Fractal Analysis as Method for studying Social Hierarchy in Prehistoric Settlement Plans

with case studies from the Linear Pottery and Trypillia cultures (5.500 - 3.500 cal. BCE)

Hallvard Bruvoll   Department of Archaeology, Conservation and History
University of Oslo

2023-07-21

Contents

Preface

Thanks to everyone.

Remember to add abstracts

Furholt, Grier, et al. (2020), Furholt, Müller, et al. (2020), Furholt, Müller-Scheeßel, et al. (2020) are three different publications. Furholt, Grier, et al. (2020), Furholt, Müller, et al. (2020), Furholt, Müller-Scheeßel, et al. (2020) are the same three publications in the same order. gfdsdfgs , Get on with it!

Part I

Frameworks

Chapter 1

Introduction

1.1 Background of the study

- Studying social hierarchy in prehistory through fractal analysis of settlement plans

- In Europe, the Neolithic is the long and messy transition period between mobile hunter-gatherer groups in the Palaeo- and Mesolithic, and the first city states in the Bronze (Aegean) and Iron Ages (Mediterranean and Central Europe)

- Early farming economy, influence from the Near East.

- Large variety in scale and content of archaeologically defined culture groups. Single farmsteads and small hamlets in many phases – some with hardly any settlement evidence at all **examples mid-neo, use** Shennan (2018). Other phases include exceptionally large settlements, probably hosting populations of several thousand inhabitants, like at Maidanetske in central Ukraine around 3.800 BCE (see 3.3). While in some settings, like in Linear Pottery society in much of continental Europe north of the Alps towards 5.100 BCE (see 3.2), the dead were buried in simple pit graves, with very little distinction in treatment between individuals, in other phases some individuals were buried with tremendous amounts of precious goods like in Varna, Romania, or under colossal burial mounds like in Carnac, France, both in the mid-5th millennium (Shennan 2018)**check ref**.

- Seen at a very large scale – across the continent and through the Holocene – the development of society from small scale and relatively egalitarian towards large scale and more hierarchical seems evident (though not to everyone, see Graeber and Wengrow (2021) and Section **??**). When we look more closely however, this evolution is anything but linear, as both population sizes and levels of hierarchical organisation seem to fluctuate considerably, sometimes over short time spans as from the Trypillia **C2 to D1??, check this**, when the so-called mega-sites are abandoned and their former inhabitants regroup into much smaller settlements during a transition of maybe only **check and ref** years.

- In many cases, the level of social hierarchisation and complexity in a given Prehistoric society is very hard for researchers to evaluate, since many indicators of such structures are either lost from the archaeological record, or were never included in the first place [Perreault (2019); Section 3.1]. Archaeological traces that are often interpreted as signs of social complexity and hierarchy may furthermore be deceiving. Seemingly monumental structures were in many cases built through small additions over centuries, rather than in one colossal construction campaign [**example, Carnac alignments? Danish megaliths?**]. In many megalithic burial contexts, it may be impossible to know how large a segment of the society that had access to such inhumations (**rephrase**).

- New methods for investigating hierarchy: fractal analysis. Borrowed from other disciplines, not much tested in archaeology. Example from human geography, use:(Batty 2005; Batty and Longley 1994; D'Acci 2019; Jahanmiri and Parker 2022; Lagarias and Prastacos 2021; Tannier and Pumain 2005) (say what fractal analysis does, don't explain what it is here)(explain just the word fractal and cite section).

1.2   Research question and objectives

The overall goal of the present study is to test and assess the utility of fractal analysis techniques as tools for studying hierarchical social organisation in prehistoric societies. Two methodological approaches are under special scrutiny: the distribution fitting approach and the image analysis approach (Brown and Liebovitch 2010; Brown, Witschey, and Liebovitch 2005). These are ap-

plied to architectural data series from well-preserved and documented archaeological samples within Neolithic Linear Pottery and Trypillia contexts, as well as to synthetic data series. This thesis is thus not to be considered a culture-historic study of Linear Pottery or Trypillia society, but mainly a methodological study. However, results from the proposed analyses of these case studies may also contribute, as side-effects, to their respective fields of research.

For the distribution fitting approach, house-size distributions within settlements are modelled following a given procedure, and the retained model (the best fit) is interpreted in terms of social generating mechanisms. In particular, it is argued here that so-called power-law distributions reflect hierarchical structure, so that the identification of these within the studied samples may indicate the presence of some social hierarchisation process which warrants further interpretation.

With the image analysis approach, the spatial layouts of archaeological and synthetic settlement plans are analysed through the calculation of fractal dimension and lacunarity – summary statistics which serve as quantifications of irregular spatial patterns or image textures. I argue here that geometrically irregular settlement plans are indicative of relative independence between households, while settlements that develop within geometrically regular grids indicate stronger overarching social structures, with a continuous range of possibilities in-between. The goal here is to test to which extent quantitative measures like fractal dimension and lacunarity may help differentiating between varying degrees of planning in prehistoric settlements.

While both these methodological approaches are well developed and integrated to other disciplines (e.g. Lagarias and Prastacos 2021), their usage in archaeology have so far remained anecdotal (Diachenko 2018). A further overall goal of this thesis is to identify and explore possible limitations in the nature of archaeological data that may limit the applicability of fractal analysis methods within this discipline. For example, does fractal analysis of settlement plans require a data quality that would be practically unattainable in archaeological settings? But also inversely, as it is impossible to explore all potential applications within the framework of one doctoral thesis, suggestions for future research are provided in the last chapter.

1.3   Defining hierarchies

The term *hierarchy* is central to the present study. Though commonly used in daily speech, defining the word is not as straight-forward as one might think, so some clarification on how it is understood here might be needed. In a volume dedicated to exploring the meanings and uses of hierarchy as a study object within a range of natural and social sciences, Denise Pumain provides a panorama of definition nuances, but also highlighting the characteristics that are commonly found in most cases (Pumain 2006). Among these characteristics are:

- A pyramidal organisation of elements, ordered by a very unequal size distribution of a certain quality or variable, from a few large elements on top to many small elements at the bottom

- When seen as a system, the whole is constituted of sub-systems, which are again constituted of sub-sub-systems, and so forth. These can either be ordered into clearly distinguished levels (stratified), or in other cases be scaled in a continuum (branched or tree structure)

- In physical, biological and social hierarchical systems, the structure is often accompanied by a flow of energy, material, information or control in one or both directions between the top and bottom levels

Hierarchies are found in humanly constructed classification and taxonomical systems, where morphological distinctions are considered more important or fundamental at the higher end of the hierarchy, while being more detailed or specific at the lower end. Many hierarchical social systems, like religious (from Greek *hieros* – sacred, and *archê* – government), military or corporate organisations, include strongly reinforced regulations of subordination, which in modern society has led to somewhat negative connotations to both hyper-rational and despotic rule(Pumain 2006:5–6). While one prevalent explanation for the frequency of hierarchical structures in nature and society is indeed that they "represent the best solution for many optimisation problems" (*ibid*. p.7), that does in no way mean they need to be consciously planned. On the contrary, in most cases hierarchies seem to emerge spontaneously, often from growth processes with systems

splitting into sub-systems once they reach a certain critical size limit. There is also no compulsory link between social hierarchies and despotism, as it matters little to the overall structure whether the top element is elected for a limited period or born into an inherited leading position. More detail on how hierarchical structures emerge and how they can be described as fractals, is given in Chapter 2. A further discussion on the specifics of social hierarchies is given below in Section 1.4, and on the differences between spontaneously emergent versus consciously planned structures in Chapter 7.

A possible confusion with a somewhat different meaning of the term hierarchy should however be mentioned already here. If hierarchical structures are abundant in nature in both physical (inert) and biological systems, social hierarchies on the other hand – understood as intra-species populations of individuals organised in pyramidal hierarchical, i.e. multi-level relationships to each other – seems to be almost exclusively found among human groups. At the same time, a different type of hierarchy is frequently described by biologists which is common among animals, namely *dominance hierarchies*, also known as pecking order (Strauss et al. 2022). These structures are hierarchical in the sense that there is difference in rank between group members, and they also seem to emerge spontaneously in the animal populations where they are described. But unlike social hierarchies, these are purely *linear*, in the sense that each group member is situated in rank above one part of the group and below the rest, so that the whole group forms a rank chain in the form A > B > C … n. The rank of an individual will typically decide their access to food and reproduction relative to the other members, and may be settled and resettled in a number of ways depending on the species and population under study, but typically involving some level of violence or threat and subordination in face-to-face encounters (Strauss et al. 2022).

A classic example – perhaps most of all in popular culture – is dominance hierarchy among wolves, led by an alpha male (e.g. Cafazzo, Lazzaroni, and Marshall-Pescini 2016; Packard 2003). Though it has been much discussed whether or not this trope model actually fits wolves (see Mech 1999; Muro et al. 2011), any reported dominance hierarchies among larger groups of wolves and stray dogs are linear rather than pyramidal, even when they are illustrated as pyramidal (e.g. Fig.1 in Rodríguez et al. 2017). Similar social organisation systems are found among a wide variety of species – mammals, birds, fish, particularly but not only among group-hunting

carnivores – and are generally interpreted as an evolutionary mechanism (Strauss et al. 2022, with references). Cases of branching, multi-level social hierarchies among animals are on the other hand extremely rare, but have been reported to operate among hamadryas baboons in Ethiopia, with "clan leaders" forming relays of information flow and decision making between the "one-male units" within a total population ("band") of about 200 individuals (Schreier and Swedell 2009). Eusocial insects like ants and wasps provide a more well-known example of hierarchical organisation among animals (Shimoji and Dobata 2022), also indicating – if it should be necessary – that it is not a matter of cognitive abilities or brain size, but rather of social function (e.g. building a hive or a village together) and population size above a certain threshold.

The reason to dwell upon this qualitative distinction between linear and pyramidal hierarchies, is that the prevalence of the former in nature is sometimes put forward as an argument for social hierarchies among hunter-gatherer groups in the Palaeo- and Mesolithic [find precise reference here –> Graeber, if else strawman, Hayden 2007, pp. 231-6] **fill in this one**. While it is a good point that there is no reason there couldn't be linear dominance hierarchies among small forager groups, and that such systems hardly can be described as egalitarian by those who live them, we cannot assume that pyramidal social hierarchies have always been part of human culture, but rather that they – much like agriculture – at some point came to be as historical phenomena.

Bottom line hierarchy: pecking order/linear dominance hierarchy may be prevalent in any part of human prehistory, but could also be very hard to study from material remains. Branching (i.e. fractal) hierarchies as well as social stratification should leave material traces, and are most probably related to population size (see below), and thus less likely to be found within pre-Neolithic societies. They should be regarded as historically situated phenomenons – much like agriculture, pottery, writing or the steam engine – justifying the search for and explanation of their possible origins (*contra* e.g. Graeber and Wengrow 2021). The issue of the universality or particularity of social inequality and hierarchy is discussed in more detail **in the final chapters of this thesis (specify).**

- Inequality Kohler and Smith (2018), Midlarsky (1999), Price and Feinman (2010), Price and Feinman (1995). The difference between the concepts of social hierarchy and social inequality lies in the former relating to social and/or political organisation and the dis-

tribution of power, while the latter to economy and the distribution of wealth or income. Societies where these two are completely unrelated may be theoretically conceivable – that is, where wealth does not entail power and *vice versa* – but historically they have tended to go hand in hand, albeit not necessarily in a straightforward way. HERE discuss a bit on delegated power, democracy, taking turns etc., and such systems being impeded by wealth (lobbying, corruption++). Keep it simple, point is to prepare for the interpretation of Trypillia mega-sites as *both* egalitarian and hierarchical (and maybe LBK as *both* unequal and unstratified..).

1.4 What is social hierarchy?

- Political assumptions – all hierarchical social structures are not despotic top-down rule. Democracies can also be very hierarchical. Matter of scale rather than political system Graeber and Wengrow (2021), Zimmermann (2012), p. 255. But, tendencies? Use Pumain (2006), also Furholt, Grier, et al. (2020). Use Redhead and Power (2022): status and leadership, multiple overlapping networks. Emergence of hierarchies: "Positions in the dominance hierarchy is determined by a combination of attributes of individuals, stochastic processes, and social context" Strauss et al. (2022).

- Nested and non-nested social hierarchies, hierarchical hunter-gatherers? Hamilton et al. (2007), Whitridge (2016) (or use in Chapter 11?)

- Biologically defined thresholds to group size? Dunbar's number and controversies, Dunbar (2022), West et al. (2023) (add published papers). Scalar stress Johnson (1982), Alberti (2014), Zhou et al. (2005). Also Carneiro (1986), Dubreuil (2010)

- Temporal dynamics of social hierarchies: cyclicity (Peters and Zimmermann 2017), sawtooth waves (Scheidel 2017), punctuated equilibrium [Gould (2007); zimmermann?]. Archaeo. example South Sweden Neo/BA: Nordquist (2001). Transitions villages to urban: Birch (2014)

- Tools for classifying societies, or evolutionary model? Discussion of Johnson and Earle (1987), Testart (2005), Service (1971), Earle (1997).

    – Lineage and Chiefdom societies Earle (2002), Sahlins (2020)

An additional category of social structure, which has had a certain success in archaeology, was described by Claude Lévi-Strauss as *house societies* (Lévi-Strauss 1982). Here a house is a social unit, often centred around a material estate but also involving titles, heirlooms, land ownership, rights to hunting grounds etc. and where membership is not determined from genealogy in a systematic manner, as is the case with more regularly structured lineage or clan societies. House membership may be gained through more competitive social action, privileging those that possess the resources to engage in activities such as gift exchange and *potlatch*-style feasting. This opens up for more complex configurations, and house membership will often entail a certain level of prestige. Inheritance may furthermore follow (male) descent or (female) affinity – that is from grand-father to grand-son via mother – in a pragmatic way depending on which option is in the best interest of the house, as long as it can be justified in more or less precise kinship-related terms. A segment of society of varying size will not afford to partake in this competition, and as a result, house societies can be viewed as being in a somewhat unstable transition state between lineage and class societies. Lévi-Strauss originally pinned the term on Kwakiutl society in the American north-west coast (as described by Franz Boas), but argued for its generality by associating it with the feudal system of medieval Europe. The concept has since been applied to a wide range of cases from ethnography (especially in the South-East Asian and Pacific regions), as well as to prehistoric contexts **insert examples**, reflecting a relevance that extends far beyond European feudalism. Ian Kuijt (2018) recently provided an interpretation of Neolithic Çatalöyük where he proposed that

Refs. house societies [Joyce and Gillespie (2000), Bnf; Beck (2007); Kuijt (2018); Lévi-Strauss (1982); Bickle et al 2016 (reload in Zotero)].

-

- Anti-evolution critique in Yoffee (1993), Yoffee (2005), Fontijn and Brück (2013), Kienlin and Zimmermann (2012), Lund, Furholt, and Austvoll (2022), Furholt, Grier, et al. (2020).

Review in Dubreuil (2010). Anarchistic critique/heterarchy: Crumley (1995), Haude and Wagner (2019), Graeber and Wengrow (2021). Paleaeolithic social inequality: Honoré (2022) ++.

- I need to address gender inequality somewhere, mainly to say that this thesis is not about gender inequality. Start from Gaydarska et al. (2023), and cite A. Augereau for the LBK.

1.5   Main findings here?

Maybe leave to the end, and fill in, like abstract.

1.6   Research ethics

- Social complexity and evolution. Are less complex societies simple? Is that a bad thing?

- Open science and open-source scripts

- Terminology and spelling (British English for text. For geographical place names, Slovak special characters are kept as far as possible, even though it can be a pain in the xxx to render in Rmarkdown on Windows OS, and the 2010 Ukranian National transliteration system with only ASCII characters and no soft sign)

- Abstracts in Slovak and Ukranian (and not only in Norwegian)

1.7   Structure of the thesis

This thesis is structured as a monograph in four parts. In the first part the overall framework of the study is exposed, with the general introduction above, the overarching theoretical framework in Chapter 2, and the background of the study material in Chapter 3. Parts II and III are devoted to each their methodological approach to the material: Part II to the study of hierarchy in size distributions, and Part III to the quantification of image textures. Each of these parts consists of

three chapters, the first of which – Chapters 4 and 7 – exposes the theoretical and interpretative background of the applied methods and their relevance to archaeology. The following chapters – Chapters 5 and 8 – detail the technical specifics of the two approaches and their implementation in this study, and the last chapters within these parts – Chapters 6 and 9 – provide the actual analyses and summaries of results. In Part IV, the findings are summarised and further discussed. Chapter 10 gives an attempt of interpreting the results in the context of the culture-historical setting of the European Neolithic, while Chapter 11 reviews the possibilities and limitations of the fractal analysis framework in archaeology. Concluding remarks and suggestions for further study are given in Chapter 12.

Some readers might react to an apparent deviation from the academic tradition of devoting a separate chapter to research history. This is a deliberate choice, not to suggest that historiography is unimportant, but rather as a result of the fundamentally interdisciplinary scope of the study. In fact, there is very little extant history of applying fractal analysis in archaeology – the few studies that, to my knowledge, have been done in this direction are discussed primarily in Sections 4.3 and 7.3. Fractal analysis itself holds a research history of its own (see Section 2.2), and so does the study of Linear Pottery (Section 3.2) and Trypillia societies (Section 3.3), not to mention the general study of social complexity in prehistory (mainly Section 3.1). In short, instead of trying to shoehorn these parallel histories into a clearly delimited but rather hybrid chapter, I have opted for what I believe to be a more useful approach, namely to fit them in more seamlessly where they belong, in the various associated theory and methods chapters.

An additional note should be made here regarding the writing style of the different parts and chapters. It is my belief that a major obstacle for fractal analysis methods to become more integrated into the standard tool kits of archaeological research, is the excessively technical nature of much of the associated literature. archaeology as a discipline remains profoundly rooted in the Humanities, as seen in the inbuilt structure of teaching, research and funding institutions in most (at least European) countries. Fractal analysis is derived from pure mathematics, and most applications so far have been developed within the natural sciences (**ref necessary?**). Archaeologists who are trained within a humanistic scholarly tradition cannot be expected to hold a skill level of mathematics more advanced than what is achieved in high school, and code programming

is hardly taught at all within the walls of Humanities faculties. Technical details regarding the methods and analyses applied in this thesis are therefore – as much as possible – limited to the devoted Chapters 5 and 8, and readers who are interested in these may also refer to the online code repository for more details and reproducibility **cite repository**. In the rest of the thesis I have opted for a more narrative approach, in an attempt to invite a somewhat larger audience of archaeologists into the fascinating complexities of fractals.

END Chapter

Here I skipped a line.

Here I added a backslash.

Figures and tables with captions will be placed in `figure` and `table` environments, respectively.
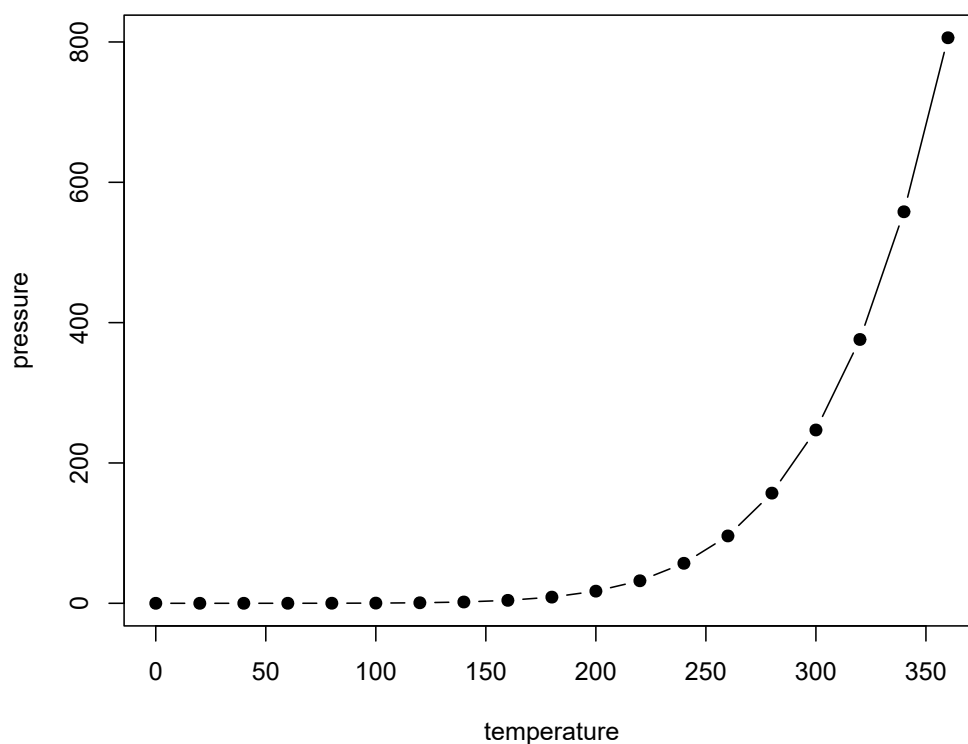
Figure 1.1: *Here is a nice figure!*

Reference a figure by its code chunk label with the `fig:` prefix, e.g., see Figure 1.1. Similarly, you can reference tables generated from `knitr::kable()`, e.g., see Table 1.1.

You can write citations, too. For example, we are using the **bookdown** package (**R-bookdown?**) in this sample book, which was built on top of R Markdown and **knitr** (Xie 2015).

Table 1.1: *Here is a nice table!*

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
5.0	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa
4.9	3.1	1.5	0.1	setosa
5.4	3.7	1.5	0.2	setosa
4.8	3.4	1.6	0.2	setosa
4.8	3.0	1.4	0.1	setosa
4.3	3.0	1.1	0.1	setosa
5.8	4.0	1.2	0.2	setosa
5.7	4.4	1.5	0.4	setosa
5.4	3.9	1.3	0.4	setosa
5.1	3.5	1.4	0.3	setosa
5.7	3.8	1.7	0.3	setosa
5.1	3.8	1.5	0.3	setosa

Test: Morgan (1965).

Test: Mandelbrot (2021). jløkj fdsasdfadfasdfadf new text

Chapter 2

Theoretical framework: Complexity and Fractals

2.1 Very short introduction to Complexity Theory / Dynamical Systems Theory

Lit. use Daems (2021), Baden and Beekman (2016), Ross and Steadman (2017), Smith (2011) and Bentley and Maschner (2008).

For Dynamical Systems, use Devaney (2020), but don't go into detail.

Describe complexity, dynamical systems, chaos, feedback loops, criticality, emergence. Scale West (2017) NA

Social complexity is more than social hierarchy, and all societies – whether they are classified as simple or complex within a social evolutionary framework – can arguably be studied as complex systems, since they always consist of various sub-systems and populations of individuals acting and interacting in a variety of ways (Daems 2021:6). The sense of social complexity that is traditionally understood in archaeology, derived from social evolution theory, is somewhat narrower and relates more to the specific characteristics of organisational scale (i.e. hierarchy) than to complexity per se. Studying the scale of social organisation from a complex systems

perspective rather than from a social evolutionist one holds several advantages:

1. The switch from discrete evolutionary stages to continuous spectra allows for more nuanced evaluation of the society in question, avoiding false binaries like simple-complex

2. The complex systems approach arguably has a stronger explanatory power than the more classificatory social evolution schemes

3. Complexity theory offers a better alternative for ethical reasons, as it avoids the underlying colonial and eurocentric connotations associated with classifying societies into simple and complex

As a further elaboration on the latter point, an analysis of a society from a complexity theoretical viewpoint is not a matter of establishing whether or not the society can be characterised as a complex system. When the scale of social organisation is the object of study – as in the present thesis – the word hierarchy is both more accurate and, if not neutral, at least more balanced than social complexity, since it is not obvious whether or not it is a good thing for a society to be characterised as hierarchical. The study of the dynamics of social hierarchy over time is thus not a story of progress, as the 19th century studies of social evolution too often were.

[@hamilton2007].

Mention the most common applications in archaeology, as discussed in Daems (2021), p. 13: network science, settlement scaling, cultural evolution and agent-based modelling.

2.2   Very short introduction to Fractals and Fractal Analysis

Lit. use Mandelbrot (2021), Falconer and Falconer (2013) (general), also Brown and Liebovitch (2010), Brown et al. (2005) and Diachenko (2018) (for archaeology)

The term *fractal* was pinned by the French-American mathematician Benoît Mandelbrot (1924-2010), from the latin word *fractus* meaning broken or irregular, to describe patterns that because of their apparent limitless complexity defied concise description within the framework of Euclidean

geometry (Falconer and Falconer 2013:116–20). Such patterns – both theoretical and empirical – had been described and analysed by mathematicians and researchers within other disciplines since the end of the 19th century, but were mostly regarded as curiosities and exceptions, and Mandelbrot was the first to link all these previous studies within a unified theoretical framework (Mandelbrot 1975, 2021).

In one influential paper, drawing on previous work by mathematician Lewis Fry Richardson, Mandelbrot (1967) argued that a rugged linear feature like a coastline cannot be fully described through traditional geometry with a set of line segments, since this would result in a curve of infinite complexity. More importantly, he showed that the traditional measure of lines – the length – will inevitably depend on the scale of observation when applied to a coastline. If measured in kilometres, a coastline will always appear shorter in total length, than if it is measured in metres, since smaller bays and inlets can then also be accounted for. But this phenomenon continues seemingly without limit, since the same coastline measured in centimetres will appear much longer, and in millimetres far longer again, and so on. Length as a measure of rugged linear features thus seems inadequate, which may become a problem in practical settings when comparing coastlines between countries that operate with different measurement units and procedures. The same problem occurs when describing irregular patterns in the plane (like island or continent outlines) with area or in three-dimensional space (like clouds or galaxies) with volume. As a solution, Mandelbrot proposed the use of the *fractal dimension* as a descriptive tool for characterising such patterns.

