

SEBASTIAN M. RASINGER

# La investigación cuantitativa en Lingüística

---

## Una introducción

Traducción de Ana Useros Martín y Gema Sanz Espinar



**akai**

ARGENTINA

ESPAÑA

MÉXICO

# Índice general

---

<b>Agradecimientos</b> .....	<b>6</b>
<b>Figuras y tablas</b> .....	<b>11</b>
<b>1. Introducción</b> .....	<b>15</b>

## PRIMERA PARTE

<b>2. La investigación cuantitativa: temas básicos</b> .....	<b>23</b>
2.1. Datos cualitativos y cuantitativos .....	23
2.2. Variables y mediciones .....	32
2.3. Definiciones, conceptos y operacionalización .....	34
2.4. Variables independientes y dependientes .....	37
2.5. Causalidad y variables ocultas .....	38
2.6. Escalas de medición .....	40
Escala categórica o nominal .....	40
Escala ordinal .....	41
Escala de intervalo .....	41
Escala de razón .....	41
2.7. Datos continuos y discretos .....	42
2.8. Fiabilidad y validez .....	43
2.9. Hipótesis, leyes y teorías .....	46
<b>3. Diseño y muestreo de la investigación</b> .....	<b>53</b>
3.1. Diseños transversales .....	53
3.2. Diseños longitudinales: paneles y cohortes .....	54
3.3. Diseños pseudolongitudinales: tiempo real y tiempo aparente .....	57
3.4. Diseños experimentales y cuasi experimentales .....	58
3.5. Resumen: diseños de investigación .....	61
3.6. Muestreo .....	62

3.7. Un apunte sobre ética .....	69
3.8. Gestionar la investigación: levantar un proyecto (y mantenerlo en pie) .....	73
<b>4. Diseño de cuestionario y codificación .....</b>	<b>80</b>
4.1. Punto de partida: saber lo que se quiere y preguntar lo que se necesita.....	81
4.2. Preguntas de selección múltiple/de opción múltiple .....	83
4.3. Medición de conceptos abstractos: actitudes y creencias .....	85
4.4. Formular las preguntas .....	88
4.5. Prueba piloto.....	92
4.6. Presentación del cuestionario .....	94
4.7. Número y secuencia de las preguntas .....	95
4.8. Codificación.....	96
4.9. Un cuestionario “de verdad” .....	100
Ejemplo de cuestionario.....	106
4.10. Resumen: cuestionarios .....	108
<b>Primera parte. Para leer más.....</b>	<b>109</b>
Lecturas generales .....	109
Acerca de los cuestionarios/encuestas.....	110
Sobre ética .....	110
Sobre gestión de proyectos.....	110

## SEGUNDA PARTE

<b>5. Un primer vistazo a los datos .....</b>	<b>113</b>
5.1. Vamos paso a paso: sumas.....	114
5.2. Frecuencias absolutas y relativas .....	115
5.3. Clases, anchura de intervalo y frecuencias acumuladas.....	119
5.4. Visualización: diagramas y gráficos .....	124
5.5. Resumen .....	135
<b>6. Describir adecuadamente los datos: centralización y dispersión .....</b>	<b>137</b>
6.1. Estadísticas: ¿qué y por qué? .....	137
6.2. Medidas de centralización .....	140
6.2.1 La media aritmética y la media aritmética acotada .....	140
6.2.2 La mediana .....	145
6.2.3. La moda.....	147
6.2.4 Cuartiles, quintiles y percentiles.....	148
6.2.4. Resumen: media, mediana y moda .....	150
6.3. Medidas de dispersión.....	151
6.3.1 Rango .....	152
6.3.2 Varianza y desviación típica.....	153
6.3.3 Unidades tipificadas .....	155
6.4. La distribución normal .....	157
6.5. El error estándar .....	162
6.6. Resumen: estadística descriptiva.....	165

<b>7. Análisis de los datos – avanzando un poco más.....</b>	<b>168</b>
7.1 Probabilidad: nociones básicas.....	168
7.2 Probabilidades: más de un suceso.....	170
7.3. La prueba $\chi^2$ .....	174
7.4. La correlación de Pearson.....	178
7.5 Correlación parcial.....	185
7.6. Causalidad.....	188
7.7. Un apunte acerca de la significancia.....	189
7.8. Regresión lineal simple: pronósticos.....	191
7.9. Regresión lineal múltiple.....	198
7.10 Correlación y fiabilidad.....	201
<b>8. Comprobación de hipótesis.....</b>	<b>205</b>
8.1. Hipótesis, causalidad y colas.....	205
8.2. La prueba <i>t</i> . Preliminares.....	208
8.3. La prueba <i>F</i> .....	210
8.4. Prueba <i>t</i> para muestras independientes.....	212
8.5 Prueba <i>t</i> dependiente.....	216
8.6. Comprobación de hipótesis, pruebas <i>t</i> y datos no experimentales.....	220
8.7 Comprobación de hipótesis con la prueba $\chi^2$ .....	222
8.8. Tres son multitud: ANOVA.....	225
8.8.1 ANOVA de un factor/unifactorial.....	225
8.8.2 ANOVA de dos factores.....	229
<b>9. Análisis de datos no paramétricos: cuando las cosas no son del todo normales.....</b>	<b>233</b>
9.1. La prueba del coeficiente de correlación de Spearman.....	235
9.2. La <i>tau</i> de Kendall.....	239
9.3. Prueba de los rangos con signo de Wilcoxon.....	239
9.4. Prueba U de Mann-Whitney.....	243
<b>Segunda parte. Para leer más.....</b>	<b>245</b>

