

Víctor M. Fernández Martínez

Arqueo-Estadística

**Métodos cuantitativos
en Arqueología**

Alianza Editorial

Reservados todos los derechos. El contenido de esta obra está protegido por la Ley, que establece penas de prisión y/o multas, además de las correspondientes indemnizaciones por daños y perjuicios, para quienes reprodujeran, plagiaran, distribuyeran o comunicaran públicamente, en todo o en parte, una obra literaria, artística o científica, o su transformación, interpretación o ejecución artística fijada en cualquier tipo de soporte o comunicada a través de cualquier medio, sin la preceptiva autorización.

© Víctor M. Fernández Martínez, 2015

© Alianza Editorial, S. A., 2015

Calle Juan Ignacio Luca de Tena, 15; 28027 Madrid; telef. 91 393 88 88

www.alianzaeditorial.es

ISBN: 978-84-206-9761-1

Depósito legal: M. 186-2015

Printed in Spain

SI QUIERE RECIBIR INFORMACIÓN PERIÓDICA SOBRE LAS NOVEDADES DE ALIANZA EDITORIAL,
ENVÍE UN CORREO ELECTRÓNICO A LA DIRECCIÓN:

alianzaeditorial@anaya.es

Índice

Prólogo	7
1. Introducción	9
1. ¿Por qué necesitamos la estadística?.....	9
2. Un ejemplo	12
3. El problema de las muestras	16
4. Comenzar a entender la probabilidad y la inferencia	20
5. Los programas informáticos	21
6. Estadística y teoría arqueológica	24
2. Las variables arqueológicas y su cuantificación.....	29
1. ¿Qué medimos y contamos? Objetos, datos y variables	29
2. Variables cualitativas	30
3. Variables cuantitativas: ordinales, de intervalo y de razón	35
4. Validez de los datos: relación entre lo que medimos y lo que buscamos.....	44
5. Ejercicios	48
3. Estadística descriptiva univariante I: gráficos y Análisis Exploratorio de Datos	49
1. Gráficos para variables cualitativas y ordinales	49
2. Agrupando datos numéricos: primeros pasos	50
3. Histogramas	53
4. Diagramas de frecuencias acumuladas.....	59
5. Cuartiles y diagramas de caja	62
6. El Análisis Exploratorio de Datos	66
7. Ejercicios	67

4. Estadística descriptiva univariante II: tendencia central y dispersión.	
La curva normal.....	69
1. Medidas de la centralidad: media, mediana, moda	69
2. Medidas de la dispersión: rango, desviación típica, varianza, rango intercuartílico.....	72
3. La curva normal	76
4. Ejercicios	79
5. Estadística descriptiva bivalente: diagramas de dispersión y tablas de contingencia	81
1. Variables cuantitativas: diagramas de dispersión o nubes de puntos	81
2. El coeficiente de correlación	88
3. La regresión lineal	93
4. Correlación no lineal y correlación múltiple	97
5. Variables cualitativas: tablas de contingencia.....	99
6. Ejercicios	100
6. Introducción al cálculo de probabilidades y la inferencia estadística	103
1. Probabilidad simple y compleja	103
2. Probabilidades y frecuencias	105
3. Probabilidad y ciencia: contraste de hipótesis y significación estadística	107
4. Ejercicios	111
7. Estadística inferencial I: el muestreo probabilístico	113
1. Población y muestra. Tipos de muestreo.....	113
2. Fórmulas de inferencia	117
3. El muestreo en prospección arqueológica	121
4. Ejercicios	127
8. Estadística inferencial II: el contraste de hipótesis.....	129
1. Variables cuantitativas: contraste de la t de Student.....	129
2. Variables cualitativas: el chi-cuadrado	134
3. Ejercicios	139
9. Estadística multivariante I: la clasificación automática	141
1. Clasificación por conglomerados jerárquicos: <i>cluster</i>	141
2. Clasificación no jerárquica: K -medias.....	150
3. Ejercicios	152
10. Estadística multivariante II: reducción de datos por análisis factorial .	153
1. Análisis de componentes principales.....	153
2. Análisis factorial de correspondencias.....	162
3. Escalograma multidimensional.....	171
4. Ejercicios	177
Bibliografía	179
Índice analítico	183

Prólogo

Este libro pretende, espero que de forma realista y no pecando de optimismo, enseñar los métodos cuantitativos básicos a estudiantes y profesionales de la arqueología con un conocimiento elemental de matemáticas. Por razones de espacio, y existir buenos manuales en castellano (Hodder y Orton, García Sanjuán), no se ha tratado el importante tema de la estadística aplicada al análisis espacial y la arqueología del paisaje. Mi mayor esfuerzo se ha dirigido a explicar los modelos y fórmulas clásicas de la estadística de la manera más sencilla en que soy capaz. Como se verá, en el volumen no aparece ninguna fórmula matemática, algo bastante extraño en un libro sobre estadística. Solo en el caso de las operaciones más básicas he explicado textualmente cómo efectuarlas de forma manual, indicando a continuación la manera práctica de calcularlas con un programa informático de estadística. En la bibliografía de los trabajos concretos que se citan en el volumen y en la lista al final del libro, quienes tengan interés en profundizar más podrán conocer otros modelos y contrastes estadísticos diferentes, así como introducirse en sus fundamentos matemáticos. La experiencia me ha enseñado que una vez comprendida la filosofía de los métodos más simples, avanzar en otros más complejos es una tarea relativamente fácil.

Decía Clive Orton, el matemático que durante decenios enseñó la materia en el Instituto de Arqueología de Londres, que siempre aconsejaba a sus alumnos que tuvieran la estadística en mente pero que luego busca-

ran un estadístico amigable y comprensivo. En vista de la escasa cantidad de estadísticos (y de matemáticos, por más señas) que un arqueólogo puede llegar a conocer a lo largo de su vida, mi consejo es que con un pequeño esfuerzo se introduzcan en los métodos que aquí se describen, y que los practiquen mínimamente con cualquier programa informático de estadística, algo en lo que por otro lado encontrarán no poca diversión. Todo un mundo nuevo de apreciaciones e hipótesis sobre la cultura material del pasado se hará presente ante sus ojos.