Fractals as hierarchy

Self-similarity and scale invariance

Processes/mechanisms that produce fractals:

- Cascading bifurcations and confluences (splitting or merging - tree structure/arborescence/branching, and relation to size. Terminology borrowed from biology and fluid dynamics (including turbulence/turbulent flow)

The role of randomness - tidy and messy fractals (romanesco broccolis are not more fractal than regular broccolis, only more regular. Also: a system may be both random and deterministic, often

only depending on scale of observation (chaos).

The relationship with (self-organised) criticality and chaos: deterministic *and* unpredictable

Fractals embedded in

- Space (hence "fractal geometry"): geomorphology, plants, ocean and wind currents, galaxies, also human constructed features (see Chapter xx for details)

- Time series: earthquakes, finance (not applied in this study, though should be done later)

- Networks/abstract: hierarchical organisations, income distributions, word counts, 1/f or pink noise, www. Barabási and Albert (1999) etc. (see Chapter xxx for details). Fractal social networks: West et al. (2023).

- Pure mathematics: Julia and Mandelbrot sets, strange attractors (don't go into details!)

No, not everything is fractal: e.g. Central Limit Theorem

Fractal analysis for studying irregular phenomena (methods described in more detail Chapters xxx), and thus as a tool for quantitative empirical research.

2.3   Very short introduction to micro-macro approaches in social theory

- Lévi-Strauss and Structuralism

- Braudel and World Systems

- Giddens and Structuration Theory

- Delanda and Assemblage Theory

For social hierarchies, he refers to Weber's classification of legitimation strategies, as being founded on sacred tradition, personal charisma or rational bureaucracy (see also Graeber and

Wengrow 2021, pp.). Delanda furthermore places Bourdieu's concept of *habitus* on the mid-range between micro and macro processes, as an explanation of social action that is not entirely individual dependent, but not emergent properties of society-as-a-whole either.

- Latour and Actor-Network Theory

What these approaches all have in common, is that they are entirely qualitative (**check**).

- That's not a problem in itself.

- Quant approach is both possible for the stated purpose, and desirable for reasons of comparative analysis.

- Data deluge (refer to chap. on geomagn data). The goal here is to establish a quantitative framework for studying social complexity and hierarchy in archaeological/prehistoric settings. Further articulating fractal analysis with existing social theoretical approaches is not the primary goal here, as it could constitute a separate research project. In the present thesis, bla bla.

END chapter.

Math can be added in body using usual syntax like this

p is unknown but expected to be around 1/3. Standard error will be approximated

$$SE = \sqrt{\left(\frac{p(1-p)}{n}\right)} \approx \sqrt{\frac{1/3(1-1/3)}{300}} = 0.027$$

You can also use math in footnotes like this[1].

We will approximate standard error to 0.027[2]

---

[1] where we mention $p = \frac{a}{b}$

[2] p is unknown but expected to be around 1/3. Standard error will be approximated

$$SE = \sqrt{\left(\frac{p(1-p)}{n}\right)} \approx \sqrt{\frac{1/3(1-1/3)}{300}} = 0.027$$

Chapter 3

Material and data: social complexity in the European Neolithic

3.1 Studying social hierarchy in archaeology and prehistory.

Grave goods

Burial monuments

Architecture Rathje and McGuire (1982) – social stratification (qualitatively) through architecture and burials (and life expectancy). Short but to-the-point article.

The denominator problem

The use of ethnography

Other approaches (osteological, isotopes, craft specialisation **refs**)

This project: house-size distributions and settlement layouts (details in subsequent chapters), just very short argumentation.

Comparative approach: Neolithic technology (not bronze axes), wood and wattle-and-daub architecture, (near) complete settlement plans/extensive documentation

For this project I opted for the use of houses and built environments as proxy for social hierarchy. In this way I hope to largely avoid the denominator problem associated with burials. While it is

true that for many archaeological culture groups in late prehistory habitats are poorly preserved and hard to discover, leaving us still with a limited understanding of them (the case in many Michelsberg, Corded Ware and Bell Beaker groups, only to mention a few **check**), in groups where habitats are well preserved, there is little reason to suspect that the available record would not cover the whole range of social statuses if these societies were hierarchical. Unlike burials, every individual in a sedentary society – with few exceptions like homeless persons in more recent urban contexts – will normally have at least one fixed place to stay overnight, and these homes will in most cases be constructed within the same fundamental framework of techniques and building materials, depending more on culture specific traditions and environmental factors than on social status (**citation?**). As an example, in a society where mudbrick is the main building material, like in the Neolithic Near East and Anatolia, nearly all constructions are made in mudbrick, regardless of the social status of the inhabitants. In Europe north of the Alps, wattle-and-daub construction was the almost exclusive building technique for any architectural feature from the early Neolithic until the Roman conquest, and well into the Middle Ages north of the *limes*. One can of course enumerate exceptions, but more importantly houses are in any stratified society also a marker of social status, which can be exhibited in a range of ways, from decorations, use of more precious raw materials as well as size. That is precisely the reason for using houses as a proxy for social status and hierarchy in archaeological settings. But the point here is that there should be little taphonomic differentiation between groups of high and low status within a given archaeological context, at least in prehistory, and at least not as much as can be expected for burials, meaning that we can expect to find samples of houses that are representative of the social structure of the archaeological culture in question. On sites where there is taphonomical loss of architectural structures, as long as the overall building tradition is homogeneous, there is no reason this loss should affect one segment of the society more or less than others.

Some caveats do remain, however, for the use of houses as proxy of social status. Firstly, there may be a documentation bias favouring larger houses, since they may be easier to discover both during excavation and in remote-sensing surveys (**citation?**). In samples with very skewed house-size distributions, there may also be a further taphonomic bias towards large houses, since smaller houses – being far more numerous – are statistically in greater risk of being affected by post-depositional disturbances. Both of these biases are hard to evaluate empirically, though computer

modelling could potentially give indications of their importance. This, however, is not within the scope of the present study.

A second, and maybe more important issue, is that of the contemporaneity of houses. When the goal is to investigate the social structure of a settlement as reflected in its architecture – be it through the size distribution of buildings or their spatial layout – all the analysed features should ideally have been in use at the same point in time. This is however very hard to achieve in most archaeological settings, and many researchers choose to either ignore the issue, or to accept a temporal resolution that is far wider than what their research questions should logically allow for (Perreault 2019). One way of limiting this problem is to select study samples with little to no stratigraphic overlap, which might indicate short occupation span, though as shown below this indicator can also be deceiving. For both study areas selected for this thesis – the Linear Pottery in south-west Slovakia and the B2/C1 Trypillia of central Ukraine – settlement plans show very little overlap between houses, even though some of them probably developed over more than three centuries, as shown by radiocarbon dates and modelling (see below). Such settlement plans may be impossible to differentiate into separate coeval time samples without precise dating of construction and abandon of every individual house, counting in the thousands on the Trypillia mega-sites (an alternative method is presented only for the Linear Pottery settlements in Section 3.2). On the other hand, the fact that there is so little overlap between houses despite temporal differences clearly illustrates how these settlement plans emerged over time, not as a *tabula rasa* in each generation but rather with new constructions respecting the location and orientation of older ones long after abandonment. Though such practices are indeed interesting, it is not at all obvious, however, to which extent they may reflect or even relate to social factors such as hierarchy. With settlement types with much higher degree of stratigraphic overlap, like the tell sites of the Balkan and Near Eastern Neolithic traditions, it may be easier to distinguish more or less coeval occupation phases, but they are again harder to document extensively – because of the high density of archaeological finds and features, excavation surfaces typically cover only very small portions of the settlement, while remote sensing performs less well and not allowing for temporal disentanglement of constructed features (**citation**). In any case, the issue of temporal resolution of the data and its influence on analytical results is crucial, and will be discussed repeatedly throughout this thesis, with a summary in Chapter 11.

Comparative technology

Meaning of "house"

3.2 The Linear Pottery culture complex

General intro to the culture

Formation, characteristics – striking difference from earlier Mesolithic cultures, hence the term "Neolithic Revolution" pinned by Vere Gordon Childe for the seemingly abrupt transition to agriculture in continental Europe with what he called the Danubian Ia **ref**. Whittle (2022), check with Trigger!

expansion trajectory: east: Saile (2020),

Questions of contact and even creolisation with local hunter-gatherer groups in the margins of the Linear Pottery settlement area – Limbourg and La Hoguette pottery groups in northern France and Belgium, Swifterbandt in the Netherlands??, Ertebølle in northern Germany??, Bug-Dniestr in Ukraine and Moldova (Saile 2020) – as well as with other already established Neolithic groups in central France (Roussot-Larroque 2022) and the lower Danube (Saile 2020), remain hotly debated

End of LBK as collapse? The whole culture as an adaptive cycle: Gronenborn et al. (2014).

Organisation of Linear Pottery society: egalitarian or hierarchic? Cemeteries: Jeunesse (2022), check Gaydarska et al. (2023) and Augereau (2021). Linear Pottery architecture and house construction

origin of the longhouse: Bánffy and Höhler-Brockmann (2020), Coudart (2015), Last (2015)

house typology, construction, and use: Modderman et al. (1970), Coudart (1998)

house orientation: "origin" hypothesis: Bradley (2001)

Number of coeval houses in LBK villages, evaluated to a mean of 5, sd. 3 (Coudart 1998:91). Furthermore, based on ethnographic analogies she estimated a typical Linear Pottery village of 3 to 8 houses (of which 1 or 2 longer than 22 m), to include on average approximately 150-260 inhabitants, though fewer (up to around 80 people) for the smallest hamlets of 3 houses.

Patrilocality in LBK based on Sr isotopes: Hrnčíř, Vondrovský, and Květina (2020) (read this more carefully), and other data (check) Schiesberg (2016).

The Žitava valley and research project

3.3   The Cucuteni-Trypillia culture complex

General intro to the culture

Recent dating, see Harper et al. (2023), Shatilo (2021).

Trypillia architecture and house construction

Pickartz et al. (2019), pp. 1649-50: "unburnt" and "eroded" are misleading labels, since highly burnt and even vitrified material regularly found during excavation of these. Probably reflecting houses that are constructed with less daub, without an oven installation, and/or without raised clay platform, thus causing significantly weaker magnetic signals. These houses may in fact represent entirely different social functions from regular houses, but there is as for now no consensus as to what this may be (e.g. artisanal, storage etc.)

Trypillia social organisation: current debate

Side note on Varna: Lichardus (1991; also Gimbutas), as discussed in Kadrow (2013), proposed that the social inequality observed at Varna in the mid-5th millennium, was already then adopted from pastoral North Pontic steppe cultures (*check this and delete if unsure*); cf. Chapman et al. (2006) – three social levels interpreted from grave goods, but not clear if these are really discrete or constructed on a continuum. The term *fractal* is here used purely as a metaphor for complex personhood structures, rather in opposition to the described social hierarchy.

The B2/C1 and the mega-sites of the Southern Bug – Dnieper interfluve The urbanism debate: Comparison with Uruk in Müller and CHECK; low-density urbanism for Nebelivka: check for refs. in Ohlrau (2022),p .91 cc. Mega-sites as *anomalocivitates*: Ohlrau (2022) and Sindbæk (2022).

3.4 Reading site plans from geomagnetic imagery

Caveats: fill in here.

Pickartz et al. (2021) fds

3.5 Synthetic data

And why I'm not (this time) relying on ethnographic data.

Don't go into technicalities here, just the reasoning.

Part II

Size distributions

Chapter 4

House sizes and social meaning

4.1 Interpreting house-size differences

In archaeology there are two recurrent and seemingly contradicting assumptions underlying interpretations of house-size differences within a society. In studies where the goal is to provide population size estimates, this is often calculated from living area, where the square metres per inhabitant is modelled from ethnographic analogies. The population of a village is then found by summing together those of every house (e.g. Coudart 1998:79–80). A simpler and probably more simplistic version of this is to consider a m^2/inhabitant proportion that is constant no matter the size, and thus calculate the population directly from the total living area of the village (**give examples**). This mean value is generally obtained from multiple ethnographic parallels. With this assumption – that every inhabitant requires a similar amount of living space – the size of a house effectively reflects its number of inhabitants. The second assumption, which is more frequently seen in studies focussing on wealth inequality and social stratification, is that house-size differences are expressions of differences in some sort of wealth or power (**give examples**). In this view, a larger house would have belonged to a wealthier household, capable of procuring more raw materials and activating a larger labour force for its construction and maintenance, in which case there would be significant disparities in living space per individual. Even though there is not necessarily any contradiction between these two interpretations from an anthropological point of view, most archaeologists seem to be unable to consider both possibilities simultaneously (Wilk

1983). From a methodological point of view, each of these assumptions will tend to mask our ability to see traces in the archaeological data relating to the other assumption – that is, with existing methods we cannot convincingly provide estimations of both population size and level of social inequality from the same data, even though house sizes frequently form the basis of both argumentations. Far from proposing a solution to this issue, my argument here is that a variety of social institutions known from ethnography and historical sources can explain some level of correlation between the two variables. Dowry and bride price are geographically and temporally widespread practices that link number of offspring (daughters, sons or both, depending on cultural context) with wealth. Clan leaders may draw upon kinship ties and dependencies in order to obtain the workforce needed to construct a larger house (**citation needed**). In agricultural societies, land ownership is often seen to correlate with household size, since land owners tend to attach workers and servants to them, people who themselves in turn tend to come from landless households (Netting 1982; Wilk and Rathje 1982:629). Nevertheless, if the notion of wealth is at all to be applied meaningfully to non-capitalistic societies, it should designate cases where there is significant and persistent material disparities between members of a population, and not simply point to different household sizes where wealth is proportional. House sizes could thus potentially reflect a somewhat more complex culture-specific interplay between household size and wealth, meaning that for a given house size, one could assume a range of possible values of the two parameters (Figure 4.1). This relationship between household size, wealth and house size should be studied more in detail empirically through the available ethnographic data, rather than reducing its complexity to a mean surface area per inhabitant for the entire population. Such a study, however, lies beyond the scope of the present thesis. Here I will largely leave aside the question of population and household size, focussing on distribution types of house-size data, arguing that the most unequal distribution type considered here – the power-law distribution – is unlikely to emerge only from random differences in household size and standard marital patterns, favouring thus interpretations relating to systemic wealth and/or power differences.

Even though the goal here is not to investigate household sizes but to focus on the material aspect of house-size distributions, some fundamental issues of terminology should be addressed. The use of the word *house* (in the wider material sense rather than the Lévi-Straussian sense, see Section 1.4) is indicative of an underlying assumption that the building in question was in use primarily
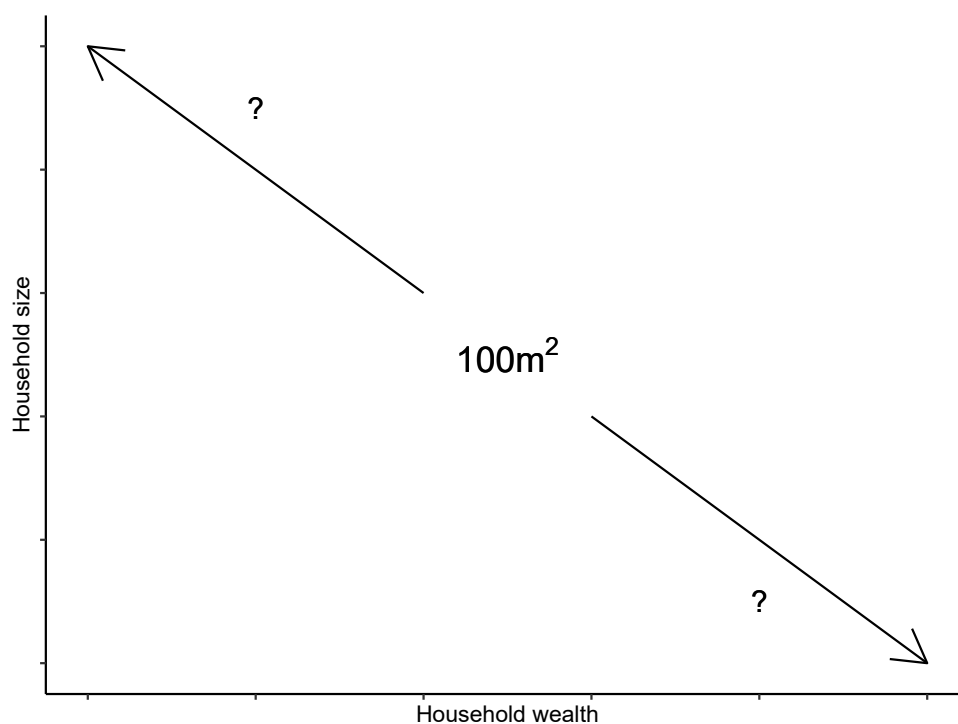
Figure 4.1: *House size can be interpreted in terms of number of inhabitants (household size) or material wealth of the inhabitants, but the exact relationship between these two variables remains poorly understood and is probably both complex and culturally contingent. Any given house size within a distribution can thus result from a combination of effects from the two. In many cultural contexts household size and wealth may furthermore be directly correlated*

for domestic purposes – essentially, a fixed architectural unit where someone would spend their nights at least most of their time, and in many cases also cook and eat their main meal during the day – though proving this directly is not always straightforward in archaeology. For the cultural contexts discussed here – the Linear Pottery and Trypillia groups in the Neolithic-Chalcolithic – there is however little evidence for buildings with specific non-domestic (e.g. economic, religious or administrative) purposes, with the probable exception of the so-called mega-structures or as-sembly houses in the Trypillia mega-sites, which are discussed more in detail below. This lack of evidence does not imply that there was no specialised economic, religious or administrative activity in these societies, as ethnography and history clearly shows that such activity must reach a certain degree of specialisation before it materialises in distinct buildings devoted exclusively for these functions. Artisans, shamans and chiefs could be specialised to some extent but still perform their activity at their domestic home or more diffusely outdoors or without any fixed location (e.g. Costin 1991:25; Kahn 2021). It is in any case of common usage to speak of houses

when discussing architectural units in Neolithic Europe and other prehistoric contexts, maybe because of a lack of a better generic word, but this usage should not prevent archaeologists from recognising other non-domestic functions of buildings whenever there is evidence for it. The *household* is furthermore the designation of all the people, genealogically related or not, usually (though not necessarily) living under the same roof or within the same architectural unit, constituting a functional whole economically and socially, and potentially including more than a single family unit, as well as servants or slaves, depending on the context (Wilk and Rathje 1982:620). The emphasis for defining a household is thus more on its economic and social function than on co-residence and kinship relations, which are somewhat more variable. In the following, I go into some more detail as to how household organisation as well as wealth are known to influence house size.

4.1.1   House size and household organisation

The focus on houses and households in archaeology – as opposed to larger units of analysis like whole settlements, cultures and periods – started to attract momentum by the end of the 1970s, with the work of Wilk and Rathje (1982) often cited as the original manifesto of its validity and importance, pinning "Household Archaeology" as an independent genre of study. In their view, the household could be understood as the most abundant activity group in any society, despite considerable variations as to its importance relative to other types and scales of social groupings. They defined it by their *social*, *material* and *behavioural* constituent elements, that is, its members and their relations, the dwelling and other possessions, and the activities it performs. The material element is of course the only one directly accessible to archaeologists, and the task for researchers would be to reconstruct the social and behavioural elements from the material. This realisation led to ambitious comparative projects of mapping the variability of houses and their occupants in different parts of the world and socio-economical contexts, with inferences onto prehistoric contexts built upon existing ethnographic literature (e.g. Murdock 1949) as well as new ethno-archaeological observations (Blanton 1994; e.g. Wilk 1983).

One of the main characteristics of households is their *size*, which, according to Wilk and Rathje (1982), to some extent is determined by the scale of the production activities that fall within their

organisational sphere. In societies where large scale complex tasks necessitating the simultaneous cooperation of many hands are organised at the household level, the optimal household size will be accordingly large. Such activities can include agricultural tasks like irrigation or terracing, as well as house construction. However, when large-scale activities are necessary only once or twice a year, as with seasonal large game hunting or intensive fisheries, they tend to be organised at a community level by many households working together temporally. In such contexts, households can be smaller as their daily activities can be performed by a smaller number of people (Wilk and Rathje 1982:623; see also Hamilton et al. 2007 CHECK#). Furthermore, ethnography has repeatedly shown that large households performing complex activities often need a head coordinator for the activities to run smoothly. However, the set of activities organised at the household level in a society and thus determining the optimal household size, will affect all households similarly unless there is some economic differentiation between them. In societies where households are economically more or less self-sufficient, their size differences should be expected to be random (i.e. normally distributed, see Section 4.2.1).

Kinship studies within anthropology have shown over the last decades that in most societies, contrarily to common misconception, kin affiliation is *not* simply a matter of biological relatedness. In an attempt of grouping together all possible justifications for kinship ties, Marshall Sahlins (Sahlins 2013) defined kinship as "mutuality of being" – that some real or imagined substance is shared, and that this substance is not necessarily genes, as is mostly the case in modern Western societies (with adoption as the main exception). A variety of non-genetic foundations of kinship are widely accepted in different cultures, like sharing of name, time or place of birth or childhood, food source, shared experiences, blood ties, and so on. At the same time, archaeology as a whole has arguably been very slow in taking this diversity of kinship configurations into account, far too often taking for granted the modern Western (especially post-war 20th century) ideal of patrilineal nuclear families as the default configuration for all of human history (Ensor 2021). That being said, anthropological kinship studies such as that represented by Sahlins typically shows little concern with material culture, and are mostly silent on the question of who – at the end of the day, quite literally – sleeps under the same roof, making it hard to interpret domestic architecture in terms of kinship structure and social organisation. The recent comeback of kinship studies in archaeology has been far more focussed on linking isotope and aDNA data with social structures,

largely fuelled by the rapid developments of the related methodologies.

• Kinship and households: who lives in a house? Ensor (2013), Ensor, Irish, and Keegan (2017).. archaeology: , Madella (2013), Joyce and Gillespie (2000), Carsten and Hugh-Jones (1995), Hofmann and Smyth (2013), Kramer (1982)

4.1.2   House size and wealth

Wilk and Rathje pointed to the important role of households in inter-generational transmission of wealth in many societies (1982:627–31). Specifically, following Murdock (1949) and other ethnographers, they argued that as populations grow, land tenure tends to institutionalise at the household – and later individual – level around the moment when agricultural land becomes more scarce than labour. Before this – in pioneer phases of agricultural development – land is readily available and if there is any concept of land ownership at all, it usually lays at the community level. Once the agricultural land in a region is saturated, rights to use it will tend to be transferred within households, and children from households with extensive rights have a greater incentive to stay within or close to it, while children from households with less land rights are more likely to emigrate. Further population pressure will tend to limit partition of inheritance between siblings, so that land ownership over time is transmitted within a smaller and smaller fraction of the population. This situation is also shown to entail stratified (and parentally arranged) marriage, further entrenching social stratification.

Maya here!

Kahn (2021) describes an example from eastern Polynesia of agricultural expansion and intensi-fication developing over 400 years from the initial colonisation before social stratification starts to become materialised in differential architecture. In later phases, she links the appearance of specialised buildings devoted to communal assemblies, rituals as well as residential and ritual buildings for elites with the emergence of complex chiefdoms. Though quantitative measures of house size are not presented in the study, it clearly demonstrates that elite architecture was consistently larger than that of commoners in the prehistoric Society Islands. However, she also identifies types of houses of special economic function like workshops, and points out that these

can be hard to distinguish from common residential houses archaeologically simply by looking at size (Kahn 2021:90–91). In such cases – where there is a range of non-residential building types – it may be of little use to analyse house-size distributions as single blocks. It may then be necessary to first group buildings into functional types based on internal structure and finds inventories, given that this information is available. This issue concerning the prerequisites of data input for the analyses done in this thesis, is further discussed in Chapter 11. It should also be noted that interpretations regarding the social structure of pre-contact Polynesian societies are greatly aided by near-contemporary written accounts from the first European explorers in the region, and house floor levels are often well-preserved due to the relatively recent dates of the structures (10th to 18th centuries CE) – both factors being in stark contrast to the central European Neolithic, where the spatial organisation and sizes of houses are often nearly the only information available.

Regarding the domestic architecture of clan leaders and chiefs, several studies have highlighted how elites both have the material means and the socio-economic incentives to build houses that are more monumental than those of commoners HAYDEN1997++. Others have warned that this should not lead archaeologists to systematically interpret large buildings as evidence of elite-based top-down social hierarchies, since there are also ethnographic examples of monumental buildings being constructed collectively by more egalitarian communities for assemblies or other communal activities. Referencing a range of ethnographic and archaeological examples, Goodale, Quinn, and Nauman (2021) point out that the act of constructing a monument may be organised and to some extent coerced by a leader, or equally by a collective group. In both cases the intended function of the finished building will serve in the interest of the one or those who initiated the project. This is furthermore valid at all social scales – the constructuion of a residential building is initiated and driven by those who wish to live in it, and the construction of a ritual building for the village like an assembly house is initiated and driven by the entire village. Furthermore, the construction of a palace in a state society is initiated by the leader and financed through the tribute collected throughout their effective territory. Consequently, buildings should be expected to somewhat reflect the scale of their social importance through the level of effort put into their construction, so that in a hierarchically organised society (be it top-down coercive or bottom-up collective), this effort should also be hierarchically distributed in its buildings. I

argue here that building size is the best (and often the only) available proxy to this construction effort (see Section 4.2.4 below for more discussion on this point).

Do clan leaders have bigger houses? check Haude and Wagner (2019), Bradley (2013), Wilk (1983). Smith (1987).

- General tendencies for Pueblo contexts: Dohm (1990)

P. Květina and J. Řídký point out both architecture (construction, size, orientation) and settlement layout as possible distinctive features between Big Men (achievement-based) and Chief societies, arguing that the former type may be recognised by a dispersed intra-settlement layout combined with uniform architecture, while the latter type would tend towards more regular settlement layout and more marked differences in architecture (Květina and Řídký 2019:13). EXPAND UPON THIS.

- Schiesberg 2010 2016, go through refs in Zotero, family size and houses for the LBK.

