando sólo los símbolos relevantes. Se utilizó la fórmula (15), dando el valor $PG_1 = 11,11\%$, para el cálculo de la frecuencia esperada.

Tabla 15. Adhesión: Aplicación de χ^2 cuadrado post-análisis semántico

Adhe- sión	Cantidad de símbolos relevantes observados	Complemento de cantidad de símbolos	Total de símbolos relevantes	Porcentaje de Observados / Total	Frecuencia esperada	Complemento de Frecuencia esperada	χ^2
Grupo 1	6	45	51	11,76%	5,67	45,33	0,02
Grupo 2	4	47	51	7,84%	5,67	45,33	0,49
Grupo 3	6	45	51	11,76%	5,67	45,33	0,02
Grupo 4	6	45	51	11,76%	5,67	45,33	0,02
Grupo 5	7	44	51	13,73%	5,67	45,33	0,31
Grupo 6	4	47	51	7,84%	5,67	45,33	0,49
Grupo 7	7	44	51	13,73%	5,67	45,33	0,31
Grupo 8	4	47	51	7,84%	5,67	45,33	0,49
Grupo 9	7	44	51	13,73%	5,67	45,33	0,31
Total	51	408	459				2,47

Se puede aplicar la estadística chi-cuadrado dado que todas las frecuencias esperadas obtenidas son mayores a 5. Dado que el valor calculado $\chi^2_1 = 2,47$ es menor que el valor de $\chi^2 = 15,51$ obtenido de la Tabla correspondiente (ver Anexo B) para el nivel de significación 5% y 8 grados de libertad, entonces no se rechaza la igualdad. Por ende, podemos concluir que los grupos han observado el mismo universo para la categoría Adhesión considerando los símbolos relevantes.

En esta categoría se identificaron en conjunto un 35% de símbolos no relevantes del total detectado (20 símbolos sin repetición). Por otro lado, cada grupo identificó en promedio 6 símbolos, esto es menos del 50% del total de símbolos relevantes detectados en conjunto. En la Figura 13 se graficaron los nueve grupos con los símbolos del LEL detectados por cada uno, identificando símbolos relevantes y no relevantes. Se puede observar visualmente que hubo 4 símbolos compartidos por todos los grupos. Hubo 2 grupos que identificaron solo estos 4 símbolos: Grupos 2 y 8. El Grupo 9 también solo identificó estos 4 símbolos, más otros 4 símbolos no relevantes.

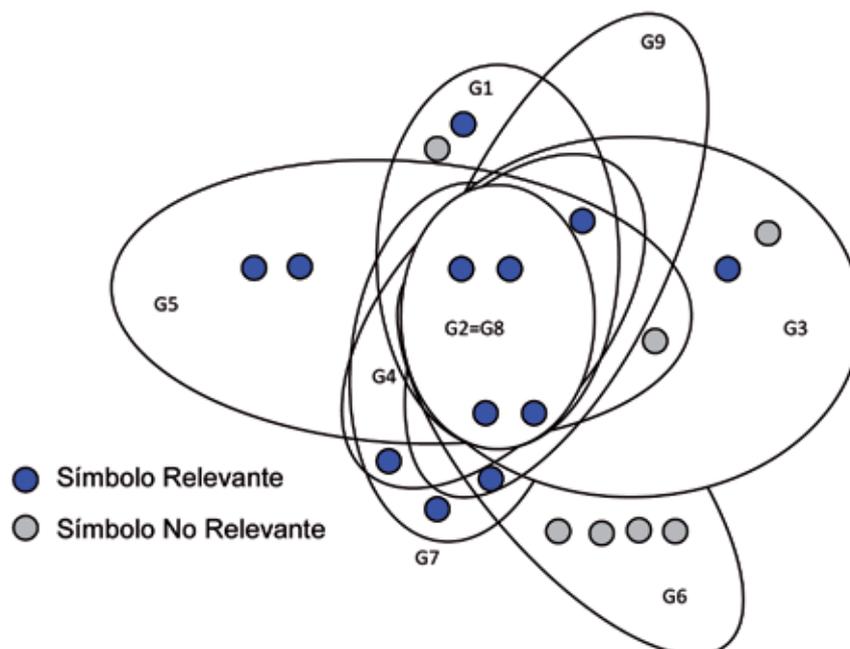


Figura 13. Distribución visual de símbolos por grupos para Adhesión

Cabe aclarar que hubo 8 símbolos detectados por 1 o 2 grupos todos diferentes, representando el 62 % del total de símbolos relevantes. Es decir, se aprecia una dispersión interesante en la identificación de símbolos para esta categoría.