El presente manual también culmina mi experiencia como profesor de la materia en la Universidad Complutense de Madrid, desde los lejanos años ochenta cuando empecé a explicarla en los cursos de doctorado a pequeños grupos de estudiantes licenciados, hasta mi docencia actual de tres años ya en la asignatura obligatoria de «Arqueología Cuantitativa» dentro del nuevo grado de Arqueología. Durante todo este tiempo he sido testigo, muchas veces impotente, de la dificultad que suponía comprender la estadística para jóvenes con muy pobre formación matemática. El comienzo del grado de Arqueología, con todo, y gracias al talento e interés que he visto en su nuevo alumnado, ha acabado restaurando mi confianza y empujándome a escribir este libro.

Dejo aquí constancia de mi agradecimiento a las personas que me animaron a escribir este volumen, la mayoría miembros del Departamento de Prehistoria de la Complutense, y sobre todo a quienes tuvieron la paciencia de leer y criticar con acierto muchos de sus aspectos: el ingeniero Jaime del Rey Tapia, mi primer amigo en la escuela de aeronáuticos, que lo sigue siendo a pesar del mucho tiempo transcurrido y de nuestros caminos profesionales tan separados, y que me hizo ver no pocas irregularidades técnicas y lógicas en el texto; la historiadora Carmen Ortiz García, todavía mi compañera después de solo un poco menos de tiempo, que afinó mi argumentación como he tenido la fortuna de que lo haya hecho también con todos mis anteriores libros; los arqueólogos Jorge de Torres Rodríguez y Jesús Álvarez Sanchís, que adoptaron el papel de lectores críticos externos con sugerencias que me fueron de gran utilidad en muchos apartados. Finalmente, los ya mencionados estudiantes del grado de Arqueología, en especial los famosamente habladores del curso 2012-2013, que preguntaron sin cesar sobre todo lo que yo con torpeza les contaba y así me ayudaron a entenderlo mejor, y también a volver a soñar con una universidad en que profesorado y alumnado participemos por igual en la generación del conocimiento.

1. Introducción

1. ¿Por qué necesitamos la estadística?

En la fachada de la Escuela de Investigación Social de la Universidad de Chicago, levantada en 1929, hay una inscripción con esta frase del eminente físico William Thomson, lord Kelvin: «Si no puedes medir, tu conocimiento es escaso e insatisfactorio». Por entonces empezaba la controversia, que de diversas maneras dura hasta hoy mismo, entre sociólogos más y menos «cientificistas» («duros» y «blandos»), según apoyaran una investigación cuantitativa (con datos concretos y medibles de muestras amplias) o cualitativa (con datos más generales de muestras más pequeñas o incluso de un solo caso concreto, como las biografías o «historias de vida»).

En un sentido parecido al anterior, cuando poco después de la Segunda Guerra Mundial el entomólogo y sexólogo estadounidense Alfred Kinsey publicó su entonces escandaloso análisis del comportamiento sexual de los varones norteamericanos, fue criticado por limitarse a observar la parte más física de la relación entre hombres y mujeres, olvidando la parte espiritual y supuestamente más «elevada», lo que se conoce como «amor». Su respuesta sobre el asunto fue muy clara: no lo había analizado porque, al contrario de lo que ocurre con el sexo, que consiste en unas actividades físicas y concretas, el amor no se puede medir.

Mi opinión es que ambos tipos de conocimiento son necesarios en el intento de comprender al ser humano, tarea cuya principal dificultad proviene de que observador y observado son la misma cosa. Aunque siento mayor simpatía hacia las aproximaciones cualitativas —y así, por ejemplo, me interesan más las aportaciones del psicoanálisis que las de la psicología conductista o las neurociencias—, como científico considero una especie de suicidio intelectual renunciar a las percepciones de la realidad que se pueden adquirir a través de los números.

De entre todas las ciencias sociales y humanas, pocas como la arqueología muestran una necesidad tan grande de los tratamientos cuantitativos. Por dedicarse al análisis de los restos materiales de la actividad humana, tanto en el pasado como en el presente (la nueva rama de la «arqueología contemporánea»), es la misma concreción de los datos arqueológicos la que parece estar pidiendo a voces que los midamos, los contemos, los clasifiquemos... Puesto que de esas piedras y huesos que descubrimos los arqueólogos es preciso extraer la mayor información posible sobre la cultura completa de sus antiguos propietarios —al igual que a través del sexo podemos saber algo del amor—, si los datos materiales están bien presentados (ordenados, resumidos, clasificados...) antes de su extrapolación al orden social, esta labor, que alguien llamó el «salto de la imaginación», será sin duda más directa y estará mejor fundamentada.

Hubo un tiempo en la investigación arqueológica cuando en las excavaciones se buscaban sobre todo piezas excepcionales, bien por su valor artístico (pensemos en el descubrimiento de la tumba de Tutankamón) o por su información cronológica o cultural (como los «fósiles directores», piezas características que definen una época o civilización concretas). Lo normal entonces era que la inmensa mayoría de los objetos encontrados simplemente se arrojaran de nuevo a la excavación una vez terminada o quedasen olvidados en las «terreras» o montículos de tierra extraída, creados alrededor. Cuando vemos, por ejemplo en los primeros decenios del siglo xx, cómo excavaciones de grandes yacimientos que duraban varios años, empleando a un gran número de obreros, se publicaban en escritos de poco más de cien páginas y algunas decenas de fotos y dibujos, mientras que hoy los informes de excavación son mucho más extensos y detallados, comprendemos el progreso analítico que ha experimentado la arqueología. Hoy en día, el principio general es que en una excavación se recoge *todo*, lo cual no es totalmente cierto porque ello haría la labor imposible, pero sí se aproxima mucho a la realidad: cada vez existen más tipos de datos que pueden proporcionar información valiosa; por ejemplo, los fragmentos de pequeños carbones (por lo general de madera quemada) que antes no se separaban del sedimento y solo se registraban cuando eran anormalmente abundantes, ahora empiezan a contabilizarse uno por uno en las excavaciones más avanzadas.