Functional difference:

- Ethnography of initiation houses, communal/assembly houses, ritual houses, including Barley (2011), Godelier (1986), Wilk (1983), Fraser (1968), Haude and Wagner (2019)

Some factors other than household size and wealth have been suggested and empirically reported to systematically influence house size (Wilk 1983:101). Among these are:

- Mobility – Seasonally mobile groups tend to build smaller dwellings than more sedentary groups (Porčić 2012)

- Post-marital residence – Houses in matrilocal societies tend to be larger on average than in patrilocal societies (Hrnčíř, Duda, et al. 2020; Porcic 2010)

- Climate – Houses are smaller in cold circumpolar or mountainous regions because of the cost of heating

- Duration of residence – Households that have been established at a location for a long time or are well integrated in the community tend to have larger houses (Wilk 1983)

- Material use and technology – All building materials have associated constraints and costs, and innovations can lead to larger constructions at lower costs

Of these factors, only the duration of residence will have a direct influence on house-size difference *within* communities. Mobility, residence patterns and climate are more constant factors affecting entire communities and will thus affect the average house size, but not the level of inequality. Access to building materials and technology on the other hand may be differentiated in stratified societies, and lead to unequal house sizes as a materialisation of wealth. A related factor that may lead to a certain under-representation of the wealthiest households, is the construction of multi-floored houses, unless this feature is recognised archaeologically and included in the calculation of house sizes. For the case studies presented in this thesis however, building materials and techniques are not noticeably differentiated between small and large houses, and the possibility of multi-floored houses is only a minor issue that has been discussed in Chapter 3.

Lastly, in his ethnographic comparison between egalitarian and more ranked contemporary Maya village communities in Belize, Wilk (1983:111–14) also pointed to social norms potentially *preventing* differences in public display of wealth even when such differences existed. He linked such norms to the openness of the village economies – in closed self-sufficient villages there was greater mutual dependence between households, and wealth display was strongly discouraged, potentially sanctioned with witchcraft, whereas in villages with more open economies household wealth was more readily displayed through house size and the quality of construction materials.

4.2 Distribution types and their underlying mechanisms

The main question in this part of the thesis is on the nature of house-size differences in the studied contexts, and to explain them either as being due to random fluctuations or as a material expression of a more structural inequality, or something in-between. In statistical terms, this is a question of distributions. A distribution is a mathematical model of the spread of data along

a variable or axis, and it can be modelled on empirical data as a succinct description, or used to predict unobserved data (e.g. future developments, or, in the case of archaeology, data that is lost to taphonomy). In theory, there is an infinite number of possible distributions, as there is no limit to how many parameters one can include to fit the data. It is however generally considered good practice in statistics to limit the number of parameters and identify the simplest possible model that gives a good fit, since a larger number of parameters can often lead to a better fit, but at the same time be harder to explain in terms of underlying mechanisms. Adding many parameters only to achieve a marginally better fit is referred to as *overfitting* (#ref?), and for most real-world contexts there is a limited number of model types that can be considered reasonable candidates. In cases where we are interested in inequality or uneven distributions of data (so-called *skewed* distributions), the most likely distribution models are those that can be described as *heavy-tailed*, meaning that on typical graphical representations (like density/PDF plots or histograms) they will show a characteristic stretch of some of the data towards the right end of the x-axis, while most of the data remains on the left side (Figure 4.2). The opposite orientation is also possible in theory, in which case the distribution can be referred to as left-tailed. It is important to keep in mind however, that not all distributions are heavy-tailed, and that other distributions also model the spread of data across a variable, but result from very different underlying mechanisms. Identifying the most likely distribution model for a data series is therefore crucial for understanding how the data could have been generated. And though it is true that one model type can have multiple different explanations – in an archaeological setting for example, many different social behaviours can lead to the same material outcome, an issue known as *equifinality* – excluding one or more model types for the observed data can help limiting the number of plausible interpretations considerably.

Distributions can be represented graphically in a number of ways, most commonly in a coordinate system with the variable on the x-axis and its density, or *probability density function* (PDF), on the y-axis (Figure 4.2). Often, and in many of the sources cited in this part of the thesis (e.g. Clauset, Shalizi, and Newman 2009), the PDF is denoted by $p(x)$, reading as *the probability of x*, or simply $f(x)$ (*the function of x*). It gives the probability for a drawn sample (X) of falling within a given arbitrarily short range of the distribution, written $Pr(x \leq X < x + dx)$. By definition, the area between the PDF curve and the x-axis sums to 1. For reasons that are further discussed in Chapter 5, heavy-tailed distributions, and power laws in particular, are instead

often represented with their *cumulative distribution function* (CDF), which is the integral of the PDF (inversely the PDF is the derivative of the CDF; Figure 4.3). Similarly to the PDF, the CDF is often denoted as $P(x)$, depending on disciplinary tradition. It indicates the probability of a random sample value being equal to or lower than the function value, or $Pr(X \leq x)$. Furthermore, the specific version of the CDF used for plotting heavy-tailed distributions, is the *complementary* or upper tail CDF (sometimes referred to as the *survival function*, or denoted cCDF), which is $1 - CDF$ or $P(X \geq x)$, indicating the probability that a random sample is higher than the function value. Both axes on such cCDF plots are traditionally set on logarithmic scales, usually $\log_{10}$ for readability. To avoid confusion, in this thesis I refer to PDF and cCDF for density and distribution functions respectively (except in equations, where I use $p(x)$ and $P(x)$ respectively), and all PDFs are plotted on linear scales and cCDFs on $\log_{10}$ scales, unless otherwise stated. Apart from scales on plots, whenever I refer to logarithms (i.e. in calculations), I imply natural logarithms, that is ln or $\log_e$ where the base number $e \approx 2.718$.
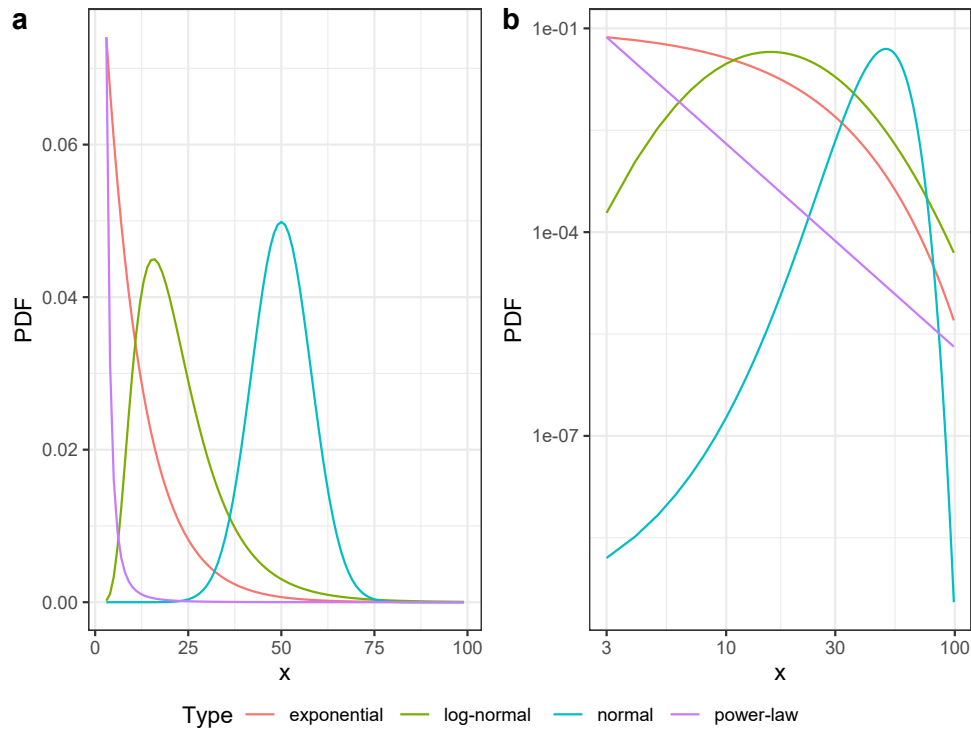
Figure 4.2: *Example curves of the probability density function (PDF) of four common distribution types: normal (blue, $\mu = 50$, $\sigma = 8$), exponential (red, $\lambda = 0.1$), log-normal (green, $\mu = 3$, $\sigma = 0.5$) and power-law (purple, $\alpha = 3$, $x_{min} = 1$), in linear (a) and logarithmic scales (b). Parameter values are arbitrary and x-axis is truncated at $2 < x < 100$ for readability. The power-law distribution is the only to form a straight line when both scales are logarithmic*
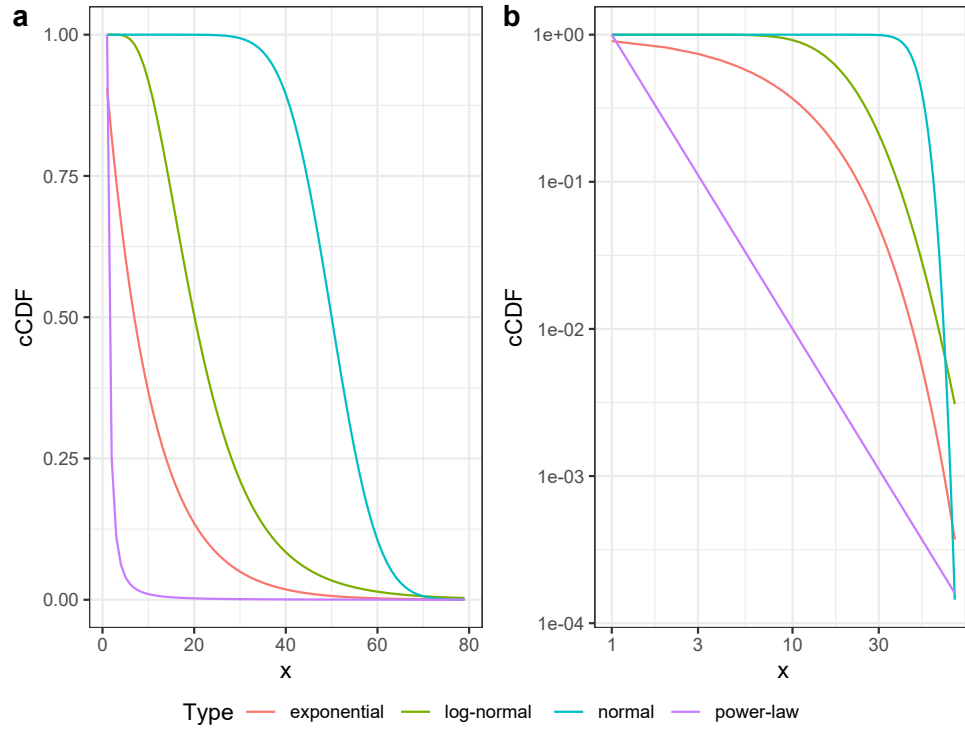
Figure 4.3: *The same distributions as in Figure 4.2, but with the complementary (right-tail) cumulative distribution function (cCDF), in linear (a) and logarithmic scales (b), with $0 < x < 80$ for readability*

In the following, I present briefly the main distribution types that will be discussed further in the following chapters, with special focus on the *power-law distribution*, which is the model type associated with fractals and structural hierarchy. All of these distributions are modelled on continuous univariate data series (see Johnson 1994 for more detailed presentations). Even though their mathematical definitions may seem complicated to non-initiated readers, most of the distribution types discussed in this thesis are readily implemented in standard statistical software, including Microsoft Excel, allowing for a more straight-forward use of them. For this thesis, I used base *R* functions for calculating PDFs and cCDFs, and for random number generation for normal, log-normal and exponential distributions (R Core Team 2023), and equivalent functions from the *poweRlaw* package for power-law and Weibull/stretched exponential distributions (Gillespie 2015).

4.2.1 Normal distributions and the Central Limit Theorem

One of the distribution models that are the most commonly referred to and well-known, is the *normal distribution*, also known as the "Bell Curve" due to the characteristic bell shape of its PDF (Figure 4.2a), or Gaussian after mathematician C.F. Gauss who contributed to its exploration in the early 19th century. Its characterising parameters are the *mean* (denoted μ, *mu*) and *standard deviation* (σ, *sigma*), and the PDF is defined mathematically as

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} \tag{4.1}$$

where e (*Euler's number*) and π (*pi*) are mathematical constants (Johnson 1994:80–88). This distribution is centred around its mean, with dwindling amounts of data spread outwards to either side. The mean – commonly known and widely used in daily speech – is the sum of all observations divided by number of observations, or

$$\mu = (\sum_{i=1}^{n} x_i)\frac{1}{n}. \tag{4.2}$$

The spread of the data from the mean is defined by the standard deviation, which is the square root of the mean of all squared deviations from the overall mean, or

$$\sigma = \sqrt{(\sum_{i=1}^{n}(x_i - \mu)^2)\frac{1}{n}}. \tag{4.3}$$

Mathematically, the square of the standard deviation, σ^2, or *variance*, is simpler, but since it is also less intuitive, I refer here to the former whenever possible. The standard deviation can be thought of as the mean of all deviations, positive or negative, from the overall mean.

The great importance the normal distribution has to a wide range of phenomena is explained through the *Central Limit Theorem* (CLT), according to which the sum of random variables tends to a normal distribution as the number of variables increases towards infinity, under certain conditions (Johnson 1994:85–88). More specifically, if $X_1, X_2, \ldots, X_n$ are independently drawn and identically distributed (condition referred to as *i.i.d*) random variables or samples, their sum

will be normally distributed in the limit as n tends to infinity, or

$$\lim_{n\to\infty} \sum_{i=1}^{n} X_i = \mathcal{N}. \tag{4.4}$$

The sum may be standardized in some way to avoid infinite numbers, but in practice n will always be limited, so the resulting distribution will always also be approximately normal at best. The original distribution of the random variables (i.e. that of the *population*) does not need to be normal for the theorem to hold, but can be any distribution as long as its variance is finite. The normal distribution (referred to as the *sample distribution* in this context) is related to the population distribution of the random variables (samples), in that the mean of the population equals the mean sum of the sample distribution divided by n, and the standard deviation of the sample distribution, referred to as the standard error s, equals the standard deviation of the population divided by the square root of n, or $s = \frac{\sigma}{\sqrt{n}}$, meaning that s is reduced with higher n (i.e. larger sample size). A common variant of the CLT is that the distribution of sample means, rather than sums, are normally distributed, in which case the mean of the sample distribution equals the mean of the population distribution.

In more laypersons' terms, the total weight of a box of strawberries is the sum of the weights of the strawberries it contains (subtract the weight of the box itself, and assume the same number of berries per box for the sake of argument). In the packaging facility, a large population of strawberries are continuously distributed randomly into boxes of the same size. The berries in the factory (the population) have weights that follow some distribution with limited variance – some are bigger than others, but there are upper and lower limits to how much a strawberry can weigh – and all the berries in the boxes are drawn from this same population (so identically distributed). New berries are shipped to the facility all the time, and berries are not sorted according to size but randomly mixed, so if one box by chance gets filled with only very large berries, that does not affect the weight of subsequent boxes (each box is the sum of n independently drawn sample berries X). Under these circumstances, the weight of boxes of n berries (i.e. the sum of the berries' weights, the sample distribution) will be normally distributed by the Central Limit Theorem.

Some details here are crucial: the distribution of the strawberry weights themselves will, with higher n, tend towards that of the population, which is not necessarily normal. The theorem is only valid for summary measures like sums or means, and not for the observations directly. If the strawberries at the facility are mostly large (heavy) and with only smaller proportions of small berries, this skewed distribution will be reflected in strawberry boxes of a certain size (n), but since the influence of this distribution is the same on all boxes of the same size, their overall weights (or mean weights) will be normally distributed. Furthermore, if we draw some boxes from one producer and some from another producer who has significantly larger or smaller berries, the samples are then drawn from different populations and are thus not identically distributed, and the sample distribution will not necessarily be normal (colloquially referred to as "comparing apples and oranges"). Similarly, if there is an overall trend of boxes becoming heavier over the season, then samples from across the season will not be normally distributed. If however, we compare the mean box weights from multiple seasons, these will again be normally distributed, unless there is also a multi-year trend of strawberries becoming larger or smaller. In other words, the i.i.d. condition of the CLT means that the samples are drawn at random, with no important underlying trends. As a side note, there is a substantial body of research on the conditions under which normal distributions can emerge *without* the i.i.d. condition being met (see Johnson 1994:87–88 for details and further references).

Also important to note is that the number of samples X_n (boxes of sample size n in the example above) does not matter to the shape of the sample distribution, other than to the resolution of the curve or binwidth of the histogram when plotting, and the weight of a single sample can be modelled as a probability following a distribution. The normal curve of the sample distribution can be entirely defined by the population μ and σ, and sample size n. In many practical settings however, the parameter values of the population distribution (and even its type of distribution) are unknown, in which case more samples are needed in order to model a sample distribution and from there infer the parameter values of the overall population. This is the case when statistical tests like Student's t-test or ANOVA are applied for examining the relations between samples and populations.

In other cases, n (sample size) can be unknown, but assumed to be approximately the same for all

X_n (samples), like in the somewhat more abstract cases where each of the n variables contained in X_n are of different nature – or stated otherwise, when the size of X_n is the sum of many different and independent causes. In the case of a normal house-size distribution within a given cultural setting (like a village), one can assume that the same number of causes affect the size each house takes when constructed, but to varying degrees (colloquially we often say *factors* for such causes, though in this setting one should strictly speaking prefer *terms*, *summands* or *addends*, since they add up rather than multiply). Say that house size in a given context results from the cumulative effects of household size, inherited wealth, soil quality, exposure to sun, wind or flooding, artisan specialisation, raw material availability in the year of construction and many more variables. Each of these may have separate probability distributions – e.g. the wealth distribution may be heavy-tailed while household size may be normal and symmetric – but as long as the overall population, so to speak, of contributing causes is distributed in the same way to all households, and that none of the variables dominates the effect of the others, and the value of each variable is independent from the values of the other variables, their sums expressed in house sizes should be normally distributed by the CLT. Such a distribution is then an expression of *random difference*, and in the case of house sizes, there would be no particular reason as to why a few houses would be bigger than all the others, just as a few houses would also be smaller, while most would be centred around the mean size. But again, if samples are drawn from *different* populations, e.g. houses from different villages or cultural contexts – where the probability distributions of the underlying variables are categorically different – the i.i.d. condition is not met, and the distribution of house sizes is unlikely to be normal. Likewise if the population contains grades, i.e. is grouped, or where the different variables are correlated between them, e.g. if wealthier households are also larger households and have more access to raw material and better quality soils, and so on, their house sizes may also become disproportionately large and deviate from normal expectations.

A final caveat for normal house-size distributions that may be mistaken for skewed ones, is the case when there are no significant differences between house sizes, except for one that has a clearly different function – as in the ethnographic cases of community houses and men's houses discussed above. Then it makes little sense to interpret the resulting slight skewness of the distribution as a sign of social inequality in itself (notwithstanding gender inequalities). As a rule of thumb, to avoid such misinterpretations, it is useful to test whether isolating the single largest

house changes the retained model when performing distribution fitting.

4.2.2 Exponential distributions and constant rates of growth and decay

The exponential distribution is – next to the normal – one of the most widely applicable distribution models. In its simplest form, it is a function of a positive random variable x where some base number (usually $e \approx 2.718$, for compound continuous growth) is raised to the power of $-x$, in other words when x has the probability density

$$p(x) = e^{-x}.$$

This simple form is called the *standard* exponential distribution, and in most practical applications there will also be a *rate* parameter λ (*lambda*), so that the density function becomes

$$p(x) = \lambda e^{-\lambda x}. \tag{4.5}$$

In the case of the standard version, $\lambda = 1$ and can be left out. The negative rate in the exponent is the actual rate that determines the shape of the distribution, whereas the rate multiplier to the base is a *normalising constant* which assures that the area under the curve adds up to 1, and thus that the values shown on the y-axis are probabilities. This constant can be though of as the y-intercept of a linear model, since at $x = 0$ the function gives $\lambda e^0 = \lambda 1$. This is seen more clearly if we take the logarithm of the exponential density function, $p(log(x)) = log(e)\,(-\lambda x) + log(\lambda) = -\lambda x + log(\lambda)$, which is a linear model with slope $-\lambda$ and $log(\lambda)$ as y-intercept. As a rule of thumb, an exponential distribution can thus also be recognised as a straight line on a plot with one linear and one logarithmic axis. A wide variety of more complex forms have been formulated, and the distribution type is furthermore generalisable to both the Gamma and Weibull distributions (Johnson 1994:494–99). Note that the simple form presented here may also have more complex notations in specialised statistical literature, *cf.* Eq. 19.1 in Johnson (1994), which is equivalent to the density function above given that $\sigma^{-1} = \lambda$ and $\theta = 0$ (Clauset et al. 2009:664; see also Johnson 1994:522–23).

Exponential distributions have the highest probability (so the most data) at low values of x, and ever lower probabilities towards the right end of the curve (Figures 4.2a and (**ref?**)(fig:04-cCDF)a), with the rate of decrease determined by λ – the higher the rate value the steeper the curve falls off from left to right, and inversely low λ values give more heavy-tailed distributions. The type of setting which is most commonly modelled as an exponential distribution, is that of "events recurring at random in time" (Johnson 1994:494). If x represents the the duration in time between events (or duration of single events) that occur continuously and independently from each other, with a constant average rate of λ events per unit of time, it can be modelled as an exponential distribution. An important feature of the underlying process (a so-called Poisson point process), is that it is memoryless, meaning that the duration between events X_1 and X_2 does in no way affect the duration between X_2 and X_3, as all durations are drawn independently from the distribution with the same average rate λ. The example of such a process that may be the most familiar to archaeologists, is the radioactive decay of the ^{14}C isotope with its average decay rate $\lambda \approx 0.00012$ or 0.012% per year. With a rate that low and using years as the time unit, it is more useful and intuitive to work with the *half-life* measure, or the time it will take on average for the initial quantity to be halved, which for ^{14}C is about 5730 years. The half-life (the median of the distribution) is given by $\log(2)/\lambda$, solving for x in the cCDF function $e^{-\lambda x} = 1/2$, where the normalising constant in the PDF is replaced with 1 (and therefore omitted in multiplication) for the initial quantity or y-intercept. If we replace $1/2$ with the remaining proportion of ^{14}C in the organic material of an ancient artefact, compared to the expected amount in the same material when alive, we can use the same equation to solve for the approximate year when the organic material died, which is the principle behind radiocarbon dating. In a radioactive decay process, a total amount of individual unstable isotopes are present from the start (here the death of an organic material), and their individual lifetimes until decay are exponentially distributed. As an additional metric of exponential random variables, the mean μ or *expected value* is given as the reciprocal of λ, so that $\mu = 1/\lambda$ and $\lambda = 1/\mu$. The mean is larger than the half-life, and for ^{14}C decay, this corresponds to $\mu \approx 1/0.00012 \approx 8267$ years.

The exponential distribution can also model many processes that are closer to an everyday human scale than ^{14}C decay. Expanding from the example used for normal distributions, let λ be the average risk for a strawberry of being harvested within a week x. As the weeks go by (as x

increases), the cumulative risk of being picked grows exponentially, so that there are very few berries that are older than a few weeks, and the mean age of the berries in the field is $1/\lambda$. The berries are furthermore picked by a harvesting machine that is unable to aim for a certain size category, and the berries grow linearly (which may be a rather poor approximation, but for the sake of argument). The harvested berries are also continuously replaced by new berries which start growing at the same pace, so the field is always renewed. Under these circumstances, the lifetime of a strawberry during which it grows is exponentially distributed, and the probability of surviving x weeks follows Equation (4.5). The example illustrates how exponential functions model repeated multiplication or multiplicative processes, since the x in the exponent means "multiplied x number of times". The rate λ (technically $e^{-\lambda}$ in the case of continuous decay) is multiplied with itself x number of times (keep in mind that $(a^b)^c = a^{bc}$). Multiplying the rate for each new step in time means that the value of λ is applied to the current value of x rather than to the initial value. For example, let λ be $1/5$ or 0.2, so that there is a 20% risk of being harvested within a week, and thus 80% chance of being left in the field. The expected lifetime of a strawberry is then 5 weeks, and the probability of surviving 6 weeks or more is $P_X(6) = e^{-(1/5)6} \approx 0.301$ or 30,1%, while that of surviving 7 weeks or more is $P_X(7) = e^{-(1/5)7} \approx 0.247$ or 24,7%, corresponding to a relative decrease in probability of $\frac{(0.301-0.247)\,100}{0.301} \approx 18\,\%$, equal to $1-e^{-1/5}$. Note that using the rate directly as the base, rather than as exponent of e, will give the same results as a (discrete) geometric distribution, in which case the reduction would be exactly the rate between each period. The difference lies in what is termed compound and simple interest in economics. For all the purposes discussed in this thesis, the continuous exponential distribution (with $e^{-\lambda}$ as base to x) is deemed more appropriate than the discrete geometric distribution.

From an archaeological point of view, the essential point from the strawberry example above is that this process will also materialise in the size distributions of each single harvest, of all the strawberries at the depot, as well as the strawberries in a finished box for sale, all of which will be exponentially distributed (unless there is some additional sorting process involved). Of course, this does not change the fact that the box weights, as well as the mean strawberry size per box will be normally distributed, as previously shown. When it comes to house sizes, several scenarios involving growth could explain an exponential distribution. Let house size (x) be directly dependent on household size, so that each inhabitant has a constant average number of m^2 of roofed

space. Say that the households grow exponentially at some rate, that is they increase in size by a factor of λ_1 each cycle of some unit length (y). At the same rate, the households also extend their houses proportionally to their growth, or replace their house altogether with a bigger house, and lastly, for each new cycle a new household is added to the village with newcomers, so that the total number of households follows y linearly. If all new households start at some minimal size x_{min}, the size x of a house after y cycles will equal $x_{min}e^{\lambda_1 y}$, and we can solve for its age (time since establishment of the household) as $y = \frac{\log(x) - \log(x_{min})}{\lambda_1}$. Here I use x_{min} for θ in Johnson (1994), p. 494, following Clauset et al. (2009), p. 664. In other words, this context would generate an exponential house-size distribution of coeval houses in a village, where the largest houses are those of the households which were the first to settle in the village.