### TERCERA PARTE

<b>10. Nivel avanzado: otros métodos y herramientas.....</b>	<b>249</b>
10.1. Interpretar MANOVA: una guía rápida.....	249
10.2 Usar los datos existentes: una introducción al metaanálisis.....	252
10.2.1. El reciclaje y la búsqueda del efecto “verdadero”.....	252
10.2.2. Metaanálisis. Primeros pasos.....	252
10.2.3. Metaanálisis. La parte matemática.....	255
10.2.4. Metaanálisis. Resumen.....	259
Para leer más.....	260
10.3. Balance de Excel y otras alternativas.....	261
10.3.1 SPSS.....	261

Para leer más .....	263
10.3.2 R .....	263
Para leer más .....	265
<b>Capítulo 11. Apéndices y soluciones .....</b>	<b>266</b>
11.1. Fórmulas de Excel .....	266
11.2. Valores críticos para $\chi^2$ .....	268
11.3. Valores críticos para la correlación de Pearson .....	268
11.4. Valores críticos para la $\rho$ de Spearman .....	270
11.5. Valores críticos para la prueba de Wilcoxon .....	271
11.6. Valores críticos para la prueba U Mann-Whitney .....	271
11.7. Soluciones para los ejercicios (no resueltos en los capítulos correspondientes) .....	274
11.7.1. Capítulo 2 .....	274
11.7.2. Capítulo 5 .....	274
11.7.3. Capítulo 6 .....	276
11.7.4. Capítulo 7 .....	277
11.7.5. Capítulo 8 .....	280
11.7.6. Capítulo 9 .....	284
<b>Bibliografía.....</b>	<b>285</b>

Hay cosas que no cambian: el sol sale por la mañana, el sol se pone por la tarde. Muchos lingüistas odian las estadísticas. Llevo una década enseñando lengua y lingüística inglesa en diversas universidades e, invariablemente, me encuentro con una aversión generalizada a todo lo que sea numérico. La queja más habitual es que la estadística es “¡taaan difícil!”. Pero no lo es. O, al menos, no tendría por qué serlo.

Este libro quiere ser varias cosas a la vez. No debería sorprender a nadie que sea una introducción a los métodos de investigación cuantitativos. Como tal es un silabario: un libro que se dirige especialmente a aquellos cuyo conocimiento de la materia sea mínimo. De hecho, la mayoría de quienes lean este libro, deberían poder seguirlo sin ningún conocimiento previo o familiaridad con sus contenidos. Es también una guía, que nos conducirá en el proceso, en ocasiones abrumador, de articular y llevar a cabo un estudio cuantitativo desde el principio hasta el final. Y, en último lugar, es un manual y, como tal, proporciona consejos prácticos sobre la investigación cuantitativa en general y sobre los métodos estadísticos en particular.

La idea original que impulsó la primera edición de este libro, publicada en 2008, surgió tras un debate con algunos de mis estudiantes de máster. Después de hablar de los factores que influían en el éxito o fracaso de la adquisición de un segundo idioma, empezamos a centrarnos en el tema de la motivación del aprendizaje y acabé preguntando a) cómo podríamos demostrar empíricamente la motivación del aprendiz y b) cómo podríamos concluir con fundamento que la motivación y los resultados estaban causalmente relacionados. El resultado fue desalentador. Varios estudiantes, de un grupo compuesto en su totalidad por docentes de idiomas con una experiencia de varios años, argumentaron que “un buen profesor puede ver que esto es así”. Y, finalmente, una de ellas me acusó de tratar de medirlo todo: “Tienes una mente demasiado

científica”, me decía. Me fue preciso emplear métodos no demasiado amables para convencerlos de que cualquier trabajo académico sólido debería inevitablemente emplear métodos y técnicas metodológicamente adecuadas, ¡o de lo contrario...! Han pasado cinco años y aún sigo teniendo estas discusiones, pero al menos ahora tengo un libro que recomendar a los alumnos. ¡La investigación cuantitativa no tiene por qué ser difícil!