Una vez que la arqueología empezó a superar aquellos años en que sus profesionales vivían preocupados solo por los «descubrimientos», la necesidad de registrar tantos artefactos (objetos modificados por el ser humano) y ecofactos (restos biológicos asociados a la actividad humana) obligó inevitablemente al uso de la cuantificación. Hoy los arqueólogos dedicamos mucho tiempo a contar (fragmentos cerámicos, útiles y desechos líticos, huesos, semillas, casas, tumbas...), medir y comparar (longitudes, anchuras, grosores, profundidades, pesos, áreas, ángulos, volúmenes, colores, grados de dureza...), etc., y el producto de ese trabajo, cuando por fin está terminado, son *números*.

Y precisamente la ciencia que trabaja con números (no con los números teóricos como la matemática pura, sino con los que provienen de la observación de la realidad) es la *estadística*. Por ello sería absurdo que los arqueólogos, amparándose en que su formación es fundamentalmente humanística, renunciaran a conocer, apreciar y sobre todo beneficiarse de los métodos y técnicas desarrollados por los estadísticos desde hace mucho tiempo. Es cierto que esos avances tuvieron lugar en disciplinas muy diferentes de la nuestra (biología, sociología, etc.), lo que plantea algunos problemas y cautelas sobre las que volveremos más adelante.

Una buena definición de estadística puede ser la siguiente: la ciencia matemática que trata de la recolección, análisis, presentación e interpretación de todo tipo de datos, sobre todo los susceptibles de adoptar, directa o indirectamente, una forma numérica, con el fin de ayudar en la toma de decisiones o explicar la regularidad o irregularidad de algún fenómeno. La lista de disciplinas concretas que utilizan la estadística es demasiado larga para resumirla aquí, pero abarca casi todos los campos científicos, con especial incidencia, por su aplicación social práctica, en economía, medicina o ecología. Hasta la rama del conocimiento que podríamos considerar menos matemática, la filosofía, incluye en sus apartados de más reciente desarrollo una «filosofía de la estadística» que analiza y critica los modelos profundos que subyacen a sus aplicaciones habituales.

Que nadie se asuste al oír hablar de matemáticas, pues por suerte las que se aplican en estadística están entre las más simples, y en este libro he puesto especial interés en presentar los sistemas más sencillos y eficaces de todos los que se usan y emplear las expresiones más accesibles que he podido encontrar. Puedo adelantar que además de las cuatro operaciones básicas, bastará con conocer lo que es elevar a una potencia (multiplicar un número por sí mismo varias veces), hallar una raíz (proceso inverso al anterior) y calcular un logaritmo (el número al que hay que elevar un número fijo, llamado base, para obtener el valor que tenemos). También por fortuna hoy día prácticamente todas las operaciones y gráficos estadísticos se realizan utilizando programas de ordenador, lo que

las hace de manejo muy simple y rápido. Esto es una ventaja pero también un inconveniente: de cualquier serie de números sometidos a un programa, adecuado o no para ellos pero que se maneja con un simple clic del ratón del ordenador, surgen unos resultados que pueden parecer significativos aunque en realidad carezcan totalmente de sentido. Por ello es necesario que los arqueólogos dispongamos de al menos una idea básica sobre los modelos, en la práctica muy pocos, que son necesarios para analizar nuestros datos, además de saber a qué tipo de estos se aplica cada modelo, y ser capaces de interpretar en líneas generales los resultados de los análisis una vez aplicados. De lo contrario nos moveremos a ciegas sin llegar a ningún resultado práctico o, lo que es peor, alcanzando conclusiones completamente equivocadas.

2. Un ejemplo

En este pequeño capítulo introductorio voy a presentar un caso concreto de datos arqueológicos cuantitativos y su tratamiento estadístico gráfico, tomado del trabajo de campo que mejor conozco, el que he llevado a cabo desde hace ya varias décadas en el noreste de África. Al igual que en otros ejemplos del libro, introduciré brevemente los datos arqueológicos y su contexto cultural, para detenerme con más detalle en el objetivo principal de la estadística: la forma en que se pueden desarrollar hipótesis, contrastables más adelante con nuevos datos, partiendo de los resultados del análisis cuantitativo.

Uno de los apartados arqueológicos donde la estadística comenzó antes y es más necesaria es el análisis lítico, tanto de las culturas más antiguas del Paleolítico como de las recientes a partir del Neolítico que siguieron utilizando la piedra para buena parte de sus herramientas. Entre 2001 y 2005 dirigí una prospección en la región de Benishangul, un área remota y boscosa del oeste de Etiopía lindando con la frontera de Sudán, donde nunca antes se había realizado ninguna investigación arqueológica. Como es habitual en este tipo de medios, los mejores, o casi los únicos, yacimientos antiguos que han conservado restos y son localizables entre la espesa vegetación son las pequeñas cuevas o abrigos rocosos. En dos de ellos la excavación reveló niveles correspondientes a un Paleolítico final con muchos caracteres arcaicos de la *Middle Stone Age* africana (Paleolítico Medio), junto con otros más recientes que ya iban acompañados de cerámicas similares a las del Sudán central, evidencia de los contactos que tuvieron que producirse entre los últimos cazadores del Altiplano etíope y los pastores saharianos que se movían hacia el sur huyendo de la progresiva sequía de las llanuras sudanesas a mediados del Holoceno.

Pues bien, en apenas unos metros cuadrados encontramos decenas de miles de fragmentos de piedra de calidad más bien inferior (cuarzo), en su mayoría correspondientes a la actividad humana (talla y retoque líticos). Por suerte, no me correspondió a mí la pesada labor de clasificar, contar y pesar todos esos restos, que ocupó muchos días de trabajo al especialista del equipo, Ignacio de la Torre. En la figura 1.1 se puede ver a Nacho trabajando, con bastante buena cara a pesar del mucho tiempo que ya llevaba empleado en el recuento: útiles retocados, lascas o láminas, núcleos, fragmentos de desecho menores de 2 cm, entre 2 y 5 cm, mayores de 5 cm...

Figura 1.1. Ignacio de la Torre clasificando restos líticos del abrigo de Bel K'urk'umu, Benishangul (Asosa, Etiopía, febrero de 2002). Fotografía de Luis Luque.