This model is of course not very realistic. For example, households cannot grow without limit, so the larger they become, the higher the probability that they split into two or more factions (see Alberti 2014 and Johnson 1982, and Section 1.4). The splitting of households itself can in fact also generate an exponential distribution. Let the households in a village grow linearly – say they each have a constant surplus of 1 person per year (persons who arrive or are born - persons who leave or die = 1) – but they also run a risk of λ_2 (e.g. of 5%) of disintegrating and being replaced by a minimal sized household (x_{min}) each year (y), no matter their current size. House sizes are then linearly correlated with their age (or time since establishment of the household) so that $y = x - x_{min}$ (disregarding the constant of m^2/inhabitant). However, over time, the probability that a household continues to exist without splitting, will decrease exponentially, and we can write the survival function for households (and therefore their sizes) as $P(X > x) = e^{-\lambda_2 y}$, or $e^{-\lambda_2 (x - x_{min})}$.

But again this model is not very satisfactory, particularly since it assumes purely linear population growth, which is unlikely in most cases. It would also seem likely that the probability of splitting of households would not be the same across the range of sizes, but rather be concentrated around some upper threshold, determined by the social structure between the inhabitants or by material or ecological constraints (or a combination). Common for both of the models above, is that one aspect is well described as exponential, but this aspect is only one out of many in the complex process which may underlie the house-size distribution of a settlement. Intuitively, it would also

seem strange to have the whole range of houses in a settlement to scale exponentially, since it would imply that single house sizes would be ever closer and closer to each other all the way down to the smallest house in the village. Adding the x_{min} parameter does change this situation though, as it could then apply to cases within some upper class in which household size and/or wealth would increase to some fixed rate over time, not affecting the size of the main part of the settlement's households, which could well be normally distributed. Or the two models above could be combined to one, pulling exponentially in opposite directions (note that the two rates, λ_1 and λ_2 are positive and negative respectively). In both cases, the resulting house-size distribution would no longer be exponential, and these common combination distributions – the log-normal and the power law, both of which are heavy-tailed – are presented in more detail below.

Another issue with exponential distributions resulting from growth or decay in time, from the archaeological point of view, is that the time-averaging that so often infiltrates our analyses because of the difficulty of distinguishing temporally coeval data sets, may very well influence the observed data distribution. This influence is much more challenging to evaluate theoretically however, so in this thesis it is instead addressed empirically through simulation in Section 6.1.

4.2.3 Log-normal distributions and Gibrat's law

One of the main candidate models for heavy-tailed continuous distributions is the log-normal, defined as a variable x of which the logarithm is normally distributed. Adapted from Equation (4.1), its density function can be written (following the notation in Mitzenmacher 2004:229) as

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma x}e^{-(\log(x)-\mu)^2/2\sigma^2}. \tag{4.6}$$

The μ (mean) and σ (standard deviation) parameters are usually understood as the equivalent values associated with the normal distribution of $\log(x)$. It can be thought of as a combination of the normal and the exponential distributions, and like these, it has been shown to apply well to a wide range of natural and social phenomena, from the growth of organisms in biology to the pricing of options in finance Johnson (1994). One implication of $\log(x)$ being normally distributed, is that the density curve of x will appear as a normal bell curve when plotted with

a logarithmic scale on the x-axis (Figure 4.5). With linear scales, the curve is skewed with the mode to the left and a tail of high values to the right. More technically, the (natural) logarithm of a variable (x) is the variable of exponents that may raise e to the values of x. A linear increment in a variable of exponents – say from 1 to 2 – will, with the same base, correspond to an exponential increment in powers (the result of exponentiation), as $e^1 \approx 2.718$ and $e^2 \approx 7.389$. Thus, if the exponents of e that correspond to the values of x are normally distributed, then x itself will resemble an exponentially stretched normal distribution, which is a log-normal distribution.

Since exponents represent repeated multiplication and normal distributions result from random additive processes (see above), log-normal distributions may be most easily understood as resulting from random multiplicative processes. The product rule of logarithms states that the logarithm of a product of numbers equals the sum of the logarithms of those same numbers. This can be expressed as $\log(ab) = \log(a) + \log(b)$. Then, since the sums of many random numbers are normally distributed according to the Central Limit Theorem, the logarithm of the product of many random numbers (that is, these numbers multiplied together) should also be normally distributed (see e.g. Newman 2005:347–48 for more elaboration).

The product of many random numbers is typically the result of a growth or decay process where the rate fluctuates randomly. A convenient example, following Newman (2005:348), is that of a financial investment. If an initial value (a) is invested in stocks that generate a return (e^λ) which fluctuates randomly from year to year with a finite variance, the return y after x years will follow a wiggly exponential curve, or $y = ae^{\lambda x}$ (Figure 4.4). After some years, the value of y will follow a log-normal probability distribution (Figure 4.5). If several persons start investing the same amount in stocks at the same time, then after a period, say of 10 years, most of them will have earned returns of comparable size, centred around some mean return, while the earnings of the top investor may be several orders of magnitude higher, simply by chance. This assumes however that everyone invests randomly in the stock market, which is rarely the case. More scrupulous investment strategies may reduce the effects of chance and thus the spread of final returns, but this effect may again be countered by the risk-willingness of investors. In either case the resulting distribution after a given time period will be log-normal, which explains why this model is a central tool in financial analyses (e.g. see Mitzenmacher 2004:236 for its use in the
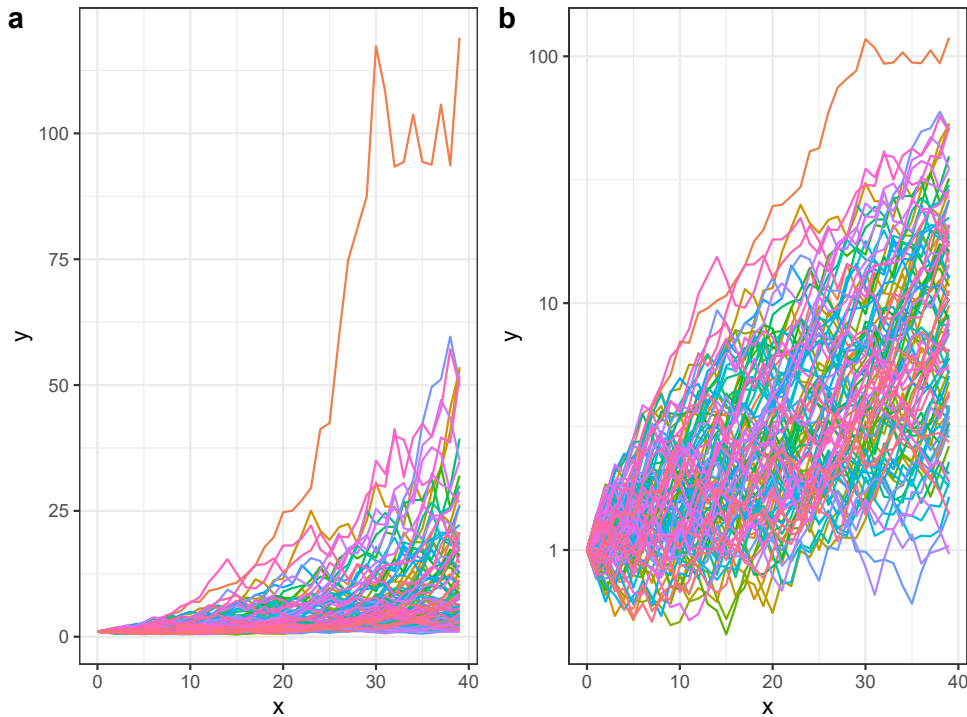
Black-Scholes option pricing model).

Figure 4.4: *Exponential distribution of $y = \lambda^x$ with rate (λ) fluctuating randomly and uniformly between 0.75 and 1.4, i.e. with a mean rate of approx. 1.08, over 40 periods (x) from an initial value of 1. The plots figure 100 individual runs of the distribution, with linear (a) and logarithmic (b) y-axis. Over time, y-values at any given x are expected to be log-normally distributed by the Central Limit Theorem*

This process is known as *Gibrat's law*, after the French engineer Robert Gibrat who was the first to demonstrate its wide applicability, offering a mathematical explanation of the skewed size distributions that had frequently been observed by economists (Gibrat 1930). Gibrat argued that the log-normal distribution was better fit for modelling firm sizes and salaries than the one already proposed by Pareto (i.e. the power law, see below), since it could account for the entire size range and not only the tail, and since it was theoretically better founded, as Pareto's original model was purely empirical. He termed the model the *law of proportionate effect*, since it emerges – as shown above – when proportionate growth rates are independent of absolute size. This essentially means that growth is exponential (size at any given time is multiplied with a rate, so proportional) rather than linear (additive), and that the range of possible rates is not determined by absolute size. Note however that there must be some randomness in the rate for the growth to result in a log-normal distribution. If the rate is *exactly* the same for all samples, or if it follows
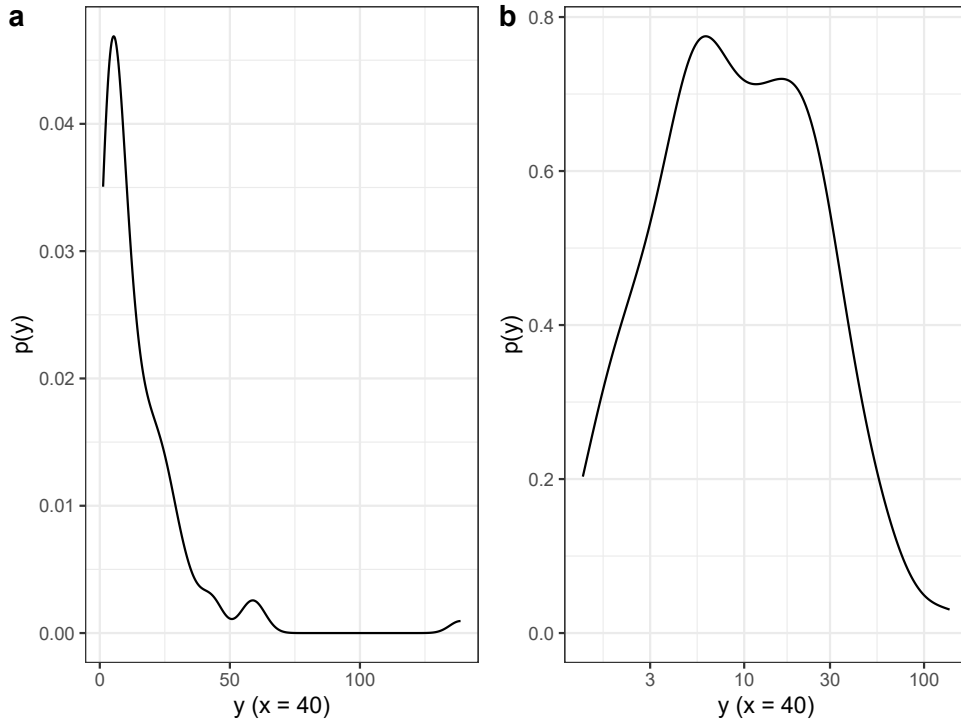
Figure 4.5: *Density function of y-values from Figure 4.4 at $x = 40$, with linear (a) and logarithmic (b) x-axis, following a typical log-normal distribution*

the exact same sequence of random rates, the initial distribution will remain the same over time, even though any initial spread will be scaled up or down according to the rates (Figure 4.6). In most economic settings, though there are overall trends that may affect everyone in a population (of firms, employees, cities etc.) at large scales, there are also many smaller factors that will affect individuals differently, causing random variation.

The modern financial market is of course not directly applicable to the Neolithic, but a number of multiplicative processes involving random fluctuations may also be relevant to Neolithic social structure or economy. One obvious example would be that of crop yields over time. Assuming that crop cultivation in a village is organised at a household level, and that households grow their crops at separate locations in the vicinity, random differences in soil quality, sunlight, water, exposure to disease and so on, would arguably generate normally distributed yields in one year (the yield volume being the sum of many random effects). But given that the yield the year after also depends on the current yield through the size of the surplus that will be available for sowing, a randomly large yield one year will have better chances of producing an even larger yield the
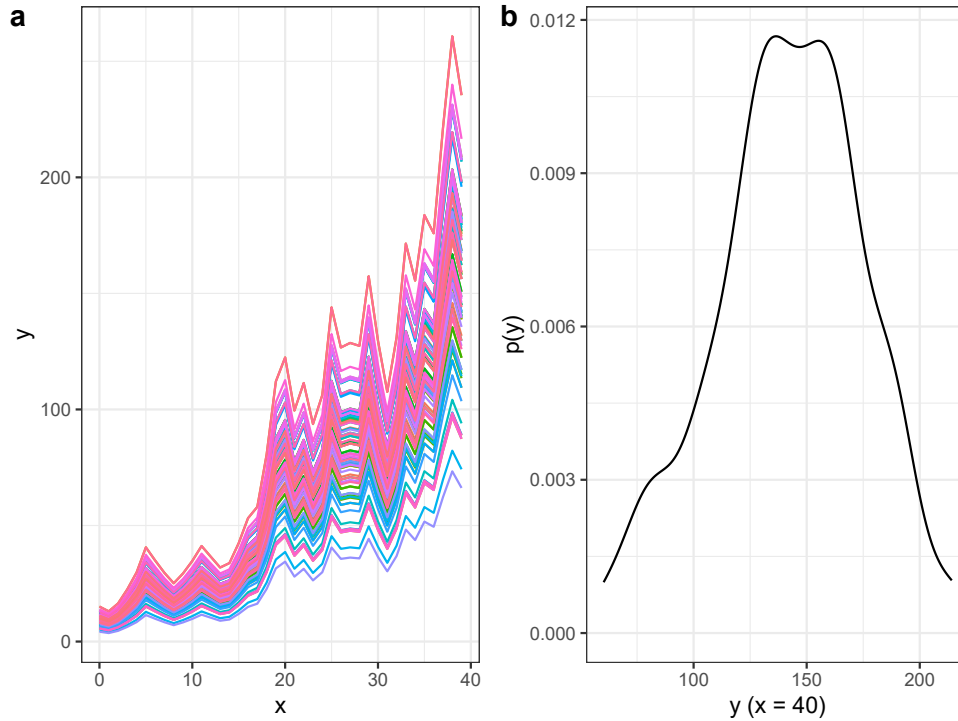
Figure 4.6: *a: Eponential growth of 100 samples drawn from a normally distributed initial population ($\mu = 10$, $\sigma = 2$), all following the same sequence of uniformly distributed rates ($0.75 < \lambda < 1.4$). b: Over time, μ and σ values change, but the distribution remains normal. Scales are linear*

next year, and so on, in the same way a large financial return of a lucky investor one year is more likely to reach an even higher return later if reinvested. Supposing again that house size reflects household size linearly, and that larger yields can sustain larger households, this process could explain the emergence of a log-normal house-size distribution in a village over time.

For Neolithic crop cultivation it is in many cases more reasonable to assume that cultivation took place very close to or even within the village, in which case the conditions for yield volume would be very similar between households (#ref needed?). Cultivation could also be organised collectively rather than at the single household level, or as a combination depending on the crop. In such cases, a good or a bad harvest would affect all households equally, and skewed distributions would not emerge easily (#Ref. to Kohler 2018). However, even in such scenarios, small random differences in the initial sowing volumes of individual households could grow exponentially over time and produce a log-normal distribution of household yields, though with smaller spread between the highest and lowest values.

A somewhat different mechanism for explaining skewed house-size distributions within Linear

Pottery settlements in particular, was proposed by Sara Schiesberg (2010). Assuming a stable population over time, with number of children per woman surviving to reproductive age being Poisson distributed with $\lambda = 2$ (with ever decreasing probability of larger numbers of surviving siblings), and the probabilities of having given ratios of male and female children following a binomial distribution, the combined probability of having *x* male children surviving to reproductive age would follow a skewed log-normal-like (though discrete) distribution. Schiesberg observed this theoretical distribution of male siblings to be analogous to the empirical house-size distribution of excavated Linear Pottery settlements on the Aldenhoven plateau in North Rhine-Westphalia, Germany, arguing that there were gaps in the continuous size distribution that fitted with the discrete limits between numbers of male siblings. The correlation was furthermore explained as a result of a mainly patrilocal residence pattern, where house size would be a function of number of sons in an extended family (patrilocality being the most widely accepted residence pattern for the Linear Pottery culture, see Section 3.2). The same pattern could have emerged as a function of number of daughters in the case of matrilocality. Though Schiesberg did not model the observed house-size distribution explicitly as a log-normal, her model and the underlying process closely resemble the Gibrat's law described above. A (discrete) Poisson distribution with low rate can well approximate a (continuous) exponential, and a (discrete) binomial can approximate a (continuous) normal – so a combination of the two (by multiplication of exponential and normal probabilities) will equally approximate a log-normal. Thus, simply by the social practice of patrilocal post-marital residence, a skewed house-size distribution would emerge spontaneously, without the presence of any additional structural inequality between community members. In this case, a log-normal distribution would be present in a settlement from its onset, and the skewness could or could not become more pronounced over time, depending on other factors like whether crop surplus would be distributed within the community or kept within households.

A last mechanism that is sometimes referred to in statistical literature as causing log-normal distributions, is that of random additive processes involving variables that by their nature cannot take on negative values, such as weights, heights or densities of physical entities (e.g. Johnson 1994:239). While (two-parameter) log-normal distributions only can have positive values (Johnson 1994:208), normal distributions will often also have positive probabilities below zero, depending on μ and σ values, and are then poor models of such quantities. A house, as an ex-

ample, cannot have negative size, but will always be larger than some lower threshold above zero, and should thus also be more adequately modelled as log-normal than normal, even if the distribution looks symmetrical. This will in turn allow for more accurate estimates of other derived parameters, like the confidence limits for the coefficient of variation. However, in cases of symmetrical distributions where μ is much higher than σ, there is little practical reason to prefer a log-normal model over a normal one, as the probabilities of values below zero will be infinitesimally low.

4.2.4 Power-law distributions, preferential attachment and hierarchy

A variable x is power-law distributed when its probability follows the power of itself with a fixed exponent α, so that $p(x) \propto Cx^{-\alpha}$ (Clauset et al. 2009:662). Such distributions decrease very quickly as x increases, and are thus highly skewed, but the probability never reaches 0 – it is said to be *asymptotic* – meaning that they are also very heavy-tailed (Figures 4.2a and 4.3a). Furthermore, a power law can only take positive values, and there is always a minimal threshold $x_{min} > 0$ above which the function holds. The exponent α, often termed *scaling exponent* or *scaling parameter*, will usually lie in the range $0 < \alpha < 3$, though values below 1 are considered rare special cases, when considering size or frequency distributions (Newman 2005:331–32). Low exponent values give more heavy-tailed distributions and *vice versa*, so power laws with a high scaling exponent (around 3 or above) are those that in practice will be more easily mistaken for other less skewed distributions like exponential or log-normal distributions (Figure 4.7. As with the previously discussed distribution models, the power law is also usually associated with a normalising constant (here denoted C), a factor that ensures that the area under the curve of the PDF sums to 1, and which here is defined as $(\alpha - 1)x_{min}^{\alpha-1}$ (Clauset et al. 2009:664–65). Applying the product rule of exponents, the power-law PDF or density function can be expressed as

$$p(x) = \frac{\alpha - 1}{x_{min}}(\frac{x}{x_{min}})^{-\alpha}. \tag{4.7}$$

In the cCDF or survival function, again the normalising constant is replaced with 1, but rather

than left out it is written in the exponential form $(\frac{x}{x_{min}})^1$ which equals 1 when $x = x_{min}$, so that

$$P(x) = (\frac{x}{x_{min}})^{-\alpha+1} \tag{4.8}$$

in the notation of Clauset et al. (2009), Eq. 2.6, equivalent to the more complex notation in Newman (2005), Eq. 4.

As shown in Figures 4.2b and 4.3b above, power-law PDFs and cCDFs hold the special property of appearing linear when plotted with logarithmic x and y axes (or equivalently when x and y values are log-transformed). This can be shown by log-transforming the simple functional form above, so that $log(p(x)) = (-\alpha)log(x) + log(C)$, which is a linear model with y-intercept $log(C)$ and slope $-\alpha$. Power functions are in this way similar to exponential functions, with the difference that the variable x is here in the base rather than in the exponent, causing the graph to be linear only when *both* axes are logarithmic (as opposed to one axis for exponential functions). Because of this property, the most common method for estimating α since the first formulation of the model and until recently (Clauset et al. 2009; Stumpf and Porter 2012) was to plot the data on logarithmic axes, perform a least squares linear regression and estimate the slope (see Chapter 5).

The main reason for the great attention that has been given to power-law distributions over the years (see below), is its property of *scale invariance*, meaning that the distribution will appear the same no matter the scale in which it is being observed (see chapters 1 and 2 of West 2017:1–78 for a general non-technical introduction to scaling in science). Figure 4.7 shows the relation between change in the variable being modelled (x) and its function value (here the PDF function) as a constant proportion of exponents (i.e. a linear relation between two vectors of logarithms). This means that for a power-law distributed variable, a change of scale (zooming in or out on the x-axis, changing order of magnitude no matter the base number) will always result in a proportional change of order of magnitude of the probability or frequency modelled on the y-axis. If x represents a size related variable, and y a frequency (probability in the PDF, rank in the CDF, or absolute frequency), this proportional relation in logarithms generates a self-similar hierarchy, where the same structure of sizes and frequencies is repeated across different scales (Figure 4.8).

Thus, even though fractal structures are most often associated with geometric shapes in physical space – like the shapes of plants, rivers or clouds – they can be equally present in non-geometric variables, like income or wealth, magnitude of events, sizes of cities or nodes in networks. Such structures are, as with physical fractals, recognised by their abrupt skewness or high level of inequality, by their hierarchically distributive mode of functioning, and their statistical signature which is the power law. This connection between power-law distributions in general and fractals was first recognised and expanded upon by Mandelbrot (e.g. 1997; short summary in 1982:341–48).
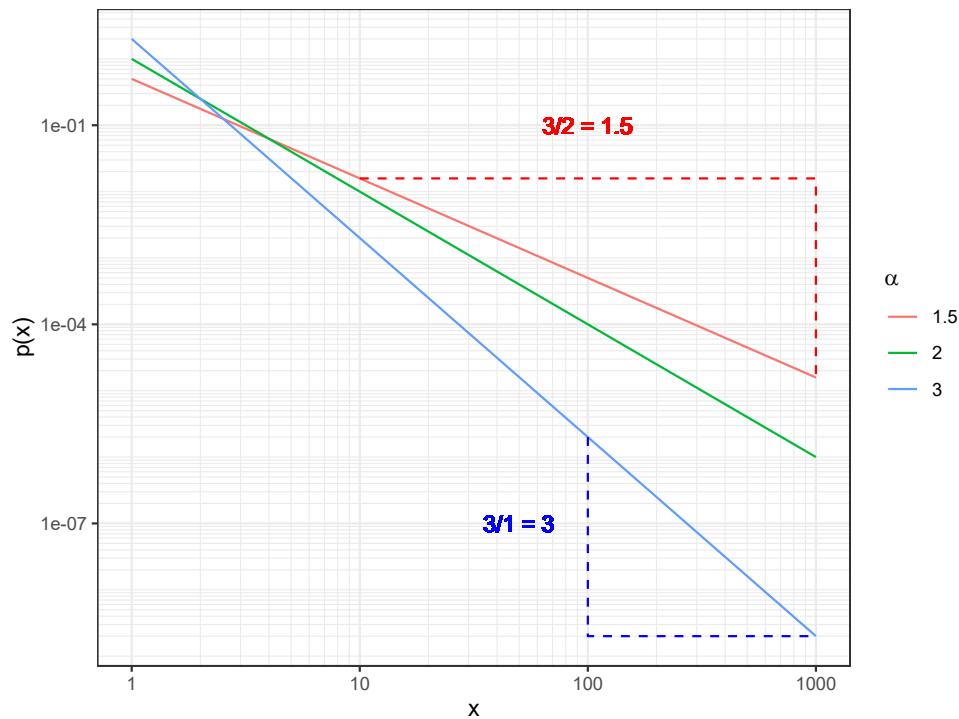

Figure 4.7: *Examples of power-law distributions with different scaling exponent (α) values on logarithmic axes, showing how this parameter reflects change over orders of magnitude. For a model with $\alpha = 3$ (blue), a decrease in probability p(x) of 3 orders of magnitude (powers of 10), e.g. from 0.1 to 0.0001 corresponds to an increase in the size of x of 1 order. For a model with $\alpha = 1.5$, the same decrease in probability corresponds to an increase in x of 2 orders of magnitude. The models appear linear in logarithmic space, but are in reality highly non-linear, as illustrated by the grid*

As already mentioned, the power-law distribution was first formulated by Italian economist Vilfredo Pareto as a model of wealth inequality (1896). Over the 20th century it saw an increasing number of applications in a wide variety of fields in natural and social sciences, modelling phenomena from magnitudes of earthquakes and sizes of moon craters to the intensity of wars,
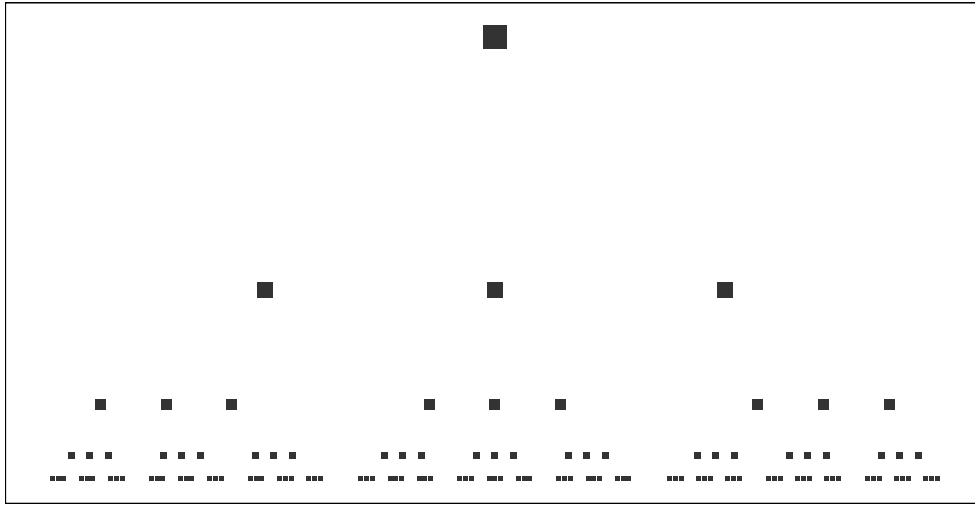
Figure 4.8: *A power-law distribution of sizes arranged in discrete levels, illustrating its characteristic scale invariance. From the largest element on top and downwards, sizes decrease while numbers (frequencies) increase, both exponentially but in opposite direction, generating a hierarchical fractal structure where the same shapes are recognised at different scales*

frequency of family names and citations of scientific papers (see Mitzenmacher 2004; Newman 2005 for detailed overviews). The idea that the power law could be a suitable model for city sizes was seemingly first proposed in a short paper by Auerbach (1913), though it is often attributed to Zipf (1949), through which it has become known to and sometimes applied by archaeologists (see Section 4.3). The first influential attempt at an explanatory model of power-law distributions – which at first were primarily descriptive, one of the main critiques against Pareto – was made by Udny Yule (1925), who sought to explain the observed size distribution of genera by number of species. Yule's model (here following the notation in Newman 2005:340–42) involves a set of n genera consisting of a variable k number of species each. During a discrete time step interval, a constant $m < n$ number of new species are added to the existing genera by speciation, so that some but not all genera will grow to $k + 1$ in each time step. The probability for each genus of receiving a new species in a given time step is proportional to k or the number of species it already includes, since a speciation is more likely to happen in an already large group of species

than in a small group. Finally, for each time step one new species is sufficiently different to be considered a new genus on its own, so that n increases linearly by 1 for each step. Under these conditions, over time k will follow a heavy-tailed distribution with a power-law tail – or strictly speaking a discrete version of it known as the Yule distribution (see Simon 1955).