6. Observaciones post análisis semántico

Respecto a las observaciones netamente estadísticas iniciales, podemos confirmar ahora en base al análisis semántico y posterior estudio estadístico realizado que:

- El Grupo 5 profundizó en el estudio de las categorías Gestión de Cuotas, Administración de Bien Tipo y Administración de Plan.
- El Grupo 9 profundizó en el estudio de la categoría Administración de Plan.
- Los Grupos 2, 3, 6 y 8 estudiaron débilmente la categoría Administración de Bien Tipo.
- Los Grupos 1, 4 y 7 elicitaron con un mismo nivel de detalle en todas las categorías.
- Las categorías Adhesión y Adjudicación fueron vistas con un nivel de detalle homogéneo por todos los grupos, es decir observaron el mismo UdeD. Este resultado ya se había obtenido con el estudio estadístico preliminar para la categoría Adhesión, no así para Adjudicación, donde se habían percibido dos universos diferentes de este sub-problema.
- Para las otras tres categorías, una vez aplicado el análisis semántico sobre los símbolos y rehaciendo el estudio estadístico, se confirmó que los grupos observaron distintos universos en cada sub-problema. Esto está en concordancia con la hipótesis de trabajo planteada.

Otras observaciones que podemos mencionar en base al análisis semántico y posterior estudio estadístico realizado son:

- Se puede afirmar desde una perspectiva semántica que el Grupo 6 observó débilmente todas las categorías, pues no aportó ningún símbolo único relevante y en todas las categorías siempre identificó menos símbolos que el resto de los grupos (considerando exclusivamente símbolos relevantes). Del total de 26 elicitados por este grupo, solo tuvo 14 símbolos relevantes. Además, identificó 7 símbolos cuyos nombres no existían en el UdeD. Es decir, en general tuvo una visión diferente del UdeD dada una captura de símbolos muy distorsionada frente al universo que conocemos.
- En el otro extremo, el Grupo 5 fue el que identificó más símbolos relevantes en todas las categorías, excepto en Adjudicación aunque allí elicitó una cantidad por encima de la media. Del total de 67 símbolos identificados por este grupo, solo 5 fueron calificados como no relevantes. Este grupo encontró el 62% de símbolos del total identificados (relevantes) entre todos los grupos.
- Aún cuando este Grupo 5 puede ser calificado como el mejor grupo de elicitadores por detectar la mayor cantidad de símbolos distintos relevantes, le han faltado encontrar una buena cantidad de símbolos reconocidos por los otros grupos. Esto podría indicar un nivel de incompletitud en base a los símbolos reconocidos.
- Se observó que en Adhesión y en Adjudicación hubo mucha dispersión en cuanto a la cantidad de símbolos detectados por los grupos. En general, más del 60% de los símbolos en cada una de estas categorías fue encontrada por uno o dos grupos, mientras que tuvieron 31% y 25% respectivamente de símbolos comunes encontrados por la mayoría de los grupos (estos valores son muy superiores a los alcanzados por otras categorías). Es decir, hubo un fuerte núcleo común de símbolos.
- Del total de símbolos distintos relevantes (97), hubo 9 símbolos comunes a todos los grupos y 10 símbolos detectados en común por 7 u 8 grupos diferentes. Es decir, se determinó que hubo casi un 20% de símbolos comunes detectados entre 7 y 9 grupos, valor que puede considerarse relativamente bajo.
- Se comprobó semánticamente la existencia de 10 símbolos cuyos nombres no pertenecían al UdeD, pero el significado (denotación y connotación) de 4 de ellos era relevante, es decir, debería estar contenida en otros símbolos. Efectivamente, esa información estaba contenida en otros símbolos de otros grupos.
- Se identificaron en total 32 símbolos no relevantes de un total de 129 símbolos distintos identificados entre todos los grupos. Es decir, se descartó un 25% de los símbolos hallados por los grupos para realizar el estudio estadístico sobre la visión de los universos.