Al final del recuento por fin dispusimos del resultado que importaba para el análisis, las tablas con el resumen de los hallazgos en los niveles artificiales de excavación. Como se aprecia en la tabla 1.1 aquí presentada, correspondiente al sitio de Bel K'urk'umu, el número total de piezas analizadas fue de 25.679, con un peso cercano a los 70 kilogramos (en el otro yacimiento, K'aaba, más antiguo y sin cerámica, se contaron 5.280 piezas). Aunque puede sorprender que al final una gran parte de esa información se haya reducido a una simple tabla de menos de una página, esta es precisamente la ventaja principal de la estadística, resumir y presentar los datos en forma manejable.

Tabla 1.1. Frecuencias de las diferentes categorías de lascado en los niveles artificiales (1-18) y naturales (A-F) del abrigo de Bel K'urk'umu (Etiopía).

	Nivel	Lascas	Láminas	Frag. 2-5 cm	Frag. >5	Frag. <2	Peso gramos
1	A-B	17	0	138	2	231	1.340
2	C	42	2	135	8	153	1.705
3	D	35	1	146	0	269	915
4	D	19	0	186	0	304	965
5	D	27	0	157	3	290	1.221
6	D	34	3	254	0	523	1.805
7	D	35	6	393	10	630	2.730
8	E	36	2	325	14	700	3.090
9	E	78	2	578	22	2.280	5.985
10	E	91	9	828	21	2.264	6.342
11	E	7	0	309	28	929	3.215
12	F	63	3	827	36	4.056	16.895
13	F	72	5	1.839	29	3.634	14.020
14	F	43	3	175	7	433	2.315
15	F	71	5	286	10	796	4.615
16	F	12	1	80	8	350	1.635
17	F	6	1	23	0	145	610
18	F	3	1	0	0	0	20
TOTAL		711	44	6.679	198	17.987	69.423

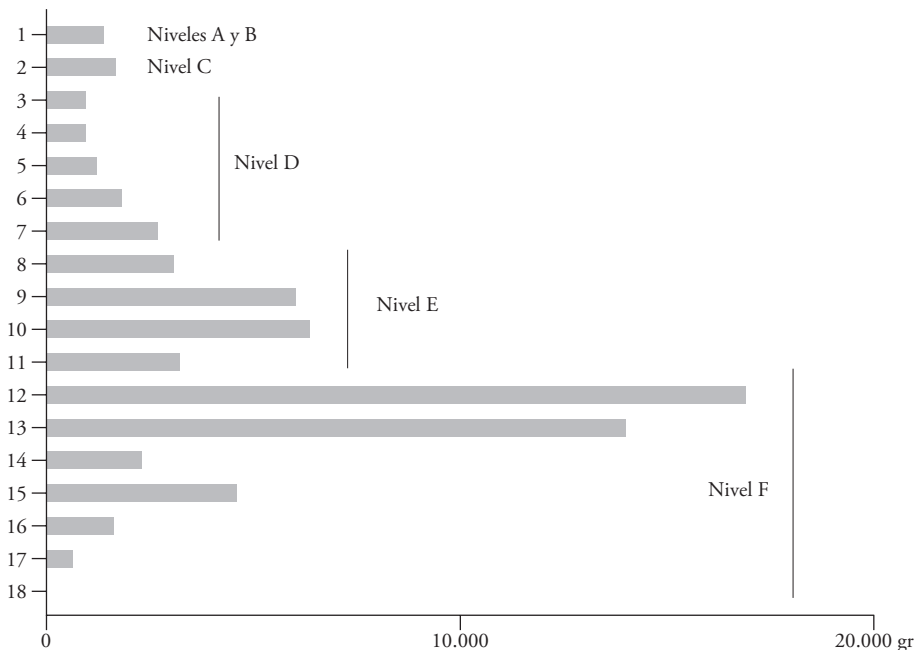
Si observamos con atención la tabla, podemos ver cosas interesantes, como la poca frecuencia de lascas y láminas completas frente a la abundancia de fragmentos de ellas y *chunks*, simples trocitos que saltan al tallar o retocar un núcleo o lasca. También observamos que hay dos niveles artificiales (12 y 13) que concentran la mayor parte de los restos, en torno a un 44% del peso total.

Seguro que a más de uno se le ocurriría emplear algún gráfico como manera simple de reducir la información de la tabla a un formato toda-

vía más fácilmente apreciable («un gráfico vale más que mil palabras... o mil números»). Un sencillo diagrama de barras como el mostrado en la figura 1.2, donde la longitud de las barras es proporcional al peso del material lítico en cada nivel respectivo, nos muestra en un instante el citado fenómeno de concentración de restos líticos en esos dos niveles de la parte inferior de la estratigrafía.

Y ahora viene la parte más peliaguda del análisis, lo que ya no es estadístico, lo que no dicen los números, sino que tenemos que afirmar nosotros. ¿Qué significa esa anómala concentración de útiles? ¿Fue la época en que se depositaron esos niveles la principal de uso del abrigo, siendo los niveles superiores fruto de una menor ocupación en tiempos posteriores? ¿O, por el contrario, son esos niveles inferiores producto de una única fase de ocupación, y luego los procesos posdeposicionales (las abundantes madrigueras, raíces, movimientos de tierra de todo tipo, etc.) se encargaron de ir subiendo hacia arriba poco a poco los artefactos hasta llegar a la distribución actual?

Figura 1.2. Diagrama de barras de los datos de peso total en gramos de útiles líticos (eje horizontal) en los 18 niveles artificiales (eje vertical) del abrigo de Bel K'urk'umu. El modelo es unimodal, con un solo máximo concentrado en la parte superior del nivel inferior del abrigo (nivel natural F).