The precise mathematics involved in this model are quite complex for non-specialists, especially in the version presented by Yule before modern stochastic theory was developed, but even more recent and concise formulations will involve some level of calculus (Mitzenmacher 2004:230–33; Newman 2005). The essential is however to note that the mechanism it describes – now known as *the Yule process* – is relatively simple and transferable to many natural as well as social settings. The heavy-tailed distribution of species will emerge even in the simplest scenario starting with one genus n_1 consisting of a single species ($k = 1$), and one added species to existing genera per time step ($m = 1$). Then the single genus will have probability $p(n_1) = 1$ of receiving the new species m so that it gets $k = 2$, while the new genus n_2 starts at $k = 1$. In the next time step a new species will be given to one of these two genera, but n_1 has twice as high probability of receiving it as n_2 – that is, probabilities are $2/3$ and $1/3$ respectively. In other words, the probability for any genus n_i with k_i species of winning the round (so to speak) and being attributed with the extra species, is given by $k_i / \sum k$, or the fraction of the total amount of species (in the given time step) that the genus already has. The main difference here with the above described Gibrat's law – which also involves proportional growth – is that in the Yule process the growth is not distributed evenly across the system. There is an additional selecting process that over time gives more to those that already have, thus according a disproportionate advantage to anyone who gets even the slightest advantage by chance from the offset – which is why this is often referred to as a *rich-get-richer* process when applied to economics.

In fact, the Yule process having been recognised more or less independently within a number of disciplines over the years, it has come to be known by a plethora of different appellations, sometimes hiding the fact that they describe similar underlying mechanisms. The term *feedback loop* is derived from acoustics, describing the bothering situation when the sound from an on-stage monitor feeds into the microphones and back to the monitors, and so on, thus very quickly generating a sound so strong and high that it overturns the system. The sound going into the

microphones is proportionally amplified to the volume of the sound coming out of the monitor, so that the stronger the input, the stronger the output and by consequence the next input, and so on. Another term, derived from attempts in sociology to explain the power-law distribution of citation frequencies of academic papers, and more generally the rewarding systems in academia, is the *Matthew effect*, alluding to a passage in the Gospel of Matthew (25:29): "*For unto every one that hath shall be given, and he shall have abundance; but from him that hath not shall be taken away even that which he hath*" (Merton 1968). A typical manifestation of such a process is when a prestigious research grant is attributed to a researcher based on academic merit, whereupon this in turn opens up numerous doors allowing for even higher career achievements, while the second best candidate, altough having arbitrarily close merits to the first before the grant, afterwards will have disproportional difficulties of following their pace in career development. For paper citations, one analogous explanation is that a paper that already has many citations, is more likely to be found in literature searches and be cited again than a similar paper with less citations (Newman 2005:341). In other words, the probability of a new citation is proportional to the citations the paper (or its author) already has, but disproportional to the quality of the paper when compared to other papers of similar quality. The equivalent of the Yule process that has possibly received the most attention within social sciences since the turn of the millennium, is that of *preferential attachment*, first defined in a widely discussed study in network analysis by Barabási and Albert (1999). They were the first to identify scale-free networks of links between websites in the then young World Wide Web, where a few sites were reported to have very high numbers of links to them, while the vast majority of sites only have very few. Their explanation of how such networks emerge – strikingly similar to Yule's explanation of speciation in biological genera – was that as new websites are made and the internet grows, these will tend to link (or preferentially attach) to existing websites that already have many links to them. It should be noted however that around 2010 there was a significant shift towards a more rigorous methodology that would be expected to accompany claims of power-law distributions in empirical data, leading to many previous claims being either weakened or fully rejected [Clauset et al. (2009); Stumpf and Porter (2012); Chapter 5].

A range of other more or less related power-law generating mechanisms and processes have been proposed (Mitzenmacher 2004; Newman 2005). Those that specifically relate to the magnitudes

of events distributed in time may be of particular use in archaeology, though they have seemingly received little attention so far. Among these are approaches focussed on phase transitions and critical phenomena, like *self-organised criticality* or SOC (Bak 2013; Bak, Tang, and Wiesenfeld 1987). This process models dynamical systems that continuously grow up until a certain critical point, at which it self-regulates downwards to stability through a collapse that is power-law distributed in size and frequency, before resuming the growth process again. The classic illustration of this model is that of a sand pile with a continuous addition of grains on top of it. Once the pile grows to its critical point where its slope becomes too steep, a sand avalanche is triggered and the pile stabilises again. The vast majorities of these avalanches are expected to be small in scale, but from time to time much larger avalanches appear. As the pile grows bigger, the critical point is also gradually raised, since a larger system can generally tolerate larger stresses before it needs to self-regulate. According to the theory behind such systems, the location of the critical point can be predicted with some accuracy, but the magnitude of the stabilising event that occurs when the system reaches this point, is impossible to predict beyond the power-law probability distribution it follows, i.e. that most events will be small but there is also a chance they will be several orders of magnitude bigger. The magnitude of the event is determined by the exact grain of sand in the pile that is the first to yield and thus triggering the chain reaction, but the scale of the potential consequences for the different grains varies greatly. The mechanism is said to be highly sensitive to initial conditions, or *chaotic*, similarly to how negligible initial differences in paper citations or website links over time can lead to disproportionally large differences, while it being virtually impossible to predict at the onset exactly *which* paper or website will come out on top Gleick (1987).

There are of course many ways in which the phenomena described here can be transposed to prehistoric social settings (see Section 4.3 below for examples of how this has been done). One important aspect of social processes that are known to generate power-law distributions, is that they often involve some sort of repeated competition, where actors who win once are more likely to win again later, thus cumulating their advantage. When the variable being modelled is house size, with the underlying assumption that this is a material reflection of wealth and/or power, and the distribution is recognised as a power law, the perhaps most straightforward interpretation is that of the emergence of a social hierarchy. The temporal process of this emergence could look

something like the following. Let there be an initial population of households living within a social system that for this purpose can be described as egalitarian (that is, ignoring factors like gender inequality, random differences between households, heterarchical differentiations between kin groups etc.). If, for any reason, a competitive dominance process is triggered, a household that initially holds a random slight advance over others may gain a dominant position over them, as clan leader or chief household. The leader or leading house may then gain certain privileges and duties related to tribute and redistribution of goods, thus strengthening the ties of dominance between them and the group of dominated households. There is however an upper limit to how many households that can be effectively dominated by a single household, so this relation would be unlikely to define the entire population of households. On the other hand, the same process could play out throughout the population, generating a set of leaders or leading houses each dominating a similar group of households – a situation that could be compared to the *house societies* described by Lévi-Strauss [-Lévi-Strauss (1982); Section 1.4]. If the social, and by extension political, competition continues, it will play out between these house or clan leaders, so that the one that by chance has an arbitrarily small advantage over the others may gain dominance over all of them and thus also their dominated households, and so on up to state societies (e.g. Earle 1997; Johnson and Earle 1987). As with the examples of the Yule process and preferential attachment discussed above, at any level of competition – e.g. between chiefs to become king – actors who find themselves further down the hierarchy, like clan leaders, have disproportionally lower chances of reaching that level, though in theory it is not impossible. Furthermore, the power-law relationship between leaders at different levels is recognised in that from any level to the next the number of leaders decreases exponentially while their dominance (the sum of households they each dominate) increases exponentially. The fact that most households remain at a bottom level with more random (*sensu* normal or log-normal) differences between them, does not change the fact that the leadership structure is hierarchical and power-law distributed – meaning that the entire distribution of dominance in the population has a power-law *tail*, as expected for most empirical settings.

There are some critical issues with such a model of the emergence of social hierarchy. Firstly, if house sizes are only *symbolic* materialisations of dominance – meaning that they do not in practice need to be large enough to actually fit all the people they dominate hierarchically – the distribution

of these house sizes would have a much shorter spread than what is expected in systems where there is a more concrete physical flow between elements, like oxygen in blood vessels. There is no natural proportion between decision-making power and square metres of floor space, and it seems likely that absolute house sizes in practice would be more constrained by building materials and techniques, which are also subject to change over time. Furthermore, several authors mention as a rule of thumb that a power-law tail should span at least two orders of magnitude in order to be accepted (Brown and Liebovitch 2010:53; e.g. Stumpf and Porter 2012:666). For a settlement where the bulk of houses are around 50 m^2 on average this would imply that the largest house (in a hierarchy of largest houses) would need to be in the order of 5.000 m^2 – comparable to the Pantheon in Rome or the Hagia Sophia in Istanbul – which, in addition to effectively excluding any prehistoric context, also seems as an unnecessary strict requirement for recognising hierarchy through house sizes. It is easily conceivable that smaller size differences could equally well be perceived as expressions of large social difference, e.g. if houses that are double the size of the average house (so 100 m^2) belong to clan leaders and double that again (200 m^2) to the chief. In that case, the size distribution would be more difficult to distinguish from a log-normal one, and if still interpreted as a power law, the scaling exponent (α) should be expected to take on a value considerably higher that what is usually expected for natural systems.

Another issue in need of empirical investigation, is how large a social hierarchy needs to be before it can be recognised as a power law, or over how many levels it needs to span. The rule of thumb of an interval of two orders of magnitude could be interpreted as two levels of scale (the base number of 10 is of course entirely arbitrary), which in social terms could translate to at least a complex chiefdom (i.e. a system of commons, intermediate chiefs and a chief of chiefs, Johnson and Earle 1987, #ADD PAGES). A system with only a single leader or leading house per settlement as described above, could potentially not be recognised as hierarchical within a distribution fitting framework, even though it may have been lived and felt very much as a hierarchy for the people involved. A large social hierarchy will furthermore often span all settlements across a geographical area forming a settlement hierarchy, in which case studying house-size distributions within single settlements may be misleading. The overall scale of the hierarchy – whether it concerns houses in a settlement or settlements in a regional polity – is best understood from the largest element, and if this is missing from the data for whatever reason, interpretations in

terms of social system will tend to underestimate the scale at hand. Luckily for archaeologists however, the largest settlements, as well as the largest houses, are usually the ones that are both best preserved and the easiest to discover, so that more often it is rather the lower end of the distribution that suffers from missing data, which has less importance to the interpretation of hierarchy (#REF?). It should also be kept in mind that in many – maybe most – hierarchical social contexts, the hierarchy may be highly organic and volatile, and there does not always need to be any discernible discrete levels. I would argue that even in social systems where discrete levels of hierarchy are clearly defined – like in the feudal system of medieval Europe – the hierarchy expressed through the material culture of the involved actors does not always need to tell the same story as the social hierarchy expressed through their titles of nobility.

It should also be noted that there are social situations that do not involve competition of power but that still can generate power-law size distributions. Most importantly is what can in a sense be seen as the opposite of the top-down hierarchy model described above, namely a bottom-up hierarchy, which structurally speaking can be very similar, or even exist as part of the former in a continuous dialectic tension (Furholt, Grier, et al. 2020). Anarchic or democratic societies can form complex (i.e. multi-scale) community structures based on clan or kinship structures (Hamilton et al. 2007; Haude and Wagner 2019:89–100), and if these structures are materialised in communal houses devoted to political assemblies or rituals, there is little reason to assume that the house-size distributions of such societies should be distinguishable from those of more top-down hierarchical contexts. An archetypical (pre-modern) example of such hierarchically structured democracies is the Iroquois League in the Great Lakes region during early European colonisation (Graeber and Wengrow 2021:481–92; e.g. Haude and Wagner 2019:127–31), but contemporary Western bureaucratic democracies are also highly hierarchical (De Landa 2006:67–91). In any such cases, distribution fitting can allow for identifying the hierarchical structure, but not the actual type of government and to which extent its authority is based on bottom-up or top-down legitimacy. This qualitative aspect must still be investigated through other strands of evidence, like find inventories and symbolic representation within the large buildings. Building size itself will play a different role in democratic systems compared to autocratic ones, though the outcome will often be similar. In the former, communal buildings are by their very nature meant to be accessible and used by large parts of the community, and will therefore often need

to be larger than common houses, while in the latter, leaders and high ranking people will be motivated to build houses for their private use that are larger than what they actually need, simply as a means to express their authority. In both cases the hierarchy is materialised in the house-size distribution, even though the absolute relationships between size and social importance may be culturally contingent. Exploring such specificities with hierarchical scaling in material culture has been one of the main motivations for this thesis.

In sum, the important characteristics of power-law generating systems, are that they are *complex*, meaning that they function or operate over several *scales* or orders of magnitude, which again implies that they exhibit *self-similarity* and can be described as *fractals* or *hierarchies*. Furthermore, they are rarely consciously planned, but rather tend to *emerge* spontaneously with growth, as a result of self-regulation and iterated responses to growth, like splitting or feedback that are proportional to size. Finally, even though such systems can be described as fully deterministic, they are also highly sensitive to initial conditions, i.e. they exhibit *chaos*, so that the magnitudes of single outcomes or trajectories can in practice be impossible to predict or forecast into the future, while the system as a whole can be well described, as well as explained backwards in time. Evidence of power-law distributed house sizes should be considered as a clear sign of a hierarchical social structure, though the exact nature of the structure – and specifically whether it is autocratic or democratic, or something in-between – needs to be argued from complementary evidence.

4.2.5 Some variants of power-law distributions

A note must be made regarding terminology on power laws and some closely related distributions. Contrarily to normal, log-normal and exponential distributions, power-law distributions are not defined in a uniform way across all the disciplines where they are applied. Despite all claims of ubiquity, they remain special cases in many contexts, are not systematically taught in basic introductions to mathematics or statistics, and their broadened understanding seems to have suffered from long-standing discipline-specific traditions of defining them. Archaeologists who borrow theory and method from different fields therefore run the risk of talking past each other and not seeing the bigger picture of common phenomena described in different ways. Two alter-

native and more or less parallel ways of describing power-law distributions run under the names *Pareto* and *Zipf distributions*, after the first researchers who became known to define them under their specific parametrisations.

The so-called Pareto distributions form a group of distribution types with varying numbers of parameters, and are most often associated with applications in economics and finance (Pareto was an economist and his major works were on wealth and income distributions, see Chapter 20 in Johnson 1994; Pareto 1896). Comparing with the power-law definitions given above (Equations (4.7) and (4.8)), it is important to note that Pareto distributions are defined as survival functions, i.e. cCDFs (complementary cumulative distribution functions) or the proportion of the distribution that is *higher* than a sample X. The scaling exponent for a Pareto distribution (Johnson 1994 confusingly note this as a, p. 573, but here I prefer to use β for clarity) thus has the value of the power-law exponent $\alpha - 1$, or in negative values $-\beta = -\alpha + 1$, meaning that Pareto plots have a less steep slope than their power-law counterparts. Accompanying their use in applied economics is the appellation of the "80-20 rule", a rule of thumb stating e.g. that in a company 80% of sales will typically go to 20% of clients as a result of the very skewed distribution. This specific rule corresponds to a power-law distribution where $\alpha \approx 2.1$ or $\beta \approx 1.1$ (Newman 2005:334).

Zipf law distributions, often referred to as "rank-frequency" or "rank-size" plots or rules, are most associated with linguistics, due to Zipf's work showing that word counts (as well as city sizes) tend to follow power laws (Arshad, Hu, and Ashraf 2018; Zipf 1949). Only some minor plotting conventions differ Zipf plots from Pareto plots, but again this influences systematically the value of the calculated scaling exponent. While Zipf distributions are also cumulative, they are not normalised to 1, meaning that absolute size ranks are plotted instead of probabilities. Furthermore, these are plotted on the x axis rather than the y axis, with the implication that the distribution is *discrete* rather than continuous. The y axis (the actual variable being measured) can in principle be discrete or continuous, but the most commonly cited cases are discrete, i.e. count data, like the number of inhabitants in a city or the number of times (frequency) a word appears in a text. The scaling exponent of a Zipf law (or a Zipf exponent, denoted q e.g. in Arshad et al. 2018:78) can thus be expressed as

$$q = 1/\beta = 1/(\alpha - 1). \tag{4.9}$$

The fact that there does not seem to be any consensus on which character to use for the different parameters is an additional difficulty to handle when comparing studies that apply these different approaches. In his extensive review of these three separate research traditions, Mark Newman (a physicist; the power law is the most common variant within the natural sciences) admits that this inconsistent nomenclature "causes much confusion in the literature" (2005:327), which I will claim is a polite understatement. For clarity, throughout this thesis I only refer to the scaling parameter of the power law as defined in Equation (4.7), which I denote α, no matter the graphical representation of the distribution (most of the plots here are cCDF plots, and could thus be qualified as Pareto plots). The value of α is furthermore only calculated directly on the data using the maximum likelihood method described in Chapter 5, and is thus not dependent on the type of plot chosen for graphical illustration.

As already mentioned, regular power laws are asymptotic so that as x approaches infinity, y goes arbitrarily close to 0 without ever reaching it. This in itself acts as a limit to its usefulness for modelling most real-world phenomena which are of finite size, even though the power law may be a good approximation over some range of the system. In other words, in a power-law model of house sizes, no matter the value of α, the probability of a house measuring a trillion square metres is certainly negligible, but still tecnhically above 0. Several solutions have been proposed to this problem, allowing the model to end somewhere at $x < \infty$. Among the main candidates are the *power law with exponential cutoff*, the *stretched exponential* and the *parabolic fractal* distributions. The first one is – as the name indicates – a power law with an additional exponential factor, which turns the function increasingly (exponentially) downwards as x gets higher (see Clauset et al. 2009:664 for precise definition). The second is a type of exponential distribution, similar to the one described above (Equation (4.5)), but where the x exponent is raised to a power constant $0 < c < 1$, giving a survival function $P(x) = e^{-(x/x_{min})^c}$, resulting in a tail which is heavier than that of a regular exponential distribution but thinner than that of a regular power law (Clauset et al. 2009:664; Laherrère and Sornette 1998). This survival function or cCDF of a stretched expoential is furthermore equivalent to that of the *Weibull distribution*,

a well studied model type that is widely implemented in statistical software, making it easier to include in analyses compared to the more piecewise power law with exponential cutoff (comp. Johnson 1994:629 ff.).

Finally, the so-called parabolic fractal distribution has been proposed as a quadratic function model on log-transformed values of size to rank (Zipf plot), to account for cases where the distribution follows a parabola of constant curvature on a log-log plot (Laherrère 2000). Jean Laherrère – working for a large oil company modelling the size distribution of oil reserves – noted how studies that claim to show power-law distributions in empirical data often tend to explain away the curvature of the distribution in log-log plots, e.g. by focussing on shorter ranges of the tail, or by constructing composite distribution models. While the log-normal model forms a parabola on log-log plots of its PDF (Figure 4.2), none of the above discussed models have this characteristic for their survival functions. A quadratic function is a simple polynomial including a squared term of the variable, of the type $f(x) = ax^2 + bx + c$. The use of the term *fractal* here relates to the shape of the distribution being evaluated with log-log transformations, so that the function is interpreted as a power law with an additional squared factor, a and b both representing scaling exponents. With quadratic functions, the coefficient to the squared term (here a) determines the curvature of the model, while c has the role of y-intercept as with linear models. The model can furthermore be interpreted as a power law where the scaling exponent α is continuously increased with higher x, in other words where α corresponds to the derivative of the parabola at any given x, forming a continuous spectrum characteristic of multifractals (see e.g. Harte 2001). Though this model seems to have much potential both for modelling and explaining self-affine phenomena within finite systems, it is relatively recent and not easily applicable without specialist knowledge in statistics and programming, and is not further included in this thesis. Future studies would be warranted, e.g. applying the analyses proposed in Laherrère (2000) and Laherrère and Sornette (1998) for modelling sizes and frequencies of undiscovered and/or lost archaeological sites within given regions.

4.3 Fitting heavy-tailed distributions in archaeology

The distribution fitting approach adopted in this thesis is inspired by a relatively small number of previous studies in archaeology. Though the relationship between households and house sizes had already been thoroughly studied more generally by archaeologists and anthropologists since at least the 1970s [see Section 4.1], the earliest quantitative study of prehistoric house-size distributions accompanied with theoretical arguments for interpreting power laws as signatures of multi-level social complexity, was perhaps a study by Herbert Maschner and Alexander Bentley published 20 years ago (Maschner and Bentley 2003). They presented a well-argued case for hierarchical scaling between households in a study area on the Alaska Peninsula, apparently discernible in several periods of the region's prehistory, and explained this as the result of an elite emerging from various (though unidentified) competitive socio-economic practices. The study could today be criticised for methodological shortcomings – the authors relied on least squares fitting on log-transformed binned data, they did not systematically propose more than one model for the data, nor propose any quantitative way of selecting the best fits, and their data sets seem to have been severely time-averaged, rendering any claims of hierarchical scaling between households potentially meaningless. All of these issues have been thoroughly addressed in subsequent studies (see Chapter 5). The general analytical procedure and rationale however, remains highly innovative and would merit far more attention than it has received. In my view, this study represented the first turn beyond the study of size *averages* towards a theoretically informed study of size *distributions* in household archaeology (#Check this sentence again after writing the top of the chapter).

This approach was further explored by Brown et al. (2012) and expanded upon by Strawinska-Zanko et al. (2018), who identified a shift from less heavy-tailed distributions (exponential) to power-law distributions (Pareto) of house sizes in the Maya region approximately coinciding with the pre-Classic/Classic transition, which is traditionally considered the onset of state-level organisation. Following political scientist Manus Midlarsky (1999), Strawinska-Zanko et al. (2018) argued that this transition to a power-law distribution of wealth (proxied through house sizes) could be explained as resulting from increasing competition for agricultural land ownership following population growth. Furthermore, they identified a trend towards more pronounced inequality,

both through lower α values and higher Gini indices, until the end of the Classic period, with an abrupt shift to more equal distribution in the post-Classic (higher α and lower Gini). Despite the inclusion of only four settlements in this case study, it remains highly interesting since the overall economic and demographic development of the Classic Maya is very well documented from a range of other approaches and intensive research. The study also contributed with detailed discussion of methodological issues, comparing the performance of different procedures. In a study by Crabtree et al. (2017) a similar analysis was performed on data from the Mesa Verde region in the American South-West over the 7th to 13th centuries CE. This cultural context is also very extensively documented, and holds the additional advantage of a fine-grained temporal sequence supported by dendrochronology, allowing for detailed analyses of near-coeval features. With the same motivation of using distribution models as proxies for underlying generative mechanisms, the authors systematically compared best fit log-normal and power-law models of settlement sizes as well as the sizes of *kivas* – a special category of communal ritual structures – across the study area. Though the results were not entirely unanimous, both indicators pointed towards a settlement hierarchy in the Pueblo II phase from ca. 1030-1140, centred around Chaco Canyon receiving tribute from surrounding areas, which again is coherent with the current understanding of the period based on other strands of evidence. Furthermore, they implemented preferential attachment mechanisms into an agent-based model of the regional socio-demographic development accordingly, largely reproducing the observed temporal patterns. This also seems to have been the first archaeological study consistently performing distribution fitting by maximum likelihood estimation rather than least squares, and comparing the fit of different models quantitatively. For this they used the same R package *poweRlaw* that is used here in the following chapters (Gillespie 2015).

The analysis of the size distribution of special communal structures across a region, as was done by Crabtree et al. (2017), is in some sense a bridge between analyses of house-size and settlement-size distributions. #GROVE2011 from here.

- Zipf law in archaeology, then A law (distribution) is not a law (of nature), see Grove (2011) for review of the long-lasting confusion in archaeology (e.g. Hodder (1979)), also "rank-size rule"

- Zipf law and Settlement Scaling theory, Bettencourt (2021), Gomez-Lievano, Youn, and Bettencourt (2012), Lobo et al. (2020), Duffy (2015) Connection with Central Place Theory, e.g. Müller-Scheeßel (2007), Chen (2011). Why I'm not doing settlement scaling in this study.