Luego del análisis semántico hemos vuelto a confirmar la hipótesis de trabajo planteada, pues se ha podido constatar que varios grupos observaron distintos UdeD para una misma categoría. Esto se dio en tres de las cinco categorías establecidas.

7. Conclusiones y Trabajos Futuros

Basados en el trabajo desarrollado por Doorn y Ridao [Doorn 03] sobre el grado de completitud del modelo léxico extendido del lenguaje, se analizaron los resultados estadísticos que habían obtenido aplicando DPM [Wohlin 98], y se sospechó que los resultados no serían los mismos si se aplicaba semántica al determinar los elementos observados (símbolos del LEL).

Según este estudio [Doorn 03], habría faltado encontrar aproximadamente 9 términos del LEL considerando el total de términos distintos encontrados entre todos los grupos. Ahora bien, si se considera al grupo que más símbolos detectó por sí solo (Grupo 5 detectó 67 símbolos), a este grupo le estaría faltando encontrar 60 símbolos, es decir, habría alcanzado un nivel de completitud del 53%. Este valor parecería indicar que este grupo construyó un modelo con notorio exceso de elementos omitidos.

Sobre esta base, se elaboraron algunas hipótesis que pudieran dar una explicación a tan bajo nivel de completitud alcanzado en forma independiente por los grupos de ingenieros de requisitos. Entonces se planteó la hipótesis “Los grupos de ingenieros tratan problemas diferentes”. Esto indicaría que no hay un exceso de términos omitidos sino que los grupos independientes tuvieron visiones muy disímiles de ese UdeD.

Las diferencias obtenidas por los grupos respecto a la cantidad de elementos observados del mundo real podrían provenir de la visión que cada grupo tenía sobre ese mundo real. Esas diferentes visiones podrían ocurrir por manejar distintos límites en el mundo real o por estudiar con distinto grado de detalle el mundo real.

Para comprobar esta hipótesis, se estudió estadísticamente la visión que tenía cada grupo de elicidores haciendo una posterior corrección semántica para comparar las distintas visiones. Para facilitar las comparaciones entre grupos, se dividió el problema en 5 categorías de sub-problemas.

Entonces, inicialmente se aplicó un estudio netamente estadístico considerando exclusivamente los nombres de los símbolos y sus sinónimos expresamente identificados por nombres idénticos o gramaticalmente idénticos (flexiones verbales, plurales, formas sustantivas para verbos). Este fue el mismo criterio aplicado en [Doorn 03] para identificar el conjunto de símbolos distintos detectados por todos los grupos.

Luego, se estudiaron las denotaciones y connotaciones de cada símbolo, para establecer efectivamente los sinónimos y los símbolos realmente importantes del UdeD. De la cantidad inicial de 129 símbolos diferentes identificados en conjunto por los nueve grupos, se determinó que se habían identificado en total 97 símbolos diferentes relevantes. La diferencia de 32 símbolos correspondía a términos cuyo nombre solo sería necesario mencionar en otros símbolos o cuyo contenido podría incluirse en otros símbolos, y un pequeño número correspondía a nombres de términos inexistentes en el UdeD.

Se pudo establecer estadísticamente, analizando adicionalmente la semántica de los elementos observados, que los grupos tuvieron distintas percepciones del dominio del problema según las categorías de sub-problemas.

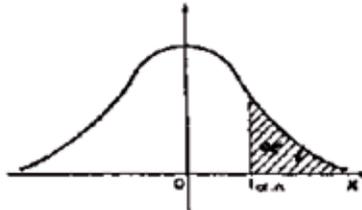
Además, se observó que aplicando el método de captura y recaptura para estimar el grado de completitud de un modelo de requisitos escrito en lenguaje natural, sin evaluar inicialmente la semántica de los elementos del modelo puede provocar una distorsión importante en los resultados obtenidos. Pues en el caso del modelo LEL estudiado, luego del análisis semántico, se descartaron el 25% de los elementos observados (términos léxicos) para realizar el estudio estadístico.

Se espera refinar el análisis semántico realizado sobre el modelo léxico, elaborando con mayor precisión criterios semánticos de comparación. Esto permitirá definir algunas heurísticas para mejorar la construcción del LEL, principalmente cómo determinar si un símbolo es parte del UdeD o no, es decir, esto llevará a precisar mejor los límites del UdeD.