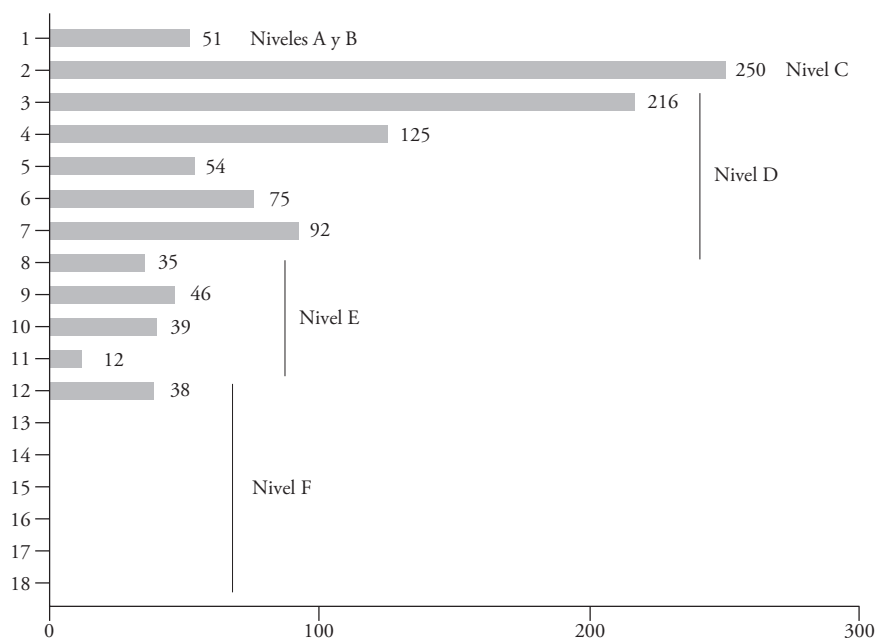


Si no optamos directamente por la primera opción, la más habitual, es porque la forma de la cultura lítica (tipos de núcleos y talla, principales útiles, etc.) apenas cambia de abajo arriba, lo que podría explicarse por la segunda hipótesis. Una forma de «hacer tiempo» hasta tomar una decisión correcta en este tema es comparar la distribución lítica con la que presentan otros tipos de hallazgos. Por desgracia no contábamos con un tipo muy frecuente en yacimientos prehistóricos, los huesos, porque el carácter muy ácido de los suelos los había destruido por completo (algo frecuente en África), pero sí se registraron en el sitio numerosos fragmentos cerámicos, un total de 1.033, de los que algo menos de un centenar llevaban algún tipo de decoración. La distribución de cerámica, tal como se ve en la figura 1.3, es opuesta a la lítica, con una concentración mucho mayor en la parte superior de la estratigrafía (niveles artificiales 2-4), sugiriendo que el abrigo se siguió ocupando tiempo después de la formación del intenso depósito inferior y que la llegada de la cerámica fue mayoritariamente posterior a ese momento. El mismo modelo surge de la distribución de las cuatro fechas de carbono-14 obtenidas para el sitio, «bien» ordenadas de abajo arriba, con valores medios de 4.965 y 4.470 bp (nivel E), 2.020 y 875 bp (nivel D). Por suerte, casi nunca estamos obligados a decidir nada a partir de una sola información o conjunto de datos, aunque también es cierto que incluso cuando contamos con una información abundante la posibilidad de equivocarse sigue estando presente. (Véase informe completo sobre el yacimiento en *Journal of African Archaeology*, 5-1, 2007, pp. 91-126).

3. El problema de las muestras

Como ya vimos, la estadística es una especialidad que surgió en el terreno de las ciencias naturales, para luego aplicarse prácticamente a todas las demás disciplinas. Dentro de las ciencias sociales, la estadística es de enorme utilidad en sociología o economía, donde ha alcanzado niveles muy altos de sofisticación y exactitud. Que este terreno, con todo, no es idéntico al de la física o la medicina lo demuestra el hecho de que la estadística tenga una cierta mala fama en algunos terrenos, como por ejemplo los que afectan a la economía política, es decir, allí donde intervienen intereses personales o de grupo. Hay un dicho muy conocido sobre cuantificación en economía que afirma que «un economista es un experto que sabrá mañana por qué las cosas que predijo ayer no sucedieron hoy». De hecho, la gran recesión económica actual fue prevista por muy pocos estudiosos, curiosamente aquellos «aguafiestas» que se pasaban el tiempo advirtiendo de desastres futuros, y que como es lógico alguna vez tenían que acertar. Otra sentencia famosa, originada en el siglo XIX (se atribuyó

Figura 1.3. Diagrama de barras del número de fragmentos cerámicos (eje horizontal) en los 18 niveles artificiales (eje vertical) del abrigo de Bel K'urk'umu. Al igual que en el caso de los útiles líticos (figura 1.2), el modelo es unimodal, pero ahora el máximo está situado en la parte superior de la estratigrafía, niveles naturales D y C.



por error a Mark Twain), es la que afirma que hay tres clases de mentiras: «mentiras, malditas mentiras y estadísticas», poniendo a estas en el ápice de la negatividad expresiva. Muy parecido es el conocido aforismo de que hay «mentirosos, malditos mentirosos y expertos». Según se ha visto hace poco, fue la opinión autorizadísima de los expertos de la agencia de calificación Standard & Poor, presumiblemente entre los mejores economistas del mundo pero que dieron una nota muy alta a los «activos crediticios tóxicos» (las hipotecas *subprime* en Estados Unidos), una de las primeras causas de la crisis actual.

Pero es evidente que en los ejemplos anteriores intervienen factores adicionales externos a la estadística misma, como pueden ser, por un lado, la extrema complicación de la realidad presente y la imperfecta calidad de los datos que tenemos sobre ella, y por otro, algo tan importante y al tiempo tan complejo como la ética personal. Los resultados estadísticos no suelen admitir una sola interpretación y el experto está obligado a elegir, debiendo ser lo más honrado posible y escoger la op-

ción que cree más probable. Cuando yo empecé a estudiar estadística todavía se citaba un libro humorístico de la década de 1950, escrito por el periodista estadounidense Darrell Huff, *How to lie with Statistics* [Cómo mentir con la estadística]. Es curioso que el propio Huff fuera uno de los expertos que testificó entonces, ante el Congreso de Estados Unidos, en contra de los datos que relacionaban el consumo de tabaco con enfermedades mortales, tal vez aplicando alguno de los trucos de su famoso libro. De todas formas, en uno de los manuales que yo seguía por entonces, de N. Downie y R. Heath, se decía también que la mayoría de los fallos de la estadística no provienen de oscuros e inconfesables motivos..., sino de simples errores de aritmética, ¡de sumas y multiplicaciones mal hechas!