- SOC in archaeology, see (Bentley and Maschner 2001; Diachenko and Sobkowiak-Tabaka 2022; Zhukov, Kanishchev, and Lyamin 2016).

- Not fitting distributions in archaeology, just assuming they are heavy-tailed, or avoiding the question: ex. Brink (2013) (could include lots more!)

END chapter

Chapter 5

Methods: Distribution fitting

5.1 Modelling heavy-tailed distributions

```
## -- Attaching core tidyverse packages ------ tidyverse 2.0.0 --
## v dplyr     1.1.1      v readr     2.1.4
## v forcats   1.0.0      v stringr   1.5.0
## v ggplot2   3.4.2      v tibble    3.2.1
## v lubridate 1.9.2      v tidyr     1.3.0
## v purrr     1.0.1
## -- Conflicts ---------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflict
```

In this chapter I go into some more detail around the methods used in Chapter 6, and the reasoning underlying my choices of methods. As mentioned earlier (Sections 1.2 and 4.2.4), I consider power-law distributions as a statistical signature of hierarchical structures, and wish to test whether such structures may be reasonably shown to exist among houses in European Neolithic villages, or if house sizes in these contexts are better explained by other non-hierarchical models. Results from these analyses add to current debates surrounding the development of social

and political organisation in the Neolithic, and to the question of the emergence of stratified societies more broadly. Furthermore, and as also mentioned in Section 4.2.4, the methodological procedure leading to claims of power-law distributed data is not entirely straightforward, and is an issue that has undergone important developments in recent years, often leading to refutations of earlier claims. It is therefore critical to be explicit as to the methods being applied in studies like this one, and not simply report results obtained in some unspecified way through obscurely documented software. In the following I will present #FILL IN AFTERWARDS#. Due to limited space and for simplicity, I will concentrate on the choices of methods and procedure, and not on the under-the-hood functioning of different statistical tools like maximum likelihood estimation and calculation of the Akaike information criterion. For more details on these there is a number of good introductory volumes, some of which – like Shennan (2008) and Baxter (2003) – are also specifically aimed at archaeologists.

The standard method for fitting power-law models to empirical data throughout the 20th century was through least squares linear regression on log-transformed x and y values (e.g. Harrison 1981; Mitzenmacher 2004), the same way exponential and log-normal models could be fit more easily to data by log-transforming x values. Conscious about the still frequent lack of statistical training among archaeologists, Brown et al. (2005) and Brown and Liebovitch (2010) presented the log-linear regression method as sufficient because of its simplicity of application compared to more sophisticated methods. Brown and Liebovitch (2010) furthermore provided a detailed discussion around how to plot the data in order to obtain the most accurate parameter estimates. The central problem with fitting power-law models to data, is that power-law distributions characteristically have an overwhelmingly large proportion of the data at lower values, while the scarce high values are typically several orders of magnitude higher. Density plots of empirical data (with the PDF on the y axis) require binning, so that the plotted data points in reality correspond to bar heights in a histogram. The applied bin width will furthermore have a heavy influence on the appearance of the plot, where small bin widths are best to represent the many low values and large bin widths are best for the few high values. Brown and Liebovitch proposed multi-scale PDFs, combining histograms of different bin widths before performing regression, which they showed give better results on synthetic data (2010:2). Logarithmic binning – increasing bin width exponentially so that points appear to be spaced as constant increments on a log-transformed x axis is also a

possibility that has been proposed (Newman 2005:325–26). Plotting the cCDF instead of the PDF has the advantage of avoiding the bin-width issue altogether, since y values then are a function of the rank of each data point. This also allows for using all the data and not reducing it into bins, and is shown on synthetic data to give more accurate α estimates (i.e. absolute slope of cCDF +1, see Eq. (4.9)). However, the inconvenience with fitting the power-law model to the cCDF, as pointed out by Brown and Liebovitch (2010), is that it does not necessarily form a straight line, but will in particular be curved when $\alpha \leq 1$, making it more difficult to distinguish visually from other heavy-tailed distributions.

From the early 2000s, physicists and mathematicians started to criticise the frequent use of log-linear regression methods for modelling power laws, since they were shown to introduce systematic biases to the parameter estimates no matter the adopted plotting method, and calls for the use of more robust methods like maximum likelihood estimation (MLE) were put forth (e.g. Newman 2005:325–27; Stumpf and Porter 2012). A new methodological tool kit was proposed by Clauset et al. (2009), which has since seemingly become the new gold standard for fitting heavy-tailed distributions. One of their main critiques of earlier practices, came from the recognition that in nearly all real-world contexts where power laws are claimed to exist, this behaviour only kicks in from some lower threshold or x_{min} in the terminology of Clauset et al. (2009). In earlier studies the value of this threshold was simply set by guessing from the looks of the plot and trying to fit a line covering as much as possible of the data. More formally, this also impeded proper normalising of the distribution. The method proposed by Clauset et al. (2009) consisted in testing a range of different x_{min} values and picking the one that gave the best MLE fit to the data by minimising the KS or Kolmogorov-Smirnov statistic (the largest observed distance between the model and the data), and was reported to perform very well on synthetic data. Next they tested the plausibility of the power-law model through bootstrapping, i.e. generating a large number of synthetic random data sets with the same estimated parameter values, each time measuring the KS statistic compared to the ideal model. The fraction of runs giving a KS statistic higher than that of the empirical data gives the *p*-value, which they argued should lead to a rejection of the power-law hypothesis when $p < 0.1$. They estimated the number of bootstrapping runs necessary for robust p-values being between 1.000 and 10.000, which for a few data series is not dramatic, but for larger numbers of data series quickly become computationally intensive. But

most importantly, they argued that the bulk of previous studies claiming to find power laws in empirical data never actually tested and compared their model with alternative models, which they argued should be done even in convincing cases where a power-law model could not be excluded as a good fit through bootstrapping. The method they proposed for comparison and selection between competing models, was Vuong's log-likelihood pairwise comparison test, though pointing out that any good statistical model selection method could serve this purpose (see Clauset et al. 2009:663 for an overview of their "recipe for analysing power-law distributed data"). These methods were later implemented with functions and documentation in the R package *poweRlaw* (Gillespie 2015), which has been used and cited in at least some archaeological studies since (Crabtree et al. 2017; Haas et al. 2015).

For the present study I have largely chosen to follow the instructions advocated by Clauset et al. (2009), but with a few modifications, for reasons that are discussed in more detail below. Early experiments with the *poweRlaw* package indicated that the proposed bootstrapping procedure possibly represented a slight overkill, requiring much computing time with relatively limited gains. Furthermore, the pairwise model comparison using the Vuong's log-likelihood test appeared good but somewhat tedious, requiring a nested algorithm eliminating competing models one by one. Therefore I decided instead to opt for testing all candidate model simultaneously using the Akaike Information Criterion (AIC), or rather the version of it designed to correct for small sample sizes – the AICc. The AIC score is also calculated from the log-likelihood of each model, and indicates which of the candidates accounts for the given data with the most weight. While not implemented in the *poweRlaw* package, it is a frequently used statistical tool implemented in a number of accessible R packages. Here I used the *AICcmodavg* package (Mazerolle 2023) because of its useful functions for manually tackling the differences in how base R and *poweRlaw* models are coded. I also tested using the Bayesian Information Criterion (BIC) on the data sets, but quickly came to the conclusion that the AICc was sufficient for the types of models being used here, with maximum two parameters.

5.2 Testing for false positive power-law tails

The difficulty of comparing power-law models with other common candidate models (like log-normal or exponential), is that they, unlike the others, by definition need a specified lower bound above 0, denoted x_{min}. Comparison of multiple models with AIC scores is only meaningful when done over the same range of data (this also applies to the Vuong's log-likelihood test for pairs of models proposed by Clauset et al. 2009). However, comparing multiple models over the range in a data set which has already been recognised as providing the best possible fit for a power-law model, gives this latter model a potential advantage over the other ones. Log-normal models, for instance, can explain the entire range of a data distribution, where a power law can in most cases only explain the highest values in the distribution tail. The fact that these two model types have frequently and for a long time represented competing explanations for the same empirical data sets, may reflect this apparent incomparability between them (e.g. Bee, Riccaboni, and Schiavo 2011; Gibrat 1930; Harrison 1981; Mitzenmacher 2004; Sheridan and Onodera 2018). One can suspect then that this procedure of distribution fitting and model selection would favour power-law models unreasonably. At the same time, one of the main findings of the Clauset et al. (2009) study, was that power-law behaviour was only confirmed beyond reasonable doubt in one out of 24 empirical data sets which had been reported as power-law distributed in earlier studies, leaving the impression that the methodology would be conservative rather than lenient. In many cases however, the study remained inconclusive, especially regarding comparisons between power-law and log-normal models to empirical data sets (comparisons between power-law and other models were generally more conclusive). The authors admitted "*In general, we find that it is extremely difficult to tell the difference between log-normal and power-law behaviour. Indeed, over realistic ranges of* x *the two distributions are very close, so it appears unlikely that any test would be able to tell them apart unless we had an extremely large data set*" (Clauset et al. 2009:689). Extremely large data sets are of course a luxury that is rarely afforded in archaeology, and if these two models are that close in many situations, one can ask whether picking one over the other really matters in the end. This question is further developed in Chapter 11.

The reliability of this methodology can to some extent be assessed using synthetic data. A first question to address is whether sample size affects the selected distribution model for the tail,

and if so what size should be considered a minimum for the results to be reliable. In Figure 5.1, using random number generator functions in base R (R Core Team 2023) and with the *poweRlaw* package (Gillespie 2015), I reproduced Fig. 5a in Clauset et al. (2009), namely examples of a power-law, a log-normal and an exponential distribution, with the addition here of a stretched exponential, illustrating how they all can look roughly linear on log-log plots with their survival functions/cCDFs. Using the same parameter values, but with four different sample sizes (10, 100, 1.000 and 10.000) on each model, these test distributions were run through the distribution fitting algorithm described above (Figs. 5.2 and 5.3). The power law was correctly identified no matter the sample size, but for the smallest sample size ($n = 10$) all other distribution types also gave power-law tails. For $n \geq 1000$ all distribution types were correctly identified also in their tails, while for $n = 100$ this was only the case for the stretched exponential and the power law. These results are in agreement with the analysis based on p-values obtained from bootstrapping presented by Clauset et al. (2009:676 ff.), but the method opted for here is far less computationally intensive. Selecting the best model alternative directly based on AICc is also a less complex operation compared to the sequence of first bootstrapping and then performing pairwise comparisons of log-likelihood as proposed by Clauset et al. (2009). The inconvenience with the method proposed here, is of course that there is no guarantee that any of the models tested for are in reality appropriate – we only find out *which* one of them is the *most* appropriate. The p-value approach in (Clauset et al. 2009) does allow for positively rejecting hypotheses that clearly do not fit the data. However, early experiences (not presented in further detail here) gave the impression that this made little practical difference, at least on the data sets analysed in this thesis. Most sample sizes are in the order of 100 or lower, in which cases bootstrapping remained inconclusive, while it was still interesting to have an indication of which model that gave the best fit. The results shown in Figure 5.2 indicate that for sample sizes below ca. 100 power-law interpretations should be treated with care, and should not be trusted as n approaches ca. 10.

However, it must be noted that in order to generate the non-power-law distributions with a defined lower threshold as done here (and in Clauset et al. 2009), a much larger number of data points is in reality needed if we also consider those falling below that same threshold. For example, to get 1000 data points with values above $x = 15$ following the log-normal distribution shown in Figure 5.1, they need to be filtered out from a total distribution almost ten times larger.

When considering entire data distributions without lower bounds, as is usually the case e.g. when analysing archaeological house-size distributions, the sample size will potentially also need to be much larger for correct model selection, although exactly *how* much larger should depend on the type of distribution and parameter values of the data. Similarly, when Clauset et al. argued that with the MLE method for estimating α in a power-law model, sample sizes around $n \geq 50$ would usually be enough for the estimates to be within 1% accurate (2009:669), sample size is here referring to the number of data points actually being considered when fitting, which is $n > x_{min}$ only. For this number to be 50 or higher, the total size of the distribution could often need to be 500 or higher, which is far more than most of the house counts per village in this study. This matter of sample size is a question that is perhaps less relevant to physicists and mathematicians, but that may be of crucial importance to archaeologists who regularly suffer from limited amounts of data.
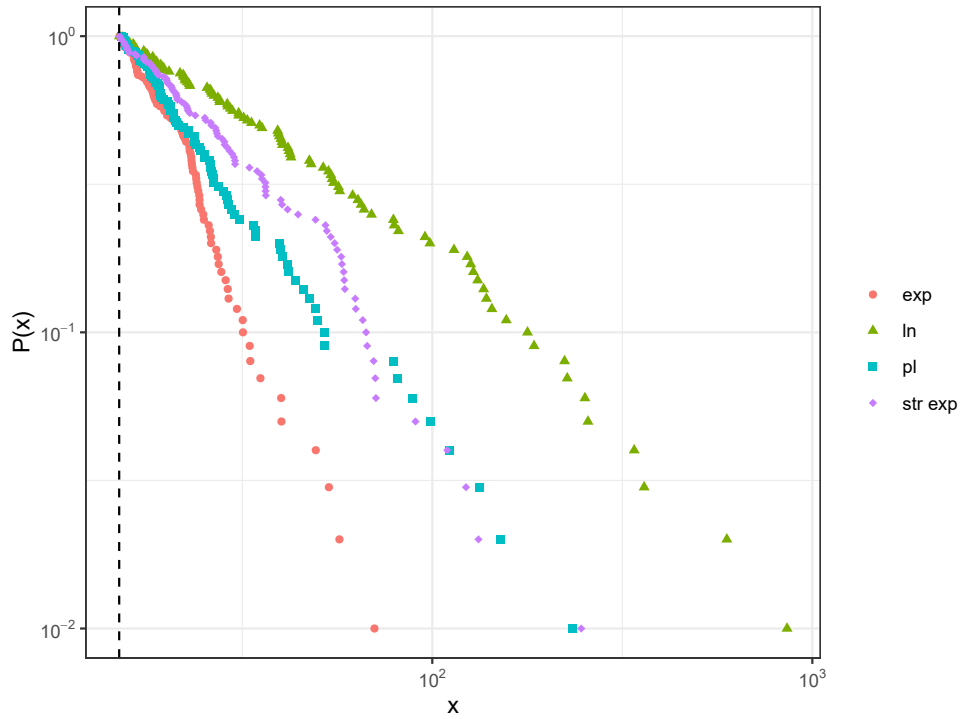

Figure 5.1: *Synthetic data series drawn from four different distribution types: exponential ($\lambda = 0.125$), log-normal ($\mu = 0.3$, $\sigma = 2$), power-law ($\alpha = 2.5$) and stretched exponential/Weibull (shape $= 0.5$ and scale $= 3$), all with $n = 100$ data points and $x_{min} = 15$. Plot equivalent to Fig.5a in Clauset et al. (2009), with deviations due to random fluctuations only. Scales are logarithmic, and all four series appear as roughly straight lines, though only one is a true power law*
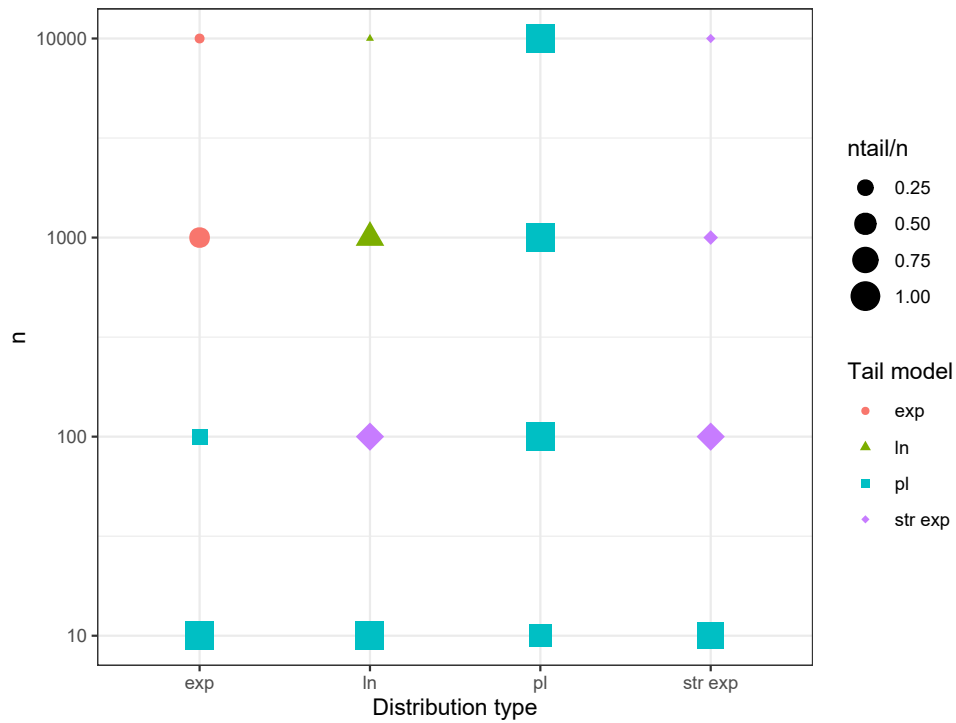
Figure 5.2: *Selected tail models for the same synthetic data sets, each with four sample sizes ($n = 10^1, 10^2, 10^3, 10^4$). For each tail model, x_{min} is set at the value which gives the best power-law fit. Point size indicates fraction of data points thus included in the tail model. For power-law distributions, all samples are correctly identified, while this is the case only for large samples ($n > 10^2$) of log-normal and exponential samples, smaller samples being interpreted as having power-law or stretched exponential tails*
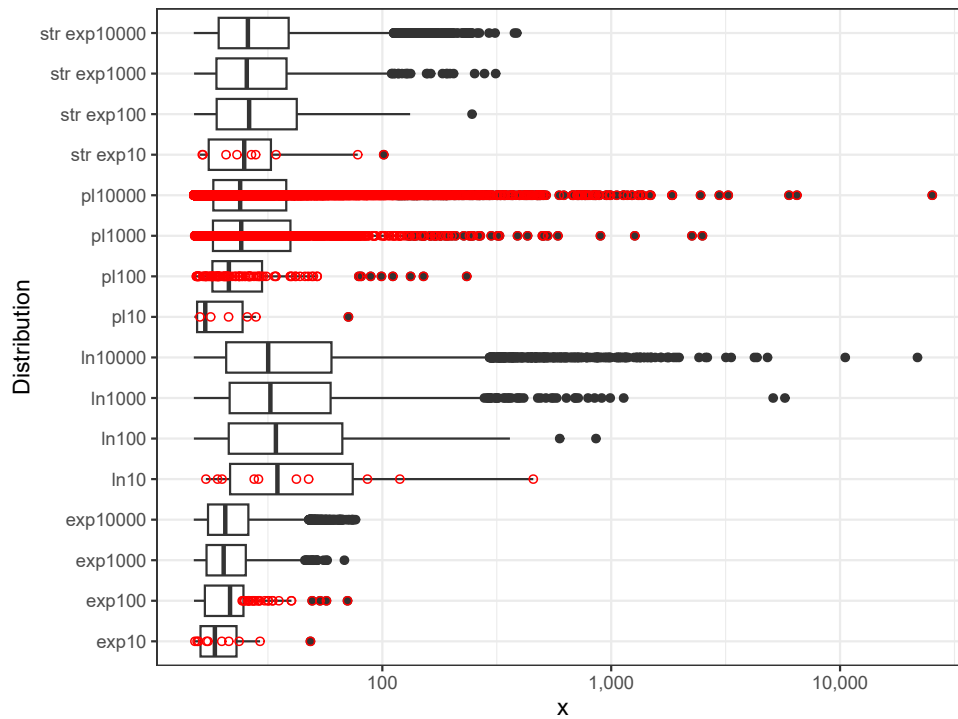
Figure 5.3: *Boxplot of all the synthetic data sets, overlaid (in red) with the data points interpreted as power-law tails. X axis is logarithmic – however the log-normal distributions do not appear symmetric since they are truncated with a lower threshold. Note that especially for log-normals and power laws, larger samples give longer tails. If the model predicts a probability of having a value of 10.000 or more as only 1 in 10.000 or 0.01%, a sample size of 10.000 will probably allow for one such value*

A second question more specifically related to the selection between log-normal and power-law models, is whether certain parameter combinations increase the likelihood of log-normal distributions being incorrectly interpreted as power laws. In his extensive review of power-law generating mechanisms, Newman (2005:347–48) showed algebraically how log-normal distributions can be mistaken for power laws especially when the range of the data that is being analysed is short, and when the value of σ is high. More specifically, since the PDF of a log-normal on log scales is a quadratic function – i.e. a parabola – sufficiently smaller sections of this will be nearly indistinguishable from straight lines, and can thus be well modelled as a power law. The curvature of the function is characterised by its quadratic term, which is a fraction with x in the numerator and σ in the denominator, written $-\frac{(\ln x)^2}{2\sigma^2}$, essentially causing a flatter curvature with higher values of σ since this term then will vary more slowly with x (see Eq. 84 in Newman 2005 for more details). The difficulty of distinguishing the two model types empirically undoubtedly lies in the close relationship between them, both being defined as some enhanced exponential distribution (Mitzenmacher 2004). Figure 5.4 shows the cCDF plot of 36 synthetically constructed log-normal distributions, with μ values ranging from 1 to 6 in integer increments, and σ values from 0.5 to 3 in increments of 0.5, each with sample size $n = 1000$ and no truncation (i.e. $x > 0$). When plotted this way (with the survival function of the variable), low σ values generate angular curves, while high σ values generate more parabolic curves.

Running these distributions through the distribution-fitting and model-selecting algorithm, more than half of them are interpreted with a power-law tail (19 of 36, Figures 5.5 and (**ref?**)(fig:05-ln-pl). These are seemingly spread throughout the parameter space, with the only clear pattern being that for the highest σ values, the power-law tails cover only smaller fractions of the data.

What are we then to conclude from these preliminary tests? Clearly, it is a challenge to confidently distinguish between log-normal and power-law distributed data in the high ends of distributions, as noted in the beginning of this section. If the proposed methodological procedure seemingly serves well to identify power-law distributions when that is what they really are, it also seemingly identify these erroneously in the tails of log-normal distributions half of the time, irrespectively of the log-normal parameter values, and with (for archaeologists) optimistic sample sizes. A more positive way to look at it, is to acknowledge that in truly log-normally distributed data,
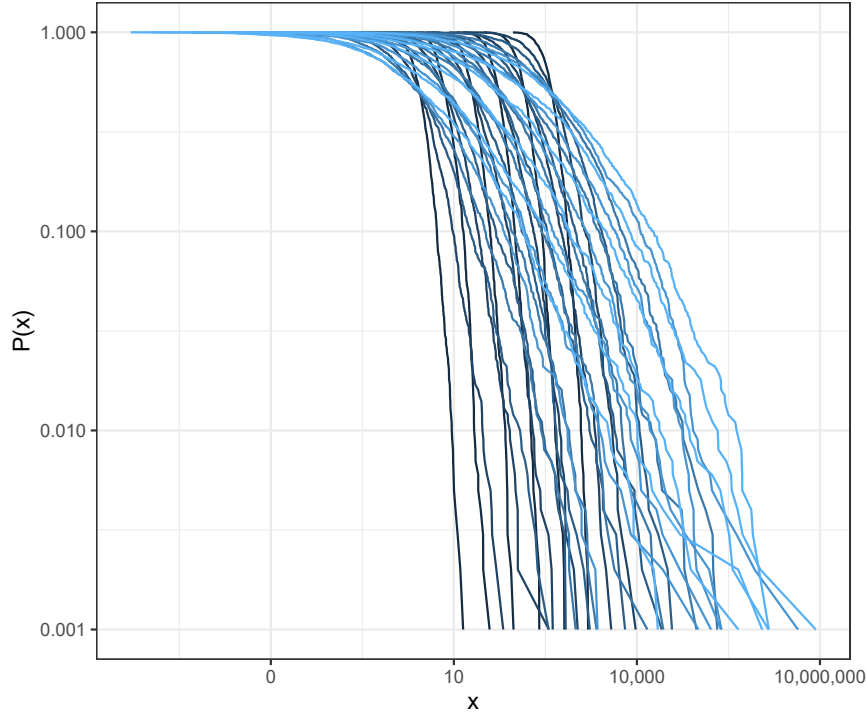
Figure 5.4: *cCDF plot of 36 synthetic log-normal distributions with parameter values $1 \leq \mu \leq 6$ and $0.5 \leq \sigma \leq 3$. Each distribution is generated with $n = 1000$ data points, but rendered here as lines for clarity. Scales are logarithmic*