8. Referencias

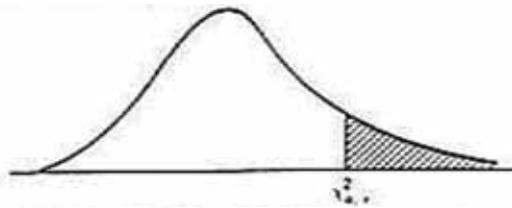
- [Barnett 94] Barnett, V., Lewis, T.: *Outliers in Statistical Data*. John Wiley & Sons, 3ª edición. ISBN 0471930946 (1994)
- [Briand 00] Briand, L., El Emam, K., Freimut, B., Laitenberger, O.: A Comprehensive Evaluation of Capture-Recapture Models for Estimating software Defects Contents. *IEEE TSE*, 26(6):518--540 (2000)
- [Cavazos 88] Cavazos, G.: *Probabilidades y Estadística. Aplicaciones y métodos*, 1ª edición. Mc. Graw Hill. ISBN: 968-451-856-0 (1988)
- [Doorn 03] Doorn, J.H., Ridao, M.: Completitud de Glosarios: Un Estudio Experimental. In: VI Workshop on Requirements Engineering (WER'03), Brasil, pp. 317--328 (2003)
- [Firesmith 05] Firesmith, D.: Are Your Requirements Complete? *Journal of Object Technology*, vol. 4, nº 1, pp. 27--43. doi:10.5381/jot.2005.4.1.c3 (2005)
- [Goguen 93] Goguen, J.A., Linde, Ch.: Techniques for Requirements Elicitation. In: IEEE First International Symposium on Requirements Engineering (RE'93), IEEE Computer Society Press, Los Alamitos, CA, pp.152--164 (1993)
- [Hadad 12] Hadad, G.D.S., Litvak, C.S., Doorn, J.H.: Agregando semántica a técnicas de predicción de completitud en modelos de requisitos. In: XIV Workshop de Investigadores en Ciencias de la Computación (WICC 2012), Universidad Nacional de Misiones, Posadas. ISBN: 978-950-766-082-5 (2012)
- [Jackson 95] Jackson, M.: *Software Requirements & Specifications. A lexicon of practice, principles and prejudices*. Addison-Wesley, Reading, MA/ACM Press, Nueva York (1995)
- [Kotonya 98] Kotonya, G., Sommerville, I.: *Requirements Engineering: Process and Techniques*. John Wiley & Sons (1998)
- [Leite 93] Leite, J.C.S.P., Franco, A.P.M.: A Strategy for Conceptual Model Acquisition. In: IEEE First Intl Symposium on Requirements Engineering (RE'93), IEEE Computer Society Press, EEUU, pp 243--246 (1993)
- [Leite 01] Leite JCSP (2001) Gerenciando a Qualidade de Software com Base em Requisitos. *Qualidade de Software: Teoria e Prática*, Rocha A, Maldonado J, Weber K (eds), Prentice-Hall, cap 17, pp 238-246
- [Loucopoulos 95] Loucopoulos, P., Karakostas, V.: *System Requirements Engineering*. McGraw-Hill, Londres (1995)
- [Mauco 97] Mauco, V., Ridao, M., del Fresno, M., Rivero, L., Doorn, J.H.: *Ingeniería de Requisitos, Proyecto: Sistema de Planes de Ahorro*. Reporte técnico, ISISTAN, UNICEN, Tandil (1997)
- [Otis 78] Otis, D.L., Burnham, K.P., White, G.C., Anderson, D.R.: Statistical inference from Capture on Closed Animal Populations. *Wildlife Monograph*, 62 (1978)
- [Pértega 01] Pértega Díaz, S., Pita Fernández, S.: Métodos paramétricos para la comparación de dos medias t de Student. *Unidad de Epidemiología Clínica y Bioestadística. Complejo Hospitalario Universitario de Coruña, España. CAD ATEN PRIMARIA*, vol. 8, pp.37--41 (2001)
- [Petersson 04] Petersson, H., Thelin, T., Runeson, P., Wohlin, C.: Capture-Recapture in Software Inspections after 10 Years Research - Theory, Evaluation and Application. *Journal of Software and Systems*, vol. 72, nº 2, pp.249--264 (2004)
- [Ridao 06] Ridao, M., Doorn, J.H.: Estimación de Completitud en Modelos de Requisitos Basados en Lenguaje Natural. In: IX Workshop on Requirements Engineering (WER'06), Brasil, pp. 151--158. ISSN: 1413-9014 (2006)
- [Rivero 98] Rivero, L., Doorn, J., del Fresno, M., Mauco, V., Ridao, M., Leonardi, M.C.: Una Estrategia de Análisis Orientada a Objetos basada en Escenarios: Aplicación en un Caso Real. En: Workshop en Engenharia do Requisitos (WER'98), Maringá, Brasil, pp. 79--90 (1998)
- [Walia 08] Walia, G.S., Carver, J.C.: Evaluation of Capture-Recapture Models for Estimating the Abundance of Naturally Occurring Defects. In: 2nd ACM-IEEE Intl Symposium of Empirical Software Engineering and Measurement, Alemania, pp.158--167. ISBN: 978-1-59593-971-5 (2008)
- [White 82] White, G.C., Anderson, D.R., Burnham, K.P., Otis, D.L.: Capture-Recapture and Removal Methods for Sampling Closed Populations. Technical Report LA-8787-NERP, Los Alamos Nat'l Laboratory (1982)
- [Wohlin 98] Wohlin, C., Runeson, P.: Defect content estimations from Review Data. In: 20th Intl Conference on Software Engineering, Japón, pp. 400--409 (1998)