En general, y aparte de los comprensibles «fallos» humanos, lo que más afecta a la calidad de la estadística es la calidad original de los datos de partida, lo que a partir de ahora llamaremos la *muestra*. Aunque veremos este tema con más detalle en otro capítulo, es necesario empezar diciendo que las muestras con que contamos en arqueología no son demasiado buenas. Al igual que ocurre en otras ciencias que estudian el pasado, lo que ha llegado hasta nosotros de otras épocas no solo es una pequeña parte de lo que existió realmente, sino que además es probable que sea poco representativa del conjunto o, lo que es peor aún, que no seamos capaces de evaluar, ni siquiera de manera aproximada, su grado real de representatividad.

La razón por la cual las encuestas de opinión pública —sean para conocer los hábitos de consumo o la intención de voto político— suelen acertar en líneas generales (siempre hay un margen de error por la falta de sinceridad de algunas respuestas, aparte del «error estadístico» que al menos es evaluable, como veremos más adelante) es porque las muestras están bien elegidas, utilizando *técnicas de muestreo* que aseguren que las personas encuestadas sean una fracción *representativa* del conjunto total que nos interesa, lo que en estadística se llama *población* o *universo*. El muestreo estadístico se basa en dar a cualquier parte de la población total de datos una igualdad de oportunidades de ser incluida en la muestra. Lo que se obtiene así es una muestra *aleatoria*, que en teoría tiene una gran probabilidad de ser representativa, aunque nunca la seguridad absoluta, siendo esto algo fácil de comprender sin necesidad de entrar en su fundamento matemático.

Como ejemplo podemos pensar en la composición por sexos del alumnado de una facultad, que supongamos que en total es de la mitad aproximadamente de varones y la mitad de mujeres. Si nos ponemos en la puerta y empezamos a apuntar el sexo de cada persona que pasa, por ejemplo durante una hora, tendremos una muestra tanto más representativa cuanto más parecidos sean nuestros porcentajes a los que sabemos

que existen en toda la facultad. Si nos sale 48% de chicos y 52% de chicas, la muestra será aceptable, pero si tenemos 40% y 60% respectivamente la cosa no irá tan bien, y no digamos si nos sale el doble de un sexo que del otro. ¿De qué dependerá esa calidad de la muestra? Se entiende rápidamente que si estamos en la puerta durante dos horas tendremos un resultado más fiable que si permanecemos una hora, es decir, el tamaño de la muestra influye en su representatividad (figura 1.4). También la aleatoriedad, y por eso quizás hubiéramos hecho mejor tomando nota durante una hora por la mañana y otra por la tarde, de manera que muestreáramos de forma más repartida a lo largo del día.

Pasando a nuestro campo, sabemos que muchos datos proporcionados por la arqueología (antiguas excavaciones, materiales de museos, publicaciones de tiempo atrás, etc.) difícilmente se pueden considerar muestras representativas. Afortunadamente, la investigación actual mejora su calidad al aplicar cada vez más las técnicas de muestreo, por ejemplo, al seleccionar partes del terreno para prospectar, materiales almacenados para análisis, zonas de un yacimiento para excavar, etc. De momento, nos quedaremos con la idea de que cuanto más dudoso sea el valor de las muestras mayor cuidado deberemos tener al aplicar los métodos de *inferencia* que veremos a continuación.

Figura 1.4. La importancia del tamaño de la muestra en arqueología es ridiculizada en este chiste «anacrónico» de M. Pickering, publicado en la revista británica *Antiquity* (vol. 67, p. 708, diciembre, 1993).



4. Comenzar a entender la probabilidad y la inferencia

Una dificultad habitual para entender la estadística es la de introducirse en el mundo de la probabilidad. Según mi experiencia docente, este es el caballo de batalla de la mayoría de los estudiantes. Estamos acostumbrados a que los números que conocemos sean algo fijo y conocido en cada momento (como por ejemplo el peso o la estatura de una persona) y cuesta comprender que a menudo del valor de algo solo sabemos que se sitúa entre tal y cual límite con una probabilidad dada. Evidentemente, es mejor saber eso que no tener ni idea de por dónde anda la variable que nos interesa.

Un ejemplo muy conocido en arqueología son las fechas de carbono-14, cuyo sistema es similar al de muchos otros métodos radiométricos. Al estar calculada la fecha en función de la radiación de ese isótopo que todavía emite la muestra (tanto menor cuanto más antigua) y ser esa radiación un fenómeno aleatorio, cada vez que la medimos nos da un valor parecido pero distinto, y lo que tenemos al final es una distribución de probabilidad muy parecida a la famosa curva normal o «campana de Gauss». La fecha real, que por supuesto nunca conoceremos, está comprendida dentro de ese intervalo, tanto más amplio cuanto mayor queramos que sea la probabilidad de acertar. La forma de campana también nos dice que esa probabilidad es tanto mayor cuanto más cerca estemos de la media, y por eso normalmente usamos el intervalo de la media más/menos una desviación típica (el «error» de la medida), que tiene una probabilidad de algo más de dos tercios (68,26%). Si queremos mayor seguridad (95,45%), usaremos un intervalo con dos desviaciones típicas a cada lado, aunque lo que ganamos en probabilidad lo perderemos en exactitud al movernos en un lapso de tiempo más amplio. Un problema añadido con las fechas de carbono-14 es que para calcular la edad hay que ver la cantidad de isótopo que existía entonces (para compararla con la que todavía conserva la muestra y ver el tiempo transcurrido), y dado que esa cantidad varió de forma muy irregular en el pasado, al hacer la corrección («calibración») la curva normal, perfecta y simétrica, se convierte por desgracia en un histograma de probabilidad irregular que puede tener varios máximos e incluso espacios vacíos (véase figura 2.3 en el capítulo siguiente).