some definable portion of the upper tail is in many cases indistinguishable from a power law, and actually best modelled as such. There may well be precise mathematical reasons behind this, but further insight to whether such power-law tails are confidently indicative of social hierarchies when observed on material culture proxies such as house sizes, would perhaps require large-scale systematic testing on ethnographically documented cases, which would clearly go beyond the scope of this thesis.
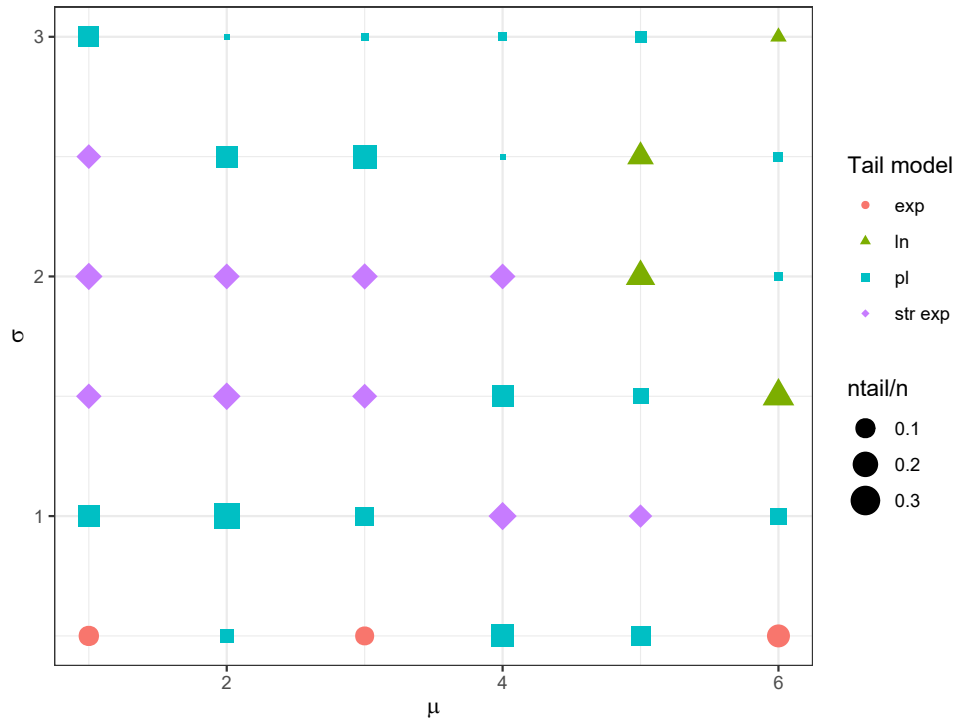
Figure 5.5: *Interpreted tail models of the same log-normal distributions. 19 of 36 distributions have tails that are best modelled as power laws. Symbol size indicates fraction of the data included in the tail, with x_{min} parameter set for best possible power law fit. See text for details*
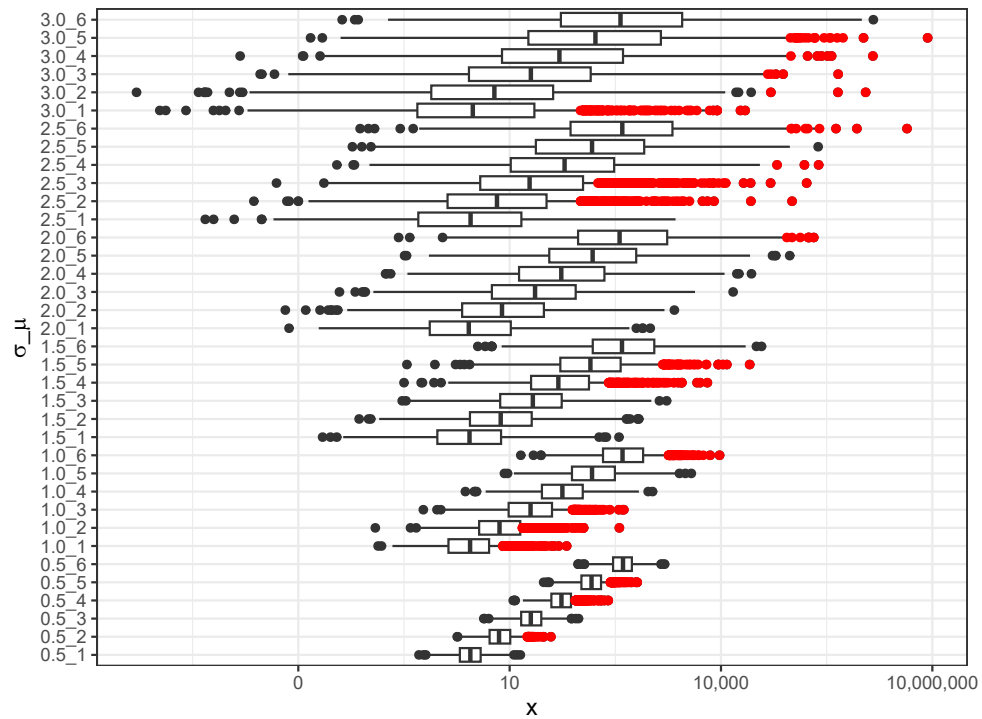
Figure 5.6: *Boxplot of the same 36 synthetic log-normal distributions, overlaid (in red) with data points included in tails interpreted as power laws. The power-law tails stretch across the log-normal data in a range from 0.3% (3 data points out of 1000) to 23.5%*

5.3   Methodological procedure

- Reminder of main goal for this part of the study: identify power-law structures in the house-size distributions of the Linear Pottery and Trypillia samples.

- Synthetic data genaration:

  - Why?

    1. Process: how long/much does it take for a normal distribution to become power-law? And reverse sense? (multiplicative process)

    2. Temporal resolution issue: does the temporal palimpsest of several phases with e.g. log-normal distributions produce false power-law signals? (additive process)

  - How?

    * Random number generation and iterated multiplicative (1.) or additive (2.) sequences, with K-S testing (Gillespie 2015) at each stage. Report when the distributions become power laws.

- Present data set with categories (settlements, quarters/neighbourhoods, time samples for Vráble)

- Parameter settings: dist. types, xmin, testing the pl hypothesis etc. Minimal house-count cutoff (min. sample size). Isolating top house.

END chapter

Chapter 6

Results: Distribution fitting

6.1 Synthetic distributions

6.1.1 Multiplicative process

A multiplicative process is one in which a term is multiplied with itself a number of times, and can be mathematically described through an exponential function. This is commonly denoted

$$f(x) = ab^x$$

read as "the function of x equals a times b to the power of x", where b is the main term which is multiplied with itself x number of times, and a being a constant of proportionality. In such functions, b determines the *shape* of the function – if between 0 and 1, the function models exponential decrease or decay as x grows larger, and if $b > 1$ the function models exponential growth. This term corresponds to the *rate* parameter for exponential distributions detailed in Section 4.2. The a term determines the *scale* of the function, similarly to the y-intercept of a linear function of the type $f(x) = ax + b$ (in this case b), and though in an exponential function this parameter is multiplied rather than added to the variable term (b^x), its value gives the y-intercept since $ab^x = a$ when $x = 0$. Furthermore, the link between exponential and linear functions is illustrated more clearly when we log-transform the former, that is: $f(x) = log(b)x + a$, which

is essentially a linear function where *a* is the y-intercept and *b* is the slope (# Show with fig! and go on).

#Insert figure, ill. exponential functions.

6.1.2   Additive process

fdsa

Lumping of samples may be done for two main reasons (in archaeology): 1 - the sample size is too small, and lumping together several contemporary samples can raise the sample size to an acceptable level (spatial lumping), or 2 - it is impossible to temporally disentangle elements within a settlement, i.e. reduce temporal resolution to coeval elements, in which case one has no option but temporal lumping (accepting a low temporal resolution). The problem can be further broken down to two case types: a) all lumped samples are really drawn from the same distribution, and b) they are not similarly distributed (**illustrate with table**). For spatial lumping this degree of similarity can be assessed, but not necessarily for temporal lumping. But even when it can be assessed, the lumping needs to be justified in social terms. As an example, several spatial samples (e.g. villages in a region) can have house-size distributions with no significant differences between them (ANOVA), but at the same time be functionally entirely independent, in which case it is more logical to augment the sample size of a single one through simulation and evaluate plausibility through bootstrapping, rather than by lumping of all samples, **rewrite this more clearly**. On the other hand, house sizes can be significantly different between quarters or suburbs within a city or metropolitan area, or even between cities in a region or country, but if they all function together in a coherent system, they may reflect a spatial segregation between different strata in the society, and it makes much sense to lump and analyse them together.

When it comes to temporal lumping, given that a settlement does not undergo substantial cultural changes during its timespan, a workaround to evaluate whether the house-size distribution evolves significantly over time may be to target a number of size categories for ^{14}C dating, and to check that they all stretch over the entire range of the settlement's duration, and if yes, accept to analyse the whole distribution as one. **recheck this after I have my results, maybe move it all to**

summary. □

6.2   Settlements

House-size distributions of whole settlements (both Linear Pottery and Trypillia)

6.3   Quarters/neighbourhoods

House-size distributions of separate quarters (Nebelivka) and neighbourhoods (Vráble). Can I include separate Nebelivka neighbourhoods as well? (prob. too small, but check)

6.4   Temporal samples (Vráble)

If time, I can add more Žitava sites here, but not necessary. I can also analyse temporal samples of Vráble neighbourhoods separately.

6.5   Summary of findings

END Chapter

Part III

Settlement Plans

Chapter 7

Village planning in Prehistory

7.1 Settlement layout and social structure

Or the social organisation of village layout. Research background:

Lit. use Furholt (2016), Fraser (1968), Ensor (2017), Ensor (2013), also check Souvatzi (2017).

For cities, see Kostof (1991)

Artursson et al. (2010) (Bronze Age, descriptive/interpretive approach). Cleuziou, Braemer, and Coudart (1999) (book in MAE)

Use the Trypillia volumes. Also Müller-Scheeßel (2019), Trebsche, Müller-Scheeßel, and Reinhold (2010)

Transition from village to urban (again): Birch (2014).

Factors affecting village layout:

- Political structure (but, as with hierarchy, an organised layout does not necessarily equate top-down despotic decision making).

- Kinship, matrimonial and locality structures

- Cosmology (e.g. Linear Pottery house orientations)

- Economic and ritual functions of village elements (constructed and non-constructed)

- Local landscape setting (to be factored out)

7.2   The geometries of conscious planning vs. emergent behaviour

- I need to find some references here! Eglash (1999)

- Euclid: grids, lines, circles – how humans think in shapes. Social settings: architect/planner, strong common institutions/ideals (examples?)

- Mandelbrot: irregular, self-similar, scale independent (i.e. fractal) shapes – emergent, not consciously preconceived. Self-organisation. Does the "no pattern" case exist? Emergence from repetitive sequences of simple choices/mechanisms. Examples.

- Binary or continuum? Needs to be studied empirically.

7.3   Fractal image analysis in archaeology

END Chapter

Chapter 8

Methods: Fractal image analysis

fdasdf

8.1 Fractal dimension and lacunarity

- Box-counting, lit. Mancuso (2021), Li, Du, and Sun (2009), Klinkenberg (1994)

- Gliding-box algorithmsc, Allain and Cloitre (1991), Hingee et al. (2019), Cheng (1997), Plotnick et al. (1996)

- Caveats:

    - Fractional box-counting dimension does not equal self-similarity in a simple way

    - My summary L is not equal to the one used in *FracLac* and thus by Farías-Pelayo (2017)

8.2 Image preparation

- Procedure for archaeological samples, same as in article

- Procedure for synthetic sample

END Chapter

Chapter 9

Results: Image analysis

9.1 Synthetic settlement plans

Relationship between image density (built-up area) and fractal dimension, evaluated in Thomas, Frankhauser, and De Keersmaecker (2007), where they show that these two parameters, under certain conditions (constant observation window, prefactor values close to 1), are exponentially correlated, which is also what I found in my article. They show that observation window size and shape, as well as centroid placement, have little influence on D, while they have more influence on density when the pattern is not homogeneous. They do show, however, that images with the same density may have quite much variation in D, which is reflected in the layouts. Judging from their examples, more clustered layouts give higher D values, while more dispersed or dusty layouts give lower D, when density is constant. I ignore the use of prefactor values. According to Thomas et al. (2007), density is a crude measure of the overall intensity of the pattern, while fractal dimension is characterises the morphological structure, though it is not directly descriptive.

9.2 Settlements

9.3 Quarters/neighbourhoods

9.4 Temporal samples (Vráble)

9.5 Summary of findings

END Chapter

Part IV

Synthesis

Chapter 10

Discussion: Social complexity in Linear Pottery and Trypillia settlements

fdsasdf

END Chapter

Chapter 11

Discussion: Fractal Analysis and Archaeological data

fdsasdf

Also mention here fractals or related concepts used as a metaphor, with no mathematics involved (e.g. Chapman et al. 2006; Sherratt 2004; Sindbæk 2022; Whitridge 2016).

11.1 Distribution fitting

- Multiplicative processes: power laws vs. log-normal, long debates Sheridan and Onodera (2018). Does it really matter in the end?

-

11.2 Image analysis

fdsa

11.3   Concluding remarks?

Check for "fractal + archaeology" in WorldCat.

END Chapter

Chapter 12

Conclusion and Outlook

12.1 Things I would like to have done, but that didn't fit into this study

- Ethnoarchaeology: Measure house sizes and settlement layouts in contemporary settings, and relate to social organisation (largely overlooked by ethnographers)
- Test distributions and settlement layout analysis on other settings: Lake dwellings, later/historic periods, other materials (e.g. megaliths)… Add more complex distribution models, add observation windows on images. Try on remote sensing imagery.
- Settlement Scaling on Neolithic settings
- Time series: Hurst exponent and scale invariance in temporal development of e.g. regional settlement or population
- Integrate – bridge the gap – between opposite theoretical (nat. and soc./hum.) approaches to the same phenomena
- Chaos and strange attractors in Archaeology
- More?

12.2   Concluding remarks

END Thesis

References

Alberti, Gianmarco. 2014. "Modeling Group Size and Scalar Stress by Logistic Regression from an Archaeological Perspective." *PLOS ONE* 9(3):e91510. doi: 10.1371/journal.pone.0091510.

Allain, C., and M. Cloitre. 1991. "Characterizing the Lacunarity of Random and Deterministic Fractal Sets." *Physical Review A* 44(6):3552–58. doi: 10.1103/PhysRevA.44.3552.

Arshad, Sidra, Shougeng Hu, and Badar Nadeem Ashraf. 2018. "Zipf's Law and City Size Distribution: A Survey of the Literature and Future Research Agenda." *Physica A: Statistical Mechanics and Its Applications* 492:75–92. doi: 10.1016/j.physa.2017.10.005.

Artursson, Magnus, Jens-Henrik Bech, Timothy Earle, Dániel Fuköh, Michael J. Kolb, Kristian Kristiansen, Johan Ling, Christian Mühlenbock, Christopher Prescott, Chris Sevara, Sebastiano Tusa, Claes Uhnér, and Magdolna Vicze. 2010. "Settlement Structure and Organisation." Pp. 87–121 in *Organizing Bronze Age Societies: The Mediterranean, Central Europe, and Scandanavia Compared*, edited by K. Kristiansen and T. Earle. Cambridge: Cambridge University Press.

Auerbach, Felix. 1913. "Das Gesetz der Bevölkerungskonzentration." *Dr. A. Petermanns Mitteilungen aus Justus Perthes' Geographischer Anstalt* LIX:73–76.

Augereau, Anne. 2021. *Femmes Néolithiques - Le Genre Dans Les Premières Sociétés Agricoles*. Illustrated édition. Paris: Cnrs.

Baden, William W., and Christopher S. Beekman. 2016. *Nonlinear Models for Archaeology and Anthropology: Continuing the Revolution*. London: Routledge.

Bak, Per. 2013. *How Nature Works: The Science of Self-Organized Criticality*. New York, NY: Springer.

Bak, Per, Chao Tang, and Kurt Wiesenfeld. 1987. "Self-Organized Criticality: An Explanation

of the 1/f Noise." *Physical Review Letters* 59(4):381–84. doi: 10.1103/PhysRevLett.59.381.

Bánffy, Eszter, and Hajo Höhler-Brockmann. 2020. "Burnt daub talking: The formation of the LBK longhouse (a work hypothesis)." *Quaternary international* 560-561:179196. doi: 10.1016/j.quaint.2020.06.007.

Barabási, Albert-László, and Réka Albert. 1999. "Emergence of Scaling in Random Networks." *Science* 286(5439):509–12. doi: 10.1126/science.286.5439.509.

Barley, Nigel. 2011. *The Innocent Anthropologist: Notes from a Mud Hut*. Eland.

Batty, Michael. 2005. *Cities and complexity: understanding cities with cellular automata, agent-based models, and fractals*. Cambridge, Mass.: MIT Press.

Batty, Michael, and Paul Longley. 1994. *Fractal Cities - a Geometry of Form and Function*. London: Academic Press.

Baxter, Michael. 2003. *Statistics in archaeology*. London: Arnold.

Beck, Robin A., ed. 2007. *The Durable House: House Society Models in Archaeology*. Center for Archaeological Investigations, Southern Illinois University.

Bee, Marco, Massimo Riccaboni, and Stefano Schiavo. 2011. "Pareto Versus Lognormal: A Maximum Entropy Test." *Physical Review E* 84(2):026104. doi: 10.1103/PhysRevE.84.026104.

Bentley, R. Alexander, and Herbert D. G. Maschner. 2001. "Stylistic Change as a Self-Organized Critical Phenomenon: An Archaeological Study in Complexity." *Journal of Archaeological Method and Theory* 8(1):35–66. doi: 10.1023/A:1009521915258.

Bentley, R. Alexander, and Herbert D. G. Maschner. 2008. "Complexity Theory." edited by R. A. Bentley, H. D. G. Maschner, and C. Chippindale. Lanham, Md: AltaMira Press.

Bettencourt, Luís M. A. 2021. *Introduction to Urban Science: Evidence and Theory of Cities as Complex Systems*. Cambridge, MA, USA: MIT Press.

Birch, Jennifer. 2014. *From prehistoric villages to cities: settlement aggregation and community*. Vol. 10. Routledge.

Blanton, Richard E. 1994. *Houses and Households*. Boston, MA: Springer US.

Bradley, Richard. 2001. "Orientations and origins: a symbolic dimension to the long house in Neolithic Europe." *Antiquity* 75(287):5056. doi: 10.1017/S0003598X00052704.

Bradley, Richard. 2013. "Houses of Commons, Houses of Lords: Domestic Dwellings and Mon-

umental Architecture in Prehistoric Europe." *Proceedings of the Prehistoric Society* 79:1–17. doi: 10.1017/ppr.2013.1.

Brink, Kristian. 2013. "Houses and Hierarchies: Economic and Social Relations in the Late Neolithic and Early Bronze Age of Southernmost Scandinavia." *European Journal of Archaeology* 16(3):433–58. doi: 10.1179/1461957113Y.0000000033.

Brown, Clifford T., Walter RT Witschey, and Larry S. Liebovitch. 2005. "The Broken past: Fractals in Archaeology." *Journal of archaeological method and theory* 12(1):3778. doi: 10.1007/s10816-005-2396-6.

Brown, Clifford, and Larry Liebovitch. 2010. *Fractal Analysis*. 2455 Teller Road, Thousand Oaks California 91320 United States of America: SAGE Publications, Inc.

Brown, Clifford, April Watson, Ashley Gravlin-Beman, and Larry Liebovitch. 2012. "Poor Mayapán." Pp. 306–24 in. London: Equinox Publishing.

Cafazzo, Simona, Martina Lazzaroni, and Sarah Marshall-Pescini. 2016. "Dominance Relationships in a Family Pack of Captive Arctic Wolves (Canis Lupus Arctos): The Influence of Competition for Food, Age and Sex." *PeerJ* 4:e2707. doi: 10.7717/peerj.2707.

Carneiro, Robert L. 1986. "On the Relationship Between Size of Population and Complexity of Social Organization." *Journal of Anthropological Research* 42(3):355–64.

Carsten, Janet, and Stephen Hugh-Jones. 1995. *About the house: Lévi-Strauss and beyond*. Cambridge: University Press.

Chapman, John, Tom Higham, Vladimir Slavchev, Bisserka Gaydarska, and Noah Honch. 2006. "The social context of the emergence, development and abandonment of the Varna cemetery, Bulgaria." *European journal of archaeology* 9(2-3):159183. doi: 10.1177/1461957107086121.

Chen, Yanguang. 2011. "Fractal Systems of Central Places Based on Intermittency of Space-Filling." *Chaos, Solitons & Fractals* 44:619–32.

Cheng, Qiuming. 1997. "Multifractal Modeling and Lacunarity Analysis." *Mathematical geology* 29(7):919932. doi: 10.1023/A:1022355723781.

Clauset, Aaron, Cosma Rohilla Shalizi, and M. E. J. Newman. 2009. "Power-Law Distributions in Empirical Data." *Society for Industrial and Applied Mathematics Review* 51(4):661–703.

Cleuziou, Serge, Frank Braemer, and Anick Coudart. 1999. *Habitat et société*. Antibes:

APDCA.

Costin, Cathy Lynne. 1991. "Craft Specialization: Issues in Defining, Documenting, and Explaining the Organization of Production." *Archaeological Method and Theory* 3:1–56.

Coudart, Anick. 1998. *Architecture Et Société Néolithique : L'unité Et La Variance de La Maison Danubienne*. Paris: Éditions de la Maison des Sciences de l'Homme.

Coudart, Anick. 2015. "The Bandkeramik Longhouses: A Material, Social, and Mental Metaphor for Small-Scale Sedentary Societies." Pp. 309–25 in, edited by C. Fowler, J. Harding, and D. Hofmann. Oxford: Oxford University Press.

Crabtree, Stefani A., R. Kyle Bocinsky, Paul L. Hooper, Susan C. Ryan, and Timothy A. Kohler. 2017. "How to Make a Polity (in the Central Mesa Verde Region)." *American Antiquity* 82(1):71–95. doi: 10.1017/aaq.2016.18.

Crumley, Carole L. 1995. "Heterarchy and the Analysis of Complex Societies." *Archaeological Papers of the American Anthropological Association* 6(1):1–5. doi: https://doi.org/10.1525/ap3a.1995.6.1.1.

D'Acci, Luca, ed. 2019. *The Mathematics of Urban Morphology*. Cham: Springer International Publishing.

Daems, Dries. 2021. *Social complexity and complex systems in archaeology*. London, New York: Routledge.

De Landa, Manuel. 2006. *A new philosophy of society: assemblage theory and social complexity*. London: Continuum.

Devaney, Robert L. 2020. *A First Course In Chaotic Dynamical Systems: Theory And Experiment*. Milton: CRC Press.

Diachenko, Aleksandr, and Iwona Sobkowiak-Tabaka. 2022. "Self-Organized Cultural Cycles and the Uncertainty of Archaeological Thought." *Journal of Archaeological Method and Theory*. doi: 10.1007/s10816-022-09548-8.

Diachenko, Alexandr. 2018. "Fractal Models" edited by S. López Varela.

Dohm, Karen. 1990. "Effect of population nucleation on house size for pueblos in the American Southwest." *Journal of anthropological archaeology* 9(3):201239. doi: 10.1016/0278-4165(90)90007-Z.

Dubreuil, Benoît. 2010. *Human evolution and the origins of hierarchies: the state of nature.*

Cambridge: University Press.

Duffy, Paul R. 2015. "Site Size Hierarchy in Middle-Range Societies." *Journal of Anthropological Archaeology* 37:85–99. doi: 10.1016/j.jaa.2014.12.001.

Dunbar, Robin. 2022. *Annales Academiciae Scientarum Fennicae [Forthcoming] Reconstructing Dunbar's Number: A Guide for the Perplexed to Some Pitfalls, Indiscretions and Fallacies*.

Earle, Timothy. 2002. *Bronze Age Economics: The Beginnings of Political Economies*. Boulder, CO: Westview Press.

Earle, Timothy K. 1997. *How chiefs come to power: the political economy in prehistory*. Stanford, Calif: Stanford University Press.

Eglash, Ron. 1999. *African Fractals: Modern Computing and Indigenous Design*. Rutgers University Press.

Ensor, Bradley E. 2013. *The Archaeology of Kinship: Advancing Interpretation and Contributions to Theory*. University of Arizona Press.

Ensor, Bradley E. 2017. "Testing Ethnological Theories on Prehistoric Kinship." *Cross-Cultural Research* 51(3):199–227. doi: 10.1177/1069397117697648.

Ensor, Bradley E. 2021. "Making aDNA Useful for Kinship Analysis." *Antiquity* 95(379):241–43. doi: 10.15184/aqy.2020.234.

Ensor, Bradley E., Joel D. Irish, and William F. Keegan. 2017. "The Bioarchaeology of Kinship: Proposed Revisions to Assumptions Guiding Interpretation." *Current Anthropology* 58(6):739–61. doi: 10.1086/694584.

Falconer, Kenneth, and Kenneth Falconer. 2013. *Fractals: A Very Short Introduction*. Oxford, New York: Oxford University Press.

Farías-Pelayo, Sabrina. 2017. "Characterization of Cultural Traits by Means of Fractal Analysis." edited by F. Brambila. IntechOpen.

Fontijn, David, and Joanna Brück. 2013. "The Myth of the Chief: Prestige Goods, Power, and Personhood in the European Bronze Age." in *The Oxford Handbook of the European Bronze Age*, *Oxford Handbooks in Archaeology*. Oxford University Press.

Fraser, Douglas. 1968. *Village Planning in the Primitive World*. New York: George Braziller.

Furholt, Martin. 2016. "Settlement Layout and Social Organisation in the Earliest European

Neolithic." *Antiquity* 90(353):1196–1212.

Furholt, Martin, Colin Grier, Matthew Spriggs, and Timothy Earle. 2020. "Political Economy in the Archaeology of Emergent Complexity: A Synthesis of Bottom-Up and Top-Down Approaches." *Journal of Archaeological Method and Theory* 27(2):157–91. doi: 10.1007/s10816-019-09422-0.

Furholt, Martin, Johannes Müller, Alena Bistáková, Maria Wunderlich, Ivan Cheben, and Nils Müller-Scheeßel. 2020. *Archaeology in the Zitava Valley 1 - the LBK and Zeliezovce Settlement Site of Vráble*. Leiden: Sidestone Press.

Furholt, Martin, Nils Müller-Scheeßel, Maria Wunderlich, Ivan Cheben, and Johannes Müller. 2020. "Communality and Discord in an Early Neolithic Settlement Agglomeration: The LBK Site of Vráble, Southwest Slovakia." *Cambridge Archaeological Journal* 30(3):469–89.

Gaydarska, Bisserka, Katharina Rebay-Salisbury, Paz Ramirez Valiente, Jana Esther Fries, Daniela Hofmann, Anne Augereau, John Chapman, Maria Mina, Eleonore Pape, Nicola Ialongo, Daniela Nordholz, Penny Bickle, Mark Haughton, John Robb, and Oliver Harris. 2023. "To Gender or not To Gender? Exploring Gender Variations through Time and Space." *European journal of archaeology* 128. doi: 10.1017/eaa.2022.51.