Anexo A: Tabla t de Student



$\alpha/2$ df	0,40	0,30	0,20	0,10	0,050	0,025	0,010	0,005	0,001	0,0005
1	0,325	0,727	1,376	3,078	6,314	12,71	31,82	63,66	318,3	636,6
2	0,289	0,617	1,061	1,886	2,920	4,303	6,965	9,925	22,33	31,60
3	0,277	0,584	0,978	1,638	2,353	3,182	4,541	5,841	10,22	12,94
4	0,271	0,569	0,941	1,533	2,132	2,776	3,747	4,604	7,173	8,610
5	0,267	0,559	0,920	1,476	2,015	2,571	3,365	4,032	5,893	6,859
6	0,265	0,553	0,906	1,440	1,943	2,447	3,143	3,707	5,208	5,959
7	0,263	0,549	0,896	1,415	1,895	2,365	2,998	3,499	4,785	5,405
8	0,262	0,546	0,889	1,397	1,860	2,306	2,896	3,355	4,501	5,041
9	0,261	0,543	0,883	1,383	1,833	2,262	2,821	3,250	4,297	4,781
10	0,260	0,542	0,879	1,372	1,812	2,228	2,764	3,169	4,144	4,587
11	0,260	0,540	0,876	1,363	1,796	2,201	2,718	3,106	4,025	4,437
12	0,259	0,539	0,873	1,356	1,782	2,179	2,681	3,055	3,930	4,318
13	0,259	0,538	0,870	1,350	1,771	2,160	2,650	3,012	3,852	4,221
14	0,258	0,537	0,868	1,345	1,761	2,145	2,624	2,977	3,787	4,140
15	0,258	0,536	0,866	1,341	1,753	2,131	2,602	2,947	3,733	4,073
16	0,258	0,535	0,863	1,337	1,746	2,120	2,583	2,921	3,686	4,015
17	0,257	0,534	0,863	1,333	1,740	2,110	2,567	2,898	3,646	3,965
18	0,257	0,534	0,862	1,330	1,734	2,101	2,552	2,878	3,611	3,922
19	0,257	0,533	0,861	1,328	1,729	2,093	2,539	2,861	3,579	3,883
20	0,257	0,533	0,860	1,325	1,725	2,086	2,528	2,845	3,552	3,850
21	0,257	0,532	0,859	1,323	1,721	2,080	2,518	2,831	3,527	3,819
22	0,256	0,532	0,858	1,321	1,717	2,074	2,508	2,819	3,505	3,792
23	0,256	0,532	0,858	1,319	1,714	2,069	2,500	2,807	3,485	3,767
24	0,256	0,531	0,857	1,318	1,711	2,064	2,492	2,797	3,467	3,745
25	0,256	0,531	0,856	1,316	1,708	2,060	2,485	2,787	3,450	3,725
26	0,256	0,531	0,856	1,315	1,706	2,056	2,479	2,779	3,435	3,707
27	0,256	0,531	0,855	1,314	1,703	2,052	2,473	2,771	3,421	3,690
28	0,256	0,530	0,855	1,313	1,701	2,048	2,467	2,763	3,408	3,674
29	0,256	0,530	0,854	1,311	1,699	2,045	2,462	2,756	3,396	3,659
30	0,256	0,530	0,854	1,310	1,697	2,042	2,457	2,750	3,385	3,646
40	0,255	0,529	0,851	1,303	1,648	2,021	2,423	2,704	3,307	3,551
50	0,255	0,528	0,849	1,298	1,676	2,009	2,403	2,678	3,262	3,495
60	0,254	0,527	0,848	1,296	1,671	2,000	2,390	2,660	3,232	3,460
80	0,254	0,527	0,846	1,292	1,664	1,990	2,374	2,639	3,195	3,415
100	0,254	0,526	0,845	1,290	1,660	1,984	2,365	2,626	3,174	3,389
200	0,254	0,525	0,843	1,286	1,653	1,972	2,345	2,601	3,131	3,339
500	0,253	0,525	0,842	1,283	1,648	1,965	2,334	2,586	3,106	3,310
∞	0,253	0,524	0,842	1,282	1,645	1,960	2,326	2,576	3,090	3,291