Los criterios de evaluación que marca la UK Quality Assurance Agency —la organización que supervisa los estándares académicos de la educación superior— para los grados de lingüística dictan que:

6.13. Para licenciarse en lingüística, los estudiantes deberán:

- demostrar una comprensión de las técnicas básicas de análisis de datos y sopesar y elegir las técnicas adecuadas, como *criterios de distribución*, análisis espectrográfico, el uso de herramientas de TI para la investigación de bases de datos informatizadas, el empleo de programas informáticos para el análisis de los fenómenos acústicos, el empleo de técnicas de laboratorio para la investigación de los fenómenos articulatorios, *pruebas estadísticas relevantes*, el empleo del material de audio y vídeo para el análisis de la interacción oral...;
- *demostrar comprensión de datos y análisis presentados mediante diagramas (incluidos diagramas de árbol), tablas, matrices y otros diagramas, así como presentar datos de manera adecuada usando estos medios con una supervisión mínima.*

(The Quality Assurance Agency for Higher Education, 2007. Linguistics. Subject Benchmarks [222.qaa.uk/publications/InformationAndaGuidance/Documents/Linguistics07.pdf, pp. 14-15]. La cursiva es mía).

Puede parecer una extravagancia empezar un manual introductorio con citas de un documento político; pero esta cita muestra lo importante que es una comprensión exhaustiva de los métodos de investigación. A lo largo de los años he calificado una buena cantidad de trabajos de estudiantes, en diversas universidades y en diversos niveles, que empezaban siendo excelentes pero que se desplomaban en cuanto llegaban a la parte del análisis de datos, en muchas ocasiones porque no se había sabido manejar adecuadamente los datos recogidos, porque los análisis eran equivocados y porque se cometían fallos en operaciones matemáticas básicas. Y, lo que es peor, algunos trabajos seguían defendiendo el argumento que habían planteado con indiferencia de si los resultados empíricos lo respaldaban o no. Hay pocas cosas peores que obtener (o dar) una mala nota por algo que podría haberse resuelto con tanta facilidad. Lo digo por experiencia.

También tenemos el caso contrario. Puede darse el caso de que no queramos emprender un estudio cuantitativo, por la razón que sea; sin embargo, eso no quiere decir que no surja la ocasión en la que tenga-

mos que asimilar las líneas básicas de una investigación cuantitativa. La mayoría de los estudios de psicolingüística y sociolingüística —especialmente los que tratan de la variación lingüística— se basan en datos y análisis cuantitativos. Si queremos comprender por completo estos estudios, tenemos al menos que hacernos una idea de lo que significan todos esos términos, índices, ilustraciones y gráficos que contiene. Puede que no sea necesario conocer todas las complejidades matemáticas que implica una prueba *t* o un Análisis de la Varianza (ANOVA) (véase Capítulo 8), pero es evidente que leer en un artículo que “una prueba *t* mostró significativas diferencias estadísticas entre los dos grupos” y no saber qué es una prueba *t* ni qué es estadísticamente significativo no es algo positivo.

Hay en el mercado, sin duda alguna, una miríada de libros acerca de los métodos cuantitativos y sobre estadística. Habitan predominantemente en las estanterías de disciplinas como las ciencias sociales o naturales, psicología o economía. Hay también, qué duda cabe, libros de estadística que atienden específicamente a la lingüística, como, por ejemplo, Butler (1985) o Wood *et al.* (1986) o, más recientemente, los manuales de Johnson (2008) o Gries (2009), por mencionar solamente unos pocos. Todos ellos son obras muy completas a las que poco se puede añadir. Pero mi experiencia como docente de estadística y métodos cuantitativos, tanto a estudiantes como a lingüistas profesionales, me ha enseñado que lo que mucha gente necesita no es únicamente una introducción a la estadística (la parte de mascar números) sino también el enfoque general de la investigación cuantitativa: con frecuencia el problema no son únicamente las matemáticas, sino toda la concepción metodológica. Y, por lo tanto, este libro no trata únicamente de estadística, sino que busca ser una introducción a la investigación cuantitativa en general, escrita específicamente para estudiantes e investigadores en lengua y lingüística.

Así pues, este libro está escrito teniendo en mente tres tipos especiales de lectores que, de un modo u otro, forman parte de la vida de su autor:

- a los colegas y estudiantes a quienes les entra el pánico cuando se les pide que ejecuten análisis estadísticos sencillos como calcular medias y, por lo tanto, se niegan en redondo a acercarse a nada remotamente relacionado con los números;
- al lingüista que, a pesar de estar versado en métodos estadísticos, insiste en hacer complicados cálculos con lápiz y papel, y se niega a usar un programa informático accesible, haciendo, por lo tanto, de la estadística un asunto aún más penoso;
- al estudiante de grado o posgrado, o cualquier otro investigador en lingüística y otras disciplinas relacionadas con el lenguaje, que se da cuenta en algún momento de que necesita una introducción global a los métodos cuantitativos en su campo para que su proyecto investigador sea metodológicamente sólido.

Este libro no exige requisitos previos, aparte de unas destrezas numéricas elementales y un conocimiento básico de cómo usar un