El tema de la probabilidad se complica, como veremos con mayor detalle más adelante, en la estadística inferencial, cuando nos atrevemos a evaluar la probabilidad que tiene una determinada hipótesis de ser cierta o falsa. Lo que se calcula realmente es la probabilidad que tendría una muestra (o un determinado valor en ella) de salir en nuestro muestreo aleatorio si la hipótesis que estamos contrastando fuera cierta. Si esa probabilidad es grande, aceptamos la hipótesis (o, mejor, no la podemos

rechazar, que no es lo mismo), y si es pequeña, la rechazamos. Desde hace tiempo, se ha tomado por acuerdo entre los estadísticos el valor del 5% como el valor límite entre la aceptación y el rechazo, de forma que si la probabilidad es mayor de ese valor aceptamos la hipótesis y lo contrario si es menor. De aquí viene esa expresión tan repetida en estadística de que algo es «significativo al 5%» (o al 1% si queremos ser más estrictos). Como veremos, lo más habitual es que deseemos rechazar una hipótesis que no nos interesa (la famosa «hipótesis nula») y ansiemos que nos salga un valor de la probabilidad (p) menor del 5% (0,05 expresado en tanto por uno o proporción, la forma más habitual).

¿Y qué hacemos si nos sale un 0,051, o todavía peor, un 0,05001? Pues aguantarnos, naturalmente. Por suerte, en arqueología no ocurre como en medicina cuando hay que decidir si un medicamento es apropiado o no para curar cierta enfermedad; a nosotros solo nos gustaría saber cosas como hasta qué punto en una necrópolis es fuerte la asociación del género masculino o femenino con los ajuares funerarios que indican poder, o si la diferencia entre el tamaño de los fragmentos cerámicos de un recinto habitacional y otro usado como basurero es «significativa», por lo que nuestra preocupación debería ser mucho menor.

A lo largo de este volumen iremos empleando progresivamente los conceptos probabilísticos para conseguir asimilarlos, y sobre todo para entender su problemática utilización en muchos casos arqueológicos. También veremos que la mayoría de los métodos estadísticos se pueden aplicar en un sentido «exploratorio», es decir, solo para observar cómo son y se distribuyen los datos sin preocuparnos de momento de contrastar ninguna hipótesis (véase epígrafe 6 del capítulo 3).

5. Los programas informáticos

Cuando comencé a utilizar la estadística en mi trabajo arqueológico, me temo que hace ya mucho tiempo, los ordenadores que se usaban eran aquellos enormes IBM que recibían los datos mediante tarjetas de cartón perforadas (cada columna de agujeros correspondía a un dígito). Usando ese sistema con los medios de la Universidad Complutense, hice una clasificación automática de ochenta vasijas de una necrópolis sudanesa que era la parte principal de mi tesis doctoral, con resultados por cierto no demasiado satisfactorios. Recuerdo que tuve que hacer varios viajes al Centro de Cálculo de la universidad y que el proceso me llevó unas cuantas semanas.

Poco después compré mi primer ordenador, un pequeño Spectrum que usaba una televisión en blanco y negro como pantalla. Utilizando el

sencillo lenguaje Basic, pasé muchas horas entretenido escribiendo programas para hacer cálculos estadísticos sencillos. Más adelante, advertido de que había salido un libro francés de Thierry Foucart con programas en Basic para análisis multivariantes en lo que entonces se llamaban «microordenadores», empleé todavía más horas (de agradable recuerdo) adaptando esa programación al Spectrum. Uno de los resultados de aquel trabajo fue un artículo sobre la seriación cronológica en arqueología que vio la luz en *Trabajos de Prehistoria* (vol. 42, 1985). Los programas de Foucart me sirvieron después, pasados en 1987 a mi primer PC, para seguir llevando a cabo análisis factoriales de mis propios datos y de los de otros investigadores de mi departamento.

Unos años después empezamos a utilizar en los PC un potente programa estadístico cuyos derechos tenía la universidad, el *Statistical Package for the Social Sciences* (SPSS, ahora propiedad de IBM). Aunque fue extremadamente útil, tenía algunos problemas, como un interfaz con el usuario bastante lento y complicado (que no parece haber mejorado en las últimas versiones), un enorme número de aplicaciones de las que en arqueología se usaban solo unas pocas, ausencia hasta hace poco de análisis muy necesarios para nosotros, como el de Correspondencias y el MDS, y un alto precio para el uso individual. Su equivalente gratuito, PSPP, aunque más simple de utilizar, permite una gama de análisis demasiado pequeña.

El problema con la mayoría de los programas estadísticos actuales es que casi todos han sido diseñados para el trabajo en ciencias físico-naturales o en determinadas ciencias sociales, lejos por lo tanto de los intereses arqueológicos, y que los mejores son de «software propietario» y es necesario pagar por su licencia de uso. Otros inconvenientes son que a veces están especializados solo en determinados análisis o que los comandos deben ser introducidos línea a línea (como en el conocido paquete libre «R»), complicando y haciendo muy lento su uso para personas poco expertas.

De los programas estadísticos especializados para arqueología, el más famoso es el gratuito desarrollado por la Universidad de Bonn (*Win-Basp*), aunque tiene la desventaja de que cada análisis se divide en varias partes, debiéndose guardar la salida de cada una para ser luego el fichero de entrada de la siguiente. No obstante, sus gráficos son de gran calidad y aparte de la estadística habitual dispone de análisis muy concretos para arqueología (matriz de Harris, reconocimiento de estructuras en aglomeraciones de agujeros de poste, rectificación de fotos aéreas oblicuas, etc.). Lo mismo ocurre con un programa comercial desarrollado por el arqueólogo estadounidense Keith Kintigh (TFQA, *Tools for Quantitative Archaeology*), que encuentro demasiado especializado y que carece de análisis más comunes para los que es necesario contar con otro progra-

ma más general. Consultando un listado razonado que incluye la mayoría de aplicaciones existentes para arqueología (archaeologicaltraces.org), se advierte el mismo problema de excesiva particularidad en la mayoría de ellas.