Anexo B: Tabla chi-cuadrado χ^2



ν	α								
	0.995	0.990	0.975	0.950	0.500	0.050	0.025	0.010	0.005
1	0.004	0.005	0.008	0.010	0.45	3.84	5.02	6.63	7.88
2	0.010	0.015	0.020	0.025	1.39	5.99	7.38	9.21	10.60
3	0.074	0.084	0.101	0.125	2.37	7.81	9.35	11.34	12.84
4	0.215	0.237	0.270	0.312	3.36	9.49	11.14	13.28	14.86
5	0.411	0.446	0.486	0.540	4.35	11.07	12.38	15.09	16.75
6	0.688	0.717	0.756	0.809	5.35	12.59	14.45	16.81	18.55
7	0.989	1.024	1.064	1.124	6.35	14.07	16.01	18.48	20.28
8	1.344	1.380	1.420	1.476	7.34	15.51	17.53	20.09	21.96
9	1.735	1.771	1.811	1.867	8.34	16.92	19.02	21.67	23.59
10	2.160	2.197	2.237	2.293	9.34	18.31	20.48	23.21	25.19
11	2.603	2.640	2.680	2.736	10.34	19.68	21.92	24.72	26.76
12	3.078	3.115	3.155	3.211	11.34	21.03	23.34	26.22	28.30
13	3.572	3.609	3.649	3.705	12.34	22.36	24.74	27.69	29.82
14	4.076	4.113	4.153	4.209	13.34	23.68	26.12	29.14	31.32
15	4.600	4.637	4.677	4.733	14.34	25.00	27.49	30.58	32.80
16	5.143	5.180	5.220	5.276	15.34	26.30	28.85	32.00	34.27
17	5.705	5.742	5.782	5.838	16.34	27.59	30.19	33.41	35.72
18	6.276	6.313	6.353	6.409	17.34	28.87	31.53	34.81	37.16
19	6.854	6.891	6.931	6.987	18.34	30.14	32.85	36.19	38.58
20	7.438	7.475	7.515	7.571	19.34	31.41	34.17	37.57	40.00
25	10.52	10.59	10.66	10.73	24.34	37.65	40.65	44.31	46.93
30	13.78	13.87	13.95	14.03	29.34	43.77	46.98	50.89	53.67
40	20.71	20.81	20.90	21.00	39.34	55.76	59.34	63.69	66.77
50	27.99	28.10	28.19	28.29	49.33	67.50	71.42	76.15	79.49
60	35.53	35.65	35.76	35.87	59.33	79.08	83.30	88.38	91.95
70	43.28	43.41	43.52	43.63	69.33	90.53	95.02	100.42	104.22
80	51.17	51.30	51.41	51.52	79.33	101.88	106.63	112.33	116.32
90	59.20	59.33	59.44	59.55	89.33	113.14	118.14	124.12	128.30
100	67.33	67.46	67.57	67.68	99.33	124.34	129.56	135.81	140.17