Gibrat, Robert. 1930. "Une Loi Des Répartitions Économiques : L'effet Proportionnel." *Bulletin de La Statistique Générale de La France* 19(3):469–513.

Gillespie, Colin S. 2015. "Fitting Heavy Tailed Distributions: The poweRlaw Package." *Journal of Statistical Software* 64(2).

Gleick, James. 1987. *Chaos: Making a New Science*. New York, N.Y: Viking.

Godelier, Maurice. 1986. *The making of great men: male domination and power among the New Guinea Baruya*. Vol. 56. Cambridge: University Press.

Gomez-Lievano, Andres, Hyejin Youn, and Luís Bettencourt. 2012. "The Statistics of Urban Scaling and Their Connection to Zipf's Law." *PloS One* 7:e40393. doi: 10.1371/journal.pone.0040393.

Goodale, Nathan, Colin P. Quinn, and Alissa Nauman. 2021. "Monumentality of Houses: Collective Action, Inequality, and Kinship in Pithouse Construction." Pp. 177–203 in *Archaeology of Households, Kinship, and Social Change*, edited by L. B. Carpenter and A. M. Prentiss. Routledge.

Gould, Stephen Jay. 2007. *Punctuated equilibrium*. Cambridge, Mass: Belknap Press of Harvard University Press.

Graeber, David, and David Wengrow. 2021. *The dawn of everything: a new history of humanity*. New York: Farrar, Straus; Giroux.

Gronenborn, Detlef, Hans-Christoph Strien, Stephan Dietrich, and Frank Sirocko. 2014. "'Adaptive Cycles' and Climate Fluctuations: A Case Study from Linear Pottery Culture in Western Central Europe." *Journal of Archaeological Science* 51:73–83. doi: 10.1016/j.jas.2013.03.015.

Grove, Matt. 2011. "An Archaeological Signature of Multi-Level Social Systems: The Case of the Irish Bronze Age." *Journal of Anthropological Archaeology* 30(1):44–61. doi: 10.1016/j.jaa.2010.12.002.

Haas, W. Randall, Cynthia J. Klink, Greg J. Maggard, and Mark S. Aldenderfer. 2015. "Settlement-Size Scaling among Prehistoric Hunter-Gatherer Settlement Systems in the New World." *PloS One* 10(11):e0140127. doi: 10.1371/journal.pone.0140127.

Hamilton, Marcus, Bruce Milne, Robert Walker, Oskar Burger, and James Brown. 2007. "The Complex Structure of Hunter-Gatherer Social Networks." *Proceedings of the Royal Society B* 274:2195–2202.

Harper, Thomas K., Aleksandr Diachenko, Sergei N. Ryzhov, Yuri Y. Rassamakin, Laurie R. Eccles, Douglas J. Kennett, and Elena V. Tsvek. 2023. "ASSESSING THE SPREAD OF ENEOLITHIC AGRICULTURAL COMMUNITIES IN THE FOREST-STEPPE OF UKRAINE USING AMS RADIOCARBON DATING." *Radiocarbon* 1–21. doi: 10.1017/RDC.2023.28.

Harrison, Alan. 1981. "A Tale of Two Distributions." *The Review of Economic Studies* 48(4):621–31.

Harte, David. 2001. *Multifractals: Theory and Applications*. New York: Chapman; Hall/CRC.

Haude, Rüdiger, and Thomas Wagner. 2019. *Herrschaftsfreie Institutionen : Texte zur Stabilisierung staatsloser, egalitärer Gesellschaften*. Heidelberg: Verlag Graswurzelrevolution.

Hingee, Kassel, Adrian Baddeley, Peter Caccetta, and Gopalan Nair. 2019. "Computation of Lacunarity from Covariance of Spatial Binary Maps." *Journal of Agricultural, Biological and Environmental Statistics* 24(2):264–88. doi: 10.1007/s13253-019-00351-9.

Hodder, Ian. 1979. "Simulating the Growth of Hierarchies." Pp. 117–44 in, edited by C. Renfrew

and K. L. Cooke. New York: Academic Press.

Hofmann, Daniela, and Jessica Smyth, eds. 2013. *Tracking the Neolithic House in Europe*. New York, NY: Springer New York.

Honoré, Emmanuelle. 2022. "Différences de richesse et inégalités sociales au Paléolithique: éléments pour une discussion." *Paléo (Les Eyzies de Tayac-Sireuil)* 5059.

Hrnčíř, Václav, Pavel Duda, Gabriel Šaffa, Petr Květina, and Jan Zrzavý. 2020. "Identifying Post-Marital Residence Patterns in Prehistory: A Phylogenetic Comparative Analysis of Dwelling Size" edited by P. F. Biehl. *PLOS ONE* 15(2):e0229363. doi: 10.1371/journal.pone.0229363.

Hrnčíř, Václav, Václav Vondrovský, and Petr Květina. 2020. "Post-Marital Residence Patterns in LBK: Comparison of Different Models." *Journal of Anthropological Archaeology* 59:101190. doi: 10.1016/j.jaa.2020.101190.

Jahanmiri, Fatemeh, and Dawn Cassandra Parker. 2022. "An Overview of Fractal Geometry Applied to Urban Planning." *Land* 11(4):475. doi: 10.3390/land11040475.

Jeunesse, Christian. 2022. "Vertical Social Differentiation and Social Competition in the Lbk: "Rich" and "Poor" Cemeteries." Pp. 33–42 in, *Universitätsforschungen zur prähistorischen Archäologie*, edited by M. Dębiec, J. Górski, J. Müller, M. Nowak, A. Pelisiak, T. Saile, and P. Włodarczak. Bonn: Verlag Dr. Rudolf Habelt GmbH.

Johnson, Allen, and Timothy Earle. 1987. *The Evolution of Human Societies : From Foraging Group to Agrarian State*. Stanford: Stanford University Press.

Johnson, Gregory A. 1982. "Organizational Structure and Scalar Stress." Pp. 389–421 in, edited by C. Renfrew, M. J. Rowlands, and B. A. Segraves. New York: Academic Press.

Johnson, Norman L. 1994. *Continuous univariate distributions: Vol. 1*. Vols. Vol. 1. 2nd ed. New York: Wiley.

Joyce, Rosemary A., and Susan D. Gillespie. 2000. *Beyond kinship: social and material reproduction in house societies*. Philadelphia, Pa: University of Pennsylvania Press.

Kadrow, Sławomir. 2013. "Social Organization and Change." P. 0 in, edited by A. Gardner, M. Lake, and U. Sommer. Oxford University Press.

Kahn, Jennifer G. 2021. "Houses of Power: Community Houses and Specialized Houses as Markers of Social Complexity in the Pre-Contact Society Island Chiefdoms." Pp. 82–110 in

Archaeology of Households, Kinship, and Social Change, edited by L. B. Carpenter and A. M. Prentiss. Routledge.

Kienlin, Tobias L., and Andreas Zimmermann. 2012. *Beyond Elites - Alternatives to Hierarchical Systems in Modelling Social Formations: International Conference at the Ruhr-Universität Bochum, Germany, ... Zur Prähistorischen Archäologie)*. Bonn: Habelt, R.

Klinkenberg, Brian. 1994. "A Review of Methods Used to Determine the Fractal Dimension of Linear Features." *Mathematical Geology* 26(1):23–46. doi: 10.1007/BF02065874.

Kohler, Timothy A., and Michael E. Smith. 2018. *Ten Thousand Years of Inequality: The Archaeology of Wealth Differences*. Tucson: University of Arizona Press.

Kostof, Spiro. 1991. *The city shaped: urban patterns and meanings through history ; original drawings by Richard Tobias*. London: Thames; Hudson.

Kramer, Carol. 1982. *Village ethnoarchaeology: rural Iran in archaeological perspective*. New York: Academic Press.

Kuijt, Ian. 2018. "Material Geographies of House Societies: Reconsidering Neolithic Çatalhöyük, Turkey." *Cambridge Archaeological Journal* 28(4):565–90. doi: 10.1017/S0959774318000240.

Květina, Petr, and Jaroslav Řídký. 2019. "Traditional and Archaic Societies - Problems Linked to the Search for Social and Power Attributes in the Archaeological Record." Pp. 7–14 in *Big men or chiefs? Rondel builders of neolithic Europe*. Oxford: Oxbow Books.

Lagarias, Apostolos, and Poulicos Prastacos. 2021. "Fractal Dimension of European Cities: A Comparison of the Patterns of Built-up Areas in the Urban Core and the Peri-Urban Ring." *Cybergeo: European Journal of Geography*. doi: 10.4000/cybergeo.37243.

Laherrère, Jean. 2000. "Distribution of Field Sizes in a Petroleum System: Parabolic Fractal, Lognormal or Stretched Exponential?" *Marine and Petroleum Geology* 17:539–46.

Laherrère, J., and D. Sornette. 1998. "Stretched Exponential Distributions in Nature and Economy: "Fat Tails" with Characteristic Scales." *The European Physical Journal B - Condensed Matter and Complex Systems* 2(4):525–39. doi: 10.1007/s100510050276.

Last, Jonathan. 2015. "Longhouse Lifestyles in the Central European Neolithic." P. 0 in, edited by C. Fowler, J. Harding, and D. Hofmann. Oxford University Press.

Lévi-Strauss, Claude. 1982. *The way of the masks*. Seattle, Wash: University of Washington Press.

Li, Jian, Qian Du, and Caixin Sun. 2009. "An Improved Box-Counting Method for Image Fractal Dimension Estimation." *Pattern Recognition* 42(11):2460–69. doi: 10.1016/j.patcog.2009.03.001.

Lobo, Jose, Luis MA Bettencourt, Michael E. Smith, and Scott Ortman. 2020. "Settlement Scaling Theory: Bridging the Study of Ancient and Contemporary Urban Systems." *Urban Studies* 57(4):731–47. doi: 10.1177/0042098019873796.

Lund, Julie, Martin Furholt, and Knut Ivar Austvoll. 2022. "Reassessing Power in the Archaeological Discourse. How Collective, Cooperative and Affective Perspectives May Impact Our Understanding of Social Relations and Organization in Prehistory." *Archaeological Dialogues* 1–18. doi: 10.1017/S1380203822000162.

Madella, Marco. 2013. *The archaeology of household*. Oxford: Oxbow books.

Mancuso, Francesco Paolo. 2021. "fractD: Estimation of Fractal Dimension of a Black Area in 2D and 3D (Slices) Images."

Mandelbrot, Benoit B. 1975. *Les objets fractals: forme, hasard et dimension*. Paris: Flammarion.

Mandelbrot, Benoit B. 1982. *The fractal geometry of nature*. San Francisco: Freeman.

Mandelbrot, Benoit B. 1997. *Fractals and scaling in finance: discontinuity, concentration, risk : selecta volume E*. New York: Springer.

Mandelbrot, Benoit B. 2021. *The Fractal Geometry of Nature*. Brattleboro: Echo Point Books & Media, LLC.

Mandelbrot, Benoît. 1967. "How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension." *Science* 156(3775):636–38.

Maschner, Herbert, and Alexander Bentley. 2003. "The Power Law of Rank and Household on the North Pacific." Pp. 47–143 in, edited by A. Bentley and H. Maschner. Salt Lake City: The University of Utah Press.

Mazerolle, Marc J. 2023. *AICcmodavg: Model Selection and Multimodel Inference Based on (q)AIC(c)*.

Mech, L. David. 1999. "Alpha Status, Dominance, and Division of Labor in Wolf Packs." *Canadian Journal of Zoology* 77(8):1196–1203. doi: 10.1139/z99-099.

Merton, Robert K. 1968. "The Matthew Effect in Science." *Science* 159(3810):56–63.

Midlarsky, Manus I. 1999. *The Evolution of Inequality: War, State Survival, and Democracy in*

Comparative Perspective. Stanford: Stanford University Press.

Mitzenmacher, Michael. 2004. "A Brief History of Generative Models for Power Law and Log-normal Distributions." *Internet Mathematics* 1(2):226–51. doi: 10.1080/15427951.2004.10129088.

Modderman, P. J. R., R. R. Newell, J. E. Brinkman, and C. C. Bakels. 1970. *Linearbandkeramik aus Elsloo und Stein*. Vol. 3. Leiden University Press.

Morgan, Lewis H. 1965. *Houses and House-Life of the American Aborigines*. Chicago: The University of Chicago Press.

Müller-Scheeßel, Niels. 2007. "Further Thoughts on the Hidden Settlement Structures of the Hallstatt Period in Southern Germany: Testing for Auto-Correlation." *Archäologisches Korrespondenzblatt* 37(1):57–66.

Müller-Scheeßel, Nils. 2019. "A New Approach to the Temporal Significance of House Orientations in European Early Neolithic Settlements." doi: 10.17026/DANS-XU6-C6YR.

Murdock, George Peter. 1949. *Social structure*. New York: Macmillan.

Muro, C., R. Escobedo, L. Spector, and R. P. Coppinger. 2011. "Wolf-Pack (Canis Lupus) Hunting Strategies Emerge from Simple Rules in Computational Simulations." *Behavioural Processes* 88(3):192–97. doi: 10.1016/j.beproc.2011.09.006.

Netting, Robert McC. 1982. "Some Home Truths on Household Size and Wealth." *American Behavioral Scientist* 25(6):641–62. doi: 10.1177/000276482025006004.

Newman, M. E. J. 2005. "Power Laws, Pareto Distributions and Zipf's Law." *Contemporary Physics* 46(5):325–51. doi: 10.1080/00107510500052444.

Nordquist, Pär. 2001. "Hierarkiseringsprocesser: om konstruktion av social ojämlikhet i Skåne, 5500-1100 f.Kr." PhD thesis, Umeå.

Ohlrau, René. 2022. "Trypillia Mega-Sites: Neither Urban nor Low-Density?" *Journal of Urban Archaeology* 5:81–100. doi: 10.1484/J.JUA.5.129844.

Packard, Jane. 2003. "Wolf Behavior: Reproductive, Social and Intelligent." Pp. 35–65 in *Wolves: Behavior, ecology, and conservation*, edited by L. D. Mech and L. Boitani. Chicago and London: The University of Chicago Press.

Pareto, Vilfredo. 1896. *Cours d'économie politique*. Vol. 1. Éd. F. Rouge. Lausanne.

Perreault, Charles. 2019. *The quality of the archaeological record*. Chicago: The University of Chicago Press.

Peters, Robin, and Andreas Zimmermann. 2017. "Resilience and Cyclicity: Towards a Macrohistory of the Central European Neolithic." *Quaternary International* 446:43–53. doi: 10.1016/j.quaint.2017.03.073.

Pickartz, Natalie, Robert Hofmann, Stefan Dreibrodt, Knut Rassmann, Liudmyla Shatilo, René Ohlrau, Dennis Wilken, and Wolfgang Rabbel. 2019. "Deciphering Archeological Contexts from the Magnetic Map: Determination of Daub Distribution and Mass of Chalcolithic House Remains." *The Holocene* 29(10):1637–52. doi: 10.1177/0959683619857238.

Pickartz, Natalie, Robert Hofmann, Tina Wunderlich, Knut Rassmann, Johannes Müller, René Ohlrau, Dennis Wilken, and Wolfgang Rabbel. 2021. "Magnetization Patterns as Indicator of Settlement Structures." *ArcheoSciences* 45-1(1):187–90.

Plotnick, Roy E., Robert H. Gardner, William W. Hargrove, Karen Prestegaard, and Martin Perlmutter. 1996. "Lacunarity Analysis: A General Technique for the Analysis of Spatial Patterns." *Physical Review E* 53(5):5461–68. doi: 10.1103/PhysRevE.53.5461.

Porcic, Marko. 2010. "House Floor Area as a Correlate of Marital Residence Pattern: A Logistic Regression Approach." *Cross-cultural research* 44(4):405424. doi: 10.1177/1069397110378839.

Porčić, Marko. 2012. "Effects of Residential Mobility on the Ratio of Average House Floor Area to Average Household Size: Implications for Demographic Reconstructions in Archaeology." *Cross-Cultural Research* 46(1):72–86. doi: 10.1177/1069397111423889.

Price, T. Douglas, and Gary M. Feinman. 1995. *Foundations of social inequality*. New York: Plenum Press.

Price, T. Douglas, and Gary M. Feinman. 2010. *Pathways to Power: New Perspectives on the Emergence of Social Inequality*. New York, NY: Springer.

Pumain, Denise. 2006. "Introduction." Pp. 1–12 in *Hierarchy in Natural and Social Sciences*, *Methodos Series*, edited by D. Pumain. Dordrecht: Springer Netherlands.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rathje, William L., and Randall H. McGuire. 1982. "Rich Men... Poor Men." *American Behavioral Scientist* 25(6):705–15. doi: 10.1177/000276482025006008.

Redhead, Daniel, and Eleanor A. Power. 2022. "Social Hierarchies and Social Networks

in Humans." *Philosophical Transactions of the Royal Society B: Biological Sciences* 377(1845):20200440. doi: 10.1098/rstb.2020.0440.

Rodríguez, Luis, Oscar Castillo, José Soria, Patricia Melin, Fevrier Valdez, Claudia I. Gonzalez, Gabriela E. Martinez, and Jesus Soto. 2017. "A Fuzzy Hierarchical Operator in the Grey Wolf Optimizer Algorithm." *Applied Soft Computing* 57:315–28. doi: 10.1016/j.asoc.2017.03.048.

Ross, Jennifer C., and Sharon R. Steadman. 2017. *Ancient complex societies*. New York: Routledge.

Roussot-Larroque, Julia. 2022. "Relations sud-nord en Europe occidentale au Néolithique ancien." P. 10-40 in *Le Néolithique du nord-est de la France et des régions limitrophes : Actes du XIIIe colloque interrégional sur le Néolithique*, *Documents d'archéologie française*, edited by C. Guillaume. Paris: Éditions de la Maison des sciences de l'homme.

Sahlins, Marshall. 2013. *What kinship is - and is not*. Chicago: University of Chicago Press.

Sahlins, Marshall. 2020. *Stone Age Economics*. Routledge.

Saile, Thomas. 2020. "On the Bandkeramik to the east of the Vistula River: At the limits of the possible." *Quaternary international* 560-561:208227. doi: 10.1016/j.quaint.2020.04.036.

Scheidel, Walter. 2017. *Great leveler: violence and the history of inequality from the stone age to the twenty-first century*. Princeton: University Press.

Schiesberg, Sara. 2010. "Von Häusern Und Menschen. Das Beispiel Bandkeramik." Pp. 53–69 in Vol. 1, *Fokus jungsteinzeit. Berichte der AG neolithikum*, edited by E. Claßen, T. Doppler, and B. Ramminger. Loogh: Welt und Erde.

Schiesberg, Sara. 2016. "Post-Marital Residence Patterns and Descent Rules in Prehistory." in Vol. 285, *Universitätsforschungen zur prähistorischen archäologie*, edited by T. Kerig, K. Nowak, and G. Roth. Bonn: Dr. Rudolf Habelt.

Schreier, Amy L., and Larissa Swedell. 2009. "The Fourth Level of Social Structure in a Multi-Level Society: Ecological and Social Functions of Clans in Hamadryas Baboons." *American Journal of Primatology* 71(11):948–55. doi: 10.1002/ajp.20736.

Service, Elman. 1971. *Primitive Social Organization : An Evolutionary Perspective*. 2nd ed. New York: Random House.

Shatilo, Liudmyla. 2021. *Tripolye Typo-Chronology. Mega and Smaller Sites in the Sinyukha*

River Basin. Sidestone Press.

Shennan, Stephen. 2008. *Quantifying archaeology*. Second edition. Edinburgh: University Press.

Shennan, Stephen. 2018. *The first farmers of Europe: an evolutionary perspective*. Cambridge: University Press.

Sheridan, Paul, and Taku Onodera. 2018. "A Preferential Attachment Paradox: How Preferential Attachment Combines with Growth to Produce Networks with Log-Normal In-Degree Distributions." *Scientific Reports* 8(1):2811. doi: 10.1038/s41598-018-21133-2.

Sherratt, Andrew. 2004. "Fractal Farmers: Patterns of Neolithic Origin and Dispersal." Pp. 53–63 in, edited by J. Cherry, C. Scarre, and S. Shennan. Cambridge: McDonald Institute for Archaeological Research.

Shimoji, Hiroyuki, and Shigeto Dobata. 2022. "The Build-up of Dominance Hierarchies in Eusocial Insects." *Philosophical Transactions of the Royal Society B: Biological Sciences* 377(1845):20200437. doi: 10.1098/rstb.2020.0437.

Simon, Herbert A. 1955. "On a Class of Skew Distribution Functions." *Biometrika* 42(3-4):425440. doi: 10.1093/biomet/42.3-4.425.

Sindbæk, Søren M. 2022. "Weak Ties and Strange Attractors: Anomalocivitas and the Archaeology of Urban Origins." *Journal of Urban Archaeology* 5:19–32. doi: 10.1484/J.JUA.5.129841.

Smith, Michael E. 1987. "Household Possessions and Wealth in Agrarian States: Implications for Archaeology." *Journal of Anthropological Archaeology* 6(4):297–335. doi: 10.1016/0278-4165(87)90004-3.

Smith, Michael E., ed. 2011. *The Comparative Archaeology of Complex Societies*. New York: Cambridge University Press.

Souvatzi, Stella. 2017. "Kinship and Social Archaeology." *Cross-cultural research* 51(2):172195. doi: 10.1177/1069397117691028.

Strauss, Eli D., James P. Curley, Daizaburo Shizuka, and Elizabeth A. Hobson. 2022. "The Centennial of the Pecking Order: Current State and Future Prospects for the Study of Dominance Hierarchies." *Philosophical Transactions of the Royal Society B: Biological Sciences* 377(1845):20200432. doi: 10.1098/rstb.2020.0432.

Strawinska-Zanko, Urszula, Larry Liebovitch, April Watson, and Clifford Brown. 2018. "Capital in the First Century: The Evolution of Inequality in Ancient Maya Society." Pp. 161–92 in, *Computational Social Sciences*, edited by U. Strawinska-Zanko and L. Liebovitch. Cham: Springer International Publishing.

Stumpf, Michael P. H., and Mason A. Porter. 2012. "Critical Truths about Power Laws." *Science* 335(6069):665–66. doi: 10.1126/science.1216142.

Tannier, Cécile, and Denise Pumain. 2005. "Fractals in Urban Geography: A Theoretical Outline and an Empirical Example." *Cybergeo : European Journal of Geography*. doi: https://doi.org/10.4000/cybergeo.3275.

Testart, Alain. 2005. *Eléments de Classification Des Sociétés*. Paris, France: Éditions Errance : ill. en noir, tabl., cartes, schémas, couv. ill. en coul.

Thomas, Isabelle, Pierre Frankhauser, and Marie-Laurence De Keersmaecker. 2007. "Fractal Dimension Versus Density of Built-up Surfaces in the Periphery of Brussels*." *Papers in Regional Science* 86(2):287–308. doi: 10.1111/j.1435-5957.2007.00122.x.

Trebsche, Peter, Nils Müller-Scheeβel, and Sabine Reinhold. 2010. *Der gebaute Raum: Bausteine einer Architektursoziologie vormoderner Gesellschaften*. Münster: Waxmann.

West, Bruce, Robin Dunbar, Culbreth Garland, and Paolo Grigolini. 2023. *Fractal Structure of Human and Primate Social Networks Optimizes Information Flow*.

West, Geoffrey B. 2017. *Scale: the universal laws of growth, innovation, sustainability, and the pace of life in organisms, cities, economies, and companies*. New York: Penguin Press.

Whitridge, Peter. 2016. "Fractal Worlds: An Archaeology of Nested Spatial Scales." *ARCTIC* 69(5):1–10. doi: 10.14430/arctic4659.

Whittle, Alasdair. 2022. "The Danube in Prehistory." *European Journal of Archaeology* 25(1):133137. doi: 10.1017/eaa.2021.59.

Wilk, Richard R. 1983. "Little house in the jungle: The causes of variation in house size among modern Kekchi Maya." *Journal of anthropological archaeology* 2(2):99116. doi: 10.1016/0278-4165(83)90009-0.

Wilk, Richard R., and William L. Rathje. 1982. "Household Archaeology." *American Behavioral Scientist* 25(6):617–39. doi: 10.1177/000276482025006003.

Xie, Yihui. 2015. *Dynamic Documents with R and Knitr*. 2nd ed. Boca Raton, Florida: Chap-

man; Hall/CRC.

Yoffee, Norman. 1993. "Too Many Chiefs? (Or, Safe Texts for the '90s)." Pp. 60–78 in, edited by N. Yoffee and A. Sherratt. Cambridge University Press.

Yoffee, Norman. 2005. *Myths of the Archaic State : Evolution of the Earliest Cities, States, and Civilizations*. Cambridge: Cambridge University Press.

Yule, George Udny. 1925. "A Mathematical Theory of Evolution, Based on the Conclusions of Dr. J. C. Willis." *Philosophical Transactions of the Royal Society of London. Series B* 213(402-410):21–87. doi: 10.1098/rstb.1925.0002.

Zhou, W.-X., D. Sornette, R. A. Hill, and R. I. M. Dunbar. 2005. "Discrete Hierarchical Organization of Social Group Sizes." *Proceedings of the Royal Society B: Biological Sciences* 272(1561):439–44. doi: 10.1098/rspb.2004.2970.

Zhukov, Dmitry S., Valery V. Kanishchev, and Sergey K. Lyamin. 2016. "Application of the Theory of Self-Organized Criticality to the Investigation of Historical Processes." *SAGE Open* 6(4):2158244016683216. doi: 10.1177/2158244016683216.

Zimmermann, Andreas. 2012. "Cultural Cycles in Central Europe During the Holocene." *Quaternary International* 274:251–58. doi: 10.1016/j.quaint.2012.05.014.

Zipf, George Kingsley. 1949. *Human behavior and the principle of least effort: an introduction to human ecology*. Cambridge: Addison-Wesley Press.

Appendix A

This is my first appendix

Jada jada jada

Appendix B

This is my second one

Bla bla bla