El programa que llevo usando durante los últimos años en los cursos de doctorado y de arqueología cuantitativa en el Grado de Arqueología de la Complutense, y del que estoy medianamente satisfecho, es el PAST (acrónimo de PAleontological STatistics), desarrollado por Øyvind Hammer del Museo de Historia Natural de Oslo y David A.T. Harper del Museo Geológico de Copenhague, y basado en un programa anterior diseñado por Paul D. Ryan de la Universidad de Galway. Creado para datos muy próximos a la arqueología, como son los geológicos y paleontológicos, es además de uso muy fácil con interfaz gráfico, salidas gráficas bastante adecuadas y sobre todo de uso libre (se puede descargar de <http://folk.uio.no/ohammer/past/>). El programa funciona como una hoja de cálculo donde se introducen o copian los datos, con una serie de ventanas desplegables que agrupan los análisis parecidos (transformación de datos, gráficos, estadígrafos y contrastes, multivariante, etc.). La mayoría de las aplicaciones en este libro se hicieron con la versión 2.17, y aunque al terminar el borrador ya existía una versión 3.02, la antigua me sigue pareciendo mejor en algunas cosas. Uno de los cambios en la reciente es que los análisis aparecen agrupados por tipos, lo que a veces retarda el encontrar cada uno de ellos en los diferentes desplegables.

En PAST los datos se pueden pegar en filas y columnas tras haberlos copiado de diversos programas comunes (Word, Excel, etc.). Para seleccionar una columna no hay más que pinchar encima de ella, al mismo tiempo que la tecla de mayúsculas si ya existe otra previamente seleccionada, y en la esquina superior izquierda si se quieren escoger todas. El programa avisa cuando hay que seleccionar más de una columna o un número determinado de ellas. Para analizar dos columnas concretas es preciso que estén juntas, lo que se consigue arrastrando una de ellas. En el desplegable «Edit» están las instrucciones habituales de edición, destacando «Remove» para borrar y «Transpose» para cambiar filas por columnas, algo muy útil en ocasiones. En el desplegable «Transform» hay cosas interesantes para hallar los logaritmos, restar de la media, pasar a porcentajes o cambiar datos numéricos en presencia/ausencia y ordenar las columnas por orden ascendente o descendente. A lo largo del libro iremos viendo la forma en que se pueden realizar los análisis estadísticos en cada caso.

6. Estadística y teoría arqueológica

La historia de la estadística aparece mezclada con la teoría social desde sus orígenes, en una manera que vista desde la actualidad resulta, cuanto menos, inquietante. En el ambiente de la ciencia británica de comienzos del siglo xx, más específicamente el campo biológico, bajo la sombra de Darwin y Galton y la idea clave de que la herencia genética determinaba de forma absoluta los rasgos biológicos e incluso los morales, uno de los principales fundadores de la estadística, Karl Pearson (1857-1936) (figura 1.5), a quien citaremos a menudo en este libro, era un ardiente defensor de la Eugenesia, es decir, del perfeccionamiento de la raza humana mediante la promoción de los individuos más aptos y la menor reproducción de los más defectuosos, o su eliminación directa en las versiones más extremistas de la teoría. Resulta hoy casi extravagante que las inocentes operaciones matemáticas de la estadística se hayan ideado entre otras cosas para aplicar una teoría que fomentaba el racismo y que fue universalmente rechazada y olvidada tras la conmoción que supuso el holocausto judío en la Segunda Guerra Mundial. Pero como un ligero descargo en su favor podemos recordar esta pequeña instantánea que nos permite comprender mejor aquella época: Pearson fue también un ardiente defensor del socialismo e incluso en cierta medida del feminismo (sufragismo), y llegó a rechazar todos los honores aristocráticos que la monarquía británica quiso imponerle como premio a su labor científica; es decir, que era posible compaginar un evolucionismo extremo con una posición progresista.

Las primeras aplicaciones numéricas en arqueología, todavía no propiamente estadísticas, fueron las complicadas comparaciones que el famoso egiptólogo británico W. Flinders Petrie realizó con los ajuares de las tumbas predinásticas de Nagada para conseguir su ordenación cronológica en grupos que llamó «fechas secuenciales» (*sequence dates*). Sus resultados se publicaron a finales del siglo xix, y aunque hubo numerosos ensayos posteriores, no fue hasta 1951 cuando se publicó el que se considera primer trabajo estadístico en arqueología, usando matrices numéricas, cálculo de coeficientes y permutación de filas y columnas, todo ello realizado entonces de forma manual por el arqueólogo estadounidense W. S. Robinson para ordenar cronológicamente varios niveles de un yacimiento maya. Es interesante que ambos trabajos, lo que hoy se conoce como «seriación», se concentraran en resolver el problema cronológico, justo en una época en que la teoría arqueológica predominante era el «historicismo cultural» que veía el pasado prehistórico como una serie de culturas consecutivas, cada una con sus rasgos culturales y períodos cronológicos concretos e influyéndose mutuamente entre sí (difusionismo).

Figura 1.5. Karl Pearson en 1890, National Portrait Gallery. Foto: Elliot y Fry.



Decía en una entrevista Lewis R. Binford, el principal fundador de la Nueva Arqueología en la década de 1960, que el descubrimiento hacia 1950 del método del carbono-14 libró a los arqueólogos de la obsesión por averiguar la fecha de sus hallazgos (que ahora les proporcionaba rápidamente un laboratorio físico-químico). Por ello pudieron ocuparse de cosas en principio más importantes, como la organización social o la funcionalidad de los restos arqueológicos. Lo que luego se llamó Arqueología Procesual, que parece seguir siendo hoy la postura teórica predominante a nivel global, empezó efectivamente con una orientación antropológica, buscando conocer las culturas antiguas en sí mismas y no solo la época en que vivieron. También se definió enseguida como una arqueología «científica» que aspiraba a seguir idénticos procedimientos que las ciencias «duras», tales como el famoso método hipotético-deductivo.

Ese método, considerado como el procedimiento científico por excelencia, consta de varias fases: recogida de muestras iniciales, formulación de hipótesis, deducciones teóricas de consecuencias prácticas de esas hipótesis, recogida de nuevas muestras y contrastación de las deducciones con los resultados experimentales, y toma de decisión sobre las hipótesis (corroboración, rechazo, modificación). La secuencia anterior está muy ligada a la estadística en varios de sus apartados, como el muestreo representativo, la probabilidad de la muestra si la hipótesis fuera cierta, la significación estadística o importancia del resultado, etc.